

Separation von $t\bar{t} + Z$ - und
 $t\bar{t} + H$ -Ereignissen im Rahmen einer
 $t\bar{t} + Z(b\bar{b})$ -Analyse am CMS-Experiment
(Separation of $t\bar{t} + Z$ - and $t\bar{t} + H$ -events in a
 $t\bar{t} + Z(b\bar{b})$ analysis at CMS)

Bachelorarbeit
von

Lennart Hendrik Bosch
Matrikelnummer: 2051972

am Institut für Experimentelle Teilchenphysik

Referent:	Prof. Dr. U. Husemann
Korreferent:	Dr. M. Schröder
Betreuender Doktorand:	J. van der Linden

Bearbeitungszeitraum: 04. November 2019 bis 10. März 2020

Erklärung zur Selbstständigkeit

Ich versichere, dass ich diese Arbeit selbstständig verfasst habe und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe, die wörtlich oder inhaltlich übernommenen Stellen als solche kenntlich gemacht und die Satzung des KIT zur Sicherung guter wissenschaftlicher Praxis in der gültigen Fassung vom 07.05.2018 beachtet habe.

Karlsruhe, den _____, _____
Datum Lennart Hendrik Bosch

Als Ansichtsexemplar genehmigt von

Karlsruhe, den _____, _____
Datum Prof. Dr. U. Husemann

Inhaltsverzeichnis

1. Einleitung	1
2. Teilchenphysikalische Grundlagen	3
2.1. Standardmodell der Teilchenphysik	3
2.2. Elementarteilchendynamik	4
2.3. $t\bar{t}$ +Z-Produktion und -Zerfall	4
2.4. Untergrundprozesse	5
3. CMS-Experiment	7
3.1. Large Hadron Collider	7
3.2. Bestandteile des CMS-Detektors	7
3.3. Kinematik im CMS-Detektor	8
3.4. Objektrekonstruktion	9
3.5. Berechnung von Ereignisvariablen	10
4. Analyseprozess	11
4.1. Simulation	11
4.2. Ereignisselektion	12
4.3. Numerische Anpassung	13
4.3.1. Neuronales Netz als Klassifizierer	13
4.3.2. Integration systematischer Unsicherheiten	14
5. Bisheriger Stand	17
5.1. Diskriminatoren	17
5.1.1. Ereignisrekonstruktion	17
5.1.2. Auslegung der ANNs	18
5.1.3. Variablenselektion	19
5.1.4. Bingrenzen	19
5.2. Unsicherheiten	19
5.3. Ergebnis	19
5.4. Ausblick	20
6. Entwicklung der Diskriminatoren	21
6.1. χ^2 -Rekonstruktion des Bosons	21
6.1.1. Abgeleitete Variablen	21
6.1.2. Erfolgsrate der χ^2 -Rekonstruktion	22
6.2. Optimierung der Leistung des Klassifizierers	23
6.2.1. Netzarchitektur	23
6.2.2. Studie der Regionen	24
6.2.3. Auswahl der Variablen	25
6.3. Evaluierung der Klassifizierung	26

6.4.	Evaluierung in physikalischem Kontext	27
6.4.1.	Methodik	27
6.4.2.	Asimov-Tests	28
7.	Ergebnisse	31
7.1.	Mehrwert der Rekonstruktionsvariablen	31
7.2.	Evaluierung der Variablensätze	32
7.3.	Evaluierung in physikalischem Kontext	34
7.3.1.	Unsicherheiten der Anpassung	34
7.3.2.	Ausschlussgrenze	35
7.3.3.	Korrelation bei zweidimensionaler Anpassung	35
7.4.	Fazit über Wahl der Regionen	38
8.	Zusammenfassung und Ausblick	39
	Literaturverzeichnis	41
	Anhang	45
A.	Verteilungen der Variablen aus der χ^2 -Rekonstruktion der Bosonen	45
B.	Übersicht der Variablensätze	51
C.	Unsicherheiten mit internen Bezeichnungen	51
D.	Ergebnisse der Asimov-Tests	53

Abbildungsverzeichnis

2.1.	Feynman-Graph der $t\bar{t}$ +Z-Produktion in führender Ordnung	5
2.2.	Feynman-Graphen der $t\bar{t}$ +H- sowie $t\bar{t}$ + $b\bar{b}$ -Prozesse in führender Ordnung .	6
3.1.	Schnitt des CMS-Detektors	8
3.2.	Beispiel der Verteilung einer Variablen	10
4.1.	Exemplarische Darstellung der Diskriminatoren	14
6.1.	Vergleich der Verteilungen für χ^2 und $\ln(\chi^2)$	22
6.2.	Exemplarische Darstellung einer ROC-Kurve und einer Verwechslungsmatrix	28
7.1.	Übersicht der ROC-Integralwerte für binäre Trennung	32
7.2.	Übersicht der ROC-Integralwerte für Multiklassifizierung	33
7.3.	Übersicht der Klassifizierungseffizienzen bei Multiklassifizierung	34
7.4.	Unsicherheiten des Signalstärkeparameters	36
7.5.	Erwartete Ausschlussgrenzen für den Signalstärkeparameter	37
7.6.	Übersicht der Korrelationen bei multidimensionaler Anpassung	38
A.1.	Verteilungen der Variablen aus der χ^2 -Rekonstruktion des Z-Bosons (1) . .	45
A.2.	Verteilungen der Variablen aus der χ^2 -Rekonstruktion des Z-Bosons (2) . .	46
A.3.	Verteilungen der Variablen aus der χ^2 -Rekonstruktion des Z-Bosons (3) . .	47
A.4.	Verteilungen der Variablen aus der χ^2 -Rekonstruktion des Higgs-Bosons (1)	48
A.5.	Verteilungen der Variablen aus der χ^2 -Rekonstruktion des Higgs-Bosons (2)	49
A.6.	Verteilungen der Variablen aus der χ^2 -Rekonstruktion des Higgs-Bosons (3)	50
B.1.	Übersicht der Zusammensetzung der Variablensätze	51

Tabellenverzeichnis

2.1.	Eichbosonen und das Higgs-Boson des Standardmodells und ihre Eigenschaften	4
4.1.	Generator und Größe der Datensätze vor Selektion	12
5.1.	Für die χ^2 -Rekonstruktion verwendete Massen	18
6.1.	Erfolgsrate der einfachen χ^2 -Rekonstruktion	23
6.2.	Konfiguration der ANNs für das Training	24

6.3. Übersicht der studierten Sätze von Variablen	27
C.1. Berücksichtigte Unsicherheiten mit internen Bezeichnungen	52
D.1. Erwartete Unsicherheiten und Limits nach numerischer Anpassung in Asimov- Tests	53
D.2. Unsicherheiten und Korrelation bei zweidimensionaler Anpassung an Asimov- Pseudodaten	54

1. Einleitung

Das Standardmodell der Teilchenphysik fasst die aktuelle Kenntnis über Materie und deren Wechselwirkung ausgenommen Gravitation zusammen. In zahlreichen Experimenten konnten die Vorhersagen des Standardmodells bestätigt und anhand gewonnener Erkenntnisse verfeinert werden. Mit der Entdeckung des Higgs-Bosons [1, 2] und der Bestimmung dessen Masse zu rund 125 GeV im Jahre 2012 durch die Experimente ATLAS und CMS am CERN LHC wurden alle bislang freien Parameter festgelegt und das Standardmodell vollständig etabliert. Mit dieser Erkenntnis gelang enormer Fortschritt in der Suche nach den Bausteinen des Universums.

Seitdem durchgeführte Analysen widmen sich in großem Umfang der genauen Vermessung der durch das Standardmodell vorhergesagten physikalischen Sachverhalte, insbesondere den Eigenschaften des Higgs-Bosons. Die Suche nach Anzeichen für die Existenz von physikalischen Effekten über das Standardmodell hinaus (Beyond Standard Model, BSM) rückt zunehmend in das Zentrum aktueller Forschung. Die Vermessung der Produktion eines Top-Quark-Antiquark-Systems ($t\bar{t}$) unter assoziierter Produktion eines Z-Bosons trägt gleichwohl zur Forschung zum Standardmodell und zu BSM-Theorien bei. Der Wirkungsquerschnitt der $t\bar{t}$ +Z-Produktion wird von dem Standardmodell vorhergesagt und mittels Hypothesentests untersucht. Die Messung der direkten elektroschwachen Kopplung des Z-Bosons an das Top-Quark im dominanten Produktionskanal offenbart für die Suche nach BSM-Wechselwirkungen relevante Informationen, welche beispielsweise in effektiven Feldtheorien (EFT) interpretiert werden können.

Der Wirkungsquerschnitts der $t\bar{t}$ +Z-Produktion wurde in der Vergangenheit in Multilepton-Zerfallskanälen vermessen [3–5] und dabei die für die Sicherung der Existenz des Prozesses konventionell notwendige Signifikanz von über 5σ erreicht. Der Zerfallskanal $Z \rightarrow b\bar{b}$ wurde im Rahmen der Masterarbeit von van der Linden [6] mit Daten zu $59,7 \text{ fb}^{-1}$ bei $\sqrt{s} = 13 \text{ TeV}$ des CMS-Experiments aus dem Jahr 2018 untersucht. Im Zerfallskanal $Z \rightarrow b\bar{b}$ teilt sich die $t\bar{t}$ +Z-Produktion den Phasenraum mit der $t\bar{t}$ +H-Produktion im Zerfallskanal $H \rightarrow b\bar{b}$. Die Prozesse weisen ähnliche Produktionswahrscheinlichkeiten und kinematische Charakteristika auf, sodass sich die Trennung in Detektordaten anspruchsvoll gestaltet. Langfristig soll durch die genaue Kenntnis des $t\bar{t}$ +Z-Prozesses eine simultane Vermessung beider Prozesse ermöglicht und unter Betrachtung der Daten des vollständigen LHC Run-II durchgeführt werden.

Die nachfolgende Arbeit präsentiert eine Studie zur Trennung der $t\bar{t}$ +Z- und $t\bar{t}$ +H-Prozesse im Hinblick auf die $t\bar{t}$ +Z-Analyse. Zu diesem Zweck werden multivariate Analysemethoden auf simulierten Ereignissen des Jahres 2017 entwickelt und erprobt. Künstliche neuronale Netze (Artificial Neural Networks, ANN) werden für die simultane Klassifizierung von $t\bar{t}$ +Z-, $t\bar{t}$ +H- sowie weiteren $t\bar{t}$ -Ereignissen mit zusätzlichen Teilchenjets eingesetzt. Den ANNs wird zu diesem Zweck eine quantifizierte Beschreibung der untersuchten Ereignisse zugeführt. Zu dieser gehört neben der Beschreibung einzelner rekonstruierter Zerfallsprodukte auch

aus einer Rekonstruktion des physikalischen Prozesses abgeleitete Variablen. Der jeweilige Einfluss wird bezüglich der Ergebnisse in Asimov-Tests untersucht und eine Grundlage für die Auswahl der Variablen für die Analyse geschaffen. Die Studie legt den Grundstein für die Messung des Wirkungsquerschnitts für $t\bar{t}+Z$ -Produktion der Messdaten des CMS-Experiments aus dem Jahr 2017 und trägt dadurch zum Fortschritt zur Kombination der Messungen der $t\bar{t}+Z$ -Produktion sowie $t\bar{t}+H$ -Produktion für den LHC Run-II bei.

Kapitel 2 stellt die Elementarteilchen sowie die fundamentalen Kräfte des Standardmodells vor. In diesem Rahmen erfolgt die Vorstellung des Phasenraums der $t\bar{t}+Z$ -Produktion und der dominanten Untergrundprozesse. Die Gewinnung der für die Analyse notwendigen Daten im CMS-Experiment werden in Kapitel 3 erläutert. Kapitel 4 präsentiert einen Überblick über den Analyseprozess am Beispiel der $t\bar{t}+Z$ -Analyse. Nach Vorstellung der Analyse des semileptonischen Zerfallskanals mit Daten aus dem Jahr 2018 in Kapitel 5 erfolgt eine detaillierte Beschreibung des Vorgehens für die Studie in Kapitel 6. Hierzu gehört die Vorstellung der angewandten Rekonstruktionsmethode sowie der unternommenen Schritte zur Optimierung des Klassifizierers. Kapitel 7 stellt die Ergebnisse der Studie vor und zieht ein Fazit hinsichtlich der verwendeten Rekonstruktionsmethode. In Kapitel 8 erfolgt eine Zusammenfassung der Arbeit sowie ein Ausblick für weitere Untersuchungen.

2. Teilchenphysikalische Grundlagen

In der experimentellen Teilchenphysik überwiegend angewandte Forschungsmethodik besteht in der Überprüfung von Vorhersagen theoretischer Modelle. Das Standardmodell der Teilchenphysik als derzeit gängiges Modell wird zu Beginn dieses Kapitels in Abschnitt 2.1 knapp vorgestellt. Nach einer Diskussion teilchenphysikalischer Prozesse und der zugrunde liegenden Phänomene in Abschnitt 2.2 wird in Abschnitt 2.3 die $t\bar{t}+Z$ -Produktion und der untersuchte Zerfallskanal vorgestellt. In Abschnitt 2.4 werden zu erwartende Untergrundprozesse erläutert.

2.1. Standardmodell der Teilchenphysik

Die heutige Kenntnis der Elementarteilchen und fundamentaler Kräfte wird zusammengefasst im Standardmodell der Teilchenphysik. Theoretische Vorhersagen sind weitestgehend vereinbar mit experimentellen Befunden. Nachfolgende Erläuterungen sind in Struktur und Inhalt an Griffiths [7] angelehnt. Nach dem Standardmodell ist alle bekannte Materie aus 61 Elementarteilchen aufgebaut, welche in oberster Gliederungsebene in drei Typen unterteilt werden können.

Leptonen sind Spin-1/2-Teilchen und existieren in drei Generationen. Jede der Generationen wird vertreten von einem elektrisch negativ geladenen Teilchen und dessen zugehörigem elektrisch neutralem Neutrino. Die geladenen Teilchen tragen die Namen Elektron, Myon und Tau-Lepton, welche gleichermaßen für die Bezeichnung der Neutrinos eingesetzt werden. Gemeinsam mit den zugehörigen Antiteilchen ergeben sich insgesamt zwölf Leptonen im Standardmodell, deren elektrische Ladungen ausnahmslos ganzzahlig sind.

Quarks werden klassifiziert durch ihren Flavor: Upness, Downness, Strange, Charm, Beauty und Truth. Analog zu den Leptonen werden jeweils zwei Flavor einer Generation zugeordnet. Jedes Quark kann eine von drei unterschiedlichen Farbladungen tragen. Aufgrund des *Confinements* der Quantenchromodynamik (QCD) [8] streben Quarks stets einen farbneutralen Zustand an und verbinden sich zu Mesonen und Baryonen [9,10]. Unter letzteren finden sich das Neutron und das Proton wieder. Die elektrische Ladung einzelner Quarks ist, im Gegensatz zu der von Leptonen, nicht ganzzahlig.

Eichbosonen bilden die Grundlage zur Beschreibung der fundamentalen Wechselwirkungen im Standardmodell. Zwei W-Bosonen, das Z-Boson sowie das Photon gehen aus spontaner Symmetriebrechung in der Quantenfeldtheorie bei Vereinigung von elektromagnetischer und schwacher Wechselwirkung hervor. Bei Interaktion von Quarks im Rahmen der starken Wechselwirkung bilden sich Gluonen als Mediatoren aus.

Das Standardmodell wird vervollständigt durch das Higgs-Boson. Als Konsequenz des Higgs-Mechanismus' [11], durch welchen den W-Bosonen und dem Z-Boson ihre Masse verliehen wird, nimmt das Higgs-Boson eine gesonderte Position im Standardmodell ein. Eine gemeinsame Aufstellung der Eichbosonen und des Higgs-Bosons findet sich in Tabelle 2.1.

Tabelle 2.1.: Eichbosonen und das Higgs-Boson des Standardmodells und ihre Eigenschaften. Die Werte der Massen sind Referenz [12] entnommen.

	Wechselwirkung	Masse	Ladung
Photon (γ)	elektromagnetisch	—	—
Gluon (g)	stark	—	Farbladung
W^\pm -Boson	elektroschwach	$(80,38 \pm 0,01) \text{ GeV}$	elektrische und schwache Ladung
Z-Boson	elektroschwach	$(91,188 \pm 0,002) \text{ GeV}$	schwache Ladung
Higgs-Boson		$(125,1 \pm 0,1) \text{ GeV}$	—

2.2. Elementarteilchendynamik

Zentrale Größe zur Charakterisierung teilchenphysikalischer Wechselwirkungsprozesse ist der Wirkungsquerschnitt σ eines spezifischen Prozesses. Er kann als Wahrscheinlichkeit für Ablauf des jeweiligen Prozesses interpretiert werden und ergibt sich aus der Berechnung des Matrixelementes für den Übergang zwischen Anfangs- und Endzustand. Mit der instantanen Luminosität L ergibt sich die Interaktionsrate zu

$$\frac{dN}{dt} = \sigma \cdot L, \quad (2.1)$$

wobei L den Fluss kollidierender Teilchen quantifiziert. Für die Diskussion der formalen mathematischen Behandlung grundlegender teilchenphysikalischer Prozesse sei verwiesen auf Griffiths [7]. Nachfolgende Erläuterungen geben eine phänomenologische Interpretation der ablaufenden Mechanismen wider.

Die Quantenelektrodynamik (QED) beschreibt jegliche Wechselwirkungen unter Austausch von Photonen. Ein simples Beispiel für einen Prozess der QED ist die Møller-Streuung, bei welcher zwei einlaufende Elektronen an dem elektrischen Potential streuen. Ferner beschreibt die QED den Effekt der Vakuumpolarisation, durch deren Einfluss in der Hochenergiephysik die elektromagnetische Kopplungsstärke variiert.

Das Pendant der QED für die starke Wechselwirkung ist in den Gesetzen der Quantenchromodynamik (QCD) zusammengefasst. Sie verlangt farbgeladene Teilchen für die gegenseitige Wechselwirkung und ermöglicht somit Modelle zur Beschreibung der Interaktion von Quarks untereinander. Da die Gluonen als Mediatoren selbst farbgeladen sind, ist eine Selbstwechselwirkung zwischen drei und mehr Gluonen möglich. Innerhalb eines Kollisionsprozesses veranlasst die QCD die zufällige Abstrahlung von Gluonen, welche anschließend zu Teilchenjets hadronisieren. Entsprechende Phänomene können wegen der relativ hohen Kopplung der starken Wechselwirkung unterhalb einer typischen Energie von etwa 200 MeV nicht störungstheoretisch behandelt werden und werden stattdessen mit phänomenologischen Modellen beschrieben.

Sämtliche weitere teilchenphysikalische Prozesse basieren auf der schwachen Wechselwirkung. Sie ermöglicht die Interaktion von Leptonen mit Quarks durch Austausch der schwach geladenen Eichbosonen und ist dadurch essentieller Bestandteil vieler tiefinelastischer Prozesse. Die Wechselwirkung eines Quarks über W-Bosonen kann eine Flavoränderung nach dem CKM-Mechanismus [13] verursachen.

2.3. $t\bar{t}+Z$ -Produktion und -Zerfall

Gegenstand der $t\bar{t}+Z$ -Analyse ist die assoziierte Produktion eines Z-Bosons unter Bildung eines Paares aus Top-Quark und eines Top-Antiquark aus Protonkonstituenten

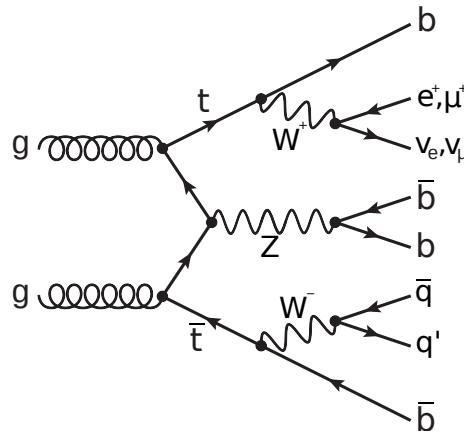


Abbildung 2.1.: Feynman-Graph der $t\bar{t}+Z$ -Produktion in führender Ordnung.
Die Abbildung ist entnommen aus [6].

bei anschließendem Zerfall in ein Paar aus Bottom-Quark und Bottom-Antiquark. Das Feynman-Diagramm der führenden Ordnung für den dominanten Anfangszustand ist in Abbildung 2.1 aufgeführt.

Für die Entstehung des $t\bar{t}$ -Systems kommen unterschiedliche Anfangszustände in Frage. Der dominante Kanal sieht einen Ausgangszustand aus zwei Gluonen vor. Die Top-Quarks besitzen eine mittlere Lebensdauer im Bereich von 10^{-25} s und ändern jeweils mit hoher Wahrscheinlichkeit unter Abstrahlung eines W-Bosons den Flavor zu einem Bottom-Quark. W-Bosonen zerfallen mit einem Anteil von 33 % in ein Paar aus Lepton und zugehörigem Neutrino. Etwa 66 % der W-Bosonen hingegen zerfallen hadronisch in ein Paar aus Quark und Antiquark. Die resultierende Wahrscheinlichkeit für jeweils einen leptonischen und einen hadronischen Zerfall der aus dem $t\bar{t}$ -Systems hervorgehenden W-Bosonen beläuft sich zu 44 %.

Im dominanten Anfangszustand aus zwei Gluonen koppelt das assoziierte Z-Boson direkt an das Top-Quark. Mit einer mittleren Lebensdauer im Bereich von 10^{-25} s zerfällt das Z-Boson in etwa 10 % der Fälle in ein geladenes Lepton und dessen Antiteilchen. In etwa 20 % der Fälle ist der Endzustand des Zerfalls eines Z-Bosons nicht messbar und wird deshalb mit zwei Neutrinos assoziiert. Etwa 70 % der Z-Bosonen zerfallen in ein Paar aus Quark und Antiquark, 15 % zerfallen explizit in ein Paar aus Bottom-Quark und Bottom-Antiquark.

Die aktuellste Berechnung des vorhergesagten Wirkungsquerschnitts für $t\bar{t}+Z$ -Produktion [14] bei einer Schwerpunktenenergie von $\sqrt{s} = 13$ TeV beläuft sich zu

$$\sigma(t\bar{t}+Z)^{\text{theo}} = 0,86_{-0,08}^{+0,07}(\text{scale}) \pm 0,03(\text{PDF} + \alpha_s) \text{ pb}. \quad (2.2)$$

Erstere Unsicherheit ergibt sich aus der Variation der Energieskala bei der perturbativen Berechnung des Wirkungsquerschnitts. Letztere entspringt der Unsicherheit auf die Partonverteilungsfunktionen (PDF) und auf die Kopplungskonstante der starken Wechselwirkung.

2.4. Untergrundprozesse

Unter Proton-Proton-Kollisionen werden in demselben Phasenraum weitere Prozesse erwartet, deren Beitrag für die Messung bekannt sein muss. Der dominante Produktionskanal des $t\bar{t}$ -Systems unter assoziierter Produktion eines Higgs-Bosons ist in führender Ordnung

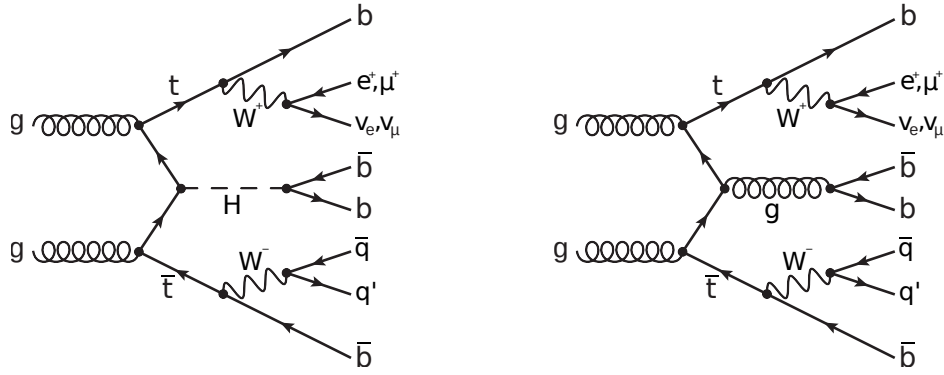


Abbildung 2.2.: Feynman-Graphen der $t\bar{t}+H$ - sowie $t\bar{t}+b\bar{b}$ -Prozesse in führender Ordnung. Der Prozess für die Produktion von $t\bar{t}+H$ ähnelt dem für $t\bar{t}+Z$ stark. Die Abbildungen sind entnommen aus [6].

in Abbildung 2.2 (links) aufgeführt und weist unverkennbare Ähnlichkeit zu dem $t\bar{t}+Z$ -Prozess in Abbildung 2.1 auf. Für die Analyse wesentliche Unterschiede bestehen in der Masse der beteiligten Resonanz und deren Spin. Der Wirkungsquerschnitt dieses Prozesses entspricht etwa dem Doppelten des Wirkungsquerschnitts für $t\bar{t}+Z$ -Produktion. Somit stellt die $t\bar{t}+H$ -Produktion unter konsekutivem Zerfall in Bottom-Quarks einen schwierig von $t\bar{t}+Z$ -Ereignissen zu trennenden Untergrund dar.

Die Gesamtheit unterschiedlicher Prozesse zur Produktion eines $t\bar{t}$ -Systems unter Emission zusätzlicher Teilchenjets leistet in der zu untersuchenden Region den größten Beitrag zur Gesamtrate. Der semileptonische Zerfallskanal des $t\bar{t}$ -Systems ist in diesem Fall weiterhin identisch zum Signalprozess. Zusätzliche Teilchenjets entstehen durch unterschiedliche Mechanismen und entwickeln sich aus der Abstrahlung beliebiger Quarks ausgenommen dem Top-Quark. Auf Grundlage der Flavor der zusätzlichen Quarks werden die Prozesse kategorisiert als die Klassen $t\bar{t}+b\bar{b}$, $t\bar{t}+c\bar{c}$ und $t\bar{t}+\text{light flavor}$. Erstere enthalten mindestens einen zusätzlichen Teilchenjet aus einem Bottom-Quark. Ereignisse mit mindestens einem zusätzlichen Teilchenjet aus einem Charm-Quark und keinem zusätzlichen Bottom-Quark werden der Klasse $t\bar{t}+c\bar{c}$ zugeordnet. Die Prozesse mit zusätzlichen Teilchenjets aus weder Bottom-Quarks noch Charm-Quarks werden als $t\bar{t}+\text{light flavor}$ -Prozesse bezeichnet. Eine exemplarische Darstellung eines $t\bar{t}+b\bar{b}$ -Ereignisses ist in führender Ordnung in Abbildung 2.2 (rechts) aufgeführt.

In der untersuchten Region sind zudem Ereignisse zahlreicher weiterer Prozesse zu erwarten. Im Verhältnis zu den bisher vorgestellten fünf Prozessklassen ist der Beitrag weiterer Prozesse zur erwarteten Gesamtrate sehr gering. Zudem sind Ereignisse dieser Prozesse im Vergleich mit hoher Effizienz von Signalereignissen zu unterscheiden und sind deshalb nahezu unbedeutend für die Entwicklung einer Analysestrategie.

3. CMS-Experiment

Zahlreiche Forschungszweige der modernen Teilchenphysik stützen sich auf die Ergebnisse von Streuexperimenten. Streuexperimente bei hohen Energien sind derzeit ein vielversprechender Ansatz zur Überprüfung von Hypothesen bezüglich des Standardmodells sowie Theorien über das Standardmodell hinaus. Die notwendigen Ausmaße werden bei genauerer Betrachtung des Large Hadron Colliders (LHC) und des Experiments Compact Muon Solenoid (CMS) deutlich.

Das Kapitel beginnt mit der Vorstellung des LHC und der Verbindung zum CMS-Experiment in Abschnitt 3.1. Der nachfolgende Abschnitt 3.2 diskutiert den Aufbau des Detektors im Detail, bevor in Abschnitt 3.3 eine kurze Einführung in die Kinematik im Detektor erfolgt. Die übrigen Abschnitte 3.4 und 3.5 widmen sich der Prozedur zur Interpretation der Messdaten und deren Aufbereitung für die Analyse.

3.1. Large Hadron Collider

Der bislang größte Teilchenbeschleuniger LHC [15] ist entworfen, um sowohl Protonen als auch Blei-Ionen zu beschleunigen und deren Strahlen in seinem System aufrechtzuerhalten. Der Beschleunigerring basiert auf dem Prinzip eines Synchrotrons und unterhält zwei Strahlen positiv geladener Teilchen für Beschleunigung in entgegengesetzter Richtung. Supraleitende Magnete halten die Strahlen in zwei isolierten Strahlrohren auf Kurs.

An den Orten der vier Experimente ALICE [16], ATLAS [17], CMS [18] und LHCb [19] werden die Strahlen innerhalb der jeweiligen Detektoren zur Kollision gebracht. Instantane Luminosität sowie Schwerpunktenergie wurden seit Inbetriebnahme stufenweise erhöht und erreichten für die Kollision von Protonen Jahr 2018 Maximalwerte von $L_{\text{inst}} = 21,40 \text{ nb}^{-1} \text{ s}^{-1}$ bei $\sqrt{s} = 13 \text{ TeV}$.

3.2. Bestandteile des CMS-Detektors

Der Name des CMS-Detektors rührt von seiner kompakten Bauweise, seiner außergewöhnlich guten Auflösung bezüglich Myonen-Spuren und dem starken Solenoid-Magneten. Er ist als Vielzweckdetektor konstruiert, der möglichst alle Produkte der unzähligen denkbaren physikalischen Prozesse am LHC auffängt. Gemeinsam mit dem LHC ging der Detektor 2008 in Betrieb und erhielt zwischen den Läufen in 2016 und 2017 eine Verbesserung des Pixeldetektors [20]. Gemeinsam mit ATLAS war das CMS-Experiment wesentlich an der Entdeckung eines Higgs-artigen Teilchens im Jahre 2012 [1] beteiligt. Struktur und Inhalte nachfolgender Erläuterungen sind an einen Bericht der CMS-Kollaboration [18] angelehnt.

Sowohl der Mantel als auch die Endflächen des zylinderförmigen Detektors sind bis auf das Strahlrohr vollständig mit Detektorelementen instrumentiert. Der Kollisionspunkt ist umgeben von Schichten aus mehreren Detektoren unterschiedlicher Funktionalität, welche

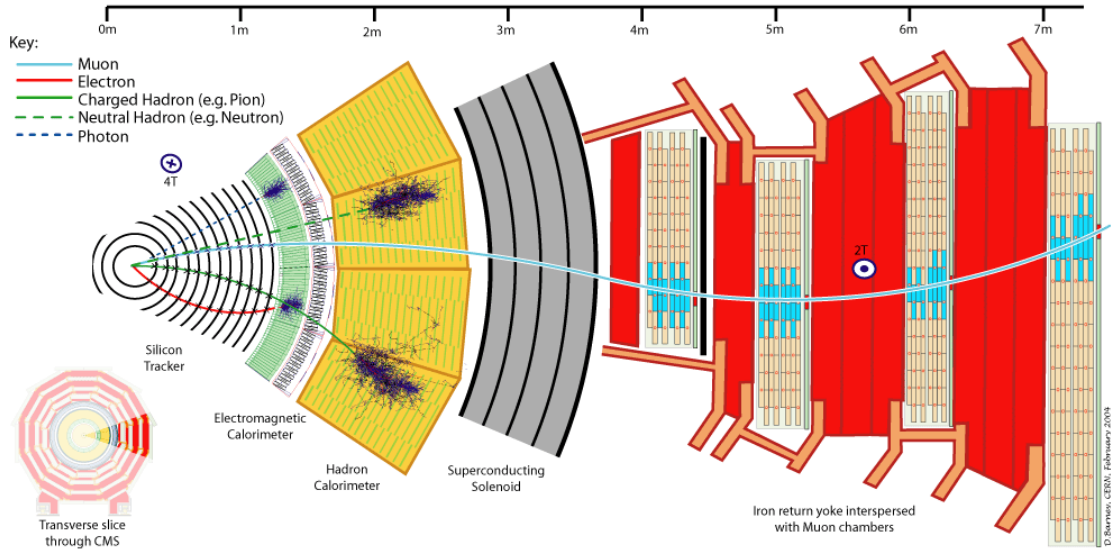


Abbildung 3.1.: Schnitt des CMS-Detektors. Abbildung unverändert übernommen [21].

jeweils auf die Detektion spezifischer Zerfallsprodukte ausgelegt ist. Der namensgebende supraleitende Magnet erzeugt Magnetfelder mit Stärken bis zu 3,8 T bei einem Spulendurchmesser von 6 m. Innerhalb des Magneten befinden sich Silizium-Pixeldetektoren und Silizium-Streifendetektoren zur Detektion des Durchgangs jeglicher elektrisch geladener Kollisionsprodukte. Der innere Spurdetektor ist umgeben vom elektromagnetischen Kalorimeter, in welchem Photonen, Elektronen und Tau-Leptonen unter Schauerbildung aufgefangen werden. Hadronische Zerfallsprodukte werden ebenfalls unter Schauerbildung im darauf folgenden hadronischen Kalorimeter aufgefangen. Aus der jeweiligen Energiedeposition in den Kalorimetern ist ein Rückschluss auf die Energie der Teilchen und Teilchenjets möglich. Außerhalb der Magnetspule befinden sich Gasionisationskammern zur Detektion von Myonen. Eine detailliertere Beschreibung des Detektors zum Zeitpunkt der Inbetriebnahme kann Referenz [18] entnommen werden. Der Schnitt des Detektors ist schematisch in Abbildung 3.1 aufgeführt.

Aufgrund der Rotationssymmetrie erweist sich eine Beschreibung von Positionen innerhalb des CMS-Detektors mit Zylinderkoordinaten als zweckmäßig. Der Azimutalwinkel ϕ beschreibt den Winkel in der Ebene orthogonal zur Strahlachse ausgehend vom Zentrum des Beschleunigerrings. Der Polarwinkel θ bezeichnet den Winkel zur Strahlachse und nimmt Werte zwischen 0 und π an. Der Ursprung des Koordinatensystems wird in dem angestrebten Kollisionspunkt gewählt.

3.3. Kinematik im CMS-Detektor

Da die Impulsanteile der Protonkonstituenten vor Kollision nicht bekannt sind, ist ein Rückschluss auf den Impuls der Kollisionsprodukte nicht möglich. Die Impulskomponente orthogonal zur Strahlachse jedoch kann durch Messung bestimmt werden. Diese Größe wird als Transversalimpuls p_T bezeichnet.

Die fehlende Transversalenergie \cancel{E}_T ist definiert als

$$\cancel{E}_T = \left| - \sum_j \vec{p}_{T,j} \right|, \quad (3.1)$$

wobei j über alle Teilchen läuft, die eine Signatur im Detektor hinterlassen. Da der Transversalimpuls der Protonen vor der Kollision annähernd 0 GeV beträgt, ist für die Summe der Transversalimpulse der Kollisionsprodukte ebenfalls annähernd 0 GeV zu erwarten. Weicht dieser wesentlich von 0 GeV ab, ist davon auszugehen, dass weitere Produkte mit der fehlenden Transversalenergie den Detektor wechselwirkungsfrei passieren. Zu diesen möglichen Produkten gehören beispielsweise Neutrinos.

Die Rapidität y eines Teilchens gilt als Maß einer relativistischen Geschwindigkeit. Sie ist definiert als

$$y = \frac{1}{2} \ln \left(\frac{E + cp_z}{E - cp_z} \right), \quad (3.2)$$

wobei E die Energie und p_z die z-Komponente des Impulses eines Teilchens bezeichnen. Bei Verwendung in der experimentellen Teilchenphysik wird y zweckmäßig entlang der Strahlachse untersucht. Der LHC operiert in Bereichen von Energien, für welche $pc \gg mc^2$ angenommen werden kann, sodass sich für y durch Näherung

$$y \approx \frac{1}{2} \ln \left(\frac{p(1 + \cos(\theta))}{p(1 - \cos(\theta))} \right) = -\ln \tan \frac{\theta}{2} =: \eta \quad (3.3)$$

ergibt. Die Pseudorapidität η ist eine Funktion des Polarwinkels θ und kann direkt aus der Messung bestimmt werden. Sie erweist sich als zweckmäßige Winkelgröße für Analysen in der experimentellen Teilchenphysik aufgrund der erhöhten Auflösung für kleine Winkel zur Strahlachse.

Die Winkelöffnung ΔR zwischen zwei Kollisionsprodukten ist definiert als

$$\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}. \quad (3.4)$$

Sie fasst das Verhältnis der Ausbreitungsrichtungen zweier an den Orten mit $\eta_{1/2}, \phi_{1/2}$ im Detektor registrierten Kollisionsprodukten zusammen.

3.4. Objektrekonstruktion

Wechselwirkungen von physikalischen Objekten mit dem Detektormaterial äußern sich in elektrischen Spannungen, welche von aufwendigen Analysesystemen mit enormer Geschwindigkeit ausgelesen werden. Für die Analyse verwertbare Information wird erst durch die Interpretation der Signale und Rekonstruktion der physikalischen Objekte gewonnen.

Im CMS-Experiment erfolgt die Rekonstruktion physikalischer Objekte mit dem eigens entwickelten Particle-Flow-Algorithmus [22]. Er kombiniert die Information aus allen Bestandteilen des Detektors und vorausgehender Spurrekonstruktion zur Berechnung der jeweiligen Viererimpulse rekonstruierter Objekte inklusive etwaiger Korrekturen. Unter Verwendung des Anti- k_t -Algorithmus [23, 24] werden im Anschluss daran aus dem Ergebnis der Particle-Flow-Rekonstruktion Teilchenjets rekonstruiert. Der Distanzparameter für den verwendeten Algorithmus wird typischerweise zu 0,4 gewählt.

Geladene Teilchen folgen in dem starken Magnetfeld einer helixförmigen Spur. Die Rekonstruktion des Kollisionspunktes [25] geschieht für jedes Teilchen durch numerische Anpassung einer helixförmigen Kurve an Ausschläge im Spurdetektor unter Berücksichtigung der Streuung an Detektormaterial und damit einhergehendem Energieverlust. Der Schnittpunkt der Spuren der jeweiligen Teilchen mit der höchsten Summe aus p_T^2 wird als der primäre Kollisionspunkt der Proton-Proton-Wechselwirkung aufgefasst.

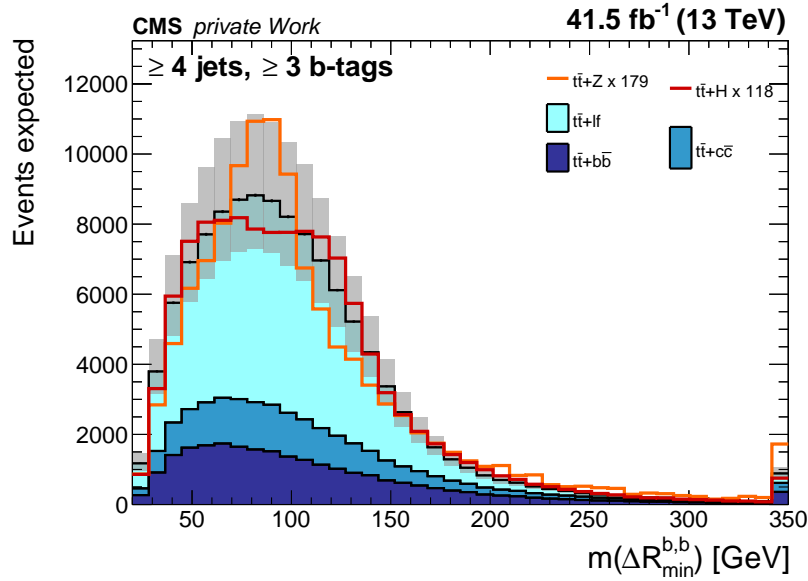


Abbildung 3.2.: Beispiel der Verteilung einer Variablen. Im Schaubild enthalten sind alle wesentlichen in der Analyse berücksichtigten Prozesse. Es zeigt die Verteilung der invarianten Masse der zwei Teilchenjets mit einem b-Tag-Wert über 0,303 und der geringsten Winkelöffnung zueinander. Die Beiträge der $t\bar{t}+Z$ - und der $t\bar{t}+H$ -Prozesse sind zur besseren Erkennbarkeit skaliert. Das Unsicherheitsband ergibt sich aus den in der Studie berücksichtigten systematischen Unsicherheiten.

Der DeepJet-Algorithmus [26] extrahiert aus den Ergebnissen der Particle-Flow-Berechnung weiterführende Information über die rekonstruierten Teilchen und Teilchenjets. Die Berechnung erfolgt durch geschickt kombinierte künstliche neuronale Netze für unterschiedliche Informationsquellen. Das Ergebnis beinhaltet unter anderem sogenannte b-Tag-Information. Als Maß für die Wahrscheinlichkeit für Entstehung aus einem Bottom-Quark wird jedem der Teilchenjets ein Wert zwischen Null und Eins zugeordnet. Aus Studien mit simulierten Daten [26] werden Schwellwerte ermittelt, bei welchen jeweils 10 %, 1 % und 0,1 % aller untersuchten Teilchenjets fälschlicherweise als Teilchenjets aus Bottom-Quarks identifiziert werden. Der mittlere Schwellwert beläuft sich für die Studie zu 0,303.

3.5. Berechnung von Ereignisvariablen

Aus den Ergebnissen der Interpretation der in Abschnitt 3.4 vorgestellten Prozedur werden charakteristische Variablen eines jeden Ereignisses berechnet. Hierbei kann beispielsweise die gesamte vom Detektor aufgenommene Energie E oder die skalare Summe der Transversalimpulse H_T bestimmt werden, welche das Ereignis als Gesamtes beschreiben. Neben allgemeinen Größen werden bei einer Vielzahl von Teilchen und Teilchenjets entsprechende Variablen für jedes Objekt gesondert berechnet und gesichert. Zu diesen gehören beispielsweise sämtliche in Abschnitt 3.3 vorgestellte Größen. Es existieren unzählige weitere Ereignisvariablen, die potentiell in einer Analyse verwendet werden können. Abbildung 3.2 zeigt exemplarisch die Verteilung einer ausgewählten Variable für alle in der Studie berücksichtigten Prozesse.

4. Analyseprozess

Die Prüfung der theoretischen Vorhersage des Wirkungsquerschnitts eines bestimmten Prozesses erfolgt durch den Vergleich einer Simulation mit der Messung mit dem Detektor. In der Messung zu erwartende Prozesse werden ausgehend von theoretischen Berechnungen simuliert und daraus vielfach entsprechend Kollisionsereignisse generiert. Für den Vergleich werden alle simulierten Ereignisse mit entsprechender Gewichtung bezüglich einer festgelegten Variable gemeinsam mit Detektordaten histogrammiert. Das Verhältnis zwischen der Simulation und der Messung wird anschließend in einer numerischen Anpassung optimiert. Der Signalstärkeparameter μ mit

$$\mu = \frac{\sigma_{\text{obs.}}}{\sigma_{\text{SM}}} \quad (4.1)$$

ist der für die Optimierung variierte Skalierungsfaktor des Signalprozesses und quantifiziert die Abweichung des beobachteten Wirkungsquerschnitts $\sigma_{\text{obs.}}$ von dem vorhergesagten σ_{SM} . Die Auswahl und Konstruktion der sogenannten Diskriminatoren, bezüglich welchen die Histogramme aufgetragen werden, ist entscheidend für das Ergebnis und die Unsicherheit der Messung. Nur unter kritischer Betrachtung von Unsicherheiten können überhaupt signifikante Aussagen getroffen und eine präzise Interpretation formuliert werden.

Am Beispiel der $t\bar{t}+Z$ -Analyse im $Z \rightarrow b\bar{b}$ -Kanal gibt dieses Kapitel eine Einführung in die einzelnen Schritte eines Analyseprozesses. Die Kenntnis des Phasenraums des untersuchten Prozesses wird für die Diskussion vorausgesetzt. Zu Beginn des Kapitels wird in Abschnitt 4.1 der Simulationsprozess und dessen Bedeutung für die Analyse auf abstrakter Ebene erläutert. Der Abschnitt 4.2 führt die zur Ereignisselektion für die $t\bar{t}+Z$ -Analyse unternommenen Schritte auf. Abschnitt 4.3 widmet sich der numerischen Anpassung der Simulation an Detektordaten. Dies beinhaltet die Konstruktion der Diskriminatoren und die Behandlung systematischer Unsicherheiten.

4.1. Simulation

Die bei Hochenergiephysik am LHC ablaufenden teilchenphysikalischen Prozesse können nur auf sehr abstrakter Ebene beobachtet und müssen anschließend auf ihren Ursprung zurückgeführt werden. Eine wesentliche Herausforderung besteht in der Zuordnung eines Ereignisses anhand der gemessenen Signatur im Detektor zu einem bestimmten Prozess. Etwaige Klassifizierungsverfahren werden auf Basis von simulierten Ereignissen entwickelt und an diesen getestet.

Die Simulation wird ausgehend von Erwartungen und Berechnungen durchgeführt, welche aus grundlegenden theoretischen Modellen abgeleitet und an die Charakteristika des jeweiligen Beschleunigers sowie Detektors angepasst werden. Zu diesen Charakteristika zählt beispielsweise die Partonverteilungsfunktionen in Abhängigkeit der Schwerpunktsenergie zur Beschreibung inelastischer Proton-Proton-Kollisionen, wie sie am LHC stattfinden. Nach

Tabelle 4.1.: Generator und Größe der Datensätze vor Selektion. Die Datensätze sind von der CMS-Kollaboration zentral generiert. Die Anzahl zu erwartender Ereignisse in Detektordaten ergibt sich aus Anwendung von Gewichten. Alle mit POWHEG generierten Datensätze enthalten NLO-Simulationen.

Prozess	Generator	Ereignisse	#Ereignisse
Signalprozesse			
$t\bar{t}+Z, Z \rightarrow q\bar{q}$	MADGRAPH5_AMC@NLO	6 161 354	810,6
$t\bar{t}+Z, Z \rightarrow ll/\nu\nu$	MADGRAPH5_AMC@NLO	5 321 218	355,8
Untergrundprozesse			
$t\bar{t}+b\bar{b}$ (semileptonisch)	POWHEG+PYTHIA 8.2	5 667 979	1161,7
$t\bar{t}$ (semileptonisch)	POWHEG+PYTHIA 8.2	102 330 938	490 657,5
$t\bar{t}+H, H \rightarrow b\bar{b}$	POWHEG+PYTHIA 8.2	4 794 126	180,6
$t\bar{t}+H, H \rightarrow \text{non-}b\bar{b}$	POWHEG+PYTHIA 8.2	5 276 848	143,1

dem Faktorisierungstheorem [27] ist die Simulation einzelner Kollisionsereignisse gegliedert in die Interaktion der Protonkonstituenten bei hohen Energien und den Zerfall sowie das Aufschauern der Teilchen bei niedrigen Energien. Die technische Umsetzung erfolgt in Ereignisgeneratoren, welche nach der Montecarlo-Methode (MC) aus Zufallszahlen einzelne Ereignisse generieren. Ferner berücksichtigt die vollständige Simulation die Wechselwirkung mit dem Detektormaterial und erzeugt Signaturen, welche anschließend mit Software analog zu der des Detektors interpretiert werden. Dieser Schritt beinhaltet gleichwohl die Abfrage von Triggern, welche während der tatsächlichen Messung nach vorgegeben Kriterien eingesetzt werden, um uninteressante Ereignisse zu verwerfen.

Die zu simulierenden Prozesse werden anhand der Erwartungen in der zu untersuchenden Region ausgewählt und dem Generator vorgegeben. Abhängig von dem vorgegebenen Prozess werden unterschiedlich viele Ordnungen störungstheoretischer Korrekturen berücksichtigt. Beispielsweise werden Simulationen bis zu einer Ordnung über die führende hinaus mit NLO (next-to-leading order) gekennzeichnet. Auch nach der Simulation werden vereinzelt Korrekturen an den Ereignissen in Form von theoretisch und experimentell motivierten Skalierungsfaktoren vorgenommen, um das Verhältnis von Simulation zu Realdaten zu verbessern. Zu diesem Zweck werden in den Daten Regionen und Variablen untersucht, welche zwar ausreichend Information über die Qualität der Simulation liefern, jedoch keinen tiefen Einblick in die tatsächlich für die Anpassung verwendete Region gewähren, sodass die spätere Analyse unvoreingenommen durchgeführt werden kann.

Die CMS-Kollaboration führt die Simulation von Ereignissen zentral organisiert durch. Dafür kommen je nach zu untersuchendem Prozess unterschiedliche MC-Ereignisgeneratoren zum Einsatz. Vielfach verwendete Generatoren sind beispielsweise POWHEG [28] und MADGRAPH5_AMC@NLO [29]. Alle für die Studie verwendeten Datensätze enthalten Ereignisse aus Simulationen des Detektors im Jahr 2017 bei einem Betrieb des LHC bei einer Schwerpunktsenergie von $\sqrt{s} = 13$ TeV. Der verwendete Generator und die Größe der Datensätze sind in Tabelle 4.1 zusammengefasst.

4.2. Ereignisselektion

Durch die Selektion von Ereignissen anhand theoretisch motivierter Kriterien werden Untergrundereignisse in großem Maß entfernt. Der dadurch erhöhte Anteil der Signalereignisse an der gesamten Rate erhöht die Sensitivität der Messung. Entsprechend der Ausführungen in Abschnitt 2.3 wird für $t\bar{t}+Z$ -Produktion mit anschließendem Zerfall $Z \rightarrow b\bar{b}$ und

semileptonischen Zerfällen des $t\bar{t}$ -Systems im Detektor ein Elektron oder Myon erwartet. Dieses Kriterium legt in erster Instanz die untersuchte Region fest. Die eingesetzten Trigger prüfen auf die Anwesenheit von

- mindestens einem Elektron mit $p_T \geq 32 \text{ GeV}$ oder $p_T \geq 28 \text{ GeV}$ während $|\eta| \leq 2,1$ und einer Transversalenergie für den gesamten Prozess von $H_T \geq 150 \text{ GeV}$, oder alternativ
- mindestens einem isolierten Myon mit $p_T \geq 24 \text{ GeV}$.

Aus dem leptonischen Zerfall des W-Bosons wird ein Neutrino erwartet, sodass ausschließlich Ereignisse mit einer fehlenden Transversalenergie $\cancel{E}_T \geq 20 \text{ GeV}$ selektiert werden.

Die zu untersuchende Region wird weiter eingegrenzt durch die Abfrage von mindestens vier Teilchenjets, wovon mindestens drei der Jets einen b-Tag-Wert von mindestens 0,303 aufweisen. Aus dem Anti- k_t -Algorithmus hervorgehende Teilchenjets werden nur als solche gewertet, wenn deren Transversalimpuls $p_T \geq 20 \text{ GeV}$ und deren Pseudorapidität $|\eta| \leq 2,1$ ist.

4.3. Numerische Anpassung

Zu den untersuchten Prozessen existieren zahlreiche Ereignisvariablen, deren Verteilungen nur geringe Unterschiede zwischen Signal- und Untergrundprozessen aufweisen. Die Unterschiede in den Verteilungen sind im Allgemeinen zu gering, um in einer numerischen Anpassung der Simulation an Detektordaten signifikante Aussagen über den Beitrag der Signalereignisse treffen zu können. Aus diesem Grund erfolgt für die $t\bar{t}+Z$ -Analyse wie für viele weitere Analysen in Experimenten am LHC eine Konstruktion speziell ausgelegter künstlicher Variablen, welche Signal- und Untergrundprozesse bestmöglich trennen. Die Verteilungen dieser sogenannten Diskriminatoren können anschließend an die Detektordaten angepasst werden und so in der Regel die Sensitivität auf den Signalstärkeparameter wesentlich erhöht werden. Im Rahmen der $t\bar{t}+Z$ -Analyse erfolgt die Anpassung mehrerer Verteilungen gleichzeitig, um die Sensitivität weiter zu erhöhen und potentiell den Einfluss einzelner Quellen von Unsicherheiten zu begrenzen.

Die Parameterschätzung sowie die Extraktion der 1σ - und 2σ -Unsicherheitsintervalle geschieht unter Anwendung der Maximum-Likelihood-Methode [30]. Das Minimum des negativen Logarithmus der Likelihood-Funktion legt den Wert des Signalstärkeparameters fest, während die Gestalt des Minimums Aufschluss über die Unsicherheiten gibt. Die technische Umsetzung erfolgt in dem Framework `combine` [31].

Für die Konstruktion der Diskriminatoren bietet sich eine multivariate Analyse-methode an, welche die Gesamtheit der Unterschiede in den Verteilungen von Ereignisvariablen kombiniert und dadurch die Konzentration der Information erhöht. Die Rückgabewerte eines Artificial Neural Networks (ANN) können direkt als Diskriminator eingesetzt werden.

4.3.1. Neuronales Netz als Klassifizierer

Künstliche neuronale Netze sind dem Aufbau eines komplexen Nervensystems nachempfunden. Wesentlicher Bestandteil sind beliebig viele einzelne Neuronen, welche in der Verarbeitung von Informationen als Knoten fungieren. Informationen werden in numerischer Form in das Netz eingespeist und entlang Verbindungen zwischen den Neuronen propagiert. Das Resultat wird in numerischer Form zurückgegeben und muss entsprechend des Anwendungsfalles interpretiert werden. Die Gewichtung der Verbindungen zwischen Neuronen wird in einem Prozess des überwachten Lernens auf Grundlage von Beispieldaten

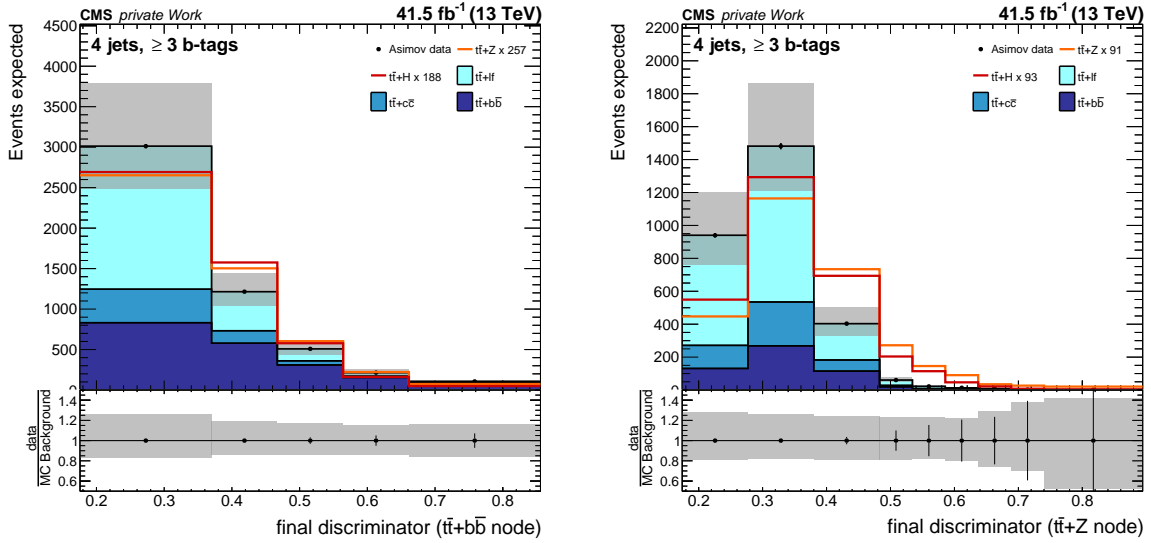


Abbildung 4.1.: Exemplarische Darstellung der Diskriminatoren. Der jeweils höchste Rückgabewert des ANN wird im jeweiligen Histogramm aufgetragen. Die Abbildungen zeigen die finale Verteilung der Werte an dem Neuron des $t\bar{t}+b\bar{b}$ -Prozesses (links) und des $t\bar{t}+Z$ -Prozesses (rechts). Jeweils enthalten sind Ereignisse aller fünf in der Studie berücksichtigten Prozesse, wobei $t\bar{t}+Z$ - und $t\bar{t}+H$ -Ereignisse auf das Integral der übrigen skaliert sind. Das Fehlerband ergibt sich aus den berücksichtigten systematischen Unsicherheiten. Die dargestellten Datenpunkte entstammen der Simulation.

angepasst und optimiert. Für eine ausführliche Diskussion der Umsetzung von neuronalen Netzen in elektronischen Rechensystemen sei auf Lämmel [32] verwiesen.

Die ANNs werden dazu ausgelegt, anhand der Werte der Ereignisvariablen eines jeden Ereignisses ein Maß für die Zuordnung zu einem bestimmten Prozess zu berechnen. Die aufzulösenden Prozesse sind davor festzulegen und jeweils Beispiele zu hinterlegen, welche dem Netz für das Training zugeführt werden. Da für Realdaten keine sichere Zuordnung existiert, stehen für diesen Schritt ausschließlich simulierte Ereignisse zur Auswahl, deren zugrunde liegende Prozessklasse per Konstruktion bekannt ist. Bei der Trennung von mehreren Prozessen in demselben Netz liefert das ANN einen Vektor, dessen Komponenten jeweils das Maß für die Zuordnung des Ereignisses zu jedem der potentiellen Prozessklassen enthalten. Die Zuordnung zu einer Prozessklasse erfolgt nach dem höchsten Wert unter allen Komponenten des Rückgabektors. Dieser höchste Wert wird in dem Histogramm der jeweiligen Prozessklasse aufgetragen. Es ergeben sich daraus bei K aufzulösenden Prozessen K Diskriminatoren, in deren Gesamtheit jedes Ereignis nur einfach berücksichtigt wird. Die gewonnenen K Verteilungen eignen sich für die numerische Anpassung. Eine exemplarische Darstellung dieser Auftragung ist in Abbildung 4.1 aufgeführt.

Die Auswahl der zu klassifizierenden Prozesse wird unter Berücksichtigung der jeweiligen vorhergesagten Beiträge zu den Detektordaten getroffen. Signalprozesse werden sinnvollerweise auch bei geringem Beitrag als zu klassifizierenden Prozess aufgenommen.

4.3.2. Integration systematischer Unsicherheiten

Die Teilchenphysik widmet sich dem Studium von Quantenobjekten, welche per Definition ausschließlich mit probabilistischen Mitteln beschrieben werden. Bei mehrfacher Wiederholung derselben Messung sind stets Streuungen um das wahre Ergebnis zu erwarten. Die Streuung wird zusammengefasst als statistische Unsicherheit, welche durch Erhöhung der Menge an Daten reduziert werden. Aus diesem Grund ist der LHC darauf ausgelegt, möglichst viele Kollisionsereignisse in möglichst kurzer Zeit stattfinden zu lassen.

Systematische Unsicherheiten folgen aus der ungenauen Kenntnis der Messapparatur sowie physikalischer Größen. Im Fall des CMS-Detektors zählen hierzu beispielsweise eine begrenzte Genauigkeit bei der Kalibrierung des hadronischen Kalorimeters, sodass sich Unsicherheiten bezüglich der Energie der aufgefangenen Teilchenjets ergeben. Die Wirkungsquerschnitte potentieller Untergrundprozesse haben direkten Einfluss auf die Inhalte der Diskriminatoren, sodass deren Unsicherheit ebenfalls berücksichtigt werden muss.

Unsicherheiten unterschiedlichen Ursprungs nehmen unterschiedlichen Einfluss auf die Verteilungen der Diskriminatoren. Für die korrekte Integration in der Analyse wird die Größe sowie das Verhalten einzelner Unsicherheiten in isolierten Studien abgeschätzt. In diesen Studien werden Einflüsse getrennt nach Unsicherheiten, welche die Verteilungen der Diskriminatoren in der Normierung variieren und Unsicherheiten, welche potentiell Einfluss auf die Gestalt der Diskriminatoren nehmen. Erstere werden im Kontext der Parameterschätzung mittels Likelihood-Methode als Ratenunsicherheiten, letztere als Gestaltunsicherheiten referenziert. Störfaktoren, welche sowohl die Form als auch die Rate variieren, werden nach Möglichkeit in Anteile separiert, welche wiederum jeweils einer der beiden Kategorien zugeordnet werden können.

Jegliche Unsicherheiten fließen in die Parameterschätzung mit der Likelihood-Methode als Störparameter in der Zielfunktion der Anpassung ein. Die Werte der Störparameter nach der Anpassung sind von geringem Interesse, finden jedoch Beachtung bei Evaluierung der Anpassung hinsichtlich ihrer Plausibilität. Für eine ausführliche Diskussion der verwendeten Methoden sei verwiesen auf die Beschreibung der $t\bar{t}+H$ -Analyse der CMS-Kollaboration [33], an welcher sich die $t\bar{t}+Z$ -Analyse orientiert. Die in bestehender $t\bar{t}+Z$ -Analyse [6] für den Zerfallskanal $Z \rightarrow b\bar{b}$ berücksichtigten Unsicherheiten sind im Wesentlichen deckungsgleich. Für die Studien werden anhand des Beitrags zum Rechenaufwand und des potentiellen Beitrags zur resultierenden Unsicherheiten einzelne bedeutende Quellen ausgewählt. Eine knappe Beschreibung und eine Aufstellung unter Nutzung der internen Bezeichnungen der berücksichtigten Unsicherheiten ist in Anhang C aufgeführt.

5. Bisheriger Stand

Das Vorgehen und die Ergebnisse der Untersuchung des Wirkungsquerschnitts für die $t\bar{t}+Z$ -Produktion im Zerfallskanal $Z \rightarrow b\bar{b}$ [6] für die im Jahr 2018 vom CMS-Detektor aufgezeichneten Daten dient als Basis für nachfolgend durchgeführte Studien. Aus diesem Grund werden in diesem Kapitel die angewandte Methodik zur Trennung von Signal und Untergrundereignissen sowie die Ergebnisse explizit besprochen.

Zu Beginn werden die eingesetzte Rekonstruktionsmethode und die Konstruktion der verwendeten Diskriminatoren in Abschnitt 5.1 rekapituliert. Anschließend wird in Abschnitt 5.2 der Einfluss der berücksichtigten Unsicherheiten kurz diskutiert und die bestimmten Ergebnisse werden in Abschnitt 5.3 vorgestellt. Das Kapitel schließt mit dem Ausblick für das weitere Vorgehen in Abschnitt 5.4.

5.1. Diskriminatoren

Wie in Abschnitt 4.3.1 vorgestellt, werden die Diskriminatoren für die numerische Anpassung aus neuronalen Netzen gewonnen. Die wichtigsten Stellschrauben zur Optimierung der Leistung ist die Architektur der ANNs sowie die Auswahl der zugeführten Variablen. Zur Verfügung stehende Variablen sind Kinematiken der in der Particle-Flow-Rekonstruktion bestimmten physikalischen Objekte, die b-Tag-Information der Teilchenjets und weitere Ereignisvariablen wie beispielsweise die fehlende Transversalenergie und die gesamte vom Detektor aufgenommene Energie. Zur Bündelung der für die Klassifizierung relevanten Information wird die Rekonstruktion der am $t\bar{t}+Z$ -Prozess beteiligten Objekte durchgeführt.

5.1.1. Ereignisrekonstruktion

Die Leistung des Klassifizierers soll durch explizit für den untersuchten Prozess entworfene Variablen erhöht werden. Zu diesem Zweck eingeführte Variablen basieren auf der Rekonstruktion physikalischer Objekte wie beispielsweise den Top-Quarks, welche Teil der aufzulösenden Ereignisse sind und nicht direkt im Detektor gemessen werden. Werden alle aus dem Zerfall des jeweiligen Objektes stammende Teilchenjets im Detektor registriert, kann aus den gemessenen Größen Information über das ursprüngliche Objekt gewonnen werden. Zu dieser Information gehören beispielsweise die ursprüngliche Energie, die Masse und Ausbreitungsrichtung der zerfallenen Top-Quarks sowie des Z-Bosons. Desweiteren ist in vielen Fällen die relative Winkelverteilung ΔR der Zerfallsprodukte interessant. Die wesentliche Herausforderung in der Rekonstruktion von $t\bar{t}+Z$ -Ereignissen besteht in der Zuweisung der Jets zu den jeweiligen ursprünglichen Teilchen, sodass eine genaue Kenntnis der Topologie des Prozesses essenziell für die Entwicklung der Rekonstruktionsmethode ist.

Durch Forderung eines einzelnen Leptons ist dieses pro Ereignis einzigartig, sodass dessen Ursprung für die Rekonstruktion als eindeutig angenommen wird. Für die Zuweisung der übrigen Teilchenjets wird aus Minimierung eines χ^2 -Maßes eine Zuweisungshypothese

Tabelle 5.1.: Für die χ^2 -Rekonstruktion verwendete Massen. Die Werte stammen aus der Dokumentation zur bestehenden Analyse [6]. Sie werden durch die Anpassung einer Gauß-Funktion an die simulierten Objekte gewonnen. Die Masse ergibt sich aus dem Mittelwert, die Breite der Verteilung wird als Unsicherheit übernommen.

Objekt	$m^{\text{exp}} \text{ (GeV)}$	$\sigma^{\text{exp}} \text{ (GeV)}$
t	173,1	18,5
W	83,7	11,9
H	116,3	17,6
Z	86,8	13,5

generiert. Für jede Permutation der Teilchenjets wird das potentielle ursprüngliche Objekt i mit Masse $m_i^{\text{reco.}}$ rekonstruiert und der χ^2 -Wert als Maß für die Abweichung zur Masse $m_i^{\text{exp.}}$ des erwarteten Objekts berechnet zu

$$\chi_i^2 = \frac{(m_i^{\text{reco.}} - m_i^{\text{exp.}})^2}{(\sigma_i^{\text{exp.}})^2}. \quad (5.1)$$

Der Index i steht für ein beliebiges am Prozess beteiligtes physikalisches Objekt, wie beispielsweise ein Boson. Die Breite der Verteilung der erwarteten Massen geht als $\sigma^{\text{exp.}}$ in die Berechnung mit ein. Bei der Permutation mit dem geringsten Wert für χ^2 wird davon ausgegangen, dass die bestmögliche Zuweisungshypothese gefunden wurde. Die bisher für die Berechnung verwendeten Werte der erwarteten Massen $m_i^{\text{exp.}}$ sind in Tabelle 5.1 zusammengefasst und werden aus der Verteilung der Massen der generierten Objekte der Simulation gewonnen [6]. Die Abweichung dieser Werte von den vergleichsweise genau bekannten Massen [12] ist den komplexen Hadronisierungsprozessen der Quarks und Detektoreffekten geschuldet.

Für die Analyse werden für jedes Ereignis sämtliche am Prozess $t\bar{t}+Z$, $Z \rightarrow b\bar{b}$ beteiligte Objekte rekonstruiert. Dies beinhaltet die Zuordnung von Teilchenjets aus Bottom-Quarks zum Zerfall der Top-Quarks sowie dem Zerfall des Z-Bosons und die Zuordnung von Teilchenjets aus beliebigen Quarks aus dem hadronischen Zerfall des W-Bosons. Für die Zuweisung von Teilchenjets ohne b-Tag zu einem Jet aus einem Bottom-Quark ist ein Mechanismus integriert, der abhängig von der Gesamtzahl an gemessenen Teilchenjets unterschiedlich viele Fehlzuweisungen zulässt. Die Abweichung χ_{total}^2 wird berechnet als Summe der χ^2 für alle erwarteten Objekte nach

$$\chi_{\text{total}}^2 = \chi_{\text{had}}^2 + \chi_{\text{lep}}^2 + \chi_{\text{W}_{\text{had}}}^2 \left(+\chi_Z^2 \right). \quad (5.2)$$

Die Klammer um den letzten Term in Gleichung (5.2) kennzeichnet eine Differenzierung bei der Berechnung für Ereignisse mit unterschiedlich vielen rekonstruierten Teilchenjets. χ_{total}^2 verlangt in vollständiger Ausführung insgesamt sechs isolierte Teilchenjets. Signalereignisse werden jedoch auch in Regionen mit weniger als sechs Jets erwartet, sodass die Berechnung in diesem Fall nicht möglich ist und sich die Rekonstruktion auf die Rekonstruktion der Top-Quarks und des W-Bosons beschränkt. Für Ereignisse mit mindestens sechs Jets wird χ_{total}^2 vollständig berechnet und zusätzlich das Z-Boson rekonstruiert. Bei der Rekonstruktion wird für tatsächliche $t\bar{t}+Z(b\bar{b})$ -Ereignisse ein kleinerer Wert für χ_{total}^2 erwartet, als für Untergrundprozesse, sodass dieser Wert selbst Anhaltspunkt für die Klassifizierung bildet.

5.1.2. Auslegung der ANNs

Abhängig von der Anzahl der gemessenen Teilchenjets stehen beispielsweise durch die χ^2 -Rekonstruktion unterschiedliche Variablen zur Beschreibung des Ereignisses zur Verfügung.

In bestehender Analyse [6] werden sämtliche Ereignisse nach Anzahl der enthaltenen Teilchenjets kategorisiert und diese separat einem speziell ausgerichteten ANN für die Klassifizierung zugeführt. Die Aufteilung erfolgt in Ereignisse mit genau vier Teilchenjets, genau fünf Teilchenjets und in Ereignisse mit sechs und mehr Teilchenjets. Die ANNs werden jeweils gespeist mit unterschiedlichen Ereignisvariablen, besitzen jeweils zwei verdeckte Schichten mit jeweils 50 Neuronen und eine Ausgangsschicht von identischem Aufbau. Weitere Informationen bezüglich des Trainings und des Aufbaus der Netze sind in Tabelle 6.2 aufgeführt. Die Rückgabewerte der ANNs werden separat aufgetragen und die Informationen der Verteilungen der insgesamt $3 \cdot K$ Diskriminatoren bei K aufzulösenden Prozessen erst in der numerischen Anpassung kombiniert. Für die Analyse werden den ANNs die in Abschnitt 2.3 und Abschnitt 2.4 explizit vorgestellten fünf Prozesse vorgegeben.

5.1.3. Variablenselektion

Zur Auswahl der letztendlich verwendeten Variablen werden für jede Region zwanzigfach Netze mit allen zur Verfügung stehenden Variablen trainiert. Die Gewichte zu den jeweiligen Neuronen werden nach dem Training jeweils aufsummiert und dadurch ein Maß für den Einfluss auf die Klassifizierung gewonnen. Für jede Region werden die 30 Variablen mit den durchschnittlich höchsten Gewichten ausgewählt und in weiteren Tests fortführend selektiert. Für eine Aufstellung und Diskussion der für die Analyse verwendeten Variablen sei verwiesen auf die Dokumentation der bestehenden Analyse [6].

5.1.4. Bingrenzen

Für die Auftragung der Diskriminatoren werden Bingrenzen zwischen dem maximalen und minimalen erhaltenen Rückgabewert der ANNs gewählt. Die untere Grenze liegt per Definition oberhalb von $1/K$, während die obere Grenze sich üblicherweise zwischen 0,8 und 0,9 befindet. Die Verteilungen bezüglich der als $t\bar{t}+c\bar{c}$ und $t\bar{t}+\text{light flavor}$ klassifizierten Ereignisse werden zur Begünstigung der Eingrenzung systematischer Unsicherheiten jeweils in einem einzelnen Bin zusammengefasst. Die Verteilungen der übrigen Prozesse werden mit höchstens acht Bins im mittleren Wertebereich jeweils etwas feiner aufgelöst, als in Wertebereichen nahe der oberen sowie unteren Grenze.

5.2. Unsicherheiten

Die Analyse berücksichtigt eine Vielzahl theoretisch und experimentell motivierter Unsicherheiten. Den größten Einfluss auf das nachfolgend vorgestellte Ergebnis entstammt der Unsicherheit auf die Normierung der Untergrundraten sowie der Unsicherheit auf die Messung der Energie von Teilchenjets und der damit verbundenen Auflösung. Ferner leistet aus b-Tag-Information hervorgehende Unsicherheit einen nicht zu vernachlässigbaren Beitrag. Eine Diskussion und Aufführung aller berücksichtigten Unsicherheiten ist in der Dokumentation bestehender Analyse [6] zu finden.

5.3. Ergebnis

Die numerische Anpassung der Simulation an aus der Simulation generierten Asimov-Pseudodaten ergibt einen Signalstärkeparameter für die $t\bar{t}+Z$ -Produktion von

$$\mu(t\bar{t}+Z)^{\text{exp.}} = 1,00 \pm 1,19.$$

Die obere Ausschlussgrenze ergibt sich zu

$$\mu(t\bar{t}+Z)^{\text{exp.}} < 2,32^{+1,03}_{-0,67} \text{ (95 \%CL)}.$$

Die numerische Anpassung der Simulation erfolgt an Daten einer integrierten Luminosität von $59,7 \text{ fb}^{-1}$ bei einer Schwerpunktsenergie von $\sqrt{s} = 13 \text{ TeV}$ aus dem Lauf des LHC im Jahr 2018. Das beobachtete Ergebnis für den Signalstärkeparameter der $t\bar{t}+Z$ -Produktion beläuft sich zu

$$\mu(t\bar{t}+Z)^{\text{obs.}} = -0,77^{+1,19}_{-1,26}.$$

Das 2σ -Intervall der Unsicherheit enthält den erwarteten Wirkungsquerschnitt des Standardmodells. Die beobachtete Ausschlussgrenze für den Signalstärkeparameter liegt bei

$$\mu(t\bar{t}+Z)^{\text{obs.}} < 1,90 \text{ (95 \%CL)}.$$

5.4. Ausblick

In Kombination mit der Analyse leptonischer Zerfallskanäle des Z-Bosons [4, 5] leistet die Messung der $t\bar{t}+Z$ -Produktion im Zerfallskanal $Z \rightarrow b\bar{b}$ einen potentiell relevanten Beitrag zur Erhöhung der Präzision. Die Signifikanz der Ergebnisse soll durch Kombination mit einer Messung mit den in den Jahren 2016 ($35,9 \text{ fb}^{-1}$) und 2017 ($41,5 \text{ fb}^{-1}$) während des Betriebs des LHC bei 13 TeV mit dem CMS-Detektor aufgezeichneten Daten erhöht werden. Zu diesem Zweck erfolgen im weiteren Verlauf Untersuchungen zu Verbesserung der Signifikanz für Messung der Daten aus dem Jahr 2017.

6. Entwicklung der Diskriminatoren

Ausgehend von dem Vorgehen für die Analyse der Daten aus dem Lauf des LHC im Jahr 2018 werden die Diskriminatoren für Analyse der Daten aus dem Jahr 2017 vorbereitet. Zu diesem Zweck erfolgt eine Studie des Einflusses der Auswahl der Variablen auf die Leistung des Klassifizierers und die Wahl der für die Anpassung verwendeten Diskriminatoren auf die zu erwartenden Ergebnisse für den Signalstärkeparameter. Die bestehende χ^2 -Methode wird ergänzt um eine auf die Rekonstruktion der Bosonen konzentrierte Version, welche unabhängig von der Anzahl der Teilchenjets durchgeführt werden kann. Die zusätzlichen Variablen können die Trennung von $t\bar{t}+Z$ - und $t\bar{t}+H$ -Ereignissen verbessern.

Dieses Kapitel stellt zu Beginn in Abschnitt 6.1 eine gegenüber früherer Analyse ergänzte Rekonstruktionsmethode im Detail vor. Im Anschluss daran wird in Abschnitt 6.2 das Vorgehen zum Studium der Einflüsse auf die Diskriminatoren diskutiert. Zum Vergleich werden diese sowohl hinsichtlich der Leistung des Klassifizierers als auch in physikalischem Kontext evaluiert. Die Vorstellung der jeweiligen Methodik erfolgt in Abschnitt 6.3 und Abschnitt 6.4.

6.1. χ^2 -Rekonstruktion des Bosons

In Ereignissen mit weniger als sechs isolierten Teilchenjets erfolgt unter der in Abschnitt 5.1.1 vorgestellten Methode ausschließlich eine Rekonstruktion des $t\bar{t}$ -Systems. Die Rekonstruktion liefert in dieser Form für die jeweiligen Ereignisse keine Information bezüglich potentieller Bosonen als Mediatoren, welche eine Trennung dieser begünstigen könnte. Die ausschließliche Rekonstruktion des Bosons soll explizit die Trennung von $t\bar{t}+Z$ - und $t\bar{t}+H$ -Ereignissen verbessern.

Sowohl mit der Masse des Z-Bosons als auch der Masse des Higgs-Bosons wird eine Rekonstruktion nach Gleichung (5.1) durchgeführt. Von der Berechnung werden Permutationen ausgeschlossen, bei welchen Teilchenjets als mutmaßliche Jets aus Bottom-Quark identifiziert werden, welche einen b-Tag-Wert kleiner als 0,303 besitzen. Diese Methode eignet sich für alle Prozesse in der für die Analyse ausgewählten Region, da bei mindestens vier Teilchenjets und darunter mindestens drei Teilchenjets mit b-Tag immer eine bestmögliche Permutation gefunden werden kann. Demnach liegen für alle Ereignisse in diesem Bereich Werte für die abgeleiteten Variablen vor, welche in der Studie und späteren Analysen berücksichtigt werden können.

Bei Ereignissen, an welchen tatsächlich ein Z-Boson beteiligt ist, ist für das kleinste $(\chi^2)_{\text{Higgs}}^{\min}$ ein insgesamt größerer Wert zu erwarten als für $(\chi^2)_Z^{\min}$ und umgekehrt.

6.1.1. Abgeleitete Variablen

Für das rekonstruierte Boson werden aus dem Viererimpuls die Energie E , die Masse M und der Transversalimpuls p_T des Bosons abgeleitet. Die Ausbreitungsrichtung des Bosons wird

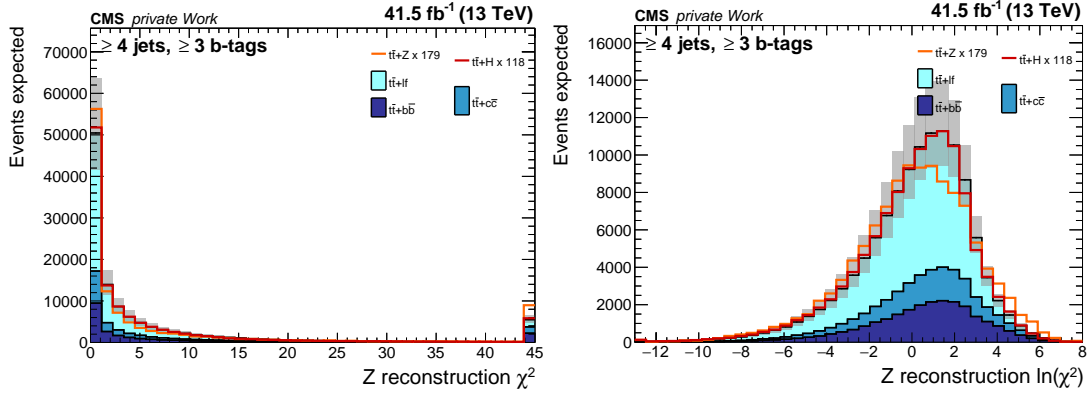


Abbildung 6.1.: Vergleich der Verteilungen für χ^2 und $\ln(\chi^2)$. In der Verteilung für χ_Z^2 (links) der verwendeten simulierten Datensätze unterscheiden sich die unterschiedlichen Prozessklassen nur marginal. Nach Anwendung des Logarithmus (rechts) sind Maximum und Mittelwert für $\ln(\chi_Z^2)$ deutlich unterschiedlich für unterschiedliche Prozesse. Die aufgetragenen Unsicherheiten entspringen den in Anhang C aufgeführt Störparametern.

mit dem Azimutalwinkel ϕ und dem Polarwinkel θ quantifiziert. Aus letzterem wird durch Transformation zusätzlich die Pseudorapidität η gewonnen. Die gleichen Informationen wird für beide zum Zerfall des Bosons assoziierten Teilchenjets extrahiert. Zusätzlich werden die Differenzen der Azimutalwinkel $\Delta\phi$ und der Pseudorapiditäten $\Delta\eta$, sowie die relative Winkelöffnung nach ΔR Gleichung (3.4) bestimmt. Ferner erfolgt eine Berechnung des Kosinus der Winkeldifferenz $\Delta\theta$.

Verteilungen von Variablen mit weiten Ausläufern zu hohen Werten und hohe Einträge für geringe Werte können durch Anwenden des Logarithmus besser aufgelöst werden. In Abbildung 6.1 sind die entsprechenden Verteilungen für χ^2 für das Z-Boson vor und nach Logarithmieren exemplarisch aufgetragen. Die Prozesse für Produktion von $t\bar{t}+Z$ und von $t\bar{t}+H$ sind hier als Signalprozesse aufgetragen und weisen nach Logarithmieren einen erkennbaren Unterschied auf, nach welchem das Maximum der Verteilung für $t\bar{t}+Z$ - bei kleineren Werten liegt als für $t\bar{t}+H$ -Produktion. Ein ähnlicher Effekt stellt sich für die Masse M des jeweiligen rekonstruierten Bosons ein.

6.1.2. Erfolgsrate der χ^2 -Rekonstruktion

Für die Überprüfung der Zuweisungshypothese aus der χ^2 -Rekonstruktion werden die den Bosonen zugewiesenen Teilchenjets aus vermeintlichen Bottom-Quarks hinsichtlich der Ausbreitungsrichtung mit den generierten Bottom-Quarks verglichen. Bei einem Wert der Winkelöffnung ΔR nach Gleichung (3.4) von kleiner als 0,4 wird davon ausgegangen, dass der Teilchenjet aus der Hadronisierung des jeweiligen generierten Quarks hervorgeht. Die Rekonstruktion wird als erfolgreich gewertet, sofern beide Teilchenjets der Zuweisungshypothese dieses Kriterium erfüllen.

Für die Evaluierung der χ^2 -Rekonstruktion des gesamten Prozesses $t\bar{t}+Z$, $Z \rightarrow b\bar{b}$ wird auf die bestehende Analyse [6] verwiesen. Die Ergebnisse der Evaluierung der einfachen Rekonstruktion sind in Tabelle 6.1 aufgeführt. Die Erfolgsrate der Rekonstruktion des Higgs-Bosons ist in allen Regionen der Rekonstruktion des Z-Bosons überlegen. In der Region mit mindestens vier Teilchenjets bei drei b-Tags ist die Effizienz der letzteren auffällig gering. Die kombinierte Erfolgsrate ist in absoluten Zahlen nur mäßig hoch, bietet jedoch ausreichend Anlass für die Berücksichtigung in der Analyse.

Tabelle 6.1.: Erfolgsrate der einfachen χ^2 -Rekonstruktion. Die Jets aus der Zuweisungshypothese werden mit der Generatorinformation der Bottom-Quarks verglichen. Aus der Anzahl der Ereignisse mit zwei korrekt zugewiesenen Teilchenjets wird die Erfolgsrate bestimmt.

	$4j, \geq 3t$	$5j, \geq 3t$	$\geq 6j, \geq 3t$	$\geq 4j, 3t$	$\geq 4j, \geq 4t$	kombiniert
Z Rekonstr.	26,1 %	27,2 %	25,4 %	19,8 %	31,0 %	26,1 %
H Rekonstr.	29,7 %	31,7 %	32,8 %	28,8 %	33,6 %	31,8 %

6.2. Optimierung der Leistung des Klassifizierers

Die Trennkraft des Klassifizierers wird optimiert durch Variation der Eingangsvariablen und Kombinationen unterschiedlicher Regionen, definiert durch die Anzahl der Jets und Anzahl der b-Tags. Eine Reduktion der Anzahl der Eingangsvariablen kann zudem wegen Eliminierung möglicher Störquellen zu einem besseren Ergebnis bei besser nachvollziehbarem Verhalten des ANNs führen. Die letztendlich für eine Analyse verwendeten Variablen und Parameter werden auf Grundlage der präsentierten Untersuchungen bezüglich Verteilung und Verhalten des ANNs sowie theoretischer Überlegungen ausgewählt.

Zur Optimierung der Leistung des Klassifizierers wird das Training des ANN mit den in Tabelle 4.1 aufgeführten Prozessen durchgeführt. Die Untersuchungen erfolgen zu Beginn für die binäre Trennung von $t\bar{t}+H$ - und $t\bar{t}+Z$ -Ereignissen, um etwaige Erkenntnisse im weiteren Verlauf zu berücksichtigen. Der Fokus liegt auf der Optimierung der simultanen Trennung der fünf wesentlichen Prozesse. Für die Diskussion der Ergebnisse werden die Fläche unter der Receiver Operating Characteristic (ROC) Kurve sowie die Klassifizierungseffizienz als Maß zur Beurteilung der Leistung eines ANN herangezogen.

Für die technische Umsetzung werden Skripte des Wrap-Ups DRACO-MLfoy [34] verwendet, welches auf Funktionalitäten von TensorFlow [35] basiert.

6.2.1. Netzarchitektur

Zur besseren Vergleichbarkeit der Ergebnisse werden die Studien mit ANNs derselben Architektur durchgeführt. Die vorliegenden Variablen stehen nur in indirekter Abhängigkeit, sodass sich die Einspeisung in das Netz über eine einzelne Eingangsschicht ohne Verbindung zwischen den Neuronen zweckmäßig erweist. Nach dem Prinzip eines Feed-Forward-Netzes [36] folgen verdeckte Schichten mit vollständiger Verknüpfung aller Neuronen zwischen benachbarten Schichten. Die eingesetzte Aktivierungsfunktion **LeakyReLU** der jeweiligen Neuronen ist definiert als

$$f(z) = \begin{cases} \alpha z & \text{für } z < 0, \\ z & \text{für } z \geq 0 \end{cases} \quad (6.1)$$

wobei der positive Parameter α auf $\alpha = 0,1$ gesetzt wird.

Für die binäre Trennung zwischen ausschließlich Signal- und Untergrundprozessen besteht die Ausgangsschicht aus einem einzelnen Neuron mit der Aktivierungsfunktion **Sigmoid**. Die Ausgabewerte bewegen sich zwischen Null und Eins, sodass ein Ereignis mit einem Ausgangswert von mindestens 0,5 als der zuvor festgelegte Signalprozess klassifiziert wird. Wegen der sigmoiden Transformation des Rückgabewertes auf den Wertebereich zwischen Null und Eins, kann der Rückgabewert als Wahrscheinlichkeit für die Zuordnung interpretiert werden.

Tabelle 6.2.: Konfiguration der ANNs für das Training. Enthalten sind die Konfigurationen für ANNs ausgelegt für binäre Klassifikation und Multiklassifikation. Abgesehen von der Anzahl der Trainingsepochen ist die Konfiguration für ANNs zur Multiklassifizierung gleich der Konfiguration der in der bestehenden Analyse [6] verwendeten ANNs.

Parameter	binäres Training	Multiklassifikation
Anzahl der verdeckten Schichten	2	
Anzahl Neuronen pro Schicht	50	
Aktivierungsfunktion	LeakyReLU, $\alpha = 0,1$	
Aktivierungsfunktion (Ausgang)	Sigmoid	Softmax
Dropout [37]	25 %	30 %
Zielfunktion	Binary Crossentropy	Crossentropy
Anzahl der Epochen	300	
L_2 -Norm [38]	10^{-5}	
Batch-Größe	200 Ereignisse pro Anpassung	
Optimierer	Adagrad [39] (Standardeinstellungen)	

Für die simultane Klassifizierung von Ereignissen zu K unterschiedlichen Prozessen ist die Anzahl der Neuronen in der Ausgangsschicht gleich der Anzahl aufzulösender Prozesse. Sodass auch in diesem Fall eine simple Interpretation der Ausgangswerte möglich ist, wird die Summe der Ausgangswerte für jedes Ereignis auf Eins normiert. Zu diesem Zweck ist die Aktivierungsfunktion **Softmax** konstruiert zu

$$\sigma(z)_j = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}} \text{ für } j = 1, \dots, K. \quad (6.2)$$

Das zu klassifizierende Ereignis wird dem jeweiligen Prozess zugeordnet, dessen Ausgangsneuron den höchsten Wert als Rückgabewert liefert. Zur Vorbeugung von Übertraining wird während des Trainings der ANNs der **Dropout**-Mechanismus [37] an jeder der verdeckten Schichten eingesetzt. Zudem erhält die Zielfunktion nach dem Prinzip der L_2 -Regularisierung [38] einen zusätzlichen Term, welcher einer zu hohen Gewichtung einzelner Variablen vorbeugt. Eine Übersicht der verwendeten Netzarchitektur ist in Tabelle 6.2 aufgeführt.

6.2.2. Studie der Regionen

Für den zu untersuchenden Prozess $t\bar{t}+Z$, $Z \rightarrow b\bar{b}$ werden im Idealfall sechs rekonstruiert Teilchenjets pro Ereignis erwartet. Ereignisse zu höheren Ordnungen des Prozesses enthalten abhängig von dem zugrundeliegenden Effekt oftmals mehr als sechs Teilchenjets. Den Simulationen zufolge enthält der größere Teil der gemessenen Ereignisse fünf oder weniger Jets. Mögliche Ursachen hierfür sind ein zu geringer Transversalimpuls p_T einzelner Teilchenjets oder die Interpretation zweier im hadronischen Kalorimeter nah beieinander aufgefangener als ein höherenergetischer Teilchenjet. Ebenso werden Teilchenjets bei Koordinaten in der Nähe des Strahlrohres mit hoher Pseudorapidität η wegen geringer Rekonstruktionseffizienz in der Analyse nicht berücksichtigt. Unabhängig von der Ursache dieses Verhaltens steigt für derartige Ereignisse wegen potentiell fehlender Information der Schwierigkeitsgrad für die korrekte Zuordnung. Unter diesen Umständen liefert eine andere Auswahl von Variablen für den Klassifizierer möglicherweise besser verwertbare Information.

Aufgrund der Variation der Signatur eines Ereignisses in Abhängigkeit der rekonstruierten Teilchenjets werden Ereignisse entsprechend der Anzahl der Teilchenjets und Anzahl der b-Tags unterschiedlich behandelt. Die Separierung resultiert in der Erstellung jeweils getrennter Verteilungen für die numerische Anpassung. Da die Anteile der Prozessklassen an den Raten in den unterschiedlichen Regionen variieren, wird durch Konstruktion der jeweiligen Verteilungen die Festlegung bestimmter Störparameter, wie beispielsweise der Untergrundnormierung begünstigt.

Folglich werden für jede Kategorie von ausgewählten Kombinationen von Anzahl der Teilchenjets und Anzahl der b-Tags dedizierte ANNs trainiert und evaluiert. Die Netze besitzen eine identische Architektur und unterscheiden sich ausschließlich in den übergebenen Eingangsvariablen.

Die isolierte Behandlung der Region mit mindestens sechs Teilchenjets erweist sich wegen der verfügbaren Rekonstruktionsvariablen aus der erweiterten χ^2 -Rekonstruktion als sinnvoll. In demselben Zug wird eine Trennung der Regionen mit fünf beziehungsweise vier Jets eingeführt. Dadurch ergibt sich für die Trennung der Ereignisse nach der Anzahl der Teilchenjets ein ANN für jede der drei Regionen. Neben der Trennung nach der Anzahl der Teilchenjets eignet sich ebenfalls eine Trennung nach der Anzahl der Teilchenjets mit einem b-Tag-Wert über 0,303, sodass Ereignisse mit drei b-Tags und mit mindestens vier b-Tags jeweils isoliert behandelt werden.

Ausgehend von der Aufteilung in zwei Gruppen wird für jedes der insgesamt fünf ANNs nach den Eingangsvariablen gesucht, welche die bestmöglichen Ergebnisse für die Auflösung der Prozesse erzielen. Innerhalb der beiden Gruppen wird jedes Ereignis nur einfach durch eines der ANNs propagiert. Der Vergleich der Leistungen soll eine Entscheidung ermöglichen, welche der beiden Gruppen für die Analyse die besseren Ergebnisse versprechen.

6.2.3. Auswahl der Variablen

Die Verteilungen der Variablen in Abbildung 3.2 zeigen deutliche Unterschiede zwischen den Prozessen, an welchen weder ein Higgs-Boson, noch ein Z-Boson beteiligt ist, gegenüber denjenigen Prozessen mit der entsprechenden Resonanz. Die erste Untersuchung für jede der im vorherigen Abschnitt definierten Regionen widmet sich deshalb der Identifizierung der zur Trennung der Prozesse $t\bar{t}+Z(b\bar{b})$ und $t\bar{t}+H(b\bar{b})$ effektivsten Variablen. Im Anschluss daran wird die Untersuchung auf eine simultane Klassifizierung zu weiteren Prozessklassen ausgeweitet.

Einen wesentlichen Anhaltspunkt für die Auswahl von Variablen bildet die Summe der Gewichte eines jeden Neurons der Eingangsschicht nach dem Training des ANNs. Neuronen zu Parametern mit betragsmäßig hohen Gewichten liefern im entsprechenden ANN einen größeren Beitrag als Neuronen mit betragsmäßig geringeren Gewichten. Bei Training mit vielen Variablen kann davon ausgegangen werden, dass in dem gefundenen Minimum der Zielfunktion der Wert der Zielfunktion entlang der Dimension von Parametern mit geringen Gewichten kaum variiert. Den jeweiligen Parametern werden deshalb keine für die Klassifizierung relevante Information zugeschrieben und sie werden für das weitere Vorgehen aus der Auswahl entfernt. Zufällig zugewiesene hohe Gewichte werden durch Training mehrerer Netze mit identischem Aufbau und Vergleich der Gewichte als solche erkannt und statistische Fluktuation ausgeschlossen. Dieser Fall tritt beispielsweise auf, wenn gewählte Anfangsparameter zufällig nahe einem lokalen Minimum der Zielfunktion liegen.

Die untersuchten Sätze von Variablen werden den Gruppen G (Gesamtheit), S (Selektion) und V (Vereinigung) zugeordnet. Sätze der Gruppe G enthalten alle jeweils verfügbaren

Variablen. Die Inhalte der Gruppe S werden abgeleitet aus den Sätzen der Gruppe G nach zehnfachem Training und Auswahl der Variablen mit den durchschnittlich höchsten Gewichten. Das Training beinhaltet die Untersuchung für binäre Trennung von $t\bar{t}+Z$ - und $t\bar{t}+H$ -Ereignissen sowie die Multiklassifizierung. Gestrichene Sätze sind Sätze von reduziertem Umfang, während Sätze mit dem Zusatz (B) Informationen enthalten, die aus der binären Trennung gewonnen wurden. Die gestrichenen Sätze der Gruppe S werden untereinander vereinigt zu den Sätzen der Gruppe V .

Innerhalb der jeweiligen Gruppen sind die Variablensätze nach Inhalt gegliedert. Das Kürzel A steht repräsentativ für Sätze mit sowohl gewöhnlichen Ereignisvariablen und Rekonstruktionsvariablen, E für Ereignisvariablen ohne jegliche Rekonstruktionsvariablen nach der χ^2 -Methode und das Kürzel R bezeichnet Sätze mit ausschließlich Rekonstruktionsvariablen. Die Vereinigung der Sätze der Gruppe S zu Sätzen der Gruppe V erfolgt jeweils nach dem Schema A aus allen Variablensätzen oder nach dem Schema M aus einer Mischung der Sätze mit Inhalt R und E . Der Zusatz B bezeichnet die Zugabe von Variablen, welche einen hohen Beitrag zur binären Trennung liefern. Die Bezeichnung der Sätze setzt sich zusammen aus dem Gruppenkürzel und dem Kürzel für den Inhalt und eventuellen Zusätzen. Der in der bestehenden Analyse [6] verwendete Variablensatz wird mit L bezeichnet. Wegen der geringeren Leistung der Rekonstruktionsvariablen bei Aufteilung der Ereignisse nach Anzahl der Teilchenjets, wird die Strategie zur Bildung des Satzes K nur in Regionen mit drei und mehr als vier b-Tags verfolgt.

6.3. Evaluierung der Klassifizierung

Für den objektiven Vergleich der Leistung eines Klassifizierers werden unterschiedliche Maße herangezogen, welche unterschiedliche Information bezüglich der Leistung in numerischer Form wiedergeben. Zu dieser Information gehört beispielsweise die Effizienz der Klassifizierung einzelner Prozessklassen und ob bei Fehlklassifikationen etwaige Tendenzen zu bestimmten anderen Klassen bestehen. Die Vereinigung in einem einzelnen Wert erweist sich in Anbetracht der zahlreichen unterschiedlichen Informationsquellen als unzureichend.

Aus der Auftragung der Effizienz ε_{TP} für die korrekte Erkennung von Signalereignissen über der Effizienz ε_{FP} für fälschliche Zuordnung von Untergrund- zu Signalereignissen ergibt sich die Receiver Operating Characteristic (ROC) Kurve. Für die Trennung der beschriebenen Prozesse ist die ROC-Kurve exemplarisch in Abbildung 6.2 aufgetragen. Der ROC-Integralwert bezeichnet die Fläche unterhalb der ROC-Kurve und ist ein viel referenziertes Maß für die Evaluierung der Leistung eines Klassifizierers. Der maximal mögliche Wert beläuft sich zu Eins und ist gleichbedeutend mit einer perfekten Auflösung. Für die simultane Klassifizierung von mehreren Prozessen wird der ROC-Integralwert nach Hand [40] bestimmt. Im Kontext der experimentellen Teilchenphysik können die jeweiligen Interpretationen [40, 41] für die Trennung von Signal und Untergrund übernommen und als Maß für den Vergleich der Trennkraft unterschiedlicher Klassifizierer herangezogen werden.

Zur Prüfung vermehrter fälschlicher Klassifizierung bestimmter Klassen werden die Ergebnisse der Klassifizierung von Testdaten in einer Verwechslungsmatrix aufgetragen. Entlang der Diagonalen in Abbildung 6.2 ist der Anteil der jeweiligen korrekt zugeordneten Ereignisse einer Prozessklasse aufgetragen. In dieser Darstellung werden Tendenzen bezüglich Fehlklassifizierungen erkennbar. Der Mittelwert entlang der Diagonalen der Verwechslungsmatrix bündelt Information über den Anteil korrekter Zuweisungen entlang aller Prozessklassen unabhängig von Differenzen in den Größen der verwendeten Datensätze. Dieser Wert korreliert demnach direkt mit der Leistung des Klassifizierers und bietet möglichen Anhaltspunkt für dessen Evaluierung. Ferner gibt die Streuung um diesen Mittelwert Aufschluss, ob die

Tabelle 6.3.: Übersicht der studierten Sätze von Variablen. Abhängig von möglichen Variablen mit Standardwerten kann die Anzahl der Variablen zwischen den Kategorien innerhalb eines Satzes variieren. Bei Auswahl von Teilmengen mit fixierter Mächtigkeit werden als Kriterium jeweils die Gewichte nach dem Training herangezogen. Eine graphische Aufbereitung findet sich in Anhang B.

Kürzel	Zusatz	Beschreibung	Mächtigkeit
GA	G	alle verfügbaren Variablen ohne Standardwerte	200 bis 276
GE		$GE \subset GA$, ausschließlich Ereignisvariablen	106 bis 147
GR		$GR \subset GA$, ausschließlich Rekonstruktionsvariablen	109 bis 140
SA SE SR	S	$\{SA \subset GA : SA = 20\}$ $\{SE \subset GE : SE = 20\}$ $\{SR \subset GR : SR = 20\}$	20
SA' SE' SR'	S	$\{SA' \subset GA : SA' = 10\}$ $\{SE' \subset GE : SE' = 10\}$ $\{SR' \subset GR : SR' = 10\}$	10
$SA(B)$ $SE(B)$ $SR(B)$	S	$\{SA(B) \subset GA : SA(B) = 10\}$ $\{SE(B) \subset GE : SE(B) = 10\}$ $\{SR(B) \subset GR : SR(B) = 10\}$	10
$SA'(B)$ $SE'(B)$ $SR'(B)$	S	$\{SA'(B) \subset GA : SA'(B) = 5\}$ $\{SE'(B) \subset GE : SE'(B) = 5\}$ $\{SR'(B) \subset GR : SR'(B) = 5\}$	5
$V(AB)$ $V(A)$ $V(MB)$ $V(M)$	M	$SA' \cup SE' \cup SR' \cup SA'(B) \cup SE'(B) \cup SR'(B)$ $SA' \cup SE' \cup SR'$ $SE' \cup SR' \cup SE'(B) \cup SR'(B)$ $SE' \cup SR'$	24 bis 31 20 bis 22 24 bis 26 20
L K		Variablensatz aus Analyse von van der Linden [6] $SR' \cup SR'(B)$, nur für Regionen nach Anzahl der b-Tags	17 bis 22 25

Effizienz der Klassifizierung zwischen den Klassen variiert und der Klassifizierer folglich voreingenommen operiert.

6.4. Evaluierung in physikalischem Kontext

Für die Untersuchung der Auswirkungen auf das physikalisch interpretierbare Ergebnis werden die Ansätze bis zur numerischen Anpassung propagiert. Die berücksichtigten Unsicherheiten werden gegenüber einer vollständigen Analyse stark reduziert, sodass eine tatsächlich physikalische Interpretation nicht möglich ist. Stattdessen dienen die Ergebnisse ausschließlich zum Vergleich der Ansätze untereinander. Eine Aufführung der berücksichtigten Unsicherheiten im Jargon der Analysegruppe sowie eine knappe Erläuterung finden sich in Anhang C.

6.4.1. Methodik

Für die Prüfung von Hypothesen bezüglich des Signalstärkeparameters bieten sich zwei Methoden an.

Aus der Anpassung der Verteilungen der Simulation an die Detektordaten wird ein Wert für den Signalstärkeparameter μ sowie dessen Unsicherheit gewonnen. Bei Berücksichtigung

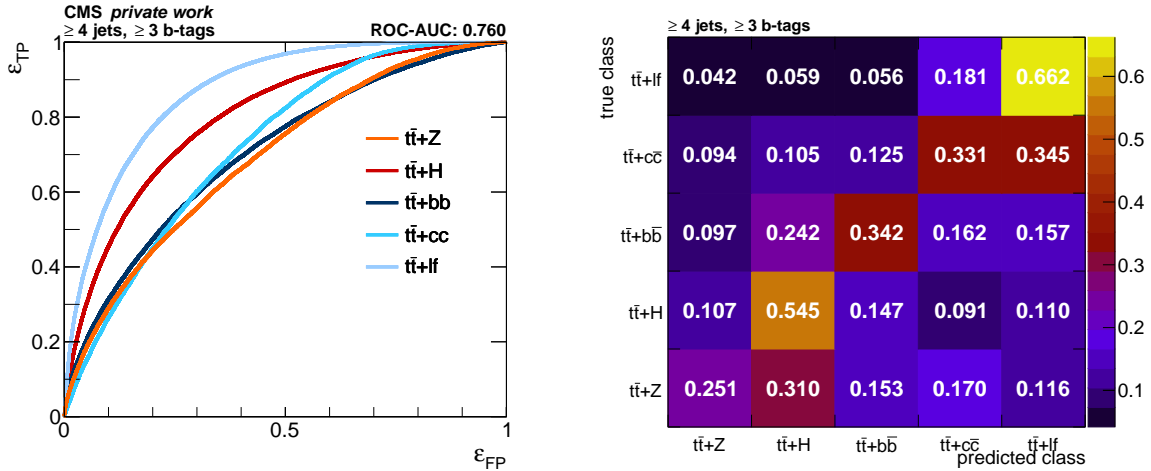


Abbildung 6.2.: Exemplarische Darstellung einer ROC-Kurve und einer Verwechslungsmatrix. Die ROC-Kurve (links) kann jeweils nur bezüglich eines einzelnen Prozesses gegenüber den übrigen Prozessen als Untergrund aufgetragen werden. In der Verwechslungsmatrix (rechts) ist die häufige Klassifikation von $t\bar{t}+Z$ -Ereignissen als $t\bar{t}+H$ -Prozess deutlich erkennbar.

systematischer Unsicherheiten ergibt sich die kombinierte Unsicherheit, von welcher die rein statistische Unsicherheit durch Anpassung ohne Berücksichtigung systematischer Unsicherheiten isoliert werden kann.

In der zweiten Methode wird unter statistischer Variation aus der Anpassung eine obere Ausschlussgrenze für den Signalstärkeparameter entsprechend des 95 % Konfidenzniveaus bestimmt. Die Unsicherheitsintervalle folgen erneut aus der Gestalt des Minimums der Zielfunktion der Anpassung. Der Wert der Ausschlussgrenze legt den maximal beobachteten Beitrag des Signalprozesses an der Gesamtrate fest. Analog zur ersten Methode kann die Anpassung unter Ausschluss der systematischen Unsicherheiten durchgeführt werden und dadurch eine Abschätzung der Ergebnisse unter rein statistischer Unsicherheit gewonnen werden.

Die Anpassung wird üblicherweise durch Variation eines einzelnen Signalstärkeparameters durchgeführt. In begründeten Fällen gibt die simultane Anpassung der Signalstärkeparameter zweier Prozesse ebenfalls interessante Information preis und kann zur Verbesserung der Analyse beitragen. Unter den betrachteten Prozessen ist die simultane Anpassung für $t\bar{t}+Z$ - und $t\bar{t}+H$ -Ereignisse von Interesse, da aufgrund der Ähnlichkeit der beiden Prozesse eine hohe negative Korrelation der Signalstärkeparameter zu erwarten und die Unsicherheit auf die Signalstärkeparameter nach der Anpassung entsprechend groß ist. Bei Erhöhung des Signalstärkeparameters der einen Prozessklasse muss der Beitrag des anderen Prozesses reduziert werden, um ein gutes Verhältnis zu den Messdaten wiederherzustellen. Die Konstruktion von Diskriminatoren, welche eine Messung beider Prozesse bei geringer Korrelation der Signalstärkeparameter erlauben, sind erstrebenswert für weitere Analysen.

6.4.2. Asimov-Tests

Bevor die Anpassung der Verteilung der Diskriminatoren an Detektordaten erfolgt, wird diese bezüglich der physikalischen Konsistenz und der zu erwartenden Unsicherheiten getestet. Gleichmaßen erfolgt in diesem Test die Validierung der Verteilungen der Diskriminatoren. Zu diesem Zweck werden die Verteilungen der Diskriminatoren an sich selbst angepasst und die daraus hervorgehenden Unsicherheiten betrachtet. Wenn bereits vor der Anpassung die Simulation hinreichend mit den Detektordaten übereinstimmt, kann

über diese sogenannten Asimov-Tests eine Abschätzung für die zu erwartenden Ergebnisse bei der tatsächlichen Messung erfolgen.

Mit aus den ANNs generierten Verteilungen der Diskriminatoren werden Asimov-Tests durchgeführt und dadurch ein Vergleich hinsichtlich der Leistung im physikalischen Kontext ermöglicht. Zu diesem Zweck werden die resultierenden Unsicherheiten als Maß herangezogen. Die Begrenzen der Histogramme werden für die Anpassung von der bestehenden Analyse [6] übernommen.

7. Ergebnisse

Die Ergebnisse der Studie werden hinsichtlich der in Abschnitt 6.3 und Abschnitt 6.4 vorgestellten Parameter untersucht und verglichen. Zu Beginn des Kapitels erfolgt in Abschnitt 7.1 der Vergleich der Klassifizierer für unterschiedliche Variablensätze bezüglich der Trennkraft bei binärer Trennung von $t\bar{t}+Z$ - und $t\bar{t}+H$ -Ereignissen. Die Ergebnisse schaffen eine Grundlage für die Validierung des positiven Einflusses der Ereignisrekonstruktion. Im Anschluss daran werden in Abschnitt 7.3 die Ergebnisse der Asimov-Tests vorgestellt und die Zusammensetzung der Variablensätze im Kontext der Multiklassifizierung evaluiert. Das Fazit bezüglich der Wahl der Methode zur Trennung der Regionen wird in Abschnitt 7.4 diskutiert.

7.1. Mehrwert der Rekonstruktionsvariablen

Die Mittelwerte des ROC-Integralwertes aus zehn ANNs für binäre Klassifizierung von $t\bar{t}+Z$ - und $t\bar{t}+H$ -Ereignissen sind in Abbildung 7.1 aufgetragen. In der Region mit mindestens vier b-Tags und in der Region mit mindestens sechs Teilchenjets erreicht der Klassifizierer innerhalb eines Variablensatzes jeweils die höchste Trennkraft. Da die genannten Regionen jeweils den Signalregionen entsprechen, deckt sich dieses Verhalten mit der Erwartung.

Die Variablensätze mit gewöhnlichen Ereignisvariablen und Rekonstruktionsvariablen (A) ermöglichen über alle Regionen hinweg gegenüber den Sätzen mit ausschließlich Ereignisvariablen (E) eine bessere Auflösung bei einer Steigerung des ROC-Integralwertes von bis zu 3,0 %. Die Aufstellung der Gewichte nach dem Training wird angeführt durch den Logarithmus der χ^2 -Werte der einfachen Boson-Rekonstruktion für das Z-Boson. Dieses Ergebnis bestätigt den Mehrwert der Rekonstruktionen, wenn auch der Wert für χ^2 vornehmlich Informationen über die Qualität der Minimierung liefert, statt das tatsächliche Charakteristika des Bosons zu beschreiben. Kinematische Größen oder Massen der rekonstruierten Objekte hingegen haben im Vergleich dazu keinen derart hohen Einfluss auf die binäre Trennung.

Variablensätze mit ausschließlich Rekonstruktionsvariablen (R) schneiden bei der Klassifizierung im Vergleich zu den übrigen Sätzen unterdurchschnittlich ab. Folglich ist der Informationsgehalt trotz wesentlicher Unterschiede in den Verteilungen der Variablen der beiden Prozessklassen dem der gewöhnlichen Ereignisvariablen unterlegen.

Die Reduktion der Mächtigkeit der Variablensätzen GA zu SA und GR zu SR übt einen geringen negativen Einfluss auf die Trennkraft des Klassifizierers bei einer maximalen Reduktion des ROC-Integralwertes von etwa 2 %. Für die Sätze mit gewöhnlichen Ereignisvariablen (E) stellt sich ein gegensätzlicher Effekt ein, durch welchen die zugehörigen Werte um bis zu 2,1 % erhöht werden. Bei deutlich erkennbaren Unterschieden in den Verteilungen der Rekonstruktionsvariablen geht durch Reduktion der Ordnung ein nicht unwesentlicher Anteil der Information verloren. Die Unterschiede in den Verteilungen der

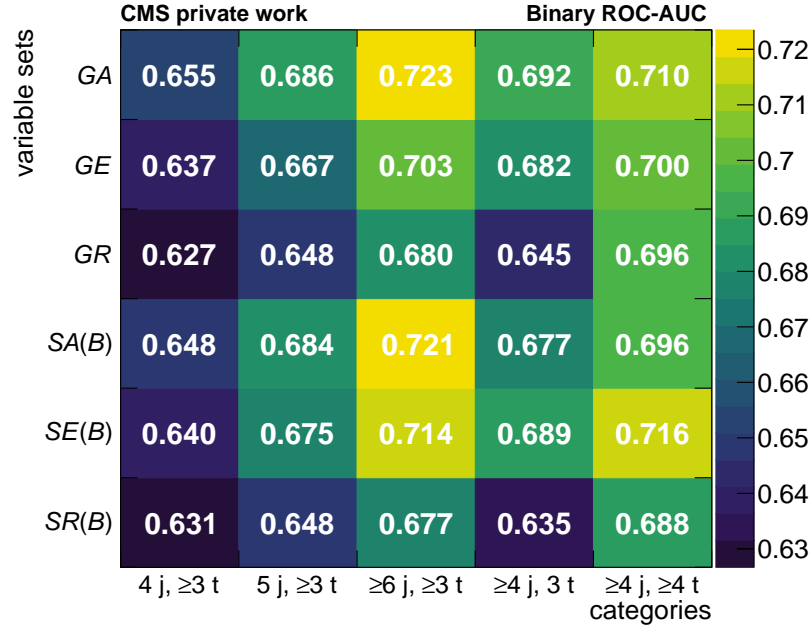


Abbildung 7.1.: Übersicht der ROC-Integralwerte für binäre Trennung. Die Ergebnisse folgen aus mehrfachem Training unter derselben Voreinstellung der Netze zur Trennung von $t\bar{t}+Z$ - und $t\bar{t}+H$ -Ereignissen. Legende: G Gesamtheit jeweils aller Variablen, S Selektion der jeweils am höchsten gewichteten Variablen, E gewöhnliche Ereignisvariablen, R Rekonstruktionsvariablen, A gewöhnliche Ereignis- und Rekonstruktionsvariablen.

gewöhnlichen Ereignisvariablen hingegen sind geringer, sodass sich ein Überfluss an für die Separation unbedeutenden Variablen negativ auswirkt. Dieser kann durch Aussortieren entsprechender Ereignisvariablen verringert werden.

7.2. Evaluierung der Variablensätze

Die Leistung des Klassifizierers für Multiklassifizierung unter Verwendung der in Tabelle 6.3 vorgestellten Variablensätze werden untereinander verglichen. Die Mittelwerte des ROC-Integralwerts aus der Evaluation mehrerer gleicher Netze sind in Abbildung 7.2 aufgeführt. Anschließend werden diese Ergebnisse mit den in Abbildung 7.3 aufgetragenen Mittelwerten der Diagonalen der Verwechslungsmatrix verglichen. Für beide Maße ist jeweils auf der rechten Seite der Abbildungen die isolierte Evaluierung des Neurons zum $t\bar{t}+Z$ -Prozess aufgeführt.

Analog zu den Ergebnissen für die binäre Trennung zeigt sich der negative Ausfall der Variablensätze mit Rekonstruktionsvariablen in den ROC-Integralwerten in Abbildung 7.2 (links) sehr deutlich. Die Variablensätze mit ausschließlich gewöhnlichen Ereignisvariablen bewegen sich in Bezug auf die Trennkraft zwischen den Sätzen mit ausschließlich Rekonstruktionsvariablen und den gemischten Sätzen.

Bezüglich des ROC-Integralwerts liefert der Variablensatz mit allen Variablen die besten Ergebnisse, gefolgt von dem Variablensatz $V(AB)$. Unter aus Vereinigung gewonnenen Variablensätzen bilden sich abgesehen von in der Region mit vier Teilchenjets nur geringe Unterschiede aus. Die Zugabe der jeweils am höchsten gewichteten Variablen in der binären Klassifizierung erhöht die Leistung des ANN leicht. Die Strategie zur Rekombination der am höchsten gewichteten Variablen unterschiedlichen Ursprungs zu einem neuen Variablensatz erweist sich als effektiv. Dieser Ansatz erzielt bessere Ergebnisse als der in bestehender

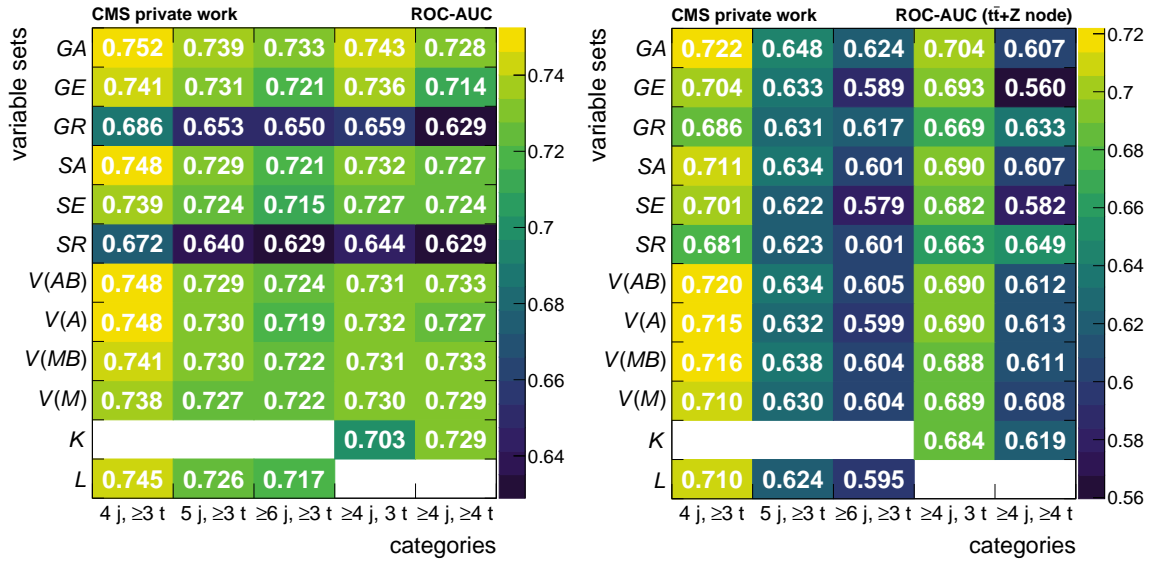


Abbildung 7.2.: Übersicht der ROC-Integralwerte für Multiklassifizierung. Die Rückgabewerte des ANN wurden sowohl in ihrer Gesamtheit (links) als auch ausschließlich bezüglich des Neurons für $t\bar{t}+Z$ -Ereignisse (rechts) untersucht. Legende: G Gesamtheit jeweils aller Variablen, S Selektion der jeweils am höchsten gewichteten Variablen, V Vereinigung der am höchsten gewichteten Variablen unterschiedlichen Ursprungs, E gewöhnliche Ereignisvariablen, R Rekonstruktionsvariablen, A gewöhnliche Ereignis- und Rekonstruktionsvariablen.

Analyse verfolgte Ansatz der Auswahl der am höchsten gewichteten Variablen nach dem Training mit allen Variablen.

Gegenüber dem mit demselben Ansatz gewonnene Variablensatz SA weist der Satz aus der bestehenden Analyse (L) leichte Defizite auf. Unter Annahme einer sehr ähnlichen Verteilung der Variablen in den jeweils verwendeten simulierten Daten, ist davon auszugehen, dass der Unterschied durch die ergänzte χ^2 -Rekonstruktion für das Boson alleine zustande kommt.

Bei isolierter Betrachtung des ROC-Integralwerts des Neurons für $t\bar{t}+Z$ -Prozesse in Abbildung 7.2 (rechts) fallen die Variablensätze aus ausschließlich Rekonstruktionsvariablen nur leicht negativ auf. Gegenüber des kombinierten ROC-Integralwertes zeigen sich die Unterschiede zwischen den unterschiedlichen Regionen noch deutlicher. Die ROC-Integralwerte in der Region mit genau vier Teilchenjets und in der Region mit genau drei b-Tags sind höher als in den übrigen Regionen. Dieses Verhalten rührt möglicherweise von den verhältnismäßig großen Mengen von Untergrundereignissen, welche in den jeweiligen Regionen mit hoher Effizienz richtig identifiziert werden.

Die relativen Unterschiede der in Abbildung 7.3 (links) aufgeführten Mittelwerte der Diagonalen der Verwechslungsmatrix sind insgesamt geringer als zwischen den ROC-Integralwerten. Auch hier fallen die Variablensätze mit ausschließlich Rekonstruktionsvariablen negativ auf. Die Tendenzen bei aus Vereinigung gewonnenen Variablensätzen sowie die Dominanz des Variablensatzes mit allen Variablen GA spiegeln sich in diesem Maß ebenfalls wider. Wesentliche Differenz in der Auftragung besteht in der unterschiedlichen Bewertung der Regionen innerhalb eines Variablensatzes. Während die Aufstellung der ROC-Integralwerte durch die Region mit genau vier Teilchenjets dominiert wird, wird die Aufstellung der Klassifizierungseffizienzen angeführt durch die Region mit mehr als vier b-Tags. Der höchste Wert diesbezüglich wird von dem Variablensatz $V(AB)$ generiert. Bezüglich der isolierten Klassifizierungseffizienz in Abbildung 7.3 (rechts) werden mit

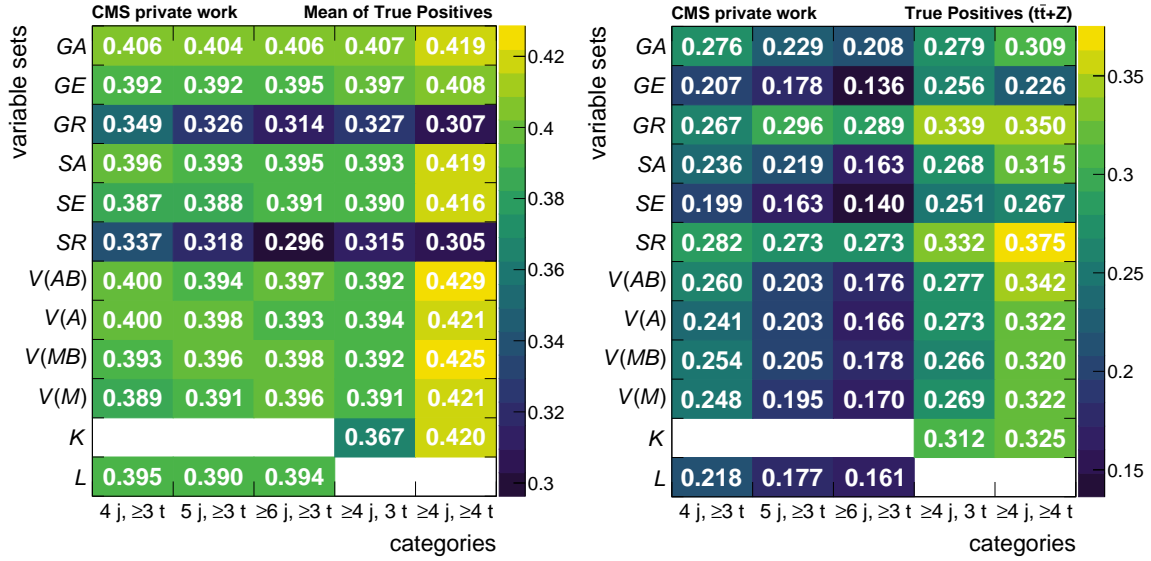


Abbildung 7.3.: Übersicht der Klassifizierungseffizienzen bei Multiklassifizierung. Die Effizienzen sind über alle Klassen gemittelt (links) und für den $t\bar{t}+Z$ -Prozess isoliert (rechts) aufgeführt. Legende: G Gesamtheit jeweils aller Variablen, S Selektion der jeweils am höchsten gewichteten Variablen, V Vereinigung der am höchsten gewichteten Variablen unterschiedlichen Ursprungs, E gewöhnliche Ereignisvariablen, R Rekonstruktionsvariablen, A gewöhnliche Ereignis- und Rekonstruktionsvariablen.

Variablensätzen aus ausschließlich Rekonstruktionsvariablen eindeutig die besten Ergebnisse erreicht. Dies bestätigt den positiven Einfluss der Rekonstruktionsvariablen auch in der Multiklassifizierung. Während in allen übrigen Fällen die isolierte Effizienz für $t\bar{t}+Z$ -Ereignisse unter dem Durchschnitt aller Prozessklassen liegen, wird unter Einsatz von ausschließlich Rekonstruktionsvariablen in der Region mit mindestens vier b-Tags das Verhältnis umgekehrt.

Insbesondere bezüglich der isolierten Betrachtung des Neurons für $t\bar{t}+Z$ -Ereignisse bestehen wesentlich Unterschiede zwischen den Aussagen der Klassifizierungseffizienz und des ROC-Integralwertes hinsichtlich unterschiedlicher verwendeter Variablensätze. Welches Maß die Leistung in physikalischem Kontext besser wiedergibt, entscheiden die Ergebnisse der Asimov-Tests.

7.3. Evaluierung in physikalischem Kontext

Die Ergebnisse der in Abschnitt 6.4 vorgestellten Methoden werden nacheinander vorgestellt. Eine tabellarische Auftragung ist in Anhang D aufgeführt. Bereits in der Evaluierung der Trennkraft erkannte Tendenzen setzen sich in nachfolgenden Erläuterungen fort.

7.3.1. Unsicherheiten der Anpassung

Der Signalstärkeparameter für die $t\bar{t}+Z$ -Produktion beläuft sich bei einfacher Anpassung ausnahmslos zu Eins, sodass die numerische Anpassung als stabil angenommen werden kann. Abhängig vom verwendeten Variablensatz und den für die Anpassung gewählten Regionen bewegen sich die näherungsweise symmetrischen Unsicherheiten zwischen 0,74 und 1,18. Eine grafische Aufbereitung ist in Abbildung 7.4 aufgeführt.

Der Einfluss der Reduktion der Mächtigkeit fällt im physikalischen Kontext wesentlich größer aus, sodass die Unsicherheiten der Sätze mit jeweils allen Variablen nach Auswahl der jeweils 20 am höchsten gewichteten Variablen erheblich zunehmen. Die mit dem

Variablensatz aus der bestehenden Analyse resultierende Unsicherheit ist marginal größer als die Unsicherheit des Variablensatzes SA , der nach demselben Schema gewonnen wird.

Aus Vereinigung gewonnene Variablensätze beweisen sich im Vergleich in der Anpassung der Diskriminatoren unter Aufteilung nach Anzahl der Teilchenjets durch geringere Unsicherheiten gegenüber den übrigen Variablensätzen. Der Satz $V(AB)$ hebt sich dabei nochmals deutlich von den übrigen ab und erreicht nahezu dieselbe Unsicherheit wie der Variablensatz GA . Somit setzt sich das Ergebnis der Evaluierung der Leistung des Klassifizierers auch in physikalischem Kontext fort. Angesichts des wesentlich reduzierten Umfangs ist der Variablensatz $V(AB)$ vorzuziehen.

Bei Anpassung der Diskriminatoren in den Regionen aufgeteilt nach Anzahl der b-Tags stellen sich Variablensätze mit ausschließlich Rekonstruktionsvariablen als effektiv heraus. Im Gegensatz dazu liefern Variablensätze aus ausschließlich gewöhnlichen Ereignisvariablen größere Unsicherheiten. Der Variablensatz GA erzielt mit etwa 0,724 die niedrigste Unsicherheit, gefolgt von Variablensätzen aus Vereinigung mit etwa 0,80.

Die relativen Unterschiede der isolierten statistischen Unsicherheit sind wesentlich geringer gegenüber den Unterschieden der Unsicherheiten unter Berücksichtigung systematischer Unsicherheiten. Unter ihnen sind Muster schwierig zu identifizieren, jedoch zeigen die Werte ähnliche Tendenzen wie die Unsicherheiten mit Systematiken.

Es ist zu erwarten, dass die Unsicherheiten bei Berücksichtigung aller zur Verfügung stehender Störparameter weiter steigen und das Verhältnis zum Signalstärkeparameter dadurch möglicherweise größer als Eins wird. Unter diesen Umständen können aus der Bestimmung der oberen Grenze, korrespondierend zum 95 % Signifikanzniveau, physikalisch signifikantere Aussagen getroffen werden.

7.3.2. Ausschlussgrenze

Die Ergebnisse aus der Bestimmung der erwarteten Ausschlussgrenze für den Signalstärkeparameter für die $t\bar{t}+Z$ -Produktion sind in Abbildung 7.5 aufgetragen und bewegen sich im Wertebereich zwischen 1,4 und 2,3. Sowohl bei Anpassung der Regionen aufgeteilt nach Anzahl der Teilchenjets als auch bei Aufteilung nach Anzahl der b-Tags zeigen sich sehr ähnliche Tendenzen wie bei den Unsicherheiten auf den Signalstärkeparameter.

Die Anpassung der Diskriminatoren aufgeteilt nach Anzahl der Teilchenjets resultiert in starken Schwankungen der Ausschlussgrenzen für unterschiedliche Variablensätze. Dennoch wird mit dem Satz $V(AB)$ die niedrigste Schranke von 1,45 erreicht. Variablensätze aus ausschließlich Rekonstruktionsvariablen liefern eine hohe Ausschlussgrenze bei hoher Streuung und erzielen damit erneut keine sinnvoll verwertbaren Ergebnisse.

Bei Anpassung der Diskriminatoren aufgeteilt nach Anzahl der b-Tags erzielen die Variablensätze aus ausschließlich Rekonstruktionsvariablen wesentlich niedrigere Werte für die Ausschlussgrenze. Aus Vereinigung gewonnene Sätze liefern durchgehend gute Ergebnisse und konkurrieren mit dem Variablensatz mit ausnahmslos allen Variablen GA , aus welchem in diesen Regionen die niedrigste obere Schranke von 1,39 hervorgeht.

7.3.3. Korrelation bei zweidimensionaler Anpassung

Nach simultaner numerischer Anpassung betragen die Signalstärkeparameter der $t\bar{t}+Z$ - und $t\bar{t}+H$ -Produktion gleichermaßen ausnahmslos Eins. Jegliche Unsicherheiten der Signalstärkeparameter und die Korrelationen, die aus der Anpassung hervorgehen, sind in Anhang D aufgeführt. Die Korrelationen allein sind in Abbildung 7.6 aufgetragen.

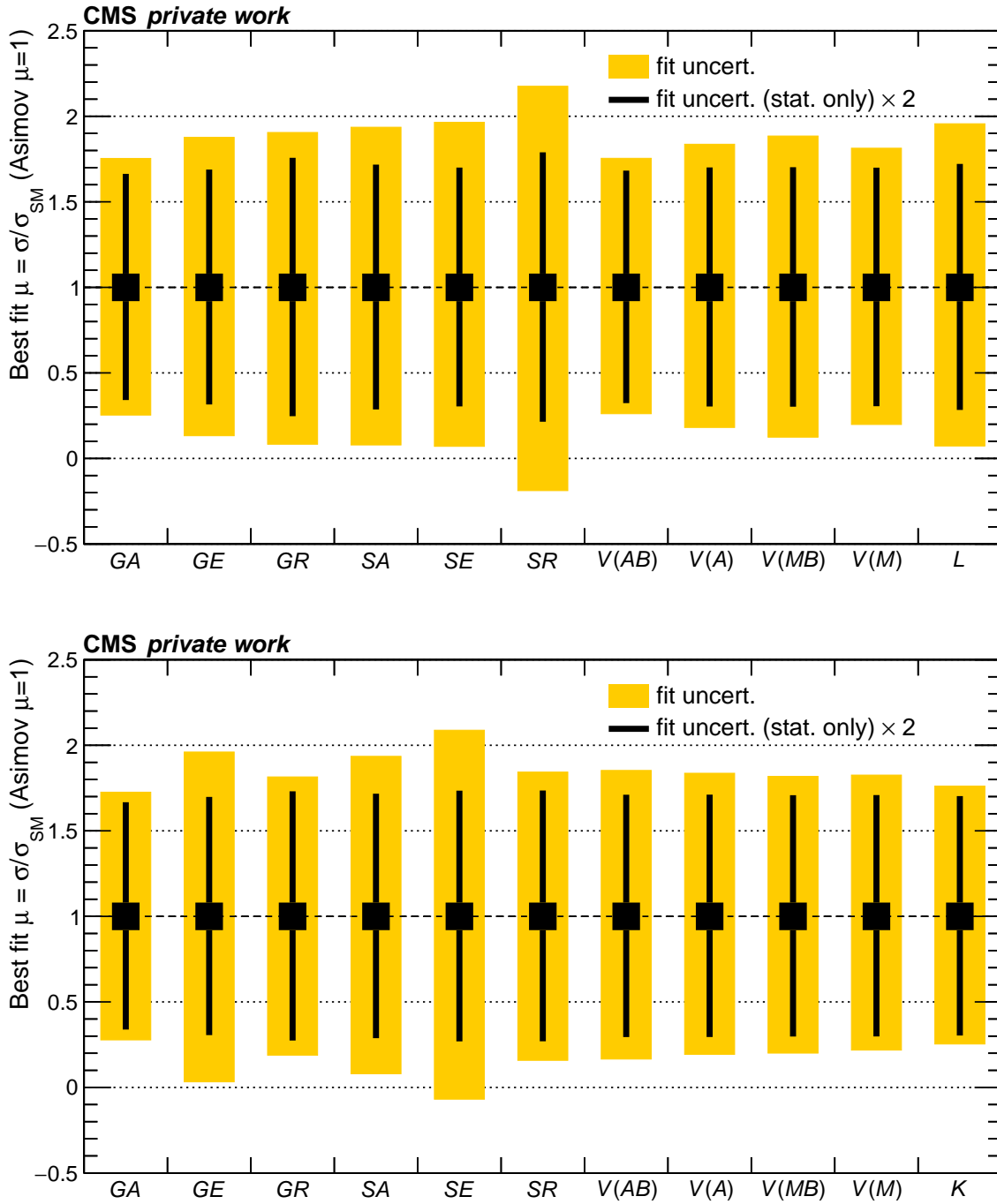


Abbildung 7.4.: Unsicherheiten des Signalstärkeparameters. Die Ergebnisse sind getrennt für Aufteilung der Diskriminatoren nach Anzahl der Teilchenjets (oben) und Anzahl der b-Tags (unten). Farbige Fehlerbänder markieren das 1σ -Intervall der Unsicherheit auf den Signalstärkeparameter unter Berücksichtigung systematischer Unsicherheiten. Die schwarzen Fehlerbalken entsprechen der isolierten Unsicherheit statistischer Unsicherheit. Legende: *G* Gesamtheit jeweils aller Variablen, *S* Selektion der jeweils am höchsten gewichteten Variablen, *V* Vereinigung der am höchsten gewichteten Variablen unterschiedlichen Ursprungs, *E* gewöhnliche Ereignisvariablen, *R* Rekonstruktionsvariablen, *A* gewöhnliche Ereignis- und Rekonstruktionsvariablen.

Die Werte der Korrelationen sind ausnahmslos negativ und decken sich in diesem Verhalten mit der Erwartung. Bei Erhöhung des Signalstärkeparameters eines Prozesses wird die Änderung der gesamten Rate durch Verringerung des Signalstärkeparameters des jeweils

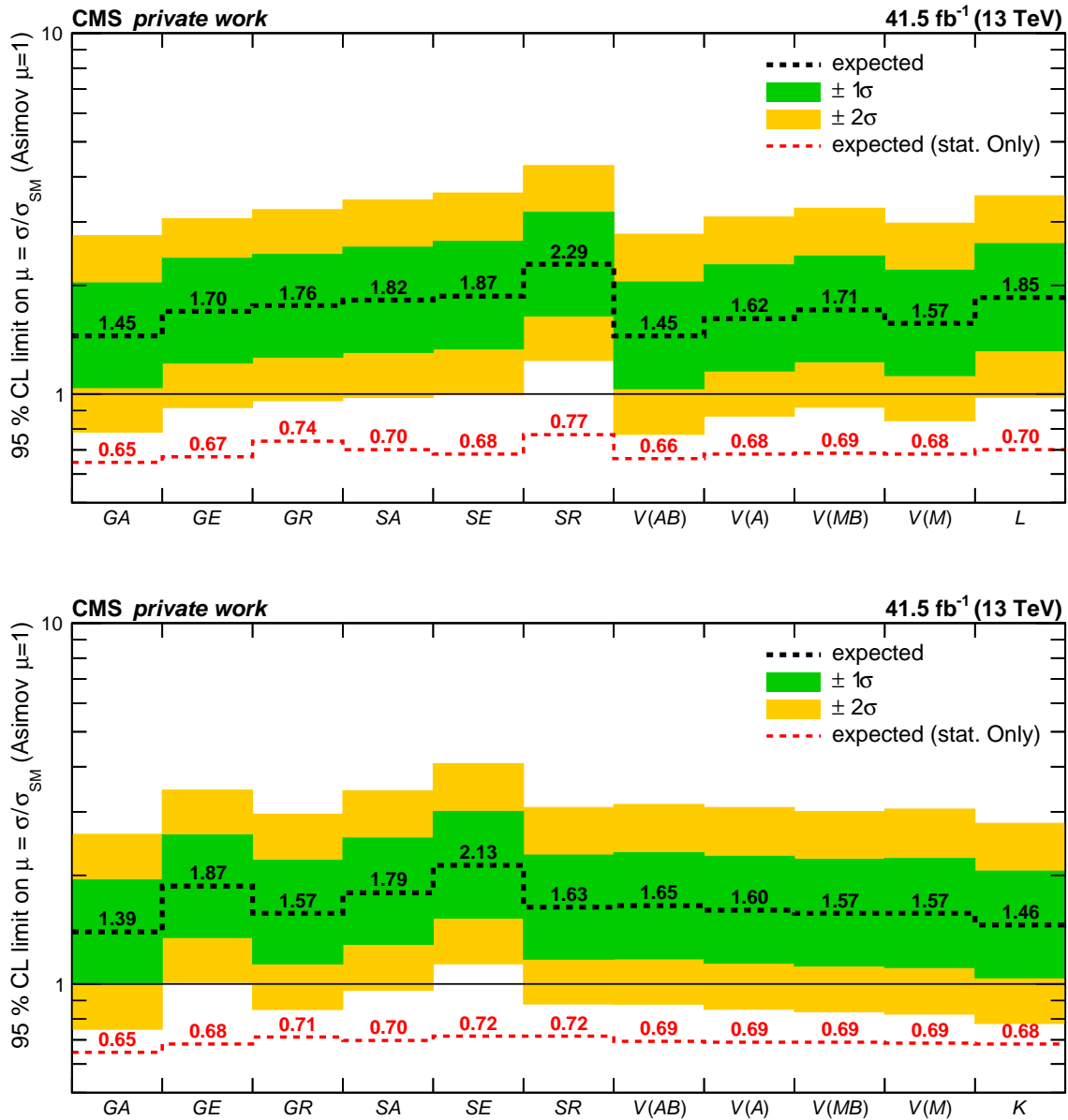


Abbildung 7.5.: Erwartete Ausschlussgrenzen für den Signalstärkeparameter.

Die Ergebnisse sind getrennt für Aufteilung der Diskriminatoren nach Anzahl der Teilchenjets (oben) und Anzahl der b-Tags (unten). Farbige Fehlerbänder markieren das 1σ und das 2σ -Standardabweichung der Streuung um die Ausschlussgrenze bei Anpassung unter Berücksichtigung der Störparameter. Ebenfalls aufgetragen ist die erwartete Ausschlussgrenze bei Anpassung ohne Berücksichtigung systematischer Unsicherheiten. Legende: *G* Gesamtheit jeweils aller Variablen, *S* Selektion der jeweils am höchsten gewichteten Variablen, *V* Vereinigung der am höchsten gewichteten Variablen unterschiedlichen Ursprungs, *E* gewöhnliche Ereignisvariablen, *R* Rekonstruktionsvariablen, *A* gewöhnliche Ereignis- und Rekonstruktionsvariablen.

anderen Prozesses ausgeglichen. Die Unsicherheiten auf den Parameter für $t\bar{t}+Z$ -Produktion bewegen sich in ähnlichen Dimensionen wie bei der alleinigen Anpassung.

Bei Aufteilung der Diskriminatoren nach Anzahl der Teilchenjets erzielen aus Vereinigung gewonnene Variablensätze gegenüber den übrigen betragsmäßig niedrigere Werte für die Korrelation. Der Variablensatz aus der bestehenden Analyse [6] erweist sich in diesem Kontext als mäßig effektiv. Der Satz $V(AB)$ fällt dabei durch besonders geringe Korrelation

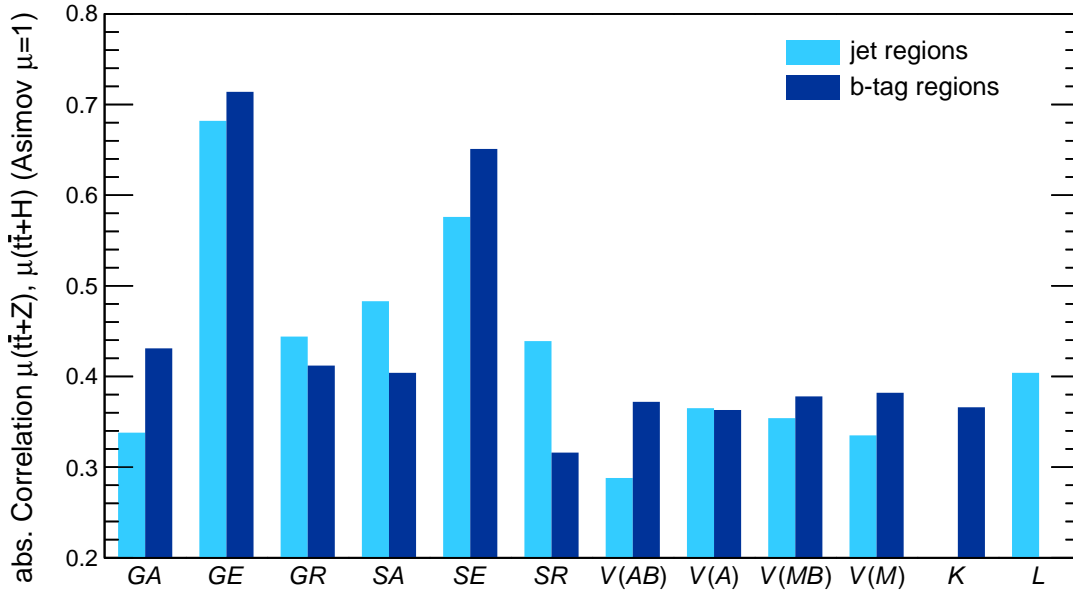


Abbildung 7.6.: Übersicht der Korrelationen bei multidimensionaler Anpassung. Aufgetragen ist der Betrag der Korrelation zwischen den Signalstärkeparametern des $t\bar{t}+Z$ - und des $t\bar{t}+H$ -Prozesses für Anpassung der Diskriminatoren bei Aufteilung nach Anzahl der Jets und Anzahl der b-Tags. Legende: *G* Gesamtheit jeweils aller Variablen, *S* Selektion der jeweils am höchsten gewichteten Variablen, *V* Vereinigung der am höchsten gewichteten Variablen unterschiedlichen Ursprungs, *E* gewöhnliche Ereignisvariablen, *R* Rekonstruktionsvariablen, *A* gewöhnliche Ereignis- und Rekonstruktionsvariablen.

von $-0,288$ auf, sodass im Kontext der zweidimensionalen Anpassung der Ansatz zur Vereinigung der Variablensätze erneut als effektiv herausstellt.

Sowohl bei Aufteilung nach Anzahl der Teilchenjets als auch nach Anzahl der b-Tags resultiert die ausschließliche Nutzung von gewöhnlichen Ereignisvariablen in hohen Korrelationen der Signalstärkeparameter, sodass auch in diesem Kontext der positive Einfluss der χ^2 -Rekonstruktion bestätigt ist. Gegenüber der Aufteilung nach Anzahl der Teilchenjets weisen die Korrelationen bei Aufteilung der Diskriminatoren nach Anzahl der b-Tags für sämtliche Variablensätze eine Tendenz zu betragsmäßig höheren Werten auf.

7.4. Fazit über Wahl der Regionen

Die Ergebnisse der Untersuchungen der Auswahl der Variablen weisen sowohl hinsichtlich der Trennkraft des Klassifizierers als auch hinsichtlich der Anpassung ähnliche Tendenzen auf. Die ROC-Integralwerte in Abbildung 7.2 (links) sind in Regionen aufgeteilt nach Anzahl der Teilchenjets geringer als bei Aufteilung nach Anzahl der b-Tags. Die Klassifizierungseffizienzen in Abbildung 7.3 (links) zeigen gegensätzliches Verhalten. Aus der Leistung im Training geht die zu präferierende Methode nicht eindeutig hervor.

Bei Aufteilung der Diskriminatoren nach Anzahl der b-Tags stützen sich niedrige Unsicherheiten überwiegend auf Rekonstruktionsvariablen, während in den Regionen aufgeteilt nach Anzahl der Teilchenjets der Einfluss mit gewöhnlichen Ereignisvariablen ausgeglichen ist. Die niedrigste Unsicherheit unter den aus Vereinigung gewonnenen Variablensätzen ergibt sich aus Anpassung der ersteren. Derselbe Variablensatz $V(AB)$ resultiert bei simultaner Anpassung der Signalstärkeparameter für $t\bar{t}+Z$ - und $t\bar{t}+H$ -Produktion in der betragsmäßig geringsten Korrelation. Vor diesem Hintergrund sprechen die Ergebnisse dafür, weiterhin die Aufteilung der Ereignisse nach Anzahl der enthaltenen Teilchenjets vorzunehmen und zur Konstruktion der Diskriminatoren mit dedizierten ANNs zu behandeln.

8. Zusammenfassung und Ausblick

Die Messung des Wirkungsquerschnitts der $t\bar{t}$ +Z-Produktion am CMS-Experiment im Zerfallskanal $Z \rightarrow b\bar{b}$ stellt eine wichtige Prüfung der Vorhersagen des Standardmodells dar. Die genaue Kenntnis des Wirkungsquerschnitts sowie weiterer Charakteristika des $t\bar{t}$ +Z-Prozesses wirkt sich zukünftig positiv auf die Messung der $t\bar{t}$ +H-Produktion im Zerfallskanal $H \rightarrow b\bar{b}$ aus. Mit dem Rückschluss auf die Kopplungsstärke des Z-Bosons zum Top-Quark liefert die Analyse ferner Hinweise auf BSM-Wechselwirkungen und leistet damit langfristig einen bedeutenden Beitrag zu aktuellen Forschungsthemen.

Zur Verbesserung des bestehenden Prozesses für die $t\bar{t}$ +Z-Analyse wurden unterschiedliche Ansätze zur Konstruktion der Diskriminatoren aus ANNs erprobt. Für den Vergleich wurden die ANNs hinsichtlich der Trennkraft der dominanten teilchenphysikalischen Prozesse evaluiert und in Asimov-Tests mit den generierten Verteilungen erwartete Unsicherheiten bestimmt.

Gegenüber einer bestehenden Analyse [6] wurde eine Rekonstruktionsmethode für das erwartete Z-Boson sowie das erwartete Higgs-Bosons implementiert. Die Zuweisungshypothese stützt sich auf die quadratische Abweichung der Massen zweier ausgewählter Teilchenjets von der des jeweils erwarteten Bosons und wurde zur Ermittlung der Erfolgsrate mit Generatorinformation verglichen. Die aus der Rekonstruktion gewonnenen Ereignisvariablen wurden an ANNs übergeben und die Leistung des Klassifizierers hinsichtlich unterschiedlicher Parameter untersucht. Der Mehrwert der ergänzten Rekonstruktion wurde sowohl für die binäre Klassifikation von $t\bar{t}$ +Z- und $t\bar{t}$ +H-Ereignissen, als auch für die Multiklassifizierung der dominanten Prozesse der $t\bar{t}$ +Z-Analyse explizit nachgewiesen.

Bisher wurden für die Analyse verwendete Ereignisvariablen anhand der jeweiligen Gewichte nach dem Training eines ANN mit allen verfügbaren Variablen ausgewählt. Für den untersuchten Ansatz wurden die Variablen kategorisiert als gewöhnliche Ereignisvariablen und den aus der χ^2 -Rekonstruktion gewonnen Variablen. Eine erste Auswahl aus den jeweiligen Gruppen wurde anhand der Gewichte der Variablen nach Training eines dedizierten ANNs getroffen. Die Vereinigung der ausgewählten Rekonstruktions- und einfachen Ereignisvariablen in gleichem Verhältnis zur Übergabe an ANNs erweist sich als vielversprechender Ansatz für die Konstruktion der Diskriminatoren. Bezüglich der Auswahl der Variablen wurde in Asimov-Tests eine Verringerung der erwarteten Unsicherheit um etwa 20 % gegenüber dem bisher verwendeten Ansatz festgestellt.

Zuletzt wurden die Ereignisse nach der jeweiligen Anzahl der rekonstruierten Teilchenjets aufgeteilt und jeweils mit dedizierten ANNs behandelt. Die Anpassung der Diskriminatoren bei Aufteilung nach Anzahl der b-Tags wurde in der Studie gleichermaßen untersucht und die Ergebnisse untereinander verglichen. Ersterer Ansatz beweist sich weiterhin als effektiv.

Für die Integration systematischer Unsicherheiten für die Asimov-Tests wurde eine Auswahl aus den bisher berücksichtigten Unsicherheiten getroffen, um den Rechenaufwand zu

reduzieren. Die resultierenden Unsicherheiten nach der numerischen Anpassung sind deshalb als unvollständig anzunehmen und dienen im Wesentlichen dem Vergleich unterschiedlicher Ansätze.

Aktuell leistet der Wert des χ^2 unter aus der Rekonstruktion gewonnenen Variablen den größten Beitrag zur Klassifizierung, sodass dessen Berücksichtigung auch in Zukunft sinnvoll ist. Eine Erhöhung der Erfolgsrate der Rekonstruktion durch einen anderen Ansatz zur Gewinnung der Zuweisungshypothese, beispielsweise mithilfe eines ANNs oder eines Boosted Decision Trees, könnte den Mehrwert für die Klassifizierung weiter erhöhen. Unter diesen Umständen kann möglicherweise auch aus den kinematischen Verteilungen der physikalischen Objekte für die Klassifizierung wertvolle Information gewonnen werden. Zu derzeitigem Stand tragen diese Ereignisvariablen nur marginal zur Sensitivität der Analyse bei.

Die zuletzt verwendete Netzarchitektur der ANNs ist sehr pragmatisch gewählt. Durch systematische Optimierung sämtlicher in Tabelle 6.2 aufgeführter Parameter kann die Trennkraft des Klassifizierers sehr wahrscheinlich weiter gesteigert werden. Die Zugabe weiterer Informationen für den Klassifizierer über Schichten anderer Funktionalität, wie beispielsweise Faltungsschichten, resultiert möglicherweise ebenfalls in einer Erhöhung der Trennkraft.

Im weiteren Vorgehen wird ausgehend von den Ergebnissen der Untersuchung eine finale Auswahl von Variablen für die Analyse getroffen und hinsichtlich der Anpassung validiert. Zu diesem Zweck empfiehlt sich die Berücksichtigung sämtlicher für die Analyse vorgesehenen Unsicherheiten. Die Kombination der Messdaten aus den Jahren 2016 bis 2018 verspricht eine präzise Bestimmung des Wirkungsquerschnitts der $t\bar{t}+Z$ -Produktion.

Literaturverzeichnis

- [1] CMS Kollaboration, *Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC*, *Phys. Lett. B* **716** (2012) 30, [1207.7235].
- [2] ATLAS Kollaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett. B* **716** (2012) 1, [1207.7214].
- [3] CMS Kollaboration, *Measurement of the cross section for top quark pair production in association with a W or Z boson in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **2018** (2018) 11.
- [4] ATLAS Kollaboration, *Measurement of the $t\bar{t}Z$ and $t\bar{t}W$ cross sections in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Phys. Rev. D* **99** (2019) 072009.
- [5] CMS Kollaboration, *Measurement of top quark pair production in association with a Z boson in proton-proton collisions at $\sqrt{s} = 13$ TeV*, Berichtnr. CMS-PAS-TOP-18-009, CERN, Genf, 2019.
- [6] J. van der Linden, *Limit on $t\bar{t}+Z$ production in the $Z \rightarrow b\bar{b}$ channel at the CMS experiment*, Masterarbeit, Karlsruher Institut für Technologie (KIT), 2019.
- [7] D. J. Griffiths, *Introduction to elementary particles*. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, 2., überarbeitete Aufl., 2008.
- [8] F. Wilczek, *Quantum Chromodynamics (QCD): The Modern Theory of the Strong Interaction*, *Ann. Rev. Nucl. Part. Sci.* **32** (1982) 177.
- [9] M. Gell-Mann, *Symmetries of baryons and mesons*, *Phys. Rev.* **125** (1962) 1067.
- [10] M. Gell-Mann, *A schematic model of baryons and mesons*, *Physics Letters* **8** (1964) 214.
- [11] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, *Phys. Rev. Lett.* **13** (1964) 508.
- [12] PARTICLE DATA GROUP, *Review of particle physics*, *Phys. Rev. D* **98** (2018) 030001.
- [13] L.-L. Chau, *Quark mixing in weak interactions*, *Physics Reports* **95** (1983) 1.
- [14] A. Kulesza et al., *Associated production of a top quark pair with a heavy electroweak gauge boson at NLO+NNLL accuracy*, *The European Physical Journal C* **79** (2019) 249.
- [15] L. Evans und P. Bryant, *LHC Machine*, *JINST* **3** (2008) S08001.
- [16] ALICE Kollaboration, *The ALICE experiment at the CERN LHC*, *JINST* **3** (2008) S08002.

- [17] ATLAS Kollaboration, *The ATLAS experiment at the CERN large hadron collider*, *JINST* **3** (2008) S08003.
- [18] CMS Kollaboration, *The CMS Experiment at the CERN LHC*, *JINST* **3** (2008) S08004.
- [19] LHCb Kollaboration, *The LHCb detector at the LHC*, *JINST* **3** (2008) S08005.
- [20] CMS Kollaboration, *CMS Technical Design Report for the Pixel Detector Upgrade*, Berichtnr. CERN-LHCC-2012-016, CMS-TDR-011, CERN, 2012. DOI: 10.2172/1151650.
- [21] S. R. Davis, *Interactive Slice of the CMS detector*, Aug., 2016, <https://cds.cern.ch/record/2205172>, abgerufen 09.03.2020.
- [22] CMS Kollaboration, *Particle-flow reconstruction and global event description with the CMS detector*, *JINST* **12** (2017) P10003, [1706.04965].
- [23] M. Cacciari, G. P. Salam und G. Soyez, *The anti- k_t jet clustering algorithm*, *JHEP* **04** (2008) 063, [0802.1189].
- [24] M. Cacciari, G. P. Salam und G. Soyez, *FastJet user manual*, *Eur. Phys. J. C* **72** (2012) 1896, [1111.6097].
- [25] CMS Kollaboration, *Description and performance of track and primary-vertex reconstruction with the CMS tracker*, *JINST* **9** (2014) P10009.
- [26] CMS Kollaboration, *Performance of the DeepJet b-tagging algorithm using 41.9/fb of data from proton-proton collisions at 13 TeV with Phase 1 CMS detector*, *CMS intern* (2018) .
- [27] J. C. Collins, D. E. Soper und G. F. Sterman, *Factorization of Hard Processes in QCD*, *Adv. Ser. Direct. High Energy Phys.* **5** (1989) 1, [hep-ph/0409313].
- [28] S. Alioli et al., *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, *JHEP* **06** (2010) 043, [1002.2581].
- [29] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *JHEP* **2014** (2014) 79.
- [30] G. Cowan et al., *Asymptotic formulae for likelihood-based tests of new physics*, *The European Physical Journal C* **71** (2011) 1554.
- [31] CMS Kollaboration, *CMS combine tool*, <http://cms-analysis.github.io/HiggsAnalysis-CombinedLimit/>.
- [32] U. Lämmel, *Künstliche Intelligenz*. Hanser Verlag, München, 4., aktualisierte Aufl., 2012.
- [33] LHC HIGGS COMBINATION GROUP, *Procedure for the LHC Higgs boson search combination in Summer 2011*, Berichtnr. CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11, CERN, Genf, Aug., 2011.
- [34] J. van der Linden et al., *DRACO-MLfoy*, 2020, <https://github.com/kit-cn-cms/DRACO-MLfoy>.
- [35] M. Abadi et al., *TensorFlow: Large-scale machine learning on heterogeneous systems*, 2015, Software verfügbar auf tensorflow.org.

- [36] D. Svozil, V. Kvasnicka und J. Pospichal, *Introduction to multi-layer feed-forward neural networks*, *Chemometrics and Intelligent Laboratory Systems* **39** (1997) 43.
- [37] N. Srivastava et al., *Dropout: A simple way to prevent neural networks from overfitting*, *Journal of Machine Learning Research* **15** (2014) 1929.
- [38] A. Y. Ng, *Feature Selection, L1 vs. L2 Regularization, and Rotational Invariance*, in *Proceedings of the Twenty-First International Conference on Machine Learning*, ICML '04, (New York, NY, USA), S. 78, Association for Computing Machinery, 2004, DOI: 10.1145/1015330.1015435.
- [39] J. Duchi, E. Hazan und Y. Singer, *Adaptive subgradient methods for online learning and stochastic optimization*, *Journal of Machine Learning Research* **12** (2011) 2121.
- [40] D. J. Hand und R. J. Till, *A Simple Generalisation of the Area Under the ROC Curve for Multiple Class Classification Problems*, *Mach. Learn.* **45** (2001) 171.
- [41] D. Bamber, *The area above the ordinal dominance graph and the area below the receiver operating characteristic graph*, *Journal of Mathematical Psychology* **12** (1975) 387.
- [42] CMS Kollaboration, *CMS luminosity measurement for the 2017 data-taking period at $\sqrt{s} = 13$ TeV*, Berichtnr. CMS-PAS-LUM-17-004, CERN, Genf, 2018.

Anhang

A. Verteilungen der Variablen aus der χ^2 -Rekonstruktion der Bosonen

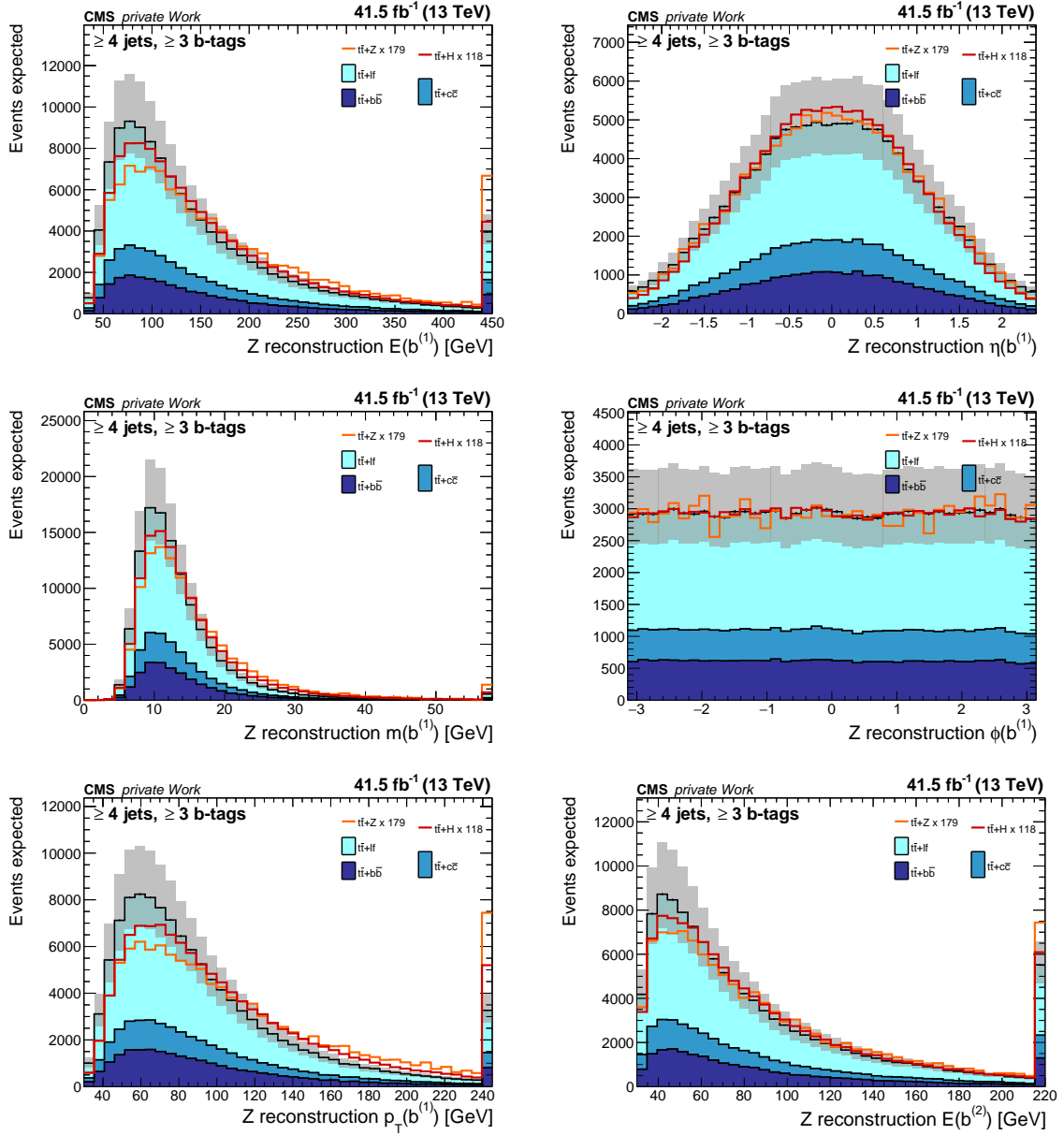


Abbildung A.1.: Verteilungen der Variablen aus der χ^2 -Rekonstruktion des Z-Bosons (1).

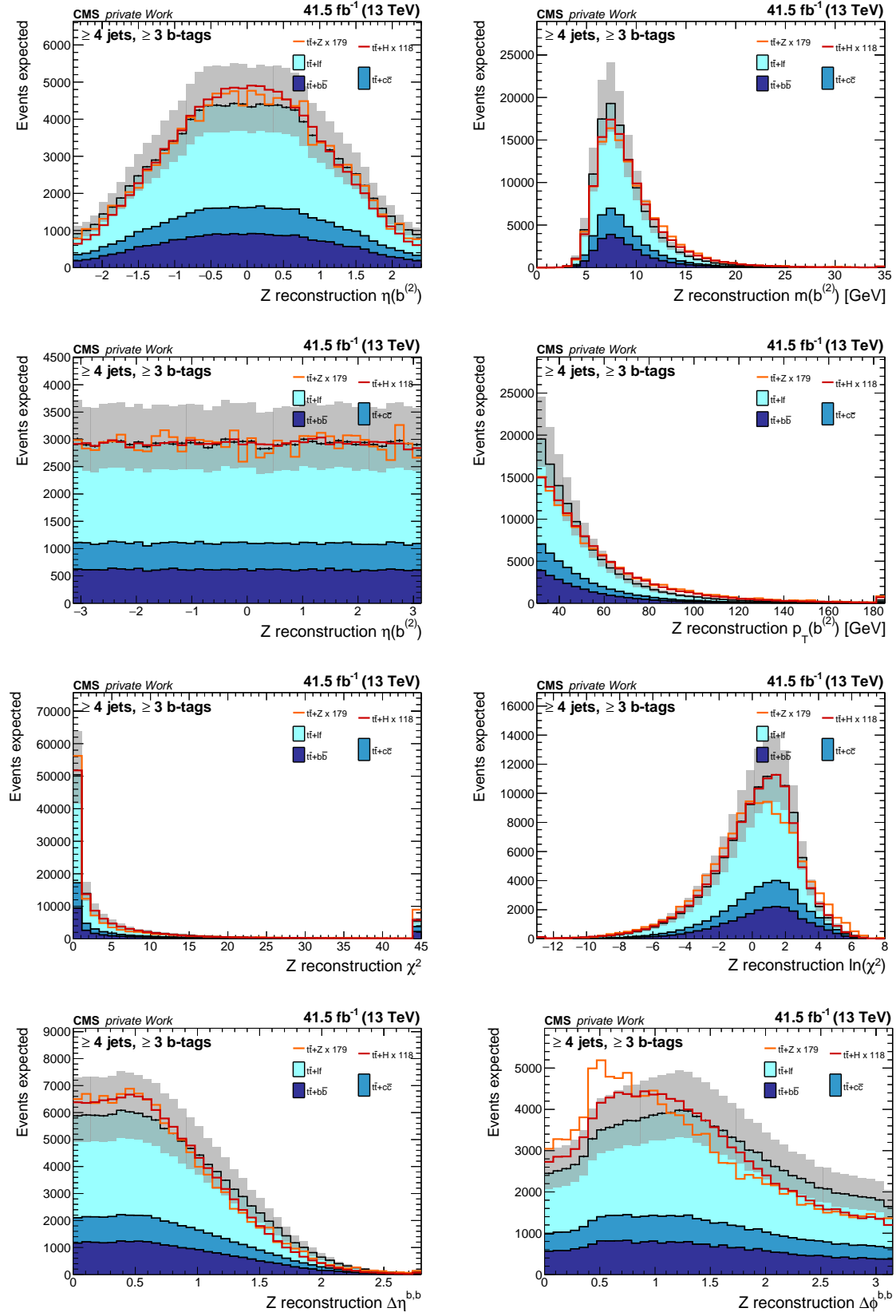


Abbildung A.2.: Verteilungen der Variablen aus der χ^2 -Rekonstruktion des Z-Bosons (2).

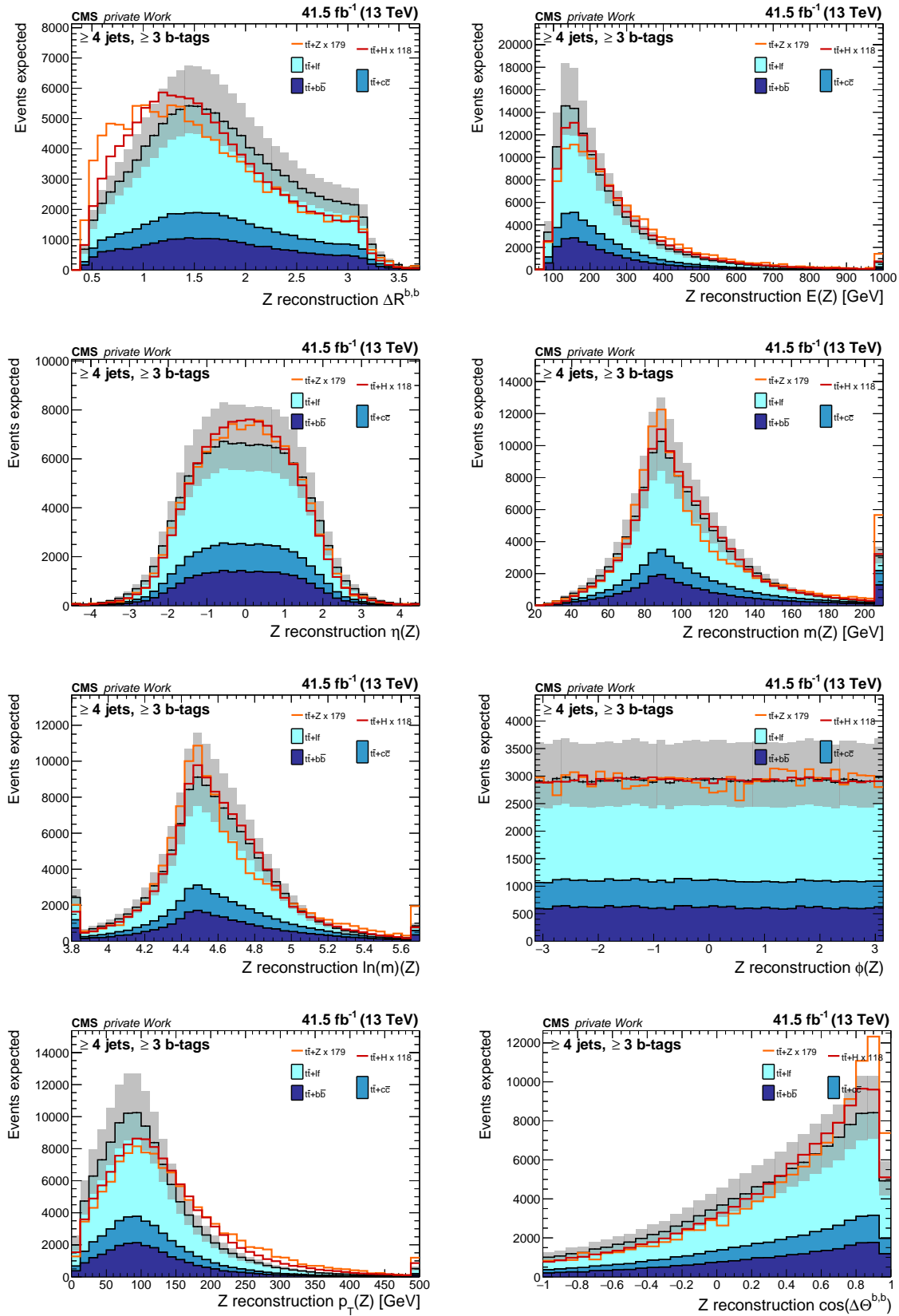


Abbildung A.3.: Verteilungen der Variablen aus der χ^2 -Rekonstruktion des Z-Bosons (3).

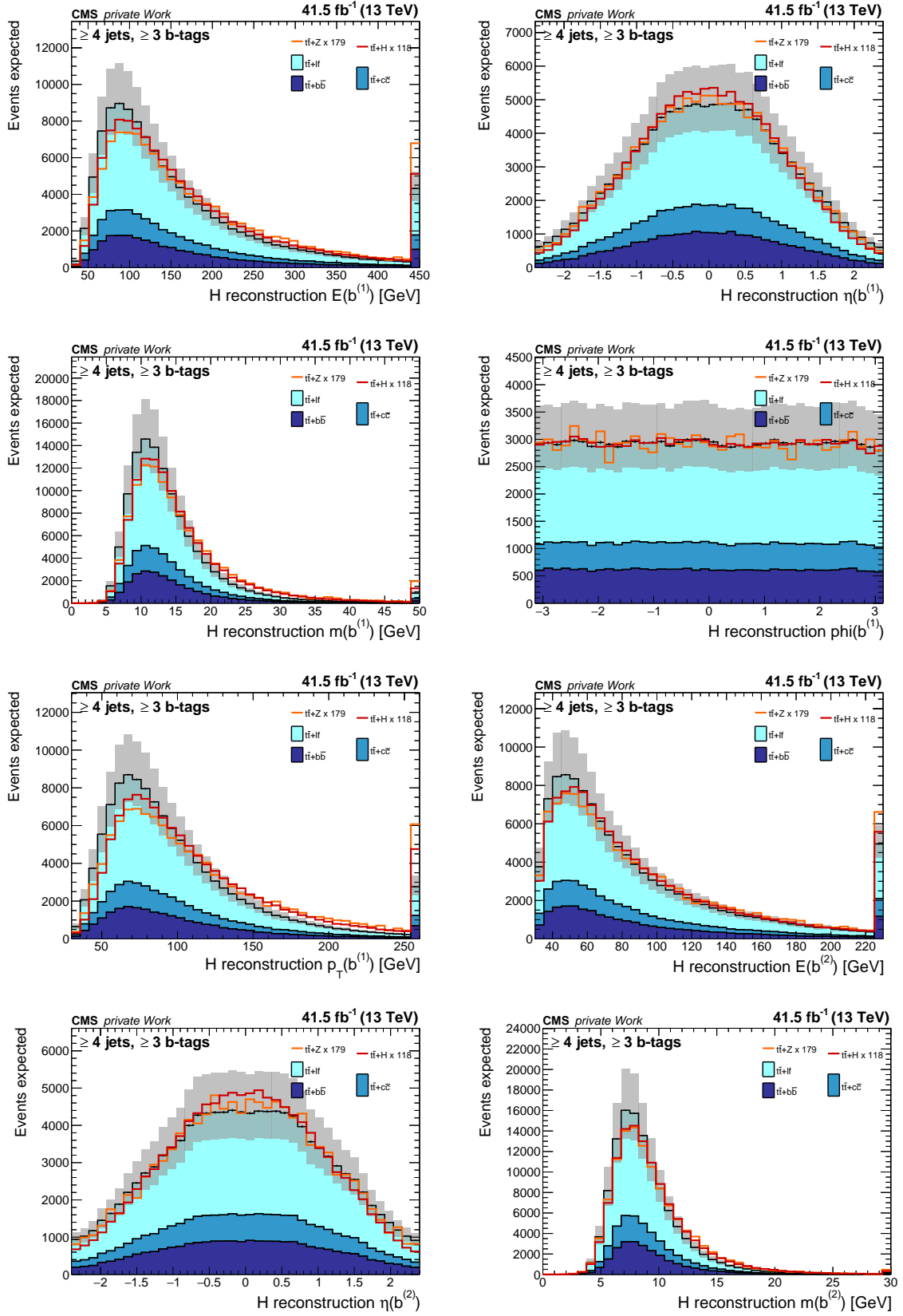


Abbildung A.4.: Verteilungen der Variablen aus der χ^2 -Rekonstruktion des Higgs-Bosons (1).

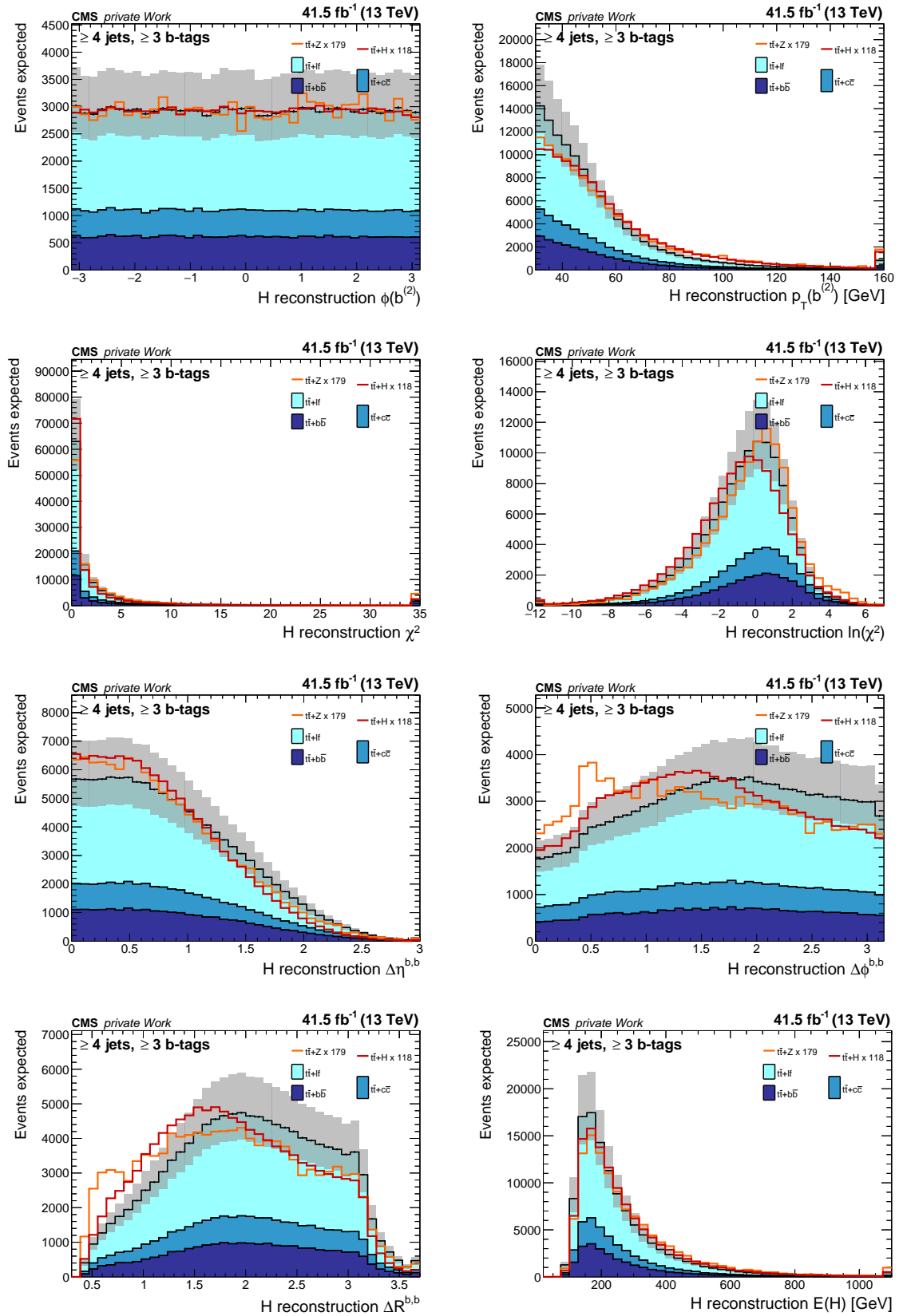


Abbildung A.5.: Verteilungen der Variablen aus der χ^2 -Rekonstruktion des Higgs-Bosons (2).

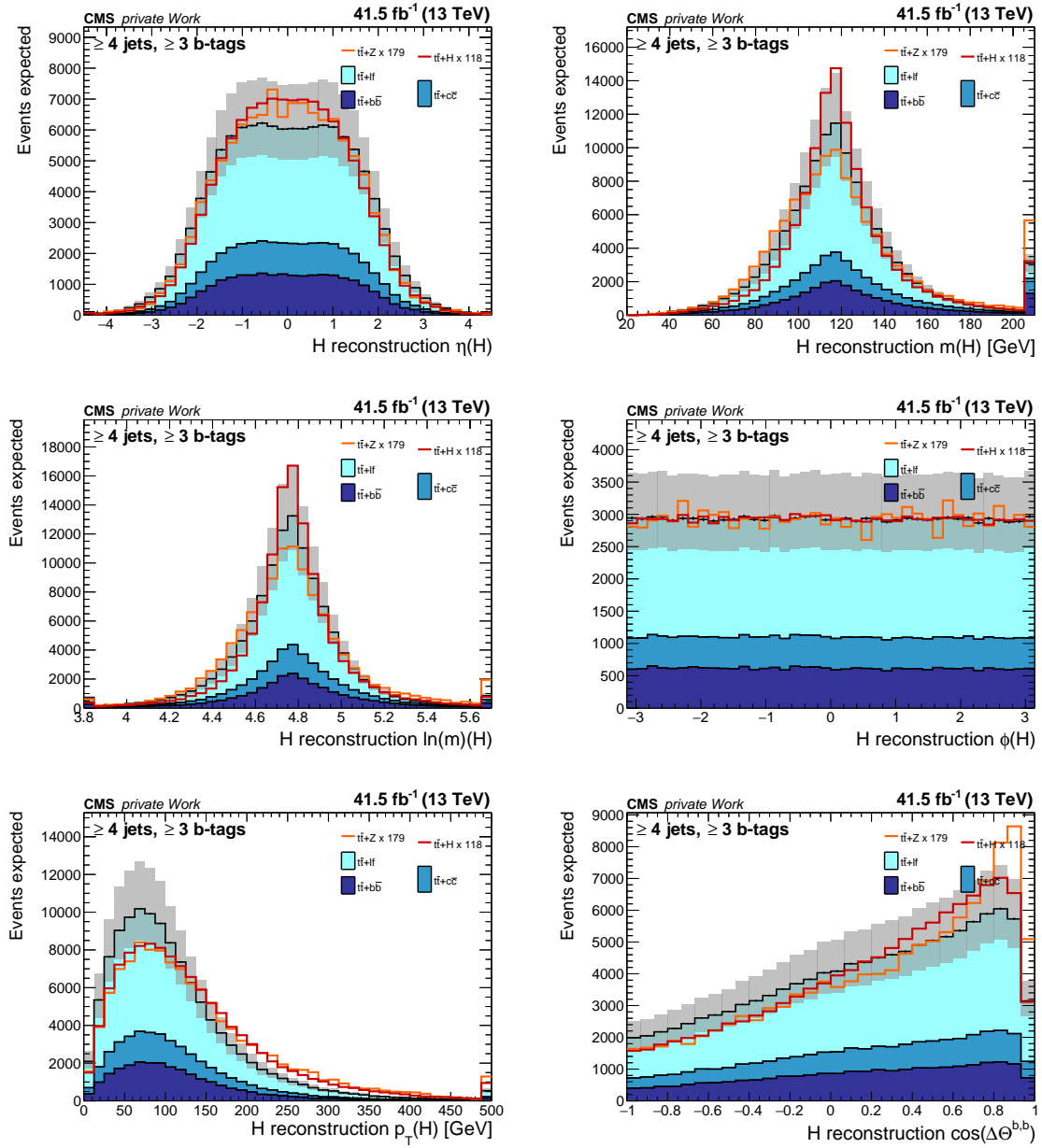


Abbildung A.6.: Verteilungen der Variablen aus der χ^2 -Rekonstruktion des Higgs-Bosons (3).

B. Übersicht der Variablensätze

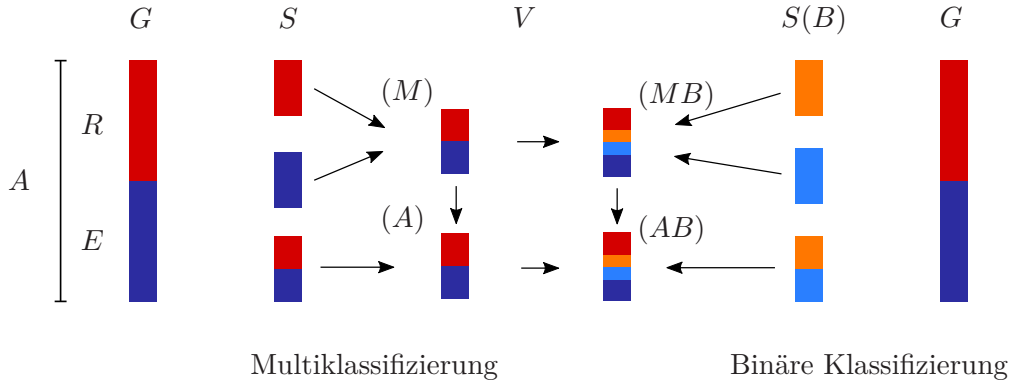


Abbildung B.1.: Übersicht der Zusammensetzung der Variablensätze. Aus dem Training mit Variablensätzen der Gruppe G ergeben sich aus den am höchsten gewichteten Variablen die Sätze mit Bezeichnung S . Die Pfeile kennzeichnen die Rekombination der Variablensätze zur den Sätzen der Gruppe V .

C. Unsicherheiten mit internen Bezeichnungen

Die berücksichtigten Unsicherheiten sind in internem Jargon in Tabelle C.1 aufgeführt und werden nachfolgend knapp motiviert. Für eine ausführlichere Diskussion sei verwiesen auf die Dokumentation der bestehenden Analyse [6].

Integrierte Luminosität

Die Unsicherheit auf die integrierte Luminosität des Laufs des LHC im Jahr 2017 beläuft sich zu 2,3 % [42].

Partonverteilungsfunktionen (PDF)

Für die Simulation verwendete Annahmen für die Partonverteilungsfunktionen werden generiert aus Analysen der an Experimenten am LHC gewonnenen Daten. Aus der Anpassung resultierende Unsicherheiten systematischer und statistischer Natur werden als Gestaltungsunsicherheit berücksichtigt.

b-Tag-Unsicherheiten

Für bessere Übereinstimmung von simulierten Verteilungen bezüglich der Identifikation von Teilchenjets aus Bottom-Quarks zu Detektordaten werden Skalierungsfaktoren angewandt. Die Ableitung der Skalierungsfaktoren ist mit Unsicherheiten auf die Ergebnisse verbunden. Die Unsicherheiten werden als Ratenunsicherheiten in weitere Berechnungen propagiert.

Untergrundnormierung

Unsicherheiten auf die Raten der Untergrundprozesse werden als Ratenunsicherheit in der Anpassung berücksichtigt. Auf diese Weise werden potentielle Unterschiede aus Variation der Generatoren und damit den zugrundeliegenden Prinzipien berücksichtigt.

Ausgangs- und Endzustand (ISR, FSR)

Die Kopplungskonstante α_s der starken Wechselwirkung hat wesentlichen Einfluss auf das Verhalten der Simulation der Schauerbildung mit PYTHIA 8.2. Die Unsicherheit α_s wird getrennt behandelt für die Abstrahlung aus Eingangszuständen (ISR) und Endzuständen (FSR).

Pile-Up (PU)

Der Betrieb des LHC mit Proton-Bunches hat zur Folge, dass bei Kreuzung der Strahlen

jeweils mehrere Protonen der kollidierenden Bunches wechselwirken. In Folge der zusätzlichen Interaktionen wird zusätzliche Strahlung (Pile-Up) im Detektor registriert, welche die tatsächliche Messung der Produkte der primären Interaktion verfälscht. Die zugehörige Unsicherheit wird als Gestaltunsicherheit in der Anpassung berücksichtigt.

Tabelle C.1.: Berücksichtigte Unsicherheiten mit internen Bezeichnungen.

Bezeichnung	Typ	Konstruktion	Wertebereich (%)
lumi_13TeV	lnN	Rate	2,3
QCDscale_ttH	lnN	Rate	-9,2/+5,8
QCDscale_ttbar			-3,5/+2,4
QCDscale_singlet			-2,1/+3,1
QCDscale_V			—
QCDscale_VV			—
pdf_Higgs_ttH	lnN	Rate	+3,6
pdf_gg			+4,2
pdf_qqbar			—
pdf_qg			+2,8
CMS_btag_lf	Gestalt	Gewicht	—
CMS_btag_lfstats1			—
CMS_btag_lfstats2			—
CMS_btag_hf			—
CMS_btag_hfstats1			—
CMS_btag_hfstats2			—
CMS_btag_cferr1			—
CMS_btag_cferr2			—
CMS_ttHbb_tt2b_glusplit	Gestalt	Gewicht	—
CMS_ttHbb_bgnorm_ttbb	lnN	Rate	—
CMS_ttHbb_bgnorm_ttcc	lnN	Rate	—
CMS_ttHbb_PU	Gestalt	Gewicht	—
CMS_ttHbb_PDF	Gestalt	Gewicht	—
CMS_ttHbb_PDF_ttbbNLO	Gestalt	Gewicht	—
CMS_ISR_ttbb	Gestalt	Gewicht	—
CMS_ISR_ttcc			—
CMS_ISR_ttlf			—
CMS_FSR_ttbb	Gestalt	Gewicht	—
CMS_FSR_ttcc			—
CMS_FSR_ttlf			—

D. Ergebnisse der Asimov-Tests

Tabelle D.1.: Erwartete Unsicherheiten und Limits nach numerischer Anpassung in Asimov-Tests. Die Signalstärkeparameter nach der Anpassung betragen ausnahmslos Eins. Die Zeilen mit den jeweils geringsten Unsicherheiten auf die Anpassung sind farbig hinterlegt.

Satz	Unsicherheit $\mu(t\bar{t}+Z)$	Unsicherheit $\mu(t\bar{t}+Z)$ (stat.)	Limit 95 %CL	Limit 95 %CL (stat.)
Aufteilung nach Anzahl der Teilchenjets				
<i>GA</i>	+0,754 / -0,747	+0,332 / -0,329	1,45	0,65
<i>GE</i>	+0,877 / -0,867	+0,344 / -0,342	1,70	0,67
<i>GR</i>	+0,906 / -0,917	+0,379 / -0,376	1,76	0,74
<i>SA</i>	+0,936 / -0,921	+0,359 / -0,357	1,82	0,70
<i>SE</i>	+0,965 / -0,929	+0,350 / -0,348	1,87	0,68
<i>SR</i>	+1,176 / -1,188	+0,395 / -0,393	2,29	0,77
<i>V(AB)</i>	+0,755 / -0,739	+0,341 / -0,338	1,45	0,66
<i>V(A)</i>	+0,837 / -0,819	+0,350 / -0,348	1,62	0,68
<i>V(MB)</i>	+0,885 / -0,876	+0,351 / -0,349	1,71	0,69
<i>V(M)</i>	+0,813 / -0,802	+0,350 / -0,347	1,57	0,68
<i>K</i>	—	—	—	—
<i>L</i>	+0,956 / -0,927	+0,361 / -0,358	1,85	0,70
Aufteilung nach Anzahl der b-Tags				
<i>GA</i>	+0,726 / -0,722	+0,334 / -0,331	1,39	0,65
<i>GE</i>	+0,961 / -0,967	+0,349 / -0,347	1,87	0,68
<i>GR</i>	+0,815 / -0,812	+0,365 / -0,363	1,57	0,71
<i>SA</i>	+0,935 / -0,920	+0,358 / -0,356	1,79	0,70
<i>SE</i>	+1,088 / -1,069	+0,367 / -0,365	2,13	0,72
<i>SR</i>	+0,843 / -0,842	+0,368 / -0,365	1,63	0,72
<i>V(AB)</i>	+0,853 / -0,834	+0,355 / -0,353	1,65	0,69
<i>V(A)</i>	+0,837 / -0,807	+0,356 / -0,353	1,60	0,69
<i>V(MB)</i>	+0,818 / -0,799	+0,354 / -0,351	1,57	0,69
<i>V(M)</i>	+0,826 / -0,782	+0,354 / -0,351	1,57	0,69
<i>K</i>	+0,762 / -0,746	+0,351 / -0,348	1,46	0,68
<i>L</i>	—	—	—	—

Tabelle D.2.: Unsicherheiten und Korrelation bei zweidimensionaler Anpassung an Asimov-Pseudodaten. Die Signalstärkeparameter für die $t\bar{t}+Z$ - und die $t\bar{t}+H$ -Produktion wurden für die Anpassung jeweils als frei variierbar angenommen. Die Zeilen mit der geringsten Korrelation sind im jeweiligen Bereich farbig hinterlegt.

Satz	Unsicherheit $\mu(t\bar{t}+Z)$	Unsicherheit $\mu(t\bar{t}+H)$	Korrelation
Aufteilung nach Anzahl der Teilchenjets			
<i>GA</i>	+0,804 / -0,797	+0,441 / -0,413	-0,338
<i>GE</i>	+1,194 / -1,200	+0,598 / -0,565	-0,682
<i>GR</i>	+1,024 / -1,003	+0,633 / -0,645	-0,444
<i>SA</i>	+1,070 / -1,062	+0,547 / -0,520	-0,483
<i>SE</i>	+1,183 / -1,149	+0,629 / -0,600	-0,576
<i>SR</i>	+1,333 / -1,296	+0,853 / -0,869	-0,439
<i>V(AB)</i>	+0,791 / -0,774	+0,469 / -0,444	-0,288
<i>V(A)</i>	+0,903 / -0,884	+0,506 / -0,484	-0,365
<i>V(MB)</i>	+0,953 / -0,937	+0,487 / -0,463	-0,354
<i>V(M)</i>	+0,865 / -0,851	+0,497 / -0,479	-0,335
<i>K</i>	—	—	—
<i>L</i>	+1,050 / -1,020	+0,533 / -0,505	-0,404
Aufteilung nach Anzahl der b-Tags			
<i>GA</i>	+0,809 / -0,798	+0,388 / -0,358	-0,431
<i>GE</i>	+1,359 / -1,379	+0,537 / -0,496	-0,714
<i>GR</i>	+0,901 / -0,882	+0,608 / -0,604	-0,412
<i>SA</i>	+1,027 / -1,005	+0,467 / -0,438	-0,404
<i>SE</i>	+1,425 / -1,427	+0,524 / -0,486	-0,651
<i>SR</i>	+0,897 / -0,880	+0,631 / -0,626	-0,316
<i>V(AB)</i>	+0,925 / -0,897	+0,432 / -0,406	-0,372
<i>V(A)</i>	+0,904 / -0,864	+0,455 / -0,434	-0,363
<i>V(MB)</i>	+0,892 / -0,862	+0,426 / -0,400	-0,378
<i>V(M)</i>	+0,901 / -0,840	+0,452 / -0,429	-0,382
<i>K</i>	+0,823 / -0,798	+0,444 / -0,413	-0,366
<i>L</i>	—	—	—