

# Search for MSSM Higgs Bosons in Tau Final States in $p\bar{p}$ Collisions at DØ

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

|          |  |           |
|----------|--|-----------|
| <b>1</b> | <b>Theoretical Background</b>                              | <b>3</b>  |
| 1.1      | The Standard Model . . . . .                               | 3         |
| 1.1.1    | Matter in the Standard Model . . . . .                     | 4         |
| 1.1.2    | Force Mediating Particles in the Standard Model . . . . .  | 6         |
| 1.1.3    | Electroweak Theory and the Higgs Mechanism . . . . .       | 7         |
| 1.2      | Why Go Beyond the Standard Model? . . . . .                | 9         |
| 1.3      | Supersymmetry . . . . .                                    | 12        |
| 1.3.1    | The Higgs Sector in SUSY . . . . .                         | 13        |
| <b>2</b> | <b>Experimental Apparatus</b>                              | <b>23</b> |
| 2.1      | The Tevatron . . . . .                                     | 23        |
| 2.1.1    | Proton Source . . . . .                                    | 23        |
| 2.1.2    | Antiproton Source . . . . .                                | 24        |
| 2.1.3    | Tevatron Collider . . . . .                                | 25        |
| 2.2      | The DØ Detector . . . . .                                  | 26        |
| 2.2.1    | Central Tracking System . . . . .                          | 28        |
| 2.2.2    | Preshower Detectors . . . . .                              | 31        |
| 2.2.3    | DØ Calorimeter . . . . .                                   | 32        |
| 2.2.4    | Intercryostat Detectors . . . . .                          | 33        |
| 2.2.5    | Muon System . . . . .                                      | 34        |
| 2.2.6    | Luminosity Monitors . . . . .                              | 36        |
| 2.2.7    | Trigger System . . . . .                                   | 37        |
| <b>3</b> | <b>Object Reconstruction</b>                               | <b>40</b> |
| 3.1      | Tracks and Vertex Reconstruction . . . . .                 | 40        |
| 3.2      | Electron Identification . . . . .                          | 40        |
| 3.3      | Jet Identification and Missing Transverse Energy . . . . . | 43        |
| 3.4      | Muon Identification . . . . .                              | 44        |

|          |   |           |
|----------|---|-----------|
| 3.4.1    | Hadronic Tau Identification . . . . .                     | 46        |
| <b>4</b> | <b>Triggers and Data Samples</b>                          | <b>50</b> |
| 4.1      | Trigger and Efficiency Measurements . . . . .             | 50        |
| 4.2      | Muon Triggers . . . . .                                   | 51        |
| 4.2.1    | Level 1 Muon Trigger (L1MU) . . . . .                     | 53        |
| 4.2.2    | Level 2 Muon Trigger (L2MU) . . . . .                     | 54        |
| 4.2.3    | Level 3 Muon Trigger (L3MU) . . . . .                     | 54        |
| 4.3      | Run IIa Electron Triggers . . . . .                       | 54        |
| 4.4      | Data Samples . . . . .                                    | 55        |
| 4.4.1    | MUinclusive Skims . . . . .                               | 56        |
| 4.4.2    | EMinclusive skims . . . . .                               | 56        |
| <b>5</b> | <b>Monte Carlo Simulation and Corrections</b>             | <b>58</b> |
| 5.1      | Monte Carlo Simulation Method . . . . .                   | 58        |
| 5.2      | Electron Efficiency Correction . . . . .                  | 61        |
| 5.3      | Muon Efficiency Correction . . . . .                      | 64        |
| 5.4      | Tau Corrections . . . . .                                 | 66        |
| 5.4.1    | Tau Efficiency Correction . . . . .                       | 66        |
| 5.4.2    | Tau Energy Scale Correction . . . . .                     | 68        |
| 5.5      | MC Samples . . . . .                                      | 77        |
| 5.5.1    | Higgs Signal Monte Carlo . . . . .                        | 77        |
| 5.5.2    | Background Monte Carlo . . . . .                          | 77        |
| <b>6</b> | <b>Background Estimation</b>                              | <b>84</b> |
| 6.1      | Multi-jet (QCD) Background Estimation . . . . .           | 85        |
| 6.2      | Background Modeling in $\tau_\mu\tau_h$ Channel . . . . . | 87        |
| 6.2.1    | Multi-jet in $\tau_\mu\tau_h$ Channel . . . . .           | 87        |
| 6.3      | Background Modeling in $\tau_e\tau_h$ Channel . . . . .   | 91        |
| 6.3.1    | Multi-jet in $\tau_e\tau_h$ Channel . . . . .             | 92        |
| <b>7</b> | <b>Event Selections</b>                                   | <b>95</b> |
| 7.1      | Preselection Criteria . . . . .                           | 95        |
| 7.1.1    | $\tau_\mu\tau_h$ Channel . . . . .                        | 95        |
| 7.1.2    | $\tau_e\tau_h$ Channel . . . . .                          | 97        |
| 7.2      | Final Selection Cuts . . . . .                            | 98        |
| 7.2.1    | $\tau_\mu\tau_h$ Channel . . . . .                        | 99        |

|          |  |            |
|----------|--|------------|
| 7.2.2    | $\tau_e\tau_h$ Channel . . . . .                                   | 102        |
| <b>8</b> | <b>Results</b>   | <b>106</b> |
| 8.1      | Final Discriminant . . . . .                                       | 106        |
| 8.2      | Cross Section Limits . . . . .                                     | 107        |
| 8.2.1    | Limit Setting Technique and Collie . . . . .                       | 107        |
| 8.2.2    | Systematic Uncertainties . . . . .                                 | 112        |
| 8.2.3    | $\tau_e\tau_h$ Channel with $1.08\text{ fb}^{-1}$ Data . . . . .   | 114        |
| 8.2.4    | $\tau_\mu\tau_h$ Channel with $5.36\text{ fb}^{-1}$ Data . . . . . | 115        |
| 8.3      | Translation into the MSSM . . . . .                                | 117        |
| 8.4      | Comparison with Previous Results and Other Searches . . . . .      | 120        |
| <b>9</b> | <b>Conclusions</b>   | <b>125</b> |

# Chapter 1

## 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) and is 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 supersymmetric extensions of the Standard Model, which will be introduced in Sect. 1.3. Further details concerning Supersymmetry can be found in many references, for example, see [1]. In this thesis, the  $c = 1$  is chosen for all units used.

### 1.1 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 [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 are explained by grouping two major theories, *Quantum Electrodynamics* (QED) and *Quantum Chromodynamics* (QCD). Each type of elementary particle is described in terms of a field  $\Phi$  via *quantum field theories* (QFT). 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 a phase change,  $\phi$ , 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  $U(1)_{\text{em}}$ . A symmetry group is called “Abelian” if it is commutative. The symmetry of the  $U(1)_{\text{em}}$  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  $U(1)_{\text{em}}$  symmetry after Spontaneous Symmetry Breaking (SSB). However, the Standard Model is a non-abelian gauge theory described by the symmetry group  $U_Y(1) \otimes SU_L(2) \otimes SU(3_c)$ ;  $SU(3_c)$  is the gauge group of QCD, the theory of strong interactions which pertain to interactions due to colour charges of quarks. In particular,  $U_Y(1) \otimes SU_L(2) \otimes SU_c(3)$  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. 1.1.1 and 1.1.2.

### 1.1.1 Matter in the Standard Model

The twelve predicted spin  $\frac{1}{2}$  fermions within the SM are classified by the type of interactions and therefore are categorized 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 the other with charge  $Q = -1$ , and a pair of quarks, one with charge  $Q = \frac{2}{3}$  and one with charge  $Q = -\frac{2}{3}$ . Each pair corresponds to a weak isospin doublet. The first generation included the electrons, electron neutrinos, up quarks and the down quarks, while the second compromises muons, muon neutrinos, charm quarks and strange quarks. Particles within the third generation

| Generation | Quarks          |                                     |                | Leptons                        |                        |            |
|------------|-----------------|-------------------------------------|----------------|--------------------------------|------------------------|------------|
|            | 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 ( $\nu_e$ )         | $< 2.2 \times 10^{-6}$ | 0          |
| Second     | strange ( $s$ ) | $105^{+25}_{-35}$                   | $-\frac{1}{3}$ | muon ( $\mu$ )                 | 105.66                 | -1         |
|            | charm ( $c$ )   | $1270^{+70}_{-110}$                 | $\frac{2}{3}$  | $\mu$ neutrino ( $\nu_\mu$ )   | $< 0.17$               | 0          |
| Third      | bottom ( $b$ )  | $4200^{+170}_{-70}$                 | $-\frac{1}{3}$ | tau ( $\tau$ )                 | $1776.84 \pm 0.17$     | -1         |
|            | top ( $t$ )     | $171.3 \pm 1.1 \pm 1.2 \times 10^3$ | $\frac{2}{3}$  | $\tau$ neutrino ( $\nu_\tau$ ) | $< 15.5$               | 0          |

Table 1.1: The matter fermions of the Standard Model [5].

are  $\tau$  leptons,  $\tau$  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 about 171.3 GeV [5], 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 1.1 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 1.2 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]:

$$Q = I_z + \frac{B + S + C + B' + T}{2}. \quad (1.1)$$

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 ( $\mu^-$ ) and taus ( $\tau^-$ ) interact electromagnetic-ally, gravitationally and weakly. Since the  $\tau$  leptons are of particular importance to this thesis, more details about  $\tau$  properties and the reconstruction methods of  $\tau$  lep-

| 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 1.2: Additive quantum numbers of the quarks. [5]

tons at DØ will be described further in Sect. 3.4.1. Neutrinos ( $\nu_e$ ,  $\nu_\mu$  and  $\nu_\tau$ ) do not carry electric charge and interact only through the weak nuclear force making the detection of them difficult in typical collider detectors.

### 1.1.2 Force Mediating Particles in the Standard Model

Table 1.3 lists the four fundamental forces existing in Nature. Here, gravity is the weakest one of the four and is normally not included within the SM, so only three are explained by the SM. Within the SM, fundamental particles interact with each other by exchanging force carriers with the values of their spin quantum number equal to one. G

Its force carrier the gravitons, has spin quantum number equal to two.

| Interaction           | Describing Theory  | Mediator(s)          | Relative strength | Acting range (m) |
|-----------------------|--------------------|----------------------|-------------------|------------------|
| Gravity               | General Relativity | gravitons            | 1                 | $\infty$         |
| Weak Force            | Electroweak Theory | $W^-$ , $W^+$ , $Z$  | $10^{25}$         | $10^{-18}$       |
| Electromagnetic Force | QED                | photons ( $\gamma$ ) | $10^{36}$         | $\infty$         |
| Strong Force          | QCD                | gluons               | $10^{38}$         | $10^{-15}$       |

Table 1.3: The forces and the corresponding mediators. The mediating particle of the gravitation, gravitons, have not been observed yet.

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 ( $\gamma$ ) 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 corre-



sponding neutrino via the exchange of the  $W^\pm$  bosons. The  $Z$  and  $W$  bosons have a mass of  $91.1876 \pm 0.0021$  GeV and  $80.398 \pm 0.025$  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, whose value 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. 1.1.3. The strong force acts on colour charged particles. It 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.

### 1.1.3 Electroweak Theory and the Higgs Mechanism

Electroweak theory describes the unification of the weak with the electromagnetic interactions. In this theory, 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 ( $\gamma$ ),  $W^\pm$  and  $Z$ . The weak hypercharge,  $Y$ , is defined through the Gell-Mann-Nishijima formula,

$$Q = L + \frac{1}{2}Y \quad (1.2)$$

by the generator of  $U(1)_{\text{em}}$ , 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 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

$$\phi \equiv \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \quad (1.3)$$

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

$$\langle\phi_0\rangle = \begin{pmatrix} 0 \\ v/\sqrt{2} \end{pmatrix} \quad (1.4)$$

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

$$\begin{aligned} M_W^2 &= \frac{g^2 v^2}{2} \\ M_Z^2 &= \frac{M_W^2}{\cos^2 \theta_W}, \end{aligned} \quad (1.5)$$

where  $\theta_W$  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  $U(1)_{\text{em}}$  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  $134 \text{ GeV} \leq m_H \leq 177 \text{ GeV}$  [9]. Figure 1.1 illustrates the upper and lower bounds on the Higgs boson mass at different energy scale from the following conditions:

- Triviality condition which keeps the coupling  $\lambda$  finite.
- The requirement that the EW vacuum is absolute minimum, which promises the stability of the Higgs potential.

Precision electroweak measurements are also shown in Fig. 1.1 and provide a upper bound on the Higgs boson mass at lower energy scale.

The direct search of 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 extend 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 is  $5.9 \text{ fb}^{-1}$  data [12]. Results are shown in Fig. 1.2.

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. Figure 1.3(a) shows indirect experimental constraints from the top mass,  $m_t$ , and  $W$  mass,  $M_W$ , for different Higgs masses. Further, a  $\Delta\chi^2$  fit for  $m_H$  using precision EW measurements is given in Fig 1.3(b). The 95% CL exclusions performed by LEP2 are also shown. Both direct and indirect determinations are consistent with each other and favour a light Higgs boson.

## 1.2 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 mass has the following form with first-order corrections [13]:

$$m_H^2 = (m_H^2)_0 + \frac{kg^2\Lambda^2}{16\pi^2}. \quad (1.6)$$

Here, the first term  $(m_H)_0$  is a fundamental parameter of theory, and the second term represents a one-loop corrections for the EW gauge coupling constant,  $g$ . The parameter  $\Lambda$  defines the reference energy scale for new physics beyond the SM and is sensitive to fermion masses and dimensionless couplings. Within nature,  $\Lambda$  is generally chosen as the Planck mass given by [13]:

$$M_{Planck} = \left(\frac{\hbar c}{G}\right)^{1/2} \approx 1.22 \times 10^{19} \text{ GeV}, \quad (1.7)$$

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,  $\Lambda$ , is much larger than the EW scale, resulting in large Higgs boson. In particular, the SM is expected to be embedded in a more fundamental theory between the

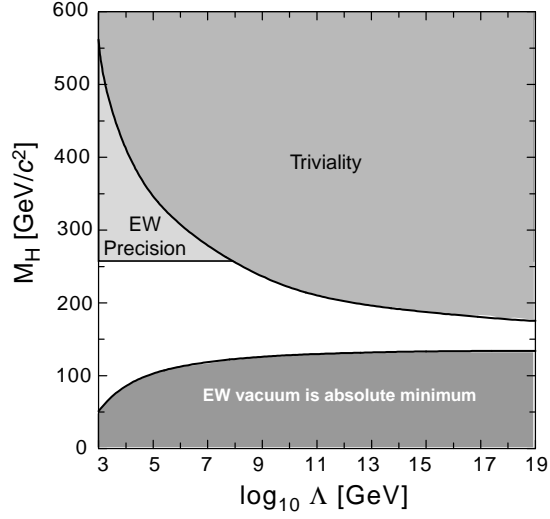


Figure 1.1: Theoretical upper bound from triviality conditions and the lower bound follows the requirement that  $V(v) < V(0)$ . Also shown as the upper bound in low energy scale is the limit from precision EW measurement with 95% confidence level (CL). [10]

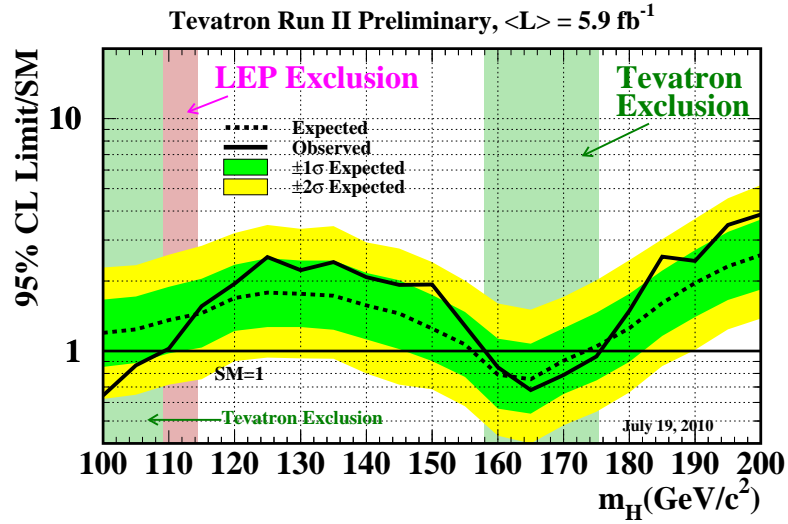


Figure 1.2: Upper bound on the SM Higgs boson cross section obtained by the Tevatron experiment and the direct exclusion from the LEP2 experiment. The coloured bands show the 68% and 95% CL probability bands of the expected upper bound from the Tevatron data.

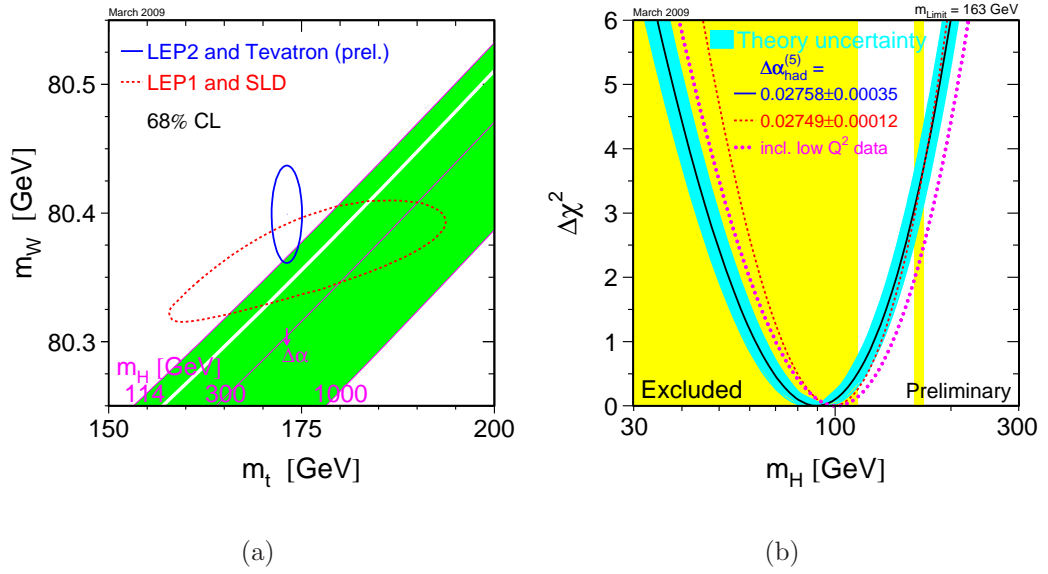


Figure 1.3: (a) Comparison of the indirect and the direct measurements of the Higgs boson mass. The relationship of  $m_t$  and  $m_W$  in the SM is also shown as a function of the Higgs boson mass. (b)  $\Delta\chi^2$  (black) for the fit to the Higgs boson mass from the electroweak measurements. The excluded region by LEP2 experiment at 95% CL is shown as yellow band and the blue band provides the theory uncertainty on the prediction.

EW energy scale to the Planck scale in order to stabilize 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 there is no particle in the SM that can be a candidate of dark matter and explains this cosmological observations yet.

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

## 1.3 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 an unit. This is defined by an operator,  $Q$ , 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 elementary particles and adding possible new interactions 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 1.4 illustrates the one-loop corrections of fermions and those scalar partners to the Higgs boson mass. This naturally provides a solution of the hierarchy problem keeping the existence of a low mass Higgs boson reasonable. However, the cancellation of the quadratical  $\Lambda$  terms in the Higgs boson mass leaves a dependency 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 constraints 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. 1.3.1.

Not only being one of the solutions of the hierarchy problem, MSSM also

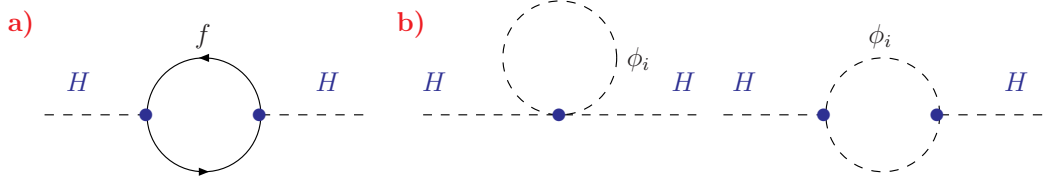


Figure 1.4: Diagrams for the one-loop contributions of (a) fermions and (b) SUSY scalar partners to the Higgs boson mass. [14]

provides good explanations for dark matter. In MSSM,  $R$ -parity is defined as

$$R = (-1)^{3(B-L)+2S}, \quad (1.8)$$

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 interaction. Therefore, it is possible that a stable LSP can be promising candidate for dark matter.

### 1.3.1 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 scalar  $H, h$  where conventionally  $M_h < M_H$  after electroweak symmetry breaking [16]. The masses  $H$  and  $h$  are degenerated with  $A$  in most parameter space, and all three neutral Higgs bosons, including both the CP-odd and CP-even states, are denoted as  $\phi$ . 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  $\tan\beta$ ;
- 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:

$$v_u^2 + v_d^2 = v^2 \approx (174 \text{ GeV})^2, \quad (1.9)$$

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

$$\tan\beta \equiv v_u^2/v_d^2. \quad (1.10)$$

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

$$m_{h,H}^2 = \frac{1}{2}(m_A^2 + m_Z^2) \mp \sqrt{(m_A^2 - m_Z^2)^2 + 4m_Z^2 m_A^2 \sin^2(2\beta)}, \quad [16] \quad (1.11)$$

$$m_{H^\pm}^2 = m_A^2 + m_W^2. \quad (1.12)$$

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

$$m_h \leq m_Z |\cos 2\beta|. \quad (1.13)$$

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]:

$$m_h \leq 135 \text{ GeV}. \quad (1.14)$$

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  $\Delta_B$  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  $\sim \tan^2 \beta$  compare to the Higgs cross section predicted by the SM.

Although the Higgs sector in 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  $M_{\text{SUSY}}$ ;
- $X_b \equiv A_b - \mu \tan \beta$ ;
- $X_t \equiv A_t - \mu / \tan \beta$ , where  $A_{t/b}$  are the trilinear Higgs sfermion couplings;
- the higgsino mass parameter,  $\mu$ , which is also know 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 benchmark scenarios,  $m_h^{\text{max}}$  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]. The  $m_h^{\text{max}}$  scenario maximizes the mass of the lightest MSSM Higgs boson for a given  $\tan \beta$  and  $M_A$  and hence provides 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 and with a higher SUSY mass scale to avoid the LEP Higgs boson mass bounds. Varying the value and sign of  $\mu$  demonstrates the effect of radiative corrections on the production and decay processes, therefore, the dependency on  $\mu$  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  $(M_A\text{-}\tan \beta)$  plane for the two scenarios for both positive and negative values of  $\mu$ . The common choices for the values of the SUSY parameters on each scenarios are listed in

| Parameter         | $m_h^{\max}$  | No-mixing     |
|-------------------|---------------|---------------|
| $X_t$             | 2 TeV         | 0 TeV         |
| $\mu$             | $\pm 0.2$ TeV | $\pm 0.2$ TeV |
| $M_2$             | 0.2 TeV       | 0.2 TeV       |
| $m_{\tilde{g}}$   | 0.8 TeV       | 1.6 TeV       |
| $M_{\text{SUSY}}$ | 1 TeV         | 2 TeV         |

Table 1.4: The MSSM parameters for the  $m_h^{\max}$  and *no – mixing* scenarios [18].

Table 1.4. **FeynHiggs** is a program used in this analysis to interpret cross-section limits into limits on MSSM  $M_A$ - $\tan\beta$  parameter space using different scenarios ??.

The production processes of the MSSM neutral Higgs bosons,  $\phi$  ( $=h, H$  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 relevant Feynman diagrams at tree-level are shown in Fig. 1.5.

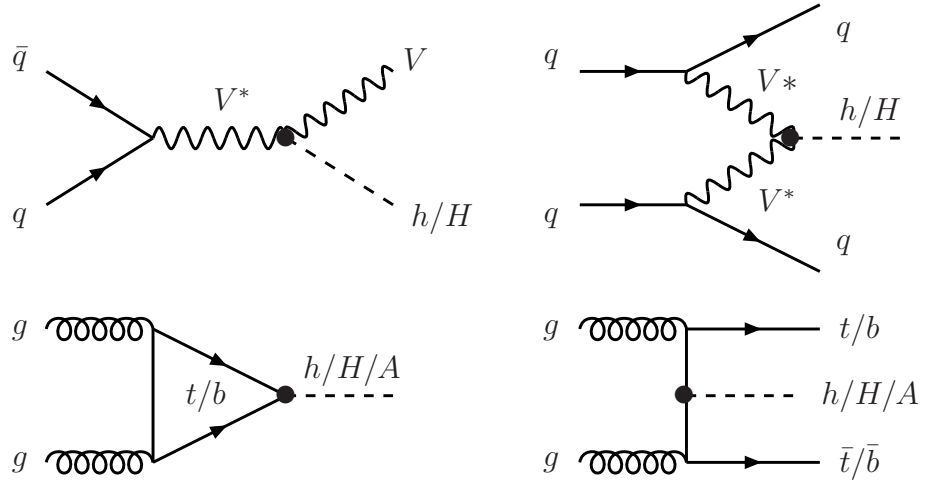


Figure 1.5: The dominant MSSM neutral Higgs production mechanisms in hadronic collisions.

Figure 1.6 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. As apparent from Fig. 1.6, the dominant production processes at the Tevatron are gluon fusion ( $gg \rightarrow \phi$ ) and the production of Higgs from  $b$  quarks ( $bb \rightarrow \phi$ ).

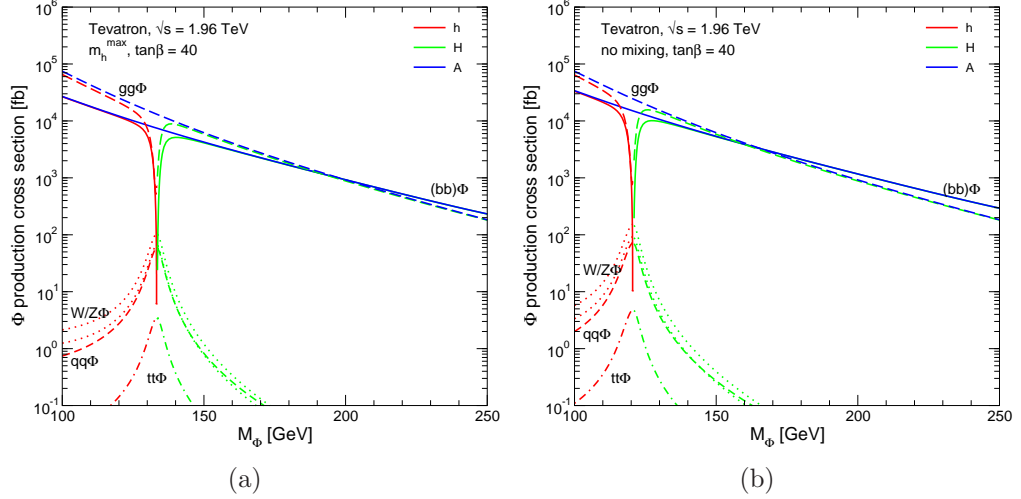


Figure 1.6: Higgs boson production cross sections as a function of Higgs boson mass at the Tevatron, (a) in the  $m_h^{\max}$  MSSM scenario and (b) in the no-mixing MSSM scenario, for  $gg \rightarrow \phi$  and  $b\bar{b} \rightarrow \phi$  [19].

Indirect search using EW constraints is studied at the Tevatron. It scans the MSSM parameter space and shows where the Higgs boson is most likely to be favoured as a function of the  $W$  boson and top quark masses at the 68% CL is given in Fig. 1.7. 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, such as MSSM model.

Considering the upper constrain on the masses Higgs bosons for  $m_h \ll W^+W^+$  threshold, the lightest Higgs boson is expected to decay mainly to fermion pairs. Therefore, the most dominant channels are  $\phi \rightarrow b\bar{b}, \tau\bar{\tau}$ . Figure 1.8 shows the decay branching ratios of the SM Higgs. Such 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. 1.9 to 1.11 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 1.12 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  $\tau$  pairs are strongly enhanced [14].

As evident from the decays shown in Fig. 1.9 to 1.11 in most parameter space, the tree-level branching ratio of the neutral MSSM  $h, H$  and  $A$  decays into pair of  $b$ -quarks is approximately 90% and 10% for  $\tau^+\tau^-$  decays. The two relevant Feynman diagrams for the decay of a CP-odd Higgs boson into  $\tau$  pairs are shown in Fig. 1.13.

A direct search for  $p\bar{p} \rightarrow \phi \rightarrow b\bar{b}$  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  $\tau$  pairs mediated by the production of a CP-odd Higgs boson with large  $\tan\beta$  for fixed masses of  $b$  and  $\tau$ . The relation is given by [17]:

$$\sigma(gg, b\bar{b} \rightarrow A) \times BR(A \rightarrow \tau^+\tau^-) \approx \sigma(gg, b\bar{b} \rightarrow A)_{SM} \frac{\tan^2\beta}{(1 + \Delta_B)^2 + 9}, \quad (1.15)$$

where  $\sigma(gg, b\bar{b} \rightarrow A)_{SM}$  is the value of the corresponding SM Higgs production cross sections, for cases the Higgs boson mass is  $M_A$ .

The program FeynHiggs [22, 21, 23, 24] is based on the results obtained in the Feynman-diagrammatic (FD) approach, which takes into account most of the first-order loop corrections. This program is used to evaluate the exclusions in the  $M_A$ - $\tan\beta$  plane in this analysis.

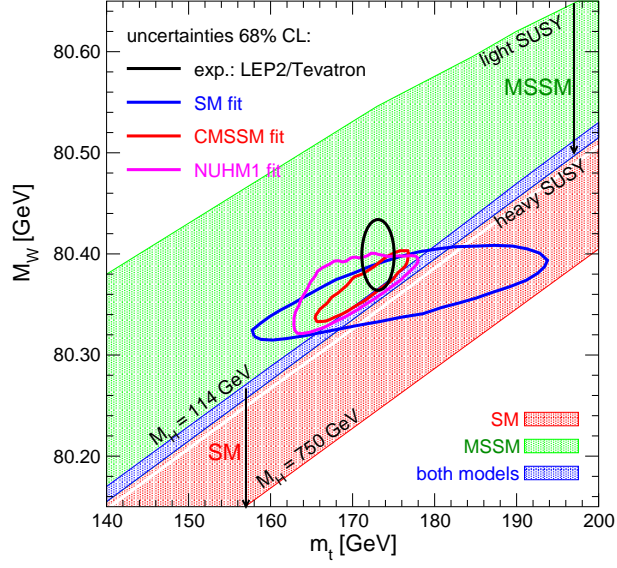


Figure 1.7: The current experimental measurements for  $W$  boson mass,  $M_W$ , and top quark mass,  $m_t$ , is shown as the black ellipse. The result is compared with the SM prediction region shown as red band and the green band for the region allowed in MSSM. The blue, red and pink ellipses are predicted by SM, CMSSM and NUHM1 fit with 68% CL. [29]

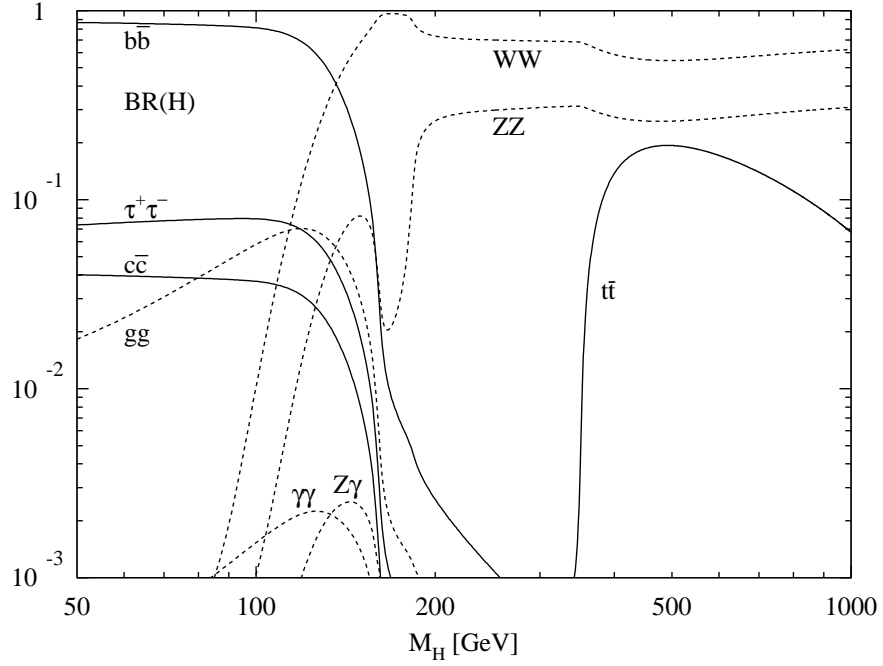


Figure 1.8: Main branching ratios for the SM Higgs boson [20] as a function of the Higgs boson mass.

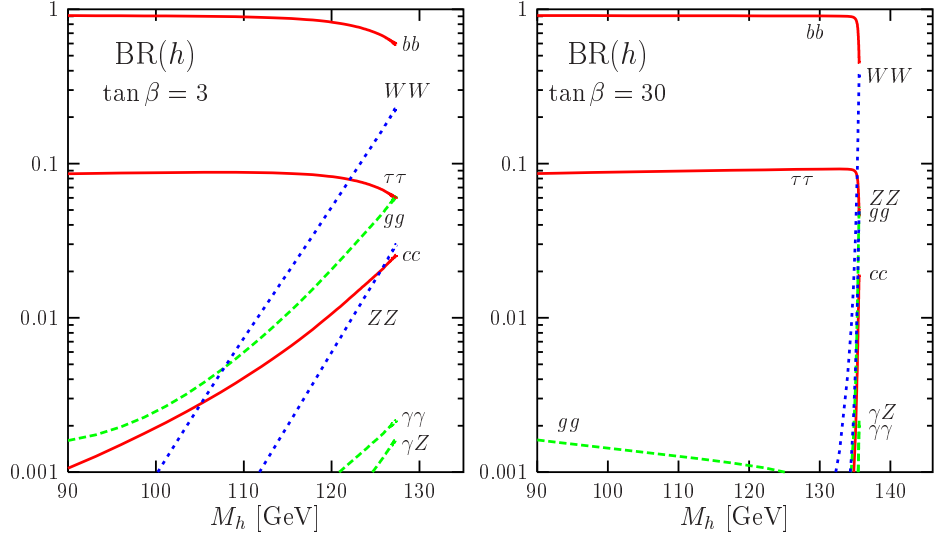


Figure 1.9: The decay branching ratios of the lighter CP-even MSSM  $h$  boson as a function of its mass for  $\tan\beta = 3$  (left) and  $\tan\beta = 30$  (right).

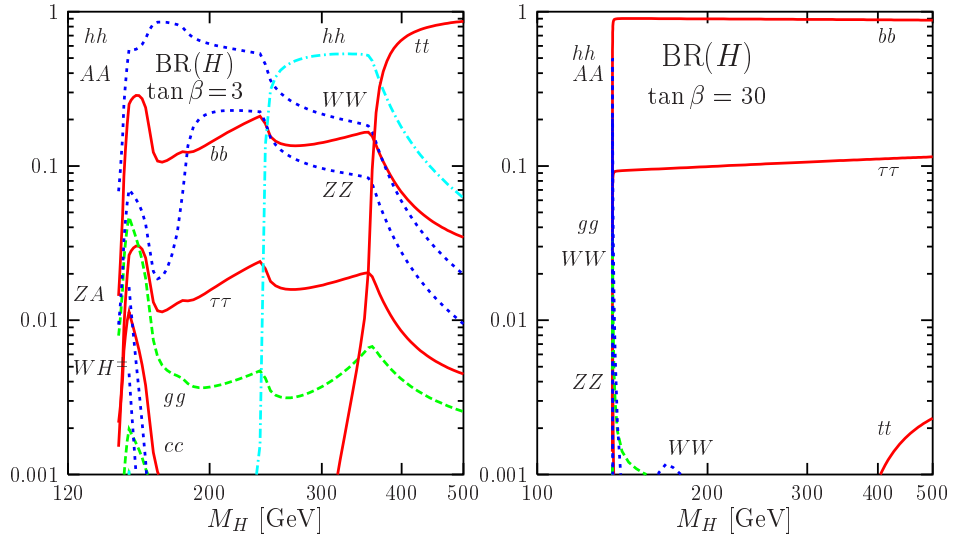


Figure 1.10: The decay branching ratios of the heavy CP-even MSSM  $H$  boson as a function of its mass for  $\tan\beta = 3$  (left) and  $\tan\beta = 30$  (right).

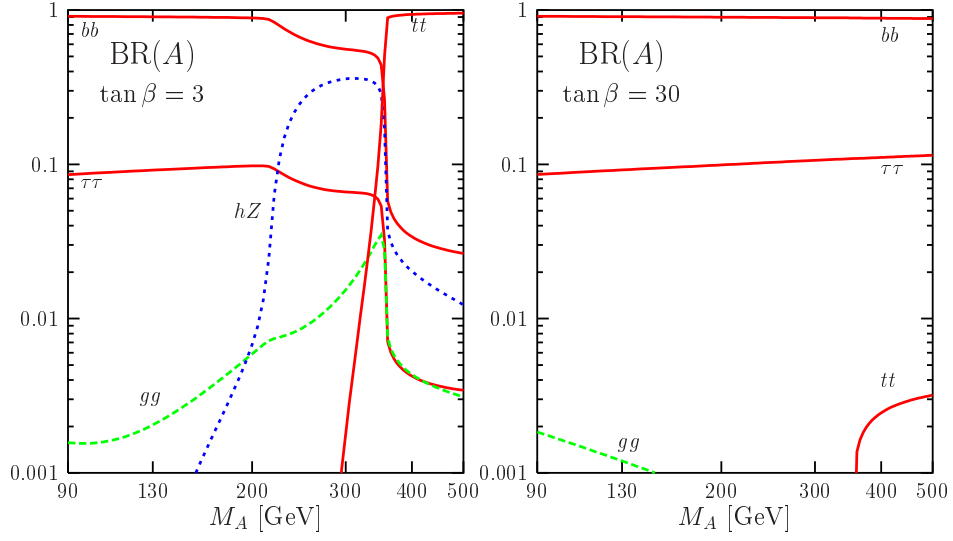


Figure 1.11: The decay branching ratios of the CP-odd MSSM  $h$  boson as a function of its mass for  $\tan\beta = 3$  (left) and  $\tan\beta = 30$  (right).

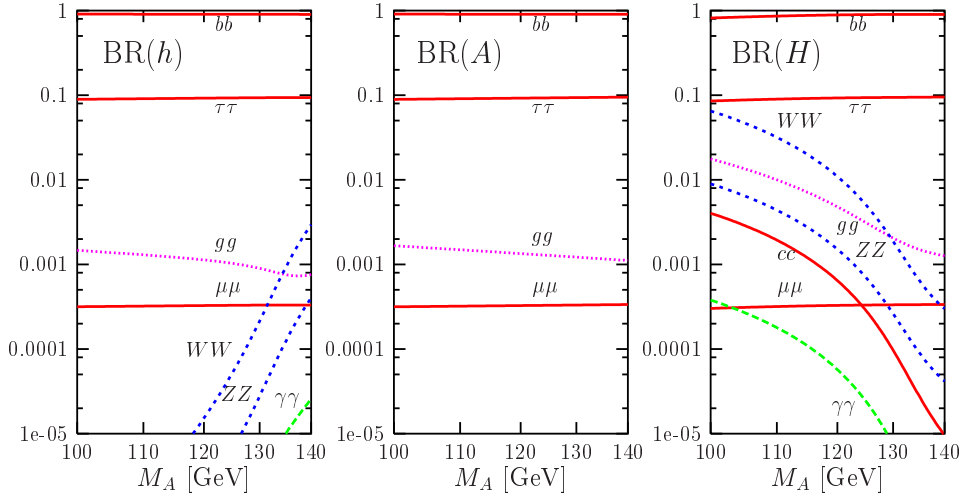


Figure 1.12: The decay branching ratios of the MSSM neutral Higgs bosons,  $h$ ,  $H$  and  $A$  as a function of  $M_A$  in the regime  $M_A \sim 120 - 140$  GeV with  $\tan\beta = 30$ .

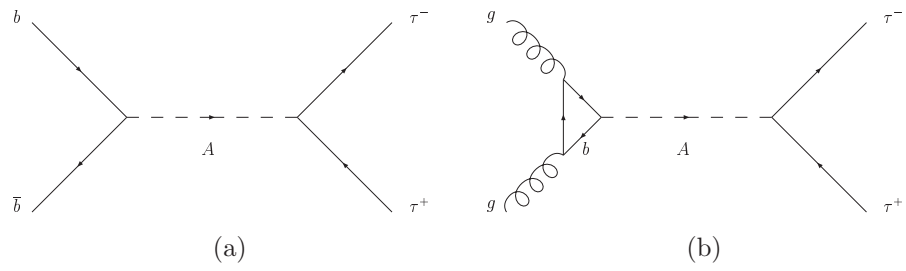


Figure 1.13: The two main production modes for the pseudo-scalar Higgs boson  $A$  in the MSSM, (a) annihilation of a pair of  $b$  quarks and (b) gluon fusion, via a loop of  $b$  quarks.



# Chapter 2

## Experimental Apparatus

The Tevatron Collider at the Fermi National Accelerator Laboratory near Chicago USA, was the world's highest energy particle collider until the Larger Hadron Collider (LHC) began operations in 2009. It enables head-on collisions of protons ( $p$ ) and anti-protons ( $\bar{p}$ ) at a centre-of-mass energy of 1.96 TeV. Operations for Run II started in early 2001 and by end-2009, the Tevatron reached integrated luminosities up to  $3.5 \times 10^{32} \text{ cm}^{-2}$ . The data produced in  $p\bar{p}$  collisions are collected by two major experiments, CDF and DØ. Both experiments are general-purpose particle detectors designed to identify particles created in the collisions, which include muons, photons or electrons,  $\tau$  leptons, particle jets and the presence of invisible neutrinos. A short description of the acceleration procedure mechanism is given in Section 2.1 [31]. Section 2.2 follows with a summary of each major component of the DØ detector [33].

### 2.1 The Tevatron

An overview of the acceleration chain in the Tevatron is shown in Figure 2.1. More detailed descriptions of each components can be found in [31].

#### 2.1.1 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 Cockcroft-Walton and injected into a 150-meter-long linear accelerator, *Linac*, through a transport line. Linac further accelerate these  $H^-$  ions to higher

energy of 400 MeV using Alvarez drift-tubes with an alternating electric field and a series of Radio Frequency (RF) accelerating cavities. The Booster is a synchrotron approximately 150 metres in diameter where a synchrotron is a circular accelerator that confines particles using magnetic field and allow them to accelerate by RF in a round orbit. Inside the Booster,  $H^-$  ions are passed through a carbon foil which strips electrons from the negative hydrogen ions [31], 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 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 accelerated to 150 GeV and stored in the Main Injector alone with the anti-proton beams described in the next section waiting for injection to the Tevatron.

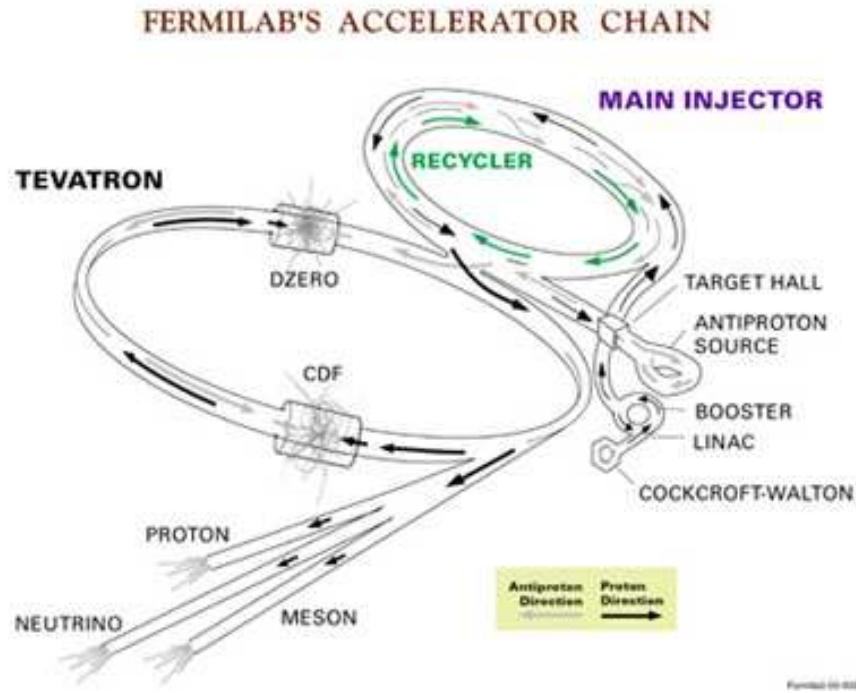


Figure 2.1: An overview of the acceleration system. [31]

### 2.1.2 Antiproton Source

The anti-proton production happens mainly in the Main Injector, and there are three main parts of the source: the Target Station, the Debuncher, and the Accumulator, plus the transport lines associated with each of these three devices.

Figure 2.2 provides a general layout of the Antiproton Source. The 120 GeV protons are first injected to small spot size of 0.1 mm on a nickel target, then a large amount of particles are produced during the collision. Anti-protons with the energy about 8 GeV are collected and focused by 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 not selected will pass to a beam dump. The Debuncher was 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 stores temperately the anti-protons. Once a sufficient number of anti-protons in the Accumulator is accumulated, they are passed to the Main Injector and accelerated to 150 GeV.

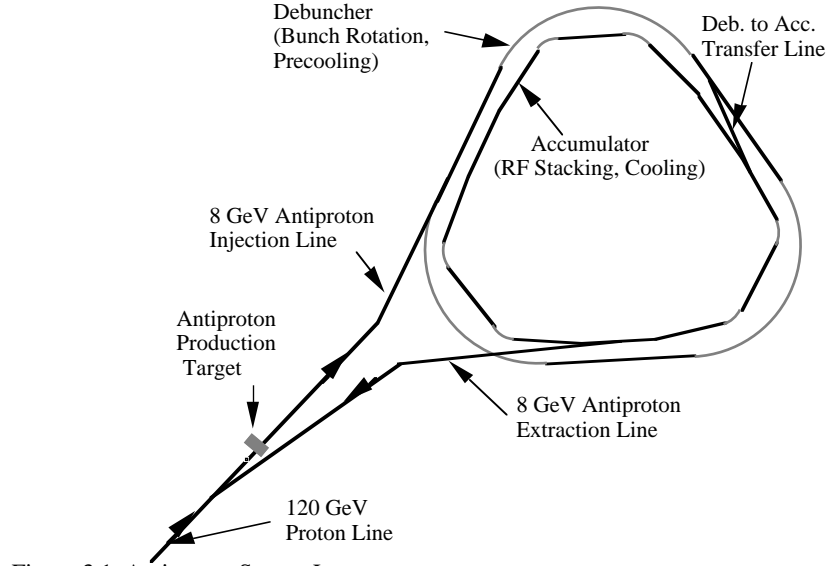


Figure 3.1. Antiproton Source Layout

Figure 2.2: Anti-proton Source layout [31].

### 2.1.3 Tevatron Collider

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 which fully replaces the Main Ring present during Run I. Superconducting magnets located in series within the Tevatron tunnel allow the  $p\bar{p}$  beam to be simultaneously accelerated to a centre-of mass energy of 1.8 TeV. During each store, the thirty-six proton bunches ( $150 \times 10^9$  protons each) are

injected into the Tevatron followed by the same amount of anti-proton bunches ( $50 \times 10^9$  anti-protons each). The beams circulate the Tevatron in opposite direction and are organized in three groups known as super-bunches. The beams are accelerated to reach instantaneous luminosities of up to  $\sim 350 \times 10^{30} \text{ cm}^{-2}\text{s}^{-1}$ . Finally, the beams are focused by special superconducting quadrupoles, and the protons and anti-protons are collided at the beam crossing points. Each interaction point is surrounded by a general purpose particle detector, CDF 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

$$\mathcal{L}(t) = \frac{fN^2}{A} \quad (2.1)$$

where  $f$  is the revolution frequency,  $N$  is the number of particles, and  $A$  is the cross-section area. The form used for 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, typically around 24-36 hours.

## 2.2 The DØ Detector

The DØ detector is a multi-purpose detector designed to capture the signature of particles spraying radially out at the collision points. Approximately  $120 \text{ pb}^{-1}$  of integrated luminosity was collected at DØ during Run I of the Fermilab Tevatron collider from 1992 to 1996 and Run II started in 2001 and runs to present. The average data taken efficiency is  $\sim 90\%$  and about  $10 \text{ fb}^{-1}$  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 Figure 2.3. [32]

The DØ detector is divided into three major parts: 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-fibre 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 most out layer of the DØ detector comprises the muon system, which is designed to detect the path and momentum of the muons.

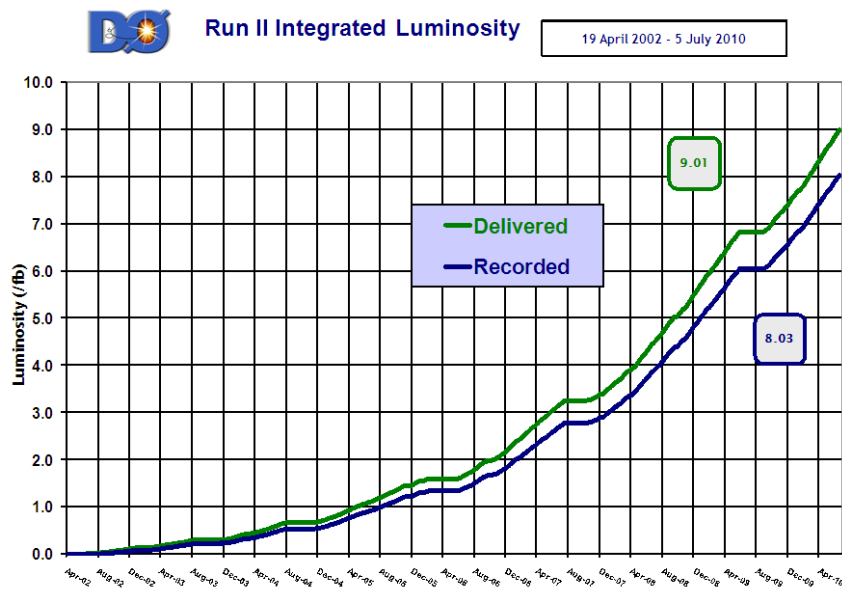


Figure 2.3: Run II integrated delivered (green) and recorded (blue) integrated luminosity.

Figure 2.4 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 helps select events that are of most interests to high-energy physics states. The following sections give a brief description of each part of the DØ detector, using a right-handed coordinate system in which the  $z$ -axis is along the proton direction and the  $y$ -axis is upward as shown in Figure 2.4. The angles  $\phi$  and  $\theta$  are the azimuthal and polar angles, respectively. The  $r$  coordinate denotes the perpendicular distance from the  $z$  axis. The pseudorapidity,

$$\eta = -\ln(\tan(\theta/2)) \quad (2.2)$$

approximates the true rapidity,

$$y = \frac{1}{2} \ln\left(\frac{E + p_z}{E - p_z}\right), \quad (2.3)$$

for finite angles in the limit that  $(m/E) \rightarrow 0$ . The term “forward” is used to describe the regions at large  $|\eta|$  whereas “central” corresponds to the small  $|\eta|$  region.

A more complete description of the detector can be found in [33].

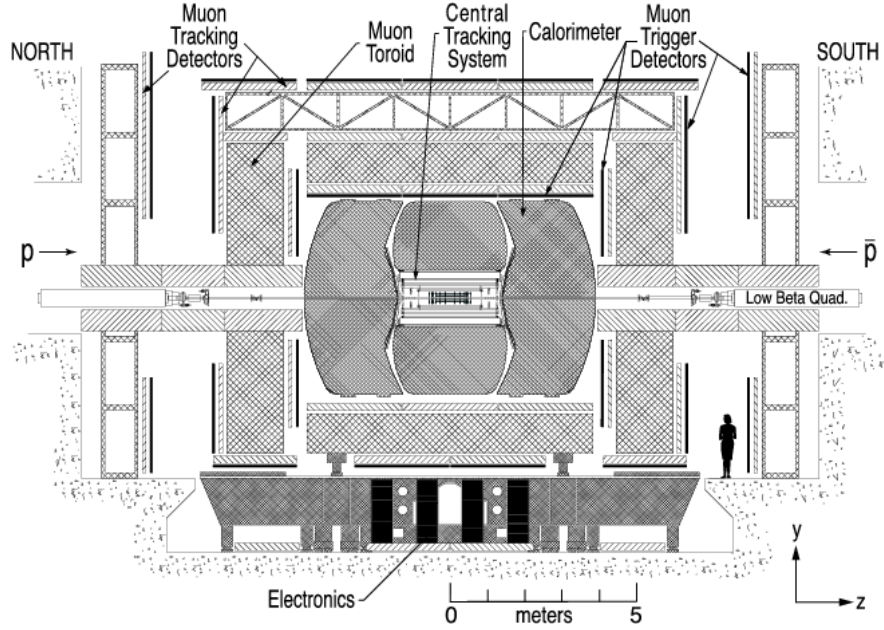


Figure 2.4: Schematic y-z view of the DØ detector. [33]

### 2.2.1 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 Figure 2.5.

The tracker comprises two sub-systems, which include the Silicon Microstrip Tracker (SMT), and the Central Fiber Tracker (CFT). A 2 T solenoidal magnet surrounds these two systems and bends charged particles in the  $r - \phi$  plane and allows the measurement of the energy to momentum ratio ( $E/P$ ) for efficient particle identification, calibration, and precise momentum measurement. The resolution of the tracker is about  $35 \mu m$  and better than  $15 \mu m$  for b-tagged jets with transverse momentum,  $p_T > 10$  GeV at normal incidence (with pseudorapidities  $|\eta| = 0$ ).

The SMT [34] provides tracking and vertexing information with the  $\eta$  coverage up to 3 but is limited at high  $\eta$ . It consists of six barrel modules in the central region, each having four silicon readout layers. The centres of the barrels are at  $|z| = 6.2, 19.0, 31.8$  cm. There are 72 silicon module, called “ladders” installed inside a barrel. The barrels are capped at high  $|z|$  by a disk called “F-disk”. The respective location of “F-disks” is at  $|z| = 12.5, 25.3, 38.2, 43.1, 48.1, \text{ and } 53.1$  cm. The F-disks have an outer radius of about 10 cm and are made up by 12 double-

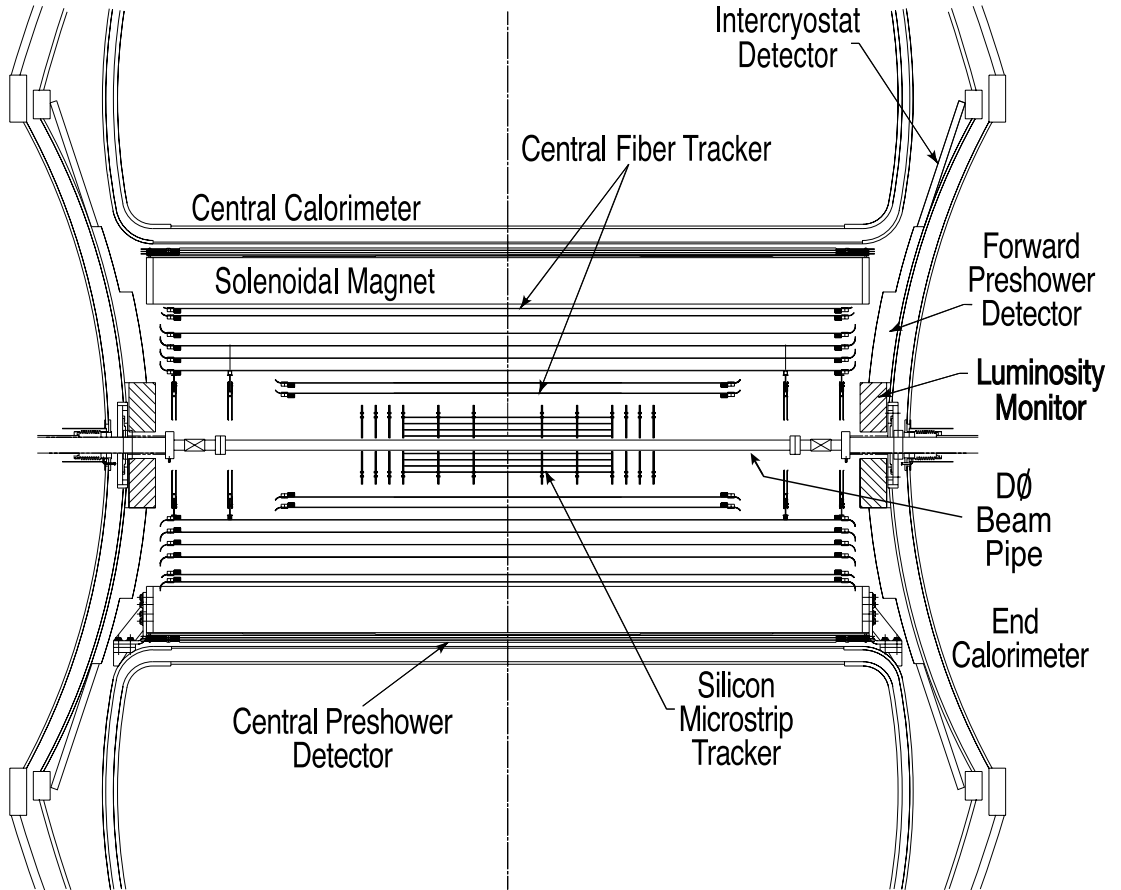


Figure 2.5: A schematic view of the central tracking system in the  $r-z$  plane. The geometrical view of the solenoid, the preshower detectors, luminosity monitor relative to the calorimeters are also shown. [33]

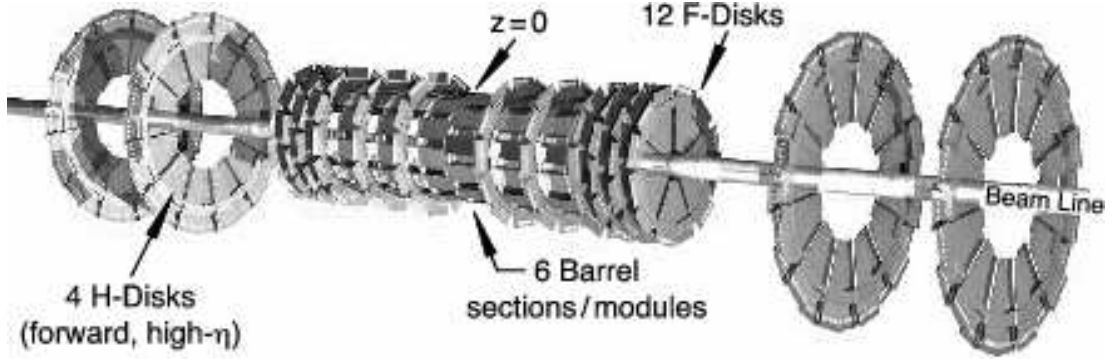


Figure 2.6: Geometrical design of the Silicon Microstrip Tracker (SMT). [33]

sided wedge detectors. In addition, two larger-diameter “H-disks” are located in each end at  $|z| = 100.4$  and  $121.0$  cm to extend the high  $|\eta|$  coverage. Each H-disk consists of 24 wedges and has an inner radius of 9.5 cm and outer radius of 26 cm. The main reason for this barrel-disk design is to capture the 3-dimension information of tracks transversing perpendicular to the detector surface. Such an arrangement allows the  $r - \eta$  coordinates to be measured by the barrels while the  $r - \phi$  coordinates are measured by disks.

The SMT readout strips consist of p-n junction diodes operated at reverse biases and 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. While a charged particle passes through the detector, the particle induces a charge separation within the material, creating an electron-hole pair, which is then collected by a capacitor in the readout. Figure 2.6 shows the layout of SMT detector.

The CFT [35] radially located outside the SMT measures the position of charged particles. It consists of scintillating fibres 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. Each fibre is  $835 \mu\text{m}$  in diameter and is capable of providing a resolution of about  $100 \mu\text{m}$ . As a charged particle passes through a fibre, the material will be ionized and thereby emit photons. The photon is then converted into electrical signals by visible light photon counters (VLPCs) located between the DØ detector at the DØ platform. A fast and continuous readout from the CFT is sent to the Level 1 trigger system. Signals that passed Level 1 are be used by Level 2 trigger decisions, while Level 3 uses full CFT readout information. Such a trigger system is described in Section 2.2.7



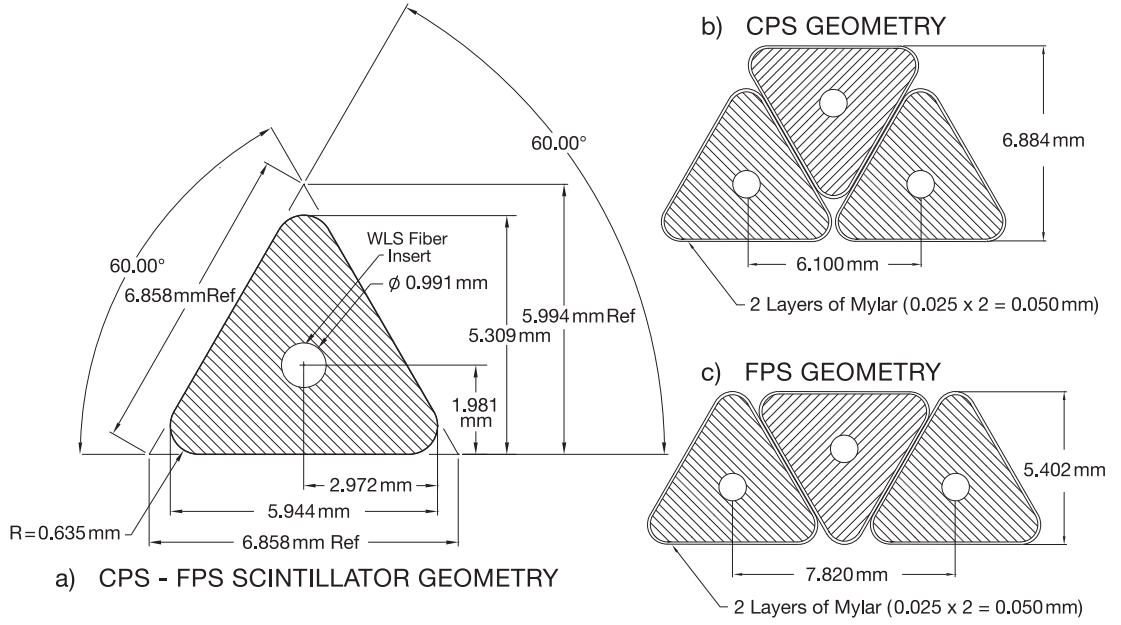


Figure 2.7: Cross sectional view and layout geometry of the triangular scintillator strips of both CPS and FPS. The central hole allows for insertion of the  $835 \mu\text{m}$  wave length shifting fibre described in the text.

### 2.2.2 Preshower Detectors

The preshower detectors help correct for the energy losses of particles due to material and thus alter the electromagnetic energy measurement in the central and end calorimeter. They further help improve the electron and photon identification at DØ. These detectors function as both calorimeters and tracking detectors. Information is also included in the Level 1 trigger because of their fast energy and position measurement. There are two preshower systems: one is the Central Preshower detector (CPS), which is located between the solenoid and the central calorimeter and covers the region  $|\eta| < 1.1$ . The other is the Forward Preshower detector (FPS), which is made up of two main modules, each located on the faces of the end calorimeters, and covering  $1.5 < |\eta| < 2.5$ . The CPS and FPS are both made from triangular strips of scintillator with inserted wavelength shifting fibres as shown in Figure 2.7

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 fibres and the waveguides are unique to the preshower detectors as are the gains and biases for the electronics.

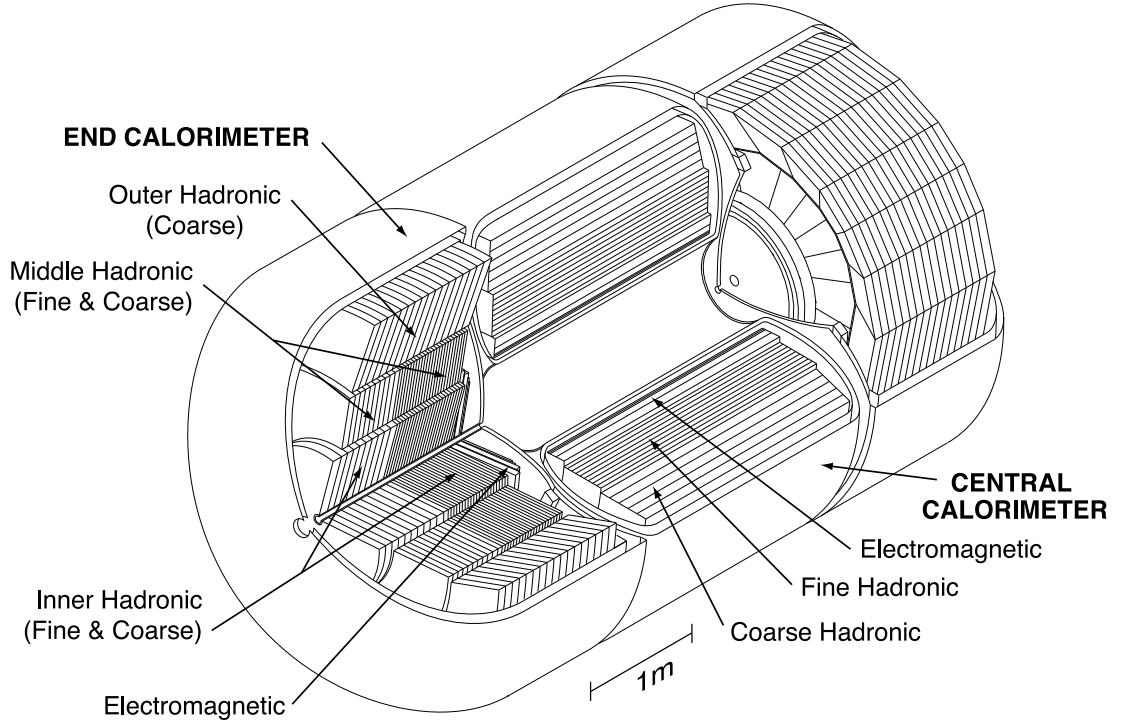


Figure 2.8: Overview of the central and end calorimeters. [33]

### 2.2.3 DØ Calorimeter

The DØ calorimeter system remains unchanged from Run I, and consists of three sampling calorimeters (primarily uranium/liquid-argon). It provides the energy measurements for different objects, electrons, photons, muons and jets. There are three main parts of the DØ calorimeter system as shown in Figure 2.8.

The central calorimeter (CC) covers the pseudorapidities up to  $|\eta| \simeq 1.1$ , and the 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. Within the calorimeter layers, an electromagnetic section is situated closest to the interaction point, and followed by the fine and coarse hadronic sections. The liquid argon sampling material is used as the active medium for the ionization to occur. The choices of the material or thickness of 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 finally, relatively thick (46.5 mm) plates of copper (CC) or stainless steel (EC) for the coarse hadronic modules.

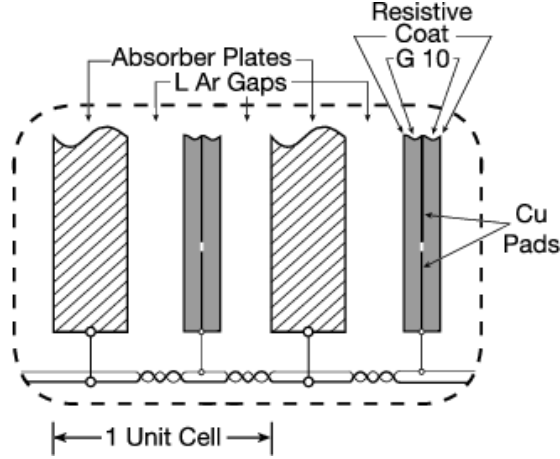


Figure 2.9: Layout of a unit cell for the calorimeter. [33]

The basic unit of the calorimeter cell is shown in Figure 2.9 which consists of plate absorber material, a 2.3 mm liquid Argon gap and a signal board.

The metal absorber is grounded and the connected resistive surfaces are kept at positive high voltage of 2000 V to establish the electric field; the typical electron drift time across the liquid argon gap is about 450 ns. Because the centres of the shower cells lie on the rays projecting from the centre of the interaction region, the calorimeter readout cells are arranged in pseudo-projective towers as shown in Figure 2.10.

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  $\eta$  and  $\phi$  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 Section 3.2. 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 resolution of the EM calorimeters is  $\simeq 5 - 7\%$  measured for electrons having energy above 20 GeV, and the energy resolution is approximately 30% for the hadronic calorimeters measured for jets.

#### 2.2.4 Intercrystat Detectors

The intercryostat detector (ICD) covers the region  $1.1 < |\eta| < 1.4$  in order to compensate for the incomplete coverage arising from the three separate cryostates

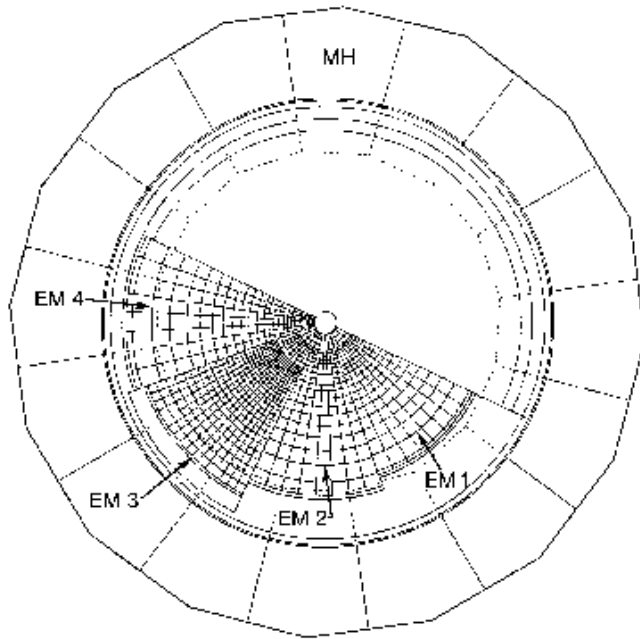


Figure 2.10: Schematic view of the EC electromagnetic readout cells. EM1 indicates EM layer 1 which is the layer nearest to the DØ interaction region, and MH is the EC middle hadronic cell.

of the calorimeter system described in Section 2.2.3. The ICD acts as a scintillator sampler and is mounted to the exterior surfaces of the end cryostats. It is made of a series of scintillating tiles, which are subdivided into twelve straight-edged trapezoidal modules each extending a region of  $\Delta\phi \times \Delta\eta \simeq 0.1 \times 0.1$  rad. A scintillator detector absorbs energy of charged particles and re-emits the energy as photons. The photons are detected by photomultiplier tubes (PMTs) after being shifted in wavelength. As 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.

### 2.2.5 Muon System

Most particles deposit their energy within the calorimeter, except neutrinos and muons that do not interact via strong forces. The general design of the muon system is to provide efficient muon triggering and measurement but low background to detect muons after travelling through the calorimeter. Figure 2.11 describes the upgraded DØ Run II muon system, and the four main components are toroidal magnets, the original central muon system proportional drift tubes

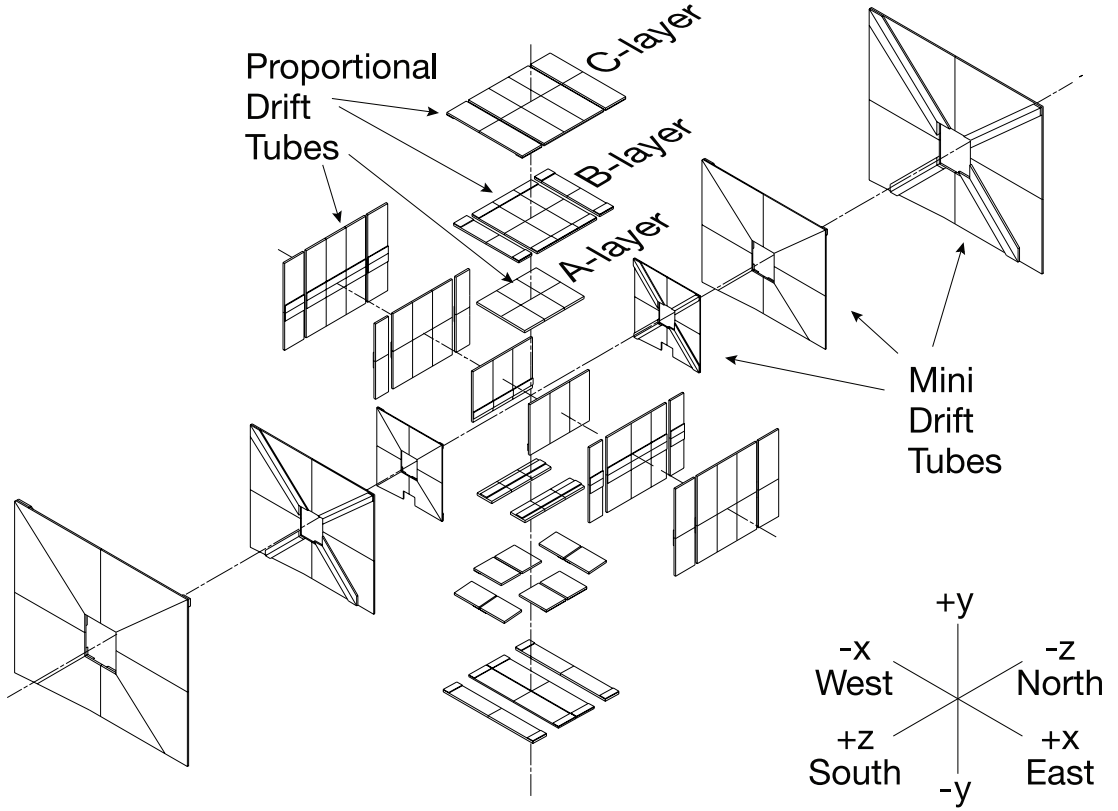


Figure 2.11: Expanded view of the muon drift tubes in the A, B and C layers of the muon system

(PDTs), central scintillation counter, and a complete new forward muon system uses mini drift tubes (MDTs).

The central toroidal magnets provides a stand-alone muon-system momentum measurement. This can not only help rejects  $\pi/K$  decays, but also allows a low- $p_T$  cutoff in the Level 1 muon trigger, cleaner matching with central detector tracks and the improvement on the momentum resolution of high momentum muons. During Run II, the magnets are operated in series and at a current of 1500 A.

The central muon system is instrumented with a toroidal magnet, drift chambers (PDTs), the cosmic cap and bottom scintillation counters, and the  $A\phi$  scintillation counters which cover the A-layer PDTs in the azimuthal angle  $\phi$ . It 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 outside. The layers provide the 55% coverage of the central region by all three layers, and 90% by any of the two. The electron drift time and the charge deposition on the inner and outer vernier pads are recorded by each PDT hit, and the information is used

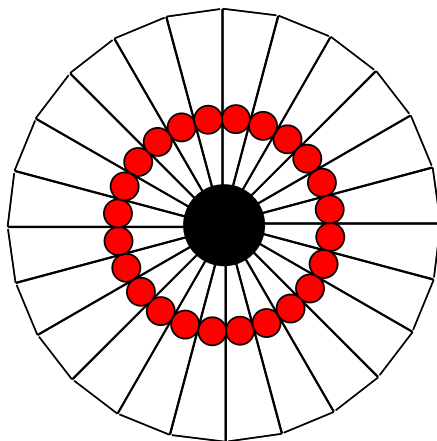


Figure 2.12: The layout geometry of the LM counters and the locations of the PMTs (shown as solid dots).

to determine the hit position. The fast timing signal provided by the cosmic cap and the bottom counters located outside the PDFs layers can help identify a muon from the cosmic ray background. The main difference between the cosmic cap and the bottom counters is that the orientation of the bottom counters match better in  $\phi$  with the central fibre tracker trigger.

The forward muon system extends the coverage to  $\eta \simeq 2.0$  and consists scintillator counters and three layers of MDTs. The closest layer to the interaction point is called A-layer and instrumented by four plans of tubes mounted along magnetic field lines, where C is furthest away with three plans of tubes as B-layer. The muon trigger scintillation counters are also mounted on all three layers, and each layer contains about 96 counters used for precision measurement of the timing of the arrival muons. The  $\phi$  segmentation is approximately  $4.5^\circ$  as with the central system and matches the CFT trigger sectors.

The region at the bottom of the A layer is not covered but occupied by the calorimeter support structure.

### 2.2.6 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 at  $z = \pm 140$  cm. Figure 2.12 shows the schematic drawing for the LM counters and the locations of the PMTs.

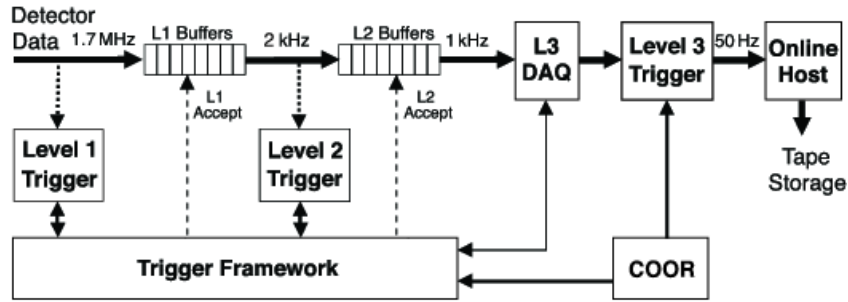


Figure 2.13: A diagram of the design of the DØ trigger system.

The arrays are situated at either side of the beam pipe in front of the end calorimeter to detect the low momenta particles from soft collisions. The length of the counters is 15 cm, and the covering range is  $2.4 < |\eta| < 4.4$ . The measurement of the average number of inelastic collisions determines the luminosity,

$$L = \frac{f N_{LM}}{\sigma_{LM}}, \quad (2.4)$$

where  $f$  is the beam crossing frequency and  $\sigma_{LM}$  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 make measurements of the beam halo rates and the  $z$  coordinate of the interaction vertex.

### 2.2.7 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 higher level given more sophisticated examination to select fewer but interesting enough events for further investigation, and an overview of the system is shown in Figure 2.13.

The global coordination of the DØ triggering is achieved via the COOR package on the online host system.

The first stage, Level 1 (L1), collects hardware trigger elements and provides fast decisions with a trigger accept rate of 2 kHz. The L1 trigger system consists the calorimeter trigger (L1Cal), the central track trigger (L1CTT), the muon system trigger (L1Muon), and the L1 forward proton detector trigger (L1FPD). Each of the each L1 trigger device processes the input from the corresponding

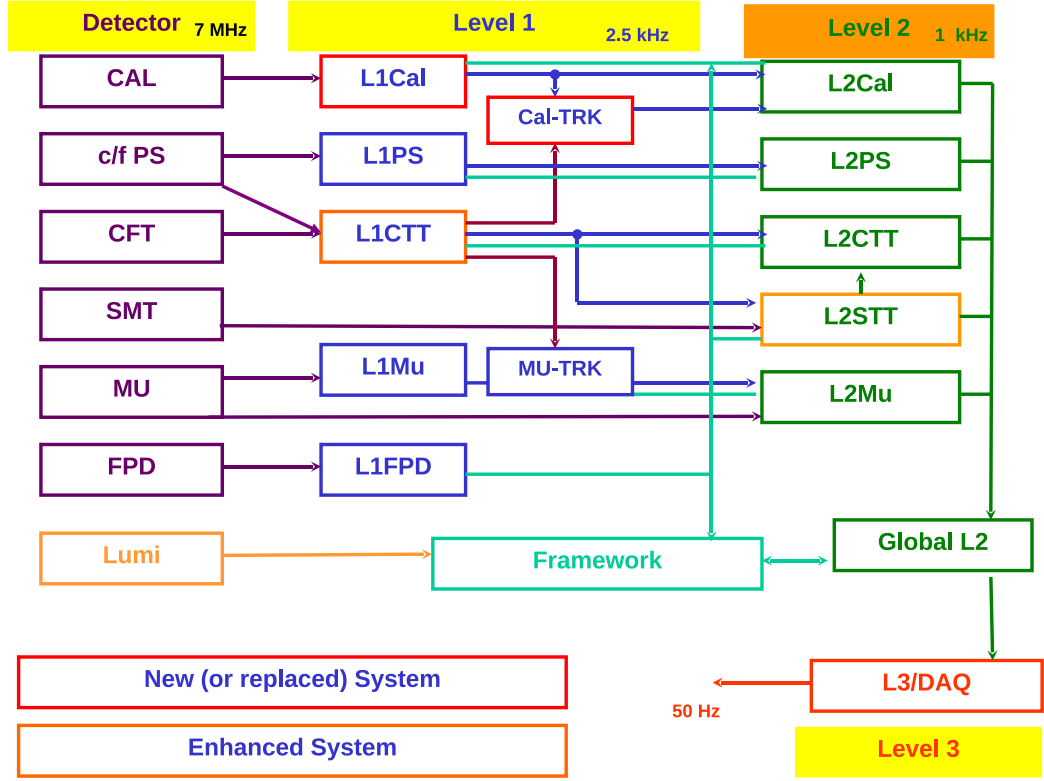


Figure 2.14: A block diagram for the DØ trigger system. [36]

sub-detector, and the digital information of the each passing event is collected the by L1 framework (L1FW) to make a global decision whether a event should be accepted for further examination. The L1 trigger decision from the subsystem is required to arrive at the L1FW in  $3.5\mu$  s or less.

The Level 2 (L2) system is designed to construct trigger decision based on individual objects and reduces the trigger rate by a factor of two to approximately 1 kHz. A L2 decision is made in global stage (L2Global), it tests the correlation in physics objects. These objects are reconstructed in the preprocessing by collecting the signature of candidates passed from the front-ends and L1 system. Each preprocessor corresponds to a detector subsystem, and preprocessing is made either with serial CPU-based cards or with CPU cards plus highly parallelized DSP or programmable logic-based cards.

Figure 2.14 is a block diagram shows how each level of DØ trigger system works.

The events passed by L1 and L2 are fully digitized and sent to Level 3 (L3)



farm. FIFO storage provided by L1 and L2 buffers can hold event data awaiting a L2 decision or transferring to L3, which reduces the experiment's dead time.

L3 trigger system is made by programmable software with sophisticated 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 such as the instantaneous luminosity, the trigger list used at DØ is changed over time. The rejection factor of L3 is large of around 20 to reduce the input rate from 1 kHz to 50 Hz. This is also the final rate of data to be recorded on tape for offline analysis.

# Chapter 3

## Object Reconstruction

### 3.1 Tracks and Vertex Reconstruction

As charged particles pass through the central tracking system and generate hits in its different layers, tracks are formed by comparison of the hit pattern. 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 can also provide for a good momentum measurement for charged particles. This is very useful for the reconstruction of charged particles such as  $\tau$  leptons.

The primary vertex (PV) is defined as the interaction point of the  $p\bar{p}$  collision, and therefore, is defined as the originating point of tracks. At DØ, multiple reconstructed primary vertices are allowed to be found in a single event. However, to separate the primary vertex from other vertices, a vertexing algorithm [37] is developed, which relies on clustering tracks with fixed  $p_T$  and additional requirements imposed on hits within the SMT.

### 3.2 Electron Identification

Electrons interact electromagnetically so that they are expected to deposit most of the energy in the EM calorimeter. Therefore, the reconstruction and identification of electrons are based on tracks from central tracking system with a cluster of energy in the EM calorimeter [38].

The reconstruction of EM cluster uses *CelNN* algorithm [40], which reconstructs particle showers in complex events. The cone size for the EM clustering is

defined with a radius  $R_{corr} = 0.2$ . CelNN algorithm begins by clustering nearby calorimeter cells in  $\eta - \phi$  space within the cone and form “Floor-Cluster” in different layers. Figure 3.1 shows the definition of “neighbor” cells in electromagnetic and hadronic layers. Cells are only considered to be “neighbors” in EM if they share a face in order to achieve the best angular separation. Because the granularity in the third EM layer (EM3) is the best 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, but 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 determine the direction of the electron momentum vector. The energy of an electron is measured using all EM and Fine hadronic layers.

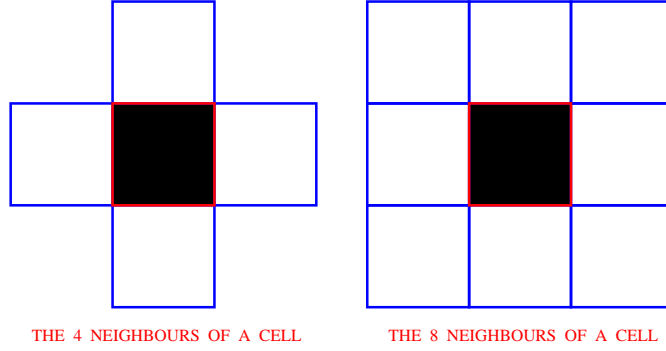


Figure 3.1: Neighboring definitions for EM cells (left) and for HAD cells (right). [40]

The shower shapes of electrons or photons are expected to be narrow, and this is very important in the reconstruction algorithm to distinguish electrons from wider hadron jets which are expected to have significant deposits outside the signal cone. A track matching is applied to the reconstructed electrons to reject photons which do not carry charges. The major variables used for electron ID are:

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

$$\text{Iso} = \frac{E_{tot}(R < 0.4) - E_{EM}(R < 0.2)}{E_{EM}(R < 0.2)}, \quad (3.1)$$

where  $E_{EM}(R < 0.2)$  is the energy deposited in the defined electron cone,  $R < 0.4$

defines the EM cluster isolation cone, therefore  $E_{tot}(R < 0.4)$  is the total energy deposited in the isolation cone.

**EMfr:** Ratio of the EM cluster energy deposited in the EM calorimeter to the total energy including clusters in the Coarse hadronic layers.

**IsoHC4:** 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.

**HMx7:** Characterizes lateral and longitudinal shapes of the EM showers

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

$\chi^2$ : Probabilities of spatial track match to EM clusters.

$\mathbf{p_T}$ : Transverse momentum of electrons.

$\mathbf{p_T^{trk}}$ : Transverse momentum of tracks associated to electrons.

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

There are multiple *qualities* defined for electrons, while the looser quality always includes the tighter set. Table 3.1 listed the cuts for *top\_tight* and *top\_loose* electrons used in this analysis. The *top\_loose* quality is applied in the sample skimming level to select a looser sample which allows the studies of finer tuning of energy and efficiency of the background. The *top\_tight* quality is required for the events passing final selections (see Chapter 7).

| Variables                         | <i>top_loose</i> | <i>top_tight</i> |
|-----------------------------------|------------------|------------------|
| <b>Iso</b> $\leq$                 | 0.15             | 0.15             |
| <b>EMfr</b> $\geq$                | 0.9              | 0.9              |
| <b>HMx7</b> $\leq$                | 50               | 50               |
| $\mathbf{p_T^{trk}}$ (GeV) $\geq$ | 5                | 5                |
| $\mathbf{p_T}$ (GeV) $\geq$       | 15               | 15               |
| $\chi^2 <$                        | 0.0              | 0.0              |
| <b>Lhood8</b> $\geq$              | –                | 0.85             |
| $\mathbf{E/P_{trk}}$ $\leq$       | 2.5              | 2.5              |

Table 3.1: Selection cuts for electron definitions *top\_loose* and *top\_tight*.

The electron efficiency in Run IIa data has been studied in  $Z \rightarrow e^+e^-$  events selected by *tag and probe* method and will be described in details in Section 5.2.

### 3.3 Jet Identification and Missing Transverse Energy

Hadronic decayed particles deposit energy as clusters in the calorimeter which travel roughly in the same direction as hadrons 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 [41]. The energy loss of the reconstructed jets due to the insufficient energy collection using the cone algorithm in the detector is corrected by the Jet Energy Scale (JES) [42] based on measurements of photon plus jet data. The corrected jet energy,  $E_{jet}^{corr}$ , is given by:

$$E_{jet}^{corr} = \frac{E_{jet}^{meas} - E_O}{R_{jet}S_{jet}} \quad (3.2)$$

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

The Missing Energy ( $\cancel{E}$ ) is defined as energy of particles not captured by the detector. This includes the energy of the neutrinos which do not interact with any of the detector elements, and the energy loss from the detectable particles. Because the boost in the  $z$  direction is unknown at the Tevatron, the calculation of missing energy can only be done in the transverse plane in DØ. The reconstruction of the missing transverse energy ( $\cancel{E}_T$ ) uses a collection of the cell energy of other well reconstructed objects in the  $x - y$  plane, and it is calculated in the event reconstruction stage using the reconstructed raw cell energy. This calculation assumes the conservation of momentum, and the momentum imbalance results in  $\cancel{E}_T$ . Because the raw energy simulated by the MC is usually not enough precise,  $\cancel{E}_T$  is re-computed at the analysis stage once the energy from

other objects have been corrected properly.

### 3.4 Muon Identification

Muons deposit only a small amount of ionization energy in the calorimeter, therefore, they travel through most of the inner detector components and are detected by the muon system, typically the most outer layers of a detector at a collider. The identification of muons at the DØ detector [43] relies on the information provided by three independent subsystems:

- three-layer muon detector system,
- central tracking system,
- *MIP* signature in the calorimeter which relies on the “Muon Tracking in the Calorimeter (MTC)” algorithm which provides a typical efficiency of  $\approx 50\%$ .

*Local muons* are muons detected by the muon detector with a coverage of  $\eta < 2.0$ . If central tracks are found to match the local muon in  $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$ , the muon is defined as a “central track-matched muon”, and the  $n_{seg}$  variable is introduced to categorize the track-match muon types. A negative  $n_{seg}$  value is assigned for muons with no associate track-match. 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 Section 2.2.5) muons will be assigned  $|n_{seg}| = 3$ , however, the requirement of BC layer hit is dropped for central  $|n_{seg}| = 3$  muons. 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 misidentified muons. In addition to the muon transverse momentum,  $p_T$ , measured in the muon detector, the central tracker can provide accurate momentum resolution for the central track-matched muons.

A separate variable, *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. In this analysis, a **loose**  $n_{seg} = 3$  muon is selected in the sample-skimming (preselection) level, and **medium**  $n_{seg}$  quality is required for the muons passing the final selections. The definition for **medium**  $n_{seg} = 3$  is given below [43]:

- 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 this requirement is dropped).

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

- one of the *medium*  $n_{seg} = 3$  tests to fail,
- the A wire and scintillator are treated as one test as requiring always at least one scintillator hit.

Moreover, in order to control the quality of central tracks match muons, three quality definitions **trackloose**, **tracknewmedium**, **trackmedium** and **tracktight** are defined and rely on [43]:

- number of hits in SMT or CFT system,
- $\chi^2$  per degrees of freedom of the central track fit ( $\chi^2/\text{dof}$ ),
- distance of closest approach with respect to the beam spot location (dca) in the x-y plane.

**tracknewmedium** quality is required for muon candidates in this analysis; the definitions of track qualities are listed below [43]:

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

**trackmedium:** *trackloose* tracks with  $\chi^2/\text{dof} < 4$  in addition.

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

**tracktight:** *trackmedium* tracks with at least 1 SMT hits.

Muon *isolation* variables are designed to separate signal muons from heavy flavor backgrounds, such as muons produced in heavy flavor quark decays  $b, c \rightarrow \mu + X$  which topologically tend to be embedded inside a jet. Isolation qualities rely on the tracks near the muon track or on the calorimeter clusters around the muon momentum vector. The relevant isolation quality for the analysis is *NPTight* which defined by the variables:

The isolation is with respect to other tracks and energy in the calorimeter. The sum of transverse energies of the cells, measured with respect to the beam direction, in an annulus around the muon direction. For *NPTight* quality, it is required to be

$$\mathcal{I}_{\text{cal}} = \sum_{\text{cells}, i} E_T^i < 2.5 \text{ GeV} \quad \text{for } 0.1 < R < 0.4, \quad (3.3)$$

where  $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$  is the distance in azimuth  $\phi$  and pseudorapidity  $\eta$  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

$$\mathcal{I}_{\text{trk}} = \sum_{\text{tracks}, i} p_T^i < 2.5 \text{ GeV} \quad \text{for } 0.0 < R < 0.5. \quad (3.4)$$

Further details on such requirements will be described in Chapter 7.

### 3.4.1 Hadronic Tau Identification

The  $\tau$  lepton is the heaviest lepton with a mass of 1.78 GeV. With a lifetime of  $2.9 \times 10^{-13}$ s,  $\tau$ s are expected to decay after a very short distance before reaching any element of the DØ detector.

Figure 3.2 shows example Feynman diagrams for a  $\tau$  decaying to a muon and a pion, and Table 3.2 lists the most significant (branching ratio  $> 2\%$ )  $\tau$  decays.

As listed in Table 3.2, the total branching ratio of leptonic and hadronic  $\tau$  decay channels is about 35% and 65%, respectively. Electron and muon identification methods as described in Sections 3.2 and 3.4 are used to detect the leptonic  $\tau$  decay products. For hadronic  $\tau$  decays, one needs dedicate  $\tau$  identification tools to reconstructed the  $\tau$  leptons, therefore, it is very crucial to the search presented



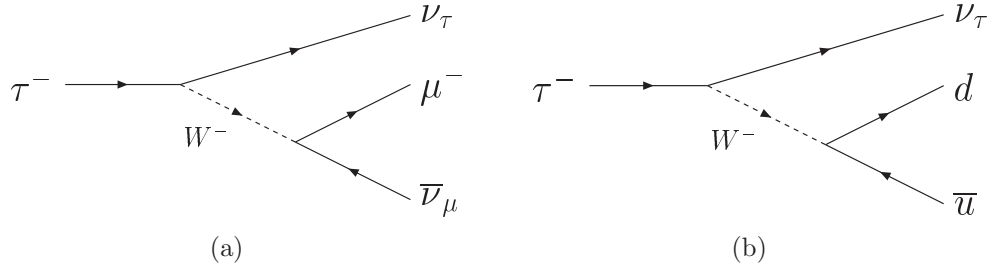


Figure 3.2: Feynman diagram for a) a  $\tau$  decaying to a muon and b) a  $\tau$  decaying into a pion.

| 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 | $\pi(K) + \geq 1\pi^0 + \nu_\tau$    | 1-prong (48.7%)  | 38.98  |
| Hadronic type 3 | $\pi\pi\pi + \geq 0\pi^0 + \nu_\tau$ | Hadronic 3-prong | 14.23  |

Table 3.2: The products in the final state of main  $\tau$  decay modes and the branching ratios (BRs).

in this thesis. The products of hadronic  $\tau$  decays are mostly neutral and charged pions so that they mostly appear to be narrow jets with low track multiplicities. The hadronic  $\tau$  reconstruction [?], [45] rely on the following elements:

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

**EM subcluster:** Reconstructed by the *CelNN* algorithm in the EM3 layer of the calorimeter(see Section 2.2.3 and See Section 3.2). Subsequently, the EM cells in other layers and preshower hits are attached if such cluster is found. These are useful for finding neutral  $\tau$  decay products such as those from  $\pi^0$  particles or  $\gamma$  radiation which deposit significant energy in the EM layers.

**Associate tracks:** Any  $\tau$  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  $R = 0.3$  centered on the calorimeter cluster. In addition, the associate 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 ( $\Delta z(second or third track, leading track) \leq 2$  cm). A second track will be associ-

ated 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,

The final states of  $\tau$  candidates can be split into three types categorized 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  $\tau^\pm \rightarrow \pi^\pm \nu$  ( $\pi\nu$ -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  $\tau^\pm \rightarrow \pi^\pm \pi^0 \nu$  ( $\rho\nu$ -like).

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

In particular, types 1 and 2 belong to the “1-prong”  $\tau$  decay mode while type 3 is considered as “3-prong”. The classification is based on only the detector signature, therefore, no direct relationship to the physics channels, and the correlations listed above can be broken. Take the decay  $\tau^\pm \rightarrow \pi^\pm \nu$  for example, this  $\tau$  lepton decay should be reconstructed as  $\tau$ -type 1 based on the signature of one track and one calorimeter without EM sections. However, if the charged pion products shower early in the calorimeter which leave associated EM clusters, this decay will be reconstructed as  $\tau$ -type 2. if the charged pion product showers early in the calorimeter. Also, a  $\tau^\pm \rightarrow \rho^\pm \nu \rightarrow \pi^\pm \pi^0 \nu$  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, the EM energy can be lost if the  $\tau$  lays in the ICD region.

The key to  $\tau$  identification at DØ is the use of a multivariate technique, , where Neural Network becomes the default choice for final  $\tau$  identification [46]. Three separate Neural Networks ( $NN_\tau$ ) are built, one for each  $\tau$  type, to separate  $\tau$  decays from the misidentified jets. The  $NN_\tau$ s are trained using  $Z \rightarrow \tau^+ \tau^-$  PYTHIA MC as signal and jets that recoil against non-isolated muons taken directly from data as background. Two  $NN_\tau$ s are trained separately for type 1 candidates

for  $\tau$ s detected in the ICD or non-ICD region. A set of discriminating variables between taus and jets are chosen as the inputs to  $NN_\tau$  training to give the best separation between signal and background. These variables are constructed from isolation within the calorimeter and tracking system, shower shape parameters and correlations built between the calorimeter and track.

The output of  $NN_\tau$  is a single value between zero and one, while value one is assigned for pure  $\tau$  signals while zero is for jet backgrounds. The  $NN_\tau$  applied currently for tau candidate selection was trained and tested in DØ Run IIa data. There are different  $NN_\tau$ s trained and tested respect to Run IIa and Run IIb data sets, and the corresponding one should be applied to the physics analysis. Most of the  $\tau$  leptons reconstructed at DØ fell in the energy region of  $20 < E_T^\tau < 40 GeV$ . The signal acceptance of  $NN_\tau$  for those taus with  $|\eta^\tau| < 2.0$  is about 66% from  $Z \rightarrow \tau^+\tau^-$  MC events, while the rejection rate of jet background is kept as high as 98%. The correction to compensate the different efficiencies in data and MC while applying a  $NN_\tau$  selection cut is required for Run IIb data set. More details about the correction can be found in Section 5.4.

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

In order to control the quality of the  $\tau$  tracks, three qualities **tracktight\_tau**, **trackmedium\_tau** and **trackloose\_tau** are defined using the same definitions as those for muon track (see Section 3.4). The  $\tau$  track qualities rely on number of hits,  $\chi^2$  per degrees of freedom of the central track fit ( $\chi^2/dof$ ), and distance of closest approach with respect to the beam spot location (dca) in the  $r$ - $\phi$  plane.

The main requirements for *trackloose\_tau* are:

- $dca < 0.2cm$  if track has no SMT hit,
- $dca < 0.04cm$  if track has SMT hits.

Tracks with quality **trackmedium\_tau** must fulfill **trackloose\_tau** selections, and in addition the  $\chi^2/dof < 4$  requirement. A **tracktight\_tau** track must pass those cuts of a **trackmedium\_tau** track and also contain at least one SMT hit.

# Chapter 4

## Triggers and Data Samples

### 4.1 Trigger and Efficiency Measurements

The DØ trigger system consists of three trigger levels (L1, L2, L3) (see Sect. 2.2.7) helps to quickly select and record events with a special interest to physics analysis. The relevant triggers used in the inclusive search for a neutral Higgs boson decaying to two tau leptons are single muon triggers for  $\tau_\mu\tau_h$  channel and single electron triggers for  $\tau_e\tau_h$  channel. These triggers select possible 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 possible for effects of the trigger to 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 [49] 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. 4.1. The method starts by selecting a dilepton sample from the process  $p\bar{p} \rightarrow Z \rightarrow l^+l^-$ , where  $l = e, \mu$  and requires two leptons. One of the selected leptons, known as the “tag” lepton, has to match all the trigger terms at each relevant trigger levels. The other lepton per event is the “probe” lepton. Assuming  $n_p$  is the number of events passing the requirement that the probe has been matched to all relevant trigger terms at each trigger level, the resulting efficiency,  $\epsilon$ , is calculated by

$$\epsilon = \frac{2n_p}{n_T + n_p}, \quad (4.1)$$

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 parameterized in terms of the most sensitive variables (e.g.  $p_T$  or detector pseudorapidity).

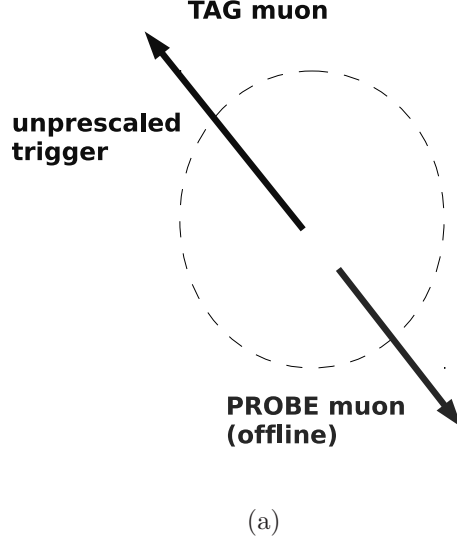


Figure 4.1: Schematic view of the tag-and-probe method for muons. Here, a reconstructed probe muon is found to match an tagged muon triggered by the muon triggers ???. Similar diagrams exist for cases with other leptons.

It is necessary to split the efficiencies into different trigger epochs (periods) due to run configurations, changing luminosities or hardware or software changes. The muon triggers are discussed in greater details in Sect. 4.2; Sect. 4.3 describes the single electron triggers.

## 4.2 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 4.1. The muon trigger terms in different levels will be discussed separately in the following subsections.

Run IIa data events are required to pass one of the muon triggers used in the periods listed in Table 4.2. In addition, a matching is done for the data events which triggered by several trigger, and this criteria removes about 4% of the data events (see Ref.[51] for complete details). Similar trigger requirements are applied to Run IIb data events, and the selected data sample contains only events from a

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

Table 4.1: Muon trigger terms used in different trigger level (L1, L2, L3).

logical ‘OR’ of the 19 single muon triggers, and data events are required to pass one of these muon triggers used in the periods listed in Table 4.2.

|         | 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 4.2: 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  $Z \rightarrow \mu^+\mu^-$  tag-and-probe events, and without a requirement on the invariant mass of this two muon sample [50]. The following selections are required for the tag muon in the measuring sample:

- loose muon quality (see Sect. 3.4) ;
- A-layer scintillator  $|\text{time}| < 7$  ns (use timing from B-layer if no A-scintillator hit is found);
- matched to a central track with medium quality;
- muon  $p_T > 30$  GeV;
- isolated using the selections  $\mathcal{I}_{\text{trk}} < 3.5$  GeV and  $\mathcal{I}_{\text{cal}} < 2.5$  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. 3.4. The probe muon must have a 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 parameterized in terms of  $\eta$  and  $\phi$ . The dependency on luminosity has been introduced in Run IIb, and the efficiency corrections for MC are therefore parameterized in terms of  $\eta$ ,  $\phi$  and luminosity.

#### 4.2.1 Level 1 Muon Trigger (L1MU)

Level 1 muon object is formed by tracks from the central fiber tracker (CFT), and scintillators 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 carried 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 heavy flavors and cosmic rays; this is necessary because the bunch crossing time is smaller than the proportional drift time (600ns).

At this level, two regions are defined in terms of detector pseudorapidity ( $\eta$ ) of the trigger system: the “wide” and “all” regions. The “all” region is defined as  $|\eta| < 2.0$ , whereas the definition for “wide” region changes from  $|\eta| < 1.5$  to  $|\eta| < 1.6$  for the triggers used post 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. *Prescaler* at level 1 is used to filter events and allow only a electronically acceptable amount 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 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 L1 loose wire term, relative to the L1 tight scintillator term in this sample is 95%.

### 4.2.2 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 transverse momentum  $p_T$  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  $p_T$  requirement 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.

### 4.2.3 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. It also includes 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 programme such as C++ subroutines 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.

## 4.3 Run IIa Electron Triggers

In the  $\tau_e\tau_h$  channel, tight electrons selected from data are required to pass least one of the unscaled electron triggers in the logical “OR”, at all three levels of the trigger. To select events fired by a particular single electron trigger, the reconstructed electron must match a Level 1 trigger tower above a given transverse momentum 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 cut in addition [52]. Table 4.3 lists the run dates for each trigger epoch.

Single electron efficiencies of Run IIa data are measured using a tag-and-probe method  $Z \rightarrow e^+e^-$  sample along with tracking information introduced for



| 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 4.3: List of run dates of electron triggers for each trigger epoch.

electron triggers in order to gain improvements at low  $p_T$  ( $15 \text{ GeV} < p_T < 25 \text{ GeV}$ ) compare to the electron trigger efficiency based on calorimeter solely information [53]. 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 **EMfr** > **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<sub>T</sub>** > **15 GeV**.
- Transverse momentum of associated tracks **p<sub>T</sub><sup>trk</sup>** > **10 GeV**.
- Match a single-electron trigger tower within  **$\Delta R$**  < **0.4**.

The definitions of these variables are given in Sect. 3.2. The probe electron is required to have  $p_T > 15 \text{ GeV}$  and must pass either loose or tight selections depending on the applied electron ID. A mass window of  $70 < m_Z < 110 \text{ GeV}$  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 preselection stage varies from 74% for a Higgs boson of 90 GeV mass to 87% for a Higgs boson mass of 200 GeV.

## 4.4 Data Samples

The analysis presented in this thesis is based up to the Run IIb-2 dataset 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 [48] for the  $\tau_\mu + \tau_h$  and  $\tau_e + \tau_h$  channel, respectively.

#### 4.4.1 MUinclusive Skims

The MUinclusive skim accepts events that fulfill 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. 3.4 with a transverse momentum  $p_T > 8$  GeV.
- At least one “loose” muon with  $p_T > 5$  GeV and at least two tracks with  $p_T > 5$  GeV and  $p_T > 8$  GeV, respectively, and one  $\tau$  candidate with  $p_T^\tau > 5$  GeV.

The muon transverse momentum,  $p_T^\mu$ , is measured using the central tracking system whereas the  $\tau$  transverse momentum  $p_T^\tau$  is measured using the calorimeter except for  $\tau$ -type 1. Details of the  $\tau$  reconstruction algorithm can be found in [44].

The Run IIa MUinclusive dataset is divided into five trigger epochs from trigger list v8 to v14.99. The total integrated luminosity for the Run IIa data used for the analysis is  $1.08 \text{ fb}^{-1}$ . The Run IIb MUinclusive dataset is divided into six periods corresponding to the v15 and v16 trigger lists. The total integrated luminosity for the Run IIb data used for the analysis is  $4.28 \text{ fb}^{-1}$ . Table 4.4 provides a breakdown of the integrated luminosity during each period. Therefore, a total of  $5.36 \text{ fb}^{-1}$  of Run IIa plus Run IIb data is used in the  $\tau_\mu + \tau_h$  search channel in this analysis.

#### 4.4.2 EMinclusive skims

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

- At least one reconstructed EM object with  $p_T > 20$  GeV.
- At least one reconstructed EM object with  $p_T > 8$  GeV associated with a track with  $p_T^{trk} > 5$  GeV.

| Trigger Epoch   | Luminosity [ pb <sup>-1</sup> ] |
|-----------------|---------------------------------|
| Run IIa Total   | 1014.25                         |
| v8 - v12        | 72                              |
| v12 - v12.99    | 226.07                          |
| v13 - v13.2     | 31.84                           |
| v13.2 - v13.99  | 327.41                          |
| v14 - v14.6     | 142.1                           |
| v14.6 - v14.99  | 190.64                          |
| Run IIb Total   | 4281.66                         |
| v15.00 - v15.22 | 227.21                          |
| v15.22 - v15.50 | 307.29                          |
| v15.50 - v15.60 | 232.48                          |
| v15.60 - v15.90 | 461.32                          |
| v15.91 - v16.00 | 397.31                          |
| v16.00 - v16.22 | 2656.05                         |

Table 4.4: Description of the recorded integrated luminosity after applying data quality requirements given as a function of trigger list versions.

| Trigger Epoch(s) | Trigger(s) | Luminosity [pb <sup>-1</sup> ] |
|------------------|------------|--------------------------------|
| v8 - v11.34      | EM_MX      | 106.9                          |
| v12 - v12.37     | E1_SH30    | 231.0                          |
| v13 - v13.90     | E1_SH30    | 378.2                          |
| v14 - v14.93     | E1_SHT25   | 333.8                          |
| v8 - v14.93      | All        | 1050                           |

Table 4.5: 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.

The electron and the  $\tau$  transverse momentum, are both measured using the calorimeter. Table 4.5 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<sup>-1</sup>.

# Chapter 5

## Monte Carlo Simulation and Corrections

### 5.1 Monte Carlo Simulation Method

In order to correctly model physics background or any potential signal events in the data, the Monte Carlo (MC) methods are used in high-energy physics analysis such as the one described in this thesis. Figure 5.1 provides a general example flowchart for the event-generation chain, which provides a proper simulation for background events and modelling of signal process. The are two main steps taken in  $D\bar{O}$  to build up MC events are:

- simulate physics processes at either leading order (LO), next-to-LO (NLO), or next-to-NLO (NNLO) using MC generators, PYTHIA 6.3 [54] or ALPGEN 1.2 [55];
- perform a simulation of particle interactions with  $D\bar{O}$  detector components.

During step 1 for generation, a general-purpose generator such as PYTHIA takes into account the following main steps sequentially to build up the structure for 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;

- hadronization process using non-perturbative models;
- simulate decay products except for  $b$  quarks and  $\tau$  leptons.

The simulation of the  $\tau$  lepton decays is done using the TAUOLA [57] package. Here, it provides not only the correct branching ratios for the tau decays, but also accounts for the polarization of the tau 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 of lower energy without explicit matrix element calculations. It begins at the energy scale of the hard process and evolves down to a lower scales until the showering ends at an energy cutoff point. The momentum distribution of partons is described as the parton distribution functions (PDFs) by using CTEQ6.1M [58]. The alternate MC generator used at DØ, ALPGEN, is a parton-level event generator and provides a better modeling of multi-parton hard processes 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.

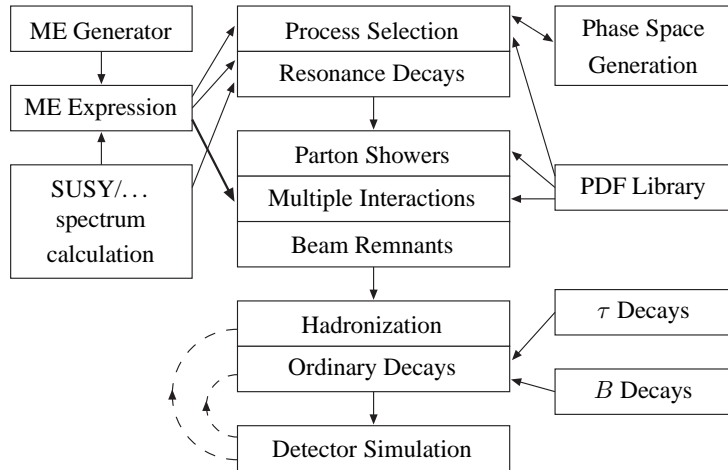


Figure 5.1: An example flowchart for the event generation chain used at hadron collider experiment. ME describes the Matrix Element method while PDFs represents the parton distribution functions. [56]

The simulation of the DØ detector is done by a GEANT3-base software, **D0gstar** [59], 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 in other words. The **D0Sim** software is used along with **D0gstar** to simulate the digitization of the signal outputs from subsequently different detector layers. Zero Biased (ZB) data events from elastic and/or diffractive interactions at low  $p_T$  that 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** for **D0Sim** to pileup [60].

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 run on the production farms. There are two output formats from **d0reco**, Data Summary Tape (DST) and Thumbnail (TMB) [61]. DSTs contain all necessary information to perform physics analysis with 150kb event size, while TMBs are summaries of DST with only 15kb event size and 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 [62]. CAF Trees are onputs object-oriented and ROOT based software framework used in DØ and contain the basic information on the properties and kinematic 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 provided by the VJets Working Group used in this analysis, and the factors are applied via CAF framework to the MC events and tune following items:

**Longitudinal beam position  $z$ :** It is found that data and MC do not match in the distribution of  $z$  coordinate. Therefore, it is re-weighted for MC in terms of instantaneous luminosity and data taking periods [63].

**Instantaneous Luminosity:** The luminosity profile used in the MC does not describe the really data taken in the detector. Therefore, a re-weighting is performed in the MC to make sure that instantaneous luminosity of MC events match data. In Run IIb analysis, this re-weighting is more significant since the data is taken from wider range of instantaneous luminosity values.

**$p_T$  distribution of  $W$  and  $Z$  bosons:** Because the higher order effects is missing in the samples generated from PYTHIA or ALPGEN, the vector boson trans-

verse momentum ( $p_T^{Z/W}$ ) is not represented correctly in MC. The transverse momentum of the  $Z$  bosons from unfolded data agrees with the prediction provided by the ResBos MC generator [64] in 0-30 GeV range, a re-weighting formula in terms of  $p_T^Z$  is derived directly from ResBos MC generator and extrapolated to higher  $p_T^Z$  region [65]. Since there is no direct measurement for  $p_T^W$  available due to the present of neutrinos in the final state, the  $p_T^Z$  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) [66].

There are a few more corrections required, such as trigger efficiency described in Chapter 4 and the efficiency corrections of lepton identification discussed in the following sub-sections.

## 5.2 Electron Efficiency Correction

The reconstruction efficiency is generally overestimated 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 [38] are required for MC. If  $\epsilon_{MC}$  and  $\epsilon_{data}$  are electron efficiencies of data and MC, respectively, the efficiency correction factor  $f_e$  is given by

$$f_e = \frac{\epsilon_{data}}{\epsilon_{MC}}. \quad (5.1)$$

Figures 5.2 and 5.3 show the dependencies of electron efficiency correction factors on  $p_T$  and detector pseudorapidity,  $\eta_{det}$ . From these measurements, the correction factors needed for MC vary for electrons detected in different parts of the detector, but the  $p_T$  dependence is small compare to the dependence on  $\eta_{det}$ . A two-dimensional correction depending on both  $p_T$  and  $\eta$  is applied in the analysis presented in this thesis.

Further, the search selections for the  $\tau_e\tau_{had}$  channel described in this thesis reject events if they contain a second electron passing the tight electron selection (see Sect. 7.1). Assuming  $N_{MC}^{2-tight}$  is the number of MC events with two electrons reconstructed as “tight”, and  $N_{MC}^{1-tight}$  is the number of events with only one electron reconstructed as “tight”. The expected number of MC events which

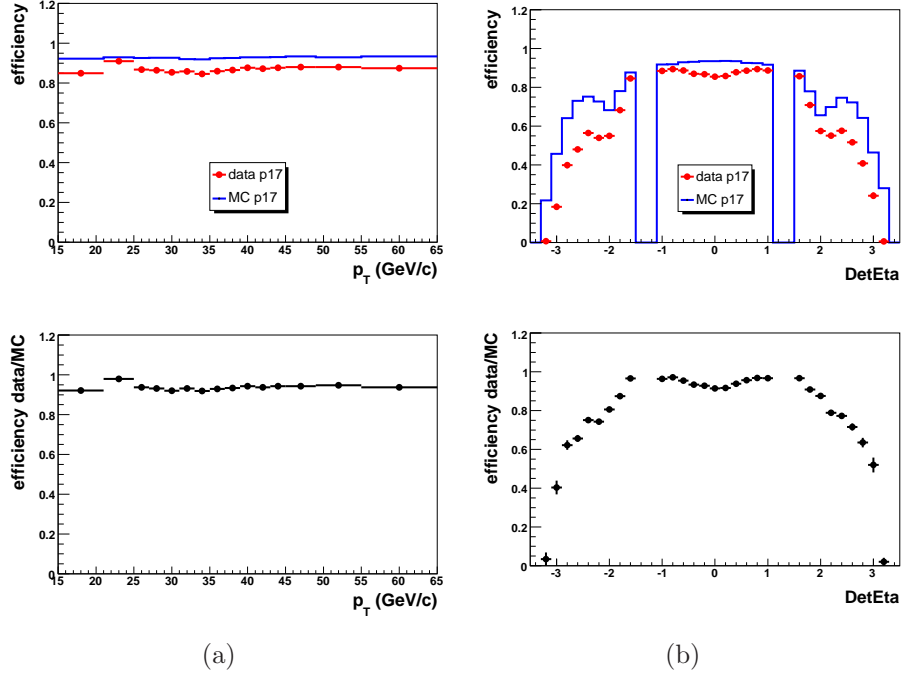


Figure 5.2: Upper figure shows the loose electron efficiencies measured in Run IIa data (red dots) and MC (blue line) with a dependency on (a) $p_T$  (b) $\eta_{det}$ . The correction factors  $f_e$  derived for both cases are shown in the lower distributions.

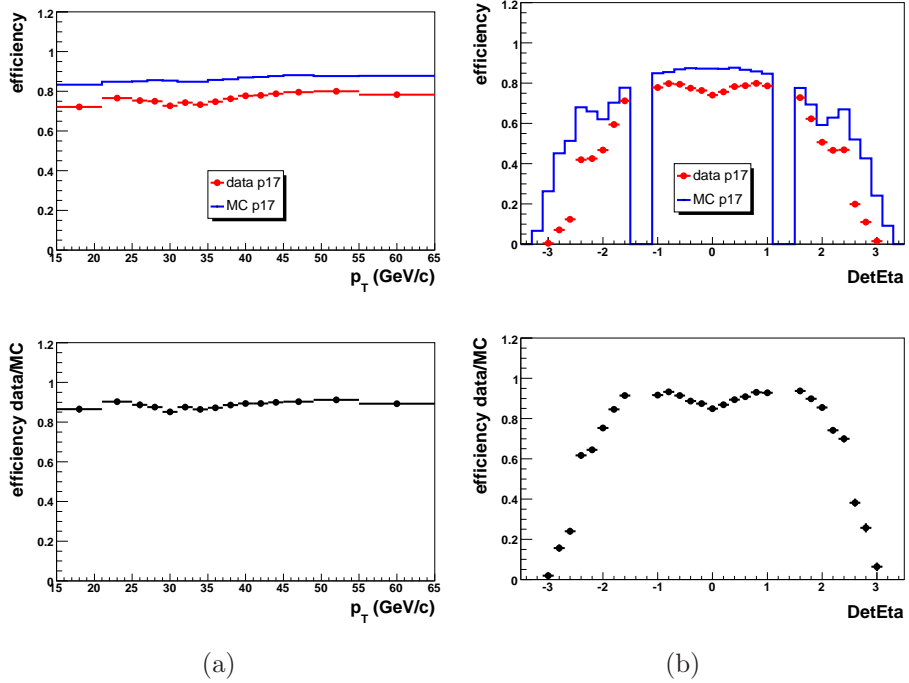


Figure 5.3: Upper figure shows the tight electron efficiencies measured in Run IIa data (red dots) and MC (blue line) with a dependency on (a) $p_T$  (b) $\eta_{det}$ . The correction factors  $f_e$  derived for both cases are shown in the lower distribution.



describes the data with only one tight electron should be given by

$$N_{MC} = f_e N_{MC}^{1-tight} + 2N_{MC}^{2-tight} f_e \left( \frac{1 - \epsilon_{data}}{\epsilon_{MC}} \right). \quad (5.2)$$

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  $2N_{MC}^{2-tight} f_e \left( \frac{1 - \epsilon_{data}}{\epsilon_{MC}} \right)$  is the contribution from “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  $Z \rightarrow e^+e^-$  sample with two electrons reconstructed as tight, the factors shown in Fig. 5.3 do not take into account the inefficiency term. This is readily observed in the invariant mass distributions of electron and  $\tau$  pair taken from  $\tau_e\tau_h$  channel with preselections are shown in Fig. 5.4. These distributions indicate that an additional factor is needed on top of the standard correction as expected. Therefore, electrons in the  $Z \rightarrow e^+e^-$  MC sample should be corrected by a scaling factor,  $k$ , given by

$$k = \frac{f_e N_{MC}^{1-tight} + 2N_{MC}^{2-tight} f_e \left( \frac{1 - \epsilon_{data}}{\epsilon_{MC}} \right)}{f_e N_{MC}^{1-tight}}, \quad (5.3)$$

, and has an estimated value of  $k$  of 1.26.

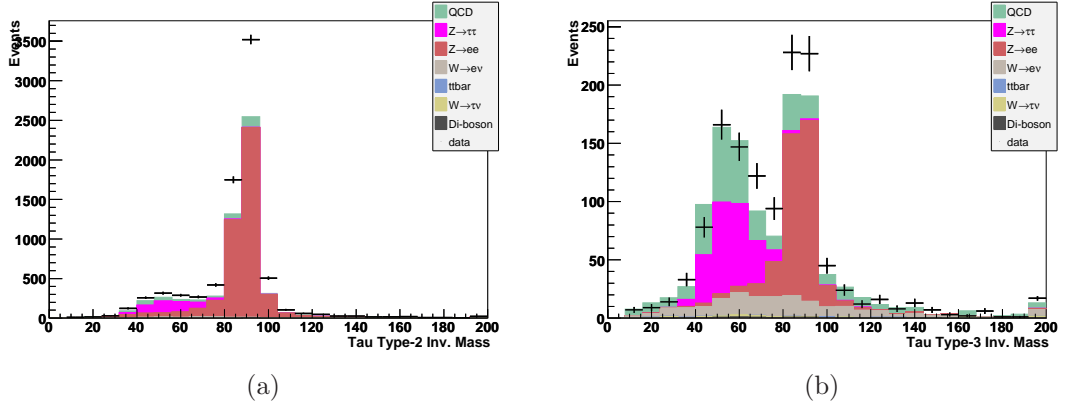


Figure 5.4: The distributions of the invariant mass in data and MC samples of tau (a) type-2 and (b) type-3 with standard efficiency correction factors applied.

In this analysis, the scaling factor  $k$  in Eq. 5.3 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

$$k(X) = \frac{N(X)_{data_{Zee}}}{N(X)_{MC_{Zee}}}, \quad (5.4)$$

while  $N(X)_{data_{Zee}}$  is the total data number after subtracting the predicted background except the  $Z \rightarrow e^+e^-$  MC, and  $N(X)_{MC_{Zee}}$  is the number of real MC electrons. Here, events for estimating  $k(X)$  are selected within a mass window  $60 < M_{ee} < 130$ . This measurement depending on  $NN_e$  provides a good estimation for the number of events expected from “only-one-tight-electron” data sample. After a fitting 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. 3.4.1) and  $M_{e\tau}$  for  $\tau$  type 2 and type 3 events, respectively. The dependence on  $NN_e$  of this correction for  $\tau$ -type 2 also compensates the mis-modeling in this variable for MC. The masses of type 2 and type 3 electron and  $\tau$  pairs after the correction are shown in Fig. 5.5, and the predicted background describes data well.

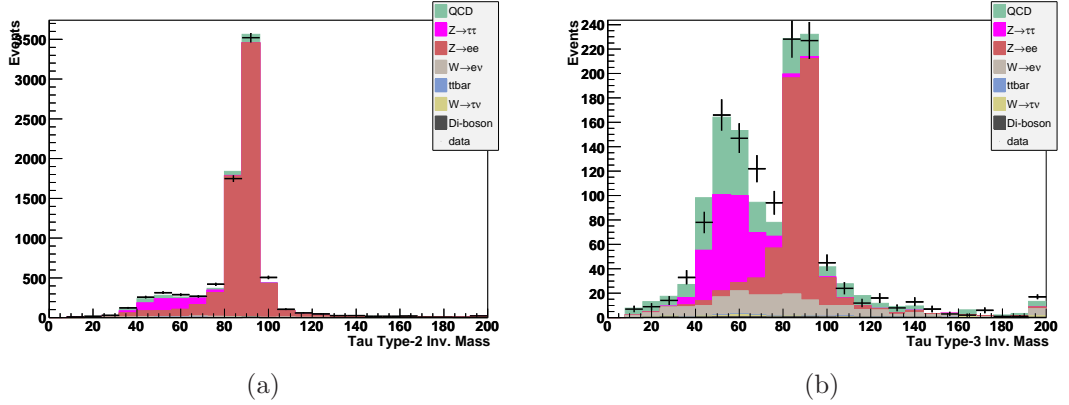


Figure 5.5: The distributions of the invariant mass in data and MC samples of tau (a) type-2 and (b) type-3 after the additional electron correction performed.

### 5.3 Muon Efficiency Correction

The tag-and-probe method is also used to determine the efficiency in the  $Z \rightarrow \mu^+\mu^-$  high  $p_T$  ( $p_T > 30$  for tag muons) sample for the reconstruction of muon objects and central tracks matched to muons. In order to compensate the difference in efficiencies between the data and MC, for example, the tracking efficiency correction, is due to the fact that Monte Carlo 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 [43]:

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

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

**muon isolation efficiency** which is parametrised in terms of  $|\eta_{CFT}|$  and  $\mathcal{L}$  for  $\Delta R(\mu, \text{closest jet}) > 0.5$ , or in  $|\eta_{CFT}|$ ,  $p_T$  and  $\Delta R$  for posterior isolation.

The average muon identification efficiency is 91.2% for *loose* muons and 73.6% for *medium* muons with  $|n_{seg}| = 3$ . The dependence on  $\mathcal{L}$ ,  $\phi$ ,  $\eta$  and  $|\eta_{CFT}|$  for *medium* muons with  $|n_{seg}| = 3$  are shown separately in Fig. 5.6. The efficiencies are 90.8% of associated tracks of *mediumnewtrack* quality and 97.2% of muons with *NPTight* isolation quality.

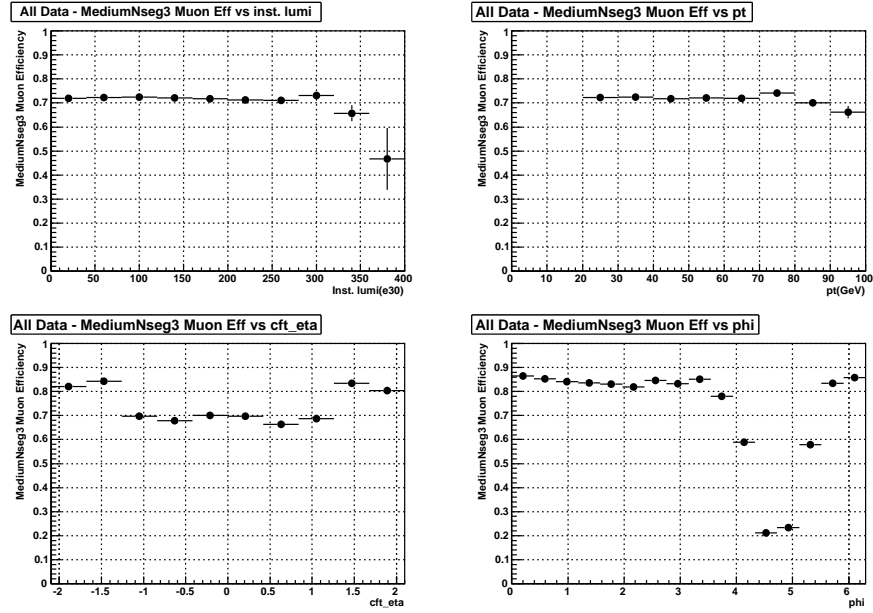


Figure 5.6: Muon reconstruction efficiencies of *medium* quality with  $|n_{seg}| = 3$  versus instantaneous luminosity  $\mathcal{L}$  (top left),  $p_T$  (top right),  $|\eta_{CFT}|$  (bottom left), and muon  $\phi$  (bottom right).

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, instantaneous luminosity dependence, correlations

between different corrections, finite bin size and choices of variables necessary to parameterize 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 [43].

## 5.4 Tau Corrections

### 5.4.1 Tau Efficiency Correction

The efficiency corrections for tracks associated to the hadronic  $\tau$  leptons,  $\tau_h$ , are taken from the measurements for muon tracks as described earlier in Sect. 5.3. The same track quality definition can also be used for tracks associated to  $\tau_h$  objects, and the relevant one used in this analysis is *tau\_trackloose* which have the same requirements as *trackloose* defined for muon tracks. Figure 5.7 shows the efficiency of data and MC having a *tau\_trackloose* requirement on the associated tracks as a function of luminosity. The average correction applied to  $\tau_h$  objects is 91.7% for typical instantaneous luminosities during Run II.

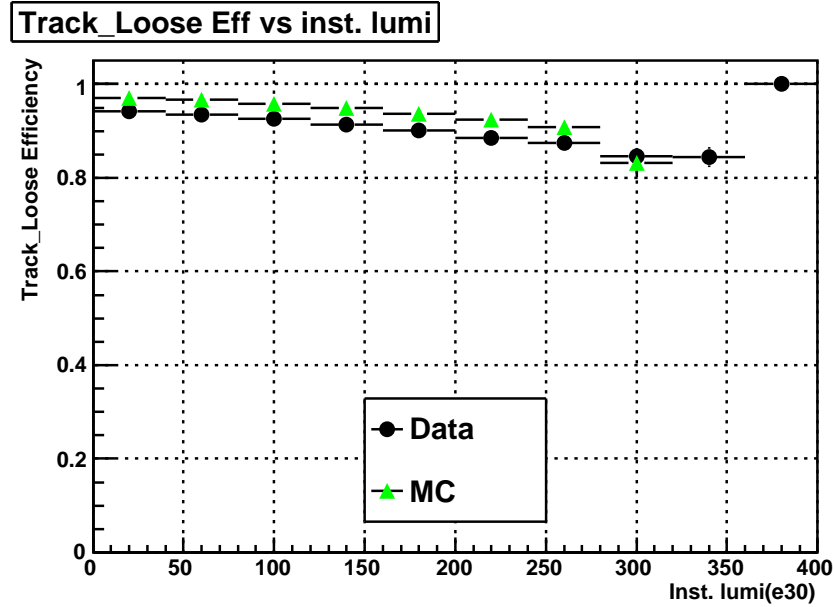


Figure 5.7: The efficiency of applying *trackloose* quality cut to data and MC as a function of instantaneous luminosity.

Because MC simulates the quality of taus than real data, the peak in  $NN_\tau$

distribution of MC is therefore narrower and with improved resolution than of data. This will cause poor modeling of the efficiency for MC while cutting on  $NN_\tau$  to select well reconstructed  $\tau$  leptons. Therefore, a re-normalization of  $NN_\tau$  variable based on a bin-by-bin re-weighting method measured in  $Z \rightarrow \tau^+\tau^-$  sample is done in addition to the track correction. In order to derive this correction, a probability distribution function, PDF, 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  $NN_\tau$  values in a certain, the  $NN_\tau$  PDF for MC is given by

$$P_{MC}(X) = \frac{N_{MC}(X)}{N_{MC}} \quad (5.5)$$

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

$$P_{data}(X) = \frac{N_{data}^{NoZee}(X)}{N_{data}^{NoZee}} \quad (5.6)$$

, where  $N_{data}^{NoZee}$  is the total data 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  $NN_\tau$  bin  $X$  is then given by

$$f_{NN}(X) = \frac{P_{data}(X)}{P_{MC}(X)}. \quad (5.7)$$

This provides a good way to probe the differences of data and MC events in each given  $NN_\tau$  bin. Figure 5.8 shows the correction factor  $f_{NN}(X)$  break down in  $\tau$ -types as a function of  $NN_\tau$  bins. For those bins with medium  $NN_\tau$  values, the correction factors vary a lot because the statistics of the sample is low. However, this uncertainties are considered in the analysis which study in the hadronic  $\tau$  final states. The correction factors for high  $NN_\tau$  bins where most of the  $\tau$  leptons peak are derived in the sample of much larger statistics. Therefore, the PDF of data and MC are close to each other as one does not expect large correction factors on the efficiency. Since events peak at low  $NN_\tau$  are mostly from multi-jet backgrounds and populated by not real  $\tau$  candidates, such a correction is only available for  $\tau$ -type 1, 2 events with  $NN_\tau > 0.3$  and for  $\tau$ -type 3 events with  $NN_\tau > 0.4$ . After the re-normalization of MC events, one can calculate the efficiency  $\epsilon_{NN}(X)$  with a cut  $NN_\tau > X$  applied, where  $X$  is an arbitrary non-zero choice corresponding to the selection cut used in the analysis. The systematic

uncertainty  $\sigma_{NN}$  is then derived by three steps. In step 1, the uncertainty caused by different choices of the values of  $X$ ,  $\sigma_{NN}^X$  is given by:

$$\sigma_{NN}(X) = (1 - \epsilon_{NN}(X))/3. \quad (5.8)$$

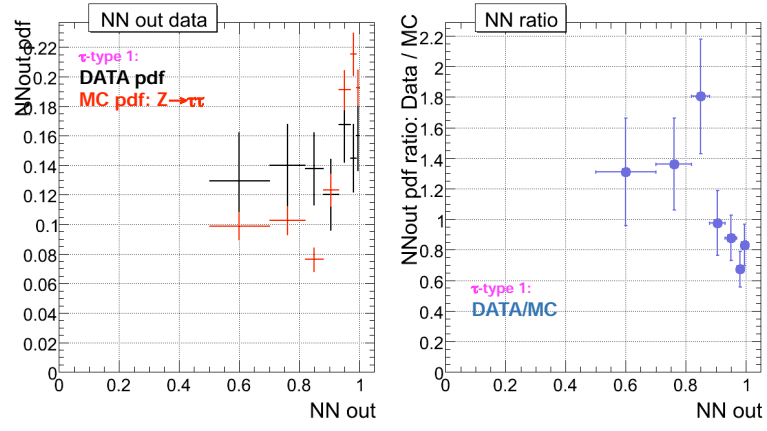
Further, the systematic uncertainty due to multi-jet background estimation,  $\sigma_{NN}^{QCD}$ , is handled by applying two estimating methods. Finally,  $\sigma_{NN}$  is calculated using the quadrature sum of  $\sigma_{NN}^X$ ,  $\sigma_{NN}^{QCD}$  and the uncertainty caused by  $\tau$  kinematics,  $\sigma_{NN}^{kin}$ , based on two transverse momentum selections,  $p_T^\tau > 25$  GeV and  $p_T^\tau > 25$  GeV. Table 5.1 lists the MC corrections and systematics for each  $\tau$ -type corresponding to the  $NN_\tau$  cut used in this analysis. The uncertainties of  $\tau$ -type 1 and 3 are

| $\tau$ -type | $NN_\tau$ cut (X) | $f_{NN}$ | $\sigma_{NN}$ (%) |
|--------------|-------------------|----------|-------------------|
| 1            | $\geq 0.9$        | 0.84     | 12                |
| 2            | $\geq 0.9$        | 0.95     | 4.2               |
| 3            | $\geq 0.95$       | 0.83     | 7                 |

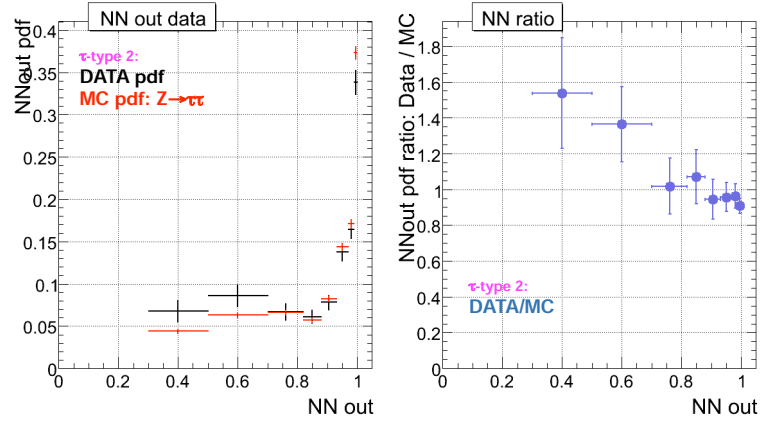
Table 5.1: Reconstructed  $\tau$  correction factors to MC events and the corresponding systematics.

### 5.4.2 Tau Energy Scale Correction

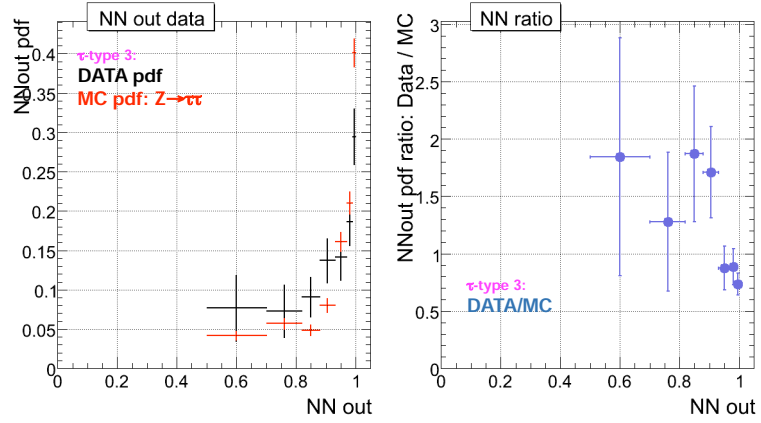
In Run IIb, the EM calibration is not properly propagated to  $\tau$ s in the reconstruction level, therefore, this has to be done in the analysis level to take into account the correct calibration for both data and MC events. Also, the difference between data and MC in the energy scale of hadronic taus is probed using the variable  $E^{cal}/P^{trk}$ , where  $E^{cal}$  is the energy measured in the calorimeter and  $P^{trk}$  is the track momentum. In this section, Run IIb  $\tau_\mu\tau_h$  channel with preselections as described in Chapter 7 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 can be due to a difference in the calorimeter energy scale between data and MC. In particular, Fig. 5.9 and Fig. 5.10 shows examples in  $p_T$  and  $E^{cal}/P^{trk}$  distributions for each  $\tau$ -type before any correction is applied to the energy scale. These variable imply that the energy scale is not fully described in MC, and the difference in  $p_T$  between data and MC can result in shifts for either the invariant mass of the  $\mu$ - $\tau$  pair, as illustrated in



(a)



(b)



(c)

Figure 5.8: Efficiencies of data and MC (left) and the correction factors  $f_{NN}(X)$  (right) derived in  $Z \rightarrow \tau^+\tau^-$  sample for  $\tau$ -type a) 1, b) 2 and c) 3 [67].

Fig. 5.11 a), and a miscalculated  $\cancel{E}_T$ , as shown in Fig. 5.11 b). Table 5.2 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 on the energy scale.

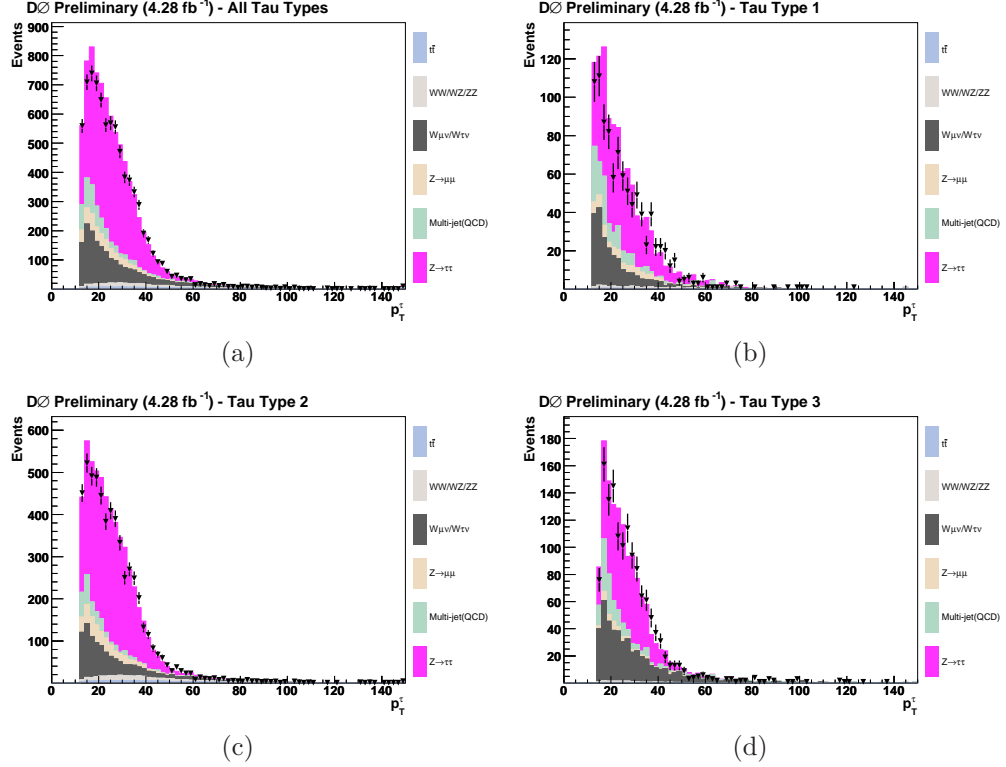


Figure 5.9:  $p_T$  distribution for a) all type and b) type 1 c) type 2 and d) type 3 taus before correcting the energy scale.

| Tau Type | $\bar{X}_{data}$ | $\bar{X}_{bkg}$ | $\Delta\bar{X}$ |
|----------|------------------|-----------------|-----------------|
| 1        | 1.17             | $1.12 \pm 0.02$ | 0.05            |
| 2        | 1.59             | $1.58 \pm 0.01$ | 0.01            |
| 3        | 0.97             | $0.96 \pm 0.01$ | 0.01            |

Table 5.2: Table showing the mean of the  $E^{cal}/P^{trk}$  distribution for the prediction from the MC and QCD  $\bar{X}_{bkg}$  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.

The corrections for hadronic  $\tau$  energy scale are derived in terms of bins of the ratio of energy clusters of  $\tau$  leptons found in EM to the total energy ( $r_{EM}$ ) [68]. Clusters with low  $r_{EM}$  correspond to taus with large contribution from charged pions, whereas clusters with high  $r_{EM}$  correspond to taus with a large contribution



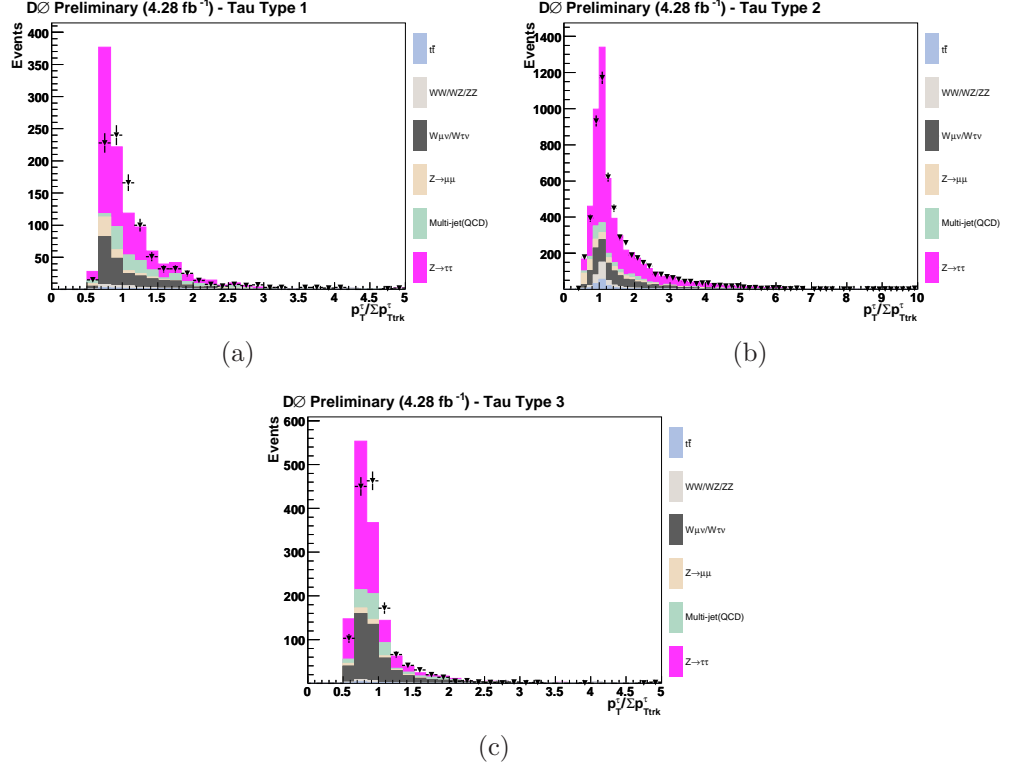


Figure 5.10:  $E^{cal}/P^{trk}$  distribution for a) type 1 b) type 2 and c) type 3 taus before correcting the energy scale.

from neutral pions. Test energy of each MC  $\tau$  lepton is allowed to vary in the region

$$0.8 \leq \frac{E_{MC}^{corr}}{E_{MC}^{ori}} \leq 1.2 \quad (5.9)$$

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

$$LL(corr) = \sum_{bin} N_{MC}^{corr}(bin) - N_{data}(bin) \times \log[N_{MC}^{corr}(bin)], \quad (5.10)$$

where  $N_{MC}^{corr}(bin)$  and  $N_{data}(bin)$  are the numbers of MC and data events in each  $r_{EM}$  bin. The correction of each  $r_{EM}$  bin is then derived by the minimum value of  $-LL(corr)$  which reflects the best fit of MC to data. There is only one  $r_{EM}$  bin used for  $\tau$ -types type 1 and 3 which means the energy of type 1 and 3 taus are corrected by multiplying the predicted  $\tau$  energy by a single number  $f_{E_\tau}$ . The formula of energy correction factor  $f_{E_\tau}(r_{EM})$  for  $\tau$ -type 2 is studied in four  $r_{EM}$

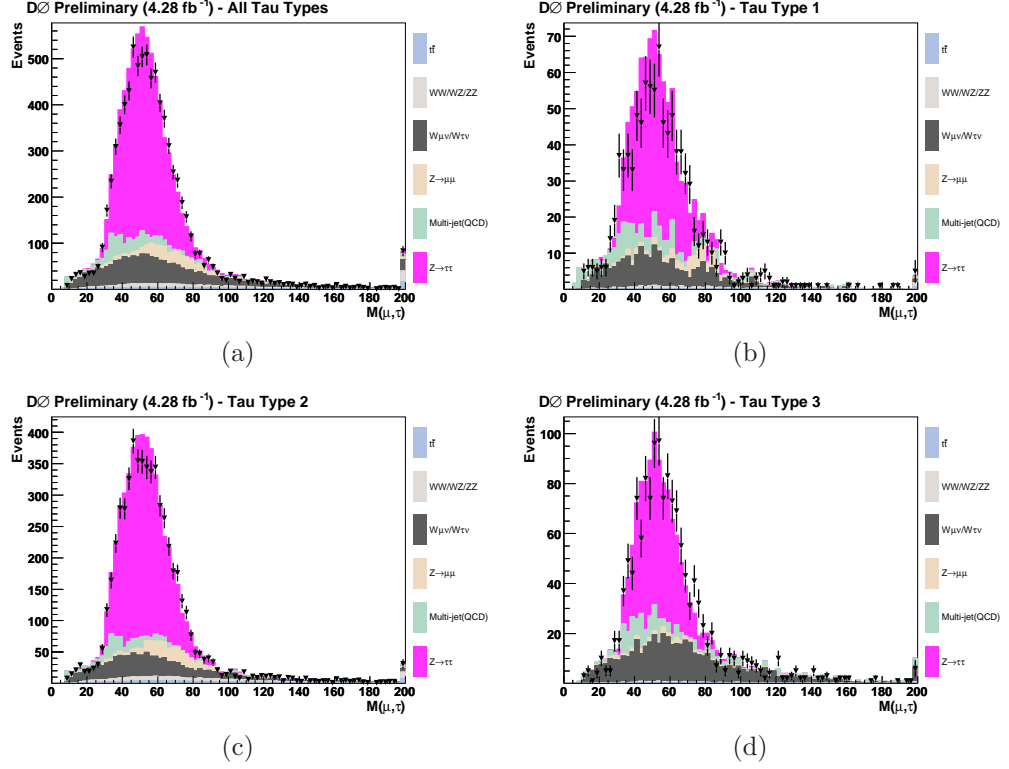


Figure 5.11: Invariant mass between  $\mu$  and  $\tau$  pairs for a) sum of all types, b) type 1, c) type 2 and d) type 3 taus before correcting the energy scale.

bins given by

$$f_{E_\tau}(r_{EM}) = a \times r_{EM} + b, \quad (5.11)$$

, and the corresponding numbers  $a$  and  $b$  are shown in Fig. 5.13 for each  $\tau$ -type. Since this correction can pull low  $p_T$  events, which originally failed the  $p_T$  cuts, back into the sample, the correction is applied prior to final selections on the hadronic taus. Figures 5.14, 5.15, and 5.16(a) show the  $p_T$  distributions,  $E^{cal}/P^{trk}$ , and invariant mass after the energy scale correction is applied. In these figures, MC events show a better description of data after the energy is corrected compare to Fig. 5.9, 5.10, 5.11, and 5.12.

The corrected calorimeter energy is used for  $\cancel{E}_T$  computation for  $\tau$ -type 2 and 3. Since  $\tau$ -type 1 are mainly from charge pion decay, the tracking momentum can be used alone with the calorimeter energy to provide a better resolution respect to the momentum measured from calorimeter only. There is no energy correction for the data, however, one can recompute the type 1  $\cancel{E}_T$  using tracking momentum as what has been done for background MC. The  $\cancel{E}_T$  distribution of sum of all  $\tau$ -types is shown in Fig. 5.17 (a), and Fig. 5.17 (b) shows the recomputed type

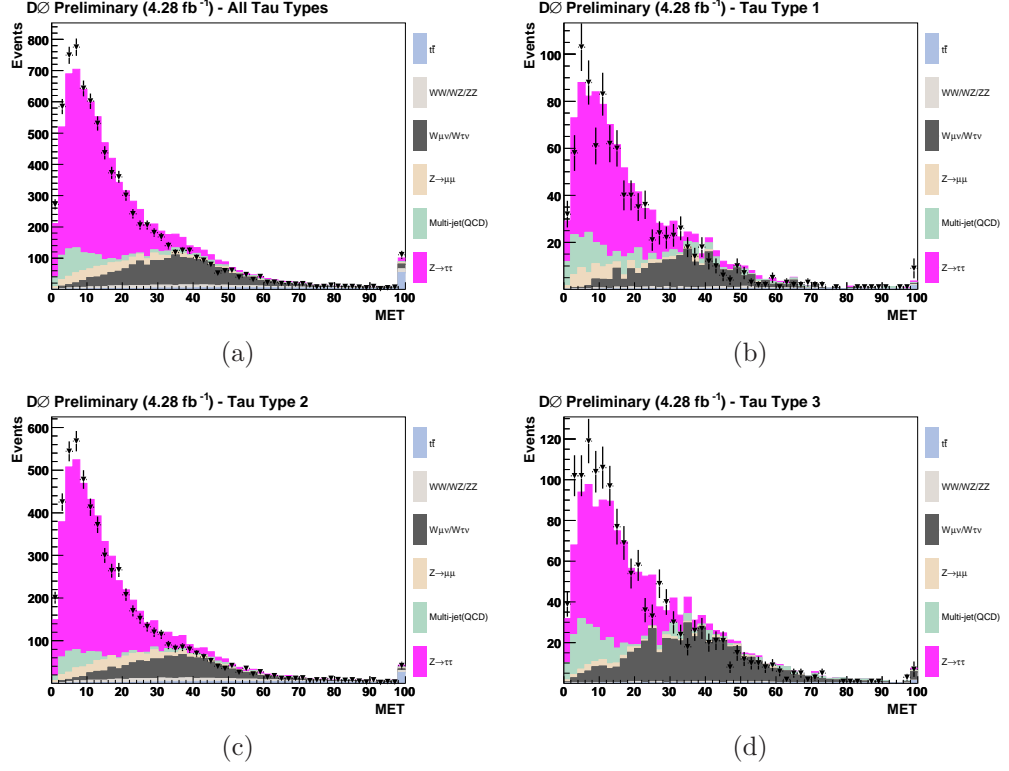


Figure 5.12:  $\cancel{E}_T$  distribution for a)sum of all types, b) type 1, c) type 2 and d) type 3 taus before correcting the energy scale.

1  $\cancel{E}_T$  from tracking momentum. A better representation in MC is observed also in  $\cancel{E}_T$  after correcting the  $\tau$  energy. Further, Table 5.3 lists the mean values of  $E^{cal}/P^{trk}$  between data and MC after applying the energy correction.

| Tau Type | $\bar{X}_{data}$ | $\bar{X}_{bkg}$ | $\Delta\bar{X}$ |
|----------|------------------|-----------------|-----------------|
| 1        | 1.18             | $1.19 \pm 0.02$ | 0.01            |
| 2        | 1.59             | $1.61 \pm 0.01$ | 0.02            |
| 3        | 0.97             | $1.01 \pm 0.01$ | 0.03            |

Table 5.3: Table showing the mean of the  $E^{cal}/P^{trk}$  distribution for the prediction from the MC and QCD  $\bar{X}_{bkg}$  and the data  $\bar{X}_{data}$  after applying the energy correction. The difference between the means in data and predicted,  $\Delta\bar{X}$ , is also listed.

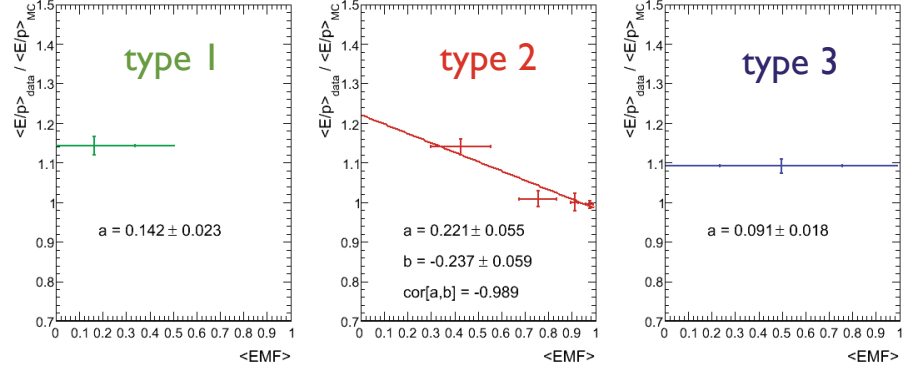


Figure 5.13: Energy correction factors as a function of  $r_{EM}$  bin for  $\tau$ -type 1 (left),  $\tau$ -type 2 (middle) and  $\tau$ -type 3 (right) [68]

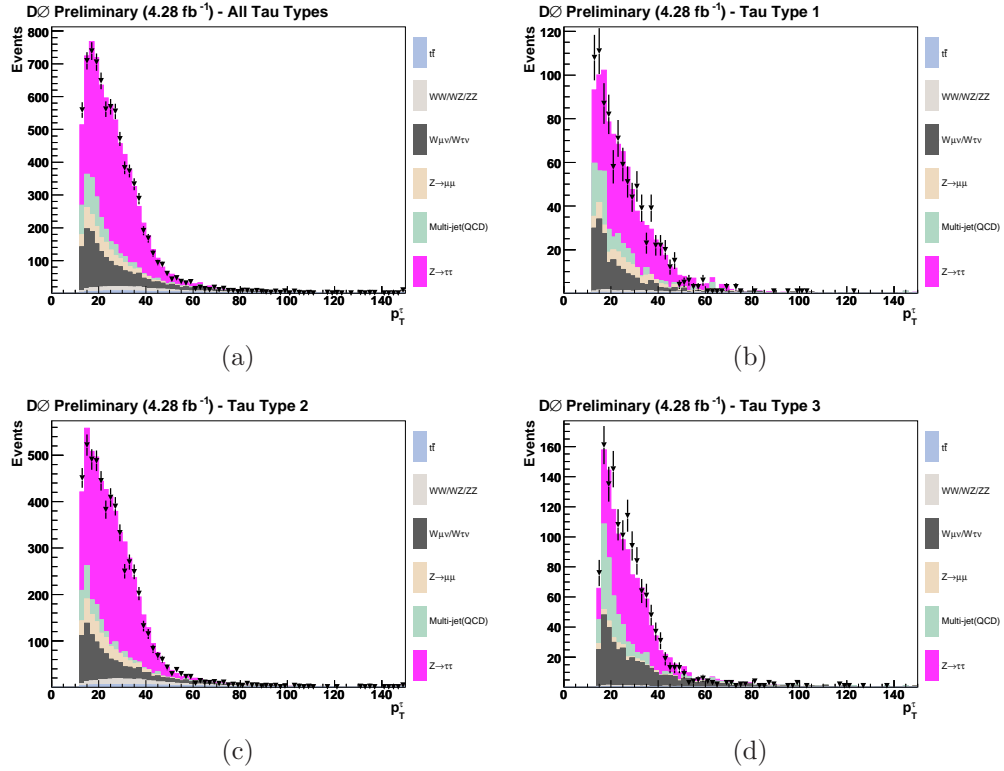


Figure 5.14:  $p_T$  distribution for a) sum of all types, b) type 1, c) type 2 and d) type 3 taus after correcting the energy scale.

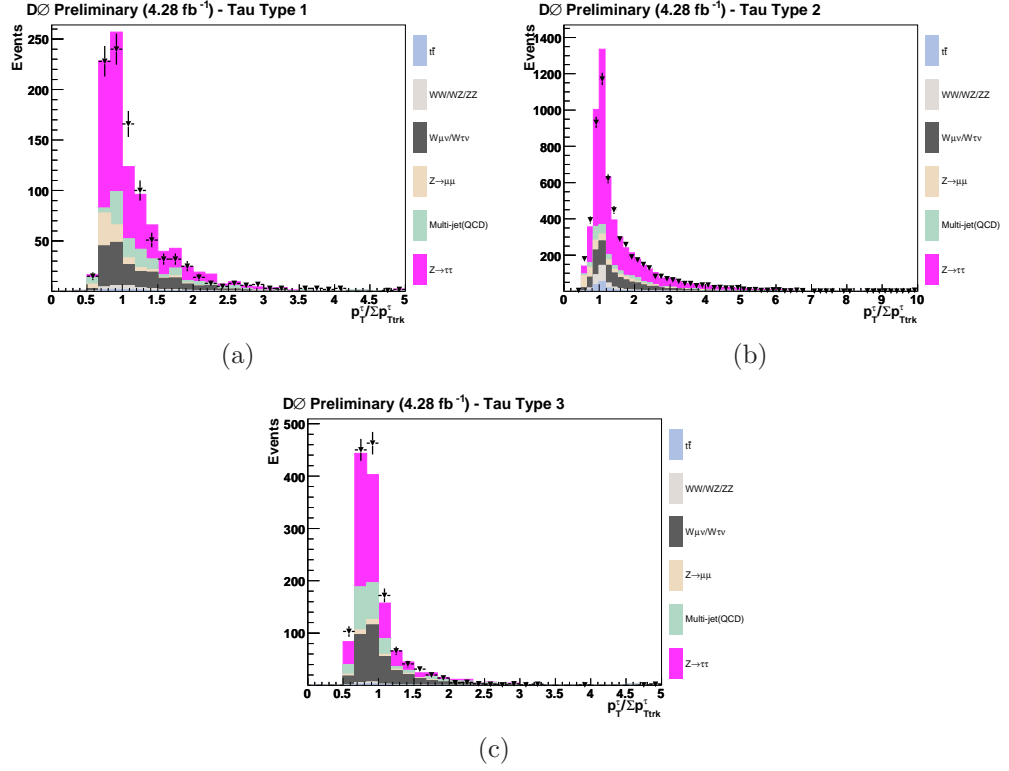


Figure 5.15:  $E^{cal} / P^{Trk}$  distribution for a) type 1 , b) type 2 and c) type 3 taus after correcting the energy scale.

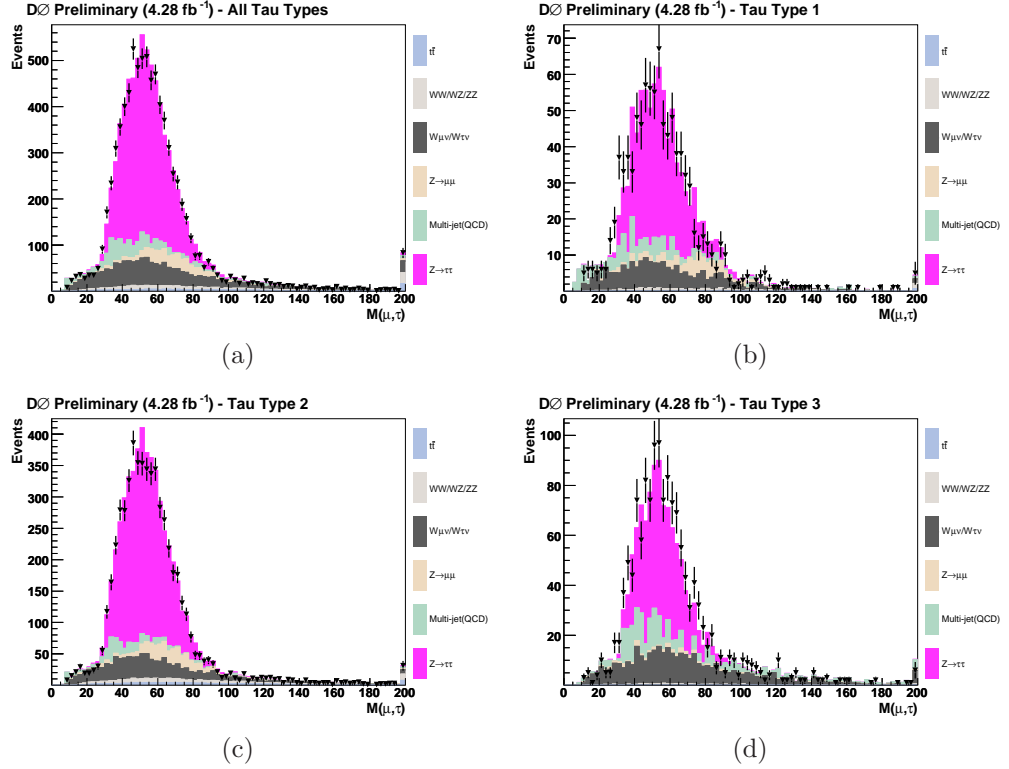
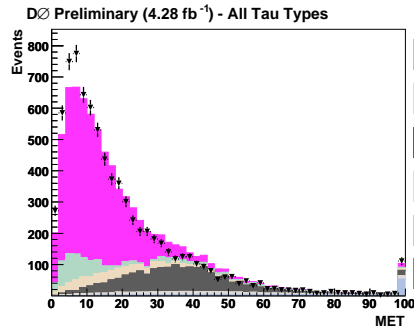
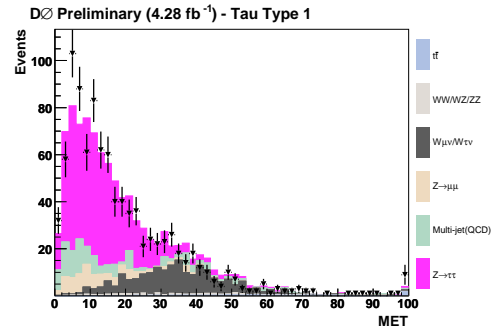


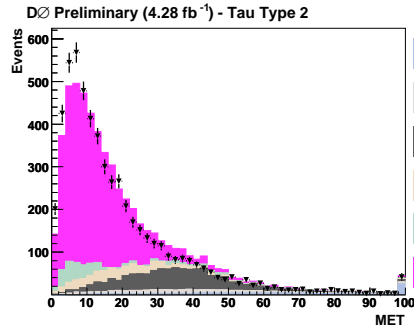
Figure 5.16: Invariant mass between  $\mu$  and  $\tau$  pairs for a)sum of all types, b) type 1, c) type 2 and d) type 3 taus after correcting the energy scale.



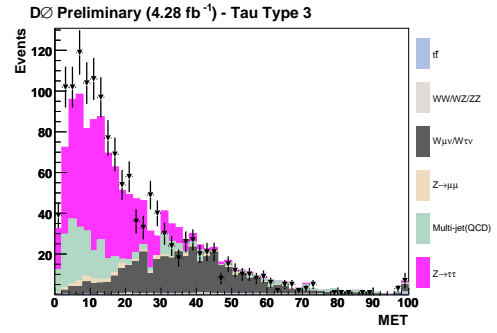
(a)



(b)



(c)



(d)

Figure 5.17:  $\cancel{E}_T$  distribution for a)sum of all types, b) type 1, c) type 2 and d) type 3 taus after correcting the energy scale.

## 5.5 MC Samples

### 5.5.1 Higgs Signal Monte Carlo

The signal samples used for different Higgs boson mass values from 90 to 320 GeV separated by intervals of 10 GeV were generated for  $\tau_\mu\tau_h$  Channel with PYTHIA 6.413 [54] using the CTEQ6L1 leading order (LO) parton distribution functions. Signal MC samples for  $\tau_e\tau_h$  channel from 90 to 200 GeV are also generated by intervals of 10 GeV, but by 20 GeV interval above 200 GeV signal mass. No signal sample used in  $\tau_e\tau_h$  channel has a mass above 300 GeV. Each sample was subsequently reconstructed using tools described in Sect. 5.1. Simulated signal samples of different masses is summarized in Table 5.5.1 with number of events after data quality selection.

The width of SM Higgs boson increases sharply at 160 GeV Higgs mass where the effect of  $W^+W^-$  decay start coming in, and the width of SM Higgs boson become greater than of MSSM Higgs bosons above 200 GeV Higgs mass as shown in Fig. 5.18. The MC Higgs signal samples used in this analysis were generated from both  $gg$  and  $b\bar{b}$  production for the standard model-like Higgs,  $h_0$ . However, for higher mass samples, a wide-width structure is observed when analyzing the MC truth information which matches the expectation (see Fig 5.19(a)). This wide-width structure appears to degrade the limit for the higher Higgs masses by yielding a wider reconstructed visible mass. In order to reduce the effect for MSSM Higgs boson search, a mass window is applied to select events only appear in  $\pm 80$  GeV region of the generated mass point. Those “fixed” true mass peak of different mass points are shown in Fig 5.19(b).

### 5.5.2 Background Monte Carlo

#### Background Monte Carlo in $\tau_\mu\tau_h$ Channel

The majority of background Monte Carlo events used in the  $\tau_\mu\tau_h$  channel were generated either with ALPGEN 2.11 [55] or PYTHIA version 6.413 [54] using the CTEQ6L1 parton distribution functions as used for the signal samples. The main background processes for the search are  $Z$  boson and Drell-Yan production decaying to  $\tau\tau$ ,  $ee$  and  $\mu\mu$ , electroweak  $W$ +jets, di-boson ( $WW$ ,  $WZ$ , and  $ZZ$ ),  $t\bar{t}$  and QCD multijet production. In this channel, the  $W$ +jet,  $Z/\gamma^*$ , and  $t\bar{t}$  MC samples have been generated using ALPGEN with parton showering and hadronization

| Higgs Mass       | $\tau_\mu\tau_h$ Channel     |                              | $\tau_e\tau_h$ Channel       |
|------------------|------------------------------|------------------------------|------------------------------|
| $m_\phi$ , [GeV] | $N_{\text{event}}$ (Run IIa) | $N_{\text{event}}$ (Run IIb) | $N_{\text{event}}$ (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 5.4: Signal Monte Carlo samples used in the analysis, with the number of events after duplicate event removal and data quality requirements,  $N_{\text{event}}$ .



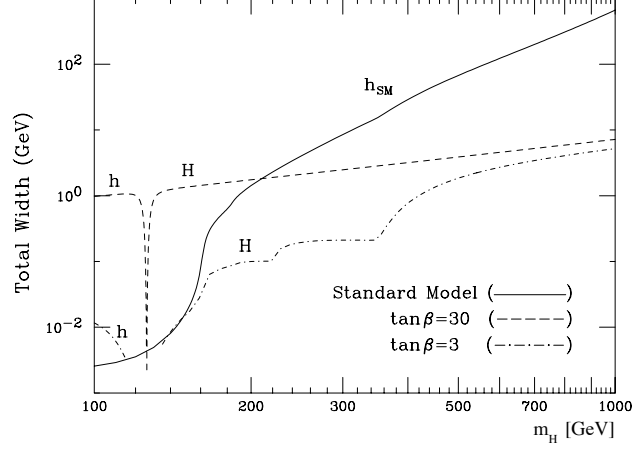


Figure 5.18: The total decay width of the SM Higgs boson, shown as a function of  $m_H$ . Also shown are the decay widths for the CP-even neutral Higgs bosons,  $h$  and  $H$ , for two choices of  $\tan\beta$ , in the MSSM benchmark scenario  $m_h^{\max}$  described in Chapter 1 [13].

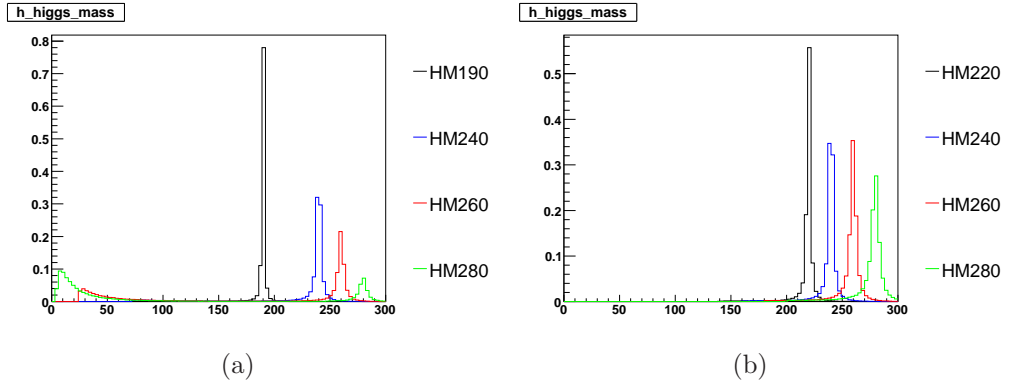


Figure 5.19: Distribution for signal true mass for mass points 190 (black), 240 (blue), 260 (red), and 280 (green) for (a) without any mass cut (b) with selections on signal true mass. All distributions are normalized to the sum of events equal to one in order to compare the differences in shapes.

provided by PYTHIA, while Di-boson processes are simulated using PYTHIA. Estimation of the QCD multi-jet background for both channels is taken directly from data and complete details on the method used is given in Sect. 6.1.

The  $Z/\gamma^* \rightarrow ll$  cross section is determined by  $\sigma(Z/\gamma^* \rightarrow ll) = \sigma_{LO} \times K_{NNLO}(Q^2)$  using CTEQ6.1M PDFs. Here, the LO cross section,  $\sigma_{LO}$  uses LO CTEQ6.1L PDF where the  $K_{NNLO}$ -factor is introduced with respect to NNLO using NLO CTEQ6.1L/M PDF to account for effects of higher-order QCD 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  $Z \rightarrow \mu\mu$  ( $Z \rightarrow \tau\tau$ ) processes generated in different invariant mass bins. The complete list is summarized in Table 5.5. Similarly, the  $W \rightarrow l\nu$  cross section is determined with CTEQ6.1M, and as indicated in Table 5.5,  $K_{NNLO}=1.266$  is applied 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 using CTEQ6.1M PDFs.

Each generated background sample is passed through the standard  $D\bar{O}$  detector simulation, digitization, and reconstruction programs. Duplicate events and events which fail the data quality check are removed prior to normalization. The number of events available for each background sample with their cross sections are listed in Tables 5.6 and 5.7.

### Background Monte Carlo in $\tau_e\tau_h$ Channel

The background Monte Carlo events used in  $\tau_e\tau_h$  channel were only generated with PYTHIA v6.323 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 consistent with each other. Table 5.8 lists the background MC samples used and the theoretical cross-section ( $\sigma$ ) used for sample normalization. Also shown in the table is the number of generated events ( $N_{\text{gen}}$ ), events after removal of duplicate events and events fail the data quality check ( $N_{\text{DQ}}$ ).

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

Table 5.5: NNLO K-factors applied during normalization of the generated  $W$ +jets and  $Z$ +jets background processes to take into account effects from higher-order QCD multijet production. The values have been determined and certified by the DØ V+jets group. For the  $W$  generated samples,  $l = \mu, \tau$ .

| Process  | Mass Range [GeV]            | $N_{\text{DQ}}$ (Run IIa) | $N_{\text{DQ}}$ (Run IIb) | $\sigma \times \text{BR}$ [pb] |
|--|-----------------------------|---------------------------|---------------------------|--------------------------------|
| $Z/\gamma^* \rightarrow \tau^+\tau^- + 0\text{lp}$ | $15 < M_{\tau\tau} < 75$    | 1,659,261                 | 1,666,177                 | 338                            |
| $Z/\gamma^* \rightarrow \tau^+\tau^- + 1\text{lp}$ | $15 < M_{\tau\tau} < 75$    | 592,821                   | 594,906                   | 39.9                           |
| $Z/\gamma^* \rightarrow \tau^+\tau^- + 2\text{lp}$ | $15 < M_{\tau\tau} < 75$    | 328,600                   | 329,981                   | 10.0                           |
| $Z/\gamma^* \rightarrow \tau^+\tau^- + 3\text{lp}$ | $15 < M_{\tau\tau} < 75$    | 340,452                   | 341,658                   | 2.77                           |
| $Z/\gamma^* \rightarrow \tau^+\tau^- + 0\text{lp}$ | $75 < M_{\tau\tau} < 130$   | 1,556,389                 | 1,562,018                 | 131                            |
| $Z/\gamma^* \rightarrow \tau^+\tau^- + 1\text{lp}$ | $75 < M_{\tau\tau} < 130$   | 595,169                   | 596,880                   | 40.3                           |
| $Z/\gamma^* \rightarrow \tau^+\tau^- + 2\text{lp}$ | $75 < M_{\tau\tau} < 130$   | 305,312                   | 306,144                   | 9.99                           |
| $Z/\gamma^* \rightarrow \tau^+\tau^- + 3\text{lp}$ | $75 < M_{\tau\tau} < 130$   | 205,365                   | 205,809                   | 3.10                           |
| $Z/\gamma^* \rightarrow \tau^+\tau^- + 0\text{lp}$ | $130 < M_{\tau\tau} < 250$  | 366,567                   | 367,982                   | 0.92                           |
| $Z/\gamma^* \rightarrow \tau^+\tau^- + 1\text{lp}$ | $130 < M_{\tau\tau} < 250$  | 180,807                   | 181,529                   | 0.38                           |
| $Z/\gamma^* \rightarrow \tau^+\tau^- + 2\text{lp}$ | $130 < M_{\tau\tau} < 250$  | 181,640                   | 182,360                   | 0.10                           |
| $Z/\gamma^* \rightarrow \tau^+\tau^- + 3\text{lp}$ | $130 < M_{\tau\tau} < 250$  | 188,620                   | 189,168                   | 0.04                           |
| $Z/\gamma^* \rightarrow \tau^+\tau^- + 0\text{lp}$ | $250 < M_{\tau\tau} < 1960$ | 546,767                   | 548,792                   | 0.07                           |
| $Z/\gamma^* \rightarrow \tau^+\tau^- + 1\text{lp}$ | $250 < M_{\tau\tau} < 1960$ | 488,524                   | 490,870                   | 0.04                           |
| $Z/\gamma^* \rightarrow \tau^+\tau^- + 2\text{lp}$ | $250 < M_{\tau\tau} < 1960$ | 371,192                   | 372,497                   | 0.01                           |
| $Z/\gamma^* \rightarrow \tau^+\tau^- + 3\text{lp}$ | $250 < M_{\tau\tau} < 1960$ | 178,014                   | 178,756                   | 0.004                          |
| $Z/\gamma^* \rightarrow \mu^+\mu^- + 0\text{lp}$   | $15 < M_{\mu\mu} < 75$      | 1,868,435                 | 1,877,017                 | 344                            |
| $Z/\gamma^* \rightarrow \mu^+\mu^- + 1\text{lp}$   | $15 < M_{\mu\mu} < 75$      | 640,691                   | 642,803                   | 40.1                           |
| $Z/\gamma^* \rightarrow \mu^+\mu^- + 2\text{lp}$   | $15 < M_{\mu\mu} < 75$      | 324,113                   | 325,330                   | 9.87                           |
| $Z/\gamma^* \rightarrow \mu^+\mu^- + 3\text{lp}$   | $15 < M_{\mu\mu} < 75$      | 328,210                   | 329,589                   | 2.84                           |
| $Z/\gamma^* \rightarrow \mu^+\mu^- + 0\text{lp}$   | $75 < M_{\mu\mu} < 130$     | 1,553,222                 | 1,565,523                 | 134                            |
| $Z/\gamma^* \rightarrow \mu^+\mu^- + 1\text{lp}$   | $75 < M_{\mu\mu} < 130$     | 639,392                   | 641,154                   | 41.4                           |
| $Z/\gamma^* \rightarrow \mu^+\mu^- + 2\text{lp}$   | $75 < M_{\mu\mu} < 130$     | 446,737                   | 447,655                   | 9.91                           |
| $Z/\gamma^* \rightarrow \mu^+\mu^- + 3\text{lp}$   | $75 < M_{\mu\mu} < 130$     | 172,628                   | 173,038                   | 3.25                           |
| $Z/\gamma^* \rightarrow \mu^+\mu^- + 0\text{lp}$   | $130 < M_{\mu\mu} < 250$    | 359,102                   | 360,533                   | 0.89                           |
| $Z/\gamma^* \rightarrow \mu^+\mu^- + 1\text{lp}$   | $130 < M_{\mu\mu} < 250$    | 179,991                   | 180,775                   | 0.36                           |
| $Z/\gamma^* \rightarrow \mu^+\mu^- + 2\text{lp}$   | $130 < M_{\mu\mu} < 250$    | 179,228                   | 179,895                   | 0.98                           |
| $Z/\gamma^* \rightarrow \mu^+\mu^- + 3\text{lp}$   | $130 < M_{\mu\mu} < 250$    | 170,045                   | 170,809                   | 0.03                           |
| $Z/\gamma^* \rightarrow \mu^+\mu^- + 0\text{lp}$   | $250 < M_{\mu\mu} < 1960$   | 560,971                   | 564,091                   | 0.07                           |
| $Z/\gamma^* \rightarrow \mu^+\mu^- + 1\text{lp}$   | $250 < M_{\mu\mu} < 1960$   | 466,903                   | 469,071                   | 0.03                           |
| $Z/\gamma^* \rightarrow \mu^+\mu^- + 2\text{lp}$   | $250 < M_{\mu\mu} < 1960$   | 367,301                   | 368,836                   | 0.01                           |
| $Z/\gamma^* \rightarrow \mu^+\mu^- + 3\text{lp}$   | $250 < M_{\mu\mu} < 1960$   | 366,606                   | 368,377                   | 0.004                          |

Table 5.6:  $Z$  boson Background Monte Carlo samples used in the  $\tau_\mu\tau_h$  channel given with the number of events passed the data quality checks,  $N_{\text{DQ}}$ , and the production of cross section and branching ratio  $\sigma \times \text{BR}$  for each process. For each  $Z/\gamma^*$  sample, the cross section is calculated at NNLO.

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

Table 5.7: Other background Monte Carlo samples used in the  $\tau_\mu\tau_h$  channel given with the number of events passed the data quality checks,  $N_{\text{DQ}}$ , and the product of cross section and branching ratio  $\sigma \times \text{BR}$  for each process. Both the  $W$  and  $t\bar{t}$  cross section is 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.

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

Table 5.8: Background Monte Carlo samples used in the  $\tau_e\tau_h$  channel, with the number of events generated,  $N_{\text{gen}}$ , the number of events after duplicate event removal and data quality requirements,  $N_{\text{DQ}}$ , and the cross-section  $\sigma$  used for normalization.

# Chapter 6

## Background Estimation

Good background modeling is very important to this search to probe signal events on top of the predicted background. Multi-jet background is one major contribution which will be discussed in Sect. 6.1, and the other main sources are listed below:

**$Z \rightarrow \tau\tau$ :** Two  $\tau$  leptons decayed from neutral  $Z$  boson, where one  $\tau$  decays leptonically to an electron or a muon and the other  $\tau$  decays hadronically. It is almost impossible to separate this background from signal apart from using spin or mass. Therefore, the strategy of this analysis is to reduce the other background. Signal events are then searched on top of the  $Z \rightarrow \tau^+\tau^-$  background.

**$Z \rightarrow \mu\mu$  and  $Z \rightarrow ee$ :** Events contain one well reconstructed electron or muon candidate, and the second one is either missing but the events is associated with a jet, or is mis-identified as a  $\tau$  lepton. In the  $\tau_\mu\tau_h$  channel, background events from  $Z \rightarrow \mu^+\mu^-$  process can be highly suppressed by matching the  $\tau$  object and the muon, and only events with  $\Delta R(\mu, \tau) \geq 0.5$  are kept. The modeling and rejection of  $Z \rightarrow e^+e^-$  background in  $\tau_e\tau_h$  channel is more complicated and are discussed in Sect. 6.3.

**$W + \text{jets}$ :** There are two possible ways that this background can fake the signal events. If  $W$  boson decays to an electron or a muon and neutrinos, the events can fake signals by mis-identifying of jets as  $\tau$  leptons. Since  $W$  boson also decays to  $\tau$  leptons and neutrinos, this background will also be mis-identified as signal events if jets which fake 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 and  $\cancel{E}_T$  or on  $m_T$  variable. More details about the selection are discussed in

Chapter 7, and re-weighting factor applied to MC events from this background in  $\tau_e\tau_h$  channel can be found in Sect. 6.3.

**WW, ZZ, ZW and  $t\bar{t}$ :** These background sources contain one well reconstructed electron or muon plus a  $\tau$  candidate, therefore, it is not easy to distinguish signal events from these background events. However, since the production cross-sections are all very small, the contribution is almost negligible compared to the contribution from  $Z \rightarrow \tau^+\tau^-$  events.

## 6.1 Multi-jet (QCD) Background Estimation

One contribution to multi-jet background is expected from heavy flavor multi-jet events, where a lepton from the heavy flavor decay passes the isolation requirement and a jet is misidentified as a  $\tau$ . Additionally, a contribution is expected from light quark multi-jet events where the jets fake both the  $\tau$  lepton and the muon or electron. These background source are referred to as the 'QCD' or 'multi-jet' background. It is difficult to be simulated, and therefore, is estimated directly from data and normalized to number of events collected in signal regions.

The primary method used to estimate multi-jet events relies on defining three different control regions determined by  $NN_\tau$ :

- **High  $NN_\tau$  region (hNN):**

$NN_\tau > 0.9$  ( $\tau$ -types 1 and 2);  $NN_\tau > 0.95$  ( $\tau$ -type 3). Since the  $\tau$  identification algorithm selects events based on  $NN_\tau$  bins for discriminating  $\tau$  leptons from jet backgrounds and  $\tau$  leptons should have values peak close to one, this region is also termed as the “signal region”.

- **Low  $NN_\tau$  region (lNN):**

$NN_\tau \leq 0.10$  (all  $\tau$ -types). Composed primarily of  $W$ +jet and QCD multi-jet events and as described in the text below, used to derive the proper normalization of multi-jet events in the signal region.

- **Medium  $NN_\tau$  region (mNN):**

$0.25 \leq NN_\tau \leq 0.75$  (all  $\tau$ -types). Relatively flat region of  $NN_\tau$  phase space and composed of  $W$ +jets and QCD multi-jet events. Since multi-jet background events also exist in this region, it can be used as a suitable cross check on the QCD background composition in the search.

Events from each of the above three  $NN_\tau$  regions are further separated by selecting events with the opposite-sign charge (OS) or same-sign (SS) electric charge of electron (e) or muon ( $\mu$ ) of  $\tau$  pairs. Since multi-jet background has a random selection of the signs of two leptons originating from the jets, one expects to collect equal number of OS and SS multi-jet backgrounds. However, muon and  $\tau$  which decay from a neutral Higgs boson should be OS in order to yield the conservation of electromagnetic charge. Consequently, this selection yields a total of five different control regions as listed in Table 6.1 plus a signal region,  $N_{hNN}^{OS}$ , where events are selected by high  $NN_\tau$  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 | $N_{lNN}^{OS}$     | $N_{mNN}^{OS}$                | —                        |

Table 6.1: Six control regions defined by the value of  $NN_\tau$  and the sign of  $\mu/e - \tau$  pair. Notations for number of events yield in each region are also.

Next, the transverse mass ( $m_T$  variable) is defined as:

$$m_T = \sqrt{2p_T^{\mu/e} \cancel{E}_T (1 - \cos \Delta\phi)} \quad (6.1)$$

where  $p_T^{\mu/e}$  is the transverse momentum of the lepton,  $\cancel{E}_T$  is the missing transverse energy of the event and  $\Delta\phi$  is the angle between  $\cancel{E}_T$  and the lepton. This variable is used to separate regions which are enriched with QCD processes or  $W$ +jet events. Studies have shown that  $W$ +jet events tend to peak at higher  $M_T$  values while multi-jet production populate lower  $M_T$  bins. Therefore, events are selected by requiring  $M_T \leq 50$  GeV for the multi-jet studies for all  $\tau$ -types in order to properly define a multi-jet enriched sample. Within this method, the shape for the multi-jet contribution is taken from a sample which requires the  $\mu/e - \tau$  pairs to have the same sign to keep orthogonal to the signal sample. All other selections for such a sample are kept the same as preselections on  $\tau$  leptons lie in high  $NN_\tau$  region. This multi-jet sample is eventually normalized to OS  $\mu/e - \tau$  pairs using a normalization factor  $\rho_{QCD}$  derived from events within the low  $NN_\tau$  region for each  $\tau$ -type.

The number of QCD multi-jet events in the signal region,  $N_{hNN}^{OSQCD}$ , is assumed to be given by

$$N_{hNN}^{OSQCD} = \rho_{QCD} * (N_{hNN}^{SSdata} - N_{hNN}^{SSMC}), \quad (6.2)$$

where  $N_{hNN}^{SSdata}$  and  $N_{hNN}^{SSMC}$  are the total data and MC events found in the SS, high



$NN_\tau$  region and normalized by  $\rho_{QCD}$ . Numerical values for  $\rho_{QCD}$  are determined by selecting multi-jet events within the low  $NN_\tau$  region. Mathematically,  $\rho_{QCD}$  is given by

$$\rho_{QCD} = \frac{N_{INN}^{OSdata} - N_{INN}^{OSMC}}{N_{INN}^{SSdata} - N_{INN}^{SSMC}}, \quad (6.3)$$

where  $N_{INN}^{OSdata}$  and  $N_{INN}^{OSMC}$  are the numbers of data and MC events expected in the opposite sign sample with a low  $NN_\tau$  cut. And  $N_{INN}^{SSdata}$  and  $N_{INN}^{SSMC}$  are the total data and MC events expected in the same sign sample in low  $NN_\tau$  region. Since there is no “sign preference” of the multi-jet events, the value of  $\rho_{QCD}$  is expected to be very close to 1. As indicated in Equation 6.2 and Equation 6.3, all respective backgrounds, other than multi-jet events being estimate here, are subtracted using the corresponding MC background samples listed in Sec. 5.5.2 using K-factors described in the same section.

Discussions about the verification of this method to estimate the multi-jet background in the  $\tau_\mu\tau_h$  search channel can be found in Sect. 6.2.1. Similar verification and the cross-checks with the second method using Run IIa data in the  $\tau_e\tau_h$  search channel are discussed in Sect. 6.3.1.

## 6.2 Background Modeling in $\tau_\mu\tau_h$ Channel

Most sources of background in this search channel are generated PYTHIA or ALPGEN, and the sample used are listed in Sect. 5.5.2. The predicted backgrounds are able to describe data well by including multi-jet background estimated from data and MC with K-factors used for higher order corrections and the other sources of MC corrections (see Chapter 5) applied. Therefore, no additional treatment to the background is required. Cross-check of multi-jet background is discussed in the following Sect. 6.2.1, and more comparison of data and predicted background with preselections and final selections can be found in Chapter 7.

### 6.2.1 Multi-jet in $\tau_\mu\tau_h$ Channel

Values for  $\rho_{QCD}$  derived in this channel for each  $\tau$ -type are listed in Table 6.2, and the definition of this variable can be found in Sect. 6.1. In order to verify if the method estimates the QCD multi-jet background well, basic kinematic distributions for the  $\tau$  and muon candidates are studied in the control region dominated by this background. In  $\tau_\mu\tau_h$  channel, this control sample is chosen to contain

events with medium  $NN_\tau$  value and the  $\mu - \tau$  pairs are required to have opposite signs of the electric charges. Figure 6.1 shows examples from Run IIB data set of the missing energy ( $\cancel{E}_T$ ), transverse mass ( $m_T$ ), muon transverse momentum ( $p_T^\mu$ ) and the mass of  $\mu - \tau$  pairs ( $M_{\mu\tau}$ ) distributions, while Fig. 6.2 shows the transverse momentum distribution of  $\tau$  lepton ( $p_T^\tau$ ) by  $\tau$ -type. These distributions show good agreement between data and expected backgrounds which indicates the multi-jet background method is well-modeled by taking shape from SS region and using  $\rho_{QCD}$  for normalization.

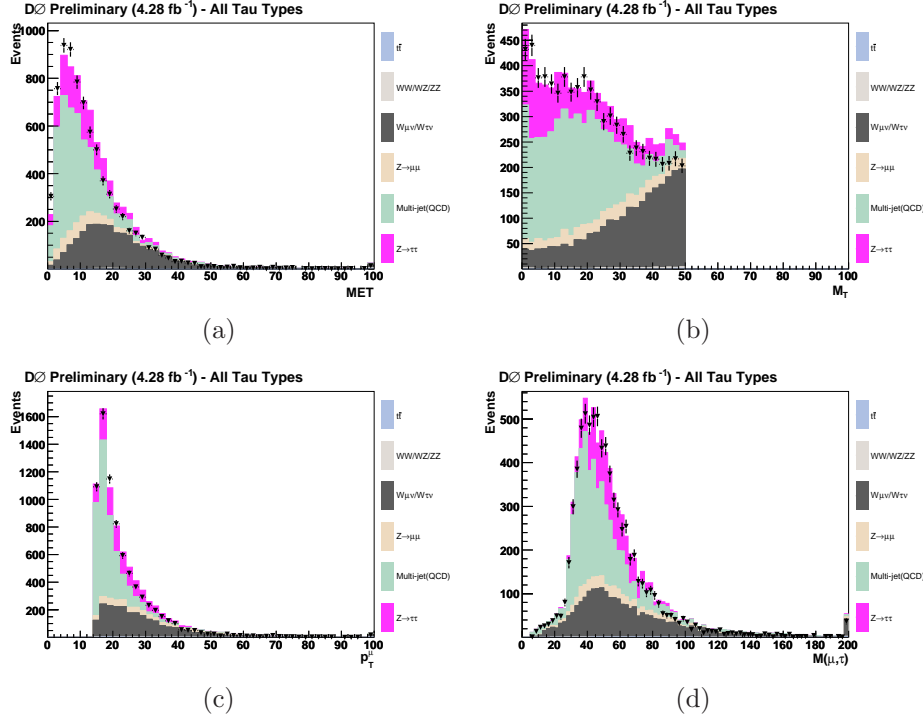


Figure 6.1: (a)  $\cancel{E}_T$ , (b)  $m_T$  (c)  $p_T^\mu$  and (d)  $M_{\mu\tau}$  for the data compared to the sum of the expected backgrounds in mNN sample with QCD shape estimated from SS data.

As a cross-check as well as determining the systematic uncertainty for the multi-jet background, a second method is considered. While main method primarily relies on the shape taken from the SS sample, the cross-check method uses events with medium  $NN_\tau$  values to model the shape. The strategy is to determine the shape of the multi-jet background by selecting OS candidates within the mNN region. The ratio of SS events with high  $NN_\tau$  values to events with medium  $NN_\tau$  values is used to normalize this selected mNN sample to predict the total number of multi-jet events contribute to the signal region. The same requirements as those for the preselection (see Chapter 7) are applied to select

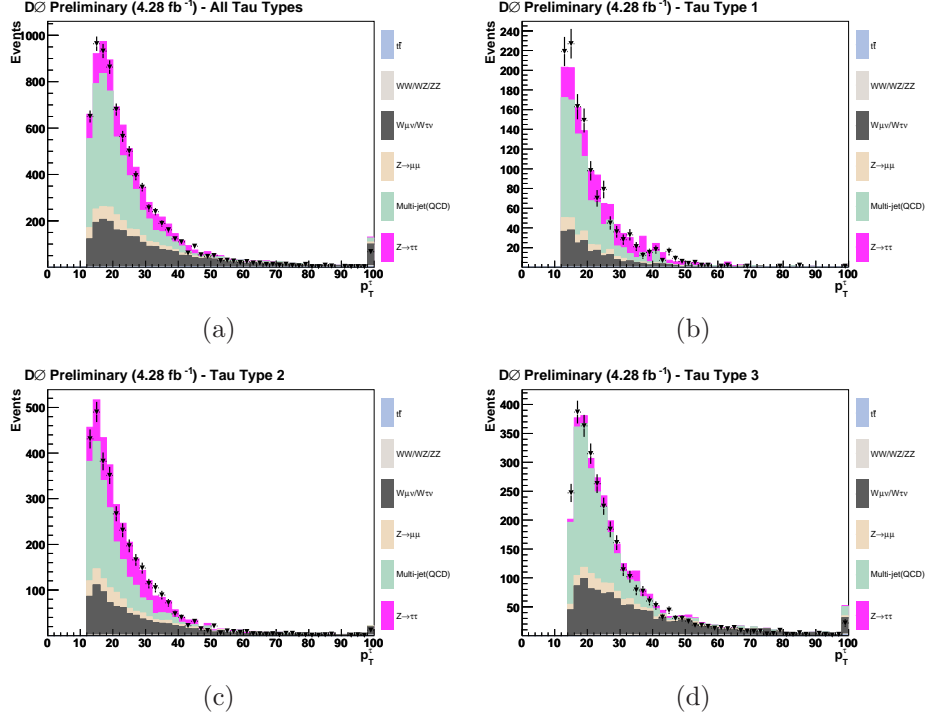


Figure 6.2:  $p_T^\tau$  for data compared to the sum of the expected backgrounds for (a) sum of all  $\tau$ -types, (b) type 1, (c) type 2 and (d) type 3 in mNN sample with QCD shape estimated from SS data.

mNN multi-jet events.

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

$$N_{hNN}^{OSQCD} = \rho_2 \times (N_{mNN}^{OSdata} - N_{mNN}^{OSMC}), \quad (6.4)$$

where  $N_{hNN}^{OSQCD}$  is the number of the QCD events expected in the OS sample within the high  $NN_\tau$  region, the signal region.  $N_{mNN}^{OSdata}$  and  $N_{mNN}^{OSMC}$  are the total data and MC events expected in the OS medium  $NN_\tau$  region enriched sample, normalized by the factor  $\rho_2$ . In such a calculation, all physics background MC samples are used in the estimation. Further,  $\rho_2$  is determined by selecting events enriched with multi-jet processes from a SS data sample and taking the ratio:

$$\rho_2 = \frac{N_{hNN}^{SSdata} - N_{hNN}^{SSMC}}{N_{mNN}^{SSdata} - N_{mNN}^{SSMC}}, \quad (6.5)$$

where  $N_{mNN}^{SSdata}$  and  $N_{mNN}^{SSMC}$  are the total SS data and MC events, respectively, expected in the medium  $NN_\tau$  region. Similarly, the total SS data and MC events

in high  $NN_\tau$  region,  $N_{hNN}^{SSdata}$  and  $N_{hNN}^{SSMC}$ , are used to extrapolate the number of multi-jet events in the signal region for the final sample. and these values are summarized by  $\tau$ -type in Table 6.2. Table 6.3 lists the number of multi-jet events estimated from the two methods. The difference in the yield per  $\tau$ -type is subsequently used to set the systematic uncertainty on the multi-jet background estimation. In particular, a systematic uncertainty of 8.3%, 16.2%, and 13.8% is assigned  $\tau$ -types 1, 2, and 3, respectively.

|              | Type 1            | Type 2            | Type 3            |
|--------------|-------------------|-------------------|-------------------|
| $\rho_{QCD}$ | $1.020 \pm 0.008$ | $1.026 \pm 0.005$ | $1.103 \pm 0.019$ |
| $\rho_2$     | $0.182 \pm 0.018$ | $0.245 \pm 0.015$ | $0.150 \pm 0.012$ |

Table 6.2: Normalization factors of two methods by each  $\tau$  type in  $\tau_{mu}\tau_h$  channel. The errors are taken statistically.

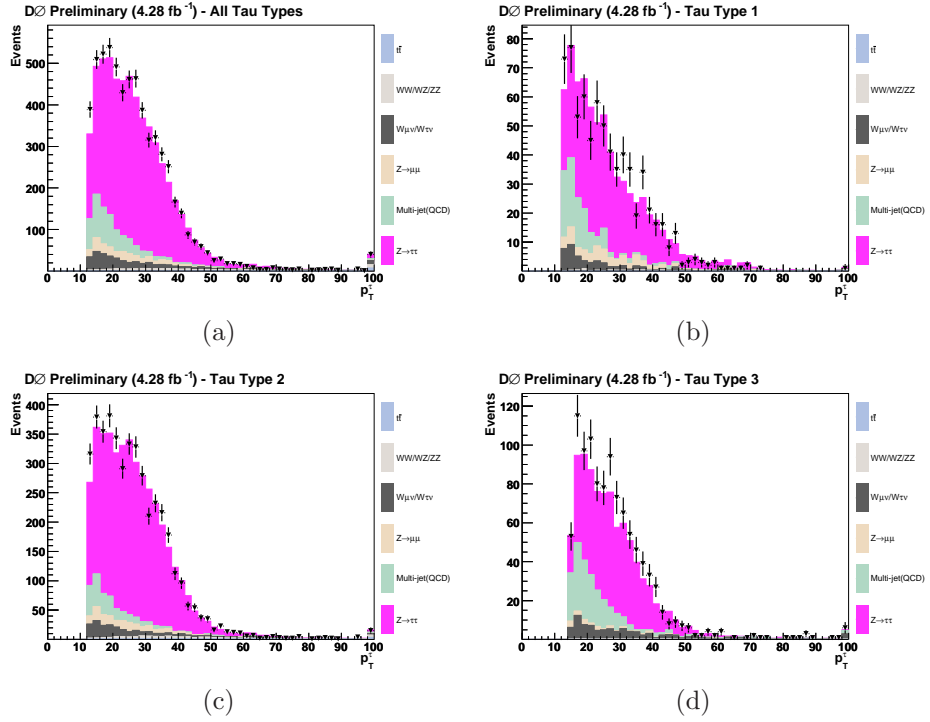


Figure 6.3: Tau transverse momentum for data compared to the sum of the expected backgrounds for (a) sum of all types, (b) type 1 taus, (c) type 2 taus and (d) type 3 taus in SS sample with QCD shape estimated from mNN data.

| Number of Multi-jet Events | $\tau$ -type 1 | $\tau$ -type 2 | $\tau$ -type 3 |
|----------------------------|----------------|----------------|----------------|
| Main Method                | $121 \pm 11$   | $351 \pm 19$   | $224 \pm 16$   |
| Cross-check Method         | $110 \pm 10$   | $289 \pm 17$   | $196 \pm 14$   |

Table 6.3: Numbers of QCD multi-jet events for the two different methods described in the text.

## 6.3 Background Modeling in $\tau_e\tau_h$ Channel

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

- **$W$ +jets:** Since PYTHIA is a leading order generator, higher order effects have to be taken into account by applying normalization factors for this background. These factors are derived in a OS sample enriched by  $W$ +jets events. The main selections to suppress multi-jet contribution in this sample are:
  - a series of preselections as used to select the signal sample (Sect. 7.1),
  - $\cancel{E}_T \geq 20$  GeV,
  - transverse momentum of electrons  $p_T^e \geq 20$  GeV.

There is no  $m_T$  cut applied, therefore, multi-jet background is expected to contribute a lot within low  $m_T$  region. A fit of the multi-jet background with low values of  $m_T$  is performed and extrapolated to high  $m_T$  region. The  $W$ +jets normalization factors  $\rho_W$  based on the comparison of data and MC are derived after multi-jet and other sources of background are subtracted from data events by

$$\rho_W = \frac{N_{data} - N_{bkg}^{NoW}}{N_{MC}^W}, \quad (6.6)$$

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 data events. This factor derived in OS region with the same selections listed above is then applied to the events passing final selections. The same study is also done

in SS region as a cross-check, and the normalization factors derived in both SS and OS samples are listed in Table 6.4 [69].

| $\tau$ -Type | $\rho_W^{SS}$   | $\rho_W^{OS}$   |
|--------------|-----------------|-----------------|
| 1            | $2.99 \pm 0.23$ | $1.88 \pm 0.07$ |
| 2            | $1.73 \pm 0.03$ | $1.36 \pm 0.02$ |
| 3            | $0.96 \pm 0.01$ | $0.94 \pm 0.01$ |

Table 6.4: Table showing the normalization factors for the W background for both opposite ( $\rho_W^{OS}$ ) and same ( $\rho_W^{SS}$ ) sign events. Errors are only taken statically.

- $Z \rightarrow e^+e^-$ : Since reconstructed electrons are required to have only one track and narrow calorimeter clusters, they are frequently also reconstructed as  $\tau$ -type 2 candidates. In addition, when electrons fall in the ICD region which has no EM coverage, they will mis-identified as  $\tau$ -type 1 candidates. Further, if electrons are associated with other tracks, they can also be reconstructed as  $\tau$ -type 3 candidates. In order to well model the electron background, the efficiency correction described in Sect. 5.2 is firstly applied to MC. Secondly, a series of selections is needed to remove the electron contamination in the tau candidates. More details about these anti-electron selections can be found in Sect. 7.2.2.

Cross-check of multi-jet background is discussed in the following Sect. 6.3.1, and more comparison of data and predicted background with preselections and final selections can be found in Chapter 7. All the rest background sources can be well presented by PYTHIA with a normalization to the luminosity  $\mathcal{L}$  and number of generated events passed the data quality checks ( $N_{DQ}$ ) as described in Sect. 5.5.2.

### 6.3.1 Multi-jet in $\tau_e\tau_h$ Channel

The same estimation method of the multi-jet background is also applied to the  $\tau_e\tau_h$  channel as well as for the  $\tau_\mu\tau_h$  channel (see Sect. 6.2.1). It is derived independently to each  $\tau$ -type, and the values for  $\rho_{QCD}$  in this channel are listed in Table 6.5. The definition of  $\rho_{QCD}$  can be found in Sect. 6.1. These factors are calculated using the sample with the other sources of background are already well modeled. In order to check the shape of the estimated multi-jet background describes data well, data is again compared to the expectation from MC and

multi-jet events in a sample collected from low  $NN_\tau$  region which is enriched by multi-jet. These distributions are shown in Fig. 6.4 for the transverse momentum of  $\tau$  ( $p_T^\tau$ ) by each type and in Fig. 6.5 for the transverse momentum ( $p_T^e$ ) and pseudo-rapidity ( $\eta^e$ ) of the electron, the missing transverse energy ( $\cancel{E}_T$ ) and the transverse mass  $m_T$ . As seen from the figures, these samples are dominated by multi-jet events shown in green colour, and all show good agreement between the background estimation and the data.

|              | Type 1 | Type 2 | Type 3 |
|--------------|--------|--------|--------|
| $\rho_{QCD}$ | 1.17   | 1.05   | 1.03   |

Table 6.5: Normalization factors by each  $\tau$  type applied in Run IIa  $\tau_e\tau_h$  channel.

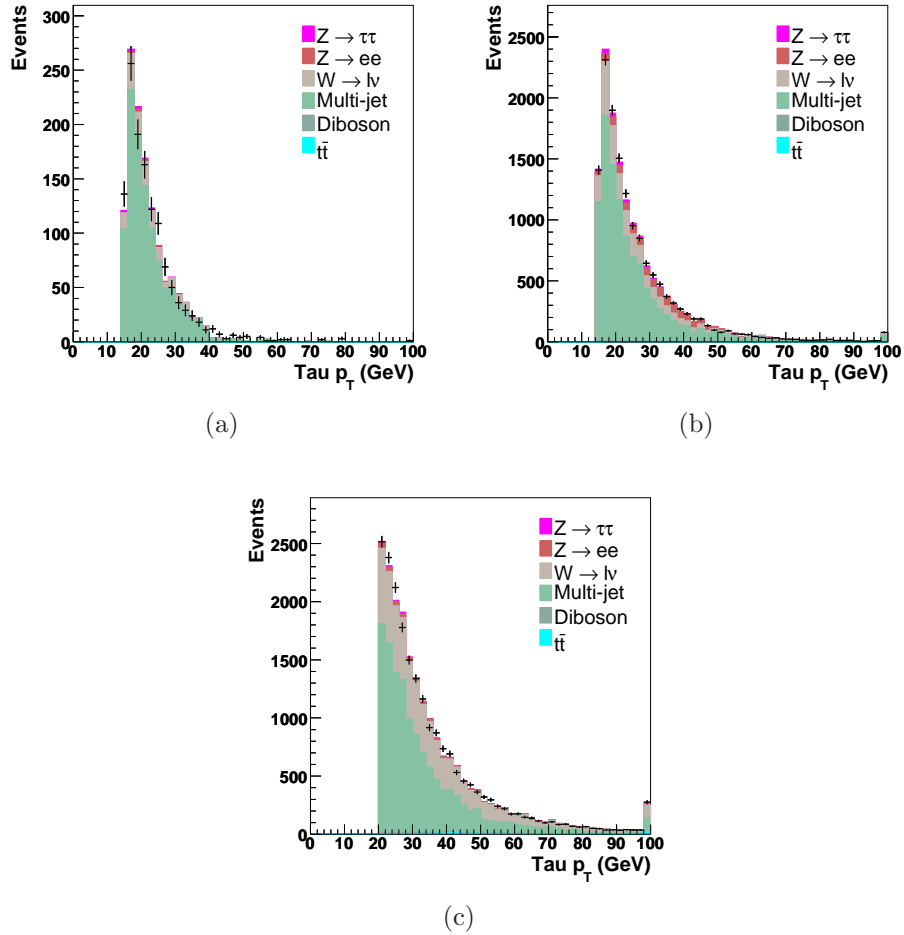


Figure 6.4:  $p_T\tau$  of events in the low  $NN_\tau$  region for (a)  $\tau$ -type 1, (b)  $\tau$ -type 2 and (c)  $\tau$ -type 3 events.

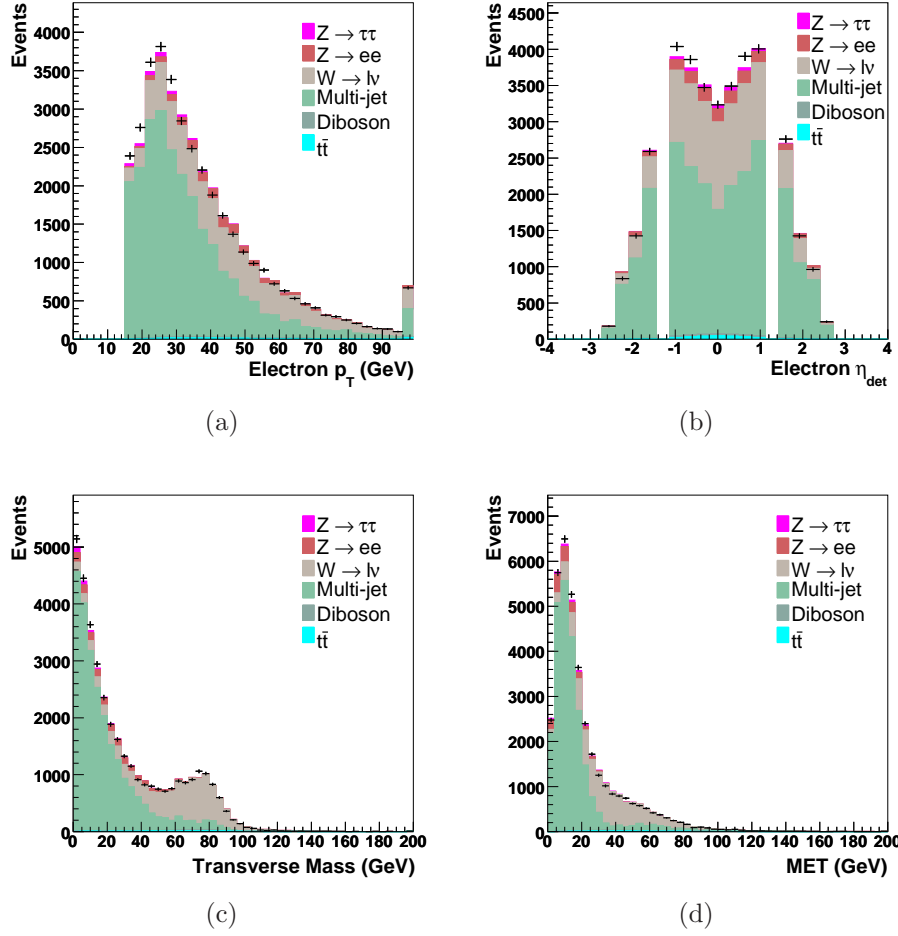


Figure 6.5: (a)  $p_T^e$  and (b)  $\eta_e$  (c)  $m_T$  (d)  $\cancel{E}_T$  of events in low  $NN_\tau$  region.



# Chapter 7

## Event Selections

### 7.1 Preselection Criteria

At the preselection stage, one electron or muon alone with one hadronically decayed  $\tau$  candidate are selected. The criteria for both  $\tau_\mu\tau_h$  and  $\tau_e\tau_h$  channels are described separately in the following sections, and Table 7.1 provided a summary of the preselections used in both channels. Many studies about the background modeling, including the efficiency and energy corrections described in Chapter 5 and 6 are done at this stage. Final selection cuts (see Section 7.2) are applied on top of those well-understanding data and background prediction.

#### 7.1.1 $\tau_\mu\tau_h$ Channel

The preselection requires one isolated muon arising from the leptonic decay of one  $\tau$  candidate plus another one hadronic decayed  $\tau$ . The muon is reconstructed by associating hits in the muon detector and a central track. It is required to be of `medium` quality and matched to a `trackmedium` (Run IIa) or `tracknewmedium` (Run IIb) quality tracks as defined in Section 3.4. At the preselection stage, muon candidates with transverse momentum  $p_T > 15$  GeV are kept and must lie within the range  $|\eta| < 2.0$ , where  $\eta$  is the detector pseudorapidity of the muon system as defined previously. This fulfills the muon trigger requirements (see Section 4.2). The muon must be isolated according to the `NPTight` quality requirement. Only one muon with the above quality is allowed in each event. Events are vetoed if they contain an additional `NPTight` isolated, loose muon with  $p_T > 12$  GeV. Such a requirement helps suppress contributions from  $Z/\gamma^* \rightarrow \mu\mu$  background processes.

|                                 | $\tau_\mu\tau_h$ Channel |                |                | $\tau_e\tau_h$ Channel |                |                |
|---------------------------------|--------------------------|----------------|----------------|------------------------|----------------|----------------|
|                                 | $\tau$ -type 1           | $\tau$ -type 2 | $\tau$ -type 3 | $\tau$ -type 1         | $\tau$ -type 2 | $\tau$ -type 3 |
| e/ $\mu$ Quality                | <b>medium</b>            |                |                | <b>top_tight</b>       |                |                |
| Number of e/ $\mu$              | Exact one                |                |                | Exact one              |                |                |
| Number of $\tau$                | Exact one                |                |                | Exact 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                     |                |                |
| $\tau$ $p_T$ (GeV) $\geq$       | 12.5                     | 12.5           | 15             | 15                     | 15             | 20             |
| $\tau$ $p_T^{trk}$ (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 7.1: The summary table of preselections.

An event is required to contain a hadronic decayed  $\tau$  candidate at a distance  $\Delta R > 0.4$  from the muon direction and with a sign of electric charge opposite to of the muon charge. The charge of the  $\tau$  candidate is the sum over the charges measured from the curvature of the tracks associated with the  $\tau$  candidate. Candidates are required to be found within  $|\eta| < 1.6$ , which is measured with respect to the center of the detector. The transverse momentum  $p_T^\tau$  of a  $\tau$  candidate must be greater than 12.5 GeV for  $\tau$ -types 1 and 2, and greater than 15 GeV for  $\tau$ -type 3. At the same time the transverse momentum of the track associated with the  $\tau$  candidate is required to be  $p_T > 7$  GeV for  $\tau$ -type 1 and  $p_T > 5$  GeV for  $\tau$ -type 2. In the case of  $\tau$ -type 3, one of the associated tracks must have  $p_T > 7$  GeV and the scalar sum of the transverse momenta of all associated tracks is required to be greater than 10 GeV. At least one of the  $\tau$ -tracks must have at least one SMT hit (this requirement is for all  $\tau$ -types). The distance along the  $z$  direction,  $\Delta z(\mu/\tau, PV)$ , between the track associated with muon or  $\tau$  lepton and the primary vertex must be less than 1 cm. Any contamination from cosmic ray muons is eliminated by requiring  $\tau$  candidates to have  $E_T^\tau/p_T^{trk} > 0.65$  (0.5) for  $\tau$ -types 1 (2, 3). In order to remove mis-identified muons with  $\tau$ -type 1 candidates, a longitudinal shape variable,  $R_\mu = (E^\tau - E_{CH}^{trk})/p_T^{trk}$ , is defined where  $E_{CH}^{trk}$  is the energy deposited around the  $\tau$ -track extrapolated to the coarse hadronic section of the calorimeter. Because muons have a distribution that peaks at much lower values than type 1  $\tau$  leptons, this search requires  $R_\mu > 0.3$  for all events.

Additionally, an energy correction for  $\tau$  candidates, derived from a pure  $Z \rightarrow$

$\tau\tau$  sample in data and based on the ratio of the calorimeter energy,  $E_T^{cal}$  to track momentum,  $p_T^{trk}$ , is applied on MC to better match the  $\tau$  energy spectrum in data. Complete details on the energy scale correction are given in Section 5.4 of this thesis. A selection is made on  $NN_\tau$  to separate real  $\tau$ s from those that originate from jet fakes as  $NN_\tau > 0.9$  for type 1 & 2,  $NN_\tau > 0.95$  for type 3.

The missing energy ( $\cancel{E}_T$ ) is corrected for the selected muon and any jets with  $p_T^{jet} > 20$  GeV, where  $p_T^{jet}$  is the transverse momentum of the jet. Jets that match  $\tau$  candidates passing the above requirements are not used in the missing energy correction to avoid using the corrections for jets on real  $\tau$  candidates.

Additionally, events are required to have no electrons with  $p_T > 12$  GeV that pass the `top_loose` requirements. This ensures this analysis 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  $W$ +jets events in association with a jet and QCD background.

Shown in Fig. 7.1 are kinematic distributions from Run IIb data at the preselection stage, including muon transverse momentum ( $p_T^\mu$ ), transverse mass ( $m_T$ ), the cosine value of the angle between muon and missing energy ( $\cos\Delta\phi(\mu, \cancel{E}_T)$ ), and the invariant mass of all tracks ( $M^{trk}$ ). All distributions show good agreement between data and predicted background. The same check is also done in Run IIa data to make sure predicted background well describes data after preselections applied. The distribution incorporates the proper MC efficiencies and  $\tau$  energy scale corrections described in Chapter 5. The QCD multi-jet events are estimated using the method described in Sec. 6.1 of this thesis.

### 7.1.2 $\tau_e\tau_h$ Channel

The preselections applied to  $\tau_e\tau_h$  channel are very similar to those to  $\tau_\mu\tau_h$  channel but overall tighter. One electron with transverse momentum  $p_T > 15$  GeV is selected if it fulfills the `top_tight` definition. This electron must lie within detector pseudo-rapidity  $|\eta| < 2.5$ , and it is mentioned as “tight electron” in the following sections. The event is required to fire any of the single electron triggers in the logical “OR” of single electron triggers described in Section 4.3 and the tight electron is 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 events, this event is rejected.

A  $\tau$  candidate is required to be at a distance  $\Delta R > 0.5$  from the electron

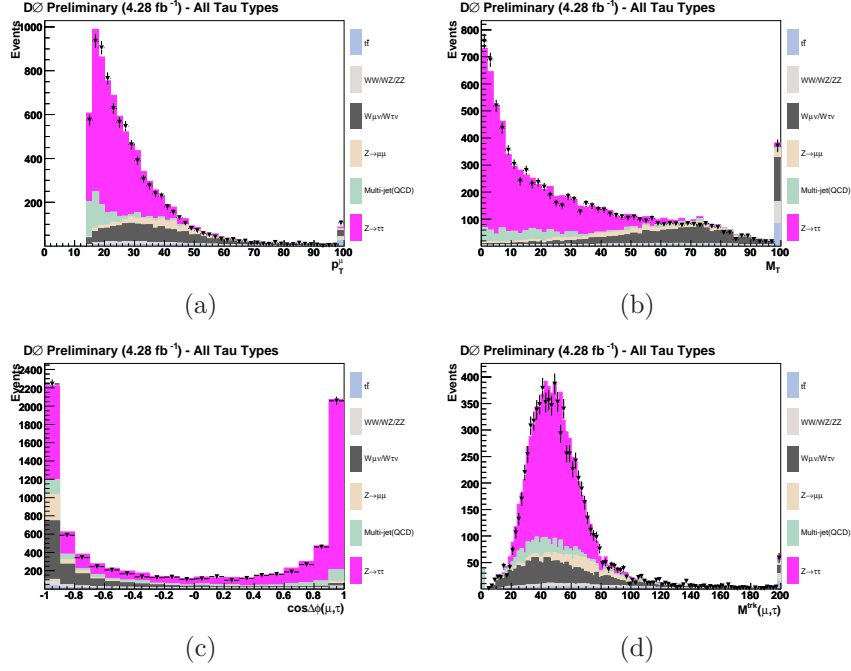


Figure 7.1: Distribution for data compared to the sum of the expected backgrounds for (a)  $p_T^\mu$  (b)  $m_T$  (c)  $\cos\Delta\phi(\mu, \cancel{E}_T)$  (d)  $M^{trk}$  at the preselection stage.

direction. The transverse momentum of the  $\tau$  candidate,  $p_T^\tau$ , as reconstructed by the calorimeter, must be greater than 15 GeV for  $\tau$ -type 1 and 2, and greater than 20 GeV for  $\tau$ -type 3. At the same time the transverse momentum of the track associated with the  $\tau$  candidate is required to be  $p_T > 15$  GeV for  $\tau$ -type 1 and  $p_T > 5$  GeV for  $\tau$ -type 2. In the case of  $\tau$ -type 3, one of the associated tracks must have  $p_T > 5$  GeV and the scalar sum of the transverse momenta of all associated tracks must be greater than 15 GeV. Type 3  $\tau$  candidates with only two associated tracks are rejected. The electric charge of the  $\tau$  candidate is the sum over the charges measured from the curvature of the tracks associated with the  $\tau$  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 good jet. The missing transverse energy  $\cancel{E}_T$  is corrected with the tight electron and jets reconstructed in cone size 0.5 with  $p_T > 20$  GeV, and these jets must not be matched to the tight electrons and  $\tau$  candidates with the requirement  $\Delta R < 0.5$ .

## 7.2 Final Selection Cuts

Final selection cuts are applied to select events used as results, and these cuts are chosen to reduce  $Z \rightarrow \ell\ell$  ( $\ell = e$  or  $\mu$ ),  $W$ +jets backgrounds as much as possible.

Since background from  $Z/\gamma^* \rightarrow \tau\tau$  is almost irreducible except for differences in its mass and spin, *i.e.*, angular distribution with respect to the Higgs boson, no attempt is made to remove contribution from  $Z/\gamma^* \rightarrow \tau\tau$ ,

### 7.2.1 $\tau_\mu\tau_h$ Channel

The background events from  $Z \rightarrow \mu\mu$  can be suppressed efficiently by applying a matching and rejecting events with  $\Delta R(\mu, \tau) < 0.5$ , therefore, only a cut on  $m_T$  is applied as an additional cut after preselection stage. In this channel, events with the transverse mass  $m_T$ , as defined in Eq. 6.1, to be less than 50 GeV for each  $\tau$ -type. This selection removes a significant portion of the  $W$ +jets background.

Table 7.2 lists the yields from data compared to the expected backgrounds for each of the three  $\tau$ -types and the sum of all types, respectively. Further, Fig. 7.2 shows the detector pseudorapidity and azimuthal  $\phi$  distributions for the tau candidates after the transverse mass cut. Figures 7.3 and 7.4 provide the missing transverse energy and the  $\tau$ -track  $p_T$  distribution for the  $\tau$  candidates after imposing the transverse mass cut, where the scalar sum of track momenta is used for  $\tau$ -type 3 candidates. All these figures are presented for Run IIb data set. Within statistics, the predicted backgrounds are consistent with the number of events in data for each of the  $\tau$ -types. The same checks for the data-background consistency are also done for the Run IIa data set, and the yields are provided in Table 7.5.

| $\tau$ -type                 | type 1       | type 2        | type 3        | all types     |
|------------------------------|--------------|---------------|---------------|---------------|
| Data                         | 744          | 4477          | 1049          | 6270          |
| Total Prediction             | $789 \pm 28$ | $4642 \pm 68$ | $1021 \pm 32$ | $6452 \pm 81$ |
| $Z \rightarrow \tau^+\tau^-$ | $567 \pm 24$ | $3741 \pm 61$ | $683 \pm 26$  | $4991 \pm 71$ |
| $Z \rightarrow \mu^+\mu^-$   | $52 \pm 7$   | $194 \pm 14$  | $18 \pm 4$    | $264 \pm 16$  |
| $W$ +jets                    | $47 \pm 7$   | $220 \pm 15$  | $90 \pm 9$    | $357 \pm 19$  |
| $WW/WZ/ZZ$                   | $7 \pm 2$    | $88 \pm 10$   | $3 \pm 1$     | $98 \pm 10$   |
| $t\bar{t}$                   | $3 \pm 1$    | $47 \pm 7$    | $2 \pm 1$     | $52 \pm 7$    |
| Multi-jet                    | $121 \pm 11$ | $351 \pm 19$  | $224 \pm 16$  | $696 \pm 27$  |

Table 7.2: Prediction from MC and QCD backgrounds compared to the observation in Run IIb data for each  $\tau$ -type and sum of all types. Only statistical errors on the MC and QCD are given.

| $\tau$ -type                  | type 1       | type 2        | type 3       | all types     |
|-------------------------------|--------------|---------------|--------------|---------------|
| Data                          | 357          | 1746          | 482          | 2585          |
| Total Prediction              | $330 \pm 18$ | $1756 \pm 42$ | $484 \pm 22$ | $2576 \pm 51$ |
| $Z \rightarrow \tau^+ \tau^-$ | $228 \pm 15$ | $1429 \pm 38$ | $322 \pm 18$ | $1979 \pm 45$ |
| $Z \rightarrow \mu^+ \mu^-$   | $15 \pm 4$   | $71 \pm 8$    | $11 \pm 3$   | $97 \pm 10$   |
| $W$ +jets                     | $12 \pm 4$   | $65 \pm 8$    | $39 \pm 6$   | $116 \pm 11$  |
| $WW/WZ/ZZ$                    | $2 \pm 1$    | $31 \pm 6$    | $2 \pm 1$    | $35 \pm 6$    |
| $t\bar{t}$                    | $0 \pm 0$    | $1 \pm 1$     | $0 \pm 0$    | $1 \pm 1$     |
| Multi-jet                     | $73 \pm 9$   | $159 \pm 12$  | $110 \pm 11$ | $342 \pm 18$  |

Table 7.3: Prediction from MC and QCD backgrounds compared to the observation in Run IIa data for each  $\tau$ -type and sum of all types. Only statistical errors on the MC and QCD are given.

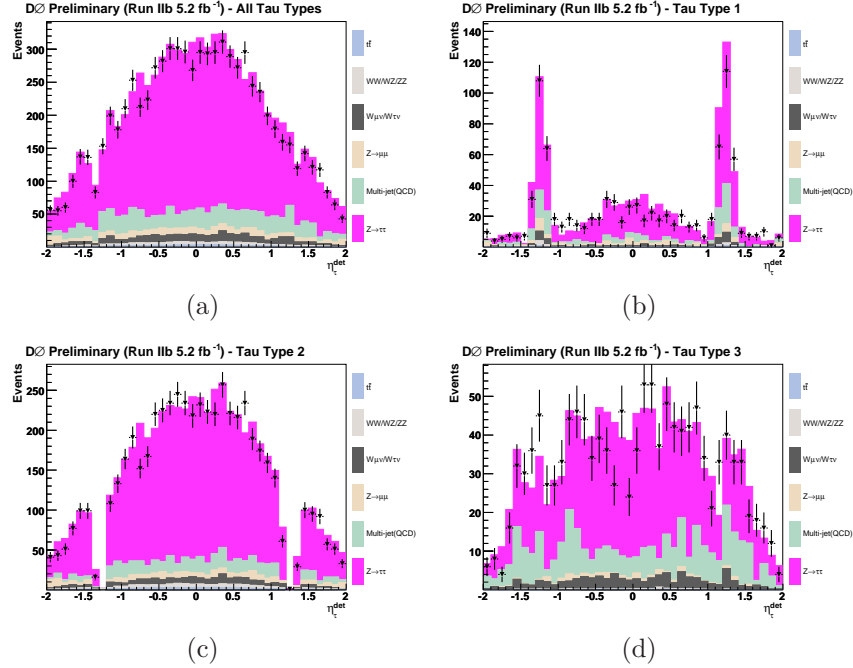


Figure 7.2: Distribution in Run IIb  $\tau_\mu \tau_h$  channel for data compared to the sum of the expected backgrounds for tau detector pseudorapidity for (a) sum of all  $\tau$ -types, (b)  $\tau$ -type 1, (c)  $\tau$ -type 2 and (d)  $\tau$ -type 3 candidates after the transverse mass cut.

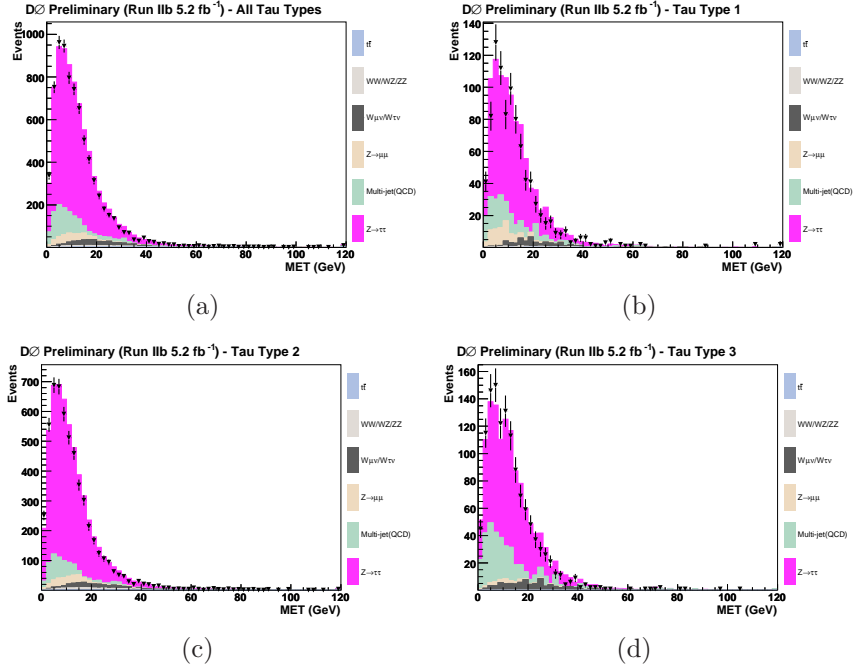


Figure 7.3: Distribution in Run IIb  $\tau_\mu\tau_h$  channel for data compared to the sum of the expected backgrounds for  $\cancel{E}_T$  for (a) sum of all  $\tau$ -types, (b)  $\tau$ -type 1, (c)  $\tau$ -type 2 and (d)  $\tau$ -type 3 candidates after the transverse mass cut.

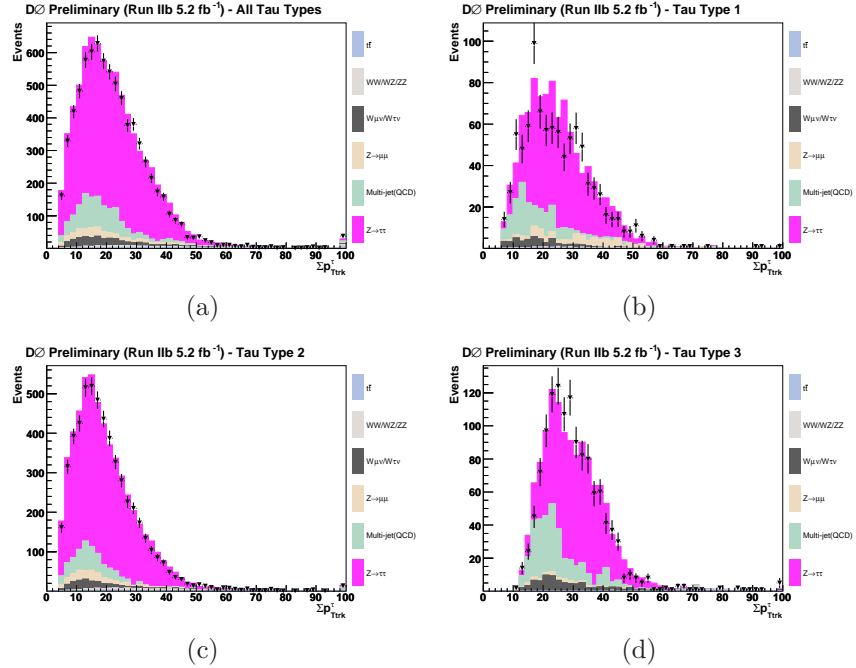


Figure 7.4: Distribution in Run IIb  $\tau_\mu\tau_h$  channel for data compared to the sum of the expected backgrounds for tau track  $p_T$  for (a) sum of all  $\tau$ -types, (b)  $\tau$ -type 1, (c)  $\tau$ -type 2 and (d)  $\tau$ -type 3 candidates after the transverse mass cut. For  $\tau$ -type 3 candidates, the scalar sum of the transverse momentum of the three tracks is used.

### 7.2.2 $\tau_e\tau_h$ Channel

The distributions show reasonable agreement between data and the expected backgrounds at the preselection stage, a series cuts are applied to further suppress  $Z \rightarrow e^+e^-$  and  $W + \text{jets}$  background.

Since there is no EM coverage in the ICD region, falling electrons with reconstructed signature of tracks and hadronic clusters can easily be mis-identified as  $\tau$  leptons. Therefore,  $\tau$  candidates in the region  $1.05 < |\eta| < 1.55$  are not considered, and this cut applied to type 1 and 2. Also, it is found that the efficiency of detecting electrons near the boundary of modules of the EM calorimeter is not as good as of the other parts of the calorimeter. This increase the amount of electrons faking  $\tau$  candidates in these regions. The variable  $f_{tPhiPS}$  is defined as:

$$f_{tPhiPS} = \text{mod} \left( \frac{\phi_{\tau\text{trk}}^{PS}}{\left(\frac{2\pi}{32}\right)} \right) \quad (7.1)$$

where  $\phi_{\tau\text{trk}}^{PS}$  is the azimuthal angle of the track associated with the  $\tau$  at the pre-shower detector. Values of this variable fall in the range from 0 to 1, and events peak at both ends are found close to the module boundary of the detector. Type 2  $\tau$  candidates are required to have  $0.1 < f_{tPhiPS} < 0.9$  to avoid collecting electrons in inefficient parts of the detector. As described in Section 3.4.1, electrons should have low  $NN_{elec}$  values peak close to 0, so  $\tau$ -type 2 candidates are required to have  $NN_{elec} < 0.9$ . With these three selection cuts, ICD,  $f_{tPhiPS}$  and  $NN_{elec}$ , background events from  $Z \rightarrow e^+e^-$  can be efficiently suppressed. Since electrons are expected to deposit most of the energy in EM layers, the final electron veto cut 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. This cut is only applied to  $\tau$ -type 2 and 3 because  $\tau$ -type 1 events have no associated EM clusters.

Figure 7.5 shows the distribution of transverse momentum ( $p_T^\tau$ ), detector pseudo-rapidity ( $\eta^\tau$ ), and the ratio of energy to sum of the transverse momentum of tracks ( $E/p_T^{\text{trk}}$ ) of the  $\tau$  candidates. Similarly, Figure 7.6 shows the distribution of transverse momentum ( $p_T^e$ ) and detector pseudo-rapidity ( $\eta^e$ ) of electrons. After above anti-electron cuts applied,  $Z \rightarrow e^+e^-$  (red) now contribute little, and the rest events are driven by Multi-jet (green),  $W$ +jets (grey) and  $Z \rightarrow \tau^+\tau^-$  (pink).

After vetoing events contain electrons faking  $\tau$  objects, a selection cut is made on the transverse mass to reject much of the remaining  $W$ +jets background by



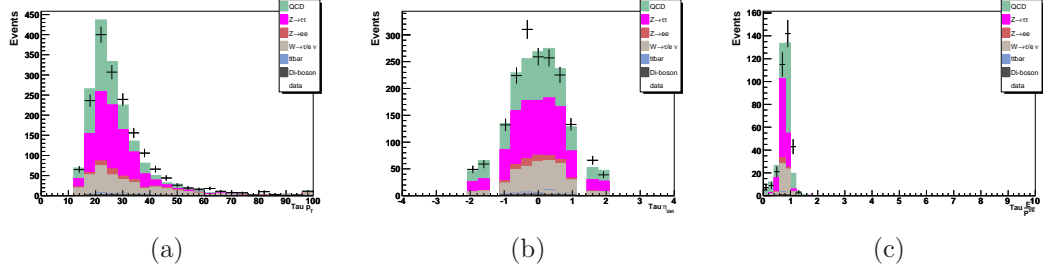


Figure 7.5: Distribution in Run IIa  $\tau_e \tau_h$  channel for data compared to the sum of the expected backgrounds for (a)  $p_T^\tau$  (b)  $\eta^\tau$  (c)  $E/p_T^{trk}$  for sum of all  $\tau$ -type events after the anti  $Z \rightarrow e^+e^-$  cuts.

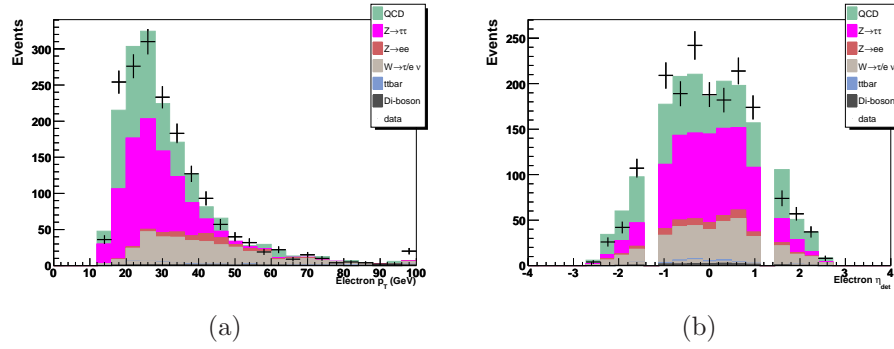


Figure 7.6: Distribution in Run IIa  $\tau_e \tau_h$  channel for data compared to the sum of the expected backgrounds for (a) electron  $p_T$  and (b) electron  $\eta_{det}$  after the anti  $Z \rightarrow e^+e^-$  cuts.

a lot requiring  $M_T < 50$  GeV for all three  $\tau$ -types. In addition, Figure 7.7 shows the azimuthal angle between the electron and  $\tau$ ,  $\Delta\phi(e, \tau)$ , after the transverse mass cut. The signal peaks strongly at  $\pi$ , whereas the multi-jet background is less strongly peaked. Events with  $\Delta\phi(e, \tau) < 1.6$  are removed to reject some of the rest multi-jet background without significantly reducing the signal efficiency. Finally, a cut is placed in the 2D  $\Delta\phi(e, \cancel{E}_T) - \Delta\phi(\tau, \cancel{E}_T)$  plane to further remove  $W$ +jets events. The distributions for  $W$ +jets events and signal events in this plane, along with the selection used, are shown in Figure 7.8. Events appear above the line will be rejected. Table 7.4 summaries these final selections applied in this channel.

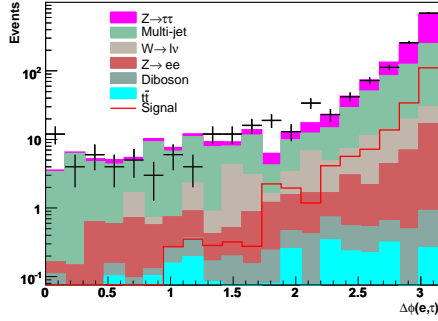


Figure 7.7: in Run IIa  $\tau_e \tau_h$  channel for data compared to the sum of the expected backgrounds for the azimuthal angle between the electron and tau, after the transverse mass cut.

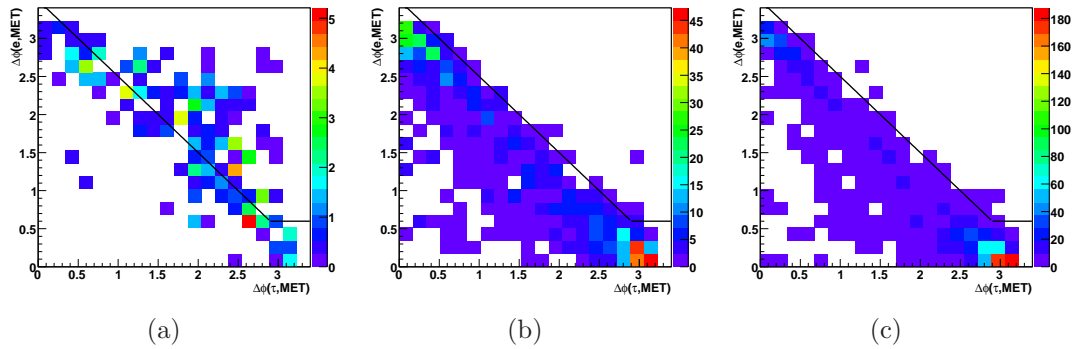


Figure 7.8: Distribution of events in the  $\Delta\phi(e, \cancel{E}_T) - \Delta\phi(\tau, \cancel{E}_T)$  plane for (a)  $W$  MC events, (b)  $m_h = 120$  GeV signal MC events and (c)  $m_h = 200$  GeV signal MC events. Events above the black line are rejected.

| Selection Cut  | Reduced Source         | $\tau$ -type 1                     | $\tau$ -type 2          | $\tau$ -type 3 |
|--|------------------------|------------------------------------|-------------------------|----------------|
| Removal of ICD   | $Z \rightarrow e^+e^-$ | $ \eta  > 1.55$ OR $ \eta  < 1.05$ | –                       | –              |
| Crack Removal  | $Z \rightarrow e^+e^-$ | –                                  | $0.1 < f_{tPhPS} < 0.9$ | –              |
| $NN_e \geq$  | $Z \rightarrow e^+e^-$ | –                                  | 0.9                     | –              |
| $f_{em} \leq$  | $Z \rightarrow e^+e^-$ | –                                  | 0.9                     | 0.9            |
| $m_T \leq$ (GeV)   | $W$ +jets              |                                    | 50                      |                |
| $\Delta\phi(e, \tau) \leq$                                     | Multi-jet              |                                    | 1.6                     |                |
| $\Delta\phi(e, \cancel{E}_T) - \Delta\phi(\tau, \cancel{E}_T)$ | $W$ +jets              | As shown in Figure 7.8             |                         |                |

Table 7.4: Table listing the selections which the  $\tau$  candidates are required to pass in order to suppress  $Z \rightarrow e^+e^-$  and  $W + \text{jets}$  background.

| $\tau$ -type                 | type 1       | type 2       | type 3       | all types    |
|------------------------------|--------------|--------------|--------------|--------------|
| Data                         | 192          | 490          | 352          | 1034         |
| Total Prediction             | $192 \pm 10$ | $453 \pm 13$ | $344 \pm 14$ | $989 \pm 22$ |
| $Z \rightarrow \tau^+\tau^-$ | $100 \pm 2$  | $305 \pm 4$  | $176 \pm 3$  | $581 \pm 5$  |
| $Z \rightarrow e^+e^-$       | $5 \pm 2$    | $20 \pm 2$   | $6 \pm 1$    | $31 \pm 2$   |
| $W$ +jets                    | $7 \pm 2$    | $12 \pm 3$   | $22 \pm 4$   | $42 \pm 5$   |
| $WW/WZ/ZZ$                   | $0 \pm 0$    | $1 \pm 1$    | $1 \pm 1$    | $2 \pm 1$    |
| $t\bar{t}$                   | $0 \pm 0$    | $0 \pm 0$    | $1 \pm 1$    | $1 \pm 1$    |
| Multi-jet                    | $80 \pm 10$  | $115 \pm 12$ | $138 \pm 13$ | $332 \pm w0$ |

Table 7.5: Prediction from MC and QCD backgrounds compared to the observation in data for each  $\tau$ -type and sum of all types for Run IIa data of  $\tau_e\tau_h$  channel. Only statistical errors on the MC and multi-jet are given.

# Chapter 8

## Results

### 8.1 Final Discriminant

The visible mass,  $M_{vis}$ , variable defined as:

$$M_{vis} = \sqrt{(P_{\tau_1} + P_{\tau_2} + \cancel{E}_T)^2}, \quad (8.1)$$

is used to search for the signal in the data sample. It is calculated using the four vectors of the visible  $\tau$  decay products  $P_{\tau_{1,2}}$  and of the missing momentum  $\cancel{P}_T = (\cancel{E}_T, \cancel{E}_x, \cancel{E}_y, 0)$ .  $\cancel{E}_x$  and  $\cancel{E}_y$  indicate the components of  $\cancel{E}_T$ . This variable is meant to include the contributions from neutrinos and to provide a better way to reconstruct full masses. For  $\tau$ -type 1, which by definition contain charged single pion decays, one can make more precise measurement for the visible mass for the lower mass backgrounds (lower compares to the signal) by using the track momentum. However, it is found that calorimeter energy can provide a much better confinement on the higher mass signals than the track momentum. Therefore, the energy of  $\tau$  candidates will be replaced by tracking momentum if the calorimeter energy is less than 50 GeV ( $p_T^{cal} < 50 \text{ GeV}$ ). This replaced energy is used for final  $M_{vis}$  calculation and the  $\cancel{E}_T$  re-computation which also affects the value of  $M_{vis}$ . For  $\tau$ -types 2 and 3, where a contribution from neutral pions can exist, the calorimeter energy is used without any replacement.

The visible mass for events passing all final selections in Run IIa  $\tau_\mu\tau_h$  channel is shown in Fig. 8.1 and 8.2, and Fig. 8.3 and 8.4 for Run IIa  $\tau_\mu\tau_h$  channel. Also shown in the figure is the predicted Higgs signal has been multiplied by a factor of 50. Similarly,  $M_{vis}$  distribution in Run IIa  $\tau_e\tau_h$  channel on linear scale is shown in Fig. 8.5. There is no significant excess in signal over background is observed,

and data matches the background prediction.

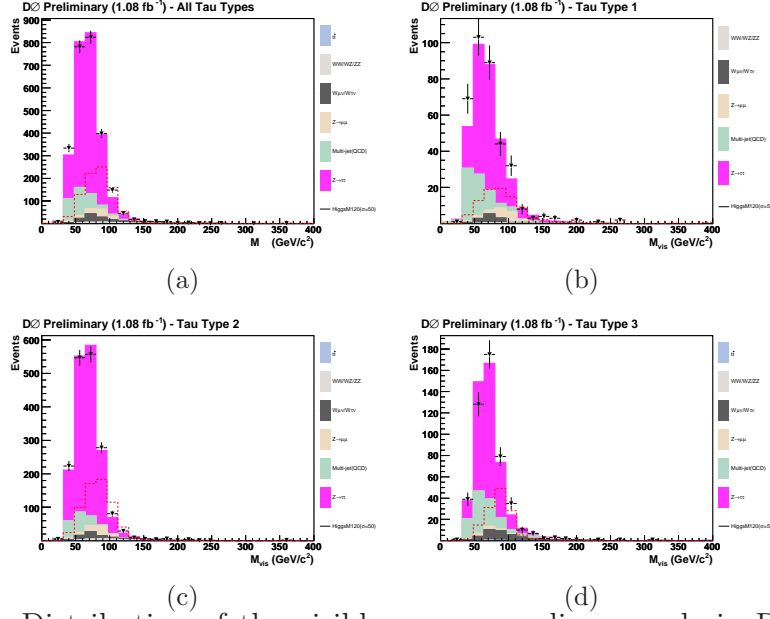


Figure 8.1: Distribution of the visible mass on a linear scale in Run IIa  $\tau_\mu\tau_h$  channel for (a) all types (b)  $\tau$ -type 1 (c)  $\tau$ -type 2 and (d)  $\tau$ -type 3. The data, shown with error bars, is compared to the sum of the expected backgrounds.

## 8.2 Cross Section Limits

Since no significant excess in signal over background is observed, limits on the production cross section times branching ratio for a neutral Higgs decaying to  $\tau^+\tau^-$  pairs is calculated. Specifically, the visible mass variable is used as the input to the limit calculator.

### 8.2.1 Limit Setting Technique and Collie

Limit of the production cross section times branching ratio is calculated using the **Collie** (Confidence Level Limit Evaluator) package [70] based on  $CL_s$  method [71] with ‘as-Frequentist-as-possible’ treatment on all calculation. It accepts ROOT histograms separately for data and for the background and the signal models, and uses a negative Poisson Log-Likelihood Ratio (LLR) test statistic to evaluate statistical significance. There are two hypotheses defined, the signal-plus-background (S+B) hypothesis and the background-only (B-Only) hypothesis. For example, in this analysis, ‘S+B’ corresponds to MSSM hypothesis for the

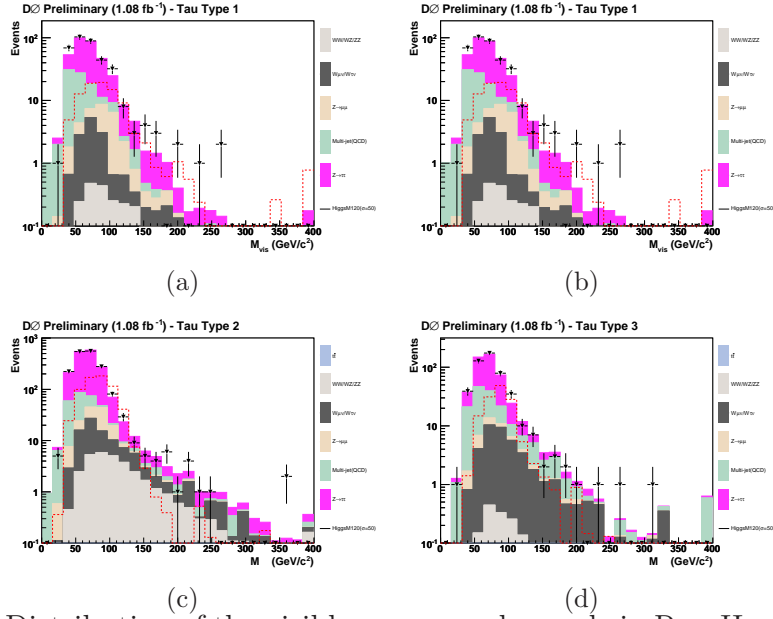


Figure 8.2: Distribution of the visible mass on a log scale in Run IIa  $\tau_\mu \tau_h$  channel for (a) all types (b)  $\tau$ -type 1 (c)  $\tau$ -type 2 and (d)  $\tau$ -type 3. The data, shown with error bars, is compared to the sum of the expected backgrounds.

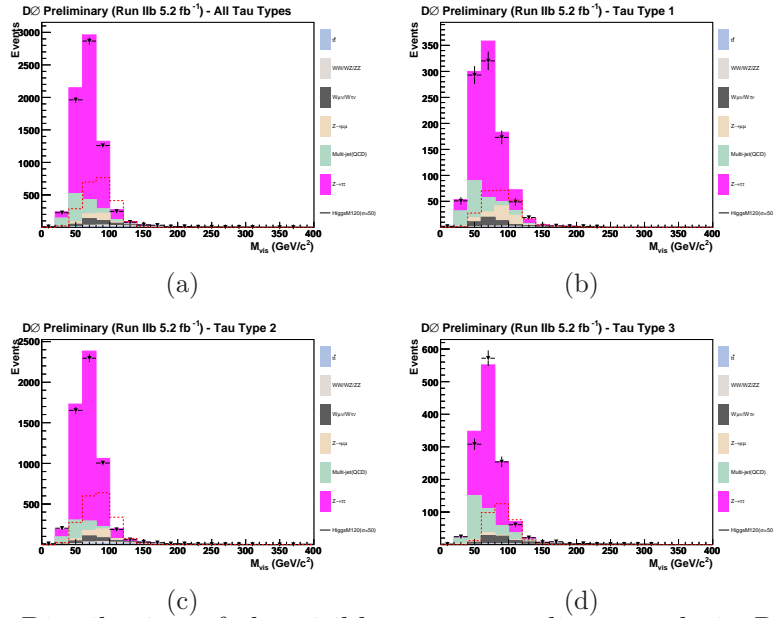


Figure 8.3: Distribution of the visible mass on a linear scale in Run IIb  $\tau_\mu \tau_h$  channel for (a) all types (b)  $\tau$ -type 1 (c)  $\tau$ -type 2 and (d)  $\tau$ -type 3. The data, shown with error bars, is compared to the sum of the expected backgrounds.

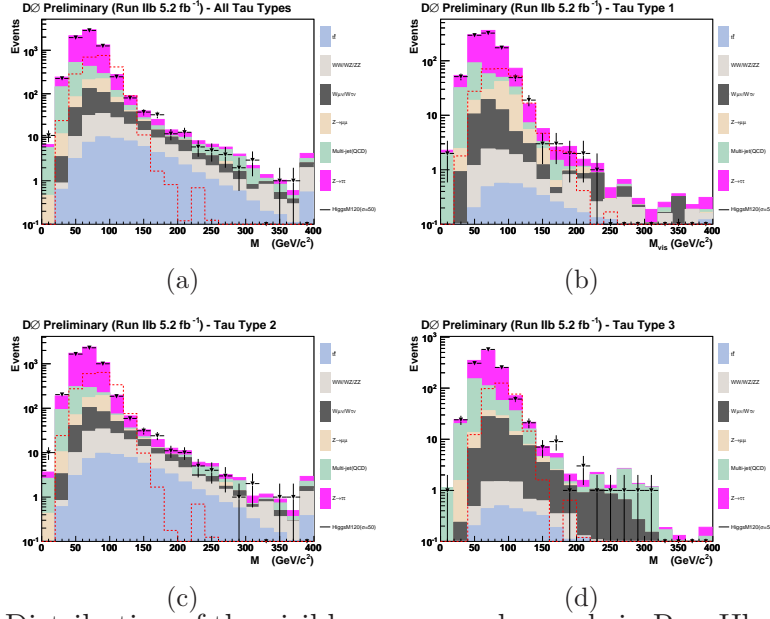


Figure 8.4: Distribution of the visible mass on a log scale in Run IIb  $\tau_\mu\tau_h$  channel for (a) all types (b)  $\tau$ -type 1 (c)  $\tau$ -type 2 and (d)  $\tau$ -type 3. The data, shown with error bars, is compared to the sum of the expected backgrounds.

Higgs bosons while ‘B-Only’ points to the SM background prediction. For each bin of the given histogram, the LLR is evaluated by

$$\text{LLR} = -2 \sum_i^{\text{channel}} \sum_j^{\text{bin}} \left( s_{ij} - d_{ij} \ln \left( 1 + \frac{s_{ij}}{b_{ij}} \right) \right), \quad (8.2)$$

where  $i$  is the  $i^{\text{th}}$  channel analyzed,  $j$  is the  $j^{\text{th}}$  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 [70]. This process is repeated many times (about 10k) for both ‘S+B’ and ‘B-Only’ hypothesis, by generating data 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, a LLR value is evaluated, and the frequency of relative outcomes defines confidence levels. Figure 8.6 shows an example of LLR distributions for ‘S+B’ (red) and ‘B-Only’ (blue) hypothesis, and the observed LLR evaluated using data is also shown in black line. Give the outcomes of LLR, one can calculated a confidence level for ‘S+B’ ( $\text{CL}_{\text{s+b}}$ ) and for ‘B-Only’ ( $\text{CL}_{\text{b}}$ ) by integrating the shaded areas of the distributions as shown in the figures. The value of the signal confidence

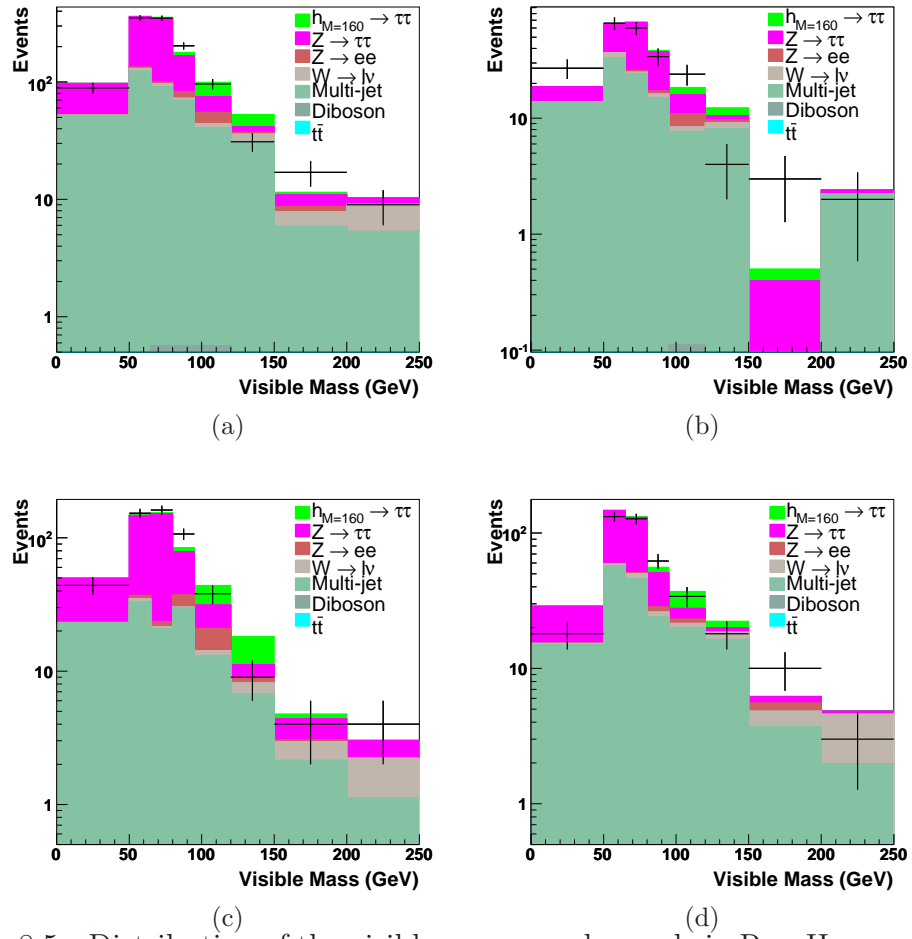


Figure 8.5: Distribution of the visible mass on a log scale in Run IIa  $\tau_e \tau_h$  channel for (a) all types (b)  $\tau$ -type 1 (c)  $\tau$ -type 2 and (d)  $\tau$ -type 3. The data, shown with error bars, is compared to the sum of the expected backgrounds.



level,  $CL_s$ , is then defined as

$$CL_s = \frac{CL_{s+b}}{CL_b} \quad [70]. \quad (8.3)$$

and the exclusion condition is made such that  $CL_s < \alpha$ . The desired value of  $\alpha$  used in this analysis 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 times branching ratio of the predicted signal. Therefore, the observed limit is calculated by varying the cross-section times branching ratio of the signal until 95% CL is reached. The expected limit is calculated in the same way by assuming the data is all background-like by setting  $d_{ij} = b_{ij}$ ; in the other word, the median outcome from ‘B-Only’ hypotheses is taken as the ‘expected data’.

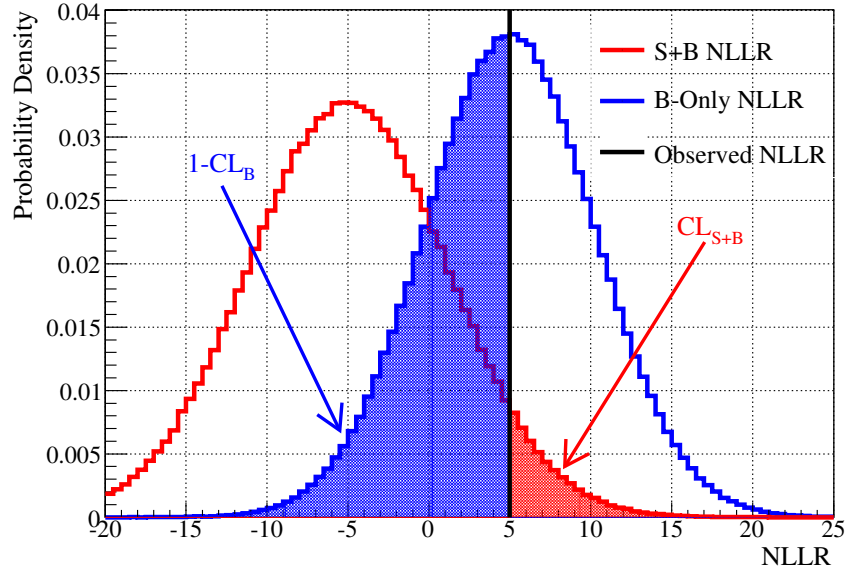


Figure 8.6: LLR distributions for the expectation in the background-only hypothesis (blue line) and in the signal-plus-background hypothesis (red line) and the observation in the data (blue line). Also shown are the regions integrated to obtain  $CL_b$  and  $CL_{s+b}$  [70].

The sensitivity of the analysis is determined by both the peak and the width of LLR distributions. Without considering any systematics, the width of LLR distributions are determined by Poisson uncertainty of the samples. Systematic uncertainties are treated by allowing events in each bin of a pseudo-experiment to vary according a random number taken from a Gaussian distribution with a

mean of zero and width of the size of the systematic uncertainty. This widens the LLR distributions and makes the width proportional to the quadrature sum of Poisson and systematic uncertainties. Figure 8.7 shows an example of LLR distributions before (left) and after applying the Gaussian smearing of the systematics (right). Means and the observed values of LLR remain the same, but the widths increase after applying the smearing. This effect makes the ability to distinguish two hypothesis decrease and therefore degrades the outcomes of limit calculation.

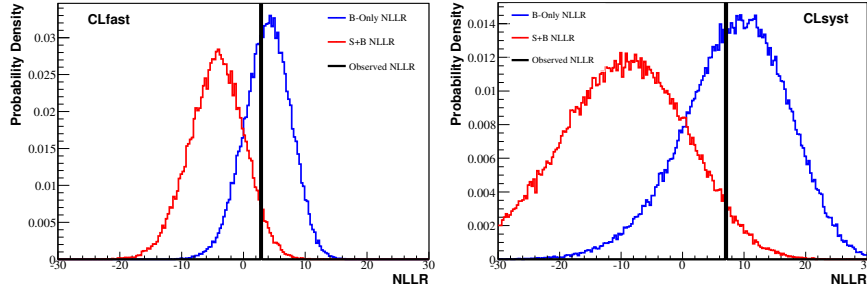


Figure 8.7: LLR distributions for ‘B-Only’ (blue) and ‘B+S’ (red) hypothesis. The left figure shows LLR for the case of no systematics and the right one shows LLR after applying Gaussian smearing of the systematic uncertainties [70].

In order to reduce the degrading effects of systematics, a maximum likelihood fit of total prediction to data on the final variable distribution is introduced in **Collie**. The value of Poisson  $\chi^2$  is determined by the numbers of data ( $d_i$ ) and total predicted ( $p_i$ ) events by

$$\chi^2 = 2 \sum_b \ln \left( (p_i - d_i) - d_i \ln \frac{p_i}{d_i} \right) + \sum_k S_k^2 [70], \quad (8.4)$$

where  $S_k^2$  is the squared deviation of the uncertainty from nominal in units normalized by the  $\pm 1$  standard deviation. Each background source is allowed to vary independently according to uncertainties to minimize the  $\chi^2$  value to achieve the ‘best fit’. More details about the techniques used in **Collie** can be found in [70].

### 8.2.2 Systematic Uncertainties

Various sources of systematic error affect both the signal efficiency and background estimations. The systematics from each source are summarized in Table 8.1. Listed below are descriptions of all sources.

- The uncertainty on the efficiencies for muon ID are due to the tracking reconstruction requirement (1.4%), muon isolation (0.9%), and the identification method (1.2%) and assigned only for  $\tau_\mu\tau_h$  channels. These are all taken from the discussion in Ref. [43] and added in quadrature to yield an overall 2.1% uncertainty for muon ID.
- The uncertainty on the electron identification is 3.3% [38], and the uncertainty on the  $Z \rightarrow e^+e^-$  background re-weighting is 13% derived from the efficiency correction factors as described in Sect. 5.2. These two uncertainties are negligible in  $\tau_\mu\tau_h$  channels and therefore assigned only for Run IIa  $\tau_e\tau_h$  channel.
- Since tracks from a tau candidate are treated similar to those of muon tracks, an additional uncertainty on the tau track must be assigned. This is taken at 1.4% as described in Ref. [72].
- The uncertainty on the luminosity is taken at 6.1% [73].
- Varying the Jet Energy Scale (JES) by  $\pm 1\sigma$  on the signal and background shapes. During limit calculation, this uncertainty modifies the shape of the visible mass distribution and therefore, it is parameterized as a function of  $M_{vis}$ . Hence in Table 8.1, it is listed as “shape”. This uncertainty is found to be  $< 1\%$ , and therefore is negligible in  $\tau_e\tau_h$  channel.
- The uncertainty on the theoretical  $Z$  cross section considered for the background contribution from  $Z/\gamma^*$  is taken to be 5% in order to account for scale and PDF uncertainties.
- As described in Secs. 6.1, two methods are used to determine the multi-jet background contribution where a jet fakes a hadronic decayed  $\tau$ . In order to assign the systematic uncertainty in the method, the difference between the primary method and the cross-check method is taken per  $\tau$  type.
- The uncertainty on the trigger efficiency is estimated by varying the trigger probabilities of the single muon or electron triggers by  $\pm 1\sigma$  [75]. During limit calculation, this uncertainty modifies the shape of the visible mass distribution and therefore, it is parameterized as a function of  $M_{vis}$ . Hence in Table 8.1, it is also listed as “shape”.

- The error on the signal acceptance due to PDF uncertainties was 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, where each set contains the  $\pm 1\sigma$  variations. 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 shown in Eq. 8.5 [74]. A 4.6% uncertainty is subsequently assigned for the PDF variation.

$$\Delta x^\pm = \sqrt{\sum_{i=0}^{20} (x_0 - x_i^\pm)^2} \quad (8.5)$$

- As described in Sect. 5.4, the systematics for the  $\tau$  energy scale is determined by varying the relative hadronic  $\tau$  energy scale by the uncertainties given in Ref [68], which depend on  $\tau$ -type. Since the scale depends on the shape of the visible mass distribution, the uncertainty is taken as shape-dependent.
- Next, the uncertainty on the overall  $\tau$  ID efficiency was estimated by selecting  $\tau$ s from  $Z \rightarrow \tau_\mu \tau_h$  decays. The study was performed and certified by the  $\tau$ -ID Algorithm group [68]. Systematic contributions arising from both multi-jet estimation in the dataset and the kinematic selections for the hadronic  $\tau$  are considered. The systematics per  $\tau$ -type is derived as a function of  $NN_\tau$  cut and for the selections given in Chapter 7 is assigned a 8% for  $\tau$ -type 1, 4% for type 2 and 5% for type 3 for Run IIa analysis. Since the  $NN_\tau$  training samples are different in Run IIa and Run IIb, a different set errors of 12%, 4.2%, and 7% for  $\tau$ -types 1, 2, and 3, respectively, are assigned for Run IIb analysis. This uncertainty is in addition to the one arising from the  $\tau$  track efficiency quoted earlier.

The entire series of systematics listed above are also assumed to apply for the Higgs signal.

### 8.2.3 $\tau_e \tau_h$ Channel with $1.08 \text{ fb}^{-1}$ Data

In this channel, the results of cross-section limits are computed by two modules provided by `Collie`:

**CLfit1:** A 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 systematics in both the signal plus background and background only hypotheses.

Figure 8.8(a) shows the LLR distributions expected for the ‘B-Only’ hypothesis, the ‘S+B’ hypothesis and the observation in the data using **CLfit2**. Also shown are the one (blue) and two (green) sigma variations for the expectation from the background only hypothesis. A deviation beyond the two sigma background LLR contour is observed around  $m_h = 130$  GeV. However, this is less than a 3 sigma departure from the background only hypothesis, as can be seen by the observed  $CL_b$  curve in Fig. 8.8(b). Figure 8.9 shows the expected and observed limits on the cross-section times branching ratio for  $\phi \rightarrow \tau_e \tau_h$  as a function of the test Higgs boson mass, from both **CLfit** and **CLfit2** modules. The observed limits evaluated by two fit methods follow the prediction within  $2\sigma$  standard deviation.

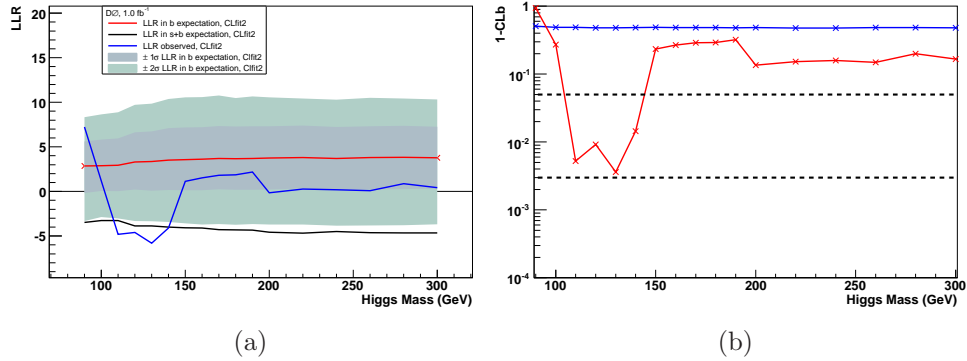


Figure 8.8: Outcomes by **CLfit2** module for (a) LLR for the data, the expectation from the ‘B-Only’ and the ‘S+B’ hypothesis, (b) CL in the ‘B-Only’ hypothesis ( $CL_b$ ) as a function of the test Higgs boson mass. The blue line shows the expectation in the ‘B-Only’ hypothesis and the red line shows the observation in the data. The dotted lines correspond to 0.05 ( $2\sigma$ ) and 0.003 ( $3\sigma$ ).

#### 8.2.4 $\tau_\mu \tau_h$ Channel with $5.36 \text{ fb}^{-1}$ Data

A fast approximation to **CLfit2** module is used to calculate the cross-section limits of this channel. It provides precise results within 2 – 3% difference from **CLfit2** but takes only 1/10 of running time [70]. The LLR distribution of the

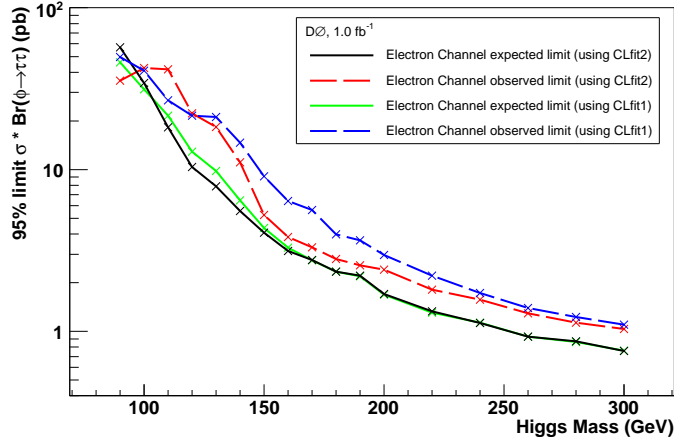


Figure 8.9: Expected and observed limits for the cross-section times branching ratio for the electron channel. The limit has been calculated with CLfit and CLfit2.

combined  $5.36 \text{ fb}^{-1}$  data of  $\tau_\mu\tau_h$  channel is shown in Fig. 8.10 with one and two sigma variations in green and yellow bands, respectively. The search results shown that data follows the ‘B-Only’ expectation very well. Figure 8.11 shows the expected and observed limits on the cross-section times branching ratio for  $\phi \rightarrow \tau_\mu\tau_h$  as a function of the test Higgs boson mass calculated by the fast approximation module in Collie. Across the whole scanned mass region, there is no deviation more than 2 sigma observed.

### 8.3 Translation into the MSSM

In this analysis, the production cross sections, widths and branching ratios are determined using FEYNHIGGS 2.6 [30] to set MSSM exclusion limits. It takes the theoretical cross section  $\sigma(p\bar{p} \rightarrow \phi) \times Br(\phi \rightarrow \tau\tau)$  at each  $M_A$  mass point, then increases the value of  $\tan\beta$  until the cross section from FeynHiggs matches the observed or expected limit discussed in the previous section. This value of  $\tan\beta$  for that value of  $M_A$  represents the observed or expected limit, respectively. Figures 8.12 and 8.13 show the combined expected and observed exclusions in the  $m_h^{\max}$  and *no-mixing* scenarios for  $\mu = \pm 200 \text{ GeV}$  using data from the searches in DØ Run IIa data set including  $\tau_e\tau_h$ ,  $\tau_e\tau_\mu$  and  $\tau_\mu\tau_h$  channels. This result has been published in PRL in 2008 [26]. Similar plots for the results in only  $\tau_\mu\tau_h$  channel with DØ Run IIa + Run IIb data are shown in Fig. 8.14 and 8.15. These

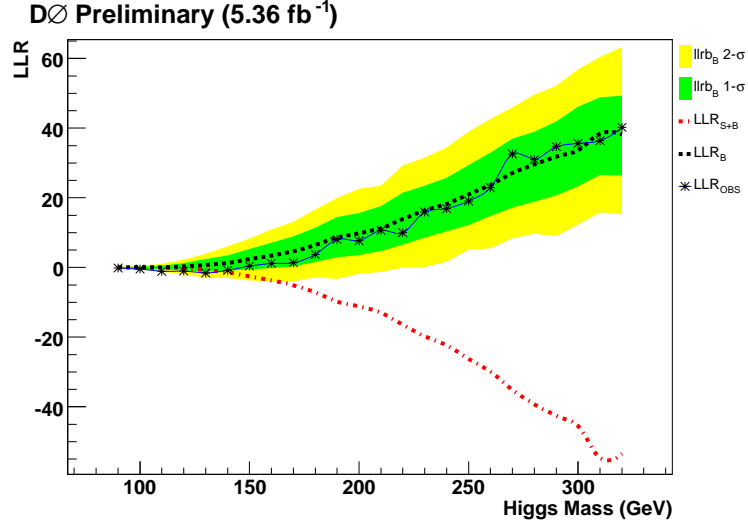


Figure 8.10: LLR distribution for data (blue line) and the expectations. Black and red lines represent the results from ‘B-Only’ and ‘S+B’ hypothesis, respectively. Also shown the green and yellow bands are one and two sigma standard deviation from the ‘B-Only’ expected values.

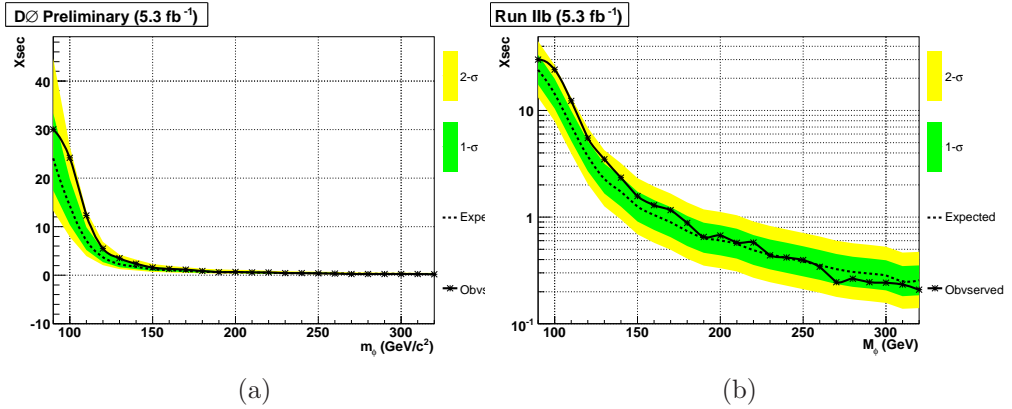


Figure 8.11: Cross-section limits evaluated by the fast approximation to CLfit2 module of  $\tau_\mu\tau_h$  channel for  $5.36\text{ fb}^{-1}$  data in linear scale (a) and in log scale (b). The green and yellow bands are one and two sigma standard deviation from the ‘B-Only’ expected limits.

exclusions assume the Higgs boson widths predicted in MSSM with low  $\tan\beta$  values are very similar to the widths of SM Higgs bosons. Discussions of the results comparing to LEP result and other Tevatron MSSM Higgs boson searches can be found in the next Sect. 8.4.

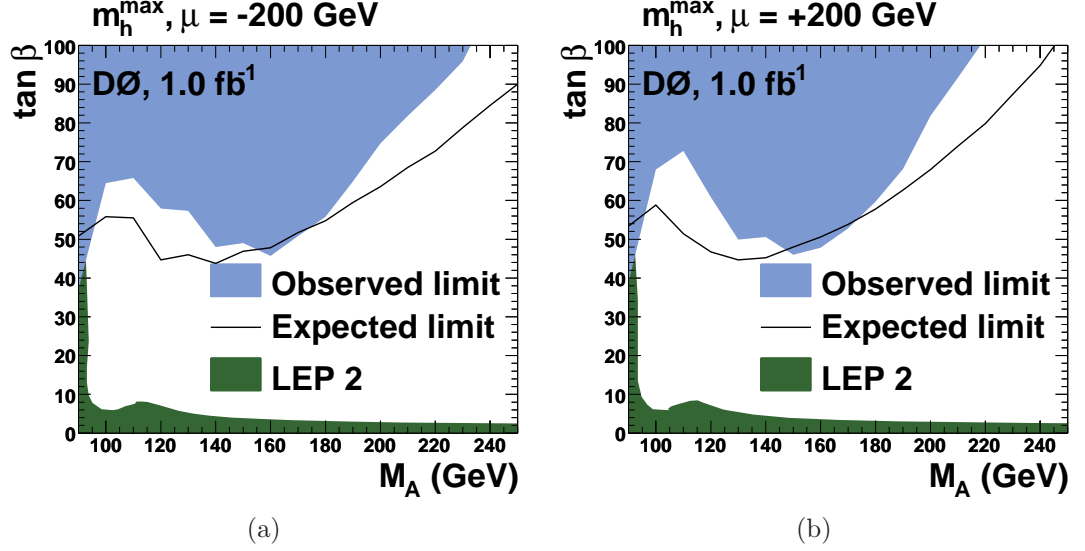


Figure 8.12: Expected and observed exclusion in  $\tau_e\tau_h$  channel for the  $m_h^{\max}$  scenario in the  $(\tan\beta - M_A)$  plane, for a)  $\mu = -200$  GeV and b)  $\mu = +200$  GeV.

## 8.4 Comparison with Previous Results and Other Searches

The LEP experiments performed searches for MSSM Higgs bosons mainly in Higgs-strahlung,  $e^+e^- \rightarrow Zh, ZH$ , channels. No signal above the expected backgrounds was found [25], therefore the results excluded  $M_A$  below approximately 93.4 GeV, and bound  $m_h > 92.8$  GeV for higher values of  $\tan\beta$  as shown in the green shaded area in Fig. 8.14 and 8.15. Low  $\tan\beta$  region with  $0.7 < \tan\beta < 2.0$  were excluded by fixing  $m_t = 174.3$  GeV for all  $M_A$  based on the fact that the lightest Higgs boson  $h$  with such low  $\tan\beta$  values should be SM-like such that it would have been observed in the LEP searches [25].

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. The most promising channels studied at



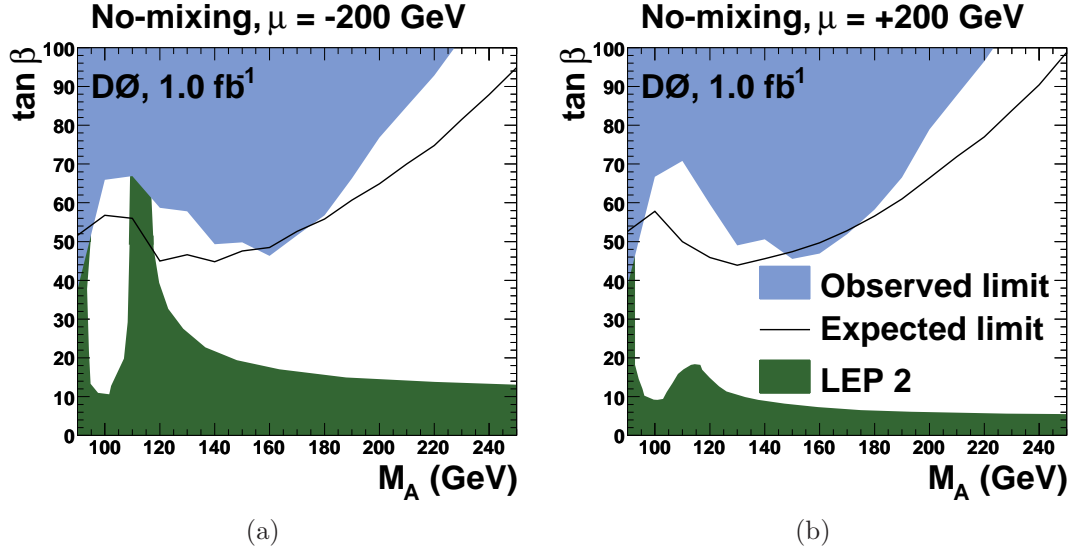


Figure 8.13: Expected and observed exclusion  $\tau_e \tau_h$  channel for the *no-mixing* scenario in the  $(\tan \beta - M_A)$  plane, for a)  $\mu = -200$  GeV and b)  $\mu = +200$  GeV.

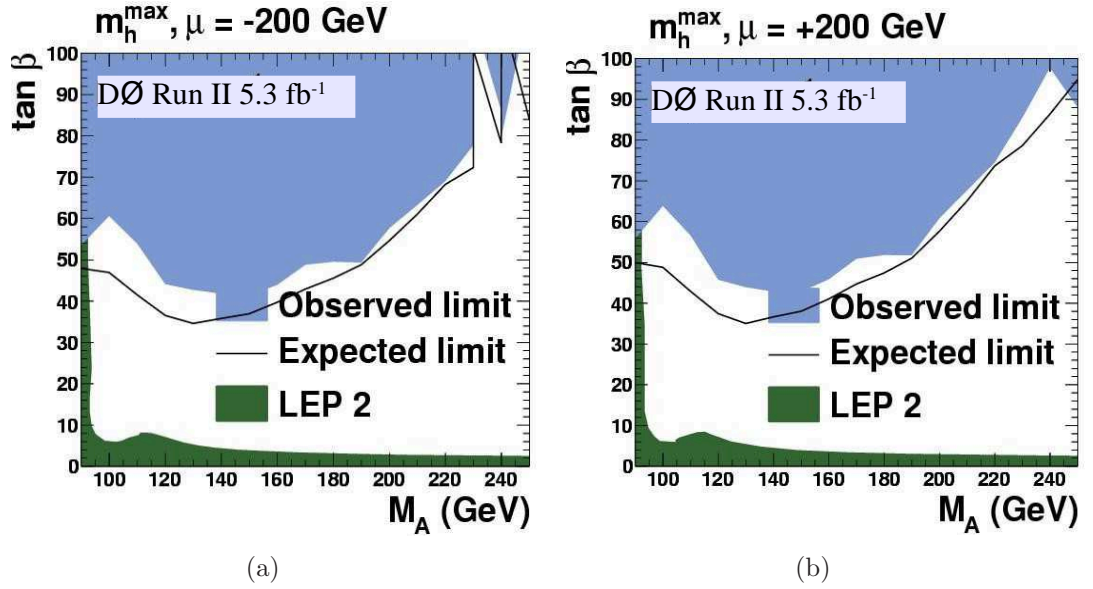


Figure 8.14: Expected and observed exclusion  $\tau_\mu \tau_h$  channel for the  $m_h^{\max}$  scenario in the  $(\tan \beta - M_A)$  plane, for a)  $\mu = -200$  GeV and b)  $\mu = +200$  GeV.

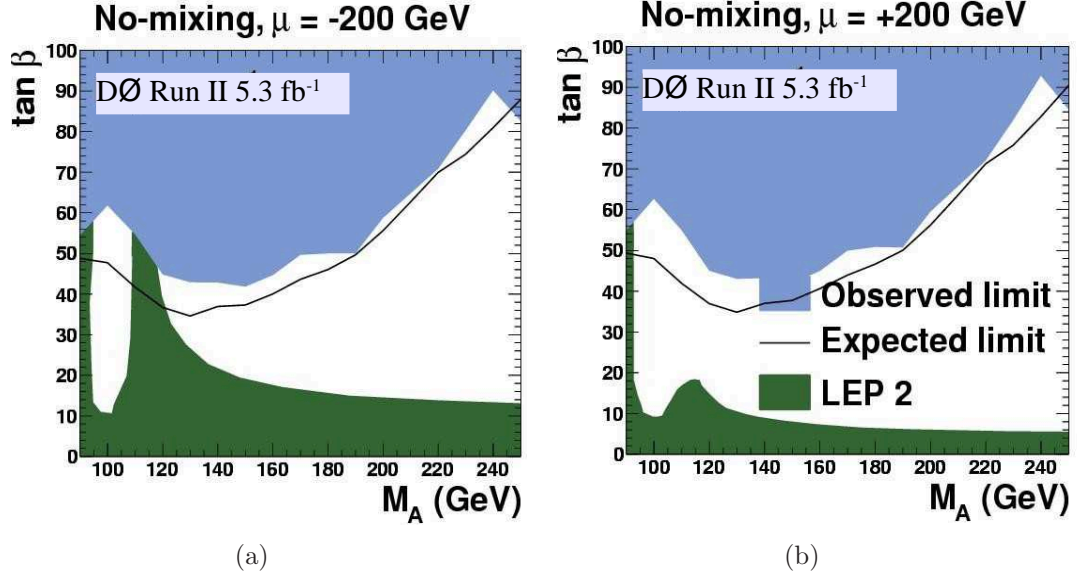


Figure 8.15: Expected and observed exclusion  $\tau_\mu\tau_h$  channel for the *no-mixing* scenario in the  $(\tan\beta - M_A)$  plane, for a)  $\mu = -200$  GeV and b)  $\mu = +200$  GeV.

the Tevatron are  $b\bar{b}\phi$ , where  $\phi \rightarrow b\bar{b}$  with at least three  $b$  quarks present in the final state or  $\phi \rightarrow \tau\bar{\tau}$ , and the inclusive  $\phi \rightarrow \tau\tau$  channel studied in this thesis. Inclusive searches for  $\phi \rightarrow \tau\tau$  have been performed with integrated luminosities of  $\mathcal{L} = 1.08 \text{ fb}^{-1}$  by DØ [26] in Run IIa and  $\mathcal{L} = 1.8 \text{ fb}^{-1}$  by CDF [27] in Run IIa and IIb. The result of  $\tau_e + \tau_h$  search channel performed in this thesis is also included. These searches require the  $\tau$  lepton pairs to decay into three final states:  $\tau_e\tau_h$ ,  $\tau_\mu\tau_h$ , and  $\tau_e\tau_\mu$ , where  $\tau_e$  and  $\tau_\mu$  are the leptonic decays of the tau and  $\tau_h$  is the hadronic decaying mode. Subsequently, DØ has updated its  $\phi \rightarrow \tau_\mu\tau_h$  search using  $\mathcal{L} = 1.2 \text{ fb}^{-1}$  of the Run IIb dataset, deriving 95% C.L. exclusion limits to include a total of  $2.2 \text{ fb}^{-1}$  data in the  $\tau_\mu\tau_h$  mode. The results performed in this thesis re-analyzed the  $\phi \rightarrow \tau_\mu\tau_h$  search channel within DØ Run IIa dataset and extends the search further to DØ Run IIb dataset for a total of  $5.36 \text{ fb}^{-1}$  of integrated luminosity. Details of these data sets are given in Chapter 4. Further, DØ has updated its  $b\bar{b}\phi, \phi \rightarrow \tau\tau$  search in summer 2010. This result is only considered in tree level and shown in Fig. 8.16 [77].

The three channels explored so far at the Tevatron have different features that is different depending on the mass of interest and the MSSM scenario. Therefore, the combinations in the future of all the channels will provide the best sensitivity to the MSSM Higgs boson. First, the search for  $\phi \rightarrow \tau\tau$  channel presented in this thesis is limited by many Drell-Yan background events around 90 GeV

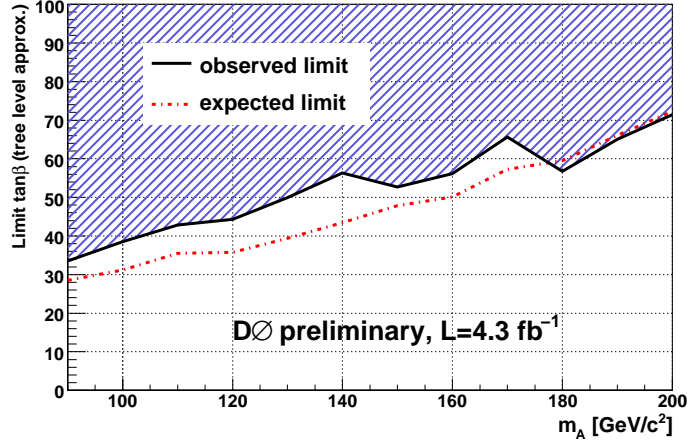


Figure 8.16: The exclusion in MSSM  $M_A$ - $\tan\beta$  plane of the  $b\bar{b}\phi$ ,  $\phi \rightarrow \tau\bar{\tau}$  channel with  $4.3 \text{ fb}^{-1}$  DØ Run IIb data updated in summer 2010. The limit has been calculated at tree level, therefore, no MSSM scenario is considered [77].

Higgs mass point, and therefore, the performance is poor in the low mass regions. However, due to sufficient statistics, this search channel is able to provide good sensitivities for higher Higgs mass points. For example, the most recent results for  $\tau_\mu\tau_h$  channel as shown in Fig. 8.14 probe  $\tan\beta < 40$  regions in a mass window of  $115 \leq M_A \leq 160 \text{ GeV}$  for the  $m_h^{\text{max}}$  scenario with a value of negative  $\mu$ . In addition, because this channel is not sensitive to the signs of higgsino mass parameter,  $\mu$ , it provides exclusions down to 40 in a similar mass window for  $\mu = +200$  as well as for  $\mu = -200$ .

Second, the search channel for Higgs bosons from the production associated with a  $b$  quark makes it possible to use the dominant  $b\bar{b}$  decaying mode of MSSM neutral Higgs bosons. As discussed in Chapter 1, the dependence on higgsino mass parameter,  $\mu$ , is strong in this decaying mode, and therefore, the  $bbb$  channel is especially sensitive to results for negative values of  $\mu$  [76]. Subsequently, the Drell-Yan backgrounds shown in lower masses can be significantly suppressed by adding a  $b$ -tagging in the  $b\bar{b}\phi$ ,  $\phi \rightarrow \tau\tau$  search channel. At  $M_A = 90 \text{ GeV}$  Higgs mass point, this channel excludes the possibility of  $\tan\beta \geq 34$  with 95% CL, which is about 1.5 times better than the results shown in Fig. 8.14 and Fig. 8.15 from the inclusive  $\phi \rightarrow \tau\tau$  searches.

The Tevatron combined results made in 2009 for  $\tau^+\tau^-$  final state including  $1\text{-}2.2 \text{ fb}^{-1}$  from DØ and  $1.8 \text{ fb}^{-1}$  from CDF inclusive searches for  $m_h^{\text{max}}$  and *no-mixing* benchmark scenarios for  $\mu = +200$  are shown in Fig. 8.17; the exclusions

set by the LEP experiment for both scenarios are also shown. This combination exclude  $\tan\beta \geq 30$  for  $M_A = 90$  and  $130 \leq M_A \leq 150$ . The most recent updates in  $\tau_\mu\tau_h$  channel presented in this thesis is not included in the combination, but this channel alone have exceeded the sensitivity in many of the Higgs masses to combined results. In addition, the later results extend the searches for Higgs mass up to 320 GeV. Table 8.2 listed the sensitivities for the MSSM Higgs searches described above.

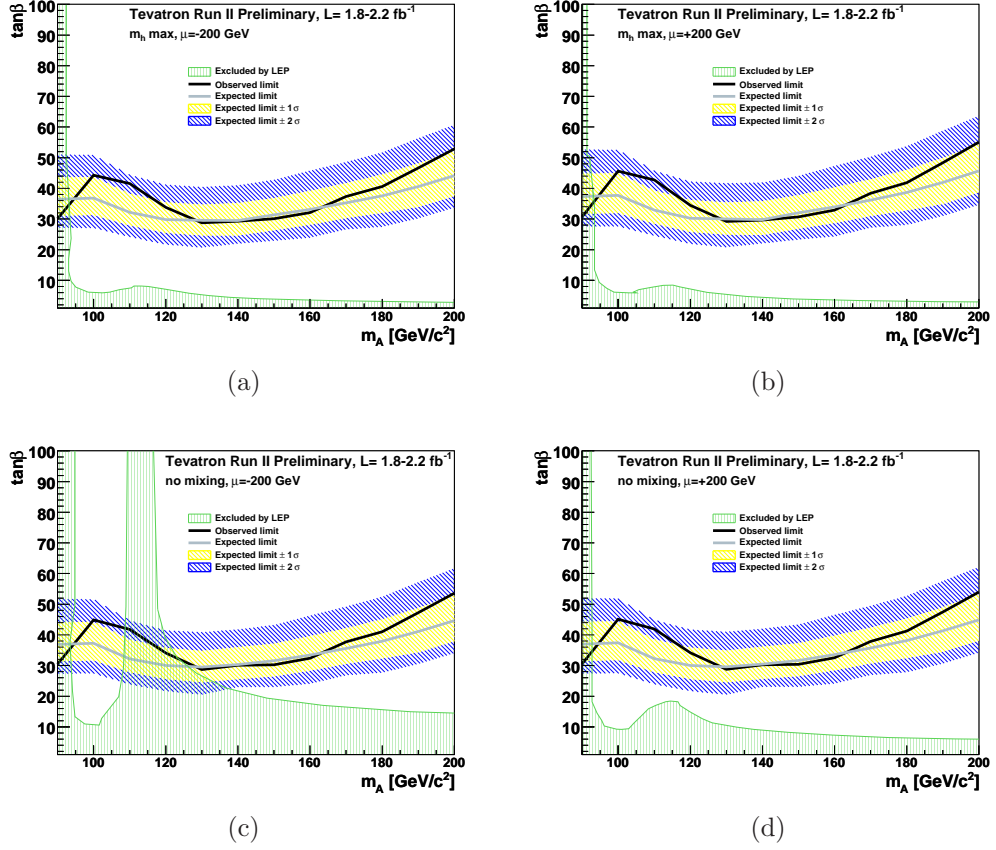


Figure 8.17: The exclusion in MSSM  $M_A$ - $\tan\beta$  plane of the Tevatron combined results in  $\tau^+\tau^-$  final state with (a)  $m_h^{\max}$  scenario with  $\mu = -200$  (b)  $m_h^{\max}$  scenario with  $\mu = +200$  (c) *no-mixing* scenario with  $\mu = -200$  (d) *no-mixing* scenario with  $\mu = +200$  [28].

| 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                    |
| $Z/\gamma^* \rightarrow ll$ cross section | —                                | —                      | —                      | 5                                    | 5                      | 5                      |
| PDF variation                             | 4.6                              | 4.6                    | 4                      | —                                    | —                      | —                      |
| $\tau$ -Track Reconstruction              | 1.4                              | 1.4                    | 4                      | 1.4                                  | 1.4                    | 4                      |
| Muon Identification                       | 2.1                              | 2.1                    | —                      | 2.1                                  | 2.1                    | —                      |
| $Z \rightarrow e^+e^-$ Background         | —                                | —                      | —                      | —                                    | —                      | 13                     |
| EM Identification                         | —                                | —                      | 3.3                    | —                                    | —                      | 3.3                    |
| Errors Assigned by $\tau$ -type           |                                  |                        |                        |                                      |                        |                        |
| $\tau$ Identification                     | 4-8                              | 4.2-12                 | 4-8                    | 4-8                                  | 4.2-12                 | 4-8                    |
| $\tau$ 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 8.1: Summary of the various sources of systematic errors for both the signal and background described in the text.

| Detector | Channel                            | Integrated<br>luminosity (fb <sup>-1</sup> ) | Observed $\tan \beta$ limit<br>at $M_A = 150$ GeV |                                 | Reference         |
|----------|------------------------------------|--|---|---------------------------------|-------------------|
|          |                                    |  | <i>no - mixing</i><br>$\mu = +200$ GeV            | $m_h^{max}$<br>$\mu = -200$ GeV |                   |
| CDF      | $b\phi \rightarrow bbb$            | 1.9  | —   | 90                              | [79]              |
| DØ       | $b\phi \rightarrow bbb$            | 2.6  | —   | 50                              | [76]              |
| CDF      | $\phi \rightarrow \tau\tau$        | 1.8  | 40  | 41                              | [78]              |
| DØ       | $\phi \rightarrow \tau\tau$        | 1.0  | 47  | 50                              | This thesis, [26] |
| DØ       | $\phi \rightarrow \tau_\mu \tau_h$ | 5.3  | 38  | 36                              | This thesis       |
| Tevatron | $\phi \rightarrow \tau\tau$        | 1.8-2.2                                      | 32  | 31                              | [28]              |
| DØ       | $b\phi \rightarrow b\tau\tau$      | 4.3  | 52 (tree-level)                                   |                                 | [77]              |

Table 8.2: 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.

# Chapter 9

## Conclusions

The direct searches at the Tevatron extend the search regions covered by LEP. The regime with low to moderate  $M_A$  with high  $\tan\beta$  provides best sensitivity due to the enhancement of the Higgs boson to down-type fermions. The analysis performed in this thesis extends the  $\phi \rightarrow \tau_\ell \tau_{had}$  search with  $5.36 \text{ fb}^{-1}$  data, where  $\ell = e, \mu$ . It includes  $4.28 \text{ fb}^{-1}$  of the Run IIb data for  $\tau_\mu \tau_{had}$  channel and results from  $1.08 \text{ fb}^{-1}$  Run IIa data for both  $\tau_\mu \tau_{had}$  and  $\tau_e \tau_{had}$  channels.

A search for a neutral Higgs boson that decays into tau leptons is presented using  $1.0 \text{ fb}^{-1}$  of collision data recorded from 2002 to 2006 with the DØ detector at the the Fermilab Tevatron. No significant signal is found and the resulting cross section limits are interpreted within the minimal supersymmetric Standard Model (MSSM).

Two independent channels are analysed where the events are required to contain a muon or electron with transverse momentum  $p_T > 15 \text{ GeV}$  and a hadronically decaying tau candidate. The main backgrounds to the signal process are  $Z \rightarrow \tau^+ \tau^-$ , W boson production in association with jets (W + jets),  $Z \rightarrow e^+ e^-$  and multijet events. Two independent methods for each of the channels have been developed to estimate the multijet background. The W + jets background shape is estimated using Monte Carlo (MC) but the normalisation is derived in a data sample dominated by W + jets events. The description of both these backgrounds is verified in control samples in the data. The remaining backgrounds are estimated using MC. The simulation is corrected for differences between the MC and the data, including a correction to the tau energy scale in the calorimeter that is derived in this thesis.

The data events are required to pass one of the single electron or muon trig-

gers. The MC simulation does not include a simulation of the trigger system and so trigger efficiencies for electrons and muons are derived from  $Z \rightarrow l^+l^-$  data events, where  $l = \mu$  or  $e$ . In particular, including electron triggers that use track information resulted in efficiencies around 50% at electron  $p_T \sim 20$  GeV in the central calorimeter, which represents a significant improvement over the calorimeter based electron triggers that have zero efficiency at  $p_T \sim 20$  GeV.

The backgrounds from  $W + \text{jets}$ ,  $Z \rightarrow e^+e^-$  and multijet events are largely reduced with a set of selections to retain a data sample that is dominated by  $Z \rightarrow \tau^+\tau^-$  events. The visible mass variable is used to search for a Higgs signal above the remaining background. No significant excess of events above the predicted background is found and 95% confidence level limits are set on the cross section times branching ratio  $\sigma \times \text{BR}(\phi \rightarrow \tau\tau)$  for Higgs bosons in the mass range  $90 \text{ GeV} < M_\phi < 300 \text{ GeV}$ . The observed limits range from 15 pb at  $M_\phi = 90 \text{ GeV}$  to 0.75 pb at  $M_\phi = 300 \text{ GeV}$  and the corresponding expected limits from 27 pb to 0.38 pb. The cross section limits are translated into exclusions in the  $(\tan\beta - M_A)$  plane for the  $m_h^{max}$  and no-mixing MSSM scenarios using FeynHiggs [22, 21, 23, 24], resulting in exclusions in the range  $40 < \tan\beta < 70$  for  $90 \text{ GeV} < M_A < 200 \text{ GeV}$ . These results [?] represent a significant extension of sensitivity in the  $(\tan\beta - M_A)$  plane compared to the previous searches for MSSM Higgs bosons [25, ?, ?, ?] and have comparable sensitivity to recent preliminary CDF results [78].

For certain MSSM parameter values, the width of the Higgs boson can become significantly larger than in the SM. In this thesis, the effect of the Higgs boson width is studied and a correction for the cross section limits as a function of Higgs boson width is derived. This allows the correction to be used independent of assumptions on the MSSM scenario for the first time. The correction for a Higgs boson of mass  $M_\phi$  and total decay width  $\Gamma_\phi$  ranges from 10% to 20% for  $\Gamma_\phi/M_\phi = 0.2$ .

The DØ experiment has now accumulated over  $3 \text{ fb}^{-1}$  of integrated luminosity and is expected to accumulate between  $6 \text{ fb}^{-1}$  and  $8 \text{ fb}^{-1}$ . The analyses presented here can be repeated with the increased dataset, allowing the Tevatron to extend its sensitivity to the MSSM Higgs bosons. In this thesis a projection of the final sensitivity for the Tevatron experiments in the MSSM benchmarks with the  $\phi \rightarrow \tau\tau$  final state is presented, which predicts sensitivity down to  $\tan\beta \sim 20$  at  $m_A = 100 \text{ GeV}$  and  $\tan\beta \sim 30$  at  $m_A = 200 \text{ GeV}$ .

The energy frontier in high energy particle physics will soon move to CERN,



with the Large Hadron Collider (LHC) scheduled to have first collisions in 2008. Searches for Higgs bosons at the LHC will also rely on decays to tau leptons, for the SM Higgs boson ( $q\bar{q}h \rightarrow q\bar{q}\tau\tau$ ) [?], the neutral MSSM Higgs boson ( $\phi \rightarrow \tau\tau$ ) [?] and charged Higgs bosons ( $H^\pm \rightarrow \tau^\pm\nu$ ) [?]. The techniques used in this thesis will therefore be of great importance as Higgs searches move into the LHC era.

# Appendix A

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

This section shows a larger range of distributions with final selections (see Sect. 7.2) from  $\tau_\mu\tau_h$  channel that were not included in the main text. These are used for checking the background modeling and shown by each  $\tau$ -type and are listed below:

$\cos(\Delta\phi(\mu/\tau, \cancel{E}_T))$ : Cosine value of the difference between azimuthal angle  $\phi$  of  $\cancel{E}_T$  and muon or  $\tau$  objects.

$\Delta\phi(\mu, \tau)$ : Difference between azimuthal angle  $\phi$  of muon and  $\tau$ .

$\cos(\Delta R(\mu, \tau))$ :  $\Delta R$  distribution between selected muon and  $\tau$  objects.

$\mathcal{L}$ : Luminosity distributions.

$\eta^{\mu/\tau}$ : Detector pseudo-rapidity of muons or  $\tau$  candidates.

$\phi^{\mu/\tau}$ : Azimuthal angle distribution of muons or  $\tau$  candidates.

$p_T^\mu$ : Transverse momentum of muons.

$N_{jets}$ : Jet multiplicity of the events.

$M_T$ : Transverse mass distribution.

$M_{trk}$ : Invariant mass distribution of associated tracks.

## Run IIa

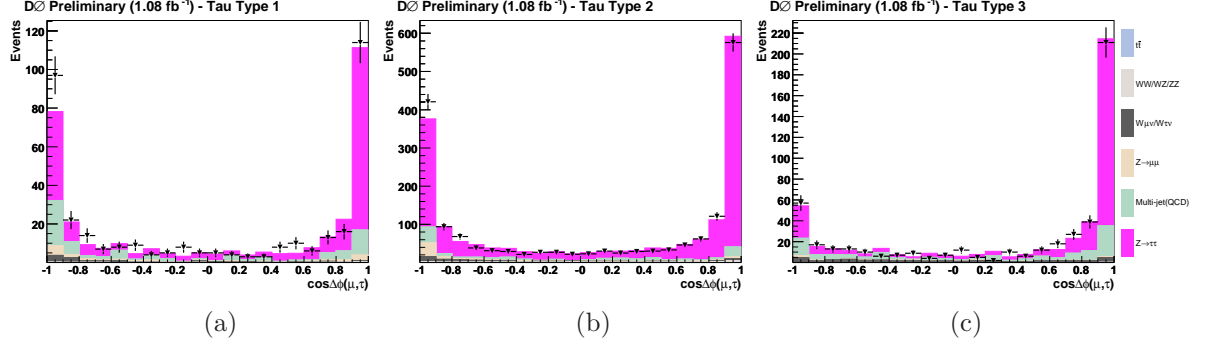


Figure 1:  $\cos(\Delta\phi(\mu, \cancel{E}_T))$  for  $\tau$ -type (a) 1, (b) 2 and (c) 3.

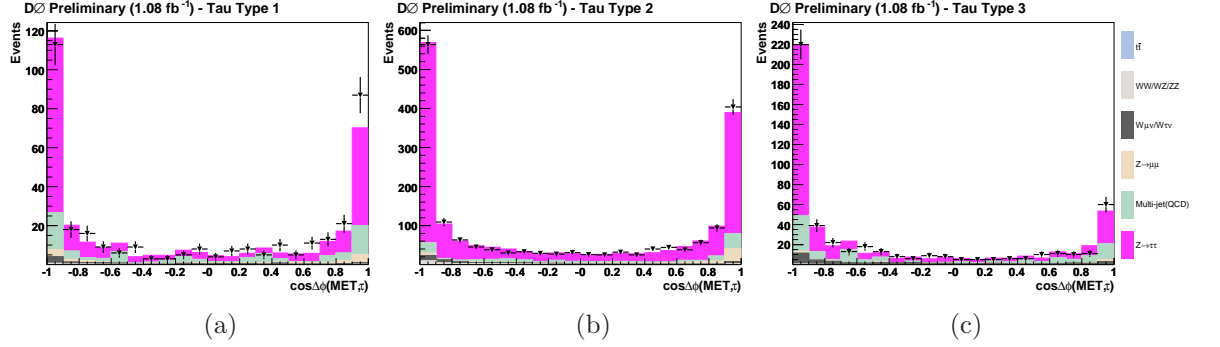


Figure 2:  $\cos(\Delta\phi(\tau, \cancel{E}_T))$  for  $\tau$ -type (a) 1, (b) 2 and (c) 3.

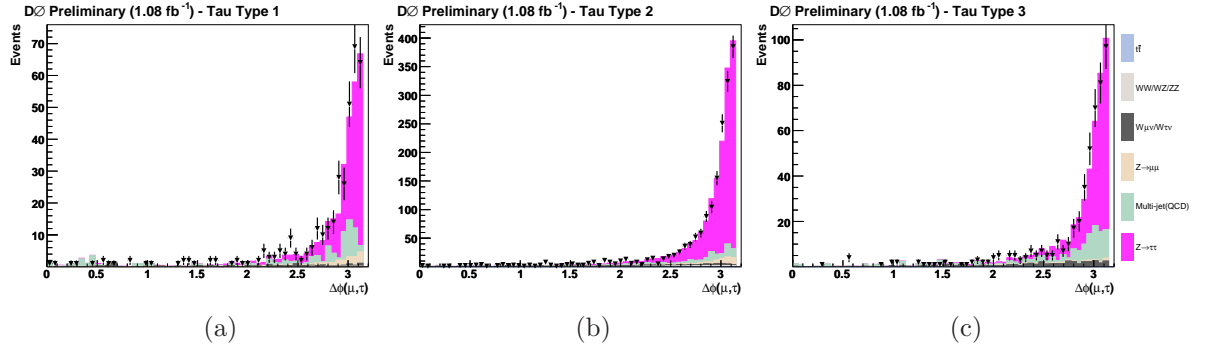


Figure 3:  $\Delta\phi(\mu, \tau)$  for  $\tau$ -type (a) 1, (b) 2 and (c) 3.

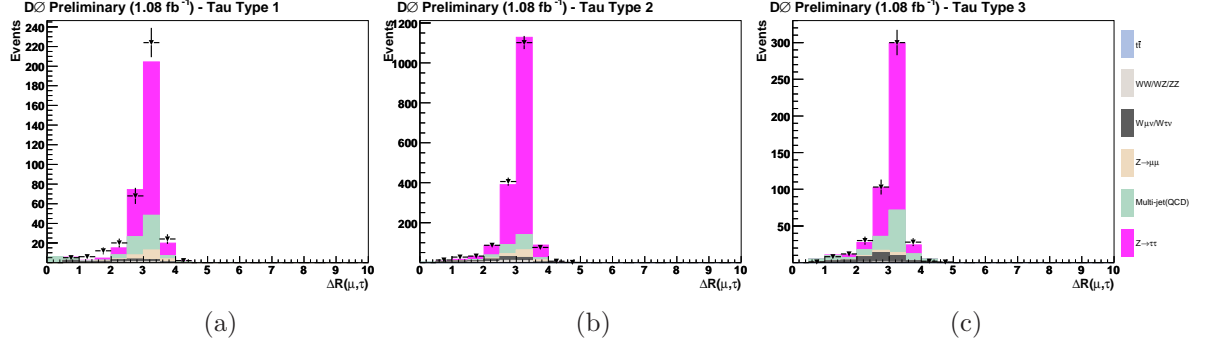


Figure 4:  $\Delta R(\mu, \tau)$  for  $\tau$ -type (a) 1, (b) 2 and (c) 3.

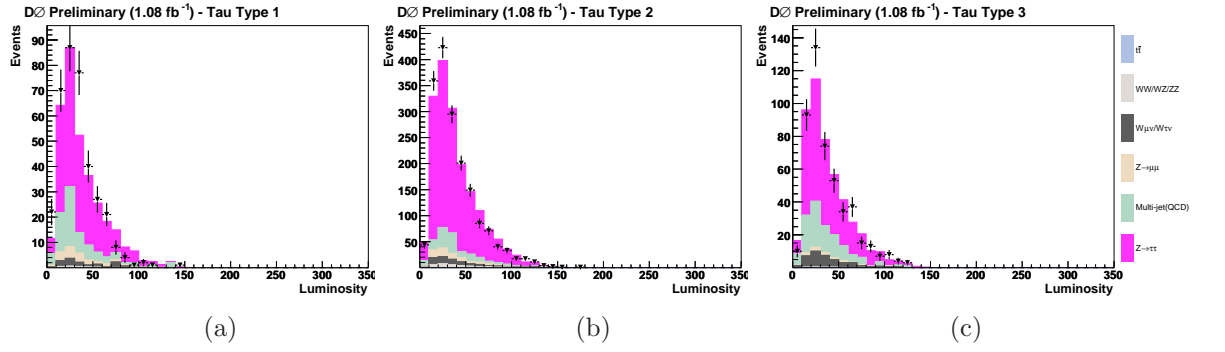


Figure 5:  $\mathcal{L} (\times 10^{32})$ , for  $\tau$ -type (a) 1, (b) 2 and (c) 3.

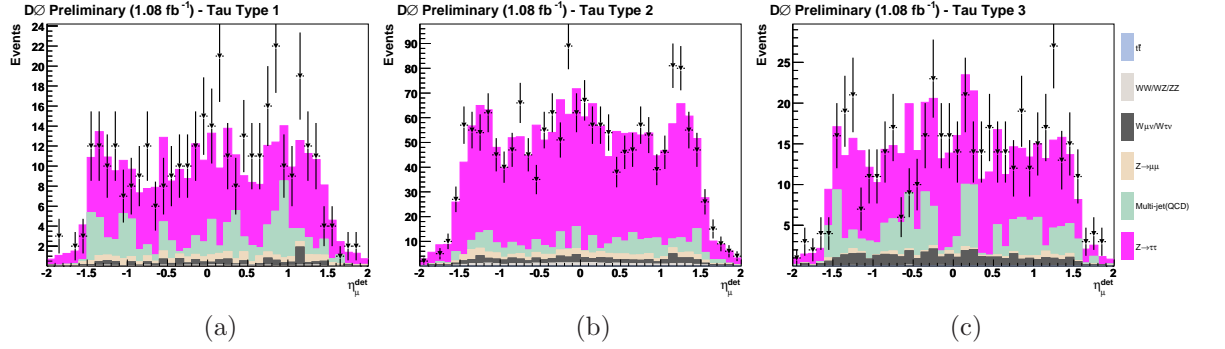


Figure 6:  $\eta^\mu$  for  $\tau$ -type (a) 1, (b) 2 and (c) 3.

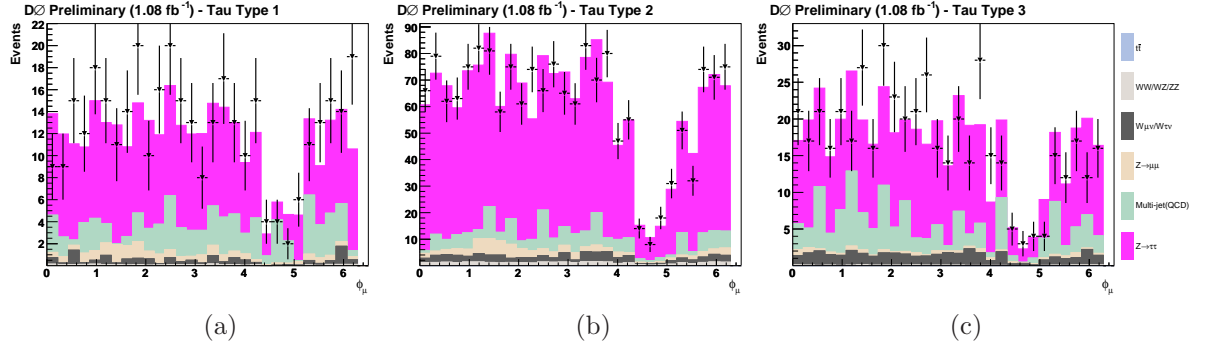


Figure 7:  $\phi^\mu$  for  $\tau$ -type (a) 1, (b) 2 and (c) 3.

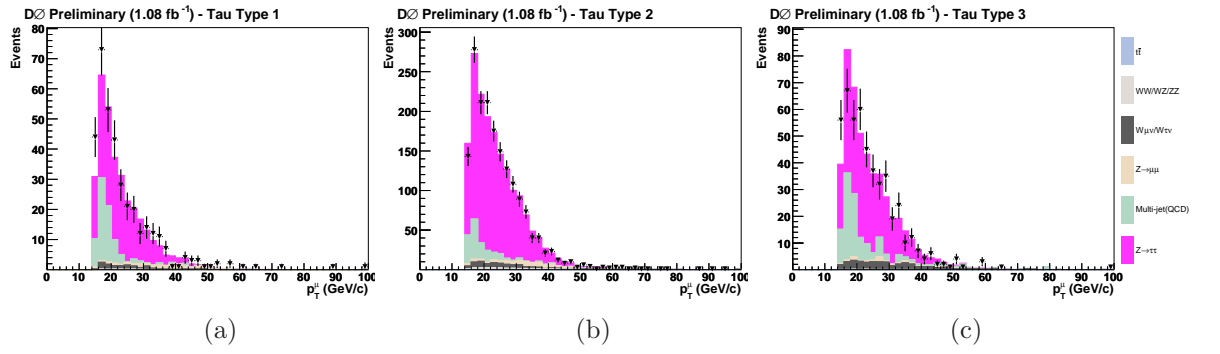


Figure 8:  $p_T^\mu$  for  $\tau$ -type (a) 1, (b) 2 and (c) 3.

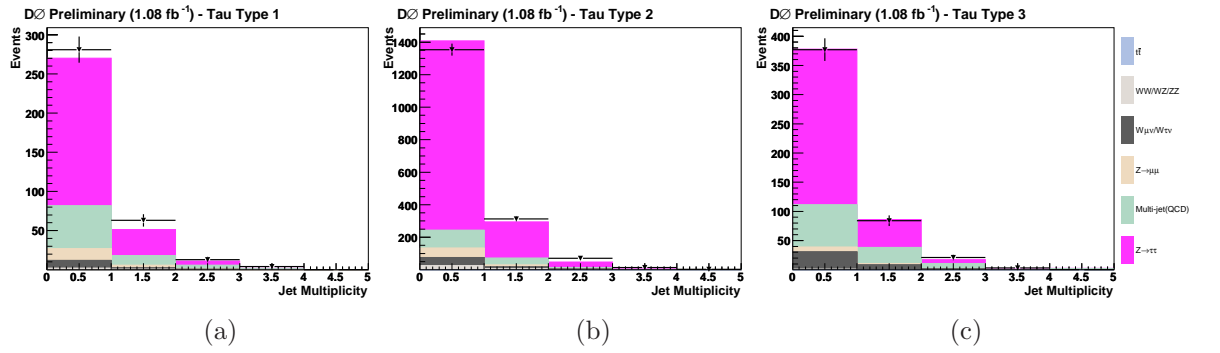


Figure 9:  $N_{jets}$  for  $\tau$ -type (a) 1, (b) 2 and (c) 3.

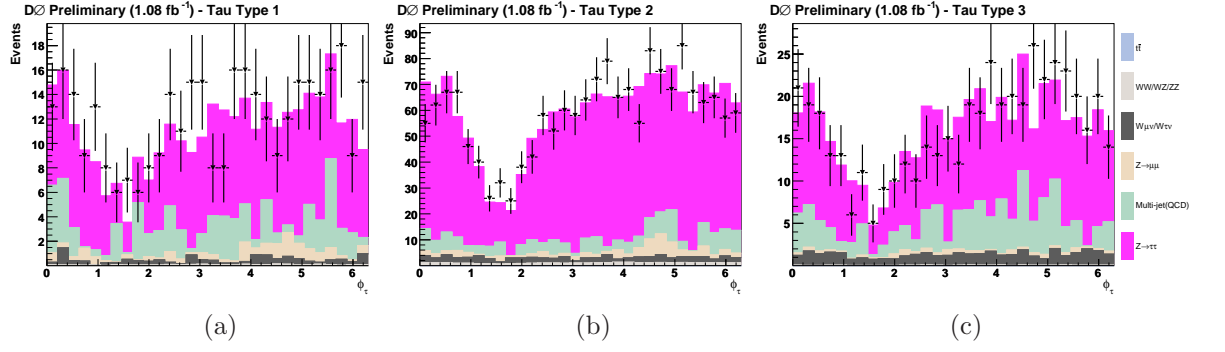


Figure 10:  $\phi_\tau$  for  $\tau$ -type (a) 1, (b) 2 and (c) 3.

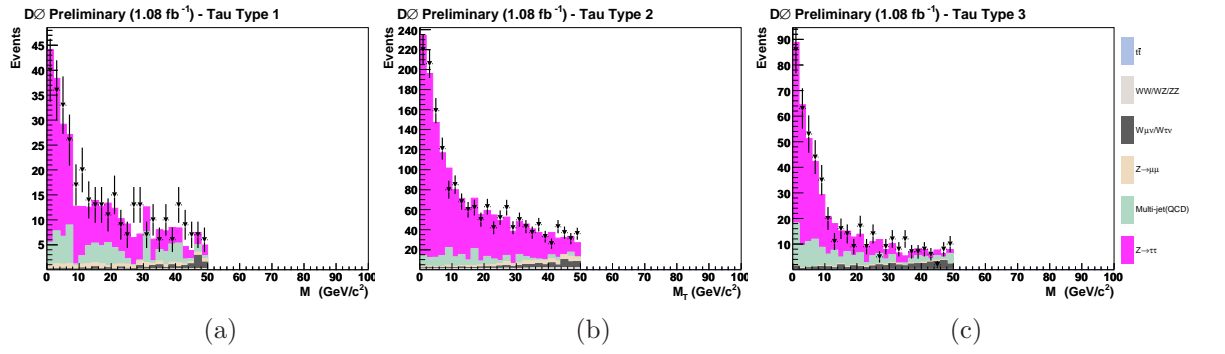


Figure 11:  $M_T$  for  $\tau$ -type (a) 1, (b) 2 and (c) 3.

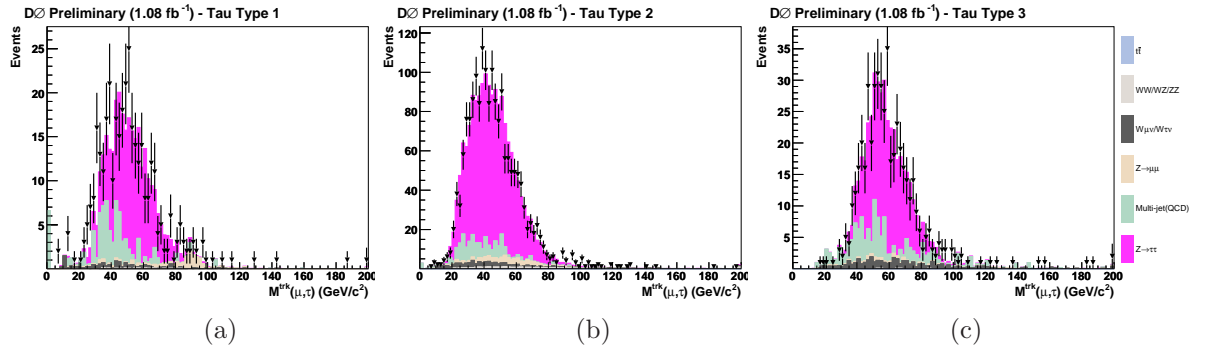


Figure 12:  $M_{trk}$  for  $\tau$ -type (a) 1, (b) 2 and (c) 3.

## Run IIb

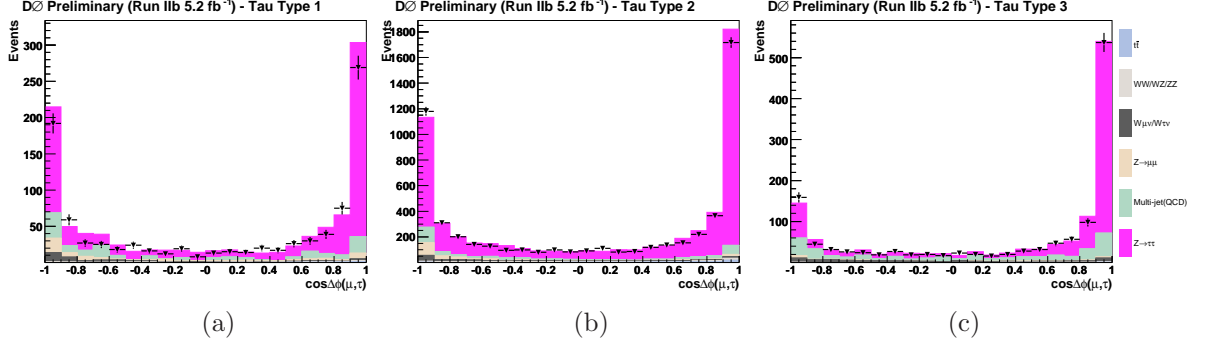


Figure 13:  $\cos(\Delta\phi(\mu, E_T))$  for  $\tau$ -type (a) 1, (b) 2 and (c) 3.

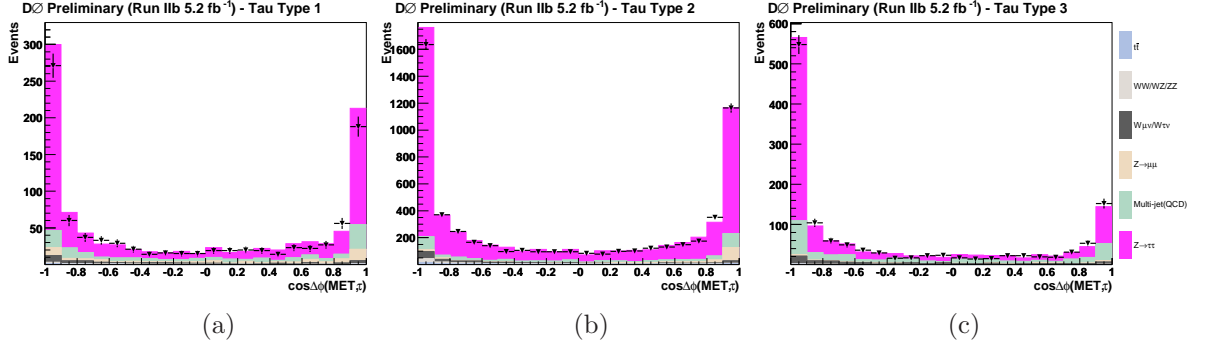


Figure 14:  $\cos(\Delta\phi(\tau, E_T))$  for  $\tau$ -type (a) 1, (b) 2 and (c) 3.

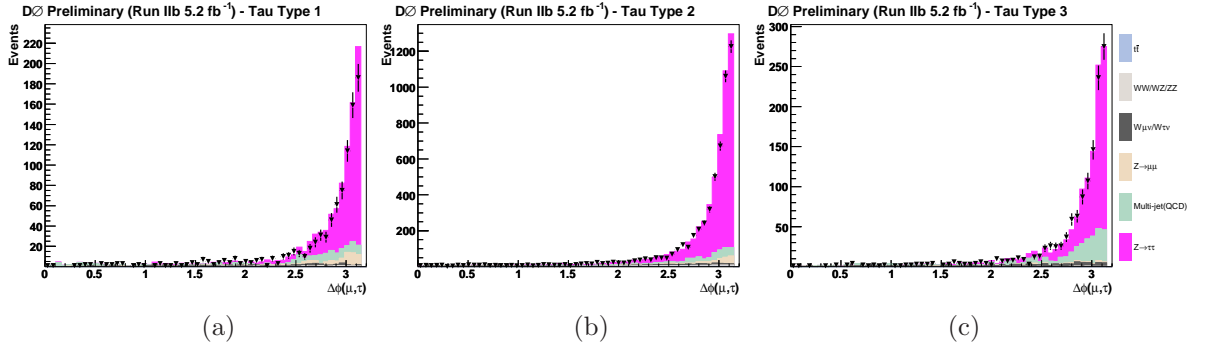


Figure 15:  $\Delta\phi(\mu, \tau)$  for  $\tau$ -type (a) 1, (b) 2 and (c) 3.

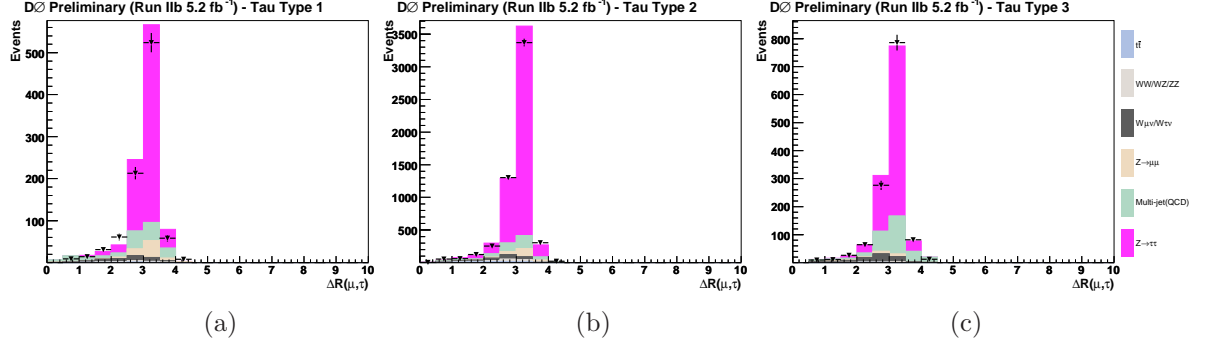


Figure 16:  $\Delta R(\mu, \tau)$  for  $\tau$ -type (a) 1, (b) 2 and (c) 3.

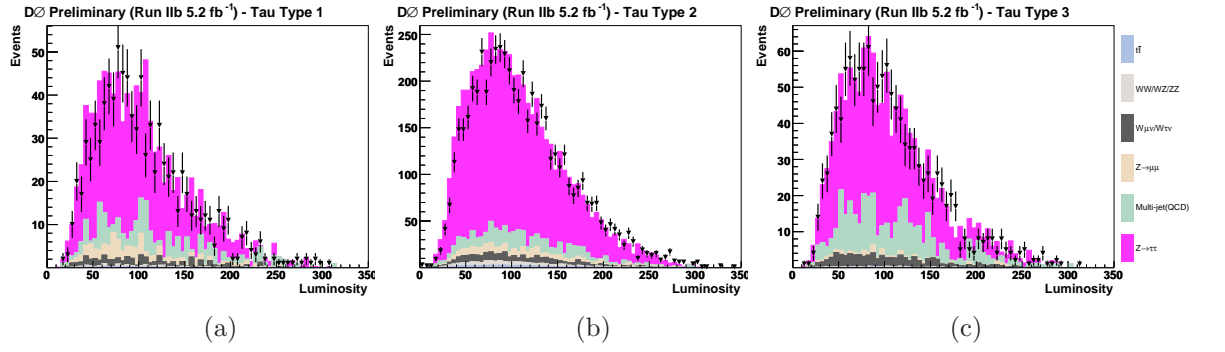


Figure 17:  $\mathcal{L} (\times 10^{32})$  for  $\tau$ -type (a) 1, (b) 2 and (c) 3.

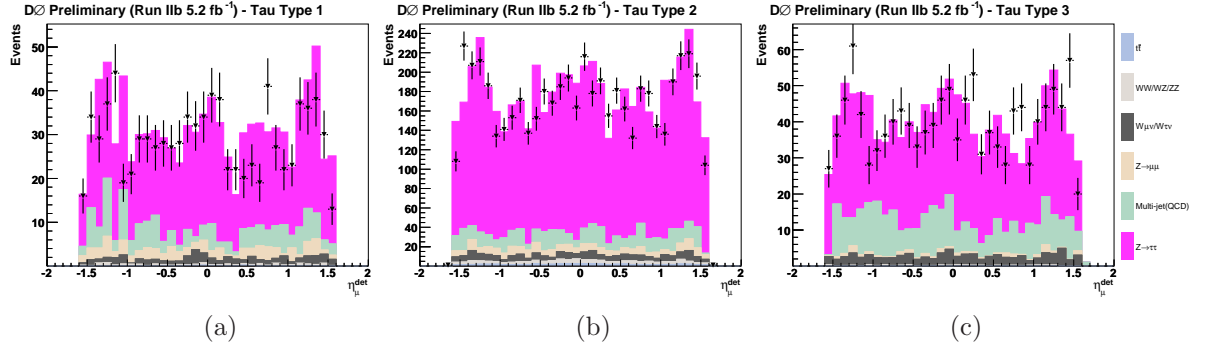


Figure 18:  $\eta^\mu$  for  $\tau$ -type (a) 1, (b) 2 and (c) 3.



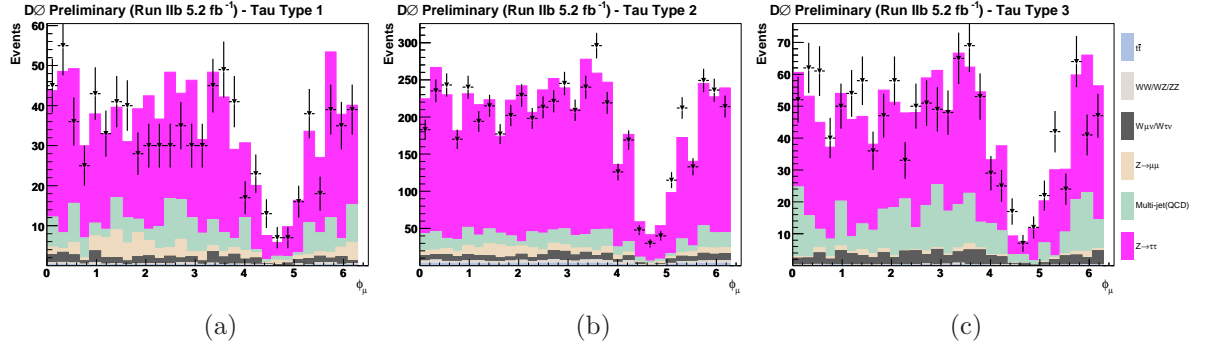


Figure 19:  $\eta^\mu$  for  $\tau$ -type (a) 1, (b) 2 and (c) 3.

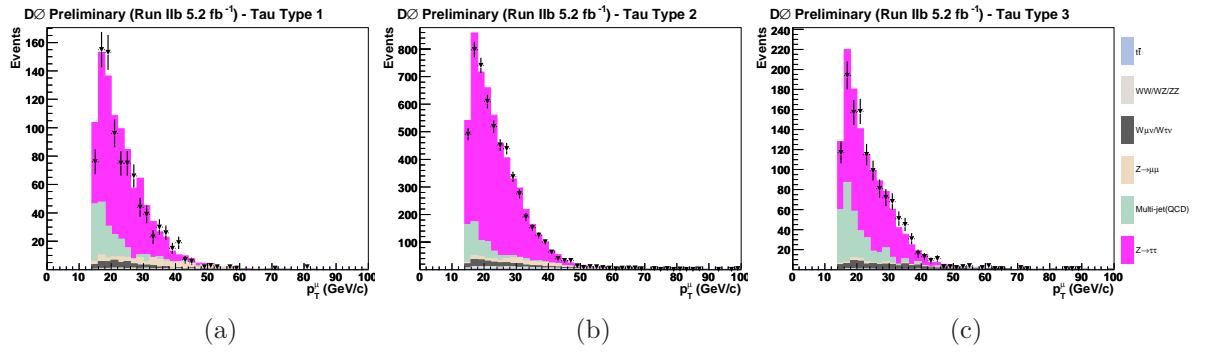


Figure 20:  $p_T^\mu$  for  $\tau$ -type (a) 1, (b) 2 and (c) 3.

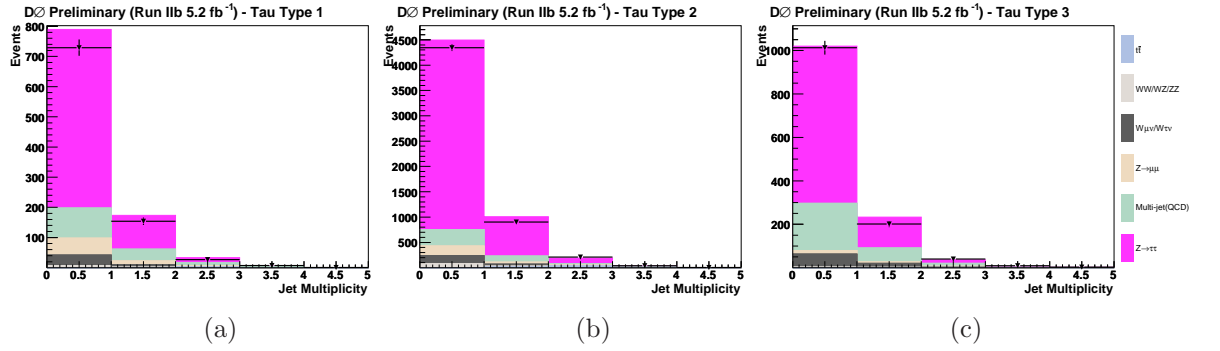


Figure 21:  $N_{jets}$  for  $\tau$ -type (a) 1, (b) 2 and (c) 3.

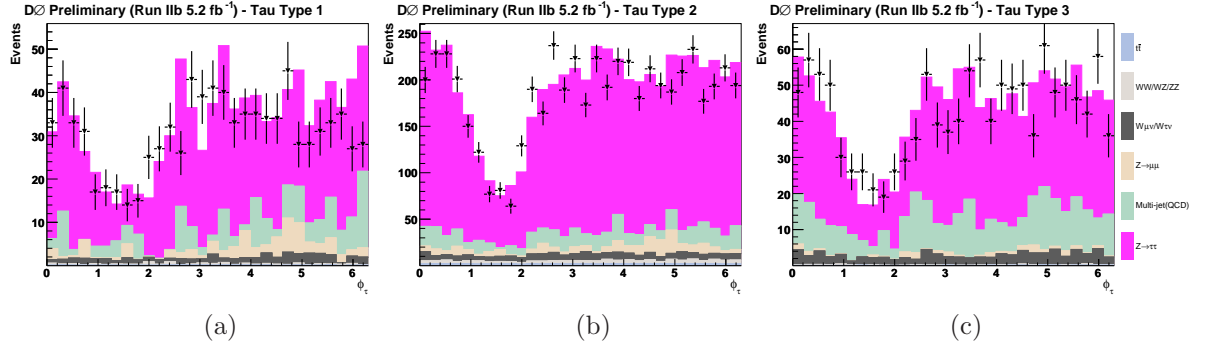


Figure 22:  $\phi^\tau$  for  $\tau$ -type (a) 1, (b) 2 and (c) 3.

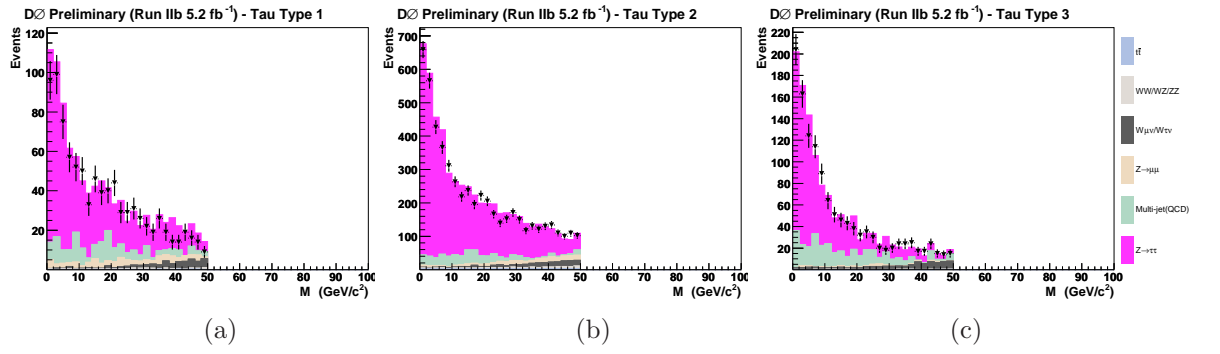


Figure 23:  $M_T$  for  $\tau$ -type (a) 1, (b) 2 and (c) 3.

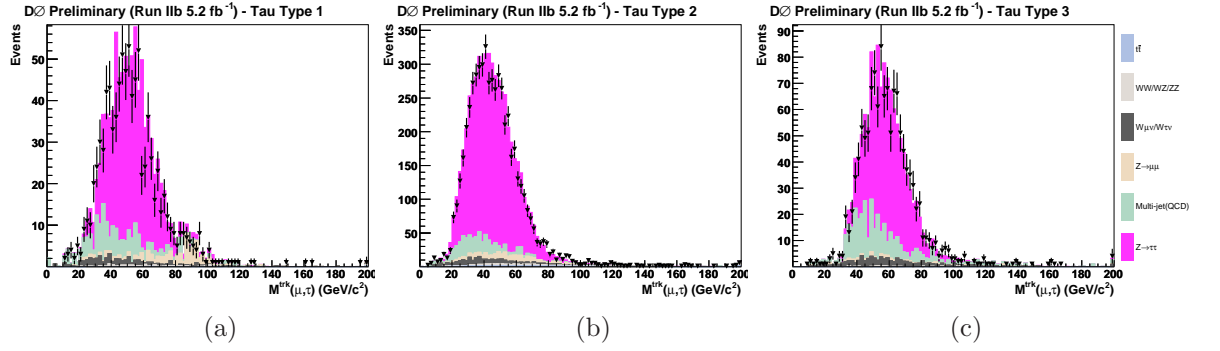


Figure 24:  $M_{trk}$  for  $\tau$ -type (a) 1, (b) 2 and (c) 3.

# Appendix B

## Cross checks for $Z \rightarrow e^+e^-$ Background

In  $\tau_e\tau_h$  channel, several anti-electron selections are applied after correcting the efficiencies (see Section 5.2 and 7.1.2). Data and predicted background in kinematic distributions are checked after the efficiency correction and also at each selection stage.

**Preselection:** Events after pre-selections are dominated by  $Z \rightarrow e^+e^-$  background (red), but big discrepancy is observed due to the mis-modeling of the electron efficiency in MC. Figure 9 shows the detector pseudo-rapidity distributions ( $\eta^\tau$ ) for  $\tau$ -type 1, 2 and 3 (from left to right).

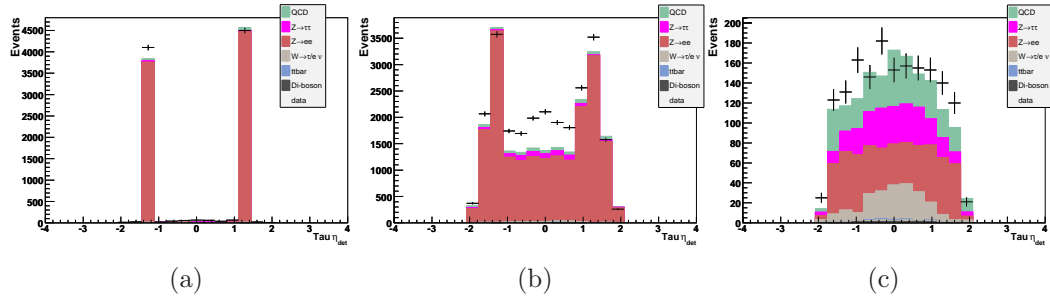


Figure 25: ( $\eta^\tau$  for  $\tau$ -type (a) 1, (b) 2 and (c) 3.

**ICD and EM Calorimeter Crack Removal:** After ICD and EM calorimeter crack Removal,  $NN_{elec}$  for type 2 events and invariant mass distribution of  $e - \tau$  pair ( $M_{e\tau}$ ) for type 3 are shown in Figure 9 which are later used for deriving efficiency correction. At this stage, there is already no many electrons in type 1 events, therefore, no attempt to further reduce electrons from type 1.

**Efficiency Correction for Only-one-tight-electron Sample:** Details about this correction is described in Section 5.2. Figure 27 show transverse momentum ( $p_T^e$ ) and pseudo-rapidity distribution ( $\eta^e$ ) of electrons. Three kinematic distributions for  $\tau$  candidates,

- transverse momentum ( $p_T^\tau$ )

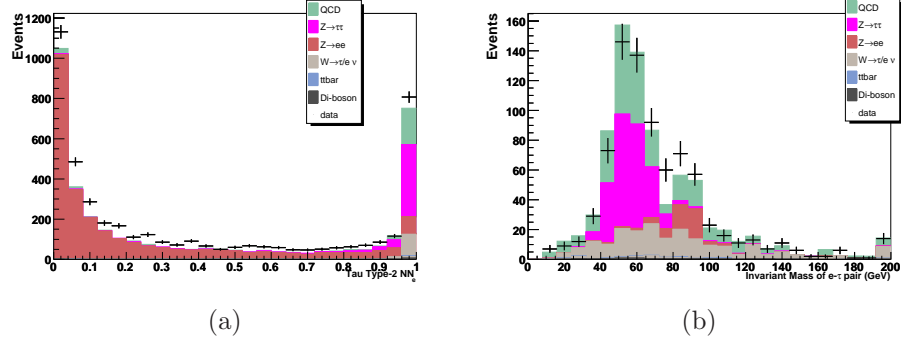


Figure 26: (a)  $NN_{elec}$  for type 2 (b)  $M_{e\tau}$  for type 3.

- pseudo-rapidity distribution ( $\eta^\tau$ )
- ratio of calorimeter cluster to transverse momentum of tracks ( $E/P_T^{trk}$ )

are shown by  $\tau$ -type in Figure 28 and 30.

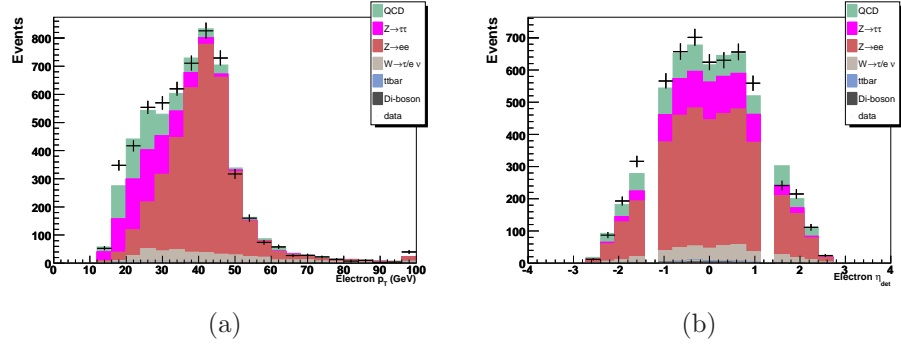


Figure 27: (a)  $p_T^e$  (b)  $\eta^e$ .

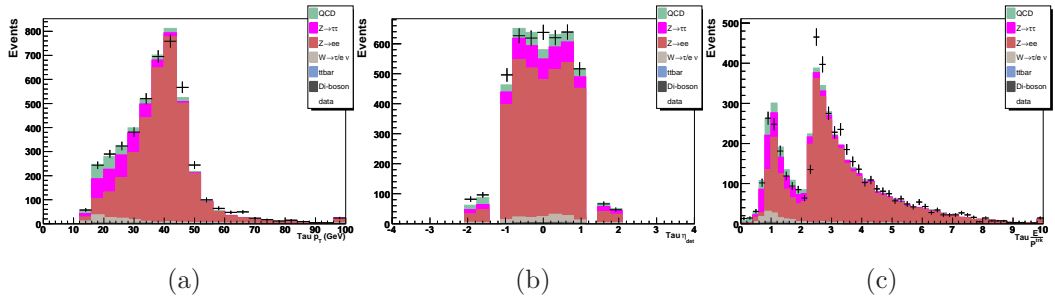


Figure 28: (a)  $p_T^\tau$  (b)  $\eta^\tau$  and (c)  $E/P_T^{trk}$  for  $\tau$ -type 2.

**$NN_{elec}$  Cut:** After correcting the efficiency, total prediction is able to describe data very well. Type 2  $Z \rightarrow e^+e^-$  events (red) are further reduced by applying  $NN_e > 0.9$ . Figure 9 shows the ratio of EM cluster to total calorimeter energy ( $f_{em}$ ) for  $\tau$ -type 2 (left) and 3 (right). The rest electron content in type 2 and 3 events can be rejected efficiently by requiring  $f_{em} \geq 0.9$

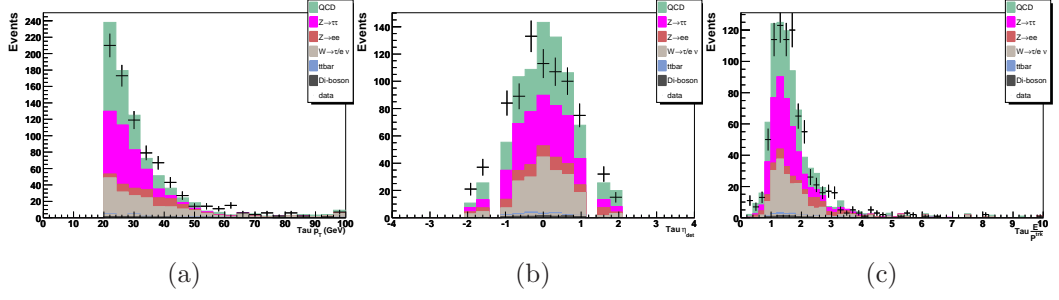


Figure 29: (a)  $p_T^{\tau}$  (b)  $\eta^{\tau}$  and (c)  $E/P_T^{\text{trk}}$  for  $\tau$ -type 3.

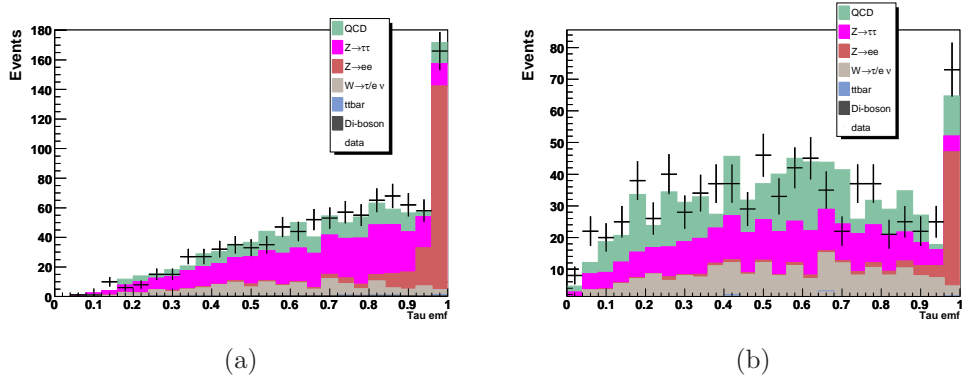


Figure 30:  $f_{em}$  after correcting electron efficiency.

# Appendix C

## Shape-dependent Systematic Uncertainties in Run IIb $\tau_\mu\tau_h$ Channel

In this section, Figures show the distributions of systematic uncertainties of each background source which have shape-dependency on  $M_{vis}$  as described in Section 8.2.2. These are then used in `Collie` to help find the ‘best fit’ for each histogram bin.

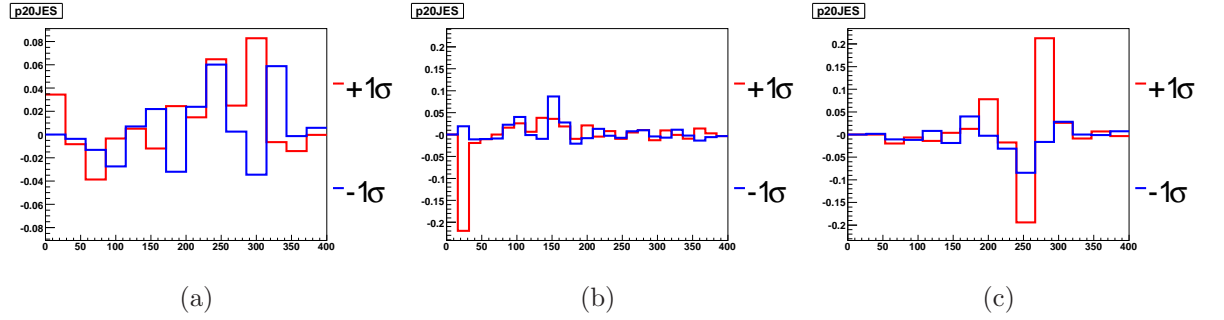


Figure 31: Shape dependent JES systematics of  $Z \rightarrow \tau^+\tau^-$  background for (a) type 1 taus, (b) type 2 taus and (c) type 3 taus.

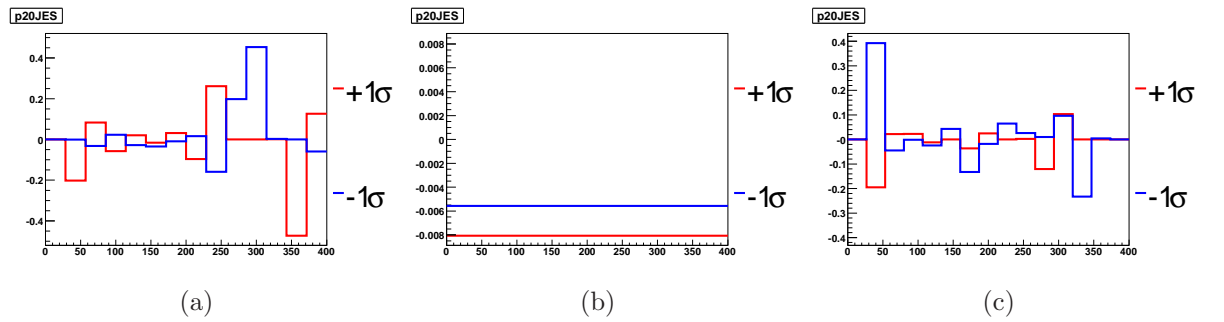


Figure 32: Shape dependent JES systematics of  $Z \rightarrow \mu^+\mu^-$  background for a) type 1 taus, b) type 2 taus and c) type 3 taus.

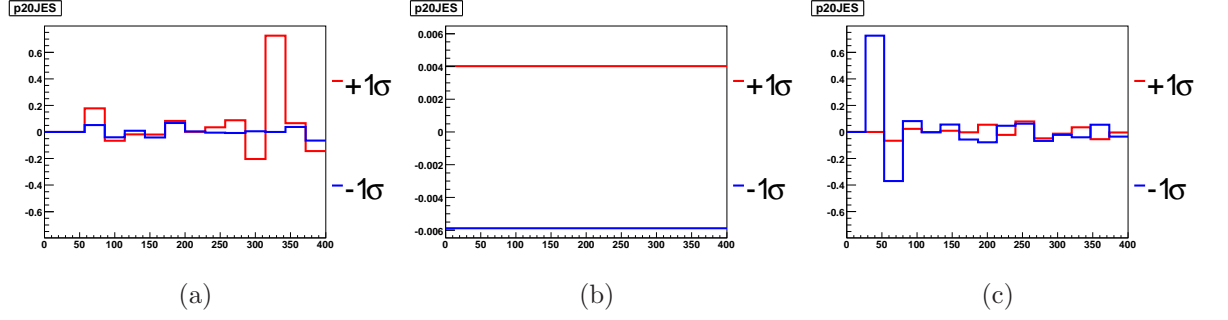


Figure 33: Shape dependent JES systematics of  $W$ +jet background for a) type 1 taus, b) type 2 taus and c) type 3 taus.

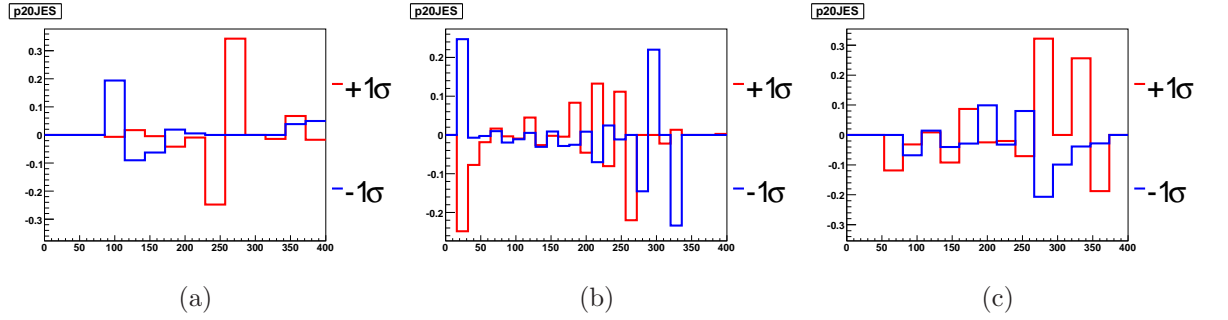


Figure 34: Shape dependent JES systematics of di-boson background for a) type 1 taus, b) type 2 taus and c) type 3 taus.

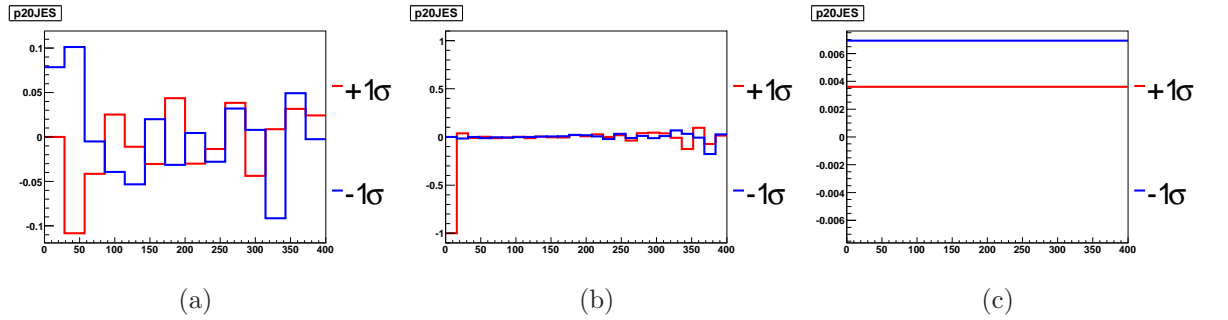


Figure 35: Shape dependent JES systematics of  $t\bar{t}$  background for a) type 1 taus, b) type 2 taus and c) type 3 taus.

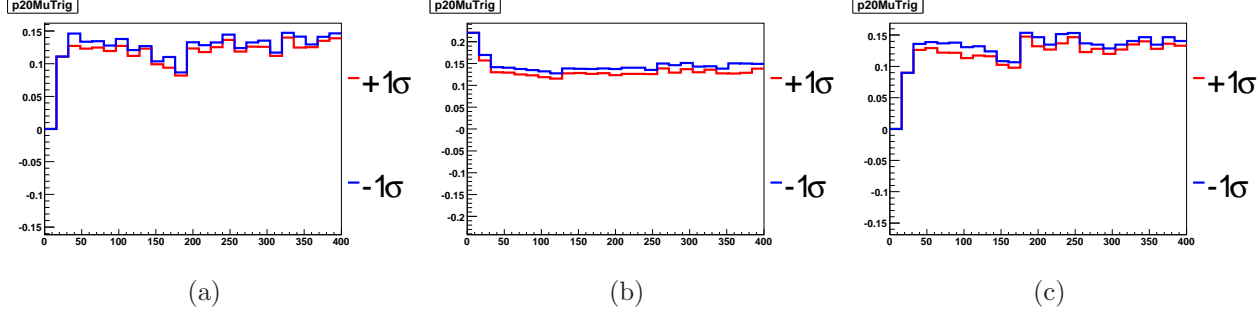


Figure 36: Shape dependent trigger systematics of  $Z \rightarrow \tau^+\tau^-$  background for a) type 1 taus, b) type 2 taus and c) type 3 taus.

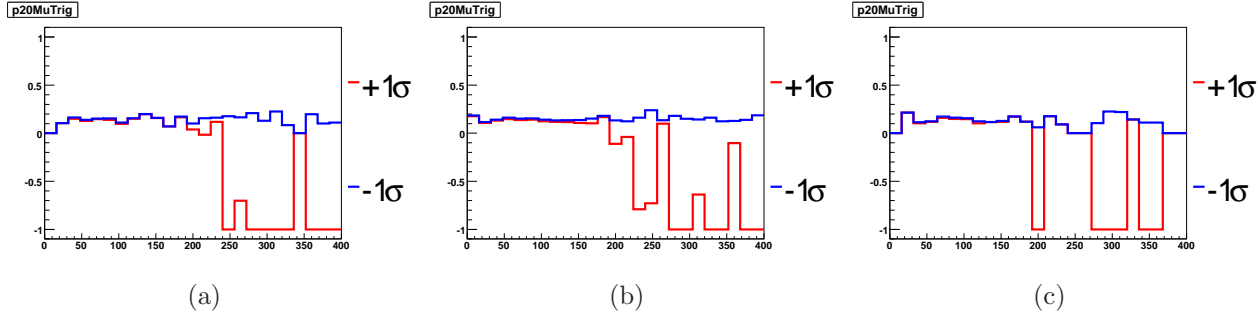


Figure 37: Shape dependent trigger systematics of  $Z \rightarrow \mu^+\mu^-$  background for a) type 1 taus, b) type 2 taus and c) type 3 taus.

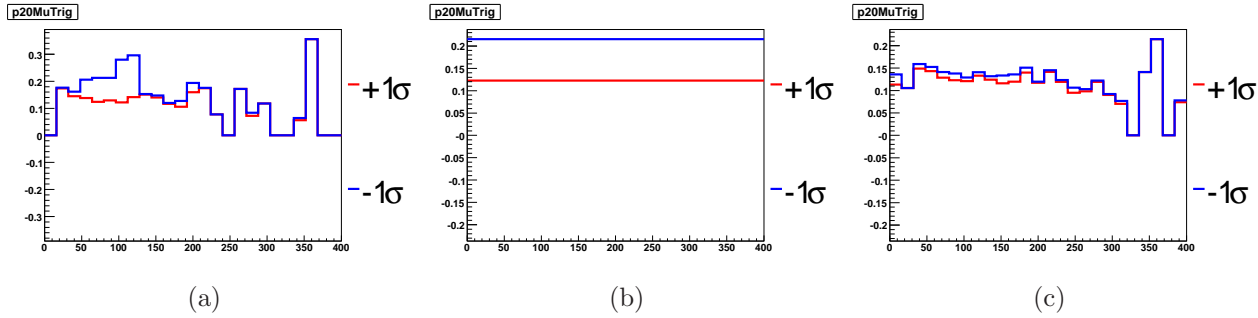


Figure 38: Shape dependent trigger systematics of  $W+\text{jet}$  background for a) type 1 taus, b) type 2 taus and c) type 3 taus.



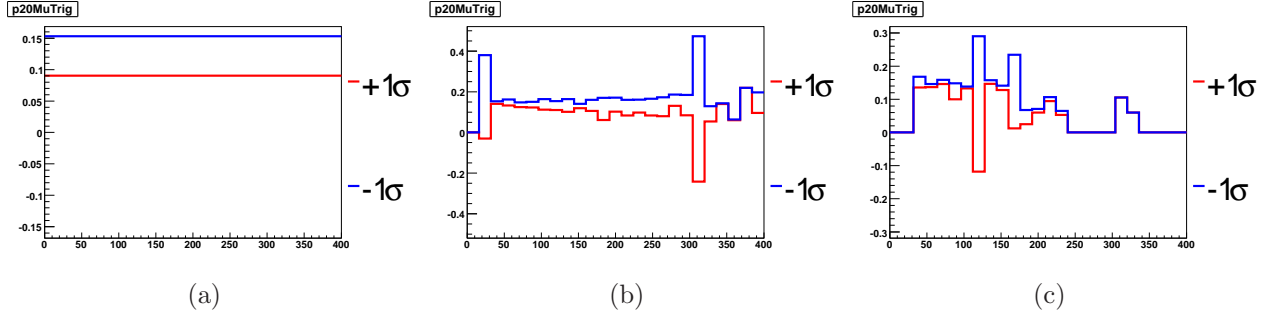


Figure 39: Shape dependent trigger systematics of di-boson background for a) type 1 taus, b) type 2 taus and c) type 3 taus.

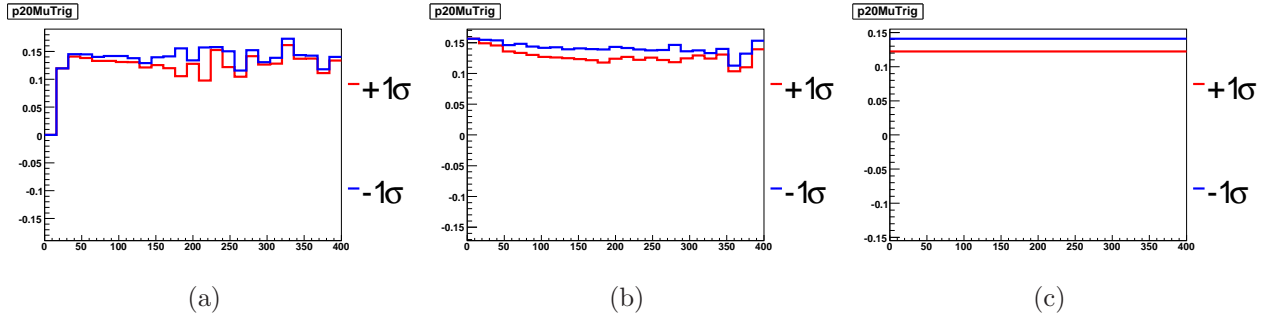


Figure 40: Shape dependent trigger systematics of  $t\bar{t}$  background for a) type 1 taus, b) type 2 taus and c) type 3 taus.

# Bibliography

- [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., *Observation of the top quark*, Phys. Rev. Lett. 74 (1995), 2632-2637, hep-ex/9503003.
- [4] F. Abe et al., *Observation of top quark production in  $p\bar{p}$  collisions*, Phys. Rev. Lett. 74 (1995), 2626-2631, hep-ex/0612052.
- [5] W.-M. Yao *et al.*, [Particle Data Group], Journal of Physics G **33** update for 2009 edition.
- [6] P.W. Higgs, *Broken Symmetries and the Masses of Gauge Fields*, Phys. Rev. Lett. 13 (1964), 508.
- [7] A. Sirlin, *Radiative Corrections in the  $SU(2)_L \times U(1)$  theory: A simple renormalization framework*, Phys. Rev. D22 (1980), 971-981.
- [8] C. Quigg, *The Electroweak Theory*, Flavor Physics for the Millennium: TASI 2000, edited by Jonathan 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), hep-ph/9808396.
- [11] R. Barate *et al.*, [ALEPH, DELPHI, L3 and OPAL Collaborations, LEP Working Group for Higgs Boson Searches], Phys. Lett. **B565**, 61 (2003), hep-ex/0306033.

- [12] T. Aaltonen *et al.* [CDF and D0 Collaborations], Combined CDF and D Upper Limits on Standard-Model Higgs-Boson Production. arXiv:1007.4587 [hep-ex].
- [13] C. Amsler *et al.*, [Particle Data Group], PL **B667**, 1 (2008).
- [14] Abdelhak Djouadi [LPTO] Phys. Rept. **459** 1-241 (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**, 797-814 (2006), hep-ph/0503173
- [18] M. Carena, S. Heinemeyer, C.E.M. Wagner, G. Weiglein Eur. Phys. J. **C26**, 601-607 (2003), hep-ph/0202167
- [19] T. Hahn, S. Heinemeyer, F. Maltoni, G. Weiglein, and S. Willenbrock, (2006), hep-ph/0607308.
- [20] A. Djouadi, J. Kalinowski, and M. Spira, Comput. Phys. Commun. **108**, 56 (1998), hep-ph/9704448.
- [21] G. Degrandi, S. Heinemeyer, W. Hollik, P. Slavich, and G. Weiglein, Eur. Phys. J. **C28**, 133 (2003), hep-ph/0212020.
- [22] M. Frank *et al.*, JHEP **02**, 047 (2007), hep-ph/0611326.
- [23] S. Heinemeyer, W. Hollik, and G. Weiglein, Eur. Phys. J. **C9**, 343 (1999), hep-ph/9812472.
- [24] S. Heinemeyer, W. Hollik, and G. Weiglein, Comput. Phys. Commun. **124**, 76 (2000), hep-ph/9812320.
- [25] S. Schael *et al.*, [ALEPH, DELPHI, L3 and OPAL Collaborations, LEP Working Group for Higgs Boson Searches], Eur. Phys. J. **C47**, 547 (2006), hep-ex/0602042.
- [26] DØ Collaboration, FERMILAB-PUB-08/132-E, Submitted to PRL (2008).

- [27] CDF Collaboration, CDF Note 9071, October 2007.  
[http://www-cdf.fnal.gov/physics/new/hdg/results/htt\\_070928](http://www-cdf.fnal.gov/physics/new/hdg/results/htt_070928)
- [28] CDF and DØ Collaborations [The TEVNPH Working Group], DØ Note 6036-CONF
- [29] O. Buchmueller, R. Cavanaugh, A. De Roeck, J.R. Ellis, H. Flacher, S. Heinemeyer, G. Isidori, K.A. Olive, F.J. Ronga, G. Weiglein Phys. Rev. **D81**, 035009 (2010), [arXiv.org:0912.1036v2](http://arxiv.org/abs/0912.1036v2).
- [30] 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), [arXiv:1007.0956v1](http://arxiv.org/abs/1007.0956v1)
- [31] Run II Handbook [http://www-bd.fnal.gov/lug/runII\\_handbook/RunII\\_index.html](http://www-bd.fnal.gov/lug/runII_handbook/RunII_index.html)
- [32] [http://d0server1.fnal.gov/Projects/Operations/D0RunII\\_DataTaking.htm](http://d0server1.fnal.gov/Projects/Operations/D0RunII_DataTaking.htm)
- [33] DØ Collaboration, *The Upgraded D0 Detector* Nucl. Instrum. Meth. **A565**: 463-537 (2006), [arXiv:physics/0507191v1](http://arxiv.org/abs/physics/0507191v1)
- [34] E. Kajifasz [DØ Collaboration], *DØ Silicon Microstrip Tracker for RunIIa*, Nucl. Instrum. Meth. **A511** 16-19 (2003).
- [35] *The DØUpgraded Central Fibre Tracker, Technical Design Report*, <http://d0server1.fnal.gov/projects/SciFi/cfthome.html>.
- [36] Darie Wood, [http://d0server1.fnal.gov/projects/run2b/Management/Posted\\_Info](http://d0server1.fnal.gov/projects/run2b/Management/Posted_Info)
- [37] Ariel Schwartzman, Chris Tully DØ Note **5157** (2007).
- [38] J. Hays, J. Mitrevski, C. Schwanenberger, and T. Toole, DØ Note **5105** (2006).
- [39] J. Hays, J. Mitrevski, C. Schwanenberger, and L. Wang, DØ Note **5114** (2006).
- [40] Laurent Duflot and Melissa Ridet DØ Note **3923**, December 2001.
- [41] A. Harel, DØ Note **4919** (2005).
- [42] DØ Jet Energy Scale Group, DØ Note **5382** (2007).

- [43] Oleg Brandt, SungWoong Cho, Michael Cooke, Michael Eads, Dave Hedin, Angelo Santos, Boris Tuchming Yuriy Yatsunenko, SungWoo Youn DØ Note **6025** (2010).
- [44] D. Chakraborty *et al.*, DØ Note **4210** (2003).
- [45] M. Arov *et al.*, DØ Note **4773** (2005).
- [46] S. Protopopescu and P. Svoisky, DØ Note **5094** (2006).
- [47] V. M. Abazov *et al.*, [DØ Collaboration], Nucl. Instrum. Meth. A **565**, 463 (2006), arXiv:physics/0507191.
- [48] Common Sample Group <http://www-d0.fnal.gov/Run2Physics/cs/index.html>
- [49] E. Nurse, *A measurement of the inclusive  $Z/\gamma^* \rightarrow \mu\mu$  cross section and study of W and Z events in pp collisions at DØ*, PhD thesis, The University of Manchester, 2005.
- [50] P. Calfayan, DØ Note **5329** (2008).
- [51] K. Herner, A. Patwa, W.C. Yang DØ Note **5699** (in preparation).
- [52] M. Hohlfeld and M. Owen, DØ Note **5409** (2007).
- [53] J. Hays *et al.*, DØ Note **5138** (2006).
- [54] T. Sjostrand *et al.*, *PYTHIA 6.3 Physics and manual*, hep-ph/0308153.
- [55] 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).
- [56] Torbjørn Sjostrand *Monte Carlo Generators*, hep-ph/0611247v1 (2006)
- [57] R. D. S. Jadach, Z. Was and J. H. Kuhn, Comput. Phys. Commun. **73**, 361 (1993).
- [58] D. Stump *et al.*, JHEP **10**, 046 (2003), hep-ph/0303013.
- [59] R. Brun *et al.*, *Geant3*, CERN-DD/EE/81-1 (1987).
- [60] <http://www-d0.fnal.gov/computing/MonteCarlo/MonteCarlo.html>

- [61] <http://www-d0.fnal.gov/computing/algorithms/howto/howtoreco.html>
- [62] <http://www-d0.fnal.gov/Run2Physics/cs/caf/>
- [63] H. Schellman, DØ Note **5142** (2006).
- [64] C. Balazs and C. P. Yuan, Phys. Rev. **D56**, 5558 (1997), hep-ph/9704258.
- [65] M. Shamim, T. Bolton, DØ Note **5565** (2008).
- [66] G. Hesketh DØ Note **5786** (2008).
- [67] <http://www-d0.hef.kun.nl//askArchive.php?base=agenda&categ=a09917&id=a09917>
- [68] <http://www-d0.hef.kun.nl//fullAgenda.php?ida=a092001&fid=57>
- [69] M. Owen, S. Soldner-Rembold, and W.-C. Yang, DØ Note **5457** (2008).
- [70] Wade Fisher, DØ Note **5595** (2010).
- [71] T. Junk, Nucl. Instrum. Meth **A434** 435 (1999)
- [72] DØ Collaboration, DØ Note **5484-CONF** (2007).
- [73] B. Casey *et al.*, DØ Note **4958** (July 2006) and G. Snow for the D0 Luminosity Group, DØ Note **5139** (October 2006).
- [74] V. Buscher, J.-F. Grivaz, T. Nunnemann, and M. Wobisch, DØ Note **4618** (2004).
- [75] K. Harder, private communication, 2006.
- [76] P. Jonsson *et al.*, [D0 Collaboration], DØ Note **5726-CONF** (2008).
- [77] F. Couderc *et al.*, [D0 Collaboration], DØ Note **6083-CONF** (2010).
- [78] A. Anastassov *et al.*, CDF Note **9071** (2007).
- [79] T. Wright and D. Amidei, CDF Note **9284** (2008).