

TECHNISCHE
UNIVERSITÄT
DRESDEN

Optimierung der Suche nach CP-ungeraden Higgs-Bosonen im Zerfall zu Tau-Lepton-Paaren bei ATLAS

Bachelor-Arbeit
zur Erlangung des Hochschulgrades
Bachelor of Science
im Bachelor-Studiengang Physik

vorgelegt von

Valentin Rothert
geboren am 23.11.2000 in Görlitz

Institut für Kern- und Teilchenphysik
Fakultät Physik
Bereich Mathematik und Naturwissenschaften
Technische Universität Dresden
2023

Eingereicht am 05. Juni 2023

1. Gutachter: Prof. Dr. Arno Straessner
2. Gutachter: Prof. Dr. Dominik Stöckinger

Zusammenfassung

Diese Arbeit beschäftigt sich mit der Suche nach dem CP-ungeraden A -Boson, dessen Existenz vom Zwei-Higgs-Dublett-Modell (2HDM) vorhergesagt wird. Dieses Modell ist eine Erweiterung des Standardmodells der Teilchenphysik (SM) und könnte die Abweichung des gemessenen anomalen magnetischen Moments des Myons von der SM-Vorhersage erklären. In der Analyse werden die Daten des ATLAS Detektors des zweiten LHC-Laufs verwendet, um den Signalprozess $A \rightarrow \tau\tau \rightarrow e 2\nu \mu 2\nu$ nachzuweisen. In dieser Arbeit wird untersucht, wie sich das zur Verbesserung der Signifikanz des Signals untersuchte neuronale Netz für die Klassifizierung von Signal- und Untergrundprozessen optimieren lässt. Dafür werden die kinematischen Variablen der Kollisionsprodukte auf Korrelation und Separationskraft untersucht und die verwendeten Variablen sowie verschiedene Netzwerkparameter angepasst. Mit der Überarbeitung des Netzwerks wird in den untersuchten Konfigurationen keine signifikante Verbesserung der bisherigen Leistung erreicht.

Abstract

This thesis presents the search for the CP-odd A -Boson, which is predicted by the two-Higgs-doublet model (2HDM). This model is an extension of the Standard Model of particle physics (SM) and can explain the measured deviation of the anomalous magnetic moment of the muon from the SM prediction. Therefore the Run 2 data of the ATLAS detector is used to search for the signal process $A \rightarrow \tau\tau \rightarrow e 2\nu \mu 2\nu$. The optimization possibilities of the neural Network, which is used for the classification of signal and background processes to improve the data significance, are examined. Therefore, the correlation and separation power of the kinematic variables of particles at collisions are determined and the input variable and network parameters are adapted. With the new configurations, no significant performance improvement is achieved.

Inhaltsverzeichnis

1. Einleitung	1
2. Theoretischer Hintergrund	3
2.1. Das Standardmodell der Teilchenphysik	3
2.1.1. Die elektroschwache Wechselwirkung	4
2.1.2. Der Higgs-Mechanismus	5
2.2. Erweiterungen des Standardmodells	8
2.2.1. Das Zwei-Higgs-Dublett-Modell	9
2.2.2. Das Flavour-aligned Zwei-Higgs-Dublett-Modell	10
2.2.3. Parameter-Beschränkungen des 2HDM	10
3. Der LHC und das ATLAS Experiment	17
3.1. Der Large Hadron Collider	17
3.2. Der ATLAS Detektor	18
3.2.1. Das ATLAS Koordinatensystem	18
3.2.2. Der Innere Detektor	20
3.2.3. Das Kalorimetersystem	21
3.2.4. Das Myonen-System	22
4. Daten und Monte-Carlo-Simulationen	23
5. Analyse	25
5.1. Signal- und Untergrundprozesse	25
5.1.1. Der Signalprozess	25
5.1.2. Untergrundprozesse	26
5.2. Ereignisauswahl	29
5.2.1. Grundlegende Auswahlkriterien	29
5.2.2. Signal- und Validierungsregionen	29
6. Klassifikation von Ereignissen mit neuronalen Netzwerken	31
6.1. Neuronale Netzwerke	31
6.1.1. Neuronen und der Aufbau von Netzwerken	31

6.1.2. Training neuronaler Netzwerke	32
6.1.3. Wichtige Parameter	35
6.1.4. Leistungsmaße	37
6.2. Verwendete Trainingsdaten	39
6.3. Das Referenznetzwerk	40
6.4. Neue Massenhypothesen	43
6.5. Analyse der kinematischen Variablen	47
6.5.1. Separationskraft	48
6.5.2. Korrelation	50
6.6. Training in der MVA-Region	54
6.7. Erhöhung der Netzwerkkapazität	58
6.8. Weitere Untersuchungen	61
7. Zusammenfassung	65
8. Literaturverzeichnis	67
A. Zusätzliche Plots in der MVA-Region	73
B. Tabelle der Separationskraft	77
C. Tabelle der Korrelationskoeffizienten	79

1. Einleitung

Das Standardmodell der Teilchenphysik (SM) ist eine der größten Errungenschaften der modernen Physik. Es enthält alle bisher beobachteten Fermionen sowie die Bosonen der elektromagnetischen, schwachen und der starken Wechselwirkung und ist in der Lage, deren Interaktionen zu beschreiben. Die Grundlage dieses Modells wurde durch die erste Entdeckung eines Elementarteilchens im Jahr 1897 gegeben. In diesem Jahr wies J.J. Thompson die Existenz des Elektrons nach und legte damit den Grundstein für die moderne Teilchenphysik [44]. Über ein Jahrhundert später lässt sich eine beeindruckende Bilanz über die Errungenschaften in diesem Teilgebiet der Physik ziehen. Die Beschreibung der fundamentalen Wechselwirkungen mit Hilfe von Quantenfeldtheorien postulierte die zugehörigen Bosonen, deren Existenz in Kollisionsexperimenten nachgewiesen wurde. Die Einführung des Higgs-Mechanismus durch P. Higgs [33, 34], F. Englert und R. Brout [24] und die damit einhergehende Vorhersage des Higgs-Bosons wurde unter anderem zum Anlass genommen, mit dem Large Hadron Collider (LHC) den bisher größten Teilchenbeschleuniger der Welt zu errichten. Mit dessen Hilfe konnte im Jahr 2012 die Entdeckung des Higgs-Bosons als einer der größten Erfolge der Teilchenphysik erreicht werden [11]. Mit dieser Errungenschaft wurden alle Teilchen des SM nachgewiesen, was der Suche nach weiteren Elementarteilchen dennoch kein Ende setzt. Neben der fehlenden Beschreibung der Gravitation besitzt das SM zum Beispiel ebenfalls keine Erklärung für die dunkle Materie [27], was dessen Unvollständigkeit impliziert.

Weiterhin weist zum Beispiel die Messung des anomalen magnetischen Moments des Myons a_μ auf eine Unvollständigkeit des SM hin: die aktuellen Messwerte weichen signifikant von der Prognose des SM ab [5]. Eine mögliche Erklärung dafür liefert das Zwei-Higgs-Dublett-Modell (2HDM), das 4 weitere Higgs-Bosonen postuliert [46]. Neben einer Erklärung für a_μ würde die Bestätigung dieses Modells die Tür für supersymmetrische Theorien öffnen, die eine „Große vereinheitlichte Theorie“ der im SM enthaltenen Wechselwirkungen oder eine Erklärung der dunklen Materie ermöglichen könnten [7].

Motiviert durch die Abweichung von a_μ beschäftigt sich diese Arbeit mit der Suche nach dem vom 2HDM postulierten A -Boson anhand der Daten aus dem zweiten Lauf des LHCs im Zeitraum von 2015 bis 2018. Der Fokus liegt dabei auf der Optimierung der Nutzung parametrisierter Neuronaler Netzwerke zur Klassifizierung von Untergrund-

und Signalprozessen anhand kinematischer Variablen der Kollisionsprodukte. Diese Netzwerke wurden in der Masterarbeit von Jannik Friese [28] in die Analyse eingeführt und sollen im Rahmen dieser Arbeit auf mögliche Verbesserungen untersucht werden, die aus Veränderung von Trainingsparametern, der Netzwerkarchitektur und der Auswahl der Variablen resultieren können.

In Kapitel 2 dieser Arbeit werden die theoretischen Hintergründe dieser Analyse vorgestellt. Anschließend wird in Kapitel 3 der experimentelle Aufbau des ATLAS Experiments beschrieben. Kapitel 4 erläutert, welche Daten und Monte-Carlo-Simulationen verwendet werden. In Kapitel 5 wird ein Überblick über die relevanten Streuprozesse und die Ereignisauswahl gegeben. Die Ergebnisse der Arbeit mit den neuronalen Netzwerken werden in Kapitel 6 dargelegt und abschließend in Kapitel 7 zusammengefasst.

In dieser Arbeit werden natürliche Einheiten ($c = \hbar = 1$) verwendet.

2. Theoretischer Hintergrund

In diesem Teil der Arbeit werden zunächst die theoretischen Hintergründe der Suche nach dem A -Boson erklärt. Dafür wird ein Überblick über das SM gegeben und die Einführung des Higgs-Bosons über die elektroschwache Symmetriebrechung nachvollzogen. Anschließend wird, anhand des experimentell gemessenen Wertes des anomalen magnetischen Moments des Myons, die dieser Arbeit zu Grunde liegende Erweiterung des SM motiviert und beschrieben. Im letzten Teil dieses Abschnitts werden die Charakteristiken des A -Bosons und der daraus resultierende untersuchte Signalprozess $A \rightarrow \tau\tau$ dargelegt.

2.1. Das Standardmodell der Teilchenphysik

Das Standardmodell der Teilchenphysik ist eine Quantenfeldtheorie (QFT), die auf der Symmetriegruppe $SU(3)_C \times SU(2)_L \times U(1)_Y$ basiert, und deren Untergruppen die drei fundamentalen Wechselwirkungen beschreiben. Die $SU(3)_C$ Symmetrie korrespondiert mit der starken Wechselwirkung, die zwischen den farbgeladenen Quarks wirkt. Die zugehörigen, masselosen Eichbosonen werden als Gluonen g bezeichnet. Aus Superposition der Eichbosonen der $SU(2)_L \times U(1)_Y$ Symmetriegruppe resultieren die physikalischen Bosonen Z^0 , W^\pm als Trägerteilchen der schwachen, sowie das Photon γ als das der elektromagnetischen Wechselwirkung (siehe Unterabschnitt 2.1.1). Die zugehörigen Ladungen sind der schwache Isospin und die schwache Hyperladung. Neben den Bosonen der Eichsymmetrien mit Spin 1 enthält das SM die Fermionen mit Spin 1/2, welche die Materie bilden. Diese werden in Leptonen l und die stark wechselwirkenden Quarks q unterteilt und bestehen jeweils aus drei Teilchengenerationen, dargestellt in Abbildung 2.1. Jedes dieser Fermionen besitzt ein entsprechendes Antiteilchen mit gleicher Masse und entgegengesetzter elektrischer Ladung [17]. Vervollständigt wird das SM durch das skalare Higgs-Boson, dessen Existenz 2012 im Large Hadron Collider (LHC) am CERN nachgewiesen wurde [11].

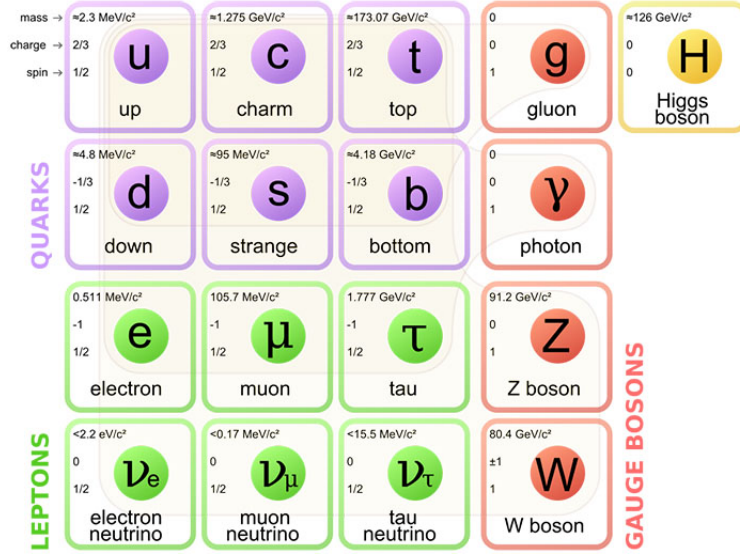


Abbildung 2.1.: Überblick über die im Standardmodell der Teilchenphysik enthaltenen Teilchen sowie deren Spin, Masse und elektrische Ladung [3].

2.1.1. Die elektroschwache Wechselwirkung

Die mit der $SU(2)_L \times U(1)_Y$ korrespondierende elektroschwache Wechselwirkung führt den schwachen Isospin I_W und die schwache Hyperladung [38]

$$Y = Q - I_W^3 \quad (2.1)$$

ein. Q ist hierbei die elektrische Ladung.

Die Fermionen des SM besitzen einen linkshändigen und einen rechtshändigen Anteil Ψ_L und Ψ_R , definiert gemäß

$$\Psi_L = \left[\frac{1}{2}(1 + \gamma_5) \right] \Psi \quad \Psi_R = \left[\frac{1}{2}(1 - \gamma_5) \right] \Psi. \quad (2.2)$$

Die linkshändigen Teilchen besitzen einen Isospin von $I_W = \frac{1}{2}$, dessen dritte Komponente I_W^3 die Werte $\pm\frac{1}{2}$ annehmen können und werden daher als Isospin-Dubletts aus Lepton- oder Quark-Paaren beschrieben. Deren obere Komponente repräsentiert eine positive, die untere eine negative Ausrichtung der dritten Spinkomponente. Im Fall der Quarks besteht ein solches Dublett $Q = (u, d)_L^T$ aus einem Up-Type Quark u und einem Down-Type Quark d , im Fall eines Lepton-Dubletts $L = (l, \nu_l)_L^T$ aus einem Lepton l und einem Neutrino ν_l [47]. Eine Eichtransformation in diesem zweidimensionalen Spin-Raum entspricht daher einer Rotation des Dubletts, was sich durch dessen Multiplikation mit einer 2×2 Matrix ausdrücken lässt. Die rechtshändigen Teilchen haben einen Isospin

von $I_W = 0$ und koppeln daher nicht an die Bosonen der schwachen Wechselwirkung und werden als Lepton- l_R oder Quark-Singlett q_R beschrieben. Beide Teilchen unterliegen aber der mit der $U(1)_Y$ korrespondierenden Wechselwirkung und lassen sich daher lokal eichtransformieren

$$L \rightarrow L' = e^{i\alpha(x)Y_L + i\sum_{j=1}^3 \theta_j(x)T^j} L \quad (2.3)$$

$$R \rightarrow R' = e^{i\alpha(x)Y_R} R, \quad (2.4)$$

wobei die $T^i = \frac{1}{2}\sigma_i$ die über die Pauli-Matrizen σ_i definierten Generatoren der zweidimensionalen Rotation sind [48]:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (2.5)$$

Damit die Lagrangedichte bezüglich dieser Transformation invariant ist, muss eine kovariante Ableitung

$$D_\mu L = \left(\partial_\mu + ig_W \sum_{j=1}^3 T^j \cdot W_\mu^j + ig_Y \frac{Y_L}{2} B_\mu \right) L \quad (2.6)$$

$$D_\mu R = \left(\partial_\mu + ig_Y \frac{Y_R}{2} B_\mu \right) R \quad (2.7)$$

eingeführt werden, in der die drei Eichbosonen W_μ^i der $SU(2)_L$ und das Eichboson B_μ der $U(1)_Y$ definiert werden. Die Kopplungskonstanten der jeweiligen Wechselwirkung werden mit g_W und g_Y bezeichnet [47].

2.1.2. Der Higgs-Mechanismus

Um die Massen der Eichbosonen und der Fermionen unter Erhalt der Eichsymmetrien zu erklären, wird der Higgs-Mechanismus in den elektroschwachen Sektor des SM eingebettet. Dieser Mechanismus beschreibt die spontane Symmetriebrechung eines skalaren Feldes unter Anwesenheit eines Potentials. In diesem Fall resultieren daraus die Masseneigenzustände W^\pm , Z^0 und A^μ der Eichbosonen und das skalare Higgs-Boson.

Zuerst wird dafür ein Higgs-Isospin-Dublett

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \quad (2.8)$$

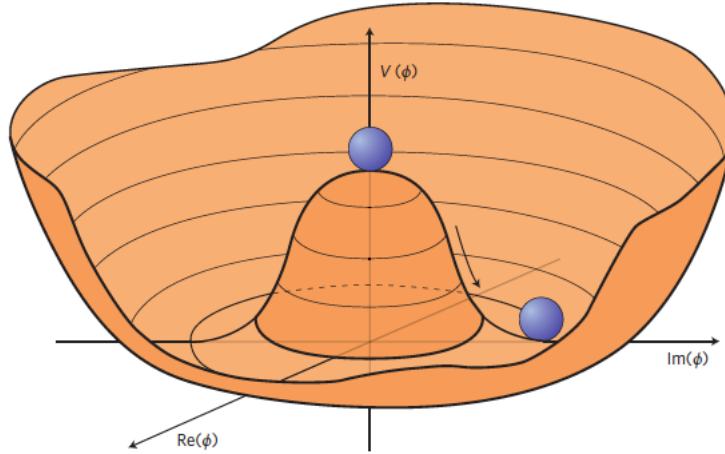


Abbildung 2.2.: Das Higgspotential für den Fall $\mu^2 < 0$. Das Higgsfeld nimmt einen neuen Grundzustand in der Potentialmulde ein.[23]

mit $Y = 1$ eingeführt, das aus zwei komplexen Skalarfeldern besteht [47]. Dabei ist das obere Feld elektrisch geladen, während das untere elektrisch neutral ist. Die Lagrange-dichte für dieses Dublett lautet

$$\mathcal{L}_{\text{Higgs}} = (\partial_\mu \phi)^\dagger (\partial^\mu \phi) - V(\phi) \quad (2.9)$$

mit dem Higgs-Potential

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2 \quad (2.10)$$

mit $\lambda > 0$ aufgrund der Vakuumstabilität [32]. Für $\mu^2 > 0$ gibt es nur ein globales Minimum bei $\phi = 0$, was in einem verschwindenden Higgs-Feld resultiert. In diesem Fall würden keine Massenterme für die Eichbosonen generiert werden. Im Fall $\mu^2 < 0$ bildet das Potential eine rotationssymmetrische Mulde aus, welche in Abbildung 2.2 dargestellt ist.

Das Feld ϕ nimmt ein beliebiges Minimum in der Mulde als Grundzustand ein und bricht damit spontan die Rotationssymmetrie [33]. Um im Folgenden ein masseloses Photon zu erhalten, wird der Grundzustand ohne Beschränkung der Allgemeinheit als

$$\langle 0 | \phi | 0 \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \quad (2.11)$$

gewählt und das Dublett in der unitären Eichung linear expandiert [35]:

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}. \quad (2.12)$$

Die Massenterme der Eichbosonen lassen sich nun aus dem kinematischen Term $(D_\mu\phi)^\dagger(D^\mu\phi)$ der Lagrangedichte bestimmen. Dabei wird die kovariante Ableitung der $SU(2)_L \times U(1)_Y$ Symmetriegruppe aus Gleichung 2.6 benutzt und für die Lagrangedichte ergibt sich

$$\begin{aligned}\mathcal{L}_{\text{Higgs}} &= (D_\mu\phi)^\dagger(D^\mu\phi) - V(\phi) \\ &= \frac{1}{2}(\partial_\mu h(x))^\dagger(\partial^\mu h(x)) + \frac{1}{8}g_W^2(W_\mu^1 + iW_\mu^2)(W^{1\mu} - iW^{2\mu})(v + h(x))^2 \\ &\quad + \frac{1}{8}(g_W W_\mu^3 - g_Y B_\mu)(g_W W^{3\mu} - g_Y B^\mu)(v + h(x))^2 - V(\phi).\end{aligned}\quad (2.13)$$

Führt man folgende Substitutionen ein und definiert damit die im SM enthaltenen Masseneigenzustände der Bosonen

$$W^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2) \quad (2.14)$$

$$\begin{pmatrix} A^\mu \\ Z^\mu \end{pmatrix} = \begin{pmatrix} \cos\Theta_W & \sin\Theta_W \\ -\sin\Theta_W & \cos\Theta_W \end{pmatrix} \begin{pmatrix} B^\mu \\ W_3^\mu \end{pmatrix} \quad (2.15)$$

$$\sin\Theta_W = \frac{g_Y}{\sqrt{g_W^2 + g_Y^2}} \quad \cos\Theta_W = \frac{g_W}{\sqrt{g_W^2 + g_Y^2}}, \quad (2.16)$$

lässt sich das Resultat folgendermaßen umschreiben:

$$\begin{aligned}\mathcal{L}_{\text{Higgs}} &= \frac{1}{2}(\partial_\mu h(x))^\dagger(\partial^\mu h(x)) + \frac{1}{4}g_W^2 W_\mu^+ W^{-\mu}(v + h(x))^2 \\ &\quad + \frac{1}{8}(g_W^2 + g_Y^2)Z_\mu Z^\mu(v + h(x))^2 - V(\phi).\end{aligned}\quad (2.17)$$

Die in den quadratischen Termen auftauchenden Vorfaktoren werden als Massenterme der Form $\frac{1}{2}m^2 X_\mu X^\mu$ interpretiert [34], daher lassen sich die Massen der Bosonen aus der Lagrangedichte ablesen, wobei das Photon aufgrund des fehlenden Terms masselos bleibt [47]:

$$\begin{aligned}m_W &= \frac{1}{2}g_W v \\ m_Z &= \frac{1}{2}v\sqrt{g_W^2 + g_Y^2} = \frac{1}{2}v\frac{g_W}{\cos\Theta_W} \\ m_H &= \sqrt{2}\mu \\ m_\gamma &= 0.\end{aligned}\quad (2.18)$$

Die Massen der Fermionen werden über zusätzliche Yukawa-Kopplungs-Terme der Form

$$\mathcal{L}_{\text{Yuk}} = -\lambda_e (\bar{L}\phi e_R + \bar{e}_R\phi^\dagger L) - \lambda_u (\bar{Q}\phi u_R + \bar{u}_R\phi^\dagger Q) - \lambda_d (\bar{Q}\phi_c d_R + \bar{d}_R\phi_c^\dagger Q) \quad (2.19)$$

(hier nur für eine Teilchengeneration) in der Lagrangedichte generiert und resultieren mit dem Higgs-Feld im neuen Grundzustand aus Gleichung 2.12 zu

$$\mathcal{L}_{\text{Yuk}} = -\frac{v + h(x)}{\sqrt{2}} \left(\lambda_e (\bar{e}_L e_R + \bar{e}_R e_L) + \lambda_u (\bar{u}_L u_R + \bar{u}_R u_L) + \lambda_d (\bar{d}_L d_R + \bar{d}_R d_L) \right) \quad (2.20)$$

wobei die Neutrinos nicht berücksichtigt werden. Für die Masse eines Fermions f ergibt sich daher [45]

$$m_f = \frac{v\lambda_f}{\sqrt{2}}. \quad (2.21)$$

2.2. Erweiterungen des Standardmodells

Trotz der vielen über Messungen bestätigten Prognosen des Standardmodells, wie zum Beispiel der Messung der Z -Resonanz im $e^+e^- \rightarrow q\bar{q}$ Wirkungsquerschnitt am Large Electron-Positron Collider (LEP) [6], kann das SM bei weitem nicht alle beobachteten Phänomene im Universum erklären, was eine Erweiterung oder sogar eine komplette Umformulierung dieser Theorie erzwingt.

Dazu gehört zum einen die Gravitation als vierte fundamentale Wechselwirkung, deren Effekte zwar erst bei Energien in Höhe der Planck-Skala $\Lambda_P \sim 10^{19}$ GeV relevant werden, im Rahmen dieser Arbeit also vernachlässigbar sind, die jedoch in einer allgemeingültigen Theorie enthalten sein muss. Ebenso wird die Existenz der dunklen Materie nicht durch das SM prognostiziert. Als Beweis für deren Existenz wird unter anderem die Abweichung der gemessenen radialen Geschwindigkeitsverteilungen von Sternen in Galaxien von der theoretischen Vorhersage gesehen. Das deutet auf die Existenz weiterer Teilchen hin, die nur über die Gravitation oder schwach mit der sichtbaren Materie wechselwirken [45]. Die in dieser Arbeit thematisierte Erweiterung des SM wird durch die Abweichung des gemessenen anomalen magnetischen Moments des Myons

$$a_\mu = \frac{(g-2)_\mu}{2} \quad (2.22)$$

von der Vorhersage des SM motiviert. Gemittelt über die Messwerte des Fermi National Accelerator Laboratory (FNAL) und des Brookhaven National Laboratory (BNL) resultiert für die Abweichung

$$a_\mu(\text{Exp}) - a_\mu(\text{SM}) = (251 \pm 59) \times 10^{-11} \quad (2.23)$$

mit einer Signifikanz von 4.2σ [5]. Die Resultate der Messungen sind in Abbildung 2.3 dargestellt. Eine mögliche Erklärung dieser Abweichung liefert das Zwei-Higgs-Dublett-Modell, welches im folgenden Abschnitt zusammengefasst wird.

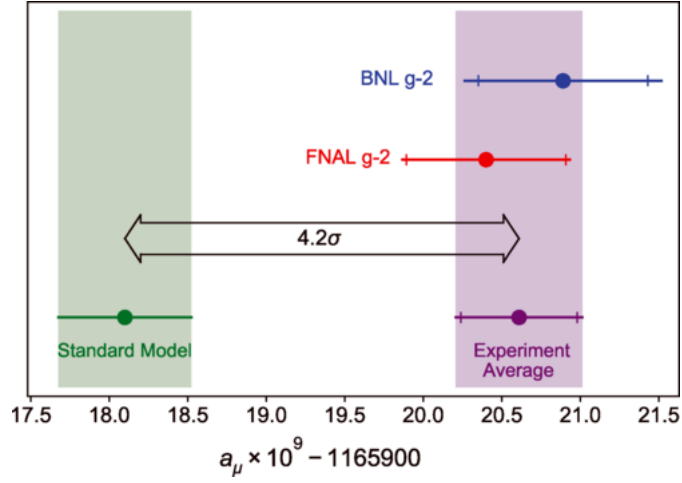


Abbildung 2.3.: Die Messwerte für a_μ des FNAL und BNL und die Vorhersage des SM inklusive der statistischen Unsicherheiten. Der gemittelte experimentelle Wert weicht um 4.2σ vom SM ab [5]

2.2.1. Das Zwei-Higgs-Dublett-Modell

Das Zwei-Higgs-Dublett-Modell (2HDM) ist eine der einfachsten Erweiterungen des SM [46]. Statt wie in Unterabschnitt 2.1.2 beschrieben werden nicht ein, sondern zwei Higgs-Dubletts aus komplexen Skalarfeldern in ein neues eichinvariantes skalares Potential

$$\begin{aligned}
 V = & m_{11}^2 \phi_1^\dagger \phi_1 + m_{22}^2 \phi_2^\dagger \phi_2 - [m_{12}^2 \phi_1^\dagger \phi_2 + h.c.] \\
 & + \frac{1}{2} \lambda_1 (\phi_1^\dagger \phi_1)^2 + \frac{1}{2} \lambda_2 (\phi_2^\dagger \phi_2)^2 + \lambda_3 (\phi_1^\dagger \phi_1)(\phi_2^\dagger \phi_2) + \lambda_4 (\phi_1^\dagger \phi_2)(\phi_2^\dagger \phi_1) \\
 & + \left\{ \frac{1}{2} \lambda_5 (\phi_1^\dagger \phi_2)^2 + [\lambda_6 (\phi_1^\dagger \phi_1) + \lambda_7 (\phi_2^\dagger \phi_2)] \phi_1^\dagger \phi_2 + h.c. \right\}
 \end{aligned} \quad (2.24)$$

mit den freien Parametern m_{ij} und λ_l eingeführt, in dem die Dubletts die Grundzustände

$$\langle 0 | \phi_1 | 0 \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v_1 \end{pmatrix} \quad \langle 0 | \phi_2 | 0 \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v_2 \end{pmatrix} \quad (2.25)$$

mit den Vakuumerwartungswerten (VEV) v_1 und v_2 einnehmen. Mit folgenden neuen Definitionen

$$\tan \beta = \frac{v_2}{v_1} \quad \begin{pmatrix} H_1 \\ H_2 \end{pmatrix} = \begin{pmatrix} \cos \beta & \sin \beta \\ -\sin \beta & \cos \beta \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} \quad (2.26)$$

nimmt H_1 den VEV $v = \sqrt{v_1^2 + v_2^2}$ an und der des H_2 Dubletts wird 0 [31].

Das zweite Dublett korrespondiert mit dem Higgsfeld A und den geladenen Feldern H^\pm .

Das A -Feld besitzt unter der CP-Transformation

$$P\Psi(x, y, z) = \Psi(-x, -y, -z) \quad C|\Psi\rangle = |\bar{\Psi}\rangle \quad (2.27)$$

mit dem Paritätsoperator P und dem Ladungs-Konjugations-Operator C den Eigenwert -1 und wird als CP-gerade bezeichnet. Aus der Mischung der zwei Dubletts um einen Winkel $(\alpha - \beta)$ resultieren die CP-geraden Higgsfelder H und h mit CP-Eigenwert $+1$, wobei h üblicherweise mit dem Higgsfeld des Standardmodells identifiziert wird [19].

2.2.2. Das Flavour-aligned Zwei-Higgs-Dublett-Modell

Im Flavour-aligned 2HDM werden die Kopplungen der Higgsfelder an die verschiedenen Fermion-Typen in neuen Yukawa-Termen beschrieben. Die Kopplungen von h entsprechen dabei denen des SM in Gleichung 2.19. Die Kopplungskonstanten der anderen 4 Felder werden über die Proportionalität zur SM-Kopplung durch Proportionalitätskonstanten ζ_l für geladene Leptonen, ζ_u für Up- und ζ_d für Down-Typ Quarks beschrieben. Das allgemeine Flavour-aligned 2HDM enthält die Modelle I, II, X und Y als Spezialfälle, deren Kopplungs-Konfigurationen in Tabelle 2.1 zu finden sind [19, 42].

Tabelle 2.1.: Relation der Kopplungskonstanten zwischen den Modellen I, II, X und Y und dem generellen Flavour-aligned 2HDM [42]

	Typ I	Typ II	Typ X	Typ Y
ζ_u	$\cot\beta$	$\cot\beta$	$\cot\beta$	$\cot\beta$
ζ_d	$\cot\beta$	$-\tan\beta$	$\cot\beta$	$-\tan\beta$
ζ_l	$\cot\beta$	$-\tan\beta$	$-\tan\beta$	$\cot\beta$

2.2.3. Parameter-Beschränkungen des 2HDM

Den größten Einfluss auf die mögliche Erklärung von a_μ durch dieses Modell haben die Parameter ζ_l , ζ_u und m_A , die Masse des A -Bosons. Deshalb werden im Folgenden die Beschränkungen der Wertebereiche sowie deren Ursachen zusammengefasst.

Da die vier neuen Felder zu Prozessen beitragen, deren Charakteristiken bereits mit hoher Präzision gemessen wurden, dürfen deren Beiträge zu den entsprechenden Matrixelementen die Prognosen des SM nur in einem geringen Maß verändern, sodass diese noch im Unsicherheitsbereich der Messungen liegen. Im Fall der Kopplungstärke ζ_l beschränkt der am LEP gemessene Prozess $ee \rightarrow \tau\tau \rightarrow \tau\tau(\tau\tau)$, gezeigt in Abbildung 2.4, den Betrag

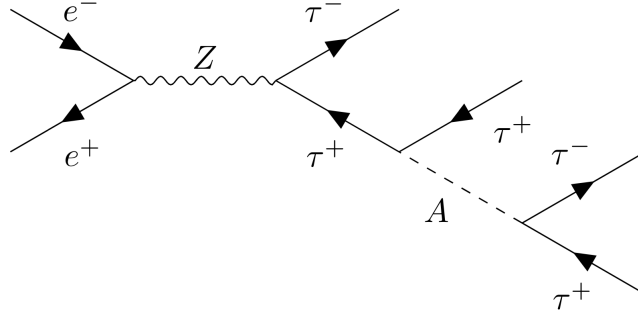


Abbildung 2.4.: Feynman-Diagramm des Prozesses $ee \rightarrow \tau\tau(A) \rightarrow \tau\tau(\tau\tau)$, der $|\zeta_l|$ primär für kleine Massen $m_A < 20$ GeV beschränkt [37].

von ζ_l für kleine Massen $m_A < 20$ GeV [21]. Für höhere Massen von A würden zu hohe Werte von $|\zeta_l|$ den Messungen des Zerfalls von τ -Leptonen im Prozess $\tau \rightarrow \mu\nu_\tau\bar{\nu}_\mu$ und des leptonischen Z -Boson-Zerfalls widersprechen. Die daraus resultierenden maximalen Werte für $|\zeta_l|$ sind in Abbildung 2.5 dargestellt. Man sieht für kleine m_A den Einfluss des am LEP gemessenen Prozesses. Für höhere Werte für m_A und $m_H = m_{H^\pm}$ dominiert die Beschränkung durch τ -Zerfälle. Hohe Werte für $m_H = m_{H^\pm}$ lockern die Beschränkungen etwas, für Werte ab 250 GeV werden aber die Einschränkungen durch Z -Zerfälle relevant [19].

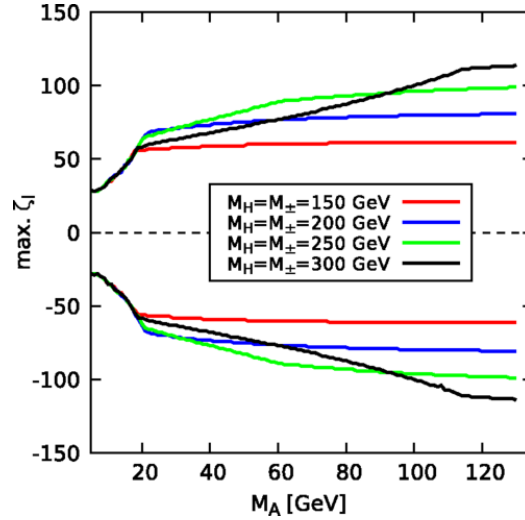


Abbildung 2.5.: Maximal mögliche Werte für $|\zeta_l|$ in Abhängigkeit der Massen m_A , m_H und m_{H^\pm} . Die Beschränkungen folgen aus τ - und Z -Zerfällen sowie aus Messwerten des Prozesses $ee \rightarrow \tau\tau \rightarrow \tau\tau(\tau\tau)$ [19].

Beschränkungen für ζ_u folgen zum einen aus der B -Physik, der Physik der B -Hadronen

mit mindestens einem Bottom-Quark b . Die relevantesten Prozesse bezüglich dieser Analyse sind in Abbildung 2.6 dargestellt. Die zugehörigen Matrixelemente sind abhängig von den Kopplungskonstanten für Leptonen, Up- und Down-Typ Quarks.

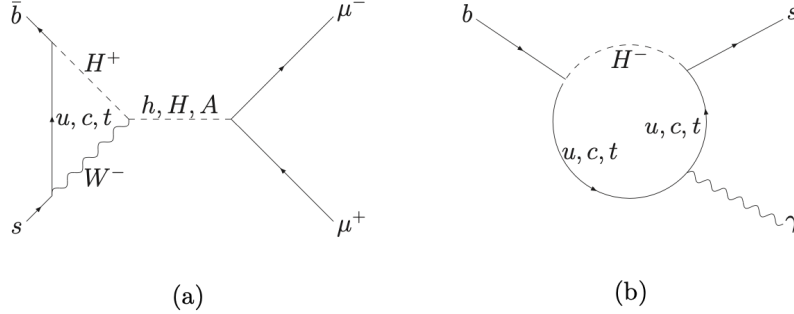


Abbildung 2.6.: Feynman-Diagramme der Prozesse (a) $B_s \rightarrow \mu^+ \mu^-$ und (b) $B_s \rightarrow \gamma$, die von den Kopplungskonstanten ζ_u , ζ_d und ζ_l abhängen [19].

Auf der anderen Seite wird ζ_u durch die LHC-Higgs-Physik beschränkt. Wichtige Prozesse hierbei sind $pp \rightarrow H \rightarrow \tau\tau$ und der in dieser Arbeit relevante Signalprozess $pp \rightarrow A \rightarrow \tau\tau$, dargestellt in Abbildung 2.7, in denen die Higgs-Bosonen durch Gluon-Fusion über eine Top-Quark Schleife erzeugt werden. Bei dem erstgenannten Prozess unterscheiden sich die Einschränkungen stark dadurch, dass der Zerfall $H \rightarrow AA$ für $m_A < \frac{m_H}{2}$ kinematisch erlaubt, sonst aber verboten ist. Die Analyse dieser dreifachen Higgs-Kopplung resultiert in einer geringen Beschränkung von ζ_u für $\frac{m_h}{2} < m_A < m_Z$ [19] mit der SM-Higgs-Masse $m_h = 126$ GeV [11] und der Z -Boson-Masse $m_Z = 91.2$ GeV [41].

Aufgrund des Einflusses der Parameter ζ_u , ζ_l sowie m_A und m_{H^\pm} auf die Beiträge dieser Prozesse, lässt sich der maximale Wert für ζ_u in Abhängigkeit dieser Parameter bestimmen.

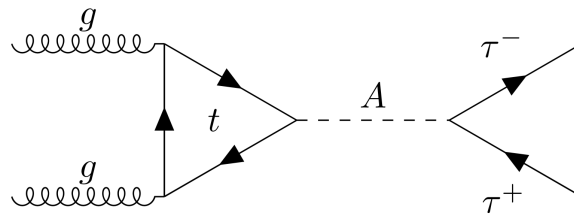


Abbildung 2.7.: Feynman-Diagramm des Prozesses $pp \rightarrow A \rightarrow \tau\tau$ [37].

In Abbildung 2.8 sind die Ergebnisse dieser Analyse zusammengefasst. Zu sehen sind die oberen Grenzen von ζ_u in Abhängigkeit von m_A , dargestellt für verschiedene Konfigurationen der Parameter ζ_l , m_H und m_{H^\pm} . Die durchgezogenen Linien zeigen die Be-

schränkungen der B -Physik, die maximale Werte für ζ_u zwischen 0.3 und 0.6 erlaubt. Die gestrichelten Linien markieren die Grenzen der LHC-Higgs Physik. Für $m_H = 150$ GeV wird ζ_u stark von $pp \rightarrow H \rightarrow \tau\tau$ für $\frac{m_H}{2} < m_A$ beschränkt. Bei höheren Werten für m_H resultiert das Maximum für $60 \text{ GeV} < m_A < 90 \text{ GeV}$ aus der dreifachen Higgs-Kopplung. Ab $m_A > 100$ GeV bestimmt $pp \rightarrow A \rightarrow \tau\tau$ die obere Grenze für ζ_u auf ca. 0.2 [19].

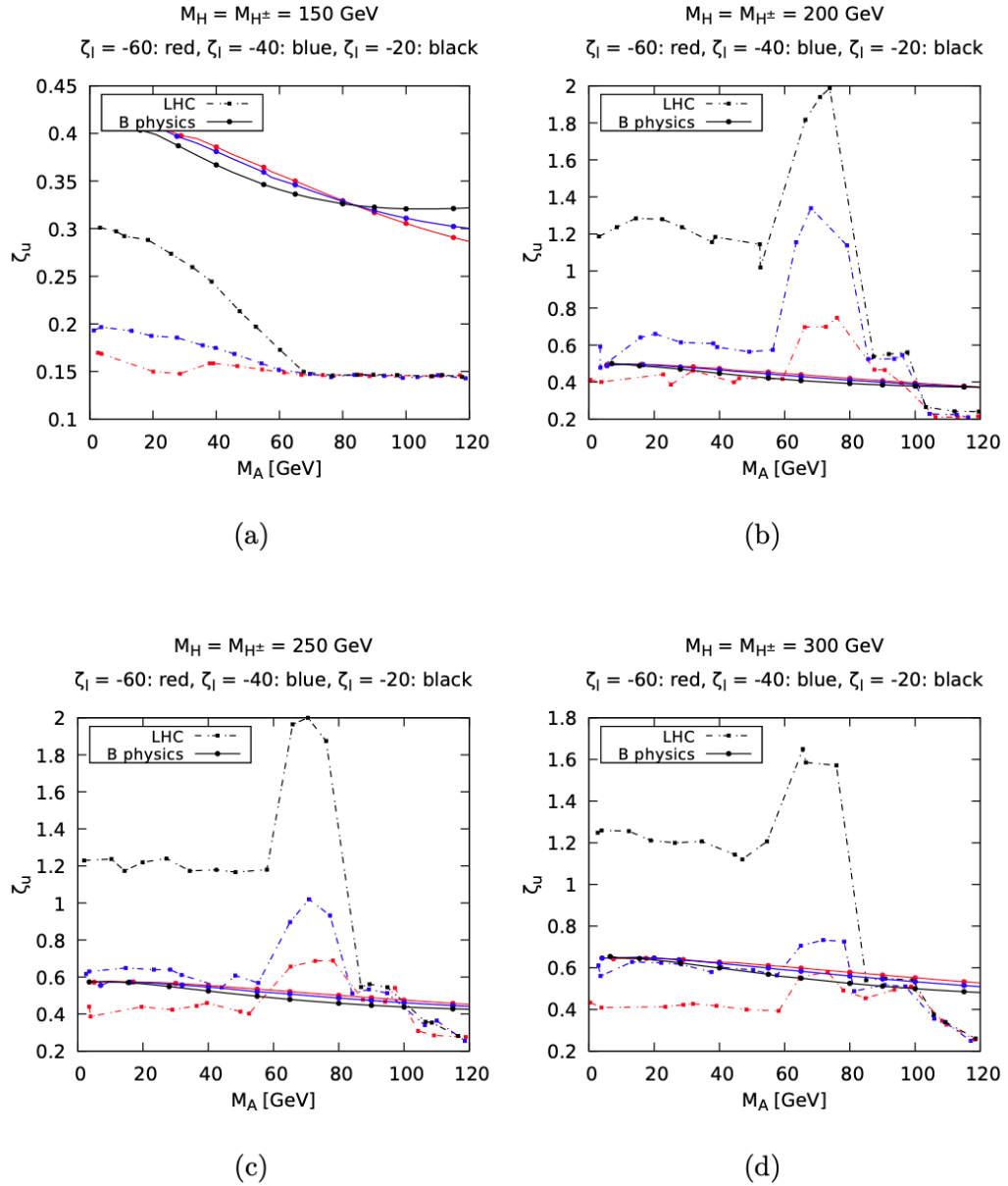


Abbildung 2.8.: Maximal mögliche Werte für ζ_u in Abhängigkeit der Masse m_A für verschiedene Konfigurationen der Parameter ζ_l , m_H und m_{H^\pm} . Die Beschränkungen der B -Physik werden mit durchgezogenen Linien dargestellt, die der LHC-Higgs-Physik mit gestrichelten Linien [19].

Eine detailliertere Analyse der Beschränkungen weiterer Parameter des 2HDM können in Quelle [19] nachgelesen werden. Die Ergebnisse dieser Publikation sind in Abbildung 2.9 zusammengefasst. Das Diagramm zeigt den maximal möglichen Wert für Δa_μ in Abhängigkeit der Masse m_A mit der Kopplungskonstante $\zeta_l = -60$. Der schwarze, blaue und rote Graph entsprechen jeweils dem Wert 150 GeV, 200 GeV und 250 GeV für m_{H, H^\pm} . Das gelbe Band stellt den Wert für Δa_μ des BNL sowie dessen 1σ Umgebung dar, das grüne Band den neuen Mittelwert aller Messungen.

Die Auswertung in Quelle [19] zeigt, dass eine hohe Kopplung an geladene Leptonen sowie Massen m_A unterhalb der Z -Masse ein hohes Δa_μ ermöglichen, weshalb für $|\zeta_l|$ ein möglichst hoher Wert gewählt wurde. Für $m_A \sim 20$ GeV hat jede Verteilung ein stark ausgeprägtes Maximum. Der starke Abfall für niedrigere Werte von m_A resultiert aus einem minimal möglichen Wert von m_A für ein bestimmtes ζ_l (siehe Abbildung 2.5). Für höhere Werte wird Δa_μ mit $\frac{1}{m_A^2}$ unterdrückt. Anhand des Plots in Abbildung 2.9 lassen sich zwei Parameterbereiche für m_H und m_{H^\pm} unterscheiden: niedrige Werte $m_H < 150$ GeV und hohe Werte $m_H > 150$ GeV, die höhere Werte für ζ_u zulassen, was ebenfalls notwendig ist, um a_μ erklären zu können [19].

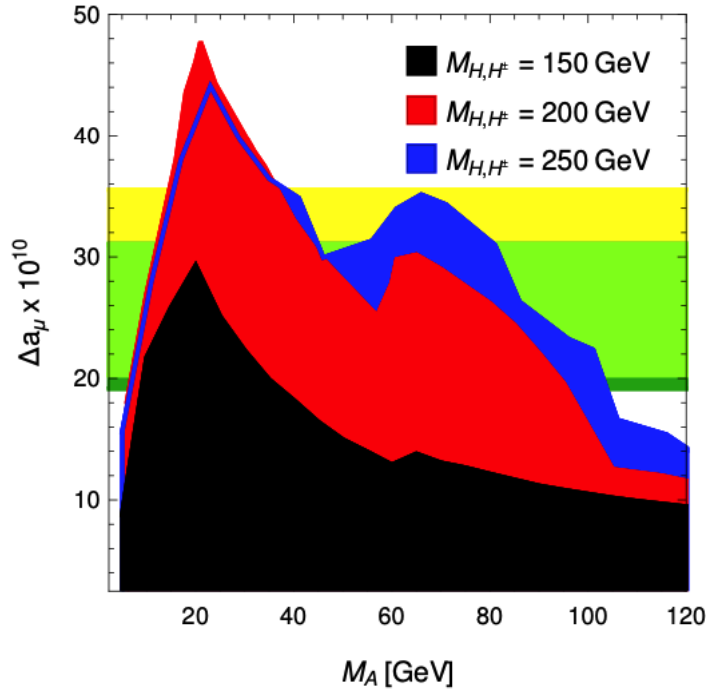


Abbildung 2.9.: Die im generellen Flavour-aligned 2HDM maximal möglichen Werte für Δa_μ in Abhängigkeit von m_A für verschiedene Werte von $m_H = m_{H^\pm}$. Der gelbe Bereich markiert die Abweichung von a_μ inklusive des 1σ Bandes. Für jede Konfiguration von Parametern wurde ζ_u maximiert[8]

Anhand dieser Ergebnisse lassen sich die Modelle II und X als mögliche Erklärung für die a_μ Abweichung ausschließen, da $\zeta_u = \cot\beta$ und $\zeta_l = -\tan\beta$ nicht gleichzeitig groß sein können. Die Messungen des B -Physik-Prozesses $b \rightarrow s\gamma$ schränken das Modell Y stark ein, da das Produkt $|\zeta_u\zeta_d| = \cot\beta \tan\beta = 1$ nicht klein sein kann. Im Fall von Modell I wäre das zwar möglich, würde aber mit kleinen Werten für ζ_l einhergehen, was allerdings das maximal mögliche a_μ zu stark einschränken würde. Als mögliche Erklärung bleibt ein allgemeines Flavour-aligned 2HDM mit voneinander unabhängigen Kopplungsparametern und möglichst hohen Werten für ζ_l und ζ_u [19].

3. Der LHC und das ATLAS Experiment

In diesem Kapitel wird der experimentelle Aufbau dieser Analyse beschrieben. Die in dieser Arbeit genutzten Daten wurden mit dem ATLAS Detektor (**A** Toroidal LHC **A**pparatu**S**) aufgenommen, welcher im Large Hadron Collider eingegliedert ist.

3.1. Der Large Hadron Collider

Der LHC ist ein in einen 26,7 km langen Tunnel eingebauter Ring-Beschleuniger, der zum CERN bei Genf gehört. Er besteht aus zwei Ringen, in denen in entgegengesetzter Richtung Protonen oder Blei-Ionen zirkulieren. Die Protonen werden auf über 99% der Lichtgeschwindigkeit beschleunigt und treffen in einem der vier Detektoren ATLAS, CMS, ALICE oder LHCb mit einer Schwerpunktsenergie von bis zu $\sqrt{s} = 13,6$ TeV aufeinander. Um solche hohen Energien zu ermöglichen, werden die Teilchen in anderen Beschleunigern vorbeschleunigt, bevor sie in den LHC geleitet werden. Zusammen bilden diese miteinander verbundenen Beschleuniger den CERN Beschleuniger-Komplex, der in Abbildung 3.1 abgebildet ist [25].

Um die Protonen auf den gekrümmten Bahnen zu halten, werden 1232 supraleitende Dipolmagnete verwendet, die mit supraflüssigem Helium auf 1.9 K herabgekühlt werden. Damit lassen sich Magnetfelder mit einer Stärke von bis zu 8.3 T erzeugen[18].

Die Protonen zirkulieren in $n_b = 2808$ Paketen aus je $N_b = 1.1 \times 10^{11}$ Protonen pro Paket. Mit der Umlauffrequenz $f_{\text{rev}} = 11.245$ kHz, dem relativistischen Gammafaktor γ_r , der normalisierten transversalen Strahlemission ϵ_n , dem Wert β^* der beta-Funktion am Kollisionspunkt und dem geometrischen Luminositäts-Reduktionsfaktor F lässt sich die Luminosität

$$L = \frac{N_b^2 n_b f_{\text{rev}} \gamma_r}{4\pi \epsilon_n \beta^*} F \quad (3.1)$$

berechnen. Für die mittlere Anzahl der Ereignisse pro Sekunde $N_{\text{Ereignisse}}$ lässt sich mit dem differentiellen Wirkungsquerschnitt σ_{Ereignis} folgende Relation aufstellen [25]:

$$N_{\text{Ereignisse}} = L \sigma_{\text{Ereignis}} \quad (3.2)$$

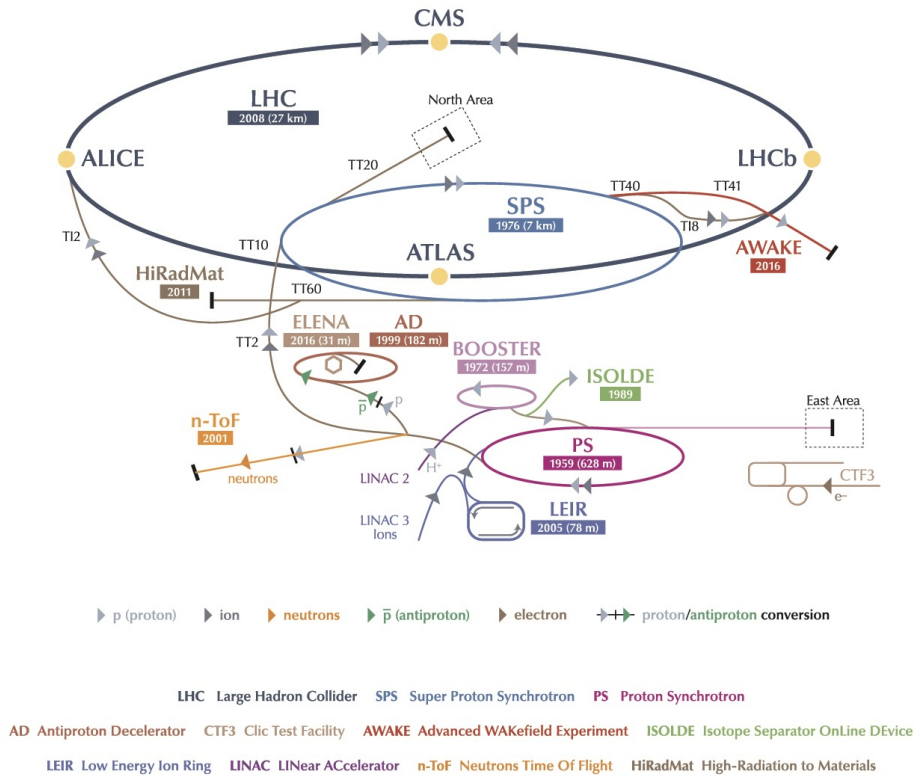


Abbildung 3.1.: Der Beschleunigerkomplex am CERN [20].

3.2. Der ATLAS Detektor

Der zylindrische ATLAS Detektor ist mit einer Länge von 44 m und einem Durchmesser von 25 m der größte Kollisionsdetektor, der bisher gebaut wurde. Er besteht aus mehreren Detektorschichten, die unterschiedliche physikalische Eigenschaften der Kollisionsprodukte analysieren. Eine schematische Darstellung ist in Abbildung 3.2 zu sehen. Da die entstehenden Datenmengen bei den hohen Kollisionsfrequenzen zu groß sind, um schnell genug ausgelesen zu werden, wird im Detektor ein Triggersystem genutzt, das in Echtzeit die Daten nach Ereignissen mit interessanten Charakteristiken filtert. Genauere Informationen zum Triggersystem können in Quelle [9] nachgelesen werden.

3.2.1. Das ATLAS Koordinatensystem

Der Ursprung des Koordinatensystems im ATLAS liegt im Kollisionspunkt im Zentrum des Detektors. Die z-Achse zeigt entlang der Strahlachse, die x-Achse in Richtung des Zentrums des LHC und die y-Achse in Richtung der Erdoberfläche. Die x-y-Ebene wird

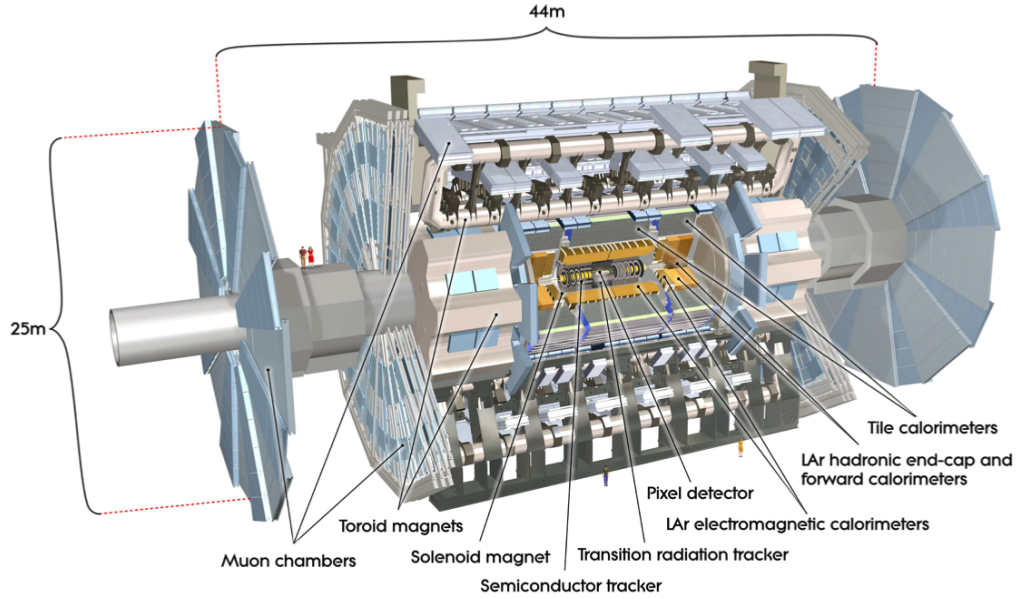


Abbildung 3.2.: Schematischer Aufbau des ATLAS Detektors [9]

als transversale Ebene bezeichnet, da sie orthogonal zur Strahlachse liegt. Die Koordinaten eines Teilchens werden mit den sphärischen Koordinaten (r, ϕ, θ) beschrieben. Der Radius r ist definiert als $r = \sqrt{x^2 + y^2 + z^2}$. Der Winkel θ wird von der Strahlachse aus gemessen, allerdings wird in Kollisionsexperimenten an Stelle des Winkels die Pseudorapidität η

$$\eta = -\ln \tan \frac{\theta}{2} \quad (3.3)$$

genutzt, da die Größe $\Delta\eta$ lorentzinvariant ist. Der azimutale Winkel ϕ mit $\tan \phi = \frac{x}{y}$ ist der von der x-Achse aus gemessene Winkel in der transversalen Ebene. Aus diesen beiden Größen lässt sich das Winkelement $R = \sqrt{\phi^2 + \eta^2}$, sowie die Winkelseparation

$$\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} = \sqrt{(\phi_1 - \phi_2)^2 + (\eta_1 - \eta_2)^2} \quad (3.4)$$

zwischen zwei detektierten Objekten berechnen [9].

In der transversalen Ebene lässt sich die Impulserhaltung anwenden, da sich die einfallenden Teilchen entlang der z-Achse bewegen. Deswegen wird der transversale Impuls

$$p_T = \sqrt{p_x^2 + p_y^2} \quad (3.5)$$

definiert. Da Neutrinos keine Signale im Detektor hinterlassen, aber der Impulserhaltung unterliegen, wird deren Entstehung während eines Ereignisses über den fehlenden

transversalen Impuls

$$p_{\text{T}}^{\text{miss}} = - \sum \vec{p}_{\text{T}} \quad (3.6)$$

nachgewiesen. In dieser Analyse wird unter der Annahme, dass die Neutrinos hoch relativistisch sind, d.h. $p_{\nu} \gg m_{\nu}$, nur die fehlende transversale Energie

$$E_{\text{T}}^{\text{miss}} = |p_{\text{T}}^{\text{miss}}| \quad (3.7)$$

verwendet.

3.2.2. Der Innere Detektor

Der Innere Detektor, dargestellt in Abbildung 3.3, besteht aus einem Pixeldetektor, Silizium-Mikrostreifen-Tracker (SCT) und einem Übergangsstrahlungs-Tracker (TRT), die in konzentrischen Zylindern um die Strahlachse angeordnet sind. Er besitzt einen Durchmesser von 2.1 m und ist 6.2 m lang und deckt den Winkelbereich $|\eta| < 2.5$ ab. Eine Magnetspule mit 2.5 m Durchmesser erzeugt im Inneren ein Magnetfeld der Stärke 2 T, das die Bahnen der geladenen Teilchen durch die Lorentzkraft krümmt, woraus der Impuls der Teilchen berechnet werden kann [9].

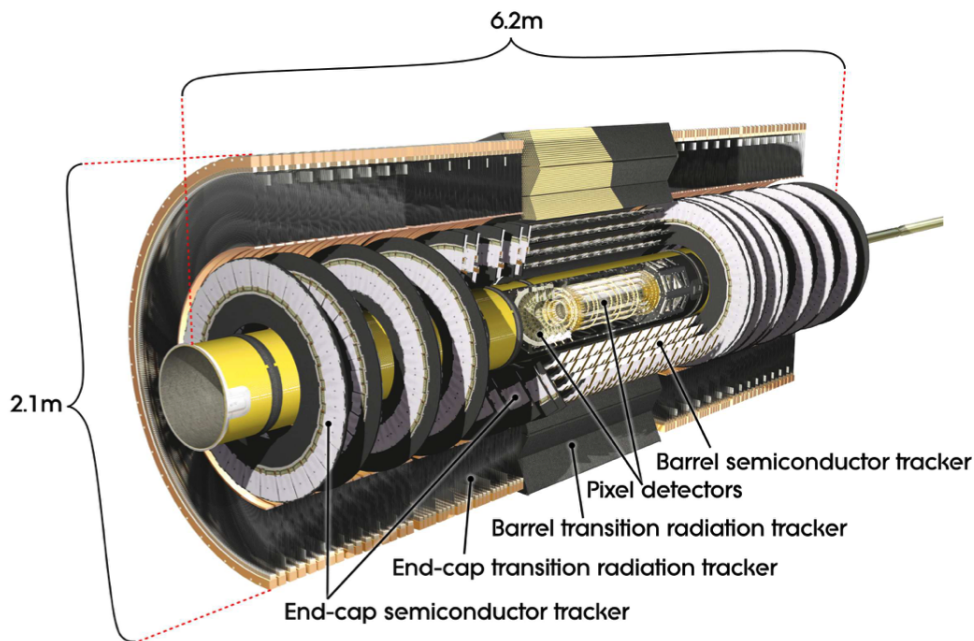


Abbildung 3.3.: Schematischer Aufbau des Inneren Detektors des ATLAS Experiments [9]

Die höchste Auflösung wird mit einem Silizium-Pixeldetektor erreicht, der sich daher am nächsten an der Vertex-Region befindet. Mit einer minimalen Pixelgröße von $R-\phi \times z$ von $50 \times 400 \mu\text{m}^2$ besitzt er ungefähr 80.4 Millionen Auslesekanäle [9]. Erweitert wurde der Pixeldetektor während der Abschaltung des LHC 2013 bis 2014 durch einen einführbaren B-Layer (IBL), der eine bessere Rekonstruktion der Teilchenbahnen ermöglicht [12]. Die nächste Detektorschicht bildet der SCT, der, wie der Pixeldetektor, auf dem Prinzip der Elektron-Loch-Paarbildung in dotiertem Silizium basiert, das hier in Streifen um die Vertexregion angeordnet ist. Der SCT besitzt 6.3 Millionen Auslesekanäle. Die äußerste Schicht des Inneren Detektors bildet der TRT. Dieser ermöglicht eine Spurverfolgung der Kollisionsprodukte mit 351000 Auslesekanälen im Bereich $|\eta| < 2.0$ [9].

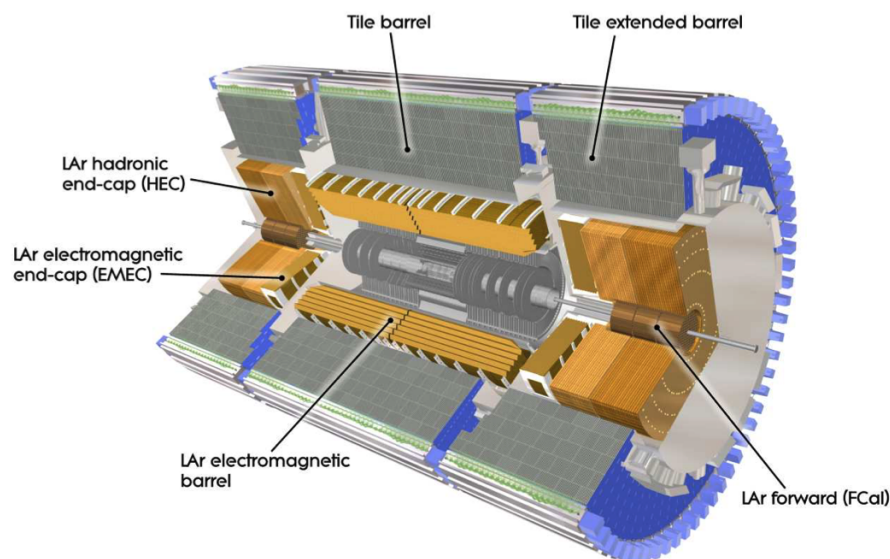


Abbildung 3.4.: Schematischer Aufbau des ATLAS-Kalorimetersystems [9].

3.2.3. Das Kalorimetersystem

Das Kalorimetersystem des ATLAS Detektors misst die Energien der Kollisionsprodukte. Es bildet die nächste Detektorschicht um den Inneren Detektor und ist in das innere elektromagnetische und das äußere hadronische Kalorimeter aufgeteilt. Der schematische Aufbau ist in Abbildung 3.4 zu sehen.

Das elektromagnetische flüssig-Argon-Kalorimeter deckt Pseudorapiditäten $|\eta| < 3.2$ ab. Das Messprinzip beruht darauf, dass hochenergetische Elektronen in Blei-Schichten Bremsstrahlung aussenden. Die Photonen bilden dann Positron-Elektron-Paare, die wiederum Bremsstrahlung emittieren, was nach häufiger Verkettung dieser beiden Prozesse in einem elektromagnetischen Schauer resultiert. Die einzelnen Energien der Schauer-

Teilchen werden in Flüssigargon (LAr) Detektoren gemessen, woraus die Energie des ursprünglichen Teilchens rekonstruiert werden kann [9].

Das hadronische Kalorimeter beruht auf dem gleichen Messprinzip, misst aber die Energie der bei den Kollisionen entstandenen Hadronen. Es besteht aus einem inneren Platten-Kalorimeter aus szintillatorischen Elementen und einem äußeren LAr-Hadronen-Kalorimeter. Zusammen decken sie Pseudorapiditäten bis $|\eta| < 4.9$ ab [9].

3.2.4. Das Myonen-System

Aufgrund ihres geringen Wirkungsquerschnitts fliegen Myonen fast ungehindert durch die erwähnten Detektoren hindurch, weshalb die äußerste Detektorschicht aus einem Myon-Spektrometer besteht. Dieses besteht aus mehreren Toroidmagneten, die die Spuren der Myonen im Bereich $|\eta| < 2.7$ krümmen. Die Spuren werden in überwachten Drift-Röhren (MDT) und Kathoden-Streifen-Kammern (CSC) gemessen. Ebenso werden die Koordinaten der Myonen in den Widerstands-Platten-Kammern (RPC) und den Thin-Gap-Kammern (TGC) bestimmt [9].

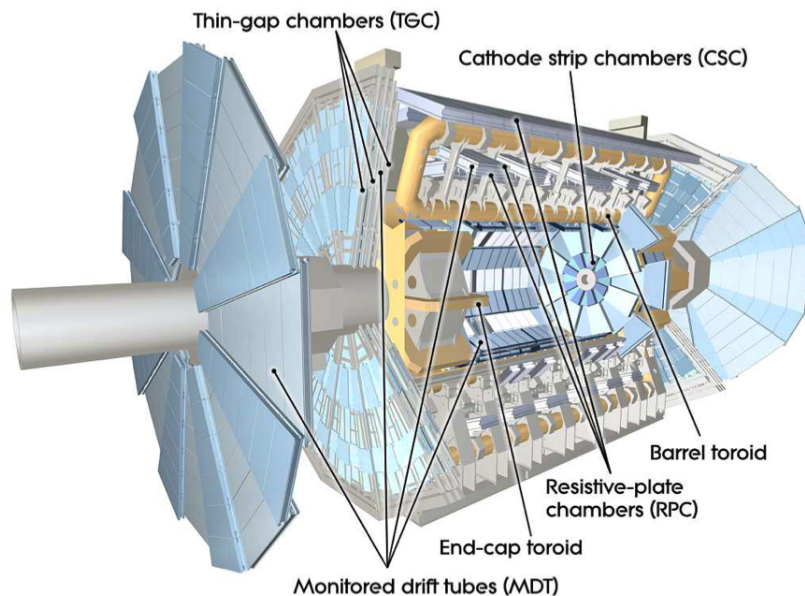


Abbildung 3.5.: Schematischer Aufbau des Myonen-Systems im ATLAS Detektor [9].

4. Daten und Monte-Carlo-Simulationen

Die in dieser Arbeit genutzten Daten wurden während des LHC Lauf-2 im Zeitraum 2015 bis 2018 bei einer Schwerpunktsenergie von 13 TeV aufgenommen. Der Verlauf der integrierten Luminositäten ist in Abbildung 4.1 dargestellt. Für diese Analyse werden 140.1 fb^{-1} genutzt [15].

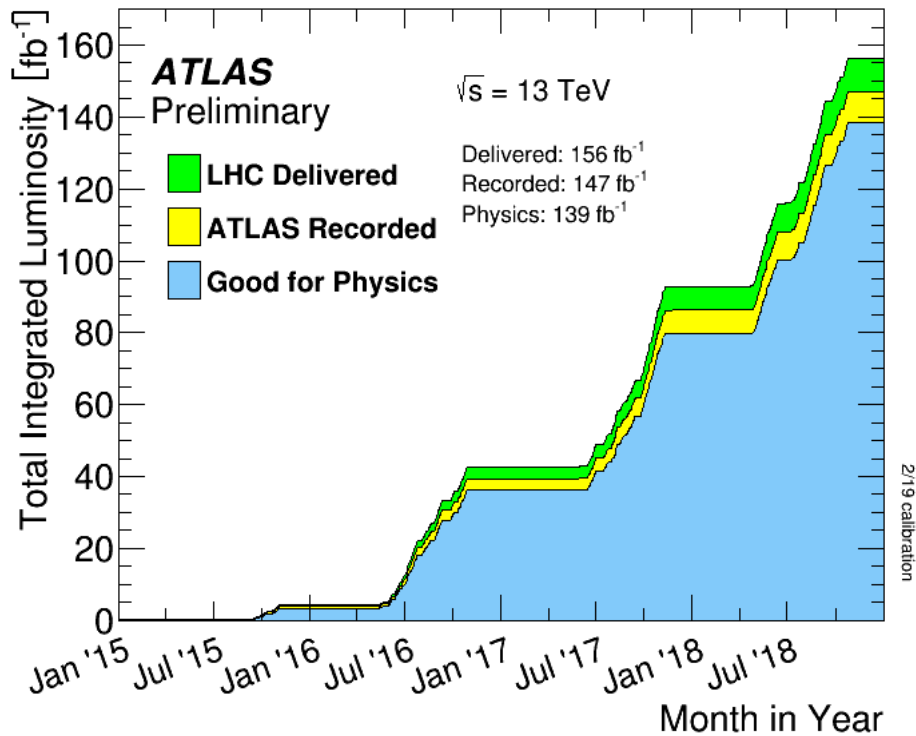


Abbildung 4.1.: Integrierte Luminosität des LHC über den Zeitraum 2015–2018 [2]. Im Dezember 2022 wurde die Messung auf 140.1 fb^{-1} korrigiert [15].

Für die Datenanalyse sind neben den Messdaten Simulationen essentiell. Sowohl Daten für Untergrund- als auch Signalprozesse werden mit Monte-Carlo-Generatoren (MC-Generatoren) erzeugt. Diese generieren die Matrixelemente (ME) der bei den Kollisionen stattfindenden Streuprozessen sowie die entstehenden Partonenschaufen (PS). Diese entstehen, wenn Quarks Gluonen aussenden, welche wiederum über Effekte der Quanten-

Chromo-Dynamik (QCD) weitere Quark-Paare erzeugen. Die Verkettung dieser Prozesse resultiert in annähernd konischen Strukturen, die als QCD Jets bezeichnet werden. In dieser Arbeit wird POWHEG-BOX [40] als ME Generator, sowie PYTHIA [43] und SHERPA [16] als Generatoren für MEs und PS genutzt. Die für die jeweiligen Untergrund- und Signalprozesse genutzten MC-Generatoren sind in Tabelle 4.1 aufgelistet.

Tabelle 4.1.: Liste der Monte-Carlo-Generatoren der einzelnen Prozesse.

Prozess	ME Generator	PS Generator
Signal		
$A \rightarrow \tau\tau$	PYTHIA v8.2	
Untergrund		
$H_{125} \rightarrow \tau\tau$	POWHEG-BOX v2	PYTHIA v8.2
$H_{125} \rightarrow WW$	POWHEG-BOX v2	PYTHIA v8.2
Diboson (semileptonisch)	SHERPA v2.2.1	
Diboson (vollleptonisch)	SHERPA v2.2.1	
$Z \rightarrow ll$	SHERPA v2.2.1	
$W \rightarrow l\nu$	SHERPA v2.2.1	
$t\bar{t}$	POWHEG-BOX v2	PYTHIA v8.2
einzel Top-Quark	POWHEG-BOX v2	PYTHIA v8.2

Zu den MC Simulationen werden ebenfalls die entsprechenden Detektorsignale im ATLAS Detektor mittels GEANT4 [29] simuliert. Für die Signalproben wird zusätzlich die ALTFast-II [10] Simulation verwendet, um die Signale der Kalorimeter zu simulieren.

5. Analyse

In diesem Kapitel werden zunächst die allgemeinen Grundzüge der Suche nach dem A -Boson zusammengefasst. Dabei werden der Signalprozess sowie die relevanten Untergrundprozesse beschrieben. Ebenso werden die Auswahlkriterien für die verschiedenen Signal- und Validierungsregionen dargelegt.

5.1. Signal- und Untergrundprozesse

Neben dem Signalprozess finden im ATLAS-Detektor weitere Streuprozesse statt, die ähnliche Signaturen im Detektor hinterlassen. Über die für diese Arbeit relevanten Prozesse wird hier ein Überblick gegeben.

5.1.1. Der Signalprozess

Der Signalprozess wird vor allem durch die hohe Kopplung des A -Bosons an massereiche Teilchen und die hohe Kopplung an geladene Leptonen und Up-Type-Quarks bestimmt (siehe Unterabschnitt 2.2.3). Die Produktion des A -Bosons läuft daher über Gluon-Fusion mit einer Top-Quark-Schleife. Die hohe Kopplungskonstante ζ_l und die geringe Masse m_A bestimmen den Zerfallskanal. Das A -Boson zerfällt fast ausschließlich in ein Paar aus τ -Leptonen, welche hadronisch und leptonisch zerfallen können. Da die Triggerschwellen für leichte Leptonen geringer sind, wird in der Analyse nur der leptonische Zerfall $\tau \rightarrow l\nu$ betrachtet [36]. Bei diesem zerfällt das τ in ein τ -Neutrino ν_τ und ein W -Boson, welches anschließend in ein Lepton-Neutrino-Paar zerfällt. Um den Signalprozess besser von dem Untergrundprozess $Z \rightarrow ll$ unterscheiden zu können, werden nur Ereignisse betrachtet, bei denen eines der W -Bosonen in ein Elektron e und das andere in ein Myon μ zerfällt. Der Signalprozess lässt sich also zu $gg \rightarrow A \rightarrow \tau\tau \rightarrow 2\nu_\tau e\nu_e \mu\nu_\mu$ zusammenfassen. Das zugehörige Feynman-Diagramm ist in Abbildung 5.1 zu sehen. Da die Masse des A -Bosons ein freier Parameter des 2HDM ist, werden in dieser Analyse Massenhypothesen von 20 GeV bis 110 GeV in 10 GeV Schritten simuliert.

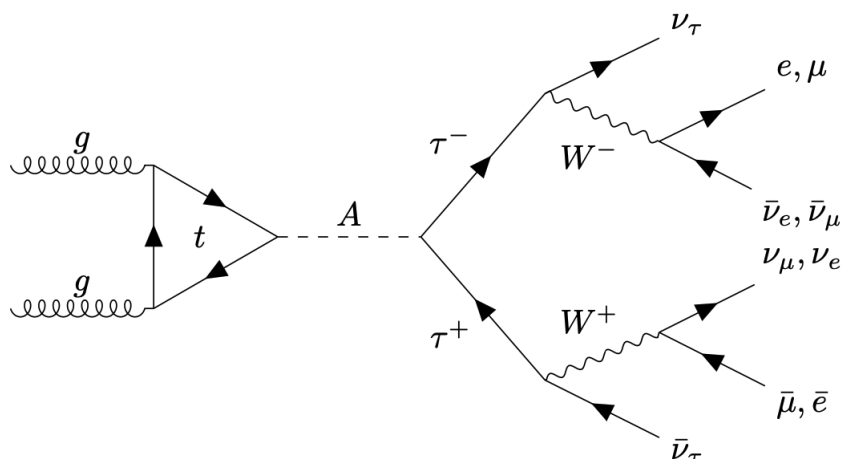


Abbildung 5.1.: Vollständiges Feynman-Diagramm des Signalprozesses mit voll-leptonischem Zerfall [37].

5.1.2. Untergrundprozesse

Die Abschätzung des Untergrundes ist essentiell für die Aussagekraft der statistischen Auswertung. Im folgenden Abschnitt werden die Untergrundprozesse beschrieben, die in dieser Analyse berücksichtigt werden müssen.

$Z \rightarrow ll$

Bei diesem Prozess zerfällt das Z -Boson in ein Paar aus Positron und Elektron, beziehungsweise Anti-Myon und Myon. Mit l sind hier und im weiteren Verlauf der Arbeit nur Elektronen und Myonen gemeint. Um diesen Untergrund zu unterdrücken, werden nur Ereignisse betrachtet, in denen ein Elektron und ein Myon im Endzustand vorhanden sind. Durch diese Bedingung gehen allerdings auch Teile des Signals verloren. Dennoch kann der Prozess Beiträge zum Untergrund leisten, falls im Detektor ein Elektron als Myon, oder umgekehrt, falsch identifiziert wird. Im Vergleich zum Signalprozess entstehen bei diesem Prozess keine Neutrinos, was in deutlich geringeren Werten für E_T^{miss} resultiert.

$Z \rightarrow \tau\tau$

Dieser Prozess stellt den Hauptuntergrund der Analyse dar. Wie im Signalprozess zerfällt das Z -Boson in zwei τ -Leptonen, die in Elektronen und Myonen zerfallen können. Dadurch, dass die Masse des Z -Bosons in dem Bereich der Massenhypothesen des Signalprozesses liegt, wird so eine Unterscheidung durch Massenrekonstruktion erschwert. Ein Unterschied zwischen dem Z -, A - und dem im nächsten Absatz beschriebenen SM-

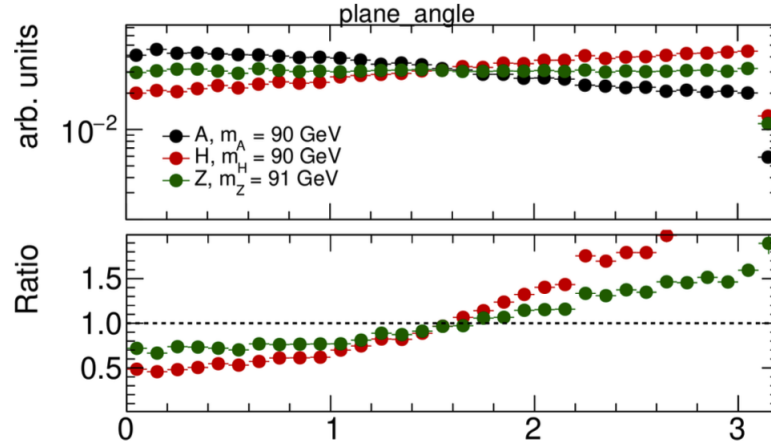


Abbildung 5.2.: Verteilung des Winkels ϕ zwischen den Zerfallsebenen der τ -Leptonen im Zerfall eines CP-ungeraden A -Higgs, CP-geraden SM-Higgs H und des Z -Bosons. Im unteren Plot werden die Werte für H und Z durch die des A -Bosons geteilt [37].

Higgs-Zerfall, der ebenfalls ein Untergrundprozess ist, liegt in den unterschiedlichen CP-Eigenwerten und Spins der Teilchen. Diese sorgen für die unterschiedlichen Winkelverteilungen zwischen den Leptonen, die in Abbildung 5.2 gezeigt sind. Es ist zu sehen, dass in dieser Verteilung ein deutliches Separationspotential zwischen den Prozessen besteht, das in Form der Variable ΔR (siehe Gleichung 3.4) betrachtet wird.

Standardmodell-Higgs Prozesse

Für das SM-Higgs-Boson existieren zwei Zerfallskanäle, bei denen ein Myon und ein Elektron im Endzustand vorhanden sind. Zum einen ist der Prozess $H \rightarrow \tau\tau$ dem Signalprozess sehr ähnlich. Unterscheiden lässt sich dieser allerdings anhand der hohen Masse des Higgs-Bosons im Vergleich zum A -Boson. Der zweite relevante Prozess ist $H \rightarrow WW$, bei dem beim Zerfall der W -Bosonen ebenfalls zwei Leptonen entstehen können.

Top-Quark Prozesse

Top-Quarks können einzeln oder als Top-Antitop-Paare bei den Kollisionen entstehen und zerfallen mit 95.7% Wahrscheinlichkeit [41] in ein W -Boson und ein Bottom-Quark. Das W -Boson kann in ein Quark-Paar oder ein Lepton-Neutrino Paar zerfallen. Dieser Zerfallskanal ist in Abbildung 5.3 dargestellt. Ist Letzteres der Fall für beide W -Bosonen, liefert auch dieser Prozess eine dem Signal sehr ähnliche Signatur im Detektor. Eine mögliche Unterscheidung ist durch weniger entstandene Neutrinos über E_T^{miss} möglich. Beim Zerfall eines einzelnen Top-Quarks wie in Abbildung 5.3, kann das Bottom-Quark

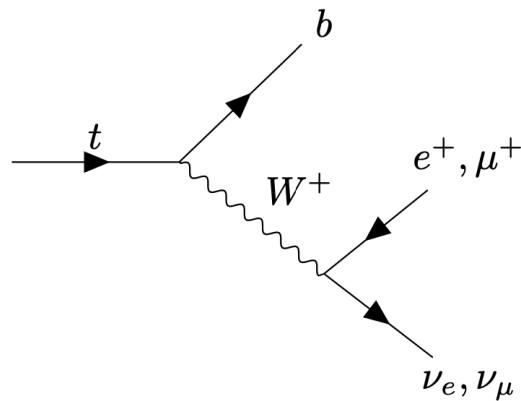


Abbildung 5.3.: Feynman-Diagramm eines Top-Zerfalls mit leptonischem W -Boson Zerfall [37].

einen Jet erzeugen, welcher als Lepton identifiziert werden kann. Um diesen Untergrund zu unterdrücken werden nur Ereignisse ausgewählt, die keine b -Jets enthalten.

$W \rightarrow l\nu_l$

Bei leptonischem Zerfall von W -Bosonen entsteht über die Kanäle $W \rightarrow l\nu_l$ und $W \rightarrow \tau\nu_\tau \rightarrow l\nu_l\nu_\tau$ ein Lepton im Endzustand. Durch falsche Identifikation zusätzlicher Detektorsignale wie Jets als Elektron oder Myon, kann dieser Prozess zum Untergrund beitragen.

Diboson-Prozesse

Mit der Bezeichnung Diboson-Prozesse werden alle Prozesse zusammengefasst, bei denen die Kombinationen WW , WZ , und ZZ auftreten. Die Endzustände können dabei dem Signalprozess stark ähneln, aber über bestimmte Massenvariablen als Untergrund identifiziert werden. Zusätzliche Beiträge liefern Diboson-Prozesse durch fehlerhafte Identifikation von Zerfalls- oder Kollisionsprodukten, wodurch fälschlicherweise der Endzustand des Signalprozesses gemessen werden kann.

QCD-Prozesse

Bei Proton-Proton-Kollisionen entstehen viele farbgeladene Teilchen, die im Detektor Jet-Signaturen oder isolierte Hadronen-Signale hinterlassen. Werden diese im Detektor als Elektronen oder Myonen misidentifiziert, können auch diese Prozesse zum Untergrund beitragen.

5.2. Ereignisauswahl

Die Auswahl der Kollisionsereignisse ist für die Auswertung sehr relevant. Dabei gibt es grundlegende Auswahlkriterien für die Gesamtstatistik sowie weitere Schnitte in den Wertebereichen verschiedener Variablen zur Definition von Kontroll- oder Signalregionen.

5.2.1. Grundlegende Auswahlkriterien

Eine der grundlegendsten Annahmen in dieser Analyse ist das Auftreten genau eines Elektrons und eines Myons im Endzustand um den $Z \rightarrow ll$ Untergrund zu reduzieren. Zusätzlich müssen die beiden Leptonen eine entgegengesetzte elektrische Ladung aufweisen, da das A -Boson elektrisch neutral ist.

Die in der Messung verwendeten Trigger sind in Tabelle 5.1 aufgelistet. Die Trigger speichern die Ereignisse ab bestimmten Werten der transversalen Impulse der Leptonen. Die Kurzbezeichnungen kodieren die Triggerschwellen. Die Bezeichnung e7mu24 gibt beispielsweise an, dass der Impuls des Elektrons $p_T^e > 7 \text{ GeV}$ und der des Myons $p_T^\mu > 24 \text{ GeV}$ sein muss. Da die Effizienz der Trigger zur Schwelle hin abnimmt, werden zusätzlich Schnitte auf minimale transversale Impulse von 10 GeV angewendet.

Eine weitere Beschränkung folgt aus dem Detektoraufbau und betrifft die Pseudorapiditäten. Myonen müssen $|\eta| < 2.7$ genügen, für Elektronen gilt $|\eta| < 2.47$. Für letztere werden Ereignisse mit $1.37 < |\eta| < 1.52$ zusätzlich ausgeschlossen, da die Energie-Auflösung in diesem Bereich niedriger ist. Für Jets gelten die Kriterien $p_T > 20 \text{ GeV}$ und $|\eta| < 4.5$. Weiterhin müssen die Leptonen dem *tight* Isolations- sowie dem *medium* Identifikationskriterium genügen, um in der Analyse berücksichtigt zu werden. Die Isolation beschreibt den Winkelabstand des Leptons zu anderen Objekten im Detektor, die Identifikation die Erkennung der Leptonen aus den Detektorsignalen. Genaueres zur Identifikation kann für Myonen in Quelle [13] und für Elektronen in Quelle [14] nachgelesen werden.

5.2.2. Signal- und Validierungsregionen

Im Folgenden wird ein Überblick über die verschiedenen Regionen dieser Analyse gegeben. Dazu werden zunächst zwei weitere Massenvariablen eingeführt.

Die totale transversale Masse

$$m_T^{\text{tot}} = \sqrt{(p_T^e + p_T^\mu + E_T^{\text{miss}})^2 - (\vec{p}_T^e + \vec{p}_T^\mu + \vec{E}_T^{\text{miss}})^2} \quad (5.1)$$

Tabelle 5.1.: Liste der verwendeten Trigger für die einzelnen Jahre der Messungen im ATLAS Detektor.

Jahr	Kurzbezeichnung	ATLAS-Trigger Bezeichnung
2015	e7mu24	HLT_e7_lhmedium_mu24
	e17mu14	HLT_e17_lhloose_mu14
	e26mu8	HLT_e24_lhmedium_nod0_L1EM20VHI_mu8noL1
2016	e7mu24	HLT_e7_lhmedium_nod0_mu24
	e17mu14	HLT_e17_lhloose_nod0_mu14
	e26mu8	HLT_e26_lhmedium_nod0_L1EM22VHI_mu8noL1
2017–18	e7mu24	HLT_e7_lhmedium_nod0_mu24
	e17mu14	HLT_e17_lhloose_nod0_mu14
	e26mu8	HLT_e26_lhmedium_nod0_mu8noL1

setzt sich aus den transversalen Impulsen und E_T^{miss} zusammen. Die zweite Variable wird mit dem *Missing Mass Calculator* (MMC) berechnet. Dieser Algorithmus berechnet anhand statistischer Methoden die wahrscheinlichste invariante Masse m_{MMC} des Ereignisses. Eine ausführlichere Erklärung ist in Quelle [22] zu finden.

Mit diesen Variablen lassen sich die verschiedenen Regionen der Analyse definieren: Die Low-Mass-Signalregion für geringe Massenhypothesen für m_A von 20 GeV bis 80 GeV, die High-Mass-Signalregion für Massenhypothesen von 80 GeV bis 110 GeV, die Top-Validierungsregion für den Top-Untergrund und die Z -Validierungsregion für den Untergrund durch den Prozess $Z \rightarrow \tau\tau$. Die Schnitte der Regionen sind in Tabelle 5.2 aufgelistet. Die Variable $n_{b\text{-jets}}$ bezeichnet die Anzahl der b -Jets eines Ereignisses.

Tabelle 5.2.: Überblick über die Schnitte der verschiedenen Regionen

Variable	Signalregionen		Validierungs-Regionen		MVA
	Low-Mass	High-Mass	Top	Z	
E_T^{miss}	$> 50 \text{ GeV}$	$> 30 \text{ GeV}$	$> 30 \text{ GeV}$	-	-
m_T^{tot}	$< 45 \text{ GeV}$	$< 65 \text{ GeV}$	$< 65 \text{ GeV}$	$< 65 \text{ GeV}$	-
ΔR_{ll}	< 0.7	< 1.0	< 1.0	> 1.4	< 1.4
m_{MMC}	$> 0 \text{ GeV}$	$> 35 \text{ GeV}$ $< 130 \text{ GeV}$	$> 0 \text{ GeV}$	$> 0 \text{ GeV}$	$> 0 \text{ GeV}$
$n_{b\text{-jets}}$	0	0	> 0	0	-

Zu den bisherigen Regionen wird für die Arbeit mit den neuronalen Netzen die MVA-Region eingeführt, welche im Vergleich zur kombinierten Signalregion mehr Ereignisse enthält, die später zum Training der Netze verwendet werden können.

6. Klassifikation von Ereignissen mit neuronalen Netzwerken

In dieser Analyse werden parametrisierte neuronale Netzwerke genutzt, um Signal- von Untergrundprozessen zu unterscheiden. Dafür wird im ersten Teil dieses Kapitels die Funktionsweise dieser Netzwerke erklärt. Danach wird das Ausgangs-Setup dieser Arbeit vorgestellt, das von Jannik Friese in seiner Masterarbeit [28] erarbeitet wurde. Anschließend wird die Auswahl neuer Variablen motiviert und die Leistung von Netzwerken verschiedener Architektur verglichen.

6.1. Neuronale Netzwerke

Neuronale Netzwerke (NN) sind in ihrer Funktionsweise eine Art mathematische Analogie zum menschlichen Gehirn: Ein komplexes Netz aus miteinander verbundenen Neuronen, das in der Lage ist, anhand einer Menge von Informationen eine Aussage zu treffen. Diese Aussage kann je nach Anwendungsgebiet stark unterschiedlich sein. Im Rahmen dieser Arbeit soll ein NN anhand verschiedener kinematischer Variablen entscheiden, ob ein Kollisionsereignis dem Signal- oder einem Untergrundprozess entspricht.

6.1.1. Neuronen und der Aufbau von Netzwerken

Die Neuronen sind der Grundbaustein eines NN. Diese erhalten als Eingabe einen Satz von Parameter, von denen jeder mit einem zugehörigen Gewicht multipliziert wird. Die Produkte werden aufsummiert und mit einem weiteren Summanden, dem Bias b addiert. In vektorieller Schreibweise mit den Vektoren \vec{a} der Eingabeparameter und der Gewichte \vec{w} lässt sich diese Operation schreiben als $z = \vec{w}\vec{a} + b$. Anschließend wird auf diese Summe eine Aktivierungsfunktion $A(x)$ angewendet. Für die in dieser Arbeit genutzten Netzwerke wird dafür die Sigmoidfunktion

$$\sigma(z) = \frac{1}{1 + e^{-z}} \quad (6.1)$$

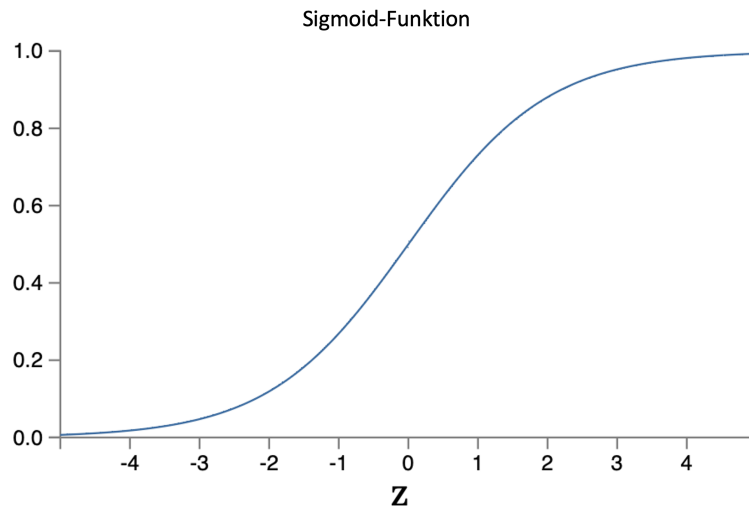


Abbildung 6.1.: Die Sigmoidfunktion in Abhängigkeit der Summe z [39].

genutzt, die einen Ausgabewert des Neurons im Bereich $0 < \sigma(x) < 1$ erzeugt. Der Verlauf der Sigmoidfunktion ist in Abbildung 6.1 dargestellt [39].

Ein NN besteht aus einer Eingabeschicht (engl. input layer), versteckten Schichten (engl. hidden layer) und einer Ausgabeschicht (engl. output layer). Die Eingabeschicht besteht aus sogenannten Eingabeneuronen. Diese speziellen Neuronen übergeben der ersten versteckten Schicht die Eingabeparameter, ohne vorher Berechnungen mit ihnen durchzuführen. Für einen Satz von n Parametern werden also n Eingabe-Neuronen benötigt.

In einem voll verbundenen NN werden jedem Neuron in einer versteckten Schicht die Ausgabewerte aller Neuronen in der Schicht davor übergeben. Aus diesen Eingaben berechnet jedes Neuron wieder eine Ausgabe mit einem eigenen Satz an Gewichten und einem eigenen Bias. Die Ausgabe a des Neurons n in der Schicht l berechnet sich gemäß

$$a_n^l = A \left(\sum_m w_{nm}^l a_m^{l-1} + b_n^l \right) = A(z_j^l) . \quad (6.2)$$

Die Ausgabe-Schicht besteht für diese Analyse aus einem Neuron und berechnet den endgültigen Wert für den Parametersatz [39]. Eine schematische Darstellung eines NN mit den verschiedenen Schichten ist in Abbildung 6.2 zu sehen.

6.1.2. Training neuronaler Netzwerke

In der Praxis werden die Gewichte eines NN mit normalverteilten Werten initialisiert. Die eigentliche Spezialisierung auf die vorgegebene Aufgabe erfolgt im Training. Dazu wird

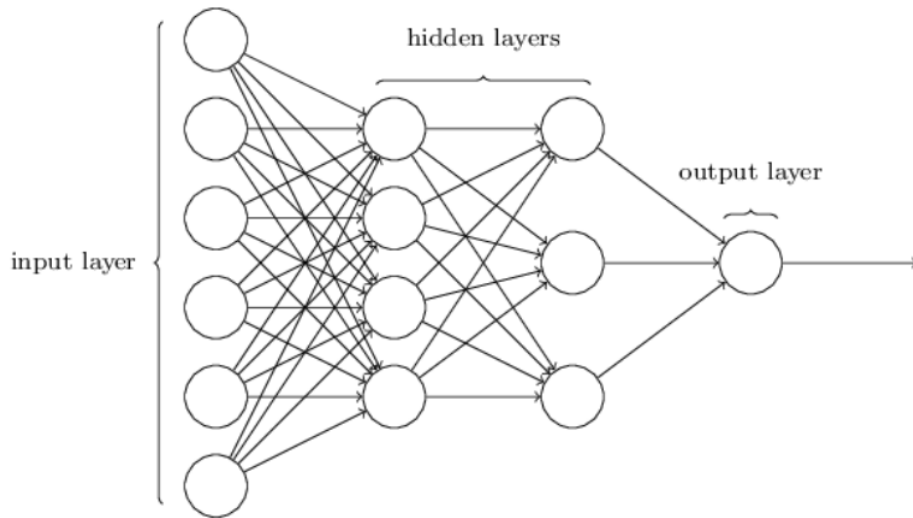


Abbildung 6.2.: Schematische Darstellung eines voll verbundenen neuronalen Netzes. Die Pfeile geben die Übergabe der Ausgabewerte der jeweiligen Neuronen an [39].

ein Datensatz benötigt, der in Trainings-, Validierungs- und Testdatensatz aufgeteilt wird. Mit dem Trainingsdatensatz wird das eigentliche Training durchgeführt. Die anderen beiden Datensätze dienen zur Validierung des Netzes bei Anwendung auf unbekannte Daten, um eine Voreingenommenheit auf die Trainingsdaten zu vermeiden. Sind die gewünschten Ausgaben des Datensatzes bekannt, kann die überwachte Trainingsmethode angewendet werden [39].

Dafür wird zunächst eine Verlustfunktion definiert. Da in dieser Analyse eine binäre Klassifizierung vorgenommen werden soll, wird die binäre Kreuzentropie

$$L = y \ln a + (1 - y) \ln(1 - a) \quad (6.3)$$

verwendet. Dabei ist y die gewünschte Ausgabe des Netzwerks und a die Ausgabe des Ausgabeneurons. Signalprozesse sollen hierbei die Wertung 1, Untergrundprozesse die Wertung 0 erhalten. Diese Funktion charakterisiert die Abweichung des Netzwerks von der Erwartung. Für einen ganzen Trainingsdatensatz lässt sich damit die Verlustfunktion

$$C = \frac{1}{N} \sum_n L_n \quad (6.4)$$

berechnen, die den Verlust über N Proben eines Trainingsdatensatzes mittelt.

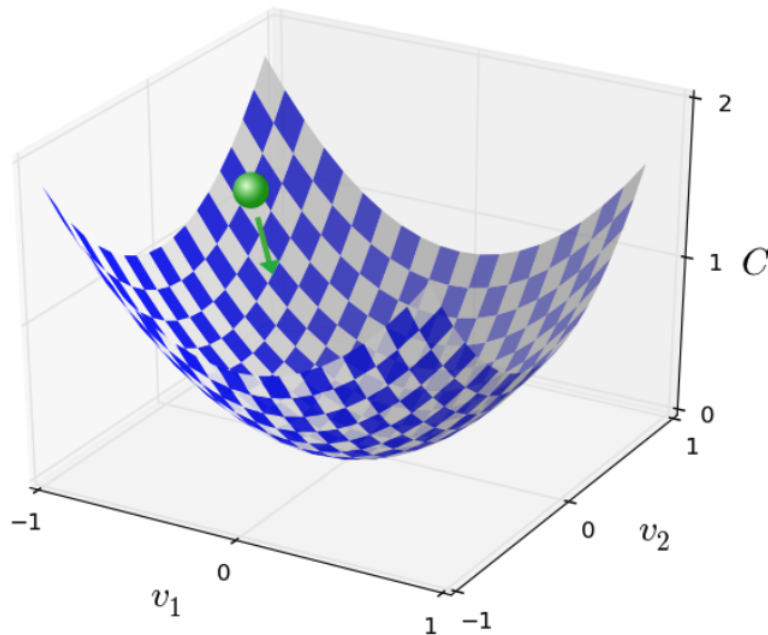


Abbildung 6.3.: Beispiel einer Verlustfunktion in Abhängigkeit der Parameter v_1 und v_2 . Das Netzwerk (grüne Kugel) soll sich ins Minimum bewegen [39].

Der Gradientenabstieg

Die Aufgabe des Netzwerktrainings ist es, den Verlust C zu minimieren, indem die Gewichte und Biases des Netzes angepasst werden. Veranschaulichen lässt sich das, indem man sich die Verlustfunktion als hochdimensionale Landschaft im Raum der Netzwerkparameter vorstellt, in der das Netzwerk durch seine Initialisierung zufällig platziert wird (siehe Abbildung 6.3). Durch gezielte Variation der Parameter soll sich das Netz im Minimum der Funktion einfinden.

Die einfachste Methode um das zu erreichen, ist der Gradientenabstieg. Dafür wird der Gradient

$$\nabla C = \left(\frac{\partial C}{\partial \theta_1}, \dots, \frac{\partial C}{\partial \theta_p} \right)^T \quad (6.5)$$

aus den Ableitungen von C nach den Gewichten und Biases θ_i des Netzes gebildet, der in Richtung des größten Anstiegs zeigt. Anschließend werden die Parameter mit der Vorschrift

$$\theta \rightarrow \theta' = \theta - \eta \nabla C \quad (6.6)$$

aktualisiert. Der Parameter η ist dabei die Lernrate des Netzwerks und kann als dessen zurückgelegte Distanz in der Landschaft pro Aktualisierungsschritt gesehen werden. Dieser Vorgang wird theoretisch so lang wiederholt, bis der Gradient verschwindet.

Der Gradientenabstieg birgt allerdings ein Problem in sich. Die Verlustfunktion besitzt im Regelfall neben dem globalen Minimum viele weitere lokale Minima, in denen sich das Netzwerk mit dieser Methode festsetzen kann [39].

Der Adam-Optimierer

Der Adam-Optimierer ist eine Methode um dieses Problem zu verhindern und zeichnet sich durch eine komplexere Aktualisierungsvorschrift für die Netzwerkparameter aus. Der “Bewegung” des Netzwerks in der Landschaft der Verlustfunktion werden ein Momentum der ersten und der zweiten Ordnung, s_t und r_t hinzugefügt, die mit den exponentiellen Zerfallsraten $\beta_1, \beta_2 \in [0, 1)$ definiert sind als

$$\begin{aligned} s_t &= \beta_1 s_{t-1} + (1 - \beta_1) \nabla C \\ r_t &= \beta_2 r_{t-1} + (1 - \beta_2) \nabla C \odot \nabla C . \end{aligned} \quad (6.7)$$

Dabei bezeichnet \odot die elementweise Multiplikation. Die Initialisierung der Momente bei 0 führt zu einer Voreingenommenheit, die im nächsten Schritt korrigiert wird:

$$\begin{aligned} \hat{s}_t &= \frac{s_t}{1 - \beta_1^t} , \\ \hat{r}_t &= \frac{r_t}{1 - \beta_2^t} . \end{aligned} \quad (6.8)$$

Die Netzwerkparameter werden anschließend elementweise aktualisiert:

$$\theta_t = \theta_{t-1} - \eta \frac{\hat{s}_t}{\sqrt{\hat{r}_t} + \epsilon} . \quad (6.9)$$

Dabei ist ϵ eine kleine Konstante, die numerische Stabilität gewährleistet [30].

Damit ist das Netzwerk in der Lage, beim Durchlaufen einer Mulde Momentum zu erhalten und sich aus einem Minimum heraus zu bewegen.

Das Training wird in dieser Analyse mittels des `Keras` Python-Moduls [1] durchgeführt, in dem der Adam-Optimierer bereits integriert ist.

6.1.3. Wichtige Parameter

Die Leistung von NNs lässt sich anhand verschiedener Parameter optimieren. Über die für diese Analyse Relevanten wird hier ein kurzer Überblick gegeben.

Schichtanzahl und Schichtgröße

Diese Parameter sind essenziell für die Kapazität, die dem Netzwerk zur Verfügung steht, um komplexe Aufgaben zu bewältigen. Sie charakterisieren die Anzahl der versteckten Schichten und die Anzahl der darin enthaltenen Neuronen. Sind die Parameter zu klein gewählt, kann das Netzwerk die Aufgabe nicht bewältigen. Sind sie zu groß, lernt das Netz die Trainingsdaten auswendig und kann schlechter mit unbekanntem Daten umgehen, was als *Overfitting* bezeichnet wird.

Lernrate

Die Lernrate η ist das Maß für die Stärke der Änderung der Parameter. Sie kann sowohl konstant sein, als auch epochenabhängig variiert werden. In dieser Analyse werden Cosinus- und exponentiell modulierte Lernraten verwendet.

Mini-Batches und Epochen

Um für eine Aktualisierung der Netzwerkparameter nicht immer den Gradienten für alle Trainings-Proben bestimmen zu müssen, wird das Trainingsset in *mini-batches* aufgeteilt, deren Größe als *mini-batch-Größe* bezeichnet wird. Die Verlustfunktion wird dann nur über die Proben in dieser *mini-batch* gemittelt, was die Berechnungszeit deutlich verkürzt. Der Nachteil ist, dass der so ermittelte Gradient nur näherungsweise berechnet wird und die Aktualisierung der Parameter weniger genau sein kann.

Für jede *mini-batch* werden die Parameter des Netzes aktualisiert, sodass innerhalb eines Durchlaufs über das Trainingsset bereits mehrere Optimierungsschritte vorgenommen werden. Dann ist eine Epoche des Trainings abgeschlossen [39]. Die Anzahl der Trainingsepochen ist ebenfalls maßgeblich für die Leistung des Netzwerks. Trainiert man über zu viele Epochen, kann das Netz *overfitten*.

k-Fold-Kreuzvalidierung

Diese Methodik soll verhindern, dass das Netz eine Voreingenommenheit gegenüber den Trainings- und Validierungsdaten entwickelt, da die Anpassung der Netzwerkparameter auf Basis dieser Daten geschieht.

Bei der k-Fold-Methode wird der zur Verfügung stehende Datensatz in k Datensätze aufgeteilt, genannt k-Folds, mit denen k separate Netzwerke trainiert werden. Dabei werden bei jedem Netz unterschiedliche k-Folds als Trainings-, Validierungs- und Test-

Tabelle 6.1.: Überblick über die Funktion der k-Folds im Training der verschiedenen Netzwerke bei Anwendung der k-Fold Kreuzvalidierung.

k-Fold	Netz 1	Netz 2	Netz 3	Netz 4	Netz 5
1	Test	Training	Training	Training	Validierung
2	Validierung	Test	Training	Training	Training
3	Training	Validierung	Test	Training	Training
4	Training	Training	Validierung	Test	Training
5	Training	Training	Training	Validierung	Test

datensatz verwendet. Bei der Anwendung eines Netzes, das mit der Kreuzvalidierung trainiert wurde, werden alle k Netze auf die Daten angewendet und anschließend die gemittelte Wertung als Ausgabe betrachtet.

In dieser Analyse werden immer 5 k-Folds pro Netzwerkkonfiguration genutzt. Die Funktion der k-Folds in den verschiedenen Netzwerken ist in Tabelle 6.1 zu sehen.

6.1.4. Leistungsmaße

Neben dem Verlust (siehe Gleichung 6.4) kann die Leistung eines NN anhand anderer Maße wie der *receiver operating characteristic* Kurve (ROC-Kurve), der Fläche unter der ROC-Kurve, die als *area under curve* (AUC) bezeichnet wird, und, im Zusammenhang mit der Suche nach dem A -Boson, anhand der resultierenden Signifikanz gemessen werden. Diese Maße werden im Folgenden erklärt und werden später genutzt, um die Leistung verschiedener NNs miteinander zu vergleichen.

ROC-Kurve und AUC

Um der Wertung eines Ereignisses durch das Netzwerk überhaupt erst eine Aussagekraft zu geben, muss ein Schwellenwert g festgelegt werden, der, wie die Ausgabe des Netzes, im Bereich $0 \leq g \leq 1$ liegt. Ereignisse mit Wertungen höher als der Schwellenwert werden als Signal, Ereignisse mit niedrigeren Wertungen als Untergrund klassifiziert angenommen.

Für die Bestimmung der ROC-Kurve wird g von 0 bis 1 variiert und für jeden Wert die *false positive rate*, die relative Anzahl an Untergrundereignissen, die als Signalereignis klassifiziert werden, und die *true positive rate*, die relative Anzahl an Signalereignissen, die als solche erkannt werden, bestimmt. Trägt man diese Werte gegeneinander auf, erhält man die ROC-Kurve.

Beispielhaft sind die ROC-Kurven eines perfekten (grün), eines zufälligen (rot) und eines

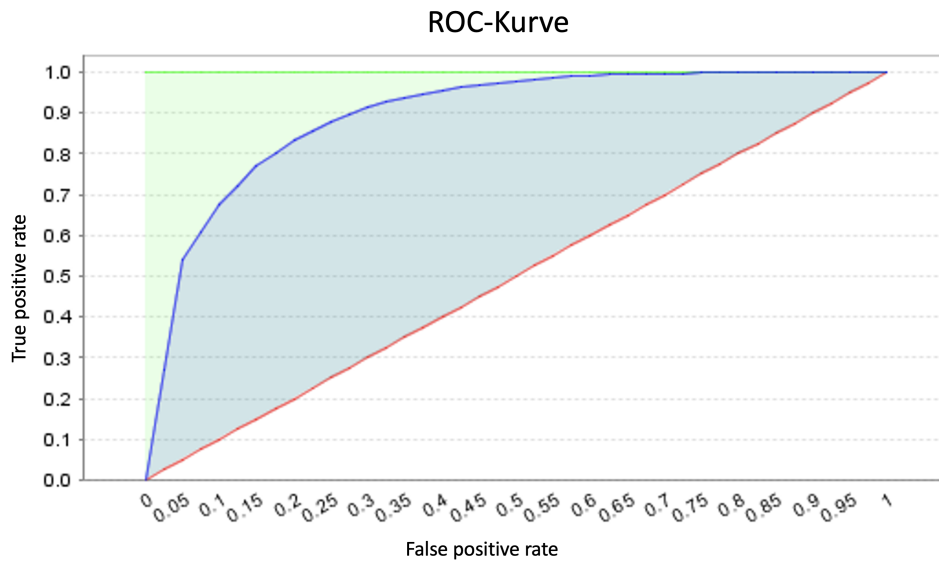


Abbildung 6.4.: Beispiele für ROC-Kurven eines perfekten (grün), eines zufälligen (rot) und eines möglichen realen Klassifizierers (blau). Die eingefärbten Flächen sind die AUCs der jeweiligen Klassifizierer [4].

möglichen realen Klassifizierers (blau) in Abbildung 6.4 dargestellt. Der zufällige Klassifizierer entspricht einer Diagonalen, der perfekte einer Konstanten bei 1. Je besser der Klassifizierer wird, desto mehr wölbt sich die ROC-Kurve nach oben hin aus.

Die Kurven werden in der Analyse berechnet, nachdem die trainierten Netzwerke auf alle Simulationsdaten angewendet und für jedes Ereignis eine Wertung ermittelt wurde. Da die Netzwerke sensitiv gegenüber den verschiedenen Massenhypothesen trainiert werden, werden für jede A -Boson-Masse Wertungen bestimmt, sodass sich aus den daraus resultierenden später gezeigten Werteverteilungen wie Abbildung 6.9 und Abbildung 6.10 eine ROC-Kurve für jede Masse erstellen lässt.

Die AUC ist das Integral der ROC-Kurve und ist in der Grafik für den realen und den perfekten Klassifizierer farblich markiert, schließt die Fläche unter der roten Kurve allerdings mit ein. Für den zufälligen Klassifizierer beträgt sie 0.5 und nimmt für einen perfekten Klassifizierer den Wert 1.0 an.

Signifikanz

Die Signifikanz ist ein Maß für das Verhältnis der Anzahl von Signalereignissen N_{Signal} und Untergrundereignissen $N_{\text{Untergrund}}$ und ist näherungsweise als

$$\sigma = \frac{N_{\text{Signal}}}{\sqrt{N_{\text{Untergrund}}}} \quad (6.10)$$

definiert. Die Leistung des Netzes lässt sich anhand des Verhältnisses der Signifikanzen vor und nach der Anwendung des Grenzwertes g auf die Wertungen des Netzwerkes abschätzen und steht über die Relation

$$\frac{\sigma_{\text{NN}}}{\sigma} = \frac{\text{True positive rate}}{\sqrt{\text{False positive rate}}} \quad (6.11)$$

mit der ROC-Kurve in Verbindung.

6.2. Verwendete Trainingsdaten

Für die in dieser Arbeit trainierten Netzwerke wurden zwei verschiedene Datensätze verwendet. Das erste Set besteht aus den Monte-Carlo-Simulationen in der gesamten Signalregion, das zweite aus den Simulationen, die in der neu eingeführten MVA-Region liegen. Die Verteilung von p_{T}^{μ} in der High-Mass-Signalregion ist in Abbildung 6.5 dargestellt.

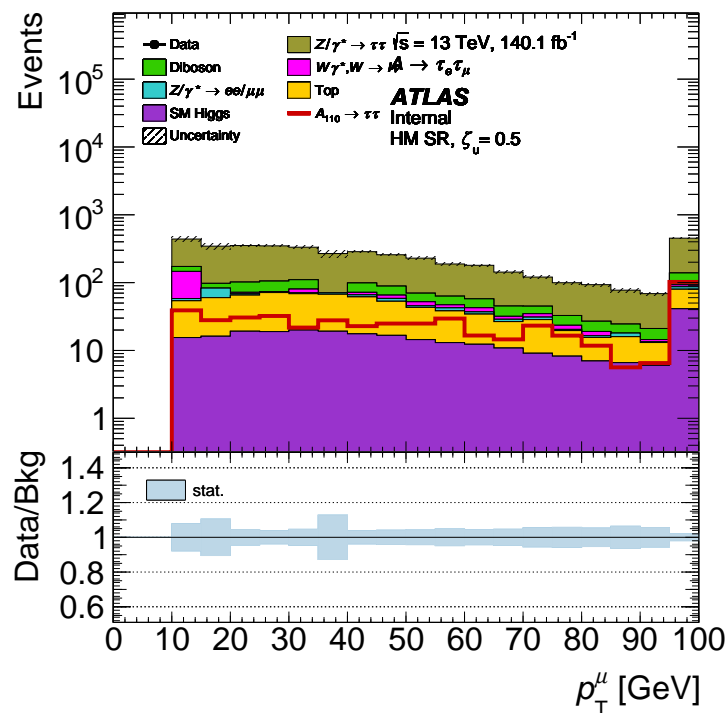


Abbildung 6.5.: Verteilung von p_{T}^{μ} der Untergründe und des Signals der Massenhypothese 110 GeV in der High-Mass-Signalregion.

Beide Datensätze enthalten sowohl die Signalereignisse als auch die in Unterabschnitt 5.1.2

aufgeführten Untergrundprozesse. Der QCD-Untergrund wird in beiden Sätzen gleichermaßen durch aufgenommene Daten in der B-Region der Fake-Faktor-Methode abgeschätzt und über die Gesamtanzahl von Multijet Ereignissen in der MVA-Region gewichtet. Diese Methode ermittelt den Beitrag von Leptonen aus unzureichend gut modellierten MC-Simulationen der QCD-Prozesse zum Untergrund. Die B-Region enthält Ereignisse, bei denen das Isolationskriterium des Myons und die Identifikation oder Isolation des Elektrons nicht erfüllt sind. Genaueres kann dazu in der Masterarbeit von Jannik Friese [28] nachgelesen werden.

Um das Netzwerk zusätzlich auf die verschiedenen Massenhypothesen zu sensitivieren, wird neben den kinematischen Parametern auch ein Wert für die Masse des A -Bosons in der Form $\frac{m_A}{100}$ GeV übergeben, da das Netzwerk mit kleinen Werten besser umgehen kann. Da die Charakteristik der Untergrundprozesse unabhängig von der A -Boson-Masse ist, wird diesen ein zufälliger Wert für m_A zugewiesen, um keine Voreingenommenheit des Netzwerks auf verschiedene Untergründe oder Massenhypothesen zu erzeugen.

6.3. Das Referenznetzwerk

Um eine Referenz für die Leistung der in dieser Arbeit trainierten Netzwerke zu haben, wird zunächst das Netzwerk aus der Masterarbeit von Jannik Friese [28] reproduziert. Dieses Netz, im Folgenden immer als *NNI* bezeichnet, besteht aus drei versteckten Schichten mit je 10 Neuronen. Trainiert wurde es mit dem Datensatz der Signalregion, das in *mini-batches* von 100 Proben aufgeteilt wurde. Als Eingabe wurden dem Netzwerk die Variablen m_T^{tot} , m_{MMC} , m_{ll} , E_T^{miss} , p_T^{Higgs} , p_T^e , p_T^μ , η_e , η_μ , ϕ_e , ϕ_μ , ΔR_{ll} , $\Delta\phi_{E_T^{\text{miss}}}^{\text{leadlep}}$ und $\Delta\phi_{E_T^{\text{miss}}}^{\text{subleadlep}}$, sowie die zusätzliche Massenvariable im damaligen Bereich von 20 GeV bis 90 GeV übergeben. Die Größen E_T^{miss} und m_T^{tot} werden verwendet, um die Beiträge von Neutrinos zur Energie- und Impulserhaltung in der transversalen Ebene der verschiedenen Prozesse zu unterscheiden. In Kombination mit den transversalen Impulsen der Leptonen und p_T^{Higgs} ist es dem Netzwerk möglich, anhand dieser Variablen die Masse und Energie des Teilchens abzuschätzen, aus dessen Zerfall die gemessenen Teilchen entstanden sein können. Für die verschiedenen Prozesse werden dafür unterschiedliche Charakteristiken erwartet, da die intermediären Teilchen verschiedene Massen haben. Die MMC-Masse bietet einen weiteren Anhaltspunkt dafür.

Die Variable m_{ll} beschreibt die invariante Masse der Leptonen und ist definiert als

$$m_{ll} = \sqrt{(E_T^e + E_T^\mu)^2 - (\vec{p}_T^e + \vec{p}_T^\mu)^2}. \quad (6.12)$$

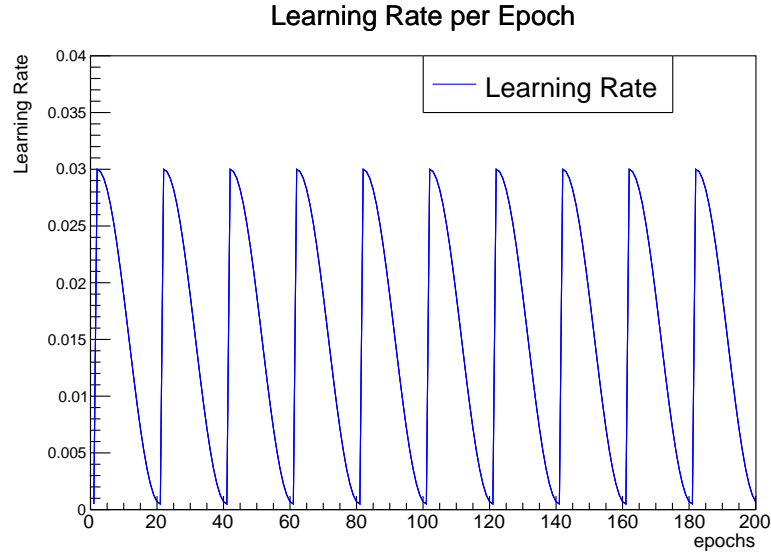


Abbildung 6.6.: Lernrate der Netzwerke NNI und $NNII$ in Abhängigkeit der Trainingsepoche.

Mit den Winkelvariablen η , ϕ und ΔR_{ll} wird dem Netz ermöglicht, die in Absatz 5.1.2 erwähnten Unterschiede in den Winkelverteilungen der Leptonen zu erkennen. Die Variablen $\Delta\phi_{E_T^{\text{miss}}}^{\text{leadlep}}$ bzw. $\Delta\phi_{E_T^{\text{miss}}}^{\text{subleadlep}}$ beschreiben die Winkeldifferenz von ϕ zwischen der Richtung von E_T^{miss} und dem Lepton mit dem höchsten bzw. zweithöchsten Impuls.

Trainiert wurde das Netz mit einer Cosinus-modulierten Lernrate, die in Abbildung 6.6 zu sehen ist. Das periodische Erhöhen der Lernrate ermöglicht es dem Netz, einfacher aus Minima der Verlustfunktion zu gelangen.

In Abbildung 6.7 sind die ROC-Kurven des Netzwerks für die verschiedenen Massenhypothesen dargestellt. Es ist eindeutig zu sehen, dass das Netzwerk für alle Massen außer für 90 GeV eine gute Klassifizierung vornimmt. Die Ursache für die schlechte Leistung für diese Masse ist der $Z \rightarrow \tau\tau$ Untergrund. Das A -Boson hat hier fast die gleiche Masse wie das Z -Boson, weswegen der Signalprozess von diesem Untergrund kaum zu unterscheiden ist.

Dieses Problem wird ebenfalls bei den Signifikanzen in Abbildung 6.8 deutlich. Für 90 GeV wird eine maximale Signifikanzverbesserung von ca. 1.2 erreicht, während diese für andere Massen Werte bis zu 2.2 annimmt.

Die Leistung des Referenznetzes weicht etwas von der in der Masterarbeit von Jannik Friese [28] ab, was auf statistischen Faktoren wie der Initialisierung der Anfangsparameter und der Zusammensetzung der einzelnen *mini-batches* zurückzuführen ist.

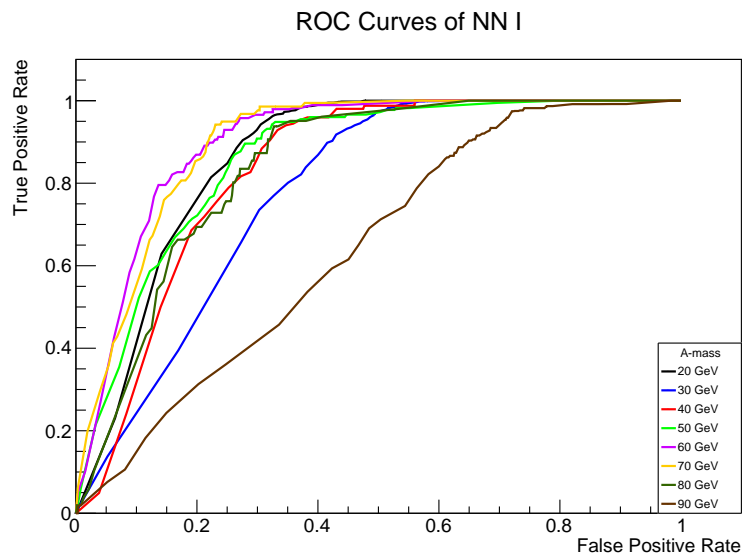


Abbildung 6.7.: ROC-Kurven des Netzwerks *NNI* für die verschiedenen Massenhypothesen.

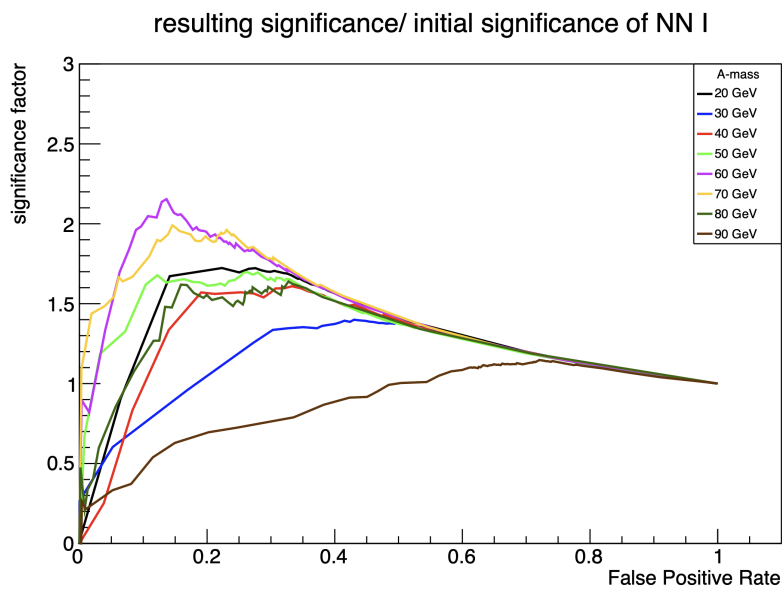


Abbildung 6.8.: Signifikanz-Faktoren des Netzwerks *NNI* für die verschiedenen Massenhypothesen.

6.4. Neue Massenhypothesen

Die erste Veränderung am Training der Netzwerke ist die Erweiterung des Massenspektrums des A -Bosons um die Massen 100 GeV und 110 GeV. Dafür werden in den Trainingsdatensatz die Monte-Carlo-Simulationen der Signalprozesse der entsprechenden m_A übernommen, sowie die zufällige Massenzuweisung auf diesen Bereich erweitert. Anhand dieses Netzwerks wird untersucht, wie sich die Erweiterung des Massenbereichs auf die Leistung der anderen Massen auswirkt. Die anderen Netzwerkparameter bleiben unverändert.

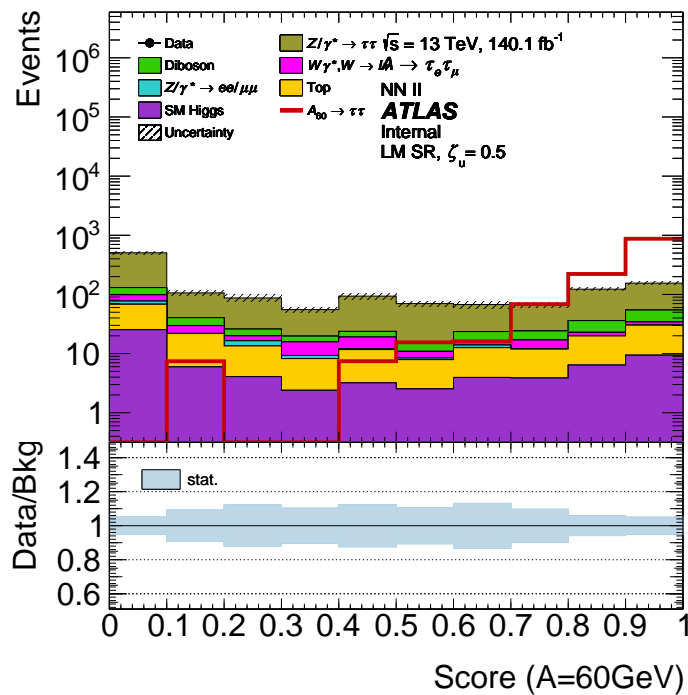


Abbildung 6.9.: Wertungsverteilung des Netzwerks *NN II* für $m_A = 60$ GeV.

Die Verteilungen der Wertungen, die das Netzwerk den Simulationen zuweist, werden in Abbildung 6.9 für $m_A = 60$ GeV und in Abbildung 6.10 für $m_A = 90$ GeV gezeigt. In Abbildung 6.9 ist zu sehen, dass das Netzwerk den Signalprozess sehr gut erkennt, da die Anzahl der Signalereignisse für hohe Wertungen zunimmt. Der niedrigste Wertebereich weist die höchste Anzahl an Untergrundereignissen auf, in den höheren Wertebereichen bleibt der Untergrund auf einem niedrigeren Niveau ungefähr gleichverteilt. Das kann dadurch erklärt werden, dass es für das Netz aufgrund der Gleichbehandlung aller Untergründe schwer ist, eine eindeutige Charakteristik festzustellen.

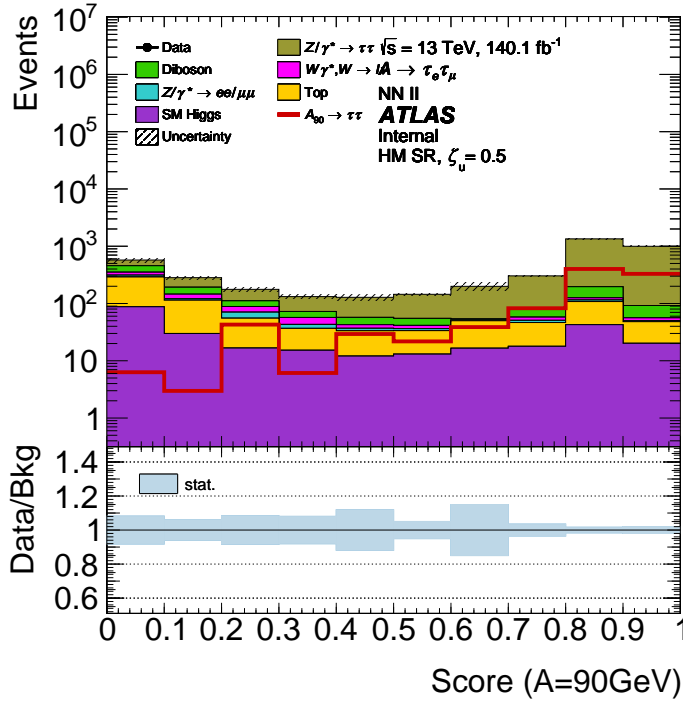


Abbildung 6.10.: Wertungsverteilung des Netzwerks *NN II* für $m_A = 90$ GeV.

In der Verteilung in Abbildung 6.10 ist das Problem der Klassifizierung für $m_A = 90$ GeV deutlich zu sehen. Die Erkennung des Signals gelingt dem Netzwerk gut, allerdings weist die Verteilung der $Z \rightarrow \tau\tau$ Ereignisse die gleiche Charakteristik auf. Das Netzwerk erkennt aufgrund der Ähnlichkeit der Ereignisse diesen Untergrund ebenfalls als Signal. Anhand einer feiner eingeteilten Variante dieser Plots werden die ROC-Kurven und Signifikanzen der Netzwerke für den Leistungsvergleich berechnet.

Die ROC-Kurven sind für das Netz *NN II* in Abbildung 6.11 abgebildet. Die Erweiterung des Massenspektrums hat eine geringe Auswirkung auf die Leistung des Netzwerks, die auf die Zusammensetzung der Trainingsdaten zurückzuführen ist. Durch die neuen Massenhypothesen reduziert sich der Anteil an Untergrundproben für das Training jeder Massenhypothese.

Vor allem für die Massen 20 GeV bis 80 GeV scheint das Netzwerk davon zu profitieren. Die ROC-Kurven steigen etwas steiler an und die AUC nimmt etwas zu. Da beim Training alle Untergründe gleich klassifiziert werden, muss das Netzwerk viele verschiedene Charakteristiken lernen, die alle zur gleichen Klassifizierung führen sollen. Der höhere relative Anteil an Signalproben sorgt für eine Fokussierung auf das Erkennen von Signalprozessen und es wird weniger Kapazität für die Mustererkennung in den verschiedenen Untergründen verwendet, die sich in diesem Massenbereich stärker vom Signal unterscheiden.

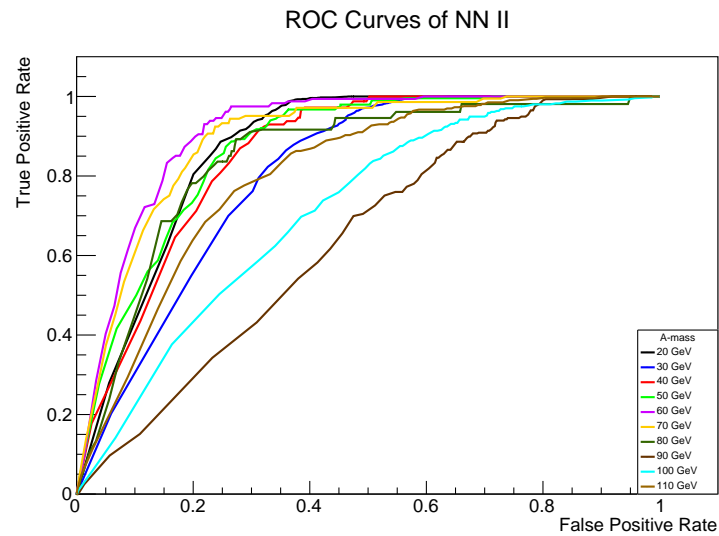


Abbildung 6.11.: ROC-Kurven des Netzwerks *NNII* für die verschiedenen Massenhypothesen.

Für die Klassifizierung bei 90 GeV wirkt sich die veränderte Zusammensetzung negativ aus. Das Netzwerk nähert sich noch weiter der Charakteristik eines zufälligen Klassifizierers an, da die Proben des $Z \rightarrow \tau\tau$ Untergrunds nicht ausreichen, um die Unterschiede zum Signal zu erkennen.

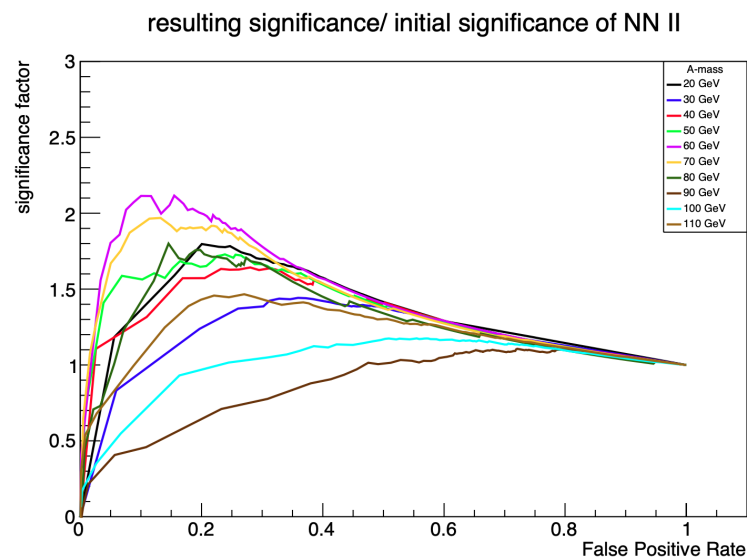


Abbildung 6.12.: Signifikanz-Faktoren des Netzwerks *NNII* für die verschiedenen Massenhypothesen.

Der gleiche Trend ist in Abbildung 6.12 an den Signifikanzfaktoren zu erkennen. Die Graphen der Massen 20 GeV bis 80 GeV steigen etwas steiler an, die maximale Signifikanzerhöhung bleibt aber annähernd unverändert. Für $m_A = 90$ GeV ist diese etwas zurückgegangen.

Addiert man die Wertungsverteilungen aller Massenhypothesen aufeinander und berechnet danach die ROC-Kurven und Signifikanzen, lässt sich die Leistung der beiden Netzwerke kompakter vergleichen, allerdings auf Kosten der Information über den Beitrag der einzelnen Massen.

Für die Netzwerke *NNI* und *NNII* werden die vollen ROC-Kurven in Abbildung 6.13 abgebildet. Der steilere Anstieg der Kurve von *NNII* ist auch hier zu erkennen. Für *false positive rates* ab ca. 0.15 sind die Unterschiede zwischen den Netzen deutlich geringer und können durch statistische Einflüsse entstanden sein.

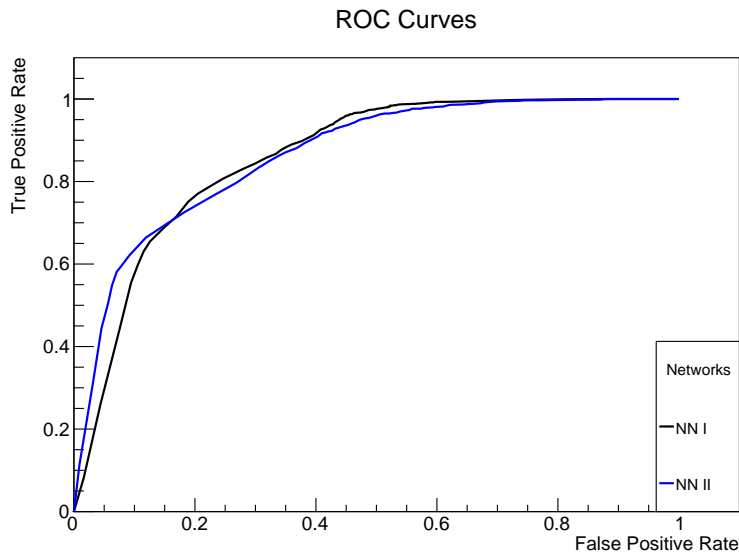


Abbildung 6.13.: Vergleich der vollen ROC-Kurven der Netzwerke *NNI* und *NNII*.

In den vollen Signifikanzfaktoren in Abbildung 6.14 sieht man, dass das *NNII* die Signifikanz deutlich erhöht. Im Mittel steigt die Leistung des Netzes und kompensiert die Verschlechterung für die Massenhypothese 90 GeV.

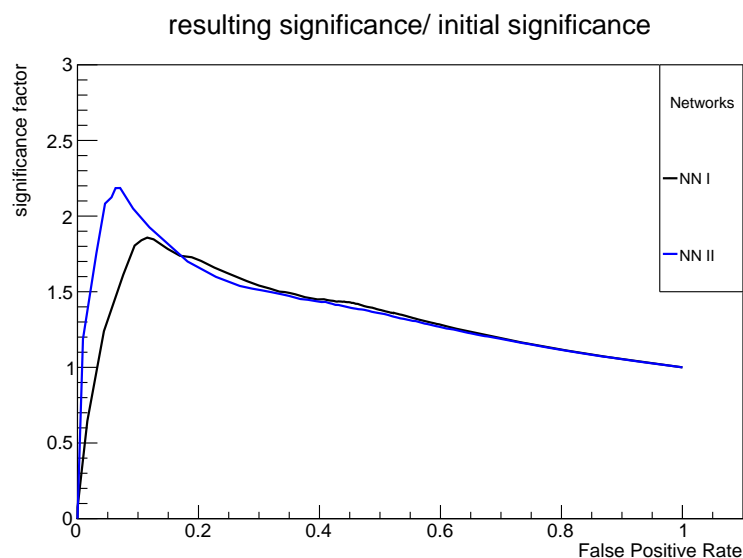


Abbildung 6.14.: Vergleich der vollen Signifikanzen der Netzwerke *NN I* und *NN II*.

6.5. Analyse der kinematischen Variablen

Um mehr Proben für das Netzwerktraining zu erhalten, werden für die folgenden Netzwerke die Ereignisse in der MVA-Region (mit Schnitten wie in Tabelle 5.2) als Datensatz verwendet. In Abbildung 6.15 ist die Verteilung der transversalen Impulse der Myonen in der neuen Region zu sehen. Im Vergleich zu Abbildung 6.5 ist die Zahl der Ereignisse deutlich gestiegen. Alle weiteren Variablenverteilungen sind in Anhang A zu finden.

Anhand der Daten in dieser Region werden die kinematischen Variablen der Simulationen auf ihre Eignung für die Klassifizierung durch das Netz überprüft und die Auswahl überarbeitet. Als Maß für die Güte einer Variable werden hier die Separationskraft der Verteilungen von Untergrund und Signal sowie die Korrelation zwischen den Variablen untereinander verwendet.

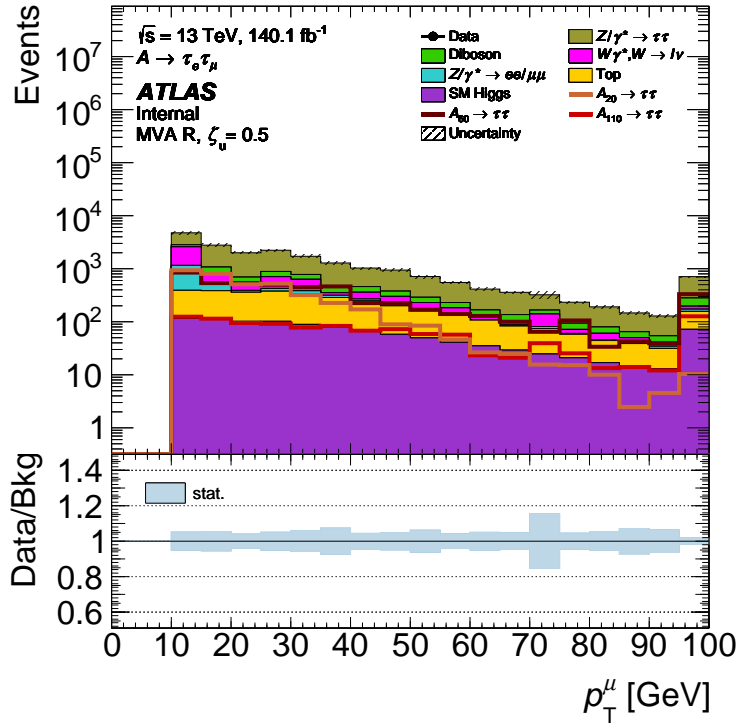


Abbildung 6.15.: Verteilung von p_T^μ der Untergründe und Signale der Massen-hypothesen 20 GeV, 60 GeV und 110 GeV in der MVA-Region.

6.5.1. Separationskraft

Die Separationskraft ist ein Maß für die Unterscheidbarkeit zweier Variablen-Verteilungen und wird in dieser Studie für die Verteilung von Untergrund und Signal berechnet. Je größer der Wert, desto einfacher ist es für das Netzwerk, die Unterschiede in den Charakteristiken zwischen Signal- und Untergrundprozessen zu erkennen. Die Verteilungen von Signal und Untergrund werden für die Berechnung auf die Gesamtzahl der Signal- bzw. Untergründereignisse normiert. Berechnet wird sie anhand der Formel

$$S = \frac{1}{2} \sum_{i=0}^N \frac{(s_i - b_i)^2}{s_i + b_i}. \quad (6.13)$$

Die Summe läuft über die einzelnen Bins der Variablenverteilungen. Die Variablen s_i und b_i sind die relativen Häufigkeiten des Signals und des gesamten Untergrundes in den einzelnen Bins, woraus eine Abhängigkeit der Separationskraft von der Einteilung der Verteilung resultiert. Im Limes einer infinitesimalen Bin-Breite würde der Wert gegen 1 konvergieren. Aus Konsistenzgründen wird deshalb für die Berechnung das Binning der Hauptanalyse verwendet.

Die über alle Massenhypothesen gemittelte Separationskraft für die verschiedenen Variablen ist in Tabelle 6.2 aufgelistet. Eine vollständige Tabelle der Werte für die einzelnen Massenhypothesen ist in Anhang B zu finden.

Tabelle 6.2.: Separationswerte der kinematischen Variablen in der MVA-Region gemittelt über alle Massenhypothesen.

Variable	Mittelwert
E_T^{miss}	0.127
p_T^e	0.026
η_e	0.009
ϕ_e	0.009
p_T^μ	0.030
η_μ	0.008
ϕ_μ	0.008
m_T^{tot}	0.108
m_{MMC}	0.413
p_T^{Higgs}	0.153
m_{ll}	0.109
ΔR_{ll}	0.118
$\Delta\phi_{E_T^{\text{miss}}}^{\text{leadlep}}$	0.007
$\Delta\phi_{E_T^{\text{miss}}}^{\text{subleadlep}}$	0.008
$d_0 \text{ Sig } \mu$	0.013
$d_0 \text{ Sig } e$	0.016
p_T^{leadjet}	0.115
ϕ_{jet}	0.004
η_{jet}	0.003

Zu dem bisherigen Satz von Variablen werden zusätzlich p_T^{leadjet} , der transversale Impuls des Jets mit dem höchsten Impuls, sowie die zugehörigen Winkel ϕ_{jet} und η_{jet} in die Studie aufgenommen. Ebenfalls kommen die d_0 Signifikanzen von Elektron und Myon hinzu, die ein Maß für die transversale Komponente des minimalen Abstands der Teilchenspur zum Primärvertex sind.

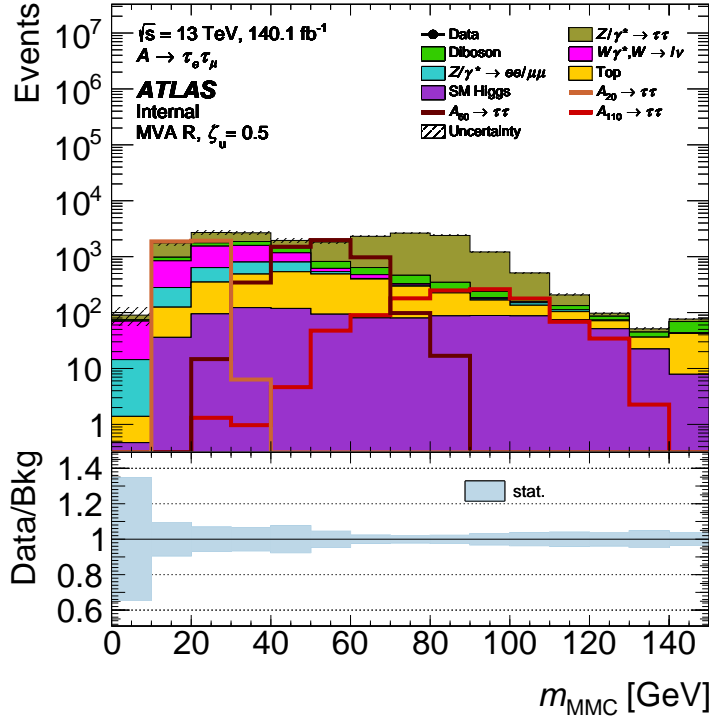


Abbildung 6.16.: Verteilung von m_{MMC} in der MVA-Region für den Untergrund und 3 Massenhypothesen. Für niedrige Massen unterscheiden sich die Graphen stark und sind für das Netzwerk einfacher zu unterscheiden.

Die höchste Separationskraft weist m_{MMC} auf, was an der Verteilung in Abbildung 6.16 gut zu erkennen ist. Die Signale für die verschiedenen Massenhypothesen unterscheiden sich in ihrer Charakteristik deutlich von der Gesamtheit des Untergrundes, für hohe A -Boson-Massen ähneln sich die Verteilungen mehr als für niedrige.

Abbildung 6.17 zeigt die Verteilung von η_{jet} , der Variable mit dem niedrigsten Separationswert. Das Ergebnis der Berechnung spiegelt sich in dem Graphen wider, da die Ereignisse aller Prozesse annähernd gleichverteilt sind. Allgemein ist erkennbar, dass vor allem die Winkelverteilungen der verschiedenen Objekte schlecht separierbar sind. Gut separierbar sind die Massenvariablen, ΔR_{ll} und die transversalen Impulse $p_{\text{T}}^{\text{Higgs}}$ und $p_{\text{T}}^{\text{leadjet}}$.

6.5.2. Korrelation

Der Korrelationskoeffizient K ist eine dimensionslose Größe, die die Stärke des linearen Zusammenhangs zwischen zwei Variablen beschreibt und Werte im Bereich $-1 \leq K \leq +1$ annehmen kann. Wird ein neuronales Netz mit untereinander korrelierten Va-

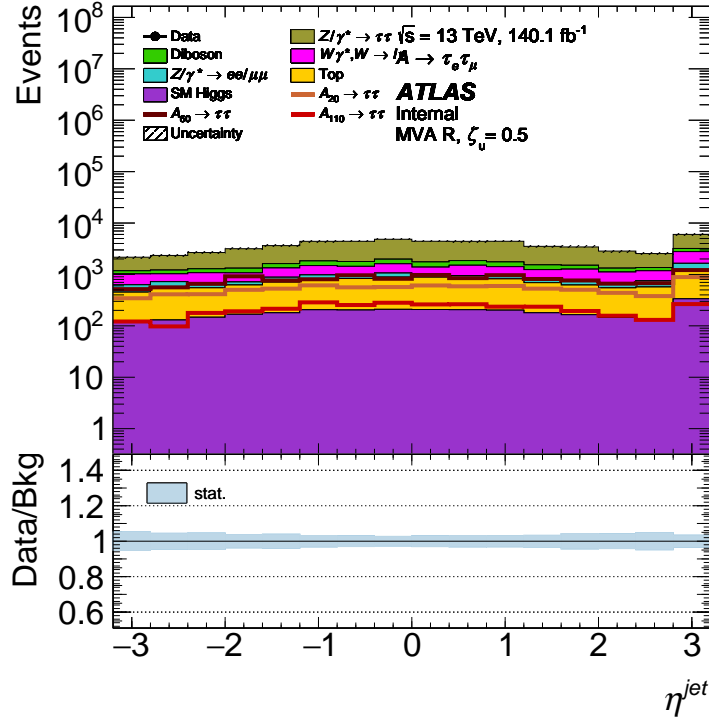


Abbildung 6.17.: Verteilung von η_{jet} in der MVA-Region für Untergrund und 3 Massenhypothesen. Die Ereignisse sind für alle Prozesse ähnlich verteilt und daher schlecht separierbar.

riablen trainiert, wird dessen Kapazität teilweise dafür verwendet, diese Korrelationen zu erkennen. Die Analyse der Korrelationskoeffizienten der kinematischen Variablen soll ein weiterer Anhaltspunkt sein, mit einer optimierten Auswahl der Eingabevariablen die Effizienz und Leistung des Netzwerks zu steigern. Berechnet wird der empirische Korrelationskoeffizient zwischen zwei Variablen x und y mit der Formel

$$K_{xy} = \frac{\langle xy \rangle - \langle x \rangle \langle y \rangle}{\sigma_x \sigma_y}. \quad (6.14)$$

Dabei bezeichnet σ die Standardabweichung der Variablen [26]. Die Koeffizienten werden wie die Separationskraft in der MVA-Region für die erweiterte Variablenauswahl berechnet, allerdings nicht separat für die einzelnen Massenhypothesen, sondern für die Gesamtheit der Signal- und Untergrundproben im Datensatz.

Da nur die Stärke der Korrelation und nicht das Vorzeichen der Koeffizienten relevant ist, sind in Abbildung 6.18 die Beträge der Korrelationskoeffizienten in einer Matrix farblich codiert dargestellt. Die genauen Werte können in Anhang C eingesehen werden. Die Einträge der Hauptdiagonalen sind per Definition 1. Ihre Bestimmung dient daher als Verifizierung der implementierten Berechnungsmethode. Die Variablen der Winkel

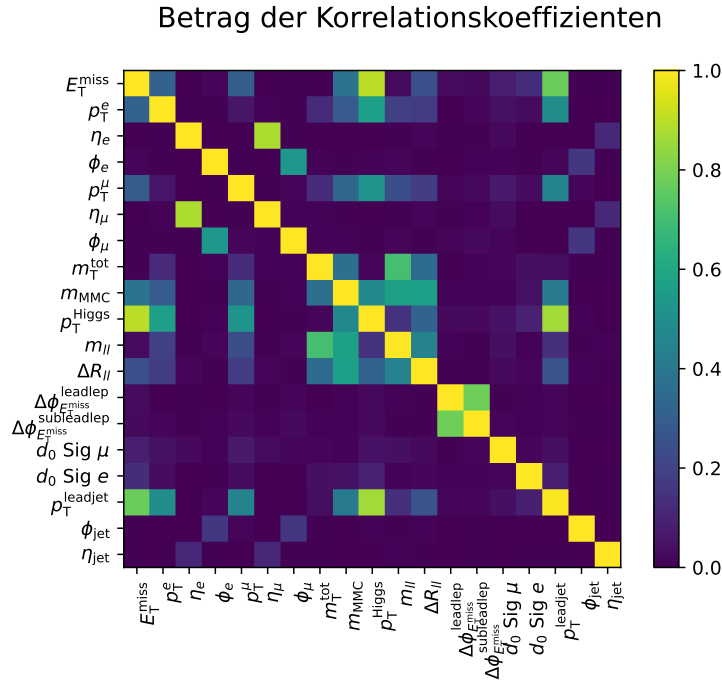


Abbildung 6.18.: Die Beträge der Korrelationskoeffizienten der erweiterten Variablenauswahl.

ϕ und η weisen eine hohe Korrelation zwischen den beiden Leptonen und eine geringe zwischen den der Leptonen und des führenden Jets auf. Das lässt sich auf die Impulserhaltung zurückführen, da für diese die Bewegungsrichtung der Teilchen zueinander relevant ist. Die Impuls- und Massenvariablen sind nahezu alle mindestens geringfügig miteinander korreliert, was sich allgemein durch die Energieerhaltung bei den Kollisionen begründen lässt.

Eine besonders hohe Korrelation liegt zwischen E_T^{miss} , p_T^{Higgs} und p_T^{leadjet} vor, da letztere die größten Impulse in der transversalen Ebene aufweisen. Im Fall von E_T^{miss} und p_T^{Higgs} wird die hohe Korrelation anhand des Plots in Abbildung 6.19 deutlich. Zwischen den Verteilungen beider Variablen liegt ein annähernd linearer Zusammenhang vor.

Im Gegensatz dazu zeigt Abbildung 6.20 einen 2D-Plot für die unkorrelierten Variablen p_T^e und η_e . Beide Verteilungen streuen um einen Häufungspunkt und weisen keine Abhängigkeit zueinander auf.

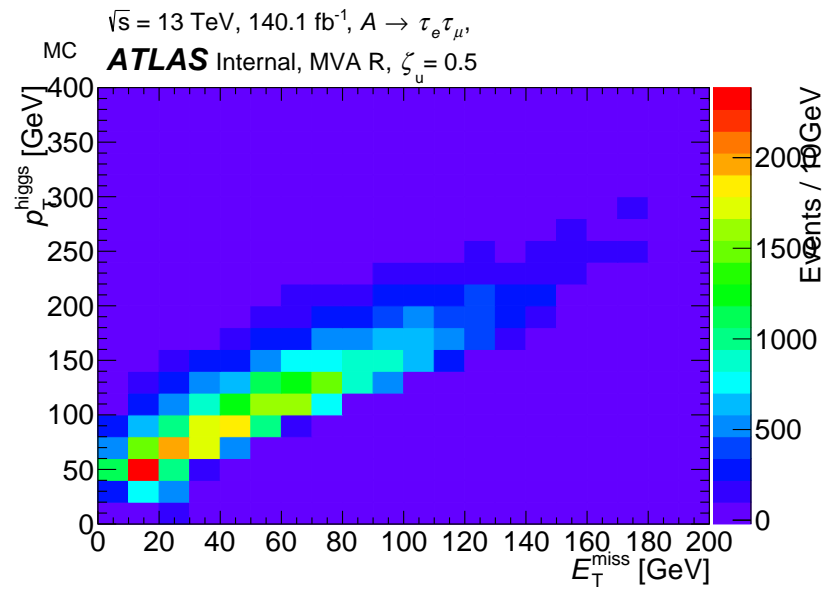


Abbildung 6.19.: 2D-Plot der Variablen E_T^{miss} und p_T^{Higgs} . Die hohe Korrelation der Variablen ist anhand des annähernd linearen Zusammenhangs der Verteilungen zu erkennen.

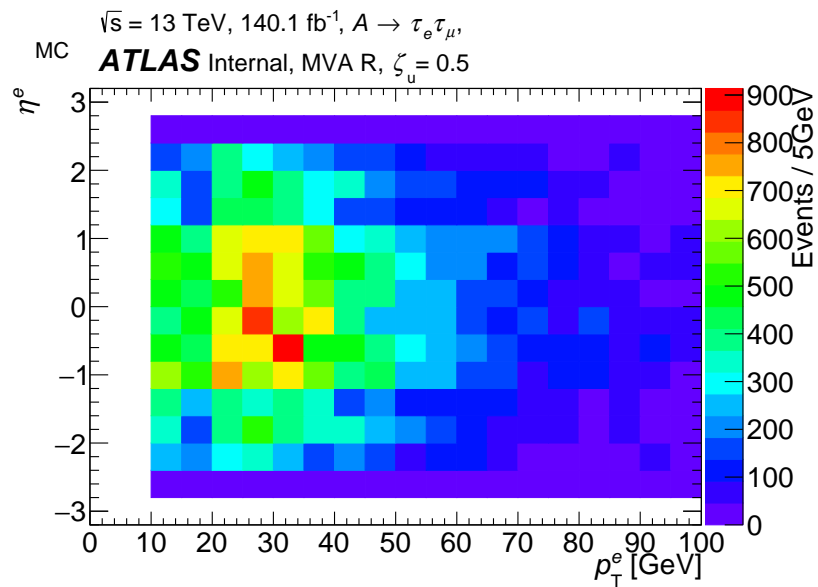


Abbildung 6.20.: 2D-Plot der Variablen p_T^e und η_e . Beide Variablen streuen unabhängig voneinander und es ist keine Tendenz zu einem linearen Zusammenhang erkennbar.

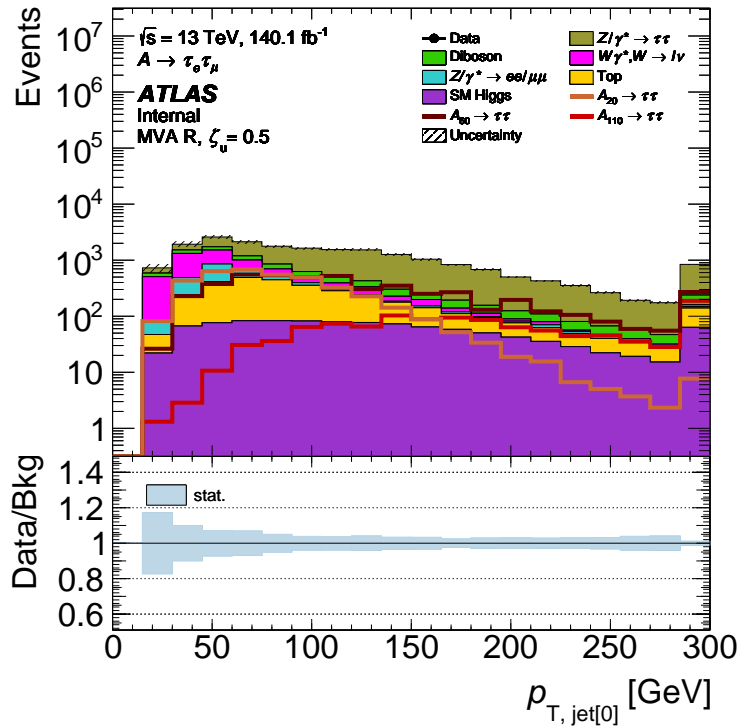


Abbildung 6.21.: Verteilung der Variable p_T^{leadjet} in der MVA-Region. Für steigende A -Boson-Massen unterscheiden sich die Verteilungen der Signale stärker vom Untergrund.

Trotz der hohen Korrelation mit anderen Impulsen und den Massenvariablen wird die Variable p_T^{leadjet} in das Training der Netzwerke aufgenommen, da die mittlere Separationskraft im Vergleich zu den meisten anderen Variablen sehr hoch ist. Die Verteilung der Variable ist in Abbildung 6.21 abgebildet. Weitere Möglichkeiten, die durch die Analyse aufgezeigt werden, wären die Aufnahme der d_0 Signifikanzen und die Entfernung von schlecht separablen Winkelvariablen aus der Variablenauswahl.

6.6. Training in der MVA-Region

Im Folgenden werden zwei weitere Netzwerke mit den Ereignissen der MVA-Region und der um p_T^{leadjet} erweiterten Variablenauswahl trainiert, um zu überprüfen, ob die Leistung des Netzwerks mit den Veränderungen optimiert werden kann.

Das Netzwerk *NNIII* nutzt die gleiche Architektur wie die vorherigen Netzwerke und wurde ursprünglich auch mit der gleichen Lernrate trainiert. Dabei gelangten die Netz-

werke der verschiedenen k-Folds allerdings in lokale Minima mit stark unterschiedlicher Netzwerkleistung, sodass das Netz im Mittel eine geringe Leistung aufwies. Die beste Leistung wurde bei diesem Netzwerk erreicht, in dem die bisherige Lernrate (siehe Abbildung 6.6) mit einer abfallenden Exponentialfunktion multipliziert wurde. Die resultierende Lernrate ist in Abbildung 6.22 dargestellt. Mit dieser Lernrate kann sich das Netzwerk vor allem am Anfang des Trainings aus Minima herausbewegen und ist im späteren Verlauf des Trainings nur noch in der Lage flachere Mulden zu überwinden. Dadurch wird das Training der k-Folds, das für höhere Epochen instabiler wurde, in diesem Bereich kontrollierter.

Abbildung 6.23 zeigt die ROC-Kurven des Netzwerks für die verschiedenen Massen-

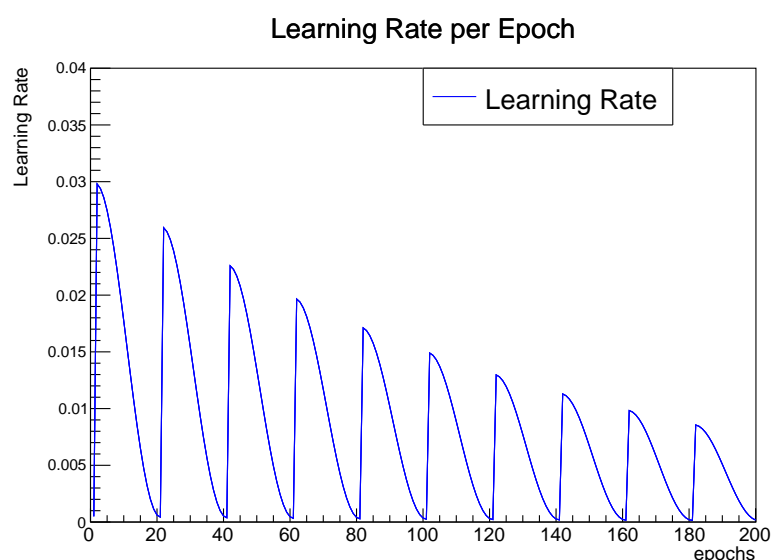


Abbildung 6.22.: Lernrate des Netzwerks *NNIII* in Abhängigkeit der Trainingsepoche.

hypothesen. Die Kurven von 20 GeV, 70 GeV und 110 GeV haben sich im Vergleich zu *NNII* leicht verbessert, für alle anderen *A*-Boson-Massen findet eine leicht verschlechterte Klassifizierung statt. Diese Tendenz wird auch anhand der Signifikanzfaktoren in Abbildung 6.24 deutlich.

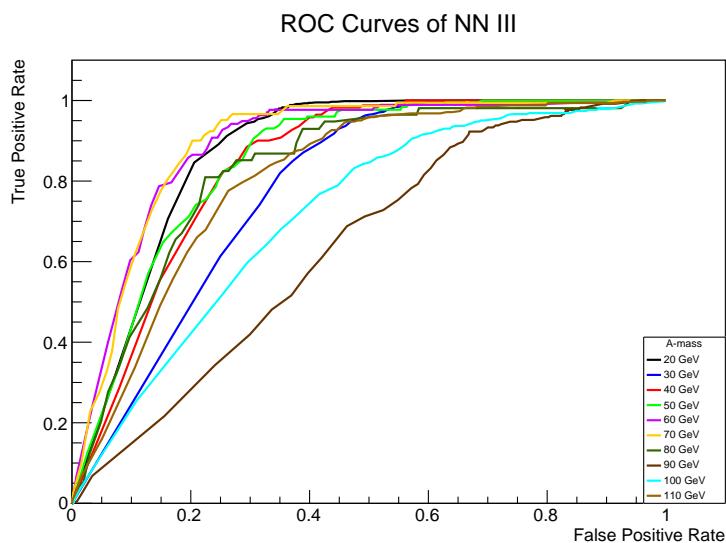


Abbildung 6.23.: ROC-Kurven des Netzwerks *NN III* für die verschiedenen Massenhypothesen.

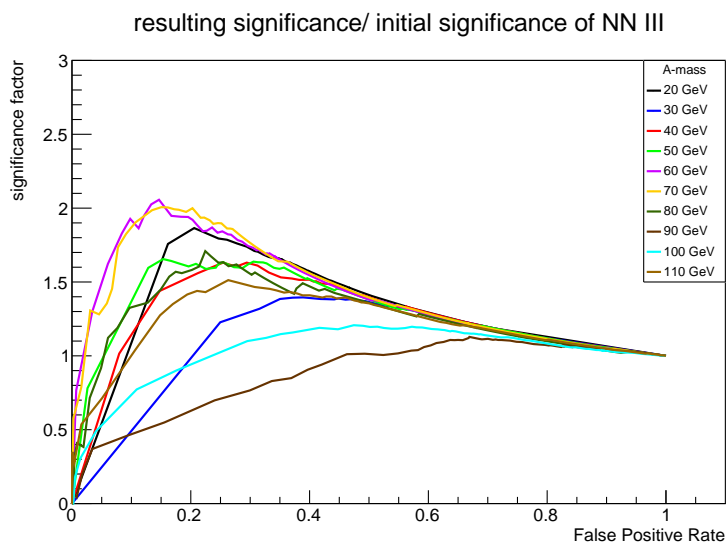


Abbildung 6.24.: Signifikanzfaktoren des Netzwerks *NN III* für die verschiedenen Massenhypothesen.

In Abbildung 6.25 ist die volle ROC-Kurve im Vergleich zu den anderen Netzwerken abgebildet. Diese weist eine leicht größere Wölbung in Richtung eines perfekten Klassifizierers auf, stellt aber keine signifikante Verbesserung der Leistung dar. Das wird auch im Vergleich der vollen Signifikanzfaktoren deutlich, die in Abbildung 6.26 dargestellt sind. Der größte erreichte Signifikanzfaktor ist im Vergleich zu *NN II* fast unverändert.

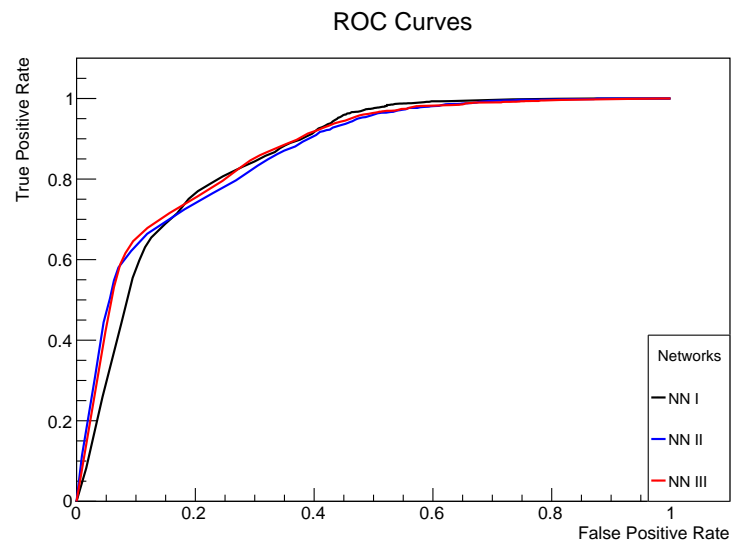


Abbildung 6.25.: Vergleich der vollen ROC-Kurven der Netzwerke *NN I*, *NN II* und *NN III*.

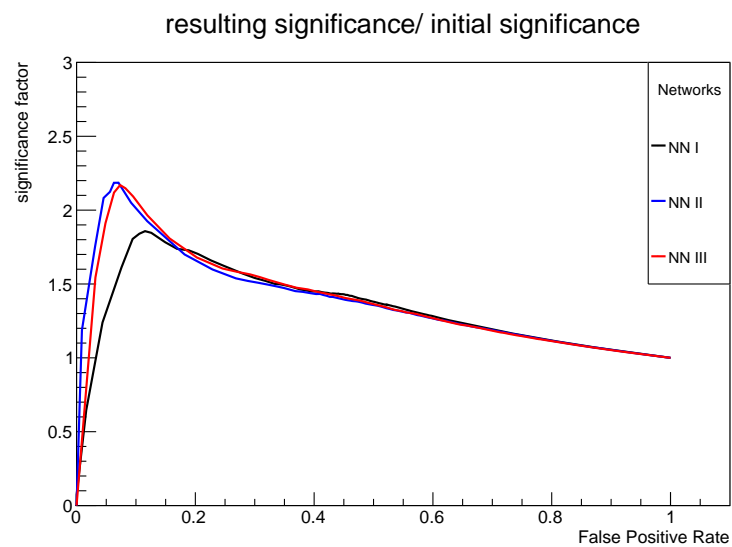


Abbildung 6.26.: Vergleich der vollen Signifikanzfaktoren der drei Netzwerke. Netzwerk *NN III* weist an dessen bestem Arbeitspunkt keine Erhöhung der Signifikanz auf.

6.7. Erhöhung der Netzwerkkapazität

Obwohl dem Netzwerk mehr Informationen für eine Klassifizierung zur Verfügung stehen, verbessert sich dessen Leistung nicht, weshalb im Folgenden die Architektur des Netzes verändert wird. Statt den bisher drei versteckten Schichten mit je zehn Neuronen werden nun vier versteckte Schichten mit je 15 Neuronen verwendet um zu überprüfen, ob das Netz *NNIV* mit erhöhter Kapazität die zusätzlichen Informationen besser verarbeiten kann.

Mit der Konfiguration der anderen Parameter wie bei *NNIII* war das Training erneut zu instabil und mindestens einer der 5 k-Folds entwickelte sich zu einem zufälligen Klassifizierer. Um die Aktualisierung der Parameter gerichteter und stabiler durchzuführen, wurde in weiteren Versuchen die Größe *mini-batches* auf 150 Proben angehoben, was das Problem jedoch nicht beseitigen konnte. Das einzige stabile Ergebnis wurde erzielt, indem das Netz im lokalen Minimum der Initialisierung gehalten wurde. Dabei wurde eine exponentiell abfallende Lernrate verwendet, die in Abbildung 6.27 zu sehen ist. Zusätzlich wurde die Zahl der Trainingsepochen auf 80 reduziert, da das Netzwerk nach der „Bewegung“ in das lokale Minimum zu *overfitten* begann.

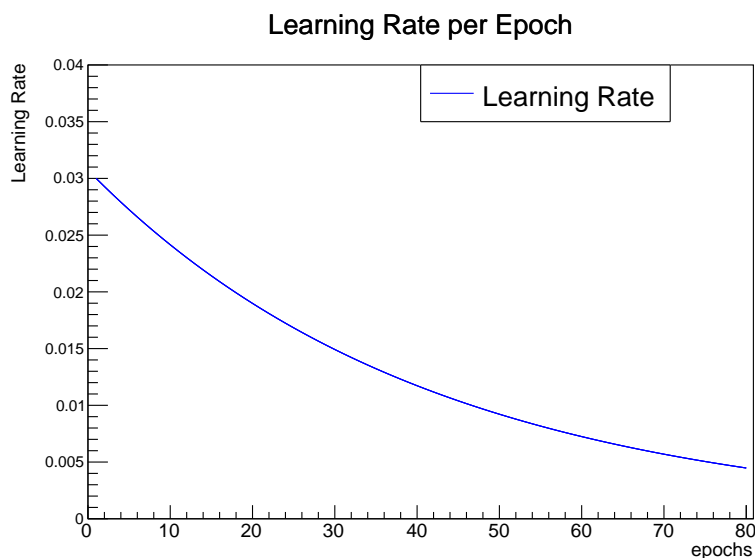


Abbildung 6.27.: Lernrate von *NNIV* in Abhängigkeit der Trainingsepoch.

In diesem lokalen Minimum bietet die erweiterte Kapazität des Netzes aufgrund der schlechten Optimierung keinen Vorteil. Die Form des Trainings sorgt sogar für eine Verschlechterung der Leistung für alle Massenhypothesen, was an den ROC-Kurven in Abbildung 6.28 und den Signifikanzfaktoren in Abbildung 6.29 zu erkennen ist.

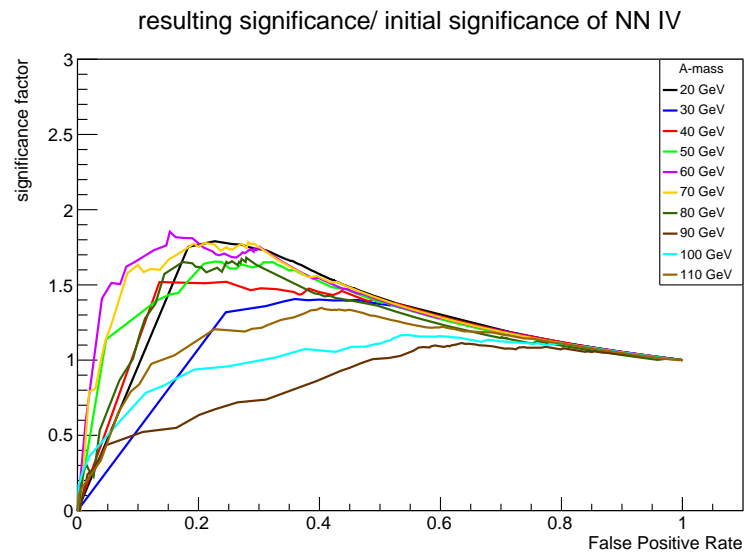


Abbildung 6.29.: Signifikanzfaktoren von *NNIV* für die verschiedenen Massenhypothesen.

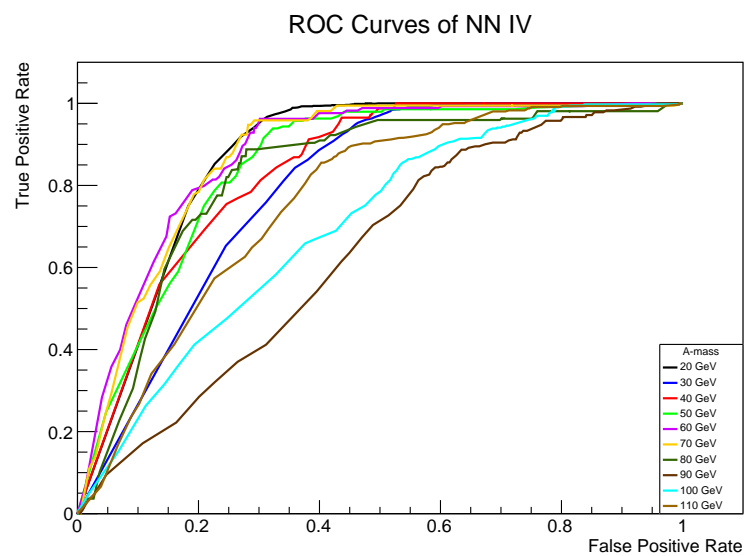


Abbildung 6.28.: ROC-Kurven von *NNIV* für die verschiedenen Massenhypothesen.

Im Vergleich der vollen ROC-Kurven, zu sehen in Abbildung 6.30, sieht der Unterschied von *NNIV* zu *NNII* und *NNIII* gering aus. Anhand der vollen Signifikanzfaktoren in Abbildung 6.31 sieht man hingegen, dass die maximale Signifikanzhöhung von *NNIV*

im Vergleich zu den beiden vorherigen Netzwerken um etwa 10% abgenommen hat. Die Erhöhung der Netzwerkkapazität konnte hier den gewünschten Effekt nicht erzielen.

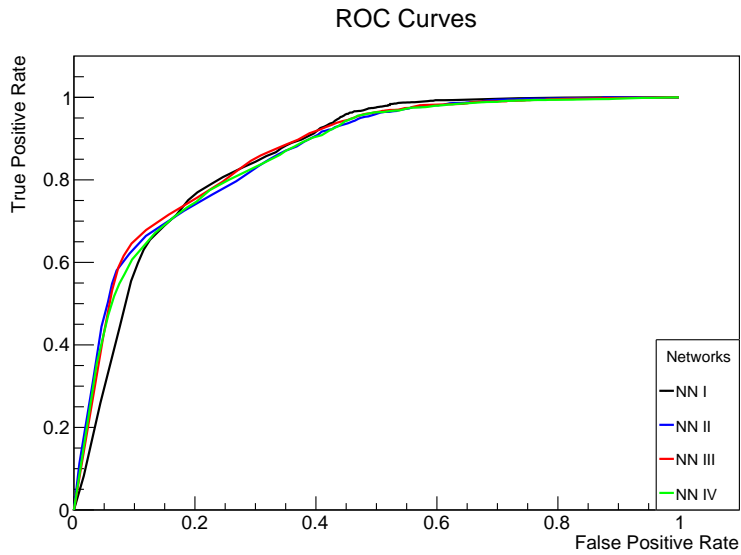


Abbildung 6.30.: Die vollen ROC-Kurven aller Netzwerke im Vergleich miteinander.

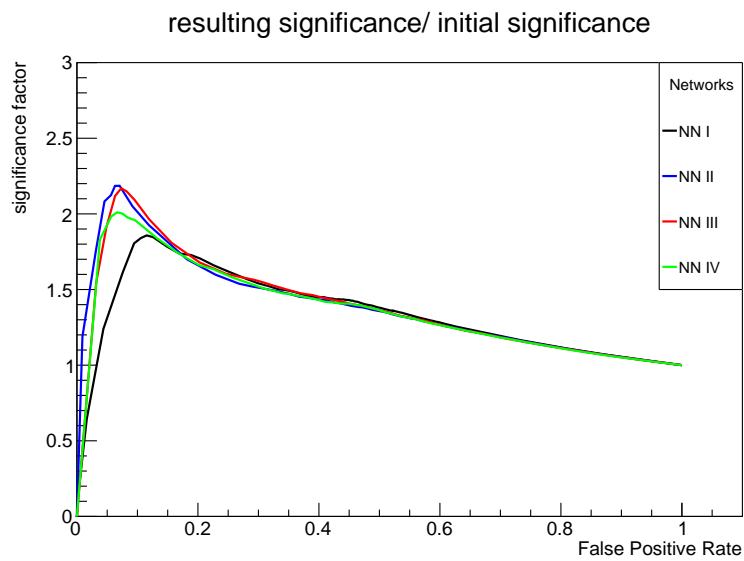


Abbildung 6.31.: Die vollen Signifikanzfaktoren aller Netzwerke im Vergleich miteinander.

6.8. Weitere Untersuchungen

In diesem Kapitel wird kurz auf zwei weitere Optimierungsversuche eingegangen, die aufgrund der schlechten Netzwerkleistung nicht vollständig analysiert wurden. Motiviert wurden diese Untersuchungen durch die Probleme bei der Klassifizierung für hohe Werte für m_A durch die bisherigen Netzwerke.

Im ersten Versuch wurde das Massenspektrum des A -Bosons entsprechend der Low-Mass und High-Mass Signalregionen aufgeteilt und ein Netzwerk trainiert, das nur auf die Massen 80 GeV, 90 GeV, 100 GeV und 110 GeV sensitiv sein soll, um eine bessere Klassifizierung in diesem Massenbereich zu ermöglichen. Aus dem Trainingsdatensatz wurden deshalb die MC-Proben der anderen Massen entfernt. Es wurde die gleiche Lernrate wie für $NNIV$ verwendet (siehe Abbildung 6.27), allerdings wurde das Netzwerk nur über 100 Epochen trainiert.

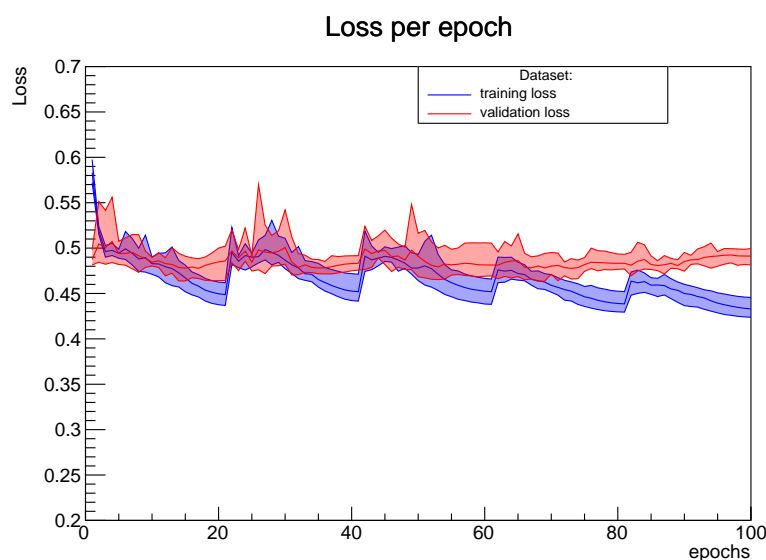


Abbildung 6.32.: Verlust des High-Mass Netzwerks im Trainingsverlauf.

Die Werte der Verlustfunktion des Netzwerks während des Trainings sind in Abbildung 6.32 zu sehen. Der rote Graph beschreibt den Verlust bezüglich des Validierungs-, der blaue bezüglich des Trainingsdatensatzes. Die obere und untere Begrenzung der eingefärbten Flächen repräsentieren den besten und schlechtesten k -Fold des Netzes. Der Graph im Zentrum der Flächen ist der Mittelwert über alle k Netzwerke. Während der Verlust bezüglich des Trainingsdatensatzes leicht abnimmt, hat die Veränderung in der Validierung einen ungerichteten, zufälligen Charakter. Zudem setzt nach Epoche 60 das *overfitting* ein. Das Netzwerk lernt nur die Trainingsdaten auswendig, klassifiziert den

Untergrund annähernd zufällig und zeigt keine Tendenzen eines Lerneffekts.

Im zweiten Versuch wurde ein Netzwerk trainiert, das nur auf die Massenhypothese $m_A = 90 \text{ GeV}$ sensitiviert werden sollte. Dafür wurde die Auswahl der kinematischen Variablen nicht verändert, aber keine A -Boson-Masse mehr an das Netz übergeben. Als Datensatz wurde die Gesamtheit der Untergrundproben sowie die MC-Simulationen des Signalprozesses für $m_A = 90 \text{ GeV}$ verwendet. Es wurde erneut die Lernrate des Netzwerks $NNIV$ verwendet.

Abbildung 6.33 zeigt die AUC des Netzwerks während des Trainings. Innerhalb weniger Epochen ist die AUC des Trainingsdatensatzes deutlich höher als die der Validierung, was auf *overfitting* hindeutet. Für die Validierungsdaten findet im Mittel kaum eine Optimierung statt. Ab der ersten Erhöhung der Lernrate wird das Training instabil und mindestens einer der k-Folds wird zum zufälligen Klassifizierer. Im weiteren Verlauf verschlechtert sich die Leistung bezüglich der Trainingsdaten ebenfalls. Beim Versuch, das Netz im Minimum der Initialisierung zu optimieren, begann schon ab wenigen Epochen extremes *overfitting*.

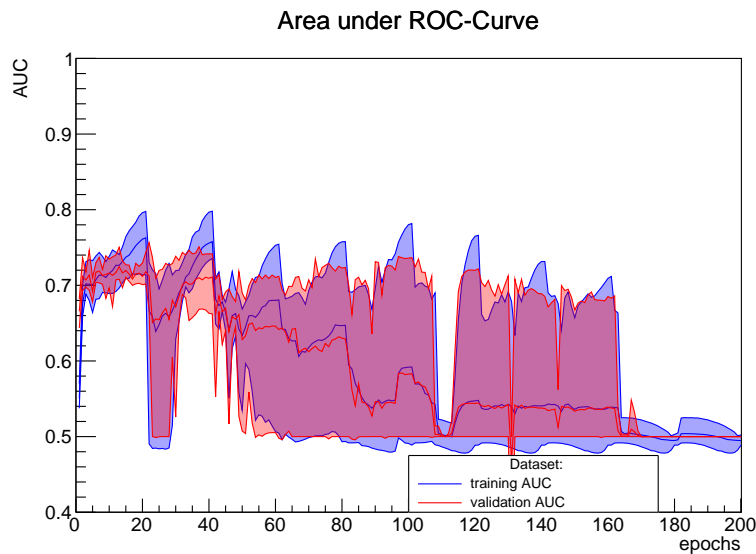


Abbildung 6.33.: AUC eines Netzwerks, das den Signalprozesse für $m_A = 90 \text{ GeV}$ vom Untergrund unterscheiden soll.

Die beiden Herangehensweisen an eine Verbesserung der Klassifizierung bei hohen A -Boson-Massen sind anhand der in dieser Arbeit untersuchten Konfigurationen nicht als aussichtslos anzusehen. Die gezeigten Grafiken stellen die besten erreichten Ergebnisse

der Studien in dieser Arbeit dar, könnten aber mit einer gezielte Anpassung der Netzwerkparameter oder der kinematischen Variablen verbessert werden.

7. Zusammenfassung

Diese Arbeit beschäftigt sich mit der Optimierung der Suche nach dem vom 2HDM vorhergesagten A -Boson mit Hilfe von neuronalen Netzwerken. Die Erweiterung des Massenspektrums, auf das das Netzwerk sensitiviert werden soll, ist die erste vorgenommene Änderung. Diese hat im Rahmen dieser Analyse nur einen geringen Einfluss auf die mittlere Leistung, bestätigt aber ein Problem, das schon in der Masterarbeit von Jannik Friese [28] deutlich wurde: die schlechte Unterscheidung von Z -Untergrund und Signal für die Massen um 90 GeV. Die bisherige Konfiguration an Netzwerkparametern und Trainingsdaten lässt keine ausreichend gute Unterscheidung zwischen den Prozessen zu. Die Analyse der kinematischen Variablen der Kollisionsprodukte zeigt, dass in der Variablenauswahl noch Potential für Verbesserung steckt. Vor allem ϕ und η weisen eine geringe Separationskraft auf. Für die Massenvariablen wird in der Analyse jedoch auch klar, dass die Separationskraft abhängig von dem untersuchten Signalprozess ist. Eine Aufteilung in Netzwerke für niedrige und hohe Massenhypothesen mit einer separaten Variablenauswahl könnte primär die Leistung für hohe Massen verbessern.

Eine weitere Erkenntnis dieser Arbeit ist die Sensitivität der Leistung auf die Zusammensetzung der Trainingsdaten. Der niedrigere relative Anteil an Proben für Untergrundprozesse hat die Leistung für $m_A = 90$ GeV ebenfalls verschlechtert. Eine Verbesserung könnte man erreichen, indem beim Netzwerktraining die Beiträge der einzelnen Proben zum Gradienten und zum Verlust mit dem zugehörigen Gewicht der MC-Simulation skaliert werden. Das würde die Zusammensetzung der Trainingsdaten an die Daten anpassen, auf die das Netzwerk später angewendet wird.

Die Erweiterung der Variablenauswahl führte im Fall von $NNIII$ zu keiner Erhöhung der Netzwerkleistung. Allgemein kann aber noch nicht davon ausgegangen werden, dass in dieser Veränderung kein Verbesserungspotential liegt, da das Training von $NNIV$ bisher noch nicht effektiv durchgeführt werden konnte.

Diese Arbeit liefert zwar keine deutliche Verbesserung des bisherigen Netzwerks, motiviert jedoch weitere Anpassungen und ermöglicht eine Orientierung an den Ergebnissen der Variablenanalyse für weitere Studien. Signifikante Verbesserungen könnten erzielt werden, wenn das Problem der Klassifizierung für $m_A = 90$ GeV und $m_A = 100$ GeV gelöst wird.

8. Literaturverzeichnis

- [1] Keras: Deep Learning for humans, . URL <https://keras.io/>. Aufgerufen am 10.05.2023.
- [2] LuminosityPublicResultsRun2 < AtlasPublic < TWiki, . URL https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2#Online_Luminosity_Summary_Plots. Aufgerufen am: 04.05.2023.
- [3] The Standard Model of particle physics is brilliant and completely flawed - ABC News, . URL <https://www.abc.net.au/news/science/2017-07-15/the-standard-model-of-particle-physics-explained/7670338>. Aufgerufen am: 10.04.2023.
- [4] Fläche-unter-der-ROC-Kurve (AUC) | SAP Help Portal. URL https://help.sap.com/docs/SAP_ANALYTICS_CLOUD/00f68c2e08b941f081002fd3691d86a7/235c79933a7b4f398369e23a04520a3e.html. Aufgerufen am: 09.05.2023.
- [5] B. Abi et al. Measurement of the positive muon anomalous magnetic moment to 0.46 ppm. *Physical Review Letters*, 126(14), apr 2021. doi: 10.1103/physrevlett.126.141801. URL <https://doi.org/10.1103%2Fphysrevlett.126.141801>.
- [6] ALEPH Collaboration. Measurement of the Z resonance parameters at LEP. *The European Physical Journal C - Particles and Fields*, 14(1):1–50, May 2000. ISSN 1434-6052. doi: 10.1007/s100520000319. URL <https://doi.org/10.1007/s100520000319>.
- [7] U. Amaldi, W. de Boer, and H. Fürstenau. Comparison of grand unified theories with electroweak and strong coupling constants measured at LEP. *Phys. Lett. B*, 260:447–455, 1991. doi: 10.1016/0370-2693(91)91641-8. URL <https://cds.cern.ch/record/217892>.
- [8] P. Athron, C. Balázs, D. H. J. Jacob, W. Kotlarski, D. Stöckinger, and H. Stöckinger-Kim. New physics explanations of a in light of the FNAL muon g - 2 measurement. *Journal of High Energy Physics*, 2021(9), sep 2021. doi: 10.1007/jhep09(2021)080. URL <https://doi.org/10.1007%2Fjhep09%282021%29080>.

- [9] ATLAS Collaboration. The atlas experiment at the cern large hadron collider. *Journal of Instrumentation*, 3(08):S08003, aug 2008. doi: 10.1088/1748-0221/3/08/S08003. URL <https://dx.doi.org/10.1088/1748-0221/3/08/S08003>.
- [10] ATLAS Collaboration. The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim. Technical report, CERN, Geneva, 2010. URL <https://cds.cern.ch/record/1300517>.
- [11] ATLAS Collaboration. Observation of a new particle in the search for the standard model higgs boson with the ATLAS detector at the LHC. *Physics Letters B*, 716(1):1–29, sep 2012. doi: 10.1016/j.physletb.2012.08.020. URL <https://doi.org/10.1016%2Fj.physletb.2012.08.020>.
- [12] ATLAS Collaboration. ATLAS Insertable B-Layer Technical Design Report Addendum. Technical report, 2012. URL <https://cds.cern.ch/record/1451888>. Addendum to CERN-LHCC-2010-013, ATLAS-TDR-019.
- [13] ATLAS Collaboration. Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s} = 13$ TeV. *The European Physical Journal C*, 76(5), may 2016. doi: 10.1140/epjc/s10052-016-4120-y. URL <https://doi.org/10.1140%2Fepjc%2Fs10052-016-4120-y>.
- [14] ATLAS Collaboration. Electron reconstruction and identification in the atlas experiment using the 2015 and 2016 LHC proton–proton collision data at $\sqrt{s} = 13$ TeV. *The European Physical Journal C*, 79(8), august 2019. doi: 10.1140/epjc/s10052-019-7140-6. URL <https://doi.org/10.1140/epjc/s10052-019-7140-6>.
- [15] ATLAS Collaboration. Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC, 2022.
- [16] E. Bothmann, G. S. Chahal, S. Höche, J. Krause, F. Krauss, S. Kuttimalai, S. Liebschner, D. Napoletano, M. Schönherr, H. Schulz, S. Schumann, and F. Siegert. Event generation with Sherpa 2.2. *SciPost Phys.*, 7:034, 2019. doi: 10.21468/SciPostPhys.7.3.034. URL <https://scipost.org/10.21468/SciPostPhys.7.3.034>.
- [17] C. P. Burgess and G. D. Moore. *The standard model: a primer*. Cambridge University Press, 2007. ISBN 9780521860369.
- [18] CERN Collaboration. LHC the guide. 2021. URL <https://cds.cern.ch/record/2809109/files/CERN-Brochure-2021-004-Eng.pdf>.
- [19] A. Cherchiglia, D. Stöckinger, and H. Stöckinger-Kim. Muon g-2 in the 2hdm: Maximum results and detailed phenomenology. *Physical Review D*, 98(3), aug 2018.

- doi: 10.1103/physrevd.98.035001. URL <https://doi.org/10.1103%2Fphysrevd.98.035001>.
- [20] C. De Melis. The CERN accelerator complex. Complexe des accélérateurs du CERN. 2016. URL <https://cds.cern.ch/record/2119882>. General Photo.
- [21] DELPHI Collaboration. Searches for neutral higgs bosons in extended models. *The European Physical Journal C - Particles and Fields*, 38:1–28, dec 2004. doi: 10.1140/epjc/s2004-02011-4. URL <https://doi.org/10.1140/epjc/s2004-02011-4>.
- [22] A. Elagin, P. Murat, A. Pranko, and A. Safonov. A new mass reconstruction technique for resonances decaying to . *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 654(1):481–489, 2011. ISSN 0168-9002. doi: <https://doi.org/10.1016/j.nima.2011.07.009>. URL <https://www.sciencedirect.com/science/article/pii/S0168900211014112>.
- [23] J. Ellis, M. K. Gaillard, and D. V. Nanopoulos. A historical profile of the higgs boson. In *Advanced Series on Directions in High Energy Physics*, pages 255–274. WORLD SCIENTIFIC, aug 2016. doi: 10.1142/9789814733519_0014. URL https://doi.org/10.1142%2F9789814733519_0014.
- [24] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.*, 13:321–323, Aug 1964. doi: 10.1103/PhysRevLett.13.321. URL <https://link.aps.org/doi/10.1103/PhysRevLett.13.321>.
- [25] L. Evans and P. Bryant. LHC Machine. *Journal of Instrumentation*, 3(08):S08001, aug 2008. doi: 10.1088/1748-0221/3/08/S08001. URL <https://dx.doi.org/10.1088/1748-0221/3/08/S08001>.
- [26] L. Fahrmeir, C. Heumann, R. Künstler, I. Pigeot, and G. Tutz. *Statistik*. Springer-Lehrbuch. Springer Berlin Heidelberg, Berlin, Heidelberg, 2016. ISBN 978-3-662-50371-3 978-3-662-50372-0. doi: 10.1007/978-3-662-50372-0. URL <http://link.springer.com/10.1007/978-3-662-50372-0>.
- [27] K. Freese. Status of dark matter in the universe. *International Journal of Modern Physics D*, 26(06):1730012, May 2017. ISSN 0218-2718, 1793-6594. doi: 10.1142/S0218271817300129. URL <https://www.worldscientific.com/doi/abs/10.1142/S0218271817300129>.
- [28] J. Friese. Search for a light CP-odd Higgs boson decaying into a pair of τ leptons. Masterarbeit, TU Dresden, 2022. URL <https://iktp.tu-dresden.de/IKTP/pub/>

22/DPG_2022_Jannik_Friese.pdf.

- [29] GEANT4 Collaboration. Geant4—a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250–303, 2003. ISSN 0168-9002. doi: [https://doi.org/10.1016/S0168-9002\(03\)01368-8](https://doi.org/10.1016/S0168-9002(03)01368-8). URL <https://www.sciencedirect.com/science/article/pii/S0168900203013688>.
- [30] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. MIT Press, 2016. <http://www.deeplearningbook.org>.
- [31] J. F. Gunion and H. E. Haber. iCP/i-conserving two-higgs-doublet model: The approach to the decoupling limit. *Physical Review D*, 67(7), apr 2003. doi: 10.1103/PhysRevD.67.075019. URL <https://doi.org/10.1103/PhysRevD.67.075019>.
- [32] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. Global conservation laws and massless particles. *Phys. Rev. Lett.*, 13:585–587, Nov 1964. doi: 10.1103/PhysRevLett.13.585. URL <https://link.aps.org/doi/10.1103/PhysRevLett.13.585>.
- [33] P. W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13:508–509, Oct 1964. doi: 10.1103/PhysRevLett.13.508. URL <https://link.aps.org/doi/10.1103/PhysRevLett.13.508>.
- [34] P. W. Higgs. Spontaneous symmetry breakdown without massless bosons. *Phys. Rev.*, 145:1156–1163, May 1966. doi: 10.1103/PhysRev.145.1156. URL <https://link.aps.org/doi/10.1103/PhysRev.145.1156>.
- [35] T. W. B. Kibble. Symmetry breaking in non-abelian gauge theories. *Phys. Rev.*, 155:1554–1561, Mar 1967. doi: 10.1103/PhysRev.155.1554. URL <https://link.aps.org/doi/10.1103/PhysRev.155.1554>.
- [36] T. Kresse. Search for a light CP-odd Higgs boson decaying into a pair of τ -Leptons with the ATLAS detector. Masterarbeit, TU Dresden, 2020. URL https://iktp.tu-dresden.de/IKTP/pub/20/Masterarbeit_Tom_Kresse.pdf.
- [37] P. Moder. Search for a light cp-odd higgs boson decaying into a $\tau\tau$ pair. Masterarbeit, TU Dresden, 2018. URL https://iktp.tu-dresden.de/IKTP/pub/18/Masterarbeit_Moder.pdf.
- [38] T. Nakano and K. Nishijima. Charge Independence for V-particles*. *Progress of Theoretical Physics*, 10(5):581–582, 11 1953. ISSN 0033-068X. doi: 10.1143/PTP.10.581. URL <https://doi.org/10.1143/PTP.10.581>.

- [39] M. Nielsen. *Neural Networks and Deep Learning*. 2015.
- [40] C. Oleari. The powheg box. *Nuclear Physics B - Proceedings Supplements*, 205-206:36–41, 2010. ISSN 0920-5632. doi: <https://doi.org/10.1016/j.nuclphysbps.2010.08.016>. URL <https://www.sciencedirect.com/science/article/pii/S0920563210001994>. Loops and Legs in Quantum Field Theory.
- [41] C. Patrignani. Review of particle physics. *Chinese Physics C*, 40(10):100001, oct 2016. doi: 10.1088/1674-1137/40/10/100001. URL <https://dx.doi.org/10.1088/1674-1137/40/10/100001>.
- [42] A. Pich and P. Tuzó n. Yukawa alignment in the two-higgs-doublet model. *Physical Review D*, 80(9), nov 2009. doi: 10.1103/physrevd.80.091702. URL <https://doi.org/10.1103%2Fphysrevd.80.091702>.
- [43] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands. An introduction to pythia 8.2. *Computer Physics Communications*, 191:159–177, 2015. ISSN 0010-4655. doi: <https://doi.org/10.1016/j.cpc.2015.01.024>. URL <https://www.sciencedirect.com/science/article/pii/S0010465515000442>.
- [44] J. Thomson. Cathode rays. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 44(269):293–316, 1897. doi: 10.1080/14786449708621070. URL <https://doi.org/10.1080/14786449708621070>.
- [45] M. Thomson. *Modern particle physics*. Cambridge University Press, Cambridge, United Kingdom ; New York, 2013. ISBN 9781107034266.
- [46] M. Trodden. Electroweak baryogenesis. *Reviews of Modern Physics*, 71(5):1463–1500, oct 1999. doi: 10.1103/revmodphys.71.1463. URL <https://doi.org/10.1103%2Frevmodphys.71.1463>.
- [47] S. Weinberg. A model of leptons. *Phys. Rev. Lett.*, 19:1264–1266, Nov 1967. doi: 10.1103/PhysRevLett.19.1264. URL <https://link.aps.org/doi/10.1103/PhysRevLett.19.1264>.
- [48] C. N. Yang and R. L. Mills. Conservation of isotopic spin and isotopic gauge invariance. *Phys. Rev.*, 96:191–195, Oct 1954. doi: 10.1103/PhysRev.96.191. URL <https://link.aps.org/doi/10.1103/PhysRev.96.191>.

A. Zusätzliche Plots in der MVA-Region

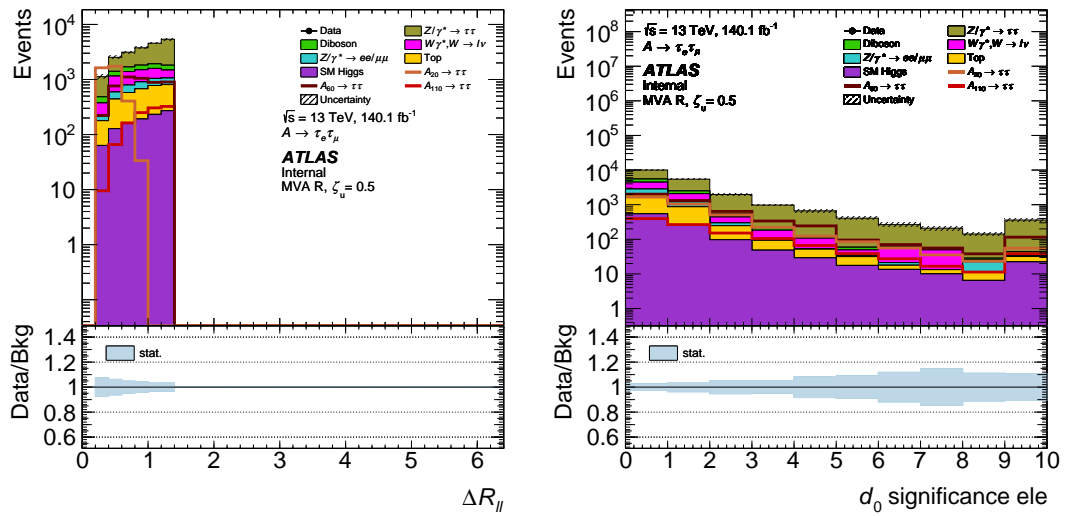


Abbildung A.1.: Verteilung von $\Delta R_{\ell\ell}$ (links) und d_0 Signifikanz e (rechts) in der MVA-Region.

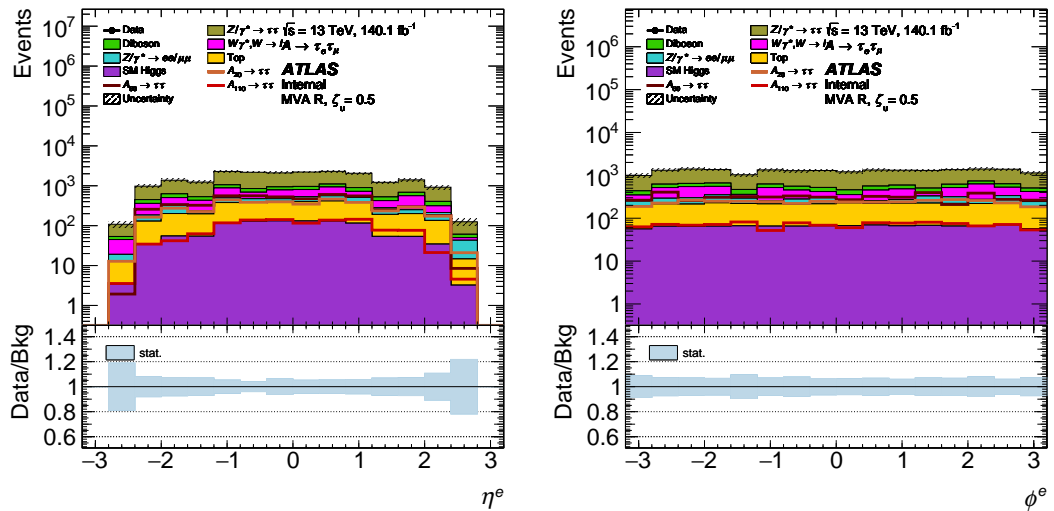


Abbildung A.2.: Verteilung von η_e (links) und ϕ_e (rechts) in der MVA-Region.

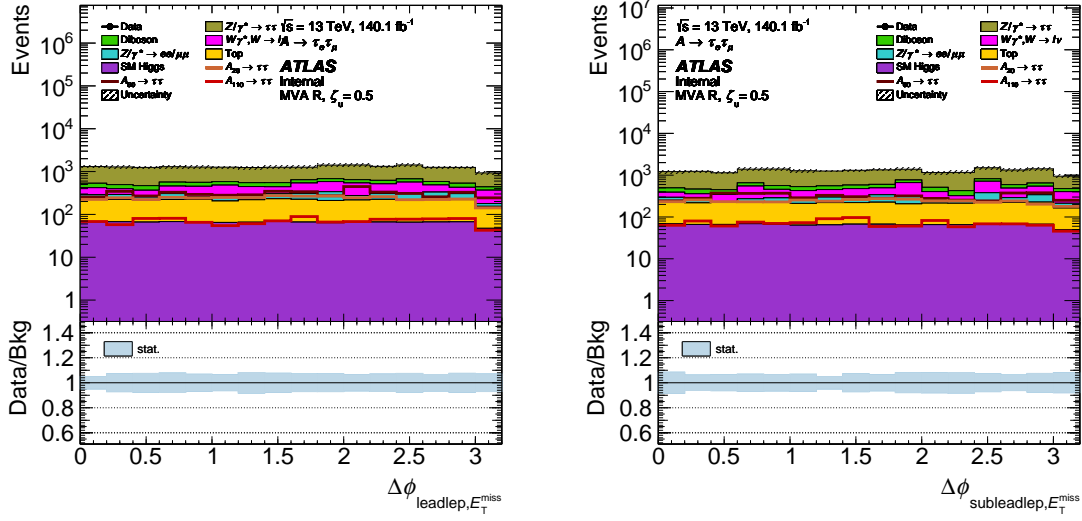


Abbildung A.3.: Verteilung von $\Delta\phi_{E_T^{\text{miss}}}^{\text{leadlep}}$ (links) und $\Delta\phi_{E_T^{\text{miss}}}^{\text{subleadlep}}$ (rechts) in der MVA-Region.

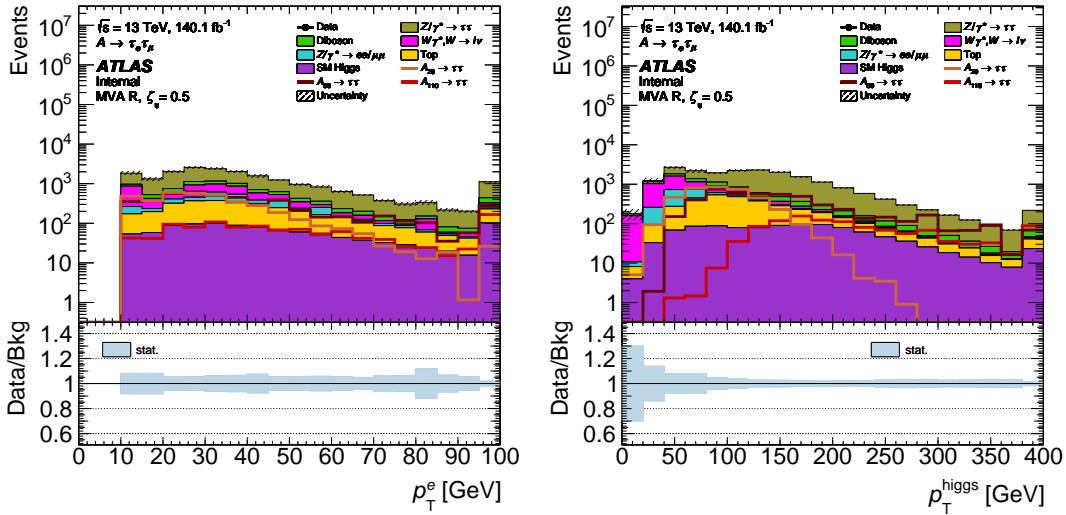


Abbildung A.4.: Verteilung von p_T^e (links) und p_T^{Higgs} (rechts) in der MVA-Region.

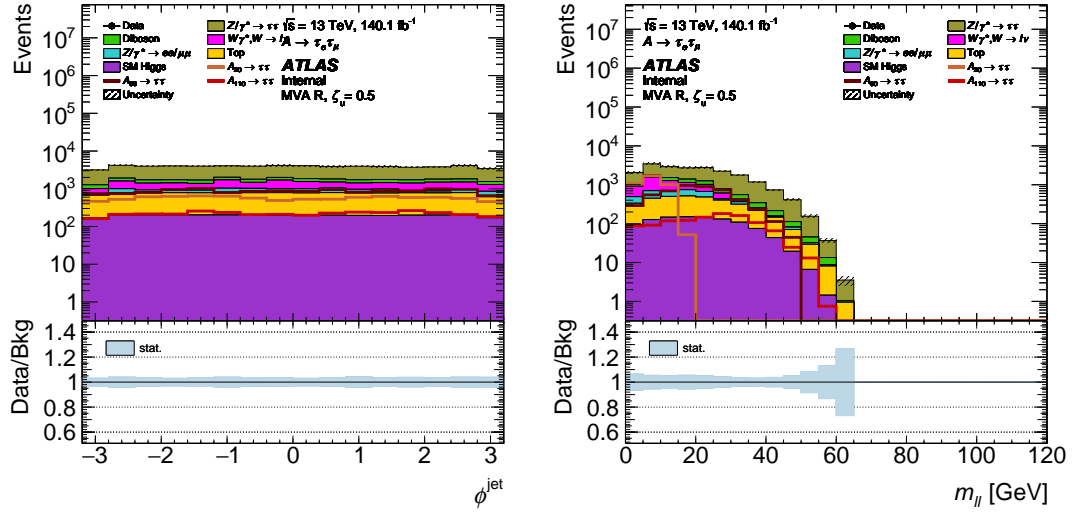


Abbildung A.5.: Verteilung von ϕ^{jet} (links) und m_{lj} (rechts) in der MVA-Region.

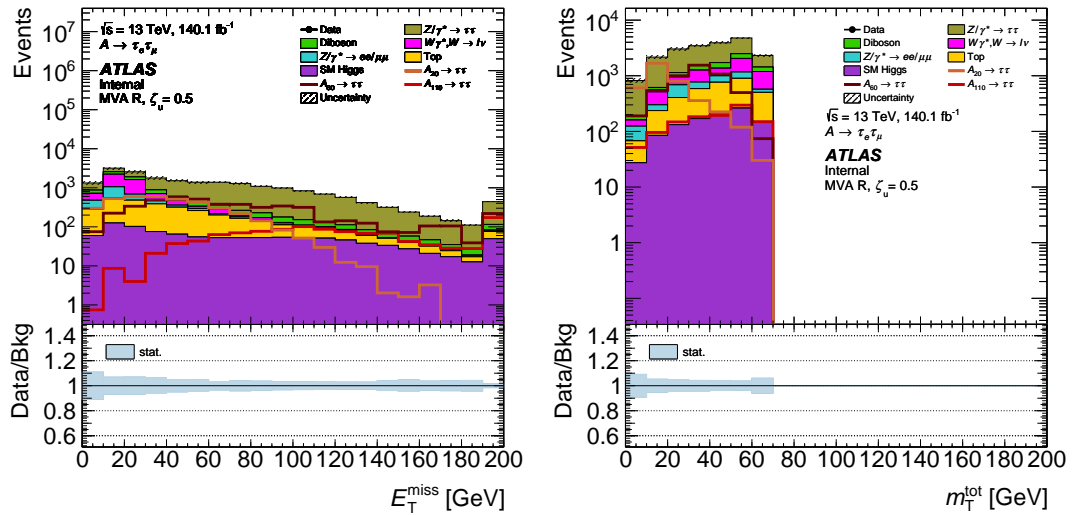


Abbildung A.6.: Verteilung von E_T^{miss} (links) und m_T^{tot} (rechts) in der MVA-Region.

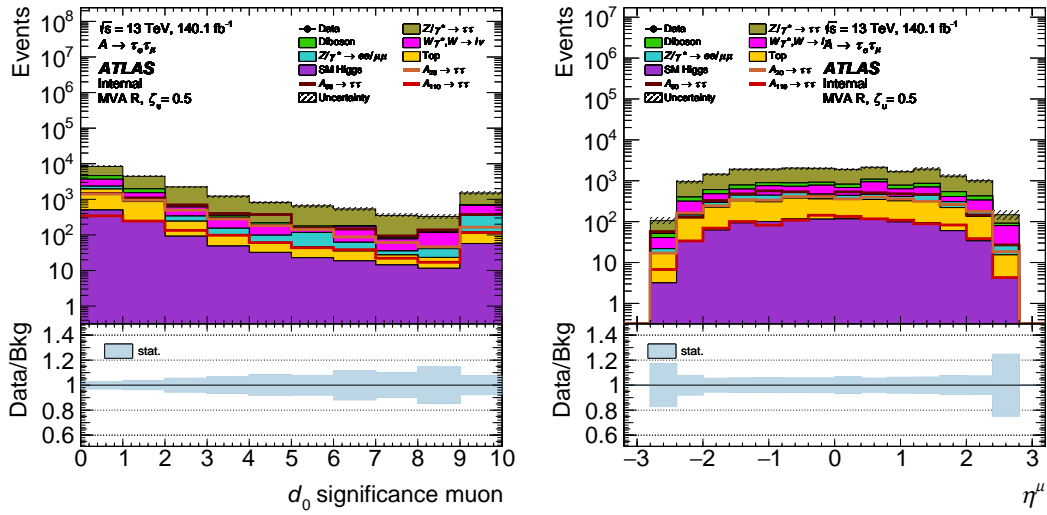


Abbildung A.7.: Verteilung von d_0 Signifikanz μ (links) und η_μ (rechts) in der MVA-Region.

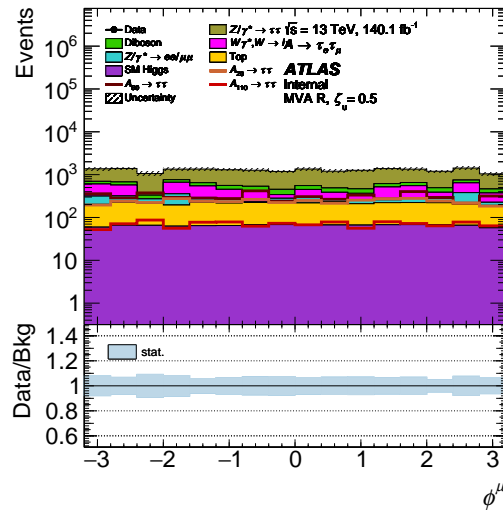


Abbildung A.8.: Verteilung von ϕ_μ in der MVA-Region

B. Tabelle der Separationskraft

Tabelle B.1.: Tabelle der Separationskraft für alle Massenhypothesen in Prozent.

Variable	$m_A[\text{GeV}]$									
	20	30	40	50	60	70	80	90	100	110
E_T^{miss}	9.0	2.6	2.4	3.6	8.0	12.5	14.8	19.0	25.2	30.3
p_T^e	4.0	1.1	2.5	1.0	1.0	1.4	2.5	4.0	3.9	4.2
η_e	0.2	0.3	1.2	0.9	0.6	0.5	1.3	1.2	0.9	1.6
ϕ_e	0.2	0.2	3.0	1.1	0.8	1.0	0.5	0.9	0.5	0.6
p_T^μ	3.9	1.0	2.4	1.4	1.3	2.6	3.6	3.8	5.2	4.6
η_μ	0.2	0.2	1.2	0.8	1.0	0.6	1.3	1.1	1.0	1.1
η_ϕ	0.3	0.3	1.9	0.9	1.1	0.8	0.5	0.8	0.4	0.8
m_T^{tot}	31.2	26.4	18.1	16.3	9.1	3.8	1.5	0.9	0.3	0.4
m_{MMC}	64.1	51.8	48.3	47.6	42.7	34.6	27.1	27.4	31.9	37.4
p_T^{Higgs}	15.8	4.1	4.4	4.9	9.2	14.2	17.6	20.9	27.9	33.8
m_{ll}	39.8	24.6	14.3	6.3	3.2	2.6	4.1	6.1	4.2	4.1
ΔR_{ll}	56.6	30.9	118	4.7	2.0	1.2	2.2	2.5	2.5	3.4
$\Delta\phi_{E_T^{\text{miss}}}^{\text{leadlep}}$	0.2	0.1	2.5	0.6	0.5	0.5	1.1	0.6	0.3	0.5
$\Delta\phi_{E_T^{\text{miss}}}^{\text{subleadlep}}$	0.3	0.1	3.0	0.7	0.6	0.5	0.6	0.9	0.5	0.8
$d_0 \text{ Sig } \mu$	0.5	0.6	0.5	1.3	2.1	1.5	0.9	1.9	1.9	1.5
$d_0 \text{ Sig } e$	0.6	0.8	1.7	2.3	1.0	1.4	0.9	1.9	2.4	3.2
p_T^{leadjet}	1.7	1.8	6.0	4.9	7.3	13.8	15.0	20.7	21.0	23.2
ϕ_{jet}	0.2	0.2	0.7	0.6	0.3	0.4	0.4	0.4	0.3	0.3
η_{jet}	0.2	0.1	0.5	0.7	0.4	0.4	0.2	0.3	0.2	0.2

C. Tabelle der Korrelationskoeffizienten

Tabelle C.1.: Korrelationskoeffizienten der kinematischen Variablen

E_{miss}^I	1,000	0,315	0,000	-0,016	0,296	0,001	0,005	-0,004	0,380	0,896	0,029	-0,245	-0,029	-0,026	0,083	0,129	0,771	-0,006	0,005
p_T^{miss}	0,315	1,000	0,007	-0,008	-0,059	0,008	-0,006	0,118	0,288	0,565	0,190	-0,179	-0,029	-0,026	0,129	0,493	0,771	-0,006	0,003
η_e	0,000	0,007	1,000	0,002	-0,003	0,878	0,005	0,003	0,878	0,008	-0,006	-0,015	-0,006	-0,006	0,000	0,000	0,000	-0,006	0,115
ϕ_e	-0,016	-0,008	0,002	1,000	0,006	0,003	0,531	0,011	0,005	0,011	0,013	0,006	-0,002	-0,018	-0,018	-0,018	-0,018	-0,002	0,005
p_T^{μ}	0,296	-0,059	-0,003	0,006	1,000	-0,007	0,013	0,128	0,328	0,518	0,240	-0,180	-0,007	-0,024	0,448	0,001	-0,012	-0,002	-0,002
η_{μ}	0,001	0,008	0,878	0,003	-0,007	1,000	0,008	0,008	-0,006	0,001	0,007	-0,012	0,001	0,007	0,001	0,005	0,007	0,001	0,112
ϕ_{μ}	0,005	-0,006	0,005	0,531	0,013	0,008	1,000	0,009	0,012	0,007	0,012	-0,005	-0,018	-0,024	0,092	-0,002	-0,003	-0,001	0,005
m_{T}^{tot}	-0,004	0,118	0,003	0,011	0,128	0,008	0,009	1,000	0,371	0,019	0,704	0,351	0,000	0,009	-0,014	-0,041	0,003	0,007	0,005
m_{MNC}	0,380	0,288	-0,003	0,005	0,328	-0,006	0,012	0,371	1,000	0,468	0,569	0,564	-0,011	-0,010	0,414	0,002	-0,002	0,004	0,004
H_{miss}	0,896	0,565	0,002	-0,011	0,518	0,001	0,007	0,019	0,468	1,000	0,145	-0,313	-0,026	-0,024	0,049	0,092	0,002	0,002	0,002
$p_T^{H_{miss}}$	0,029	0,190	-0,006	0,013	0,240	-0,007	0,012	0,704	0,569	0,145	1,000	0,443	-0,011	-0,007	-0,008	-0,019	-0,017	0,009	0,008
m_{ll}	-0,245	-0,179	-0,015	0,006	-0,180	-0,012	0,005	0,351	0,564	0,443	0,443	1,000	0,005	0,006	0,133	0,009	-0,001	0,000	-0,003
ΔR_{ll}	-0,029	0,003	-0,006	-0,002	-0,018	-0,006	-0,018	0,000	-0,011	-0,011	0,005	1,000	0,778	0,778	-0,005	-0,009	-0,017	0,001	-0,003
$\Delta\phi_{E_{miss}^{lepton}}$	-0,026	0,012	0,007	-0,008	-0,024	0,011	-0,024	0,009	-0,010	-0,024	0,006	0,006	1,000	1,000	-0,016	-0,010	-0,014	0,004	0,004
$\Delta\phi_{E_{miss}^{sublepton}}$	0,083	0,045	0,027	0,012	-0,073	0,031	0,015	-0,014	-0,009	-0,042	-0,026	-0,005	-0,016	-0,016	1,000	0,014	0,040	-0,001	0,003
d_0 Sig μ	0,129	-0,028	0,000	-0,006	0,032	0,000	-0,002	0,041	0,044	-0,008	-0,019	-0,009	-0,010	-0,010	0,014	1,000	0,086	-0,002	0,006
d_0 Sig e	0,771	0,493	0,000	-0,018	0,448	-0,003	-0,003	0,039	0,414	0,133	0,257	-0,017	-0,014	-0,014	0,040	0,086	0,996	0,000	0,004
p_T^{lepton}	-0,006	0,001	0,002	-0,012	-0,012	0,001	-0,158	0,007	-0,002	-0,009	-0,005	0,009	-0,001	-0,004	-0,001	-0,002	0,000	1,000	-0,003
ϕ_{jet}	0,005	0,003	0,115	0,005	-0,002	0,112	0,005	0,005	0,004	0,002	-0,008	0,006	-0,003	0,004	0,003	-0,006	0,004	-0,003	1,000
η_{jet}																			

Erklärung

Hiermit erkläre ich, dass ich diese Arbeit im Rahmen der Betreuung am Institut für Kern- und Teilchenphysik ohne unzulässige Hilfe Dritter verfasst und alle Quellen als solche gekennzeichnet habe.

Valentin Rothert

Dresden, Juni 2023