

Dissertation ETH Nr. 14373

Studies of Resonant and Non-Resonant Weak Boson Pair Production at LHC

A dissertation submitted to the
Swiss Federal Institute of Technology Zürich

for the degree of
Doctor of Natural Sciences

presented by

Dario E.G. Zürcher

Dipl.-Phys. ETH-Zürich

born January 28th, 1971

in Lugano, Switzerland

Citizen of Switzerland (from Thalwil ZH)

Accepted on the recommendation of

Prof. Dr. F. Pauss	examiner
Prof. Dr. Z. Kunszt	co-examiner
Dr. D. Renker	co-examiner

Abstract

The weak boson pair production at LHC is especially important for the study of the electroweak symmetry breaking mechanism. A resonant diboson signature could reveal the presence of a Standard Model Higgs boson. In this case, precise predictions for the Higgs signal and for the non-resonant diboson background are mandatory. In the absence of a Higgs scalar, the dynamics of the interactions between weak bosons must reveal a new structure and some new physics should contribute to the normalization of the diboson cross section in the TeV range.

In this work, two experimental aspects are discussed first. The knowledge of the expected WW , WZ and ZZ rates at LHC is the starting point of each diboson analysis. For this purpose, a method to measure parton luminosities and parton distribution functions with unprecedented accuracy is proposed. The method is especially valid as monitor for diboson events. At LHC, W and Z bosons will be preferentially measured through their leptonic decays. Using test beam data and GEANT simulations, the probability of $\pi^\pm - e^\pm$ misidentification is evaluated to be between 0.2% and 0.01% for pion momenta between 5 and 50 GeV.

Simulations of weak boson pair signatures, computed with a new parton level Monte Carlo program, allow to investigate the effects of higher order ($\mathcal{O}(\alpha_s)$) corrections on WW , WZ and ZZ kinematics. It is found that the implementation of the NLO diagrams results in a huge increase of diboson events with high- p_t jets, leptons or missing transverse energy. Large corrections are therefore expected for an analysis which uses these observables to discriminate between signal and background.

All these studies improve our understanding of weak boson pair signatures at LHC. A Standard Model Higgs boson produced via gluon fusion should be visible from 100 GeV up to 600 GeV, while the weak boson fusion production mode provides signals in the 300 GeV to 1 TeV range. In this work, Higgs production via the fusion of vector bosons is studied in detail also for masses between 110 and 300 GeV. It is shown that these signatures provide excellent signals. The required luminosity to observe a 5σ signal ranges from 3 fb^{-1} , for a 170 GeV Higgs, up to about 25 fb^{-1} , for a Higgs mass of 125 or 300 GeV. Furthermore, using both the vector boson fusion and the gg fusion production modes, the ratio between the Higgs couplings to gauge bosons and fermions can be measured with an accuracy of the order of 15% for an integrated luminosity of 100 fb^{-1} .

Zusammenfassung

Die Paarproduktion von Vektorbosonen ($V = W, Z$) am LHC spielt eine wichtige Rolle für das Studium des Mechanismus der elektroschwachen Symmetriebrechung. Die Existenz des damit verbundenen Higgs Teilchens könnte mit Hilfe der resonanten VV -Produktion nachgewiesen werden. Voraussetzung dafür ist eine genaue Vorhersage sowohl für das Higgs Signal als auch für den nicht-resonanten VV Untergrund. Falls kein skalares Higgs Teilchen existiert, sollte die Dynamik der Wechselwirkung zwischen Vektorbosonen Rückschlüsse über neue Strukturen erlauben und somit Informationen über neue Physik im TeV-Energiebereich ermöglichen.

In der vorliegenden Arbeit werden zuerst zwei experimentelle Aspekte diskutiert. Die Zahl der zu erwartenden WW , WZ und ZZ Ereignisse am LHC ist ein wichtiger Input für jede Bosonpaar Analyse. Dazu wird eine neue Methode vorgeschlagen, wie man Partonluminositäten und Partonverteilungsfunktionen mit sehr guter Genauigkeit messen kann. Diese Methode eignet sich auch ausgezeichnet die Zahl der Bosonpaarereignisse zu bestimmen. Am LHC werden W und Z Boson vor allem via dem leptonischen Zerfall nachgewiesen. Mit Hilfe von Teststrahlraten und GEANT Monte Carlo Studien wird die Wahrscheinlichkeit der $\pi^\pm - e^\pm$ Falschidentifikation ermittelt. Sie ergibt einen Beitrag im Bereich von 0.2% bis 0.01% für Pionimpulse zwischen 5 und 50 GeV.

Ein neues Parton-level Monte Carlo Programm für VV -Produktion erlaubt die Effekte höherer Ordnung (next-to-leading order (NLO), $\mathcal{O}(\alpha_s)$) in Bezug auf WW , WZ und ZZ Kinematik zu studieren. Diese Simulationen zeigen, dass Ereignisse mit sehr hohen transversalen Energien der Jets, Leptonen und grosser fehlender transversaler Energie viel grössere Ereignisraten vorhersagen, falls NLO Diagramme mitberücksichtigt werden. Analysen, welche diese Observable als diskriminierende Signaturen zwischen Signal und Untergrund verwenden, erfahren daher grosse Korrekturen.

Diese durchgeführten Studien verbessern somit unser Verständnis der VV Produktion am LHC. Frühere Studien zeigten, dass ein Standardmodell Higgs Signal im Massenbereich von 100 bis 600 GeV via dem Gluon-Gluon Fusionsprozess beobachtbar ist. Hingegen spielt der Vektorboson Fusionsprozess eine wichtige Rolle für beobachtbare Higgs Signale im Massenbereich von 300 bis 1000 GeV. In der vorliegenden Arbeit wird die Higgsproduktion via dem Vektorboson Fusionsprozess im Detail auch für den Higgsmassenbereich zwischen 110 und 300 GeV studiert, und gezeigt, dass ein Signal beobachtbar ist. Um ein 5σ Signal für eine Higgsmasse von 170 GeV zu erhalten, ist eine integrierte Luminosität von 3 fb^{-1} notwendig. Für eine Higgsmasse von 125 GeV oder 300 GeV hingegen sind 25 fb^{-1} nötig. Zusätzlich, unter Verwendung von Vektorboson und Gluon-Gluon Fusionsproduktion, kann man mit einer Genauigkeit von 15% (100 fb^{-1}) die Kopplung des Higgs Teilchens zu Eichbosonen und Fermionen messen.

Contents

1	Introduction	7
2	Theoretical motivation	12
2.1	Mass generation within the Standard Model	12
2.2	Standard Model 2001: status report	17
2.3	The Challenge: finding the Standard Model Higgs Boson at LHC	20
2.3.1	The “well known”	20
2.3.2	Additional signatures: towards a better determination of the Standard Model Higgs couplings	25
3	Experimental aspects of a 14 TeV pp-collider experiment	27
3.1	The LHC project	28
3.1.1	The pp collider	28
3.1.2	ATLAS, LHC-b and ALICE	29
3.1.3	CMS: the “Compact” Muon Solenoid	31
3.1.4	Time schedule	41
3.2	Event simulation and reconstruction	42
3.2.1	The complexity of high-energy processes	42
3.2.2	Reconstruction parameters	43
3.3	The simulation frame	46
3.3.1	CMS detector simulations	46
3.3.2	Physics simulations	47
4	A parton luminosity monitor and applications for the diboson SM cross section prediction	64
4.1	The internal structure of the proton	64
4.2	Luminosities and event rates at the LHC	67
4.3	Rates of $pp \rightarrow W^+, W^-$ and Z^0 events	69
4.4	Using the W^\pm and Z^0 production to constrain the q and \bar{q} structure functions	71
4.5	Gluon and heavy quarks distribution functions	81
4.6	Precise prediction of diboson events	83

5	π-e misidentification as new background source	90
5.1	The energy measurement in the electromagnetic calorimeter . . .	90
5.1.1	Light Yield	91
5.1.2	The Avalanche Photo Diodes	91
5.1.3	The Nuclear Counter Effect	92
5.2	Charged pions crossing the ECAL	93
5.2.1	Nuclear Counter Effect contribution to the ECAL response	93
5.3	Results from the 1998 test beam runs	97
5.3.1	Measurement of the Nuclear Counter Effect	97
5.3.2	Pions as “fake” electrons candidates	100
5.4	Results from 1999 test beam runs	103
5.5	Discrimination methods for isolated pions and electrons	105
5.5.1	Shower shape selection of electrons	106
5.6	Expected pion-electron misidentification and possible consequences for the CMS physics program	110
5.6.1	Expected ratios of “fake electron” candidates	110
5.6.2	Misidentified pions as source of new backgrounds?	112
6	Next-to-leading order (NLO) effects in the vector boson pair production	115
6.1	NLO generators	116
6.2	Results	116
6.2.1	NLO effects in the WW production	118
6.2.2	NLO effects in the WZ production	122
6.2.3	NLO effects in the ZZ production	124
6.3	Discussion	125
7	Observing light Standard Model Higgs through weak boson fu- sion signatures	129
7.1	Towards a determination of the Higgs couplings	129
7.2	Event selection	131
7.2.1	$140 \text{ GeV} \leq M_{\mathbf{H}} \leq 200 \text{ GeV}$	131
7.2.2	$110 \text{ GeV} \leq M_{\mathbf{H}} \leq 140 \text{ GeV}$	141
7.2.3	$200 \text{ GeV} \leq M_{\mathbf{H}} \leq 300 \text{ GeV}$	146
7.3	Results	150
7.3.1	$110 \text{ GeV} \leq M_{\mathbf{H}} \leq 200 \text{ GeV}$: PYTHIA vs. original studies [25, 26]	150
7.3.2	$110 \text{ GeV} \leq M_{\mathbf{H}} \leq 200 \text{ GeV}$: a full PYTHIA analysis	158
7.3.3	$200 \text{ GeV} \leq M_{\mathbf{H}} \leq 300 \text{ GeV}$	164
7.4	Required integrated luminosity for a Higgs discovery: an update .	166
8	Conclusions and outlook	169

A	Hardware related topics	174
A.1	PSI 1997: crystal and APD tests	174
A.1.1	R&D for the PbWO_4 crystals	175
A.1.2	R&D for the APD's	179
A.2	CERN 1998-2000: ECAL tests	180
A.2.1	1998	180
A.2.2	1999	181
A.2.3	2000	182

"Dissertations are not finished; they are abandoned."
- Fred Brooks

Chapter 1

Introduction

The Standard Model (SM) [1, 2] answers many of the questions about the structure and stability of matter at the smallest scales (10^{-18} m) and highest energies (~ 200 GeV) accessible to current experiments. This quantum field theory describes the behavior of spin-1/2 point-like fermions, interacting through spin-1 gauge bosons. Within this representation, matter is seen as the result of the interactions of six types of quarks and six types of leptons through three fundamental forces.

While the Standard Model provides a very good description of phenomena observed by experiments, it is still an incomplete theory. One problem is that the Standard Model can not explain why particles exist as they do. Within this theory, some parameters, like the particle masses, are still free. For instance, given that both the photon and the W bosons are force carrier particles: why is the photon massless and the W massive? Furthermore, the electron mass (511 keV) and the top quark mass (~ 175 GeV) differ by more than 5 orders of magnitude. What is the origin of this mass scale? These problems arise because the Standard Model lacks a mathematical model for finding a pattern for particle masses.

As a possible explanation of this mass hierarchy, the existence of the so-called Higgs field was proposed [3]. Within the Electro Weak Symmetry Breaking mechanism, the presence of the Higgs field leads to the emergence of massive vector bosons, which mediate the weak interactions, while the photon remains massless. This hypothesis was supported by the experimental discovery of both W and Z bosons (the carriers of the Electro Weak force) at masses of the order of 100 GeV. Within this mechanism, the Higgs field interacts then with other particles to give them mass. In the minimal scenario, the field should manifest itself as a neutral scalar called “Higgs boson”.

Up to now, the Standard Model provided a very successful description of the interactions of the components of matter. However something important is still missing: the Higgs particle has not yet been observed! This leaves many open questions about the exact way in which the symmetry breaking mechanism is realized; *i.e.* how the field couples to fermions and gauge bosons, whether it will

show up as a resonance and, within this last scenario, what is the mass of the possible Higgs boson.

In fact, the Standard Model cannot predict the Higgs boson mass. Nevertheless, the requirement for the Higgs couplings to remain finite and positive up to an energy scale Λ put some constraints on the possible mass range [4]. For example, according to the measurements of the top mass m_t and of the strong coupling α_s and assuming the validity of the Standard Model up to Planck scale $\Lambda_{\text{Pl}} \sim 10^{19}$ GeV, the Higgs mass should then lie in the range between 130 and 190 GeV [5]. On the contrary, for smaller or larger Higgs masses, there should be some new physics which set in well below Λ_{Pl} .

From the experimental side, direct searches at LEP [6, 7] set a lower bound on the Higgs mass. The present limit is $M_H > 114.1$ GeV [8] at the 95% CL¹. The mass of the Higgs boson is also constrained by precise measurements of the electro-weak data, which are sensitive to loop corrections. Assuming the overall validity of the Standard Model, a global fit to all the present data leads to $M_H = 88^{+53}_{-33}$ GeV [9], with an upper limit of 196 GeV at the 95% CL.

Other crucial considerations can be made without reference to the scale Λ . The Higgs boson self couplings and its couplings to the W and Z bosons grow as the Higgs mass increases [10]. Moreover, without the presence of an Higgs resonance, also the amplitudes of the weak boson scatterings ($VV \rightarrow VV$, with $V \equiv W^\pm$ or Z^0) are expected to rise with growing diboson mass. These features have very important consequences: either the Higgs mass lies below an upper limit of about 800 GeV or there should be some new mechanism which enters in the interactions between weak boson pairs to control the cross-section behavior. Without a Higgs scalar, the dynamics of the interactions between massive vector bosons must then reveal a new structure and, to respect the unitarity limit, this should happen at center-of-mass energies of the order of 1 TeV. The ability to reach this energy scale and to detect diboson events ensures a valuable collection of informations on the nature of the electroweak symmetry breaking.

The LHC [11] project has been developed with the aim not only to make a significant step in the understanding of the Higgs mechanism but also in general in the explanation of physics processes which occur at the smallest scales and highest energies, allowing also to make further progress in the understanding of the early universe.

This machine will be operational in the year 2006 at CERN in Geneva and it will bring protons into head-on collisions at very high center-of-mass energies (14 TeV). As next research instruments in world's particle physics domain, the LHC detectors will allow to penetrate still further into the structure of matter and

¹These last combined results of the four LEP experiments show also an excess of events which can be interpreted as production of a Standard Model Higgs boson with a mass (115.6 GeV) higher than the quoted limit. However the observed 2.1σ deviation is not sufficient to claim an "evidence".

will recreate the conditions prevailing in the Universe just about 10^{-12} second after the “Big Bang”, revealing the behavior of fundamental particles of matter at distances/energies never studied before.

The aim of the project is to produce high center-of-mass energies combined with high luminosity, which is designed to reach $\mathcal{L}_{\text{Design}} = 10^{34} \text{ cm}^{-2}\text{s}^{-1}$. The LHC fulfills these requirements, but the high luminosity leads to difficult experimental conditions. The total cross-section will be very large, *i.e.* about 100 mb, which results in about 20 interactions per crossing at the design luminosity. With a crossing frequency of ~ 40 MHz, this is equivalent to an interaction rate of $\sim 10^9$ Hz!

The detection of processes with signal to total cross-section ratios of about 10^{-11} , as for example for a 150 GeV Higgs boson, will be a difficult experimental challenge.

In this work I present some aspects of the problems related to the Higgs search at LHC. Particularly, I concentrate on the approach to the Higgs mechanism via the detection of weak boson pairs. This is justified by the fact that, within the diboson interactions, something must happen beyond the current knowledge of the high energy physics. Consequently, these processes constitute the ideal start point for a study of the electroweak symmetry breaking mechanism.

I discuss arguments which cover several aspects of this problems, starting from topics concerning the cross section measurement and concluding with Monte Carlo studies of Higgs discovery channels, treating also hardware related topics like backgrounds arising from detector effects.

The work described in this paper can not give a complete coverage of all experimental difficulties for a Higgs boson detection. The aim is however to give an objective view, at different levels, of the possible problems related to the Higgs discovery.

Topics overview

After a theoretical introduction to the Standard Model (chapter 2), the LHC project, one of the general purpose detectors (CMS), as well as the simulation set-up, are described in more details in chapter 3.

In chapter 4 problems related to the cross section measurements are discussed. Before starting to study the production of vector bosons, some crucial input parameters are needed. One of them is the effective amount of diboson events which, according to the theory, are expected to be produced. The event rate for a process results from the combination of the proton–proton luminosity of the collider with the probability that this process occurs within a $p - p$ collision. However, by inelastic scattering of two 7 TeV protons, the interesting physics processes arise from a direct interaction between the quasi-free quarks and gluons within the proton. The interpretations of essentially all proposed measurements

require thus a good knowledge of the parton distribution functions (PDF) at the relevant Q^2 . With the help of the PDF's, one can pass from the event cross section given in a proton–proton collision to the one given in a parton–parton collision.

In chapter 4 a new approach to determine the LHC luminosity is investigated [12, 13]: instead of employing the p – p luminosity measurement, a direct measurement of the different parton–parton luminosities is suggested. In addition, the aim is to study to what extent the single boson productions $q\bar{q} \rightarrow W^\pm, Z^0$ are able to constrain the parton distribution functions.

The goal is an accurate prediction of $q\bar{q}$ related scattering processes like $q\bar{q} \rightarrow WW, WZ, ZZ$. A precise estimate of these processes enables to better control the background for a Standard Model Higgs in the WW or ZZ decay channels. Moreover, in case that any light Higgs scalar would have been found, the Standard Model diboson expectations would be the basis for further studies in the higher masses domain.

One problem of today's particle physics experiments lies in the ability to relate the signatures left in the detector with the true physical process. The difficult task is to reconstruct what happened at the interaction point, starting from the different energy depositions within its dedicated parts.

In chapter 5 I study the interactions of electrons and pions with one of the components of the CMS detector [14]. As the detection of vector bosons, with subsequent decay into leptons $W^\pm \rightarrow l^\pm\nu$, $Z^0 \rightarrow l^+l^-$, is a crucial point for my analysis, an essentially background free identification of isolated, high- p_t electrons is mandatory. Given the huge amount of π^\pm produced at LHC, other sources of background for the vector boson production could arise in case that a charged pion is misidentified as an isolated electron.

Using GEANT Monte Carlo simulations [15] and test-beam results, the amount of "fake" $W \rightarrow e\nu$ and $WW \rightarrow l\nu e\nu$ events can be estimated [16].

Besides the luminosity, the knowledge of the event cross section is necessary to obtain the probability that a physics process occurs. This knowledge depends on the accuracy of the theoretical calculations.

For the description of the weak boson pair productions W^+W^- , $W^\pm Z^0$, $Z^0 Z^0$, with their subsequent decays into leptons, two numerical parton-level Monte Carlo programs have become available during 1999: MCFM[17] and DKS[18]. These packages consider the hadronic diboson production by including the complete next-to-leading-order (NLO, or $\mathcal{O}(\alpha_s)$) corrections.

On the other hand, most of the physics studies carried out so far for LHC have been made using the PYTHIA [19] Monte Carlo, which is a LO simulation tool for the diboson production. In the search for massive, highly energetic and peculiar diboson events, it is often suggested to optimize cuts which use large values of the boson transverse momenta, $p_t(V)$, and large values of the diboson invariant mass M_{VV} . The problem is that these kinematical regions might be affected by

large $\mathcal{O}(\alpha_s)$ corrections and therefore the `PYTHIA` estimations might be wrong. This topic is investigated in chapter 6, by comparing predictions from `PYTHIA` with a NLO (DKS) Monte Carlo.

As the search for the origin of electroweak symmetry breaking is one of the main tasks of LHC, many studies have been devoted to find methods which assure a Higgs boson discovery over the entire allowed mass range. These analysis have demonstrated the capability of this $p-p$ collider to discover the SM Higgs. More recent studies have concentrated on a larger variety of other Higgs production and decay channels, with the aim to determine the Higgs boson couplings to known fermions and gauge bosons [20]. The production mode via the annihilation of vector bosons (the so-called Weak Boson Fusion $VV \rightarrow H$) promises to be a copious source of Standard Model Higgs and, at the same time, it gives access to another production mode compared to the dominant gluon-gluon fusion process ($gg \rightarrow t\bar{t} \rightarrow H$). Early studies [21] took the weak boson fusion into account only for the high Higgs mass domain (*i.e.* $M_H \gtrsim 800$ GeV). During the past years, other studies extended the analysis of these signatures down to about $M_H \sim 300$ GeV [22, 23, 24].

Recently, some new studies [25, 26] reconsidered this production mode, with subsequent Higgs decay to pairs of W bosons, also for much lower Higgs masses ($115 \text{ GeV} \leq M_H \leq 200 \text{ GeV}$). In these publications, the signatures were analyzed at “parton level”, namely, using directly the kinematics of the quarks involved in the weak boson fusion process and of the different Higgs decay products. The selection cuts were subsequently developed starting from these variables. The authors of these studies estimated that these signatures are visible in the Higgs mass range 115-200 GeV. Furthermore, it is concluded that such channels are also competitive for a Higgs discovery.

In chapter 7 these parton level studies are reconsidered using `PYTHIA`. This program allows us to take the full hadronization processes into account and, starting from the “detectable” particles (*i.e.* those consistent with the CMS detector performances), to reconstruct jets and isolated leptons. The gluon radiation pattern from the initial state quarks is also considered, as well as the possible presence of underlying events. In addition, the analysis is extended also to the mass range $200 \text{ GeV} \leq M_H \leq 300 \text{ GeV}$.

Chapter 2

Theoretical motivation

2.1 Mass generation within the Standard Model

One of the main challenges in high energy physics is the discovery of the last important missing piece of the Standard Model: the Higgs boson. Its existence is related to the generation of elementary particle masses.

The history of the Standard Model starts during the sixties, when Glashow, Weinberg and Salam [27, 28] proposed a gauge theory unifying weak and electromagnetic interactions. In 1972 a similar theory for the strong interactions was first suggested by Fritzsche and Gell-Mann [2]. Finally, in 1974, during a summary talk for a conference in London, Iliopoulos presented for the first time in the same report the new quantum field theory, unifying electroweak and strong interactions within an single model, since then known under the name of “Standard Model” (SM) of particle physics [29].

What emerges from this historical process is a new representation of matter, now seen as the result of the interactions of spin-1/2 point-like fermions, whose interactions are mediated by spin-1 gauge bosons. This theory describes the spin-1 gauge bosons as a consequence of local gauge invariance applied to the fermion fields, where they arise as a manifestation of the symmetry group of the theory, *i.e.* $SU(3) \times SU(2) \times U(1)$ [1, 2].

The fundamental spin-1/2 fermions are leptons and quarks, where the strong force is a prerogative of the latter (*i.e.* leptons lack the “strong charge”). The left-handed states are doublets under the $SU(2)$ “weak isospin” group I , while the right-handed states are singlets under the $U(1)$ group of “weak hyper-charge” Y . The theory of strong interaction (QCD¹) resumes the structure of quantum electrodynamics (QED). The spin-1/2 quarks are triplets under the $SU(3)$ group and thus carrying an additional color charge. Eight massless spin-1 gluons are the quanta of the strong-interaction field. In contrast to the neutral photons of

¹Since strong interaction deals with the so-called “color” charge this theory was named quantum chromo-dynamics (QCD).

2.1 Mass generation within the Standard Model

electromagnetism, the gluons carry color charges and are thus self-interacting. The implication of this important difference is that the QCD coupling α_s is small for large momentum transfers but large for small momentum transfers. This is the reason why quarks, unlike leptons, are not directly observed as free objects: attempting to free a quark produces a jet of hadrons through the generation of quark-antiquark pairs and gluons. In other words, the form of the strong potential leads to the confinement of quarks inside color-neutral hadrons.

generations	1 st	2 nd	3 rd
quarks	$\begin{pmatrix} u \\ d \end{pmatrix}$	$\begin{pmatrix} c \\ s \end{pmatrix}$	$\begin{pmatrix} t \\ b \end{pmatrix}$
leptons	$\begin{pmatrix} \nu_e \\ e \end{pmatrix}$	$\begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}$	$\begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}$
mass range for charged fermions ($m_\nu = 0$)	0.5~5 MeV	0.1~1 GeV	2~200 GeV

Table 2.1: Fermion families.

As table 2.1 shows, there are three generations of fermions, which can be distinguished only by the different mass. The leptons have either electrical charge -1 or 0 . The charged leptons are the electron e , the muon μ and the τ lepton. These latter are paired to three neutrinos ν_e , ν_μ and ν_τ . Similarly, there are three quarks with electric charge $2/3$, up u , charm c and top t , paired with three quarks with electric charge $-1/3$, down d , strange s and bottom b .

The presence of at least three generations had not only already been predicted before the important discovery of the third and heaviest one, but even before the discovery of the charm quark! After Christenson, Cronin, Fitch and Turlay in 1964 discovered that the CP-symmetry was broken in the decay of neutral kaons, in 1972 Kobayashi and Maskawa showed that the Standard Model with three families of fermions could explain the observed CP-violating phenomena. In fact, between the three generations of quarks there is mixing, parameterized by the so-called Cabibbo-Kobayashi-Maskawa (CKM) [30] matrix. This means that in the quark sector there are couplings between different quark families, mediated by the weak interactions.

Neither the existence of this mixing nor the huge mass difference between electrons (511 keV) and top quarks (~ 175 GeV) is explained by the Standard Model. Actually, the origin of the family structure remains a mystery.

For the three fundamental forces included in the Standard Model, table 2.2 sum-

Force	Gauge symmetry	Bosons	Symbol	Strength ($Q^2 = M_Z^2$)
Weak	SU(2)	vector bosons	W^\pm, Z^0	$\alpha_{\text{weak}} \approx 0.03$
Electromagnetic	U(1)	photons	γ	$\alpha_{\text{em}} = 1/128$
Strong	SU(3)	gluons	g	$\alpha_s = 0.118$

Table 2.2: Fundamental forces, with their associated symmetry, bosons and strength (at $Q^2 = M_Z^2$).

marizes the symmetry group, the carrier of the force and its strength.

To explain the observed mass spectrum of bosons and fermions, a theory called “spontaneous symmetry breaking mechanism” or “Higgs mechanism”² was developed during the years 1967/68 by Weinberg and Salam [28], in parallel with the unification of electromagnetic and weak interactions. Within the $SU(2) \times U(1)$ symmetry group, the mechanism gives a mass to the gauge bosons W and Z , keeps $m_\gamma = 0$ and at the same time makes the theory renormalizable. This may be achieved with the help of fundamental scalar Higgs fields, which interact with each other [31, 32] such that they acquire a nonzero vacuum expectation value and the $SU(2)_L \times U(1)_Y$ symmetry is spontaneously broken down to the electromagnetic $U(1)_{\text{em}}$ symmetry. This leads to the emergence of massive vector bosons, the W and Z , which mediate the weak interactions, while the photon of electromagnetism remains massless. The fermions obtain their masses by interacting with the vacuum Higgs field.

In the minimal version ($\hat{=}$ the Standard Model), the theory requires one Higgs field doublet and predicts a single neutral Higgs boson. This scalar couples to the weak bosons W and Z with strength g (the coupling constant of the $SU(2)$ gauge theory) and to quarks and leptons of mass m_f with the strength $gm_f/2M_W$. Therefore, the Higgs coupling to W , Z and to t -quarks is substantial, while the coupling to light fermions is very small.

Up to now, the Standard Model provided a very good description of observed high energy processes and obtained important successes with the discovery of the massive weak bosons³. Nevertheless, the way in which the Higgs mechanism realizes itself remains a mystery and there is no direct evidence either for or against the simplest scenario. New informations would come from the direct observation of one or more Higgs bosons.

Its properties, decay widths and production mechanisms, can be predicted if the

²Often referred to as “Electro Weak Symmetry Breaking” (EWSB).

³C. Rubbia and S. van der Meer received the Nobel Prize for Physics for their work which culminated in the discovery of the W boson and Z boson at CERN in 1983 (UA1 and UA2 experiments).

2.1 Mass generation within the Standard Model

Higgs mass is given.

In the Standard Model, the Higgs mass $M_H = \sqrt{2\lambda}v$ is proportional to the vacuum expectation value v of the Higgs field, which is fixed by the Fermi coupling, while the quartic Higgs coupling λ , and thus M_H , is not determined. The Higgs width is $\Gamma_H \sim G_F M_H^3$, where G_F is the Fermi constant which describes an apparent four-fermion coupling $G_F = g^2/M_W^2$. It is thus expected that the Higgs width Γ_H becomes large for masses M_H above $150 \text{ GeV}/c^2$. For example: at $M_H = 100 \text{ GeV}$ one obtains only $\Gamma_H \approx 2.5 \text{ MeV}$, while at $M_H = 200 \text{ GeV}$ this width becomes already 1.5 GeV large. For $M_H = 400 \text{ GeV}$ its value increases further to 30 GeV and would eventually reach $\sim 450 \text{ GeV}$ at $M_H = 900 \text{ GeV}$, which can be more likely considered as an “excess” than as a “peak”.

Nevertheless some constraints on the Higgs mass can be put using self-consistency arguments [4]. Since the running coupling λ rises as the energy grows, the theory would eventually become non perturbative. The requirement that in the Standard Model this does not occur at a scale lower than Λ defines an upper bound on M_H . In other words, assuming that the Standard Model is an effective theory up to an energy scale Λ (*i.e.* that the Higgs boson is the only missing particle), leads to an upper constraint on its mass. On the other hand, the requirement that the electroweak minimum is an absolute minimum up the scale Λ yields a “vacuum stability” condition which limits M_H from below.

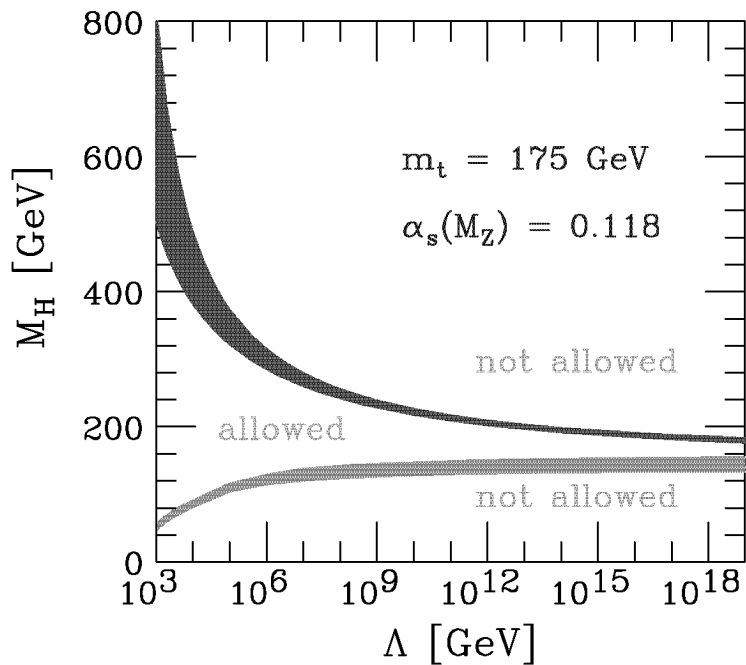


Figure 2.1: Assuming the validity of the Standard Model up to a scale Λ , the allowed Higgs mass range reduces itself to the area between the two curves. [33]

The so-obtained theoretical bounds depend strongly on the top quark mass and are shown in figure 2.1. According to the current values for the top mass ($m_t = 175$ GeV) and of the strong coupling constant ($\alpha_s(M_Z) = 0.118$), if one requires the Standard Model to be valid up to the Plank scale Λ_{Pl} (*i.e.* $\approx 10^{19}$ GeV), then the above arguments tends to favor the $130 \text{ GeV} < M_H < 190 \text{ GeV}$ mass range [5]. On the other hand, if the Higgs mass lies outside this range, then some new physics should appear well below Λ_{Pl} . In fact, figure 2.1 also shows that, in case an Higgs would be found for example at 800 GeV, there should be something new already in the TeV range.

In the Standard Model, as the Higgs mass grows, the couplings to itself and to longitudinally polarized gauge bosons (W_L , Z_L) increase. This makes the resonance wider and the interaction stronger, eventually leading to a violation of the unitarity bound⁴ at the order of 1 TeV. In case the Higgs scalar exists, it should therefore appear within this mass scale.

Not only does the coupling of the Higgs field to W_L and Z_L grow, but also the couplings between W_L and Z_L themselves increase with growing diboson mass. The presence of a regularizing resonance is not the only possible scenario. In case that no fundamental scalar particle exists, some new phenomena must set in to account for the breaking of electroweak symmetry, to control the vector boson coupling and to generate fermion masses. Consequently, there is either a Higgs boson with a mass less than about 1 TeV, or the dynamics of WW and ZZ interactions with center-of-mass energies of the order of 1 TeV will reveal new structure. The so-called “no lose” strategy is based on these simple arguments: the ability for the LHC collider to reach TeV energy scales and to detect diboson events should ensure that the experiments will be able to provide valuable informations on the nature of the electroweak symmetry breaking.

Despite the great success of the Standard Model in describing the strong, weak and electromagnetic interactions, there is also an increasing attention among the physics community about its limitations. For example, it makes *ad hoc* assumptions about the shape of the potential, responsible for the EWSB mechanism, and provides no explanation for the values of the parameters. The instability of the Higgs mass under radiative corrections at larger energy scales (for example the Plank scale Λ_{Pl}) leads to the *naturalness/hierarchy* problem. Questions arise also about the origin of the fermion generations or of the CKM matrix.

The model, in its simplest version, has 19 parameters:

- the 3 coupling constants,
- the masses of the 3 leptons and 6 quarks ($m_\nu = 0$),
- the Z boson mass, which sets the scale of the weak interaction,

⁴Set by the condition that the scattered intensity cannot exceed the incident intensity in any partial wave.

2.2 Standard Model 2001: status report

- 3 mixing angles and 1 phase, describing the rotation from the weak to the mass eigenstates of the charge $-1/3$ quarks (CKM matrix).

These parameters have been constrained and are known with varying errors. For example $\Delta m_Z/m_Z \approx 0.002\%$, while the strong coupling constant $\alpha_s(m_Z)$ has been measured within a $\pm 5\%$ error. Two parameters are still missing:

- a CP -violating parameter associated with the strong interaction (which should be very small),
- a parameter related to the spontaneous symmetry breaking mechanism.

This last unknown quantity can be associated with the mass of the Higgs boson. It has to be pointed out that, beyond the Standard Model, there are other possibilities, which can explain the current data and in the same time solve part of the problems quoted before. However, in addition, they also lead to a proliferation of particles and parameters. For example, the introduction of supersymmetry cures the mass divergence problems of the SM Higgs⁵, but at the expense of introducing supersymmetric partners to every known boson and fermion and at least an additional Higgs super-multiplet [34]. This work is limited to the study of the Standard Model, leaving to the interested reader the study of other possible scenario (see for example [35]).

2.2 Standard Model 2001: status report

The precise measurements performed at LEP [6], SLC [36], Tevatron [37] and elsewhere can be used to check the validity of the Standard Model and, in its framework, to provide information about the Higgs mass and the missing fundamental parameters.

The Higgs boson mass M_H is sensitive to the electroweak parameters through loop corrections. In particular, measurements of the weak boson masses m_Z and m_W , the weak mixing angle $\sin^2 \Theta_W$, together with the obtained values for the top quark mass m_t and the strong coupling α_s , can be used to constrain the Higgs mass.

The most stringent experimental limit on M_H in the Standard Model has been set by LEP. During the 12 years of LEP operation (1989-2000), the 4 experimental collaborations [7] delivered an impressive amount of precise measurements for various Standard Model parameters. These values have been completed by measurements from the SLD experiment at SLC, from the $p\bar{p}$ experiments CDF and D0 at Tevatron and from the NuTeV neutrino experiment. Figure 2.2 shows the

⁵All the diagrams, in which the Higgs boson dissociates into a virtual fermion-antifermion pair, contribute to the Higgs boson's own mass. There are infinitely many such diagrams, involving more than one such fermion loops, and by trying to calculate the Higgs mass corrections, one gets a value which diverges to infinity.

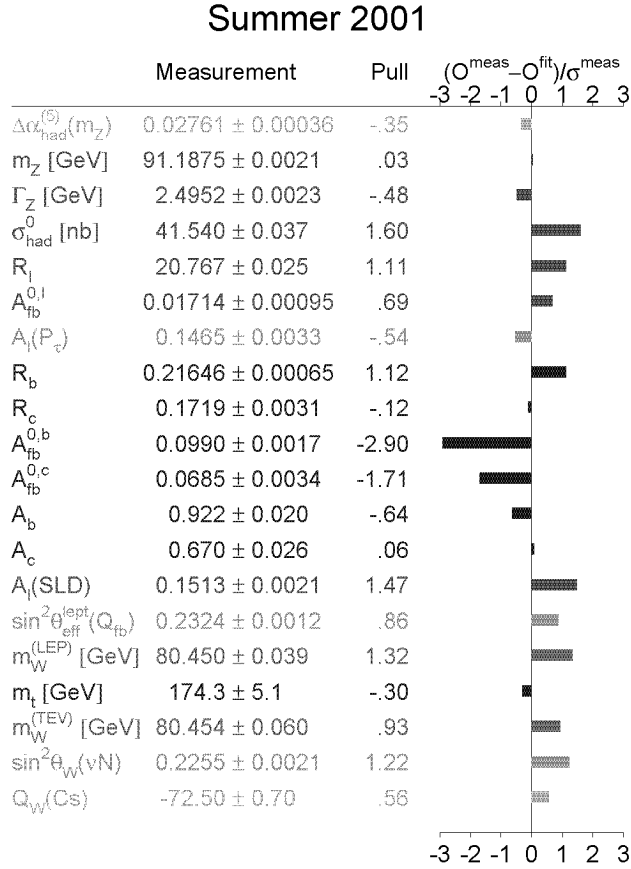


Figure 2.2: The up-to-date status of the Standard Model precision measurements [9].

present combined values for a selected sample of electroweak observables. Most of the data are consistent with the SM predictions. In general, the relative errors are smaller than 1%. For example, the W boson mass has been constrained with an error of about 40 MeV, *i.e.* with a precision of better than 1 per mil, while the uncertainty for the Z mass is even smaller (about 2 MeV)!

A large number of improvements were made to the LEP collider during its life, increasing the Higgs discovery potential and extending the sensitivity of the direct search of a Standard Model Higgs up to approximately 115 GeV. The combination of the last results of the four LEP experiments showed signs for a possible 115.6 GeV Higgs boson [8]. An observed 2.1σ deviation could be related to the production of a Higgs boson with Standard Model like couplings. However, the statistical significance of these results is not sufficient to claim an

“evidence”. Nevertheless, such an excess of candidates can be an indication for future experiments, like Tevatron [37] and LHC, which could have the last word on this possible signal.

The present lower mass limit for the Higgs direct search delivered by the four LEP collaborations corresponds to $M_H > 114.1$ GeV [8] at the 95% CL. This limit is indicated in figure 2.3 as the excluded area on the left.

Moreover, through the sensitivity of electro-weak data to radiative corrections and assuming the overall validity of the Standard Model, a global fit to the present values ($p\bar{p}$ colliders and LEP II) leads to the mass expectation $M_H = 88^{+53}_{-33}$ GeV [9], with an upper limit of 196 GeV at the 95% CL. This is shown in figure 2.3, where the minimal χ^2 of the fit is given as a function of M_H .

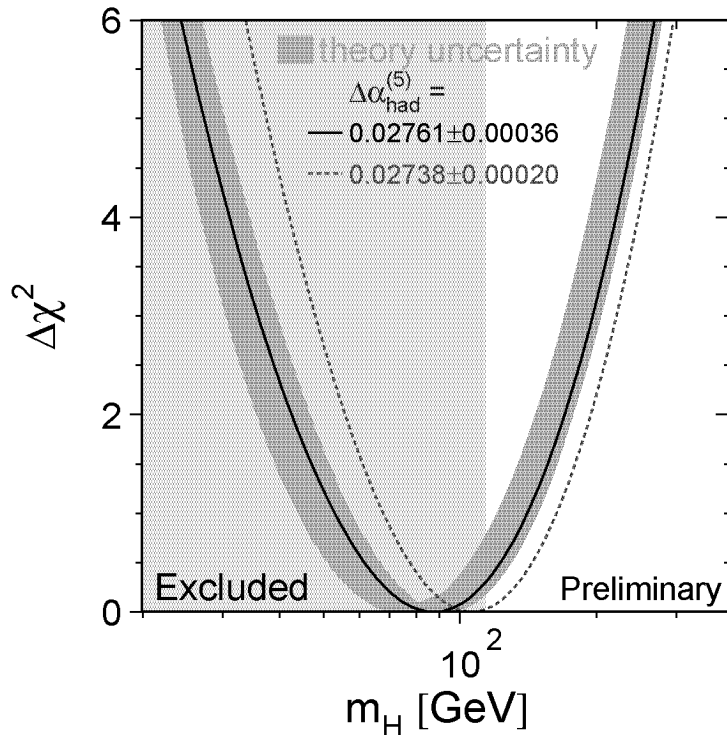


Figure 2.3: Dependence of the χ^2 of the data fit as a function of the unknown Higgs boson mass M_H . The curve results from a combination of all data ($p\bar{p}$ colliders and LEP II) [9]. The band represents an estimate of the theoretical error due to missing higher order corrections. The vertical band shows the 95% CL exclusion limit on M_H from the direct search. The dashed curve is the result obtained using the evaluation of $\Delta\alpha_{\text{had}}^{(5)}(m_Z^2)$ from reference [38].

2.3 The Challenge: finding the Standard Model Higgs Boson at LHC

2.3.1 The “well known”

As pointed out in section 2.1, one of the main goal of the LHC Project is to assure a Higgs discovery for the entire mass range allowed in the Standard Model. The exploitable domain ranges from the highest mass excluded by the LEP experiment (~ 115 GeV) up to about 1 TeV.

Within the Standard Model framework there are several Higgs production mechanisms which lead to detectable cross sections at LHC. Each of them exploits the preference of the SM Higgs to couple to heavy particles: *i.e.* either massive vector bosons (W and Z) or massive quarks (especially t -quarks).

The main production mechanisms at a 14 TeV pp -collider are then (see the feynmann diagrams (a)-(d) in figure 2.4):

- (a) The so called gluon–gluon fusion $gg \rightarrow H$, where the coupling is mediated primarily by heavy top quark loops [39]. In this case the cross section is directly sensitive to the $t\bar{t}H$ coupling.
- (b) The Weak Boson Fusion mode $qq \rightarrow qqH$, where two W/Z bosons are emitted by incoming quarks and interact to create a Higgs [40]. This production mode provides informations about the Higgs coupling to weak vector bosons.
- (c) The associated production of a Higgs together with a W or a Z boson [41]. Also in this case, the cross section depends on the WWH/ZZH couplings.
- (d) The associated production with $t\bar{t}$ pairs [42]. As for the gluon–gluon fusion process, this mode is sensitive to the Higgs coupling to fermions.

Figure 2.5 shows the cross section related to these production modes as a function of the Higgs mass. The gluon–gluon fusion $gg \rightarrow H$ dominates over the entire Higgs mass range, followed by the WW, ZZ fusion, which becomes comparable in magnitude to the gluon–gluon fusion for very large Higgs masses (~ 1 TeV). The cross sections of the associated Higgs productions $q\bar{q}' \rightarrow WH, ZH$ and $gg, q\bar{q} \rightarrow t\bar{t}H$ are much smaller: about a factor of 10 for $M_H \approx 100$ GeV and decreasing to $\sim 1/1000$ for $M_H \approx 1$ TeV.

2.3 The Challenge: finding the Standard Model Higgs Boson at LHC

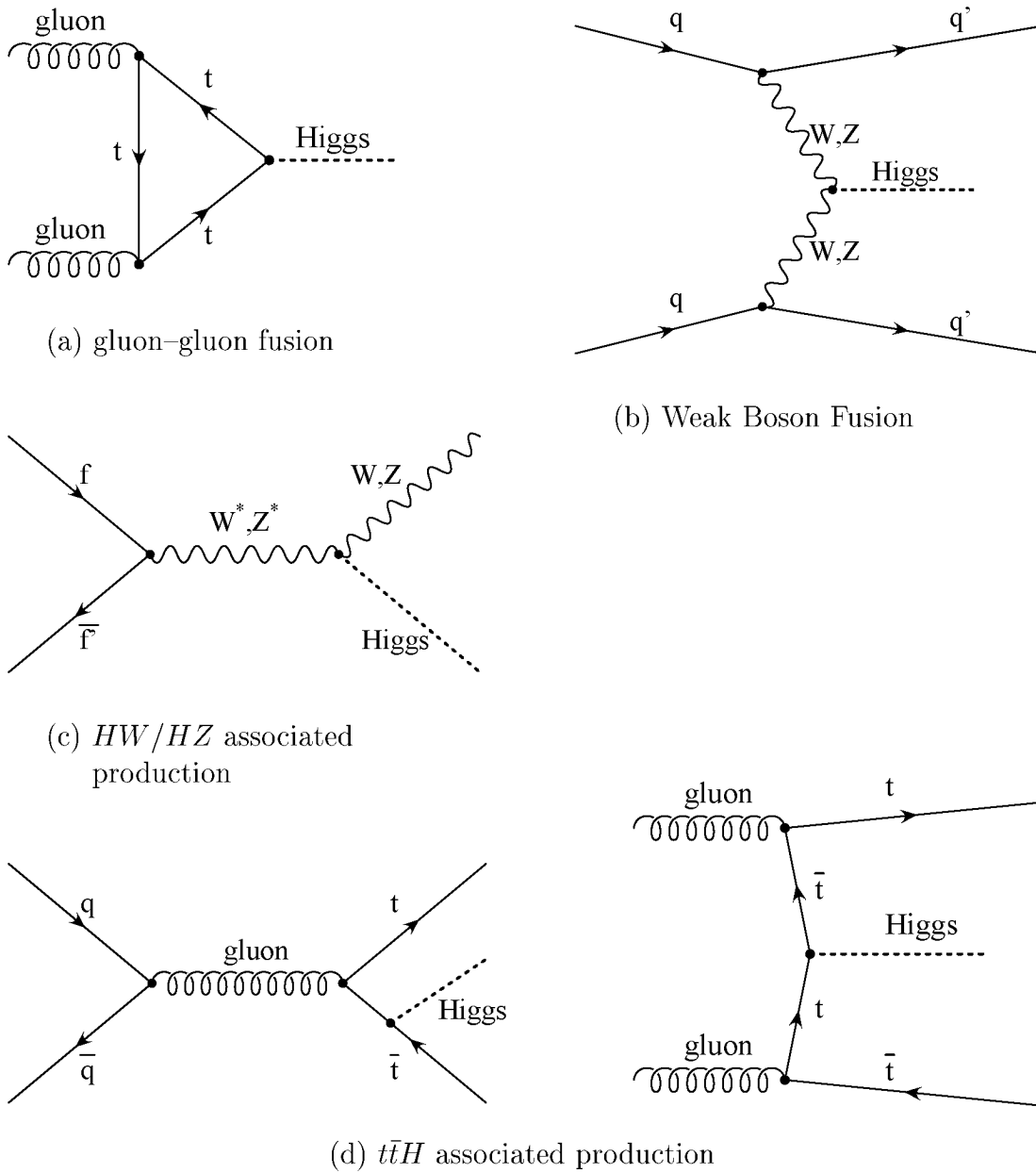


Figure 2.4: Main Higgs production modes at LHC.

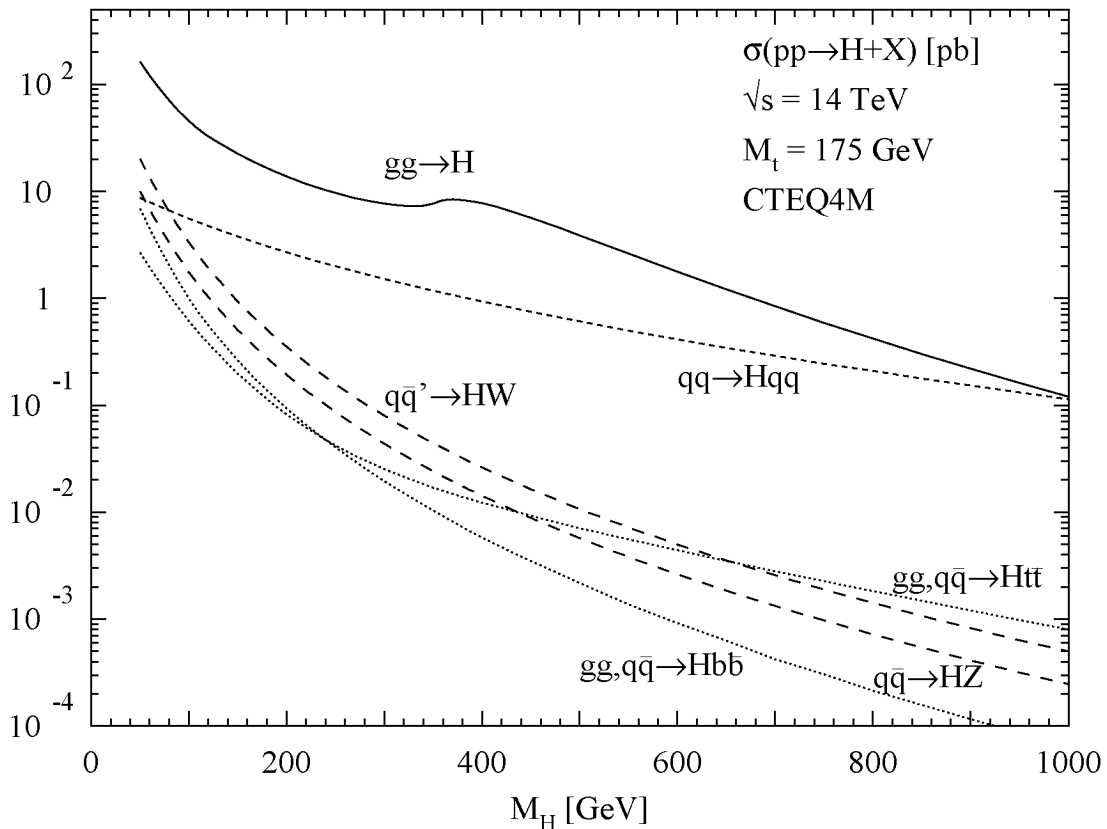


Figure 2.5: Standard Model expectations for the Higgs production Cross Section at LHC [43, 44].

The production cross section has then to be combined with the Branching Ratios (BR) for the considered decay modes. For the mass range consistent with a Standard Model Higgs, the various BR 's to pairs of fermions and gauge bosons are shown in figure 2.6. Well below the WW threshold, the various BR 's to pairs of heavy fermions dominate, while the decays to gluon or photon pairs have smaller probabilities. Above about 100 GeV, the Higgs Branching Ratios to pairs of massive vector bosons ZZ^* , WW^* increase rapidly, till they become the most relevant ones at ~ 150 GeV. Above the $t\bar{t}$ threshold, the WW and ZZ decay modes are then slightly reduced as the $H \rightarrow t\bar{t}$ decay starts to contribute.

One has then to add the Branching Ratios which lead to detectable products. The vector bosons decay either hadronically: $Z \rightarrow q_i \bar{q}_i$, $W \rightarrow q_i q_j$, with branching ratios of about 70%, or leptonically: $Z \rightarrow l^+ l^-$, $Z \rightarrow \nu \bar{\nu}$, $W \rightarrow l^\pm \nu$, with branching ratios, respectively, of about 10%, 20% and 30% (τ leptons included). However, at the LHC, the presence of a huge irreducible hadronic background turns out to be a handicap to the decays to quark pairs. In the middle-low Higgs mass range (*i.e.* for $\lesssim 300$ GeV), this favors the full leptonic channels $ZZ^{(*)} \rightarrow 4l^\pm$ and $WW^{(*)} \rightarrow l^+ \nu l^- \bar{\nu}$ ($l \equiv e, \mu$), despite the smaller branching

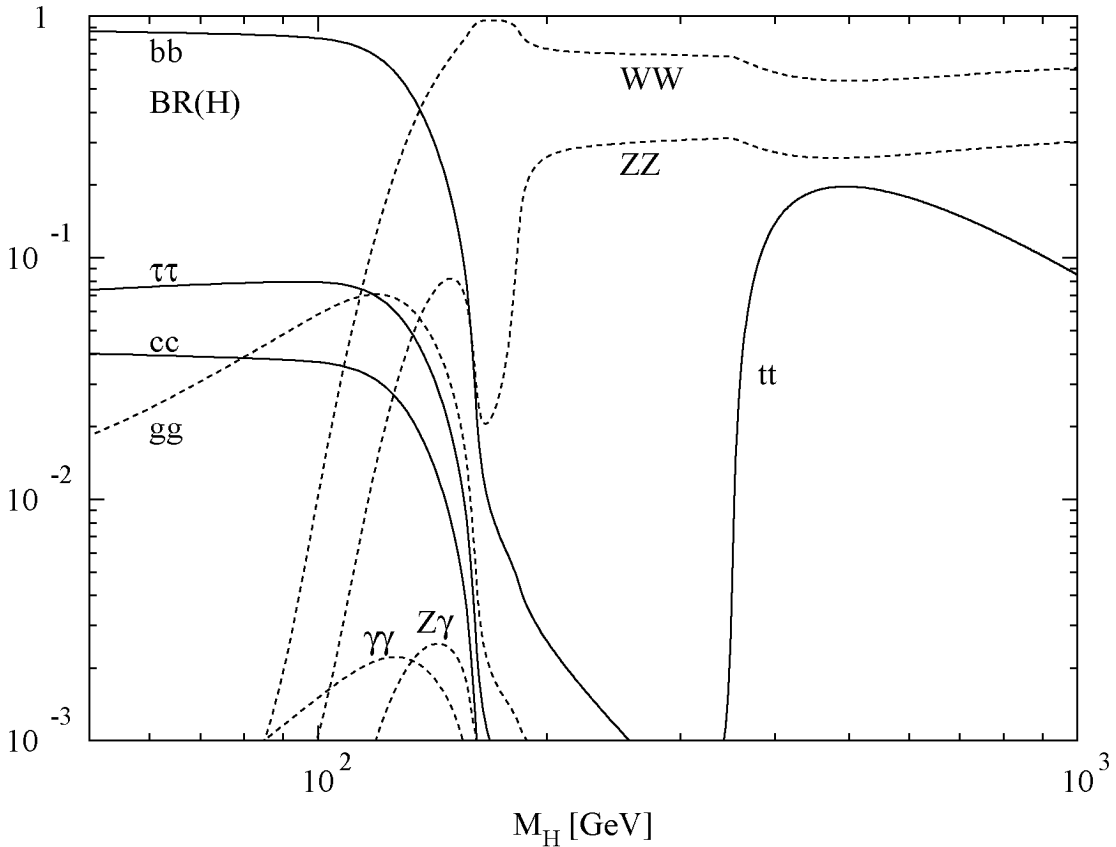


Figure 2.6: Higgs Branching Ratios as a function of the mass of a Standard Model Higgs [45].

ratios of, respectively, $< 0.5\%$ and $< 5\%$. These channels are characterized by isolated, high- p_t leptons and, for the WW , by large missing energy. For larger Higgs masses, also the Z decays to neutrinos and the W, Z decays to quark pairs are considered.

For the most promising Higgs search modes, the detectable $\sigma \times BR$ are shown in figure 2.7. The two decays $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^{(*)} \rightarrow 4l^\pm$ can be seen as narrow mass peaks. On the other hand, these signatures have a much smaller detectable cross section compared to the $H \rightarrow WW^{(*)} \rightarrow l^+\nu l^-\bar{\nu}$, which however can not be fully reconstructed because of the presence of two neutrinos.

By performing a full visibility study, that is by taking into account the different specific backgrounds, one obtains then the significance values for each of the above modes. The so-obtained "discovery channels" can be seen in figure 2.8, where the required luminosity for a 5σ significance is shown as a function of M_H . In the $80 < M_H < 140$ GeV mass domain, the $H \rightarrow \gamma\gamma$ decay mode provides the best significance. Despite its low branching ratio, the $\gamma\gamma$ channel profits from the small width of a low mass Higgs. The possibility to precisely measure the

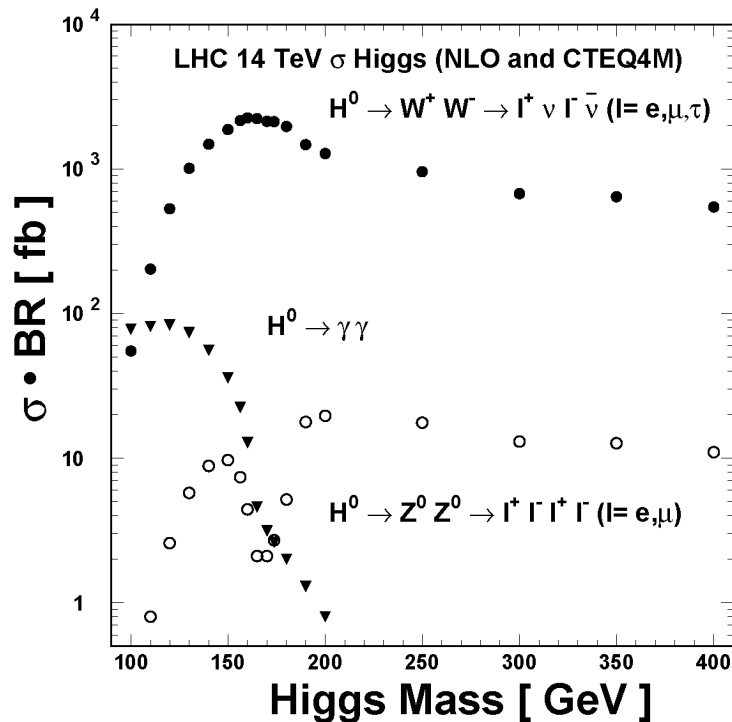


Figure 2.7: Expected $\sigma \times BR$ for the detectable SM Higgs decay modes, which offer the best significance [46].

photon energy, associated to a good vertex recognition, will allow to reconstruct a Higgs peak, giving to the $\gamma\gamma$ mode a better discovery potential than all other decay channels in this mass range. On the contrary, the $b\bar{b}$, $c\bar{c}$ and gg decay modes are believed to be hidden in the huge hadronic background and despite the much larger BR 's (see figure 2.6) do not offer an useful signal (the $H \rightarrow b\bar{b}$ decay channel, associated with the $t\bar{t}H$ production mode, could give eventually significance for $M_H < 115$ GeV).

For a Higgs mass between 130 and 180 GeV, the $H \rightarrow WW^{(*)} \rightarrow l^+\nu l^-\bar{\nu}$ has a substantial cross section. Its large event rate, together with the exploitation of the WW spin correlations [47], compensate the absence of a narrow mass peak. In case of an Higgs with mass $2 \times m_Z \leq M_H \leq 400$ GeV, the $H \rightarrow ZZ \rightarrow 4l^\pm$ decay mode profits of the presence of four isolated high- p_t leptons. By means of the so-obtained mass peak, this channel should provide the easiest discovery signature in this mass range.

In addition, if the Standard Model Higgs mass lies above 400 GeV, also the decay channel $H \rightarrow ZZ \rightarrow l^+l^-\nu\bar{\nu}$, involving an invisible Z decay, could give a good significance.

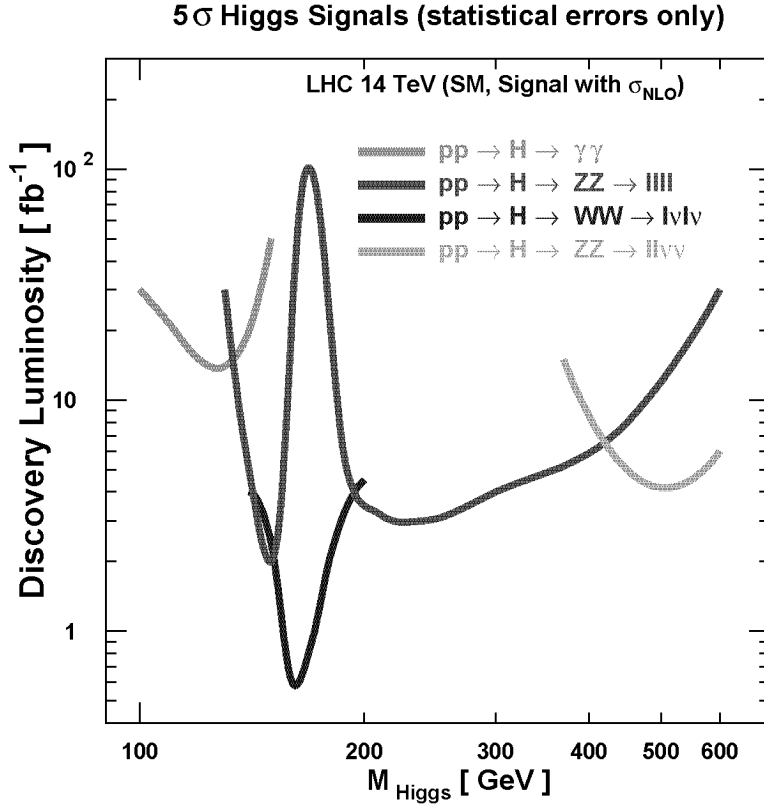


Figure 2.8: Required luminosity for a 5σ discovery of a SM Higgs through the most promising channels at LHC [46].

2.3.2 Additional signatures: towards a better determination of the Standard Model Higgs couplings

Additional informations on the Higgs parameters (mass, width, production rates, branching ratios) could be obtained by means of other production and/or decay modes. Such measurements should give further insights into the electroweak symmetry breaking mechanism and into the way the Higgs couples to fermions and bosons.

In fact, as was already said before, the Weak Boson Fusion mode $qq \rightarrow qqH$, which is directly sensitive to the Higgs coupling to vector bosons, has also an important production cross section (see figure 2.5), from the lowest Higgs masses up to the highest domain covered by the LHC experiment. Moreover, the related Feynmann diagram (figure 2.4b) put in evidence the presence of the two scattered quarks, which give rise to two very energetic forward jets. By tagging these two jets and reconstructing the Higgs decay products, one obtains a very distinctive signature, which should be exploitable to eliminate most of the back-

ground. These channels have already been proposed for the higher mass domain ($M_H > 300$ GeV), where the existing well-established studies [21, 22, 23] concluded that good signatures are the ones where both vector bosons decay either to charged leptons (only electrons and muons are considered) or the mixed ones with one W or Z decaying to leptons and the other one to hadrons. Recently, new studies [25, 26] proposed to use this production mode with the fully leptonic decay $H \rightarrow WW \rightarrow \ell^+ \nu \ell^- \bar{\nu}$ also in the mass range $110 < M_H < 200$ GeV.

Finally, figure 2.5 shows that, in the lower Higgs mass domain, the associated productions of an Higgs boson together with a vector boson $q\bar{q}' \rightarrow WH, ZH$ (figure 2.4c) or with a pair of top quarks $gg, q\bar{q} \rightarrow t\bar{t}H$ (figure 2.4d) have a competitive cross section. These signatures could play an important role in presence of an Higgs with mass below 120 GeV, while at larger M_H , as said before, the cross sections drop dramatically. For the $WH, ZH, t\bar{t}H$ ($H \rightarrow \gamma\gamma$) processes, the additional presence of the decay products of the associated W/Z or $t\bar{t}$ pair could enhance the visibility.

To summarize, for the detection of a Standard Model Higgs, the established signatures which should give clean signals at the LHC are the following:

$$\begin{array}{ll}
 H \rightarrow \gamma\gamma & \text{for } 80 \text{ GeV} \lesssim M_H \lesssim 140 \text{ GeV} \\
 H \rightarrow ZZ^{(*)} \rightarrow 4\ell^\pm & \text{for } 130 \text{ GeV} \lesssim M_H \lesssim 700 \text{ GeV} \\
 H \rightarrow WW^{(*)} \rightarrow \ell^+ \nu \ell^- \bar{\nu} & \text{for } 140 \text{ GeV} \lesssim M_H \lesssim 200 \text{ GeV} \\
 H \rightarrow ZZ \rightarrow \ell^+ \ell^- \nu \bar{\nu} & \text{for } 400 \text{ GeV} \lesssim M_H \lesssim 1 \text{ TeV}
 \end{array}$$

In addition, in the low Higgs mass domain, the $H \rightarrow \gamma\gamma$ channel could be integrated by the associated productions:

$$\left. \begin{array}{l}
 q\bar{q}' \rightarrow WH, ZH \\
 gg, q\bar{q} \rightarrow t\bar{t}H
 \end{array} \right\} \text{ with } H \rightarrow \gamma\gamma \text{ for } 80 \text{ GeV} \lesssim M_H \lesssim 120 \text{ GeV}$$

The $H \rightarrow b\bar{b}$ could give significance up to 115 GeV:

$$gg, q\bar{q} \rightarrow t\bar{t}H \quad \text{with } H \rightarrow b\bar{b} \quad \text{for } 80 \text{ GeV} \lesssim M_H \lesssim 115 \text{ GeV}$$

Furthermore, for a high-mass Higgs, well visible signatures should also be provided by the fusion of longitudinally polarized vector bosons:

$$qq \rightarrow qqH \quad \text{with } \begin{cases} H \rightarrow WW \rightarrow \ell\nu jj & \text{for } 300 \text{ GeV} \lesssim M_H \lesssim 1 \text{ TeV} \\ H \rightarrow ZZ \rightarrow \ell^+ \ell^- jj & \text{for } 700 \text{ GeV} \lesssim M_H \lesssim 1 \text{ TeV} \end{cases}$$

Finally, this same production mode should also give some sensitivity for an intermediate mass Higgs:

$$qq \rightarrow qqH \quad \text{with } H \rightarrow WW \rightarrow \ell^+ \nu \ell^- \bar{\nu} \quad \text{for } 110 \text{ GeV} \lesssim M_H \lesssim 300 \text{ GeV}$$

This last scenario will be the topic of chapter 7.

Chapter 3

Experimental aspects of a 14 TeV pp -collider experiment

The proposed Large Hadron Collider (LHC) [11] at CERN will provide proton–proton collisions at a center-of-mass energy of $\sqrt{s} = 14$ TeV with a design luminosity of $\mathcal{L} = 10^{34}$ cm⁻² s⁻¹.

For a multi-purpose LHC detector, the ability to explore the entire SM Higgs mass range plays a central role among the design criteria. Such a detector should have enough flexibility to take the large variety of expected Higgs signatures into account. In addition, the detector should also be sensitive to the other possible signatures as predicted by physics processes beyond the Standard Model.

High \sqrt{s} and high luminosity are the key to reach higher mass domains and to allow the discovery of processes with very rare signatures. However, these requirements lead also to difficult experimental conditions. At LHC, the protons will be set into ~ 3000 bunches, producing a pp crossing every 25 ns at each of the 4 interaction points. Moreover, the very large non-diffractive inelastic cross-section of about 70 mb results in an average of ~ 20 minimum bias interactions per beam crossing interval. Combined with the 40 MHz, this corresponds to a total of $\sim 10^9$ events per second, while the important physics signals have much smaller rates. For example, about one $H \rightarrow \gamma\gamma$ event per day is expected for $M_H = 100$ GeV. In other words, the interesting signatures have to be found among a huge background!

To achieve this goal, the online event selection must first reduce the 40 MHz to about 100 events/s (the maximal rate which can be accepted for the data storage and subsequent analysis). One has also to take into account that at design luminosity every signal will be superimposed with the ~ 20 minimum bias events produced in parallel, resulting in a large amount of “unwanted” particles within the interesting kinematical regions. These huge event sizes (≈ 1 MByte), combined with the very selective acquisition, require enormous efforts in the domains of data storage, data handling and in the development of algorithms to

select the interesting physics events. High particle densities will make the pattern recognition difficult and the application of selection algorithms based on track reconstruction will not be trivial.

The high particle fluxes coming from the interaction region will also lead to high radiation levels in the experimental area. Therefore, radiation-hard detectors and front-end electronics are required.

All the above problems will not be directly treated in this work, where the availability of a detector according to the design performances is assumed. However the reader should keep in mind the huge effort required to make interesting signals measurable at LHC.

Signatures for new physics can show up in a number of final states of isolated leptons, jets and missing energy. For a multi-purpose detector the following requirements are mandatory: good particle identification, good energy, momentum and angle resolutions for charged leptons, jets, photons and missing transverse energy and this over a large solid angle. All these requirements are also valid to fully exploit the production of weak bosons.

In this chapter, I first introduce the LHC collider and the 4 foreseen detectors, with a more detailed discussion of CMS. Then I discuss the simulation frames that will be used for the different studies reported in this work and describe the algorithms for the selection of isolated leptons and gammas, as well as for the reconstruction of jets.

3.1 The LHC project

3.1.1 The pp collider

The effective parton–parton center-of-mass energies available in the pp collisions will reach the TeV range, about 5 times more than LEP and ~ 7 times more than Tevatron. In order to maintain an equally effective physics program at a higher energy E , the luminosity of a collider should increase in proportion to E^2 . The reason lies in the $1/E$ decrease of the De Broglie wavelength associated to a particle, giving a $1/E^2$ reduction in cross section. The foreseen luminosity at LHC is $\mathcal{L} = 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$. This will be achieved by filling each of the two rings with 2835 bunches of 10^{11} particles each.

The LHC is designed to use the 27 km LEP tunnel and will be fed by existing CERN particle sources and pre-accelerators. Having a fixed tunnel radius, large magnetic bending fields are needed to achieve highest possible energies. Consequently, to run 7 TeV protons around the ring, the LHC dipoles must be able to produce fields of 8.36 T, over five times those used a few years ago at the SPS proton–antiproton collider. Superconductivity makes this possible. LHC magnets will be operated at 1.9 K above absolute zero. This low temperature puts new demands on cable quality and coil assembly.

LHC superconducting magnets will sit at atmospheric pressure in a 1.9 K bath of superfluid helium, which has very efficient heat transfer properties, allowing kilowatts of refrigeration to be transported over more than a kilometer with a temperature drop of less than 0.1 K.

In electron–positron colliders, particles lose energy through synchrotron radiation¹. This phenomenon limits the attainable energy. Given the larger proton mass, in the LHC these effects are less important and the energy radiated by a proton during the same time corresponds only to a tiny fraction of its energy. However the power emitted, about 3.7 kW, cannot be neglected as it has to be absorbed by the beam pipe at cryogenic temperature. A challenging task will be to manage a machine made of delicate superconducting magnets, operating at cryogenic temperatures and at the resulting large beam current ($I_b = 0.53$ A).

The LHC will consist of two "colliding" synchrotrons installed in the 27 km LEP tunnel. They will be filled with protons delivered from the SPS at 0.45 TeV. Two superconducting magnetic channels will then accelerate the protons to 7 TeV, after which the beams will counter-rotate for several hours with colliding points at the experiments.

The LHC is a versatile accelerator and, besides protons, can also collide beams of heavy ions (like lead) with a total collision energy per nucleon in excess of 5.5 TeV.

Two luminosity regimes are planned: an initial phase at "low" luminosity $\mathcal{L} = 10^{33}$ cm⁻² s⁻¹, where the experiments should provide evidence for the Higgs in certain mass ranges. In addition, experiments may access new physics, as well as act also as factories for QCD processes, heavy flavor and gauge bosons. During a second phase at the design luminosity $\mathcal{L} = 10^{34}$ cm⁻² s⁻¹, the full physics performances of LHC will be then exploited.

3.1.2 ATLAS, LHC-b and ALICE

The four foreseen detectors are ATLAS [48], CMS [14], LHC-b [49] and ALICE [50]. The first two will be so-called "general purpose" detectors, while the others two are more dedicated projects.

Figure 3.1 offers a general view of the existing and future LHC facilities, indicating the position on the LEP/LHC ring of the four detectors and the status of the civil engineering works.

The central feature of ATLAS (A Toroidal Lhc ApparatuS) is a large air-core toroidal magnet muon spectrometer, consisting of an 8-coils barrel toroid with 0.8 T average field and two end-cap toroids. Each coil measures 26 m in length and is 4.7 m wide, defining the overall dimensions of the ATLAS detector, with a total radius of about 11 m. Using the external muon system alone, this set-up will allow precise measurements of μ 's even at the highest luminosities.

¹With a rate proportional to $1/m^4$.

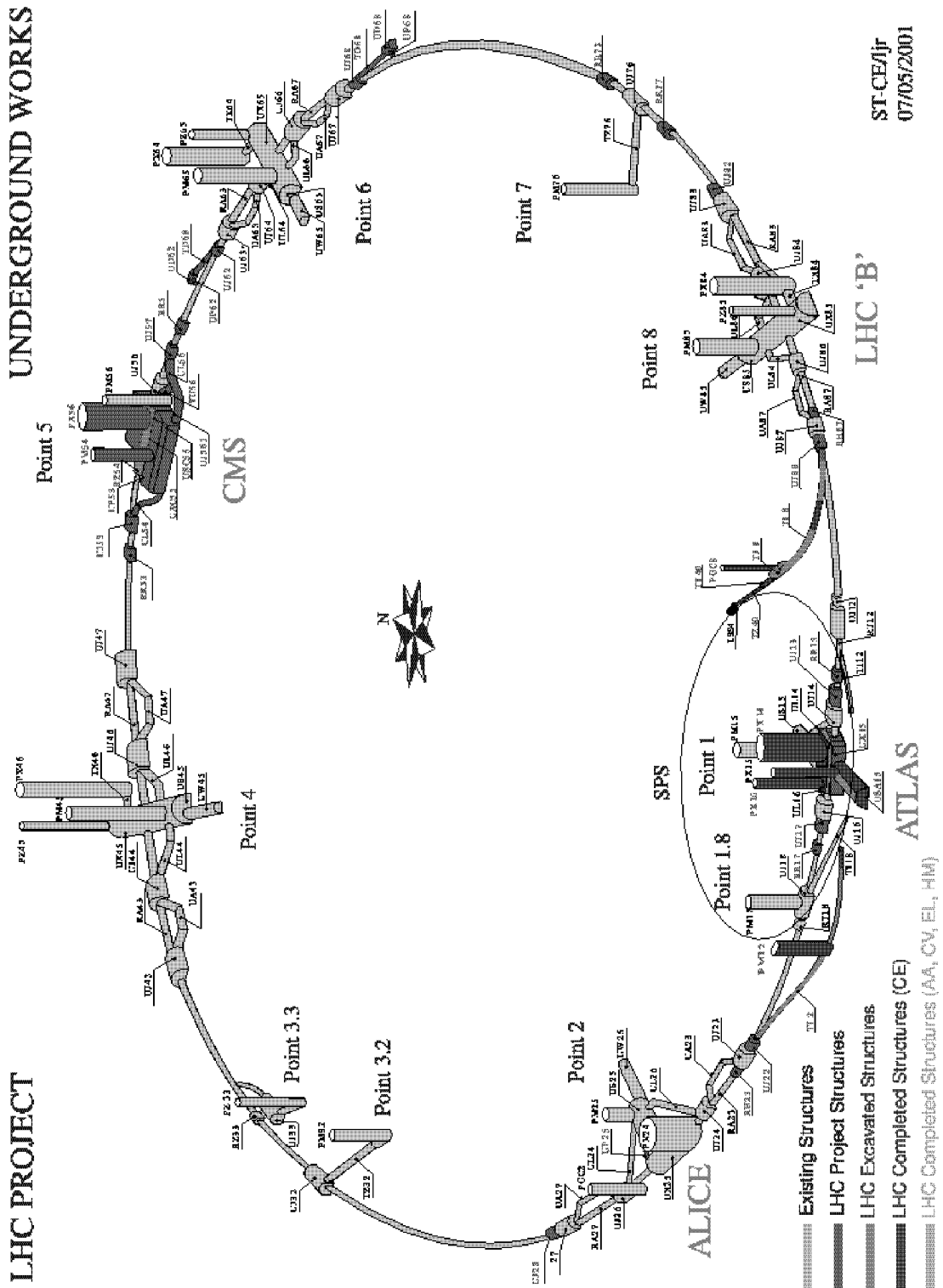


Figure 3.1: Layout of the LEP/LHC tunnel including present and future LHC infrastructures.

The measurements are done in the chambers behind the calorimeters that have absorbed all hadrons. The electromagnetic calorimeter is filled with liquid argon and provides excellent performances in terms of energy and position resolution. A 2 T solenoid is placed in front of the EM calorimeter to allow momentum measurements for the charged "stable" particles in the inner tracker.

LHC-b is a dedicated B-physics detector and will be a single arm spectrometer. B mesons are most likely to emerge from collisions close to the beam direction, so the detector is designed to catch low-angle particles, covering the angular range from ~ 10 mrad to ~ 400 mrad with respect to the beam line. The key elements of LHC-b will be a vertex detector, which will measure charged particle tracks, and Ring-Imaging Cherenkov (RICH) detectors, that will identify the different kinds of particles.

Part of the LHC physics program will be devoted to heavy ion collisions. The aim is to study the physics of strongly interacting matter at extreme energy densities, where the formation of a new phase of matter, the quark-gluon plasma, is expected. The foreseen luminosities range from 10^{27} cm^{-2} s^{-1} in ^{208}Pb – ^{208}Pb collisions to 3×10^{31} cm^{-2} s^{-1} in ^{16}O – ^{16}O collisions, with a nucleon-nucleon center-of-mass of about 5.5 TeV. The purpose of the ALICE (A Large Ion Collider Experiment) project is to detect a broad range of possible signatures of quark-gluon plasma formation and thus to carry out a comprehensive study of the hadrons, electrons, muons and photons produced in the collision of heavy nuclei. The central part, which covers $\pm 45^\circ$ over the full azimuth, is embedded in a large magnet with a weak solenoidal field. Outside of the inner tracking system, there are a cylindrical time projection chamber and a large area particle identification array of time-of-flight counters. In addition, there are two small-area single-arm detectors: an electromagnetic calorimeter and an array of RICH counters, optimized for high-momentum inclusive particle identification.

3.1.3 CMS: the "Compact" Muon Solenoid

CMS (Compact Muon Solenoid) [14] will be one of two general purpose detectors at the CERN Large Hadron Collider. The collaboration consists of ~ 2000 scientists from more than 150 institutions in 31 countries. Within the project, ETH Zürich plays an important role in the design and construction of the detector. Like ATLAS, CMS will be operational at the startup and therefore, in order to fully exploit the physics potential offered by LHC, it is also well adapted for studies at the initial lower luminosities.

Given the broad range of possible Higgs signatures, the largest flexibility is required. Other important goals are the detailed studies of beauty and top quark decays, as well as searches for super-symmetry and compositeness.

One of the design goals for the CMS detector is the ability to identify and measure isolated muons, photons and electrons with high precision.

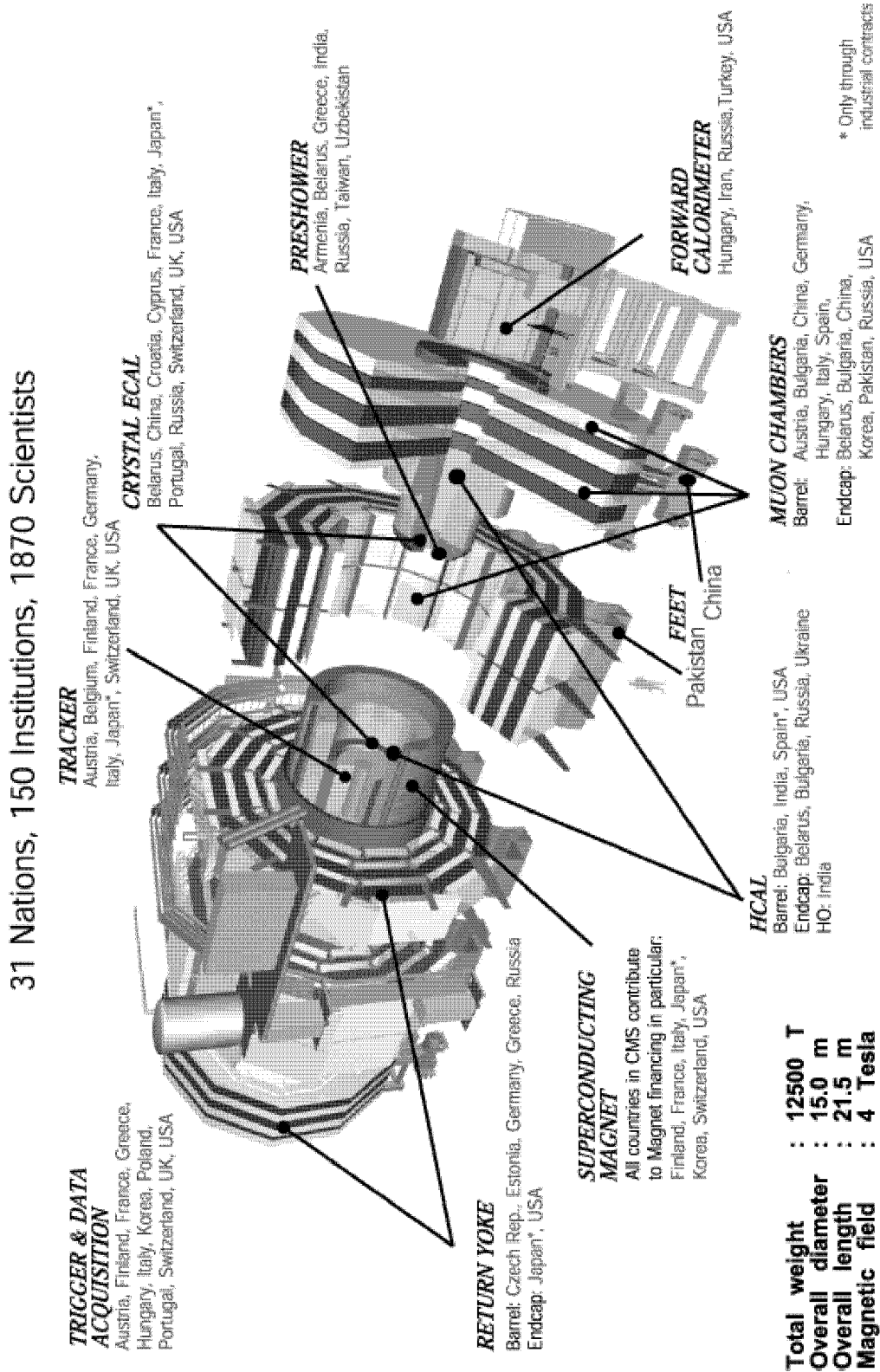


Figure 3.2: Perspective view of the CMS detector.

3.1 The LHC project

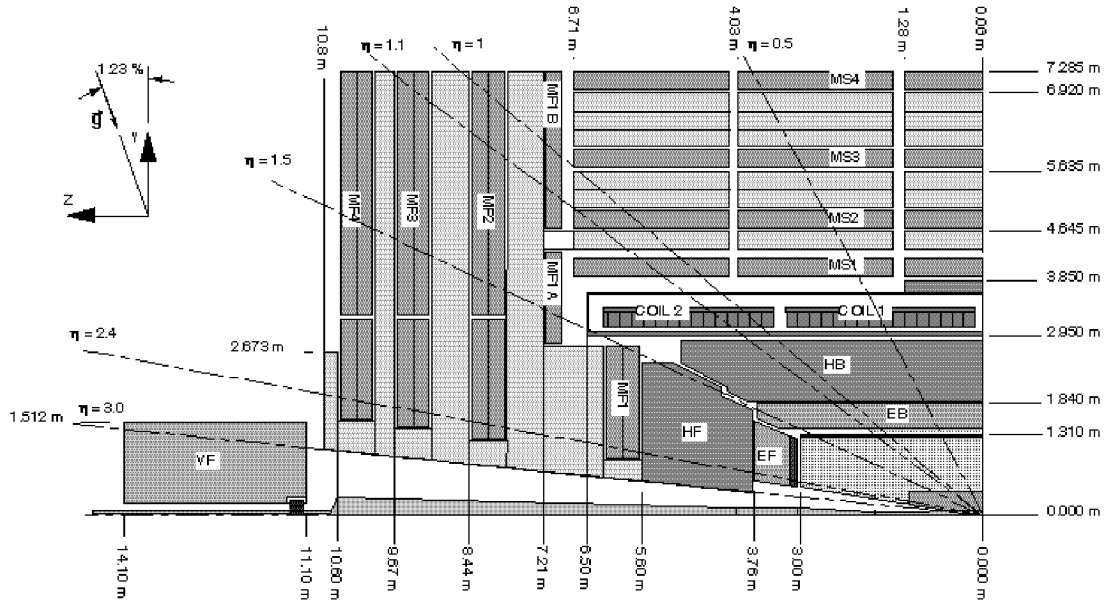


Figure 3.3: Longitudinal one quarter view of the CMS detector, with the pseudorapidity ranges η covered by the different calorimeters.

The energy resolution for the above particles should be 1% or better at 100 GeV. Also jets and missing energy should be measured with good accuracy, leading to an overall tight design within the rapidity region with $|y| \lesssim 3$ (for a definition of rapidity and pseudorapidity see equations 3.3-3.4 in section 3.2.2).

To achieve these objectives, the following sub-detectors are foreseen: a high performance and redundant muon system, a high resolution electromagnetic calorimeter, a high quality central tracking system and full calorimetric coverage.

The overall dimensions of the so-called "Compact Muon Solenoid" (a length of 21.6 m and a diameter of 14.6 m, resulting in a total weight of 14500 tones) are especially impressive, when one compares it with the dimensions of $\approx 10^{-18}$ m one wants to exploit. That means a difference of about 20 orders of magnitude, which is about the ratio between the size of an elephant and the size of the milky way!

Figure 3.2 offers a three dimensional view of CMS, with a description of its main parts, the involved countries and some overall dimensions. Figures 3.3 gives a longitudinal one quarter view, indicating the initials of the different detector parts and the covered pseudorapidity regions. Figure 3.4 shows the transverse view. Starting from the center, the tracking system, the electro-magnetic- (EB,EF) and hadron- (HB,HF) calorimeters are located inside the coil, while the layers of the muon system, alternated with the plates of the return yoke of the magnet, make up the external part of the detector.

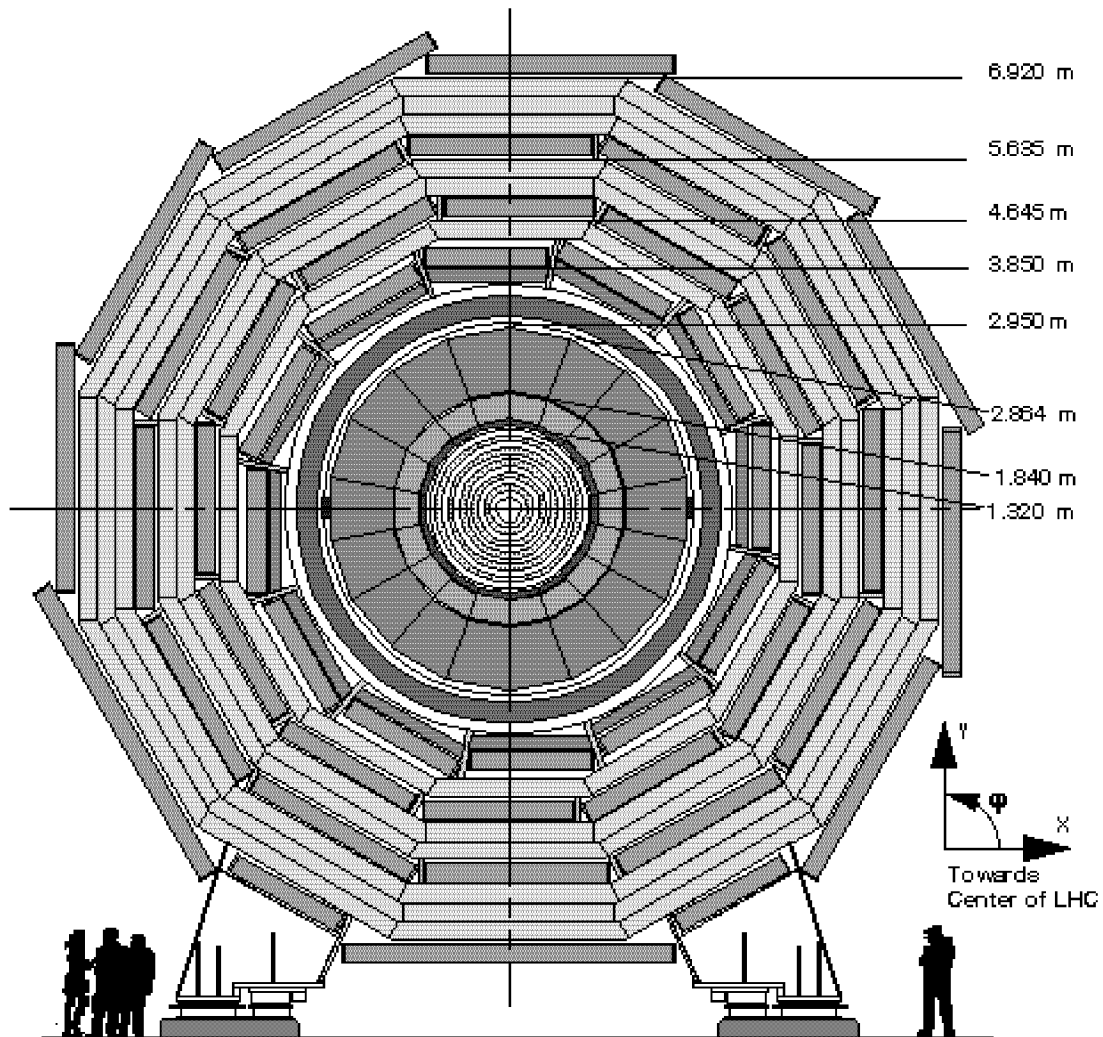


Figure 3.4: Transverse view of the CMS detector.

The superconducting magnet

The CMS detector is built around a large, 13 m long, 6 m diameter, high field superconducting solenoid [51] (COIL1 and COIL2 on figure 3.3), leading to a compact design for the muon spectrometer (hence the name Compact Muon Solenoid). The magnetic flux is returned through 1.8 m thick saturated iron yoke (1.8 T), which is split into five barrel rings and two end-cap disks housing the muon chambers. The magnet is in fact the main component of CMS in terms of size, weight and structural rigidity and plays a structural role by supporting all the other components of the detector. It generates an uniform magnetic field of 4 T and provides the bending power for precise tracking, assuring a good momentum measurement of charged particles up to rapidities of 2.5.

The muon system

The ability for the detector to measure high- p_t muons provides an important instrument to analyze diboson events. The ZZ mass resolution, for Z decays to μ 's, depends directly on the achievable energy and momentum resolution.

The objectives for the muon system [52] can be sub-divided into three basic tasks: (i) μ -identification, (ii) μ -momentum measurement and (iii) triggering of physics events. Given the enormous event rate in hadron collider experiments, the latter will be very challenging. The 4 T solenoidal field leads to an excellent momentum resolution and facilitates the momentum thresholds at the trigger level. The possibility to identify μ 's also inside jets allows an utilization of these particles as b-tagging ($b \rightarrow \mu + X$) as well as, with the help of the calorimeters, for measuring the energy flow around leptons inside jets (and thus permit to evaluate the efficiency of isolation cuts).

Starting from the primary vertex, muons are first measured in the inner tracker, then traverse the calorimeters, the coil and the return yokes. They are identified and measured in the four muon stations, inserted in the return yokes over most of the solid angle. In the barrel, each station (MS1 to MS4 in figure 3.4) consists of twelve planes of aluminum drift tubes (DTBX) arranged in twelve azimuthal sectors, such that there are no cracks pointing to the primary vertex. It is designed to have a bunch crossing identification capability and to give a muon vector in space with 100 μm precision in position and better than 1 mrad in direction. The end-cap muon system consists also of four muon station MF1 to MF4, where each one is made of sectors of cathode strip chambers (CSC) overlapping in azimuth to maintain full coverage. Two stations MF1A and MF1B are added to ensure that all muons traverse four stations at all rapidities, including the transition region between the barrel and the end-caps ($1 < |\eta| < 1.5$). Resistive Plate Chambers (RPC) are used in both the barrel and the end-cap. The last muon stations are placed after a total of more than 20 interaction lengths of absorber, so that only muons with $p_t \gtrsim 4$ GeV can reach them.

The DTBX and CSC detectors are used to obtain precise measurements of the position and thus of the muon momenta, whereas the RPC chambers, with their excellent time resolution of ~ 2 ns, will provide fast information for the level-1 trigger.

Efficient muon detection is guaranteed up to $|\eta| = 2.5$, for $p_t > 4$ GeV/c. Figure 3.5 shows, for the muon system alone, the expected transversal momentum resolution $\Delta p_t/p_t$ for 10, 100 and 1000 GeV muons as a function of their pseudorapidity. After matching with the central tracker, the global momentum resolution should improve to about 1-1.5% at 10 GeV, and 6-17% at 1 TeV.

CMS Muon System Alone

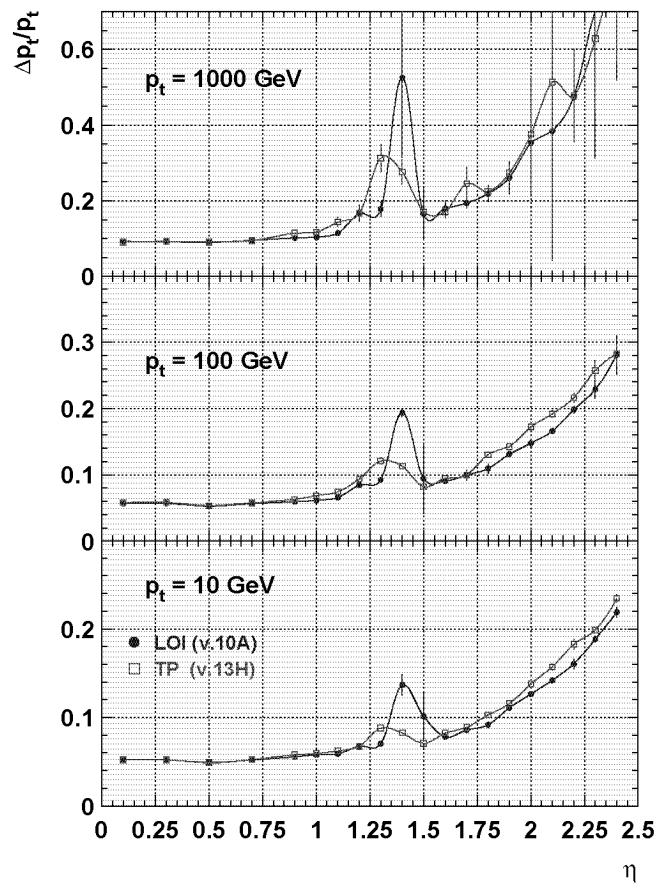


Figure 3.5: Muon system resolution vs pseudorapidity (given by the muon system alone). The two curves represent the values given in the CMS “Technical Proposal” and “Letter of Intent”.

The hadron calorimeter

The hadron calorimeter (HCAL) [53], combined with the electromagnetic one (ECAL), will measure jets, which are crucial for the discovery of many physics processes (including the detection of high mass vector boson pairs and of weak boson fusion processes).

The performances of the hadron calorimeter can be characterized by the jet–jet mass resolution, which depends also on fragmentation, energy pile-up and on the algorithms for the jet reconstruction. A calorimetry granularity of $\Delta\eta \times \Delta\phi \sim 0.09 \times 0.09$ has been chosen so that highly boosted di-jets from W, Z decays can still be distinguished. A good E_t^{miss} resolution and the detection of forward jets require an hermetic calorimeter coverage up to $|\eta| = 5$. This will improve the

detection capability of weak boson fusion processes, characterized by the presence of two forward jets, and for signatures with missing transverse energy.

The barrel and end-cap parts of the HCAL (HB and HF) are sampling calorimeters made of copper absorber plates, interleaved with 4 mm thick plastic scintillator tiles. The produced blue scintillation light is captured and shifted towards green in wavelength shifting fibers embedded in the scintillator and then transported through clear fibers to hybrid photodiodes that can operate in high axial magnetic fields. The active depth of the HCAL is everywhere larger than 9 interaction lengths λ .

The VF very forward calorimeter ($3 < |\eta| < 5$), which has to withstand high radiation doses, uses steel absorbers with quartz fibers, as active medium parallel to the beam. Charged particles produce Cherenkov light as they cross the quartz, some of which are captured in the fibers and channeled to photomultipliers.

The hadron calorimeter is insensitive to low energy neutrons, is extremely fast and yields narrow showers. The combined ECAL-HCAL single pion energy resolution is expected to be $\sigma/E \approx 110\%/\sqrt{E} \oplus 3.6\%$ (E in GeV). For a $W \rightarrow q_i q_j$ event, the obtained dijet mass resolution should be about 12 GeV with pile-up and 8 GeV without.

The electromagnetic calorimeter

The main function of the electromagnetic calorimeter (ECAL) [54] is to measure precisely the energy of electrons and photons. In addition, combined with the hadron calorimeter, it contributes to the measurement of jets. Both calorimeters will ensure a good detector hermeticity for the missing transverse energy measurements. This is particularly important for processes where the missing E_t is on the order of few tens of GeV, like for example $H \rightarrow \tau\tau$, $H \rightarrow WW$, $W \rightarrow l\nu$ decays.

The design choices for the ECAL are to a large extent determined by the requirements imposed by the $H \rightarrow \gamma\gamma$ channel. As is shown in the last chapter, the most promising channel for the search of a Standard Model Higgs in the 80-140 GeV mass range is the $\gamma\gamma$ decay mode. This signature profits of the very small natural width (few MeV) of a low mass Higgs. Therefore, the expected signal width is entirely dominated by the experimental $\gamma\gamma$ mass resolution. This width, together with the level of the irreducible $\gamma\gamma$ background, determine the signal significance and thus the detector capabilities to discover a Higgs in this mass range.

In order to achieve a high mass resolution, an excellent electromagnetic energy resolution σ/E is required. For the energy range of about 25 to 500 GeV, the energy resolution can be parameterized as (\oplus are quadratic sums):

$$\frac{\sigma}{E} = \frac{a}{\sqrt{E}} \oplus \frac{\sigma_n}{E} \oplus c, \quad (3.1)$$

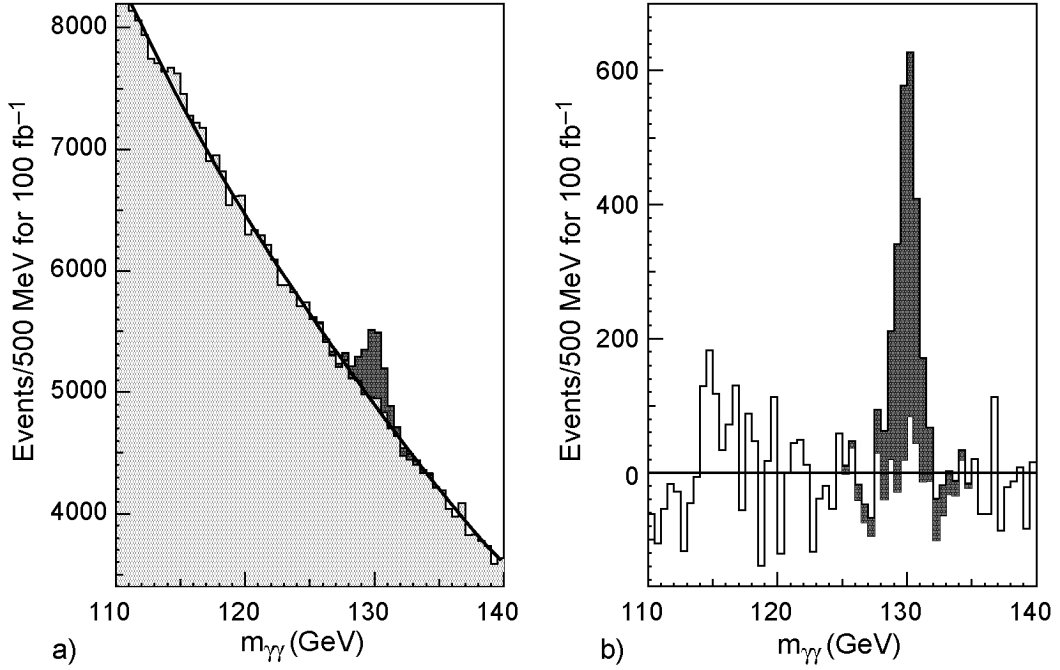


Figure 3.6: Simulated signal by a $H \rightarrow \gamma\gamma$ event, for $M_H = 130$ GeV (100 fb $^{-1}$ collected at high luminosity). (a) before, and (b) after background subtraction [54].

where a is the stochastic term, σ_n the noise and c the constant term. The expected values for the barrel (end-cap) are: $a = 2.7\%$ (5.7%), $\sigma_n = 155$ MeV (205 MeV) at low luminosity, $\sigma_n = 210$ MeV (245 MeV) at high luminosity and $c = 0.55\%$ (0.55%) [54].

In addition, the $\gamma\gamma$ mass resolution σ_M/M depends also on the two-photons angular separation θ :

$$\frac{\sigma_M}{M} = \frac{1}{2} \left(\frac{\sigma_{E_1}}{E_1} \oplus \frac{\sigma_{E_2}}{E_2} \oplus \frac{\sigma_\theta}{\tan \frac{\theta}{2}} \right). \quad (3.2)$$

At the low luminosity regime ($\mathcal{L} = 10^{33}$ cm $^{-2}$ s $^{-1}$) the event vertex will be indicated by other tracks in the event and the angular resolution will have only a minor effect on the $\gamma\gamma$ mass resolution. At high $\mathcal{L} = 10^{34}$ cm $^{-2}$ s $^{-1}$, about 20 minimum bias events are superimposed to the triggered $\gamma\gamma$ event and one expects the vertex definition to be much more difficult. Some precision can be obtained from the shower position measurements in the calorimeter. Depending on the luminosity, a mass resolution of 650 to 700 MeV is expected from detailed simulations of a 100 GeV $H \rightarrow \gamma\gamma$ [55]. Figure 3.6 shows the two-photon signal from a 130 GeV Higgs, after collecting 100 fb $^{-1}$ at high luminosity (before and after background subtraction).

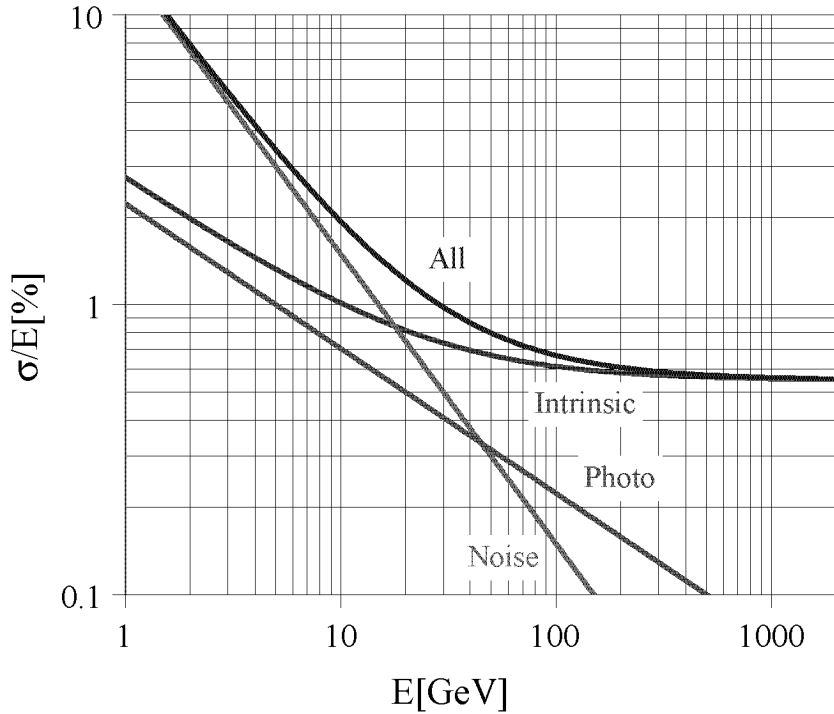


Figure 3.7: Different contributions to the e, γ energy resolution of the PbWO_4 calorimeter [54].

Figure 3.7 summarizes the different contributions expected for the energy resolution. The noise term contains the contributions from electronic noise (important at low energy) and pile-up energy (negligible at low luminosity). The “intrinsic” term includes the shower containment and a constant term. The curve labeled “Photo” gives the stochastic term of PbWO_4 .

The ECAL has a geometrical coverage up to $|\eta| = 3$ and consists of $\approx 76'000$ scintillating lead tungstate (PbWO_4) crystals with a thickness of more than 25 radiation length and a front face size of $2.2 \times 2.2 \text{ cm}^2$. PbWO_4 has been chosen because of its short radiation length, its small Molière radius and short light decay time, allowing for a compact ECAL design with narrow and fast-recovering showers. The scintillating light is detected using avalanche photodiodes (APD) in the barrel region and vacuum photo-triodes (VPT) in the end-cap region.

A scheme of the PbWO_4 crystals, together with the APD’s and the readout chain, is shown in figure 3.8.

For a detector facing an enormous hadronic background, an excellent electron identification is particularly important if one wants to exploit the various signatures associated to electrons. Problems could arise in case that charged pions will be misidentified as electrons within the ECAL. This problem is studied in chapter 5.

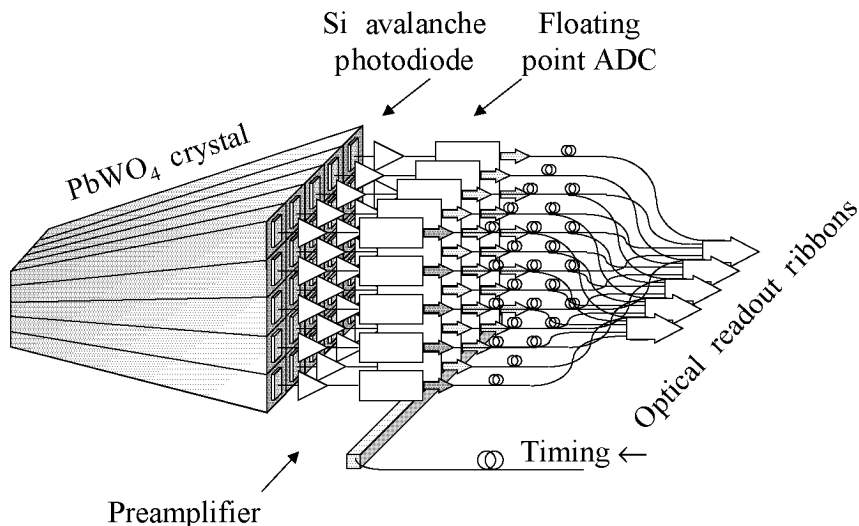


Figure 3.8: The ECAL readout chain [54].

Inner tracking detectors

The inner tracking system of CMS [56], consisting of silicon pixel and silicon strip detectors, is designed to reconstruct high- p_t muons, isolated electrons and hadrons in the range $|\eta| < 1.6$ with a momentum resolution of $\delta p_t/p_t \approx (15 \cdot p_t \oplus 0.5)\%$ (p_t in TeV/c), gradually degrading to $\delta p_t/p_t \approx (60 \cdot p_t \oplus 0.5)\%$ as $|\eta|$ approaches 2.5 [56]. An important design goal is the ability to tag and reconstruct b -jets, using the secondary vertex from b -decays, as well as to identify tau leptons. The tracker should also have a fast response (reducing pile-up effects) and all this without putting a too large amount of material in front of the calorimeters.

In total, the CMS tracker will implement 24328 silicon sensors, which cover an area of 206 m^2 .

The current layout is schematized in figure 3.9. Starting from the inner part and moving outwards, the tracker will be composed by silicon pixel and silicon micro-strip devices, distributed in a cylindrical volume of 6 m length and with an outer diameter of 2.4 m. In the central rapidity region ($\eta \lesssim 1.2$) the detectors are arranged in a barrel geometry, while in the forward regions they are arranged as disks, segmented into radial petals.

Data acquisition and trigger

The LHC pp collision rate will be 40 MHz. A total amount of ~ 16 million read-out channels will have to be processed, producing events containing an average of 1 MByte of data (most of the information being provided by the tracker). A challenge will be the reduction of the events rate (to be written in a storage

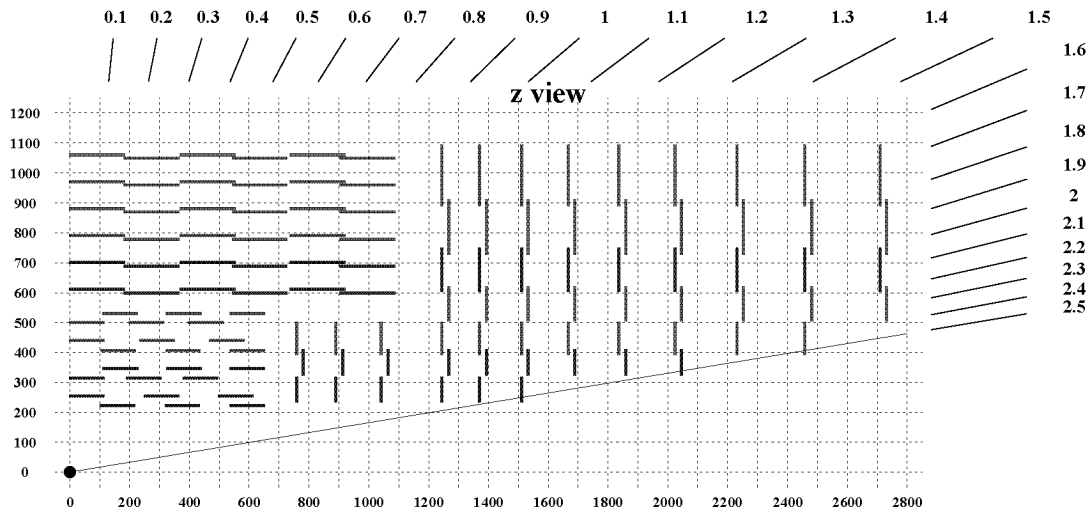


Figure 3.9: A transversal view of the tracker.

medium) to a frequency of ~ 100 Hz.

The trigger and data acquisition [57] consists of four parts: the front-end detector electronics, the calorimeter and muon first level trigger processors, the readout network and an on-line event filter system.

The level-1 trigger decision is based on the presence of physics objects such as high- p_t muons, photons, electrons and jets, as well as global sums of E_t and missing E_t . Each of these physics objects is required to pass a series of p_t or E_t thresholds, giving an output rate of 100 kHz. At this stage the data will be stored in 500 readout memories. The event builder is a large switching network with a total throughput of about 500 Gbit per second. The following event filter consists of a set of high performance commercial processors organized into many farms convenient for on-line and off-line applications. One event will be processed by a single CPU. The farms replace the traditional second and higher level triggers and will reduce the level-1 trigger rate from 100 kHz to a final event rate of about 100 Hz.

3.1.4 Time schedule

The choice of using a surface hall rather than the underground area, allows the construction and installation of the detectors in parallel with the civil engineering works. For example the assembly and testing of the CMS magnet take place in a large surface hall before lowering it into the underground experimental cavern. The present plans for the LHC collider, foresee the first beam injection and a sector test in 2005. For the detectors, the installation should be completed by January 2006, followed by the LHC beam commissioning (February-March 2006).

According to this time schedule, the first collisions and pilot run should take place in April 2006 at an initial luminosity of $\mathcal{L} \cong 5 \times 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$. After about 4 weeks at this regime, the first physics run will start in August 2006 at a luminosity $\mathcal{L} \cong 2 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$ and will last about 7 months [58].

3.2 Event simulation and reconstruction

Before introducing the simulation set-up, I start with a description of the general characteristics of an high-energy physics event.

As its name indicates, the output of an event generator should be in the form of an “event”, with the “same” average behavior and the same fluctuations as the data. This randomness can be simulated by the use of Monte Carlo techniques [59].

3.2.1 The complexity of high-energy processes

From the detector point of view, a pp -event can be subdivided in following processes:

- a. hard scattering processes between partons, like for example:

$$\begin{aligned} u\bar{d} &\rightarrow W^+ \rightarrow l^+\nu \quad \text{or} \\ gg &\rightarrow H \rightarrow ZZ \rightarrow l^+l^- q\bar{q}, \end{aligned}$$

which constitute the “key” of the event;

- b. the underlying event from the debris of the proton, resulting in many low p_t particles;
- c. multiple proton–proton interactions, relevant especially at the highest luminosities.

In point **a** the hard physics process was given in its simplest form, that is, as an interaction between fundamental objects of nature, *i.e.* quarks, leptons and gauge bosons. This is valid only as a first approximation. Corrections to this picture can be subdivided into three main classes:

- i. Bremsstrahlung-type modifications, *i.e.* the emission of additional final-state particles by branchings such as $q \rightarrow qg$. Because of the largeness of the strong coupling constant α_s , and because of the presence of the triple gluon vertex, emissions of quarks and gluons are especially prolific. These are simulated using so called “parton shower” techniques.
- ii. Higher order corrections, which involve a combination of loop graphs with the soft bremsstrahlung parts.

- iii. For quarks and gluons this must be complemented with a picture for the hadronization process (fragmentation and decay), wherein the colored partons are transformed into jets of colorless hadrons, photons and leptons.

The simple structure that we started with has now become considerably more complex. The original physics is not gone, but the skeleton process has been dressed and a comparison between theory and experiment has become more complicated.

It is at this point that event generators give their most important contribution. In order to analyze complex event structures, experimentalists can use computer simulations to generate events as detailed as can be modeled from the known theory.

Given the different levels one can reach in the description of high-energy processes, different Monte Carlo generators, with different simulation targets, are available. For example there are event generators dedicated to the simulation of only few hard scattering processes at parton level (point **a**). To this class belong the DKS generator [18], which is specialized in the simulation of the production of electroweak boson pairs. An interesting characteristic of this program is that it includes also complete next-to-leading-order (NLO, or $\mathcal{O}(\alpha_s)$) corrections (point **ii**). Other event generators offer a broader choice of different production mechanisms and subsequent decay modes. The PYTHIA Monte Carlo [19] is one of these simulation programs. It includes also hard and soft interactions, parton distributions, initial and final state parton showers, multiple interactions, fragmentation and decays, *i.e.* essentially all the points listed before. Nevertheless, for the production of weak boson pairs, it lacks the full NLO corrections. This makes a comparison of the PYTHIA and DKS generators interesting. In chapter 6 I proceed then to an analysis of the two program outputs. Finally there are generators dedicated to the simulation of the interactions between high-energy particles and stable matter. These programs find their application in the description of the signatures left by the different particles in the different detector parts. GEANT [15] belongs to this class of simulation tools, and will be applied in chapter 5 for the study about $\pi^\pm - e$ confusion in the CMS ECAL.

In general, by combining several types of event generators, one can reach a complete description of the physics processes which occur in the detectors, allowing a full comparison between data and expectations.

3.2.2 Reconstruction parameters

Multi-particle production is a characteristic feature of current high-energy physics. Depending on the center-of-mass energy and on the involved particles, the observed multiplicities range typically between two (as for example by an $e^+e^- \rightarrow \mu^+\mu^-$ event at LEP) and a hundred (like for a $p\bar{p} \rightarrow t\bar{t}$ event at Tevatron). With future machine (e.g. LHC) this range will be extended upwards.

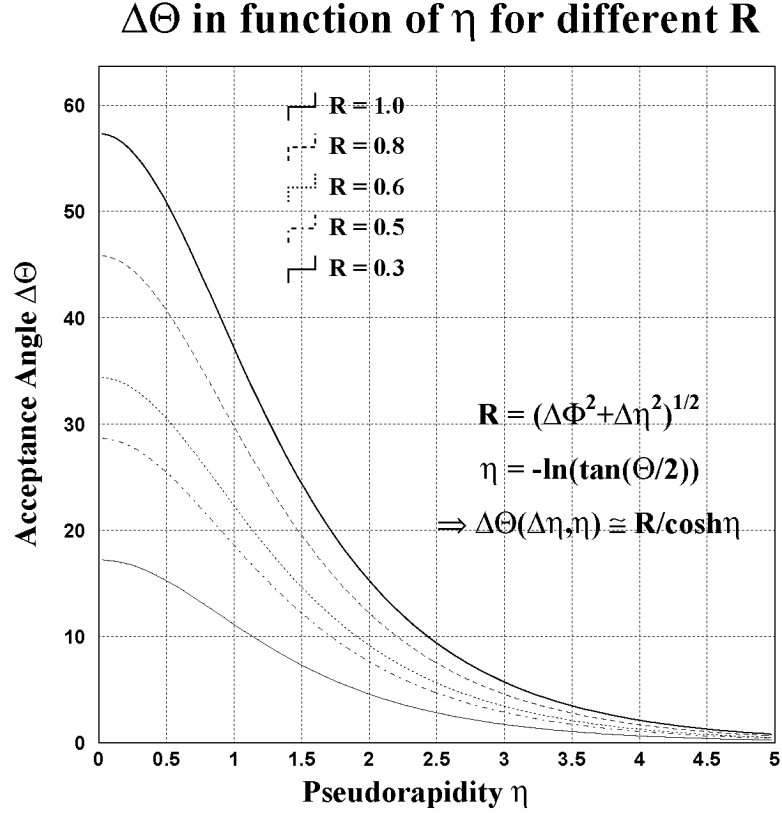


Figure 3.10: Acceptance polar angle $\Delta\Theta$ as a function of the pseudorapidity η and for different values of the acceptance R .

One goal for a multi-purpose detector is the ability to reconstruct jets present in the event: the parton shower has to be subdivided in the jet-constituents. As mentioned before, besides the jets, the detector should also detect isolated high- p_t e , μ and γ .

In order to identify a jet or a lepton as isolated, some selection conditions are required.

A widely used variable is the cone-size $R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$, where ϕ is the angle in the transverse plane and η is the pseudorapidity. The rapidity y of a scattered particle is defined by:

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right), \quad (3.3)$$

where E and p_z are respectively the particle energy and the momentum in the beam direction. For $p \gg m$ the rapidity can be substituted with the so called *pseudorapidity* η :

$$\eta \equiv -\ln[\tan(\Theta/2)] \approx y, \quad (3.4)$$

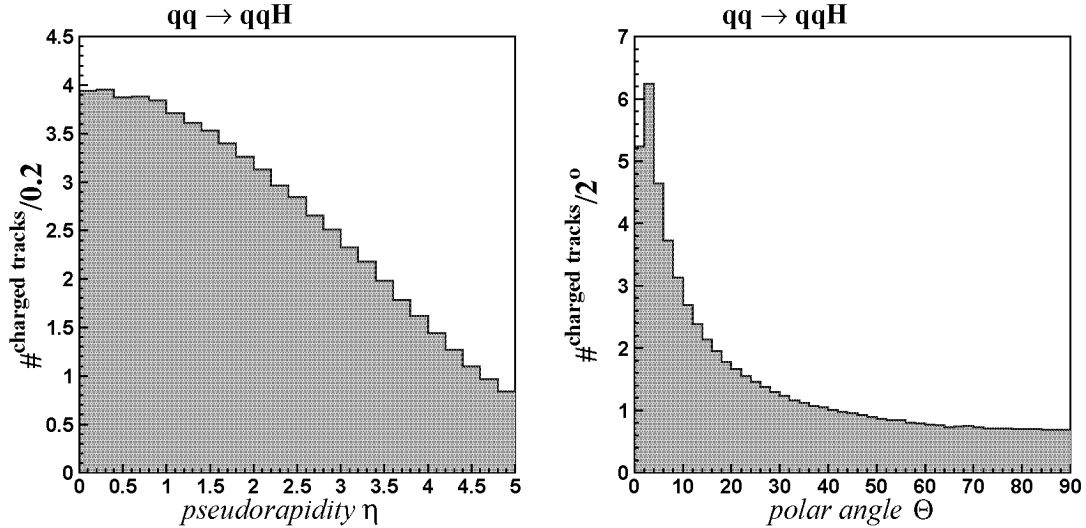


Figure 3.11: Average number of charged tracks (which consist in $e^\pm, \mu^\pm, \pi^\pm, K^\pm, p$ and \bar{p} with $p_t > 0.5$ GeV) expected in every $qq \rightarrow qqH$ event as a function of the pseudorapidity $|\eta|$ (left) and of the polar angle Θ (right). Typically one observes between 1 and 4 charged tracks in every interval $\Delta\eta = 0.2$. For small scattering angles Θ , this corresponds to an average of 5 charged tracks within a polar angle $\Delta\Theta = 2^\circ$, while only about 1 charged track is expected in the same interval in the barrel region of the detector.

where $\cos\Theta = p_z/p$ is the scattering angle with respect to the beam direction. In order to belong to a jet, a particle is required to have a distance in ϕ and η in comparison to the jet direction, which is consistent with the maximal value of R . In the same way, an isolated lepton should not have too many particles inside the cone defined by the acceptance R , besides eventually bremsstrahlung γ 's. A consequence of the utilization of the variable R is that, for the cone defined by $\sqrt{\Delta\eta^2 + \Delta\phi^2}$, the acceptance $\Delta\Theta$ in the polar angle Θ depends on the pseudorapidity. Figure 3.10 shows $\Delta\Theta$ as a function of η , for different acceptance values R . As can be observed, at larger η values (that is by smaller scattering angles) the "polar" acceptance $\Delta\Theta$ is smaller than by centrally emitted jets, leptons or photons. This choice is consistent with the topology of an high-energy collision: at small angles with respect to the beam direction, the decaying partons or leptons are normally boosted in the beam direction, giving rise to narrow jets or very close decay products. Moreover, at LHC these space regions are typically filled by large amounts of scattered particles (see figure 3.11). In general, for the number n_i of scattered particles, it is observed that:

$$\frac{dn_i}{d\Theta} \neq \text{const.} \quad \text{while} \quad \frac{dn_i}{d\eta} \approx \text{const.} \quad (3.5)$$

Therefore, for smaller angle Θ , the selection requires smaller acceptance angles $\Delta\Theta$, while the number of accepted particles in the cone defined by R remains about constant. The highest detector granularity is then needed to separate the different energy depositions.

The case is different for a high energy process produced almost "at rest": if its invariant mass is not too large, in the laboratory frame its decay products are less boosted and therefore spread in all directions, eventually entering the small rapidity region. For this reason, for large angles Θ (*i.e.* for $\eta \approx 0$), one normally observes broader jets and larger opening angles between decay products. Larger acceptance angles $\Delta\Theta$ are thus needed.

A typical value for R is 0.5. As shown in figure 3.10, this value corresponds to an acceptance angle $\Delta\Theta \approx 30^\circ$ at $\eta = 0$ and to $\Delta\Theta \approx 5^\circ$ at $\eta = 2.5$.

3.3 The simulation frame

As mentioned in the previous section, in this work Monte Carlo generators are used to simulate processes at two different levels: at detector level, where the goal is to describe the signature given by charged pions and electrons within the CMS electromagnetic calorimeter, and at the interaction level, with the purpose of studying high-energy physics processes. For these two specific applications, different kind of event generators, as well as two simulation set-ups, are used.

3.3.1 CMS detector simulations

Once a particle track has been reconstructed and once this track has been correlated to an energy deposition, the next experimental question is then to relate these signatures with the right particle. In other words this particle has to be identified correctly, otherwise the physics meaning of the whole event could be lost.

In chapter 5 I investigate the case where charged π 's leave a large fraction of their energy in the crystals of the Electromagnetic Calorimeter (ECAL). The effect could then be enhanced if several charged particles cross the Avalanche Photo Diodes (APD). These devices are supposed to measure the light produced in the crystals, but a signal can also be produced by a charged particle which cross their layers.

The combination of these two effects could give rise to an observed large energy deposition in the ECAL. This would be associated with the momentum of a charged particle measured in the tracker. The obtained signature is consistent with the one expected from an electron and therefore could lead to a possible particle misidentification.

A $\pi^\pm - e$ confusion would be especially dangerous, because of the enormous amount of π^\pm 's produced at each pp collision.

In order to perform this study, a simplified ECAL simulation set-up is used. High-energy π^\pm 's and e 's are first produced by PYTHIA and then inserted in the calorimeter, which is described with the help of GEANT. Since this simulation frame is only used in chapter 5, it will be described there in more detail.

3.3.2 Physics simulations

After having identified particles, together with their kinematics, the physics analysis can begin.

The studies carried out in this work are mostly done by means of the PYTHIA Monte Carlo. In the following I describe the methods employed to select isolated leptons, photons and jets, which will then be used as starting variables for the physics studies. This is done by first selecting the "detectable" particles (*i.e.* compatible with the CMS detector performances) among those produced by PYTHIA and then by applying different selection algorithms.

pp events generated with the PYTHIA Monte Carlo program

The PYTHIA program has been developed in order to generate complete events, in as much detail as experimentally observed ones, within the bounds of our current understanding of the underlying physics.

The output of the program is in the form of a list of the involved particles, including all their kinematical structure. The event can be analyzed at different stages: right after the hard-scattering process, before the full hadronization/decay process or starting from the "stable" particles given at the end. This allows to perform the analysis starting from different levels and also to check the validity of the reconstruction algorithms.

Moreover, the user has the full choice on the processes to be simulated: colliding particles, production and decay mechanisms, models and so on, allowing a broad range of possible analysis.

Table 3.1 shows an example of the PYTHIA output for the production of a Standard Model Higgs via the fusion of two vector bosons. The related feynmann diagram is shown in figure 3.12.

The first column gives simply the position of the particle in the tree. The label "status" is used to express the current condition of the parton/particle stored in the line. The entries labeled with "status = 2" corresponds for example to particles which have subsequently decayed or fragmented. "ID" is an identity code according to the PDG standard (e.g. "11" are electrons). "parent" and "child" are pointers to the positions where, respectively, the first "mother" and "daughter" of the particle are stored. Finally, in the last five columns, the kinematics of the particle are given. Clearly all the 4-momentum components can be accessed and this table represents only one output option.

event dump p_t -cut: 2.5 η -cut: 5										
tree	status	name	ID	parent	child	y	Φ	p_t	m	p_z
1	3	p^+	2212	0	0	9.6	0	0	0.938	$7 \cdot 10^3$
2	3	p^+	2212	0	0	-9.6	0	0	0.938	$-7 \cdot 10^3$
3	3	d	1	1	0	6.1	90.3	1.27	-0	298
4	3	u	2	2	0	-6.8	256	0.48	-0	-217
5	3	d	1	3	0	4.6	26.5	4.14	0	213
6	3	u	2	4	0	-3.8	317	5.72	0	-129
7	3	d	1	5	0	1.5	92.6	42.1	0.33	86
8	3	u	2	5	0	-0.33	306	68.7	0.33	-22.9
9	3	h^0	25	6	0	0.12	159	32.5	170	21.2
10	3	W^+	24	9	0	0.092	43.2	14.6	81.9	7.68
11	3	W^-	-24	9	0	0.18	175	42	80	16.4
12	3	e^+	-11	10	0	0.79	5.95	36.3	$5 \cdot 10^{-4}$	31.6
13	3	ν_e	12	10	0	-0.82	166	26.2	0	-23.9
14	3	e^-	11	11	0	1.3	213	25.8	$5 \cdot 10^{-4}$	45.5
15	3	$\bar{\nu}_e$	-12	11	0	-0.94	138	26.8	0	-29.1
16	2	h^0	25	9	18	0.14	156	34.1	170	24
18	2	W^+	24	10	20	0.092	43.2	14.6	81.9	7.68
19	2	W^-	-24	11	22	0.18	175	42	80	16.4
20	1	e^+	-11	12	0	0.79	5.95	36.3	$5 \cdot 10^{-4}$	31.6
21	1	ν_e	12	13	0	-0.82	166	26.2	0	-23.9
22	1	e^-	11	19	0	1.3	213	25.8	$5 \cdot 10^{-4}$	45.5
23	1	$\bar{\nu}_e$	-12	19	0	-0.94	138	26.8	0	-29.1
84	1	π^-	-211	64	0	0.52	182	2.57	0.14	1.4
92	1	π^+	211	64	0	-0.27	309	9.6	0.14	-2.6
224	1	π^+	211	77	0	-0.76	318	4.62	0.14	-3.85
240	1	\bar{n}^0	-2112	89	0	-0.3	290	3.61	0.94	-1.12
242	1	π^-	-211	90	0	-0.42	307	4.44	0.14	-1.92
243	1	π^+	211	90	0	-0.22	306	2.57	0.14	-0.578
244	1	n^0	2112	91	0	-0.31	306	37.8	0.94	-11.9
245	1	π^-	-211	91	0	-0.29	304	9.23	0.14	-2.7
248	1	π^+	211	95	0	-1.2	83.4	2.51	0.14	-3.66
271	1	π^-	-211	113	0	1.4	93	2.7	0.14	5.38
281	1	p^+	2212	124	0	3.4	112	3.7	0.938	60
445	1	γ	22	272	0	1.4	106	2.87	0	5.32

Table 3.1: Example of PYTHIA output by a $qq \rightarrow qqH \rightarrow qqWW \rightarrow jj e \nu e \nu$ event (Higgs production via weak boson fusion, $M_H = 170$ GeV).

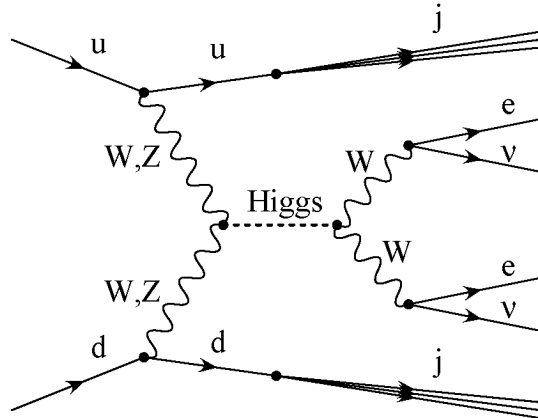


Figure 3.12: Feynmann diagram of the weak boson fusion $qq \rightarrow qqH \rightarrow qqWW \rightarrow jj e \nu e \nu$ event listed in table 3.1.

To reduce the size of the list, I put two cuts on the “stable” particles (status = 1) by requiring a p_t larger than 2.5 GeV and a pseudorapidity $|\eta|$ smaller than 5. Starting from the diagram of figure 3.12, one can then find at the beginning of the list (status = 3) all the involved particles, namely: the two colliding protons (1 and 2 in the tree), the two quarks which give rise to the hard scattering process (5-6 before and 7-8 after the W, Z emission), the produced Higgs (9) with the two W boson daughters (10-11) and the decayed leptons (12-15). This first part of the list gives a documentation of the event at parton level, allowing an analysis restricted to the hard scattering process.

Particles labeled with “status = 1” remain undecayed or unfragmented and represent the “final state” given by the generator. Starting from this part of the event list, one can select the “detectable particles”, that is the “stable” ($\tau_i \gtrsim 10^{-10}$ s) hadrons, leptons and photons which are consistent with the detector acceptance and resolution. Isolated e, μ, γ , as well as jets, are subsequently selected starting from this list of detectable particles.

Selection algorithms for isolated leptons and γ 's

An important feature of the PYTHIA-level analysis are the algorithms which define isolated leptons, isolated photons, τ decay products and which reconstruct jets originating from gluons and quarks.

I first select particles which fall within the CMS acceptance, taking all $e, \mu, \gamma, n, p, \pi^\pm, K^\pm, K_L^0$ with a minimal p_t of 500 MeV and pseudorapidities $|\eta|$ smaller than 4.5 into account. These detectable particles are then used for the different reconstruction algorithms.

Starting with electrons or muons which have a p_t larger than 10 GeV and which are consistent with the ECAL/muon system acceptance (*i.e.* $|\eta| < 2.5$), the

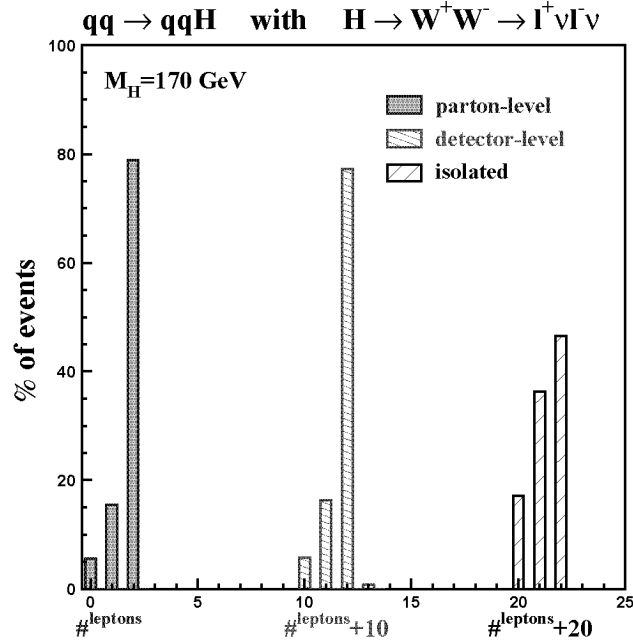


Figure 3.13: $qq \rightarrow qqH$ with $H \rightarrow WW \rightarrow l\nu l\nu$: number of selected leptons at parton-level (left columns), after the hadronization/decay (middle columns) and after the application of the isolation algorithm (right columns).

energy of all the detectable particles with $p_t > 1$ GeV, lying within a cone of radius $R = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.5$, is added. A lepton is assumed to be isolated if the following conditions are fulfilled:

- its energy is at least 90% of the total cone energy,
- the total mass of this cone is smaller than 2 GeV,
- there is no other detectable particle within a smaller cone of $R = 0.15$ around the lepton.

The weak boson fusion process $qq \rightarrow qqH$, with the subsequent Higgs decay $H \rightarrow WW \rightarrow l\nu l\nu$ ($l \equiv e, \mu$), is used to test the isolation algorithm. At parton-level (*i.e.* by the hard scattering process), the total amount of events where both leptons are consistent with the rapidity and minimal transverse momentum cuts ($|\eta| < 2.5$, $p_t > 10$ GeV) is about 79%, as shown in the figure 3.13 (left columns). About the same percentage of events is accepted after the the full hadronization/decay process (columns in the middle). Once the isolation algorithm is applied for these leptons, one obtains two isolated leptons in about 47% of the events (right columns). By comparing this value with the previous one

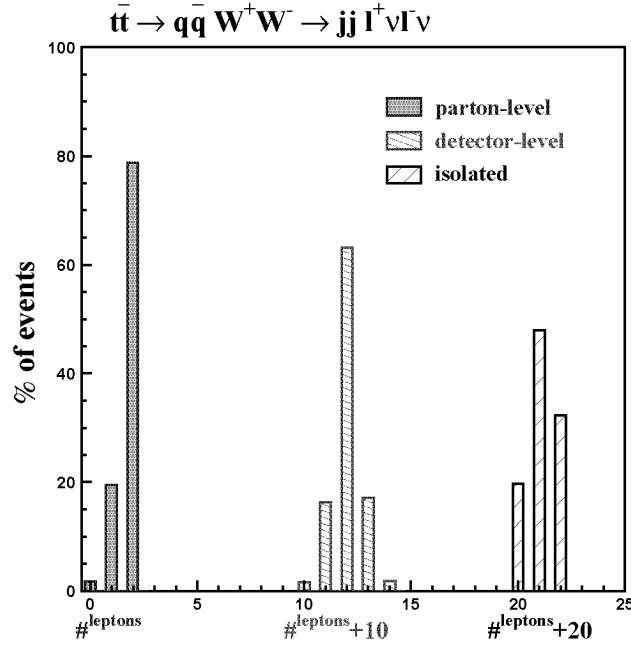


Figure 3.14: Same as figure 3.13, but for the background process $t\bar{t} \rightarrow bW^+ \bar{b}W^- \rightarrow jj l\nu l\nu$.

(events with two “detectable” leptons), one obtains:

$$\frac{0.47}{0.79} \approx 0.6,$$

which gives the probability that both detectable e ’s or μ ’s are isolated. From this value one can extract the probability for a single detectable electron or muon to be isolated in a $H \rightarrow WW \rightarrow l\nu l\nu$ event, which is $\sqrt{0.6} \approx 77\%$ (before any other selection cut and assuming an absence of correlations between the isolation of the two leptons).

Figure 3.14 gives the number of accepted leptons per event for the process $t\bar{t} \rightarrow bW^+ \bar{b}W^-$, with $W \rightarrow l\nu$ ($l \equiv e, \mu$). At parton level, one observes a similar percentage of events with 2 selectable leptons as in the $qq \rightarrow qqH \rightarrow jj l\nu l\nu$ processes: $\sim 79\%$. If the η - and p_t -acceptance cuts are applied on the “detectable” particles, two or more leptons are observed in about 82% of the events. The increase is due to the presence of other leptons produced during the hadronization process. After the application of the isolation algorithm, this percentage falls then to only $\sim 32\%$. By comparing events with at least two detectable leptons and events with two isolated leptons, one obtains:

$$\frac{0.32}{0.79} \approx 0.4.$$

The probability that both detectable e/μ are isolated is thus of the order of 40%. Therefore, the probability for a muon or an electron to be isolated in a $t\bar{t}$ -event is found to be $\sqrt{0.4} \approx 63\%$, which is a little lower compared to 77% for a lepton in a $H \rightarrow WW \rightarrow l\nu l\nu$ event.

In order to find isolated photons, the same criteria as for leptons are used, where in addition also other γ 's close to the tested one are accepted (more precisely, the γ 's should be found within 2 neighboring crystals).

Regarding the τ leptons, I first eliminate from the list of detectable particles all the leptons and photons, which have already been identified as isolated. As "build axis" for the reconstruction of the tau event, I look for a charged pion, kaon or proton with $p_t > 5$ GeV and $|\eta| < 2.5$. Other particles are then added to this axis if they fall within a inner cone of radius $R = 0.15$ around it. The mass of the obtained vector should be smaller than 2 GeV and its p_t larger than 20 GeV. Moreover, the multiplicity of its constituents should be less than six and among them, one or three should be charged (π^\pm , K^\pm or protons). If these requests are fulfilled, then the total energy of all particles lying in a outer cone of $R = 0.5$ around the vector is collected. In order to identify this vector as a τ -jet, the energy in the inner cone should constitute more than 95% of the total energy in the outer cone.

Jet reconstruction

If we are interested in the measurements of quarks and gluons, we cannot access the information about them directly. We have to extract the q/g properties from the properties of jets. Unfortunately, in most cases, the energy of a jet is not the same as the q/g energy. It depends strongly on the definition of jet. In case of a jet reconstruction via cone algorithm, the reconstruction fidelity depends on the cone size in which the energy of the jet is collected (and also a little on the splitting or merging parameters for jet candidate pairs). If the cone size is too large and thus a lot of additional energy not coming from the q/g is included, then the energy resolution is bad. On the other hand, a small cone has the disadvantage of leaving a fraction of the q/g energy outside the cone.

The real reconstruction algorithms, which will be used in the data analysis, are developed with full detector simulations. For the fast PYTHIA simulations used in this work, some aspects of the problems are ignored. In particular, those affecting the particle momentum resolution and the particle identification (like magnetic field effects, electronic noise, particle separation in the tracker and so on). However, we take into account the spatial coverage of the sub-detectors and, to some extent, also the granularity of the calorimeters.

Similarly to isolated leptons and gammas, jets are reconstructed starting from all detectable particles. From this list, isolated e, μ and γ as well as τ -jets are eliminated. At the end of the reconstruction process, the obtained jets should be

3.3 The simulation frame

consistent with the following transverse momentum and rapidity values:

$$\begin{cases} p_t^{\text{jet}} > 20 \text{ GeV} \\ |\eta|^{\text{jet}} < 4.5 . \end{cases}$$

In this report, two different algorithms for the jet reconstruction are investigated. After having ordered the detectable particles by their transverse momentum, the first one starts from the particle with largest p_t , to which the momentum vectors of particles lying within a cone of $R < 0.5$ are added. The reference vector obtained from this first iteration is assumed to be about in the center of the jet. Therefore, to obtain the final jet, a second iteration is performed by adding all particles found in $R < 0.5$ with respect to this reference axis.

In the second algorithm, with the same aim to start the jet reconstruction from its center, I look for particles ($p_{t,\text{min}} > 5 \text{ GeV}$), which have the maximal amount of other particles within a cone of $R < 0.4$ around themselves. Starting from this “central particle”, all other members of the jet are then added in two iterations: the first one using a cone of $R < 0.3$ and the second one of $R < 0.5$. The second iteration is done in order to avoid the mis-inclusion of particles, due to the movement of the jet direction during the reconstruction.

The two algorithms produce very similar results in the central detector region (*i.e.* for $|\eta|$ smaller than 2.5), while in the forward pseudorapidity range $2.5 < |\eta| < 4.5$ the first one tends to identify also isolated electrons and γ as jets². The reason of this mis-identification is simple: in these $|\eta|$ -regions the isolation algorithms have not been used. In fact, without the help of ECAL, HCAL and Tracker, it is not possible to apply the isolation cuts and one is left with only the information about the energy deposition in the Very Forward Calorimeters. Above $|\eta| \sim 3$, jets will be identified as “clusters” of energy depositions and about the same signature is expected from e, τ and γ . From this point of view the second algorithm is “cheating”, since for every particle it uses also in this $|\eta|$ domain the information about the number of surrounding particles. The first jet algorithm is thus more realistic than the second one. However, for the investigated processes, this difference has practically no effect: for the signals all the leptons are always required to lie within the ECAL acceptance ($|\eta| < 2.5$). Moreover, the amount of new backgrounds arising from the identification of a lepton or a photon as a jet is very small. For example, for the weak boson fusion process $qq \rightarrow qqH$ with $H \rightarrow WW \rightarrow l\nu l\nu$, both leptons must have $|\eta| < 2.5$, and thus should not be identified as jets. In addition, the signatures with two isolated leptons, missing energy, a forward jet and another high- p_t e, μ or γ , which is misidentified as jet and matches the selection cuts, are very rare compared to other backgrounds, as for example the $t\bar{t}$ production.

As mentioned, in this work the reconstruction of jets is especially important

²In contrast to electrons and photons, μ 's with $|\eta| > 2.5$ will not be measured and should thus not contribute to the observable particles.

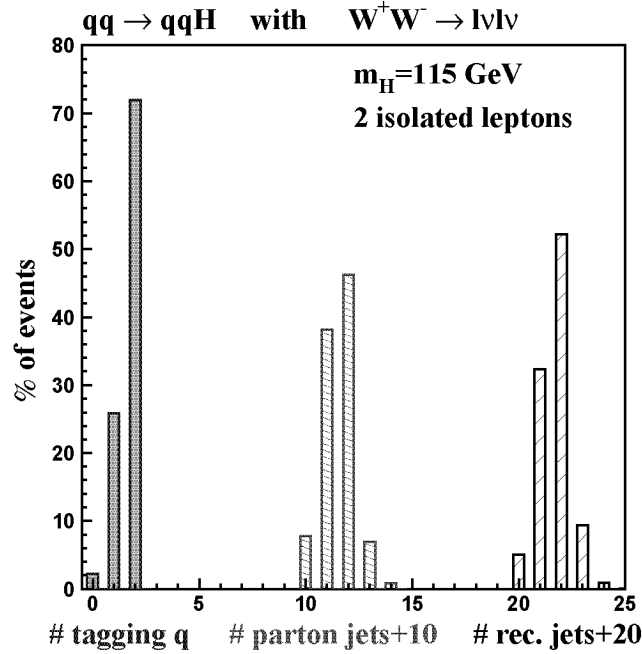


Figure 3.15: Example of the efficiency of the jet reconstruction algorithm applied on the weak boson fusion process $qq \rightarrow qqH \rightarrow qqWW \rightarrow jjl\nu l\nu$. After some simple selection cuts (2 jets with $|\eta^{\text{jet}}| < 4.5$, $p_t^{\text{jet}} > 20$ GeV, 2 isolated leptons with $|\eta^l| < 2.5$, $p_t^l > 20$ GeV), the histogram shows the obtained number of tagging quarks (left columns), reconstructed jets before (middle columns) and after the hadronization (right columns).

for the study of the weak boson fusion processes. The selection of these events requires the presence of two tagging jets, arising from the two scattered quarks (see for example figure 3.12). Therefore, as applied examples for the jet algorithm, in this section I make use mostly of the signature $qq \rightarrow qqH \rightarrow qqZZ \rightarrow 2j4\nu$, which is not affected by the Higgs decay products, or directly of the signal process $qq \rightarrow qqH \rightarrow qqWW \rightarrow jjl\nu l\nu$.

To analyze the effectiveness of the algorithm through the different reconstruction steps, the jet selection is performed at different levels. Following the arguments introduced in section 3.3.2, one can either directly select the scattered quarks involved in the weak boson fusion process (tagging quarks), or use the fully reconstructed jets. Inbetween, I add also an “hybrid” solution, whose result is labeled “parton jets”. These latter are generated by simply applying the jet algorithm to the quarks and gluons before the full hadronization process, instead of using the detectable particles. The motivation behind this choice is to see if, after having selected both tagging quarks, the parton jets are still consistent with the jet selection criteria and it is possible to reconstruct them after hadronization.

3.3 The simulation frame

The left columns of figure 3.15 show the number of tagging quarks per $qq \rightarrow qqH \rightarrow qqWW$ event, which fulfill simple selection cuts (2 quarks with $|\eta^q| < 4.5$, $p_t^q > 20$ GeV and 2 isolated leptons with $|\eta^l| < 2.5$, $p_t^l > 20$ GeV). In about 72% of the events, both tagging quarks are consistent with these acceptance criteria. Starting instead from the quarks and gluons before the hadronization, the amount of events with 2 or more parton jets drops to about 54% (46% have exactly 2 parton jets), as shown in the middle columns. There are at least two reasons which contribute to this loss: (1) the splitting of the tagging quarks into two or more jets (*i.e.* a large opening angle between the produced quarks and gluons) and (2) the presence of q/g outside the η and/or p_t acceptance. In both cases, the obtained parton jets could be then rejected by the selection criteria. Finally, after the full hadronization, 63% of the events possess 2 or more reconstructed jets (52% have 2 jets). The increase in the final state is due to several reasons. The first one is related to the fact that the parton jets are reconstructed starting only from quarks and gluons but not leptons. Furthermore, the reconstructed jets gain some contributions from the presence of other particles which “push” the jets above the p_t -threshold. These “intruder” particles can come from underlying events (mostly at high rapidities). Finally there could be some other jets arising from minimum bias events or from the products of the rest of the two involved protons.

In the following, when I speak about a “tagging quark”, a “parton jet” or a “reconstructed jet”, it always means that this q /jet is compatible with the η and p_t selection (*i.e.* $|\eta| < 4.5$, $p_t > 20$ GeV). In order to obtain an impression about the efficiency of the jet algorithm, I proceed in the following way: I count the number of $qq \rightarrow qqH$ events with two or more reconstructed jets ($\sim 63\%$) and I compare it with the number of events with two tagging quarks ($\sim 72\%$). This means that about

$$\frac{0.63}{0.72} \approx 87\%$$

of the selected parton level events are still compatible with the cuts also after the hadronization. Following these simple arguments, the selection efficiency per jet would be of $\sqrt{87\%} \approx 94\%$.

The number of jets is only a first indication of the efficiency of the reconstruction algorithm. To learn something more about its “quality”, figure 3.16 shows the pseudorapidity distributions of, respectively, tagging quarks, parton jets and reconstructed jets. The η -distributions of the reconstructed jets (before and after the hadronization) show very similar shapes. However, compared to the original tagging quarks, they tend to have a slightly smaller rapidity. This difference arises from the fact that, at large pseudorapidities, more particles belonging to the jets are outside the detectable η -region and below the p_t -threshold. Therefore, these jets will be either reconstructed with a smaller pseudorapidity compared to the original quark or could fall below the p_t -cut.

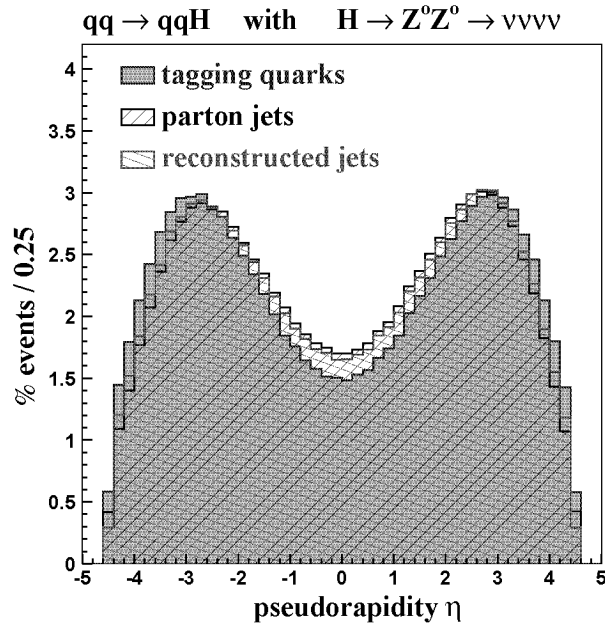


Figure 3.16: $qq \rightarrow qqH$: pseudorapidity distributions of tagging quarks and reconstructed jets (before and after the hadronization).

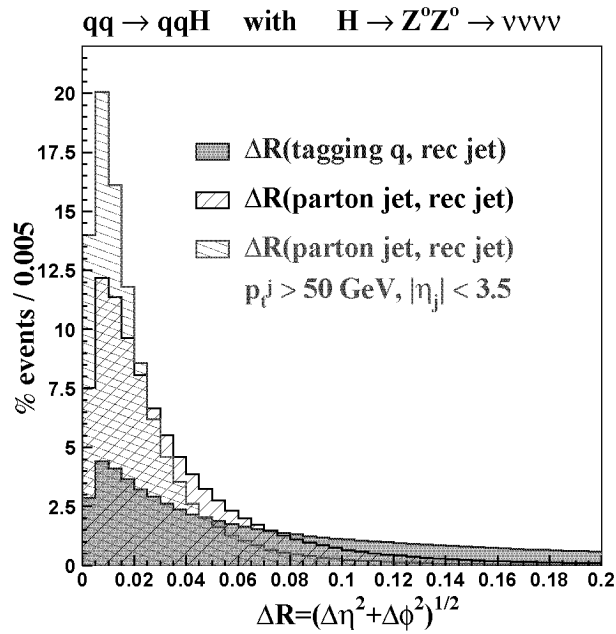


Figure 3.17: For weak boson fusion processes $qq \rightarrow qqH$, the histograms show the spatial distances (ΔR) between reconstructed jets and tagging quarks or between reconstructed jets and parton jets (with or without stronger selection cuts).

3.3 The simulation frame

Another parameter one can check is the direction of the jet. If one takes the vector of every tagging quark or parton jet and, in order to find its detectable product, one looks for the closest reconstructed jet (by minimizing $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$), the histograms shown in figure 3.17 is obtained. The result of these comparisons are summarized in table 3.2.

For about 45% of the tagging quarks, a reconstructed jet can be found in a cone of $\Delta R = 0.1$ (corresponding to roughly 5.7° or 100 mrad at $\eta = 0$). The remaining 55% cases include either events where the tagging quark is at the limit of the acceptance and after the hadronization it was not compatible with the η/p_t cuts, or mostly events where the quark splits into two or more jets. In fact at least this latter effect is no longer present, if instead of the tagging quarks, one starts from the parton jet and compares it with the closest reconstructed jet (in ΔR). In this case, after the hadronization, 49% of the parton jets can still be found within a ΔR of 0.025 and 88% in a ΔR of 0.1. If stronger selection criteria ($|\eta^{\text{jet}}| < 3.5$, $p_t^{\text{jet}} > 50$ GeV) are used, a further efficiency improvement (see table 3.2) can be observed: 71% of the initial parton jets lie in a cone of $\Delta R = 0.025$ with respect to the closest reconstructed jet and 97% within $\Delta R = 0.1$! From these results, one can conclude that the largest difference in direction between the tagging quarks involved in the hard scattering process and the reconstructed jets arises by the first step, that is as the quarks are subdivided into other quarks and gluons. Starting from this level (*i.e.* from the parton jets), a much better compatibility with the final state jets is instead observed.

As third criterion to judge the reconstruction algorithm, I use the transverse momentum of the jets. For this purpose, figure 3.18 shows the relative $\Delta p_t/p_t$ difference between every tagging quark/parton jet and the closest reconstructed jet (in ΔR). Table 3.3 shows the percentages of tagging quarks/parton jets, whose p_t is compatible with those of the final reconstructed jets.

Starting from a:	% of reconstructed jet found within:	
	$\Delta R < 0.025$	$\Delta R < 0.1$
tagging quark	18 %	45 %
parton jet	49 %	88 %
parton jet with $ \eta^{\text{jet}} < 3.5$, $p_t^{\text{jet}} > 50$ GeV	71 %	97 %

Table 3.2: According to the simulations shown in figure 3.17: percentage of tagging quarks or of parton jet, for which a reconstructed jet is found within a cone of $R < 0.025$ (first column) or of $R < 0.1$ (second column).

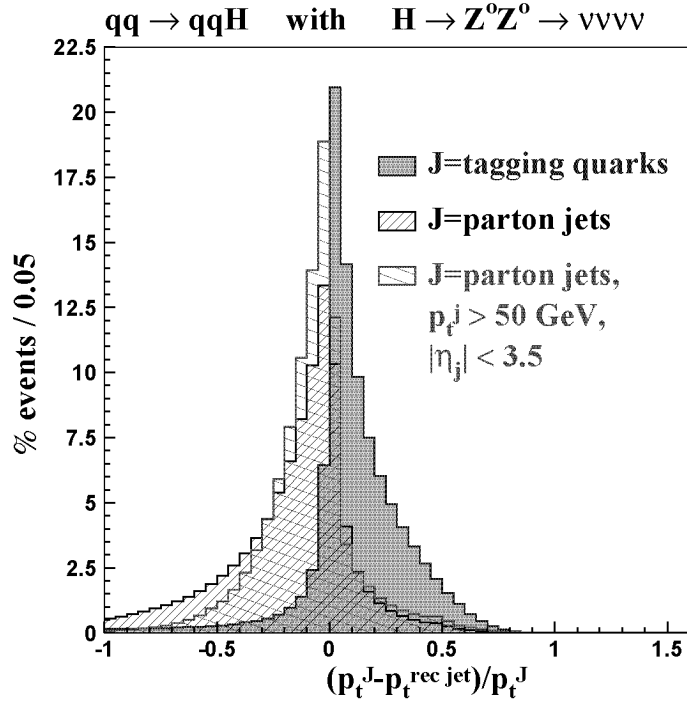


Figure 3.18: Relative transverse momentum difference $\Delta p_t/p_t$ between reconstructed jets, the quarks originating from $qq \rightarrow qqH \rightarrow qqZZ$ and the jets at parton level.

Starting from a:	% of reconstructed jet found within:	
	$\Delta p_t/p_t < 0.1$	$\Delta p_t/p_t < 0.25$
tagging quark	44 %	70 %
parton jet	38 %	63 %
parton jet with $ \eta^{\text{jet}} < 3.5$, $p_t^{\text{jet}} > 50 \text{ GeV}$	48 %	78 %

Table 3.3: According to the simulations shown in figure 3.18: percentage of tagging quarks or of parton jet, for which a reconstructed jet is found within a $\Delta p_t/p_t < 0.1$ (first column) or of $\Delta p_t/p_t < 0.25$ (second column).

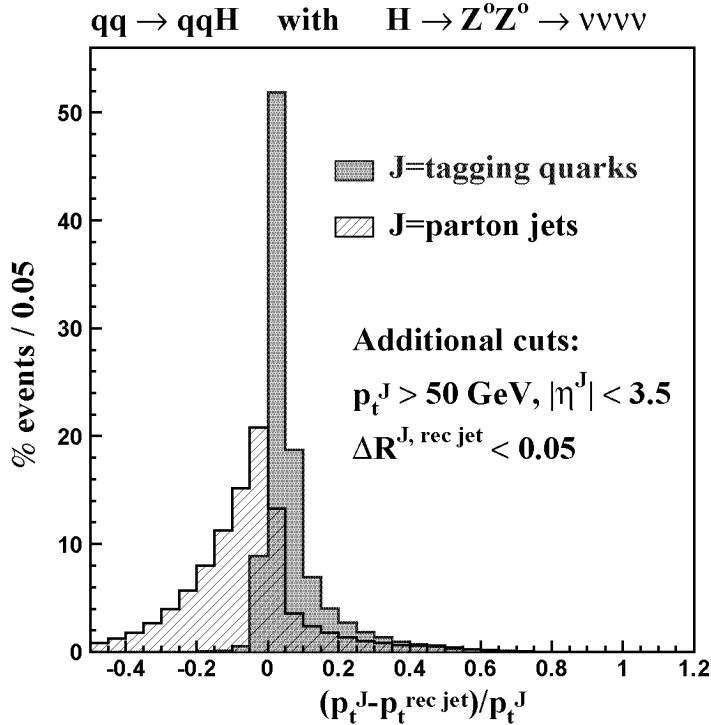


Figure 3.19: Relative transverse momentum difference $\Delta p_t/p_t$ between reconstructed jets, the quarks originating from $qq \rightarrow qqH \rightarrow qqZZ$ and the so-called “parton jets”. The comparison is made only for high- p_t jets, which have already been found to be collinear (see text).

As a result one finds that 44% of the tagging quarks are reconstructed within a $\Delta p_t/p_t = 0.1$. Allowing a larger difference $\Delta p_t/p_t = 0.25$, a reconstructed jet is found for 70% of the tagging quarks. As shown in the figure 3.18, generally the reconstructed jets have smaller transverse momenta as the original tagging quarks.

This is probably due to the presence of undetectable particles, to the splitting of the quark into two or more jets and to the acceptance cuts, which eliminate the soft part. That’s not the case if one compares the parton jets with the closest reconstructed jet. As said before, reconstructed jets have normally larger momenta than parton jets.

Of course the transverse momentum difference should be smaller if one compares the p_t of jets, which we know to have been reconstructed with the correct direction. If the same comparison is repeated, but by requiring the collinearity between the tagging quark/parton jet and the reconstructed jet (*i.e.* they should not be separated by a ΔR larger than 0.05), one obtains the histograms shown in figure 3.19 (for central, high- p_t jets). Table 3.4 summarizes these results. Starting

Starting from a collinear ($\Delta R < 0.05$):	% of reconstructed jet found within:	
	$\Delta p_t/p_t < 0.1$	$\Delta p_t/p_t < 0.25$
tagging quark	80 %	94 %
parton jet	53 %	83 %

Table 3.4: Same as table 3.3, but by requiring the collinearity between the tagging quark/parton jet and the reconstructed jet ($\Delta R < 0.05$) and only for more energetic and central jets ($|\eta^{\text{jet}}| < 3.5$, $p_t^{\text{jet}} > 50$ GeV).

from the tagging quarks, about 94% of the reconstructed jets lie within a $\Delta p_t/p_t$ difference of 0.25! This means that, if the reconstructed jet matches the direction of the tagging quark, then also its momentum is in very good agreement. Another interesting fact is that the relative p_t -difference between parton jets and reconstructed jet does not show a similar improvement as for the tagging quarks (cf. tables 3.3 and 3.4). This result confirms the previous arguments: a precise reconstruction of a tagging quark is hindered by its possible splitting into several jets and by edge-effects.

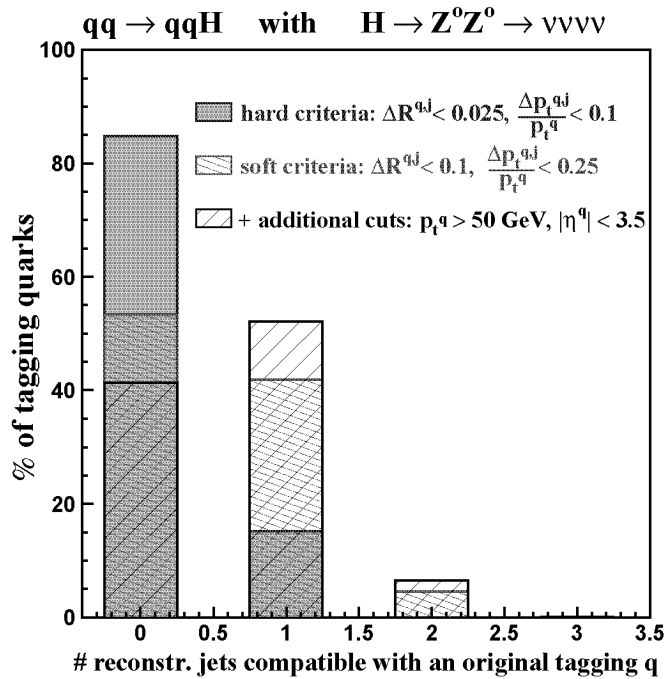


Figure 3.20: $qq \rightarrow qqH \rightarrow qqZZ$: starting from a tagging quark, how many reconstructed jets are compatible with it after the hadronization?

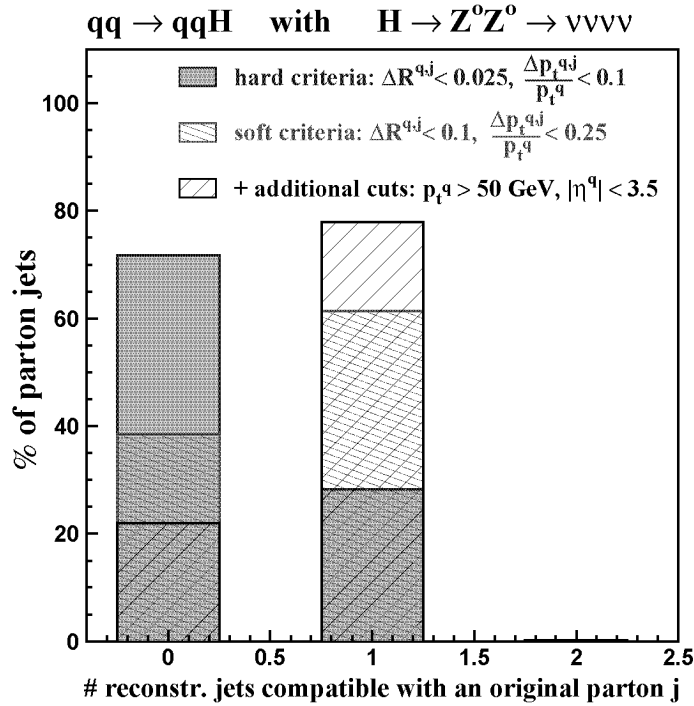


Figure 3.21: $qq \rightarrow qqH \rightarrow qqZZ$: when a jet is reconstructed at parton level, how many jets are compatible with it after the hadronization?

Finally, for the $qq \rightarrow qqH \rightarrow qqZZ \rightarrow jj\nu\nu\nu\nu$ process, the efficiency of the jet reconstruction algorithm can be extracted from figures 3.20 and 3.21. Starting from every tagging quark (fig. 3.20) or parton jet (fig. 3.21), the amount of compatible reconstructed jets are shown. What I do is the following: I take every tagging quark/parton jet and then I look for a reconstructed jet that matches its direction and transverse momentum. If any candidate is found, I look for a pair of reconstructed jets, and so forth. This means that the original tagging quark/parton jet can either evolve in a single reconstructed jet or in a combination of several jets, which reproduce its kinematics. The “compatibility” is judged with the help of two criteria: a “hard” one, where original and reconstructed jets should not be separated by more than $\Delta R = 0.025$ and $\Delta p_t/p_t = 0.1$, and a “soft” one, where the maximal ΔR and $\Delta p_t/p_t$ are respectively 0.1 and 0.25. In other words, I combine the same criteria regarding the direction and the transverse momentum precision as before and I summarize the results in a single table (3.5).

Starting from a tagging quark and by using the so-called “hard criteria” only for about 15% of the q ’s a matching reconstructed jet can be found (see table 3.5). This must affect the selection of $qq \rightarrow qqH$ events and result in a different accep-

% of reconstructed jets compatible with:	Starting from a:	
	tagging quark	parton jet
$\Delta R < 0.025, \Delta p_t/p_t < 0.1$	15 %	28 %
$\Delta R < 0.1, \Delta p_t/p_t < 0.25$	1:1	$\geq 2:1$
	40 %	5 %
+($ \eta^{\text{jet}} < 3.5, p_t^{\text{jet}} > 50 \text{ GeV}$)	1:1	$\geq 2:1$
	52 %	7 %
		60 %
		80 %

Table 3.5: According to the histograms shown in figures 3.20 and 3.21: percentages of “well reconstructed” tagging q /parton jets. The tagging quarks, for which a compatible reconstructed jet is found, are labeled as “1:1”. The label “ $\geq 2:1$ ” means instead that the tagging q ’s split into a combination of at least 2 jets.

tance between the PYTHIA analysis and the parton level studies [25, 26]. Using the “soft criteria” about 40% of the tagging quarks result in a single compatible reconstructed jet and slightly more than 5% split into two or more jets. However, in case of a splitting, the ≥ 2 reconstructed jets possess different kinematics in comparison with the original tagging quark (mostly less p_t) and add a new jet into the event, compromising its possibility to get through the selection cuts³.

As table 3.5 shows, for “harder” jets ($p_t > 50 \text{ GeV}$, $|\eta| < 3.5$), the efficiency increases slightly: for example for about 52% of the tagging quarks, a single matching reconstructed jet is found.

Further improvements are expected starting from a parton jet: in fact, with the “hard criteria”, for 28% of them a compatible reconstructed jet can be found, while with softer criteria more than 60% of the parton jets have a matching reconstructed jet ($\sim 80\%$ if only the “hard” jets are taken into account).

In conclusion, the algorithm reconstructs more than 60% of the “parton jets” with good precision. However, it has to be taken into account that the correlation between tagging quarks and gluon/quarks before the hadronization is not always straightforward: frequently the “parton jets” do not reproduce exactly the kinematics of the original tagging quark. On the other hand, the fact that the reconstructed jet does not have precisely the same momentum and direction as the tagging quark does not mean that the event will not get through the selection cuts. For examples, for the selection of $qq \rightarrow qqH$ events, besides the $|\eta^{\text{jet}}|$ and p_t^{jet} cuts, a cut on the dijet mass is also used. The M_{jj} distributions of, respectively, tagging quarks, parton jets and reconstructed jets, are shown in figure

³The selection cuts for the Higgs production via the fusion of weak boson pairs will be discussed in detail in chapter 7.

3.3 The simulation frame

3.22. The choice to use, in the analysis, the tagging quarks or the reconstructed jets does not give rise to large differences in the M_{jj} -cut efficiency.

Finally, it has to be pointed out that the goal of this section was not to give a full efficiency study of the jet selection, depending for instance on the jet p_t and η , but rather to introduce the reconstruction algorithm and to investigate the specific case of a jet selection in a $qq \rightarrow qqH$ event.

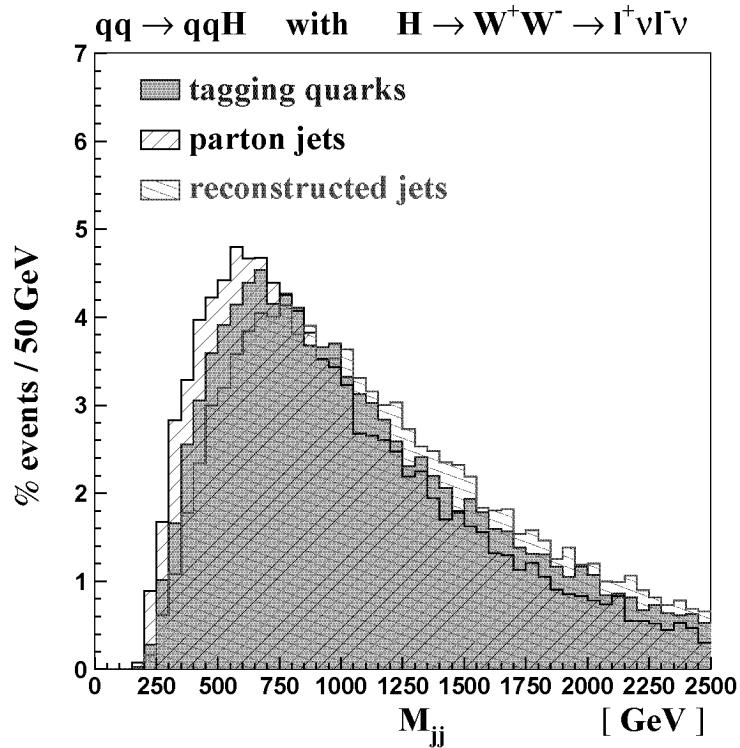


Figure 3.22: $qq \rightarrow qqH \rightarrow qqWW \rightarrow jjl\nu l\nu$: dijet mass M_{jj} reconstructed starting either from 2 tagging quarks, from 2 parton jets or from 2 reconstructed jets. Some simple selection criteria have been used, requiring well separated and isolated leptons and jets.

Chapter 4

A parton luminosity monitor and applications for the diboson SM cross section prediction

Interpretations of essentially all proposed measurements at the LHC require a good knowledge of the parton distribution functions at the relevant Q^2 and the collected integrated luminosity. Both omni-purpose experiments, ATLAS [48] and CMS [14], consider a luminosity accuracy of $\pm 5\%$ as their goal [60].

The aim of this chapter is to show that the electron and muon pseudorapidity distributions, originating from the decay of W^+ , W^- and Z^0 bosons, can constrain the internal momentum distributions of quarks and antiquarks within the protons, at a scale Q^2 of about 10^4 GeV². The fractional momentum x ranges from $\approx 3 \times 10^{-4}$ up to $\approx 10^{-1}$ and covers most of the high- Q^2 LHC processes. One can then relate the rate of l^\pm events from W/Z decays to the quark and antiquark luminosity at the LHC. Furthermore, using the improved information about the quark momentum distributions and the information about the quark luminosity, other $q\bar{q}$ related scattering processes like $q\bar{q} \rightarrow W^+W^-$, $W^\pm Z^0$, $Z^0 Z^0$ should then be predicted more accurately. A good knowledge of these diboson event rates is a first step towards the understanding of the electroweak symmetry breaking mechanism: it enhances the discovery potential for the Standard Model Higgs in the W^+W^- decay channel for $120 \text{ GeV} \leq M_H \leq 500 \text{ GeV}$ and, in case that any light Higgs scalar would have been found, it will be the basis for further studies in the higher mass domain.

4.1 The internal structure of the proton

Nucleons are composed of fundamental spin- $\frac{1}{2}$ fermion constituents, the quarks, and these constituents interact by exchange of spin-1 gluons.

4.1 The internal structure of the proton

As regards to the proton, its structure can be illustrated as follows (see figure 4.1): three valence quarks (two up- and one down-quark), interacting through gluons, constitute the “skeleton” of the nucleon; besides these partons, there are other slow debris consisting of quarks-antiquarks pairs and their associated gluons (the so-called “sea”).

All these components contribute to the creation of a “model” of the proton which takes into account the momentum distributions of the quarks and gluons. More precisely, the structure function $f_i(x_i, Q^2)$ gives, at the scale Q^2 , the probability density to find a parton i with a fraction of the proton momentum $x_i = |\mathbf{p}_i|/E_{\text{beam}}$ (in a frame where this momentum is large). The variable x is also known under the name of “Bjorken x ”.

The knowledge of the parton distribution function (PDF) is of fundamental importance for physics processes at pp -colliders, because it determines, together with the pp -luminosity, the interaction rate between the different quarks and gluons. The cross sections of processes produced at the interaction point are then strongly related to this function.

There is currently no theory which can calculate the parton distribution function and the only solutions today are QCD interpretations of experimental data combined with sum rules. Perturbative QCD does not predict the absolute value of the parton densities within the proton but determines their Q^2 -evolution. For an initial distribution, given at a particular scale Q_0^2 , the so-called Altarelli-Parisi evolution [61] enables the distributions at higher Q^2 to be determined.

These “global fitting” of current high-energy data are performed by different research collaborations. New experimental inputs improve continually the knowledge of the parton densities and new sets replace the older ones. Most of the new sets are next-to-leading order (NLL) evolutions of the Altarelli-Parisi equations. Among the research groups which produce the most up to date sets, we find:

- Martin, Roberts and Stirling (MRS) [62],
- The CTEQ collaboration [63],
- Glück, Reya and Vogt (GRV) [64].

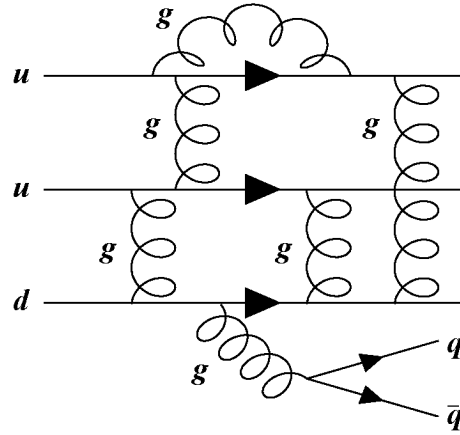


Figure 4.1: Scheme of the internal composition of the proton, where the three valence quarks u, u, d interact via exchange of gluons. In addition, some q, \bar{q} -loops produce other low momentum quarks.

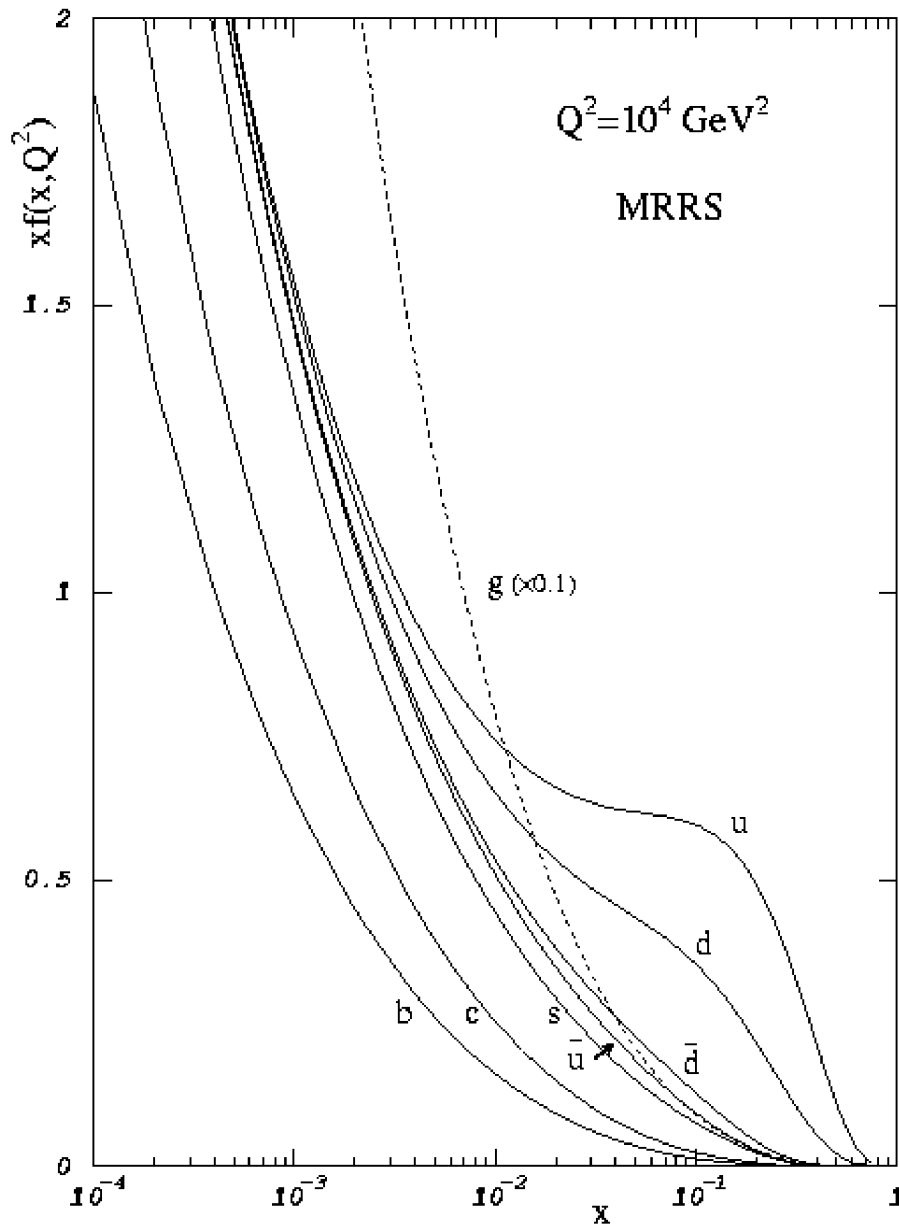


Figure 4.2: Probability densities $f_i(x_i, Q^2)$ for the different partons i to carry fractions x_i of the total proton momentum ($Q^2 \approx m_Z^2$), parameterized by the MRS group [62].

These parton distributions are currently determined from experimental observables in lepton–hadron scattering (DIS processes from fixed target and HERA experiments), Drell–Yan lepton pair production processes (q/\bar{q} -distributions) and γ -jet rates (for the g -contribution) at hadron colliders [65].

The results, obtained at different Q^2 , have then to be extrapolated to the relevant Q^2 scale of the studied process by means of the Altarelli-Parisi evolution. Figure 4.2 gives a quantitative example of the parton distributions for quark and gluons within protons at the scale $Q^2 \approx m_Z^2$. The u and d quarks are mostly valence quarks for large x and mostly sea quarks for small x . Therefore, for $x \gtrsim 10^{-2}$, the difference between u and d comes from the 2:1 ratio of the valence quarks. All the other quarks belong to the proton sea and have mostly smaller x . The figure put also in evidence the large dominance of gluons in these low- x regions.

4.2 Luminosities and event rates at the LHC

The luminosity, together with the cross section, defines the expected event rate N_X :

$$N_X = \sigma_{pp \rightarrow X} \cdot \mathcal{L}_{pp}, \quad (4.1)$$

where $\sigma_{pp \rightarrow X}$ is the cross-section for a given product X , starting from a pp collision, and \mathcal{L}_{pp} is the proton–proton luminosity. This latter is defined as:

$$\mathcal{L}_{pp} = \frac{N^2 k f \gamma}{4\pi \varepsilon_n \beta^*}, \quad (4.2)$$

where N is the number of protons in each of the k circulating bunches, f is the revolution frequency, β^* is the value of the betatron function at the crossing point and ε_n is the emittance corresponding to the 1σ contour of the beam, normalized by multiplying by the Lorentz factor $\gamma = (E/mc^2)$.

In order to determine the pp luminosity at the LHC, preliminary estimates had been made by members of the CERN AT group using the Van der Meer method [66] (*i.e.* size and intensity measurements of the beams at the interaction point). The ATLAS [48] and CMS [14] collaborations have investigated physics processes, to obtain the relative and absolute luminosity. They considered mainly:

- elastic pp -scattering: simultaneous measurement of the differential cross-section $d\sigma/d\theta$ of elastic scatterings at small angles [67]. Such an experiment has been proposed by the TOTEM collaboration [68].
- QED processes like $pp \rightarrow pp e^+e^-$ can be calculated with good precision. The measurement of such reactions has therefore been suggested [69].

Accurate measurements of the above processes, especially at the high luminosity phase of LHC, appear to be very difficult and it is not obvious that the proton–proton luminosity can be measured even with a $\pm 5\%$ accuracy.

To obtain the event cross section at LHC, one has to combine all the possibilities in which a 14 TeV pp -collision, through the combination of the partons i and j , can give rise to the product X :

$$\sigma_{pp \rightarrow X} = \sum_{i,j} \int_{x_i} \int_{x_j} f_i(x_i, Q^2) \cdot f_j(x_j, Q^2) \cdot \sigma_{ij \rightarrow X}(x_i, x_j, \alpha_s(Q^2)) dx_i dx_j \quad (4.3)$$

Therefore, the knowledge of the event rate N_X depends on the knowledge of the parton density functions $f_i(x_i, Q^2)$, of the proton–proton luminosity \mathcal{L}_{pp} and of the point cross section $\sigma_{ij \rightarrow X}(x_i, x_j, \alpha_s(Q^2))$.

While the x distributions of the valence quarks are now quite well constrained, uncertainties for the x distributions of sea quarks and antiquarks and gluons remain more important. As a result of this approximate knowledge of the structure function, total cross section uncertainties for the W^+ , W^- and Z^0 boson production at 14 TeV pp collisions are currently of the order of 5% [70]. Even though the experimental errors are expected to decrease during the next years, cross section uncertainties related to structure functions will remain important. Older estimates of the PDF uncertainties, combined with the unknown contributions from higher order QCD corrections, in the past were usually considered to limit the use of the reaction $pp \rightarrow W^\pm(Z^0)$ as an absolute proton–proton luminosity monitor at very high center of mass energies [71].

Cross-section measurements would therefore be limited to $\pm 5\%$. This is somewhat depressing, when these uncertainties are compared with the possible small statistical errors for many LHC measurements.

Even though the LHC is a proton–proton collider, experiments will mostly study the interactions between their fundamental constituents, the quarks and gluons, at energies where these partons can be considered as quasi free. Ideally [72], it would thus be better to start directly from the parton–parton luminosities $\mathcal{L}_{ij}(\mathbf{p}_i, \mathbf{p}_j) d\mathbf{p}_i d\mathbf{p}_j$, which give the luminosity for a collision between two partons i and j at the center-of-mass energy $\sqrt{s'} = \sqrt{(\mathbf{p}_i + \mathbf{p}_j)^2}$. The rate for a $pp \rightarrow X$ event would then be given by:

$$N_X = \frac{1}{E_{\text{beam}}^2} \sum_{i,j} \int_{\mathbf{p}_i} \int_{\mathbf{p}_j} \mathcal{L}_{ij}(\mathbf{p}_i, \mathbf{p}_j) \cdot \sigma_{ij \rightarrow X}(\mathbf{p}_i, \mathbf{p}_j, \alpha_s(Q^2)) d\mathbf{p}_i d\mathbf{p}_j, \quad (4.4)$$

where one has to integrate over the possible partons i, j and their momenta $\mathbf{p}_{i,j}$. Assuming a head-on collision between the two partons (*i.e.* $\mathbf{p}_{i,j} \parallel \mathbf{p}_{\text{beam}}$, which is well justified for 7 TeV protons), $\mathcal{L}_{ij}(\mathbf{p}_i, \mathbf{p}_j)$ can be obtained from \mathcal{L}_{pp} through:

$$\mathcal{L}_{ij}(\mathbf{p}_i, \mathbf{p}_j) d\mathbf{p}_i d\mathbf{p}_j \approx E_{\text{beam}}^2 \cdot \mathcal{L}_{pp} \cdot f_i(x_i, Q^2) \cdot f_j(x_j, Q^2) dx_i dx_j, \quad (4.5)$$

4.3 Rates of $pp \rightarrow W^+, W^-$ and Z^0 events

with $\sqrt{s'} \approx \sqrt{s \cdot x_i \cdot x_j}$ [$s = (14 \text{ TeV})^2$].

Clearly, if the parton-parton luminosities $\mathcal{L}_{ij}(\mathbf{p}_i, \mathbf{p}_j)$ have to be extracted using equation 4.5, they would keep exactly the same uncertainties related to the PDF's and to \mathcal{L}_{pp} as before. The challenge is then to find a method, which allows a direct measurement of the different $\mathcal{L}_{ij}(\mathbf{p}_i, \mathbf{p}_j)$.

In the following sections, we want to show that the rapidity distributions of the weak bosons are directly related to the fractional momenta x of the quarks and antiquarks. Consequently, the observable pseudorapidity distributions of the charged leptons from the decays of W^\pm and Z^0 bosons are also related to the x distributions of quarks and antiquarks. Therefore, the shape of the lepton η -distributions provide actually the key to precisely constrain the quark and antiquark structure functions, while the lepton rates would then allow a monitoring of the corresponding q/\bar{q} -luminosities.

In addition to the previous listed methods to measure the LHC pp -luminosity and the parton distribution functions, there is thus a promising alternative:

- the production of W and Z bosons can be used to measure the q -, \bar{q} -density functions and the effective q -, \bar{q} -luminosities [12].

A similar strategy can be imagined also for gluons:

- gluon dominated scattering processes, as for example $gq \rightarrow \gamma(Z^0)q$, can constrain the g -density function and the corresponding luminosity [13, 73].

Moreover, a recent paper [74] proposed a equivalent approach for measuring specifically the density functions and luminosity of heavy quarks.

The importance of such measurements is enormous: in order to find the Higgs or other new physics phenomena with small signal to background ratios, the event rate of the background reactions should be known as accurately as possible. With these methods, uncertainties of perhaps 1% can be imagined.

4.3 Rates of $pp \rightarrow W^+, W^-$ and Z^0 events

The expected initial pp -luminosity at LHC of $10^{33} \text{ cm}^{-2} \text{ s}^{-1}$ will result in an integrated luminosity of about 10 fb^{-1} per year. It is worth pointing out, that already at this initial luminosity the event rate of weak bosons is huge: about $10^6 W^\pm \rightarrow \ell^\pm \nu$ and about $500 W^+W^- \rightarrow \ell^+ \nu \ell^- \bar{\nu}$ per day are expected¹! The non-resonant W^+W^- production is especially important as it is the dominant background for the Standard Model Higgs search in the $H \rightarrow W^+W^-$ decay channel [47].

The production of $pp \rightarrow W^+, W^-$ and Z^0 and their identification using the leptonic decays have been discussed extensively in the literature [75]. Assuming collisions of essentially free partons, the production of weak bosons, $u\bar{d} \rightarrow W^+ \rightarrow \ell^+ \nu$,

¹The about 10^5 seconds in a day result in $\mathcal{L}_{\text{day}} \approx 100 \text{ pb}^{-1}$.

Process	$\sigma \times BR$			
	MRS(R1)	CTEQ 4L	CTEQ 4M	GRV 94HO
$q_i \bar{q}_j \rightarrow W^+ \rightarrow \ell^+ \nu$	19.65 nb	20.21 nb	21.05 nb	21.55 nb
$q_i \bar{q}_j \rightarrow W^- \rightarrow \ell^- \bar{\nu}$	14.11 nb	14.43 nb	15.42 nb	15.27 nb
$q_i \bar{q}_i \rightarrow Z^0 \rightarrow \ell^+ \ell^-$	3.188 nb	3.279 nb	3.479 nb	3.481 nb
$q_i \bar{q}_i \rightarrow W^+ W^- \rightarrow \ell^+ \nu \ell^- \bar{\nu}$	3.378 pb	3.323 pb	3.677 pb	3.606 pb
$q_i \bar{q}_j \rightarrow W^\pm Z^0 \rightarrow \ell^\pm \nu \ell^+ \ell^-$	0.4027 pb	0.3908 pb	0.4347 pb	0.428 pb
$q_i \bar{q}_i \rightarrow Z^0 Z^0 \rightarrow \ell^+ \ell^- \ell^+ \ell^-$	46.8 fb	45.93 fb	51.15 fb	49.76 fb

Table 4.1: Examples of estimated weak boson production cross sections at the LHC for three different sets of structure functions using PDFLIB and PYTHIA programs [76, 19]. In all cases, the leptonic branching ratios into electrons and muons are included.

$d\bar{u} \rightarrow W^- \rightarrow \ell^- \bar{\nu}$ and $u\bar{u}(d\bar{d}) \rightarrow Z^0 \rightarrow \ell^+ \ell^-$ are in lowest order understood to at least a percent level. Moreover it is a well known fact that the W^\pm and Z^0 production rates at the LHC, including their leptonic branching ratios into electrons and muons, are huge and provide clean and well measurable signals with isolated, high- p_t electrons or muons. These signature are therefore often considered as a clean and excellent calibration tool at the LHC [77]. However, following the arguments of the previous section, past studies concluded that their use as a pp -luminosity monitor was limited to relative luminosity measurements only [71]. Using different sets of structure functions [76], today's cross section variations are about 5-10%, as can be seen from table 4.1. The cross section predictions as well as the following simulation results are obtained using the PYTHIA Monte Carlo program [19].

These cross section variations for single W^\pm , Z^0 production are strongly correlated with the cross section predictions for other $q_i \bar{q}_j$ related processes. As an example, the corresponding cross sections for the reactions $q_i \bar{q}_j \rightarrow W^+ W^-$, $W^\pm Z^0$, $Z^0 Z^0$ are also given in table 4.1. Thus, the uncertainties for multi boson production cross sections at the LHC are already reduced to about 3-5% if event rates are estimated relative to the production rates of single W^\pm , Z^0 events:

$$\frac{N_{VV}}{N_V} = \frac{\mathcal{L}_{pp} \otimes PDF(x_k, x_l, Q^2) \otimes \sigma_{kl \rightarrow VV}}{\mathcal{L}_{pp} \otimes PDF(x_i, x_j, Q^2) \otimes \sigma_{ij \rightarrow V}}, \quad (4.6)$$

where V indicates a weak boson. Furthermore, such relative measurements reduce also errors arising from uncertainties in branching ratios, detection efficiency, while systematic errors from $\Delta \mathcal{L}_{pp} / \mathcal{L}_{pp}$ completely disappear as the pp -luminosities cancel. We expect also a cancellation of some errors given by higher order corrections, factorization, scale dependence and choice of α_s . Therefore,

4.4 Using the W^\pm and Z^0 production to constrain the q and \bar{q} structure functions

by taking the ratios between $q_i\bar{q}_j$ -related processes, besides to the experimental accuracy by the measurement of the "monitoring" $q_i\bar{q}_j \rightarrow W, Z$ process, we are now left mainly with part of the uncertainties related to the PDF (*i.e.* about its x -evolution, its Q^2 -evolution and about the relations between the different parton contributions) and, in addition, with a smaller error given by the cross section calculations.

4.4 Using the W^\pm and Z^0 production to constrain the q and \bar{q} structure functions

At the LHC, in contrast to proton-antiproton colliders, the antiquarks have to come from the sea. The feynmann diagrams related to the W and Z production at LHC are shown in figure 4.3. The production of a W^+ proceeds mainly through the annihilation of an u -quark and a \bar{d} -quark. A W^- instead arises normally from the combination of a d -quark with a \bar{u} -quark, while Z^0 bosons are produced largely through $u\bar{u}$ and $d\bar{d}$ annihilations.

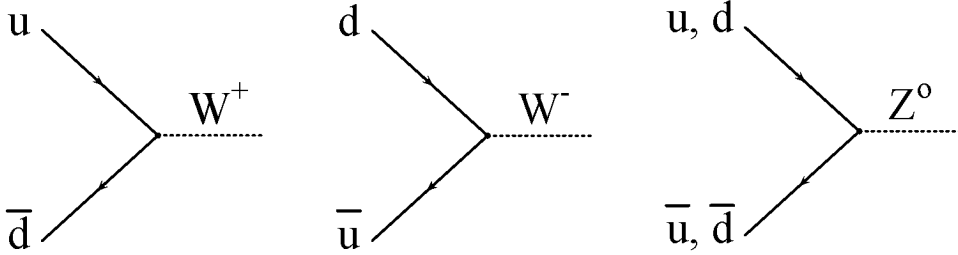


Figure 4.3: Main production modes for W and Z bosons at LHC.

With the known W^\pm and Z^0 masses and by using the well justified approximations for the quark momenta $m_q \approx 0 \Rightarrow E_i \approx |\mathbf{p}_i|$, one obtains:

$$m_{W,Z}^2 \approx 2|\mathbf{p}_1||\mathbf{p}_2| - 2|\mathbf{p}_1||\mathbf{p}_2|\cos\vartheta, \quad (4.7)$$

where ϑ is the angle between the incoming quarks. In a head-on collision between 7 TeV protons, this angle is $\approx \pi$. Consequently, the known weak bosons masses are related to the $x_i = |\mathbf{p}_i|/E_{\text{beam}}$ according to:

$$m_{W,Z}^2 \approx s \cdot x_1 \cdot x_2 \quad (4.8)$$

with $s = 4E_{\text{beam}}^2$. The product $x_1 \cdot x_2$ for quarks-antiquarks at LHC is therefore constrained by equation 4.8: $\approx 3.3 \times 10^{-5}$ for W and $\approx 4.2 \times 10^{-5}$ for Z bosons. The mean fractional momentum is thus $\bar{x}_i \approx 6 \times 10^{-3}$.

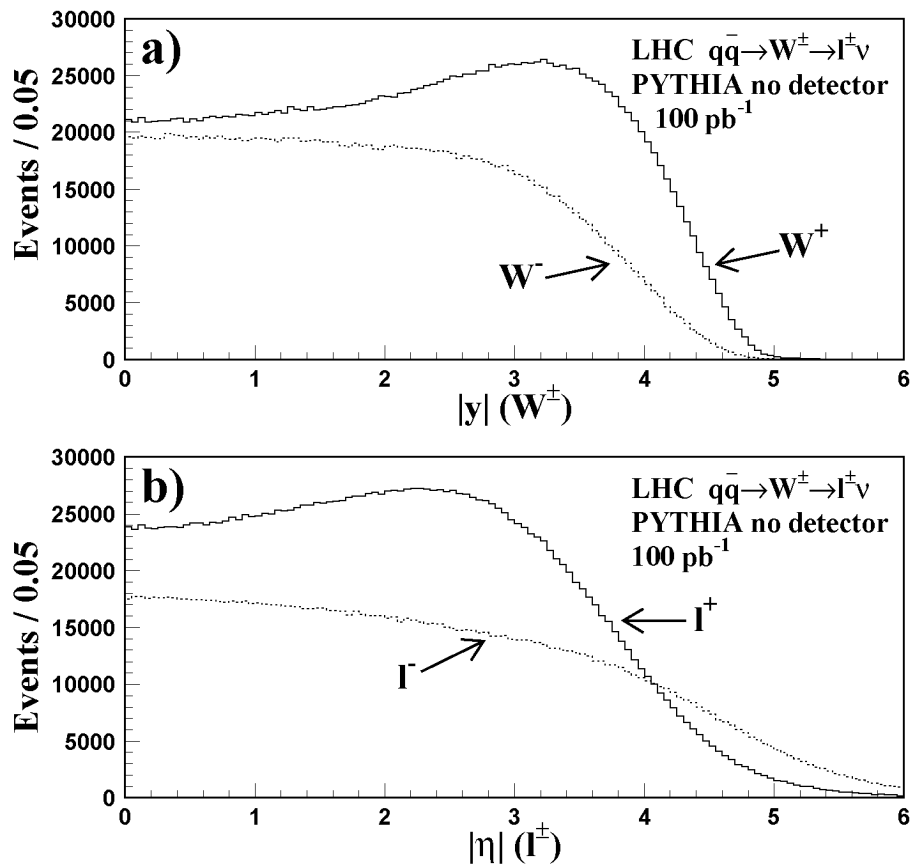


Figure 4.4: Expected rapidity (pseudorapidity) distribution of W^\pm (a) and ℓ^\pm (b), originating from the reaction $q\bar{q} \rightarrow W^\pm \rightarrow \ell^\pm \nu$ at the LHC ($\sqrt{s} = 14 \text{ TeV}$ and according to the MRS(A) structure function [76]). The assumed luminosity of 100 pb^{-1} corresponds to about one day of data taking with a luminosity $\mathcal{L}_{pp} = 10^{33} \text{ s}^{-1} \text{ cm}^{-2}$.

Figure 4.4a shows the expected rapidity distributions of W^+ and W^- , which directly reflect the difference between the x distributions of the u , d valence quarks and the sea quark or antiquarks. Using the definition of rapidity (equation 3.3) and with the help of:

$$\begin{cases} E_{W,Z} \approx (x_1 + x_2) \cdot E_{\text{beam}} \\ (p_z)_{W,Z} \approx (x_1 - x_2) \cdot E_{\text{beam}} \end{cases}$$

one obtains:

$$y = \frac{1}{2} \ln \frac{x_1}{x_2}, \quad (4.9)$$

4.4 Using the W^\pm and Z^0 production to constrain the q and \bar{q} structure functions

which connects the fractional momenta x of the involved quarks with the rapidity of the produced boson. Therefore we have access to a system of two equations (4.8 and 4.9) for two unknowns x_1 and x_2 and we are thus able to precisely reconstruct both fractional momentum values. What is still missing is the knowledge of the type of quark related to each x_i . However, also this information can be partially inferred.

For example, small W^\pm rapidities mean that the boson was produced almost at rest, *i.e.* both incoming quarks had about the same momenta. In this case, we find $x_{1,2}$ values of $\approx 6 \times 10^{-3}$. Therefore, for these η -regions the most likely combination corresponds to W^\pm which originate from the annihilation of sea quark–antiquarks (see figure 4.2) and thus only small differences between W^+ and W^- are expected.

On the contrary, for larger rapidities, the W^\pm 's originate normally from the annihilation of quarks and antiquarks with very different x values. In fact, to produce a W^\pm at a rapidity of about 2.5, one finds from equations 4.8 and 4.9 the corresponding $x_{1,2}$ values to be, respectively, $x_1 \approx 0.1$ and $x_2 \approx 3 \times 10^{-4}$. Following the probability densities shown in figure 4.2, we can then argue that the first one would be more likely a valence quark and the second one a sea antiquark. Given the 2:1 ratio between valence quarks and according to the main production modes $u\bar{d} \rightarrow W^+$, $d\bar{u} \rightarrow W^-$, at large rapidities a positive

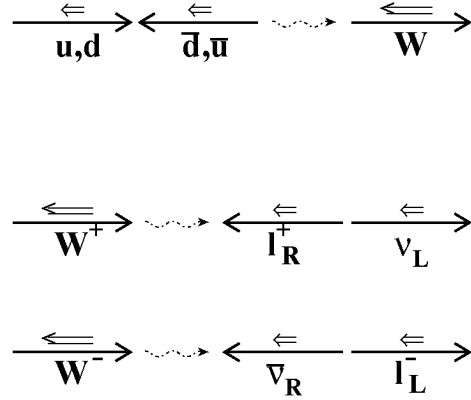


Figure 4.5: Spin conservation rules in the W production and decay.

W^+ boson is then much more likely produced than a negative W^- one.

Figure 4.4b shows the pseudorapidity distributions of the charged leptons originating from the W^\pm decays. Because of the V–A interaction, the differences between the pseudorapidity distributions of ℓ^+ and ℓ^- , especially at large η values, are larger than the ones for the W^+ and W^- . The reason is that the left handed lepton (ℓ^-) is emitted preferentially in the direction of the incoming quark and the right handed anti-lepton (ℓ^+) is emitted opposite to the quark direction. These spin combinations are illustrated in figure 4.5.

Thus, to summarize, the observable charged lepton pseudorapidities reflect not only the x distributions of quarks and antiquarks but allow also to some extent a distinction between valence and sea quarks at a given x .

Following the main W, Z -production modes shown in figure 4.3, the pseudora-

pseudorapidity distribution of the positive charged leptons, $u\bar{d} \rightarrow W^+ \rightarrow \ell^+\nu$ contains mostly the information about the sea \bar{d} -quarks and the valence or sea u -quarks:

$$\sigma_{W^+}[\eta] \sim f_u(x, Q^2 = m_W^2) \otimes f_{\bar{d}}(x, Q^2 = m_W^2).$$

The negative charged leptons, $d\bar{u} \rightarrow W^- \rightarrow \ell^-\bar{\nu}$ carry consequently the information about the sea \bar{u} quarks and the valence or sea d quarks:

$$\sigma_{W^-}[\eta] \sim f_d(x, Q^2 = m_W^2) \otimes f_{\bar{u}}(x, Q^2 = m_W^2).$$

The predominance of these modes derives directly from the probability densities of these partons within protons. As shown on figure 4.2, the high- x region is dominated by the presence of two valence u -quarks and one valence d -quark. Moreover, within $q\bar{q}$ -loops, lighter quarks arise more likely than the heavier ones (c and b quarks have thus smaller probabilities than u, d and s at the scale $Q^2 = m_{W,Z}^2$). The presence of other quarks is not only disfavored by the smaller density functions, but is also a consequence of the ‘‘charged’’ weak couplings to quarks, which are dominated by exchanges between same generations. For example, despite the similar x -distribution, the production mode $d\bar{s} \rightarrow W^-$ is suppressed against the $d\bar{u} \rightarrow W^-$ one, which proceeds via the annihilation of two quarks belonging to the first generation. The coupling $s\bar{c} \rightarrow W^-$ is instead large, but it suffers from the lower parton densities of the s - and c -quarks at large x fractional momentum.

These arguments are also valid for the Z^0 production, where in addition the mixing between generations is totally excluded. The rapidity distribution of charged lepton pairs, from $Z^0, (Z^*, \gamma^*) \rightarrow \ell^+\ell^-$, provide thus the information about the sum of sea $\bar{u}+\bar{d}$ antiquarks and the sum of valence and sea $u+d$ quarks:

$$\sigma_{Z^0}[\eta] \sim \begin{cases} f_u(x, Q^2 = m_Z^2) \otimes f_{\bar{u}}(x, Q^2 = m_Z^2) \\ f_d(x, Q^2 = m_Z^2) \otimes f_{\bar{d}}(x, Q^2 = m_Z^2) \end{cases}$$

Consequently, the combination of the different observable lepton pseudorapidity distributions should provide some sensitivity to the u, d, \bar{u} and \bar{d} parton densities over a large x range.

As equations 4.8 and 4.9 suggest, the different parameterization for the parton momentum distributions provided by the different groups should appear then in a different kinematical distribution of the W, Z decay products. We study therefore here the supposed effects of these parameterizations on the measured ℓ^\pm pseudorapidity distributions ($q\bar{q} \rightarrow W^\pm \rightarrow \ell^\pm\nu$), and on the reconstructed Z^0 rapidity distribution ($q\bar{q} \rightarrow Z^0 \rightarrow \ell^+\ell^-$). For this purpose simple event selection criteria are used (for details see [78]). These criteria closely follow the design

4.4 Using the W^\pm and Z^0 production to constrain the q and \bar{q} structure functions

characteristics of the CMS experiment [14].

Leptons are selected according to:

- Electrons and Muons are required to be isolated, to have $p_t > 30$ GeV and to lie within a pseudorapidity of $|\eta| < 2.4^2$.
- To reduce possible backgrounds from heavy quark decays and to reject high- p_t W^\pm and Z^0 production due to initial state radiation, events with reconstructed jets with $E_t > 20$ GeV are removed.

Using these charged lepton selection criteria, $pp \rightarrow W^\pm \rightarrow \ell^\pm \nu$ events are required to have exactly one isolated charged lepton with $30 < p_t < 50$ GeV. The resulting p_t spectra of ℓ^\pm and their pseudorapidity distributions are shown in figure 4.6a and 4.6b respectively. The used kinematic and geometric event selection criteria, result in an event detection efficiency of about 25% for $W^+ \rightarrow \ell^+ \nu$, and about 28% for $W^- \rightarrow \ell^- \bar{\nu}$.

To select events of the type $pp \rightarrow Z^0 \rightarrow \ell^+ \ell^-$, we require the presence of a pair of isolated leptons with opposite charge ($e^+ e^-$ or $\mu^+ \mu^-$). The $\ell^+ \ell^-$ invariant mass should be consistent with the Z^0 mass: $m_{\ell^+ \ell^-} = m_{Z^0} \pm 2$ GeV. In addition, the opening angle between the two leptons in the plane transverse to the beam has to be larger than 135° and $p_t(Z^0) < 20$ GeV. Using the above kinematical and geometrical event selection criteria, the efficiency to detect both leptons from Z^0 decays is about 16%, and increases to about 23% for dilepton masses in the range between 150–200 GeV. As these dilepton events are usually considered to be background free, Z^0 events with large p_t can be used to constrain the x distribution of gluons, as will be discussed later.

Events of the type $q\bar{q} \rightarrow (Z^*, \gamma^*) \rightarrow \ell^+ \ell^-$ with dilepton masses above 100 GeV have a much lower rate. However, these events can be used to study the Q^2 evolution, up to masses where the neutral current sector is well understood (e.g. up to masses of about 200 GeV). At least up to these dilepton masses, a measurement of the lepton forward backward charge asymmetry, following the method of reference [79], constrains the ratio of valence and sea u and d quarks at different x values and different Q^2 . Using both informations about the x -distribution and their Q^2 evolution, it should be possible to obtain precise predictions for much higher dilepton masses.

One can now ask how well q and \bar{q} structure function can be constrained from the observable weak boson rapidities. For this purpose the different ℓ^\pm cross sections are studied relative to a pair of reference structure functions, arbitrarily chosen to be the MRS(A) and CTEQ 4M sets.

²Despite the interest in the very forward region, lepton detection up to much larger $|\eta|$ values appears to be very difficult for the CMS and ATLAS detectors. Perhaps LHC-b could cover also these regions, allowing a larger sensitivity in the low- x domain.

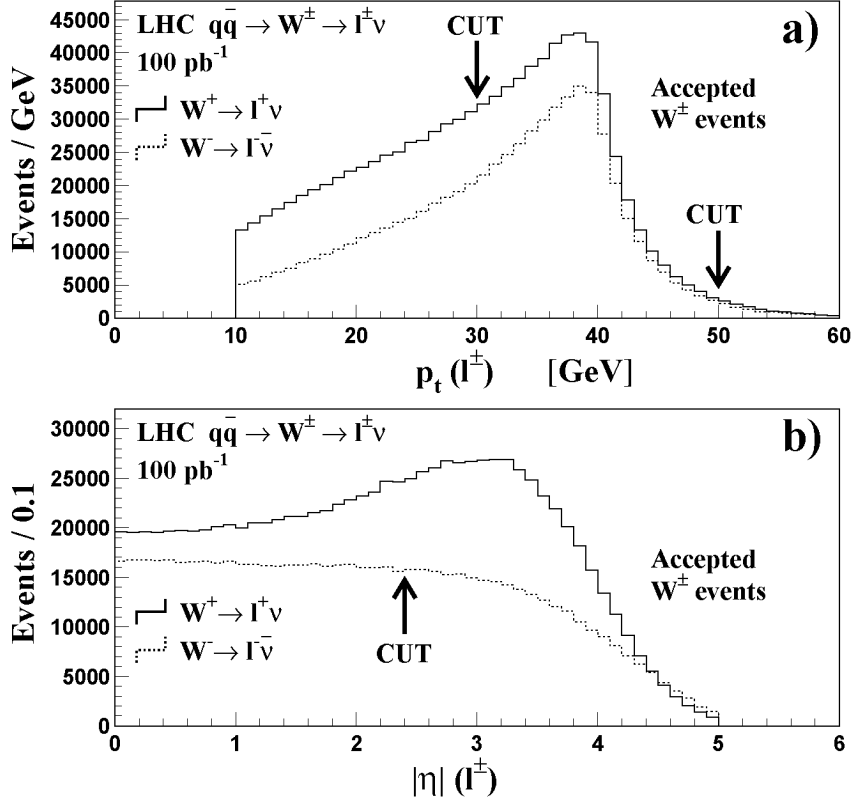


Figure 4.6: The observable (a) charged lepton p_t and (b) pseudorapidity η distributions originating from the reaction $q\bar{q} \rightarrow W^\pm \rightarrow \ell^\pm \nu$ at the LHC ($\sqrt{s} = 14$ TeV and according to the MRS(A) structure function [76]), including all selection criteria discussed in the text.

This sensitivity is first investigated by comparing the weak boson production, using two quite similar structure functions: MRS(A) and MRS(H) [76]. The main difference between these two sets lies in the x parameterization for the light sea quarks. While the older MRS(H) set uses u, d flavor symmetric sea distributions, the MRS(A) set includes a fine tuning of the sea quark parton densities with some isospin symmetry breaking, required to describe the observation of Drell-Yan asymmetries of $A_{DY} = (\sigma_{pp} - \sigma_{pn}) / (\sigma_{pp} + \sigma_{pn})$ from the NA 51 experiment [80].

Figure 4.7a shows the ratio of $\sigma(W^+) / \sigma(W^-)$ as a function of the charged lepton pseudorapidity for the two structure function sets. The different parameterizations thus lead, depending on the lepton pseudorapidity, to a cross section

4.4 Using the W^\pm and Z^0 production to constrain the q and \bar{q} structure functions

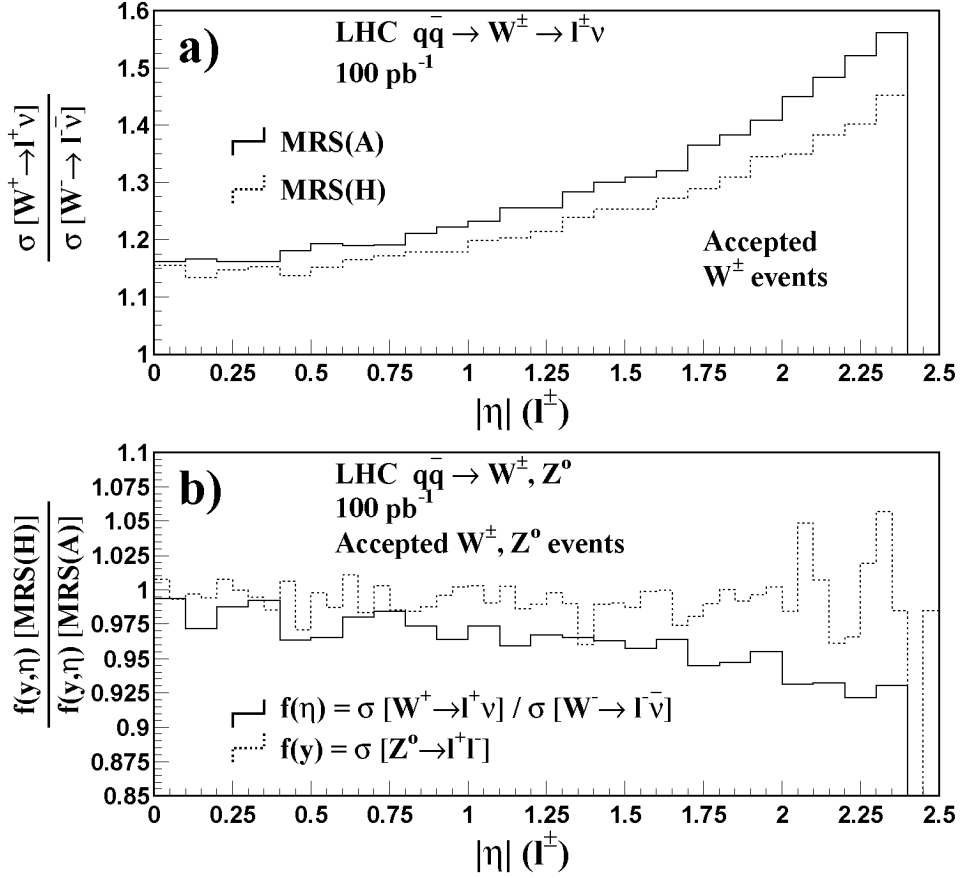


Figure 4.7: a) The detected charged lepton cross section ratio, $\sigma(\ell^+\nu)/\sigma(\ell^-\bar{\nu})$, originating from the reaction $q\bar{q} \rightarrow W^\pm \rightarrow \ell^\pm\nu$ as a function of the lepton pseudorapidity for the MRS(H) and MRS(A) structure function parameterization. b) The relative changes of the charged lepton ratios of 4.7a) between the MRS(H) and MRS(A) parameterizations and also the cross section ratio of the Z^0 production using both parameterizations.

variation of up to about 10%. The double ratio

$$\frac{(\sigma_{W^+}/\sigma_{W^-})^{\text{MRS(H)}}}{(\sigma_{W^+}/\sigma_{W^-})^{\text{MRS(A)}}$$

is shown in figure 4.7b. The differences of about 5-10% between the two sets should be compared with the statistical precision, which is already smaller than 1% per bin for an integrated luminosity of only 100 pb^{-1} . Furthermore, both sets of structure functions predict almost identical Z^0 cross sections. The ratio between the Z cross sections from the two sets, $\sigma_{Z^0}^{\text{MRS(H)}}/\sigma_{Z^0}^{\text{MRS(A)}}$, is also shown in figure 4.7b. Following these arguments, by combining the obtainable information from W^+ , W^- and Z^0 production, the “fine tuned” isospin splitting of u and d sea quarks between MRS(A) and MRS(H) should be detectable with good accuracy.

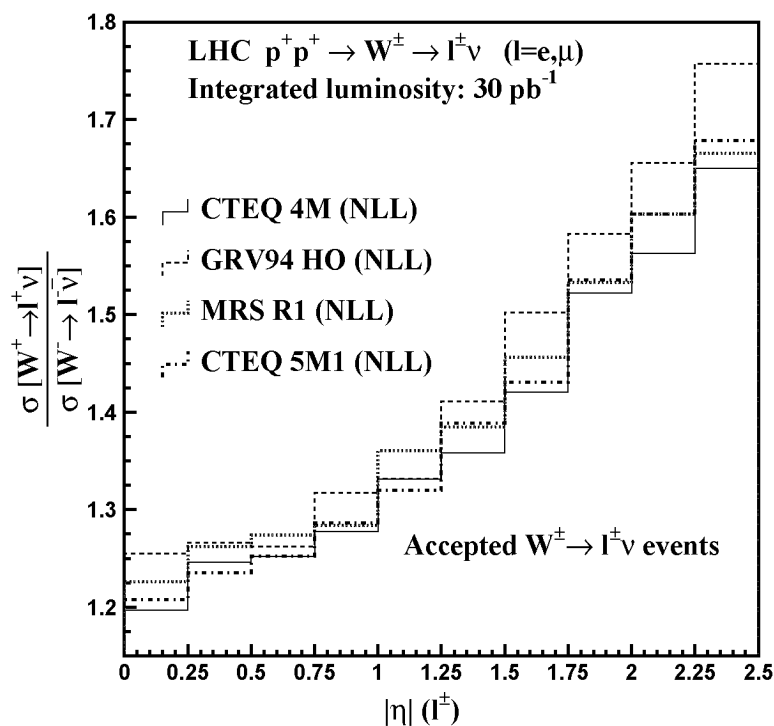


Figure 4.8: The ratio of the accepted cross sections $\sigma(u\bar{d} \rightarrow W^+ \rightarrow \ell^+\nu)$ and $\sigma(d\bar{u} \rightarrow W^- \rightarrow \ell^-\bar{\nu})$ as a function of the lepton pseudorapidity for four different structure functions [76].

4.4 Using the W^\pm and Z^0 production to constrain the q and \bar{q} structure functions

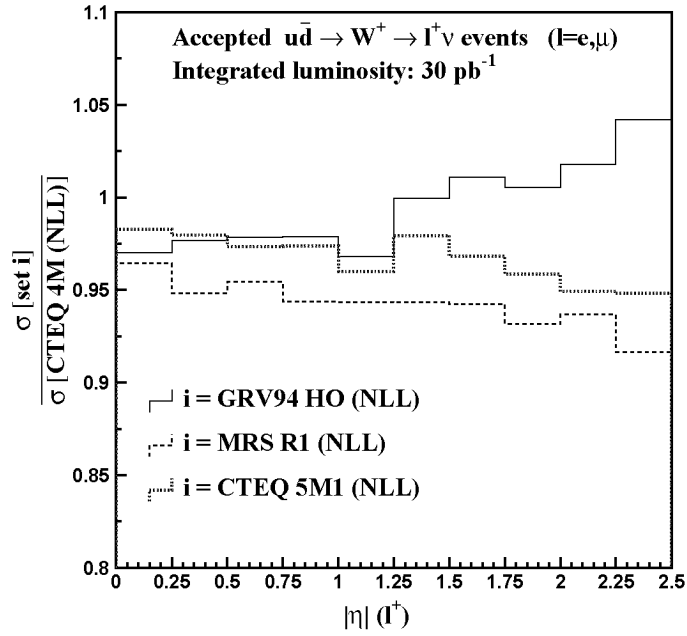


Figure 4.9: Rapidity dependence of the l^+ cross section predictions from different sets of NLL structure functions relative to the one obtained from the CTEQ 4M parameterization.

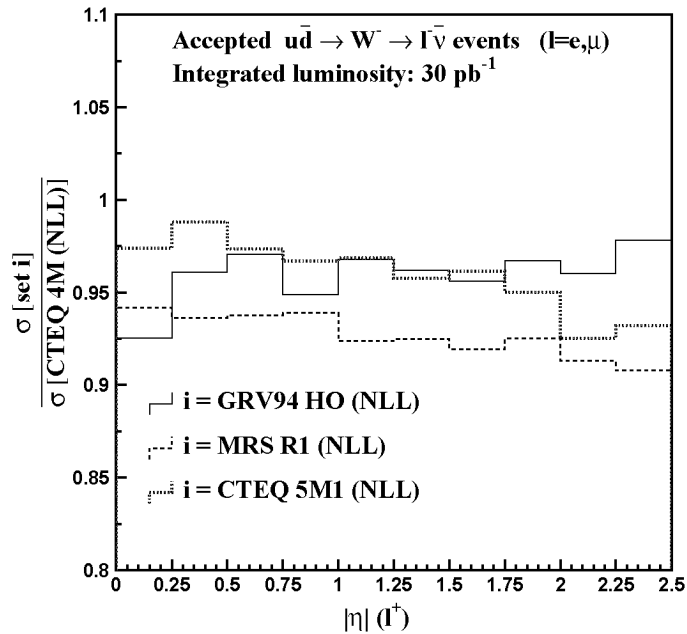


Figure 4.10: Same as figure 4.9, but for the $W^- \rightarrow l^-\bar{\nu}$ production.

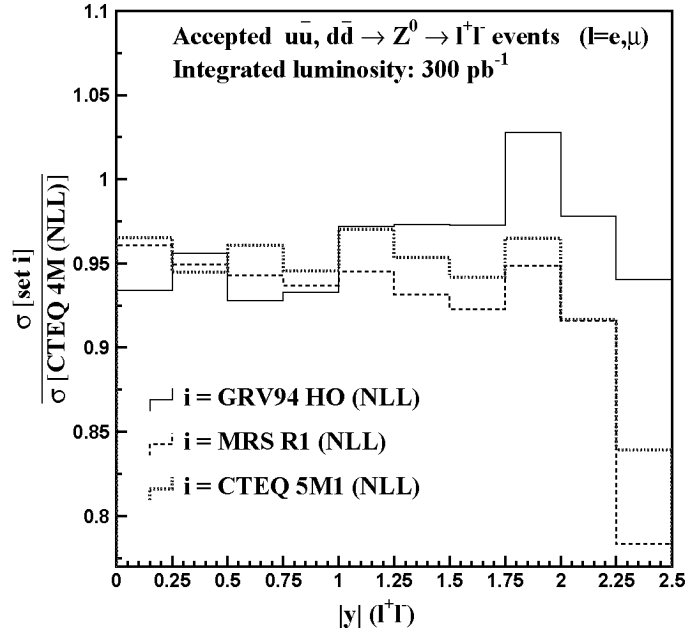


Figure 4.11: Rapidity dependence of the reconstructed Z^0 cross section. The predictions from different sets of structure functions are used and plotted relative to the CTEQ 4M parameterization.

Through this example, we have shown that the weak boson rapidity distributions are sensitive even to small differences between the x distribution of u and d sea quarks and antiquarks.

The fraction of weak bosons which are produced from the annihilation of valence quarks and low x antiquarks increases strongly with increasing rapidity. The valence quark x distribution is already quite well constrained. The main difference between the various structure functions comes especially from q and \bar{q} parameterizations at low x (sea). Thus precise measurements of the charged lepton pseudorapidity distributions from W^\pm decays and the rapidity distribution of the Z^0 events constrain the low x domain of \bar{u} and \bar{d} .

The sensitivity of the measurable lepton rapidity distribution to the different sets of structure functions is shown in figures 4.8-4.11. Figure 4.8 shows the observable ratio of the ℓ^+ to ℓ^- event rates for four different sets of structure functions and an integrated luminosity of 30 pb^{-1} . The difference between the various low x sea quark parameterizations are thus reflected in the observable lepton pseudorapidities. Consequently, the shape of the ℓ^\pm pseudorapidity distributions provide a strong constraint on the underlying x distribution of quarks and antiquarks with x between $\approx 3 \times 10^{-4}$ and $\approx 10^{-1}$.

4.5 Gluon and heavy quarks distribution functions

Figures 4.9-4.11 show the ratio of the predicted ℓ^\pm cross sections from different NLL structure functions relative to the reference CTEQ 4M set. The statistical fluctuations shown in figures 4.9 and 4.10 correspond to the errors from roughly 10 hours of data taking at the initial LHC luminosity of $10^{33}\text{s}^{-1}\text{cm}^{-2}$. The Z^0 event rates are roughly a factor of 10 smaller and the errors shown in figure 4.11 correspond to about 3 days of data taking.

As expected, at large pseudorapidities the uncertainties by small x -values produce larger deviations among the different PDF sets.

The sensitivity of the η_{ℓ^\pm} distributions (originating from weak boson decays) to details of the q, \bar{q} density functions can be translated into a direct measurement of the q, \bar{q} luminosities. The rate of ℓ^\pm events, in a selected pseudorapidity interval $\Delta\eta$, can be actually related to the quark and antiquark luminosity at the given x , allowing also to some extent to distinguish between the contribution of the sea q, \bar{q} and of the valence u - and d -quarks. Obviously, once the shape of the pseudorapidity distribution is accurately known, the ℓ^\pm event rates need to be measured only for a small pseudorapidity interval. For example, counting of ℓ^\pm events from the process $pp \rightarrow W^\pm \rightarrow \ell^\pm \nu$ could be restricted to the pseudorapidity range of $|\eta| < 0.5$. Including all selection criteria one would observe roughly 150 000 “clean” luminosity events, corresponding to a statistical error of 0.3% per day at the initial LHC luminosity ($\approx 100 \text{ pb}^{-1}/\text{day}$), which is already much smaller than the optimistic systematic errors of perhaps $\pm 1\%$.

4.5 Gluon and heavy quarks distribution functions

Up to now, we have only discussed the ways in which the q, \bar{q} distribution function can be extracted from experimental observables. The gluon x distribution and the corresponding gluon luminosity could also be constrained in a similar way, *i.e.* by measuring the rapidity distribution of gluon dominated scattering processes.

In fact, as the q, \bar{q} luminosity can accurately be measured from the weak boson rapidity distribution, the rapidity distribution of these gluon related processes has only to be measured relative to the weak boson y -distributions. Once the gluon density function is known relative to the density function of quarks, the weak boson rate also provides the luminosity monitor for gluon related signal and background processes.

The possible experimental accuracy thus depends mainly on how accurate the rapidity and Q^2 distributions of gluon related scattering processes can be measured.

Possible gluon related scattering processes are $gg \rightarrow X$ and $gq(\bar{q}) \rightarrow X$. As these processes involve jets, measurement problems should be minimized by us-

Process	$50 < p_t < 100$ [GeV]	$100 < p_t < 200$ [GeV]	$p_t > 200$ [GeV]
$q\bar{q} \rightarrow Z^0(\rightarrow \ell^+\ell^-)g$	36.4 pb	6.01 pb	0.71 pb
$qg \rightarrow Z^0(\rightarrow \ell^+\ell^-)q$	150 pb	34.8 pb	4.08 pb
$qq \rightarrow \gamma g$	717 pb	74.5 pb	7.45 pb
$qg \rightarrow \gamma q$	6590 pb	615 pb	49.3 pb

Table 4.2: PYTHIA cross section estimates for high- p_t final states of the type $Z^0(\rightarrow \ell^+\ell^-)q(g)$ and $\gamma q(g)$ (according to the MRS(A) set) [13, 73].

ing processes with small backgrounds and well measurable p_t . Candidates for such processes are high- p_t events with one or more jets and an isolated γ or a $Z^0(\rightarrow \ell^+\ell^-)$. As the energy and momentum of isolated photons and leptons can be measured very accurately, the p_t of the jets, assuming transverse momentum conservation, can also be determined. Thus, the observables are well measured and should provide accurate Q^2 measurements.

The production of events with isolated high p_t photons or Z^0 are dominated by $gq \rightarrow \gamma(Z^0)q$, and $q\bar{q} \rightarrow \gamma(Z^0)g$. As shown in table 4.2, the expected cross sections for these reactions, including the branching ratios $Z^0 \rightarrow \ell^+\ell^-$, and relatively high p_t of γ and Z^0 are still quite large. Furthermore, the calculable background corrections from the process $q\bar{q}$ are expected to be small as the cross sections are dominated by the qg scattering process.

Previous studies of γ -jet final states have shown that jet events with isolated π^0 's provide a considerable background [81]. This large background will therefore limit the achievable accuracy of such a final state. However, the leptonic Z^0 decays provide an excellent signature and should allow the selection of essentially background free Z^0 -jet events.

Unfortunately, the accurate Q^2 determination due to inherent uncertainties of jet energy measurements especially at large rapidities will probably limit the interpretation of the observable rapidity distribution with respect to the gluon x -distribution. Nevertheless, such direct measurements of the gluon structure function will provide the highest possible accuracy for the x distribution of gluons and might eventually lead to cross section predictions with perhaps 1-2% accuracies for other gluon related scattering processes.

Finally, a recent publication [74] proposed to measure specifically the parton distribution functions and luminosities of charm, beauty and strange quarks and antiquarks, using events with an isolated high- p_t photon or lepton and exactly one jet which contains an inclusive muon. The studied processes are $cg \rightarrow \gamma c$, $bg \rightarrow \gamma b$ and $sg \rightarrow Wc \rightarrow (l^\pm\nu)\mu X$, which constrain the corresponding x in the range $\approx 5 \times 10^{-4} - 10^{-1}$ and at a $Q^2 > (100 \text{ GeV})^2$. Systematic uncertain-

4.6 Precise prediction of diboson events

ties from the charm fragmentation function and the semi-leptonic charm decay branching ratios will limit the possible accuracy to perhaps $\pm 5\text{--}10\%$, which represents nevertheless a drastical improvement in comparison with any pre LHC measurement.

4.6 Precise prediction of diboson events

Since the production of weak boson pairs is the major focus of interest of this work, we concentrate our analysis on the reactions $q\bar{q} \rightarrow W^+W^-$, $W^\pm Z^0$ and $Z^0 Z^0$.

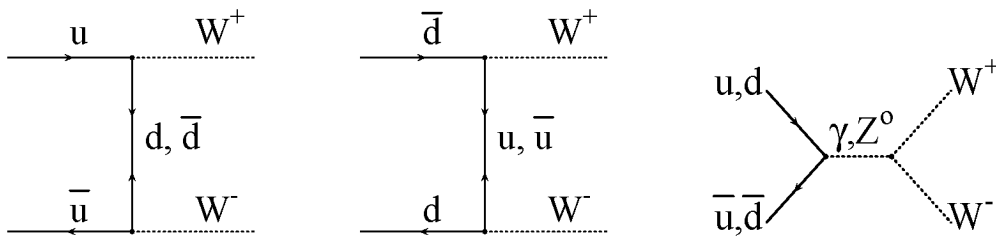


Figure 4.12: Lowest order $pp \rightarrow W^+W^-$ production modes at LHC.

As the feynmann diagrams shown in figures 4.3 and 4.12 suggest, the combinations of quarks which lead, in the lowest order, to a W^+W^- pair are very similar to the ones for the production of a Z^0 boson:

$$\left. \begin{array}{l} u\bar{u} \\ d\bar{d} \\ (s\bar{s}) \end{array} \right\} \longrightarrow \left\{ \begin{array}{l} Z^0 \\ W^+W^- \end{array} \right.$$

As indicated, a smaller contribution may also come from $s\bar{s}$ annihilation, which however does not involve any valence quark, while the $c\bar{c}$ one is further suppressed by the smaller parton c, \bar{c} -density functions (see figure 4.2).

Regarding the $pp \rightarrow W^+$ and $pp \rightarrow W^+Z^0$ processes (see the feynmann diagrams of figures 4.3 and 4.13), the main quark combinations are the following ones:

$$\left. \begin{array}{l} u\bar{d} \\ (c\bar{s}) \end{array} \right\} \longrightarrow \left\{ \begin{array}{l} W^+ \\ W^+Z^0 \end{array} \right.$$

Therefore, the related production cross sections depend mainly on the parton distribution functions of the u - and \bar{d} -quarks, with a smaller contribution given

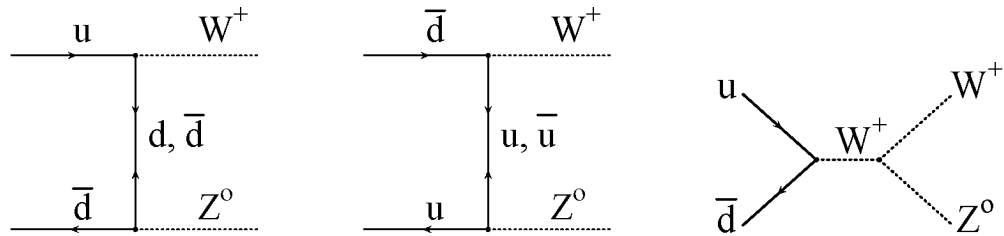


Figure 4.13: Lowest order $pp \rightarrow W^+ Z^0$ production modes at LHC.

by $c\bar{s}$. Other $q_i\bar{q}_j$ combinations, which mix the quark generations, are disfavored by the very small Cabibbo mixing angles and give thus much smaller contributions for the production of weak bosons.

In case of the $pp \rightarrow W^-, W^- Z^0$ processes, the main production modes are shown in figures 4.3 and 4.14. The quarks which lead to these weak boson(s) are mostly:

$$\left. \begin{array}{l} d\bar{u} \\ (s\bar{c}) \end{array} \right\} \longrightarrow \left\{ \begin{array}{l} W^- \\ W^- Z^0 \end{array} \right.$$

with much smaller contributions given by Cabibbo suppressed modes.

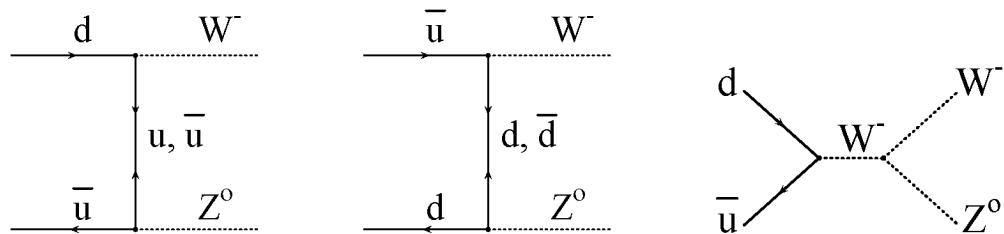


Figure 4.14: Lowest order $pp \rightarrow W^- Z^0$ production modes at LHC.

Finally, the available modes for the $pp \rightarrow Z^0, Z^0 Z^0$ productions are:

$$\left. \begin{array}{l} u\bar{u} \\ d\bar{d} \\ (s\bar{s}) \end{array} \right\} \longrightarrow \left\{ \begin{array}{l} Z^0 \\ Z^0 Z^0 \end{array} \right.$$

which are essentially the same as for the $pp \rightarrow W^+ W^-$ production.

To summarize, the single Z^0 production proceeds via the same quarks as the $W^+ W^-$ and $Z^0 Z^0$ productions and should therefore provide an efficient tool to

4.6 Precise prediction of diboson events

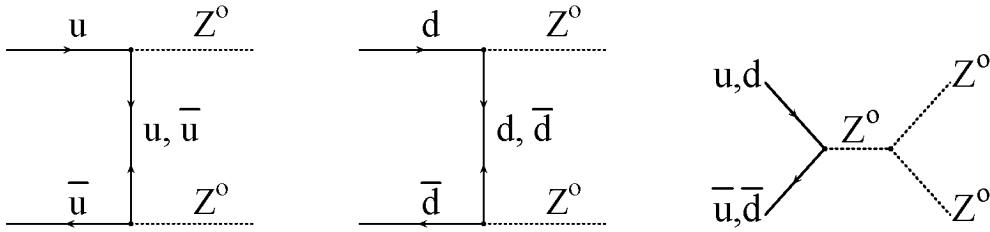


Figure 4.15: Main $pp \rightarrow Z^0 Z^0$ production modes at LHC.

control the cross section of these diboson reactions. The quarks involved in the production of positive W^+ bosons are the same as for the $W^+ Z^0$ production, while the negative W^- bosons are given through the annihilation of the same quarks which give rise to $W^- Z^0$ bosons. Therefore, by taking the ratios:

$$\frac{W^+ W^-}{Z^0}, \quad \frac{Z^0 Z^0}{Z^0}, \quad \frac{W^+ Z^0}{W^+}, \quad \frac{W^- Z^0}{W^-}$$

one can eliminate part of the uncertainties connected to the quark distribution functions.

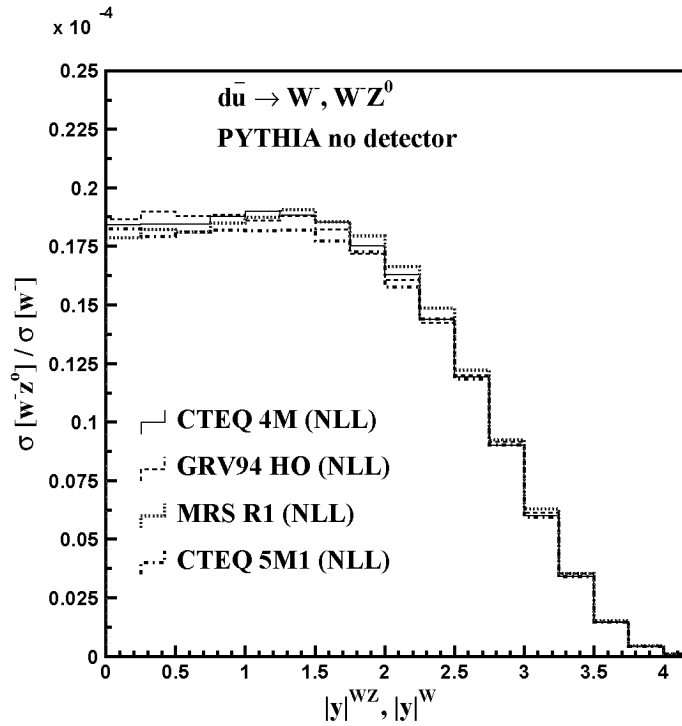


Figure 4.16: Ratios between $W^- Z^0$ and W^- cross sections as a function of the rapidity and according to four different NLL sets of structure functions.

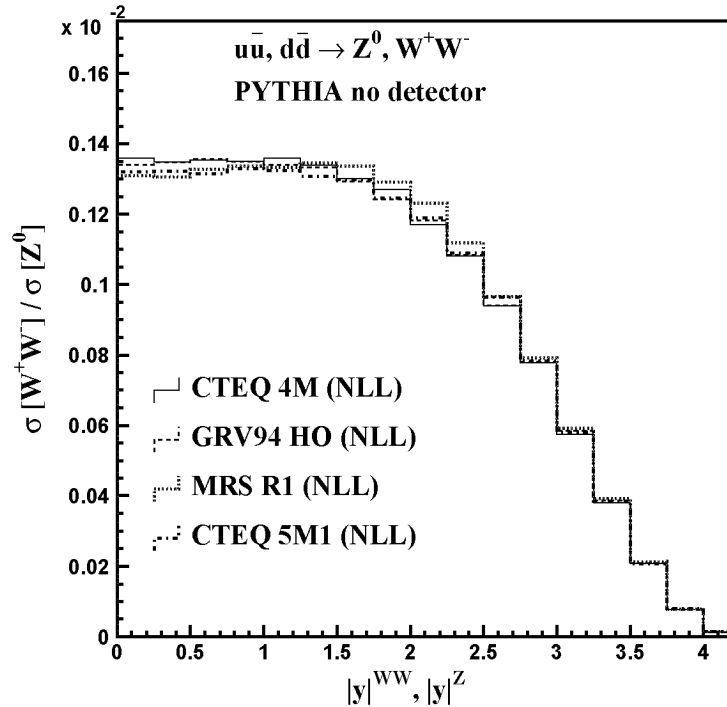


Figure 4.17: Ratios between W^+W^- and Z^0 cross sections as a function of the rapidity and according to four different NLL sets of structure functions.

The correlation between cross section predictions for single and pair production of weak bosons has been pointed out already in section 4.3 (see table 4.1). Now we repeat the comparison between sets of structure functions, however taking into account the ratios introduced in this section. Figures 4.16-4.17 show two of these ratios, each given by four different sets of parton distribution functions. These curves put in evidence the remaining uncertainties. If we show the behavior of these ratios, relative to a reference PDF set (CTEQ 4M), we can then better see the size of the uncertainties. This is done in figure 4.18 using the curves of figure 4.17. For example, the total cross section predictions for the process $pp \rightarrow W^\pm$ between the MRS(R1) and the GRV 94HO parameterizations differ by about 8%. However, as we suggest to use the processes $q\bar{q} \rightarrow W^\pm, Z^0$ as references, one has to relate for instance $\sigma(q\bar{q} \rightarrow W^-Z^0)$ to $\sigma(q\bar{q} \rightarrow W^-)$. Comparing now the prediction for the relative cross sections between MRS(R1) and GRV 94HO one finds that the difference is reduced to $\approx 2\%$.

The different cross section ratios are given in tables 4.3 and 4.4, where the values are obtained with a couple of older and some more recent sets of PDF's.

4.6 Precise prediction of diboson events

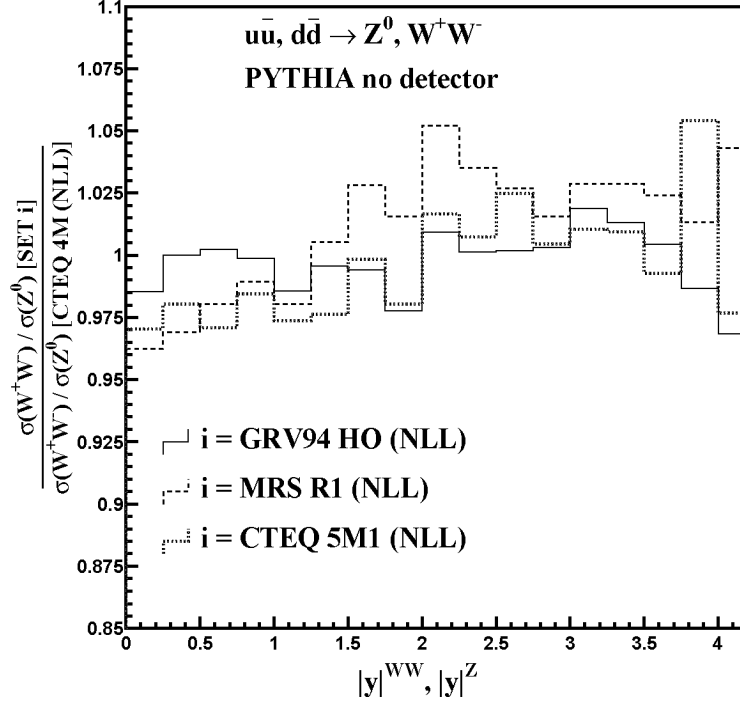


Figure 4.18: Ratios between the curves shown in figure 4.17.

PDF sets	“ \mathcal{L}_{ij} –calibration”	total cross section ratios	
	$\sigma_{Z^0 \rightarrow l+l^-}$	$\frac{\sigma_{W^+W^- \rightarrow l\nu l\nu}}{\sigma_{Z^0 \rightarrow ll}}$	$\frac{\sigma_{Z^0 Z^0 \rightarrow 4l}}{\sigma_{Z^0 \rightarrow ll}}$
MRS(A)	3.24 nb	1.089×10^{-3}	1.501×10^{-5}
MRS(H)	3.197 nb	1.098×10^{-3}	1.506×10^{-5}
MRS(R1)	3.188 nb	1.06×10^{-3}	1.468×10^{-5}
CTEQ 4M	3.479 nb	1.057×10^{-3}	1.47×10^{-5}
GRV 94HO	3.481 nb	1.036×10^{-3}	1.429×10^{-5}
CTEQ 5M1	3.375 nb	1.052×10^{-3}	1.458×10^{-5}
Uncertainty	$\sim 5\%$	$< 2\%$	

Table 4.3: Only decays to electrons and muons are taken into account, the uncertainties are given for the last four (more recent) sets of structure functions.

PDF sets	“ \mathcal{L}_{ij} –calibration”	total cross section ratios	
	$\sigma_{W^\pm \rightarrow l^\pm \nu}$	$\frac{\sigma_{W^+ Z^0 \rightarrow l \nu ll}}{\sigma_{W^+ \rightarrow l \nu}}$	$\frac{\sigma_{W^- Z^0 \rightarrow l \nu ll}}{\sigma_{W^- \rightarrow l \nu}}$
MRS(A)	34.35 nb	1.282×10^{-5}	1.107×10^{-5}
MRS(H)	34.13 nb	1.294×10^{-5}	1.084×10^{-5}
MRS(R1)	33.76 nb	1.289×10^{-5}	1.059×10^{-5}
CTEQ 4M	36.47 nb	1.277×10^{-5}	1.076×10^{-5}
GRV 94HO	36.82 nb	1.247×10^{-5}	1.043×10^{-5}
CTEQ 5M1	35.53 nb	1.266×10^{-5}	1.07×10^{-5}
Uncertainty	$\sim 5\%$	$< 2\%$	

Table 4.4: Only decays to electrons and muons are taken into account, the uncertainties are given for the last four (more recent) sets of structure functions.

As a next step, the parameterizations of the q, \bar{q} structure functions, especially at low x , should be adjusted such that the observed ℓ^\pm pseudorapidity distributions are described.

Once these better determinations of the parton density functions are achieved, one can safely use the $W^\pm \rightarrow l^\pm \nu$ signatures (which have the largest cross sections) to monitor all diboson events. Figure 4.19 shows, for example, the remaining uncertainties in the WW/W ratios as a function of rapidity. The visible differences (of few %) should then be eliminated, as soon as a better knowledge of the q, \bar{q} -densities will be provided by the $\sigma(q\bar{q} \rightarrow W^\pm, Z^0)$ processes.

As the final experimental accuracy for the lepton pseudorapidity distributions will be limited by systematics, the limitations of the structure function “fine tuning” are difficult to estimate. It is nevertheless worth pointing out that neither the ℓ^\pm momentum and charge determination nor differences between ℓ^+ and ℓ^- detection are expected to be problematic. Furthermore, backgrounds from different sources and efficiency uncertainties can be controlled by the simultaneous analysis of the W^\pm and Z^0 events with isolated electrons and muons. We therefore do not expect any principle problem of measuring the shape and the rate of the charged lepton pseudorapidity distribution with a $\pm 1\%$ accuracy. Thus even small differences for the sea quark parameterization, like those between MRS(A) and MRS(H), as shown in figure 4.7a and 4.7b, should be detectable. One could thus use the difference in cross section for the two sets as a pessimistic limitation of the proposed method. Differences between relative cross section predictions for different $q\bar{q}$ scattering processes and the two parameterizations indicate therefore the size of the remaining uncertainties. For example the cross section ratios

4.6 Precise prediction of diboson events

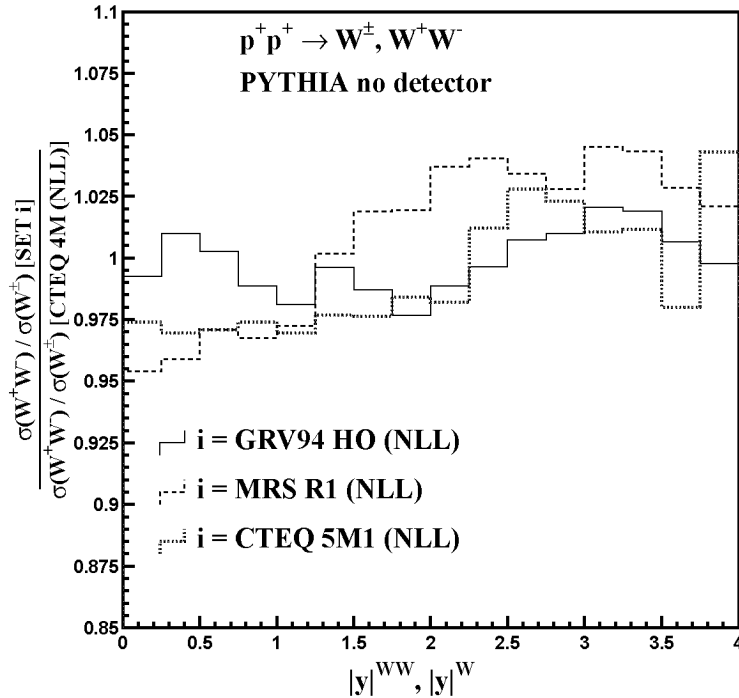


Figure 4.19: Same as figure 4.18, but for the ratios between W^+W^- and W^\pm cross sections.

$\sigma(q\bar{q} \rightarrow W^+W^-)/\sigma(q\bar{q} \rightarrow W^\pm)$ are 4.74×10^{-4} for MRS(A) and 4.76×10^{-4} for MRS(H). Other $q\bar{q}$ scattering processes like $\sigma(q\bar{q} \rightarrow W^\pm Z^0)/\sigma(q\bar{q} \rightarrow W^\pm)$ show similar stabilities.

Following the above procedure, *i.e.* constraining the q, \bar{q} structure functions and the corresponding parton luminosities with the W^\pm and Z^0 lepton decays, the uncertainties for the diboson event rates coming from PDF's and luminosity appear to be optimistically reduced to about $\pm 1\%$.

We have not investigated the achievable theoretical accuracies, but believe that many theoretical uncertainties, like the $\alpha_s(Q^2)$ uncertainties or still unknown higher order QCD corrections, contribute in very similar ways to the single and pair production of weak bosons. Furthermore, the possibility to measure the x distributions of sea and valence quarks and the corresponding luminosities within $\pm 1\%$ should encourage our theoretical colleagues to match this experimental accuracy. Eventually, this study needs to be repeated once complete next-to-leading order (NLO) Monte Carlo will be available.

Chapter 5

π - e misidentification as new background source

The Compact Muon Solenoid (CMS)[14] is designed to measure the energy and momentum of photons, electrons, muons and other charged particles with high precision. Among the detector requirements, an essentially background free identification of isolated, high- p_t electrons is of crucial importance for many physics channels to be measured at LHC. The electron energy will be measured by the electromagnetic calorimeter. In this chapter we discuss the case, where charged pions produce a large signal within the ECAL and could therefore be misidentified as electrons. This could result in new sources of background for the vector boson productions $W^\pm \rightarrow e^\pm \nu$, $Z^0 \rightarrow e^+ e^-$. At LHC the amount of π^\pm produced by the scattering of colored partons is huge, therefore it is particularly important to have good estimates of this effect.

The electron/charged pion discrimination is investigated using **GEANT** Monte Carlo simulations and the 1998/1999 test beam results [16]. We first study the interactions between π^\pm and the PbWO_4 crystals, and especially the events where a large energy deposition is observed. To this energy one has then to add the contribution of the so-called Nuclear Counter Effect observed in the Avalanche Photo-Diodes (APD).

5.1 The energy measurement in the electromagnetic calorimeter

We have already introduced the CMS electromagnetic calorimeter in section 3.1.3. The energy left in the ECAL produces scintillation light within the lead tungstate (PbWO_4) crystals and these photons have to be collected and measured by a photo-detector. The relatively low light yield of PbWO_4 , along with the high magnetic field and radiation environment in the CMS detector, severely limited the choice of these photo-devices. For the barrel part, new silicon Avalanche

Photo-Diodes (APD) were especially developed and optimized to measure wavelengths corresponding to the crystal scintillation light (which peaks at around 440 nm). Each ECAL barrel-crystal will be read out by two APD's.

5.1.1 Light Yield

The light yield (LY) of a crystal is measured by a photo-detector, which relates the number of generated photoelectrons to the scintillation light produced in the crystal, as given by [82]:

$$N_{pe} = LY \cdot \epsilon_Q \cdot f, \quad (5.1)$$

where ϵ_Q is the quantum efficiency of the photo-detector, which covers the fraction f of the crystal rear face.

Light yield values of 10 photoelectrons/MeV have been systematically observed using a photo-multiplier (XP2262B with $\epsilon_Q \approx 15 - 20\%$) covering all the back face of the 23 cm long crystals (*i.e.* $f = 1$) [83]. Consequently, with a perfect photo-detector ($\epsilon_Q = f = 1$), for each MeV left within the PbWO_4 crystals, about 50-60 photoelectrons should be observed.

5.1.2 The Avalanche Photo Diodes

Contrary to classical PIN photo-diodes, the Avalanche Photo-Diodes have an internal amplification (see for example reference [84]). Figure 5.1 shows a scheme of the APD internal structure and of its working principle.

In these sensors the light enters via a Si_3N_4 window and it is absorbed in a p^{++} layer ($\sim 2 \mu\text{m}$ thick), where the photons are converted into electron/hole pairs.

Electrons drift then in a high electric field region ($\sim 6 \mu\text{m}$ thick), where they are accelerated and multiplied by impact ionization (the gain is of about 50 to 200). Subsequently, they continue in a drift region of low doped intrinsic silicon and finally a highly doped "collection layer" precedes the ohmic contact.

The number of photoelectrons generated in the conversion layer by an energy

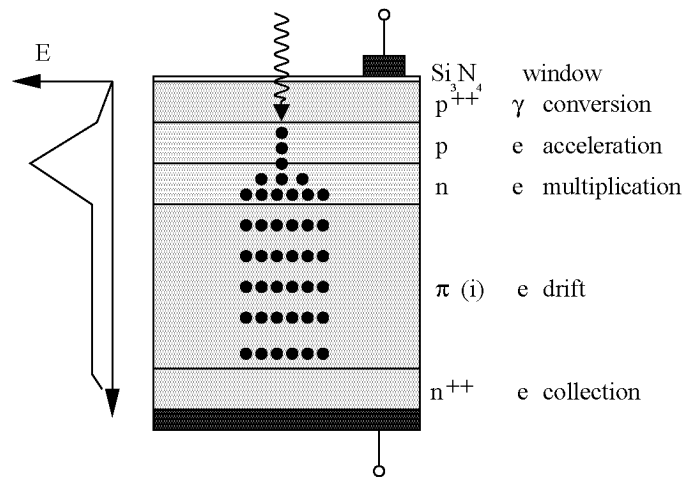


Figure 5.1: APD working principle.

deposition within the crystals is given by equation 5.1. Compared to a photomultiplier, the quantum efficiency ϵ_Q of an APD is larger and is of the order of 70%. However these sensors are small and one APD covers 5% of the crystal rear face. For the CMS ECAL the readout will be performed by two APD's on each crystal and therefore 10% of the rear surface will be covered. Consequently, for each MeV left within a PbWO_4 crystal, one expects to see approximately 4 photoelectrons. In the 1998 test-beam set-up (see appendix A), the readout was instead performed with a single APD and thus the covered area had to be divided by two, giving rise to about 2 photoelectrons per MeV.

5.1.3 The Nuclear Counter Effect

In addition to the photoelectrons produced in the APD's by an energy deposition in the crystal, electron/hole pairs can also be produced by the passage of ionizing radiation through the APD. This is the so-called ‘‘Nuclear Counter Effect’’ (NCE), which is responsible for a deterioration of the energy measurement, as one can not distinguish between the signal arising from ‘‘fake’’ e /hole pairs and those arising from a real energy deposition in the crystal. Charged particles, which cross a silicon layer, generate a number of electron/hole pairs per μm according to (see ref. [84]):

$$\frac{dn}{dx} = \frac{dE}{dx} \cdot \rho \cdot \frac{1}{E_{e/h}}, \quad (5.2)$$

where dE/dx is the deposited energy per unit of length, ρ is the density of the involved silicon layer and $E_{e/h}$ is the energy needed to create an e /hole pair. One expects a value dn/dx of about 80-100 e /hole pairs per μm [82, 84].

The total amount of produced electron/hole pairs can be obtained once the ‘‘active’’ thickness of the APD is known. In contrast to the photoelectrons coming out of the conversion layer which always experience full amplification, only electrons created before the high field region and holes created after this region are amplified. The so-called *effective thickness* d_{eff} of the APD can be quantified in terms of the real thickness d_{PIN} of a silicon PIN diode:

$$d_{eff} = \frac{d_{PIN}}{Q_{PIN}} \cdot \frac{Q_{APD}}{M}, \quad (5.3)$$

where Q_{PIN} and Q_{APD} are, respectively, the charges collected by the PIN diode and the APD, whose gain is M . The value of d_{eff} for the APD's depends mostly on the thickness of the p^{++} layer. In order to reduce the Nuclear Counter Effect it is clear that the effective thickness has to be reduced, causing however an increase of the capacitance of the device. For the selected Hamamatsu devices d_{eff} is supposed to be about 5 μm [82].

A comparison of the number of e /holes pairs produced by the Nuclear Counter

5.2 Charged pions crossing the ECAL

Effect with the number of e /holes pairs produced by scintillation light, allows to calculate the corresponding fake energy signal.

One “Minimum Ionizing Particle” (MIP) crossing the APD will then give a signal equivalent to the one given by a photon in the crystal with energy:

$$E_{\text{NCE}} = \frac{\frac{dn}{dx} \cdot d_{\text{eff}}}{N_{pe}} . \quad (5.4)$$

For example a μ (MIP) deposits an energy of about 280 MeV within the crystal. Therefore, with a two APD read-out (*i.e.* $N_{pe} \approx 4 \text{ pe/MeV}$), one gets about 1120 photoelectrons from scintillation light. In addition about 500 pe are produced if the μ crosses the APD. Consequently one expects the size of the Nuclear Counter Effect to be about 100-150 MeV, according to:

$$E_{\text{NCE}} = \frac{500 \text{ pe}}{1120 \text{ pe}} \cdot 280 \text{ MeV} = 125 \text{ MeV} . \quad (5.5)$$

For a '98 test beam-like set-up (with one APD per crystal and thus $N_{pe} \approx 2 \text{ pe/MeV}$), the scintillation light produces half of the photoelectrons, while the number of those produced by a charged particle crossing the APD remains constant. Therefore, one obtains instead the corresponding E_{NCE} -value of about 200-250 MeV.

5.2 Charged pions crossing the ECAL

To study the Nuclear Counter Effect and its consequences for the pion-electron misidentification, we make use of **GEANT** simulations with a test beam-like set-up (without magnetic field). For this purpose we defined a matrix of 9×9 crystals, with the dimensions of $2.2 \times 2.2 \times 23 \text{ cm}^3$, as for the barrel part of the ECAL. Following the foreseen CMS design, we first use a set-up with two APD's on each crystal rear-face to simulate the consequences of a charged particle leakage on the pion-electron confusion.

A set-up with one APD per crystal will then be instead used to compare the simulations with the 1998 test-beam data.

5.2.1 Nuclear Counter Effect contribution to the ECAL response

The interactions between charged pions and the lead tungstate (PbWO_4) crystals can be schematically subdivided in two main possibilities. Either the particles produce only simple ionization in the material or they interact hadronically, resulting in a larger amount of deposited energy.

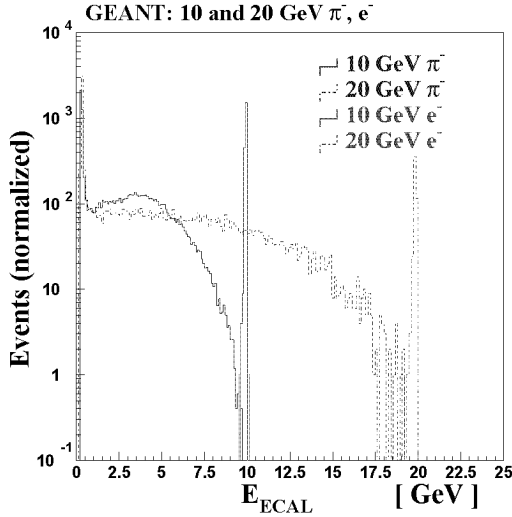


Figure 5.2: GEANT simulation of the total amount of energy deposited by charged pions and electrons within the ECAL.

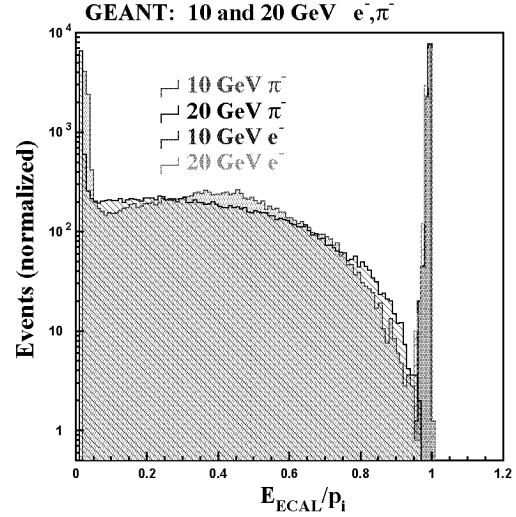


Figure 5.3: The curves shown in figure 5.2 divided by the initial momentum $|\mathbf{p}_i|$.

In the first case the charged pions behave like “Minimum Ionizing Particles” (MIP) and, like muons, deposit about 280 MeV of energy within the PbWO_4 crystals.

If some hard interaction like $\pi^\pm A \rightarrow A^* \pi^0 \dots \pi^\pm \dots$ occurs (where A means an atom and A^* a modified one), the presence of π^0 's gives instead rise to electro-magnetic showers within the crystal.

Figure 5.2 shows the energy deposited by π^\pm and e^\pm within the 9×9 matrix of crystals. If this deposited energy is subsequently divided by the initial momentum $|\mathbf{p}_i|$, one obtains the curves shown in figure 5.3. It appears that, for different initial momenta, pions give rise to about the same $E_{\text{ECAL}}/|\mathbf{p}_i|$ shape. The e^\pm 's leave all their energy within the ECAL and produce a narrow Gaussian peak at $\approx 100\%$ fractional energy. On the contrary, non interacting (MIP) π^\pm 's leave only a small amount of energy, giving rise to the peak on the left side of the plots. Furthermore, a small fraction of the “hard interacting” pions, shows a deposited energy $E_{\text{ECAL}}/|\mathbf{p}_i|$ in excess of 0.95.

The momentum resolution of the tracking system, the ECAL energy resolution and the wanted electron efficiency define the lower $E_{\text{ECAL}}/|\mathbf{p}_i|$ -cut for an electron selection (*i.e.* the threshold for a possible $\pi^\pm - e$ misidentification). For the simulations with a complete energy reconstruction of the event (*i.e.* using all the 9×9 crystals), a lower $E_{\text{ECAL}}/|\mathbf{p}_i|$ -cut of 0.95 is used.

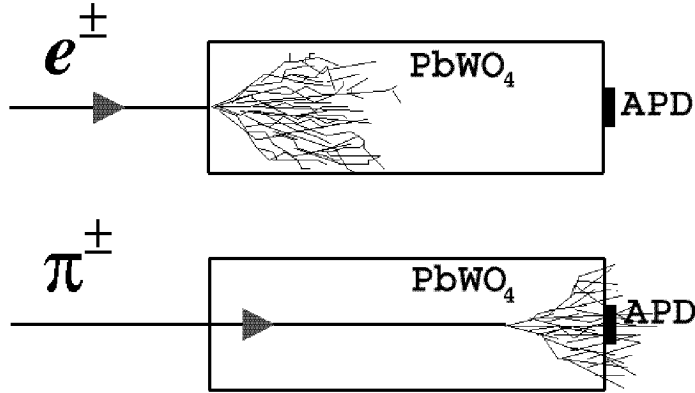


Figure 5.4: Schematic view of the electron and pion showers produced in the PbWO_4 crystals. A charged pion shower can start from any position, given the much smaller probability.

The **GEANT** simulations show that about 0.015% of the π^\pm 's, with an initial momentum p_i in the range 5-30 GeV, give an energy deposition $E_{\text{ECAL}}/|\mathbf{p}_i|$ larger than 0.95. For 50 GeV π^\pm 's, we found one event in 10^5 and for p_i larger than 50 GeV no candidate was found within 2×10^4 simulated events (the simulations at high energies are much slower, giving less statistics).

One has then to add the “fake”-energy contribution from the Nuclear Counter Effect, which is simulated in the following way. The crystals are coupled with two $5 \times 5 \text{ mm}^2$ pieces of silicon to simulate the APD's. In a first simple method, for every charged particles entering the silicon, we calculate the incidence angle, in order to have the path length within the APD's. We assume that, if the path is perpendicular to the surface, we would obtain exactly $E_{\text{NCE}} = 100 \text{ MeV}$. This method has no spread in the energy deposition and, for a given angle with respect to the APD, gives always the same value. For this reason we developed a second method, where signal fluctuations are included. In fact, the energy lost within the silicon by charged particles extends over a broad range with a Landau tail at large values. As given in equation 5.2, this deposited energy produces then the electron/hole pairs which are accelerated and multiplied in the high electric field. Therefore, in the second method, we multiply the deposited energy with a factor to obtain a Nuclear Counter Effect with statistical fluctuations close to reality. This factor is chosen by forcing the mean deposited energy to match a Nuclear Counter Effect of 100 MeV.

Both methods produce similar results for $E_{\text{ECAL}}/|\mathbf{p}_i| < 1$. However, the second method produces a longer tail at large $E_{\text{ECAL}}/|\mathbf{p}_i|$ values, as expected by a statistical distribution of the deposited energies. In the following studies, we always use the second method.

For crystal sizes of about 25 radiation lengths, very small shower leakage is ex-

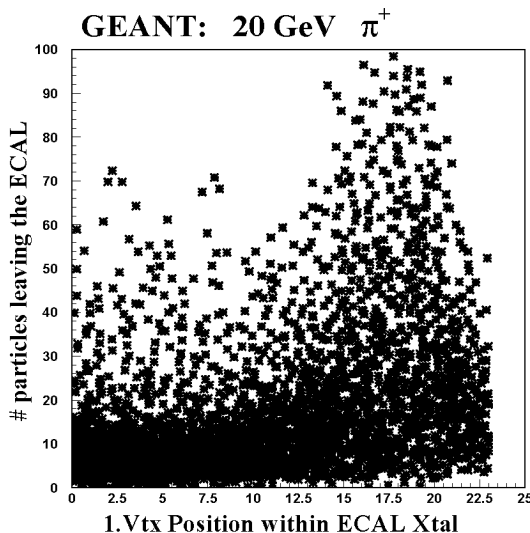


Figure 5.5: Number of charged particles exiting via the rear face of the crystal as a function of the position of the first hard interaction $\pi^+A \rightarrow A^*\pi^0 \dots \pi^+ \dots$

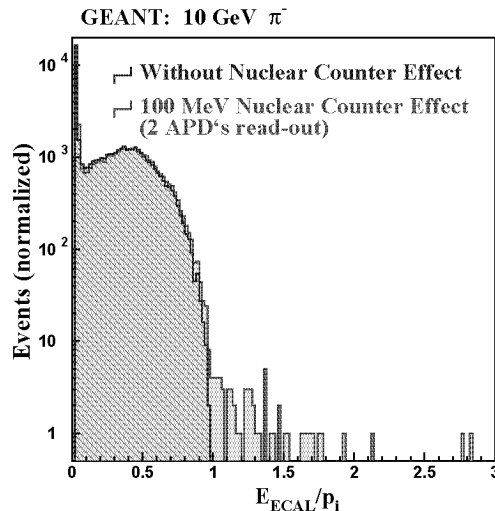


Figure 5.6: GEANT simulation showing the $E_{\text{ECAL}}/|\mathbf{p}_i|$ signal generated by 10 GeV π^- 's. The two curves show the results obtained with or without Nuclear Counter Effect.

pected for electrons and photons with energies below 1 TeV. On the contrary, for charged pions, a hard interaction is less likely and the shower can start all along the crystal length (see figure 5.4). The leakage of ionizing particles could then result in problems for the corresponding energy measurement.

In figure 5.5 a GEANT simulation shows the number of charged particles exiting via the crystal rear face as a function of the position of the first hard process $\pi^\pm A \rightarrow A^*\pi^0 \dots \pi^\pm \dots$ (*i.e.* the origin of the shower). In case that an hard interaction occurs in the last part of the PbWO_4 crystal, a large number of particles might eventually pass through the APD's (as shown schematically in figure 5.4) and produce a sizeable fake energy signal. For example, if 100 particles cross these photo-devices and each one gives a 100 MeV large Nuclear Counter Effect, then a 10 GeV fake signal is observed (starting for instance from a 20 GeV π^\pm , as shown in figure 5.5!). Then, adding the energy deposited in the crystals, these pions can give rise to signals of more than 20 GeV, that is to $E_{\text{ECAL}}/|\mathbf{p}_i|$ ratios in excess of 1.

Starting with the estimated parameters, that is from an effective thickness of 5 μm and from a yield of 4 photoelectrons per MeV, one finds the $E_{\text{ECAL}}/|\mathbf{p}_i|$ distributions shown in figures 5.6-5.8 for, respectively, 10, 20 and 50 GeV negative pions. Equivalent $E_{\text{ECAL}}/|\mathbf{p}_i|$ distributions are obtained with positively charged

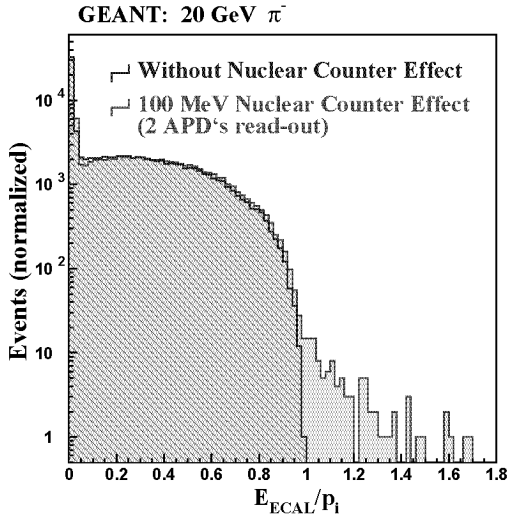


Figure 5.7: Same as figure 5.6, but for 20 GeV π^- 's.

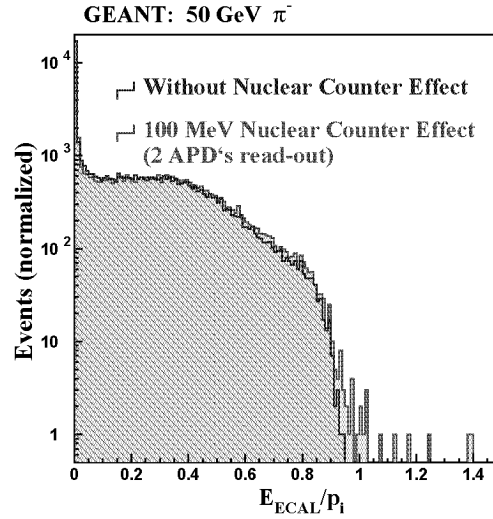


Figure 5.8: Same as figure 5.6, but for 50 GeV π^- 's.

pions. The Nuclear Counter Effect produces a tail at large $E_{\text{ECAL}}/|\mathbf{p}_i|$, enhancing the probability of a pion–electron confusion.

As a result one finds that the fraction of π^\pm 's with an $E_{\text{ECAL}}/|\mathbf{p}_i|$ -value larger than 0.95 increases, for example, from 0.006% to 0.16% for 5 GeV π^\pm 's and from 0.018% to 0.07% for 20 GeV π^\pm 's.

The possible consequences for physics are discussed in section 5.6.

5.3 Results from the 1998 test beam runs

During July 1998, test beam data with electrons, muons and pions were collected. A set-up (PROTO97) made-up of 6×6 crystals, each coupled with a single APD, was used to perform different measurements with different particle energies (for details see appendix A or reference [85]). The results of these test runs can be compared to our simulations.

5.3.1 Measurement of the Nuclear Counter Effect

First, one has to determine the size of the Nuclear Counter Effect (NCE). For this purpose the data collected with muons are analyzed. As these particles behave essentially like Minimum Ionizing Particles, they produce a peak in the energy plot at $E \approx 280$ MeV [86].

To measure the NCE one has to compare the observed energy of tracks directed towards the APD's with those of tracks directed outside this area.

Once the position of the APD's is isolated (as shown in figure 5.9), we perform final scans in x and y -direction by taking only data within, respectively, the 5 mm-wide y and x -regions corresponding to the APD position. The results are shown in figures 5.10 and 5.11, where a clear shift in energy in correspondence with the APD can be observed. The size of this shift can be extracted from figure 5.12, where the energy signals of muons are subdivided depending on the track direction (*i.e.* either towards or outside the APD position). One finds that the peak position for tracks directed towards the APD is between 700 and 800 MeV, compared to the about 300 MeV obtained for tracks not cross-

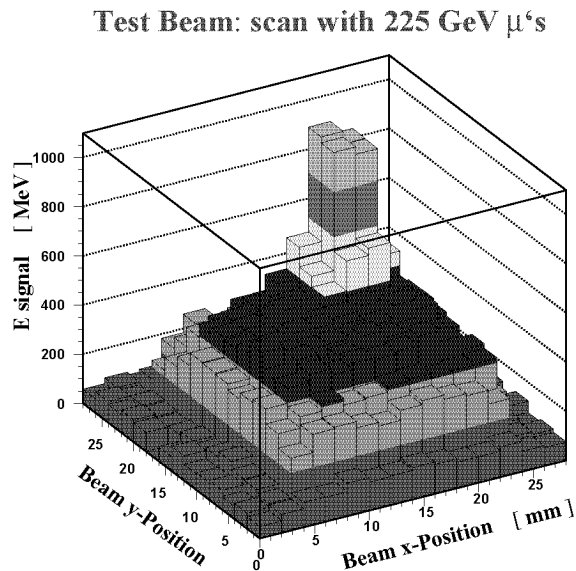


Figure 5.9: Measured energy during the $x - y$ -scan. The APD position can be clearly isolated as well as the size of the Nuclear Counter Effect.

ing the APD. The separation between the two peaks is thus approximately 450 MeV. Similar results are obtained for all crystals in the 1998 test beam set-up. Therefore, for muons, a Nuclear Counter Effect between 400 and 500 MeV instead of the expected 200–250 MeV is found (for a single APD readout).

The same effect is measured with charged pions. We select MIP-like π^\pm 's by requiring a "zero-energy" deposition in the neighboring crystals (*i.e.* we veto any "strongly interacting" pions). Figure 5.13 shows the results taken with different pion initial momenta. The space between the two peaks is about 440 MeV, compared with the expected Nuclear Counter Effect of 200–250 MeV. This value is thus consistent with the one observed with muons.

There are two main possibilities which can explain this larger Nuclear Counter Effect:

- (1) either less light reached the photo-diodes than expected,
- (2) or there was a larger production of electron/hole pairs due to the charged particles crossing the APD's.

To measure the absolute number of photoelectrons produced at the cathode per MeV one can use a calibration method based on the LED line width, which is proportional to the number of photons reaching the APD's.

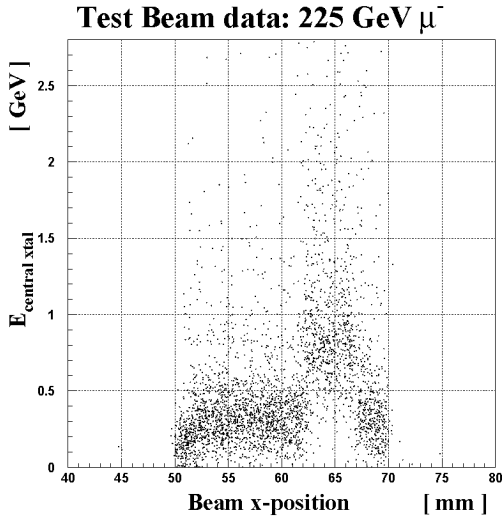


Figure 5.10: Muon energy left in the central crystal as a function of the track position on the x -axis.

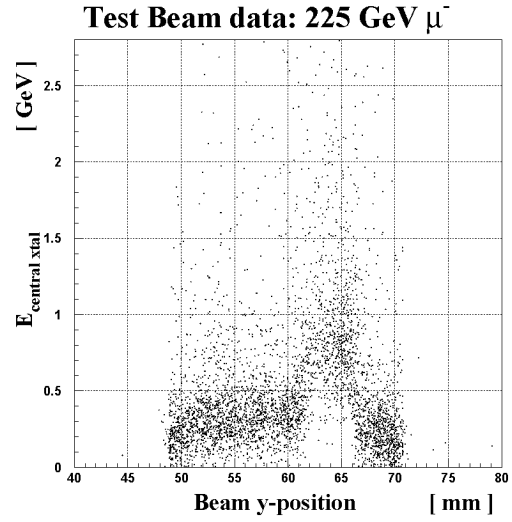


Figure 5.11: Same as fig. 5.10, but as a function of the track y -position. Due to the Nuclear Counter Effect there is a larger signal when the track crosses the APD.

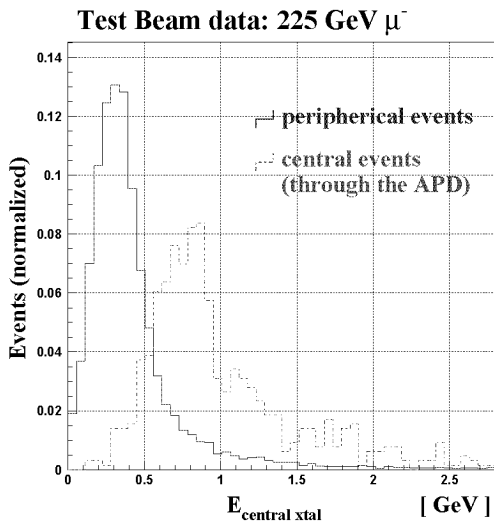


Figure 5.12: Energy signal of 225 MeV muons. The “central events” have the track directed towards the APD. The difference between the two peaks gives the size of the Nuclear Counter Effect.

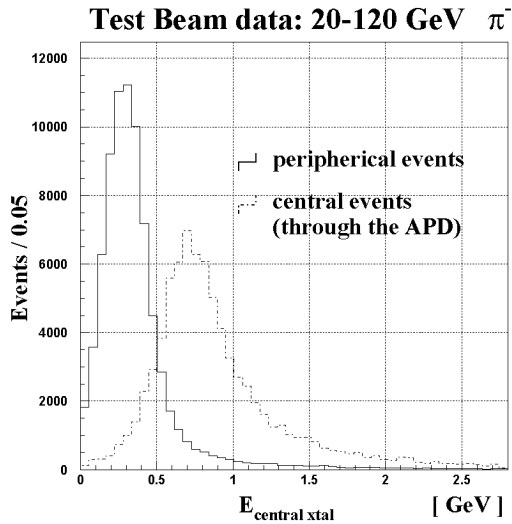


Figure 5.13: Same as figure 5.12, but for pions (20, 50 and 120 GeV) behaving like Minimum Ionizing Particles.

The analysis of data, taken by irradiating the crystal matrix with lasers, shows in fact that less light was collected by the APD's in comparison with the expectations (about 1 pe /MeV instead of 2 pe /MeV [87]). A checkup of the matrix after the tests indicated that this lack of photoelectrons depended on a bad contact between crystal rear faces and APD's and not on a lower light yield of the $PbWO_4$ crystals. The gluing material between them was filled with bubbles, so that the light transmission was difficult. By taking into account this lower transmitted light (*i.e.* using equation 5.4 with $N_{pe} = 1$ pe /MeV), one obtains a value of E_{NCE} between 400 and 500 MeV, in good agreement with the data. Therefore the measured Nuclear Counter Effect could be fully explained by an inefficient gluing between APD's and crystals and their properties were then not called in question.

5.3.2 Pions as “fake” electrons candidates

To compare simulations with the 1998 test beam data, we run **GEANT** with a Nuclear Counter Effect of 450 MeV instead of the estimated 200-250 MeV [82].

To reduce the noise effects originating from pedestal fluctuations, a 3×3 crystal matrix is used to reconstruct the particle energy. This noise (of the order of 1 GeV for the 3×3 crystal sum) causes also a broadening of the electron peak and of the pion MIP peak. In figures 5.14-5.16, the fractional energy collected in the 3×3 crystals around the hit point is shown for both e^- and π^- and with beam energies of, respectively, 20, 50 and 120 GeV. The pion data samples consist in 154×10^3 events at 20 GeV, 193×10^3 events at 50 GeV and 217×10^3 events at 120 GeV.

Electrons leave about 90% of their initial energy in the 3×3 crystal matrix, producing the Gaussian peaks shown in the figures. Events with

lower $E_{3 \times 3}/|\mathbf{p}_i|$ indicate a pion contamination in the electron beam (of the order of 0.4% for 20 and 50 GeV and of 3.5% for 120 GeV).

In contrast to e^\pm , charged pions produce a long tail at high $E_{3 \times 3}/|\mathbf{p}_i|$, as shown in figures 5.14-5.16. This tail is well reproduced by **GEANT** simulations, which include a Nuclear Counter Effect of 450 GeV. A comparison between simulations and data is shown figures 5.17-5.19.

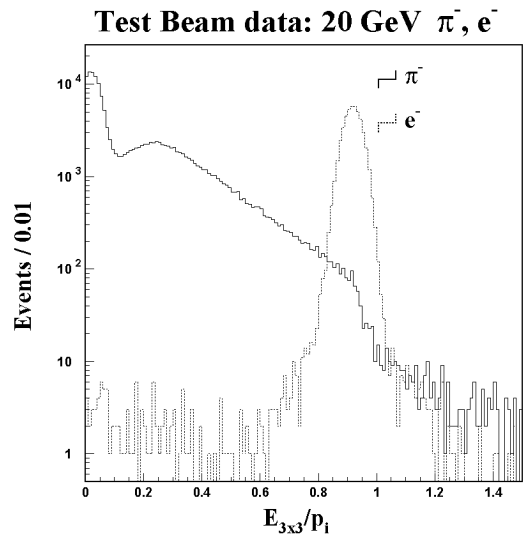


Figure 5.14: 1998 test beam data: fractional amount of energy left by charged pions and electrons with an initial momentum of 20 GeV.

5.3 Results from the 1998 test beam runs

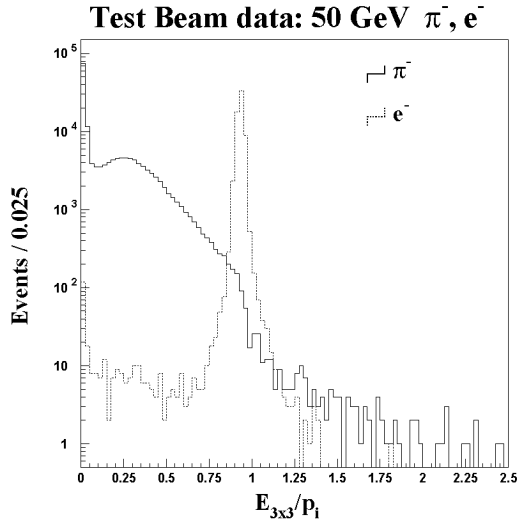


Figure 5.15: '98 test beam data: fractional amount of energy left by charged pions and electrons with an initial momentum of 50 GeV.

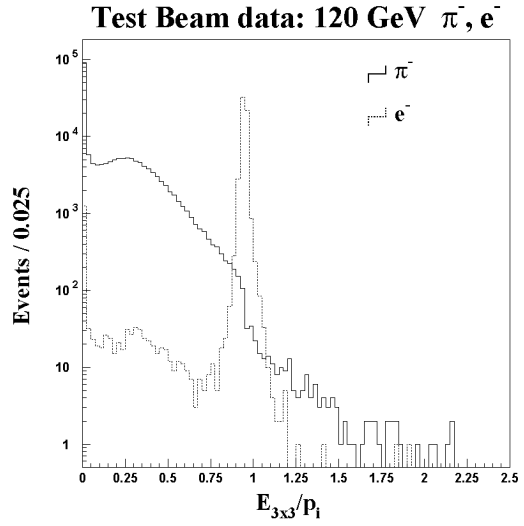


Figure 5.16: Same as figure 5.15, but for 120 GeV charged pions.

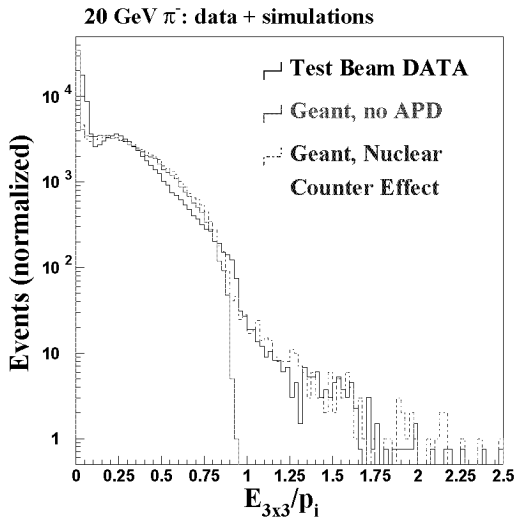


Figure 5.17: 20 GeV π^- 's: 1998 test beam data compared with GEANT simulations with and without Nuclear Counter Effect (the noise is not simulated).

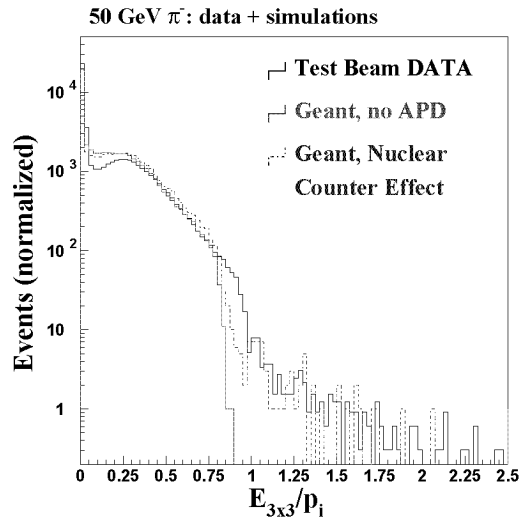


Figure 5.18: Same as figure 5.17, but for 50 GeV charged pions.

A small electron contamination in the pion beam has also to be taken into account. It produces a small excess close to $E_{3\times 3}/|\mathbf{p}_i| \approx 0.9$ (put in evidence in figure 5.20). However this contamination does not affect the measurement of events with $E_{3\times 3}/|\mathbf{p}_i| > 1.2$, as the electron data do not show a tail.

Table 5.1 summarizes the test beam results and the **GEANT** simulations with a 450 GeV Nuclear Counter Effect. By comparing simulations and data in the $E_{3\times 3}/|\mathbf{p}_i|$ -region between 0.85 and 1.2, the electron contamination can be extrapolated and it appears to be less than one per-mil for 20 GeV and of the order of few per-mil for 50 and 120 GeV.

The simulations show that the fraction of π^- 's with a value $E_{3\times 3}/|\mathbf{p}_i|$ larger than 0.85 vary between 0.7% (for 20 GeV) and about 0.2% (for 50 and 120 GeV). These percentages depend partly also on the choice of **GEANT** parameters¹.

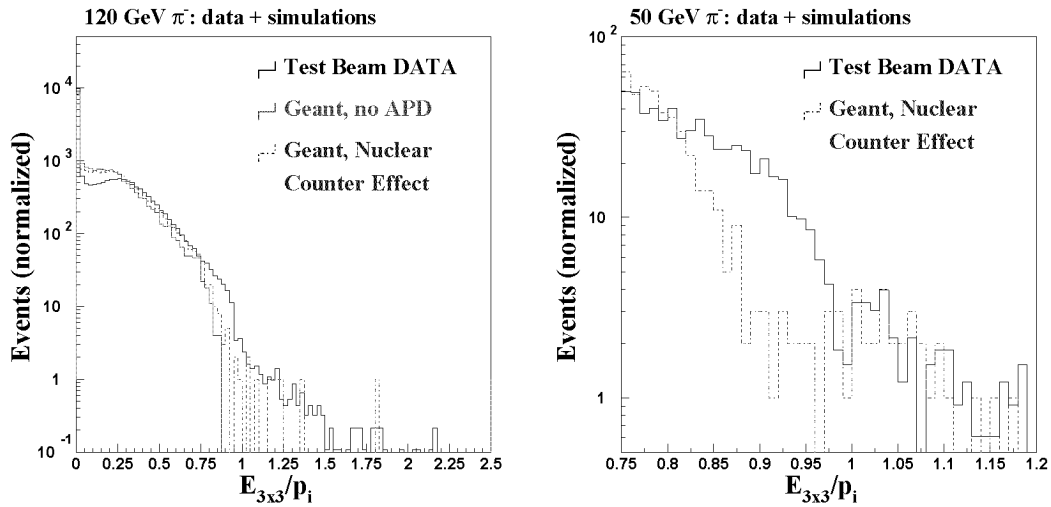


Figure 5.19: Same as figure 5.17, but for 120 GeV charged pions.

Figure 5.20: 50 GeV π^- 's: comparison between data and simulations, showing the $E_{3\times 3}/|\mathbf{p}_i|$ -region where an electron contamination is expected.

¹In **GEANT** the different particles are only transported above a minimal energy. For π/e momenta smaller than 50 GeV, this limit is set at 100 keV. For the simulations with initial π/e momenta of 50 and 120 GeV, these minimal energies are instead put at higher values to increase the simulations speed. This causes a deficit of pion events with large $E_{3\times 3}/|\mathbf{p}_i|$, as charged particles stop more frequently in the crystals and therefore can not cross the APD's. Consequently, the simulations with p_i larger than 50 GeV tend to underestimate the data in the $E_{3\times 3}/|\mathbf{p}_i|$ -region larger than 1.2. Furthermore, the production of electron-hole pairs within the APD's is subject to a saturation at high energy depositions. We introduced therefore a cut-off in the deposited energy within the silicon when simulating the Nuclear Counter Effect. The results of the simulations show a small dependence on this cut-off.

5.4 Results from 1999 test beam runs

E_{π^-}	1998 Test beam data (PROTO 97)		GEANT (450 MeV NCE)	
	$0.85 < \frac{E_{3 \times 3}}{ \mathbf{p}_i } < 1.2$	$\frac{E_{3 \times 3}}{ \mathbf{p}_i } > 1.2$	$0.85 < \frac{E_{3 \times 3}}{ \mathbf{p}_i } < 1.2$	$\frac{E_{3 \times 3}}{ \mathbf{p}_i } > 1.2$
20 GeV	0.56 %	0.08 %	0.58 %	0.12 %
50 GeV	0.41 %	0.056 %	0.15 %	0.05 %
120 GeV	0.39 %	0.041 %	0.12 %	0.02 %

Table 5.1: Charged pions as “fake electron” candidates: comparison between ’98 test beam data and Monte Carlo simulations (**GEANT**), which include the large Nuclear Counter Effect observed during these test runs (450 MeV).

5.4 Results from 1999 test beam runs

During the test beam runs performed with a new ECAL set-up in summer 1999 (see appendix A), also some (few) data were collected with charged pions. The results of these runs are shown in figures 5.21-5.22 (for 50 and 120 GeV π^- ’s). In comparison with the 1998 data, the large tail in the high $E_{3 \times 3}/|\mathbf{p}_i|$ -region disappeared. Figure 5.23 shows indeed that the **GEANT** simulations with the expected 100 MeV Nuclear Counter Effect are consistent with the 1999 data. However, also in this case, an electron contamination can be inferred from the excess of events within the $0.8 < E_{5 \times 5}/|\mathbf{p}_i| < 1$ region. This contamination should be accounted for the discrepancy between **GEANT** simulations and data reported in the central column of table 5.2.

Concerning the electrons, the improvement in resolution with respect to the previous year is shown in figure 5.24. For 50 GeV e^- ’s, the width of the Gaussian $E_{3 \times 3}/|\mathbf{p}_i|$ -peak dropped from approximately $\pm 1.5\%$ to $\pm 0.75\%$.

$E_{\pi^-} = 50$ GeV	$0.85 < \frac{E_{5 \times 5}}{ \mathbf{p}_i } < 1.2$	$\frac{E_{5 \times 5}}{ \mathbf{p}_i } > 1.2$
1999 Test beam data	0.62 %	3×10^{-5}
GEANT (100 MeV NCE)	0.3 %	4×10^{-5}

Table 5.2: Charged pions as “fake electrons” candidates: comparison between 1999 test beam data and Monte Carlo simulations (**GEANT**), with the expected 100 MeV Nuclear Counter Effect.

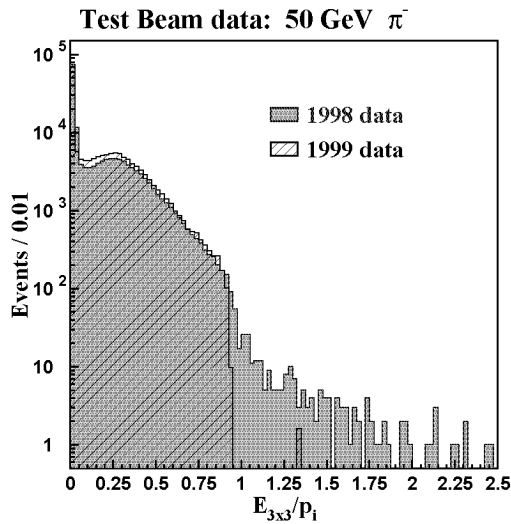


Figure 5.21: 50 GeV pions: comparison between test beam data taken in 98 and 99.

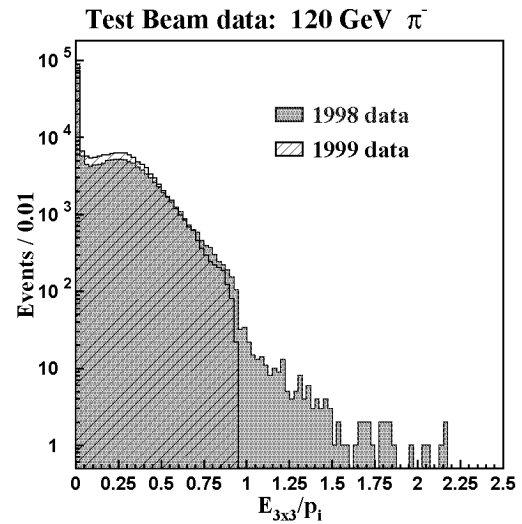


Figure 5.22: 120 GeV: comparison between test beam data taken in 98 and 99.

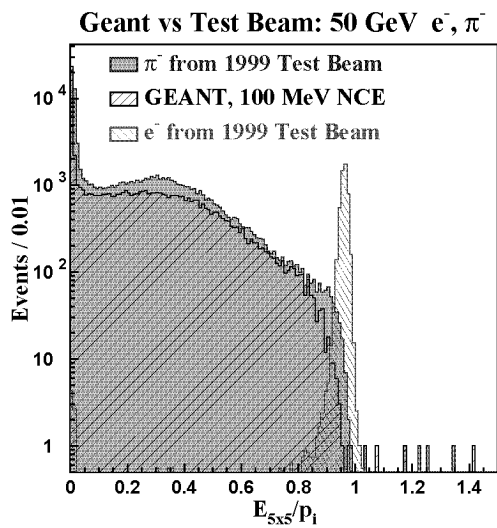


Figure 5.23: 50 GeV pions: test beam data collected in 1999 compared with GEANT simulations (with the expected 100 MeV Nuclear Counter Effect).

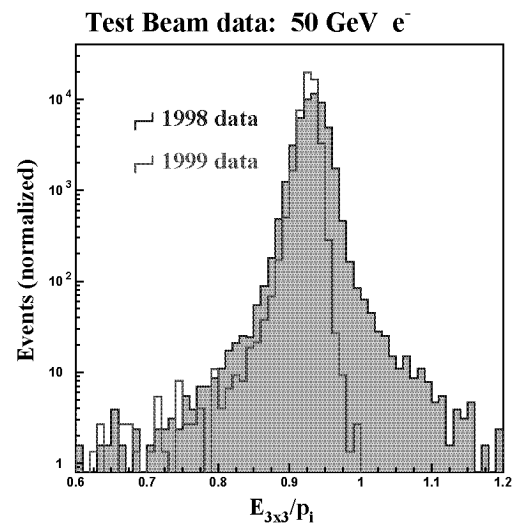


Figure 5.24: Test beam data: electron resolution improvement.

5.5 Discrimination methods for isolated pions and electrons

In contrast to electrons, which deposit almost 100% of their energy in the ECAL, the $E_{\text{ECAL}}/|\mathbf{p}_i|$ ratio for charged pions is generally very small. Consequently the first measure to limit the pion-electron discrimination consists in putting a lower cut on the $E_{\text{ECAL}}/|\mathbf{p}_i|$ ratio.

However, due to the interactions with the material within the inner detector, electrons can lose some of their initial energy by bremsstrahlung before reaching the ECAL. These radiated γ 's enter the ECAL crystals near the original electron and their energy can contribute to the measured electron energy. Nevertheless, as the photon momenta are not determined, the tracker could measure a lower momentum compared to the deposited energy, giving rise to $E_{\text{ECAL}}/|\mathbf{p}_i|$ ratios larger than 1. Therefore a too restrictive $E_{\text{ECAL}}/|\mathbf{p}_i|$ -upper cut should be avoided.

Besides the $E_{\text{ECAL}}/|\mathbf{p}_i|$ -based selection, we take into account the following possibilities² :

- HCAL veto

- ECAL shower shape.

The first one consists in putting a veto if the hadronic energy measured in the HCAL exceeds a given value [86]. The idea behind is that, for electron-like π^\pm -events (especially those with a remarkable Nuclear Counter Effect contribution), a large number of hadrons survives the passage through the ECAL and enters in the successive HCAL, while for electrons with energies below 1 TeV essentially no leakage of particles is observable.

Such a criteria can be studied with **GEANT** simulations: we require that the energy sum for hadrons exiting the rear side of the hit 3×3 crystal matrix does not exceed 1 GeV, for initial energies below 20 GeV, or 5% of the particle initial momentum, for initial energies above 20 GeV.

The efficiency of the ‘‘HCAL veto’’ cut is not comparable to the one obtained with a shower shape selection (see next section). In fact, hadrons which leave the ECAL often have already energies below 1 GeV and are thus undetectable in the hadron calorimeter. Nevertheless, for our fast detector simulations, such HCAL cut gives good results for pions with very large fractional energy depositions within the ECAL ($E_{\text{ECAL}}/|\mathbf{p}_i| > 2$), succeeding in eliminate events where hundreds of particles cross the APD's and subsequently enter the HCAL.

²Another selection method (signal timing) has also been briefly reviewed. It is based on the assumption that the APD pulse shape differ, depending if it has been produced by scintillation light or by the Nuclear Counter Effect. Unfortunately no difference has been observed with the 1998 test beam data [88].

5.5.1 Shower shape selection of electrons

A “shower shape cut” has the advantage to be directly applicable to the 1998 test beam data. The method is based on the assumption that the deposited energy distribution over the crystals is different for charged pions and electrons [89]. Its applicability depends on the track reconstruction precision, on the ECAL granularity and on the knowledge of the radiative processes within the tracker. For the test beam data, the absence of magnetic field and of material in front of the test set-up, made things easier. Anyway, we do not expect that, under real conditions, the use of the shower shape cuts will be prevented.

As discussed in section 5.2.1, a shower arising from π^\pm -tracks can start all along the crystal and its cone-shape could therefore be different compared to electrons.

Moreover, electron-like π^\pm -events often arise from a particle shower with a tail entering one APD, producing a large Nuclear Counter Effect and thus an abnormal energy deposition in one crystal.

An electron selection using shower shape cuts seems thus promising. Figure 5.25 shows, for 20 GeV electrons and pions, the energy collected in the central crystal divided by the energy deposited in the 3×3 crystal matrix. We call this ratio “E1/E9” and we measure it only for tracks directed towards the central part of the crystal face. For tracks directed towards the crystal sides (see figure 5.26) we add to the energy of the hit crystal also the energy of the closest one (E2/E9). Finally, for tracks directed towards the crystal corners, we use the energy of the hit crystal plus the one collected in the closest three crystals (E4/E9), as shown in figure 5.27.

In order to tune the cuts to the relevant events, we select only charged pions which produced a fractional energy deposition larger than 0.8. To extract the E_n/E_9 -boundaries, the results given by electrons and pions are compared. The goal is to eliminate the “dangerous” pions without affecting too much the electron efficiency. Electron-like pions tend more frequently to leave a large amount of

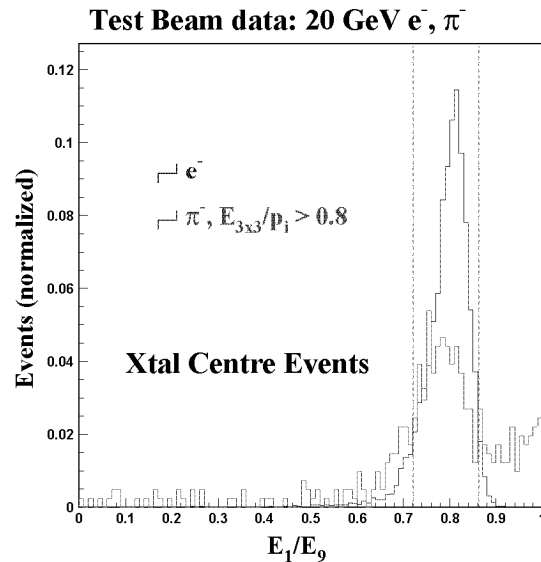


Figure 5.25: For 20 GeV e -like pions ($E_{3\times 3}/|\mathbf{p}_1| > 0.8$) and electrons: deposited energy within the hit crystal divided by the deposited energy in the 3×3 crystal matrix. Only tracks directed towards the center of the crystal front face have been selected.

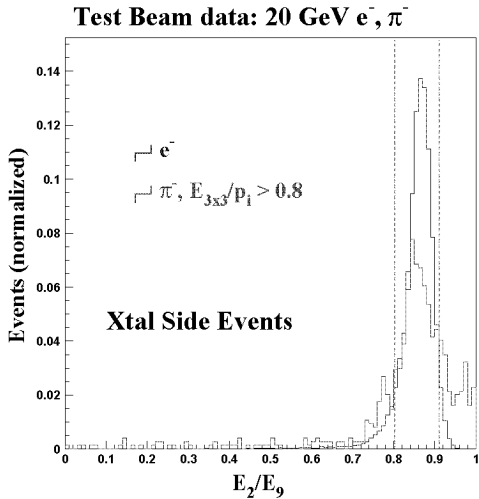


Figure 5.26: For 20 GeV pions and electrons: deposited energy in the two crystals closest to the entering track, divided by the deposited energy in the 3×3 crystal matrix. Only tracks directed towards the sides of a crystal front face have been selected.

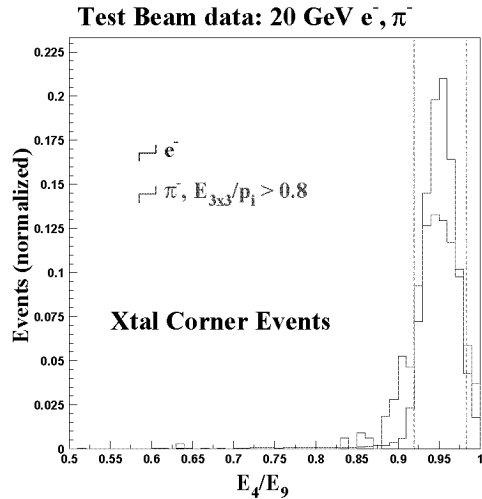


Figure 5.27: For 20 GeV pions and electrons: deposited energy in the four crystals closest to the entering track, divided by the deposited energy in the 3×3 crystal matrix. Only tracks directed towards crystal corners have been selected.

energy in a single crystal tower in comparison to electrons. This can be observed in figures 5.25 and 5.26: e -like pions more often have $E1/E9$ - and $E2/E9$ -values larger than 0.9 in comparison to electrons. Such peculiar events arise as a large number of particles crosses the central APD producing a sizeable “fake” signal. The shower shape cut should be then able to eliminate at least part of the events with a large Nuclear Counter Effect, *i.e.* the events with a large fractional energy deposition $E_{\text{ECAL}}/|\mathbf{p}_i|$. For example we have chosen, for the 20 GeV data sample, to limit the $E1/E9$ values into the 0.72-0.86 range³ (see figure 5.25, where the chosen boundaries are shown with dotted lines).

These boundaries are fixed for the particular case of the test beam and would have to be re-adapted, as the real CMS set-up with magnetic field will be taken into account. Nevertheless, this study can give a hint of the possible improvements, obtained by making use of these shower shape cuts.

³These selection cuts have to be adjusted to the different beam energies. Moreover, we observed that in the GEANT Monte-Carlo, the shower shape boundaries can change depending on the choice of simulation parameters. Also the absence of an APD noise simulation can produce some differences.

E_{e^\pm}	20 GeV	50 GeV	120 GeV
$\epsilon_{\text{sh.sh.cuts}}$	88.6 %	92.2 %	95.5 %

Table 5.3: Electron selection efficiency $\epsilon_{\text{sh.sh.cuts}}$ after the shower shape cuts (obtained using the 1998 test beam data).

The electron efficiency after these cuts (applied to the '98 test beam data) is given in table 5.3. Figures 5.28-5.30 show the effectiveness of the shower shape cuts. The fraction of “fake electrons” candidates which are eliminated by the selection, increases strongly with increasing pion initial momentum. In addition, also an improvement in electron efficiency is observed by growing E_e (see table 5.3).

The results of the cuts can be seen also in table 5.4 for both test beam data and GEANT simulations. We chose to use a 450 MeV large Nuclear Counter Effect, because of the larger statistics given by the larger amount of e -like events. Differences in the cut effectiveness between simulations and data can be seen in the problematic $E_{3\times 3}/|\mathbf{p}_i|$ -region $0.85 - 1$. This is again due to an electron contamination of the pion beam, which is, of course, not rejected by the shower shape cuts.

Table 5.4 shows that the cuts eliminate more than 50% of the candidates with $E_{3\times 3}/|\mathbf{p}_i|$ -values larger than 0.85. Furthermore, above the value $E_{3\times 3}/|\mathbf{p}_i| = 1.2$, the selection succeeds even in eliminating more than 90% of these fake- e candidates (the percentage depends on the initial momentum).

We can therefore conclude that, with a 450 MeV Nuclear Counter Effect, between 0.03% (for 50 GeV) and 0.16% (for 20 GeV) of the charged pions still survive the selection cuts and can be confused as “electrons”, while keeping the electron efficiency above 85-90%.

With the expected Nuclear Counter Effect and a selection with the real CMS set-up, if one assumes the percentages of electron-like pions to be reduced by the same amount as with a 450 MeV NCE, then one would obtain few candidates in 10^4 events (cf. tables 5.4 and 5.5).

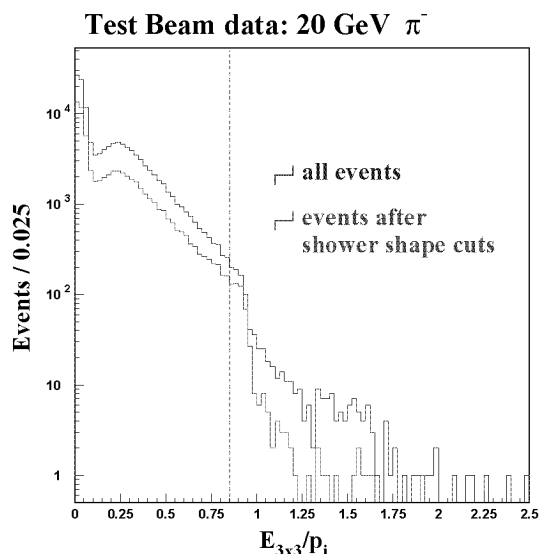


Figure 5.28: 1998 test beam data: fractional amount of energy left by charged pions with an initial momentum of 20 GeV, for events before and after the shower shape cuts.

5.5 Discrimination methods for isolated pions and electrons

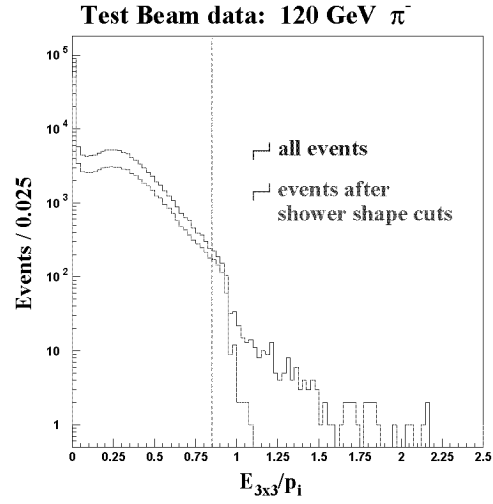
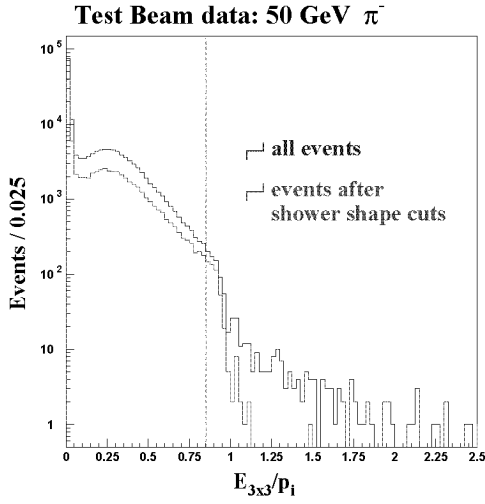


Figure 5.29: '98 test beam data (50 GeV): fractional amount of energy left by charged pions, for events before and after the shower shape cuts.

Figure 5.30: Same as figure 5.29, but for 120 GeV pions.

E_{π^-}	1998 Test beam data		GEANT (450 MeV NCE)	
	$0.85 < \frac{E_{3 \times 3}}{ \mathbf{p}_i } < 1.2$	$\frac{E_{3 \times 3}}{ \mathbf{p}_i } > 1.2$	$0.85 < \frac{E_{3 \times 3}}{ \mathbf{p}_i } < 1.2$	$\frac{E_{3 \times 3}}{ \mathbf{p}_i } > 1.2$
20 GeV	0.56 %	0.08 %	0.58 %	0.12 %
+sh.sh.cuts	0.34 %*	0.008 %	0.16 %	0.008 %
50 GeV	0.41 %	0.056 %	0.15 %	0.05 %
+sh.sh.cuts	0.25 %*	0.0005 %	0.03 %	$< 8 \times 10^{-5}$
120 GeV	0.39 %	0.041 %	0.12 %	0.02 %
+sh.sh.cuts	0.24 %*	$< 5 \times 10^{-6}$	0.04 %	$< 8 \times 10^{-5}$

Table 5.4: “Fake electrons” candidates before and after the shower shape cuts (labeled “sh.sh.cuts”). *affected by a small electron contamination.

5.6 Expected pion-electron misidentification and possible consequences for the CMS physics program

In this section we concentrate our analysis on the pion-electron misidentification obtained with the expected Nuclear Counter Effect within the CMS detector, that is ~ 100 MeV, with a 2-APD readout and with 4 photoelectrons per MeV produced in these photo-devices.

5.6.1 Expected ratios of “fake electron” candidates

Table 5.5 summarizes the amounts of “fake electrons” candidates originating from pion tracks for different pion initial momenta. These values are obtained from simulations with and without APD’s, *i.e.* either by taking the crystal energy alone or by adding also the Nuclear Counter Effect contribution.

The table shows clearly an increase of events with electron-like signals (*i.e.* $0.95 < E_{\text{ECAL}}/|\mathbf{p}_i| < 1.2$) when the Nuclear Counter Effect is included (the amount of events is, for instance, 3-times larger at 20 GeV). Moreover, a long tail appears, representing events with $E_{\text{ECAL}}/|\mathbf{p}_i|$ -values larger than 1.2.

As shown in section 5.1, with a two APD’s readout, the Nuclear Counter Effect is half of those obtained with a single APD per crystal. Nevertheless, the active surface of the sensors is doubled and a larger number of charged particles in the shower can then cross these photo-devices. One could expect a compensation between these two effects and is therefore not clear whether the 2-APD readout will diminish the probability for a $\pi - e$ confusion.

E_{π^-}	No Nuclear Counter Effect	100 MeV Nuclear Counter Effect (2 APD’s per crystal)	
	$\frac{E_{\text{ECAL}}}{ \mathbf{p}_i } > 0.95$	$0.95 < \frac{E_{\text{ECAL}}}{ \mathbf{p}_i } < 1.2$	$\frac{E_{\text{ECAL}}}{ \mathbf{p}_i } > 1.2$
5 GeV	0.006 ± 0.004 %	0.11 ± 0.02 %	0.05 ± 0.01 %
10 GeV	0.01 ± 0.005 %	0.06 ± 0.01 %	0.02 ± 0.01 %
20 GeV	0.018 ± 0.005 %	0.06 ± 0.01 %	0.007 ± 0.003 %
50 GeV	0.001 ± 0.001 %	0.011 ± 0.005 %	0.001 ± 0.001 %
100 GeV	< 0.005 %	0.008 ± 0.005 %	< 0.005 %

Table 5.5: GEANT (CMS design): charged pions as “fake electrons” candidates (the errors are purely statistical).

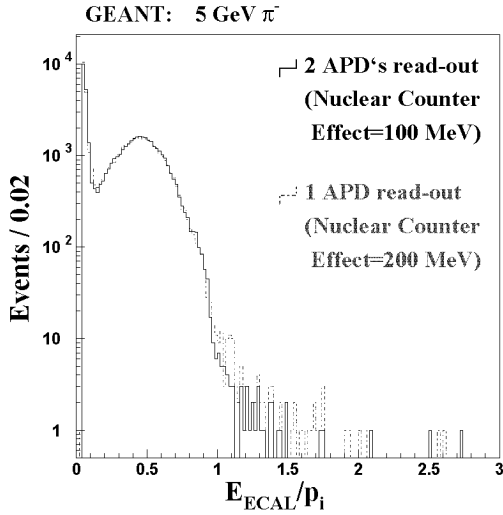


Figure 5.31: Comparison between a one- and a two-APD readout: the first case gives two times more events consistent with $E_{\text{ECAL}}/|\mathbf{p}_i| > 1$.

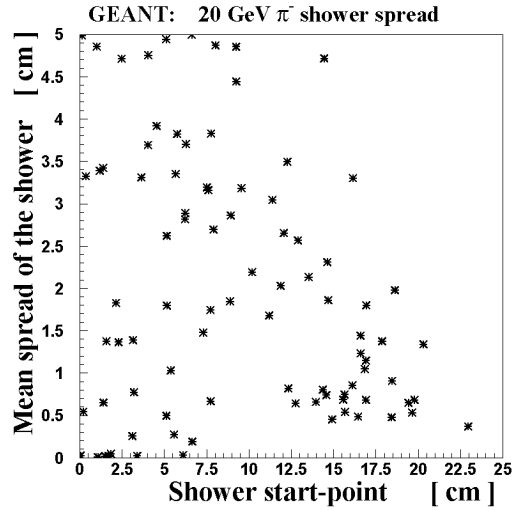


Figure 5.32: Spread of the π^- -showers in correspondence with the crystal rear face as a function of its start-point within the crystal (for events with $E_{\text{ECAL}}/|\mathbf{p}_i| > 0.8$).

Figure 5.31 shows a comparison between these two alternatives. Using one APD per crystal, for 5 GeV pions, 0.22% of the events have a fractional energy deposition $E_{\text{ECAL}}/|\mathbf{p}_i|$ larger than one, while with two APD's the percentage drops to 0.11%. The fact that the one-APD readout produces a longer tail at large $E_{\text{ECAL}}/|\mathbf{p}_i|$ -values could indicate that the shower spread of these rare events is not large enough to cross uniformly two APD's. Therefore, with two APD's per crystal, we would not get twice as much charged particles which cross these photo-devices and, since the Nuclear Counter Effect is half compared to the single APD case, its contribution to the measured energy will then be smaller.

To study this in more details we proceed in the following way: for every charged particle which leaves the ECAL, we plot its position in the $x - y$ -plane (*i.e.* the plane transverse to the beam). Then we find the center-of-mass of the obtained distribution and we calculate the mean distance of all points from this shower center. This value represents the width of the shower in correspondence with the crystal rear face (*i.e.* with the APD's position).

The result is shown in figure 5.32, where the shower spreads for events with a large energy deposition ($E_{\text{ECAL}}/|\mathbf{p}_i| > 0.8$), are plotted as a function of the shower start-point within the crystal. As this figure indicates, some electron-like events have shower spreads of about 0.5 cm (mostly events, whose shower started

very late (in the last centimeters of the crystal), and which are supposed to lead to the largest amounts of Nuclear Counter Effect, as shown in figure 5.5). This spread could be insufficient to cover the surface of both APD's, explaining the smaller size of the tail in the two-APD readout case.

5.6.2 Misidentified pions as source of new backgrounds?

After the probability of having a particular amount of energy left in the CMS ECAL by charged pions has been estimated, the possible consequences for the physics program can be studied.

Figure 5.33 shows a PYTHIA simulation of two possible physics channels, which could be affected by this background. The upper curve shows the transverse momentum spectrum of charged pions produced in one day (10^5 seconds) at low luminosity ($10^{33} \text{ cm}^{-2}\text{s}^{-1}$) via processes like $q_i q_j \rightarrow q_i q_j$, $q_i g \rightarrow q_i g$, $gg \rightarrow gg$ [90]. Their huge cross-sections (even larger if the Minimum Bias Events are included) represent a dangerous background for almost all physics signatures. If the isolation cuts used to select electrons are applied to charged pions, one gets

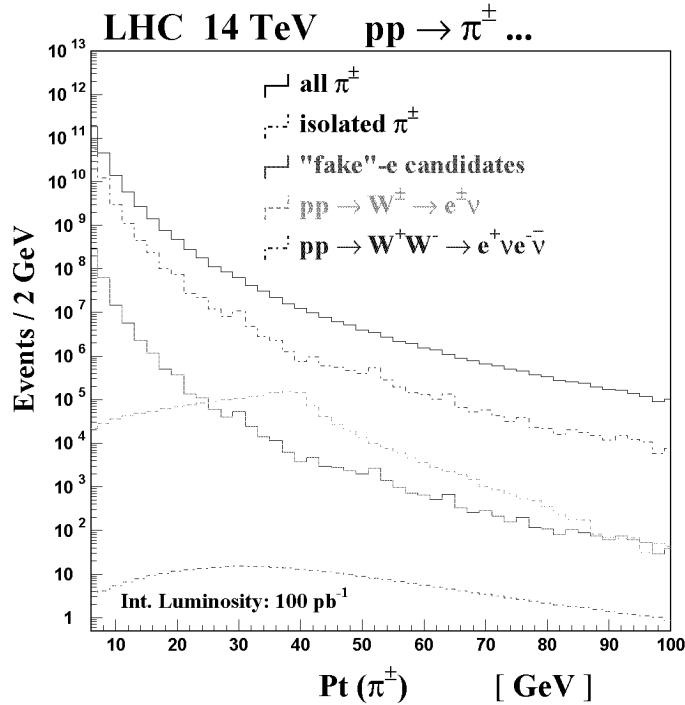


Figure 5.33: p_t -spectrum for different processes at LHC, where charged pions and electrons are generated (PYTHIA). These π and e have pseudorapidities $|\eta|$ smaller than 2.5 and one isolated pion out of thousand is considered to be a “fake” electron.

5.6 Expected pion-electron misidentification and possible consequences for the...

# of isolated leptons	physics channels [1 $l^\pm \equiv e^\pm$]	possible backgrounds	
1	$W^\pm \rightarrow e^\pm \nu$	QCD- $jet(s)$	+ Multiple Interactions + Minimum Bias Events
	$t\bar{t} \rightarrow e^\pm \nu b jj\bar{b}$	$b\bar{b} + jet(s)$	
2	$W^+W^- \rightarrow l^\pm \nu e^\pm \nu$	$W^\pm \rightarrow l^\pm \nu + jet(s)$	
3	$W^\pm Z^0 \rightarrow l^\pm \nu l^+ l^-$	$Z^0 \rightarrow l^+ l^- + jet(s)$	
		$t\bar{t} \rightarrow l^+ \nu l^- \bar{\nu} + jet(s)$	
4	$H \rightarrow ZZ^* \rightarrow l^+ l^- e^+ e^-$	$W^\pm Z^0 \rightarrow l^\pm \nu l^+ l^- + jet(s)$	

Table 5.6: Scheme of some physics channels and their possible backgrounds, arising from the eventuality of a $\pi^\pm - e$ misidentification.

the second curve shown in figure 5.33. With these criteria, about one pion out of hundred is found to be isolated. As pointed out in the previous section, among the charged pion tracks, the dangerous e -like candidates are about 0.1% of the total (without any further selection cut). If we apply this conservative value over the entire p_t -spectrum, we obtain the one labeled as “fake- e candidates”.

The total rate of these events above a p_t of 20 GeV (the threshold for the selection of isolated objects) is about 10 Hz. This is negligible compared to the “fake” trigger rate coming from π^0 's (which however are not charged and could be thus rejected with the help of the tracker). For charged pions the situation is different: since the tracker will not be able to separate them from electrons, they could still survive and become a real dangerous background for the off-line analysis, where a maximum rate of 100 Hz will be accepted (of course there will be other possible rejection cuts, like for example those presented in section 5.5). In figure 5.33 we include therefore also the event rates of electrons originating from two important physics processes like $pp \rightarrow W^\pm \rightarrow e^\pm \nu$ and $pp \rightarrow W^+W^- \rightarrow e^+ \nu e^- \bar{\nu}$ [91] to show qualitatively the size of the problem. The first one, for example, above a p_t of 20 GeV has a rate of few Hz, so that one obtains about the same amount of isolated electrons as for the “fake-electrons” which arise from QCD processes. If one then wants to set the p_t threshold at 10 GeV, the rate of e -like π^\pm 's increases to hundreds of Hz! Moreover, one has also to consider that the tail in the $E_{\text{ECAL}}/|\mathbf{p}_i|$ distribution (see figures 5.14-5.16) allows for example a 10 GeV π^\pm (which has a much larger production cross section) to have the same ECAL energy as a 20 GeV electron. In principle, with the information provided by the tracker, almost all these candidates should be rejected. However, these pions could at least affect the event selection of the first level trigger.

In general, the pion-electron misidentification should be considered as an addi-

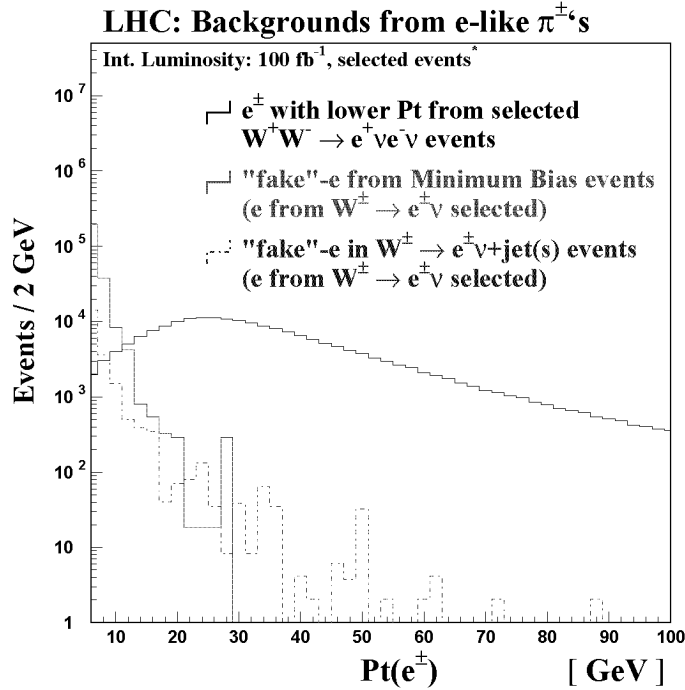


Figure 5.34: p_t -spectrum of the electron with lower transverse momentum within a $WW \rightarrow eevv$ event. In addition, the figure includes two possible "fake- e " backgrounds, which produce a similar signature.

tional source of background for several types of events which require isolated electrons. Table 5.6 shows some of these physics signatures which, depending on the number of isolated leptons in the event, could be affected by a "fake- e " background. As indicated in the table, the number of e -like pions can be also increased with the contributions of multiple interaction processes within the pp -collisions and of minimum bias events. One example is given in figure 5.34, where two possible backgrounds for the $WW \rightarrow eevv$ process are shown: in both cases one electron comes directly from a $W \rightarrow e\nu$ event, while the second one arises either from minimum bias events or from the high- p_t debris in the $W \rightarrow e\nu + \text{jet}(s)$ event. Also in this case, the size of the background depends strongly on the choice of the lower p_t -threshold for the electron selection: above 20 GeV one should be "safe", while by $p_{tmin}^e \approx 5$ GeV the rate of "fake WW events" would be huge.

As these studies give only an indication of the possible new backgrounds arising from the $\pi^{\pm} - e$ misidentification, we encourage further analysis which would take into account the exact amount of Nuclear Counter Effect and which would include a full detector simulation for signal and background processes. The presence of the 4 Tesla magnetic field could further complicate the task of a $\pi^{\pm} - e$ discrimination.

Chapter 6

Next-to-leading order (NLO) effects in the vector boson pair production

The importance of the weak boson pair production for the search of a Standard Model Higgs was already pointed out in section 2.3. In addition, the diboson production is also an interesting physics channel, since it provides a way to test the vector boson self-interactions, where, for example, anomalous effects might appear.

For each of these physics signatures, the starting point of the analysis is the Standard Model expectation for the non-resonant weak boson pair production. Any deviation from this continuum diboson background could indicate a signal for physics beyond the Standard Model. The success of the above-listed physics studies depends therefore on the accuracy of the Standard Model cross sections calculations for W^+W^- , $W^\pm Z$ and ZZ events.

Vector bosons can only be measured through those detectable decay products that are compatible with the experimental acceptance. To estimate the number of events which survive the experimental cuts, it is important to properly treat the production and decay of vector bosons, including all the spin correlations and the next-to-leading-order (NLO, or $\mathcal{O}(\alpha_s)$) QCD corrections.

Past NLO calculations of the hadronic diboson production [92, 93] have shown that the $\mathcal{O}(\alpha_s)$ corrections could be large in some specific kinematical regions. It is thus mandatory to include higher order calculations, not only to get a precise estimate of the overall production cross-section, but also to obtain realistic values for the significance of several specific signal processes.

For the description of the weak boson pair productions W^+W^- , $W^\pm Z^0$ and $Z^0 Z^0$, with their subsequent decays into leptons, two numerical parton-level Monte Carlo programs have become available during 1999: MCFM [17] and DKS [18]. These packages consider the production of the leptons in the double resonance approximation with complete next-to-leading-order corrections.

To investigate the accuracy of PYTHIA (which is based on LO matrix elements [94] with parton shower) describing the diboson cross sections at the LHC, a comparison is performed with the DKS parton level NLO Monte Carlo. In particular, we want to investigate if, for some optimization cuts (e.g. high diboson p_t or diboson mass regions), the corrections provided by NLO diagrams are large, and thus PYTHIA predictions are not reliable.

6.1 NLO generators

At Born-level, or leading order (LO), the cross sections for W^+W^- , $W^\pm Z$ and ZZ pair production were computed twenty years ago by treating the W and Z as stable particles and summing over their polarization states. These cross sections were evaluated using completeness relations to simplify the sum, thus spin and decay correlations were neglected. The Monte Carlo program DKS [18] includes at $\mathcal{O}(\alpha_s)$ the full lepton decay correlations in the narrow-width approximation for the hadronic production of W^+W^- , $W^\pm Z$ and ZZ pairs. The implemented $\mathcal{O}(\alpha_s)$ corrections can be subdivided in real (tree graphs with an additional gluon in either the initial or the final state) and virtual corrections (one-loop amplitudes that interfere with the Born amplitude). They rely on reference [95] for all required matrix elements.

Other Monte Carlo programs also include all spin correlations to the next-to-leading order in $\mathcal{O}(\alpha_s)$ for the vector boson pair production. For example, singly-resonant diagrams are implemented also in the MCFM [96] program, but without assuming the narrow-width approximation.

The effects of anomalous couplings on various distributions in the WW and WZ production have been carried out at $\mathcal{O}(\alpha_s)$ also in similar studies [97, 98]. However, in the DKS Monte Carlo the spin correlation effects from the finite virtual contributions are added.

6.2 Results

In all results presented in this analysis, it is assumed that the vector bosons always decay leptonically. In both Monte Carlo programs we use the CTEQ 4M parton distribution set [99] and the cross section values are given for the scale $\mu = (M_{V_1} + M_{V_2})/2$, where M_{V_i} are the masses of the two produced vector bosons. If the DKS Monte Carlo is run at Born-level (*i.e.* by including only the LO diagrams), we obtain very good agreement with the total cross sections given by PYTHIA (see, for instance, figure 6.1 or the $\sigma \times BR$ values in tables 6.1-6.3). On the other hand, as shown in this same figure, the one-loop corrections to the total cross sections are of the order of 50% of the leading-order term.

A more accurate comparison between PYTHIA and DKS requires some remarks.

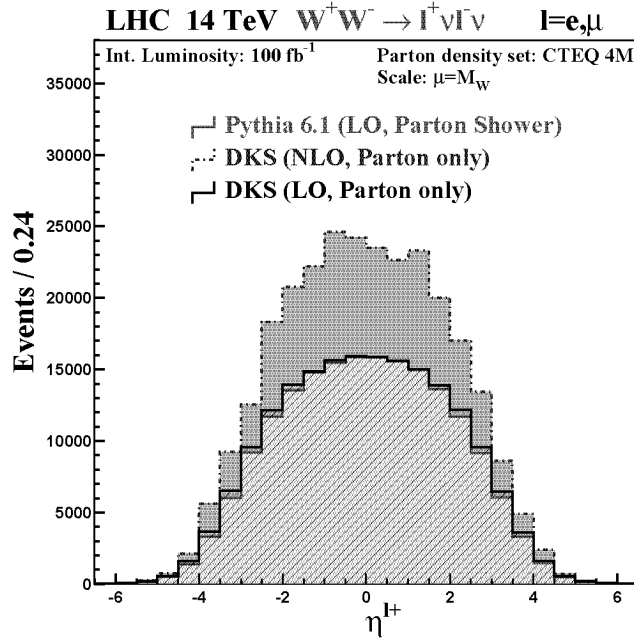


Figure 6.1: Lepton η -distribution in $WW \rightarrow l\nu l\nu$ events. While at LO the PYTHIA results are consistent with the DKS ones, the K-factor introduced by one-loop corrections is of the order of 1.5.

The DKS Monte Carlo produces only events at parton level, that is without including the showering process and initial state radiation. Therefore, the only “jet” generated by DKS coincides always with the hard gluon/quark emitted within the bulk of the event. This jet arises thus from those events which have been added by including the NLO diagrams. In contrast, PYTHIA does not simulate these NLO processes but features instead the complete parton showering machinery. The soft jet(s), observed in WW , WZ or ZZ events generated by PYTHIA, originate from initial state radiations and/or from other QCD processes in the pp -event (multiple scattering or minimum bias events). Consequently, these soft emissions give rise to a Gaussian smearing of the diboson transverse momenta around zero, while in DKS generated events the $p_t(V_1V_2)$ is in average much harder and is directly produced by the NLO q/g emission. A comparison between the two Monte Carlo is therefore possible for large diboson transverse momenta (typically $p_t(V_1V_2) > 100$ GeV) and instead difficult at small values ($p_t(V_1V_2) \lesssim 20$ GeV), where the distribution is dominated by “soft jets”, which are not included in the DKS generator.

In events simulated by PYTHIA, jets are reconstructed following the criteria presented in subsection 3.3.2 but are limited to a maximal pseudorapidity $|\eta| = 3$, while leptons are selected following the CMS acceptance criteria (see subsection 3.3.2), where a p_t larger than 20 GeV and a pseudorapidity $|\eta| < 2.5$ are required.

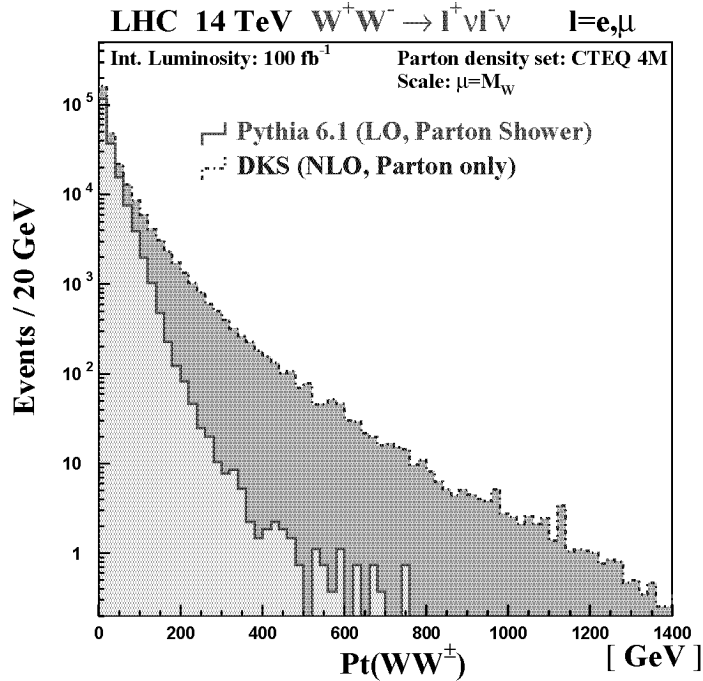


Figure 6.2: Transverse momentum distribution of WW bosons pairs, simulated with both the PYTHIA and the DKS Monte Carlo generators and using the CTEQ 4M structure function set.

In the tables given in this chapter, the statistical errors are not given. This is consistent with the aim of this study, which is not to give the exact K factors for all possible selection cuts, but rather to show qualitatively what could be the size of the NLO corrections for some specific cuts.

6.2.1 NLO effects in the WW production

To investigate the possible effects induced by the introduction of next-to-leading order diagrams, we look for kinematical regions which could be “NLO-sensitive”. First of all, the additional emission of hard quarks or gluons in processes like $qg \rightarrow V_1 V_2 q$ should lead to enhanced transverse momenta for the diboson system [97, 100]. Figure 6.2 shows, for example, the p_t distribution of WW pairs. A comparison between PYTHIA and DKS indicates a large difference for large p_t^{WW} values. This diboson transverse momentum is balanced by an equivalent p_t for the jet(s). The numerical results are reported in table 6.1. The K-factor increases from 1.5, without any $p_t^{jet(s)}$ -cut, up to values of about 60, if the jets are required to have a p_t larger than 150 GeV.

	PYTHIA [pb]	DKS (LO) [pb]	DKS (NLO) [pb]	K-factor
$\sigma \times BR$	3.704	3.79	5.56	1.5
selected leptons	1.704	1.71	2.58	1.54
jet veto	1.125	-	1.49	1.32
$20 \text{ GeV} < p_t^{jet} \leq 150 \text{ GeV}$	0.568	-	0.942	1.66
$150 \text{ GeV} < p_t^{jet} \leq 400 \text{ GeV}$	2×10^{-3}	-	0.135	67
$p_t^{jet} > 400 \text{ GeV}$	2.8×10^{-4}	-	1.69×10^{-2}	~ 60

Table 6.1: Evolution, through the different cuts, of the $W^+W^- \rightarrow l^+\nu l^-\bar{\nu}$ cross sections (in pb) and of the corresponding K-factors.

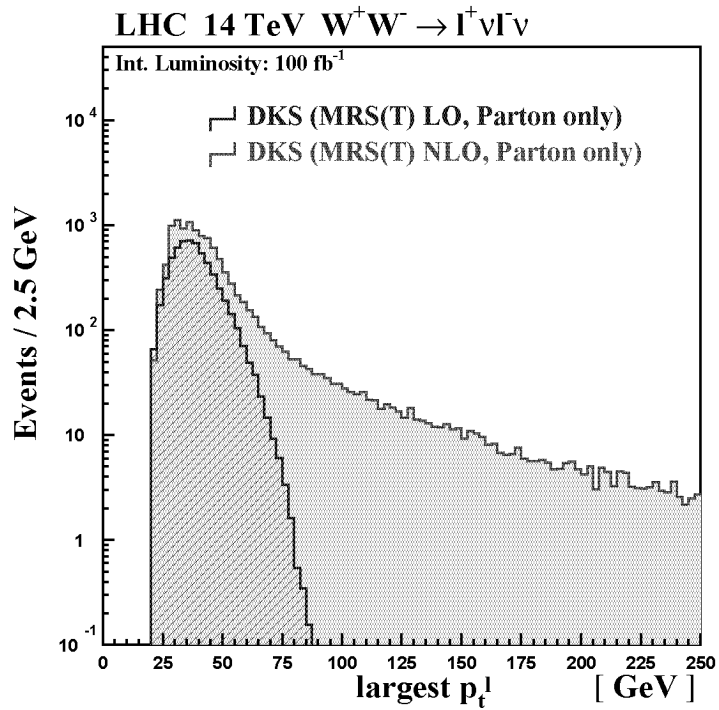


Figure 6.3: Largest lepton p_t in WW events. The results are obtained by running the DKS generator either with or without including the NLO corrections.

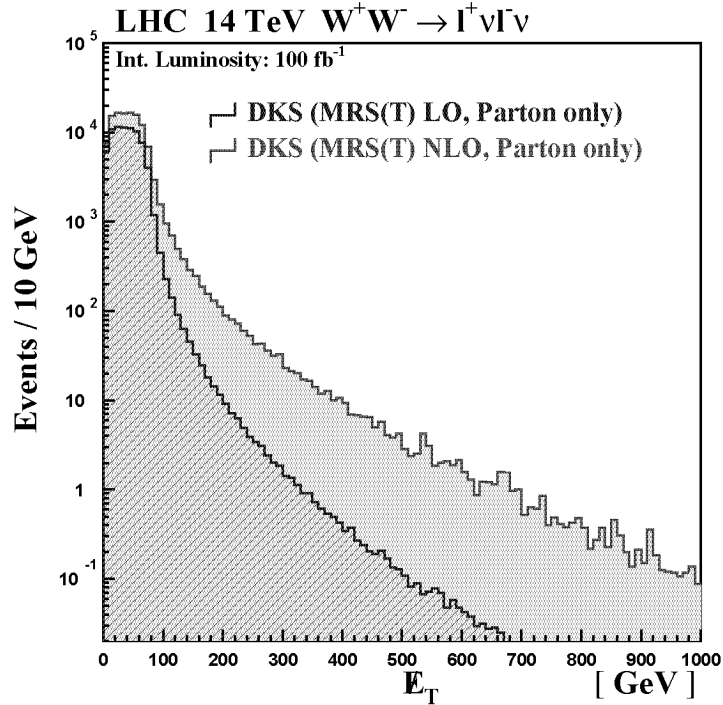


Figure 6.4: Missing transverse energy in the WW production. The events are obtained by running the DKS generator either with or without including the NLO corrections.

The transverse momentum of the diboson system (or of the jets) is not the only variable affected by large NLO corrections. Significant differences between NLO and LO expectations could show up also in other kinematical regions. For instance, in “common” variables like the lepton p_t , the lepton pair invariant mass m_{ll} , the missing transverse energy \cancel{E}_t , the maximal transverse momentum of the two charged leptons p_{tmax}^l , the lepton pseudorapidities η^l and so on.

Figures 6.3 and 6.4 show, respectively, the p_{tmax}^l and \cancel{E}_t distributions. These histograms are simulated with only the DKS generator, either with or without including the next-to-leading order diagrams. Therefore, in this example, the differences between the two curves are produced exclusively by the $\mathcal{O}(\alpha_s)$ corrections, while the effects given by showering and initial state radiation are not included. As for the p_t^{WW} distribution, the K-factor corresponds to about 1.5, for small p_{tmax}^l or \cancel{E}_t values, and reaches very large values for large p_{tmax}^l or \cancel{E}_t . Differences could also appear in some more specific variables. Very often, the cuts which are used in physics studies involving the diboson production are rather elaborated. For example, one can select a region in the distribution of the difference in pseudorapidity $\Delta\eta^l = \eta^{l^-} - \eta^{l^+}$ between leptons, of their distance in polar angle $\cos\theta_{ll}$ or of their relative azimuthal angle $\cos\phi_{ll}$.

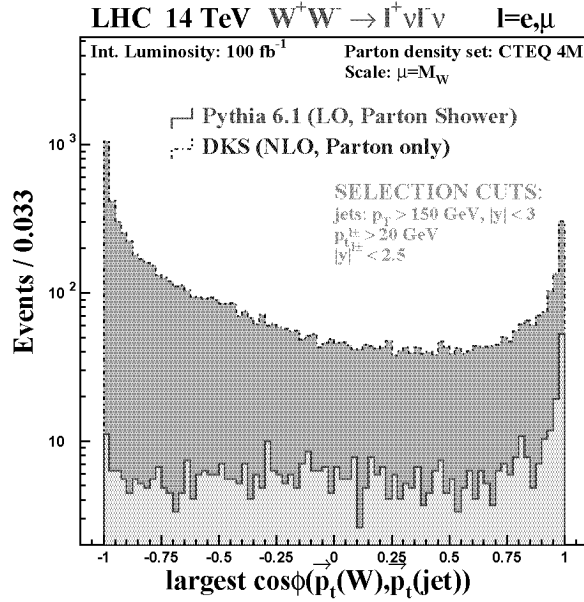


Figure 6.5: PYTHIA-DKS comparison which shows the smallest angle between one of the W 's and the jet in WW events. The two leptons are required to be consistent with the detector acceptance and the jet should have a p_t larger than 150 GeV.

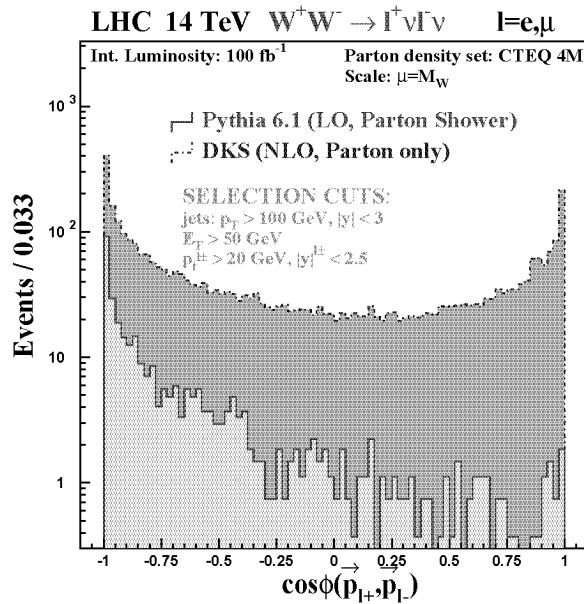


Figure 6.6: PYTHIA-DKS comparison of the azimuthal angle between the two leptons produced in WW events. The following additional cuts are used: $p_t^{\text{jet}} > 100 \text{ GeV}$ and $\cancel{E}_T > 50 \text{ GeV}$.

Figures 6.5 and 6.6 show two possible examples of such event selections. In the first figure, after having selected only WW events with two detectable leptons and a jet with $p_t > 150$ GeV, we plot the smallest angle between one of the W 's and the jet (as schematically illustrated in figure 6.7).

As can be observed, the effect of an additional cut on $\cos(\vec{p}_t^W, \vec{p}_t^{jet})$ would be very different, if the NLO diagrams are included in the simulation. The same remark applies to figure 6.6: if in a WW event we require the presence of two leptons, a jet with p_t larger than 100 GeV and a missing transverse energy of at least 50 GeV, the azimuthal angle between the two leptons is affected by large NLO corrections.

Both examples show that analysis cuts, which use such variables, need to employ NLO Monte Carlo programs for the diboson production.

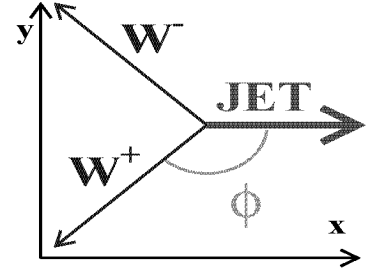


Figure 6.7: Azimuthal angle between one of the W 's and the jet in WW events.

6.2.2 NLO effects in the WZ production

For WZ events, similar effects as for the WW production can be observed. If the DKS generator is run at LO, it reproduces the total WZ -cross section given by PYTHIA with a difference of about 1%. If instead the NLO corrections are taken into account, the $\sigma \times BR$ value increase to 171% of the LO one.

	PYTHIA [pb]	DKS (LO) [pb]	DKS (NLO) [pb]	K-factor
$\sigma \times BR$	4.35×10^{-1}	4.34×10^{-1}	7.42×10^{-1}	1.71
selected leptons	1.45×10^{-1}	1.48×10^{-1}	2.77×10^{-1}	1.91
jet veto	9.47×10^{-2}	-	1.31×10^{-1}	1.39
$20 \text{ GeV} < p_t^{jet} \leq 150 \text{ GeV}$	4.91×10^{-2}	-	1.27×10^{-1}	2.3
$150 \text{ GeV} < p_t^{jet} \leq 400 \text{ GeV}$	9.33×10^{-4}	-	2.8×10^{-2}	30
$p_t^{jet} > 400 \text{ GeV}$	6.5×10^{-6}	-	4.63×10^{-3}	~ 700

Table 6.2: Evolution, through the different cuts, of the $W^\pm Z^0 \rightarrow l^\pm \nu l^+ l^-$ cross sections (in pb) and of the corresponding K-factors.

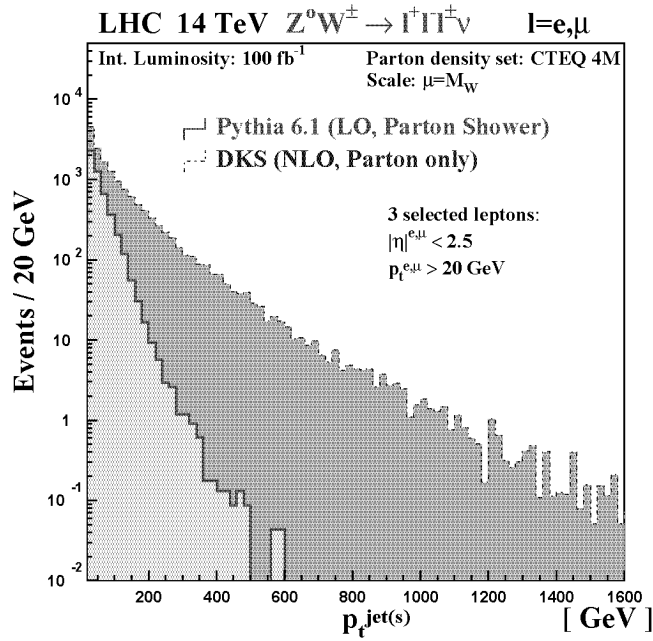


Figure 6.8: Transverse momentum distribution of jets produced in WZ events. All 3 leptons decayed from both weak bosons have been selected.

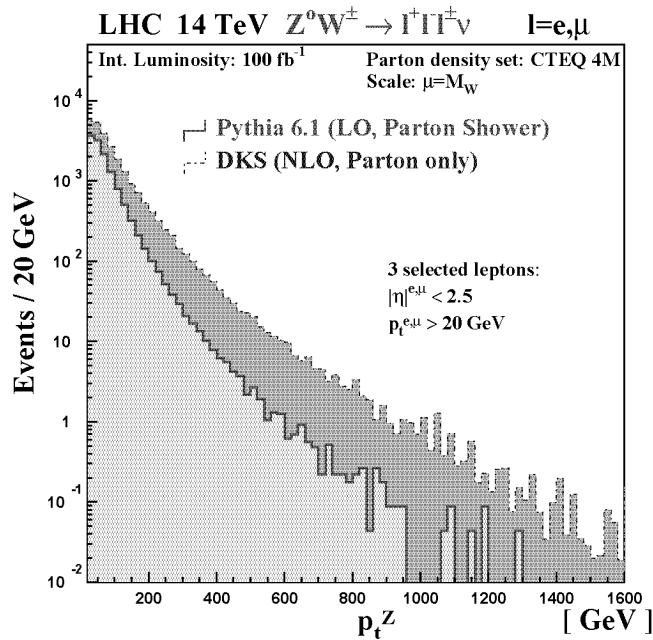


Figure 6.9: p_t -distribution of the Z bosons produced in WZ events. All 3 leptons decayed from both weak bosons have been selected.

As regards to some common variables, figure 6.8 shows, for example, the p_t distribution of jets (which balances the transverse momentum of the diboson system). In case of the WZ production, the K-factors at large p_t^{jet} -values are even larger than in WW events.

The WZ cross sections observed in some selected p_t^{jet} -regions are reported in table 6.2.

The ratio between DKS and PYTHIA generated events (*i.e.* between σ_{NLO} and σ_{LO}) leads for instance to K-factors larger than 500 above a p_t of 400 GeV¹.

Smaller NLO effects are also visible in the transverse momentum distribution of the Z bosons, as shown in figure 6.9. The K-factor corresponds to ~ 1.5 at small p_t^Z values, and reaches $K \approx 10$ at larger p_t^Z .

6.2.3 NLO effects in the ZZ production

The non-resonant ZZ production was already known to be subject to higher-order QCD corrections [101]. The evolution of the ZZ -cross sections, through some illustrative selection cuts, is reported in table 6.3. While at LO the DKS expectations are very close to the PYTHIA ones, the addition of the NLO diagrams leads to an inclusive K-factor of about 1.4. Furthermore, like for the WW or WZ signatures, if the analysis is restricted to the higher p_t^{jet} (or p_t^{ZZ}) domain, much larger K-values are reached.

	PYTHIA [pb]	DKS (LO) [pb]	DKS (NLO) [pb]	K-factor
$\sigma \times BR$	5.13×10^{-2}	5.31×10^{-2}	7.07×10^{-2}	1.38
selected leptons	1.79×10^{-2}	1.84×10^{-2}	2.55×10^{-2}	1.42
jet veto	1.15×10^{-2}	-	1.58×10^{-2}	1.38
$20 \text{ GeV} < p_t^{jet} \leq 150 \text{ GeV}$	6.26×10^{-3}	-	8.79×10^{-3}	1.4
$150 \text{ GeV} < p_t^{jet} \leq 400 \text{ GeV}$	1.33×10^{-4}	-	8.23×10^{-4}	6.2
$p_t^{jet} > 400 \text{ GeV}$	1.5×10^{-6}	-	7.78×10^{-5}	~ 50

Table 6.3: Same as table 6.2, but for the $ZZ \rightarrow l^+l^-l^+l^-$ process.

¹The results given for the high- p_t regions are affected by the small statistical significance.

6.3 Discussion

The simulations of diboson events presented in this chapter point out the importance of taking care of the possible NLO corrections in the weak boson pair production. This is especially true at least in some specific kinematical regions, where the $\mathcal{O}(\alpha_s)$ diagrams have sizeable effects. Therefore, for signal optimizations where the weak boson pair production is an important background process, one should be aware of the possible inaccuracies due to the use of a LO generator like PYTHIA. This is especially true for complicated cuts, where it is difficult to judge whether the effects are large or not.

An example, where important $\mathcal{O}(\alpha_s)$ effects could eventually appear, is the Higgs search through the decay mode $H \rightarrow ZZ \rightarrow 4l$. Past studies [102] proposed to improve the significance of the signal by using a minimal cut on the p_t of the harder Z boson. A similar approach is also used by the ATLAS collaboration [103], which proposes, as a moderate requirement, to set $p_t^{max}(Z_1, Z_2) > M_H/3$. On the basis of the available simulations, they also stated that a harder cut would improve even further the signal significance. However, the efficiency of such selection cuts for the signal events is subject to possibly significant higher-order

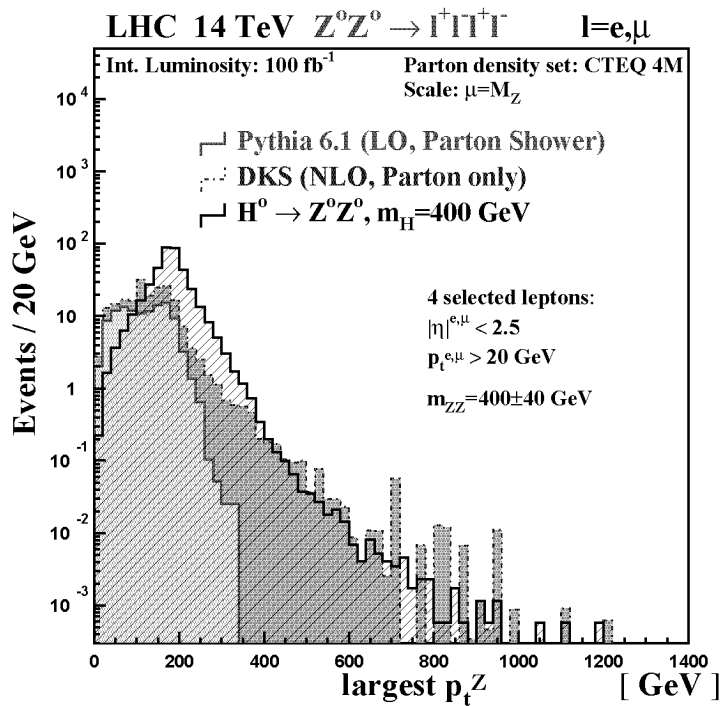


Figure 6.10: p_t -distribution of the harder Z among ZZ events. Both the resonant $H \rightarrow ZZ$ ($M_H = 400$ GeV) and the continuum ZZ productions are shown (the latter is simulated either with PYTHIA or with DKS).

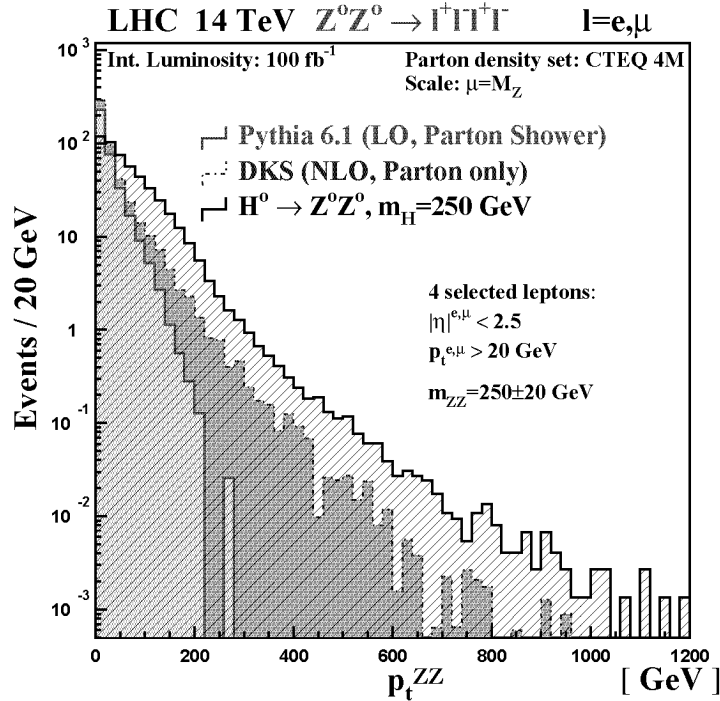


Figure 6.11: Transverse momentum distribution of ZZ pairs, originating from 250 GeV Higgs bosons (for events where M_{ZZ} is consistent with the Higgs mass). The figure includes both the DKS and PYTHIA expectations for the non-resonant ZZ background, which is thus simulated either with or without NLO corrections.

corrections, which increase substantially the continuum ZZ -background by high $p_t^{max}(Z_1, Z_2)$ values. This is shown in figure 6.10 for a 400 GeV Higgs. If the NLO corrections are included, the $p_t^{max}(Z_1, Z_2)$ distribution for the non-resonant background follows much more closely those of the signal, reducing considerably the gain of a harder $p_t^{max}(Z_1, Z_2)$ -cut. In other words, the benefits of this strategy may not be as large as one would expect by fully trusting the PYTHIA results.

Figures 6.11 and 6.12 show the p_t^{ZZ} distribution in $H \rightarrow ZZ \rightarrow 4l$ events, for Higgs masses equal to, respectively, 250 and 400 GeV. In both figures, only ZZ masses consistent with the corresponding M_H have been selected.

As largely expected from the previous results, in the high- p_t^{ZZ} domain the continuum ZZ -background generated by DKS is in average much more important than what PYTHIA predicts. In particular, above $p_t^{ZZ} \approx 100$ GeV, the weak boson fusion signatures $qq \rightarrow qqH \rightarrow jjZZ$ show a p_t^{ZZ} distribution, which is almost identical to those expected from the NLO-background (see figure 6.12). Again, the efficiency of a p_t^{ZZ} -selection cut could be very different if the $\mathcal{O}(\alpha_s)$ corrections are included.

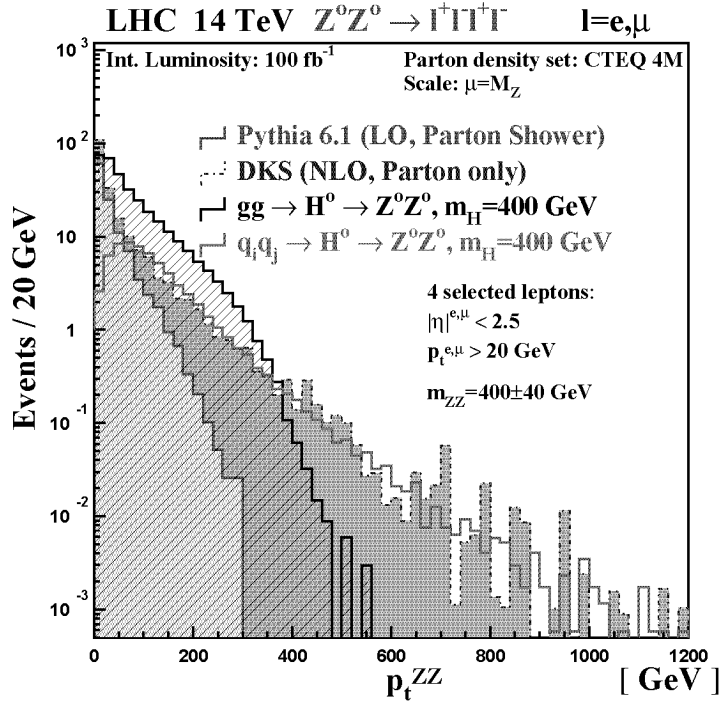


Figure 6.12: Same as figure 6.11, but for 400 GeV Higgs bosons. Both the expectations for the inclusive $gg \rightarrow H$ and the weak boson fusion $qq \rightarrow qqH$ production modes are shown. The latter gives a p_t^{ZZ} distribution which is almost identical to those expected from the continuum NLO-background.

Finally, to extend the visibility of inclusive $H \rightarrow WW \rightarrow l\nu l\nu$ signatures into the $200 \text{ GeV} \leq M_H \leq 400 \text{ GeV}$ domain, one could select events featuring reconstructed jet(s) and use the p_t^{jet} and spin correlations to put new kinematics constraints. Following this strategy, for the simulations shown in figure 6.14 we set a minimal p_t^{jet} -cut of 100 GeV, as well as some additional standard acceptance cuts (given in the figure).

As can be observed, large NLO corrections affect the m_C distribution of the continuum WW background. Nevertheless, the Higgs signal events could be well visible above both LO and NLO non-resonant backgrounds. The $t\bar{t}$ background has however not been included.

In conclusion, for diboson events, the simulations presented in this chapter show qualitatively the size of K factors found in some specific kinematical regions. Better NLO expectations of the event rates for these regions should help to better understand the Standard Model non-resonant production of vector boson pairs.

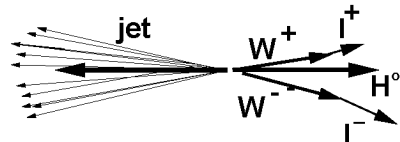


Figure 6.13: Schematic presentation of a $H \rightarrow WW \rightarrow l\nu l\nu + jet$ event.

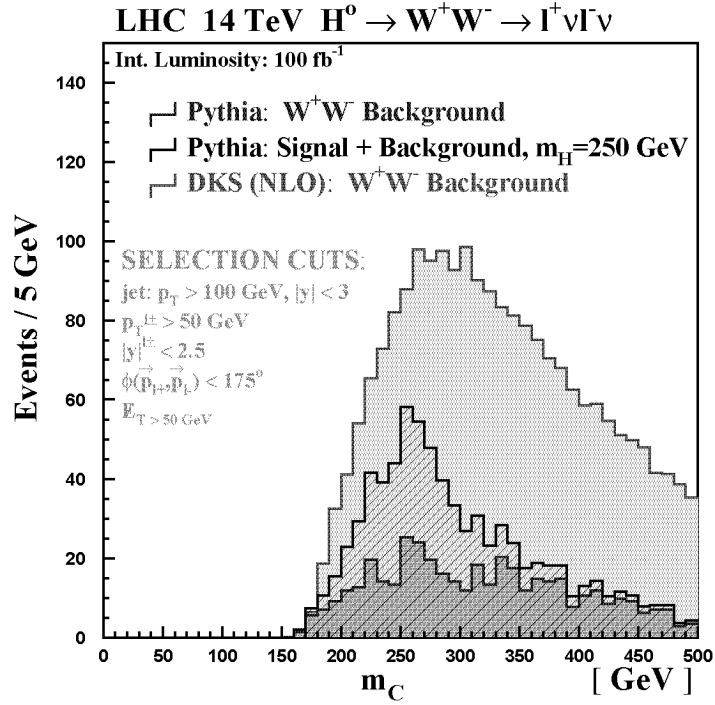


Figure 6.14: Reconstructed transverse mass $m_C = \sqrt{p_{tU}^2 + m_U^2} + \cancel{E}_t$. The three curves show (from bottom to top) the WW background (PYTHIA), the sum of signal $H \rightarrow WW \rightarrow l\nu l\nu$ ($M_H = 250$ GeV) and continuum WW background (PYTHIA), or the NLO WW background alone (DKS).

Resonant diboson signatures should then profit of the new background knowl-edges.

Chapter 7

Observing light Standard Model Higgs through weak boson fusion signatures

The discovery of one or more Higgs scalar(s) at LHC will be followed by other important tasks. Among them, there will be a more systematic investigation of the Higgs boson properties, *i.e.* the determination of its couplings to fermions and gauge bosons.

Recent parton level studies [25, 26] concluded that a $115 \text{ GeV} \leq M_H \leq 200 \text{ GeV}$ Higgs can be separated from background by demanding the tagging of the two additional quark jets in the reaction $qq \rightarrow qqH$ and by central jet vetoing techniques. These analysis showed that, with the additional use of the $H \rightarrow WW \rightarrow l\nu l\nu$ event characteristics, this signature yields a significant Higgs boson signal with an integrated luminosity of few fb^{-1} .

In this chapter, we repeat these studies using PYTHIA and in addition, we propose a similar set of cuts to cover also the $200 \text{ GeV} \leq M_H \leq 300 \text{ GeV}$ mass domain.

7.1 Towards a determination of the Higgs couplings

The Weak Boson Fusion $qq \rightarrow qqH$ promises to be a copious source of Standard Model Higgs at LHC. Together with the Higgs production in the gluon fusion process, $gg \rightarrow H$, it will allow a measurement of the $t\bar{t}H/WWH$ coupling ratio [20], as schematically illustrated in figure 7.1. In the Standard Model, the fermion triangles leading to an effective Hgg coupling are expected to be dominated by the top-quark (see figure 2.4a). Therefore, the ratio between the above processes probes the Higgs coupling to quarks relative to the HWW ones. This measurement constitutes a first step towards a determination of the Higgs properties.

Other channels, like for instance WH or $t\bar{t}H$ associated production, with subsequent decay $H \rightarrow b\bar{b}$ or $H \rightarrow \gamma\gamma$, would provide additional informations on Higgs coupling ratios, allowing a measurement of the Htt , Hbb , $H\tau\tau$ and HWW , HZZ , $H\gamma\gamma$ couplings.

Similarly to previous studies regarding weak boson fusion signatures [21, 22, 23], the cited studies [25, 26] proposed to exploit the forward jet tagging to select signal events. Such tagging is also required to reduce possible backgrounds from diboson and $t\bar{t}$ productions to negligible levels. Another feature of the weak boson fusion signal is the lack of color

exchange between the initial-state quarks, which leads to suppressed hadron production in the region between the two tagging jets. Most background processes, instead, involve typically a color exchange in the t -channel and feature therefore a larger hadronic activity between the tagging jets. Via a veto of soft jet activity in the central region [104], one can thus achieve a significant background reduction. Moreover, the anti-correlation of the W spins becomes an important additional tool for separating the signal from various backgrounds.

Previous studies [25, 26] were performed using full tree-level matrix elements for the signal and backgrounds, but only at parton-level. In other words, the final state partons have been identified with observable jets. The goal of our study is then to perform the analysis at PYTHIA-level, using selection criteria which are consistent with the CMS design. As this Monte Carlo includes the full showering processes as well as the gluon radiation pattern from the initial state quarks, it allows us to take into account hadronization and decay effects. The analysis proceeds using cone algorithms for the reconstruction of jet and isolated leptons. The drawback of using PYTHIA is that in this Monte Carlo generator some possible background processes are not included. Particularly, for the $t\bar{t}$ production, all QCD corrections which lead to additional real parton emissions (*i.e.* which deliver the $t\bar{t}$ -pair plus one or more hard jets) are not well described. This is especially regrettable, since the QCD $t\bar{t} + jets$ production becomes the largest source of background for the weak boson fusion process. The same happens for the W pair production with additional hard jets, which is also missing in the PYTHIA “menu”. For this reason, I am not presenting a complete “visibility” study, *i.e.* the goal is not to give the significance of the signal at PYTHIA level, but

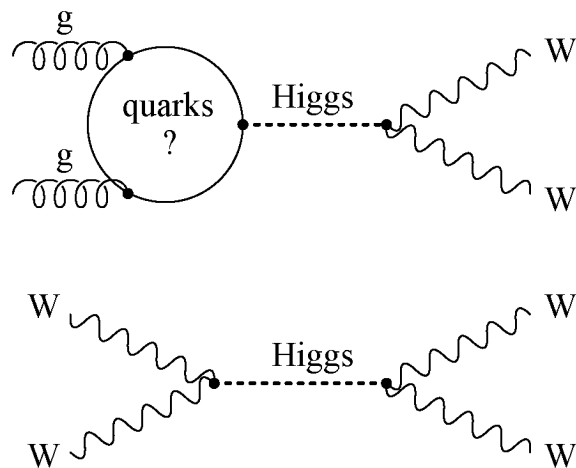


Figure 7.1: Ratios between such Higgs signatures allow to test the Higgs couplings to quarks and gauge bosons.

rather to check how jet reconstruction, isolation criteria, initial state radiation and underlying events are affecting the signal and background efficiency.

7.2 Event selection

In lowest order, the Feynmann-diagram related to the vector boson fusion ($qq \rightarrow qqH$) process is shown in figure 7.2. Both weak bosons are emitted by the incoming quarks, which give rise to two energetic jets with large rapidity, while the Higgs and its decay products will be more central.

For this study, W decays to electrons and muons are used.

First, we want to reproduce the analysis given in reference [25], which was optimized for a Higgs boson in the range $130 \text{ GeV} \leq M_H \leq 200 \text{ GeV}$. Subsequently, we deal with the second study [26], which focused on the mass region $110 \text{ GeV} \leq M_H \leq 140 \text{ GeV}$.

Two slightly different set of cuts are followed: the first one is applied for masses above 140 GeV and the latter up to this value. In addition, the Higgs range is extended upwards, for a Higgs mass between 200 GeV and 300 GeV.

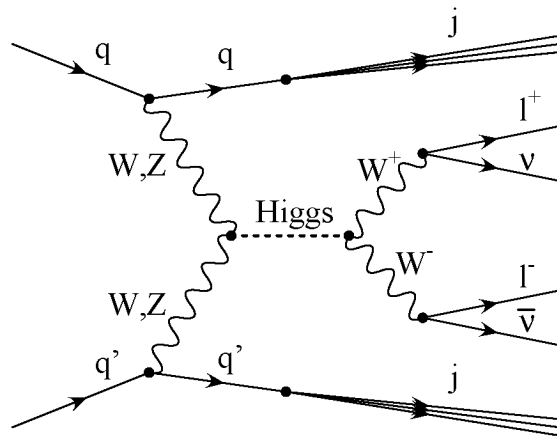


Figure 7.2: Feynmann diagram of the $qq \rightarrow qqH \rightarrow jj l\nu l\nu$ signal process.

7.2.1 $140 \text{ GeV} \leq M_H \leq 200 \text{ GeV}$

Any processes giving two jets, two oppositely charged leptons and missing transverse momentum are potential background candidates to this particular Higgs signal.

The main physics background arises from the $t\bar{t}(+jets)$ production. This is due to the large top production cross section at LHC and to the branching ratio $BR(t \rightarrow Wb)$, which is essentially 100%. Other important background processes are the $WWjj$ production, which includes the electroweak boson fusion with a momentum and color structure identical to the signal, and the $\tau^+\tau^- + jet(s)$ background. For the same flavor decays $H \rightarrow WW \rightarrow e^+\nu e^-\bar{\nu}$ or to $\mu^+\nu\mu^-\bar{\nu}$, one has to take into account also the background process $ZZjj \rightarrow l^+l^-\nu\bar{\nu}$ and the EW $lljj$ final states, where the missing transverse momentum is entirely due to detector effects. The expected cross sections at parton level are given in table 7.1 for both the signal (with a different flavor decay $H \rightarrow WW \rightarrow e^\pm\nu\mu^\mp\nu$) and

cuts	Hjj	$t\bar{t} + jets$	$WWjj$		$\tau\tau jj$		S/B
			QCD	EW	QCD	EW	
forward tagging I-V	17.1	1080	4.4	3	15.8	0.8	$\approx 1/65$
b -jet rejection VI		64					1/5.1
m_{jj} & “V-A” VII-IX	11.8	5.5	0.54	0.5	3.6	0.4	1.1/1
τ rejection X	11.4	5.1	0.5	0.45	0.6	0.08	1.7/1

Table 7.1: Cross section $\sigma \times BR(H \rightarrow WW \rightarrow e^\pm \nu \mu^\mp \nu)$ (in fb) through the cuts I-X (described in this subsection). The rates are given at parton level, for the signal process $qq \rightarrow qqH \rightarrow qqW^{(*)}W^{(*)} \rightarrow l^+ \nu l^- \bar{\nu}$ ($M_H=160$ GeV) and for the corresponding backgrounds (values taken from reference [25]).

the main backgrounds (values taken from reference [25]).

As the Feynmann-diagram 7.2 suggests, the $qq \rightarrow qqH \rightarrow qqW^{(*)}W^{(*)} \rightarrow l^+ \nu l^- \bar{\nu}$ signal is characterized by two forward jets, two central leptons coming from the W bosons and missing transverse energy. Therefore, the basic acceptance requirements should not only ensure that two jets and two leptons are observed inside the detector¹ but also that the Higgs decay products and the forward jets are well-separated from each other. These requirements lead to the following selection cuts:

$$p_{t_j} \geq 20 \text{ GeV}, \quad |\eta_j| \leq 4.5, \quad \Delta R_{jj} \geq 0.7, \quad (\text{I})$$

$$p_{t_l} \geq 20 \text{ GeV}, \quad |\eta_l| \leq 2.5, \quad \Delta R_{jl} \geq 0.7. \quad (\text{II})$$

The spatial distances ΔR_{jj} between jets and ΔR_{jl} between jets and leptons are shown in figures, respectively, 7.3 and 7.4 (for signal and the background process $t\bar{t}$).

For the background which arises from QCD corrections to the W^+W^- production, the weak boson bremsstrahlung occurs at small angles with respect to the parent quarks. Therefore, for this background, the W bosons have generally higher rapidity than the Higgs signal. This background can thus be strongly reduced by requiring that both leptons lie between the tagging jets and are well separated

¹In the reference study [25], the higher bound in the jet pseudorapidity was set to $|\eta|=5$. We follow the kinematical acceptance criteria of the CMS detector and we set this value to $|\eta|=4.5$.

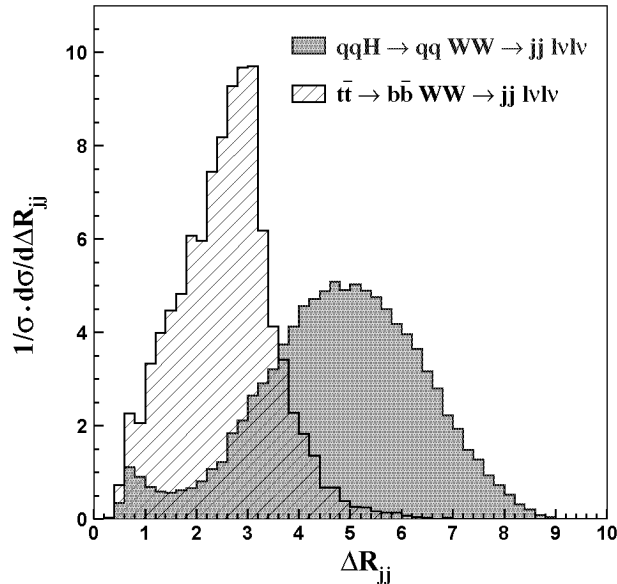


Figure 7.3: PYTHIA simulation of the spatial distance ΔR_{jj} between found jets for the signal process $qqH \rightarrow qqWW$ and the $t\bar{t}$ background ($M_H=170$ GeV).

from them. Furthermore, both jets have to be found in opposite hemispheres:

$$\eta_{j_{min}} + 0.7 < \eta_{1,2} < \eta_{j_{max}} - 0.7, \quad (\text{III})$$

$$\eta_{j_1} \cdot \eta_{j_2} < 0. \quad (\text{IV})$$

The two forward tagging jets should also be well separated in pseudorapidity:

$$\Delta\eta_{tags} = |\eta_{j_1} - \eta_{j_2}| \geq 4.4, \quad (\text{V})$$

leaving a gap of at least 3 units of pseudorapidity in which the leptons can be observed. The expected separation in pseudorapidity for the tagging jets is shown in figures 7.5 and 7.6 for, respectively, signal and $t\bar{t}$ background. The histograms confirm the differences between signal and background.

After these first cuts, the $t\bar{t}$ cross section remains very large (see table 7.1). The largest contribution ($\sim 70\%$) comes from $t\bar{t}jj$, where the additional jets are identified as the forward tagging jets, while the $t\bar{t}j$ constitutes roughly 30% of this background and the $t\bar{t}$ only about 1%. At this level the $t\bar{t}jj$ dominates, because the $t\bar{t}$ pair is not required to have as large an invariant mass as in the other two cases, where one or both b -jets from the top decay are required to be the tagging jets. However, in case that one of the tagging jet is generated by an additional parton in the event, the b -jet which is not “tagged” will most frequently lie between the tagging jets. Vetoing any additional identifiable jet

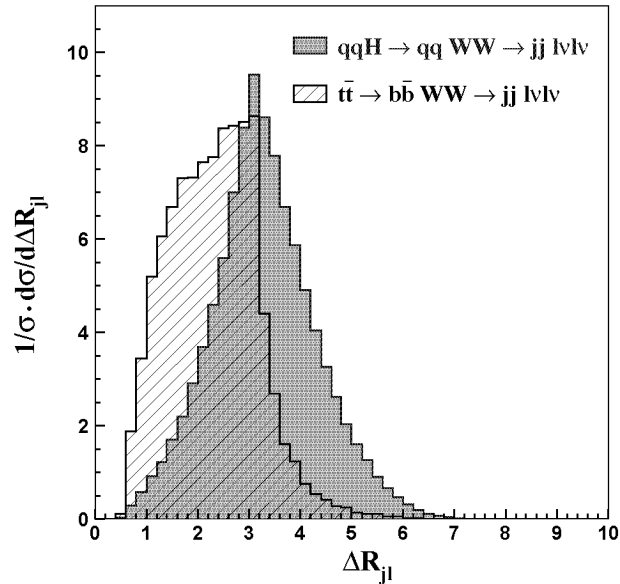


Figure 7.4: PYTHIA simulation of the spatial distance ΔR_{jl} between leptons and jets for signal and $t\bar{t}$ background ($M_H=170$ GeV).

activity in the central region provides thus a powerful suppression tool against the top background. Events with additional jet(s) consistent with:

$$p_{tb} > 20 \text{ GeV}, \quad \eta_{j_{min}} < \eta_b < \eta_{j_{max}} \quad (\text{VI})$$

are therefore rejected. In addition, this central jet veto exploits the different gluon radiation patterns of the Weak Boson Fusion process and of t -channel gluon exchange processes. For the signal, the color coherence between initial- and final-state gluon bremsstrahlung leads to suppressed hadron production in the central region and additional (soft) partons are mostly emitted in the far forward and backward directions. This is in contrast to most background processes, where the color exchange in the t -channel leads instead to enhanced hadronic activity between the tagging jets.

Due to the momentum distribution of gluons within the incoming protons, which is concentrated at small Bjorken x (see for instance figure 4.2), QCD processes typically occur at smaller invariant masses than EW processes. A large dijet invariant mass is therefore required:

$$m_{jj} > 650 \text{ GeV}. \quad (\text{VII})$$

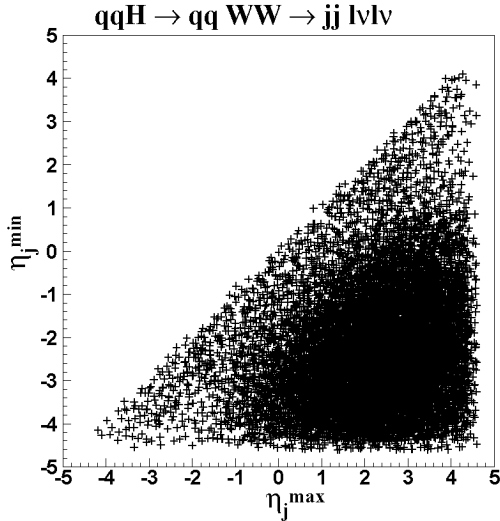


Figure 7.5: PYTHIA: maximal vs minimal jet rapidities in $qqH \rightarrow qqWW \rightarrow jj\ell\nu\ell\nu$ events ($M_H=170$ GeV).

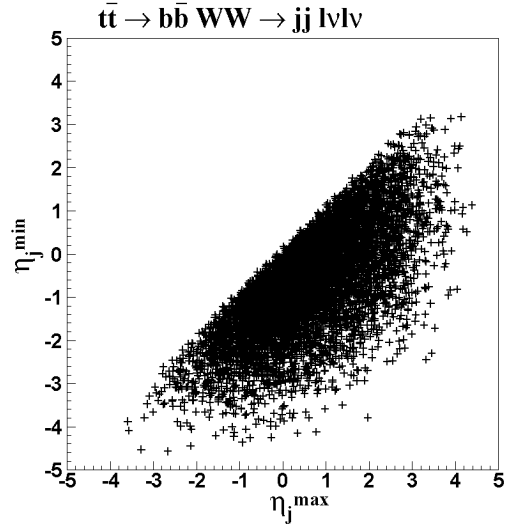


Figure 7.6: PYTHIA: same as figure 7.5, but for the $t\bar{t} \rightarrow qqWW \rightarrow jj\ell\nu\ell\nu$ background.

The first set of cuts I-VII (except the central jet veto VI) could be labeled “forward jet tagging” and constitutes a known effective technique to separate weak boson scattering from various backgrounds. Together with the jet veto, these cuts become a formidable tool to suppress QCD background to weak boson fusion processes.

Other significant differences between signal and various non-resonant scenarios can be found in the angular distributions of the decayed leptons [47]. For the considered mass range $140 \text{ GeV} \leq M_H \leq 200 \text{ GeV}$, in the Higgs rest frame, the two (virtual) W bosons are produced close to the threshold with very small transverse momenta. Moreover, in the case of the Higgs signal, the W s have opposite spins. The V-A coupling, which describe the charged current weak interactions, produce then preferentially two charged leptons propagating in the same direction and thus separated

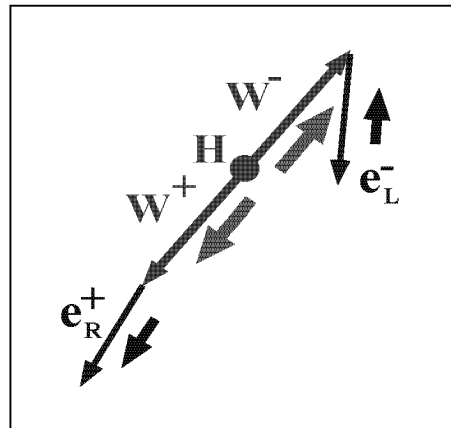


Figure 7.7: Schematic view of the $H \rightarrow WW \rightarrow \ell\nu\ell\nu$ decay in the Higgs rest frame: two W s with opposite spin are produced. Because of spin conservation, the two leptons propagate preferentially in the same direction.

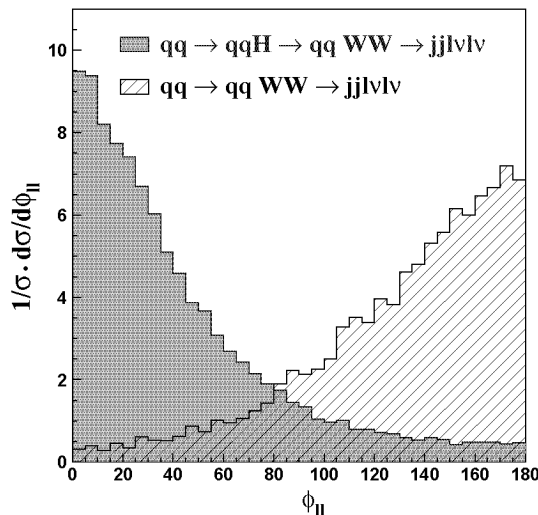


Figure 7.8: PYTHIA simulation of the angle ϕ between the two lepton directions in the transverse plane (in the laboratory frame). The results are shown after the cuts I-VII for resonant and non-resonant weak boson scattering processes $qq \rightarrow qqWW$ ($M_H = 170$ GeV).

by a small angle (see figure 7.7). This favors also small dilepton invariant masses, while the centrally produced Higgs leads to small lepton transverse momenta. These features are exploited in the cuts:

$$\begin{aligned} \phi_U < 105^\circ, \quad \cos \theta_U > 0.2, \quad \Delta R_U < 2.2, & \quad \text{(VIII)} \\ m_U < 110 \text{ GeV}, \quad p_{t_{e,\mu}} < 120 \text{ GeV}, & \quad \text{(IX)} \end{aligned}$$

where ϕ_U and θ_U are, respectively, the azimuthal and polar angles between the two charged leptons (with $z \equiv$ beam axis). These cuts are efficient only if the Higgs boson is produced approximately at rest (*i.e.* if laboratory frame \approx center-of-mass frame) and if the W bosons get a small momenta in the Higgs rest frame. Both conditions are fulfilled for small Higgs masses (typically $M_H \lesssim 180$ GeV). These cuts (labeled “V-A”) are very effective against the top background, since the large t -quark mass allows for very high- p_t leptons, and also against the EW $WWjj$ backgrounds, where the W spins (and therefore the leptons directions) are differently correlated (see table 7.1). This is especially interesting for the non-resonant weak boson scattering processes $qq \rightarrow qqVV$, whose tagging jets have the same kinematical properties as the signal and therefore could constitute a very dangerous source of background. Due to the spin and direction correlations for the signal, cuts VIII-IX prove to be very effective, as shown in figures 7.8, 7.9 and 7.10. Large differences in the ϕ_U -, $\cos \theta_U$ - and ΔR_U -spectrum between resonant and non-resonant weak boson fusion processes can be clearly observed.

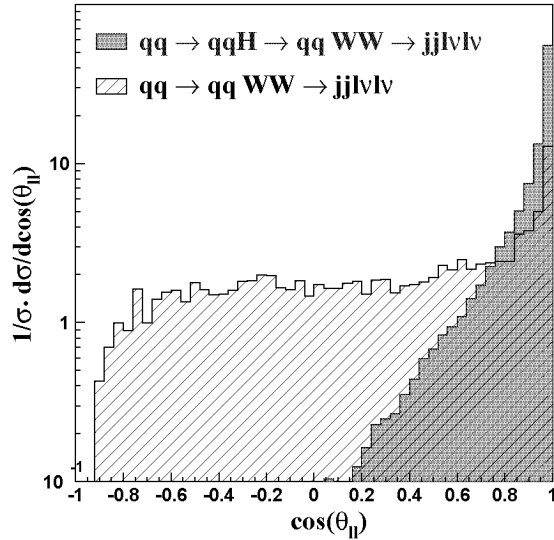


Figure 7.9: Cosine of the polar angle θ between the two lepton directions (PYTHIA). The results are shown after the cuts I-VII for the Higgs signal and for the non-resonant weak boson scattering process $qq \rightarrow qqWW$ ($M_H = 170$ GeV).

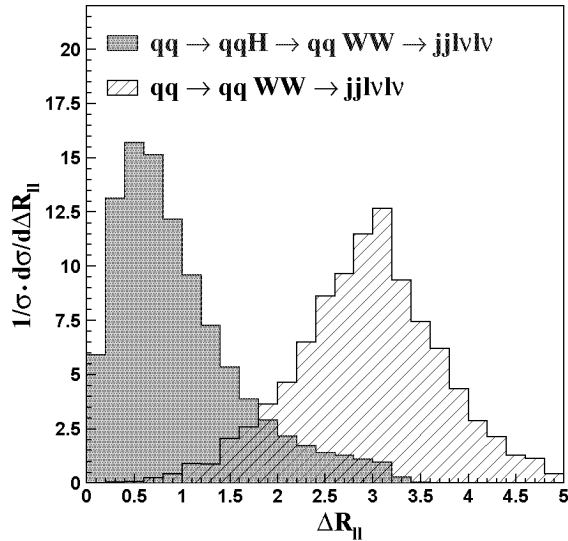


Figure 7.10: Spatial distance $\Delta R_{ll} = \sqrt{\Delta \eta_{ll}^2 + \Delta \phi_{ll}^2}$ between the two leptons for the Higgs signal and the weak boson scattering process $qq \rightarrow qqWW$ (PYTHIA). The cuts I-VII have been applied ($M_H = 170$ GeV).

After applying these cuts, the background from τ -pairs still remains large. For the $Z/\gamma \rightarrow \tau\tau$ processes, the Z^0 or γ is produced with a quite large p_t (~ 100 GeV), giving a large boost to the emitted τ 's. As $\mathbf{p}_\tau \gg m_\tau$, the decay products are then nearly collinear to the original τ (as schematically shown in figure 7.11). Therefore, with the information about the charged lepton momenta and the missing transverse energy, one can approximately reconstruct the momenta of the original τ 's. In detail, for a $\tau^+\tau^- \rightarrow l^+l^- \cancel{p}_t$ event, if one assumes a collinearity between each τ and its decay products ($\mathbf{p}_{t\tau} \parallel \mathbf{p}_{t\nu_\tau} \parallel \mathbf{p}_{tl} \parallel \mathbf{p}_{t\nu_l}$), the fractions $x_{\tau_i} \equiv \mathbf{p}_{tl_i}/\mathbf{p}_{t\tau_i}$ of τ -energy carried by each l^\pm can be determined from the transverse momentum balance:

$$\mathbf{p}_{t\tau_1} + \mathbf{p}_{t\tau_2} = \frac{\mathbf{p}_{tl_1}}{x_{\tau_1}} + \frac{\mathbf{p}_{tl_2}}{x_{\tau_2}} = \mathbf{p}_{tl_1} + \mathbf{p}_{tl_2} + \cancel{p}_t \quad (7.1)$$

$$\Rightarrow \mathbf{p}_{tl_1} \left(\frac{1}{x_{\tau_1}} - 1 \right) + \mathbf{p}_{tl_2} \left(\frac{1}{x_{\tau_2}} - 1 \right) = \cancel{p}_t \quad (7.2)$$

By solving the 2 equations for the x, y transverse directions, one finds for the two unknowns $x_{\tau_{1,2}}$:

$$\begin{cases} x_{\tau_1} = \frac{p_{xl_1} \cdot p_{yl_2} - p_{yl_1} \cdot p_{xl_2}}{\cancel{p}_x \cdot p_{yl_2} - \cancel{p}_y \cdot p_{xl_2} + p_{xl_1} \cdot p_{yl_2} - p_{yl_1} \cdot p_{xl_2}} \\ x_{\tau_2} = \frac{p_{yl_1} \cdot p_{xl_2} - p_{xl_1} \cdot p_{yl_2}}{\cancel{p}_x \cdot p_{yl_1} - \cancel{p}_y \cdot p_{xl_1} + p_{yl_1} \cdot p_{xl_2} - p_{xl_1} \cdot p_{yl_2}} \end{cases} \quad (7.3)$$

Thus, with the help of the found variables $x_{\tau_{1,2}}$, the τ -pair invariant mass $m_{\tau\tau}$ can be reconstructed:

$$m_{\tau\tau} \cong \frac{m_{ll}}{\sqrt{x_{\tau_1} x_{\tau_2}}}, \quad (7.4)$$

where the well-justified approximation $E_{\tau,l} \cong |\mathbf{p}_{\tau,l}|$ has been used.

In case of a real $Z^0/\gamma \rightarrow \tau\tau \rightarrow l^+l^- \cancel{p}_t$ decay, normally the reconstruction of the τ momenta should yield to values between 0 and 1 for both x_{τ_1} and x_{τ_2} (see figure 7.12). On the other hand, for the signal process $qqH \rightarrow qqWW \rightarrow jj l\nu l\nu$

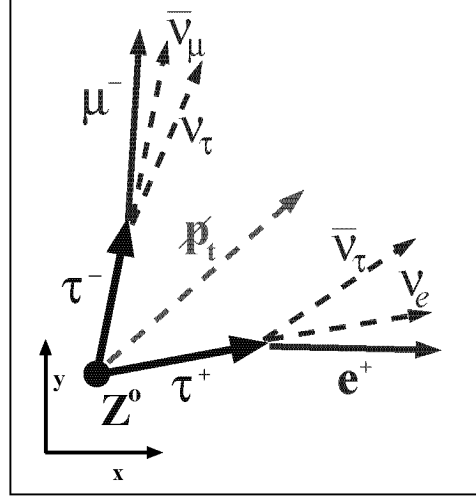


Figure 7.11: Schematic view of a $Z^0 \rightarrow \tau^+\tau^- \rightarrow l^+l^- \cancel{p}_t$ process, where the decay products are nearly collinear to the original τ .

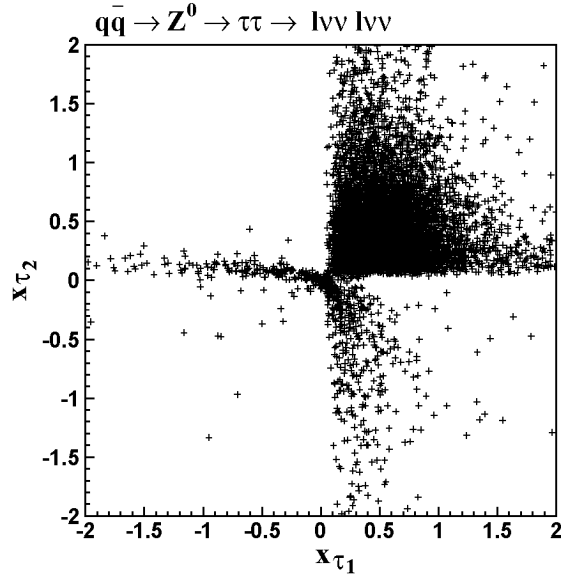


Figure 7.12: x_{τ_1} vs. x_{τ_2} distribution for the $q\bar{q} \rightarrow Z^0 \rightarrow \tau^+\tau^- \rightarrow l^+l^- \cancel{p}_t$ process (PYTHIA). The fractions x_{τ_i} of τ energy carried by each charged lepton are reconstructed assuming that the two l^\pm are coming from a τ decay (*i.e.* using the equations 7.3).

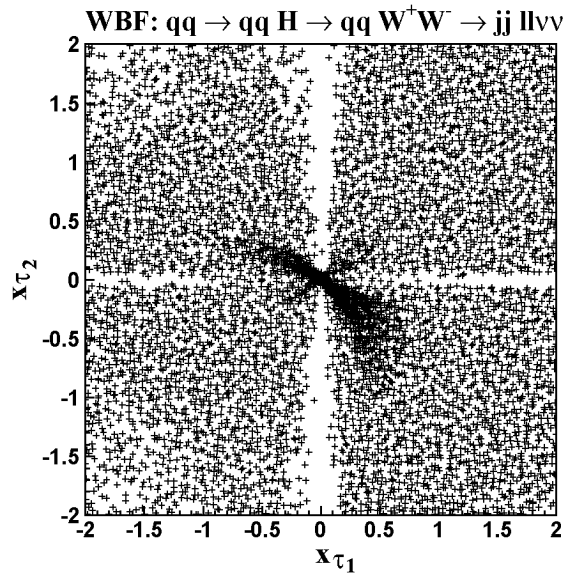


Figure 7.13: Same as figure 7.12, but for the signal process $H \rightarrow WW \rightarrow l^+\nu l^-\bar{\nu}$. In this case, the wrong assumption of τ decays leads to negative $x_{\tau_{1,2}}$ -values, which are without any physical meaning.

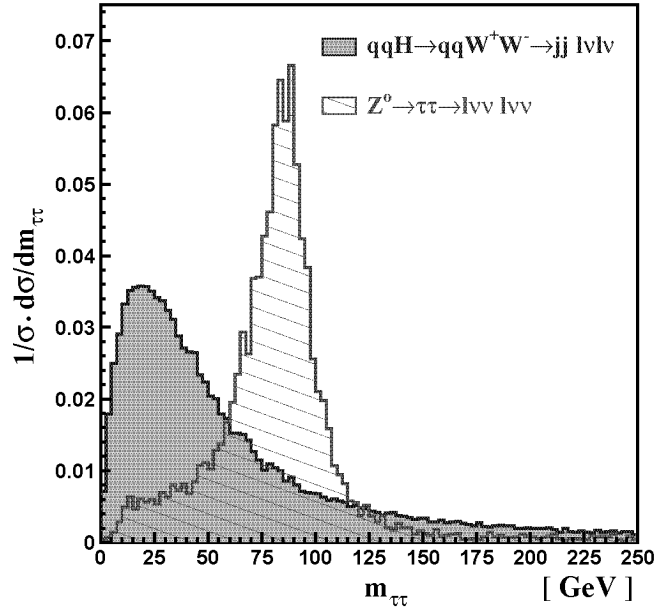


Figure 7.14: PYTHIA-reconstruction of $m_{\tau\tau}$ (using equation 7.4): in contrast to the signal $H \rightarrow WW \rightarrow l^+\nu l^-\bar{\nu}$ ($M_H = 160$ GeV), the background process $Z^0 \rightarrow \tau^+\tau^- \rightarrow l^+\nu l^-\bar{\nu}$ peaks at m_Z .

and for other backgrounds, the collinear approximation used in equation 7.1 is almost never valid, because the W s receive modest boosts and the W mass is much larger than the τ one, giving rise to larger opening angles between the produced l 's and ν 's. Therefore, for about 95% of the signal events, an attempt to reconstruct the τ momenta results in $x_{\tau_{1,2}}$ values without physical meaning (see figure 7.13). Even when both $x_{\tau_{1,2}}$ are positive, the reconstructed $m_{\tau\tau}$ does not peak at m_Z , as shown in figure 7.14. Thus, by vetoing events consistent with $Z^0 \rightarrow \tau\tau$ decays :

$$x_{\tau_1} > 0, \quad x_{\tau_2} > 0, \quad |m_{\tau\tau} - m_Z| < 25 \text{ GeV}, \quad (\text{X})$$

the $\tau\tau$ background is strongly reduced.

Up to now, all the cuts are perfectly consistent with a signal selection with two different lepton flavors, that is $H \rightarrow WW \rightarrow e^\pm\mu^\mp \not{p}_t$. In reference [25] (*i.e.* in the analysis which focused on the Higgs mass domain $150 \text{ GeV} \leq M_H \leq 200 \text{ GeV}$) the “same flavor” cases were not taken into account. The analysis is extended by including these events as well (doubling the statistics). For this purpose, the same minimum m_{ll} and missing p_t cuts as for $110 \text{ GeV} \leq M_H \leq 140 \text{ GeV}$ is used. If one wants to include also the $e^+e^- \not{p}_t$ and $\mu^+\mu^- \not{p}_t$ signatures, a minimum m_{ll}

7.2 Event selection

cut is needed in order to get rid of the large $\gamma^* \rightarrow ll$ background:

$$m_{ll} > 10 \text{ GeV}. \quad (\text{XI})$$

To reject $lljj$ events, where the missing transverse momentum is generated by detector effects, an additional cut on \cancel{p}_t is used:

$$\cancel{p}_t > 30 \text{ GeV}. \quad (\text{XII})$$

Finally, the reconstructed m_{ll} should not be compatible with m_Z :

$$|m_{ll} - m_Z| > 10 \text{ GeV}. \quad (\text{XIII})$$

To summarize, for Higgs masses between 140 and 200 GeV and for events with two different flavor leptons, the same cuts I-X as in the original study are used [25]², while to include also the same flavor cases three additional cuts XI-XIII are used.

7.2.2 110 GeV $\leq M_H \leq$ 140 GeV

For this lower Higgs mass domain, the selection criteria given in reference [26] are followed. The first part of the cuts are very similar to the previous case, with some minor modifications required by the different event kinematics:

$$p_{tj} \geq 20 \text{ GeV}, \quad |\eta_j| \leq 4.5, \quad \Delta R_{jj} \geq 0.6, \quad (\text{i})$$

$$p_{tl_1} \geq 20 \text{ GeV}, \quad p_{tl_2} \geq 10 \text{ GeV}, \quad |\eta_l| \leq 2.5, \quad \Delta R_{jl} \geq 1.7, \quad (\text{ii})$$

$$\eta_{j_{min}} + 0.6 < \eta_{l_{1,2}} < \eta_{j_{max}} - 0.6, \quad (\text{iii})$$

$$\eta_{j_1} \cdot \eta_{j_2} < 0, \quad (\text{iv})$$

$$\Delta\eta_{tags} = |\eta_{j_1} - \eta_{j_2}| \geq 4.2, \quad (\text{v})$$

$$p_{tb} > 20 \text{ GeV}, \quad \eta_{j_{min}} < \eta_b < \eta_{j_{max}}, \quad (\text{vi})$$

$$m_{jj} > 600 \text{ GeV}, \quad (\text{vii})$$

$$m_{ll} < 60 \text{ GeV}, \quad \phi_{ll} < 140^\circ, \quad (\text{viii})$$

with a veto for events consistent with:

$$x_{\tau_1} > 0, \quad x_{\tau_2} > 0, \quad m_{\tau\tau} > m_Z - 25 \text{ GeV}. \quad (\text{ix})$$

Clearly, in order to identify a possible Higgs boson, the reconstruction of the mass distribution would be desirable. However, the leptonic decay of the two W bosons to electrons or muons implies the presence of two neutrinos (*i.e.*, even with some restrictive assumptions, it would not be possible to reconstruct the Higgs mass without a two-fold ambiguity).

²Except for the smaller $|\eta|$ -acceptance for jets (cut I).

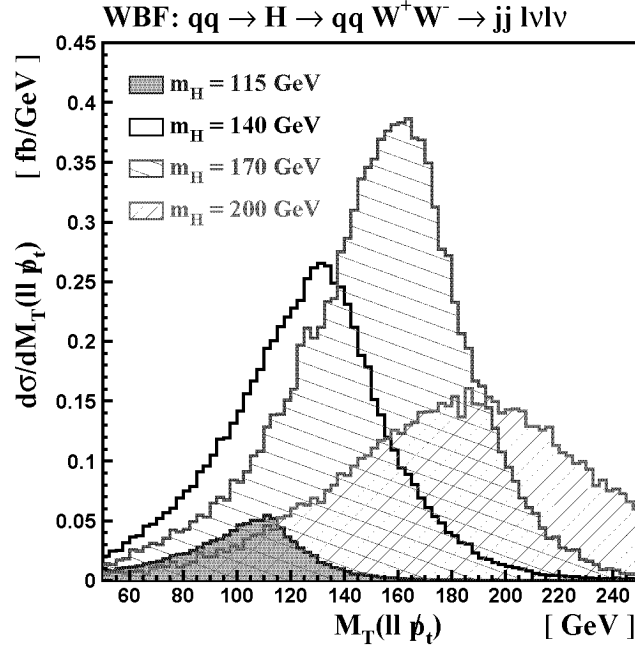


Figure 7.15: PYTHIA simulation of the transverse mass distribution of the dilepton- \cancel{p}_t system, for a Higgs boson of mass $M_H = 115, 140, 170$ and 200 GeV (events with 2 jets and 2 isolated leptons have been selected).

Anyhow, we can instead reconstruct the transverse mass M_T of the $ll \cancel{p}_t$ system (using the threshold relationship $m_{\nu\nu} \approx m_{ll}$, which is justified by the very small W momenta in the Higgs frame):

$$M_T(ll \cancel{p}_t) = \sqrt{(\cancel{E}_t + E_{t,ll})^2 - (\mathbf{p}_{t,ll} + \cancel{\mathbf{p}}_t)^2}, \quad (7.5)$$

where the two transverse energies are defined as follow:

$$\cancel{E}_t = \sqrt{\cancel{\mathbf{p}}_t^2 + m_t^2} \quad \text{and} \quad E_{t,ll} = \sqrt{\mathbf{p}_{t,ll}^2 + m_{ll}^2}. \quad (7.6)$$

The transverse mass distribution for four different Higgs masses is shown in figure 7.15. Above the threshold for the $H \rightarrow WW$ process, the relationship $m_{\nu\nu} = m_{ll}$ becomes only a rough approximation and a broadening of the Higgs mass peak can be clearly observed. Nevertheless the signal is largely concentrated in the region

$$50 \text{ GeV} < M_T(ll \cancel{p}_t) < M_H + 20 \text{ GeV}, \quad (\mathbf{x})$$

and a good improvement of the signal significance is obtained by selecting only events in this mass region.

After all previous cuts, a large contribution to the total background is given by

7.2 Event selection

cuts	Hjj	$t\bar{t} + jets$	$b\bar{b}jj$	$WWjj$		$\tau\tau jj$	
				QCD	EW	QCD	EW
i-x	1.04	1.9	20.8	0.26	0.19	0.67	0.118
xi	0.93	1.65	0.48	0.23	0.17	0.11	0.026
xii	0.92	1.64	0.013	0.23	0.17	0.10	0.024

Table 7.2: Parton level cross sections $\sigma \times BR(H \rightarrow e^\pm \mu^\mp \cancel{p}_t)$ for a signal with $M_H = 115$ GeV and for the corresponding backgrounds. The values are given in fb and are taken from reference [26].

$bbjj$ events (see table 7.2). It arises from semi-leptonic $b \rightarrow cl\nu$ decays in which the charm quark and other possible hadronization products carry only little energy. As proposed in reference [26], an efficient way to eliminate this background (and also part of the surviving $\tau\tau jj$ events), is to exploit the correlations between lepton azimuthal angles and the reconstructed Higgs transverse momentum $p_{tH} = |\mathbf{p}_{t_{l_1}} + \mathbf{p}_{t_{l_2}} + \cancel{\mathbf{p}}_t|$. For the background processes $b\bar{b}$ and $\tau^+\tau^-$, the neutrinos are emitted almost parallel to the observed charged lepton. This means that the $\cancel{\mathbf{p}}_t$ -vector will probably lie between the p_t -vectors of the two charged leptons and, hence, $\cancel{\mathbf{p}}_t$ and $\mathbf{p}_{t,ll} = \mathbf{p}_{t_{l_1}} + \mathbf{p}_{t_{l_2}}$ will be approximately parallel to each other. In the Higgs rest frame, due to the anti-correlation of the W spins and to the V-A structure of the W -fermion coupling, the two leptons are preferentially emitted in the same direction with the neutrinos recoiling against the ll -system.

These differences can be exploited by using the azimuthal angle $\Delta\phi(ll, \cancel{p}_t)$ between $\mathbf{p}_{t,ll}$ and $\cancel{\mathbf{p}}_t$. The $bbjj$ and $\tau\tau jj$ backgrounds favor small angles, while in the signal the separation between dilepton and \cancel{p}_t is large, except when these decay products are boosted by a large Higgs transverse momentum.

These correlations can be clearly seen in the scatter plots of figures 7.16-7.18³. By imposing the so-called 'contour-cut' (*i.e.* by eliminating the events below the line shown in the figures), one can effectively get rid of these backgrounds without seriously affecting the signal (see table 7.2):

$$\Delta\phi(ll, \cancel{p}_t) + \frac{15^\circ}{\text{GeV}} \cdot p_{tH} > 180^\circ, \quad 12 \cdot \Delta\phi(ll, \cancel{p}_t) + \frac{1^\circ}{\text{GeV}} \cdot p_{tH} > 360^\circ, \quad (\mathbf{xi})$$

with $\Delta\phi(ll, \cancel{p}_t)$ in degrees and p_{tH} measured in GeV.

This 'contour-cut' should also eliminate part of the background which arises from mismeasurements in the detector: in that case the missing p_t is small and uncorrelated to the lepton direction.

After the last cut, the remaining $bbjj$ events are characterized by the softness

³One has to proceed cautiously with the PYTHIA simulation shown in figure 7.16, since the results are given without the previous cuts. However, these cuts do not seem to produce relevant changes (see [26]).

of their missing p_t . Moreover, their “ p_{tH} ” (that is the transverse momentum of the jet sum) is also small in comparison with real qqH events. Therefore, a large fraction of this background can be eliminated by imposing:

$$\cancel{p}_t > 20 \text{ GeV}, \quad \text{provided that } p_{tH} < 50 \text{ GeV}. \quad (\text{xii})$$

It has to be pointed out, that the cuts **x-xii** are only used for the lower Higgs mass domain $110 \text{ GeV} \leq M_H \leq 140 \text{ GeV}$.

Finally, to include also the “same-flavor” final states $e^+e^- \cancel{p}_t$ and $\mu^+\mu^- \cancel{p}_t$, the following cuts are used⁴:

$$m_{ll} > 10 \text{ GeV}, \quad (\text{xiii})$$

$$\cancel{p}_t > 30 \text{ GeV}. \quad (\text{xiv})$$

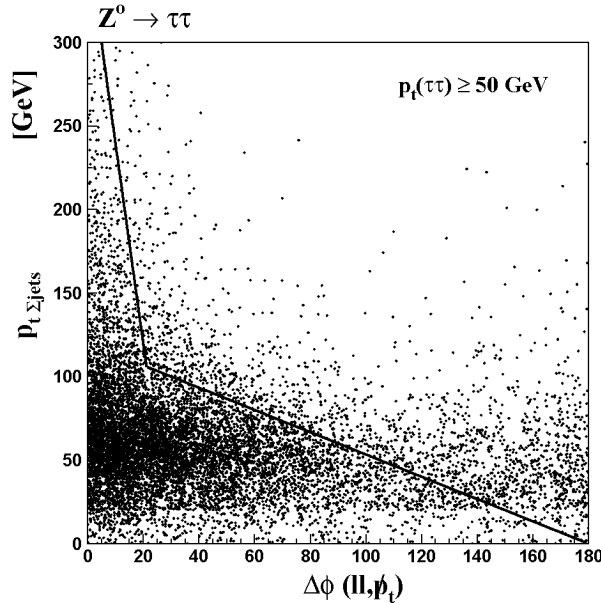


Figure 7.16: $Z^0(q/g) \rightarrow \tau\tau(jets)$ (PYTHIA): p_{tjets} as a function of the azimuthal angle between dilepton and missing transverse momenta ($\Delta\phi(ll, \cancel{p}_t)$). The boosted τ 's give rise to a lepton and a neutrino which are almost parallel to each other. Consequently, ll system and \cancel{p}_t are also parallel. For the signal, this is valid only for large p_{tjets} ($\approx p_{tH}$). Therefore, by selecting the events above the line, one can get rid of this background (the same arguments are valid also for the $b\bar{b}$ background). In order to tune the simulation to the relevant events, a minimal cut of 50 GeV on the Z^0 - p_t was imposed (otherwise the only constraints are to have 2 isolated leptons and at least one jet in the event).

⁴The invariant mass cut **viii** ($m_{ll} < 60 \text{ GeV}$) already eliminated the leptonic Z decays.

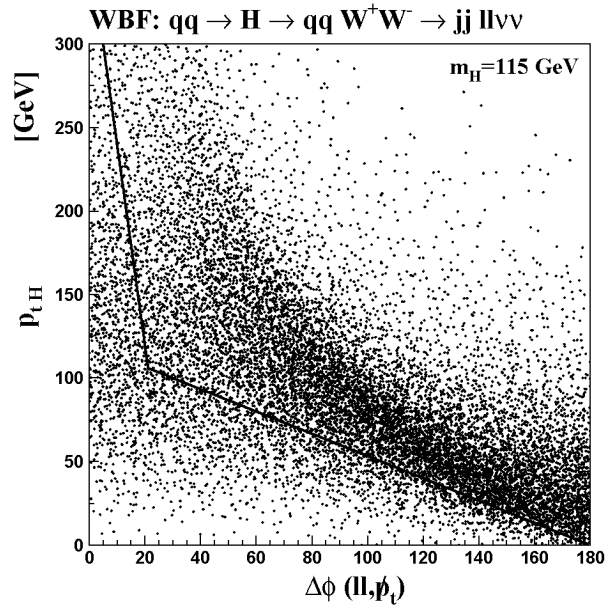


Figure 7.17: $qqH \rightarrow qqWW \rightarrow jjl\nu l\nu$ (PYTHIA): p_{tH} as a function of the azimuthal angle between dilepton momentum and missing transverse momentum ($\Delta\phi(l, \cancel{p}_t)$). Only events above the line are selected by cut **xi**.

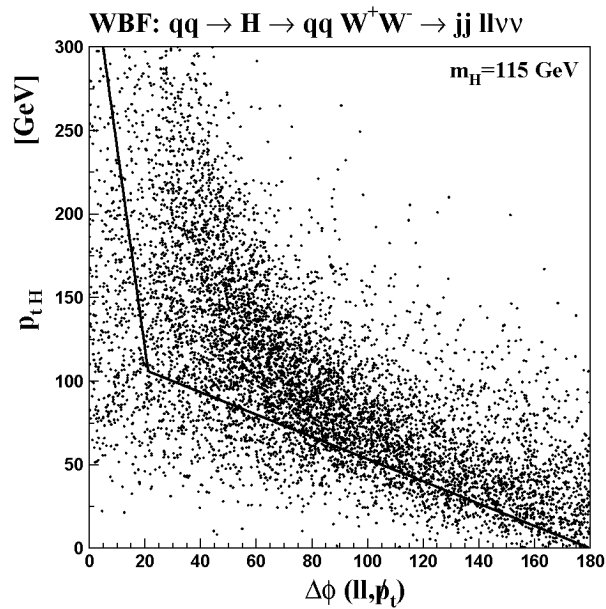


Figure 7.18: Same as figure 7.17, but after the cuts **i-x**. The cut **xi** selects only events above the line.

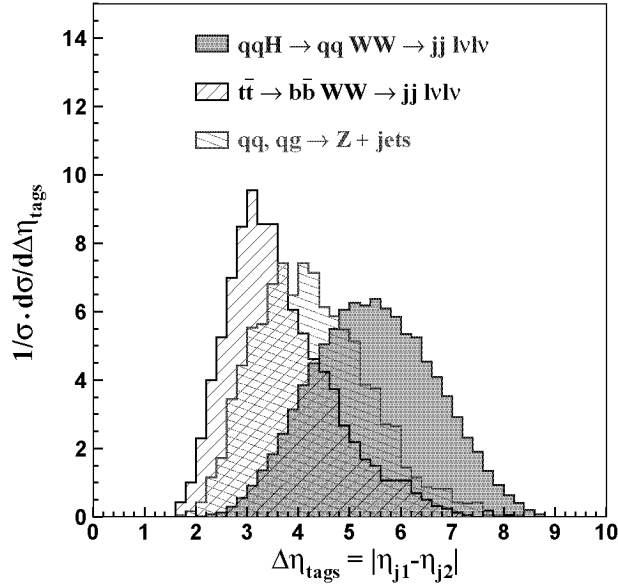


Figure 7.19: PYTHIA simulation of the η -separation between the two tagging jets. The figure shows the distribution for the signal ($M_H = 260$ GeV) and for the background processes $t\bar{t} \rightarrow q\bar{q} WW \rightarrow jj l\nu l\nu$ and $q_i\bar{q}_i, q_i g \rightarrow Z + jets$.

7.2.3 $200 \text{ GeV} \leq M_H \leq 300 \text{ GeV}$

Similar cuts are used to extend the analysis into the 200-300 GeV mass range:

$$p_{t_j} \geq 20 \text{ GeV}, \quad |\eta_j| \leq 4.5, \quad \Delta R_{jj} \geq 0.7, \quad (\text{i})$$

$$p_{t_l} \geq 20 \text{ GeV}, \quad |\eta_l| \leq 2.5, \quad \Delta R_{jl} \geq 1.5, \quad (\text{ii})$$

$$\eta_{j_{min}} + 0.7 < \eta_{l_{1,2}} < \eta_{j_{max}} - 0.7, \quad (\text{iii})$$

$$\eta_{j_1} \cdot \eta_{j_2} < 0. \quad (\text{iv})$$

As for the previous mass ranges, a wide separation in pseudorapidity between the tagging jets is required:

$$\Delta\eta_{tags} = |\eta_{j_1} - \eta_{j_2}| \geq 4.6, \quad (\text{v})$$

For the signal and two important backgrounds, the expected $\Delta\eta_{tags}$ distribution is shown in figure 7.19.

Figure 7.20 shows the mass distribution of the jj -system, which has to be larger than:

$$m_{jj} > 650 \text{ GeV}. \quad (\text{vi})$$

Also in this case, the background processes $t\bar{t} \rightarrow q\bar{q} WW \rightarrow jj l\nu l\nu$ and $q_i\bar{q}_i, q_i g \rightarrow$

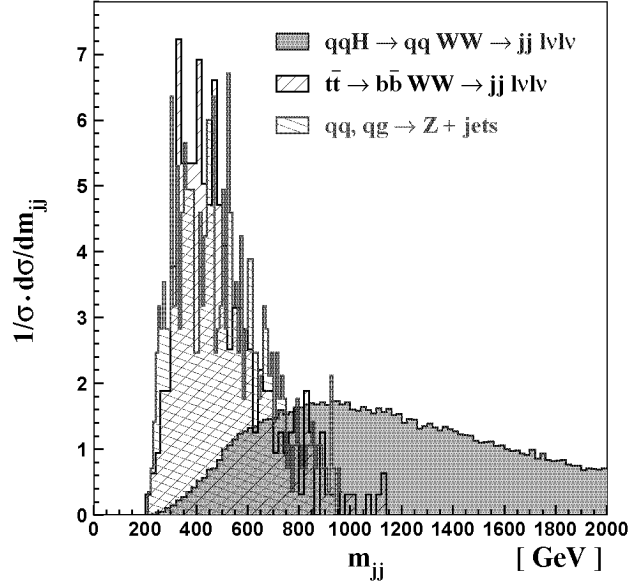


Figure 7.20: PYTHIA simulation of the mass of the tagging jet system (after cut \mathbf{v}). The two background processes $t\bar{t} \rightarrow q\bar{q} WW \rightarrow jj l\nu l\nu$ and $q_i\bar{q}_i, q_i g \rightarrow Z + jets$ have generally smaller m_{jj} -values in comparison with the weak boson fusion signal ($M_H = 260$ GeV).

$Z + jets$ give rise to smaller m_{jj} -values in comparison with the $qqH \rightarrow qq WW \rightarrow jj l\nu l\nu$ signal.

$t\bar{t} + jets$ events are further suppressed by vetoing additional b -jets consistent with:

$$p_{tb} > 20 \text{ GeV}, \quad \eta_{j_{min}} < \eta_b < \eta_{j_{max}}. \quad (\text{vii})$$

For the background processes $q\bar{q} \rightarrow WW \rightarrow ll \not{p}_t$ and $t\bar{t} \rightarrow jj l\nu l\nu$, the spatial distance between leptons is larger than for the signal (see figure 7.21). This feature is exploited using:

$$\cos \theta_{ll} > 0.2, \quad \Delta R_{ll} < 3.2. \quad (\text{viii})$$

In addition, the m_{ll} and $p_{t_{e,\mu}}$ regions are selected according to:

$$50 \text{ GeV} < m_{ll} < 200 \text{ GeV}, \quad p_{t_{e,\mu}} < 150 \text{ GeV}. \quad (\text{ix})$$

To suppress $Z \rightarrow \tau\tau$ decays, we veto any event consistent with:

$$x_{\tau_1} > 0, \quad x_{\tau_2} > 0, \quad |m_{\tau\tau} - m_Z| < 25 \text{ GeV}. \quad (\text{x})$$

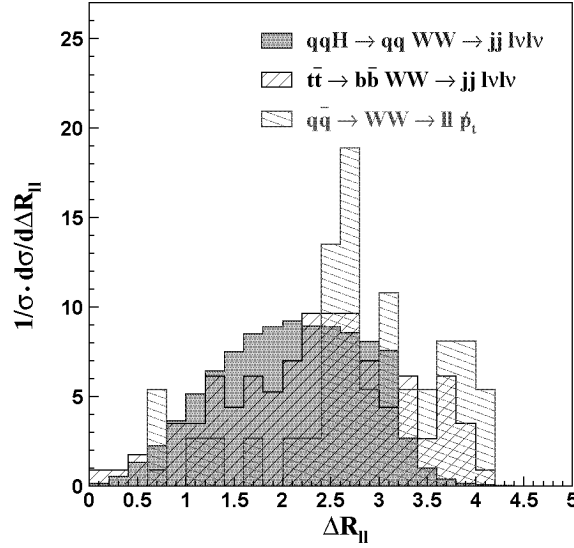


Figure 7.21: Spatial distance $\Delta R_{\ell\ell} = \sqrt{\Delta\eta_{\ell\ell}^2 + \Delta\phi_{\ell\ell}^2}$ between the two leptons produced in the Higgs signal and in the background processes $q\bar{q} \rightarrow WW \rightarrow \ell\ell \cancel{p}_t$ and $t\bar{t} \rightarrow jj \ell\nu\ell\nu$ (PYTHIA). The cuts **i-vii** have been applied ($M_H = 260$ GeV).

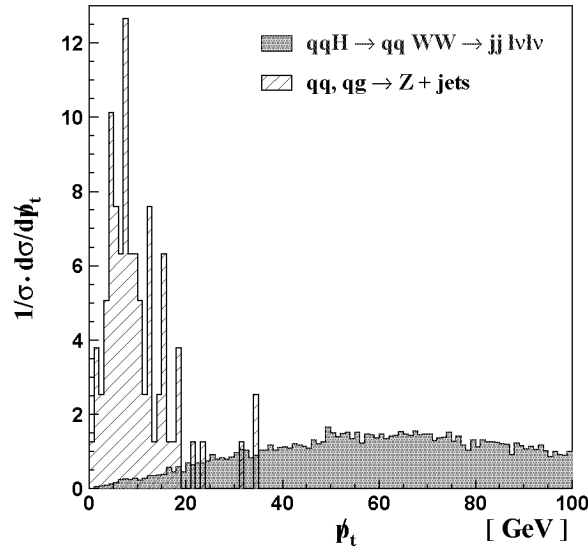


Figure 7.22: For e^+e^- and $\mu^+\mu^-$ events: expected missing transverse momentum for the Higgs signal ($M_H = 260$ GeV) and the $q_i\bar{q}_i, q_i g \rightarrow Z + jets$ background process (after cuts **i-x**, simulated with PYTHIA).

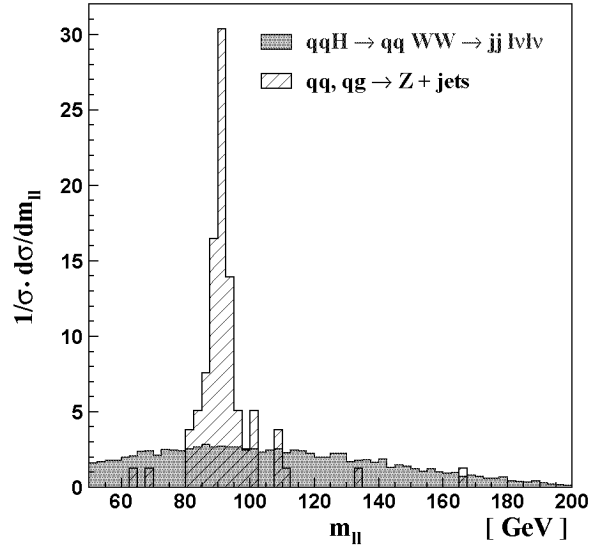


Figure 7.23: PYTHIA simulation of the dilepton mass distribution after cuts **i-x**. Only $e^+e^- \cancel{p}_t$ and $\mu^+\mu^- \cancel{p}_t$ final states are shown ($M_H = 260$ GeV).

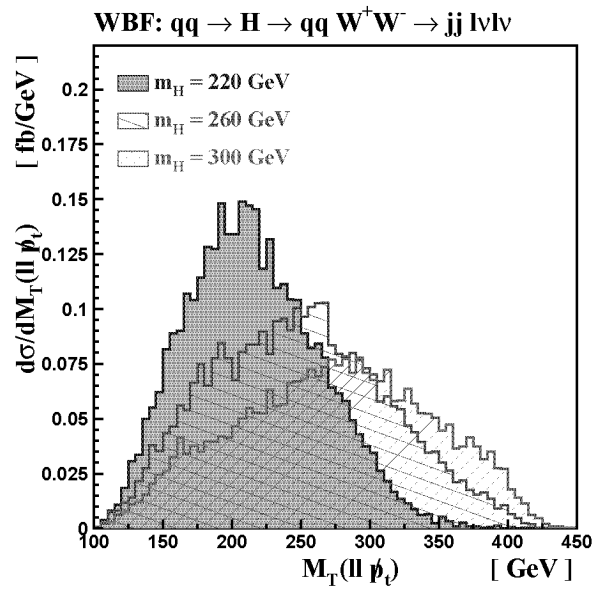


Figure 7.24: PYTHIA simulation of the transverse mass distribution of the dilepton- \cancel{p}_t system, for a Higgs boson of mass $M_H = 220, 260$ and 300 GeV (after cuts **i-xii**).

The “same-flavor” final states $e^+e^- \cancel{p}_t$ and $\mu^+\mu^- \cancel{p}_t$ are selected with the additional cuts:

$$\cancel{p}_t > 30 \text{ GeV}, \quad (\text{x i})$$

$$|m_{ll} - m_Z| > 10 \text{ GeV}. \quad (\text{x ii})$$

These latter selection cuts allow a very effective suppression of $q_i\bar{q}_i, q_i g \rightarrow Z + jets$ events, as shown in figures 7.22 and 7.23.

Finally, the transverse mass $M_T(ll \cancel{p}_t)$ should be approximately consistent with the Higgs mass (see figure 7.24):

$$M_H - \left(\frac{M_H}{4} + 25\right) \text{ GeV} < M_T(ll \cancel{p}_t) < M_H + 100 \text{ GeV}. \quad (\text{x iii})$$

7.3 Results

Our PYTHIA analysis follows a series of kinematic cuts proposed in references [25] and [26]. Similarly to these studies, the signatures are analyzed at parton-level. In addition, also the full showering process, initial state radiation and the detector acceptance are included, allowing the analysis at several levels. Following this procedure, we can obtain informations regarding the efficiency for the identification and reconstruction of leptons and jets, as reported in section 3.3.2. Efficiency factors have also been taken into account in both studies [25] and [26], but were found to be too optimistic by a complete “hadronization” study and with a more realistic detector set-up.

As was pointed out at the beginning of this chapter, not all the main backgrounds relevant to the $qq \rightarrow qqH$ signal are included in the PYTHIA program, precluding the possibility to perform a complete visibility study with this generator. In order to overcome this problem, for the signal the idea is to use the cross section obtained from the PYTHIA analysis, while for the background we take the results given in the two above listed references and we use the efficiency factors we get in our PYTHIA analysis of the signal. In other words, we eliminate the $\epsilon_{\text{tag}, j}, \epsilon_{\text{iso}, l}$ given in references [25, 26] and we multiply instead the background cross sections by our $\epsilon_{\text{tag}, j}, \epsilon_{\text{iso}, l}$ values.

All the PYTHIA cross sections given in this chapter have been corrected with the “K-factors” given in reference [105], which include the NLO corrections for the SM Higgs production $\sigma(q\bar{q} \rightarrow q\bar{q}H)$ and for the branching ratios $BR(H \rightarrow WW^{(*)})$.

7.3.1 110 GeV $\leq M_H \leq$ 200 GeV: PYTHIA vs. original studies [25, 26]

For a 160 GeV Higgs, the behavior of the cross sections through the selection cuts is listed on table 7.3. These values cover only the different lepton flavor case

7.3 Results

selection cuts	$\sigma \times BR$	2 isolated leptons & ≥ 2 jets	forward jet tagging I-VI	m_{jj} & "V-A" VII-IX	real τ rejection X
R&Z [25]	~ 125	-	17.1	11.8	11.4
PYTHIA parton level	125.6	37.6	14.8	10.5	10.0
PYTHIA full showering	125.6	14.3	6.6	4.0	3.9

Table 7.3: Signal rates (in fb) for different selection cuts ($M_H = 160$ GeV). In the first line we report the results given in reference [25] for the process $qqH \rightarrow qq WW \rightarrow jj e^\pm \mu^\mp \not{p}_t$ (labeled as "R&Z"). For our PYTHIA analysis we show the results obtained either at parton level or after the hadronization.

$$H \rightarrow WW \rightarrow e^\pm \mu^\mp \not{p}_t.$$

The first important result is the good agreement (within a 15% difference) between the parton level⁵ cross sections given by PYTHIA⁶ and the original simulations of reference [25]. On the other hand, if one reconstructs jets and leptons starting from the "detectable" particles, the cross sections drop to about 1/3 of the parton level ones. In the previous analysis [25, 26], these "showering" and "detector" effects were taken into account by multiplying the cross sections by the following factors:

- 1) an efficiency factor $\epsilon_{\text{minijet veto}}$ for the "minijet veto"⁷,
- 2) one factor $\epsilon_{\text{iso. } l}$ for the lepton identification efficiency (used only for $M_H \leq 140$ GeV),
- 3) and a last one ($\epsilon_{\text{tag. } j}$) for the reconstruction efficiency of the tagging jets.

After these factors are applied, the obtained signal cross sections are given in the last column of the table 7.4. This value (7.5 fb) is still about twice as large compared to our "full showering" result (3.9 fb). In other words, we obtain smaller values for the different efficiency factors!

The strategy is then to use these different efficiencies in order to obtain new values also for the background. For the "minijet veto" we rely on the factors given in the references [25, 26]. This choice is justified by the fact that these values

⁵*i.e.* obtained by using directly the kinematics of the partons shown on figure 7.2.

⁶In this case we used the same higher bound in the jet pseudorapidity as in reference [25], *i.e.* $|\eta|_{\text{max}}=5$.

⁷This factor simulated the probability to survive the cut which eliminates events with other central low- p_t jets (arising from soft gluon radiation).

	all selection cuts	minijet veto	tag ID efficiency
$qqH \rightarrow qqWW \rightarrow jj e^\pm \mu^\mp \cancel{p}_t$	11.4	10.1	7.5
$t\bar{t} (+jets)$	5.1	1.48	1.1

Table 7.4: Rates (in fb) given in reference [25] for the signal process ($M_H = 160$ GeV) and the $t\bar{t}$ -background. The values are given before and after the application of the different efficiency factors.

vary strongly among the different processes and depend on the gluon radiation patterns.

In contrast to the “minijet veto”, the remaining efficiency factors should be more stable among the signal and the different backgrounds. The reason is simple: after all selection cuts, the lepton and jet kinematics should be very similar for both signal and backgrounds.

For the remaining efficiency factors, the previous studies gave: $\epsilon_{\text{tag},j} = 0.74$ (for $M_H \geq 140$ GeV⁸) and $\epsilon_{\text{iso},l} \cdot \epsilon_{\text{tag},j} = 0.67$ (for $M_H \leq 140$ GeV). We eliminate them and we multiply instead the background cross sections by our values. These latter are obtained by first multiplying our signal cross section at parton level by the “minijet veto” efficiency. For a 160 GeV Higgs we get:

$$10.0 \text{ fb} \cdot \epsilon_{\text{minijet veto}} = 10.0 \text{ fb} \cdot \frac{10.1 \text{ fb}}{11.4 \text{ fb}} = 8.9 \text{ fb}, \quad (7.7)$$

Then by comparing the obtained value with our “full showering” signal cross section:

$$\epsilon_{\text{iso},l} \cdot \epsilon_{\text{tag},j} = \frac{3.9 \text{ fb}}{8.9 \text{ fb}} = 0.44. \quad (7.8)$$

This efficiency factor is what we obtain from the combination of our lepton isolation and jet tagging efficiencies. We have to point out that, for the high mass domain 140-200 GeV and compared to ref. [25], we changed the maximal jet $|\eta|$ from 5 to 4.5. However, for our “parton level” PYTHIA analysis, we kept the old value in order to reproduce the original study. As we use the ratio between our “parton level” and “full showering” results to obtain the correct efficiency (see equation 7.8), the $\epsilon_{\text{tag},j}$ in this case includes also the change in pseudorapidity and a part of the efficiency loss is to be ascribed to the different η -acceptance. One could argue, that the jet pseudorapidity distributions of signal and QCD background are different and therefore one should be careful in multiplying the background by this factor. For $M_H \leq 140$ GeV, the η_j -acceptance in reference

⁸No efficiency for the lepton selection was given in reference [25].

7.3 Results

[26] is instead identical to ours.

If we multiply the obtained efficiency factor with, for example, the $t\bar{t} + jets$ cross section given in reference [25] after the minijet cut, we obtain:

$$1.48 \text{ fb} \cdot \epsilon_{\text{iso.}l+\text{tag.}j} = 0.65 \text{ fb.} \quad (7.9)$$

This is “our” $t\bar{t} + jets$ background value, which substitutes the “R&Z” one given in table 7.4. The same procedure has to be followed for all background contributions.

selection cuts	R& Z		PYTHIA signal	
	signal	Σ B's	parton level	full showering
$\sigma \times \text{BR}$	-	-	14.3	14.3
2 isolated leptons & ≥ 2 jets	-	-	8.42	5.06
forward tagging i-vi	-	-	2.49	1.34
m_{jj} & "V-A" vii-viii	-	-	1.93	0.89
τ rejection & $m_T(WW)$ ix-x	2.08	-	1.56	0.71
$\Delta\phi(l, \not{p}_t), \not{p}_t$ & p_{tH} xi-xii	1.84	-	1.40	0.63
same flavor xiii-xiv	1.56	-	1.21	0.54
$\epsilon_{\text{minijet veto}}$	1.41	1.4		
$\epsilon_{\text{tag } j}$ & $\epsilon_{\text{iso } l}$	0.94	0.94		

Table 7.5: Cross section rates (in fb) for different selection cuts, for the Higgs signal ($M_H = 115$ GeV) and for the total background. In the first two columns we report the results given in reference [26] for the process $H \rightarrow WW \rightarrow l^\pm l^\mp \not{p}_t$ (labeled as "R&Z"). For our "PYTHIA" analysis we show the signal rates obtained either at parton level or after the hadronization. The "R&Z" values reported in the bottom lines were obtained by multiplying the cross sections by the different efficiency factors given in reference [26].

For $M_H \leq 140$ GeV we proceed in the same way, but in this case we take advantage from the identical upper $|\eta_j|$ -bounds used either in the reference study [26] and in our analysis. In table 7.5, for instance, we show the results obtained with a 115 GeV Higgs and by taking the $e^\pm\mu^\mp$, $e^\pm e^\mp$ and $\mu^\pm\mu^\mp$ decays into account. The background rates given in ref. [26] are also reported. In this case, the parton level cross sections obtained with PYTHIA (reported in the third column) are already about 20-25% smaller (depending on the cut) than the “R&Z” ones (see first column). Starting instead from the “detectable” particles, after all cuts the PYTHIA result drops then to 0.54 fb, that is $0.54/0.94 = 57\%$ of the rate given in ref. [26] after applying the efficiency factors (see last line in table 7.5).

M_H [GeV]	signal		Σ backgrounds		required $\int \mathcal{L} dt$ for a $5\sigma_{\text{Gauss}}$ significance	
	R&Z [fb]	PYTHIA [fb]	R&Z [fb]	PYTHIA [fb]	R&Z [fb $^{-1}$]	PYTHIA [fb $^{-1}$]
110	0.51	0.27	0.82	0.61	95	210
115	0.94	0.54	0.94	0.70	35	60
120	1.62	0.92	1.04	0.77	15	23
125	2.5	1.46	1.15	0.86	8	14
130	3.7	2.2	1.27	0.95	4	8
140	6.1	3.8	1.5	1.1	2	3.7
150	4.8	4.6	1.2	1.4	2.7	3.1
160	7.5	7.1	1.6	1.9	1.7	1.9
170	7.3	7.3	1.6	1.9	1.5	1.8
180	6.0	6.2	1.5	1.8	2.2	2.2
190	4.2	4.4	1.6	1.9	3.7	3.9
200	3.3	3.45	1.6	1.9	6.2	5.7

Table 7.6: Comparison between our PYTHIA analysis and the expected cross sections given in references [26] and [25] (labeled “R&Z”). The PYTHIA results include also same flavor decays, while for the “R&Z” analysis this is valid only up to 140 GeV. In the last columns we report the minimal integrated luminosity required for a 5σ signal (using Poisson statistics).

7.3 Results

Following the same procedure as for a 160 GeV Higgs, we obtain an efficiency $\epsilon_{\text{iso.}l} \cdot \epsilon_{\text{tag.}j} = 0.5$. If we multiply the total “R&Z” background (after the “minijet veto”) by this factor, we get a rate of 0.7 fb, which is the value for the background that we use to obtain the signal significance.

For the Higgs mass ranges $110 \text{ GeV} \leq M_H \leq 200 \text{ GeV}$, the results are summarized in table 7.6. The PYTHIA background has been extracted from the “R&Z” one, as explained before for a 115 and a 160 GeV Higgs.

However, an additional remark has to be made: for the $140 \text{ GeV} \leq M_H \leq 200 \text{ GeV}$ domain, we include also the “same flavor” WW decays to $e\nu e\nu$ and $\mu\nu\mu\nu$, which were instead excluded in the original study [25]. In order to obtain the correct rates, we multiply the background cross sections given in [25] by the same factor which differentiates the “different flavor” and the “both flavors” backgrounds in the other study [26] at lower Higgs masses. This factor appears to be very stable, with values oscillating between 1.9 and 1.92⁹.

In the last two columns of table 7.6, the minimal integrated luminosity required for a 5σ signal (using Poisson statistics) is listed for both the original reference [25, 26] and our PYTHIA studies.

For Higgs masses below 140 GeV, the “ 5σ ” luminosities given in our PYTHIA simulations are about a factor 2 larger than the ones from reference [26]. For higher masses, the inclusion of the “same flavor” decays in our PYTHIA analysis compensates essentially the efficiency loss.

Why PYTHIA produces lower rates?

What is the source of the discrepancy in efficiency between the previous works [25, 26] and our PYTHIA studies? Part of the loss appears already by comparing our parton level study with the “R&Z” cross sections: the PYTHIA expectations are about 15-25% smaller. The difference between a parton level and a “full showering” analysis has to be accounted for the remaining loss.

As can be observed in both tables 7.5 and 7.9, the 1 to 2 ratio between “full showering” and parton level cross sections is already present in events with two isolated leptons and two or more jets, *i.e.* before the selection cuts. Thus, once leptons and jets have been selected, the “full showering” events pass through the cuts with about the same efficiency as the parton level ones. One could therefore conclude that the source of the different efficiency lies only in the lepton and jet reconstruction phase prior to the selection cuts. However, this simple analysis seems not to be accurate enough. In fact, there are several factors which contribute to the difference between a “parton level” and a “full hadronization” analysis. In section 3.3.2 we described the algorithms for the selection of isolated leptons and for the reconstruction of jets. For instance, before any other selection cut, the probability that both leptons are isolated in a $qq \rightarrow qqH \rightarrow qqWW$ event

⁹We do not expect this value to change much, since the additional cuts added in order to include the “same flavor” cases are nearly the same in both studies.

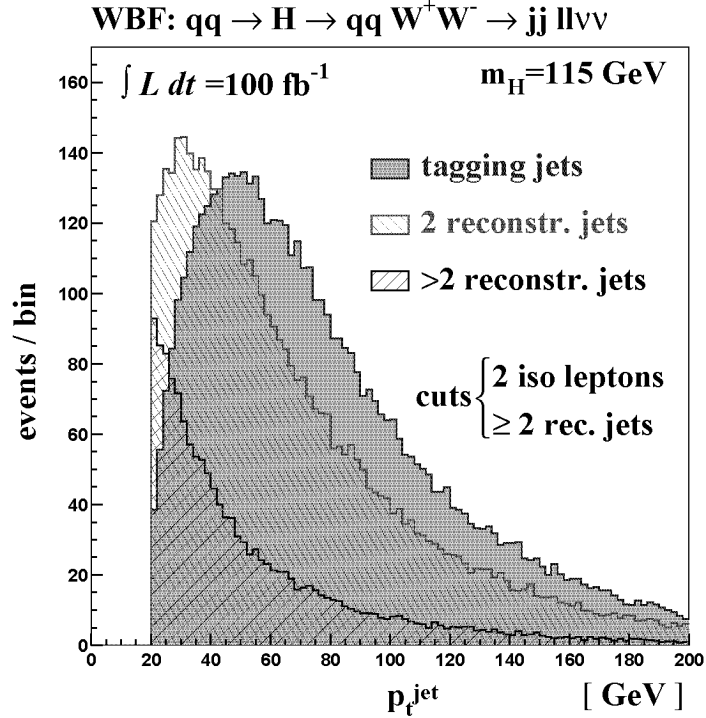


Figure 7.25: Transverse momentum of tagging quarks and reconstructed jets in $qq \rightarrow qqH \rightarrow qqWW$ events.

is about 60% (and is stable for the mass range $110 \text{ GeV} \leq M_H \leq 300 \text{ GeV}$). If we require in addition the presence of at least two jets, this probability increase to more than 70%. Finally, after all selection cuts, in about 97% of the events both leptons are consistent with our isolation criteria! The reason of this increased efficiency is simple: the cuts favor the selection of high- p_t leptons which are well separated to other objects. Thus, after all cuts, both leptons are mostly automatically isolated.

Coming back to the ratio $\sigma_{\text{full showering}}/\sigma_{\text{parton level}} \approx 0.5$ observed at the beginning of the event selection, we can assert that, if the lepton isolation contributes with a 0.7 factor, the selection of two or more reconstructed jets has to be accounted for about the same contribution, *i.e.*

$$\epsilon_{jets} = \frac{\epsilon_{\text{tot}}}{\epsilon_l} = \frac{0.5}{0.7} \approx 0.7.$$

Through the different cuts, the importance of the lepton efficiency drops, but this change is compensate by a smaller efficiency for the jet reconstruction, leading to the observed stability for the ratio $\sigma_{\text{full showering}}/\sigma_{\text{parton level}}$. At the end of the selection, the lepton isolation plays only a negligible role and the efficiency depends

instead almost only on the difference between tagging quarks and reconstructed jets.

In figure 7.25, the transverse momentum distribution of tagging quarks and reconstructed jets is shown in case that 2 isolated leptons have been selected (for $M_H = 115$ GeV). After the hadronization, the presence of quarks and gluons outside the pseudorapidity or p_t acceptance (and, eventually, of some invisible contribution) gives rise to reconstructed jets with smaller p_t in comparison with the original quarks. Therefore, the probability for these events to survive the cuts becomes smaller than in a “parton level” analysis.

In addition, before the cuts, about 20% of the events have 3 or more jets (either coming from tagging quarks which split into several jets or else arising from other partons). As shown in figure 7.25, in this case the p_t of the reconstructed jets is much smaller than those of the tagging quarks. Furthermore, the cuts VI or vi (depending on the mass domain) eliminate all the events with additional jets lying between the tagging jets.

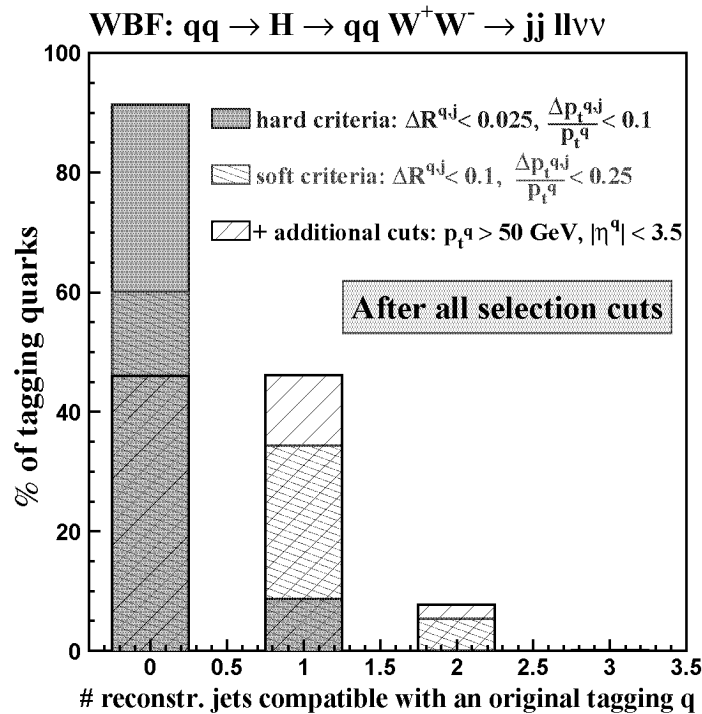


Figure 7.26: $qq \rightarrow qqH \rightarrow qqWW$: by applying all selection cuts I-XIII, the histograms show the percentages of tagging quarks, whose kinematics is compatible with, respectively, 0, 1, 2 or 3 reconstructed jets after the hadronization. The percentages are calculated for three different sets of criteria.

Figure 7.26 shows, starting from each tagging quark, the number¹⁰ of reconstructed jets which are found within a maximal ΔR and $\Delta p_t/p_t$ (after all cuts). For only about 35% of the tagging quarks (and thus $0.35^2 \approx 10\%$ of the events), one finds a reconstructed jet which is compatible with $\Delta R < 1$ and $\Delta p_t/p_t < 0.25$. This means that a large fraction of the reconstructed jets do not reproduce exactly the kinematics of the original tagging quarks.

Given the smaller jet- p_t observed in the “full showering” analysis and, in general, the different kinematics of the reconstructed jets in comparison with the tagging quarks, the “full showering” events are more likely eliminated by the different selection cuts as the “parton level” ones. Therefore, the jet efficiency factor ϵ_{jets} decreases from 0.7 before the cuts to about 0.5 at the end of the selection.

7.3.2 $110 \text{ GeV} \leq M_H \leq 200 \text{ GeV}$: a full PYTHIA analysis

The results presented in section 7.3.1 are based on the backgrounds given in references [25, 26]. As for the $200 \text{ GeV} \leq M_H \leq 300 \text{ GeV}$ mass domain (see next section), we perform the full analysis with PYTHIA also for masses between 110 and 200 GeV, keeping in mind that an important source of background, namely the $t\bar{t} + jet(s)$, is not properly included in this Monte Carlo generator. These events arise from QCD corrections to the basic $t\bar{t}$ -production, which lead to additional hard parton emissions:

$$gq \rightarrow t\bar{t}q, \quad g\bar{q} \rightarrow t\bar{t}\bar{q}, \quad q\bar{q} \rightarrow t\bar{t}g, \quad gg \rightarrow t\bar{t}g,$$

and similarly for $t\bar{t}jj$ subprocesses. The omission of these higher order processes affects to some extent the completeness of this study and our analysis should be repeated, once more accurate estimates of this background will be available.

As background, we take the following signatures into account:

- (a) $q_i\bar{q}_i, gg \rightarrow t\bar{t} \rightarrow WbWb \rightarrow jj\,l\nu l\nu$
- (b) $q_i\bar{q}_i \rightarrow WW \rightarrow ll\, \cancel{p}_t$
- (c) $q_i\bar{q}_i, q_i g \rightarrow Z + jets$
- (d) $q_i q_j \rightarrow q_i q_j WW \rightarrow jj\, ll\, \cancel{p}_t$
- (e) $q_i q_j \rightarrow q_i q_j (W, Z)Z \rightarrow jj\, ll(l)\, \cancel{p}_t$

The simulations include “same flavor” decays and multiple interactions¹¹. Furthermore, also $W \rightarrow \tau\nu$ decays are included.

¹⁰If this number is equal or larger than 2, it means that the original tagging quark is compatible with a combination of several jets (*i.e.* the quark splits into two or more jets).

¹¹The so called “Minimum Bias Events”. We use the simplest (default) scenario, with an abrupt $p_{t\,min}$ cut-off by 1.4 GeV [106].

7.3 Results

M_H [GeV]	qqH signal		$t\bar{t}$ background		Σ B's [fb]	required $\int \mathcal{L} dt$ for a $5\sigma_{\text{Gauss}}$ significance [fb $^{-1}$]
	parton level [fb]	full showering [fb]	parton level [fb]	full showering [fb]		
110	0.62	0.31	0.39	1.96	~ 3	780
115	1.22	0.61				202
120	2.09	1.07				65
125	3.36	1.71				26
130	5.00	2.55				12
140	8.47	4.47				6
150	11.6	5.35	0.92	4.45	5.3	6
160	17.9	8.11				3
170	18.6	8.31				3
180	15.7	7.25				4
190	11.1	5.03			5.7	6
200	8.67	4.03			5.8	12

Table 7.7: PYTHIA expectations for the $qqH \rightarrow qqWW$ signal, the $t\bar{t}$ background and the sum of all included backgrounds. In the last columns we report the minimal integrated luminosity required for a 5σ signal (using Poisson statistics).

The results of this “full PYTHIA” study are shown in table 7.7. The $t\bar{t}$ background is given both at parton level and after the full showering process. In agreement with the expectations, the parton level value given in reference [25] for $M_H = 160$ GeV (5.1 fb, as shown in table 7.4) is larger than the PYTHIA one (0.92 fb). However, after the application of the “minijet veto” and of the lepton/jet efficiency factors, the remaining $t\bar{t}$ background cross section reported in [25] is only about 1.1 fb. In our PYTHIA analysis we observe instead an increase and the “full showering” $t\bar{t}$ cross section is found to be about a factor of 4 larger, e.g. 4.45 fb. Multiple interactions or minimum bias events contribute only marginally to this difference and we obtain actually about the same increase also in simulations where these processes are not included. In table 7.8 we show the rejection powers of cuts I-IX for the $t\bar{t}$ background. The goal is to find out to what extent the “full showering” analysis differs from the parton level one, *i.e.* if the selection cuts, which have been developed at parton level, lose part of their effectiveness when applied on the leptons and jets found in the complete PYTHIA analysis. As shown in table 7.8, once events with two isolated leptons and at least two jets have been selected, the parton level PYTHIA cross sections are more than a factor 2 larger than the “full showering” ones. The contribution given by events with more than 2 jets is similar to those with exactly 2 jets.

selection cuts	R&Z	parton level	PYTHIA		
			2 jets	> 2 jets	≥ 2 jets
2 isolated leptons & ≥ 2 jets	-	16130	3898	+ 2998	= 6897
η & p_t I-II	-	$\rightarrow \times 79\%$ 12797	$\rightarrow \times 96\%$ 3729	-	
hemisphere III-IV	-	$\rightarrow \times 4.5\%$ 570	$\rightarrow \times 4.9\%$ 183	-	
$\Delta\eta_{tags}$ V	1080	$\rightarrow \times 7.0\%$ 39.8	$\rightarrow \times 15\%$ 26.9	-	
m_{jj} VII ^a	-	$\rightarrow \times 18\%$ 7.1	$\rightarrow \times 43\%$ 11.6	+ 195.5	= 207.1
b_{jets} VI ^a	64	7.1	11.6	+ 6.4	= 18.1
"V-A" VIII	-	$\rightarrow \times 17\%$ 1.2	3.7	+ 2.6	$\rightarrow \times 35\%$ = 6.3
$p_{t_{e,\mu}}$ & m_{ll} IX	5.5	1.2	3.5	+ 2.5	= 6.0

Table 7.8: Rejection power of cuts I-IX against the $t\bar{t}$ background. All the cross section values are given in fb. The arrows indicate the percentage of events given in the preceding rows which survive the corresponding cut.

^aIn this table, we inverted the order of cuts VI and VII, because the first one has no effect on the PYTHIA 2-jet events, while cuts I-V and VII are used in our analysis to select the tagging jets in events with more than 2 jets.

As indicated in percentage, the “full showering” events pass more easily through the first cuts I-II in comparison to the parton level events. This difference in efficiency (about 20%) arises mainly from the minimal ΔR_{jj} and ΔR_{jl} separations. This happens because parton level leptons and quarks could still have close directions, while the reconstruction and isolation algorithms gave only leptons and jets already separated from each other. The “hemisphere” cuts III-IV show similar rejection powers for both parton level and “full showering” events with 2 jets.

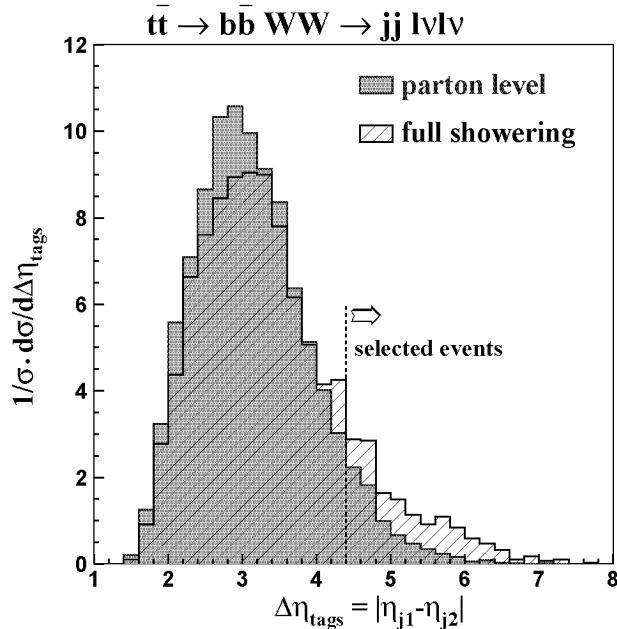


Figure 7.27: PYTHIA expectation of the separation in η between the two tagging jets, which arise in the background process $t\bar{t} \rightarrow q\bar{q}WW \rightarrow jj\,l\nu l\nu$ (after cut IV).

One observes instead different efficiencies by cut V. In this case, the number of “full showering” 2-jet events which survive the cut is twice as large as for parton level events. The related difference in pseudorapidity $\Delta\eta$ between the tagging jets is shown in figure 7.27. As can be observed, the “full showering” events have a larger percentage of events above the cut-threshold as the parton level ones. This results from the presence of other jets besides those arising from top decays. For “full showering” events, these additional jets can be identified as tagging jets, while at parton level the two tagging jets coincide always with the two b quarks coming from the $t \rightarrow Wb$ decay.

At this level of analysis, the $t\bar{t}$ background cross section given in reference [25] is about 25 times larger than our parton level $\sigma_{t\bar{t}}$, as expected from the lack of higher order QCD $t\bar{t}$ processes in the PYTHIA Monte Carlo.

The larger η -separation between the “full showering” tagging jets (see figure 7.27) produces also larger m_{jj} values in comparison to the parton level analysis. The rejection power of cut VII is therefore much larger at parton level, as shown in table 7.8.

After cut VII (but without cut VI, for the reason explained in the footnote of table 7.8), the PYTHIA cross section for events with more than 2 jets is much larger than the 2-jet one. The two b -quarks coming from the $t \rightarrow Wb$ decay can split into two or more jets and eventually other jets can arise from initial state radiation. The

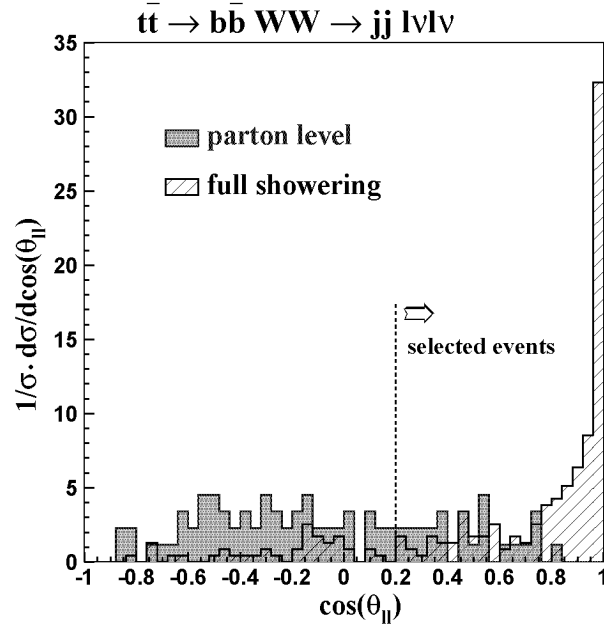


Figure 7.28: PYTHIA expectation for the cosine of the polar angle between the two leptons. These $t\bar{t} \rightarrow q\bar{q}WW \rightarrow jj l\nu l\nu$ background events are shown after cut VII. The request for high-mass dijets (cuts V and VII) should force the two leptons coming from the $t \rightarrow Wb \rightarrow l\nu jet$ decays to be well separated. At parton level, since both jets are directly identified with the two b quarks, this constrain is always successful. It is instead no longer valid in the “full showering” analysis if one of the two tagging jets comes from initial state radiation or from underlying events. In this case, the b jets, and thus the two leptons, can have close directions.

soft radiation processes present in PYTHIA reproduce to some extent the “soft” part of the next-to-leading order $t\bar{t}$ processes. Therefore, in one sense, the “full showering” analysis includes also part of the higher order diagrams present in the original study [25], while our parton level results are limited to the leading order diagram. This difference produces the large discrepancy between our “parton level” and “full showering” analysis.

After having applied cut VI, which eliminates a large fraction of PYTHIA events with more than 2 jets, the “full showering” cross section is now found to be more than twice as large as the parton level one. In other words, we have a reversed situation in comparison to the one observed before the cuts. The $t\bar{t}$ cross section given in ref. [25] is still larger compared to PYTHIA, but the lepton and jet efficiency as well as the “minijet veto” have not yet been applied. With these factors ($0.74 \times 0.29 = 0.21$), the “R&Z” value after cut VII drops to 13.7 fb, that is already less than the PYTHIA expectation with full showering! The situation

change further with cut VIII. As shown in table 7.8, the cut eliminates 83% of the events at parton level, while “only” 65% of “full showering” events are rejected. The reason of this different efficiency can be extracted from figure 7.28. A cut on the cosine of the polar angle between the two leptons seems to be very efficient at parton level. This happens because the wide opening angle requested for the tagging jets (through $\Delta\eta_{tags} \geq 4.4$ and $m_{jj} > 650$ GeV) forces the two b -jets to be also well separated and thus the two leptons, which arise from the $t \rightarrow Wb \rightarrow l\nu jet$ decays, are consequently well separated too. However, if one of the tagging jets comes from initial state radiation, this constraint disappears and the amount of “full showering” events which survive the cut becomes clearly larger.

The rejection power of the cuts proposed in ref. [25] against the $t\bar{t}$ background is thus found to be too optimistic. The additional emission of partons at initial state and the hadronization process seem to clearly change the acceptance. After cut IX and by taking the l 's, jets and “minijet veto” efficiencies into account, the “R&Z” expectation is $5.5 \text{ fb} \times 0.21 = 1.2 \text{ fb}$, while what we obtain with PYTHIA is about 6 fb (see table 7.8). Therefore, even if the $\mathcal{O}(\alpha_s)$ QCD corrections are not fully included in PYTHIA, the $t\bar{t}$ background cross section is a factor 5 larger! Besides the sizeable $t\bar{t}$ contribution, table 7.7 shows also that the sum of all included backgrounds becomes larger with increasing Higgs mass. This is mostly due to the non-resonant weak boson fusion background (processes (d) and (e)), which depends strongly on the Higgs mass. Its rate is negligible below $M_H = 190$ GeV, whereas it becomes very important above 200 GeV, as shown in the next section.

Finally, the contribution of the $q_i\bar{q}_i, q_i g \rightarrow Z + jets$ background is very difficult to estimate. The very large cross section implies a large statistical uncertainty for the predicted number of events from this background process. Furthermore, this background could get a significant contribution from additional jets in the underlying events and from “pile-up”¹² or detector effects. However, it affects only the “same flavor” signatures ee and $\mu\mu$ and a comparison with the $e\mu$ rates should provide a powerful means to control the $Z + jets$ background.

In conclusion, the “full PYTHIA” study presented in this section confirms the observability of a 130-200 GeV Higgs produced via weak boson fusion. The minimal integrated luminosity required for a 5σ discovery remains below 20 fb^{-1} . However, the signal to background ratios are found to be clearly smaller in comparison with the previous expectations [25, 20]. S/B values of the order of 1-1.5 could be affected by the background uncertainties and consequently limit the discovery potential of these weak boson fusion signatures. The different rejection powers against the $t\bar{t}$ background, found at parton level or with the full showering analysis, put for instance in evidence the need of a more complete PYTHIA-like study,

¹²Several independent hadron-hadron interactions generated in the same bunch-bunch crossing.

which should also include all $\mathcal{O}(\alpha_s)$ QCD corrections to the $t\bar{t}$ production.

7.3.3 200 GeV $\leq M_H \leq$ 300 GeV

For a Higgs in the 200 GeV $\leq M_H \leq$ 300 GeV mass domain, we can not compare with previously estimated values. For the background, we include the same processes (a)-(e) as in section 7.3.2.

The cross section values for the different selection criteria, is reported in table 7.9 for a 260 GeV Higgs and the corresponding backgrounds.

The largest background comes from $t\bar{t}$ production. As was pointed out in the previous section, precise estimations of this background would require a “full showering” analysis, using the complete next-to-leading order diagrams.

selection cuts	signal		$t\bar{t}$	WW	Zj	$qqWW$	other $qqVV$
	parton level	full showering	(a)	(b)	(c)	(d)	(e)
$\sigma \times \text{BR}$	115.3	115.3	65400	8060	2.46×10^6	53.17	41.45
$2l + \geq 2j$	29.97	15.49	6899	25.7	36950	7.89	4.01
i-ii	20.56	10.67 ^a	1538 ^a	13.6 ^a	17840 ^a	5.50 ^a	2.53 ^a
iii-iv	16.35	8.06 ^a	104.8 ^a	1.93 ^a	2173 ^a	3.88 ^a	1.75 ^a
v	12.50	6.12 ^a	18.0 ^a	0.90 ^a	778 ^a	2.76 ^a	1.26 ^a
vi	11.11	6.35	156.9	0.48	564	2.74	1.34
vii	11.11	5.81	11.4	0.31	287	2.57	1.18
viii	9.86	5.19	8.5	0.17	263	2.36	1.11
ix	8.43	4.36	5.8	0.10	233	1.95	1.02
x	7.15	3.66	5.0	0.08	117	1.65	0.88
xi-xii	6.15	3.12	4.2	0.07	≈ 1.5	1.41	0.07
xiii	5.54	2.71	2.9	0.05	≈ 1	1.23	0.06

Table 7.9: PYTHIA cross sections (in fb) for different analysis cuts, either for a signal $qq \rightarrow qqH$, with $H \rightarrow WW \rightarrow ll \not{p}_t$ ($M_H = 260$ GeV), or for the different backgrounds (see text). The signal rates are reported both at parton level and after the hadronization.

^aThese values include only events with exactly 2 jets (since the combination of these cuts is used to select the tagging jets in events with more than 2 jets).

7.3 Results

In fact, already for the PYTHIA results, events with more than two reconstructed jets have large rates. These events would be further increased once the full higher order diagrams would be properly included.

After all cuts, the non-resonant weak boson fusion signatures $q_i q_j \rightarrow q_i q_j WW \rightarrow jj ll \cancel{p}_t$ produce also a large rate (despite their relatively small cross section). This is expected, since this process is kinematically very similar to the weak boson fusion signal.

An important source of error arises from the uncertainties related to $q_i \bar{q}_i$, $q_i g \rightarrow Z + jets$ processes. Because of the large cross section, the $Z + jets$ productions are affected by small statistics and by uncertainties related to the presence of additional jets.

Finally, in table 7.10 we summarize the cross sections obtained after all selection cuts. The values are given for signal and total background, and give rise to a 5σ significance for the integrated luminosity shown in the last column. The increasing importance of the non-resonant weak boson fusion background is remarkable. For a 300 GeV Higgs it provides the largest contribution to the overall background.

M_H [GeV]	qqH signal [fb]	$t\bar{t}$ [fb]	WW [fb]	Zj [fb]	$qqWW$ [fb]	other $qqVV$ [fb]	Σ B's [fb]	required $\int \mathcal{L} dt$ for a $5\sigma_{\text{Gauss}}$ significance [fb $^{-1}$]
220	3.14	3.54	0.06	~ 1	0.61	0.05	~ 5.3	14
240	2.93	3.27	0.06	~ 1	0.74	0.05	~ 5.1	15
260	2.71	2.90	0.05	~ 1	1.23	0.06	~ 5.3	16
280	2.44	2.35	0.03	~ 0.5	1.84	0.06	~ 4.8	20
300	2.20	2.02	0.03	~ 0.5	2.06	0.07	~ 4.7	24

Table 7.10: Expected signal and background cross sections for $200 \text{ GeV} \leq M_H \leq 300 \text{ GeV}$. In the last column we report the minimal integrated luminosity required for a 5σ signal (using Gaussian statistics).

7.4 Required integrated luminosity for a Higgs discovery: an update

The results given in this chapter contribute to improve the overall discovery potential for a Standard Model Higgs at LHC. What can be obtained, using only the weak boson fusion signatures discussed so far, is summarized in figure 7.29. The required luminosity needed to reach a 5σ signal is shown either for the reference studies [25, 26], for our “hybrid” PYTHIA study with modified “R&Z” backgrounds (results of section 7.3.1) or for the analysis exclusively performed with PYTHIA (see sections 7.3.2 and 7.3.3).

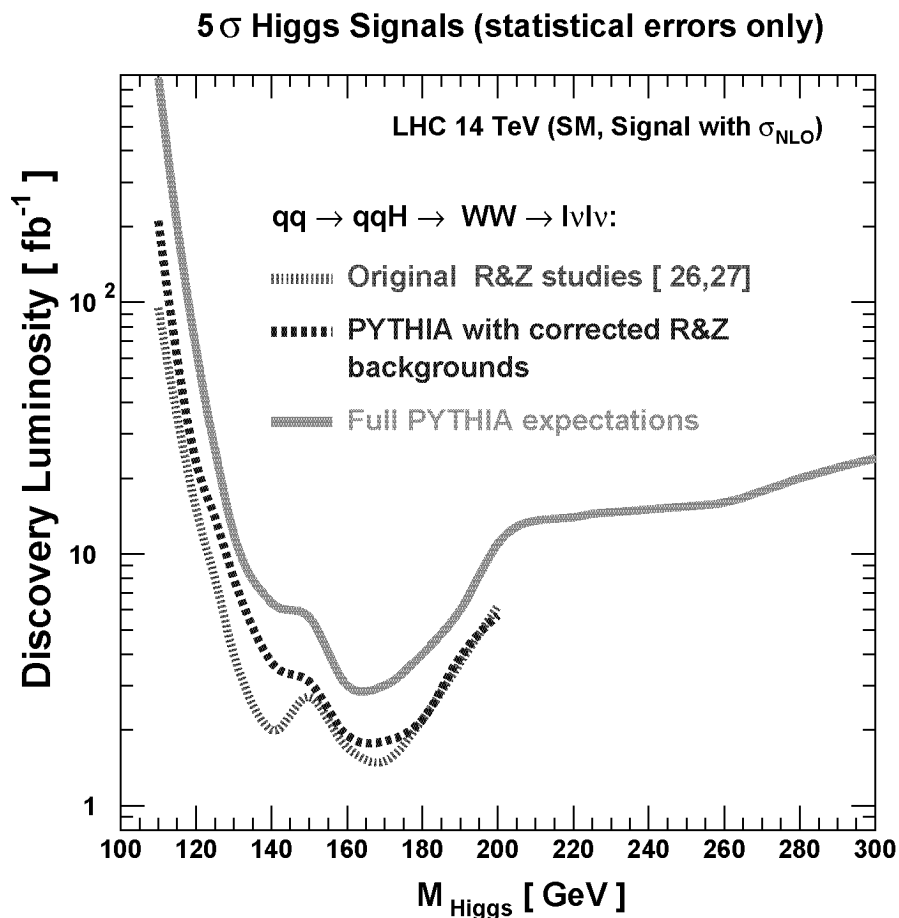


Figure 7.29: Required luminosity for a 5σ discovery of a SM Higgs through weak boson fusion signatures (for $110 \text{ GeV} \leq M_H \leq 300 \text{ GeV}$). The curve obtained from ref. [25, 26] has a larger step in comparison with both PYTHIA studies. This is due to the lack (in [25]) of the “same flavor” $ee, \mu\mu$ signatures above 140 GeV.

The “step” present in all curves in correspondence with the 140-150 GeV mass region is due to the switching between the two analysis, which are optimized for, respectively, the $110 \text{ GeV} \leq M_H \leq 140 \text{ GeV}$ and the $140 \text{ GeV} \leq M_H \leq 200 \text{ GeV}$ mass domains. The larger step, observable between the reference studies [25] and [26], is due to the omission of the “same flavor” $ee, \mu\mu$ signatures in the $140 \text{ GeV} \leq M_H \leq 200 \text{ GeV}$ domain. Above 200 GeV, the switch between l-XIII and **i-xiii** cuts produces another “step” in the “full PYTHIA” curve.

Given the absence of next-to-leading order $t\bar{t}$ diagrams in PYTHIA, which lead to additional hard jet emissions, one would expect larger significances (and thus smaller discovery luminosities) from the “full PYTHIA” analysis in comparison with the “hybrid” PYTHIA study, which inherits the NLO $t\bar{t}$ -backgrounds from references [25, 26]. On the contrary, the required discovery luminosity, obtained with a full PYTHIA study for both signal and background, is larger, as shown in figure 7.29. As reported in sections 7.3.2 and 7.3.3, the inclusion of multiple interactions, initial state radiation and full hadronization process leads actually to enhanced rates for some backgrounds and, eventually, to the emergence of new ones. This fact points out the need for an improved analysis, based on NLO matrix elements with parton shower. Nevertheless, the analysis proposed in references [25, 26], as well as our PYTHIA extensions and the recent study [24], show that the weak boson fusion signatures $qq \rightarrow qqH \rightarrow WW \rightarrow \nu\nu\nu$ should provide well visible signals in at least the 115-600 GeV mass region.

As pointed out at the beginning of this chapter, the availability of several Higgs signatures would allow to extract more informations about the nature of the Higgs couplings to gauge bosons and fermions. Figure 7.30 shows the updated status for a Standard Model Higgs discovery at LHC. The different curves are obtained from PYTHIA simulations and show the required integrated luminosity for a 5σ signal. As can be observed, the additional presence of the new qqH results allows an Higgs discovery in both Higgs production modes¹³ from about 115 GeV up to ~ 600 GeV and with integrated luminosities of less than 100 fb^{-1} . Moreover, the measurement of the Higgs couplings can be further improved, if channels with identical decay modes are used (see for instance figure 7.1). The updated $qq \rightarrow qqH \rightarrow WW \rightarrow \nu\nu\nu$ curve covers practically the whole mass domain shown in figure 7.30, while the $pp \rightarrow H \rightarrow WW \rightarrow \nu\nu\nu$ channel, dominated by the inclusive $gg \rightarrow H$ Higgs production mode, is restricted to $140 \text{ GeV} \lesssim M_H \lesssim 200 \text{ GeV}$. In this mass region, both channels will provide well visible signatures with few fb^{-1} ! Previous preliminary studies [107] have suggested the possibility to extend the $gg \rightarrow H \rightarrow WW \rightarrow \nu\nu\nu$ channel up to $M_H \approx 400 \text{ GeV}$, with equivalent discovery luminosities as the $qq \rightarrow qqH$ one. However, similarly to the weak boson fusion channel, also the gluon-gluon fusion signatures suffer from a poor signal to background ratio and are thus sensitive

¹³For the channels labeled as $pp \rightarrow H \rightarrow \dots$, the Higgs bosons are mostly produced via gg -fusion processes.

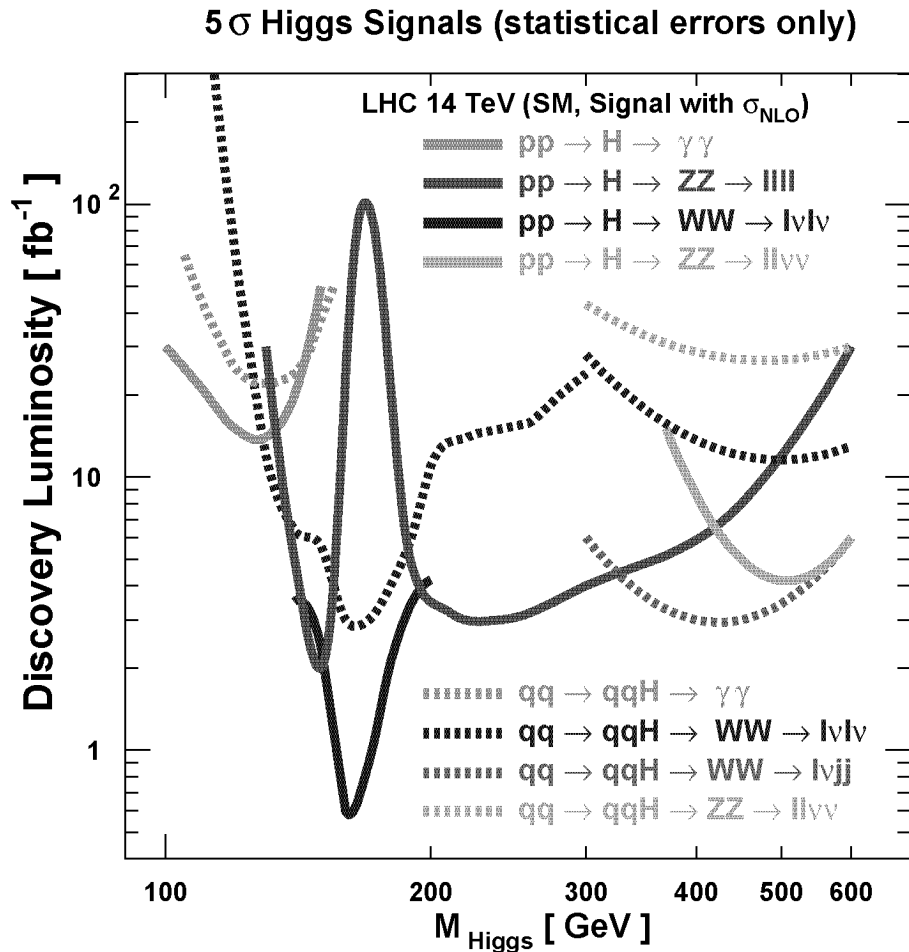


Figure 7.30: Required luminosity for a 5σ discovery of the SM Higgs at LHC (PYTHIA).

to possible large background uncertainties. Nevertheless, if the Higgs mass is in the range of $200 \text{ GeV} \lesssim M_H \lesssim 400 \text{ GeV}$, the Higgs will be probably first discovered through the $H \rightarrow ZZ \rightarrow ll ll$ channel. This would allow to better control the signal and background uncertainties for other channels. Both the $gg \rightarrow H \rightarrow WW \rightarrow l\nu l\nu$ and $qq \rightarrow qqH \rightarrow WW \rightarrow l\nu l\nu$ signatures could profit of this knowledge and become visible also in the 200-400 GeV mass domain.

Chapter 8

Conclusions and outlook

This work focused on the reactions $pp \rightarrow WW, WZ, ZZ$ at LHC. Boson pair productions with leptonic decays are expected to play a central role for many measurements which will be performed with the new 14 TeV pp collider.

First of all, the discovery of the exact nature of the electro weak symmetry breaking mechanism could likely emerge from a resonant production of dibosons. In that case, the simplest scenario would be the discovery of a Standard Model Higgs boson through the $H \rightarrow WW, ZZ$ decay modes. To ensure the visibility of these signatures, precise estimations for the signal and for the non-resonant WW and ZZ backgrounds are mandatory.

Alternatively, in case that no light scalar exists, some new physics should contribute to the normalization of the diboson cross section in the TeV range.

Furthermore, apart for the solution of the symmetry breaking mechanism, these signatures provide a way to test the vector boson self-interactions, where, for example, anomalous effects might appear.

The starting point of such analysis will be a comparison of the data with the Standard Model expectations for the WW , WZ and ZZ cross sections. This is valid also for many other new physics, which could reveal signatures similar to diboson events.

In order to achieve a precise estimation of the resonant and non-resonant production of dibosons, several experimental and theoretical problems need to be solved. In this work, some aspects of the overall problematics have been discussed.

First, a new approach to the LHC luminosity measurement was investigated. In this study it was shown that the pseudorapidity distributions of ℓ^\pm 's originating from weak boson decays are directly related to the fractional momenta x of quarks and antiquarks within the protons. The shape and rate of the η_ℓ distributions provide therefore the key to precisely constrain the quark and antiquark structure functions and to monitor their corresponding luminosities. With the well known W^\pm and Z^0 masses, possible x values of quarks and antiquarks are obtained from $M_{W^\pm, Z^0}^2 = s \cdot x_q \cdot x_{\bar{q}'}$, where $s = 4E_{\text{beam}}^2$. The product $x_q \cdot x_{\bar{q}'}$ at the LHC is thus

fixed to $\approx 3 \times 10^{-5}$.

A similar approach, using gluon related scattering processes, might eventually lead to similar accuracies also for the x distribution of gluons.

Once the quark and antiquark luminosities at $Q^2 \approx 10^4 \text{ GeV}^2$ and in the x range between $\approx 5 \times 10^{-4}$ and 10^{-1} are determined, accurate rate predictions of other $q\bar{q}$ scattering processes are possible. Assuming collisions of essentially free partons, the production of weak bosons, $u\bar{d} \rightarrow W^+ \rightarrow \ell^+\nu$, $d\bar{u} \rightarrow W^- \rightarrow \ell^-\bar{\nu}$ and $u\bar{u}(d\bar{d}) \rightarrow Z^0 \rightarrow \ell^+\ell^-$ are in lowest order understood to at least a percent level. Cross section uncertainties from higher order QCD corrections are certainly larger, but are obviously included in the measured weak boson event rates. Similar higher order QCD corrections to other $q\bar{q}$ scattering processes at different Q^2 , like $q\bar{q} \rightarrow W^+W^-$, can be expected. Thus, assuming that the Q^2 dependence can in principle be calculated, very accurate theoretical predictions for cross section ratios like $\sigma(pp \rightarrow W^+W^-)/\sigma(pp \rightarrow W^\pm)$ should be possible. For example, the strong correlation between the weak boson pair production and the single boson production leads to an estimated parton luminosity accuracy at the $\pm 1\%$ level. This should be compared to the often considered optimistic goal of $\pm 5\%$ accuracy.

This ability to monitor the expected rate of diboson events will improve the potential of LHC for any study which involves the resonant and/or non resonant production of weak boson pairs.

In the subsequent chapter, the probability for charged pions to be identified as electrons was investigated using GEANT simulations and test beam results.

If one takes into account the energy left in the electromagnetic calorimeter alone (*i.e.* without read-out effects), the probability that a charged pion leaves more than 95% of its initial energy is about 0.01% for 10 GeV, 0.02% for 20 GeV and about 0.001% for 50 GeV.

The fraction of electron-like pions is then clearly increased by the so-called “Nuclear Counter Effect” within the Avalanche Photo-Diodes. With the expected value of this effect ($\approx 100 \text{ MeV}$), the probability of a $\pi^\pm - e^\pm$ misidentification appears then to be between 0.2% and 0.01% for pion momenta varying, respectively, between 5 and 50 GeV.

The test beam results collected in 1998 have shown an even larger Nuclear Counter Effect, which has been qualitatively explained as an effect of the bad contacts between crystals and APD’s. However these results allowed a comparison between simulations and data, confirming the accuracy of the GEANT program and allowing to study some possible methods which reduce the possibility of a $\pi^\pm - e^\pm$ confusion. Using the distribution of the deposited energy within the ECAL and by requiring a small signature within the HCAL, one can eliminate more than 60% of the “fake” electrons.

Important consequences of the pion-electron misidentification might appear in the form of new backgrounds for physics channels. For instance, the presence

of fake electrons, associated with the leptons coming from a real W or Z decay, could give rise to a new source of background for the WW , WZ and ZZ signatures. The size of this background depends strongly on the choice of the lower p_t -threshold for the electron selection. Above $p_t \sim 20$ GeV, this background should be negligible, while for $p_t \sim 5$ GeV the amount of e -like π^\pm 's is very large and could seriously compromise the success of a study.

In chapter 6 the WW , WZ and ZZ productions have been analyzed by including the complete next-to-leading order corrections. These simulations, computed with a new parton level Monte Carlo program (DKS), showed an overall cross sections increase at LHC energies of about a factor 1.5 in comparison with the LO expectations. However, it was demonstrated that these inclusive K-factors can be extremely different to K-factors obtained after some typical selection cuts. In other words, for specific cuts (for instance, searches with high- p_t jets, leptons or \cancel{E}_t) the NLO corrections can be much larger. The analysis performed with the current version of PYTHIA could therefore strongly underestimate the diboson rates, making this generator unsuited for these simulations.

Finally, in the last chapter a visibility study for a Standard Model Higgs in the $110 \text{ GeV} \leq M_H \leq 300 \text{ GeV}$ mass range was performed with PYTHIA. For $M_H \leq 200$ GeV we followed (with some minor modifications) the selection criteria proposed in the previous studies [25, 26], which were done at parton level. It was found that the weak boson fusion process, together with the $H \rightarrow W^+W^- \rightarrow l^+l^- p_t$ decay mode, permits to achieve a very clean and pure signal. This is made possible thanks to a multitude of characteristics which distinguish signal from backgrounds. One of the main features of the $qq \rightarrow qqH$ process is the presence of two forward tagging jets inside the acceptance of the CMS detector. Basic properties of these jets are the sizeable p_t and the large dijet invariant mass ($\sim \mathcal{O}(\text{TeV})$). Besides the two jets, the signal can rely on the Higgs products, which consist in two central, isolated and high- p_t leptons together with a missing transverse energy.

This contrasts with the situation for the typical QCD backgrounds, where the isolated charged leptons and the dijet masses are much softer. Moreover, in these backgrounds leptons normally arise from W bremsstrahlung off forward scattered quarks and have thus high rapidities.

On the other hand the EW processes include not only the signal but also the EW $WWjj$ production, which proceeds also via weak boson fusion. Despite its lower cross section compared to the QCD processes, the kinematical similarities with the signal make these non-resonant weak boson fusion events a dangerous background.

The study resulted in a required integrated luminosity for a 5σ discovery which ranges from 3 fb^{-1} , for a 170 GeV Higgs, up to about 25 fb^{-1} , for $M_H = 125$ or 300 GeV.

These results need eventually to be confirmed by further analysis, as soon as

more precise background expectations will be available. Particularly, higher order corrections to the basic $pp \rightarrow t\bar{t}$ process, which lead to additional real parton emissions (*i.e.* to additional jets), are expected to give the largest background contribution. In PYTHIA, only the “soft” part of these $t\bar{t} + jet(s)$ processes are included. Another important source of error arises then from the uncertainties related to $q_i\bar{q}_i$, $q_i g \rightarrow Z + jets$ processes. Their very large cross section implies a large statistical uncertainty for our simulation. Moreover, the $Z + jet(s)$ background could be enhanced by the presence of additional jets (from underlying events, “pile-up” or detector effects). However, it affects only the ee and $\mu\mu$ signatures and a comparison with the $e\mu$ cross section should provide a powerful tool to control this background.

The availability of more than one channel, in which a Higgs could show up, allows a deeper understanding of the Higgs properties. For instance, the fact that both the weak boson fusion and the gluon fusion production modes are well visible between 140 and 200 GeV, assures a determination of the Higgs couplings to gauge bosons and fermions with an accuracy at the 15% level (with 100 fb^{-1} integrated luminosities).

In the coming years, many problems will still have to be solved in order to fully exploit the weak boson pair signatures at LHC. What has been studied here could serve as incentive for further investigations. For instance, in this work we always made use of fast physics and detector simulations. In general, to improve the accuracy of the different studies, one could instead use (wherever possible) full detector simulations, which would eventually include also higher order QCD corrections to the involved physics processes, minimum bias events, pile-up effects and so on.

Regarding the method to monitor the parton luminosities presented in chapter 4, the achievable statistical accuracy will be probably better than the present theoretical uncertainties. The possibility to measure the x distributions of sea and valence quarks and the corresponding luminosities within $\pm 1\%$ should therefore encourage our theoretical colleagues to match this experimental accuracy.

In the study about the $\pi^\pm - e^\pm$ confusion, it was shown that new kind of backgrounds could emerge from this misidentification. However, to obtain more precise estimations, the knowledge of exact amount of the Nuclear Counter Effect, as well as its effects on the ECAL response, should be improved. A full detector simulation, which would include the magnetic field, could also answer some questions related to Bremsstrahlung effects, which affect the electron identification. In addition, possible consequences for the whole data acquisition chain, by the different trigger levels, should also be investigated in detail.

The study concerning the NLO diboson productions (chapter 6) points out the need of important corrections for some PYTHIA motivated cuts. This is especially important, since a complete multi-jet or no-jet simulation of the WW , WZ and ZZ productions at $\mathcal{O}(\alpha_s)$ is essential for many physics studies. All the analysis

which are affected by large NLO corrections will have then to be repeated, once more precise simulation tools will be available!

Finally, the Higgs signatures discussed in the last chapter rely on the presence of two forward jets. In the future, the jet reconstruction algorithm could be tested with more realistic simulation tools. Furthermore, the uncertainties related to some backgrounds are large and need to be reduced by additional analysis. For example, the implementation of NLO diagrams for the $t\bar{t} + jets$ production and studies with larger statistical accuracy of the $Z + jet(s)$ background will provide a more precise estimation of the signal significance.

The final message could be that, even after having discussed all problems related to the diboson production, one should nevertheless be ready for surprises!

Appendix A

Hardware related topics

Almost all the arguments discussed in this work are related to physics simulations. Chapter 5 represents the only exception: simulations of the energy deposited in a simple ECAL set-up by charged pions and electrons are compared with test beam data collected in 1998 and 1999. In fact, during the last 4 years as a PhD student, I participated to different test beam runs, performed either at CERN in Geneva or at the Paul Scherrer Institute¹ (PSI). For all these tests, the aim was the development of different components of the CMS Electromagnetic Calorimeter (ECAL).

A.1 PSI 1997: crystal and APD tests

During 1997, I spent 6 months at PSI. For the ECAL, this institute is responsible for the APD's and takes part in the crystal and electronics R&D.

At that time, the main problems related to the APD's were the excessive radiation damage suffered by silicon and the temperature dependence of the gain. On the other hand, the PbWO_4 crystals were not well understood and not yet optimized for the CMS ECAL requirements. To solve these problems, an intensive R&D program together with the producers needed to be followed.

The experiments were conducted at PSI, which features of an high intensity proton accelerator complex (see figure A.1). Particularly, we made use of πM1 , which is a high resolution pion beam line with a momentum range between 100 and 500 MeV (see figure A.2). To test the radiation hardness of the APD's, we used also the OPTIS facility, which was originally developed (and is still successfully used) for the treatment of ocular tumors through proton radiotherapy.

¹the Swiss National Laboratory in Villigen, Switzerland.

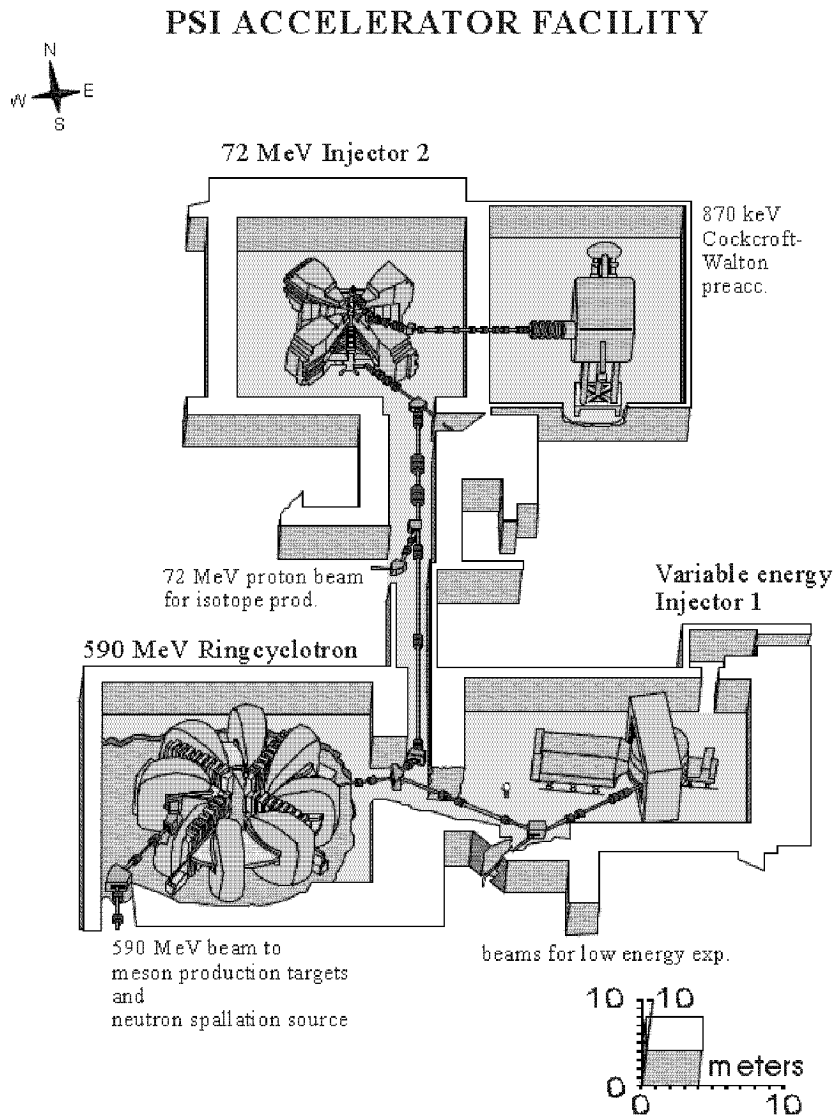


Figure A.1: PSI accelerator overview: a proton energy of 870 keV is obtained from the 60 keV extraction voltage at the ion source and the 810 keV DC acceleration of the Cockcroft-Walton accelerator. The 870 keV beam enters then in the Injector 2 cyclotron, which provides high intensity and quality beams of 72 MeV protons. These are then injected into the 590 MeV Ring cyclotron. This accelerator facility delivers in four experimental areas a large variety of beams for a broad research program.

A.1.1 R&D for the PbWO_4 crystals

During the R&D period, all relevant crystal parameters, like the absolute amount of scintillation light, the decay time of the light, the wavelength of the emitted

light, the uniformity of the light yield and the radiation hardness, had to be measured. The crystals were exposed to the π M1 proton beam with an energy of 84 MeV. To test the radiation hardness, the crystals were irradiated with pions from the π M1 beam and/or photons coming from a Co^{60} source.

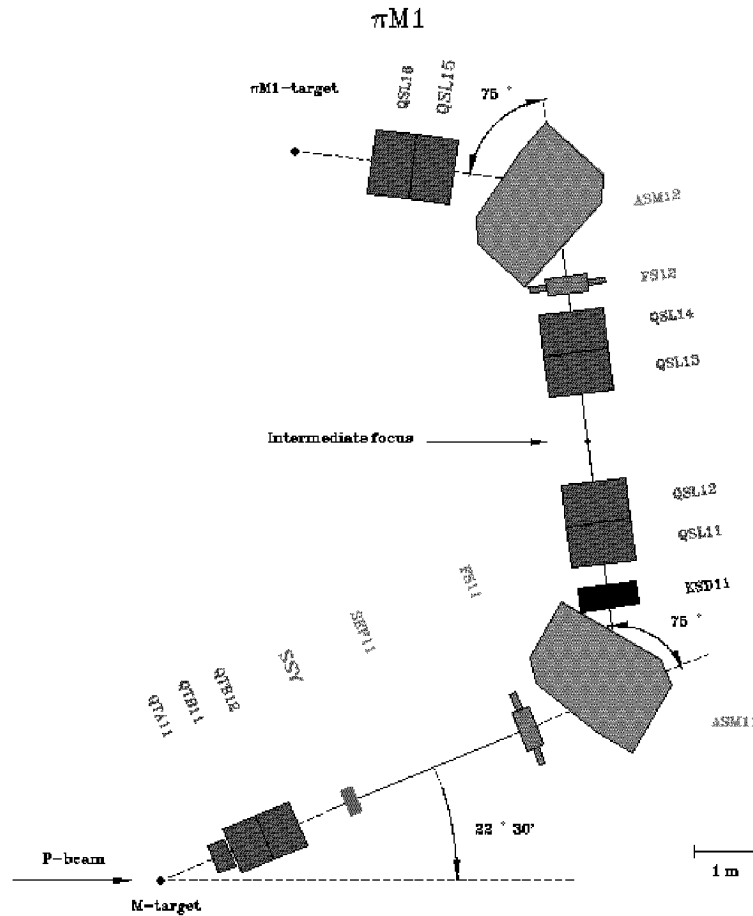


Figure A.2: Overview of the π M1 facility (PSI), which features a momentum resolution of better than 0.1%.

π M1 set-up

Figure A.3 shows the set-up which was used to measure the crystal parameters in the test beam. The particle beam, which could consist of protons, pions or electrons, entered the test facility from the left side. Four planes of Multi-Wire-Proportional-Chambers measured the trajectory of the particle. Its entrance point into the crystal could be determined with a precision of 2 mm in the x

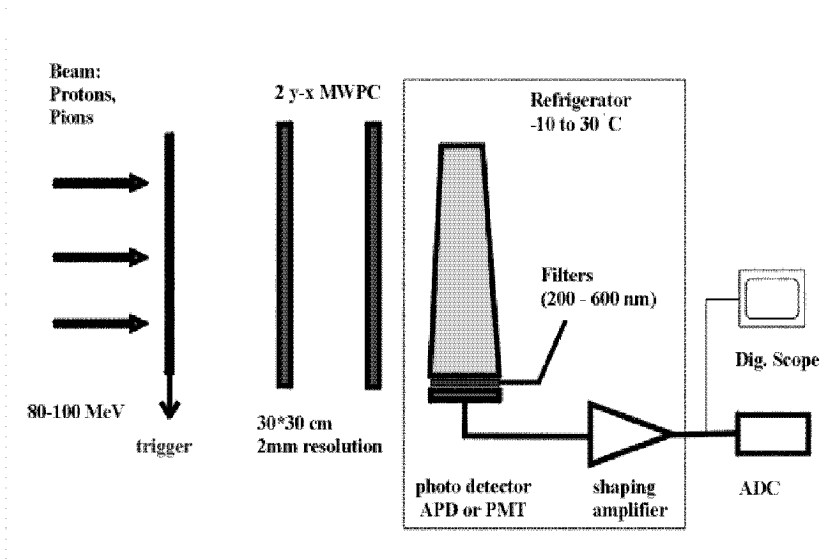


Figure A.3: Set-up for the crystal tests at the π M1 facility (PSI) [108].

and y coordinates. The crystal was installed in a refrigerator which guaranteed a temperature stability of $\pm 0.1^\circ\text{C}$. This was essential because the PbWO_4 light yield varies with $-2\%/^\circ\text{C}$ [108].

The passing particle produced scintillation light, which was converted into an electrical signal by a photo multiplier (PMT) or an Avalanche Photo Diode (APD). This signal was then amplified and its amplitude was measured by an Analog to Digital Converter (ADC). The ADC converted signals only during a time interval (gate) of 100 ns. This interval duration was chosen to match the time structure at the LHC.

Results

The knowledge of the absolute light yield and its longitudinal uniformity is very important. The light yield determination was done with protons and showed an increase with decreasing temperature. From this test, a temperature coefficient of $-2\%/^\circ\text{C}$ was calculated (in a good agreement with the literature). Moreover, the relative uniformity was not changing with temperature.

Between the PMT and the crystal, different optical filters were introduced to measure the spectral distribution of the emitted light. The obtained spectrum matched very well the quantum efficiency of the APD's.

The decay time of the scintillation light was determined using the same set-up with a PMT read-out. The tested crystals fulfilled the request for short decay time.

The radiation hardness of the PbWO_4 crystals was tested in two different ways.

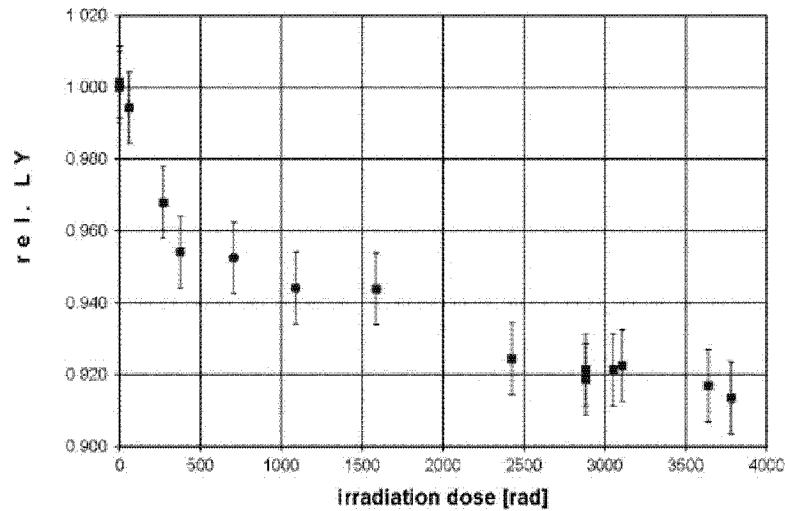


Figure A.4: Light Yield decrease of a PbWO_4 crystal after pion irradiation [108].

The first set-up was identical to figure A.3. A 10 cm long lucite light guide was inserted between the PMT and the crystal to avoid a radiation damage to the PMT. Instead of protons, the crystal was exposed to a 300 MeV pion beam. The pion rate could be adjusted to reach a dose rate of up to 100 mGy/h. The irradiation was stopped from time to time to determine the light yield and uniformity with protons. The total doses was of the order of 20-40 Gy. The light yield loss seemed to saturate after a dose of about 20 Gy at 93% of its original light yield (see figure A.4). To recover, the crystal was kept in the dark refrigerator at a temperature of about 19°C. Only a marginal recovery was detected.

To have a complementary method, an irradiation with neutral γ 's emitted by a Co^{60} source was used (with a dose rate of 25 mGy/h to 50 mGy/h). The set-up is sketched in figure A.5. The crystal was installed in an air conditioned room which houses the high intensity source. The produced light was again readout by a photo multiplier. The light signals induced by a second Co^{60} source with low intensity was used to follow the damage. A comparison of both measurements demonstrated the compatibility of the two methods. The irradiation of crystals with protons and low energy γ 's showed the same damage to the light yield [108].

Conclusions

The method used in 1997 to measure the PbWO_4 crystal parameters was fast and very precise. There was considerable progress in the reproducibility of the crystal quality. Moreover, the light output and radiation hardness reached a good performance level. Previous problems such as long tails to the pulses were eliminated.

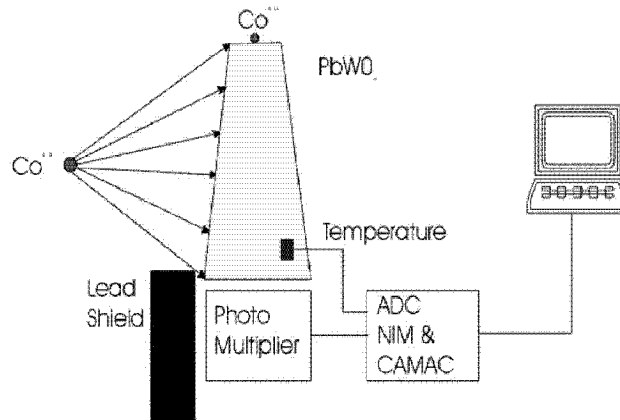


Figure A.5: Set-up for the crystal irradiation tests with Co^{60} sources [108].

A.1.2 R&D for the APD's

As proposed by PSI, the readout of the ECAL crystals (in the barrel part) will be performed by Avalanche Photo Diodes (APD). The first tests, in which the PbWO_4 scintillation light was collected by APD's, were performed in 1993 at CERN. These tests with a prototype provided by Hamamatsu Photonics were very promising, but the device showed also a large instability. The problems were the large dark current and the excessive capacitance. This resulted in a large statistical broadening of the amplified signal. The quantum efficiency of the APD's had to be also adjusted to the spectrum of the PbWO_4 emitted light. Furthermore, the production technique had to be optimized in order to guarantee a higher radiation hardness of the devices. Finally, the temperature dependence of the gain had to be fully understood.

To solve all these problematics, Hamamatsu Photonics, in tight collaboration with PSI, started a long R&D phase. In the meanwhile, also the APD's produced by another manufacturer (EG&G) were tested at PSI.

Results

In 1997, the radiation hardness of the APD's was tested in the OPTIS facility, where the devices could be bombarded with 64 MeV protons. With this proton beam, integrated doses equivalent to 10 years at LHC could be reached in about 2 hours!

By the irradiation, only a change in the dark current could be observed, while all the other parameters remained constant. These measurements allowed an estimate of the radiation-induced noise given by these devices to 70 MeV per crystal after 10 years of operation, which is within the CMS prescriptions.

Conclusions and schedule

One of the main progress on the APD's in 1997 was a halving of the price. The two candidate manufacturers (Hamamatsu Photonics and EG&G) were preparing the APD's for the final selection. The choice between the two vendors was made in favor of Hamamatsu in 1998. The definitive contract about the delivery of 130'000 APD's was subsequently signed. In the next 2 years, the 65'000 APD-pairs will be calibrated at CERN, while the radiation hardness will be further investigated at PSI.

A.2 CERN 1998-2000: ECAL tests

During the 1998-2000 period, I took part to some test beam runs performed at CERN. In particular, the goal of the experiments was the development of the CMS ECAL, including the PbWO_4 crystals, the APD's and the electronic readout.

One of the tasks of these test beam activities is to get an initial set of calibration constants to be used as starting point for the physics analysis. This is considered as an essential input for the possibility of discovering the Higgs decay into two photons. The clean situation of an electron test beam permits a measurement of these constants with high precision. Additional studies of overall linearity, calorimeter crack effects, temperature effects, e^\pm/γ and e^\pm/π^\pm separation on full size modules are unavoidable to understand the performances of the calorimeter. During this 3-years period, many test beam experiments have been performed on different set-ups. In the following, the tests in which I participated are described.

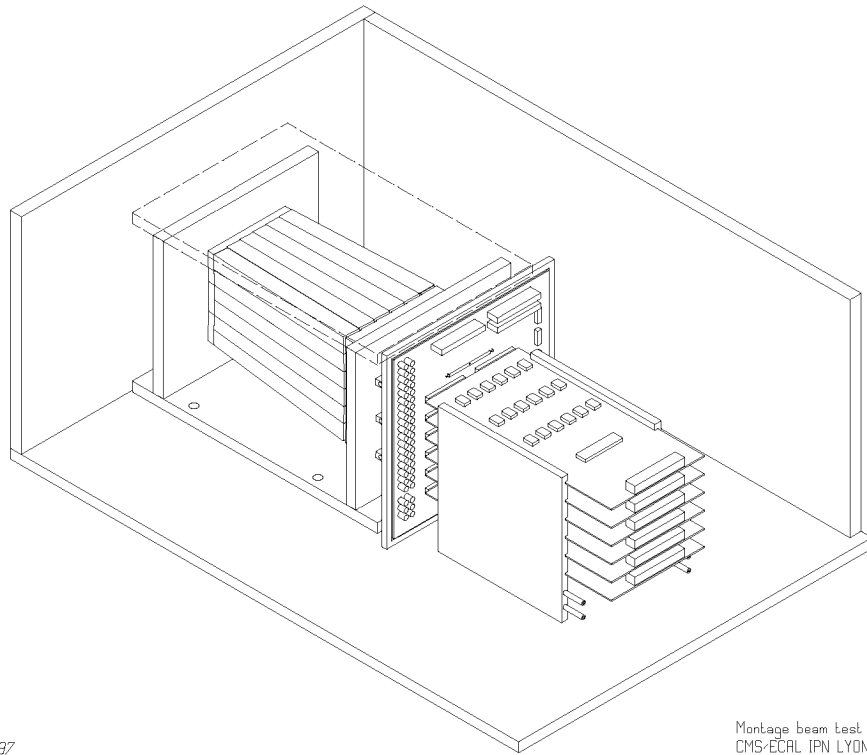
A.2.1 1998

In 1998, different prototypes were tested in high-energy beams (electrons and pions up to 150 GeV) at CERN. In particular, experiments on some parts of the readout, as well as radiation hardness tests, were performed.

A prototype (PROTO97, see figure A.6) corresponding to a 7×7 PbWO_4 crystal matrix, with each crystal read by one APD, was extensively used in the tests performed in the H4 facility in Prévessin. The electronic chain was composed of an optimized low-noise preamplifier and a charge ADC. A cooling system stabilized the temperature at 18°C with a precision of $\pm 0.1^\circ\text{C}$.

An energy resolution of approximately $\sigma/E = 0.6\%$ at 100 GeV was measured for a sum of 3×3 crystals within this configuration of 7×7 [109]. This result was reproduced many times with several combinations of 3×3 crystals.

This set-up showed also a lower Light Yield and a larger Nuclear Counter Effect in comparison with the expectations. After a checkup of the matrix, a bad contact between the crystal rear faces and the APD's was found. The gluing material was filled with bubbles and affected the light transmission. Nevertheless, these



lyocms,ecal97,skproto97

Montage beam test PROTO97
CMS-ECAL IPN LYON 15-04-97

Figure A.6: Schematic view of the PROTO97 prototype [110].

tests validated the concept of PbWO_4 crystals readout by APD's. In addition, other prototypes were built to test the electronic parts of the readout (preamplifier, compressor, etc).

A.2.2 1999

At the end of June 1999, a test of a preshower prototype, equipped with real-size detectors and LHC-style electronics, was tested in the H4 beam in front of a 5×5 matrix of endcap crystals. Data were taken with a variety of incident electron energies and three angles of incidence (to simulate different regions of the CMS endcaps). The prototype functioned well, with a very small startup period and operated successfully for the duration of the test (about 1 week). Good agreement was found between data and a GEANT based simulation.

In August 1999, data were taken in the H4 test beam facility with a matrix of 6×5 crystals with final electronics (PROTO99). The crystals dimensions were ones expected for the ECAL: $22 \times 22 \text{ mm}^2$ for the front face, $26 \times 26 \text{ mm}^2$ for the rear face and 230 mm in length. Each crystal was in an alveolar structure. The temperature of the 30 crystals was controlled via a cooling system and was fixed at $18^\circ\text{C} \pm 0.1^\circ\text{C}$. Capsules containing 2 APD's in parallel were glued on the

rear crystal faces. The August 1999 results demonstrated that a lead tungstate ECAL is able to exhibit very good performances in terms of noise and energy resolution. In these beam tests, excellent energy resolution was currently found with 11 crystals at the center of 9 crystal sub-matrices. Moreover, simulations and data showed approximately the same reconstructed energy. In contrast to the 1998 results, the crystal Light Yield and the Nuclear Counter Effect were consistent with the expectations.

For 1999, a final running period in H4 was performed between the 23. and the 30. September. The same prototype (PROTO99), but with sampling ADC readout was tested. This set-up represented the more or less full and final readout chain foreseen for the CMS ECAL (in its barrel part).

To summarize, the test beam periods of 1998/99 have lead to the following results:

- The long term stability of the monitoring system under test beam conditions was obtained at the 0.1% level.
- Laser light injection system was used successfully with ECAL prototypes to correct for radiation damages.
- Despite the effects of radiation damage due to the beam and recovery afterwards, the excellent resolution of the calorimeter could be maintained over long period of data taking.

A.2.3 2000

In year 2000, the R&D for the different ECAL sub-detectors involved several test beam periods. One of the main goals was the performance monitoring and the quality control for the PbWO_4 crystals produced by the two manufacturers (China and Russia). Another focus of interest was the development of the VFE electronics and of the upper level read out (the so-called Read-Out System Ecal (ROSE)).

I personally participated to the following sessions:

- The first test beam (in May) was dedicated to the monitoring and irradiation tests of some Russian crystals. For these runs, a new matrix was introduced (PROTO2000).
- The targets for the next period (June) were the following of calibration coefficients, the recovery of irradiated crystals and some other precise studies (position scan, high energy scan and so on). The same prototype (PROTO2000) with Russian crystals was used.
- Two very important test beam runs were undertook in August. The first one was dedicated to performance studies of new Chinese crystals (using the PROTO2000 matrix).

- In the second one, the old crystal matrix (PROTO99) served as test bed for the sampling ADC, with the new version of the VFE electronics and of the board based digital readout (ROSE).
- Finally, during October-November 2000, additional test runs on the Chinese crystals were performed with the PROTO2000 matrix. In particular, further informations about the absolute calibration and the recovery after irradiation were needed by the producer as a feed-back for future optimizations.

Bibliography

- [1] S.L. Glashow, Nucl. Phys. **22**, 579 (1961);
S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967);
A. Salam, Elementary Particle Theory, Almquist and Wiksells (1968);
H.D. Politzer, Phys. Rev. Lett. **30**, 1346 (1973);
D.J. Gross and F.E. Waltzed, Phys. Rev. Lett. **30**, 1343 (1973).
- [2] H. Fritzsche and M. Gell-Mann, Proc. XVI Int. Conf. on High Energy Physics,
eds. J.D. Jackson and A. Roberts (Fermilab 1972).
- [3] P.W. Higgs, Phys. Rev. Lett. **12**, 132 (1964); Phys. Rev. **145**, 1156 (1966);
F. Englert and R.Brout, Phys. Rev. Lett. **13**, 321 (1964);
G.S. Guralnik, C.R. Hagen and T.W. Kibble, Phys. Rev. Lett. **13**, 585
(1964).
- [4] L. Maiani, G. Parisi and R. Petronzio, Nucl. Phys. **B136**, 115 (1979);
N. Cabibbo, L. Maiani, G. Parisi and R. Petronzio, Nucl. Phys. **B158**, 295
(1979);
R. Dashen and H. Neuberger, Phys. Rev. Lett. **50**, 1897 (1983);
D.J.E. Callaway, Nucl. Phys. **B233**, 189 (1984);
M.A. Beg, C. Panagiatakopolus and A. Sirlin, Phys. Rev. Lett. **52**, 883
(1984);
M. Lindener, Z. Phys. **C31**, 295 (1986).
- [5] T. Hambye and K. Riesselmann, Phys. Rev. **D55**, 7255 (1997).
- [6] The LEP collaborations, CERN-EP/2001-021.
- [7] For more informations see the experiments web-sites:
<http://alephwww.cern.ch/>
<http://delphi.web.cern.ch/Delphi/>
<http://l3www.cern.ch/>
<http://opal.web.cern.ch/Opal/> .
- [8] ALEPH, DELPHI, L3 and OPAL Collaborations, CERN-EP/2001-055, July
2001

- [9] The LEP Electroweak Working Group, public page, <http://lepewwg.web.cern.ch/LEPEWWG/> (updated July 20. 2001).
- [10] C. Quigg, B.W. Lee and H. Thacker, Phys. Rev. **D16**, 1519 (1977);
M. Veltman, Acta Phys. Polon. **B8**, 475 (1977).
- [11] Design Study of the Large Hadron Collider (LHC), The LHC Study Group, CERN (1991).
- [12] M. Dittmar, F. Pauss and D. Zürcher, Phys. Rev. **D56**, 7284 (1997).
- [13] F. Behner, M. Dittmar, F. Pauss and D. Zürcher, contributed paper 418, EPS 97 Jerusalem (1997).
- [14] CMS Collaboration, G.L. Bayatian *et al.*, Technical Proposal, CERN Report No. CERN/LHCC 94-38, LHCC/P1, 15. Dec. 1994.
- [15] Application Software Group, CERN Computing and Networks Division, CERN Program Library Long Writeup W5013.
- [16] D. Zürcher, CMS Note 1999/040.
- [17] J.M. Campbell and R.K. Ellis, Phys. Rev. **D60**, 113006 (1999).
- [18] L. Dixon, Z. Kunszt and A. Signer, Phys. Rev. **D60**, 114037 (1999), [hep-ph/9907305].
- [19] T. Sjöstrand, JETSET 7.4, Comput. Phys. Commun. **82**, 74 (1994); *ibid.* Comput. Phys. Commun. **82**, 74 (1994); *ibid.* "Recent progress in PYTHIA", hep-ph/0001032.
- [20] D. Rainwater and D. Zeppenfeld, Phys. Rev. **D62**, 13009 (2000).
- [21] R. N. Cahn *et al.*, Phys. Rev. **D35**, 1626 (1987);
V. Barger, T. Han and R.J.N. Phillips, Phys. Rev. **D37**, 2005 (1988);
R. Kleiss and W. J. Stirling, Phys. Lett. **B200**, 193 (1988);
D. Froideveaux, in [111], Vol. II, p. 444;
M. H. Seymour, *ibid.*, p. 557;
U. Baur and E. W. N. Glover, Nucl. Phys. **B347**, 12 (1990); Phys. Lett. **B252**, 683 (1990).
- [22] S. Abdullin and N. Stepanov, CMS TN/94-178 and CMS TN/94-179 (1994).
- [23] See for example:
"ATLAS detector and physics performance TDR", CERN/LHCC 99-15, 716-729 (1999).

- [24] A.S. Nicollerat, "*Prospects of Measuring $qq \rightarrow qqH$ and $gg \rightarrow H$ Cross Sections at the LHC with CMS*" or "the Higgs files", Diploma Thesis (2001), (unpublished).
- [25] D. Rainwater and D. Zeppenfeld, hep-ph/9906218 (1999).
- [26] N. Kauer, T. Plehn, D. Rainwater and D. Zeppenfeld, hep-ph/0012351 (2000).
- [27] S. L. Glashow, Nucl. Phys. **22**, 579 (1961).
- [28] S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967),
A. Salam, "*Elementary Particle Theory*", Almquist and Wiksells (1968).
- [29] J. Iliopoulos, talk given at the London Conference, June 1974.
- [30] M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 652 (1973);
N. Cabibbo, Phys. Rev. Lett. **10**, 531 (1963).
- [31] P.W. Higgs, Phys. Lett. **12**, 132 (1964).
- [32] P.W. Higgs, Phys. Rev. **145**, 1156 (1966).
- [33] T. Hambye and K. Riesselmann, Phys. Rev. **D55**, 7255 (1997), [hep-ph/9708416].
- [34] See for example Z. Kunszt and F. Zwirner, Nucl. Phys. **B385**, 3 (1992).
- [35] K. Inoue, A. Kakuto, H. Komatsu and S. Takeshita, Prog. Theor. Phys. **68**, 927 (1982); *ibid.* **70**, 330 (1983); *ibid.* **71**, 413 (1984);
H.E. Haber and G.L. Kane, Phys. Rep. **C117**, 75 (1985).
- [36] R. Assmann *et al.*, "*SLC The End Game*", Proceedings of EPAC 2000, Vienna, Austria.
- [37] See for example the Run II Handbook on the web:
<http://www-bd.fnal.gov/lug/runIIhandbook/RunIIindex.html>.
- [38] B. Pietrzyk, talk presented at ICHEP2000, Osaka (2000).
- [39] H. Georgi, S.L. Glashow, M.E. Maccahek and D.V. Nanopoulos, Phys. Rev. Lett. **40**, 692 (1978).
- [40] R.N. Cahn and S. Dawson, Phys. Lett. **B136**, 196 (1984).
- [41] S.L. Glashow, D.V. Nanopoulos and A. Yildiz, Phys. Rev. **D18**, 1724 (1978);
Z. Kunszt, Z. Trocsanyi and W.J. Stirling, Phys. Lett. **B271**, 247 (1991).

- [42] Z. Kunszt, Nucl. Phys. **B247**, 339 (1984);
 J.F. Gunion, Phys. Lett. **B253**, 269 (1991);
 W.J. Marciano and F.E. Paige, Phys. Rev. Lett. **66**, 2433 (1991);
 J.F. Gunion, H.E. Haber, F.E. Paige, W.K. Tung and S.S.D. Willenbrock,
 Nucl. Phys. **B294**, 621 (1987);
 D.A. Dicus and S.S.D. Willenbrock, Phys. Rev. **D39**, 751 (1989).
- [43] M. Spira, Fortsch. Phys. **46**, 203 (1998), [hep-ph/9705337].
- [44] Z. Kunszt, S. Moretti and W.J. Stirling, hep-ph/9611397, ETH-TH-96/48 (1996).
- [45] M. Spira and M. Zerwas, hep-ph/9803257 (1998).
- [46] M. Dittmar, Pranam journal of physics **53**, No.6, 1 (1999).
- [47] M. Dittmar, H. Dreiner, Phys. Rev. **D55**, 167 (1997), [hep-ph/9608317].
- [48] ATLAS Collaboration, W. W. Armstrong *et al.*, Technical Proposal, CERN Report No. CERN/LHCC 94-43, LHCC/P2, December 15. 1994.
- [49] LHC-b technical proposal, LHC-b Collaboration, CERN/LHCC 98-4, February 1998.
- [50] ALICE Collaboration, "*ALICE technical proposal for a large ion collider experiment at the CERN LHC*", CERN/LHCC/95 71, LHCC/P3, December 15. 1995.
- [51] The CMS Magnet Project, Technical Design Report, CERN/LHCC 97-10, May 1997.
- [52] The CMS MUON Technical Design Report, CERN/LHCC 97-32, December 1997.
- [53] The CMS HCAL Technical Design Report, CERN/LHCC 97-31, June 1997.
- [54] The CMS ECAL Technical Design Report, CERN/LHCC 97-33, December 1997.
- [55] K. Lassila-Perini, Reconstruction of Higgs $\rightarrow \gamma\gamma$ in CMS, CMS CR-1997/006 (1997).
- [56] The CMS Tracker Project, Technical Design Report, CERN/LHCC 98-6, April 1998.
- [57] The CMS TriDAS Project, Technical Design Report, Volume 1: The Trigger Systems, CERN/LHCC 2000 - 38, December 2000.

- [58] T. Nakada, talk presented at the 4th International Workshop on B physics & CP violation, Ise-Shima, Japan, February 2001.
- [59] Particle Data Group, R.M. Barnett *et al.*, Phys. Rev. **D54**, 1 (1996).
- [60] See for example reference [48] p. 177 and reference [14] p. 211.
- [61] G. Altarelli and G. Parisi, Nucl. Phys. **126**, 297 (1977);
V.N. Gribov and L.N. Lipatov, Sov. J. Nucl. Phys. **15**, 438 and 675 (1972);
Yu. L. Dokshitzer, Sov. Phys. JETP **46**, 641 (1977).
- [62] A.D. Martin, R.G. Roberts, W.J. Stirling, Phys. Lett. **B306**, 147 (1993).
- [63] CTEQ Collaboration, Phys. Lett. **B304**, 159 (1993).
- [64] M. Glück, E. Reya, A. Vogt, Ph.D. Thesis, 1994.
- [65] For more details see for example: P.N. Harriman *et al.*, Phys. Rev. **D42**, 798 (1990) and A.D. Martin *et al.*, Phys. Rev. **D50**, 6734 (1994).
- [66] S. van der Meer, "*Calibration of the effective beam height in the ISR*", Internal Report 68-31, CERN/ISR-PO, 1968.
- [67] TOTEM Collaboration, M. Bozzo *et al.*, CERN Report No. CERN/LHCC 93-47 (1993) (unpublished) and CERN/LHCC 94-39 (1994) (unpublished).
- [68] M. Bozzo *et al.* (TOTEM Collaboration), Internal report, CERN/LHCC, 1993.
- [69] V. M. Budnev *et al.*, Phys. Lett. **B39**, 526 (1972) and Nucl. Phys. **B63**, 519 (1973).
- [70] A.D. Martin, R.G. Roberts, W.J. Stirling and R.S. Thorne, contribution to the "*Proceedings of the Workshop on Standard Model Physics (and more) at the LHC*", 10, CERN 2000-004.
- [71] See for example: SDC Collaboration, E. L. Berger *et al.*, "*Technical Design of a Detector*", Report No. SDC-92-201, SSCL-SR-1215 (unpublished) section 3.7.4 and reference [48] p. 177.
- [72] R. K. Ellis, W. J. Stirling and B. R. Webber, "*QCD and Collider Physics*", Cambridge University Press, ISBN 0 521 58189 3, 1996.
- [73] M. Dittmar, talk given at the LHC Workshop "*Theory of LHC Processes*", CERN, February 1998.
- [74] M. Dittmar and K. Mazumdar, CMS Note 2001/002.

- [75] For a recent theoretical analysis and further references therein see for example: R.K. Ellis *et al.*, hep-ph/9704239 FERMILAB-PUB-97/082-T and SHEP-96/37 (unpublished).
- [76] For the simulations we have used the PYTHIA 6.1 frame [19] and the structure functions MRS(A), MRS(H), MRS(R1), CTEQ4M, CTEQ4L, CTEQ5L, CTEQ5M1 and the GRV 94 HO as implemented within the PDFLIB. For details and further references see H. Plochow-Besch, Comput. Phys. Commun. **75**, 396 (1993); PDFLIB (version 8.04) W5051 CERN Computer Library.
- [77] See for example reference [48] page 50 and reference [14] page 172.
- [78] D. Zürcher, "*A Method to Measure Parton Luminosities at the LHC: a Tool to determine Higgs Backgrounds accurately*", Diploma Thesis (1997), (unpublished).
- [79] M. Dittmar, Phys. Rev. **D55**, 161 (1997).
- [80] NA51 Collaboration, A. Baldit *et al.*, Phys. Lett. **B332**, 244 (1994).
- [81] See for example: P. Aurenche *et al.*, in [111], Vol. II.
- [82] "*The Electro-magnetic Calorimeter Project, Technical Design Report*", CERN/LHCC 97-33, CMS TDR 4 (1997).
- [83] See chapter 2 of reference [82].
- [84] F. Cavallari, "*Progress on Avalanche Photo-diodes as photon detectors for PbWO₄ crystals in the CMS experiment*", CMS Conference Report 1997/010.
- [85] C. Seez, "*Test Beam Irradiation Catalogue for 1997 and 1998*", CMS Internal Note 1998/026.
- [86] M. F. Gautheron, "*Etudes de cristaux scintillants de tungstate de plomb et de prototypes de calorimètre électromagnétique pour le projet CMS au LHC*", These de doctorat N° 13097, Université Claude Bernard, Lyon (1997).
- [87] C. Seez, private communication.
- [88] C. Tully, private communication.
- [89] C.J. Purves, "*New values for electron / charged-pion Discrimination from the '95 CMS Electro-magnetic Calorimeter Prototype*", CMS Technical Note 1996/006.
- [90] PYTHIA subprocesses codes 11, 12, 13, 28, 53 and 68.

- [91] PYTHIA subprocesses codes 2,16, 20, 31 and 36 (for W) and 25 (for WW).
- [92] J. Ohnemus, Phys. Rev. **D44**, 1403 (1991).
- [93] S. Frixione, Nucl. Phys. **B410**, 280 (1993).
- [94] J.F. Gunion and Z. Kunszt, Phys. Rev. **D33**, 665 (1986).
- [95] L. Dixon, Z. Kunszt and A. Signer, Nucl. Phys. **B531**, 3 (1998), [hep-ph/9803250].
- [96] J.M. Campbell and R.K. Ellis, preprint hep-ph/9905386.
- [97] U. Baur, T. Han and J. Ohnemus, Phys. Rev. **D53**, 1098 (1996) [hep-ph/9507336].
- [98] U. Baur, T. Han and J. Ohnemus, Phys. Rev. **D51**, 3381 (1995) [hep-ph/9410266].
- [99] H.L. Lai *et al.* (CTEQ Collaboration), Phys. Rev. **D55**, 1280 (1997), [hep-ph/9606399].
- [100] S. Frixione, Nucl. Phys. **B410**, 280 (1993).
- [101] U. Baur and E.W.N. Glover, in [111], Vol.II, p. 570.
- [102] V. Barger, T. Han, R.J.N. Phillips, Phys. Lett. **B206**, 339 (1998).
- [103] E. Richter-Was *et al.*, ATLAS Internal Note ATL-PHYS-95-048 (1995).
- [104] V. Barger, R.J.N. Phillips and D. Zeppenfeld, Phys. Lett. **B346**, 106 (1995).
- [105] M. Spira and M. Dittmar, CMS Note 1997/080.
- [106] A. Kupčo, "*MC event generators and their predictions for minimum bias events at LHC energies*", ATLAS note ATL-PHYS-99-019 (1999).
- [107] M. Dittmar and H. Dreiner, CMS Note 1997/083.
- [108] K. Deiters, "*Methods to measure parameters of PbWO₄ Crystals at PSI*", Proceedings of the SINT97 Conference in Shanghai, China.
- [109] F. Martin, CMS Conference Report 1999/003.
- [110] Courtesy of P. Ingenito (ETH Zürich).
- [111] Proceedings of the Large Hadron Collider Workshop, Aachen 1990, edited by G. Jarlskog and D. Rein, CERN 90-10/ECFA 90-133.

Acknowledgments

It's hard to believe, but... "That's all folks!"

The citation quoted at the beginning of this work finally assumes its full meaning: at some point one says "that's enough!" and suddenly, almost magically, the work becomes complete, even if only few days before the goal seemed far away. My first thought was: "I did it!". However, this egocentric vision does not do justice to all the people who, each with his different contribution in terms of means, knowledge or friendship, contributed to the success of this work.

First of all, I owe my sincere gratitude to Professor Felicitas Paus, Head of the Institute for Particle Physics, for support, excellent conditions and working environment within our group. She also gave me the opportunity to present my work at international conferences, workshops and summer schools. From this point of view, I am also especially indebted to Kurt Gabathuler, who supported me on many occasions when I wanted to participate in events outside CERN.

I would like also to thank Prof. Zoltan Kunszt, who kindly accepted to act as co-examiner in my defense.

I wish to express my warmest thanks to Michael Dittmar. He supported me through many valuable discussions full of wit and invested a lot of time to share his knowledge with me. His critical comments were always inspiring and substantially contributed to this thesis. I cannot omit also the rest of the Dittmar family: Françoise, Lea and Marie, for the friendship demonstrated throughout all these years.

I am indebted to Dieter Renker, who succeeded in importing some "Münich atmosphere" into the Paul Scherrer Institut, making my stay in Villigen more enjoyable. Furthermore I owe a debt of gratitude to him also for having accepted to act as co-examiner.

A particular thank goes to André Holzner. He was always there to help whenever problems came up during my work. His incomparable competence regarding computing has been crucial in many difficult situations.

Special thanks go also to Gabriele Kogler, the secretary of the ETH CERN group, for the many times she kindly assisted me in the complicated paths of the bureaucracy and for all the interesting discussions.

I would like to particularly express my gratitude to Artur Barczyk: it was particularly pleasant to share my office with him. Moreover, we spent great times

together outside CERN and, last but not least, he perfectly fulfilled the task of bringing the first copy of this work to the secretariat in Zürich. This last exploit would not have been possible without the precious help of his wife: thank you Marina!

Many other friends enriched my social life inside and outside CERN. Special thanks go to Vuko Brigljević, Frank Behner, Konrad Deiters, Renzo Ramelli, Stève Udriot, Anne-Sylvie Nicollerat, Radek Ofierzynski, Stefano Villa, Luca Malgeri, Sonia Natale and Mimmo Della Volpe.

I want also to express my gratitude towards the members of the L3 and Fairboots football teams, who helped me to follow the motto “mens sana in corpore sano” (at least in the intentions...).

My parents Giorgio and Inge Zürcher deserve my deepest gratitude for their unselfish support and encouragement throughout my life.

Finally, I wish to express my deep love to my girlfriend Cristina (“Titti”), who supported me unconditionally all over these years and especially in the last crucial weeks. Now it’s your turn... ;-)

Curriculum Vitæ

PERSONAL DATA:

Date and place of birth: January 28. 1971, Lugano (Switzerland)

Nationality: Swiss

EDUCATION:

1977-1982 Primary school in Viganello, Switzerland

1982-1986 Secondary school in Viganello, Switzerland

1986-1990 High school in Lugano, Switzerland

1990-1997 Study of physics at the Swiss Federal Institute of Technology (ETHZ), Zürich.

Diploma in physics (1997): "A Method to Measure Parton Luminosities at the LHC: a Tool to determine Higgs Backgrounds accurately",

performed under the supervision of Prof. Dr. Felicitas Pauss and Dr. Michael Dittmar at the Institute for Particle Physics (IPP) of ETH Zürich, March 17.

1997-2001 PhD thesis (this work) in particle physics.

Work started (Juli 1997) at the Paul Scherrer Institut under the supervision of Dr. Dieter Renker and continued (1998-2001) at the European Centre for Nuclear Research (CERN) within the group of the Institute for Particle Physics of the ETH Zürich, under the supervision of Prof. Dr. Felicitas Pauss and Dr. Michael Dittmar.

Member of the CMS collaboration (which develops one of the multi purpose detectors for the future Large Hadron Collider). Working at the simulations of Vector Boson pair production and, from the hardware side, testing crystals and APD's for the CMS ECAL.