

**TECHNISCHE
UNIVERSITÄT
DRESDEN**

Identifikation hadronischer Tau-Leptonzerfälle mit neuronalen Netzen und Substrukturvariablen bei ATLAS

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

vorgelegt von

Anne Kristin Dietert
geboren am 09.12.1993 in Berlin

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

2016

Eingereicht am 04. März 2016

1. Gutachter: Prof. Dr. Arno Straessner
2. Gutachter: Prof. Dr. Kai Zuber

Zusammenfassung

Zusammenfassung

Deutsch

In dieser Arbeit werden hadronische Tau-Zerfälle am ATLAS-Detektor untersucht. Die Identifizierung dieser spielt in der Forschung zum Standardmodell und dem Higgs-Boson eine entscheidende Rolle. Für die Identifikation hadronisch zerfallender Tau-Leptonen werden multivariate Methoden verwendet. Die derzeit übliche Methode sind Boosted Decision Trees (BDT). Im Zuge dieser Arbeit sollen die kürzlich als Alternative entwickelten Künstlichen Neuronalen Netze verbessert werden. Hierfür werden die Standardvariablen, die Variablen, die normalerweise zur Tau-Identifizierung verwendet werden, durch Hinzufügen von Substrukturvariablen erweitert beziehungsweise ersetzt. Dies soll eine Verbesserung neuronaler Netze gegenüber dem BDT erbringen.

Abstract

English:

This thesis will examine hadronic tau decays at the ATLAS detector. Their identification of them plays an important role in the research of the Standard Model and the Higgs boson. For the identification of hadronic tau decays, multivariate methods will be used. Currently, boosted decision trees (BDT) are used for that purpose. In this thesis artificial neural networks are studied as an alternative. Therefore the standard variables, which are used for the tau identification, will be extended respectively replaced by substructure variables. This shall bring an improvement of neural networks towards BDTs.

Inhaltsverzeichnis

1	Einleitung	1
2	Theoretische Grundlagen	3
2.1	Entstehung von Tau-Leptonen	3
2.2	Standardvariablen	4
2.3	Neuronale Netze	6
2.4	LHC und ATLAS	9
3	Auswertung	12
3.1	Substrukturvariablen	12
3.2	Bestätigung vorangegangener Ergebnisse	14
3.3	Untersuchung des Einflusses der Standardvariablen	16
3.4	Hinzufügen von Substrukturvariablen bei Entfernen von Standardvariablen	17
3.5	Vollständig ersetzte Variablenkonfigurationen	19
3.6	Hinzufügen von Substrukturvariablen zu den Standardvariablen	24
3.7	Einfluss diskreter Variablen	26
3.8	Stabilitätsuntersuchung	27
4	Zusammenfassung und Ausblick	29
A	Verteilungen der Variablen	30
A.1	Standardvariablen	30
A.2	Substrukturvariablen	33
B	Konfigurationen	45
C	Literaturhinweise	46
	Abbildungsverzeichnis	49
	Tabellenverzeichnis	51

1 Einleitung

In der Teilchenphysik wird der Frage nach den Bausteinen unserer Natur und unseres Universums nachgegangen. Dabei werden nicht nur die Teilchen der Materie, sondern auch die Kräfte, die auf diese wirken, untersucht.

Die grundlegenden Bausteine der Materie sind im Standardmodell der Teilchenphysik zusammengefasst, kurz auch Elementarteilchen genannt. Abbildung 1 zeigt die Quarks, Leptonen und Bosonen des Standardmodells.

Aus Quarks sind Hadronen zusammengesetzt wie beispielsweise das Proton oder das Neutron. Leptonen sind in geladene und ungeladene Leptonen eingeteilt, wobei die ungeladenen Neutrinos jeweils einem geladenen Lepton zuzuordnen sind. Die Eichbosonen werden auch Austauschteilchen genannt, da sie die Wechselwirkungen zwischen den anderen Teilchen ermöglichen. Das Photon ist das Austauschteilchen der elektromagnetischen Wechselwirkung, die Gluonen die der starken Wechselwirkung. Die W- und das Z-Boson sind die Austauschteilchen der schwachen Wechselwirkung. Hadronen sind Teilchen, die an der starken Wechselwirkung teilnehmen [2]. Das Higgs-Boson ist das derzeit schwerste Teilchen im Standardmodell.

Am LHC (Large Hadron Collider) [13] werden verschiedene Experimente durchgeführt. Eines der Ziele des ATLAS-Experiments [11] am LHC ist die Suche nach dem Higgs-Boson des Standardmodells und nach neuer Physik. Das Vorhandensein des Higgs-Bosons wurde 1964 von Brout, Englert und Higgs [16] vorhergesagt und 2012 am oben genannten Experiment nachgewiesen [17][18]. Dieses zerfällt unter anderem in zwei Tau-Leptonen [3]. Der Zerfall ist wegen der hohen Masse der Taus ($m = 1.777 \text{ GeV}$) [3] bevorzugt gegenüber anderen leptonischen Zerfällen. Sie sind Fermionen, also Spin-1/2-Teilchen und haben eine elektrische Ladung von ± 1 . Tau-Leptonen zerfallen wiederum zum einen hadronisch, zum anderen leptonisch (siehe Kapitel 2.1). Ihre mittlere Lebensdauer beträgt $c\tau = 87 \mu\text{m}$ [3]. Die Zerfallsprodukte der Tau-

Drei Generationen der Materie (Fermionen)					
	I	II	III	Eichbosonen	
Masse	2,3 MeV	1,275 GeV	173,07 GeV	0	125,9 GeV
Ladung	$\frac{2}{3}$	$\frac{2}{3}$	$\frac{2}{3}$	0	0
Spin	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	0
Name	u up	c charm	t top	γ Photon	H Higgs Boson
Quarks	4,8 MeV	95 MeV	4,18 GeV	0	
	$-\frac{1}{3}$	$-\frac{1}{3}$	$-\frac{1}{3}$	0	
	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	
	d down	s strange	b bottom	g Gluon	
Leptonen	<2 eV	<0,19 MeV	<18,2 MeV	0	91,2 GeV
	0	0	0	0	0
	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	1
	ν_e Elektron-Neutrino	ν_μ Myon-Neutrino	ν_τ Tau-Neutrino	Z^0 Z Boson	
	0,511 MeV	105,7 MeV	1,777 GeV	80,4 GeV	
	-1	-1	-1	± 1	
	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	
	e Elektron	μ Myon	τ Tau	W^\pm W Boson	

Abbildung 1: Die elementaren Teilchen im Standardmodell [1].

Leptonen sind im ATLAS-Detektor nachweisbar. In Modellen jenseits des Standardmodells werden weitere Higgs-Bosonen vorhergesagt [12], die bevorzugt in Tau-Leptonen zerfallen. Bei der Suche nach diesen Higgsbosonen ist der effiziente Nachweis der Tau-Leptonen bei hoher Untergrundunterdrückung besonders wichtig.

2 Theoretische Grundlagen

2.1 Entstehung von Tau-Leptonen

Bei der Kollision von Protonen im LHC entstehen W- und Z-Bosonen. Diese zerfallen unter anderem in Tau-Leptonen, welche dann wiederum zerfallen. $W \rightarrow \tau\nu_\tau$ und $Z \rightarrow \tau\tau$ sind für die Suche nach $H \rightarrow \tau\tau$ Prozessen fundamentale Hintergrundprozesse [3]. $Z' \rightarrow \tau\tau$ [3] ist ein Prozess, der ebenfalls jenseits des Standardmodells vorhergesagt wird. Sie müssen indirekt über ihre Zerfallsprodukte bestimmt werden. In Tabelle 1 ist eine Übersicht mit den wichtigsten Zerfällen der Tau-Leptonen angegeben.

	Zerfallskanal	Verzweigungsverhältnis
leptonisch	$\tau^+ \rightarrow e^+ + \nu_e + \bar{\nu}_\tau$	17,82 %
	$\tau^+ \rightarrow \mu^+ + \nu_\mu + \bar{\nu}_\tau$	17,41 %
hadronisch	$\tau^+ \rightarrow \pi^+ + \bar{\nu}_\tau$	10,83 %
	$\tau^+ \rightarrow \pi^+ + \pi^0 + \bar{\nu}_\tau$	25,52 %
	$\tau^+ \rightarrow \pi^+ + 2\pi^0 + \bar{\nu}_\tau$	9,30 %
	$\tau^+ \rightarrow \pi^+ + 3\pi^0 + \bar{\nu}_\tau$	1,05 %
	$\tau^+ \rightarrow 2\pi^+ + \pi^- + \bar{\nu}_\tau$	8,99 %
	$\tau^+ \rightarrow 2\pi^+ + \pi^- + \pi^0 + \bar{\nu}_\tau$	2,70 %

Tabelle 1: Zerfälle der Tau-Leptonen [3] für den Fall positiv geladener Tau-Leptonen, für negativ geladene Tau-Leptonen sind die Produkte entsprechend durch die jeweiligen Antiteilchen zu ersetzen.

Hierbei ist erkennbar, dass die hadronischen Zerfälle in drei Gruppen unterscheidbar sind, je nach Anzahl der geladenen Zerfallsprodukte. Bei 1-Prong-Zerfällen gibt es ein geladenes Hadron im Endzustand. 3-Prong-Zerfälle haben drei geladene Hadronen. Zerfälle in fünf und mehr geladene Teilchen sind auch möglich, haben aber eine sehr niedrige Zerfallsbreite und werden daher im Folgenden ignoriert.

Bei der Rekonstruktion von Tau-Zerfällen werden im vorliegenden Fall hadronische Tau-Prozesse betrachtet, da Zerfälle in Elektronen und Myonen nur schwer von denen anderer Prozesse zu unterscheiden sind, bei denen diese direkt entstehen, z.B. $Z \rightarrow ee$ oder $Z \rightarrow \mu\mu$. Die Hadronen sind die beim Zerfall sichtbaren Objekte, weshalb diese auch mit $\tau_{had-vis}$ bezeichnet werden. Die Hadronen werden als Jets rekonstruiert. Im Schritt der Rekonstruktion werden diese Jets nicht von Quark- und Gluon-Jets, sogenannten QCD-Jets, unterschieden.

Diese sind somit in diesem Fall also der Untergrund. Aufgrund der großen Ähnlichkeit von QCD- und Tau-Jets ist nach der ersten Trennung ein weiterer Schritt notwendig, die Tau-Identifikation. In dieser Arbeit werden hierfür künstliche neuronale Netze verwendet, welche in Abschnitt 2.3 genauer beschrieben werden.

2.2 Standardvariablen

Bisher wurden zur Identifikation der Jets aus hadronischen Tau-Zerfällen die in diesem Abschnitt beschriebenen sogenannten Standardvariablen verwendet. Sie haben sich bisher für die Identifikation mit BDTs (Boosted Decision Trees) [7], welche die zur Zeit übliche Methode zur Tau-Identifikation sind, als beste Kombination erwiesen. Die mit 1- bzw. 3-Prong überschriebenen Variablen sind nur für diese Art von Zerfall definiert.

Zur Einteilung der Bereiche gibt es folgende Definition:

cent = $\Delta R < 0,1$ ¹ ist der Zentralbereich.

core = $\Delta R < 0,2$ beschreibt den Kernbereich.

iso = $\Delta R < 0,4$ charakterisiert den Isolationsbereich.

$$f_{core}^{corr} = \frac{\sum_i^{\Delta R_i < 0,1} E_{T,i}^{EM}}{\sum_j^{\Delta R_i < 0,2} E_{T,j}^{EM}}$$

beschreibt das Verhältnis der Transversalenergien im Zentralbereich zum Kernbereich. In der Verteilung erwartet man, dass die Signalverteilung bei kleineren Werten konzentriert ist als die des Hintergrunds ist.

$$f_{track}^{corr} = \frac{p_T^{leadtrk}}{\sum_j^{\Delta R_i < 0,2} E_{T,j}^{EM}}$$

beschreibt den Anteil des Transversalimpulses der impulsreichsten Spur an der Energie im Kernbereich der Interaktion. Da die Anzahl an Teilchen pro Jet bei Tau-Zerfällen geringer ist als bei Jets des Hintergrunds, erwartet man eine Verteilung nahe 1 für Tau-Zerfälle.

$$R_{track} = \frac{\sum_i^{\Delta R_i < 0,4} p_{T,i} \Delta R_i}{\sum_i^{\Delta R_i < 0,4} p_{T,i}}$$

¹ $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ definiert in Kapitel 2.4

ist der Abstand aller Spuren zur τ -Achse gewichtet mit dem Transversalimpuls. Hier erwartet man einen kleineren Wert für Jets aus Tau-Zerfällen, da diese stärker kollimiert sind.

3-Prong-Variablen:

ΔR_{max} ist der größte Abstand zwischen einer Spur im Kernbereich und der Jetachse. Genau wie R_{track} wird hier die Kollimation der Jets beschrieben, was ein ähnliches Ergebnis erwarten lässt.

$$S_T^{flight} = \frac{L_T^{flight}}{\delta L_T^{flight}}$$

ist das Verhältnis zwischen der Zerfallslänge und ihrer Unsicherheit. Tau-Zerfälle haben eine positive Zerfallslänge, wohingegen Jets vom Hintergrund in der Regel nur einen Vertex, den sogenannten Primärvertex besitzen und deshalb hier einen Wert bei 0 erwarten lassen.

$$m_{tracks} = \sqrt{\left(\sum_{i \in (tracks)}^{\Delta R_i < 0,4} E_i\right)^2 - \left(\sum_{i \in (tracks)}^{\Delta R_i < 0,4} \vec{p}_i\right)^2}$$

beschreibt die Masse aller betrachteten geladenen Spuren. Für die Jets von Tau-Zerfällen wird hier ein Peak etwas unterhalb der Tau-Masse erwartet, da neutrale Teilchen hier nicht berücksichtigt werden. Die Jets des Hintergrunds erzeugen ein breiteres Spektrum.

1-Prong-Variablen:

$$S_{leadtrk}^{IP} = \frac{d_0}{\delta d_0}$$

ist der Stoßparameter der impulsreichsten Spur im Verhältnis zu dessen Unsicherheit. Hier wird für Jets aus Tau-Zerfällen ein höherer Absolutwert erwartet als für den Hintergrund.

N_{track}^{iso} beschreibt die Anzahl der Spuren auf dem Isolationsring, der zwischen $\Delta R = 0,2$ und $\Delta R = 0,4$ liegt. Der Wert von N_{track}^{iso} ist für Tau-Zerfälle geringer als für den Hintergrund.

π^0 -Variablen:

$$f_{vis-p_T} = \frac{\sum_i^{\Delta R_i < 0,2} p_{T,i}}{p_T^\tau}$$

ist der Anteil des Transversalimpulses der π^0 -Teilchen am Gesamtimpuls. Hier erwartet man für Jets aus Tau-Zerfällen einen Peak bei 1, für QCD-Jets eine breitere Verteilung.

$$m_{\tau}^{vis} = \sqrt{\left(\sum_{i \in (tracks), \pi_0}^{\Delta R_i < 0,2} E_i\right)^2 - \left(\sum_{i \in (tracks), \pi_0}^{\Delta R_i < 0,2} \vec{p}_i\right)^2}$$

beschreibt die Masse aller sichtbaren π^0 -Teilchen. Diese Variable zeigt eine starke Korrelation zu m_{tracks} auf.

N_{π^0} ist die Anzahl der π^0 Teilchen eines τ -Kandidaten.

Für jede Variable gibt es Verteilungskurven, die die Anteile von Signal und Untergrund an den Variablen zeigen. Anhand dieser kann die Separationskraft der Variable beurteilt werden. Alle Verteilungen der Standardvariablen sind im Anhang A.1 zu finden.

2.3 Neuronale Netze

Künstliche neuronale Netze werden in der Datenverarbeitung verwendet, und sind eine Methode zur multivariaten Analyse von Daten. Dabei dienen Neuronen in der Biologie als Modell [4][8][9][10].

Der Aufbau dieser Netze ist in Abbildung 2 gezeigt. Wie in der Biologie gibt es einzelne Neuronen. Diese sind in verschiedener Schichten angeordnet. Zwischen den Neuronen der einzelnen Schichten gibt es Verbindungen, deren Relevanz im Verlauf bestimmt werden soll. Die Relevanz der Verbindung wird über die Gewichte repräsentiert. Diese geben den Einfluss dieses Neurons auf ein anderes wieder. Ist das Gewicht null, besteht auch keine Verbindung zwischen den Neuronen. Bestimmt werden Gewichte mit Hilfe eines Trainings, in dem Daten verwendet werden, deren Ergebnis bereits bekannt ist. Zwischen den Schichten wird nur in eine Richtung gearbeitet. Das heißt, in der Abbildung nur von links nach rechts, weshalb sie auch Feed-Forward-Networks genannt werden.

Die erste Schicht heißt Eingangsschicht und wird mit den Eingangsvariablen gespeist. Demnach entspricht die Anzahl der Neuronen dieser Schicht der Anzahl der Eingangsvariablen. Die letzte Schicht ist das Ausgabeschicht. Deren Ausgabe ist das Ergebnis des gesamten Netzes. Alle Schichten zwischen der Eingabe- und der Ausgabeschicht nennt man verborgene Schichten. Diese benötigt man bei komplexeren Klassifikationsaufgaben, um das neuronale Netz flexibler auf das Problem reagieren zu lassen.

Zur Informationsverarbeitung im Netz wird in diesem Fall eine gewichtete Summe verwendet:

$$z(x_1, \dots, x_n | b, w_1, \dots, w_n) = \sum_{i=1}^n x_i w_i + b$$

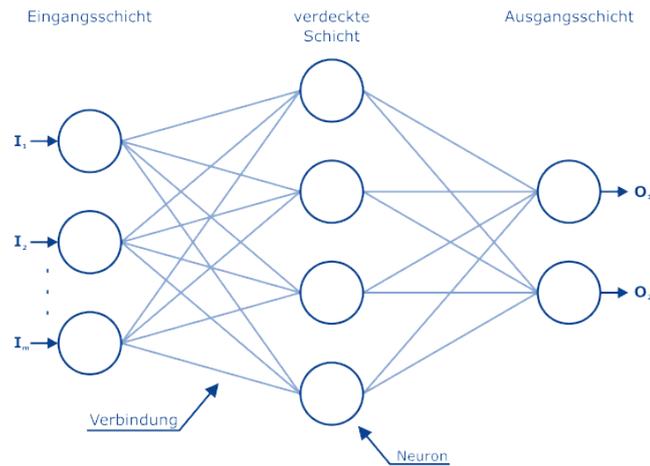


Abbildung 2: Dargestellt ist das Schema eines neuronalen Netzes [14].

wobei x_i die Eingangsvariablen, w_i die Gewichte und b die Bias sind. z ist die Netzeingabe. Bias-Neuronen erzeugen ein konstantes Signal und haben keine Eingabe. Sie dienen der Darstellung eines Schwellenwertes. Der Ausgabewert eines vorangegangenen Neurons wird mit der Gewichtung der Verbindung multipliziert. Über alle eingehenden Neuronen wird dann summiert. Zur Verarbeitung der Eingabe wird eine Aktivierungsfunktion verwendet. Die Aktivierungsfunktion ist für alle Neuronen gleich, nur der Wert, bei dem die einzelnen Neuronen die größte Aktivität aufweisen, variiert. Die in diesem Fall verwendete Aktivierungsfunktion ist der Tangens Hyperbolicus:

$$\tanh(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}.$$

Beim Training werden in das neuronale Netz Datensätze gespeist, deren erwarteter Ausgabewert bekannt ist (Kapitel 3.2).

Beim Training sollen die Gewichte der Neuronen so genau wie möglich angepasst werden, so dass die Netzfunktion möglichst gut an die Zielfunktion angenähert wird. Diese ist in diesem Fall für das Signal der Wert 1 und für den Hintergrund den Wert 0. Die Übereinstimmung beider wird mit der Fehlerfunktion $E(w)$ beschrieben. Erreicht diese ihr Minimum, so ist die beste Übereinstimmung gefunden. Die hier verwendete Fehlerfunktion ist die Kreuzentropie[19].

Zunächst wird die Datensatz in je zwei Teile geteilt, ein Satz für das Training und ein Satz für die Auswertung. Dann werden die Anfangsgewichte festgelegt. Nun beginnt das Training, bei dem mit einer Anzahl von n Epochen folgende Schritte n -fach durchgeführt werden.

1. Eine bestimmte Menge an Daten des Trainingsatzes werden durch das Netzwerk geschickt, um die Netzwerkfunktion zu ermitteln.
2. Die Abweichung der Berechnung wird bestimmt und läuft rückwärts durch das Netz. So

werden die Gewichte der Synapsen ermittelt.

3. Mithilfe der zuvor bestimmten Abweichung werden die Gewichte aktualisiert.
4. Das Training wird beim Erreichen eines vorgegebenen Kriteriums gestoppt.

Wenn das Training abgeschlossen ist, wird der Teil der Daten, der zur Auswertung zur Verfügung steht, zur Bestimmung der Fehlerfunktion verwendet. Diese wird mit der Fehlerfunktion der Trainingsdaten verglichen, um zu bestimmen, wie gut das neuronale Netz das Problem lösen kann. Aufgrund der Eigenschaft neuronaler Netze, sich an einen bestimmten Datensatz anzupassen, auch Überanpassung genannt, ist es notwendig, das Training auf einem separaten Datensatz durchzuführen, und erst danach die Daten zur Auswertung hinein verwenden.

Ein häufiges Problem neuronaler Netze ist, dass bei den verwendeten Algorithmen zur Minimierung der Fehlerfunktion ein lokales Minimum anstelle des globalen als Ergebnis erreicht wird. Des Weiteren werden auf Plateaus sehr viele Schritte durchgeführt, ohne ein Minimum zu identifizieren. Es kann ein Minimum übersprungen werden, oder eine Schleife entstehen. Eine weitere Schwierigkeit ist die Unterscheidung zwischen Minima und Sattelpunkten. Des Weiteren führen zu viele Eingangsvariablen zu einer starken Vergrößerung des Phasenraums, was die Identifikation von Minima erschwert.

Ensembles

In der Analyse mit neuronalen Netzen hat sich herausgestellt, dass diese sehr stark auf Anfangsgewichte und Strukturveränderungen reagieren. Beispielsweise haben die Anzahl der Neuronen pro verborgener Schicht oder die Anzahl der Schichten einen starken Einfluss. Aber auch die zufälligen Anfangsgewichte aller Schicht spielen eine entscheidende Rolle. Um diese Effekte zu vermeiden, werden Ensembles gebildet. Die dabei verwendeten Netze werden alle mit dem gleichen Datensatz trainiert, allerdings variiert die Eingabe der zufälligen Gewichte der Neuronen am Anfang. Die Ausgabe aller Netze wird dann kombiniert. Die einfachste Variante ist das arithmetische Mittel der Ausgangswerte.[21]

Für das Training wurden die folgende zwei Algorithmen verwendet.

SARPROP

SARPROP (Simulated annealing resilient back-propagation) ist einer der hier verwendeten Algorithmen für das Training[8]. Zwei große Vorteile von SARPROP sind die Geschwindigkeit, mit der dieser neuronale Netze trainiert, und die Fähigkeit, aus lokalen Minima herauszukommen. Damit werden häufiger bessere lokale Minima identifiziert.

SARPROP basiert auf RPROP (Resilient back-propagation). Dabei wird der Gradient der

Fehlerfunktion ∇E aus $\frac{\partial E}{\partial w_{ij}}$ bestimmt. Für jedes Gewicht wird zuerst ∇E bestimmt. Dazu werden mit Hilfe des Back-Propagation-Algorithmus die partiellen Ableitungen gebildet. Ist das Vorzeichen von ∇E identisch wie im vorherigen Schritt wird die Schrittweite $\Delta_i(\tau)$ vergrößert, da kein Minimum überschritten worden ist. Bei Änderung des Vorzeichens wird die Schrittweite verkleinert. Nun kann das Gewicht angepasst werden, indem zum vorherigen die neue Schrittweite aufaddiert oder subtrahiert wird, je nachdem ob ein Vorzeichenwechsel stattgefunden hat oder nicht.

SARPROP erweitert diese Methode um die Möglichkeit, lokalen Minima zu entkommen. Hierzu wird bei SARPROP ein Temperaturparameter eingeführt $T = 2^{-\lambda\tau}$, mit λ als Abkühlgeschwindigkeit und τ als die aktuelle Epoche. Mit diesem statistischen Rauschen wird in lokalen Minima die Möglichkeit geschaffen, diese zu verlassen. Er wird zur Schrittweite hinzugefügt, wenn sich das Vorzeichen des Gradienten geändert hat und der Wert der Schrittweite unter einem vorgegeben Minimum liegt. Des Weiteren kommt ein neuer Term zum Gradienten hinzu. Dieser ist von den Gewichten und der Temperatur abhängig. Der Einfluss der beiden hinzugefügten Terme wird aufgrund der Abkühlung weniger, sodass nach einer bestimmten Zeit die Spezialisierung auf ein Minimum ermöglicht wird.

BFGS

BFGS (Broyden-Fletcher-Goldfarb-Shanno-Algorithmus) ist ein Algorithmus, der auf dem Quasi-Newton-Verfahren basiert [20]. Hierbei wird nicht die Hesse-Matrix, die die zweite Ableitung ist, direkt bestimmt, sondern iterativ approximiert. Ein entscheidender Vorteil ist, dass die zweite Ableitung der Fehlerfunktion zusätzlich bestimmt wird und diese weitere Informationen zur Form der Fehlerfunktion liefert.

Vom Prinzip her wird eine Taylor-Entwicklung der Fehlerfunktion um einen bestimmten Punkt durchgeführt. Dabei erhält man einen Ausdruck für die Extrema. Diese könnten aber in lokale Minima führen, da man die Inverse der Hesse-Matrix benötigt und diese nicht unbedingt positiv definit ist.

Dieses Problem wird durch das Quasi-Newton-Verfahren gelöst. Es wird von dem Sekantenverfahren abgeleitet, welches die Näherung aufstellt, dass die Hesse-Matrix zwischen zwei Trainingsepochen konstant ist. Die positive Definitheit der Approximation der Hesse-Matrix bleibt erhalten und somit wird die Fehlerfunktion minimiert.

2.4 LHC und ATLAS

Die Tau-Erkennung erfolgt mit dem ATLAS-Detektor[12] am LHC. Der Large Hadron Collider (LHC) ist ein Synchrotron-Beschleuniger am CERN in Genf. Der Ring hat einen Umfang von 26,7 km. Der LHC ist der derzeit stärkste Teilchenbeschleuniger weltweit, mit einer Schwer-

punktenergie von nominell 13 TeV. Die verwendeten Daten wurden bei einer Schwerpunktsenergie von 8 TeV aufgezeichnet. Die Teilchen, die im Ring kollidieren sind Protonen und werden in mehreren Vorbeschleunigern auf die nötige Energie gebracht, sodass sie im LHC auf die Endenergie beschleunigt werden können. Im Ring befinden sich vier Detektoren: LHCb[23], CMS[22], ALICE[24] und ATLAS. ATLAS (A Toroidal LHC ApartuS), welcher für die hier diskutierte Untersuchung verwendet wird, ist in Abbildung 3 im Querschnitt zu sehen.

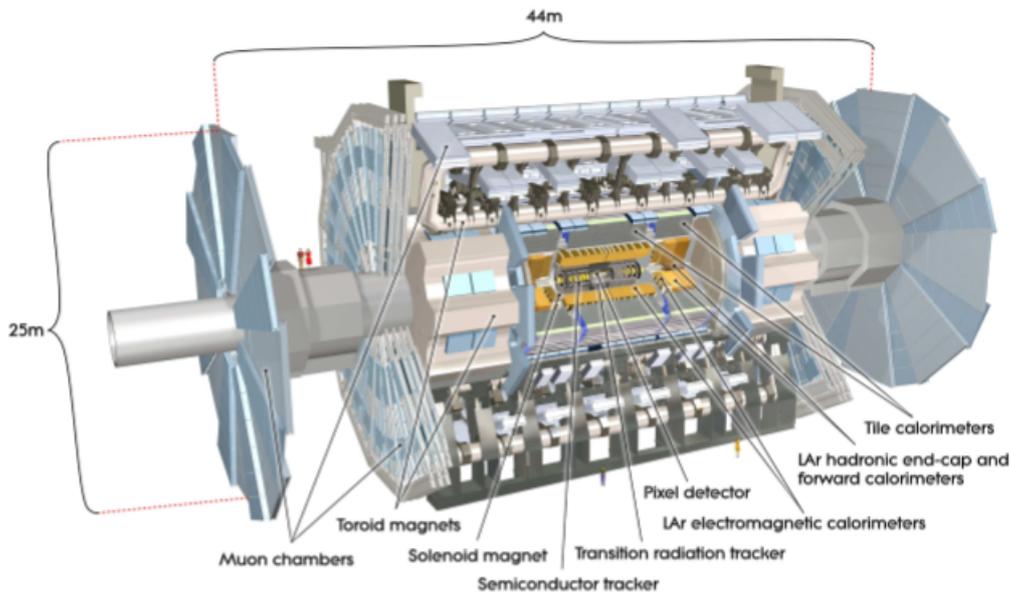


Abbildung 3: Querschnitt des ATLAS-Detektors[11]

Die Detektoren sind zylindrisch um die Strahlrichtung aufgebaut. Der innere Detektor besteht aus einem Pixel-Detektor, einem Silizium-Streifen-Detektor und einem Übergangsstrahlungsdetektor. Des Weiteren befindet sich darum ein Spulenmagnet. Dieser wird benötigt, um geladene Teilchen auf eine gekrümmte Bahn zu zwingen, sodass ihre Impulse gemessen werden können. Im Inneren Detektor werden geladene Teilchen registriert, wie beispielsweise geladene Pionen nach einem Tau-Zerfall. Im Pixel- und SCT-Halbleiter-Detektor erzeugen geladene Teilchen Elektron-Loch-Paare, welche ein messbares Signal erzeugen. Im Übergangsstrahlungsdetektor wird durch Ionisation ein Spursignal erzeugt, welches nachgewiesen wird. Bei der Wechselwirkung der Teilchen an der Grenzfläche zwischen einem Gas und Kunststoff wird zusätzlich Übergangsstrahlung erzeugt. Diese erlaubt nicht nur eine Detektion, sondern auch eine Unterscheidung zwischen leichten Elektronen und schwereren Teilchen.

Die Kalorimeter, ein elektromagnetisches Kalorimeter und ein hadronisches Kalorimeter befinden sich weiter außen im Detektor. Hier werden die Teilchen gestoppt, sodass ihre Energie und ihre Flugrichtung bestimmt werden können. Dabei gibt es zwei Ausnahmen, die Myonen und Neutrinos. Erstere wechselwirken aufgrund ihrer vergleichsweise hohen Masse und der Tatsache, dass sie nicht der starken Wechselwirkung unterliegen nur wenig

und werden im Myonen-Spektrometer detektiert. Dazu wird ein Magnetfeld mit Hilfe von supraleitende Toroidspulen erzeugt. So können sowohl Impuls als auch Ladung der Teilchen bestimmt werden. Letztere werden im ATLAS-Detektor gar nicht erfasst, können jedoch durch die Rekonstruktion fehlender Energie und Anwendung der Energie-Impuls-Erhaltung indirekt nachgewiesen werden. Im elektromagnetischen Kalorimeter werden Elektronen und Photonen, im hadronischen alle stark wechselwirkenden Teilchen detektiert.

Für die Analyse wird ein Koordinatensystem festgelegt. Das Zentrum des Detektors wird als Koordinatenursprung festgelegt und die Strahlachse als z-Achse definiert. Die x-y-Ebene ist senkrecht zur Strahlrichtung. Aufgrund des konzentrischen Aufbaus der Detektoren um die Strahlachse, sind viele Variablen in Zylinderkoordinaten angegeben. ϕ wird wie üblich definiert und θ gemessen. Des weiteren ist die Pseudorapidity wie folgt definiert: $\eta = -\ln(\tanh \frac{\theta}{2})$ und folgt der Winkelabstand zwischen zwei Teilchen als: $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

3 Auswertung

Im Folgenden soll die Identifikation der Tau-Leptonen mit neuronalen Netzen und unter Verwendung der Substrukturvariablen im Fokus stehen. Die zur Auswertung verwendete Software sowie die Daten wurden zur Verfügung gestellt [4][7]. Für die Tau-Kandidaten wurden Monte-Carlo-Simulationen genutzt. Die Daten für den QCD-Hintergrund, bei denen es sich um 2-Jet-Ereignisse bei $\sqrt{s} = 8 \text{ TeV}$ handelt, wurden im Jahr 2012 mit dem ATLAS-Detektor aufgezeichnet.

3.1 Substrukturvariablen

Aufgabe dieser Arbeit ist die Untersuchung der Substrukturvariablen mit Hilfe neuronaler Netze. Dabei sind Substrukturvariablen größtenteils abgewandelte Standardvariablen, die durch ihre Ladung und den betrachteten Öffnungswinkel unterschieden werden. Die Substrukturvariablen werden im Folgenden näher betrachtet.

$N_{\text{chrg/neut/all}}^{\text{cent/core/iso}}$ ist die Anzahl der geladenen, neutralen und aller Teilchen im Zentral-, Kern- und Isolationsbereich.

Exemplarisch ist die Verteilung für $N_{\text{neut}}^{\text{core}}$ in Abbildung 4 dargestellt.

$\Delta R_{\text{max, chrg}}^{\text{core/iso}}$ beschreibt den maximalen Abstand zwischen geladenen Pionen im Kernbereich bzw. im Isolationsbereich.

$\Delta R_{\text{pT, minmax}}^{\text{iso}}$ ist der Abstand zwischen der impulsreichster und impulsärmster Spur.

Der p_T -gewichtete Abstand zur τ -Achse ist

$$R_{\text{cal, chrg/neut/all}}^{\text{core/iso}} = \frac{\sum_{\substack{\Delta R_i < 0,2/0,4 \\ i \in (\pi^\pm)(\pi^0)/(\pi^0, \pi^\pm)}} p_{T,i} \Delta R_i}{\sum_{\substack{\Delta R_j < 0,2/0,4 \\ j \in (\pi^\pm)(\pi^0)/(\pi^0, \pi^\pm)}} p_{T,j}}.$$

Die Variable $R_{\text{cal, chrg}}^{\text{iso}}$ ist in Abbildung 5 gezeigt.

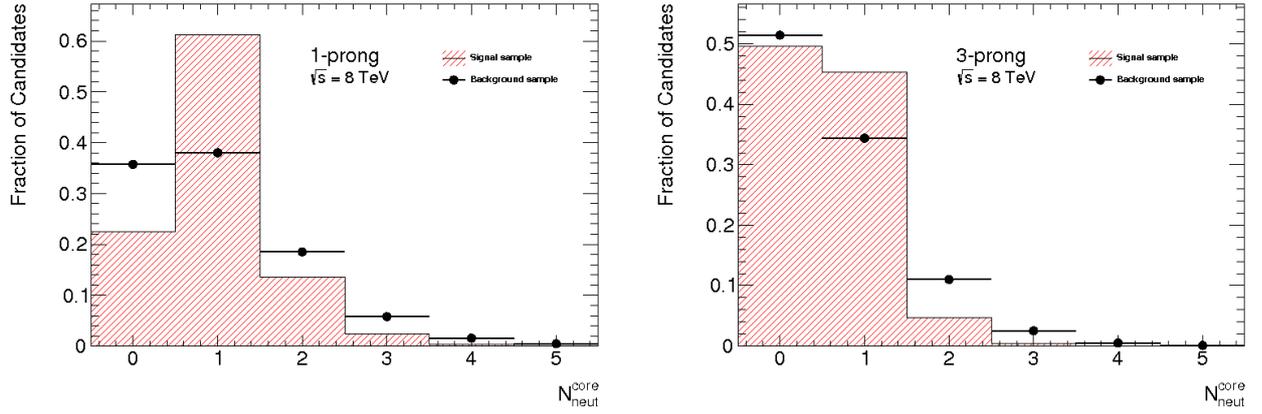


Abbildung 4: Darstellung der Variable N_{neut}^{core} für 1-Prong-Ereignisse (links) und 3-Prong-Ereignisse (rechts), wobei das Signal durch Balken und der Hintergrund durch Striche dargestellt ist.

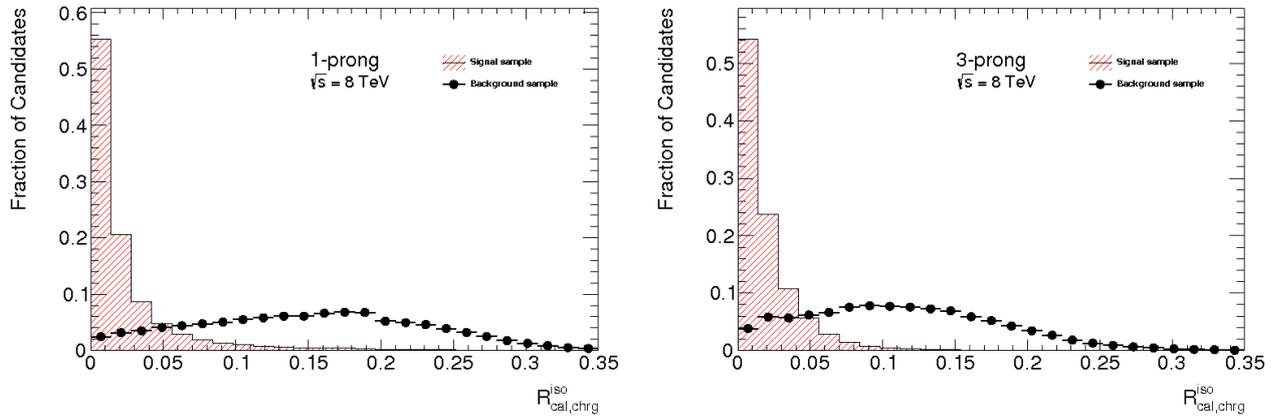


Abbildung 5: Darstellung der Variable $R_{cal,charg}^{iso}$ für 1-Prong-Ereignisse (links) und 3-Prong-Ereignisse (rechts), wobei das Signal durch Balken und der Hintergrund durch Striche dargestellt sind.

$$m_{chg/neut/all}^{cent/core/iso} = \sqrt{\left(\sum_{i \in (\pi^\pm)(\pi^0)/(\pi^0, \pi^\pm)}^{\Delta R_i < 0,1/0,2/0,4} E_i \right)^2 - \left(\sum_{i \in (\pi^\pm)(\pi^0)/(\pi^0, \pi^\pm)} \vec{p}_i \right)^2}$$

ist die invariante Masse der Zerfallsprodukte mit der jeweiligen Unterscheidung nach Teilchensorte.

$$f_{E_T, chg/neut/all}^{cent/core/iso} = \frac{\sum_{i \in (\pi^\pm)(\pi^0)/(\pi^0, \pi^\pm)}^{\Delta R_i < 0,1/0,2/0,4} E_{T,i}}{\sum_{j \in (\pi^\pm)(\pi^0)/(\pi^0, \pi^\pm)} E_{T,j}}$$

beschreibt den Energieanteil der jeweiligen Zerfallsprodukte.

$$f_{lead,chg}^{cent/core/iso} = \frac{\Delta R_i < 0,1/0,2/0,4}{\Delta R_i < 0,1/0,2/0,4} \frac{p_{lead,chg/neut/all}}{\sum_{i \in (\pi^\pm)(\pi^0)/(\pi^0,\pi^\pm)} E_{T,i}}$$

beschreibt den Anteil des Transversalimpulses der impulsreichsten Spur an der Transversalenergie.

$$f_{p_T,chg/neut/all}^{cent/core/iso} = \frac{\Delta R_i < 0,1/0,2/0,4}{\sum_{i \in (\pi^\pm)(\pi^0)/(\pi^0,\pi^\pm)} p_{T,i}} \frac{p_T}{p_T^T}$$

ist der Anteil des Transversalimpulses in der jeweiligen Region am Gesamtimpuls. Alle weiteren Variablenverteilungen sind im Anhang A.2 zu finden.

3.2 Bestätigung vorangegangener Ergebnisse

Zunächst wird überprüft, ob die letzten Ergebnisse [4], die bei der Identifikation von hadronischen Tau-Zerfällen mit neuronalen Netzen erbracht wurden, wieder erzeugt werden können. Dafür wird eine Netzwerkkonfiguration mit dem SARPROP-Algorithmus verwendet (MLP(001), Anhang B). Im Vergleich zu den aktuellsten Ergebnissen liegt hier eine deutliche Verschlechterung vor, wie man in Abbildung 6 sehen kann.

Die dargestellte ROC-Kurve (Receiver-Operation-Characteristic) ist eine Grenzwertoptimierungskurve, in diesem Fall wird die Untergrundunterdrückung über die Signalidentifikationswahrscheinlichkeit aufgetragen. Sie kann als Maß für die Qualität eines Netzes verwendet werden. Wie dem linken Diagramm von Abbildung 6 entnommen werden kann, wird am Punkt (0,62|10) 62% des Signals richtig identifiziert, aber auch jedes zehnte Untergrundereignis dem Signal zugeordnet. Das heißt 10% des Untergrunds würden falsch klassifiziert werden.

Die Bezeichnung MLP (Multi-Layer Perceptron) ist eine allgemeine Bezeichnung für ein Feed-Forward-Netzwerk. Im unteren Teil des Diagramms ist das Verhältnis der Kurven zur Vergleichskurve angegeben. Bei der Auswertung ist zu beachten, dass der hier relevante Bereich zwischen 35% und 70% Signaleffizienz liegt. Ein weiteres Maß für die Güte der Ergebnisse ist das Integral unter der Kurve, auch AUC (Area under Curve) genannt. Auch hierfür wird nur das Effizienzintervall zwischen 0,35 und 0,7 betrachtet.

Aufgrund der schlechteren Ergebnisse wird der Trainingsalgorithmus von SARPROP auf BFGS (Konfiguration Anhang B) umgestellt. Dies erzielt die Ergebnisse aus Abbildung 7, die vergleichbar mit dem aktuellen Stand bei Verwendung der Standardvariablen mit dem BDT sind.

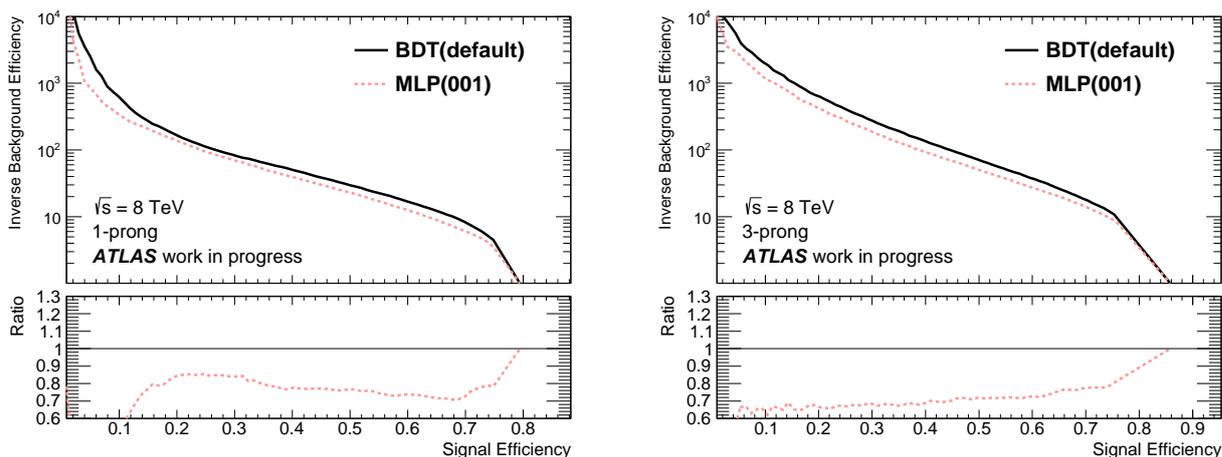


Abbildung 6: Dargestellt sind die ROC-Kurven des BDT und des MLP unter Verwendung des SARPROP-Algorithmus von 1-Prong- (links) und 3-Prong-Ereignissen (rechts).

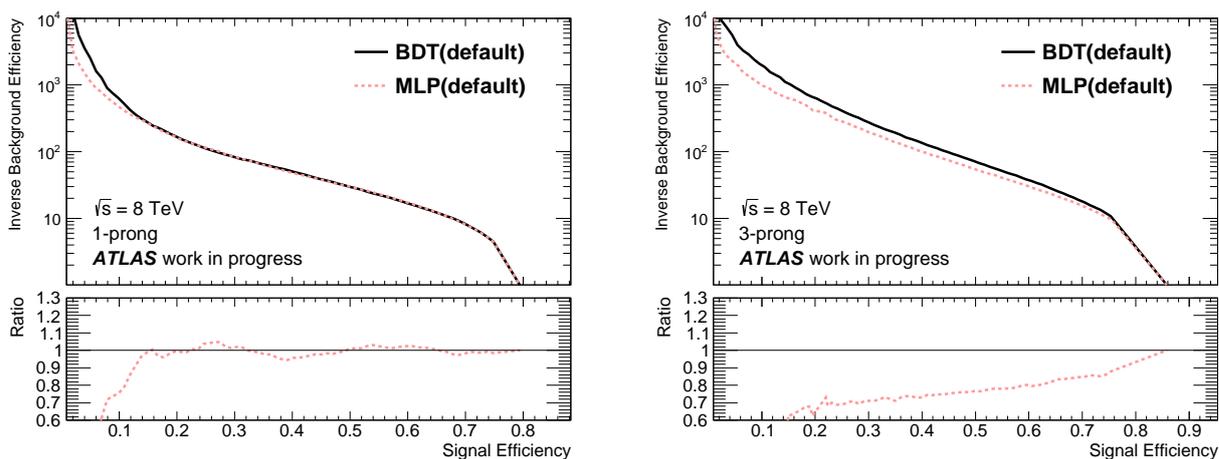


Abbildung 7: Dargestellt sind die ROC-Kurven des BDT und des MLP unter Verwendung des BFGS-Algorithmus von 1-Prong- (links) und 3-Prong-Ereignissen (rechts).

In Abbildung 7 kann man sehen, dass bei Verwendung der Standardvariablen der MLP für 1-Prong-Ereignisse nahezu vergleichbar mit dem BDT ist. Dies wird auch durch die Werte der AUCs bestätigt, welche für 1-Prong-Ereignisse für den BDT 10,10 und das neuronale Netz 9,97 beträgt. Wohingegen für 3-Prong-Ereignisse der MLP um 20% bis 30% vom BDT abweicht. Der AUC ist in diesem Fall für den BDT 23,92 und das neuronale Netz 19,51. Wie in der Abbildung sichtbar, werden die mit den Standardvariablen erzeugten Ergebnisse für den BDT und MLP mit BDT(default) und MLP(default) bezeichnet.

Aufgrund der besseren Ergebnisse wird im Folgenden mit dem BFGS-Algorithmus gearbeitet.

3.3 Untersuchung des Einflusses der Standardvariablen

Um den Einfluss der Standardvariablen bestimmen zu können, werden diese einzeln weggelassen. In Tabelle 2 ist ein Ranking der Standardvariablen anhand der AUCs dargestellt. Dabei ist zu beachten, dass der niedrigste AUC-Wert der Indikator für die stärkste Variable ist.

Tabelle 2: Ranking der Standardvariablen anhand der AUCs.

Ranking	1-Prong		3-Prong	
	Variable	AUC	Variable	AUC
201	f_{core}^{corr}	6.22	S_{flight}^{track}	9.64
202	f_{track}^{corr}	6.59	m_{track}	14.35
203	R_{track}	7.15	R_{track}	15.83
204	S_{lead}^{IP}	8.06	f_{track}^{corr}	17.87
205	N_{track}^{iso}	8.82	f_{pT}	18.61
206	m_{τ}^{vis}	9.29	f_{core}^{corr}	20.03
207	N_{π^0}	9.56	N_{π^0}	20.37
208	f_{pT}	9.8	m_{τ}^{vis}	22.31
209	-	-	ΔR_{max}	23.54

In Abbildung 8 ist die ROC-Kurve für die besten und schlechtesten Ergebnisse mit 1-Prong- und 3-Prong-Ereignissen dargestellt. Für 1-Prong-Ereignisse bringt das Weglassen von f_{pT} ein vergleichbares Ergebnis wie mit dem MLP(default), allerdings kann das MLP in keinem Fall vom Weglassen einer Variable profitieren. Wohingegen das MLP für 3-Prong-Ereignisse ohne m_{τ}^{vis} und ΔR_{max} eine Verbesserung erzielen kann.

Das Weglassen einer Variable kann eine Verbesserung bewirken, da die Größe des Phasenraums von großer Bedeutung ist. Dieser wird mit einer fehlenden Variable um eine Dimension kleiner, wodurch das Lokalisieren globaler Minima leichter möglich ist.

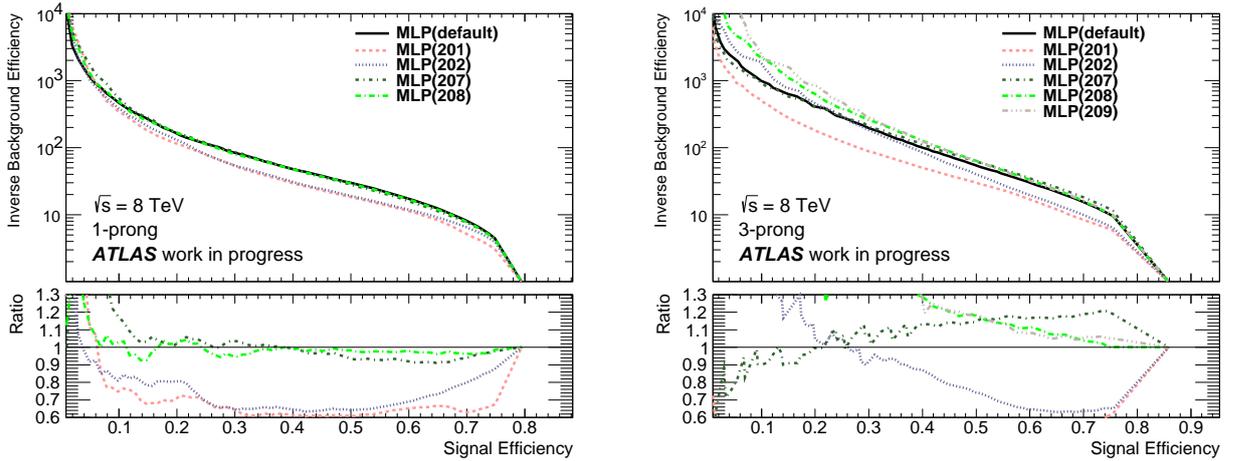


Abbildung 8: Dargestellt sind die ROC-Kurven der MLPs bei Weglassen einer Standardvariable (Tabelle 2) von 1-Prong- (links) und 3-Prong-Ereignissen (rechts).

3.4 Hinzufügen von Substrukturvariablen bei Entfernen von Standardvariablen

Aufgrund dieser Ergebnisse wird N_{π^0} durch Substrukturvariablen ersetzt, auch wenn dies nicht die Variable mit dem schlechtesten Ergebnis ist.. Sie hat sowohl für 1-Prong-, als auch für 3-Prong-Ereignisse ein schlechtes Ergebnis erzielt. Zur Auswahl der hierfür passenden Variablen wird die Separationsfähigkeiten der Variablen verwendet. Die hierbei überprüften Variablen sind in Tabelle 3 mit den jeweiligen AUCs aufgelistet.

Tabelle 3: Hinzugefügte Substrukturvariablen und die Nummer des neuronalen Netzes und der jeweilige AUC.

Neuronales Netz	Variable	AUC	
		1-Prong-Ereignisse	3-Prong-Ereignisse
301	ΔR_{max}^{core}	9.78	23.52
302	ΔR_{max}^{iso}	9.69	20.423
303	$f_{ET}^{cent-iso}$	9.74	20.17
304	$f_{ET}^{core-iso}$	9.81	20.07
305	m_{chg}^{iso}	10.02	20.63
306	$N_{chg}^{core-iso}$	9.83	21.5
307	f_{pT}^{cent}	9.57	18.25
308	f_{pT}^{core}	9.28	19.25
309	$f_{pT,chg}^{core}$	9.79	19.89
310	R_{cal}^{core}	9.57	19.54
311	R_{cal}^{iso}	7.48	18.80
312	$R_{cal,chg}^{core}$	9.39	22.56
313	$R_{cal,chg}^{iso}$	10.01	18.85

Dabei wird für 3-Prong-Ereignisse das beste Ergebnis mit ΔR_{max}^{core} erzielt und für 1-Prong-

Ereignisse mit m_{chrg}^{iso} , wie in Abbildung 9 dargestellt ist. Wie die AUC-Werte in Tabelle 3 zeigen, sind die Ergebnisse alle im Bereich der neuronalen Netze, die mit den Standardvariablen trainiert wurden. In den ROC-Kurven ist ebenfalls sichtbar, dass die Ergebnisse nur gering voneinander abweichen. Es ist anzumerken, dass man nicht klar sagen kann, ob sich ausschließlich durch die Substrukturvariablen Verbesserungen ergeben bzw. wie groß der Einfluss der fehlenden N_{π^0} -Variable ist. Die Ergebnisse geben einen ersten Hinweis darauf, dass Variablen geladener Teilchen gerade für 1-Prong-Ereignisse sehr geeignet sind.

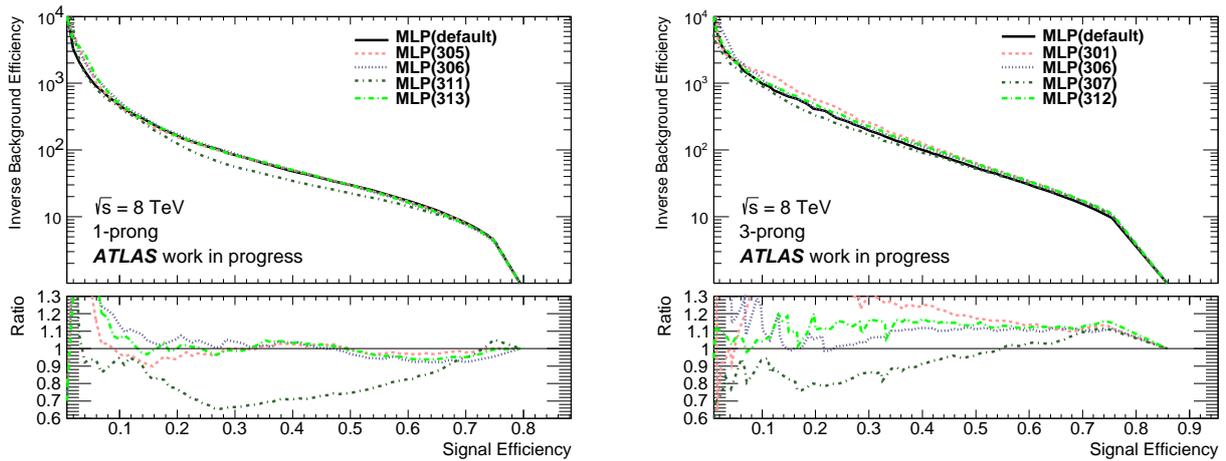


Abbildung 9: Dargestellt sind die ROC-Kurven der MLPs mit den hinzugefügten Substrukturvariablen mit den besten und schlechtesten Ergebnissen (Tabelle 3) für 1-Prong- (links) und 3-Prong-Ereignisse (rechts).

Nachfolgend wird untersucht, ob starke Korrelationen eine negative Auswirkung haben, wie sich das Ergebnis ändert, wenn eine Variable zu dem gesamten Satz von Standardvariablen hinzugefügt wird und wenn die Variable ersetzt wird, die im Ranking am schlechtesten abgeschnitten hat. Die dabei verwendeten Neuronale Netze sind in Tabelle 4 zusammengefasst. Die entsprechenden ROC-Kurven für 1- und 3-Prong-Ereignisse sind in Abbildung 10 dargestellt.

Tabelle 4: Verwendete Variablen der MLPs mit ΔR_{max}^{core} und der jeweilige AUC.

	MLP(301)	MLP(401)	MLP(402)	MLP(403)
Standardvariable 1-Prong	kein N_{π^0}	alle	kein ΔR_{max}	kein m_{τ}^{vis}
Standardvariable 3-Prong	kein N_{π^0}	alle	kein R_{track}	kein f_{p^T}
Substrukturvariable	ΔR_{max}^{core}	ΔR_{max}^{core}	ΔR_{max}^{core}	
AUC für 1-Prong	9.78	9.56	9.72	9.78
AUC für 3-Prong	23.52	19.09	15.57	20.82

Wie in Abbildung 10 deutlich erkennbar ist, ist der Einfluss der unterschiedlichen Variablen auf 3-Prong-Ereignisse größer als auf 1-Prong-Ereignisse. Für 1-Prong-Ereignisse ist deutlich

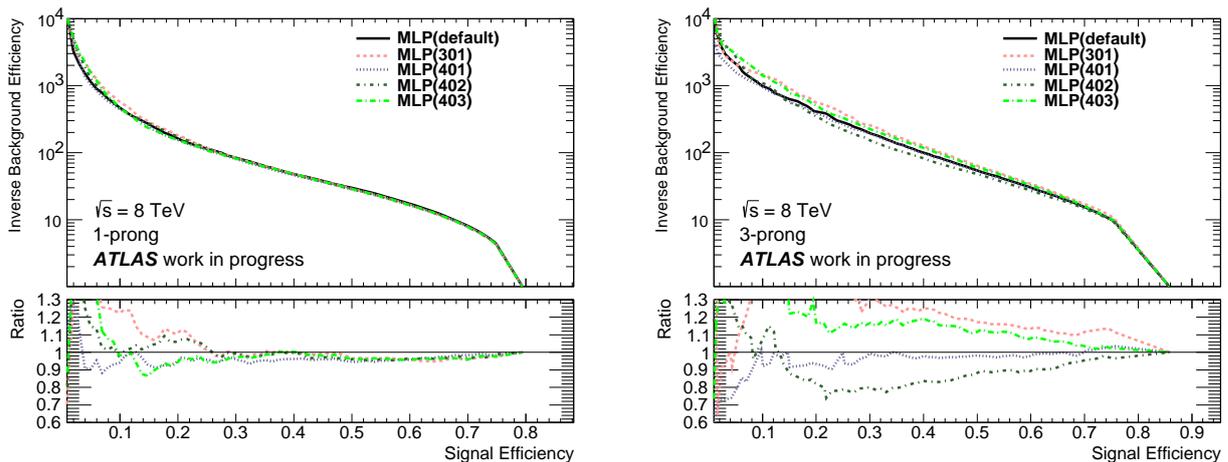


Abbildung 10: Ergebnisse der Neuronalen Netze mit Untersuchungen zu den Auswirkungen von Korrelationen, dem Separationsranking und der Größe des Netzes. Dargestellt sind die ROC-Kurven der MLPs mit ΔR_{max}^{core} (Tabelle 3) für 1-Prong- (links) und 3-Prong-Ereignisse (rechts).

erkennbar, dass alle Kurven, nahezu gleich sind. Eine leichte Verschlechterung weist die Kurve auf, bei der ΔR_{max}^{core} hinzugefügt wird. Des Weiteren sind alle Kurven geringfügig schlechter als das MLP mit den Standardvariablen.

Bei 3-Prong-Ereignissen treten die Unterschiede deutlicher hervor. Durch das Ersetzen von N_{π^0} und von m_{τ}^{vis} zeigt sich eine Verbesserung gegenüber dem MLP(default). Wobei das Ersetzen der Variable m_{τ}^{vis} ein schlechteres Ergebnis erzielt als das Ersetzen von N_{π^0} , obwohl erstere schwächer eingestuft wurde (siehe Tabelle 2). Im Gegensatz dazu verursachte das Hinzufügen von ΔR_{max}^{core} eine leichte Verschlechterung, sodass mit dem Ersetzten von ΔR_{max} ein Defizit von 25% bis 5% entsteht.

Demnach führen starke Korrelationen nicht ausschließlich zu schlechteren Ergebnissen. Aus den vorliegenden Unterschieden gewinnt das neuronale Netz Informationen, die zu verbesserten Ergebnissen führen.

3.5 Vollständig ersetzte Variablenkonfigurationen

Im Folgenden wird auf vorangegangene Arbeiten [5][7] zurückgegriffen, die sich mit ähnlichen Problemstellungen auseinandersetzen. Aus ihnen werden drei getestete Variablenkonfigurationen gewählt, die Tabelle 5 entnommen werden können. MLP(501) verwendet die Variablen, die für den BDT und MLP(502) die, die allgemein am stärksten eingeschätzt werden. Bei der dritten Konfiguration werden drei bzw. zwei Variablen zu den Standardvariablen hinzugefügt. Abbildung 11 zeigt die Ergebnisse.

Es ist deutlich erkennbar, dass die neuronalen Netze für 1-Prong-Ereignisse mit den obigen

Tabelle 5: Variablen der höchsten Rankings für den BDT(501) und generell(502) und der Variablenkonfiguration, die durch Hinzufügen zweier Substrukturvariablen (503) profitieren kann.

MLP(501)		MLP(502)		MLP(503)	
1-Prong-Var.	3-Prong-Var.	1-Prong-Var.	3-Prong-Var.	1-Prong-Var.	3-Prong-Var.
$R_{cal,all}^{iso}$	$R_{cal,chg}^{iso}$	$R_{cal,all}^{iso}$	$R_{cal,chg}^{iso}$	Standardvar.	Standardvar.
$f_{cent-iso}^{E,chg}$	$\Delta R_{max,all}^{iso}$	N_{chg}^{iso}	$f_{core-iso}^{E,chg}$	m_{neut}^{iso}	$f_{pT,chg}^{iso}$
$\Delta R_{max,all}^{iso}$	m_{chg}^{iso}	$\Delta R_{max,all}^{iso}$	S_T^{flight}	$f_{lead,chg}^{iso}$	f_{pT}^{iso}
$f_{pT,all}^{cent}$	m_{chg}^{cent}	$f_{E,chg}^{cent-iso}$	N_{chg}^{iso}	$f_{pT,chg}^{iso}$	
m_{all}^{iso}	m_{all}^{cent}	$S_{lead,track}^{IP}$	$f_{pT,chg}^{core}$		
$f_{lead,chg}^{iso}$	$f_{pT,all}^{cent}$	$f_{pT,all}^{core}$	$f_{pT,neut}^{iso}$		
	S_T^{flight}	m_{all}^{iso}	m_{all}^{core}		
	$f_{lead,chg}^{iso}$	m_{neut}^{core}	$S_{lead,track}^{IP}$		
		$f_{pT,neut}^{iso}$			

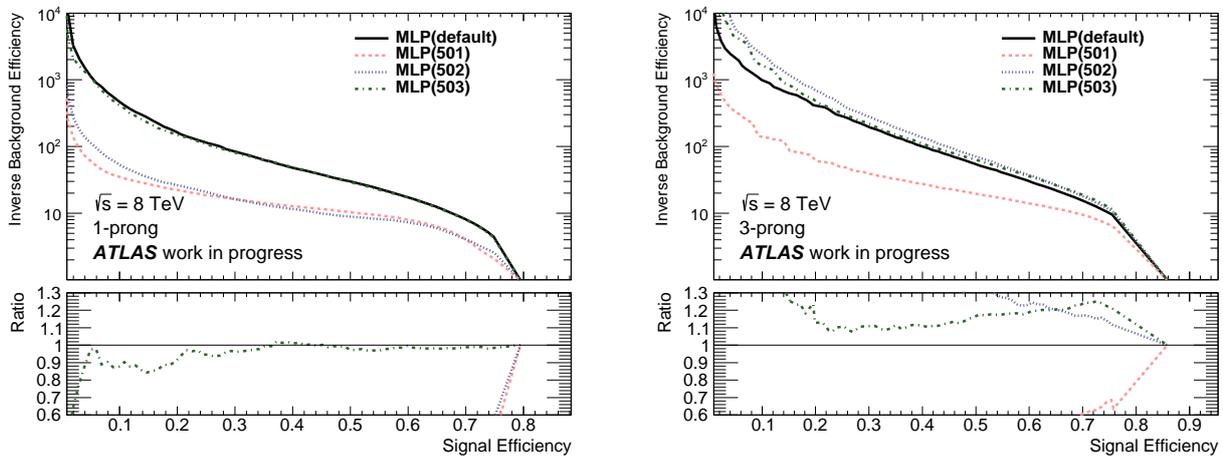


Abbildung 11: Dargestellt sind die ROC-Kurven der MLPs der vollständig ersetzten Variablenkonfigurationen und die hinzugefügten Variablen für 1-Prong- (links) und 3-Prong-Ereignisse (rechts).

Kombinationen keine verbesserten Ergebnisse liefern. Lediglich der Satz mit den hinzugefügten (MLP(503)) und nicht komplett ausgetauschten Variablen ist vergleichbar mit dem normalen Satz.

Bei den für den BDT als geeignet eingeschätzte Variablen haben die Ergebnisse für 3-Prong-Ereignisse mit deutlichem Abstand die größte Abweichung zum MLP(default). Der Variablensatz, bei dem den Standardvariablen zwei weitere hinzugefügt werden, ist leicht besser als die Standardvariablen allein. Somit ist anzumerken, dass die Vergrößerung des Phasenraums bei der hier verwendeten Anzahl von Variablen noch keine starken Auswirkungen hat. Bei den komplett ersetzten Standardvariablen durch Substrukturvariablen (MLP(502)) gibt es eine erhebliche Steigerung. Der AUC liegt hier bei 24,04, was im Bereich des BDT(default) liegt. Des Weiteren ist zu beobachten, dass die Verbesserung, die das neuronale Netz gewinnt, im BDT für die Konfiguration 502 nicht vorhanden ist.

Die für 3-Prong-Ereignisse gute Ergebnisse erzielende Variablen werden auch für die 1-Prong-Ereignisse eingesetzt und getestet (MLP(5021)). Während diese Konfiguration für 3-Prong-Ereignisse sehr gute Ergebnisse erzielt, ist für 1-Prong-Ereignisse keine Verbesserung sichtbar, wie Abbildung 12 zeigt. Weiterhin werden diese Variablen mit dem SARPROP-Algorithmus trainiert (MLP(5022)), wobei in Abbildung 12 deutlich erkennbar ist, dass dies eine starke Verschlechterung bewirkt.

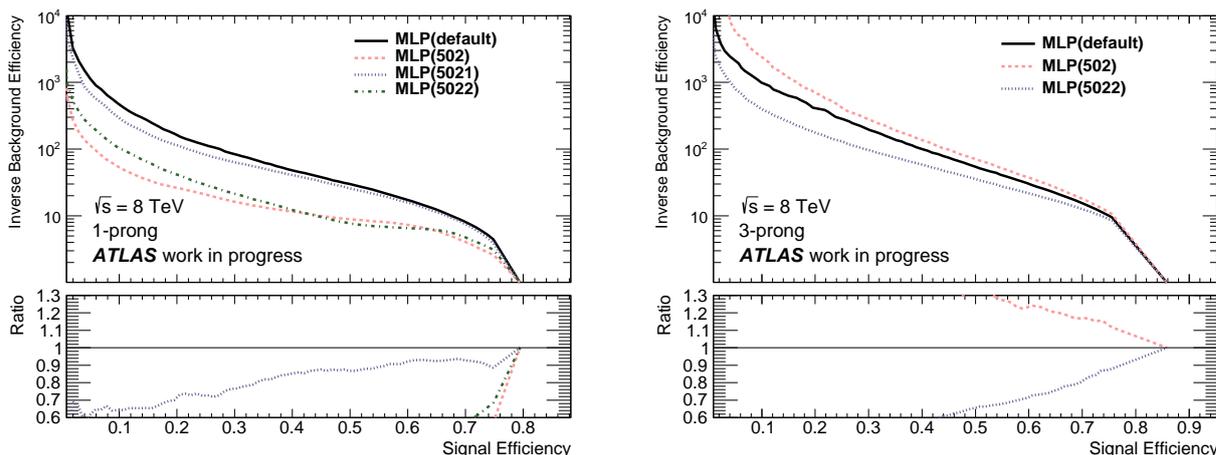


Abbildung 12: Dargestellt ist das MLP bei dem die Variablen, die für 3-Prong-Ereignisse gute Ergebnisse erzielten, für 1-Prong-Ereignisse verwendet werden. Des Weiteren wird ein erneuter Test mit SARPROP durchgeführt, wobei 1-Prong- (links) und 3-Prong-Ereignisse (rechts gezeigt sind).

Im Folgenden wird das auf der Separationskraft der Variablen beruhende Ranking untersucht. Die Liste der verwendeten Variablen ist in Tabelle 6 gezeigt (MLP(701)). Außerdem wird ein Ranking mit zusätzlicher Berücksichtigung der Korrelationen erstellt (MLP(702)). Die Korrelationen sind in Abbildung 13 und 14 dargestellt. In diesem Diagramm werden allerdings nur Korrelationen erster Ordnung betrachtet.

Abbildung 15 zeigt die ROC-Kurven dieser zwei Variablensätze. Im Vergleich zum MLP(default) kann man für beide Konfigurationen eine deutliche Verschlechterung sehen. Es kann im Ergebnis darauf geschlossen werden, dass eine gute Separationskraft allein unzureichend ist, um einen effektiven Informationsgewinn in neuronalen Netzen zu gewährleisten. Auch Korrelationen scheinen nur einen untergeordneten Einfluss zu haben, denn die Verbesserung mit deren Berücksichtigung ist minimal. Auf den Einfluss diskreter Variablen wird in Abschnitt 3.7 eingegangen.

Tabelle 6: Variablen der höchsten Rankings allgemein (MLP(701)) und mit Berücksichtigung der Korrelationen (MLP(702)).

MLP(701)		MLP(702)	
1-Prong-Var.	3-Prong-Var.	1-Prong-Var.	3-Prong-Var.
$R_{cal,chg}^{iso}$	$\Delta R_{max,all}^{iso}$	$R_{cal,chg}^{iso}$	$\Delta R_{max,all}^{iso}$
R_{track}	$f_{cent-iso}^{E,chg}$	$f_{core-iso}^{E,chg}$	$f_{cent-iso}^{E,chg}$
$f_{cent-iso}^{E,chg}$	$f_{cent}^{pT,chg}$	$f_{core-iso}^{E}$	$f_{cent}^{pT,chg}$
$\Delta R_{max,all}^{iso}$	$\Delta R_{minmax,pT,chg}^{iso}$	$f_{lead,chg}^{iso}$	m_{track}
$R_{cal,chg}^{core}$	m_{track}	f_{cent}^{corr}	$\Delta R_{max,all}^{core}$
$\Delta R_{max,all}^{core}$	m_{chg}^{iso}	m_{T}^{iso}	R_{track}
$f_{core-iso}^{E,chg}$	ΔR_{max}^{core}	$f_{cent-core}^{E}$	$n_{chg}^{core-iso}$
R_{cal}^{iso}		m_{chg}^{iso}	$R_{cal,chg}^{core}$

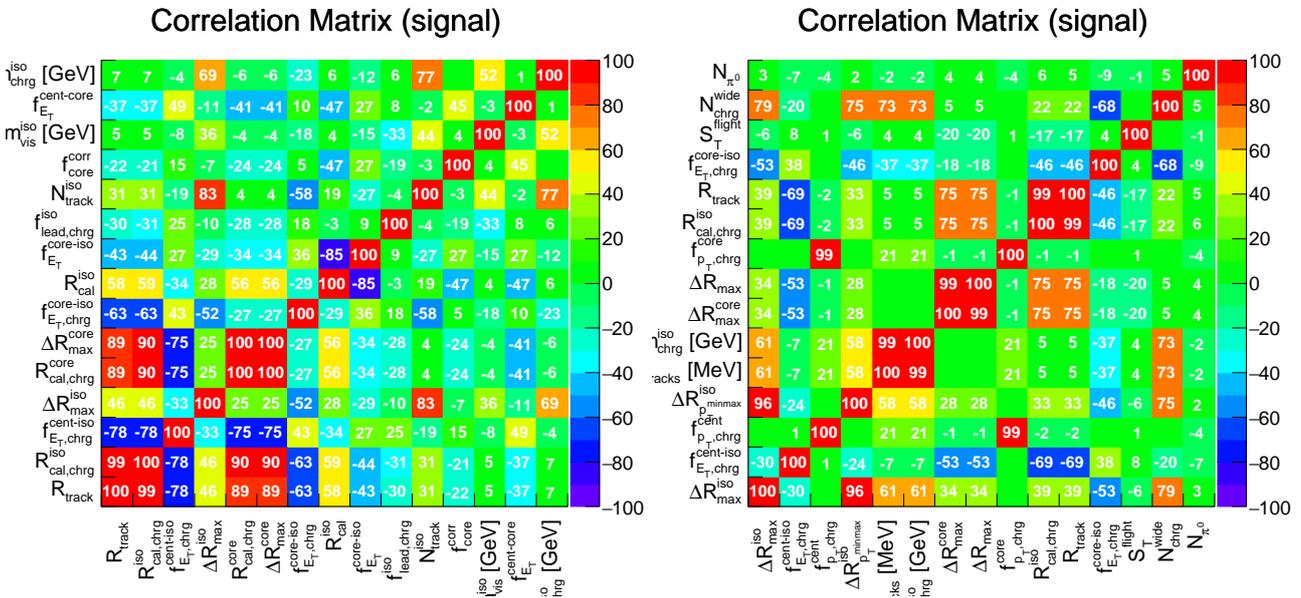


Abbildung 13: Korrelationsplots des Signals der am höchsten eingestuft Variablen für 1-Prong- (links) und 3-Prong-Ereignisse (rechts), wobei die Korrelationen in Prozent angegeben sind.

3.6 Hinzufügen von Substrukturvariablen zu den Standardvariablen

Alle aus Tabelle 5 noch nicht einzeln untersuchten Variablen werden, wie in Tabelle 7 gezeigt, einzeln zu den Standardvariablen hinzugefügt. Außerdem werden aus dem anhand der Separationskraft erstellten Ranking alle fehlenden Variablen hinzugefügt. Das ergibt sechs neue Untersuchungen folgender Substrukturvariablen (Tabelle 7).

Tabelle 7: Übersicht über die zu den Standardvariablen hinzugefügten Substrukturvariablen aufgelistet mit den AUCs.

	MLP(601)	MLP(602)	MLP(603)	MLP(604)	MLP(605)	MLP(606)
1-Prong-Var.	$\Delta R_{minmax,p_T,chg}^{iso}$	$f_{E_T}^{cent-core}$	$f_{E_T,chg}^{cent-iso}$	$f_{E_T,chg}^{core-iso}$	$f_{lead,chg}^{iso}$	m_{vis}^{iso}
3-Prong-Var.	$\Delta R_{minmax,p_T,chg}^{iso}$	$f_{p_T,chg}^{iso}$	$f_{E_T,chg}^{cent-iso}$	$f_{E_T,chg}^{core-iso}$	$f_{p_T,chg}^{cent}$	m_{vis}^{iso}
AUC 1-Prong	6.85	9.97	2.96	2.17	10.05	10.14
AUC 3-Prong	15.42	25.63	9.86	2.79	25.16	19.66

Die gesamten Ergebnisse sind in Abbildung 16 dargestellt. $\Delta R_{minmax,p_T,chg}^{iso}$, $f_{E_T,chg}^{cent-iso}$ und $f_{E_T,chg}^{core-iso}$ erzeugen hierbei die schlechtesten Ergebnisse. Für 3-Prong-Ereignisse ergaben alle anderen Netze eine Verbesserung und für 1-Prong-Ereignisse waren sie mindestens genau so gut wie das MLP(default).

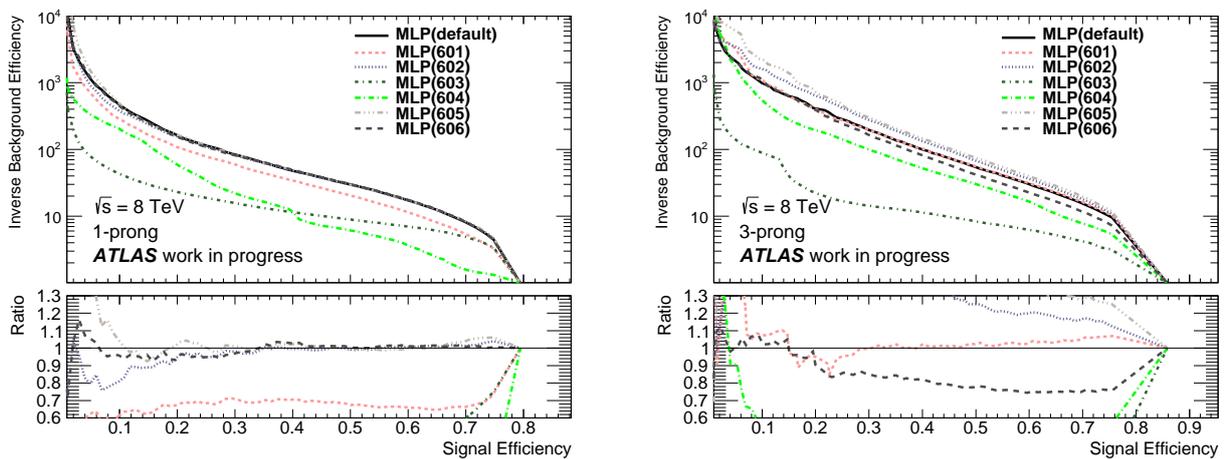


Abbildung 16: Dargestellt sind die ROC-Kurven der sechs verschiedenen MLPs mit jeweils einer hinzugefügten Variable aus Tabelle 6 für 1-Prong-Ereignisse(links) und 3-Prong-Ereignisse(rechts).

In Abbildung 17 stehen BDT und MLP für die bisher besten Ergebnisse im Vergleich. Für 1-Prong-Ereignisse ist ersichtlich, dass das MLP mit keiner Konfiguration profitiert, während die Verbesserung für den BDT(606) bei bis zu 9% und für den BDT(605) bei bis zu 8% liegt.

Diese Verbesserung ist für 3-Prong-Ereignisse bei dem BDT(602) stärker sichtbar, allerdings liegt die Steigerung bei dem MLP(602) bei über 30%. Das Ergebnis des MLP's mit der hinzugefügten Variable $f_{pT,chg}^{iso}$ ist zwar besser als der BDT(default), allerdings bleibt das MLP(602) hinter dem BDT mit dem gleichen Variablensatz zurück. Der AUC liegt für das neuronale Netz (602) bei 25,63 und für den vergleichbaren BDT(602) bei 28,60.

Des Weiteren liegt im Vergleich zu den bisher besten Ergebnissen sowohl für 1-Prong- als auch für 3-Prong-Ereignisse eine Verbesserung der neuronalen Netze vor.

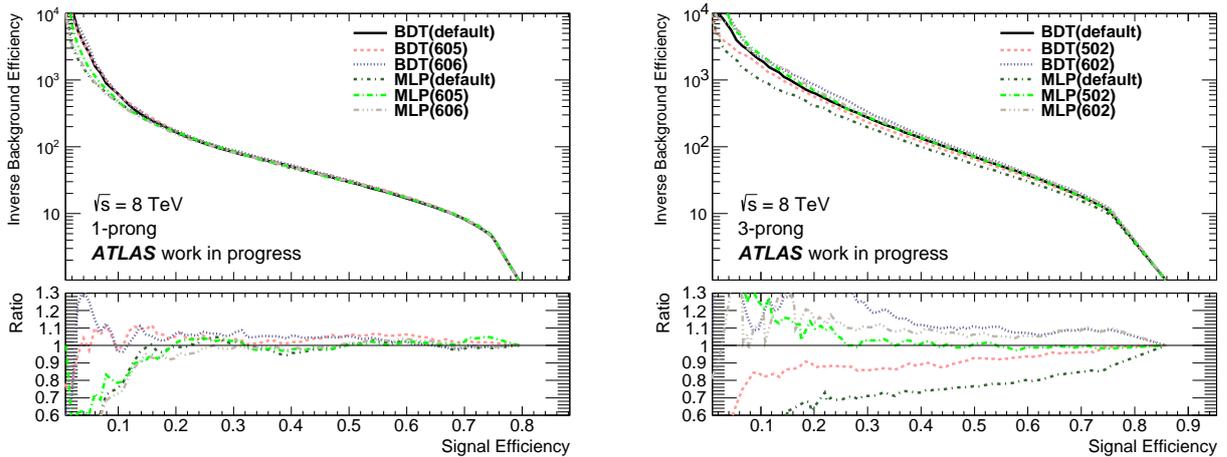


Abbildung 17: Dargestellt sind die ROC-Kurven der jeweils besten Kombination vom MLP und BDT im Vergleich zum BDT(default) für 1-Prong- (links) und 3-Prong-Ereignisse (rechts).

Darüber hinaus wird der Einfluss des Weglassens von N_{π^0} bei dem Hinzufügen von m_{vis}^{iso} (MLP(6061)) für 1-Prong-Ereignisse und $R_{pT,chg}^{iso}$ (MLP(6021)) für 3-Prong-Ereignisse untersucht. Wie in Abbildung 15 zu sehen ist, ergibt diese Modifikation noch einmal eine leichte Verbesserung gegenüber dem Ergebnis mit N_{π^0} bei 1-Prong-Ereignissen, wohingegen bei 3-Prong-Ereignissen keine weitere Verbesserung erzielt werden kann. Der AUC für erstere liegt bei 10,41 für das MLP und bei 10,33 für den BDT. Auffällig ist weiterhin, dass der BDT für 1-Prong-Ereignisse sogar eine Verschlechterung verzeichnet. Die Verschlechterung des BDT für 3-Prong-Ereignisse ist jedoch geringer als die des MLP's.

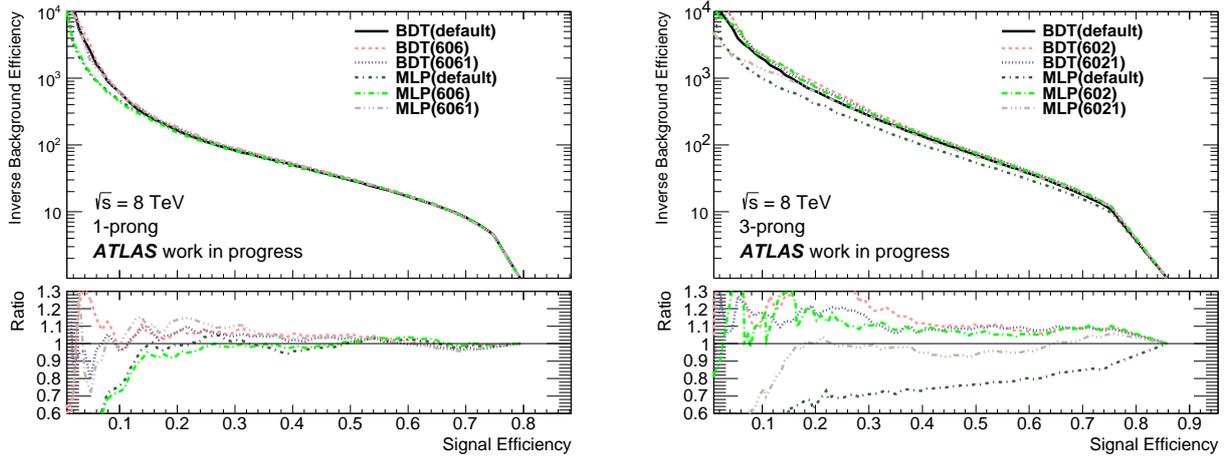


Abbildung 18: Dargestellt sind die ROC-Kurven der jeweils besten Kombinationen ohne die Variable N_{π^0} vom MLP und BDT im Vergleich zum BDT(default) und MLP(default) für 1-Prong- (links) und 3-Prong-Ereignisse (rechts)

3.7 Einfluss diskreter Variablen

Dafür kann es verschiedene Gründe geben. Eine mögliche Ursache, die auch aufgrund der ersten Tests zu vermuten wäre, ist der negative Einfluss diskreter Variablen. Sie schaffen Plateaus, die es dem neuronalen Netz schwerer machen das globale Minimum der Fehlerfunktion zu finden. Aufgrund dieser Annahme wird aus dem MLP(502) für 3-Prong-Ereignisse die diskrete Variable $N_{chr}^{core-iso}$ weggelassen (MLP(901)). In Abbildung 19 ist eine deutliche Verbesserung im Vergleich zum Ergebnis mit der diskreten Variable sichtbar. Es zeigt sich noch immer ein um 50 % schlechteres Ergebnis als das MLP(default).

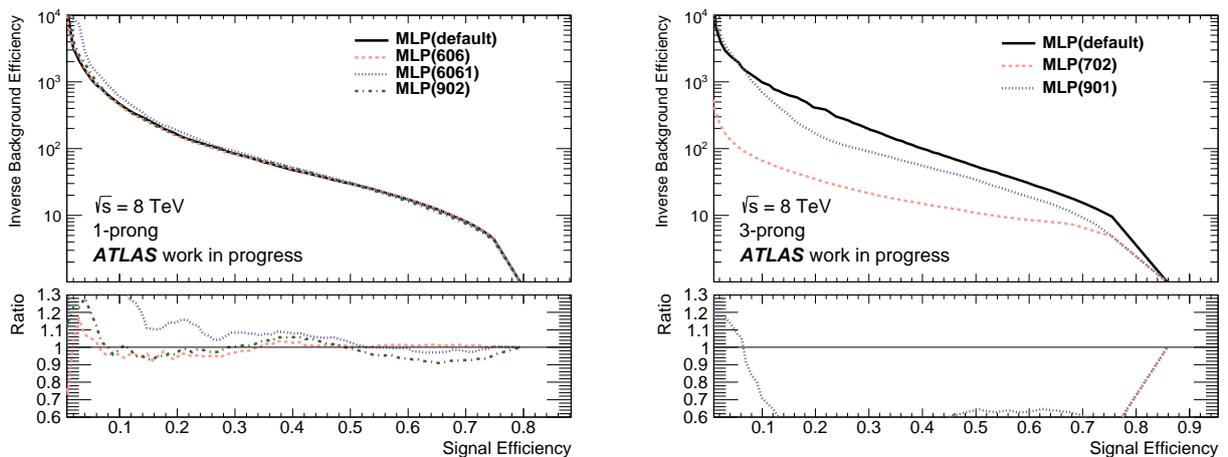


Abbildung 19: ROC-Kurve der neuronalen Netze mit (MLP(701)) und ohne (MLP(901)) diskrete Variable für 3-Prong-Ereignisse (rechts). ROC-Kurve der neuronalen Netze mit (MLP(606)) und ohne (MLP(6061), MLP(902)) diskrete Variable für 1-Prong-Ereignisse (links).

Um das bisher beste Ergebnis unter diesem Aspekt zu untersuchen, wird das MLP(606) für 1-Prong Ereignisse ohne diskrete Variablen betrachtet. Dies bedeutet, dass $N_{\tau 0}$ und zusätzlich N_{chrg}^{iso} weggelassen werden (MLP(902)). Die daraus resultierende ROC-Kurve ist ebenfalls in Abbildung 19 sichtbar. Es zeigt sich, dass das Weglassen der diskreten Variablen eine Verschlechterung hervorgerufen hat. Somit zeigt sich, dass diskrete Variablen keine negativen Auswirkungen auf die neuronalen Netze haben.

3.8 Stabilitätsuntersuchung

Zusätzlich zu den Analysen der Substrukturvariablen wird eine Stabilitätsuntersuchung durchgeführt, um die Ergebnisse vergleichbar bewerten zu können und Abweichungen einschätzen zu können. Hierfür werden 30 Netze mit gleicher Architektur, aber unterschiedlichen Anfangsgewichten trainiert. Die dazu verwendeten Variablen waren die Standardvariablen erweitert um m_{τ}^{iso} für 1-Prong-Ereignisse und f_{pTchrg}^{iso} für 3-Prong-Ereignisse. Danach werden jeweils fünf Netze zu einem Ensemble zusammengefasst, also sechs Ensembles erstellt. Daraus ergeben sich die Werte aus Tabelle 8 für den AUC und dessen Abweichungen.

Tabelle 8: Mittelwert und Standardabweichung der AUCs für 1- und 3-Prong-Zerfälle der einzelnen neuronalen Netze und der Ensembles.

	1-Prong AUC	3-Prong AUC
Neuronales Netz	8.80 ± 0.37	19.70 ± 2.46
Ensemble	10.11 ± 0.20	24.44 ± 0.78

In Abbildung 20 sind die sechs Ensembles im Vergleich gegenübergestellt. Für 1-Prong-Ereignisse kann man eine geringere Abweichung sehen. Bei 3-Prong-Ereignissen ist die Abweichung im Bereich unter einer Signaleffizienz von 0.5 sehr stark und insgesamt sehr viel größer als bei 1-Prong-Ereignissen. Dies sieht man auch anhand unterschiedlicher Abweichungen der Mittelwerte der AUCs. Wichtig ist, dass für 3-Prong-Zerfälle weniger Ereignisse vorhanden sind, sodass eine größere Schwankung erklärt werden kann. Dies führt zu einer geringeren Genauigkeit der Daten.

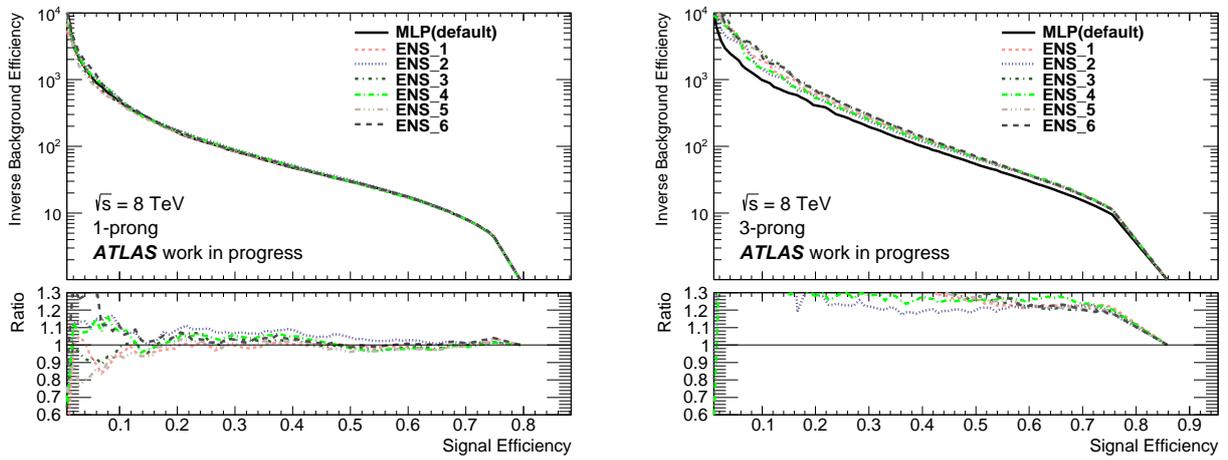


Abbildung 20: Dargestellt sind die ROC-Kurven der sechs Ensembles für 1-Prong- (links) und 3-Prong-Ereignisse (rechts).

4 Zusammenfassung und Ausblick

Es hat sich gezeigt, dass künstliche neuronale Netze von den Substrukturvariablen profitieren können. Sowohl das Hinzufügen einzelner Variablen, als auch das Austauschen des gesamten Variablensatzes erzielte Verbesserungen.

Abschließend lässt sich sagen, dass die Verbesserungen der Signalerkennung mit Hilfe von neuronalen Netzen, die im Zuge dieser Arbeit erzielt wurden, auch Verbesserungen des konkurrierenden BDTs mit sich brachten. Dieser profitiert von geringeren statistischen Schwankungen und einer weitaus geringeren Trainingszeit. Die reine Berücksichtigung von Korrelationen und der Separationskraft der Einzelvariablen erreicht keine starke Verbesserung. Außerdem erbrachten die neuronalen Netze mit komplett ausgetauschten Variablensätze keine Verbesserung.

Anders ist es beim Hinzufügen von einzelnen Substrukturvariablen. Hier waren für einige Konfigurationen Verbesserungen erkennbar. Für 3-Prong-Ereignisse ist $f_{pT,chr}^{iso}$ die vielversprechendste Variable. Für 1-Prong-Ereignisse ergab sich mit dem Weglassen von N_{π^0} und dem Hinzufügen von m_{vis}^{iso} eine Verbesserung.

Die Möglichkeiten in der vorgegebenen Arbeitszeit sind begrenzt gewesen. Die Studien zu den neuronalen Netzen mit Substrukturvariablen können noch sehr viel weiter betrieben werden. Ansätze hierfür könnte das Testen bestimmter Variablen sein, die bereits logarithmiert wurden. So würden die Verteilungen angepasst werden, um die Verarbeitung dieser im neuronalen Netz zu optimieren. Weiterhin sollten noch weitere Konfigurationen getestet werden, die eine weitere Mischung von Standard- und Substrukturvariablen untersuchen.

Alles in allem hat diese Arbeit gezeigt, dass die Substrukturvariablen eine Verbesserung bei der Tau-Identifizierung ermöglichen. Ob eine größere Steigerung der Untergrundunterdrückung als bei den BDTs möglich ist, müssen allerdings weitere Untersuchungen zeigen.

A Verteilungen der Variablen

A.1 Standardvariablen

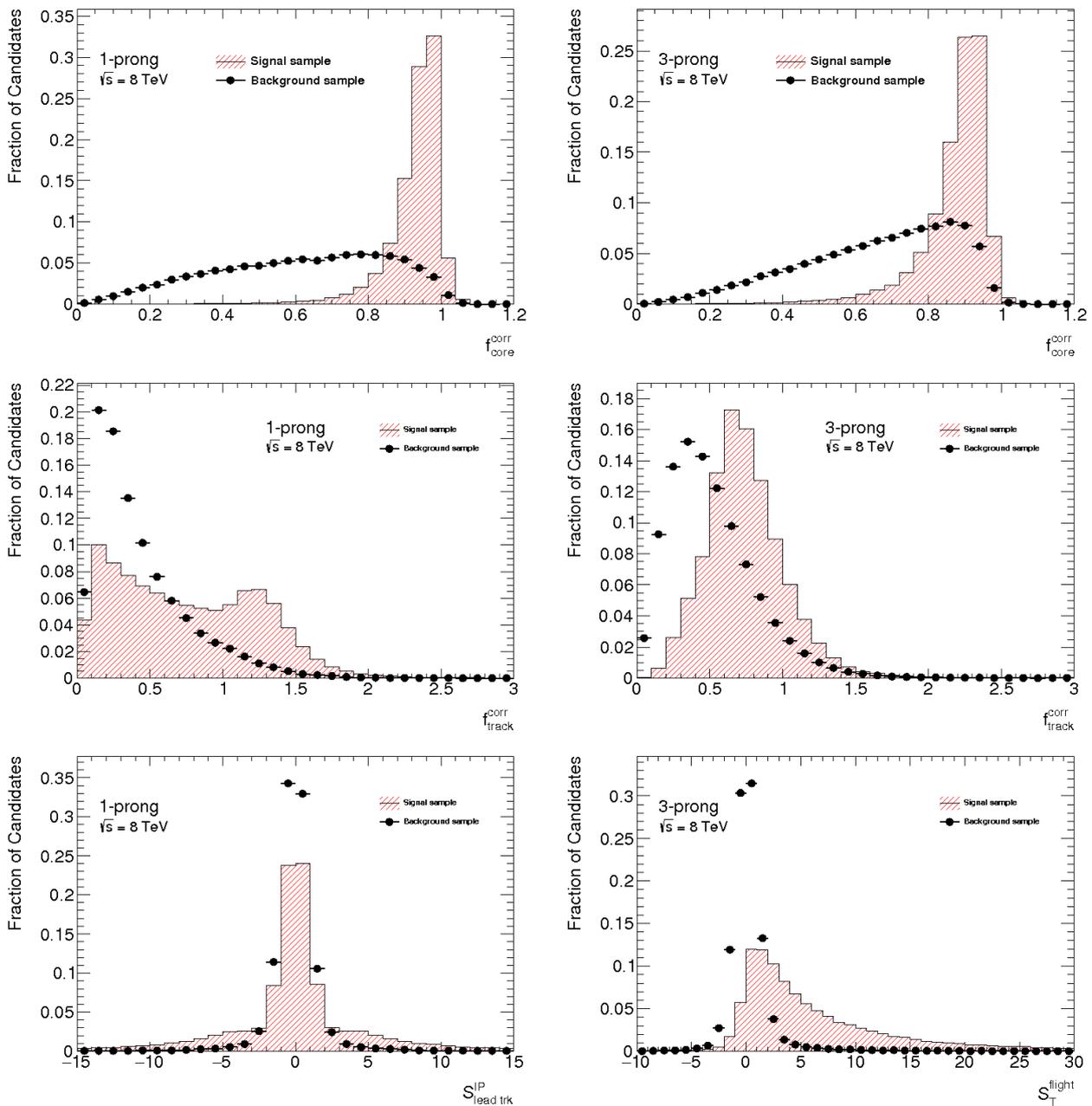


Abbildung 21: Verteilungen der Standardvariablen für 1-Prong- (links) und 3-Prong-Ereignisse (rechts).

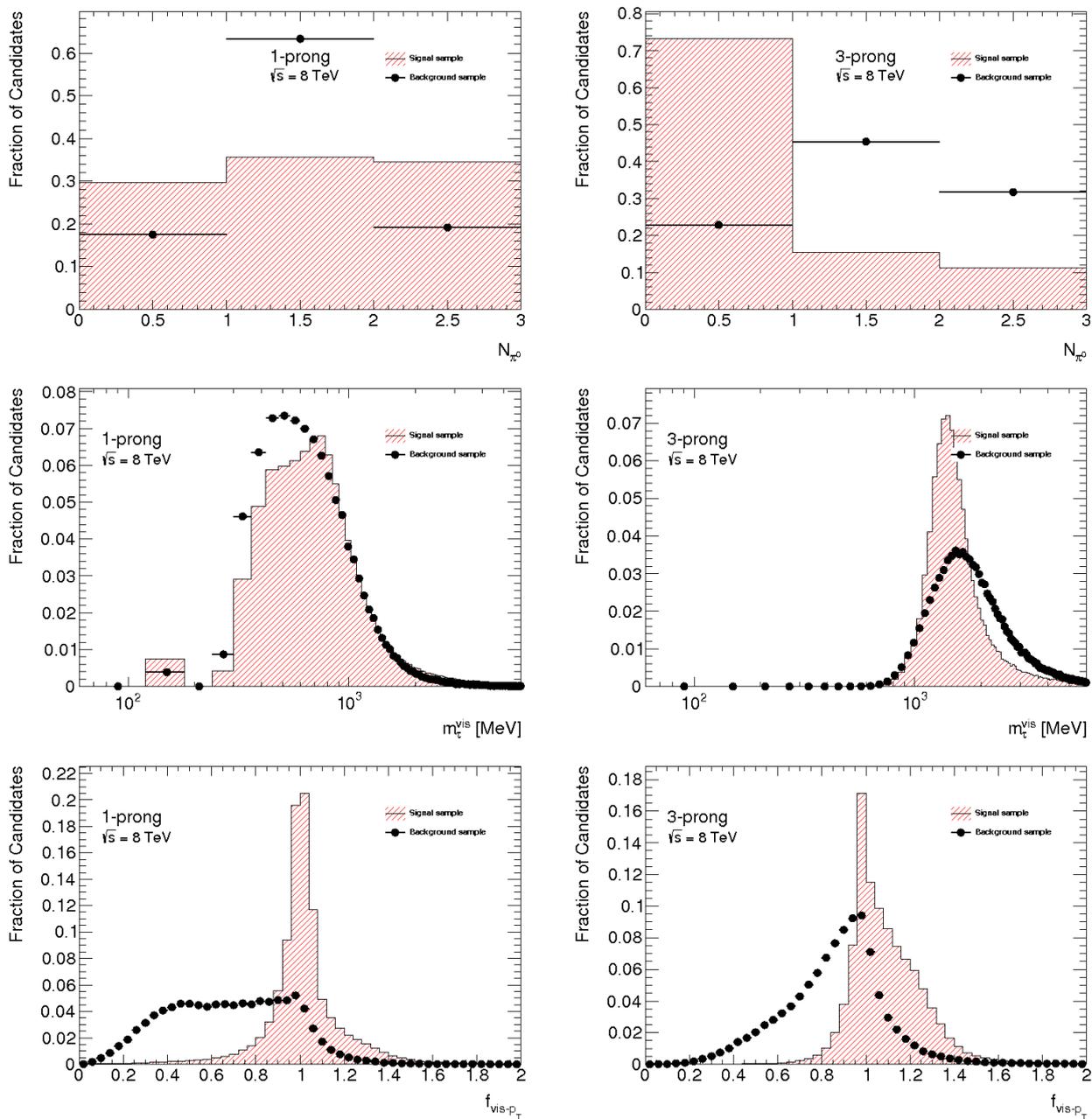


Abbildung 22: Verteilungen der Standardvariablen für 1-Prong- (links) und 3-Prong-Ereignisse (rechts).

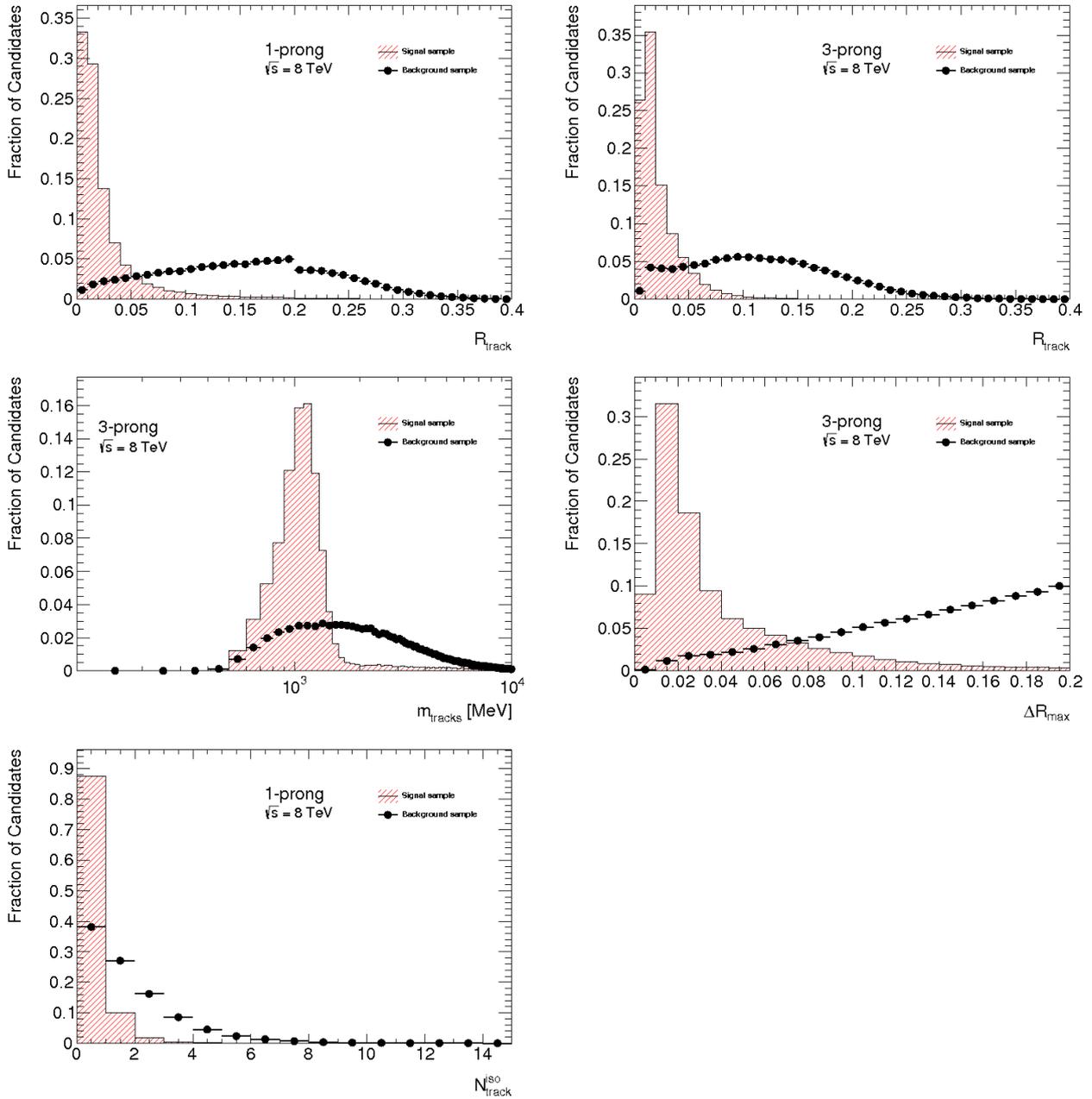


Abbildung 23: Verteilungen der Standardvariablen für 1-Prong- (links) und 3-Prong-Ereignisse (rechts).

A.2 Substrukturvariablen

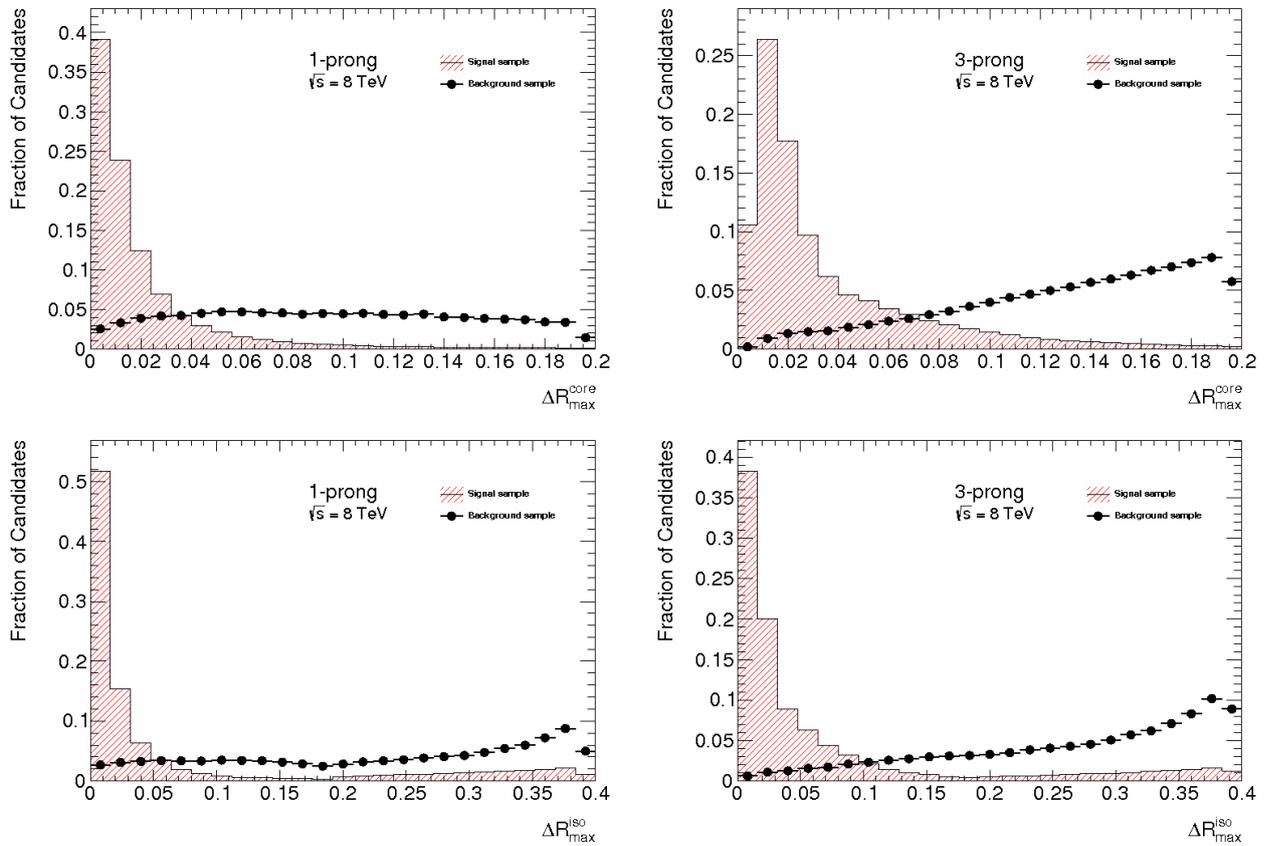


Abbildung 24: Verteilungen der Substrukturvariablen für 1-Prong-Ereignisse (links) und 3-Prong-Ereignisse (rechts).

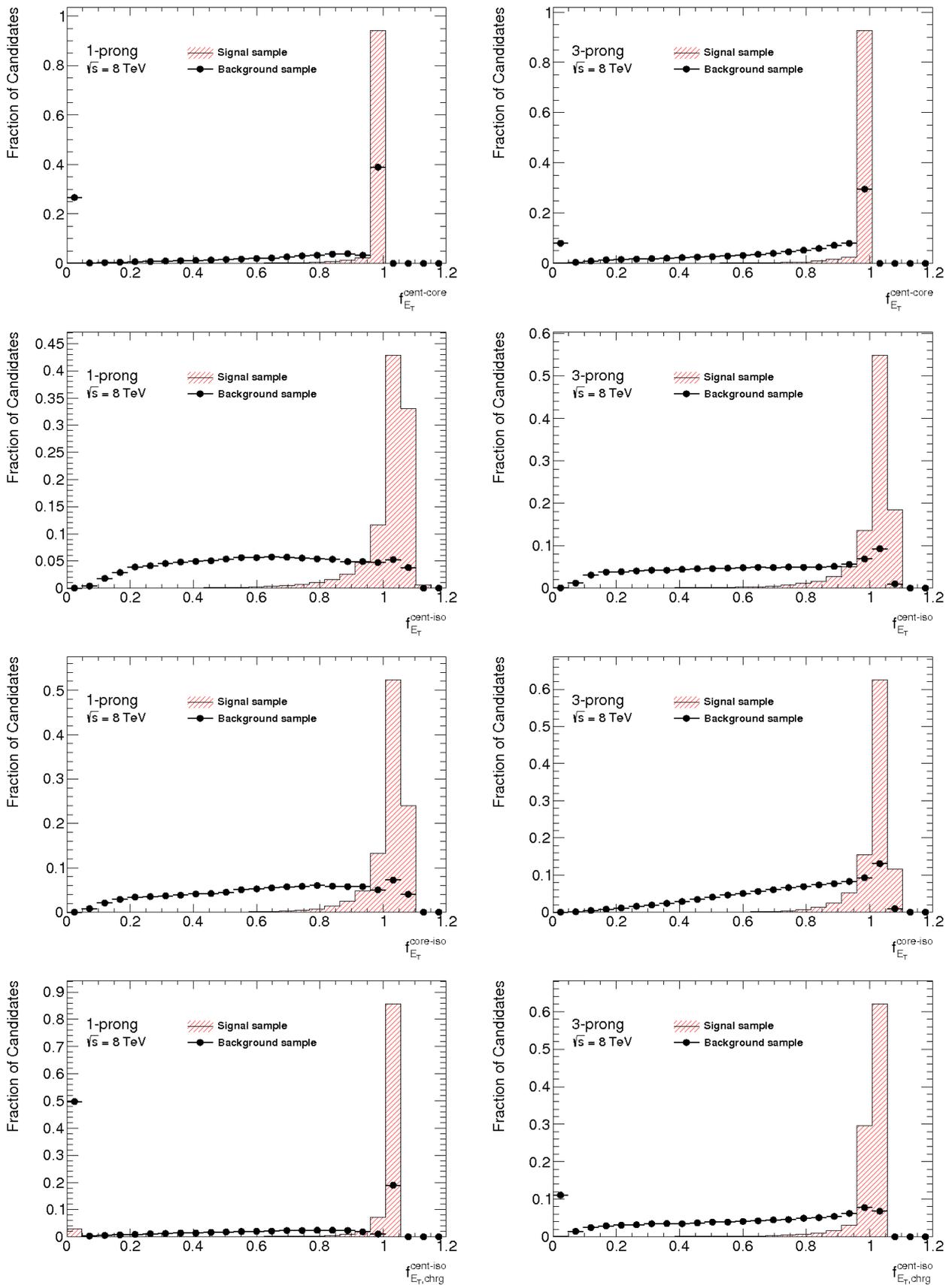


Abbildung 25: Verteilungen der Substrukturvariablen für 1-Prong-Ereignisse (links) und 3-Prong-Ereignisse (rechts).

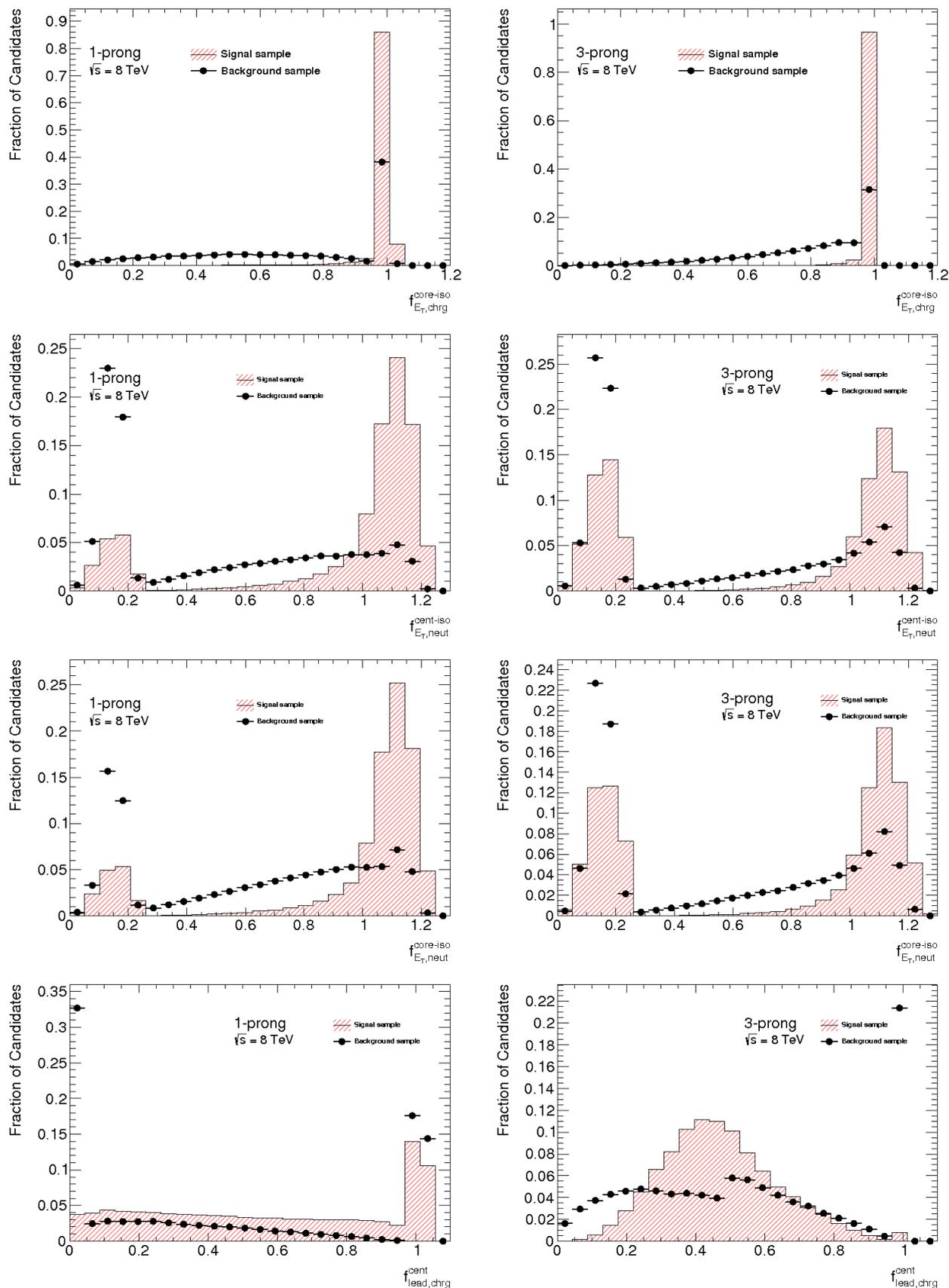


Abbildung 26: Verteilungen der Substrukturvariablen für 1-Prong-Ereignisse (links) und 3-Prong-Ereignisse (rechts).

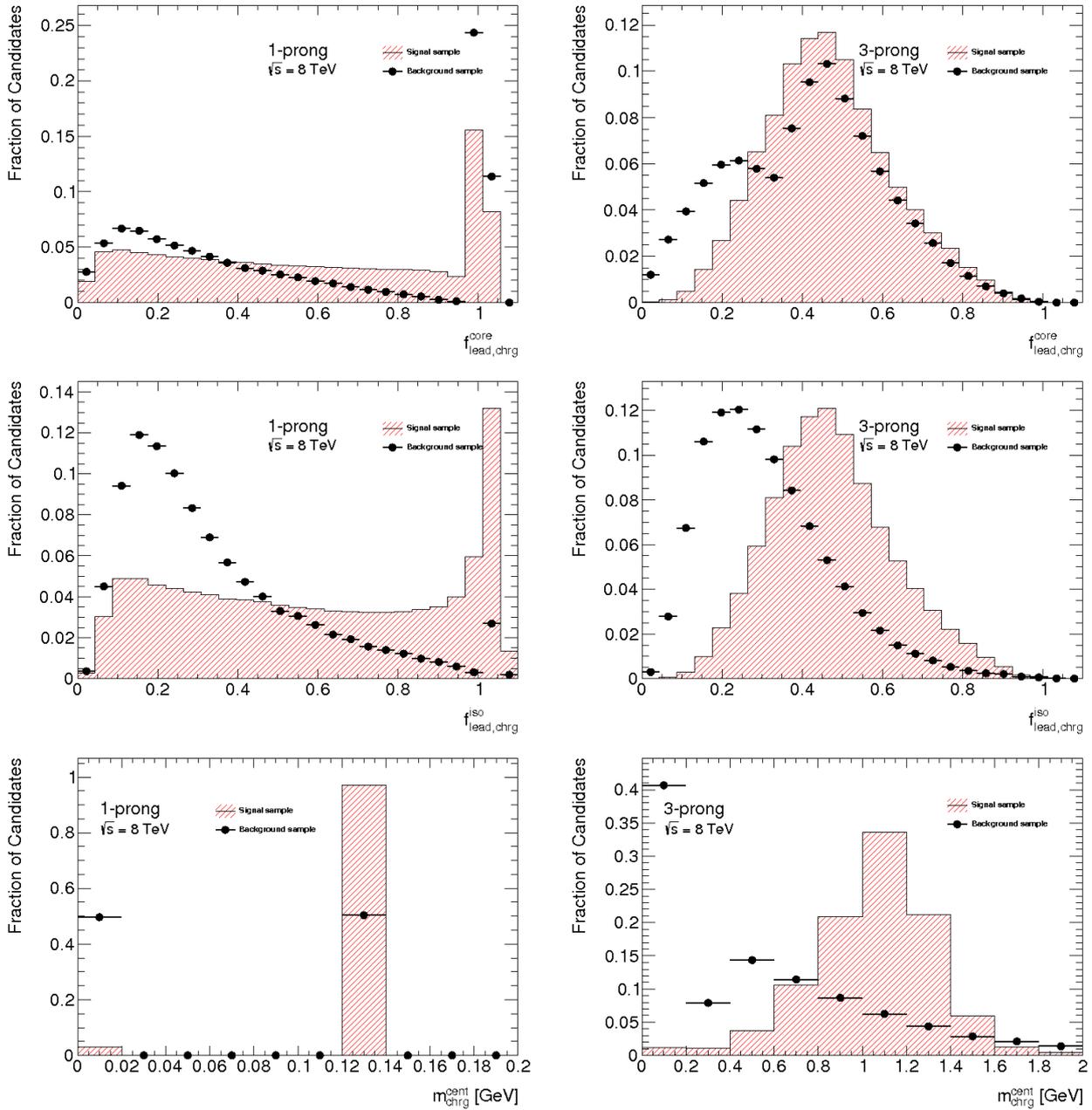


Abbildung 27: Verteilungen der Substrukturvariablen für 1-Prong-Ereignisse (links) und 3-Prong-Ereignisse (rechts).

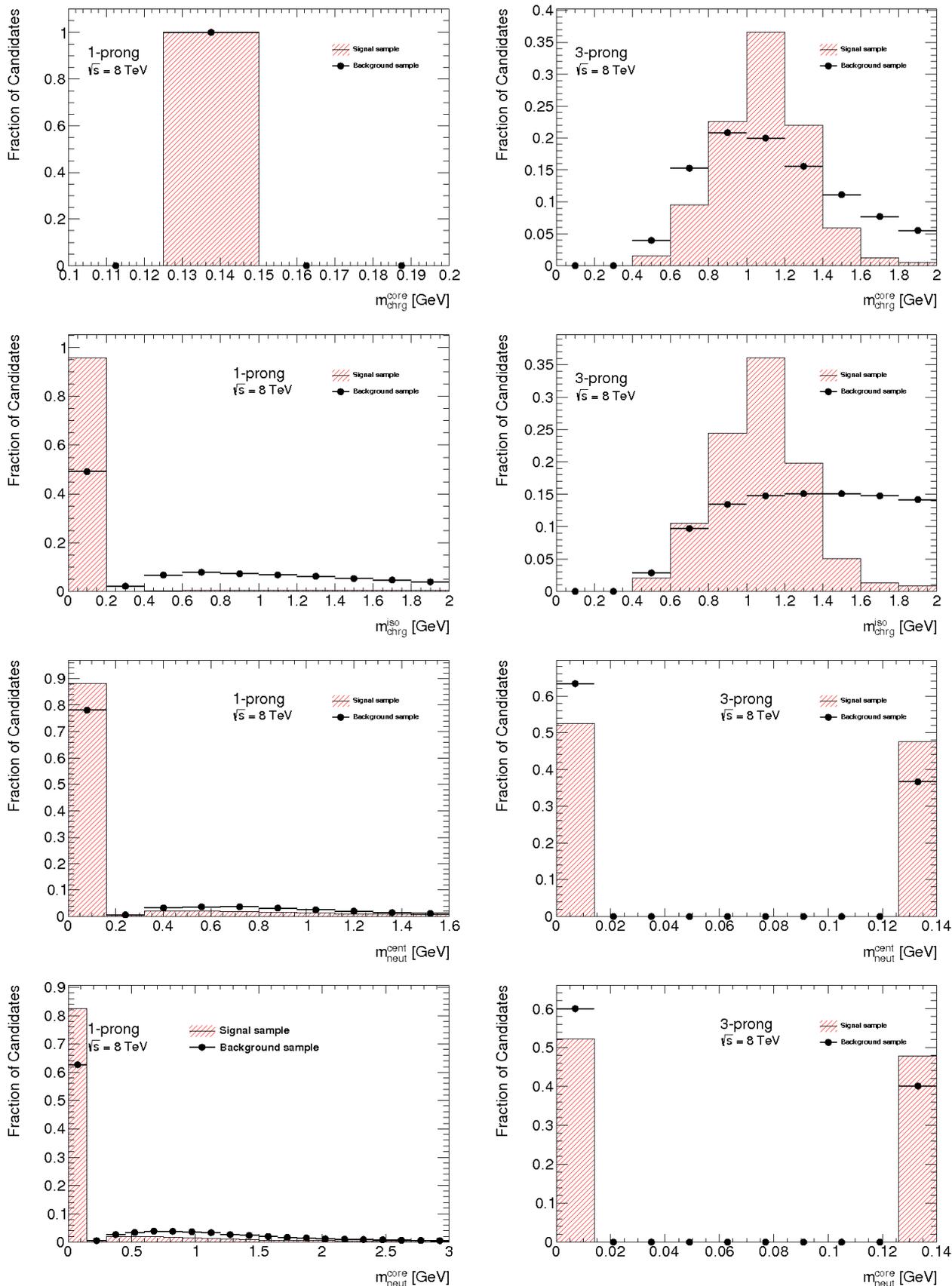


Abbildung 28: Verteilungen der Substrukturvariablen für 1-Prong-Ereignisse (links) und 3-Prong-Ereignisse (rechts).

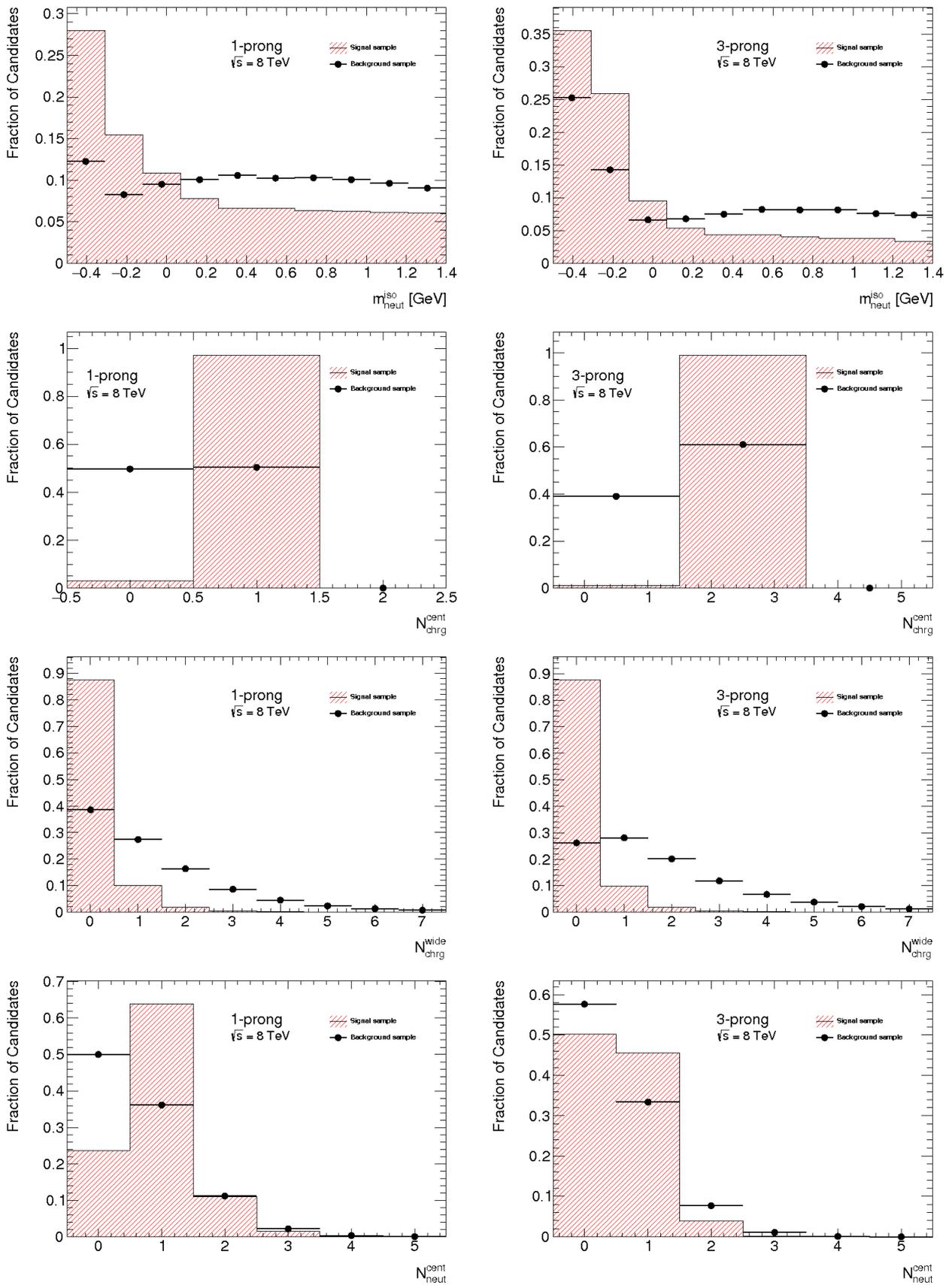


Abbildung 29: Verteilungen der Substrukturvariablen für 1-Prong-Ereignisse (links) und 3-Prong-Ereignisse (rechts).

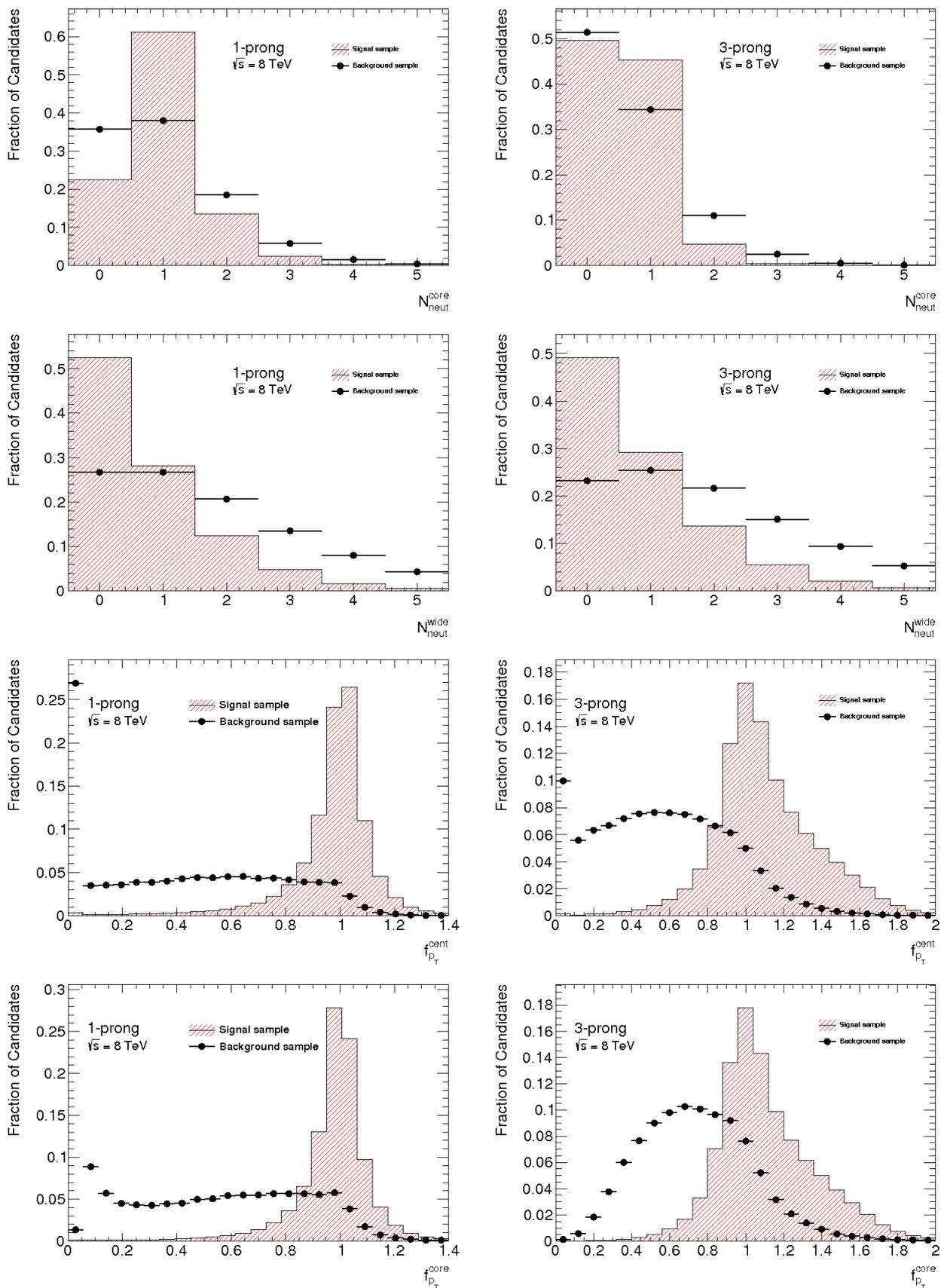


Abbildung 30: Verteilungen der Substrukturvariablen für 1-Prong-Ereignisse (links) und 3-Prong-Ereignisse (rechts).

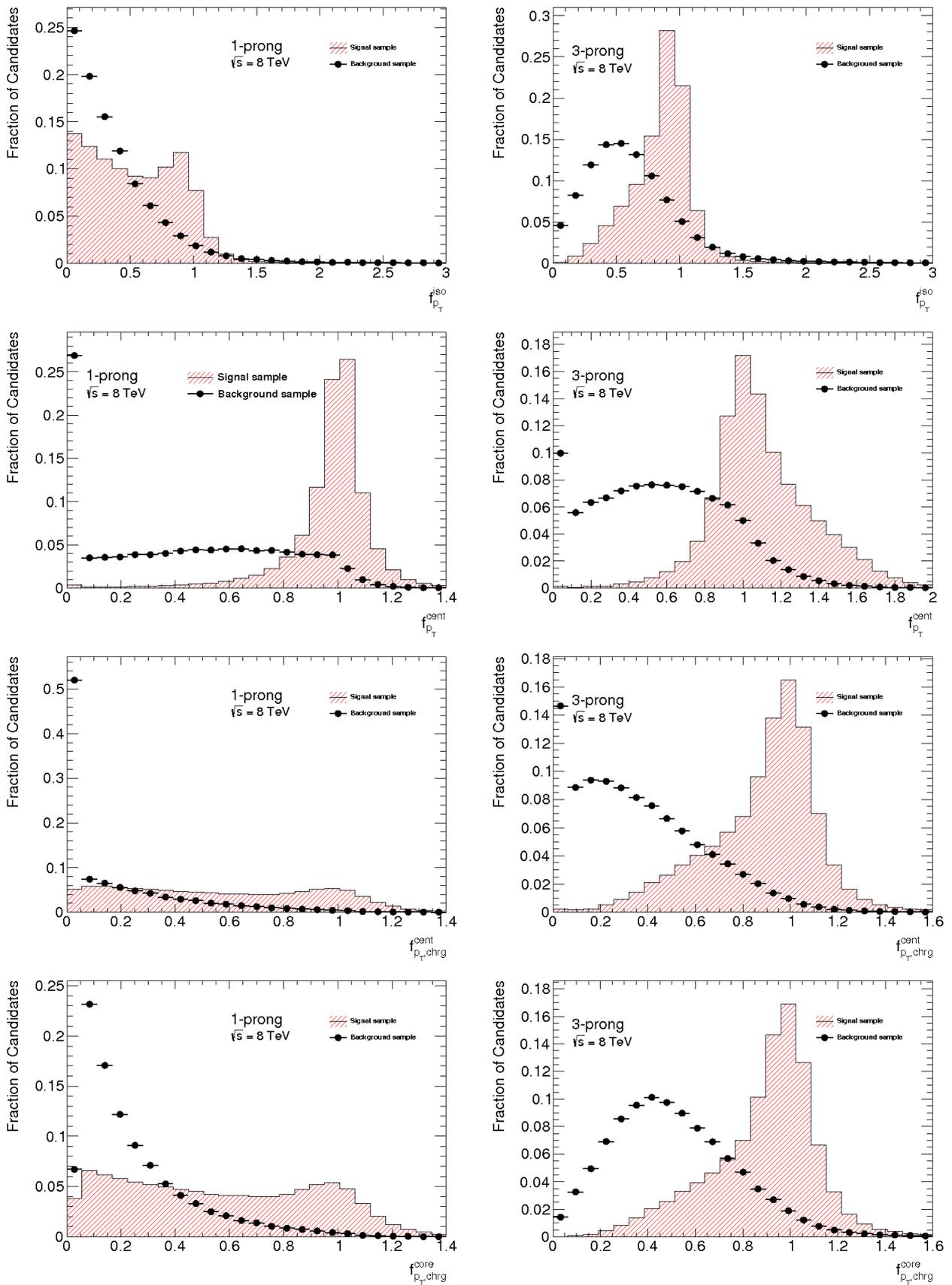


Abbildung 31: Verteilungen der Substrukturvariablen für 1-Prong-Ereignisse (links) und 3-Prong-Ereignisse (rechts).

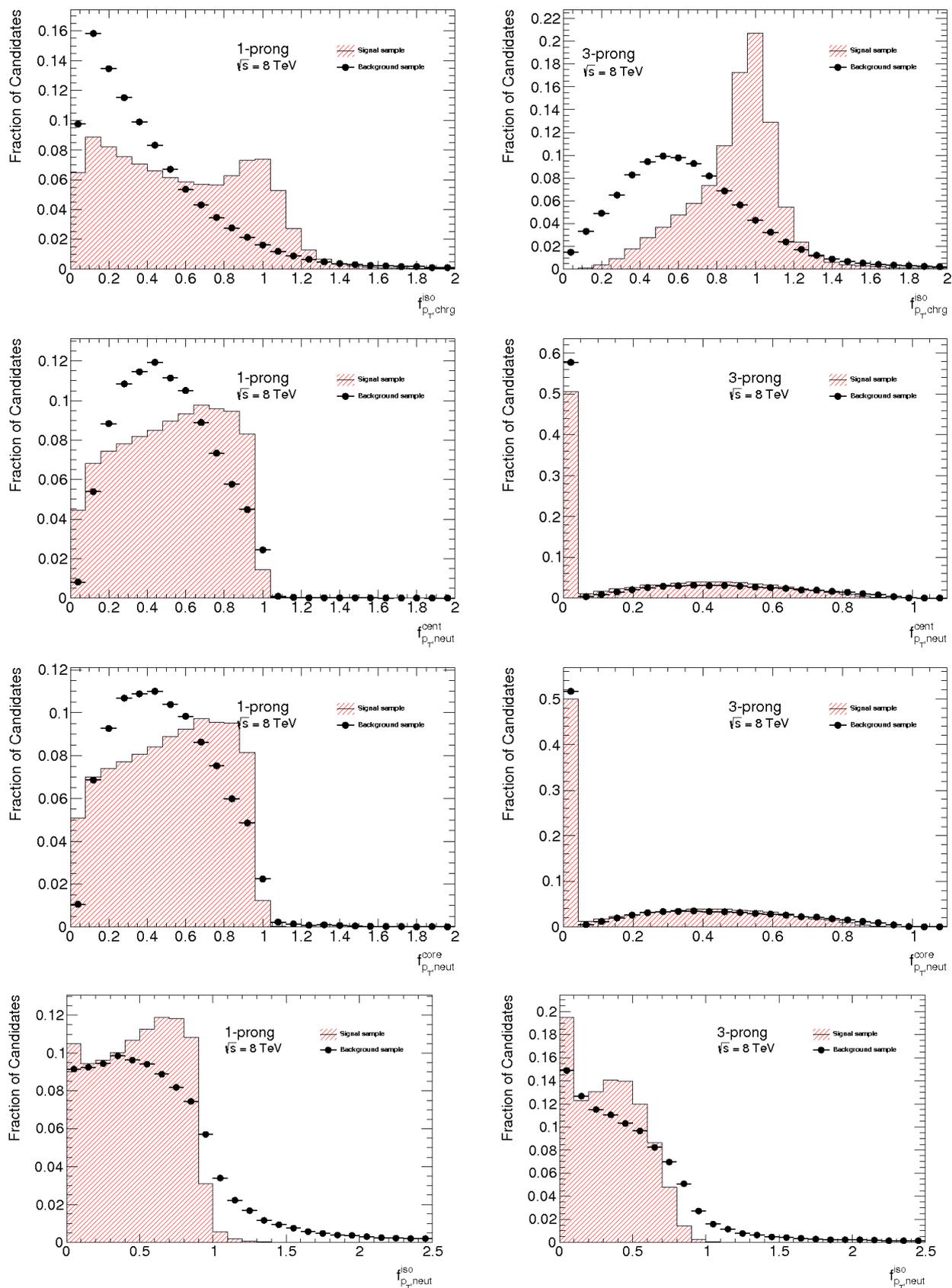


Abbildung 32: Verteilungen der Substrukturvariablen für 1-Prong-Ereignisse (links) und 3-Prong-Ereignisse (rechts).

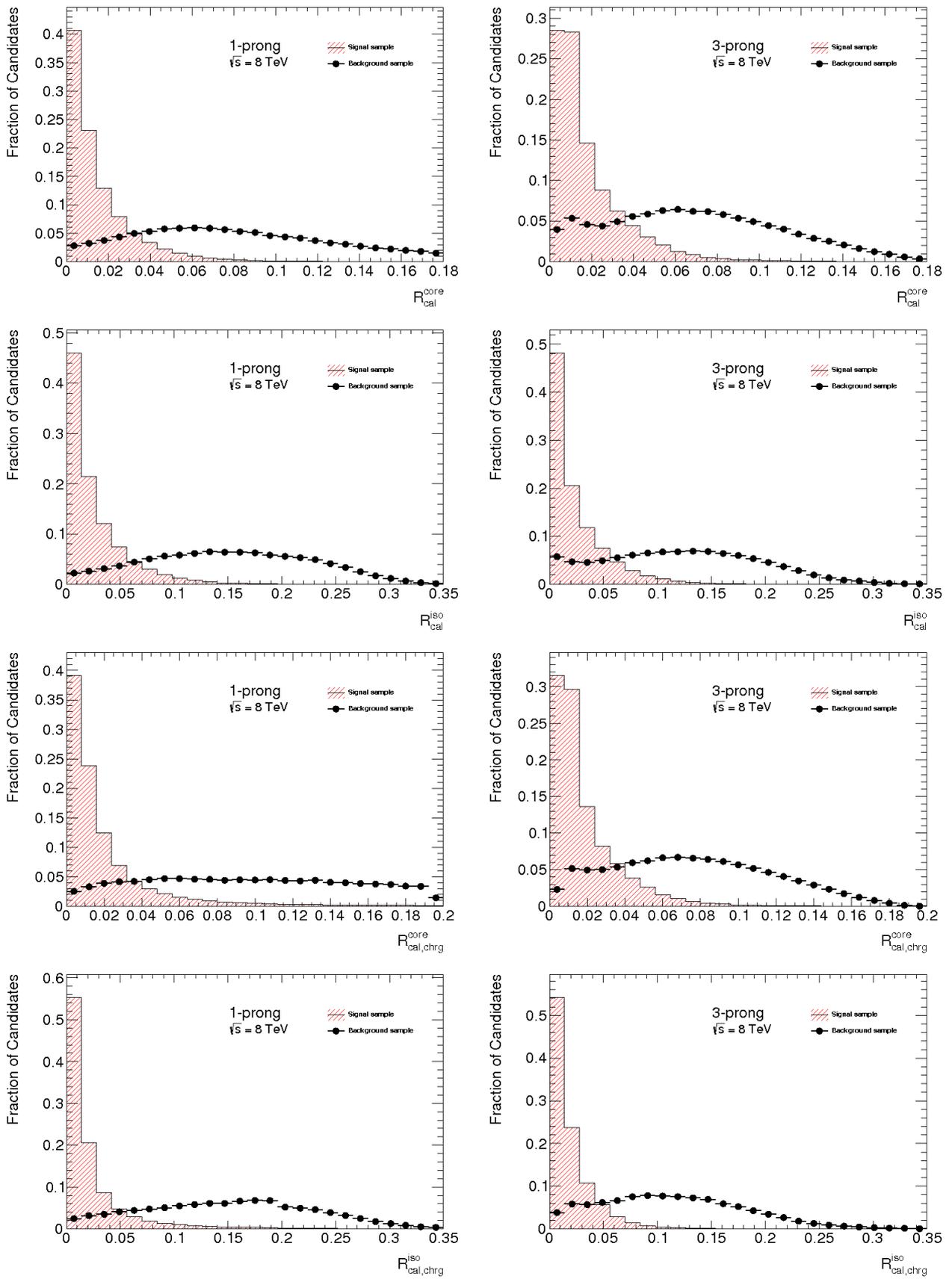


Abbildung 33: Verteilungen der Substrukturvariablen für 1-Prong-Ereignisse (links) und 3-Prong-Ereignisse (rechts).

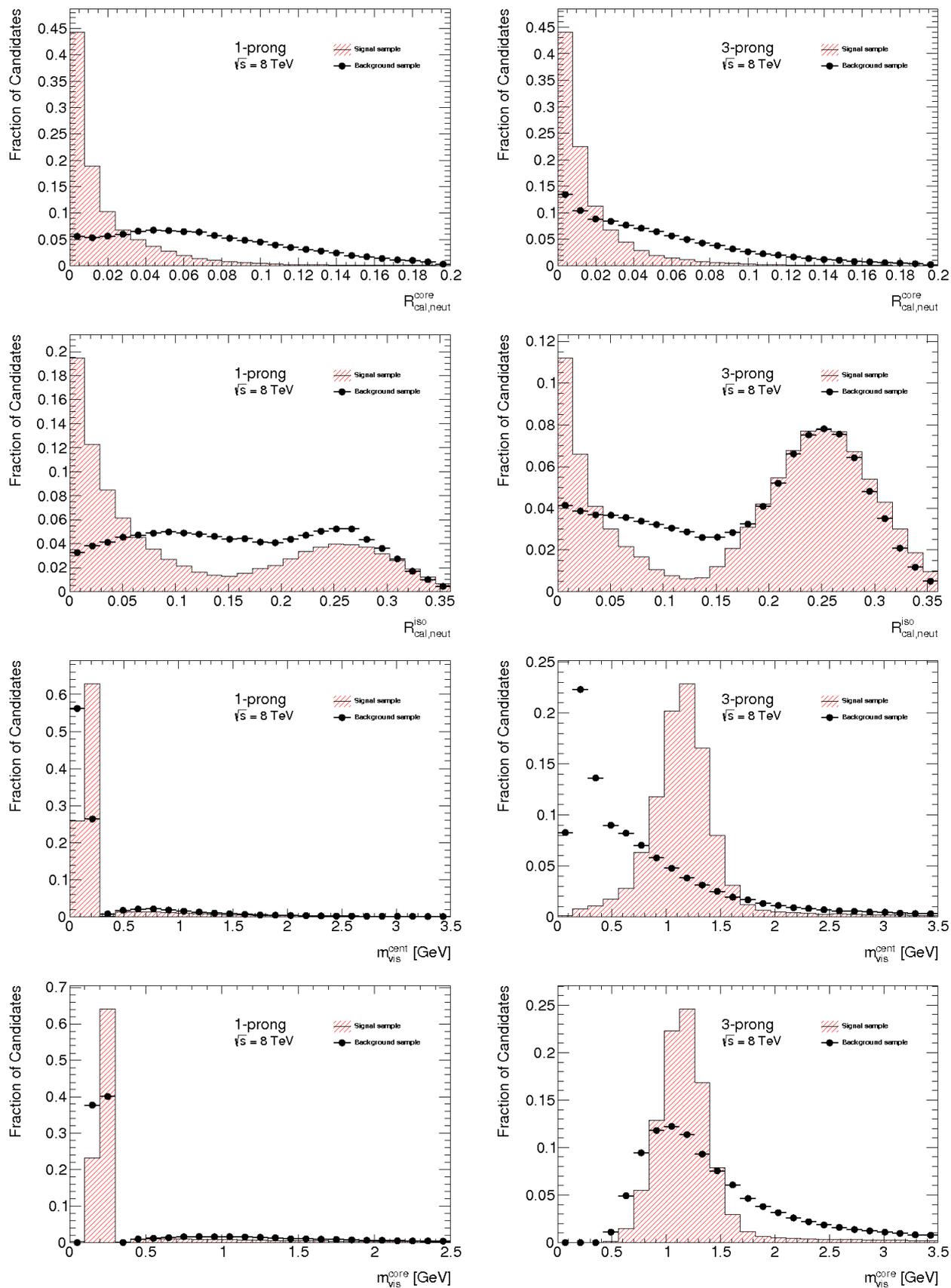


Abbildung 34: Verteilungen der Substrukturvariablen für 1-Prong-Ereignisse (links) und 3-Prong-Ereignisse (rechts).

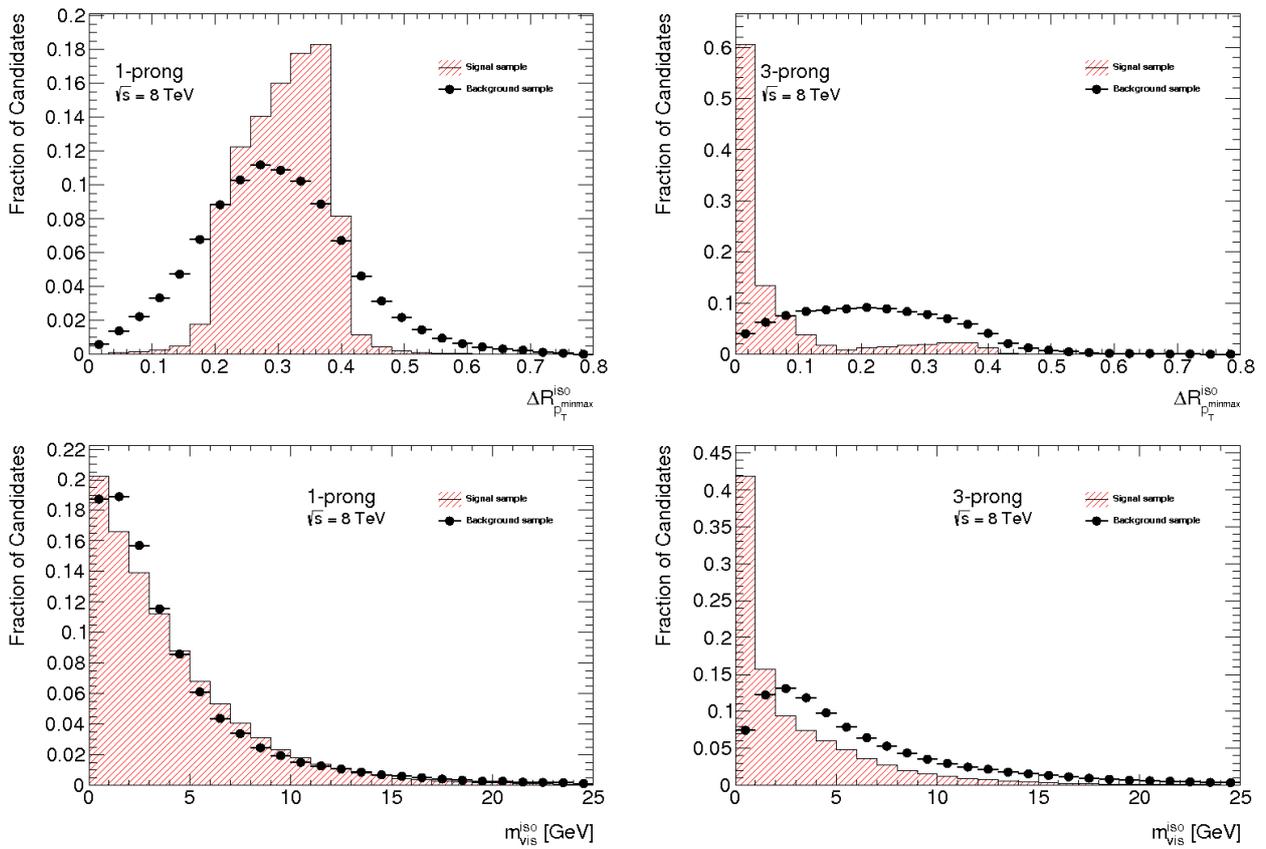


Abbildung 35: Verteilungen der Substrukturvariablen für 1-Prong-Ereignisse (links) und 3-Prong-Ereignisse (rechts).

B Konfigurationen

Type	MLP Konfiguration	Type	MLP Konfiguration
VarTransform	kMLP	VarTransform	kMLP
RandomSeed	N	RandomSeed	N
NCycles	[41, 42, 43, 44, 45]	RandomSeed	[1, 2, 3, 4, 5]
TrainingMethod	3000	NCycles	2000
Temperature	SARP	TrainingMethod	BFGS
HiddenLayers	0.02	HiddenLayers	20,20,20,20
NeuronType	20,20,20,20	NeuronType	tanh
EstimatorType	tanh	EstimatorType	CE
TestRate	CE	TestRate	50
	50		

Tabelle 9: Konfigurationen der zwei verwendeten Trainingskonfigurationen mit SARPROP (links) bzw. BFGS (rechts) als Trainingsalgorithmus

C Literaturhinweise

- [1] Wikipedia, https://de.wikipedia.org/wiki/Standardmodell#/media/File:Standard_Model_of_Elementary_Particles-de.svg (abgerufen: 20.01.2016, 15:18 Uhr)
- [2] Povh, Rith, Scholz, Zetsche; *Teilchen und Kerne - Eine Einführung in die physikalischen Konzepte*, 8. Auflage, Springer-Verlag (2009)
- [3] Particle Data Group Collaboration; *Review of Particle Physics*, Chin. Phys. C38 (2014) 090001
- [4] Madysa, Nico; *Identification of Hadronic Tau Lepton Decays at the ATLAS Detector Using Artificial Neural Networks*, Masterarbeit, TU Dresden 2015
- [5] Henseleit, Tony; *Nachweis hadronisch zerfallender Tau-Leptonen mit multivarianten Methoden im ATLAS-Detektor am LHC*, Bachelorarbeit, TU Dresden 2015
- [6] Thomson, Mark David; *Modern particle physics*, Cambridge University Press 2013
- [7] Hanisch, Stefanie; *Optimisation of the Hadronic Tau Identification Based on the Classification of Tau Decay Modes with the ATLAS Detector*, Diplomarbeit, TU Dresden, 2015
- [8] Treadgold, Gedeon; *Simulated Annealing and Weight Decay in Adaptive Learning: The SARPROP Algorithm*, IEEE Transactions on Neural Networks **9** (1998) URL: <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=701179&tag=1>
- [9] Kiesel, David; *Ein kleiner Überblick über Neuronale Netze*, URL: http://www.dkriesel.com/_media/science/neuronalenetze-de-zeta2-2col-dkrieselcom.pdf
- [10] Rey, Wender; *Neuronale Netze: eine Einführung in die Grundlagen, Anwendungen und Datenauswertung*, 2. Auflage, Bern: Huber (2011)
- [11] ATLAS Collaboration; G. Aad et al.; *The ATLAS Experiment at the CERN Large Hadron*

Collider, Journal of Instrumentation **3** (2008) S08003

[12] Glashow, S. L.: *Elementary Particle Theory*; Nucl. Phys. **22** (1961) 579-588

[13] Evans, Bryant; *LHC Machine*, JINST 3 (2008) S08001

[14] <https://www.lfi.rwth-aachen.de/uploads/images/knn.png> (Abgerufen am 21.1.2016, 13:18 Uhr)

[15] Fritzsche, H.; Gell-Mann, M.; Leutwyler, H., *Advantages of the color octet gluon picture*, Phys. Lett. **B 47** (1973) 365

[16] Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321-323;

[17] ATLAS Collaboration, G. Aad et al., *Observation of a new particle in the search for the Standard Model Higgs Boson with the ATLAS detector at the LHC*, Phys. Lett. **B716** (2012) 1-29

[18] CMS Collaboration, S. Chatrchyan et al., *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, **B716** (2012) 30-61

[19] Hinton, G. E.; Salakhutdinov, R. R.; *Reducing the Dimensionality of Data with Neural Networks*, Science (28.7.2006)

[20] Morales, J.; *A numerical Study of limited memory BFGS method*, Applied Mathematics Letters **15** (2002) no. 4

[21] Hansen, L. K.; Salamon, P.; *Neural Network Ensembles*, IEEE Transactions on Pattern Analysis and Machine Intelligence **12** (1990) no. 10

[22] CMS Collaboration, S. Chatrchyan et al., *The CMS experiment at the CERN LHC*, Journal of Instrumentation **3** (2008) no. 8

[23] LHCb Collaboration, A. A. Alves J. et al., *The LHCb detector at the LHC*, Journal of Instrumentation **3** (2008) no. 8

[24] ALICE Collaboration, K. Aamodt et al., *The ALICE experiment at the CERN LHC*,

Abbildungsverzeichnis

1	Standardmodell	1
2	Schema Künstliches Neuronales Netz	7
3	Querschnitt des ATLAS-Detektors	10
4	Verteilung der Variable N_{neut}^{core}	13
5	Verteilung der Variable $R_{cal,chg}^{iso}$	13
6	ROC-Kurve des BDT und MLP mit SARPROP	15
7	ROC-Kurven des BDT und des MLP mit BFGS	15
8	ROC-Kurve der MLP durch Weglassen der Standardvariablen	17
9	ROC-Kurven der MLPs durch Ersetzen von N_{π^0}	18
10	ROC-Kurven der MLPs mit ΔR_{max}^{core}	19
11	ROC-Kurven der vollständig mit Substrukturvariablen trainierten MLPs	20
12	ROC-Kurven der MLPs bei Austausch aller Variablen für 1-Prong-Ereignisse	21
13	Korrelationsplots des Signals	22
14	Korrelationsplots des Hintergrunds	23
15	ROC-Kurven des MLP's bei Ersetzen aller Variablen	23
16	ROC-Kurve der MLPs bei Hinzufügen einer Substrukturvariable	24
17	ROC-Kurve im Vergleich von BDT und MLP für die besten Ergebnisse	25
18	ROC-Kurve der BDT und MLP bei ersetzen einer Standardvariable	26
19	ROC-Kurve des MLP's ohne diskrete Variable	26
20	ROC-Kurve der sechs Ensembles	28
21	Verteilung der Standardvariablen I	30
22	Verteilung der Standardvariablen II	31
23	Verteilung der Standardvariablen III	32
24	Verteilungen der Substrukturvariablen I	33
25	Verteilungen der Substrukturvariablen I	34
26	Verteilungen der Substrukturvariablen III	35
27	Verteilungen der Substrukturvariablen IV	36
28	Verteilungen der Substrukturvariablen V	37
29	Verteilungen der Substrukturvariablen VI	38

30	Verteilungen der Substrukturvariablen VII	39
31	Verteilungen der Substrukturvariablen VIII	40
32	Verteilungen der Substrukturvariablen IX	41
33	Verteilungen der Substrukturvariablen X	42
34	Verteilungen der Substrukturvariablen XI	43
35	Verteilungen der Substrukturvariablen XI	44

Tabellenverzeichnis

1 Zerfälle der Tau-Leptonen	3
2 Variablenkonfiguration zur Zuordnung der MLPs	16
3 Variablenkonfiguration zur Zuordnung der MLPs mit AUCs	17
4 Variablenkonfiguration der MLPs mit ΔR_{max}^{core}	18
5 Variablenkonfiguration der vollständig ersetzten Variablensätze	20
6 Variablenkonfiguration der am stärksten eingestuften Variablen	22
7 Variablenkonfiguration bei Hinzufügen einer Substrukturvariable	24
8 Mittelwert und Standardabweichung der Ensembles	27
9 Konfiguration der neuronalen Netze	45

Erklärung

Hiermit erkläre ich, Anne Kristin Dietert, 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.

Anne Kristin Dietert
Dresden, März 2016