



**Universiteit  
Antwerpen**

# Studie van het Higgsboson verval naar een charm quark-antiquark paar aan de hand van CMS Run 3 MC data

Faculteit Wetenschappen  
Departement Fysica  
Academiejaar 2024-2025

Proefschrift ter verkrijging van de graad van Master in de Fysica.

**Auteur:** Abigail Van Asch  
**Promotor:** Prof. Dr. Pierre Van Mechelen  
**Begeleider:** Dr. Tahys Janssen

18 augustus 2025

# Samenvatting

Het Higgsboson is een fundamenteel deeltje dat ontdekt is door de ATLAS- en CMS-experimenten in 2012. [1] De interacties van het Higgsboson met fermionen waaronder bottom quarks, top quarks, tau leptonen, en muonen zijn reeds waargenomen en opgemeten. [2] Een volgende stap is het onderzoeken van de sterkte van de interactie tussen het Higgsboson en charm quarks. Dit is echter een uitdaging. Het verval van het Higgsboson naar een charm quark-antiquarkpaar heeft namelijk een kleine waarschijnlijkheid, en daarbij zijn jets afkomstig van charm quarks ook moeilijk te identificeren en te onderscheiden van jets geproduceerd door bottom quarks. [3]

Deze thesis bestudeert het verval van het Higgsboson naar een charm quark-antiquarkpaar, waarbij het Higgsboson geproduceerd wordt in associatie met een Z-boson. Specifiek wordt er gekeken naar het proces waarbij het Z-boson vervalt naar een paar muonen of elektronen en het Higgsboson opgebouwd wordt uit twee goed afgezonderde jets. Voor deze studie wordt er gebruikgemaakt van CMS Run 3 2022postEE data bij een centre-of-mass energie van  $\sqrt{s} = 13,6$  TeV. Dit komt overeen met een geïntegreerde luminositeit van  $26,7 \text{ fb}^{-1}$ .

Om het te onderzoeken signaal te onderscheiden van de achtergrond worden enkele methodes toegepast. Ten eerste worden er selecties uitgevoerd op enkele observabelen. Deze selecties zorgen ervoor dat regio's met veel achtergrond en weinig signaal onderdrukt worden. Vervolgens wordt er onderzocht of de resolutie van de invariante massapijk van de jets, die het Higgsboson opbouwen, verbeterd kan worden. Zowel de invloed van jet energieregressie als kinematic fitting worden bestudeerd. Deze technieken trachten de waargenomen parameters van de jets (bijvoorbeeld transversale impuls) te corrigeren richting de werkelijke waarden. Jet energieregressie zorgt voor een duidelijke verbetering van de resolutie en het gemiddelde van de invariante massapijk van de jets. Kinematic fitting geeft enkel een verbetering van het gemiddelde wanneer het wordt toegepast samen met jet energieregressie. Bijkomend worden er ook enkele controle regio's bepaald waarbij het signaal zo goed als afwezig is. Dit stelt een beperking op de achtergronden in de finale analyse. Verder wordt een boosted decision tree (BDT) getraind. De getrainde BDT heeft een AUC van 0,91 voor zowel het muon- als elektronvervalkanaal en kan het signaal dus goed onderscheiden van de achtergrond aan de hand van de gekozen inputvariabelen. In de laatste stap wordt er een statistische analyse gedaan om een verwachte bovenlimiet op de signaalsterkte te bekomen. Dit wordt gedaan met behulp van het software framework Combine [4]. De voornaamste systematische onzekerheden worden hierbij in rekening gebracht. Eén van de belangrijkste systematische onzekerheden die nog niet is toegevoegd, is de onzekerheid op de c-tagging efficiënties.

Er wordt een blinde analyse uitgevoerd voor de twee onafhankelijke vervalkanalen, namelijk het verval naar muonen en elektronen. De resultaten worden gecombineerd voor het finale resultaat. Er wordt een verwachte bovenlimiet gevonden van  $81_{-26}^{+44}$  voor de signaalsterkte bij een betrouwbaarheidsinterval van 95 %.

# Abstract

The Higgs boson is a fundamental particle that was discovered by the ATLAS and CMS experiments in 2012. [1] The interactions of the Higgs boson with fermions, including bottom quarks, top quarks, tau leptons, and muons, have already been observed and measured. [2] The next step is to investigate the strength of the interaction between the Higgs boson and charm quarks. However, this is a challenge, since the decay of the Higgs boson into a charm quark-antiquark pair has a low probability, and jets originating from charm quarks are also difficult to identify and distinguish from jets produced by bottom quarks. [3]

This thesis studies the decay of the Higgs boson into a charm quark-antiquark pair where the Higgs boson is produced in association with a Z boson. Specifically, this study looks at the process in which the Z boson decays into a pair of muons or electrons and the Higgs boson is constructed from two well separated jets. For this CMS Run 3 2022postEE data at a centre-of-mass energy of  $\sqrt{s} = 13.6$  TeV is used. This corresponds to an integrated luminosity of  $26.7 \text{ fb}^{-1}$ .

Several methods are applied to extract the investigated signal from the background. Firstly, selections are made on several observables. These selections will suppress regions with high background yields and low signal yields. Next, it is investigated whether the resolution of the invariant mass peak of the jets that build up the Higgs boson can be improved. Both the effect of jet energy regression and kinematic fitting are studied. These techniques attempt to correct the observed parameters of the jets (for example transverse momentum) towards the true values. Jet energy regression provides a clear improvement in the resolution and average of the invariant mass peak of the jets. Kinematic fitting only improves the average when applied together with jet energy regression. In addition, several control regions are also determined where the signal is almost absent. This gives a constraint on the backgrounds in the final analysis. Furthermore, a boosted decision tree (BDT) is trained. The trained BDT has an AUC of 0.91 for both the muon and electron decay channels and can therefore distinguish the signal well from the background based on the chosen input variables. In the final step, a statistical analysis is performed to obtain an expected upper limit on the signal strength. This is done by using the Combine software framework [4]. The main systematic uncertainties are taken into account. One of the most important systematic uncertainties that has not yet been added, is the uncertainty on the c-tagging efficiencies.

A blind analysis is performed for the two independent decay channels, namely the decay to muons and electrons. The results are combined for the final result. An expected upper limit of  $81_{-26}^{+44}$  is found for the signal strength at a confidence level of 95 %.

# Inhoudsopgave

<b>1</b>	<b>Inleiding</b>	<b>1</b>
<b>2</b>	<b>Theoretische achtergrond</b>	<b>2</b>
2.1	Het standaardmodel . . . . .	2
2.2	Het Brout-Englert-Higgs mechanisme . . . . .	3
2.3	Higgs processen . . . . .	7
<b>3</b>	<b>Experimentele opstelling</b>	<b>11</b>
3.1	LHC . . . . .	11
3.2	CMS . . . . .	11
3.3	Event reconstructie . . . . .	14
<b>4</b>	<b>Methodologie</b>	<b>16</b>
4.1	Achtergrond en datasets . . . . .	17
4.2	Object en event selecties . . . . .	17
4.2.1	Plots van observabelen . . . . .	20
4.3	Controle regio's . . . . .	30
4.4	Resolutie massapieken . . . . .	33
4.4.1	Jet energieregressie . . . . .	33
4.4.2	Kinematic fitting . . . . .	33
4.5	Boosted decision tree . . . . .	37
4.6	Statistische analyse . . . . .	41
4.7	Systematische onzekerheden . . . . .	42
<b>5</b>	<b>Resultaten</b>	<b>44</b>
<b>6</b>	<b>Conclusie en vooruitblik</b>	<b>45</b>
<b>7</b>	<b>Dankwoord</b>	<b>47</b>
<b>A</b>	<b>Plots van observabelen elektronkanaal</b>	<b>48</b>
<b>B</b>	<b>Controle regio's elektronkanaal</b>	<b>58</b>
<b>C</b>	<b>BDT elektronkanaal</b>	<b>60</b>

# 1 Inleiding

Het Higgsboson is een fundamenteel deeltje dat ontdekt is door de ATLAS- en CMS-experimenten in 2012. Het is van groot belang om alle eigenschappen van dit deeltje waar te nemen en op deze manier de voorspellingen gemaakt door het standaardmodel experimenteel te bevestigen en het model verder te testen. [1]

Fermionen verkrijgen hun massa doordat ze interageren met het Higgsboson. De interactie van het Higgsboson met bottom en top quarks, tau leptonen, en muonen zijn reeds waargenomen en opgemeten. [2] Een volgende stap is het onderzoeken van de sterkte van de interactie tussen het Higgsboson en tweede generatie quarks, waaronder charm quarks. Dit is echter een uitdaging. Dit komt doordat het verval van het Higgsboson naar een charm quark-antiquarkpaar een kleine waarschijnlijkheid heeft, en daarbij zijn charm jets ook moeilijk te identificeren en te onderscheiden van bottom jets. [3]

In deze thesis wordt het verval van het Higgsboson naar een charm quark-antiquarkpaar bestudeerd, waarbij het Higgsboson geproduceerd is in associatie met een Z-boson. Specifiek wordt er gekeken naar het proces waarbij het Z-boson vervalt naar twee muonen of elektronen en waarbij het Higgsboson opgebouwd wordt uit twee goed afgezonderde jets. Er wordt CMS Run 3 2022postEE data gebruikt bij een centre-of-mass energie van  $\sqrt{s} = 13,6$  TeV dat overeenkomt met een geïntegreerde luminositeit van  $26,7 \text{ fb}^{-1}$ .

In sectie 2 wordt de theoretische achtergrond gegeven. Onder meer wordt het standaardmodel beschreven en ook de voorspellingen ervan relevant voor het te onderzoeken Higgs proces. In sectie 3 wordt er meer informatie gegeven over het CMS-experiment waarvan de data afkomstig is voor deze studie. Er wordt beschreven hoe CMS botsingen observeert aan de hand van de verschillende subdetectoren en hoe een botsing wordt gereconstrueerd. In sectie 4 wordt een methodologie gegeven voor de analyse. Om achtergrond te onderdrukken en het signaal te onderscheiden, worden enkele technieken toegepast: er worden selecties op enkele waargenomen variabelen gemaakt, de invloed van jet energieregressie en kinematic fitting worden onderzocht, en ook een boosted decision tree wordt getraind. In sectie 5 worden ten slotte de resultaten van de statistische analyse gevonden, waarbij een bovenlimiet op de signaalsterkte wordt bepaald.

## 2 Theoretische achtergrond

### 2.1 Het standaardmodel

Het standaardmodel (SM) is een beschrijving van de fundamentele deeltjes en hun interacties, zoals weergegeven wordt in figuur 1. De deeltjes interageren met elkaar via vier fundamentele krachten. Deze vier fundamentele krachten in het standaardmodel zijn de elektromagnetische kracht, de sterke wisselwerking, de zwakke wisselwerking en de zwaartekracht. Krachten tussen de deeltjes worden beschreven door de uitwisseling van een ijkdeeltje. [1]

De materie in ons universum wordt opgebouwd uit fermionen. Dit zijn spin-halfdeeltjes. De twaalf fermionen worden opgedeeld in quarks en leptonen. Quarks hebben zowel een elektrische als een kleur-lading. Leptonen zijn kleurneutrale fermionen en kunnen opnieuw opgedeeld worden op basis van hun elektrische lading. Elektrisch neutrale leptonen worden neutrino's genoemd. Voor elk van de twaalf fermionen bestaat er ook een anti-deeltje dat dezelfde massa maar tegengestelde ladingen heeft. De fermionen worden ook opgedeeld in drie generaties. Bij elke generatie horen twee quarks, een geladen lepton en een neutrino. [1]

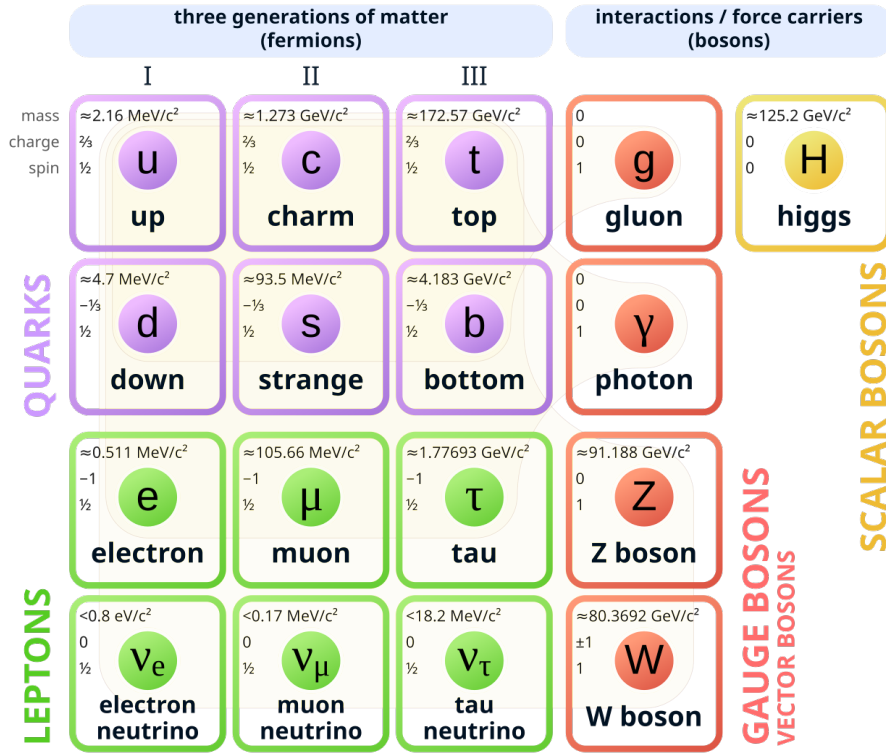
Alle deeltjes voelen de zwaartekracht en de zwakke wisselwerking. Alle elektrisch geladen deeltjes interageren met de elektromagnetische wisselwerking. Ten slotte voelen alle kleurgeladen deeltjes de sterke wisselwerking. Quarks zijn de fermionen die wel deelnemen aan de sterke wisselwerking en worden enkel waargenomen in gebonden toestanden van quarks die kleurneutraal zijn. Deze gebonden toestanden worden hadronen genoemd. Dit betekent dat quarks niet als vrije deeltjes propageren en hadroniseren. [1]

De zwakke, de sterke en de elektromagnetische krachten worden beschreven door quantum field theory (QFT) als de uitwisseling van een spin-1 ijkdeeltje. Elektromagnetisme wordt beschreven door Quantum Electrodynamics (QED) waarbij het foton wordt uitgewisseld. De sterke wisselwerking wordt beschreven door Quantum Chromodynamics (QCD), hierbij wordt een gluon uitgewisseld. Het foton en gluon zijn beide massaloos. De zwakke wisselwerking wordt opgedeeld in twee soorten interacties. De geladen stroom wisselwerking gebeurt door de uitwisseling van een massief W-boson en de neutrale stroom wisselwerking gebeurt door de uitwisseling van een massief en elektrisch neutraal Z-boson. [1]

Het standaardmodel wordt aangevuld door het Higgsboson. Dit deeltje is ontdekt door de ATLAS- en CMS-experimenten in 2012. Het Higgsboson wordt verwezenlijkt door het Higgsveld en dit veld heeft een vacuümverwachtingswaarde dat niet gelijk is aan nul. Alle fundamentele deeltjes zijn initieel massaloos en door hun interactie met het Higgsveld kunnen ze massa verwerven. [1]

Het Higgsboson is een spin-0 deeltje [1] en heeft een massa van 125,35 GeV [5] en een levensduur van  $2,1 \times 10^{-22}$  s [6].

# Standard Model of Elementary Particles



Figuur 1: Schematisch overzicht van het standaardmodel [7]

## 2.2 Het Brout-Englert-Higgs mechanisme

In het standaardmodel verkrijgen zowel de elektrozwakke ijkbosonen, als de fundamentele fermionen hun massa aan de hand van het Higgs mechanisme. Dit gebeurt op een manier dat de lokale ijk-symmetrie spontaan gebroken wordt en de theorie renormaliseerbaar blijft. Het Higgs mechanisme wordt ingebed in de elektrozwakke theorie. De elektrozwakke sector van het standaardmodel wordt verkregen door het vereisen van een  $U(1)_Y \times SU(2)_L$  lokale ijk-symmetrie. [1] [8]

Het concept van het Higgs mechanisme wordt eerst aan de hand van leptonen uitgelegd als voorbeeld, later worden de resultaten voor charm quarks neergeschreven. De vrije Lagrangiaanse dichtheid voor leptonen, waarbij de massaterm later nog toegevoegd moet worden, wordt gegeven door: [8]

$$\mathcal{L}_0 = i[\bar{\psi}_l(x)\gamma^\mu\partial_\mu\psi_l(x) + \bar{\psi}_{\nu_l}(x)\gamma^\mu\partial_\mu\psi_{\nu_l}(x)] \quad (1)$$

De linkshandige en rechtshandige leptonische velden worden gedefinieerd als volgt: [8]

$$\begin{cases} \psi_L = P_L\psi \\ \psi_R = P_R\psi \end{cases} \quad (2)$$

Hierbij geldt er: [8]

$$\begin{cases} P_L = \frac{1}{2}(\mathbf{1} - \gamma_5) \\ P_R = \frac{1}{2}(\mathbf{1} + \gamma_5) \end{cases} \quad (3)$$

De linkshandige leptonische velden kunnen samengevoegd worden in een doublet en de rechtshandige velden worden elk in een singlet geplaatst. [8]

$$\Psi_l^L = \begin{pmatrix} \psi_{\nu_l}^L \\ \psi_l^L \end{pmatrix}, \quad \psi_l^R, \quad \psi_{\nu_l}^R \quad (4)$$

Om de leptonische Lagrangiaanse dichtheid te bekomen, kan er gelijktijdig de  $SU(2)$  en  $U(1)$  transformatie toegepast worden. De Lagrangiaanse dichtheid voor vrije leptonen is niet invariant onder deze transformaties. De Lagrangiaanse dichtheid van de elektrozwakke sector wordt verkregen door de afgeleiden te vervangen door covariante afgeleiden. De covariante afgeleiden zijn afhankelijk van de generatoren van de bijhorende symmetriegroepen en introduceren enkele ijkvelden. [1]

$$\begin{cases} \partial_\mu \Psi_l^L \rightarrow D_\mu \Psi_l^L = \partial_\mu \Psi_l^L + ig_W \mathbf{T} \cdot \mathbf{W}_\mu(x) \Psi_l^L + ig' \frac{Y}{2} B_\mu(x) \Psi_l^L \\ \partial_\mu \psi_l^R \rightarrow D_\mu \psi_l^R = \partial_\mu \psi_l^R + ig' \frac{Y'}{2} B_\mu(x) \psi_l^R \\ \partial_\mu \psi_{\nu_l}^R \rightarrow D_\mu \psi_{\nu_l}^R = \partial_\mu \psi_{\nu_l}^R + ig' \frac{Y''}{2} B_\mu(x) \psi_{\nu_l}^R \end{cases} \quad (5)$$

Op deze manier kan er een Lagrangiaanse dichtheid opgesteld worden die invariant is onder  $SU(2) \times U(1)$  transformaties. [8]

$$\mathcal{L} = i[\bar{\Psi}_l^L \gamma^\mu D_\mu \Psi_l^L + \bar{\psi}_l^R \gamma^\mu D_\mu \psi_l^R + \bar{\psi}_{\nu_l}^R \gamma^\mu D_\mu \psi_{\nu_l}^R] \quad (6)$$

Indien hier echter een massaterm aan toegevoegd wordt, wordt de lokale ijksymmetrie gebroken en is de theorie niet meer renormaliseerbaar. Opdat de ijkbosonen massa kunnen verkrijgen, wordt het Higgs mechanisme geïntroduceerd. Het eenvoudigste Higgs model dat zal zorgen voor de benodigde massatermen, bestaat uit twee complexe scalaire velden die geplaatst worden in een zwak isospin doublet. [1] [8]

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \quad (7)$$

De Lagrangiaanse dichtheid voor dit nieuw geïntroduceerde scalaire veld wordt gegeven door: [1]

$$\begin{cases} \mathcal{L} = (\partial_\mu \phi)^\dagger (\partial^\mu \phi) - V(\phi) \\ V(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2 \end{cases} \quad (8)$$

Hierbij is  $V(\phi)$  de Higgs potentiaal. Om een eindig minimum te hebben, moet  $\lambda$  groter dan nul zijn. Voor  $\mu^2$  is er echter geen restrictie. Als deze groter dan nul gekozen wordt, komt de laagste energietoestand van het veld (de vacuümverwachtingswaarde) overeen met nul. Indien deze echter kleiner dan nul gekozen wordt, kan  $\mu^2$  niet meer

geïnterpreteerd worden als een massa en heeft het veld nu een niet-nul vacuümverwachtingswaarde. Er bestaat dan een oneindige set van ontaarde minima voor de potentiaal. [1]

$$\phi^\dagger\phi = \frac{1}{2}(\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2) = \frac{v^2}{2} = -\frac{\mu^2}{2\lambda} \quad (9)$$

Er wordt een minimum gekozen en de velden kunnen dan beschreven worden als expansies rond het minimum. [1]

$$\langle 0|\phi|0\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \quad (10)$$

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ v + \eta(x) + i\phi_4 \end{pmatrix} \quad (11)$$

De symmetrie wordt spontaan gebroken door een bepaalde vacuümtoestand te kiezen; de theorie is niet meer ijk invariant, maar is afkomstig van een ijk invariante theorie. Goldstone bosonen zijn de deeltjes beschreven door de scalaire velden  $\phi_i$  en komen overeen met expansies in de richting waar de potentiaal niet verandert. [1] [8]

Het Higgs doublet kan dan nog vereenvoudigd worden door het wegwerken van de drie niet-fysische en massaloze Goldstone velden ( $\phi_i$ ) aan de hand van een gepaste ijktransformatie. Deze ijktransformatie wordt de unitaire ijk genoemd. Dit komt er op neer het veld op de volgende manier te kiezen: [1] [8]

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \quad (12)$$

Er kan nu een kinetische term met het Higgs doublet opgeschreven worden. Omdat deze term invariant is onder de lokale ijsymmetrie van de elektrozwakke theorie, moet de afgeleide vervangen worden door de covariante afgeleide. Het Higgs doublet kan ook vervangen worden door het doublet in de unitaire ijk te kiezen. Hierbij wordt de lokale ijsymmetrie gebroken. [1]

$$\begin{aligned} D_\mu\phi(x) &= [\partial_\mu + ig_W\sigma_a W_\mu^a + ig' \frac{Y}{2} B_\mu]\phi(x) \\ &= \frac{1}{2\sqrt{2}} \begin{pmatrix} 2\partial_\mu + ig_W W_\mu^{(3)} + ig' \frac{Y}{2} B_\mu & ig_W(W_\mu^{(1)} - iW_\mu^{(2)}) \\ ig_W(W_\mu^{(1)} + iW_\mu^{(2)}) & 2\partial_\mu - ig_W W_\mu^{(3)} + ig' \frac{Y}{2} B_\mu \end{pmatrix} \begin{pmatrix} 0 \\ v + h \end{pmatrix} \\ &= \frac{1}{2\sqrt{2}} \begin{pmatrix} ig_W(W_\mu^{(1)} - iW_\mu^{(2)})(v + h) \\ (2\partial_\mu - ig_W W_\mu^{(3)} + ig' \frac{Y}{2} B_\mu)(v + h) \end{pmatrix} \end{aligned} \quad (13)$$

$$\begin{aligned}
(D_\mu\phi)^\dagger(D^\mu\phi) &= \frac{1}{2}(\partial_\mu h)(\partial^\mu h) + \frac{1}{8}g_W^2(W_\mu^{(1)} + iW_\mu^{(2)})(W^{\mu(1)} - iW^{\mu(2)})(v+h)^2 \\
&+ \frac{1}{8}(g_W W_\mu^{(3)} - g' \frac{Y}{2} B_\mu)(g_W W^{\mu(3)} - g' \frac{Y}{2} B^\mu)(v+h)^2
\end{aligned} \tag{14}$$

De Lagrangiaanse dichtheid die bekomen wordt door deze substitutie te maken, is afkomstig van een ijk-invariante theorie maar nu geschreven in functie van expansies rond de vacuümtoestanden. De ijkvelden in de Lagrangiaanse dichtheid kunnen in functie van de fysische ijkvelden geschreven worden. [1]

$$W^+ = W_\mu^{(1)} + iW_\mu^{(2)} \tag{15}$$

$$W^- = W_\mu^{(1)} - iW_\mu^{(2)} \tag{16}$$

$$A_\mu = \frac{g'W_\mu^{(3)} + g_W B_\mu}{\sqrt{g_W^2 + (g')^2}} \tag{17}$$

$$Z_\mu = \frac{g_W W_\mu^{(3)} - g' B_\mu}{\sqrt{g_W^2 + (g')^2}} \tag{18}$$

Aan de hand van deze identificatie kan de zwakke hyperlading,  $Y$ , van het Higgsveld bepaald worden. Er wordt geëist dat het fotonveld niet koppelt aan het Higgsveld. Dit is nodig opdat het fotonveld massaloos is. Er wordt gevonden dat de zwakke hyperlading van het Higgsveld gelijk is aan  $\frac{1}{2}$ . [1]

De massatermen voor de ijk-bosonen komen nu overeen met de termen die kwadratisch zijn in de ijk-bosonvelden. Drie fysische bosonvelden verkrijgen een massa en één ijkveld blijft massaloos. Op deze manier kunnen deeltjes in het standaardmodel massa verkrijgen door te interageren met het Higgsveld. De bekomen theorie wordt de standaard elektrozwakke theorie genoemd. [1] [8]

$$m_W = \frac{1}{2}g_W v \tag{19}$$

$$m_A = 0 \tag{20}$$

$$m_Z = \frac{1}{2}v\sqrt{g_W^2 + g'^2} \tag{21}$$

De fermionische velden worden gekoppeld aan het Higgsveld via de Yukawa interacties. Er kan op twee manieren een Lagrangiaanse dichtheid opgeschreven worden die de fermionen koppelt aan het Higgsveld en voldoet aan de lokale ijk-symmetrie van de elektrozwakke theorie. [1] [8]

$$\mathcal{L} = -g_f[\bar{L}\phi R + (\bar{L}\phi R)^\dagger] \tag{22}$$

$$\mathcal{L} = g_f[\bar{L}\phi_c R + (\bar{L}\phi_c R)^\dagger] \tag{23}$$

Hierbij is  $\phi_c$  de geconjugeerde van het Higgs doublet en geldt er: [1]

$$\begin{aligned}\phi_c &= -i\sigma_2\phi^* \\ &= \frac{1}{\sqrt{2}} \begin{pmatrix} -\phi_3 + i\phi_4 \\ \phi_1 - i\phi_2 \end{pmatrix}\end{aligned}\quad (24)$$

Charm quarks zijn up-type quarks, en de koppeling tussen charm quarks en het Higgsveld wordt gegeven door de Lagrangiaanse dichtheid in vergelijking 23. Hierbij is  $L$  een doublet met de linkshandige fermionen en transformeert volgens  $SU(2)$ .  $R$  is een singlet dat de rechtshandige fermionen bevat en transformeert als gevolg niet volgens  $SU(2)$ . [1]

Voor het charm quark geldt er: [1]

$$L = \begin{pmatrix} \psi_c^L \\ \psi_s^L \end{pmatrix}\quad (25)$$

$$R = \psi_c^R\quad (26)$$

Door opnieuw de unitaire ijk te kiezen, wordt er een uitdrukking bekomen voor de massa van het charm quark en ook de Yukawa koppeling. De Yukawa koppeling geeft de sterkte aan van de driepuntsinteractie tussen charm-anticharm en een Higgsboson. [1]

$$\mathcal{L} = -\frac{g_c}{\sqrt{2}}v(\bar{\psi}_c^L\psi_c^R + \bar{\psi}_c^R\psi_c^L) - \frac{g_c}{\sqrt{2}}h(\bar{\psi}_c^L\psi_c^R + \bar{\psi}_c^R\psi_c^L)\quad (27)$$

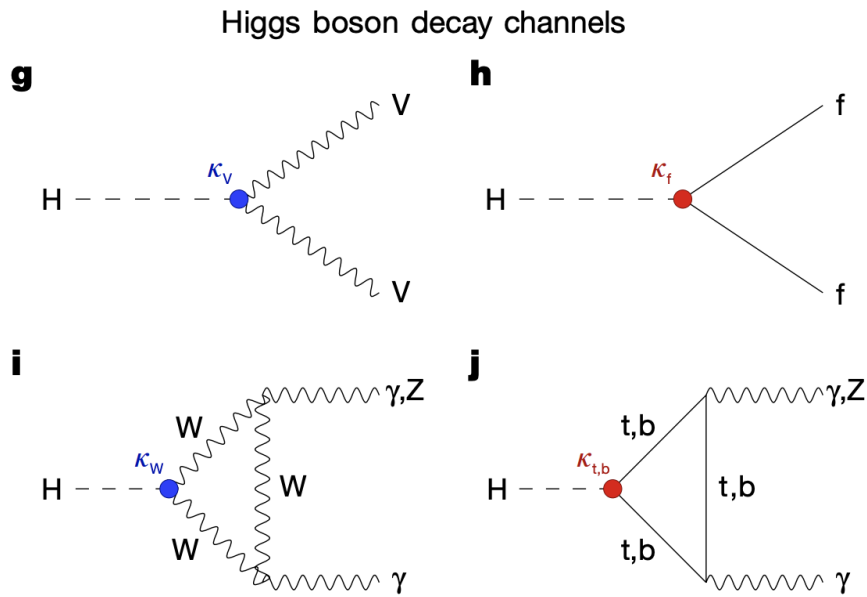
De eerste twee termen geven de massatermen voor de charm quarks en de laatste twee termen geven de interactie tussen het Higgsboson en de charm quarks. De Yukawa koppeling voor het charm quark wordt gegeven door: [1]

$$g_c = \sqrt{2}m_c/v\quad (28)$$

Hierbij heeft de vacuümverwachtingswaarde een waarde van  $v = 246$  GeV. [1]

## 2.3 Higgs processen

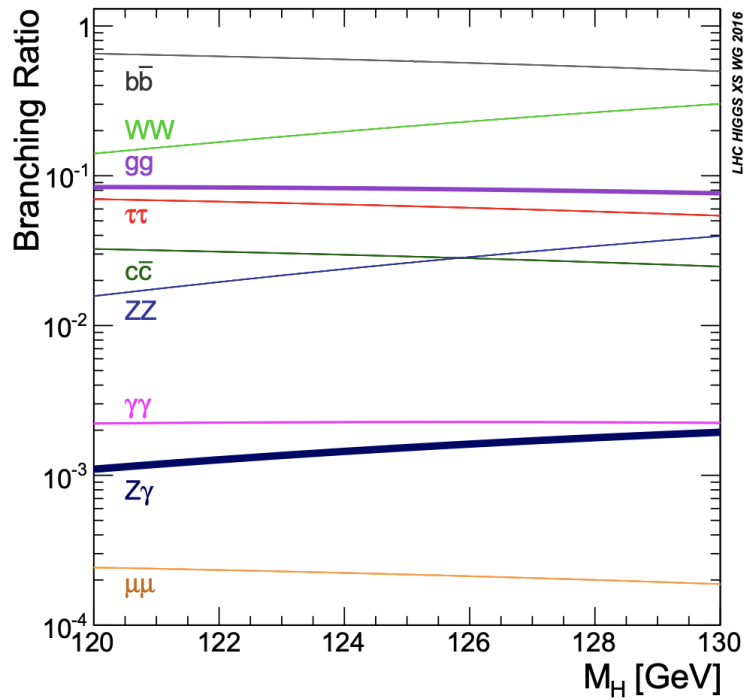
Het Higgsboson is een neutraal, scalair deeltje. Het kan in principe naar alle SM deeltjes vervallen. Aangezien het foton en het Higgsboson niet koppelen, kan het Higgsboson echter enkel indirect naar fotonen vervallen via kwantumloops. De Feynmandiagrammen voor het verval van het Higgsboson (naar fermionen ( $f$ ), vectorbosen ( $V$ ),  $Z$ -bosen ( $Z$ ) en fotonen ( $\gamma$ )) worden weergegeven in figuur 2. [1]



Figuur 2: Vervalkanalen Higgsboson [2]

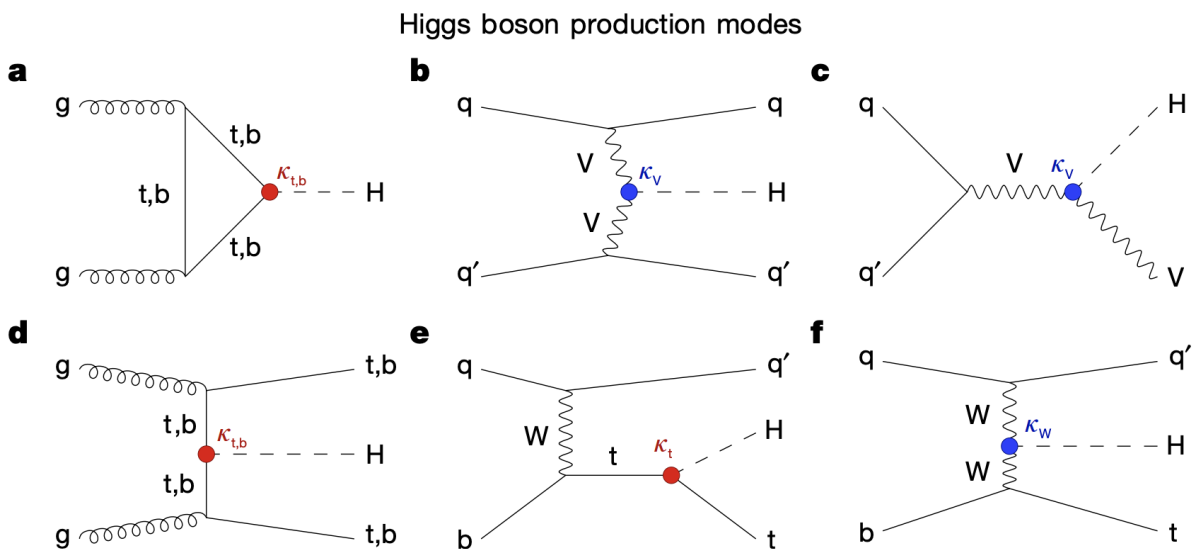
De observatie van het verval van het Higgsboson naar verschillende SM deeltjes en de meting van de branching ratio's zijn belangrijk om de theorie te testen. Door de kleine levensduur van het Higgsboson, vervalt het direct naar een paar fermionen of bosonen na productie. De amplitude van de koppeling tussen het Higgsboson en fermionen is evenredig met de massa van het fermion. Anderzijds is de amplitude van de koppeling tussen het Higgsboson en bosonen evenredig met de massa kwadraat van het boson. [2]

Doordat de koppeling tussen het Higgsboson en fermionen proportioneel is aan de massa van het fermion, hebben de vervallen van het Higgsboson naar zwaardere fermionen grotere branching ratio's. Hierbij wordt met de branching ratio de relatieve frequentie van een bepaald vervalkanaal ten opzichte van alle mogelijke vervalkanalen bedoeld. Het Higgsboson vervalt voornamelijk naar een bottom quark-antiquarkpaar, zoals weergegeven in figuur 3. [1]



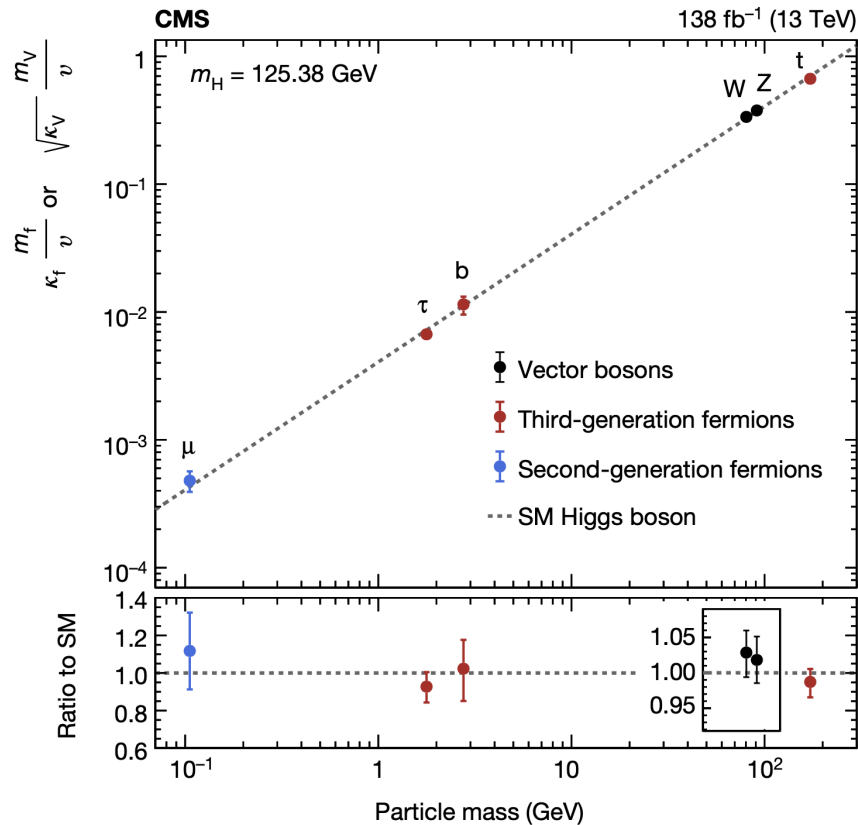
Figuur 3: Branching ratio's voor de belangrijkste vervalkanalen van het Higgsboson [9]

De voornaamste productieprocessen voor het Higgsboson worden weergegeven in figuur 4 en zijn de gluon fusie (ggF) (a), vector boson fusie (VBF) (b), geassocieerde productie met een ijkboson (VH) (c), en geassocieerde productie met een top-antitoppaar ( $t\bar{t}H$ ) (d). [1] [10] Het productieproces via gluon fusie is het meest voorkomend bij de LHC. [2]



Figuur 4: Productieprocessen Higgsboson [2]

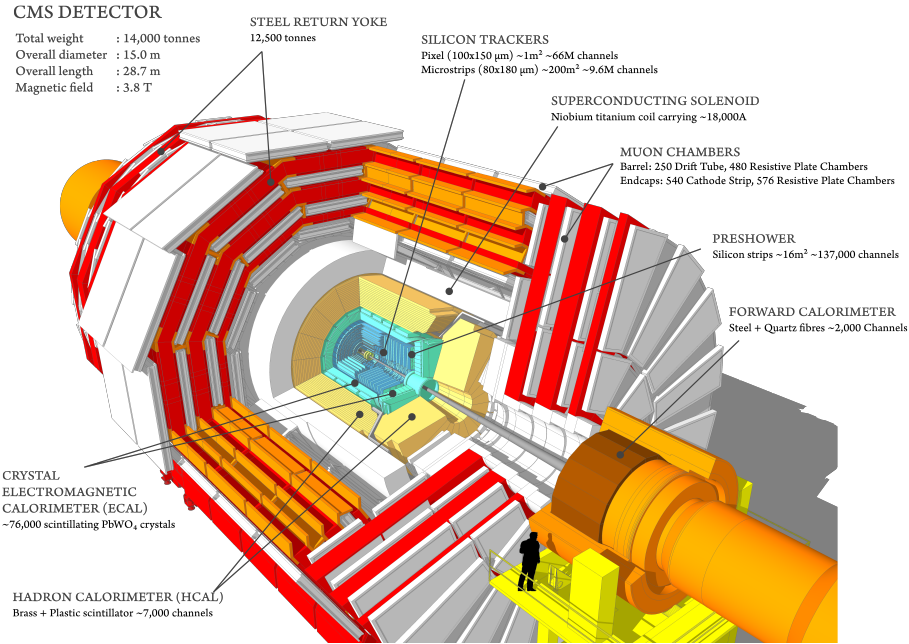
Bosonische vervallen van het Higgsboson  $H \rightarrow \gamma\gamma$ ,  $H \rightarrow ZZ \rightarrow 4l$ ,  $H \rightarrow WW \rightarrow \nu l \nu l$  en  $H \rightarrow Z\gamma$  werden geobserveerd. Ook de fermionische vervalkanalen  $H \rightarrow \tau\tau$ ,  $H \rightarrow bb$  en  $H \rightarrow \mu\mu$  zijn waargenomen. Hierbij zijn  $H \rightarrow \gamma\gamma$  en  $H \rightarrow ZZ \rightarrow 4l$  de meest gevoelige vervalkanalen. Deze vervalkanalen hebben een kleine branching ratio, maar zijn makkelijk waar te nemen en onderscheidbaar van de achtergrond. [1] [2]



Figuur 5: Waargenomen koppelingen voor het Higgsboson met fermionen en bosonen in functie van de massa [2]

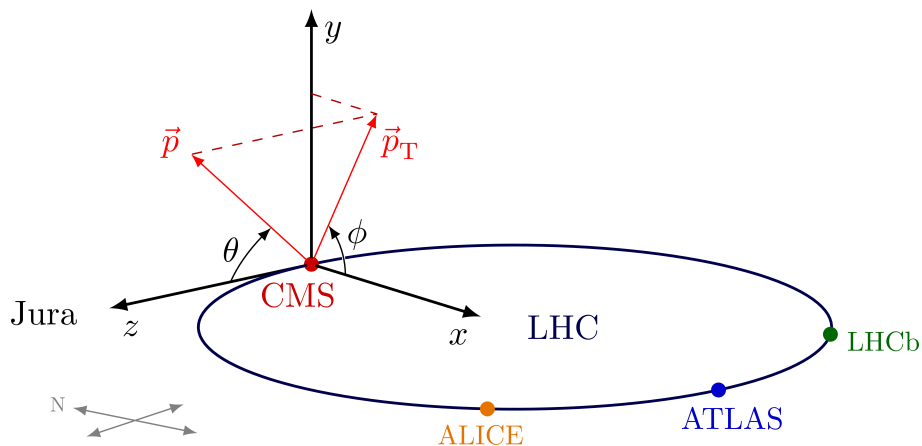
De koppeling van het Higgsboson met bottom quarks, top quarks, tau leptonen, en muonen zijn reeds waargenomen, zie figuur 5. Met hogere luminositeiten verhoogt de productie van het Higgsboson en is het CMS experiment ook in staat om de koppeling van het Higgsboson met charm quarks te observeren. [2] Hierbij is de luminositeit een maat voor de frequentie aan potentiële botsingen tussen deeltjes in de deeltjesbundels. [11] De geïntegreerde luminositeit geeft het totale aantal potentiële botsingen gedurende een bepaalde periode van metingen. [11]





Figuur 7: CMS detector [15]

Het coördinatenstelsel bij CMS heeft zijn oorsprong in het interactiepunt en wordt weergegeven in figuur 8. De x-as is horizontaal naar het middelpunt van de LHC-ring gericht, de y-as is verticaal naar boven gericht en de z-as ligt volgens de bundelrichting. De azimuthale hoek  $\phi$  is de hoek in het x-y-vlak vertrekkende van de x-as met als radiale coördinaat  $r$  in dit vlak. De transversale impuls en energie,  $p_T$  en  $E_T$ , worden berekend aan de hand van de x- en y-componenten. Het behoud van impuls geldt voor het transversale vlak en het geobserveerde tekort aan energie in dit vlak, is de missende transversale energie  $E_T^{\text{miss}}$ . De polaire hoek  $\theta$  is de hoek vertrekkende van de z-as en hiervoor wordt ook de pseudorapiditeit gedefinieerd  $\eta = -\ln(\tan(\frac{\theta}{2}))$ . [14]



Figuur 8: CMS coördinaten [16]

CMS heeft een 4 T supergeleidende solenoïde, gemaakt van NbTi geleidermateriaal. Elektrisch geladen deeltjes ondergaan een afbuiging door de aanwezigheid van een

magnetisch veld. De afbuiging van het deeltje is afhankelijk van de grootte van de impuls dat het heeft. Een sterk magnetisch veld is dus belangrijk voor nauwkeurige metingen van de impuls. [14] [17]

Er is ook een return yoke aanwezig die nodig is voor de controle van het veld buiten de solenoïde. Hierin zitten de muonkamers vervat. Deze bestaan uit enkele lagen opgebouwd uit aluminium drift tubes en resistive plate chambers in de barrel-regio, en cathode strip chambers en resistive plate chambers voor de endcap-regio. Het muonenstelsel zorgt voor de identificatie, meting van de impuls en ook de triggering van muonen. Het muonenstelsel is in staat precieze metingen te doen van de impuls van muonen in de regio tot  $|\eta| < 2,4$ . [14] [17]

De binnenste tracker en de calorimeters worden vervat door de solenoïde. Het tracker systeem zal zorgen voor zowel precieze metingen van de trajecten van geladen deeltjes als precieze metingen van de secundaire vertices. De trackers zijn gemaakt van concentrische lagen van silicium sensoren. Ze bestaan voor het grootste deel uit microstrip detectoren, en bijkomend pixel detectoren dichtbij het interactiepunt. Deze zorgen voor een hoge granulariteit en snelle respons, wat vereist is bij de hoge frequentie aan interacties in CMS. De binnenste tracker dekt de regio tot  $|\eta| < 2,5$ . [14] [17]

De hermetische, homogene ECAL (elektromagnetische calorimeter) bevat  $PbWO_4$  kristallen. Wanneer fotonen en elektronen/positronen door de kristallen gaan, verliezen ze energie waardoor scintillatielicht gegenereerd wordt. Dit licht wordt gedetecteerd door silicium avalanche photodiodes in de barrel-regio en vacuüm phototriodes in de endcap-regio. De hoeveelheid opgemeten licht is een maat voor de energie van de deeltjes. De ECAL kan energieën opmeten van fotonen en leptonen tot  $|\eta| < 3$ . De ECAL heeft een hoge granulariteit, een snelle respons en is resistent tegen straling door de hoge dichtheid aan kristallen die geplaatst zijn. [14] [17]

Rondom de ECAL bevindt zich een brass/scintillator sampling HCAL (hadronische calorimeter). Deze zorgt voor metingen van de richting en de energie van jets. Jets bestaan uit deeltjes die in min of meer dezelfde richting bewegen en geproduceerd zijn door een initieel quark of gluon. [2] De energie van jets is belangrijk om informatie te krijgen over de quarks, gluonen en neutrino's. De HCAL kan de energieën van jets opmeten tot  $|\eta| < 3$ . Een tail-catcher wordt ook nog toegevoegd buiten de solenoïde voor de hadronische cascades die nog niet volledig geabsorbeerd zijn door de HCAL. Ook in de voorwaartse regio's wordt de HCAL gecomplementeerd door een quartz-fibre calorimeter. De Cherenkov straling die geproduceerd wordt in de quartzfibras, wordt opgemeten door photomultipliers. Deze dekt de regio's tot  $|\eta| < 5$ . [14] [17]

Events bestaan uit de opgemeten data van één getriggerde bundelinteractie samen met informatie afgeleid uit deze interactie. Door de hoge frequentie aan interacties is er een triggersysteem nodig dat het aantal events sterk vermindert. Dit is nodig omdat niet alle data opgeslagen kan worden. De online event selectie door het trig-

gersysteem gebeurt in twee stappen. Ten eerste, maakt de Level-1 (L1) trigger een selectie op basis van informatie van de calorimeters en muonkamers. De L1 trigger is opgebouwd uit hardware systemen bestaande uit programmeerbare elektronica. Vervolgens, maakt de High-level trigger (HLT) een selectie gebruikmakend van complexe calculaties en de informatie van alle subdetectoren. Deze wordt opgebouwd uit een software systeem bestaande uit processors. Er moet elke 25 ns een beslissing gemaakt worden of het event al dan niet opgeslagen wordt. [14] [17]

De trigger zorgt samen met het data-acquisitiesysteem ervoor dat alle informatie van alle afzonderlijke detectoren verzameld en geanalyseerd wordt. Hierna worden de events opgeslagen en kunnen later gebruikt worden voor de offline analyse. [14]

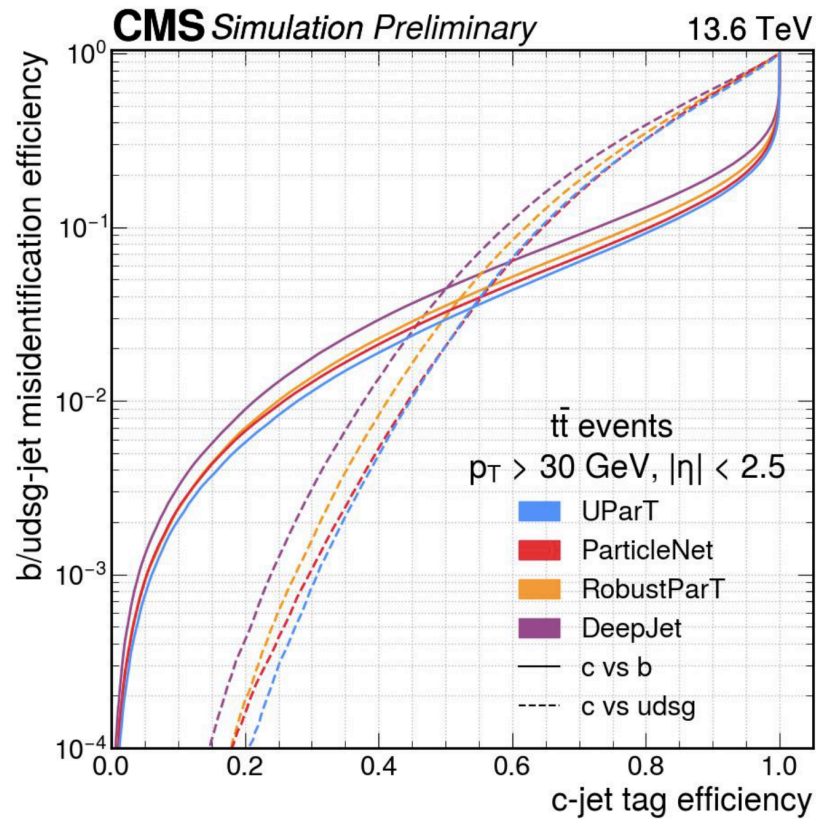
### 3.3 Event reconstructie

Aan de hand van de informatie van alle subdetectoren kan er een identificatie van de deeltjes en reconstructie van hun trajecten gedaan worden. De identificatie gebeurt op basis van de producten van de interactie en wordt gedaan aan de hand van het CMS particle-flow (PF) algoritme [18]. Dit geeft een set van PF kandidaat-objecten terug, en kan bestaan uit geladen en neutrale hadronen, elektronen, muonen en fotonen. [19] [20]

Voor de reconstructie van de elektrontrajecten wordt zowel informatie van de tracker als van de ECAL gebruikt. [21] Zowel de waarnemingen van de tracker als van het muonenstelsel worden gebruikt voor de reconstructie van de trajecten van de muonen. Jets kunnen geïdentificeerd en gereconstrueerd worden aan de hand van het anti- $k_T$  algoritme [22]. Dit wordt gedaan door alle objecten met een kleine afstand van elkaar te clusteren, waarbij een afstandsparameter  $R = 0,4$  gebruikt wordt. De afstandsparameter wordt gedefinieerd als  $\sqrt{(\Delta\eta)^2 + \Delta\phi^2}$ . De impuls van een jet wordt bepaald door de impulsen van de deeltjes die de jet opbouwen, vectorieel te sommeren. De energie van de jets wordt dan ook nog gecorrigeerd voor pileup effecten. Wanneer deeltjesbundels kruisen, gebeurt er meer dan één interactie. [23] Dit zorgt ervoor dat er naast de te onderzoeken interactie bijkomende events overlappen, en deze worden pileup events genoemd. [23] De correctie hiervoor kan gedaan worden met het PUPPI (pileup per particle identification) algoritme [24]. De jets afkomstig van bottom quarks kunnen onderscheiden worden aan de hand van b-tagging. Typisch voor bottom jets zijn secundaire vertices die samen met de impactparameters gebruikt worden voor b-tagging. De proton-proton interactievertext kan bepaald worden door de trajecten te extrapoleren. [19] [20]

Het onderscheiden van jets afkomstig van charm quarks ten opzichte van jets afkomstig van andere quarks, is een uitdaging. Om charm jets te taggen wordt er gebruikgemaakt van ParticleNet [25], een algoritme gebaseerd op een dynamisch grafisch convolutioneel neuraal netwerk. Hiermee kunnen charm (c) jets onderscheiden worden van bijvoorbeeld bottom (b) jets of jets geproduceerd door lichte quarks of gluonen (udsg). [3]

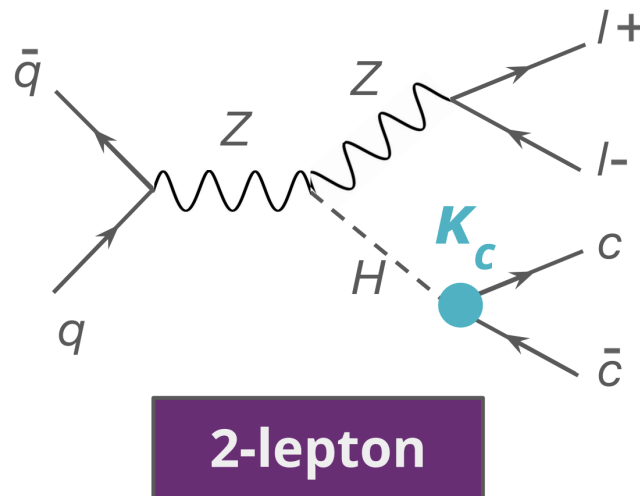
In figuur 9 wordt de prestatie van ParticleNet weergegeven. De figuur toont ook hoe er een afweging moet gemaakt worden tussen het goed onderscheiden van charm jets ten opzichte van ofwel bottom jets ofwel lichtere jets. ParticleNet presteert beter dan het ouder algoritme DeepJet dat gebruikt werd in Run2. Er is nu ook een nieuw model beschikbaar, UPart, dat nog een betere prestatie vertoont. Deze zal eventueel gebruikt worden in latere Run3 analyses. [26]



Figuur 9: Prestatie ParticleNet c-tagging [26]

## 4 Methodologie

Het proces dat onderzocht wordt, is een productie van een Higgsboson in associatie met een Z-boson (VH productie) door een proton-proton botsing, waarbij het Z-boson vervalt naar twee leptonen en het Higgsboson vervalt naar een charm quark-antiquarkpaar. Voor het verval van het Z-boson naar twee leptonen, worden enkel twee vervalkanalen beschouwd, namelijk het verval naar elektronen en muonen. Met tau leptonen wordt hier geen rekening gehouden. Het onderzochte proces wordt weergegeven in figuur 10.



Figuur 10: Overzicht onderzochte Higgs proces [27]

De productie van een Higgsboson in associatie met een Z-boson heeft een cross-sectie van  $\sigma(pp \rightarrow ZH) = 0,8839$  pb. [28] De cross-sectie van een proces is een maat voor de quantummechanische waarschijnlijkheid dat het proces plaatsvindt. [1] Ondanks dat de productie cross-sectie voor VH productie kleiner is dan de cross-secties voor andere productiekanalen, helpt het leptonische verval van het Z-boson ervoor dat achtergronden sterk verminderd kunnen worden. [29] Het verval van een Higgsboson naar een charm quark-antiquarkpaar heeft een branching ratio van  $BR(H \rightarrow c\bar{c}) = 0,0289$ . [30] Het verval naar een charm quark-antiquarkpaar heeft een kleine branching ratio, en daarbij zijn charm jets ook moeilijk te identificeren en te onderscheiden van bottom jets. [3]

Er wordt gewerkt in de "resolved-jet" topologie ten opzichte van de "merged-jet" topologie. Hiermee wordt er bedoeld dat de Higgsbosonkandidaat wordt gereconstrueerd van twee goed afgezonderde charm jets in de plaats van één grote jet. Dit is het geval bij Higgsbosonkandidaten met relatief lage  $p_T$  (kleiner dan 200 GeV). De meeste events voldoen aan de resolved-jet topologie, maar deze events bevatten wel meer achtergrond ten opzichte van de merged-jet topologie events. [3] [19]

Voor deze analyse wordt er onder andere gebruikgemaakt van het ROOT [31] en ook H+c [32] framework.

## 4.1 Achtergrond en datasets

Er wordt CMS Run 3 2022postEE data gebruikt bij een centre-of-mass energie van  $\sqrt{s} = 13,6$  TeV dat overeenkomt met een geïntegreerde luminositeit van  $26,7 \text{ fb}^{-1}$ .

De achtergrond voor het onderzochte proces komt voornamelijk van de productie van een  $W$ - of  $Z$ -boson samen met één of meerdere jets ( $V + \text{Jets}/ \text{DY} + \text{Jets}$ ), enkele ( $t$ ) of in paar geproduceerde ( $tt$ ) top-quarks, en de productie van een Higgsboson in associatie met een vector boson, waarbij het Higgsboson vervalst naar een bottom quark-antiquarkpaar ( $bb$ ). De achtergrond afkomstig van diboson ( $VV$ ) events is meestal heel klein door de kleine productie cross-sectie. Er is ook een achtergrond aanwezig afkomstig van jets geproduceerd via de sterke wisselwerking, namelijk quantum chromodynamics (QCD) multijet events. Door het vervalkanaal te kiezen waarbij het  $Z$ -boson naar twee leptonen vervalst, wordt deze laatste achtergrond sterk onderdrukt. [19]

Signaal en achtergrond events worden gesimuleerd met verscheidene Monte Carlo (MC) event generatoren. De respons van de detector wordt gesimuleerd met Geant4 [33]. Het signaal en de  $bb$  achtergrond (ZH samples) worden gegenereerd door de POWHEG v2 [34] [35] [36] event generator tot next-to-leading order (NLO) voor producties met quarks en tot leading order (LO) voor producties met gluonen. De  $V + \text{Jets}$  achtergrond samples worden gegenereerd met MADGRAPH5\_aMC@NLO v2.6.0 [37] tot NLO. Zowel de  $tt$  en  $t$  samples worden gegenereerd aan de hand van POWHEG v2 tot NLO. Ten slotte worden de  $VV$  achtergrond samples gegenereerd tot NLO met POWHEG v2 voor de  $WW$  samples en MADGRAPH5\_aMC@NLO v2.4.2 voor de  $WZ$  en  $ZZ$  samples. De parton showering (cascade van straling door QCD interacties [38]) en hadronisatie worden gesimuleerd met PYTHIA v8.230 [39]. Daarnaast worden pileup interacties ook gegenereerd met PYTHIA. De MC events krijgen dan een gewicht opdat de pileup van de geobserveerde data matcht met de simulaties. [3]

Voor de analyse wordt de (MC) data geraadpleegd in het NANO-AOD (Analysis Object Data) formaat. Dit is een compact formaat dat onder andere de high-level objecten bevat. [14]

## 4.2 Object en event selecties

De geselecteerde events moeten voldoen aan de HLT criteria. Voor het muonkanaal is deze trigger "IsoMu24". Hierbij moet er minstens één muon getriggerd zijn met een  $p_T$  hoger dan 23 GeV. Voor het elektronkanaal is dit "Ele30WPTightGsf" voor de HLT. Hiervoor moet er minstens één elektron getriggerd zijn met een  $p_T$  hoger dan 28 GeV.

De events en objecten (leptonen, jets, dileptonen, en dijets) moeten voldoen aan enkele offline criteria. Er worden onder andere eisen gesteld op de  $p_T$ ,  $\eta$  en identificatie en isolatie van de objecten. De Higgsbosonkandidaten (H-kandidaten) en Z-bosonkandidaten (Z-kandidaten) worden respectievelijk opgebouwd uit de jets en leptonen met de hoogste (leading) en tweede hoogste (subleading)  $p_T$ -waarden.

Jets ( $j$ ) worden ook opgeschoond door overlappings met geselecteerde elektronen ( $e$ ) of muonen ( $\mu$ ) te verwijderen (op basis van  $\Delta R$ ). [40] Voor de dilepton (Z-kandidaat) en dijet (H-kandidaat) objecten wordt er een criterium geplaatst op de invariante massa ( $m$ ). De leptonen die de Z-bosonkandidaat opbouwen, moeten bovendien een tegengestelde elektrische lading ( $q$ ) hebben. De object selecties worden opgelijst in tabel 1. [19]

Er is een bepaalde onzekerheid op de reconstructie van objecten in de detector. De identificatie (ID) is een maat voor hoe waarschijnlijk het waargenomen type object (muon/elektron/...) ook daadwerkelijk dat type object is. De isolatie (ISO) van een object is een maat voor hoe geïsoleerd een object is ten opzichte van andere objecten. De isolatie van een object wordt bepaald door de energieën binnen een kegel rond het object op te tellen en de pileup energie ervan af te trekken. [41] [42]

Bij de object selecties wordt er een selectie geplaatst op de identificatie en isolatie van muonen/elektronen/jets (Tight ID, Muon tight ISO, Wp80iso ID, tightlepveto). Hierbij betekent een sterker criterium dat echte muonen/elektronen/jets ook goed geselecteerd worden als muonen/elektronen/jets, maar dat er eveneens veel werkelijke muonen/elektronen/jets weggefilterd worden. Anderzijds betekent een lossier criterium dat veel van de werkelijke muonen/elektronen/jets behouden worden, maar dat veel van de muonen/elektronen/jets die geselecteerd zijn, fout geïdentificeerd worden. [42] [43]

Object selecties		
Muon object selecties	Elektron object selecties	Jet object selecties
$p_T(\mu) > 10 \text{ GeV}$ $ \eta(\mu)  < 2, 4$ Tight ID Muon tight ISO	$p_T(e) > 10 \text{ GeV}$ $ \eta(e)  < 2, 4$ Wp80iso ID	$p_T(j) > 20 \text{ GeV}$ $ \eta(j)  < 2, 4$ Tightlepveto ID $\Delta R(j, \mu) > 0, 4$ $\Delta R(j, e) > 0, 4$
Dilepton object selecties		Dijet object selecties
Muonkanaal	Elektronkanaal	$50 \text{ GeV} < m(H) < 200 \text{ GeV}$
$q(\mu_1)q(\mu_2) < 0$ $60 \text{ GeV} < m(Z) < 120 \text{ GeV}$	$q(e_1)q(e_2) < 0$ $60 \text{ GeV} < m(Z) < 120 \text{ GeV}$	

Tabel 1: Object selecties

Hierbovenop wordt er een selectie van de events gedaan onder andere op basis van het aantal ( $N$ ) dat aanwezig is van een bepaald object in een event. Er worden bijkomend criteria geplaatst op de leading ( $l_1, j_1$ ) en subleading ( $l_2, j_2$ ) leptonen en

jets. De event selecties worden opgelijst in tabel 2. Doordat de trigger-efficiëntie niet 100 % is, wordt de offline drempelwaarde voor de  $p_T$  van de leading leptonen lager gekozen dan de drempelwaarde voor de HLT.

Een groot deel van de achtergrond door  $V + \text{Jets}$  events kan verminderd worden door een eis te stellen op de invariante massa van de Higgsbosonkandidaat. De achtergrondevents waarbij een Higgsboson vervalt naar bottom quarks kan onderdrukt worden door gebruik te maken van jetflavour taggers. Er wordt een selectie gemaakt op basis van twee discriminatoren voor charm jets ten opzichte van andere jets. De eerste discriminator  $CvsL = p(c)/[p(c) + p(udsg)]$  maakt een onderscheid tussen charm jets en lichte quark of gluon jets. De tweede discriminator maakt een onderscheid tussen charm jets en bottom jets  $CvsB = p(c)/[p(c) + p(b)]$ . Hierbij stellen  $p(c)$ ,  $p(udsg)$  en  $p(b)$  de waarschijnlijkheden voor dat een jet komt van respectievelijk een charm quark, een lichte quark of gluon, en een bottom quark. De c-tagging waarden, de twee discriminatoren, worden bekomen aan de hand van een dynamisch grafisch convolutioneel neuraal netwerk, ParticleNet [25]. Hierbij worden jets beschouwd als "particle clouds" die bestaan uit PF kandidaten en gereconstrueerde secundaire vertices. [3] [19] [44]

De c-taggerwaarden van de jets waardoor het Higgsboson wordt opgebouwd, moeten de "loose" working points voor CvsL en CvsB behalen. [19] Dit betekent dat er een drempelwaarde wordt geplaatst bij de c-taggerwaarden met de focus op het onderscheiden van charm jets ten opzichte van bottom jets. [45] Zo is er een betere onderdrukking van bottom jets en worden er meer misidentificaties van charm jets als lichte jets toegelaten.

Event selecties	
Muonkanaal	Elektronkanaal
$p_T(\mu_1) > 17 \text{ GeV}$	$p_T(e_1) > 23 \text{ GeV}$
$N(\mu) = 2$	$p_T(e_2) > 12 \text{ GeV}$
$N(j) \geq 2$	$N(e) = 2$
$N(Z) = 1$	$N(j) \geq 2$
$N(H) = 1$	$N(Z) = 1$
$CvsL(j_1) > 0,054$	$N(H) = 1$
$CvsB(j_2) > 0,182$	$CvsL(j_1) > 0,054$
	$CvsB(j_2) > 0,182$

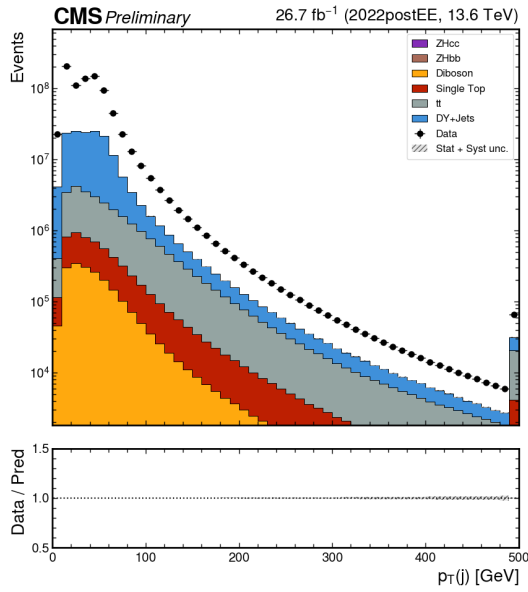
Tabel 2: Event selecties

Op deze manier wordt er getracht zoveel mogelijk achtergrond te onderdrukken, terwijl zoveel mogelijk van het signaal behouden blijft. Ook worden regio's die niet goed beschreven worden door de MC simulaties verwijderd. Dit kan gebeuren door bijvoorbeeld ontbrekende achtergronden in de MC simulaties. [19]

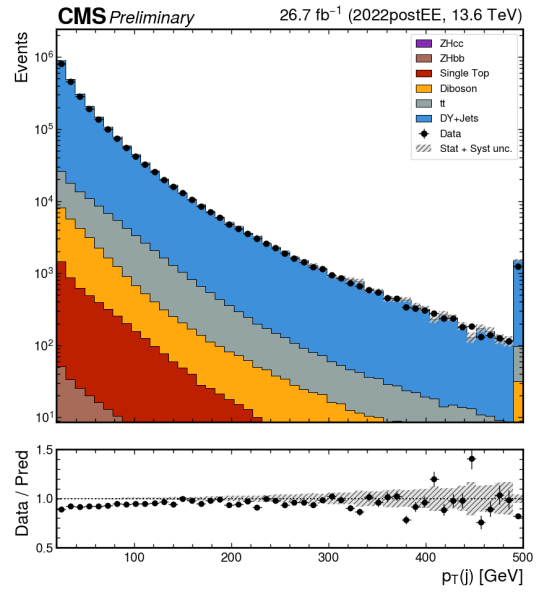
### 4.2.1 Plots van observabelen

In deze sectie worden enkele observabelen voor en na de selecties geplot in logaritmische schaal. Op deze manier wordt het effect van de selecties getoond: achtergrond wordt onderdrukt en regio's die niet goed beschreven worden door de MC simulaties, worden verwijderd. Het signaal wordt aangegeven door het "ZHcc" label. Onder elke plot wordt de ratio van de geobserveerde data op MC verwachting weergegeven. De schaalfactoren die verkregen worden aan de hand van controle regio's (zie sectie 4.3), zijn hierbij nog niet aanwezig.

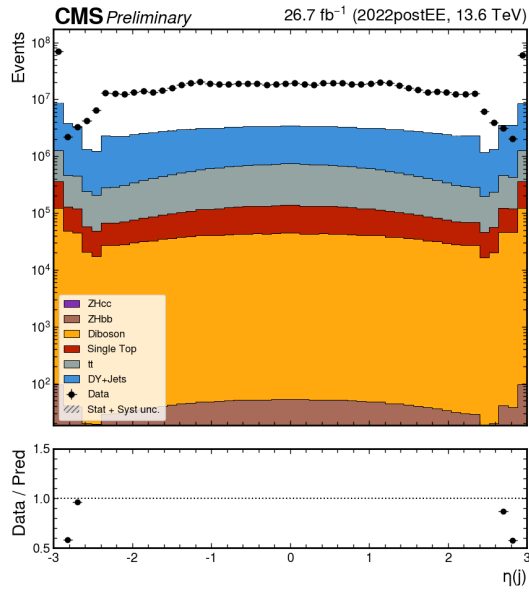
De plots die getoond worden in deze sectie behoren tot het muonvervalkanaal. Dezelfde plots voor het elektronvervalkanaal kunnen teruggevonden worden in appendix A.



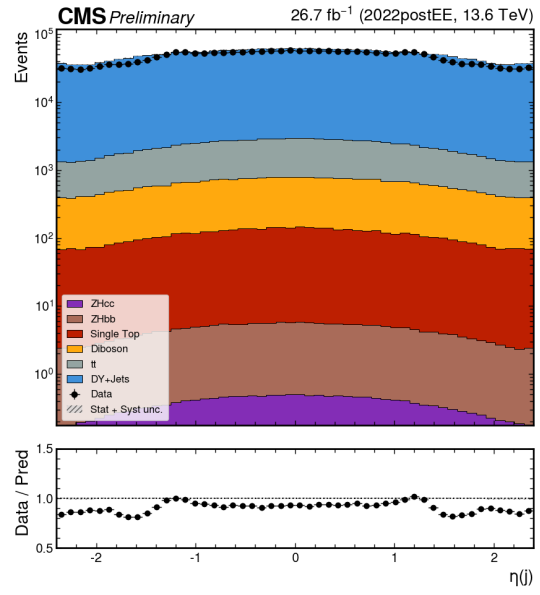
Figuur 11: Transversale impuls jets zonder selecties



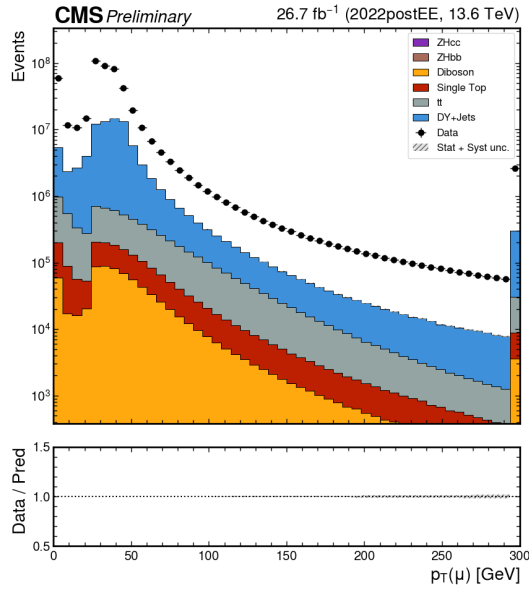
Figuur 12: Transversale impuls jets met selecties



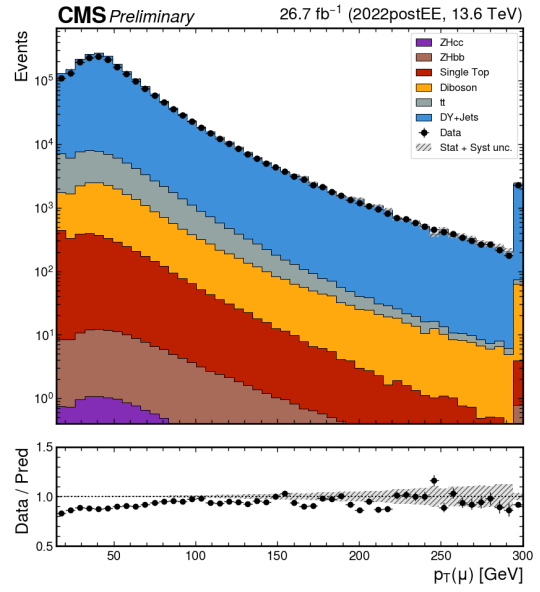
Figuur 13: Pseudorapiditeit jets zonder selecties



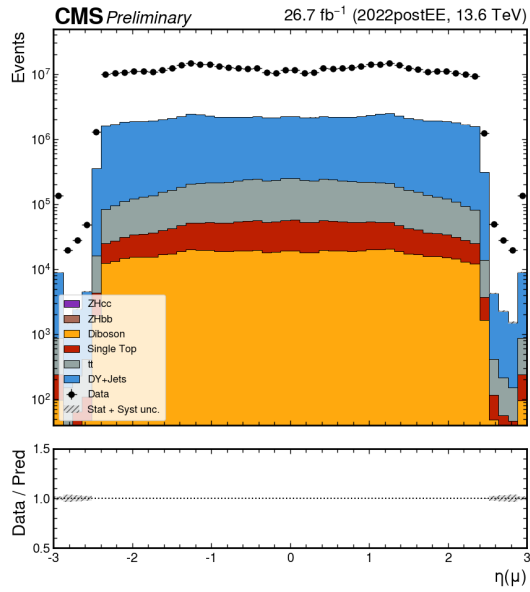
Figuur 14: Pseudorapiditeit jets met selecties



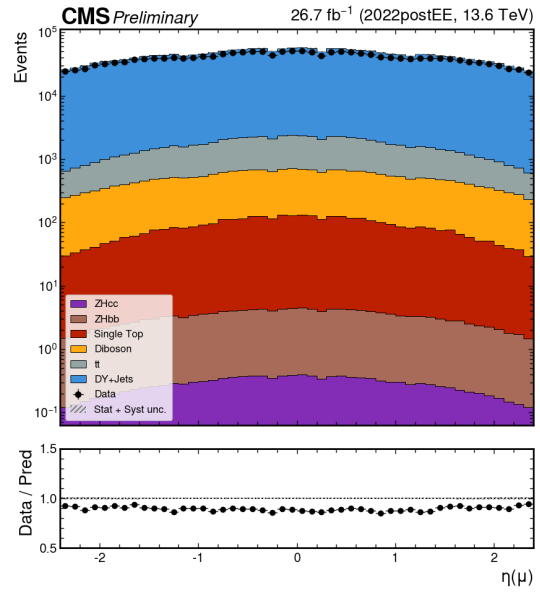
Figuur 15: Transversale impuls muonen zonder selecties



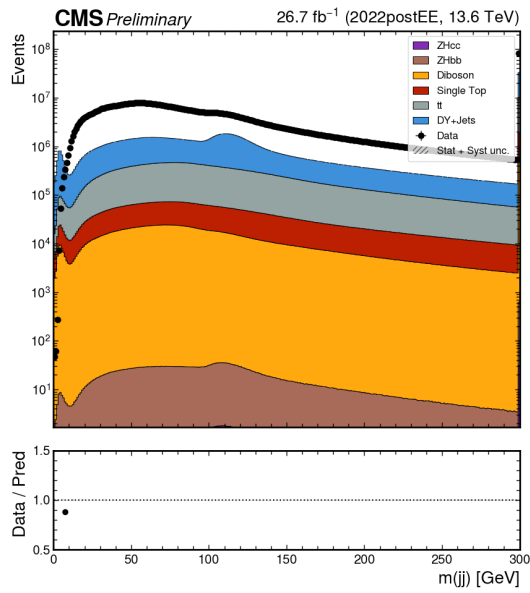
Figuur 16: Transversale impuls muonen met selecties



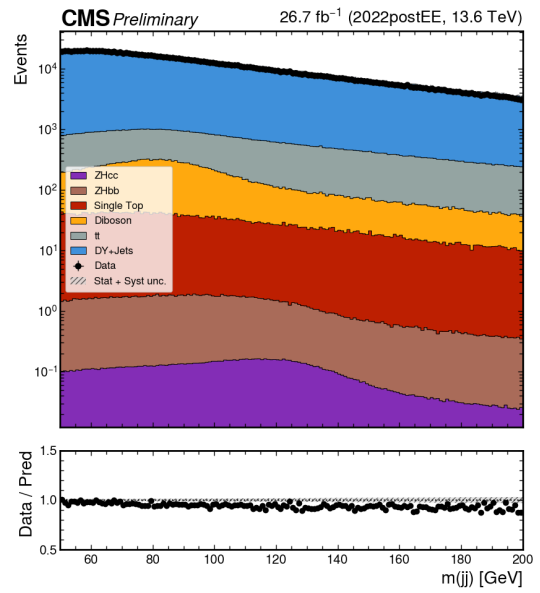
Figuur 17: Pseudorapiditeit muonen zonder selecties



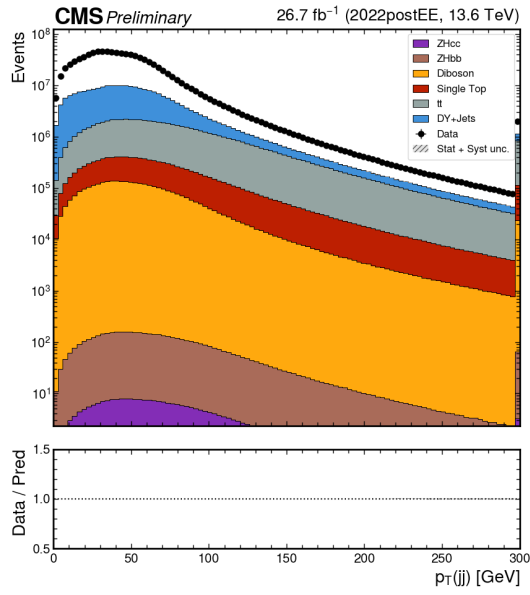
Figuur 18: Pseudorapiditeit muonen met selecties



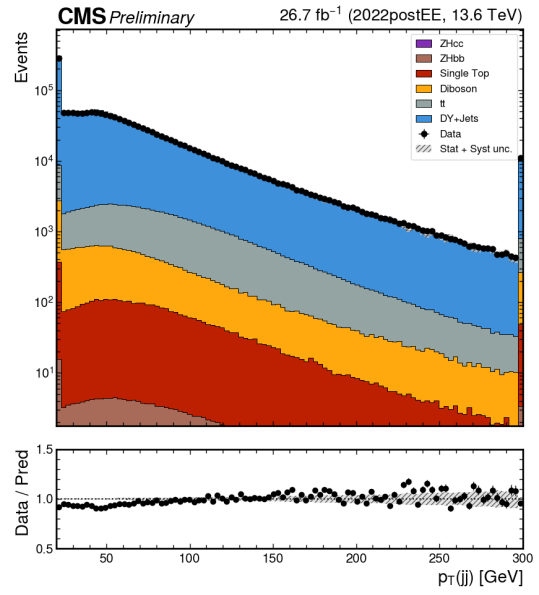
Figuur 19: Invariante massa dijets zonder selecties



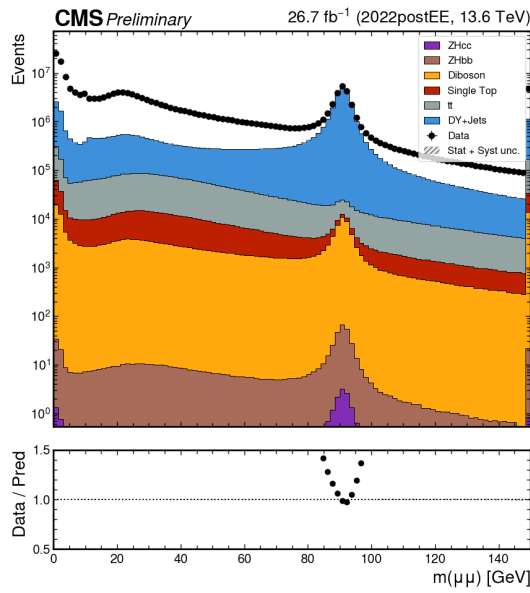
Figuur 20: Invariante massa dijets met selecties



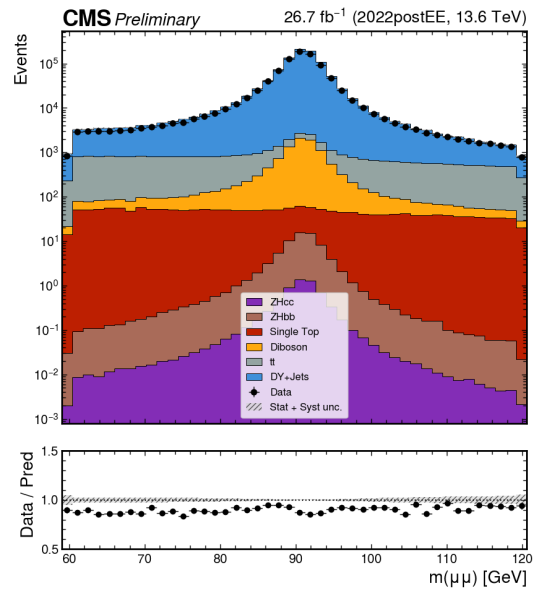
Figuur 21: Transversale impuls dijets zonder selecties



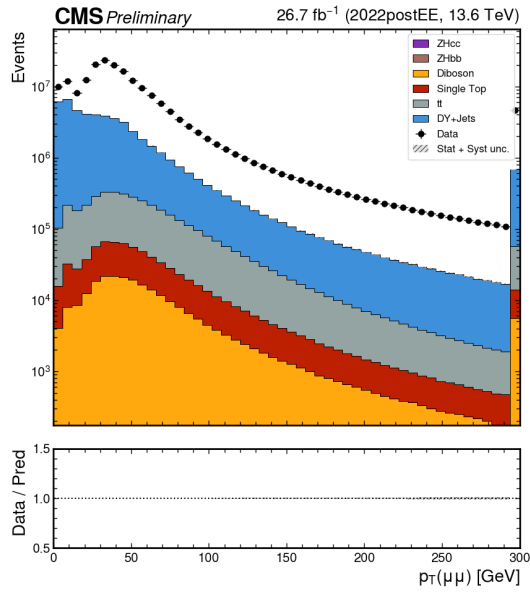
Figuur 22: Transversale impuls dijets met selecties



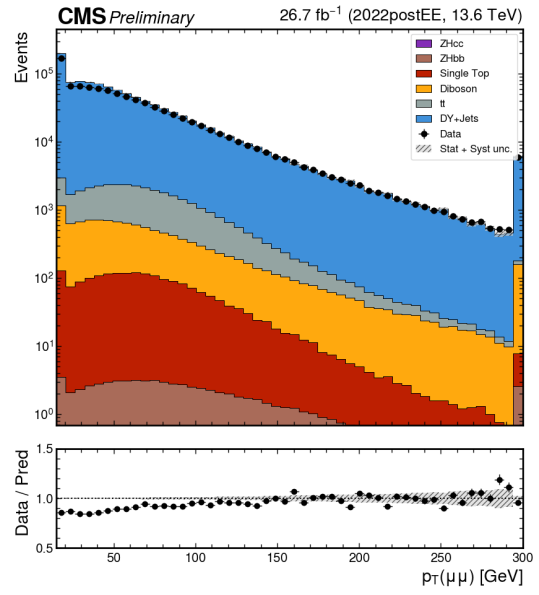
Figuur 23: Invariante massa dimuonen zonder selecties



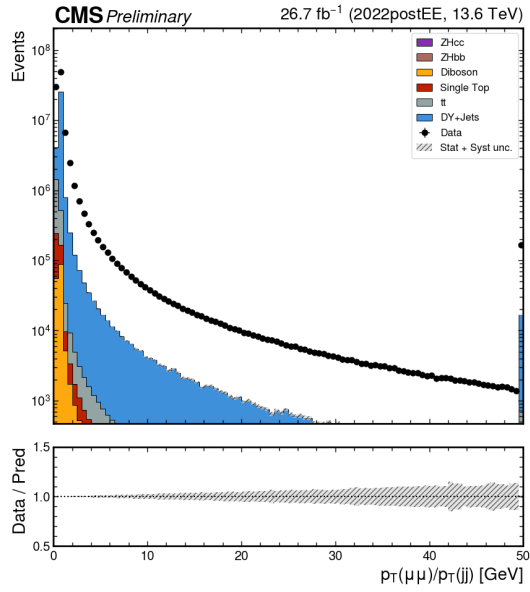
Figuur 24: Invariante massa dimuonen met selecties



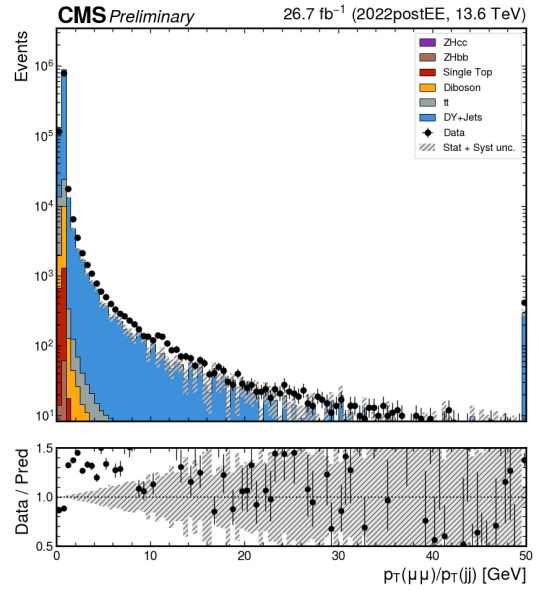
Figuur 25: Transversale impuls dimuonen zonder selecties



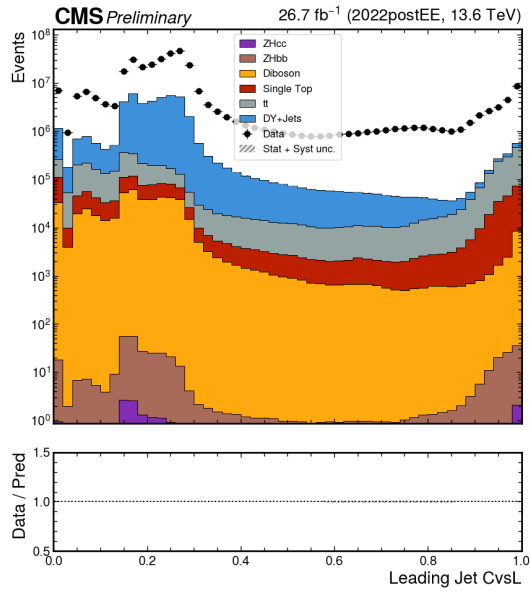
Figuur 26: Transversale impuls dimuonen met selecties



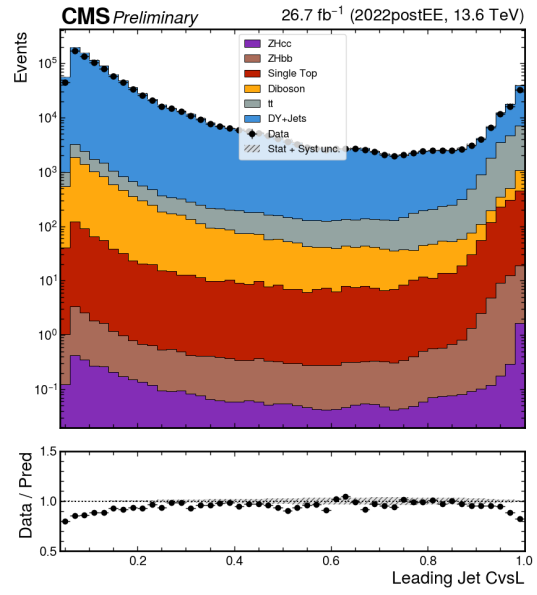
Figuur 27: Ratio van transversale impuls van dileptonen en dijets zonder selecties



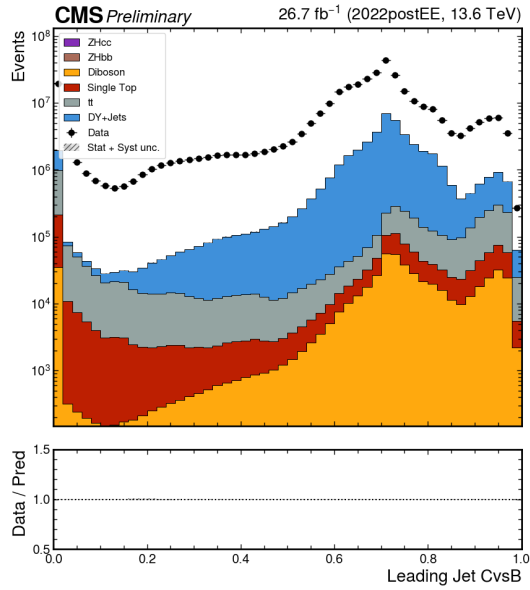
Figuur 28: Ratio van transversale impuls van dileptonen en dijets met selecties



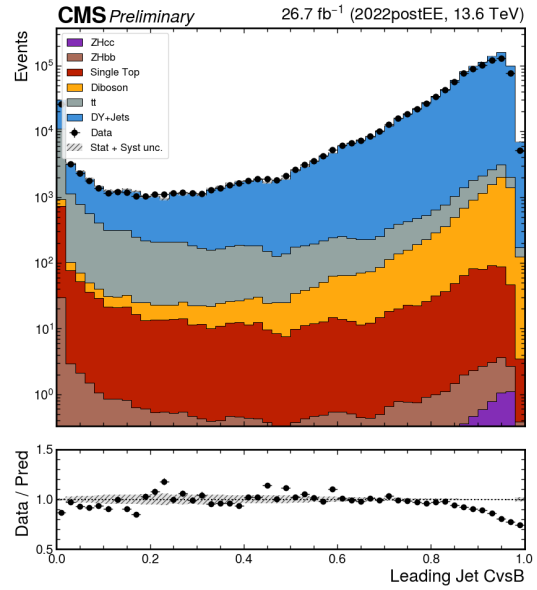
Figuur 29: CvsL-waarde van leading jet zonder selecties



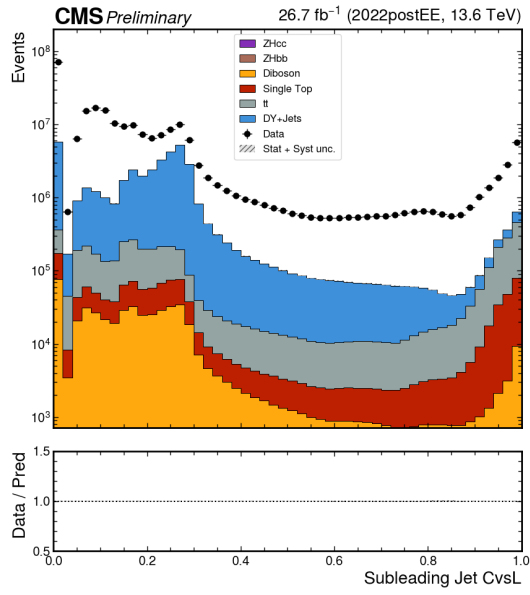
Figuur 30: CvsL-waarde van leading jet met selecties



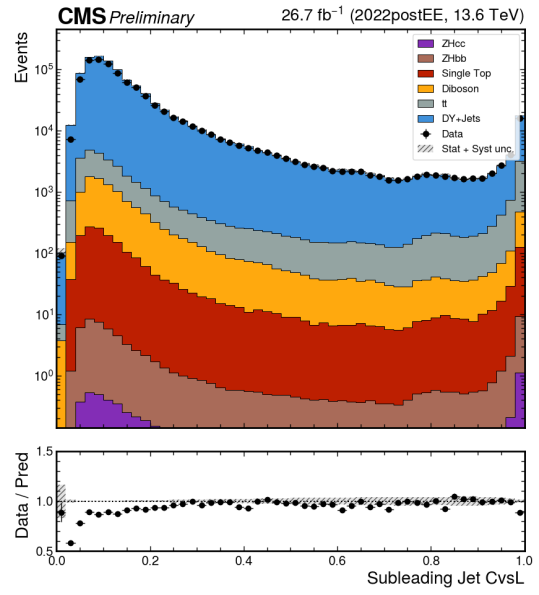
Figuur 31: CvsB-waarde van leading jet zonder selecties



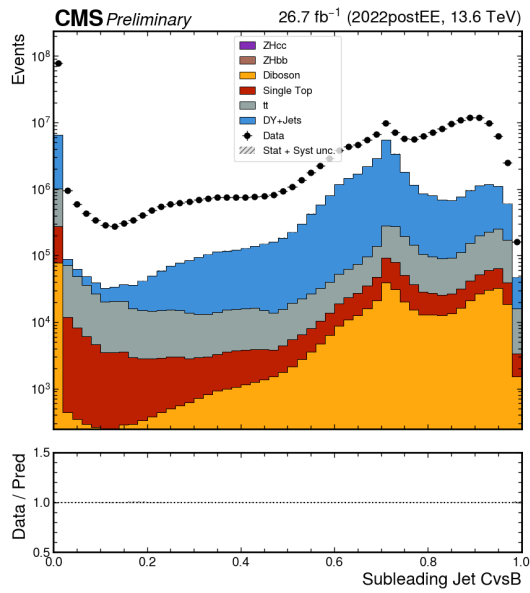
Figuur 32: CvsB-waarde van leading jet met selecties



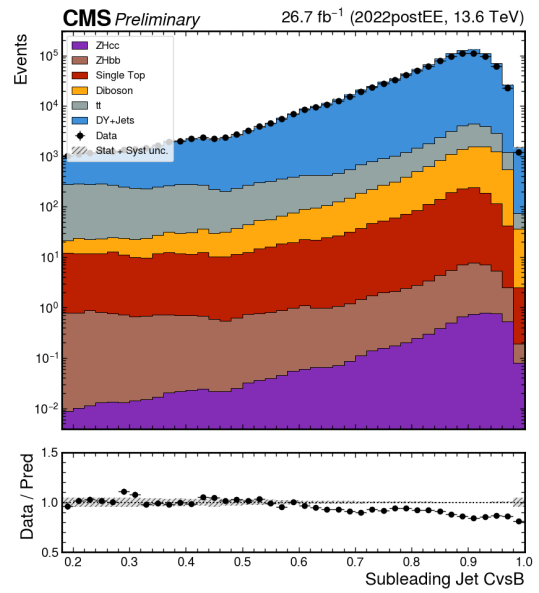
Figuur 33: CvsL-waarde van subleading jet zonder selecties



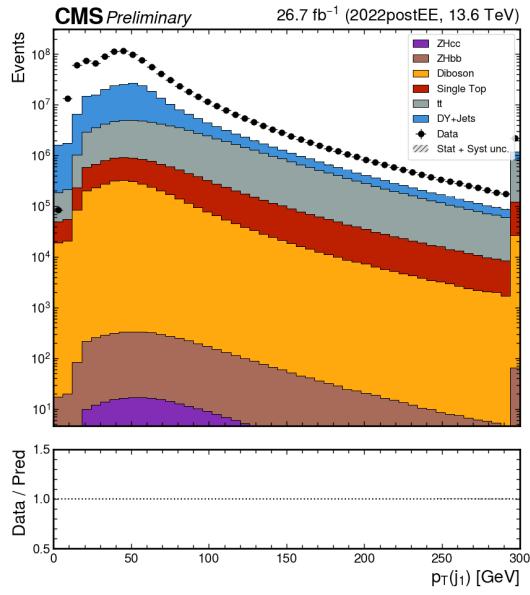
Figuur 34: CvsL-waarde van subleading jet met selecties



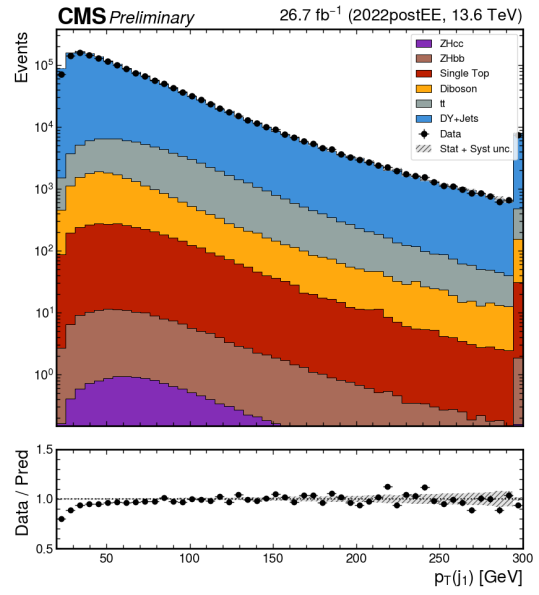
Figuur 35: CvsB-waarde van subleading jet zonder selecties



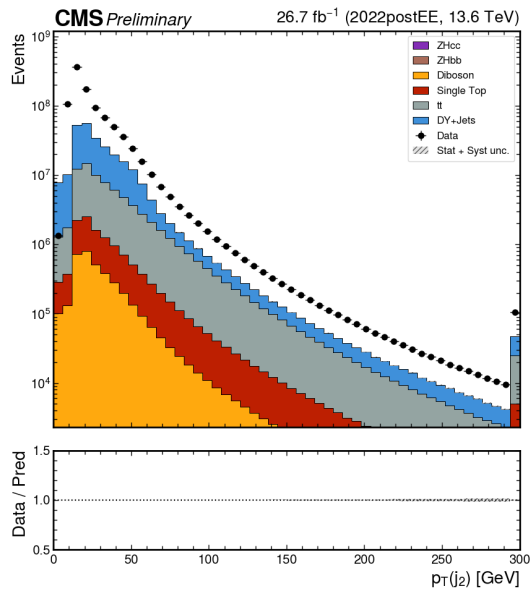
Figuur 36: CvsB-waarde van subleading jet met selecties



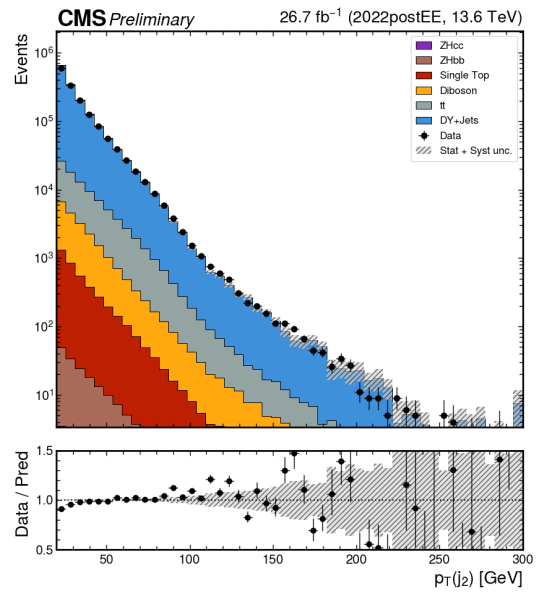
Figuur 37: Transversale impuls leading jet zonder selecties



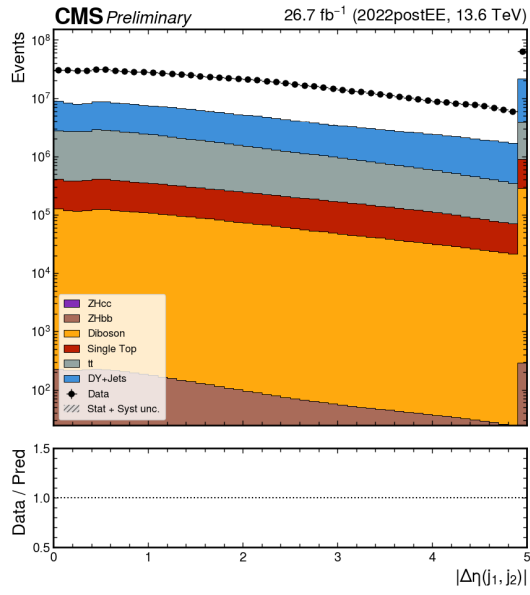
Figuur 38: Transversale impuls leading jet met selecties



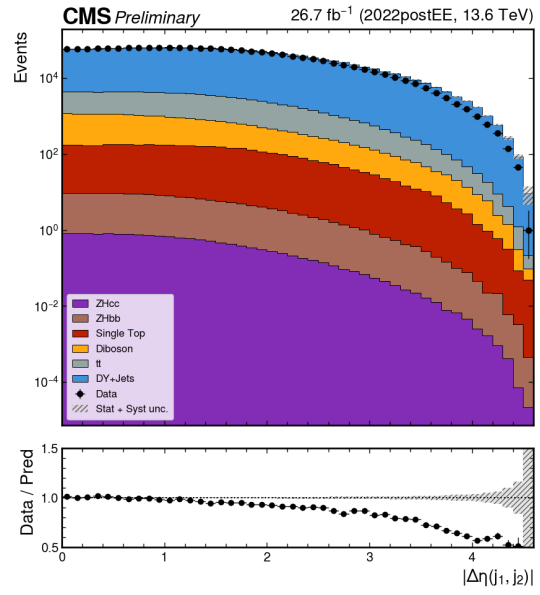
Figuur 39: Transversale impuls subleading jet zonder selecties



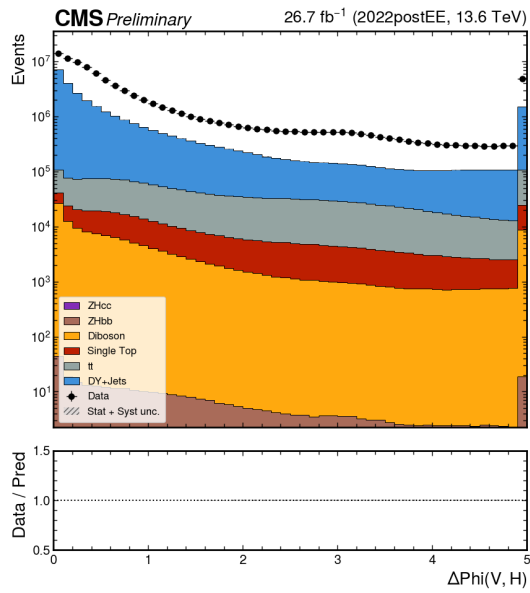
Figuur 40: Transversale impuls subleading jet met selecties



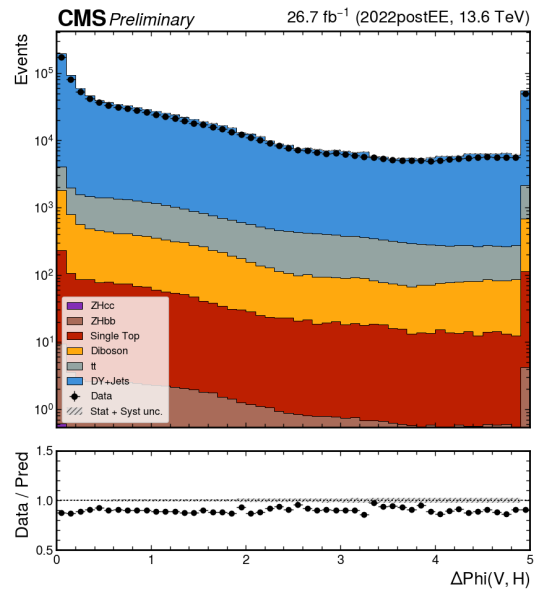
Figuur 41: Verschil in pseudorapiditeit tussen leading en subleading jet zonder selecties



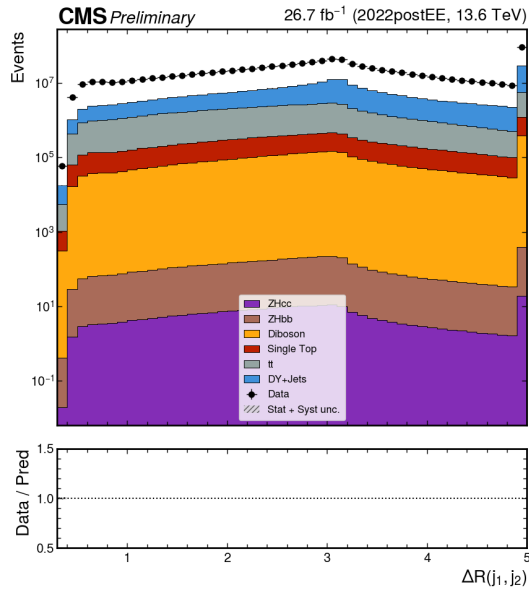
Figuur 42: Verschil in pseudorapiditeit tussen leading en subleading jet met selecties



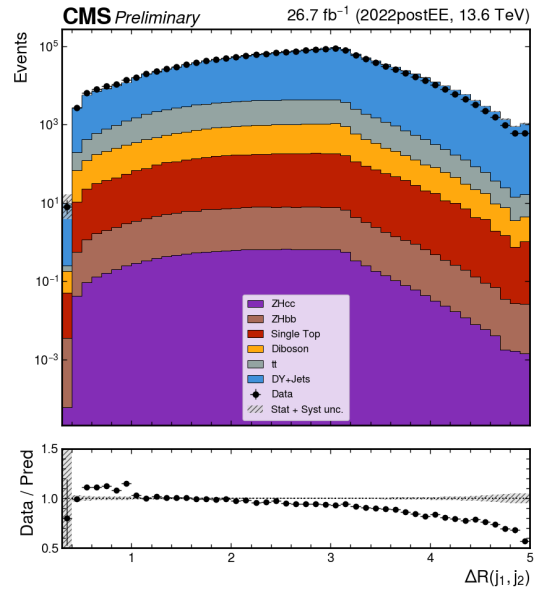
Figuur 43: Azimutale hoek tussen Z-bosonkandidaat en Higgsbosonkandidaat zonder selecties



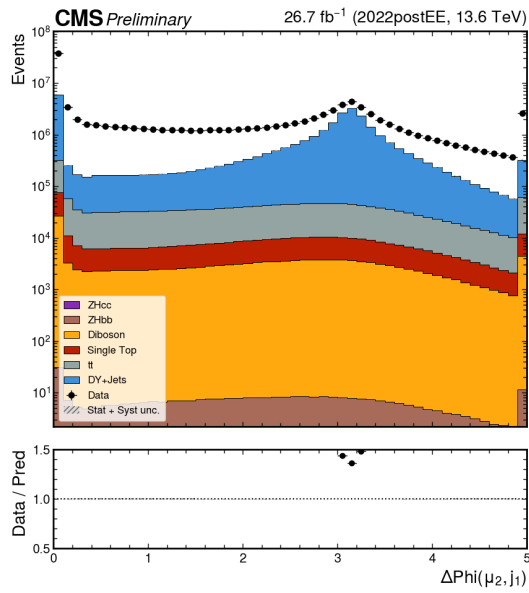
Figuur 44: Azimutale hoek tussen Z-bosonkandidaat en Higgsbosonkandidaat met selecties



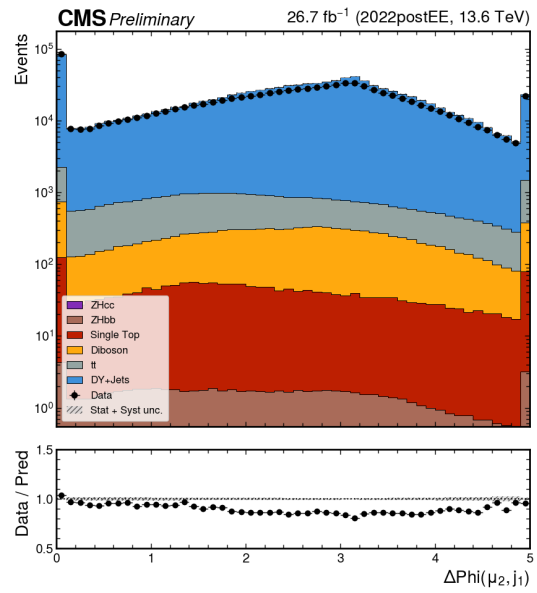
Figuur 45:  $\Delta R$  tussen leading en subleading jet zonder selecties



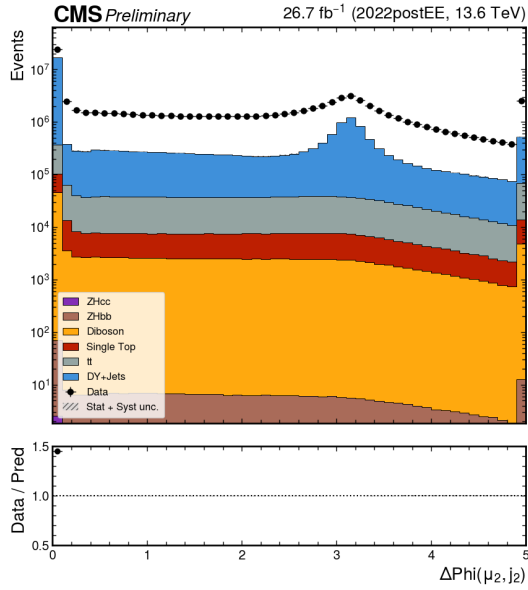
Figuur 46:  $\Delta R$  tussen leading en subleading jet met selecties



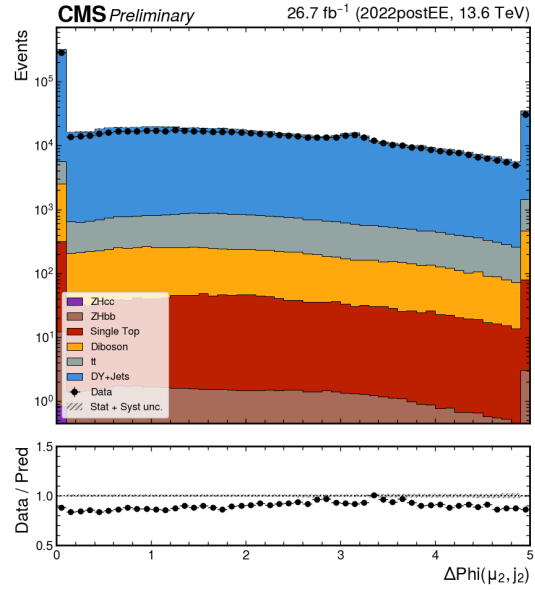
Figuur 47: Azimutale hoek tussen subleading lepton en leading jet zonder selecties



Figuur 48: Azimutale hoek tussen subleading lepton en leading jet met selecties



Figuur 49: Azimutale hoek tussen subleading lepton en subleading jet zonder selecties



Figuur 50: Azimutale hoek tussen subleading lepton en subleading jet met selecties

Voor sommige plots valt er op dat het signaal niet zichtbaar is. Dit betekent dat het signaal niet goed onderscheidbaar is van de achtergrond op basis van die observabele. Er geldt dat de verdeling van die observabele er hetzelfde uitziet voor het signaal en de achtergrond.

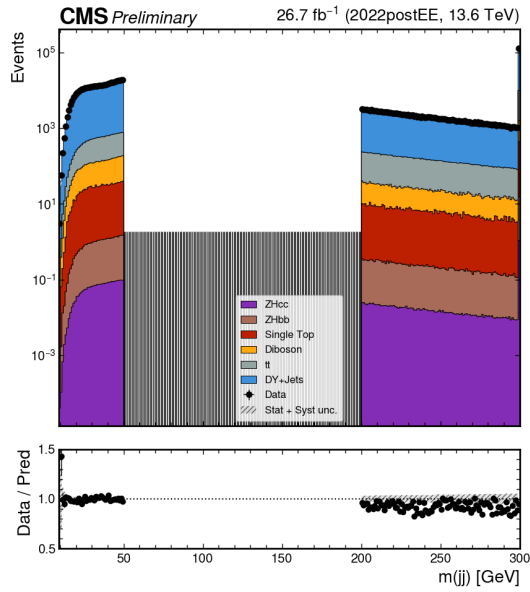
### 4.3 Controle regio's

Controle regio's (CR) zijn regio's die zo goed als geen signaal bevatten en waar één achtergrond domineert. Het construeren van controle regio's is nuttig voor de statistische analyse die later gebeurt. Deze regio's geven sterke beperkingen mee op de dominante achtergronden in de signaal regio (SR). [19]

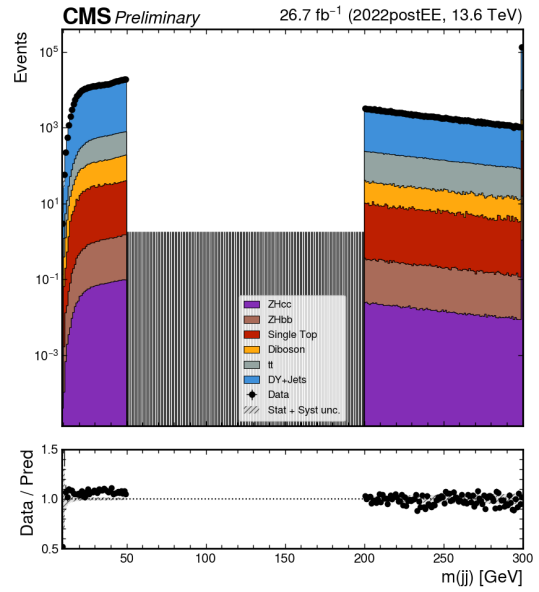
Er kunnen potentiële fouten zijn voor de normalisatie van de achtergronden in de MC samples. Door de bepaling van twee controle regio's kan de normalisatie van de voornaamste achtergronden ( $V + \text{Jets}$  en  $tt$ ) verkregen worden. [3] Voor elke controle regio kan een schaalfactor ( $SF$ ) op basis van het aantal events van de dominerende achtergrond berekend worden. [19] Dit wordt gedaan door te integreren over de volledige controle regio. Het totaal aantal MC events wordt geschaald opdat dit overeenkomt met het totaal aantal data events.

$$SF = \frac{|\# \text{ data events} - \# \text{ MC events niet-dominante achtergrond}|}{\# \text{ MC events dominante achtergrond controle regio}} \quad (29)$$

Voor de eerste controle regio (controle regio 1) is de dominerende achtergrond  $V + \text{Jets}$ . Deze wordt bekomen door een inverse selectie te maken op de invariante massa van de Higgsbosonkandidaat:  $50 > m(H)$  of  $200 < m(H)$ , zie figuur 51 en 52. [19]

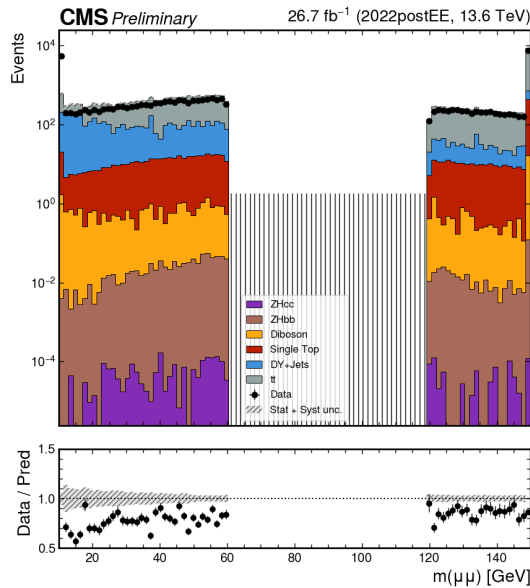


Figuur 51: Invariante massa dijets controle regio 1 zonder normalisatie

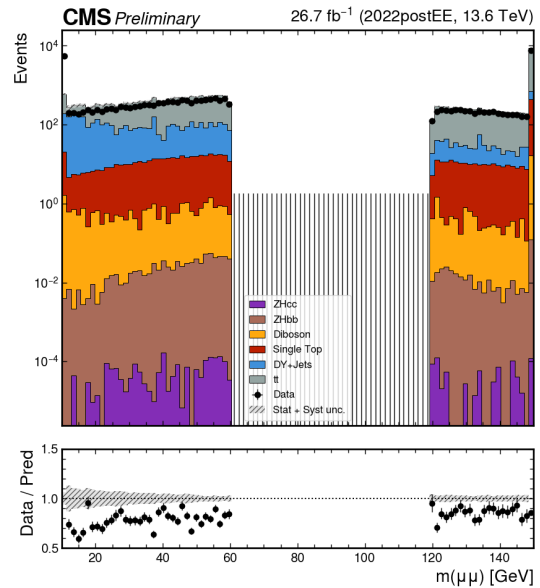


Figuur 52: Invariante massa dijets controle regio 1 met normalisatie

De tweede controle regio (controle regio 2) wordt bekomen door een selectie op de  $c$ -taggerwaarden en ook een inverse selectie op de invariante massa van de  $Z$ -bosonkandidaat:  $60 > m(Z)$  of  $120 < m(Z)$ ,  $C_{vsL}(j_1) > 0,054$  en  $C_{vsB}(j_2) < 0,182$ , zie figuur 53 en 54. De  $tt$  achtergrond is de dominerende achtergrond hier. [19]



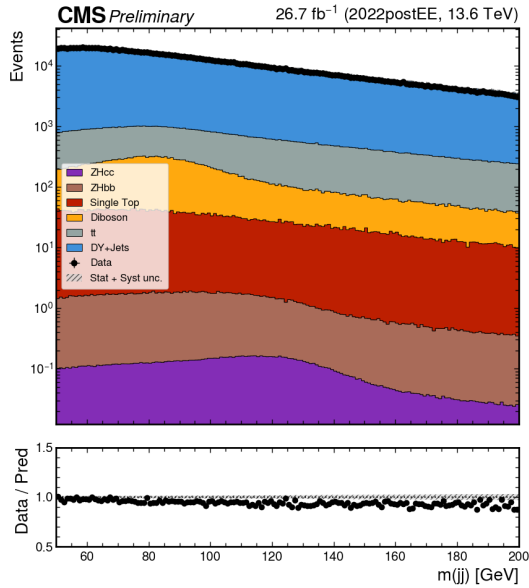
Figuur 53: Invariante massa dimuonen controle regio 2 zonder normalisatie



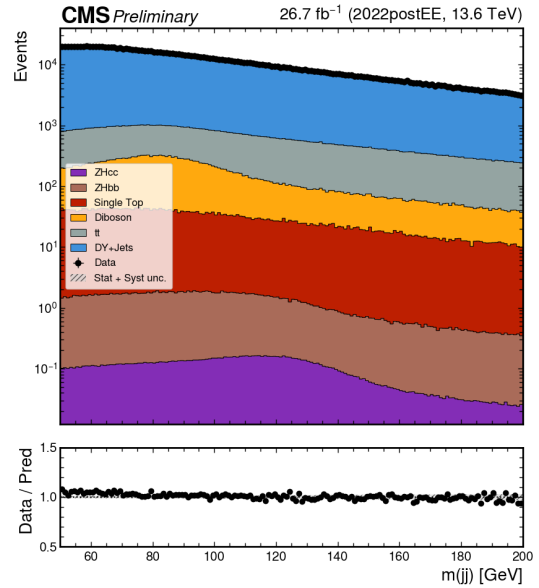
Figuur 54: Invariante massa dimuonen controle regio 2 met normalisatie

De gevonden schaalfactoren voor de dominerende achtergronden worden vervolgens

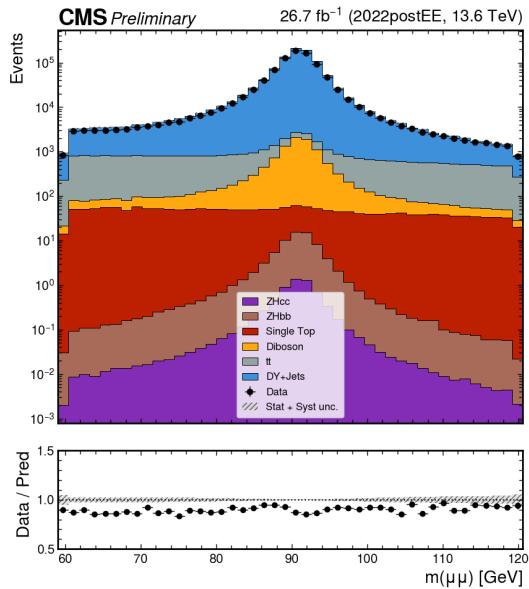
ook toegepast op de bijhorende achtergronden in de signaal regio, zie figuur 56 en 58.



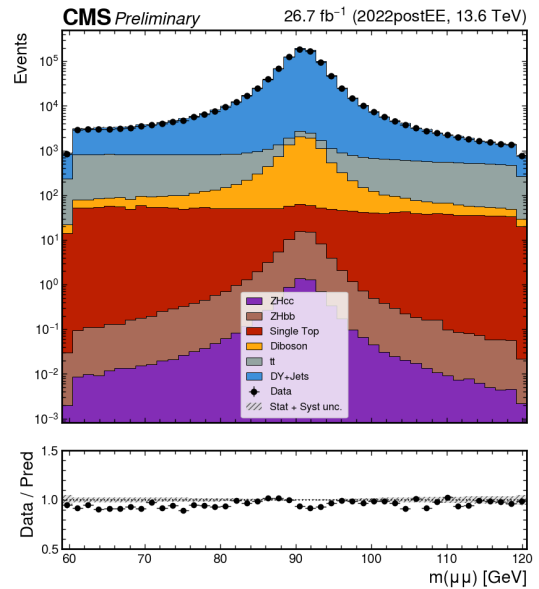
Figuur 55: Invariante massa dijets signaal regio zonder normalisatie



Figuur 56: Invariante massa dijets signaal regio met normalisatie



Figuur 57: Invariante massa dimuonen signaal regio zonder normalisatie



Figuur 58: Invariante massa dimuonen signaal regio met normalisatie

De bovenstaande figuren tonen de controle regio's met en zonder normalisatie voor het muonkanaal. Hetzelfde wordt gedaan voor het elektronkanaal in appendix B.

## 4.4 Resolutie massapieak

In deze sectie wordt er onderzocht of de resolutie van de invariante massapieak van de Higgsbosonkandidaat verder verbeterd kan worden. Er wordt jet energieregressie en kinematic fitting geïmplementeerd. Deze technieken kunnen gebruikt worden om de resolutie van de invariante massapieak van het Higgsboson te verbeteren. [3] [44] [46]

### 4.4.1 Jet energieregressie

Bij jet energieregressie wordt geprobeerd om een schatting te maken van de werkelijke  $p_T$  van de jet en tegelijk de  $p_T$  resolutie te verbeteren. Dit wordt gedaan door de jet  $p_T$  te vermenigvuldigen met correctiefactoren. Deze corrigeren de energie van de charm jets. Zo wordt bijvoorbeeld de energie onderschat door de aanwezigheid van ongedetecteerde neutrino's in de jets. Deze correcties worden bekomen aan de hand van ParticleNet [25], een dynamisch grafisch convolutioneel neurale netwerk. [3] [44]

Deze correctiefactoren zijn aanwezig in de NANO-AOD files.

### 4.4.2 Kinematic fitting

Aan de hand van een kinematic fitting methode kan een set van traject parameters geoptimaliseerd worden opdat ze zo dicht mogelijk liggen bij hun werkelijke waarden en ook voldoen aan de opgelegde beperkingen (behoudsvoorwaarden).  $\chi^2$  is de waarde die hierbij geminimaliseerd wordt: [46]

$$\chi^2 = ( \vec{y} - \vec{\eta} )^T V^{-1} ( \vec{y} ) ( \vec{y} - \vec{\eta} ) \quad (30)$$

Hierbij is  $\vec{y}$  de set opgemeten traject parameters,  $\vec{\eta}$  is de set van geschatte traject parameters, en  $V( \vec{y} )$  is de covariantiematrix. Daarbij moet er ook aan een set van beperkingen  $\vec{f}$  in functie van de geschatte traject parameters  $\vec{\eta}$  voldaan worden: [46]

$$\vec{f} = \vec{f}( \vec{\eta} ) = 0 \quad (31)$$

Dit kan gedaan worden aan de hand van de techniek van Lagrange multipliers. Op deze manier kan het minimalisatieprobleem neergeschreven worden als een minimalisatie van de Lagrange functie  $\mathcal{L}$ : [46] [47]

$$\mathcal{L} = ( \vec{y} - \vec{\eta} )^T V^{-1} ( \vec{y} ) ( \vec{y} - \vec{\eta} ) + 2 \vec{\lambda} \vec{f}( \vec{\eta} ) \quad (32)$$

Hierbij wordt er een ongekende set parameters  $\vec{\lambda}$  geïntroduceerd, die de Lagrange multipliers genoemd worden. De minimalisatie van deze functie kan gedaan worden door ze af te leiden naar de onbekende parameters  $\vec{\eta}$  en  $\vec{\lambda}$ , en daarna de afgeleiden gelijk te stellen aan nul. Daarna kunnen de bekomen gelijkheden opgelost worden

naar de gezochte  $\vec{\eta}$ : [46] [47]

$$\begin{cases} \nabla_{\eta}\chi^2 = -2V^{-1}(\vec{y} - \vec{\eta}) + 2F_{\eta}^T \vec{\lambda} = \vec{0} \\ \nabla_{\lambda}\chi^2 = 2\vec{f}(\vec{\eta}) = \vec{0} \end{cases} \quad (33)$$

Hierbij is  $F_{\eta}$  de Jacobiaan, de afgeleide van de beperking naar  $\vec{\eta}$ . Dit leidt tot de volgende set van gelijkheden die opgelost moet worden: [46] [47]

$$\begin{cases} V^{-1}(\vec{\eta} - \vec{y}) + F_{\eta}^T \vec{\lambda} = \vec{0} \\ \vec{f}(\vec{\eta}) = \vec{0} \end{cases} \quad (34)$$

Dit is echter een niet-lineair probleem en de oplossing kan in het algemeen gevonden worden door een iteratief proces toe te passen. Aan de hand van de waarden bekomen in de  $\nu$ -de iteratiestap ( $\vec{\eta}^{\nu}$ ,  $\vec{\lambda}^{\nu}$ ), kunnen de waarden voor de  $(\nu + 1)$ -ste iteratiestap bekomen worden ( $\vec{\eta}^{\nu+1}$ ,  $\vec{\lambda}^{\nu+1}$ ). Het iteratief proces blijft doorgaan tot er aan één van de convergentiecriteria voldaan wordt. [46] [47]

De kinematic fitting wordt gedaan aan de hand van opgemeten parameters, namelijk de traject parameters van zowel de twee jets ( $j_1, j_2$ ) en de twee muonen ( $\mu_1, \mu_2$ ) die respectievelijk het Higgsboson en het Z-boson opbouwen. De traject parameters die worden meegegeven zijn  $p_T$ ,  $\eta$  en  $\phi$ . Ook wordt de impuls in de x- en y-richting meegegeven voor de missende energie ( $p(miss)$ ) en voor de remnant jets ( $p(r)$ ). De remnant jets zijn alle overige jets naast de leading en subleading jet die een  $p_T$  groter dan 20 GeV hebben.

Voor de beperkingen geldt er dat het Z-boson on-shell moet zijn, en ook het behoud van vierimpuls in de x- en y-richting moet gelden.

$$\begin{cases} p_x(\mu_1) + p_x(\mu_2) + p_x(j_1) + p_x(j_2) + p_x(r) + p_x(miss) = 0 \\ p_y(\mu_1) + p_y(\mu_2) + p_y(j_1) + p_y(j_2) + p_y(r) + p_y(miss) = 0 \\ 2p_T(\mu_1)p_T(\mu_2)[\cosh(\eta(\mu_1) - \eta(\mu_2)) - \cos(\phi(\mu_1) - \phi(\mu_2))] - m_Z^2 = 0 \\ m_Z = 91,1876 \end{cases} \quad (35)$$

Hierbij geldt er: [48]

$$\begin{cases} p_x = p_T \cos \phi \\ p_y = p_T \sin \phi \end{cases} \quad (36)$$

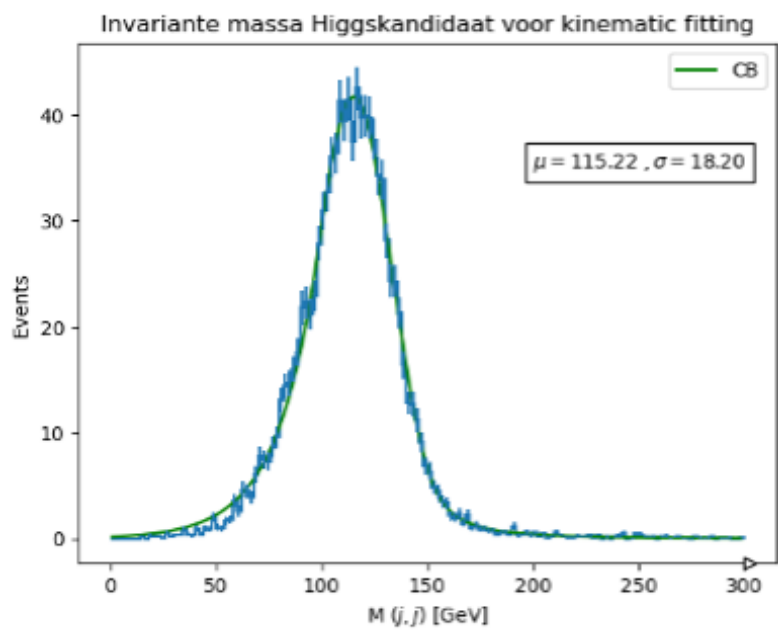
De invloed van de kinematic fit en de jet energieregressie wordt onderzocht voor de MC data voor het signaal (met een charm quark-antiquarkpaar). Dit wordt gedaan met behulp van de Python KinFit implementatie. [49] Er wordt telkens voor en na de implementatie een asymmetrische dubbelzijdige crystal ball functie ( $f_{CB}$ ) gefit voor de piek van de invariante massa. Deze functie wordt beschreven door een Gaussfunctie in het centrum (met gemiddelde  $\mu$  en standaarddeviatie  $\sigma$ ) en een machtsfunctie bij de staarten (met macht  $n$ ).  $\alpha$  beschrijft telkens de positie

waar de machtsfunctie overgaat naar een Gaussfunctie en omgekeerd. [50] [51]

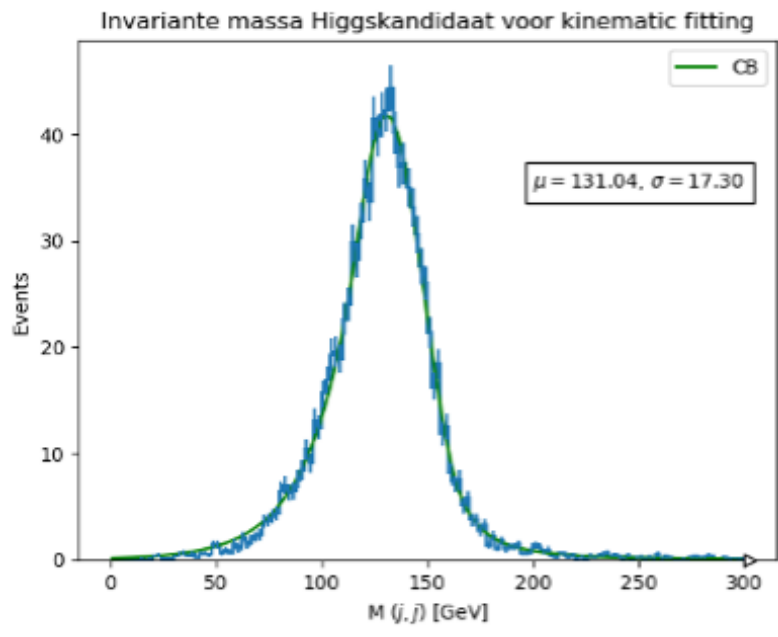
$$f_{CB} = \begin{cases} N \cdot A_L \cdot (B_L - \frac{x-\mu}{\sigma})^{-n_L}, & \text{voor } \frac{x-\mu}{\sigma} < -\alpha_L \\ N \cdot \exp(-\frac{(x-\mu)^2}{2\sigma^2}), & \text{voor } -\alpha_L \leq \frac{x-\mu}{\sigma} \leq \alpha_R \\ N \cdot A_R \cdot (B_R + \frac{x-\mu}{\sigma})^{-n_R}, & \text{voor } \frac{x-\mu}{\sigma} > \alpha_R \end{cases} \quad (37)$$

Hierbij geldt er: [50] [51]

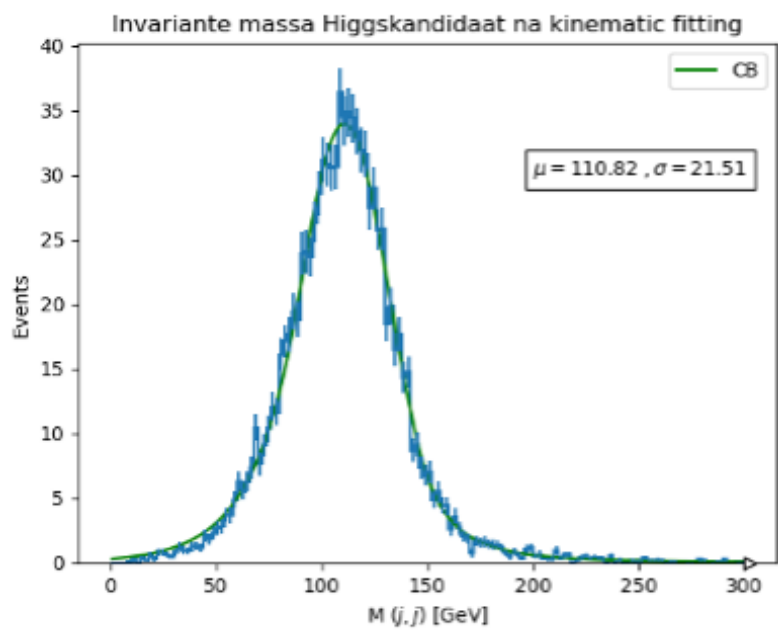
$$\begin{cases} A_{L/R} = (\frac{n_{L/R}}{|\alpha_{L/R}|})^{n_{L/R}} \cdot \exp(-\frac{|\alpha_{L/R}|^2}{2}) \\ B_{L/R} = \frac{n_{L/R}}{|\alpha_{L/R}|} - |\alpha_{L/R}| \end{cases} \quad (38)$$



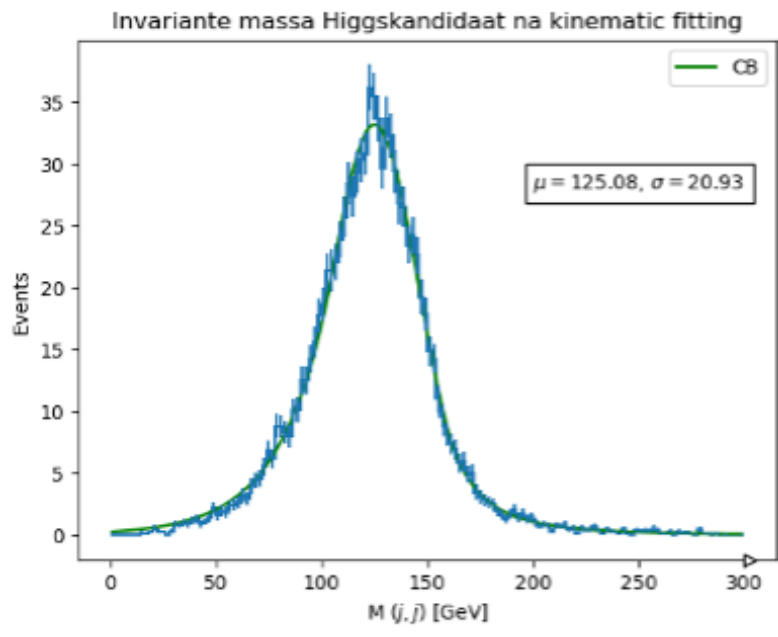
Figuur 59: Dijet invariante massa



Figuur 60: Dijet invariante massa met jet energieregressie



Figuur 61: Dijet invariante massa met kinematic fitting



Figuur 62: Dijet invariante massa met jet energieregressie en kinematic fitting

	gemiddelde ( $\mu$ )	afwijking ( $\sigma$ )
zonder regressie, zonder kinematic fitting	115,22	18,20
met regressie, zonder kinematic fitting	131,04	17,30
zonder regressie, met kinematic fitting	110,82	21,51
met regressie, met kinematic fitting	125,08	20,93

Tabel 3: Resultaten jet energieregressie en kinematic fitting

De invloed van het toepassen van de jet energieregressie en de kinematic fitting wordt weergegeven in tabel 3. Het toepassen van de jet energieregressie geeft een verbetering van zowel het gemiddelde (deze ligt dicht bij de werkelijke massa van het Higgsboson) en ook de resolutie van de invariante massapijk. Het toepassen van kinematic fitting zorgt telkens voor een verslechtering van de resolutie van de invariante massapijk. Echter, samen met de jet energieregressie levert kinematic fitting het beste resultaat voor het gemiddelde van de invariante massa. Op basis hiervan is er besloten om de jet energieregressie toe te passen bij de berekening van alle observabelen. De resultaten van de kinematic fitting worden niet meegenomen bij het trainen van de BDT.

## 4.5 Boosted decision tree

Voor de signaal regio wordt er een boosted decision tree (BDT) getraind om het signaal beter te onderscheiden van de achtergrond. [19] De samples worden geschaald aan de hand van hun bijhorende cross-sections en opgesplitst in twee subsets. [19] Namelijk 30% van zowel signaal als achtergrond samples worden gebruikt voor het

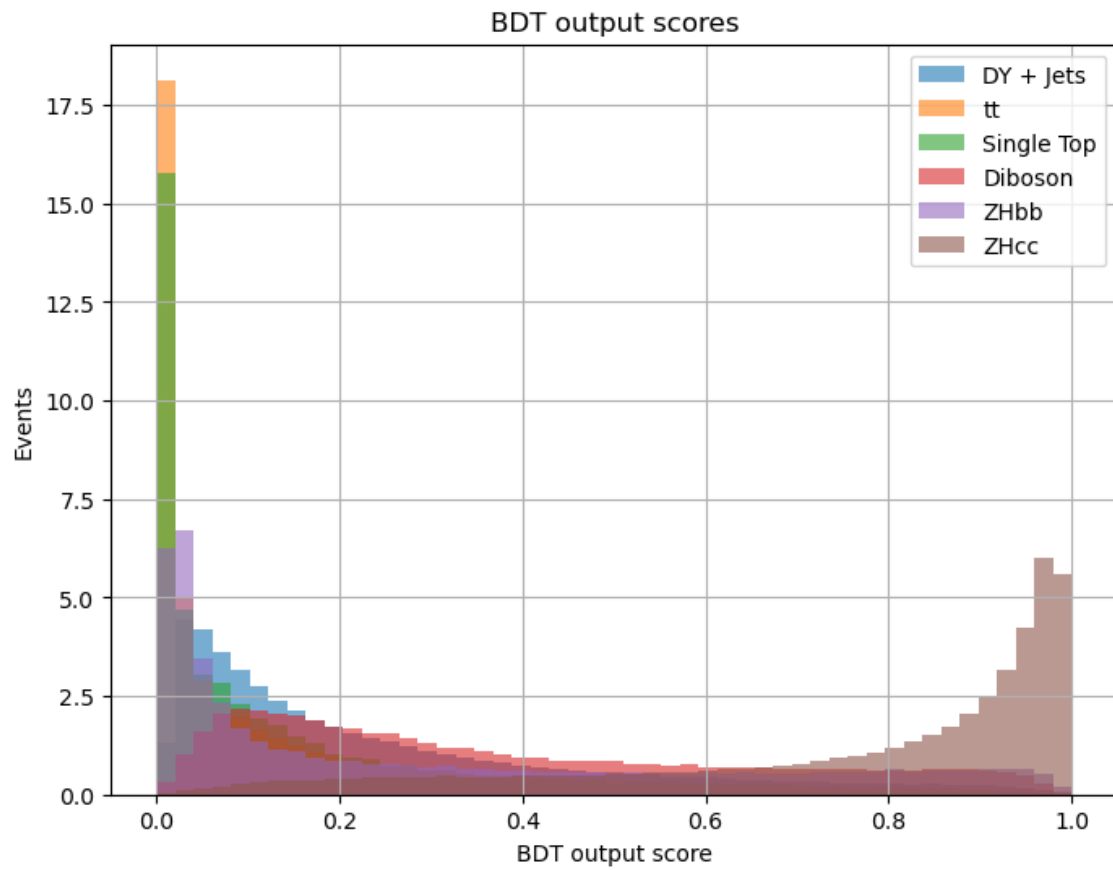
trainen, en 70% wordt gebruikt voor het testen van de BDT. De vorm van de BDT output kan dan gebruikt worden in de finale analyse. [52]

Decision trees zijn binaire classificatoren die het signaal en achtergrond kunnen onderscheiden. Ze hebben een boomstructuur en bij elk knooppunt worden de samples gescheiden op basis van een bepaald criterium op de inputvariabelen; bij elk knooppunt wordt er een beslissing gemaakt of een event eerder lijkt op signaal of achtergrond. De output van de BDT is dan een score horende bij het betreffende event en is een maat voor hoe waarschijnlijk het event signaal of achtergrond is, zie figuur 63. Wanneer de outputscore van een event dichtbij 1 is, is het eerder signaal, en als het dichtbij 0 is, is het eerder achtergrond. Bij boosted decision trees worden meerdere van deze decision trees gecombineerd. Op deze manier kan de prestatie verbeterd worden. [52]

Om de BDT te construeren wordt lightGBM [53] gebruikt. De inputvariabelen die gebruikt worden om de BDT te trainen worden opgesomd in tabel 4.

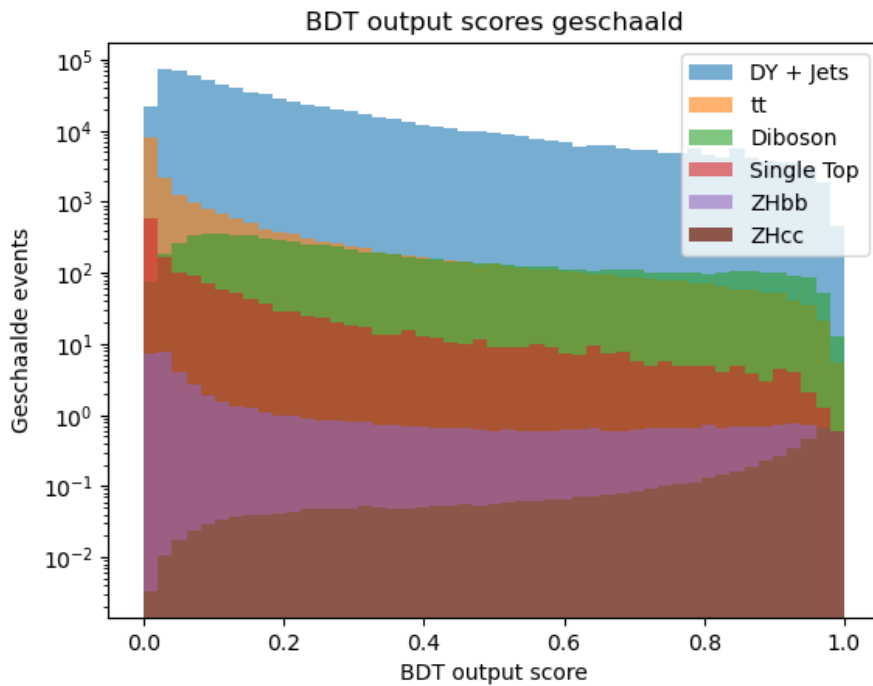
Variabele	Beschrijving
$m(jj)$	Invariante massa dijets
$p_T(jj)$	Transversale impuls dijets
$m(ll)$	Invariante massa dileptonen
$p_T(ll)$	Transversale impuls dileptonen
$p_T(ll)/p_T(jj)$	Ratio van transversale impuls van dileptonen en dijets
CvsL ( $j_1$ )	CvsL-waarde van leading jet
CvsB ( $j_1$ )	CvsB-waarde van leading jet
CvsL ( $j_2$ )	CvsL-waarde van subleading jet
CvsB ( $j_2$ )	CvsB-waarde van subleading jet
$p_T(j_1)$	Transversale impuls leading jet
$p_T(j_2)$	Transversale impuls subleading jet
$\Delta\phi(V, H)$	Azimutale hoek tussen Z-bosonkandidaat en Higgsbosonkandidaat
$\Delta R(j_1, j_2)$	$\Delta R$ tussen leading en subleading jet
$\Delta\eta(j_1, j_2)$	Verschil in pseudorapiditeit tussen leading en subleading jet
$\Delta\phi(l_2, j_1)$	Azimutale hoek tussen subleading lepton en leading jet
$\Delta\phi(l_2, j_2)$	Azimutale hoek tussen subleading lepton en subleading jet

Tabel 4: BDT inputvariabelen



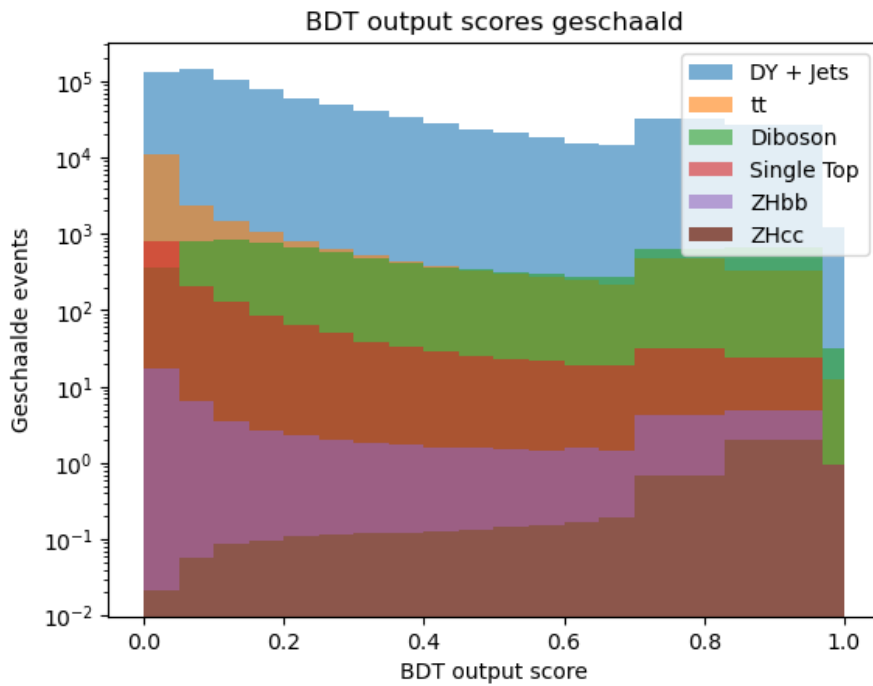
Figuur 63: Ongeschaalde outputscores getrainde BDT

Er moet nog rekening gehouden worden met de cross-sectie en gewichten voor elke achtergrond. De geschaalde BDT outputscores worden weergegeven in figuur 64.



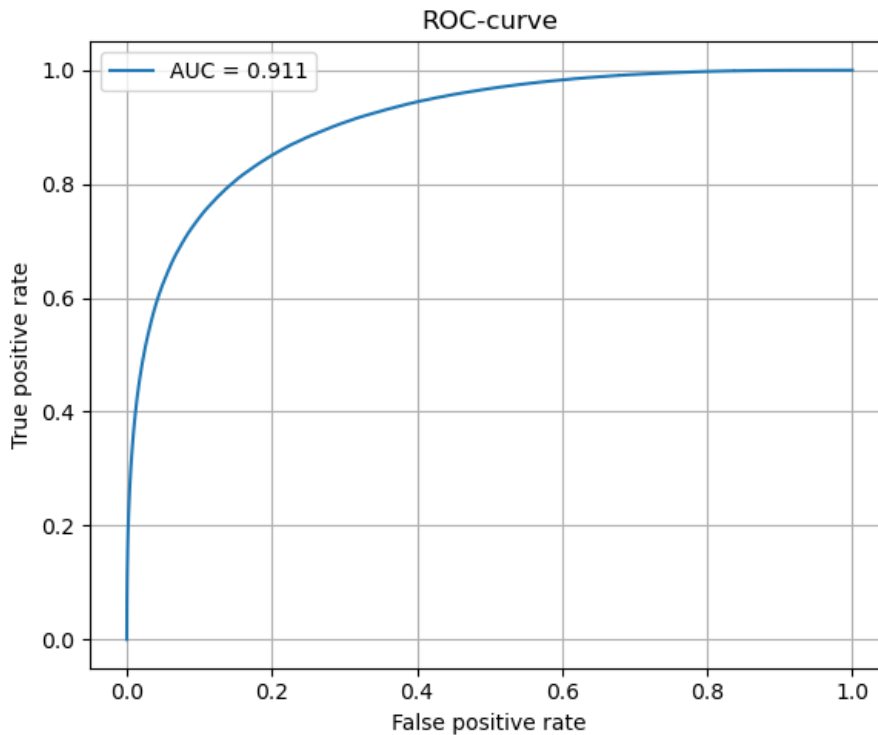
Figuur 64: Geschaalde outputscores getrainde BDT in logaritmische schaal

De breedte van de laatste bins (dichtbij outputscore 1) worden aangepast, zie figuur 65. De breedtes worden gekozen door naar de verhouding van het signaal op de achtergrond te kijken en deze proberen te verbeteren. In principe zou in de toekomst hiervoor een betere optimalisatie moeten gebeuren.



Figuur 65: Geschaalde outputscores getrainde BDT met aangepaste binning in logaritmische schaal

De ROC (receiver operating characteristic) -curve geeft het onderscheidend vermogen van de BTD weer, zie figuur 66. De ROC-curve plot de true positive rate (een maat voor de proportie van echte signaal events die de BDT ook classificeert als signaal) ten opzichte van de false positive rate (een maat voor de proportie van echte achtergrond events die de BDT echter classificeert als signaal). Voor ROC-curves kan de oppervlakte onder de curve (AUC) berekend worden. Voor een BDT die het signaal en achtergrond perfect kan onderscheiden geeft dit 1, voor een random gok geeft dit 0,5. [52]



Figuur 66: ROC-curve getrainde BDT

De BDT heeft een AUC van 0,91 en kan het signaal dus goed onderscheiden van de achtergrond aan de hand van de gekozen inputvariabelen.

De figuren voor de getrainde BDT in deze sectie behoren tot het muonkanaal. Er is op dezelfde manier ook een BDT getraind voor het elektronkanaal en de resultaten hiervoor zijn terug te vinden in appendix C.

## 4.6 Statistische analyse

De BDT die getraind is op de signaal regio, wordt toegepast op de twee controle regio's. Er wordt dan een binned maximum likelihood fit uitgevoerd op de BDT outputscores van tegelijk de signaal regio en de controle regio's. [19]

Voor de statistische analyse wordt het software framework Combine [4] gebruikt. AsymptoticLimits is de statistische methode die hierbij gebruikt wordt. Deze me-

thode bepaalt de bovenlimiet op de signaalsterkte en wordt berekend aan de hand van de ratio tussen de likelihood met signaalsterkte 0 (enkel achtergrond) ten opzichte van een signaalsterkte die vrij kan variëren, zie vergelijking 39. [54]

$$-2 \ln \frac{\mathcal{L}(\mu = 0, \hat{\nu}(0))}{\mathcal{L}(\hat{\mu}, \hat{\nu})} \quad (39)$$

Hierbij is  $\mu$  de signaalsterkte en  $\nu$  zijn de modelparameters, de systematische onzekerheden. [55] De signaalsterkte schaal het geobserveerde signaal ten opzichte van de voorspelling door SM. [2] De likelihood  $\mathcal{L}$  geeft de waarschijnlijkheid dat een dataset wordt geobserveerd voor een bepaald model. [55] Voor een binned likelihood wordt dit gegeven door een Poisson verdeling per bin. [55] De likelihoodfunctie wordt weergegeven in vergelijking 40. Hierbij is  $N_i$  het totaal aantal opgemeten events,  $s_i$  is het verwachte aantal signaal events en  $b_i$  is het verwachte aantal achtergrond events in bin  $i$ . De signaalsterkte  $\mu$  is dus een parameter van de likelihoodfunctie. [56]

$$\mathcal{L}(\mu) = \prod_{i \in \text{bins}} \text{Pois}(N_i | \mu s_i + b_i) = \prod_{i \in \text{bins}} \frac{(\mu s_i + b_i)^{N_i}}{N_i!} e^{-(\mu s_i + b_i)} \quad (40)$$

Voor een binned shaped analyse wordt een histogram, dat de outputscores van de BDT bevat, als input meegegeven. Er wordt een blinde analyse gedaan. Dit betekent dat de geobserveerde data wordt afgedekt en er enkel gebruikgemaakt wordt van MC samples.

## 4.7 Systematische onzekerheden

Naast statistische onzekerheden zijn er ook enkele systematische onzekerheden die in rekening gebracht moeten worden. Systematische effecten kunnen de vorm of de normalisatie van de BDT outputscores beïnvloeden. [19] Ze worden toegevoegd in de analyse via nuisance parameters. [3]

Een belangrijke experimentele systematische onzekerheid komt van het gelimiteerde aantal MC samples. Er zijn ook experimentele effecten van de onzekerheid op de jet energieschaal en de jet energieresolutie. Deze onzekerheden variëren typisch tussen 5 en 10 %. Er zijn ook systematische onzekerheden op de efficiënties van de reconstructie en identificatie van muonen en elektronen. Deze onzekerheden variëren tussen 1 en 4 %. Gelijkaardig zijn er ook onzekerheden voor de efficiëntie van c-tagging. Over het algemeen zijn experimentele onzekerheden afkomstig van onzekerheden op de simulaties van de detector respons en uiteindelijk de reconstructie van objecten en observabelen, de onzekerheid op de geschatte geïntegreerde luminositeit, en ook onzekerheden op de modellering van pileup interacties. Voor de geïntegreerde luminositeit wordt er een onzekerheid van 1,4 % in rekening gebracht. [3] [19] [29]

Een andere bron van systematische onzekerheden zijn theoretische onzekerheden voor de cross-secties en de vorm van  $p_T$ -spectra voor signaal en achtergrond samples. De onzekerheid op de cross-sectie wordt bepaald door onzekerheden op de schaal en de PDFs (parton distributie functies, de waarschijnlijkheidsdichtheid om

een deeltje met een bepaalde fractie van het impuls van een hadron aan te treffen bij een bepaalde resolutie [38]). Schaal onzekerheden zijn afkomstig van missende hogere orde correcties in de theoretische berekeningen. Theoretische onzekerheden worden veroorzaakt door de gelimiteerde precisie bij modelleringen en worden gelijkgesteld aan hun theoretische verwachtingen. Voor Higgsboson producties in associatie met een Z-boson (ZHcc en ZHbb) variëren deze onzekerheden tussen 1,6 en 2,4 % voor de onzekerheden op de PDFs en tussen 19 en 25 % voor de schaal. [3] [19] [29]

Er zijn twee types van systematische onzekerheden. Een eerste type (shape) beïnvloedt de vorm. Hiervoor worden er twee histograms meegegeven in de analyse. De BDT histograms worden verkregen door de parameter die onderhevig is aan een bepaalde systematische onzekerheid met één standaarddeviatie te verschuiven naar zowel boven als beneden. Een tweede type (rate) beïnvloedt enkel de normalisatie. Voor deze multiplicatieve correcties wordt de onzekerheid in rekening gebracht via de lognormale waarschijnlijkheidsverdeling. [57]

De systematische onzekerheden worden in rekening gebracht aan de hand van nuisance parameters door hun waarschijnlijkheidsdichtheden (afhankelijk van het type, shape of rate) toe te voegen als factor in de likelihoodfunctie. [29] De systematische onzekerheden die toegevoegd zijn aan de analyse, worden opgesomd in tabel 5. De systematische onzekerheden voor de c-tagging efficiënties zijn nog niet beschikbaar en zijn niet aanwezig in deze analyse. Deze onzekerheden bestaan doordat de c-tagger een verschillende prestatie heeft op data ten opzichte van MC samples. [29]

Bron van systematische onzekerheid	Type
Pileup herschaling	Shape
Lepton trigger efficiëntie	Shape
Lepton identificatie efficiëntie	Shape
Geïntegreerde luminositeit	Rate
Jet energieschaal	Shape
Jet energieresolutie	Shape
PDF	Rate
Schaal	Shape
ZH cross-sectie (PDF)	Rate
ZH cross-sectie (schaal)	Rate

Tabel 5: Systematische onzekerheden

In de praktijk zijn er nog meer systematische onzekerheden die toegevoegd moeten worden in de analyse, maar hier is er enkel rekening gehouden met de meest dominante systematische onzekerheden. Door de gelimiteerde tijd van dit project ontbreken enkele onzekerheden en valt dit buiten het opzet van deze studie.

## 5 Resultaten

Een bovenlimiet (UL) op de signaalsterkte wordt bepaald bij een betrouwbaarheidsinterval van 95 %. [19] Dit wordt gedaan voor de twee afzonderlijke vervalkanalen en de resultaten worden achteraf gecombineerd. De resultaten hiervoor worden opgelijst in tabel 6.

	UL $-1\sigma$	UL	UL $+1\sigma$
Verwachting muonkanaal	77	123	216
Verwachting elektronkanaal	118	184	306
Verwachting gecombineerd	55	81	125

Tabel 6: Resultaten statistische analyse

In dit geval is er een bovenlimiet op de signaalsterkte bepaald ervan uitgaande dat er enkel achtergrond is en geen signaal. Er is dus duidelijk dat de analyse een relatief lage gevoeligheid heeft: de bovenlimiet op de signaalsterkte is relatief groot. Er is dus een grote signaalsterkte nodig om een signaal te kunnen waarnemen dat niet wordt veroorzaakt door achtergrondfluctuaties. Deze lage gevoeligheid is voornamelijk te wijten aan het kleine aantal events voor het signaal.

Het verschil in gevoeligheid voor het muonkanaal en elektronkanaal kan verklaard worden doordat er minder events zijn voor het elektronkanaal. Dit komt door een lagere efficiëntie voor de reconstructie van ID en ISO voor elektronen. [58] De resultaten voor de combinatie van het muon- en elektronkanaal kunnen nu vergeleken worden met de bekomen waarden in eerdere studies. In tabel 7 worden de bekomen resultaten voor deze analyse (CMS Run3 bij  $26,7 \text{ fb}^{-1}$ ) vergeleken met de resultaten van een VHcc studie voor CMS Run2 data bij  $35,9 \text{ fb}^{-1}$  [19].

	UL $-1\sigma$	UL	UL $+1\sigma$
Verwachting (Run3 bij $26,7 \text{ fb}^{-1}$ )	55	81	125
Verwachting Observatie (Run2 bij $35,9 \text{ fb}^{-1}$ )	42 /	59 116	84 /

Tabel 7: Resultaten vergeleken met literatuurwaarden [19]

Het minder gevoelig resultaat voor deze analyse (Run3) ten opzichte van de vorige studie (Run2) kan gedeeltelijk verklaard worden door het verschil in geïntegreerde luminositeit. De verwachte bovenlimieten voor Run2 en Run3 lijken statistisch compatibel te zijn binnen een onzekerheidsinterval van één standaarddeviatie. Een betere gevoeligheid kan bekomen worden aan de hand van een grotere geïntegreerde luminositeit. Zo kan in een eerste volgende stap de data van 2022preEE gecombineerd worden met 2022postEE.

## 6 Conclusie en vooruitblik

In deze studie wordt het verval van het Higgsboson naar een charm quark-antiquarkpaar onderzocht, waarbij het Higgsboson geproduceerd wordt in associatie met een Z-boson. Specifiek wordt er gekeken naar het proces waarbij het Z-boson vervalst naar een paar muonen of elektronen en de jets afkomstig van de charm quarks voldoen aan een "resolved" topologie. Er wordt gebruikgemaakt van CMS Run 3 2022postEE data bij een centre-of-mass energie van  $\sqrt{s} = 13,6$  TeV. Dit komt overeen met een geïntegreerde luminositeit van  $26,7 \text{ fb}^{-1}$ .

Om het te onderzoeken signaal te onderscheiden van de achtergrond worden er enkele methodes toegepast. Ten eerste worden er selecties gedaan op enkele observabelen. Deze selecties zorgen ervoor dat regio's met veel achtergrond en weinig signaal onderdrukt worden. Ook worden regio's die niet goed beschreven worden door MC simulaties verwijderd. Vervolgens wordt er onderzocht of de resolutie van de invariante massapijk van de jets verbeterd kan worden met jet energieregressie en kinematic fitting. Jet energieregressie zorgt voor een duidelijke verbetering van de resolutie en het gemiddelde van de invariante massapijk van de dijets. Kinematic fitting geeft enkel een verbetering van het gemiddelde wanneer het wordt toegepast samen met jet energieregressie. Daarnaast worden ook enkele controle regio's bepaald waarbij het signaal zo goed als afwezig is. Dit stelt een beperking op de achtergrond in de finale analyse. Verder wordt er een BDT getraind dat gebruikmaakt van enkele observabelen als input. De BDT heeft een AUC van 0,91 voor zowel het muon- als elektronkanaal en kan het signaal dus goed onderscheiden van de achtergrond aan de hand van de gekozen inputvariabelen. In de laatste stap wordt er een statistische analyse gedaan om een verwachte bovenlimiet op de signaalsterkte te bekomen. De voornaamste systematische onzekerheden worden hierbij in rekening gebracht. Eén van de belangrijkste systematische onzekerheden die nog niet is toegevoegd, is de onzekerheid op de c-tagging efficiëntie.

De analyse wordt blind uitgevoerd voor de twee onafhankelijke vervalkanalen, namelijk het verval naar muonen en elektronen. En de resultaten worden op het einde gecombineerd. Er wordt een verwachte bovenlimiet gevonden van  $81_{-26}^{+44}$  voor de signaalsterkte bij een betrouwbaarheidsinterval van 95 %. Dit lijkt een realistisch resultaat te zijn wanneer het vergeleken wordt met de resultaten van een eerdere studie voor CMS Run2 data bij een geïntegreerde luminositeit van  $35,9 \text{ fb}^{-1}$ .

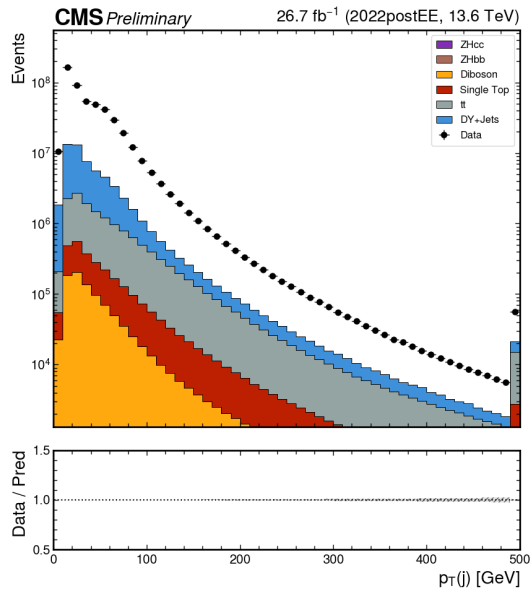
In toekomstige studies kunnen zowel 2022postEE als 2022preEE data gecombineerd worden. Dit zal zorgen voor een verhoogde geïntegreerde luminositeit wat zal leiden tot een betere gevoeligheid. Bijkomend kunnen er ook nog andere vervalkanalen toegevoegd worden voor een verbetering van de gevoeligheid. Enkele verbeteringen in de analyse zullen daarnaast ook zorgen voor een gevoeliger resultaat. Zo kan er bijvoorbeeld een verbetering van c-tagging zijn aan de hand van het nieuwe c-tagging model UPart. Het is ook belangrijk dat er in de toekomst een betere optimalisatie gebeurt voor de breedtes van de bins van de BDT output. Vervolgens kan een meer accuraat resultaat bekomen worden door het toevoegen van alle relevante systema-

tische onzekerheden. Onder andere de onzekerheid op de efficiëntie van c-tagging is een belangrijke onzekerheid die in rekening gebracht moet worden. Ten slotte kan in een volgende stap de geobserveerde data toegevoegd worden en kunnen de resultaten vergeleken worden met de SM voorspellingen.

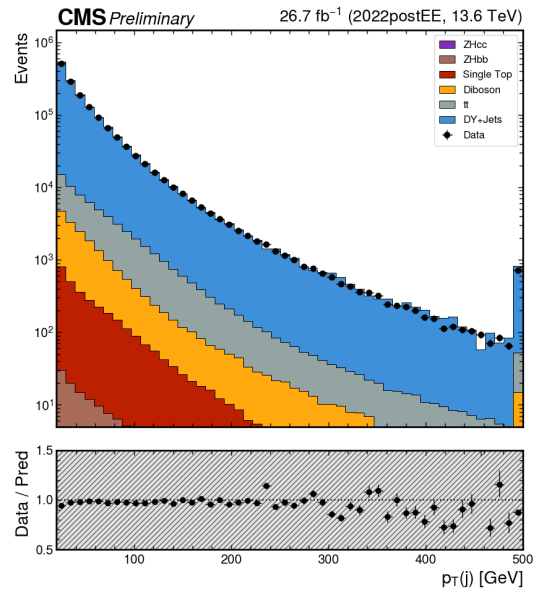
## 7 Dankwoord

Ik zou graag Prof. Dr. Pierre Van Mechelen en Dr. Tahys Janssen willen bedanken voor de mogelijkheid om onder hun begeleiding aan deze thesis te werken. Daarnaast zou ik ook graag Daniel Ocampo Henao willen bedanken voor de hulp bij het uitvoeren van deze analyse.

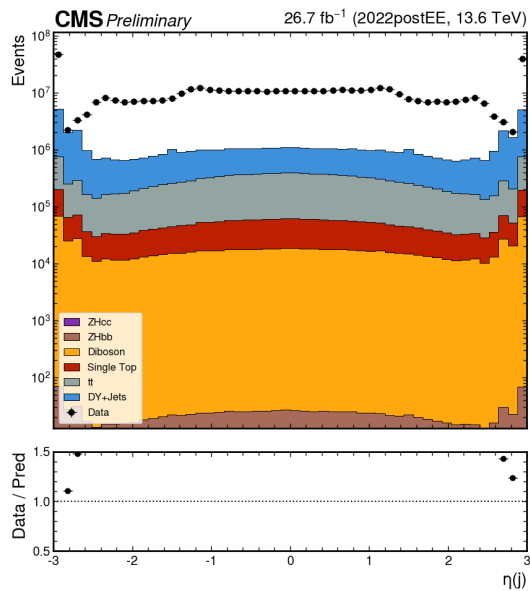
# A Plots van observabelen elektronkanaal



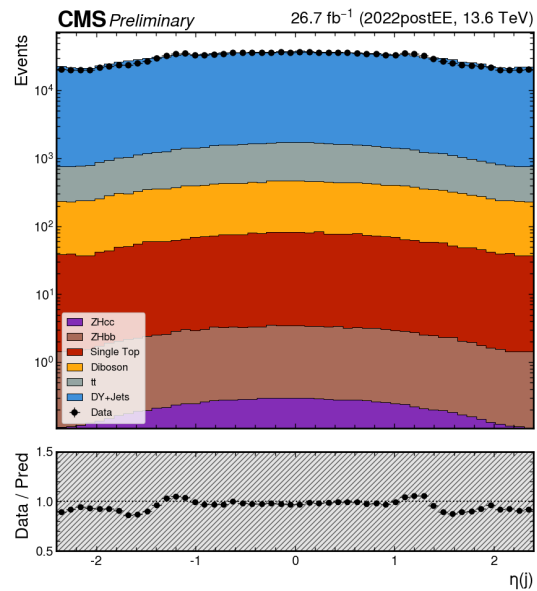
Figuur 67: Transversale impuls jets zonder selecties



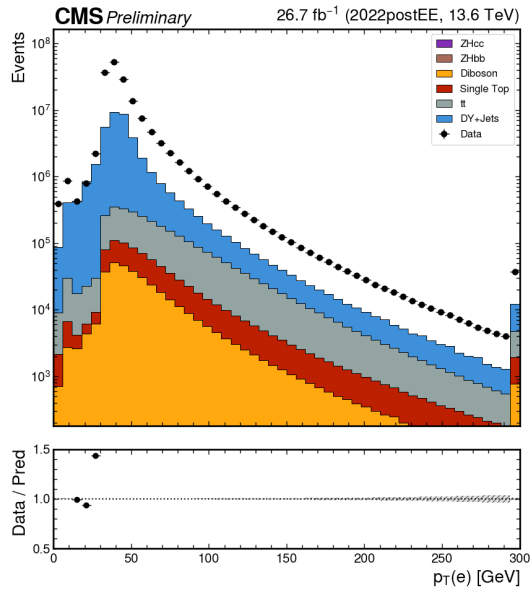
Figuur 68: Transversale impuls jets met selecties



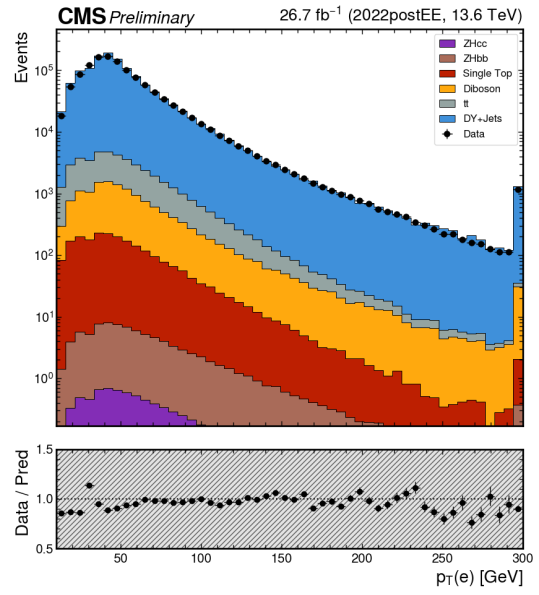
Figuur 69: Pseudorapiditeit jets zonder selecties



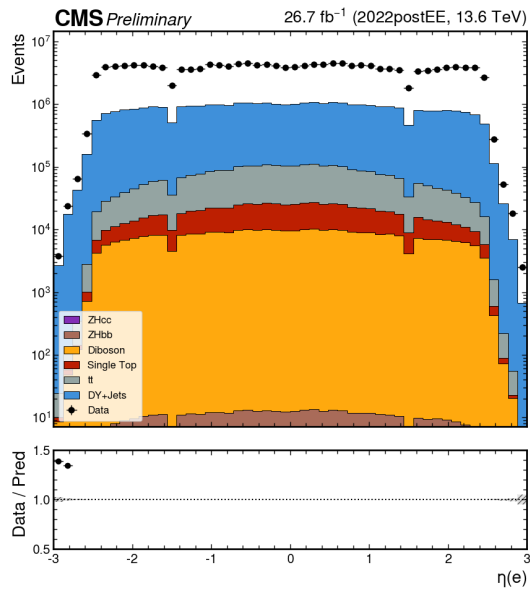
Figuur 70: Pseudorapiditeit jets met selecties



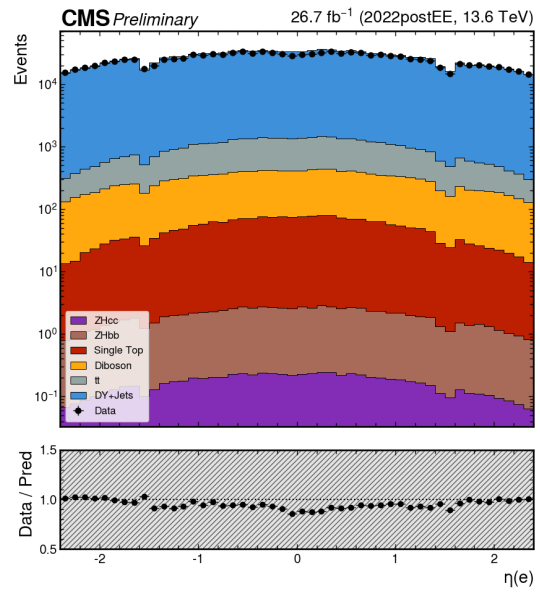
Figuur 71: Transversale impuls elektronen zonder selecties



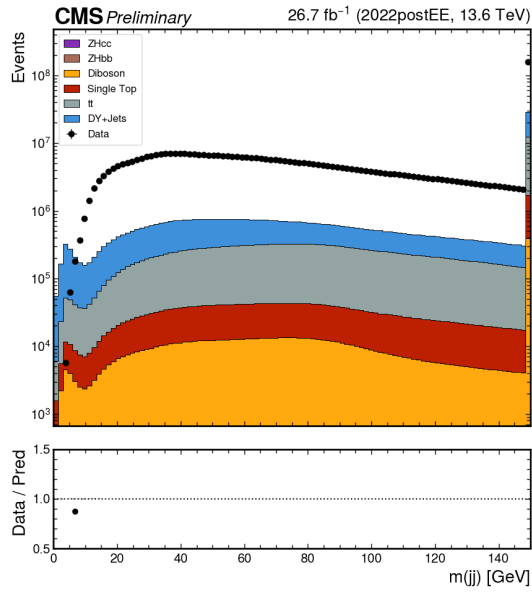
Figuur 72: Transversale impuls elektronen met selecties



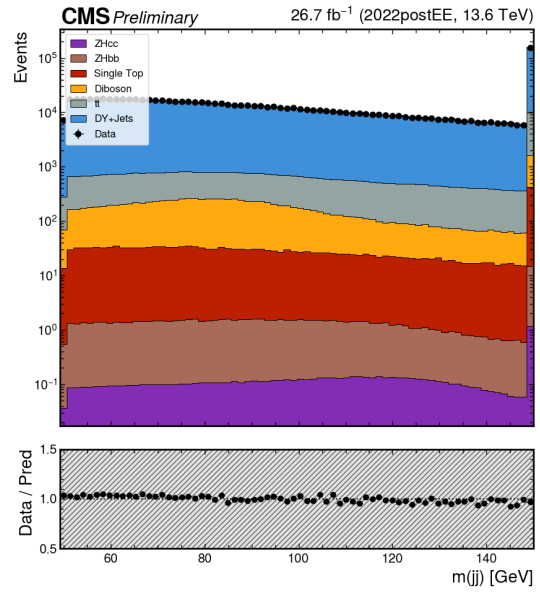
Figuur 73: Pseudorapiditeit elektronen zonder selecties



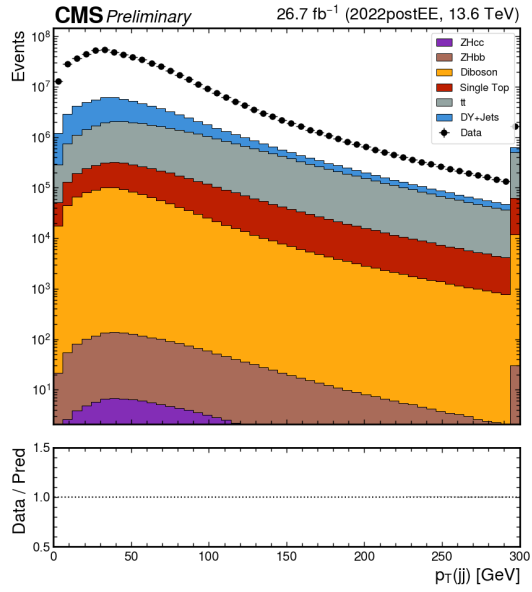
Figuur 74: Pseudorapiditeit elektronen met selecties



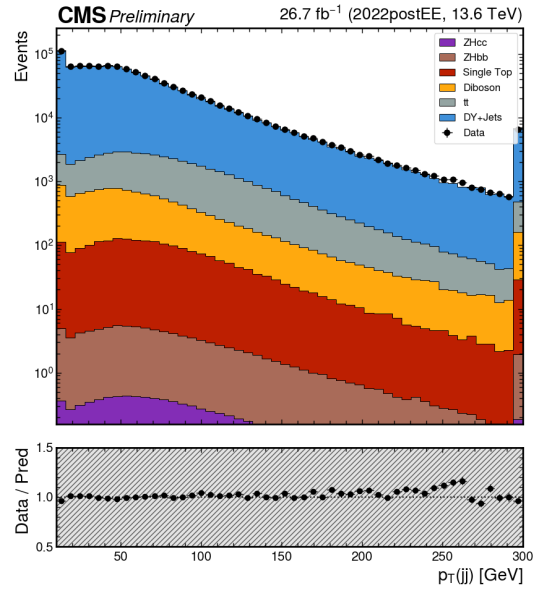
Figuur 75: Invariante massa dijets zonder selecties



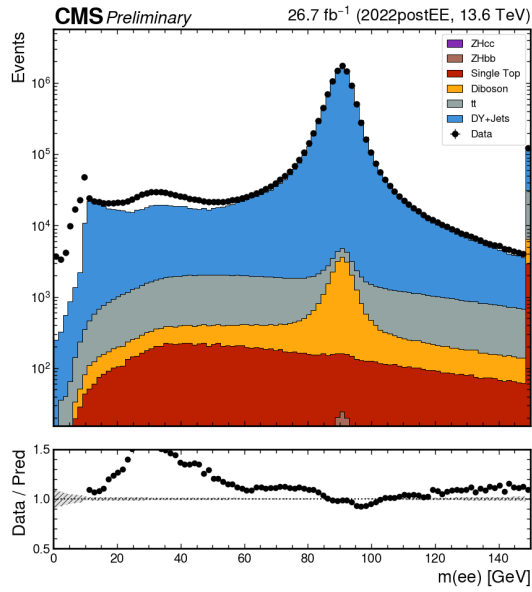
Figuur 76: Invariante massa dijets met selecties



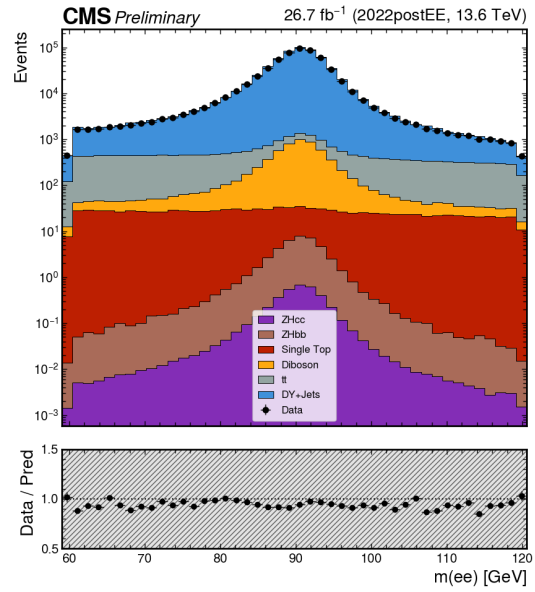
Figuur 77: Transversale impuls dijets zonder selecties



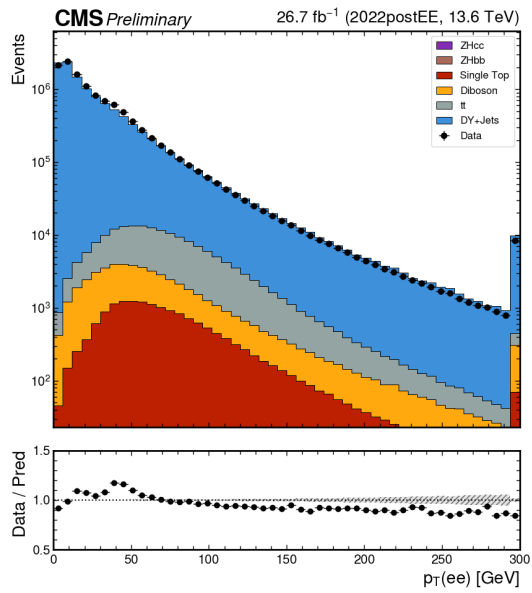
Figuur 78: Transversale impuls dijets met selecties



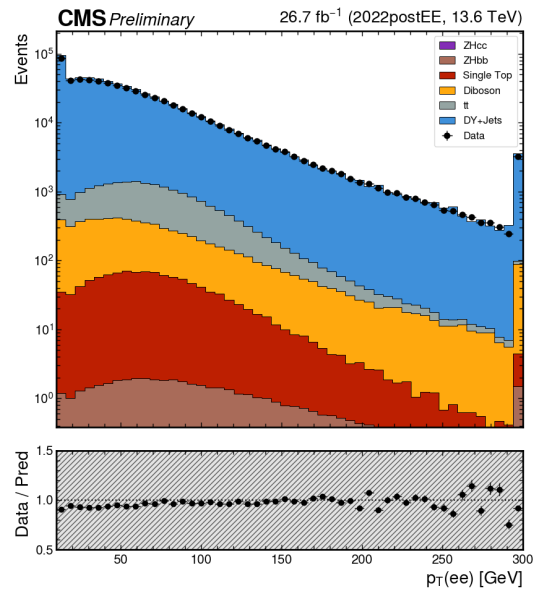
Figuur 79: Invariante massa dielektronen zonder selecties



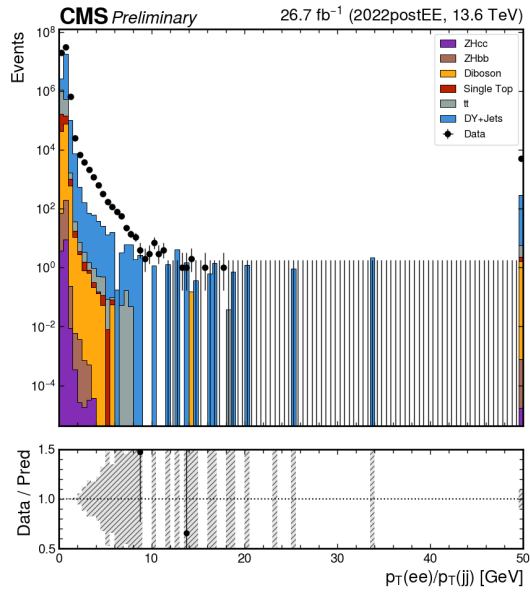
Figuur 80: Invariante massa dielektronen met selecties



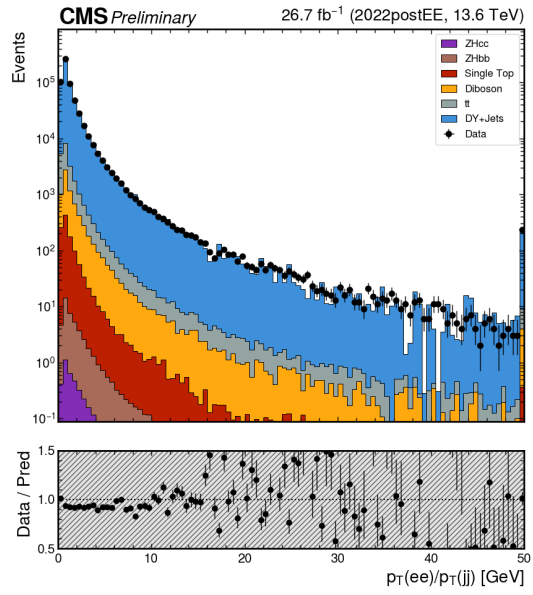
Figuur 81: Transversale impuls dielektronen zonder selecties



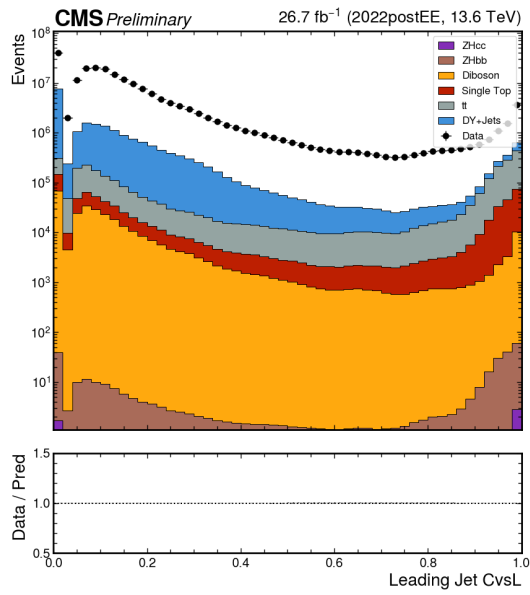
Figuur 82: Transversale impuls dielektronen met selecties



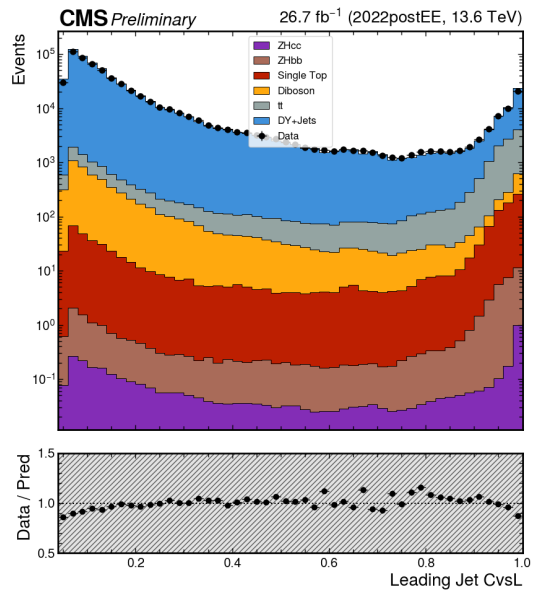
Figuur 83: Ratio van transversale impuls van dileptonen en dijets zonder selecties



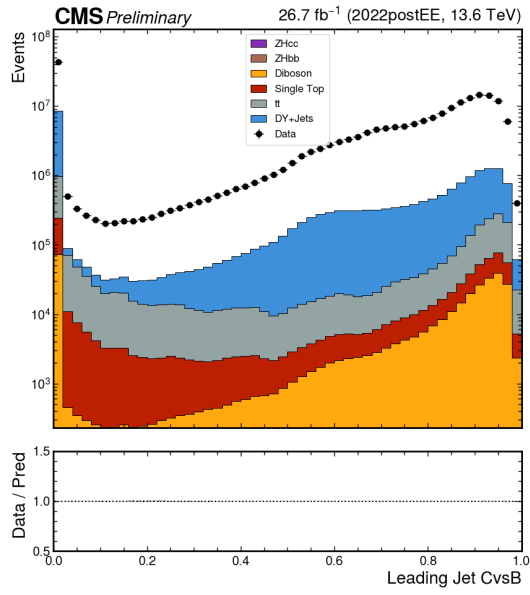
Figuur 84: Ratio van transversale impuls van dileptonen en dijets met selecties



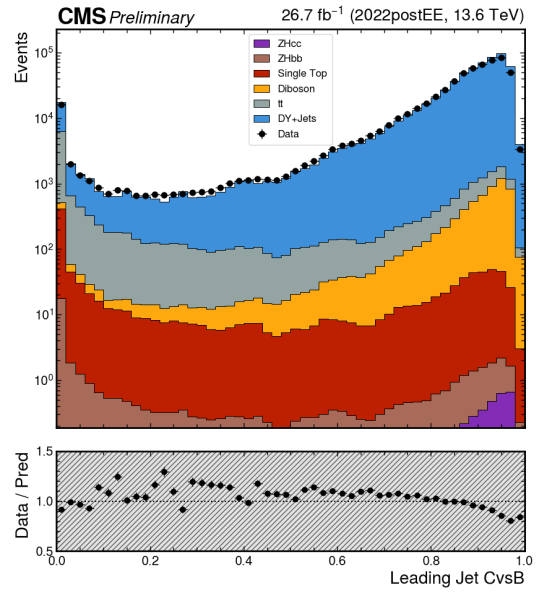
Figuur 85: CvsL-waarde van leading jet zonder selecties



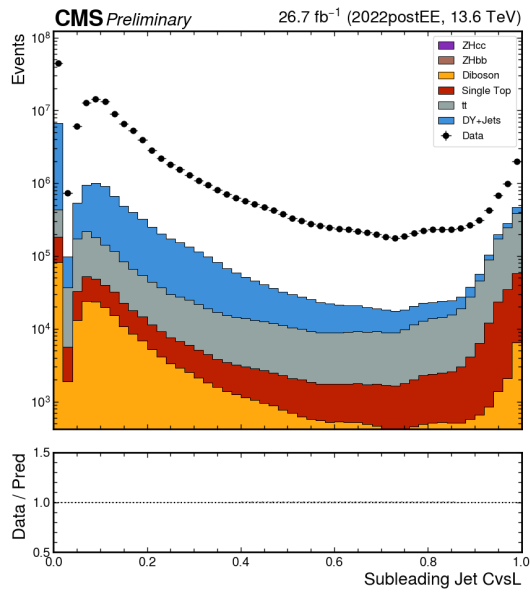
Figuur 86: CvsL-waarde van leading jet met selecties



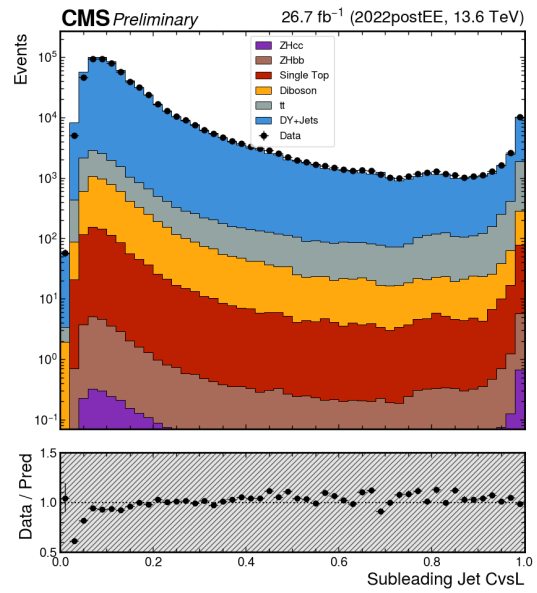
Figuur 87: CvsB-waarde van leading jet zonder selecties



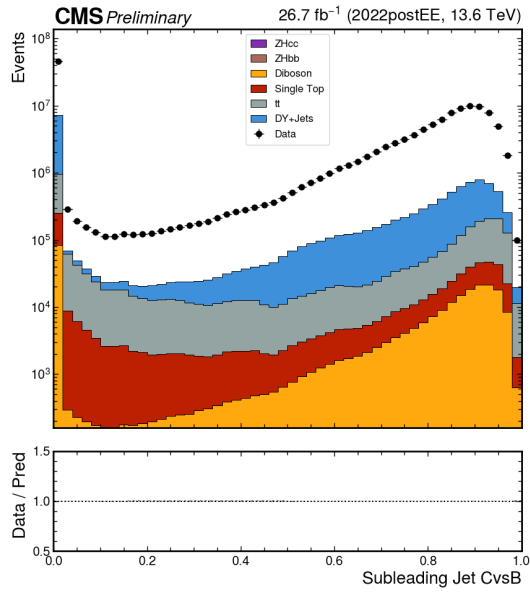
Figuur 88: CvsB-waarde van leading jet met selecties



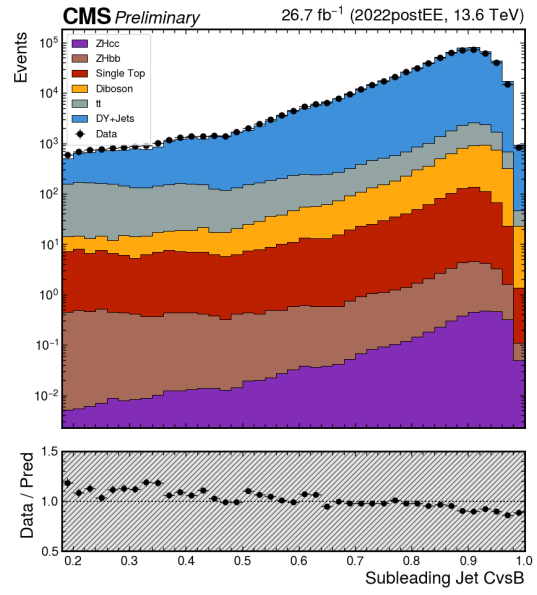
Figuur 89: CvsL-waarde van subleading jet zonder selecties



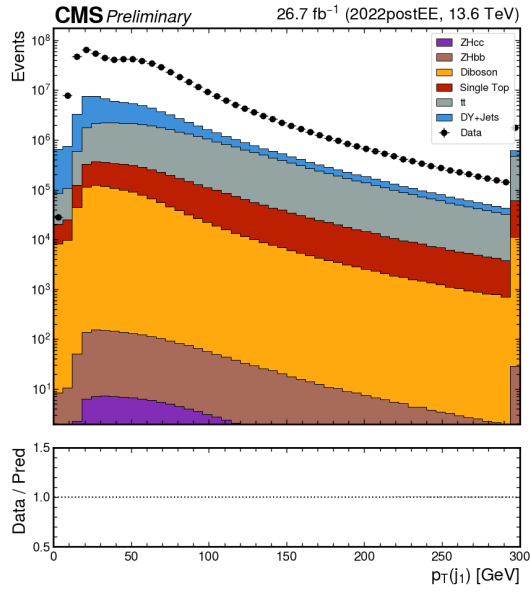
Figuur 90: CvsL-waarde van subleading jet met selecties



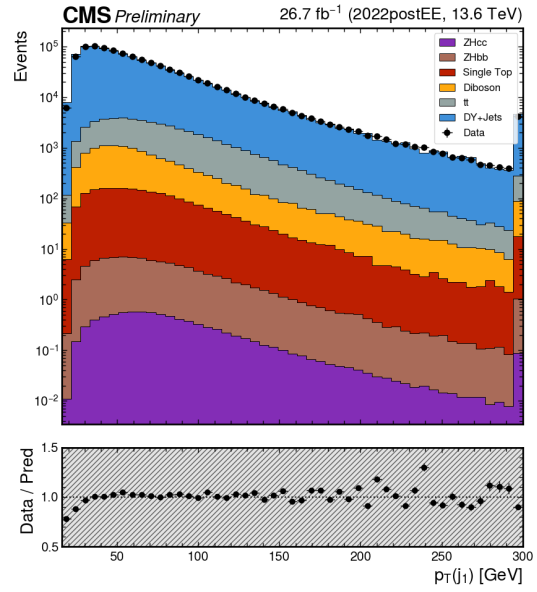
Figuur 91: CvsB-waarde van subleading jet zonder selecties



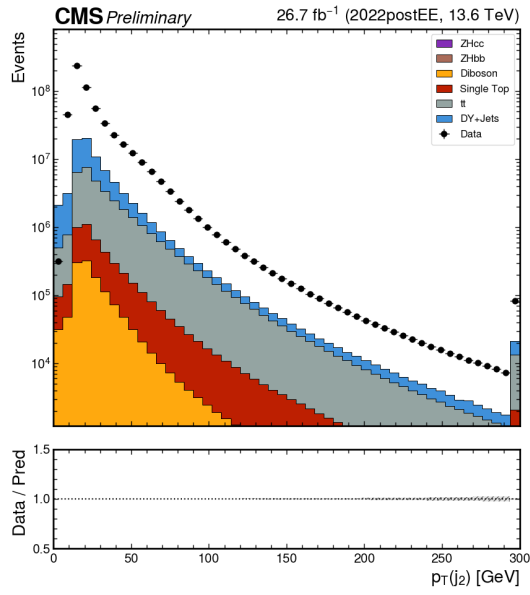
Figuur 92: CvsB-waarde van subleading jet met selecties



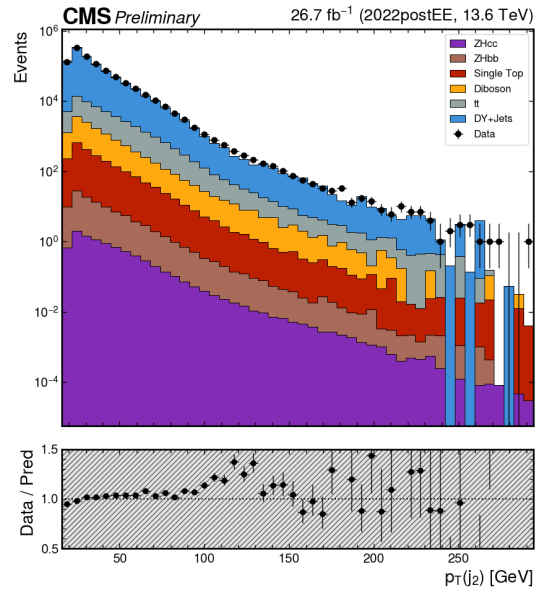
Figuur 93: Transversale impuls leading jet zonder selecties



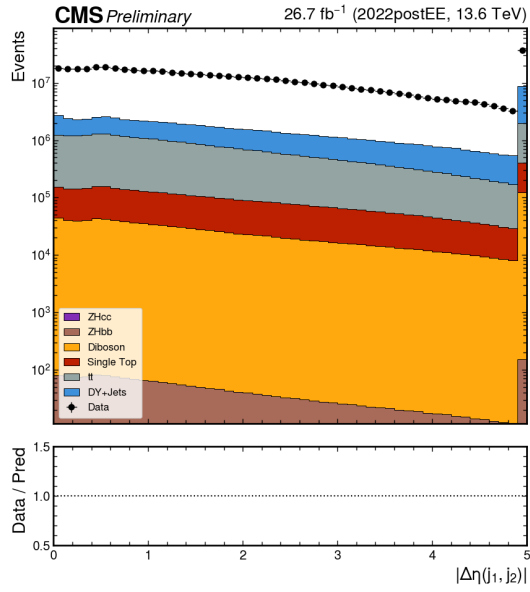
Figuur 94: Transversale impuls leading jet met selecties



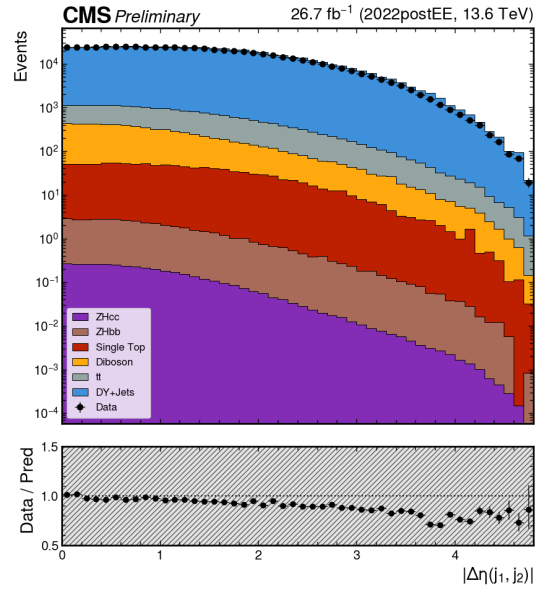
Figuur 95: Transversale impuls subleading jet zonder selecties



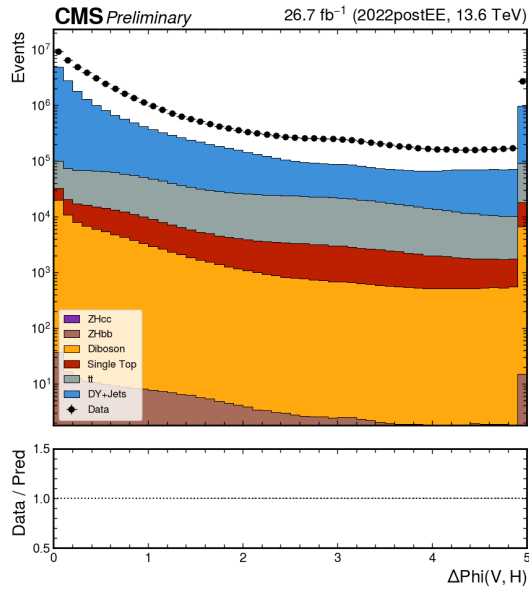
Figuur 96: Transversale impuls subleading jet met selecties



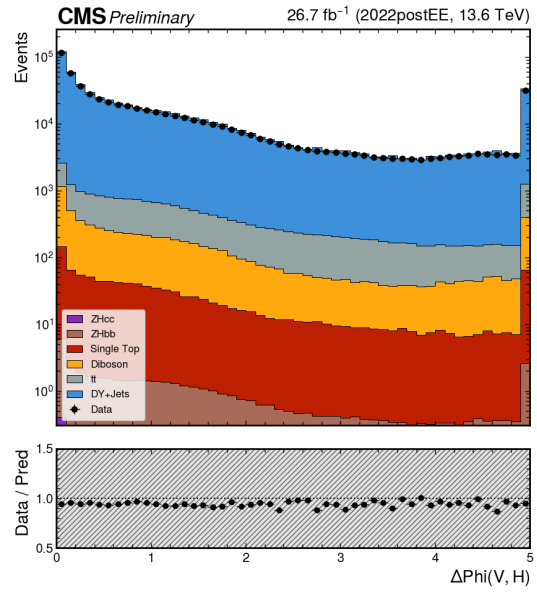
Figuur 97: Verschil in pseudorapiditeit tussen leading en subleading jet zonder selecties



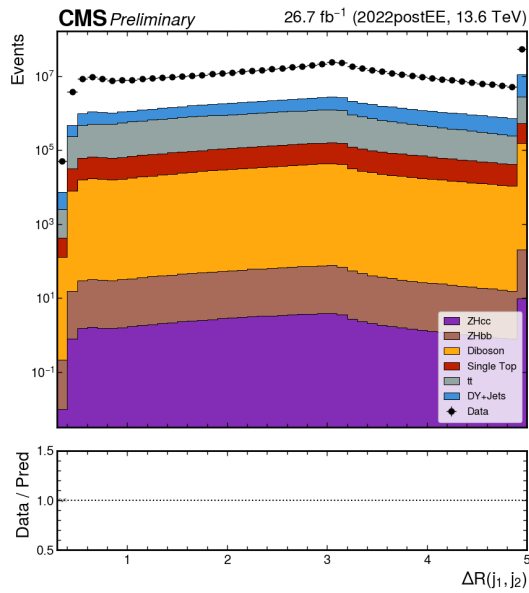
Figuur 98: Verschil in pseudorapiditeit tussen leading en subleading jet met selecties



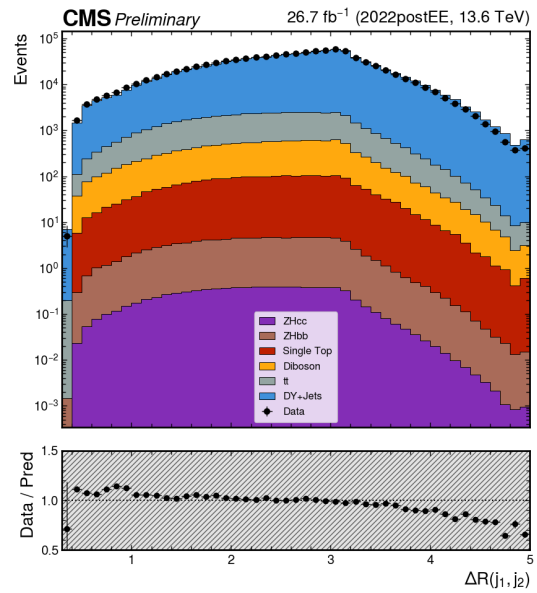
Figuur 99: Azimutale hoek tussen Z-bosonkandidaat en Higgsbosonkandidaat zonder selecties



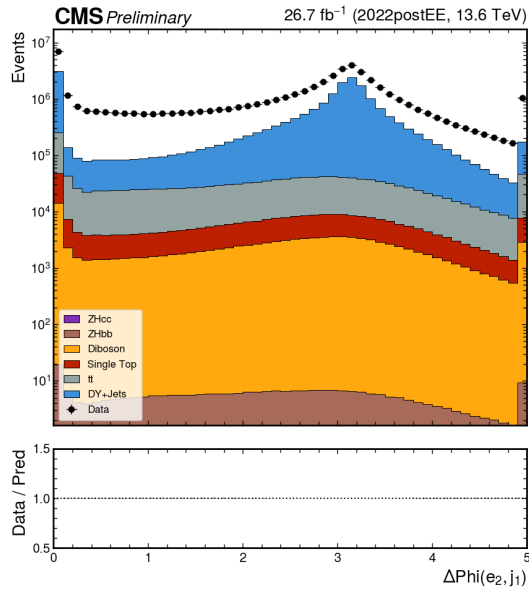
Figuur 100: Azimutale hoek tussen Z-bosonkandidaat en Higgsbosonkandidaat met selecties



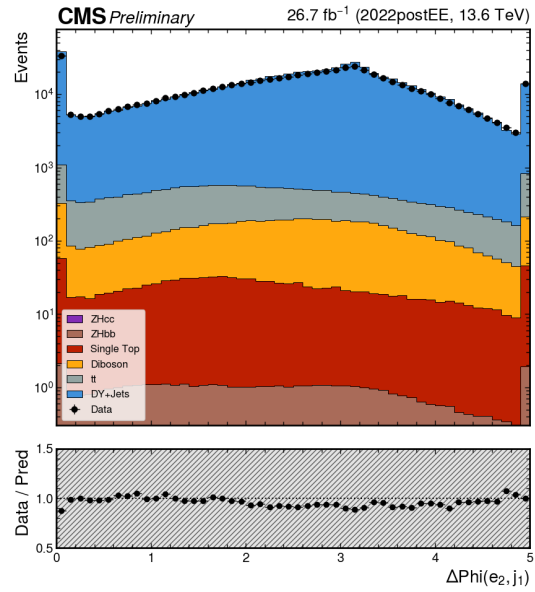
Figuur 101:  $\Delta R$  tussen leading en subleading jet zonder selecties



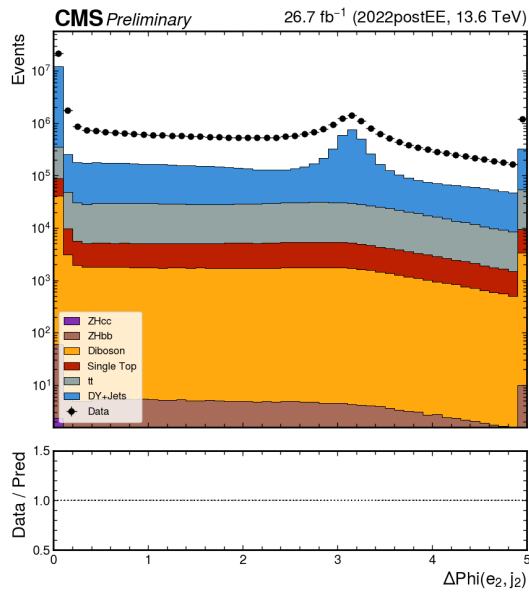
Figuur 102:  $\Delta R$  tussen leading en subleading jet met selecties



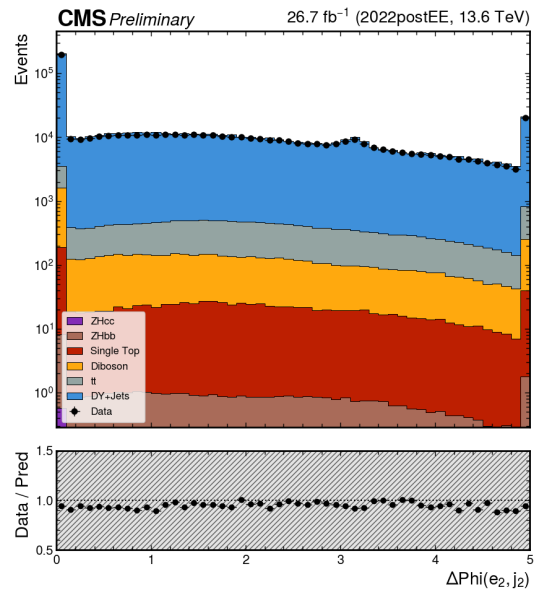
Figuur 103: Azimutale hoek tussen subleading lepton en leading jet zonder selecties



Figuur 104: Azimutale hoek tussen subleading lepton en leading jet met selecties

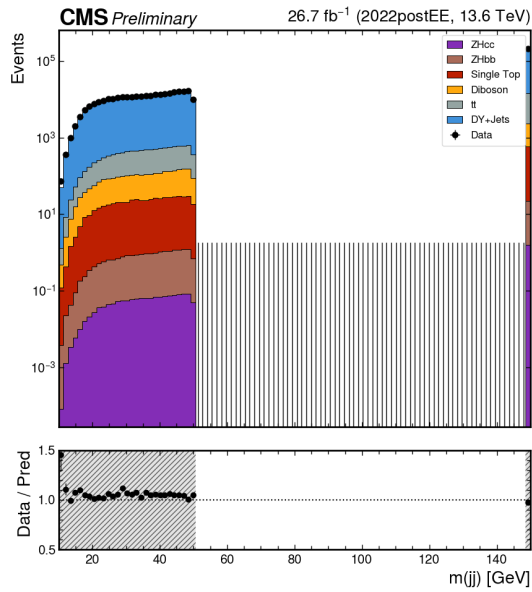


Figuur 105: Azimutale hoek tussen subleading lepton en subleading jet zonder selecties

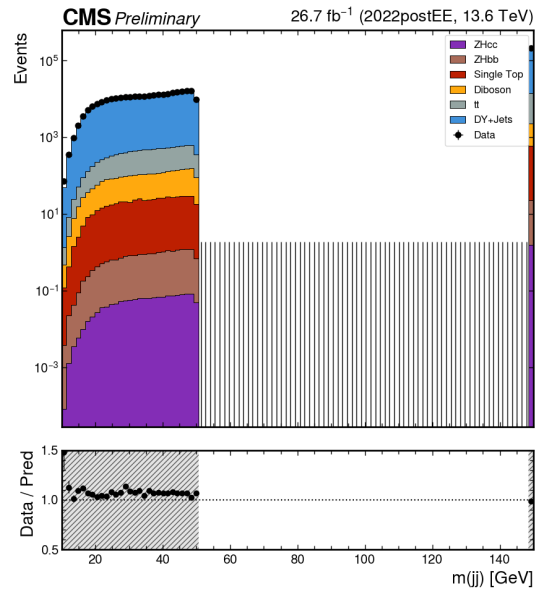


Figuur 106: Azimutale hoek tussen subleading lepton en subleading jet met selecties

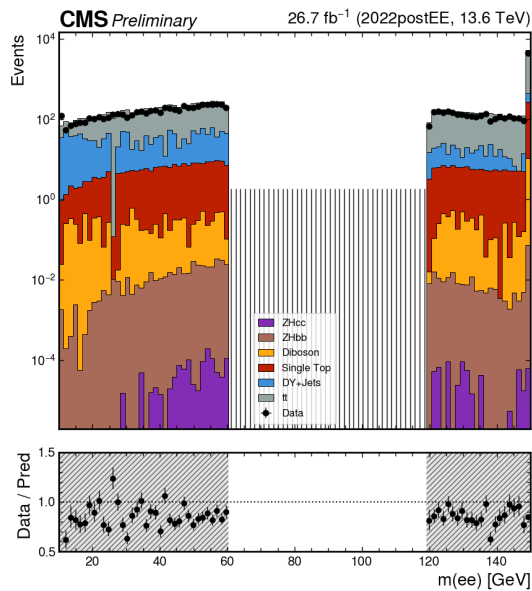
## B Controle regio's elektronkanaal



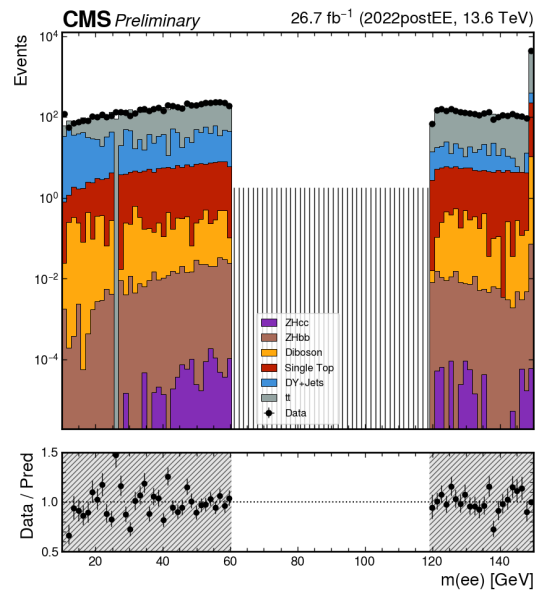
Figuur 107: Invariante massa dijets controle regio 1 zonder normalisatie



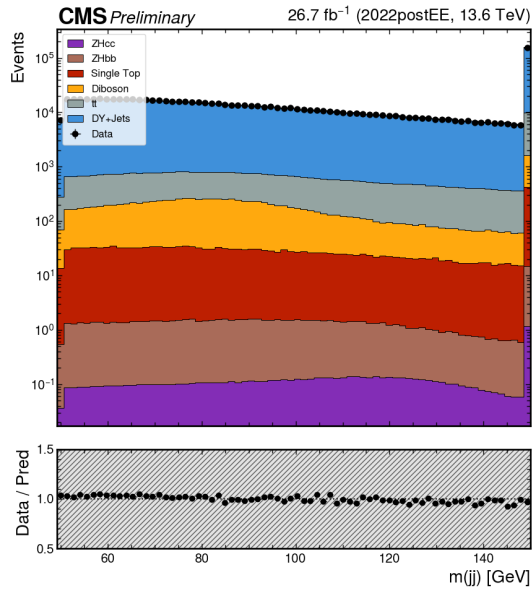
Figuur 108: Invariante massa dijets controle regio 1 met normalisatie



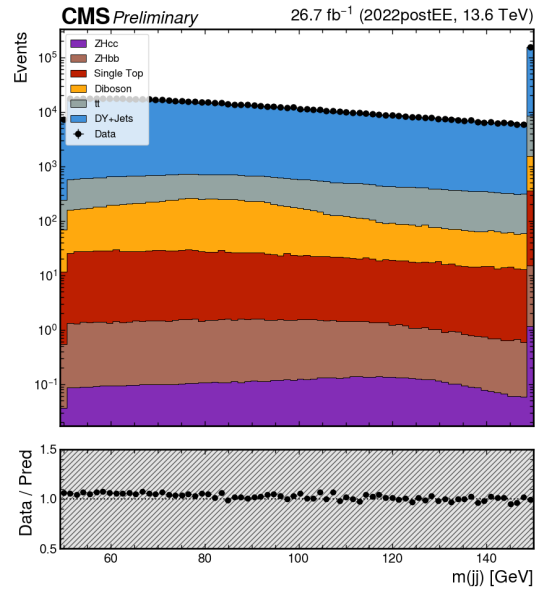
Figuur 109: Invariante massa dielektronen controle regio 2 zonder normalisatie



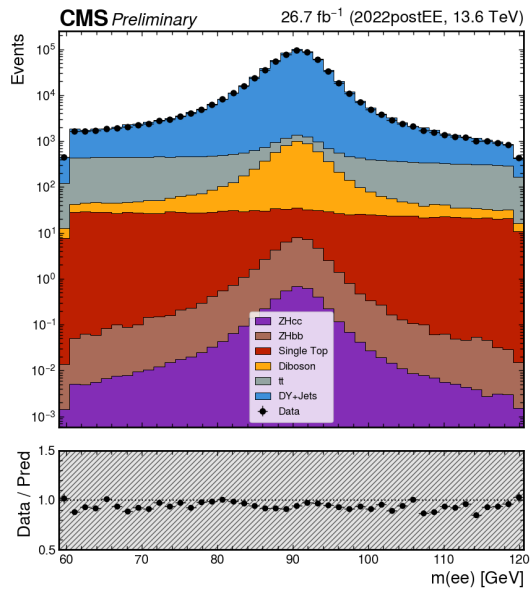
Figuur 110: Invariante massa dielektronen controle regio 2 met normalisatie



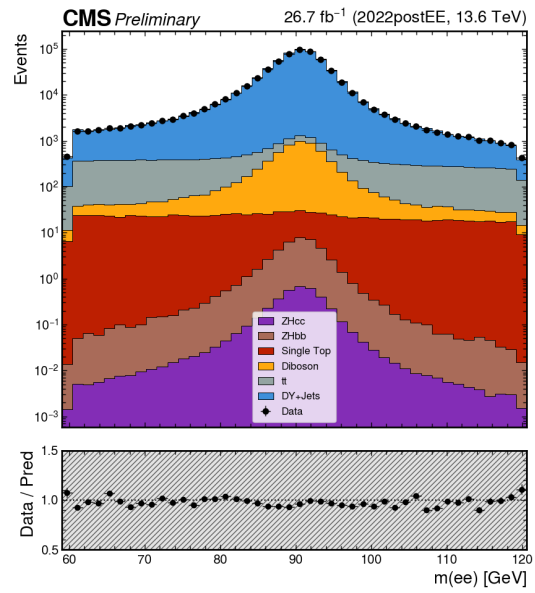
Figuur 111: Invariante massa dijets signaal regio zonder normalisatie



Figuur 112: Invariante massa dijets signaal regio met normalisatie

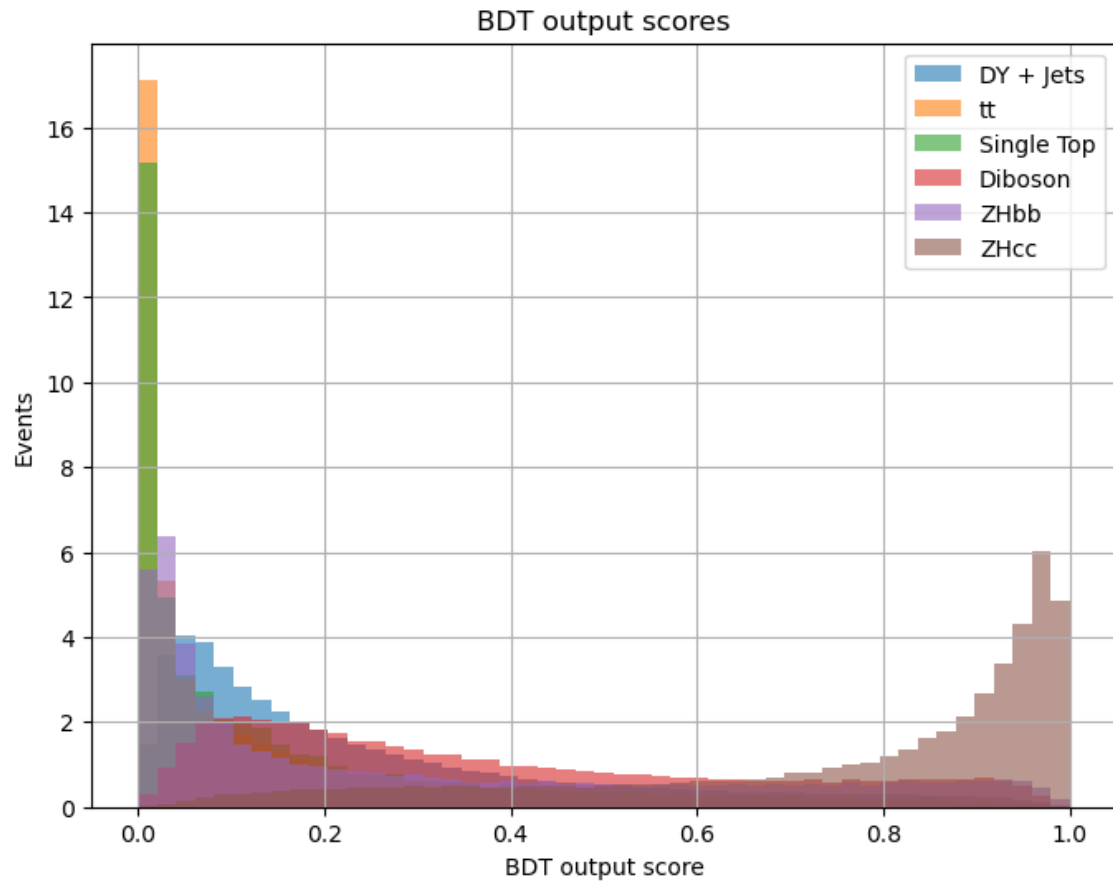


Figuur 113: Invariante massa dielektronen signaal regio zonder normalisatie

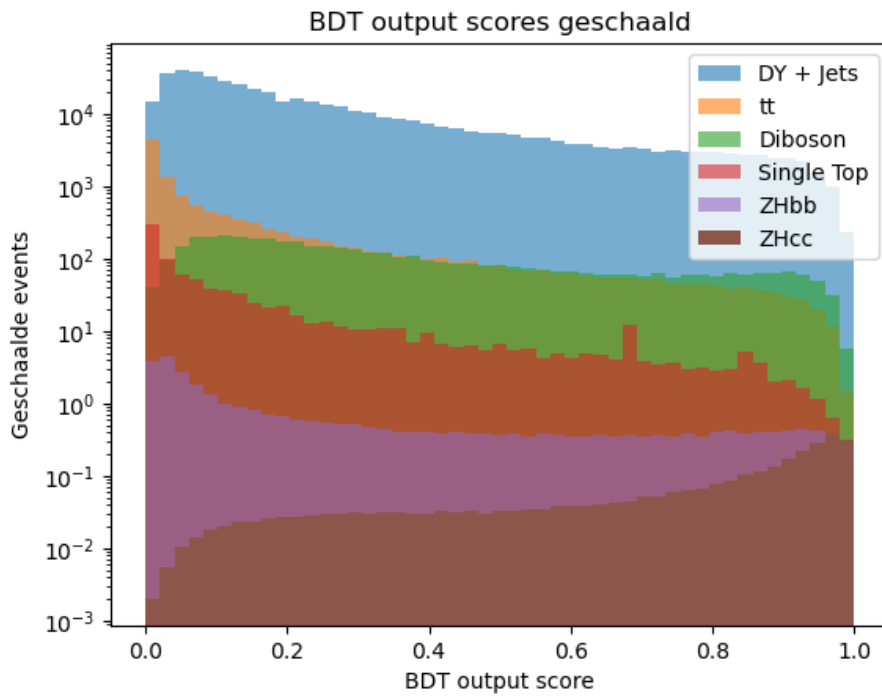


Figuur 114: Invariante massa dielektronen signaal regio met normalisatie

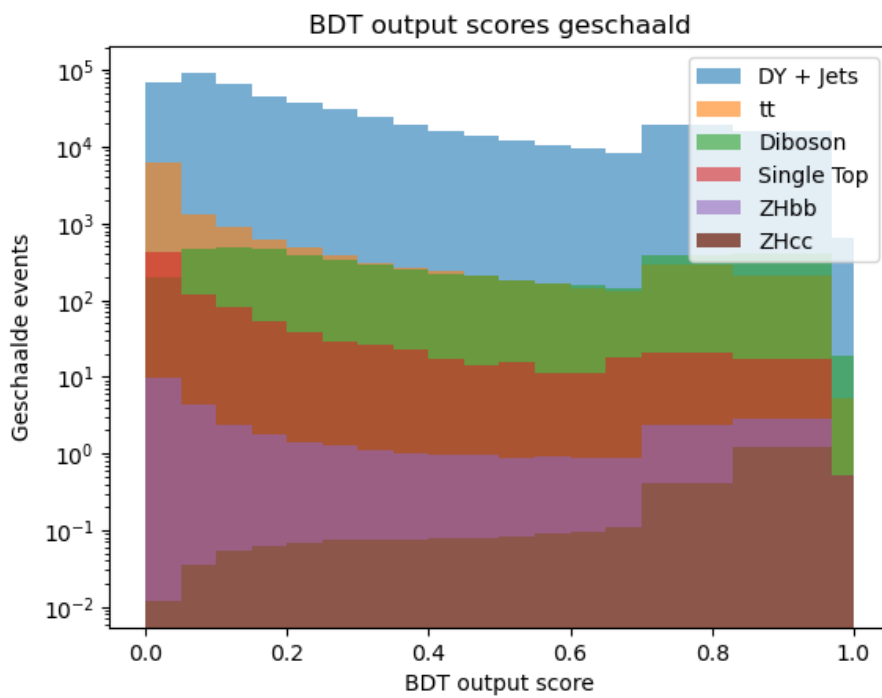
## C BDT elektronkanaal



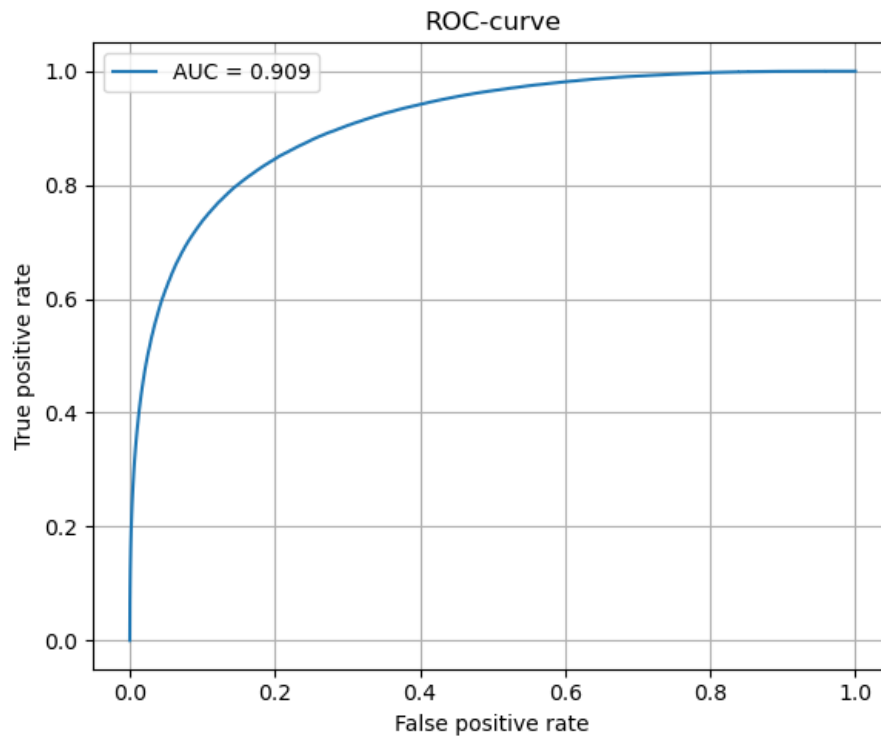
Figuur 115: Ongeschaalde outputscores getrainde BDT



Figuur 116: Geschaalde outputscores getrainde BDT in logaritmische schaal



Figuur 117: Geschaalde outputscores getrainde BDT met aangepaste binning in logaritmische schaal



Figuur 118: ROC-curve getrainde BDT

## Referenties

- [1] Mark Thomson. *Modern Particle Physics*. Cambridge University Press, 5 sep 2013. DOI: 10.1017/cbo9781139525367.
- [2] A. Tumasyan e.a. “A portrait of the Higgs boson by the CMS experiment ten years after the discovery”. In: *Nature* 607.7917 (4 jul 2022), p. 60–68. DOI: 10.1038/s41586-022-04892-x.
- [3] A. Tumasyan e.a. “Search for Higgs Boson Decay to a Charm Quark-Antiquark Pair in Proton-Proton Collisions at s=13 TeV”. In: *Physical Review Letters* 131.6 (7 aug 2023). DOI: 10.1103/physrevlett.131.061801.
- [4] The CMS Collaboration. *The CMS statistical analysis and combination tool: COMBINE*. 12 nov 2024. DOI: 10.1007/s41781-024-00121-4.
- [5] *CMS precisely measures the mass of the Higgs boson — CMS Experiment*. URL: <https://cms.cern/news/cms-precisely-measures-mass-higgs-boson>.
- [6] *Life of the Higgs boson — CMS Experiment*. URL: <https://cmsexperiment.web.cern.ch/news/life-higgs-boson>.
- [7] *File:Standard Model of Elementary Particles.svg - Wikimedia Foundation Governance Wiki*. 17 sep 2019. URL: [https://foundation.wikimedia.org/wiki/File:Standard\\_Model\\_of\\_Elementary\\_Particles.svg](https://foundation.wikimedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg).
- [8] Franz Mandl en Graham P. Shaw. *Quantum Field Theory*. Wiley, 17 mei 2010.
- [9] C. Grojean e.a. *Higgs Physics*. 2 aug 2017. URL: <https://arxiv.org/pdf/1708.00794>.
- [10] The CMS Collaboration e.a. *A New Boson with a Mass of 125 GeV Observed with the CMS Experiment at the Large Hadron Collider*. 2012. DOI: 10.1126/science.1230816.
- [11] *Luminosity*. URL: <https://atlas.cern/Glossary/luminosity>.
- [12] *The Large Hadron Collider*. 9 jul 2025. URL: <https://home.cern/science/accelerators/large-hadron-collider>.
- [13] *File:Cern-accelerator-complex.svg - Wikimedia Commons*. 23 mei 2011. URL: <https://commons.wikimedia.org/wiki/File:Cern-accelerator-complex.svg>.
- [14] Cms Collaboration e.a. “The CMS experiment at the CERN LHC”. In: *Journal of Instrumentation* 3.08 (14 aug 2008), S08004. DOI: 10.1088/1748-0221/3/08/s08004.
- [15] *Detector — CMS Experiment*. 17 mrt 2025. URL: <https://cms.cern/index.php/detector>.
- [16] Izaak Neutelings. *CMS coordinate system*. 18 feb 2025. URL: [https://tikz.net/axis3d\\_cms/](https://tikz.net/axis3d_cms/).
- [17] Dave Barney. *An overview of the CMS experiment for CERN guides*. 27 okt 2003. URL: <https://cds.cern.ch/record/2629323/files/CMSdocumentforGuides.pdf>.

- [18] A.M. Sirunyan, A. Tumasyan en given-i=W. “Particle-flow reconstruction and global event description with the CMS detector”. In: *Journal of Instrumentation* 12.10 (6 okt 2017), P10003. DOI: 10.1088/1748-0221/12/10/p10003.
- [19] A. M. Sirunyan e.a. “A search for the standard model Higgs boson decaying to charm quarks”. In: *Journal of High Energy Physics* 2020.3 (1 mrt 2020). DOI: 10.1007/jhep03(2020)131.
- [20] S. Chatrchyan e.a. *Observation of a new boson with mass near 125 GeV in pp collisions at  $\sqrt{s} = 7$  and 8 TeV*. Jun 2013. DOI: 10.1007/jhep06(2013)081.
- [21] S. Baffioni e.a. *Electron reconstruction in CMS*. 2007. DOI: <https://doi.org/10.1140/epjc/s10052-006-0175-5>.
- [22] Matteo Cacciari, Gavin P Salam en Gregory Soyez. “The anti-ktjet clustering algorithm”. In: *Journal of High Energy Physics* 2008.04 (16 apr 2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063.
- [23] *Reconstructing a multitude of particle tracks within CMS — CMS Experiment*. URL: <https://cms.cern/news/reconstructing-multitude-particle-tracks-within-cms>.
- [24] Daniele Bertolini e.a. “Pileup per particle identification”. In: *Journal of High Energy Physics* 2014.10 (1 okt 2014). DOI: 10.1007/jhep10(2014)059.
- [25] Huilin Qu en Loukas Gouskos. “Jet tagging via particle clouds”. In: *Physical review. D/Physical review. D*. 101.5 (26 mrt 2020). DOI: 10.1103/physrevd.101.056019.
- [26] Uttiya Sarkar. “Run 3 performance and advances in heavy-flavor jet tagging in CMS”. In: *arXiv (Cornell University)* (8 dec 2024). DOI: 10.48550/arxiv.2412.05863.
- [27] Zhuoran Feng On behalf of ATLAS and CMS collaborations. *Full Run 2 VH, H→cc Results from ATLAS and CMS collaborations*. 5 dec 2024.
- [28] LHCPhysics Web. *SM Higgs production cross sections at  $s = 13$  TeV*. URL: [https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageAt13TeV#ZH\\_Process](https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageAt13TeV#ZH_Process).
- [29] A. Hayrapetyan e.a. “Measurement of the Higgs boson production via vector boson fusion and its decay into bottom quarks in proton-proton collisions at  $\sqrt{s} = 13$  TeV”. In: *Journal of High Energy Physics* 2024.1 (30 jan 2024). DOI: 10.1007/jhep01(2024)173.
- [30] LHCPhysics Web. *SM Higgs Branching Ratios and Total Decay Widths*. URL: <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageBR>.
- [31] Rene Brun en Fons Rademakers. “ROOT — An object oriented data analysis framework”. In: *Nuclear Instruments and Methods in Physics Research Section A Accelerators Spectrometers Detectors and Associated Equipment* 389.1-2 (1 apr 1997), p. 81–86. DOI: 10.1016/s0168-9002(97)00048-x.
- [32] Deoache. *GitHub - deoache/higgscharm*. URL: <https://github.com/deoache/higgscharm.git>.

- [33] S. Agostinelli e.a. “Geant4—a simulation toolkit”. In: *Nuclear Instruments and Methods in Physics Research Section A Accelerators Spectrometers Detectors and Associated Equipment* 506.3 (21 jun 2003), p. 250–303. DOI: 10.1016/s0168-9002(03)01368-8.
- [34] Paolo Nason. “A New Method for Combining NLO QCD with Shower Monte Carlo Algorithms”. In: *Journal of High Energy Physics* 2004.11 (15 nov 2004), p. 040. DOI: 10.1088/1126-6708/2004/11/040.
- [35] Stefano Frixione, Paolo Nason en Carlo Oleari. “Matching NLO QCD computations with parton shower simulations: the POWHEG method”. In: *Journal of High Energy Physics* 2007.11 (23 nov 2007), p. 070. DOI: 10.1088/1126-6708/2007/11/070.
- [36] Simone Alioli e.a. “A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX”. In: *Journal of High Energy Physics* 2010.6 (1 jun 2010). DOI: 10.1007/jhep06(2010)043.
- [37] J. Alwall e.a. “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”. In: *Journal of High Energy Physics* 2014.7 (jul 2014). DOI: 10.1007/jhep07(2014)079.
- [38] Wikipedia contributors. *Parton (particle physics)*. 26 jan 2025. URL: [https://en.wikipedia.org/wiki/Parton\\_\(particle\\_physics\)](https://en.wikipedia.org/wiki/Parton_(particle_physics)).
- [39] Torbjörn Sjöstrand e.a. “An introduction to PYTHIA 8.2”. In: *Computer Physics Communications* 191 (11 feb 2015), p. 159–177. DOI: 10.1016/j.cpc.2015.01.024.
- [40] CMSPublic Web. *PAT Cross Cleaning*. URL: <https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuidePATCrossCleaning>.
- [41] *Physics Objects - CMS Open Data Guide*. URL: <https://cms-opendata-guide.web.cern.ch/analysis/selection/objects/objects/>.
- [42] Pedro J. Fernández Manteca en Instituto de Física de Cantabria (Universidad de Cantabria - CSIC). *Muon Identification and Isolation efficiencies on Run II data with the CMS experiment*. 2018. URL: <https://pos.sissa.it/321/068/pdf>.
- [43] *Uncertainties in CMS analyses: Physics object corrections*. 26 jul 2024. URL: <https://cms-opendata-workshop.github.io/workshop2024-lesson-uncertainties/instructor/03-corrections.html>.
- [44] “Jet energy scale and resolution of jets with ParticleNet pT regression using data collected by the CMS experiment for partial Run 3”. In: *Comput Softw Big Sci* 4 (7 nov 2024). URL: [https://indico.cern.ch/event/1386125/contributions/6183557/attachments/2958626/5210964/ML4Jets\\_Malucchi\\_PNetCalibration.pdf](https://indico.cern.ch/event/1386125/contributions/6183557/attachments/2958626/5210964/ML4Jets_Malucchi_PNetCalibration.pdf).
- [45] Joshua Thomas-Wilsker en IHEP. *HF Tagging in CMS*. URL: [https://indico.ihep.ac.cn/event/12369/contributions/17839/attachments/8567/9778/HF\\_Tagging\\_in\\_CMS.pdf](https://indico.ihep.ac.cn/event/12369/contributions/17839/attachments/8567/9778/HF_Tagging_in_CMS.pdf).

- [46] Waleed Esmail e.a. “KinFit: a kinematic fitting package for Hadron Physics experiments”. In: *Computing and Software for Big Science* 8.1 (7 jan 2024). DOI: 10.1007/s41781-023-00112-x.
- [47] Dan Lurie e.a. “Probability and statistics in Particle Physics.” In: *Journal of the American Statistical Association* 77.378 (1 jun 1982), p. 497. DOI: 10.2307/2287284.
- [48] *Taking a closer look at LHC - Momentum*. URL: [https://www.lhc-closer.es/taking\\_a\\_closer\\_look\\_at\\_lhc/0.momentum](https://www.lhc-closer.es/taking_a_closer_look_at_lhc/0.momentum).
- [49] Waleed Esmail e.a. “KinFit – A Kinematic Fitting Package for Hadron Physics Experiments”. In: (aug 2023). arXiv: 2308.09575 [physics.data-an].
- [50] *DoubleCB — zfit 0.6.4 documentation*. URL: [https://zfit.readthedocs.io/en/0.6.4/user\\_api/\\_generated/pdf/zfit.pdf](https://zfit.readthedocs.io/en/0.6.4/user_api/_generated/pdf/zfit.pdf). DoubleCB.html.
- [51] *ROOT: ROOCrystalBall Class Reference*. URL: <https://root.cern/doc/master/classRooCrystalBall.html>.
- [52] Yann Coadou en Centre de physique des particules de Marseille (CPPM), Aix Marseille Université, CNRS/IN2P3, Marseille, France. *Boosted decision trees*. Red. door P. Calafiura, D. Rousseau en K. Terao. 2022. URL: <https://arxiv.org/pdf/2206.09645>.
- [53] *lightgbm.LGBMClassifier — LightGBM 4.6.0.99 documentation*. URL: <https://lightgbm.readthedocs.io/en/latest/pythonapi/lightgbm.LGBMClassifier.html>.
- [54] *Common statistical methods - Combine*. URL: <https://cms-analysis.github.io/HiggsAnalysis-CombinedLimit/part3/commonstatsmethods/#compute-the-observed-significance>.
- [55] *Model and Likelihood - Combine*. URL: [https://cms-analysis.github.io/HiggsAnalysis-CombinedLimit/latest/what\\_combine\\_does/model\\_and\\_likelihood/#sets-of-observation-models](https://cms-analysis.github.io/HiggsAnalysis-CombinedLimit/latest/what_combine_does/model_and_likelihood/#sets-of-observation-models).
- [56]  *$H \rightarrow b\bar{b}$  decays in associated  $VH$  production using data from the ATLAS detector - Enlighten Theses*. URL: <https://theses.gla.ac.uk/84731/>.
- [57] *Preparing the datacard - Combine*. URL: <https://cms-analysis.github.io/HiggsAnalysis-CombinedLimit/latest/part2/settinguptheanalysis/#a-simple-counting-experiment>.
- [58] *Electron and photon reconstruction and identification performance at CMS in 2022 and 2023*. 13 jul 2024. URL: <https://cds.cern.ch/record/2904365>.