

Studie zur Unterscheidung von Quark- und Gluon-Jets in Ereignissen der Streuung schwacher Eichbosonen W und Z mit dem ATLAS Detektor

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

vorgelegt von

Anna Sophie Tegetmeier
geboren am 09.07.1997 in Hamburg

Institut für Kern- und Teilchenphysik
Fakultät Physik
Bereich Mathematik und Naturwissenschaften
Technische Universität Dresden
2019

Eingereicht am 14. Oktober 2019

1. Gutachter: Prof. Dr. Michael Kobel
2. Gutachter: Dr. Frank Siegert

Zusammenfassung

Zusammenfassung

Ein Prozess der Vektorbosonstreuung (VBS) ist die Streuung von zwei volleptonisch zerfallenden gleich geladenen W-Bosonen. Der größte Untergrund zu diesem Prozess sind Prozesse mit leptonisch zerfallenden Z- und W-Bosonen. Eine Möglichkeit zur Unterscheidung dieser Prozesse bildet der höhere Anteil an Gluon-initiierten Jets im letzteren. In dieser Studie wird ein Quark-Gluon-Tagging Tool, das Boosted Decision Trees (BDTs) als multivariaten Ansatz nutzt, untersucht. Dazu wird dieses auf Monte-Carlo-Simulationen der Proton-Proton-Streuung bei einer Schwerpunktsenergie von $\sqrt{s} = 13$ TeV angewendet. Als erster Schritt wird das Potential eines perfekten Quark-Gluon-Taggers mithilfe der Wahrheitswerte der Jet-initiiierenden Teilchen auf „truth level“ untersucht. Im Fall, dass die Gluon-Jets in allen η -Bereichen perfekt identifiziert werden können, ergibt sich eine maximale Steigerung der Signifikanz um 1,3 Standardabweichungen (σ). Für den Fall, dass die Separation nur im Zentralbereich möglich ist, kann die Signifikanz um maximal 0,89 σ erhöht werden. Bei der Anwendung des BDT Tools kann keine Verbesserung der Signifikanz gefunden werden. Auch die Anwendung in einem erweiterten Phasenraum führt zu keinen Verbesserungen.

Abstract

The scattering of two leptonically decaying same-sign W bosons is a process of the vector boson scattering (VBS). The largest background of this process are processes with leptonically decaying Z and W bosons. In order to separate the signal from this background the larger fraction of gluon initiated jets in the latter can be used. In this study a quark-gluon tagging tool which uses boosted decision trees as a multivariate approach is investigated by applying it on Monte-Carlo simulations of these processes with a centre-of-mass energy of $\sqrt{s} = 13$ TeV. As a first step the potential of a perfect quark-gluon tagger is estimated on truth level to gain information about the limits of this approach. If gluon jets can be identified in all η regions the significance can be increased by a maximum of 1,3 standard deviations (σ). In the case that gluon jets can only be separated in the forward region the maximum increase of the significance is 0.89 σ . With the BDT tool no improvement of the significance can be found. The investigation of a looser phase space does not result in any improvements either.

Inhaltsverzeichnis

1	Einleitung	5
2	Theoretischer Hintergrund	6
2.1	Das Standardmodell der Teilchenphysik	6
2.2	Streuung von Eichbosonen	7
2.3	Die Signalprozesse - ssWWjj EW	7
2.4	Die Untergrundprozesse - WZjj QCD	8
3	Das Experiment	9
3.1	Der Large Hadron Collider	9
3.2	Der ATLAS-Detektor	9
3.2.1	Das Koordinatensystem	9
3.2.2	Die Subdetektoren	10
4	Unterscheidung von Quark- und Gluon-Jets	10
4.1	Das N_{track} Tool	11
4.2	Das BDT-Tool	11
4.2.1	Boosted Decision Trees	11
4.2.2	Einführung in das BDT Tool für Quark-Gluon-Tagging	12
5	Statistische Methoden	12
5.1	Berechnung der Signifikanz	12
5.2	Berechnung der Unsicherheit von Effizienten	13
6	Ereignisauswahl	14
7	Studie im VBS Phasenraum	15
7.1	Unterscheidung auf Truth Level - Die PartonTruthLabelID	15
7.2	Das Potential eines perfekten Quark-Gluon-Taggers	19
7.2.1	Das Potential eines perfekten Quark-Gluon-Taggers in allen η -Bereichen	19
7.2.2	Das Potential eines perfekten Quark-Gluon-Taggers im Zentralbereich .	20
7.3	Das BDT-Tool	22
7.3.1	Differentielle Verteilung der BDT-Variablen	22
7.3.2	Differentielle Verteilung des BDT-Wertes	26
7.4	Abhängigkeit vom Transversalimpuls der Jets	28
7.5	Abhängigkeit von der durchschnittlichen Anzahl an Interaktionen pro Teilchen- bündel	32

8 Studie im Phasenraum ohne die VBS Schnitte	34
8.1 Die PartonTruthLabelID	36
8.2 Differentielle Verteilung der BDT Variablen	36
8.3 Differentielle Verteilung des BDT-Wertes	39
8.4 Anwendung von einzelnen VBS Schnitten	42
9 Zusammenfassung und Ausblick	46
A Verwendete Monte-Carlo Dateien	48
B Vergleich von Quark-Jets im Signal und Untergrund	48
C Vergleich von Quark-Jets mit großen zu Quark-Jets mit kleinen p_T-Werten	51
D Signifikanz der BDT-Werte für kleine und große p_T-Werte	52
E BDT Variablen für unterschiedliche avIntPerX Regionen	52
F BDT-Variablen bzw. BDT-Wert der ersten Jets ohne VBS Schnitte	57
G BDT-Wert nach einzelnen VBS Schnitten	59
Literatur	60

1 Einleitung

Die erfolgreichste Theorie, die die Elementarteilchen und ihre Wechselwirkungen beschreibt, ist das Standardmodell der Teilchenphysik. Dessen Vorhersagen wurden bereits in vielen Experimenten bestätigt. Zur immer genaueren Überprüfung werden Teilchenbeschleuniger gebaut, die die Teilchen auf immer höhere Energien beschleunigen können. Der momentan leistungsstärkste Teilchenbeschleuniger ist der Large Hadron Collider am CERN, der bisher Proton-Proton-Kollisionen von bis zu 13 TeV realisieren konnte.

Ein Prozess, der vom Standardmodell vorhergesagt wird ist die Streuung von zwei massiven Eichbosonen. Mit diesem als Vektorbosonstreuung (VBS) bezeichneten Prozess kann die elektroschwache Symmetriebrechung untersucht werden [1]. Die VBS beinhaltet die elektroschwache (EW) Streuung zwischen zwei leptonisch zerfallenden gleich geladenen W-Bosonen. Diese Erste Hinweise auf diesen Prozess wurden bei einer Luminosität von $20,3 \text{ fb}^{-1}$ bei Proton-Proton-Kollisionen mit einer Schwerpunktsenergie von $\sqrt{s} = 8 \text{ TeV}$ mit einer Signifikanz von $3,6 \sigma$ von der ATLAS-Kollaboration [2] und mit einer Signifikanz von $2,0 \sigma$ von der CMS-Kollaboration [3] gefunden. Die Beobachtung dieses Prozesses konnte mit einer Signifikanz von $5,5 \sigma$ [4] mit dem CMS-Detektor und mit einer Signifikanz von $6,5 \sigma$ mit dem ATLAS-Detektor [1] bei pp -Kollisionen mit $\sqrt{s} = 13 \text{ TeV}$ und Luminosität $36,1 \text{ fb}^{-1}$ erreicht werden. Der größte Untergrund zu diesem Prozess sind quantenchromodynamische (QCD) Prozesse der Produktion von einem W- und einem Z-Boson. Diese unterscheiden sich von Signal-Prozessen durch einen größeren Anteil an Gluon-initiierten Jets im Endzustand. Somit bietet die Unterscheidung zwischen Quark-initiierten (Quark-Jets) und Gluon-initiierten Jets (Gluon-Jets) eine Möglichkeit, Signal-Prozesse von Untergrundprozessen zu unterscheiden. Für das Quark-Gluon-Tagging wurden unterschiedliche Tools entwickelt, die die unterschiedlichen topologischen Eigenschaften von Quark-Jets und Gluon-Jets ausnutzen. Eines davon nutzt den multivarianten Ansatz mit Boosted Decision Trees, mit dem die Informationen aus mehreren Tagging-Variablen kombiniert werden können. Dieses Tool wird in der folgenden Studie auf sein Potential und die tatsächliche Leistung beim Quark-Gluon-Tagging bei Anwendung auf Monte-Carlo Simulationen bezüglich der pp -Kollisionen bei $\sqrt{s} = 13 \text{ TeV}$ bei einer Luminosität von $36,1 \text{ fb}^{-1}$ untersucht.

Die Arbeit ist wie folgt aufgebaut: Zuerst werden die grundlegenden theoretischen Konzepte und die untersuchten Prozesse eingeführt (Abschnitt 2) und anschließend ein Überblick über den ATLAS-Detektor gegeben (Abschnitt 3). Das BDT Tool für das Quark-Gluon-Tagging wird in Abschnitt 4 beschrieben. In Abschnitt 5 werden statistische Methoden, die in dieser Arbeit verwendet werden, eingeführt. Die Auswahl an Ereignissen wird in Abschnitt 6 beschrieben. In Abschnitt 7 wird das Potential eines perfekten Quark-Gluon-Taggers berechnet und die Anwendung des BDT Tools auf die Monte-Carlo Simulation untersucht. Die Betrachtung des BDT Tools in einem erweiterten Phasenraum erfolgt in Abschnitt 8.

2 Theoretischer Hintergrund

2.1 Das Standardmodell der Teilchenphysik

Das Standardmodell der Teilchenphysik ist die beste Beschreibung von Elementarteilchen und deren Wechselwirkungen, die es momentan gibt. Es ist eine relativistische Quantenfeldtheorie, die hauptsächlich zwischen 1961 und 1973 entwickelt wurde.

Grundlegend für das Standardmodell ist die Forderung nach lokalen Eichsymmetrien. In der Lagrangedichte, die die Prozesse der Teilchenphysik beschreibt, tauchen durch diese Forderung Terme auf, die den Austauschteilchen der jeweiligen Wechselwirkung entsprechen. Die Ladungsoperatoren der jeweiligen Wechselwirkungen sind die Generatoren der Symmetrie. Die Anzahl der Austauschteilchen entspricht dabei der Anzahl der Generatoren. Als Ladungen können die Eigenwerte der Generatoren gemessen werden. Nur Teilchen, die die entsprechende Ladung tragen, können bei einer bestimmten Wechselwirkung mitwirken. Das Standardmodell beschreibt mit dem Elektromagnetismus, der schwachen Wechselwirkung und der starken Wechselwirkung drei der vier fundamentalen Wechselwirkungen. Der Elektromagnetismus folgt aus der $U(1)$ Symmetrie. Ihm werden die elektrische Ladung Z und das Photon γ als Austauschteilchen zugeordnet. Aus der $SU(2)$ Symmetrie folgt die schwache Wechselwirkung, die die drei Austauschteilchen W^+ , W^- und Z besitzt. Ihre Ladung ist der schwache Isospin $\vec{I}_W$. Für die starke Wechselwirkung folgen aus der $SU(3)$ Symmetrie acht Austauschteilchen, die Gluonen genannt werden. Zu ihr gehört die Farbladung $\vec{C}$. Die Gravitation ist die einzige grundlegende Wechselwirkung, die im Standardmodell nicht berücksichtigt wird.

Die bisher bekannten Elementarteilchen lassen sich in zwei Kategorien Bosonen und Fermionen einteilen. Bosonen sind Teilchen mit ganzzahligen Spin. Zu dieser Kategorie gehören die Austauschteilchen, die daher auch Eichbosonen genannt werden und alle den Spin 1 besitzen, und das Higgs-Boson, das den Spin Null besitzt. Die Wechselwirkungen mit den zugehörigen Eichbosonen sind in der folgenden Tabelle dargestellt:

Wechselwirkung	Eichboson	Ladung
Elektromagnetismus	Photon γ	Z
schwache Wechselwirkung	W^-, W^+, Z	$\vec{I}^W$
starke Wechselwirkung	Gluonen g	$\vec{C}$

Die Fermionen (Teilchen mit halbzahligen Spin) lassen sich weiter in Leptonen und Quarks unterteilen. Bei den Leptonen kann man die elektrisch geladenen (e , μ und τ) von den elektrisch neutralen (ν_e , ν_μ und ν_τ) unterscheiden. Die Quarks tragen als einzige Materieteilchen eine Farbladung. Auch hier kann eine Unterscheidung aufgrund der elektrischen Ladung vorgenommen werden. Es gibt die Quarks mit $Z = 2/3$ (u, c, t) und mit $Z = -1/3$ (d, s, b). Alle diese Fermionen besitzen den Spin $1/2$. Die Teilchen innerhalb der gerade beschriebenen

Kategorien unterscheiden sich in ihrer Masse. Damit lassen sich die Teilchen in drei Generationen unterscheiden, wobei eine höhere Generation eine größere Masse bedeutet. Nur für die Neutrinos ist dies nicht sicher. Die Teilchen mit den entsprechenden Ladungen sind in der folgenden Tabelle dargestellt:

	Generation			Ladungen		
	1	2	3	Z	I_3^W	$\vec{C}$
Quarks	u	c	t	$+\frac{2}{3}$	$\frac{1}{2}$	r, g, b
	d	s	b	$-\frac{1}{3}$	$-\frac{1}{2}$	r, g, b
Leptonen	e	μ	τ	-1	$-\frac{1}{2}$	$\vec{0}$
	ν_e	ν_μ	ν_τ	0	$\frac{1}{2}$	$\vec{0}$

Laut dem Standardmodell gibt es außerdem zu jedem Teilchen noch ein Antiteilchen mit entgegengesetzter Ladung.

2.2 Streuung von Eichbosonen

Aus der Lagrangedichte für die jeweiligen Wechselwirkungen kann abgelesen werden, ob es Wechselwirkungen zwischen den Austauscheteilchen einer bestimmten Wechselwirkung gibt. Die U(1) Symmetrie des Elektromagnetismus ist abelsch, was dazu führt, dass keine Terme für die Selbstwechselwirkung von Photonen in der Lagrangedichte auftauchen. Im Gegensatz dazu führen die nicht-abelschen Symmetrien SU(2) und SU(3) zum Auftreten solcher Terme. Ein interessanter Prozess ist dabei die Vektorbosonstreuung (VBS) der Austauscheteilchen der schwachen Wechselwirkung. Dieser Prozess beinhaltet die Streuung von zwei massiven Eichbosonen $VV \rightarrow VV$, wobei $V = W^\pm, Z$.

2.3 Die Signalprozesse - ssWWjj EW

Der in dieser Studie untersuchte Signalprozess ist der elektroschwache (EW) Prozess der Vektorbosonstreuung (VBS) zwischen zwei gleich geladenen W-Bosonen. In der Proton-Proton-Streuung am LHC werden diese von zwei Quarks abgestrahlt. Die aus der VBS hervorgehenden W-Bosonen zerfallen in ein Lepton-Anti-Lepton-Paar. In dem hier betrachteten Prozessen kann das W^- in ein Elektron und ein Anti-Elektron-Neutrino oder das entsprechende Paar mit Myonen zerfallen. Die W^+ -Bosonen zerfallen wiederum in ein Positron und ein Elektron-Neutrino oder die entsprechenden Gegenstücke mit Myonen. Die Quarks können als Träger der starken Ladung nicht isoliert vorkommen und daher nicht direkt gemessen werden. Stattdessen treten

sie im Detektor als Teilchenschauer auf. Diese Quark-initiierten Jets werden im folgenden als Quark-Jets bezeichnet.

Neben dieser Vektorbosonstreuung können weitere Prozesse mit den gleichen Endzuständen auftreten. Diese können nicht von dem eigentlichen Prozess unterschieden werden und werden somit im Signal mit berücksichtigt. Prozesse mit diesen Endzuständen werden als ssWWjj EW-Prozesse bezeichnet. Einige Feynmandiagramme der Signalprozesse sind in Abbildung 2.1 dargestellt.

In ssWWjj EW-Prozessen kann durch Prozesse höherer Ordnung auch ein kleiner Anteil an Gluon-Jets vorkommen.

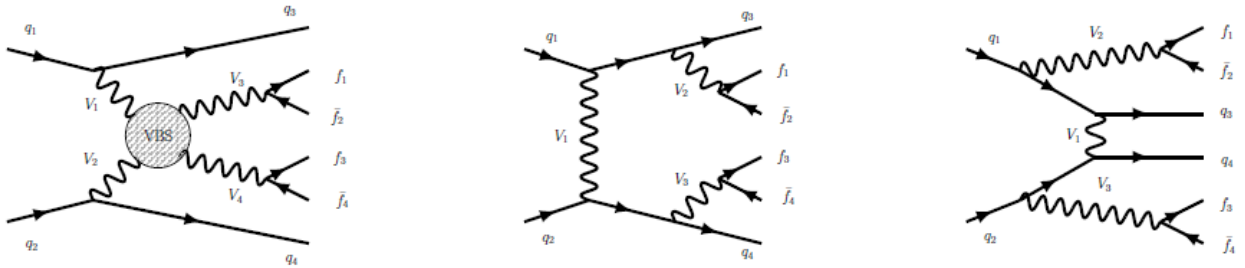


Abbildung 2.1: Feynmandiagramme der ssWWjj-Signalprozesse. Der Kreis im linken Diagramm steht dabei für alle möglichen Prozesse der Vektorbosonstreuung.

2.4 Die Untergrundprozesse - WZjj QCD

Der größte Untergrund zum ssWWjj EW-Prozess sind quantenchromodynamische (QCD) Prozesse mit einem $W^\pm$ - und einem Z-Boson (WZjj QCD). Diese werden im folgenden einfach als Untergrund bezeichnet. Feynmandiagramme, die diesen Prozess abbilden, sind in Abbildung 2.2 dargestellt. Wie bei ssWWjj EW-Prozessen werden auch hier die Prozesse mit vier Leptonen im Endzustand betrachtet. Das Z-Boson zerfällt dabei in ein Elektron-Positron oder Myon-Anti-Myon-Paar. Neben Quark-Jets treten in diesen Prozessen auch Gluon-Jets in führender Ordnung auf.

Der somit größere Anteil an Gluon-Jets im Untergrund bietet eine Möglichkeit zur Unterscheidung von Signal und Untergrund.

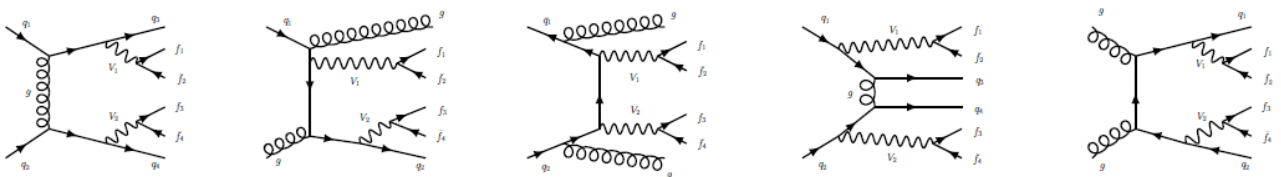


Abbildung 2.2: Beispiele von Feynmandiagrammen für den WZjj QCD Untergrund.

3 Das Experiment

3.1 Der Large Hadron Collider

Der Large Hadron Collider (LHC) ist der weltweit größte Teilchenbeschleuniger. Er liegt am Kernforschungszentrum CERN an der Grenze von Schweiz zu Frankreich zwischen 27 und 170 m unter der Erde. Die Teilchen passieren zunächst mehrere vorgeschaltete Beschleuniger, bevor sie in den ca. 27 km langen Hauptring, in dem die Teilchenkollisionen stattfinden, gelangen. Zum LHC gehören die vier Detektoren ATLAS, CMS, ALICE und LHCb. Im LHC können sowohl Protonen als auch schwere Ionen zur Kollision gebracht werden. Bei den Proton-Proton-Kollisionen wurden bisher Schwerpunktsenergien von bis zu $\sqrt{s} = 13$ TeV erreicht.

3.2 Der ATLAS-Detektor

Der ATLAS Detektor ist mit einer Länge von 44 m, einer Höhe von 23 m und einem Gewicht von ca. 7000 Tonnen der momentan größte Teilchendetektor weltweit. Er ist neben CMS einer der Universaldetektoren und ist zylinderförmig um die Strahlenachse angeordnet. Um die vielfältigen Teilchen, die bei den Teilchenkollisionen entstehen, zu messen, besteht er aus mehreren Subdetektoren. Die folgenden Angaben lassen sich in Referenz [5] finden.

3.2.1 Das Koordinatensystem

Am ATLAS wird ein rechtshändiges kartesisches Koordinatensystem verwendet. Der Ursprung ist dabei der Kollisionspunkt der Teilchen. Die x-Achse zeigt vom Kollisionspunkt aus zum Mittelpunkt des Beschleunigerrings. Die y-Achse zeigt senkrecht dazu nach oben und die z-Achse ist entlang der Strahlrichtung angeordnet. Anstelle der kartesischen Koordinaten werden oft auch der Azimuthwinkel Φ und der Polarwinkel Θ verwendet. Φ gibt dabei den Winkel von der x-Achse zu der Projektion auf die yz-Ebene an, während Θ den Winkel zwischen der z-Achse und dieser Projektion misst. Anstelle von Θ wird oft die Pseudorapidität η genutzt:

$$\eta = -\ln \left(\tan \frac{\Theta}{2} \right) \quad (3.1)$$

Diese ist der Grenzfall der Rapidität y für $m \ll p$:

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

Dabei ist E die Teilchenenergie und p_z die z-Komponente des Teilchenimpulses. Für Teilchen in Strahlrichtung gilt $|\eta| \rightarrow \infty$, während Teilchen senkrecht dazu $\eta = 0$ haben. Mit der Pseudorapidität lassen sich die zwei Bereiche Vorwärtsbereich ($|\eta| > 2,5$) und Zentralbereich ($|\eta| \leq 2,5$) aufteilen. Eine häufig verwendete Größe ist der Winkelabstand ΔR zwischen zwei

Objekten:

$$\Delta R = \sqrt{(\Delta\Phi)^2 + (\Delta\eta)^2} \quad (3.3)$$

In der xy-Ebene kann der Transversalimpuls definiert werden:

$$p_T^2 := p_x^2 + p_y^2 \quad (3.4)$$

3.2.2 Die Subdetektoren

Der ATLAS-Detektor kann in die vier Subdetektoren innerer Detektor, elektromagnetisches Kalorimeter, hadronisches Kalorimeter und Myon-Spektrometer aufgeteilt werden.

Direkt um die Strahlröhre liegt der innere Detektor, der der Rekonstruktion von Teilchenspuren dient. Er ist von einem 2 T starken Magnetfeld umgeben, das von Solenoidmagneten erzeugt wird. In diesem werden die Teilchen abhängig von ihrer Ladung und ihrem Impuls abgelenkt. Er besteht aus einem Pixeldetektor, einem Halbleiterdetektor (SCT) und einem Transition Radiation Tracker (TRT). Der Pixeldetektor und der SCT dienen dabei der präzisen Messung der Teilchenspuren und decken eine Region von $|\eta| < 2,5$ ab. Mit dem TRT können elektrisch geladene Teilchen bis zu $|\eta| = 2,0$ identifiziert werden.

Der innere Detektor ist von dem elektromagnetischen Kalorimeter umgeben. Mit diesem kann die Energie von elektromagnetischen Teilchen gemessen werden. Das Kalorimeter kann in die Barrel-Region ($|\eta| < 1,475$) und die Endkappenregion ($1,375 < |\eta| < 3,2$) aufgeteilt werden. Er besteht aus abwechselnden Schichten von Blei als Absorbermaterial und flüssigem Argon als Nachweismedium.

Mit dem hadronischen Kalorimeter, das das elektronische Kalorimeter umgibt, soll die Energie von Hadronen gemessen werden. Im Bereich bis $|\eta| = 1,7$ besteht er aus dem Absorbermaterial Stahl und Szintillatoren. Mit einem Detektor aus Kupfer-Wolfram und flüssigen Argon kann die Reichweite bis zu $|\eta| = 4,9$ erweitert werden.

Der äußerste Detektor ist das Myon-Spektrometer, das dem Nachweis von Myonen dient.

4 Unterscheidung von Quark- und Gluon-Jets

Zur Unterscheidung von Quark- und Gluon-Jets kann ausgenutzt werden, dass diese unterschiedliche topologische Eigenschaften besitzen. Aufgrund der unterschiedlichen Farbfaktoren von Quarks und Gluonen (Quarks: $Q_F = 4/3$, Gluonen: $Q_A = 3$) wird erwartet, dass Gluon-Jets ungefähr $C_A/C_F = 9/4$ mehr Teilchen als Quark-Jets besitzen und die Gluon-Jets durch die Teilchenschauer außerdem breiter sind [6].

4.1 Das N_{track} Tool

Ein bereits existierendes Quark-Gluon-Tagging Tool ist in Referenz [7] für $jj \rightarrow pp$ bei $\sqrt{s} = 13$ TeV dargestellt. Dieses beruht allein auf der Anzahl von Teilchenspuren N_{tracks} . Zur Unterscheidung von Quark- und Gluon-Jets wird ein Grenzwert N_{tracks}^{max} festgelegt und alle Jets mit $N_{tracks} < N_{tracks}^{max}$ als Quark-Jets angesehen. Der Grenzwert soll dabei so gewählt werden, dass möglichst viele Quark-Jets den Schnitt passieren, während möglichst viele Gluon-Jets ausgeschlossen werden. Die Betrachtung der Quark-Jets Effizient ($\text{Pr}(\text{tagged}|\text{Quark-Jet})$) zur Rate der beibehaltenen Gluon-Jets ($\text{Pr}(\text{tagged}|\text{Gluon-Jet})$) in drei p_T Bins zwischen $p_T = 50$ GeV und $p_T = 1500$ GeV zeigt, dass die Unterscheidbarkeit von Quarks und Gluon-Jets mit zunehmenden p_T steigt.

4.2 Das BDT-Tool

Um die Unterscheidung von Quark- und Gluon-Jets zu verbessern, erscheint es sinnvoll, die Informationen aus mehreren Variablen zu kombinieren. Eine Möglichkeit, dies zu erreichen sind Boosted Decision Trees (BDTs). Das mit dieser Methode entwickelte Tool zum Quark-Gluon-Tagging wird im Folgenden beschrieben.

4.2.1 Boosted Decision Trees

Boosted Decision Trees sind eine Form des maschinellen Lernens, in der Daten eine Reihe von Aufspaltungen in zwei Kategorien durchlaufen. So können Ereignisse durch die Kombination von mehreren Variablen dem Signal oder dem Untergrund zugeordnet werden. Bei dem Training eines BDT wird zunächst die beste Variable und der beste Ort zum Aufspalten der Daten in Signal und Untergrund gefunden und die Trainingsdaten damit in zwei Kategorien (Knoten) aufgeteilt. Auf diese Knoten wird dann wieder eine Aufspaltung vorgenommen und der Prozess so lange wiederholt, bis ein Kriterium für den Abbruch erfüllt wird. Ein Beispiel für die Entscheidung, welche Variable und welcher Ort zum Aufteilen ausgewählt wird, ist in [8] dargestellt. Den Schlussknoten wird dann die Reinheit P zugeordnet, welche sich aus der Summe der Gewichte des Signals dividiert durch die Summe der Gewichte des Signals und des Untergrundes berechnet. Knoten mit $P \geq 0.5$ werden dann als Signal und die übrigen als Untergrund angesehen.

Da bei einem einzelnen Baum kleine Schwankungen in den Daten große Veränderungen an dem Ergebnis hervorrufen können, wird sogenanntes Boosting angewendet. Dabei wird das Gewicht der den falschen Kategorien zugeordneten Ereignisse erhöht und mit diesem Prinzip weitere Bäume gebildet.

Nach dem Training können Daten durch die Bäume geschickt werden, wobei sie für jeden Baum entweder den Wert 1 oder -1 bekommen, je nach dem, in welche Kategorie sie gelangen. Aus der gewichteten Summe dieser Werte wird dann der BDT-Wert gebildet, welcher zwischen 1

und -1 liegt.

4.2.2 Einführung in das BDT Tool für Quark-Gluon-Tagging

Zur Unterscheidung von Quark- und Gluon-Jets wurden mehrere potentielle Variablen gefunden, die auf Spur- oder Kalorimeter-Daten basieren ([6], [9]). Die Spur-Variablen berücksichtigen dabei nur geladene Teilchen, haben dabei gegenüber denen vom Kalorimeter aber den Vorteil, dass auch Teilchen mit kleinen p_T , die vom Kalorimeter nicht nachgewiesen werden können, berücksichtigt werden. Des Weiteren werden die Teilchen im magnetischen Feld abgelenkt und können gegebenenfalls nicht mehr im Kalorimeter-Jet verwendet werden. Außerdem können die Teilchenspuren leicht ihrem jeweiligen Vertex zugeordnet werden [6].

Für das BDT Tool für das Quark-Gluon-Tagging werden die folgenden fünf Variablen im Bezug auf die Tagging-Jets untersucht:

- N_{tracks}
- $w_{tracks} = \frac{\sum_{i \in jet} p_{T,i} \Delta R_{i,Jet}}{\sum_{i \in jet} p_{T,i}}$
- $C_{1,tracks}^{\beta=0.2} = \frac{\sum_{i,j \in jet} p_{T,i} p_{T,j} (\Delta R_{i,j})^\beta}{\sum_{i,j \in jet} p_{T,i} p_{T,j}}$
- p_T
- η

Hierbei ist N_{tracks} die Anzahl der Teilchenspuren in dem Jet und w_{tracks} die Breite des Jets. Bei $C_{1,tracks}$ werden die Transversalimpulse der Teilchenspuren unter Berücksichtigung ihres Winkelabstandes miteinander in Beziehung gebracht. Die Summen bei w_{tracks} bzw. C_{tracks} laufen dabei jeweils über die Teilchenspuren, die mit dem Jet assoziiert sind. Neben den Spur-Variablen werden auch der Transversalimpuls und die Pseudorapidität der Tagging-Jets verwendet. Das BDT-Tool ist für $p_T \in [20, 1500] \text{ GeV}$ und $|\eta| \in [0, 2.5]$ trainiert. Damit werden die Grenzen des inneren Detektors berücksichtigt, der nur Informationen über den Zentralbereich liefern kann. Durch einen Schnitt können Bedingungen an Ereignisse gestellt werden. Der Schnitt auf die η - und p_T -Werte, auf die das BDT-Tool trainiert ist, wird im folgenden als BDT Training Schnitt bezeichnet.

5 Statistische Methoden

5.1 Berechnung der Signifikanz

Die Signifikanz gibt an, wie unwahrscheinlich es ist, die vom Signal erwarteten zusätzlichen Ereignisse zu messen. Damit ist die Signifikanz ein Maß dafür wie sicher es etwas wie Signalprozesse geben muss. Für eine große Anzahl an Ereignissen kann die Näherung einer gaußförmigen Verteilung der Ereignisse, die nur statistische Unsicherheiten besitzt, angenommen

werden. Bei S gemessenen Signalereignissen und B Untergrundereignissen ergibt sich damit für die Signifikanz Z :

$$Z = \frac{S}{\sqrt{B}} \quad (5.1)$$

Für die genaue Berechnung der Signifikanz muss beachtet werden, dass die Ereignisse tatsächlich Poissonförmig verteilt sind und weitere Unsicherheiten wie Monte-Carlo-Unsicherheiten und Detektorunsicherheiten existieren. Bei einer Unsicherheit dB auf die Ereignisse des Untergrundes kann die Signifikanz folgendermaßen berechnet werden ([10]):

$$Z = \sqrt{2 * (N * (\ln(N) - 1 - \ln(b_{hh})) + b_{hh}) + \frac{(B - b_{hh})^2}{dB}} \quad (5.2)$$

wobei:

$$b_{hh} = 0.5 * (B - (dB)^2) + \sqrt{0.25 * (B - (dB)^2)^2 + N * (dB)^2}. \quad (5.3)$$

5.2 Berechnung der Unsicherheit von Effizienten

Bei der Bestimmung des Anteils z.B. der Gluon-Jets an allen Jets im Signal, muss bei der Berechnung der Unsicherheiten beachtet werden, dass die beiden Größen nicht statistisch unabhängig sind. Dazu kann das Bayes-Theorem verwendet werden. Dies wird in Referenz [11] beschrieben.

Mit dem Bayes Theorem kann man die Wahrscheinlichkeit, dass sich bei k von n Ereignissen, die einen Schnitt passieren, die tatsächliche Effizienz zwischen ϵ und $\epsilon + d\epsilon$ befindet, berechnen durch:

$$P(\epsilon; k, n) = \frac{P(k; \epsilon, n)P(\epsilon; n)}{C} \quad (5.4)$$

Dabei ist $P(k; \epsilon, n)$ die Wahrscheinlichkeit von k Erfolgen bei einer Binomialverteilung, $P(\epsilon, n)$ ist die Wahrscheinlichkeit von ϵ bei einem gegebenen n bevor die Daten berücksichtigt werden und C ist eine Normalisationskonstante. Für $P(\epsilon; n)$ kann eine Gleichverteilung in $[0, 1]$ gewählt werden, womit sich $C = \frac{1}{n+1}$ ergibt. Damit gilt:

$$P(\epsilon; k, n) = \frac{(n+1)!}{k!(n-k)!} \epsilon^k (1-\epsilon)^{n-k} \quad (5.5)$$

Der wahrscheinlichste Wert für ϵ ($dP/d\epsilon = 0$) ist dann $\text{mode}(\epsilon) = \frac{k}{n}$ und für die Varianz gilt

$$\sigma_\epsilon^2 = \frac{(k+1)(n-k+1)}{(n+2)^2(n+3)} \quad (5.6)$$

Die geschätzte Unsicherheit auf $\frac{k}{n}$ ist somit die Wurzel aus diesem Wert.

6 Ereignisauswahl

Für die Studie werden Monte-Carlo-Simulationen genutzt, die von Sherpa erzeugt wurden und die ATLAS-Detektor-Simulation durchlaufen haben. Es werden Simulationen von Proton-Proton-Kollisionen bei einer Schwerpunktsenergie von $\sqrt{s} = 13$ TeV aus 2015 und 2016 bei einer Luminosität von $36,07456 \text{ fb}^{-1}$ benutzt.

Die in diesem Abschnitt beschriebenen Kriterien können Referenz [12] entnommen werden. Da bei der Kreuzung der Protonenpakete mehr als eine Proton-Proton-Kollision stattfindet, kann mehr als ein Vertex rekonstruiert werden. Als primärer Vertex wird dabei derjenige mit dem größten Transversalimpuls der assoziierten Teilchenspuren bezeichnet. Damit Ereignisse berücksichtigt werden, muss der primäre Vertex mindestens drei zugeordnete Teilchenspuren besitzen. Die betrachteten Jets sollen vom primären Vertex stammen. Jets, die von anderen Vertices kommen, werden als Pileup bezeichnet.

Die Jets werden aus den topologischen Kalorimeter-Clustern mithilfe des anti- k_T -Algorithmus [13] mit dem Resolution-Parameter $\Delta R = 0,4$ rekonstruiert. Die ausgewählten und kalibrierten Jets müssen $p_T > 30$ GeV für $2,4 < |\eta| < 4,5$ und $p_T > 25$ GeV für $|\eta| < 2,4$ erfüllen. Für Jets im Bereich $|\eta| < 2,4$ gilt zusätzlich die Bedingung $JVT > 0,59$. Diese Variable wird genauer in Referenz [14] beschrieben. Mit ihr werden Jets, die nicht vom primären Vertex kommen, unterdrückt, womit sie der Reduzierung von Pileup dient.

Neben diesen Bedingungen für die Rekonstruktion werden noch weitere Schnitte angewendet, die der Reduzierung des Untergrundes dienen. Diese sind in Tabelle 6.1 dargestellt.

Die Bedingung von exakt zwei gleich geladenen Leptonen im Endzustand kann dabei aus der Betrachtung der Ereignisse in Unterabschnitt 2.2 gerechtfertigt werden. Im Gegensatz zu den ssWWjj EW-Prozessen treten bei den WZjj QCD-Prozessen auch Endzustände mit mehr als zwei nachweisbaren Leptonen auf. Die nicht nachweisbaren Neutrinos werden durch den Schnitt auf den fehlenden Transversalimpuls berücksichtigt.

Die beiden Jets mit den höchsten p_T werden als Tagging-Jets verwendet. Der Jet mit dem höchsten Transversalimpuls wird im Folgenden als erster Jet und der mit dem zweithöchsten Transversalimpuls als zweiter Jet bezeichnet. Die beiden Größen m_{jj} und Δy_{jj} werden mit diesen beiden Jets berechnet.

Die Schnitte auf den Transversalimpuls der Tagging Jets, auf m_{jj} und $|\Delta y_{jj}|$ werden im folgenden als VBS Schnitte bezeichnet.

Tabelle 6.1: Tabelle mit Schnitten für die Ereignisauswahl

Genau zwei Signal-Leptonen mit $p_T > 27$ GeV und der gleichen elektrischen Ladung mit $ \eta < 2,5$ für Myonen
mit $ \eta < 2,47$ ohne $1,37 \leq \eta \leq 1,52$ für Elektronen
mit $ \eta < 1,37$ in dem ee Kanal
$m_{ll} \geq 20$ GeV
Ereignisse mit drei oder mehr vorselektierten Leptonen werden aussortiert
$ m_{ee} - m_Z > 15$ GeV in dem ee Kanal
$E_T^{miss} \geq 30$ GeV
mindestens zwei Jets
Die zwei höchstenergetischsten Jets müssen $p_T > 65$ GeV bzw. $p_T > 35$ GeV erfüllen
$m_{jj} \geq 500$ GeV
b-Jet Veto durch den MV2c10 Tagger mit 85% working point
$ \Delta y_{jj} > 2$

7 Studie im VBS Phasenraum

Als ersten Schritt wird das Potential eines perfekten Quark-Gluon Taggers untersucht und das BDT-Tool auf die Monte-Carlo Simulationen angewendet. Dabei wird der Phasenraum nach den VBS Schnitten betrachtet.

7.1 Unterscheidung auf Truth Level - Die PartonTruthLabelID

Um die tatsächliche Verteilung von Quark- und Gluon-Jets zu bestimmen, wird zunächst die PartonTruthLabelID (PTL ID) betrachtet. Diese sagt aus, aus welchem Teilchen die Jets ursprünglich entstehen. Die Zuordnung ist in Tabelle 7.1 dargestellt.

Tabelle 7.1: Zuordnung der PTL ID zu den jeweiligen Teilchen

PTL ID	-1	1	2	3	4	5	6	21
Teilchen	<i>Pileup</i>	$d, \bar{d}$	$u, \bar{u}$	$s, \bar{s}$	$c, \bar{c}$	$b, \bar{b}$	$t, \bar{t}$	g

Die Verteilung der nach der PTL ID sortierten Jets werden in Abbildung 7.1 dargestellt. Die Anzahl der Ereignisse und die berechneten Anteile der Gluon- bzw. Pileup-Jets können Tabelle 7.2 für den ersten und Tabelle 7.3 für den zweiten Jet entnommen werden. Die Unsicherheiten auf die Anteile der Gluon- bzw. Pileup-Jets werden dabei mithilfe von Gleichung 5.2 berechnet. Insgesamt ist der Anteil an Quark-Jets sowohl in den ersten als auch in den zweiten Jets größer als der der Gluon-Jets. Die beobachtbaren größeren Anteile an Down- und Up-Quark-Jets lassen sich aus der Kollision von Protonen, die aus zwei Down- und einem Up-Quark bestehen,

erklären. Bei der Betrachtung der Gluon-Jets fällt auf, dass der Anteil von diesen sowohl im Vorwärtsbereich als auch im Zentralbereich im Untergrund größer ist als im Signal. So machen die Gluon-Jets des Signals nur ca. 4% bzw. 6% der ersten bzw. zweiten Jets aus, während es im Untergrund 21% bzw. 35% sind. Dies bestätigt das in Unterabschnitt 2.3 vorhergesagte Verhalten. Hiermit kann außerdem gesehen werden, dass der Anteil an Gluon-Jets in den zweiten Jets größer ist als in den ersten. Des Weiteren fällt auf, dass der Anteil der Gluon-Jets des Untergrundes im Zentralbereich etwas höher ist als im Vorwärtsbereich.

Der Anteil an Pileup-Jets im Untergrund ist mit ca. 9% im ersten und 13% im zweiten Jet recht klein. Im Signal macht er weniger als einen Prozent der Ereignisse aus. Im Gegensatz zu den Gluon-Jets kann festgestellt werden, dass der Anteil der Pileup-Jets des Untergrundes im Vorwärtsbereich größer ist als im Zentralbereich. Eine Ursache hierfür könnte sein, dass wie in Abschnitt 6 beschrieben bei der Auswahl der Ereignisse die Variable JVT, die zur Reduzierung von Pileup dient, nur im Zentralbereich angewendet wird.

Da die p_T -Werte aller Jets innerhalb der Trainings-Region des BDT-Tools liegen, ist die Betrachtung des Zentralbereichs gleichbedeutend mit der Anwendung des BDT-Training-Schnitte. In Abbildung 7.2 ist dargestellt, in welchen Kombinationen der erste und zweite Jet vorkommen. Im Signal tritt fast ausschließlich der Fall auf, dass beide Jets Quark-Jets sind. Im Untergrund ist diese Anordnung ebenfalls der häufigste Fall. Es können jedoch auch Ereignisse, in denen einer der Jets ein Quark-Jet und der andere ein Gluon- oder Pileup-Jet ist, beobachtet werden. Besonders häufig ist dabei der Fall, dass der erste Jet ein Quark-Jet und der zweite ein Gluon-Jet ist. Dies lässt sich auf die Beobachtung zurückführen, dass der Anteil an Gluon-Jets in den zweiten Jets insgesamt größer ist. Der Fall, dass beide Jets Gluon-Jets sind, tritt dagegen kaum ein.

7.2 Das Potential eines perfekten Quark-Gluon-Taggers

Bevor das BDT-Tool selbst getestet wird, ist es interessant herauszufinden, welches Potential ein perfekter Quark-Gluon-Tagger hat und welche Verbesserungen somit prinzipiell überhaupt erreicht werden können. Dazu werden die Signifikanzen mit und ohne einen Schnitt auf die Gluon-Jets berechnet.

7.2.1 Das Potential eines perfekten Quark-Gluon-Taggers in allen η -Bereichen

Zunächst wird das Potential eines perfekten Quark-Gluon Taggers untersucht, der keinen Einschränkungen im η Bereich unterliegt, der also Gluonen sowohl in der Zentral- als auch in der Vorwärtsregion perfekt identifizieren kann. Dazu wird die Signifikanz ohne einen Schnitt auf Gluon-Jets und nach einem Schnitt auf Gluon-Jets für unterschiedliche Szenarien berechnet. Für die Identifizierung der Jets wird dabei die PTL ID verwendet. Für die Berechnung der Signifikanz wird Gleichung 5.1 verwendet. Die Szenarien Schnitt auf Gluon-Jets der ersten Jets, Schnitt auf Gluon-Jets der zweiten Jets, Schnitt auf Ereignisse, bei denen mindestens

Tabelle 7.2: Tabelle mit der Anzahl an Ereignissen in allen η -Bereichen und den Anteilen an Gluon- und Pileup-Jets in unterschiedlichen η -Bereichen für den ersten Jet. Die Angaben in den Klammern hinter den Prozessen geben an, von welcher Jetsorte der erste Jet sein muss.

Anzahl an Ereignissen:		
Signal	53.41	± 0.90
WZ QCD	17.39	± 0.27
Signal (Gluon)	1.88	± 0.17
Signal (Quark)	51.45	± 0.90
Signal (Pileup)	0.08	± 0.03
WZ QCD (Gluon)	3.71	± 0.11
WZ QCD (Quark)	13.68	± 0.25
WZ QCD (Pileup)	1.63	± 0.12
Alle η-Bereiche:		
Signal(Gluon) / Signal	0.035	± 0.003
WZ QCD (Gluon) / WZ QCD	0.213	± 0.008
Signal(Pileup) / Signal	0.0014	± 0.0003
WZ QCD (Pileup) / WZ QCD	0.094	± 0.006
Zentralregion:		
Signal (Gluon) / Signal	0.036	± 0.004
WZ QCD (Gluon) / WZ QCD	0.221	± 0.009
Signal(Pileup) / Signal	0.001	± 0.004
WZ QCD (Pileup) / WZ QCD	0.033	± 0.004
Vorwärtsregion:		
Signal (Gluon) / Signal	0.034	± 0.006
WZ QCD (Gluon) / WZ QCD	0.186	± 0.002
Signal(Pileup) / Signal	0.003	± 0.001
WZ QCD (Pileup) / WZ QCD	0.304	± 0.018

Tabelle 7.3: Tabelle mit der Anzahl an Ereignissen in allen η -Bereichen und den Anteilen an Gluon- und Pileup-Jets in unterschiedlichen η -Bereichen für den zweiten Jet. Die Angaben in den Klammern hinter den Prozessen geben an, von welcher Jetsorte der zweite Jet sein muss.

Anzahl an Ereignissen:		
Signal	53.41	± 0.90
WZ QCD	17.39	± 0.27
Signal (Gluon)	3.17	± 0.21
Signal (Quark)	49.99	± 0.90
Signal (Pileup)	0.25	± 0.06
WZ QCD (Gluon)	6.04	± 0.15
WZ QCD (Quark)	11.35	± 0.23
WZ QCD (Pileup)	2.32	± 0.15
Alle η-Bereiche:		
Signal (Gluon) / Signal	0.059	± 0.004
WZ QCD (Gluon) / WZ QCD	0.347	± 0.009
Signal(Pileup) / Signal	0.0046	± 0.0003
WZ QCD (Pileup) / WZ QCD	0.133	± 0.006
Zentralregion:		
Signal (Gluon) / Signal	0.066	± 0.006
WZ QCD (Gluon) / WZ QCD	0.391	± 0.012
Signal(Pileup) / Signal	0.0006	± 0.0005
WZ QCD (Pileup) / WZ QCD	0.052	± 0.005
Vorwärtsregion:		
Signal (Gluon) / Signal	0.052	± 0.020
WZ QCD (Gluon) / WZ QCD	0.288	± 0.013
Signal(Pileup) / Signal	0.0097	± 0.0006
WZ QCD (Pileup) / WZ QCD	0.244	± 0.012

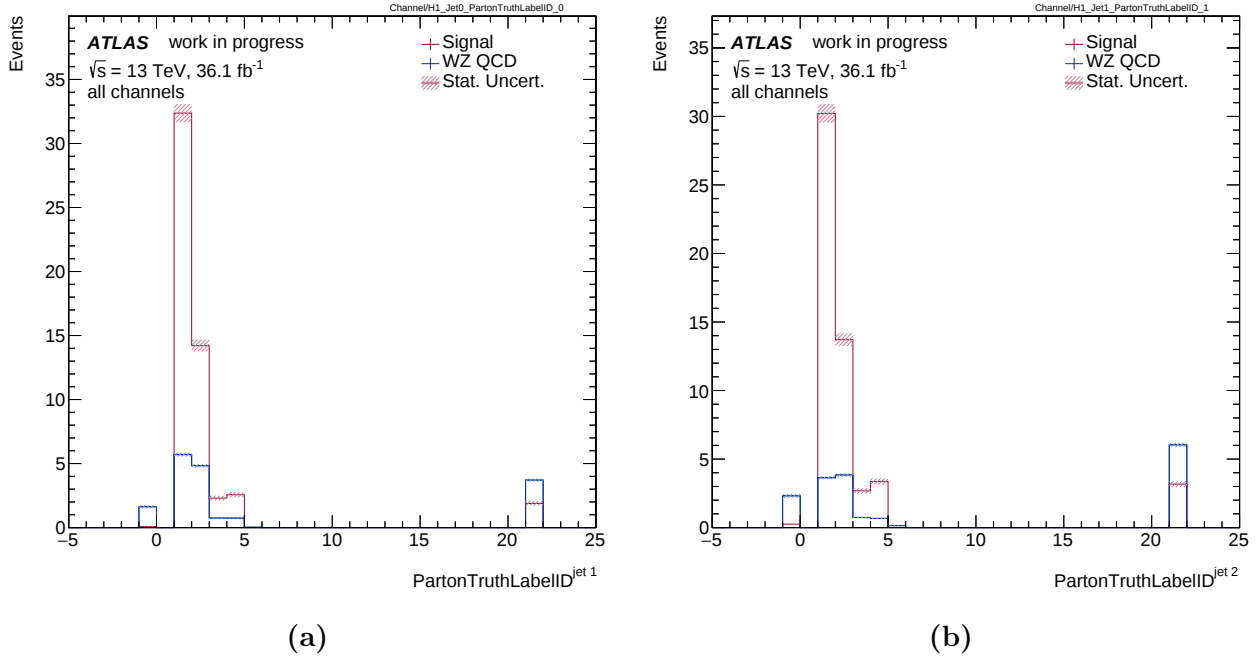


Abbildung 7.1: PTL ID des (a) ersten Jets und (b) zweiten Jets in allen η -Bereichen.

einer der Jets Gluon-initiiert ist, und Schnitt auf Ereignisse, in denen beide Jets Gluon-Jets sind, werden getestet. Die Ergebnisse sind in Tabelle 7.4 dargestellt. Es ist zu erkennen, dass ein Schnitt auf die Gluon-Jets in allen Fällen die Signifikanz erhöht. Der Ausschluss von Ereignissen mit einem Gluon-Jet als zweiten Jet führt zu einem höheren Anstieg der Signifikanz als der Ausschluss von Ereignissen mit einem Gluon-Jets als ersten Jet. Dies liegt vermutlich an dem höheren Gluon-Anteil im zweiten Jet und weist darauf hin, dass der zweite Jet besser zur Unterscheidung geeignet ist. Insgesamt erhöht sich die Signifikanz umso mehr, je mehr Gluon-Jets ausgeschlossen werden. Wie im vorherigen Abschnitt festgestellt, ist der Anteil an Ereignissen, bei denen mindestens einer der Jets ein Gluon-Jet ist, im Untergrund deutlich höher als im Signal, womit sich dieses Verhalten erklären lässt. Die höchste Signifikanz liefert insofern ein Schnitt auf alle Ereignisse mit mindestens einem Gluon-Jet.

Somit kann gesehen werden, dass im Idealfall der perfekten Identifizierung von Gluon-Jets in allen $|\eta|$ -Bereichen eine Verbesserung der Signifikanz um maximal $1,3\sigma$ erreicht werden kann.

Tabelle 7.4: Signifikanzen eines perfekten Quark-Gluon-Taggers, der Gluon-Jets in allen η Bereichen identifizieren kann.

Schnitt auf	$PTL ID < 21$ (Gluon-Jets)	Kein Schnitt
Gluon-Jets der ersten Jets	9.99σ	9.55σ
Gluon Jets der zweiten Jets	10.33σ	
Ereignisse mit mindestens einen Gluon-Jet	10.85σ	
Ereignisse mit zwei Gluon-Jets	9.71σ	

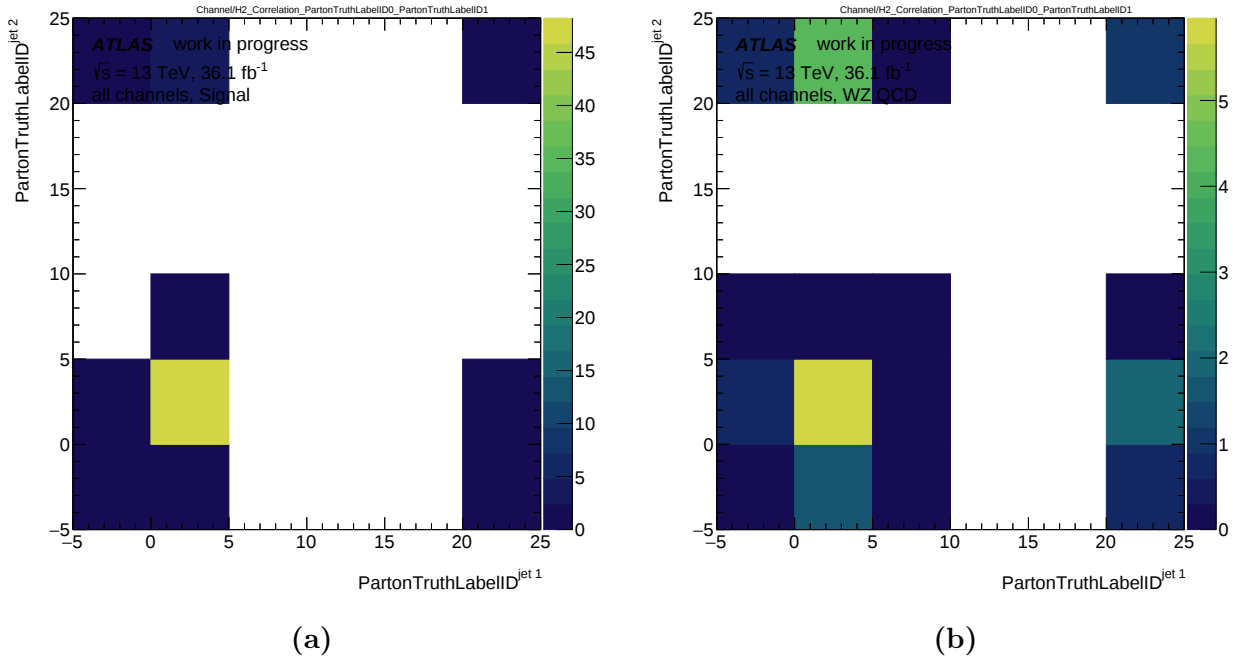


Abbildung 7.2: Korrelation der PTL ID des ersten und zweiten Jets im (a) Signal und (b) Untergrund in allen η -Bereichen. Es werden nur Bins mit einer Ereignisanzahl größer als Null gezeigt.

7.2.2 Das Potential eines perfekten Quark-Gluon-Taggers im Zentralbereich

Wie in Abschnitt 3 dargestellt, liefert der innere Detektor, aus dem die Spur-Informationen stammen, nur im Zentralbereich Informationen. Insofern können die Spur-Variablen w_{tracks} , $C_{1,tracks}$ und N_{tracks} nur für $|\eta| < 2,5$ verwendet werden. Damit ist für die Abschätzung des Potentials des Tagging-Tools die Untersuchung eines perfekten Quark-Gluon Taggers, der die unterschiedlichen Jets nur im Zentralbereich unterscheidet, praxisnäher.

In diesem Fall können für Ereignissen, in denen beide Jets im Zentralbereich zu finden sind, dieselben Schnitte auf Gluon-Jets angewendet werden wie im Abschnitt zuvor. Für Ereignisse mit mindestens einem Jet im Vorwärtsbereich gibt es dagegen mehrere Möglichkeiten diese zu behandeln. Im Fall eines Schnittes auf Gluon-Jets der ersten bzw. zweiten Jets werden alle Ereignisse mit einem Gluon-Jet im ersten bzw. zweiten Jet im Zentralbereich ausgeschlossen. In dem Fall, dass die Information von beiden Jets verwendet wird, werden folgende Kombinationen verwendet:

- Ereignisse, in denen mindestens einer der Zentralbereich-Jets ein Gluon-Jet ist, werden ausgeschlossen und für die Jets im Vorwärtsbereich gilt:
 - (a) keine weiteren Einschränkungen
 - (b) Ereignisse, in denen beide Jets im Vorwärtsbereich sind, werden ausgeschlossen
 - (c) Ereignisse, in denen nur einer der Jets im Vorwärtsbereich ist, werden ausgeschlossen

- Ereignisse, in denen beide Zentralbereich-Jets ein Gluon-Jet sind werden ausgeschlossen und für Jets im Vorwärtsbereich gilt:
 - (a) Ereignisse mit einem Gluon-Jet im Zentralbereich und einem Vorwärts-Jet werden ausgeschlossen
 - (b) keine weiteren Einschränkungen

Diese Fälle sind in Abbildung 7.3 bildlich dargestellt.

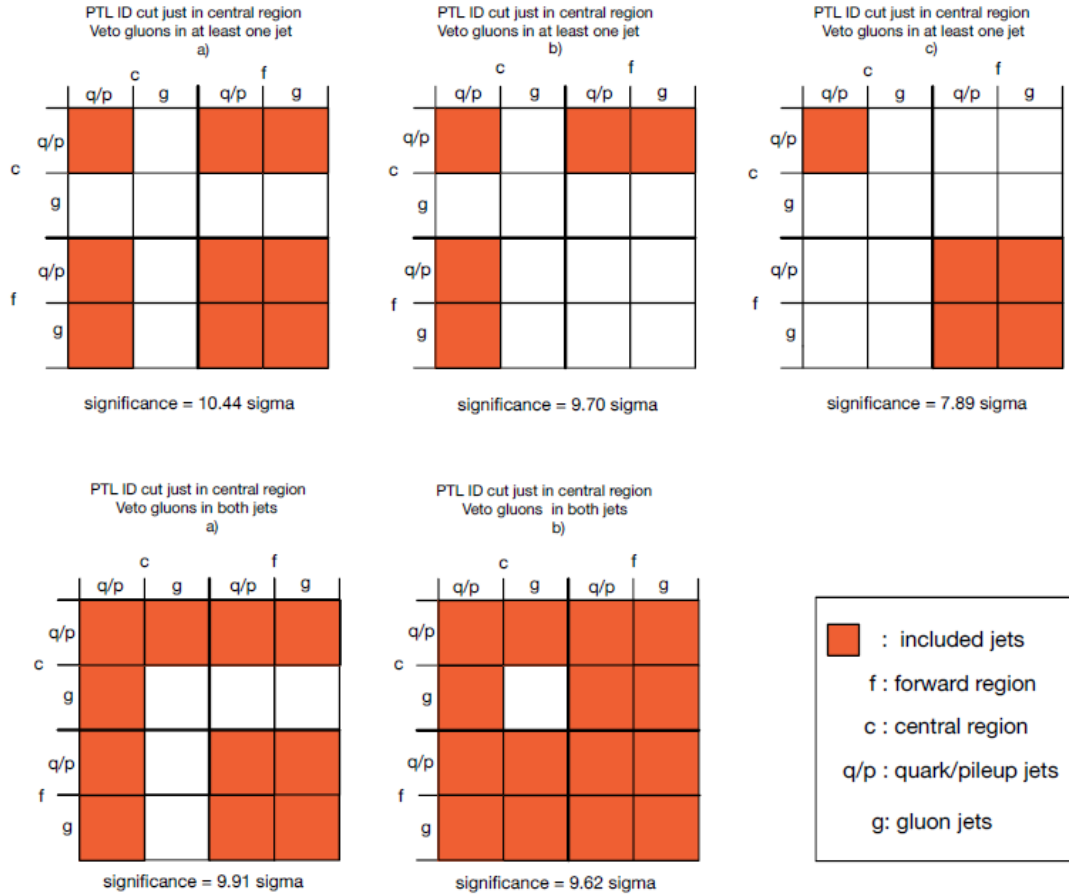


Abbildung 7.3: Veranschaulichung der Jets, die bei einem perfekten Quark-Gluon-Tagger im Zentralbereich betrachtet werden

Die so berechneten Signifikanzen sind in Tabelle 7.5 zusammengefasst. Wie im Abschnitt zuvor führt ein Schnitt auf die Gluon-Jets der zweiten Jets zu einer leicht höheren Signifikanz als ein Schnitt auf die Gluon-Jets der ersten Jets. Durch die Kombination der Information aus beiden Jets kann die Signifikanz weiter verbessert werden. Die höchste Signifikanz wird erreicht, wenn alle Ereignisse, bei denen mindestens einer der Jets ein Gluon-Jet im Zentralbereich ist, ausgeschlossen werden. Damit wird ein Anstieg der Signifikanz um $0,89 \sigma$ erreicht. Eine weitere Einschränkung der Jets im Vorwärtsbereich verschlechtert die Signifikanz wieder. Dies liegt vermutlich daran, dass durch diese Schnitte auch Ereignisse, in denen beide Jets Quark-Jets sind, ausgeschlossen werden. Somit wird durch diese Schnitte auch viel Signal weggeschnitten.

Dies trifft insbesondere auf den Fall c) zu, der sich durch eine deutlich geringere Signifikanz als in den anderen Fälle hervorhebt. Die Weiterverwendung der Fälle, in denen einer der Jets im Zentralbereich ein Gluon-Jet ist, erhöht die Signifikanz ebenfalls nicht weiter. Auch dies liegt vermutlich daran, dass diese Fälle vor allem im Untergrund vorkommen.

7.3 Das BDT-Tool

In diesem Abschnitt geht es darum, herauszufinden, ob das Potential für einen perfekten Quark-Gluon-Tagger, dass im Unterabschnitt 7.2 gefunden wurde, von dem BDT-Tool genutzt werden kann. Dazu werden die BDT-Variablen und der BDT-Wert aufgrund der Einschränkungen des Detektors nach dem BDT Training Schnitt auf den ersten oder zweiten Jet betrachtet. Wie schon zuvor beschrieben, ist dies gleichbedeutend mit der Betrachtung der Jets im Zentralbereich.

7.3.1 Differentielle Verteilung der BDT-Variablen

Für die Untersuchung der BDT-Variablen werden diese im Signal und Untergrund verglichen. Für die Unterscheidung zwischen Signal und Untergrund sollten vor allem die Gluon-Jets im Untergrund verantwortlich sein. Daher werden die Untergrundereignisse in Gluon-Jets und Quark- oder Pileup-Jets aufgeteilt. Zur genaueren Untersuchung wird außerdem das Verhältnis von Signalereignissen zu Untergrundereignissen und die Signifikanz betrachtet.

Die Ergebnisse für w_{tracks} und $C_{1,tracks}$ sind in Abbildung 7.4 und für N_{tracks} in Abbildung 7.5 jeweils für den ersten und zweiten Jet dargestellt. Es ist zu erkennen, dass die Gluon-Jets im Untergrund dazu tendieren, größere Werte der Spur-Variablen anzunehmen als die übrigen Jets. Bei dem Vergleich zwischen Signalereignissen und Untergrundereignissen kann man bei w_{tracks} und $C_{1,tracks}$ ebenfalls beobachten, dass die Untergrundereignisse eher höhere Werte annehmen als die Signalereignisse. Dies wird durch den Verhältnis-Plot bestätigt, wo das Verhältnis von Signal- zu Untergrundereignissen für große Werte abnimmt. Als Ursache hierfür kann das zuvor beschriebene Verhalten der Gluon-Jets angesehen werden. N_{tracks} hingegen weist nur einen geringen Unterschied zwischen Signal- und Untergrundereignissen auf. Insbe-

Tabelle 7.5: Signifikanzen eines perfekten Quark-Gluon-Taggers, der Gluon-Jets nur in der Zentralregion identifizieren kann.

Schnitt auf	$PTL ID < 21$ (Gluon-Jets)	Kein Schnitt
Gluon-Jets der ersten Jets	9.91 σ	9.55 σ
Gluon-Jets der zweiten Jets	10,02 σ	
Ereignisse mit mindestens einem Gluon-Jet	a) 10.44 σ	
	b) 9.70 σ	
	c) 7.89 σ	
Ereignisse mit zwei Gluon-Jets	a) 9.91 σ	
	b) 9.62 σ	

sondere für große Werte, bei denen man aufgrund des höheren Anteils von Gluon-Jets eine Unterscheidung erwarten würde, bleibt das Verhältnis von Signal- zu Untergrundereignissen fast konstant. Dies scheint zunächst den Erwartungen zu widersprechen, da diese Variable bereits zur Unterscheidung von Quark- und Gluon Jets angewendet wird (siehe Unterabschnitt 4.1). In dem Tool, das allein auf N_{track} basiert, wird allerdings eine Unterteilung in unterschiedliche p_T -Bereiche vorgenommen. Die Betrachtung der Korrelationsplots zwischen p_T und N_{tracks} , die in Abbildung 7.6 dargestellt werden, zeigt, dass auch in diesem Fall eine p_T -Abhängigkeit vorliegt. Sowohl im Signal als auch im Untergrund steigt die Anzahl von Teilchenspuren mit p_T . Dabei ist festzustellen, dass N_{tracks} im Signal schneller mit p_T steigt als im Untergrund. Insofern ist eine weitere Betrachtung der p_T -Abhängigkeit sinnvoll, um zu sehen, ob dadurch die Unterscheidungskraft weiter verbessert werden kann.

Für die Gluon-Jets im Untergrund und die Untergrundereignisse selber kann für den ersten und zweiten Jet ein vergleichbares Verhalten festgestellt werden. Der Vergleich der Verhältnissplots im ersten und zweiten Jet (Abbildung 7.4) zeigt, dass im Fall von w_{tracks} und $C_{1,tracks}$ das Verhältnis von Signal- zu Untergrundereignissen im zweiten Jet stärker abnimmt als im ersten Jet. Dies weist darauf hin, dass der zweite Jet besser zur Unterscheidung von Signal und Untergrund geeignet sein könnte. Der Grund hierfür ist vermutlich der höhere Gluon-Anteil im zweiten Jet, der in Unterabschnitt 7.1 festgestellt wurde.

Um zu sehen, ob es auch Unterschiede zwischen gleichartigen Jets im Signal und Untergrund gibt, werden die Quark-Jets im Signal mit denen im Untergrund verglichen. Das gleiche wird mit Gluon-Jets gemacht. Bei den Gluon-Jets tritt dabei das Problem auf, dass aufgrund der geringen Anzahl an diesen Ereignissen die Unsicherheiten sehr hoch sind. Innerhalb von diesen kann kein Unterschied zwischen den Gluon-Jets im Signal und Untergrund gefunden werden. Bei den Quark-Jets treten kleinere Unterschiede auf. Diese sind in Anhang B dargestellt. So scheinen die Quark-Jets im Untergrund etwas höhere w_{tracks} -Werte zu haben als im Signal. Die Signifikanz-Subplots der Spur-Variablen zeigen keine, bzw. nur vernachlässigbar kleine Maxima der Signifikanz. Insofern können die Variablen separat nicht sinnvoll zur Unterscheidung zwischen Signal und Untergrund beitragen.

In Abbildung 7.7 sind die BDT-Variablen η und p_T dargestellt. Es ist zu erkennen, dass die Lücke für η nahe Null für die Untergrundereignisse weniger stark ausgeprägt ist als für das Signal. Des Weiteren fällt auf, dass diese Lücke für den zweiten Jet stärker ausgeprägt ist als für den ersten Jet. Die Verteilung der Gluon-Jets im Untergrund ist bei dieser Variable sehr ähnlich zu jener der Quarks. Dafür kann festgestellt werden, dass die Quark-Jets im Untergrund deutlich gleichmäßiger in den η -Werten verteilt sind als im Signal. Insofern ist der Unterschied zwischen Signal und Untergrund in dieser Variable eher auf letzteren Unterschied zurückzuführen. Auch in dieser Variable kann kein Schnitt, der die Signifikanz erhöhen würde, gefunden werden. Die Betrachtung von p_T zeigt, dass Gluon-Jets vor allem bei kleineren p_T -Werten zu finden sind. Damit zeigt sich auch in Untergrundereignissen die Tendenz, eher kleine

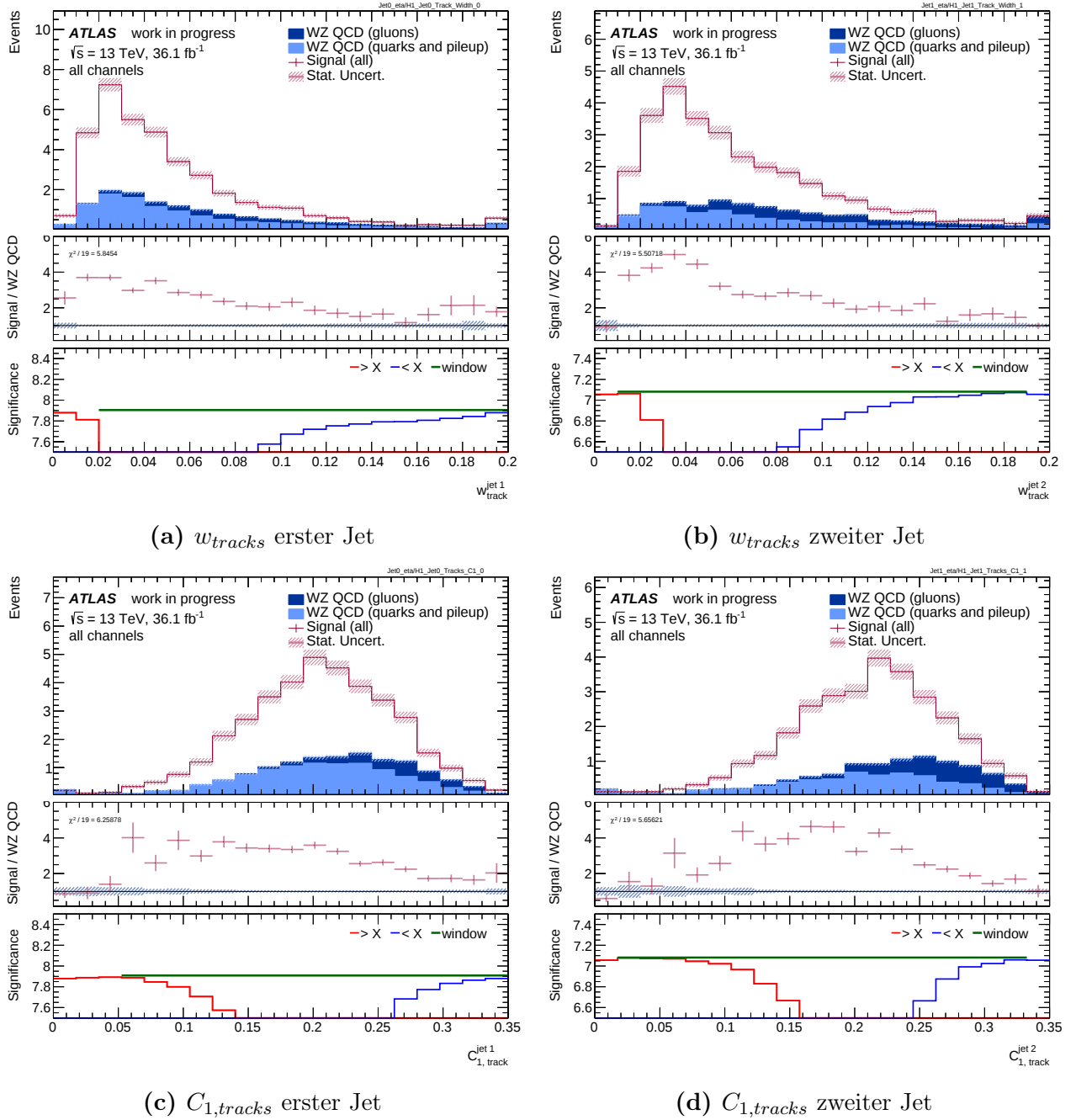


Abbildung 7.4: Plots mit den Spur-Variablen (a)(b) w_{tracks} und (c)(d) $C_{1,tracks}$ der ersten bzw. zweiten Jets nach dem BDT Training Schnitt. Die Angaben in den Klammern hinter den Prozessen geben dabei an, von welcher Jetsorte der erste bzw. zweite Jet sein muss. Die gefüllten Einträge werden aufaddiert. In den oberen Subplots ist das Verhältnis von Signal- zu Untergründereignissen dargestellt. Die unteren Subplots zeigen die Signifikanz bei unterschiedlichen Schnitten. Die rote Linie zeigt dabei die Signifikanz, die erreicht wird, wenn alle Ereignisse größer als x berücksichtigt werden. Die blaue Linie gibt die Signifikanz an, die erreicht wird, wenn nur die Ereignisse kleiner als x berücksichtigt werden. Die grüne Linie gibt ein Fenster an, in dem die Ereignissen berücksichtigt werden. Da der Fenster-Schnitt anfällig für statistische Fluktuationen ist, wird er im Folgenden nicht weiter beachtet.

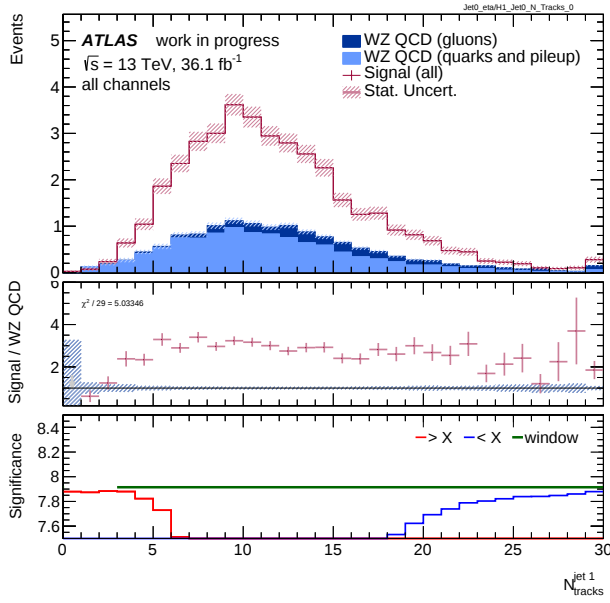
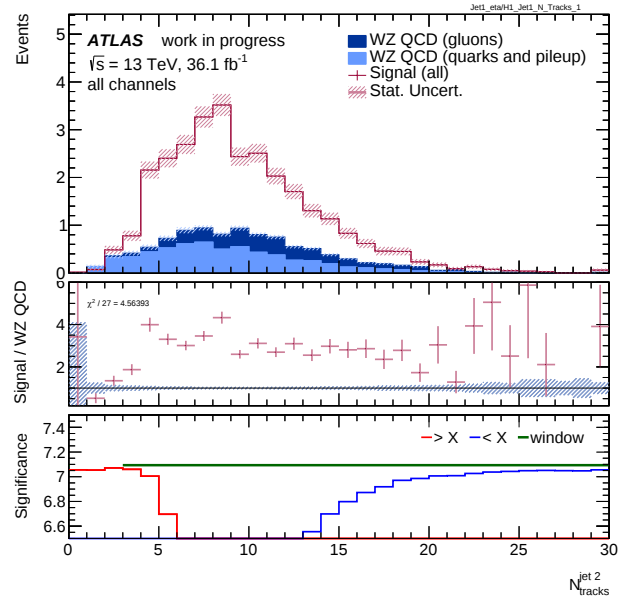
(a) N_{tracks} erster Jet(b) N_{tracks} zweiter Jet

Abbildung 7.5: N_{tracks} für die ersten und zweiten Jets nach dem BDT Training Schnitt auf den jeweiligen Jet. Die Angaben in den Klammern hinter den Prozessen geben dabei an, von welcher Jetsorte der erste bzw. zweite Jet sein muss. Die gefüllten Einträge werden aufaddiert. In den oberen Subplots ist das Verhältnis von Signal- zu Untergrundereignissen dargestellt. Die unteren Subplots zeigen die Signifikanz bei unterschiedlichen Schnitten.

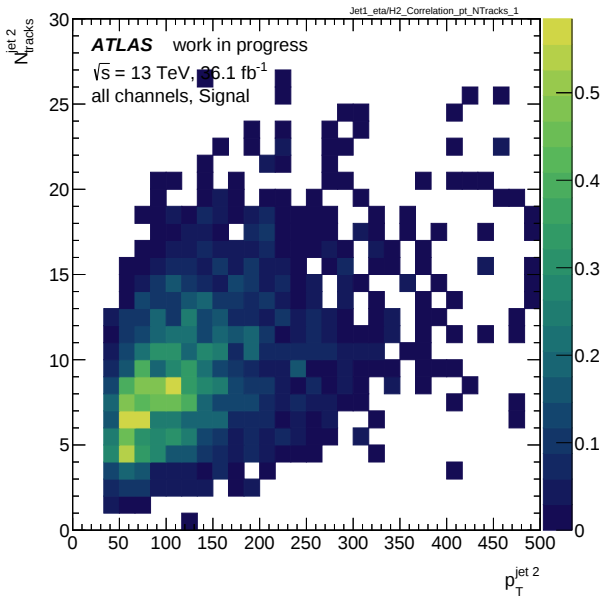
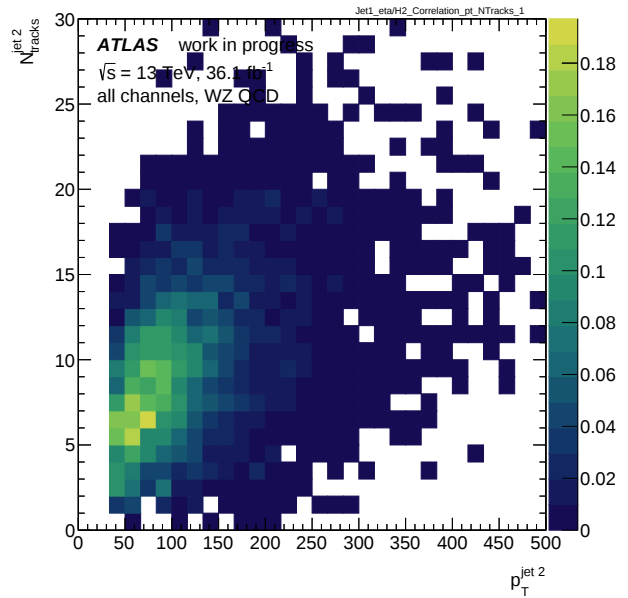
(a) Korrelation zwischen p_T und N_{tracks} im Signal des zweiten Jets(b) Korrelation zwischen p_T und N_{tracks} im Untergrund des zweiten Jets

Abbildung 7.6: Korrelationsplot zwischen p_T und η des zweiten Jets für Signal (a) und Untergrund (b) nach dem BDT Training Schnitt auf den zweiten Jet. Es werden nur Bins mit einer Ereignisanzahl größer als Null dargestellt.

Werte anzunehmen. Wie bei den Variablen zuvor scheint auch hier der zweite Jet besser zur Unterscheidung von Signal und Untergrund geeignet zu sein, da das Verhältnis bei diesem etwas stärker ansteigt. Bei dem Vergleich der Quark-Jets im Signal und Untergrund weisen diese im Untergrund leicht geringere p_T -Werte auf. Die Betrachtung des Signifikanz-Subplots (Abbildung 7.7) liefert auch hier keine Region, in der ein Schnitt signifikanzsteigernd wäre.

7.3.2 Differentielle Verteilung des BDT-Wertes

Bei den einzelnen BDT-Variablen werden zwar Unterschiede zwischen Signal und Untergrund gefunden, diese sind allein jedoch nicht ausreichend, um einen sinnvollen Beitrag zur Unterscheidung von Signal und Untergrund zu liefern. Nun wird untersucht, ob die Kombination der Variablen durch den BDT dazu ausreicht, einen Schnitt zu rechtfertigen, und ob das BDT-Tool somit zur Unterscheidung des Signals vom Untergrund geeignet ist.

In Abbildung 7.8 sind die BDT-Werte für den ersten und zweiten Jet dargestellt. Bei den Untergrundereignissen ist zu erkennen, dass die Gluon-Jets eher höhere BDT-Werte als die Quark-Jets annehmen. Insofern tendieren auch Untergrundereignisse zu höheren Werten als die Signalereignisse. Der Vergleich von Quark- bzw. Gluon-Jets im Signal mit denen im Untergrund zeigt nur geringe Unterschiede. Wie aus den Ergebnissen der einzelnen BDT-Variablen zu erwarten, scheint der zweite Jet wieder etwas besser für die Unterscheidung von Signal und Untergrund geeignet zu sein, wie man aus dem steileren Abfall im Verhältnisplot schließen kann.

Anhand des fehlenden Maximums im Signifikanz-Subplot sieht man, dass auch hier keine Verbesserung der Signifikanz durch einen Schnitt auf die einzelnen BDT-Werte erreicht werden kann.

Es besteht die Möglichkeit, dass die Kombination der beiden BDT-Werte zu einem sinnvollen Schnitt führt. Um dies herauszufinden, werden zunächst die beiden Korrelationsplots in Abbildung G.1 betrachtet. Um die Ereignisse, in denen sich der Untergrund vom Signal unterscheidet, hervorzuheben, werden im Untergrund nur die Ereignisse dargestellt, in denen mindestens einer der Jets kein Quark-Jet ist. Da kein Unterschied zwischen Quark-Jets im Signal und im Untergrund gefunden wird, sollte deren Darstellung nicht dazu beitragen, eine Region zu finden, in denen ein Schnitt möglicherweise sinnvoll wäre. Der Korrelationsplot vom Untergrund unterscheidet sich von dem im Signal vor allem durch die Ereignisse, in denen beide Jets einen vergleichsweise sehr hohen BDT-Wert haben. Damit scheint es sinnvoll, auf die Summe der BDT-Werte zu schneiden, da damit viele Ereignisse des Untergrundes ausgeschlossen werden, während viel vom Signal beibehalten wird. Der Plot mit der Summe der BDT-Werten ist in Abbildung 7.10 zu sehen. Zunächst kann erkannt werden, dass auch hier kein Maximum in der Signifikanz zu finden ist. Für Werte der Summe der BDT-Werte von ca. 0,3 steigt die Signifikanz nicht weiter. Ein Schnitt auf diese Werte würde eine merkbare Anzahl an Ereignissen reduzieren. Wenn es möglich wäre, durch eine leichte Verbesserung des BDT-

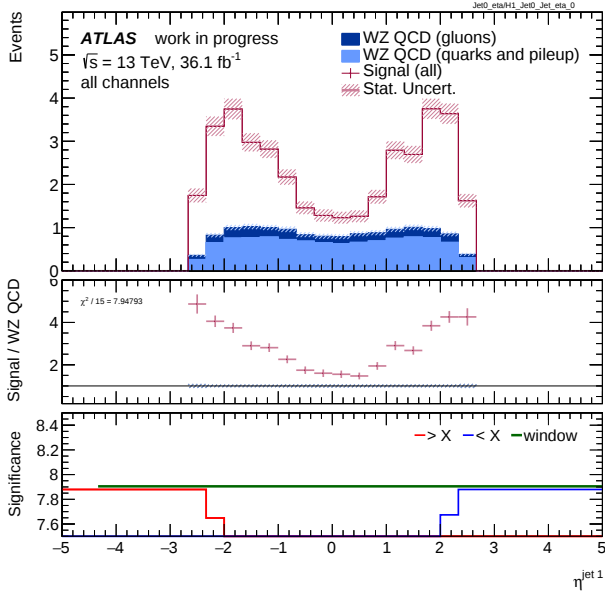
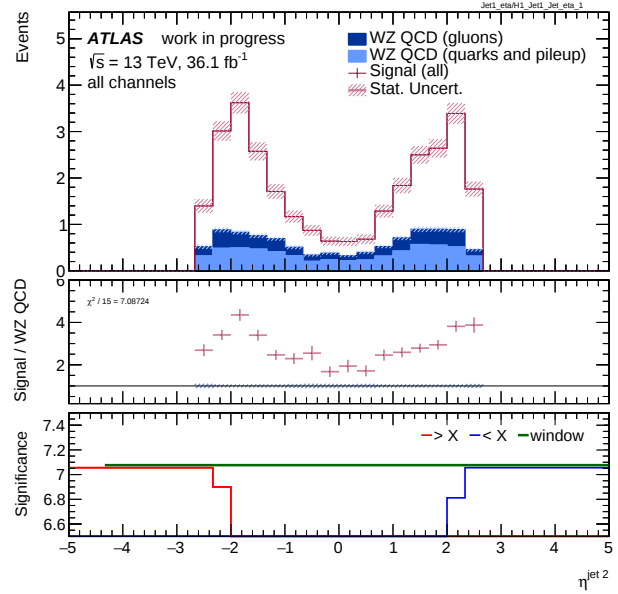
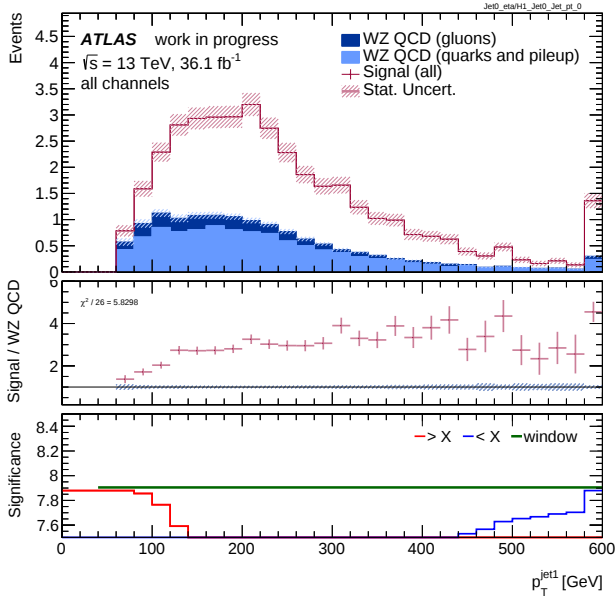
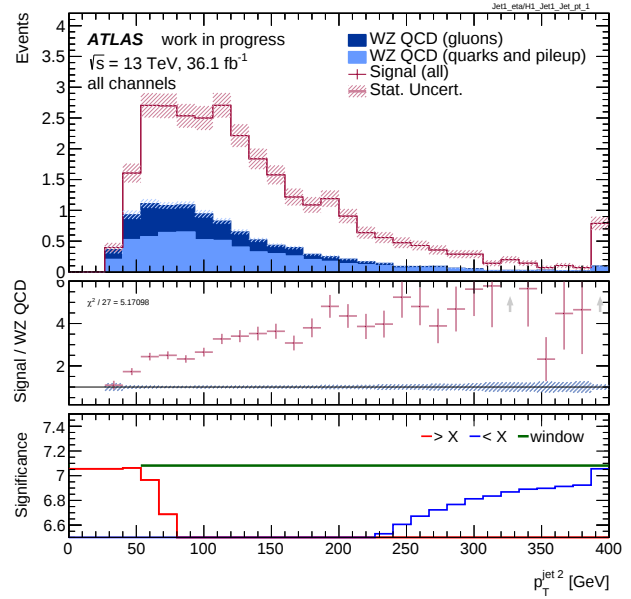
(a) η erster Jet(b) η zweiter Jet(c) p_T erster Jet(d) p_T zweiter Jet

Abbildung 7.7: η und p_T für die ersten und zweiten Jets nach dem BDT Training Schnitt auf die jeweiligen Jets. Die Angaben in den Klammern hinter den Prozessen geben dabei an, von welcher Jetsorte der erste bzw. zweite Jet sein muss. Die gefüllten Einträge werden aufaddiert. In den oberen Subplots ist das Verhältnis von Signal- zu Untergrundereignissen dargestellt. Die unteren Subplots zeigen die Signifikanz bei unterschiedlichen Schnitten.

Tools dort ein Maximum zu erzeugen, wäre ein Schnitt auf diese Werte also gegebenenfalls signifikanzsteigernd. Mit den bisherigen Ergebnissen kann jedoch kein Schnitt gerechtfertigt werden.

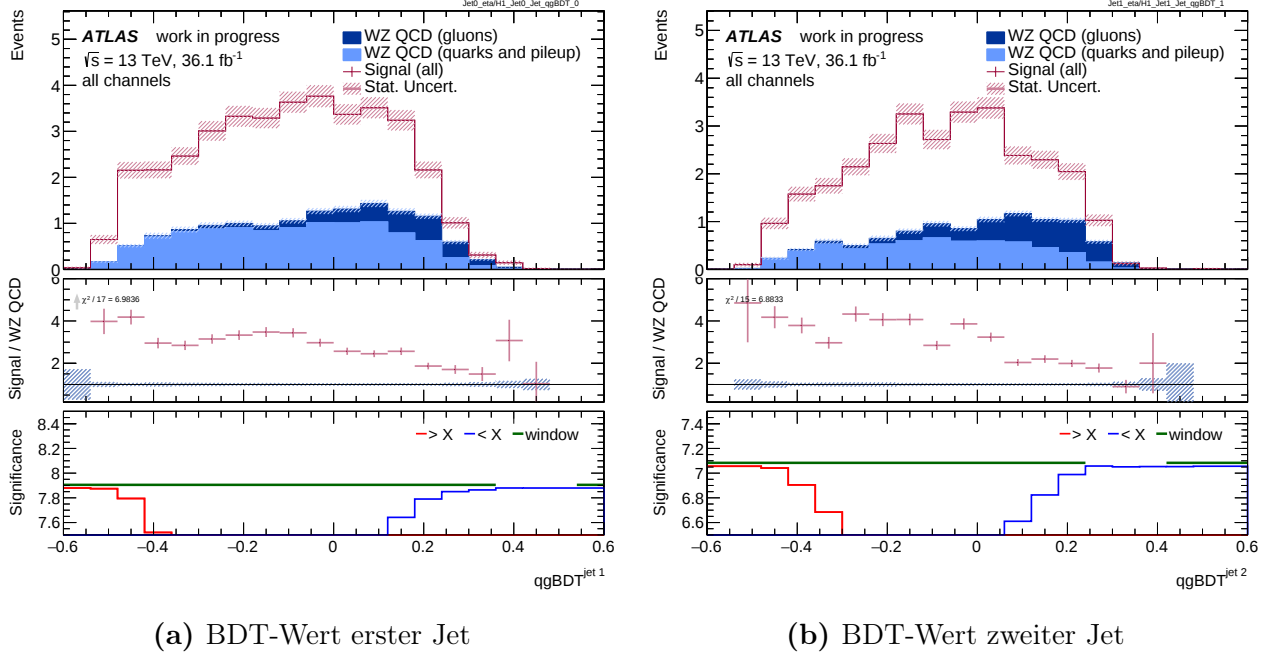
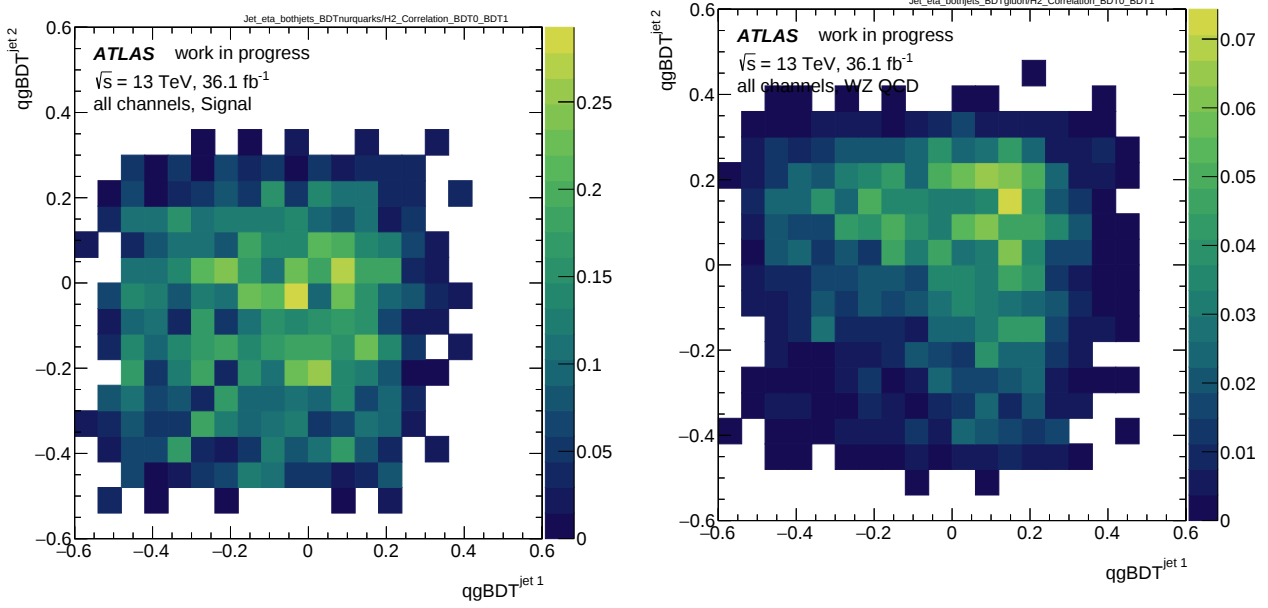


Abbildung 7.8: BDT-Werte für die ersten und zweiten Jets nach dem BDT Training Schnitt auf den jeweiligen Jets. Die Angaben in den Klammern hinter den Prozessen geben dabei an, von welcher Jetsorte der erste bzw. zweite Jet sein muss. Die gefüllten Einträge werden aufaddiert. In den oberen Subplots ist das Verhältnis von Signal- zu Untergrundereignissen dargestellt. Die unteren Subplots zeigen die Signifikanz bei unterschiedlichen Schnitten.

7.4 Abhängigkeit vom Transversalimpuls der Jets

Wie in Unterunterabschnitt 7.3.1 festgestellt, ist N_{tracks} von p_T abhängig. Insofern wird diese Variable in unterschiedlichen Bereichen von p_T betrachtet. Da der zweite Jet besser zur Unterscheidung von Quark- und Gluon-Jets geeignet scheint, wird dieser im Folgenden betrachtet. Für den zweiten Jet liegen in den betrachteten Ereignissen die p_T -Werte hauptsächlich zwischen 0 und 250 GeV. Daraus ergibt sich eine Einteilung in einem Bereich mit kleinen p_T -Werten ($50 < p_T < 100 \text{ GeV}$) und in einem mit hohen p_T -Werten ($150 < p_T < 250 \text{ GeV}$). In Abbildung 7.11 sind die N_{tracks} -Werte von Quark- und Gluon-Jets der zweiten Jets im Untergrund dargestellt. Ein Vergleich der beiden Größen im Signal wird aufgrund der geringen Anzahl von Gluon-Jets nicht durchgeführt. Zunächst ist zu erkennen, dass es einen klaren Unterschied zwischen kleinen und großen p_T -Werten von Quark- und Gluon-Jets gibt. Wie schon aus dem Korrelationsplot in Abbildung 7.6 zu erkennen ist, haben die Jets für größere p_T -Werte mehr Teilchenspuren. Besonders stark ausgeprägt ist dabei der Unterschied bei den Gluon-Jets, wie



(a) Korrelation des BDT-Werts für den ersten und zweiten Jet im Signal, wobei beide Jets ein Quark-Jet sein müssen.

(b) Korrelation des BDT-Wertes des ersten und zweiten Jets im Untergrund, wobei die Ereignisse, in denen beide Jets ein Quark-Jets sind, ausgeschlossen werden.

Abbildung 7.9: Korrelationshistogramme der BDT-Werte der ersten und zweiten Jets nach dem BDT Training Schnitt auf beide Jets. Es werden nur Bins mit einer Ereignisanzahl größer als Null betrachtet.

man in den Subplots, die die Anteile von hohen zu niedrigen p_T -Werten der jeweiligen Jets darstellen, erkennen kann. Es ist zu erkennen, dass die Unterscheidung zwischen Quark- und Gluon-Jets für höhere p_T -Werte zunimmt. Damit besteht die Möglichkeit, dass auch der BDT-Wert eine p_T -Abhängigkeit aufweist und somit gegebenenfalls ein p_T abhängiger Schnitt die Signifikanz steigern kann.

In Abbildung D.1b sind die BDT-Werte von Quark- und Gluon-Jets im Untergrund der zweiten Jets zu sehen. In den Subplots, die das Verhältnis von Quark-Jets zu Gluon-Jets für kleine und große p_T -Werte zeigen, ist zu sehen, dass der Nulldurchgang für größere p_T -Werte leicht nach rechts verschoben ist. Des Weiteren scheinen die Gluon-Jets bei größeren p_T -Werte leicht höhere Werte anzunehmen als bei kleinen p_T -Werten, während dies für die Quark-Jets nicht zu sehen ist. Diese Tendenz ist jedoch nicht sehr stark ausgeprägt und aufgrund der Unsicherheiten auch nicht eindeutig feststellbar. Der in Anhang C zu sehende Plot mit dem Verhältnis der Quark- bzw. Gluon-Jets mit kleinen p_T -Werten zu denen mit großen p_T -Werten kann dieses Verhalten ebenfalls nicht zeigen, auch wenn eine leichte Tendenz in diese Richtung erkennbar ist. Insofern scheint eine p_T -Abhängigkeit, falls vorhanden, nur leicht ausgeprägt zu sein.

Die Betrachtung der Signifikanz-Subplots des Signals und des Untergrundes für kleine bzw. großer p_T -Werte in Anhang D zeigt, dass auch hier kein Maximum der Signifikanz zu finden ist. Also kann auch ein p_T -abhängiger Schnitt nicht zur Separation von Gluon-Jets beitragen.

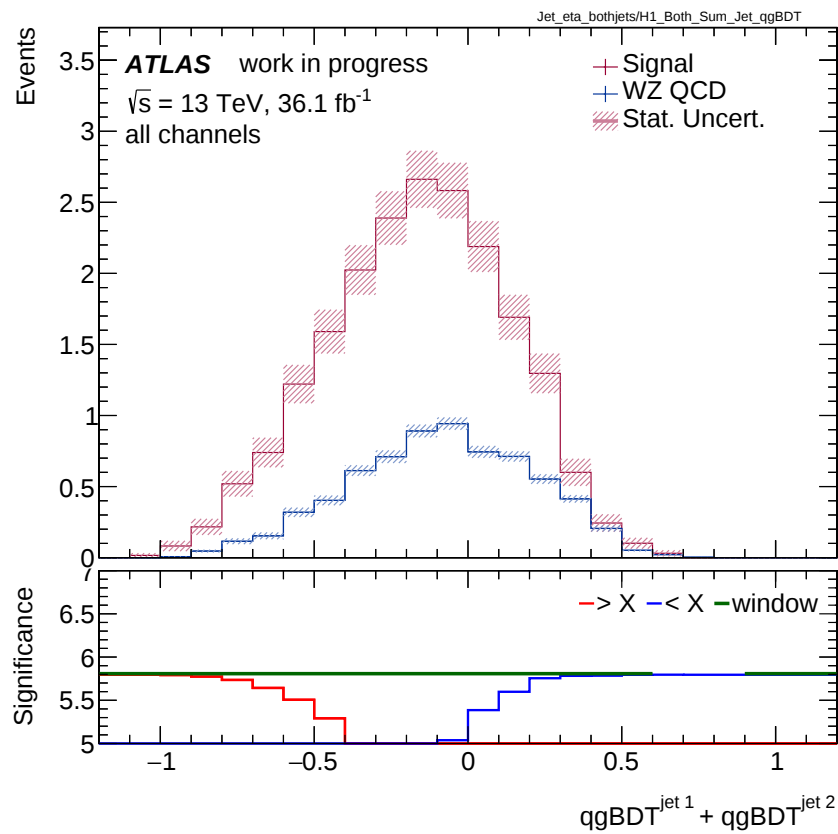


Abbildung 7.10: Summe der BDT-Werte der beiden Tagging-Jets nach den BDT Training Schnitt auf beide Jets. In dem Subplot ist die Signifikanz bei unterschiedlichen Schnitten dargestellt.

Der geringe Nutzen einer Betrachtung unterschiedlicher p_T -Bereiche lässt sich vermutlich damit erklären, dass die Jets vor allem relativ kleine p_T -Werte annehmen. Es ist zu vermuten, dass die Unterscheidungskraft für größere p_T -Werte noch zunehmen würde.

7.5 Abhängigkeit von der durchschnittlichen Anzahl an Interaktionen pro Teilchenbündel

In diesem Abschnitt soll überprüft werden, ob es einen Zusammenhang zwischen den BDT-Variablen bzw. dem BDT-Wert und der Anzahl von Pileup gibt. Dazu wird die durchschnittliche Anzahl von Interaktionen pro Kreuzung der Teilchenbündel (avIntPerX, aus englischem average interactions per bunch crossing) betrachtet. Mit einer zunehmenden Anzahl an Interaktionen kommt es dabei erwartungsgemäß auch zu einer höheren Anzahl an Pileup. Zur Untersuchung wird die avIntPerX in vier Bins mit ungefähr gleich vielen Ereignissen eingeteilt. Diese sind in Abbildung 7.13 dargestellt. Für diese Bins werden dann Quarks-Jets und die Gluon-Jets im Signal bzw. Untergrund betrachtet. Dabei kann keine Abhängigkeit der BDT-Variablen und des BDT-Wertes von der avIntPerX gefunden werden. Als Beispiel ist der BDT-Wert der Gluon-Jets im Untergrund für die zweiten Jets in Abbildung 7.14a dargestellt. Die anderen Variablen können im Anhang E gefunden werden. Die mangelnde Abhängigkeit kann damit erklärt werden, dass die Teilchenspuren sich gut mit ihrem jeweiligen Vertex assoziieren lassen und die Spur-Variablen somit kaum vom Pileup abhängen sollten.

Der Anteil an Gluon-Jets könnte ebenfalls vom Pileup abhängen. Bei den betrachteten Prozessen entstehen vor allem Quark-Jets. Allgemein gibt es hingegen viele Prozesse, bei denen Gluon-Jets entstehen, sodass der Anteil an Gluon-Jets bei mehr Pileup steigen sollte. Mit der PTL ID werden die Pileup-Jets separat von den Gluon-Jets identifiziert. Somit werden durch Pileup hinzukommende Gluon-Jets nicht als solche identifiziert. Damit ist zu erwarten, dass der berechnete Anteil an Gluon-Jets sinkt. Die berechneten Werte sind in Tabelle 7.6 für den ersten und in Tabelle 7.7 für den zweiten Jet dargestellt. Im Untergrund kann die Tendenz zu den erwarteten Verhalten beobachtet werden. Es muss jedoch beachtet werden, dass aufgrund der großen Unsicherheiten keine eindeutige Aussage getroffen werden kann. In Abbildung 7.14b ist die PTL ID für die unterschiedlichen avIntPerX-Bins dargestellt. Hier kann gesehen werden, dass die Anzahl an Tagging-Jets, die als Pileup-Jets identifiziert werden, nicht von der avIntPerX abhängt. Damit kann die beobachtbare Tendenz nicht auf die Pileup-Jets zurückgeführt werden. Im Signal scheint es den Trend zu geben, dass mit höherem avIntPerX der Anteil an Gluon-Jets steigt. Es muss jedoch beachtet werden, dass im Signal die Anzahl der Gluon- und Pileup-Jets sehr gering ist und insofern nur schwer Aussagen über deren Verhalten getroffen werden können. Insbesondere die Pileup-Jets machen weniger als einen Prozent der Jets aus, womit deren Einfluss vernachlässigbar sein sollte.

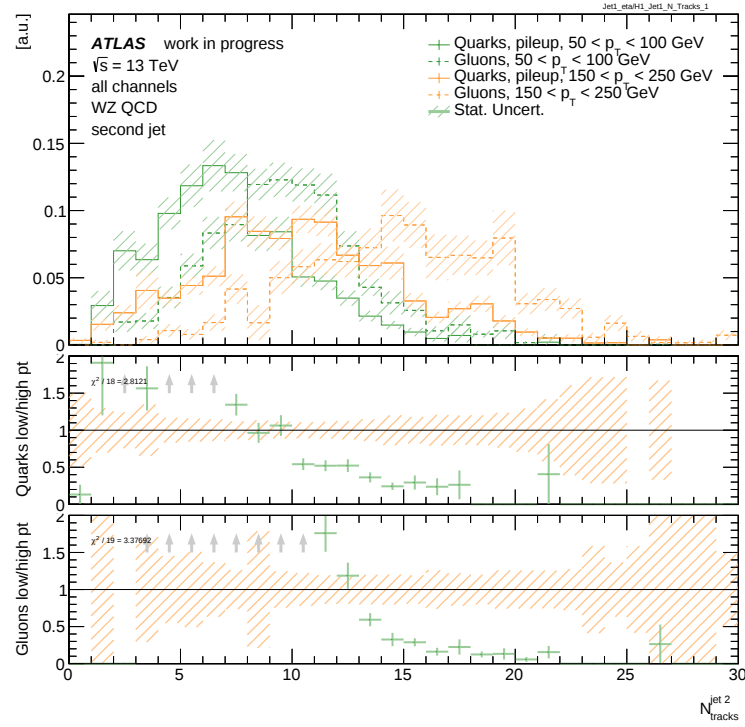


Abbildung 7.11: Normierte N_{tracks} -Werte der zweiten Jets in unterschiedlichen p_T -Regionen nach dem BDT Training Schnitt auf die zweiten Jets. Die Bezeichnungen Quarks, Pileup und Gluon steht dabei für die jeweiligen Jets. Der obere Subplot zeigt das Verhältnis von Quark-Jets dieses Verhältnis für Gluon-Jets.

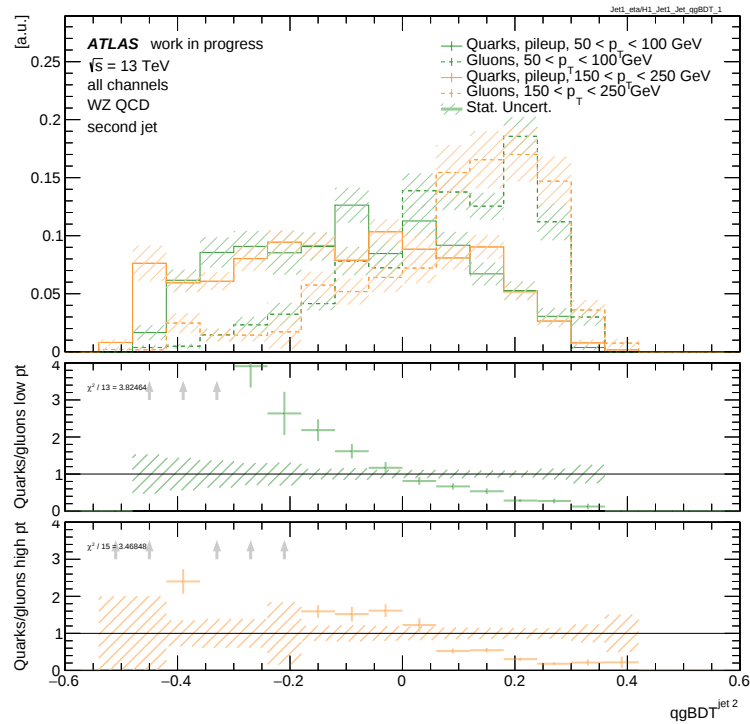


Abbildung 7.12: Normierte BDT-Werte des zweiten Jets in unterschiedlichen p_T -Regionen nach dem BDT Training Schnitt auf den zweiten Jet. Die Bezeichnungen Quarks, Pileup und Gluon steht dabei für die Jetssorte der zweiten Jets. Der obere Subplot zeigt das Verhältnis von Quark- zu Gluon-Jets bei kleinen p_T -Werten und der untere dieses Verhältnis für große p_T -Werte.

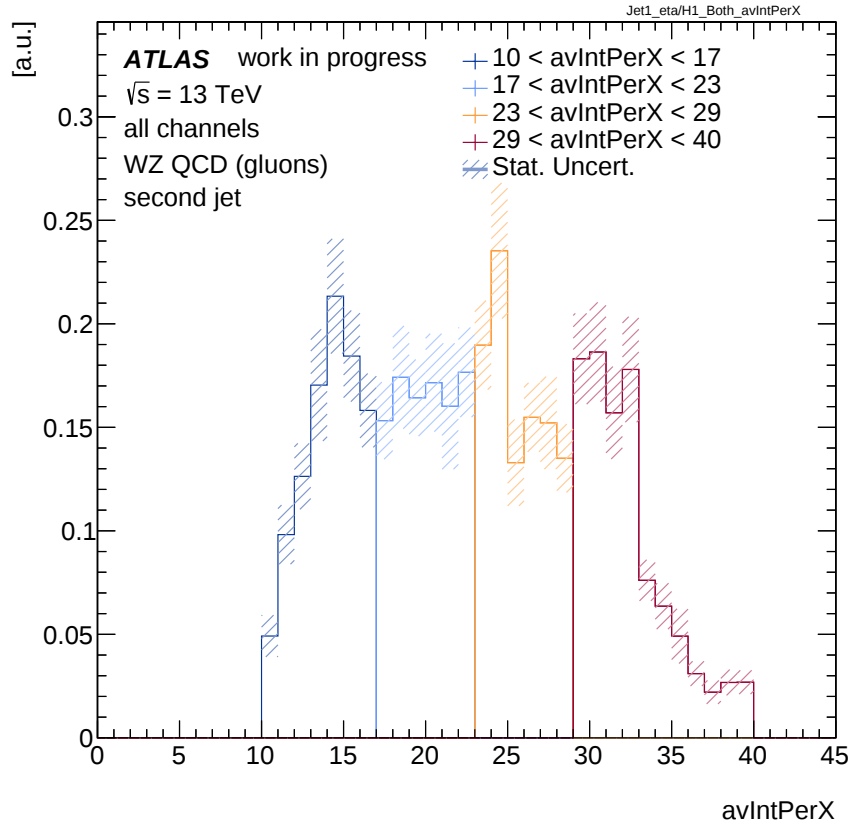
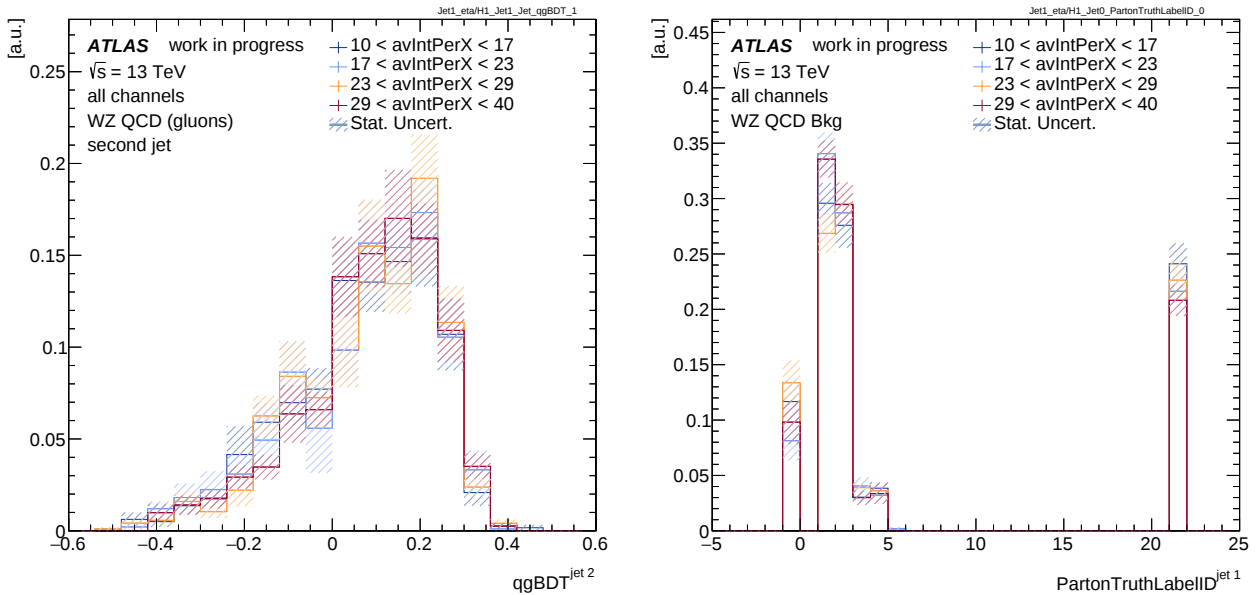


Abbildung 7.13: Unterteilung der avIntPerX in unterschiedliche Bins



(a) Normierter Plot der BDT-Werte der Gluon-Jets der zweiten Jets im Untergrund nach dem BDT Training Schnitt auf die zweiten Jets.

(b) Normierter Plot mit der PTL ID der zweiten Jets im Untergrund nach dem BDT Training Schnitt auf die zweiten Jets.

Abbildung 7.14: Variablen für unterschiedliche avIntPerX Bins.

Tabelle 7.6: Anteil an Gluon-Jets in unterschiedlichen avIntPerX Gebieten nach dem BDT Training Schnitt auf die ersten Jets. Die Angaben in den Klammern hinter den Prozessen geben an, von welcher Jetsorte der erste Jet sein muss. Dabei gilt: Region 1 : $10 < \text{avIntPerX} < 17$, Region 2: $17 < \text{avIntPerX} < 23$, Region 3 : $23 < \text{avIntPerX} < 29$, Region 4 : $29 < \text{avIntPerX} < 40$.

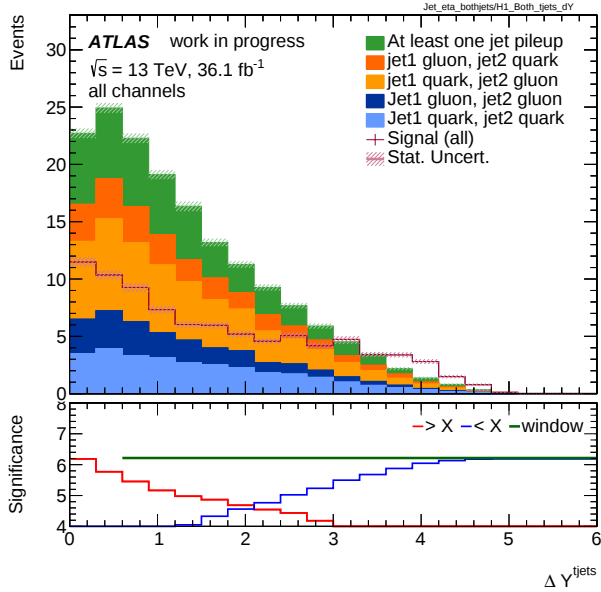
Signal(Gluon) / Signal, Region 1	0.026 ± 0.007
Signal(Gluon) / Signal, Region 2	0.039 ± 0.007
Signal(Gluon) / Signal, Region 3	0.033 ± 0.007
Signal(Gluon) / Signal, Region 4	0.047 ± 0.010
WZ QCD(Gluon)/ WZ QCD, Region 1	0.251 ± 0.011
WZ QCD(Gluon)/ WZ QCD, Region 2	0.203 ± 0.010
WZ QCD(Gluon)/ WZ QCD, Region 3	0.219 ± 0.010
WZ QCD(Gluon)/ WZ QCD, Region 4	0.227 ± 0.015

Tabelle 7.7: Anteile an Gluon-Jets in unterschiedlichen avIntPerX Gebieten nach dem BDT Training Schnitt auf die zweiten Jets. Die Angaben in den Klammern hinter den Prozessen geben an, von welcher Jetsorte der zweite Jet sein muss. Dabei gilt: Region 1 : $10 < \text{avIntPerX} < 17$, Region 2: $17 < \text{avIntPerX} < 23$, Region 3 : $23 < \text{avIntPerX} < 29$, Region 4 : $29 < \text{avIntPerX} < 40$.

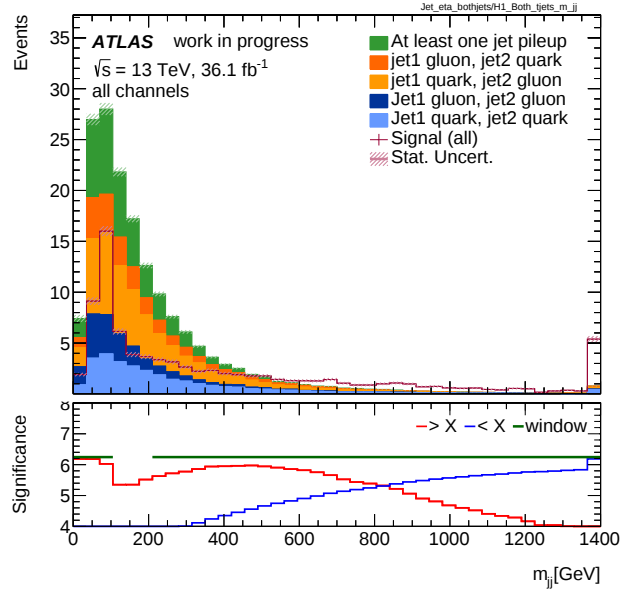
Signal(Gluon) / Signal, Region 1	0.051 ± 0.020
Signal(Gluon) / Signal, Region 2	0.066 ± 0.016
Signal(Gluon) / Signal, Region 3	0.065 ± 0.017
Signal(Gluon) / Signal, Region 4	0.083 ± 0.020
WZ QCD(Gluon)/ WZ QCD, Region 1	0.419 ± 0.027
WZ QCD(Gluon)/ WZ QCD, Region 2	0.396 ± 0.023
WZ QCD(Gluon)/ WZ QCD, Region 3	0.386 ± 0.022
WZ QCD(Gluon)/ WZ QCD, Region 4	0.374 ± 0.026

8 Studie im Phasenraum ohne die VBS Schnitte

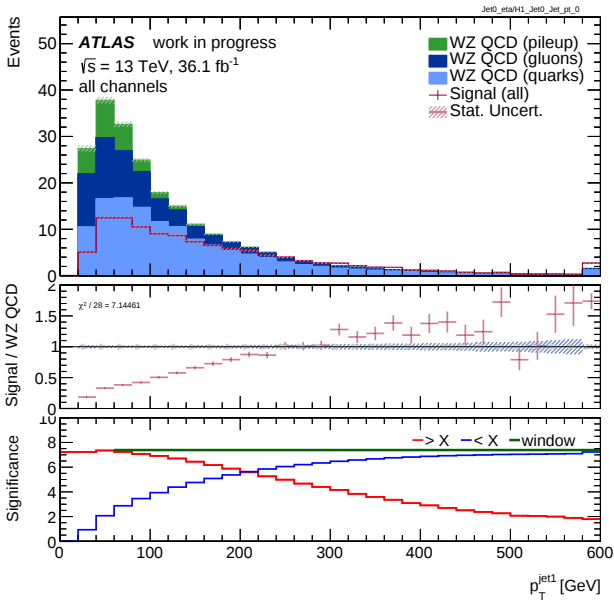
In dem bisherigen Phasenraum kann kein Schnitt zur signifikanzsteigernden Unterscheidung zwischen Quark- und Gluon-Jets gefunden werden. Es besteht die Möglichkeit, dass durch die Betrachtung eines erweiterten Phasenraums Verbesserungen gefunden werden können. Daher wird im Folgenden die Untersuchung des Tagging-Tools ohne die VBS Schnitte durchgeführt. Um den Einfluss der reduzierten Schnitte zu sehen, sind in Abbildung 8.1 die Variablen, auf die die VBS Schnitte wirken, dargestellt. Für p_T wird dabei immer noch ein Schnitt auf 20 GeV angewendet. Insbesondere bei $|\Delta y_{jj}^{tjets}|$ und m_{jj}^{tjets} werden durch die fehlenden Schnitte viele Ereignisse beibehalten. Besonders die Anzahl der Untergrundereignisse steigt dadurch stark an. Betrachtet man die Kombinationen, in denen der erste und der zweite Jet vorkommen, kann festgestellt werden, dass viele Ereignisse aus mindestens einem Gluon- oder Pileup-Jet bestehen.



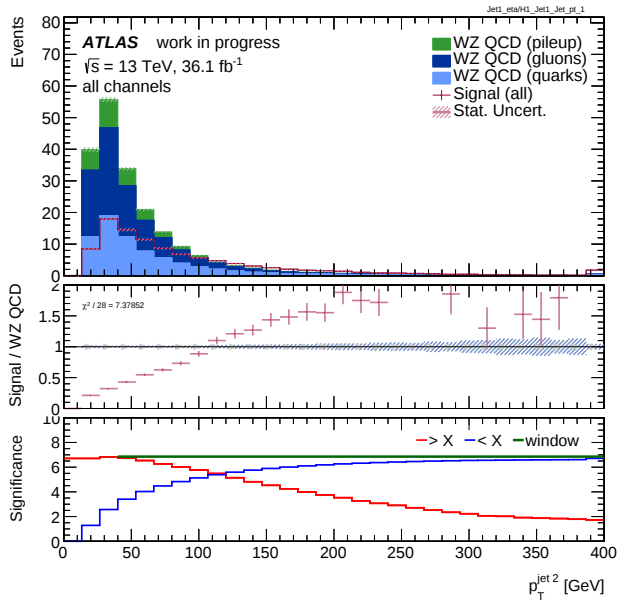
(a) Unterschied der Rapidität der Tagging-Jets



(b) Invariante Masse der beiden Tagging-Jets



(c) Transversalimpuls des ersten Tagging-Jets



(d) Transversalimpuls des zweiten Tagging-Jets

Abbildung 8.1: VBS-Variablen ohne die VBS-Schnitte. Die ausgefüllten Einträge werden aufaddiert. ΔY und m_{jj} sind nach dem BDT Training Schnitt auf beide Jets und p_T nach dem BDT Training Schnitt von dem betrachteten Jet dargestellt.

8.1 Die PartonTruthLabelID

Die Verteilung der PTL ID ohne VBS Schnitte ist in Abbildung 8.2 für den ersten und zweiten Jet dargestellt. Die Anzahl an Ereignissen und die berechneten Anteile an Gluon- und Pileup-Jets können in Tabelle 8.1 für die ersten Jets und in Tabelle 8.2 für die zweiten Jets gefunden werden. Wie zu erwarten steigt der Anteil der Gluon-Jets sowohl im Signal als auch im Untergrund im Vergleich zu den Ereignissen mit den VBS Schnitte deutlich an. Bei den zweiten Jets macht er hier beispielsweise ca. 48% statt 35% der Jets aus. Des Weiteren kann erkannt werden, dass der Anteil an Pileup-Jets im Untergrund deutlich ansteigt. Bei den zweiten Jets macht er nun z.B. ca. 22% der Jets aus. Im Signal ist der Anteil an Pileup-Jets weiterhin sehr gering. Der Gluon-Anteil ist in den zweiten Jets weiterhin größer als in den ersten. Das gleiche gilt für die Pileup-Jets im Untergrund. Wie bei der Betrachtung mit den VBS Schnitte ist der Gluon-Anteil im Zentralbereich größer als im Vorwärtsbereich, während für die Pileup-Jets das Gegenteil gilt.

8.2 Differentielle Verteilung der BDT Variablen

Die BDT-Variablen der zweiten Jets ohne VBS Schnitte sind in Abbildung 8.3 nach dem BDT Training Schnitt auf die zweiten Jets dargestellt. Die Variablen der ersten Jets weisen ein vergleichbares Verhalten auf und können in Anhang F gefunden werden. Der Plot für p_T kann aus Abbildung 8.1 entnommen werden. Da im Abschnitt zuvor festgestellt wird, dass der Anteil an Pileup-Jets ohne die VBS Schnitte relativ groß ist, werden diese im Folgenden separat von den Quark-Jets betrachtet. Somit werden die Untergrundereignisse in Quark-, Gluon- und Pileup-Jets aufgeteilt. Bei den Spur-Variablen kann erkannt werden, dass wie zuvor die Gluon-Jets im Untergrund zu höheren Werten der Variablen tendieren als die Quark-Jets. Die Pileup-Jets hingegen treten vor allem bei kleinen Werten auf. Damit verringert sich insgesamt der Unterschied zwischen Signal und Untergrund. Dementsprechend müssen Gluon- und Pileup-Jets durch separate Schnitte reduziert werden. Besonders auffällig ist die Anhäufung von Ereignissen bei $C_{1,tracks} = 0$, die durch die Pileup-Jets verursacht wird. Durch einen Schnitt auf diese Werte könnten somit viele Pileup-Jets ausgeschlossen werden. Durch das leichte Maximum im Signifikanz-Subplot für kleine $C_{1,tracks}$ -Werte bestätigt sich, dass ein Schnitt hier die Signifikanz erhöhen könnte. Es sollte jedoch vorher geprüft werden, ob das Pileup im Vergleich zu den Daten gut modelliert ist. Bei der Betrachtung der Signifikanz bei rechtsseitigen Schnitten kann nur bei w_{tracks} ein leichtes Maximum bei ca. 0,2 gefunden werden.

Für N_{tracks} kann festgestellt werden, dass hier die Untergrundereignisse kleinere Werte als die Signalereignisse annehmen. Da der Untergrund mehr Gluon-Jets enthält und Gluon-Jets eher zu höheren N_{tracks} -Werten neigen, scheint dies zunächst widersprüchlich zu sein. Dies kann auch nicht allein durch die Pileup-Jets, die vor allem bei kleinen N_{tracks} -Werten auftreten, erklärt werden. Wie schon bei der Betrachtung mit den VBS-Schnitte muss hier wieder

Tabelle 8.1: Tabelle mit der Anzahl an Ereignissen in allen η -Bereichen und den Anteilen an Gluon- und Pileup-Jets in unterschiedlichen η -Bereichen für die ersten Jets ohne die VBS Schnitte. Die Angaben in den Klammern hinter den Prozessen geben an, von welcher Jetsorte der erste Jet sein muss.

Anzahl an Ereignissen:		
Signal	140.5	± 1.5
WZ QCD	267.9	± 1.7
Signal(Gluon)	10.9	± 0.4
Signal(Quark)	129.1	± 1.4
Signal(Pileup)	0.47	± 0.08
WZ QCD(Gluon)	76.0	± 0.9
WZ QCD(Quark)	142.2	± 1.0
WZ QCD(Pileup)	49.7	± 1.0
Alle η-Bereiche:		
Signal(gluon) / Signal	0.078	± 0.003
WZ QCD(gluon) / WZ QCD	0.284	± 0.002
Signal(Pileup) / Signal	0.0003	± 0.0006
WZ QCD(Pileup) / WZ QCD	0.186	± 0.002
Zentralregion:		
Signal(gluon) / Signal	0.080	± 0.003
WZ QCD(gluon) / WZ QCD	0.300	± 0.002
Signal(Pileup) / Signal	0.0007	± 0.0003
WZ QCD(Pileup) / WZ QCD	0.118	± 0.002
Vorwärtsregion:		
Signal(gluon) / Signal	0.069	± 0.006
WZ QCD(gluon) / WZ QCD	0.214	± 0.005
Signal(Pileup) / Signal	0.016	± 0.003
WZ QCD(Pileup) / WZ QCD	0.481	± 0.006

Tabelle 8.2: Tabelle mit der Anzahl an Ereignissen in allen η -Bereichen und den Anteilen an Gluon- und Pileup-Jets in unterschiedlichen η -Bereichen für die zweiten Jets ohne VBS Schnitte. Die Angaben in den Klammern hinter den Prozessen geben an, von welcher Jetsorte der zweite Jet sein muss.

Anzahl an Ereignissen		
Signal	140.5	± 1.5
WZ QCD	267.9	± 1.7
Signal(Gluon)	19.81	± 0.56
Signal(Quark)	118.1	± 1.3
Signal(Pileup)	0.47	± 1.08
WZ QCD(Gluon)	116.0	± 1.0
WZ QCD(Quark)	92.7	± 0.9
WZ QCD(Pileup)	59.2	± 0.9
Alle η-Bereiche:		
Signal(gluon) / Signal	0.141	± 0.004
WZ QCD(gluon) / WZ QCD	0.433	± 0.002
Signal(Pileup) / Signal	0.018	± 0.001
WZ QCD(Pileup) / WZ QCD	0.221	± 0.002
Zentralbereich:		
Signal(gluon) / Signal	0.156	± 0.001
WZ QCD(gluon) / WZ QCD	0.480	± 0.003
Signal(Pileup) / Signal	0.018	± 0.002
WZ QCD(Pileup) / WZ QCD	0.219	± 0.002
Vorwärtsregion:		
Signal(gluon) / Signal	0.100	± 0.006
WZ QCD(gluon) / WZ QCD	0.296	± 0.004
Signal(Pileup) / Signal	0.018	± 0.003
WZ QCD(Pileup) / WZ QCD	0.230	± 0.005

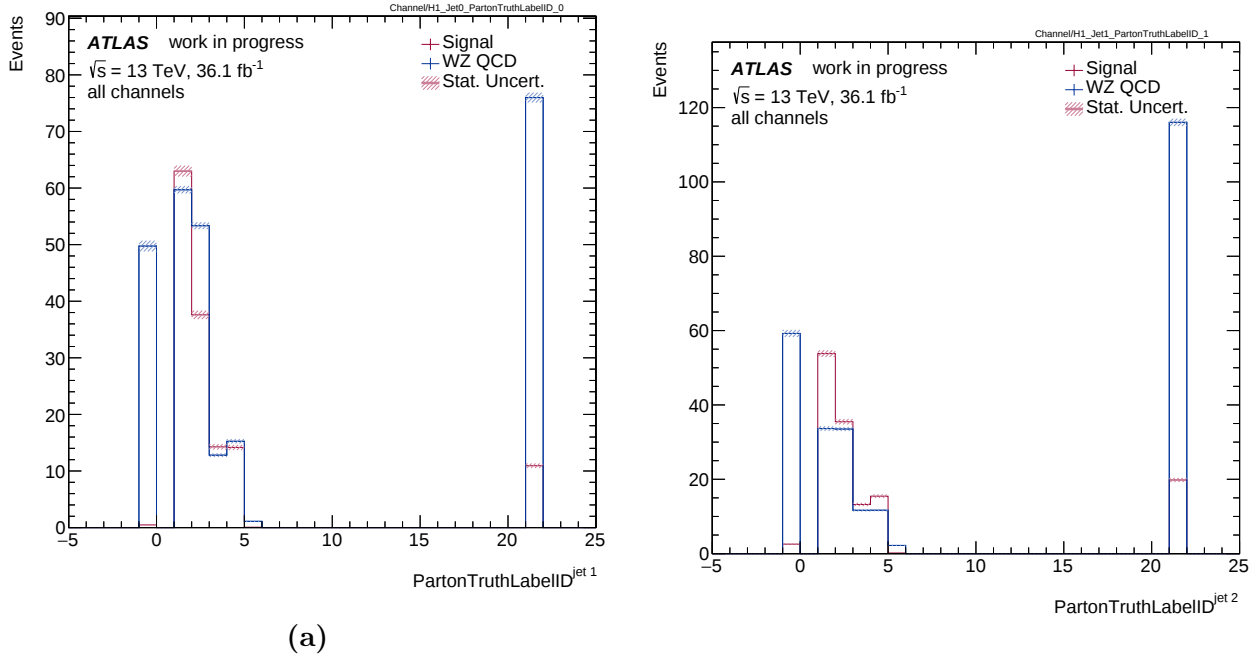


Abbildung 8.2: PTL ID des (a) ersten Jets und (b) zweiten Jets in allen η -Bereichen.

die p_T -Abhängigkeit dieser Variable berücksichtigt werden. Wie in Abbildung 8.4 zu sehen ist, tendieren die Quark-Jets im Signal zu höheren p_T -Werten als im Untergrund. Durch die Zunahme der Anzahl an Teilchenspuren mit p_T weisen damit auch die Quark-Jets im Signal höhere N_{tracks} -Werte auf als im Untergrund. Gleiches kann für die Gluon-Jets beobachtet werden. Der Vergleich der N_{tracks} -Werte im Signal und Untergrund für kleine und große p_T -Werte zeigt, dass hier die Untergrundereignisse etwas größere Werte annehmen. Dies ist in Abbildung 8.5 dargestellt. Der Unterschied zwischen Signal und Untergrund ist jedoch nicht sehr groß. Dies liegt vermutlich wieder daran, dass die Jets vor allem relativ kleine p_T -Werte annehmen. Des Weiteren ist weiterhin der Einfluss der Pileup-Jets zu berücksichtigen.

Bei der Betrachtung von η fällt auf, dass hier die Lücke für $|\eta|$ nahe Null, die mit den VBS Schnitten beobachtet werden konnte, nicht mehr auftaucht. Im Signal bleiben die η -Werte für einen Großteil der Werte nahezu konstant, während im Untergrund für $|\eta| \approx 0$ ein Maximum angenommen wird. Im Gegensatz zu den anderen Variablen kann zwischen den Quark-, Gluon- und Pileup-Jets im Untergrund kein Unterschied in der Verteilung festgestellt werden.

Bei den p_T -Werten scheinen sowohl Gluon- als auch Pileup-Jets eher bei niedrigen Werten vorzukommen. Damit tendieren auch die Untergrundereignisse zu kleineren p_T -Werten als Ereignisse im Signal.

8.3 Differentielle Verteilung des BDT-Wertes

Der BDT-Wert ohne die VBS Schnitte ist in Abbildung 8.6 für die zweiten Jets dargestellt. Der Plot für die ersten Jets kann in Anhang F gefunden werden. Es kann das gleiche Verhalten wie

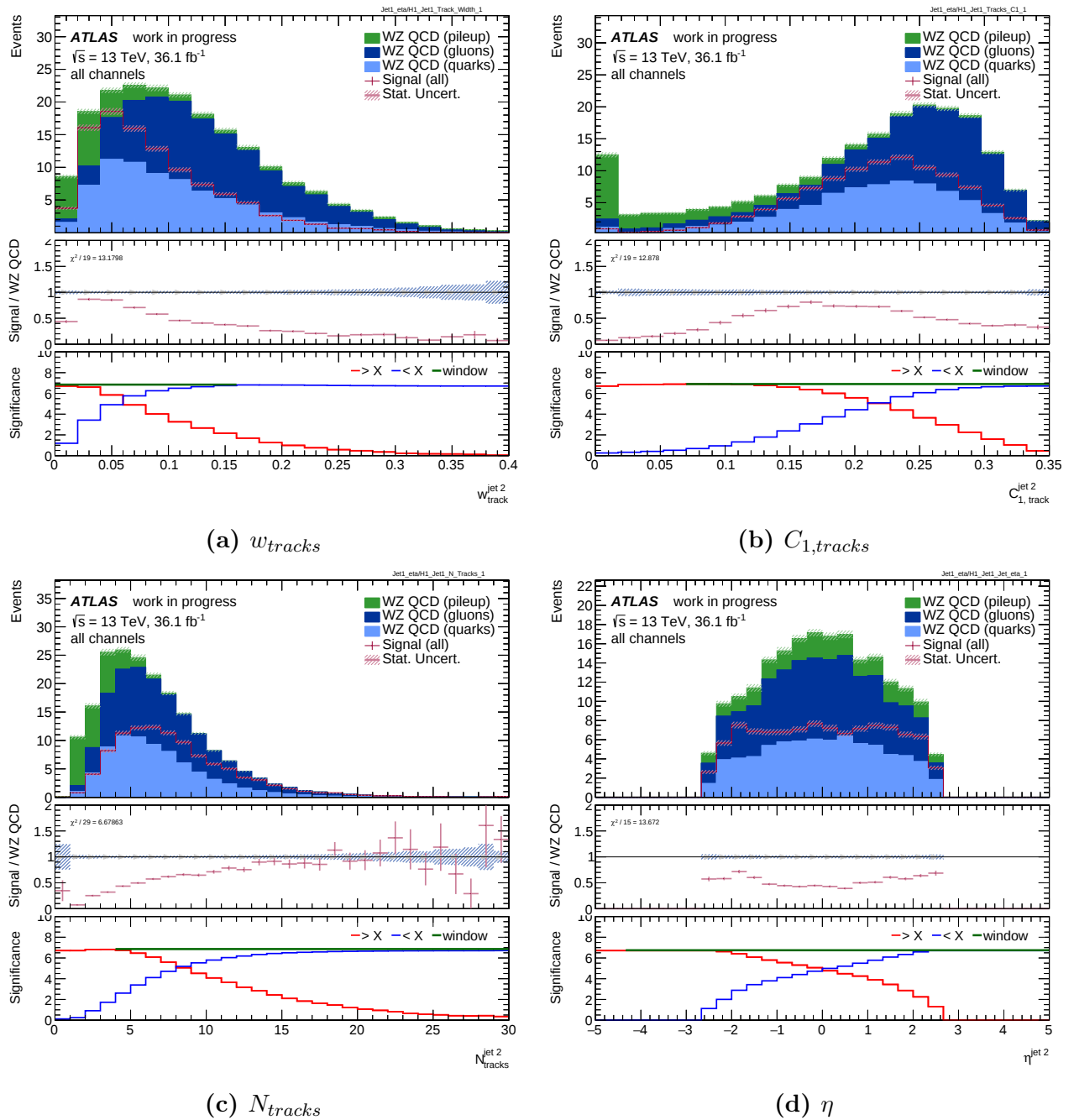


Abbildung 8.3: BDT Variablen des zweiten Jets ohne VBS Schnitte nach dem BDT Training Schnitt auf die zweiten Jets. Die Angaben in den Klammern hinter den Prozessen geben an, von welcher Jetsorte der zweite Jet sein muss. Die gefüllten Einträge werden aufaddiert. In den oberen Subplots ist das Verhältnis vom Signal- zu Untergrundereignissen dargestellt. Die unteren Subplots zeigen die Signifikanz bei unterschiedlichen Schnitten.

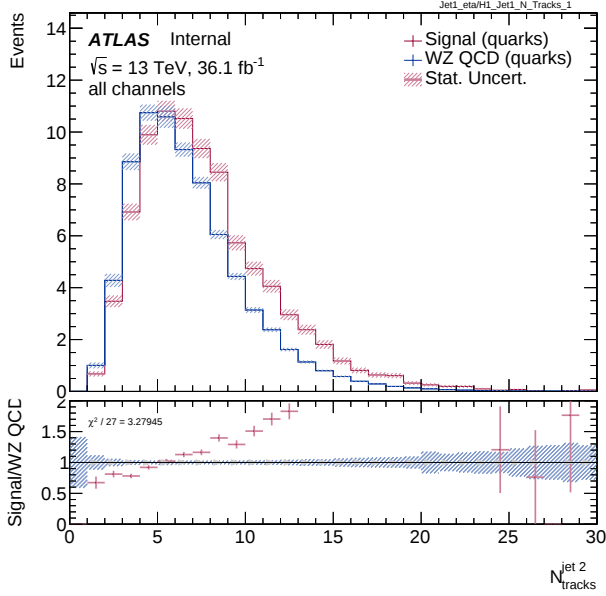
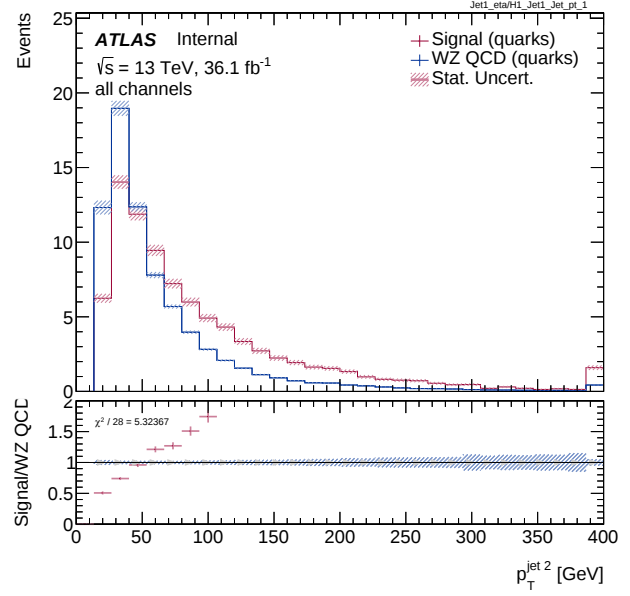
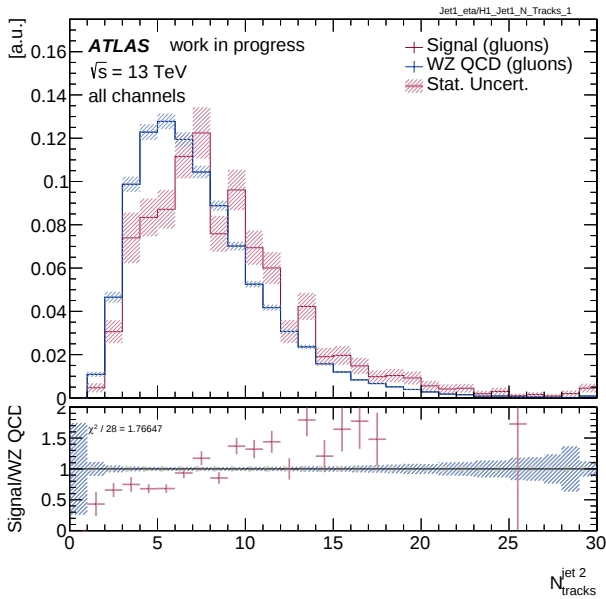
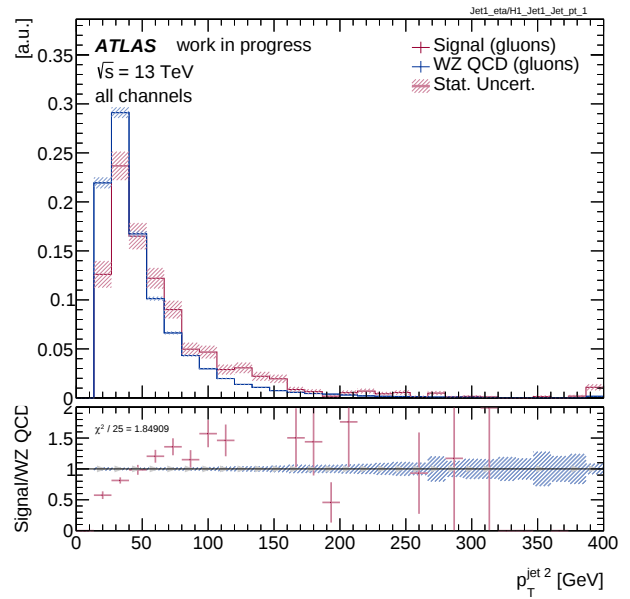
(a) N_{tracks} der Quark-Jets(b) p_T der Quark-Jets(c) N_{tracks} der Gluon-Jets(d) p_T der Gluon-Jets

Abbildung 8.4: Normierte Plots N_{tracks} - und p_T -Werte von den Quark- oder Gluon-Jets der zweiten Jets nach dem BDT Training Schnitt auf die zweiten Jets. Die Subplots zeigen die Signifikanz bei unterschiedlichen Schnitten.

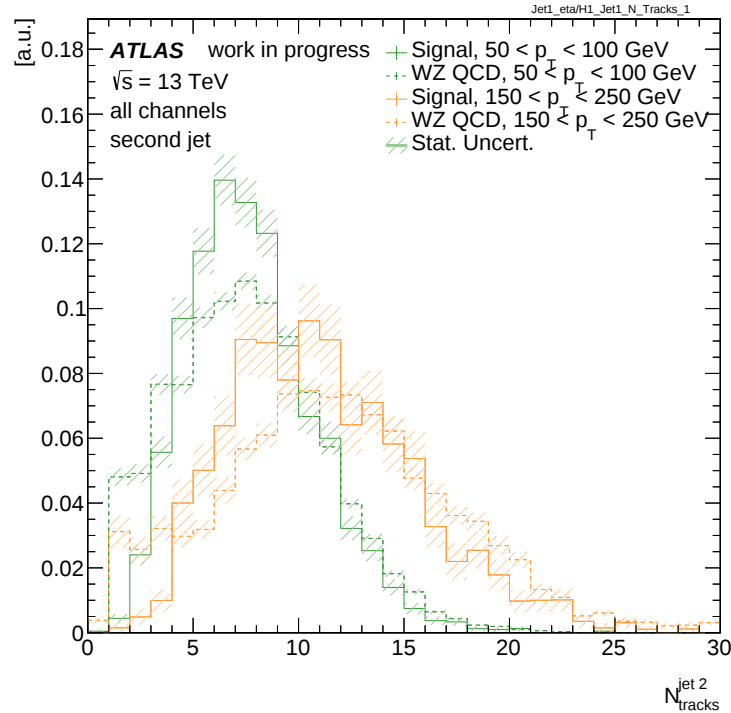


Abbildung 8.5: Normierter Plot der N_{tracks} im Signal und Untergrund der zweiten Jets nach dem BDT Training Schnitt auf die zweiten Jets für unterschiedliche p_T -Werte.

bei den zweiten Jets gefunden werden. Im Verhältnisplot kann gesehen werden, dass zwischen Signal und Untergrund nur geringe Unterschiede bestehen. Dies liegt zum Teil vermutlich wieder daran, dass die Gluon-Jets zu höheren und die Pileup-Jets zu niedrigeren BDT-Werten tendieren und sich damit in etwa ausgleichen. Wie somit zu erwarten tritt auch im Signifikanz-Subplot kein Maximum auf. Die Korrelationsplot zwischen den BDT-Werten der ersten und zweiten Jets scheinen wieder einen Blick auf die Summe der BDT-Werte zu rechtfertigen (Abbildung 8.7). Wie in Abbildung 8.8 zu sehen ist, kann jedoch auch hier kein Maximum der Signifikanz gefunden werden. Somit ist ein Schnitt auch hier nicht sinnvoll. Möglicherweise könnte ein Schnitt auf kleine Werte von $C_{1,track}$, der die Anzahl an Pileup-Jets reduziert, dazu führen, dass durch die BDT-Werte der Untergrund besser vom Signal unterschieden werden kann. Dazu sollte jedoch erst untersucht werden, ob das Pileup im Vergleich zu den Daten gut modelliert ist.

8.4 Anwendung von einzelnen VBS Schnitten

Im bisherigen Kapitel kann gesehen werden, dass durch das Weglassen der VBS Schnitte keine Verbesserungen in den Ergebnissen des BDT-Tools gefunden werden kann. Im Folgenden wird untersucht, ob durch das Anwenden einzelner dieser Schnitte eine Verbesserung gefunden werden kann. Dazu werden alle Kombinationen zwischen den vier Schnitten nach dem BDT Training Schnitt auf beide Jets angewendet und sich der BDT-Wert und die Summe der

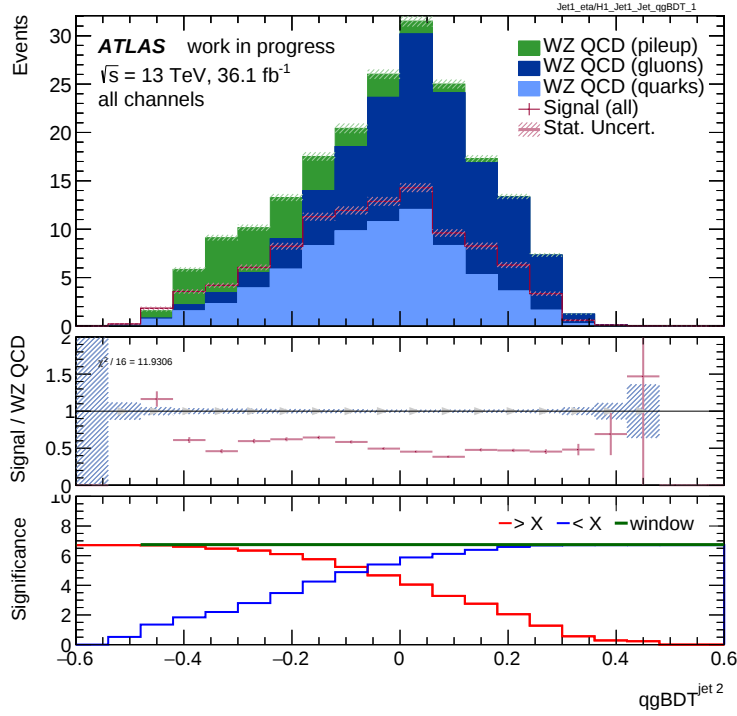
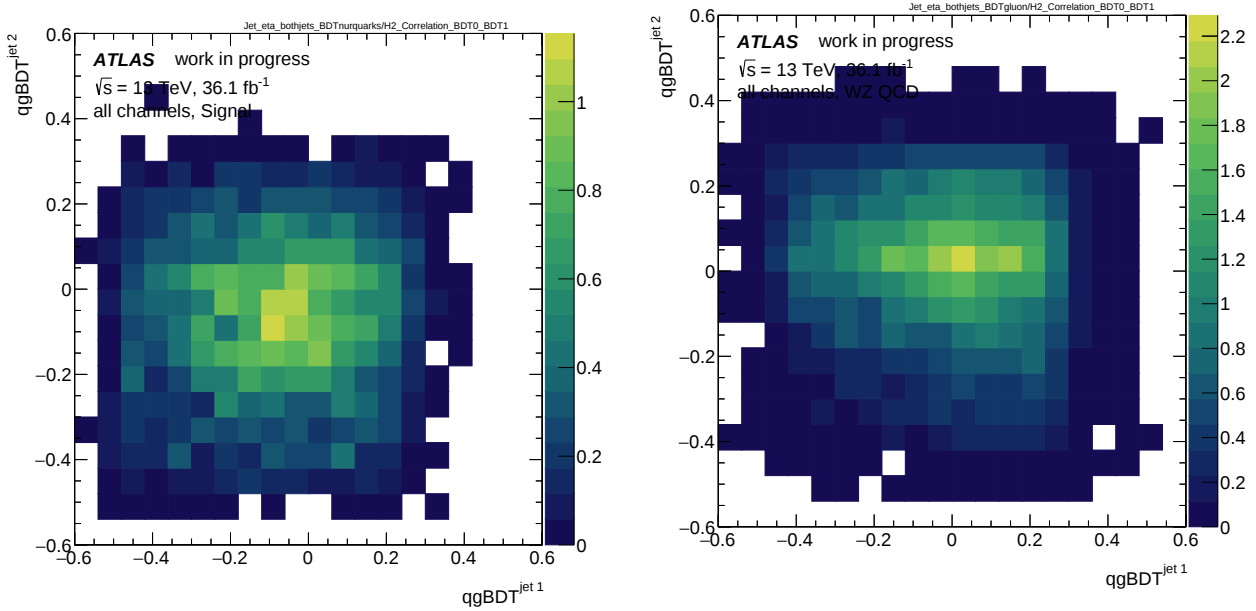


Abbildung 8.6: BDT-Werte der zweiten Jets nach dem BDT Training Schnitt auf die zweiten Jets. Die Angaben in den Klammern hinter den Prozessen geben dabei an, von welcher Jetsorte die zweiten Jets sein müssen. Die gefüllten Einträge werden aufaddiert.



(a) Korrelation des BDT-Werts für den ersten und zweiten Jet im Signal, wobei beide Jets ein Quark-Jet sein müssen.

(b) Korrelation des BDT-Wertes des ersten und zweiten Jets im Untergrund, wobei die Ereignisse, in denen beide Jets ein Quark-Jets sind, ausgeschlossen werden.

Abbildung 8.7: Korrelationsplots der BDT-Werte für den ersten und zweiten Jet nach dem BDT Training Schnitt auf beide Jets ohne die VBS Schnitte. Es werden nur Bins mit einer Ereignisanzahl größer als Null dargestellt.

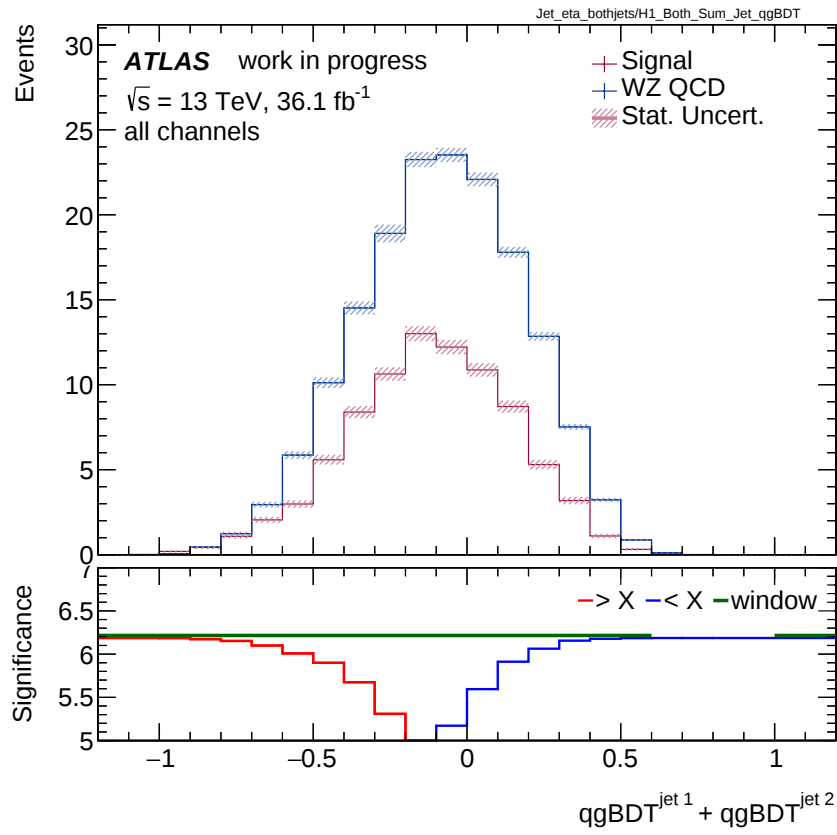


Abbildung 8.8: Summe der BDT-Werte nach dem BDT Training Schnitt auf beide Jets. Der Subplot zeigt die Signifikanz bei unterschiedlichen Schnitten.

BDT-Werte angeschaut. Bei all diesen Kombinationen kann kein Maximum der Signifikanz gefunden werden. Insofern können auch einzelne VBS-Schnitte nicht zur Verbesserung der Unterscheidungskraft beitragen. In Anhang G sind beispielhaft die Plots nach den einzelnen VBS-Schnitten dargestellt.

9 Zusammenfassung und Ausblick

Bei der Vektorbosonstreuung von zwei gleich geladenen leptonisch zerfallenden W-Bosonen kommt es in erster Ordnung zu Endzuständen mit zwei Quark-Jets und zwei gleich geladenen Fermionen im Detektor. Der größte Untergrund zu solchen Prozessen sind WZjj QCD-Prozesse die sich durch einen größeren Anteil an Gluon-Jets auszeichnen. Ein BDT-Tool zur Unterscheidung von Quark- und Gluon-Jets wird somit zur Unterscheidung von Signal und Untergrund untersucht.

Mit der Betrachtung des Potentials eines perfekten Quark-Gluon-Taggers wird die prinzipiell mögliche Verbesserung der Signifikanz durch einen Schnitt auf die Gluon-Jets mithilfe der Wahrheitswerte der Jet-initiiierenden Teilchen untersucht. Dazu werden verschiedene Bedingungen an die Wahrheitswerte ausprobiert. Im Fall der perfekten Separation in allen η -Bereichen kann eine maximale Steigerung der Signifikanz bei einem Schnitt auf alle Ereignisse mit mindestens einen Gluon Jet erreicht werden. Damit kann die Signifikanz um $1,3\sigma$ gesteigert werden. Bei Betrachtung eines perfekten Quark-Gluon-Taggers, der die Gluon-Jets nur im Zentralbereich separieren kann und der somit die Beschränkungen des inneren Detektors berücksichtigt, kann die Signifikanz maximal um $0,89\sigma$ gesteigert werden. Dies wird für einen Schnitt auf alle Ereignisse, in denen mindestens einer der Jets ein Gluon-Jets im Zentralbereich ist, erreicht.

Die Anwendung des BDT Tools zeigt zwar, dass die Untergrundereignisse dazu tendieren, größere BDT-Werte als die Signalereignisse anzunehmen, diese Unterschiede aber nicht zu einer Steigerung der Signifikanz durch einen Schnitt auf die BDT-Werte der jeweiligen Jets führen. Auch ein Schnitt auf die Summe der BDT-Werte kann die Signifikanz nicht steigern. Möglicherweise könnte durch eine kleine Verbesserung des Tools ein solcher Schnitt zu einer signifikanten Unterscheidung zwischen Signal und Untergrund führen.

Bei der BDT-Variable N_{tracks} kann eine Verbesserung der Unterscheidung zwischen Quark- und Gluon-Jets für höhere p_T -Werte gefunden werden. Die Untersuchung des BDT-Tools zeigt jedoch keine Verbesserung des BDT-Tools durch p_T abhängige Schnitte. Dies liegt wahrscheinlich daran, dass die p_T -Werte der Jets überwiegend kleiner als 250 GeV und somit relativ klein sind.

Es kann keine Abhängigkeit der BDT-Variablen und des BDT-Wertes von der durchschnittlichen Anzahl an Interaktionen pro Kreuzung der Teilchenbündel und somit vom Pileup gefunden werden.

Die Untersuchung des Tools in einem erweiterten Phasenraum, in dem die VBS Schnitte nicht angewendet werden, zeigt ebenfalls keine Verbesserung der Signifikanz. In dieser Betrachtung kommen viele Pileup-Jets im Untergrund hinzu, die entgegengesetzte Tendenzen zu den Gluon-Jets zeigen und somit die Unterscheidung zwischen Signal und Untergrund erschweren. Ein Schnitt auf kleine Werte von $C_{1,tracks}$ könnte nützlich sein, um das Pileup zu reduzieren. Um zu sehen, ob ein solcher Schnitt sinnvoll wäre, muss zunächst noch geprüft werden, ob das

Pileup im Vergleich zu den Daten gut modelliert ist.

Für die Anwendung einzelner VBS Schnitte kann ebenfalls kein Schnitt auf die BDT-Werte gefunden werden, der die Signifikanz erhöht.

Für weitere Studien wäre es interessant zu sehen, ob die Nutzung von Kalorimeterinformationen erfolgreich ist, um auch Jets im Vorwärtsbereich zu berücksichtigen und die Unterscheidung im Zentralbereich zu verbessern. Bisher wurde der Effekt des Quark-Gluon-Tagger nur auf den WZ QCD-Untergrund untersucht. Die Untersuchung von anderen Untergründen könnte ebenfalls interessant sein.

A Verwendete Monte-Carlo Dateien

Die Auswertung wurde mit Software basierend auf ATLAS Software Release 21 durchgeführt. Die verwendeten Monte-Carlo-Dateien sind:

- mc16_13TeV.364253.Sherpa_222_NNPDF30NNLO_lllv.deriv.DAOD_STDM3.e5916_s3126_r9364_p3897
- mc16_13TeV.364287.Sherpa_222_NNPDF30NNLO_llvvjj_ss_EW6.deriv.DAOD_STDM3.e6055_s3126_r9364_p3897

B Vergleich von Quark-Jets im Signal und Untergrund

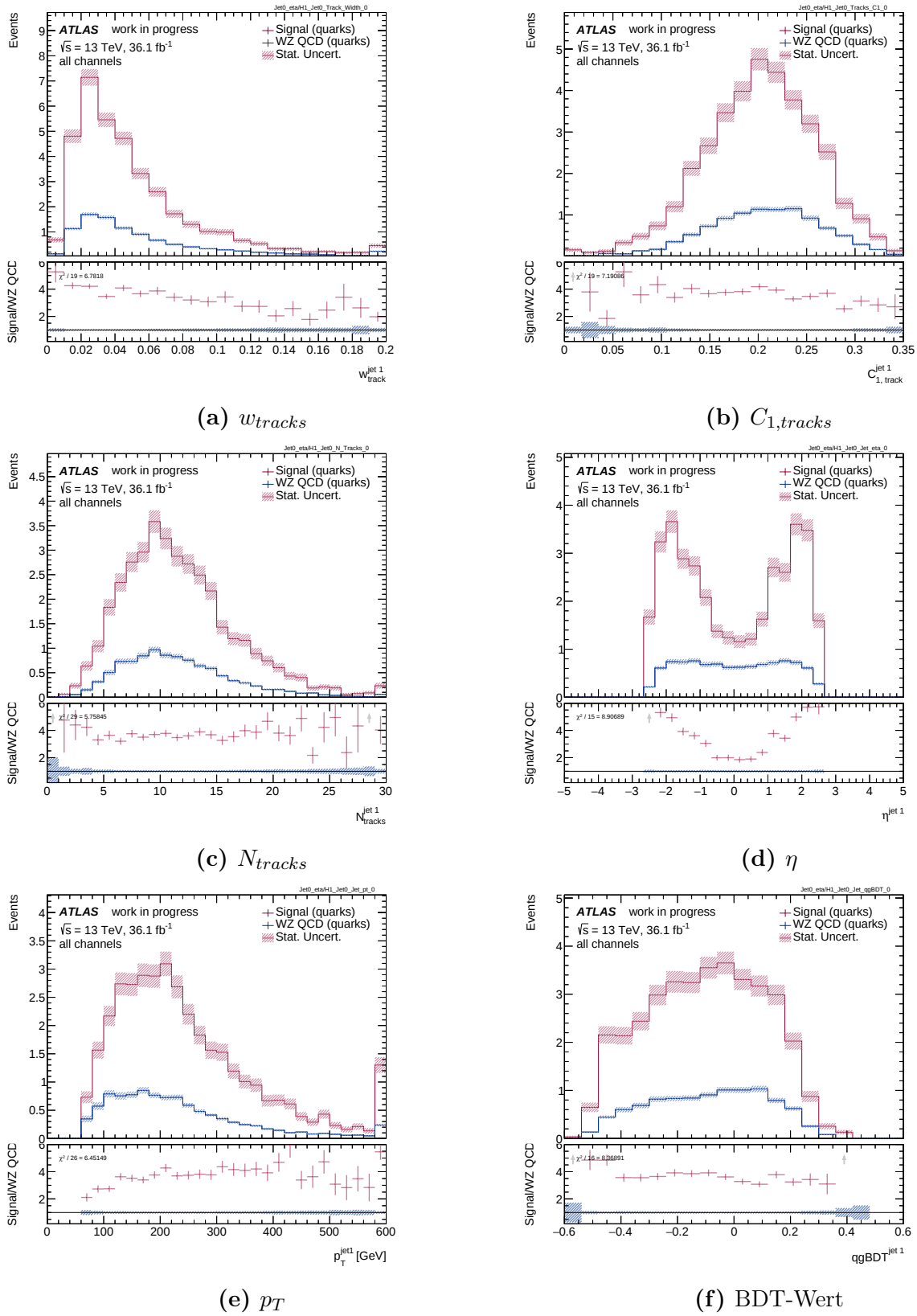
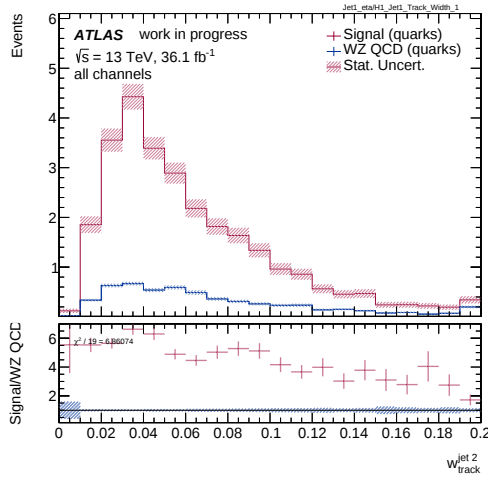
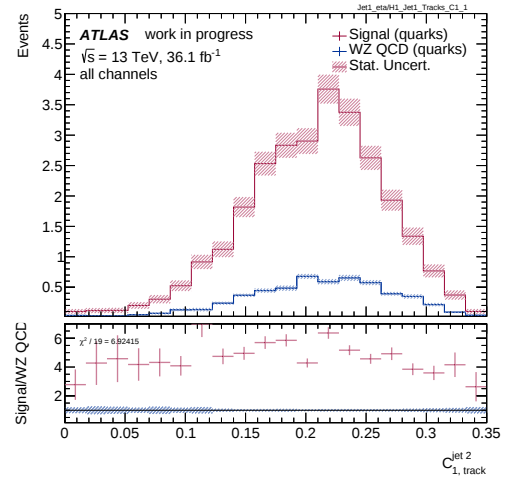
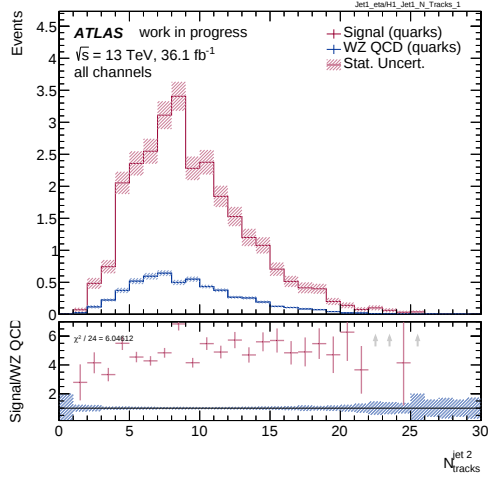
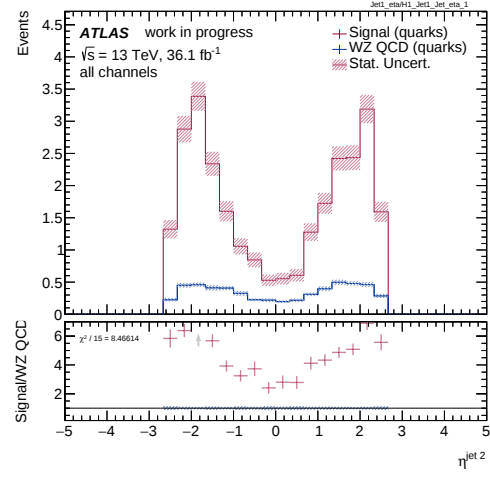
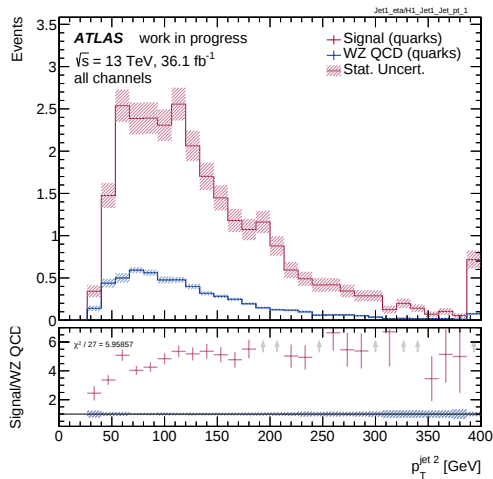
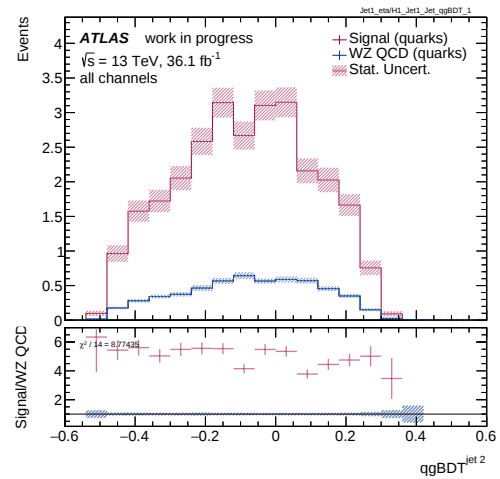


Abbildung B.1: Quark-Jets der ersten Jets im Signal und Untergrund nach dem BDT Training Schnitt auf die ersten Jets

(a) w_{tracks} (b) $C_{1,tracks}$ (c) N_{tracks} (d) η (e) p_T 

(f) BDT-Wert

Abbildung B.2: Quark-Jets der zweiten Jets im Signal und Untergrund nach dem BDT Training Schnitt auf die zweiten Jets

C Vergleich von Quark-Jets mit großen zu Quark-Jets mit kleinen p_T -Werten

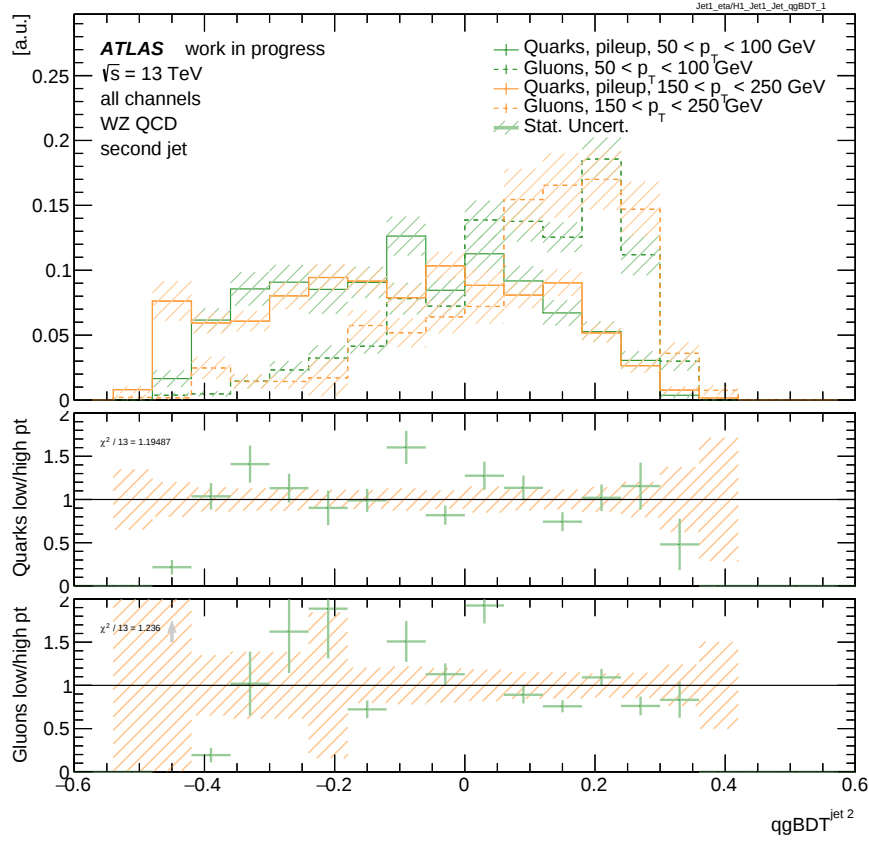


Abbildung C.1: Normierte Verteilung der BDT-Werte der zweiten Jets in unterschiedlichen p_T -Regionen nach dem BDT Training Schnitt auf die zweiten Jets. Die Bezeichnung Quarks, Pileup und Gluon steht dabei für Art der zweiten Jets. Der obere Subplot zeigt das Verhältnis von Quark-Jets mit kleinen zu Quark-Jets mit großen p_T -Werten und der untere das gleiche Verhältnis für Gluon-Jets.

D Signifikanz der BDT-Werte für kleine und große p_T -Werte

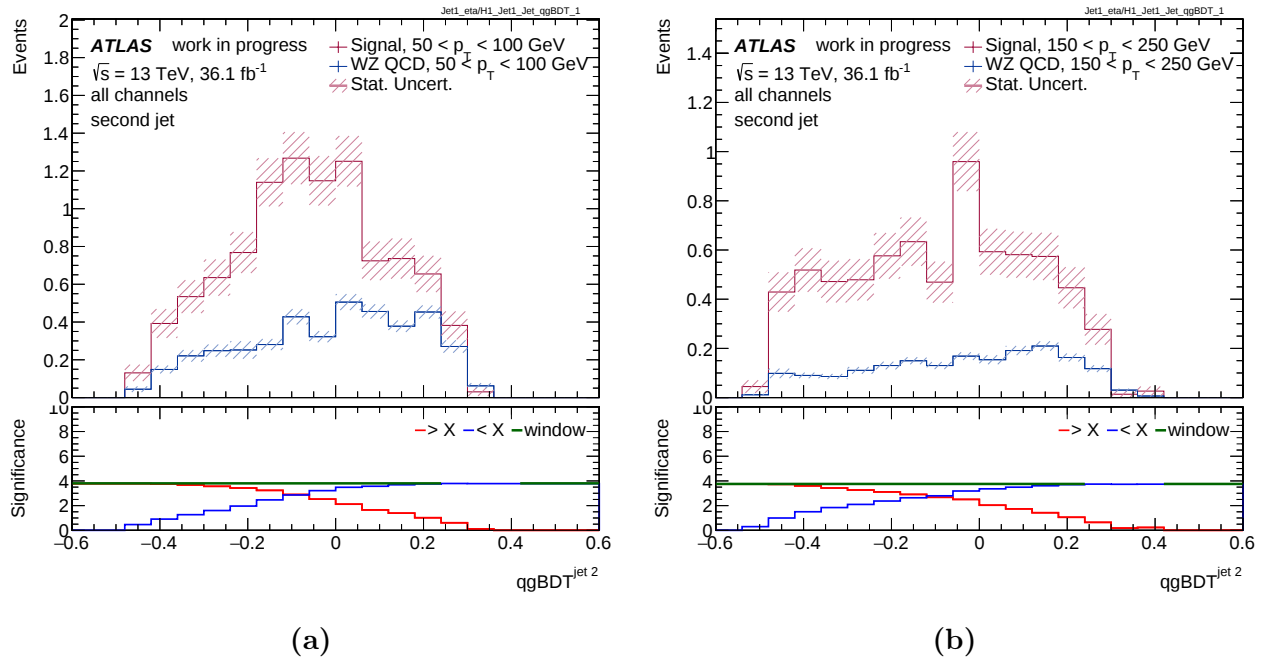


Abbildung D.1: Verteilung der BDT-Werte der zweiten Jets für (a) kleine und (b) große p_T -Werte nach dem BDT-Training Schnitt auf die zweiten Jets. Der Subplot zeigt die Signifikanz bei unterschiedlichen Schnitten.

E BDT Variablen für unterschiedliche avIntPerX Regionen

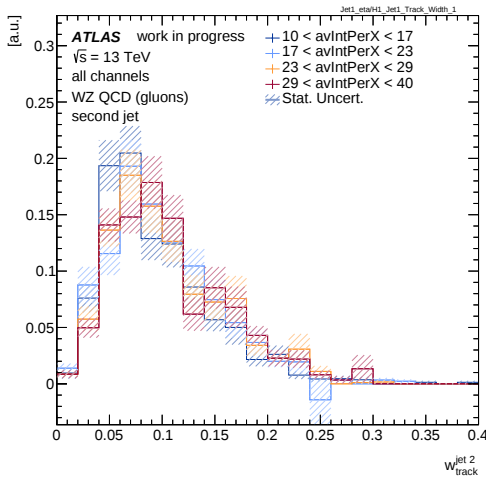
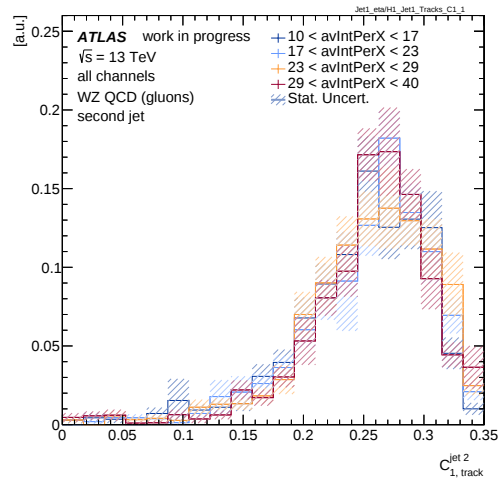
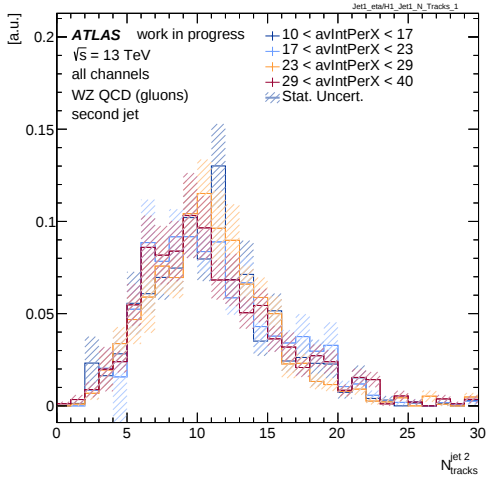
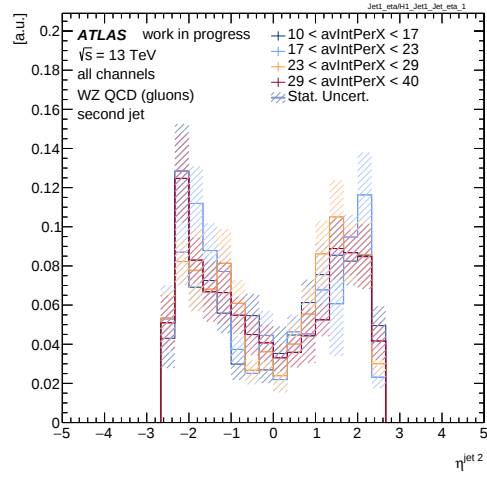
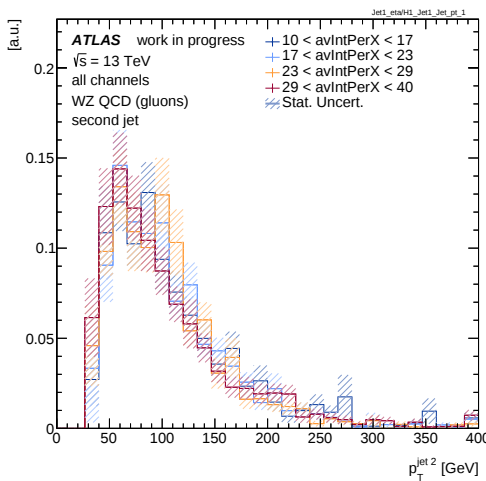
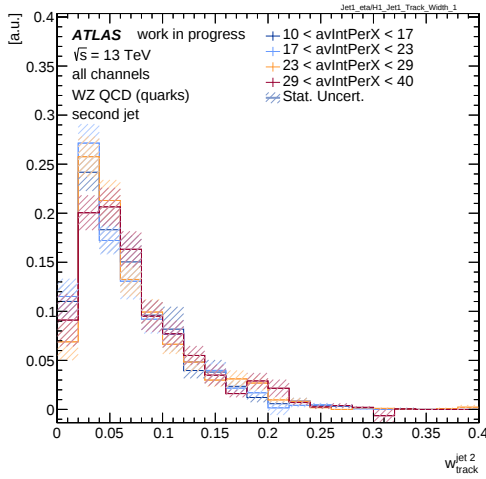
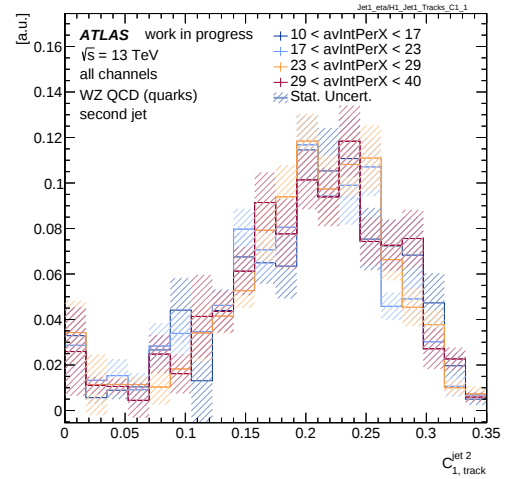
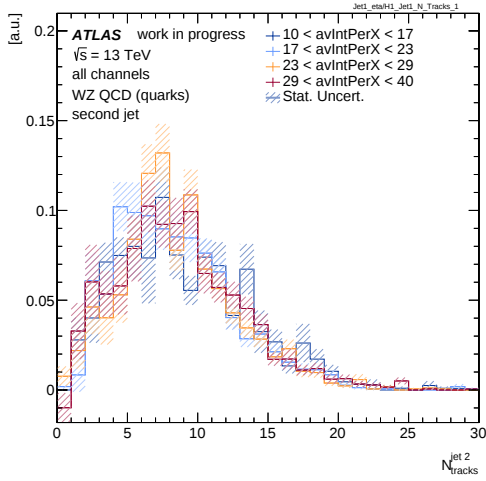
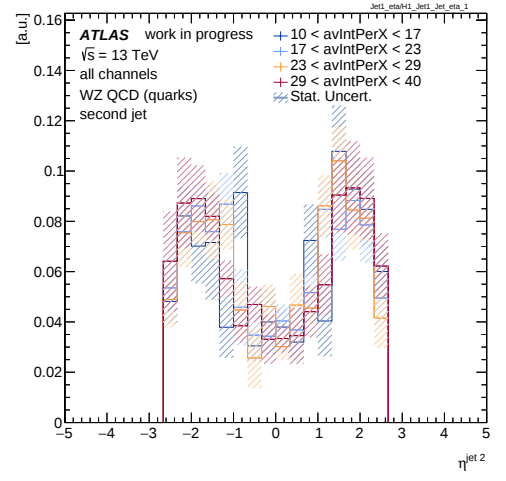
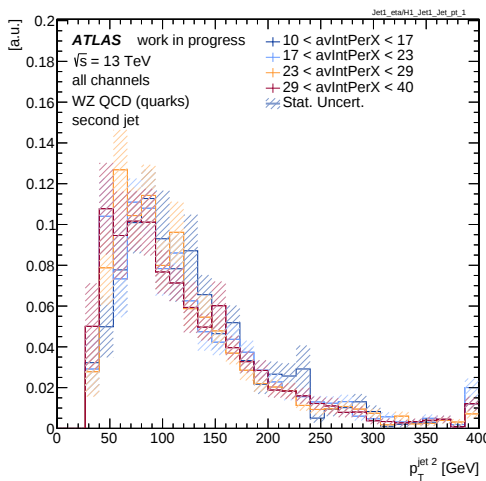
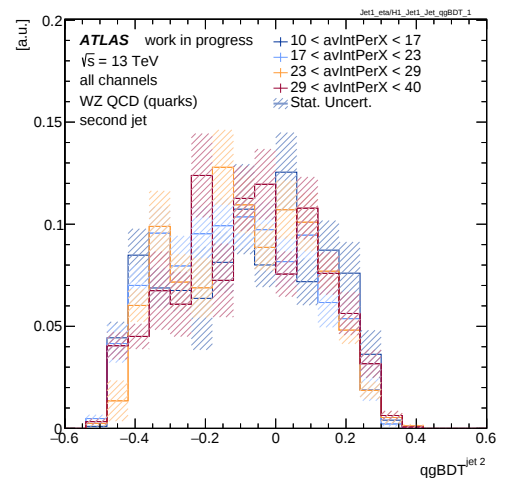
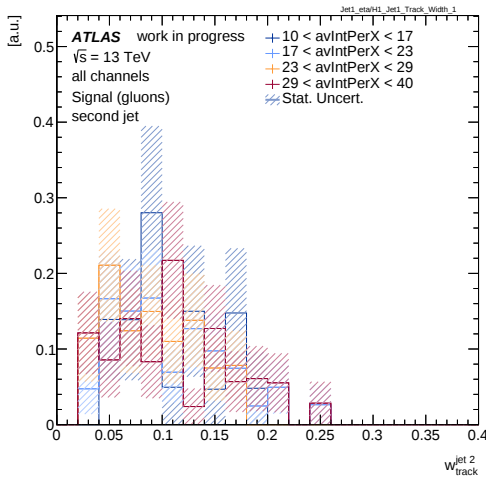
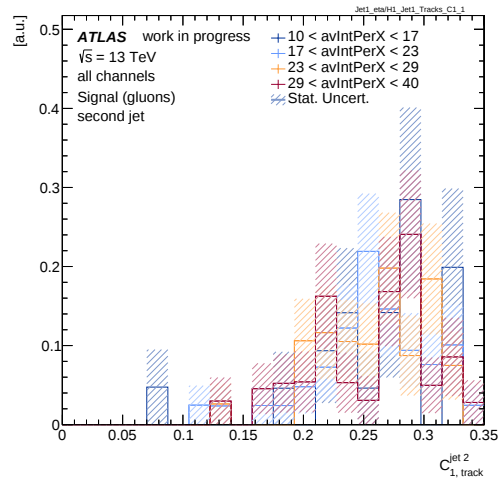
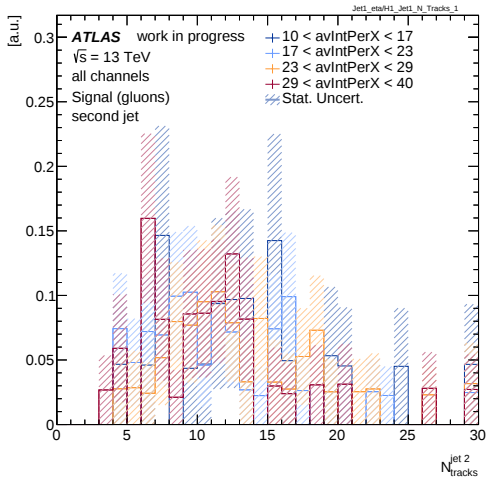
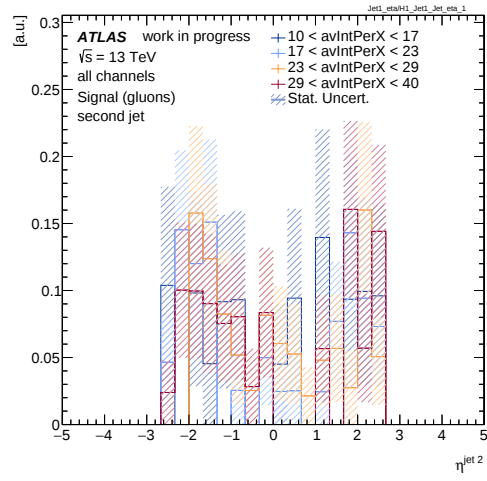
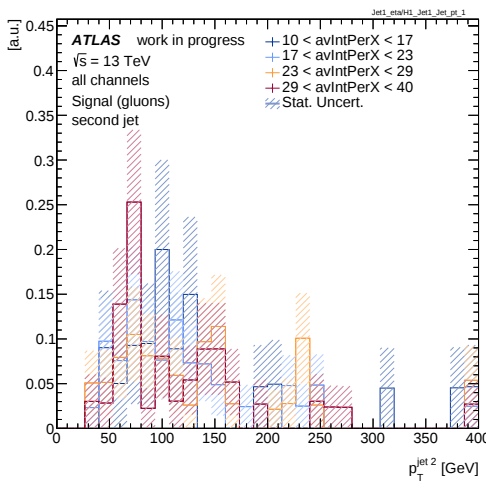
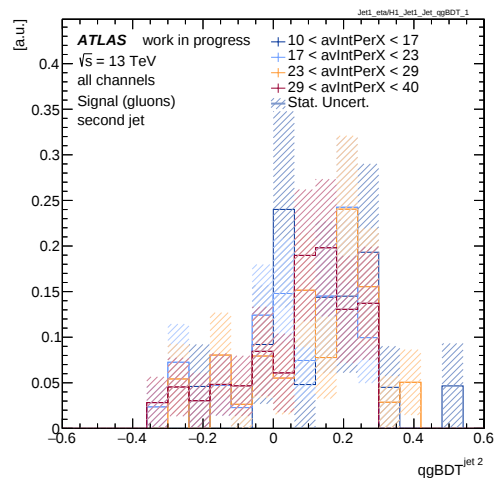
(a) w_{tracks} (b) $C_{1,tracks}$ (c) N_{tracks} (d) η (e) p_T

Abbildung E.1: Gluon-Jets der zweiten Jets im Untergrund nach dem BDT Training
Schnitt auf die zweiten Jets für unterschiedliche avIntPerX-Bereiche

(a) w_{tracks} (b) $C_{1,\text{tracks}}$ (c) N_{tracks} (d) η (e) p_T 

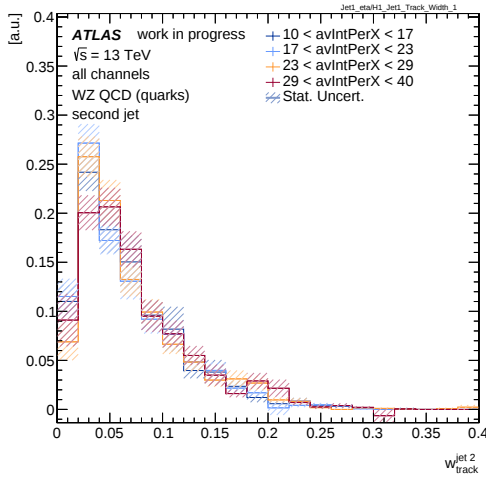
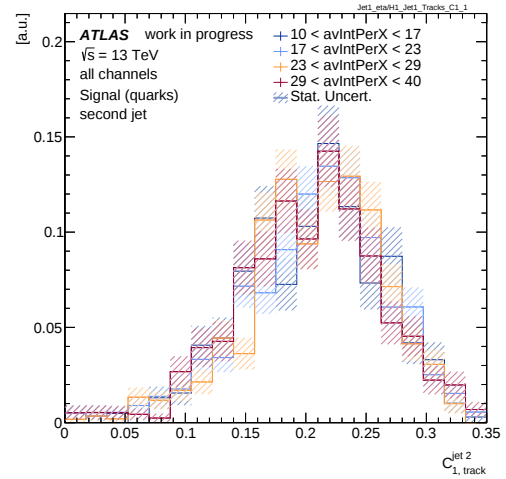
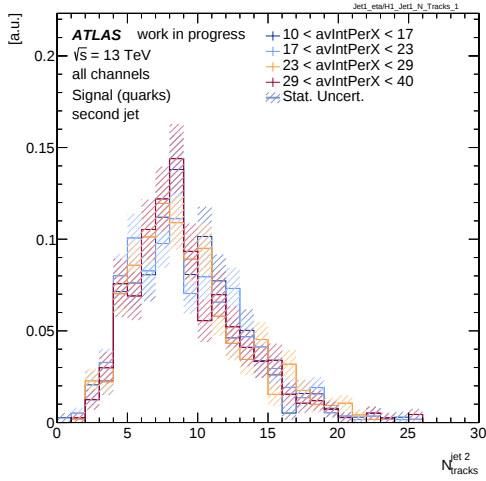
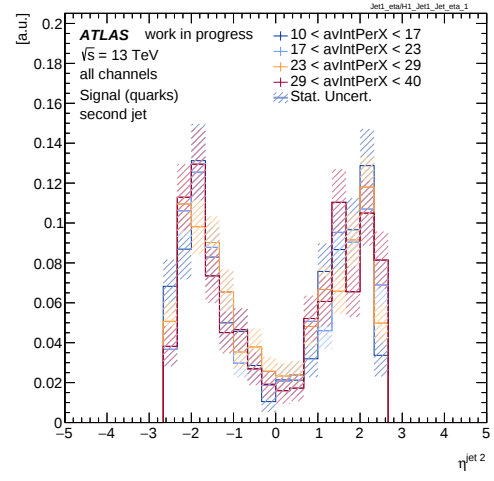
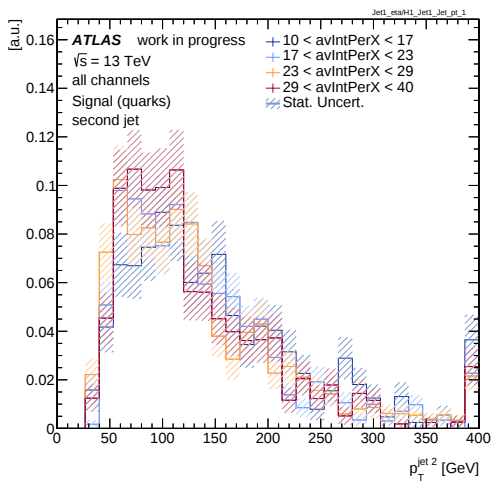
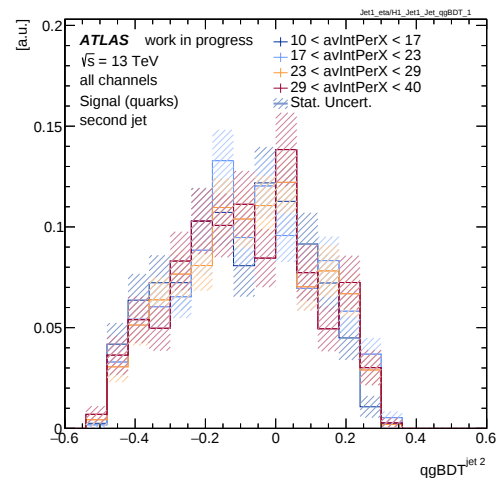
(f) BDT-Wert

Abbildung E.2: Quark-Jets der zweiten Jets im Untergrund nach dem BDT Training
 Schnitt auf die zweiten Jets für unterschiedliche avIntPerX-Bereiche

(a) w_{tracks} (b) $C_{1,tracks}$ (c) N_{tracks} (d) η (e) p_T 

(f) BDT-Wert

Abbildung E.3: Gluon-Jets der zweiten Jets im Signal nach dem BDT Training Schnitt auf die zweiten Jets für unterschiedliche avIntPerX-Bereiche

(a) w_{tracks} (b) $C_{1,tracks}$ (c) N_{tracks} (d) η (e) p_T 

(f) BDT-Wert

Abbildung E.4: Quark-Jets der zweiten Jets im Signal nach dem BDT Training Schnitt auf die zweiten Jets für unterschiedliche avIntPerX-Bereiche

F BDT-Variablen bzw. BDT-Wert der ersten Jets ohne VBS Schnitte

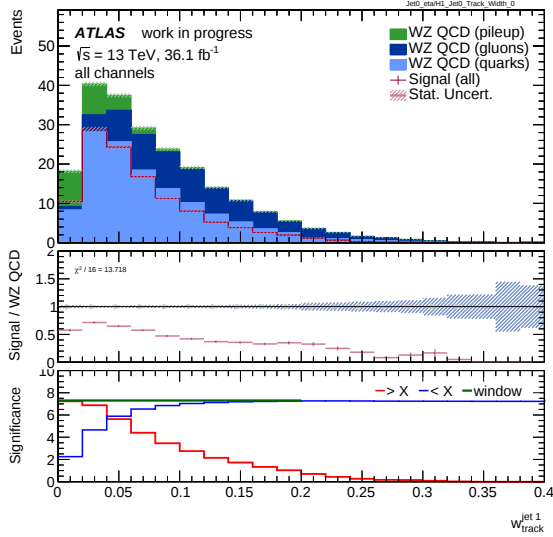
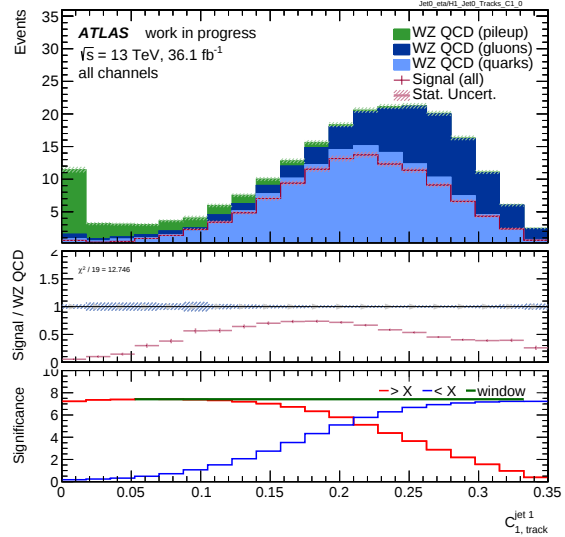
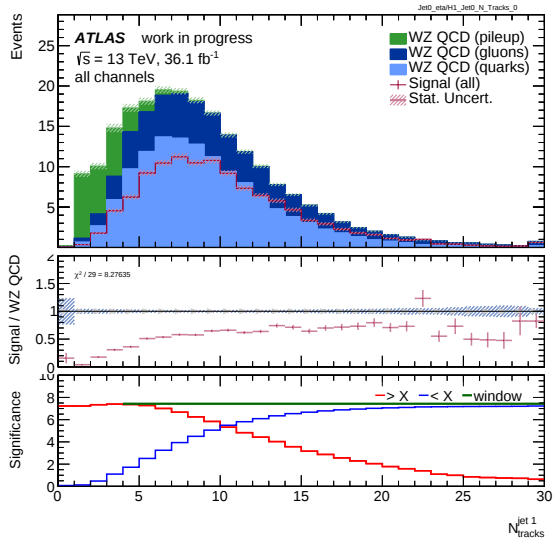
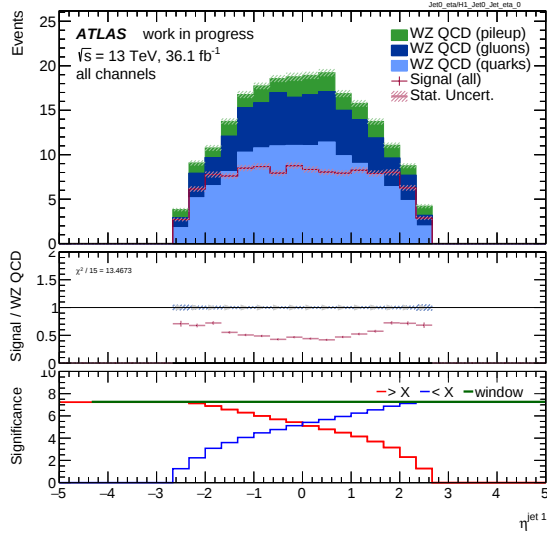
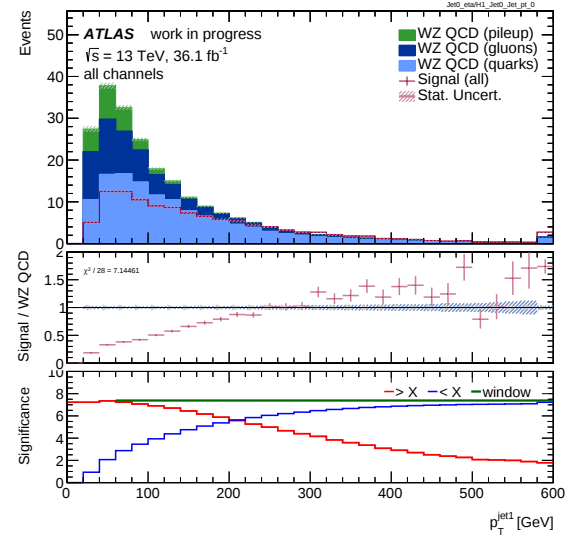
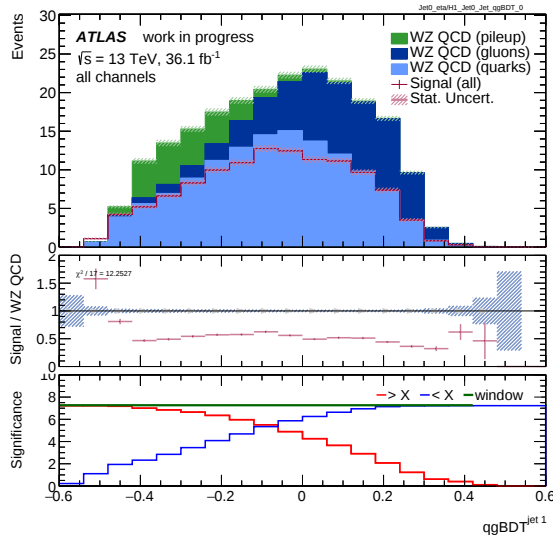
(a) w_{tracks} (b) $C_{1,tracks}$ (c) N_{tracks}

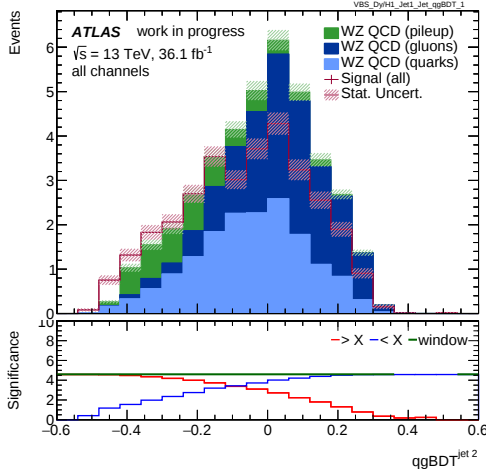
Abbildung F.1: BDT-Variablen der ersten Jets nach dem BDT Training Schnitt auf die ersten Jets ohne die VBS Schnitte. Die Angaben in den Klammern hinter den Prozessen geben an, von welcher Jetsorte der erste Jet sein muss. Die gefüllten Einträge werden aufaddiert. In den oberen Subplots ist das Verhältnis von Signal zum Untergrund dargestellt. Die unteren Subplots zeigen die Signifikanz bei unterschiedlichen Schnitten.

(a) η (b) p_T 

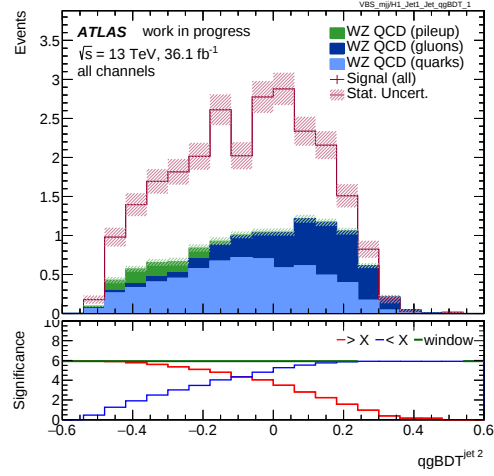
(c) BDT-Wert

Abbildung F.2: BDT-Variablen und BDT-Wert der ersten Jets nach dem BDT Training Schnitt auf die ersten Jets ohne die VBS Schnitte. Die Angaben in den Klammern hinter den Prozessen geben an, von welcher Jetsorte der erste Jet sein muss. Die gefüllten Einträge werden aufaddiert. In den oberen Subplots ist das Verhältnis von Signal zum Untergrund dargestellt. Die unteren Subplots zeigen die Signifikanz bei unterschiedlichen Schnitten.

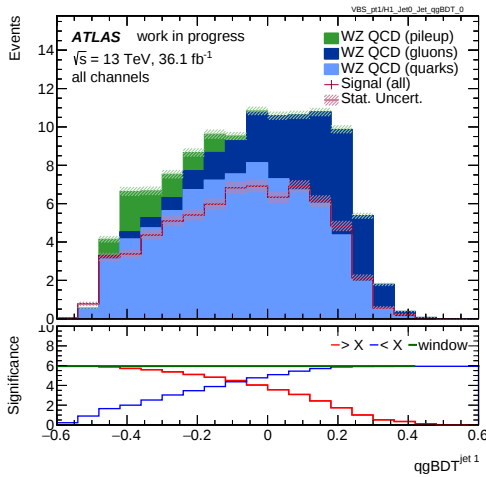
G BDT-Wert nach einzelnen VBS Schnitten



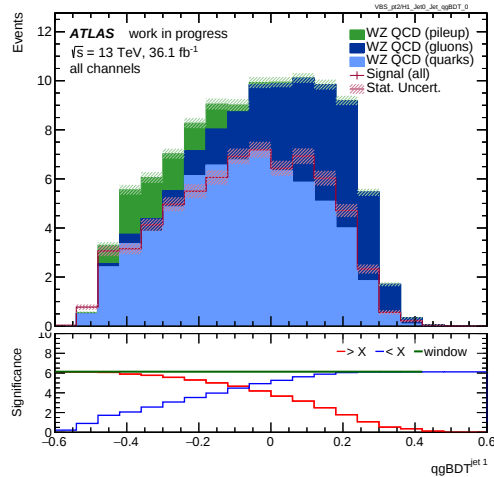
(a) BDT-Werte der zweiten Jets nach dem VBS Schnitt auf ΔY der Tagging-Jets.



(b) BDT-Werte der zweiten Jets nach dem VBS Schnitt auf m_{jj} der Tagging-Jets



(c) BDT-Werte der ersten Jets nach dem VBS Schnitt auf die p_T -Werte der ersten Jets.



(d) BDT-Werte der zweiten Jets nach dem VBS Schnitt auf die p_T -Werte der zweiten Jets.

Abbildung G.1: BDT-Werte nach einzelnen VBS Schnitten und nach dem BDT Training Schnitt auf beide Jets. Die Angaben in den Klammern hinter den Prozessen geben an, von welcher Jetsorte der erste bzw. zweite Jet sein muss. Die gefüllten Einträge werden aufaddiert. Die Subplots zeigen die Signifikanz bei unterschiedlichen Schnitten.

Literatur

- [1] Aaboud et al. Observation of electroweak production of a same-sign W boson pair in association with two jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. 2019.
- [2] Aad et al. Evidence for Electroweak Production of $W^\pm W^\pm jj$ in pp Collisions at $\sqrt{s} = 8$ TeV with the ATLAS Detector. *Phys. Rev. Lett.*, 113:141803, Oct 2014.
- [3] Khachatryan et al. Study of Vector Boson Scattering and Search for New Physics in Events with Two Same-Sign Leptons and Two Jets. *Phys. Rev. Lett.*, 114:051801, Feb 2015.
- [4] Sirunyan et al. Observation of Electroweak Production of Same-Sign W Boson Pairs in the Two Jet and Two Same-Sign Lepton Final State in Proton-Proton Collisions at $\sqrt{s} = 13$ TeV. *Phys. Rev. Lett.*, 120:081801, Feb 2018.
- [5] Aad et al. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008.
- [6] Aad et al. Light-quark and gluon jet discrimination in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *The European Physical Journal C*, 74(8), Aug 2014.
- [7] The ATLAS collaboration. Quark versus Gluon Jet Tagging Using Charged Particle Multiplicity with the ATLAS Detector. Technical Report ATL-PHYS-PUB-2017-009, CERN, Geneva, May 2017.
- [8] Roe et al. Boosted decision trees, a powerful event classifier. In *Statistical Problems in Particle Physics, Astrophysics and Cosmology (PHYSTAT 05): Proceedings, Oxford, UK, September 12-15, 2005*, pages 139–142, 2005.
- [9] The ATLAS collaboration. Discrimination of Light Quark and Gluon Jets in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS Detector. Technical Report ATLAS-CONF-2016-034, CERN, Geneva, Jul 2016.
- [10] Buttinger et al. Formulae for Estimating Significance. Technical Report ATL-COM-GEN-2018-026, CERN, Geneva, Oct 2018.
- [11] Casadei. How to measure efficiency. Technical Report arXiv:0908.0130, Aug 2009.
- [12] Bittrich et al. Support note for measurement of electroweak $W^\pm W^\pm jj$ production at $\sqrt{s} = 13$ TeV. Technical Report ATL-COM-PHYS-2018-252, CERN, Geneva, Mar 2018.
- [13] Cacciari et al. The anti-ktjet clustering algorithm. *Journal of High Energy Physics*, 2008(04):063–063, Apr 2008.

- [14] The ATLAS Collaboration. Tagging and suppression of pileup jets with the ATLAS detector. Technical Report ATLAS-CONF-2014-018, CERN, Geneva, May 2014.

Erklärung

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

Anna Sophie Tegetmeier
Dresden, Oktober 2019