

Bestimmung von Trigger-Effizienzen und zugehörigen Daten-Simulation- Skalenfaktoren im Rahmen einer Monotop-Analyse am CMS-Experiment

Determination of Trigger Efficiencies and Associated
Data Simulation Scale Factors in the Scope of a
Monotop Analysis at the CMS Experiment

Bachelorarbeit

Jonas Janik

An der Fakultät für Physik
Institut für Experimentelle Teilchenphysik

Erstgutachter:

Prof. Dr. Ulrich Husemann

Zweitgutachter:

Dr. Michael Waßmer

Karlsruhe, 20. Juli 2022

Inhaltsverzeichnis

1	Einleitung	1
2	Grundlagen	3
2.1	Standardmodell der Teilchenphysik	3
2.2	Large Hadron Collider	6
2.3	Compact Muon Solenoid	6
2.4	Ereignisrekonstruktion und -generation	7
3	Suche nach Dunkler Materie in Monotop-Prozessen mit CMS	9
3.1	Dunkle Materie	9
3.1.1	Evidenz	9
3.1.2	Kandidaten für Dunkle Materie	11
3.2	Hadronische Monotop-Analyse	13
3.2.1	Monotop-Prozess	13
3.2.2	Untergrundprozesse	14
3.2.3	Ereignisselektion	16
4	CMS-Trigger und ihre Effizienzen	19
4.1	Das CMS-Triggersystem	19
4.2	Trigger-Effizienzen	20
4.3	Auswahl der Trigger	22
5	Studien zu MET-Trigger-Effizienzen	23
5.1	Vergleich von Monte-Carlo-Datensätzen	24
5.2	Auswirkungen der Ereignisgewichtung	26
5.3	Vergleich von Kontrollregionen	28
6	Bestimmung von Daten-Simulation-Skalenfaktoren	35
7	Zusammenfassung und Ausblick	41
	Literaturverzeichnis	43

1. Einleitung

Ein Rätsel der modernen Physik ist die Dunkle Materie. Ihre Identität ist bislang ungeklärt, da sie nur gravitativ (und womöglich auch schwach) wechselwirkt und daher zum heutigen Stand nicht direkt detektiert werden kann. Ihre Existenz wird jedoch durch astronomische und kosmologische Beobachtungen stark motiviert.

In einigen Theorien wird angenommen, dass der heute vorhandene Anteil an Dunkler Materie im Universum aus einer Zeit kurz nach dem Urknall stammt. Um bei der Entschlüsselung die Auswahl an Kandidaten für die Elementarteilchen der Dunklen Materie einzugrenzen, sind demnach Experimente notwendig, in denen unter Laborbedingungen und in möglichst hoher Anzahl Ereignisse hervorgerufen werden können, bei denen ähnliche Bedingungen herrschen wie kurz nach dem Urknall. Ein solches Experiment ist das Compact-Muon-Solenoid-Experiment (CMS) [1], das am Large Hadron Collider (LHC) [2] am Kernforschungszentrum CERN bei Genf betrieben wird. Hierbei werden im LHC gegenläufig beschleunigte Protonen paarweise bei einer Schwerpunktsenergie von $\sqrt{s} = 13 \text{ TeV}$ zur Kollision gebracht. Bei einer solchen Kollision kann im Rahmen des Energieerhaltungssatzes eine Vielzahl verschiedener - oft kurzlebiger - Teilchen entstehen, deren Zerfallsprodukte im schichtförmigen Detektorsystem des CMS-Detektors rekonstruiert werden. Da die Bewegung der Protonen annähernd auf einer Achse stattfindet, müssen aufgrund des Impulserhaltungssatzes alle Kollisionsprodukte in der Summe einen vernachlässigbaren Impuls senkrecht zur Strahlachse (Transversalimpuls) besitzen. Wird ein Transversalimpuls gemessen, der sich wesentlich von Null unterscheidet, so sind Teilchen entstanden, die den Detektor unregistriert passiert haben. Der unbeobachtete Transversalimpuls wird fehlender Transversalimpuls bzw. fehlende Transversalenergie (MET) genannt. Bei solchen Teilchen könnte es sich in seltenen Fällen hypothetisch um Dunkle Materie handeln. Auf dieser Idee basieren Analysen wie die sogenannte Monotop-Analyse [3].

Um die Dichte an interessanten seltenen Ereignissen für eine Analyse zu erhöhen und wegen Beschränkungen in der Datentransferrate und der Speicherkapazität, wird von Trigger-Systemen Gebrauch gemacht [4]. Ein solcher Trigger selektiert nur Ereignisse, die ein bestimmtes Mindestkriterium erfüllen. Im Fall der Monotop-Analyse ist dies ein MET-Trigger mit einer mindestens geforderten fehlenden Transversalenergie. Wie später gezeigt wird, ist eine Eigenschaft von Triggern, dass Ereignisse, die beispielsweise eine fehlende Transversalenergie nahe der geforderten Schwelle aufweisen, nicht immer wie erwünscht verworfen bzw. akzeptiert werden. Um einen negativen Einfluss auf die Analyse zu vermeiden, ist es notwendig, das Verhalten des Triggers genau zu kennen. Zu diesem Zweck werden in der folgenden Thesis für eine Monotop-Analyse Effizienz-Kurven von MET-Triggern bestimmt und studiert.

In der Monotop-Analyse werden darüber hinaus nicht nur aufgezeichnete sondern auch simulierte Daten eingesetzt. Da die Trigger-Effizienzen von aufgezeichneten und simulierten Daten im Allgemeinen nicht identisch sein müssen, besteht das Ziel insbesondere darin,

Daten-Simulation-Skalenfaktoren für die MET-Trigger-Effizienzen abzuleiten. Damit können simulierte Effizienzen in Übereinstimmung mit datenbasierten Effizienzen gebracht werden.

Zu Beginn werden in Kap. 2 die wichtigsten theoretischen und experimentellen Grundlagen für die Teilchenphysik am CMS-Experiment präsentiert. Kap. 3 thematisiert sowohl die Dunkle Materie als auch die Monotop-Analyse, mit der nach Dunkler Materie gesucht wird. Danach wird in Kap. 4 das CMS-Triggersystem sowie Trigger-Effizienzen im Allgemeinen skizziert. Ebenso werden die in dieser Arbeit verwendeten Trigger und eine Methode zur Effizienzbestimmung vorgestellt. Der Hauptteil besteht aus Studien zu MET-Trigger-Effizienzen, die in Kap. 5 diskutiert werden. Mit den daraus gewonnenen Erkenntnissen werden schließlich in Kap. 6 Daten-Simulation-Skalenfaktoren für die Monotop-Analyse abgeleitet. Zum Abschluss folgt in Kap. 7 eine Zusammenfassung inklusive Ansätze für weiterführende Studien zu Fragestellungen, die bis dahin nicht beantwortet worden sind.

2. Grundlagen

Im diesem Kapitel wird die theoretische Grundlage und das experimentelle Umfeld eingeführt, aus dem die Daten, auf denen diese Arbeit basiert, stammen. Die theoretische Basis ist mit dem Standardmodell der Teilchenphysik gegeben. Es umfasst alle derzeit bekannten Elementarteilchen sowie ihre Wechselwirkung untereinander und wird in Kap. 2.1 zusammengefasst. Da jedoch Indizien für das Vorhandensein von Teilchen existieren, die sich nicht allein mit Hilfe des Standardmodells erklären lassen, wie beispielsweise die sogenannte Dunkle Materie (s. Kap. 3.1), liegt es nahe, dass das Standardmodell noch nicht vollständig ist. Dies motiviert die Konstruktion geeigneter Experimente, um jene Theorien, die sich mit der Physik außerhalb des Standardmodells beschäftigen, zu prüfen. Die Grundlage für einige solcher Experimente stellen Teilchenbeschleuniger wie der LHC dar. Prägnante Beschreibungen dieses Beschleunigers sowie des am LHC befindlichen CMS-Experiments, mit dem unter anderem auch nach Dunkler Materie gesucht wird, folgen in Kap. 2.2 und Kap. 2.3. Zuletzt wird in Kap. 2.4 die Rekonstruktion und Generation von Ereignissen thematisiert.

2.1 Standardmodell der Teilchenphysik

Das Standardmodell der Teilchenphysik fasst die bekannten elementaren Teilchen und deren Wechselwirkungen unter Vernachlässigung der Gravitation zusammen. Elementar bedeutet in diesem Kontext, dass diese Teilchen nach heutigem Wissensstand nicht weiter teilbar sind.

Eine Übersicht der Teilchen im Standardmodell ist in Abb. 2.1 zu sehen. Teilchenmassen m werden in diesem Abschnitt aus Gründen der Lesbarkeit nur vereinzelt angegeben und können der Übersicht entnommen werden. Zudem wird die elektrische Ladung q von Teilchen in Vielfachen der Elementarladung e angegeben.

Die folgende Beschreibung des Standardmodells basiert auf [5]. Das Modell umfasst zwölf Fermionen (Spin- $\frac{1}{2}$ -Teilchen), vier Eichbosonen (Spin-1-Teilchen) und ein Skalarboson (Spin-0-Teilchen). Weiter lassen sich die zwölf Fermionen in zwei Gruppen mit sechs Quarks bzw. sechs Leptonen teilen. Sowohl bei Quarks als auch bei Leptonen unterscheidet man zwischen drei Generationen. Für jedes Teilchen der ersten Generation existiert je ein Teilchen der zweiten und dritten Generation, die sich nur in ihren Massen unterscheiden. Die Masse nimmt dabei mit der Generation zu.

Die erste Generation der Quarks besteht aus dem Up-Quark ($q = \frac{2}{3}$) und dem Down-Quark ($q = -\frac{1}{3}$). Die Up-artigen Quarks der höheren Generationen sind das Charm- und das Top-Quark. Die Down-artigen Quarks der höheren Generationen sind das Strange- und das Bottom-Quark.

Die Leptonen der ersten Generation sind das Elektron ($q = -1$) und das Elektron-Neutrino ($q = 0$). Die zugehörigen Leptonen höherer Generationen sind das Myon(-Neutrino) und das Tau(-Neutrino). Im Standardmodell sind Neutrinos masselos¹.

Im Standardmodell werden Wechselwirkungen zwischen Teilchen durch den Austausch von Eichbosonen (Gluonen, Photonen, Z-Bosonen und W-Bosonen) vermittelt.

Gluonen ($m = 0$, $q = 0$) sind die Austauschteilchen der starken Wechselwirkung und werden ausschließlich zwischen Teilchen mit den möglichen Farbladungen rot, grün und blau ausgetauscht. Die zugehörige Quantenfeldtheorie ist die Quantenchromodynamik (QCD). Da nur Partonen (Quarks und Gluonen) selbst Farbladungen tragen, wirkt die starke Wechselwirkung auch nur zwischen Quarks und/oder Gluonen. Quarks tragen dabei je eine Farbladung und Gluonen eine Kombination aus Farbladung und Antifarbladung. Somit ergeben sich unter Berücksichtigung der Farbladung für jede Quarkart drei mögliche Zustände. Um die Anzahl verschiedener Gluonen zu bestimmen wird von der Gruppentheorie Gebrauch gemacht. Das Ergebnis ist ein Oktett und ein Singulett, wobei der Singulett-Zustand $\sqrt{1/3}(\bar{r}\bar{r} + \bar{g}\bar{g} + \bar{b}\bar{b})$ aufgrund der Invarianz unter Rotationen im Farbraum nicht zwischen Farbladungen ausgetauscht werden kann. Das Oktett stellt somit die Menge der acht austauschbaren Gluonen dar. Teilchen, die eine Verbindung von Quarks oder Gluonen durch die starke Wechselwirkung darstellen, werden Hadronen genannt. Alle bisher beobachteten Hadronen sind Mesonen aus einem Quark und einem Anti-Quark oder Baryonen aus drei Quarks, jedoch gibt es auch Anzeichen für die Existenz von Tetra- und Pentaquarks [6, 7].

Eine Besonderheit der starken Wechselwirkung zwischen zwei Teilchen ist, dass ihre Stärke mit dem Abstand der Teilchen zunimmt. Die Wechselwirkung kann durch ein flussröhrenförmiges Feld beschrieben werden. Um den Abstand zwischen zwei Teilchen eines Hadrons zu erhöhen, muss dem Feld Energie zugeführt werden. Ab Abständen von ungefähr 1-2 fm werden aus der Feldenergie neue Teilchen gebildet. Das Ergebnis dieses als Hadronisierung bezeichneten Prozesses sind zwei Hadronen. Hadronisierung kann bei hochenergetischen Teilchenkollisionen, an denen mindestens ein Hadron beteiligt ist, auftreten. Sind die Teilchenkollisionen sehr hochenergetisch, dann kann sich der Vorgang mehrfach wiederholen, sodass sich Schauer an Hadronen bilden. Bei einem solchen Schauer bewegen sich eine Vielzahl an Hadronen und weiterer Teilchen annähernd in dieselbe Richtung und können als Jet zusammengefasst werden. Eine Folge der Hadronisierung von Partonen ist außerdem, dass sie nicht isoliert beobachtbar sind (*Confinement*).

Die Bosonen Z^0 ($m = 91,2 \frac{\text{GeV}}{c^2}$, $q = 0$) und W^\pm ($m = 80,4 \frac{\text{GeV}}{c^2}$, $q = \pm 1$) vermitteln die schwache Wechselwirkung, indem sie an die schwache Ladung koppeln. Es wird zwischen neutralen Strömen bei Z^0 -Austausch und geladenen Strömen bei W^\pm -Austausch unterschieden.

Zur Beschreibung der schwachen Wechselwirkung werden den Quarks (und analog den Leptonen) verschiedene Flavour-Quantenzahlen zugeordnet. Beim Up- und Down-Quark ist diese Quantenzahl die dritte Komponente des Isospins I_z . Die übrigen Flavours werden nach den Quarks Charm, Topness, Strangeness und Bottomness benannt. Bei geladenen Strömen der schwachen Wechselwirkung sind die Flavour-Quantenzahlen nicht erhalten. Flavourverändernde neutrale Ströme (FCNC) sind dagegen stark unterdrückt und nur in höheren Ordnungen möglich, weshalb der Flavour bei neutralen Strömen im Standardmodell erhalten ist. Auch bei der starken Wechselwirkung ist der Flavour eine Erhaltungsgröße. Photonen ($m = 0$, $q = 0$) sind die Eichbosonen der elektromagnetischen Wechselwirkung, die zwischen Teilchen mit einer elektrischen Ladung wirkt. Im Gegensatz zu den W-Bosonen, Z-Bosonen und Gluonen können Photonen jedoch nicht an andere gleichnamige

¹Aufgrund der Beobachtung von Neutrinooszillationen am Super-Kamiokande[8] müssen Neutrinos massesehaftet sein. Die genauen Massen sind bislang nicht bekannt, jedoch sollen mit Experimenten wie KATRIN[9] die aktuellen Obergrenzen der Neutrinomassen weiter reduziert werden.

Eichbosonen koppeln.

Im Standardmodell werden darüber hinaus die elektromagnetische und die schwache Wechselwirkung in der elektroschwachen Wechselwirkung zusammengefasst. In dieser Theorie sind das Photon γ und das Z^0 -Boson zueinander orthogonale Linearkombinationen der Eichbosonen B^0 und W^0 :

$$|\gamma\rangle = \cos \theta_W |B^0\rangle + \sin \theta_W |W^0\rangle \quad (2.1)$$

$$|Z^0\rangle = -\sin \theta_W |B^0\rangle + \cos \theta_W |W^0\rangle. \quad (2.2)$$

Hierbei ist θ_W der sogenannte Weinberg-Winkel.

Das Higgs-Boson ($m = 125,10 \frac{\text{GeV}}{c^2}$, $q = 0$)[11] ist das Quant des Higgs-Felds. In der Theorie der elektroschwachen Vereinheitlichung wird angenommen, dass alle Vektorbosonen oberhalb einer bestimmten Temperatur masselos sind. Erst unterhalb dieser Temperatur erhalten W^\pm - und Z^0 -Boson ihre Massen durch Wechselwirkung mit dem Higgs-Feld.

Die elementaren Fermionen bilden die Materie und für jede Art dieser Fermionen gibt es ein Antifermion, dass sich nur in elektrischer Ladung, Farbladung und dritter Komponente des schwachen Isospins vom dazugehörigen Fermion unterscheidet. Die elementaren Antifermionen bilden zusammen die Antimaterie. Ein Teilchen und sein Antiteilchen werden in den nachfolgenden Kapiteln auch als zueinander (ladungs-)konjugierte Teilchen bezeichnet.

Standard-Modell der Elementarteilchen

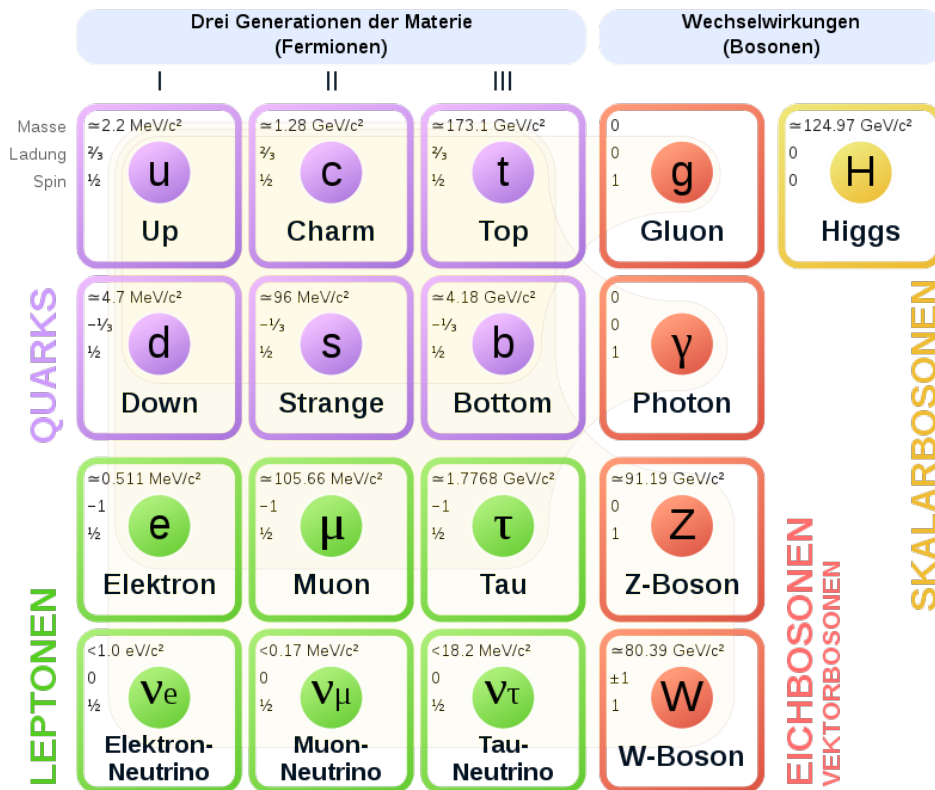


Abbildung 2.1: Eine Übersicht des Standardmodells der Teilchenphysik. Entnommen aus [10].

2.2 Large Hadron Collider

Die folgenden Ausführungen zum LHC basieren auf Informationen aus [2, 12].

Beim Large Hadron Collider (LHC) am Europäischen Kernforschungszentrum CERN bei Genf handelt es sich mit einem Umfang von 26,7 km um das derzeit größten Synchrotron der Welt. Über einen Linearbeschleuniger und zusätzliche Synchrotrone werden Protonen vorbeschleunigt, bevor sie in den LHC geführt werden. Dort werden die Protonenstrahlen in zwei separaten Röhren entgegengesetzt auf Geschwindigkeiten nahe der Lichtgeschwindigkeit beschleunigt. In gegenläufigen Strahlen werden an den vier Teilchendetektoren ALICE, ATLAS, LHCb und CMS dann bündelweise Protonen zur Kollision gebracht. Eine Kollision von zwei Teilchen wird in diesem Kontext auch als Ereignis bezeichnet. Die mittlere Rate $\frac{dN_P}{dt}$, mit der ein bestimmter Prozess P mit Wirkungsquerschnitt σ_P stattfindet, ist gegeben durch

$$\frac{dN_P}{dt} = \sigma_P L, \quad (2.3)$$

wobei die Luminosität L die Anzahl an Ereignissen pro Zeit und Fläche angibt. Innerhalb der zweiten Laufzeit des LHC von 2016 bis 2018² kreuzten sich Protonenbündel mit einer Rate von 40 MHz bei Proton-Proton-Schwerpunktsenergien von $\sqrt{s} = 13$ TeV. Harte Protonenkollision führen zur Produktion verschiedener Elementarteilchen durch Wechselwirkung zwischen Partonen. Viele der Kollisionsprodukte können dann von den unterschiedlichen Detektoren registriert werden. Je nach Detektor werden damit Evidenzen zu teils unterschiedlichen teilchenphysikalischen Theorien gesucht.

2.3 Compact Muon Solenoid

Das Compact Muon Solenoid (CMS) ist einer der Teilchendetektoren, mit dem Produkte von Proton-Proton-Kollisionen am LHC studiert werden. Den Namen trägt der Detektor, da seine Bauweise kompakt ist und Myonenspuren mit Hilfe von Myonenkammern im äußeren Magnetfeld eines supraleitenden Solenoids präzise gemessen werden können. Neben der erfolgreichen Entdeckung des Higgs-Bosons zeitgleich mit dem ATLAS-Experiment [13, 14] werden unter anderem auch verschiedene Modelle der Supersymmetrie und der Dunklen Materie untersucht.

Die nun folgende Beschreibung vom Aufbau des CMS-Detektors orientiert sich an [1, 15]. Der CMS-Detektor ist schichtförmig konstruiert. In Abb. 2.2 ist eine Skizze des Detektorquerschnitts zu sehen.

Die innerste Komponente ist der Silizium-Spurdetektor. Schichten aus Silizium-Pixel- und Silizium-Streifen-Detektoren dienen dort der genauen Ortsauflösung hochenergetischer geladener Teilchen. Fliegen solche Teilchen durch mehrere Schichten des Silizium-Spurdetektors, so lassen sich damit ihre Trajektorien nahe der Strahlachse rekonstruieren.

Der Silizium-Spurdetektor ist umgeben vom elektromagnetischen Kalorimeter (ECAL) zusammengesetzt aus Bleiwolframat-Kristallen, mit denen die Energie von Elektronen und Photonen durch Entstehen elektromagnetischer Schauer bestimmt wird.

Es folgt das hadronische Kalorimeter (HCAL), das wiederum aus abwechselnden Lagen von Messing als Absorbermaterial und Szintillatoren bestehen. Dadurch lassen sich die Energien von Hadronen ermitteln.

Die nächste Schicht stellt der supraleitende Solenoidmagnet mit einer magnetischen Flussdichte von 3,8 T im Innenbereich der Spule dar. Das Magnetfeld verursacht eine Krümmung

²In diesem Zeitraum sind die Daten für die Monotop-Analyse entnommen worden, zu der in Kap. 5 Studien durchgeführt und in Kap. 6 Skalenfaktoren berechnet werden.

der Trajektorien geladener Teilchen, was ihrer Identifikation dient: Aus der Krümmungsrichtung kann das Ladungsvorzeichen und aus der Krümmungsstärke der Impuls eines Teilchens bestimmt werden.

Alle bereits genannten Komponenten des CMS-Detektors werden von einem Eisenjoch umhüllt, das einerseits dem Zurückführen des Solenoid-Magnetfeldes und andererseits dem Beherbergen des Myonen-Spurdetektors dient [17]. Für den Myonen-Spurdetektor werden in der Region des Detektormantels Driftröhrendetektoren enthaltende Driftkammern verwendet - in der Region der Endkappen werden Kathoden-Streifenkammern eingesetzt. In beiden Regionen sind zusätzlich Widerstandsplattenkammern verbaut. Als minimal ionisierende Teilchen, die die Kalorimeter und das Solenoid fast ungehindert durchdringen, sind Myonen die einzigen Teilchen, die vom Myonen-Spurdetektor erfasst werden.

Insgesamt lassen sich durch diesen Aufbau wichtige Eigenschaften von Elektronen, Photonen, Myonen und verschiedenen Hadronen bestimmen, wodurch die genannten Teilchen mit hoher Präzision rekonstruiert werden können. Dies ermöglicht eine Analyse von Prozessen, die bei einer hochenergetischen Proton-Proton-Kollision stattfinden.

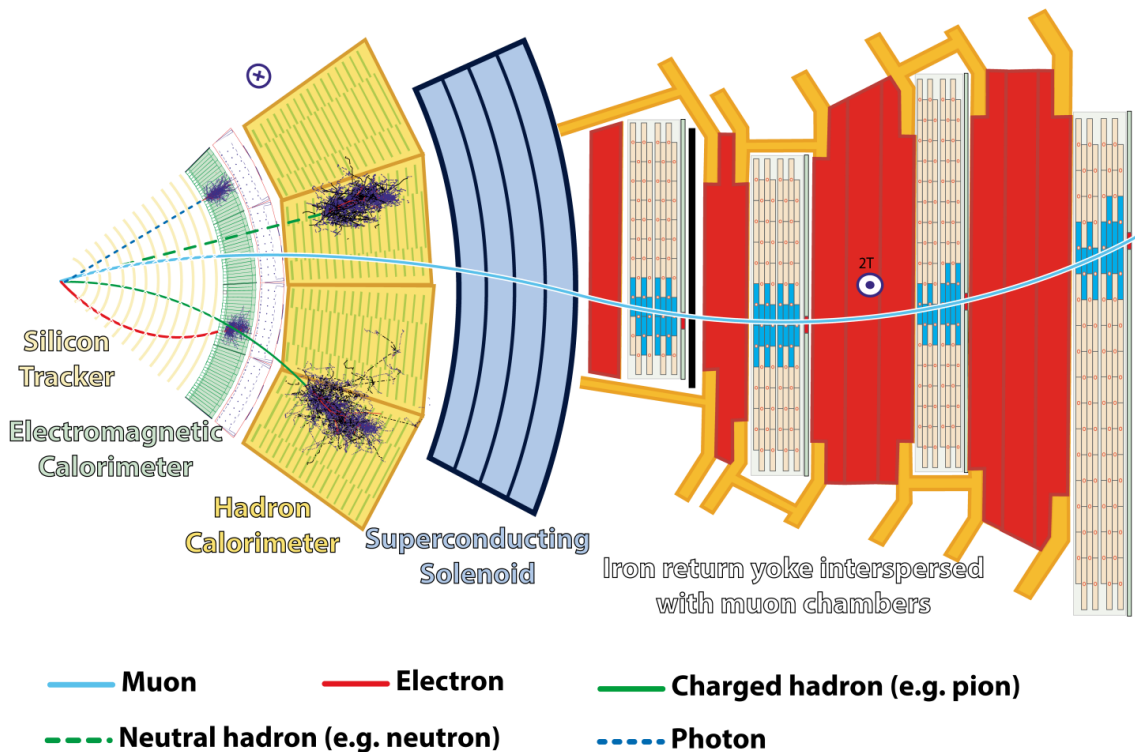


Abbildung 2.2: Querschnitt eines Teils des CMS-Detektors. Die dargestellten Komponenten entsprechen von links nach rechts den Schichten des Detektors von innen nach außen: 1. Silizium-Spurdetektor, 2. Elektromagnetisches Kalorimeter, 3. Hadronisches Kalorimeter, 4. Supraleitender Solenoid, 5. Eisernes Rückführjoch mit Myonenkammern. Entnommen aus [16].

2.4 Ereignisrekonstruktion und -generation

Bei der Rekonstruktion von Ereignissen wird zwischen Online- und Offline-Niveau unterschieden. Die Online-Rekonstruktion findet durch Trigger und unmittelbar nach einer Proton-Proton-Kollision statt und dient der Auswahl zu speichernder Ereignisse. Da der Fokus dieser Bachelorarbeit auf der Bestimmung von Trigger-Effizienzen liegt, wird dem CMS-Trigger-System in Kap. 4 ein eigenes Kapitel gewidmet und daher an dieser Stelle nicht weiter diskutiert. Die Rekonstruktion auf Offline- und Online-Niveau erfolgt mit Hilfe

des *Particle-Flow*-Algorithmus (PF-Algorithmus) [18]. Dabei werden Informationen der einzelnen Detektorteile genutzt, um einzelne Teilchen im Endzustand zu identifizieren.

Neben CMS-Datensätzen werden auch nach der Monte-Carlo-Methode simulierte Ereignisse verwendet: Die Fragmentation und Hadronisierung wird durch Pythia 8 simuliert [19]. Mit Hilfe des Matrixelement-Generators MadGraph5aMC@NLO [20] erfolgt die Generation von Ereignissen. Durch Geant4 [21] wird die Detektorantwort bei Durchdringen von Teilchen simuliert.

3. Suche nach Dunkler Materie in Monotop-Prozessen mit CMS

Eine Form von Materie, die nicht mit dem Standardmodell der Teilchenphysik erklärt werden kann, ist die Dunkle Materie (DM). Der erste Teil (Kap. 3.1) dieses Kapitels thematisiert die Dunkle Materie im Allgemeinen: Die wichtigsten Hinweise für ihre Existenz werden in Kap. 3.1.1 diskutiert. Im Zusammenhang mit Messungen der kosmischen Hintergrundstrahlung wird gezeigt, dass es im Standardmodell der Kosmologie eine nicht-baryonische Form der Dunklen Materie geben muss. Im Anschluss daran folgt in Kap. 3.1.2 eine kurze Auflistung der populärsten Kandidaten für die nicht-baryonischen Elementarteilchen der Dunklen Materie.

Der zweite Teil (Kap. 3.2) befasst sich mit der Monotop-Analyse - einem möglichen Ansatz, Dunkle Materie potentiell nachzuweisen. Er basiert auf dem Modell, dass Dunkle Materie in einem Monotop-Prozess produziert wird. Das Modell und die Analysestrategie zum hadronischen Monotop-Prozess werden in Kap. 3.2.1 diskutiert. Danach werden in Kap. 3.2.2 die wichtigsten Untergrundprozesse erklärt und für deren Untersuchungen Kontrollregionen definiert. Im Kontext des Monotop-Prozesses werden mit der fehlenden Transversalenergie und dem Hadronischen Rückstoß zwei wichtige Messgrößen bei der Monotop-Analyse eingeführt. Zuletzt wird in Kap. 3.2.3 eine Übersicht der Ereignisselektionen, die durch die Monotop-Analyse motiviert ist, für alle Studien in Kap. 5 dargestellt.

3.1 Dunkle Materie

Der Begriff der Dunklen Materie drückt ihre Eigenschaft aus, nicht elektromagnetisch zu wechselwirken. Auch eine starke Wechselwirkung der Dunklen Materie kann ausgeschlossen werden. Der folgende Abschnitt beschäftigt sich in strenger Anlehnung an [22] mit den Evidenzen und einigen nicht-baryonischen Elementarteilchen-Kandidaten für Dunkle Materie.

3.1.1 Evidenz

Historisch findet die Theorie einer Dunklen Materie ihren Ursprung in astronomischen Beobachtungen. Die Beobachtungen vereint die Tatsache, dass in den betrachteten Galaxien und Galaxienhaufen Widersprüche auftreten, für die es im Wesentlichen zwei unterschiedliche Erklärungen gibt: Entweder es existiert mit der Dunklen Materie eine Form von Materie, die bisher nicht beobachtet worden ist oder das heutige Verständnis der Gravitation ist nicht vollkommen richtig. Es werden nun beispielhaft einige der wichtigsten Indizien für die Existenz von Dunkler Materie angeführt.

Bereits 1933 ist nach Messungen der Rotverschiebungen im Coma-Galaxienhaufen abgeschätzt worden, dass unter der Annahme einer homogenen Massendichte des Galaxienhaufens

diese Dichte mindestens 400 mal größer sein müsste als die mittlere Dichte der leuchtenden Materie [23].

Ein noch bedeutenderes Indiz zeigt sich allerdings in den Rotationskurven von Galaxien. Die Rotationskurve einer Spiralgalaxie beschreibt die Rotationsgeschwindigkeit von Sternen in Abhängigkeit vom Abstand zum Zentrum der Galaxie. Rotationskurven lassen sich bei bekannter Massendichte der Galaxie über die Gravitationskraft, die als Zentripetalkraft wirken muss, ableiten. Wird als Massendichte nur die Masse der beobachtbaren Materie hinzugezogen, so stimmt die resultierende theoretische Rotationskurve nicht mit der beobachteten Rotationskurve überein. Ein Beispiel für eine solche Diskrepanz zeigt sich bei der Spiralgalaxie Messier 33 in Abb. 3.1.

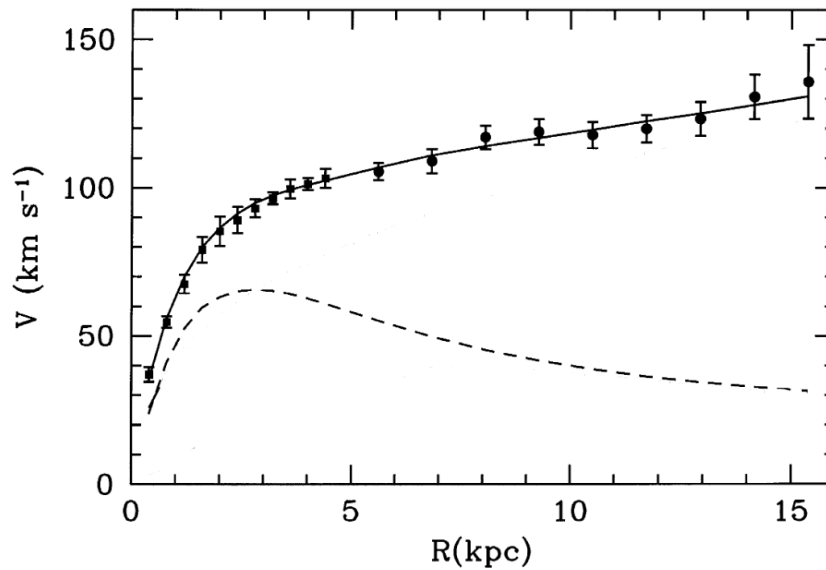


Abbildung 3.1: Vergleich der theoretischen Rotationskurve für die beobachtbare Sternscheibe von Messier 33 (gestrichelte Linie) mit der gemessenen Rotationskurve (Datenpunkte inklusive dem besten Fit für ein Modell aus der beobachtbaren Sternscheibe, einer Gaswolke und einem Dunkle-Materie-Halo). Entnommen und bearbeitet aus [24].

Sowohl in Galaxien als auch in Galaxienhaufen zeigen sich darüber hinaus auch Anzeichen von Dunkler Materie in Form des Gravitationslinseneffekts¹. Es wird dazu die Verzerrung des Bilds eines hell leuchtenden Objekts wie etwa einer Galaxie untersucht, deren emittierte elektromagnetische Strahlung relativ knapp eine Gravitationslinse passiert, bevor sie ein Teleskop erreicht. Anhand der Stärke der Verzerrung lässt sich dann die Masse der Gravitationslinse bestimmen und mit der Masse aller direkt beobachtbaren Komponenten vergleichen. Im Fall vom Galaxienhaufen CL0024+1654 ergibt sich so innerhalb eines festgelegten Radius ein Massenanteil der Dunklen Materie von 83 % der Gesamtmasse des Galaxienhaufens [25], um nur ein Beispiel zu nennen.

Bis hierhin wurden nur Hinweise für ein lokales Vorhandensein von Dunkler Materie diskutiert. Um jedoch Vermutungen über den Anteil der Dunklen Materie im gesamten Universum formulieren zu können, wird die Anisotropie der kosmischen Mikrowellenhintergrundstrahlung untersucht. Es folgt eine kurzer Einschub zur Entstehung der Hintergrundstrahlung als eine Zusammenfassung aus [26]: Die Hintergrundstrahlung stammt aus der Epoche

¹Unter dem Gravitationslinseneffekt versteht man die Ablenkung von elektromagnetischer Strahlung nahe massereichen Objekten aufgrund der Raumkrümmung durch die Gravitation nach der allgemeinen Relativitätstheorie. Das massereiche Objekt wird in diesem Kontext auch Gravitationslinse genannt.

„Photon-Entkopplung“ ca. 350.000 Jahre nach dem Urknall. Zuvor lag ein thermisches Gleichgewicht aus freien Protonen, freien Elektronen und Photonen vor. Die mittlere freie Weglänge der Photonen war aufgrund von Streuung an den freien Elektronen vergleichsweise gering, sodass die elektromagnetische Strahlung nicht frei im Universum propagieren konnte. Erst durch die Abkühlung der Teilchen infolge der Expansion des Universums rekombinierten Elektronen und Protonen zu Wasserstoffatomen, wodurch der Wirkungsquerschnitt für Photonstreuung abnahm. Damit fand in der Epoche „Photon-Entkopplung“ eine thermische Entkopplung der Photonen von den Protonen und Elektronen statt. Gleichzeitig führte die Abnahme des Wirkungsquerschnitts für Photonstreuung auch zu einer Zunahme der mittleren freien Weglänge der Photonen, wodurch das Universum transparent für elektromagnetische Strahlung wurde. Die kosmische Mikrowellenhintergrundstrahlung ist demnach jene elektromagnetische Strahlung, die während der Rekombinationsphase das letzte Mal gestreut wurde.

Bis auf Schwankungen in der Größenordnung von 0,001 % ist die Strahlung isotrop und ihr Spektrum entspricht dem eines schwarzen Strahlers mit der Temperatur 2,726 K. Eine Anisotropie der Strahlung kann einen weiteren Hinweis für Dunkle Materie liefern. Zur genaueren Untersuchung werden die gemessenen Temperaturschwankungen $\frac{\delta T}{T}(\theta, \phi)$ nach den Kugelflächenfunktionen $Y_{lm}(\theta, \phi)$ mit den Multipolen l entwickelt:

$$\frac{\delta T}{T}(\theta, \phi) = \sum_{l=2}^{\infty} \sum_{m=-l}^l a_{lm} Y_{lm}(\theta, \phi). \quad (3.1)$$

In einem Leistungsspektrum wird dann $\mathcal{D}_l = l(l+1)C_l/(2\pi)$ als Maß für die Temperaturschwankung in Abhängigkeit vom Multipolindex l aufgetragen, wobei C_l die Varianzen der Entwicklungskoeffizienten a_{lm} angibt. Ein Leistungsspektrum, das sich aus Daten des Planck-Weltraumteleskops bis 2018 ergibt, ist in Abb. 3.2 dargestellt. Durch eine Anpassung des Spektrums an das Λ CDM-Modell (Standardmodell der Kosmologie) lässt sich so die Dichte von Baryonen und Dunkler Materie im Universum bestimmen. Mitunter kann aus dem zweiten Maximum auf die baryonische Materie und aus dem dritten Maximum auf die Dunkle Materie geschlossen werden [28]. Das Ergebnis für die Planck-Daten von 2018 ist eine Dichte der Dunklen Materie von $\Omega_c h^2 = 0,1200 \pm 0,0012$ und eine Baryondichte von $\Omega_b h^2 = 0,02237 \pm 0,00015$ [27]. Die angegebenen Dichten sind hierbei proportional zum relativen Anteil der Dichten an der gesamten Energiedichte im Universum.

3.1.2 Kandidaten für Dunkle Materie

Wegen der unterschiedlichen Dichten von Baryonen und Dunkler Materie im Universum (s. Kap. 3.1.1) muss implizit auch eine nicht-baryonische Form von Dunkler Materie existieren, was der Anlass für einige Theorien mit teilweise exotischen Kandidaten für die Elementarteilchen der Dunklen Materie ist. In unterschiedlichen Theorien stimmen die vorausgesagten Dichten der DM-Teilchenarten teils gut und teils weniger gut mit der kosmologisch berechneten Dichte Dunkler Materie überein. Darüber hinaus ist nicht bekannt, aus wie vielen verschiedenen Teilchenarten sich die nicht-baryonische Dunkle Materie zusammensetzt. Die einzige bekannte Tatsache ist, dass die Teilchen lediglich der gravitativen Wechselwirkung unterliegen, wobei in manchen Theorien auch eine schwache Wechselwirkung angenommen wird. Für eine kurze Übersicht, die keinen Anspruch auf Vollständigkeit erhebt, wird nun basierend auf [29, 22] eine Auswahl an Kandidaten aufgelistet:

- WIMPs (*weakly interacting massive particles*) wechselwirken, wie der Name suggeriert, nicht nur gravitativ sondern auch schwach. Ihre Masse kann in einem Bereich zwischen 10 GeV und einigen TeV liegen. WIMPs können in hochenergetischen Teilchenkollisionen von Standardmodell-Teilchen zu frühen Zeiten im Universum oder in Teilchenbeschleunigern produziert werden und durch Annihilation zweier zueinander

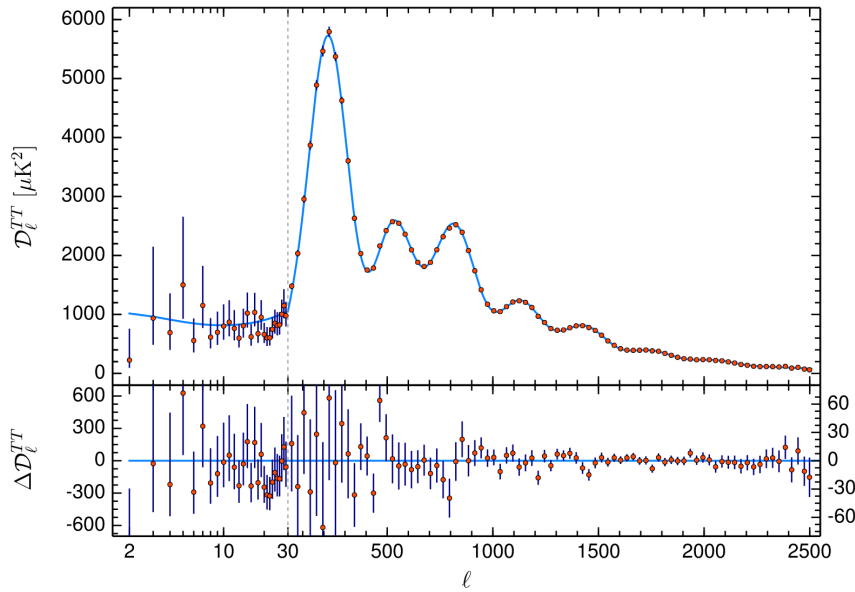


Abbildung 3.2: Leistungsspektrum aus Daten des Planck-Weltraumteleskops von 2018 inklusive Residuen. Dargestellt ist $\mathcal{D}_l = l(l+1)C_l/(2\pi)$ in Abhängigkeit der Multipolordnung l . Entnommen aus [27].

konjugierter WIMPs vernichtet werden. In der Theorie befanden sich WIMPs früh im Universum im thermischen Gleichgewicht mit den übrigen Teilchen. Zu diesem frühen Zeitpunkt herrschen so hohe Temperaturen, dass sich die WIMPs zudem im chemischen Gleichgewicht mit den Teilchen, die für ihre Produktion verantwortlich sind, befinden. Durch die Abkühlung des Universums in Folge seiner Expansion wird einerseits die Produktion der WIMPs unterdrückt. Andererseits nimmt die Wahrscheinlichkeit für WIMP-Annihilation mit der Expansion des Universums ab und die Dichte der WIMPs konvergiert gegen einen konstanten Wert, der Reliktdichte genannt wird. Dieser Prozess wird meist als *freeze-out* bezeichnet. Beispiele für WIMPs sind Neutralinos und Kaluza-Klein-Teilchen. Neutralinos kommen in supersymmetrischen Theorien² als Masseneigenzustände von Mischungen der Superpartner neutraler Bosonen vor. Die Bosonen sind B^0 , W^0 sowie verschiedene in supersymmetrischen Theorien benötigte neutrale Higgs-Bosonen. Kaluza-Klein-Teilchen sind das Resultat einer Theorie mit zusätzlichen Raumdimensionen.

- SuperWIMPs wechselwirken noch schwächer als WIMPs. In der dazugehörigen Theorie werden SuperWIMPs durch den Zerfall von WIMPs spät nach deren *freeze-out* gebildet. Beispiele für SuperWIMPs sind Gravitinos - die Superpartner des Gravitons, das ein hypothetisches Eichboson der Gravitation ist.
- Leichte Gravitinos sind analog zu den SuperWIMP-Gravitinos die Superpartner des Gravitons. Der Unterschied besteht darin, dass leichte Gravitinos maximal eine Masse in der Größenordnung von einigen keV besitzen können.
- Sterile Neutrinos sind Masseneigenzustände überwiegend rechtshändiger Neutrinos und wechselwirken daher bis auf die Neutrinomischung nicht schwach. Im Gegensatz dazu sind Standardmodell-Neutrinos Masseneigenzustände überwiegend linkshändiger Neutrinos.
- Axionen sind leichte und sehr schwach wechselwirkende Teilchen, die auch mit Gluonen

²Die Supersymmetrie ist ein Überbegriff für Theorien, in denen jedes Boson bzw. Fermion in ein dazugehöriges bisher unbeobachtetes Fermion bzw. Boson transformiert werden kann. Zwei Teilchen, die sich im Rahmen der Supersymmetrie ineinander transformieren können, werden Superpartner genannt.

und Photonen wechselwirken. Darüber hinaus stellen sie einen Ansatz zur Erklärung der CP-Verletzung in der QCD dar [30].

3.2 Hadronische Monotop-Analyse

Während Evidenzen für die Existenz von Dunkler Materie auf der astronomischen und kosmologischen Skala gegeben sind, erfolgt die Suche danach oft auf der Skala der kleinsten Teilchen. An Teilchendetektoren wie dem CMS finden solche Suchen nach den Elementarteilchen der Dunklen Materie statt. In einem Monotop-Prozess wird die gemeinsame Produktion von Dunkler Materie und einem einzelnen Top-Quark angenommen, wobei das Top-Quark hadronisch oder leptonisch zerfallen kann. Basierend auf [31, 3] wird nun auf die Theorie des Monotop-Prozesses und auf die wichtigsten Untergrundprozesse eingegangen. Ausgehend von der Theorie ist in [31] eine Analysestrategie abgeleitet worden, deren Grundzüge hier beschrieben werden.

3.2.1 Monotop-Prozess

Der Monotop-Prozess beschreibt die Produktion von DM-Teilchen und einem einzelnen Top-Quark und kann durch resonante oder nicht-resonante Modelle erklärt werden. Feynman-Diagramme für ein nicht-resonantes und ein resonantes Modell sind in Abb. 3.3 dargestellt. Die Analyse, die den Studien in Kap. 5 zugrunde liegt, beschränkt sich auf das vereinfachte nicht-resonante Modell. Im Wesentlichen wird dabei angenommen, dass sich ein Up-Quark durch Emission eines neutralen Spin-1-Bosons V in ein Top-Quark transformiert. Das Boson zerfällt anschließend in zwei zueinander konjugierte DM-Dirac-Fermionen. Da V elektrisch neutral ist und sich der Flavour des Up-Quarks bei der Transformation zum Top-Quark ändert, repräsentiert V einen flavourverändernden neutralen Strom und ist damit im Standardmodell stark unterdrückt (s. Kap. 2.1).

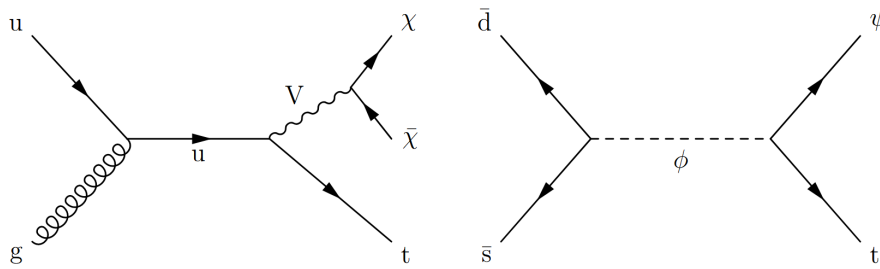


Abbildung 3.3: Zwei Modelle zum Monotop-Prozess: Links wird im nicht-resonanten Modell durch einen FCNC ein Top-Quark und ein Vektorboson V produziert. Rechts wird im resonanten Modell das Top-Quark durch den Zerfall einer skalaren Resonanz ϕ (Spin 0) erzeugt. Entnommen aus [3].

Die entstehenden DM-Teilchen wechselwirken nur gravitativ und womöglich schwach, wie in Kap. 3.1.2 beschrieben ist, weshalb sie nicht direkt vom CMS-Detektor detektiert werden können. Die grundlegende Idee, um dennoch potentiell Dunkle Materie nachzuweisen, ist, dass die am LHC beschleunigten Protonen und deren Konstituenten sehr hohe Impulse entlang der Strahlachse, aber vernachlässigbar kleine Transversalimpulse senkrecht dazu besitzen. Nach dem Impulserhaltungssatz muss die Vektorsumme der Transversalimpulse aller Produkte einer Proton-Proton-Kollision damit nahezu Null sein. Bei der Berechnung der Impulssumme können allerdings nur rekonstruierte Teilchen berücksichtigt werden. Das bedeutet auch, dass im Fall einer wesentlich von Null verschiedenen Impulssumme unerkannte Teilchen den Detektor passiert haben müssen. Die Impulssumme der undetektierten Teilchen wird auch fehlender Transversalimpuls bzw. fehlende Transversalenergie

(MET) \vec{E}_T genannt und ergibt sich aus der Impulssumme $\sum_i \vec{p}_{T,i}$ aller rekonstruierten Teilchen durch eine Rotation von 180° in der Transversalebene zu

$$\vec{E}_T = - \sum_i \vec{p}_{T,i}. \quad (3.2)$$

Fehlende Transversalenergie wird häufig durch Neutrinos verursacht. Vor allem größere fehlende Transversalenergien können allerdings hypothetisch auch durch massereiche DM-Teilchen wie WIMPs hervorgerufen werden. Daher ist vor allem eine große fehlende Transversalenergie interessant bei der Suche nach Dunkler Materie und wird bei der Ereignisselektion der Analyse berücksichtigt.

Im vereinfachten nicht-resonanten Modell entsteht neben den DM-Teilchen, die für die große fehlende Transversalenergie verantwortlich sind, auch ein Top-Quark, das mit einem Verzweigungsverhältnis von $(66,5 \pm 1,4) \%$ hadronisch zerfällt [11] gemäß

$$t \rightarrow q\bar{q}b. \quad (3.3)$$

Die drei Quarks bilden durch iterative Hadronisierung Teilchenschauer, die als Jets zusammengefasst werden können, wie in Kap. 2.1 beschrieben ist. Da das Top-Quark aufgrund des Impulserhaltungssatzes den Rückstoß des massereichen Eichbosons V kompensieren muss, erhalten die drei Zerfallsquarks einen Boost in dieselbe Richtung. Sie können daher als ein Jet mit vergrößertem Radius rekonstruiert werden. Die Rekonstruktion von Jets erfolgt mit Hilfe des Anti- k_T -Algorithmus [32]. Für die Selektion wird ein durch den Anti- k_T -Algorithmus rekonstruierter Jet mit einem Radiusparameter ΔR von 1,5 in der η - ϕ -Ebene (AK15 Jet) gefordert mit

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

Dabei gibt η die Pseudorapidity und ϕ den Azimutwinkel an. Es gilt

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right), \quad (3.5)$$

mit dem Polarwinkel θ .

3.2.2 Untergrundprozesse

Monotop-Prozesse sind im Standardmodell stark unterdrückt. Im nicht-resonanten Modell kann Dunkle Materie jedoch über einen FCNC auf *tree-level*³ erzeugt werden, wodurch der Wirkungsquerschnitt im nicht-resonanten Modell höher als im Standardmodell ist. Das potentielle Signal ist praktisch erst dadurch messbar, wenn auch eine relativ schwache Signalstärke zu erwarten ist. Folglich werden bei der Forderung nach einem Mindestwert an fehlender Transversalenergie und einem AK15-Jet vom CMS hauptsächlich Untergrundprozesse registriert. Die fehlende Transversalenergie bei Untergrundprozessen stammt von undetektierten Neutrinos und der AK15-Jet von Quarks, die Teilchenschauer bilden. Zur Einschränkung der Untergrundprozesse wird eine Signalregion definiert, in der keine Elektronen, Myonen und Photonen vorhanden sein sollen. Die wichtigsten Untergrundprozesse sind Z+Jets, W+Jets und $t\bar{t}$:

- Z+Jets ist ein Prozess, bei dem ein Z-Boson und zusätzliche Jets gebildet werden. Das Z-Boson kann weiter in zwei geladene Leptonen, zwei Neutrinos oder zwei Quarks zerfallen. Der Zerfall in Neutrinos resultiert in fehlender Transversalenergie. Ein

³Feynman-Diagramme ohne geschlossene Schleifen (*Loops*) werden auch als *Tree-Level*-Feynman-Diagramme bezeichnet.

dazugehöriges Feynman-Diagramm ist in Abb. 3.4 zu sehen. Die zusätzlichen Jets im Z+Jets-Prozess können auch von anderen Quarks als dem Top-Quark verursacht und unerwünscht als AK15-Jet rekonstruiert werden. Eine Methode, um hadronisch zerfallende Top-Quarks mit Boost zu identifizieren ist das sogenannte *Top-Tagging* [33]. In dieser Arbeit werden allerdings keine *Top-Tagger* verwendet.

- W+Jets beschreibt die Produktion eines W-Bosons und zusätzlicher Jets, wobei das W-Boson leptonisch zerfällt. Ein Feynman-Diagramm zum Prozess ist in Abb. 3.4 dargestellt. Beim leptonischen W-Zerfall entsteht ein Neutrino, das eine fehlende Transversalenergie verursacht und ein geladenes Lepton. Trotz der Forderung, dass keine geladenen Leptonen in der Signalregion vorhanden sein dürfen, trägt W+Jets zum Untergrund in dieser Region bei, da Elektronen und Myonen gelegentlich nicht als solche rekonstruiert werden. Wie beim Prozess Z+Jets können Quarks auch hier AK15-Jets verursachen und damit nicht immer von Top-Quarks mit Boost unterschieden werden.
- $t\bar{t}$ -Produktion gefolgt von einem semileptonischen Zerfall ist ein weiterer wichtiger Untergrund, da ein hadronisch zerfallendes Top-Quark bei ausreichendem Rückstoß zu einem AK15-Jet führen kann. Den Rückstoß liefert das konjugierte Top-Quark, das leptonisch zerfällt. Dabei entsteht ein Neutrino, das zu einer fehlenden Transversalenergie führt. Abb. 3.5 zeigt ein Feynman-Diagramm zur Top-Quark-Antiquark-Produktion durch Wechselwirkung zweier Gluonen.

Für die Monotop-Analyse und die Suche nach dem Monotop-Prozess als Signalprozess sind auch genauere Untersuchungen der Untergrundprozesse notwendig. Im Rahmen der Studien zu MET-Trigger-Effizienzen in Kap. 5 und der Bestimmung von Daten-Simulation-Skalenfaktoren für diese Effizienzen in Kap. 6 liegt das Hauptaugenmerk auf den Untergrundprozessen W+Jets und Z+Jets. Der Zerfall des Z-Bosons in zwei Neutrinos besitzt allerdings keine eindeutige Signatur, weshalb stattdessen der Z-Zerfall in zwei geladene Leptonen untersucht wird. Um zur Untersuchung eines bestimmten Prozesses geeignete Ereignisse anzureichern, werden Kontrollregionen (in den Abbildungen in Kap. 5 und Kap. 6 mit CR bezeichnet) definiert. Für W+Jets wird die SingleMuon-Kontrollregion und für Z+Jets die DiMuon-Kontrollregion eingeführt. Genaue Selektionskriterien zu den Kontrollregionen können der Übersicht in Kap. 3.2.3 entnommen werden.

Um den Transversalimpuls der Eichbosonen für verschiedene Signal- und Kontrollregionen mit einer einheitlichen und damit vergleichbaren Methode zu bestimmen, wird der Transversalimpuls des hadronischen Rückstoßes $\vec{\cancel{U}}_T$ rekonstruiert. Er ist neben der fehlenden Transversalenergie \cancel{E}_T eine weitere wichtige Größe bei der Monotop-Analyse und unterscheidet sich nur durch Addition der Impulse aller Elektronen, Myonen und Photonen:

$$\vec{\cancel{U}}_T = \vec{\cancel{E}}_T + \sum_{i=e,\mu,\gamma} \vec{p}_{T,i}. \quad (3.6)$$

Dadurch kann beispielsweise für den Vergleich des Transversalimpulses eines Z-Bosons in der Signalregion und der DiMuon-Kontrollregion in beiden Fällen der Hadronische Rückstoß bestimmt werden.

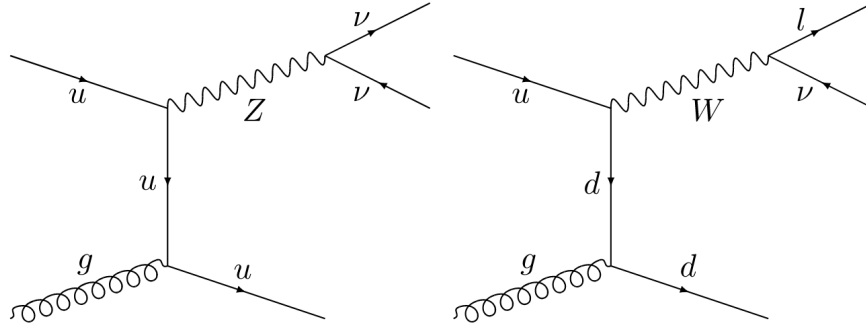


Abbildung 3.4: Feynman-Diagramme zweier Untergrundprozesse bei der Monotop-Analyse. Links wird ein Z-Boson und zusätzliche Jets produziert, wobei das Z-Boson in zwei Neutrinos zerfällt. Rechts wird ein W-Boson und zusätzliche Jets erzeugt, wobei das W-Boson in ein Lepton und ein Neutrino zerfällt. Entnommen aus [31].

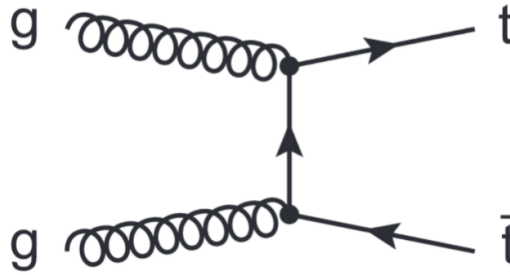


Abbildung 3.5: Feynman-Diagramm zum Untergrundprozess $t\bar{t}$ bei der Monotop-Analyse. Ein Top-Quark-Antiquark-Paar entsteht durch Wechselwirkung zweier Gluonen. Entnommen aus [31].

3.2.3 Ereignisselektion

Es wird nun ausgehend von [31] auf die Ereignisselektion eingegangen, die bei den Studien zu MET-Trigger-Effizienzen in Kap. 5 und bei der Bestimmung von Skalenfaktoren in Kap. 6 Anwendung finden. Hierbei wird zwischen den Ären⁴ 2016preVFP, 2016postVFP, 2017 und 2018 unterschieden. Für die Effizienz-Studien werden zwei verschiedene Grundselektionen verwendet, die der Übersicht in Tabelle 3.1 zu entnehmen sind. Die erste Selektion liegt nah an der Analysestrategie und wird für einen Vergleich verschiedener Monte-Carlo-Simulationen in der SingleMuon-Kontrollregion eingesetzt (Kap. 5.1). Die zweite Selektion ist dadurch motiviert, dass die Daten-Simulation-Skalenfaktoren von MET-Trigger-Effizienzen zu einem möglichst geringen hadronischen Rückstoß zu bestimmen sind. Deshalb wird die zweite Selektion für die übrigen Studien in Kap. 5 und die Bestimmung der Skalenfaktoren in Kap. 6 benutzt. Es folgt eine kurze Beschreibung der einzelnen Selektionsschritte. Eine detailliertere Schilderung ist [31] zu entnehmen.

- In der ersten Selektion ist ein hadronischer Rückstoß von mindestens 250 GeV gefordert. In der zweiten Selektion wird der Schnitt auf 70 GeV reduziert.
- Es soll mindestens ein AK15-Jet rekonstruiert werden. Diese Bedingung wird für die zweite Selektion fallengelassen, da AK15-Jets meist zu einem deutlich höheren hadronischen Rückstoß als 70 GeV führen. Damit die Verteilung des hadronischen Rückstoßes möglichst wenig durch die Selektionen der AK-Jets eingeschränkt wird, werden in der zweiten Selektion darüber hinaus alle weiteren Selektionsschritte,

⁴Eine Ära stellt einen längeren Zeitabschnitt der Datenentnahme dar. Die verschiedenen Ären spiegeln geänderte Detektorbedingungen wieder.

die AK-Jets enthalten, ignoriert. Diese Schritte betreffen die Winkelverteilungen zwischen AK-Jets und dem hadronischen Rückstoß sowie die Rekonstruktion von *Soft-Drop*-Jets [34].

- Der letzte Selektionsschritt betrifft nur die Ära 2018. Aufgrund einer defekten HCAL-Einheit während der Messung werden alle Ereignisse mit Jets innerhalb von $-3,0 \leq \eta \leq -1,3$ und $-1,57 \leq \phi \leq -0,87$ verworfen [31]. Der Defekt und seine Konsequenzen werden auch als HEM-Problem bezeichnet.

Übersichten zu den Selektionskriterien in der SingleMuon- und DiMuon-Kontrollregion sind in Tabellen 3.2 und 3.3 zu sehen. In der SingleMuon-Kontrollregion wird genau ein Myon, aber keine Elektronen und Photonen gefordert. Gleichzeitig soll die fehlende Transversalenergie mindestens 100 GeV betragen. In der ersten Selektion sollen zudem neben dem führenden AK15-Jet keine weiteren Jets rekonstruiert werden.

Dagegen sollen sich in der DiMuon-Kontrollregion genau zwei Myonen aber keine Elektronen und Photonen befinden. Mindestens eines davon soll einen Transversalimpuls von mehr als 30 GeV besitzen. Der Grund dafür ergibt sich in Kap. 4.3, denn es werden Referenztrigger mit einer Schwelle von 24 / 27 GeV eingesetzt und es soll gewährleistet werden, dass nur im Plateau des Referenztriggers selektiert wird. Die fehlende Transversalenergie soll außerdem 120 GeV übersteigen und die Schwerpunktsenergie der zwei Myonen soll zu einem Wert zwischen 60 GeV und 120 GeV bestimmt werden, da die Masse des Z-Bosons in diesem Bereich liegt.

Tabelle 3.1: Eine Übersicht beider Grundselektionen unter Verwendung der in der Analysesoftware verwendeten Variablennamen zur Bezeichnung von Größen. Der hadronische Rückstoß ist in GeV und die Winkel ϕ in Bogenmaß gegeben.

Ära	Selektionsschritt	Selektion 1	Selektion 2
Alle	Hadr_Recoil_Pt_nom \geq 250	✓	
Alle	Hadr_Recoil_Pt_nom \geq 70		✓
Alle	N_AK15Jets_nom \geq 1	✓	
Alle	N_Jets_nom \geq 1	✓	✓
Alle	DeltaPhi_AK15Jet_Hadr_Recoil_nom $>$ 1.5	✓	
Alle	All(DeltaPhi_AK4Jet_Hadr_Recoil_nom $>$ 0.5)	✓	
Alle	N_AK15SoftDropJets_nom == N_AK15Jets_nom	✓	
2018	N_HEMJets_nom = 0	✓	✓
2018	\neg (RawJet_Eta_nom $<$ -1.3 \wedge RawJet_Eta_nom $>$ -3.0 \wedge RawJet_Phi_nom $<$ -0.87 \wedge RawJet_Phi_nom $>$ -1.57)	✓	✓

Tabelle 3.2: Eine Übersicht der Selektion für die SingleMuon-Kontrollregion unter Verwendung der in der Analysesoftware verwendeten Variablennamen zur Bezeichnung von Größen. Massen M und Transversalimpulse P_T sind in GeV gegeben.

Selektionsschritt	Selektion 1	Selektion 2
N_LooseMuons == 1	✓	✓
N_TightMuons == 1	✓	✓
N_LooseElectrons == 0	✓	✓
N_LoosePhotons == 0	✓	✓
MET_T1_PT_nom $>$ 100	✓	✓
Sum(Jets_outside_lead_AK15Jet_taggedL_nom) == 0	✓	

Tabelle 3.3: Eine Übersicht der Selektion für die DiMuon-Kontrollregion unter Verwendung der in der Analysesoftware verwendeten Variablennamen zur Bezeichnung von Größen. Massen M und Transversalimpulse P_T sind in GeV gegeben.

Selektionsschritt	Selektion 2
N_LooseMuons == 2	✓
N_LooseMuons_Pt[0] \geq 30	✓
N_LooseElectrons == 0	✓
N_LoosePhotons == 0	✓
DiMuon_M $>$ 60	✓
DiMuon_M $<$ 120	✓
MET_T1_PT_nom $<$ 120	✓

4. CMS-Trigger und ihre Effizienzen

Wie in Kap. 2.2 beschrieben ist, kollidieren am LHC Bündel an Protonen mit einer Rate von 40 MHz. Aufgrund von Einschränkungen bezüglich der vorhandenen Datenspeicherkapazität und Datentransferrate können jedoch nicht alle Ereignisse aufgezeichnet werden und selbst wenn diese Begrenzung nicht vorhanden wäre, sind für eine bestimmte Analyse nicht alle auftretenden Ereignisse interessant. Daher wird bei der Datenaufnahme von Triggern Gebrauch gemacht. In Kap. 4.1 wird die Realisierung des Triggersystems beim CMS-Detektor nach [4, 35] zusammengefasst. Das Selektionsverhalten von Triggern lässt sich durch Trigger-Effizienzen charakterisieren, die in Kap. 4.2 diskutiert werden. Es wird außerdem eine mögliche Methode zur Bestimmung von Trigger-Effizienz-Kurven unter der Verwendung eines Referenztriggers aufgezeigt. Anschließend wird mit den Daten-Simulation-Skalenfaktoren die Lösung für ein Problem bei der Verwendung von Triggern für Analysen mit Simulationsdaten präsentiert. Da die Monotop-Analyse sowohl MET-Trigger als auch Simulationsdaten nutzt, ist es notwendig solche Skalenfaktoren zu bestimmen (s. Kap. 6). Zuletzt enthält Kap. 4.3 eine Übersicht der MET-Trigger, die bei der Monotop-Analyse eingesetzt werden (Probentrigger) und der Referenztrigger, mit denen die Effizienz-Kurven der Probentrigger bestimmt werden.

4.1 Das CMS-Triggersystem

Das CMS-Triggersystem dient der Auswahl, welche Ereignisse interessant für eine spezifische Analyse sind und daher gespeichert werden sollen. Der Aufbau ist zweischichtig in den L1-Trigger und den HLT einteilbar.

Der L1-Trigger (*Level-1 Trigger*) stellt die erste Schicht dar. Er besteht aus einer speziell für diesen Zweck gefertigten Hardware. Informationen aus den elektromagnetischen und hadronischen Kalorimetern sowie den Myonenspurdetektoren werden auf Konsistenz mit den Signaturen einiger zu konstruierenden Objekte geprüft. Zu den Objekten gehören Elektronen, Photonen, Myonen, Jets, Energiesummen und Tauonen. Innerhalb von 4 μ s nach Eintreten eines Ereignisses wird auf diese Weise entschieden, ob es von Relevanz ist oder nicht. Es wird damit signifikant weniger Zeit in Anspruch genommen, als beispielsweise für eine Offline-Rekonstruktion benötigt wird. Irrelevante Ereignisse werden verworfen und relevante Ereignisse mit einer Rate von maximal 100 kHz an die zweite Schicht des Triggersystems geleitet.

Der HLT (*High Level Trigger*) repräsentiert die zweite Schicht des CMS-Triggersystems. Sie umfasst Software, die auf Computern mit insgesamt 13000 Single-Core-Prozessoren läuft. Beim HLT werden die Objekte mit derselben Software wie auf dem Offline-Level mit Hilfe von Algorithmen rekonstruiert. In der Software sind mehrere sogenannte HLT-Pfade (*HLT paths*) implementiert, die in einer festen Reihenfolge Algorithmen zur Objektrekonstruktion und anschließend Algorithmen zur Objektselektion anwenden. Um die Zahl an benötigten

Prozessoren zu minimieren, werden relevante Ereignisse nur teilweise rekonstruiert. Dies geschieht, indem nur Informationen über diejenigen Detektor-Regionen hinzugezogen werden, in denen eine gemessene Signatur zu einer Akzeptanz durch den L1-Trigger geführt hat. Die Rate an ausgewählten Ereignissen zur Speicherung und Offline-Rekonstruktion liegt bei bis zu 1 kHz.

Bei den in Kap. 5 untersuchten Triggern handelt es sich um verschiedene HLT-MET-Trigger mit konstanten Schwellen. Als Beispiel akzeptiert ein HLT-MET-Trigger mit einer Schwelle von 120 GeV nur Ereignisse, bei denen vom zugehörigen HLT-Algorithmus eine fehlende Transversalenergie von mindestens 120 GeV berechnet wird.

4.2 Trigger-Effizienzen

Sowohl beim HLT als auch auf Offline-Level werden mittels Algorithmen Objekte wie Myonen, Jets, Energiesummen etc. rekonstruiert und dabei deren Eigenschaften berechnet, wie bereits in Kap. 4.1 erwähnt worden ist. Der wesentliche Unterschied besteht darin, dass beim HLT nur ausgewählte Detektorregionen zur Rekonstruktion von Objekten verwendet werden - auf Offline-Level jedoch alle Detektorregionen. Eine Konsequenz ist, dass der Wert einer vom HLT berechneten Selektionsgröße¹ im Allgemeinen nicht mit dem Wert, wie er aus der Offline-Rekonstruktion resultiert, übereinstimmt. Hierbei ist die Offline-Rekonstruktion die bessere Beschreibung des realen Ereignisses. Folglich werden manche interessanten Ereignisse verworfen und andere uninteressante Ereignisse vom Trigger akzeptiert, falls der reale Wert der Selektionsgröße nah an der Schwelle liegt. Um in diesem Zusammenhang quantitative Aussagen treffen zu können, wird die (Selektions-)Effizienz ϵ des Triggers betrachtet. Sie gibt an, wie hoch der Anteil akzeptierter Ereignisse an der Gesamtzahl an Ereignissen ist. Üblicherweise wird die Trigger-Effizienz als Funktion der offline-rekonstruierten Selektionsgröße untersucht. Der Grund dafür ist, dass nur der funktionale Zusammenhang zwischen der Trigger-Effizienz und der HLT-rekonstruierten Selektionsgröße X wie erwünscht durch eine Heaviside-Funktion Θ beschrieben werden kann:

$$\epsilon = \Theta(X - X_S), \quad (4.1)$$

wobei X_S den Schwellenwert angibt, ab dem eine Akzeptanz des Ereignisses erfolgen soll. Hingegen ergibt sich bei Abhängigkeit der Trigger-Effizienz von der offline-rekonstruierten Selektionsgröße X' eine Verschmierung nahe der Schwelle. Das Ergebnis ist eine Trigger-Effizienz-Kurve. Häufig kann der funktionale Zusammenhang durch eine modifizierte Fehlerfunktion parametrisiert werden

$$\epsilon(X'; a_0, a_1, a_2) = \frac{a_2}{2} \cdot \left[\operatorname{erf}\left(\frac{1}{\sqrt{2}a_0}(X' - a_1)\right) + 1 \right]. \quad (4.2)$$

Die Parametrisierung beruht auf der Annahme, dass gaußförmige Verschmierung an der Schwelle stattfindet, denn die kumulative Verteilungsfunktion einer Gaußfunktion kann durch die Fehlerfunktion ausgedrückt werden. Der Parameter a_1 gibt hier die Trigger-Schwelle an.

Es wird nun die verallgemeinerte Methode beschrieben, mit der alle Trigger-Effizienz-Kurven in Kap. 5 bestimmt werden. Der Trigger, dessen Effizienz-Kurve es zu bestimmen gilt, wird als Probentrigger bezeichnet. Ausgehend von einer Basisselektion, die sich aus der zugrundeliegenden Analyse ergibt, wird durch Filtern mit einem Referenztrigger eine reduzierte Ereignismenge erzeugt. Diese reduzierte Menge ist vollständig bezüglich der Ereignisse, die

¹In diesem Abschnitt wird mit der Selektionsgröße die Größe bezeichnet, nach der bei Überschreitung eines bestimmten Triggerschwellenwerts ein Ereignis selektiert wird.

der Probentrigger prüft und wird daher im Weiteren totale Ereignismenge genannt. Der Sinn des Referenztriggers ist es, die Dichte an relevanten Ereignissen zu erhöhen. Dabei sollte darauf geachtet werden, dass der Referenztrigger möglichst unkorreliert zum Probentrigger ist. Anschließend kann der Probentrigger auf die totale Ereignismenge angewandt und daraus die akzeptierte Datenmenge generiert werden. Die Trigger-Effizienz kann dann durch Division der Anzahl an Elementen in der akzeptierten und totalen Ereignismenge berechnet werden

$$\epsilon = \frac{N_{\text{akzeptiert}}}{N_{\text{total}}}, \quad (4.3)$$

mit der Anzahl an Ereignissen N_{total} in der totalen Ereignismenge und $N_{\text{akzeptiert}}$ in der akzeptierten Ereignismenge.

Um die Effizienz-Kurve in Abhängigkeit der offline-rekonstruierten Selektionsgröße X' des Probentriggers zu erhalten, werden sowohl aus der totalen Ereignismenge als auch aus der akzeptierten Ereignismenge Histogramme in Bezug auf X' erstellt. Eine binweise Division beider ermittelten Histogramme gemäß Gl. (4.3) liefert eine gute Näherung für die Effizienz-Kurve. Mit größeren Binzahlen wird die Effizienz-Kurve tendenziell glatter. Gleichzeitig nimmt jedoch die statistische Unsicherheit auf die einzelnen Bins zu, weshalb die Verteilung der Anzahl an Ereignissen in den Bins für verschiedene Binzahlen betrachtet und eine geeignete Binzahl abgeschätzt werden sollte. Um eine noch glattere Effizienz-Kurve zu erhalten, kann in vielen Fällen die Funktion in Gl. (4.2) beispielsweise mit der Methode der kleinsten Fehlerquadrate an die genäherte Effizienz-Kurve angepasst werden. Falls durch Anpassungen keine zufriedenstellenden Ergebnisse erzielt werden können, erweist sich die Wahl der Anzahl an Bins und damit auch deren Breite als umso wichtiger. Eine konkretere Beschreibung der angewandten Methodik in dieser Arbeit erfolgt zu Beginn von Kap. 5.

Es ist erwähnenswert, dass Methoden entwickelt worden sind, mit denen die Trigger-Effizienz-Kurve ohne einen Kalibrationsdatensatz, wie er hier mittels Referenztrigger erzeugt worden ist, abgeschätzt werden kann. Eine solche Methode ist in [36] beschrieben. Methoden ohne Verwendung von Kalibrationsdatensätzen können sich als nützlich erweisen, wenn keine oder nur limitierte Kalibrationsdatensätze vorliegen. Durch diese Methode sind jedoch größere Unsicherheiten als bei der Benutzung eines großen Kalibrationsdatensatzes zu erwarten.

Es existieren mehrere mögliche Herangehensweisen bei der Bestimmung von Konfidenzintervallen für Effizienzmessungen. Von der Particle Data Group werden Clopper-Pearson-Intervalle empfohlen [37, 38]. Es wird angenommen, dass die beobachtete Trigger-Effizienz ϵ_0 einer Binomialverteilung mit der wahren Effizienz ϵ als Wahrscheinlichkeitsparameter folgt. So ergibt sich für eine beobachtete Effizienz-Kurve aus N Bins je eine beobachtete Effizienz für N Binomialverteilungen. Clopper-Pearson-Intervalle resultieren aus der Inversion zweier einseitiger Binomialtests [39]: Die obere Intervallgrenze ϵ_{\uparrow} für ein gefordertes Signifikanzniveau α ist gegeben durch die Lösung der Gleichung

$$\sum_{k=n_{\text{akz}}}^{n_{\text{tot}}} \binom{n_{\text{tot}}}{k} \epsilon_{\uparrow}^k (1 - \epsilon_{\uparrow})^{n_{\text{tot}}-k} = \frac{\alpha}{2}, \quad (4.4)$$

mit der Anzahl an akzeptierten Ereignissen n_{akz} aus der totalen Anzahl an Ereignissen n_{tot} in einem Bin. Analog ist die untere Intervallgrenze ϵ_{\downarrow} die Lösung von

$$\sum_{k=0}^{n_{\text{akz}}} \binom{n_{\text{tot}}}{k} \epsilon_{\downarrow}^k (1 - \epsilon_{\downarrow})^{n_{\text{tot}}-k} = \frac{\alpha}{2}. \quad (4.5)$$

Eine Eigenschaft von Clopper-Pearson-Intervallen ist, dass das geforderte Konfidenzniveau von $1 - \alpha$ stets überabgedeckt wird. Die Berechnung der Intervallgrenzen kann beispielsweise

Tabelle 4.1: Liste aller in Kapitel 5 verwendeten Proben- und Referenztriggerpfade. Die in den Ären 2016preVFP und 2016postVFP ausgewählten Triggerkombinationen sind identisch und deshalb als Ära 2016 zusammengefasst. Sind mehrere Triggerpfade für eine Ära genannt, so sind sie durch eine logische Adjunktion || verknüpft.

Ära	Probentrigger	Referenztrigger
2016	HLT_PFMETNoMu90_PFMHTNoMu90_IDTight	HLT_IsoMu24
	HLT_PFMETNoMu100_PFMHTNoMu100_IDTight	HLT_IsoTkMu24
	HLT_PFMETNoMu110_PFMHTNoMu110_IDTight	
	HLT_PFMETNoMu120_PFMHTNoMu120_IDTight	
2017	HLT_PFMETNoMu120_PFMHTNoMu120_IDTight_PFHT60	HLT_IsoMu27
	HLT_PFMETNoMu120_PFMHTNoMu120_IDTight	
2018	HLT_PFMETNoMu120_PFMHTNoMu120_IDTight_PFHT60	HLT_IsoMu24
	HLT_PFMETNoMu120_PFMHTNoMu120_IDTight	

mit Hilfe der Beta-Verteilung erfolgen. Genauere Details zu dieser und anderen äquivalenten Beschreibungen sind in [39] zu finden.

Im Allgemeinen können sich die daten- und simulationsbasierten Effizienzkurven unterscheiden. Dies ist ein unerwünschter Effekt und führt zu einer Verzerrung der Ergebnisse in Analysen, die simulierte Ereignisse nutzt. Um dem Effekt entgegenzuwirken, können Daten-Simulation-Skalenfaktoren ermittelt und damit die Simulation in bessere Übereinstimmung mit den Daten gebracht werden. Die Berechnung einiger solcher Skalenfaktoren erfolgt in Kap. 6.

4.3 Auswahl der Trigger

Wie in Kap. 3.2 diskutiert worden ist, sind in der Monotop-Analyse Ereignisse mit hoher fehlender Transversalenergie von Interesse. Ausgehend von der Analysestrategie ist auch beschrieben worden, dass für eine bessere Vergleichbarkeit verschiedener Kontrollregionen der hadronische Rückstoß Untersuchungsgegenstand in der Analyse ist. Daher werden für die Ären 2016preVFP, 2016postVFP, 2017 und 2018 HLT_PFMETNoMu-Triggerpfade eingesetzt, wobei sich die Schwellen von Ära zu Ära teilweise unterscheiden. Es handelt sich also um das nach dem PF-Algorithmus bestimmte MET ohne die Transversalimpulse rekonstruierter Myonen und stimmt daher mit dem hadronischen Rückstoß überein, sofern ein Ereignis keine Photonen und Elektronen enthält. In den Trigger-Studien dieser Arbeit sind die genannten Triggerpfade die Probentrigger.

Die verwendeten Referenztrigger sind HLT_IsoMu-Triggerpfade. Das bedeutet, dass der Referenztrigger nur Ereignisse akzeptiert, die mindestens ein isoliertes² Myon mit einem bestimmten minimalen Transversalimpuls besitzen. Diese Trigger bieten sich deshalb an, da in Kap. 5 nur Untergrundprozesse verglichen werden, bei denen Eichbosonen myonisch zerfallen. Die Eichbosonen stammen aus dem harten Prozess der Proton-Proton-Kollision, sodass Myonen, die aus dem Eichboson-Zerfall entstehen, prompt sind. Die Isolation ist ein gutes Kriterium, um solche prompten Myonen festzustellen, denn nicht-prompte Myonen, die beispielsweise erst durch den Zerfall eines schweren Mesons in einem Jet produziert werden, sind nicht isoliert.

Die Triggerpfade aller in den Trigger-Effizienz-Studien verwendeten Proben- und Referenztrigger sind Tabelle 4.1 zu entnehmen.

²Gemeint ist eine räumliche Isolation - also eine Abwesenheit anderer Teilchen in der näheren Umgebung des Myons.

5. Studien zu MET-Trigger-Effizienzen

Vor dem Hintergrund, dass im Rahmen der Monotop-Analyse Daten-Simulation-Skalenfaktoren für MET-Trigger-Effizienzen in der SingleMuon- und DiMuon-Kontrollregion zu bestimmen sind, werden im Folgenden vorab einige Studien zu den Effizienzen durchgeführt. Das generelle Vorgehen in den Studien besteht darin, dass Trigger-Effizienz-Kurven für HLT_PFMETNoMu-Trigger in Abhängigkeit des Hadronischen Rückstoßes erstellt und miteinander verglichen werden. Da das Ziel die Bestimmung von Daten-Simulation-Skalenfaktoren ist, wird bei den Eingangsdaten der zu ermittelnden Effizienz-Kurven zwischen Daten und simulierten Daten unterschieden. Es werden SingleMu-Daten im Dateiformat NanoAOD-v9 [40] eingesetzt. Die simulierten Daten wurden nach der Monte-Carlo-Methode erzeugt: Zur Simulation der Fragmentation und Hadronisierung wird Pythia 8 verwendet. Die Eventgeneration erfolgt mit MadGraph5aMC@NLO. Die Antwort des Detektors wird mit Geant4 simuliert.

Darüber hinaus wird in allen Studien strikt zwischen Daten und simulierten Daten aus den Ären¹ 2016preVFP, 2016postVFP, 2017 und 2018 unterschieden, da sich bei den zugehörigen Effizienz-Kurven systematische Abweichungen ergeben und die Skalenfaktoren folglich für alle Ären einzeln zu bestimmen sind. Die über den Zeitraum einer Ära integrierte Luminosität variiert von Ära zu Ära und kann daher im folgenden Kapitel den Abbildungen, die datenbasierte Effizienzen enthalten, entnommen werden.

Allen in diesem Kapitel berechneten Effizienzen liegt die in Kap. 4.2 eingeführte allgemeine Methodik zugrunde, die nun spezifiziert wird. Die Basisselektion orientiert sich an der Monotop-Analyse. Es werden zwei verschiedene Selektionen verwendet; Selektion 1 wird in Kap. 5.1 und Selektion 2 in Kap. 5.2, 5.3 und 6 eingesetzt. Nähere Details sowie die exakten Selektionsschritte können Kap. 3.2.3 entnommen werden. Daraus wird die totale Ereignismenge in Bezug auf die Probentrigger mittels der HLT_IsoMu-Referenztrigger aus Kap. 4.3 erzeugt. Die untersuchten HLT_PFMETNoMu-Probentrigger, mit denen die akzeptierte Ereignismenge selektiert wird, sind ebenfalls in Kap. 4.3 gegeben. Die Erzeugung der Histogramme erfolgt mit Hilfe eines gruppeninternen Software-Frameworks aus den Eingangsdaten. Aus den Histogrammen werden die Effizienzen mit der in der Datenanalyseumgebung ROOT [41] enthaltenen TEfficiency-Klasse [37] berechnet. Mit TEfficiency können auch verschiedene Konfidenzintervalle von Effizienzen bestimmt werden. Als Standardeinstellung werden aufgrund der Empfehlung durch die Particle Data Group Clopper-Pearson-Intervalle berechnet. In den nachfolgenden Studien werden aus demselben Grund ebenfalls Clopper-Pearson-Intervalle ermittelt. Wenn nicht anders angegeben, entspricht das geforderte Konfidenzniveau $1 - \alpha = 68,27\%$.

In Kap. 5.1 werden die Effizienz-Kurven von simulierten Datensätzen verschiedener Prozesse in der SingleMuon-Kontrollregion verglichen. Im Anschluss daran wird auf das sogenannte

¹Eine Ära stellt einen Zeitabschnitt der Datenentnahme dar. Die verschiedenen Ären spiegeln geänderte Detektorbedingungen wieder.

Sample-Stitching eingegangen, das bei den verwendeten Monte-Carlo-Datensätzen Anwendung findet. Danach wird in Kap. 5.2 untersucht, inwiefern sich die Ereignisgewichtung bei den Monte-Carlo-Simulationen auf die Effizienzen auswirkt. Die letzte Studie in Kap. 5.3 befasst sich mit den Unterschieden der Trigger-Effizienzen in der SingleMuon- und DiMuon-Kontrollregion.

5.1 Vergleich von Monte-Carlo-Datensätzen

Wie bereits beschrieben macht der Untergrundprozess W+Jets den Hauptbestandteil aller Ereignisse in der SingleMuon-Kontrollregion aus (s. Kap. 3.2.2). Demnach ist davon auszugehen, dass der W+Jets-Prozess den Verlauf der Effizienz-Kurve in dieser Kontrollregion vorgibt.

Zur Überprüfung werden die in der Monotop-Analyse eingesetzten Monte-Carlo-Simulationen betrachtet. Die Untergrundprozesse umfassen die Produktion von Z+Jets, W+Jets, $t\bar{t}$, Bosonenpaaren² und einzelnen Top-Quarks. Die Simulation für den Prozess Z+Jets geht von einem Drell-Yan-Prozess mit zusätzlichen Jets aus. Für die Ären 2017 und 2018 wird jeweils mit der zu Beginn von Kap. 5 beschriebenen Methodik eine Effizienz-Kurve mit simulierten W+Jets-Ereignissen als Eingangsdaten und eine Effizienz-Kurve mit allen genannten Prozessen erstellt. Die Ergebnisse sind in Abb. 5.1 dargestellt. Es ist zu beachten, dass aufgrund des ausgewählten Impulsbereichs nur Effizienzen nahe des Plateaus abgebildet sind. Des Weiteren sind die angegebenen Unsicherheiten keine Clopper-Pearson-Intervalle, da Monte-Carlo-Datensätze gewichtete Ereignisse enthalten und TEfficiency keine Clopper-Pearson-Intervalle für gewichtete Ereignisse unterstützt. Der Grund dafür ist, dass sich jedes Clopper-Pearson-Intervall aus zwei einseitigen Binomialtests ableitet (s. Kap. 4.2) und die totalen und akzeptierten Ereignisse deshalb ganzzahlig sein müssen, um Clopper-Pearson-Intervalle zu berechnen. Daher wird bei den Effizienzen von Simulationen eine Normal-Approximation angewandt.

Es ist zu sehen, dass für jede Ära einzeln betrachtet die Effizienzen in jedem Bin nicht um mehr als $\Delta\epsilon = 0,002$ voneinander abweichen. Daran und insbesondere auch am beinahe identischen statistischen Verhalten der Kurven ist erkennbar, dass die Ereignismenge der simulierten W+Jets-Ereignisse tatsächlich ausschlaggebend für den Verlauf der Effizienz-Kurve in der SingleMuon-Kontrollregion ist.

Eine weitere Auffälligkeit des Kurvenverhaltens ist ein scheinbarer Sprung im Hinblick auf die statistischen Unsicherheiten bei einem Hadronischen Rückstoß von ungefähr 450 GeV. Um dies zu verstehen, muss die verwendete Ereignismenge der in die Histogramme eingehenden Monte-Carlo-Simulation genauer untersucht werden: Sowohl beim Prozess W+Jets als auch Z+Jets werden je zwei verschiedene Ereignismengen zusammengefügt (*Sample-Stitching*). Ihr Unterschied ist, dass bei der einen Menge eine Einteilung anhand der Zahl der Jets („Jet-Binnung“) und bei der anderen eine Einteilung anhand des Transversalimpulses („ P_T -Binnung“) vorliegt. Ab einem Transversalimpuls des rekonstruierten Bosons von 400 GeV findet ein Wechsel von der Jet-Binnung zur P_T -Binnung statt.

Um zu testen, ob *Sample-Stitching* wirklich die Ursache für die abrupte Abnahme der statistischen Unsicherheit ist, werden für die simulierten W+Jets-Ereignisse (2018) die Effizienz-Kurven der Mengen mit Jet-Binnung und P_T -Binnung einzeln in der SingleMuon-Kontrollregion ermittelt und mit der Effizienzkurve der *Sample-Stitching*-Menge verglichen. Der Vergleich ist in Abb. 5.1 einsehbar.

Es zeigen sich eindeutige Übereinstimmungen des *sample-stitched* W+Jets-Datensatzes mit dem W+Jets-Datensatz mit Jet-Binnung bis ca. 450 GeV und mit dem W+Jets-Datensatz

²Genauer werden zwei W-Bosonen, zwei Z-Bosonen oder je ein W-Boson und ein Z-Boson erzeugt.

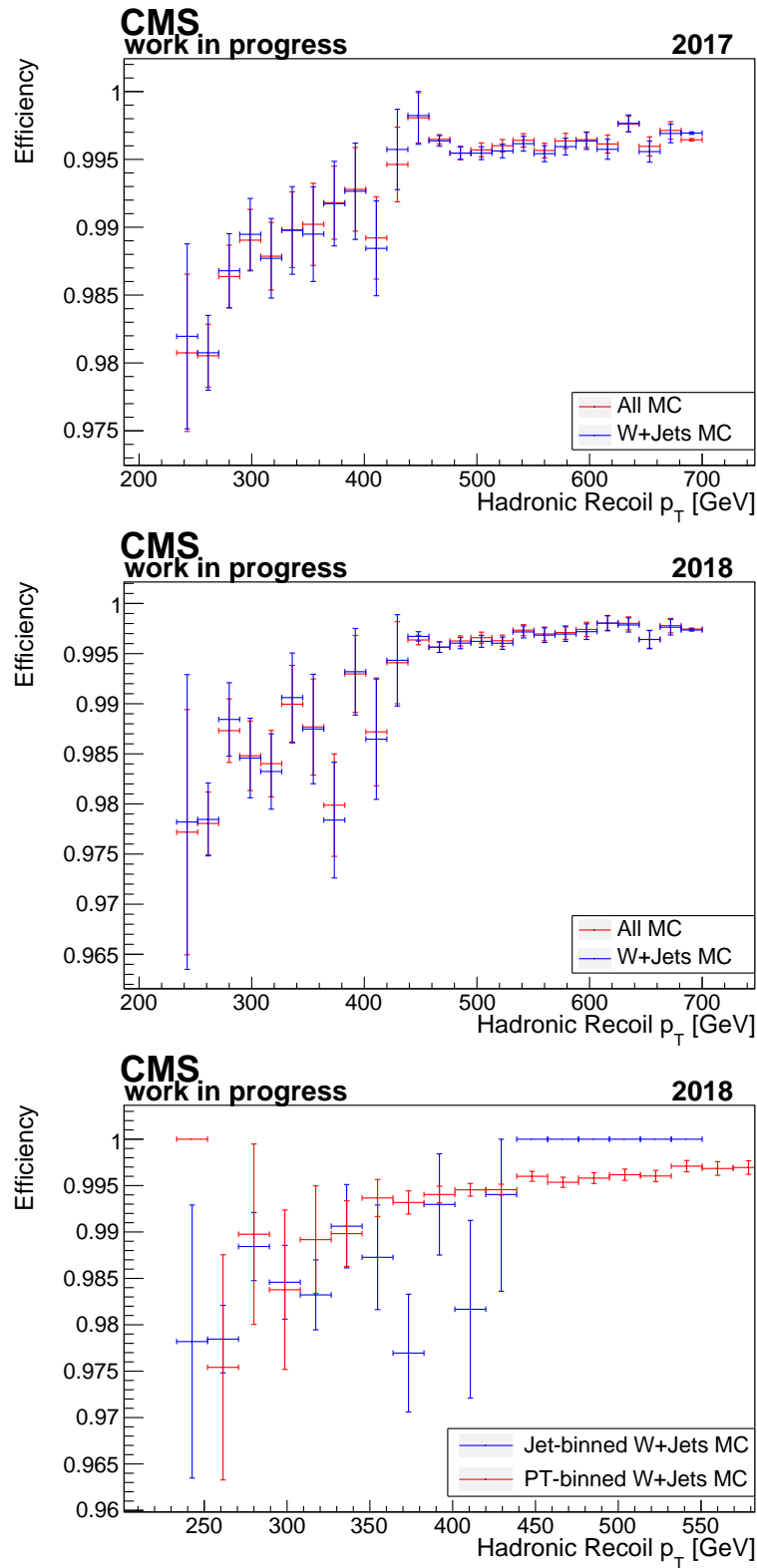


Abbildung 5.1: Simulierte Effizienzkurven für den W+Jets-Prozess und für die kumulierten Ereignisse der Prozesse Z+Jets, W+Jets, $t\bar{t}$, Diboson und SingleTop (All MC) in der SingleMuon-Kontrollregion. Es ist die Effizienzkurve zu Simulationen mit *Sample-Stitching* in den Ären 2017 (oben) und 2018 (in der Mitte) abgebildet. Unten sind die Effizienzen aus simulierten Datensätzen mit Jet-Binnung und P_T -Binnung dargestellt.

mit P_T -Binnung ab 450 GeV. Damit ist gezeigt, dass die plötzliche Abnahme der statistischen Unsicherheiten bei der W+Jets-Simulation tatsächlich vom angewandten *Sample-Stitching* stammt.

Für alle weiteren Kapitel werden nun erweiterte Eingangsdaten mit Ereignisselektion 2 (s. Kap. 3.2.3) verwendet, um die Effizienzen in der Nähe der Triggerschwelle zu untersuchen.

5.2 Auswirkungen der Ereignisgewichtung

Es ist bereits diskutiert worden, dass Clopper-Pearson-Intervalle im Sinne der zugrundeliegenden Binomialtests nicht für gewichtete Ereignisse, wie sie bei Monte-Carlo-Simulationen auftreten, berechnet werden können. Als Alternative kann für eine Simulation wie bei den Effizienz-Kurven in Kap. 5.1 eine Normal-Approximation angewandt werden. Da Clopper-Pearson-Intervalle dennoch zu bevorzugen sind, stellt sich die Frage, ob es sich bei der Ereignisgewichtung in der Monte-Carlo-Simulation um einen Einflussfaktor auf die Trigger-Effizienz handelt. Falls dies nicht der Fall ist, kann für simulierte Daten anstatt der Normal-Approximation bei gewichteten Ereignissen auch die Clopper-Pearson-Methode für die ungewichteten Ereignisse eingesetzt werden. Es scheint plausibel, dass die Ereignisgewichtung keinen großen Einfluss auf die Effizienz hat, falls die Anzahl an Ereignissen groß oder die Schwankung der Gewichte klein ist. In diesen Fällen ist zu erwarten, dass die mittlere Gewichtung der Ereignisse im akzeptierten Datensatz nahezu der mittleren Gewichtung im totalen Datensatz entspricht.

Die Methode, um Auswirkungen der Ereignisgewichtung auf die Trigger-Effizienz zu testen, ist ein simpler Vergleich. Zuerst werden aus W+Jets-Simulationen der Ären 2017 und 2018 mit gewichteten Ereignissen Trigger-Effizienz-Kurven bestimmt. Die dazugehörigen Unsicherheiten werden durch eine Normal-Approximation von TEfficiency berechnet. Danach werden zu denselben Simulationen mit ungewichteten Ereignissen Effizienz-Kurven mit Clopper-Pearson-Intervallen ermittelt. Die direkten Vergleiche sind für die Ären 2017 und 2018 in Abb. 5.2 dargestellt.

In beiden Ären unterscheiden sich die Effizienz-Kurven aus gewichteten und ungewichteten Ereignissen hauptsächlich zu Beginn des *Turn-On*. Die fehlende Transversalenergie ist in diesem Bereich allerdings gering, weshalb der Bereich von keiner großen Bedeutung bei der Monotop-Analyse ist. Im Plateau sind die Abweichungen der Effizienz mit Ausnahme eines Ausreißers bei einem Hadronischen Rückstoß von ca. 750 GeV verschwindend klein mit

$$\left| \frac{\epsilon_{\text{gewichtet}}}{\epsilon_{\text{ungewichtet}}} - 1 \right| \ll 1\%, \quad (5.1)$$

wobei $\epsilon_{\text{gewichtet}}$ die gewichtete und $\epsilon_{\text{ungewichtet}}$ die ungewichtete Effizienz in einem Bin bezeichnet. Der unerwünschte Ausreißer wird womöglich durch einen Effekt hervorgerufen, der aufgrund der unterschiedlichen Wirkungsquerschnitte von Jet-Binnung und P_T -Binnung beim *Sample-Stitching* auftritt. Die starke Übereinstimmung der Effizienzen rechtfertigt die Vernachlässigung von Gewichten zur Bestimmung der Daten-Simulation-Skalenfaktoren. Darüber hinaus kann das Vorhandensein des Ausreißers im Plateau als zusätzliches Argument dafür angeführt werden, dass ungewichtete Simulationen für die Ableitung von Skalenfaktoren zu bevorzugen sind. Noch wichtiger ist jedoch der Umstand, dass bei den benutzten Simulationen auch negative Ereignisgewichte entstehen können, was in einigen Fällen dazu führt, dass beim gewichteten Fall in einem Bin mehr Ereignisse in der akzeptierten Menge vorkommen als in der totalen Menge. In solchen Situationen werden Effizienzen $\epsilon > 1$ ohne physikalischen Sinn berechnet. Zudem lässt TEfficiency solche Effizienzen unter Ausgabe einer Fehlermeldung nicht zu. Der Effekt ist auch bei manchen der in dieser Arbeit eingesetzten W+Jets-Datensätzen in den Ären 2016 beobachtet worden.

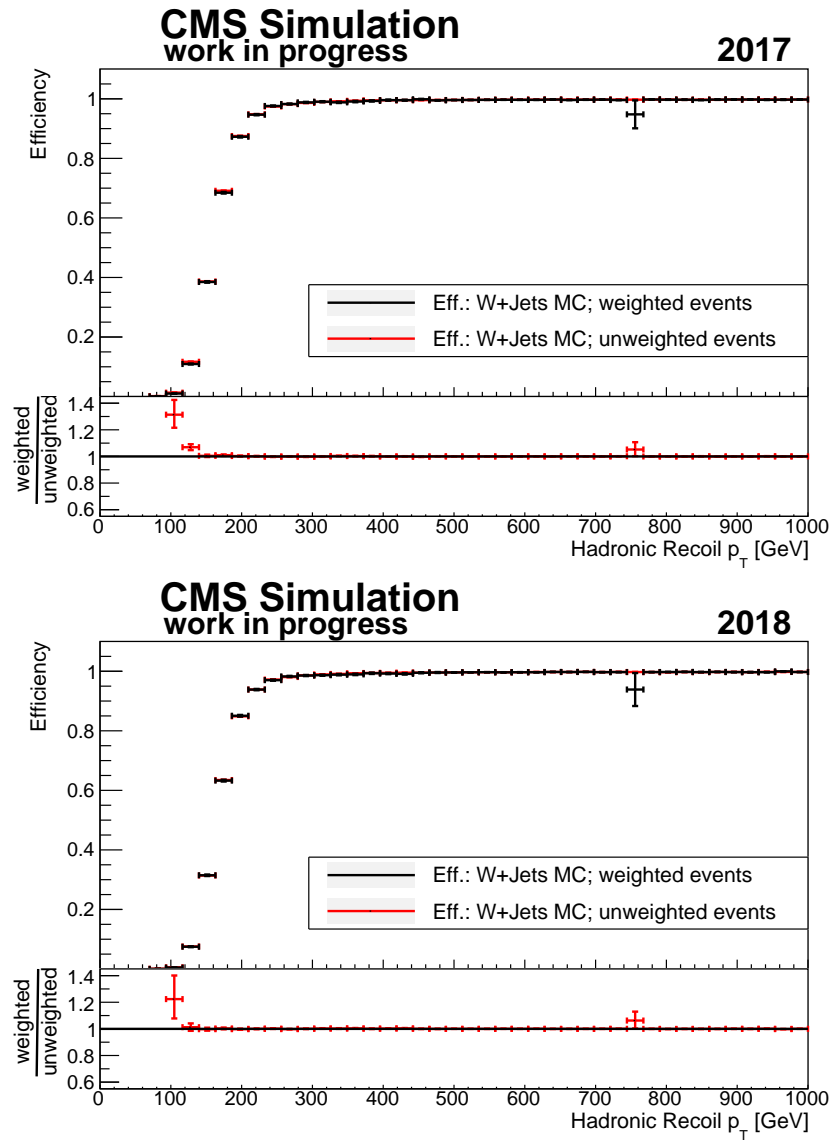


Abbildung 5.2: Effizienzkurven für simulierte W+Jets-Datensätze unter Verwendung von gewichteten und ungewichteten Ereignissen für die Ära 2017 (oben) und die Ära 2018 (unten) in der SingleMuon-Kontrollregion.

Aus den drei genannten Gründen werden in den weiteren Studien und bei der finalen Skalenfaktorenbestimmung Trigger-Effizienzen von Simulationen stets für ungewichtete Ereignisse berechnet.

Der Frage, inwiefern die Wirkungsquerschnitte beim *Sample-Stitching* tatsächlich für die Ausreißer im Plateau verantwortlich sind, kann in weiteren Untersuchungen nachgegangen werden, indem beispielsweise die Effizienz-Kurven für unterschiedliche Kombinationen an Wirkungsquerschnitten erstellt werden.

5.3 Vergleich von Kontrollregionen

Als letzte Studie vor der Berechnung der Daten-Simulation-Skalenfaktoren wird geprüft, ob sich die Effizienzen in den verschiedenen Kontrollregionen innerhalb einer Ära unterscheiden. Dies ist einerseits wichtig, um festzustellen, ob die zu berechnenden Skalenfaktoren auch einheitlich für die SingleMuon- und die DiMuon-Kontrollregion bestimmt werden können oder ob eine Differenzierung notwendig ist. Gleichzeitig wird eine Art Plausibilitätsprüfung durchgeführt, indem getestet wird, ob sich die Ergebnisse in Einklang mit früheren Studien bringen lassen.

Zu diesem Zweck werden die Effizienz-Kurven in allen vier Ären jeweils für Daten und simulierte Daten und in beiden Kontrollregionen separat ermittelt. Die verwendeten Daten sind SingleMuon-Daten, die in Form von NanoAOD-v9 vorliegen. Als simulierte Daten in der SingleMuon-Kontrollregion werden Ereignisse aus der W+Jets-Simulation genutzt, da der W+Jets-Prozess der dominante Untergrund in dieser Kontrollregion ist und daher den Verlauf der Effizienz-Kurve vorgibt, wie in Kap. 5.1 gezeigt worden ist. Analog wird angenommen, dass die Z+Jets-Datensätze den Verlauf der Effizienz-Kurve in der DiMuon-Kontrollregion vorgeben, weil der Prozess Z+Jets der Hauptuntergrund in dieser Kontrollregion ist. Die Selektionen der Kontrollregionen sind in Kap. 3.2.3 zu sehen. Motiviert durch die Ergebnisse aus Kap. 5.2 werden für alle Simulationen ungewichtete Ereignisse eingesetzt, sodass sowohl für Daten als auch für simulierte Daten Clopper-Pearson-Intervalle angegeben werden können. Die resultierenden Effizienzkurven sind zu der SingleMuon-Kontrollregion in Abb. 5.3 und zu der DiMuon-Kontrollregion in Abb. 5.4 dargestellt.

Zunächst kann festgestellt werden, dass die Effizienz-Kurven in allen vier Ären, für beide Kontrollregionen und sowohl für Daten als auch für Simulation mit dem Ergebnis einer Monojet-Analyse kompatibel sind [42]: Ähnlich wie bei der Monotop-Analyse wird eine Signalregion mit einem hohen fehlenden Transversalimpuls und Jets hoher Transversalimpulse gefordert. Zusätzlich wird eine *single-muon*- sowie eine *dimuon*-Kontrollregion definiert. Untersucht wurde ein Trigger, der ab einem fehlenden Transversalimpuls von mindestens 120 GeV auf Online-Level triggert, wobei Myonen bei der Berechnung nicht berücksichtigt worden sind. Es ist damit derselbe Trigger eingesetzt worden wie in der Monotop-Analyse. Bei der Monojet-Analyse ist festgestellt worden, dass ab einem fehlenden Transversalimpuls von 250 GeV der untersuchte Trigger eine Effizienz von mindestens 95 % aufweist. Als Vergleich weisen die 16 Effizienzkurven aus Abb. 5.3 und 5.4 bei einem Hadronischen Rückstoß von 250 GeV eine durchschnittliche Effizienz von 96 % auf.

Wie zu erwarten war, nimmt darüber hinaus in beiden Kontrollregionen die Wahrscheinlichkeit für Ereignisse mit zunehmendem hadronischen Rückstoß ab, was an den Histogrammen der totalen Ereignismengen (in Abb. 5.3 und 5.4 schwarz dargestellt) eingesehen werden kann. Daher nehmen die statistischen Unsicherheiten mit dem hadronischen Rückstoß zu. Dies wird besonders gut an den höheren Impulsbereichen in den Plateaus der Daten-Effizienz-Kurven deutlich. Im Gegensatz dazu können mit Hilfe von Monte-Carlo-Simulationen praktisch beliebig viele Ereignisse generiert werden, was sich an den Unsicherheiten der simulierten Effizienzen in beiden Kontrollregionen bemerkbar macht. Ein Vergleich zwischen den datenbasierten Effizienzen in der SingleMuon- und DiMuon-Kontrollregion zeigt zudem signifikant geringere Unsicherheiten in der SingleMuon-Kontrollregion. Die Ursache für die größeren Unsicherheiten der datenbasierten Effizienzen in der DiMuon-Kontrollregion ist das geringe Verzweigungsverhältnis des Z-Zerfalls in zwei Myonen von $(3,3662 \pm 0,0066) \%$ [11]. Dennoch sind die Unsicherheiten aller Effizienzen bis zu einem Hadronischen Rückstoß von ungefähr 700 GeV verhältnismäßig gering. Diese Erkenntnis beeinflusst die Wahl der Methode zur Unsicherheitsabschätzung der Skalenfaktoren in Kap. 6 insofern, dass für einen Hadronischen Rückstoß von bis zu 700 GeV in beiden Kontrollregionen eine konservative

Abschätzung der Unsicherheiten gerechtfertigt ist. Für die Abschätzung der Unsicherheiten von Skalenfaktoren in der SingleMuon-Kontrollregion kann zudem eine wesentlich konservativere Methode verwendet werden als in der DiMuon-Kontrollregion.

Ein direkter Vergleich der Kontrollregionen wird durch das Bilden der Quotienten k von Effizienzen ϵ_{DiMuon} in der DiMuon-Kontrollregion zu Effizienzen $\epsilon_{\text{SingleMuon}}$ in der SingleMuon-Kontrollregion erreicht:

$$k = \frac{\epsilon_{\text{DiMuon}}}{\epsilon_{\text{SingleMuon}}} . \quad (5.2)$$

Mit einem geforderten Konfidenzniveau von $1 - \alpha = 95\%$ bei der Berechnung der Clopper-Pearson-Intervalle resultieren die in den Abb. 5.5 und 5.6 zu sehenden Verhältnisse der Kontrollregionen für Daten und Simulation. Die Unsicherheiten der Quotienten sind mittels Größtfehlerabschätzung errechnet worden. Eine genaue Formulierung dieser Abschätzung sowie eine Untersuchung der zugehörigen Abdeckungswahrscheinlichkeit erfolgt in Kap. 6.

Die Abhängigkeit aller so berechneten Quotienten vom Hadronischen Rückstoß lässt sich auf konsistente Weise wie folgt beschreiben: Für kleine Impulse von ca. 100 GeV erreicht k sein Maximum mit einem Wert von bis zu 2. Bis zu Impulsen von höchstens 150-200 GeV fällt k auf Werte von $0,99 < k < 1,01$ ab und bleibt im gesamten Plateau auf diesem Niveau. Allgemein ist die Beziehung $|k - 1| \ll 1$ nicht für alle k erfüllt, weshalb es nötig ist, Skalenfaktoren für beide Kontrollregionen separat abzuleiten.

In der Monotop-Analyse wird ein Hadronischer Rückstoß von mindestens 250 eV gefordert. Da der Unterschied zwischen den Effizienzen beider Kontrollregionen im Bereich $200 \text{ GeV} < P_T < 1000 \text{ GeV}$ in einer Größenordnung von 1 % liegt, kann alternativ eine systematische Unsicherheit von 1 % auf die Effizienz angenommen werden.

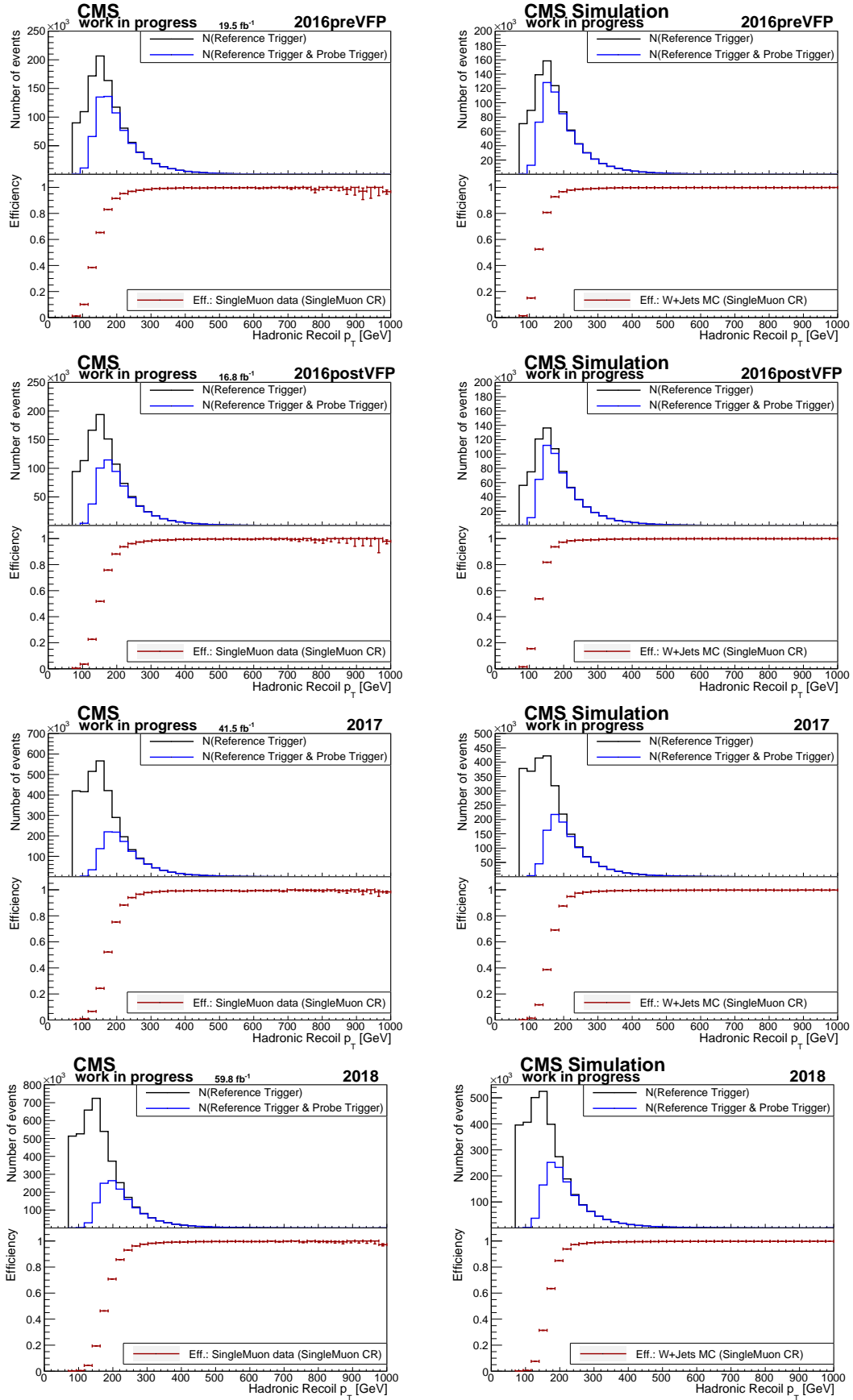


Abbildung 5.3: Effizienzkurven der SingleMu-Daten (links) und der simulierten W+Jets-Datensätze unter Verwendung von ungewichteten Ereignissen (rechts) für die Ären 2016preVFP, 2016postVFP, 2017 und 2018 (von oben nach unten) in der **SingleMuon-Kontrollregion**. Ebenfalls sind die Histogramme abgebildet, aus denen die Effizienzen berechnet worden sind, wobei für simulierte Daten die Histogramme aus gewichteten Ereignissen zu sehen sind. Das geforderte Konfidenzniveau der Clopper-Pearson-Intervalle beträgt $1 - \alpha = 68,27\%$.

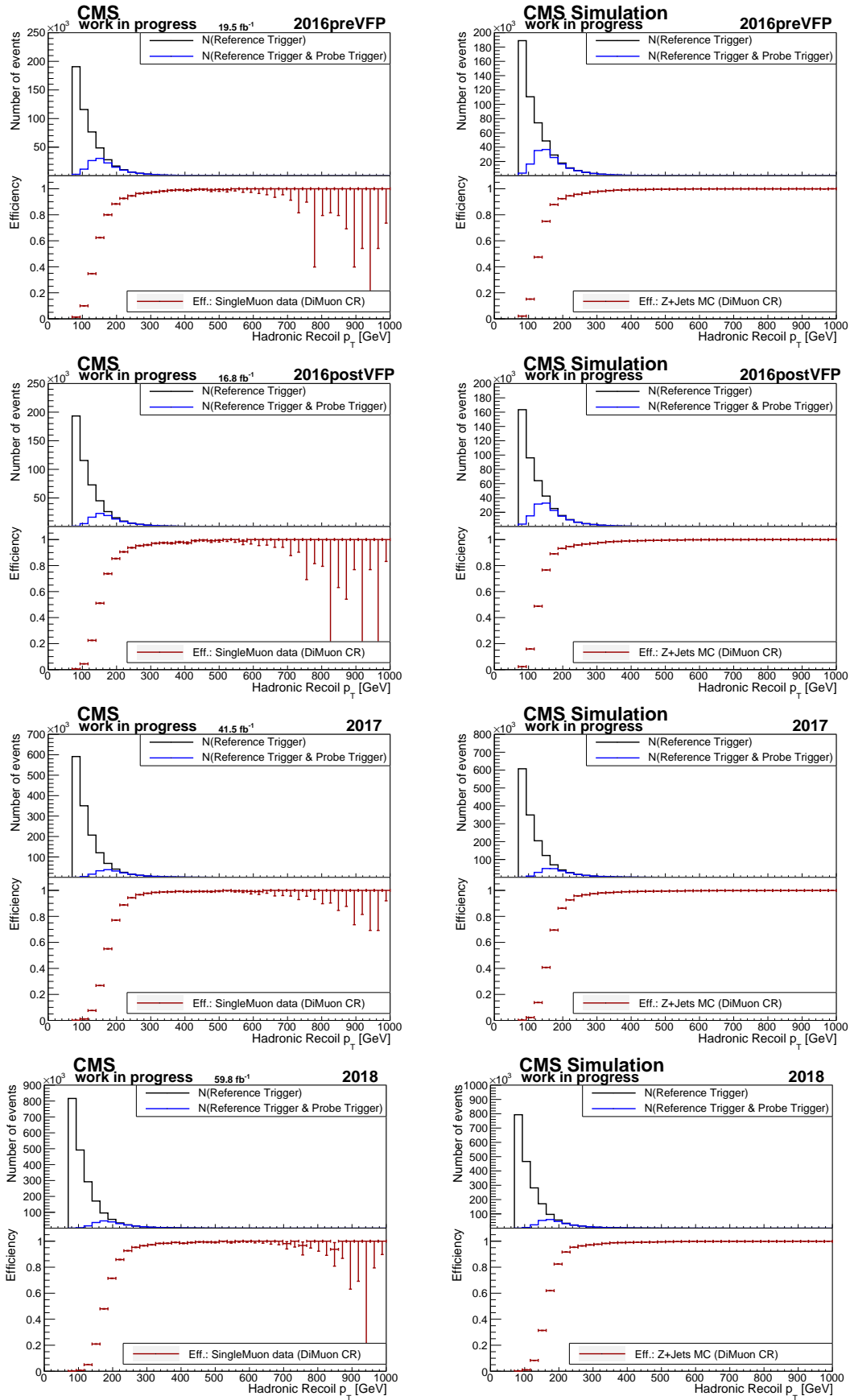


Abbildung 5.4: Effizienzkurven der SingleMu-Daten (links) und der simulierten Z+Jets-Datensätze unter Verwendung von ungewichteten Ereignissen (rechts) für die Ären 2016preVFP, 2016postVFP, 2017 und 2018 (von oben nach unten) in der **DiMuon-Kontrollregion**. Ebenfalls sind die Histogramme abgebildet, aus denen die Effizienzen berechnet worden sind, wobei für simulierte Daten die Histogramme aus gewichteten Ereignissen zu sehen sind. Das geforderte Konfidenzniveau der Clopper-Pearson-Intervalle beträgt $1 - \alpha = 68,27\%$.

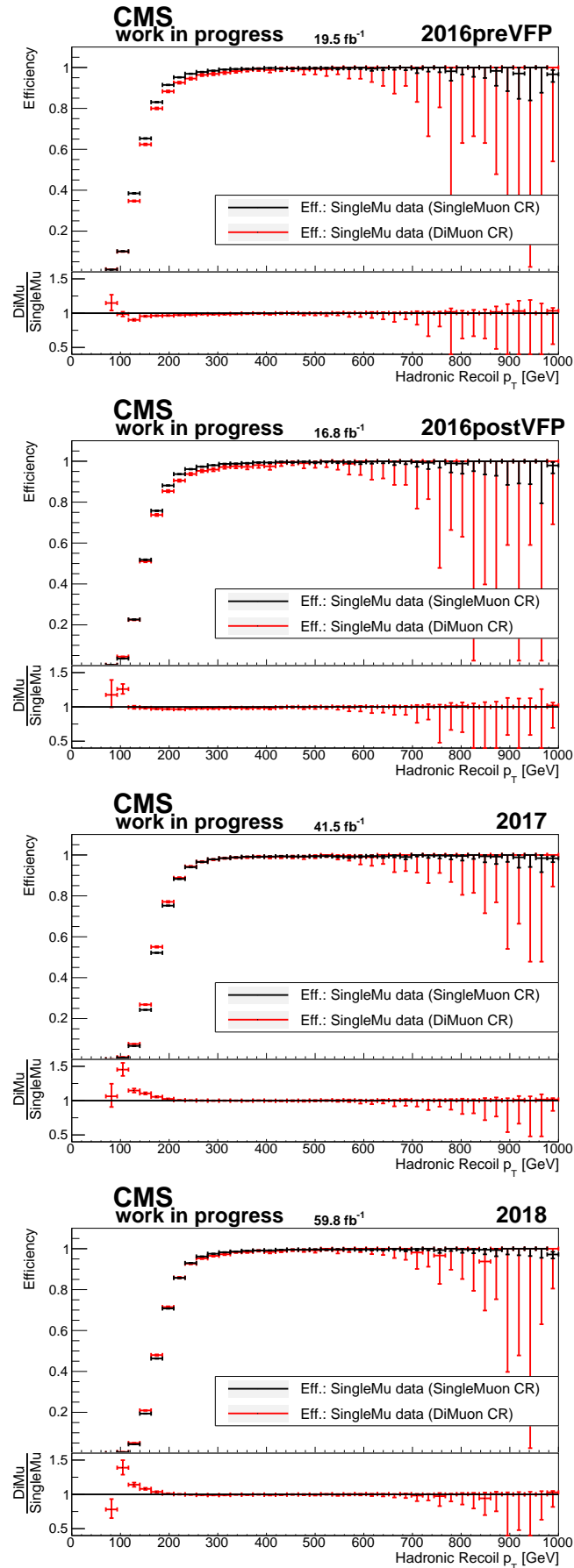


Abbildung 5.5: Datenbasierte Effizienzkurven in den SingleMuon- und DiMuon-Kontrollregionen inklusive deren Verhältnisse für die Ären 2016preVFP, 2016postVFP, 2017 und 2018 (von oben nach unten). Das geforderte Konfidenzniveau der Clopper-Pearson-Intervalle beträgt $1 - \alpha = 95,0\%$.

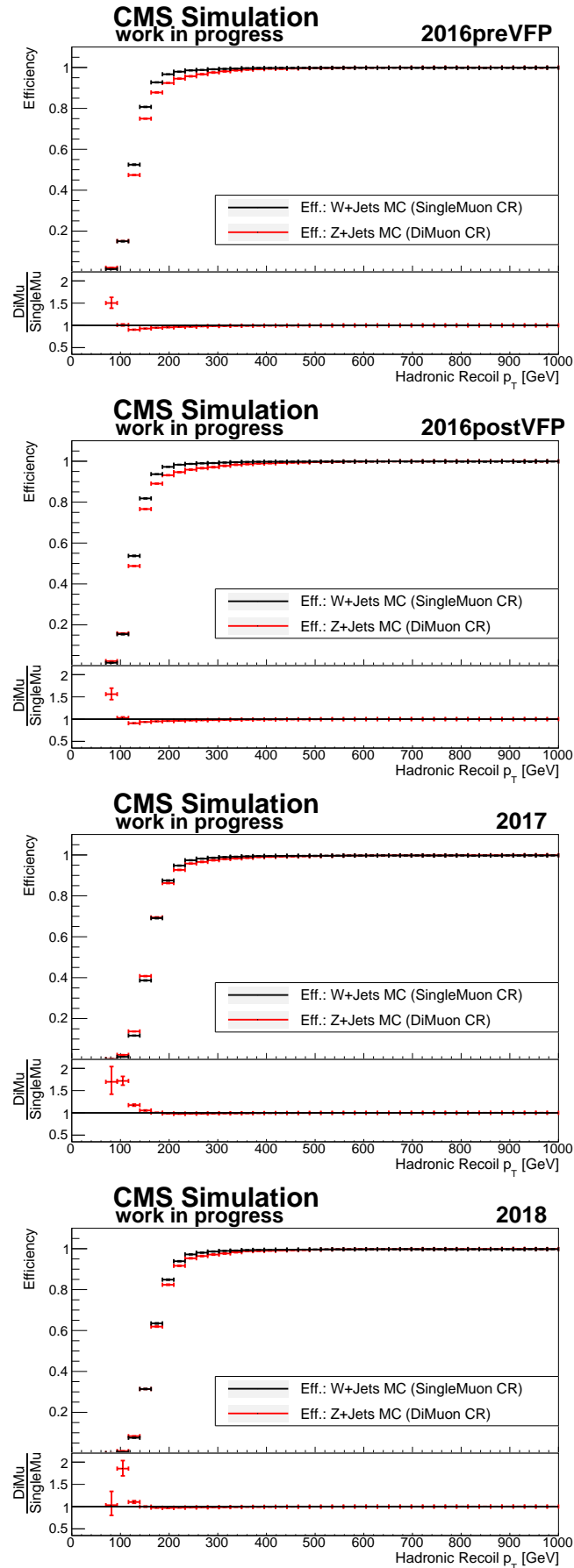


Abbildung 5.6: Simulationsbasierte Effizienzkurven in den SingleMuon- und DiMuon-Kontrollregionen inklusive deren Verhältnisse für die Ären 2016preVFP, 2016postVFP, 2017 und 2018 (von oben nach unten). Die Effizienzen sind aus Histogrammen mit ungewichteten Ereignissen berechnet worden. Das geforderte Konfidenzniveau der Clopper-Pearson-Intervalle beträgt $1 - \alpha = 95,0\%$.

6. Bestimmung von Daten-Simulation-Skalenfaktoren

Um die bei der Monotop-Analyse verwendeten MET-Trigger-Effizienzen für simulierte Daten in bessere Übereinstimmung mit den datenbasierten Effizienzen zu bringen, werden nun Daten-Simulation-Skalenfaktoren der Effizienzen bestimmt. Dabei werden die Erkenntnisse des vorherigen Kapitels berücksichtigt, die im Folgenden zusammengefasst werden.

Der Vergleich von Effizienz-Kurven eines simulierten W+Jets-Datensatzes und der kumulierten Ereignismenge der simulierten Produktion von Z+Jets, W+Jets, $t\bar{t}$, Bosonenpaaren und einzelnen Top-Quarks in der SingleMuon-Kontrollregion zeigt, dass die W+Jets-Simulation den Verlauf der Effizienz-Kurve in dieser Kontrollregion vorgibt. Daher wird nur die W+Jets-Simulation bei der Berechnung von Effizienzen in dieser Kontrollregion verwendet. In der DiMuon-Kontrollregion wird lediglich die Simulation Z+Jets eingesetzt, da insbesondere die zusätzliche Forderung einer invarianten Masse M des Systems aus den zwei rekonstruierten Myonen innerhalb von $60 \text{ GeV} < M < 120 \text{ GeV}$ gestellt wird und sich die Masse des Z-Bosons im genannten Bereich befindet. Dadurch kann der Z+Jets-Prozess in der DiMuon-Kontrollregion angereichert werden. Um sicherzustellen, dass die Z+Jets-Simulation tatsächlich den Hauptuntergrund ausmacht, kann in Analogie zur SingleMuon-Kontrollregion ein Vergleich durchgeführt werden.

Bei allen Simulationsprozessen werden Ereignisgewichtungen vernachlässigt, da sie sich kaum auf die Trigger-Effizienz auswirken. Zudem ist aufgrund von negativen Ereignisgewichten in einigen wenigen Ereignismengen eine Bestimmung der Trigger-Effizienz-Kurve unter Verwendung von Gewichten gar nicht erst möglich.

Beim Vergleich der Effizienzen in beiden Kontrollregionen sind Diskrepanzen in der Größenordnung von 1 % erkennbar. Dies kann als systematische Unsicherheit oder durch eine separate Bestimmung der Skalenfaktoren für beide Kontrollregionen berücksichtigt werden. Die geringen Unsicherheiten aller Effizienzen (mit Ausnahme der Effizienzen bei hohem Hadronischen Rückstoß für Daten in der DiMuon-Kontrollregion) rechtfertigen außerdem eine verhältnismäßig konservative Methode zur Bestimmung der Skalenfaktor-Unsicherheiten.

Ein Daten-Simulation-Skalenfaktor SF berechnet sich pro Bin nach

$$\text{SF} = \frac{\epsilon_{\text{Daten}}}{\epsilon_{\text{Simulation}}}, \quad (6.1)$$

mit der datenbasierten Trigger-Effizienz ϵ_{Daten} und der simulationsbasierten Trigger-Effizienz $\epsilon_{\text{Simulation}}$. Wie motiviert worden ist, werden dabei für alle simulierte Datensätze ungewichtete Ereignisse verwendet.

Als Unsicherheiten der daten- und simulationsbasierten Trigger-Effizienzen werden Clopper-Pearson-Intervalle mit gefordertem Konfidenzniveau von $1 - \alpha = 95,0 \%$ berechnet, da in Kap. 5.3 größtenteils geringe Unsicherheiten der Effizienzen beobachtet worden sind.

Damit kann eine hohe Abdeckungswahrscheinlichkeit bei kleinen Unsicherheitsintervallen erreicht werden. Zur Bestimmung der Unsicherheiten wird eine Größtfehlerabschätzung durchgeführt. Die obere und untere Intervallgrenze SF_{\uparrow} und SF_{\downarrow} ist gegeben durch

$$SF_{\uparrow} = \frac{\epsilon_{\text{Daten},\uparrow}}{\epsilon_{\text{Simulation},\downarrow}}, \quad (6.2)$$

$$SF_{\downarrow} = \frac{\epsilon_{\text{Daten},\downarrow}}{\epsilon_{\text{Simulation},\uparrow}}, \quad (6.3)$$

mit der oberen und unteren Intervallgrenze $\epsilon_{\text{Daten},\uparrow}$ und $\epsilon_{\text{Daten},\downarrow}$ der datenbasierten Effizienz und der oberen und unteren Intervallgrenze $\epsilon_{\text{Simulation},\uparrow}$ und $\epsilon_{\text{Simulation},\downarrow}$ der simulationsbasierten Effizienz. Die Abdeckungswahrscheinlichkeit dieser Abschätzung der Unsicherheiten kann mittels Pseudoexperimenten ermittelt werden. Mit einer Modifizierung des Python-Skripts in [43] wird die Abdeckungswahrscheinlichkeit eines Skalenfaktors $SF = \frac{\epsilon_1}{\epsilon_2}$ für $N = 5 \cdot 10^5$ Zufallsexperimente simuliert. In jedem der Zufallsexperimente $i \in \{1, 2, \dots, N\}$ werden $n_{1,2}$ Ereignisse mit festgelegten wahren Effizienzen $\epsilon'_{1,2}$ selektiert. Daraus werden Clopper-Pearson-Intervalle berechnet und für jedes Experiment i der simulierte Skalenfaktor mit Unsicherheits-Größtfehlerabschätzung bestimmt. Die relative Anzahl aller i Größtfehlerintervalle, die den wahren Skalenfaktor $SF' = \epsilon'_1/\epsilon'_2$ enthalten, ist eine Näherung der Abdeckungswahrscheinlichkeit. Es lässt sich feststellen, dass die Abdeckungswahrscheinlichkeit für verschiedene Kombinationen von $n_{1,2}$ und $\epsilon'_{1,2}$ meist mehr als 99 % beträgt. Für $n_1 = 100$, $n_2 = 400$ und zwei verschiedene Kombinationen von $\epsilon'_{1,2}$ sind die Zufallsexperimente in Abb. 6.1 visualisiert.

Die Methode der Größtfehlerabschätzung von Clopper-Pearson-Intervallen kann damit als höchst konservativ beurteilt werden.

Die Daten-Simulation-Skalenfaktoren mit Größtfehlerabschätzung der Unsicherheiten können für die SingleMuon-Kontrollregion in Abb. 6.2 und für die DiMuon-Kontrollregion in Abb. 6.3 eingesehen werden. In allen vier Ären und beiden Kontrollregionen liegt ein Minimum der Skalenfaktoren bei ca. 120 GeV Hadronischem Rückstoß vor. Für Impulse größer als 250 GeV sind Skalenfaktoren im Bereich von ca. $0,99 \leq SF \leq 1,01$ beobachtbar. Es ist zudem erkennbar, dass die Unsicherheitsintervalle in der SingleMuon-Kontrollregion trotz der konservativen Abschätzung klein sind. Auch in der DiMuon-Kontrollregion sind die Intervalle bis zu einem Hadronischen Rückstoß von ca. 700 GeV verhältnismäßig klein. Erst für höhere Transversalimpulse sind Effizienzen mit großen Unsicherheitsintervallen insbesondere in den Ären 2016preVFP und 2016postVFP beobachtbar. Dies war auch zu erwarten, da die Bins hoher Impulse in der DiMuon-Kontrollregion meist weniger als 10 Ereignisse enthalten. Insgesamt ist die verwendete Methode zur Bestimmung von Skalenfaktor-Unsicherheiten besonders gut für die SingleMuon-Kontrollregion in der Monotop-Analyse, da hier sehr hohe Abdeckungswahrscheinlichkeiten bei verhältnismäßig geringen Unsicherheitsintervallen erzielt werden konnten. Auch in der DiMuon-Kontrollregion erweist sich die Methode als ausreichend gut. Falls weniger Wert auf eine hohe Abdeckungswahrscheinlichkeit gelegt wird, ist es ebenso möglich, Clopper-Pearson-Intervalle mit einem geforderten Konfidenzniveau von beispielsweise $1 - \alpha = 68\%$ zu berechnen, um die Unsicherheitsintervallgröße in der DiMuon-Kontrollregion zu reduzieren. Hierfür wären weitere Untersuchungen der Abdeckungswahrscheinlichkeit für verschiedene Konfidenzniveaus der Clopper-Pearson-Intervalle von Interesse.

Eine Alternative zur Bestimmung der Unsicherheiten könnte darin bestehen, die Verteilungen von Skalenfaktoren in Zufallsexperimenten zu untersuchen. Zu einem Histogramm der simulierten Skalenfaktoren, wie es in Abb. 6.1 zu sehen ist, könnte dann eine passende Wahrscheinlichkeitsdichtefunktion wie eine Gaußverteilung gesucht werden, um daraus Unsicherheiten zu definieren.

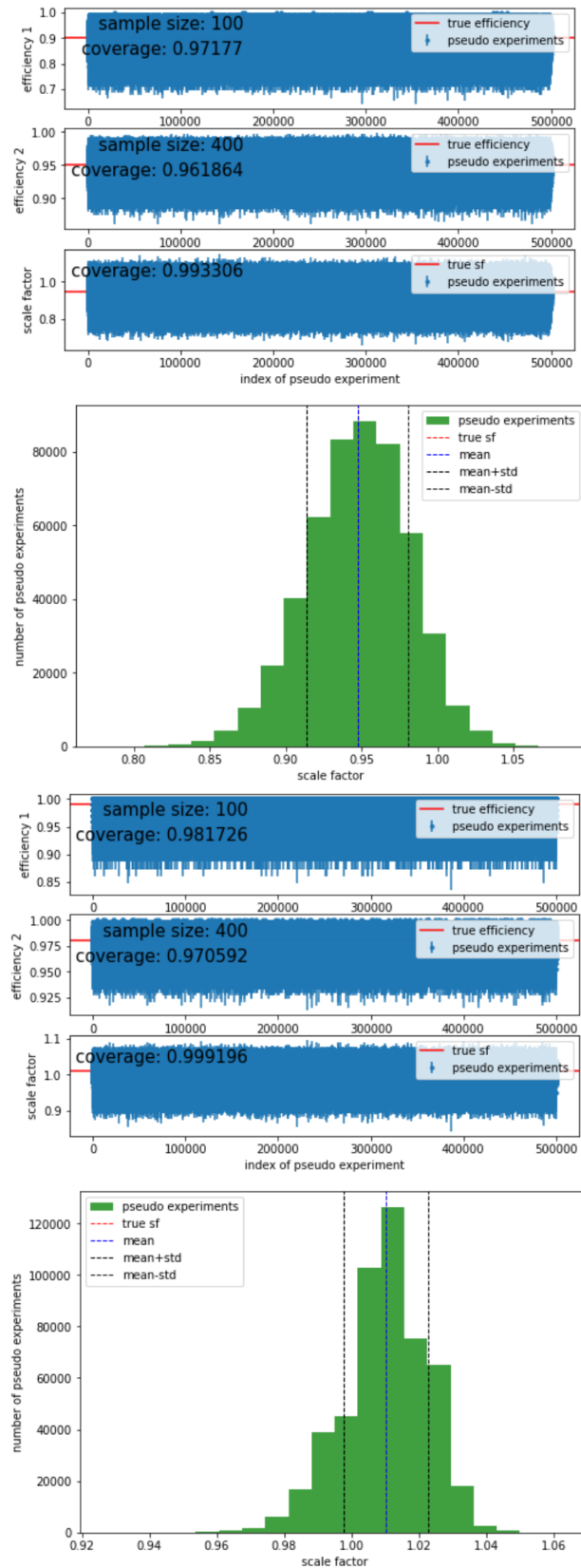


Abbildung 6.1: Untersuchung der Abdeckungswahrscheinlichkeit für eine Größtfehlerabschätzung von Skalenfaktoren aus Effizienzen ϵ_1 , ϵ_2 mit Clopper-Pearson-Intervallen mittels Zufallsexperimenten. In den oberen vier Abbildungen wird die Abdeckungswahrscheinlichkeit der Skalenfaktoren für die wahren Effizienzen $\epsilon_1 = 0.9$ und $\epsilon_2 = 0.95$ simuliert. In den unteren vier Abbildungen wird die Abdeckungswahrscheinlichkeit der Skalenfaktoren für die wahren Effizienzen $\epsilon_1 = 0.99$ und $\epsilon_2 = 0.98$ simuliert. Jeweils von oben nach unten zu sehen sind die Pseudoexperimente von ϵ_1 , ϵ_2 und der Skalenfaktoren sowie ein Histogramm aller simulierter Skalenfaktoren. Zusätzlich sind die ermittelten Abdeckungswahrscheinlichkeiten der Effizienzen und Skalenfaktoren angegeben.

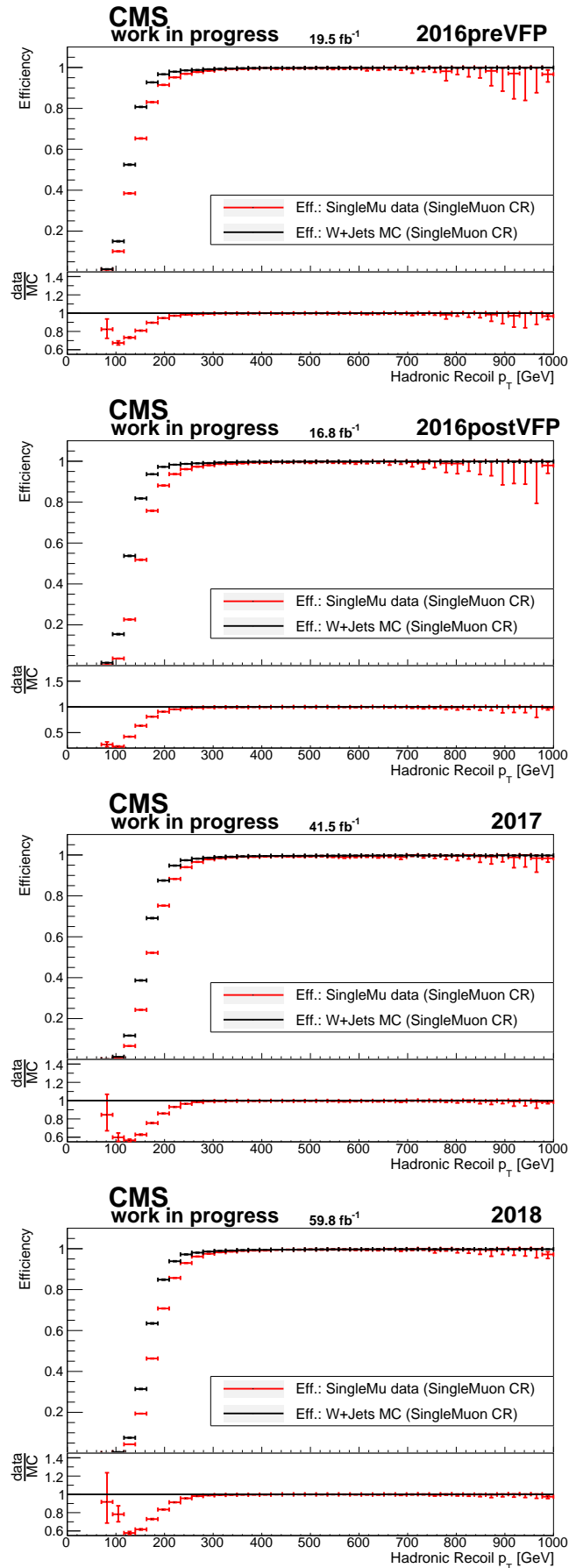


Abbildung 6.2: Daten-Simulation-Skalenfaktoren in der SingleMuon-Kontrollregionen für die Ären 2016preVFP, 2016postVFP, 2017 und 2018 (von oben nach unten). Die simulationsbasierten Effizienzen sind aus Histogrammen mit ungewichteten Ereignissen berechnet worden. Alle Effizienz-Unsicherheiten sind Clopper-Pearson-Intervalle mit gefordertem Konfidenzniveau von $1 - \alpha = 95,0\%$. Die Unsicherheiten der Skalenfaktoren stellen Größtfehlerabschätzungen dar.

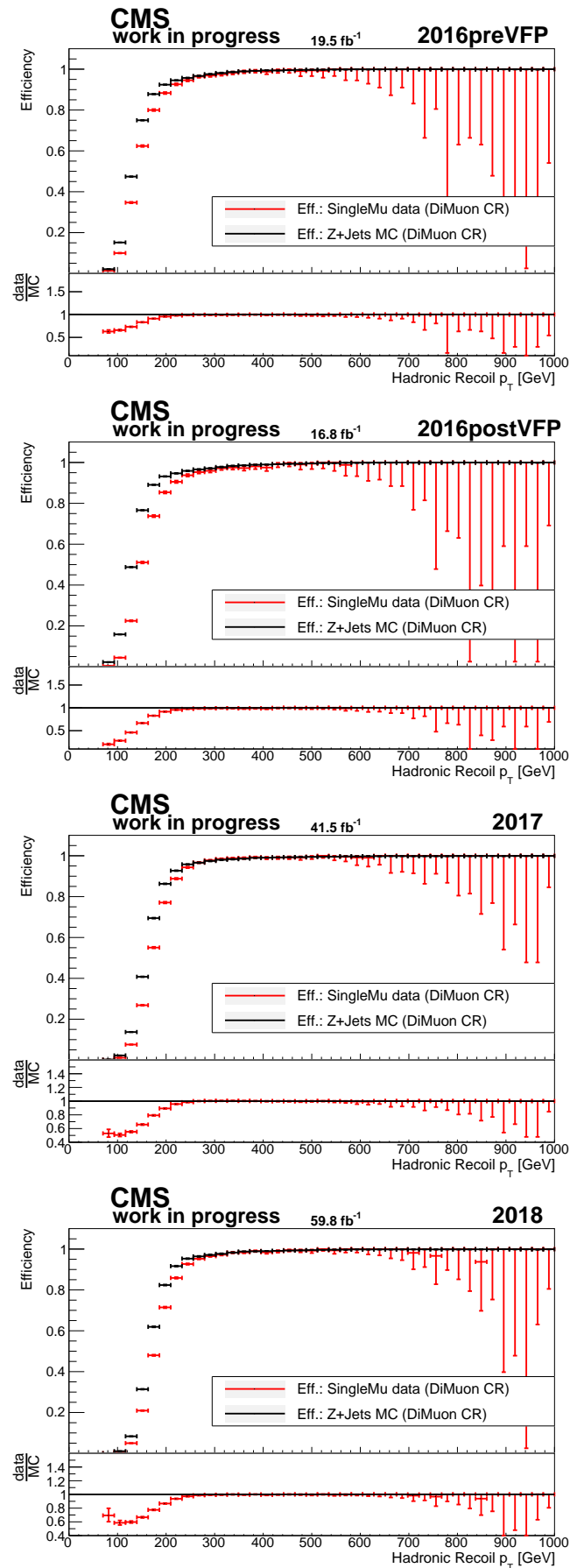


Abbildung 6.3: Daten-Simulation-Skalenfaktoren in der DiMuon-Kontrollregionen für die Ären 2016preVFP, 2016postVFP, 2017 und 2018 (von oben nach unten). Die simulationsbasierten Effizienzen sind aus Histogrammen mit ungewichteten Ereignissen berechnet worden. Alle Effizienz-Unsicherheiten sind Clopper-Pearson-Intervalle mit gefordertem Konfidenzniveau von $1 - \alpha = 95,0\%$. Die Unsicherheiten der Skalenfaktoren stellen Größtfehlerabschätzungen dar.

7. Zusammenfassung und Ausblick

Astronomische und kosmologische Beobachtungen sprechen für die Existenz einer nicht mit dem Standardmodell der Teilchenphysik zu vereinbarenden Dunklen Materie. Auf der einen Seite bestehen unterschiedliche Theorien, die diese neue Form von Materie zu erklären versuchen und auf der anderen Seite gibt es Bestrebungen, Dunkle Materie experimentell nachzuweisen. Mit dem nicht-resonanten Monotop-Prozess ist ein Modell gegeben, durch das Dunkle Materie hypothetisch in einer kontrollierten Laborumgebung wie am CMS-Experiment produziert werden kann. Basierend auf diesem Modell wird in einer Monotop-Analyse mit Hilfe von CMS-Daten und Daten aus Monte-Carlo-Simulationen nach der Produktion von Dunkler Materie gesucht. Aufgrund der Analysestrategie werden bei der Datenaufnahme Trigger verwendet, die sensitiv auf den Transversalimpuls des Hadronischen Rückstoßes sind und deren Selektionsverhalten durch Trigger-Effizienzen beschrieben werden kann. Da im Allgemeinen Unterschiede zwischen Effizienzen aus verschiedenen Datensätzen bestehen können, ist es notwendig, die Trigger-Effizienzen näher zu untersuchen. Solche Untersuchungen der Trigger-Effizienzen in Abhängigkeit des Hadronischen Rückstoßes mit dem Ziel des Ableitens von Daten-Simulation-Skalenfaktoren standen im Fokus dieser Arbeit.

In den Trigger-Effizienz-Studien zur Monotop-Analyse ist zunächst festgestellt worden, dass sich die Effizienzen des Hauptuntergrundprozesses $W+Jets$ von den Effizienzen für die kumulativen Ereignisse aller relevanten Prozesse der Analyse nahezu decken. Dabei ist auch festgestellt worden, dass das bei den Simulationen eingesetzte *Sample-Stitching* für eine abrupte Abnahme der statistischen Unsicherheiten im Plateau der Effizienz-Kurven sorgt. Im nächsten Schritt ist getestet worden, ob sich die Ereignisgewichtung von simulierten Ereignissen auf die Trigger-Effizienz auswirkt. Mit Ausnahme des Beginns vom *Turn-On* und einem Ausreißer bei höheren Impulsen stimmt die Effizienz-Kurve aus gewichteten Ereignissen mit der Effizienz-Kurve aus ungewichteten Ereignissen überein. Es ist zu vermuten, dass das *Sample-Stitching* Ursache für den Ausreißer im hohen Impulsbereich ist. Danach ist untersucht worden, wie sich die Effizienz-Kurven in den SingleMuon- und DiMuon-Kontrollregionen unterscheiden. Die Erkenntnis ist, dass sich die Effizienzen nur für einen geringen Hadronischen Rückstoß stark voneinander unterscheiden. Im Plateau liegen nur Abweichungen in der Größenordnung von 1 % vor. Diese Abweichungen können als systematische Unsicherheiten berücksichtigt werden.

Zuletzt sind für alle vier betrachteten Ären und für beide Kontrollregionen separat Daten-Simulation-Skalenfaktoren der Trigger-Effizienzen abgeleitet worden. Dabei sind Unsicherheiten mittels Größtfehlerabschätzung aus den Clopper-Pearson-Intervallen der daten- und simulationsbasierten Effizienzen bestimmt worden. Trotz der konservativen Abschätzung sind in der SingleMuon-Kontrollregion geringe Unsicherheitsintervalle bei hoher Abdeckungswahrscheinlichkeit erzielt worden. Dies gilt auch für die DiMuon-Kontrollregion bis zu einem Impuls von 700 GeV.

Um die Unsicherheitsintervalle für hohe Impulse in der DiMuon-Kontrollregion zu re-

duzieren, können andere Methoden eingesetzt werden. Ein mögliches Vorgehen besteht darin, die Verteilung der Skalenfaktoren mit Hilfe von Zufallsexperimenten zu simulieren und daraus Unsicherheiten zu definieren. Darüber hinaus kann untersucht werden, ob das *Sample-Stitching* tatsächlich verantwortlich für Ausreißer in den hohen Energiebereichen einiger simulationsbasierter Effizienzen ist. Weiterführend ist es von Interesse, die Trigger-Effizienzen zum dritten wichtigen Untergrundprozess $t\bar{t}$ in einer geeigneten Kontrollregion genauer zu untersuchen und mit den Prozessen W +Jets und Z +Jets zu vergleichen. Abschließend könnte in weiteren Studien geklärt werden, welche sonstigen systematischen Unsicherheiten einen Einfluss auf die Trigger-Effizienzen und deren Daten-Simulation-Skalenfaktoren haben könnten.

Literaturverzeichnis

- [1] The CMS Collaboration. „The CMS experiment at the CERN LHC“. In: *JINST* 3 (2008), S08004. DOI: 10.1088/1748-0221/3/08/S08004.
- [2] L. Evans und P. Bryant. „LHC machine“. In: *JINST* 3 (2008), S08001. DOI: 10.1088/1748-0221/3/08/S08001.
- [3] The CMS Collaboration. *Search for dark matter in events with energetic, hadronically decaying top quarks and missing transverse momentum at $\sqrt{s} = 13$ TeV*. In: *J. High Energ. Phys.* 2018, 27. DOI: [https://doi.org/10.1007/JHEP06\(2018\)027](https://doi.org/10.1007/JHEP06(2018)027).
- [4] V. Khachatryan et al. „The CMS trigger system“. In: *JINST* 12 (2017), P01020. DOI: <https://doi.org/10.1088/1748-0221/12/01/P01020>.
- [5] B. Povh et al. *Teilchen und Kerne: Eine Einführung in die physikalischen Konzepte*. 9. Auflage. Berlin: Springer, 2014. DOI: <https://doi.org/10.1007/978-3-642-37822-5>.
- [6] H. Chen et al. „The hidden-charm pentaquark and tetraquark states“. In: *Phys. Rept.* 639 (2016), S. 1-121. DOI: <https://doi.org/10.1016/j.physrep.2016.05.004>.
- [7] E. Klempt, A. Zaitsev. „Glueballs, hybrids, multiquarks: Experimental facts versus QCD inspired concepts“. In: *Phys. Rept.* 454 (2007), S. 1-202. DOI: <https://doi.org/10.1016/j.physrep.2007.07.006>.
- [8] Y. Fukuda et al. „Evidence for oscillation of atmospheric neutrinos“. In: *Phys. Rev. Lett.* 81 (1998), S. 1562–1567. DOI: <https://doi.org/10.1103/PhysRevLett.81.1562>.
- [9] The KATRIN Collaboration. *KATRIN: Status and Prospects for the Neutrino Mass and Beyond*. März 2022. arXiv: 2203.08059 [nucl-ex].
- [10] *Standard-Modell der Elementarteilchen*. 2010. URL: https://commons.wikimedia.org/wiki/File:Standard_Model_of_Elementary_Particles-de.svg [Zugriff am: 13.07.22].
- [11] P.A. Zyla et al. (Particle Data Group). „The Review of Particle Physics“. In: *Prog. Theor. Exp. Phys.* 2020.
- [12] *Facts and figures about the LHC*. URL: <https://home.cern/resources/faqs/facts-and-figures-about-lhc> [Zugriff am: 13.07.22].
- [13] The CMS Collaboration. „Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC“. In: *Phys. Lett. B* 716 (2012), S. 30-61. DOI: <https://doi.org/10.1016/j.physletb.2012.08.021>.
- [14] The ATLAS Collaboration. „Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC“. In: *Phys. Lett. B* 716 (2012), S. 1-29. DOI: <https://doi.org/10.1016/j.physletb.2012.08.020>.

- [15] Webseite des CMS-Detektors, URL: <https://cms.cern/detector> [Zugriff am: 13.07.22].
- [16] D. Barney. *CMS Detector Slice*. 2016. URL: <https://cds.cern.ch/record/2120661>.
- [17] The CMS Collaboration. *The CMS muon project: Technical Design Report*. Genf: CERN, 1997. URL: <https://cds.cern.ch/record/343814>.
- [18] The CMS Collaboration. „Particle-flow reconstruction and global event description with the CMS detector“. In: *JINST* 12 (2017), P10003. DOI: <https://doi.org/10.1088/1748-0221/12/10/P10003>
- [19] T. Sjöstrand. „An introduction to PYTHIA 8.2“. In: *Comput. Phys. Commun.* 191 (2015), S. 159-177. DOI: <https://doi.org/10.1016/j.cpc.2015.01.024>.
- [20] 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“. In: *J. High Energ. Phys.* 2014, 79. DOI: <https://doi.org/10.1007/JHEP07%282014%29079>.
- [21] S. Agostinelli et al. „Geant4—a simulation toolkit“. In: *Nucl. Instrum. Methods Phys. Res. A* 506(3) (2003), S. 250-303. DOI: [https://doi.org/10.1016/S0168-9002\(03\)01368-8](https://doi.org/10.1016/S0168-9002(03)01368-8).
- [22] G. Bertone, D. Hooper und J. Silk, „Particle dark matter: evidence, candidates and constraints“. In: *Phys. Rept.* 405 (2005), S. 279-390, DOI: 10.1016/j.physrep.2004.08.031, arXiv: hep-ph/0404175.
- [23] F. Zwicky. „Die Rotverschiebung von extragalaktischen Nebeln“. In: *Helvetica Physica Acta* 6 (1933), S. 110-127.
- [24] E. Corbelli, P. Salucci. „The extended rotation curve and the dark matter halo of M33“. In: *Monthly Notices of the Royal Astronomical Society* 311(2) (Januar 2000). DOI: <https://doi.org/10.1046/j.1365-8711.2000.03075.x>.
- [25] J. A. Tyson, G. P. Kochanski und I. P. Dell’Antonio. „DETAILED MASS MAP OF CL 0024+1654 FROM STRONG LENSING“. In: *The Astrophysical Journal* 498 (1998), S.L107–L110. DOI: <https://doi.org/10.1086/311314>.
- [26] B. Ryden. *Introduction to Cosmology*. Camebridge University Press, 2016. DOI: <https://doi.org/10.1017/9781316651087>.
- [27] Planck Collaboration. *Planck 2018 results. VI. Cosmological parameters*. 2021. DOI: <https://doi.org/10.48550/arXiv.1807.06209>.
- [28] W. Hu. *Intermediate Guide to the Acoustic Peaks and Polarization*. 2001. URL: <http://background.uchicago.edu/~whu/intermediate/intermediate.html> [Zugriff am: 22.06.2022].
- [29] J. L. Feng. „Dark Matter Candidates from Particle Physics and Methods of Detection“. In: *Ann. Rev. Astron. Astrophys.* 48 (2010), S. 495-545. DOI: <https://doi.org/10.1146/annurev-astro-082708-101659>.
- [30] M. Dine, W. Fischler, M. Srednicki. „A simple solution to the strong CP problem with a harmless axion“. In: *Physics Letters B* 104(3) (1981), S. 199-202. DOI: [https://doi.org/10.1016/0370-2693\(81\)90590-6](https://doi.org/10.1016/0370-2693(81)90590-6).
- [31] M. Waßmer. *Search for the production of Dark Matter in hadronic mono-top signatures with the CMS experiment*. Dissertation. Karlsruher Institut für Technologie (KIT), 2020.

- [32] M. Cacciari, G. P. Salam, G. Soyez. „The anti-k_t jet clustering algorithm“. In: *JHEP* 04 (2008), 063. DOI: <https://doi.org/10.1088/1126-6708/2008/04/063>.
- [33] D. E. Kaplan, K. Rehermann, M. D. Schwartz und B. Tweedie. „Top Tagging: A Method for Identifying Boosted Hadronically Decaying Top Quarks“. In: *Phys. Rev. Lett.* 101 (2008), 142001. DOI: [10.1103/PhysRevLett.101.142001](https://doi.org/10.1103/PhysRevLett.101.142001).
- [34] A. J. Larkoski, S. Marzani, G. Soyez, J. Thaler. „Soft Drop“. In: *JHEP* 1405 (2014), 146. DOI: <https://doi.org/10.1007/JHEP05%282014%29146>.
- [35] The CMS Collaboration. „The CMS high level trigger“. In: *Eur. Phys. J. C* 46 (2006), S. 605–667. DOI: <https://doi.org/10.1140/epjc/s2006-02495-8>.
- [36] T. R. Pollmann. „Estimating the efficiency turn-on curve for a constant-threshold trigger without a calibration dataset“. In: *Eur. Phys. J. C* 79 (2019), 322. DOI: <https://doi.org/10.1140/epjc/s10052-019-6841-1>.
- [37] *TEfficiency Class Reference*. URL: <https://root.cern.ch/doc/master/classTEfficiency.html> [Zugriff am: 13.07.22].
- [38] C. J. Clopper, E. S. Pearson. „The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial“. In: *Biometrika* 26(4) (1934), S. 404-413.
- [39] A. Agresti, B. A. Coull. „Approximate Is Better than ”Exact” for Interval Estimation of Binomial Proportions“. In: *The American Statistician* 52(2) (Mai 1998), S. 119-126.
- [40] M. Peruzzi, G. Petrucciani, A. Rizzi. „The NanoAOD event data format in CMS“. In: *J. Phys.: Conf. Ser.* 1525 (2020), 012038. DOI: [10.1088/1742-6596/1525/1/012038](https://doi.org/10.1088/1742-6596/1525/1/012038).
- [41] R. Brun, F. Rademakers. „ROOT - An Object Oriented Data Analysis Framework“. In: *Phys. Res. A* 389 (1997), S. 81-86. URL: <https://github.com/root-project/root>.
- [42] The CMS Collaboration. „Search for new particles in events with energetic jets and large missing transverse momentum in proton-proton collisions at $\sqrt{s} = 13$ TeV“. In: *J. High Energ. Phys.* 2021, 153. DOI: [https://doi.org/10.1007/JHEP11\(2021\)153](https://doi.org/10.1007/JHEP11(2021)153).
- [43] M. Waßmer. *Clopper Pearson Playground*. URL: https://github.com/michaelwassmer/useful_scripts/blob/master/Statistics/ClopperPearson.py [Zugriff am: 14.07.2022].