



Vrije
Universiteit
Brussel

Academiejaar 2006 - 2007

Faculteit Wetenschappen
Departement Natuurkunde
Michael Maes

Bachelorproject:

Onderscheiden van signaal en achtergrond
in de CMS-detector van LHC te CERN.

Promotor: Jorgen D'Hondt

Inhoudsopgave

1	Inleiding	1
1.1	Het Standaard Model	1
1.2	De Large Hadron Collider	6
1.2.1	De versneller	6
1.2.2	De CMS detector	9
2	Project	11
2.1	Productie van de Top Quark	11
2.2	Reconstructie en analyse	15
2.3	Conclusie	20

Samenvatting

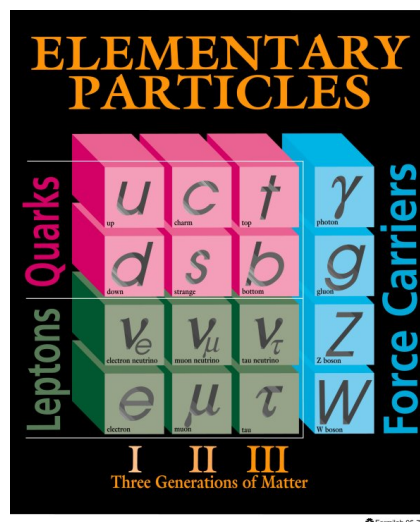
Dit project handelt over de Large Hadron Collider (LHC) die op dit moment in constructie is te CERN. Deze deeltjesversneller zal in de zomer van 2007 starten met het versnellen van bundels protonen (10^{11} protonen per bundel). Op 4 punten in deze versneller zijn op dit moment gigantische detectoren in opbouw (CMS, ATLAS, LHCb, ALICE). Binnen deze detectoren zullen twee protonenbundels elkaar kruisen en een aantal protonen tot botsing komen. Deze botsingen genereren een energie van 14TeV, en hierdoor worden top quarks (en andere deeltjes) geproduceerd. Dit maakt deze circulaire versneller de grootste (27km omtrek) en meest energierijke ter wereld. De top quark werd ontdekt bij Tevatron in 1995 en heeft een massa van $175 \frac{GeV}{c^2}$. Bij de LHC is er voor het eerst een gigantische verzameling aan top quarks. De VUB werkt mee aan de constructie van de CMS (Compact Muon Solenoid) detector. Dit is een algemene detector die onderandere gebruikt zal worden in de zoektocht naar een experimentele bevestiging van het Higgs-deeltje en de verificatie van Supersymmetrie (SUSY) en het bestaan van super-deeltjes. Van alle proton-proton botsingen die zullen plaatsvinden in CMS zijn er een aantal $pp \rightarrow t\bar{t}$ (signaal) en een aantal $pp \rightarrow W + Jets$ (achtergrond). Deze laatste gebeurtenissen zijn achtergrondsignalen en dus niet interessant indien we de fysica van de top quark wensen te bestuderen. Daarom wilt men een techniek om deze achtergrondsignalen weg te filteren. Dit moet echter zo nauwkeurig mogelijk gebeuren zodat de kans geminimaliseerd wordt dat men signalen zal wegfilteren. Mijn project bestaat erin eigenschappen te bestuderen van beide type gebeurtenissen en daarin verschillen te zoeken. De selectie van signaal en gebeurtenissen gebeurt aan de hand van selectiesneden op specifieke observabelen. Dit geeft aanleiding tot fouten van types I en II zoals bij een hypothesetoets. Aan de hand van gesimuleerde gebeurtenissen wensen we observabelen te identificeren die beide fouten minimaliseren. Het doel van mijn project is dan ook dat ik zoveel mogelijk eigenschappen van de beide types botsingen combineer in een ultieme hypothesetoets. Dit kan ik doen via Support Vector Machine algoritmen. Maar eerst moet ik uiteraard zelf op zoek gaan naar eigenschappen van de beide gebeurtenissen die verschillend zijn. Deze eigenschappen zal ik onderzoeken op gesimuleerde data van proton-protonbotsingen zoals deze in de LHC ook zullen plaatsvinden. Dit proced is belangrijk voor de analyse van LHC-data omdat er veel meer achtergrondgebeurtenissen dan signalen zullen worden geregistreerd.

Hoofdstuk 1

Inleiding

1.1 Het Standaard Model

Het *Standaard Model*[1], is een kwantumveldentheorie binnen de deeltjesfysica die begin de jaren '70 tot stand is gekomen. Deze theorie beschrijft de elektromagnetische, zwakke en sterke wisselwerkingen, drie van de vier fundamentele interacties tussen de elementaire bouwstenen van alle materie. Deze theorie is consistent zowel met de kwantummechanica als de speciale relativiteitstheorie. Toch is deze theorie geen complete theorie over de fundamentele interacties aangezien de vierde fundamentele interactie, nl. gravitatie, niet beschreven is. Een andere tekortkoming van deze theorie is dat er een aantal parameters zijn (o.a. massa) die men manueel moet ingeven. Zoals in de



Figuur 1.1: Het Standaard Model[2]

schematische voorstelling te zien is (fig. 1.1), zijn de elementaire deeltjes binnen het Standaard Model opgedeeld in twee klassen: de fermionen en de bosonen.

De fermionen zijn de bouwstenen van de materie. Deze deeltjes hebben een halfgehele spin ($\frac{1}{2}$) en zijn onderhevig aan het Pauli exclusieprincipe (geen twee fermionen gelijktijdig in dezelfde kwantumtoestand). Deze klasse bestaat uit 12 deeltjes (plus 12 anti-deeltjes) en is opgedeeld in twee groepen: zes leptonen (electron, muon, tau, en de bijhorende neutrinos). en zes quarks (up, down, strange, charm, top en bottom). Een eerste verschil tussen beide groepen is dat leptonen vrij voorkomen in de natuur, dit is niet het geval bij quarks die altijd voorkomen in een gebonden toestand die men een hadron noemt (bv. proton : up-up-down quark)

De up, charm en top quarks hebben een elektrische lading van $+2/3$ en de down, strange en bottom quarks een lading van $-1/3$ waardoor deze elektromagnetisch kunnen interageren. Dit geldt ook voor de “up-type” leptonen (elektronen, muonen en tau deeltjes) die een lading -1 bezitten, maar niet voor de neutrino’s aangezien deze elektrisch neutraal zijn. Bovenop de elektrische lading beschikken de quarks nog over een kleurlading [4] die gerelateerd is aan de sterke interactie. De quarks kunnen rood, groen of blauw geladen zijn. De benaming kleurlading heeft niets te zien met echte kleuren, maar dit is gewoon een naamgeving voor het concept lading in het licht van de kwantumchromodynamica (analoog aan de elektrische lading). De leptonen bezitten echter geen kleurlading waardoor ze geen sterke interacties kunnen ondergaan. Tot slot hebben zowel quarks als leptonen een “flavour” waardoor ze ook zwak kunnen interageren.

Op de schematische weergave van het Standaard Model (fig. 1.1) zien wij ook dat de fermionen onderverdeeld worden in drie generaties.

$$\begin{pmatrix} e^- \\ \nu_e \\ u \\ d \end{pmatrix} \begin{pmatrix} \mu^- \\ \nu_\mu \\ c \\ s \end{pmatrix} \begin{pmatrix} \tau^- \\ \nu_\tau \\ t \\ b \end{pmatrix}$$

De eerste generatie bevat de bouwstenen voor alle materie in de natuur. De andere twee generaties komen niet stabiel in de natuur voor. Deze waren aanwezig na de big bang, maar komen nu alleen nog voor in experimentele omgevingen zoals te CERN. De massa van de deeltjes neemt toe van de eerste generatie (de lichtsten) naar de derde generatie (de zwaarsten). Zoals eerder gezegd, zijn er nog twaalf anti-deeltjes die in het algemeen dezelfde massa hebben als hun partners maar een tegengestelde lading.

De tweede klasse van elementaire deeltjes in het Standaard Model wordt gevormd door de bosonen. Deze deeltjes zijn de dragers van de fundamentele

interacties en hebben een gehele spin. Hierdoor zijn ze in tegenstelling tot fermionen niet onderhevig aan het Pauli exclusieprincipe. Het eerste type bosonen zijn de fotonen¹. Deze deeltjes hebben een geen massa, bewegen zich voort aan de lichtsnelheid (in vacuüm) en staan in voor de elektromagnetische wisselwerking. Een tweede groep zijn de W^\pm en Z^0 bosonen en staan in voor de zwakke wisselwerking tussen zowel quarks als leptonen. Deze deeltjes werden in 1983 ontdekt te CERN en waren een groot succes voor het Standaard Model. Het Z^0 boson is het zwaarste deeltje van de drie² en is elektrisch neutraal. Dit deeltje werkt op zowel rechtshandige als linkshandige deeltjes, maar bij voorkeur op linkshandige. Linkshandige deeltjes zijn deze met een chiraliteit³ gelijk aan $-1/2$ en rechtshandige met een chiraliteit $+1/2$. Het W-boson heeft een lading ($+1$ voor W^+ en -1 voor W^-). Daardoor kan dit boson buiten de zwakke wisselwerking (exclusief op linkshandige deeltjes) nog elektromagnetisch interageren. De Z^0 en W^\pm bosonen samen met de fotonen dragen de elektrozwakke wisselwerking. Tot slot hebben wij nog acht gluonen. Deze gluonen dragen de sterke interactie tussen kleurgeladen deeltjes zoals quarks. Deze deeltjes zijn massaloos, maar door hun dubbele kleurlading (een kleur en een anti-kleur) kunnen ze met andere gluonen interageren.

Er ontbreekt echter nog een deeltje in dit plaatje: het Higgs deeltje (Higgs boson)[6]. Zoals eerder werd aangehaald faalt het Standaard Model in het voorspellen van de massa's van de deeltjes, waardoor men deze manueel moet invoeren. Het mechanisme waaraan dit deeltje gekoppeld is (Higgs-mechanisme) kan de niet nulle massa van de deeltjes in het Standaard Model verklaren en ook het groot massaverschil tussen het massaloze foton en de zware Z en W bosonen. Dit deeltje is nog steeds niet ontdekt, dus als men het bestaan van dit momenteel hypothetisch deeltje kan aantonen zal dit een enorme invloed hebben op de fysica. Het Standaard Model kan dus uiteraard geen massa voorspellen voor dit Higgs boson, maar de massa wordt geschat tussen 115 en 180 $\frac{GeV}{c^2}$.

Buiten de 19 vrije parameters in het model, hebben wetenschappers in 1998 nog tekortkoming gevonden in dit model. Deze tekortkoming is de neutrino oscillatie. Volgens het Standaard Model kan een massaloos neutrino niet oscilleren, maar experimenteel werd aangetoond dat een neutrino wel degelijk van flavour⁴ kan veranderen (dus oscilleren). Dit gaf aanleiding tot het bestaan van neutrino's met een niet nulle (maar zeer kleine) massa waardoor het Standaard Model moest aangepast worden (wat nog extra vrije parame-

¹1905-17 door Albert Einstein

² $m_{Z^0} = (91.1876 \pm 0.0021) \text{ GeV}/c^2$ en $m_{W^+} = m_{W^-} = (80.398 \pm 0.025) \text{ GeV}/c^2$

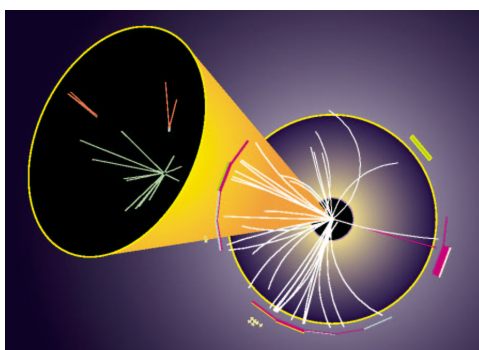
³Component van de spin parrallel aan het momentum

⁴flavour: elektron, muon of tau neutrino

ters met zich meebracht). Desondanks kent deze theorie vele successen. Het Standaard Model voorspelde in het verleden onderandere het bestaan van het W en Z boson, het gluon en de top en charm quark en deze zijn allemaal experimenteel geobserveerd.

Omdat het Standaard Model toch een aantal lancunes bevat, daarom verwacht men dat er extensies op de huidige theorieën nog te ontdekken zijn in experimenten op de TeV-schaal. Eén van de kandidaat theorieën is supersymmetrie⁵. Deze theorie voorspelt massieve supersymmetrische partners voor de elementaire deeltjes (die misschien een verklaring bieden voor donkere materie).

Een heel belangrijk deeltje in de huidige studies binnen de deeltjesfysica is de top quark. Deze up-type quark is niet alleen de zwaarste in zijn soort, maar ook met een massa van $174.2 \pm 3.3 \frac{GeV}{c^2}$ [7] het zwaarste deeltje in het huidige Standaard Model en werd pas in 1995 ontdekt bij de Tevatron⁶ versneller. Hoewel de top quark een elektrische lading bezit ($+2/3$) zal deze bijna uitsluitend sterk interageren, vervalt uitsluitend via zwak verval en meestal in een W boson en bottom quark. Dit volgt uit de korte levensduur van de top quark (10^{-25} s) die voorspeld wordt via het Standaard Model. Deze is veel korter dan het tijdsschaal voor sterke interacties waardoor de top quark niet zal hadroniseren [10]. Dit wil zeggen dat de top quark niet zal voorkomen in een gebonden toestand van quarks of hadron. Dit heeft als groot voordeel dat men de top quark op zich kan bestuderen. Dit is uniek omdat alle andere quarks altijd voorkomen gegroepeerd met twee (mesonen) of met drie (baryonen) [11]. De b quark, gevormd in het verval van de top quark, zal jets [9] voortbrengen. Jets zijn “smalle kegels“ (fig. 1.2) van hadronen en



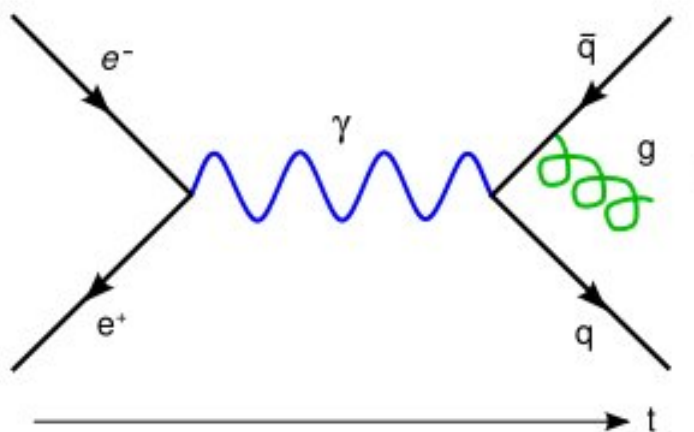
Figuur 1.2: Voorbeeld van een jet

andere deeltjes die voortkomen uit hadronisatie van quarks en gluonen.

⁵Deze theorie wordt verder in dit project niet beschouwd

⁶Bij Fermilab, VSA

Tot slot voor het Standaard Model nog een klein woordje over de representatie van interacties. Een veel gebruikte methode om de interacties voor te stellen tussen deeltjes is de voorstelling via Feynman diagrammen [3]. Figuur



Figuur 1.3: Voorbeeld van een Feynman diagram

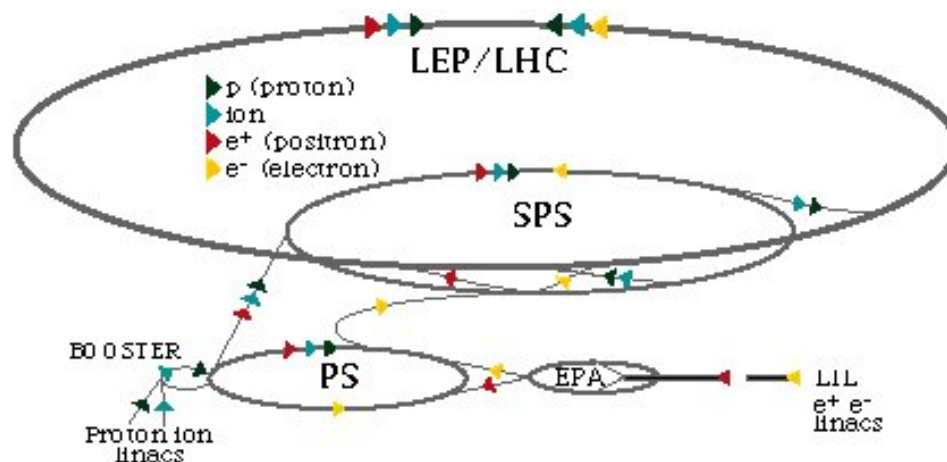
1.3 is een arbitrair voorbeeld van dergelijk diagram. In dit diagram vertolken de rechte lijnen deeltjes met spin $\frac{1}{2}$ of fermionen (quarks en leptonen), de golvende lijnen spin 1 deeltjes of bosonen (photon, W^\pm en Z^0) en de spiraalvormige lijnen spin 1 deeltjes of bosonen (gluonen). De pijlen op de lijnen geven aan of het een deeltje of anti-deeltje is (bv voor het e^- naar rechts en het e^+ naar links). Tot slot zijn er nog de punten waarin de lijnen samenkomen (interactiepunten van de deeltjes), deze punten krijgen de naam vertex. Deze diagrammen zijn niet alléén een voorstelling van de interacties! Uit deze diagrammen kan men ook de wekzame doorsnede en de waarschijnlijkheid van de interactie berekenen.

1.2 De Large Hadron Collider

1.2.1 De versneller

De Large Hadron Collider [13] (LHC) is het huidige vlaggenschip van CERN gelegen nabij Genève (Zwitserland). Deze deeltjesversneller wordt momenteel opgebouwd in de 27km lange cirkulaire tunnel van zijn voorganger de Large Electron Positron collider (LEP) en zal tegen 2008⁷ protonenbundels laten botsen met een energie van 7 TeV wat een massamiddelpuntsenergie van 14 TeV voortbrengt. Dit maakt dat de LHC, wanneer hij op volle kracht protonen zal botsen, de meest energierijke versneller ter wereld zal zijn. De tunnel ligt 50 tot 175m onder de grond op sommige plaatsen en ligt voor het grootste deel op Frans grondgebied met vier detectoren op het traject: ATLAS, CMS, ALICE en LHCb. Hierbij zijn ATLAS en CMS "algemene" detectoren en de andere twee zijn detectoren met een specifieke taak: LHCb zal gebruikt worden voor het bestuderen van onder andere interacties met b-hadronen⁸ en ALICE zal gebruikt worden voor het bestuderen van botsingen van zware kernen (Pb-kernen die de LHC gedurende bepaalde periodes zal versnellen) bij 5,5TeV massamiddelpuntsenergie.

Om tot een energie van 7 TeV te geraken, worden de protonen voorversneld in andere versnellers alvorens te worden geïnjecteerd in de LHC ring (Figuur 1.2.1).



Figuur 1.4: Het versnellercomplex van CERN

⁷Normaliter zouden er vanaf November 2007 al botsingen moeten gebeuren bij lagere energie

⁸hadronen die een bottom quark bevatten



Figuur 1.5: H_2 "fles" (rood) voor proton-productie

In figuur 1.5 zien wij waar het allemaal begint: een "fles" H_2 die gekoppeld is aan een toestel waarmee men hieruit protonen losmaakt. Deze protonen worden dan geïnjecteerd in een lineaire versneller (linac2) en verkrijgen een energie van 50 MeV. Daarna worden deze protonen in de Proton Synchrotron Booster (BOOSTER) versneld tot 1.4 GeV waarna ze in de Proton Synchrotron (PS) nog extra versneld worden tot 26 GeV. Tot slot worden ze door de laatste voorversneller, de Super Proton Synchrotron (SPS), versneld tot 450 GeV. Eens de protonen deze energie bereikt hebben, worden ze in de LHC ring gebracht en daar verder versneld tot een energie van 7 TeV. De protonen bundels zullen echter geen continue bundels zijn, maar bunches ("pakketten" van 1.15×10^{11} protonen) waarvan er zo'n 2808 per beam zijn zodat er om de 25ns twee bundels elkaar kruisen op de interactiepunten⁹. In de LHC zijn twee bundelpijpen voorzien. Dit komt omdat men een tegengestelde magneetvelden moet gebruiken om de beide bundels af te buigen en door de repulsie van de gelijk geladen protonen kan men de tegenovergesteld circulerende bundels elkaar dus niet kan laten kruisen in dezelfde bundelpijp. Een ander gevolg van de repulsie is dat de kans op frontale botsing tussen twee protonen dan ook vrij klein is. Dit is dan ook de reden dat men heel veel protonen ($\sim 10^{11}$) in elk pakket steekt. Een ander probleem is dat de protonen kunnen botsen met gasmoleculen en dus nooit het interactiepunt bereiken. Om te zorgen dat dit effect zo klein mogelijk is, moet men in de versneller een hoog vacuüm behouden ($\sim 10^{-10}$ Torr of 3 miljoen deeltjes per cm^3). Opdat de protonen circulaire banen zouden beschrijven, moet men deze afbuigen door middel van magnetische velden. Daarom zijn langs het 27km lange traject van de Large Hadron Collider 1232 dipoolmagneten [13]

⁹Wanneer de LHC op volle kracht draait, in de beginfase om de 75ns

geïnstalleerd op de ring. Deze magneten moeten een magnetisch veld van 8.4 Tesla met een stroom van 11700 A leveren om de protonen van 7 TeV in een circulaire baan te houden. Hiervoor moet men gebruik maken van supergeleidende magneten (die dus stroom geleiden zonder weerstand) die op een temperatuur van 1.9 Kelvin gehouden worden met behulp van vloeibare He. Deze magneten zijn 14.3 m lang, wegen ongeveer 35 ton en bevatten 11Gj aan energie¹⁰. Deze magneten zijn 2-in-1 ontworpen, dat wil zeggen dat ze twee openingen hebben waardoor slechts één dergelijke dipoolmagneet nodig is voor de beide bundelpijpen. Buiten de dipoolmagnetten maakt men onder andere ook nog gebruik van quadrupolen om de bundels te focuseren.

De Large Hadron Collider zal op volle kracht proton bundels leveren met een luminositeit¹¹ die 10^2 keer groter is dan bij huidige experimenten (LHC: $10^{34} \text{cm}^{-2} \text{s}^{-1}$) [14]. Als de LHC versneller op volle kracht draait (dus protonen met energie 7 TeV) zal bij deze luminositeit ongeveer 362Mj aan energie opgeslagen zijn in een bundel protonen. Dit brengt nog een technisch probleem met zich mee aangezien bij de kruising van twee bundels slechts een klein aantal protonen frontaal zal botsen. De andere protonen uit de bundel worden gedeflecteerd door het elektromagnetisch veld van de deeltjes uit de andere bundel en dit effect neemt toe met de dichtheid van de bundel. De gedeflecteerde deeltjes botsen tegen de wand van de bundelpijp. Bij deze botsing geven ze hun energie af in de vorm van warmte en dit kan ervoor zorgen dat de supergeleidende magneten een tijd lang niet meer functioneren. Daardoor moet men dit effect zo klein mogelijk houden en is het dus belangrijk ervoor te zorgen dat zo veel mogelijk protonen frontaal botsen. Dit doet men onder andere door de bundels zo veel mogelijk te focuseren (minimale diameter van de bundel ± 16 micron). Deze versneller wordt uiteraard niet zonder reden gebouwd. Deze versneller zal zoals zijn voorgangers testen in hoeverre het Standaard Model een geldige theorie is. De Large Hadron Collider zal onder andere gebruikt worden voor

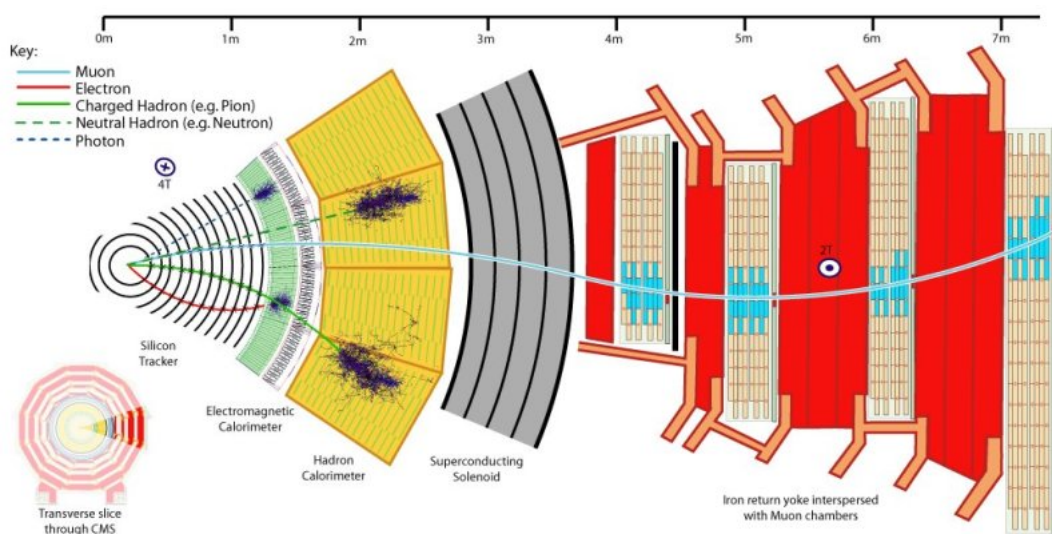
- De zoektocht naar het Higgs deeltje.
- De zoektocht naar "SUSY" (supersymmetrie).
- Bepaling van de eigenschappen van donkere materie.
- De zoektocht naar extra dimensies.

¹⁰Dit komt overeen met de explosie van 2350kg TNT

¹¹Het aantal deeltjes per oppervlakte-eenheid per tijdseenheid

1.2.2 De CMS detector

De Compact Muon Solenoid [15] (CMS) detector is één van de vier grote detectoren die geïnstalleerd zijn nabij de LHC versneller. Deze detector bevindt zich in een ondergrondse schacht te Cessy (Frankrijk) en is 21.5m lang, heeft een diameter van 15m en weegt ongeveer 12500 ton. CMS zal gebruikt worden om de fysica op de TeV-schaal te exploreren, het Higgs boson te zoeken en het bestaan van supersymmetrie en extra dimensies te onderzoeken.



Figuur 1.6: Een snede uit de CMS detector

In figuur 1.6 zien wij een stuk van de transversale doorsnede van de detector. De binnenste laag in CMS is de **spoordetector** [15]. Deze laag is zeer belangrijk in experimenten die alle fysica die mogelijk is met de LHC versneller. Het tracking systeem van CMS zal in staat zijn sporen op te meten van muonen, elektronen en hadronen als ook heel preciese metingen van de momenta. Voor deze spoordetector koos men voor silicium strips. Deze zijn door hun snelle respons en hoge resolutie geschikt voor dit experiment, waarbij 40 miljoen botsingen per seconde in de detector zullen plaatsvinden. Deze detectoren bedekken een oppervlakte van 210 m^2 . Op de figuur zien wij dat de geladen deeltjes (volle lijnen: elektron, muon, ...) gezien worden in de spoordetector en de neutrale deeltjes (stippelijnen: foton, neutraal hadron,..) worden hier niet gedetecteerd.

Een volgende onderdeel van CMS is de **calorimeter** [15]. Hierin zullen de elektronen, fotonen en hadronen gestopt worden waardoor men hun energie kan opmeten. Deze calorimeter bestaat uit twee lagen.

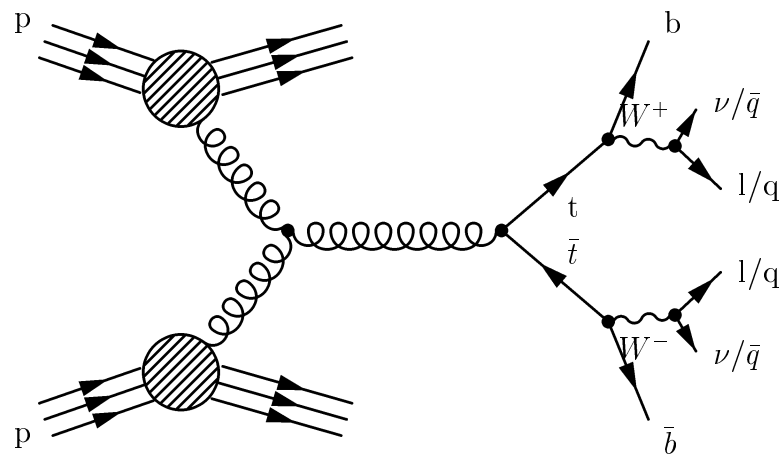
- *De elektromagnetische calorimeter* [15]. Deze dankt zijn naam aan het feit dat de deeltjes die hierin gestopt worden elektromagnetisch wisselwerken en hierin worden dus de energieën van de fotonen en elektronen met hoge precisie gemeten. Deze calorimeter zal bestaan uit 80000 $PbWO_4$ kristallen. Men heeft voor dergelijk kristal gekozen omdat ze bestand zijn tegen de omstandigheden in CMS: een magneetveld van 4T en een aanzienlijke stralingsdosis (1-2kGy/jaar). Een tweede reden is de korte stralingslengte, waardoor de nodige dikte van de kristallen kleiner wordt. De fotonen die in de spoordetector geen spoor nalaten, worden in de elektromagnetische calorimeter wel gezien omdat ze hier dus hun energie afzetten.
- *De hadronische calorimeter* [15]. Deze calorimeter speelt een belangrijke rol in het ontdekken en opmeten van jets komende van quarks of gluonen. Zoals wij zien in de figuur, worden in deze laag geladen en neutrale hadronen gestopt en de energie opgemeten.

Het neutrino momentum wordt opgemeten als het complement van de transverse energie van de jets, aangezien de neutrino's geen spoor achterlaten in de detector en dit een onevenwicht veroorzaakt in de energiesbalans. Een derde laag is de supergeleidende magneet[15]. Deze magneet wekt een magnetisch veld op van 4 Tesla. Dit veld is nodig om de geladen deeltjes af te buigen zodat men uit de buiging van het spoor het momentum kan afleiden en ster veld is nodig om een hoge precisie te bekomen. De magneet in CMS is de grootste supergeleidende magneet ter wereld met een diameter van 14m, een lengte van 21.6m en een gewicht van 12000 ton. In deze magneet zal een energie opgeslagen zijn die, moest ze vrijkomen, voldoende is om 18 ton goud te smelten. Dus men moet hier ook zoals bij de magneten in LHC voorkomen dat de magneten gaan "quenchen". Dit wil zeggen dat ze, door een plotse verandering in stroom of veldsterkte, terugkeren van supergeleidende toestand naar een toestand waar ze als normale geleider werken. Dit brengt warmteontwikkeling met zich mee wat kan leiden tot schade aan de magneet en andere componenten van de detector in de nabijheid van de magneet. Het laatste onderdeel dat te zien is in figuur 1.6 zijn de muon kamers[15]. Deze dienen in samenwerking met de spoordetector om de muonen te identificeren en hun momenta (en dus ook hun sporen) met grote precisie te bepalen. Zoals we zien in deze figuur, zal het muon door de calorimeters doorgaan zonder gestopt te worden en pas in de muon kamers tot interacties ondergaan. Deze muonkamers zullen ook snelle informatie leveren voor de triggers die de "interessante" events onderscheiden van de "niet interessante".

Hoofdstuk 2

Project

2.1 Productie van de Top Quark



Figuur 2.1: Feynmandiagram van de p-p interactie

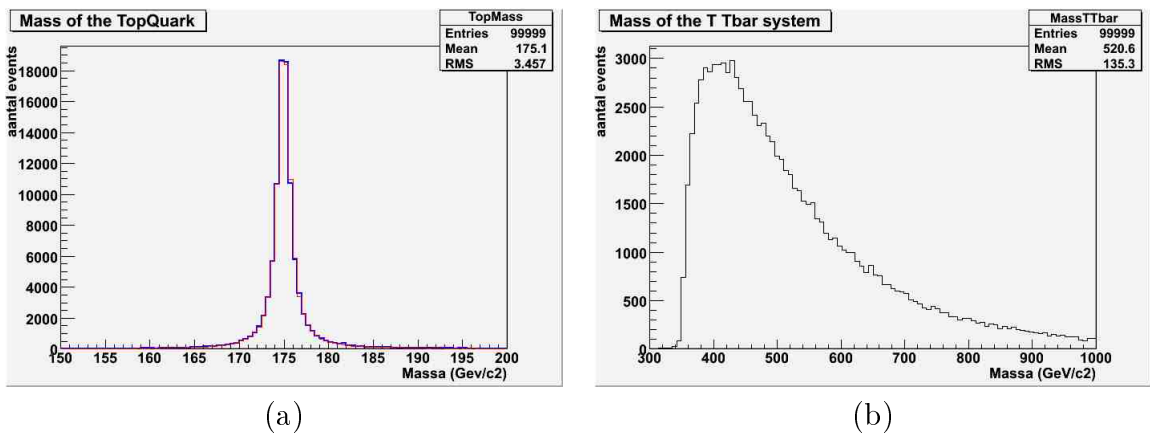
De productie van top quarks uit proton-proton botsingen in LHC zal in 90% van de gevallen gebeuren zoals weergegeven in het Feynman-diagram (Figuur 2.1). Hierin merken wij al op dat de top quarks niet geproduceerd worden door directe frontale botsing van de protonen zelf, maar door de sterke interactie tussens de gluonen die door de protonen uitgezonden worden. Er wordt bij de gluon-gluon interactie dus een $t\bar{t}$ -paar¹ gevormd die beide na ruwweg 10^{-25} [8] seconden vervallen in een W-boson en een b-quark. De b-quarks kan door "confinement" niet op zichzelf voorkomen dus zullen hadroniseren en jets vormen terwijl. Het W-boson vervalt in $\frac{1}{3}$ van de gevallen in een lepton en een neutrino en in de andere $\frac{2}{3}$ van de gevallen in een quark-anti-quark paar. Als beide W-bosonen vervallen in een lepton en een neutrino spreken we van

¹ \bar{t} is de anti-top quark

een di-leptonische gebeurtenis, als ze beide vervallen in een quark-anti-quark paar spreken we van een volledig hadronische gebeurtenis en als de ene in een lepton en neutrino en de andere in een quarkpaar vervalt spreken wij van een semi-leptonische gebeurtenis. De gevormde quarks zullen dan bijgevolg ook hadronizeren en jets vormen.

In eerste instantie gebruikten wij gegevens van de proton-proton botsingen uit een generator. Deze generator brengt data voort aan de hand van de interactie die weergegeven is in 2.1. Het verschil tussen deze gegevens en de simulatie-gegevens die zullen gebruikt worden in sectie 2.2 is dat wij hier directe informatie hebben over de neutrino's en de b-quarks, nl.: de impuls-energie viervector (E, P_x, P_y, P_z) . In de simulatie-gegevens is deze viervector alléén beschikbaar voor de leptonen. De b-quarks zullen daar gereconstrueerd moeten worden uit de jets, en de neutrino's uit de ontbrekende transverse energie.

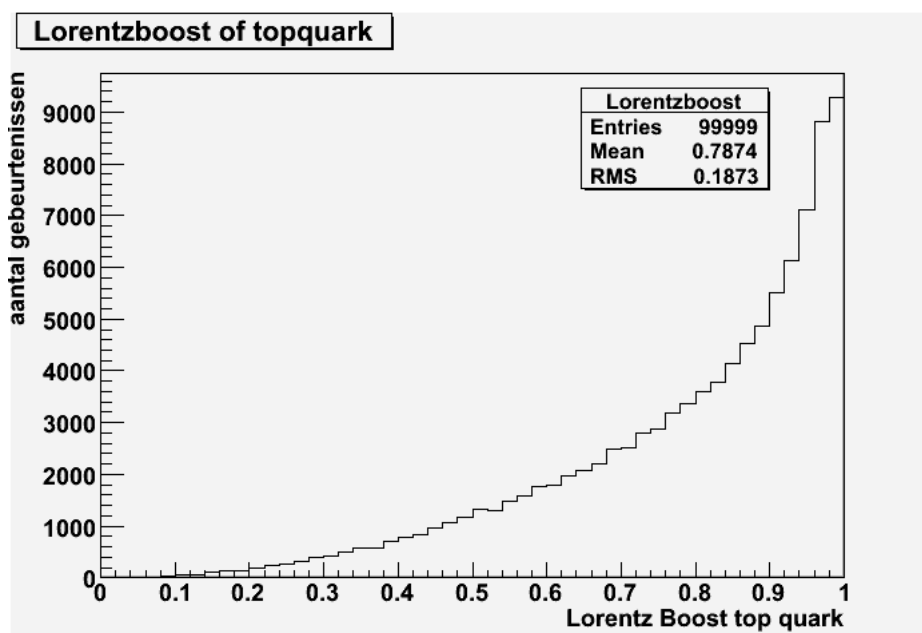
Als eerste plot zien wij in figuur 2.2 (a) de massa van de top quark (blauw) en de massa van de anti-top quark (rood). Wij zien een verdeling die scherp piet rond $175 \text{ GeV}/c^2$ ². Ook voor de anti-top quark krijgen wij dezelfde verdeling aangezien een anti-deeltje dezelfde massa heeft als zijn partner. In de (b) figuur is de massa van het $t\bar{t}$ -systeem geplott. Hierop zien wij dat dit systeem minstens een massa van twee keer de massa van de top quark heeft. De massa van de top quark bekomt men uit de relatie $E^2 = m^2 \times p^2$ met



Figuur 2.2: (a) Massa van de top quark, (b) invariante massa van het top quark-paar

$E = E_b + E_l + E_\nu$ en $p = p_b + p_l + p_\nu$, de massa van het $t\bar{t}$ -systeem bekomt men door deze vier-vectoren van beide deeltjes samen te nemen en dezelfde relatie

²dit is de waarde die aan de topquark gekoppeld werd in de generator



Figuur 2.3: Lorentz boost van de top quark

toe te passen.

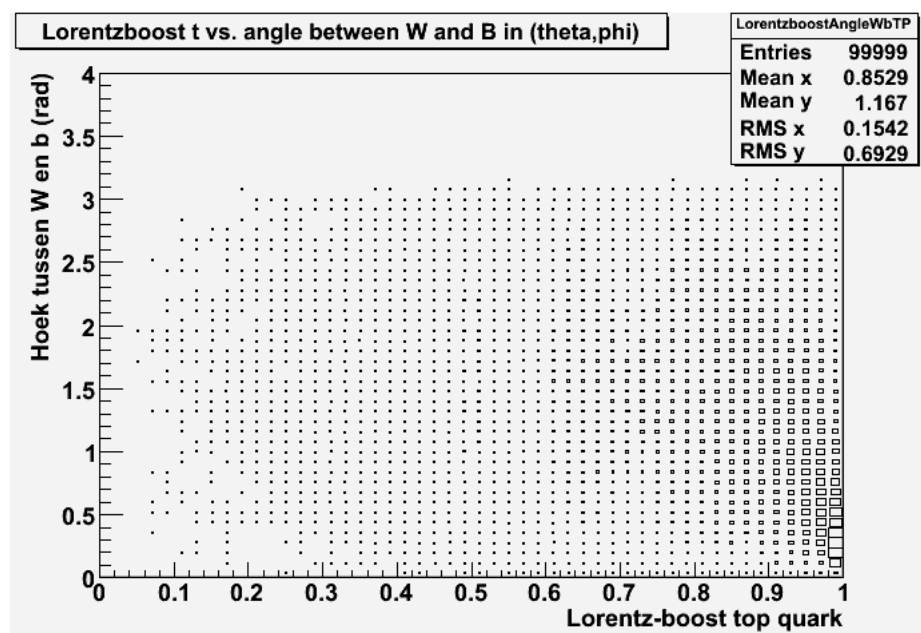
In figuur 2.3 zien wij de lorentz boost (β) van de top quark weergegeven.

$$\beta = \frac{|\vec{P}|}{E}$$

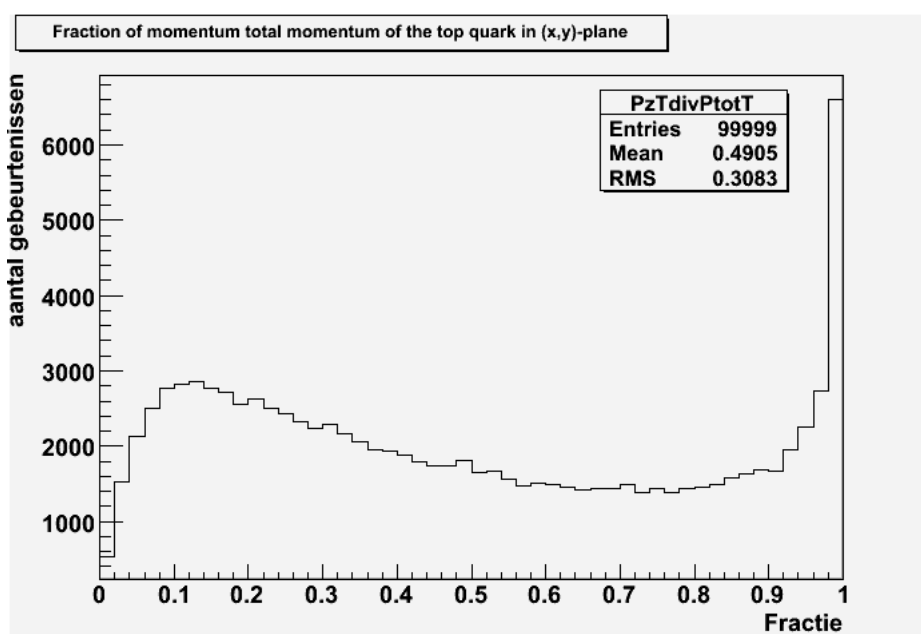
De waarden van β liggen tussen 0 en 1, waarbij $\beta = 0$ een situatie zonder boost is en $\beta = 1$ een situatie met een zeer sterke boost. Deze variabele is gekoppeld met de hoek tussen het W-boson en de b-quark (de vervalproducten van de top quark) gemeten in de (θ, ϕ) -ruimte³. Dit laatste is duidelijk te zien op de "boxplot" in figuur 2.4. Hier zien we dat de boost naar 1 gaat wanneer de hoek klein wordt (≤ 0.5 rad). In de plot van de lorentz boost van de top quark zien we dat deze boost gemiddeld groot is voor dit deeltje, waardoor we kunnen besluiten dat de hoek tussen het W-boson en de b-quark in het algemeen "klein" zal zijn.

Tot slot zien wij in figuur 2.5 een plot van de fractie van het totale momentum van de top quark die zich in het (x,y)-vlak (transverse vlak) bevindt. Hier merken wij op dat er twee maxima "pieken" zijn. Een kleine brede piek

³ θ is de hoek tussen het deeltje en de bundelpijp en ϕ is de azimutale hoek



Figuur 2.4: Lorentz boost van de top quark vs. de hoek tussen W en b



Figuur 2.5: Fractie van het totale momentum in het transverse vlak

bij kleine fracties waar de hoek tussen deze quarks en de bundelpijp klein is (en de bijdrage aan het totaal moment in het transverse vlak dus klein is). Een tweede grote en heel nauwe piek vinden we rond een fractie gelijk aan 1. Dit wil zeggen dat in dergelijke gebeurtenissen de z-component van het momentum van de top-quark verwaarloosbaar klein is en de top quark zich in het transverse vlak bevindt. Dit is ook de reden dat men bij dergelijke experimenten sterk geïnteresseerd is in wat er zich afspeelt in het transverse vlak.

2.2 Reconstructie en analyse

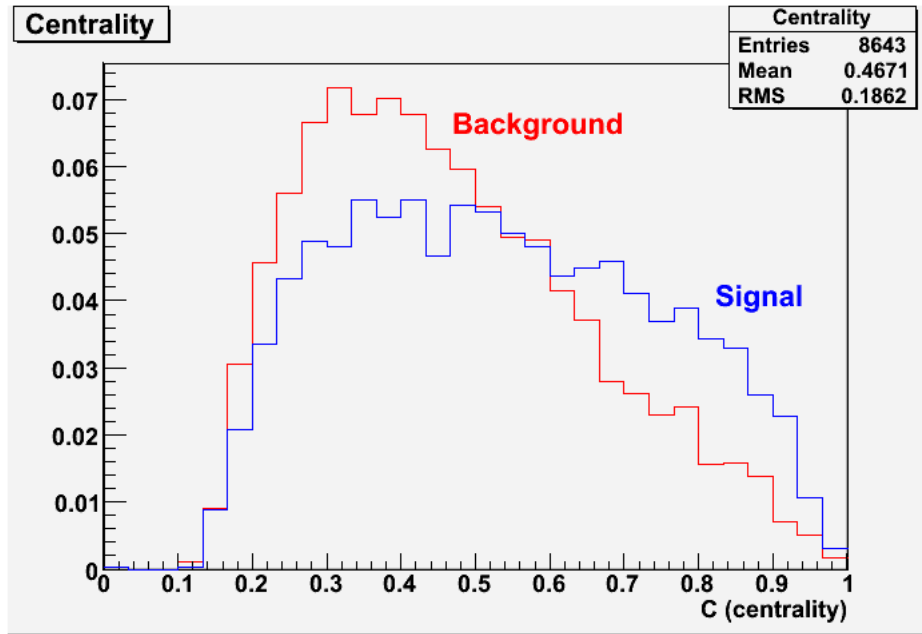
Na de studie van de kinematica van de top quarks ging ik over tot het analyseren van een gesimuleerde gegevens-set. Zoals eerder gezegd is het grote verschil tussen deze data en die uit de vorige sectie, dat deze data informatie over de deeltjes in de eindtoestand geeft zoals een detector dat zou doen. In de detector zullen steeds interacties plaatsvinden die wij niet willen bestuderen en dus achtergrond genoemd worden. Zo zullen wij dus in dergelijke experimenten nooit alleen de gewenste gebeurtenissen. Voor de studie van de top quark in CMS wil men uiteraard de goede gebeurtenissen, $pp \rightarrow t\bar{t}$ (zoals in fig. 1.3) scheiden van de achtergrond, die gegeven wordt door $pp \rightarrow W + \text{jets}$. Het is echter ook zo dat in CMS meer achtergrondgebeurtenissen zullen opgemeten worden dan de gewenste signalen omdat de werkzame doorsnede voor de achtergrond 500pb bedraagt en die voor de signaalgebeurtenissen slechts 80pb. Het doel van dit project is dan ook variabelen te zoeken waarvoor de verdelingen van signaal en achtergrond zo ver mogelijk uit elkaar liggen, wat het maken van sneden toelaat.

Voor deze analyse beschikte ik over gesimuleerde gebeurtenissen van het diagram in figuur 2.1 in het semi-leptonisch geval. Hierin had ik de volgende informatie:

- De vier hoogst energetische jets (hoogste transverse energie)
 - Energie
 - P_x, P_y, P_z
 - Waarschijnlijkheid dat ze afkomstig zijn van een b-quark
- Het lepton
 - Energie
 - P_x, P_y, P_z

- De ontbrekende transverse energie (P_x en P_y)

Mijn eerste variabele die een duidelijk verschil geeft tussen signaal en achtergrond is *centraliteit* [16] (Figuur 2.6). Deze variabele wordt bekomen door de scalaire som van de transverse momenta van de jets te delen door de scalaire som van de energieën van deze jets. De centraliteit is bijgevolg een variabele die de fractie van de totale energie van de deeltjes in het transverse vlak weergeeft.



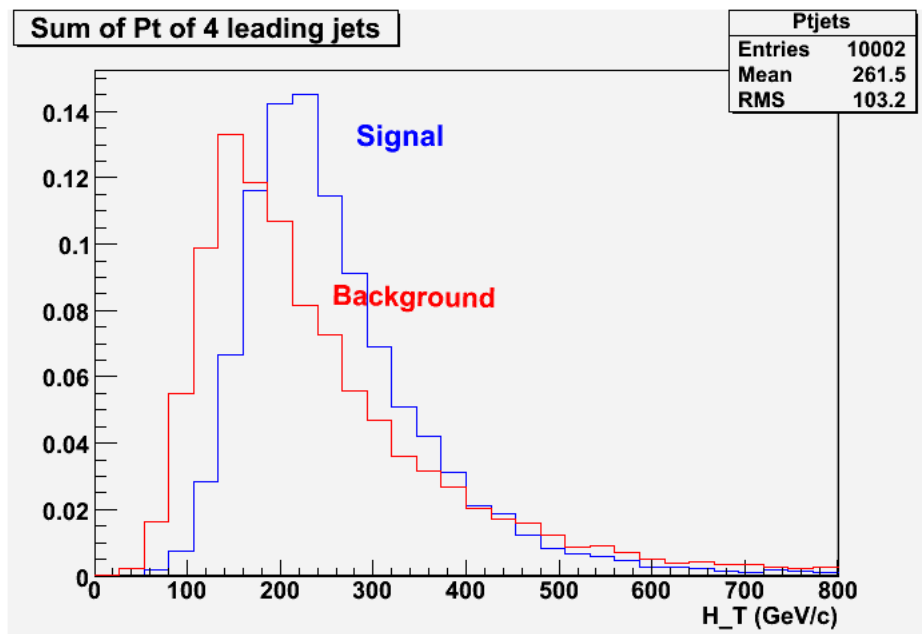
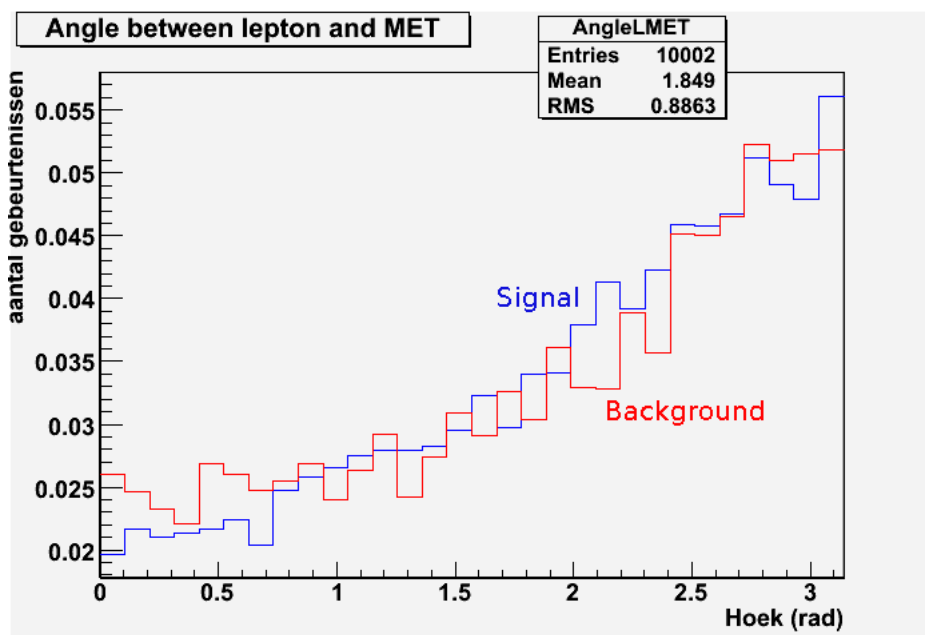
Figuur 2.6: Centraliteit

Een volgende variabele is H_T (Figuur 2.7). Dit is de scalaire som van de transverse momenta van de jets. Op deze plot zien we ook de duidelijke scheiding tussen de signaal- en achtergrondverdeling.

Een derde plot (Figuur 2.8) geeft de *azimuthale openingshoek tussen het lepton en de ontbrekende transverse energie* [16] weer. Deze variabele wordt bekomen uit

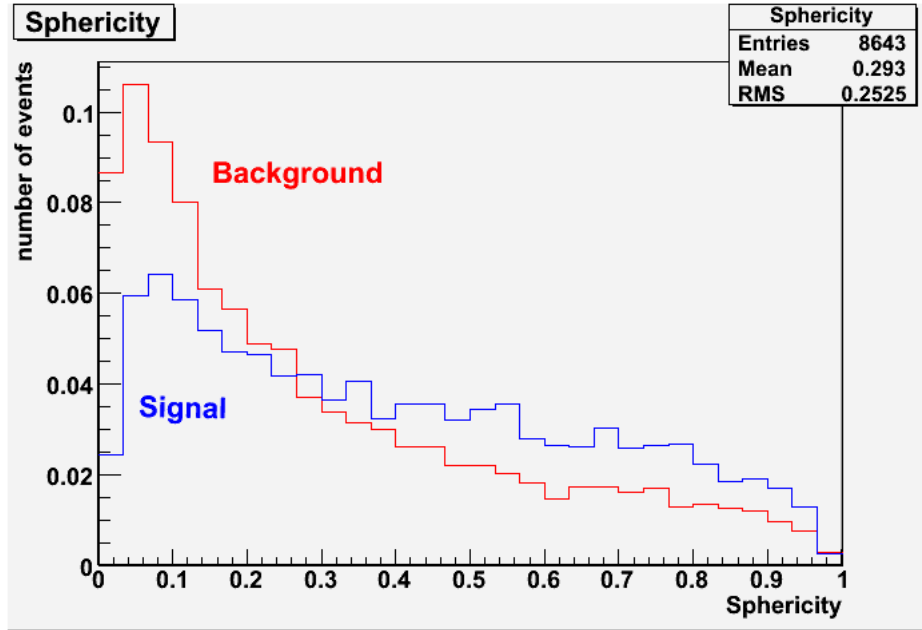
$$\vec{P}_T(l) \cdot \vec{P}_T(MET) = |\vec{P}_T(l)| \cdot |\vec{P}_T(MET)| \cdot \cos\theta_{l,MET}$$

In deze plot zijn de verschillen tussen signaal en achtergrond al minder groot als bij de vorige twee, maar toch zijn deze twee signalen hier nog duidelijk te onderscheiden.

Figuur 2.7: H_T 

Figuur 2.8: Azimuthale openingshoek tussen lepton en MET

Een volgende variabele die kan gebruikt worden om onderscheid te maken tussen signaal en achtergrond is de *sfericiteit* [16] (Fig 2.9) die weergeeft in



Figuur 2.9: Sfericiteit

hoeverre de deeltjes zich op een sfeer bevinden. Deze variabele heb ik gehaald uit de eigenwaarden van de kwadratische momentum tensor opgesteld met de momentum vier-vectoren van de vier jets en het lepton (dit resulteert hier een drie bij drie matrix).

$$\begin{pmatrix} S^{xx} & S^{xy} & S^{xz} \\ S^{yx} & S^{yy} & S^{yz} \\ S^{zx} & S^{zy} & S^{zz} \end{pmatrix}$$

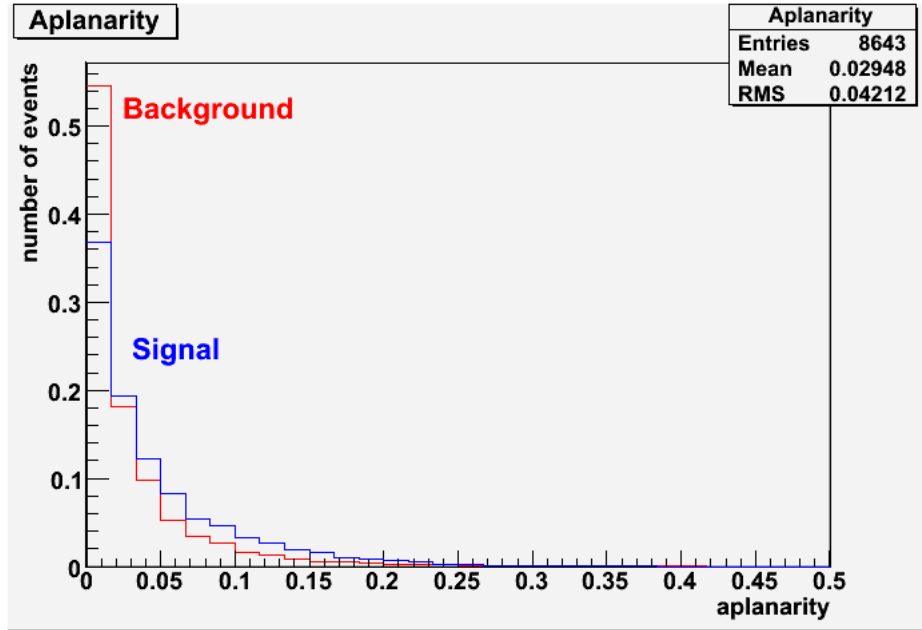
Waarbij de matrixelementen als volgt gedefiniëerd zijn:

$$S^{ab} = \frac{\sum_{i=1}^5 p_{ia} p_{ib}}{\sum_{i=1}^5 |p_i^2|}$$

De sfericiteit zelf is dan gelijk aan $\frac{3}{2}(\lambda_1 + \lambda_2)$, waarbij λ_1 en λ_2 de kleinste eigenwaarden van de matrix zijn⁴.

Uit de eigenwaarden van dezelfde matrix kan nog een andere variabele gevonden worden: *aplanariteit* [16] (Figuur 2.10). Deze is gedefiniëerd als $\frac{3}{2}\lambda_1$, waarbij λ_1 de kleinste eigenwaarde is van de matrix. Deze variabele geeft weer in hoeverre de deeltjes zich niet in hetzelfde vlak bevinden.

⁴De karakteristieke vergelijking van deze matrix is van de derde graad, wat dus voor elke gebeurtenis 3 eigenwaarden oplevert



Figuur 2.10: Aplanariteit

De verschillen tussen signaal en achtergrond zijn hier echter wel kleiner dan bij de variabele sfericiteit.

Een belangrijk punt in deze analyse is nagaan of de vijf genoemde variabelen niet gecorreleerd zijn met elkaar. Indien twee variabelen met elkaar gecorreleerd zijn, zal de tweede variabele geen extra informatie leveren en dus onbruikbaar zijn. Daarom heb ik de correlatiecoëfficiënten ρ_{ab} berekend tussen alle variabelen onderling (Tabel 2.1). We zien dus dat de grootste cor-

ρ_{ab}	Centraliteit	$\Delta\phi_{l,MET}$	H_T	Sfericiteit	Aplanariteit
Centraliteit	1	0.05	0.06	0.31	0.45
$\Delta\phi_{l,MET}$	0.05	1	0.11	0.002	0.06
H_T	0.06	0.11	1	0.07	0.03
Sfericiteit	0.31	0.002	0.07	1	0.45
Aplanariteit	0.45	0.06	0.03	0.45	1

Tabel 2.1: Correlatiecoëfficiënten tussen de variabelen

relatie tussen twee verschillende variabelen uit de groep van vijf, niet groter is dan 45%. Dit geeft aan dat de 5 variabelen onafhankelijk bijdragen tot de scheiding van signaal en achtergrond en ik dus vijf "goede" variabelen heb gevonden.

2.3 Conclusie

De opgave van dit project was het vinden van variabelen die toelaten een onderscheid te maken tussen signaal- ($pp \rightarrow t\bar{t}$) en achtergrondgebeurtenissen ($pp \rightarrow W+\text{jets}$) met behulp van een set gesimuleerde gegevens uit de CMS detector nabij de Large Hadron Collider. Ik vond vijf variabelen die, toegepast op deze data, verschillen vertonen tussen de verdelingen voor signaal en achtergrond. Dit is echter nog geen experimenteel bruikbaar selectiemechanisme. Een mogelijke methode om een selectiemechanisme op te bouwen uit deze variabelen is door deze te combineren met behulp van het *Support Vector Machine algoritme*. Dit algoritme zal de kleine verschillen in alle plots uitvergrooten en zo een éénduidige selectie te kunnen maken. Dit was oorspronkelijk gepland als het tweede deel van mijn analyse, maar wegens tijdsgebrek heb ik niet genoeg variabelen gevonden om dit algoritme te kunnen gebruiken. Het resultaat van mijn project is dus nog geen selectiemechanisme, maar een aanzet ertoe, die mits meer tijd kan uitgebouwd worden tot dergelijke selectie.

Bibliografie

- [1] Wikipedia: Standard model:
http://en.wikipedia.org/wiki/Standard_Model
- [2] Standard model:
http://www-d0.fnal.gov/Run2Physics/WWW/results/final/NP/N04C/N04C_files/standardmodel.jpg
- [3] Fundamentele krachten & Feynman diagrammen:
http://www.hep.phy.cam.ac.uk/drwl/particleandnuclear/Lectures_2_2007.pdf
- [4] Kleurlading:
http://en.wikipedia.org/wiki/Color_charge
- [5] W,Z-bosonen:
http://en.wikipedia.org/wiki/W_and_Z_bosons
- [6] Higgs boson:
http://en.wikipedia.org/wiki/Higgs_Boson
- [7] Mass top quark:
<http://pdg.web.cern.ch/pdg/2006/listings/q007.pdf>
- [8] Top Quark:
http://en.wikipedia.org/wiki/Top_quark
- [9] Jets:
http://en.wikipedia.org/wiki/Jet_%28particle_physics%29
- [10] Hadronisatie:
<http://en.wikipedia.org/wiki/Hadronization>
- [11] Quarks:
<http://en.wikipedia.org/wiki/Quark>

- [12] Large Hadron Collider (Wikipedia):
http://en.wikipedia.org/wiki/Large_Hadron_Collider
- [13] Large Hadron Collider (CERN):
<http://lhc-machine-outreach.web.cern.ch>
- [14] LHC accelerator challenges:
<http://lhc.web.cern.ch/lhc/general/acphys.htm>
- [15] Compact Muon Solenoid:
<http://cmsinfo.cern.ch/outreach>
- [16] Variabelen: CMS Note 2006/064