Search for MSSM Higgs Bosons in Tau Final States with the DØ Detector

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Engineering and Physical Sciences

2010

Wan-Ching Yang

Particle Physics Group
School of Physics and Astronomy

Contents

Abstract

Name of the University: Tha University of Manchester

Candidate Name: Wan-Ching Yang

Degree Title:
Doctor of Philosophy in the Faculty of Engineering and Physical Sciences

Thesis Title:
Search for MSSM Higgs Bosons in Tau Final States with the DØ Detector

Date: 07/09/2010

The cross-section times branching ratio of the Higgs boson decaying to $\tau^+\tau^-$ final state in the Standard Model (SM) is too small to play any role in the SM Higgs boson searches. This, however, is different in the Minimal Supersymmetric Standard Model (MSSM), which predicts two Higgs doublets leading to five Higgs bosons: a pair of charged Higgs boson ($H^{\pm}$); two neutral CP-even Higgs bosons (h,H) and a CP-odd Higgs boson (A).

A search for the production of neutral Higgs bosons decaying into $\tau^+\tau^-$ final states in $p\bar{p}$ collisions at a centre-of-mass energy of $\sqrt{s} = 1.96$ TeV is presented in this thesis. One of the two τ leptons is required to decay into a muon while the other decays hadronically. The integrated luminosity is $\mathcal{L} = 1.0$-5.36 fb^{-1}, collected by the DØ experiment at the Fermilab Tevatron Collider from 2002 to 2009 in the Run II.

Declaration

I declare that no portion of this work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Signed: Wan-Ching Yang, 7/9/2010

Copyright

The author of this thesis (including any appendices and/or schedules to this thesis) owns any copyright in it (the "Copyright")[1] and s/he has given The University of Manchester the right to use such Copyright for any administrative, promotional, educational and/or teaching purposes.

Copies of this thesis, either in full or in extracts, may be made only in accordance with the regulations of the John Rylands University Library of Manchester. Details of these regulations may be obtained from the Librarian. This page must form part of any such copies made.

The ownership of any patents, designs, trade marks and any and all other intellectual property rights except for the Copyright (the "Intellectual Property Rights") and any reproductions of copyright works, for example graphs and tables ("Reproductions"), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property Rights and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property Rights and/or Reproductions.

Further information on the conditions under which disclosure, publication and exploitation of this thesis, the Copyright and any Intellectual Property Rights and/or Reproductions described in it may take place is available from the Head of School of Physics and Astronomy.

---

[1] This excludes material already printed in academic journals, for which the copyright belongs to said journal and publisher. Pages for which the author does not own the copyright are numbered differently from the rest of the thesis.

Acknowledgement

I would like to first thank my supervisor, Stefan Söldner-Rembold, for providing advice and encouragement throughout my PhD for giving great assistance whenever I needed it, just like a superman. Many thanks to Mark Owen and Abid Patwa for their tireless ability to answer my questions, this thesis would not have been possible without their contribution. I feel very lucky to be one of the members of the DØ Manchester group, in particular thanks to Maiko Takahashi who has always given me very strong support in life. Thanks to all the members of the HEP group in Manchester, in particular to Fred Loebinger and Terry Wyatt who are always around for advice and guidance, and to Sabah Salih who always solves computer issues immediately. Special thanks to Ching Hung Lai who printed copies of this thesis and helped with the submission of the thesis because I was not physically in the UK. The results presented in this thesis are only achievable in collaboration with the Higgs and τ identification groups at DØ. Thanks to all the people I met at Fermilab who made my life in Chicago wonderful, particularly, I would like to thank Phil and Krisztian who assisted me a lot during my first trip to Fermilab. Thanks to Abid, Christian, Kostas, Maiko, Stefan, and Yvonne who gave me so much valuable advice for my thesis. Studying and living in the UK and US for the past four years would not have been possible without support from my family as well as my friend, Chun-Che. Finally, a huge thank-you to Pan who always makes me smile, gives me any kind of help even though he is sometimes thousands of miles away from me.

Chapter 1

Introduction

A search for neutral Higgs bosons in $\tau^+\tau^-$ final states, produced in the collisions of protons (p) and antiprotons $(\bar{p})$ at the Fermilab Tevatron, is presented in this thesis. For this search one of the two τ leptons is required to decay into a muon (τ_μ) or an electron (τ_e) while the other decays hadronically (τ_h). The data have been collected from 2002 to 2009 by the DØ detector at a $p\bar{p}$ centre-of-mass energy of $\sqrt{s} = 1.96$ TeV. They correspond to up to 5.36 fb^{-1} of integrated luminosity.

The first part of the thesis includes basic introductions about the theoretical background, the Tevatron and the DØ detector, and the search tools, which can be found in Chapters 2, 3, and 4. More details about analysis techniques, including studies of the trigger and the lepton identification efficiencies, are described in Chapters 5 and 6. The main background sources for the Higgs boson search, W+jets, $Z \rightarrow e^+e^-$, $Z \rightarrow \mu^+\mu^-$, and multijet events, are discussed in Chapter 7. Signal selections that are designed based on topological or kinematic properties of the different background sources and of possible signal candidates are studied in Chapter 8.

In Chapter 9 the results are used to set limits at the 95% confidence level (CL) on the product of the Higgs production cross section and branching ratio, $\sigma(p\bar{p} \rightarrow \phi) \times \mathrm{BR}(\phi \rightarrow \tau\tau)$. The cross section limits are interpreted in the Minimal Supersymmetric Standard Model (MSSM) for two MSSM benchmark scenarios, M_h^{max} and *no-mixing*, using the FEYNHIGGS program. The results are presented as exclusions in terms of two MSSM parameters, the mass of the CP-odd Higgs boson, M_A, and $\tan\beta$.

Chapter 2

Theoretical Background

This chapter describes the theoretical framework for the search performed in this thesis. It begins with a brief overview of the Standard Model (SM) followed by a brief description of the reasons why extending the theory beyond the SM is necessary. One of the most popular theories that resolve pending issues left by the SM is a supersymmetric extension of the Standard Model, which will be introduced in Sect. 2.3. Further details concerning Supersymmetry (SUSY) can be found in many references, for example, see [1]. In this thesis, the convention where speed of light, $c = 1$, and Plank's constant, $\hbar = 1$ are chosen for all units used.

2.1 The Standard Model

The Standard Model of particle physics is a widely accepted and tested theory that provides explanations to many experimental results and predicts a variety of new phenomena. A much more comprehensive description for the SM can be found in Ref. [2], and only those details relevant to the subject of this thesis will be given here. The model was developed in 20th century, and was confirmed by the discovery of W and Z bosons in 1983. The most recent discoveries such as the top quark (1995) [3] [4] gave even more credence to the SM. The SM is not only a well-tested theory in particle physics, but also explains many of the phenomena observed in other physics fields such as astrophysics and cosmology.

The fundamental interactions between quantum field Φ are explained by three local gauge invariant quantum field theories (QFT), *Quantum Electrodynam-*

ics (QED), Quantum Flavordynamics (QFD) and *Quantum Chromodynamics* (QCD). In particle physics, Gravitation does in general not play a role so it is not part of the SM.

The Lagrangian of the theory is required to be invariant under a continuous group of local transformation, $\Phi' \rightarrow \Phi e^{i\phi(\underline{x})}$. Here, the term "local" indicates that ϕ depends on the space-time position, $\underline{x}$, of the field. In order to achieve this invariance and keep the physical observables unchanged under the local group transformation, corresponding vector fields, called *gauge fields*, are included in the Lagrangian. This subsequently leads to *gauge invariance*, and the quanta of the gauge fields are called *gauge bosons*. The simplest abelian gauge theory is QED with the symmetry group $U(1)_{em}$. A symmetry group is called "abelian" if it is commutative. The symmetry of the $U(1)_{em}$ group is imposed on the Lagrangian for the electrodynamics transformations and leaves the physical action unchanged. The electromagnetic and weak forces are unified with the combined gauge group $SU_L(2) \otimes U_Y(1)$ [2] which breaks down to $U(1)_{em}$ symmetry after Spontaneous Symmetry Breaking (SSB). However, the Standard Model is a non-abelian gauge theory described by the symmetry group $SU_L(2) \otimes U_Y(1) \otimes SU(3_c)$; $SU(3_c)$ is the gauge group of QCD, the theory of strong interactions which pertain to interactions due to colour charges of quarks. In particular, $SU_L(2) \otimes U_Y(1) \otimes SU(3_c)$ gauge symmetry leads to a total of twelve gauge bosons: W^+, W^-, Z^0, eight gluons and one massless photon.

According to the Standard Model, the universe is composed of twelve point-like matter particles each with spin quantum number $\frac{1}{2}$. Each of these particle contain corresponding anti-particles with the same mass and spin quantum number but with opposite electric charge. The kinematics and interactions of matter particles are governed by exchanging gauge bosons with integral values for spin. Further details regarding these elementary particles and interactions can be found in Sect. 2.1.1 and 2.1.2.

2.1.1 Matter in the Standard Model

The twelve spin $\frac{1}{2}$ fermions within the SM are classified by the type of interactions and therefore are categorised into two sets: leptons and quarks. The six known flavours of quarks and six flavour of leptons occur in pairs called generations, with each generation containing a pair of leptons, one with zero electric charge and

11

	Quarks			Leptons		
Generation	Particle	Mass (MeV)	Charge (e)	Particle	Mass (MeV)	Charge (e)
First	up (u)	$1.5 - 3.3$	$\frac{2}{3}$	electron (e)	0.51	-1
	down (d)	$3.5 - 6.0$	$-\frac{1}{3}$	e neutrino (ν_e)	$< 2.2 \times 10^{-6}$	0
Second	strange (s)	105^{+25}_{-35}	$-\frac{1}{3}$	muon (μ)	105.66	-1
	charm (c)	1270^{+70}_{-110}	$\frac{2}{3}$	μ neutrino (ν_μ)	< 0.17	0
Third	bottom (b)	4200^{+170}_{-70}	$-\frac{1}{3}$	tau (τ)	1776.84 ± 0.17	-1
	top (t)	$173.3 \pm 1.1 \times 10^3$	$\frac{2}{3}$	τ neutrino (ν_τ)	< 15.5	0

Table 2.1: The matter fermions of the Standard Model [5].

the other with charge $Q = -1$, and a pair of quarks, one with charge $Q = \frac{2}{3}$ and one with charge $Q = -\frac{1}{3}$. Each pair corresponds to a weak isospin doublet. The first generation includes the electrons, electron neutrinos, up quarks and down quarks, while the second compromises muons, muon neutrinos, charm quarks and strange quarks. Particles within the third generation are τ leptons, τ neutrinos, bottom and top quarks, and are very different from the first two generations by masses. For example, the measured world-average mass of a top quark is 173.3 GeV [7], which is four orders of magnitude larger than the mass of an up quark with 3.3 MeV. It is even heavier than some of the baryons like the proton with a mass of 938 MeV. The huge difference in masses between generations is not yet explained by the SM, and therefore, this results in one of the biggest open questions in elementary particle physics. Table 2.1 lists the basic properties of these elementary particles including masses and charges which are referenced from [5].

Quarks carry colour charges and interact via the strong force. They are constituents of colour neutral hadrons, including mesons, made of one quark and one anti-quark, and baryons, composed of three quarks. The most basic baryons existing in nature are protons and neutrons having the smallest mass. These are made of a combination of the two lightest quarks: up and down quarks. However, heavier quarks (c, s, t, b) can only be created in high-energy collisions. Table 2.2 lists the additive quantum numbers of quarks. The baryon number B has the values of 1/3 for quarks and -1/3 for anti-quarks. The *flavours*, I_z, S, C, B' or T are related to the types of quarks with the same sign of the electric charge a quark carries. Anti-quarks have the same value of flavour quantum numbers but with opposite sign. These additive quantum numbers fully describe the electric

charge, Q, of quarks through the Gell-Mann-Nishijima formula [5]:

$$Q = I_z + \frac{B + S + C + B' + T}{2}. \tag{2.1}$$

Quark Property	d	u	s	c	b	t
Q - electric charge	$-\frac{1}{3}$	$\frac{2}{3}$	$-\frac{1}{3}$	$\frac{2}{3}$	$-\frac{1}{3}$	$\frac{2}{3}$
I_z - z-component of isospin	$-\frac{1}{2}$	$\frac{1}{2}$	0	0	0	0
S - strangeness	0	0	-1	0	0	0
C - charm	0	0	0	+1	0	0
B' - bottomness	0	0	0	0	-1	0
T - topness	0	0	0	0	0	1

Table 2.2: Additive quantum numbers of the quarks. [5]

On the other hand, leptons have no colour charge and do not participate in strong interactions. Therefore, the three leptons each in the specific generation, electrons (e^-), muons (μ^-) and taus (τ^-) interact electromagnetically and weakly. Since the τ leptons are of particular importance to this thesis, more details about τ properties and the reconstruction methods of τ leptons at DØ will be described further in Sect. 4.1.6. Neutrinos (ν_e, ν_μ and ν_τ) do not carry electric charge and interact only through the weak nuclear force making the detection of them difficult in typical collider detectors. Since they are undetected, neutrinos are only visible as missing energy in the transverse plane.

2.1.2 Force Mediating Particles in the Standard Model

Within the SM, fundamental particles interact with each other by exchanging force carriers with the values of their spin quantum number equal to one. Table 2.3 lists the three fundamental forces explained by the SM.

Interaction	Describing Theory	Mediator(s)	Relative strength	Acting range (m)
Weak Force	Electroweak Theory	W^-, W^+, Z	10^{25}	10^{-18}
Electromagnetic Force	QED	photons (γ)	10^{36}	∞
Strong Force	QCD	gluons	10^{38}	10^{-15}

Table 2.3: The forces and the corresponding mediators [5].

The electromagnetic interaction between electrically charged particles might be the most familiar force to everyday life rather than the gravitation. It is

mediated by photons (γ) and fully described by QED. The weak force is responsible for the interactions between particles of different flavours. The mediators of the weak force are massive W^-, W^+ and Z bosons. The leptons can emit or absorb a Z boson via the neutral current process of the weak interaction. The charged current process of the weak interactions allows a quark of one flavour to transform into a quark of another flavour, or a lepton to convert into a corresponding neutrino via the exchange of the $W^\pm$ bosons. The Z and W bosons have a mass of 91.1876 ± 0.0021 GeV and 80.398 ± 0.025 GeV [5], respectively. The electromagnetic interaction and weak force are merged into only one force, the electroweak force, at the energy above the unification energy, which is on the order of 100 GeV. Since the Higgs mechanism studied in this thesis is related to electroweak symmetry breaking (EWSB), more details about the electroweak interaction can be found in Sect. 2.1.3. The strong force acts on colour charged particles. It is has a range comparable to the size of an atomic nucleus and is mediated by eight independent types of massless gluons. The gluon multiplicity is labelled by a combination of colour charge and anti-colour charge. For example, the red-antiblue combination is one of the gluon colour states.

2.1.3 Electroweak Theory and the Higgs Mechanism

Electroweak theory describes the unification of the weak with the electromagnetic interactions. In QED, mediators of the weak force act similarly as photons that mediate electromagnetic interactions above the unification energy scale. In order to incorporate electromagnetism into the theory of the weak force, $SU_L(2)$ is introduced to the $U_Y(1)$ weak-hypercharge phase symmetry. The two subsequently form the electroweak gauge group, $SU_L(2) \otimes U_Y(1)$ with four massless gauge bosons, photon (γ), $W^\pm$ and Z. The weak hypercharge, Y, is defined through the Gell-Mann-Nishijima formula,

$$Q = L + \frac{1}{2}Y \tag{2.2}$$

by the generator of $U(1)_{em}$, electric charge, Q, and the weak isospin, L. However, the $W^\pm$ and Z bosons are found to be massive from the experimental measurements. Therefore, with the exception of photons, the fact that bosons formed by the electroweak gauge group are all massless violates the observations. This

14

implies that the $SU_L(2) \otimes U_Y(1)$ symmetry must be broken, through a so called Electroweak Symmetry Breaking (EWSB) mechanism. A complex doublet of scalar fields with weak hypercharge $Y_\phi = +1$ given by

$$\phi \equiv \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \qquad (2.3)$$

is added to the electroweak Lagrangian, and is known as the Higgs doublet. The vacuum or the minimum energy state of the Higgs field is chosen to correspond the vacuum expectation value, v, where $v = 246$ GeV sets the scale of EWSB [6]. This vacuum state $\langle \phi_0 \rangle$ given by

$$\langle \phi_0 \rangle = \begin{pmatrix} 0 \\ v/\sqrt{2} \end{pmatrix} \qquad (2.4)$$

simultaneously breaks the $SU_L(2) \otimes U_Y(1)$ symmetry to $U(1)_{em}$ and weak bosons acquire masses. The predicted masses are given in terms of the electroweak gauge coupling, g:

$$
\begin{aligned}
M_W^2 &= \frac{g^2 v^2}{2} \\
M_Z^2 &= \frac{M_W^2}{\cos^2 \theta_W},
\end{aligned}
\qquad (2.5)
$$

where θ_W is the weak mixing angle and determines the respective coupling strengths between the $SU_L(2)$ and $U_Y(1)$ groups in the electroweak Lagrangian [8]. Photons remain massless because the $U(1)_{em}$ phase symmetry stays unbroken after EWSB. In this theory, fermions also acquire masses through the Yukawa interactions to the Higgs doublets. The remaining real component of the complex Higgs doublet is a massive, spin-zero particle, known as the Higgs boson. There are no precise predictions of its mass provided within the SM, but theoretical bounds limit its mass, m_H, in the interval 134 GeV $\leq m_H \leq 177$ GeV [9]. Figure 2.1 illustrates the upper and lower bounds on the Higgs boson mass at different energy scale from the following conditions:

- Triviality condition, which keeps the Higgs coupling λ parameter finite.

- The requirement that the EW vacuum is an absolute minimum, which

15

promises the stability of the Higgs potential.

Precision electroweak measurements are also shown in Fig. 2.1 and provide an upper bound on the Higgs boson mass at lower energy scales.

The direct search for the SM Higgs boson was first performed in the $e^+e^- \rightarrow HZ$ channel in e^+e^- colliders at the Large Electron Positron (LEP2) collider at CERN, Geneva. Direct searches at LEP excluded the Higgs boson for masses below 114 GeV [11]. Presently, the Tevatron collider extends the search in several Higgs decay channels and excludes a SM Higgs boson in the mass range $158 - 175$ GeV and $100 - 109$ GeV at 95% confidence level (CL) using sensitivity-weighted average luminosity of 5.9 fb^{-1} of data [12]. Results are shown in Fig. 2.2.

Because the mass terms of the top quark and $W^\pm$ and Z bosons are sensitive to the mass of the Higgs boson through loop corrections, it is also possible to examine the bounds of the Higgs boson mass through precision electroweak (EW) measurements. The ellipses shown in Fig. 2.3(a) are made from direct measurements of the top mass, m_t, and W mass, M_W. The green band describes the relationship between m_t, M_W and Higgs boson masses predicted by the SM. This provides an indirect experimental constraint for different Higgs masses. Furthermore, a $\Delta\chi^2$ fit for m_H using precision EW measurements is given in Fig 2.3(b). The 95% CL exclusions from LEP2 are also shown. Both direct and indirect determinations are consistent with each other and presently favour a light Higgs boson.

2.2 Why Go Beyond the Standard Model?

The Standard Model is presently one of the most successful theories yielding predictions that have been extensively tested. However, it is still unable to answer several fundamental questions such as the so called "hierarchy problem" [13] or the existence of overwhelming "dark matter" in the universe.

As described earlier, the Higgs mass is a free parameter in the SM. However, the first-order correction to the Higgs boson mass is proportional to the square of Λ and EW gauge coupling constant, g [13]:

$$\Delta m_H^2 \approx \frac{g^2 \Lambda^2}{16\pi^2}.$$

(2.6)
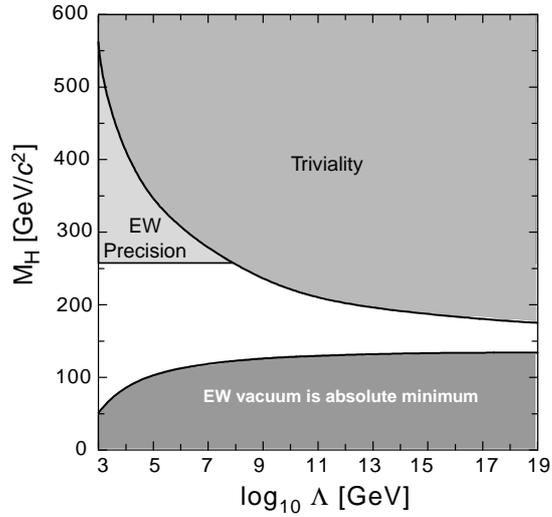
Figure 2.1: Theoretical upper bound from triviality conditions and the lower bound follows the requirement that $V(v) < V(0)$. Also shown as the upper bound in low energy scale is the limit from precision EW measurement with 95% confidence level (CL). [10]

Figure 2.2: Upper bound on the SM Higgs boson cross section obtained by the Tevatron experiment and the direct exclusion from the LEP2 experiment. The coloured bands show the 68% and 95% CL probability bands of the expected upper bound from the Tevatron data [12].

17

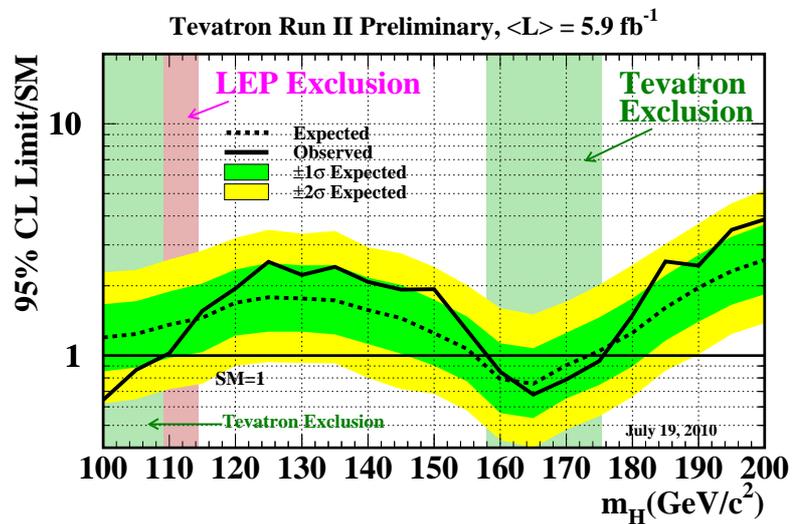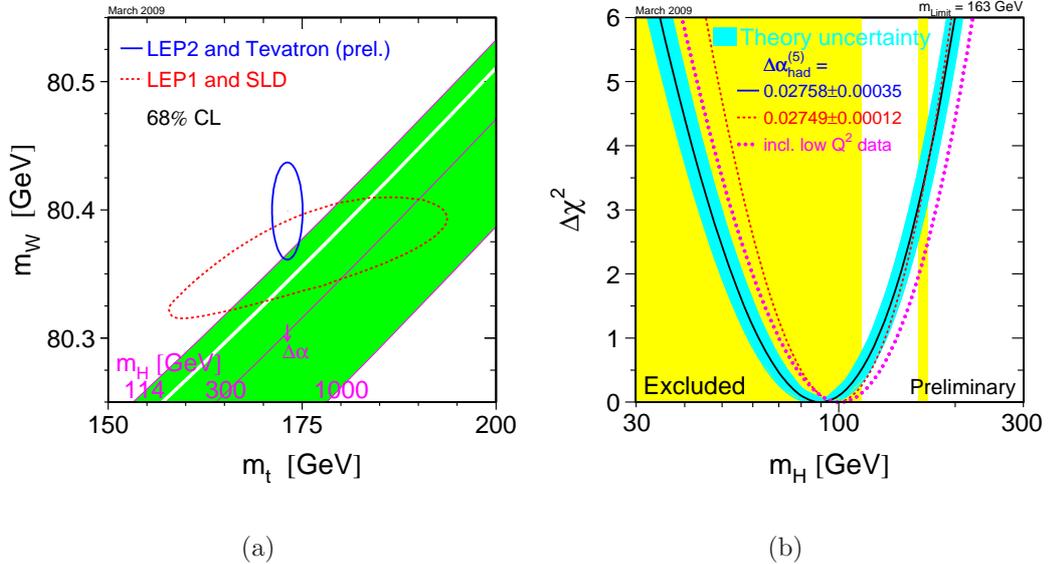(a)   (b)

Figure 2.3: (a) Experimental constraints on the Higgs masses from direct measurements of m_t and M_W. (b) $\Delta\chi^2$ (black) for the fit to the Higgs boson mass from the electroweak measurements. The excluded region by LEP2 experiment at 95% CL is shown as yellow band and the blue band provides the theory uncertainty on the prediction.

Here, Λ defines the reference energy scale for new physics beyond the SM and is sensitive to fermion masses and dimensionless couplings. Within nature, Λ is generally chosen as the Planck mass given by [13]:

$$M_{Planck} = (\frac{1}{G})^{1/2} \approx 1.22 \times 10^{19} \text{ GeV}, \tag{2.7}$$

where the constant G is the Gravitational constant. This choice is based on the fact that fundamental interactions are described by the $U_Y(1) \otimes SU_L(2) \otimes SU(3_c)$ symmetry. The known "hierarchy problem" then stresses the fact that the Higgs boson mass can diverge quadratically if the scale for new physics, Λ, is much larger than the EW scale, resulting in large Higgs boson masses. In particular, the SM is expected to be embedded in a more fundamental theory between the EW energy scale to the Planck scale in order to stabilise the mass of the Higgs boson.

Another example of the incompleteness of the SM involves the observation in astronomy in regards to why matter in the Universe is predominantly made up of non-baryonic "dark" matter. Dark matter is found to interact very weakly, and

18

there are no particles in the SM that can be candidates for dark matter in order
to explain these cosmological observations yet.

Many of these open questions left by the SM imply that there must be physics
beyond the SM. The supersymmetric (SUSY) extensions of the SM are one of the
most popular candidate theories and provide elegant solutions for those remaining
issues. The following section discusses some of these basic concepts as well as the
Higgs sector within supersymmetry.

2.3   Supersymmetry

The supersymmetric extension of the SM pairs up fermions with bosons so that
each elementary particle in the SM has a super-partner whose spin differs by half
a unit. This is defined by an operator, which transforms fermions into bosons
and vice-versa. The Minimal Supersymmetric extension to the Standard Model
(MSSM) extends the SM with a minimum set of extra parameters. By doubling
the number of leptons and quarks and adding new gauge bosons in this theory,
MSSM solves many of the open questions in the SM described in the previous
section.

Firstly, quadratical divergent contributions from fermions to quantum loop
corrections are balanced by the contributions from their SUSY scalar partners.
Figure 2.4 illustrates the one-loop corrections of fermions and those scalar part-
ners to the Higgs boson mass. This naturally provides a solution to the hierarchy
problem keeping the existence of a low mass Higgs boson reasonable. However,
the cancellation of the quadratic Λ terms in the Higgs boson mass leaves a depen-
dency on the logarithmic term: $\Delta M_H^2 \propto (m_f^2 - m_S^2)\log(\Lambda/m_S)$, where m_f and m_S
are the masses of the fermions or scalars, respectively. Such a solution to keep
the Higgs boson mass stable is only valid if $m_f \sim m_S$ which leads to constrains on
the masses of the super-particles to be less than about 1 TeV. The Higgs sector
of SUSY is important to this thesis, and therefore it is discussed in more details
in Sect. 2.3.1.

Not only being one of the solutions of the hierarchy problem, MSSM also
provides good explanations for dark matter. In MSSM, R-parity is a defined as
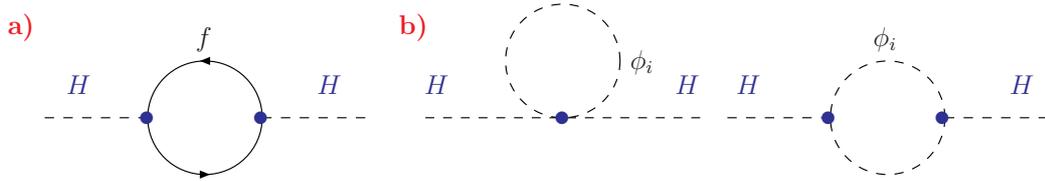
$$R = (-1)^{3(B-L)+2S},\tag{2.8}$$

Figure 2.4: Diagrams for the one-loop contributions of (a) fermions and (b) SUSY scalar partners to the Higgs boson mass. [14]

where S is the particle spin quantum number, B is the baryon number, and L is the lepton number. In the realm of the SM, all particles have $R = 1$ with their super-partners each having $R = -1$. Hence, the decay of a single SUSY particle into SM particles is not allowed if R-parity is conserved. The lightest supersymmetric particle (LSP) is then stable under this assumption of R-parity conservation. In addition, a stable LSP is required to be electrically and colour neutral [15], which implies that it can carry energy without interacting with any detector components and thereby appear as the missing energy within the inter-action. Therefore, it is possible that a stable LSP can be a promising candidate for dark matter.

2.3.1 The Higgs Sector in SUSY

The MSSM predicts a two-Higgs doublet model (2HDM) with one doublet, H_u, coupling to "up"-type fermions, while the other doublet, H_d, couples to "down"-type fermions. This physically leads to one charged pair, $H^{\pm}$ and three neutral Higgs bosons comprising one CP-odd scalar, A, and two CP-even scalars H, h where conventionally $M_h < M_H$ after electroweak symmetry breaking [16]. The masses H and h are degenerate with A in most parameter space, and all three neutral Higgs bosons, including both the CP-odd and CP-even states, are denoted as ϕ. The self-interactions for the Higgs particles depend on the gauge coupling constants, and all Higgs sector parameters can be completely specified by three parameters:

- the Z boson mass, M_Z;

- the ratio of the H_u and H_d vacuum expectation values, numerically given by $\tan\beta$;

20

- one of Higgs boson masses, conventionally chosen to be the mass of pseudoscalar A, M_A.

The MSSM vacuum expectation values are related to those within the SM by:

$$v_u^2 + v_d^2 = v^2 \approx (174 \text{ GeV})^2, \tag{2.9}$$

such that the definition of $\tan\beta$ can be written as:

$$\tan\beta \equiv v_u^2/v_d^2. \tag{2.10}$$

The other neutral Higgs boson masses at tree-level are given in terms of W, Z and A boson masses [16]:

$$m_{h,H}^2 = \frac{1}{2}(m_A^2 + m_Z^2) \mp \sqrt{(m_A^2 - m_Z^2)^2 + 4m_Z^2 m_A^2 \sin^2(2\beta)}, \tag{2.11}$$

$$m_{H^\pm}^2 = m_A^2 + m_W^2. \tag{2.12}$$

As a consequence, the mass of the lightest CP-even Higgs is constrained with an upper bound at tree-level at [13]:

$$M_h \leq m_Z |\cos 2\beta|. \tag{2.13}$$

This constraint immediately contrasts that of the SM prediction for the masses of a Higgs boson, which is only bounded by perturbative and unitarity arguments, and requires the masses must not be bigger than 1 TeV, Therefore, the lightest Higgs boson in MSSM is expected to have a mass bounded by [13]:

$$M_h \leq 135 \text{ GeV}. \tag{2.14}$$

This prediction is consistent with current observations from both direct and indirect searches.

The value of $\tan\beta$ is theoretically believed to be large and a preferred value is on the order the ratio of top to bottom quark masses, $m_t/m_b \approx 40$. At tree-level, the production cross section for a CP-odd Higgs boson produced in association with a pair of down-type fermions is proportional to $\tan\beta/(1 + \Delta_B)^2$ where the Δ_B term accounts for radiative corrections and arises from the mixing

and coupling of h and H [17]. This results in an increase in the production cross section for the Higgs boson, A, by a factor proportional to $\sim \tan^2 \beta$ compared to the Higgs cross section predicted by the SM.

Although the Higgs sector in the MSSM can be fully described by M_A and $\tan \beta$ at tree level, dependence on the other SUSY parameters must be introduced to account for higher order corrections. Because the main correction to M_h is from the $t - \tilde{t}$ sector and $b - \tilde{b}$ sector (only with large $\tan \beta$), the most important parameters relating corrections to M_h are:

- the top quark mass m_t;

- the common scalar mass M_{SUSY};

- $X_b \equiv A_b - \mu \tan \beta$;

- $X_t \equiv A_t - \mu / \tan \beta$, where $A_{t/b}$ are the trilinear Higgs sfermion couplings;

- the higgsino mass parameter, μ, which is also known as the MSSM Higgs boson mixing parameter;

- the further dependency on the SU(2) gaugino mass parameter M_2;

- the gluino mass, $m_{\tilde{g}}$, which enters at the two-loop level.

Two CP-conserving benchmark scenarios, M_h^{max} and *no-mixing*, are commonly used for the MSSM Higgs boson searches at the Tevatron in order to exemplify different limiting cases in the MSSM parameter space [18]. A M_h^{max} scenario maximises the mass of the lightest MSSM Higgs boson for a given $\tan\beta$ and M_A, and the stop (the SUSY partner of the top quark) mixing parameter is set to be a large value, $X_t = 2M_{\mathrm{SUSY}}$. Hence, it provides the largest parameter space in M_h and as a consequence, relatively conservative $\tan \beta$ exclusion bounds. On the other hand, the *no-mixing* scenario is designed to make the mixing in the stop sector vanish ($X_t = 0$). This gives a rise to a relatively restricted μ parameter space. Varying the value and sign of μ demonstrates the effect of radiative corrections on the production and decay processes, therefore, the dependency on μ is expected to be weaker in the $\tau^+\tau^-$ final state due to compensations between large corrections in the Higgs production and the decay. The limits on the cross-section of MSSM Higgs bosons are consequently translated into an exclusion in

the (M_A-tanβ) plane for the two scenarios for both positive and negative values of μ. The common choices for the values of the SUSY parameters on each scenarios are listed in Table 2.4. FEYNHIGGS is a program used in this analysis to interpret cross-section limits into limits on MSSM M_A-tanβ parameter space using different scenarios [22].

Parameter	M_h^{max}	No-mixing
X_t	2 TeV	0 TeV
μ	± 0.2 TeV	± 0.2 TeV
M_2	0.2 TeV	0.2 TeV
$m_{\tilde{g}}$	0.8 TeV	1.6 TeV
M_{SUSY}	1 TeV	2 TeV

Table 2.4: The MSSM parameters for the M_h^{max} and *no-mixing* scenarios [18].

The production processes of the MSSM neutral Higgs bosons, ϕ (=h, H and A), are very similar to those for the SM Higgs boson but differ in rates. The leading order processes are [14]:

- associated production of h/H with W/Z bosons: $q\bar{q} \rightarrow W/Z + h/H$;

- vector boson fusion: $qq \rightarrow V^*V^* \rightarrow qq + h/H$ with $V = W/Z$;

- gluon-gluon fusion: $gg \rightarrow h/H/A$;

- associated production with heavy quarks: $gg, q\bar{q} \rightarrow Q\bar{Q} + h/H/A$.

For each, the most important Feynman diagrams for hadronic collisions at tree-level are shown in Fig. 2.5.

Figure 2.6 shows the production cross sections of MSSM Higgs bosons at the Tevatron for fixed $\tan\beta = 40$ within the *no-mixing* and M_h^{max} scenarios with $\mu = +200$ GeV. [19]. The results illustrate a large increase of the cross section in the MSSM relative to the SM. A further increase on the cross section is expected at $M_A \sim 132$ GeV where the mass of the three neutral Higgs bosons become degenerate. The dominant production processes at the Tevatron are gluon fusion ($gg \rightarrow \phi$) and the production of Higgs from b quarks ($bb \rightarrow \phi$) as apparent from Fig. 2.6.

Direct measurements of m_t and M_W at the Tevatron also provide indirect constraints of the Higgs boson masses. The MSSM parameter space is scanned
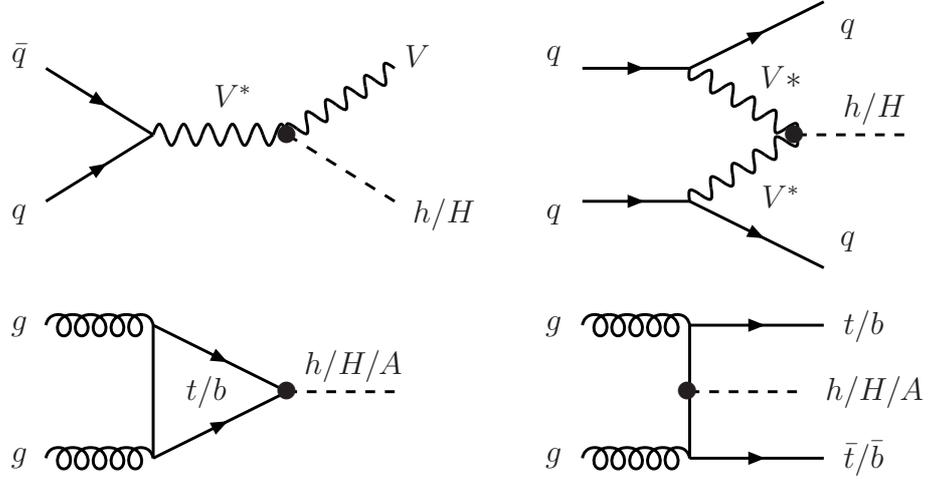
Figure 2.5: The dominant MSSM neutral Higgs production mechanisms in hadronic collisions.
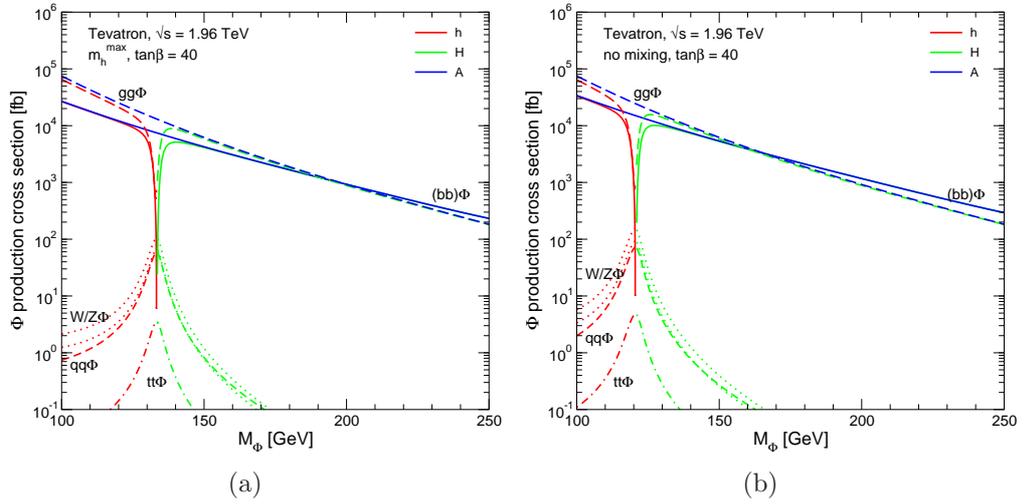

Figure 2.6: Higgs boson production cross sections as a function of Higgs boson mass at the Tevatron, (a) in the M_h^{max} MSSM scenario and (b) in the *no-mixing* MSSM scenario, for $gg \to \phi$ and $b\bar{b} \to \phi$ [19].

where the Higgs boson is most likely to be favoured as a function of m_t and M_W at the 68% CL is given in Fig. 2.7. The current measurements are consistent with the SM but have a slight preference for the MSSM. This further motivates a search for the Higgs bosons beyond SM for the MSSM model.
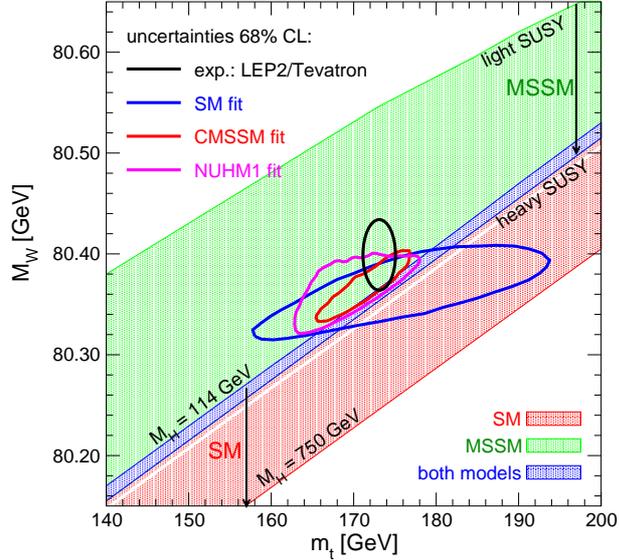

Figure 2.7: The current experimental measurements for M_W, and m_t, is shown as the black ellipse. This result is compared with the SM prediction region shown as red band and the green band for the region allowed in MSSM. The blue, red and pink ellipses are predicted by SM, Constrained Minimal Supersymmetric Standard Model (CMSSM) and Non-Universal Higgs Mass Model (NUHM1) fit with 68% CL. [20]

Considering the upper constraint on the masses of different Higgs bosons for $M_h \ll W^+W^+$, the lightest Higgs boson is expected to decay mainly to fermion pairs. Therefore, the most dominant channels are $\phi \to b\bar{b}, \tau\bar{\tau}$. Figure 2.9 shows the decay branching ratios of the SM Higgs. These should be compared to ones for the lightest CP-even Higgs boson h and heavy CP-even Higgs boson H in the MSSM as shown in Fig. 2.10 to 2.12 for $M_h^{\max}$ and $\tan\beta = 3$ for cases with $M_A \gtrsim 150$ GeV (left) as well as $\tan\beta = 30$ with $M_A \gtrsim 400$ (right) [14]. In this case, the lighter h boson reaches its maximal mass value and couples as a SM Higgs boson. Figure 2.8 shows the decay branching ratios of MSSM neutral Higgs bosons with $\tan\beta = 30$ and $M_A \sim 120 - 140$ GeV. Within this regime, the couplings to up-type fermions are suppressed and the couplings to b and τ pairs are strongly enhanced [14].

25
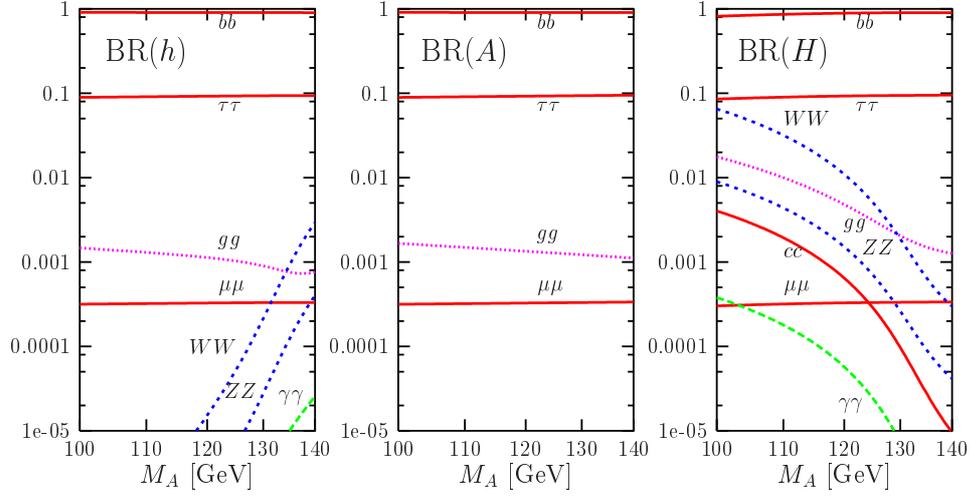
Figure 2.8: The decay branching ratios of the MSSM neutral Higgs bosons, h, H and A as a function of M_A in the regime $M_A \sim 120-140$ GeV with $\tan\beta = 30$ [14].

Figure 2.9: Main branching ratios for the SM Higgs boson [21] as a function of the Higgs boson mass.

Figure 2.10: The decay branching ratios of the lighter CP-even MSSM h boson as a function of its mass for $\tan\beta = 3$ (left) and $\tan\beta = 30$ (right) [14].

Figure 2.11: The decay branching ratios of the heavy CP-even MSSM H boson as a function of its mass for $\tan\beta = 3$ (left) and $\tan\beta = 30$ (right) [14].

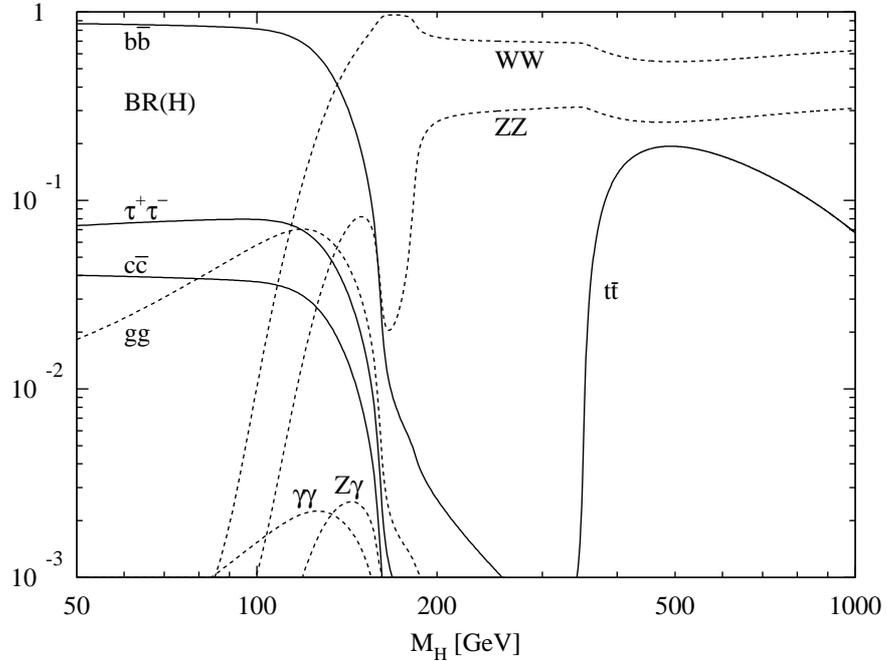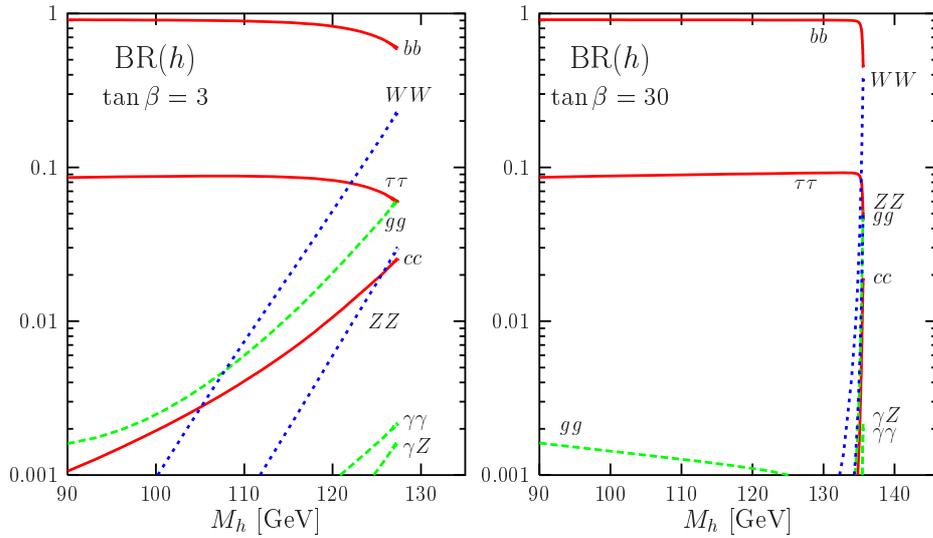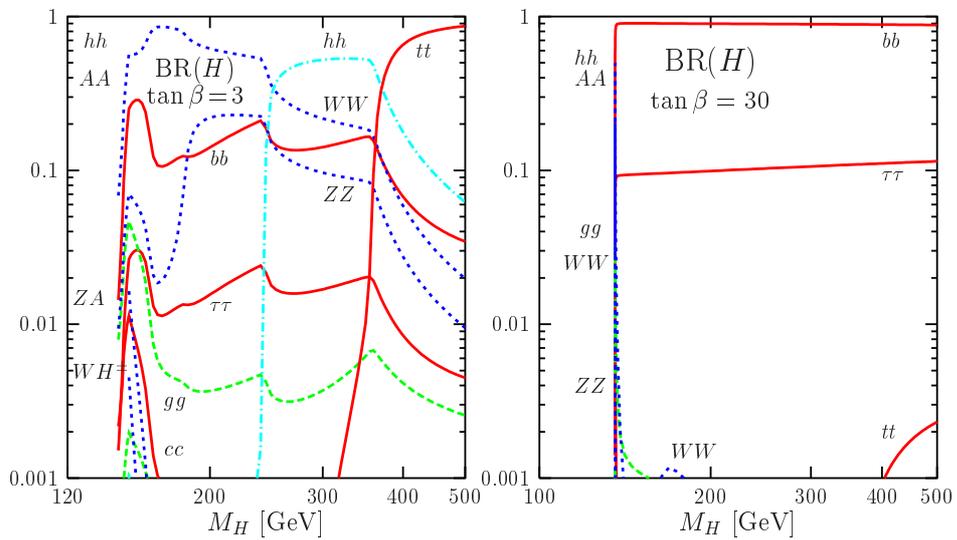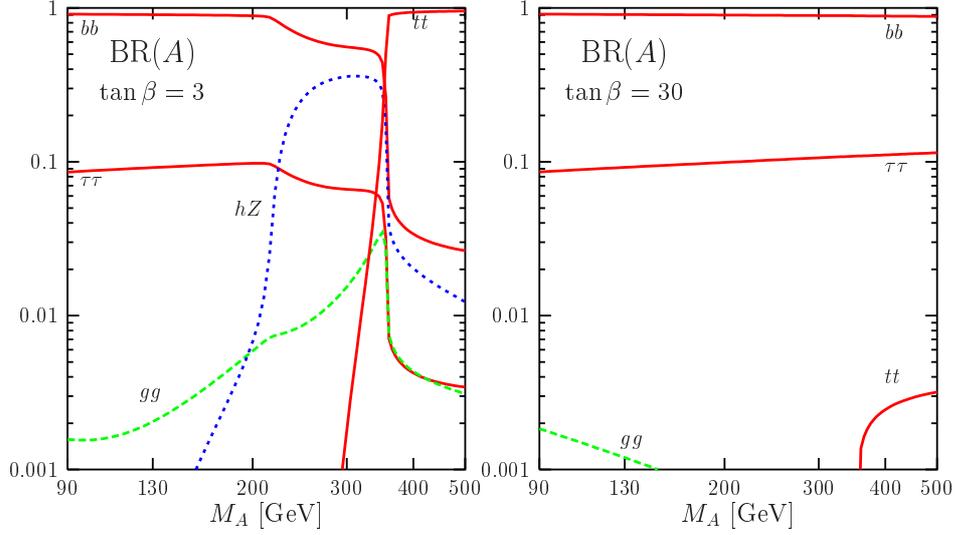Figure 2.12: The decay branching ratios of the CP-odd MSSM h boson as a function of its mass for $\tan\beta = 3$ (left) and $\tan\beta = 30$ (right) [14].

As evident from the decays shown in Figs. 2.10 to 2.12 in most regions of the parameter space, the tree-level branching ratio of the neutral MSSM h, H and A decays into a $b\bar{b}$ pair is approximately 90% and $\tau^+\tau^-$ decays $\approx 10\%$. The two relevant Feynman diagrams for the decay of a CP-odd Higgs boson into τ pairs are shown in Fig. 2.13.

A direct search for $p\bar{p} \to \phi \to b\bar{b}$ is experimentally difficult at hadron colliders due to the enormous background from heavy-flavour production. This makes the $\tau^+\tau^-$ mode the most promising inclusive search channel. One can estimate the total production rate of τ pairs mediated by the production of a CP-odd Higgs boson with large $\tan\beta$ for fixed masses of b and τ. The relation is given by [17]:

$$\sigma(gg, b\bar{b} \to A) \times BR(A \to \tau^+\tau^-) \approx \sigma(gg, b\bar{b} \to h)_{SM} \frac{\tan^2\beta}{(1+\Delta_B)^2 + 9}, \qquad (2.15)$$

where $\sigma(gg, b\bar{b} \to h)_{SM}$ is the value of the corresponding SM Higgs production cross section, for a SM Higgs boson mass of M_h.
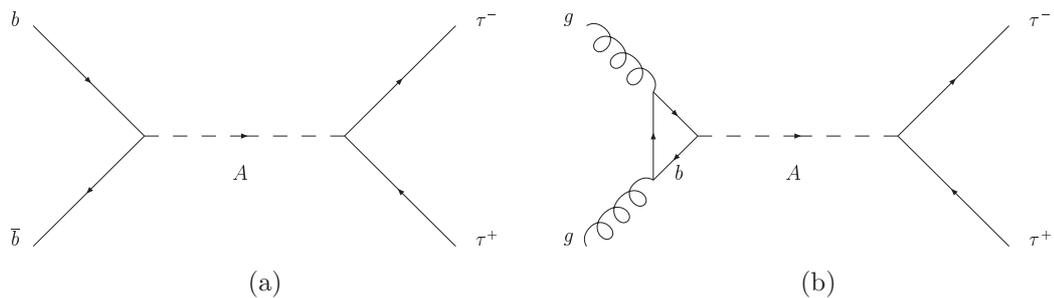
Figure 2.13: The two main production modes for the pseudo-scalar Higgs boson A in the MSSM, (a) annihilation of a pair of b quarks and (b) gluon fusion, via a loop of b quarks.

2.3.2 FEYNHIGGS Program

To translate limits of production cross sections to $(M_A, \tan\beta)$ exclusion results, the program FEYNHIGGS [22] is used. This program is based on the results obtained in the Feynman-diagrammatic approach. It is a FORTRAN program for computing masses and related observables, such as mixings, branching ratios (BR), and higher-order couplings of the Higgs bosons in the MSSM at the two-loop level. CP-violating effects are also introduced by complex parameters such as the Higgsino mass parameter, μ, and enter the calculation via loop corrections. For setting limits in the analysis presented in this thesis, the theoretical cross section is multiplied by the branching ratio $(\sigma(p\bar{p} \to \phi) \times Br(\phi \to \tau\tau))$ at each M_A mass point, and subsequently increases the value of $\tan\beta$ until the cross section from FEYNHIGGS matches the observed or expected limit discussed in the previous section. The value of $\tan\beta$ for a given value of M_A represents the observed or expected exclusion limit, respectively.

Chapter 3

Experimental Apparatus

The Tevatron Collider [23] at the Fermi National Accelerator Laboratory near Chicago, USA, was the world's highest energy particle collider until the Large Hadron Collider (LHC) at CERN began operations in 2009. The Tevatron enables head-on collisions of protons (p) with anti-protons ($\bar{p}$) at a centre-of-mass energy $\sqrt{s} = 1.96$ TeV. Operations for Run I of the Fermilab Tevatron collider lasted from 1992 to 1996 and Run II started in early 2001 and runs to present. The Tevatron reached integrated luminosities of up to 4.02×10^{32} cm^{-2} The data produced in $p\bar{p}$ collisions are collected by two major experiments, CDF and DØ. Both experiments have general-purpose particle detectors designed to identify particles created in the collisions, which include muons, photons or electrons, τ leptons, particle jets and the presence of undetected neutrinos. A short description of the particle acceleration procedure mechanism is given in Sect. 3.1. Section 3.2 follows with a summary of each major component of the DØ detector.

3.1 The Tevatron

An overview of the accelerator chain leading to the Tevatron is shown in Fig. 3.1. More detailed descriptions of each component can be found in [23].

3.1.1 Proton Source

Production of high energy protons begins at the *preaccelerator* which consists of a Cockcroft-Walton generator with an accompanying negative hydrogen ion (H^-) source. Negative ions are accelerated to 750 keV under high voltage generated by

the Cockroft-Walton and injected into a 150-meter-long linear accelerator, *Linac*, through a transport line. The *Linac* further accelerates these H^- ions to higher energies of 400 MeV using Alvarez drift-tubes with an alternating electric field and a series of Ratio Frequency (RF) accelerating cavities. There is a second part of the *Linac*, which is often called "side coupled" cavities. The *Booster* is a synchrotron approximately 150 metres in diameter; a synchrotron is a circular accelerator that confines particles using magnetic fields and allows them to accelerate by RF in a circular orbit. Inside the *Booster*, H^- ions are passed through a carbon foil which strips electrons from the negative hydrogen ions [23], and finally 8 GeV proton beams are passed into the Main Injector.

Subsequent to the *Booster* system is the *Main Injector* with a circumference of 3.3 km located adjacent to the Tevatron Collider. It is another circular synchrotron like the *Booster* and capable of accelerating the 8 GeV proton beams to energies of 150 GeV. Protons are not stored for long but transferred as soon as they are accelerated.
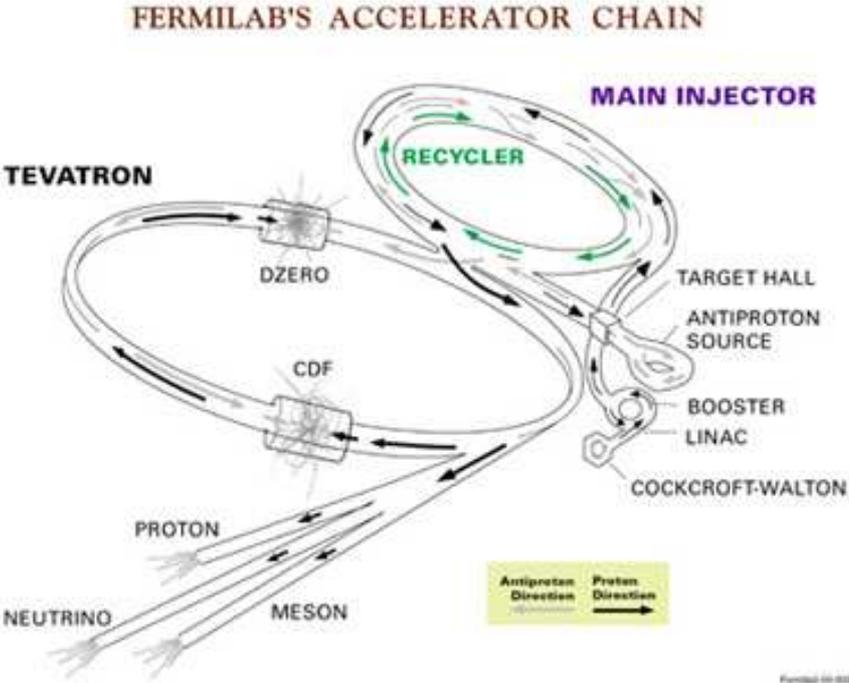

Figure 3.1: An overview of the acceleration system. [23]

3.1.2 Antiproton Source

Anti-proton beams are accelerated in the *Main Injector* but stored in the *Recycler* which is in the same tunnel as the *Main Injector* but with different systems. There are three main parts of the anti-proton production: the Target Station, the Debuncher, and the Accumulator, plus the transport lines associated with each of these three devices. Figure 3.2 provides a general layout of the Antiproton Source. The 120 GeV protons are incident with a small spot size of 0.1 mm on a nickel target, and a large number of particles are produced during the collision. Anti-protons with the energy of ≈ 8 GeV are collected and focused by a cylindrical lithium lens placed immediately after the target and passed to the Debuncher ring. Only 20 anti-protons can be collected from every 10^6 incident protons at the target, and particles that are not selected will pass to a beam dump. The Debuncher is designed to rotate and adiabatically debunch the anti-proton bunches to form DC beams. These beams are then sent into the orbit of the 520 m circumference Accumulator. The Accumulator is roughly triangular in shape and cools and temporarily stores the anti-protons. Once a sufficient number of anti-protons are accumulated, they are passed to the Main Injector and accelerated to 150 GeV.
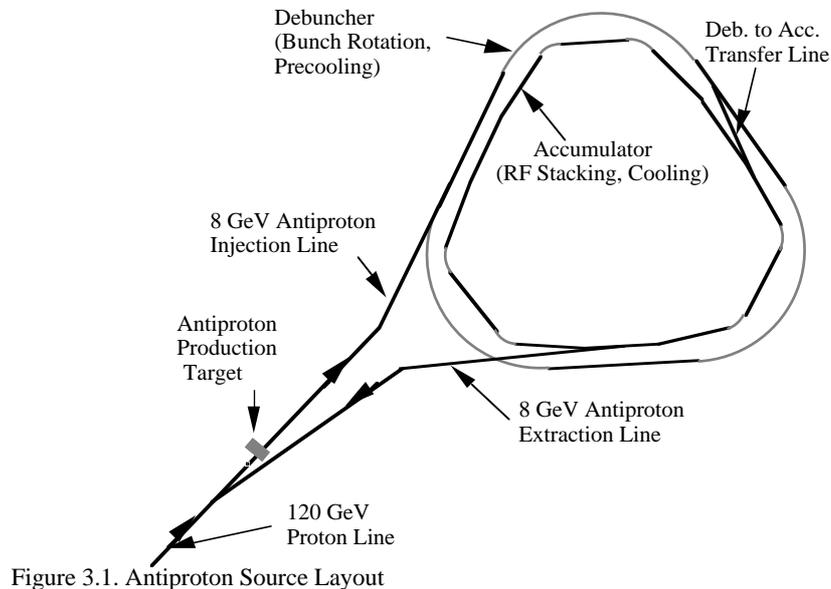

Figure 3.1. Antiproton Source Layout

Figure 3.2: Anti-proton Source layout [23].

3.1.3 Tevatron Collider

Both the proton and anti-proton beams are eventually sent to the Tevatron via the Main Injector for collisions. The Tevatron is a synchrotron accelerator of nearly 2 km diameter. Superconducting magnets located in series within the Tevatron tunnel allow the $p\bar{p}$ beam to be simultaneously accelerated to collide at a centre of mass energy of 1.96 TeV. At the start of each store, the thirty-six proton bunches (150×10^9 protons each) are injected into the Tevatron followed by the same number of anti-proton bunches (50×10^9 anti-protons each). The bunches circulate the Tevatron in opposite directions and are organised in three groups known as super-bunches. The beams are accelerated to reach an energy of 980 GeV and reach instantaneous luminosities of up to $\sim 400 \times 10^{30}$ cm^{-2}s^{-1}. However, due to beam tuning, typical stores begin with luminosities in the range of $250 - 350 \times 10^{30}$ cm^{-2}s^{-1}. Finally, the beams are focused by special super-conducting quadrupoles, and the protons and anti-protons are collided at their beam crossing points. Each interaction point is surrounded by a general purpose particle detector, CDF (BØ) or DØ. Further details about the DØ detector will be described in the next section.

The general equation for calculating instantaneous luminosity ($\mathcal{L}(t)$) is given by

$$\mathcal{L}(t) = \frac{fN^2}{A}, \tag{3.1}$$

where f is the revolution frequency, N is the number of particles, and A is the cross-section area. The form used for the Tevatron is more complicated in order to take into account the bunch size and structure. The instantaneous luminosity decreases during the lifetime of a store, which is typically around 24-36 hours.

3.2 The DØ Detector

The DØ detector [24] is a multi-purpose detector designed to capture the signature of particles spraying radially out at the collision points. Approximately 120 pb^{-1} of integrated luminosity was collected at DØ during Run I. The average data taking efficiency is $\sim 90\%$ and about 10 fb^{-1} of integrated luminosity is expected to be recorded by the end of 2011. The integrated luminosity delivered by the Tevatron and recorded by the DØ detector for Run II up to July 2010 is
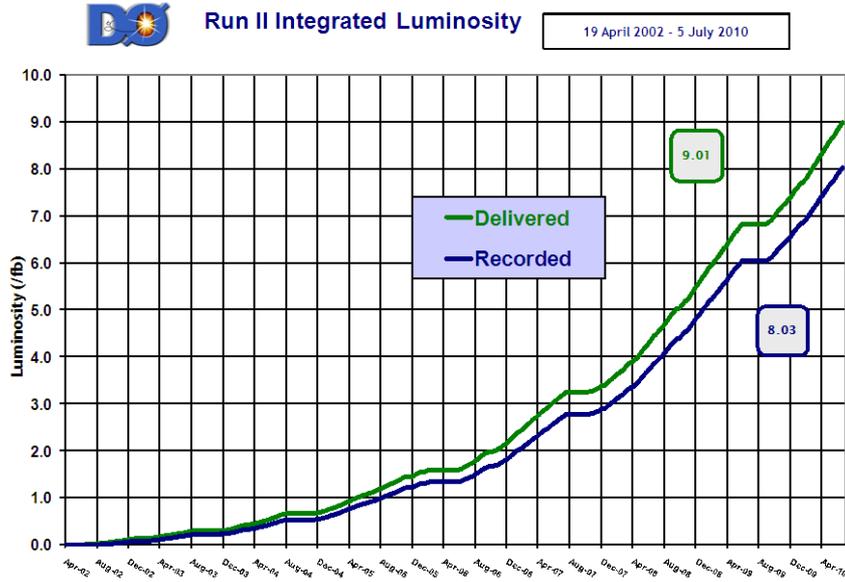
shown in Fig. 3.3 [25].

Figure 3.3: Run II integrated delivered (green) and recorded (blue) integrated luminosity.

The DØ detector is divided into three major parts: a central tracking system, uranium/liquid-argon calorimeters, and a muon spectrometer. The central tracking volume has been fully replaced in DØ Run II, and includes a silicon microstrip tracker and a scintillating-fiber tracker both located inside a 2 T solenoidal magnet. It is designed to record the trajectory of charged particles. Preshower scintillating detectors have been added in between the solenoidal magnet and the central calorimeter, and upstream of the two forward calorimeters. The calorimeter surrounds the tracker system. The outermost layer of the DØ detector comprises the muon system, which is designed to measure the path and the momentum of the muons. Figure 3.4 shows a side view of the upgraded DØ detector. There are also other important components of the detector, such as the luminosity monitor and the trigger system, which help select events that are of most interest to high-energy physics states. A right-handed coordinate system is used in which the z-axis is along the proton direction and the y-axis is upward as shown in Fig. 3.4. The angles ϕ and θ are the azimuthal and polar angles, respectively. The r coordinate denotes the radial distance from the z axis. The pseudorapidity,

$$\eta = -\ln(\tan(\theta/2)) \tag{3.2}$$

34

approximates the true rapidity,

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

for finite angles in the limit that $(m/E) \to 0$. The term "forward" is used to describe the regions at large $|\eta|$ whereas "central" corresponds to the small $|\eta|$ region. Momentum and energy are often measured in the transverse plane (x-y), and the transverse momentum is denoted as $p_T = p \sin \theta$.

The following sections give a brief description of each part of the DØ detector. A more complete description of the detector can be found in [24].
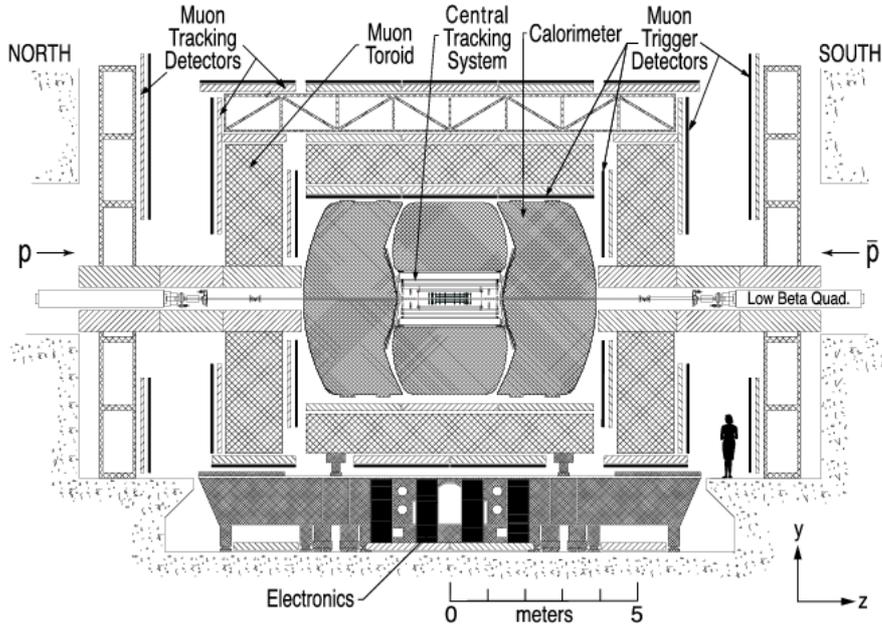
Figure 3.4: Schematic y-z view of the DØ detector. [24]

3.2.1 Central Tracking System

The central tracking system was fully replaced in DØ Run II and occupies the volume nearest to the beam pipe. A schematic view of the system is given in Fig. 3.5. The tracker comprises two sub-systems: the Silicon Microstrip Tracker (SMT) and the Central Fiber Tracker (CFT). A 2 T solenoidal magnet surrounds these two systems and consequently bends charged particles in the $r - \phi$ plane which allows precise measurements of the momentum. The ratio of calorimeter

energy to track momentum (E^{cal}/P^{trk}) is used for for efficient particle identification, and calibration. The resolution of the primary vertex position in the tracker is about 35 μm.
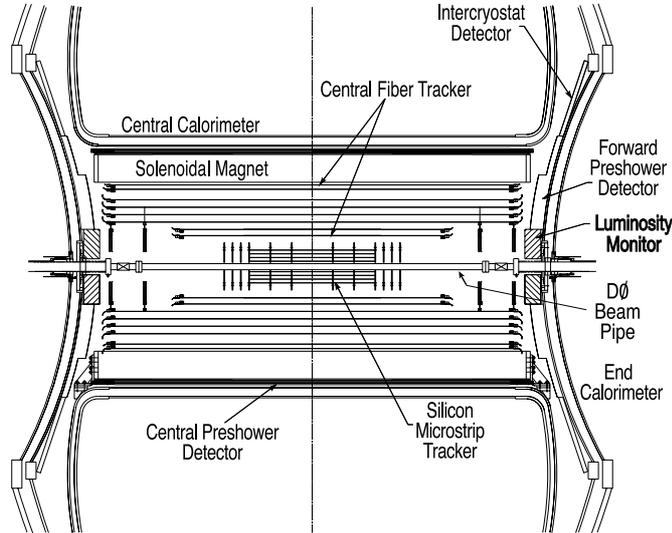
Figure 3.5: A schematic view of the central tracking system in the $r - z$ plane. The geometrical view of the solenoid, the preshower detectors, luminosity monitor relative to the calorimeters are also shown. [24]

The SMT [26] provides tracking and vertexing information with $|\eta|$ coverage up to 3 but is limited at high η. It consists of six barrel modules in the central region, each having four silicon readout layers. The centres of the barrels are located at $|z| = 6.2, 19.0, 31.8$ cm. There are 72 silicon modules, called "ladders" installed inside a barrel. The barrels are capped at high $|z|$ by a disk termed "F-disk". The respective location of each F-disks is at $|z| = 12.5, 25.3, 38.2$ 43.1, 48.1, and 53.1 cm. The F-disks have an outer radius of about 10 cm and are made up by 12 double-sided wedge detectors. In addition, two larger-diameter "H-disks" are located at each end at $|z| = 100.4$ and 121.0 cm to strengthen the coverage at high rapidity. Each H-disk consists of 24 wedges and has an inner (outer) radius of 9.5 (26) cm. The main reason for this barrel-disk arrangement is to capture the three-dimensional information of tracks traversing perpendicular to the detector surface. Such a design allows the $r - \eta$ coordinates to be measured by the barrels while the $r - \phi$ coordinates are measured by the disks. Figure 3.6 shows the layout of SMT detector.

The SMT strip sensors are p-n junction diodes operated at reverse biases and

Figure 3.6: Geometrical design of the Silicon Microstrip Tracker (SMT). [24]

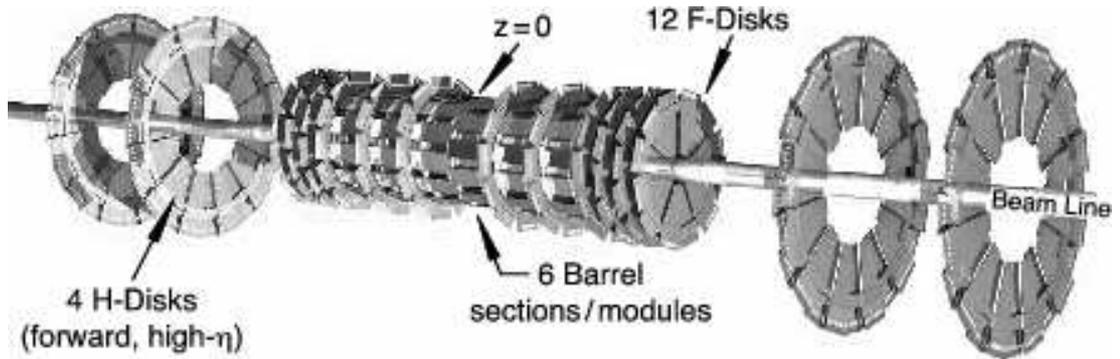are arranged parallel to the long edge of the device. These p-n junctions form a depleted zone with no free electrons in the conduction band. When a charged particle passes through the detector, it induces a charge separation within the material, creating electron-hole pair, which is then collected by a capacitor in the readout.

The CFT [27] is radially located outside the SMT and precisely measures the position of charged particles. It consists of scintillating fibers mounted on 8 concentric support cylinders and provides coverage for $|\eta|$ up to 1.7. The two innermost cylinders are 1.66 m long, while the outer cylinders are 2.52 m long. This difference is to allow readout cables from both SMT and inner layers of CFT. Each fiber is 835 μm in diameter and is capable of providing a resolution of about 100 μm. As a charged particle passes through a fiber, the material will be ionised and thereby emit photons. The photon is then converted into electrical signals by visible light photon counters (VLPCs) located at the DØ platform. A fast and continuous readout from the CFT is sent to the Level 1 trigger system. Signals that pass Level 1 are propagated to Level 2, while Level 3 uses full CFT readout information for trigger decisions. The trigger system at DØ is described in Sect. 3.3.

3.2.2   Preshower Detectors

The preshower detectors can be used to correct the energy losses of particles due to materials. They further help improve the electron and photon identification at DØ. These detectors function as both calorimeters and tracking detectors.

Information from these detectors is also included in the Level 1 trigger due to their fast energy and position measurement. There are two preshower systems at DØ: one is the Central Preshower (CPS) detector, which is located between the solenoid and the central calorimeter and covers the region $|\eta| < 1.1$, while the other is the Forward Preshower (FPS) detector, which is made up of two main modules with each located on the face of an end calorimeter to cover $1.5 < |\eta| < 2.5$. The CPS and FPS are both constructed using triangular strips of scintillator with inserted wavelength shifting fibers, as shown in Fig. 3.7
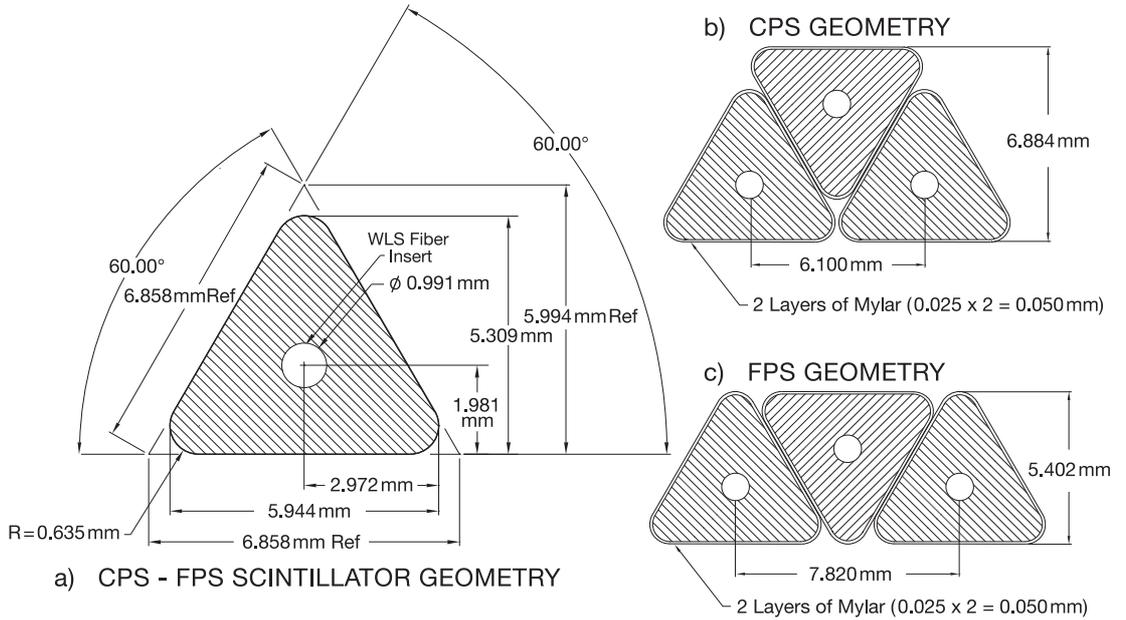
Figure 3.7: Cross sectional view and layout geometry of the triangular scintillator strips of both CPS and FPS. The central hole allows for insertion of the 835 μm wavelength shifting fiber described in the text.

The preshower detectors share common elements with the CFT such as the waveguides and the front-end readout electronics system. However, the connections between the wavelength shifting fibers and the waveguides are unique to the preshower detectors as are the gains and biases for the electronics.

3.2.3   Calorimeter

The DØ calorimeter system remains unchanged from Run I, and consists of three sampling calorimeters (primarily uranium/liquid-argon). It provides the energy

38

measurements for different objects, electrons, photons, muons and jets. There are three main parts of the DØ calorimeter system as shown in Fig. 3.8.
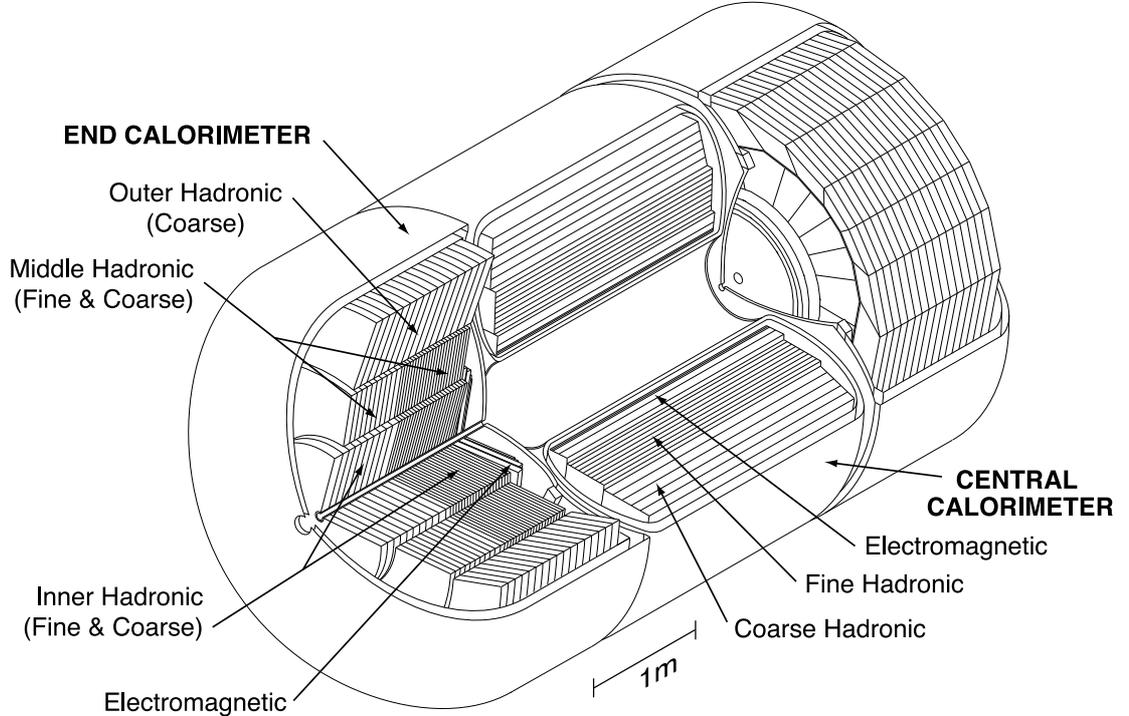
Figure 3.8: Overview of the central and end calorimeters. [24]

The central calorimeter (CC) covers pseudorapidity up to $|\eta| \simeq 1.1$, and two end calorimeters (EC) extend the coverage to $|\eta| \simeq 4.2$. All three are housed in their own cryostat, which maintains a temperature of approximately 90 K. The electromagnetic section is situated closest to the interaction point radially followed by fine and coarse hadronic sections. The liquid argon sampling material is used as the active medium for the ionisation to occur. The choice of material for the absorber plates are different in each section and are arranged as follows: thin plates made from nearly pure depleted uranium (3 mm in CC and 4 mm in EC) in the electromagnetic sections (EM), 6-mm-thick uranium for the fine hadronic sections, and relatively thick (46.5 mm) plates of copper (CC) or stainless steel (EC) for the coarse hadronic modules. The basic unit of the calorimeter cell is shown in Fig. 3.9, and consists of plate absorber material, a 2.3 mm liquid argon gap and a signal board.

The metal absorber is grounded while the connected resistive surfaces are kept at a high voltage of +2000 V to establish an electric field. Under these
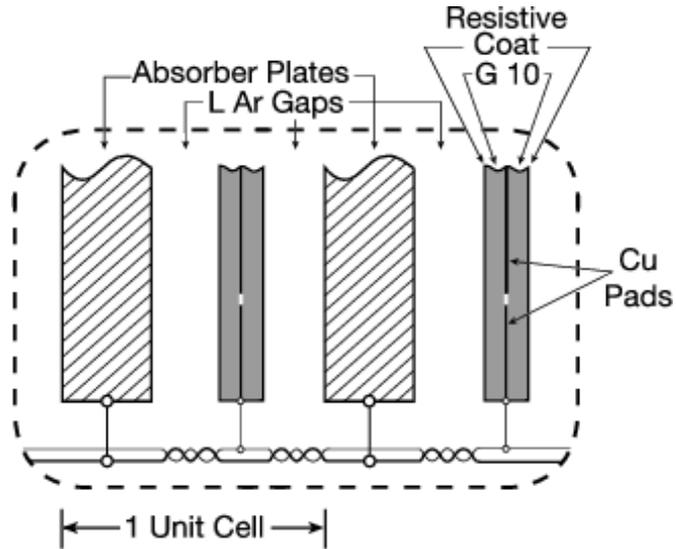
Figure 3.9: Layout of a unit cell for the calorimeter. [24]

configurations, the typical electron drift time across the liquid argon gap is about 450 ns. The calorimeter readout cells are arranged in pseudo-projective towers such that the centres of cells lie along the particle projection from the interaction point. Figure 3.10 shows a schematic view of the EC electromagnetic readout cells.

The tower of each cell in both EM and hadronic modules has a size of $\Delta\phi = 0.1$ and $\Delta\eta = 2\pi/64 \simeq 0.1$, and is twice as finely segmented in both η and ϕ at the third layer of the EM modules (EM3). Because the granularity in the EM3 is the best relative to the other layers, it is normally used as the "seed" of the energy cluster. More details about the cluster reconstruction algorithm can be found in Sect. 4.1.2. Such a segmentation is motivated by the fact that the EM tower maximum is expected within the third layer. The transverse sizes of the cells are designed to match the transverse size of the shower: $1 - 2$ cm for EM showers and about 10 cm for hadronic showers. The measured resolution of the EM calorimeters is $\simeq 5 - 7\%$ for electrons with energies above 20 GeV, while $\approx 30\%$ for jets in the hadronic calorimeters.

3.2.4 Intercryostat Detectors

Intercryostat detectors (ICD) cover the region $1.1 < |\eta| < 1.4$ in order to compensate for the incomplete coverage arising from the three separate calorime-
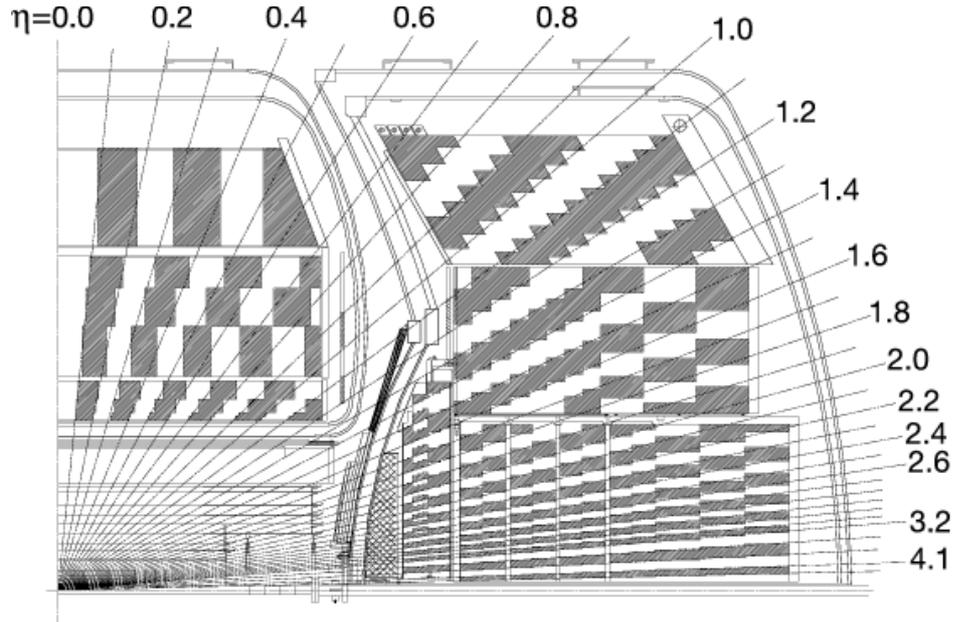
Figure 3.10: Schematic view of the EC electromagnetic readout cells. EM1 indicates EM layer 1 which is the layer nearest to the DØ interaction region, and MH is the EC middle hadronic cell.

try cryostats. The ICD is mounted to the exterior surfaces of each end of the cryostat. It is made of a series of scintillating tiles, which are subdivided into twelve straight-edged trapezoidal modules each extending over a region of $\Delta\phi \times \Delta\eta \simeq 0.1 \times 0.1$ rad. A scintillator absorbs energy of charged particles and re-emits the energy as photons. The photons are detected by photo-multiplier tubes (PMTs) after being shifted in wavelength. One half of a tile is missing at the south end of the detector to accommodate services from the solenoid magnet, the total number of the channels is 378 from the ICD system.

3.2.5 Muon System

Most particles deposit their energy within the calorimeter, except neutrinos and muons that do not interact via strong forces. The general purpose of the muon system is to provide efficient momentum measurements and effective triggering to reject backgrounds for muons. Figure 3.11 describes the upgraded DØ Run II muon system. Four main components comprise the system: toroidal magnets, central muon proportional drift tubes (PDTs), central scintillation counters, and

forward muon system composed of mini drift tubes (MDTs).
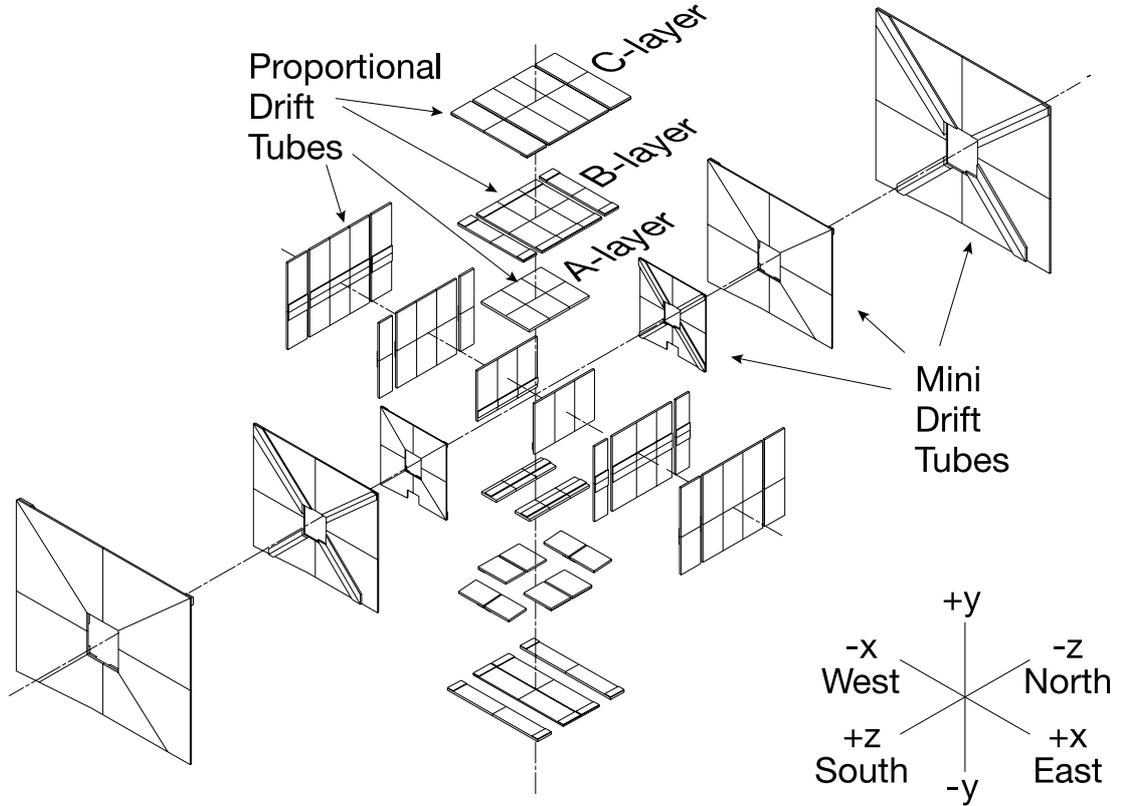
Figure 3.11: Expanded view of the muon drift tubes in the A, B and C layers of the muon system

The central toroidal magnets provide a stand-alone muon-system momentum measurement. This not only helps reject muons from pion or kaon decays, but also allows a low-p_T cutoff in the Level 1 muon trigger. To impose muon identification and momentum resolution of high p_T muons, track matches with the central inner detector are required. The magnets are operated in series and at a current of 1500 A. Together with the toroidal magnet, the central muon system is instrumented with a toroidal magnet, drift chambers (PDTs), the cosmic cap and bottom scintillation counters, and Aϕ scintillation counters which cover the PDTs of the A layer in the azimuthal angle ϕ. The system covers the range up to $|\eta| \simeq 1.0$. There are three layers of the drift chambers, the layer A inside the toroidal magnet and layer B and C located radially outside. The layers provide about 55% of the coverage of the central region with all three layers, and $\approx 90\%$ by any two. The electron drift time and the charge deposition are recorded for

each PDT hit and the information is used to determine the hit position. The fast timing signal provided by the cosmic cap and the bottom counters located outside the PDT layers can identify a muon from cosmic ray backgrounds.

The forward muon system extends the coverage to up to $\eta \simeq 2.0$ and consists of scintillator counters and three layers of MDTs. The layer closest to the interaction point is called the A-layer and is instrumented by four planes of tubes mounted along the magnetic field lines. The C layer is the furthest away. Both B and C layers have three planes of tubes. Muon trigger scintillation counters are also mounted on all three layers with each layer containing about 96 counters that are used for precision timing measurement for the arrival muons. The ϕ segmentation is approximately 4.5 degrees in order to match the CFT trigger sectors.

3.2.6   Luminosity Monitors

The Luminosity Monitor (LM) serves to measure the delivered instantaneous luminosity from the Tevatron to the experiment. This is achieved by measuring the rate of inelastic $p\bar{p}$ collisions. The DØ LM system is made up by two arrays of 24 plastic scintillation counters with PMT readout each located at $z = \pm 140$ cm around the beam pipe. Figure 3.12 shows the schematic drawing for the LM counters and the locations of the PMTs. The arrays surround the beam pipe and are located in front of the end calorimeters to detect the low momenta particles from soft collisions. The length of each counter is 15 cm, and covers the rapidity range of $2.4 < |\eta| < 4.4$. The measurement of the average number of inelastic collisions, N_{LM}, determines the luminosity given by [24]

$$L = \frac{f N_{LM}}{\sigma_{LM}}, \tag{3.4}$$

where f is the beam crossing frequency and σ_{LM} is the effective cross section for the LM. The acceptance and efficiency of the LM detector are taken into account while determining the luminosity. The LM system can also measure beam halo rates as well as the z coordinate of the interaction vertex.
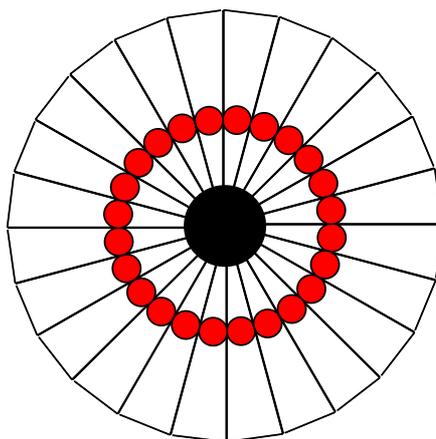
Figure 3.12: The layout geometry of the LM counters and the locations of the PMTs (shown as solid dots).

3.3   Trigger System

The trigger system is designed to pick out interesting events from approximately 2.5 million $p\bar{p}$ interactions per second. The DØ trigger system consists of three distinct levels with each successive level using more sophisticated algorithms to select fewer but more important events for further investigation. An overview of the system is shown in Fig. 3.13.
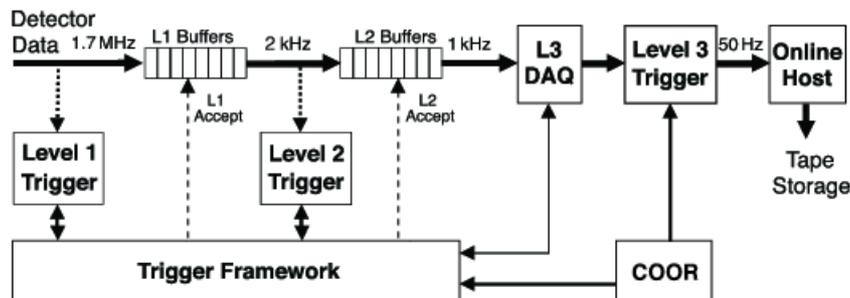
Figure 3.13: A schematic layout of the design of the DØ trigger system.

The first stage of the trigger chain, Level 1 (L1), is a hardware-based system, which recognises patterns in the detected signals and provides fast decisions with a trigger accept rate of 2 kHz. The L1 trigger system groups information from the calorimeter trigger (L1Cal), the central track trigger (L1CTT), the muon system trigger (L1Muon), and the L1 forward proton detector trigger (L1FPD).

44

Each L1 trigger device processes the input from the corresponding sub-detector, and the digital information of the passing event is collected by the L1 framework (L1FW) for a global decision on whether an event should be accepted for further processing. The L1 trigger decisions from the subsystem are required to arrive at the L1FW in 3.5 μs or less.

The Level 2 (L2) system is designed to make trigger decisions based on individual objects and reduces the trigger rate by a factor of two relative to L1 to approximately 1 kHz. A L2 decision is made at the global stage (L2Global) by studying the correlation between physics objects in different sub-detectors. These objects are reconstructed in the preprocessor by collecting the signature of candidates passed from the front-end readout and the L1 system. Each preprocessor corresponds to a detector subsystem, and the preprocessing is made either with serial CPU-based or programmable logic-based cards.

Figure 3.14 is a block diagram that shows how each level of DØ trigger system works. The events passed by the L1 and the L2 are fully digitised and sent to the Level 3 (L3) farm. The first-in-first-out (FIFO) storage provided by L1 and L2 buffers can hold event data awaiting a L2 decision before transferring to L3 and thereby reduces the experiment's dead time.

L3 trigger system is made by a programmable software with sophiscated algorithms running on a farm of Linux PCs. The L3 decisions are based on complete physics objects and the relationship between them. This is done by object-specific software algorithms with precise input physics definitions provided by programmable trigger lists. In order to match the changing of running conditions over time, such as the instantaneous luminosity, the trigger list used at DØ evolves throughout Run II. The rejection factor at L3 is around 20 and significant enough to reduce the input rate from 1 kHz to $100 - 200$ Hz. This is also the final rate to tape for offline analyses of the data. The global coordination of the DØ triggering is achieved via the main DØ run control package, `COOR`, on the online host system interfaced to the Level 3 data quality (DAQ) system.
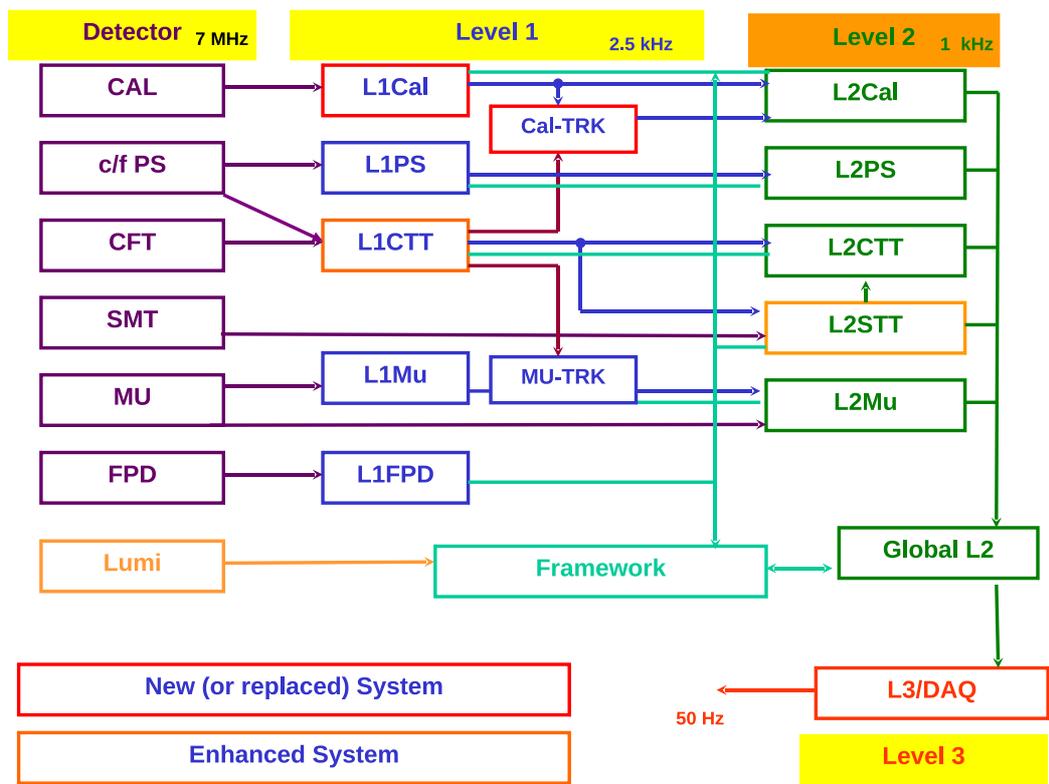
Figure 3.14: A block diagram for the DØ trigger system [28].

Chapter 4

Search Tools

4.1 Event Reconstruction

4.1.1 Tracks and Vertex Reconstruction

As charged particles pass through the central tracking system and generate hits in successive layers, tracks are formed by recognition of the hit patterns. Algorithms are developed to fit the hits to reconstruct the original particle trajectory with high precision. If the tracker volume is located within a magnetic field, the resulting curvature of the track for charged particles can provide for improved momentum measurements compared to calorimeter momentum at $p_T < 50$ GeV.

The primary vertex (PV) is defined as the interaction point of the $p\bar{p}$ collision, and therefore, is the originating point for tracks. Multiple reconstructed vertices may be found in a single event. The vertices are formed by clustering tracks with p_T above a threshold and additional requirements imposed on hits within the SMT. In order to separate the primary vertex from other vertices, an algorithm is applied to select the one corresponding to the high p_T physics interaction [29].

4.1.2 Electron Identification

Electrons interact electromagnetically so they are expected to deposit most of their energy in the EM calorimeter. Therefore, the reconstruction and identification of electrons are based on tracks from the central tracking system with a cluster of energy in the EM layers of the calorimeter [30].

The reconstruction of EM cluster uses the *CelNN* algorithm [31], which re-

constructs particle showers in events with multiple interactions. Here, *NN* stands for nearest-neighbour. The cone size for the EM clustering is defined with a radius $R = \sqrt{\eta^2 + \phi^2} = 0.2$. The *CelNN* algorithm begins by clustering nearby calorimeter cells in $\eta - \phi$ space within the signal cone and by forming "Floor-Clusters" in different layers. Figure 4.1 shows the definition of "neighbour" cells in electromagnetic and hadronic layers. Cells are only considered to be "neighbours" if they share a common tile in order to achieve the best angular separation. Because the granularity in the third EM layer (EM3) is finer relative to the other layers, each Floor-Cluster found in EM3 seeds a new global calorimeter cluster. The Floor-Clusters are then combined into global calorimeter clusters, and clusters in the preshower detectors are only added to the electron if they lie within a window of $\eta \times \phi = 0.05 \times 0.05$ of the centroid in EM3. If such a preshower cluster is matched, it will be used to help determine the direction of the electron momentum vector. The energy of an electron is measured using all EM and the hadronic layers.

THE 4 NEIGHBOURS OF A CELL   THE 8 NEIGHBOURS OF A CELL
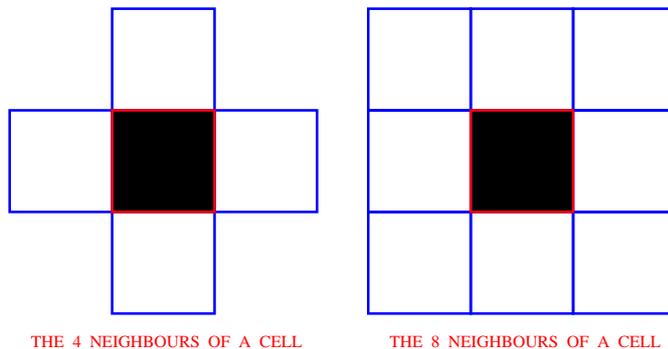
Figure 4.1: Neighbouring definitions for tiles in EM cells (left) and HAD cells (right) [31]

.

The shower shape for electrons or photons are expected to be narrow. This information is incorporated in the reconstruction algorithm to distinguish electrons from wider hadronic jets which are expected to have significant deposits extending outside the signal cone. Since photons do not carry charge, a matching to associated tracks found in the central tracker is applied to the reconstructed electrons. The primary variables used for electron identification are:

*Iso***:** Calorimeter Isolation variable. This is determined by

$$Iso = \frac{E_{tot}(R < 0.4) - E_{EM}(R < 0.2)}{E_{EM}(R < 0.2)}, \qquad (4.1)$$

where $E_{EM}(R < 0.2)$ is the energy deposited in the defined electron cone, and $E_{tot}(R < 0.4)$ is the total energy deposited in the isolation cone of a size $R = 0.4$.

f_{em}**:** Ratio of the EM cluster energy deposited in the EM calorimeter to the total energy including clusters in the Coarse hadronic layers.

*IsoHC*4**:** Track isolation for the EM cluster which is the total extrapolated transverse momentum of tracks in the hollow cone with $0.05 < R < 0.4$ around the EM cluster.

*HMx*7**:** Characterises lateral and longitudinal shapes of the EM showers

*Lhood*8**:** 8-variable electron likelihood. The resulting Lhood8 peaks at one for electrons and zero for jets.

χ^2**:** Probabilities of spatial track match to EM clusters.

p_T^e**:** Transverse momentum of electrons.

p_T^{trk}**:** Transverse momentum of tracks associated to electrons.

E/P_{trk}**:** The ratio of calorimeter energy to tracking momentum. The value for ideally reconstructed electrons should be one.

There are multiple *qualities* defined for electrons, such that looser quality always includes the tighter set. Table 4.1 lists the cuts for `top_tight` and `top_loose` electron definitions used in this analysis. The `top_loose` quality is applied in the sample skimming level to select a looser sample. This sample allows the studies of finer tuning of efficiency of the estimated background. The `top_tight` quality is required for the events passing final selections (see Chapter 8). The electron efficiency in the Run II has been studied in $Z \to e^+e^-$ events selected by *tag-and-probe* method which will be described in detail in Sect. 6.2.

Variables	top_loose	top_tight
$Iso \leq$	0.15	0.15
$f_{em} \geq$	0.9	0.9
$HMx7 \leq$	50	50
p_T^{trk} (GeV) $\geq$	5	5
p_T^e (GeV) $\geq$	15	15
$\chi^2 >$	0.0	0.0
$Lhood8 \geq$	–	0.85
$E/P_{trk} \leq$	2.5	2.5

Table 4.1: Selection cuts for electron definitions top_loose and top_tight.

4.1.3 Jet Identification

Hadrons that deposit energy as clusters in the calorimeter are collimated in the direction of the parent parton originating from a $p\bar{p}$ interaction These clusters are referred to as "jets", and are reconstructed from clustering calorimeter cells into a cone. Single cone algorithms with two radial sizes, $R = 0.5$ and $R = 0.7$, are used to combine the particles in DØ Run II [32]. The energy loss of the reconstructed jets due to insufficient energy collection using the cone algorithm is calibrated. The Jet Energy Scale (JES) [33]. is measured using a sample of photon plus jet data. The corrected jet energy, E_{jet}^{corr}, is given by

$$E_{jet}^{corr} = \frac{E_{jet}^{meas} - E_O}{R_{jet}S_{jet}}, \tag{4.2}$$

where E_{jet}^{meas} is the measured jet energy, E_O is the energy that is not associated with the hard scattering process which produced the jet, R_{jet} is the energy response of the calorimeter, and S_{jet} is the correction for the energy that is deposited outside of the jet cone. Other corrections are also applied to make more accurate measurements for the jet energy to make it closer to the original parton energy.

4.1.4 Missing Transverse Energy

The Missing Energy ($\not{E}_T$) is defined as the energy of particles not captured by the detector. This includes the energy of neutrinos which do not interact with any detector element, and the energy loss from the detectable particles. Because the

boost in the z direction is unknown, the calculation of missing energy can only be done in the transverse plane. The reconstruction of the $\not{E}_T$ uses a collection of the cell energy projected to the (x, y) plane. It is initially calculated using the raw cell energy. When the calorimeter objects such as electrons and jets are fully reconstructed, the clusters of energy are matched and replaced with the corrected object energies. The muons that are not measured in the calorimeter are separatedly added. This calculation assumes the conservation of momentum, and any momentum imbalance equates to $\not{E}_T$.

4.1.5   Muon Identification

Muons deposit only a small amount of ionisation energy in the calorimeter and travel through most of the inner detector components up to the muon system, typically the outer most radial layers of a detector. The identification of muons with the DØ detector [34] relies on the information provided by three independent subsystems:

- three-layer muon detector system,

- central tracking system,

- signature as a minimum ionising particle (MIP) in the calorimeter that relies on the "Muon Tracking in the Calorimeter (MTC)" algorithm

Local muons are muons detected by the muon detector up to $\eta < 2.0$. If central tracks are found to match the local muon successfully, this muon is defined as a "central track-matched muon". An n_{seg} variable is introduced to categorise the track-match muon types. A negative n_{seg} value is assigned for muons with no associate track. On the other hand, $n_{seg} \geq 0$ indicates that the muon has been matched to at least one central track. Subsequently, if a local muon is reconstructed by hits in all three layers, A, B and C, (see Sect. 3.2.5) muons will be assigned $|n_{seg}| = 3$, however, the requirement of BC layer hit is dropped for qualified $|n_{seg}| = 3$ muons if the matched central tracks exist. Correspondingly, values of $|n_{seg}| = 2$ indicate hits from B and C layers which are each located radially outside the magnet toroid, and $|n_{seg}| = 1$ defines a track-match muon with A-layer hit only. Muons with $|n_{seg}| = 0$ are matched to a central track with any logical "OR" of A, B or C layer hit. During the event selections,

tighter $|n_{seg}|$ requirements are assigned to improve the purity of identified muons and to reduce the number of mis-identified muons. In addition to the muon transverse momentum measured in the muon detector, the central tracker can provide accurate momentum resolution for the central track-matched muons.

A separate variable, *muon quality*, is defined for muon classification. There are three qualities defined for muons: *tight, medium* and *loose*. The definitions of each quality are based on the number of hits in the muon drift and scintillation chambers. For example, in this analysis, a `loose` $n_{seg} = 3$ muon is selected in the sample-skimming (preselection) level, and `medium` $n_{seg} = 3$ quality is required for the muons passing the final selections. The definition for `medium` $n_{seg} = 3$ is given below [34]:

- at least two A layer wire hits,

- an A layer scintillator hit,

- at least two BC layer wire hits,

- at least one BC layer scintillator hit (except for central muons where such a requirement is dropped).

A *loose* $|n_{seg}| = 3$ muon shares the same requirements as listed above but allowing the following two conditions [34]:

- one of the *medium* $n_{seg} = 3$ tests to fail,

- the A wire and scintillator are treated as one test which always requires at least one scintillator hit.

The typical reconstruction efficiencies are $\approx 89\%$ and $\approx 72\%$ of muons with `loose` and `medium` $n_{seg} = 3$ qualities, respectively [34]. Figure 4.2 shows the reconstruction efficiencies of `medium` $n_{seg} = 3$ muons as a function of time, and this quality is also used as one of the final selections in this analysis. Further details on selection requirements will be described in Chapter 8.

Moreover, in order to control the quality of central tracks matched to muons, four quality definitions `trackloose, tracknewmedium, trackmedium` and `tracktight` are defined and rely on [34]:

- number of hits in SMT or CFT system,
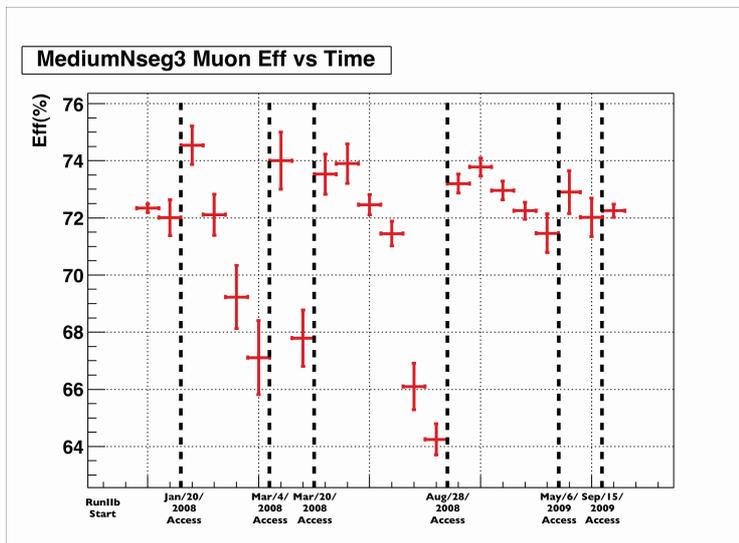
Figure 4.2: Offline reconstruction efficiencies for the `medium` quality muons with $n_{seg} = 3$ requirement. This figure is consistent with hardware repairs of the DØ detector [34].

- χ^2 per degrees of freedom of the central track fit (χ^2/dof),

- distance of closest approach with respect to the beam spot location (dca) in the (x, y) plane.

Similarly, `tracknewmedium` quality is required for muon candidates in this analysis; the definitions of track qualities are listed below [34]:

trackloose: track with $dca < 0.2$ cm, or $dca < 0.04$ cm if and only if the track has SMT hit.

trackmedium: *trackloose* track in addition to $\chi^2/dof < 4$.

tracknewmedium: *trackloose* track with $\chi^2/dof < 9.5$ and at least 2 CFT hits.

tracktight: *trackmedium* track contain at least one SMT hit.

Muon *isolation* variables are designed to separate signal muons from those produced in heavy flavour quark decays $b, c \rightarrow \mu + X$, which topologically tend to be embedded inside a jet. Isolation qualities are defined by a presence of tracks near the muon track or on the calorimeter clusters around the muon momentum vector. The relevant isolation quality for this thesis is *NPTight*. The sum of transverse energies of the cells, measured with respect to the beam position, in an annulus around the muon direction is used as the calorimeter isolation variable. For `NPTight` quality, it is required to be [34]:

$$\mathcal{I}_{cal} = \sum_{cells,i} E_T^i < 2.5 \text{ GeV} \quad \text{for } 0.1 < R < 0.4, \tag{4.3}$$

where ΔR is the distance in azimuth ϕ and pseudorapidity η between the calorimeter cluster and the muon direction. The isolation condition for the sum of the transverse momenta of all tracks within a cone of $\Delta R =0.5$ around the muon, excluding the muon track itself, is defined as [34]:

$$\mathcal{I}_{\mathrm{trk}} = \sum_{\mathrm{tracks},i} p_T^i < 2.5 \text{ GeV} \quad \text{for } 0.0 < R < 0.5. \qquad (4.4)$$

4.1.6 Hadronic Tau Identification
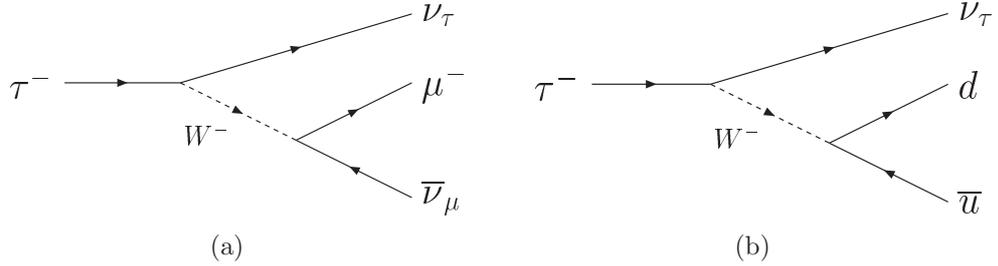
Figure 4.3: Feynman diagram for (a) a τ decaying leptonically to a muon and (b) a τ decaying hadronically into a pion.

Decay type	Tau Final State	Channel	BR (%)
Electron mode	e $+ \nu_e + \nu_\tau$	Leptonic	17.84
Muon mode	$\mu + \nu_\mu + \nu_\tau$		17.36
Hadronic type-1	$\pi(/K) + \nu_\tau$	Hadronic	11.59
Hadronic type-2	$\pi(/K) + \geq 1\pi^o + \nu_\tau$	1-prong (48.7%)	38.98
Hadronic type-3	$\pi\pi\pi + \geq 0\pi^o + \nu_\tau$	Hadronic 3-prong	14.23

Table 4.2: The products in the final state of main τ decay modes and the branching ratios (BRs) [5].

The τ lepton is the heaviest lepton and a mass of 1.78 GeV with a lifetime of the order of 2.9×10^{-13}s. The τ leptons are therefore expected to decay after a very short distance before reaching any element of the DØ detector. Shown in Fig. 4.3 are Feynman diagrams for a τ decaying to a muon and a pion, and Table 4.2 lists the primary branching ratio (BR) for τ decays to either leptonic or hadronic final states. As enumerated in Table 4.2, the total branching ratio of leptonic and hadronic τ decay channels are about 35% and 65%, respectively. Electron and muon identification methods as described in Sections 4.1.2 and 4.1.5 are

used to detect the leptonic τ decay products. The products of hadronic τ decays contain mostly neutral and charged pions, and therefore, the decay products mostly appear to be narrow jets with low track multiplicities, which require a dedicated τ identification algorithms. The hadronic τ reconstruction [35], [36] relies on the following elements:

Calorimeter cluster: Reconstructed using a single cone algorithm with a cone size of $R_{core} = 0.3$ in the r-ϕ plane, and an isolation cone size of $R_{iso} = 0.5$

EM subcluster: Reconstructed by the *CelNN* algorithm in the EM3 layer of the calorimeter(see Sect. 3.2.3 and Sect. 4.1.2). Subsequently, clusters that are found in EM layers other than EM3 are attached. The preshower of an EM subcluster is used to find neutral τ decay products such as those from π^0 particles where significant energy is deposited in the EM layers.

Associate tracks: Any τ candidate must match the calorimeter cluster with at least one track in the central tracking system. The track-matching algorithm requires track candidates to be sorted in increasing p_T within a cone with a size of $R = 0.3$ centred on the calorimeter cluster. In addition, the associated leading track must have $p_T > 1.5$ GeV, and up to two more tracks are considered if the the distance of these tracks to the leading track is less than 2 cm. A second track will be associated if the invariant mass of the first and the second track is less than 1.1 GeV. Similarly, a third track is added if the invariant mass calculated from the three tracks is less than 1.7 GeV and consistent with the mass of τ leptons. The sum of total charges of three tracks should also be either $+1$ or -1.

The final states of τ candidates can be split into three types categorised by the detector signature:

Type-1: With one associated track and one calorimeter cluster but no EM sub-cluster. This corresponds mainly to the physics decay channel $\tau^\pm \rightarrow \pi^\pm \nu$ (i.e., $\pi\nu$-like).

Type-2: With one associated track and one calorimeter cluster as for type-1, but at least one EM sub-cluster is required in addition. Type-2 corresponds mainly to decays with at least one neutral pion, such as $\tau^\pm \rightarrow \pi^\pm \pi^0 \nu$ (i.e., $\rho\nu$-like).

Type-3: With more than one associated tracks and a wide calorimeter cluster. No requirement is on the EM subcluster for type-3. This type corresponds mainly to the physics decay channel with products of three charged particles, such as, $\tau^\pm \to \pi^\pm \pi^\pm \pi^\mp (\pi^0) \nu$ (i.e., 3-prong).

The type-1 and type-2 belong to the "1-prong" τ decay mode while the type-3 is considered as "3-prong". The direction of hadronically decaying τ leptons is determined by the direction of the associated track with the τ for type-1 and type-2, whereas for type-3, the direction is taken as the vector sum of the momentum of the associated tracks for 3-prong decays. The classification is based only on the detector signature, therefore, no direct relationship to the physics channels are made, and the correlations listed above can be broken. The decay $\tau^\pm \to \pi^\pm \nu$, for example, should be reconstructed as τ type-1 based on the signature of one track and one calorimeter cluster without EM sections. However, if the charged pion showers early in the calorimeter leaving associated EM clusters, this decay will be reconstructed as τ type-2. Also, a $\tau^\pm \to \rho^\pm \nu \to \pi^\pm \pi^0 \nu$ decay will be reconstructed as type-1 instead of type-2, if the event is detected in the ICD region. Because there is only little EM coverage exists in this region, the EM energy can be lost if the τ lies in the ICD region.

The key to τ identification at DØ is the use of a multivariate technique. Neural Networks (NN_τ) are the default choice for the τ identification [37]. Three separate NN_τ are built, one for each τ type, to separate τ decays from the mis-identified jets. The NN_τs are trained using a simulated $Z \to \tau^+ \tau^-$ sample as signal and jets that recoil against non-isolated muons taken directly from data as background. Two NN_τs are trained separately for type-1 candidates for τ candidates detected in the ICD or non-ICD region. A set of discriminating variables between τ leptons and jets are chosen as the inputs to NN_τ training to give the best separation between the signal and the background. These variables are composed of isolation variables of calorimeter and tracker, shower shape parameters and correlations built between the calorimeter and track.

The output of NN_τ is a single value between zero and one, while value one is assigned for pure τ signals while zero is for jet backgrounds. The NN_τ applied currently for τ candidate selection was trained and tested in DØ Run IIa data. There are different NN_τ trained and tested with respect to Run IIa and Run IIb data sets, and the corresponding one should be applied to the physics analysis.

Most of the τ leptons reconstructed at DØ fell in the energy region of $20 < E_T^\tau < 40 GeV$. The signal acceptance of NN_τ for those taus with $|\eta^\tau| < 2.0$ is about 66% from $Z \to \tau^+\tau^-$ simulated events, while the rejection rate of jet background is kept as high as 98%. A correction to compensate the different efficiencies in data and MC while applying a NN_τ selection cut is required for the Run IIb data set. More details about the correction can be found in Sect. 6.4.

Due to the requirement on energy deposited in EM subcluster, the faking rate of electrons to type-2 τ candidates is high. In order to achieve a better separation between real electrons and type-2 τ candidates, an additional Neural Network (NN_{elec}) is trained and tested for τ type-2 using simulated $Z \to \tau^+\tau^-$ and $Z \to e^+e^-$ samples as input signal and background, respectively.

In order to control the quality of the τ tracks, three qualities `tracktight_tau`, `trackmedium_tau` and `trackloose_tau` are defined similarly to those defined for muon tracks (see Sect. 4.1.5). The τ track qualities rely on number of hits, χ^2 per degrees of freedom of the central track fit (χ^2/dof), and the distance of closest approach with respect to the beam spot location (dca) in the r-ϕ plane.

The main requirements for *trackloose_tau* are:

- $dca < 0.2$ cm if the track has no SMT hit,

- $dca < 0.04$ cm if the track has SMT hits.

Tracks with quality `trackmedium_tau` must fulfil `trackloose_tau` selections as well as an additional $\chi^2/dof < 4$ requirement. Finally, a `tracktight_tau` track must pass the requirements of a `trackmedium_tau` track but also contain at least one SMT hit.

4.2   Limit Setting Technique and Collie

Upper limits on the production cross sections for given physics processes are calculated using the `Collie` (Confidence Level Limit Evaluator) software package [38] based on the CL_s method [39]. It separately accepts ROOT [40] histograms for data and for background and signal models, and uses a negative Poisson Log-Likelihood Ratio (LLR) test statistic to evaluate the statistical significance. Two hypotheses are defined: the signal-plus-background (S+B) test hypothesis

and the background-only (B-only) null hypothesis. In this search, "S+B" corresponds to the production of MSSM Higgs bosons while "B-only" points to the SM background only prediction. The log-likelihood ratio, LLR, is evaluated by

$$\text{LLR} = 2 \sum_i \sum_j \left(s_{ij} - d_{ij} \ln(1 + \frac{s_{ij}}{b_{ij}}) \right), \tag{4.5}$$

where i is the sum over the ith channel analysed, j sums over the jth bin in the distribution, and d_{ij}, b_{ij}, s_{ij} are the numbers of data, background, and signal events observed or expected in that bin, respectively [38]. This process is repeated several times (about 10k) for both, "S+B" and "B-only", hypotheses, by generating data sets of pseudo-experiments via random Poisson trials. Each Poisson trial is seeded with a mean value taken from the sum of the contributing processes. For each pseudo-experiment, an LLR value is evaluated, and the frequency of relative outcomes defines confidence levels (CL). Figure 4.4 shows an example of negative LLR (NLLR) distributions for the "S+B" and "B-only" hypotheses, and the observed NLLR evaluated using data is also shown. Given the outcomes of LLR, one can calculate a confidence level for "S+B" (CL_{s+b}) and for "B-only" (CL_b) by integrating the shaded areas of the distributions as shown in Fig. 4.4. The value of the signal confidence level, CL_s, is then defined as [38]

$$CL_s = \frac{CL_{s+b}}{CL_b}. \tag{4.6}$$

and the exclusion condition is made such that $CL_s < \alpha$. The desired value of α used during the search is 5% with the condition that $CL = (1 - CL_s) > 95\%$ has to be reached to claim an exclusion. Although the number of background events is fixed, the LLR can be calculated for any cross section of the predicted signal. The observed limit is calculated by varying the cross section of the signal until the 95% CL is reached. The expected limit is calculated similarly by assuming the data is fully background-like and setting $d_{ij} = b_{ij}$; in other words, the median outcome from the "B-only" hypotheses is taken as the "expected data".

The sensitivity of the analysis is determined by both the peak and the width
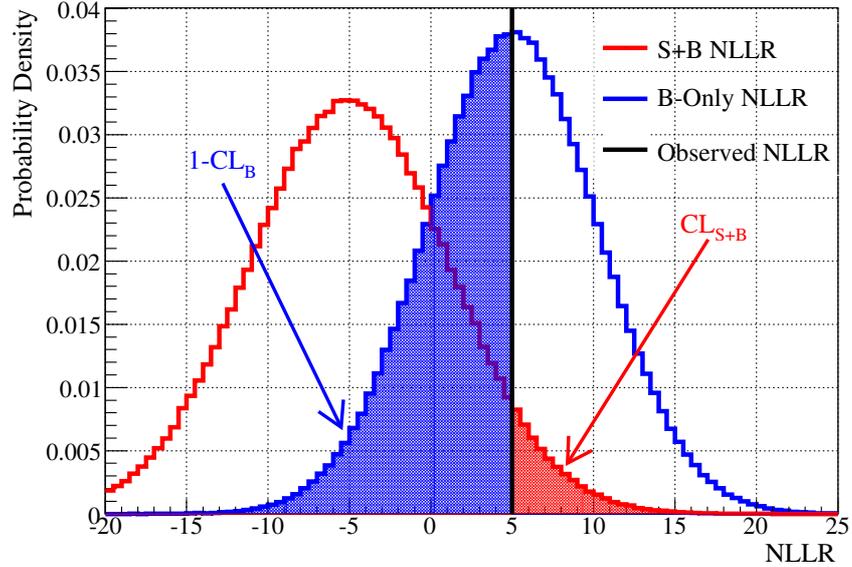
Figure 4.4: Negative LLR (NLLR) distributions for the expectation in the background-only hypothesis (blue line) and in the signal-plus-background hypothesis (red line) and the observation in the data (blue line). Also shown are the regions integrated to obtain CL_b and CL_{s+b} [38].

of the LLR distributions. Without considering any systematic effects for the measurement, the width of LLR distributions are determined by the Poisson uncertainty of the samples. Systematic uncertainties are incorporated by allowing events in each bin of a pseudo-experiment to vary according to a random number taken from a Gaussian distribution with a mean of zero and width equivalent to the size of the systematic uncertainty. Including the uncertainties widens the LLR distributions and makes the width proportional to the the quadrature sum of Poisson and systematic uncertainties. Figure 4.5 shows an example of LLR distributions before and after applying the Gaussian smearing of the systematic uncertainties. Although the means and the observed values of LLR remain the same, the corresponding width increases after applying Gaussian smearing. This effect makes it more difficult to distinguish the null and test hypothesis and therefore degrades the outcome of the limit calculation.

In order to reduce the degrading effects from including uncertainties, a maxi-
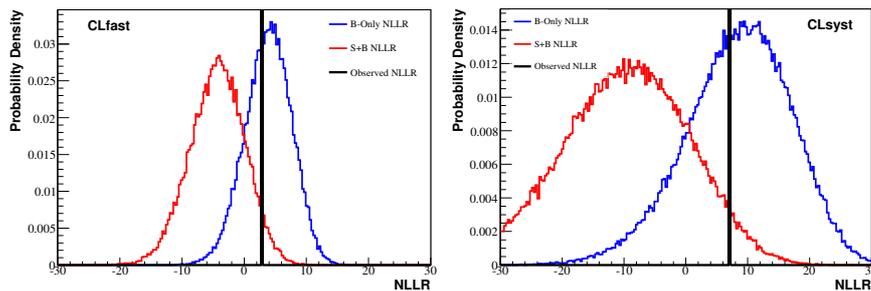
Figure 4.5: Negative LLR (NLLR) distributions for "B-only" (blue) and 'B+S' (red) hypothesis. The left figure shows LLR for the case of no systematics and the right one shows LLR after applying Gaussian smearing of the systematic uncertainties [38].

mum likelihood fit for the total prediction to data on the final variable distribution is introduced in `Collie`. The value of a Poisson χ^2 function is determined by the numbers of data (d_i) and total predicted (p_i) events by

$$\chi^2 = 2 \sum_{bin} \left((p_i - d_i) - d_i \ln \frac{p_i}{d_i} \right) + \sum_k S_k^2 \ [38], \tag{4.7}$$

where S_k^2 is the square of the deviations of the uncertainty from the nominal in units normalised by ± 1 standard deviation. Each background source is allowed to vary independently according to its uncertainties in order to minimise the χ^2 value to achieve the "best fit". More details concerning the techniques used in `Collie` along with concepts involved in hypothesis testing can be found in Ref. [38].

Chapter 5

Triggers and Data Samples

5.1 Trigger and Efficiency Measurements

The DØ trigger system consists of three trigger levels (Level 1, 2 and 3) (see Sect. 3.3). It helps to quickly select and record events of special interest to physics analysis. The relevant triggers used in the inclusive search for a neutral Higgs boson decaying to two τ leptons are single muon triggers for the $\tau_\mu \tau_h$ channel and single electron triggers for the $\tau_e \tau_h$ channel. These triggers select muon or electron candidates, and are designed in different trigger epochs known as "trigger lists" corresponding to running time periods. However, no reliable simulations are available for effects of the trigger on MC events. Therefore, it is very important to measure the corrections for trigger efficiencies in data and correct at the MC level. A *tag-and-probe* technique [41] is used at DØ to derive the efficiencies from data by considering leptonic decays of the Z boson. A schematic diagram of this method is shown in Fig. 5.1. The method starts by selecting a di-lepton sample from the process $p\bar{p} \to Z \to l^+ l^-$, where $l = e, \mu$, and requires two leptons. Both of the selected leptons can be considered as the "tag" lepton if they have matched all requirements for the "tag" leptons. The other corresponding lepton in the same event is the "probe" lepton, it is required to pass the trigger terms at each relevant trigger levels. Assuming n_p is the number of probe leptons which match all relevant trigger terms at each trigger level, the resulting efficiency, ϵ, is calculated by

$$\epsilon = \frac{n_p}{n_T},$$

(5.1)

where n_T is the total number of tag leptons. This efficiency is taken from the logical "OR" of each trigger and applied to MC events. It is generally parameterised in terms of the most sensitive variables (e.g. p_T or detector pseudorapidity).
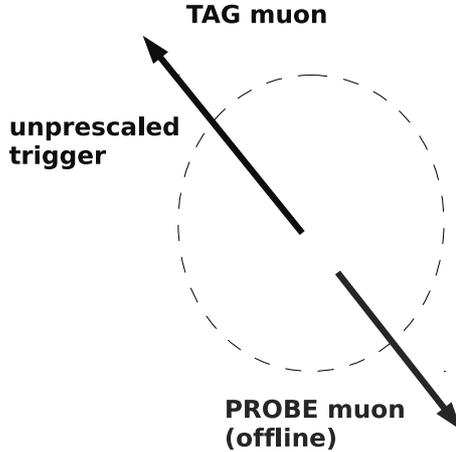
(a)

Figure 5.1: Schematic view of the tag-and-probe method for muons. Here, a reconstructed probe muon is found to match a tagged muon triggered by the muon triggers [42]. Similar diagrams exist for cases with other leptons.

It is necessary to split the efficiencies into different trigger epochs (periods) due to run configurations, changing instantaneous luminosities, hardware or software changes. The muon triggers are discussed in greater details in Sect. 5.2. Further, Sect. 5.3 describes the single electron triggers.

5.2   Muon Triggers

Muon triggers are composed of different trigger terms dedicated to muon or track detection at the three different trigger levels as listed in Table 5.1. The muon trigger terms in different levels will be discussed separately in the following subsections.

Data events are required to pass one of the muon triggers. The running periods of those muon triggers are listed in Table 5.2. In addition, a matching of the selected muon to a Level 2 triggered muon is done for the data events, and this criterion removes about 4% of the data events (see Ref.[43] for complete

Trigger Level	Trigger Term(s)
Level 1 (L1)	L1 muon scintillator terms L1 muon wire terms L1 track terms
Level 2 (L2)	L2 muon terms
Level 3 (L1)	Local muons (reconstructed only in the muon system) terms L3 tracking terms L3 local muons matched to L3 tracks

Table 5.1: Muon trigger terms used in different trigger level (L1, L2, L3).

details). The selected data sample contains only events from a logical "OR" of the 19 single muon triggers.

	Trigger List	Run Dates (Run II)
Run IIa	v8 - v10.3	2002 JUL - 2003 FEB
	v10.3 - v13	2003 FEB - 2004 JUN
	v13 - v14	2004 JUN - 2005 AUG
	v14 - v14.6	2005 JUN - 2005 OCT
	v14.6 - v15	2005 OCT - 2006 FEB
Run IIb	v15.00 - v16.22	2006 JUN - 2009 SEP

Table 5.2: Muon triggers used in Run IIa and Run IIb analysis. The table is split into the different trigger list for run dates in Run II.

The muon trigger efficiency is measured by selecting $Z \to \mu^+\mu^-$ tag-and-probe events, and without a requirement on the invariant mass of this two muon sample [42]. The following selections are required for the tag muon in the measuring sample:

- loose muon quality (see Sect. 4.1.5) ;

- A-layer scintillator |time| < 7 ns (use timing from B-layer if no A-scintillator hit is found);

- matched to a central track with medium quality;

- muon transverse momentum, $p_T^\mu > 30$ GeV;

- isolation requirement for tracks, $\mathcal{I}_{\text{trk}} < 3.5$ GeV, and for calorimeter clusters, $\mathcal{I}_{\text{cal}} < 2.5$ GeV;

- matched to all the trigger terms of at least one single muon trigger.

Definitions of the variables listed above can be found in Sect. 4.1.5. The probe muon must pass at least loose quality and match to a loose quality track with $p_T > 20$ GeV. Muon trigger efficiencies measured in the Run IIa data are parameterised in terms of η and ϕ. The dependency on instantaneous luminosity has been introduced in Run IIb, and the efficiency corrections for MC are therefore parameterised in terms of η, ϕ and instantaneous luminosity.

5.2.1 Level 1 Muon Trigger (L1MU)

A Level 1 muon object is formed by tracks from the central fiber tracker (CFT), and by scintillator and wire hits from the muon system with an acceptance rate of 10 kHz. The criteria on the Level 1 muon triggers are:

- matching the p_T information measured by the CFT tracks with hits in the muon system;

- using patterns of wire and scintillator hits from the muon detectors: a low p_T trigger uses only the innermost tracking stubs, whereas a high p_T trigger uses correlations between stubs in more than one layers;

- imposing a timing gate around the muon arrival time at the scintillator planes to suppress background from cosmic rays.

At this level, two regions are defined in terms of detector pseudorapidity (η) of the trigger system: the "wide" and "all" regions. The "all" region is defined as $|\eta| < 2.0$, whereas the definition for the "wide" region changes from $|\eta| < 1.5$ to $|\eta| < 1.6$ for the triggers used after June 2004 corresponding to trigger periods v13 and higher. Therefore, the efficiencies for the wide region are split into two periods, pre- and post-v13 trigger epoch. *Prescaling* at Level 1 is used to filter events and allow only a electronically acceptable number of events to pass the triggers. Not all triggers are prescaled, but this effect has to be taken into account in order to calculate the integrated luminosity and efficiency correctly. Therefore, the unprescaled high transverse momentum single jet trigger is a general choice for the luminosity calculation. Since the pre-scale of triggers are allowed to vary in different trigger lists, the provided "OR" trigger efficiency should only describe data which exactly contains the same runs as those used for

the efficiency calculation. The average efficiency for the Level 1 (L1) loose wire term, relative to the L1 tight scintillator term in this sample is about 95%.

5.2.2   Level 2 Muon Trigger (L2MU)

L2MU accepts L1 inputs and uses calibration and more precise timing information. Muon qualities are used to classify muons as either loose, medium or tight at this level. The p_T above a given threshold is measured with the muon system only. Most of the muon triggers use medium muons at Level 2, with either a p_T threshold of 0, 3 or 5 GeV. The overall acceptance rate in L2MU is reduced by approximately a factor of ten with respect to the L1MU outputs.

5.2.3   Level 3 Muon Trigger (L3MU)

The Level 3 muon triggers are only used for trigger lists post-v13 using more information provided from drift chambers, calorimeter, and SMT. They also include more analytical calculations on CFT tracks and PDT and A-layer hits with the calibrated data. Level 3 is different from Level 1 and 2 by implementing software based programs such as `C++` methods to make the decisions whether the event will be kept or rejected. Therefore, not only the resolution but the rejection rate is improved with respect to the L2MU output.

5.3   Run IIa Electron Triggers

In the $\tau_e \tau_h$ channel, tight electrons selected from data are required to pass at least one of the unprescaled electron triggers in the logical "OR", at all three levels of the trigger. To select events with a particular single electron trigger, the reconstructed electron must match a Level 1 trigger tower above a given p_T^e threshold within $\Delta R < 0.4$. If any Level 2 and 3 objects are found, the same match requirement is applied to these higher level objects. In addition to the transverse momentum threshold, the EM calorimeter cluster is required to be isolated at Level 2, and Level 3 applies a shower shape requirement in addition [44]. Table 5.3 lists the run dates for each trigger epoch.

Single electron efficiencies of Run IIa data are measured using a tag-and-probe method $Z \to e^+ e^-$ sample. Tracking information is introduced for efficiencies

Trigger Epoch(s)	Run Dates (IIa)
v8 - v12	2002 Jul - 2003 Jul
v12 - v13	2003 May - 2004 Jun
v13 - v13.11	2004 Jun - 2004 Aug
v13.11 - v14	2004 Aug - 2005 Aug
v14 - v15	2005 Jun - 2006 Feb

Table 5.3: List of run dates of electron triggers for each trigger epoch.

measurements in order to gain improvements at low p_T (15 GeV $< p_T < 25$ GeV) compared to the electron trigger efficiency based on calorimeter only information [45]. The tag electron is required to pass the following criteria:

- Ratio of the EM cluster energy deposited in the EM calorimeter to the total energy, $f_{em} > 0.9$.

- Calorimeter isolation, $Iso < 0.2$.

- Parameter for lateral shapes of the EM showers, $HMx7 < 50$.

- Spatial track match, $\chi^2 > 0$.

- Electron likelihood, $Lhood8 > 0.85$.

- Electron transverse momentum, $p_T^e > 15$ GeV.

- Transverse momentum of associated tracks, $p_T^{trk} > 10$ GeV.

- Match a single-electron trigger tower within $\Delta R < 0.4$.

The definitions of these variables are given in Sect. 4.1.2. The probe electron is required to have $p_T > 15$ GeV and must pass either loose or tight selections depending on the electron identification used. A mass window of $70 < M_Z < 110$ GeV is applied to the samples used for trigger efficiencies measurements in order to reduce the background contamination from multijet events. The average trigger efficiency applied to signal MC at the preselection stage varies from 74% for a Higgs boson of 90 GeV mass to 87% for a Higgs boson mass of 200 GeV.

5.4 Data Samples

The analysis presented in this thesis is based on the data set up to DØ Run IIb-2 collected between April 2002 and June 2009 by the DØ detector and corresponds to physics runs taken within the range 151817 to 215670 for Run IIa data, and 221698 to 252918 Run IIb data. The analysis uses MUinclusive and EMinclusive Common Analysis Format (CAF) Tree skims provided by the DØ Common Sample Group [46] for the $\tau_\mu\tau_h$ and $\tau_e\tau_h$ channels, respectively.

5.4.1 MUinclusive Skims

The MUinclusive skim accepts events that fulfil a range of selections. The most important for this analysis are a logical "OR" of the following:

- At least one muon of quality "loose" as defined in Sect. 4.1.5 with $p_T^\mu > 8$ GeV.

- At least one "loose" muon with $p_T^\mu > 5$ GeV and at least two tracks with $p_T^{trk} > 5$ GeV and $p_T^{trk} > 8$ GeV, respectively, and one τ candidate with $p_T^\tau > 5$ GeV.

The muon transverse momentum, p_T^μ, is measured using the central tracking system whereas the τ transverse momentum p_T^τ is measured using the calorimeter except for τ type-1. Details of the τ reconstruction algorithm can be found in [35].

The Run IIa MUinclusive dataset is divided into five trigger epochs from trigger list v8 to v14.99 while the Run IIb MUinclusive into six periods corresponding to the v15 and v16 trigger lists. The total integrated luminosity for the Run IIa data used for the analysis is 1.08 fb^{-1} and is 4.28 fb^{-1} for the Run IIb data. Table 5.4 provides a breakdown of the integrated luminosity during each period. Therefore, a total of 5.36 fb^{-1} of Run IIa plus Run IIb data is used in the $\tau_\mu\tau_h$ search channel. The total uncertainty on the integrated luminosity is about 6.1% [47].

5.4.2 EMinclusive skims

The EMinclusive skim accepts events that fulfil a range of selections. The most important for the Higgs search presented in this thesis are a logical "OR" of the following:

Trigger Epoch	Luminosity [pb^{-1}]
Run IIa Total	1014.3
v8 - v12	72.0
v12 - v12.99	226.1
v13 - v13.2	31.8
v13.2 - v13.99	327.4
v14 - v14.6	142.1
v14.6 - v14.99	190.6
Run IIb Total	4281.7
v15.00 - v15.22	227.2
v15.22 - v15.50	307.3
v15.50 - v15.60	232.5
v15.60 - v15.90	461.3
v15.91 - v16.00	397.3
v16.00 - v16.22	2656.1

Table 5.4: Recorded integrated luminosity after applying data quality requirements given as a function of trigger list versions.

Trigger Epoch(s)	Luminosity [pb^{-1}]
v8 - v11.34	106.9
v12 - v12.37	231.0
v13 - v13.90	378.2
v14 - v14.93	333.8
Total	1049.9

Table 5.5: Integrated luminosity of Run IIa data for the single electron triggers in the different trigger list versions. The integrated luminosities are given after the application of the data quality selection.

- At least one reconstructed EM object, assumed to be the electron, with $p_T^e > 20$ GeV.

- At least one reconstructed EM object, assumed to be the electron, with $p_T^e > 8$ GeV associated with a track with $p_T^{trk} > 5$ GeV.

The values of p_T^e used here are measured using the calorimeter. Table 5.5 provides a breakdown of the integrated luminosity for the Run IIa data during each electron trigger period. The total integrated luminosity used for the $\tau_e \tau_h$ Higgs search is 1.08 fb^{-1}.

Chapter 6

Monte Carlo Simulation and Corrections

6.1 Monte Carlo Simulation Method

In order to correctly model physics background or any potential signal events in the data, Monte Carlo (MC) methods are used in high energy physics analyses such as the one described in this thesis. Figure 6.1 provides an example flowchart for the event-generation chain. The two main steps in DØ to generate MC events are:

- simulate physics processes at either leading order (LO), next-to-LO (NLO), or next-to-NLO (NNLO) using MC generators, such as PYTHIA 6.3 [49] or ALPGEN 1.2 [50];

- perform a simulation of particle interactions with DØ detector components.

During step 1, a general-purpose generator such as PYTHIA takes into account the following main steps sequentially to build up the structure of an event:

- simulate collisions between two initial partons in hard processes;

- create decays of short-lived "resonances" produced in hard processes;

- associate emissions with incoming or outgoing partons;

- form multiple interactions by further parton pairs from the original hadrons;

- perform hadronisation process using non-perturbative models;

- simulate decay products except for b quark and τ lepton decays.

The simulation of the τ lepton decays is done using the TAUOLA [51] package. It provides not only the correct branching ratios for the τ decays, but also accounts for the polarisation of the τ leptons that affects the momentum distribution of the decay products.

PYTHIA simulates events based on leading order (LO) matrix elements calculated from the Lagrangian of a theory with *parton-shower* approaches, which means that the underlying physics process is produced from only the tree level Feynman diagram for that process. *Parton showering* is the model of the radiative evolution of partons without explicit matrix element calculations. It begins at the energy scale of the hard process and evolves down to a lower scale until the showering ends at an energy cutoff point. The momentum distribution of partons is described through the parton distribution functions (PDFs) by using CTEQ6.1M [52]. The alternate MC generator used at DØ, ALPGEN, is a parton-level event generator and provides a better modelling of multi-parton hard processes with respect to PYTHIA. Because the evolution of the partonic final state is not done in ALPGEN, it is interfaced to PYTHIA to perform the showering and hadronisation.
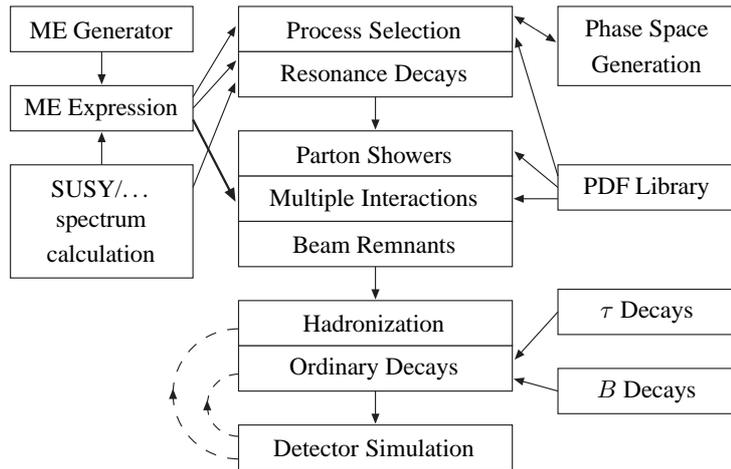

Figure 6.1: An example flowchart for the event generation chain used at hadron collider experiment. ME describes the Matrix Element method while PDFs represents the parton distribution functions [48].

The simulation of the DØ detector is done by a GEANT3-based software, `D0gstar` [54], and the program `D0Sim`. The decision of how much energy should be deposited in the active areas of the detector is made by `D0gstar`; it simulates the energy response. The `D0Sim` software is used along with `D0gstar` to simulate the digitisation of the signal outputs from different detector layers. Zero Bias (ZB) data events from elastic and/or diffractive interactions at low p_T, which are collected by special triggers, can be overlayed to the simulated events at this stage. `D0Raw2Sim` is a software to convert the ZB events from data to the same format as the outputs from `D0gstar`. These converted events are subsequently taken by `D0Sim` as inputs.

The DØ offline reconstruction program (`D0Reco`) is used for object reconstruction and processes both online data and MC simulated events. It is a CPU intensive program and runs on the production farms. There are two output formats from `D0Reco`, Data Summary Tape (DST) and Thumbnail (TMB) [56]. DSTs contain all necessary information to perform physics analysis with $\approx$ 150kb event size, while TMBs are summaries of DST with only $\approx$ 15kb event size and they are commonly used in DØ analysis. Both, TMB outputs of MC as well as data events, are transformed into Common Analysis Format (CAF) format, known as "CAF Trees" for physics analysis [57]. CAF Trees are outputs in a object-oriented and ROOT based software framework used in DØ and contain the basic information on the properties and kinematics of the event. Since the simulation output from `D0Reco` does not produce a completely accurate representation of the data, CAF provides tools to tune efficiencies or energy scale for MC events using the measurements from data. There are several standard corrections used in this analysis, which are applied via the CAF framework to the MC events:

Longitudinal beam position z: Due to small shifts in the magnets used to focus the p and $\bar{p}$ beams, the data and MC are not well matched in the distributions of z coordinate of the beam spot during different time periods. Therefore, the true MC distribution of the beam spot is re-weighted in the longitudinal direction in terms of instantaneous luminosity and different data taking periods [58].

Instantaneous Luminosity: The instantaneous luminosity profile used in the MC does not describe the data taken in the detector. Therefore, a re-weighting is performed in the MC to ensure that the instantaneous luminosity distribution of

71

MC events matches data. In Run IIb analysis, this re-weighting is more significant since the data is taken from a wider range of instantaneous luminosity values.

p_T **distribution of** W **and** Z **bosons:** Because higher order effects are missing in the samples generated from `PYTHIA` or `ALPGEN`, the vector boson transverse momentum ($p_T^{Z/W}$) is not represented correctly in MC. Since the transverse momentum of the Z bosons, p_T^Z, from unfolded data agrees with the prediction provided by the `ResBos` MC generator [59] in the 0-30 GeV range, a re-weighting formula in terms of p_T^Z is derived directly from the `ResBos` MC generator and extrapolated to the higher p_T^Z region [60]. Since there is no direct measurement for p_T^W available due to the presence of neutrinos in the final state, the p_T^Z re-weighting is applied and adjusted by the ratio of W to Z cross sections at next leading order (NLO) and next to next leading order (NNLO) [61].

Other corrections are required, such as the trigger efficiency correction described in Chapter 5 and the efficiency corrections of lepton identification discussed in the following sub-sections.

6.2   Electron Efficiency Correction

The reconstruction efficiency is generally overestimated by the simulation for physical object such as electrons, and therefore, corrections measured using independent and pure sample of *tag-and-probe* electrons from $Z \rightarrow e^+e^-$ events [30] are required for MC. If ϵ_{MC} and ϵ_{data} are electron efficiencies of data and MC, respectively, the efficiency correction factor f_e is given by

$$f_e = \frac{\epsilon_{data}}{\epsilon_{MC}}. \tag{6.1}$$

Figures 6.2 and 6.3 show the dependencies of the electron efficiency correction factors, f_e, on p_T^e and detector pseudorapidity, η^e. The correction factors needed for MC vary for electrons detected in different parts of the detector, but the p_T^e dependence is small compared to the dependence on η^e. A two-dimensional correction depending on both p_T^e and η^e is applied in the analysis.

Furthermore, the selections for the $\tau_e\tau_h$ channel reject events if they contain a second electron passing the tight electron selection (see Sect. 8.1). Assuming $N_{MC}^{2-tight}$ is the number of MC events with two electrons reconstructed as "tight",
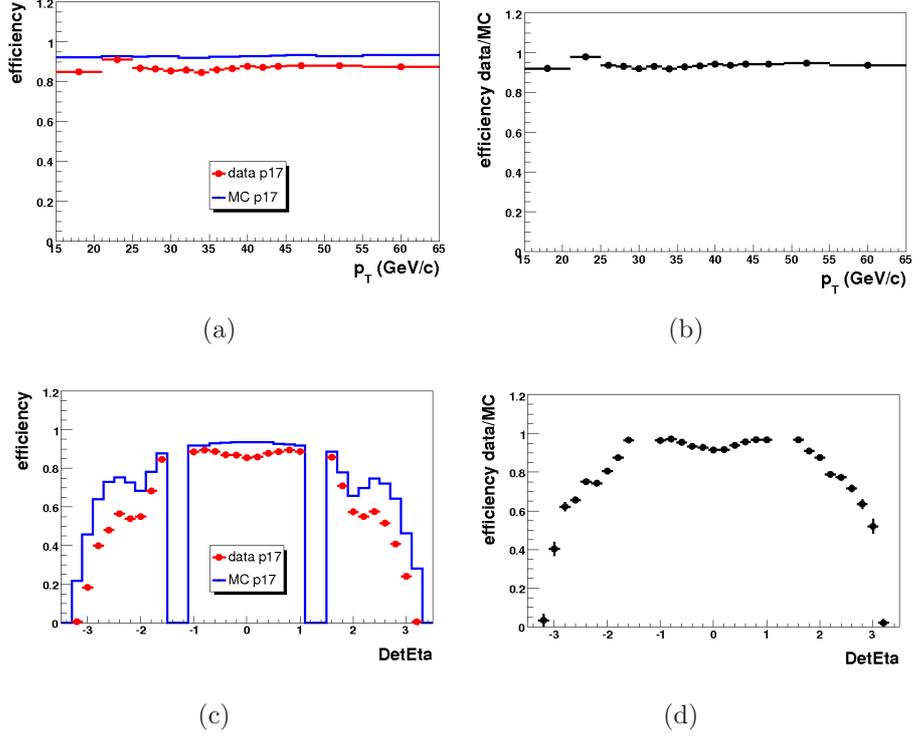
Figure 6.2: Figures (a) and (c) show the loose electron efficiencies measured in Run IIa data (red dots) and MC (blue line) as a function of p_T^e and η^e, respectively. The correction factors f_e derived for both cases are also shown in (b) and (d).

and $N_{MC}^{1-tight}$ is the number of events with only one electron reconstructed as "tight", the number of MC events with only one tight electron, $N_{MC}^{1-tight,corr}$, should be given by

$$N_{MC}^{1-tight,corr} = f_e N_{MC}^{1-tight} + 2N_{MC}^{2-tight} f_e \left(\frac{1 - \epsilon_{data}}{\epsilon_{MC}} \right). \qquad (6.2)$$

The first term, $f_e N_{MC}^{1-tight}$, describes the "one-tight-electron" data sample which is also the number of estimated events given by standard efficiency correction for MC. The second term, $2N_{MC}^{2-tight} f_e(\frac{1-\epsilon_{data}}{\epsilon_{MC}})$, is the contribution from the "two-tight-electron sample" caused by the inefficiency in data with one of the two tight electrons being mis-identified as "not tight". Because the standard efficiency corrections are derived in the $Z \to e^+e^-$ sample with two electrons reconstructed as tight, the factors shown in Fig. 6.3 do not take into account the inefficiency term. This is readily observed in the invariant mass distributions of electron and τ pair $(M(e, \tau))$ taken from the $\tau_e\tau_h$ channel after preselection that are shown in Fig. 6.4. These distributions indicate that an additional factor is needed in
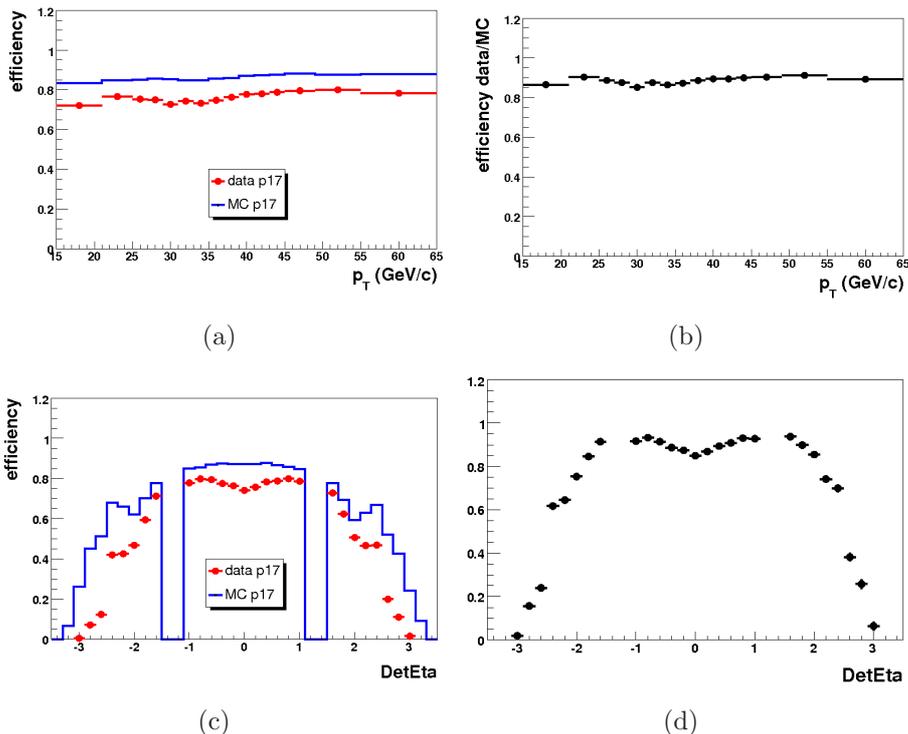
Figure 6.3: Figures (a) and (c) show the tight electron efficiencies measured in Run IIa data (red dots) and MC (blue line) as a function of p_T^e and η^e, respectively. The correction factors f_e derived for both cases are also shown in (b) and (d).

addition to the standard correction. Therefore, electrons in the $Z \to e^+e^-$ MC sample should be corrected by a scaling factor, k:

$$k = \frac{f_e N_{MC}^{1-tight} + 2N_{MC}^{2-tight} f_e\left(\frac{1-\epsilon_{data}}{\epsilon_{MC}}\right)}{f_e N_{MC}^{1-tight}}, \tag{6.3}$$

which has an estimated value of $k \approx 1.26$.

In this analysis, the scaling factor k is calculated as a function of a kinematic variable, X, by subtracting the bin value $N(X)$ of the histogram of the real data and Monte Carlo samples by

$$k(X) = \frac{N(X)_{data}^{Zee}}{N(X)_{MC}^{Zee}}, \tag{6.4}$$

where $N(X)_{data}^{Zee}$ is the total number of data events after subtracting the predicted background except the $Z \to e^+e^-$ MC, and $N(X)_{MC}^{Zee}$ is the number of real MC electrons. Events for estimating $k(X)$ are selected within a mass window $60 < M(e,e) < 130$ GeV. This measurement as a function of NN_e provides an
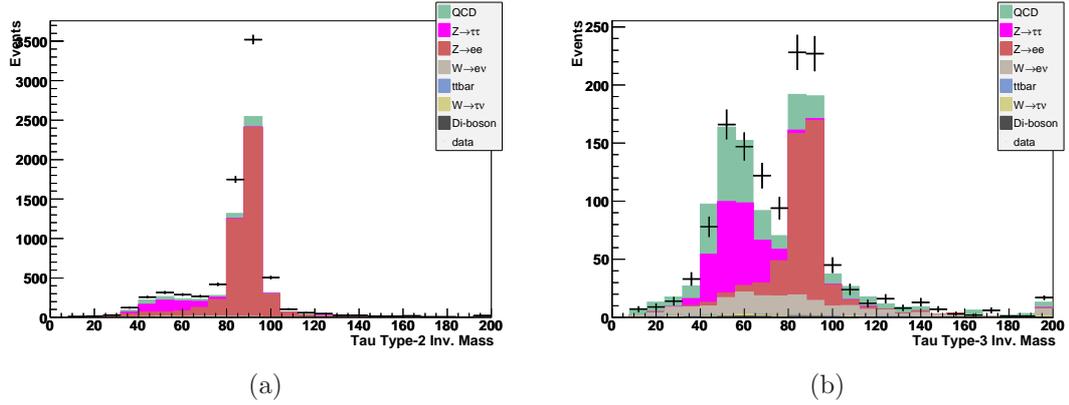
Figure 6.4: Distributions of $M(e, \tau)$ for (a) type-2 and (b) type-3 τ candidates with standard efficiency correction factors applied.

estimation for the number of events expected for the "only-one-tight-electron" data sample. After a fit is performed on the values of $k(X)$ in each bin, the result is very close to the estimated value. This correction is derived in terms of NN_e (see Sect. 4.1.6) and $M(e, \tau)$ for τ type-2 and type-3 events, respectively. The dependence on NN_e of this correction for τ type-2 also compensates for the mis-modelling in this variable for MC. The masses of type-2 and type-3 τ-e pairs after the correction are shown in Fig. 6.5, and the predicted background describes data well.
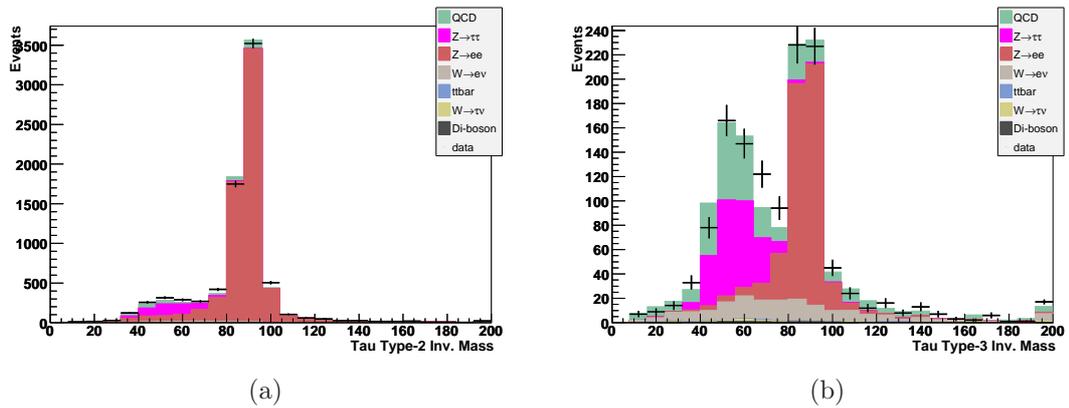

Figure 6.5: Distributions of $M(e, \tau)$ for (a) type-2 and (b) type-3 τ candidates after the additional electron correction performed.

6.3 Muon Efficiency Correction

The tag-and-probe method is also used to determine the efficiency in the high p_T^μ ($p_T^\mu > 30$ GeV for tag muons) $Z \to \mu^+\mu^-$ sample for the reconstruction of muon objects and central tracks matched to muons. The difference in tracking efficiencies between the data and MC is due to the fact that MC simulation overestimates the probability for tracks to have SMT or CFT hits associated with them. Therefore, the three corrections are applied on an muon-by-muon basis [34]:

Muon identification efficiency which is parametrised in the muon $\eta \times \phi$ plane.

Muon tracking efficiency which is parametrised in terms of the absolute pseudorapidity, $|\eta_{CFT}|$, and position , z_0, both measured at the outer boundary of the CFT, and the instantaneous luminosity, $\mathcal{L}$.

muon isolation efficiency which is parametrised in terms of $|\eta_{CFT}|$ and $\mathcal{L}$ for $\Delta R(\mu, jet) > 0.5$ measured with respect to the closest jet, or in $|\eta_{CFT}|$, p_T^μ and ΔR.

The average muon identification efficiency is 91.2% for *loose* muons and 73.6% for *medium* muons with $|n_{seg}| = 3$. The dependence on $\mathcal{L}$, ϕ^μ, η^μ and $|\eta_{CFT}|$ for *medium* muons with $|n_{seg}| = 3$ are shown separately in Fig. 6.6. The efficiencies are 90.8% of associated tracks of *mediumnewtrack* quality and 97.2% of muons with *NPTight* isolation quality.

The statistical uncertainties on these corrections are negligible due to the large number of Z boson events used to determine the efficiency corrections. Sources for systematic uncertainties on muon corrections include tag-and-probe biases, background contamination, remaining instantaneous luminosity dependence, correlations between different corrections, finite bin size and choices of variables necessary to parameterise the efficiency. These lead to a quadrative sum of total uncertainties of 1.2% $\oplus$ 1.0% $\oplus$ 0.9% = 1.86% for *medium* selected muons with $|nseg| = 3$, *NPTight* isolation, and *tracknewmedium* tracking requirements [34].
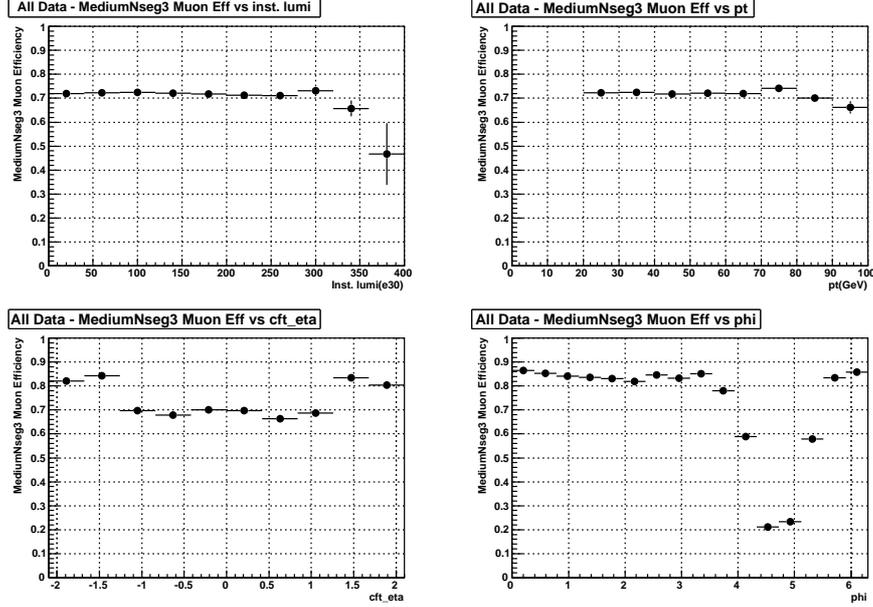
Figure 6.6: Muon reconstruction efficiencies of *medium* quality muons with $|n_{seg}| = 3$ versus instantaneous luminosity $\mathcal{L}$ (top left), p_T^μ (top right), $|\eta_{CFT}|$ (bottom left), and muon ϕ (bottom right).

6.4 Tau Corrections

6.4.1 Tau Efficiency Correction

The efficiency corrections for tracks associated to the hadronic τ leptons, τ_h, are taken from the measurements for muon tracks as described earlier in Sect. 6.3. The same track quality definition can also be used for tracks associated to τ_h candidates, and the relevant one used in this analysis is *tau_trackloose* which has the same requirements as *trackloose* defined for muon tracks. Figure 6.7 shows the efficiency of data and MC having a *tau_trackloose* requirement on the associated tracks as a function of instantaneous luminosity $\mathcal{L}$. The average correction applied to τ_h candidates is 91.7% for typical instantaneous luminosities during Run II.

Because the quality of τ candidates in the simulation is better than for real data, the peak in NN_τ distribution of MC is narrower and with better resolution than that of data. This will cause poor modelling of the efficiency for MC while cutting on NN_τ to select well reconstructed τ leptons. Therefore, a re-normalisation of the NN_τ variable based on a bin-by-bin re-weighting method measured in a $Z \rightarrow \tau^+\tau^-$ sample is performed in addition to the track efficiency
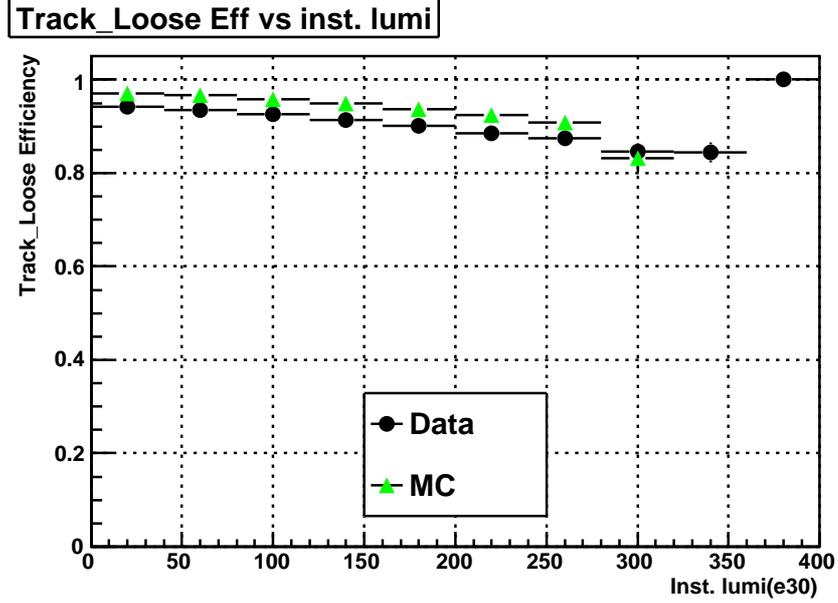
Figure 6.7: The efficiency of applying `trackloose` quality cut to data and MC as a function of $\mathcal{L} \times 10^{30}$ cm^{-2}s^{-1}.

correction. In order to derive this correction, a probability distribution function, is defined as P_{MC} for MC and P_{data} for data. For a given bin X of the histogram which contains the events with NN_τ values in a certain range, the NN_τ probability distribution function for MC is given by

$$P_{MC}(X) = \frac{N_{MC}(X)}{N_{MC}} \tag{6.5}$$

where N_{MC} is the total number of selected MC events and $N_{MC}(X)$ is the number of MC events in bin X. Similarly, for data we obtain

$$P_{data}(X) = \frac{N_{data}^{NoZee}(X)}{N_{data}^{NoZee}} \tag{6.6}$$

, where N_{data}^{NoZee} is the total number of data events after all estimated backgrounds except $Z \rightarrow e^+e^-$ are subtracted, and $N_{data}^{NoZee}(X)$ provides the number of data events in bin X after the same subtraction. The correction factors for a particular NN_τ bin X is then given by

$$f_{NN}(X) = \frac{P_{data}(X)}{P_{MC}(X)}. \tag{6.7}$$

78

This provides a good way to probe the differences of data and MC events in each given NN_τ bin. Figure 6.8 shows the correction factor $f_{NN}(X)$ broken down by each τ type as a function of NN_τ bins. For those bins with medium NN_τ values, the correction factors vary a lot because the statistics of the sample is low. These uncertainties are considered in the analysis which the hadronic τ final states are involved. The correction factors for high NN_τ bins where most of the τ leptons peak are derived in a sample of much larger statistics. Here, one does not expect large correction factors on the efficiency. Since events peaking at low NN_τ are mostly from multijet backgrounds and are not populated by real τ candidates, such a correction is only available for τ type-1, type-2 events with $NN_\tau > 0.3$ and for type-3 events with $NN_\tau > 0.4$. After the re-normalisation of MC events, one can calculate the efficiency $\epsilon_{NN}(Y)$ with a cut $NN_\tau > Y$ applied, where Y is an arbitrary non-zero choice corresponding to the selection cut used in the analysis. The systematic uncertainty σ_{NN}^{tot} is then derived by three steps. In step 1, the uncertainty caused by different choices of the values of Y, σ_{NN}^{Y} is given by:

$$\sigma_{NN}^{Y} = (1 - \epsilon_{NN}(Y))/3. \tag{6.8}$$

Further in step 2, the systematic uncertainty due to multijet background estimation, σ_{NN}^{QCD}, is handled by applying two estimating methods. Finally, σ_{NN}^{tot} is calculated using the quadrature sum of σ_{NN}^{Y}, σ_{NN}^{QCD} and the uncertainty caused by τ kinematics, σ_{NN}^{kin}, based on two transverse momentum selections, $p_T^\tau > 25$ GeV and $p_T^\tau > 25$ GeV. Table 6.1 lists the MC corrections and systematics for each τ type corresponding to the NN_τ cut used in this analysis.

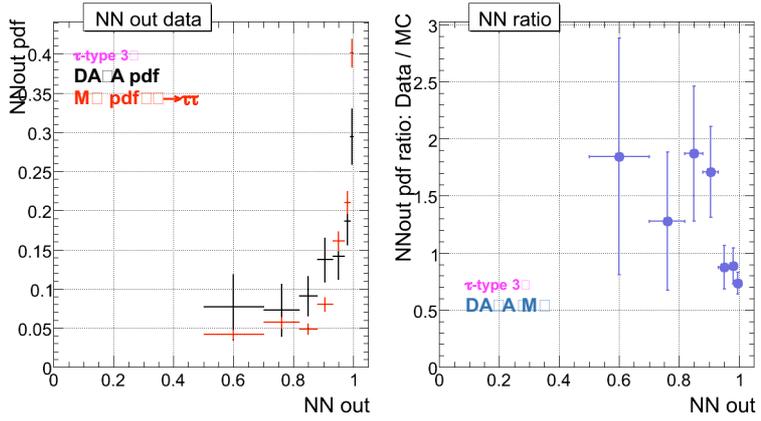| τ type	NN_τ cut	f_{NN}	σ_{NN}^{tot} (%)
1	≥ 0.9	0.84	12
2	≥ 0.9	0.95	4.2
3	≥ 0.95	0.83	7

Table 6.1: Measured τ correction factors to MC events and the corresponding systematics.
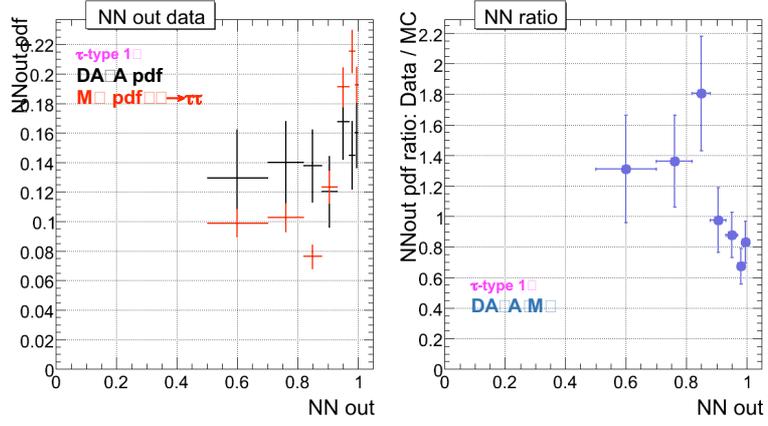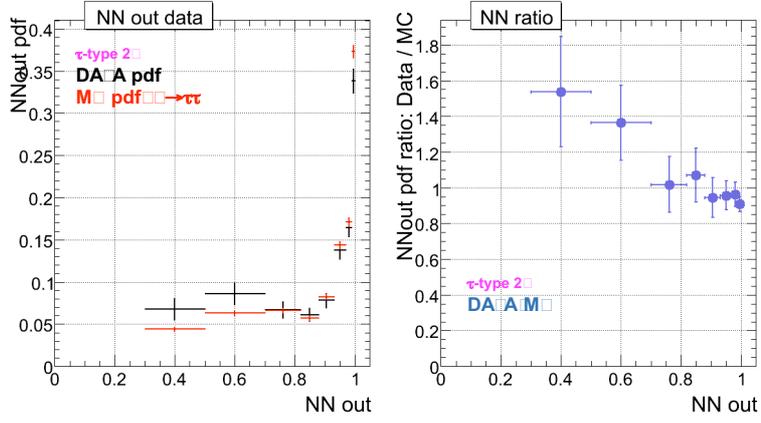
Figure 6.8: Efficiencies of data and MC (left) and the correction factors $f_{NN}(X)$ (right) derived in a $Z \to \tau^+\tau^-$ sample for (a) type-1, (b) type-2 and (c) type-3 τ candidates [64].

6.4.2 Tau Energy Scale Correction

In Run IIb, the EM calibration is not properly propagated to τ leptons in the reconstruction level. This has to be done at the analysis level to take into account the correct calibration for both data and MC events. Also, previous studies on calibrating the energy scale for τ leptons were based solely on MC studies [62]. However, efforts to calibrate the scale using data-driven techniques were made and applied here. Using τ leptons selected from an enriched dataset containing $Z \rightarrow \tau\tau$ events, the difference between data and MC in the energy scale of hadronically decaying τ leptons can be probed using the variable $p_T^\tau / \Sigma p_{Ttrk}^\tau$, where p_T^τ is the energy measured in the calorimeter and Σp_{Ttrk}^τ is the sum of the tracking momentum. In this section, the Run IIb $\tau_\mu \tau_h$ data after preselection requirements is used as examples to show the comparison between data and MC distributions. Assuming the track momentum is well measured, any difference between data and MC in this variable should be due to a difference in the calorimeter energy scale between data and MC. Deviations in the distribution of $p_T^\tau / \Sigma p_{Ttrk}^\tau$ shown in Fig. 6.9 imply that the energy scale is not fully described in MC. This causes differences in the p_T^τ distributions between data and MC as shown in Fig. 6.10, which further results in shifts for either the invariant mass of the μ-τ pair, $M(\mu, \tau)$, as illustrated in Fig. 6.11, and a miscalculated $\not{E}_T$, as shown in Fig. 6.12. Table 6.2 provides a comparison of the mean value of the E^{cal}/P^{trk} between data ($\bar{X}_{data}$) and total predicted background ($\bar{X}_{bkg}$) and the difference between the two ($\Delta \bar{X}$) without any correction applied to the τ energy scale.

τ type	$\bar{X}_{data}$	$\bar{X}_{bkg}$	$\Delta \bar{X}$
type-1	1.17	1.12 ± 0.02	0.05
type-2	1.59	1.58 ± 0.01	0.01
type-3	0.97	0.96 ± 0.01	0.01

Table 6.2: Summary of the mean of the E^{cal}/P^{trk} distribution for the prediction from the MC and multijet $\bar{X}_{bkg}$ and the data $\bar{X}_{data}$ before applying the energy correction. The difference between the means in data and predicted, $\Delta \bar{X}$, is also listed per τ type.

The corrections for the hadronic τ energy scale are derived as a function of the ratio, f_{em}, which is the ratio of the energy clusters of τ candidates found in
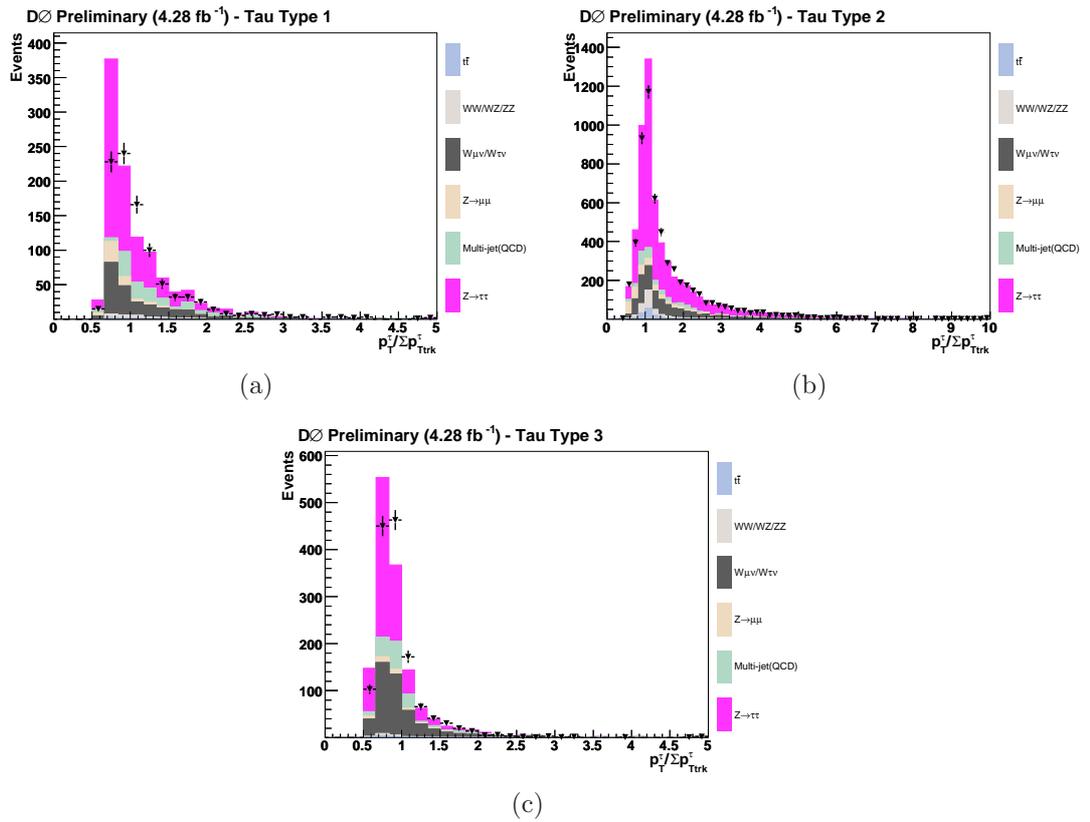
(a)

(b)

(c)

Figure 6.9: Distributions of E^{cal}/P^{trk} for (a) type-1 (b) type-2 and (c) type-3 τ candidates before correcting the energy scale.
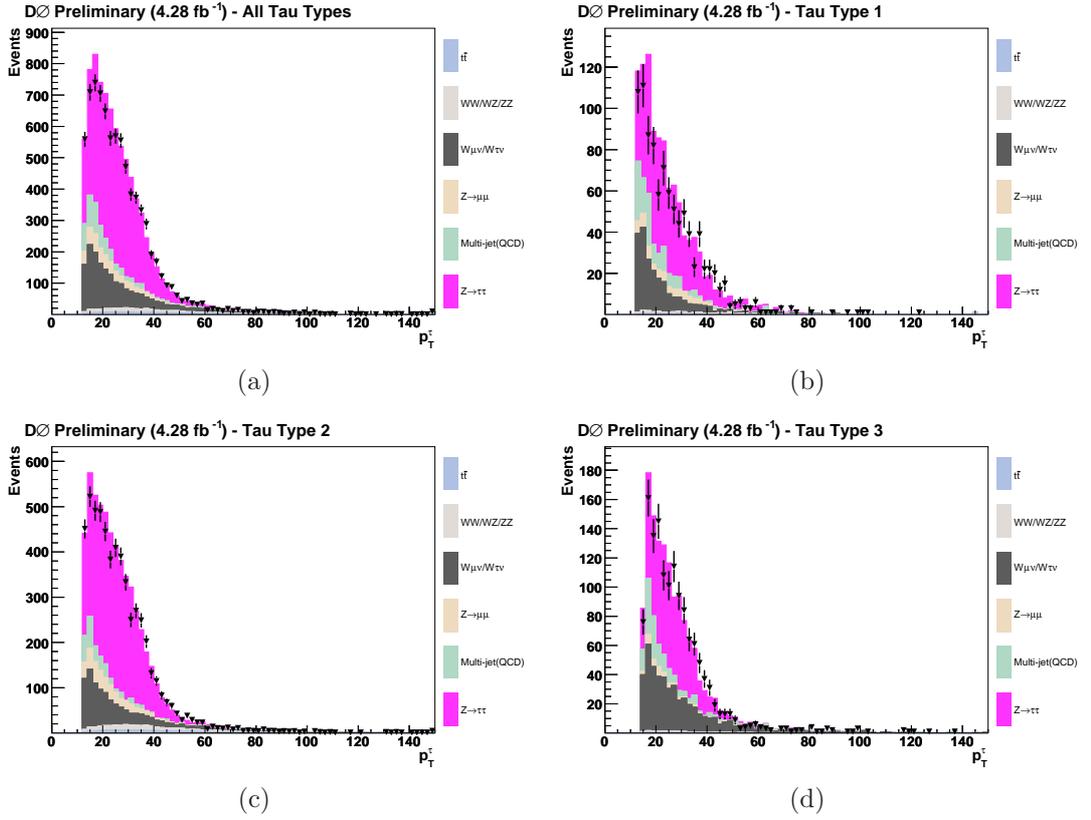
Figure 6.10: Distributions of p_T^τ for (a) all τ types and (b) type-1 (c) type-2 and (d) type-3 τ candidates before correcting the energy scale.
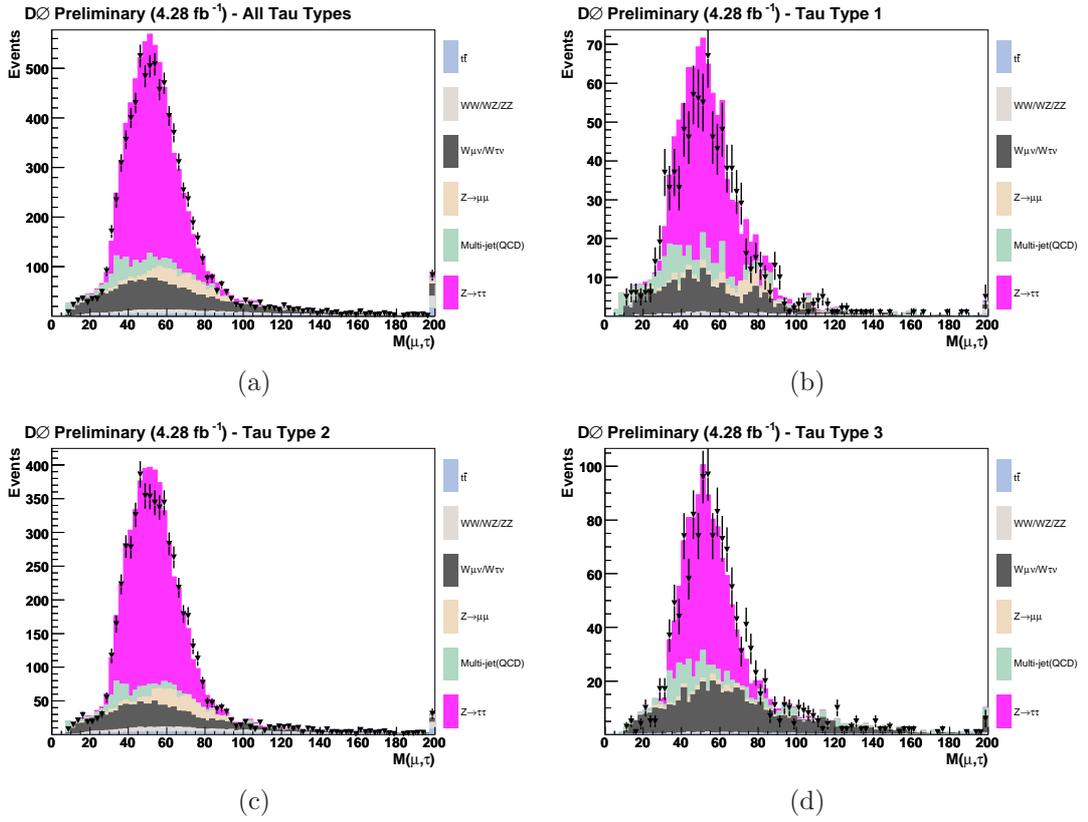
Figure 6.11: Distributions of $M(\mu, \tau)$ for (a) sum of all τ types, (b) type-1, (c) type-2 and (d) type-3 τ candidates before correcting the energy scale.
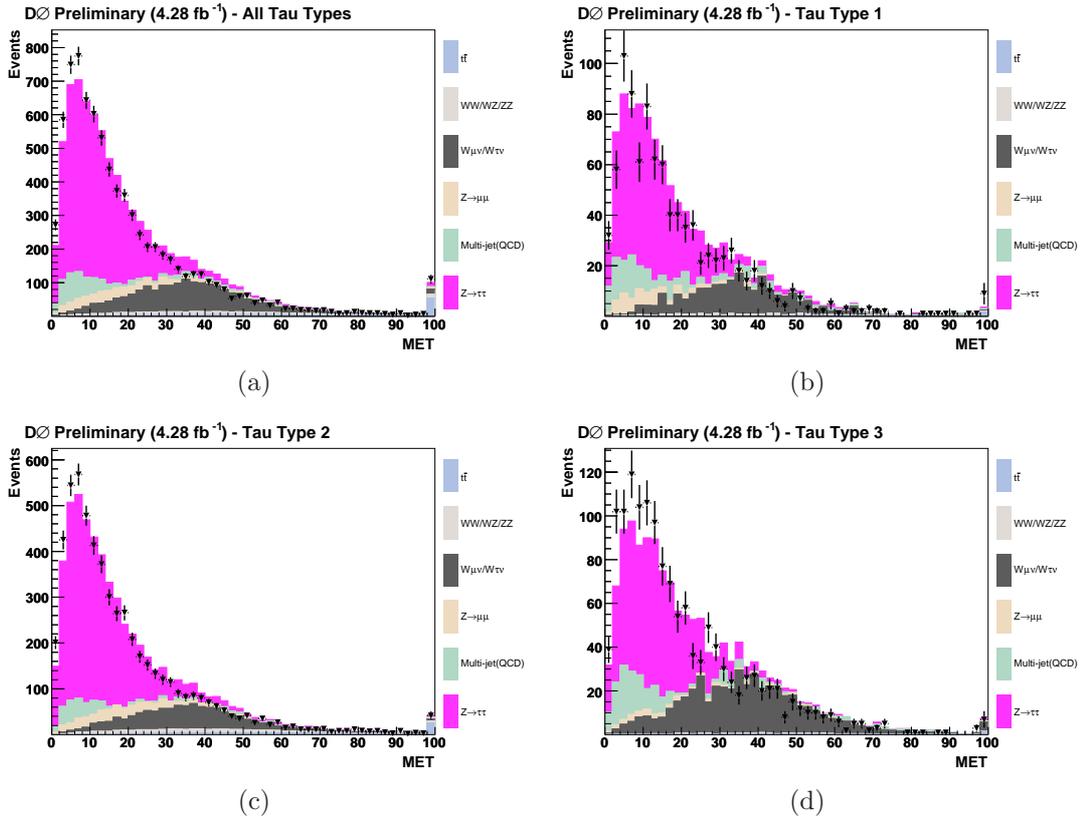
Figure 6.12: Distributions of $\not{E}_T$ for (a) sum of all τ types, (b) type-1, (c) type-2 and (d) type-3 τ candidates before correcting the energy scale.

EM layers to the total τ energy in the calorimeter [65]. Clusters with low f_{em} correspond to τ candidates with large contribution from charged pions, whereas clusters with high f_{em} correspond to τ candidates with a large contribution from $\pi^0 \to \gamma\gamma$. The energy of each MC τ lepton is allowed to vary in the region of

$$0.8 \leq \frac{E_{MC}^{corr}}{E_{MC}^{ori}} \leq 1.2 \tag{6.9}$$

where E_{MC}^{ori} is the given MC energy and E_{MC}^{corr} is the varied energy constrained by the above condition. For each f_{em} bin, a log likelihood ratio $LL(corr)$ of the corrected energy is calculated by

$$LL(corr) = \sum_i N_{MC}^{corr}(f_{em}) - N_{data}(f_{em}) \times \log[N_{MC}^{corr}(f_{em})], \tag{6.10}$$

where $N_{MC}^{corr}(f_{em})$ and $N_{data}(f_{em})$ are the numbers of MC and data events, respectively, in each f_{em} bin i. The correction for each f_{em} bin i is then derived by the minimum value of $-LL(corr)$, which reflects the best fit of MC to data. Only one f_{em} bin is used for τ type-1 and type-3, i.e., the energies for type-1 and type-3 τ candidates are corrected by multiplying the predicted τ energy by a single factor, f_{E_τ}. The energy correction factor $f_{E_\tau}(f_{em})$ for τ type-2 is given by:

$$f_{E_\tau}(f_{em}) = a \times f_{em} + b \tag{6.11}$$

using four f_{em} bins. The corresponding values for slope a and y-intercept b are shown in Fig. 6.13 for each τ type. Since this correction can pull low p_T^τ events back into the sample, which originally failed p_T^τ cuts at the selection stage, the correction must be applied prior to final selections of the hadronic τ candidates. Figures 6.14, 6.15, and 6.16(a) show distributions of p_T^τ , E^{cal}/P^{trk}, and $M(\mu, \tau)$ after the energy scale correction is applied. In these figures, MC events show a better description of data after the energy is corrected compared to Figs. 6.9, 6.10, 6.11, and 6.12. Furthermore, Table 6.3 lists the mean values of E^{cal}/P^{trk} for data and MC after applying the energy correction.

The corrected calorimeter energy is used for the $\not{E}_T$ re-computation for τ type-2 and type-3. Since type-1 τ candidates are mainly from $\pi^\pm$ decay, the tracking momentum should be used to provide a better resolution compared to the energy measured from the calorimeter. Figure 6.17 shows p_T^τ measured in the
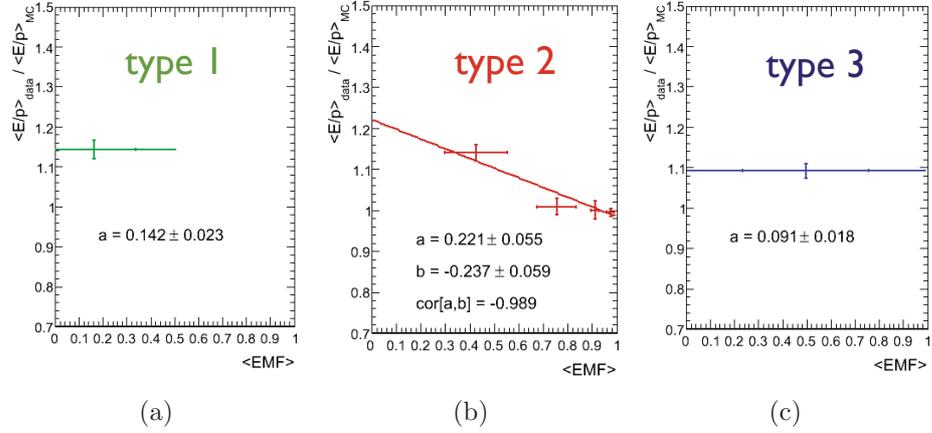
Figure 6.13: Energy correction factors as a function of f_{em} bin for (a) type-1 (b) type-2 and (c) type-3 τ candidates [65].
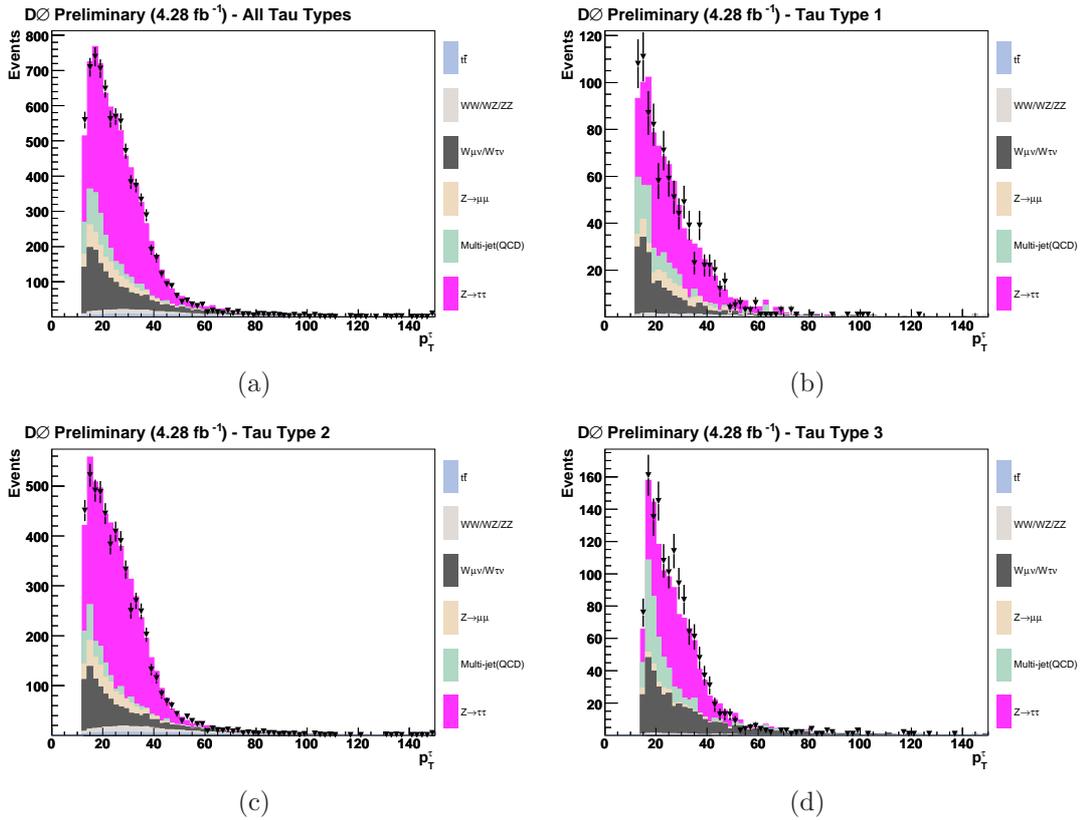

Figure 6.14: Distributions of p_T^τ for (a) sum of all τ types, (b) type-1, (c) type-2 and (d) type-3 τ candidates after correcting the energy scale.
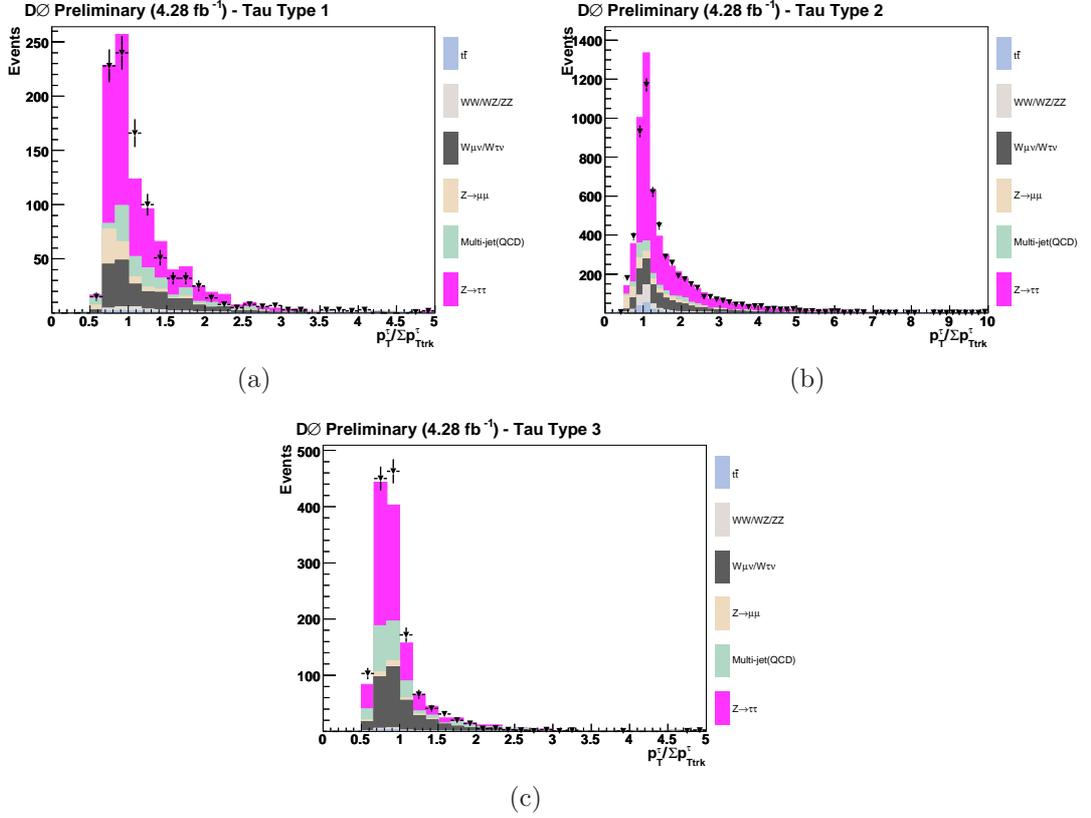
Figure 6.15: Distributions of E^{cal}/P^{trk} for (a) type-1 , (b) type-2 and (c) type-3 τ candidates after correcting the energy scale.

τ type	$\bar{X}_{data}$	$\bar{X}_{bkg}$	$\Delta\bar{X}$
type-1	1.18	1.19 ± 0.02	0.01
type-2	1.59	1.61 ± 0.01	0.02
type-3	0.97	1.01 ± 0.01	0.03

Table 6.3: Mean of the E^{cal}/P^{trk} distribution for the prediction from the MC and multijet, $\bar{X}_{bkg}$, and the data, $\bar{X}_{data}$, after applying the energy correction. The difference between the means in data and in MC, $\Delta\bar{X}$, is also given per τ type.

Figure 6.16: Distributions of $M(\mu, \tau)$ for (a) sum of all τ types, (b) type-1, (c) type-2 and (d) type-3 τ candidates after correcting the energy scale.
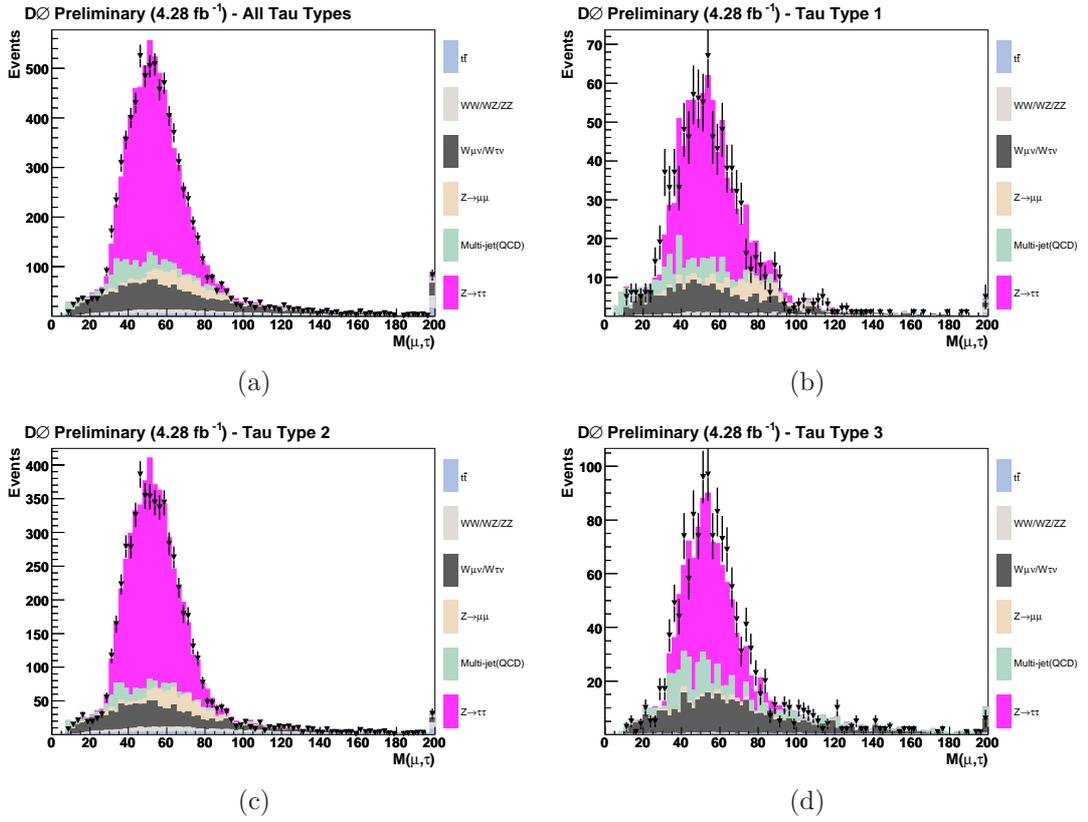
$\tau_\mu \tau_h$ channel for the Run IIb integrated luminosity from the calorimeter or from the tracking system. However, the tracking resolution for single pions is worse than the calorimeter resolution above $p_T > 50$ GeV [63], and the resolution of reconstructed $\not{E}_T$, which is based on the energies of reconstructed τ leptons also becomes worse. This effect is not observed in background events, which have relative low masses of $\mu - \tau$ pairs compare to Higgs bosons, but it becomes more significant while reconstructing $\not{E}_T$ or the visible mass M_{vis} of signal events. The variable, M_{vis}, is defined as:

$$M_{vis} = \sqrt{(P_{\tau_1} + P_{\tau_2} + \not{E}_T)^2}. \tag{6.12}$$

Here, M_{vis} is calculated using the four vectors, $P_{\tau_{1,2}}$, of the visible τ decay products, τ_1 and τ_2, and the missing momentum $\not{P}_T = (\not{E}_T, \not{E}_x, \not{E}_y, 0)$ due to neutrinos. The variables $\not{E}_x$ and $\not{E}_y$ represents the x and y components of $\not{E}_T$ while the z component, $\not{E}_z$, is unknown and therefore set to zero. Figure 6.18 shows the $\not{E}_T$ and M_{vis} distributions measured by the tracking system for type-1 background events from $Z \to \tau^+\tau^-$ decays and signal events with different Higgs boson masses, and it is observed that $\not{E}_T$ resolution of higher masses Higgs bosons is bad which makes the M_{vis} distributions flat without peaking at the generated mass points. At the same time, calorimeter energy is able to provide a much better constraint for signals with the same masses as shown in Figure 6.19. Therefore, the energy of type-1 τ candidates is replaced by track momentum if the calorimeter energy is less than 50 GeV ($p_T^{cal} < 50\ GeV$). This replacement is used during the M_{vis} calculation in the $\tau_\mu \tau_h$ channel, and during the $\not{E}_T$ re-computation which also affects the value of M_{vis}. For τ type-2 and type-3, where a contribution from π^0 exists, the calorimeter energy is used throughout and tracking information is not used for the M_{vis} calculation. The $\not{E}_T$ distribution of the sum of all τ types is shown in Fig. 6.20 (a), and Fig. 6.20 (b) shows the recomputed type-1 $\not{E}_T$ using track momentum. A better representation in MC compared to what is shown in Fig. 6.12 is observed in $\not{E}_T$ after correcting the τ energy.
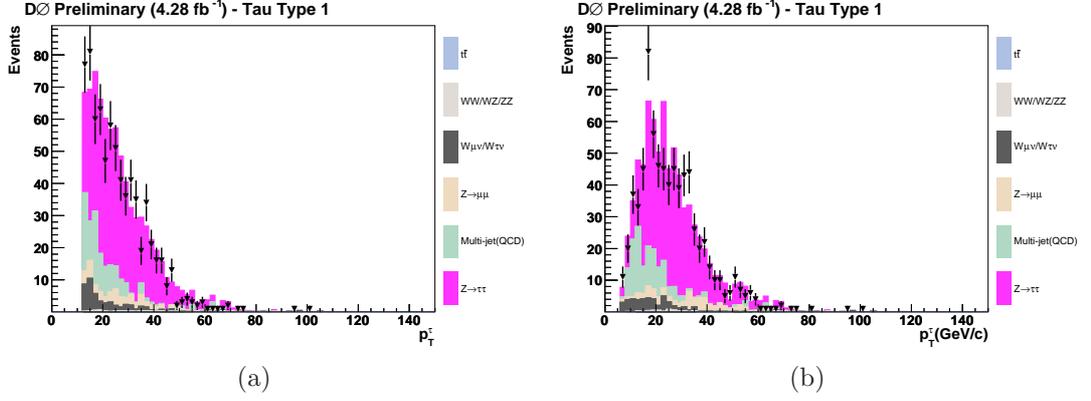
Figure 6.17: Distributions of type-1 p_T^τ measured from (a) calorimeter (b) tracking system.
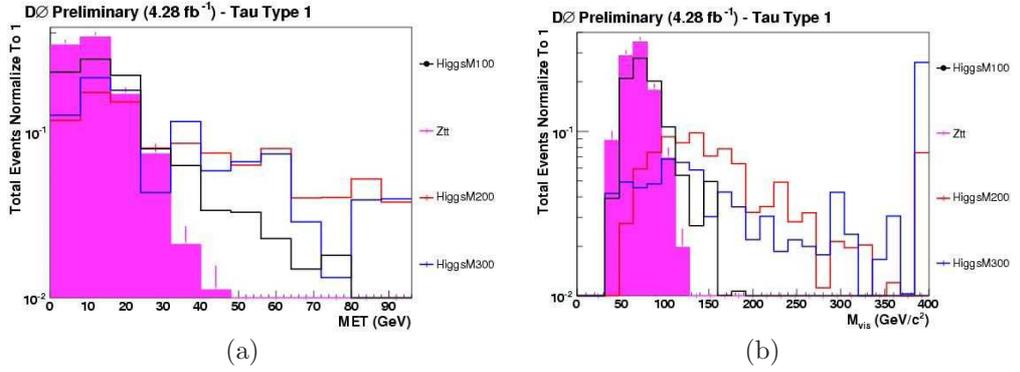
Figure 6.18: Distributions of type-1 (a) $\not{E}_T$ (b) M_{vis} measured from tracking system. Background events from $Z \to \tau^+\tau^-$ decay are shown as pink area, and signals of $M_\phi = 100$ GeV (black), $M_\phi = 200$ GeV (red) and $M_\phi = 300$ GeV (blue) are also shown.

Figure 6.19: Distributions of type-1 (a) $\not{E}_T$ (b) M_{vis} measured from tracking system. Background events from $Z \to \tau^+\tau^-$ decay are shown as pink area, and signals of $M_\phi = 100$ GeV (black), $M_\phi = 200$ GeV (red) and $M_\phi = 300$ GeV (blue) are also shown.

Figure 6.20: Distributions of $\not{E}_T$ for (a) sum of all τ types, (b) type-1, (c) type-2 and (d) type-3 τ candidates after correcting the energy scale.

6.5 MC Samples

6.5.1 Higgs Signal Monte Carlo

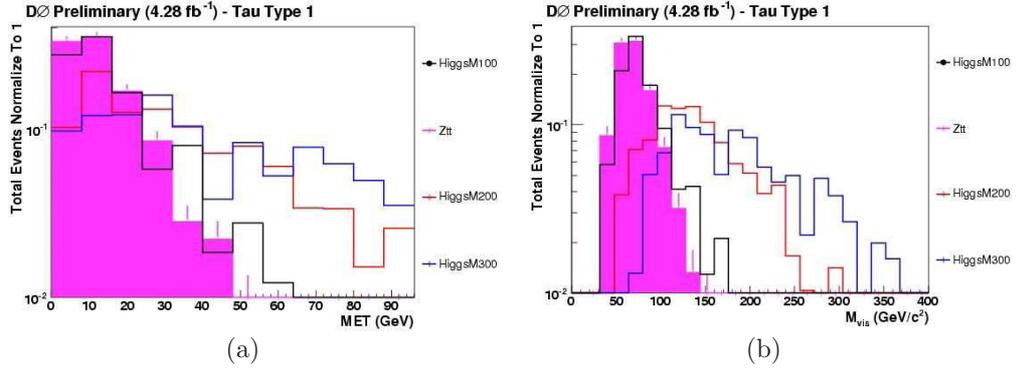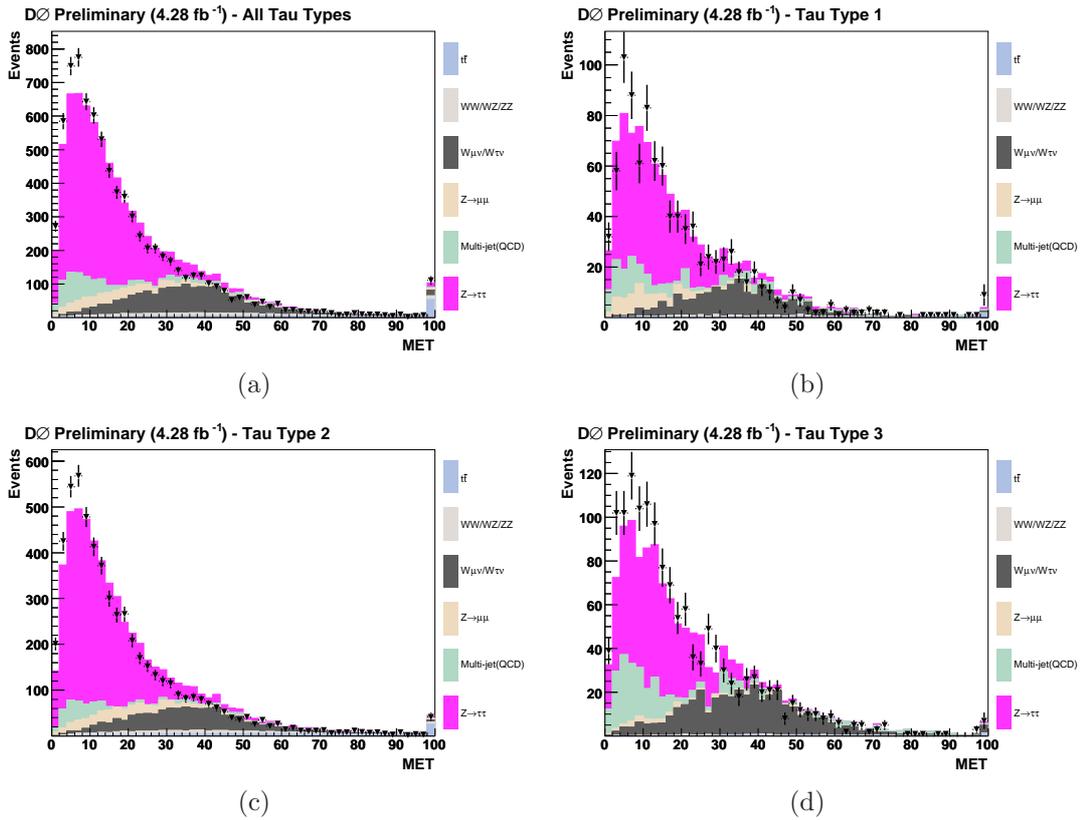The signal samples used for different Higgs boson mass values range from 90 to 320 GeV each separated by intervals of 10 GeV, and were generated for the $\tau_\mu\tau_h$ channel with PYTHIA 6.413 [49] using the CTEQ6L1 [52] leading order (LO) parton distribution functions. Signal MC samples for the $\tau_e\tau_h$ channel in the Higgs boson mass range from 90 to 200 GeV were also generated in intervals of 10 GeV, and in 20 GeV intervals above 200 GeV. No signal sample used in the $\tau_e\tau_h$ channel has a mass above 300 GeV. Each sample was subsequently reconstructed using programs described in Sect. 6.1. Simulated signal samples for different masses are summarised in Table 6.4 with the available number of events after applying data quality selections.

The MC Higgs signal samples were generated as a SM-like Higgs, which is dominated by the gluon-fusion, $gg \to \phi$, production mode. Since the rates of $gg \to \phi$ and $b\bar{b} \to \phi$ annihilation production modes depend on $\tan\beta$ in the MSSM, studies have been made on both production modes separately for two masses, $M_\phi = 120$ GeV and $M_\phi = 160$ GeV. Table 6.5 lists the reconstruction efficiencies by τ type. The $b\bar{b} \to \phi$ production mode shows higher rates across all τ types for both masses. However, the reconstructed $b\bar{b} \to \phi$ rate will be reduced by applying a b-jet veto due to the presence of an extra jet. This b-jet veto is necessary for future combinations to keep the search orthogonal to the other MSSM Higgs search channel, $b\bar{b} \to \phi$, where $\phi \to b\bar{b}$. Since the remaining difference in efficiencies is small, it is neglected in the determination of the final limits.

The width of the SM Higgs boson increases sharply near $M_h \approx 160$ GeV where the effect of the W^+W^- decay enters resulting in the width of SM Higgs boson becoming larger than that of MSSM Higgs bosons for Higgs boson masses above 200 GeV as shown in Fig. 6.21. Hence, for higher mass samples, a wide-width structure is observed when analysing the MC truth information which matches the expectation (see Fig 6.22(a)). This wide peak degrades the limit at higher Higgs masses by yielding a wider reconstructed visible mass. Hence, a mass window is applied to select events that only appearing within ± 80 GeV of the generated mass point. The "fixed" true mass peaks for different mass points are

Higgs Mass	$\tau_\mu\tau_h$ Channel		$\tau_e\tau_h$ Channel
M_ϕ, [GeV]	N_{event} (Run IIa)	N_{event} (Run IIb)	N_{event} (Run IIa)
90	294,648	287,239	98,621
100	287,403	360,159	100,348
110	192,124	282,709	101,441
120	192,434	359,429	96,098
130	192,986	270,474	98,242
140	194,791	265,329	95,321
150	181,066	360,769	98,595
160	97,141	356,882	96,747
170	54,840	355,660	97,201
180	47,686	346,325	96,181
190	99,055	347,098	97,986
200	293,568	280,915	99,934
210	100,131	173,921	–
220	96,119	141,031	97,175
230	92,637	339,898	–
240	91,934	322,308	104,899
250	83,344	291,834	–
260	61,680	214,773	113,149
270	30,890	109,708	–
280	26,677	91,911	96,756
290	22,248	77,615	–
300	285,888	64,723	97,713
310	14,738	53,273	–
320	12,257	44,352	–

Table 6.4: Signal Monte Carlo samples used in the analysis, with the number of events, N_{event}, after duplicate event removal and data quality requirements.

Mass (GeV)	Production mode	type-1	type-2	type-3
120	$b\bar{b} \to \phi$	0.112 ± 0.011	0.633 ± 0.026	0.141 ± 0.013
	$gg \to \phi$	0.109 ± 0.011	0.537 ± 0.024	0.127 ± 0.013
160	$b\bar{b} \to \phi$	0.164 ± 0.013	0.855 ± 0.030	0.189 ± 0.014
	$gg \to \phi$	0.143 ± 0.012	0.824 ± 0.031	0.169 ± 0.013

Table 6.5: Reconstruction efficiencies of $b\bar{b}$ and gg production modes by τ type using selections applied to the $\tau_\mu\tau_h$ channel.
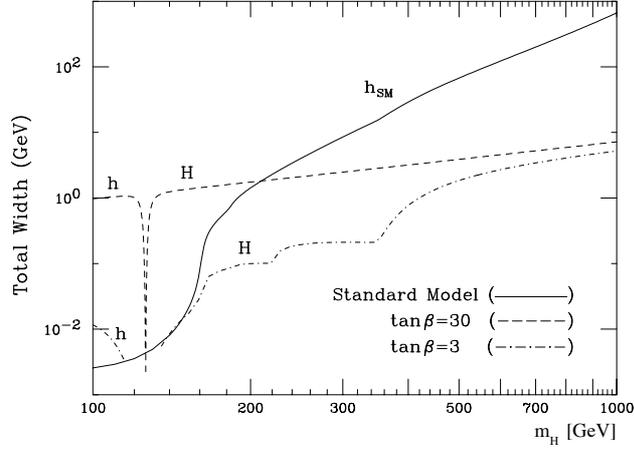
94

shown in Fig 6.22(b).

Figure 6.21: The total decay width of the SM Higgs boson, shown as a function of m_H. Also shown are the decay widths for the CP-even neutral Higgs bosons, h and H, for two choices of $\tan\beta$ in the MSSM benchmark scenario $m_h^{\max}$ described in Chapter 2 [13].
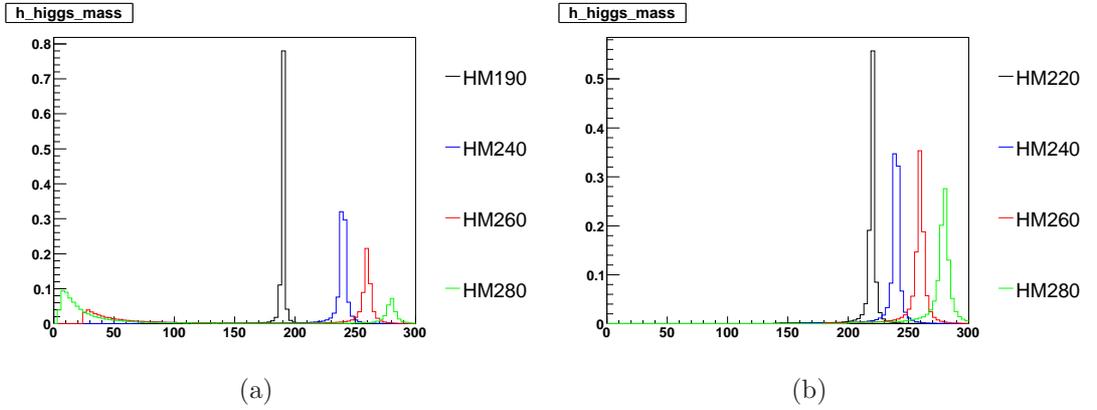
Figure 6.22: Distribution of simulated Higgs boson masses generated by PYTHIA for $M_\phi = 190$ GeV (black), $M_\phi = 240$ GeV (blue), $M_\phi = 260$ GeV (red), and $M_\phi = 280$ GeV (green) for (a) without any mass window applied (b) only events lie in the window of ± 80 GeV from the generated mass points. All distributions are normalised to the sum of events equal to one in order to compare the differences in shapes.

The dependence of the widths of MSSM Higgs bosons on $\tan\beta$ has been studied in the DØ Run IIa dataset [53]. Those results show that the effect is only significant at very high values of $\tan\beta \approx 100$, which are far away from the

95

values preferred theoretically ($\tan \beta \approx 35$) and already excluded [53]. Therefore, widths of MSSM Higgs bosons are assumed to be very close to that of the SM Higgs boson in the $\tau_\mu \tau_h$ channel.

6.5.2 Background Monte Carlo

Background Monte Carlo in the $\tau_\mu \tau_h$ channel

The majority of background Monte Carlo events used in the $\tau_\mu \tau_h$ decay channel were generated either with `ALPGEN` 2.11 [50] or `PYTHIA` version 6.413 [49] using the CTEQ parton distribution functions (PDFs) [52] as for the Higgs signal samples. The main background processes for the search are Z/γ^* decaying to either $\tau\tau$, ee and $\mu\mu$, as well as W+jets, di-boson (WW, WZ, and ZZ), $t\bar{t}$ and multijet production. The W+jet, Z/γ^*, and $t\bar{t}$ MC samples have been generated using `ALPGEN` with parton showering and hadronisation provided by `PYTHIA`, while diboson processes are simulated using `PYTHIA` alone. Estimation of the multijet background for both $\tau_\mu \tau_h$ and $\tau_e \tau_h$ channels is taken directly from data and complete details of the estimation method is given in Sect. 7.1.

The $Z/\gamma^* \to ll$ cross section is determined by $\sigma(Z/\gamma^* \to ll) = \sigma_{NLO} \times K_{NNLO}(Q^2)$ using CTEQ6.1M PDFs. Here, the NLO cross section, σ_{NLO} uses NLO PDF where the K_{NNLO}-factor is introduced with respect to NNLO using NLO to account for effects of higher-order multijet processes. Therefore, K_{NNLO} values ranging from 1.256 to 1.330 (1.275 to 1.300) are applied to take into consideration effects for $Z \to \mu\mu$ ($Z \to \tau\tau$) processes generated in different invariant mass bins. The complete list is summarised in Table 6.6. Similarly, the $W \to l\nu$ cross section is determined using K_{NNLO}=1.266 for each generated sample. The $t\bar{t}$ cross section is calculated at NNLO and each top sample is generated with a top quark mass of 172 GeV. The di-boson (WW, ZZ, and WZ) cross sections are taken at NLO. The number of events passing data quality check of each background sample with their cross sections are listed in Tables 6.7 and 6.8.

Background Monte Carlo in the $\tau_e \tau_h$ channel

The background Monte Carlo events used in the $\tau_e \tau_h$ channel were only generated with `PYTHIA` v6.3 [49] and no K_{NNLO}-factor is applied. The effect of different choice of the MC generators in this channel has been studied, and the results are

Process	Mass Range [GeV]	K_{NNLO}-factor
$Z/\gamma^* \to \mu^+\mu^- + 0\text{lp}$	$15 < M_{\mu\mu} < 75$	1.256
$Z/\gamma^* \to \mu^+\mu^- + 1\text{lp}$	$75 < M_{\mu\mu} < 130$	1.264
$Z/\gamma^* \to \mu^+\mu^- + 2\text{lp}$	$130 < M_{\mu\mu} < 250$	1.330
$Z/\gamma^* \to \mu^+\mu^- + 3\text{lp}$	$250 < M_{\mu\mu} < 1960$	1.295
$Z/\gamma^* \to \tau^+\tau^- + 0\text{lp}$	$15 < M_{\tau\tau} < 75$	1.275
$Z/\gamma^* \to \tau^+\tau^- + 1\text{lp}$	$75 < M_{\tau\tau} < 130$	1.29
$Z/\gamma^* \to \tau^+\tau^- + 2\text{lp}$	$130 < M_{\tau\tau} < 250$	1.281
$Z/\gamma^* \to \tau^+\tau^- + 3\text{lp}$	$250 < M_{\tau\tau} < 1960$	1.300
$W \to l\nu + 0\text{lp}$	–	1.266
$W \to l\nu + 1\text{lp}$	–	1.266
$W \to l\nu + 2\text{lp}$	–	1.266
$W \to l\nu + 3\text{lp}$	–	1.266
$W \to l\nu + 4\text{lp}$	–	1.266
$W \to l\nu + 5\text{lp}$	–	1.266
$t\bar{t} \to 2b + 2l2\nu$	–	1.392
$t\bar{t} \to 2b + 4lpcl$	–	1.498
$t\bar{t} \to 2b + 2lpc + l\nu$	–	1.455

Table 6.6: NNLO K-factors applied during normalisation of the generated W+jets and Z+jets background processes to take into account effects from higher-order multijet production [66]. For the W samples, $l = \mu, \tau$.

Process	Mass Range [GeV]	N_{DQ} (Run IIa)	N_{DQ} (Run IIb)	$\sigma \times \mathrm{BR}$ [pb]
$Z/\gamma^* \to \tau^+\tau^- + 0\mathrm{lp}$	$15 < M_{\tau\tau} < 75$	1,659,261	1,666,177	338
$Z/\gamma^* \to \tau^+\tau^- + 1\mathrm{lp}$	$15 < M_{\tau\tau} < 75$	592,821	594,906	39.9
$Z/\gamma^* \to \tau^+\tau^- + 2\mathrm{lp}$	$15 < M_{\tau\tau} < 75$	328,600	329,981	10.0
$Z/\gamma^* \to \tau^+\tau^- + 3\mathrm{lp}$	$15 < M_{\tau\tau} < 75$	340,452	341,658	2.77
$Z/\gamma^* \to \tau^+\tau^- + 0\mathrm{lp}$	$75 < M_{\tau\tau} < 130$	1,556,389	1562,018	131
$Z/\gamma^* \to \tau^+\tau^- + 1\mathrm{lp}$	$75 < M_{\tau\tau} < 130$	595,169	596,880	40.3
$Z/\gamma^* \to \tau^+\tau^- + 2\mathrm{lp}$	$75 < M_{\tau\tau} < 130$	305,312	306,144	9.99
$Z/\gamma^* \to \tau^+\tau^- + 3\mathrm{lp}$	$75 < M_{\tau\tau} < 130$	205,365	205,809	3.10
$Z/\gamma^* \to \tau^+\tau^- + 0\mathrm{lp}$	$130 < M_{\tau\tau} < 250$	366,567	367,982	0.92
$Z/\gamma^* \to \tau^+\tau^- + 1\mathrm{lp}$	$130 < M_{\tau\tau} < 250$	180,807	181,529	0.38
$Z/\gamma^* \to \tau^+\tau^- + 2\mathrm{lp}$	$130 < M_{\tau\tau} < 250$	181,640	182,360	0.10
$Z/\gamma^* \to \tau^+\tau^- + 3\mathrm{lp}$	$130 < M_{\tau\tau} < 250$	188,620	189,168	0.04
$Z/\gamma^* \to \tau^+\tau^- + 0\mathrm{lp}$	$250 < M_{\tau\tau} < 1960$	546,767	548,792	0.07
$Z/\gamma^* \to \tau^+\tau^- + 1\mathrm{lp}$	$250 < M_{\tau\tau} < 1960$	488,524	490,870	0.04
$Z/\gamma^* \to \tau^+\tau^- + 2\mathrm{lp}$	$250 < M_{\tau\tau} < 1960$	371,192	372,497	0.01
$Z/\gamma^* \to \tau^+\tau^- + 3\mathrm{lp}$	$250 < M_{\tau\tau} < 1960$	178,014	178,756	0.004
$Z/\gamma^* \to \mu^+\mu^- + 0\mathrm{lp}$	$15 < M_{\mu\mu} < 75$	1,868,435	1,877,017	344
$Z/\gamma^* \to \mu^+\mu^- + 1\mathrm{lp}$	$15 < M_{\mu\mu} < 75$	640,691	642,803	40.1
$Z/\gamma^* \to \mu^+\mu^- + 2\mathrm{lp}$	$15 < M_{\mu\mu} < 75$	324,113	325,330	9.87
$Z/\gamma^* \to \mu^+\mu^- + 3\mathrm{lp}$	$15 < M_{\mu\mu} < 75$	328,210	329,589	2.84
$Z/\gamma^* \to \mu^+\mu^- + 0\mathrm{lp}$	$75 < M_{\mu\mu} < 130$	1,553,222	1,565,523	134
$Z/\gamma^* \to \mu^+\mu^- + 1\mathrm{lp}$	$75 < M_{\mu\mu} < 130$	639,392	641,154	41.4
$Z/\gamma^* \to \mu^+\mu^- + 2\mathrm{lp}$	$75 < M_{\mu\mu} < 130$	446,737	447,655	9.91
$Z/\gamma^* \to \mu^+\mu^- + 3\mathrm{lp}$	$75 < M_{\mu\mu} < 130$	172,628	173,038	3.25
$Z/\gamma^* \to \mu^+\mu^- + 0\mathrm{lp}$	$130 < M_{\mu\mu} < 250$	359,102	360,533	0.89
$Z/\gamma^* \to \mu^+\mu^- + 1\mathrm{lp}$	$130 < M_{\mu\mu} < 250$	179,991	180,775	0.36
$Z/\gamma^* \to \mu^+\mu^- + 2\mathrm{lp}$	$130 < M_{\mu\mu} < 250$	179,228	179,895	0.98
$Z/\gamma^* \to \mu^+\mu^- + 3\mathrm{lp}$	$130 < M_{\mu\mu} < 250$	170,045	170,809	0.03
$Z/\gamma^* \to \mu^+\mu^- + 0\mathrm{lp}$	$250 < M_{\mu\mu} < 1960$	560,971	564,091	0.07
$Z/\gamma^* \to \mu^+\mu^- + 1\mathrm{lp}$	$250 < M_{\mu\mu} < 1960$	466,903	469,071	0.03
$Z/\gamma^* \to \mu^+\mu^- + 2\mathrm{lp}$	$250 < M_{\mu\mu} < 1960$	367,301	368,836	0.01
$Z/\gamma^* \to \mu^+\mu^- + 3\mathrm{lp}$	$250 < M_{\mu\mu} < 1960$	366,606	368,377	0.004

Table 6.7: MC samples of Z/γ^* backgrounds used in the $\tau_\mu \tau_h$ channel together with the number of events passing the data quality checks, N_{DQ}, and $\sigma \times \mathrm{BR}$ for each process. For each Z/γ^* sample, the production cross section is calculated at NNLO [66].

Process	N_{DQ} (Run IIa)	N_{DQ} (Run IIb)	$\sigma \times \mathrm{BR}$ [pb]
$W \to l\nu + 0\mathrm{lp}$	47,070,044	47,444,656	4513
$W \to l\nu + 1\mathrm{lp}$	20,683,540	20,831,562	1278
$W \to l\nu + 2\mathrm{lp}$	19,686,862	19,769,803	303.7
$W \to l\nu + 3\mathrm{lp}$	4,269,023	4,284,266	72.6
$W \to l\nu + 4\mathrm{lp}$	3,084,707	3,097,008	16.8
$W \to l\nu + 5\mathrm{lp}$	2,565,942	2,694,913	5.1
$WW \to$ incl.	709,879	712,436	11.6
$WZ \to$ incl.	632,296	634,730	3.25
$ZZ \to$ incl.	540,273	542,690	1.33
$t\bar{t} \to 2b + 2l2\nu + 0\mathrm{lp}$	749,642	752,776	0.35
$t\bar{t} \to 2b + 2l2\nu + 1\mathrm{lp}$	452,117	454,273	0.14
$t\bar{t} \to 2b + 2l2\nu + 2\mathrm{lp}$	281,453	282,447	0.07
$t\bar{t} \to 2b + 4lpcl + 0\mathrm{lp}$	2,655,792	1,534,905	3.51
$t\bar{t} \to 2b + 2lpc + l\nu + 0\mathrm{lp}$	1,534,107	1,563,304	2.40

Table 6.8: Other background Monte Carlo samples used in the $\tau_\mu \tau_h$ channel together with the number of events passing the data quality checks, N_{DQ}, and $\sigma \times \mathrm{BR}$ for each process. Both the W and $t\bar{t}$ cross section are calculated at NNLO while the di-boson cross sections are taken at NLO. The $t\bar{t}$ MC samples are generated with a top quark mass of 172 GeV [66].

consistent. Table 6.9 lists the background MC samples used and the theoretical cross-section (σ) used for sample normalisation. Also shown in the table is the number of generated events (N_{gen}), and the number of events after removal of duplicate events and events that fail the data quality check (N_{DQ}).

Process	Mass (GeV)	N_{gen}	N_{DQ}	$\sigma \times \mathrm{BR}$ [pb]
$Z/\gamma^* \to \tau^+\tau^-$	$60 - 130$	5,791,737	5,601,262	256.6
$Z/\gamma^* \to \tau^+\tau^-$	$15 - 60$	1,873,000	1,809,781	409
$Z/\gamma^* \to \tau^+\tau^-$	$130 - 250$	409,250	394,674	1.96
$Z/\gamma^* \to \tau^+\tau^-$	$250 - 500$	102,250	99,518	0.16
$Z/\gamma^* \to e^+e^-$	$60 - 130$	3,056,250	2,954,973	256.6
$Z/\gamma^* \to e^+e^-$	$15 - 60$	2,071,000	2,000,391	409
$Z/\gamma^* \to e^+e^-$	$130 - 250$	407,250	393,605	1.96
$Z/\gamma^* \to e^+e^-$	$250 - 500$	100,000	96,706	0.16
$W \to e\nu$		7,298,500	5,280,132	2600
$W \to \tau\nu$		3,411,172	3,302,101	2600
$t\bar{t} \to$ incl.		107,250	97,654	6.8
$WW \to$ incl.		722,000	700,164	12.0
$WZ \to$ incl.		209,500	201,824	3.68
$ZZ \to$ incl.		203,750	197,100	1.42

Table 6.9: Background Monte Carlo samples used in the $\tau_e\tau_h$ channel, with the number of events generated, N_{gen}, the number of events after duplicate event removal and data quality requirements, N_{DQ}, and $\sigma \times \mathrm{BR}$ used for normalisation.

Chapter 7

Background Estimation

Proper background modelling is very important for this search in order to detect signal events above the predicted background. One of the major backgrounds for the $\phi \to \tau\tau$ signal is multijet background, which will be discussed in Sect. 7.1. The other main sources are listed below:

$\boldsymbol{Z/\gamma^* \to \tau^+\tau^-}$: Two τ leptons from Z boson decays, where one τ decays leptonically to an electron or a muon and the other τ decays hadronically. It is almost impossible to separate this background from signal apart from using spin or mass of the Z and Higgs boson, and therefore, the $Z/\gamma^* \to \tau^+\tau^-$ component is referred to as "irreducible". The strategy of this analysis relies on reducing the other backgrounds. Signal events are searched for as enhancements above the $Z \to \tau^+\tau^-$ background.

$\boldsymbol{Z/\gamma^* \to \mu^+\mu^-}$ **and** $\boldsymbol{Z/\gamma^* \to e^+e^-}$**:** Events contain one well reconstructed electron or muon candidate, and the second lepton is either missing and the event contains a jet, or it is mis-identified as a τ lepton. In the $\tau_\mu\tau_h$ channel, background events from the $Z/\gamma^* \to \mu^+\mu^-$ process can be suppressed by matching the τ candidate and the muon, and only keeping events with $\Delta R(\mu, \tau) \geq 0.5$. The modelling and rejection of $Z \to e^+e^-$ background in the $\tau_e\tau_h$ channel is more complicated and discussed in detail in Sect. 7.3.

$\boldsymbol{W}$**+jets:** There are two possible ways that this background can mimic signal events. If W bosons decay to an electron or a muon and neutrinos, the events can appear to be signal if jets are mis-identified as τ leptons. Since W bosons also decay to τ leptons and neutrinos, this background will also be mis-identified

as signal events if jets faking muons or electrons are found in the same event. In order to reduce this background, one can apply cuts on the angle of electrons or muons with respect to $\not{E}_T$ or by imposing cuts on the transverse mass, M_T. More details about the selection are discussed in Chapter 8, and the re-weighting factor applied to MC events for this background in the $\tau_e \tau_h$ channel can be found in Sect. 7.3.

WW, ZZ, ZW and $t\bar{t}$: These background sources contain at least one well reconstructed electron or muon plus a real τ lepton. Therefore, it is often challenging to distinguish signal events from these background events. However, since the production cross sections are very small compared to the other background sources, the contribution is almost negligible in low mass search regions.

7.1 Multijet (QCD) Background Estimation

The primary contribution to the multijet background is expected from heavy flavour multijet events, where a lepton from the heavy flavour decay passes the isolation requirement and a jet is mis-identified as a τ lepton. Additionally, a contribution is expected from light quark multijet events where the jets are mis-identified as both a τ lepton and as a muon or an electron. This background source is often referred to as the "QCD" or "multijet" background. It is difficult to simulate such events, and therefore, these background contributions are estimated directly from data.

The primary method used to estimate multijet events relies on defining three different control regions each determined by NN_τ:

- **High NN_τ region (hNN):**
  $NN_\tau > 0.9$ (τ type-1 and type-2); $NN_\tau > 0.95$ (type-3). The τ identification algorithm selects events based on NN_τ for discriminating τ leptons from jet backgrounds. Since τ leptons should have values peaked close to one, this region is also termed the "signal region".

- **Low NN_τ region (lNN):**
  $NN_\tau \leq 0.10$ (all τ types). Composed primarily of W+jet and QCD multijet events and, as described below, used to derive the normalisation of multijet events in the signal region.

- **Medium NN_τ region (mNN):**

  $0.25 \leq NN_\tau \leq 0.75$ (all τ types). The NN_τ distribution is approximately constant in this region and composed mainly of W+jets and QCD multijet events. Therefore, it can be used to perform a cross check on the multijet background composition.

Events from each of the above three NN_τ regions are further separated by selecting events with opposite-sign (OS) or same-sign (SS) electric charge of electron (e) or muon (μ) and the hadronic τ decay. For multijet events where the two leptons are mis-identified jets, the relative sign of the two lepton charges is random and the number of SS and OS events is expected to be equal. However, the muon or electron and the τ lepton from a neutral Higgs boson decay should be OS in order to conserve charge. Consequently, this selection yields a total of five different control regions as listed in Table 7.1 plus a signal region where events are selected by high NN_τ values and OS requirements.

	$NN_\tau \leq 0.1$	$0.25 \leq NN_\tau \leq 0.75$	$NN_\tau \geq 0.9(0.95)$
Same Sign	N_{lNN}^{SS}	N_{mNN}^{SS}	N_{hNN}^{SS}
Opposite Sign	N_{lNN}^{OS}	N_{mNN}^{OS}	N_{hNN}^{OS} (signal)

Table 7.1: Five control and a signal region defined by the value of NN_τ and the relative sign of the given $\mu/e - \tau$ charges. Notations for number of events yield in each region are also listed.

The transverse mass (M_T) is defined as:

$$M_T = \sqrt{2p_T^{\mu/e} \not{E}_T (1 - \cos \Delta\phi(\mu/e, \not{E}_T))} \tag{7.1}$$

where $p_T^{\mu/e}$ is the transverse momentum of the selected muon or electron, $\not{E}_T$ is the missing transverse energy of the event and $\Delta\phi$ is the angle between $\not{E}_T$ and the lepton. This variable is used to separate regions that are enriched with multijet processes or W+jet events. Since W+jet events tend to peak at higher M_T values while multijet production populates at lower M_T, events are selected by requiring $M_T \leq 50$ GeV for all τ types in order to define a multijet enriched sample. The shape for the multijet contribution is taken from a sample which requires the μ/e-τ pairs to have the same sign in order to keep it orthogonal to the signal sample. All other selections are kept the same as for the preselections of τ leptons in the high NN_τ region. This SS sample is normalised by a factor,

ρ_{QCD}, derived from events within the low NN_τ region for each τ type in order to predict the number of multijet events of the OS signal sample. The number of multijet events in the signal region, $N_{hNN}^{OS,QCD}$, is assumed to be given by

$$N_{hNN}^{OS,QCD} = \rho_{QCD} * (N_{hNN}^{SS,data} - N_{hNN}^{SS,MC}), \qquad (7.2)$$

where $N_{hNN}^{SS,data}$ and $N_{hNN}^{SS,MC}$ are the total number of data and MC events found in the SS, high NN_τ region and normalised by ρ_{QCD}. Values for ρ_{QCD} are determined by selecting multijet events within the low NN_τ region using:

$$\rho_{QCD} = \frac{N_{lNN}^{OS,data} - N_{lNN}^{OS,MC}}{N_{lNN}^{SS,data} - N_{lNN}^{SS,MC}}, \qquad (7.3)$$

where $N_{lNN}^{OS,data}$ and $N_{lNN}^{SS,MC}$ are the numbers of data and MC events expected in the OS sample with a low NN_τ cut. Respectively, $N_{lNN}^{SS,data}$ and $N_{lNN}^{SS,MC}$ are the total number of data and MC events expected in the same sign sample in the low NN_τ region. Since there is little "sign preference" of multijet events, the value of ρ_{QCD} is expected to be very close to 1. As indicated in Eqs. 7.2 and 7.3, all backgrounds other than multijet events are subtracted using the corresponding MC background samples listed in Sec. 6.5.2 normalised by $\sigma \times$ BR described in the same section.

Discussions concerning the verification of this method for the $\tau_\mu \tau_h$ search channel can be found in Sect. 7.2.1. Similar verification and cross-checks with a second method using Run IIa data in the $\tau_e \tau_h$ search channel are discussed in Sect. 7.3.1.

7.2 Background Modelling in the $\tau_\mu \tau_h$ Channel

Most sources of background to this Higgs boson search are generated using PYTHIA or ALPGEN, and these samples are listed in Sect. 6.5.2. Overall, the total predicted background is able to describe the data well once these MC samples include the corrections described in Chapter 6. Therefore, no additional treatment to the background is required. A cross-check of the multijet background is discussed in the following Sect. 7.2.1, and further comparison of the data with predicted background with preselection and the final selection can be found in Chapter 8.

7.2.1 Multijet in the $\tau_\mu \tau_h$ Channel

As indicated by Eq. 7.2, the values for ρ_{QCD} derived in the $\tau_\mu \tau_h$ channel for each τ type are listed in Table 7.2. In order to verify if the method estimates the multijet background well, basic kinematic distributions for the τ and muon candidates are studied in control regions dominated by this background. For $\tau_\mu \tau_h$ decays, the control sample is chosen to contain events with medium NN_τ value and the $\mu - \tau$ pairs are required to have electric charges with opposite signs. Figure 7.1 shows distributions from the Run IIb data set of $\not{E}_T$, M_T, p_T^μ and the mass of $\mu - \tau$ pairs ($M_{\mu\tau}$), while Fig. 7.2 shows p_T^τ distributions by τ type. These distributions show good agreement between data and expected backgrounds which indicates the multijet background is well-modelled through the use of the SS region to determine the shape and of ρ_{QCD} for normalisation.
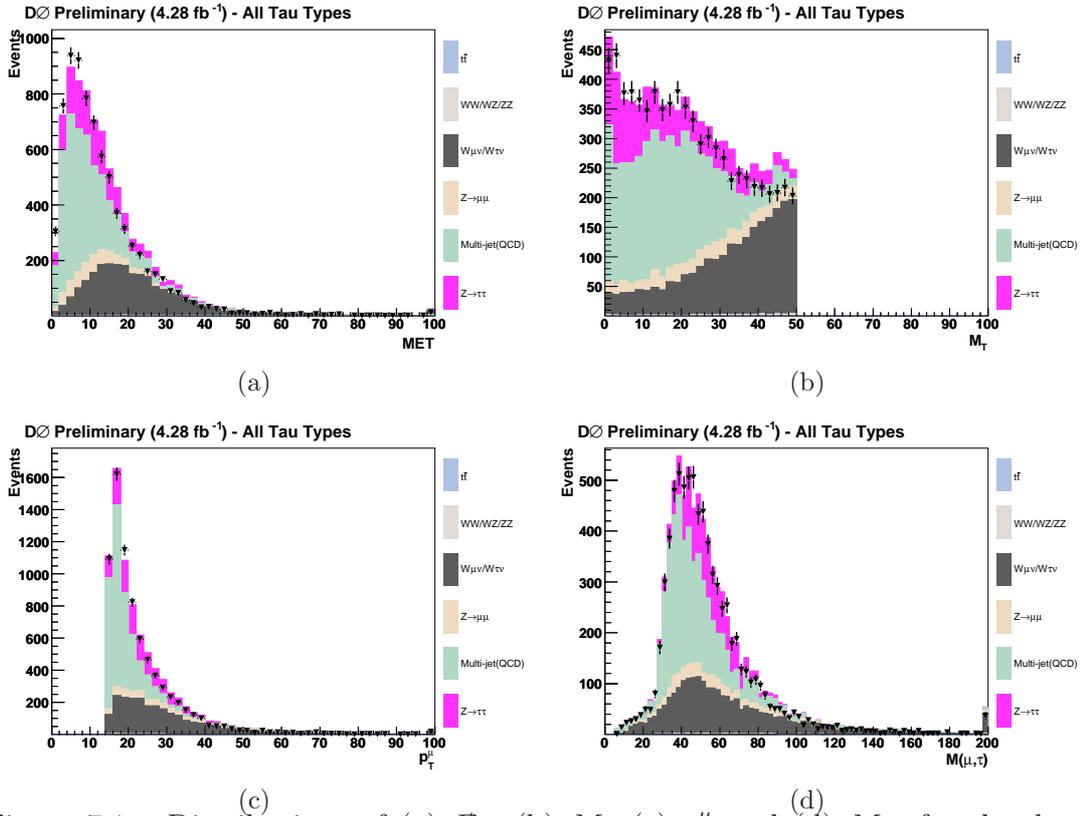

Figure 7.1: Distributions of (a) $\not{E}_T$, (b) M_T (c) p_T^μ and (d) $M_{\mu\tau}$ for the data compared to the sum of the expected backgrounds in the mNN sample with the multijet shape estimated from SS data.

For determining the systematic uncertainty of the multijet background estimation, a second method is considered. While the main method described in
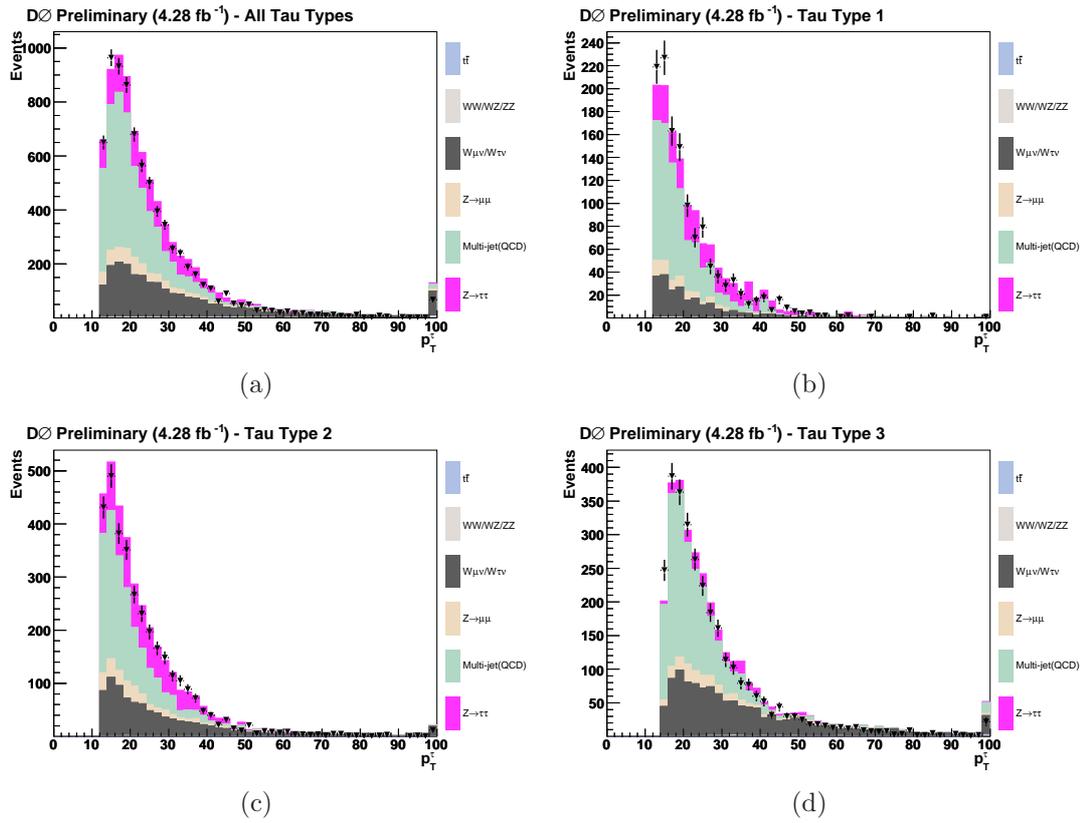
D∅ Preliminary (4.28 fb^{-1}) - All Tau Types

D∅ Preliminary (4.28 fb^{-1}) - Tau Type 1

D∅ Preliminary (4.28 fb^{-1}) - Tau Type 2

D∅ Preliminary (4.28 fb^{-1}) - Tau Type 3

(a)

(b)

(c)

(d)

Figure 7.2: Distributions of p_T^τ for data compared to the sum of the expected backgrounds for (a) sum of all τ types, (b) type-1 , (c) type-2 and (d) type-3 in mNN sample with the multijet shape estimated from SS data.

Sect. 7.2.1 relies primarily on the shape taken from the SS sample in the hNN region, the cross-check method uses OS events with medium NN_τ values to model the shape. The ratio of the number of SS events with high NN_τ values to the number of events with medium NN_τ values is subsequently used to normalise the selected mNN events to predict the total number of multijet events composed in the signal region. The same requirements as those for the preselection (see Chapter 8) are applied to select mNN multijet events.

The number of multijet events in the sample within the signal region, $N_{hNN}^{OS,QCD}$, is assumed to be given by

$$N_{hNN}^{OS,QCD} = \rho_{CC} \times (N_{mNN}^{OS,data} - N_{mNN}^{OS,MC}), \tag{7.4}$$

where $N_{hNN}^{OS,QCD}$ is the number of the multijet events expected in the OS sample within the high NN_τ region, i.e., the signal region. $N_{mNN}^{OS,data}$ and $N_{mNN}^{OS,MC}$, are the total number of data and MC events expected in the OS medium NN_τ region multijet enriched sample, normalised by the factor ρ_{CC}. All physics background MC samples are used in the estimation. Furthermore, ρ_{CC} is determined by selecting events enriched with multijet processes from a SS data sample and taking the ratio:

$$\rho_{CC} = \frac{N_{hNN}^{SS,data} - N_{hNN}^{SS,MC}}{N_{mNN}^{SS,data} - N_{mNN}^{SS,MC}}, \tag{7.5}$$

where $N_{mNN}^{SS,data}$ and $N_{mNN}^{SS,MC}$, are the total numbers of SS data and MC events, respectively, in the medium NN_τ region. Similarly, the total number of SS data and MC events in the high NN_τ region, $N_{hNN}^{SS,data}$ and $N_{hNN}^{SS,MC}$, are used to estimate the number of multijet events in the signal region. These values are summarised by τ type in Table 7.2. Furthermore, Table 7.3 lists the number of multijet events estimated by the two methods. The difference in the yield per τ type is subsequently used to set the systematic uncertainty on the multijet background estimation. A systematic uncertainty of 8.3%, 16.2%, and 13.8% is assigned for τ type 1, 2, and 3, respectively.

	type-1	type-2	type-3
ρ_{QCD}	1.020 ± 0.008	1.026 ± 0.005	1.103 ± 0.019
ρ_{CC}	0.182 ± 0.018	0.245 ± 0.015	0.150 ± 0.012

Table 7.2: Normalisation factors of the two methods by each τ type in the $\tau_{mu}\tau_h$ channel. The errors are statistical.
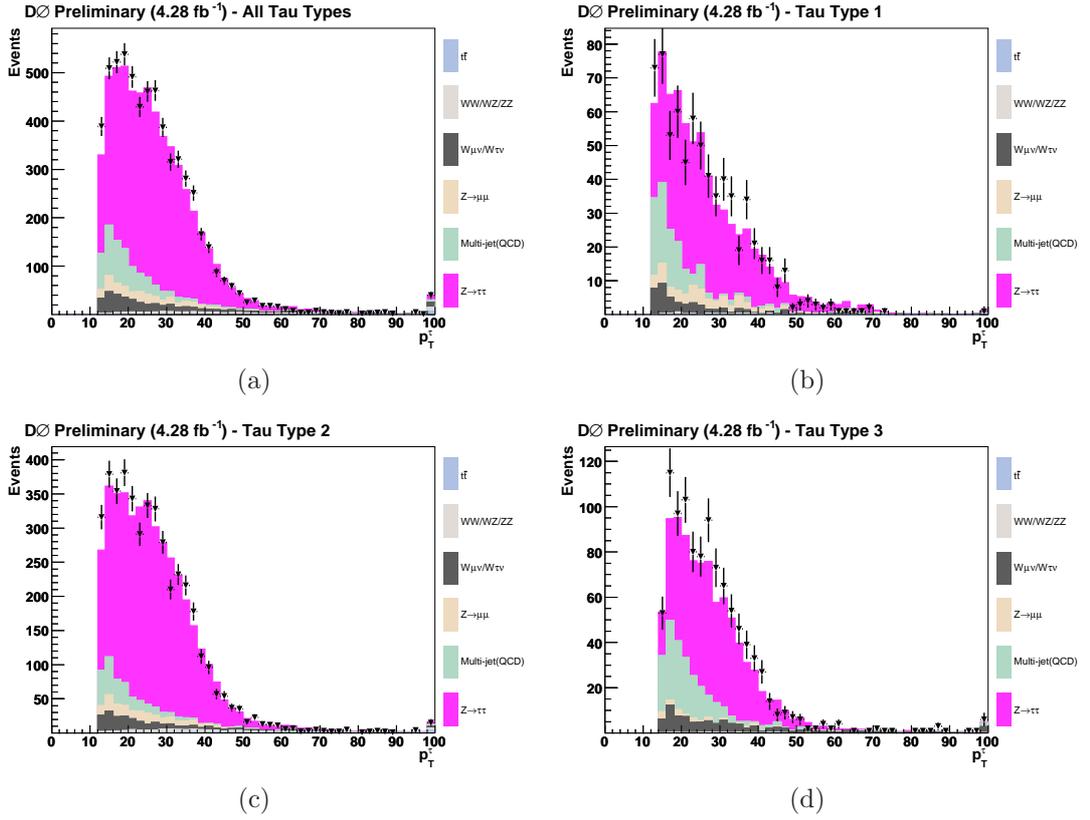

Figure 7.3: Distributions of τ transverse momentum for data compared to the sum of the expected backgrounds for (a) sum of all types, (b) type-1, (c) type-2 and (d) type-3 τ leptons in the SS sample with the multijet shape estimated from mNN data.

Number of Multijet Events	type-1	type-2	type-3
Main Method (SS-OS)	121 ± 11	351 ± 19	224 ± 16
Cross-check Method	110 ± 10	289 ± 17	196 ± 14

Table 7.3: Numbers of multijet events for the two different methods described in the text. The errors shown are statistical.

108

7.3   Background Modelling in the $\tau_e\tau_h$ Channel

In the $\tau_e\tau_h$ channel, one electron and one τ lepton are selected. However, the detector signatures of electrons and τ leptons can be easily mis-identified, and so special treatments for the W+jets and $Z \to e^+e^-$ background modelling are needed within this search channel. The method to properly handle such backgrounds is listed below:

- $\boldsymbol{W}$**+jets**: Since PYTHIA is a leading order generator, higher order effects must be considered by applying normalisation factors for W+jets background composition. These factors are derived in an OS sample enriched with W+jets events. The main selections to suppress the multijet contribution in this sample are:

    - a series of preselections as used to select the signal sample (Sect. 8.1),

    - $\not{E}_T \geq 20$ GeV,

    - transverse momentum of electrons $p_T^e \geq 20$ GeV.

    Since no M_T cut is applied, background events from multijet processes are expected to contribute significantly at low M_T even after this selection. A fit of the multijet background with low values of M_T is performed and extrapolated to the higher M_T region. The factor, ρ_W, based on the comparison of data with MC is derived in this sample after multijet and other sources of background are subtracted from data events. The formula of deriving ρ_W is given by

$$\rho_W = \frac{N_{data} - N_{bkg}^{NoW}}{N_{MC}^W}, \tag{7.6}$$

    where N_{bkg}^{NoW} is the total number of predicted events without W+jets, N_{MC}^W is the W+jets prediction from MC, and N_{data} is the total number of data events. This factor is derived in the OS region from events passing the final selections. A similar study is also done in the SS region in order to cross-check the normalisation factors derived in both SS and OS samples. The two factors per τ type are listed in Table 7.4 [67].

- $\boldsymbol{Z \to e^+e^-}$: Since reconstructed electrons are required to have only one track and narrow calorimeter clusters, they can also be reconstructed as τ

τ type	ρ_W^{SS}	ρ_W^{OS}
type-1	2.99 ± 0.23	1.88 ± 0.07
type-2	1.73 ± 0.03	1.36 ± 0.02
type-3	0.96 ± 0.01	0.94 ± 0.01

Table 7.4: The normalisation factors for the W background for both OS (ρ_W^{OS}) and SS (ρ_W^{SS}) events. Errors are only statistical.

type-2 candidates. In addition, when electrons fall in the ICD region, which has no EM coverage, they can be mis-identified as τ type-1 candidates. Furthermore, if electrons are associated with other tracks, they can also be reconstructed as type-3 candidates. In order to model the fake electron background, the efficiency correction described in Sect. 6.2 is first applied to MC samples. Next, a series of selections is needed to remove the electron contamination within the candidate sample of τ leptons. More details about these anti-electron selections can be found in Sect. 8.2.2.

Cross-checks of the multijet background for the $\tau_e \tau_h$ channel are discussed in the following Sect. 7.3.1, and further comparisons of data with the predicted background can be found in Chapter 8. The remaining physics background sources can be represented well by `PYTHIA` normalised to the integrated luminosity $\mathcal{L}$ and number of generated events passing the data quality checks, N_{DQ}, as described in Sect. 6.5.2.

7.3.1 Multijet in the $\tau_e \tau_h$ Channel

A similar method to estimate the multijet background is also applied to the $\tau_e \tau_h$ channel. It is derived independently for each τ type, and the values for ρ_{QCD} for this decay mode are listed in Table 7.5. The definition of ρ_{QCD} can be found in Sect. 7.1. These factors are calculated using `PYTHIA` MC samples for other well-modelled background sources. In order to study whether the shape of the estimated multijet background describes the data well, events are compared to the expectation from a sum of MC background and multijet events in a sample restricted to the low NN_τ region, which is highly enriched by such multijet events. These distributions are shown in Fig. 7.4 for the p_T^τ by each τ type. Furthermore, Fig. 7.5 shows the transverse momentum (p_T^e) and pseudo-rapidity (η^e) of the electron, $\not{E}_T$ and M_T. As observed in these distributions, the samples are

110

dominated by multijet events, and each shows consistent agreement between the background estimation and the data. After the checks of multijet background modelling and other sources of background simulated by MC are done in the control regions, the same method is applied to the signal regions which provides predictions for shapes and numbers of events of multijet background.

	type-1	type-2	type-3
ρ_{QCD}	1.17	1.05	1.03

Table 7.5: Normalisation factor for each τ type applied in the Run IIa $\tau_e\tau_h$ channel.

(a)

(b)

(c)
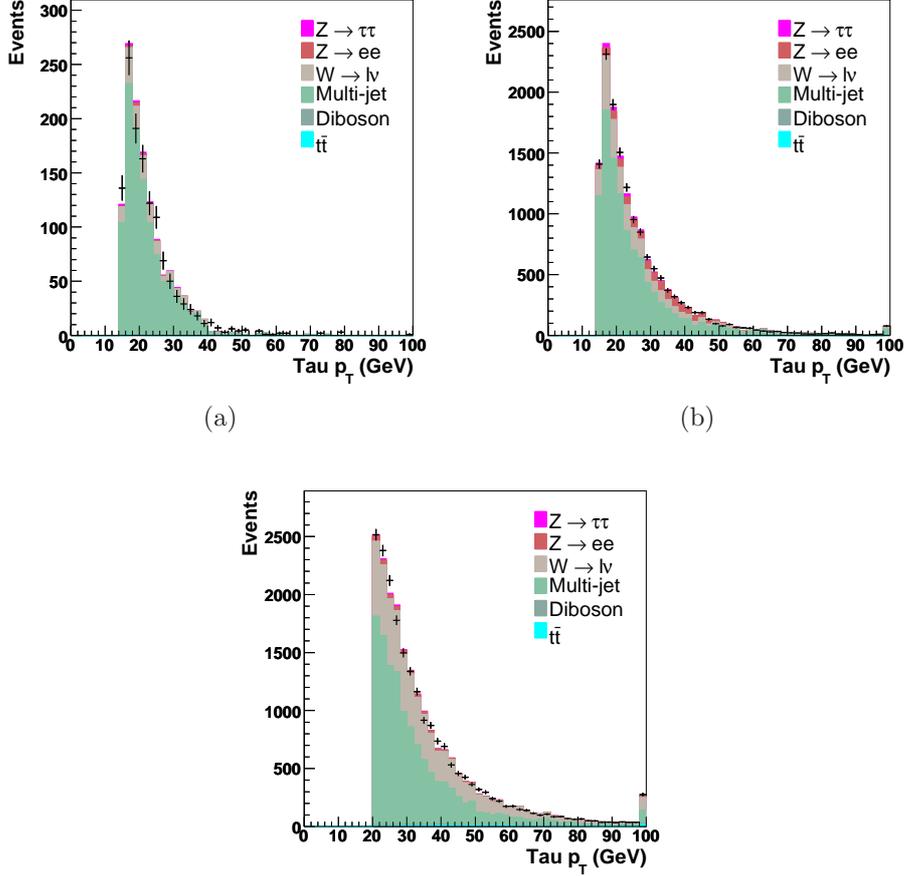
Figure 7.4: Distributions of p_T^τ of events in the low NN_τ region for (a) type-1, (b) type-2 and (c) type-3 τ candidates.

111

(a)

(b)

(c)

(d)
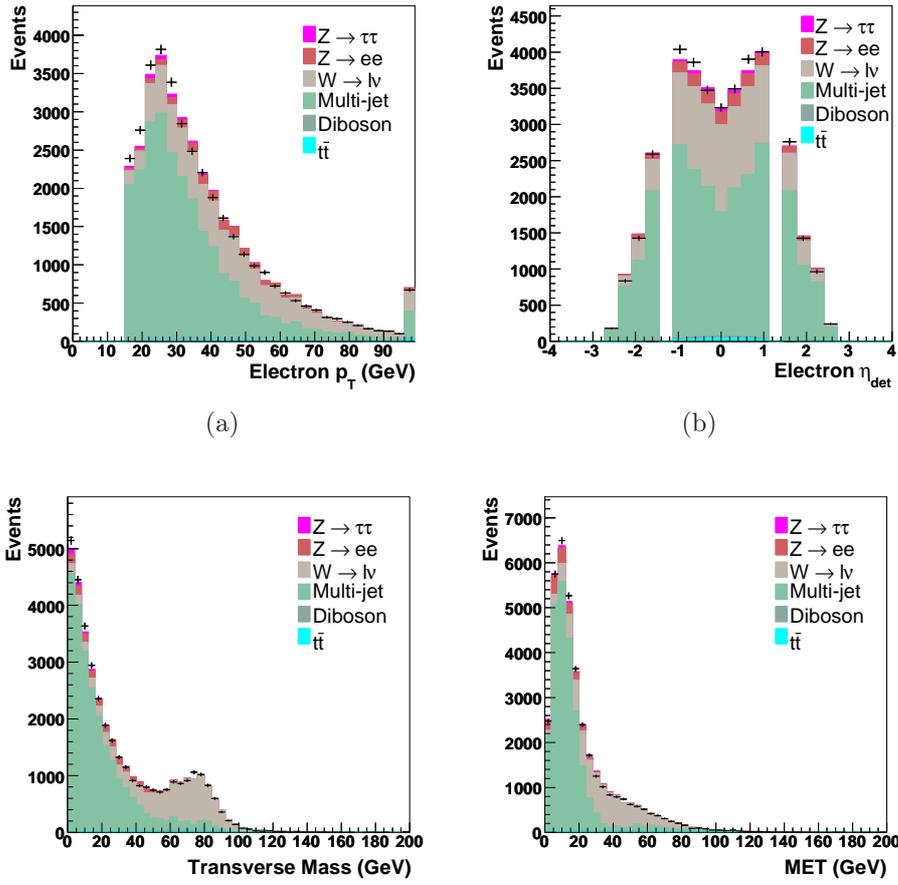
Figure 7.5: Distributions of (a) p_T^e, (b) η^e, (c) M_T, and (d) $\not{E}_T$ for events populating the low NN_τ region.

Chapter 8

Event Selection

8.1 Preselection Criteria

At the preselection stage, one electron or muon along with one hadronically decaying τ candidate are selected. The criteria for the $\tau_\mu\tau_h$ and $\tau_e\tau_h$ channels are described separately in the following sections. Table 8.1 provides a summary of the preselection requirements imposed in each channel. The studies concerning background modelling, including the efficiency and energy corrections described in Chapter 6 and 7, are incorporated at this stage. Final selections are applied in addition to the initial selections to improve signal to background yield and described in Sect. 8.2.

8.1.1 $\tau_\mu\tau_h$ Channel

The search for a Higgs boson in this final state begins with a set of preselection requirements involving one isolated muon arising from the leptonic decay of one τ candidate plus another τ decaying hadronically. The muon is reconstructed by associating hits in the muon detector with a central track. It is required to be of `medium` quality and matched to a `trackmedium` (Run IIa) or `tracknewmedium` (Run IIb) quality track as defined in Sect. 4.1.5. At the preselection stage, muon candidates with transverse momentum $p_T > 15$ GeV are kept and must lie within the range $|\eta| < 1.6$, where η is the detector pseudo-rapidity of the muon system as defined in Chapter 3. This fulfils the muon trigger requirements (see Sect. 5.2). The muon must be isolated according to the `NPTight` quality requirement. Only one muon of such quality is allowed per event. Events are vetoed if they contain

113

	$\tau_\mu \tau_h$ Channel			$\tau_e \tau_h$ Channel		
	τ type-1	τ type-2	τ type-3	τ type-1	τ type-2	τ type-3
e/μ Quality	medium			top_tight		
Number of e/μ	Exactly one			Exactly one		
Number of τ	Exactly one			Exactly one		
Lepton p_T (GeV) $\geq$	15			15		
$\Delta R(e/\mu, \tau_h) \geq$	0.4			0.5		
$\Delta R(jet, \tau_h) \geq$	0.5			0.5		
$\Delta z(\mu, PV) <$ (cm)	1			1		
Jet p_T (GeV) $\geq$	15			20		
τ p_T (GeV) $\geq$	12.5	12.5	15	15	15	20
τ p_T^{trk} (GeV) $\geq$	7	5	7	15	15	5
$NN_\tau \geq$	0.9	0.9	0.95	0.9	0.9	0.95
$\sum p_T^{trk} \geq$	–	–	10	–	–	15
$E_\tau / p_T^{trk} \geq$	0.65	0.5	0.5	–		
$R_\mu \geq$	–	–	0.3	–		

Table 8.1: Summary of preselection requirements.

an additional NPTight isolated, loose quality muon with $p_T > 12$ GeV. Such a requirement helps suppress contributions from $Z/\gamma^* \to \mu\mu$ background processes.

An event is required to contain a hadronic τ candidate at a distance $\Delta R > 0.4$ from the muon direction and with a sign for the electric charge opposite to that of the muon charge. In particular, the charge of the τ candidate is the sum over the charges measured from the curvature of the tracks associated with the τ candidate. Candidates are required to be found within $|\eta| < 1.6$, which is measured with respect to the centre of the detector. The transverse momentum of a τ candidate, p_T^τ, must be greater than 12.5 GeV for τ type-1 and type-2, and greater than 15 GeV for τ type-3. At the same time, the transverse momentum of the track associated with the τ candidate is required to be $p_T^{trk} > 7$ GeV for τ type-1 and $p_T^{trk} > 5$ GeV for τ type-2. In the case of τ type-3, one of the associated tracks must have $p_T^{sumtrk} > 7$ GeV and the scalar sum of the transverse momenta of all associated tracks is required to be greater than 10 GeV. Independent of τ types, at least one of the τ-tracks must have at least one SMT hit. The absolute distance along the z direction between the primary vertex (PV) and the track associated with either the muon or τ lepton, $|\Delta z(track_{\tau/\mu}, PV)|$, must be less than 1 cm.

Any contamination from cosmic ray muons is eliminated by requiring τ candidates to have $E_T^\tau / p_T^{trk} > 0.65$ (0.5) for τ type-1 (-2, -3). In order to remove mis-identified muons from τ type-1 candidates, a longitudinal shape variable,

114

$R_\mu = (E^\tau - E^{trk}_{CH})/p^{trk}_T$, is defined where E^{trk}_{CH} is the energy deposited around the τ-track extrapolated to the coarse hadronic (CH) section of the calorimeter. Because muons have a R_μ distribution that peaks at much lower values than type-1 τ candidates, this search requires $R_\mu > 0.3$ for all events.

Additionally, an energy correction for τ candidates, derived from a pure $Z \rightarrow \tau\tau$ sample in data and based on the ratio of the calorimeter energy, E^{cal}_T to track momentum, p^{trk}_T, is applied to MC to better match the τ energy spectrum in data. Complete details on such energy scale corrections are given in Sect. 6.4 of this thesis. A selection is made on NN_τ to separate real τ leptons from those that originate from jet fakes, with $NN_\tau > 0.9$ for types 1 and 2 , $NN_\tau > 0.95$ for type-3.

Since the calibration and therefore the energy scale differ for many of the reconstructed objects, the missing transverse energy, $\slashed{E}_T$, is corrected separately for tight electrons and jets with $p_T > 20$ GeV. Furthermore, these jets must not match to either tight electron or τ candidates within $\Delta R < 0.5$. Additionally, events are required to have no electrons with $p_T > 12$ GeV that pass the `top_loose` requirements. This ensures that this selection is orthogonal to the $\tau_e\tau_\mu$ and $\tau_e\tau_h$ decay channels for future combinations. After the preselection stage, the data sample is expected to be dominated by $W \rightarrow \mu\nu$ and $W \rightarrow \tau\nu$ events in association with a jet (W+jets) and multijet background.

Kinematic distributions from Run IIb data at the preselection stage are shown in Fig. 8.1, including muon transverse momentum (p^μ_T), transverse mass (M_T), the cosine value of the angle between muon and missing energy, $cos\Delta\phi(\mu, \slashed{E}_T)$, and the invariant mass of all tracks (M^{trk}). All distributions show reasonable agreement between data and the predicted background. The same check is also done in Run IIa data to make sure predicted background describe the data well after all preselection criteria are applied. These distributions incorporate the proper MC efficiencies and τ energy scale corrections summarised in Chapter 6. The multijet events are estimated using the method described in Sec. 7.1 of this thesis.

8.1.2 $\tau_e\tau_h$ Channel

The preselections applied to the $\tau_e\tau_h$ channel are very similar to those of the $\tau_\mu\tau_h$ channel but tighter to help suppress larger expected background levels. One
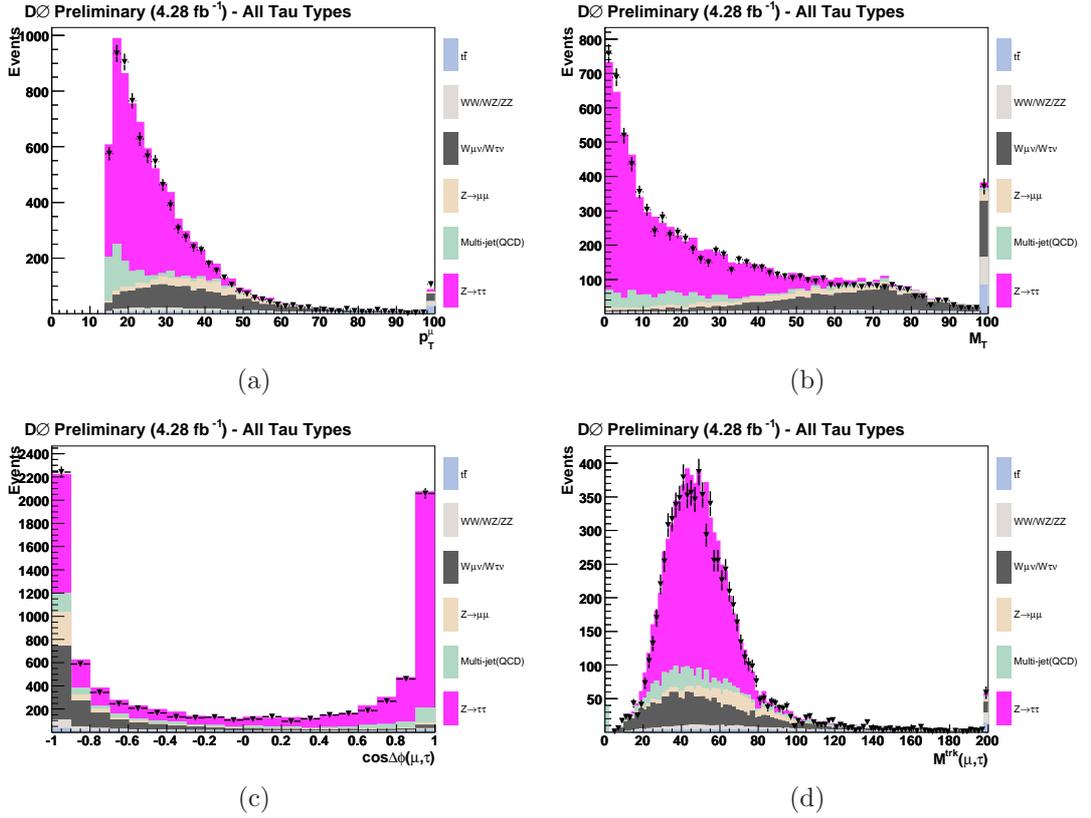
Figure 8.1: Distributions for data compared to the sum of the expected backgrounds for (a) p_T^μ (b) M_T (c) $cos\Delta\phi(\mu, \not{E}_T)$ (d) M^{trk} after the preselection.

electron with transverse momentum $p_T > 15$ GeV is selected if it fulfils the top_tight definition. This electron must lie within detector pseudorapidity $|\eta| <$ 2.5, and it is denoted in the following sections as a "tight electron". The event is required to fulfil any of the single electron trigger conditions listed in the logical "OR" of single electron triggers. The tight electron of the event is subsequently required to match one of the triggers that fired at all three trigger levels. If a second electron meeting these criteria is found in the same event, this event is rejected. The vetoing on a second tight electron is to reduce fakes from $Z \to e^+e^-$ background.

A τ candidate is required to be at a distance $\Delta R > 0.5$ from the electron direction. The transverse momentum of the τ candidate, p_T^τ, as reconstructed by the calorimeter, must be greater than 15 GeV for τ type-1 and type-2, and greater than 20 GeV for τ type-3. At the same time the transverse momentum of the track associated with the τ candidate is required to be $p_T^{trk} > 15$ GeV for τ type-1 and $p_T^{trk} > 5$ GeV for τ type-2. In the case of τ type-3, one of the associated

tracks must have $p_T^{trk} > 5$ GeV and the scalar sum of the transverse momenta of all associated tracks must be greater than 15 GeV. Type 3 τ candidates with only two associated tracks are rejected. As in the muon channel, the electric charge of the τ candidate is the sum over the charges measured from the curvature of the tracks associated with the τ candidate, and it must be opposite to the charge of the tight electron. Coarse hadronic calorimeter cells are only used if they are associated to a reconstructed jet. Since the calibration and therefore the energy scale are different for most of the reconstructed objects, the missing transverse energy, $\slashed{E}_T$, is corrected separately with energies of the tight electron and selected jets. Furthermore, these jets must not be matched to either the tight electrons and the τ candidates applying the requirement $\Delta R < 0.5$.

Events after preselections are dominated by $Z \to e^+e^-$ background, but a larger discrepancy is observed due to the mis-modelling of the electron efficiency in MC. Figure 8.2 shows the detector pseudo-rapidity distributions (η^τ) for τ types 1, 2 and 3.

8.2 Final Selection Requirements

Final selection criteria are applied to improve the expected signal -to-background yield, and in particular to reduce $Z \to \ell\ell$ ($\ell = e$ or μ) and W+jets backgrounds. Since background from $Z/\gamma^* \to \tau\tau$ is almost irreducible except for differences in its mass and spin, *i.e.,* the angular distribution, with respect to the Higgs boson, no attempt is made to remove contributions from $Z/\gamma^* \to \tau\tau$.

8.2.1 $\tau_e\tau_h$ Channel

The background events from $Z \to \mu\mu$ can effectively be suppressed by applying a matching requirement and rejecting events with $\Delta R(\mu, \tau) < 0.5$. Therefore, a cut on M_T is the only additional requirement after the preselection stage. In this channel, events with the transverse mass m_T, as defined in Eq. 7.1, must be less than 50 GeV for each τ type. This selection removes a significant portion of the W+jets background.

Table 8.2 lists the yields from data compared to the expected backgrounds for each of the three τ types and the sum of all types, respectively. Furthermore, Fig. 8.3 shows the detector pseudorapidity and azimuthal ϕ distributions for the
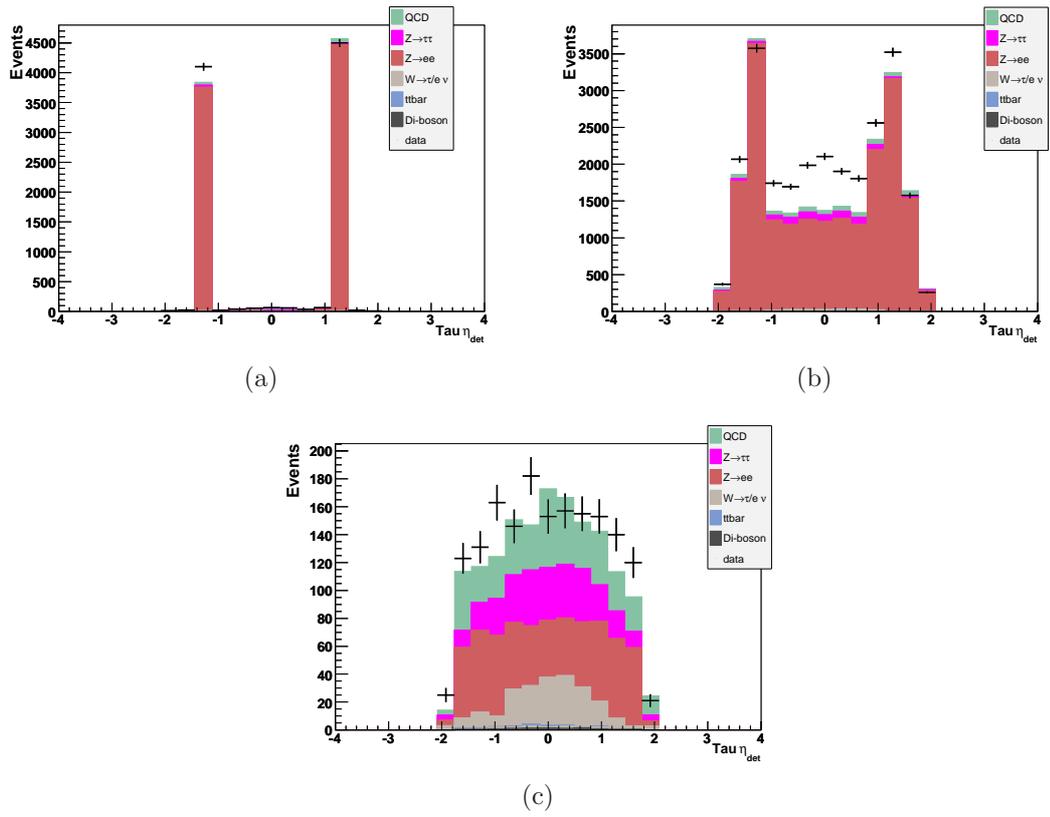
(a)

(b)

(c)

Figure 8.2: Distributions of η^τ for (a) type-1, (b) type-2 and (c) type-3 τ candidates.

τ candidates after the M_T cut. Figures 8.4 and 8.5 provide the missing transverse energy and the τ-track p_T distribution for the τ candidates after imposing the M_T cut, where the scalar sum of track momenta is used for τ type-3 candidates. All distributions are presented for the Run IIb data set. Within statistics, the predicted backgrounds are consistent with the number of events in data for each τ type. Similar checks for the data-background consistency are also done for the Run IIa data set, and the yields are provided in Table 8.3.

τ type	type-1	type-2	type-3	all types
Data	744	4477	1049	6270
Total Prediction	789 ± 28	4642 ± 68	1021 ± 32	6452 ± 81
$Z \to \tau^+\tau^-$	567 ± 24	3741 ± 61	683 ± 26	4991 ± 71
$Z \to \mu^+\mu^-$	52 ± 7	194 ± 14	18 ± 4	264 ± 16
W+jets	47 ± 7	220 ± 15	90 ± 9	357 ± 19
$WW/WZ/ZZ$	7 ± 2	88 ± 10	3 ± 1	98 ± 10
$t\bar{t}$	3 ± 1	47 ± 7	2 ± 1	52 ± 7
Multijet	121 ± 11	351 ± 19	224 ± 16	696 ± 27
Signal Eff. (%) (Mass 120 GeV)	0.109 ± 0.011	0.537 ± 0.024	0.127 ± 0.013	–

Table 8.2: Prediction from MC physics processes and multijet background taken from data compared to the observation in Run IIb data for each τ type and the sum of all types. Only statistical errors on the MC and multijet are given.

τ type	type-1	type-2	type-3	all types
Data	357	1746	482	2585
Total Prediction	330 ± 18	1756 ± 42	484 ± 22	2576 ± 51
$Z \to \tau^+\tau^-$	228 ± 15	1429 ± 38	322 ± 18	1979 ± 45
$Z \to \mu^+\mu^-$	15 ± 4	71 ± 8	11 ± 3	97 ± 10
W+jets	12 ± 4	65 ± 8	39 ± 6	116 ± 11
$WW/WZ/ZZ$	2 ± 1	31 ± 6	2 ± 1	35 ± 6
$t\bar{t}$	< 1	1 ± 1	< 1	1 ± 1
Multi-jet	73 ± 9	159 ± 12	110 ± 11	342 ± 18
Signal Eff.(%)×100 (Mass 120 GeV)	0.106 ± 0.074	0.661 ± 0.185	0.134 ± 0.083	–

Table 8.3: Prediction from MC physics processes and multijet background taken from data compared to the observation in Run IIa data for each τ type and the sum of all types. Only statistical errors on the MC and multijet are given.
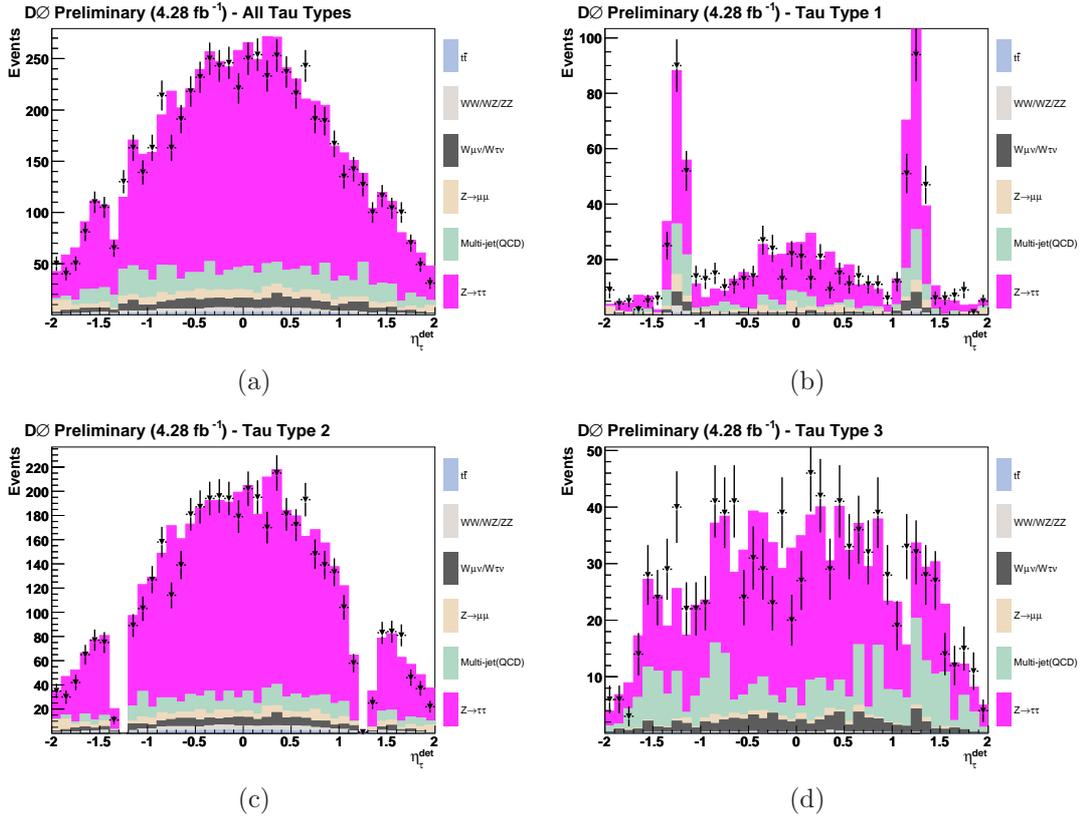
Figure 8.3: Distributions for the $\tau_\mu\tau_h$ channel for Run IIb data compared to the sum of the expected backgrounds for τ detector pseudorapidity, η^τ for (a) sum of all τ types, (b) type-1, (c) type-2 and (d) type-3 τ candidates after $M_T < 50$ GeV cut.
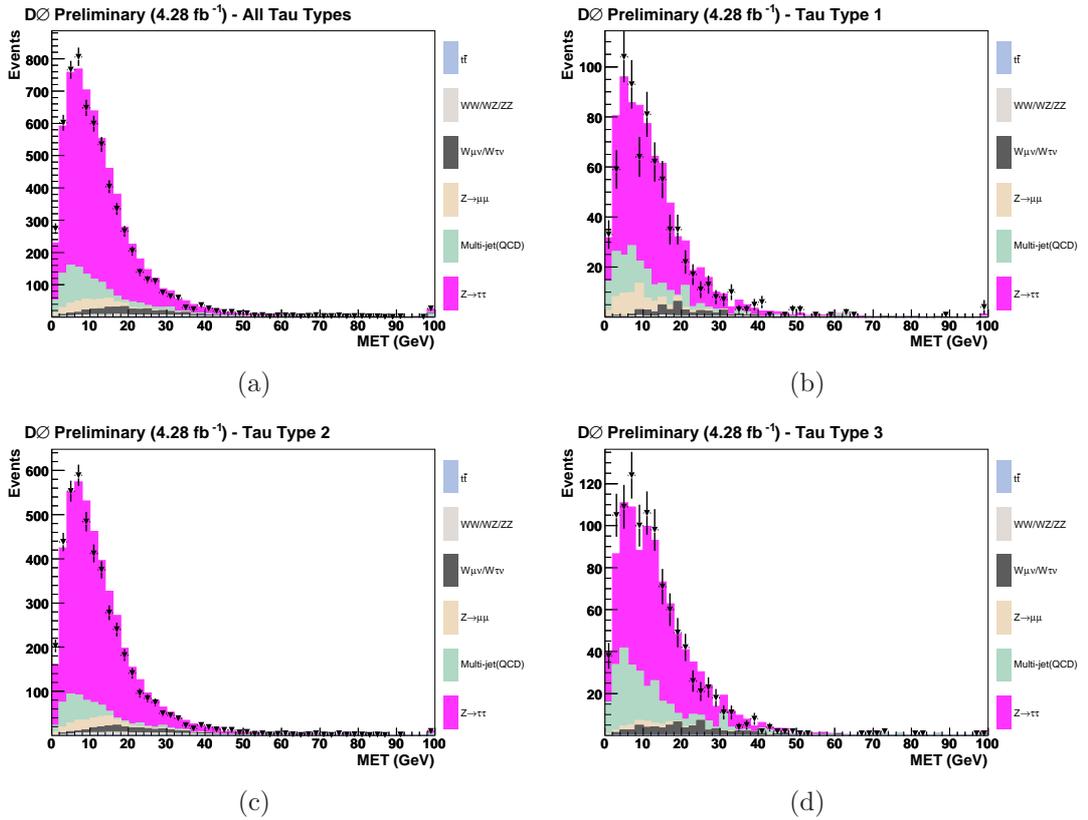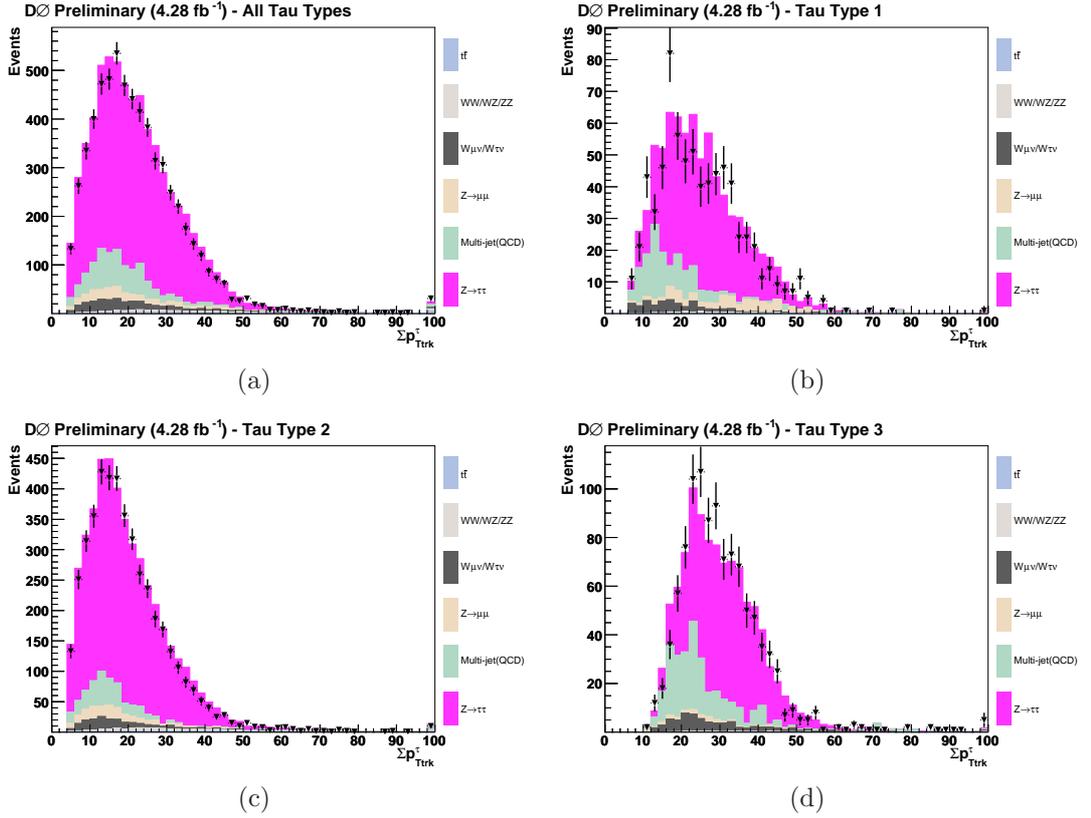
Figure 8.4: Distributions for the $\tau_\mu \tau_h$ channel for Run IIb data compared to the sum of the expected backgrounds for $\not{E}_T$ for (a) sum of all τ types, (b) type-1, (c) type-2 and (d) type-3 τcandidates after the $M_T < 50$ GeV cut.

Figure 8.5: Distributions for the $\tau_\mu\tau_h$ channel for Run IIb data compared to the sum of the expected backgrounds for τ track p_T for (a) sum of all τ types, (b) type-1, (c) type-2 and (d) type-3 τcandidates after the $M_T < 50$ GeV cut. For τ type-3 candidates, the scalar sum of the transverse momentum of the three tracks is used.

8.2.2 $\tau_e\tau_h$ Channel

A series of cuts are applied to further suppress $Z \to e^+e^-$ and $W+$jets background. Since there is no EM coverage in the ICD region, incident electrons with the reconstructed signature of tracks and hadronic clusters can easily be misidentified as τ leptons. Therefore, τ candidates in the region $1.05 < |\eta| < 1.55$ are not considered for τ type-1 and type-2. Also, it is found that the efficiency of detecting electrons near the boundary of modules of the EM calorimeter is not as effective as for fiducial towers of the calorimeter. This increases the total amount of electrons faking τ candidates. The variable $f_{\phi PS}$ is defined as [24]:

$$f_{\phi PS} = \text{mod}\left(\frac{\phi_{\tau\text{trk}}^{PS}}{\left(\frac{2\pi}{32}\right)}\right), \tag{8.1}$$

where $\phi_{\tau trk}^{PS}$ is the azimuthal angle of the track associated with the τ at the radius of the pre-shower detector. By construction, values of this variable are limited to the range from 0 to 1, where events peaking at either end are deposited close to a calorimeter module boundary. Consequently, type-2 τ candidates are required to have $0.1 < f_{\phi PS} < 0.9$ to avoid collecting electrons in inefficient parts of the detector.

After ICD and EM calorimeter crack removal, distributions for NN_{elec} for type-2 events and the invariant mass of the $e - \tau$ pair ($M(e,\tau)$) for type-3 are shown in Figure 8.6. These are subsequently used for deriving efficiency corrections. At this stage, there is a negligible number of electrons in type-1 events, and therefore, no attempt is made to further separate electrons from type-1 taus.

(a)   (b)
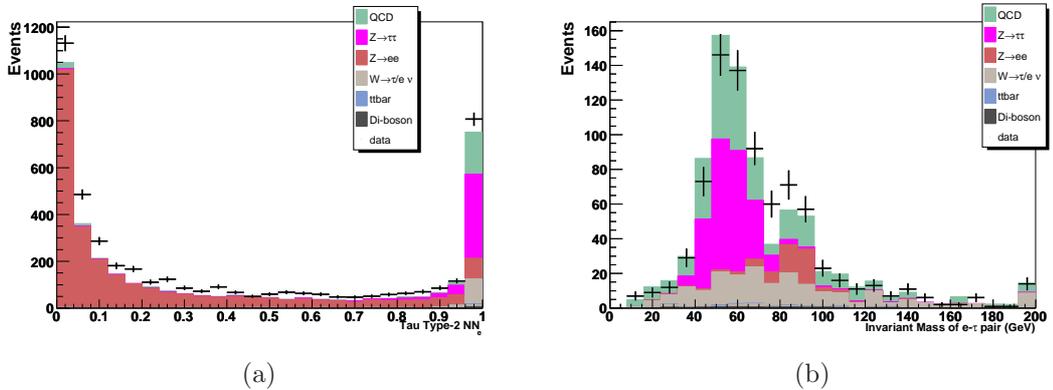
Figure 8.6: Distributions of (a) NN_{elec} for type-2 (b) $M(e,\tau)$ for type-3.

The efficiency correction for the sample containing only one tight electron sample, $k(X)$, as described in Sect. 6.2 is applied after ICD and EM calorimeter crack removal. Figure 8.7 shows electron transverse momentum (p_T^e) and pseudorapidity (η^e) distributions of electrons with correction applied. Furthermore, three kinematic distributions for τ candidates:

- τ transverse momentum (p_T^τ),

- pseudorapidity distribution for τ (η^τ),

- ratio of calorimeter cluster to transverse momentum of tracks (E^{cal}/P_T^{trk})

are shown by different τ types in Figs. 8.8 and 8.9.
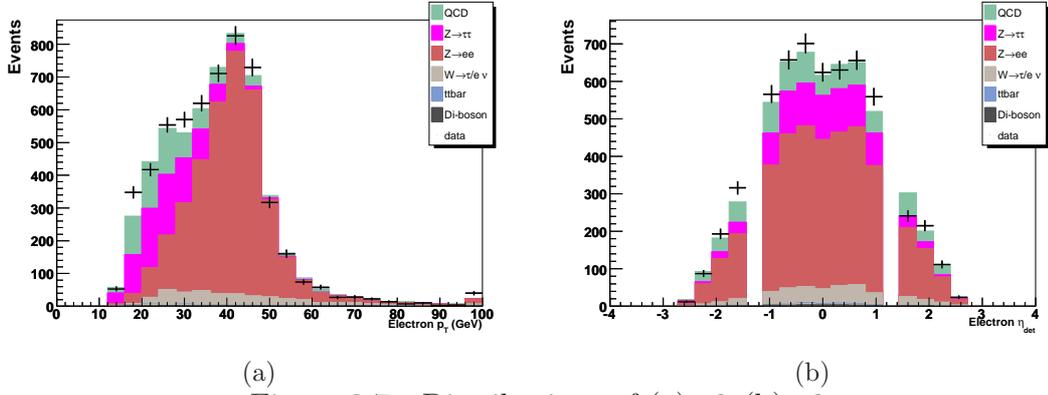
(a)   (b)

Figure 8.7: Distributions of (a) p_T^e (b) η^e.

After correcting the efficiency, the total background prediction describes the data well. As discussed in Sect. 4.1.6, electrons should peak at low NN_{elec} values towards 0, and therefore, τ type-2 candidates are required to have $NN_{elec} < 0.9$. With these three selection criteria: ICD, $f_{\phi^{PS}}$ and NN_{elec}, background events from $Z \rightarrow e^+e^-$ decays can be efficiently suppressed as shown in Fig. 8.10, which is the ratio of EM cluster to total calorimeter energy (f_{em}) for τ type-2 and 3 candidates.

Since electrons are expected to deposit most of the energy in EM layers, a final electron veto is to require events passing $f_{em} \leq 0.9$, where f_{em} is defined as the ratio of total energy from EM clusters to the sum of all clusters. The remainder of electrons in types 2 and 3 events can be rejected efficiently by requiring $f_{em} \geq 0.9$ This cut is not applied to τ type-1 events since there are no associated EM clusters. Figure 8.11 shows the distribution of transverse momentum (p_T^τ),

124

(a)

(b)

(c)
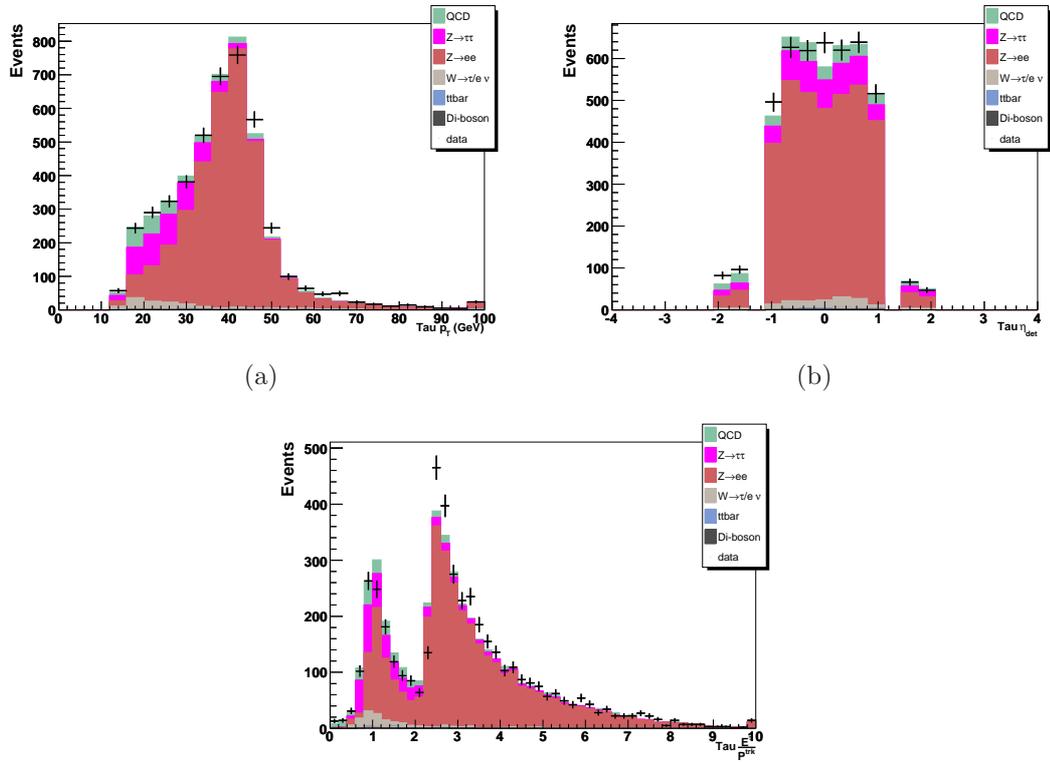
Figure 8.8: Distributions of (a) p_T^τ (b) η^τ and (c) E^{cal}/P_T^{trk} for τ type-2.

(a)

(b)

(c)
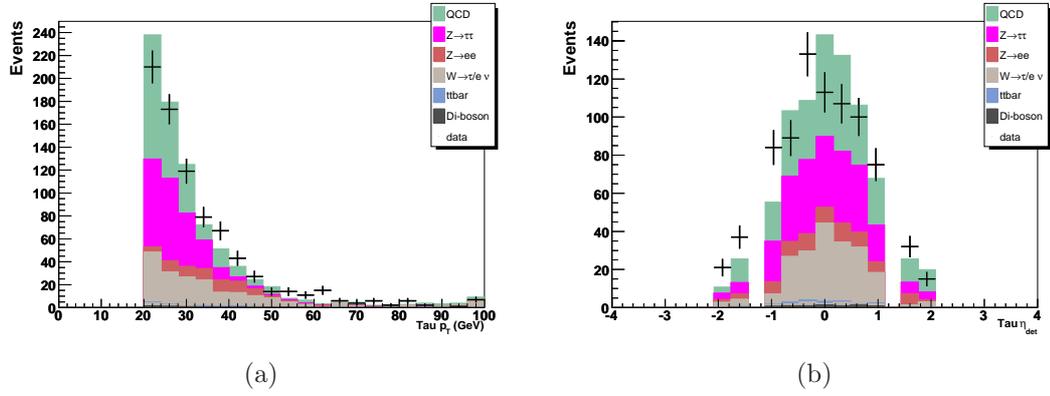
Figure 8.9: Distributions of (a) p_T^τ (b) η^τ and (c) E^{cal}/P_T^{trk} for τ type-3.
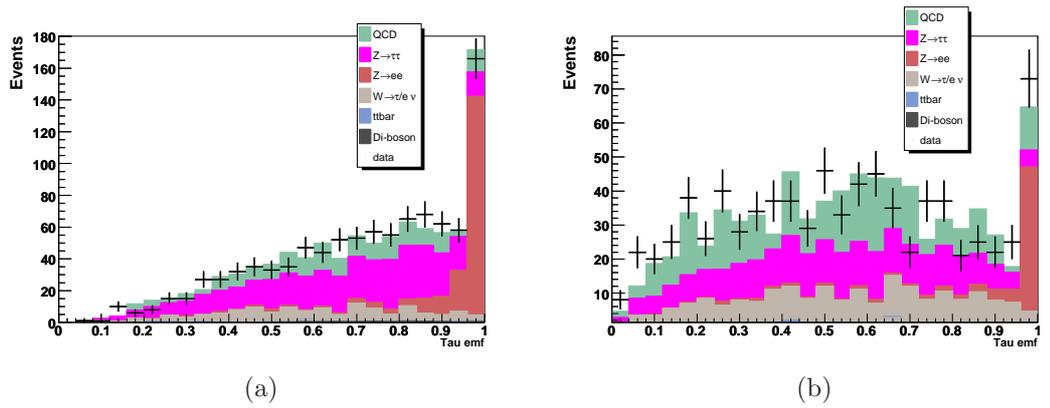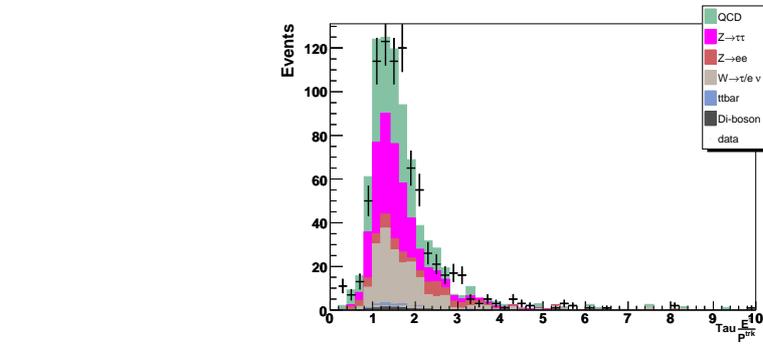

(a)

(b)

Figure 8.10: Distributions of electromagnetic fraction, f_{em}, for tau cluster depositions in the calorimeter after correcting electron efficiency.

126

detector pseudorapidity (η^τ), and the ratio of energy to the sum of the transverse momentum of tracks (E/p_T^{trk}) of the τ candidates. Similarly, Fig. 8.12 shows the distribution of transverse momentum (p_T^e) and detector pseudorapidity (η^e) of electrons. After these anti-electron selection requirements are applied, $Z \to e^+e^-$ events contribute little to the overall background content, and the remaining events are driven by multijet, W+jets and $Z \to \tau^+\tau^-$ backgrounds.

(a)

(b)

(c)
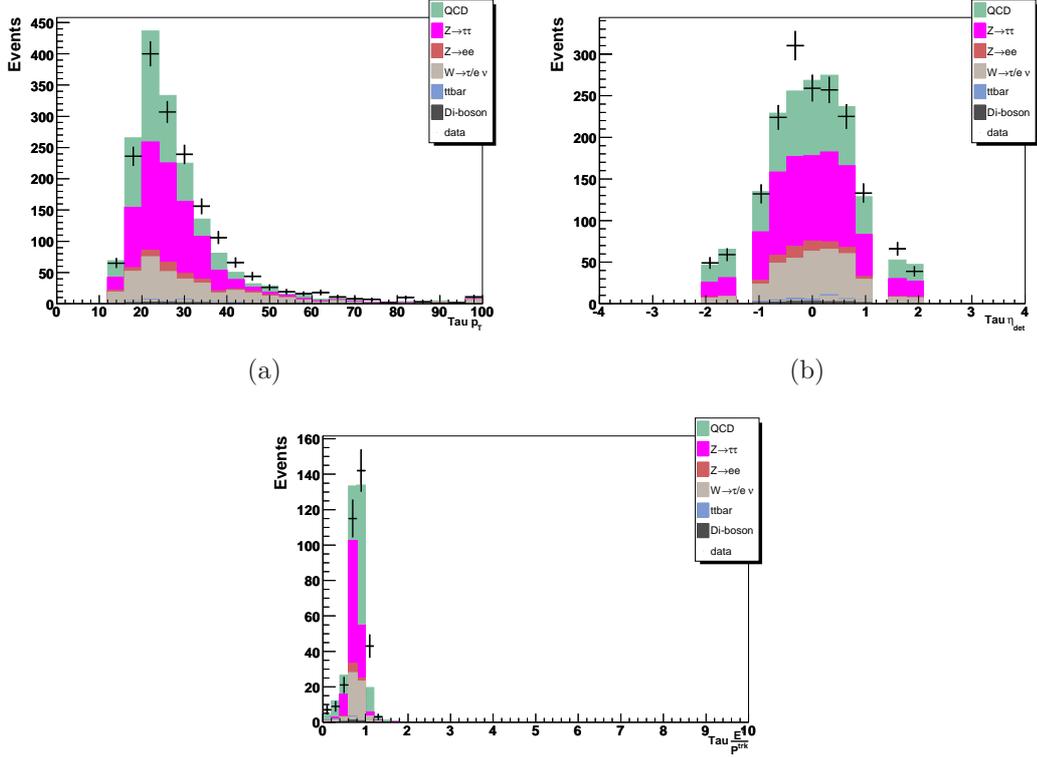
Figure 8.11: Distributions for the $\tau_e\tau_h$ channel for Run IIa data compared to the sum of the expected backgrounds for (a) p_T^τ (b) η^τ (c) E/p_T^{trk} for sum of all τ type events after the $Z \to e^+e^-$ vetos.

After vetoing events that contain electrons faking τ objects, a requirement is made on the M_T in order to reject much of the remaining W+jets background. Here, events with $M_T < 50$ GeV are kept for all three τ types. In addition, Fig. 8.13 shows the azimuthal angle between the electron and τ, $\Delta\phi(e, \tau)$, after the M_T cut. The signal peaks strongly at π, whereas the multijet background is more collimated. Events with $\Delta\phi(e, \tau) < 1.6$ are removed to further reject the multijet background without significantly reducing the signal efficiency. Finally, a cut is placed in the $\Delta\phi(e, \not{E}_T) - \Delta\phi(\tau, \not{E}_T)$ plane to further remove W+jet
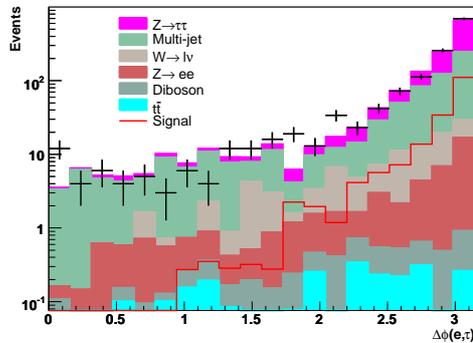
(a)	(b)
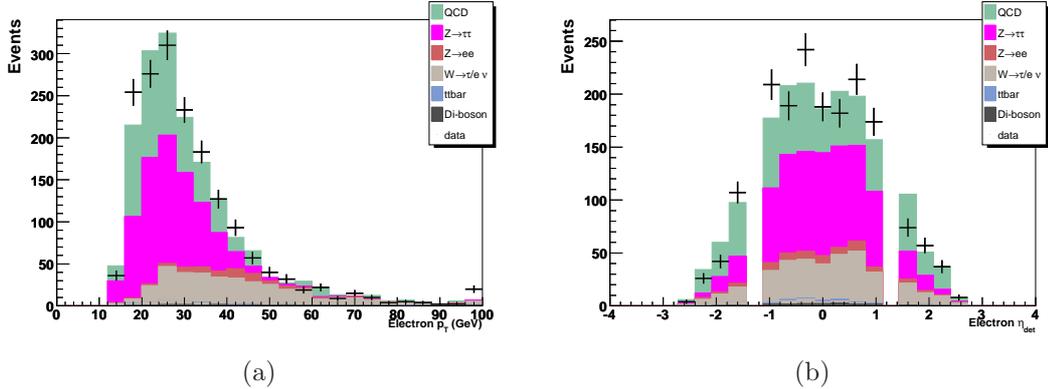
Figure 8.12: Distributions for the $\tau_e\tau_h$ channel for Run IIa data compared to the sum of the expected backgrounds for (a) electron p_T and (b) electron η_{det} after the $Z \to e^+e^-$ vetos.

events. The distributions for W+jet events and signal events at two different Higgs masses in this plane are shown in Fig. 8.14. Superimposed on each figure is a linear curve indicating the selection used to separate W+jet events from potential signal events. In particular, events appearing above the line will be rejected. Table 8.4 summaries the final selection criteria applied for this channel, and the event yields are listed in Table 8.5.

After the selection requirements are applied to both $\tau_e\tau_h$ and $\tau_\mu\tau_h$ channels, most of the background events are reduced significantly. However, events from the $Z \to \tau^+\tau^-$ process have very similar detector signatures to those of Higgs bosons, and these background events are difficult to be separated from signal. Therefore, a visible mass variable, M_{vis}, as defined in Sect. 6.4.1 is used to reconstructed the mass of the events and to search for a signal excess on top of the background mass distributions. More details about the search results can be found in Chapter 9.

Figure 8.13: Distributions in the $\tau_e\tau_h$ channel for Run IIa data compared to the sum of the expected backgrounds for the azimuthal angle between the electron and tau, $\Delta\phi(e,\tau)$, after the $M_T < 50$ GeV cut.
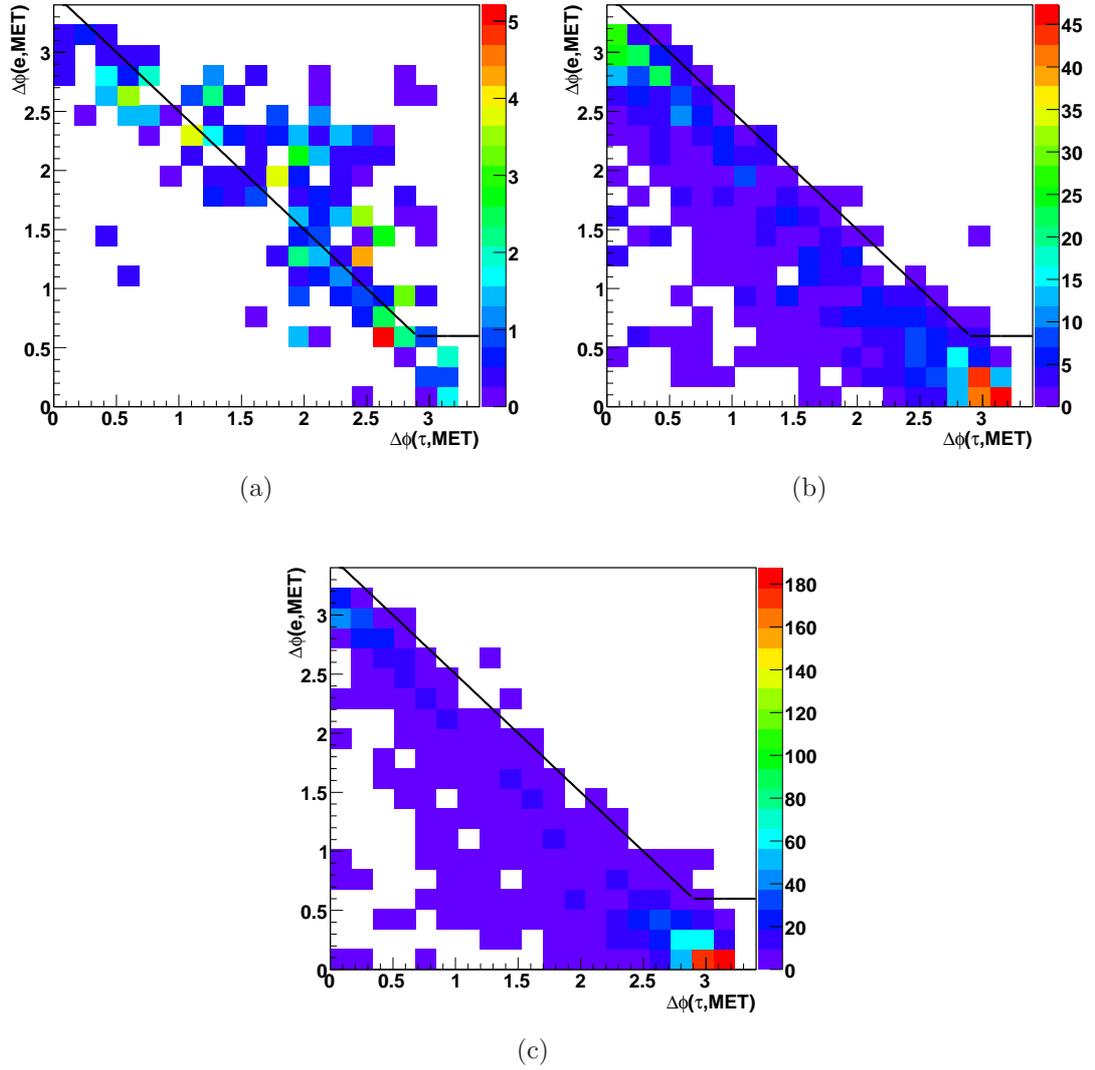
Figure 8.14: Distribution of events in the $\Delta\phi(e, \not{E}_T) - \Delta\phi(\tau, \not{E}_T)$ plane for (a) W MC events, (b) $M_h = 120$ GeV signal MC events and (c) $M_h = 200$ GeV signal MC events. Events above the black line are rejected.

Selection Cut	Reduced Source	τ type-1	τ type-2	τ type-3
Removal of ICD	$Z \to e^+e^-$	$|\eta| > 1.55$ OR $|\eta| < 1.05$	–	–
Crack Removal	$Z \to e^+e^-$	–	$0.1 < f_{\phi^{PS}} < 0.9$	–
$NN_e \geq$	$Z \to e^+e^-$	–	0.9	–
$f_{em} \leq$	$Z \to e^+e^-$	–	0.9	0.9
$M_T \leq$ (GeV)	W+jets	50		
$\Delta\phi(e,\tau) \leq$	Multi-jet	1.6		
$\Delta\phi(e,\not{E}_T) - \Delta\phi(\tau,\not{E}_T)$	W+jets	As shown in Fig. 8.14		

Table 8.4: The selection requirements which the τ candidates are required to pass in order to suppress $Z \to e^+e^-$ and W + jets backgrounds.

τ type	type-1	type-2	type-3	all types
Data	192	490	352	1034
Total Prediction	192 ± 10	453 ± 13	344 ± 14	989 ± 22
$Z \to \tau^+\tau^-$	100 ± 2	305 ± 4	176 ± 3	581 ± 5
$Z \to e^+e^-$	5 ± 2	20 ± 2	6 ± 1	31 ± 2
W+jets	7 ± 2	12 ± 3	22 ± 4	42 ± 5
$WW/WZ/ZZ$	< 1	1 ± 1	1 ± 1	2 ± 1
$t\bar{t}$	< 1	< 1	1 ± 1	1 ± 1
Multi-jet	80 ± 10	115 ± 12	138 ± 13	$332 \pm w0$

Table 8.5: Prediction from MC physics processes and multijet background events estimated from data compared to the observation in data for each τ type and sum of all types for Run IIa data of the $\tau_e\tau_h$ channel. Only statistical errors on the MC and multijet are given.

Chapter 9

Results

9.1 Final Discriminant

Due to the neutrinos present in the $\phi \to \tau^+\tau^-$ decay, it is difficult to reconstruct the full ditau mass. Furthermore, since the decay products of each of the two τ leptons originating from the Higgs boson are nearly back-to-back in (r,ϕ) plane at Tevatron energies, masses of Higgs bosons can not be fully reconstructed using the collinear approximation [68], where $\not{E}_T$ is projected alone the direction of visible τ decay products. Instead, a partial mass reconstruction is used as a final discriminating variable. The M_{vis} variable as defined in Sect. 6.4.2 allows searching for enhancements due to signal in the data sample. The M_{vis} distributions of Higgs bosons for events passing all final selections in the Run IIa data for the $\tau_\mu\tau_h$ channel are shown in Figs. 9.1 and 9.2. Similarly, Figs. 9.3 and 9.4 show M_{vis} for the Run IIb $\tau_\mu\tau_h$ channel. Also shown in the figure is the predicted Higgs boson signal multiplied by a factor of 50 in order to be visible. Similarly, M_{vis} distributions in the Run IIa data for the $\tau_e\tau_h$ channel are shown on a linear scale in Fig. 9.5. No significant excess in signal over background is observed, indicating the data are consistent with the background prediction.

9.2 Cross Section Limits

Since no significant excess in signal over background is observed, limits on the production cross section multiplied by the branching ratio for a neutral Higgs boson decaying to $\tau^+\tau^-$ pairs is calculated. The M_{vis} variable is used as the
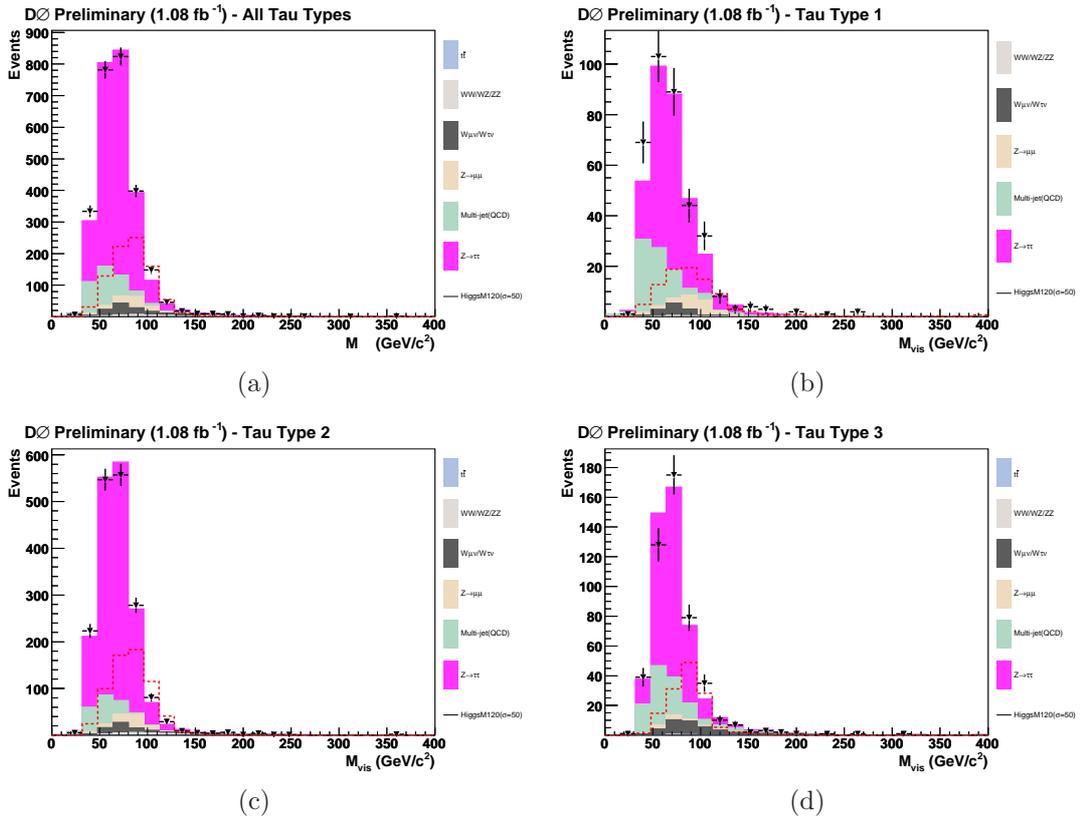
Figure 9.1: Distribution of M_{vis} on a linear scale in the Run IIa $\tau_\mu \tau_h$ channel for (a) sum of all τ types (b) type-1 (c) type-2 and (d) type-3. The data, shown with error bars, are compared to the sum of the expected backgrounds.
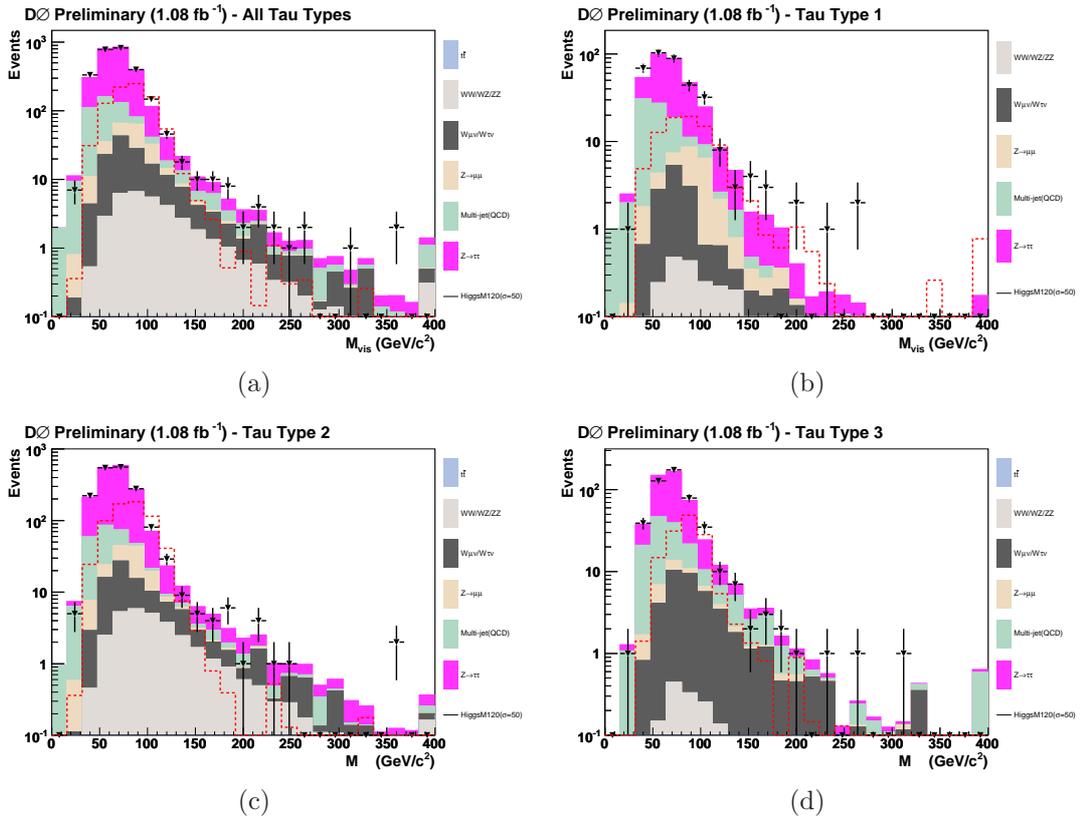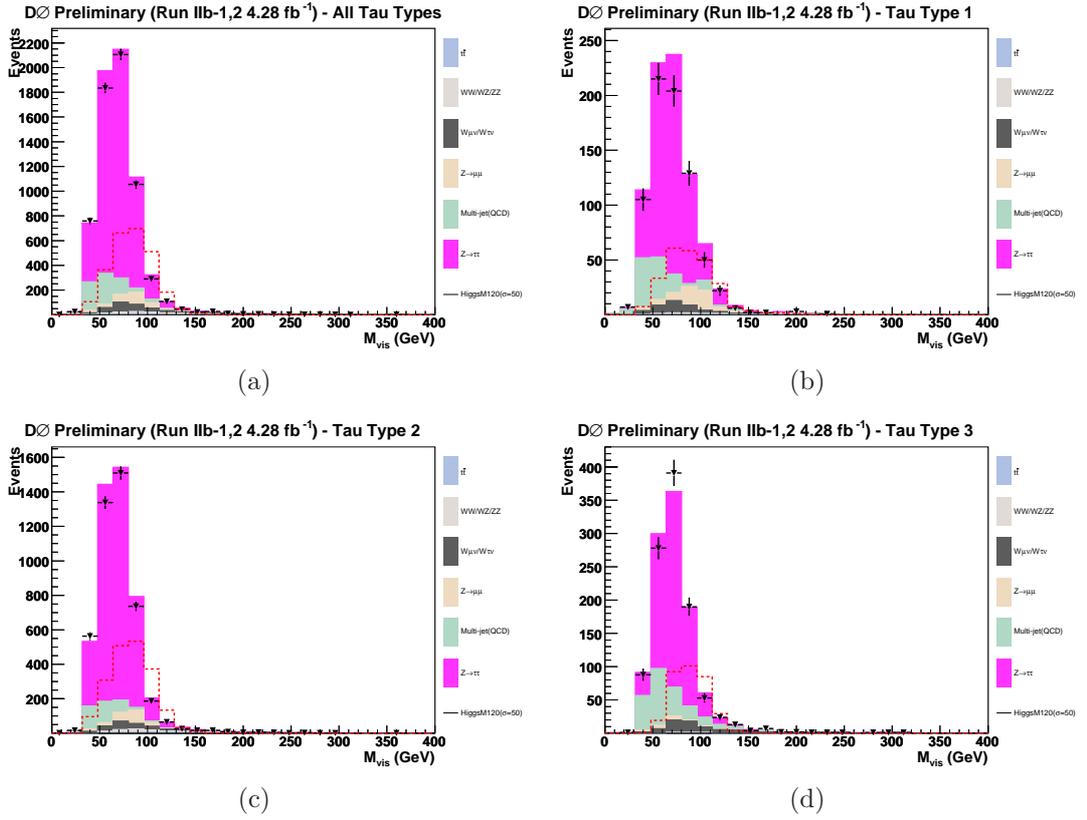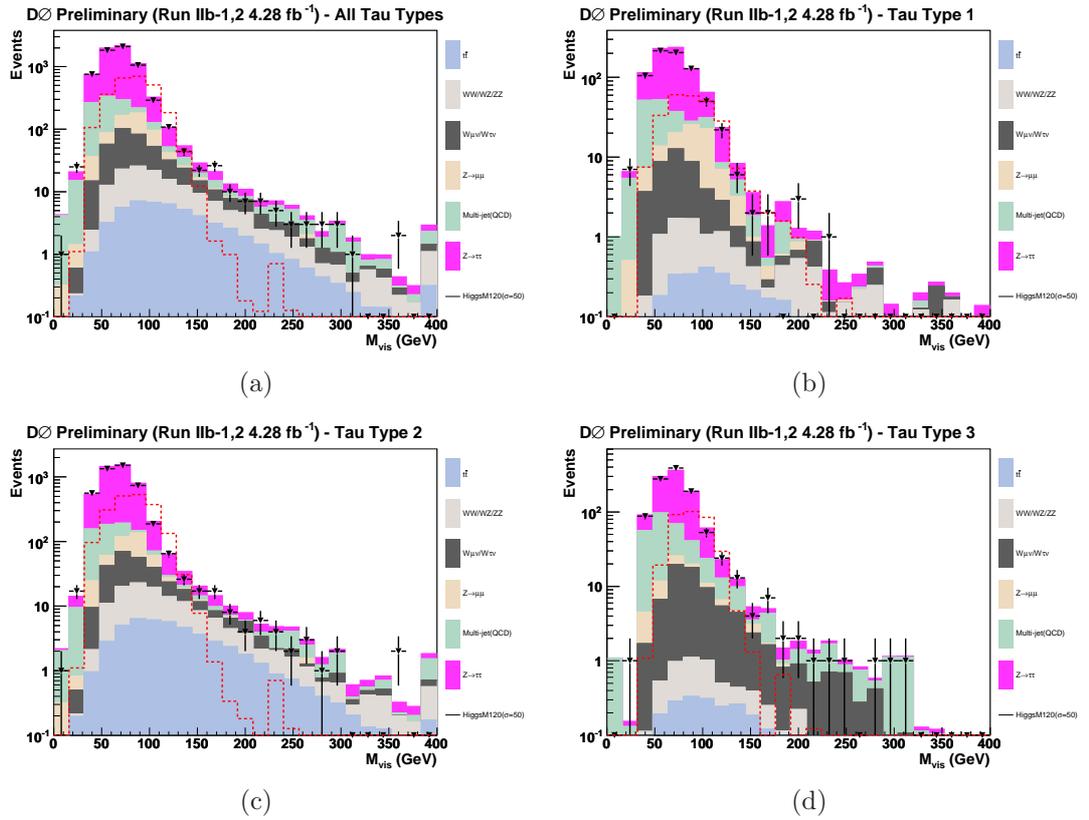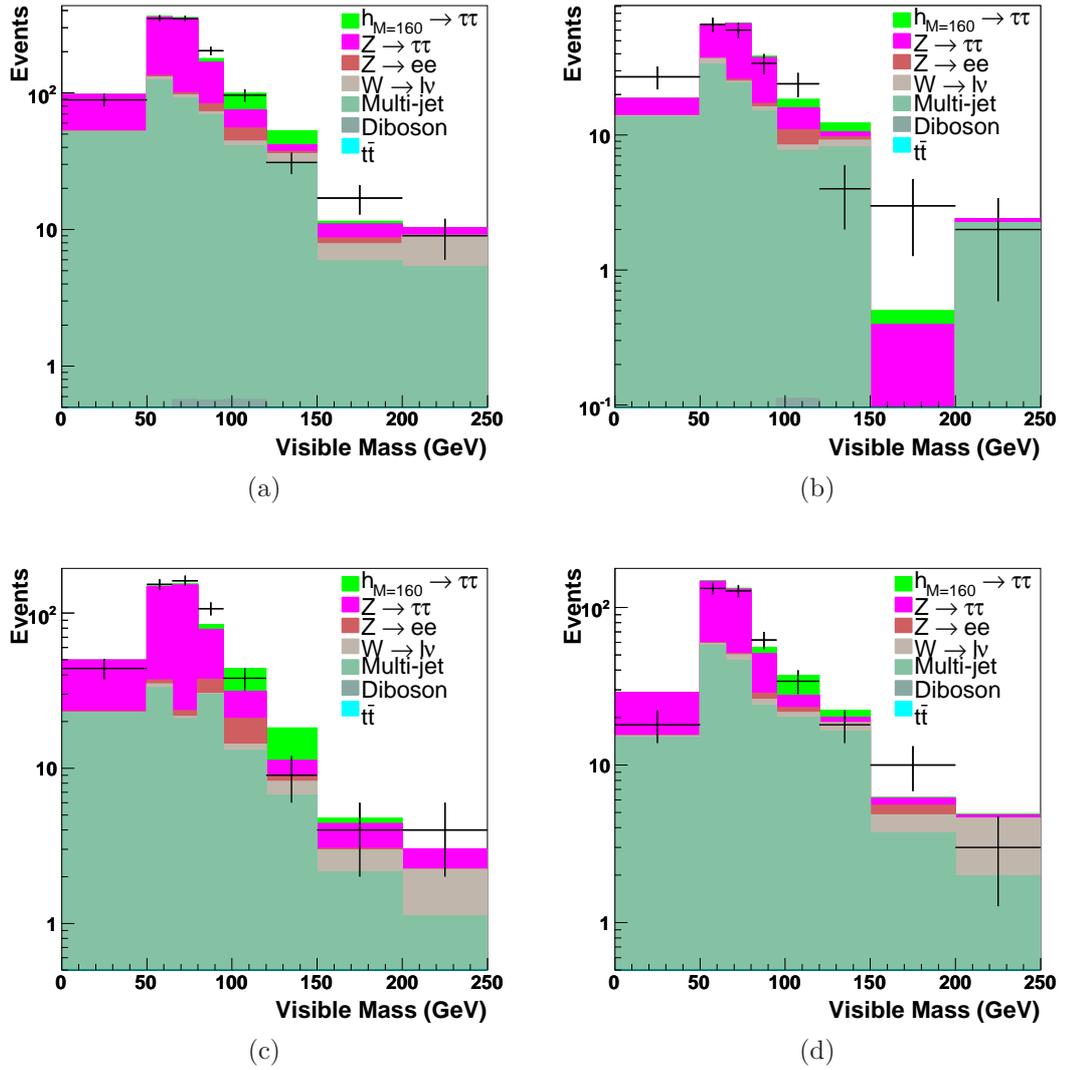
Figure 9.2: Distribution of M_{vis} on a log scale in the Run IIa $\tau_\mu \tau_h$ channel for (a) sum of all τ types (b) type-1 (c) type-2 and (d) type-3. The data, shown with error bars, are compared to the sum of the expected backgrounds.

Figure 9.3: Distribution of M_{vis} on a linear scale in the Run IIb $\tau_\mu \tau_h$ channel for (a) sum of all τ types (b) type-1 (c) type-2 and (d) type-3. The data, shown with error bars, are compared to the sum of the expected backgrounds.

Figure 9.4: Distribution of M_{vis} on a log scale in the Run IIb $\tau_\mu\tau_h$ channel for (a) sum of all τ types (b) type-1 (c) type-2 and (d) type-3. The data, shown with error bars, are compared to the sum of the expected backgrounds.

Figure 9.5: Distribution of M_{vis} on a log scale in the Run IIa $\tau_e\tau_h$ channel for (a) sum of all τ types (b) type-1 (c) type-2 and (d) type-3. The data, shown with error bars, are compared to the sum of the expected backgrounds.

input to the limit calculator, `Collie`, as described in Sect. 4.2.

9.2.1 Systematic Uncertainties

Various sources of systematic uncertainties affect both the signal efficiency and background estimation. The assigned systematics from each source present in the Higgs search described here are summarised in Table 9.1. During the limit calculation, some of the uncertainties modify the shape of the M_{vis} distribution and are parameterised as a function of M_{vis}. Hence, these uncertainties are listed as "shape" in Table 9.1. Another type of systematic uncertainty is also listed, which changes only the normalisation but not the shape of the final discriminant. Listed below are descriptions of both shape and normalisation systematic uncertainties.

- The uncertainty on the efficiencies for muon identification are due to the tracking reconstruction requirement (1.4%), muon isolation (0.9%), and the identification method (1.2%) and assigned only for the $\tau_\mu \tau_h$ channel. These are all taken from the discussion in Ref. [34] and added in quadrature to yield an overall 2.1% uncertainty for muon identification.

- The uncertainty on the electron identification is 3.3% [30], and the uncertainty on the $Z \to e^+ e^-$ background re-weighting is 13% derived from the efficiency correction factors as described in Sect. 6.2. These two uncertainties are negligible in the $\tau_\mu \tau_h$ channels and therefore assigned only for the Run IIa $\tau_e \tau_h$ channel.

- Since tracks from a τ candidate are treated similar to those of muon tracks, an additional uncertainty on the τ track must be assigned. This is taken at 1.4% as described in Ref. [69].

- The uncertainty on the integrated luminosity is taken at 6.1% [47].

- The final distributions are varied with Jet Energy Scale (JES) by ± 1 standard deviation (s.d.) on the signal and background shapes. It is parameterised as a function of M_{vis} and listed as "shape" in Table 9.1 [70]. This uncertainty is found to be $< 1\%$, and therefore is negligible in the $\tau_e \tau_h$ channel.

- The uncertainty on the theoretical Z boson cross section considered for the background contribution from Z/γ^* decays is taken to be 5% in order to account for scale and PDF uncertainties [71].

- As described in Secs. 7.1, two methods are used to determine the multi-jet background contribution where a jet fakes a hadronically decayed τ. In order to assign the systematic uncertainty in the method, the difference between the primary method and the cross-check method is taken per τ type.

- The uncertainty on the trigger efficiency is estimated by varying the trigger probabilities of the single muon or electron triggers by ± 1 s.d. [72]. It is parameterised as a function of M_{vis} and listed as "shape" in Table 9.1.

- The error on the signal acceptance due to PDF uncertainties is estimated by comparing the acceptance of the signal when the MC is re-weighted to the 20 available error sets in the CTEQ6.1 PDF [71], where each set contains plus and minus one standard deviation (± 1 s.d.). The error on the acceptance, $\Delta x^\pm$, is then given in terms of the acceptance for each error set, $x_i^\pm$, and the acceptance of the central value, x_0, as given in Eq. 9.1 [71]:

$$\Delta x^\pm = \sqrt{\sum_{i=0}^{20} (x_0 - x_i^\pm)^2}. \tag{9.1}$$

A 4.6% uncertainty is subsequently assigned for the PDF variation.

- As described in Sect. 6.4, the systematic uncertainties for the τ energy scale is determined by varying the relative hadronic τ energy scale by the uncertainties given in Ref [65], which depend on τ type. Since the scale depends on the shape of the M_{vis} distribution, the uncertainty is taken as shape-dependent.

- Next, the uncertainty on the overall τ identification efficiency was estimated by selecting τ leptons from $Z \to \tau\tau$ decays. The study was performed and certified by the τ identification algorithm group [65]. Systematic contributions arising from both multi-jet estimation in the dataset and the kinematic selections for the hadronic τ decay are considered. The systematic

138

uncertainties per τ type are derived as a function of NN_τ cut and for the selections given in Chapter 8 and assigned as the values of 8% for τ type-1, 4% for type-2 and 5% for type-3 for the Run IIa analysis. Since the data samples are different in Run IIa and Run IIb, different errors of 12%, 4.2%, and 7% for τ type-1, 2, and 3, respectively, are assigned for the Run IIb analysis. This uncertainty is in addition to the one arising from the τ-track efficiency quoted earlier.

The entire series of systematic uncertainties listed above are also assumed to apply for the Higgs boson signal across the full mass region studied.

Figures 9.6 to 9.7 summarise the distributions for systematic uncertainties of the Run IIb $\tau_\mu\tau_h$ channel of each background source which have "shape"-dependency on M_{vis}. These distributions are subsequently used in the `Collie` program to help determine the "best fit" for each histogram bin.

9.2.2 $\tau_e\tau_h$ Channel with $\mathcal{L}$=1.08 fb^{-1}

In this channel, the cross section limits are computed using two algorithms provided by `Collie`:

CLfit1: A single fit to the systematics in the background only hypothesis, and bins with $\log(1 + \frac{s}{b}) > 0.005$ are excluded.

CLfit2: A fit to the systematic uncertainties in both the signal plus background and background-only hypotheses.

Figure 9.8(a) shows the LLR distributions expected for the "B-only" hypothesis, the "S+B" hypothesis and the observation in the data using `CLfit2`. The LLR values are calculated from M_{vis} inputs as shown in Fig 9.1. Also shown are the one and two s.d. for the expectation from the background-only hypothesis. A deviation beyond the 2 s.d. background LLR contour is observed around $M_h = 130$ GeV. However, this is less than a 3 s.d. departure from the background only hypothesis, as can be seen by the observed p-value in the $1 - CL_b$ curve in Fig. 9.8(b). Figure 9.9 shows the expected and observed limits on the cross section multiplied by the branching ratio for $\phi \rightarrow \tau_e\tau_h$ as a function of the test Higgs boson mass, from both `CLfit` and `CLfit2`. The observed limits evaluated by the two `Collie` methods follow the prediction within two standard deviation, and therefore, no signal excess has been observed in this channel.

Source of Systematics	Relative Error on the Signal (%)			Relative Error on the Background (%)		
	Run IIa $\tau_\mu\tau_h$	Run IIb $\tau_e\tau_h$	Run IIa $\tau_e\tau_h$	Run IIa $\tau_\mu\tau_h$	Run IIb $\tau_e\tau_h$	Run IIa $\tau_e\tau_h$
Luminosity	6.1	6.1	6.1	6.1	6.1	6.1
$Z/\gamma^* \to ll$ cross section	–	–	–	5	5	5
PDF variation	4.6	4.6	4	–	–	–
τ-Track Reconstruction	1.4	1.4	4	1.4	1.4	4
Muon Identification	2.1	2.1	–	2.1	2.1	–
$Z \to e^+e^-$ Background	–	–	–	–	–	13
EM Identification	–	–	3.3	–	–	3.3
Errors Assigned by τ type						
τ Identification	4-8	4.2-12	4-8	4-8	4.2-12	4-8
τ Energy Scale	Shape	Shape	Shape	Shape	Shape	Shape
Multi-jet	–	–	–	8-12	8-12	3-28
JES	Shape	Shape	Shape	Shape	Shape	Shape
Trigger	Shape	Shape	Shape	Shape	Shape	Shape

Table 9.1: Summary of the various sources of systematic errors for both the signal and background described in the text.
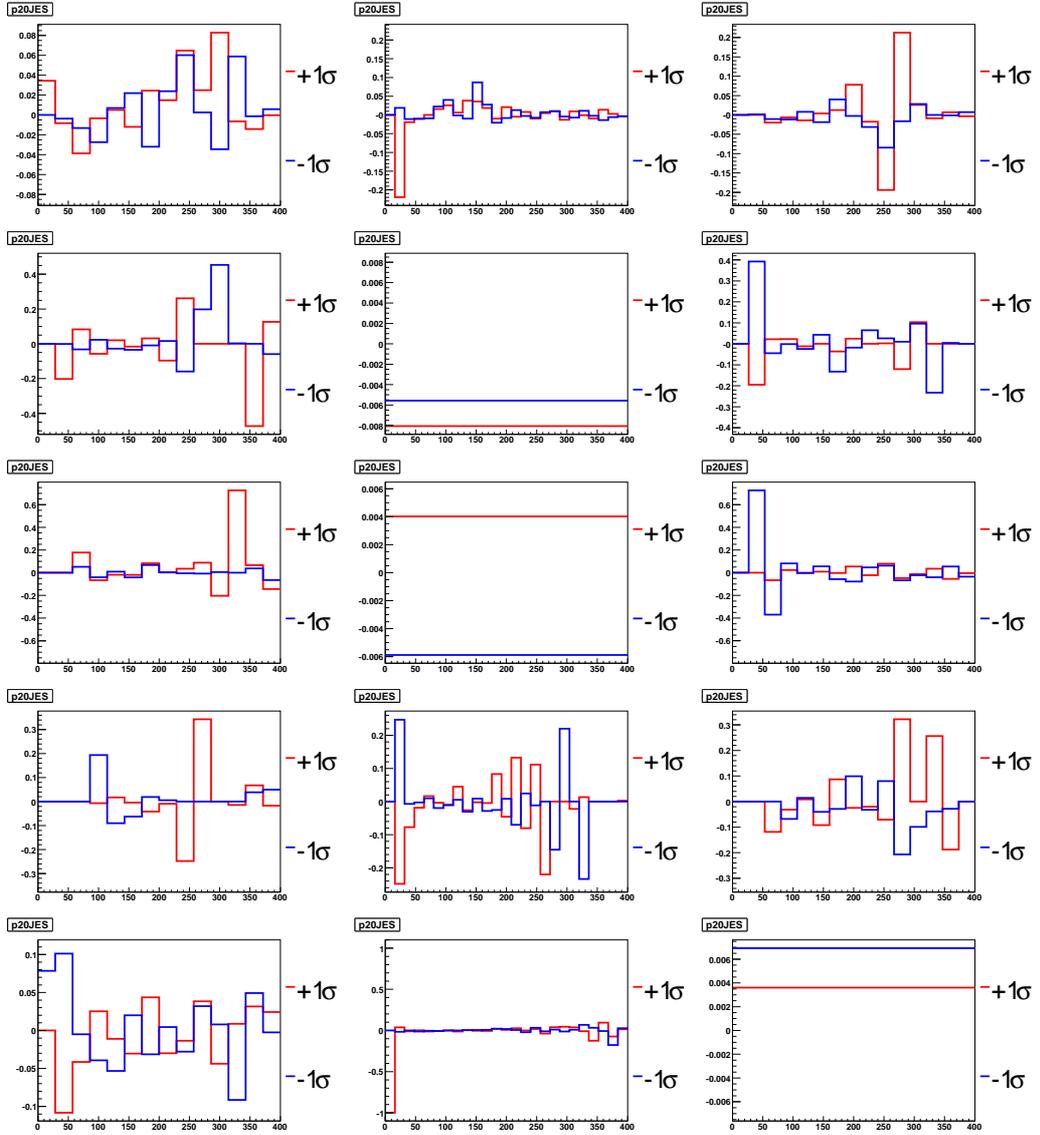
Figure 9.6: Distributions of shape dependent Jet energy scale (JES) systematics of $Z \rightarrow \tau^+\tau^-$, $Z \rightarrow \mu^+\mu^-$, W+jets, diboson and $t\bar{t}$ background sources (from top to bottom) for τ types 1, 2 and 3 (from left to right).
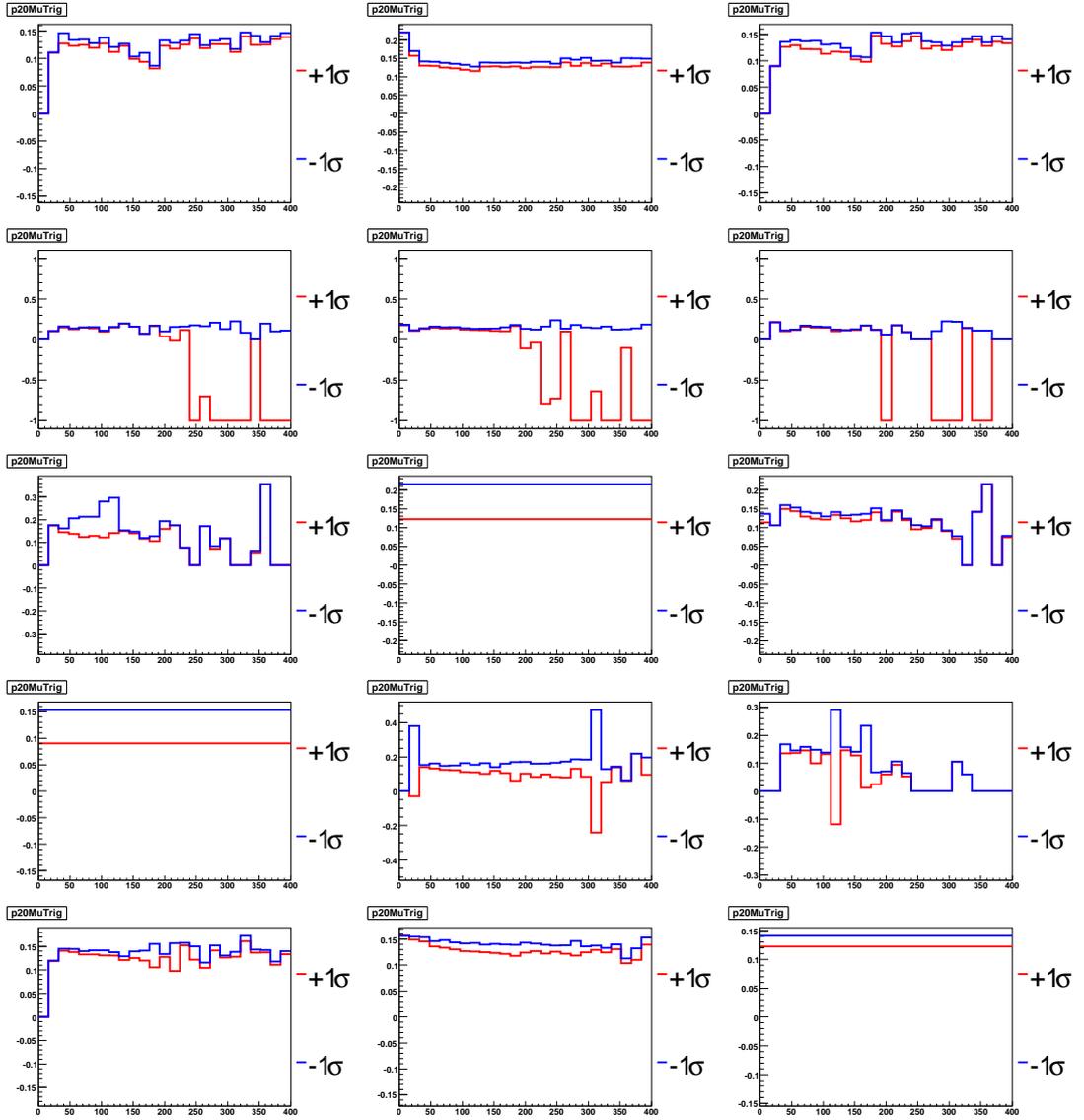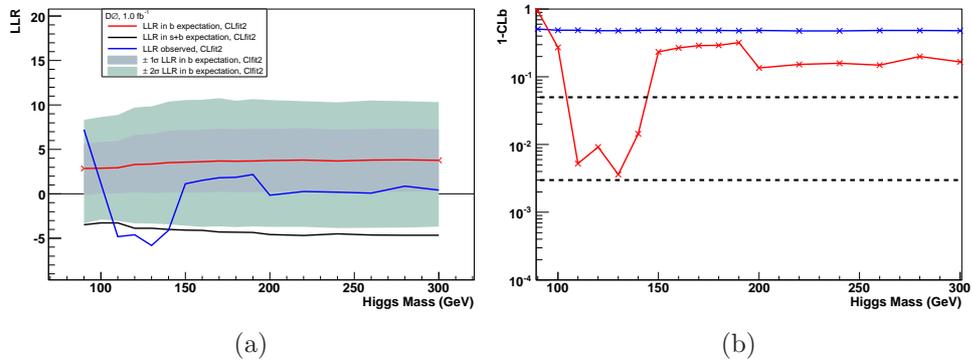
Figure 9.7: Distributions of shape dependent single muon OR trigger systematics of $Z \to \tau^+\tau^-$, $Z \to \mu^+\mu^-$, W+jets, diboson and $t\bar{t}$ background sources (from top to bottom) for τ types 1, 2 and 3 (from left to right).

Figure 9.8: Outcomes of `CLfit2` method for (a) LLR for the data, the expectation from the "B-only" and the "S+B" hypothesis, (b) CL in the "B-only" hypothesis (CL_b) as a function of the test Higgs boson mass. The blue line shows the expectation in the "B-only" hypothesis and the red line shows the observation in the data. The dotted lines correspond to $1 - CL_b$ p-values of 0.05 (2 s.d.) and 0.003 (3 s.d.).
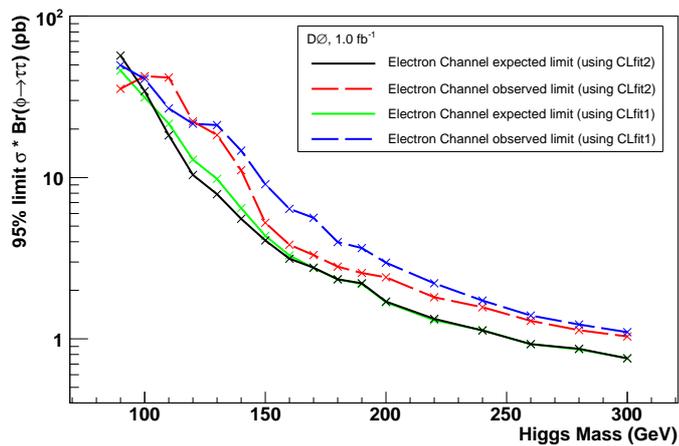
Figure 9.9: Expected and observed limits for the cross section multiplied by the branching ratio for the electron channel. The limit has been calculated with `CLfit` and `CLfit2`.

143

9.2.3 $\tau_\mu\tau_h$ Channel with $\mathcal{L}$=5.36 fb^{-1}

A fast approximation to the `CLfit2` algorithm is used to calculate the cross section limits of this channel. It provides results within $2-3\%$ precision compared to `CLfit2` and takes only 1/10 of the running time [38]. The LLR distribution of the combined 5.36 fb^{-1} data of $\tau_\mu\tau_h$ channel is shown in Fig. 9.10 with 1 s.d. and 2 s.d. variations. The data follow the "B-only" expectation very well. Figure 9.11 shows the expected and observed limits on the cross section multiplied by branching ratio $(\sigma(p\bar{p} \rightarrow \phi) \times Br(\phi \rightarrow \tau\tau))$ as a function of the test Higgs boson mass calculated by the fast approximation module in `Collie`. Across the whole scanned mass region, there are no deviations greater than 2 s.d. observed.
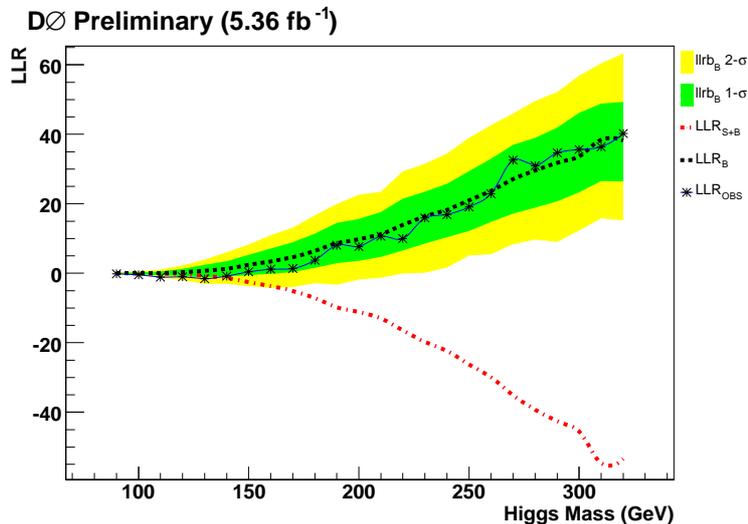
Figure 9.10: LLR distribution for data (blue line) and the expectations. Black and red lines represent the results from "B-only" and "S+B" hypothesis, respectively. Also shown the green and yellow bands are ± 1 s.d. and ± 2 s.d. from the "B-only" expected values.

9.3 Translation into the MSSM

In this analysis, the production cross sections, widths, and branching ratios are determined using `FEYNHIGGS` 2.6 [22] (see Sect. 2.3.2) to set exclusion limits. Figures 9.12 and 9.13 show the combined expected and observed exclusions of the M_h^{max} and *no-mixing* scenarios for $\mu = \pm 200$ GeV using the DØ Run IIa
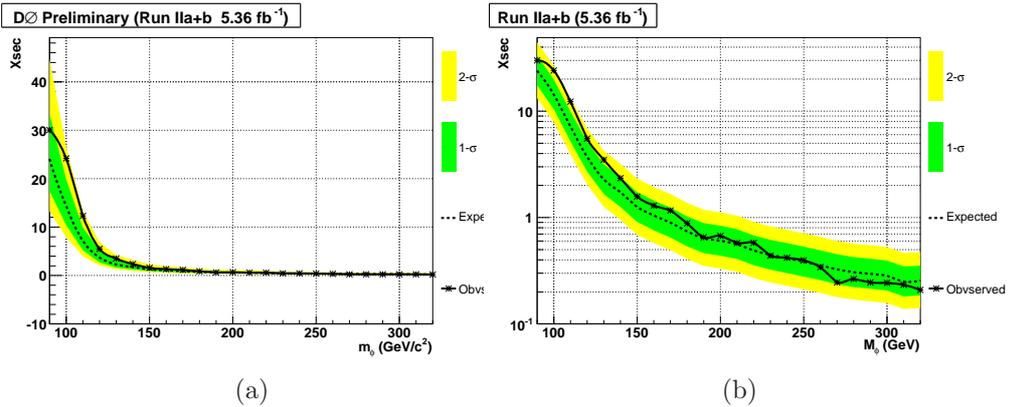
Figure 9.11: Cross section limits evaluated by the fast approximation to `CLfit2` for the $\tau_\mu\tau_h$ channel and 5.36 fb^{-1} integrated luminosity in linear scale (a) and in log scale (b). The green and yellow bands are ± 1 s.d. and ± 2 s.d. from the "B-only" expected limits.

dataset including the $\tau_e\tau_h$, $\tau_e\tau_\mu$ and $\tau_\mu\tau_h$ channels [53]. This result has been published in Phys. Rev. Lett. in 2008 [53]. Similar distributions for the results in only the $\tau_\mu\tau_h$ channel with DØ Run IIa + Run IIb data are shown in Fig. 9.14 and 9.15. Discussions of the results comparing to LEP result and other MSSM Higgs boson searches at the Tevatron can be found in the next Sect. 9.4.

9.4 Comparison with Previous Results and Other Higgs Searches

The LEP experiments at CERN performed searches for MSSM Higgs bosons mainly in Higgs strahlung, $e^+e^- \rightarrow Zh, ZH$, channels. No signal above the expected backgrounds was found [73], and the results therefore excluded all $\tan\beta$ values for M_A below approximately 93.4 GeV and bounded $M_h > 92.8$ GeV for higher values of $\tan\beta$ as shown in the green shaded area in Fig. 9.14 and 9.15. The low $\tan\beta$ region with $0.7 < \tan\beta < 2.0$ was excluded for all M_A motivated by the fact that the lightest Higgs boson (h) with such low $\tan\beta$ values should be "SM-like" since otherwise it would have been observed by the LEP searches [73].

Direct searches including the search presented in this thesis at the Tevatron are able to probe several MSSM benchmark scenarios and consequently extend the search regions covered by the LEP experiment. The most promising channels
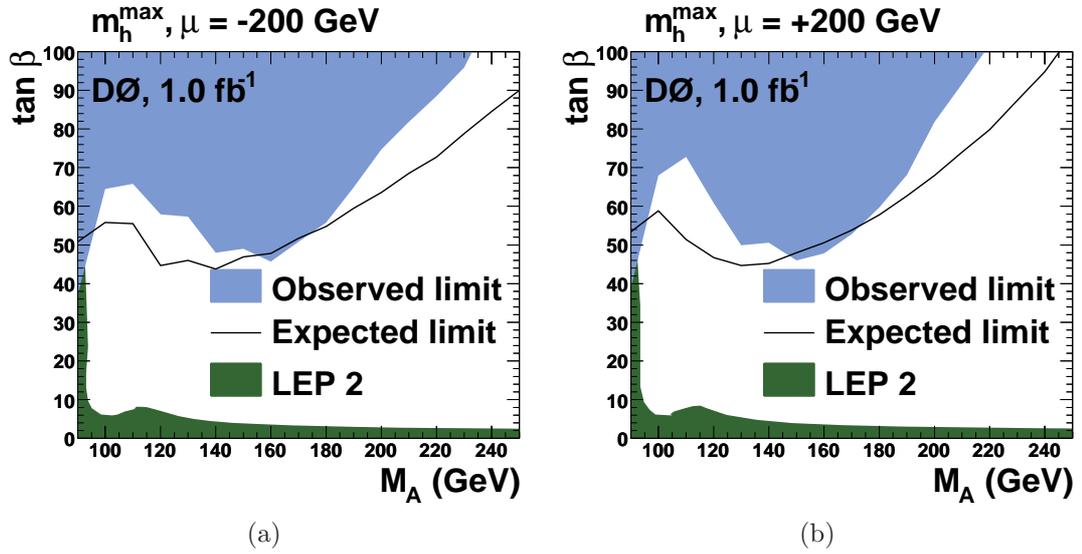
Figure 9.12: Expected and observed exclusion using Run IIa data set for the M_h^{max} scenario in the $(M_A, \tan\beta)$ plane, for (a) $\mu = -200$ GeV and (b) $\mu = +200$ GeV.
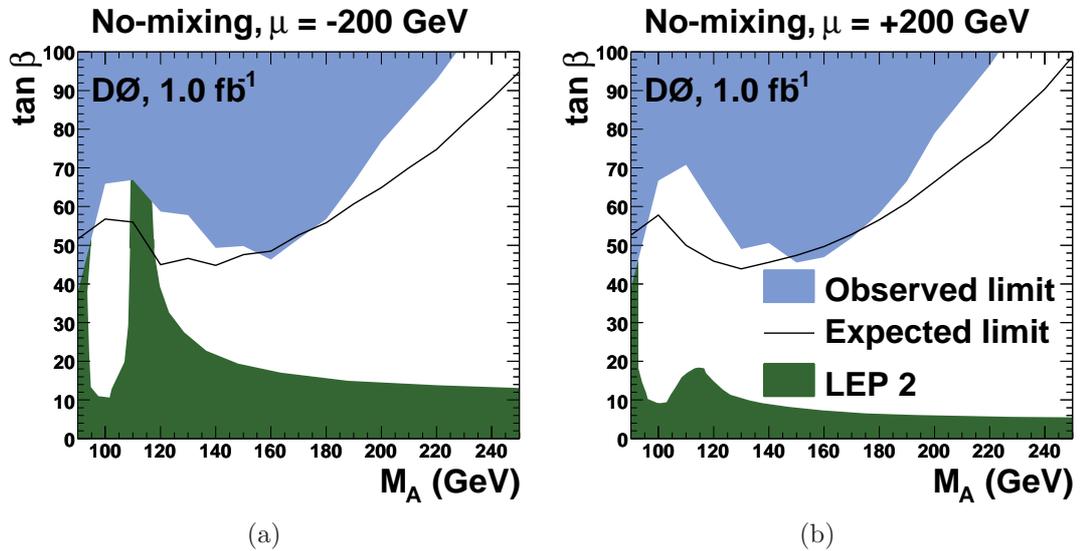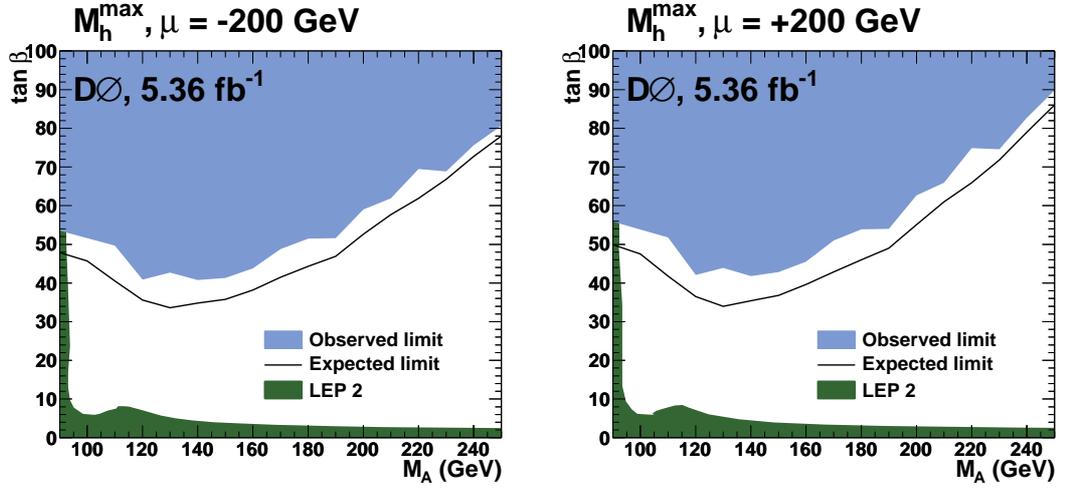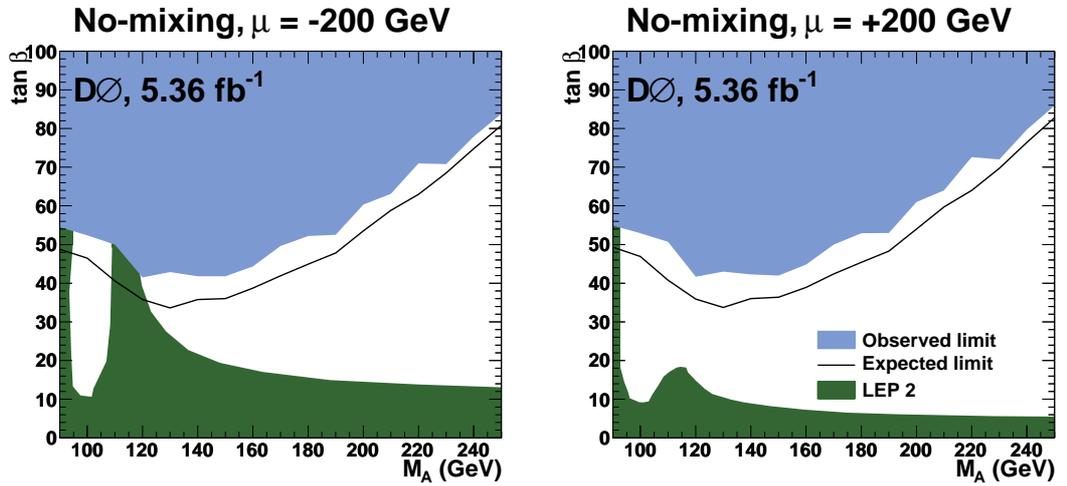

Figure 9.13: Expected and observed exclusion using Run IIa data set for the *no-mixing* scenario in the $(M_A, \tan\beta)$ plane, for (a) $\mu = -200$ GeV and (b) $\mu = +200$ GeV.

Figure 9.14: Expected and observed exclusion $\tau_\mu \tau_h$ channel for the M_h^{max} scenario in the $(M_A, \tan\beta)$ plane, for (a) $\mu = -200$ GeV and (b) $\mu = +200$ GeV.

Figure 9.15: Expected and observed exclusion $\tau_\mu \tau_h$ channel for the *no-mixing* scenario in the $(\tan\beta - M_A)$ plane, for (a) $\mu = -200$ GeV and (b) $\mu = +200$ GeV.

studied at the Tevatron are $b\bar{b}\phi$, where $\phi \to b\bar{b}$ where the final state contains at least three b quarks, $b\phi \to b\tau^+\tau^-$, and the inclusive $\phi \to \tau^+\tau^-$ channel studied in this thesis. Inclusive searches for $\phi \to \tau\tau$ have been performed with integrated luminosities of $\mathcal{L} = 1.08$ fb^{-1} by DØ [53] in Run IIa and $\mathcal{L} = 1.8$ fb^{-1} by CDF [74] in Run IIa and IIb. The result of the $\tau_e + \tau_h$ search channel performed in this thesis is also included. These searches require the τ lepton pairs to decay into three final states: $\tau_e\tau_h$, $\tau_\mu\tau_h$, and $\tau_e\tau_\mu$, where τ_e and τ_μ are the leptonic decays of the tau and τ_h is the hadronic decaying mode. Subsequently, the DØ collaboration has updated its $\phi \to \tau_\mu\tau_h$ search using $\mathcal{L} = 1.2$ fb^{-1} of the Run IIb integrated luminosity to include a total of 2.2 fb^{-1} integrated luminosity in the $\tau_\mu\tau_h$ mode. The results performed in this thesis re-analysed the $\phi \to \tau_\mu\tau_h$ search channel with the DØ Run IIa dataset and extends the search further with the DØ Run IIb data set for a total of 5.36 fb^{-1} of integrated luminosity. Details of these data sets for DØ are given in Chapter 5. Furthermore, DØ has updated its $b\phi \to b\tau^+\tau^-$ search in 2010. This result considers only tree level enhancements and is shown in Fig. 9.16 [77].

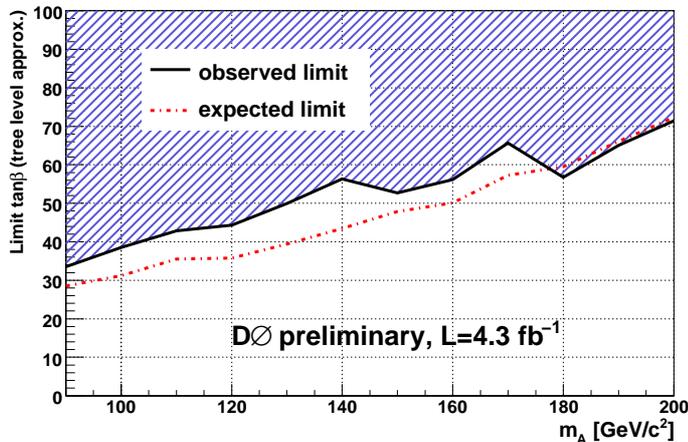Figure 9.16:   The exclusion in the MSSM $(M_A, \tan\beta)$ plane of the $b\phi \to b\tau^+\tau^-$ channel with 4.3 fb^{-1} DØ Run IIb integrated luminosity updated in summer 2010. The limit has been calculated at tree level, and therefore, no MSSM scenario is considered [77].

The three channels explored so far at the Tevatron have unique features that are different depending on the mass region of interest and the MSSM scenario. Therefore, the combinations of such search channels in the future is expected

to provide the best sensitivity to the MSSM Higgs boson. First, the search for the $\phi \to \tau\tau$ channel presented in this thesis is limited by the Z/γ^* background events particularly the 90 GeV Higgs mass point, and therefore, its performance is less sensitive in the low mass regions. However, due to sufficient statistics, this search channel is able to provide better sensitivities at intermediate and higher Higgs mass points. For example, the most recent result for $\tau_\mu\tau_h$ channel as shown in Fig. 9.14 is capable of probing regions with $\tan\beta < 40$ in a mass window of $115 \leq M_A \leq 160$ GeV for the M_h^{max} scenario and negative μ values. In addition, because this channel is not sensitive to the sign of μ, it provides exclusions down to $\tan\beta \approx 40$ in a similar mass window for both $\mu = \pm 200$ GeV.

Moreover, the search channel for Higgs bosons from the production associated with a b quark makes it possible to use the dominant $b\bar{b}$ decay mode of MSSM neutral Higgs bosons. As discussed in Chapter 2, the dependence on μ is stronger for this decay mode, and therefore, the *bbb* channel is highly sensitive to results with negative values of μ [76]. Subsequently, the Drell-Yan background appearing at lower masses can be significantly suppressed by tagging b quark jets in the $b\phi \to b\tau^+\tau^-$ search channel. For $M_A = 90$ GeV Higgs mass point, this channel excludes $\tan\beta \geq 35$, which is about 1.5 times better than the results shown in Figs. 9.14 and Fig. 9.15 from the inclusive $\phi \to \tau\tau$ searches alone.

The Tevatron combined results in 2009 for $\tau^+\tau^-$ final state including 1-2.2 fb^{-1} from DØ and 1.8 fb^{-1} from CDF inclusive searches for M_h^{max} and *no-mixing* benchmark scenarios for $\mu = +200$ GeV are shown in Fig. 9.17; the exclusions set by the LEP experiment for both scenarios are also shown. This combination excludes $\tan\beta \geq 30$ for $M_A = 90$ GeV and $130 \leq M_A \leq 150$ GeV and has reached sensitivities which are on the order of the ratio of top-quark mass to bottom-quark mass, m_t/m_b. The most recent updates in $\tau_\mu\tau_h$ channel presented in this thesis is not yet included in the combination, which has exceeded the sensitivity of the combined result at the intermediate Higgs masses. In addition, the latter results described in this thesis extend the searches for Higgs to masses up to 320 GeV. Table 9.2 listed the sensitivities for the MSSM Higgs searches summarised here.

Detector	Channel	Integrated luminosity (fb^{-1})	Observed $\tan\beta$ limit at $M_A = 150$ GeV		Reference
			no-mixing $\mu = +200$ GeV	m_h^{max} $\mu = -200$ GeV	
CDF	$b\phi \to bbb$	1.9	–	90	[79]
DØ	$b\phi \to bbb$	2.6	–	50	[76]
CDF	$\phi \to \tau\tau$	1.8	40	41	[78]
DØ	$\phi \to \tau\tau$	1.08	47	50	This thesis, [53]
DØ	$\phi \to \tau_\mu\tau_h$	5.36	38	36	This thesis
Tevatron	$\phi \to \tau\tau$	1.8-2.2	32	31	[75]
DØ	$b\phi \to b\tau\tau$	4.3	52 (tree-level)		[77]

Table 9.2: Sensitivity of the different MSSM Higgs boson searches. The expected sensitivity in terms of $\tan\beta$ in two scenarios at $M_A = 150$ GeV is shown for each channel, along with the integrated luminosity of the dataset used.
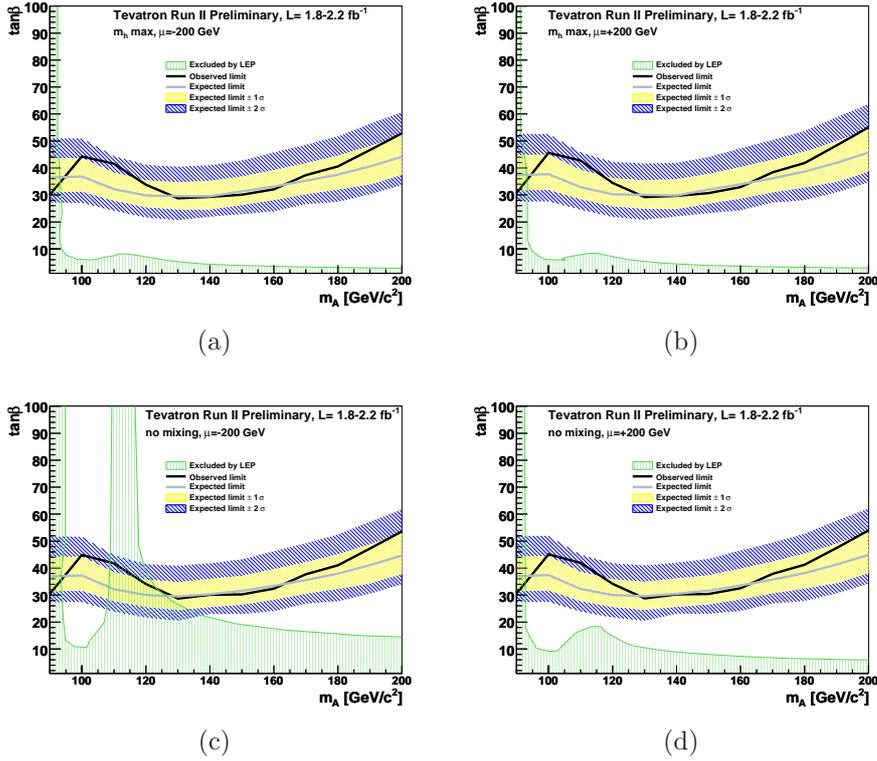

Figure 9.17: The exclusion in the MSSM $(M_A, \tan\beta)$ plane of the Tevatron combined results in $\tau^+\tau^-$ final state with (a) M_h^{max} scenario with $\mu = -200$ (b) M_h^{max} scenario with $\mu = +200$ (c) *no-mixing* scenario with $\mu = -200$ (d) *no-mixing* scenario with $\mu = +200$ [75].

Chapter 10

Conclusions

A search for neutral Higgs bosons decays into τ lepton pairs is presented in this thesis with one τ lepton decaying hadronically and the other one decaying leptonically into either an electron (e) or a muon (μ). The data used correspond to 1.08 and 5.36 fb^{-1} of integrated luminosity for the $\tau_e \tau_h$ and $\tau_\mu \tau_h$ channels, respectively, and have been collected from 2002 to 2009 by the DØ detector at the the Tevatron. No significant excess has been observed in data and the resulting cross section limits are interpreted in the Minimal Supersymmetric Standard Model (MSSM) to exclude a region in the MSSM parameter space in terms of the mass of the CP-odd Higgs boson, M_A, and $\tan\beta$.

Both search channels, $\tau_e \tau_h$ and $\tau_\mu \tau_h$, require the events to contain an electron or a muon with transverse momentum $p_T^{e/\mu} \geq 15$ GeV. The hadronically decaying τ candidate is required to pass p_T selections greater than 15-20 GeV for the $\tau_e \tau_h$ decay mode and 12.5-15 GeV for the $\tau_\mu \tau_h$ channel. These requirements are applied depending on the different hadronically decaying τ types.

Relevant physics backgrounds are modelled by MC simulations using the `PYTHIA` and `ALPGEN` generators. In order to compensate for effects which are not simulated properly, a series of efficiency and energy scale corrections are applied to identified objects. Two independent methods have been developed for each of the channels to estimate the contribution from multijet background events. The difference between the two methods is taken as the systematic uncertainty for the multijet background determination. At the preselection stage, those backgrounds simulated by MC generators or those estimated with data in control regions (multijet) are studied for their consistency with data.

Most of the backgrounds are reduced significantly by a series of selections in each decay channel. These backgrounds include $W + $ jets, $Z \to e^+e^-$ and $Z \to \mu^+\mu^-$, and multijet events. The selections are designed based on topological or kinematic properties of different backgrounds and possible signal candidates such that background is reduced while keeping the signal rate maximised. Events from $Z/\gamma^* \to \tau^+\tau^-$ processes dominate the final samples, since it is impossible to distinguish Z/γ^* decays from Higgs candidate events apart from mass and spin. Due to the presence of neutrinos in the final states, the visible mass M_{vis} is used to discriminate and search for a Higgs signal above the remaining backgrounds.

Limits at the 95% confidence level (CL) are set on $\sigma(p\bar{p} \to \phi) \times \mathrm{BR}(\phi \to \tau\tau)$ for Higgs boson masses within the range of 90 GeV $< M_\phi < 300$ GeV and 90 GeV $< M_\phi < 320$ GeV for the $\tau_e\tau_h$ and $\tau_\mu\tau_h$ channels, respectively. The $\tau_e\tau_h$ was combined with other search channels, $\tau_\mu\tau_h$ and $\tau_e\tau_\mu$ in a publication [53] using 1.08 fb^{-1} of integrated luminosity collected in DØ Run IIa. The Run IIa data are re-analysed for the $\tau_\mu\tau_h$ channel and combined with 4.28 fb^{-1} of integrated luminosity collected in DØ Run IIb. The expected (observed) limits range from 24 (30) pb at $M_\phi = 90$ GeV to 0.28 (0.2) pb at $M_\phi = 320$ GeV. The cross section limits are translated into exclusions in the $(M_A - \tan\beta)$ plane for the m_h^{max} and *no-mixing* MSSM scenarios using FEYNHIGGS [22]. The search excluded the range $45 < \tan\beta < 98$ for 90 GeV $< M_A < 250$ GeV. The results from the $\tau_\mu\tau_h$ channel alone provide comparable exclusions to the Tevatron (CDF+DØ) combinations using 1.8-2.2 fb^{-1} of data from all three $\tau\tau$ search channels [75].

The DØ experiment has recorded over 8.5 fb^{-1} of integrated luminosity. Therefore, the analyses presented in this thesis can be extended with the increased data set in the future. By combining with other MSSM Higgs boson search channels studied at the Tevatron, projections show that the Tevatron will be able to probe the most theoretically preferred value of $\tan\beta = 35$ for all M_A regions below 200 GeV with the expected 2011 data set. Although the Large Hadron Collider (LHC) at CERN has started running in 2008, stable running at the Tevatron along with sufficient understanding of the detector's performance by both CDF and DØ keeps the ability to search for Higgs bosons at low masses promising.

Appendix

Kinematic Distributions in $\tau_\mu \tau_h$ Channel

For completeness, a comprehensive list of distributions is given here after final selections (see Sect. 8.2) in the $\tau_\mu \tau_h$ channel. These distributions are used for studying of the background modelling and are shown by τ type. Good agreement between data and MC is observed in most distributions and indicating that backgrounds are well modelled, which are either estimated from data or MC generators.

1. Figures 1, 5: Cosine value of the difference between azimuthal angle ϕ of $\not{E}_T$ and muon or τ candidates, $\cos(\Delta\phi(\mu/\tau, \not{E}_T))$, and difference between azimuthal angle ϕ of muon and τ, $\Delta\phi(\mu, \tau)$.

2. Figures 2, 6: ΔR distribution between selected muon and τ candidates, $\cos(\Delta R(\mu, \tau)$, instantaneous luminosity distributions, $\mathcal{L}$, and jet multiplicity of the events, N_{jets}.

3. Figures 3, 7: Detector pseudorapidity of muons, η^μ, azimuthal angle of muon candidates, ϕ^μ, and transverse momentum of muons, p_T^μ.

4. Figures 4, 8: Azimuthal angle of τ candidates, ϕ^τ, transverse mass, M_T, and invariant mass of associated tracks, M_{trk}.

Run IIa

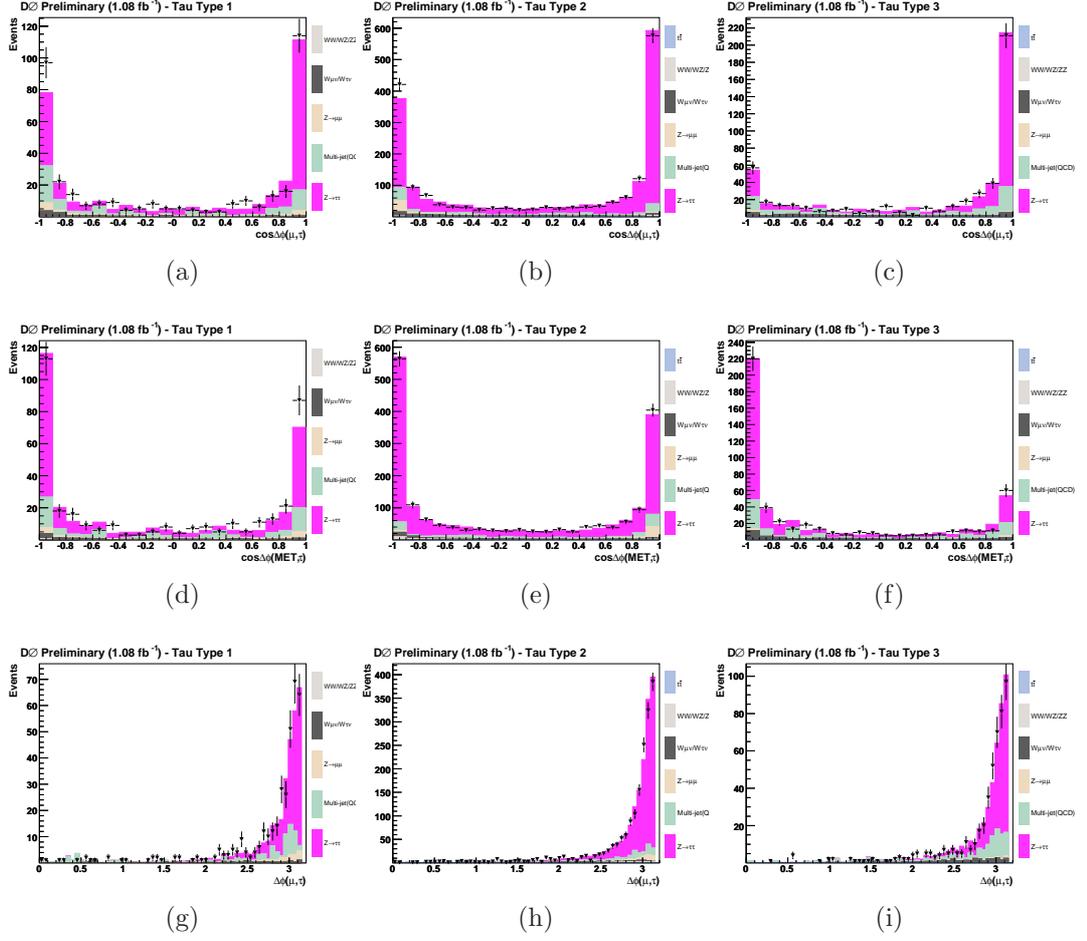Figure 1: Distributions of $\cos(\Delta\phi(\mu, \not{E}_T)$, $\Delta\phi(\mu, \tau)$, and $\cos(\Delta\phi(\tau, \not{E}_T))$ (from top to bottom) for τ types 1, 2 3 (from left to right) for the Run IIa $\tau_\mu\tau_h$ channel.
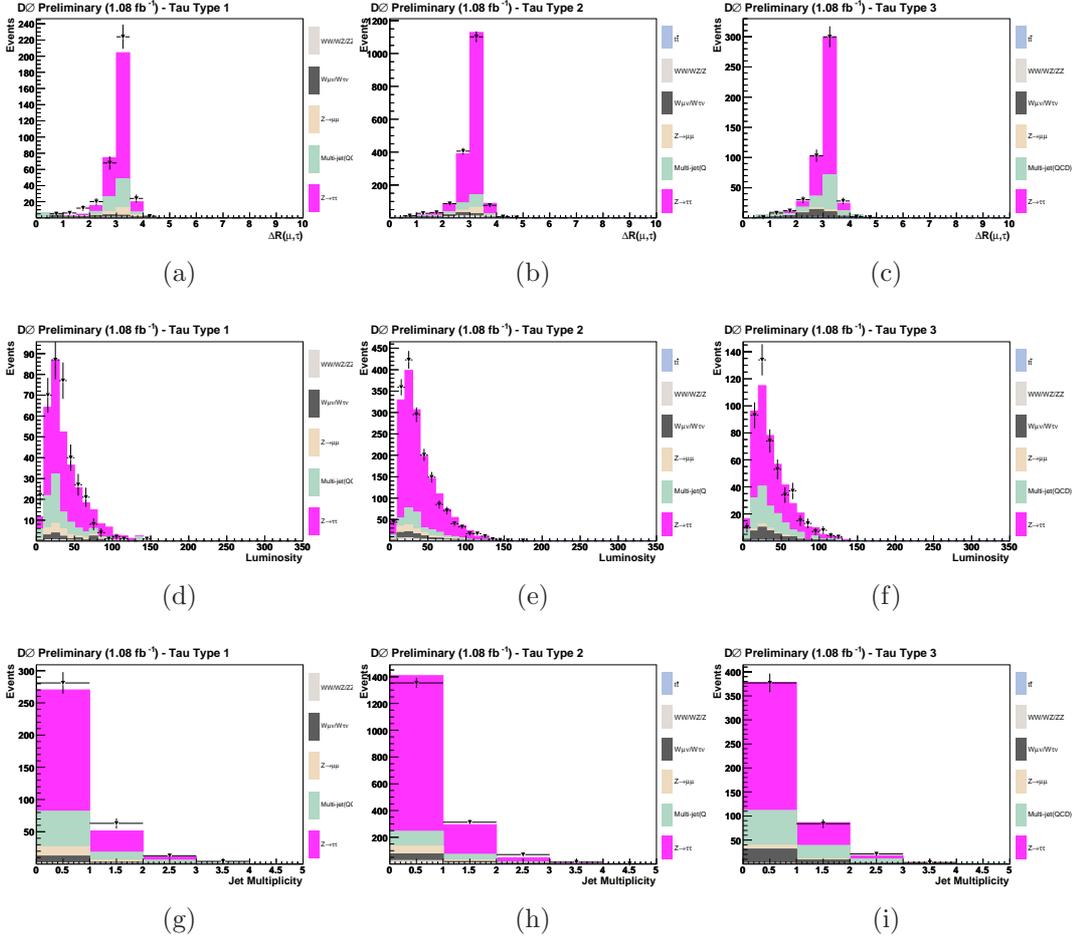
Figure 2: Distributions of $\Delta R(\mu, \tau)$, $\mathcal{L} \times 10^{30}$ cm^{-2}s^{-1}, and $N_{jets}\eta^{\mu}$ (from top to bottom) for τ types 1, 2 3 (from left to right) for the Run IIa $\tau_{\mu}\tau_h$ channel.
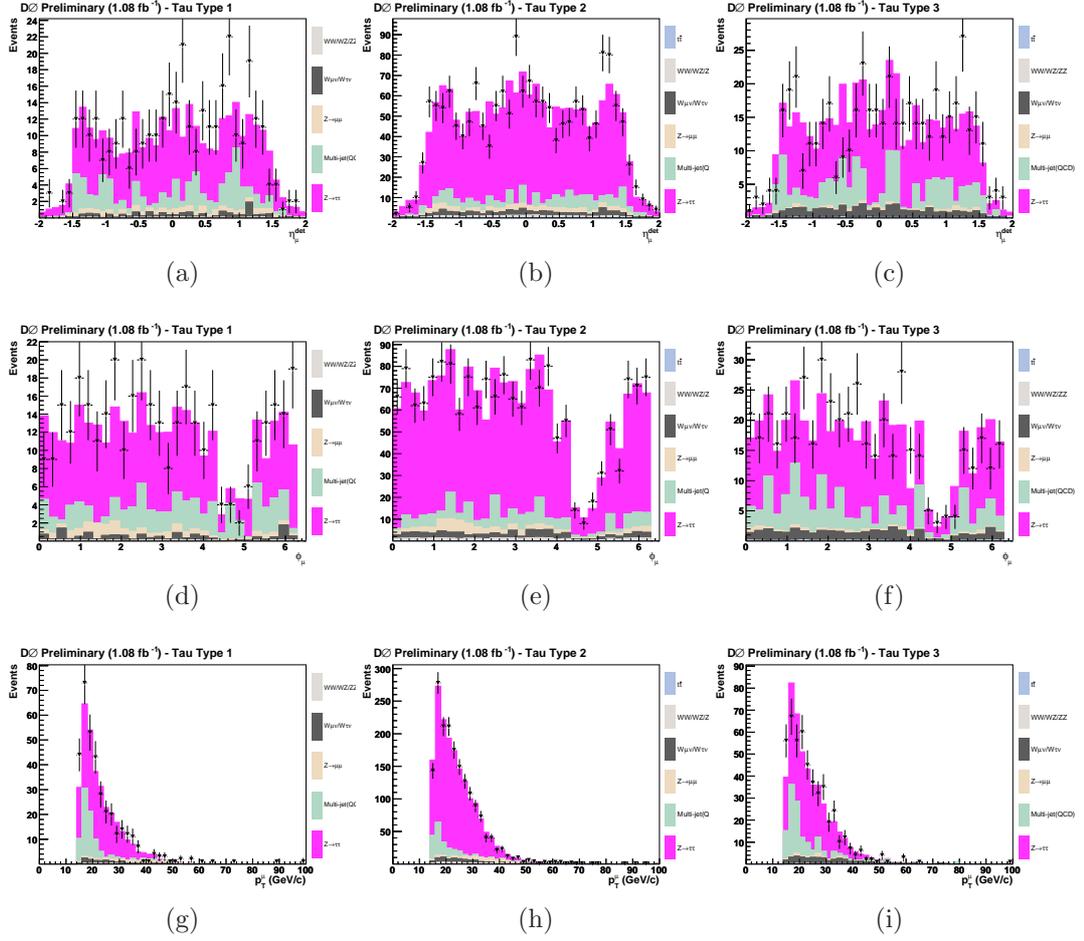
Figure 3: Distributions of η^{μ}, ϕ^{μ}, and p_T^{μ} (from top to bottom) for τ types 1, 2 3 (from left to right) for the Run IIa $\tau_{\mu}\tau_h$ channel.
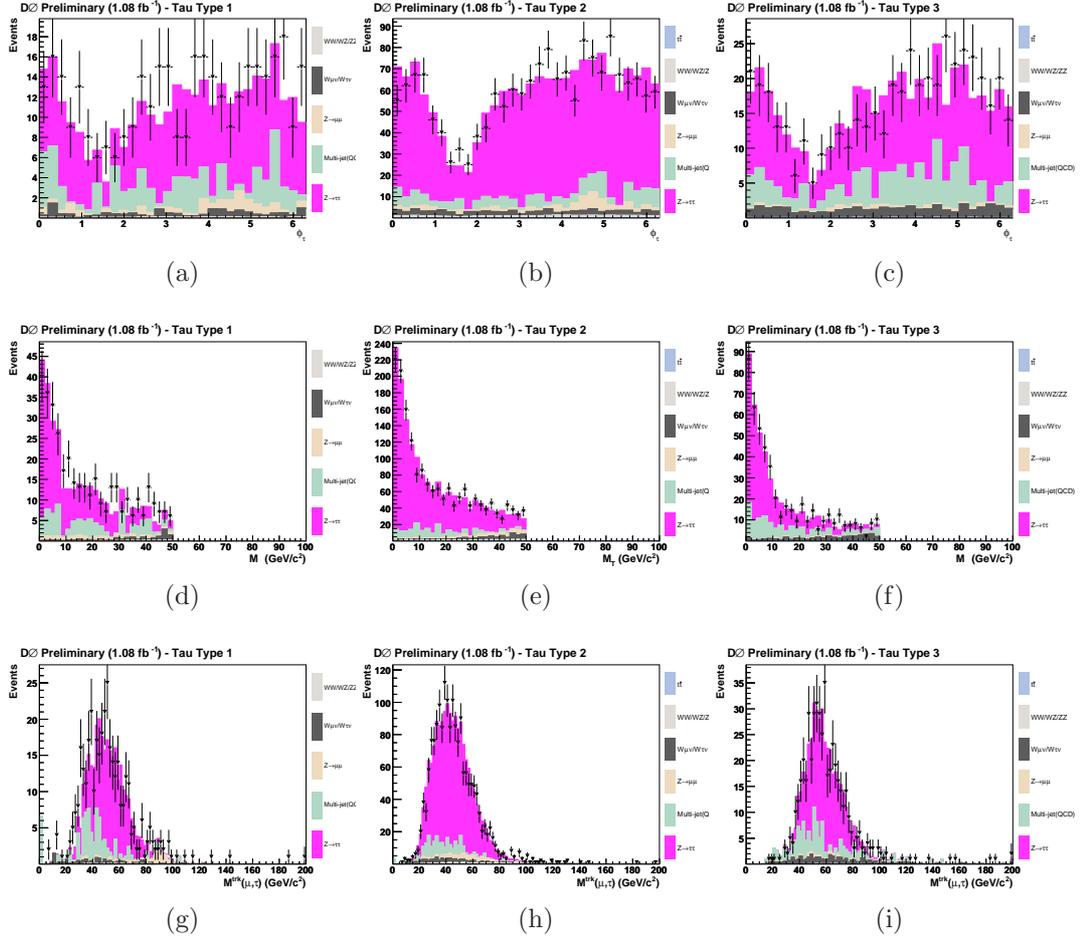
Figure 4: Distributions of ϕ^τ, M_T, and M_{trk} (from top to bottom) for τ types 1, 2 3 (from left to right) for the Run IIa $\tau_\mu\tau_h$ channel.
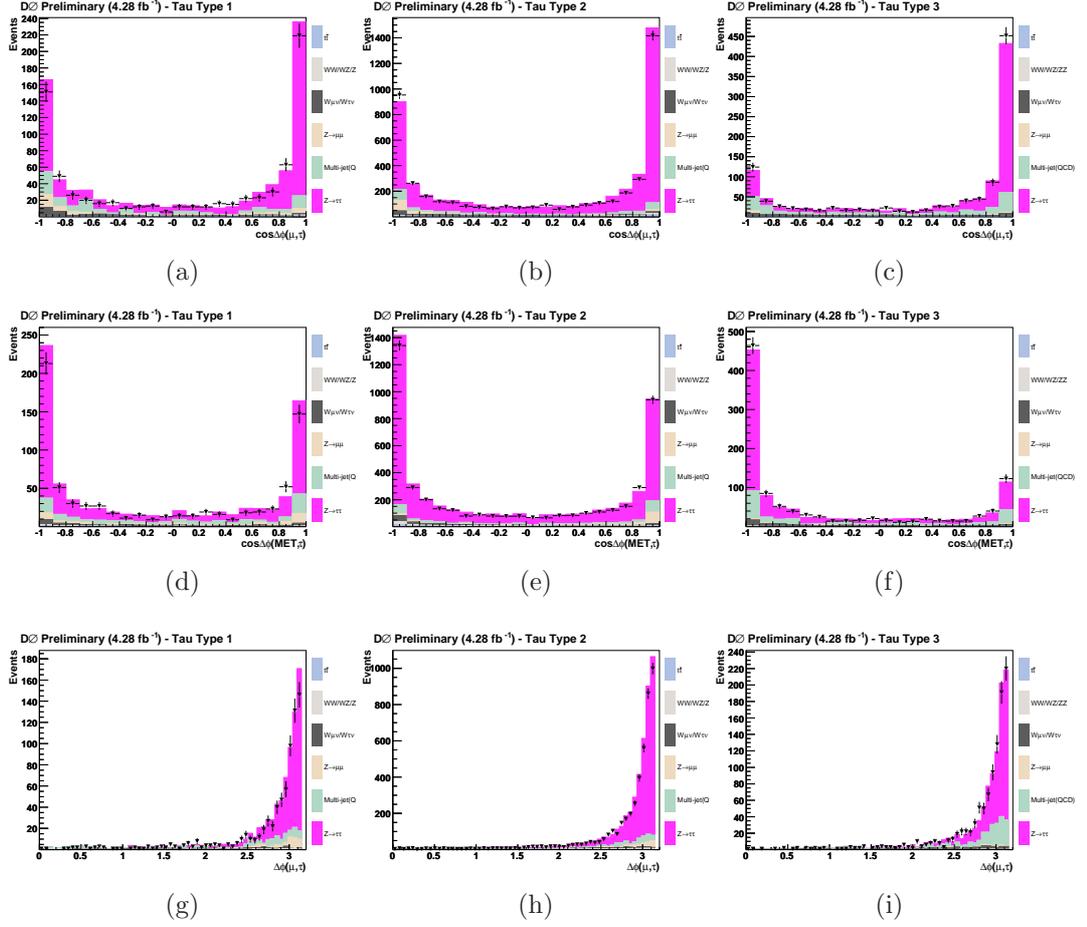
Run IIb

Figure 5: Distributions of $\cos(\Delta\phi(\mu, \not{E}_T))$, $\Delta\phi(\mu, \tau)$, and $\cos(\Delta\phi(\tau, \not{E}_T))$ (from top to bottom) for τ types 1, 2 3 (from left to right) for the Run IIb $\tau_\mu\tau_h$ channel.
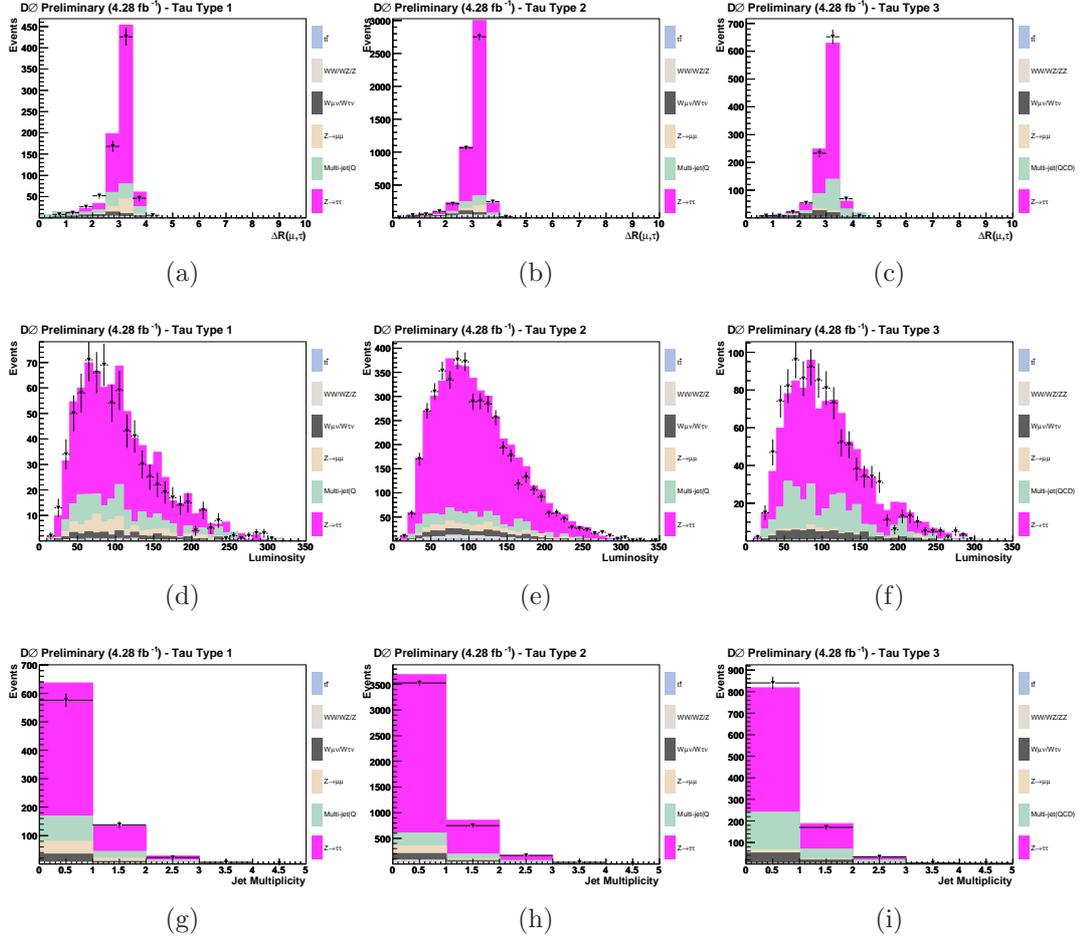
Figure 6: Distributions of $\Delta R(\mu, \tau)$, $\mathcal{L} \times 10^{30}$ cm^{-2}s^{-1}, and $N_{jets}\eta^{\mu}$ (from top to bottom) for τ types 1, 2 3 (from left to right) for the Run IIb $\tau_{\mu}\tau_{h}$ channel.
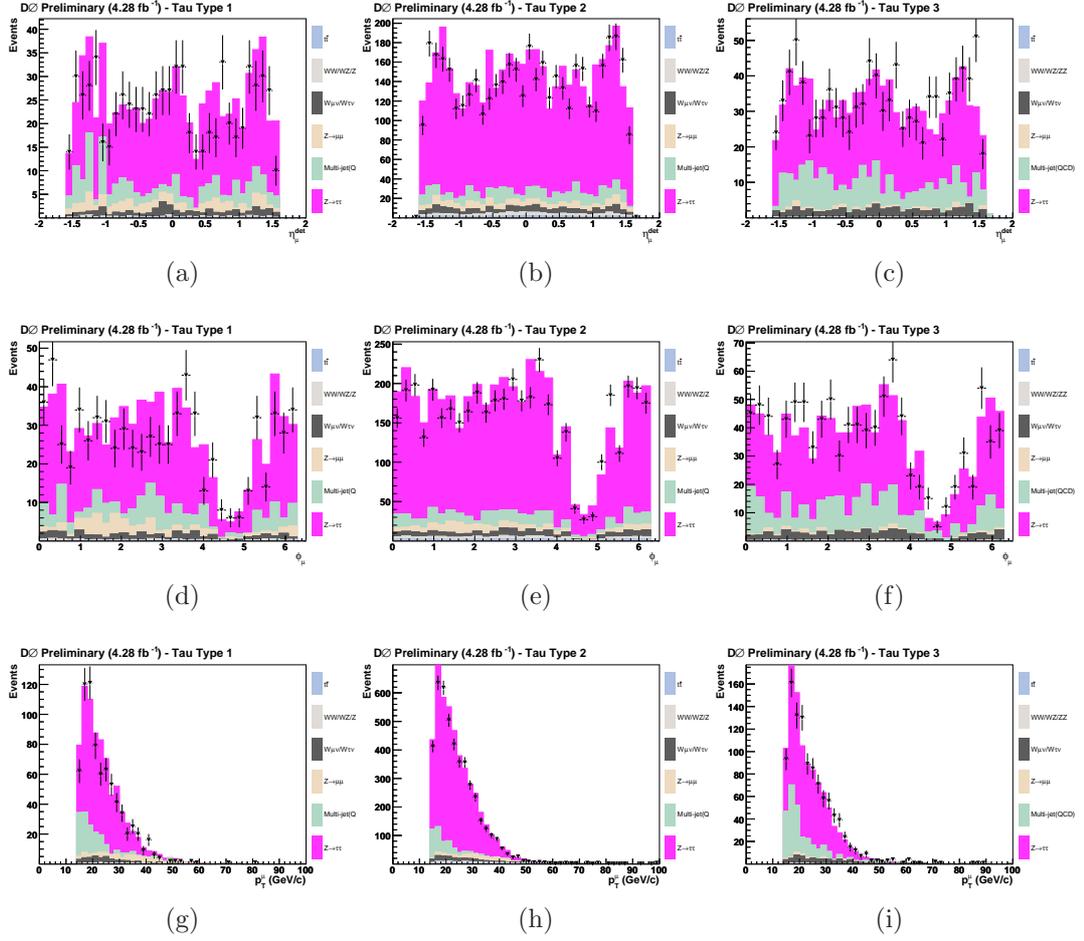
Figure 7: Distributions of η^μ, ϕ^μ, and p_T^μ (from top to bottom) for τ types 1, 2 3 (from left to right) for the Run IIb $\tau_\mu\tau_h$ channel.
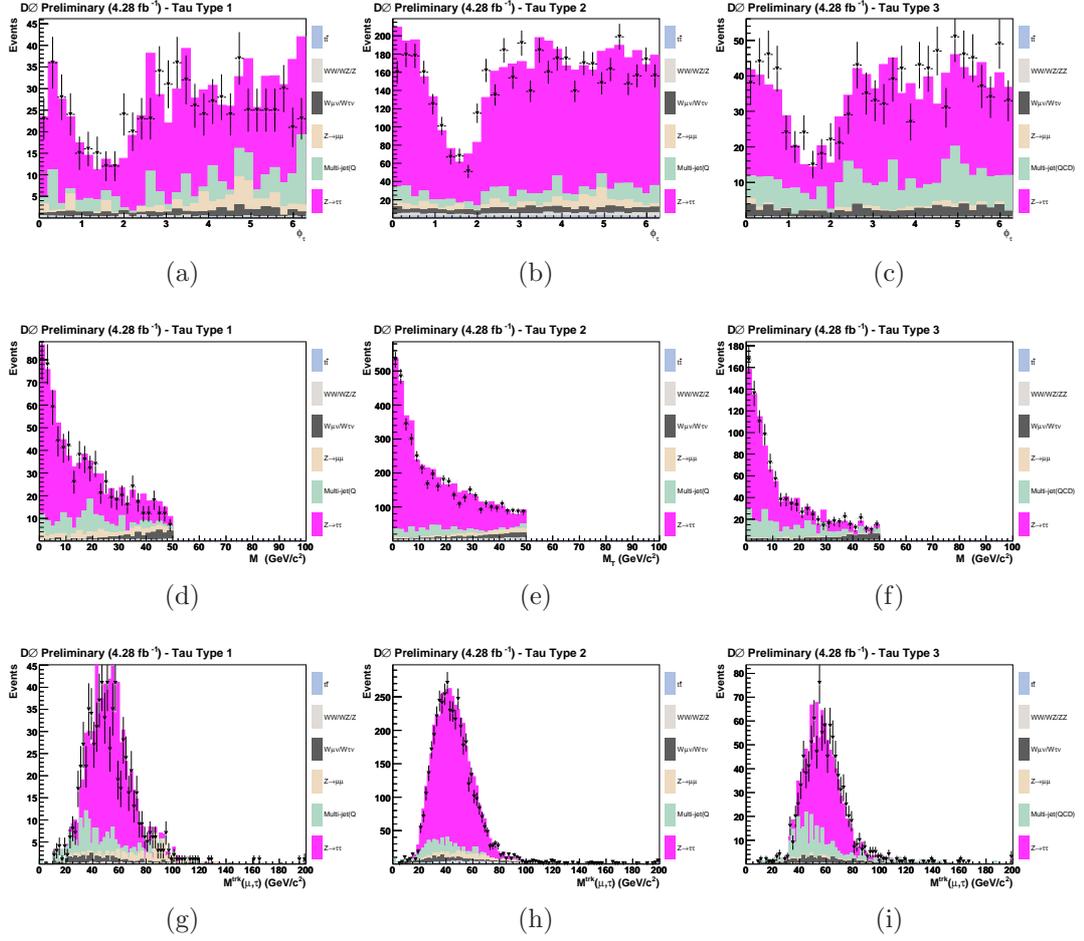
Figure 8: Distributions of ϕ^τ, M_T, and M_{trk} (from top to bottom) for τ types 1, 2 3 (from left to right) for the Run IIb $\tau_\mu\tau_h$ channel.

References

[1] H. P. Nilles, Rhys. Rept. **110**, 1 (1984).

[2] F. Halzen and A. Martin, *Quarks and Leptons* (Wiley, 1984).

[3] S. Abachi *et al.*, Phys. Rev. Lett. 74 (1995), 2632 (1995).

[4] F. Abe *et al.*, Phys. Rev. Lett. 74, 2626 (1995).

[5] W.-M. Yao *et al.*, [Particle Data Group], Journal of Physics G **33**, update for 2009 edition.

[6] P.W. Higgs, Phys. Rev. Lett. 13, 508 (1964).

[7] The Tevatron Electroweak Working Group for the CDF and DØ Collaborations, arXiv:1007.3178 [hep-ex].

[8] C. Quigg, *The Electroweak Theory*, Flavor Physics for the Millennium: TASI 2000, edited by J. L. Rosner (World Scientific, Singapore, 2001), pp. 3-67.

[9] J. Alcaraz *et al.*, [ALEPH, DELPHI, L3 and OPAL Collaborations, LEP Electroweak Working Group], updated for 2009 winter conferences, http://lepewwg.web.cern.ch/LEPEWWG/.

[10] J. A. Casas *et al.*, *Nucl. Phys.* **B436**, 3 (1995); Y.F. Pirogov and O.V.Zenin, *Eur. Phys. J.* **C10**, 629 (1999).

[11] R. Barate *et al.*, [ALEPH, DELPHI, L3 and OPAL Collaborations, LEP Working Group for Higgs Boson Searches], Phys. Lett. **B565**, 61 (2003).

[12] T. Aaltonen *et al.* [CDF and DØ Collaborations], arXiv:1007.4587 [hep-ex].

[13] G. Bernardi *et al.*, [Particle Data Group],
`http://pdg.ge.infn.it/reviews/rppref/mini/2010/pdghiggsv9.pdf`

[14] A. Djouadi Phys. Rept. **459** 1 (2008).

[15] J. Ellis *et al.*, Nucl. Phys. **B238**, 453 (1984).

[16] J.F. Gunion *et al.*, *The Higgs Hunter's Guide* (Addison-Wesley), 1990;
A. Djouadi, hep-ph/0503172, hep-ph/0503173

[17] M. Carena, S. Heinemeyer, C.E.M. Wagner, G. Weiglein Eur. Phys. J. **C45**,
797 (2006).

[18] M. Carena, S. Heinemeyer, C.E.M. Wagner, G. Weiglein Eur. Phys. J. **C26**,
601 (2003).

[19] T. Hahn *et al.*, hep-ph/0607308 (2006).

[20] O. Buchmueller, *et al.* Phys. Rev. **D81**, 035009 (2010).

[21] A. Djouadi, J. Kalinowski, and M. Spira, Comput. Phys. Commun. **108**, 56
(1998).

[22] T. Hahn, S. Heinemeyer, W. Hollik, H. Rzehak, G. Weiglein, MPP-2010-78,
KA-TP-21-2010, SFB-CPP-10-56, DESY-10-100, (2010).

[23] Run II Handbook, `http://www-bd.fnal.gov/lug/runII_handbook/`
`RunII_index.html`

[24] V.M. Abazov *et al.* DØ Collaboration, Nucl. Instrum. Meth. **A565:** 463-537
(2006).

[25] `http://d0server1.fnal.gov/Projects/Operations/`
`D0RunII_DataTaking.htm`

[26] E. Kajifasz [DØ Collaboration], Nucl. Intrum. Meth. **A511** 16 (2003).

[27] `http://d0server1.fnal.gov/projects/SciFi/cfthome.html`.

[28] D. Wood,
`http://d0server1.fnal.gov/projects/run2b/Management/`
`Posted_Info/PMG_Talks/Sep07_04/r2b_trig_PMG_07sep04_v3.ppt`.

[29] A. Schwartzman *et al.*, DØ Note **5192** (2006),
    `http://www-d0.fnal.gov/global_tracking/vertex.html`

[30] J. Hays, J. Mitrevski, C. Schwanenberger, and T. Toole, DØ Note **5105**
    (2006),
    `http://www-d0.fnal.gov/phys_id/emid/d0_private/emid.html`

[31] L. Duflot and M. Ridel, DØ Note **3923**, (2001).

[32] A. Harel, DØ Note **4919** (2005),
    `http://www-d0.fnal.gov/phys_id/jets/jetid.html`

[33] DØ Jet Energy Scale Group, DØ Note **5382** (2007),
    `http://www-d0.fnal.gov/phys_id/jes/public/plots_v7.1/index.html`

[34] O. Brandt *et al.*, DØ Note **6025** (2010),
    `http://www-d0.fnal.gov/computing/algorithms/muon/muon_algo.html`

[35] D. Chakraborty *et al.*, DØ Note **4210** (2003).

[36] M. Arov *et al.*, DØ Note **4773** (2005),
    `http://agenda.hep.manchester.ac.uk/getFile.py/access?contribId=58`
    `&sessionId=46&resId=0&materialId=slides&confId=1672`

[37] S. Protopopescu and P. Svoisky, DØ Note **5094** (2006),
    S. Levy, Springer Proceedings in Physics, **108** Part 4, 102-106 (2006)

[38] W. Fisher, DØ Note **5595** (2010),
    `https://plone4.fnal.gov/P1/D0Wiki/Members/szelitch/hwwmunujj/`
    `collie/CollieDocumentation.pdf.`

[39] T. Junk, Nucl. Instrum. Meth **A434** 435 (1999)

[40] `http://root.cern.ch/drupal/.`

[41] E. Nurse, PhD thesis, The University of Manchester, 2005.

[42] P. Calfayan, DØ Note **5329** (2008).

[43] K. Herner, A. Patwa, W.C. Yang, DØ Note **5699** (in preparation).

[44] M. Hohlfeld and M. Owen, DØ Note **5409** (2007).

[45] J. Hays *et al.*, DØ Note **5138** (2006).

[46] Common Sample Group,
`http://www-d0.fnal.gov/Run2Physics/cs/index.html`.

[47] B. Casey *et al.*, DØ Note **4958** (July 2006) and
G. Snow for the DØ Luminosity Group, DØ Note **5139** (October 2006),
`http://www-d0.fnal.gov/runcoor/RUN/run2_lumi.html`.

[48] T. Sjstrand, *Monte Carlo Generators*, hep-ph/0611247 (2006).

[49] T. Sjostrand *et al.*, *PYTHIA 6.3 Physics and manual*, hep-ph/0308153.

[50] M. L. Mangano, M. Moretti, F. Piccinini, R.Pittau and A. D. Polosa,
ALPGEN, a generation for hard multiparton processes in hadronic collisions,
hep-ph/0206193 (2002).

[51] R. D. S. Jadach, Z. Was and J. H. Kuhn, Comput. Phys. Commun. **73**, 361 (1993).

[52] D. Stump *et al.*, J. High Energy Phys. **10**, 046 (2003).

[53] V.M. Abazov *et al.* [DØ Collaboration], Phys. Rev. Lett. **101**, 071804 (2008).

[54] R. Brun *et al.*, *Geant3*, CERN-DD/EE/81-1 (1987).

[55] `http://www-d0.fnal.gov/computing/MonteCarlo/MonteCarlo.html`

[56] `http://www-d0.fnal.gov/computing/algorithms/`
`howto/howtoreco.html`

[57] `http://www-d0.fnal.gov/Run2Physics/cs/caf/`

[58] H. Schellman, DØ Note **5142** (2006).

[59] C. Balazs and C. P. Yuan, Phys. Rev. **D56**, 5558 (1997).

[60] M. Shamim, T. Bolton, DØ Note **5565** (2008).

[61] G. Hesketh, DØ Note **5786** (2008).

[62] C. Galea *et al.*, DØ Note **5533** (2007).

[63] K. Peters, DØ Note **5206** (2006).

[64] `http://www-d0.hef.kun.nl//askArchive.php?base=agenda&`
`categ=a09917&id=a09917s1t1/transparencies`

[65] `http://www-d0.hef.kun.nl//fullAgenda.php?ida=a092001&fid=57`

[66] `https://plone4.fnal.gov/P1/D0Wiki/physics/VplusJets/`

[67] M. Owen, S. Söldner-Rembold, and W.-C. Yang, DØ Note **5457** (2008), and
V. M. Abazov *et al.*, Phys. Rev. Lett. **101**, 071804 (2008), 0805.2491.

[68] G. Unal, ACTA PHYSICA POLONICA B **Vol. 38** No.3 (2006)

[69] DØ Collaboration, DØ Note **5484-CONF** (2007).

[70] The JES group, DØ Note **6011** (in preparation), and
`http://www-d0.fnal.gov/phys_id/jes/public/plots_v7.1/index.html`

[71] V. Buscher, J.-F. Grivaz, T. Nunnemann, and M. Wobisch,
DØ Note **4618** (2004),
`http://www-d0.fnal.gov/Run2Physics/WWW/documents/Reports/`
`PDFScale_Buescher.pdf`

[72] K. Harder, private communications, 2006.

[73] S. Schael *et al.*, [ALEPH, DELPHI, L3 and OPAL Collaborations, LEP
Working Group for Higgs Boson Searches], Eur. Phys. J. **C47**, 547 (2006).

[74] T. Aaltonen *et al.* [CDF Collaboration], Phys. Rev. Lett. **103**, 201801 (2009).

[75] CDF and DØ Collaborations [The TEVNPH Working Group], DØ Note
6036-CONF.

[76] V. M. Abazov *et al.*, [D0 Collaboration], DØ Note **5726-CONF** (2008).

[77] V. M. Abazov *et al.*, [D0 Collaboration], DØ Note **6083-CONF** (2010).

[78] A. Anastassove *et al.*, CDF Note **9071** (2007).

[79] T. Wright and D. Amidei, CDF Note **9284** (2008).

[80] U. Aglietti *et al.*, **FERMILAB-CONF-06-467-E-T** (2007).