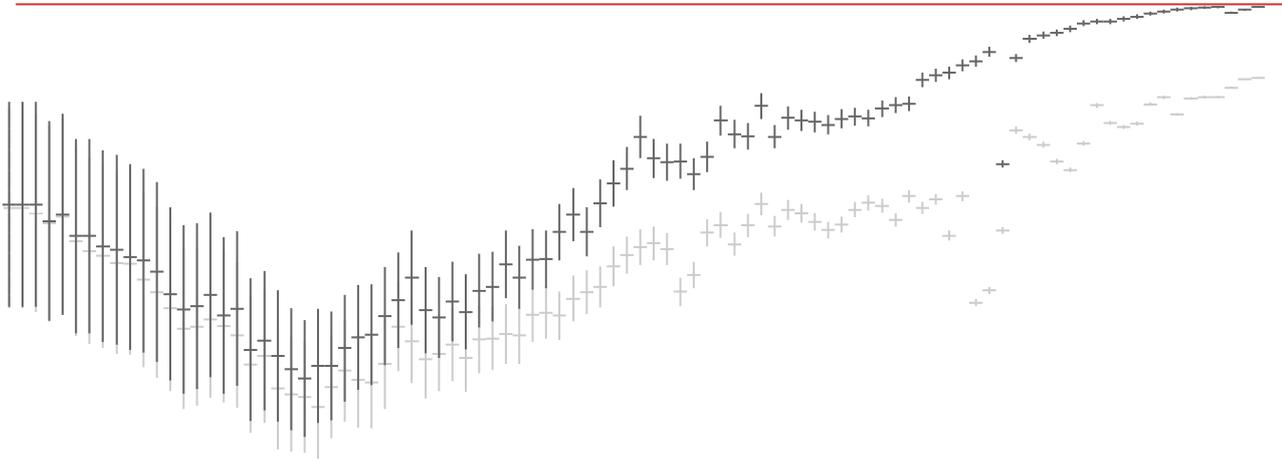


DIPLOMARBEIT

---



# Studien zur Vertex-Selektion im ATLAS-Experiment

vorgelegt von:

**Johanna Bronner**

1. Dezember 2009

Betreuer: Prof. Dr. Karl Jakobs

---

ALBERT-LUDWIGS UNIVERSITÄT FREIBURG I. BR.  
PHYSIKALISCHES INSTITUT  
2008 – 2009



Als Antwort auf die Widmung in [1]:  
Meiner Mutter Dorothea in Dankbarkeit gewidmet.



# Inhaltsverzeichnis

<b>1. Einleitung</b>	<b>1</b>
<b>2. Grundlagen</b>	<b>5</b>
2.1. Das Modell . . . . .	5
2.1.1. Elementarteilchen . . . . .	5
2.1.2. Wechselwirkungen . . . . .	8
2.1.3. Proton-Proton-Kollisionen . . . . .	15
2.2. Die Maschinen: Der LHC-Ring und der ATLAS-Detektor . . . . .	17
2.2.1. Der LHC-Ring . . . . .	17
2.2.2. Der ATLAS-Detektor . . . . .	19
2.3. Die Simulation . . . . .	27
2.3.1. Ereignisgeneration . . . . .	27
2.3.2. Detektorsimulation . . . . .	28
<b>3. <i>Pile-Up</i> und <i>Minimum-Bias</i></b>	<b>31</b>
3.1. <i>Minimum-Bias</i> -Ereignisse . . . . .	31
3.2. <i>Pile-Up</i> . . . . .	32
3.2.1. <i>Pile-Up</i> Vorhersagen für den LHC . . . . .	34
3.2.2. Die Wechselwirkungsregion . . . . .	36
<b>4. Primärvertex-Rekonstruktion in ATLAS</b>	<b>41</b>
4.1. Algorithmen der Primärvertex-Rekonstruktion . . . . .	42
4.1.1. Algorithmen zum Vertex-Fit . . . . .	43
4.1.2. Algorithmen zur Vertex-Suche . . . . .	47
4.2. Rekonstruktionsleistung der Algorithmen . . . . .	51
<b>5. Signalvertex-Selektion in Gegenwart von <i>Pile-Up</i></b>	<b>55</b>
5.1. Eigenschaften von <i>Minimum-Bias</i> - und Signalvertizes . . . . .	55
5.1.1. Die verwendeten Signalprozesse . . . . .	55
5.1.2. Eigenschaften der Spuren am Vertex . . . . .	57
5.2. Methoden zur Selektion des Signalvertex . . . . .	60
5.2.1. Summe der transversalen Impulse der Spuren am Vertex . . . . .	62

5.2.2.	Vertex-Wahrscheinlichkeit . . . . .	65
5.2.3.	Methode des Likelihood-Verhältnisses . . . . .	70
5.2.4.	Leistung der Methoden im Vergleich . . . . .	76
5.3.	Korrelationen der Spur-Eigenschaften . . . . .	78
5.4.	Vertex-Selektion durch ein künstliches neuronales Netz . . . . .	83
5.4.1.	Funktionsweise eines künstlichen neuronalen Netzes . . . . .	83
5.4.2.	Neuronale Netze zur Signalvertex-Selektion . . . . .	86
5.5.	Zusammenfassung . . . . .	100
<b>6.</b>	<b>Signalvertex-Selektion mit einem offiziellen Rekonstruktionswerkzeug</b>	<b>101</b>
6.1.	ATHENA: die Software des ATLAS-Detektors . . . . .	101
6.2.	ATHENA-Werkzeug zur Selektion des Signalvertex . . . . .	102
6.3.	Leistungstest der in ATHENA implementierten Selektionsmethoden . . . . .	104
<b>7.</b>	<b>Einfluss von <i>Pile-Up</i> auf die b-Jet-Identifikation</b>	<b>109</b>
7.1.	b-Tagging-Algorithmen bei ATLAS . . . . .	110
7.1.1.	Verwendete Rekonstruktionsobjekte . . . . .	110
7.1.2.	Impact-Parameter-Algorithmen . . . . .	112
7.1.3.	Sekundärvertex-Algorithmen . . . . .	114
7.2.	Einfluss von <i>Pile-Up</i> auf die b-Tagging-Leistung . . . . .	116
7.2.1.	Einfluss von Spuren aus <i>Minimum-Bias</i> -Wechselwirkungen . . . . .	121
7.2.2.	Einfluss der Fehlidentifikation des Signalvertex . . . . .	123
7.2.3.	Kombination der beiden <i>Pile-Up</i> -Einflüsse . . . . .	130
<b>8.</b>	<b>Zusammenfassung</b>	<b>135</b>
	<b>Anhang</b>	<b>139</b>
<b>A.</b>	<b>b-Tagging Abbildungen</b>	<b>141</b>
A.1.	Einfluss von <i>Minimum-Bias</i> -Spuren auf die b-Tagging-Leistung . . . . .	141
A.2.	Einfluss der Fehlidentifikation des Signalvertex auf die b-Tagging-Leistung . . . . .	145
A.3.	Einfluss der beiden betrachteten <i>Pile-Up</i> -Effekte . . . . .	148
	<b>Literaturverzeichnis</b>	<b>150</b>
	<b>Danksagung</b>	<b>153</b>

# 1. Einleitung

Die Bausteine der Materie und ihre fundamentalen Wechselwirkungen werden sehr erfolgreich durch die Theorie des Standardmodells beschrieben. Die theoretischen Vorhersagen des Standardmodells wurden durch viele Experimente erfolgreich bestätigt. Beispielhaft sei hier die Entdeckung  $W$ - und  $Z$ -Bosonen im Jahr 1983 am SPS<sup>1</sup> am CERN<sup>2</sup> angeführt. Die Eigenschaften dieser konnten später bei LEP<sup>3</sup> präzise bestimmt werden. In einem sehr weiten Bereich ist das Standardmodell in der Lage, Phänomene der Teilchenphysik richtig zu beschreiben. Dabei sind drei der vier fundamentalen Kräfte durch das Standardmodell beschreibbar. Die vierte Kraft, die Gravitation, konnte noch nicht eingebunden werden. Fast alle vom Standardmodell vorhergesagten Elementarteilchen sind an Experimenten nachgewiesen worden. 1995 wurde das bisher schwerste beobachtete Elementarteilchen, das top-Quark am Tevatron (TeV Synchrotron) am Fermilab in Chicago entdeckt. Alleine von einem einzigen Teilchen, dem Higgs-Boson, dessen Existenz für das Standardmodell essentiell ist, existiert bisher kein experimenteller Nachweis. Es wird erwartet, dass die Masse des Higgs-Bosons größer ist als die der meisten bisher beobachteten Elementarteilchen. Der experimentelle Nachweis eines schweren Teilchens wie beispielsweise des Higgs-Bosons benötigt einerseits sehr große Energien. Andererseits wird erwartet, dass ein Higgs-Boson nur selten auftritt. Dementsprechend sind sehr große Ereignisraten und Energien notwendig, um ein solches nachweisen zu können.

Am ATLAS-Experiment<sup>4</sup> am LHC<sup>5</sup> am CERN, der derzeit in Betrieb genommen wird, ist man potenziell in der Lage, ein Higgs-Boson zu entdecken. Neben der Suche nach dem Higgs-Boson können hier auch weitere offene Fragen der Elementarteilchenphysik untersucht werden. Im LHC sollen hierfür Protonen beschleunigt und im Zentrum von vier großen Detektoren, darunter der ATLAS-Detektor, zur Kollision gebracht werden. Mit einer Schwerpunktsenergie von bis zu geplanten 14 TeV werden am LHC Kollisionsenergien erreicht, die bisher von keinem anderen Experiment verwirklicht werden konnten. Der LHC eröffnet somit die Möglichkeit zur Untersuchung eines bisher nicht beobachtbaren Ener-

---

<sup>1</sup>Super Proton Synchrotron

<sup>2</sup>franz.: Conseil Européen pour la Recherche Nucléaire

<sup>3</sup>Large Electron Positron Collider

<sup>4</sup>A Toroidal LHC Apparatus

<sup>5</sup>Large Hadron Collider

giebereichs. Neben der sehr großen Schwerpunktsenergie zeichnet sich der LHC durch eine sehr hohe geplante Luminosität von  $\mathcal{L} = 10^{34} \text{ cm}^{-2}\text{s}^{-1}$  aus. Bündel von  $1.5 \cdot 10^{11}$  Protonen sollen alle 25 ns zur Kollision gebracht werden.<sup>6</sup> Es werden hierdurch die notwendigen Ereignisraten erreicht.

Die geplanten Energien und Luminositäten führen einerseits zu der gewünschten Ereignisrate und machen sehr seltene, hochenergetische Prozesse beobachtbar. Andererseits führen sie aber auch zu einer sehr großen Dichte von Teilchen im Endzustand und konfrontieren damit die Experimente mit einer Vielzahl von neuen Herausforderungen. Zum einen sind dies sehr hohe technische Anforderungen an die Detektoren, zum anderen aber ist die Ereignisrekonstruktion vor die Aufgabe gestellt, eine komplexe Struktur aus einer sehr großen Zahl von Detektorsignalen zu erkennen. Erschwert wird dies dadurch, dass bei der geplanten Luminosität in jeder Strahlkreuzung im Mittel ca. 20 inelastische Proton-Proton-Wechselwirkungen stattfinden werden, deren Streuteilchen sich im Detektor überlagern. Die Überlagerung mehrerer Proton-Proton-Wechselwirkungen in einer Strahlkreuzung wird *Pile-Up* genannt. Jeder interessante Prozess wird demnach mit einer Vielzahl von Teilchen aus nicht zur Untersuchung stehenden, Prozessen überlagert. Die Präzision mit der die Wechselwirkungspunkte bestimmt werden können, ist im Vergleich zur Ausdehnung des Bereichs innerhalb dessen die Wechselwirkungen stattfinden, ausreichend genau um die unterschiedlichen Wechselwirkungspositionen getrennt rekonstruieren zu können. Es ist Aufgabe der in dieser Arbeit untersuchten Signalvertex-Selektion zu erkennen, welcher Wechselwirkungspunkt derjenige ist, an dem der Prozess von Interesse stattgefunden hat.

Eine möglichst effiziente Signalvertex-Selektion ist von großer Bedeutung für Analysen, die explizit die Position der primären Wechselwirkung verwenden. Dies gilt zum Beispiel für die Identifikation von Jets aus b-Quarks, b-Jets genannt. Hier ist zu erwarten, dass die Fehlidentifikation des Signalvertex die Identifikation eines b-Jets stark beeinträchtigt. In dieser Arbeit soll der Einfluss von zusätzlichen *Pile-Up*-Wechselwirkungen auf die Leistungsfähigkeit der b-Jet-Identifikation untersucht werden.

Die Arbeit ist in folgender Weise gegliedert:

In Kapitel 2 werden die theoretischen und experimentellen Grundlagen der Arbeit beschrieben. Hier wird das Standardmodell dargestellt und der Aufbau des LHCs und des ATLAS-Detektors skizziert. Kapitel 3 behandelt die für diese Arbeit wichtigen Hintergründe von *Pile-Up* am LHC. Es werden dabei aus den technischen Gegebenheiten des LHCs Abschätzungen über die zu erwartenden *Pile-Up*-Raten gemacht. Im Anschluss widmet sich die Arbeit der Rekonstruktion und Selektion des Wechselwirkungspunkts des Signalprozesses

---

<sup>6</sup>In der vorliegenden Arbeit werden Punkte als Dezimalzeichen verwendet, wie es in der anglo-amerikanischen Zahlenschreibweise üblich ist.

und den Folgen einer Fehlselektion. Hierfür wird zunächst in Kapitel 4 die Rekonstruktion der Proton-Proton-Wechselwirkungspunkte bei ATLAS beschrieben. Das folgende Kapitel, Kapitel 5, beinhaltet die Beschreibung der wesentlichen Studien dieser Arbeit. Hier werden verschiedene Methoden der Signalvertex-Selektion untersucht und auf ihrer Selektionsleistung hin verglichen. In Kapitel 6 wird die Leistungsfähigkeit zusammengefasst dargestellt, die ein, im Rahmen dieser Arbeit neu in der ATLAS-Software implementiertes, offizielles Werkzeug zur Signalvertex-Selektion erreicht. Die hier angegebenen Ergebnisse entsprechen den in der Realität erwarteten Leistungen. Abschließend wird in Kapitel 7 der Einfluss von zusätzlichen *Pile-Up*-Wechselwirkungen auf die b-Jet-Identifikation untersucht.



## 2. Grundlagen

Zu Beginn dieser Arbeit sollen die theoretischen und experimentellen Hintergründe beschrieben werden. Die vorliegende Arbeit hat im Rahmen des ATLAS-Experiments am CERN in Genf stattgefunden, dessen Zweck es ist Phänomene der Teilchenphysik zu untersuchen. Die theoretische Basis der Teilchenphysik ist das *Standardmodell*. Das vorliegende Kapitel umfasst eine Beschreibung des Standardmodells in Abschnitt 2.1, des experimentellen Aufbaus des ATLAS-Experiments in Abschnitt 2.2 und der Ereignisrekonstruktion in Abschnitt 2.3.

### 2.1. Das Modell

Die uns bekannten Bausteine der Materie und deren Wechselwirkungen lassen sich im Standardmodell in erstaunlich einfacher Form zusammenfassen und beschreiben. Das Standardmodell umfasst alle derzeit bekannten Elementarteilchen und beschreibt deren Wechselwirkungen (ausführlichere Beschreibungen finden sich z.B. in [2] und [3]). Bis heute hat dieses Modell allen Falsifizierungsversuchen standgehalten.

#### 2.1.1. Elementarteilchen

Ein Teilchen, dem das Attribut *elementar* zugeschrieben wird, hat nach dem derzeitigen Stand der Erkenntnis keine weitere Substruktur. Eine Klassifizierung der uns bekannten Elementarteilchen kann anhand ihrer Quantenzahlen vorgenommen werden. Hierzu gehören z.B. der Spin  $s$ , die elektrische Ladung  $q$ , die Farbladung  $c$ , und der schwache Isospin  $I$ . Desweiteren lassen sich die Elementarteilchen anhand ihrer Masse unterscheiden. Während sich die oben genannten Quantenzahlen aus der zugrunde liegenden Quantenfeldtheorie ergeben, bleibt die Herkunft der Masse der Teilchen bisher unverstanden. Die Masse ist experimentell bestimmbar, nicht aber theoretisch vorhersagbar. Anhand ihres Spins lassen sich die Elementarteilchen in zwei Klassen unterteilen: *Fermionen* und *Bosonen*. Teilchen mit halbzahligem Spin sind Fermion und gehorchen der Fermi-Dirac-Statistik. Hierzu zählen im Standardmodell die Leptonen und Quarks, die unsere Materie aufbauen. Alle Teilchen mit ganzzahligem Spin dagegen sind Bosonen. Diese werden durch die Bose-

Einstein-Statistik beschrieben. Alle Austauscheteilchen im Standardmodell sind Bosonen. Austauscheteilchen sind die Vermittler der Kräfte.

### Leptonen und Quarks

Leptonen und Quarks sind Teilchen mit Spin  $s = 1/2$  und zählen daher zu den Fermionen. Beide lassen sich durch ihre Eigenschaften in sehr ähnlicher Weise in einer Drei-Generationenstruktur anordnen. Dabei sind die zweite und dritte Generation sowohl der Quarks als auch der Leptonen schwerere Kopien der ersten Generation.

Die drei Generationen der Leptonen werden von Elektron ( $e$ ), Myon ( $\mu$ ) und Tauon ( $\tau$ ) mit zugehörigen Neutrinos ( $\nu_e, \nu_\mu, \nu_\tau$ ) gebildet:

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix} \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix} \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix} \begin{matrix} q = 0 \\ q = -1 \end{matrix} . \quad (2.1)$$

Elektron, Myon und Tauon sind mit einer einfachen elektrischen Elementarladung geladen. Neutrinos tragen keine elektrische Ladung. Die geladenen Leptonen haben mit experimentellen Methoden bestimmbare und unterscheidbare Massen. Der Massenbereich der geladenen Leptonen erstreckt sich von der Elektron-Masse  $m_e = 0.511 \text{ MeV}/c^2$  über die Myon-Masse  $m_\mu = 105.659 \text{ MeV}/c^2$  bis zur Tauon-Masse  $m_\tau = (1776 \pm 0.17) \text{ MeV}/c^2$  [4]. Auch Neutrinos, so ist mittlerweile bekannt, tragen Massen und es konnte nachgewiesen werden, dass diese voneinander unterschiedlich sein müssen [5]. Innerhalb des Standardmodells die Neutrinos als masselos angenommen. Zu jedem Lepton existiert ein Antiteilchen, das die umgekehrten additiven Quantenzahlen, aber die gleiche Masse wie das zugehörige Teilchen besitzt.

Die Quarks sind in folgender Weise in einer Drei-Generationenstruktur angeordnet:

$$\begin{pmatrix} \mathbf{up} \\ \mathbf{down} \end{pmatrix} \begin{pmatrix} \mathbf{charm} \\ \mathbf{strange} \end{pmatrix} \begin{pmatrix} \mathbf{top} \\ \mathbf{bottom} \end{pmatrix} \begin{matrix} q = +\frac{2}{3} \\ q = -\frac{1}{3} \end{matrix} . \quad (2.2)$$

Die verschiedenen Quark-Typen werden Quark-*Flavour* genannt. Quarks tragen einen Bruchteil der elektrischen Elementarladung. Die Quarks u, c und t tragen einen Anteil von  $+\frac{2}{3}e$ , die Quarks d, s und b tragen einen Anteil von  $-\frac{1}{3}e$  der Elementarladung. Im Gegensatz zu den Leptonen haben Quarks die Eigenschaft, eine sogenannte *Farbladung* zu tragen. Quarks interagieren durch ihre elektrische Ladung elektromagnetisch, durch ihre Farbladung auch stark miteinander. Leptonen und Quarks nehmen ebenso an der schwachen Wechselwirkung teil. Auch zu allen Quarks existiert ein Antipartner. Die Antiquarks tragen Antifarbe an Stelle von Farbe.

Quarks sind nur in gebundenen Zuständen, genannt *Hadronen*, beobachtbar. Es existieren gebundene Zustände aus drei Quarks (*Baryonen*) wie zum Beispiel Proton ( $uud$ ) und

Neutron (ddu). Beobachtet wird auch ein Verbund aus einem Quark und einem Antiquark (*Mesonen*). Ein in Streuexperimenten sehr häufig auftretendes Meson ist das Pion, welches aus einem Quark und einem Antiquark der ersten Generation aufgebaut ist. Da Quarks nicht als freie Teilchen vorkommen, ist die Bestimmung ihrer Masse nicht ohne Weiteres möglich. Die *nackten* Massen der Quarks müssen aus den Massen der Hadronen, die aus diesen Quarks aufgebaut sind und aus theoretischen Modellen bestimmt werden. Während sich für die Quarks der ersten Generation eine Masse von einigen  $\text{MeV}/c^2$  ergibt, erreicht das b-Quark bereits eine Masse in einer Größenordnung von  $m_b \approx 4 \text{ GeV}/c^2$  [4]. Das Top-Quark ist im Vergleich zum b-Quark um fast zwei weitere Größenordnungen schwerer mit  $m_t \approx 170 \text{ GeV}/c^2$  [4].

Die uns umgebende Materie ist ausschließlich aus den Teilchen der ersten Leptonen- und Quark-Generation aufgebaut. Protonen und Neutronen sind aus u- und d-Quarks zusammengesetzt. Protonen und Neutronen wiederum bilden den Aufbau der Atomkerne, die zusammen mit den Elektronen die für uns sichtbare Materie in all ihrer Vielfalt ausmachen. Alle anderen beschriebenen Fermionen lassen sich in Experimenten nachweisen.

### Die Austauschteilchen

Neben den Leptonen und Quarks existiert eine weitere Klasse von elementaren Teilchen, die *Austauschteilchen*. Diese sind die Träger der Wechselwirkungen. Alle beobachteten Austauschteilchen im Standardmodell haben Spin  $s = 1$  und sind somit Bosonen. Die Austauschteilchen werden auch *Eichbosonen* genannt. In Tabelle 2.1 sind die Eigenschaften der im Standardmodell nachgewiesenen elementaren Bosonen aufgelistet. Es sei auf die großen Massen der  $W^\pm$ - und  $Z^0$ -Bosonen hingewiesen. Im Gegensatz hierzu sind das Photon  $\gamma$  und das Gluon  $g$  masselos. Diese Tatsache wird in der Folge eine wichtige Rolle spielen. Die Kopplung der Eichbosonen an die Fermionen wird innerhalb der Quantenfeldtheorie durch die Forderung nach einer lokalen Eichinvarianz eingeführt.

Eichboson	Ladung [e]	Masse [ $\text{GeV}/c^2$ ]
$\gamma$	0	0
$g$	0	0
$W^\pm$	$\pm 1$	$80.398 \pm 0.025$ [4]
$Z^0$	$\pm 0$	$91.1876 \pm 0.0021$ [4]

**Tab. 2.1.:** Masse und elektrische Ladung der Austauschteilchen.

### 2.1.2. Wechselwirkungen

Das Standardmodell beschreibt folgende Wechselwirkungen:

**QED:** Die **Q**uantenelektrodynamik ist die Theorie der elektromagnetischen Kraft. Ihr Vermittler ist das Photon.

**QCD:** Die **Q**uantenchromodynamik ist die Theorie der starken Kraft. Ihre Vermittler sind Gluonen.

**GSW:** Innerhalb der Theorie von **G**lashow, **S**alam und **W**einberg kann die schwache Kraft in Vereinigung mit der elektromagnetischen Kraft beschrieben werden. Die Vermittler dieser Kraft sind zusätzlich zum Photon die W- und Z-Bosonen.

Innerhalb des Standardmodells existiert für die Gravitation, die letzte fehlende fundamentale Kraft, keine theoretische Beschreibung. Für alle anderen Kräfte soll hier eine kurze Darstellung der zugrunde liegenden Quantenfeldtheorien folgen.

#### Quantenelektrodynamik (QED)

Eine freies geladenes Fermion der Masse  $m$  wird in der Quantenfeldtheorie durch das Diracfeld  $\Psi$  beschrieben. Die zugehörige Lagrangedichte  $\mathcal{L}$  lautet:

$$\mathcal{L} = i\bar{\psi}\gamma_{\mu}\partial_{\mu}\psi - m\bar{\psi}\psi . \quad (2.3)$$

Dabei sind  $\partial_{\mu}$  die partiellen Ableitungen nach den Raumzeitkoordinaten und  $\gamma_{\mu}$  die Dirac-Matrizen.

Physikalischen Prozessen liegen immer Forderungen nach Eichinvarianzen zugrunde. Mit der Forderung nach einer Eichinvarianz wird vom Prozess eine Symmetrieeigenschaft verlangt. Jede Symmetrie induziert nach dem Noether-Theorem eine Erhaltungsgröße. Für die zeitliche Translation ist dies beispielsweise die Energieerhaltung für die örtliche Translation ist es die Impulserhaltung. Die obige Lagrangedichte besitzt weitere Symmetrieeigenschaften, beispielsweise unter einer Phasentransformation der Form  $\psi \rightarrow e^{i\theta}\psi$ , wobei  $\theta$  ein beliebiger reeller Winkel ist, der nicht vom Ort oder der Zeit abhängt. Man spricht insofern von einer *globalen* Transformation. Fordert man, dass die Lagrangedichte auch invariant unter einer *lokalen* Transformation  $\psi \rightarrow e^{i\theta(x)}\psi$  ist, wobei  $\theta(x)$  nun eine vom Raumzeitpunkt  $x$  abhängige reelle Größe sei, ist dies nur unter folgenden Bedingungen möglich: Die partiellen Ableitungen  $\partial_{\mu}$  müssen durch die kovarianten Ableitungen  $D_{\mu} = \partial_{\mu} - iqA_{\mu}$  ersetzt werden, wobei  $A_{\mu}$  ein neu einzuführendes Eichfeld darstellt, welches an die Ladung  $q$  des Diracfelds koppelt. Fügt man in der Lagrangedichte einen Energieterm hinzu, wird die Interpretation von  $A_{\mu}$  als Photonenfeld zulässig. Man erhält eine Lagrangedichte die unter einer lokalen Transformation invariant ist und die Kopplung

eines freien Fermions an ein masseloses Photonenfeld beschreibt:

$$\mathcal{L}_{QED} = i\bar{\psi}\gamma_{\mu}\partial_{\mu}\psi - m\bar{\psi}\psi + q\bar{\psi}\gamma^{\mu}A_{\mu}\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \quad (2.4)$$

Hierbei ist  $F_{\mu\nu} = \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu}$  der Feldstärketensor des Photonenfelds. Aus der lokalen Invarianz resultiert die Erhaltung der Ladung  $q$ . Aus der gegebenen Lagrangedichte ist man in der Lage, über Störungsrechnung die Übergangswahrscheinlichkeiten physikalischer Prozesse zu berechnen. Auf dieser Grundlage können Vorhersagen der Theorie für Wirkungsquerschnitte und Übergangsraten gemacht werden, anhand derer die Theorie im Experiment getestet werden kann.

Die sich aus den Eichinvarianzen ergebenden Symmetrien legen eine Gruppenstruktur der Theorien zu Grunde. Einfache Phasentransformationen, wie sie in der QED vorliegen, haben eine unitäre  $1 \times 1$ -Matrix-Struktur. Die Symmetriegruppe der QED ist somit eine  $U(1)$ -Symmetrie.

### Quantenchromodynamik (QCD)

Die QCD ist die theoretische Beschreibung der Wechselwirkung von Teilchen, die Farbladung tragen. Die Wechselwirkung wird auch *starke* Wechselwirkung genannt und ist die Kraft, die Quarks in ihrem Hadronen-Verbund zusammenhält. Die starke Kraft wird von Gluonen vermittelt. Als Analogon der elektrischen Ladung wird den stark wechselwirkenden Quarks die Eigenschaft Farbladung zugeschrieben. Die additive Farbmischung aus der Optik gab das Beispiel für dieses Farb-Modell. Rot gemischt mit Grün und Blau ergibt weiß bzw. farblos. Ein Quark-Triplett mit dieser Farbkombination ist wieder farbneutral. Ein weiterer beobachteter Bindungszustand von Quarks sind die Quark-Antiquark-Paare. Auch dieser Bindungszustand ist farblos, wenn Quark und Antiquark Farbe und zugehörige Antifarbe tragen. Die zwischen zwei farbigen Teilchen wirkende Kraft hat die Eigenschaft, mit wachsendem Abstand der Teilchen stärker zu wirken. Dies hat zur Folge, dass farbige Teilchen nicht frei, sondern nur im farblosen Verbund, auftreten. Dieses Phänomen wird *confinement* genannt. Aufgrund dessen ist die experimentelle Untersuchung der QCD deutlich komplexer als die der QED.

Anders als in der QED, ist das Gluon selbst farbgeladen. Ein Gluon trägt eine Kombination aus Farbe und Antifarbe. Gluonen können somit an sich selbst koppeln. Es existiert ein Farb-Oktett aus 8 Gluonen. Das Farb-Singlett wäre denkbar, hätte aber eine weitreichende nicht beobachtete starke Kraft zur Folge.

Analog zur QED lässt sich für freie Quarks eine Lagrangedichte formulieren. Durch die Forderung nach lokaler Eichinvarianz lassen sich die Gluonen-Felder in die Lagrangedichte einführen. Die Symmetrie der Transformation hat hier eine  $SU(3)$ -Struktur. Da Quarks

nicht frei vorkommen, ist eine reine Quark-Quark-Streuung, wie sie von der QCD berechnet werden kann, nicht direkt experimentell realisierbar. Am LHC werden Protonen zur Kollision gebracht. Bei ausreichender Schwerpunktsenergie wechselwirken tatsächlich Konstituenten des Protons. Ein, an einem solchen Subprozess teilnehmendes Quark ist quasifrei. Der Subprozess, sofern er ein starker Prozess ist, was für den Großteil aller Fälle am LHC gegeben ist, kann dann anhand der QCD berechnet werden. Der gesamte Prozess einer Proton-Proton-Kollisionen wird in Abschnitt 2.1.3 beschrieben.

### Elektroschwache Vereinigung

Die elektroschwache Vereinigung ist eine einheitliche Formulierung der schwachen und der elektromagnetischen Wechselwirkung. Es sollen hier zunächst kurz die Eigenschaften der schwachen Kraft dargestellt werden.

Die Vermittler der schwachen Kraft sind die schweren Eichbosonen  $Z^0$  und  $W^\pm$ . Die  $W^\pm$ -Bosonen ändern bei der Kopplung an Quarks den Quark-Flavour, wie beispielsweise bei dem Zerfall eines b-Quarks in ein c-Quark, wie er in Kürze beschrieben wird. Bei der Kopplung an Leptonen gehen diese von geladenen in ungeladene Leptonen über und umgekehrt. Dieses Verhalten wird über eine weitere Erhaltungsgröße, die man den *schwachen Isospin* nennt, beschrieben. Die linkshändigen Quarks und Leptonen bilden je drei schwache Isospin-Dubletts. Die rechtshändigen Quarks und Leptonen sind schwache Isospin-Singletts. Den Fermionen werden folgende dritte Komponenten  $T_3$  des schwachen Isospins zugeordnet:

$$\begin{aligned} T_3\left(\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L\right) &= T_3\left(\begin{pmatrix} u \\ d \end{pmatrix}_L\right) &= \begin{matrix} +1/2 \\ -1/2 \end{matrix} \\ T_3(e_R) = T_3((\nu_e)_R) = T_3(u_R) = T_3(d_R) &= 0 \end{aligned} .$$

Die Indizes R und L stehen für die Händigkeit der Fermionen. Gleiches gilt für die zweite und dritte Generation von Leptonen und Quarks. Die dritte Komponente des schwachen Isospins ist für die geladenen schwachen Eichbosonen  $T_3(W^\pm) = \pm 1$ , für das ungeladene schwache Eichboson  $T_3(Z^0) = 0$ .

Die Gruppenstruktur des schwachen Isospins hat in dieser Form eine  $SU(2)$ -Symmetrie, die Vereinigung mit der QED ergibt dann folgende Struktur:

$$SU(2)_L \times U(1)_Y . \tag{2.5}$$

Der Index  $L$  steht für die linkshändige Dublett-Struktur.  $Y$  heisst *Hyperladung*. Auf dieser Basis lässt sich eine Lagrangedichte formulieren mit vier masselosen Eichfeldern ( $W_\mu^{1,2,3}$  und  $B_\mu$ ), von denen man naiv erwarten würde, dass sie die Felder der schwachen Eichbosonen und des Photons sind. Die zugehörigen Erhaltungsgrößen wären hier dann der

schwache Isospin und die elektrische Ladung. Es zeigt jedoch sich anderes: Die aus der Lagrangedichte gegebene erhaltene Ladung ist nicht die elektrische Ladung  $Q$  sondern die Hyperladung  $Y$  mit

$$Q = T_3 + \frac{1}{2}Y . \quad (2.6)$$

Desweiteren stellt man fest, dass die vier Eichfelder ( $W_\mu^{1,2,3}$  und  $B_\mu$ ) in der Lagrangedichte nicht den Feldern der beobachtbaren Eichbosonen ( $W_\mu^\pm$ ,  $Z_\mu$  und  $A_\mu$ ) entsprechen. Diese sind eine Mischung der vier Eichfelder:

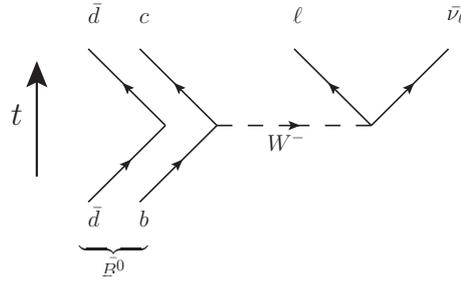
$$\begin{aligned} A_\mu &= B_\mu \cos\theta_w + W_\mu^3 \sin\theta_w \\ Z_\mu &= -B_\mu \sin\theta_w + W_\mu^3 \cos\theta_w \\ W_\mu^\pm &= \frac{1}{2}(W_\mu^1 \pm W_\mu^2) \end{aligned} .$$

Der Mischungswinkel  $\theta_w$  wird Weinbergwinkel oder schwacher Mischungswinkel genannt.<sup>1</sup> Alle Eichbosonen sind aufgrund der Eichinvarianz masselos. Aus vielen Experimenten ist aber bekannt, dass die schwachen Eichbosonen eine von null verschiedene Masse tragen (siehe Tabelle 2.1). Eine Einführung der Massen in die Lagrangedichte ist ohne die Symmetrie zu zerstören erst durch den *Higgs-Mechanismus* möglich. An dieser Stelle wird vor der Beschreibung des Higgs-Mechanismus noch eine für die vorliegende Arbeit wichtige Eigenart des Quark-Sektors der schwachen Wechselwirkung eingeführt.

**Zerfälle von b-Hadronen am Beispiel des  $\bar{B}^0$ -Meson-Zerfalls** An dieser Stelle soll beispielhaft auf einen expliziten Quark-Zerfall und seine Besonderheiten eingegangen werden. In Abbildung 2.1 ist der Zerfall eines  $\bar{B}^0$ -Mesons ( $\bar{B}^0 = b\bar{d}$ ) gezeigt, dargestellt in einem *Feynmangraphen*. Das b-Quark im  $\bar{B}^0$ -Meson zerfällt über die Abstrahlung eines  $W^-$  in ein c-Quark. Das  $W^-$  zerfällt instantan weiter in ein Lepton-Neutrino-Paar ( $\ell, \bar{\nu}_\ell$ ). Der Zerfall von Hadronen, die b-Quarks enthalten (b-Hadronen) sind innerhalb dieser Arbeit vor allem in Kapitel 7 relevant. Die Lebensdauer der b-Hadronen ist ausreichend lang, um eine, für die Detektoren der Teilchenphysikexperimente nachweisbare Flugstrecke zu hinterlassen. Die Ursache der langen Lebensdauer von b-Hadronen soll in der Folge erklärt werden.

Die QED und QCD erlauben keine Übergänge zwischen den Generationen. Im Quark-Sektor dagegen sind generationenändernde Vorgänge bekannt, seit die Quarks aus zweiter und dritter Generationen entdeckt worden sind. Anderes hätte zwangsläufig stabile schwerere Hadronen aus Quarks der zweiten und dritten Generation zur Folge. Solche werden jedoch nicht beobachtet. Auch der oben beschriebene  $\bar{B}^0$ -Meson-Zerfall wäre nicht erlaubt.

<sup>1</sup>Der schwache Mischungswinkel kann experimentell bestimmt werden:  $\theta_w \approx 0.231$  [4].



**Abb. 2.1.:** Feynmangraph des  $\bar{B}^0$ -Zerfalls.

Die Eigenzustände der starken Kraft ( $(u, c, t)$  und  $(d, s, b)$ ) sind offensichtlich nicht identisch mit den Eigenzuständen der schwachen Kraft ( $(u, c, t)$ , und  $(d', s', b')$ ). Man kann aber die schwachen Eigenzustände  $(d', s', b')$  als eine Drehung im Raum der starken Eigenzustände verstehen:

$$\begin{pmatrix} d' \\ s' \\ c' \end{pmatrix} = \underbrace{\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}}_{V_{CKM}} \cdot \begin{pmatrix} d \\ s \\ c \end{pmatrix}. \quad (2.7)$$

Die Matrix  $V_{CKM}$  wird nach **C**abibbo, **K**obayashi und **M**askawa CKM-Matrix genannt. Die Wahrscheinlichkeit, dass ein Quark des Flavours  $x$  in ein Quark des Flavours  $y$  zerfällt ist proportional zum Betragsquadrat des CKM-Matrixelements  $|V_{xy}|^2$ . Aus der Forderung nach Unitarität ( $V^\dagger V = 1$ ) werden die freien Parameter der CKM-Matrix auf drei Winkel und eine Phase eingeschränkt, sofern nur drei Generationen von Quarks existieren. Diese müssen experimentell bestimmt werden. In der sogenannten generalisierten Wolfenstein-Parametrisierung lässt sich die CKM-Matrix in einer Näherung in dritter Ordnung in  $\lambda$  auf folgende Form bringen (siehe hierzu [6]):

$$V_{CKM} = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\bar{\rho} - i\bar{\eta}) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \bar{\rho} - i\bar{\eta}) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^5), \quad (2.8)$$

wobei  $A \approx 4/5$  und  $\lambda \approx 0.22$  beträgt.<sup>2</sup> In dieser Form sieht man, dass die Diagonalelemente der CKM-Matrix nahe bei eins liegen, der Übergang von Quarks innerhalb der eigenen

<sup>2</sup> $\lambda$  stammt aus einer Formulierung der schwachen Übergänge durch *Nicola Cabibbo*, als nur die Quarks  $u$ ,  $d$  und  $s$  bekannt waren.  $|\lambda|^2$  entsprach hier der Übergangswahrscheinlichkeit von einem  $s$ -Quark in ein  $u$ -Quark. Die Parameter  $\bar{\eta}$  und  $\bar{\rho}$  sind definiert als:

$$\bar{\rho} = \rho \left(1 - \frac{\lambda}{2}\right) \quad \bar{\eta} = \eta \left(1 - \frac{\lambda}{2}\right). \quad (2.9)$$

In dieser Formulierung beinhaltet die obige Parametrisierung auch Ordnungen in  $\lambda$  von  $\mathcal{O}(\lambda^4)$ . Die Abweichungen von der Unitarität sind in dieser Form vernachlässigbar. Ein nicht verschwindendes  $\eta$  hat CP-Verletzung zur Folge.

Generation ist demnach, sofern kinematisch erlaubt, sehr wahrscheinlich. Je weiter ein Element von der Diagonalen entfernt liegt, desto größer wird die Ordnung in  $\lambda$  und desto unwahrscheinlicher wird der Übergang.

Zur Berechnung der Zerfallsraten und Lebensdauern müssen die Matrixelemente der CKM-Matrix berücksichtigt werden. Für den Zerfall des  $\bar{B}^0$ -Mesons hat dies eine Unterdrücken mit  $|V_{cb}|^2$  zur Folge. Die  $\bar{B}^0$ -Lebensdauer wird dann wie folgt aus Fermis goldener Regel berechnet:<sup>3</sup>

$$\tau_{\bar{B}^0} = \frac{192\pi^3}{G_F^2 |V_{cb}|^2 m_B^5} . \quad (2.10)$$

Hierbei ist  $G_F$  die *Fermi-Konstante*. Die Variable  $m_B$  ist die Masse des b-Hadrons. Obiges gilt für alle Zerfälle von b-Hadronen, bei denen das b-Quark in ein c-Quark übergeht. Als Konsequenz des kleinen CKM-Matrixelements für diese Übergänge ist die b-Hadron-Lebensdauer trotz der großen b-Hadron-Massen erstaunlich lang ( $\tau_{\bar{B}^0} \approx 1.5$  ps [4]). Prinzipiell kann ein b-Quark auch in ein u-Quark zerfallen. Dieser Übergang ist aber nach Gleichung 2.8 noch stärker unterdrückt. Die lange Lebensdauer ist eine Eigenart von b-Hadronen und kann zur experimentellen Identifikation dieser verwendet werden. Dies wird in Kapitel 7 ausführlicher dargestellt.

### Der Higgs-Mechanismus

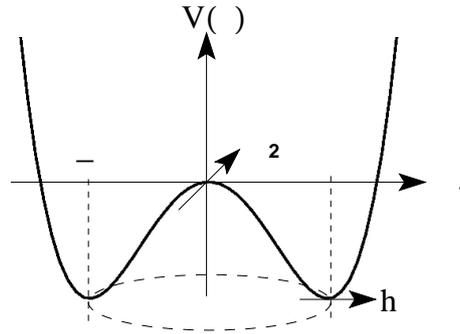
Die Theorie des Standardmodells, basierend auf der Quantenfeldtheorie beinhaltet in der bisher erläuterten Form nur masselose Teilchen. Dies widerspricht allen bisherigen Beobachtungen. In der QED lassen sich Fermionen-Massen in die Lagrangedichte einführen ohne damit die lokale Eichinvarianz zu zerstören (Gleichung 2.4 beinhaltet bereits einen Massenterm). Versucht man allerdings Gleiches für die schweren Eichbosonen der schwachen Kraft  $W^\pm$  und  $Z^0$ , zerstört dies unumgänglich die lokale Eichinvarianz und damit die Symmetrie  $U(1)_Y \times SU(2)_L$ . Das Standardmodell verliert an diesem Punkt seine Selbstkonsistenz.

Schon 1964 wurde eine Theorie zur Lösung dieses Problems entwickelt: der *Englert-Brout-Higgs-Guralnik-Hagen-Kibble-Mechanismus* kurz auch *Higgs-Mechanismus*. Bis heute konnte diese Theorie nicht experimentell bestätigt werden. Eine detaillierte Beschreibung findet sich beispielsweise in [2].

Es werden zunächst vier skalare Higgs-Felder  $\phi_i$  eingeführt, die typischerweise in einem Isospin-Dublett angeordnet werden

$$\phi = \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} . \quad (2.11)$$

<sup>3</sup>Eine Durchführung der Berechnung findet sich beispielsweise in [3].

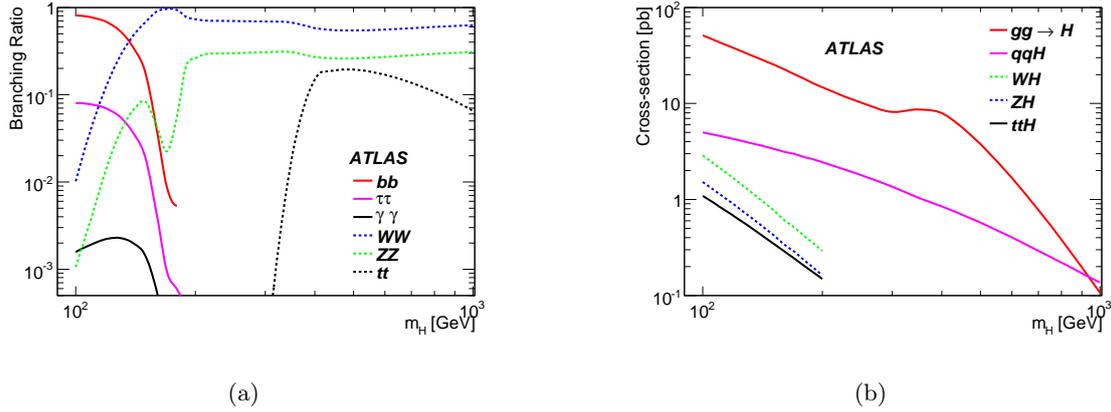


**Abb. 2.2.:** Schema des Higgs-Potenzials  $V(\phi)$ . Die Minima des Potenzials liegen in einem Kreis um den Ursprung bei einem Radius  $v$ .

Die Felder erzeugen in der Lagrangedichte einen Potenzialterm. Die Form des Potenzials kann so gewählt werden, dass die  $U(1)_Y \times SU(2)_L$ -Symmetrie erhalten bleibt:  $V = \frac{1}{2}\mu^2\phi^2 + \frac{1}{4}\lambda\phi^4$ . Wählt man  $\mu^2 < 0$  und  $\lambda > 0$  erhält man ein Potenzial der in Abbildung 2.2 gezeigten Form. Das Minimum des Potenzials ist so entartet. Die Wahl eines expliziten Minimums als Vakuumerwartungswert ist willkürlich. Durch die konkrete Wahl des Vakuumerwartungswerts wird die Symmetrie des Potenzials gebrochen, man spricht von einer spontanen Symmetriebrechung. Der Symmetriebruch hat ein sehr willkommenes Resultat zur Folge: Massenterme der Eichfelder. Für jedes der vier Higgs-Felder entsteht durch den Symmetriebruch nach dem Goldstone-Theorem ein sogenanntes Goldstone-Boson. Dies sind keine reellen Teilchen. Es ist möglich, sie als die zusätzlichen Freiheitsgrade der durch die Massen entstehenden longitudinalen Polarisierungen zu interpretieren. Drei der vier Goldstone-Bosonen werden damit von der longitudinalen Polarisierung der massebehafteten schwachen Eichbosonen absorbiert. Das masselose Photon allerdings absorbiert keinen Freiheitsgrad. Eines der Goldstone-Bosonen bleibt erhalten. Damit wird ein massives spinloses Boson vorhergesagt: Das Higgs-Boson. Dieses müsste, sofern die Theorie des Higgs-Mechanismus auf unsere Natur zutrifft, nachweisbar sein. Es existieren theoretische Argumente, die Higgs-Massen unterhalb von  $m_H < 1 \text{ TeV}$  favorisieren [7]. Von experimenteller Seite konnten bestimmte Massenbereiche ausgeschlossen werden. An Experimenten bei LEP und auch heute am Tevatron wurde und wird nach dem Higgs-Boson gesucht. Durch LEP wurde ein Massenbereich bis  $m_H = 114 \text{ GeV}/c^2$  ausgeschlossen. Das Tevatron konnte bis dato einen Bereich von  $163 \text{ GeV}/c^2 < m_H < 166 \text{ GeV}/c^2$  ausschließen [8].

Das Higgs-Boson koppelt an Masse. Es zerfällt aus diesem Grund bevorzugt in schwere Teilchen. In Abbildung 2.3(a) sind die Verzweungsverhältnisse verschiedener Zerfallskanäle in Abhängigkeit der Higgs-Masse gezeigt. Für kleine Higgs-Massen  $m_H < 135 \text{ GeV}/c^2$  dominiert der Zerfall in zwei b-Quarks. Für Higgs-Massen oberhalb dieser

Grenze dominieren Zerfälle in  $W$ - und  $Z$ -Bosonen. In Abbildung 2.3(b) sind auch die Wirkungsquerschnitte verschiedener Produktionskanäle des Higgs-Bosons in Abhängigkeit der Higgs-Masse für den LHC zu sehen. Am LHC dominiert für nahezu alle betrachteten Higgs-Massen die Gluon-Fusion ( $gg \rightarrow H$ ).



**Abb. 2.3.:** (a) Verzweungsverhältnisse verschiedener Higgs-Zerfälle, (b) Wirkungsquerschnitte der Higgs-Produktionskanäle am LHC, beides in Abhängigkeit der Higgs-Masse. Abbildungen aus [9].

Der Higgs-Mechanismus ist keine Erweiterung des Standardmodells, sondern das letzte notwendige noch fehlende Puzzleteil. Zusammen mit einem nachgewiesenen Higgs-Boson wäre das Standardmodell in einem weiteren Schritt bravourös bestätigt. Erst der Nachweis des Higgs-Bosons macht das Standardmodell vollständig. Es ist eine der Hauptaufgaben der Experimente am LHC, ein Higgs-Boson zu entdecken oder aber seine Existenz auszuschließen.

### 2.1.3. Proton-Proton-Kollisionen

Die vorgestellten Quantenfeldtheorien sind in der Lage, eine Vorhersage für Wirkungsquerschnitte von z.B. einer Quark-Quark- oder Gluon-Gluon-Streuung zu machen. Auf Grund der Bindung der Quarks in Hadronen ist eine solche Vorhersage aber nicht praktisch überprüfbar. Die Vorhersagen solcher Wechselwirkungen müssen auf im Experiment durchführbare Hadronen-Streuungen erweitert werden. Im LHC-Ring werden Protonen beschleunigt, die frontal zur Kollision gebracht werden. Theoretische Beschreibungen solcher Proton-Proton-Kollisionen erfordern eine gute Kenntnis des Protonenaufbaus.

Die Basis des Protons bilden die sogenannten *Valenzquarks uud*. Unter diesen werden ständig Gluonen ausgetauscht. Diese Gluonen können für kurze Zeit Quark-Antiquark-Paare bilden. Solche kurzlebigen Quarks im Proton heißen *Seequarks*. Alle das Proton

aufbauenden Teilchen, also Valenzquarks, Seequarks und Gluonen, werden Partonen  $q_i$  genannt. Es sei nun die folgende Streuung zweier Partonen  $q_i, q_j$  von Interesse:  $q_i q_j \rightarrow A$ .  $A$  ist der Endzustand der Parton-Parton-Wechselwirkung, dies könnte beispielsweise ein W- oder Z-Boson oder im Falle eines QCD-Prozesses ein Quark-Antiquark- oder ein Gluon-Paar sein. Der experimentell beobachtbare Prozess hat dann die Form:  $pp \rightarrow A + X$ , wobei  $X$  alle Teilchen im Endzustand zusammenfasst, die nicht durch die betrachtete Parton-Parton-Streuung entstanden sind. Der Subprozess:  $q_i q_j \rightarrow A$  bildet den *harten* Anteil der Proton-Proton-Streuung. Um den Wirkungsquerschnitt einer solchen Proton-Proton-Streuung berechnen zu können, muss auch der andere Anteil an der Streuung berücksichtigt werden, der durch die zusätzlichen Partonen im Proton hervorgerufen wird. Dieser Teil wird durch die Parton-Dichtefunktionen (PDF, engl. *parton density function*)  $f_{q_i}(x_i, Q^2)$  angegeben. Die PDFs beschreiben das Parton  $q_i$  in Abhängigkeit vom Impulsübertrag  $Q^2$  und vom Impulsanteil  $x_i$ , welches das Parton am Protonenimpuls hat. Das Integral über einen kleinen Bereich  $x_i + dx$  gibt dann die Wahrscheinlichkeit an, dass das Parton  $q_i$  einen Impulsanteil des Protonenimpulses des Intervalls  $(x_i, x_i + dx_i)$  trägt. Die Abhängigkeit vom Impulsübertrag ist eine Folge des energieabhängigen Auflösungsvermögens. Für kleine  $Q^2$  sind zunächst nur die Valenzquarks sichtbar. Für größere  $Q^2$  nimmt dagegen der Anteil der Seequarks und Gluonen am Impuls des Protons zu. Zusammen mit dem aus der QCD erhaltenen Wirkungsquerschnitt  $\sigma(q_i, q_j \rightarrow A)$  des *harten* Streuprozesses lässt sich nun der Wirkungsquerschnitt einer Proton-Proton-Streuung wie folgt aufschreiben:

$$\sigma(pp \rightarrow A + X) = \sum_{i,j} \int f_{q_i}(x_i, Q^2) f_{q_j}(x_j, Q^2) \sigma(q_i q_j \rightarrow A) dx_i dx_j . \quad (2.12)$$

Aus den Impulsanteilen der streuenden Partonen  $(x_i, x_j)$  und der vom Beschleuniger gelieferten Schwerpunktsenergie  $\sqrt{s}$  der Protonen ergibt sich die Schwerpunktsenergie des *harten* Prozesses:  $\sqrt{\hat{s}} = \sqrt{x_i x_j s}$ .

Die Impulsanteile der Partonen  $x_i$  und  $x_j$  sind dabei nicht bekannt. Desweiteren ist zu berücksichtigen, dass das Schwerpunktsystem der Parton-Parton-Streuung nicht dem Laborsystem entspricht, sondern bezogen auf dieses in longitudinaler Richtung einen unbekanntem Lorentz-Boost erhält. Man wählt zur Untersuchung solcher Ereignisse deshalb Variablen, die invariant sind unter einer Lorentztransformation. Hierzu gehört der transversale Impuls  $p_T$  und Differenzen von Pseudorapiditäten  $\Delta\eta$ . Rapiditäten ergeben sich aus der Energie  $E$  und dem longitudinalen Impuls  $p_z$  durch  $y = \ln[(E - p_z)/(E + p_z)]$ . Im Grenzfall masseloser Teilchen ergibt sich die allein vom Streuwinkel  $\theta$  abhängige Pseudorapidität  $\eta = -\ln(\tan\theta/2)$ .

Direkt nach der Streuung existieren eine Reihe von Partonen. Die Kraft zwischen diesen, sich entfernenden Partonen wird mit größer werdendem Abstand immer stärker und es entstehen Quark-Antiquark-Paare aus dem Vakuum, die sich wiederum von einander

entfernen. Dieser Vorgang wird so oft wiederholt, bis die Energie zur Paarbildung nicht mehr ausreicht. Alle entstandenen Quarks und Gluonen sind nun in Hadronen gebunden. Dieser Vorgang wird *Hadronisierung* genannt. Ein einzelnes Gluon oder Quark, das an der initialen Wechselwirkung teilgenommen hat, hinterlässt eine große Zahl an Hadronen und anderen Zerfallsprodukten wie Elektronen oder Myonen, die sich kegelförmig um die Richtung dieses Teilchens ausbreiten. Eine solche gebündelte Ansammlung von Teilchen wird Jet genannt. Je nach Ausgangsteilchen wird ein Jet als Gluon- oder Quark-Jet eines bestimmten Quark-Flavours bezeichnet.

## 2.2. Die Maschinen: Der LHC-Ring und der ATLAS-Detektor

### 2.2.1. Der LHC-Ring

Der **L**arge **H**adron **C**ollider (LHC) ist ein Beschleunigerring für Hadronen. Die Anlage befindet sich am Forschungszentrum CERN bei Genf. Mit einem Umfang von 27 km ist der LHC der größte bisher gebaute Hadronen-Beschleuniger. Der Ring liegt in einem Tunnel ca. 100 m unter der Erdoberfläche, der zuvor von LEP zur Elektron-Positron-Beschleunigung genutzt wurde. Im LHC sollen Bündel von Protonen à  $1.15 \cdot 10^{11}$  Teilchen beschleunigt werden. Protonen erreichen dabei eine Energie von 7 TeV. Um dies zu ermöglichen, sind hohe technische Anforderungen an den Beschleuniger gestellt. Eine wichtige Voraussetzung ist z.B. ein sehr gutes Vakuum in den Strahlröhren, innerhalb derer die Protonen den Ring durchlaufen. Die Beschleunigung wird durch ein hochfrequentes Wechselfeld von 5 MV/m durch acht supraleitende Kavitäten erreicht. Auch die Dipolmagnete, 1232 an der Zahl, die die Protonen auf ihrer Kreisbahn halten sollen, sind supraleitend. Es werden magnetische Feldstärken bis 8.33 T erreicht. Um die Magnete in supraleitendem Zustand halten zu können, werden sie auf bis zu 1.9 K gekühlt. Durch supraleitende Quadrupolmagnete werden die Protonenbündel fokussiert.

Die Protonen durchlaufen ein Verfahren zur Vorbeschleunigung, bevor sie in den LHC eingespeist werden. In Abbildung 2.4 ist eine schematische Darstellung der Anlagen am LHC zu sehen. Die Protonen werden zunächst im PS<sup>4</sup>, dann im SPS<sup>5</sup> bis auf 450 GeV beschleunigt. Die vorbeschleunigten Protonen werden sowohl in als auch gegen den Uhrzeigersinn in den LHC eingespeist. Innerhalb von 20 Minuten sollen die Protonen gegenläufig in zwei Strahlrohren auf ihre nominelle Endenergie gebracht und auf dieser gehalten werden. In je einem Strahlrohr sollen ca. 3000 Protonenbündel zirkulieren.

Die Protonenstrahlen kreuzen sich an vier Punkten des Rings. Hier werden nach Errei-

---

<sup>4</sup>Proton Synchrotron

<sup>5</sup>Super Proton Synchrotron

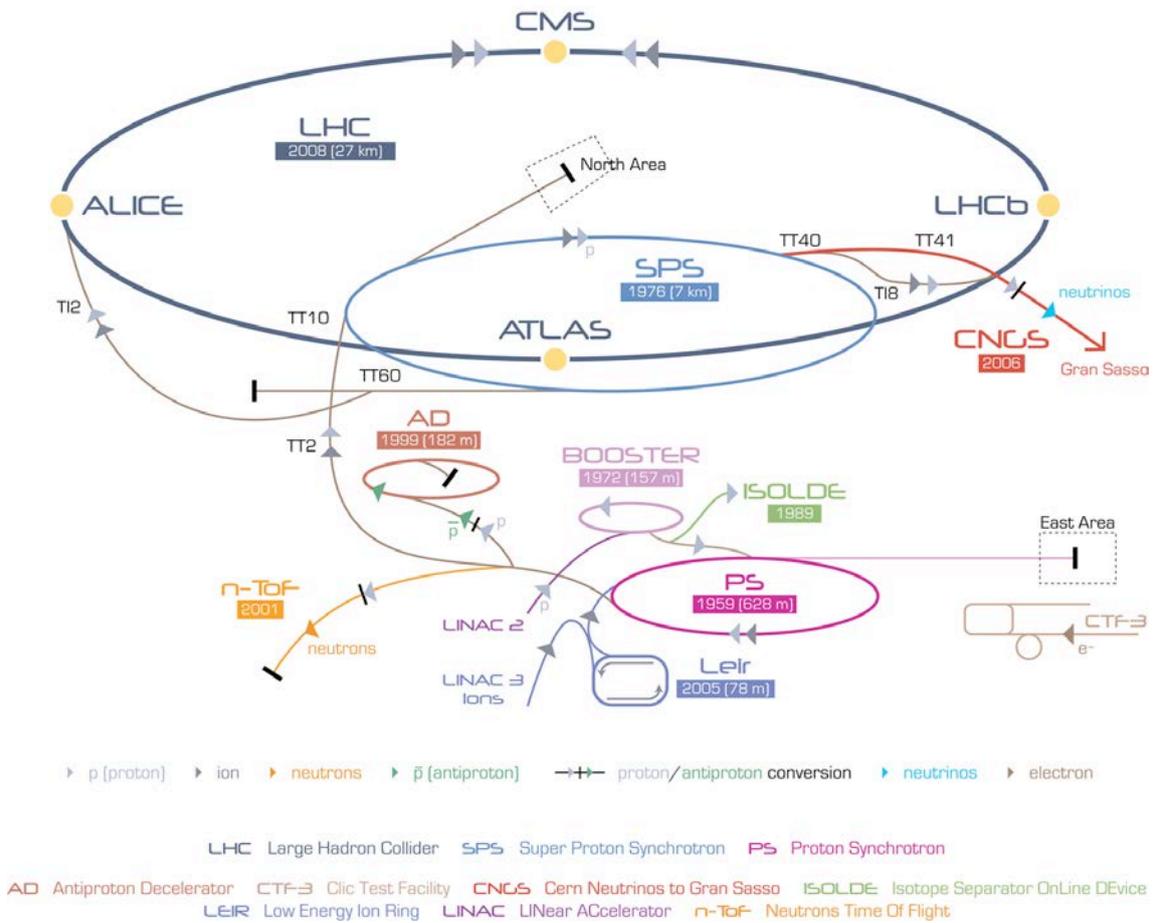


Abb. 2.4.: Schematische Darstellung der Beschleunigeranlagen des LHCs. Abbildung aus [10].

chen der gewünschten Energie die Protonenbündel zur Kollision gebracht. Für die nominelle Protonenenergie hat dies eine Schwerpunktsenergie der Kollision von  $\sqrt{s} = 14 \text{ TeV}$  zur Folge. Geplant sind Kollisionen der Bündel alle 25 ns, was einer Kollisionsfrequenz von 40 MHz entspricht. Für den LHC ist eine Luminosität von  $\mathcal{L} = 10^{34} \text{ cm}^{-2}\text{s}^{-1}$  vorgesehen. Dies führt zu sehr hohen Ereignisraten, die benötigt werden, um auch sehr selten auftretende Prozesse nachweisen zu können. In Kapitel 3 wird sehr viel genauer auf die Strahleigenschaften, die Luminosität und die damit zusammenhängenden Ereignisraten eingegangen.

Die Streuprodukte der Kollisionen werden an jedem Kreuzungspunkt von großen Detektoren nachgewiesen (siehe Abbildung 2.4). Die Experimente und die zugehörigen Detektoren sind auf unterschiedliche Forschungsziele ausgelegt:

**LHCb:** LHCb<sup>6</sup> ist auf Prozesse mit b-Quarks spezialisiert.

**ALICE:** ALICE<sup>7</sup> ist spezialisiert auf die mögliche Produktion von Quark-Gluon-Plasmen. Hierfür werden schwere Ionen im LHC beschleunigt und im Zentrum von ALICE zur Kollision gebracht.

**CMS:** CMS<sup>8</sup> ist ein Vielzweckexperiment.

**ATLAS:** ATLAS<sup>9</sup> ist ebenso ein Vielzweckexperiment.

ATLAS und CMS sind in erster Linie auf die Entdeckung des Higgs-Bosons und auf Physik, die über das Standardmodell hinaus geht, ausgelegt.

### 2.2.2. Der ATLAS-Detektor

Sehr hohe Ereignisraten und große Teilchen-Multiplizitäten bei den Ereignissen stellen einen hohen Anspruch an die Detektoren am LHC. Desweiteren fordern die Forschungsziele eine sehr detaillierte Beobachtung seltener Prozesse. Folgende Bedingungen sind damit an die Detektoren gestellt:

- Alle Bauteile des Detektors müssen einer sehr hohen Strahlenbelastung Stand halten.
- Die Totzeiten der elektronischen Bauteile müssen ausreichend kurz sein, so dass alle 25 ns ein neues Ereignis aufgezeichnet werden kann.
- Ein schnelles, effizientes Triggersystem ist notwendig um Ereignisse von Interesse zu selektieren. Ohne ein solches wäre die zu speichernde Datenmenge zu groß.

---

<sup>6</sup>Large Hadron Collider beauty

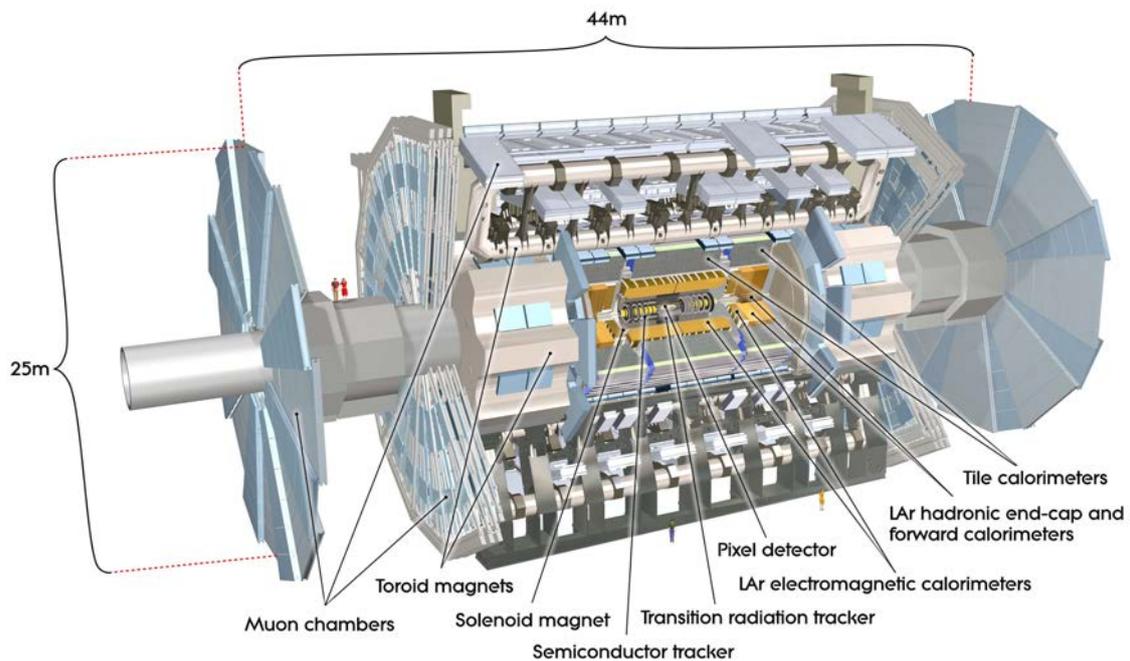
<sup>7</sup>A Large Ion Collider Experiment

<sup>8</sup>Compact Muon Solenoid

<sup>9</sup>A Toroidal LHC Apparatus

- Maximale räumliche Abdeckung, insbesondere für einen großen Pseudorapiditäts-Bereich ist notwendig.
- Ein sehr gutes Auflösungsvermögen der Impulse und eine sehr hohe Rekonstruktions-effizienz der Spuren von geladenen Teilchen nahe am Wechselwirkungsbereich wird benötigt.
- Zur Identifikation von Elektronen und Photonen ist eine sehr gute elektromagnetische Kalorimetrie nötig. Die Messung von Jets fordert eine hadronische Kalorimetrie mit möglichst voller Raumabdeckung.
- Eine gute Myon-Identifikation und eine gute Auflösung der Myon-Impulse für einen breiten Impulsbereich, ebenso wie die Bestimmung der Ladung von Myonen mit sehr großem transversalen Impuls, sind für viele Analysen fundamental.

Der ATLAS-Detektor wird diesen Anforderungen gerecht. Der Aufbau des Detektors soll in der Folge beschrieben werden. Eine detaillierte Beschreibung findet sich in [11]. Alle verwendeten Zahlen und Bilder sind, sofern nicht anders ausgezeichnet, ebenfalls aus diesem Artikel entnommen.



**Abb. 2.5.:** Schema des ATLAS-Detektors. Abbildung aus [11].

Der ATLAS-Detektor ist mit einer Höhe von 25 m und einer Breite von 44 m der Größte der vier Detektoren am LHC. Er ist zylinderförmig aufgebaut. Abbildung 2.5 zeigt den Detektor in einer schematischen Darstellung. Im Zentrum des Detektors finden die Kollisio-

nen statt. Hier liegt auch der Ursprung des ATLAS-Koordinatensystems. Der Zylinderform des Detektors angepasst werden zur weiteren Beschreibung Zylinderkoordinaten verwendet. Die Strahlrichtung definiert die  $z$ -Achse, der Streuwinkel ist der Polarwinkel  $\theta$  und der Azimutalwinkel ist  $\phi$ . An Stelle des Polarwinkels wird häufig die Pseudorapidität  $\eta$  verwendet. Der Abstand zweier Streuteilchen  $\Delta R$  im  $(\eta, \phi)$ -Raum wird als  $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$  definiert. Dabei ist  $\Delta\eta = |\eta_1 - \eta_2|$  und damit auch  $\Delta R$  lorentzinvariant.

Der Detektor ist in Schichten verschiedener Detektortypen aufgebaut, wobei jede Schicht Mantel- und Deckelflächen eines Zylinders bildet. Die Deckelflächen werden Endkappen genannt. Die Hauptkomponenten, die in Abbildung 2.5 bezeichnet sind, lassen sich in folgender Struktur vom Inneren zum Äußeren des Detektors zusammenfassen:

**Der innere Detektor:** Dieser ist der Kern des ATLAS-Detektors, bestehend aus den, im Zentrum liegenden Pixel-Detektoren (*Pixel detector*), dem Halbleiter-Spurdetektor (*Semiconductor tracker* (SCT)) und dem Übergangsstrahlungsdetektor (*Transition radiation tracker* (TRT)).

**Das Kalorimetersystem:** Den inneren Detektor umhüllend schließt sich das elektromagnetische Kalorimeter (*LAr electromagnetic calorimeters*) an. Das hadronische Kalorimeter (*Tile calorimeters*) bildet die nächste Schicht.

**Das Myon-System:** Die Detektoren des Myon-Systems (*Muon chamber*) sind die äußersten Elemente des ATLAS-Detektors.

Desweiteren befinden sich zwei große Magnetsysteme im ATLAS-Detektor. Der innere Detektor liegt in einem Magnetfeld der Stärke 2 T eines Solenoid-Magneten (*Solenoid Magnet*), das Myon-System in einem ca. 0.3 T starken Feld eines Toroid-Magnetsystems (*Toroid Magnets*). Eine Beschreibung des inneren Detektors soll später etwas ausführlicher zusammen mit einer kurzen Beschreibung der Rekonstruktion von Spuren gegeben werden. Zuvor werden hier die Komponenten der anderen Systeme des Detektors dargestellt.

**Kalorimetrie** Die Kalorimetrie dient der Energiemessung von Teilchen, wobei das elektromagnetische Kalorimeter zuständig ist für die Messung der Energie von Photonen und Elektronen, während die Energie von Teilchen, die stark wechselwirken, im hadronischen Kalorimeter bestimmt wird. Photonen und Elektronen lösen beim durchqueren von Materie eine Kaskade von Bremsstrahlungsprozessen ( $e \rightarrow e\gamma$ ) und Photon-Konversionen ( $\gamma \rightarrow e^+e^-$ ) aus. Dabei entsteht eine Vielzahl von neuen Elektronen, Positronen und Photonen, die solange die oben beschriebenen Prozesse durchführen, bis ihre kinetische Energie hierzu nicht mehr ausreicht. Die Kalorimeter besitzen aktive Schichten. Im Falle des elektromagnetischen Kalorimeters sind dies Schichten aus flüssigem Argon, wodurch die Teilchen nachgewiesen werden können. So kann im Kalorimeter ein sogenannter Schauer

beobachtet werden, anhand dessen die Energie des ins Kalorimeter einlaufenden Teilchens bestimmt werden kann. Dabei ist essentiell, dass der Schauer vollständig innerhalb des Kalorimeters stattfindet. Um trotz allem die Ausdehnung des Kalorimeters beschränken zu können sind zwischen den aktiven Schichten inaktive Bleischichten angebracht. Innerhalb dieser ist die Wahrscheinlichkeit für Bremsstrahlung und Photon-Konversion größer, so dass dies das Aufschauern beschleunigt. Auch alle anderen geladenen Teilchen nehmen am Prozess der Bremsstrahlung teil. Allerdings nimmt die Bremsstrahlungswahrscheinlichkeit mit der Masse ab. Aus diesem Grund ist die Ausdehnung des elektromagnetischen Kalorimeters ausreichend groß, um Elektronen- und Photonen-Schauer in diesem vollständig nachzuweisen. Teilchen, die schwerer sind als das Elektron, werden dagegen im Allgemeinen auch Energie hinter dem elektromagnetischen Kalorimeter deponieren. Stark wechselwirkende Teilchen verursachen auf Basis von starken Prozessen einen hadronischen Schauer, der im elektromagnetischen Kalorimeter beginnt und bis in das hadronische Kalorimeter reicht. Die Energie stark wechselwirkender Teilchen kann so aus der Information des elektromagnetischen und des hadronischen Kalorimeter bestimmt werden. Auch das hadronische Kalorimeter ist schichtweise angeordnet, wobei hier Plastikszieltoren als aktives Material und Stahl als inaktives Material verwendet wird.

**Myon-Spektrometer** Die Lebensdauer des Myons ist ausreichend lang, so dass Myonen mit genügend Energie den Rand des Detektors erreichen. Aufgrund der im Vergleich zum Elektron höheren Masse des Myons und der damit verbundenen geringeren Bremsstrahlungswahrscheinlichkeit und der Tatsache, dass es nicht stark wechselwirkt, deponiert das Myon nur wenig Energie in den Kalorimetern. Myonen erreichen somit als einzige die eigens zum Nachweis dieser gedachten Myonen-Spektrometer am Rand des ATLAS-Detektors. Eine Vielzahl von unterschiedlichen Spurdetektoren wird zur Spurmessung von Myonen verwendet. Hierzu zählen z.B. Driftröhren und Kathoden-Streifendetektoren. Durch die Bahnen, die die Myonen im Magnetfeld des Toroidmagnetsystems beschreiben, kann eine vom inneren Detektor unabhängige Bestimmung der Ladung und des Impulses von Myonen durchgeführt werden.

In Tabelle 2.2 sind die erwarteten Auflösungen für die Messung des transversalen Impulses der Myonen und für die Kalorimetrie für den jeweils gültigen Pseudorapiditätsbereich aufgetragen. Im Falle des hadronischen Kalorimeters wird der Pseudorapiditätsbereich durch einen Vorwärts-Detektor bis zu  $|\eta| < 5$  erweitert.

Detektor-Komponente	Auflösung	$\eta$ -Bereich
Elektromagnetische Kalorimetrie	$\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$	$ \eta  < 3.2$
Hadronische Kalorimetrie	$\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$	$ \eta  < 3.2$
Myonspektrometer	$\sigma_{p_T}/p_T = 10\%$ bei $p_T = 1 \text{ TeV}/c$	$ \eta  < 2.7$

**Tab. 2.2.:** Erwartete Auflösung der Energiemessung für elektromagnetische und hadronische Kalorimetrie in einem gegebenen Pseudorapiditätsbereich. Ebenso ist die Auflösung des Myonimpulses alleine durch die Messung der Myonspektrometer zu sehen.  $p_T$  und  $E$  sind dabei in  $GeV/c$  bzw. in  $GeV$  angegeben. Das  $\oplus$  steht für eine quadratische Addition. Tabelle aus [11].

**Triggersystem** An dieser Stelle soll das Triggersystem des ATLAS-Detektors kurz angesprochen werden. Es sorgt dafür, dass nur Ereignisse aufgezeichnet werden, die Hinweise darauf liefern, dass ein Prozess von Interesse vorliegt. Das Triggersystem enthält drei Ebenen, wobei in jeder Ebene härtere Kriterien an die Ereignisse gestellt werden. Der *Level1*-Trigger sucht nach Ereignissen, die Teilchen mit hohem transversalen Impuls enthalten oder aber eine große gesamte transversale oder fehlende Energie aufweisen. Es werden hierfür Informationen aus dem Myon-Spektrometer und den Kalorimetern verwendet. Die Entscheidung, ob ein Ereignis aufgezeichnet wird oder nicht, wird innerhalb von  $2 \mu s$  gefällt. Der *Level1*-Trigger markiert interessante Bereiche im Detektor. Diese werden vom *Level2*-Trigger, der nächsten Triggerstufe, genauer untersucht. Zuletzt wird durch den Ereignisfilter die aufzuzeichnende Datenmenge um ein Weiteres eingeschränkt. Hierfür stehen bereits Algorithmen zur Verfügung, wie sie in der offline-Datenanalyse angewendet werden.

### Der innere Detektor und die Rekonstruktion von Spuren

Der innere Detektor widmet sich dem Nachweis geladener Teilchen, die beim Durchqueren des Detektormaterials eine Spur hinterlassen. Bis zu 1000 Spuren müssen dabei pro Ereignis vom inneren Detektor wahrgenommen werden. Die Ansprüche an das Auflösungsvermögen sind aus diesem Grund sehr hoch. Vor allem in den Bereichen sehr nahe am Wechselwirkungspunkt liegen die Spuren sehr eng. Aus diesem Grund ist die Granularität des Detektors innen sehr groß und nimmt nach außen hin ab. Eine gute Spurauflösung ist essentiell für die Rekonstruktion der Proton-Proton-Wechselwirkungspunkte, ebenso wie zur Rekonstruktion von Zerfallspunkten von z.B. b-Hadronen. Beides spielt innerhalb dieser Arbeit eine sehr wichtige Rolle. Es soll an dieser Stelle etwas ausführlicher auf den inneren Detektor und seine Komponenten sowie die Rekonstruktion von Spuren aus den Messpunkten des Detektors eingegangen werden.

Der innere Detektor setzt sich aus drei Teilsystemen zusammen. Abbildung 2.6 zeigt

den Schnitt eines Quadranten des inneren Detektors. Die innerste Lage bilden die Pixel-Detektoren. Diese bestehen aus  $50 \times 400 \mu\text{m}$  großen Silizium-Pixeln. Sie erreichen damit eine intrinsische Auflösung von  $10 \mu\text{m}$  in der  $(R - \phi)$ -Ebene und  $115 \mu\text{m}$  in  $z$ -Richtung. Drei Lagen von Pixel-Detektoren sind in konzentrischen Zylindern in einem Radius von ca. 5 bis 12 cm um das Strahlrohr angeordnet. Ebenso befinden sich drei Lagen von Pixel-Endkappen scheibenförmig in Vorwärts- und Rückwärts-Richtung des Detektors. Sie liegen in  $z$ -Richtung zwischen 40 und 65 cm vom Ursprung des Detektors entfernt. Wie in Abbildung 2.6 zu erkennen ist, hinterlassen Teilchen, die sich in den zentralen Bereich des Detektors ausbreiten, im Allgemeinen drei Treffer in den Pixel-Detektoren. Durch den Einsatz der Endkappen ist dies bis zu einem Bereich  $|\eta| < 2.5$  möglich. Der innere Detektor ist darauf ausgelegt, Spuren innerhalb dieses Winkelbereichs präzise messen zu können. So sind auch die, den Pixel-Detektor umhüllenden, SCT-Detektoren so angeordnet, dass dieser Winkelbereich gut mit Detektormaterial abgedeckt ist. Der SCT besteht aus Silizium-Streifendetektoren. Die Streifen sind 4.6 cm lang und haben einen Abstand von  $80 \mu\text{m}$ . Zur Bestimmung der  $z$ -Position sind zwei Lagen solcher Streifendetektoren unter einem kleinen Winkel von 40 mrad gegeneinander eingebaut. In der zentralen Region des Detektors sind vier solcher Doppellagen konzentrisch um die Pixel-Detektoren angeordnet. Sie decken dabei einen radialen Bereich von ca. 30 bis 60 cm ab. Neun scheibenförmige Endkappen des SCT garantieren gute Spurerkennung für einen Bereich von  $|\eta| < 2.5$ . Die SCT-Streifendetektoren erreichen hierbei eine intrinsische Auflösung von  $17 \mu\text{m}$  in der  $(R - \phi)$ -Ebene und  $580 \mu\text{m}$  in  $z$ -Richtung. Pixel- und SCT-Detektoren bilden zusammen die Präzisions-Spurdetektoren. Die äußerste Lage des inneren Detektors ist der Übergangsstrahlungsdetektor (TRT). Dieser füllt einen Bereich bis zu einem Radius von etwas mehr als einem Meter um die Strahlachse aus. Jeder Messpunkt kann hier mit einer Auflösung von  $130 \mu\text{m}$  in der  $(R - \phi)$ -Ebene gemessen werden. Der TRT besteht aus Röhren mit einem im Zentrum liegenden feinen Draht. Die Röhren sind mit einem Xenon-Gasgemisch gefüllt. Bis zu 36 Messpunkte werden für zentrale Spuren im TRT erwartet. Die Abdeckung des TRT reicht hierbei nur bis zu einem  $\eta$ -Wert von  $|\eta| < 2.0$ .

Wie das Myon-System ist der innere Detektor durchsetzt von einem Magnetfeld. Geladene Teilchen werden hierdurch auf eine Kreisbahn in der  $(R - \phi)$ -Ebene gezwungen und beschreiben im inneren Detektor eine gekrümmte Spur. Aus der Krümmung kann der transversale Impuls bestimmt werden. Hierfür müssen aber aus allen Messpunkten im Ereignis zunächst gekrümmte Spuren rekonstruiert werden. Dies wird von der Spurrekonstruktion durchgeführt, wobei sich die Bestimmung der Spuren im Ereignis in zwei große Teilaufgaben zerlegen lässt: Unter allen Treffern im Ereignis müssen Sätze von Messpunkten herausgefiltert werden, die gute Spurkandidaten bilden. Für bis zu 1000 Spuren pro

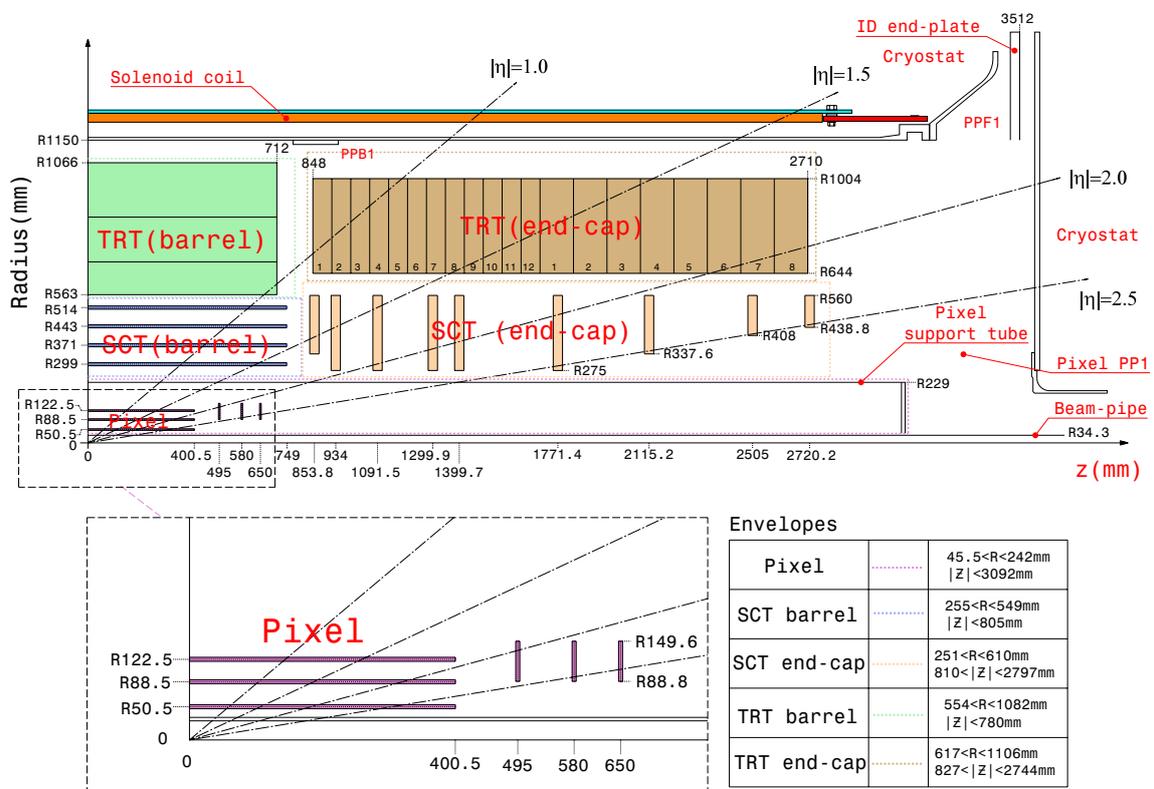


Abb. 2.6.: Schema eines Quadranten des inneren Detektors. Abbildung aus [11].

Ereignis ist dies eine anspruchsvolle Aufgabe. Alle Punkte eines Messpunktsatzes müssen desweiteren zu einer Spur verbunden werden. Verschiedene Techniken zur Bestimmung der Spurkandidaten werden verwendet, ebenso wie verschiedene Fit-Techniken zur präzisen Bestimmung des Spurverlaufs, darunter Kalman-Filter und globale  $\chi^2$ -Fits. Beide werden in Kapitel 4 im Zusammenhang mit der Rekonstruktion von Wechselwirkungspunkten erläutert. Eine ausführliche Beschreibung der Spurrekonstruktion findet sich in [9]. Die Spurrekonstruktion stellt am Ende für jede Spur die Helixparameter und die Kovarianzmatrizen aus den Spurbestimmungen zur Verfügung. Die Helixparameter sind die folgenden, wobei sie auf den Punkt der nächsten Annäherung der Spur an die Strahlachse in der  $(R - \phi)$ -Ebene P bezogen werden:

$q/p_T$ : Verhältnis aus der Ladung  $q$  des Teilchens und dem transversalen Impuls  $p_T$ , gemessen im Punkt P.

$\phi$ : Azimutal-Winkel im Punkt P.

$\cot(\theta)$ : Der Kotangens des Polarwinkels im Punkt P.

$d_0$ : Abstand zwischen P und dem Ursprung in der  $(R - \phi)$ -Ebene (*transversaler Impact-Parameter*).

$z_0$ : z-Position des Punktes P (*longitudinaler Impact-Parameter*).

Die erreichte Auflösung dieser Parameter variiert je nach Stärke des transversalen Impulses der Spur. Aus diesem Grund wird die Auflösung eines Spurparameters  $X$  in Abhängigkeit vom transversalen Impuls  $p_T$  bestimmt:

$$\sigma_X(p_T) = \sigma_X(\infty)(1 \oplus p_x/p_T) . \quad (2.13)$$

Dabei ist  $p_x$  eine Konstante,  $\sigma_X(\infty)$  ist die Auflösung für den Grenzwert eines unendlich großen transversalen Impulses und  $\oplus$  weist auf eine quadratische Addition hin. In Tabelle 2.3 sind  $p_x$  und  $\sigma_X(\infty)$  für zwei verschiedene Pseudorapiditäts-Bereiche ( $0.25 < |\eta| < 0.50$ ) und ( $1.5 < |\eta| < 0.75$ ) für die verschiedenen Spurparameter angegeben.

Spurparameter	$0.25 <  \eta  < 0.50$		$1.5 <  \eta  < 1.75$	
	$\sigma_X(\infty)$	$p_X [\text{GeV}/c]$	$\sigma_X(\infty)$	$p_X [\text{GeV}/c]$
$q/p_T$	$0.34 (\text{TeV}/c)^{-1}$	44	$0.41 (\text{TeV}/c)^{-1}$	80
$\phi$	$70 \mu\text{rad}$	39	$92 \mu\text{rad}$	49
$\cot(\theta)$	$0.7 \cdot 10^{-3}$	5.0	$1.2 \cdot 10^{-3}$	10
$d_0$	$10 \mu\text{m}$	14	$12 \mu\text{m}$	20
$z_0$	$91 \mu\text{m}$	2.3	$71 \mu\text{m}$	3.7

**Tab. 2.3.:** Erwartete Auflösung  $\sigma_X(\infty)$  der Spurparameter  $X$  für einen gegen unendlich gehenden transversalen Impuls. Die Auflösungen werden für verschiedene  $\eta$ -Bereiche betrachtet. Tabelle aus [9].

## 2.3. Die Simulation

Wie in Abschnitt 2.1.3 beschrieben, wird im ATLAS-Detektor nicht die initiale Parton-Parton-Streuung zu beobachten sein, sondern eine Fülle von Teilchen, meist gebündelt in Jets, die aus dieser resultieren. Verschiedenartige Parton-Parton-Streuungen können zum selben Endzustand führen und gleiche Parton-Parton-Interaktionen können sehr unterschiedliche Endzustände besitzen. Aus der theoretischen Kenntnis der Prozesse können Vorhersagen für die verschiedenen Endzustände eines Prozesses gemacht werden, ebenso wie für die Häufigkeiten, mit denen verschiedene Prozesse stattfinden werden. Dies kann zur Generation von simulierten Ereignissen genutzt werden, anhand derer schon im Vorfeld abgeschätzt werden kann, wie sich bestimmte Messgrößen verhalten werden, sofern sich die zugrunde liegende Theorie bestätigt. Zusätzlich zur Generation der Ereignisse ist es dann auch notwendig, die Messung eines Teilchens durch den Detektor zu simulieren. Generierte Ereignisse, die eine Detektorsimulation durchlaufen haben, können ganz analog zu tatsächlichen Daten analysiert werden. Datensätze mit simulierten Ereignissen bieten die Möglichkeit, schon im Vorfeld der Datennahme alle Rekonstruktions- und Analysealgorithmen auf ihre Leistungsfähigkeit hin zu testen, da die vollständige Information des stattgefundenen Prozesses zur Verfügung steht.

### 2.3.1. Ereignisgeneration

Eine Ereignisgeneration beginnt mit der initialen Streuung, im Falle des LHCs mit Protonen und gibt als Ergebnis die stabilen bzw. quasi-stabilen Teilchen des Endzustands aus. Von beiden Protonen werden zwei Partonen ausgewählt, die am harten Streuprozess teilnehmen sollen. Diese führen gemäß den theoretischen Wirkungsquerschnitten einen Prozess aus. Für die aus diesem Prozess entstandenen Teilchen und den nach dem Stoß noch

Prozess	Generator
$pp \rightarrow t\bar{t}$	MC@NLO
$pp \rightarrow WH_{120}(H \rightarrow b\bar{b}/u\bar{u})$	PYTHIA
$pp \rightarrow H_{120} \rightarrow \gamma\gamma$	PYTHIA
<i>Minimum-Bias</i> (siehe Kapitel 3)	PYTHIA

**Tab. 2.4.:** In dieser Arbeit verwendete simulierte Datensätze mit den jeweiligen Ereignisgeneratoren. Der Index am Higgs-Boson bezeichnet die Higgs-Masse in  $\text{GeV}/c^2$ . Beschreibungen der physikalischen Prozesse finden sich in Kapitel 3 und 5.

erhaltenen Partonen werden nun sogenannte Partonschauer simuliert. Dies beinhaltet z.B. die Abstrahlung von Gluonen. Es folgt die Simulation der Hadronisierung. Die Information über jedes in dieser Kette vorkommenden Teilchens ist nachträglich zugänglich. Insofern ist man nun in der Lage aus dem Endzustand einen Einblick in den Produktionsprozess zu erhalten. Die Simulation von Ereignissen basiert einerseits auf dem theoretischen Verständnis des Prozess, andererseits aber auch auf bereits vorhandenen Sätzen von experimentellen Daten, mit denen die Simulation abgeglichen werden kann. Die Basis der Simulationen sind *Monte-Carlo*-Generatoren.

Alle Ergebnisse in dieser Arbeit sind auf simulierten Ereignissen erarbeitet. Zwei verschiedene Ereignisgeneratoren wurden zur Simulation dieser Ereignisse verwendet: PYTHIA [12] und MC@NLO [13]. Die Simulation aller Prozesse berücksichtigt dabei im Falle von PYTHIA nur die ersten Ordnungen der Störungsrechnung. Im Falle von MC@NLO wird dies bis zur zweiten Ordnung fortgesetzt.

In den Studien dieser Arbeit sind verschiedene simulierte Datensätze für unterschiedliche physikalische Prozesse verwendet worden. Diese sollen hier aufgelistet werden. Verwendung finden sie vor allem in den Kapiteln 5, 6 und 7. Tabelle 2.4 zeigt die verschiedenen verwendeten Datensätze und die zugehörigen Generatoren.

### 2.3.2. Detektorsimulation

Für die Simulationen des ATLAS-Detektors wird auf Basis von GEANT [14] die Wechselwirkung der Teilchen mit dem Detektor simuliert. Aus der Detektorsimulation resultieren simulierte Signale, wie sie vom Detektor erwartet werden. Für den ATLAS-Detektor existiert eine sehr detaillierte Beschreibung des Detektors. Simulationen anhand dieser Beschreibung heißen *volle* Simulation. In der vorliegenden Arbeit werden ausschließlich Ereignisse mittels der vollen Detektorsimulation verwendet.

---

Anhand der Ereignis- und Detektorsimulation ist man in der Lage, schon vor der Datennahme jeden Schritt in der Rekonstruktion und in den Analysen zu testen. Mit Hilfe der Kenntnis des simulierten Prozesses kann eine Aussage gemacht werden, wie leistungsstark das untersuchte Rekonstruktions- oder Analysewerkzeug ist. Man ist so in der Lage, vor dem tatsächlichen Start des Experiments abzuschätzen, ob ein aus der Theorie vorhergesagtes Modell entdeckt werden kann oder nicht. So haben Studien auf solchen simulierten Monte-Carlo-Daten gezeigt, dass ein vom Standardmodell vorhergesagtes Higgs-Boson beispielsweise am LHC gefunden werden sollte. Es sei angemerkt, dass die Unsicherheiten, vor allem auf die Ereignisgeneration, sehr groß sind. Das Arbeiten auf simulierten Daten gibt Indizien und macht es möglich Werkzeuge zu testen. Die Realität wird aber mit großer Wahrscheinlichkeit anders aussehen. Alle Monte-Carlo-Daten müssen deshalb, sobald experimentelle Daten vorhanden sind, mit richtigen Daten verglichen und entsprechend korrigiert werden.



## 3. *Pile-Up* und *Minimum-Bias*

Am LHC sind sehr hohe Energien und Luminositäten geplant. Dies ermöglicht Ereignisraten, die auch selten auftretende Prozesse beobachtbar machen. Gleichzeitig hat dies jedoch zur Folge, dass auch die Rate von Untergrundereignissen groß ist. Untergrundereignisse, die bei allen Analysen auftreten werden, sind die so genannten *Minimum-Bias*-Ereignisse. Die Überlagerung mehrerer *Minimum-Bias*-Wechselwirkungen in einer Strahlkreuzung wird *Pile-Up* genannt. Eine anspruchsvolle Umgebung, wie sie am LHC vorliegt, lässt sehr hohe *Pile-Up*-Raten erwarten. Es ist eine wichtige Aufgabe, *Pile-Up*-Phänomene zu verstehen und eine Trennung der zur Untersuchung stehenden Prozesse vom *Minimum-Bias*-Untergrund ist unerlässlich. Die Selektion der gewünschten Wechselwirkung unter allen, in einer Strahlkreuzung stattfindenden, Wechselwirkungen ist das Hauptthema dieser Arbeit.

### 3.1. *Minimum-Bias*-Ereignisse

Im ATLAS-Detektor sollen Bündel von  $1.5 \cdot 10^{11}$  Protonen bei einer Schwerpunktsenergie von  $\sqrt{s} = 14$  TeV zur Kollision gebracht werden. Bei einer solchen Kollision werden Protonenpaare inelastisch zusammenstoßen, wobei die Anzahl der stattfindenden inelastischen Prozesse vom totalen inelastischen Wirkungsquerschnitt  $\sigma_{inel}$  und der Luminosität  $\mathcal{L}$  abhängt. Die einlaufenden Protonen können bei einem inelastischen Stoß entweder erhalten bleiben (doppelt-diffraktiv,  $\sigma_{dd}$ ), beide zerstört werden (nicht-diffraktiv,  $\sigma_{nd}$ ) oder aber einer der beiden Wechselwirkungspartner bleibt im Endzustand erhalten (einfach-diffraktiv,  $\sigma_{ed}$ ). Aus diesen Anteilen setzt sich dann der totale inelastische Wirkungsquerschnitt zusammen:

$$\sigma_{inel} = \sigma_{dd} + \sigma_{nd} + \sigma_{ed} . \quad (3.1)$$

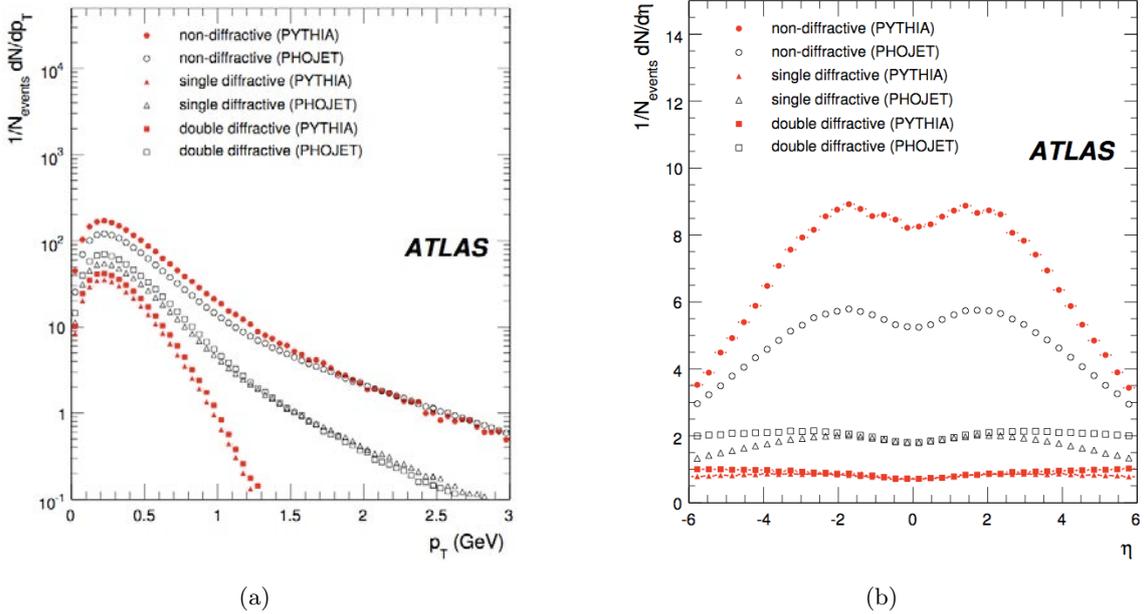
Diese inelastischen Stöße werden *Minimum-Bias*-Ereignisse genannt. Die Bezeichnung *Minimum-Bias* kommt aus dem Englischen und lässt sich mit *minimal verzerrt* übersetzen. Der Trigger für *Minimum-Bias*-Ereignisse soll einen möglichst geringen verzerrenden Einfluss auf die Selektion der aufzuzeichnenden Kollisionen haben. Man zeichnet im Fall eines *Minimum-Bias*-Triggers einerseits Ereignisse auf, die Hinweise auf einen inelastischen Stoß geben, andererseits versucht man, dabei möglichst keine weiteren Einschränkungen

zu machen.<sup>1</sup> Für einen Großteil der aufgezeichneten Ereignisse erwartet man einen relativ geringen Impulsübertrag. Ein geringer Impulsübertrag äußert sich in einer starken Vorwärts-Rückwärtsausrichtung der Kollisionsprodukte, bezogen auf die Strahlachse. Dies wird sichtbar, wenn man das Spektrum des transversalen Impulses  $p_T$  und der Pseudorapidität  $\eta$  aller Spuren dieser Ereignisse betrachtet. Abbildung 3.1 zeigt die Spektren der Kollisionsprodukte aus *Minimum-Bias*-Ereignissen zweier unterschiedlicher Ereignisgeneratoren: PHOJET[15] und PYTHIA[12]. Die Ereignisse wurden hier in die einzelnen, den totalen inelastischen Wirkungsquerschnitt zusammensetzenden Beiträge eingeteilt, also in doppelt-, einfach- und nicht-diffraktiv. Im Spektrum der transversalen Impulse (Abbildung 3.1(a)) ist für alle betrachteten Fälle die oben gemachte Aussage bestätigt. Die Verteilung fällt mit größer werdendem transversalen Impuls für einen Wert oberhalb von  $p_T > 0.5 \text{ GeV}/c$  stark ab. Neben den Unterschieden durch die verschiedenen Simulationen, die in dieser Arbeit keine weitere Rolle spielen, ist zu beobachten, dass nicht-diffraktive Ereignisse im Durchschnitt häufiger hohe transversale Impulse aufweisen als die anderen *Minimum-Bias*-Beiträge. Das  $\eta$ -Spektrum (Abbildung 3.1(b)) zeigt, dass die bei der Kollision produzierten Teilchen aus *Minimum-Bias*-Ereignissen tendenziell in Vorwärts- und Rückwärtsrichtung des Detektors gestreut werden. Die Pseudorapidität ist wie in Abschnitt 2.1.3 beschrieben eine Variable, die nicht-linear vom Streuwinkel  $\theta$  abhängt. Man erwartet für eine in  $\theta$  gleichverteilte Variable eine größere Teilchendichte bei kleinen  $\eta$ -Werten, als es das Spektrum der *Minimum-Bias*-Streuprodukte zeigt. Tendenziell streuen *Minimum-Bias*-Ereignisse daher in die Vorwärts- und Rückwärtsrichtung des Strahls.

## 3.2. *Pile-Up*

Bei jeder Kollision zweier Protonenbündel mit im Mittel  $1.5 \cdot 10^{11}$  Teilchen, in der Folge Strahlkreuzung genannt, werden im ATLAS-Detektor gemäß des Wirkungsquerschnitts  $\sigma_{inel}$  und der Luminosität  $\mathcal{L}$  mehrere Protonenpaare inelastisch zusammenstoßen. Im Detektor überlagern sich die Streuprodukte dieser Proton-Proton-Wechselwirkungen. Die Überlagerung mehrerer Proton-Proton-Wechselwirkungen aus derselben Strahlkreuzung wird *in-time-Pile-Up* genannt. Es existieren auch weitere *Pile-Up*-Arten, die in dieser Arbeit allerdings nicht betrachtet werden. Hierzu gehören beispielsweise *out-of-time-Pile-Up*, Überlagerungen mit Streuprodukten aus Kollisionen vorangegangener Strahlkreuzungen oder der *Cavern-Untergrund*, Überlagerungen mit Streuprodukten aus Kollisionen mit dem Detektormaterial. Auch die den Detektor erreichende kosmische Untergrundstrahlung

<sup>1</sup>Einfach-diffraktive Ereignisse werden in der Praxis nicht zur Klasse der *Minimum-Bias*-Ereignisse gezählt, da bei diesen ein Ungleichgewicht in Vorwärts- und Rückwärtsrichtung des Detektors vorliegt. Für solche Ereignisse wäre ein deutlich komplexeres und einschränkenderes Trigger-Kriterium notwendig.

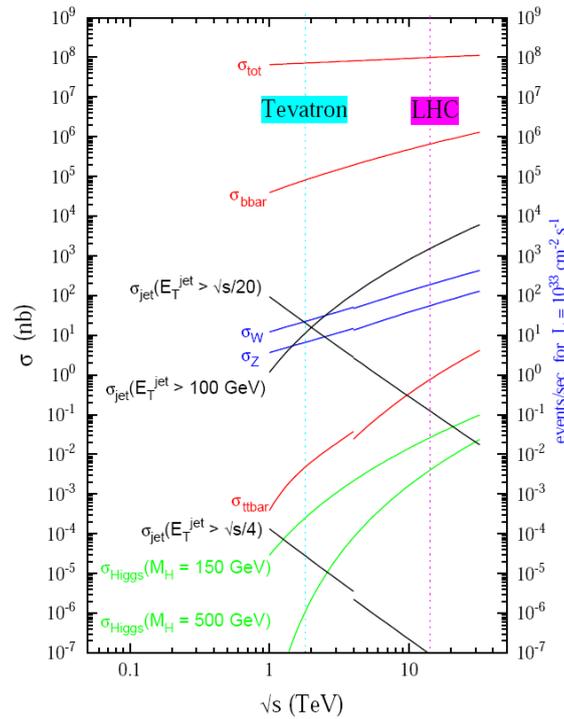


**Abb. 3.1.:** (a) Spektrum des transversalen Impulses  $p_T$  und der Pseudorapidität  $\eta$  (b) der Streuprodukte von *Minimum-Bias*-Ereignissen. Abbildungen aus [9].

zählt zu den *Pile-Up*-artigen Untergründen. *In-time-Pile-Up* stellt am LHC mit Abstand den größten Untergrund unter den oben beschriebenen *Pile-Up*-Arten dar. Der Begriff *Pile-Up* wird in dieser Arbeit ausschließlich in der Bedeutung des *in-time-Pile-Ups* verwendet.

Im Allgemeinen gilt das Interesse am LHC der Produktion schwerer Teilchen. Dazu ist ein großer Impulsübertrag notwendig. Der Wirkungsquerschnitt für solche, meist sehr harten speziellen Ereignisse ist zum Teil um viele Größenordnungen kleiner als der oben erwähnte gesamte totale inelastische Wirkungsquerschnitt. In Abbildung 3.2 sind die Wirkungsquerschnitte für verschiedene Prozesse am Tevatron und am LHC im Vergleich zu sehen. Der totale inelastische Wirkungsquerschnitt, in dieser Abbildung  $\sigma_{\text{tot}}$ , überragt mit einer Größenordnung von  $10^8$  nb alle anderen gezeigten Wirkungsquerschnitte. Dagegen haben zum Beispiel Prozesse, die ein Higgs-Boson beinhalten, am LHC einen um 10 Größenordnungen kleineren Wirkungsquerschnitt ( $10^{-7} - 10^{-2}$  nb). Diese Prozesse werden sehr viel seltener auftreten. In dieser Arbeit werden diese seltenen, sehr harten Wechselwirkungen, die von physikalischem Interesse sind, ganz allgemein als *Signalwechselwirkungen* bezeichnet. Um die aufgezeichnete Datenmenge möglichst gering zu halten, wird zur Untersuchung solcher Signalprozesse das Trigger-Kriterium so eingestellt, dass nur Ereignisse aus Strahlkreuzungen aufgezeichnet werden, in denen es Hinweise auf eine Signalwechselwirkung gibt. Falls weitere *Minimum-Bias*-Wechselwirkungen in derselben

Strahlkreuzung stattgefunden haben, überlagern diese den Signalprozess. Somit stellen *Minimum-Bias*-Prozesse für alle Analysen einen Untergrund dar und in jedem rekonstruierten Ereignis muss der Anteil des Signalereignisses am Gesamt ereignis herausgefiltert werden. Je mehr zusätzliche Wechselwirkungen stattgefunden haben, desto schwieriger wird diese Aufgabe. In den folgenden Abschnitten soll eine Abschätzung gemacht werden, wie viele *Minimum-Bias*-Wechselwirkungen man am LHC zusätzlich zum Signalprozess für die gegebene Schwerpunktsenergie und die angestrebten Luminositäten erwartet.



**Abb. 3.2.:** Erwartete Wirkungsquerschnitte (linke Achse) bzw. Ereignisraten für  $\mathcal{L} = 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$  (rechte Achse) in Abhängigkeit der Schwerpunktsenergie am Tevatron und am LHC für verschiedene Prozesse. Abbildung aus [16].

### 3.2.1. *Pile-Up* Vorhersagen für den LHC

Die Anzahl der zusätzlichen *Minimum-Bias*-Prozesse hängt vom Wirkungsquerschnitt  $\sigma_{inel}$  (Gl. 3.1), der erreichten Luminosität und dem Abstand zwischen den Protonenbündeln ab. Der Wirkungsquerschnitt kann aus der theoretischen Kenntnis des Prozesses berechnet werden. In dieser Arbeit wird nach [17] von einem totalen inelastische Wirkungsquerschnitt von  $\sigma_{inel} \approx 70 \text{ mb}$  ausgegangen.

Der LHC ist darauf ausgelegt, mit einer Luminosität von  $\mathcal{L}_{vorgesehen} = 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$

betrieben zu werden. Man erwartet jedoch für die Anfangsphase eine wesentlich geringere Luminosität. Zwei der hier betrachteten Szenarien für den Beginn des LHC-Betriebs sehen Luminositäten von  $\mathcal{L}_{A1} = 10^{32} \text{ cm}^{-2}\text{s}^{-1}$  und  $\mathcal{L}_{A2} = 10^{33} \text{ cm}^{-2}\text{s}^{-1}$  vor. Die Luminosität ergibt sich aus folgenden Parametern des Strahls, der Anzahl der Protonen pro Bündel  $N_p$ , der Umlauffrequenz des Strahls  $f_{\text{Strahl}}$ , der Anzahl der Bündel  $n_B$  im Strahl und der überlappenden Fläche  $\sigma_*^2$  der kollidierenden Bündel zu:

$$\mathcal{L} = \frac{N_p f_{\text{Strahl}} n_B}{4\pi\sigma_*^2} . \quad (3.2)$$

Die Strahlparameter werden im Laufe des Betriebs des LHCs perfektioniert und man erwartet im Laufe der Zeit eine Steigerung der Luminosität bis zur vorgesehenen Luminosität von  $\mathcal{L} = 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ .

Aus der Luminosität und dem Wirkungsquerschnitt lässt sich die Ereignisrate für inelastische Proton-Proton-Stöße berechnen:

$$\dot{N} = \frac{dN}{dt} = \mathcal{L} \cdot \sigma_{inel} . \quad (3.3)$$

Da kein kontinuierlicher Strahl von Protonen zur Kollision gebracht wird, sondern einzelne Bündel von Protonen, muss man zur Berechnung der Anzahl der Wechselwirkungen pro Strahlkreuzung den zeitlichen Abstand (Bündelabstand  $\Delta t_B$ ) zwischen den Kollisionen in Betracht ziehen. Es wird davon ausgegangen, dass 20% aller Bündel keine Protonen enthalten. Dies streckt den Bündelabstand im Mittel um einen Faktor von 1.25 (siehe [17]). Daraus ergibt sich die Anzahl  $N_{MB}$  der Wechselwirkungen pro Kollision je zweier Bündel (dies entspricht einer Strahlkreuzung) wie folgt:

$$N_{MB} = \mathcal{L} \cdot \sigma_{inel} \cdot 1.25 \cdot \Delta t_B . \quad (3.4)$$

Der Wirkungsquerschnitt entspricht dabei einer Stoßwahrscheinlichkeit. Gemäß dieser Wahrscheinlichkeit verteilt sich die Anzahl von Stößen pro Strahlkreuzung binomial um den Mittelwert  $\langle N_{MB} \rangle$ . Bei einer Anzahl von  $10^{11}$  potenziellen Proton-Proton-Kollisionen ist der Übergang zu einer *Poisson*-Verteilung legitim. Die Anzahl der inelastischen Stöße zweier Protonen bei einer Strahlkreuzung verteilt sich demnach gemäß einer *Poisson*-Kurve, wobei sich der Mittelwert  $\langle N_{MB} \rangle$  dieser Verteilung nach Gleichung 3.4 berechnen lässt.

An dieser Stelle sollen drei Szenarien mit verschiedenen Wechselwirkungsmultiplizitäten angegeben werden (siehe Tabelle 3.1). Für die geplanten Anfangsluminositäten werden im Mittel 4.1 bzw. 6.9 inelastische Prozesse pro Strahlkreuzung erwartet. Für die in der Planung vorgesehene Luminosität des LHC ist dagegen mit im Mittel über 20 solcher Prozesse zu rechnen.

$\mathcal{L}$ [ $\text{cm}^{-2}\text{s}^{-1}$ ]	$10^{32}$	$10^{33}$	$10^{34}$
$\Delta t_B$ [ns]	450	75	25
$\langle N_{MB} \rangle$	4.1	6.9	23

**Tab. 3.1.:** Erwartete mittlere Wechselwirkungsmultiplizitäten  $\langle N_{MB} \rangle$  für drei *Pile-Up*-Szenarien am LHC. Für den Wirkungsquerschnitt wurde hier  $\sigma_{inel} = 73.6$  mb angenommen.

### 3.2.2. Die Wechselwirkungsregion

Die im LHC beschleunigten Protonenbündel werden kollimiert, bevor sie in den Detektoren zur Kollision gebracht werden. Für die Verteilung der Protonen im Strahl kann in alle Raumrichtungen eine *Gauß*-Verteilung angenommen werden. In transversaler Richtung, also in der Ebene senkrecht zum Strahl, wird die Standardabweichung der Protonendichteverteilung mit  $\sigma_{Strahl,x} = \sigma_{Strahl,y} = 16.7 \mu\text{m}$  angegeben. Entlang des Strahls, in longitudinaler Richtung, ist die Verteilung deutlich breiter. Hier wird eine Standardabweichung von  $\sigma_{Strahl,z} = 7.55$  cm angegeben (Angaben aus [18]). Aus den Protonendichteverteilungen kann die Verteilung der Dichte der Wechselwirkungswahrscheinlichkeit im Zentrum des Detektors berechnet werden. In der transversalen Ebene ist die sich ergebende Verteilung bei einem frontalen Zusammenstoß der beiden Bündel eine Multiplikation zweier *Gauß*-Verteilungen und demnach wiederum eine *Gauß*-Verteilung, mit einer Standardabweichung von

$$\sigma_{WR,x} = \sigma_{WR,y} = \frac{\sigma_{Strahl,x}}{\sqrt{2}} = 11.8 \mu\text{m} . \quad (3.5)$$

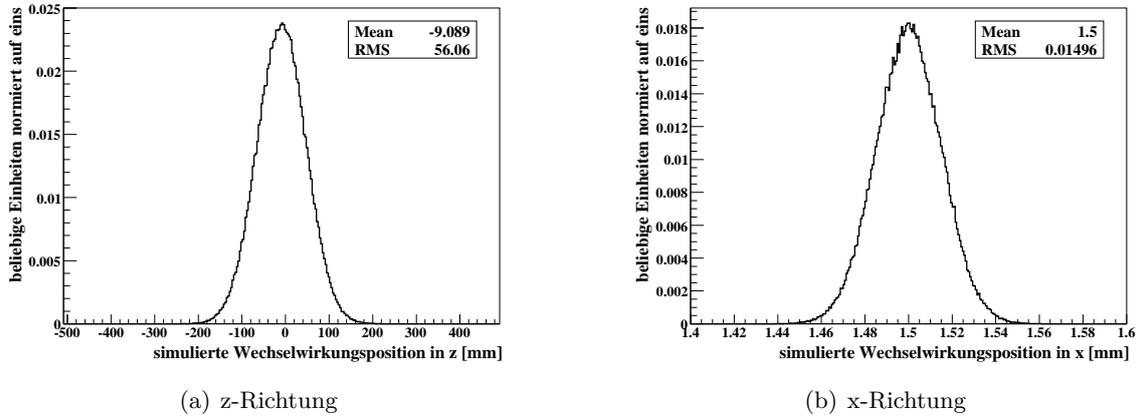
In  $z$ -Richtung bewegen sich die beiden Bündel durcheinander hindurch. Hier muss die Multiplikation einer *Gauß*-Verteilung  $g_{Strahl}$  mit einer um  $c \cdot t$  verschobenen gleichen *Gauß*-Verteilung für alle Zeiten des Durchlaufens aufaddiert werden:

$$\begin{aligned} g_{WR}(z) &= \int_{-\infty}^{\infty} g_{Strahl}(z - c \cdot t) \cdot g_{Strahl}(z + c \cdot t) dt \\ &= const \cdot \exp \left[ -\frac{z^2}{\sigma_{Strahl,z}^2} \right] . \end{aligned}$$

Die Funktion  $g_{WR}(z)$  entspricht demnach einer *Gauß*-Verteilung mit einer Standardabweichung von

$$\sigma_{WR,z} = \frac{\sigma_{z,Strahl}}{\sqrt{2}} = 53.4 \text{ mm} . \quad (3.6)$$

Die Gültigkeit dieser Werte ist nur für einen frontalen Zusammenstoß gegeben. In der Realität werden die Strahlen unter einem Winkel von  $300 \mu\text{rad}$  aufeinander stoßen. Die Ausdehnung des Wechselwirkungsbereichs wird hierdurch etwas größer. Alle bisher angegebenen Werte beruhen auf gut begründeten Annahmen, wie beispielsweise einer *Gauß*-Verteilung der Protonen in den Bündeln. Tatsächlich wird man diese Größe jedoch bei



**Abb. 3.3.:** Verteilung der Wechselwirkungen über die Wechselwirkungsregion.

Betrieb des LHCs messen müssen. Die Simulationen verwenden für die Ausdehnung der Wechselwirkungsregion etwas größere Werte von  $\sigma_{WR,x} = 15 \mu\text{m}$  und  $\sigma_{WR,z} = 56 \text{mm}$ . In Abbildungen 3.3(a) und 3.3(b) ist die Verteilung der simulierten Wechselwirkungspositionen über die Wechselwirkungsregion in  $z$  und in  $x$ -Richtung zu sehen. Um zu berücksichtigen, dass die Strahlen sich nicht genau im Ursprung des Detektors treffen werden, wurde der Kollisionsbereich in longitudinaler Richtung um 9 mm und in transversaler Richtung um 1.5 mm verschoben. Hierher rühren die von null abweichenden Mittelwerte.

Die große Ausdehnung der Wechselwirkungsregion in  $z$ -Richtung macht es möglich, die Positionen der verschiedenen Wechselwirkungen aus einer Strahlkreuzung voneinander zu unterscheiden. Je näher zwei Wechselwirkungen beieinander liegen, desto schwerer wird es sein, diese getrennt zu rekonstruieren. Liegen zwei Wechselwirkungspunkte sehr nahe beieinander, kann dies fälschlicherweise als eine einzelne Wechselwirkung interpretiert werden. Dies hat einen Einfluss auf die im Rahmen dieser Arbeit betrachtete Signalvertex-Selektion. Es ist insofern interessant, abschätzen zu können, wie wahrscheinlich es ist, bei einer gegebenen Dichteverteilung der Wechselwirkungspositionen zwei (oder mehr) Wechselwirkungspunkte in einem kleinen Intervall gegebener Größe zu finden. Bei einer Wechselwirkungsichte, die sich gemäß einer *Gauß*-Kurve verteilt, kann sehr einfach die Verteilung des Abstandes  $d_z$  zweier Wechselwirkungspunkte berechnet werden. Die Verteilung der Differenz zweier *Gauß*-verteilter Variablen ist auf eine Faltung zurückführbar. Sind die beiden *Gauß*-Verteilungen identisch, ist die Differenz der beiden Variablen ebenso gemäß einer *Gauß*-Kurve mit einer zugehörigen Standardabweichung von  $\sigma_{Dif} = \sqrt{2} \cdot \sigma_{WR,z}$  verteilt. Das Integral in einem bestimmten Intervall  $\mathcal{I}$  um den Mittelwert dieser Verteilung entspricht dann gerade der Wahrscheinlichkeit  $P_{d_z}$ , zwei Wechselwirkungen innerhalb des

Intervalls  $\mathcal{I}$  zu finden:

$$\mathcal{P}_{d_z} = \text{const} \cdot \int_{-\frac{\mathcal{I}}{2}}^{\frac{\mathcal{I}}{2}} \exp\left(-\frac{1}{2}\left(\frac{d_z}{\sigma_{Dif}}\right)^2\right) \cdot \quad (3.7)$$

Das Komplement  $(1 - \mathcal{P}_{d_z})$  entspricht dann der Wahrscheinlichkeit, dass die Wechselwirkungspunkte weiter als  $\mathcal{I}$  voneinander entfernt sind. Es sei  $N_{MB}$  die Anzahl der zusätzlichen Wechselwirkungen. Die Wahrscheinlichkeit, dass keine dieser  $N_{MB}$  Wechselwirkungspunkte im Intervall  $\mathcal{I}$  um den Signalwechselwirkungspunkt zu finden sind, ergibt sich wie folgt:

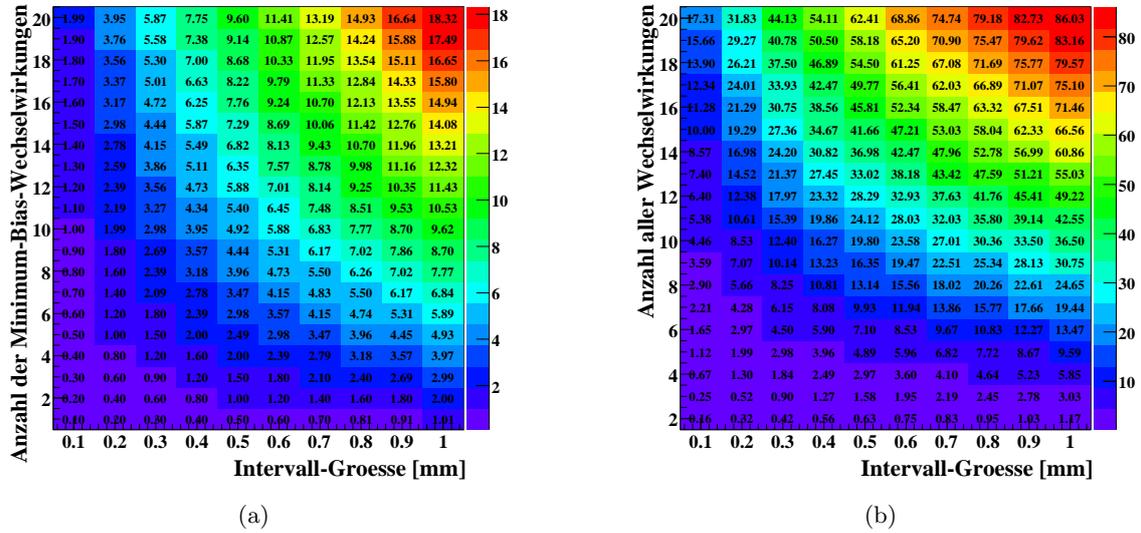
$$\mathcal{P}_{\text{kein Überlapp}} = \left(1 - \text{const} \cdot \int_{-\frac{\mathcal{I}}{2}}^{\frac{\mathcal{I}}{2}} \exp\left(-\frac{1}{2}\left(\frac{d_z}{\sigma_{Dif}}\right)^2\right)\right)^{N_{MB}} \cdot \quad (3.8)$$

Das Komplement von  $\mathcal{P}_{\text{kein Überlapp}}$  ergibt dann die gesuchte Wahrscheinlichkeit  $\mathcal{P}$ , mindestens eine *Minimum-Bias*-Wechselwirkung im Intervall  $I$  um den Signalwechselwirkungspunkt zu finden:

$$\mathcal{P} = 1 - \left(1 - \text{const} \cdot \int_{-\frac{\mathcal{I}}{2}}^{\frac{\mathcal{I}}{2}} \exp\left(-\frac{1}{2}\left(\frac{d_z}{\sigma_{Dif}}\right)^2\right)\right)^{N_{MB}} \cdot \quad (3.9)$$

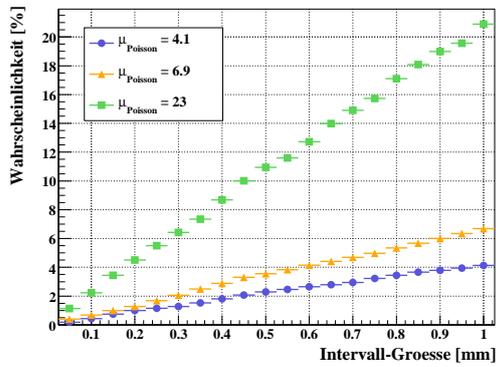
Abbildung 3.4(a) zeigt diese Wahrscheinlichkeit für verschiedene Intervallgrößen und verschiedene Anzahlen von zusätzlichen *Minimum-Bias*-Wechselwirkungen. Es ist desweiteren von Interesse, die Wahrscheinlichkeit abschätzen zu können, bei einer gegebenen Anzahl von Wechselwirkungen im Ereignis zwei oder mehr beliebige Wechselwirkungspunkte in einem Intervall  $\mathcal{I}$  zu finden, unabhängig davon, ob dabei die Signalwechselwirkung involviert ist oder nicht. Diese Wahrscheinlichkeit ist in Abbildung 3.4(b) zu sehen. Die Wahrscheinlichkeiten wurden hier anhand einer kleinen *Monte-Carlo*-Zufallsstudie ermittelt. Es wurde genügend oft eine gewisse Anzahl von Wechselwirkungspunkten zufällig aus der  $z$ -Wechselwirkungsverteilung (Abbildung 3.3(a)) gezogen und getestet, ob zwei oder mehr von diesen in einem bestimmten Intervall um einen beliebigen anderen liegen. Genügend oft bedeutet hier, dass die Anzahl der Tests groß genug ist, um ein statistisch aussagekräftiges Ergebnis zu erhalten. Die Wahrscheinlichkeiten sind natürlich deutlich höher als die in Abbildung 3.4(a) gezeigten. Man beachte die unterschiedliche Farbskala. Die Wahrscheinlichkeiten steigen hier für wachsende Wechselwirkungsmultiplizitäten und für größere Intervalle fast bis zu 90 % an. Die Überlagerung zweier *Minimum-Bias*-Wechselwirkungen im Ereignis kann eine Trennung der Signalwechselwirkung von den *Minimum-Bias*-Wechselwirkungen erschweren. Solche Effekte sind nach den obigen Studien schon für kleine Wechselwirkungsmultiplizitäten nicht zu vernachlässigen.

Wie bereits erwähnt, verteilt sich die Wechselwirkungsmultiplizität gemäß einer *Poisson*-Verteilung. Die Anzahl der *Minimum-Bias*-Wechselwirkungen wird demnach um einen, von der Luminosität und von dem Bündelabstand abhängenden, Mittelwert variieren. Die

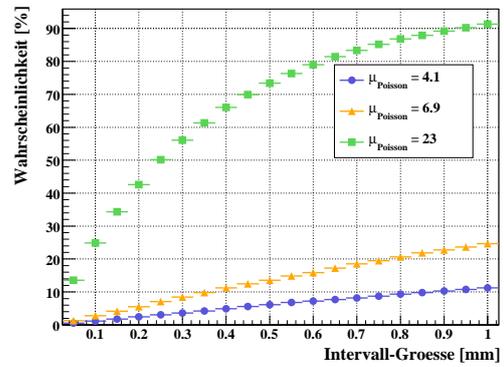


**Abb. 3.4.:** Wahrscheinlichkeit, bei gegebener Anzahl von zusätzlichen *Minimum-Bias*-Wechselwirkungen  $N_{MB}$  mindestens eine in einem gewissen Intervall um die Signalwechselwirkung zu finden (a). Wahrscheinlichkeit, bei einer gegebenen Anzahl von Wechselwirkungen ( $N_{MB} + 1$  inklusive der Signalwechselwirkung) zwei (oder mehr) beliebige Wechselwirkungen in einem gewissen Intervall zu finden (b).

oben betrachteten Wahrscheinlichkeiten lassen sich auch für solche variierenden Wechselwirkungsmultiplizitäten angeben. Es wurden hier die drei in Tabelle 3.1 beschriebenen *Pile-Up*-Szenarien betrachtet, also für  $\langle N_{MB} \rangle = 4.1, 6.9$  und  $23$  *Minimum-Bias*-Wechselwirkungen zusätzlich zum Signalprozess. In Abbildung 3.5 sind die Wahrscheinlichkeiten gegen die Intervallgröße für die drei *Pile-Up*-Szenarien zu sehen, wobei Abbildung 3.5(a) die Wahrscheinlichkeit zeigt, mindestens eine *Minimum-Bias*-Wechselwirkung innerhalb eines Intervalls um den Signalwechselwirkungspunkt zu finden. In Abbildung 3.5(b) ist hingegen die Wahrscheinlichkeit, mehrere beliebige Wechselwirkungspunkte innerhalb eines Intervalls zu finden aufgetragen. Die hier gezeigten Ergebnisse sind ebenfalls mit einer *Minimum-Bias*-Zufallsstudie berechnet worden. Die Unterschiede im Vergleich zu den Wahrscheinlichkeiten ohne Berücksichtigung der *Poisson*-Verteilung der Wechselwirkungsmultiplizität ist sehr klein. Mit Berücksichtigung der *Poisson*-Verteilung ist eine leichte Wahrscheinlichkeitssteigerung zu beobachten. Für eine Intervallgröße von  $0.4$  mm liegt die Wahrscheinlichkeit, dass zwei oder mehr beliebige Wechselwirkungen innerhalb dieses Intervalls liegen für das *Pile-Up*-Szenario mit im Mittel  $6.9$  *Minimum-Bias*-Wechselwirkungen, bei  $\mathcal{P} = 11.2\%$  (Abb. 3.5(b)). Der Vergleich mit Abbildung 3.5(a) zeigt dagegen für  $8$  Wechselwirkungen ( $6.9$  *Minimum-Bias*-Wechselwirkungen + eine Signalwechselwirkung) eine Wahrscheinlichkeit von  $10.8\%$ . Ursache für die leichte Steigerung sind die Ausläufer der *Poisson*-Verteilung zu großen Wechselwirkungsmultiplizitäten hin.



(a)



(b)

**Abb. 3.5.:** Wahrscheinlichkeit, dass eine Signalwechselwirkung mit mindestens einer *Minimum-Bias*-Wechselwirkung in einem gegebenen Intervall stattfindet (a) und Wahrscheinlichkeit, dass zwei beliebige Wechselwirkungen innerhalb eines solchen Intervalls stattfinden (b) für verschiedene Pile-Up-Szenarien. ( $\mu_{\text{Poisson}}$  entspricht hier der mittleren Anzahl der zusätzlichen *Minimum-Bias*-Wechselwirkungen ( $\langle N_{MB} \rangle$ ).

## 4. Primärvertex-Rekonstruktion in ATLAS

Im vorigen Kapitel ist aufgezeigt worden mit welchen *Pile-Up*-Raten man am LHC zu rechnen hat. Es zeigte sich, dass sich schon von Betriebsbeginn an mehrere Proton-Proton-Wechselwirkungen überlagern werden. Im Fall einer vorhandenen Signalwechselwirkung wird diese von meist mehreren *Minimum-Bias*-Wechselwirkungen überdeckt. Für die Rekonstruktion bedeutet dies, dass in jedem aufgezeichneten Ereignis zunächst eine Überlagerung der Spuren aus den Signal- und den *Minimum-Bias*-Wechselwirkungen im Detektor vorliegt. Jeweils verschiedene Sätze von Spuren gehen von verschiedenen Interaktionspunkten in der Wechselwirkungsregion aus. Die Punkte, von denen diese Spuren ausgehen werden Primärvertizes genannt. Dabei ist ein Satz von Spuren der zur Signalwechselwirkung gehörende. Die Rekonstruktion des Ausgangspunkts dieser Spuren ist essentiell für nachfolgenden Analysen. Es ist die Aufgabe der *Primärvertex-Rekonstruktion* den einzelnen Vertizes einen Satz von Spuren zuzuordnen und den Punkt des Ursprungs dieser Spuren zu rekonstruieren. Um den Ausgangspunkt der Signalwechselwirkung (Signalvertex) zu bestimmen ist es zweckmäßig alle Primärvertex-Positionen, also auch die der *Minimum-Bias*-Vertizes zu rekonstruieren, da a-priori nicht klar ist welche Spuren zu welchem Vertex gehören. Eine Zuordnung der Spuren ist leistungsstärker und flexibler wenn alle Vertex-Positionen bekannt sind. Rekonstruiert man alleine die Position des Signalvertex, ist die Wahrscheinlichkeit größer, dem Signalvertex Spuren aus *Minimum-Bias*-Wechselwirkungen zuzuordnen, bzw. Spuren, die von diesem ausgehen, nicht zu berücksichtigen. Ist hingegen bekannt, dass in der Nähe ein *Minimum-Bias*-Vertex liegt, kann über die Berücksichtigung dieser Information für jede Spur flexibel entschieden werden zu welchem Vertex sie am besten passt. Die Primärvertex-Positionen sind dabei nur entlang der Strahlrichtung unterscheidbar. Dies ist auf die, in Abschnitt 3.2.2 angesprochene Ausdehnung der Wechselwirkungsregion zurückzuführen, die in der transversalen Richtung nur  $\approx 1.5 \mu\text{m}$  in der longitudinalen Richtung jedoch  $\approx 50 \text{ mm}$  beträgt.

Die Rekonstruktion aller Primärvertizes im Ereignis hat weiterhin den Vorteil, dass erst in einem sehr späten Stadium, in dem bereits viel Information über die Vertizes zu Verfügung steht, entschieden werden muss, welcher der rekonstruierten Vertizes der Signalvertex ist. Die Bestimmung der Signalvertex-Position lässt sich daher in zwei Aufgaben zerlegen:

**Primärvertex-Rekonstruktion:** Bestimmung aller Primärvertex-Positionen im Ereignis und

die Zuordnung von Spuren zu den Vertizes.

**Signalvertex-Selektion:** Identifikation des Signalvertex unter allen rekonstruierten Primärvertizes. Hierdurch wird die  $z$ -Position bestimmt, von der die Spuren der Signalwechselwirkung ausgehen.

Eine ausführliche Studie zur Signalvertex-Selektion ist Thema dieser Arbeit und wird in Kapitel 5 und 6 vorgestellt. Dieses Kapitel widmet sich im Weiteren der Primärvertex-Rekonstruktion.

## 4.1. Algorithmen der Primärvertex-Rekonstruktion

Wie oben bereits angemerkt, beinhaltet die Primärvertex-Rekonstruktion mehrere Aufgaben. Diese sind:

**Spur-Selektion:** Dies beinhaltet die Selektion der Spuren, die von der Wechselwirkungsregion ausgehen. Hiermit werden Spuren unterdrückt, die aus Zerfällen langlebiger Teilchen herrühren und deren Extrapolationen nicht in die Wechselwirkungsregion zeigen.

**Vertex-Suche:** Es gilt Punkte im Raum bzw. in der Wechselwirkungsregion zu finden, in denen eine Häufung von Spuren auftritt. Die Vertex-Suche bestimmt demnach eine gewisse Anzahl potenzieller Vertex-Positionen mit jeweils einem Satz von zugeordneten Spuren. Jeder Satz von Spuren wird an den Vertex-Fit übergeben.

**Vertex-Fit:** Aus einem gegebenen Satz von Spuren und einer potenziellen Vertex-Position als Startwert wird durch  $\chi^2$ -Minimierung eine neue Vertex-Position bestimmt, wobei für die Spuren gefordert wird, dass sie sich in diesem Punkt treffen. Die Spurparameter werden hierdurch neu bestimmt und der Vertex-Fit gibt als Ergebnis neben der Vertex-Position mit Fehler die neuen Impulse aller Spuren am Vertex zurück.

Je nach Methode werden die Aufgaben der *Vertex-Suche* und des *Vertex-Fits* hintereinander ausgeführt und bleiben so voneinander unabhängig. Der in der Folge beschriebene *einfache* Algorithmus zur Vertex-Suche geht in dieser Weise vor. Ein weiterer, hier behandelte *adaptiver* Algorithmus, der in der Lage ist mehrere Vertizes gleichzeitig zu suchen, verwendet dagegen einen anderen Ansatz. In diesem Fall werden die potenziellen Vertex-Kandidaten erst bei der Anpassung der Spuren an den Vertex, also beim Vertex-Fit festgelegt.

Als Ausgangspunkt erhält die Primärvertex-Rekonstruktion von der Spurrekonstruktion alle  $N$  im Ereignis rekonstruierten Spuren. Jede Spur wird dabei durch ihre Helixparameter  $\vec{q}_i = (d_{0,i}, z_{0,i}, \phi_i, \theta_i, q_i/p_i)$  ( $i = 1, \dots, N$ ) bestimmt, wie sie in Abschnitt 2.2.2 be-

schrieben wurden. Zusätzlich stehen die Kovarianzmatrizen zu Verfügung, die sich aus den Unsicherheiten der Bestimmungen der Spuren ergeben. Die Primärvertex-Rekonstruktion übernimmt die Information aller  $N$  Spuren und gibt als Ergebnis eine Liste von Vertex-Positionen mit zugehörigen Spuren zurück, deren Parameter unter Berücksichtigung der Vertex-Position neu berechnet wurden.

Erst im zweiten Abschnitt 4.1.2 dieses Kapitels werden die Algorithmen zur Vertex-Suche beschrieben. Zunächst sollen im folgenden Abschnitt die Vertex-Fit-Methoden dargestellt werden, die den Algorithmen der Vertex-Suche zur Verfügung stehen. Die Vertex-Fit- und die Vertex-Suchalgorithmen sind in einer Form in der Software implementiert, die es ermöglicht, sie unabhängig voneinander zu verwenden. Auf die Implementierung von Rekonstruktionswerkzeugen wird in Kapitel 6 genauer eingegangen.

#### 4.1.1. Algorithmen zum Vertex-Fit

Der *Vertex-Fit* hat die Bestimmung einer Vertex-Position aus einem gegebenen Satz von Spuren zur Aufgabe. Es gilt den Kreuzungspunkt dieser Spuren aus den Parametern  $\vec{q}_i = (d_{0,i}, z_{0,i}, \phi_i, \theta_i, q_i/p_i)$  und deren Kovarianzmatrizen  $C_i$  zu bestimmen. Unter der Annahme, dass die Fehler einer *Gauß*-Verteilung folgen, kann aus der Information aller betrachteten Spuren eine räumliche Wahrscheinlichkeitsdichteverteilung in Abhängigkeit von der Vertex-Position  $\vec{r}$  berechnet werden. Anhand des Maximums dieser Verteilung lässt sich dann die optimale Vertex-Position für die gegebenen Spuren berechnen. Das Maximieren einer solchen Wahrscheinlichkeitsdichte ist äquivalent zu einer  $\chi^2$ -Minimierung, mit

$$\chi^2(\vec{r}) = \sum_{i=1}^N (\vec{r} - \vec{r}_i(\vec{q}_i))^T C_i^{-1} (\vec{r} - \vec{r}_i(\vec{q}_i)) . \quad (4.1)$$

Hier wird unter Berücksichtigung der Kovarianzmatrix  $C_i$  der Abstand der zu bestimmenden Vertex-Position  $\vec{r}$  zu den Punkten der nächsten Annäherung der Spuren  $\vec{r}_i(\vec{q}_i)$  minimiert.<sup>1</sup> Die  $\vec{q}_i$  entsprechen den gemessenen Spurparametern. Ohne weitere Einschränkung ist eine Minimierung von Gleichung 4.1 nicht analytisch lösbar. Es ist notwendig obige Gleichung so zu approximieren, dass nur eine lineare Abhängigkeit von den Fitparametern vorliegt. Selbst mit einer solchen Vereinfachung fordert die Lösung eine Invertierung einer Matrix der Dimension  $N$ , was eine große Rechenzeit beansprucht. Zwei Methoden sollen in der Folge vorgestellt werden, die die Invertierung einer großen Matrix umgehen und damit eine große Zeitersparnis erreichen können.

<sup>1</sup>Die Dimension der Kovarianzmatrix muss hierfür auf die Dimension des Vektors  $\vec{r}$  eingeschränkt werden. Dies ist grundsätzlich möglich, wobei die Diskrepanz der Dimensionen in den folgenden Anwendungen anders gelöst wird.

### Billoir-Fit-Methode

Anstatt direkt den Abstand zwischen der Vertex-Position und den Spuren zu minimieren ist es üblich, den Abstand zwischen den gemessenen Spurparametern  $\vec{q}_m$  und den Parametern  $\vec{q}_i(\vec{r}, \vec{p}_i)$  derselben Spur, die durch die Position  $\vec{r}$  des Vertex läuft, zu minimieren. Dabei werden die neuen Spurparameter zusätzlich in Abhängigkeit von den neuen Impulsen  $\vec{p}_i$  am Vertex geschrieben. Die  $\vec{p}_i$  bilden so zusätzliche Fitparameter. Die zu minimierende Größe schreibt sich nun wie folgt:

$$\chi^2(\vec{r}) = \sum_{i=1}^N (\vec{q}_i(\vec{r}, \vec{p}_i) - \vec{q}_{m,i})^T C_i^{-1} (\vec{q}_i(\vec{r}, \vec{p}_i) - \vec{q}_{m,i}) . \quad (4.2)$$

Fordert man von den  $\vec{q}_i$  eine lineare Abhängigkeit von  $\vec{r}$  und  $\vec{p}_i$ , muss man diese auf folgende Form bringen:

$$\vec{q} = \vec{q}_0 + A[\vec{r} - \vec{r}_0] + B[\vec{p} - \vec{p}_0] . \quad (4.3)$$

Die Matrizen  $A$  und  $B$  sind die Jakobi-Matrizen der Entwicklung des Zustands von  $\vec{r}_0$  nach  $\vec{r}$  und von  $\vec{p}_0$  nach  $\vec{p}$ . Die Bestimmung dieser Matrizen in einer Näherung erster Ordnung sind in [19] zu finden. In diesem Artikel ist die hier nur skizzierte Methodik eines Billoir-Fits detailliert beschrieben. Einsetzen der linearen Approximation für  $\vec{q}(\vec{r})$  in Gleichung 4.2 reduziert die Abhängigkeit des  $\chi^2$  von den Fitparametern auf eine lineare Abhängigkeit. Unter Anwendung des Billoir-Formalismus beschränkt sich die Notwendigkeit einer Invertierung einer  $N \times N$ -Matrix auf lediglich  $N$  Matrizen der Dimension  $5 \times 3$ . Ein solcher Fitalgorithmus ist bei ATLAS unter dem Namen *FullVertexFitter* implementiert. Unter der Annahme, dass sich die Impulse der Spuren in der Nähe des Vertex nur wenig ändern, kann man diese beim Fit unberücksichtigt lassen. Dies führt zu einer weiteren Reduktion der Rechendauer. Die zu bestimmenden variablen Spurparameter beschränken sich dann auf  $d_0$  und  $z_0$ . Eine Implementierung dessen findet sich bei ATLAS im *FastVertexFitter*.

### Kalman-Fit-Methode

Alternativ zu einem Billoir-Fit können sogenannte *Kalman-Filter* verwendet werden. Solche sind ein viel verwendetes Werkzeug auch in der Spur- und Vertex-Rekonstruktion[20]. In ihrer ursprünglichen Anwendung dienen Kalman-Filter der Entwicklung von Zuständen eines linearen, dynamischen Systems. Ein Zustand  $\vec{x}$  wird durch eine lineare Transformation von einem diskreten Schritt  $k-1$  zum nächsten Schritt  $k$  über eine Entwicklungsmatrix  $F_{k-1}$  entwickelt:

$$\vec{x}_k = F_{k-1}\vec{x}_{k-1} + \vec{w}_{k-1} . \quad (4.4)$$

Dabei wird  $\vec{w}_{k-1}$  als statistischer Rauschanteil der Entwicklung beigefügt. Durch eine jeweils neue Messung  $\vec{m}_k$  wird das System bei jedem  $k$ -ten Schritt weiter eingeschränkt,

wobei für einen Kalman-Filter die Messung linear vom Zustandsvektor abhängen muss:

$$\vec{m}_k = H_k \vec{x}_k + \vec{\epsilon}_k, \quad (4.5)$$

wobei  $\vec{\epsilon}_k$  ein zur Messung assoziiertes *Gauß*-förmiges Rauschen ist. Im Falle der Rekonstruktion von Spuren entspricht die Hinzunahme eines weiteren Treffpunktes der Spur mit einem Detektorelement je einer neuen Messung. Jeder neue Punkt  $\vec{x}_k$  der Spur wird durch die Kenntnis aller vorherigen Messpunkte genauer bestimmbar. Die letzte Messung liefert für ein lineares System dann die optimale Lösung für  $\vec{x}_{k=N}$ . Um alle anderen Spurpunkte mit äquivalenter Genauigkeit bestimmen zu können wird unter Kenntnis aller Spurpunkte jeder Punkt erneut bestimmt. Im Falle des Vertex-Fits ist die Entwicklung deutlich einfacher. Eine neue Messung  $\vec{m}_k$  entspricht hier der Hinzunahme einer weiteren Spur zum Vertex-Fit und der Zustandsvektor  $\vec{x}$  entspricht für jeden Schritt  $k$  der Vertex-Position  $\vec{r}$  und den Impulsen  $\vec{p}_k$  der hinzugenommenen Spur,

$$\vec{x}_k = (\vec{r}, \vec{p}_k). \quad (4.6)$$

Die Entwicklung ist nun schlicht eine iterative Verbesserung der Vertex-Position und die Entwicklungsgleichung 4.4 wird auf

$$\vec{x}_k = \vec{x}_{k-1} \quad (4.7)$$

reduziert. Sind alle Spuren hinzugefügt, ist ein optimaler Vertex-Fit erreicht. Die Vertex-Position wird iterativ verbessert, nicht aber die Spurparameter. Alleine für die letzte hinzugefügte Spur wird zur Berechnung der neuen Parameter die aus dem Fit resultierende Vertex-Position verwendet. Für alle anderen Spuren müssen die Parameter nachträglich unter Berücksichtigung dieser Position erneut berechnet werden. Dieser Vorgang heißt *Smoothing*.

Der Kalman-Filter verlangt Linearität zwischen Messung und Zustand. Unter Verwendung von Gleichung 4.3 wird Gleichung 4.5 zu

$$\vec{m}_k = \vec{q}_0 + A[\vec{r} - \vec{r}_0] + B[\vec{p}_k - \vec{p}_0]. \quad (4.8)$$

Die Bestimmung der Vertex-Position bei jedem Schritt  $k$ , die Impulse der hinzugefügten Spur und die zugehörigen Kovarianzen lassen sich, wie in [20] gezeigt, aus den Zuständen und Kovarianzen der  $k-1$ -ten Messung berechnen. Das Durchlaufen eines ganzen Kalman-Filters, also iteratives Hinzufügen von allen Spuren ist äquivalent zu einer  $\chi^2$ -Minimierung. Zur Vereinfachung und zur Veranschaulichung wird in der Folge ein auf eine Dimension reduzierter Formalismus einer  $\chi^2$ -Minimierung verwendet. Es sei aber zu bedenken, dass die in der Folge beschriebenen Verfahren alle auf den Formalismus des Kalman-Filters zurückführbar sind.

Ein Kalman-Filter-Zyklus kann vereinfacht wie folgt dargestellt werden:

$$\chi^2 = \sum_{i=1}^N \left( \frac{\vec{q}_i(\vec{r}) - \vec{q}_{m,i}}{\sigma_i} \right)^2, \quad (4.9)$$

wobei die  $\vec{q}_{m,i}$  wie oben die gemessenen Spurparameter der  $i$ -ten Spur darstellen. Die  $\vec{q}_i(\vec{r})$  sind die zu bestimmenden Parameter der Spur, die durch den Vertex im Punkt  $\vec{r}$  laufen. Ein solcher Kalman-Filter findet im *sequentiellen* Vertex-Fitalgorithmus Anwendung.<sup>2</sup>

### Adaptive Vertex-Fit Methoden

In der vorhandenen schlichten Form ist der Kalman-Filter wie auch der Billoir-Fitter stark von einer zuvor stattfindenden Spur-zu-Vertex-Zuordnung abhängig. Eine falsch zugeordnete Spur wie auch eine sehr schlecht gemessene Spur lenkt die Bestimmung der Vertex-Position in eine falsche Richtung. Wenn für jede Spur zuvor eine Abschätzung gemacht werden kann, wie gut sie zum betrachteten Vertex passt, kann dies über Gewichte  $w(\chi^2)$  in der folgenden Form berücksichtigt werden:

$$\chi^2 = \sum_{i=1}^N w(\hat{\chi}_i^2) \left( \frac{\vec{q}_i(\vec{r}) - \vec{q}_{m,i}}{\sigma_i} \right)^2, \quad (4.10)$$

wobei  $\hat{\chi}_i^2$  die Kompatibilität der Spur mit dem Vertex darstellt. Es kann gezeigt werden, dass im Fall einer flachen Verteilung der nicht dem Vertex zugehörigen Spuren im Raum, die Gewichte optimal über eine Fermi-Funktion von den Kompatibilitäten abhängen:

$$w(\hat{\chi}_i^2) = \frac{1}{1 + \exp\left(\frac{\hat{\chi}_i^2 - \chi_{max}^2}{2T}\right)}. \quad (4.11)$$

Über die Fermi-Kante  $\chi_{max}^2$  lässt sich festlegen wie stark eine Spur gewichtet wird, die in einem bestimmten Abstand vom zu bestimmenden Vertex entfernt liegt. Der Kalman-Filter wird nun mehrmals durchlaufen, wobei zur Bestimmung der Gewichte schrittweise der Parameter  $T$  von 64 auf 1 reduziert wird. Hierüber wird berücksichtigt, dass die Unsicherheit der Vertex-Position in jeder Iteration kleiner wird und somit die  $\hat{\chi}_i^2$ -Werte genauer bestimmbar sind. Je besser die Vertex-Position bekannt ist, desto schwieriger wird es für eine von dieser Position abweichende Spur, berücksichtigt zu werden. Eine solche Prozedur nennt sich *Annealing*-Verfahren. Die  $\hat{\chi}_i^2$ -Werte werden in jedem Iterationsschritt alleine aus dem Abstand  $\Delta R$  zur derzeitigen Vertex-Position und aus der Unsicherheit der Spur  $\sigma(\Delta R)$  bestimmt:

$$\hat{\chi}_i^2 = \left( \frac{\Delta R}{\sigma(\Delta R)} \right)^2. \quad (4.12)$$

<sup>2</sup>In der ATLAS-Software *SequentialVertexFitter* genannt.

Ein solches Vorgehen wird *adaptive* Vertex-Fit genannt.<sup>3</sup> Im Prinzip ist man nun in der Lage, alle im Ereignis vorhandenen Spuren zum Fit zuzulassen und eine eindeutige Zuordnung zum Vertex erst nach dem letzten Schritt durch einen Schnitt auf die Gewichte durchzuführen.

Die Gewichtung der Spuren hängt in dieser Form alleine vom Abstand zum Vertex und der Unsicherheit der Spuren ab. In einer anspruchsvollen Umgebung, wie sie am LHC zu erwarten ist, werden sehr hohe *Pile-Up*-Multiplizitäten auftreten. In den meisten Fällen wird die Spur, die nicht zum derzeit betrachteten Vertex gehört von einem in der Nähe liegenden anderen Vertex stammen. Die Wahrscheinlichkeit, dass dies für eine Spur der Fall ist, hängt vom Abstand zum nächsten Vertex ab. Eine sogenannte *adaptive Multi-Vertex-Fit-Methode*<sup>4</sup> ist eine Erweiterung der oben beschriebenen *adaptiven* Vertex-Fit-Methode, die die Gewichte der Spuren unter Berücksichtigung der Wahrscheinlichkeit berechnet, dass sie auch zu einem anderen Vertex im Ereignis gehören können. In jeder Iteration der Bestimmung der Vertex-Position werden die Gewichte zum  $i$ -ten Vertex für alle  $N$  Spuren mit der Information aller  $M$  Vertex-Positionen wie folgt berechnet:

$$w(\hat{\chi}_i^2) = \frac{\exp(-\frac{\hat{\chi}_i^2}{2T})}{\sum_{j=1}^M \exp(-\frac{\hat{\chi}_j^2}{2T}) + \exp(-\frac{\chi_{max}^2}{2T})} . \quad (4.13)$$

In jedem *Annealing*-Schritt bestimmen  $M$  Kalman-Filter  $M$  neue Vertex-Positionen aus allen  $N$  Spuren. Die *adaptive Multi-Vertex-Fit-Methode* ist in der Lage, unter Berücksichtigung aller anderen Vertices im Ereignis Spuren zu einem bestimmten Vertex zuzuordnen. Dies stellt eine sehr flexible Methode der Spur-zu-Vertex-Zuordnung dar, die angepasst auf das jeweilige Ereignis Spuren zwischen den Vertex-Kandidaten übergeben kann. Im Vergleich hierzu sind Spuren aus dem Billoir- und sequentiellen Fit, sobald sie als unpassend verworfen worden sind für alle Vertices im Ereignis nicht mehr zugänglich. So anpassungsfähig das gleichzeitige Bestimmen mehrerer Vertices auch sein mag, man ist in dieser Form noch nicht in der Lage die Anzahl von Vertices  $M$  für den Vertex-Fit offen zu lassen. Die Anzahl an vorhandenen Wechselwirkungen steht vor der Bestimmung der Vertex-Positionen im Idealfall noch nicht fest und soll erst durch den Vertex-Fit bestimmt werden. Eine Lösung dieses Problems wird im nächsten Abschnitt bei den Verfahren zur Vertex-Suche vorgestellt.

#### 4.1.2. Algorithmen zur Vertex-Suche

Es sollen an dieser Stelle zwei unterschiedliche bei ATLAS implementierte Methoden zur Vertex-Suche erläutert werden. Wie zu Beginn dieses Kapitels bereits erwähnt kann der

<sup>3</sup>In der ATLAS-Software *AdaptiveVertexFitter* genannt.

<sup>4</sup>In der ATLAS-Software *AdaptiveMultiVertexFit* genannt.

Vertex-Fit von der Vertex-Suche unabhängig verwendet werden. Innerhalb der Vertex-Suche werden die Spuren zu Vertex-Kandidaten gebündelt. Alle Spuren eines jeden Vertex-Kandidaten werden dann von der aufgerufenen Vertex-Fit-Methode in einem gemeinsamen Kreuzungspunkt zusammengeführt. Eine diesem Vorgehen folgende Methode zur Vertex-Suche wird in der Folge *einfache* Vertex-Suche genannt.<sup>5</sup> Anschließend soll ein Algorithmus vorgestellt werden, der *adaptiv* nach mehreren Vertices gleichzeitig sucht. Ein solcher wird in der Folge der Einfachheit halber mit dem Namen *AdaptiveMultiVertexFinder* bezeichnet.<sup>6</sup> Die Vertex-Suche und der Vertex-Fit aller Primärvertices im Ereignis wird hier in einem Schritt ausgeführt. Für die *einfache* Vertex-Suche gilt es also, zuerst den Vertex zu suchen und dann die Position zu bestimmen, dem *AdaptiveMultiVertexFinder* liegt dagegen das Prinzip zugrunde, die Vertex-Suche durch die Vertex-Fits zu erreichen.

### Die *einfache* Vertex-Suche

Jede Vertex-Suche beginnt zunächst mit der Selektion von Spuren, die von der Wechselwirkungsregion ausgehen. Dies wird durch grobe Schnitte auf den transversalen Impact-Parameter  $d_0$  erreicht. Unter den selektierten Spuren gilt es dann Häufungspunkte im Raum zu finden. Im Falle der *einfachen* Vertex-Suche wird dies durch eine Projektion auf die Strahlachse vereinfacht. Eine solche Projektion ist gleichbedeutend mit der Verwendung des longitudinalen Impact-Parameters  $z_0$  als z-Position der Spur auf der Strahlachse. Die  $z_0$ -Werte aller Spuren werden in absteigender Ordnung aufgelistet. Ein Fenster der Breite 5 mm fährt die z-Achse ab. Erreicht das Fenster eine Position, die eine bestimmte Anzahl von Spuren beinhaltet, werden diese zu einem Vertex-Kandidat gebündelt. Diese Spurbündel werden standardmäßig mit einem Billoir-Fit in einem Punkt zusammengeführt. Prinzipiell ist es hier aufgrund der unabhängigen Implementierung der Fit- und Suchalgorithmen möglich, auch jeden anderen Fitalgorithmus zu verwenden. Diese ersten Vertices werden nun von inkompatiblen Spuren bereinigt indem alle Spuren, die einen  $\chi^2$ -Wert größer als einen gewissen Schwellenwert besitzen, verworfen werden. Ein neuer Fit wird mit den übrigen Spuren durchgeführt und wenn nötig wird der Vertex von inkompatiblen Spuren erneut bereinigt. Dies wird für alle Spurbündel ausgeführt und man erhält eine Liste von Vertex-Positionen mit zugehörigen Spuren. Diese Liste wird im Falle der *einfachen* Vertex-Suche anhand des Kriteriums

$$\sum_{i=1}^{N_{\text{Spuren}}} p_T \quad (4.14)$$

nach absteigender Ordnung sortiert. An erster Stelle steht nun ein ausgewählter Primärvertex, der mit einer vom obigen Kriterium und der Signalart abhängigen Wahrscheinlichkeit

<sup>5</sup>In der ATLAS-Software *InDetPriVxFinder* genannt.

<sup>6</sup>In der ATLAS-Software *InDetAdaptiveMultiPriVxFinder* genannt.

der Signalvertex ist. Ein Leistungstest unterschiedlicher Sortier- bzw. Identifikationskriterien wird ausführlich in Kapitel 5 behandelt. An dieser Stelle werden lediglich die in den Suchalgorithmen implementierten Selektionskriterien angegeben.

Die *einfache* Vertex-Suche ist ein simples, intuitives Vorgehen zur Primärvertex-Rekonstruktion. Sie bringt jedoch einige Nachteile mit sich:

- Die Anzahl der Vertizes liegt schon sehr früh fest, da diese schon aus der ersten Bündelung der Spuren erhalten wird.
- Die Vertex-Zugehörigkeit der Spuren wird ebenso im ersten Schritt festgelegt. Nachträglich können lediglich Spuren entfernt, nicht aber hinzugefügt werden. Ist eine Spur inkompatibel zu einem Vertex, wird sie verworfen und ist für die anderen Vertizes im Ereignis nicht mehr zugänglich.

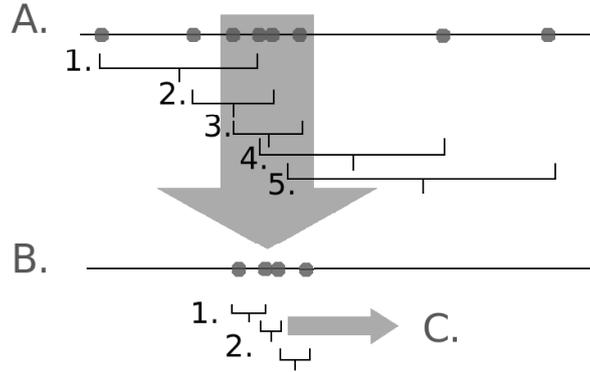
Dies wird vor allem in Szenarien mit hohen *Pile-Up*-Raten einen Einfluss haben, da hier die Wahrscheinlichkeit für nahe beieinander liegende Vertizes stark ansteigt (Abb. 3.5(b)). Die Zuordnung der Spuren zu einer Vertex-Position wird durch den Überlapp der Vertizes zu einer komplexen Aufgabe. Eine sehr pfiffige Lösung dieser Aufgabe ist der in der Folge vorgestellte *AdaptiveMultiVertexFinder*.

### **Der *AdaptiveMultiVertexFinder***

Anstatt die Anzahl der Vertizes und die zugehörigen Spuren im ersten Schritt festzulegen, beginnt der *AdaptiveMultiVertexFinder* mit einem einzelnen Vertex aus allen Spuren im Ereignis. Zunächst muss allerdings eine Position als Startpunkt gefunden werden. Diese Aufgabe wird unter dem Begriff *Seed-Finding* zusammengefasst. Der *Seed-Finder* sucht den Punkt im Raum mit der größten Spurdichte. Dies ist für gekrümmte Spuren im dreidimensionalen Raum eine anspruchsvolle Aufgabe. An dieser Stelle soll zuerst kurz das Prinzip des *Seed-Finders* skizziert werden. In der Folge wird dann die Systematik des *AdaptiveMultiVertexFinder* beschrieben.

Der *3DSeedFinder* berechnet zunächst paarweise den kleinsten Abstand  $d$  aller Spuren und bestimmt den Mittelpunkt zwischen den beiden Spuren. Nun werden die Punkte auf jede Raumachse projiziert und für jede dieser wird in folgender Weise ein Punkt der größten Häufung bestimmt. Das Verfahren ist in Abbildung 4.1 schematisch dargestellt.

1. Alle Intervalle die mehr als 50% aller Punkte enthalten, werden selektiert und das kleinste Intervall unter diesen wird weiter verwendet.
2. Die Prozedur wird erneut für die Punkte in diesem Intervall bei 1. begonnen und wird wiederholt bis nur drei oder weniger Punkte im Intervall übrig sind.
3. Der Mittelwert der übrigen Punkte wird als geschätzte Vertex-Position verwendet.



**Abb. 4.1.:** Schema zur Bestimmung eines *Vertex-Seeds*. A. entspricht der Suche nach dem kleinsten Intervall, welches 50% aller Punkte enthält. Im Schema wird das vierte Intervall ausgewählt und im nächsten Schritt (B.) wird unter den Punkten in diesem Intervall wieder nach dem kleinsten Intervall gesucht, welches 50% der Punkte enthält. Abbildung aus [21].

Jedem Punkt kann über den Abstand  $d$  zwischen den beiden Spuren ein Gewicht  $w = 1/d$  zugeordnet werden. Die Gewichtung kann durch Multiplikation der Intervalllänge mit dem mittleren Gewicht aller Spuren im Intervall berücksichtigt werden.

Durch die gute Strahlkollimierung in x- und y-Richtung ist es ausreichend einen Startpunkt in z-Richtung zu finden. Der *z-ScanSeedFinder* gibt als x- und y-Koordinate die Position der Wechselwirkungsregion zurück und berechnet die z-Position anhand des obigen Schemas.

Bevor der *Seed-Finder* vom *AdaptiveMultiVertexFinder* aufgerufen wird, werden die Spuren auf ihre Kompatibilität mit der Wechselwirkungsregion überprüft. Spuren, die aus diesem Bereich kommen, werden an den *z-ScanSeedFinder* übergeben und dieser bestimmt die erste Vertex-Position  $z_{seed}$ . Nun wird in folgender Weise fortgefahren:

1. Aus allen Spuren mit  $|z - z_{seed}| < 8 \text{ mm}$  wird mit einem *adaptiven* Multi-Vertex-Fit eine Vertex-Position bestimmt. Sollten zuvor schon Vertex-Fits stattgefunden haben, wodurch für einige der ausgewählten Spuren festgestellt werden konnte, dass diese gut zu einem anderen Vertex passen, wird dieser Vertex zum Fit hinzugenommen. Für alle  $M$  beteiligten Vertizes wird gleichzeitig ein *adaptiver* Multi-Vertex-Fit durchgeführt. In dieser Weise werden die Spuren nur dann einem Vertex zugeordnet, wenn sie nicht zu einem in der Nähe liegenden Vertex besser passen.
2. Spuren, die dem neu bestimmten Vertex zugeordnet werden können ( $\chi^2 < \chi^2_{Schnitt}$ ), werden aus der Liste der Spuren entfernt.
3. Eine neue Vertex-Position wird über den *z-ScanSeedFinder* unter Ausschluss der bereits zugeordneten Spuren gefunden und das Verfahren beginnt von vorne, wobei

wieder *alle* Spuren, inklusive der Spuren, die für den vorherigen Schritt entfernt wurden zum neuen Vertex-Fit verwendet werden.

In dieser Weise wird eine Liste von Vertizes erhalten. Wie auch im Falle der *einfachen* Vertex-Suche wird diese Liste sortiert, wobei hier das Kriterium

$$\sqrt{N_{\text{Spuren}}} \sum_{i=1}^{N_{\text{Spuren}}} p_{T,i}^2 \quad (4.15)$$

ist.

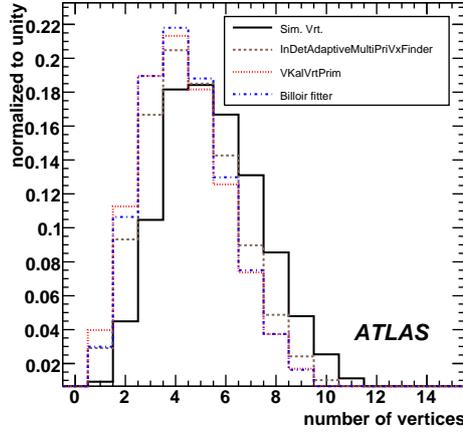
Neben dem *AdaptiveMultiVertexFinder* ist auch ein *AdaptiveVertexFinder* bei ATLAS zur Verfügung gestellt. Dieser verwendet den *AdaptiveVertexFitter* und es wird nur ein einzelner Vertex zurückgegeben. Für *Pile-Up*-Studien ist dieser Algorithmus aus diesem Grund nicht geeignet.

Der *AdaptiveMultiVertexFinder* ist eine sehr potentes Werkzeug zur Primärvertex-Rekonstruktion in Gegenwart von *Pile-Up*. Er ist das Standardwerkzeug in der ATLAS-Software. Für alle folgenden Studien wird ausschließlich diese Art der Primärvertex-Rekonstruktion verwendet.

## 4.2. Rekonstruktionsleistung der Algorithmen

Im Rahmen von Vertex-Studien wurde unter anderem der Einfluss von *Pile-Up* auf die Rekonstruktionseffizienz des Signalvertex untersucht. Zum Abschluss dieses Kapitels sollen die Ergebnisse dieser Studien kurz beschrieben werden.

Die Untersuchungen wurden auf  $WH$ -Ereignissen ( $WH(H \rightarrow b\bar{b})$ ), mit im Mittel 4.6 *Minimum-Bias*-Wechselwirkungen zusätzlich zum obigen Prozess durchgeführt. In Abbildung 4.2 ist die Anzahl der simulierten und die Anzahl der rekonstruierten Vertizes für verschiedene Rekonstruktionsalgorithmen zu sehen. Für alle zeigt sich, dass im Mittel ein bis zwei Vertizes im Ereignis nicht rekonstruiert werden. Es ist aus dieser Abbildung zunächst nicht ersichtlich, ob tendenziell die Signal- oder *Minimum-Bias*-Vertizes nicht rekonstruiert werden. Testet man selbiges auf Ereignissen ohne *Pile-Up* wird eine Rekonstruktionseffizienz von nahezu 100 % beobachtet. Die Rekonstruktion verliert demnach in erster Linie *Minimum-Bias*-Wechselwirkungen. In der betrachteten Studie wurde die Effizienz bestimmt mit der, der Signalvertex rekonstruiert und unter allen rekonstruierten Vertizes richtig selektiert wurde. Ein Vertex gilt als rekonstruiert und richtig selektiert, wenn für den Abstand zwischen der Position des simulierten Signalvertex  $z_{sim}$  und des rekonstruierten, selektierten Vertex  $z_{rek}$ ,  $|z_{rek} - z_{sim}| < 500 \mu\text{m}$  gilt. In Tabelle 4.1 ist diese Effizienz für verschiedene Algorithmen zur Vertex-Suche gezeigt. Aus den vorliegenden Studien



**Abb. 4.2.:** Simulierte (sim. Vrt) und rekonstruierte Vertex-Multiplizität für Ereignisse mit im Mittel 4.6 *Minimum-Bias*-Kollisionen. (*VKalVrtPrim* ist eine weitere Vertex-Such-Methode die hier nicht betrachtet wird.) Abbildung aus [9].

Vertex-Suche	AdaptiveMultiVertexFinder	einfache Vertex-Suche	
Vertex-Fit	AdaptiveMultiVertexFitter	FullFitter	FastFitter
Effizienz [%]	$93.7 \pm 0.1$	$89.8 \pm 0.2$	$89.8 \pm 0.2$

**Tab. 4.1.:** Effizienz der Bestimmung der Signalvertex-Position in Gegenwart von *Pile-Up* (im Mittel 4.6 *Minimum-Bias*-Vertices zusätzlich zum Signalvertex) für  $WH(H \rightarrow b\bar{b})$ -Ereignisse. Tabelle aus [9].

ist nicht ersichtlich, ob eine fehlgeschlagenen Bestimmung der Signalvertex-Position auf die Primärvertex-Rekonstruktion oder die Signalvertex-Selektion zurückzuführen ist. So ist die Effizienz der Bestimmung der Signalvertex-Position des *AdaptiveMultiVertexFinder* im Vergleich zur *einfachen* Vertex-Suche höher, wobei dies eine Folge der unterschiedlichen Selektionskriterien (vergleiche Gleichungen 4.14 und 4.15) sein kann.

Neben der Rekonstruktions- und Selektionseffizienz ist eine gute Vertex-Auflösung ein wichtiges Leistungskriterium. Die Auflösung ist eine Eigenschaft der verwendeten Vertex-Fit-Methoden. Die Gegenwart von zusätzlichen *Minimum-Bias*-Ereignissen kann einen Einfluss auf die Auflösung des Signalvertex der verschiedenen Fitalgorithmen haben. In den Tabellen 4.2 und 4.3 ist die Auflösung in longitudinaler und transversaler Richtung für verschiedene Fitalgorithmen mit und ohne *Pile-Up* gezeigt. Es sei festzuhalten, dass für sowohl longitudinale als auch transversale Raumrichtung der Fitalgorithmus des *AdaptiveMultiVertexFinder* eine bessere Auflösung erreicht als die Fitalgorithmen der *einfachen* Vertex-Suche. Für alle betrachteten Fitalgorithmen ist kein signifikanter Unter-

Finder	InDetAdaptiveMultiPriVxFinder	InDetPriVxFinder	
Fitter	AdaptiveMultiVertexFitter	FullFitter	FastFitter
nur Signal	$11.46 \pm 0.05$	$12.22 \pm 0.05$	$12.23 \pm 0.05$
Signal + <i>Pile-Up</i>	$11.66 \pm 0.05$	$12.51 \pm 0.05$	$12.50 \pm 0.05$

**Tab. 4.2.:** Vertex-Auflösung  $\sigma_x$  in [ $\mu\text{m}$ ] in transversaler Richtung für *Pile-Up*-Ereignisse (im Mittel 4.6 *Minimum-Bias*-Kollisionen) und Einzel-Kollisions-Ereignisse der Art  $WH \rightarrow b\bar{b}$ . Tabelle aus [9].

Finder	InDetAdaptiveMultiPriVxFinder	InDetPriVxFinder	
Fitter	AdaptiveMultiVertexFitter	FullFitter	FastFitter
nur Signal	$41.3 \pm 0.5$	$49.3 \pm 0.5$	$49.3 \pm 0.5$
Signal + <i>Pile-Up</i>	$40.3 \pm 0.5$	$48.6 \pm 0.5$	$48.6 \pm 0.5$

**Tab. 4.3.:** Vertex-Auflösung  $\sigma_z$  in [ $\mu\text{m}$ ] in longitudinaler Richtung für *Pile-Up*-Ereignisse (im Mittel 4.6 *Minimum-Bias*-Kollisionen) und Einzel-Kollisions-Ereignisse der Art  $WH \rightarrow b\bar{b}$ . Tabelle aus [9].

schied beim Vergleich der Leistung in Gegenwart von *Pile-Up* mit dem Fall ohne *Pile-Up* zu erkennen. Die angegebenen Werte sind die Standardabweichungen eines einfachen *Gauß*-Fits, der aus den Vertex-Fits erhaltenen Auflösungen.

Zusammenfassend ist festzuhalten, dass für niedrige Vertex-Multiplizitäten der Einfluss von *Pile-Up* auf die Vertex-Auflösung des Signalvertex klein ist, wobei eine Untersuchung für größere Vertex-Multiplizitäten unumgänglich ist, da der Überlapp mit der Anzahl deutlich zunimmt (siehe Abschnitt 3.2.2). Die transversale Auflösung eines  $WH(H \rightarrow b\bar{b})$ -Ereignisses liegt im Bereich von  $10 \mu\text{m}$ , die longitudinale Auflösung im Bereich von  $40 \mu\text{m}$ . Die Auflösung hängt von der jeweiligen Topologie der Signalwechselwirkung ab. Eine Studie über das Auflösungsvermögen der *Minimum-Bias*-Vertizes wurde an dieser Stelle noch nicht durchgeführt. Eine detailliertere Studie zur Primärvertex-Rekonstruktion wird derzeit verfasst und sollte in Kürze zugänglich sein. Hier finden sich auch ausführlichere Beschreibungen der Fit- und Such-Methoden [21].

Die Effizienz, mit der die verschiedenen Primärvertex-Rekonstruktionsalgorithmen die Signalvertex-Position bestimmen können, hängt einerseits von der Effizienz ab, mit der Primärvertizes rekonstruiert werden können, andererseits vom Selektionskriterium, anhand dessen unter allen rekonstruierten Vertizes der Signalvertex identifiziert wird. Im Rahmen der in diesem Kapitel betrachteten Studien wurden nur die Selektionskriterien auf ihre Leistungsfähigkeit hin getestet, die innerhalb der verschiedenen Rekonstruktions-

algorithmen implementiert sind. Die vorliegenden Studien können daher keine Aussage darüber machen wie leistungsfähig das Selektionskriterium unabhängig von der Methode der Primärvertex-Rekonstruktion ist. Im Folgenden wird in Kapitel 5 eine detaillierte Studie zur Selektion des Signalvertex vorgestellt, die unabhängig von der Primärvertex-Rekonstruktion bleibt. Die Signalvertex-Selektion kann von der Primärvertex-Rekonstruktion entkoppelt werden. Aus diesem Grund wurde ein Werkzeug eigens hierzu entwickelt und in die ATLAS-Software implementiert. Eine Beschreibung dessen schließt sich an die Selektionsstudien in Kapitel 6 an.

## 5. Signalvertex-Selektion in Gegenwart von *Pile-Up*

In diesem Kapitel wird eine Studie zur Signalvertex-Selektion unabhängig von der Primärvertex-Rekonstruktion vorgestellt. Dabei werden anhand eines einfachen Tests verschiedene Methoden auf ihre Selektionsleistung hin untersucht. Einige der untersuchten Methoden zur Vertex-Selektion zeichnen sich dabei durch große Schlichtheit aus, von anderen multivariaten Ansätzen ist eine sehr gute Selektionsleistung zu erwarten. Die Selektion des Signalvertex basiert dabei alleine auf der Information der Spuren an den rekonstruierten Vertices, die aus der Spur- und Primärvertex-Rekonstruktion erhalten werden.

### 5.1. Eigenschaften von *Minimum-Bias*- und Signalvertices

Drei Signalsorten werden im Folgenden zum Leistungstest der Signalvertex-Selektion verwendet. Dabei unterscheiden sich die Signalsorten in ihren Eigenschaften, sodass für die verschiedenen Signale sehr unterschiedliche Selektionsleistungen erwartet werden. Je nach Signalprozess ist die Selektion des Signalvertex unterschiedlich anspruchsvoll. Die Selektionseffizienz kann so für ein großes Spektrum von Prozessen abgeschätzt werden. Dies, mehr als die physikalische Relevanz in den Analysen, begründet die Wahl der verschiedenen Signale.

#### 5.1.1. Die verwendeten Signalprozesse

##### Top-Antitop-Ereignisse

Charakteristisch für Top-Antitop-Ereignisse, auch  $t\bar{t}$ -Ereignisse genannt, sind eine Vielzahl von Spuren und Spuren mit großem Impuls. Das top-Quark ist das schwerste aller uns bekannten Fermionen. Es zerfällt sehr schnell und fast ausschließlich in ein W-Boson und ein b-Quark ( $t \rightarrow Wb$ ). Sichtbar wird aus diesem Zerfall im Detektor ein vom b-Quark ausgehender Jet und entweder zwei weitere Jets im Fall eines hadronischen W-Boson-Zerfalls oder ein Lepton im Falle eines leptonischen W-Boson-Zerfalls. Jets enthalten im Allgemeinen viele Spuren. Da ein  $t\bar{t}$ -Prozess zwischen zwei bis sechs Jets besitzen kann, ist

hier eine große Anzahl von Spuren (hohe Spur-Multiplizität) zu erwarten. Durch die große Massendifferenz zwischen Top- und b-Quark ist von einigen Spuren ein sehr hoher Impuls zu erwarten. Dies ist vor allem in der transversalen Ebene aussagekräftig, da in Strahlrichtung alle Ereignisse Spuren mit hohen Impulsen aufweisen können.  $t\bar{t}$ -Ereignisse sind in ihrer Topologie den *Minimum-Bias*-Ereignissen sehr fremd, insofern wird die Selektion eines  $t\bar{t}$ -Vertex sehr einfach sein.

### ***WH*-Ereignisse**

Eine weitere Ereignisklasse, bei der man mit Spuren mit großem transversalen Impuls rechnet sind *WH*-Ereignisse. Die hier verwendeten Ereignisse wurden mit einer Higgs-Masse von  $120 \text{ GeV}/c^2$  simuliert und der Zerfall des Higgs-Bosons wurde auf den Zerfall in ein  $b\bar{b}$ -Paar eingeschränkt. Die beiden b-Quarks hinterlassen je einen Jets. Im vorliegenden Datensatz zerfällt das W-Boson leptonic ( $W \rightarrow \mu\nu_{\mu}$ ). Auch hier sind Spuren mit großen transversalen Impulsen und eine hohe Spur-Multiplizität zu erwarten, wobei durch die geringere Anzahl an Jets im Vergleich zu  $t\bar{t}$ -Ereignissen etwas weniger Spuren vorliegen werden. Die räumliche Struktur der *WH*-Ereignisse variiert je nachdem wie viel Impuls das Higgs- und das W-Boson tragen. Trägt das Higgs-Boson einen sehr kleinen transversalen Impuls, laufen die beiden Jets der b-Quarks annähernd entgegengesetzt auseinander, für einen größeren Impuls des Higgs-Bosons zeigen die Jets tendenziell in die gleiche Richtung. Auch von einem *WH*-Vertex wird erwartet, dass er einfach zu selektieren ist.

### **$H \rightarrow \gamma\gamma$ -Ereignisse**

$H \rightarrow \gamma\gamma$ -Ereignisse haben Eigenschaften, die eine Signalvertex-Selektion stark erschweren. Die beiden aus dem Higgs-Zerfall entstehenden Photonen bleiben für den Spurdetektor unsichtbar. Damit geht die Information der Teilchen mit dem erwartungsgemäß größten transversalen Impuls für die Signalvertex-Selektion verloren. Die Spuren an einem  $H \rightarrow \gamma\gamma$ -Ereignis stammen alleine aus den Wechselwirkungen der Partonen, die nicht an der Higgs-Boson-Produktion beteiligt waren. Dieser Anteil des Ereignisses wird *unterliegendes* Ereignis (engl. *underlying event*) genannt. Jede Proton-Proton-Wechselwirkung wird einen solchen Ereignisanteil besitzen, wobei das unterliegende Ereignis abhängig vom Produktionsprozess des Signals unterschiedliche Eigenschaften aufweisen kann. Die Spuren des unterliegenden Ereignisses kommen aus dem gleichen Wechselwirkungspunkt und werden so dem jeweiligen Vertex zugeordnet. Die Prozesse des unterliegenden Ereignisses zeichnen sich durch einen geringen Impulsübertrag aus. Es ist daher zu erwarten, dass die Eigenschaften des unterliegenden Ereignisses und damit auch eines  $H \rightarrow \gamma\gamma$ -Ereignisses den Eigenschaften einer *Minimum-Bias*-Wechselwirkung ähnlich sind. Die Selektion eines

$H \rightarrow \gamma\gamma$ -Vertex ist daher sehr anspruchsvoll.

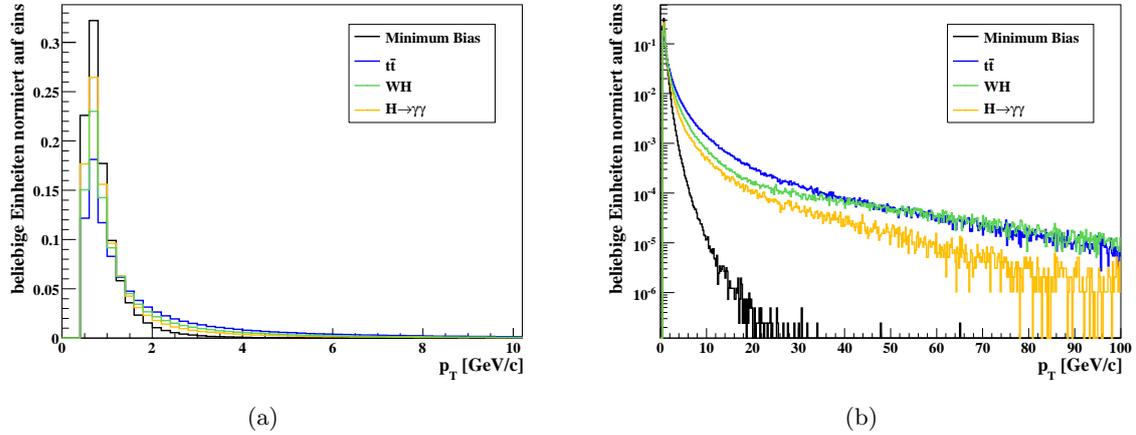
### 5.1.2. Eigenschaften der Spuren am Vertex

Eine Signalereignisse auszeichnende Eigenschaft ist, wie oben beschrieben, die Existenz von Spuren mit großem transversalen Impuls. Die Ursache dessen, ist dass im Allgemeinen für Signalereignisse ein größerer Impulsübertrag notwendig ist. Ein großer Impulsübertrag hat Streuteilchen mit großen Impulsen zur Folge. Für *Minimum-Bias*-Prozesse dagegen, ist ein beliebiger Impulsübertrag möglich. Aus diesem Grund erwartet man von diesen Ereignissen nur sehr selten eine Spur mit sehr großem transversalen Impuls. Die Abbildungen 5.1(a) und 5.1(b) zeigen die Spektren der transversalen Impulse der Spuren am Vertex für die drei verschiedenen Signal- und *Minimum-Bias*-Ereignisse. Es werden hier nur Spuren mit  $p_T > 500 \text{ MeV}/c$  betrachtet<sup>1</sup>. In Abbildung 5.1(a) sieht man die Spektren für einen Bereich kleiner transversaler Impulse bis  $10 \text{ GeV}/c$ . Das Verhalten der Spektren aller Ereignisarten ist sehr ähnlich. Die Verteilungen verdichten sich bei kleinen Werten des transversalen Impulses und fallen zu großen Werten hin ab. Für *Minimum-Bias*-Ereignisse ist dieser Abfall am stärksten, gefolgt von  $H \rightarrow \gamma\gamma$ -Ereignisse und  $WH$ -Ereignissen. Die Verteilung der  $t\bar{t}$ -Ereignisse fällt am wenigsten steil ab und besitzt lange Ausläufer zu sehr großen transversalen Impulsen hin. Dies wird noch deutlicher bei Betrachtung der Verteilung in einem größeren Bereich ( $p_T < 100 \text{ GeV}/c$  Abbildung 5.1(b)). Hier sind die Ausläufer der Verteilungen gut zu erkennen. Es sei darauf hingewiesen, dass ab einem transversalen Impuls von  $p_T > 50 \text{ GeV}/c$  die Ausläufer der  $WH$ -Ereignisse und der  $t\bar{t}$ -Ereignisse ein sehr ähnliches Verhalten aufweisen.

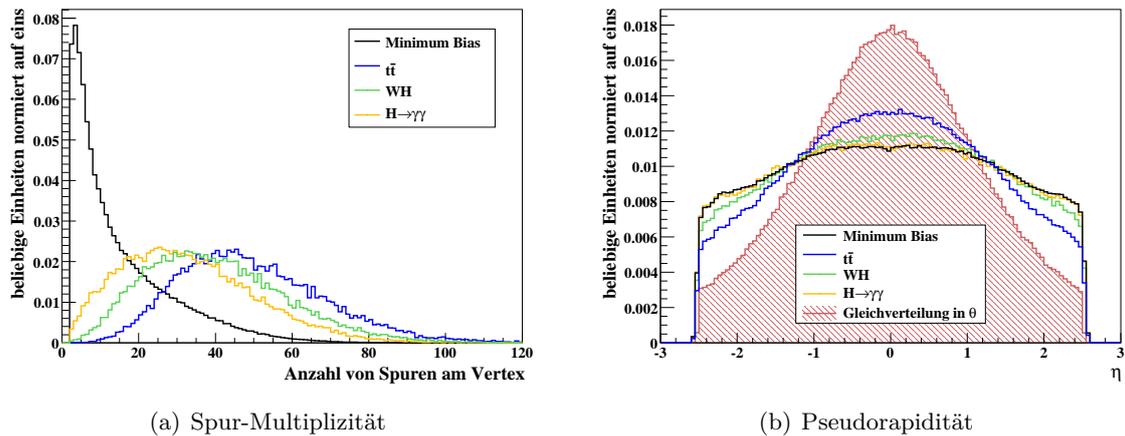
Eine weitere trennende Eigenschaft ist die Anzahl der Spuren am Vertex oder auch Spur-Multiplizität. Die Spektren der Spur-Multiplizität sind in Abbildung 5.2(a) gezeigt. Wie erwartet haben die  $t\bar{t}$ -Ereignisse die höchste Spur-Multiplizität, gefolgt von  $WH$ -Ereignissen. Die Verteilung der  $H \rightarrow \gamma\gamma$ -Ereignisse besitzen den größten Überlapp mit der Verteilung der *Minimum-Bias*-Ereignisse.

Neben dem transversalen Impuls gibt auch die Verteilung der Pseudorapidity Auskunft über die Verteilung der Spuren im Raum, bzw. darüber wie sehr das Ereignis in Vorwärts- und Rückwärtsrichtung des Detektors ausgerichtet ist. In Abbildung 5.2(b) sind die Pseudorapidity-Spektren zu sehen. Um diese besser vergleichen zu können ist schattiert eine im Polarwinkel  $\theta$  gleichverteilte Variable eingezeichnet. Ist die Verteilung in der Pseudorapidity flach verteilt, bedeutet dies starke Aktivität in Vorwärts- und Rückwärts-

<sup>1</sup>Die Spur-Rekonstruktion von ATLAS rekonstruiert bei Standardeinstellungen nur Spuren mit einem transversalen Impuls  $> 500 \text{ MeV}/c$ .



**Abb. 5.1.:** Spektren der transversalen Impulse der verschiedenen Signal- und *Minimum-Bias*-Ereignisse im Vergleich. (a)  $p_T < 10$  GeV/c, (b)  $p_T < 100$  GeV/c in einfachlogarithmischer Darstellung.



**Abb. 5.2.:** Verteilung der Spur-Multiplizität (a) und der Pseudorapidity  $\eta$  (b) verschiedener Signal- und *Minimum-Bias*-Ereignisse im Vergleich. In (b) ist deutlich die begrenzte Abdeckung des inneren Detektors bis  $|\eta| < 2.5$  zu erkennen (Vgl. Abschnitt 2.2.2).

richtung des Detektors. Die *Minimum-Bias*-Verteilung zeigt dieses Verhalten am ausgeprägtesten, gefolgt von der Verteilung der  $H \rightarrow \gamma\gamma$ -, der  $WH$ - und der  $t\bar{t}$ -Ereignisse. Prinzipiell ist eine Trennung anhand der Verteilung der Pseudorapiditäten möglich. Allerdings ist diese Größe stark mit dem transversalen Impuls korreliert. Aus diesem Grund wird die Pseudorapidität im Folgenden nicht weiter berücksichtigt.

Ereignisse können auch dadurch charakterisiert werden, wie sehr sich die Spuren der Ereignisse in alle Raumrichtungen verteilen. Ein Maß für diese Charaktereigenschaft ist die Sphärizität  $\Psi$ . Man betrachte die Durchstoßpunkte aller Spuren mit einer Kugeloberfläche um den Vertex. Die Sphärizität gibt dann an, wie stark isotrop die Durchstoßpunkte auf dieser Oberfläche verteilt sind. Je mehr ein Ereignis eine gemeinsame Vorwärts- und Rückwärtsrichtung aller Spuren aufweist, desto weniger sphärisch wird das Ereignis sein. Ereignisse mit nur zwei Spuren oder Jets die in entgegengesetzte Richtung auseinander laufen haben eine sehr kleine Sphärizität. Die Sphärizität ist wie folgt definiert: Aus der Matrix

$$S^{\alpha\beta} = \frac{\sum_i p_i^\alpha p_i^\beta}{\sum_i |p_i|^2}, \quad (5.1)$$

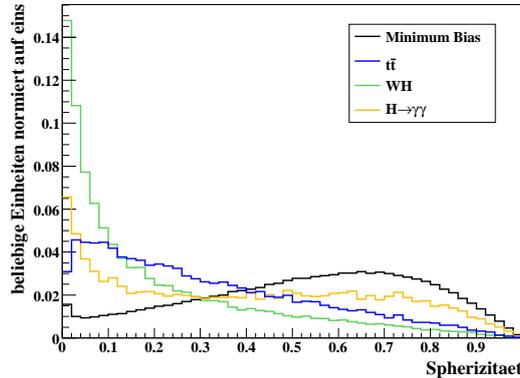
wobei  $p_i$  der Impuls der  $i$ -ten Spur ist und  $\alpha, \beta = (x, y, z)$  den Indizes der drei Raumrichtungen entsprechen, bestimme man die Eigenwerte und sortiere sie nach ihrer Größe  $\lambda_1 > \lambda_2 > \lambda_3$ . Die räumliche Sphärizität  $\Psi_{3D}$  ist dann:

$$\Psi_{3D} = \frac{3}{2}(\lambda_2 + \lambda_3). \quad (5.2)$$

Jedes Ereignis besitzt durch die Existenz des unterliegenden Ereignisanteils Spuren in Vorwärts- und Rückwärts-Richtung des Detektors. Die Sphärizität ist aus diesem Grund in einer Projektion in die transversale Ebene aussagekräftiger. Die Raumindizes laufen dann nur noch über  $x$  und  $y$  und die Sphärizität berechnet sich einfach aus dem kleineren der beiden Eigenwerte ( $\lambda_1 > \lambda_2$ )

$$\Psi_{2D} = 2 \cdot \lambda_2. \quad (5.3)$$

In der Folge bezieht sich die Sphärizität immer auf diese Projektion in die transversale Ebene ( $\Psi = \Psi_{2D}$ ). In Abbildung 5.3 ist die Verteilung der Sphärizität für Signal- und *Minimum-Bias*-Ereignisse zu sehen. *Minimum-Bias*-Ereignisse sind tendenziell sphärischer als Signalereignisse. Die Verteilung der  $H \rightarrow \gamma\gamma$ -Ereignisse besitzt einen sphärischen und einen nicht sphärischen Anteil. Auch *Minimum-Bias*-Ereignisse weisen eine Häufung von Vertices mit einer Sphärizität nahe bei null auf. Dies sind Ereignisse mit nur zwei oder sehr wenig Spuren.  $t\bar{t}$ - und  $WH$ -Ereignisse weisen tendenziell weniger stark sphärische Ereignisse auf. Im niederen Sphärizitätsbereich unterscheiden sich die Verteilungen der beiden Signalarten.  $WH$ -Ereignisse sind aufgrund ihrer Topologie Zwei-Jet-artiger als  $t\bar{t}$ -Ereignisse, da hier auch eine höhere Anzahl an Jets auftreten kann. Die Sphärizität kann



**Abb. 5.3.:** Verteilung der Sphärizität (Projektion in die transversale Ebene) der verschiedenen Signal- und *Minimum-Bias*-Ereignisse.

zusätzlich zu den transversalen Impulsen und der Spur-Multiplizität zur Selektion des Signalvertex verwendet werden. Die Trennleistung dieser Größe ist allerdings klein.

In der Folge werden zur Vertex-Selektion folgende Spur- und Vertex-Eigenschaften untersucht: Spur-Multiplizität, Sphärizität und die transversalen Impulse aller Spuren am Vertex. Jede einzelne Spur besitzt über die Größe ihres transversalen Impulses eine Information darüber, wie wahrscheinlich es ist, dass sie von einem Signal- oder *Minimum-Bias*-Vertex abstammt. Um eine Entscheidung treffen zu können, ob ein Vertex ein Signal- oder *Minimum-Bias*-Vertex ist, sollen die Informationen der einzelnen Spuren zu einer Variablen kombiniert werden. Im Folgenden werden verschiedene Kombinationsmöglichkeiten verglichen und getestet. Dabei geht es neben der Kombination der transversalen Impulse der Spuren auch darum, die Information der Spur-Multiplizität und der Sphärizität mit der Information der Impulse kombinieren zu können.

## 5.2. Methoden zur Selektion des Signalvertex

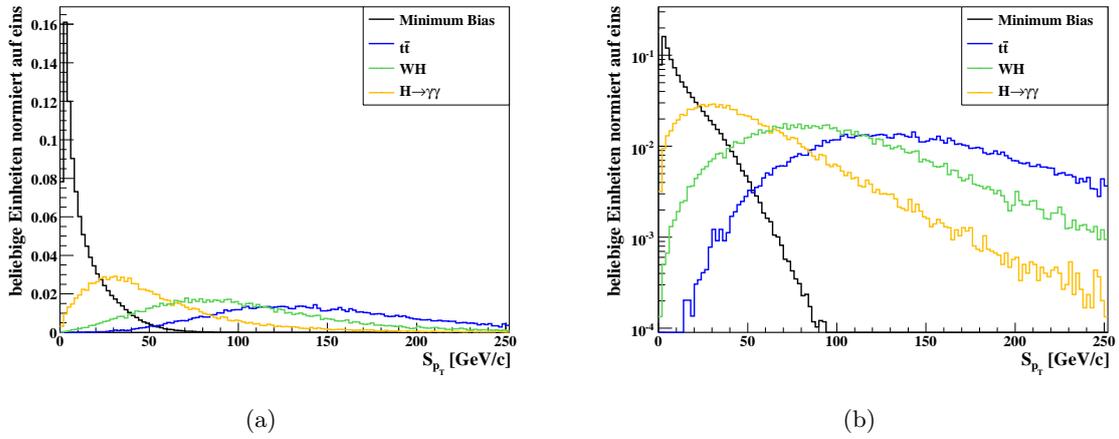
Bevor die verschiedenen Methoden beschrieben und verglichen werden, soll an dieser Stelle noch einmal kurz die Aufgabenstellung der Vertex-Selektion zusammengefasst dargestellt werden. Bei jeder Strahlkreuzung werden mehrere Protonen-Paare inelastisch wechselwirken. Bei gegebener Trigger-Einstellung werden nur Strahlkreuzungen aufgezeichnet, bei denen Hinweise auf einen Signalprozess vorliegen. Zusätzlich zum Signalereignis werden auch die in derselben Strahlkreuzung stattfindenden *Minimum-Bias*-Wechselwirkungen rekonstruiert. Dem Signalprozess und den *Minimum-Bias*-Prozessen werden durch die Vertex-Rekonstruktion Sätze von Spuren und z-Positionen zugeordnet. Man steht bei der Selektion demnach vor folgendem Problem: Jedes Ereignis besitzt eine bestimmte Anzahl

von rekonstruierten Vertizes  $N_{Vtx}$  wovon einer, der zu identifizierende Signalvertex ist. Die Verteilungen der Eigenschaften der Vertizes und der Spuren am Vertex, anhand derer man den Signalvertex identifiziert, können aus Simulationen abgeschätzt werden.

Es ist von Interesse das Verhalten der Selektionsleistung verschiedener Methoden mit steigender Vertex-Multiplizität beobachten zu können. Um dies detailliert und für beliebige Vertex-Multiplizitäten zu studieren, wurde an dieser Stelle vorerst darauf verzichtet, die Leistung an Ereignissen zu testen, die schon vor der Rekonstruktion mit *Pile-Up* überlagert wurden. Effekte auf Grund des Überlapps der Vertizes werden so bei der Rekonstruktion berücksichtigt. Studien auf solchen Ereignissen sind Inhalt von Kapitel 6. Innerhalb der Studien in diesem Kapitel wurde das *Pile-Up* von Hand nach der Primärvertex-Rekonstruktion nach folgendem Rezept zum Signal beigemischt:

1. Man nehme Signal- und *Minimum-Bias*-Ereignisse, die *ohne Pile-Up* simuliert wurden.
2. Für jedes Signalereignis ziehe man die gewünschte Anzahl an *Minimum-Bias*-Ereignissen ( $N_{MB} = N_{Vtx} - 1$ ) und berechne für alle Vertizes die Variable der Methode, die getestet werden soll.
3. Da von vornherein bekannt ist welcher Vertex der Signal-Vertex ist, muss nur festgestellt werden ob er von der jeweiligen Methode erkannt worden ist.
4. Man berechne zuletzt die Fehlidentifikationsrate, also den Anteil aller getesteten Ereignisse, bei dem die Methode den Signalvertex nicht korrekt identifiziert hat.

Anhand dieses Vorgehens ist es möglich, für verschiedene Vertex-Multiplizitäten das Verhalten zwischen den zu vergleichenden Methoden zu beobachten. Allerdings wird hier der Überlapp zwischen den Streuprodukten der verschiedenen Wechselwirkungen bei der Rekonstruktion vernachlässigt. Vor allem für große Vertex-Multiplizitäten ist nach den Studien aus 3.2.2 zu erwarten, dass der Abstand zwischen den Vertizes teilweise kleiner als das Auflösungsvermögen der Primärvertex-Rekonstruktion (siehe Abschnitt 4.2) wird. Eng beieinander liegende Vertizes können in solch einem Fall als ein gemeinsamer Vertex rekonstruiert werden. Sowohl die Rekonstruktion zweier *Minimum-Bias*-Wechselwirkungen in einem Vertex, als auch der Überlapp eines Signal- mit einem *Minimum-Bias*-Vertex wird Auswirkungen auf die Signalidentifikationseffizienz haben. Auch eine Fehlzuordnung einzelner Spuren nah beieinander liegender Vertizes kann einen Einfluss auf die Selektionsleistung nehmen. Trotz Kenntnis dieser Einflüsse ist es im Rahmen dieser Studien legitim wie oben beschrieben vorzugehen. Hier werden in erster Linie Vergleiche zwischen unterschiedlichen Selektionsmethoden betrachtet. Die verschiedenen Methoden basieren dabei auf den gleichen Spur- und Vertex-Eigenschaften. Vernachlässigung des Überlapps mag insofern einen Einfluss auf die absolute Selektionseffizienz haben, nicht aber auf die



**Abb. 5.4.:** Summe aller transversalen Impulse  $\mathcal{S}_{p_T}$  in (a) linearer Skala und (b) einfachlogarithmischer Skala.

Unterschiede zwischen den Methoden. Das oben beschriebene Vorgehen hat den Vorteil, dass beliebige Wechselwirkungsmultiplizitäten untersucht werden können, ohne dabei von vorhandenen *Pile-Up*-Datensätzen und deren Statistik abhängig zu sein.

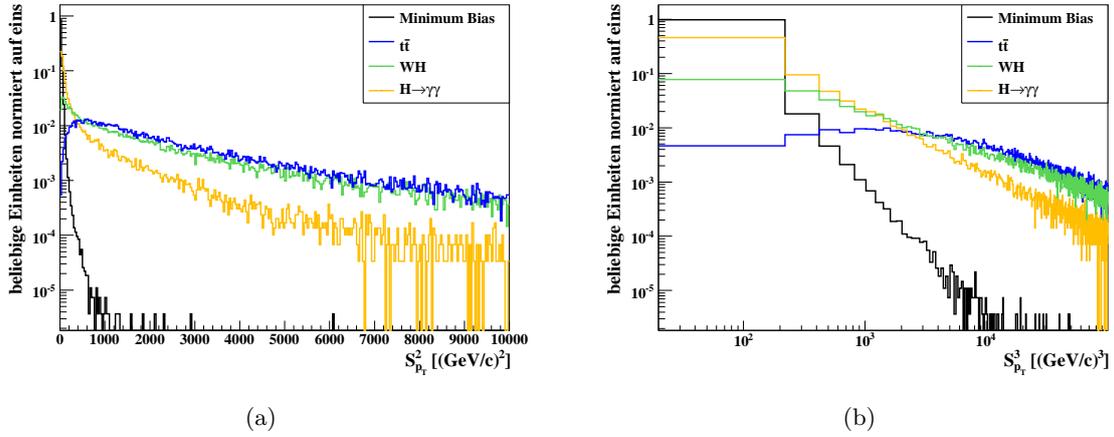
### 5.2.1. Summe der transversalen Impulse der Spuren am Vertex

Eine sehr einfache und intuitive Methode, die transversalen Impulse aller Spuren und die Spur-Multiplizität eines Vertex zu kombinieren, ist die Summe der transversalen Impulse aller Spuren ( $N_{Sp}$ ) zu bilden:<sup>2</sup>

$$\mathcal{S}_{p_T} = \sum_{N_{Sp}} p_T . \quad (5.4)$$

Abbildung 5.4 zeigt die Verteilung dieser Variablen in linearer (5.4(a)) und einfachlogarithmischer (5.4(b)) Skala für die drei Signalsorten und für *Minimum-Bias*-Ereignisse. Während die *Minimum-Bias*-Ereignisse hauptsächlich kleine  $\mathcal{S}_{p_T}$ -Werte liefern, streuen die Werte für die Signalereignisse um höhere Werte. Es ist an diesen Abbildungen schon sehr gut zu erkennen, dass es sehr einfach sein wird, ein  $t\bar{t}$ -Ereignis zu identifizieren, während es für  $H \rightarrow \gamma\gamma$ -Ereignisse einen deutlichen Überlapp mit den *Minimum-Bias*-Ereignissen gibt. Die Vertex-Identifikation benötigt keinen Schnitt auf die trennende Variable. Im vorliegenden Fall wird der Vertex mit dem größten  $\mathcal{S}_{p_T}$ -Wert im Ereignis als Signalvertex ausgewählt.

<sup>2</sup>Dieses Kriterium wird von der *einfachen* Vertex-Suche zur Selektion des Signalvertex verwendet (Abschnitt 4.1.2).



**Abb. 5.5.:** Summe der (a) quadrierten transversalen Impulse  $\mathcal{S}_{p_T}^2$  und (b) kubischen transversalen Impulse  $\mathcal{S}_{p_T}^3$ .

Um die im Falle von Signalprozessen häufiger vorkommenden Spuren mit großen transversalen Impulsen stärker hervorzuheben, kann an Stelle der einfachen Summe, die Summe über die quadrierten und kubischen transversalen Impulse berechnet werden. So erhalten die Spuren mit großen transversalen Impulsen ein stärkeres Gewicht im Selektionskriterium:

$$\mathcal{S}_{p_T}^2 = \sum_{N_{Sp}} p_T^2, \quad (5.5)$$

$$\mathcal{S}_{p_T}^3 = \sum_{N_{Sp}} p_T^3. \quad (5.6)$$

Die Verteilungen für  $\mathcal{S}_{p_T}^2$  und  $\mathcal{S}_{p_T}^3$  sind in den Abbildungen 5.5(a) und 5.5(b) zu sehen.

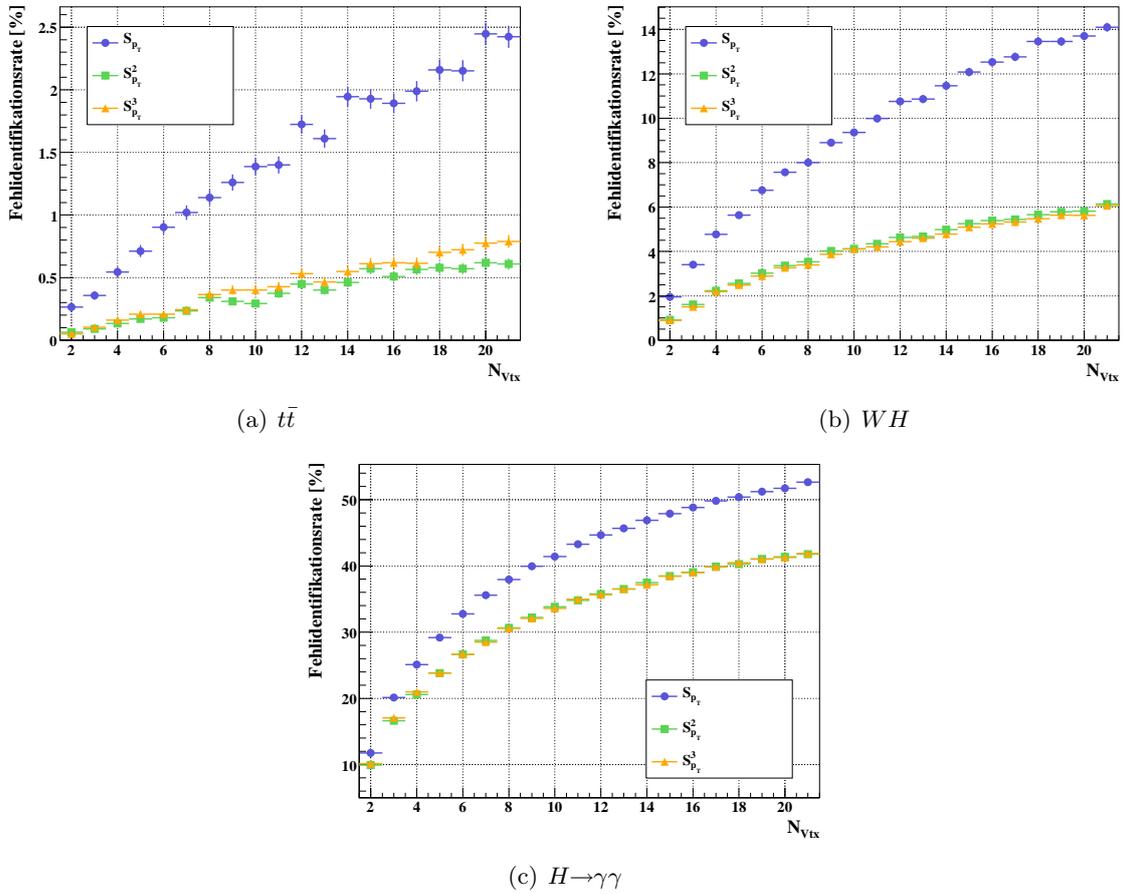
Die verschiedenen Kriterien können nun anhand des oben beschriebenen Rezepts auf ihre Selektionsleistung hin überprüft und verglichen werden. Es wurden hierfür insgesamt 30 000 Signalereignisse und 300 000 *Minimum-Bias*-Ereignisse zum Beimischen verwendet. Für Vertex-Multiplizitäten von  $N_{Vtx} = 2, \dots, 21$  (Anzahl *aller* Vertizes im Ereignis  $N_{MB} + 1$ ) wurden für die drei Signalarten die Fehlidentifikationsraten bestimmt und aufgetragen. Die Fehlidentifikationsrate  $\epsilon_{FI}$  ist dabei wie folgt definiert:

$$\epsilon_{FI} = \frac{\text{Anzahl aller Ereignisse, bei denen der Signalvertex nicht selektiert wurde}}{\text{Anzahl aller getesteten Ereignisse}}. \quad (5.7)$$

Der Fehlidentifikationsrate liegt eine Binomial-Verteilung zugrunde, daher ergibt sich ein Fehler von:

$$\sigma_{\epsilon_{FI}} = \sqrt{\frac{\epsilon_{FI} \cdot (1 - \epsilon_{FI})}{N_{test}}}. \quad (5.8)$$

Für alle Signalarten ist zu sehen, dass die Fehlidentifikationsrate mit steigender Vertex-Multiplizität, wie zu erwarten, ansteigt. Die absoluten Raten für die verschiedenen Si-



**Abb. 5.6.:** Fehlidentifikationsrate in % in Abhängigkeit der Vertex-Multiplizität ( $N_{Vtx}$  ist die Anzahl aller Vertices inklusive des Signalvertex) verschiedener summenartiger Selektionsmethoden. Verglichen werden folgende Selektionskriterien: Summen der einfachen ( $\mathcal{S}_{pT}$ ), quadratischen ( $\mathcal{S}_{pT}^2$ ) und kubischen ( $\mathcal{S}_{pT}^3$ ) transversalen Impulse für (a)  $t\bar{t}$ -, (b)  $WH$ - und (c)  $H \rightarrow \gamma\gamma$ -Ereignisse.

gnalsorten unterscheiden sich deutlich. So bewegt sich die Fehlidentifikationsrate für  $t\bar{t}$ -Ereignisse im Rahmen von 1-2%, für  $WH$ -Ereignisse bis hin zu 14% und für  $H \rightarrow \gamma\gamma$ -Ereignisse bis zu über 50%.

Vergleicht man für alle Signalsorten die Leistung der verschiedenen Summen der transversalen Impulse, ist eine deutliche Leistungssteigerung zwischen  $\mathcal{S}_{pT}$  und  $\mathcal{S}_{pT}^2$  für alle Signalsorten zu sehen. Der Vergleich von  $\mathcal{S}_{pT}^2$  und  $\mathcal{S}_{pT}^3$  zeigt keinen signifikanten Unterschied für  $H \rightarrow \gamma\gamma$ - und  $WH$ -Ereignisse. Im Falle von  $t\bar{t}$ -Ereignissen zeigt sich hier sogar ein Abfall in der Leistung durch Verwendung der dritten Potenz.  $\mathcal{S}_{pT}^2$ , die leistungsstärkste unter allen bisher betrachteten Methoden, wird in der Folge weiter berücksichtigt.

Die vorgestellten Methoden zeichnen sich durch ihre Einfachheit aus. Man erwartet

von solchen, dass sie robust sind gegen eventuell auftretende Abweichungen zwischen der Realität und der Simulation anhand der die Methoden getestet wurden. Aus diesem Grund ist es möglich Methoden dieser Art bei Datennahme unmittelbar anzuwenden. Es sind Methoden denkbar, die multivariate Ansätze verwenden. Von solchen erhofft man sich eine bessere Selektionsleistung. Im Vergleich zu den bisher beschriebenen Methoden basieren solche aber auf den Informationen aus den Verteilungen bestimmter Variablen, die vor Datennahme zunächst auf Monte-Carlo-Simulation beruhen. Man geht davon aus, dass Abweichungen zwischen Simulation und Realität einen größeren Einfluss auf die Leistung solcher Methoden haben als auf Methoden der Art von  $\mathcal{S}_{p_T}^2$ . In der Folge sollen einige multivariate Ansätze zur Vertex-Selektion beschrieben und getestet werden.

### 5.2.2. Vertex-Wahrscheinlichkeit

Anhand simulierter Datensätzen einer bestimmten Ereignisklasse kann eine Vorhersage gemacht werden, wie sich bestimmte Variablen dieser Ereignisklasse verteilen. Die Information, die aus dem Verlauf der Verteilungen erhalten wird, kann zur Trennung von einer anderen Ereignisklasse genutzt werden. Wie zuvor erwähnt basieren diese Verteilungen zunächst auf Monte-Carlo-Simulationen, werden aber, sobald Daten vorhanden sind, validiert oder durch diese ersetzt. Im Falle von *Minimum-Bias*-Ereignissen wird dies schon in sehr frühem Stadium möglich sein, da der zugrunde liegende totale inelastische Wirkungsquerschnitt groß ist (siehe Abbildung 3.2) und schon nach sehr kurzer Zeit ausreichend Daten gesammelt sein werden. Hier soll eine Methode vorgestellt werden, die eine Signalvertex-Selektion ermöglicht, alleine unter Verwendung der Verteilungen von *Minimum-Bias*-Ereignissen, nicht aber der Signalverteilungen. Diese Methode wurde zuerst im Rahmen des ALEPH-Experiments<sup>3</sup> zur b-Jet-Identifikation verwendet [22]. Heute findet diese Methode auch bei ATLAS zur b-Jet-Identifikation Anwendung. Sie ist unter dem Namen JETPROB in der ATLAS-Software implementiert.

Die hier vorgestellte Methode basiert auf der Grundidee, dass aus den Eigenschaften der Vertices eine Kompatibilität mit der *Minimum-Bias*-Hypothese angegeben werden kann. So kann beispielsweise anhand der Verteilung der transversalen Impulse von *Minimum-Bias*-Spuren abgeschätzt werden, inwieweit eine Spur, mit einem gemessenen transversalen Impuls  $p_T$ , kompatibel damit ist, von einem *Minimum-Bias*-Vertex abzustammen. Hierfür wird die Verteilung der transversalen Impulse der Spuren der *Minimum-Bias*-Ereignisse auf eins normiert. Dies kann als Schätzer für die Wahrscheinlichkeitsdichtefunktion  $\mathcal{F}_{MB}(p_T)$  verwendet werden. Das Integral über ein kleines Impulsintervall von  $\mathcal{F}_{MB}(p_T)$  gibt dann an, wie wahrscheinlich es ist, dass eine *Minimum-Bias*-Spur einen transversalen Impuls aus diesem Impulsintervall trägt. Daraus lässt sich eine Kompatibilität  $\mathcal{P}_{Spur}$  definieren,

<sup>3</sup>Apparatus for LEP Physics, Detektor bei LEP.

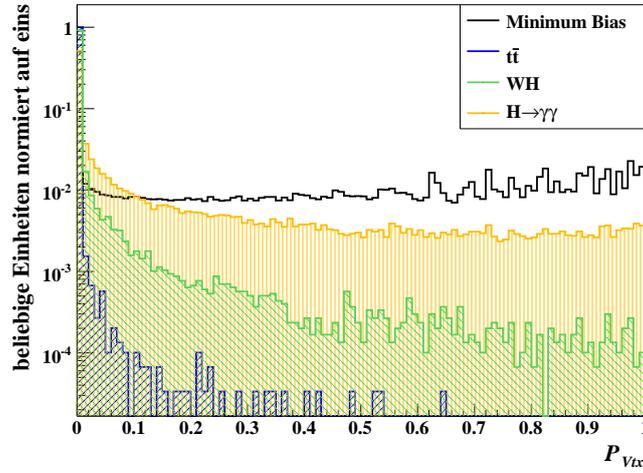


Abb. 5.7.: Vertex-Wahrscheinlichkeit  $\mathcal{P}_{Vtx}$  für *Minimum-Bias*- und Signalereignisse.

dass eine Spur bei gegebenem transversalen Impuls von einem *Minimum-Bias*-Vertex abstammt:

$$\mathcal{P}_{Spur}(p_T) = \int_{p_T}^{\infty} \mathcal{F}_{MB}(p'_T) dp'_T. \quad (5.9)$$

Die Spur-Kompatibilität ist für *Minimum-Bias*-Ereignisse flach verteilt während Signalereignisse tendenziell bei kleinen Werten liegen.

Es gilt nun die Spur-Kompatibilitäten aller Spuren am Vertex zu einer Variablen zu kombinieren, die eine Kompatibilität  $\mathcal{P}_{Vtx}$  dafür angibt, dass ein Vertex mit einem gegebenen Satz von transversalen Impulsen  $\vec{p}_T = (p_{T,1}, p_{T,2}, \dots, p_{T,N_{Sp}})$  ein *Minimum-Bias*-Vertex ist:

$$\mathcal{P}_{Vtx}(\vec{p}_T) = \mathcal{P}_0 \sum_{j=0}^{N_{Sp}-1} \frac{(-\ln \mathcal{P}_0)^j}{j!}, \quad (5.10)$$

wobei

$$\mathcal{P}_0 = \prod_{i=1}^{N_{Sp}} \mathcal{P}_{Spur}(p_{T,i}). \quad (5.11)$$

Kombiniert man die Spur-Kompatibilitäten anhand Gleichung 5.10 zu einer sogenannten *Vertex-Wahrscheinlichkeit*, erhält man, sofern die transversalen Spuren am Vertex unkorreliert sind, eine für *Minimum-Bias*-Ereignisse flach verteilte Größe zwischen null und eins. Für alle Signalereignisse sollte  $\mathcal{P}_{Vtx}$  im Großteil aller Fälle kleine Vertex-Wahrscheinlichkeiten liefern. In Abbildung 5.7 sind die  $\mathcal{P}_{Vtx}$ -Verteilungen für alle drei Signal- und für *Minimum-Bias*-Ereignisse zu sehen. Für *Minimum-Bias*-Ereignisse verteilt sich die Vertex-Wahrscheinlichkeit nahezu flach. Die Signalereignisse häufen sich dagegen bei kleinen Vertex-Wahrscheinlichkeiten. Dies gilt vor allem für  $t\bar{t}$ - und  $WH$ -Ereignisse.

Ein Vorteil der angewendeten Kombinationsmethode ist, dass durch die Art der Kombination die berechnete Wahrscheinlichkeit unabhängig bleibt von der Anzahl der Eingangsvariablen. Im vorliegenden Fall entspricht dies der Spur-Multiplizität. So kann die Information der Spur-Multiplizität explizit in die Vertex-Wahrscheinlichkeit eingebunden werden, indem sie wie eine zusätzliche Spur behandelt wird. Es wird zusätzlich die Kompatibilität  $\mathcal{P}_{Vtx}^{N_{Sp}}(N_{Sp})$  bestimmt, dass ein *Minimum-Bias*-Vertex eine bestimmte Anzahl von Spuren besitzt:

$$\mathcal{P}_{Vtx}^{N_{Sp}}(N_{Sp}) = \int_{N_{Sp}}^{\infty} \mathcal{F}_{MB}(N'_{Sp}) dN'_{Sp}, \quad (5.12)$$

wobei  $\mathcal{F}_{MB}(N_{Sp})$  das normierte Spur-Multiplizitätsspektrum der *Minimum-Bias*-Ereignisse ist (Abb. 5.2(a)). Die Vertex-Wahrscheinlichkeit wird wie zuvor berechnet (lediglich die Summe läuft über einen Index mehr als im zuvor betrachteten Fall)

$$\mathcal{P}'_{Vtx}(\vec{p}_T, N_{Sp}) = \mathcal{P}'_0 \sum_{j=0}^{N_{Sp}} \frac{(-\ln \mathcal{P}'_0)^j}{j!}. \quad (5.13)$$

Wobei diesmal das Integral über die Spur-Multiplizität explizit in folgender Weise berücksichtigt wird:

$$\mathcal{P}'_0 = \mathcal{P}_{Vtx}^{N_{Sp}}(N_{Sp}) \cdot \prod_{i=1}^{N_{Sp}} \mathcal{P}_{Spur}(p_{T,i}). \quad (5.14)$$

Ganz analog kann ebenso die Sphärizität  $\Psi$  als zusätzliche Variable verwendet werden. Es wird nun aus dem normierten Sphärizitätsspektrum der *Minimum-Bias*-Ereignisse (Abb. 5.3) die Kompatibilität berechnet, dass ein *Minimum-Bias*-Vertex eine gewisse Sphärizität aufweist. Da *Minimum-Bias*-Ereignisse tendenziell sphärischer sind, wird hier das Integral der Verteilung unterhalb der Sphärizität des betrachteten Vertex als Maß der Kompatibilität verwendet:

$$\mathcal{P}_{Vtx}^{\Psi}(\Psi) = \int_0^{\Psi} \mathcal{F}_{MB}(\Psi') d\Psi'. \quad (5.15)$$

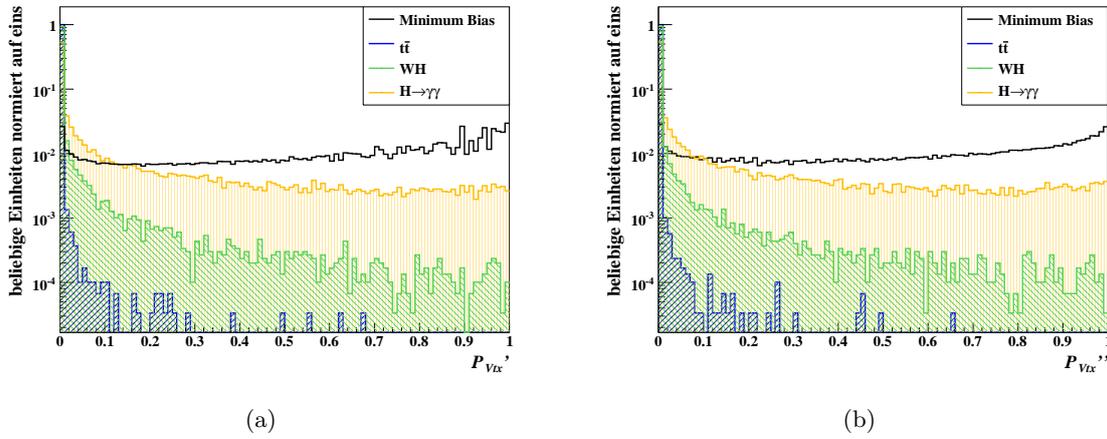
Die Vertex-Wahrscheinlichkeit unter expliziter Verwendung der Sphärizität und der Spur-Multiplizität ergibt sich durch

$$\mathcal{P}''_{Vtx}(\vec{p}_T, N_{Sp}, \Psi) = \mathcal{P}''_0 \sum_{j=0}^{N_{Sp}+1} \frac{(-\ln \mathcal{P}''_0)^j}{j!} \quad (5.16)$$

mit

$$\mathcal{P}''_0 = \mathcal{P}_{Vtx}^{\Psi}(\Psi) \cdot \mathcal{P}_{Vtx}^{N_{Sp}}(N_{Sp}) \cdot \prod_{i=1}^{N_{Sp}} \mathcal{P}_{Spur}(p_{T,i}). \quad (5.17)$$

Die Verteilung der Vertex-Wahrscheinlichkeit mit Spur-Multiplizität  $\mathcal{P}'_{Vtx}$  und die Verteilung der Vertex-Wahrscheinlichkeit mit Spur-Multiplizität und Sphärizität  $\mathcal{P}''_{Vtx}$  sind in

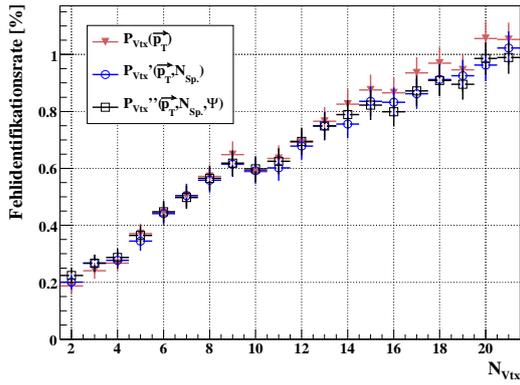
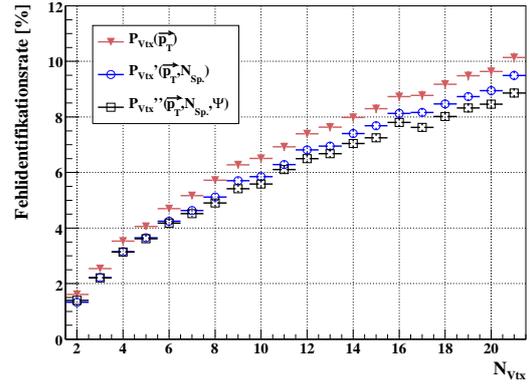
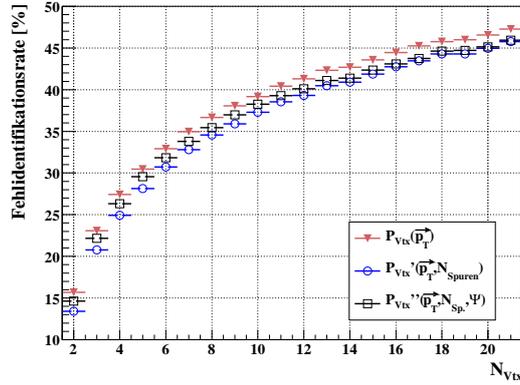


**Abb. 5.8.:** Vertex-Wahrscheinlichkeit (a)  $\mathcal{P}'_{Vtx}$  mit expliziter Einbindung der Spur-Multiplizität, (b)  $\mathcal{P}''_{Vtx}$  mit expliziter Einbindung der Spur-Multiplizität und der Sphärizität.

Abbildung 5.8 zeigt. Im Vergleich mit Abbildung 5.7 ist zu erkennen, dass die Vertex-Wahrscheinlichkeiten für *Minimum-Bias*-Ereignisse mit zusätzlichen Variablen nicht flach verteilt sind, sondern einen kleinen Anstieg bei großen Werten aufweisen. Wie bereits erwähnt ist eine flache Verteilung nur gegeben, sofern die Korrelationen zwischen den Eingangsvariablen klein sind. Korrelationen werden in Abschnitt 5.3 ausführlicher behandelt.

In Abbildung 5.9 sind die Leistungen der verschiedenen Vertex-Wahrscheinlichkeiten zu sehen. Der Leistungsvergleich zeigt, dass die Verwendung der Anzahl von Spuren leistungssteigernd auf die Vertex-Wahrscheinlichkeit wirkt. Im Falle von *WH*-Ereignissen wirkt sich auch die Sphärizität positiv aus, für die anderen beiden Signalarten dagegen ist keine Verbesserung, im Falle von  $H \rightarrow \gamma\gamma$  sogar eine Verschlechterung zu beobachten. Dies ist auf die schlechte Trennleistung der Sphärizität zurückzuführen. In der Folge werden nur die Vertex-Wahrscheinlichkeit  $\mathcal{P}'_{Vtx}$  mit Spur-Multiplizität und die allein vom Spektrum der transversalen Impulse abhängende Vertex-Wahrscheinlichkeit  $\mathcal{P}_{Vtx}$  weiterverwendet.

Wie bereits erwähnt, ist der Vorteil der vorliegenden Methode, dass lediglich die Verteilung von *Minimum-Bias*-Ereignissen verwendet wird, die schon in einem frühen Stadium des LHC-Betriebs validiert werden können. Prinzipiell erwartet man eine bessere Selektionsleistung wenn auch die Verteilungen der Signalereignisse verwendet werden. In der Folge wird eine Methode vorgestellt, die auf Signal- und *Minimum-Bias*-Verteilungen basiert.

(a)  $t\bar{t}$ (b)  $WH$ (c)  $H \rightarrow \gamma\gamma$ 

**Abb. 5.9.:** Fehlidentifikationsrate in Abhängigkeit der Vertex-Multiplizität. Verglichen werden folgende Vertex-Wahrscheinlichkeiten:  $\mathcal{P}_{Vtx}$ , basierend auf den transversalen Impulsen der Spuren,  $\mathcal{P}'_{Vtx}$  zusätzlich basierend auf der Spur-Multiplizität und  $\mathcal{P}''_{Vtx}$  zusätzlich basierend auf der Spur-Multiplizität und der Sphärizität für (a)  $t\bar{t}$ - (b)  $WH$ - und (c)  $H \rightarrow \gamma\gamma$ -Ereignisse.

### 5.2.3. Methode des Likelihood-Verhältnisses

Setzt man die Verteilungen sowohl der *Minimum-Bias*-Spektren als auch der Signalspektren als bekannt voraus, kann anhand dieser Verteilungen die Information aller Eigenschaften, solange diese von einander unabhängig sind, optimal genutzt werden. Jede gemessene Eigenschaft erhält aus der Information der Verteilungen ein Gewicht, das angibt, ob die Eigenschaft signal- oder untergrundartig (*Minimum-Bias*-artig) ist. In der sogenannten *Likelihood-Methode* wird ein solches Vorgehen angewendet. An dieser Stelle soll kurz das Prinzip der Likelihood-Methode erläutert werden, bevor eine solche Methode für die Vertex-Selektion angewendet und getestet wird.

Ein Ereignis habe  $n$  Eigenschaften (hier hat z.B. ein Vertex  $N_{Sp}$  verschiedene transversale Impulse der Spuren und somit  $N_{Sp}$  Eigenschaften) anhand derer man es von seinem Untergrund trennen kann. Es seien  $\mathcal{F}_{Sig}^i$  die Wahrscheinlichkeitsdichtefunktionen der  $i$ -ten Eigenschaft der Signalereignisse und analog  $\mathcal{F}_{MB}^i$  die der Untergrundereignisse. Sind diese bekannt, kann für jede Messung  $x_i$  der  $i$ -ten Eigenschaft eines Ereignisses ein Gewicht  $G$  berechnet werden das angibt, ob der gemessene Wert signal- oder untergrundartig ist, indem man das Verhältnis

$$G_i(x_i) = \frac{\mathcal{F}_{Sig}^i(x_i)}{\mathcal{F}_{MB}^i(x_i)} \quad (5.18)$$

aus der Wahrscheinlichkeitsdichte des Signals und der des Untergrunds bei diesem gemessenen Wert bildet. Je größer  $G$  ist, desto signalartiger ist die gemessene Eigenschaft, liegt es zwischen null und eins ist der Wert der Eigenschaft untergrundartig. Aus diesen  $n$  Gewichten  $G_i$  kann durch Multiplikation über alle Eigenschaften das Likelihood-Verhältnis berechnet werden. Es ist dabei üblich anstatt der einfachen Multiplikation den Logarithmus des Produkts zu berechnen:<sup>4</sup>

$$L = \log\left(\prod_{i=1}^n G_i(x_i)\right) = \sum_{i=1}^n \log(G_i(x_i)) = \sum_{i=1}^n \log\left(\frac{\mathcal{F}_{Sig}^i(x_i)}{\mathcal{F}_{MB}^i(x_i)}\right). \quad (5.19)$$

Die Likelihood-Methode verwendet die aus den Verteilungen zur Verfügung stehende Information optimal, wenn die verschiedenen Eingangsvariablen nicht korreliert sind (Neyman-Perason-Lemma). Es ist aus diesem Grund zu erwarten, sofern man davon ausgeht, dass die transversalen Impulse der Spuren nicht korreliert sind, dass die Likelihood-Methode die beste Selektionsleistung aufweist.

Bei der Anwendung der Likelihood-Methode wird man zunächst mit dem Problem konfrontiert, dass die Wahrscheinlichkeitsdichtefunktionen nicht bekannt sind. Aus der Monte-Carlo-Simulation stehen aber Ereignisse zur Verfügung, von denen man erwartet, dass sie

<sup>4</sup>Der Logarithmus ist eine stetige, monoton steigende Funktion und verursacht insofern nur eine Verzerrung und eine Verschiebung der Diskriminanten.

sich gemäß der theoretischen, nicht bekannten Wahrscheinlichkeitsdichtefunktion verteilen. Die auf die Anzahl der Ereignisse normierte Verteilung kann als Annäherung an die Wahrscheinlichkeitsdichtefunktion verwendet werden. Dies ist bei ausreichend Statistik und ein wenig Vertrauen in die Simulation legitim. Auch für ausreichende Statistik stößt man bei stark abfallenden Verteilungen wie den Spektren des transversalen Impulses (Abb. 5.1) trotz allem schnell auf Schwierigkeiten. Die langen Ausläufer fluktuieren sehr stark und somit auch das zu berechnende Verhältnis. Solche Schwankungen im Likelihood-Verhältnis entsprechen nicht der theoretischen Erwartung und führen zu Verzerrungen der Gewichte. Eine Möglichkeit dieses Problem zu umgehen ist es eine Funktion an einen Teil der Verteilung mit ausreichend Statistik anzupassen. Die Funktion wird dann in Bereichen, in denen starke Fluktuationen auftreten an Stelle der Verteilung verwendet. Zu diesem Zweck wurde an die Verteilungen  $\log(\mathcal{F}_{Sig}^{p_T}(p_T)/\mathcal{F}_{MB}^{p_T}(p_T))$  aller drei Signalarten folgende Funktion angepasst:

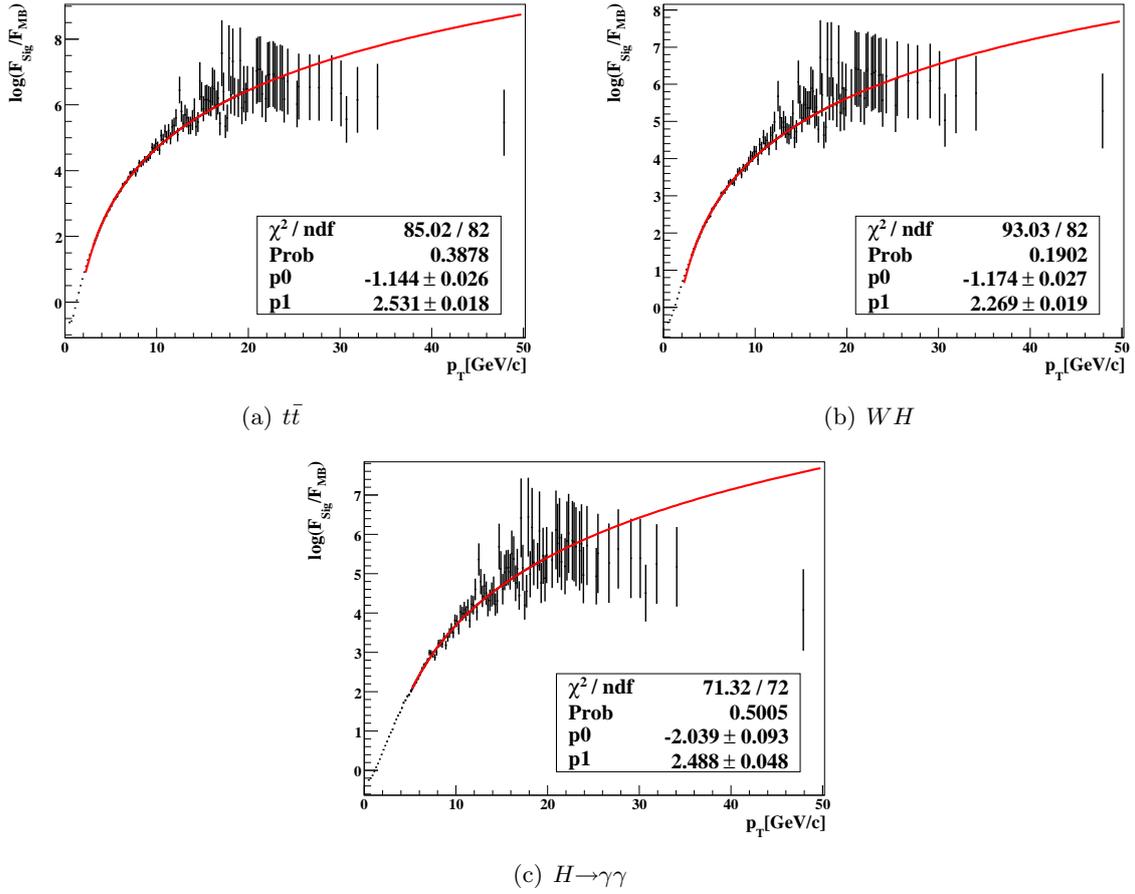
$$f(p_T) = p_0 + p_1 \cdot \log(p_T) . \quad (5.20)$$

In Abbildung 5.10 ist diese Anpassung für die drei Signalsorten gezeigt. Für transversale Impulse  $p_T > 10 \text{ GeV}/c$  wird die angepasste Funktion an Stelle der Verteilung benutzt. Es mag fraglich sein, ob die Funktion die Verteilung für große transversale Impulse richtig beschreibt. Allerdings sind Spuren mit einem transversalen Impuls dieser Stärke prinzipiell sehr signalartig und aus diesem Grund ist es faktisch irrelevant, ob dieser Bereich exakt beschrieben wird. Es ist lediglich ausschlaggebend, dass diese Spuren als stark signalartig eingestuft werden.

Die Vertex-Selektion hat eine Besonderheit gegenüber sonstigen Anwendungen der Likelihood-Methode. Die Anzahl der Eingangsvariablen ist im vorliegenden Fall nicht konstant, da die Anzahl der Spuren und damit die Anzahl der betrachteten transversalen Impulse variiert. Man vergleicht bei jeder Selektion also Diskriminanten, die von unterschiedlich vielen Eingangsvariablen abhängen. Die Likelihood-Methode garantiert in einem solchen Fall nicht die optimale Trennung. Es ist aber möglich, die Anzahl von Spuren in die Likelihood-Diskriminante einzubinden. Hiermit lässt sich der Effekt der unterschiedlichen Anzahl an Eingangsvariablen ausgleichen. Zusätzlich lässt sich in gleicher Weise die Sphärizität berücksichtigen. Folgende drei Diskriminanten wurden getestet:

$$L(\vec{p}_T) = \sum_{i=1}^{N_{Sp}} \log \left( \frac{\mathcal{F}_{Sig}^{p_T}(p_{T,i})}{\mathcal{F}_{MB}^{p_T}(p_{T,i})} \right), \quad (5.21)$$

$$L'(\vec{p}_T, N_{Sp}) = L(\vec{p}_T) + \log \left( \frac{\mathcal{F}_{Sig}^{N_{Sp}}(N_{Sp})}{\mathcal{F}_{MB}^{N_{Sp}}(N_{Sp})} \right), \quad (5.22)$$



**Abb. 5.10.:** Anpassung der Funktion  $f(p_T) = p_0 + p_1 \cdot \log(p_T)$  an die Verteilung  $\log(\mathcal{F}_{Sig}^{p_T}(p_T)/\mathcal{F}_{MB}^{p_T}(p_T))$  aus den Verteilungen der transversalen Impulse der Signal- ( $\mathcal{F}_{Sig}^{p_T}$ ) und der *Minimum-Bias*-Spuren ( $\mathcal{F}_{MB}^{p_T}$ ) für (a)  $t\bar{t}$ - (b)  $WH$ - und (c)  $H \rightarrow \gamma\gamma$ -Ereignisse. Der Bereich der Anpassung wurde im Fall von  $WH$ - und  $t\bar{t}$ -Ereignissen auf ein Intervall von  $(3 \text{ GeV}/c < p_T < 20 \text{ GeV}/c)$  eingeschränkt. Im Fall der  $H \rightarrow \gamma\gamma$ -Ereignisse ist das Anpassungsintervall  $(5 \text{ GeV}/c < p_T < 20 \text{ GeV}/c)$ .

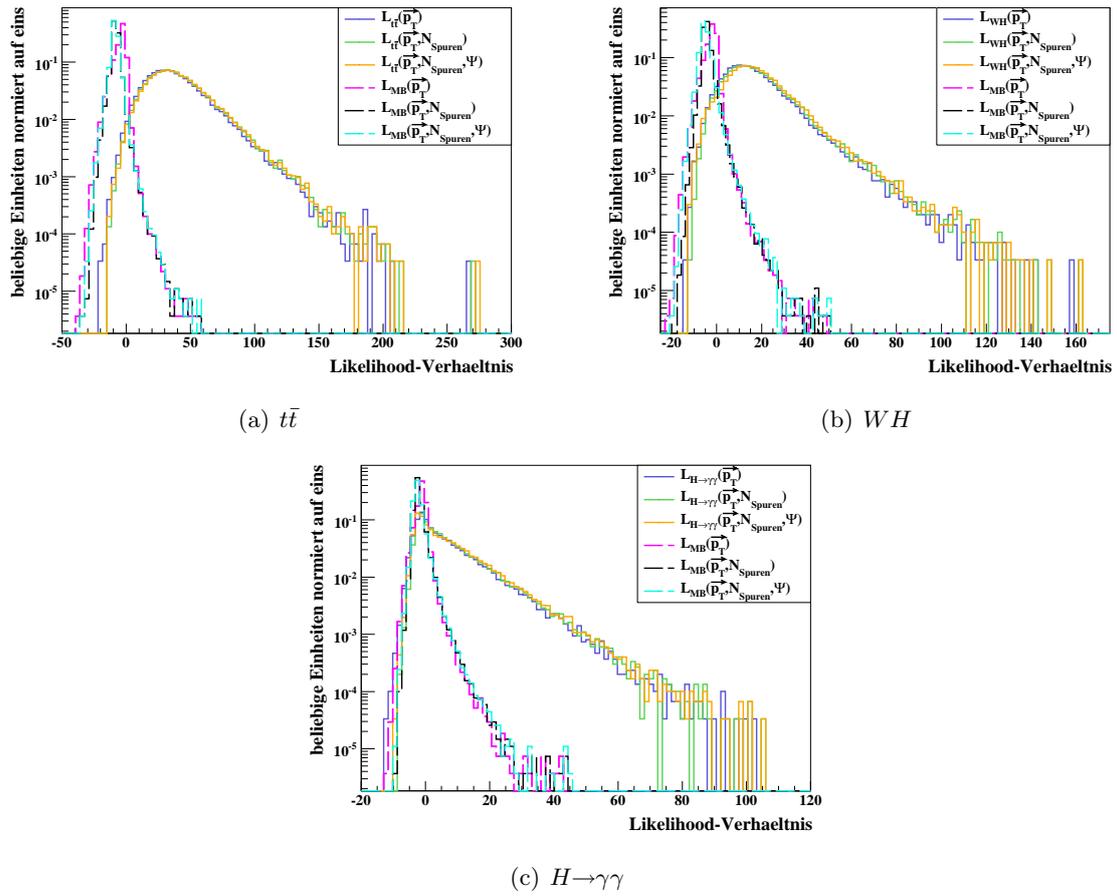
$$L''(\vec{p}_T, N_{Sp}, \Psi) = L'(\vec{p}_T, N_{Sp}) + \log \left( \frac{\mathcal{F}_{Sig}^\Psi(\Psi)}{\mathcal{F}_{MB}^\Psi(\Psi)} \right). \quad (5.23)$$

Da diese anders als für die  $\mathcal{S}_{p_T}$ - oder  $\mathcal{P}_{Vtx}$ -Methoden direkt von der Signalart abhängen, wurden für alle Signalarten signalspezifische Diskriminanten bestimmt. Zur optimalen Trennung müssen für unterschiedliche Signalarten unterschiedliche Diskriminanten verwendet werden. Dies ist prinzipiell ein Nachteil dieser Methode.

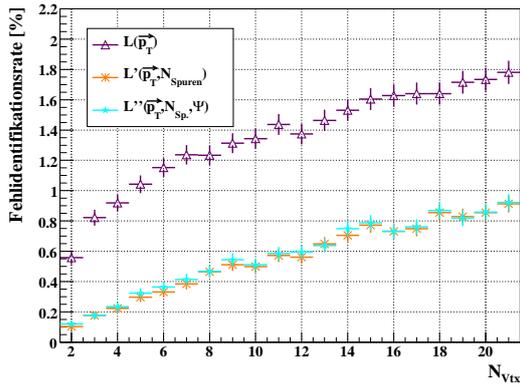
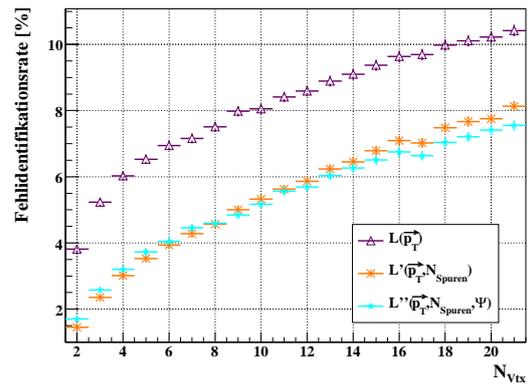
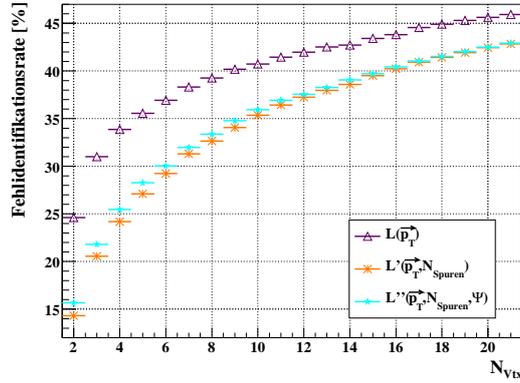
In den Abbildungen 5.11 sind für alle drei Signalsorten die drei Diskriminanten für Signal und Untergrund aufgetragen. Für  $WH$ - und  $t\bar{t}$ -Ereignisse zeichnet sich eine deutlich vom Untergrund getrennte Verteilung der Diskriminanten ab. Für  $H \rightarrow \gamma\gamma$ -Ereignisse überlappen dagegen Signal und Untergrund auch für kleine Werte der Diskriminanten. Die Unterschiede zwischen den verschiedenen Likelihood-Methoden sind in der vorliegenden Darstellung schlecht zu erkennen. Im folgenden Leistungstest werden die Unterschiede aber sichtbar.

Die Leistungen der drei Likelihood-Methoden wurden wie auch zuvor gegeneinander getestet. Das Ergebnis ist in den Abbildungen 5.12 zu sehen. Es zeigt sich für alle drei Signalsorten eine deutliche Leistungssteigerung durch die Verwendung der Spur-Multiplizität. Die zusätzliche Verwendung der Sphärizität führt zu keinem signifikanten Leistungsunterschied.

Es stehen nun drei Arten von Methoden zur Selektion des Signalvertex zu Verfügung. Im Folgenden werden die verschiedenen Methoden einander gegenüber gestellt.



**Abb. 5.11.:** Verteilungen der verschiedenen Likelihood-Diskriminanten der Signal- und Untergrundeignisse für (a)  $t\bar{t}$ , (b)  $WH$  und (c)  $H \rightarrow \gamma\gamma$ .

(a)  $t\bar{t}$ (b)  $WH$ (c)  $H \rightarrow \gamma\gamma$ 

**Abb. 5.12.:** Fehlidentifikationsrate der verschiedenen Likelihood-Methoden für (a)  $t\bar{t}$ - (b)  $WH$ - und (c)  $H \rightarrow \gamma\gamma$ -Ereignisse. Es werden folgende Likelihood-Diskriminanten verglichen:  $L$  basierend auf den transversalen Impulsen der Spuren am Vertex,  $L'$  zusätzlich basierend auf der Spur-Multiplizität und  $L''$  zusätzlich basierend auf der Spur-Multiplizität und der Sphärizität.

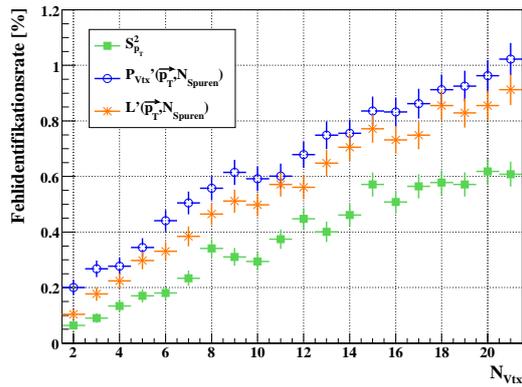
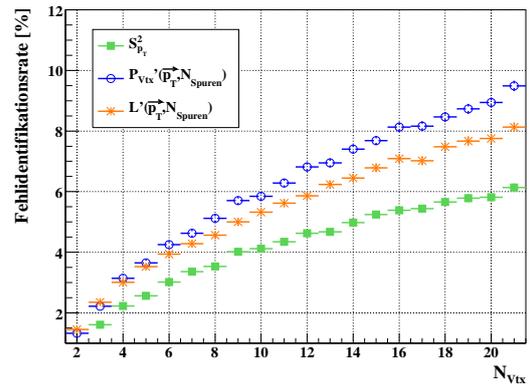
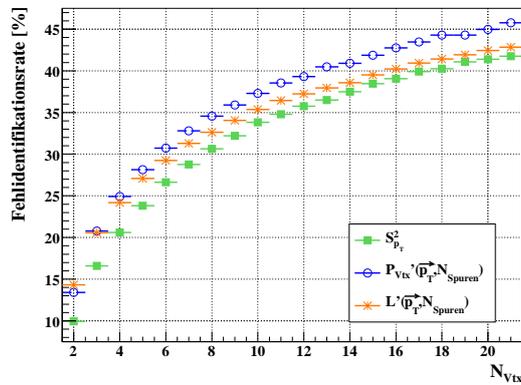
### 5.2.4. Leistung der Methoden im Vergleich

Die folgenden Methoden werden in diesem Abschnitt verglichen:

- $\mathcal{S}_{p_T}^2$ , die beste aller summenartigen Methoden.
- $\mathcal{P}'_{Vtx}$ , die Vertex-Wahrscheinlichkeit mit der höchsten Leistungsfähigkeit. Von dieser ist eine Leistungssteigerung gegenüber von  $\mathcal{S}_{p_T}^2$  zu erwarten, da die Information des Verlaufs der *Minimum-Bias*-Verteilung verwendet wird.
- $L'$ , die Likelihood-Diskriminante abhängig von den transversalen Impulsen der Spuren und von der Spur-Multiplizität, die unter den Likelihood-Methoden die höchste Leistungsfähigkeit gezeigt hat. Von dieser Methode ist die beste Leistung gegenüber allen anderen Methoden nach dem Neyman-Pearson-Lemma zu erwarten.

Es sei hier noch einmal betont, dass bei der Vertex-Wahrscheinlichkeit  $\mathcal{P}'_{Vtx}$  und der Likelihood-Methode nur dann sicher eine Verbesserung gegenüber der  $\mathcal{S}_{p_T}^2$ -Methode zu erwarten ist, wenn die zur Trennung verwendeten Variablen nicht oder nur sehr wenig korreliert sind. Sollten keine Korrelationen vorliegen, muss vor allem die Likelihood-Methode eine mindestens ebenso gute oder aber bessere Trennleistung zeigen wie die summenartigen Methoden. In den Abbildungen 5.13 sind die Vergleiche zwischen den oben angeführten Methoden zu sehen. Es ergibt sich ein den Erwartungen widersprechendes Ergebnis. Für alle Signalsorten zeigt die  $\mathcal{S}_{p_T}^2$ -Methode eine sehr gute Leistung. Die Vertex-Wahrscheinlichkeit ist auf allen Signalen die leistungsschwächste Methode. Die Tatsache, dass auch die Likelihood-Methode keine bessere Leistung liefert als die  $\mathcal{S}_{p_T}^2$ -Methode spricht dafür, dass starke Korrelationen zwischen den transversalen Impulsen der Spuren am Vertex vorliegen.

Bevor der Einfluss der Korrelation im nächsten Abschnitt betrachtet wird, soll an dieser Stelle noch ein Test durchgeführt werden, um zu berücksichtigen, dass sich die Vertex-Multiplizitäten in der Realität gemäß einer *Poisson*-Funktion um einen bestimmten Mittelwert verteilen werden. In Abbildung 5.14 sind noch einmal die Fehlidentifikationsraten der drei besten Methoden zu sehen. Diesmal ist nun aber zusätzlich für dieselben Methoden auch die Fehlidentifikationsrate in Abhängigkeit vom Mittelwert der *Poisson*-Verteilung aufgetragen. Hier wurde für jeden Test eine Vertex-Multiplizität zufällig aus einer *Poisson*-Verteilung gezogen. Die x-Achse der vorliegenden Abbildung gibt entweder die bisher betrachtete feste Vertex-Multiplizität  $N_{Vtx}$  oder den Mittelwert  $\mu_{Pois}$  der *Poisson*-Verteilung an, aus der die Vertex-Multiplizität zufällig gezogen wurde. Es ist für  $t\bar{t}$ - und  $WH$ -Ereignisse keine signifikante Änderung zu beobachten. Für  $H \rightarrow \gamma\gamma$ -Ereignisse sieht man dagegen eine kleine Vereinfachung der Selektion bei *Poisson*-verteilter Vertex-Multiplizität, vor allem für kleine Vertex-Multiplizitäten. Nach den Studien aus

(a)  $t\bar{t}$ (b)  $WH$ (c)  $H \rightarrow \gamma\gamma$ 

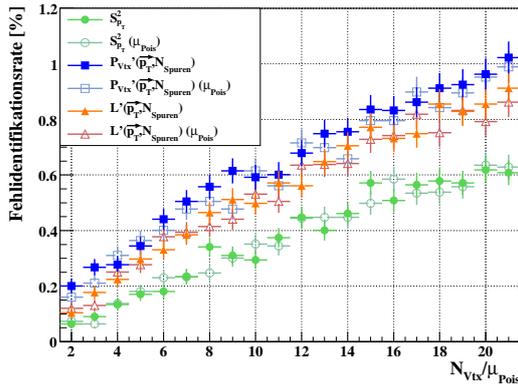
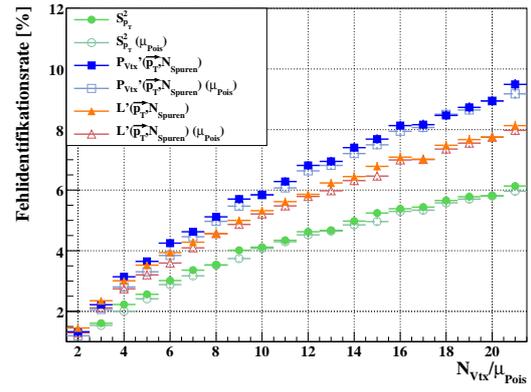
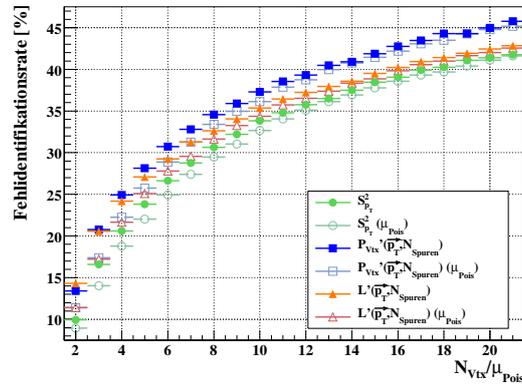
**Abb. 5.13.:** Fehlidentifikationsrate verschiedener Selektionsmethoden im Vergleich für (a)  $t\bar{t}$ - (b)  $WH$ - und (c)  $H \rightarrow \gamma\gamma$ -Ereignisse. Verglichen wird die  $S^2_{p_T}$ -, die  $\mathcal{P}'_{Vtx}$ - und die Likelihood-Methode ( $L'$ ).

dem Abschnitt 3.2.2 hatte sich gezeigt, dass die Wahrscheinlichkeit, zwei oder mehr Vertices in einem bestimmten Intervall vorzufinden, bei einer Verteilung der Vertices gemäß einer *Poisson*-Funktion leicht gegenüber der Verwendung immer einer bestimmten Vertex-Multiplizität ansteigt. Grund hierfür waren die Ausläufer der *Poisson*-Verteilung. Es wäre zu erwarten, dass seltene, sehr hohe Vertex-Multiplizitäten aus diesen Ausläufern einen Leistungsabfall bei der Vertex-Selektion verursachen. Dieser Effekt wird durch einen weiteren Einfluss überlagert. Mit wachsender Vertex-Multiplizität flacht die Kurve der Fehlidentifikation vor allem im Falle von  $H \rightarrow \gamma\gamma$  leicht ab. Eine Vertex-Multiplizität aus dem Ausläufer der *Poisson*-Verteilung hat aus diesem Grund einen weniger starken Einfluss. Prinzipiell ist festzuhalten, dass die Unterschiede klein sind und dass man anhand der bisher durchgeführten Tests ohne Berücksichtigung der *Poisson*-Verteilung der Vertex-Multiplizität das Verhalten in Abhängigkeit der Vertex-Multiplizität gut abschätzen kann.

### 5.3. Korrelationen der Spur-Eigenschaften

Wie zuvor bereits erwähnt, kann die Leistungsfähigkeit der Likelihood und die der Vertex-Wahrscheinlichkeit durch Korrelation zwischen den verschiedenen Eingangsvariablen eingeschränkt werden. Der Leistungsvergleich zwischen den verschiedenen Selektionsmethoden und die Tatsache, dass die Likelihood sich nicht als die optimale Methode herausgestellt hat, weist klar darauf hin, dass Korrelationen vorliegen. Schon die Verteilungen der Vertex-Wahrscheinlichkeit (siehe Abbildungen 5.7 und 5.8) geben Anlass zu dieser Annahme. Sollten alle hier verwendeten Eigenschaften unabhängig sein, wäre eine absolut flache Verteilung der Vertex-Wahrscheinlichkeit für *Minimum-Bias*-Ereignisse zu erwarten gewesen. Es zeigt sich aber, dass vor allem bei Hinzunahme der Spur-Multiplizität keine flache Verteilung mehr vorliegt. Dies lässt darauf schließen dass die Anzahl an Spuren am Vertex mit dem transversalen Impuls korreliert ist. Abbildung 5.15 zeigt die Verteilung der Spur-Multiplizität in Abhängigkeit des größten, am Vertex vorkommenden transversalen Impuls  $p_{T,max}$ . Es ist zu beobachten, dass mit ansteigender Anzahl von Spuren auch der maximale transversale Impuls  $p_{T,max}$  am Vertex ansteigt.

Es wäre von Interesse abschätzen zu können, wie die transversalen Impulse der Spuren am Vertex untereinander korreliert sind. Wäre dies möglich, könnte man die Eigenschaften der Vertices dekorrelieren, bevor man z.B. eine Likelihood-Diskriminante berechnet. Im Falle der transversalen Impulse ist es aber nicht mit einfachen Mitteln möglich, die Korrelationen zu erkennen. Hier liegt ein sehr komplexes System vor, da je ein beliebiger Satz von Spuren am Vertex korreliert sein kann, während der Rest davon unabhängig bleibt.

(a)  $t\bar{t}$ (b)  $WH$ (c)  $H \rightarrow \gamma\gamma$ 

**Abb. 5.14.:** Fehlidentifikationsrate verschiedener Methoden gegen eine bestimmte Anzahl von Vertices  $N_{Vtx}$  und gegen den *Poisson*-Mittelwert  $\mu_{Pois}$  der Vertex-Multiplizitätsverteilung. Es werden (a)  $t\bar{t}$ - (b)  $WH$ - und (c)  $H \rightarrow \gamma\gamma$ -Ereignisse betrachtet.

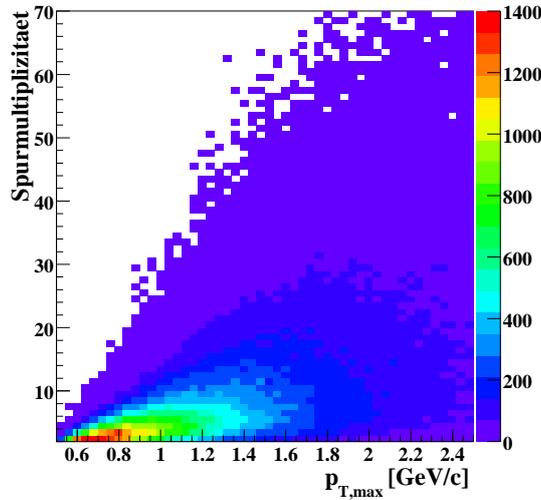


Abb. 5.15.: Spur-Multiplizität in Abhängigkeit des maximalen transversalen Impulses  $p_{T,max}$  im Ereignis.

So mag im einen Fall die Spur mit dem größten transversalen Impuls abhängig sein vom nächst stärkeren transversalen Impuls, im nächsten Fall besteht die Abhängigkeit dann aber zwischen dem zweit, viert und fünft stärksten transversalen Impuls.

Man ist zwar nicht in der Lage, mit einfachen Mitteln die Korrelation zu bestimmen, dennoch ist es möglich eine Aussage darüber zu machen, ob es auf Korrelationen zurückzuführen ist, dass der Leistungsvergleich ein nach dem Neyman-Pearson-Lemma unerwartetes Ergebnis gezeigt hat. Hierfür werden die oben gemachten Tests erneut durchgeführt, wobei diesmal *künstliche* Vertizes verwendet werden, bei denen keine Korrelationen vorliegen. Solche korrelationsfreien Vertizes werden nach dem folgenden Rezept erzeugt:

- Aus dem Spektrum der Spur-Multiplizität zieht man zufällig eine Anzahl von Spuren  $N_{Sp}$ .
- Nun zieht man  $N_{Sp}$ -mal zufällig aus dem Spektrum der transversalen Impulse einen  $p_T$ -Wert.
- Aus diesen  $N_{Sp}$  transversalen Impulsen kann nun ein *künstlicher* Vertex gebildet werden, für den alle Methoden berechnet und getestet werden können.

Für *Minimum-Bias* und für jede Signalart wurden so aus den zugehörigen Spektren Vertizes gewürfelt und ganz analog zu den richtige Vertizes wurden diese *künstlichen* Signalvertizes mit *künstlichen Minimum-Bias*-Vertizes gemischt. Nun konnte wieder die Fehlidentifikationsrate für verschiedene Vertex-Multiplizitäten berechnet werden. Im Falle von  $WH$  und  $H \rightarrow \gamma\gamma$  wurde die Fehlidentifikationsrate für jede Vertex-Multiplizität mit 100 000, im Falle von  $t\bar{t}$  mit 1 000 000 solcher Ereignisse bestimmt. Für diesen Korrelationstest wurde

auch die einfache  $\mathcal{P}_{Vtx}(p_T)$ - und die einfache Likelihood-Methode  $L(\vec{p}_T)$  erneut betrachtet. Auch von diesen Methoden ist für unkorrelierte Eigenschaften eine bessere Leistung zu erwarten als von der  $\mathcal{S}_{p_T}^2$ -Methode.

In den Abbildungen 5.16 ist auf der linken Seite das Ergebnis des Tests an Vertizes mit unkorrelierten Eigenschaften zu sehen. Zum Vergleich wurden diesen die Abbildungen für die realen Vertizes auf der rechten Seite gegenübergestellt. Für unkorrelierte transversale Impulse zeigt sich ein Verhalten, welches dem nach dem Neyman-Pearson-Lemma vorausgesagten Verhalten entspricht. Die beste Selektion wird von der Likelihood-Methode mit Verwendung der Spur-Multiplizität erreicht. Eine sehr gute Leistung zeigt auch die Vertex-Wahrscheinlichkeit, sowohl mit als auch ohne Verwendung der Spur-Multiplizität. Vergleicht man die Leistung der einfachen Likelihood  $L(\vec{p}_T)$  mit der der einfachen Vertex-Wahrscheinlichkeit  $P_{Vtx}(\vec{p}_T)$ , zeigt sich dass die Vertex-Wahrscheinlichkeit bei gleicher Eingangsinformation eine bessere Leistung aufweist. Während die Vertex-Wahrscheinlichkeit von der variierenden Anzahl an Eigenschaften nicht beeinflusst wird, scheint dies für die Likelihood-Methode eine Schwierigkeit darzustellen. Die Likelihood-Methode ist für Vergleiche mit unterschiedlich vielen Eingangseigenschaften nicht optimal. Berücksichtigt man aber die Spur-Multiplizität bei der Likelihood-Methode  $L'(\vec{p}_T, N_{Sp})$  wird die höchste hier beobachtete Leistung erreicht. Schließlich und endlich schneidet  $\mathcal{S}_{p_T}^2$  unter allen Methoden auf den *künstlichen* Vertizes am schlechtesten ab. Wären die transversalen Spuren am Vertex nicht korreliert, würde die Likelihood-Methode und die Vertex-Wahrscheinlichkeit die beste Trennung erreichen. Aus der offensichtlichen Diskrepanz zwischen der rechten und linken Seite von Abbildung 5.16 kann eindeutig geschlossen werden, dass Korrelationen zwischen den transversalen Impulsen der Spuren vorliegen, die die Vertex-Wahrscheinlichkeit und die Likelihood-Methode beeinträchtigen. Lediglich  $\mathcal{S}_{p_T}^2$  scheint mit diesen Korrelationen in der richtigen Form umzugehen.

Als Ergebnis bleibt festzuhalten, dass eine sehr simple Methode zur Vertex-Selektion den Signalvertex mit der besten Effizienz unter verschieden vielen *Minimum-Bias*-Vertizes identifizieren kann. Es stellt sich die Frage, ob eine höhere Effizienz überhaupt noch erreicht werden kann, oder ob  $\mathcal{S}_{p_T}^2$  optimal mit den Korrelationen umgeht. Im nächsten Kapitel wird eine multivariate Methode vorgestellt, die in der Lage ist, Korrelationen optimal zu nutzen, ohne dass hierfür die Korrelation selbst bekannt sein muss.

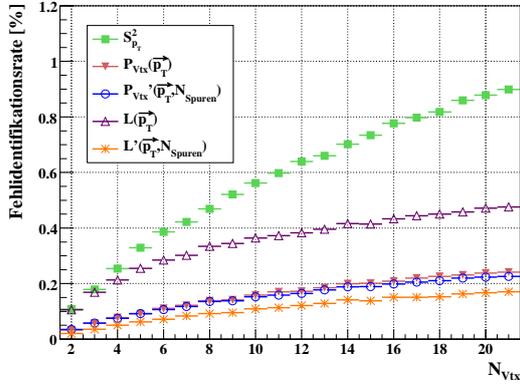
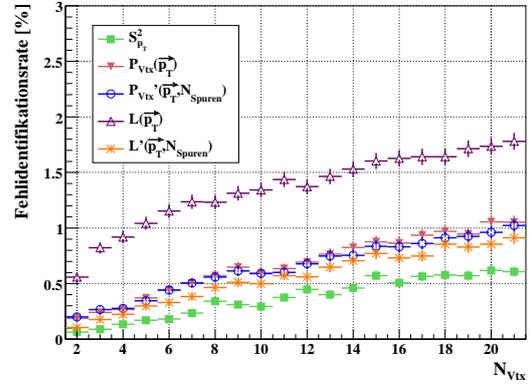
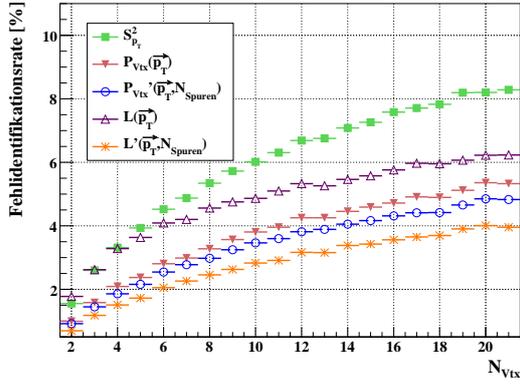
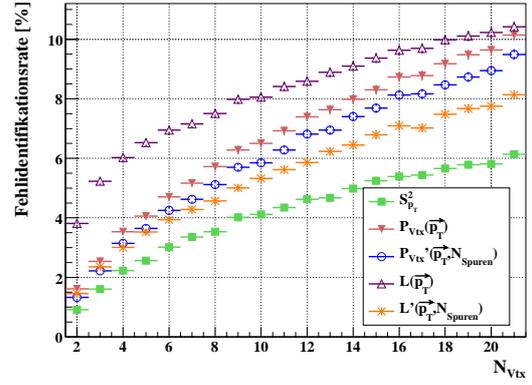
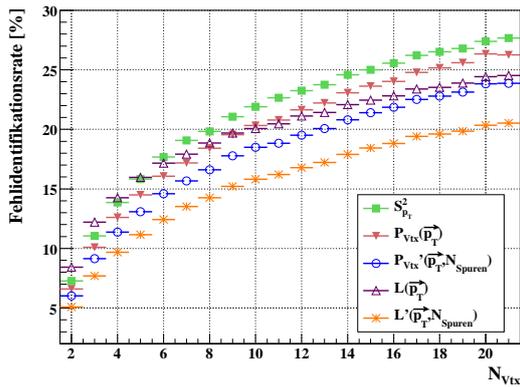
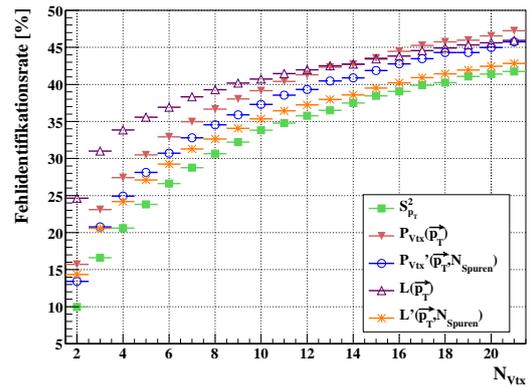
(a) künstliche  $t\bar{t}$ -Vertizes(b) reale  $t\bar{t}$ -Vertizes(c) künstliche  $WH$ -Vertizes(d) reale  $WH$ -Vertizes(e) künstliche  $H \rightarrow \gamma\gamma$ -Vertizes(f) reale  $H \rightarrow \gamma\gamma$ -Vertizes

Abb. 5.16.: Fehlidentifikationsrate verschiedener Methoden getestet an *künstlichen* Vertizes mit unkorrelierten Eigenschaften (links) und an in der Realität erwarteten Vertizes (rechts).

## 5.4. Vertex-Selektion durch ein künstliches neuronales Netz

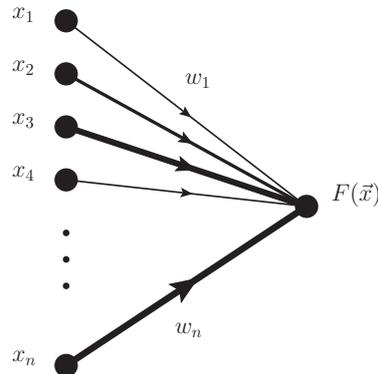
Künstliche neuronale Netze sind bei der Analyse von Daten, denen ein komplexes Muster zugrunde liegt, ein vielverwendetes und bewährtes Werkzeug. Die Analogie zu neuronalen Netzen im Gehirn gibt ihnen ihren Namen. Im Folgenden werden künstliche neuronale Netze einfach neuronale Netze (NN) genannt, es ist dabei immer von einem *künstlichen* neuronalen Netz die Rede. Ein mehrlagiges Netz, auch MLP (*Multilayer Perceptron*) genannt, ist in der Lage, komplexe Muster in Datensätzen zu erkennen. Aufgrund dieser guten Mustererkennung können neuronale Netze auch im Falle von unbekanntem Korrelationen, wie sie bei der Vertex-Selektion vorliegen, optimal zuordnen, zu welcher Art von Ereignisklasse ein bestimmter Satz an Eingangsvariablen gehört. Es existiert mittlerweile eine Vielzahl von unterschiedlichen Methoden zum Training und Aufbau solcher künstlicher neuronaler Netze. Innerhalb dieser Arbeit wird ein sogenanntes *feed-forward Multi-Layer-Perceptron* verwendet, welches durch *Rückwärts-Propagation* trainiert wird. Diese Begriffe sollen im nächsten Abschnitt mit Inhalt gefüllt werden. Die darauf folgenden Absätze widmen sich dann der Beschreibung von neuronalen Netzwerken zur Vertex-Selektion. Die Selektionsleistung der neuronalen Netze soll zum Abschluss dieses Kapitels mit der Leistungsfähigkeit der  $\mathcal{S}_{pT}^2$ -Methode verglichen werden.

### 5.4.1. Funktionsweise eines künstlichen neuronalen Netzes

Ein neuronales Netz bestimmt eine Funktion  $F(\vec{x})$ , die von einem Satz  $\vec{x} = (x_1, \dots, x_n)$  von  $n$  gemessenen Eingangsvariablen abhängt. Die  $x_i$  sind dabei die zur Trennung verwendeten Eigenschaften. Anhand des Funktionswertes  $F(\vec{x})$  kann eine Entscheidung gefällt werden, ob ein Ereignis signal- oder untergrundartig ist. Im Allgemeinen ist ein neuronales Netz auch in der Lage, mehr als nur zwei mögliche Ausgangsoptionen zuzulassen, da dies aber im vorliegenden Fall nicht nötig ist, wird sich hier auf den einfachen Fall einer Entscheidung zwischen entweder Signal oder Untergrund bzw. *Minimum-Bias* beschränkt. Das einfachste mögliche Netz bei  $n$  Eingangsvariablen ist in Abbildung 5.17 dargestellt. Es besteht lediglich aus den  $n$  Eingangsvariablen, die eine Verbindung zum Ausgabewert  $F(\vec{x})$  besitzen. Die Verbindungen zwischen den Ein- und Ausgangsvariablen heißen Neuronen und Punkte, an denen sich solche treffen, heißen Knoten. In diesem Fall ist  $F$  von der Form:

$$F(\vec{x}) = g\left(\sum_{i=1}^n w_i \cdot x_i + \theta\right). \quad (5.24)$$

Die  $w_i$  sind die Gewichtungen der Eingangsvariablen. Je nach Gewichtung wird eine bestimmte Eigenschaft mehr oder weniger stark bei der Trennung berücksichtigt. Exemplarisch ist dies in Abbildung 5.17 durch die Dicke der Neuronen angedeutet. Der Parameter



**Abb. 5.17.:** Minimales Netz für  $n$  Eingangsvariablen und einen Ausgabewert.

$\theta$  ist ein Schwellenwert, der die Grenze zwischen Signal- und Untergrundbereich der Diskriminanten festlegt. Die Schwelle und die Gewichtungen werden im Training bestimmt. In dieser einfachen Form bewegt sich die Information nur in eine Richtung, nämlich von den Eingangsvariablen zum Ausgabewert. Ein solches Netz wird *feed-forward*-Netz genannt. Die Funktionswerte der sogenannten *Aktivierungsfunktion*  $g$  bilden die Selektionsdiskriminante. Es ist daher sinnvoll eine Funktion zu wählen, die monoton steigt (oder fällt) und einen Ausgabewert zwischen null und eins besitzt. Ereignisse mit einem Ausgabewert nahe bei null sind dann z.B. untergrundartig, ein Ausgabewert nahe bei eins steht für ein signalartiges Ereignis. Eine mögliche Aktivierungsfunktion wäre die Stufenfunktion:

$$g(x) = \begin{cases} 0 & \text{für } x < 0 \\ 1 & \text{für } x > 0 \end{cases} \quad (5.25)$$

Die Stufenfunktion hat diskrete Ausgabewerte, ein Ereignis ist entweder Signal oder Untergrund. Durch eine *weichere* Aktivierungsfunktion können auch kontinuierlich verteilte Ausgabewerte erreicht werden, so dass auch Werte zwischen null und eins angenommen werden können. Eine gängige Aktivierungsfunktion für neuronalen Netzen ist die *Sigmoid*-Funktion:

$$g(x) = \frac{1}{1 + e^{-2x}} . \quad (5.26)$$

Durch Datensätze mit bekannter Klassenzugehörigkeit, im vorliegenden Fall also einem simulierten Signal- und *Minimum-Bias*-Datensatz, können die Gewichte  $w_i$  und der Schwellenwert  $\theta$  bestimmt werden. Es gilt die Gewichte  $w_i$  und die Schwelle  $\theta$  so zu bestimmen, dass der Unterschied zwischen  $F$  und der durch den Datensatz gegebenen Zielfunktion  $O$  minimal wird. Die Zielfunktion kann durch jedes Ereignis im Datensatz annähernd beschrieben werden.  $O_m$  sei das Muster des  $m$ -ten Ereignisses im Datensatz, welches zum

Abgleich verwendet wird und  $F_t$  sei ein Satz von Funktionen. Zu jeder Funktion  $F_t$  gehört jeweils ein bestimmter Vektor  $\vec{w}_t$  von Gewichten  $w_i$  und einer Schwelle  $\theta$ . Es gilt die Funktion  $F_t$  zu finden für die der Fehler,

$$E = \frac{1}{2N_M} \sum_{m=1}^{N_M} (F_t(\vec{x}_m) - O_m)^2 \quad (5.27)$$

minimal wird. Dabei wird in der Praxis nicht der komplette Datensatz, sondern nur eine bestimmte Anzahl von Mustern bzw. Testereignisse  $N_M$  durchlaufen bevor  $F_t$  variiert wird. Variiert werden immer jeweils die Gewichte und die Schwelle. Hierfür wird im vorliegenden Fall die sogenannte *Rückwärts-Propagation* verwendet. Jedes Gewicht  $w_i$  wird um ein kleines Stückchen entgegen die Richtung des Gradienten des Fehlers  $E$  verschoben:

$$w_{i,t+1} = w_{i,t} - \eta \cdot \frac{\partial E_t}{\partial w_i}. \quad (5.28)$$

Über den Parameter  $\eta$  kann die Schrittgröße festgelegt werden, um die die Gewichte variiert werden. Aus den neu bestimmten Gewichten ergibt sich die neue Funktion  $F_{t+1}$ . Hiermit wird der Fehler auf weiteren Mustern erneut bestimmt. Das Durchlaufen des ganzen Datensatzes entspricht einem Trainingszyklus des Netzes.

Die Funktion  $F_t$  definiert für ein solch einfaches Netz wie in Abbildung 5.17 nichts weiter als eine Schnittebene im  $n$ -dimensionalen Raum der Eingangsvariablen. Messwerte  $\vec{x}$ , die gleiche Funktionswerte der Funktion  $F_t$  liefern, liegen auf einer  $n-1$ -dimensionalen Ebene in diesem Raum. Die verschiedenen Ebenen gleicher Funktionswerte sind parallel. Lineare Korrelationen können mit einem solchen Netz berücksichtigt werden. Wünschenswert ist, dass beliebige auch gekrümmte Flächen im Raum der Eingangsvariablen durch die Funktion  $F_t$  definiert werden können. Erst dann ist es möglich auch komplexe Korrelationen berücksichtigen zu können. Hierfür muss das Netz um eine sogenannte verborgene Ebenen mit zusätzlichen Knotenpunkten erweitert werden (siehe Abbildung 5.18). Die zugehörige Funktion  $F$  ist dann:

$$F(\vec{x}) = g\left(\sum_{j=1}^k w_j g\left(\sum_{i=1}^n w_{ji} \cdot x_i + \theta_j\right) + \theta\right). \quad (5.29)$$

Jeder Knotenpunkt in der verborgenen Lage entspricht einer Schnittebene im Variablenraum. Es ist nun möglich, weit komplexere Zusammenhänge trennen zu können. Je nach Problem müssen mehr zusätzliche Knoten zur verborgenen Ebene bzw. weitere verborgene Lagen hinzugefügt werden. Es hat sich aber gezeigt, dass für die meisten Probleme eine verborgene Lage ausreicht, wenn in dieser eine ausreichende Anzahl an Knoten vorhanden ist. Das Training des Netzes läuft ganz analog wie oben beschrieben, nur der Vektor  $\vec{w}_i$ , der zu bestimmenden Gewichte und Schwellen ist gewachsen. Nach einer bestimmten

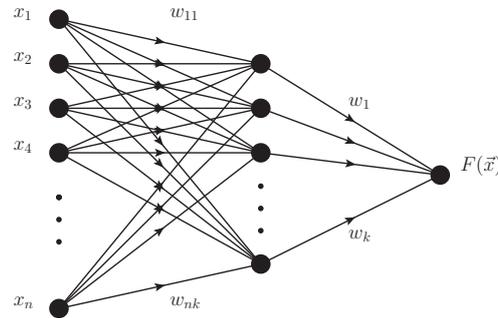


Abb. 5.18.: Neuronales Netz mit  $n$  Eingangsvariablen und einer versteckten Lage mit  $k$  Knoten.

Anzahl von Trainingszyklen ist das Netz in der Lage, den Variablenraum in einen beliebig geformten Signal- und Untergrund-Unterraum zu zerteilen. Dies kann soweit reichen, dass statistische Fluktuationen im Datensatz berücksichtigt werden. Man spricht in einem solchen Fall von *Übertraining*. Es gilt das Training rechtzeitig abzubrechen, bevor statistische Fluktuationen als Muster erkannt werden. Zu diesem Zweck wird der Datensatz in einen Trainings- und einen Testdatensatz unterteilt. Nach einer bestimmten Anzahl an Trainingszyklen, im vorliegenden Fall sind dies zehn, wird der Fehler auf dem Testdatensatz bestimmt (die Bestimmung der neuen Gewichte beruht nach wie vor nur auf dem Fehler des Trainingsdatensatzes). Dies entspricht einer Trainingsepoche. Man beobachtet sowohl den Fehler auf dem Trainings- als auch den Fehler auf dem Testdatensatz. Während der Fehler auf dem Trainingsdatensatz weiter kleiner wird, wird der Fehler auf dem Testdatensatz wieder ansteigen, wenn das Netz die statistischen Fluktuationen im Trainingsdatensatz berücksichtigt. Stellt man dies fest, kann das Training abgebrochen werden und man verwendet das Netz bei minimalem Fehler auf dem Testdatensatz.

Zum Trainieren der in der Folge betrachteten Netze wurde JETNET 3.0 verwendet, ein auf FORTRAN basiertes Paket, welches ursprünglich zur Jet-Identifizierung entwickelt wurde. Eine Beschreibung findet sich in [23].

#### 5.4.2. Neuronale Netze zur Signalvertex-Selektion

Wie bei den zuvor betrachteten Methoden stehen auch dem neuronalen Netz als Eingangsvariablen der transversale Impuls jeder Spur, die Spur-Multiplizität und die Sphärizität zur Verfügung. Unabhängig von der spezifischen Trennleistung der einzelnen Eingangsvariablen und deren Korrelationen untereinander ist ein neuronales Netz, wie oben beschrieben, in der Lage, eine optimale Trennung zwischen Signal und Untergrund zu erreichen. Aus diesem Grund wird eine zusätzliche Eingangsvariable bei richtigem Training nie zu ei-

ner Verschlechterung der Leistung führen. Ist die zusätzliche Information einer weiteren Eingangsvariablen bereits ausgeschöpft, bleibt diese vom Netz schlicht unberücksichtigt. Sphärizität und Spur-Multiplizität können demnach ohne Bedenken verwendet werden. Für die transversalen Impulse ergibt sich im Falle des neuronalen Netzes eine technische Schwierigkeit. Die Anzahl der Eingangsvariablen ist für neuronale Netze eine feste Zahl. Für das vorliegende Problem ist dies durch die variierende Spur-Multiplizität bei Verwendung der transversalen Impulse als Eingangsvariablen nicht gegeben. Aus diesem Grund wurden dem Netz an Stelle der transversalen Impulse der einzelnen Spuren die statistischen Momente der Verteilung der transversalen Impulse der Spuren am Vertex als Eingangsvariablen übergeben. In dieser Form kann die Information der transversalen Impulse aller Spuren berücksichtigt werden. Die Anzahl an Eingangsvariablen wird durch die Anzahl der verwendeten statistischer Momente festgelegt. Das nullte statistische Moment  $\mu_0$  entspricht dem Mittelwert, das erste  $\mu_1$  der Standardabweichung, das zweite  $\mu_2$  der Schiefe. Die statistischen Momente sind wie folgt definiert, wobei die höheren statistischen Momente ( $j > 0$ ) auf den Mittelwert  $\mu_0$  bezogen werden:

$$\mu_0 = \frac{1}{N_{Sp}} \sum_{i=1}^{N_{Sp}} p_{T,i} \quad (5.30)$$

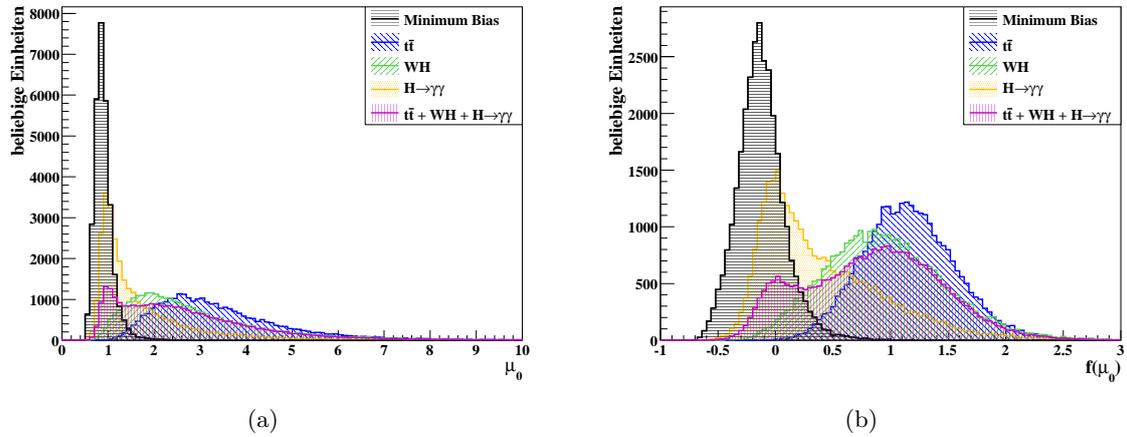
$$\mu_j = \frac{1}{N_{Sp}} \sum_{i=1}^{N_{Sp}} (p_{T,i} - \mu_0)^j \quad (\text{für } j > 0) . \quad (5.31)$$

Die Gesamtheit aller statistischen Momente gibt im Grenzwert für  $j \rightarrow \infty$  die Verteilung der transversalen Impulse im Ereignis wieder. Somit bilden die statistischen Momente eine pfiffige Lösung, um die variierende Anzahl an Eingangsvariablen zu umgehen, ohne die Information der Verteilung der transversalen Impulse am Vertex zu verlieren.

Das Training eines neuronalen Netzes wird erleichtert, wenn die Verteilungen der Eingangsvariablen möglichst kontinuierlich sind. Um statistische Fluktuationen zu minimieren wurden die Verteilungen der statistischen Momente mit folgender Funktion in den Ausläufern gestaucht und im Zentrum entzerrt:

$$f(\mu) = \begin{cases} -\log(-\mu) + 2 \cdot b & \text{für } \mu < -x_0 \\ b + a\mu & \text{für } -x_0 < \mu < +x_0 \\ \log(\mu) & \text{für } \mu > x_0 \end{cases} \quad (5.32)$$

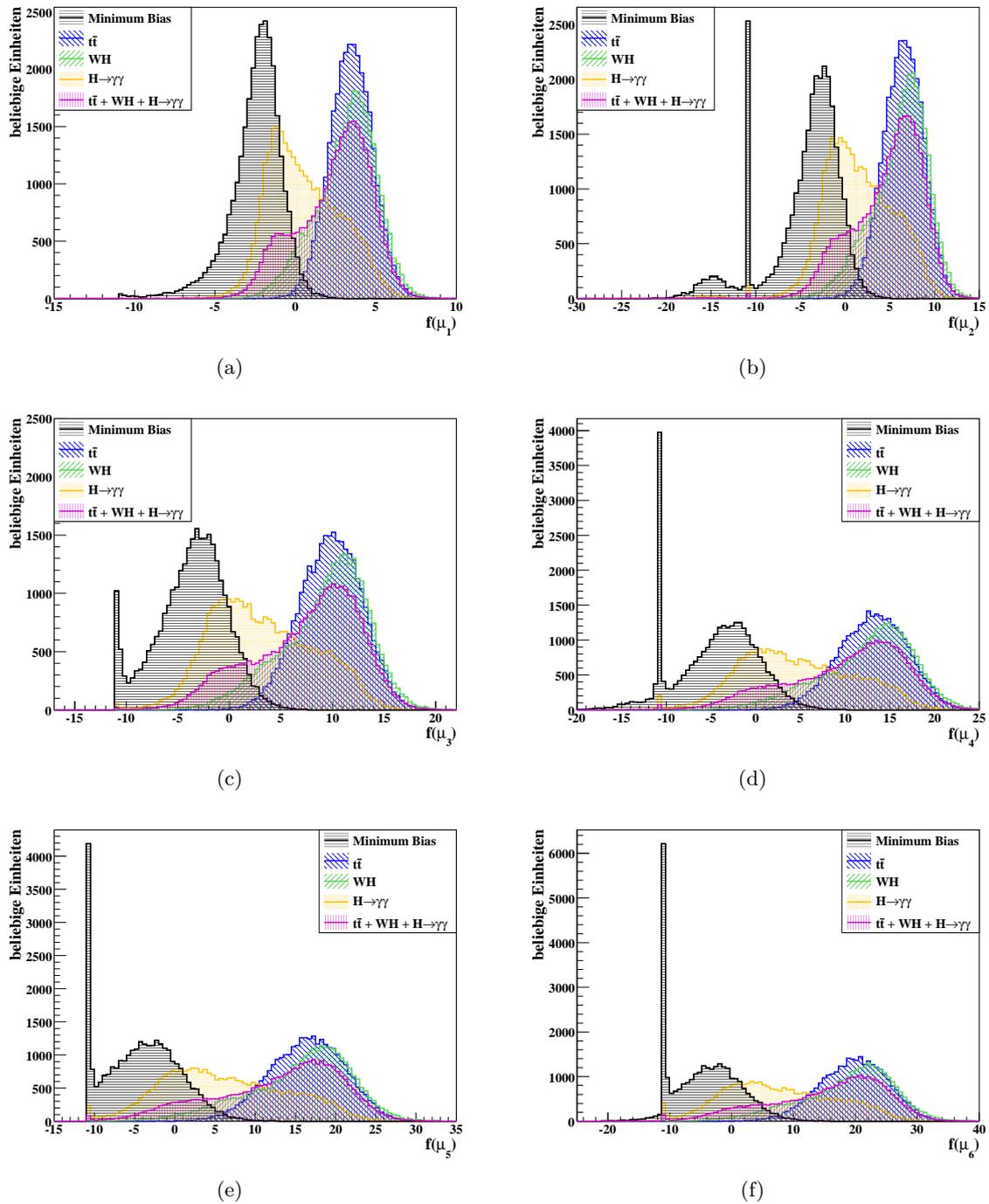
Die Funktion  $f(\mu)$  ist eine Kombination zweier am Punkt  $f(0) = b \approx -10$  gespiegelter Logarithmen. Diese werden in einem kleinen Bereich ( $|\mu| < x_0 = 5 \cdot 10^5$ ) durch eine steile Gerade ( $a = 20000$ ) stetig verbunden. Abbildung 5.19 zeigt exemplarisch die Verteilung des Mittelwerts  $\mu_0$  vor und nach der Transformation mit der obigen Funktion  $f$  für die verschiedenen Signalarten und für *Minimum-Bias*.



**Abb. 5.19.:** (a) Verteilung des Mittelwert ( $\mu_0$ ) und (b) Verteilung der Transformation des Mittelwerts ( $f(\mu_0)$ ) der Verteilung der transversalen Impulse der Spuren am Vertex, für *Minimum-Bias*- und Signalergebnisse.

Abbildung 5.20 zeigt die Verteilungen weiterer sechs transformierter statistischer Momente für Signal- und *Minimum-Bias*-Ereignisse. Unter anderem ist die Verteilung eines aus allen Signalen gemischten Datensatzes zu sehen. Die Verteilung wurde hier auf die Anzahl an *Minimum-Bias*-Ereignissen normiert. Es ergeben sich kontinuierliche Verteilungen mit kleinem Überlapp zwischen Untergrund und Signal. Die Verteilungen der Momente mit  $j > 1$  weisen eine deutliche Anhäufung der *Minimum-Bias*-Ereignisse bei  $f(0) \approx -10$  auf. Dies entspricht einem Wert von  $\mu_j = 0$ . Ursache dieser Anhäufungen sind im Falle der geraden statistischen Momente sowohl die Ein- als auch die Zwei-Spurereignisse im Falle der ungeraden nur die Zwei-Spurereignisse. Die Verteilungen der Ein-Spurereignisse weisen keine Abweichung vom Mittelwert und keine Asymmetrie auf und führen deshalb zu:  $\mu_j = 0$  (für  $j > 1$ ). Zwei-Spurereignisse haben immer eine symmetrische Verteilung. Ungerade statistische Momente, sind ein Maß für die Asymmetrie der Verteilung und verschwinden deshalb auch im Fall, dass nur zwei Spuren am Vertex vorliegen.

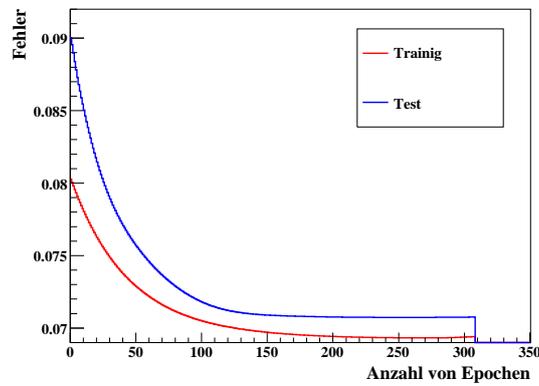
Die statistischen Momente sind schon per Definition untereinander stark korreliert und es wird mit jedem zusätzlichen Moment zum Training neben neuer Information auch bereits vorhandene Information hinzugefügt. Tabelle 5.1 zeigt beispielhaft die Korrelationskoeffizienten zwischen den ersten fünf statistischen Momenten. Die Koeffizienten liegen zwischen 0.8 und 1. Es liegen demnach sehr starke Korrelationen vor. Es zeigt sich tendenziell, dass die Korrelation zwischen größeren statistischen Momenten und benachbarten Momenten stärker ist. Es ist davon auszugehen, dass eine bestimmte Anzahl von Momenten ausreicht, um die volle Information auszuschöpfen. Anhand einer Studie an *WH*-Ereignissen wurde eine Abschätzung für die notwendige Anzahl an Momenten durchge-



**Abb. 5.20.:**  $f(\mu_j)$  für  $(j = 1, \dots, 6)$ .  $\mu_i$  sind dabei die statistischen Momente der Verteilung der transversalen Impulse am Vertex. Gezeigt sind *Minimum-Bias*-Ereignisse und Signalereignisse ( $WH$ ,  $t\bar{t}$  und  $H \rightarrow \gamma\gamma$ ), ebenso wie die Summe aller Signalereignisse (normiert auf die Anzahl an *Minimum-Bias*-Ereignissen).

	$\mu_0$	$\mu_1$	$\mu_2$
$\mu_1$	0.893		
$\mu_2$	0.819	0.977	
$\mu_3$	0.822	0.987	0.987
$\mu_4$	0.801	0.976	0.997

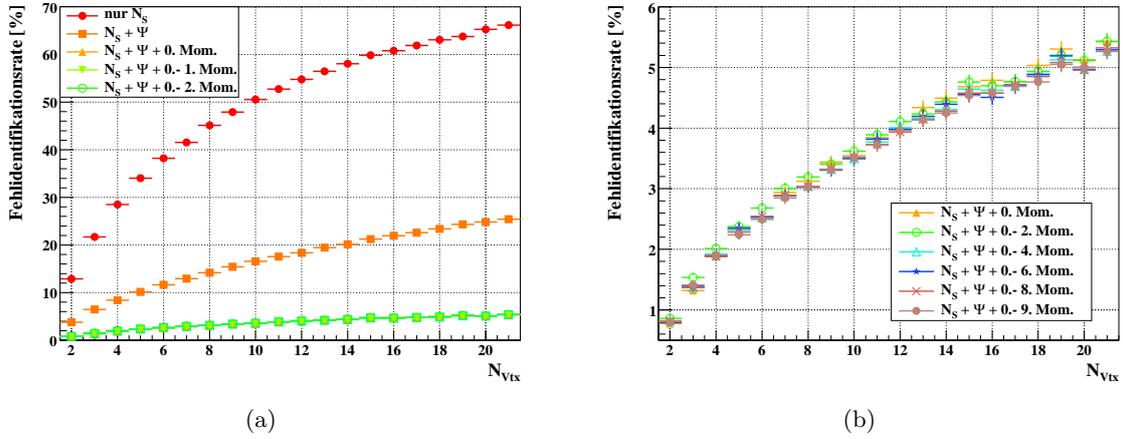
**Tab. 5.1.:** Korrelationskoeffizienten zwischen den statistischen Momenten  $\mu_i$  (für  $i = 0, \dots, 4$ ) der Verteilung transversalen Impulse der Spuren am Vertex.



**Abb. 5.21.:** Trainingsverlauf des neuronalen Netzes. Gezeigt ist der Fehler auf dem Test- und dem Trainingsdatensatz in Abhängigkeit der Anzahl an durchlaufenen Trainingsepochen.

führt. Zum Training wurde eine Netzkonfiguration mit einer versteckten Lage mit drei Knoten verwendet. Es wurde zunächst ein Netz mit lediglich der Spur-Multiplizität trainiert. In Abbildung 5.21 ist der Verlauf des Trainings durch die Fehler auf dem Test- und dem Trainingsdatensatz in Abhängigkeit der Anzahl an Trainingsepochen zu sehen. Der Fehler auf dem Testdatensatz erreicht nach 250 Trainingsepochen sein Minimum. Nach etwa weiteren 50 Epochen wird das Training abgebrochen und das Netz bei minimalem Fehler auf dem Testdatensatz wird verwendet.

Die Leistung dieses Netzes ist in Abbildung 5.22(a) zu sehen. In der Folge wurden Netze mit je einer weiteren Eingangsvariablen trainiert, wobei zunächst die Sphärizität gefolgt vom nullten bis neunten Moment hinzugefügt wurden. Abbildung 5.22(a) zeigt die Leistung der ersten fünf Netze im Vergleich. Eine deutliche Leistungssteigerung wird durch die Hinzunahme der Sphärizität und des ersten statistischen Moments erreicht. Für die weiteren Netze ist hier kein Unterschied zu sehen. Um die Unterschiede besser erkennen zu können, sind die Leistungen dieser und der weiteren Netze in Abbildung 5.22(b) gezeigt.

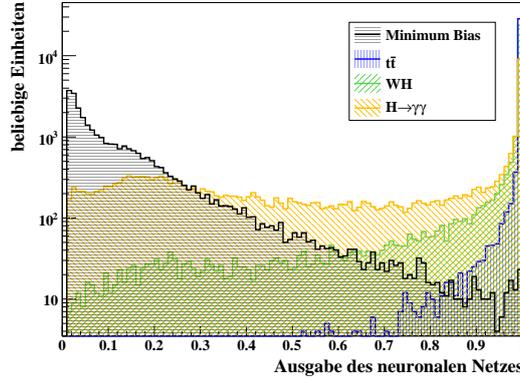


**Abb. 5.22.:** Fehlidentifikationsrate bestimmt auf  $WH$ -Ereignissen. Es wird die Leistung neuronaler Netze mit unterschiedlich vielen Eingangsvariablen verglichen.

Die Leistungszunahme ist sehr gering bei Verwendung weiterer statistischer Momente. Der Grund dafür, dass schon eine kleine Anzahl statistischer Momente ausreicht, ist auf die starken Korrelationen zurückzuführen (siehe Tabelle 5.1). Es ist ganz offensichtlich nicht notwendig, weitere Momente einzubinden.

Es ist von großem Interesse ein und dasselbe Selektionskriterium für alle Signalarten verwenden zu können. Aus diesem Grund wurde zusätzlich die Leistungsfähigkeit eines Netzes getestet, welches anstatt mit explizit einer Signalsorte mit einer Kombination aller hier betrachteten Signalereignisse trainiert wurde (in Abbildung 5.20 in violett). Hierfür wurden je 30 000 Ereignisse der verschiedenen Signaldatensätze gemischt und zusätzlich ein 90 000 Ereignisse umfassender *Minimum-Bias*-Datensatz verwendet. Als Eingangsvariablen wurden die ersten zehn statistischen Momente zusätzlich zur Spur-Multiplizität und zur Sphärizität berücksichtigt. Die Konfiguration des Netzes wurde analog zu den vorangegangenen gewählt: eine versteckten Lage mit drei Knoten. In Abbildung 5.23 ist die Verteilung des Ausgabewerts dieses neuronalen Netzes für sowohl *Minimum-Bias*-Ereignisse als auch für alle drei Signalereignistypen zu sehen.

Zum Vergleich ist in Abbildung 5.24(b) sowohl die Leistung des zuvor betrachteten neuronalen Netzes, welches spezifisch auf  $WH$ -Ereignissen trainiert wurde, als auch die Leistung eines Netzes, welches auf einem gemischten Signaldatensatz trainiert wurde zu sehen. Abbildungen 5.24(a) und 5.24(c) zeigen selbiges für  $t\bar{t}$ - und  $H \rightarrow \gamma\gamma$ -Ereignisse, wobei hier zum Training des signalspezifischen Netzes ein  $t\bar{t}$ - bzw. ein  $H \rightarrow \gamma\gamma$ -Signaldatensatz verwendet wurde. Es ist kein signifikanter Unterschied zwischen den Leistungen der signalspezifisch trainierten und des nicht signalspezifisch trainierten Netzes erkennbar. Für



**Abb. 5.23.:** Verteilung des Ausgabewerts des neuronalen Netzes (auf einem kombinierten Signaldatensatz trainiert) für *Minimum-Bias*-,  $t\bar{t}$ -,  $H \rightarrow \gamma\gamma$ - und  $WH$ -Ereignisse.

alle betrachteten Signalereignisklassen ist insofern die Verwendung eines nicht signal-spezifischen Netzes legitim. Es mag zunächst etwas verwunderlich sein, dass die auf den signal-spezifischen Datensätzen trainierten Netze nicht leistungsstärker sind, als die auf einem nicht spezifischen Signaldatensatz trainierten. Dass dies aber für die gegebenen Bedingungen durchaus nachvollziehbar ist, soll an einem kleinen Beispiel veranschaulicht werden.

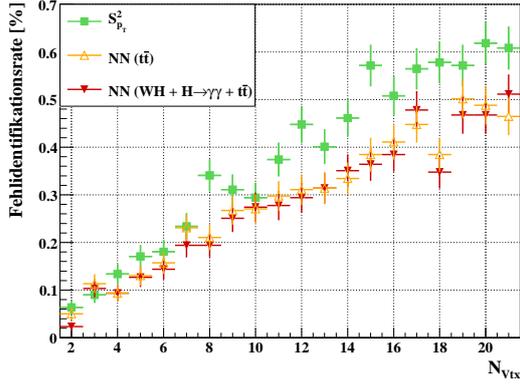
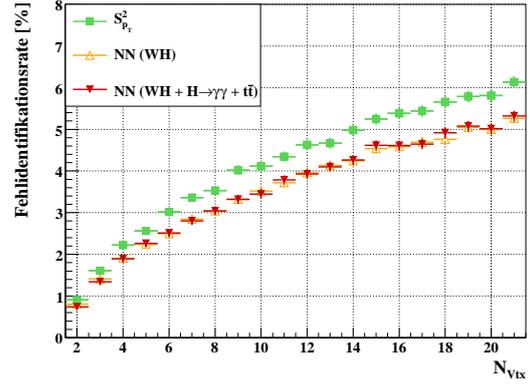
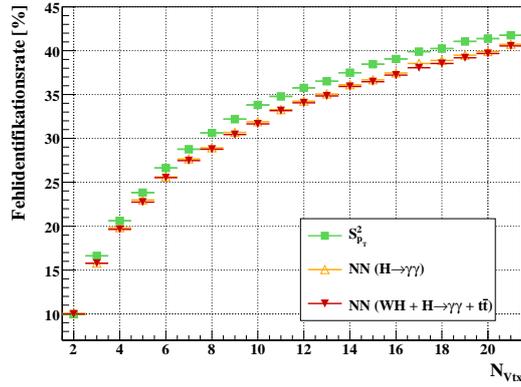
### **Einfluss der Eingangsverteilungen auf die Separationsfähigkeit eines neuronalen Netzes**

Angenommen es lägen zwei Signalsorten ( $S_1$ ,  $S_2$ ) vor, die von einem Untergrund anhand der Eigenschaft  $x$  getrennt werden sollen. Dabei soll Signal 1 in  $x$  leichter vom Untergrund trennbar sein als Signal 2. Zusätzlich soll gelten, dass der signalartige Bereich in  $x$  für beide Signalarten auf der gleichen Seite des untergrundartigen Bereichs liegt. In Abbildung 5.25(a) sind beispielhaft *Gauß*-Verteilungen für die Signale und den Untergrund gezeigt. Es ist möglich, ein neuronales Netz auf dem Datensatz von Signal 1 oder auf dem Datensatz von Signal 2 zu trainieren. Dabei berechnet das neuronale Netz die Wahrscheinlichkeit, dass bei einem bestimmten Wert der Eigenschaft  $x$  ein Ereignis zur Signalklasse  $S_1$  oder  $S_2$  gehört. Diese Wahrscheinlichkeit  $P(S_i|x)$  (für  $i = 1, 2$ ) lässt sich nach der *Bayes*-Regel berechnen:

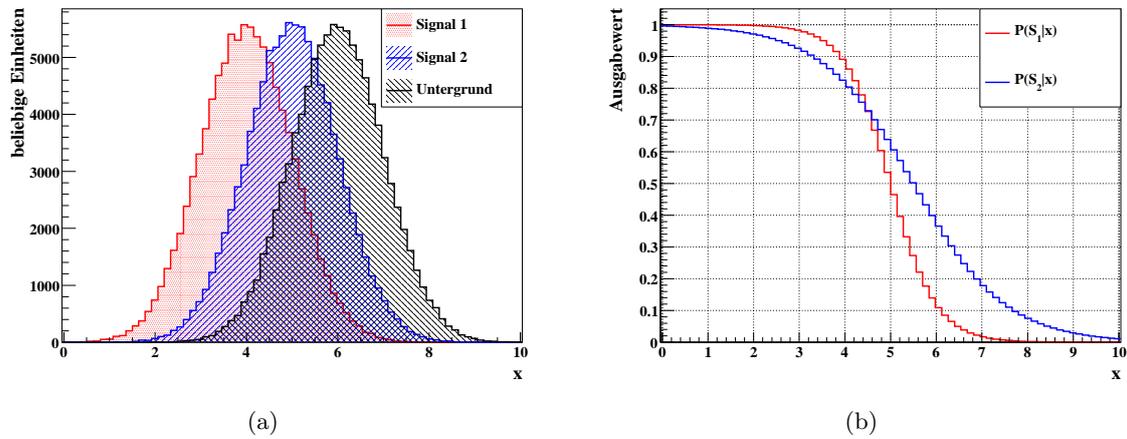
$$P(S_i|x) = \frac{P(x|S_i) \cdot P(S_i)}{P(x)} \quad (\text{für } i = 1, 2). \quad (5.33)$$

Dabei gilt folgendes für die hier verwendeten Größen:

- $P(x|S_i)$  ist die Wahrscheinlichkeit, dass ein Ereignis der  $i$ -ten Signalklasse die Eigenschaft  $x$  aufweist. Dies entspricht einfach dem Wert, der auf eine normierte Verteilung von  $S_i$  bei  $x$ .

(a)  $t\bar{t}$ (b)  $H \rightarrow \gamma\gamma$ (c)  $WH$ 

**Abb. 5.24.:** Fehlidentifikationsrate der Vertex-Selektion für (a)  $t\bar{t}$ - (b)  $WH$ - und (c)  $H \rightarrow \gamma\gamma$ -Ereignisse. Verglichen werden Methoden anhand eines signalspezifisch ( $NN(t\bar{t})$  bzw.  $NN(WH)$  bzw.  $NN(H \rightarrow \gamma\gamma)$ ) trainierten und eines nicht signalspezifisch ( $NN(WH + H \rightarrow \gamma\gamma + t\bar{t})$ ) trainierten neuronalen Netzes. Zusätzlich ist zum Vergleich die Leistung der Methode der Summe der quadratischen transversalen Impulse ( $S^2_{p_T}$ ) gezeigt.

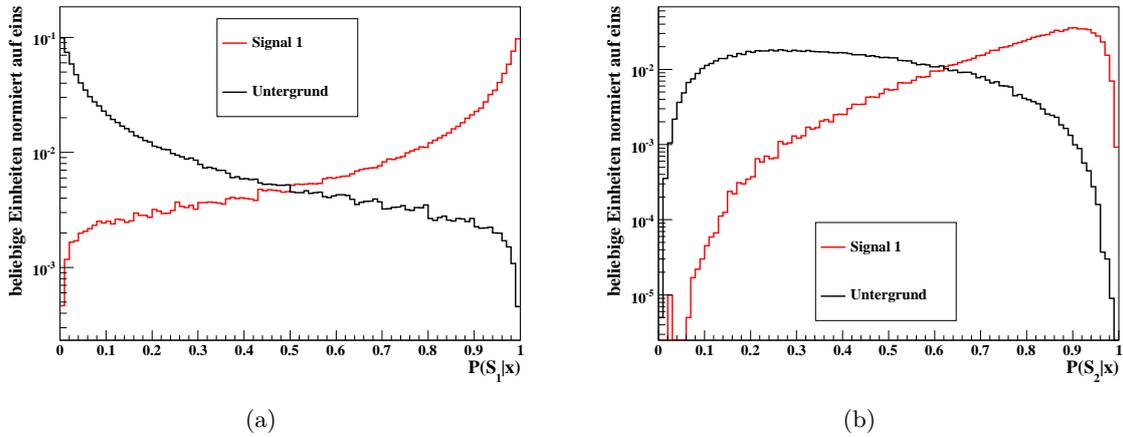


**Abb. 5.25.:** (a) Beispielhafte Verteilung einer Variablen  $x$  zweier Signalarten und eines Untergrunds; (b) Ausgabewert in Abhängigkeit von  $x$ . Der Ausgabewert entspricht der Wahrscheinlichkeit  $P(S_i|x)$  (mit  $i = 1, 2$ ), dass ein Ereignis mit der Eigenschaft  $x$  zur Klasse  $S_i$  gehört.

- $P(S_i)$  ist die a-priori Wahrscheinlichkeit, dass ein Ereignis der  $i$ -ten Signalklasse auftritt.
- $P(x)$  ist die Wahrscheinlichkeit, dass bei einem beliebigen Ereignis, unabhängig von der Ereignisklasse die Eigenschaft  $x$  gemessen wird. Dies entspricht  $P(S_i)P(x|S_i) + P(U)P(x|U)$ , wobei  $U$  für die Zugehörigkeit zu Untergrundklasse steht und im vorliegenden Fall  $P(S_1) = P(S_2) = P(U)$  gilt.

In Abbildung 5.25(a) ist der Ausgabewert  $P(S_i|x)$  in Abhängigkeit von  $x$  für die beiden unterschiedlichen zur Berechnung verwendeten Signaldatensätze gezeigt. Die Wahrscheinlichkeit  $P(S_1|x)$  entspricht der Diskriminanten eines mit Signal 1 trainierten Netzes. Die Wahrscheinlichkeit  $P(S_2|x)$  entspricht dagegen der Diskriminanten eines mit Signal 2 trainierten Netzes. Beide können zur Trennung von Signal und Untergrund angewendet werden. Im Folgenden wird die Trennleistung der beiden für  $S_1$ -Ereignisse untersucht.

Abbildung 5.26 zeigt die Verteilungen von  $P(S_1|x)$  (5.26(a)) und  $P(S_2|x)$  (5.26(b)) für Signalereignisse der Klasse  $S_1$  und für Untergrundergebnisse. Die Verteilungen unterscheiden sich grundsätzlich für die verschiedenen Kriterien. Die Separationsfähigkeit der jeweiligen Diskriminanten ist anhand dieser Abbildungen nicht erkennbar. Hierfür muss für verschiedene mögliche Schnittbedingungen betrachtet werden, wie viel Untergrundergebnisse man bei solch einer Schnittbedingung fälschlicherweise als Signalereignis erkannt hat. Eine Schnittbedingung ist ein Wert  $P_{Schnitt}$ , so dass Ereignisse mit einem gemessenen Wert von  $x'$  mit  $P(S_i|x') > P_{Schnitt}$  als signalartig eingestuft werden. Jede Schnittbedin-



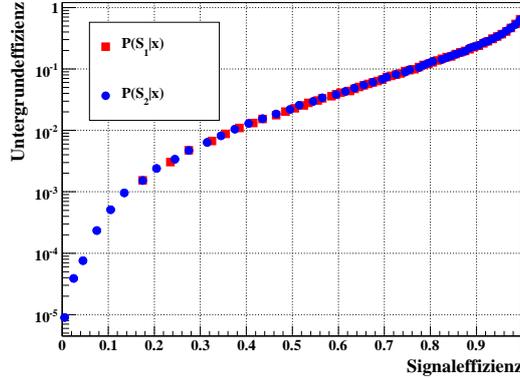
**Abb. 5.26.:** Verteilung der Wahrscheinlichkeit (a)  $P(S_1|x)$ , dass die gemessene Eigenschaft  $x$  zu einem Ereignis der Klasse  $S_1$  gehört bzw. (b)  $P(S_2|x)$ , dass sie zu einem Ereignis der Klasse  $S_2$  gehört für Ereignisse der Untergrundklasse und der Signalklasse  $S_1$ .

gung entspricht einer Signal- und einer Untergrundeffizienz  $\epsilon_{S,U}$ :

$$\epsilon_{S_j,U}^i(P_{Schnitt}) = \frac{\text{Anzahl aller Ereignisse der Klasse } S_j \text{ bzw. } U \text{ mit } P(S_i|x) > P_{Schnitt}}{\text{Anzahl aller Ereignisse der Klasse } S_j \text{ bzw. } U}.$$

Der Index  $i$  steht hier für die zum Training verwendete Signalart und  $j$  für die Signalart, auf der die Separationsleistung getestet wird. Trägt man die Untergrund- gegen die Signaleffizienz auf, kann die Leistungsfähigkeit der Trennung verglichen werden. Je kleiner die Untergrundeffizienz bei gleicher Signaleffizienz, desto höher die Separationsleistung. In Abbildung 5.27 ist ein solcher Leistungstest für die beiden, auf unterschiedlichen Signalerignisklassen trainierten Diskriminanten gezeigt. Im vorliegenden Fall ist die Trennleistung gänzlich unabhängig von dem zum Training verwendeten Datensatz. Die Punkte gleicher Signaleffizienz entsprechen dabei sehr unterschiedlichen Schnittbedingungen in  $P'(S_1|x)$  und  $P'(S_2|x)$  bzw. für eine gleiche Schnittbedingung  $P_{Schnitt}$  ergeben sich unterschiedliche Signaleffizienzen:  $\epsilon_{S_1}^1(P_{Schnitt}) \neq \epsilon_{S_1}^2(P_{Schnitt})$ . Dennoch liegen alle Ereignisse die in Abbildung 5.26(a) rechts von  $P'(S_1|x)$  liegen in Abbildung 5.26(b) rechts von  $P'(S_2|x)$ .

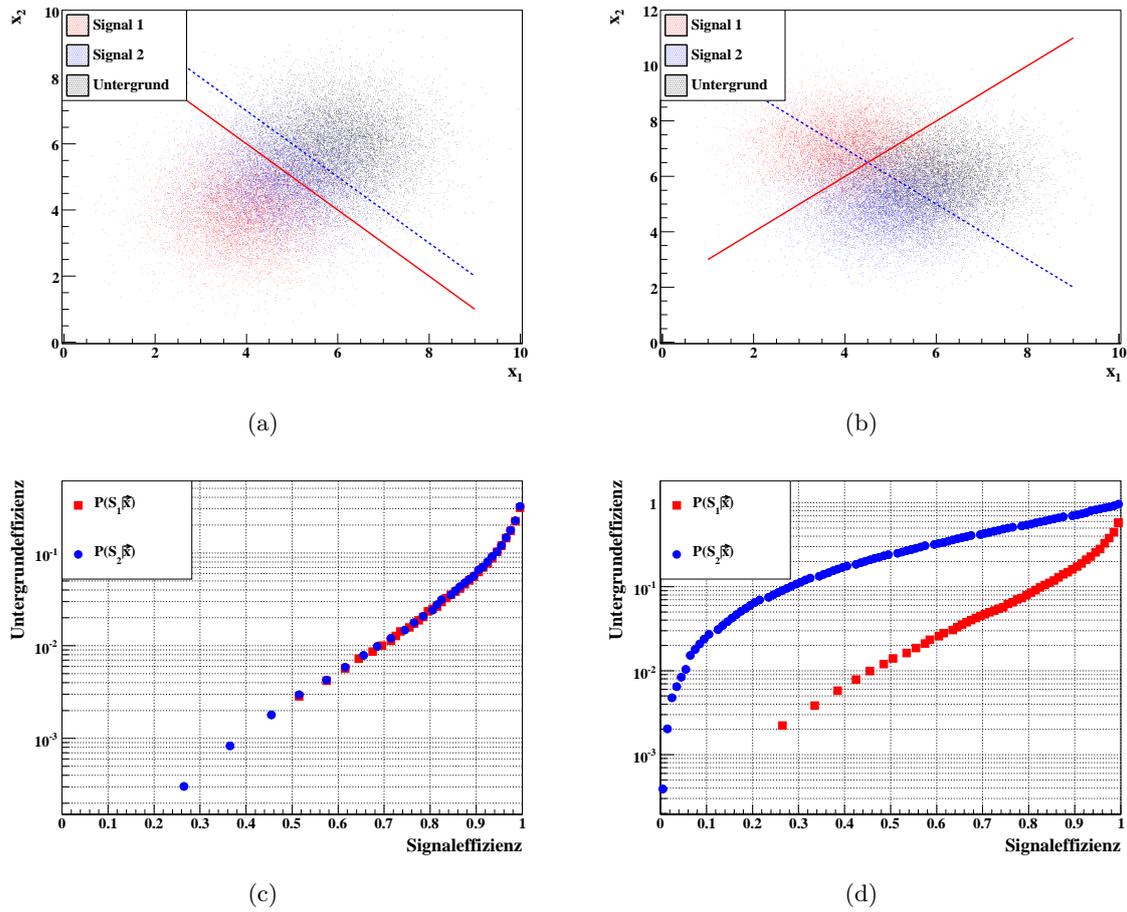
Die hier für den eindimensionalen Fall gemachten Aussagen gelten bei bestimmte Gegebenheiten auch für mehrdimensionale Probleme. Abbildung 5.28(a) zeigt die Verteilung zweier Signal- und einer Untergrundklasse in zwei Eigenschaften  $x_1$  und  $x_2$ . Wie oben wurden hier wieder *Gauß*-verteilte Ereignisse angenommen. Zusätzlich ist die Schnittgerade bei  $P(S_i|\vec{x}) = 0.5$  (für  $i = 1, 2$  und  $\vec{x} = (x_1, x_2)$ ) eingezeichnet. Je nach Trainingsdatensatz  $S_i$ , liegt die gerade genau zwischen der Verteilung des Untergrunds und Signal 1 bzw. Signal 2. Die Schnittgeraden sind für den gezeichneten Fall parallel. Um die Untergrunde-



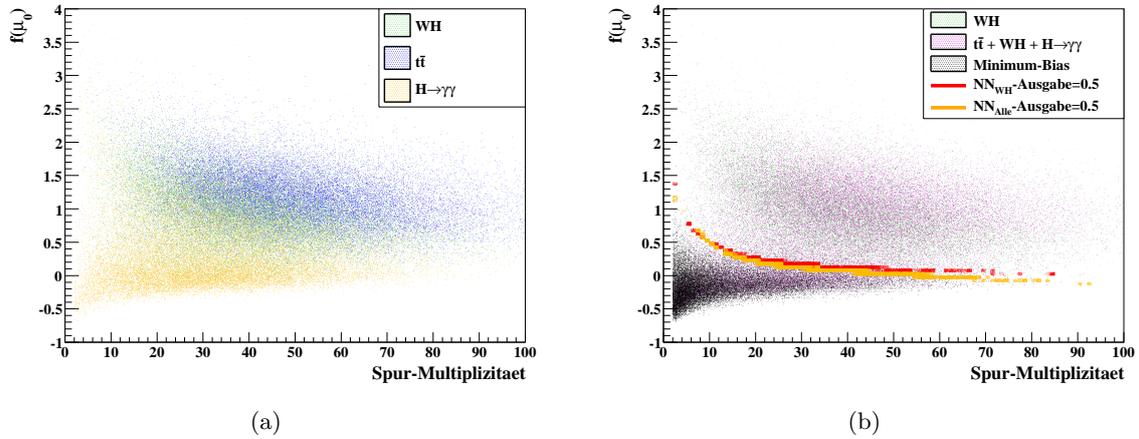
**Abb. 5.27.:** Separationsleistung zwischen Ereignissen der Klasse  $S_1$  und der Untergrundeignisklasse. Gezeigt ist die Untergrundeffizienz in Abhängigkeit der Signaleffizienz für ein Training anhand von Signal 1 ( $P(S_1|x)$ ) und anhand von Signal 2 ( $P(S_2|x)$ ).

fizienz in Abhängigkeit der Signaleffizienz zu berechnen, werden diese Geraden lediglich verschoben. Die Separationsleistung ist unabhängig vom zum Training verwendeten Signaldatensatz. Dies ist deutlich im Leistungsvergleich in Abbildung 5.28(c) zu sehen. Hier wurde wieder die Separationsfähigkeit zwischen Signal 1 und dem Untergrund der beiden Diskriminanten getestet. Abbildung 5.28(b) zeigt dagegen Verteilungen der Datensätze in  $x_1$  und  $x_2$ , für die der zum Training gewählte Signaldatensatz relevant ist. In einem solchen Fall liegen die beiden Schnittgeraden nicht parallel. Verwendung der Geraden bestimmt durch  $P(S_1|\vec{x})$  wird auf Ereignissen von Signal 2 zu einer schlechteren Leistung als Verwendung der Geraden bestimmt durch  $P(S_2|\vec{x})$  führen. Im Leistungsvergleich in Abbildung 5.28(d) ist dies offensichtlich.

Im Falle einer Verteilung wie sie Abbildung 5.28(a) zu sehen ist, ist der Trainingsdatensatz irrelevant im Falle einer Verteilung wie in Abbildung 5.28(b) aber ist er relevant. Der Grund hierfür ist, dass die Reihenfolge des Auftretens für Signal 1, Signal 2 und Untergrund im ersten Fall in beiden Eigenschaften  $x_1$  und  $x_2$  gleich ist (rot-blau-schwarz in  $x_1$  und  $x_2$ ), während im zweiten Fall die Abfolge der Ereignisklassen unterschiedlich ist (rot-blau-schwarz für  $x_1$  und blau-schwarz-rot für  $x_2$ ). Im vorliegenden Problem einer Signalvertex-Selektion ist diese Reihenfolge der verschiedenen Signalklassen und des Untergrunds für alle betrachteten Eigenschaften gleich (die Sphärizität ist hier eine Ausnahme). In allen Eigenschaften lässt sich ein  $t\bar{t}$ -Ereignis am leichtesten und ein  $H \rightarrow \gamma\gamma$ -Ereignis am schwersten von einem *Minimum-Bias*-Ereignis unterscheiden. Ein gemischter Signaldatensatz hat aus diesem Grund nur eine Verschiebung der Verteilung im Falle von  $t\bar{t}$  zum Untergrund hin, im Falle von  $H \rightarrow \gamma\gamma$  vom Untergrund weg zur Folge, was wie in Abbildung 5.28(a) und 5.28(c) zu sehen ist, keinen Unterschied der Trennleistung verursacht. In Abbildung 5.29 ist zur Veranschaulichung die Verteilung von  $f(\mu_0)$  in Abhängigkeit



**Abb. 5.28.:** Zweidimensionale Verteilung zweier Signalereignisklassen und einer Untergrundereignisklasse für zwei Eigenschaften  $x_1, x_2$ . Im Falle von Abbildung (a) sind die Schnittgeraden, die anhand der verschiedenen Signaldatensätze bestimmt werden parallel. (Die Schnittgerade trainiert auf  $S_1$  ist die rote, volle Linie und die Schnittgerade trainiert auf  $S_2$  ist die blaue, gestrichelte Linie.) Der Trainingsdatensatz ist in einem solchen Fall für die Trennleistung irrelevant. Abbildung (c) zeigt hierzu die Untergrundeffizienz in Abhängigkeit der Signaleffizienz. Im Falle von (b) sind die Schnittgeraden nicht parallel. Aus diesem Grund ergibt sich ein Leistungsunterschied je nach Trainingsdatensatz. Abbildung (d) zeigt den Leistungsvergleich für diesen Fall.



**Abb. 5.29.:** Verteilung der Transformation  $f(\mu_0)$  des Mittelwerts der Verteilung der transversalen Impulse in Abhängigkeit von der Spur-Multiplizität für die verschiedenen Signalsorten (a) und für *Minimum-Bias*, den gemischten Signaldatensatz und *WH* (b). In Abbildung (b) sind zusätzlich die Ereignisse farbig markiert, die von einem signalspezifisch trainierten ( $NN_{WH}$ ) und einem nicht signalspezifisch trainiertem Netz ( $NN_{Alle}$ ) eine Ausgabewert von  $\approx 0.5$  zugeordnet bekommen.

von der Spur-Multiplizität gezeigt. Abbildung 5.29(a) vergleicht die drei verschiedenen verwendeten Signalarten. In Abbildung 5.29(b) ist ein Vergleich zwischen  $t\bar{t}$ , *Minimum-Bias* und dem gemischten Signaldatensatz gezeigt. Die von einem signalspezifischen und einem nicht signalspezifischen Netz bestimmten Schnittlinien sind in Abbildung 5.29(b) zusätzlich eingezeichnet. Hier sind alle Ereignisse farbig markiert, die einen Ausgabewert  $F(N_{Sp}, \mu_0) \approx 0.5$  vom jeweiligen neuronalen Netz zugeordnet bekommen. Die Linien sind sehr ähnlich, wohlgeordnet nicht identisch und nicht ganz parallel, aber dennoch in ausreichendem Maße äquivalent, dass die ähnliche Leistungsfähigkeit der spezifischen und nicht spezifischen Netze nachvollziehbar ist. Zum Abschluss der hier gemachten Überlegungen bezüglich des Einflusses der Eingangsverteilungen auf die Separationsleistung eines neuronalen Netzes sei angemerkt, dass sich gleiche Überlegung auch für die Likelihood-Methode anstellen lassen, da die Wahrscheinlichkeit nach der *Bayes-Regel*, gerade einem Likelihood-Verhältnis entspricht.

In den Abbildungen 5.24 ist zusätzlich zum Vergleich die Leistung der  $\mathcal{S}_{p_T}^2$ -Methode zu sehen. Die neuronalen Netze erreichen eine Leistungssteigerung gegenüber dieser Methode. Im Prinzip verwendet das neuronale Netz sehr ähnliche Information wie die  $\mathcal{S}_{p_T}^2$ -Methode. Ein Netz, welches als Eingangsvariablen lediglich die Spur-Multiplizität und den Mittelwert der Verteilung der transversalen Impulse verwendet, ist faktisch äquivalent zur  $\mathcal{S}_{p_T}$ -Methode, also zur einfachen Summe der transversalen Impulse. Um dies vergleichen zu können ist in Abbildung 5.30(a) die Leistungsfähigkeit eines solchen Netzes und die der

	$t\bar{t}$	$WH$	$H \rightarrow \gamma\gamma$	$t\bar{t} + H \rightarrow \gamma\gamma + t\bar{t}$	<i>Minimum-Bias</i>
$\rho(N_{Sp}, f(\mu_0))$	-0.31	-0.28	-0.02	0.03	0.22

**Tab. 5.2.:** Linearer Korrelationskoeffizient  $\rho(N_{Sp}, (\mu_0))$  der Korrelation zwischen dem transformierten Mittelwert  $f(\mu_0)$  der Verteilung der transversalen Impulse und der Spur-Multiplizität  $N_{Sp}$ . Korrelationen zu Abbildung 5.29.

$\mathcal{S}_{p_T}$ -Methode gezeigt. Das neuronale Netz gewinnt deutlich an Leistungsfähigkeit gegenüber der  $\mathcal{S}_{p_T}$ -Methode. Alleine die Korrelationen zwischen der Spur-Multiplizität und dem Mittelwert stehen dem Netz explizit als zusätzliche Information zur Verfügung. Diese sind allerdings, wie in Abbildung 5.29 zu sehen ist, eher klein. Die Korrelationskoeffizienten sind in Tabelle 5.2 aufgeführt. Während  $t\bar{t}$ - und  $WH$ -Ereignisse leicht negativ in diesen Größen korreliert sind, ist im Falle der  $H \rightarrow \gamma\gamma$ -Ereignisse und der Mischung der Signale fast gar keine Korrelation zu beobachten. Die Korrelation der *Minimum-Bias*-Ereignisse ist leicht positiv. Die gute Leistung des Netzes im Vergleich zur  $\mathcal{S}_{p_T}$ -Methode ist einleuchtend, wenn man berücksichtigt, dass das Netz eine beliebige Funktion aus dem Mittelwert und der Spur-Multiplizität berechnen kann, während  $\mathcal{S}_{p_T}$  immer eine Multiplikation vornimmt.

In einer zweite Abbildung 5.30(b) ist zusätzlich zu beobachten, dass die Verwendung der Sphärizität nicht in einem Maße ausschlaggebend ist, wie man nach Abbildung 5.22(a) hätte glauben können. Vergleicht man die Leistungsfähigkeit eines Netzes, welches als Eingangsvariablen nur die Spur-Multiplizität und den Mittelwert  $\mu_0$  verwendet, mit der eines Netzes, welches zusätzlich die Information der Sphärizität ausnutzt, zeigt sich, dass nur eine minimale Leistungssteigerung durch die Sphärizität erreicht wird. Beim Vergleich mit der Leistung des Netzes trainiert auf allen Eingangsvariablen ist auch zu erkennen, dass im Prinzip deutlich weniger Momente zum Training ausreichen würden. Allerdings nimmt das Netz auch keinen Schaden an überflüssigen Eingangsvariablen. Fraglich ist, ob es nicht möglich wäre die Verwendung eines neuronalen Netzes zu umgehen und an Stelle dessen eine einfachere Diskriminante aus den vorliegenden Informationen zu bestimmen. Dies wurde im Rahmen dieser Arbeit aber nicht weiter untersucht.

Es bleibt als Ergebnis festzuhalten, das anhand eines neuronalen Netzes ein Leistungszugewinn gegenüber der  $\mathcal{S}_{p_T}^2$ -Methode erreicht wird.

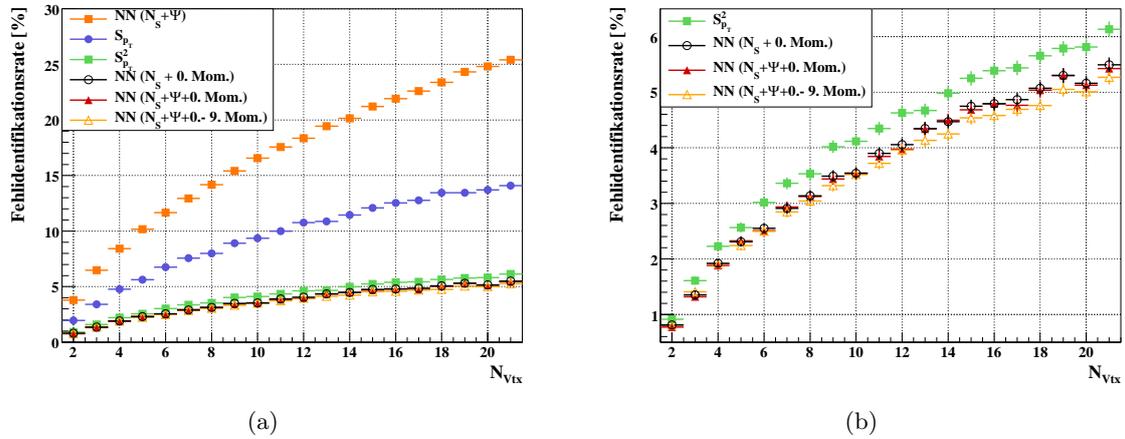


Abb. 5.30.: Fehlidentifikation von  $WH$ -Ereignissen. Es wird die Leistung neuronaler Netze mit unterschiedlichen Eingangsvariablen mit der Leistung der  $S_{p_T}$ - und der  $S_{p_T}^2$ -Methode verglichen.

## 5.5. Zusammenfassung

Abschließend sollen kurz die Ergebnisse dieses Kapitels zusammengefasst werden. Zur Signalvertex-Selektion stehen folgende Vertex- und Spureigenschaften zur Verfügung: Spur-Multiplizität, transversaler Impuls aller Spur am Vertex und die Sphärizität. Es wurden insgesamt vier Methoden zur Vertex-Selektion getestet, die auf diesen Eigenschaften basieren. Dabei zeigte die Vertex-Wahrscheinlichkeit und die Likelihood-Methode eine schlechtere Leistung als erwartet, was auf Korrelationen zwischen den transversalen Impulsen der Spuren am Vertex zurückgeführt werden konnte. Die  $S_{p_T}^2$ -Methode ist eine sehr gute Methode für den Beginn der Datenahme, da sie einfach ist und trotz allem hohe Selektionseffizienzen auf allen betrachteten Signalprozessen erreicht. Der Leistungstest einer Methode basierend auf einem neuronalen Netzes zeigte für alle getesteten Signalarten die beste Leistung. Aufgrund der direkten Abhängigkeit von den Verteilungen der Eingangsvariablen wird eine solche Methode erst im weiteren Verlauf der LHC-Datennahme Anwendung finden, da man im Falle eines neuronalen Netzes mit eventuellen Leistungsabweichungen zu rechnen hat, wenn Diskrepanzen zwischen der Simulation und der Realität auftreten. Auch im Fall der  $S_{p_T}^2$ -Methode, die ebenfalls auf simulierten Daten getestet wurde muss in einem solchen Fall mit einer Leistungsabweichung gerechnet werden, es ist aber zu erwarten, dass eine einfache Methode dieser Art robuster gegen solche Effekte ist.

## 6. Signalvertex-Selektion mit einem offiziellen Rekonstruktionswerkzeug

Die im vorangegangenen Kapitel getesteten Methoden sollen innerhalb der allgemein verwendeten Software bei ATLAS zur Verfügung stehen. Zu diesem Zwecke wurde eigens zur Selektion des Signalvertex ein Werkzeug implementiert. Es soll hier zunächst kurz die bei ATLAS verwendete Software beschrieben werden. In der Folge werden dann etwas ausführlicher das neue Werkzeug und seine Vorteile vorgestellt. Dabei wird die Leistung der Selektionsmethoden dieses Werkzeugs mit den Ergebnissen der Studien aus dem vorigen Kapitel verglichen. Innerhalb dieser Leistungstests ist es nun auch möglich, den zuvor vernachlässigten Einfluss des Überlapps der Vertices bei der Primärvertex-Rekonstruktion zu berücksichtigen.

### 6.1. ATHENA: die Software des ATLAS-Detektors

ATHENA ist die Software-Umgebung des ATLAS-Experiments. Hier sind alle gemeinschaftlich genutzten Werkzeuge für die Ereignis- und Detektorsimulation, die Ereignisrekonstruktion und für die physikalischen Analysen implementiert. Die einzelnen Werkzeuge sind in einer C++-Klassenstruktur in ATHENA angeordnet. Die von ATHENA ausgeführten Simulations-, Rekonstruktions- oder Analyseaufgaben können vom Benutzer durch PYTHON-Skripte gesteuert werden. Dabei ist ein sehr flexibles Ein- und Ausschalten der verschiedenen, zur Verfügung stehenden Werkzeuge möglich. Für die Implementierung eines Werkzeugs existieren allgemeine strukturelle Vorgaben. Alle Werkzeuge besitzen abstrakte Schnittstellen, genannt *Interface-Klassen*. Diese definieren Funktionalitäten, wobei innerhalb der Interface-Klasse nicht weiter spezifiziert wird, wie diese ausgeführt werden. Dies wird erst in einer expliziten Implementierung festgelegt. Dabei können mehrere Implementierungen die gleiche Funktionalität der Interface-Klasse auf sehr unterschiedliche Weise ausführen. Die Funktionalität wird also durch die Interface-Klasse festgelegt, die Art der Ausführung durch die Implementierungen. In der Folge wird anhand des neuen Werkzeugs zur Vertex-Selektion ein Beispiel für diese Struktur gegeben.

## 6.2. ATHENA-Werkzeug zur Selektion des Signalvertex

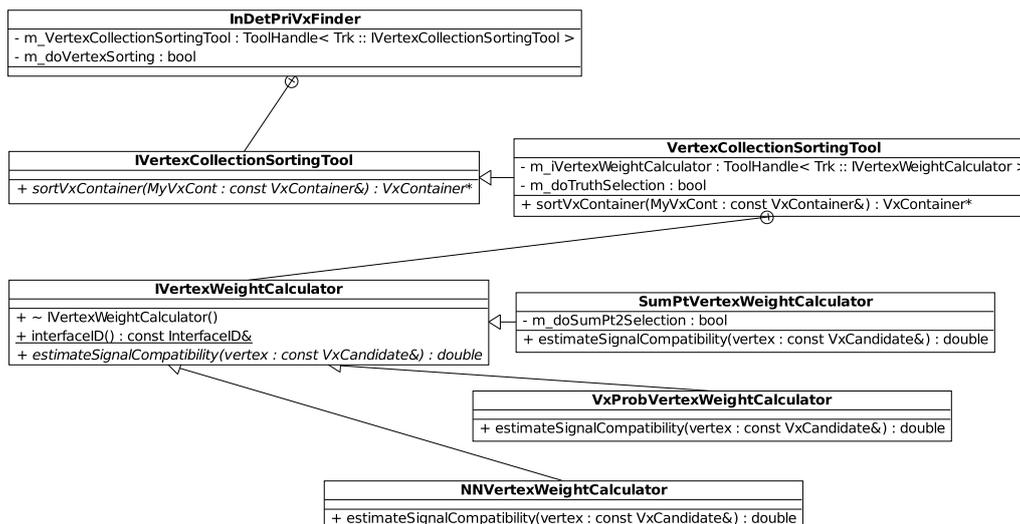
Ein Werkzeug eigens zur Signalvertex-Selektion hat den Vorteil, dass eine von der Primärvertex-Rekonstruktion unabhängige Wahl der Methode zur Signalvertex-Selektion vorgenommen werden kann. Wie bereits in Kapitel 4 angesprochen, wurde vor der Implementierung des in der Folge beschriebenen Werkzeugs die Selektion des Signalvertex innerhalb der verschiedenen Primärvertex-Rekonstruktionsalgorithmen durchgeführt. Dabei wurden unterschiedliche Selektionskriterien innerhalb der unterschiedlichen Primärvertex-Rekonstruktionsalgorithmen verwendet und die Wahl des Selektionskriteriums wurde durch die Wahl des Rekonstruktionsalgorithmus festgelegt. Am Ende der Primärvertex-Rekonstruktion steht eine Kollektion rekonstruierter Signalvertex-Kandidaten zu Verfügung. Das neue Signalvertex-Selektionswerkzeug übernimmt diese Vertex-Kollektion und berechnet für jeden Vertex ein Gewicht, welches ein Maß dafür angibt, wie wahrscheinlich es ist, dass es sich um den Signalvertex handelt. Die Vertex-Kollektion wird gemäß dieser Gewichte so umsortiert, dass der Vertex mit der größten Wahrscheinlichkeit, der Signalvertex zu sein, an erster Stelle steht. Die umsortierte Vertex-Kollektion wird nun zur weiteren Verwendung zur Verfügung gestellt. Das neue Werkzeug wird so einfach in den bisherigen Ablauf integriert, ohne dabei die vorherige Struktur zu verändern. Es soll nun genauer der Aufbau dieses Werkzeugs beschrieben werden.

Der im vorigen Abschnitt beschriebenen Struktur folgend sind zwei abstrakte Interface-Klassen neu in ATHENA implementiert worden, der *IVertexWeightCalculator* und das *IVertexCollectionSortingTool*. In zwei Abbildungen ist die Struktur des neuen Werkzeugs mit diesen Interface-Klassen veranschaulicht, dabei zeigt Abbildung 6.2 die Funktionalität der Interface-Klassen und Abbildung 6.1 die Klassenstruktur. Die Primärvertex-Rekonstruktion, bzw. der *InDetPriVxFinder*, ruft nach der Bestimmung einer Kollektion rekonstruierter Vertizes (*VxContainer*) das neue Werkzeug auf. Hierfür wird ein Objekt des *VertexCollectionSortingTools* erzeugt. Das *IVertexCollectionSortingTool* besitzt die Funktionalität *sortVxContainer*, die die Vertex-Kollektion neu sortiert. Die Sortiermethode ist in der Implementierung *VertexCollectionSortingTool* definiert (siehe Abbildung 6.2). Diese Implementierung erzeugt zur Berechnung der Gewichte, gemäß denen die Sortierung vorgenommen werden soll, ein abgeleitetes Objekt des *IVertexWeightCalculator*. Die Funktionalität des *IVertexWeightCalculators* ist die Bestimmung eines Gewichts des Vertex (*VxCandidate*), welches ein Maß für die Kompatibilität mit der Signalthypothese angibt (*estimateSignalCompatibility*). Drei Methoden zur Berechnung des Gewichts sind derzeit implementiert:

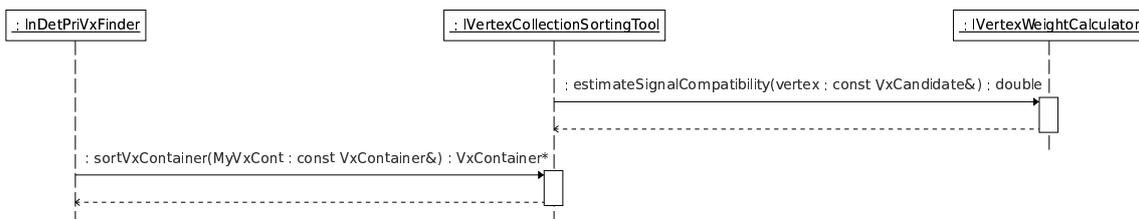
- *SumPtVertexWeightCalculator* entspricht der  $S_{p_T}^2$ -Methode, wobei wahlweise auch die  $S_{p_T}$ -Methode verwendet werden kann (`m_doSumPt2Selection=false`).

- *VxProbVertexWeightCalculator* entspricht der  $\mathcal{P}_{Vtx}$ -Methode.
- *NNVertexWeightCalculator* berechnet die Gewichte anhand des neuronalen Netzes, welches auf dem gemischten Signaldatensatz ( $WH + t\bar{t} + H \rightarrow \gamma\gamma$ ) trainiert wurde.

Für jeden Vertex der Kollektion wird ein Gewicht an das *IVertexCollectionSortingTool* gegeben, die Vertex-Kollektion wird von diesem gemäß der Größe der Gewichte in absteigender Ordnung neu sortiert und der *InDetPriVxFinder* erhält die neu sortierte Vertex-Kollektion zurück (Abb. 6.2). Die Methode zur Bestimmung der Gewichte und damit zur Identifikation des Signalvertex kann von außen durch den Benutzer in Skripten zur Steuerung des Vertex-Rekonstruktionsauftrags gewählt werden. Der Benutzer hat so die Möglichkeit, die gewünschte Rekonstruktions- und Selektionsmethode voneinander unabhängig auszuwählen. Das Selektionswerkzeug ist ab ATHENA-Version 15.0.0 zugänglich.



**Abb. 6.1.:** Klassendiagramm des Werkzeugs zur Sortierung der Vertex-Kollektion (*IVertexCollectionSortingTool*) und des Werkzeugs zur Berechnung eines Gewichts (*IVertexWeightCalculator*). Das Gewicht ist ein Maß für die Kompatibilität des Vertex mit der Signalhypothese.



**Abb. 6.2.:** Zusammenspiel der Interface-Klassen zur Signalvertex-Selektion über ihre Funktionalität.

### 6.3. Leistungstest der in ATHENA implementierten Selektionsmethoden

Innerhalb dieses Abschnitts sollen schließlich die implementierten Methoden auf ihre Selektionsleistung hin untersucht werden. Hierfür wurden Datensätze mit *simuliertem Pile-Up* verwendet. Die Leistungen sollen hier mit denen aus Kapitel 5 verglichen werden, die auf *selbstgemischtem Pile-Up* erhalten worden sind. Dabei ist Folgendes unter diesen *Pile-Up*-Spezifikationen zu verstehen:

- *Selbstgemischtes Pile-Up*: Hiermit sind Datensätze gemeint, bei denen die *Minimum-Bias*-Ereignisse gemäß des in Abschnitt 5.2 beschriebenen Rezepts beigemischt werden. Diese Methode vernachlässigt den Überlapp der Vertizes bei der Spur- und Vertex-Rekonstruktion.
- *Simuliertes Pile-Up*: Dies entspricht Datensätzen, bei denen die *Pile-Up*-Wechselwirkungen vor der Ereignisrekonstruktion, im Speziellen vor der Spur- und Primärvertex-Rekonstruktion, beigemischt werden. Somit werden Effekte durch den Überlapp der Wechselwirkungen bei der Rekonstruktion berücksichtigt. Es wird erwartet, dass die Einflüsse des Überlapps auf die Primärvertex-Rekonstruktion auch einen Einfluss auf die Selektion des Signalvertex haben.

Um die in ATHENA implementierten Methoden zu testen, wurden alle drei bisher betrachteten Signalprozesse auf der Basis von simuliertem *Pile-Up* für zwei verschiedene *Pile-Up*-Szenarien auf ihre Selektionsleistung hin untersucht. Die beiden betrachteten *Pile-Up*-Szenarien wurden bereits in Kapitel 3 (siehe Tabelle 3.1) vorgestellt. Die Ereignisse wurden mit im Mittel 4.1 und 6.9 *Minimum-Bias*-Vertizes zusätzlich zum Signalvertex überlagert. Für die drei Ereignisklassen  $WH$ ,  $t\bar{t}$  und  $H \rightarrow \gamma\gamma$  sind die Fehlidentifikationsraten aller implementierten Selektionsmethoden für die beiden *Pile-Up*-Szenarien in Tabelle 6.1 zu sehen. Zusätzlich zu den drei implementierten Methoden wurde die Methode betrachtet, die verwendet wird wenn vom Benutzer keine Neusortierung verlangt wird. Die Sortierung wird dann von dem verwendeten Vertex-Rekonstruktionsalgorithmus vorgenommen. Für die Standard-Vertex-Rekonstruktion bei ATLAS mit dem *AdaptiveMultiVertexFinder* ist das Selektions- bzw. Sortierkriterium:  $\sum p_T^2 \sqrt{N_{Sp}}$  (siehe Abschnitt 4.1.2). Der Vergleich zwischen den implementierten Selektionsmethoden zeigt das aus den Studien mit selbstgemischtem *Pile-Up* erwartete Verhalten. Auch bei diesen Ereignissen erreicht das neuronale Netz, gefolgt von der  $\mathcal{S}_{p_T}^2$ -Methode die kleinste Fehlidentifikationsrate. Die Vertex-Wahrscheinlichkeit weist die höchste Fehlidentifikationsrate auf. Das neuronale Netz und die  $\mathcal{S}_{p_T}^2$ -Methode erreichen gegenüber dem Kriterium des *AdaptiveMultiVertexFinder* eine deutliche Verbesserung der Selektion.

Es sollen an dieser Stelle Vergleiche mit den in Kapitel 5 erhaltenen Ergebnissen gemacht werden. Es gilt Folgendes dabei zu berücksichtigen: Die vorliegenden Ereignisse mit simuliertem *Pile-Up* sind mit im Mittel 4.1 bzw. 6.9 *Minimum-Bias*-Ereignissen zusätzlich zum Signalereignis simuliert. Die Primärvertex-Rekonstruktion erhält insgesamt im Mittel 5.1 bzw. 7.9 simulierte Vertizes, wobei nach Abbildung 4.2 davon ausgegangen werden kann, dass ca. ein *Minimum-Bias*-Vertex bei der Rekonstruktion verloren geht. Die Signalvertex-Selektion erhält demnach im Mittel nur  $\approx 4.1$  bzw.  $\approx 6.9$  Vertizes, wovon einer der Signalvertex ist. Dies ist aus zwei Gründen eine Näherung: Zum Einen wird nicht immer genau ein Vertex verloren gehen, zum Anderen werden vor allem *Minimum-Bias*-Vertizes mit sehr schwachen Spuren nicht rekonstruiert, die für die Signalvertex-Selektion keine große Schwierigkeit darstellen. Desweiteren ist bei dem Vergleich zu berücksichtigen, dass sich die Vertex-Multiplizitäten gemäß einer *Poisson*-Verteilung um die oben genannten Mittelwerte verteilen. Die Raten müssen für diese mittleren, *Poisson*-verteilten Vertex-Multiplizitäten mit dem selbstgemischtem *Pile-Up* verglichen werden. In Tabelle 6.2 sind die Raten dieses Vergleichsfalls angegeben. Die Fehlidentifikationsraten sind in allen Fällen kleiner als bei den mit *Pile-Up* simulierten Ereignissen (Tab. 6.1). Der Leistungsabfall ist auf den Überlapp der Vertizes bei der Rekonstruktion zurückzuführen: Liegen beispielsweise mehrere *Minimum-Bias*-Wechselwirkungen sehr nahe beieinander, ist es möglich, dass diese als ein einzelner Vertex rekonstruiert werden. Die Auflösung für den Signalvertex in z-Richtung wird mit  $\sigma_z \approx 0.04$  mm (siehe Tabelle 4.3) angegeben. Es ist davon auszugehen, dass die Auflösung für *Minimum-Bias*-Vertizes schlechter ist, da Teilchen die von diesen Vertizes ausgehen tendenziell weniger Impuls tragen und aus diesem Grund mehr Vielfachstreuung auftritt. Nach den Studien aus Kapitel 3 zeigt sich, dass für eine Vertex-Multiplizität von  $\langle N_{MB} \rangle = 4.1$  bzw.  $\langle N_{MB} \rangle = 6.9$  (hier ohne den Signalvertex) in  $\approx 0.5\%$  bzw.  $\approx 1.5\%$  der Fälle zwei oder mehr Vertizes in einem Intervall der Breite von 0.05 mm liegen (siehe Abbildung 3.5(b)). Dies ist ein grobes Maß für die Häufigkeit, dass zwei Vertizes als ein Vertex rekonstruiert werden. Durch die größere Spur-Multiplizität wird ein solcher Vertex einem Signalvertex ähnlich sein. Dies macht eine Fehlidentifikation wahrscheinlicher. Diese Effekte sind leider noch nicht detailliert untersucht worden und eine Studie dieser Einflüsse konnte im Rahmen dieser Arbeit aufgrund eines Fehlers in der Information zur Simulation nicht durchgeführt werden.

Um die Unterschiede zwischen simuliertem und selbstgemischtem *Pile-Up* für die verschiedenen Signalarten, *Pile-Up*-Szenarien und Selektionsmethoden besser vergleichen zu können, ist in Tabelle 6.3 jeweils das Verhältnis aus der Rate bei Ereignissen mit simuliertem und mit selbstgemischtem *Pile-Up* zu sehen. Im Falle des  $H \rightarrow \gamma\gamma$ -Datensatz mit  $\langle N_{MB} \rangle = 6.9$  ist kein direkter Vergleich mit den anderen Datensätzen möglich, da zur

Kriterium		$\sum p_T^2 \cdot \sqrt{N_{Sp}}$	$\mathcal{P}_{Vtx}$	$\mathcal{S}_{pT}^2$	NN
$WH$	4.1	$(3.2 \pm 0.1)\%$	$(4.0 \pm 0.1)\%$	$(2.7 \pm 0.1)\%$	$(2.5 \pm 0.1)\%$
	6.9	$(4.5 \pm 0.1)\%$	$(5.6 \pm 0.1)\%$	$(3.8 \pm 0.1)\%$	$(3.4 \pm 0.1)\%$
$t\bar{t}$	4.1	$(0.33 \pm 0.01)\%$	$(0.42 \pm 0.02)\%$	$(0.24 \pm 0.01)\%$	$(0.20 \pm 0.01)\%$
	6.9	$(0.49 \pm 0.02)\%$	$(0.63 \pm 0.02)\%$	$(0.38 \pm 0.02)\%$	$(0.32 \pm 0.02)\%$
$H \rightarrow \gamma\gamma$	4.1	$(23.1 \pm 0.2)\%$	$(28.9 \pm 0.2)\%$	$(22.2 \pm 0.2)\%$	$(21.9 \pm 0.2)\%$
	6.9	$(43.3 \pm 0.5)\%$	$(42.9 \pm 0.5)\%$	$(40.4 \pm 0.5)\%$	$(37.9 \pm 0.5)\%$

**Tab. 6.1.:** Fehlidentifikationsraten der in ATHENA implementierten Methoden zur Signalvertex-Selektion für zwei *Pile-Up*-Szenarien (4.1 und 6.9 entsprechen den mittleren Anzahlen an zusätzlichen *simulierten Minimum-Bias*-Vertexes zum Signalvertex  $\langle N_{MB,sim} \rangle$  bzw. der Anzahl aller *rekonstruierten* Vertexes  $\langle N_{Vtx,rek} \rangle$  inklusive des Signalvertex (siehe Text)).  $\sum p_T^2 \sqrt{N_{Sp}}$  (vgl. Gl. 4.15) wird als Selektionskriterium verwendet, wenn keine Neusortierung vom Benutzer verlangt wird.

Kriterium		$\mathcal{P}_{Vtx}$	$\mathcal{S}_{pT}^2$	NN
$WH$	4.1	$(3.6 \pm 0.1)\%$	$(2.3 \pm 0.1)\%$	$(2.0 \pm 0.1)\%$
	6.9	$(5.3 \pm 0.1)\%$	$(3.2 \pm 0.1)\%$	$(2.8 \pm 0.1)\%$
$t\bar{t}$	4.1	$(0.34 \pm 0.03)\%$	$(0.20 \pm 0.03)\%$	$(0.10 \pm 0.02)\%$
	6.9	$(0.51 \pm 0.04)\%$	$(0.26 \pm 0.03)\%$	$(0.23 \pm 0.03)\%$
$H \rightarrow \gamma\gamma$	4.1	$(27.6 \pm 0.3)\%$	$(21.4 \pm 0.2)\%$	$(20.4 \pm 0.2)\%$
	6.9	$(34.7 \pm 0.3)\%$	$(28.6 \pm 0.3)\%$	$(26.9 \pm 0.3)\%$

**Tab. 6.2.:** Fehlidentifikationsraten bei selbstgemischtem *Pile-Up* für  $\langle N_{Vtx} \rangle = 4.1$  und  $\langle N_{Vtx} \rangle = 6.9$  (Signal + *Minimum-Bias*) als *Poisson*-Mittelwert der Vertex-Multiplizitätsverteilung für verschiedene Selektionsmethoden.

Kriterium		$\mathcal{P}_{Vtx}$	$\mathcal{S}_{pT}^2$	NN
$WH$	4.1	$(1.11 \pm 0.04)\%$	$(1.17 \pm 0.05)\%$	$(1.22 \pm 0.05)\%$
	6.9	$(1.06 \pm 0.03)\%$	$(1.17 \pm 0.04)\%$	$(1.21 \pm 0.04)\%$
$t\bar{t}$	4.1	$(1.2 \pm 0.1)\%$	$(1.2 \pm 0.1)\%$	$(2.0 \pm 0.2)\%$
	6.9	$(1.2 \pm 0.1)\%$	$(1.5 \pm 0.1)\%$	$(1.4 \pm 0.1)\%$
$H \rightarrow \gamma\gamma$	4.1	$(1.02 \pm 0.01)\%$	$(1.04 \pm 0.01)\%$	$(1.07 \pm 0.01)\%$

**Tab. 6.3.:** Verhältnis der Fehlidentifikationsraten von Ereignissen mit simuliertem *Pile-Up* (Tabelle 6.1) zu Ereignissen mit selbstgemischtem *Pile-Up* (Tabelle 6.2) für  $\langle N_{Vtx} \rangle = 4.1$  und  $\langle N_{Vtx} \rangle = 6.9$  (Signal + *Minimum-Bias*) für verschiedene Selektionsmethoden.

Rekonstruktion der Ereignisse nur eine Version der Rekonstruktion zur Verfügung stand, die sich von der für die restlichen Datensätze verwendeten Version unterscheidet. Der Leistungsvergleich dieses Datensatzes ist aus diesem Grund nicht in Tabelle 6.3 zu finden. In allen betrachteten Fällen ist der Unterschied zwischen den Raten für simuliertes und für selbstgemischtes *Pile-Up* für das neuronale Netz am größten. Hier liegt im Falle von  $t\bar{t}$  bei  $\langle N_{Vtx} \rangle = 4.1$  ein Unterschied von einem Faktor zwei vor. Ursache dessen mag sein, dass die Datensätze zum Trainieren des Netzes kein simuliertes *Pile-Up* enthalten haben und somit der Vertex-Überlapp vernachlässigt wurde. Zusätzlich zu den oben erwähnten Effekten des Vertex-Überlapps kann in einem solchen Fall auch die Fehlzuordnung von einzelnen Spuren zu den Vertices zu einer Veränderung der Form der Verteilungen führen. Es scheint insofern sinnvoll, das Training des Netzes auf Datensätzen mit simulierten *Pile-Up* zu wiederholen, sobald die Effekte des Überlapps verstanden sind. Unter allen Signalsorten zeigt  $t\bar{t}$  die größten Unterschiede, wobei anzumerken ist, dass die Unterschiede in den meisten Fällen innerhalb der Fehler mit den Unterschieden auf den anderen Datensätzen kompatibel sind. Bis auf wenige Ausnahmen liegen die Unterschiede zwischen 10 % und 20 %.

Zusammenfassend ist festzuhalten, dass sich die Leistungen der in ATHENA implementierten Selektionsmethoden in der Weise verhalten, wie man es aus den Ergebnissen der vorangegangenen Studien erwartet. Die absoluten Raten sind dabei in allen betrachteten Fällen etwas höher. Dies kann auf den in den vorigen Studien vernachlässigten Vertex-Überlapp zurück geführt werden. Hierzu sind allerdings weitere Studien notwendig, die die Effekte des Überlapps detaillierter untersuchen. Nach wie vor wird sowohl mit Verwendung der  $\mathcal{S}_{p_T}^2$ -Methode als auch mit der Selektion anhand des neuronalen Netzes eine Verbesserung der Selektionseffizienz gegenüber der vorherigen Standardmethode ( $\sum p_T^2 \sqrt{N_{Sp}}$ ) erreicht.



## 7. Einfluss von *Pile-Up* auf die b-Jet-Identifikation

Die Selektion des Signalvertex aus allen rekonstruierten Primärvertizes im Ereignis entspricht in erster Linie der Wahl der  $z$ -Position, von der die Spuren des Prozesses von Interesse ausgehen. Wird an Stelle der Position des Signalvertex die eines *Minimum-Bias*-Vertex ausgewählt, entspricht diese Position nicht dem Ursprung der Spuren, die untersucht werden sollen. Für eine nachfolgende Analyse, die die Primärvertex-Position explizit verwendet, ist zu erwarten, dass eine Fehlidentifikation der Signalvertex-Position einen Leistungsabfall zur Folge hat. Ein Beispiel hierfür ist die Identifikation von b-Jets, *b-Tagging* genannt. Unter einem b-Jet ist ein Jet zu verstehen, der als schwerstes Quark ein b-Quark enthält. Das effiziente Identifizieren von b-Jets ist unter anderem für die Higgs-Boson-Entdeckung essentiell, da dieses für Higgs-Boson-Massen unterhalb von  $135 \text{ GeV}/c^2$  bevorzugt in ein  $b\bar{b}$ -Paar zerfällt (siehe Abbildung 2.3(a)), aber auch für viele andere Analysen spielt die b-Jet-Identifikation eine wichtige Rolle. Da Top-Quarks fast ausschließlich in b-Quarks zerfallen, ist das b-Tagging für alle Prozesse, die Top-Quarks enthalten relevant. Dies ist unter anderem für die Unterdrückung des am LHC sehr häufig auftretenden  $t\bar{t}$ -Untergrunds wichtig.

Das Identifizieren von b-Jets bedeutet, solche von Jets zu unterscheiden, die von einem c-Quark (c-Jet) oder von einem Quark des Typs u, d oder s (light-Jet) ausgehen. Die Unterscheidung zwischen c- und b-Jets ist dabei schwieriger als die Unterscheidung zwischen light- und b-Jets, da die Eigenschaften von Hadronen aus c-Quarks, in der Folge c-Hadronen genannt, den Eigenschaften der b-Hadronen ähnlich sind. Für die vorliegenden Studien ist zunächst lediglich die Unterscheidbarkeit von light-Jets und b-Jets untersucht worden.

b-Quarks hadronisieren nach ihrer Produktion sehr schnell zu einem b-Hadron. Wenn das b-Quark zerfällt geht das b-Hadron fast immer in ein oder zwei c-Hadronen, beispielsweise in D-Mesonen über. Ein solcher Zerfall ist in Abbildung 2.1 gezeigt worden. Auch das c-Hadron zerfällt nach einer gewissen Zeit. Ein b-Hadron-Zerfall hinterlässt demnach zwei charakteristische Zerfallspunkte. Viele b-Tagging-Algorithmen machen sich diese zunutze. Die mittlere Flugstrecke eines b-Hadrons lässt sich aus der b-Hadron-Lebensdauer

von  $\tau_B \approx 1.5 \text{ ps}$  (siehe Abschnitt 2.1.2) abschätzen. Für einem typischen transversalen Jetimpuls von  $p_T = 50 \text{ GeV}/c$  ergibt sich eine mittlere b-Hadron-Flugstrecke von  $\langle \ell \rangle = \beta\gamma c\tau_B \approx 3 \text{ mm}$ . Diese legt das b-Hadron zurück, bevor es zerfällt. Die Spuren des b-Hadron-Zerfalls gehen somit nicht vom Primärvertex, sondern von einem etwas vom Primärvertex entfernt liegenden Sekundärvertex aus. Gleiches gilt für den etwas später erfolgenden c-Hadron-Zerfall. Werden diese Spuren in den Wechselwirkungsbereich extrapoliert, treffen sie nicht in den Primärvertex und weisen somit im Allgemeinen größere Impact-Parameter auf, als Spuren direkt aus dem Primärvertex. Dies kann zur Identifikation eines b-Jets verwendet werden. Ebenso ist es möglich den Sekundärvertex zu rekonstruieren und Informationen über diesen zur b-Jet-Identifikation hinzuzuziehen.

Bei einer Fehlidentifikation des Signalvertex stimmt die z-Position des ausgewählten Primärvertex nicht mit dem Ursprung des b-Hadrons überein. Die zum b-Tagging verwendeten Variablen hängen unterschiedlich stark von dieser Position ab. Es ist davon auszugehen, dass eine Fehlidentifikation des Signalvertex einen Leistungsabfall der b-Jet-Identifikation zur Folge hat, wobei je nach Methode und je nachdem, welche Variablen zur Identifikation verwendet werden mit einem unterschiedlich starken Einfluss zu rechnen ist.

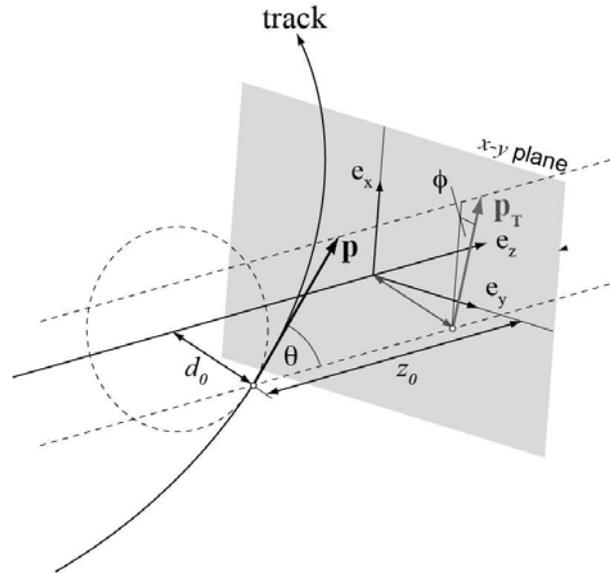
Da am LHC große *Pile-Up*-Raten erwartet werden, ist es notwendig, abschätzen zu können, welchen Einfluss *Pile-Up* auf die Leistung des b-Taggings hat. In diesem Kapitel wird neben dem Einfluss durch die Fehlidentifikation des Signalvertex auch der Einfluss alleine durch die Existenz von Spuren aus *Minimum-Bias*-Wechselwirkungen untersucht.

## 7.1. b-Tagging-Algorithmen bei ATLAS

In diesem Abschnitt werden die b-Tagging-Algorithmen bei ATLAS vorgestellt, wobei insbesondere auf die Algorithmen eingegangen werden soll, die für die nachfolgenden Studien von Bedeutung sind. Eine Beschreibung der Algorithmen findet sich, wenn nicht anderweitig angegeben, in [9].

### 7.1.1. Verwendete Rekonstruktionsobjekte

Die Identifikation von b-Jets hängt in erster Linie von drei Rekonstruktionsobjekten ab: von Spuren, von der Position  $\vec{x}_{pv}$  des ersten Primärvertex in der Liste aller rekonstruierten Primärvertizes und von der Richtung der rekonstruierten Jets im Ereignis. Anhand der Informationen über diese Objekte werden die Spuren für das b-Tagging präpariert. Dabei werden zunächst Spuren entfernt, die nicht aus dem selektierten Primärvertex stammen. Die Spurparameter enthalten die Information über sowohl den transversalen ( $d_0$ ) als auch longitudinalen Impact-Parameter ( $z_0$ ). Die Größe  $z_0$  wird im folgenden bezogen auf die



**Abb. 7.1.:** Skizze zur räumlichen Lage des transversalen und des longitudinalen Impact-Parameters. Der Primärvertex liegt in der vorliegenden Abbildung im Ursprung des Koordinatensystems. Abbildung aus [24].

z-Position des Primärvertex. In Abbildung 7.1 ist eine Skizze zu sehen, die den transversalen und longitudinalen Impact-Parameter veranschaulicht. Der Primärvertex liegt in dieser Abbildung im Ursprung. Spuren, die zum b-Tagging zugelassen werden, müssen einige Bedingungen erfüllen. Die Bedingungen für die in der Folge beschriebenen Impact-Parameter-Algorithmen sind:  $d_0 < 1 \text{ mm}$  und  $z_0 \cdot \sin \theta < 1.5 \text{ mm}$ . Die Bedingungen für die Sekundärvertex-Algorithmen sind etwas schwächer, hier gilt das Folgende:  $d_0 < 5 \text{ mm}$  und  $(z_0 \cdot \sin \theta) / \sigma_{z_0} < 6$ . Diese Bedingungen dienen der Unterdrückung von Spuren aus *Minimum-Bias*-Wechselwirkungen, deren Spuren nicht in den ausgewählten Primärvertex zeigen, der im Idealfall der Signalvertex ist. Für die folgenden Studien werden die oben angeführten Bedingungen eine wichtige Rolle spielen, da bei einer Fehlidentifikation des Signalvertex die Spuren von Interesse durch diese Bedingung nicht zum b-Tagging zugelassen werden.

Eine Reihe weiterer Qualitätskriterien wird an die Spuren gestellt, die in erster Linie dafür Sorge tragen, schlecht gemessene Spuren zu unterdrücken. Um eine hohe Spurqualität zu garantieren, werden mindestens zwei Treffer in den Pixel-Detektoren und mindestens sieben Treffer in den Präzisions-Spurdetektoren (Pixel- und SCT-Detektoren) gefordert. Hierzu sei auf die Beschreibung des inneren Detektors und der Spurrekonstruktion in Ab-

schnitt 2.2.2 hingewiesen. Aus der Information über die Richtung der Jetachse können die übrigen Spuren zu Sätzen gebündelt werden, die jeweils zu einem Jet gehören. Alle Spuren, die in einem Abstand  $\Delta R < 0.4$  um die Jetachse liegen, werden diesem Jet zugeordnet. Jedes Spurbündel wird weiterhin von Spuren bereinigt, die kompatibel mit einer Photon-Konversion oder dem Zerfall eines sogenannten  $V^0$ -Teilchens sind.  $V^0$ -Teilchen haben eine Lebensdauer, die ausreichend lang ist, um ähnlich wie ein b-Hadron einen deplatzierten Zerfallsvertex zu hinterlassen. Zu ihnen zählt beispielsweise das  $K_s$  (Meson mit s-Quark) und das  $\Lambda$  (Baryon mit s-Quark).

Die Information über die Jets dient lediglich der Bündelung der Spuren. Die eigentliche Identifikation von b-Jets basiert in erster Linie auf der Information über die Helixparameter der Spuren in einem Jet. Es sind die Eigenschaften der Spuren in einem Jet, anhand derer entschieden wird, ob ein Jet ein b-Jet ist oder nicht. Hierbei spielen vor allem die Impact-Parameter und deren Auflösungen eine zentrale Rolle.

### 7.1.2. Impact-Parameter-Algorithmen

Der Impact-Parameter ist ein Maß für die Wahrscheinlichkeit, dass eine Spur aus dem Primärvertex stammt. Dabei ist die jeweilige Auflösung des Impact-Parameters zu berücksichtigen. An Stelle der Impact-Parameter selbst werden als Variablen zur Identifikation von b-Jets daher die sogenannten Impact-Parameter-Signifikanzen  $d_0/\sigma_{d_0}$  und  $z_0/\sigma_{z_0}$  verwendet. Abbildung 7.2 zeigt eine Skizze einer typischen Spur-Konstellation für einen b-Jet. Spuren aus einem b-Hadron-Zerfall kreuzen die Jet-Achse im Allgemeinen auf der gleichen Seite, bezogen auf den Primärvertex, auf der auch der Jet selbst liegt. Spuren, die aus dem Primärvertex stammen, werden in gleichem Maße auf beiden Seiten der Jet-Achse kreuzen. Dies kann über ein Vorzeichen des Impact-Parameters berücksichtigt werden. Dabei sind die folgenden in Abbildung 7.2 bezeichneten Größen relevant: die Jet-

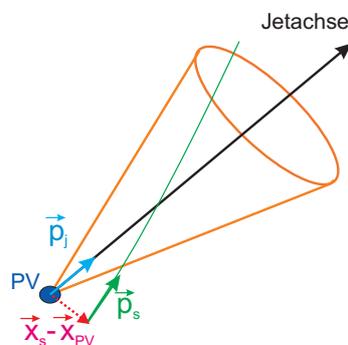
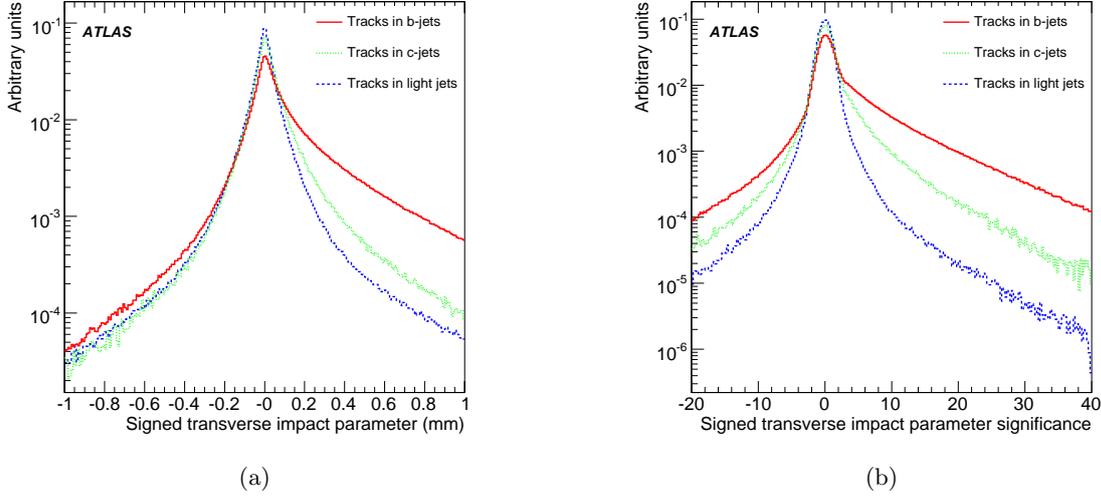


Abb. 7.2.: Skizze zum Vorzeichen der Impact-Parameter. Abbildung aus [21].



**Abb. 7.3.:** (a) Vorzeichenbehafteter transversaler Impact-Parameter  $d_0$  und (b) transversale Impact-Parameter-Signifikanz  $d_0/\sigma_{d_0}$  für Spuren von b-, c- und light-Jets. Abbildungen aus [9].

richtung  $\vec{p}_j$ , die Richtung der Spur  $\vec{p}_s$  im Punkt  $\vec{x}_s$  der nächsten Annäherung der Spur an die z-Achse und der Abstand von diesem Punkt zum Primärvertex  $|\vec{x}_{pv} - \vec{x}_s|$ . Das Vorzeichen berechnet sich dann wie folgt:

$$\text{sign}(d_0) = (\vec{p}_j \times \vec{p}_s) \cdot (\vec{p}_s \times (\vec{x}_{pv} - \vec{x}_s)). \quad (7.1)$$

Abbildung 7.3 zeigt die Verteilung des transversalen Impact-Parameters  $d_0$  und der Impact-Parameter-Signifikanz  $d_0/\sigma_{d_0}$  für Spuren von b-, c- und light-Jets.

Bei ATLAS existieren zwei Methoden, die die vorzeichenbehafteten Impact-Parameter-Signifikanzen in einer Likelihood-Methode, wie sie in Abschnitt 5.2.3 beschrieben ist, verwenden:

- *IP2D* basiert alleine auf der transversalen Impact-Parameter-Signifikanz.
- *IP3D* basiert sowohl auf der longitudinalen als auch auf den transversalen Impact-Parameter-Signifikanz, wobei zur Berücksichtigung der Korrelationen zwischen den beiden eine zweidimensionale Wahrscheinlichkeitsdichtefunktion verwendet wird.

Prinzipiell ist zu erwarten, dass *Pile-Up* einen größeren Effekt auf *IP3D* als auf *IP2D* haben sollte, da ersterer explizit und letzterer nicht explizit von der z-Position des Primärvertex abhängt. Aus diesem Grund werden diese beiden Algorithmen im Hinblick auf den Einfluss von *Pile-Up* in den folgenden Studien verglichen.

Der Vollständigkeit halber sei hier darauf hingewiesen, dass zwei weitere Impact-Para-

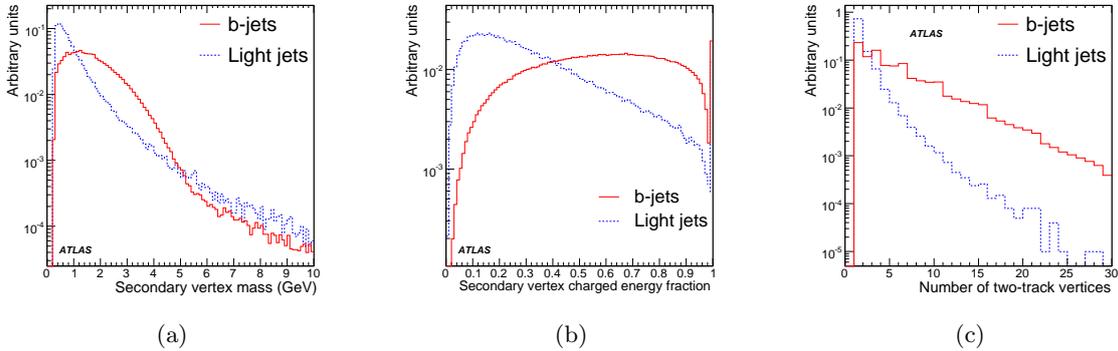
meter-Algorithmen in der ATLAS-Software zur Verfügung stehen. Dies ist einerseits *Jet-Prob*, eine Methode, die auf den gleichen Grundprinzipien beruht wie die in Abschnitt 5.2.2 beschriebene  $\mathcal{P}_{Vtx}$ -Methode [9]. Andererseits existiert eine sehr simple Methode, *Track-Counting* genannt, die zur b-Jet-Identifikation lediglich eine gewisse Anzahl an Spuren im Jet mit einem Impact-Parameter einer gewissen Größe fordert [25]. Diese beiden Methoden zeigen zwar im Vergleich zu den anderen implementierten Methoden eine geringere Identifikationsfähigkeit, sie sind aber für den Beginn der Datennahme von Interesse, da hier nur eine geringe Abhängigkeit von Verteilungen aus der Simulation vorliegt.

### 7.1.3. Sekundärvertex-Algorithmen

Für b-Tagging-Algorithmen, die die Information von rekonstruierten Sekundärvertizes im Jet verwenden, bedarf es zunächst einer Bestimmung der Sekundärvertizes. Zwei unterschiedliche Ansätze hierfür finden bei ATLAS Anwendung. Im einen Fall wird ohne Rücksicht auf den etwas später erfolgenden c-Hadron-Zerfall nur ein Sekundärvertex rekonstruiert. Im anderen Fall wird die ganze Topologie der Zerfallskette ausgenutzt. Hierfür müssen der Vertex des b- und des c-Hadron-Zerfalls getrennt rekonstruiert werden. Der erste Ansatz wird von dem b-Tagging-Algorithmus *SV1* verwendet, der zweite von einem Algorithmus genannt *JetFitter*.

Die Sekundärvertex-Rekonstruktion der *SV1*-Methode sucht nach nur einem Vertex im Jet, der die Spuren des b- und des c-Hadron-Zerfalls enthält. So geht zwar die Information über die Topologie der Zerfallskette verloren, dafür ist aber die Effizienz, überhaupt einen Sekundärvertex zu rekonstruieren, größer. Es werden hierfür zunächst alle möglichen Vertizes aus zwei Spuren im Jet gebildet, in der Folge *Zwei-Spurvertizes* genannt. Danach werden alle Spuren entfernt, bei denen der zugehörige Zwei-Spurvertex eine mit einem  $V^0$ -Teilchen kompatible Masse hat. Die Spuren an den übrigen gebildeten Vertizes werden zu einem Vertex zusammengeführt. Nun werden iterativ immer die Spuren entfernt, die den größten Beitrag zum  $\chi^2$  der Vertex-Bestimmung liefern und die Vertex-Position wird mit den übrigen Spuren erneut berechnet. Die Prozedur wird so lange wiederholt, bis das  $\chi^2$  einen akzeptablen Wert erreicht hat. Von dem resultierenden Vertex werden die folgenden Größen zur b-Jet-Identifikation genutzt: Die invariante Masse der geladenen Teilchen an diesem Vertex, das Verhältnis der gesamten Energie aller Spuren am Vertex zur gesamten Energie aller Spuren im Jet und die Anzahl der Zwei-Spurvertizes. Die Verteilungen dieser Größen für b- und light-Jets sind in Abbildung 7.4 zu sehen.

Die hier gezeigten Variablen werden mittels einer Likelihood-Methode kombiniert, wobei das Energieverhältnis und die invariante Masse als zweidimensionale und die Anzahl der Zwei-Spurvertizes als eindimensionale Wahrscheinlichkeitsdichtefunktion verwendet wird.



**Abb. 7.4.:** (a) Invariante Masse der geladenen Teilchen am Sekundärvertex. (b) Verhältnis der Energie aller Spuren am Sekundärvertex zur Energie aller Spuren im Jet. (c) Anzahl der Zwei-Spurvertices im Jet. Abbildungen aus [9].

Vom *SV1*-Algorithmus wird erwartet, dass er gegen Effekte von *Pile-Up* robuster ist als die Impact-Parameter-Algorithmen *IP2D* und *IP3D*. In Hinblick hierauf sollen diese drei Algorithmen in der Folge untersucht und verglichen werden.

Prinzipiell wäre es möglich, eine weitere Größe, die die Information über den Abstand zwischen Primär- und Sekundärvertex enthält, zu verwenden. Die Sekundärvertex-Algorithmen sind jedoch darauf ausgelegt, mit den Impact-Parameter-Algorithmen kombiniert zu werden. Der Abstand zwischen Primär- und Sekundärvertex und die Impact-Parameter enthalten teilweise die gleiche Information und sind aus diesem Grund voneinander abhängig. Es wird daher im Falle des *SV1*-Algorithmus auf die Verwendung dieser Information verzichtet. Dies macht es möglich, *SV1* und *IP3D* über eine einfache Likelihood-Methode zu kombinieren, ohne dabei allzu große Korrelationen berücksichtigen zu müssen. Ein solcher Algorithmus ist bei ATLAS unter dem Namen *COMB* implementiert. Auch der *COMB*-Algorithmus soll in der Folge im Hinblick auf den Einfluss von *Pile-Up* untersucht werden.

In den Abbildungen 7.5 sind die unterschiedlichen Methoden der Sekundärvertex-Rekonstruktion des *JetFitters* und des *SV1*-Algorithmus skizziert. Im Falle von *SV1* werden beide Zerfallspunkte in einem Sekundärvertex rekonstruiert im Falle des *JetFitters* wird nach zwei getrennten Zerfallspunkten gesucht, wobei angenommen wird, dass der Zerfall des b- und des c-Hadrons annähernd auf der b-Hadron-Flugachse bzw. auf der Jetachse zu finden sind. Die Einschränkung durch die Jetachse gibt die Möglichkeit, aus dem Kreuzungspunkt nur einer Spur mit der Jetachse einen Sekundärvertex zu bestimmen. In sehr pffiger Weise kann so die Topologie der Zerfallskette auch erhalten werden, wenn einige der Spuren nicht rekonstruiert wurden. Der *JetFitter*-Algorithmus wird in der Folge nicht



**Abb. 7.5.:** (a) Inklusiver Sekundärvertex-Fit des *SV1*-Algorithmus. Die Spuren aus b-Hadron-Zerfall (Punkt B) und aus dem c-Hadron-Zerfall (Punkt D) werden in einem Punkt (*inclusive B/D vertex*) rekonstruiert. (b) Rekonstruktion der Zerfallskette durch den *JetFitter*-Algorithmus. Der Primärvertex (*Primary Vertex*), der Vertex des b-Hadron- und des c-Hadron-Zerfalls liegen näherungsweise auf der b-Hadron-Flugachse (*B flight axis*). Abbildungen aus [9].

weiter berücksichtigt.

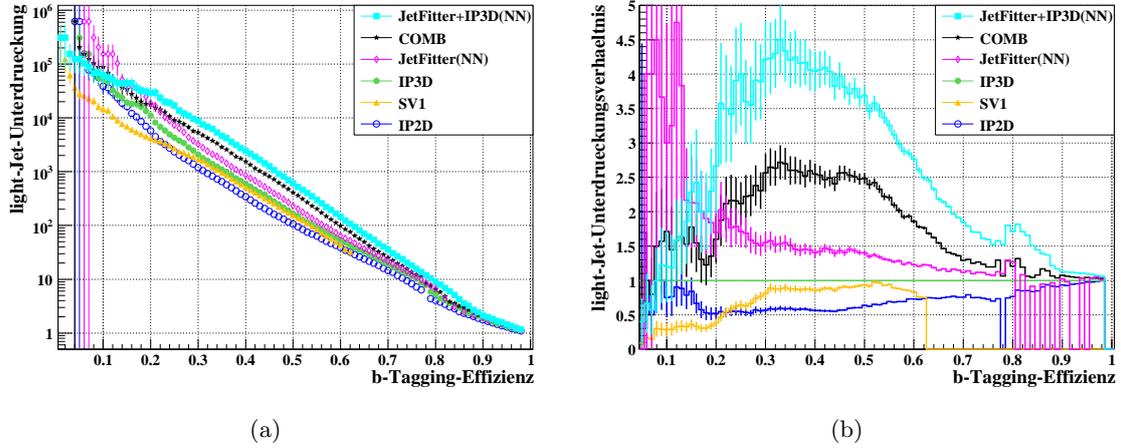
Zusätzlich zu den hier angesprochenen b-Tagging-Methoden existieren auch Algorithmen, die die Informationen über Leptonen im Jet ausnutzen. Solche Methoden bleiben in den nachfolgenden Studien unberücksichtigt. Es sei hier noch einmal zusammengefasst, welche b-Tagging-Algorithmen auf *Pile-Up*-Einflüsse untersucht werden sollen:

- *IP2D*: Ein Algorithmus, der alleine von den transversalen Impact-Parametern der Spuren abhängt.
- *IP3D*: Ein Algorithmus, der sowohl von den transversalen, als auch von den longitudinalen Impact-Parametern der Spuren abhängt.
- *SV1*: Ein Algorithmus, der die Information über einen rekonstruierten Sekundärvertex zur b-Jet-Identifikation verwendet.
- *COMB*: *IP3D* + *SV1* über eine Likelihood-Methode kombiniert.

## 7.2. Einfluss von *Pile-Up* auf die b-Tagging-Leistung

Ziel eines effizienten b-Taggings ist es, eine möglichst große b-Tagging-Effizienz bei möglichst kleiner c- und light-Jet-Fehlidentifikation zu erreichen. Die b-Tagging-Effizienz ist definiert als:

$$\epsilon_b = \frac{\text{Anzahl der als b-Jet identifizierten b-Jets}}{\text{Anzahl aller b-Jets}}. \quad (7.2)$$



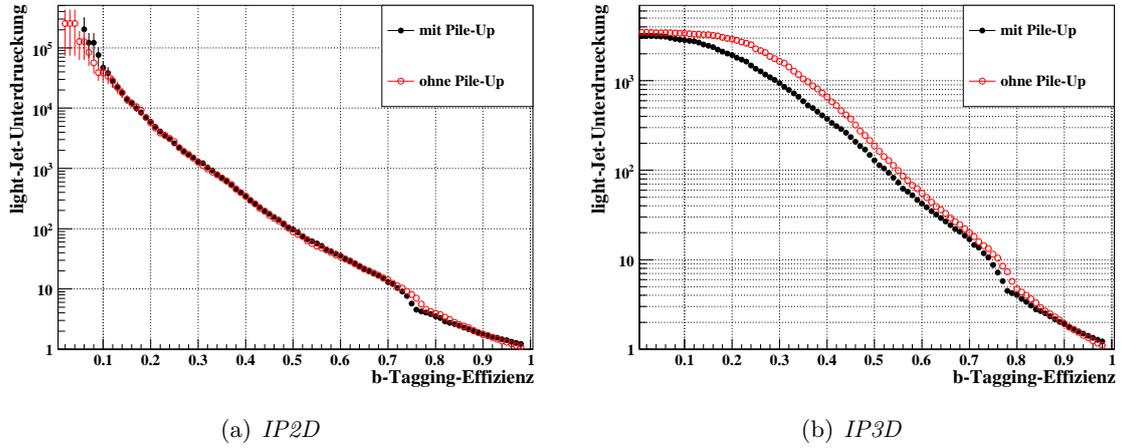
**Abb. 7.6.:** b-Tagging-Leistungsfähigkeit für verschiedene b-Tagging-Algorithmen im Vergleich. (a) Light-Jet-Unterdrückung gegen die b-Tagging-Effizienz. (b) Verhältnis der light-Jet-Unterdrückung der verschiedenen Methoden, bezogen auf die light-Jet-Unterdrückung von *IP3D*, gegen die b-Tagging-Effizienz.

Die c-Jet und light-Jet-Fehlidentifikation ist wie folgt definiert:

$$\epsilon_{c,light} = \frac{\text{Anzahl der als b- identifizierten c- bzw. light-Jets}}{\text{Anzahl aller c- bzw. light-Jets}}. \quad (7.3)$$

Der Kehrwert der Fehlidentifikationsrate ergibt die Unterdrückung  $R_{c,light} = 1/\epsilon_{c,light}$ . Als Maß für die Leistung eines b-Tagging-Algorithmus wird die light-Jet-Unterdrückung in Abhängigkeit der b-Tagging-Effizienz betrachtet. In Abbildung 7.6 ist ein solcher Leistungsvergleich für die verschiedenen b-Tagging-Algorithmen gezeigt. Zusätzlich zur light-Jet-Unterdrückung wird auch das Verhältnis von dieser zu der Unterdrückung einer Referenzmethode gezeigt. Es ist zu sehen, dass eine Kombination aus *IP3D* und *JetFitter* die beste Leistungsfähigkeit zeigt. Die Bezeichnung *NN* weist darauf hin, dass die verschiedenen Eingangsvariablen über ein neuronales Netz kombiniert wurden. Der *COMB*-Algorithmus erreicht ebenfalls eine sehr gute b-Tagging-Leistung. Auch unter den b-Tagging-Methoden, die keine Kombinationen aus mehreren Algorithmen sind (*IP2D*, *IP3D*, *SV1*, und *JetFitter*), ist der *JetFitter* der leistungsstärkste b-Tagging-Algorithmus. Leistungsvergleiche dieser Art werden in der Folge auch für die Untersuchung von *Pile-Up*-Effekten durchgeführt.<sup>1</sup>

<sup>1</sup>Die hier gezeigten Ergebnisse sind mit der ATHENA-Version 14.0.0 erhalten worden. Alle weiteren beschriebenen Studien wurden mit der Version 15.4.0 durchgeführt. In dieser neueren Version ist ein Fehler in der Kalibration für *IP3D* aufgetreten. Ein Leistungsvergleich zwischen den verschiedenen B-Tagging-Algorithmen ist aus diesem Grund in dieser Version nicht aussagekräftig. Für die in der Folge betrachteten *Pile-Up*-Vergleiche ist dies aber nicht relevant. Es sei aber darauf hingewiesen, dass für kleine b-Tagging-Effizienzen die absoluten Leistungen von *IP3D* und *COMB* in den folgenden Studien von den hier gezeigten abweichen werden.



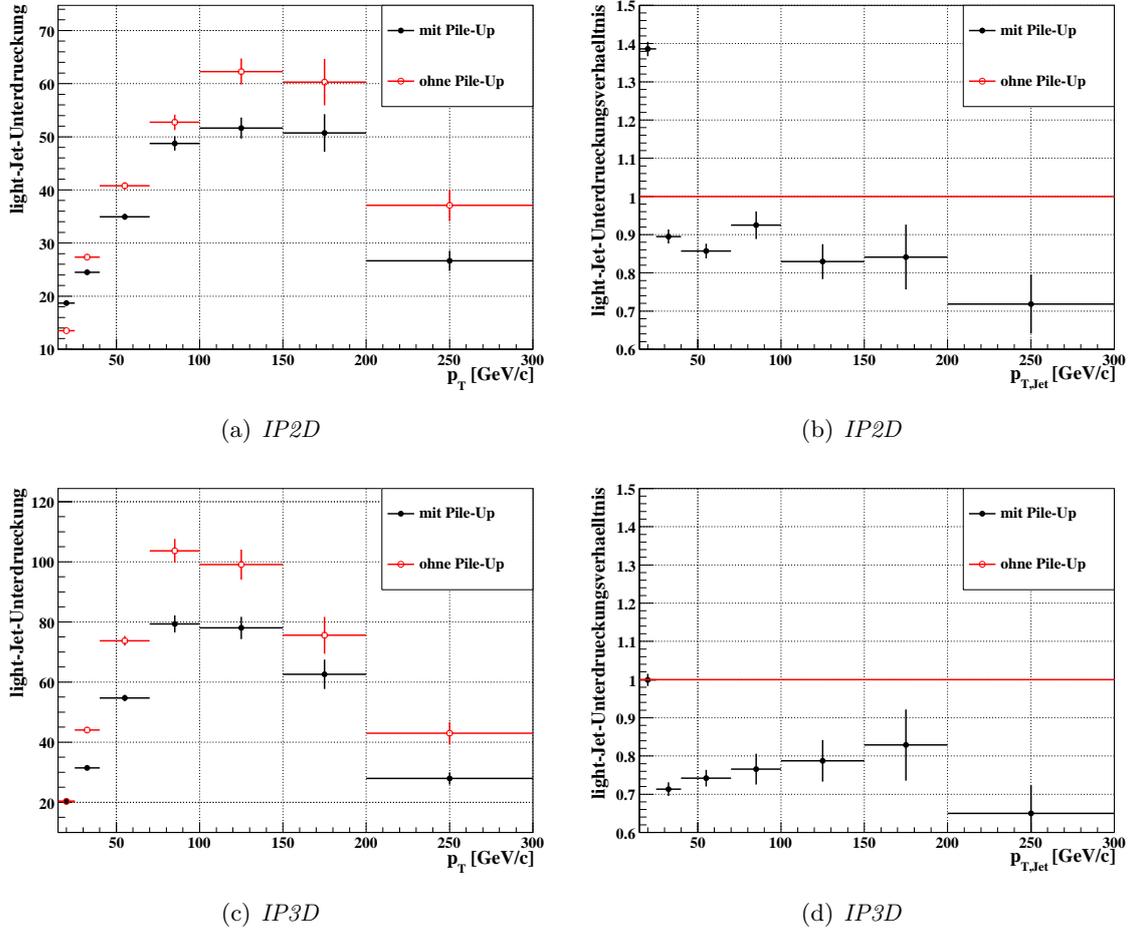
**Abb. 7.7.:** Vergleich der b-Tagging-Leistungsfähigkeit auf Datensätzen mit und ohne *Pile-Up*. Im Falle des *Pile-Up*-Datensatzes wurde der Signalvertex anhand der  $S_{pT}^2$ -Methode selektiert. Es werden alle rekonstruierten Jets berücksichtigt.

Ein naiver erster Schritt, um den Einfluss von *Pile-Up* auf die b-Tagging-Leistungsfähigkeit zu untersuchen, ist die light-Jet-Unterdrückung gegen die b-Tagging-Effizienz für Datensätze mit und ohne *Pile-Up* zu vergleichen. Hier und für alle nachfolgenden Studien wurde ein *Pile-Up*-Datensatz mit im Mittel 6.9 *Minimum-Bias*-Wechselwirkungen zusätzlich zur Signalwechselwirkung verwendet. Die Studien wurden auf *WH*-Ereignissen durchgeführt. Um ausreichend light-Jets zu erhalten, wurde ein Datensatz mit 200 000 *WH*( $H \rightarrow u\bar{u}$ ) Ereignissen zusätzlich zu einem Datensatz mit 50 000 *WH*( $H \rightarrow b\bar{b}$ ) Ereignissen verwendet. Diese Datensatzzusammensetzung wurde für alle nachfolgenden Studien genutzt. Abbildung 7.7 zeigt einen Vergleich, wie oben beschreibend. Im Falle von *IP3D* ist eine Verschlechterung durch *Pile-Up* vor allem für b-Tagging-Effizienzen zwischen 20 % und 50 % zu sehen. Für *IP2D* ist zunächst keine signifikante Änderung zu beobachten.

Betrachtet man den Leistungsvergleich in Abhängigkeit vom transversalen Jetimpuls, wird ein wichtiger Effekt sichtbar. In den Abbildungen 7.8 ist für *IP2D* und *IP3D* sowohl die light-Jet-Unterdrückung als auch das Verhältnis der light-Jet-Unterdrückung von Datensätzen mit und ohne *Pile-Up* für verschiedene Bereiche des transversalen Jetimpulses zu sehen. Die Unterdrückung wird für einen feste b-Tagging-Effizienz von 60 % betrachtet.<sup>2</sup> Während für große transversale Jetimpulse ( $p_{T,Jet} > 25 \text{ GeV}/c$ ) für beide b-Tagging-Algorithmen *Pile-Up* zu einem Leistungsabfall führt, ist für Impulse  $p_{T,Jet} < 25 \text{ GeV}/c$  im

<sup>2</sup>Hier und für alle weiteren *Pile-Up*-Vergleiche wird das Verhältnis auf den Fall ohne *Pile-Up* bezogen.

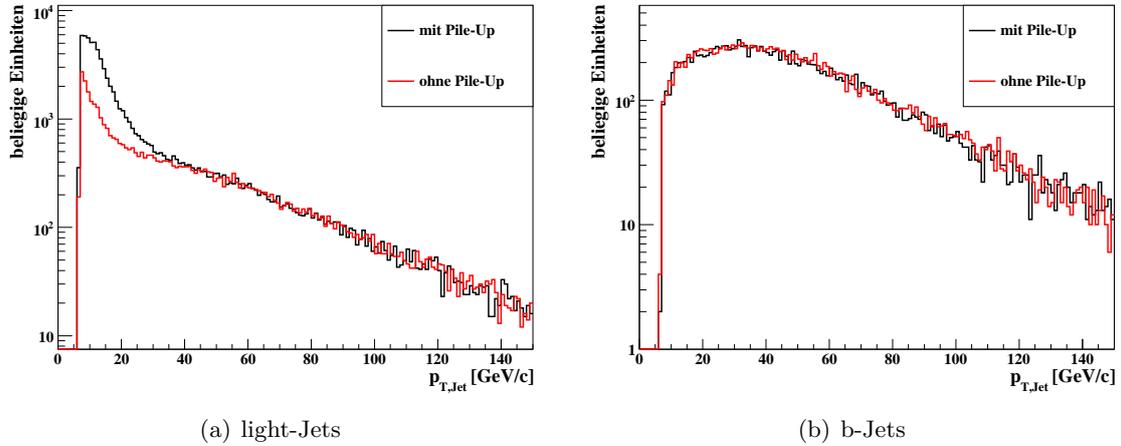
Ebenso wird für alle nachfolgenden Betrachtungen, die in Abhängigkeit des transversalen Impulses oder der Pseudorapidität gemacht werden, die Unterdrückung für eine feste b-Tagging-Effizienz von 60 % angegeben.



**Abb. 7.8.:** Vergleich der b-Tagging-Leistungsfähigkeit für Datensätze mit und ohne *Pile-Up* in Abhängigkeit des transversalen Jetimpulses  $p_{T,Jet}$  für eine b-Tagging-Effizienz von 60%. Im Falle des *Pile-Up*-Datensatzes wurde der Signalvertex anhand der  $S_{p_T}^2$ -Methode selektiert. Es werden alle rekonstruierten Jets berücksichtigt.

Falle von *IP3D* eine gleichbleibende Leistung, im Falle von *IP2D* sogar eine Leistungssteigerung durch *Pile-Up* zu beobachten. Der starke Anstieg des Verhältnisses zu kleinen transversalen Jetimpulsen hin ist eine künstliche Verbesserung, die nicht auf die Leistungsfähigkeit der b-Tagging-Methoden selbst zurückzuführen ist. Dies soll in der Folge kurz erläutert werden und es soll ein Verfahren vorgestellt werden, wie dieser Effekt unberücksichtigt bleibt.

In Abbildung 7.9 sind die Verteilungen der transversalen Jetimpulse für b- und light-Jets für Datensätze mit und ohne *Pile-Up* zu sehen. Während sich kein Unterschied für b-Jets zeigt, sieht man im Falle von light-Jets für transversale Jetimpulse unterhalb von  $p_{T,Jet} < 40$  GeV/c zusätzliche Jets. Diese sind Jets aus *Minimum-Bias*-Wechselwirkungen.



**Abb. 7.9.:** Spektrum des transversalen Jetimpulses für (a) light- und (b) b-Jets für Datensätze mit und ohne *Pile-Up*. Es wurden gleichviele Ereignisse im Falle der Datensätze mit und ohne *Pile-Up* verwendet.

Durch das Qualitätskriterium des longitudinalen Impact-Parameters ( $z_0 \cdot \sin \theta < 1.5$  mm im Falle von *IP2D* und *IP3D*) werden diesen Jets keine, oder in seltenen Fällen sehr wenige, Spuren zugeordnet. Solche *leeren* Jets bleiben von den b-Tagging-Algorithmen unberücksichtigt und werden direkt als light-Jets deklariert. Auf diese Weise, wird die Leistung des b-Taggings künstlich durch die Existenz dieser Jets verbessert. Eine solche Verbesserung zeigt nicht die tatsächliche Leistungsfähigkeit der b-Tagging-Methode, sondern bestenfalls, dass das oben erwähnte Qualitätskriterium sinnvoll *Minimum-Bias*-Jets unterdrückt. Um untersuchen zu können, welchen Einfluss *Pile-Up* unabhängig von dieser künstlichen Verbesserung hat, ist es notwendig, eine Studie durchzuführen, die die Jets aus *Minimum-Bias*-Ereignissen nicht in den Vergleich einbezieht. Aus diesem Grund wurden für alle nachfolgenden Vergleiche an Stelle der rekonstruierten Jets Jets verwendet, die aus den simulierten, bzw. generierten Teilchen der Signalwechselwirkung rekonstruiert wurden. Solche Jets werden in der Folge *wahre Signaljets* genannt. Vergleicht man nun wahre Signaljets in Datensätzen mit und ohne *Pile-Up*, ist die Jetinformation identisch, da so auch im Fall des *Pile-Up*-Datensatzes nur die Jets aus der Signalwechselwirkung berücksichtigt werden. Nach wie vor werden bei den nachfolgenden Vergleichen aber alle Spuren berücksichtigt, sowohl aus den Signal- als auch aus den *Minimum-Bias*-Wechselwirkungen.

Es sollen in der Folge zwei Einflüsse durch *Pile-Up* auf die b-Tagging-Leistungsfähigkeit untersucht werden:

- Einfluss aufgrund der Verunreinigungen der Signaljets durch Spuren aus *Minimum-Bias*-Wechselwirkungen.

- Einfluss aufgrund der Fehlidentifikation des Signalvertex.

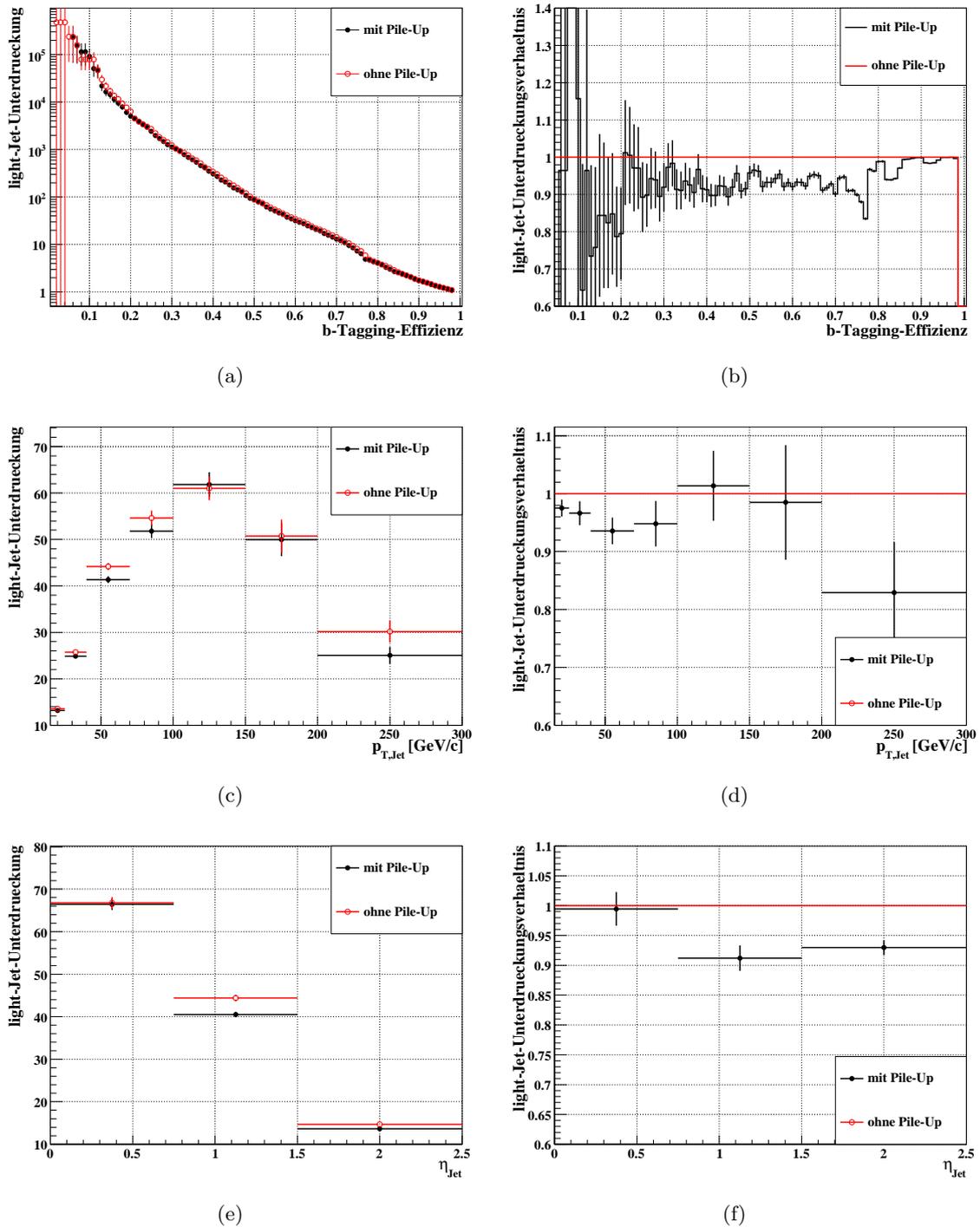
Bei einem einfachen Vergleich zwischen Datensätzen mit und ohne *Pile-Up* wird man, auch bei Verwendung der wahren Signaljets, eine Mischung aus diesen beiden Effekten beobachten. Es ist aber wünschenswert den Einfluss beider Effekte unabhängig voneinander abschätzen zu können. In den folgenden zwei Abschnitten sollen Studien vorgestellt die dies ermöglichen.

### 7.2.1. Einfluss von Spuren aus *Minimum-Bias*-Wechselwirkungen

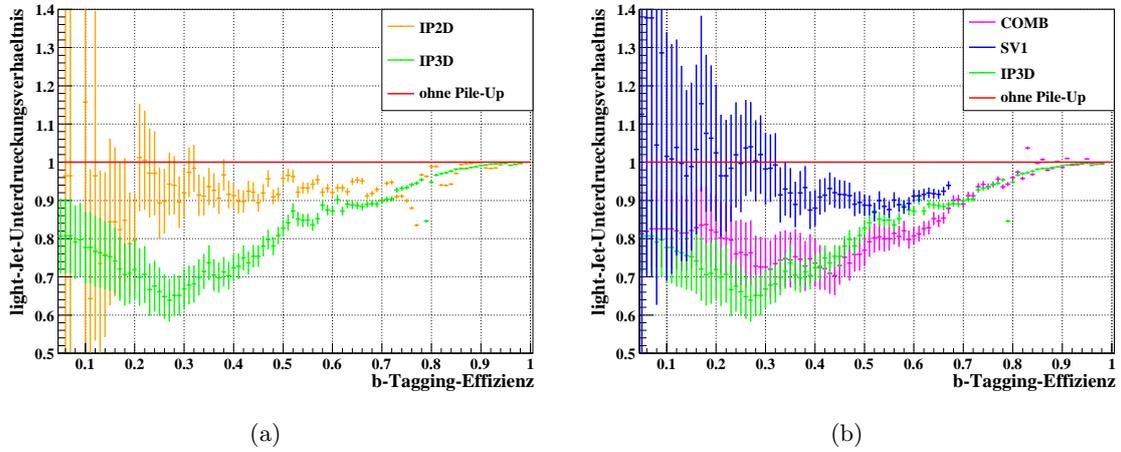
In diesem Abschnitt soll nun der Frage nachgegangen werden, welchen Einfluss Spuren aus *Minimum-Bias*-Wechselwirkungen haben, die fälschlicherweise den light-Jets und b-Jets aus der Signalwechselwirkung zugeordnet wurden. Um eine Fehlidentifikation des Signalvertex auszuschließen wurde mit Hilfe der zur Verfügung stehenden Simulationsinformation immer der tatsächliche Signalvertex selektiert.

Abbildung 7.10 zeigt einen Vergleich der b-Tagging-Leistungsfähigkeit von *IP2D* für Datensätze mit und ohne *Pile-Up*, der auf diese Weise erhalten worden ist. Die Leistungsvergleiche sind in Abhängigkeit der b-Tagging-Effizienz, des transversalen Jetimpulses und der Pseudorapidität der Jets gezeigt. Prinzipiell wäre zu erwarten, dass der Einfluss von *Pile-Up* für große Pseudorapiditäten stärker wird, da in diesen Bereichen mehr *Minimum-Bias*-Spuren erwartet werden. Für die vorliegende Statistik ist ein solcher Effekt aber nicht signifikant zu beobachten. Bei dem Leistungsvergleich in Abhängigkeit des transversalen Jetimpulses (Abb. 7.10(d)) ist zu erkennen, dass auch bei kleinen transversalen Jetimpulsen *Pile-Up* eine Verschlechterung der b-Tagging-Leistungsfähigkeit zur Folge hat. Das Leistungsverhältnis steigt zu kleinen transversalen Impulsen hin deutlich weniger stark an, als im Falle von Abbildung 7.8(b). Aus den vorliegenden Abbildungen ist zu sehen, dass eine Verschlechterung der b-Tagging-Leistungsfähigkeit durch zusätzliche Spuren aus *Minimum-Bias*-Wechselwirkungen auftritt. Gleiches gilt für alle anderen betrachteten Algorithmen. Die entsprechenden Abbildungen für diese befinden sich in Anhang A.1.

Wie in den ersten Abschnitten des Kapitels beschrieben, ist zu erwarten, dass die verschiedenen b-Tagging-Methoden unterschiedlich stark auf den Einfluss von *Pile-Up* reagieren. Um dies vergleichen zu können, wurde jeweils das Verhältnis der b-Tagging-Leistungsfähigkeit eines Datensatzes mit *Pile-Up* zur Leistungsfähigkeit eines Datensatzes ohne *Pile-Up* für verschiedene b-Tagging-Algorithmen in einer Abbildung aufgetragen. In Abbildung 7.11(a) ist der Vergleich zwischen *IP2D* und *IP3D* gezeigt. *IP3D* reagiert vor allem für kleine b-Tagging-Effizienzen stärker auf *Minimum-Bias*-Spuren als *IP2D*. Spuren aus einem benachbarten *Minimum-Bias*-Vertex, die durch das Qualitätskriterium des lon-



**Abb. 7.10.:** Vergleich der b-Tagging-Leistungsfähigkeit von *IP2D* für Datensätze mit und ohne *Pile-Up*. Es werden nur wahre Signaljets berücksichtigt und es liegt keine Fehlidentifikation des Signalvertex vor. Die Vergleiche sind in Abhängigkeit der b-Tagging-Effizienz gezeigt (a,b) und für verschiedene Bereiche des transversalen Impulses  $p_{T,Jet}$  (c,d) und der Pseudorapidität  $\eta_{Jet}$  (e, f) der Jets. Die Unterdrückung in den Abbildungen (c,d,e,f) ist für eine b-Tagging-Effizienz von 60 % angegeben.

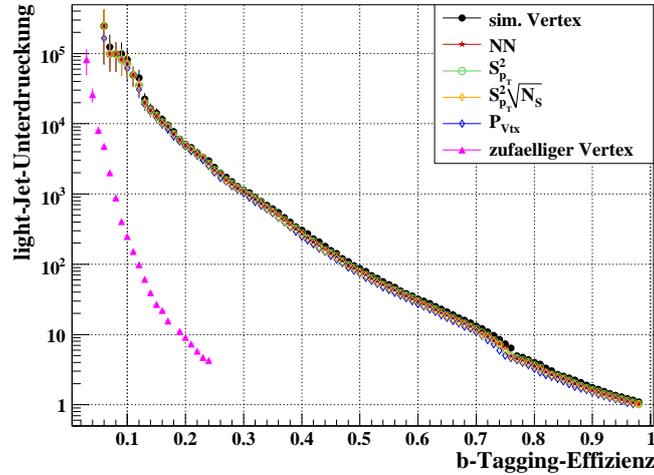


**Abb. 7.11.:** Vergleich des Einflusses von *Minimum-Bias*-Spuren auf die verschiedenen b-Tagging-Algorithmen. Gezeigt ist für jede Methode das Verhältnis der light-Jet-Unterdrückung in einem Datensatz mit *Pile-Up* zur light-Jet-Unterdrückung in einem Datensatz ohne *Pile-Up* gegen die b-Tagging-Effizienz. Es werden nur wahre Signaljets berücksichtigt und es liegt keine Fehlidentifikation des Signalvertex vor.

gitudinalen Impact-Parameters nicht entfernt worden sind, werden deutlich höhere longitudinale Impact-Parameter aufweisen als Spuren aus dem Signalvertex. Diese können sehr ähnlich aussehen wie Spuren aus einem b-Hadron-Zerfall. In transversaler Richtung wird dies nicht sichtbar, da hier aufgrund der sehr geringen Ausdehnung der Wechselwirkungsregion faktisch keine Unterscheidung zwischen den einzelnen Primärvertex-Positionen vorliegt. Somit ist *IP3D* stärker von *Minimum-Bias*-Spuren beeinflusst als *IP2D*. In Abbildung 7.11(b) werden die *IP3D*-, *SV1*- und *COMB*-Algorithmen verglichen. *SV1* zeigt die kleinste Beeinflussung durch *Minimum-Bias*-Spuren. Der *COMB*-Algorithmus ist ähnlich stark wie *IP3D* durch diesen Effekt beeinflusst.

### 7.2.2. Einfluss der Fehlidentifikation des Signalvertex

Die Unterschiede zwischen den Fehlidentifikationsraten der verschiedenen implementierten Signalvertex-Selektionsmethoden liegen im Falle eines *WH*-Datensatzes mit  $\langle N_{MB} \rangle = 6.9$  im Bereich von 1 – 2% (siehe Tabelle 6.1). Diese Unterschiede sind verhältnismäßig klein, um den Einfluss der Fehlidentifikation des Signalvertex auf die b-Tagging-Leistungsfähigkeit zu untersuchen. Um den Effekt auf den ersten Blick sichtbar zu machen, wurde neben der oben bereits erwähnten Selektion anhand der Simulationsinformation (in den Abbildungen mit *sim. Vertex* bezeichnet), die den Signalvertex immer richtig identifiziert, zusätzlich eine Selektion verwendet, die zufällig einen Vertex im Ereignis auswählt. Eine solche Vertex-Selektion hat eine sehr hohe Fehlidentifikationsrate und der Einfluss



**Abb. 7.12.:** Vergleich der Leistungsfähigkeit des *IP2D*-Algorithmus für verschiedene Signalvertex-Selektionsmethoden. Es werden nur wahre Signaljets berücksichtigt.

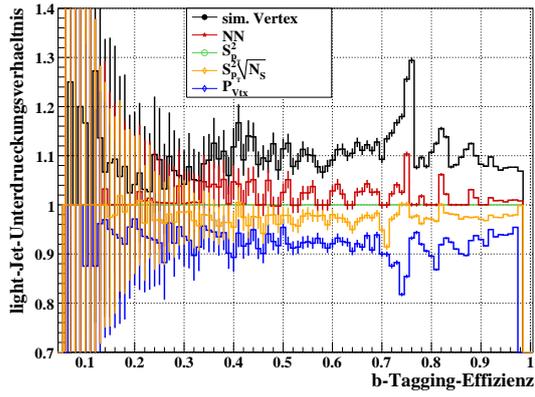
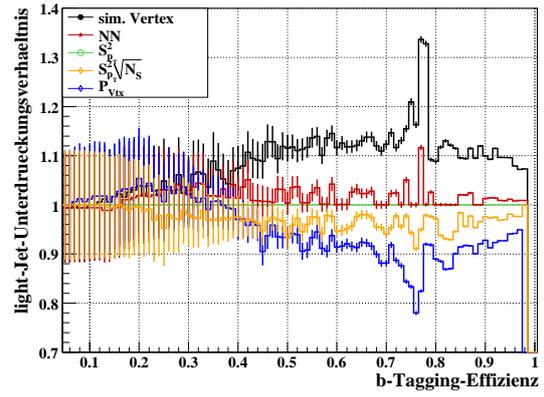
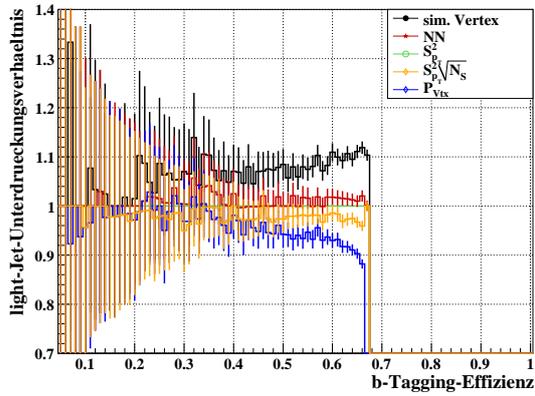
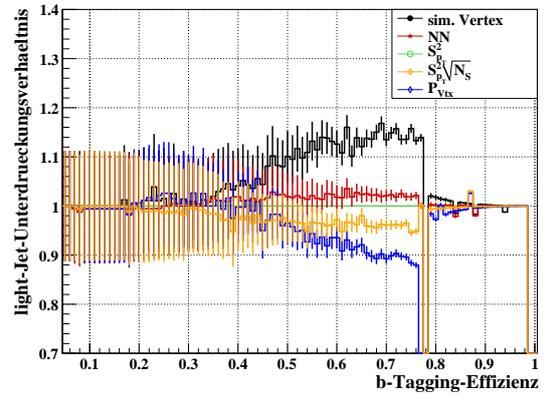
einer Fehlidentifikation sollte so offensichtlich sein.

In Abbildung 7.12 ist ein Leistungsvergleich der b-Jet-Identifikation, beispielhaft für *IP2D*, für die verschiedenen Selektionsmethoden des Signalvertex zu sehen. Die entsprechenden Abbildungen für die anderen betrachteten b-Tagging-Algorithmen befinden sich in Anhang A.2. Im Falle einer zufälligen Vertex-Selektion zeigt sich ein deutlicher Leistungseinbruch. Die Kurve bricht hier bei einer b-Tagging-Effizienz von ca. 25 % ab. Die Abbruchgrenze wird alleine durch die Vertex-Selektionseffizienz bestimmt. Hieraus lässt sich schließen, dass die b-Jet-Identifikation scheitert, wenn die falsche z-Position des Signalvertex vorliegt. Ursache dafür ist, dass bei einer Fehlidentifikation die Bedingung an die longitudinalen Impact-Parameter der Spuren auf die falsche z-Position bezogen wird, und somit die Spuren aus dem eigentlichen Signalprozess nicht mehr zur Identifikation der Jets zugelassen werden. Es sei angemerkt, dass in sehr seltenen Fällen ein fälschlicher Weise als Signalvertex selektierter *Minimum-Bias-Vertex* sehr nah am tatsächlichen Signalvertex liegt und die Spuren aus dem b-Hadron-Zerfall für das b-Tagging erhalten bleiben.

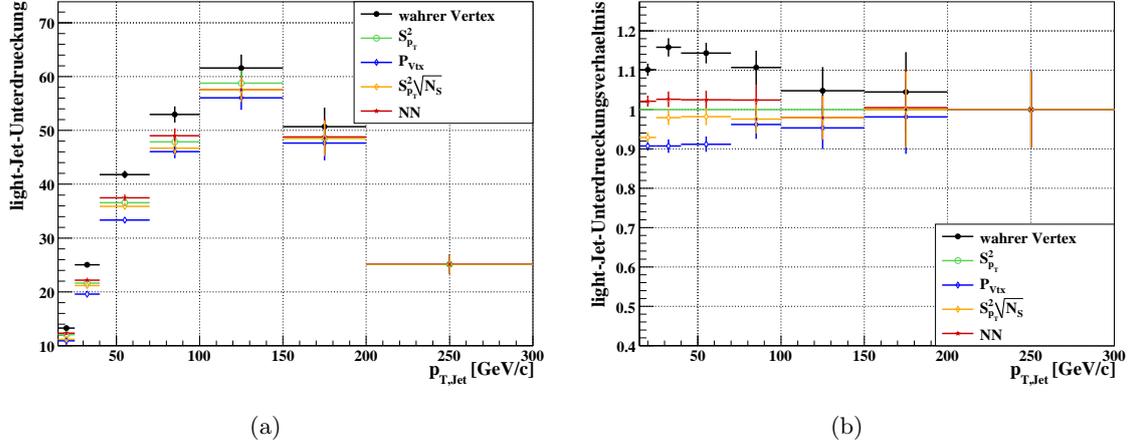
In Abbildung 7.12 sind die Unterschiede zwischen den anderen Signalvertex-Selektionsmethoden schwer zu erkennen. Diese werden besser sichtbar, wenn man die Verhältnisse der light-Jet-Unterdrückungen der verschiedenen Vertex-Selektionsmethoden zu der Unterdrückung einer Referenz-Selektionsmethode betrachtet. Als Referenz wird in den nachfolgenden Studien immer die Unterdrückung bei einer Signalvertex-Selektion mit der  $S_{p_T}^2$ -Methode verwendet. Abbildung 7.13 zeigt das Unterdrückungsverhältnis für alle zur Untersuchung stehenden b-Tagging-Algorithmen. Darin ist die Reihenfolge der Leistungs-

fähigkeiten zu beobachten, die aus den Signalvertex-Selektionseffizienzen, bzw. den Fehlidentifikationsraten (siehe Tabelle 6.1) zu erwarten ist. Die b-Tagging-Leistungsfähigkeit ist bei immer richtiger Signalvertex-Selektion am besten, gefolgt von der Vertex-Selektion anhand eines neuronalen Netzes und der  $\mathcal{S}_{pT}^2$ -Methode. Selektionen anhand der  $\mathcal{P}_{Vtx}$ -Methode und anhand des Selektionskriteriums  $\sum p_T^2 \sqrt{N_{Sp}}$  führen zu den geringsten light-Jet-Unterdrückungen. Der Abbruch der Kurven im Falle von *SV1* ist auf die limitierte Sekundärvertex-Rekonstruktionseffizienz von ca. 70 % zurückzuführen. Der Vergleich zwischen den b-Tagging-Algorithmen zeigt für alle Methoden ähnliche Leistungsunterschiede zwischen den verschiedenen Signalvertex-Selektionsmethoden. Der Einfluss einer Signalvertex-Fehlidentifikation wird durch die oben beschriebene Bedingung für die longitudinalen Impact-Parameter der Spuren dominiert. Diese Bedingung existiert für alle b-Tagging-Methoden, wobei die Bedingung im Falle von *SV1* etwas schwächer ist, als im Falle der Impact-Parameter-Algorithmen. Es wäre aus diesem Grund zu erwarten, dass der *SV1*-Algorithmus etwas weniger auf den Einfluss einer Fehlidentifikation des Signalvertex reagiert. Unterschiede dieser Art sind aber in den vorliegenden Studien nur schwer beobachtbar. In allen Fällen ist die b-Tagging-Leistung direkt abhängig von der Effizienz, mit der der Signalvertex richtig identifiziert werden kann. Im Falle der *COMB*-Methode ist für b-Tagging-Effizienzen oberhalb von  $\approx 78\%$  (Abb. 7.13(d)) zu beobachten, dass im Gegensatz zu den anderen Methoden die Unterschiede zwischen den verschiedenen Signalvertex-Selektionsmethoden abrupt kleiner werden. Die Ursache dafür ist eine ungünstige Art der Kombination von *SV1* und *IP3D*. Dies ist ein konzeptuelles Problem, welches in Verbindung mit der Fragestellung steht, wie mit Jets umgegangen wird, bei denen die Information zur Identifikation fehlt. Dies soll im nächsten Abschnitt erläutert werden.

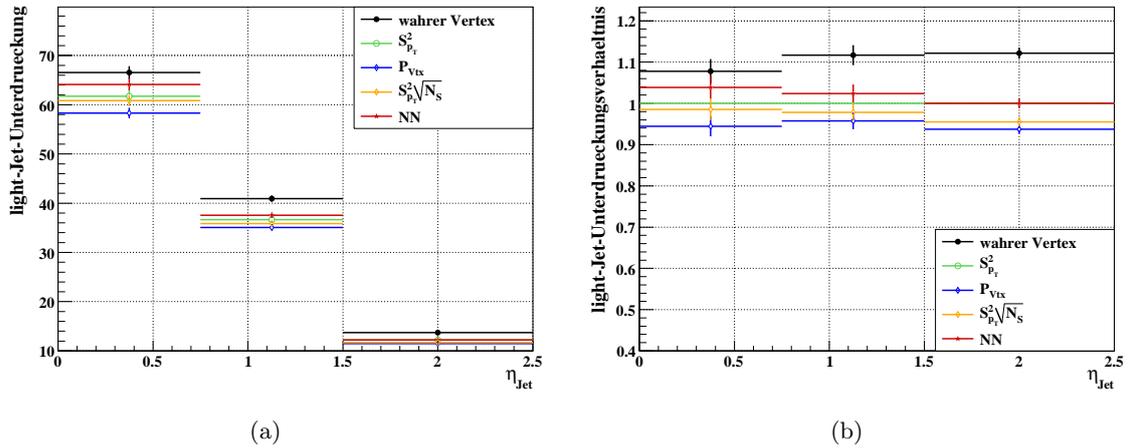
Die Abbildungen 7.14 und 7.15 zeigen die light-Jet-Unterdrückung in Abhängigkeit des transversalen Impulses und der Pseudorapidität der Jets für die verschiedenen Signalvertex-Selektionsmethoden für *IP2D*. Die entsprechenden Abbildungen für die anderen b-Tagging-Algorithmen befinden sich auch hier im Anhang A.2. Bei Betrachtung der Unterschiede in Abhängigkeit des transversalen Impulses ist zu beobachten, dass zu großen transversalen Impulsen hin die Selektionsmethode des Signalvertex nicht mehr ausschlaggebend ist. Dies ist zu erwarten, da ein Ereignis mit mindestens einem Jet mit sehr großem transversalen Impuls auch Spuren mit sehr großem transversalen Impuls aufweisen wird, und demnach von jeder Signalvertex-Selektionsmethode gut erkannt wird. Mit Ausnahme des sehr hohen transversalen Impulsbereichs sind für alle weiteren Bereiche des transversalen Impulses und für die ausgewählten Pseudorapiditätsbereiche die erwarteten Unterschiede zwischen den Signalvertex-Selektionsmethoden zu beobachten.

(a) *IP2D*(b) *IP3D*(c) *SV1*(d) *COMB*

**Abb. 7.13.:** Vergleich der Leistungsfähigkeit verschiedener b-Tagging-Algorithmen für verschiedene Vertex-Selektionsmethoden. Gezeigt ist das Verhältnis der light-Jet-Unterdrückung bezogen auf die light-Jet-Unterdrückung bei Verwendung der  $S^2_{pT}$ -Vertex-Selektion. Es werden nur wahre Signaljets berücksichtigt.



**Abb. 7.14.:** Vergleich der Leistungsfähigkeit der b-Tagging-Methode  $IP2D$  für verschiedene Vertex-Selektionsmethoden in Abhängigkeit des transversalen Jetimpulses  $p_{T,Jet}$  bei einer b-Tagging-Effizienz von 60%. Es werden nur wahre Signaljets berücksichtigt. Das Verhältnis (b) wird bezogen auf die Unterdrückung bei Verwendung einer  $S_{p_T}^2$ -Vertex-Selektion.

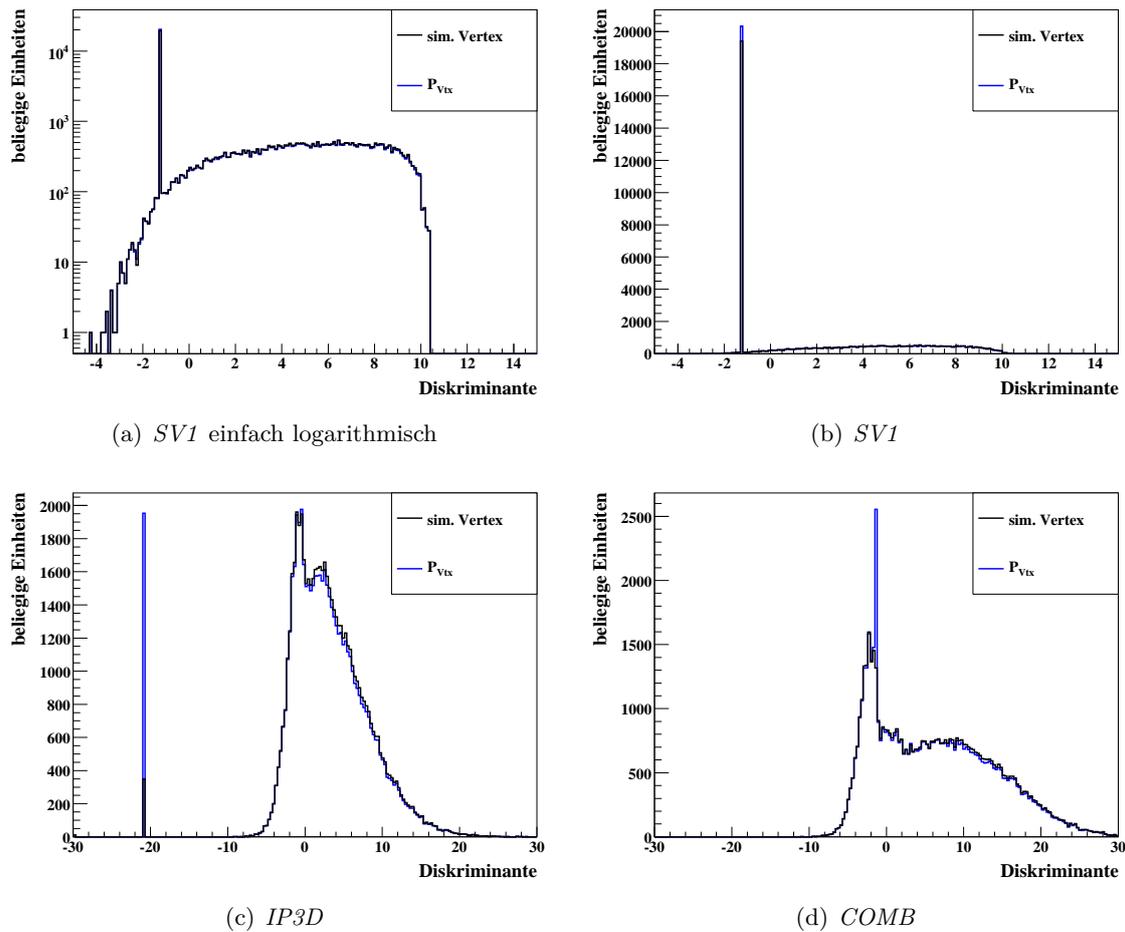


**Abb. 7.15.:** Vergleich der Leistungsfähigkeit der b-Tagging-Methode  $IP2D$  für verschiedene Vertex-Selektionsmethoden in Abhängigkeit der Pseudorapidität  $\eta_{Jet}$  bei einer b-Tagging-Effizienz von 60%. Es werden nur wahre Signaljets berücksichtigt. Das Verhältnis (b) wird bezogen auf die Unterdrückung bei Verwendung einer  $S_{p_T}^2$ -Vertex-Selektion.

## Umgang mit Jets ohne Spuren

Wie oben beschrieben, werden in allen hier betrachteten b-Tagging-Algorithmen verschiedene Eigenschaften der Jets über eine Likelihood-Methode in einer Diskriminanten kombiniert. Anhand dieser kann angegeben werden, wie wahrscheinlich ein Jet ein b- oder light-Jet ist. In Abbildung 7.16 ist die Verteilung der Diskriminanten von b-Jets für die *IP3D*-, *SV1*- und *COMB*-Algorithmen für eine Vertex-Selektion anhand der Information des simulierten Vertex und anhand der  $\mathcal{P}_{Vtx}$ -Methode verglichen. Für jede b-Tagging-Methode muss entschieden werden, wie mit Jets umgegangen wird, bei denen die Information fehlt, anhand derer ein Jet als b-Jet identifiziert werden kann. Im Falle von *IP3D* sind dies in erster Linie Jets, die keine Spuren zugeordnet bekommen und für die somit keine Impact-Parameter-Information vorliegt. Diese Jets werden im Falle von *IP3D* in einer Region der Diskriminanten angesiedelt, in der eindeutig light-Jets zu erwarten sind. Diese Ereignisse werden auf einen Wert von  $\approx -20$  gesetzt (siehe Abbildung 7.16(c)). Es zeigt sich hier ein deutlicher Unterschied zwischen den betrachteten Signalvertex-Selektionen. Fünfmal mehr Jets haben im Falle einer  $\mathcal{P}_{Vtx}$ -Signalvertex-Selektion keine Impact-Parameter-Information als im Falle einer Selektion, die keine Fehlidentifikation zulässt. Der Unterschied lässt sich durch Jets erklären, die aufgrund einer Fehlidentifikation des Signalvertex und der nachfolgenden longitudinalen Impact-Parameter-Bedingung, bezogen auf eine falsche Primärvertex-Position, keine Spuren zugeordnet bekommen. Es sei noch einmal darauf hingewiesen, dass in den vorliegenden Abbildungen lediglich b-Jets zu sehen sind. Diese kommen in erster Linie aus der Signalwechselwirkung. Die Wahrscheinlichkeit, dass diesen Jets aus anderen Gründen als der Fehlidentifikation des Signalvertex keine Spuren zugeordnet werden können, ist klein. Der große Unterschied in Abbildung 7.16(c) bei  $\approx -20$  macht deutlich sichtbar, dass Jets ohne Spuren in erster Linie auf eine solche Fehlidentifikation zurückzuführen sind.

Im Falle von *SV1* ist das Vorgehen anders. Hier werden Jets, in denen kein Sekundärvertex rekonstruiert werden kann, in einer Region angesiedelt, die innerhalb der Verteilung der Diskriminanten der b-Jets liegt (siehe hierzu Abbildung 7.16(a)). Die Ursache dafür, dass in einem Jet kein Sekundärvertex rekonstruiert werden kann, ist einerseits die limitierte Sekundärvertex-Rekonstruktionseffizienz selbst, andererseits existieren in Gegenwart von *Pile-Up* mit großer Wahrscheinlichkeit Jets, denen keine Spuren zugeordnet werden können. Auch für solche ist eine Sekundärvertex-Rekonstruktion unmöglich. Die Anhäufung bei  $\approx -1.5$  zeigt einen großen Anteil an b-Jets, die keinen rekonstruierbaren Sekundärvertex aufweisen, wobei dies nicht auf eine Fehlidentifikation des Signal-Vertex zurückgeführt werden kann, sondern von der limitierten Sekundärvertex-Rekonstruktion herrührt (siehe hierzu die Abbildung 7.16(b)). Nur ein kleiner Anteil an zusätzlichen Jets

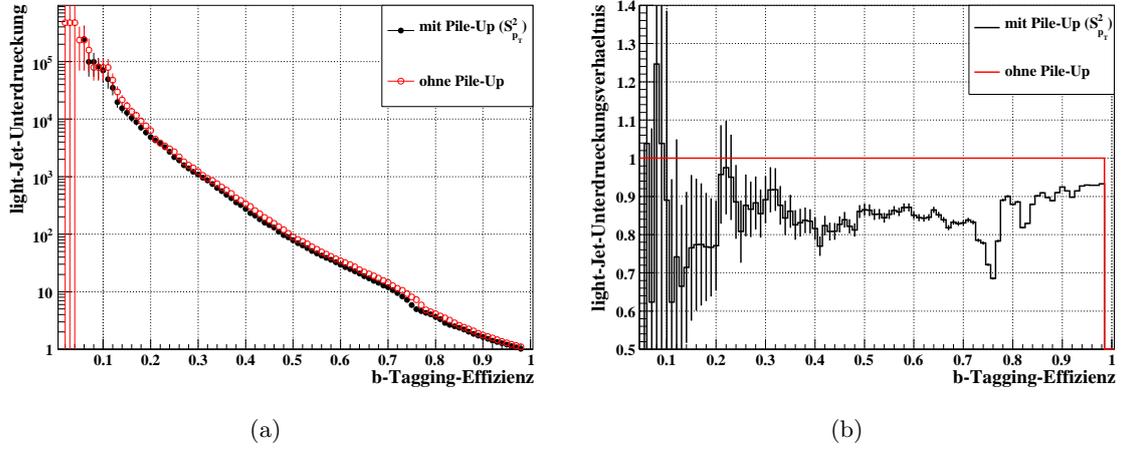


**Abb. 7.16.:** Diskriminante zur b-jet-Identifikation für verschiedene b-Tagging-Algorithmen, jeweils für eine Signalvertex-Selektion anhand der  $\mathcal{P}_{V_{tx}}$ -Methode und für die Methode anhand der Information der Simulation (*sim. Vertex*, hier ist keine Fehlidentifikation des Signalvertex möglich).

wird durch die Fehlidentifikation des Signalvertex verursacht. Kann kein Sekundärvertex rekonstruiert werden, existiert trotz allem eine gewisse Wahrscheinlichkeit, dass der Jet ein b-Jet ist. Aus diesem Grund wäre es wenig sinnvoll all diese Jets in einer Region mit großer light-Jet-Wahrscheinlichkeit anzusiedeln. Würde man alleine die *SV1*-Methode zur Identifikation von b-Jets verwenden und läge kein *Pile-Up* vor, wäre die vorliegende Taktik ein sinnvolles Vorgehen, in der Kombination mit *IP3D* entstehen aber Probleme. Derzeit wird für die *COMB*-Methode für einen Jet, der keine Impact-Parameter-Information, also in erste Linie keine zugeordneten Spuren besitzt, einfach die *SV1*-Diskriminante verwendet. Diese Jets sind aus diesem Grund nicht am linken Rand der *COMB*-Diskriminante (Abbildung 7.16(d)), sondern innerhalb der Verteilung zu finden. Es zeigt sich hier eine deutliche Anhäufung bei ca.  $-0.5$  für eine Signalvertex-Selektion durch die  $\mathcal{P}_{Vtx}$ -Methode. Diese Jets sind genau diejenigen, die aufgrund einer Fehlselektion des Signalvertex keine Spuren zugeordnet bekommen. An dieser Abbildung lässt sich leicht erklären, warum die Unterschiede der b-Tagging-Leistungsfähigkeit im Falle der *COMB*-Methode für große b-Tagging-Effizienzen verschwinden (siehe Abbildung 7.13(d)). Für sehr große b-Tagging-Effizienzen liegt die Bedingung, für eine Identifikation als b-Jets, bei einem kleinen Wert der Diskriminanten. Verschiebt man die Bedingung zu höheren Werten der Diskriminanten, wird die b-Tagging-Effizienz kleiner, da mehr b-Jets durch die Schnittbedingung nicht als b-Jets erkannt werden. Die Unterschiede zwischen den verschiedenen Signalvertex-Selektionsmethoden werden erst sichtbar, wenn die Bedingung einen Wert  $> -0.5$  der Diskriminanten überschreitet. Da dies im Falle der *COMB*-Methode erst nach einer gewissen b-Tagging-Effizienz geschieht, sind oberhalb einer bestimmten b-Tagging-Effizienz keine Unterschiede zu erkennen. Dies ist keineswegs sinnvoll, da in dieser Form im Falle von richtiger Selektion des Signalvertex die Jets aus *Minimum-Bias*-Wechselwirkungen, die aus gutem Grund keine Spuren zugeordnet bekommen, nur unter Verlust einer gewissen b-Tagging-Effizienz unterdrückt werden können. Selbiges gilt in Gegenwart von *Pile-Up* auch für *SV1*. Es wäre sehr zweckmäßig, hierfür eine andere Lösung zu finden. Es ist beispielsweise denkbar, dass die Jets, bei denen kein Sekundärvertex rekonstruiert werden kann, weil keine Spuren im Jet vorhanden sind, von *SV1* anders behandelt werden als Jets bei denen aus anderen Gründen kein Sekundärvertex rekonstruierbar ist.

### 7.2.3. Kombination der beiden *Pile-Up*-Einflüsse

In der Realität wird eine Kombination aus dem Einfluss der Signalvertex-Fehlidentifikation und dem Einfluss der Spuren aus *Minimum-Bias*-Wechselwirkungen zu beobachten sein. Um den Gesamteffekt abschätzen zu können, soll hier ein weiterer Vergleich angestellt werden. Es werden wieder Datensätze mit und ohne *Pile-Up* verglichen, wobei diesmal im Falle des *Pile-Up*-Datensatzes eine in der Realität verwendbare Signalvertex-Selektion,

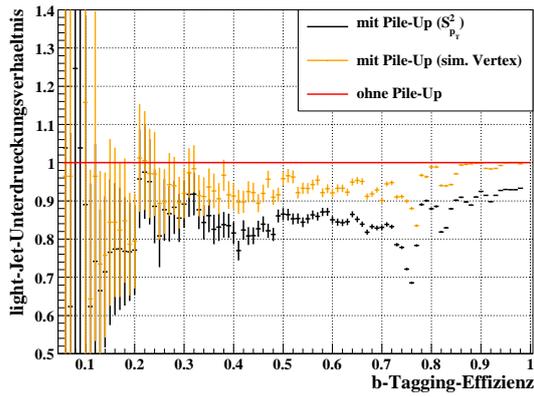
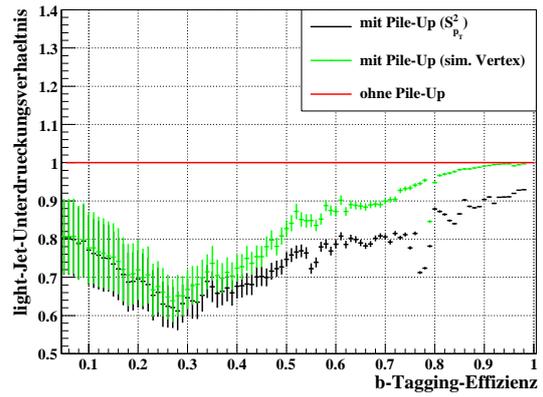
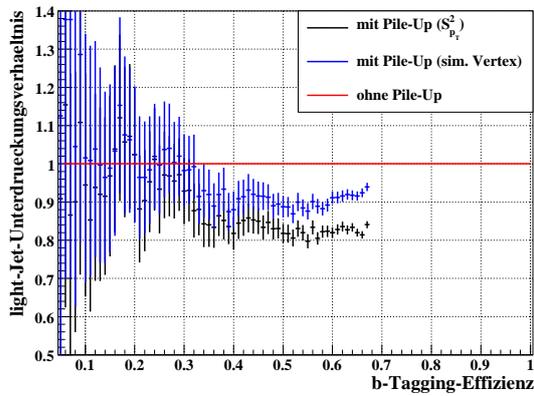
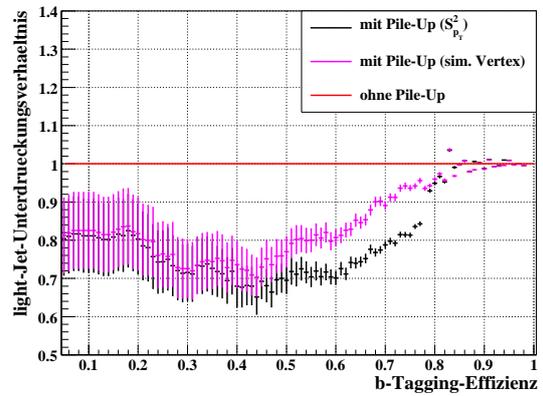


**Abb. 7.17.:** Vergleich der b-Tagging-Leistungsfähigkeit für Datensätze mit und ohne *Pile-Up*. Im Falle des *Pile-Up*-Datensatzes wurde der Signalvertex anhand der  $S_{p_T}^2$ -Methode selektiert. Es werden nur wahre Signaljets berücksichtigt.

die  $S_{p_T}^2$ -Methode, angewendet wird. Auch hier wurden nur wahre Signaljets berücksichtigt, so dass keine künstliche Verbesserung der Unterdrückung durch *Minimum-Bias*-Jets vorliegt. In Abbildung 7.17 ist beispielhaft für *IP2D* die light-Jet-Unterdrückung gegen die b-Tagging-Effizienz für diesen Vergleich gezeigt. Aufgrund der nicht verschwindenden Fehlidentifikationsrate der  $S_{p_T}^2$ -Methode und der zusätzlichen Spuren aus *Minimum-Bias*-Wechselwirkungen, die im Falle des Datensatzes ohne *Pile-Up* nicht vorhanden sind, werden so beide *Pile-Up*-Effekte berücksichtigt. In Anhang A.3 sind die gleichen Abbildungen für die anderen betrachteten b-Tagging-Algorithmen zu sehen.

Die Verhältnisse der light-Jet-Unterdrückung, wie in Abbildung 7.17(b) gezeigt, können jetzt mit den Studien aus Abschnitt 7.2.1 verglichen werden. Ein solcher Vergleich ist in Abbildung 7.18 zu sehen. Gezeigt ist hier einerseits das Verhältnis der light-Jet-Unterdrückung eines Datensatzes mit *Pile-Up*, wobei der Signalvertex immer richtig selektiert wurde (*sim. Vertex*, in schwarz), zum anderen das Verhältnis der light-Jet-Unterdrückung eines Datensatzes mit *Pile-Up*, wobei der Signalvertex anhand der  $S_{p_T}^2$ -Methode selektiert wurde und somit eine gewisse Fehlidentifikation des Signalvertex vorliegt ( $S_{p_T}^2$ , in Farbe). Die Verhältnisse werden immer in Bezug auf die Unterdrückung berechnet, die mit einem Datensatz ohne *Pile-Up* erreicht wird.

Aus den vorliegenden Abbildungen kann nun eine Abschätzung gemacht werden, welche Größenordnung die beiden betrachteten *Pile-Up*-Effekte auf die jeweiligen b-Tagging-Algorithmen haben werden. Beispielhaft sei dies hier für die *IP3D*-Methode (siehe Ab-

(a) *IP2D*(b) *IP3D*(c) *SV1*(d) *COMB*

**Abb. 7.18.:** Vergleich des Einflusses auf die b-Tagging-Effizienz alleine durch Spuren aus *Minimum-Bias*-Wechselwirkung (in Farbe) sowie durch Spuren aus *Minimum-Bias*-Wechselwirkungen und die Fehlidentifikation des Signalvertex (in schwarz).

bildung 7.18(b)) für eine b-Tagging-Effizienz von 60 % angegeben: Durch Verunreinigung der Signaljets durch Spuren aus *Minimum-Bias*-Wechselwirkungen ist zu erwarten, dass nur eine um 10 % geringere light-Jet-Unterdrückung im Vergleich zum Fall ohne *Pile-Up* erreicht werden kann. Zusätzlich wird die Unterdrückung durch die Fehlidentifikation des Signalvertex (Fehlidentifikationsrate von 3.8 % für die angewendete  $\mathcal{S}_{pT}^2$ -Methode, siehe Tabelle 6.1) in Gegenwart von *Pile-Up* um weitere 10 % verringert.

In den vorliegenden Studien ist durch die Verwendung der wahren Signaljets vernachlässigt worden, dass sich diese von Jets unterscheiden, die aus rekonstruierten Teilchen bestimmt werden, auch wenn diese Jets ebenfalls von der Signalwechselwirkung ausgehen. Es ist zu erwarten, dass in diesem Fall Spuren aus *Minimum-Bias*-Wechselwirkungen einen Einfluss auf die Bestimmung der Jetrichtung und damit auf die b-Tagging-Leistungsfähigkeit nehmen werden. Will man diese Effekte berücksichtigen, müssen die Jets aus den *Minimum-Bias*-Wechselwirkungen wieder hinzugezogen werden. Hierdurch stößt man aber schnell auf das Problem einer künstlichen Verbesserung der Unterdrückung. Der Grund, warum es mit einfachen Mitteln derzeit nicht möglich ist, den Einfluss durch *Minimum-Bias*-Spuren auf die Jetbestimmung zu untersuchen, ohne dabei die künstliche Verbesserung durch *Minimum-Bias*-Jets in Kauf nehmen zu müssen, ist, dass bei ATLAS keine Information zur Verfügung steht, die angibt, ob ein Jet ein Signal oder *Minimum-Bias*-Jet ist. Aus diesem Grund wurde dieser dritte *Pile-Up*-Effekt in dieser Arbeit nicht untersucht.

Das b-Tagging ist in einem gewissen Maß in der Lage, zu erkennen, ob der Signalvertex richtig selektiert worden ist. Prinzipiell ist es aus diesem Grund denkbar, diese Information zu nutzen, um die Signalvertex-Selektion zu verbessern und gleichzeitig das b-Tagging gegen *Pile-Up*-Effekte robuster zu machen. Beispielsweise ist es möglich, dass das b-Tagging für alle rekonstruierten Primärvertizes durchgeführt wird. Erst nachträglich könnte dann entschieden werden, welcher Vertex am wahrscheinlichsten der Signalvertex ist. Diese Information müsste dann mit der aus der spurbasierten Signalvertex-Selektion erhaltenen kombiniert werden. Stellt sich ein anderer Vertex als ein besserer Signalvertex-Kandidat heraus, ist es möglich, dass die Vertex-Kollektion durch das b-Tagging erneut umsortiert wird. Ebenso ist es denkbar, den folgenden Analysen die Information mehrerer Vertex-Kandidaten zur Verfügung zu stellen, so dass für jede Analyse individuell entschieden werden kann, welcher Vertex den besten Signalvertex-Kandidaten darstellt.



## 8. Zusammenfassung

In der vorliegenden Arbeit wurden detaillierte Studien zu Effekten von *Pile-Up* beim ATLAS-Experiment vorgestellt, wobei die Signalvertex-Selektion im Fokus der Untersuchungen stand. Erstmals wurden hier ausführlich unterschiedliche Methoden zur Signalvertex-Selektion verglichen. Die verschiedenen Selektionsmethoden wurden an  $H \rightarrow \gamma\gamma$ -,  $WH$ - und  $t\bar{t}$ -Ereignissen getestet. Diese sind aufgrund ihrer verschiedenartigen Topologien in Bezug auf die Signalvertex-Selektion unterschiedlich anspruchsvoll. So wird bei ca. sieben zusätzlichen *Minimum-Bias*-Wechselwirkungen eine  $t\bar{t}$ -Wechselwirkung fast nie falsch identifiziert. Eine  $WH$ -Wechselwirkung wird dagegen in einigen Prozent der Fälle und eine  $H \rightarrow \gamma\gamma$ -Wechselwirkung in etwa 40 % der Fälle nicht richtig erkannt. Die Identifikationseffizienz hängt zusätzlich in hohem Maße vom Kriterium ab, anhand dessen die Signalwechselwirkung unter allen Wechselwirkungen im Ereignis ausgewählt wird. In dieser Arbeit wurden Kriterien getestet, die zur Identifikation Spur- und Vertex-Eigenschaften heranziehen. Dazu werden die transversalen Impulse der Spuren am jeweiligen Vertex, die Spur-Multiplizität und die Sphärizität des Vertex ausgenutzt. Die verschiedenen Kriterien unterschieden sich dabei in der Art, wie die Information über die verschiedenen Eigenschaften kombiniert wird. Es wurden unter anderem mehrere multivariate Analysemethoden, wie ein Likelihood-Verhältnis, eine Vertex-Wahrscheinlichkeit und ein neuronales Netz getestet. Wider Erwarten war nur eine verhältnismäßig geringe Identifikationsfähigkeit durch das Likelihood-Verhältnis und die Vertex-Wahrscheinlichkeit zu beobachten. Es konnte gezeigt werden, dass dies auf die Korrelationen der verwendeten Eigenschaften zurückzuführen ist. Das neuronale Netz erreichte aus diesem Grund die höchste Identifikationsleistung. Die  $\mathcal{S}_{p_T}^2$ -Methode, die als Kriterium die Summe der quadrierten transversalen Impulse verwendet, lieferte ebenfalls eine sehr gute Leistungsfähigkeit, und wird heute als Standard-Selektionsmethode in der ATLAS-Software verwendet.

Der Anteil an falsch identifizierten Wechselwirkungen für die zuvor verwendete Standardmethode liegt im Falle einer  $WH$ -Wechselwirkung für im Mittel sieben zusätzliche *Minimum-Bias*-Wechselwirkungen bei  $(4.5 \pm 0.1) \%$ . Die  $\mathcal{S}_{p_T}^2$ -Methode erreicht hier dagegen einen Wert von  $(3.8 \pm 0.1) \%$  und das neuronale Netz reduziert die Fehlidentifikation weiter auf  $(3.4 \pm 0.1) \%$ .

Ein neues Rekonstruktionswerkzeug zur Selektion des Signalvertex wurde in die ATLAS-

Software ATHENA implementiert und ist nun Teil jeder Standard-Ereignisrekonstruktion. Dieses Werkzeug macht es dem Benutzer nun möglich, eine Selektionsmethode unabhängig vom verwendeten Primärvertex-Rekonstruktionsalgorithmus auszuwählen. Vor der Implementierung des neuen Selektionswerkzeugs war das Selektionskriterium an die jeweilige Primärvertex-Rekonstruktionsmethode geknüpft und eine unabhängige Wahl durch den Benutzer war nicht vorgesehen. Einige der zuvor getesteten Methoden sind dem Benutzer hierfür zur Verfügung gestellt. Ein weiterer Vorteil der Art der Implementierung ist, dass sie ein sehr flexibles Hinzufügen möglicher neuer Selektionsmethoden erlaubt.

Unter Verwendung des Selektionswerkzeugs ist es mit technisch einfachen Mitteln möglich, die Auswirkungen einer Fehlidentifikation des Signalvertex auf die nachfolgenden Rekonstruktionsalgorithmen und Analysen zu testen. Beispielfhaft wurden in dieser Arbeit verschiedene Methoden zur b-Jet-Identifikation daraufhin untersucht. Es gelang in diesen Studien erstmalig, den Effekt durch die Fehlidentifikation des Signalvertex von anderen *Pile-Up*-Effekten vollkommen zu entkoppeln. Es zeigte sich für alle betrachteten b-Tagging-Methoden ein eindeutiges Bild: Die Fehlidentifikation des Signalvertex hat zur Folge, dass b-Jets nur sehr eingeschränkt identifiziert werden können. Die Ursache dafür ist, dass nur Spuren aus dem selektierten Primärvertex zur Identifikation zugelassen werden und bei einer Fehlidentifikation des Signalvertex die Spuren von Interesse aus diesem Grund von den b-Tagging-Algorithmen nicht berücksichtigt werden.

Es war im Rahmen dieser Studien auch möglich, die Beeinflussung der b-Jet-Identifikation durch einen weiteren *Pile-Up*-Effekt zu untersuchen: die Verunreinigung der Jets der Signalwechselwirkung durch Spuren aus *Minimum-Bias*-Wechselwirkungen. Hier zeigten sich für die verschiedenen getesteten b-Tagging-Algorithmen unterschiedliche Reaktionen. Algorithmen, die den longitudinalen Impact-Parameter zu b-Jet-Identifikation verwenden, wie der *IP3D*- und der *COMB*-Algorithmus, sind von diesem Effekt stärker beeinflusst als solche, die von diesem unabhängig bleiben, wie der *IP2D*- und der *SV1*-Algorithmus.

In der Realität wird eine Kombination der beiden Effekte zu einem Leistungsabfall der b-Jet-Identifikation führen. So ergibt sich beispielsweise für den *IP3D*-Algorithmus aus den vorliegenden Studien folgende Abschätzung: Im Falle von *WH*-Ereignissen ist bei einer b-Tagging-Effizienz von 60 % damit zu rechnen, dass, bei ca. sieben zusätzlichen *Minimum-Bias*-Wechselwirkungen nur ca. 80 % der Unterdrückung von light-Jets erreicht wird, die für Ereignisse ohne *Pile-Up* erwartet wird. Nicht untersucht worden ist der Einfluss auf die Rekonstruktion der Jets durch zusätzlichen *Minimum-Bias*-Wechselwirkungen.

Alle hier durchgeführten Studien sind dabei für Szenarien betrachtet worden, die schon zu Beginn des LHC-Betriebs erwartet werden. Für die geplante Endenergie und Endluminosität ist mit deutlich stärkeren Auswirkungen von *Pile-Up* zu rechnen. Die vorliegende Studie ist dennoch schon in einem sehr frühen Stadium relevant, da von Beginn der Da-

tennahme an mit *Pile-Up* zu rechnen ist. Es ist daher sehr wichtig, schon im Vorfeld zu verstehen, welche Menge an *Pile-Up* man erwartet und welchen Einfluss dies auch schon auf frühe Analysen haben wird. Die Existenz von *Pile-Up* schon zu Beginn des LHC-Betriebs macht es außerdem möglich, die vorliegenden Studien sehr bald auf realen Daten zu überprüfen.

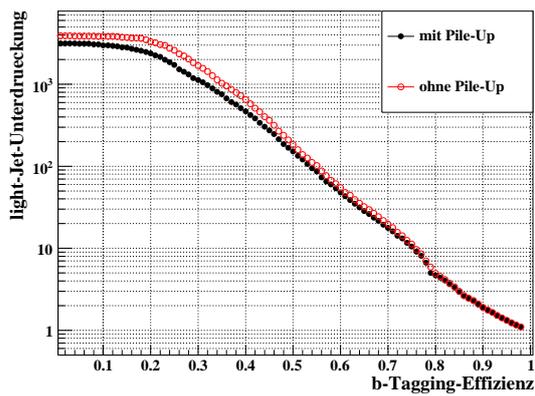


# Anhang

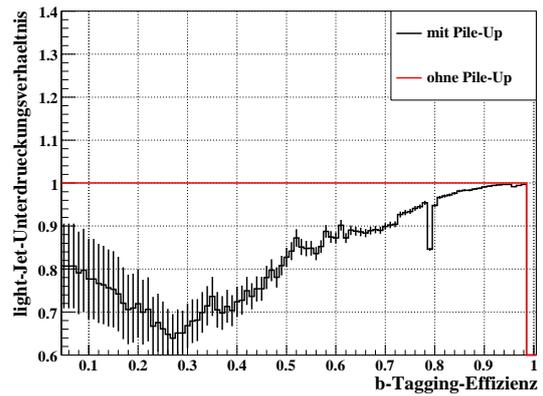


# A. b-Tagging Abbildungen

## A.1. Einfluss von *Minimum-Bias*-Spuren auf die b-Tagging-Leistung

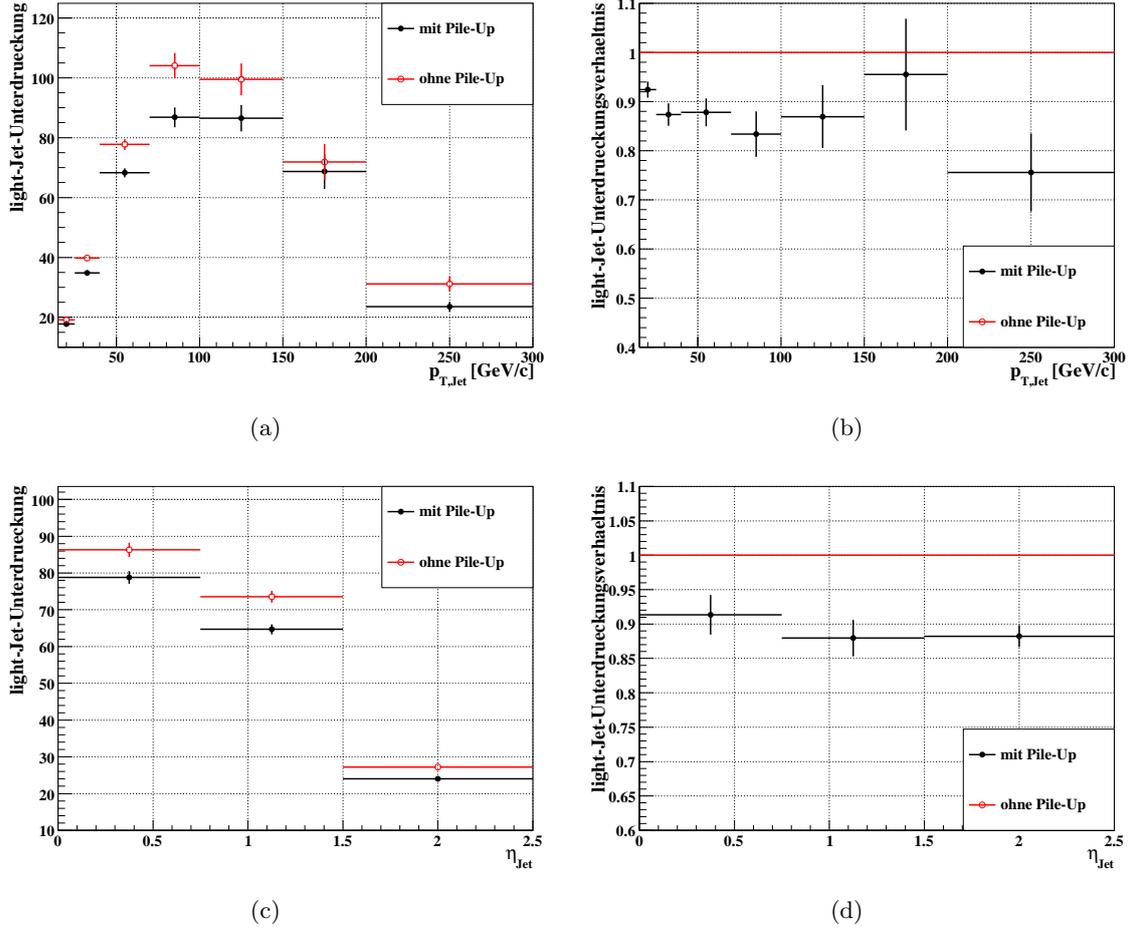


(a)

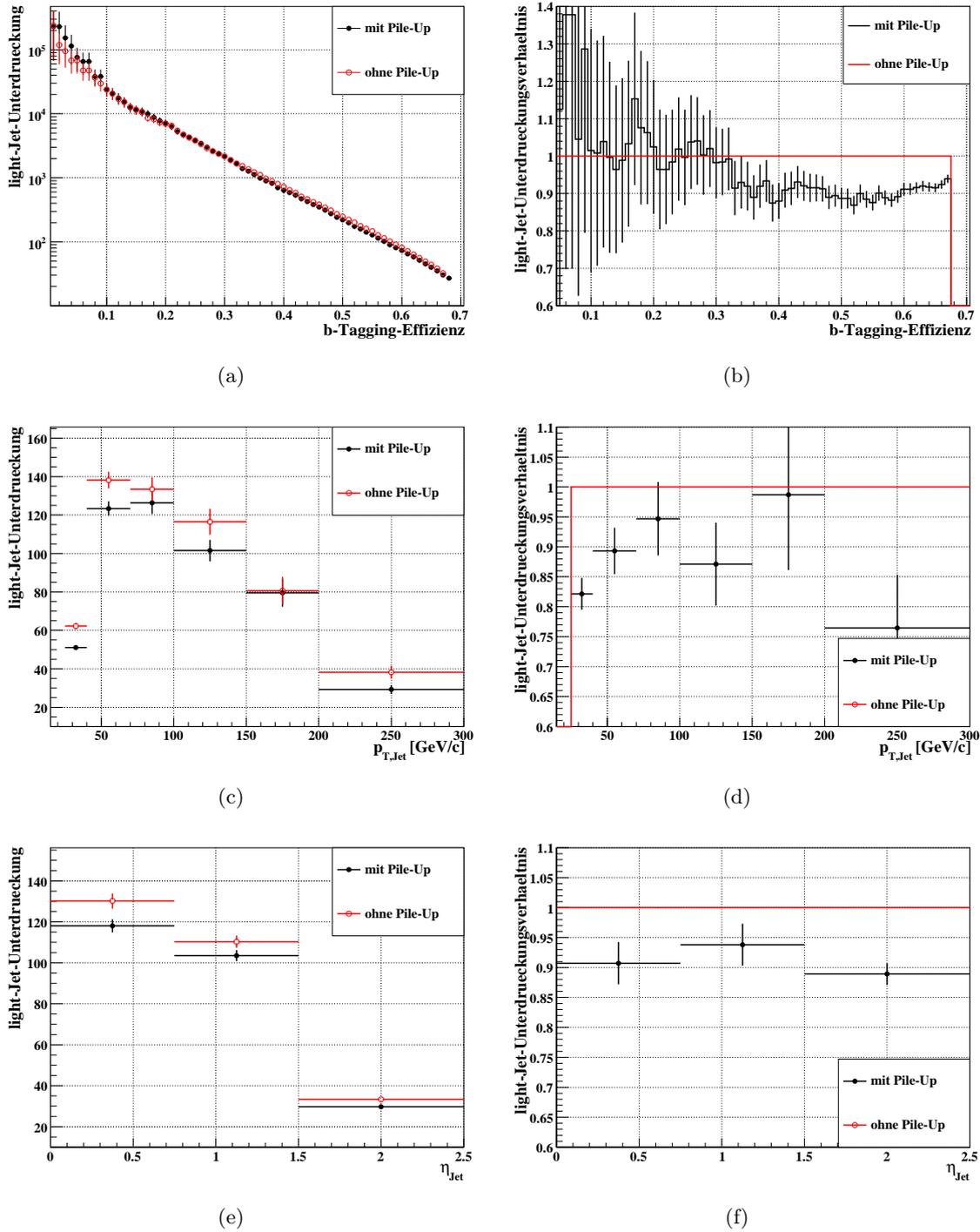


(b)

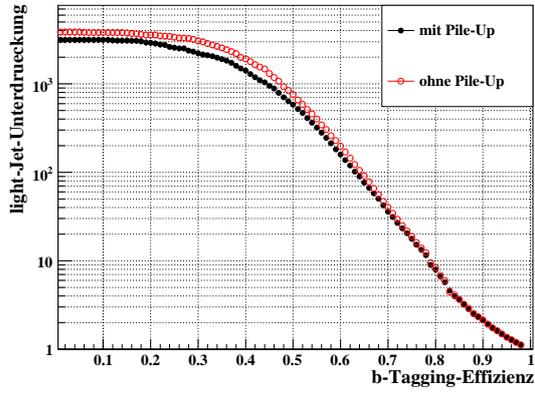
**Abb. A.1.:** Vergleich der b-Tagging-Leistungsfähigkeit von *IP3D* für Datensätze mit und ohne *Pile-Up*. Es werden nur wahre Signaljets berücksichtigt und es liegt keine Fehlidentifikation des Signalvertex vor. Die Vergleiche sind in Abhängigkeit der B-Tagging-Effizienz gezeigt.



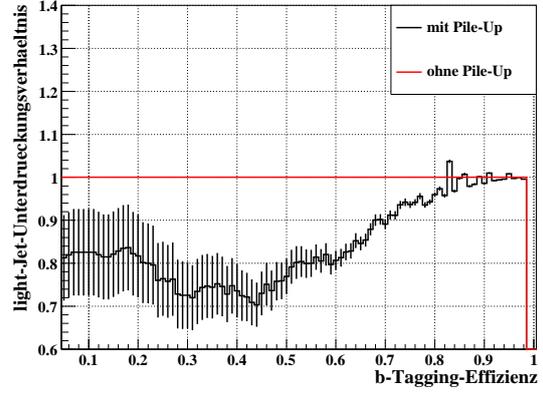
**Abb. A.2.:** Vergleich der b-Tagging-Leistungsfähigkeit von *IP3D* für Datensätze mit und ohne *Pile-Up*. Es werden nur wahre Signaljets berücksichtigt und es liegt keine Fehlidentifikation des Signalvertex vor. Die Vergleiche sind für verschiedene Bereiche des transversalen Impulses der Jets  $p_{T,Jet}$  (a,b) und der Pseudorapidität  $\eta$  (c,d). Die Unterdrückung ist für eine b-Tagging-Effizienz von 60% angegeben.



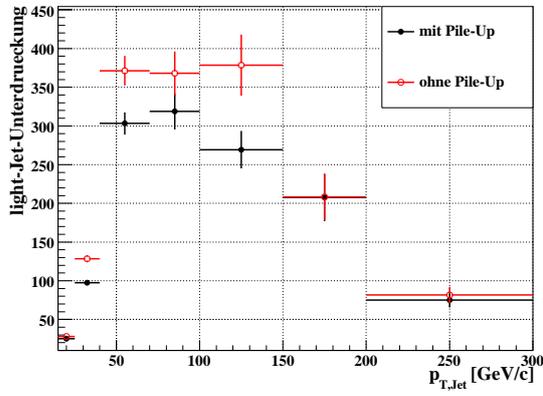
**Abb. A.3.:** Vergleich der b-Tagging-Leistungsfähigkeit von *SV1* für Datensätze mit und ohne *Pile-Up*. Es werden nur wahre Signaljets berücksichtigt und es liegt keine Fehlintifikation des Signalvertex vor. Die Vergleiche sind in Abhängigkeit der b-Tagging-Effizienz gezeigt (a,b) und für verschiedene Bereiche des transversalen Impulses  $p_{T,Jet}$  (c,d) und der Pseudorapidität  $\eta_{Jet}$  (e, f) der Jets. Die Unterdrückung in den Abbildungen (c,d,e,f) ist für eine b-Tagging-Effizienz von 60 % angegeben.



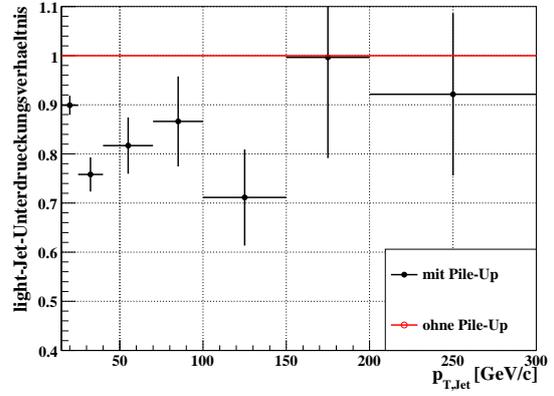
(a)



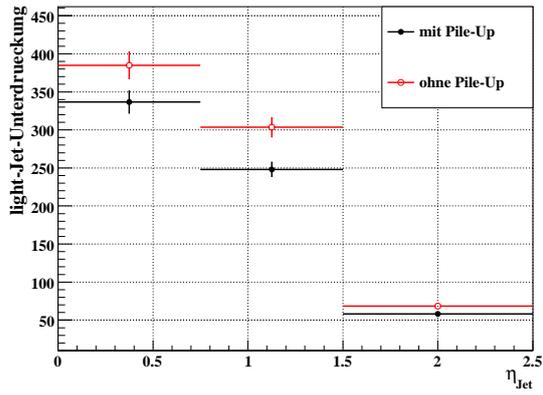
(b)



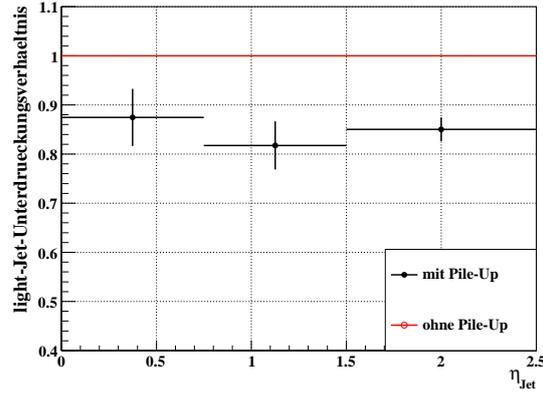
(c)



(d)



(e)



(f)

**Abb. A.4.:** Vergleich der b-Tagging-Leistungsfähigkeit von *COMB* für Datensätze mit und ohne *Pile-Up*. Es werden nur wahre Signaljets berücksichtigt und es liegt keine Fehlidentifikation des Signalvertex vor. Die Vergleiche sind in Abhängigkeit der b-Tagging-Effizienz gezeigt (a,b) und für verschiedene Bereiche des transversalen Impulses  $p_{T,Jet}$  (c,d) und der Pseudorapidität  $\eta_{Jet}$  (e, f) der Jets. Die Unterdrückung in den Abbildungen (c,d,e,f) ist für eine b-Tagging-Effizienz von 60 % angegeben.

## A.2. Einfluss der Fehlidentifikation des Signalvertex auf die b-Tagging-Leistung

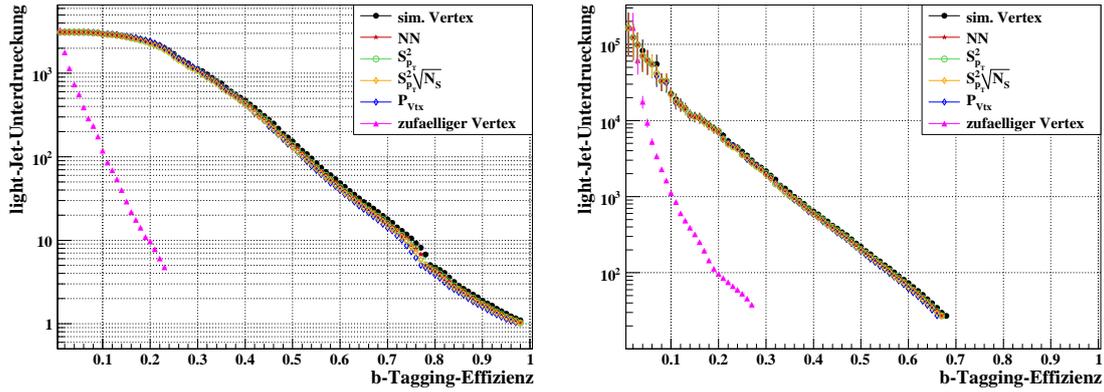
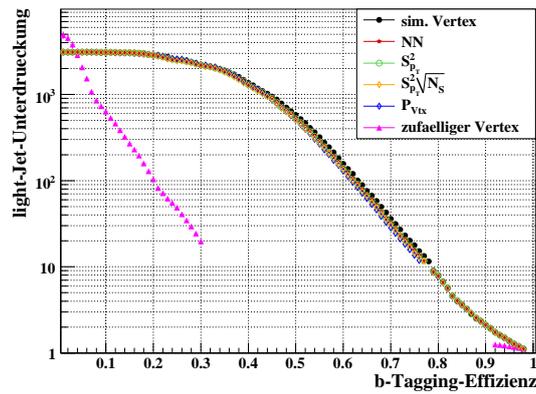
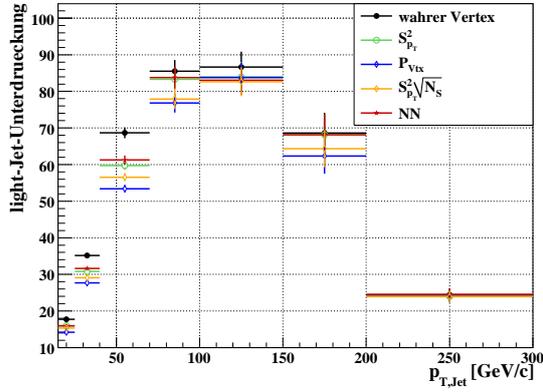
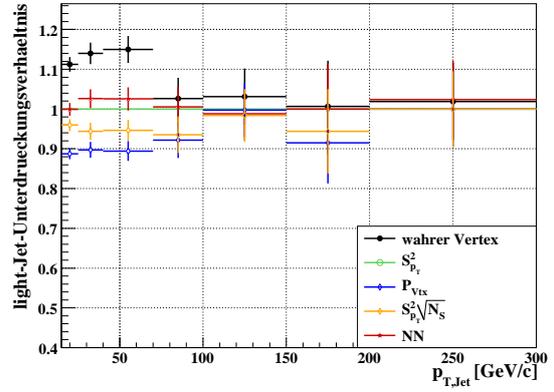
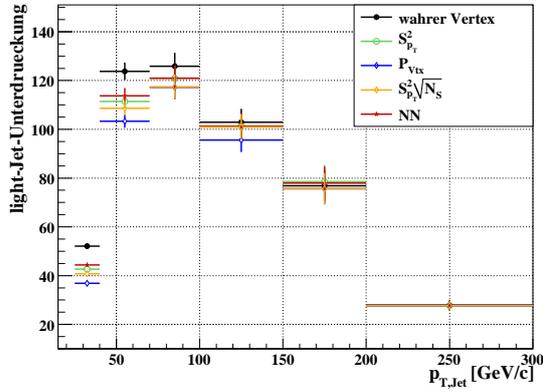
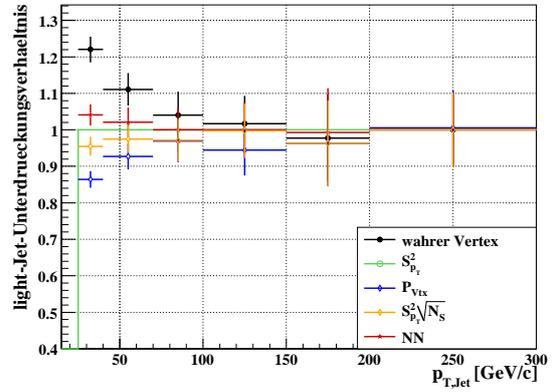
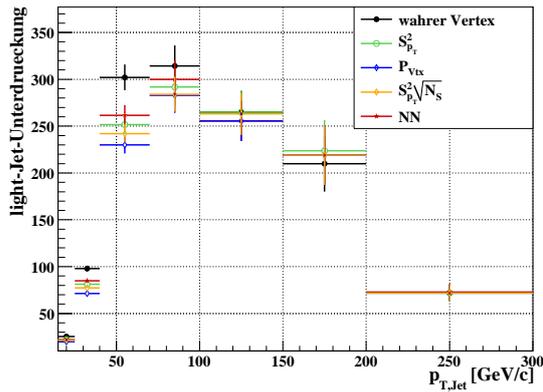
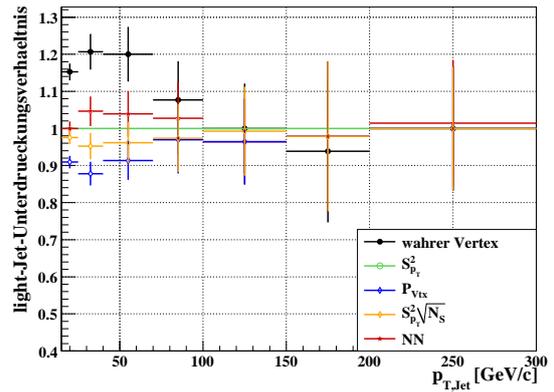
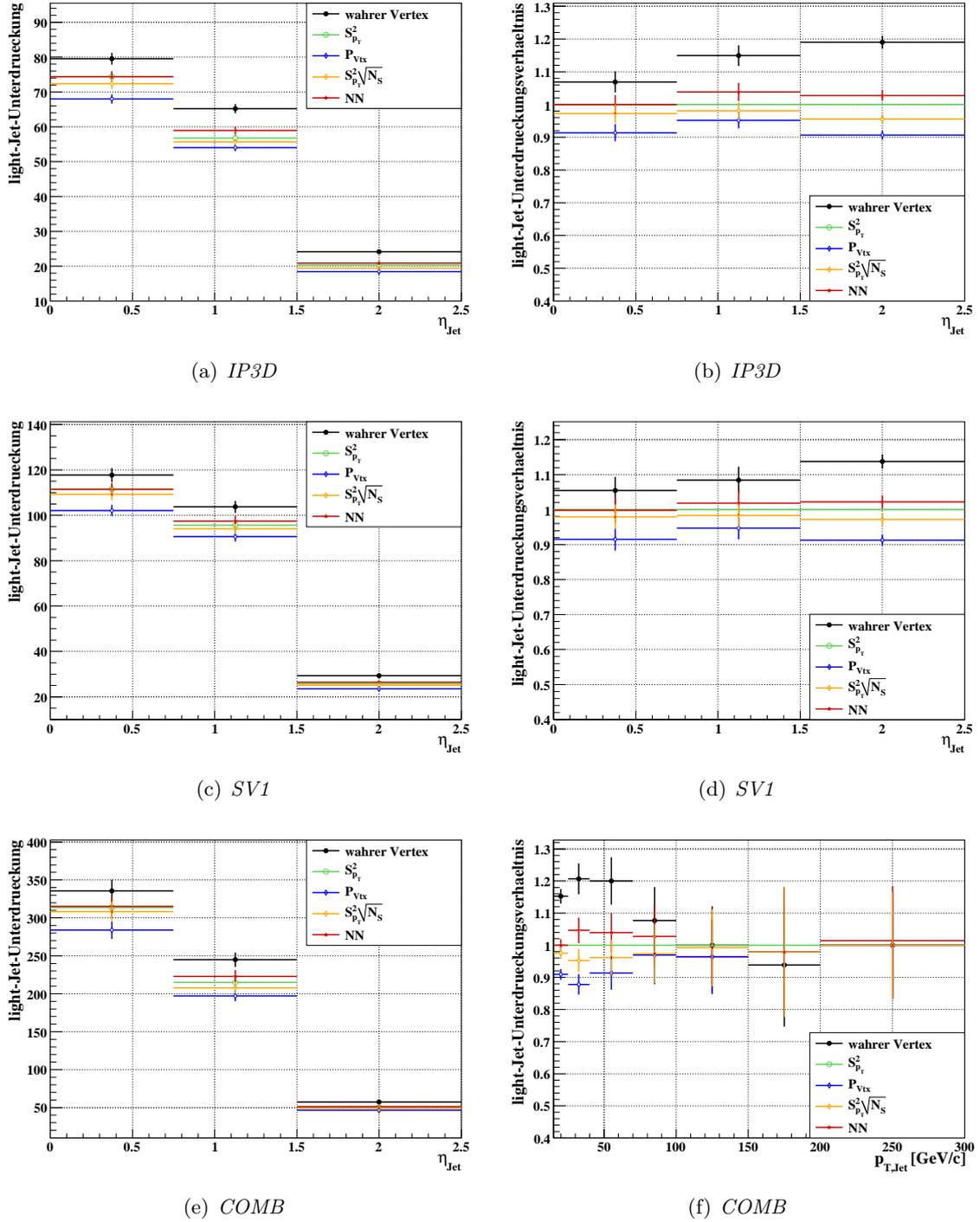
(a) *IP3D*(b) *SV1*(c) *COMB*

Abb. A.5.: Vergleich der Leistungsfähigkeit verschiedener B-Tagging-Methoden für verschiedene Signalvertex-Selektionsmethoden. Es werden nur wahre Signaljets berücksichtigt.

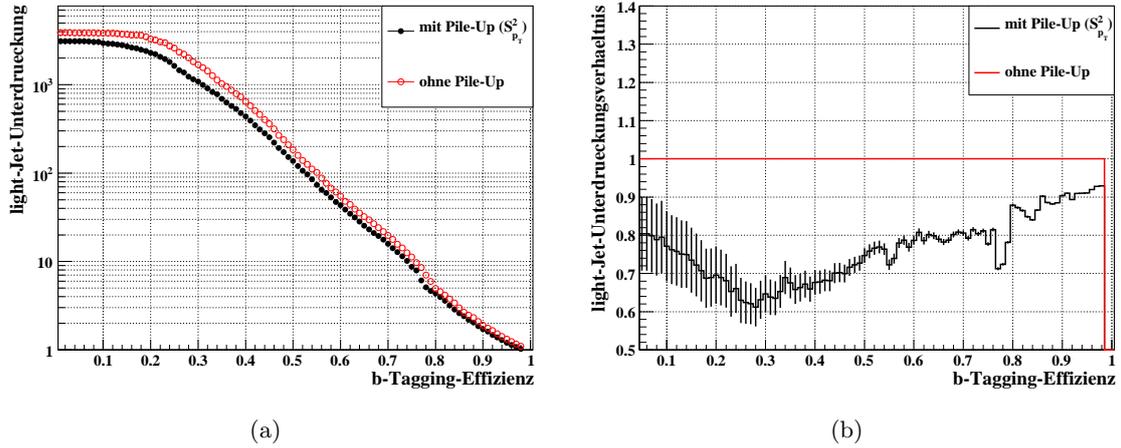
(a) *IP3D*(b) *IP3D*(c) *SV1*(d) *SV1*(e) *COMB*(f) *COMB*

**Abb. A.6.:** Vergleich der Leistungsfähigkeit verschiedener B-Tagging-Methoden für verschiedene Vertex-Selektionsmethoden in Abhängigkeit des transversalen Jetimpulses  $p_{T,Jet}$  bei einer B-Tagging-Effizienz von 60%. Es werden nur wahre Signaljets berücksichtigt. Das Verhältnis (b,d,f) wird bezogen auf die Unterdrückung bei Verwendung einer  $S_{p_T}^2$ -Vertex-Selektion.

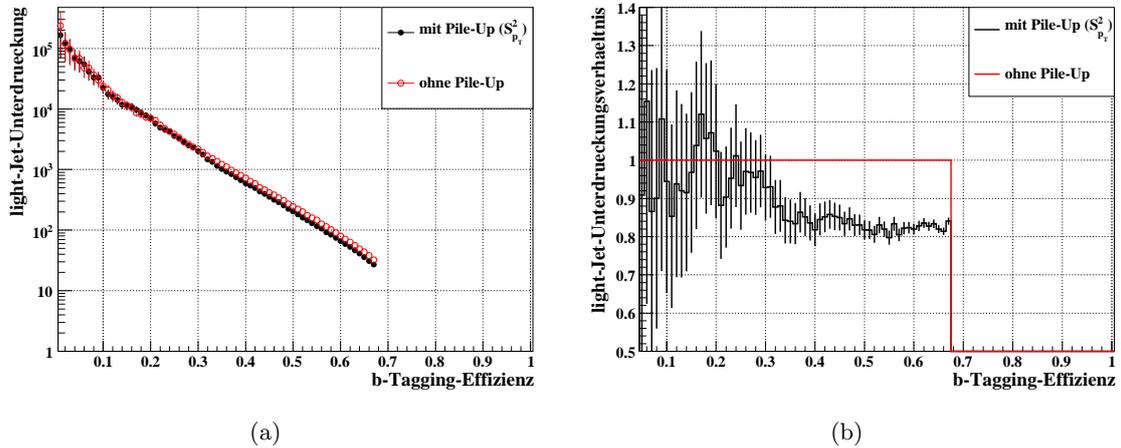


**Abb. A.7.:** Vergleich der Leistungsfähigkeit verschiedener B-Tagging-Methoden für verschiedene Vertex-Selektionsmethoden in Abhängigkeit der Pseudorapidity  $\eta$  bei einer B-Tagging-Effizienz von 60%. Es werden nur wahre Signaljets berücksichtigt. Das Verhältnis (b,d,f) wird bezogen auf die Unterdrückung bei Verwendung einer  $S_{p_T}^2$ -Vertex-Selektion.

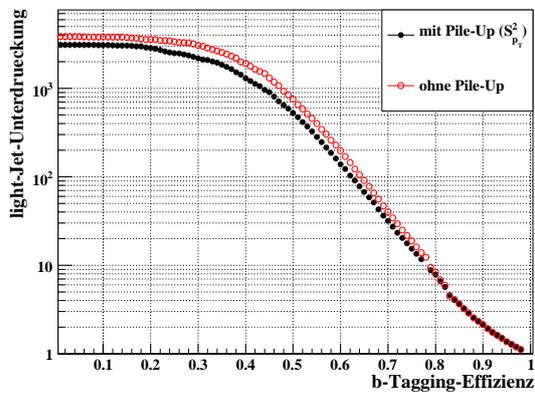
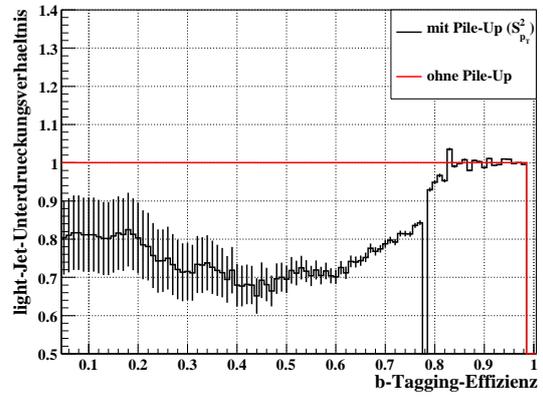
### A.3. Einfluss der beiden betrachteten *Pile-Up*-Effekte



**Abb. A.8.:** Vergleich der B-Tagging-Leistungsfähigkeit für *IP3D* auf einem Datensatz mit und ohne *Pile-Up*. Nur Jets aus der Signalwechselwirkung werden berücksichtigt. Im Falle des *Pile-Up*-Datensatzes wurde der Signalvertex an Hand der  $S_{p_T}^2$ -Methode selektiert. Die Unterdrückung in den Abbildungen (c,d,e,f) ist für eine b-Tagging-Effizienz von 60 % angegeben.



**Abb. A.9.:** Vergleich der B-Tagging-Leistungsfähigkeit für *SV1* auf einem Datensatz mit und ohne *Pile-Up*. Nur Jets aus der Signalwechselwirkung werden berücksichtigt. Im Falle des *Pile-Up*-Datensatzes wurde der Signalvertex an Hand der  $S_{p_T}^2$ -Methode selektiert. Die Unterdrückung in den Abbildungen (c,d,e,f) ist für eine b-Tagging-Effizienz von 60 % angegeben.

(a) *COMB*(b) *COMB*

**Abb. A.10.:** Vergleich der B-Tagging-Leistungsfähigkeit für *COMB* auf einem Datensatz mit und ohne *Pile-Up*. Nur Jets aus der Signalwechselwirkung werden berücksichtigt. Im Falle des *Pile-Up*-Datensatzes wurde der Signalvertex an Hand der  $S^2_{p_T}$ -Methode selektiert. Die Unterdrückung in den Abbildungen (c,d,e,f) ist für eine b-Tagging-Effizienz von 60 % angegeben.



# Literaturverzeichnis

- [1] Dorothea Bronner. *Untersuchung zur Lokalisierung der Hydroxycholanoyl-SCoA: Akzeptor N-Hydroxycholanoyl-Transferase*. Diplomarbeit, Fakultät für Chemie, Pharmazie und Geowissenschaften, Albert-Ludwigs-Universität Freiburg, Januar 1989.
- [2] Francis Halzen and Alan Douglas Martin. *Quarks and leptons: an introductory course in modern particle physics*.
- [3] David J Griffiths. *Introduction to elementary particles; 2nd ed.* Physics textbook.
- [4] C. Amsler et al. Review of particle physics. *Phys. Lett.*, B667:1, 2008.
- [5] Boris Kayser. On the Quantum Mechanics of Neutrino Oscillation. *Phys. Rev.*, D24:110, 1981.
- [6] Gerhard Buchalla, Andrzej J. Buras, and Markus E. Lautenbacher. Weak decays beyond leading logarithms. *Rev. Mod. Phys.*, 68:1125–1144, 1996.
- [7] Laura Reina. TASI 2004 lecture notes on Higgs boson physics. 2005.
- [8] The CDF Collaboration, the D0 Collaboration, the Tevatron New Physics, and Higgs Working Group. Combined CDF and D0 Upper Limits on Standard Model Higgs- Boson Production with 2.1 - 5.4 fb<sup>-1</sup> of Data. 2009.
- [9] G. Aad et al. Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics. 2009.
- [10] <https://cdsweb.cern.ch/record/979035>. CERN Document Server.
- [11] G. Aad et al. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008.
- [12] Torbjorn Sjostrand, Stephen Mrenna, and Peter Skands. PYTHIA 6.4 Physics and Manual. *JHEP*, 05:026, 2006.
- [13] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching NLO QCD computations with Parton Shower simulations: the POWHEG method. *JHEP*, 11:070, 2007.

- 
- [14] F. Carminanti R. Brun. GEANT-Detector Description and Simulation Tool. 1993. CERN Programm Library Long Writeup W5013 edition.
- [15] R. Engel. <http://www-ik.fzk.de/~engel/phoman5c.ps>. PHOJET Manual (programm version 1.05c, June 96).
- [16] F. Gianotti. Physics at the LHC. *Phys. Rept.*, 403:379–399, 2004.
- [17] The ATLAS Collaboration. ATLAS: Detector and physics performance technical design report. Volume 1. CERN-LHCC-99-14.
- [18] (Ed. ) Bruning, Oliver S. et al. LHC design report. Vol. I: The LHC main ring. CERN-2004-003-V-1.
- [19] Pierre Billoir and S. Qian. Fast vertex fitting with a local parametrization of tracks. *Nucl. Instrum. Meth.*, A311:139–150, 1992.
- [20] R. Fruhwirth. Application of Kalman filtering to track and vertex fitting. *Nucl. Instrum. Meth.*, A262:444–450, 1987.
- [21] N. Giacinto Piaquadio. *Vertex reconstruction algorithms for the identification of b-jets with the ATLAS Detector and discovery potential of a Standard Model Higgs Boson decaying into a couple of b-quarks*. PhD thesis, Albert-Ludwigs-Universität Freiburg, Germany, 2009. In Bearbeitung.
- [22] D. Buskulic et al. A Precise measurement of Gamma ( $Z \rightarrow b\bar{b}$ ) / Gamma ( $Z \rightarrow$  hadrons). *Phys. Lett.*, B313:535–548, 1993.
- [23] Carsten Peterson, Thorsteinn Rognvaldsson, and Leif Lonnblad. JETNET 3.0: A Versatile artificial neural network package. *Comput. Phys. Commun.*, 81:185–220, 1994.
- [24] T G Cornelissen, N Van Eldik, M Elsing, W Liebig, E Moyse, G Piacquadio, K Prokofiev, A Salzburger, and A Wildauer. ATL-SOFT-PUB-2007-003.
- [25] Th Goepfert. Counting signed impact parameter based b-tagging algorithm with atlas. Technical Report ATL-PHYS-INT-2009-064. ATL-COM-PHYS-2008-254, CERN, Geneva, Dec 2008.

## Danksagung

Meine Arbeit in der Abteilung Jakobs im vergangenen Jahr hat meine wissenschaftliche Entwicklung in entscheidendem Maße beeinflusst. Ich hatte hier die Möglichkeit unter sehr guter Betreuung selbst Ideen zu entwickeln und zu verwirklichen. Dabei wurde mir in sehr gesundem Maße eine große Eigenverantwortung anvertraut, ohne dass ich mich je dabei alleine gelassen gefühlt habe. Schon nach sehr kurzer Zeit wurde ich als volles wissenschaftliches und soziales Mitglied in dieser Abteilung aufgenommen. Nie habe ich zu spüren bekommen, dass ich *nur* Diplomandin war und genoss trotz allem einen gewissen Welpenschutz. Die Vollwertigkeit vor allem als wissenschaftliches Mitglied kann zunächst abschrecken und auch überfordern. Sie beinhaltet aber die große Chance sich zu messen und sich voll zu entfalten. Nur so war ich in der Lage eine Entscheidung zu fällen, ob eine weitere wissenschaftliche Laufbahn infrage kommt. Und ja, das kommt sie! Was ich im letzten Jahr in dieser Abteilung erlebt habe, bestätigt mich in höchstem Maße darin den akademischen Weg zumindest für den nächsten Schritt weiter zu gehen. Hierfür sind alle Mitglieder der Abteilung Jakobs verantwortlich und ich bedanke mich herzlich bei allen von euch für die für mich in jeglicher Hinsicht erfolgreichen vergangenen eineinhalb Jahre.

Die Aufnahme in diese Abteilung und die Vollwertigkeit ist vor allem in der zweiten Hälfte meiner Zeit hier nicht zuletzt auf das Oberhaupt der Abteilung zurückzuführen. Zeigt man Spaß und Interesse an der Arbeit, dann wird man nach vorne gestellt, dann darf bzw. muss man sich und seine Arbeit präsentieren. Dies ermöglichte es mir mich selbst in einer wissenschaftlichen Umgebung kennen zulernen und meine Scheu abzulegen. Dieser Dank geht an Herrn Prof. Karl Jakobs. Ihm habe ich neben der Bereitstellung für die Arbeit zu verdanken, dass ich einen Einblick in die wirkliche Welt der Wissenschaft erhalten haben und schon als Diplomandin daran teilnehmen durfte.

Und nun gibt es die Menschen, die da waren, Tag für Tag. Ohne diese wäre ich nicht jeden Morgen so fröhlich ins Büro gekommen. Denn da saßen immer schon zwei, die einen jeden Tag aufs neu nach dem werten Befinden befragt haben. (Einmal hab ich's geschafft vor 8 Uhr und vor den beiden da zu sein, ich bin fast geplatzt vor Stolz... und dann hatten die beiden Urlaub und keiner hat's gesehen. Gott sei Dank waren nicht alle in meinem Büro so pünktlich!) Susanne, Evelyn und Matti ich hab euch im vergangenen Jahr fast Tag täglich gegenüber gesessen und ich hab euch kein bisschen über. Vielen Dank, dass

ihr da wart!

Es gibt Menschen die hinterlassen einem etwas ohne es zu ahnen. Jochen Hartert, ich bin dir zu tiefstem Dank für einen Satz verpflichtet: 'Is mir egal, kannst auch nen anderen Editor nehmen, aber dann kann ich dir nicht helfen!'. Auch wenn ich geflucht habe wie ein Rohrspatz, heute bin ich dir sehr dankbar dafür. Du hast als mein inoffizieller Betreuer während meiner Sommerstudentenzeit am CERN, die Grundlage meiner *Programmierungskunst* gelegt. Ich bin sehr froh um alles, was du mich gelehrt hast damals und auch später hier in Freiburg. Denn: einmal Betreuer...immer Betreuer!

Zwei Menschen in dieser Abteilung gilt mein aller größter Dank. Diesen beiden ist alles zu verdanken. Alles, was in dieser Arbeit erreicht worden ist, wäre ohne diese Beiden unmöglich gewesen. Sie haben mir vertraut, sie haben mir Fehler verziehen, sie waren immer da, wenn ich Schwierigkeiten hatte, sie waren geduldig, sie hatten die Ideen, die diese Arbeit zu dem gemacht hat, was sie ist. Der Leser mag hin und wieder über das etwas unwissenschaftliche Wort *pfiffig* gestolpert sein. Dies ist ein dezenter Hinweis, dass dabei Giacinto Piacquadio seine Finger mit im Spiel hatte. Pfiffig war einfach das erste Wort, was mir zu Giacinto und seinen genialen Ideen eingefallen ist. Giacinto, du hast mich schwer beeindruckt. Es war ein unglaubliches Glück, dass du dich um mich gekümmert hast. Dr. Christian Weiser ist der zweite, dem ich herzlich für alles danken möchte, was ich im letzten Jahr lernen durfte. Neben der Tatsache, dass er ein besserer Betreuer nicht hätte sein können hat er mir etwas Besonderes vorgelebt: ehrliche Begeisterung für die Wissenschaft. In den vielen inhaltlichen Gesprächen mit ihm ging es immer um die Sache. Weil es um die Sache ging und nicht so sehr um mich habe ich die Angst ablegen können nicht gut genug zu sein. Aus Gründen, die ich selbst nicht recht erklären kann, habe ich immer mit großer wissenschaftlicher Zuversicht in mich selbst sein Büro verlassen, auch wenn ich Dinge nicht wusste. Christian, du hast viel mehr aus mir herausgeholt, als ich je erwartet hätte, dass es von mir möglich ist. Und besten Dank, dass du das ganze Ding gelesen und korrigiert hast. Giacinto und Christian ihr seit mir ein großes Vorbild geworden und ich habe sehr sehr viel von euch gelernt. Ich danke euch für jeden Satz, den ich durch eure Hilfe in diese Arbeit schreiben konnte und das ist jeder Satz.

Ich möchte allen denen danken, die sich meiner katastrophalen Rechtschreibung und Kommasetzung und meinen etwas absonderlichen deutschen Ausdrücken angenommen haben. Dieser Dank gilt vor allem Inga Ludwig, Susanne Kühn und meiner Mitbewohnerin Carola Fricke aber auch vielen anderen fleißigen Lesern. Ohne euch, wäre das gute Stück wohl unlesbar! Ich weiß schon warum ich Physik studiert habe!

Es gibt viele Menschen in meiner privaten Umgebung, die mich das letzte Jahr begleitet haben, aber um die geht es hier nicht. Diese Arbeit ist ein Produkt dieser Abteilung und der Dank dafür geht an euch alle hier, die ihr mich dabei begleitet habt.

Hiermit versichere ich, dass ich diese Arbeit selbständig verfasst und keine anderen als die angegebenen Hilfsmittel verwendet habe.

Freiburg, den \_\_\_\_\_

\_\_\_\_\_

Johanna Bronner