UNIVERSITE
PIERRE & MARIE CURIE
LA SCIENCE A PARIS

THÈSE DE DOCTORAT DE L'UNIVERSITÉ PIERRE ET MARIE CURIE

Spécialité

Physique des particules

Présentée par

Dikai LI

Pour obtenir le grade de

Docteur de l'Université "Pierre Et Marie Curie"

Search for the standard model Higgs boson in $l\nu + b\bar{b}$ final states in 9.7 fb^{-1} of $p\bar{p}$ collisions with the DØ detector

Recherche du boson de Higgs du modèle standard dans l'etat final $WH \to \ell\nu b\bar{b}$ avec 9.7 fb^{-1} de collisions $p\bar{p}$ dans l'expérience DØ.

Soutenue le 10 décembre 2013 devant le jury composé de:

M^{me}	Stephanie BEAUCERON	Examinateur
M	Gregorio BERNARDI	Directeur de thèse
M	Jacques CHAUVEAU	Président
M	Yuji ENARI	Examinateur
M	Pascal PAGANINI	Rapporteur
M	Boris TUCHMING	Examinateur
M	Laurent VACAVANT	Rapporteur

Acknowledgements

There are many people to thank for their help throughout this thesis work. First, I would like to thank my parents, Jingfeng Li, and Hongping Qi. They always encouraged me to do what I am interested in, and as a result, I never stopped pursuing what I was dedicated to, though there were difficulties on the road.

Next, I would like to thank my thesis advisor, Professor Gregorio Bernardi. Gregorio brought me the most exciting opportunity to participate in one of the most advanced high energy physics laboratory in the world and searched for the golden physics target. I will be forever grateful for all his help and support over the years in everything from answering physics question, to editing my thesis, to taking care of my living, travelling at France and the United States. I would also like to thank Dr. Yuji Enari and Dr. Lidija Živković. Yuji supervised me for two years, he was patient and helped me a lot, from helping debug code, tutoring running the framework, to explaining every physics questions, even after he left Paris. Lidija supervised me for one year, she was very kind and helped me to collaborate with colleagues from other physics groups. I owe many thanks to Dr. Xuebing Bu, he supervised me on the JetID service work, which constitutes one of the most important parts of this thesis. The D0 WH analysis group members collaborated tightly and helped me a lot, so I would like to thank the group members I worked with for the past few years: Dr. Mike Cooke, Dr. Sebastien Greder, Dr. Ken Herner, Dr. Bob Hirosky, Dr. Jonathan Brown, Dr. Savanna Shaw, Dr. Emily Johnson and Dr. Huong Nguyen.

Many thanks to the other Professors, graduate students and post-doc at D0 and at LPNHE. Professor Lydia Roos gave me the opportunity to study at the LPNHE, and helped me a lot on dealing with the China students affairs. Professors Jacques Chauveau, Jacques Dumarchez, Sophie Trincaz-Duvoid, José Ocariz, Bertrand Laforge, and Giovanni Marchiori always cared about the progress on my working and helped me on all kinds of practical doctoral students issues. Dr. Li Yuan, Dr. Liwen Yao, and Kun Liu discussed with me a lot when I stayed at LPNHE.

Many thanks go also to the members of the Jury of my thesis not directly involved during the thesis writing, i.e. Dr Stephanie Beauceron, Dr Pascal Paganini, Dr Boris Tuchming and Dr Laurent Vacavant, for their help in improving the manuscript. Special thanks to Boris, who spent lots of time to help me finalize the manuscript.

Happy thanks go to those dear friends who made my life at Fermilab enjoyable: Dr. Yunhe Xie, Dr. Jie Chen, Dr. Fan Yang, Dr. Liang Li, Dr. Hang Yin, Dr. Weigang Geng, Dr. Guo Chen, Dr. Pengfei Ding, Jiaming Yu, Siqi Yang, Peng Jiang.

Last but not the least, I must thank the China Scholarship Council and the China University of Geosciences, which gave me the opportunity to pursue this doctoral program at LPNHE.

It is impossible to thank everyone who deserves my gratitude in such a small space, to all of those who have helped me along the way, thank you all here.

Contents

Introduction

Humans have always attempted to understand the mystery of Nature, and more recently physicists have established theories to describe the observed phenomena. The most recent theory is a gauge quantum field theory framework, called Standard Model (SM), which proposes a model comprised of elementary matter particles and interaction particles which are fundamental force carriers in the most unified way. The Standard Model contains the internal symmetries of the unitary product group $SU(3)_c \otimes SU(2)_L \otimes U(1)_Y$, describes the electromagnetic, weak and strong interactions; the model also describes how quarks interact with each other through all of these three interactions, how leptons interact with each other through electromagnetic and weak forces, and how force carriers mediate the fundamental interactions. This theoretical framework will be introduced in Chapter 1.

The SM successfully predicted the W and Z particles and the top quark, and introduces the Higgs boson being responsible for generating the mass of gauge bosons and fermions. The main topic of this thesis is search for the Higgs boson. We will introduce the status of the searches when this thesis started in 2010, and the recent observation of the Higgs boson done by the ATLAS and CMS Collaborations at the Large Hadron Collider (LHC) in 2013 [1, 2], in Chapter 1. We will also introduce the report from the CDF and D0 Collaborations about the combined evidence for a particle, with a mass consistent with that of the new boson observed at LHC [3].

This thesis presents the search by the D0 collaboration for the SM Higgs boson using events containing one isolated charged lepton ($\ell = e$ or μ), a significant imbalance in transverse energy ($\not\!\!E_T$), and two or more jets, which is part of the Tevatron Higgs combination.

The instrumental tools used to conduct this search are presented in Chapter 2. At the Fermi National Accelerator Laboratory, $p\bar{p}$ beams are generated, accelerated and collided where located two detectors. $p\bar{p}$ collisions will produce different kinds of events. To probe a specific process that is relevant to our study, we need to be able to reconstruct and identify the final state objects, as described in Chapter 3.

Since we try to search the signal of the Higgs boson decaying into a pair of b quarks, verification ("tagging") of jets originating from b quarks is important to the analysis. The b tagging algorithm is described in Chapter 4.

In Chapter 5 we introduce the modelling of the simulation, to avoid potentially false result from a mismatch between data and simulation. Here we also introduce the criteria for events preselection.

The selected events are analysed and compared to the simulation of physical processes. Chapter 6 will introduce how to reduce the background to observe the potential signal, and compare this to the selected events.

Because of the low cross section of the Higgs decay channel we considered, we apply a multivariate analysis technique to increase the sensitivity, which consequently discriminate the signal-like events from background-like ones. In Chapter 7, we also present a measure-

ment of VZ production with $Z \to b\bar{b}$ to cross check the validity of our methodology.

Before going to the final result, we describe how we assign the systematic uncertainties in detail in Chapter 8.

After all the analysis steps have been applied, no obvious excess is observed. In Chapter 9, we set 95% C.L. upper limits on the Higgs boson production cross section for masses between 90 and 150 GeV. Here we also report the latest result for searching the Higgs boson by combining all the D0 channels [4], and by combining CDF and D0 results, which led to the first evidence of the Higgs boson decay in two fermions [5].

This result of searching for the standard model Higgs boson in $\ell\nu + b\bar{b}$ final states by the D0 collaboration is based on 9.7 fb^{-1} of integrated luminosity, which is the full dataset available at Tevatron. This search has been published by the D0 collaboration in Physics Review Letter in September 2012 [6], and Physics Review D in September 2013 [7].

Chapter 1

The Standard Model and the Higgs boson

The Standard Model is the quantum field theory (QFT) based on the symmetries of the $SU(3)_c \otimes SU(2)_L \otimes U(1)_Y$ gauge group and describes all the fundamental particles and their interactions (except gravity). These particles, and their interactions are described in Section 1.1. The theoretical framework of the standard model is described in Section 1.2. The recent observation of the Higgs boson was the last remaining prediction of the SM to be experimentally verified. The current status of the Higgs boson measurements is discussed in Section 1.6.2.

1.1 The Standard Model

The Standard Model, combines the special relativity and quantum mechanics, and describes the fundamental particles and the interactions between them, except gravity. The fermions, which include quarks and leptons, are described as fundamental particles in the Standard model, as no substructure to them has been found. Gauge bosons mediate the interactions between quarks, leptons and neutrinos. The Higgs mechanism, from which originates the boson is responsible for giving mass to all fermions and bosons.

The quarks and leptons are fermions, since they have 1/2 intrinsic angular momentum, or spin, and can be divided into three generations. Among generations, only mass differs, while other quantum numbers, like spin and charge, are the same. All the fundamental fermions, with their basic physics properties are listed in table 1.1.

1.1.1 The fundamental forces

The Standard Model describes three forces: the electromagnetic, weak and strong forces. Each force is relying on a gauge group, and mediated by a bosonic field.

By exchanging a massless photon with spin 1, the electrically charged particles interact electromagnetically with each other. Strong forces are the interactions between quarks and gluons, through exchanging colour charged particles. Both quarks and gluons have colour charge, and gluons are the strong force carriers. The amplitude of strong force increases as the spatial separation between particles is growing. Weak forces are responsible for nuclear decays, and mediate interactions between all fermions. All the elementary particles, except gluons, carry a weak isospin quantum number and undergo weak interactions. The carriers

of weak force include electrically neutral Z boson and charged $W^{\pm}$, which are massive force carrier bosons. Such bosons have an intrinsic spin equal to integer multiples of $\hbar$, and are summarized in Table 1.2. In the following we will use the convention $\hbar = 1$.

1.2 The Quantum Electrodynamics Field Theory

We could start with the simplest gauge group $U(1)$. For describing the motion of a free particle with spin $\frac{1}{2}$, mass m, and spinor field ψ, the Lagrangian could be expressed as

$$\mathcal{L} = \underbrace{i\bar{\psi}\gamma_\mu\partial^\mu\psi}_{kinetic\ term} - \underbrace{m\bar{\psi}\psi}_{mass\ term} \tag{1.1}$$

, where γ^μ are the 4×4 Dirac matrices defined from Pauli matrices σ^i:

$$\gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}, \tag{1.2}$$

This Lagrangian is invariant under global phase transformation, or say the Lagrangian remains unchanged if we take the following global transformation:

$$\psi(x) \to e^{i\Lambda}\psi(x), \quad \bar{\psi}(x) \to e^{-i\Lambda}\bar{\psi}(x) \tag{1.3}$$

given Λ is a free parameter.

When considering a more general case, i.e. the Λ parameter can vary as function of space-time coordinates x, then the Lagrangian is not invariant under the local transformation: $\psi(x) \to e^{i\Lambda(x)}\psi(x)$.

To restore the invariance of Lagrangian under local gauge transformation, we replace the derivative ∂_μ by a gauge covariant derivative D_μ to get:

$$\bar{\psi}(x)D_\mu\psi(x) \to \bar{\psi}'(x)D_\mu\psi'(x) = \bar{\psi}(x)D_\mu\psi(x) \tag{1.4}$$

	Generation	Particle	Mass(MeV)	Electric charge
Quarks	1^{st}	u	2.3	$2/3$
	1^{st}	d	4.8	$-1/3$
	2^{nd}	c	1.28×10^3	$2/3$
	2^{nd}	s	95	$-1/3$
	3^{rd}	t	173.2×10^3	$2/3$
	3^{rd}	b	4.18×10^3	$-1/3$
Leptons	1^{st}	e	0.511	-1
	1^{st}	ν_e	< 0.22	0
	2^{nd}	μ	105.66	-1
	2^{nd}	ν_μ	< 0.19	0
	3^{rd}	τ	1.78×10^3	-1
	3^{rd}	ν_τ	< 18.2	0

Table 1.1: Summary of Standard Model elementary fermions [8]. Their antiparticles have the same mass with an opposite electric charge, baryonic and leptonic number. The confidence level on each limit on nuetrino mass is 95% for ν_e, 90% for ν_μ, 95% for ν_τ respectively.

Force	Gauge Boson	Mass(GeV)	Interaction range	Interacts with
electromagnetic	Photon(γ)	0	infinite	charged particles
weak	$W^{\pm}$	80.399	$\sim 10^{-18}$ m	quark, leptons, $W^{\pm}$, Z
	Z	91.1876		
strong	gluons (g)	0	$\sim 10^{-15}$ m	quarks, gluons

Table 1.2: Summary of Standard Model elementary bosons [8].

We can form the covariant derivative by introducing a vector field A_μ:

$$D_\mu = \partial_\mu - ieA_\mu \tag{1.5}$$

with e being a coupling strength constant and A_μ transforming as

$$A_\mu \rightarrow A_\mu^{'} = A_\mu + \frac{1}{e}\partial_\mu \Lambda(x)$$

Then the Lagrangian can be expressed as

$$\mathcal{L} = i\bar{\psi}\gamma_\mu D^\mu \psi - m\bar{\psi}\psi \tag{1.6}$$

$$= \underbrace{\bar{\psi}(i\gamma_\mu \partial^\mu - m)\psi}_{free\ particle\ motion} + \underbrace{e\bar{\psi}\gamma_\mu \psi A_\mu}_{coupling\ to\ A_\mu} \tag{1.7}$$

If we also introduce the field tensor to include the dynamics of the gauge field A_μ: $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$

Then the Lagrangian can be expressed as:

$$\mathcal{L}_{QED} = \underbrace{\bar{\psi}(i\gamma_\mu \partial^\mu - m)\psi}_{free\ particle\ motion} + \underbrace{e\bar{\psi}\gamma_\mu \psi A_\mu}_{coupling\ to\ A_\mu} - \underbrace{\frac{1}{4}F_{\mu\nu}F^{\mu\nu}}_{photon\ kinetic\ term} \tag{1.8}$$

From this we can see three terms. The first term is the free particle Lagrangian. The second term is added to keep Lagrangian invariance, and this term describes the interaction between the spinor field ψ and gauge field A_μ. The third term describes the dynamics of the gauge field. And one could note that there is no mass term $A_\mu A^\mu$, since it would violate local gauge invariance.

1.3 The Quantum Chromodynamics Field Theory

To describe the strong interaction, the Quantum Chromodynamics (QCD) Lagrangian was introduced by Gell-Mann and Zweig in 1963, in an analogous way to the QED Lagrangian by requiring invariance under $SU(3)$ local gauge transformations:

$$\mathcal{L} = \sum_q \bar{\psi}_q(i\gamma_\mu D^\mu - m_q)\psi_q - \frac{1}{4}G^a_{\mu\nu}G^{\mu\nu}_a \tag{1.9}$$

where the index q corresponds to one of 6 quark flavours, the index a running over the 8 colours.

Here the covariant derivative is formed as:

$$D_\mu = \partial_\mu + ig_s T_a G^a_\mu \tag{1.10}$$

where g_s is the strong coupling constant and T_a are the generators of the $SU(3)$ group. $G^a_{\mu\nu}$ is the gluon field tensor:

$$G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu - g_s f^{abc} G^b_\mu G^c_\nu \tag{1.11}$$

where f^{abc} are the structure constant of the $SU(3)$ group, defined from $[T^a, T^b] = if^{abc}T^c$.

Substitute 1.10 and 1.11 into 1.9, we get the full QCD Lagrangian:

$$\mathcal{L}_{QCD} = \underbrace{\bar{\psi}(i\gamma_\mu\partial^\mu - m)\psi}_{free\ particle\ motion} \tag{1.12}$$

$$- \underbrace{\frac{1}{4}(\partial^\mu G^a_\nu - \partial^\nu G^a_\mu)}_{gluon\ kinetic} \tag{1.13}$$

$$- \underbrace{g_s\bar{\psi}T^a G^a_\mu\gamma_\mu\psi}_{quark-gluon\ coupling} + \underbrace{\frac{g_s}{2}f^{abc}(\partial^\mu G^{a\nu} - \partial^\nu G^{a\mu})(G_{b\mu}G_{c\nu}) - \frac{g_s}{2}f^{abc}f_{ade}G_{b\mu}G_{c\nu}G_{d\mu}G_{e\nu}}_{gluon\ self-interaction\ term} \tag{1.14}$$

From this we can see that there is a gluon self interaction term, which comes from the gluons themselves carrying colour charge, and does not have an equivalent in QED, since photon not having charge.

1.4 The Electroweak Sector

The electromagnetic and weak forces can be unified within the completed gauge group $SU(2)_L \otimes U(1)_Y$, with $SU(2)_L$ representing the weak isospin space and $U(1)_Y$ representing the hypercharge space. The $SU(2)_L$ group introduces three gauge fields of $W_\mu^{i=1,2,3}$ with coupling g, and $U(1)_Y$ group introduces the gauge field of B_μ with coupling $g^{'}$.

Considering that only left-handed neutrinos interact with matter, we assume that each fermion field can be decomposed in a "doublet" left-handed and singlet "right-handed" component. Then we will have the purely right handed interaction between a singlet of a lepton field ψ_R and the B_μ field:

$$L^{singlet} = i\bar{\psi_R}\gamma^\mu(\partial_\mu + ig^{'}\frac{Y}{2}B_\mu)\psi_R$$

and the interaction between the left handed doublet ψ_L and the both groups:

$$L^{doublet} = i\bar{\psi_L}\gamma^\mu(\partial_\mu + ig^{'}\frac{Y}{2}B_\mu + ig\frac{\sigma_a}{2}W^a_\mu)\psi_L$$

And by introducing the $W^a_{\mu\nu}$ and $B_{\mu\nu}$ field tensors, we can write the kinetic term of the Lagrangian as:

$$L^{kinetic} = -\frac{1}{4}W^{\mu\nu}_a W^a_{\mu\nu} - \frac{1}{4}B^{\mu\nu}B_{\mu\nu}$$

where

$$W^i_{\mu\nu} = \partial_\mu W^i_\nu - \partial_\nu W^i_\mu + g\epsilon^{ijk}A^j_\mu A^k_\nu$$

1.5 The Higgs Mechanism

The above description is only valid for massless particles and interaction bosons, since mass terms for bosons would violate their associated gauge symmetry. This is obviously not complete because the W and Z bosons have mass. The theory is also incomplete considering that the term $m(\bar{\ell}_R\ell_L + \bar{\ell}_L\ell_R)$ in the Lagrangian is not invariant under $SU(2)_L \otimes U(1)_Y$.

To generate mass terms for weak bosons and preserve local gauge invariance in the Lagrangian at the same time, the electroweak spontaneous symmetry breaking was proposed by Higgs [9], Englert and Brout [10], and Kibble-Guralnik-Hagen [11]. Such mechanism is called the Higgs mechanism, and generates massive gauge bosons within the standard model. Higgs and Englert were recently awarded the Nobel prize in physics for proposing such mechanism [12].

1.5.1 The Scalar Higgs Field

To introduce the Higgs mechanism, we consider a locally invariant complex scalar field $\phi = \frac{1}{2}(\phi_1 + \phi_2)$, in a potential $V(\phi) = \mu^2\phi^2 + \lambda\phi^4$:

$$L = D^\mu\phi^\star D_\mu\phi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} - \mu^2\phi^\star\phi - \lambda(\phi^\star\phi)^2$$

where $D_\mu = \partial_\mu + iqa_\mu$, $F_{\mu\nu} = \partial_\mu a_\nu - \partial_\nu a_\mu$, a_μ represents the gauge field, and q is the charge.

If $\mu^2 > 0$, the potential has a parabolic shape and the ground state, also called the vacuum, corresponds to $\phi = 0$. If $\mu^2 < 0$, we get the minimum as a circle in $\phi_1 - \phi_2$ space: $\phi_1^2 + \phi_2^2 = v^2 = \sqrt{\frac{-\mu^2}{\lambda}}$. It is completely equivalent to choose either ϕ_1 or ϕ_2 to be 0, and then the other to be v; and when we choose one or the other as a point to perturbatively expand about, we spontaneously break the symmetry.

The fluctuation of the vacuum around its minimum can be expanded in terms of real field θ and h, which are field excitation along the $SU(2)_L$ directions and radial direction, such as

$$\phi \to \frac{1}{\sqrt{2}}(v + h)e^{i\theta/v}$$

$$a_\mu \to a_\mu + \frac{1}{qv}\partial_\mu\theta$$

Now we have

$$\mathcal{L}'_{Higgs} = D^\mu\phi^\star D_\mu\phi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} - \mu^2\phi^\star\phi - \lambda(\phi^\star\phi)^2 \tag{1.15}$$

$$= \frac{1}{2}(\partial_\mu h^2) - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} - v^2\lambda h^2 + \frac{1}{2}q^2v^2a_\mu a^\mu - v\lambda h^3 - \frac{1}{4}\lambda h^4 + \frac{1}{2}q^2a_\mu a^\mu h^2 + vq^2a_\mu a^\mu \tag{1.16}$$

By defining $M_H^2 = 2\lambda v^2$, Equation 1.16 now reads

$$\mathcal{L}'_{Higgs} = \frac{1}{2}(\partial_\mu h^2) - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} - \frac{1}{2}M_H^2h^2 + \frac{1}{2}q^2v^2a_\mu a^\mu - v\lambda h^3 - \frac{1}{4}\lambda h^4 + \frac{1}{2}q^2a_\mu a^\mu h^2 + vq^2a_\mu a^\mu \tag{1.17}$$

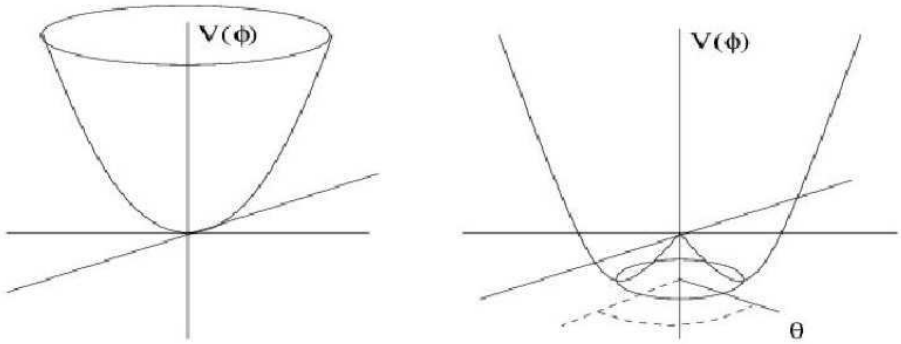


Figure 1.1: Left figure: local minimum of the Higgs potential for $\mu^2 > 0$ at $\phi = 0$. Right figure: for $\mu^2 < 0$, the vacuum acquires a non-zero value.

From here we can see that the Lagrangian have two massive fields: a massive gauge field a_μ, and a massive scalar real field h, and h is called the Higgs boson. Here $v = (\sqrt{2}G_F)^{-1/2}$, G_F is the Fermi coupling constant, measured with high precision while measuring the muon lifetime. In this equation λ is still a free parameter, which leaves the Higgs boson mass unconstrained.

1.5.2 Mass Generation for the Standard Model Particles

If we only consider kinetic term of the Higgs field Lagrangian:

$$|D_\mu \phi_0|^2 = (D^\mu \phi_0)^\dagger (D^\mu \phi_0) \tag{1.18}$$

$$= \left| (\partial_\mu - igT_a W_\mu^2 - ig' \frac{Y_H}{2} B_\mu)\phi_0 \right|^2 \tag{1.19}$$

$$= \frac{1}{2} \left| \partial_\mu - \frac{1}{2} \begin{pmatrix} gW_\mu^3 + g'B_\mu & g(W_\mu^1 - iW_\mu^2) \\ g(W_\mu^1 + iW_\mu^2) & -gW_\mu^3 + g'B_\mu \end{pmatrix} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \right|^2 \tag{1.20}$$

$$= \frac{1}{2} |\partial_\mu \phi_0|^2 \tag{1.21}$$

$$+ \frac{1}{8} v^2 g^2 \left[(W_\mu^1)^2 + (W_\mu^2)^2 \right] + \frac{1}{8} v^2 [gW_\mu^3 - g'B_\mu]^2 + 0[gW_\mu^3 + g'B_\mu]^2 \tag{1.22}$$

One can obtain mass terms by combining W_μ^i and B_μ fields:

$$W_\mu^\pm = \frac{1}{\sqrt{2}} (W_\mu^1 \mp iW_\mu^2) \tag{1.23}$$

$$Z_\mu = -sin(\theta_W) B_\mu + cos(\theta_W W_\mu^3) = \frac{gW_\mu^3 - g'B_\mu}{\sqrt{g^2 + g'^2}} \tag{1.24}$$

$$A_\mu = cos(\theta_W) B_\mu + sin(\theta_W W_\mu^3) = \frac{gW_\mu^3 + g'B_\mu}{\sqrt{g^2 + g'^2}} \tag{1.25}$$

where θ_W is known as the Weinberg mixing angle:

$$cos\theta_W = \frac{g}{\sqrt{g^2 + g'^2}} \tag{1.26}$$

$$sin\theta_W = \frac{g'}{\sqrt{g^2 + g'^2}} \tag{1.27}$$

Now using these defined fields in equation 1.22, we get

$$\left(\frac{vg}{2}\right)^2 W_\mu^+ W^{-\mu} + \frac{1}{2}\left(\frac{v}{2}\right)^2 (g^2 + g'^2) Z_\mu Z^\mu + \frac{1}{2}(0)^2 A_\mu A^\mu \tag{1.28}$$

and the mass of electroweak bosons are now:

$$M_{W^\pm} = \frac{1}{2}vg \tag{1.29}$$

$$M_Z = \frac{1}{2}v\sqrt{g^2 + g'^2} \tag{1.30}$$

$$M_\gamma = 0 \tag{1.31}$$

We can express the relation between Z_μ and A_μ by a simple rotation matrix, through equations 1.24 and 1.25:

$$\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} cos\theta_W & sin\theta_W \\ -sin\theta_W & cos\theta_W \end{pmatrix} \begin{pmatrix} B_\mu \\ W_\mu^3 \end{pmatrix} \tag{1.32}$$

The masses of fermions are introduced by a Yukawa type interaction between the fermion and the Higgs fields. This interaction is characterized by a coupling constant f between the spinor and scalar field. For example, a new term is introduced for the first generation of lepton:

$$\mathcal{L}_{FH} = f\bar{\psi}_L \psi_R \phi_0 \tag{1.33}$$

$$= f \begin{pmatrix} \bar{\nu} & \bar{e} \end{pmatrix}_L \begin{pmatrix} 0 \\ \frac{v+h}{\sqrt{2}} \end{pmatrix} e_R \tag{1.34}$$

$$= f\bar{e}_L e_R \frac{v}{\sqrt{2}} + f\bar{e}_L e_R \frac{h}{\sqrt{2}} \tag{1.35}$$

and so a mass term would appear as $M_f = \frac{f_v}{\sqrt{2}}$. It is not known if the neutrinos acquire mass within the Higgs mechanism or through another mechanism.

1.6 Higgs Searches

1.6.1 Indirect Searches

Indirect searches constrain the dependence of the Higgs boson mass through precisely measuring the mass of the W boson and top quark as done by the CDF and D0 Collaborations [13, 14], and this yields an indirect constraint on the allowed mass of the Higgs boson, $m_H < 152$ GeV [15], at the 95% confidence level.

An analysis considering the global fit over all electroweak observables is performed, which reports that $\Delta\chi^2 = \chi^2 - \chi^2_{min}$ with χ^2_{min} yielding the preferred value of the Higgs boson mass, which is

$$M_H = 94^{+29}_{-24}\ GeV \tag{1.36}$$

at 68% confidence level [15], as shown in Figure 1.2

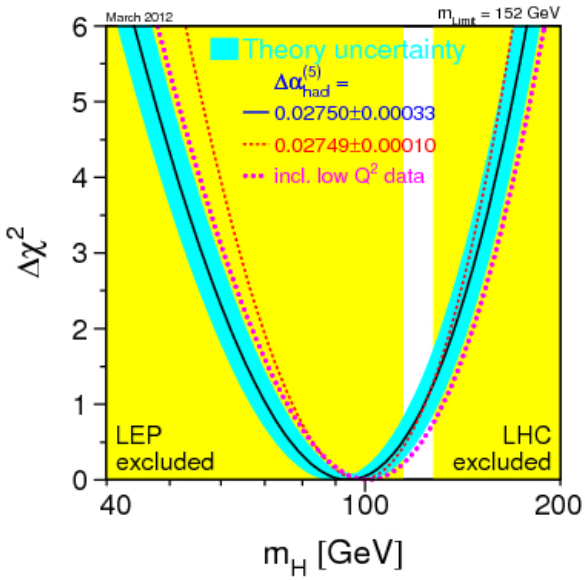


Figure 1.2: The $\Delta\chi^2$ curve derived from high-Q^2 precision electroweak measurements, performed at LEP and by SLD, CDF, and D0, as a function of the Higgs-boson mass, assuming the Standard Model to be the correct theory of nature. The preferred value for its mass, corresponding to the minimum of the curve, is at 94 GeV, with an experimental uncertainty of +29 and −24 GeV (at 68% confidence level derived from $\Delta\chi^2 = 1$ for the black line, thus not taking the theoretical uncertainty shown as the blue band into account).

1.6.2 Direct Searches

Only direct observation of the Higgs boson constitutes a proof of its existence. The LEP Higgs Working Group constrained the Standard-Model Higgs boson to be heavier than 114.4 GeV (95% confidence level limit), as indicated by the left excluded area drawn in yellow in Figure 1.2 [16].

In 2010, when I started to be involved in the search for the Higgs boson, the LHC had just started *pp* collisions and had not collected enough data to perform Higgs boson search. At ICHEP 2010, the ATLAS and CMS Higgs groups proposed simulation results, and at the same time, the Tevatron combination reported the exclusion of the Higgs mass region $158 < m_H < 175$ GeV at 95% C.L., as shown in Figure 1.3. So at the beginning of the thesis, the Higgs boson had not been discovered yet. And $114.4 < m_H < 158$ GeV was the last mass region where the Higgs could hide, if it existed.

Higgs Discovery

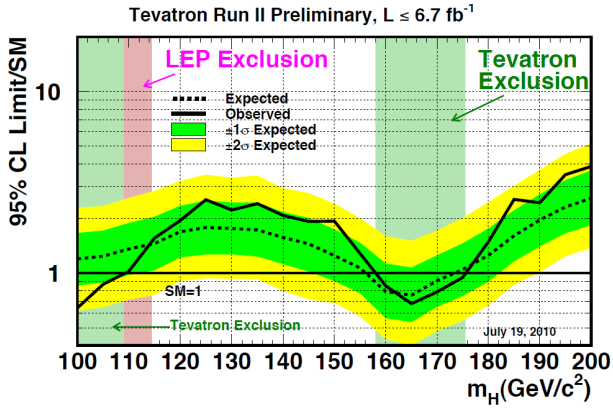


Figure 1.3: Observed and expected 95% C.L. upper limits on SM Higgs boson production, as function of the Higgs boson mass for the combined CDF and D0 analyses. The bands indicate the 68% and 95% probability regions where the limits can fluctuate, in the absence of signal.

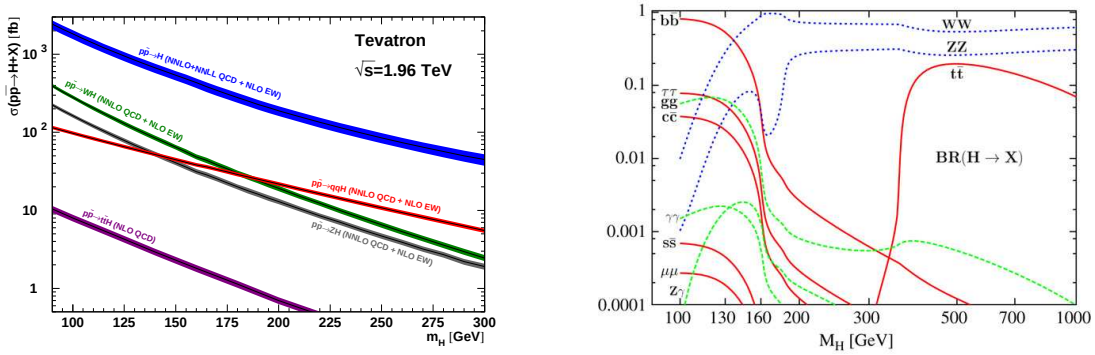
In July 2012, the ATLAS and CMS Collaborations at the Large Hadron Collider (LHC) have reported independently the observation of a new boson with mass of around 125 GeV [1, 2]. Much of the sensitivity of the LHC searches comes from gluon-gluon fusion ($gg \to H$) production with decay channels of $H \to \gamma\gamma$, $H \to ZZ^{\star}$, and $H \to W^{+}W^{-}$. Published searches for associated production $VH \to Vb\bar{b}$ at the LHC, where $V = W$ or Z [17, 18], had not yet reached sensitivity to SM Higgs boson production. In July 2012, the CDF and D0 Collaborations reported combined evidence for a particle, with a mass consistent with that of the new boson observed at LHC, produced in association with a W or Z boson and decaying to a bottom-antibottom quark pair [3].

Low Mass Higgs Search at Tevatron

Figure 1.4 show the Higgs boson cross sections and branching ratios at Tevatron [19]. The Higgs boson decays dominantly in a $b\bar{b}$ pair below 135 GeV. The highest cross section for the Higgs production is from gluon fusion (through a loop of top quarks). Since the simple $b\bar{b}$ pair final state is impossible to exploit with several orders of magnitude higher QCD production as a background, the following highest cross section modes are exploited, i.e. WH, ZH, where a W or Z boson radiates a Higgs boson. Considering that the W decays into a lepton and high transverse energy neutrino, this signature helps to exploit the WH events, the $WH \to \ell\nu b\bar{b}$ is the most advanced channel at Tevatron, and this is the major focus in this document.

The following channels are used for the search for a low mass Higgs boson at Tevatron:

- $ZH \to \nu\bar{\nu}b\bar{b}$ [20],
- $WH \to \ell\nu b\bar{b}$, which will be described in this document,
- $ZH \to \ell^{+}\ell^{-}b\bar{b}$ [21].



(a) Higgs boson cross section.

(b) Higgs boson branching ratio.

Figure 1.4: (a)Higgs boson cross sections for different production modes as function of the Higgs boson mass. The cross section values (given in pb) are reported for $p\bar{p}$ collisions at 1.96 TeV. (b)Higgs boson branching ratios for different decays modes as function of the Higgs boson mass. The search for a lower mass Higgs boson ($M_H \lesssim 135$ GeV) makes use of the $b\bar{b}$ decay mode, whereas the search for a higher mass Higgs boson ($M_H \gtrsim 135$ GeV) is more sensitive to the W^+W decay mode [22].

Chapter 2

The Tevatron and the D0 Detector

The Tevatron $p\bar{p}$ collider and the D0 detector is described in this chapter. Considering the WH analysis final states, i.e. a lepton, missing transverse energy from the escaping neutrino and jets originating from a pair b quarks, all parts of the detector are involved in the objects reconstruction and will be described here. The Tevatron acceleration chain and the production of proton and antiproton beams are described in Section 2.1. The main components of the D0 detector and data acquisition are described in Section 2.2.

2.1 The Chain of Accelerators and the Tevatron

The Tevatron is a $p\bar{p}$ collider, located at the Fermilab since 1988. Important discoveries have been made at the Tevatron such as the existence of the bottom quark from the E288 collaboration led by Leon Lederman in 1977 [23], the top quark by the CDF and D0 experiments in 1995 [24, 25] and the tau neutrino in 2000 by the DONUT collaboration [26].

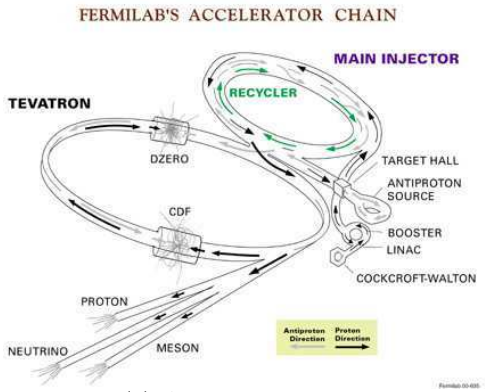
Fermilab utilizes a complex acceleration chain to provide high energy collisions, and Tevatron accelerating beams of protons and antiprotons is the most powerful one, which produces collisions at a centre-of-mass energy of 1.96 TeV. CDF and D0 are two detectors located at the interaction points where the Tevatron beams are crossing. Figure 2.1 is an aerial view, which shows the acceleration chain.

Collisions at the Tevatron occurred during two major periods:

- the "Run I" took place from 1988 to 1996, with a centre-of-mass energy of 1.8 TeV, collecting 125 pb^{-1} of data.

- the "Run II" took place from 2001 to 2011, after the Tevatron upgraded to a centre-of-mass energy of 1.96 TeV. This period is separated is two phases.

- The first part is the Run IIa data set, taken from March 2001 to March 2006 shut-down, with an integrated luminosity of about 1 fb^{-1}; the second part is the Run IIb dataset, from June 2006 to September 2011 shut-down with about 12 fb^{-1} delivered luminosity. Important detector and trigger updates have been made between Run IIa and Run IIb [28, 29], like the installation of an additional layer of detector to the Silicon Microstrip Tracker (SMT) closer to the beam pipe (more details in Section 2.2.2). The Run IIb data are collected during four main periods: the Run IIb1 dataset was taken from June 2006 to August 2007; the Run IIb2 dataset was taken from October



(a) Acceleration chain.

(b) Aerial view.

Figure 2.1: Protons are first created and accelerated at the Cockcroft-Walton pre-accelerator, LINAC and Booster. Part of the protons are used to create antiprotons. The antiprotons are stored in the Main Injector before being injected into the Tevatron. [27]

2007 to June 2009; the Run IIb3 dataset was taken from September 2009 to July 2010; and the Run IIb4 dataset was taken from August 2010 to September 2011.

The Run II dataset is summarized in Table 2.1

	Run IIa	Run IIb1	Run IIb2	Run IIb3	Run IIb4
Period	Apr. 20, 2002 - Feb. 22, 2006	Jun. 9, 2006 - Aug. 4, 2007	Oct. 28, 2007 - Jun. 13, 2009	Sept. 15, 2009 - July 18, 2010	Sept. 15, 2010 - Sept. 30, 2011
Integrated Luminosity	1.08 fb^{-1}	1.22 fb^{-1}	3.04 fb^{-1}	1.99 fb^{-1}	2.40 fb^{-1}

Table 2.1: Summary table of D0 datasets. The luminosity listed here is after requiring data quality and removing bad luminosity blocks from the luminosity system.

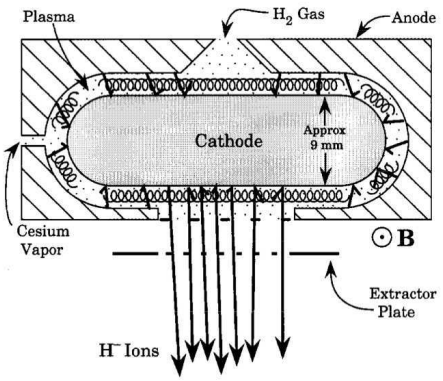
2.1.1 Proton Beam Production

The colliding protons are produced from Hydrogen gas (H_2). H_2 is injected into an ionization chamber, fully filled with plasma comprised of electrons and protons. A schematic view of a magnetron is given in Figure 2.2(a). Then the negatively charged H^- ions moves to the first step acceleration, i.e. the Cockcroft-Walton.

The Cockcroft-Walton (see Figure 2.2(b)) accelerates the H ions with 750 keV kinetic energy, in a static electric field.

The Linear Accelerator (LINAC) shown in Figure 2.3 is the next acceleration of ions. H ions are accelerated here to 400 MeV by the mean of radio frequency (RF) cavities. The motion of particles is constrained by a periodically alternated electric field, and the bunches are formed according to the frequency of the RF cavities.

The ions pass through a thin carbon foil, which strips off the loosely bound electrons, so a beam of protons are got. The protons then enter a 475 meters long synchrotron Booster, formed by 96 magnets bending the trajectory of the protons. The Booster accelerates the protons to an energy of 8 GeV, with RF cavities of 1 GHz, and it is typically filled with about 3×10^{12} protons.



(a) Basic configuration of a magnetron

(b) Cockcroft-Walton generator.

Figure 2.2: (a) Schematic view of a magnetron, producing H ions. (b) Photography of the Cockcroft-Walton. [30]

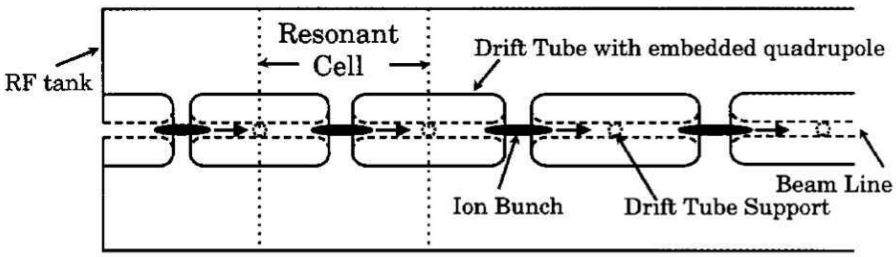


Figure 2.3: Schematic view of RF cavities in the LINAC.

2.1.2 The Main Injector and Recycler

The protons enter the Main Injector [31] after being accelerated to an energy of 8 GeV in the Booster. The main injector is a synchrotron with a 3.3 km circumference, built between the Run I and Run II. Several purposes are accomplished within the main injector: protons are accelerated to 120 GeV and partially sent to a Nickel target to produce antiprotons through the reaction $p + p \rightarrow p + p + p + \bar{p}$. The antiproton production efficiency is of the order of 3×10^6 proton on target.

Before being accelerated to 150 GeV and injected into the Tevatron, protons and antiprotons travel in the same beam pipe, surrounded by 344 dipole magnets and 208 quadrupole magnets.

Since the antiprotons' production efficiency is low, this is a limiting factor for the Tevatron integrated luminosity. The Recycler [32] was built during the Tevatron upgrade before the beginning of the Run II and was used to recover antiprotons which are present after the period, during which protons and antiprotons collide in the Tevatron.

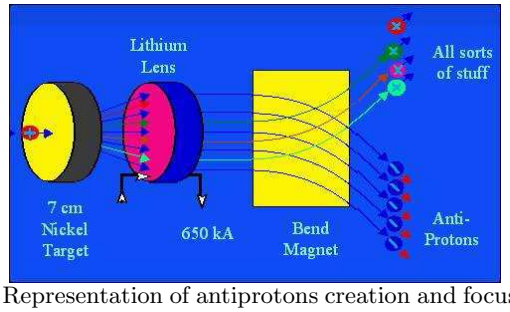
Figure 2.4: Photography of the Main Injector tunnel, where are located the Recycler (in green) and the Main Injector.

2.1.3 Antiproton Beam Production

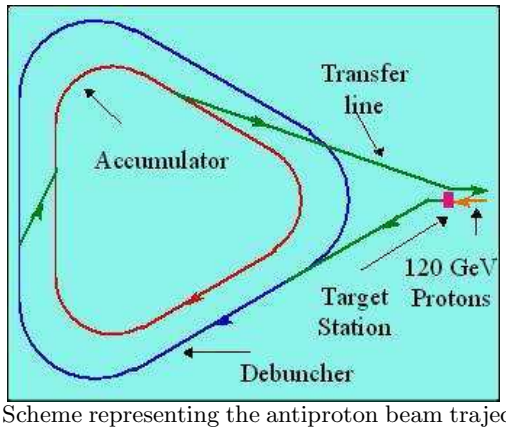
The $p\bar{p}$ Tevatron collider was designed mainly for searching top quark pair production through opposite charged particles' reactions. And antiprotons are produced using three devices: a fixed Nickel target to create, one synchrotron to accumulate, and another synchrotron to cool antiprotons.

Antiprotons are produced from collisions between protons and Nickel target, and the production rate is about one antiprotons created for three million protons on target. The

created antiprotons have 8 GeV energy and are focussed with a magnetic lens and sent to the debuncher, in order to separate the antiprotons from other particles. This process is depicted on Figure 2.5.



(a) Representation of antiprotons creation and focusing.



(b) Scheme representing the antiproton beam trajectory after production.

Figure 2.5: After protons hitting the Nickel target, produced antiprotons are focused and filtered from other particles (a). Antiprotons have then their energy spread reduced in the Debuncher.

The created antiprotons have same structure and needed to be split according to different energies. The Debuncher is a triangular synchrotron with 505 meters circumference. While travelling through the debuncher, lower energy protons moves closer to the inner part of the cavity whereas higher energy protons moves along an outer trajectory. This naturally separate the protons into different RF field intensity, with desired trajectory and speed.

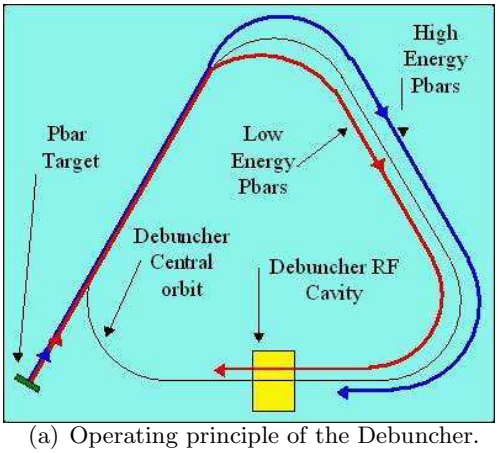
After the antiproton beams are stable, the stochastic cooling [33] process are applied. This cooling process is to reduce the transverse oscillation and consequently energy spread. The antiproton is cooled, or picked by detecting momentum fluctuations and correcting the trajectory when it travels at the vicinity of an electrode. Figure 2.6 shows a brief description of the Debuncher and how the stochastic cooling is performed.

The Debuncher and cooling process usually take 2.4 seconds, and after that antiprotons are sent to the Accumulator, where the antiproton bunch structure will be formed again and cooled further. Here the antiproton beam could stably last several days without decaying or major losses. Once the Tevatron has dumped the beam from the previous store, antiprotons stored in the Accumulator will be injected to the Tevatron.

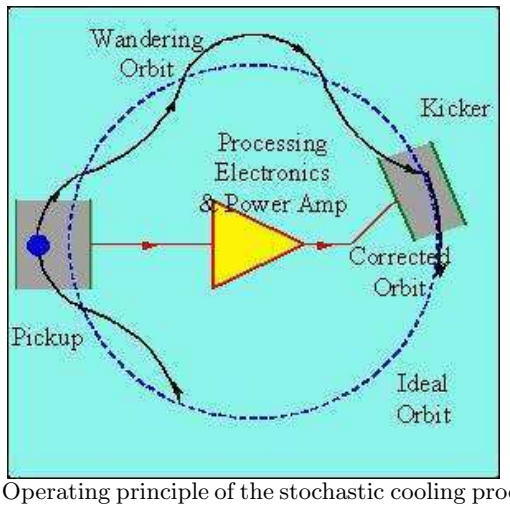
2.1.4 The Tevatron

Tevatron is the last accelerator, with CDF and D0 detector located where the p and $\bar{p}$ beams collisions take place. Tevatron is comprised of 8 accelerating cavities, 816 dipole superconducting magnets and 204 quadrupole magnets within a 1 km radius circular synchrotron. A full revolution of particles in the Tevatron is achieved in 21 μs, through all the implements: dipole magnets of a 4.2 T magnetic field made of Niobium-Titanium alloy wire, superconducting reached by cooled to 4.3 K liquid Helium temperature, and 980 GeV beams accelerated from 150 GeV by the RF cavities operating at a frequency of 53 MHz.

According to the Accelerator Division optimization studies, the most efficient way to produce highest rate collisions is to fill the Tevatron with 36 bunches of approximately



(a) Operating principle of the Debuncher.



(b) Operating principle of the stochastic cooling process.

Figure 2.6: Description of the two devices allowing the reduction of the beam spread, the Debuncher (a) and the stochastic cooling process (b).

3×10^{11} protons and about 10^{10} antiprotons at the beginning of a store, which is shown in Figure 2.7. That is separating 3 "super bunches" by 2.64 μs, each of them containing 12 bunches separated by 396 ns. p and $\bar{p}$ beams travel in the same beam pipe in an helical motion (see Figure 2.8). The collisions start when the beams reached 980 GeV and focused and halo removal is completed. 1.96 TeV $p\bar{p}$ interacts with each other at around the centre of the CDF and D0 detectors. The distribution of the luminous region along the beam axis corresponds to a Gaussian distribution with a spread $\sigma_z = 18$ cm.

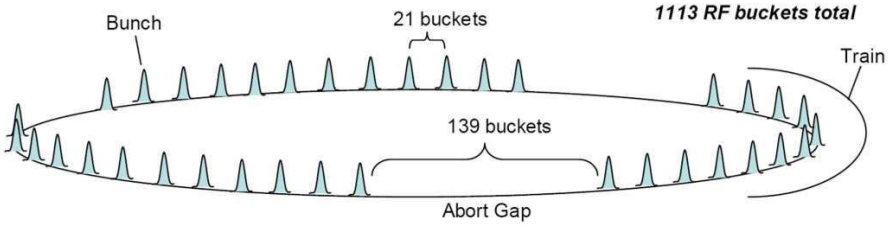


Figure 2.7: Bunch structure of the proton and antiproton beams.

Major changes have been made for continually improving the Tevatron performance, see Table 2.2. The instantaneous luminosity has been increasing also during the Run II as shown in Figure 2.9.

2.2 The D0 Detector

As one of the two detectors at Tevatron, D0 is used to record and study the outcome from the interactions of the 1.96 TeV $p\bar{p}$. The D0 detector [34, 35] is a multi-purpose detector, comprised of following parts:

- the Silicon Microstrip Tracker (SMT)
- the Central Fiber Tracker (CFT)

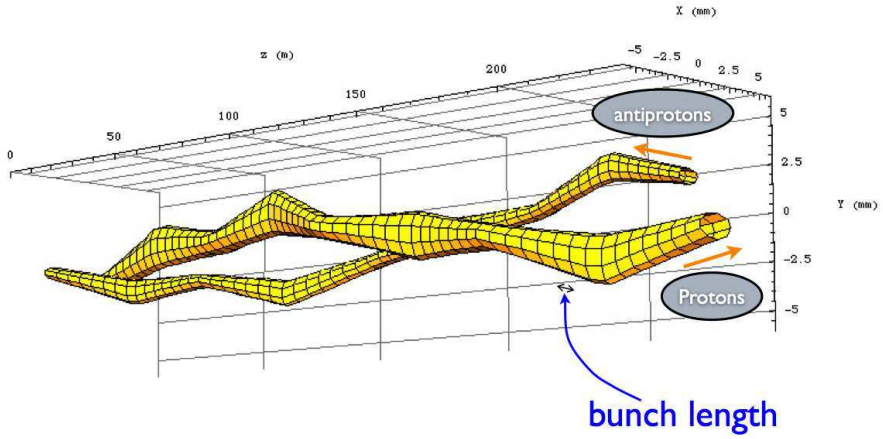


Figure 2.8: Beam envelopes along their motion axis.

	Run I	Run IIa	Run IIb
Period	1989-1996	2001-2006	2006-2011
Beam energy (GeV)	900	980	980
Bunch spacing	3.5 μs	396 ns	396 ns
Number of bunches	6×6	36×36	36×36
Protons per bunch	2.3×10^{11}	2.7×10^{11}	3×10^{11}
Antiprotons per bunch	5.5×10^{10}	3×10^{10}	7×10^{10}
Peak luminosity (cm^2s^1)	16×10^{30}	100×10^{30}	$200 - 400 \times 10^{30}$
Integrated Luminosity	0.16 fb^{-1}	1 fb^{-1}	~ 11 fb^{-1}

Table 2.2: Summary table of Tevatron characteristics for the Run I and Run II periods.

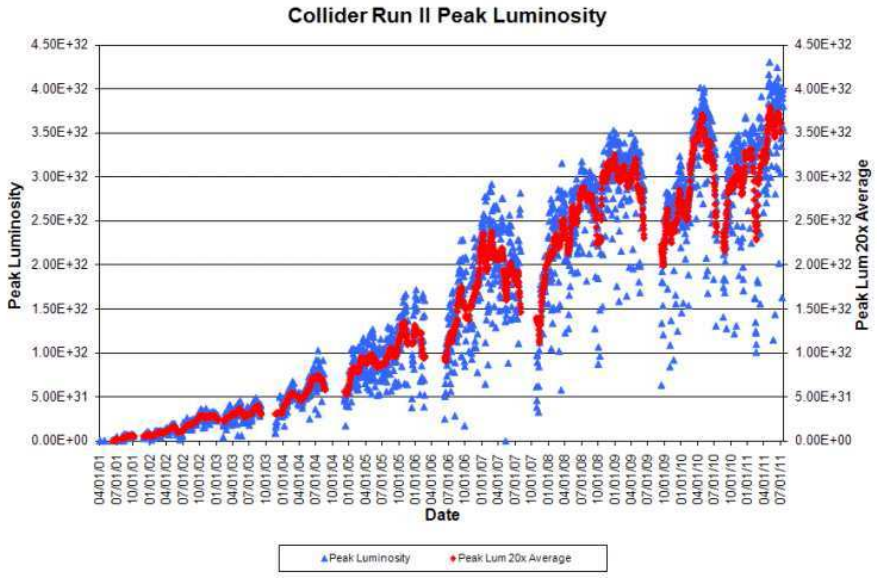


Figure 2.9: Peak instantaneous luminosity, reached in beginning of store, shown as function of time.

- the superconducting solenoid magnet
- the preshower detector
- the electromagnetic and hadronic calorimeter
- the muon spectrometer
- the muon toroid

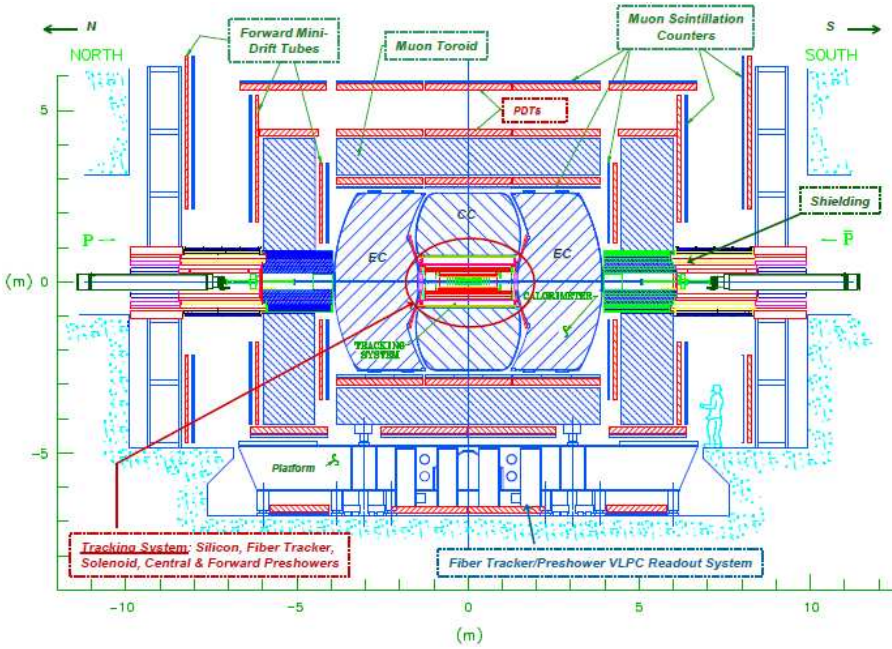


Figure 2.10: Overview of the D0 detector. [35]

The D0 detector was upgraded several times during the transition among run periods, according to new detecting and recording requirement. During the transition between the Run I and Run II, the whole central tracking system was replaced to introduce a solenoid magnet for measuring the momentum of charged tracks. During the transition between Run IIa and Run IIb, an additional layer of tracking detector was placed at the innermost part of the detector to improve track resolution and b jets identification (bID).

We use the datasets from the Common Samples Group (CSG) EMinclusive and MUinclusive skims for the electron and muon channels, respectively [36]. And to properly account for the detector responses among different running period, we separately treat the datasets, listed in Table 2.1, and use different MC samples on the corresponding datasets, owing to differences between the epochs such as vertex algorithms, instantaneous luminosity profiles, and tracking detector efficiency.

2.2.1 Coordinate System

To locate objects in the D0 detector, we introduce the following coordinate system : a right-handed coordinate system with the positive z-axis along the proton travelling direction, as shown in Figure 2.1(a), and the positive y-axis pointing upward. The (x, y) plane is called

the transverse plane, and all the "transverse" physics quantity we used in this document is defined as here. The angles ϕ and θ are the azimuthal and polar angles, respectively, and the r coordinate denotes the perpendicular distance from the z axis.

The rapidity $Y = \frac{1}{2}ln[\frac{E+p_z}{E-p_z}]$ is a useful quantity, which shows the angle between the $x-y$ plane and the direction of a flying particle. And we often use pseudorapidity $\eta = -ln[tan(\theta/2)]$ for approximation, where E and p_z are the energy and longitudinal momentum of an object, in the limit that $m/E \to 0$. And usually, $\eta \sim 0$ at the central part of D0 detector, and the term "forward" refers to objects in the regions where $|\eta| \gtrsim 1$.

2.2.2 Tracking and Vertexing System

The Silicon Microstrip Tracker (SMT) and the Central Fiber Tracker (CFT) are the two major component of the tracking system, as shown in Figure 2.11. These two tracks are embedded in a magnetic field generated by the solenoid magnet.

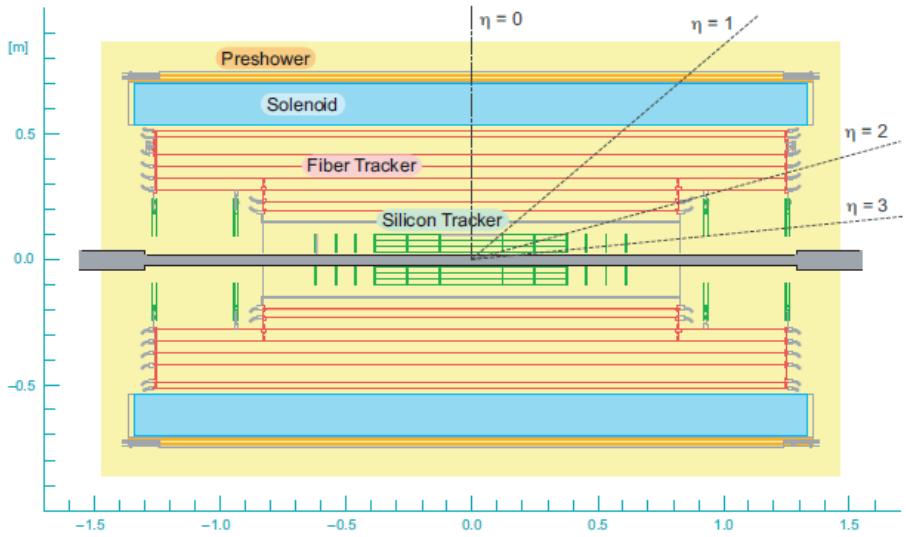


Figure 2.11: Overview of the D0 tracking system. [35]

Layer 0

During the 2006 shut-down, an additional layer (Layer 0) was introduced in the SMT [29], where is the closest point to the beam. This upgrade is to achieve three improvements:

- recover tracking performances coming from radiation damages,
- achieve better track finding efficiency, which is deteriorated at higher instantaneous luminosity,
- improve impact parameter (IP) resolution (see Figure 2.12 [37]), IP is the closest distance between the charged track and the primary vertex, impacting b-tagging performances.

There's an annular space ("gap") between the beam pipe and the first layer of the SMT, with a radius of 22.90 mm, all the detector cables and support structure placed here need to fit in this gap.

Just as the SMT, Layer 0 is also built with silicon microstrips sensors. As shown in Figure 2.12, the implementation of the Layer 0, improves the IP resolution for tracks $p_T < 5$ GeV and consequently improves 15% in b-tagging performance with respect to the beginning of Run II.

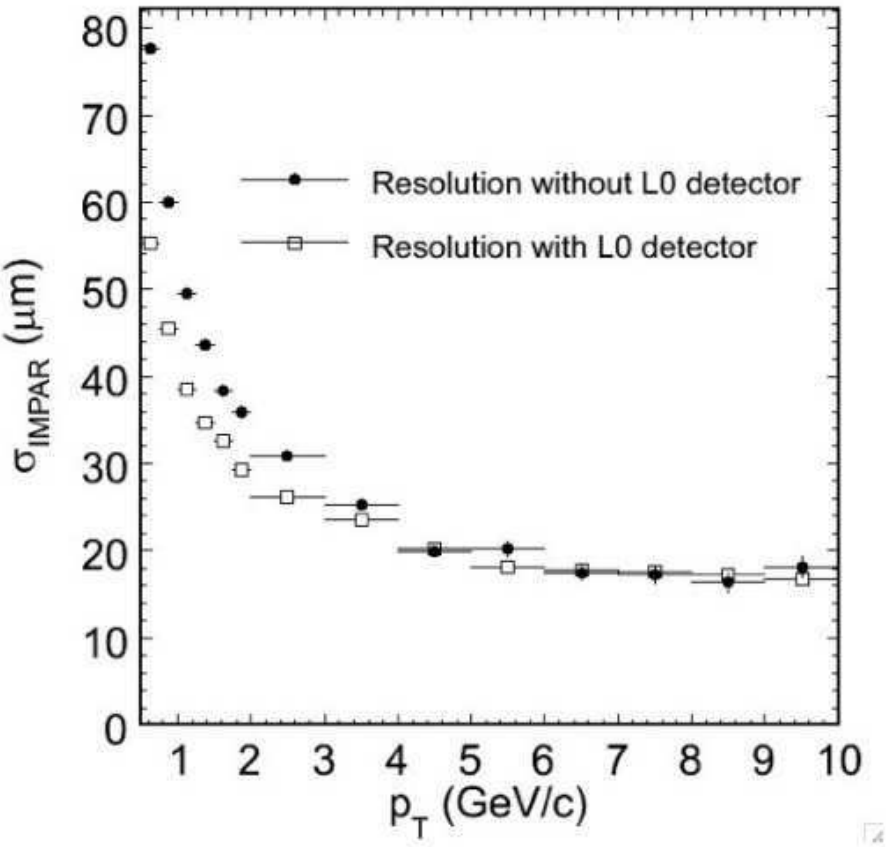


Figure 2.12: The impact parameter resolution with and without the Layer 0 detector. [37]

Silicon Microstrip Tracker

The Silicon Microstrip Tracker (SMT) is located at the innermost part of the detector (just directly outside of the beam pipe), and is bended by a 1.92 T magnetic field. The silicon-based detector SMT has a high resolution and fast response, and is crucial for reconstructing primary vertices (PV), measuring the lepton transverse momentum, identifying b quark jets, and therefore crucial for the whole range of the physics analysis, like: b physics, searching for New Physics and Higgs boson.

The SMT is constructed from barrel and disk modules, the barrel module is parallel to the beam pipe to measure the tracks perpendicular to the beam with $r-\varphi$ coordinates, and the disk modules measure $r-z$ as well as $r-\varphi$ components. The central part of the SMT is comprised of two sub-parts:

- 6 barrels, each of them comprised of 72 modules ordered in 4 layers. They are arranged along the z axis. The distance with respect to the beam pipe is 2.7, 4.5, 6.6 and 9.4

cm respectively for each layer. Silicon strips have either a 90° or 2° angle between the axial (parallel to the z axis), and the stereo sides.

- 12 "F-disks", where 6 are located between the barrels except for the two most central ($|z| = 12.5, 25.3$ and 38.2 cm) and the remaining are placed at higher $|\eta_{det}|$ region ($|z| = 43.1, 48.1$ and 53.1 cm). These disks are for detecting forward tracks.

The forward disk part of the SMT covers the $60 < |z| < 100$ cm or $|\eta_{det}| < 3$ region, and is comprised of two "H-disks" arranged on each side of the detector at $|z| = 100.4$ and 121.0 cm. Their diameter is larger than the F-disks in order to cover the $|\eta_{det}| > 2$ region. Each section of the disks covers 15° in the transverse plane.

Disks are comprised of semi-conductive doped silicon sensors. Charged particles create electron-hole pairs while passing through the sensors, and triggers signals which are detected by collecting drifting charges on corresponding voltage difference in the sensors.

The SMT achieves a spatial resolution of around 20 μm for hits. And an isometric view of the SMT is given in Figure 2.13.

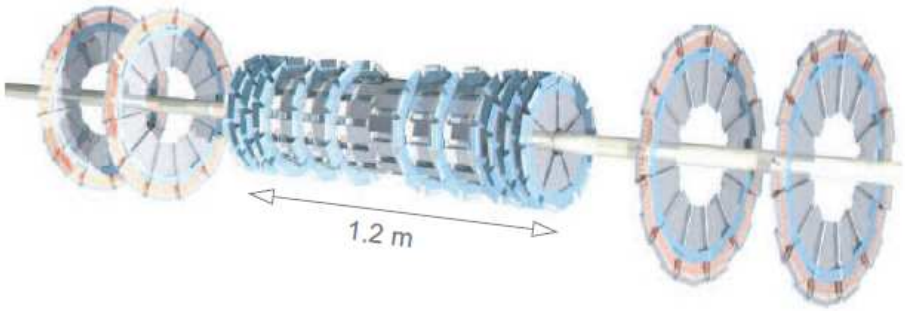


Figure 2.13: Overview of the SMT. [35]

Central Fiber Tracker

The Central Fiber Tracker (CFT) is located just outside around SMT, and used to detect the charged particles momentum and track the trajectory. CFT covers up to $|\eta_{det}| \lesssim 1.7$, with a 1.66 m long central part, and 2.52 m long outer part. The CFT is constructed from eight concentric layers of scintillating fiber, and such layers are arranged in 8 cylinders of 4 fibers, 2 of them aligned along the z direction, whereas the other layers of fibers have a stereo angle of $+3°$ or $-3°$, respectively denoted as u and v layers. And when the charged particles travels through the fibers, they excite the fiber core and photons are emitted with a wavelength of 340 nm. The photons propagate down the internally reflective fibers, where they are read out by photon counters which can detect a single photon.

Solenoid

A superconducting solenoid magnet was added after Run I to measure the momentum of tracks, this was cooled down by liquid helium and with a size of 2.73 m in length and of 1.42 m in the outer diameter. The Solenoid generates a 2 T magnet, which is uniform at the 0.5% level, this is to remove potential systematic effects in the tracking, the polarity of the D0 solenoid magnet is regularly reversed (every two weeks). The periodic reversal of

the D0 solenoid and toroid polarities also results in a cancellation at the first order of most detector-related asymmetries.

2.2.3 Preshower Detectors

The Preshower detector is a thin layer of scintillators located in front of the calorimeter, helping in electron identification and hadrons/electromagnetic particles discrimination during both triggering and offline reconstruction. There are 3 preshower detectors. The Central Preshower detector (CPS) is located in the central part ($|\eta_{det}| < 1.3$), between the solenoid magnet and the central calorimeter. Two Forward Preshower detectors (FPS) are located in the forward region ($1.5 < |\eta_{det}| < 2.5$) in front of the end calorimeters, around the luminosity monitors (see Section 2.2.6).

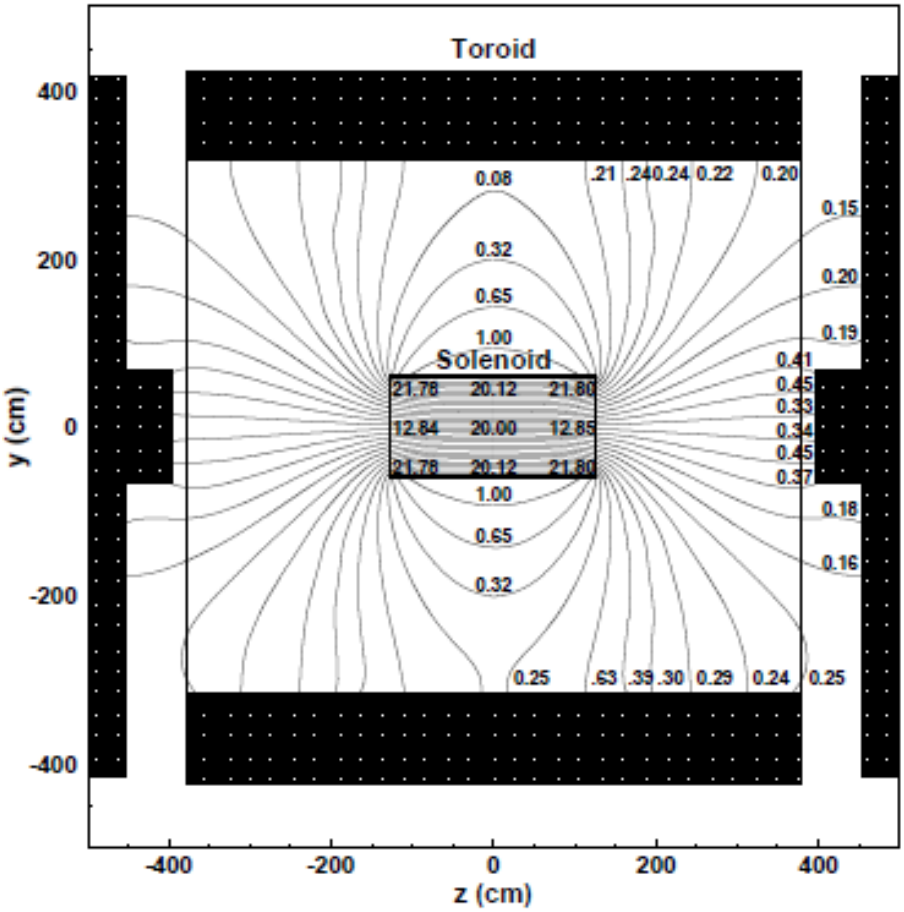


Figure 2.14: Side view of the magnetic field measured in kG. Lines are representing the field lines in $y - z$ plane. [35]

Both preshower detectors are made from triangular strips of scintillators, as shown in Figure 2.15.

2.2.4 Calorimeter

The calorimeter is used to measure the energy of photons, electrons, and hadrons, as well as any imbalance in the transverse energy from which we infer the existence of neutrinos,

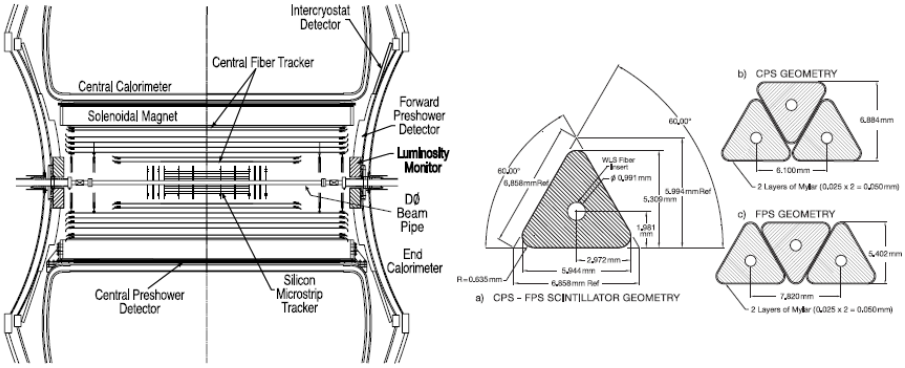


Figure 2.15: Left figure: The CPS detector is located in the gap between the solenoid magnet and the calorimeter. FPS detectors are placed between the CFT and the End Calorimeters. Right figure: Preshower detectors geometry. [35]

which are crucial for the WH analysis.

While a particle travels the material and interacts through atomic or nuclear interactions, it creates a "shower" which can be measured by the calorimeter to get the energy. The D0 calorimeter consists of Uranium and liquid Argon, as absorbing and acting material respectively. The sampling fraction of the calorimeter, defined as the ratio of energy deposited in the active medium to the total energy deposit in the calorimeter and determined by simulation, is approximately 5%, depending on the Uranium/Argon ratio in the different parts of the detector.

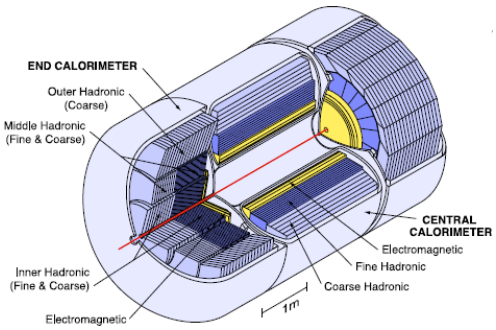
Different part of the calorimeter is composed with different materials, and corresponds to different kinds of interactions, which could differentiates different particles.

- Photons and electrons initiate electromagnetic shower, corresponds to the inner electromagnetic calorimeter (EM). Here particles are produced through either pair production ($\gamma \to e^+e^-$) or bremsstrahlung ($e^\pm \to e^\pm\gamma$). When describes the electromagnetic cascades, the radiation length X_0 is of interest, and X_0 is the length over which the electron's energy is reduced to $1/e$ of it's original energy through bremsstrahlung, or 7/9 of the mean free path for a photon to produce an electron-positron pair. The EM layers of the calorimeter have a total depth of about 20 X_0.

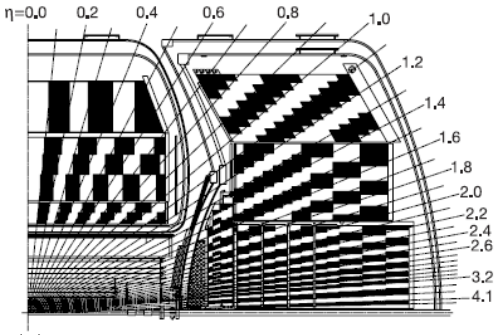
- Hadrons mainly undergo strong interactions and deposit their energy deeper in the hardonic calorimeter. The hadronic part of the calorimeter is further divided in two parts: the fine and coarse hadronic calorimeter, where the granularity is lower for the latter. Hadronic particles deposit most of their energy in the calorimeter, by hitting nuclei, a spray of pions, protons and fragments of the nuclei will be created. Similarly to X_0, the nuclear interaction length is defined as the mean free path of hadron before undergoing an inelastic collision with a nucleus. The absorption length is generally larger than the radiation length, so hadronic particles tend to deposit energy deeper in the calorimeter than electromagnetic particles do.

The calorimeter is also segmented longitudinally in two main parts: the central calorimeter (CC) and two end calorimeters (EC) located in the forward regions[1]. A unit cell of the D0 calorimeter is shown in figure 2.17.

[1]EC calorimeters are commonly referred as North and South calorimeters.



(a) Isometric view of the calorimeter.



(b) Side view of a portion of the calorimeter.

Figure 2.16: Overview of the D0 calorimeter. [35]

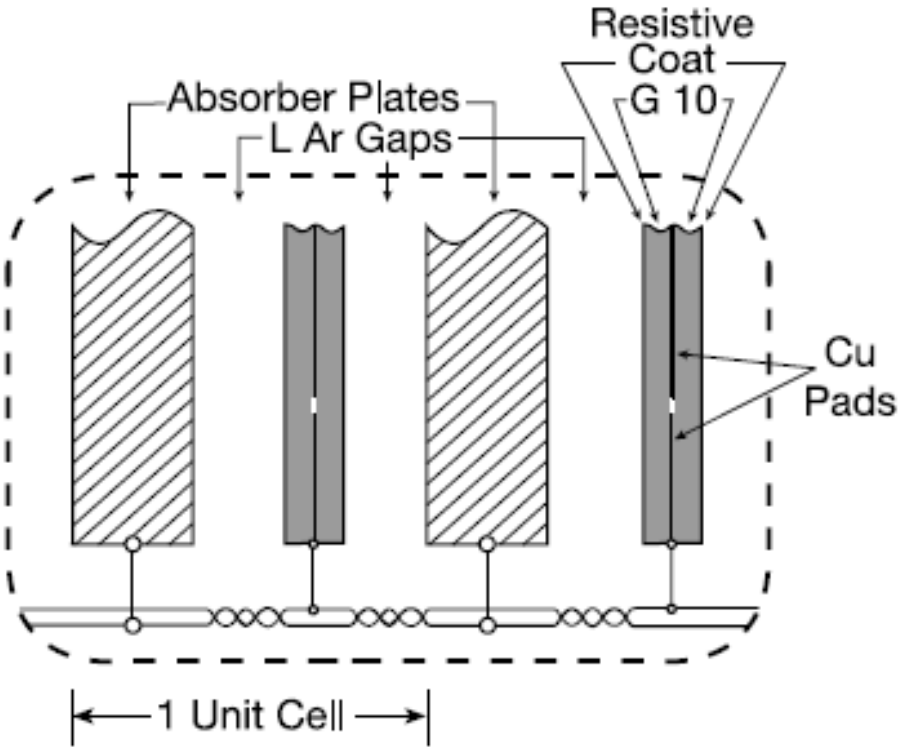


Figure 2.17: Schematic view of a unit cell of the D0 calorimeter. [35]

Central Calorimeter

The central calorimeter (CC) covers up to $|\eta_{det}| \lesssim 1$. In the electromagnetic region, 3 mm thick Uranium (short radiation length makes the calorimeter compact) plates are used as absorbers. In the fine hadronic calorimeter region, 6 mm thick plates made from Uranium-Niobium alloy, whereas in the coarse hadronic region, 46.5 mm thick Copper plates are used as absorbers.

End Cap Calorimeters

Two identical End Cap calorimeters (EC) are located on both sides of the CC, denoted as "North" and "South" respectively. EC cover up to $0.7 \lesssim |\eta_{det}| \lesssim 4$ region, and have similar structure of the CC, except stainless steel plates are used as absorber here.

Opposed to CC is subdivided as fine and coarse hadronic calorimeter, EC is subdivided into three parts:

- the inner hadronic calorimeter, with four 1.1 λ_f thick layers,
- the middle hadronic calorimeter, with four internal layers having a thickness of 0.9 λ_f and an external layers with 4.4 λ_f,
- the single-layered outer hadronic calorimeter, with 4.1 λ_f thick,

Intercryostat Detector

An intercryostat detector (ICD), consisting of scintillating tiles, is placed in the narrow $0.8 < |\eta_{det}| < 1.4$ region, where attached to the exterior surfaces of the end cryostats (see Figure 2.15). Each $\Delta\eta \times \Delta\phi = 0.3 \times 0.4$ tile is subdivided. Hence, the granularity in this region is 0.1×0.1. The cryostats are used to cool down CC EC parts to 90 K. And here cryostats are acting as absorbers.

Energy Resolution of the Calorimeter

The energy resolution σ_E in sampling calorimeter can be expressed in the following way:

$$(\frac{\sigma_E}{E})^2 = C^2 + \frac{S^2}{E} + \frac{N^2}{E^2} \tag{2.1}$$

Here the characterized parameters are:

- the constant term C, determined by non-uniformity of material thickness, non-uniformity in charge collection, mechanical imperfections, fluctuations in the amount of the upstream energy deposit and shower leakage. C is independent of the energy and dominant at the high energy regime.
- the sampling term S, determined by the absorber and active material, the thickness of sampling layer. S represents the statistical fluctuations in the amount of measured energy.
- the noise term N, determined by the electronic noise in the detector. N is dominant at low energy.

A detailed explanation of each term is given in [35] and [38].

The resolution for the D0 calorimeter are given in Table 2.3. They are determined separately for electrons and jets in different calorimeter regions [38, 39]. Figure 2.18 shows the jet energy resolution for central jets. In order to emphasize the energy resolution dependence on the parameters, we also show simple examples when one of the jet energy parameters are multiplied by a factor 2 in Figure 2.18.

	e (CC)	jet ($\|\eta\| < 0.4$)	e (EC)	jet ($2.4 < \|\eta\| < 2.8$)
C	0.030 ± 0.001	$\simeq 0.049$	0.028 ± 0.001	$\simeq 0.068$
S ($\text{GeV}^{1/2}$)	0.24 ± 0.01	$\simeq 0.757$	0.27 ± 0.01	$\simeq 0.523$
N (GeV)	0.47 ± 0.09	$\simeq 0.860$	0.73 ± 0.07	$\simeq 0.860$

Table 2.3: Energy resolution parameters for electrons and jets in Run II.

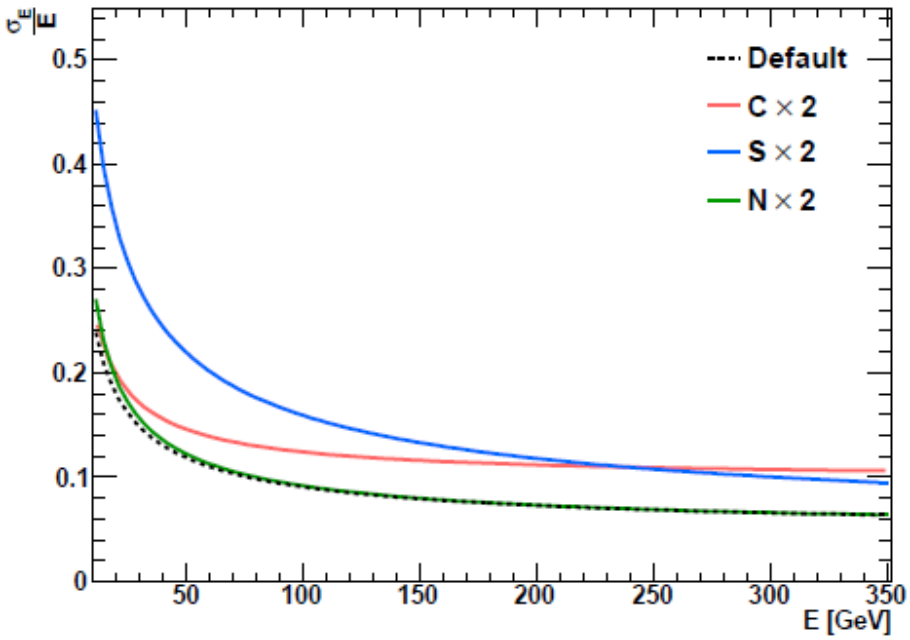


Figure 2.18: Energy resolution for jets with $|\eta| < 0.4$ as function of the jet energy. The doted black line is obtained from the parameters quoted in Table 2.2. Coloured lines are obtained by multiplying S (in red), N (in blue) and C (in green) by a factor 2 independently of the other parameters. [35]

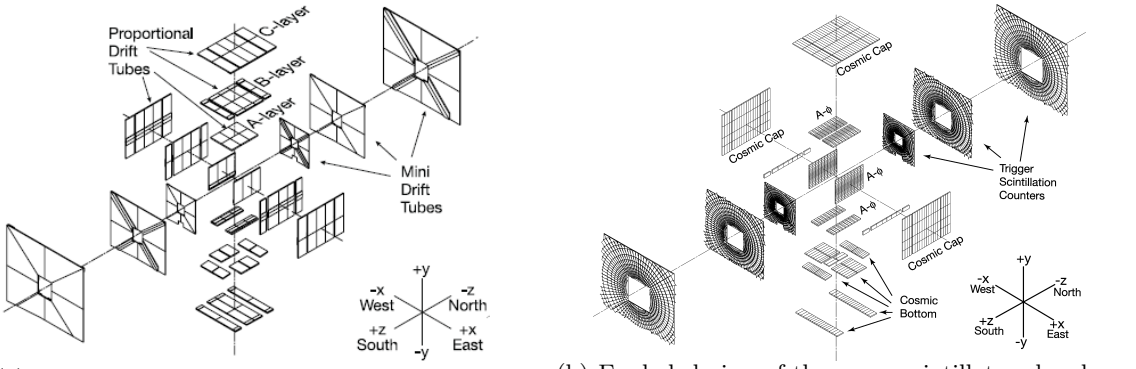
2.2.5 Muon System

A specific Muon System is necessary since muons would not be identified or measured correctly since little of their energy is deposited in the calorimeter. The muon spectrometer is located at the outmost of the detector, covers up to $|\eta| \simeq 2$, and consists of the following elements as shown in Figure 2.19:

- drift chambers,
- scintillators chambers,

- toroidal magnets generating a 1.8 T magnetic field. Similarly to the solenoid magnet, the polarity is reversed about every two weeks.

The muon system is built in 3 planes called layer A, B and C. The layer A is placed between the calorimeter and the toroidal magnets, where layers B and C are in the outermost part of the detector. There is a gap at the bottom of the layer A as visible in Figure 2.19, where the calorimeter support is located. Since the detector coverage is limited over the ϕ angle, the gap is also called "phi-hole" or "phi-crack". The counters overlap an average of about 3%in ϕ to reduce the possibility of muons passing through such crack. [35]



(a) Exploded view of the muon drift chambers.

(b) Exploded view of the muon scintillator chambers.

Figure 2.19: Exploded view of the muon drift chambers (a) and muon scintillator chambers (b). [35]

Muon Drift Chambers

Proportional Drift Tubes (PDT) and Mini Drift Tubes (MDT) are the two techniques used in the central and forward muon drift chambers.

PDTs cover up to $|\eta_{det}| \lesssim 1$ region and are divided in cells: 96 cells in layer A and 72 for layer B and C. Each cell is composed of an anode wire connect to a 4.7 kV high voltage and 2 pads operating with a high voltage of 2.3 kV. A mixture of gas (84% Argon, 8% methane, and 8% CF4) will be ionized if muons travel through, and consequently secondary electrons will be created, and collected by wires. The drift velocity is approximately 10 cm/μs, for a maximum drift time of about 500 ns.

MDTs are similar to PDTs and also arranged in ABC layers in the $1.0 < |\eta_{det}| < 2.2$ region. They consist of tubes divided in 8 cells, each cell contains a W-Au anode wire embedded in an aluminium compartment. A 90% CF_4 - 10%CH_4 gas mixture will be ionized if muons travel through and consequently created charges will be collected by the wires. The charge drift times is between 40 and 50 ns.

Muon Scintillator Chambers

Muon Scintillator Chambers (MSC) send signal to photomultipliers. The most central part of MSC is covered by the "$A\phi$" layer in the $|\eta| < 1$ region, upstream to the A layer and are used for triggering, their timing response being of the order of 2 ns. Since this part locates before the toroidal magnets, it is also helpful in identifying the stopped muons here. The fast timing detector located in layer C is used to associate a muon in a PDT with the

appropriate bunch crossing and discriminate against the cosmic ray background. Finally, scintillators in the forward region ($1 < |\eta_{det}| < 2$) are used for triggering.

2.2.6 Luminosity Monitor

Since the number of particles in each bunch cannot be measured when protons and antiprotons are present in the Tevatron, Luminosity Monitors (LM) are used to measure the integrated and delivered luminosity, by estimating the number of inelastic collisions.

The instantaneous luminosity can be expressed as function of the parameters of the accelerator:

$$L = \frac{fBN_pN_{\bar{p}}}{2\pi(\sigma_p^2 + \sigma_{\bar{p}}^2)} \cdot F \tag{2.2}$$

where $f \simeq 43.7$ kHz is rotation frequency of the bunches, B =36 being the number of bunches circulating, $N_p(N_{\bar{p}})$ the number of protons (antiprotons) per bunch, $\sigma_p(\sigma_{\bar{p}})$ the transverse proton (antiproton) beam size and F being a form factor depending on the bunch length.

Luminosity Monitors are placed around the beam pipe, at each extremity of the detector ($z = \pm 140$ cm, which corresponds to the forward region $2.7 < |\eta_{det}| < 4.4$).

The Luminosity Monitor (LM) detector consists of two arrays of 24 plastic scintillation counters with photomultiplier tube (PMT) readout, scintillation light produced in the Bicron BC-408 scintillator is detected by Hamamatsu [40] R5505Q finemesh PMTs.

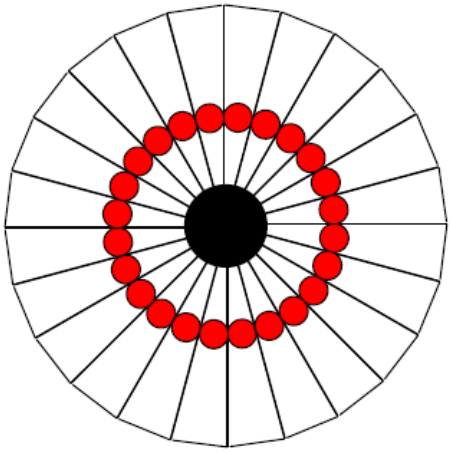


Figure 2.20: Schematic view of the luminosity monitor geometry and location of photomultipliers (red circles). [35]

A technique called the "counting empties" is used to count the number of crossings for which no collisions occurred, since it would be difficult to count the same crossing when multiple interactions may occur.

The uncertainty of the integrated luminosity includes a dominant contribution of 4.2% stemming from the uncertainty on the luminosity constant. [41]

2.2.7 Trigger and Data Acquisition System

The time between two bunch crossing being of 396 ns, the frequency at which events would be recorded by the D0 detector would be around 2.5 MHz. Recording such amount of data is not achievable and D0 adopts three levels of filters, or triggers passing about 100 events per second to be considered for analyses.

Events not used in analyses, low p_T multijet events for example, are rejected and would not pass trigger system. Therefore such multiple layers (see Figure 2.21) trigger system could reduce the acceptance and recording rate, and each layer relying on detector information as shown in Figure 2.22.

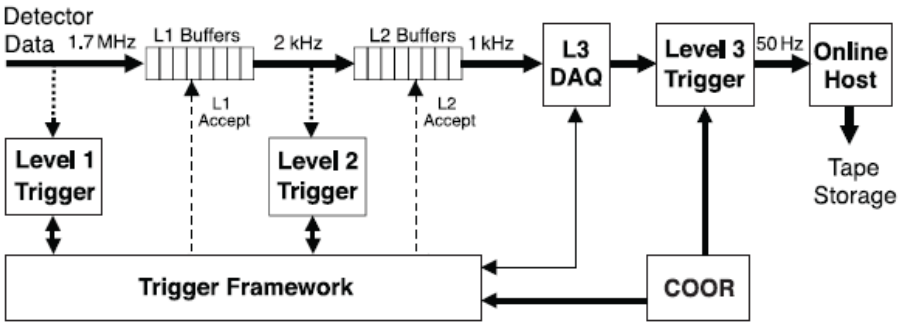


Figure 2.21: Data flow in the D0 trigger system. The event rate is reduced from 1.7 MHz to 100 Hz. [35]

Only inelastic collision events are triggered. Therefore both luminosity monitors are required to have hits in coincidence with the bunch crossing, i.e. the "Minimum Bias" events.

Sets of conditions are defined in order to "label" events. The data can be then described by logical "and" and "or" of several trigger conditions.

Level 1 Trigger

The first stage (Level 1 or L1) comprises a collection of hardware trigger elements that provide a trigger accept rate of about 2 kHz. The Level 1 relies on the fast electronics readout of parts of the detector (CFT, preshower detectors, calorimeter, muon spectrometer).

CFT information is used to reject high p_T fake tracks, by comparing recorded data to predefined models of tracks in 4.5° angular sections of the CFT.

Calorimetric triggers rely on reconstructed "trigger towers" in $\Delta\phi \times \Delta\eta = 0.2 \times 0.2$, passing above predefined energy thresholds.

Muon triggers are based on the matching of tracks produced in the CFT, given that they fulfil the above mentioned requirement, and hits obtained at least in the A or B layer of the muon system.

Level 2 Trigger

Since all parts of the detector and complex algorithms are used, and a fast reconstruction of electrons, photons and jets are performed for the Level 2 trigger, this level takes less than 100 μs and reduces the rate from 2 kHz to 1 kHz.

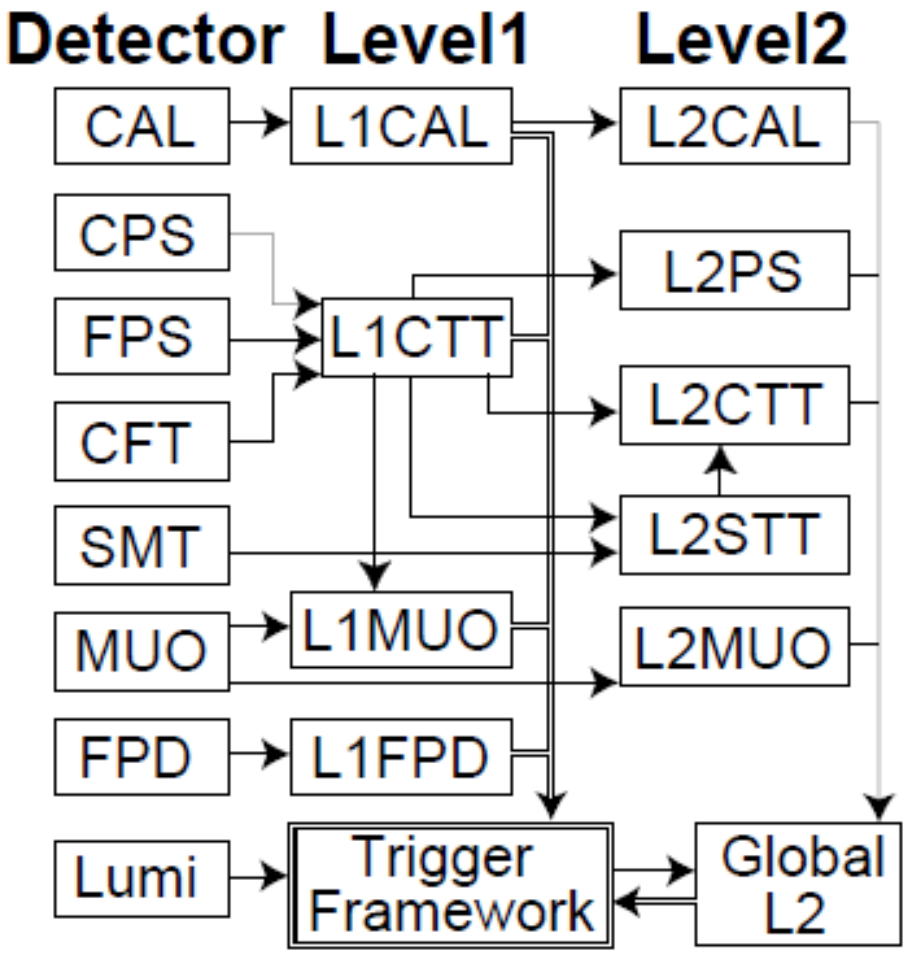


Figure 2.22: Level 1 and 2 of the D0 trigger system. Arrows represent the data stream. [35]

Jets (electromagnetic objects) are reconstructed in groups of 5×5 (3×3) trigger towers. Electrons and photons can be identified if a Level 1 track is match to clusters in the preshower detectors. An estimate of the missing transverse can be obtained by these "raw" calorimetric objects. At least 3 SMT hits around tracks found at the Level 1 are required here, and consequently gets a good impact parameter resolution. Combined with additional detector informations from the wire chambers and scintillator chambers, the muon Level 2 trigger uses information from Level 1 and gives a better precision of the object coordinates (η,ϕ) and transverse momentum.

Before events are passed to the Level 3, correlations between objects are checked.

Level 3 Trigger

The Level 3 Trigger reduces the event rate down to around 100 Hz, with decision time being less than 150 ms. Based on the events passed Level 1 and 2, a partial event reconstruction is performed as much sophistication as is done in analyses. Here tracks are used to locate the primary vertex, secondary vertices can also be identified. The output of the Level 3 trigger is the electronic signal from all detector parts, to be sent to the processor which will reconstruct objects as it is detailed in Chapter 3.

2.2.8 Data Format and Detector Simulation

The format of the data ("raw data") after passing the 3 levels of the trigger system is the electronic signal coming from all parts of the detector. To analyze the raw data, objects are reconstructed using the d0reco[1] software.

In simulated events obtained from Monte Carlo event generators, objects are reconstructed as well after detector simulation, using the d0gstar[2] software, which is based on the description of the structure and electronics of the D0 detector by Geant3. Simulated events are in the same format as raw data, and reconstructed in a similar manner. Randomly selected events from Zero Bias data are overlayed to the simulation, to reproduce multiple bunch crossing in simulation and detector noise.

A common format, *CAF* [42], is used across the D0 collaboration to store the data and simulation events in ROOT files, with sufficient informations from detector parts. To perform a specific analysis, it is unnecessary to use the big files containing irrelevant events, so a splitting using a logical OR of several basic object selection and trigger requirements would be reasonable. In the WH analysis, the *EMinclusive* and *MUinclusive* skims are used respectively for the electron and muon channels.

[1] standing for D0 offline RECOnstruction program

[2] standing for D0 Geant Simulation of the Total Apparatus Response.

Chapter 3

Reconstruction and Identification of Leptons, Jets, and $\not\!\!E_T$

In the WH analysis, the final state include one lepton, $\not\!\!E_T$, and 2 or 3 jets. Reconstruction and identification of these objects are the primary steps for reconstructing the W and H. We use the physics quantities recorded by the D0 detector to reconstruct these objects. In section 3.1 we introduce how we reconstruct particle tracks from hits in the CFT and SMT tracking detectors. In section 3.2, we describe how the vertices are constructed from tracks originating from the corresponding vertices. In following sections, we describe how we construct all the essential objects one by one from the data recorded by the D0 detector.

3.1 Tracks

Since tracks are used to reconstruct objects, like vertices and leptons, it is crucial to identify tracks properly. Two different algorithms are used for the D0 SMT and CFT systems: the Histogram Track Finder (HTF) [43] and the Alternative Algorithm (AA) [44], both using the Kalman Fitter technique [45, 46].

The HTF method is based on recognizing the motion of a particle, which can be described with its track curvature: $\rho = \frac{qB}{p_T}$. This corresponds to a particle, with charge q and transverse momentum p_T, travelling through a magnetic field B. When a hit is found in the tracking system at (x_h, y_h), while the interaction vertex is assumed at $(x, y) = (0, 0)$ in the same coordinate system. Several tracks could be drawn between the vertex and the hit, and could be transformed into (ϕ, ρ) coordinate system. Each hit could correspond to several tracks, while considering several hits, those tracks might intersect at one single point, thus defining the tracks for this particle. This procedure is illustrated in Figure 3.1.

The second algorithm used at D0 is the AA algorithm. It starts from a track comprised of three innermost hits in the SMT. Then as the track is extrapolated outwards, additional hits are added if criterias are fulfilled. So if $\chi^2 < 16$ while a hit is added, then this track is recorded. If not any one hit could be found, it is a "miss track". Also, since multiple hit could be found fitting the criteria, a removal procedure is applied on the duplicate ones.

Each of HTF and AA algorithms create one set of reconstructed tracks. These are then combined later, hence giving a final single list of tracks.

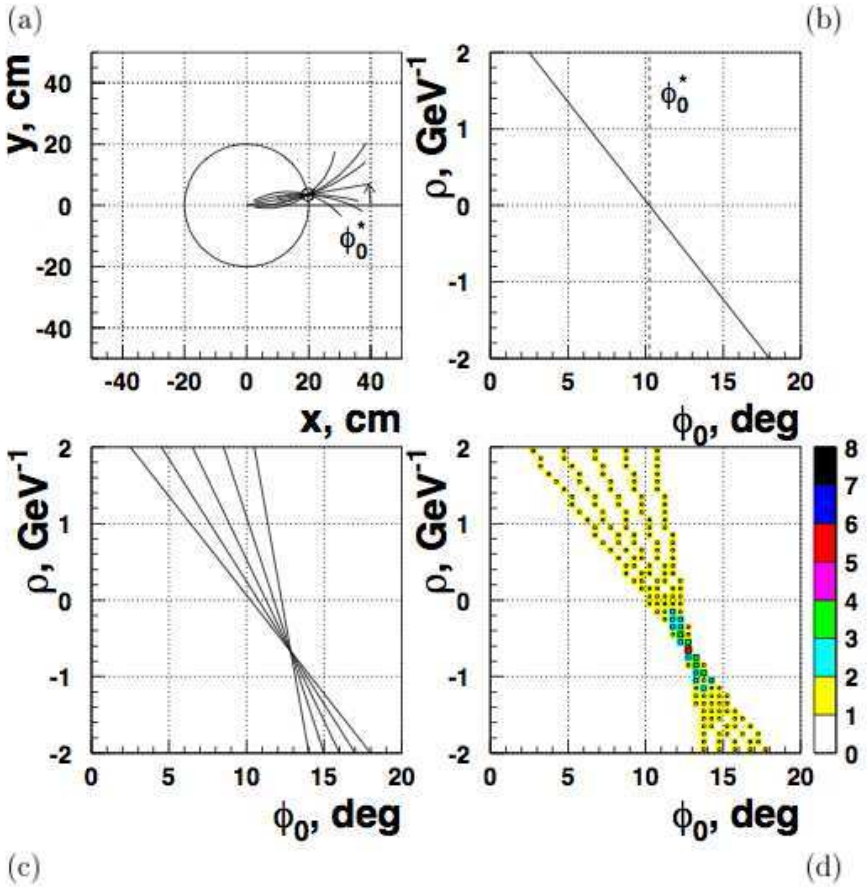


Figure 3.1: The HTF method applied to a single 1.5 GeV track with 5 hits: (a) a set of track solutions is defined in the (x, y) plane for a given hit. (b) Trajectories of tracks in the (ϕ, ρ) plane for a given hit. (c) Trajectories for several hits intersect at the same point in the (ϕ, ρ) plane. (d) The intersection can be seen as a peak in a 2D distribution.

3.2 Primary Vertex

For each single bunch crossing, multiple $p\bar{p}$ could happen, among which inelastic collisions are interesting in our analysis. The elastic interactions are refereed to as minimum bias (MB) events. The vertex selection is processed in two steps: determining all vertices from a $p\bar{p}$ collision, and then choose the vertex with least MB hypothesis. Tracks with $p_T > 0.5$ GeV and at least 2 hits in the SMT are clustered together, and the distance between a track and the cluster should be $z < 2$ cm. If the χ^2(assigned to each vertex through Kalman fits) per degree of freedom is larger than 10, the track that contributes the most to the χ^2 is removed, until $\chi^2 < 10$, or only one track is left. Each cluster is a PV candidate after such procedure. And since tracks from MB usually have low p_T, the probability can be constructed for a track to be from a MB. The probabilities from MB for both the tracks of a specific vertex, and this vertex are considered together, then the vertex with the lowest probability of being from MB is selected as the PV.

3.3 Leptons

Two categories of lepton identification criteria are applied for each lepton channel: the "loose" sample is used to estimate the multijet background from data, and the "tight" sample is used to perform the search. We will introduce these criterias together with the lepton reconstruction and identification criterias here below.

3.3.1 Electron

There are several steps for identifying the electrons: we choose electron "seeds" through clustering the towers in electromagnetic calorimeter. We apply different ceitiera to enhance the electron candidate purity.

Since electrons interact mostly in the electromagnetic calorimeter, the towers with the highest p_T are chosen to be "seeds", and all towers within $\Delta R = \sqrt{\Delta\Phi^2 + \Delta\eta^2} < 0.2$ are summed to create a cluster.

To avoid misidentify photons or jets as electrons, a Boosted Decision Tree from the TMVA package is used [47], and a detailed description of the electron identification scheme can be found in D0 note [48]. The following are part of the input variables for the BDT:

- Electromagnetic fraction (EMf). EMf is the ratio of the energy deposited in the electromagnetic calorimeter to the energy in both the electromagnetic and hadronic calorimeter, i.e. $EMf = \frac{E_{EM}(\Delta R<0.2)}{E_{totql}(\Delta R<0.4)}$. For electrons EMf should be close to 1, since electrons deposit most of their energy in electromagnetic calorimeter.

- Calorimeter isolation "Iso". Isolation is considered to avoid overlap between EM and hadronic, and is calculated as $Iso = \frac{E_{tot}(0.4)-E_{EM}(0.2)}{E_{EM}(0.2)}$. This helps removing noise or jets.

- Track isolation of the EM cluster $IsoHC4$, calculated as the total track p_T (for tracks with $p_T > 0.5$ GeV) in the hollow cone $\Delta R < 0.4$ around the EM cluster,

- "track match". χ^2 is defined to measure the fitting quality of ϕ and z coordinates in tracker and calorimeter: $\chi^2 = (\frac{\Delta\phi}{\sigma_\phi})^2 + (\frac{\Delta z}{\sigma_z})^2 + (\frac{E_T/p_T-1}{\sigma_{E_T/p_T}})^2$, where $\Delta\phi$ and Δz are the differences as measured by the tracker versus calorimeter, σ_ϕ and σ_z are the RMS of

these differences, E_T is the calorimeter energy and p_T is the transverse momentum measured by the tracker of objects in the CC.

- $HMatrix$. While comparing the shape of the EM cluster to that of an electron in Monte Carlo, one could create a covariance matrix using multivariate algorithm. H-matrix of dimension 7 comprised of seven correlated variables: (1-4) the fraction of energy deposited to the total in each of the four EM layers, (5)the shower width in ΔR, (6)the log of the total shower energy, (7)z_{PV}. And the H-matrix of dimension 8 includes the above seven variables and one more, which is the transverse shower width.

- Electron Likelihood. The 8 likelihood variables are: (1)spatial track match χ^2 probability, (2)E_T/p_T , (3)H-matrix of dimension 7, (4)H-matrix of dimension 8, (5)EMf, (6)distance of closest approach to primary vertex, (7)number of tracks in a $\Delta R = 0.5$ cone, (8)total p_T of tracks in 0.4 cone around candidate track. [49].

We form "loose" and "tight" electron samples in our WH analysis. We define such "loose" and "tight" criteria based on the different requirements on the response of this BDT discriminant, to retain high electron selection efficiencies while suppressing backgrounds at differing rates.

To account for the different electron ID efficiencies between data and MC, we derive a scale factor as a function of the electron ϕ, η_{det}, and the instantaneous luminosity, and apply it to the simulation. [50]

3.3.2 Muon

The muons that we are looking for are from the $W \to \ell\nu$, so we want to search for high transverse momentum muons, which have no extra calorimeter or tracking activity around them. Since muons are minimally ionizing particles, there is little muon radiation energy loss, and muons generally escape the detector leaving hits in the tracking, calorimeter and muon detectors.

The muon reconstruction utilizes the hits from the muon detector to form track segments, and then these segments are used to form tracks (local tracks) in the muon system. The segments which are not used to form local tracks, and all the local tracks are called the "local muons". [51]. If a track in the tracking system is found to match a local muon, then it is called a "central track-matched muon".

Based on the reconstructed information:

- the presence of a local muon;

- a local muon with a matched track in the tracking system (called a "central track-matched muon");

- the absence of significant activity around the muon trajectory, both in the calorimeter and in the central tracker;

the muon identification (ID) group defines different categories of muon identification quality.

In the WH analysis, we require the muon candidates fulfil the following criteria:

- "medium" quality as defined by the muon ID [51]

- scintillator hits timing requirement for A layer $|t_A| < 10$ ns and for BC layers $|t_{BC}| < 10$ ns (this is to reduce cosmic ray background)
- "trackmedium" for the central track match quality

As mentioned, we form "loose" and "tight" muon samples in our WH analysis. We define such "loose" and "tight" criteria based on the different muon isolation case, since muons coming from the leptonic decay of W bosons tend to be isolated from jets, while muons originating from semi-leptonic decays of heavy flavoured hadrons are non-isolated. The "loose" muon sample is required to satisfy the muon identification, track quality mentioned above, and the "loose" isolation criterion: $\Delta R(\mu, jet) > 0.5$, this limits the distance between a muon and the nearest good jet. And the "tight" muon sample is required to satisfy the muon identification, track quality mentioned above, and the "tight" isolation criterion: "TrkTightScaled", which is based on $Halo(0.1, 0.4)$ and $TrkCone(0.5)$. $Halo(0.1, 0.4)$ is scalar sum of transverse energies within the cone $0.1 < \Delta R < 0.4$, and $TrkCone(0.5)$ is the scalar sum of transverse momenta of all the tracks within $\Delta R < 0.5$ around the muon (except muon p_T). Then a "TrkTightScaled" is

- Halo$(0.1, 0.4)/p_T(\mu) < 0.4$
- TrkCone$(0.5)/p_T(\mu) < 0.12$

From these identification, track matching and isolation criteria, we also get scale factors and apply these factors to the Data/MC correction. [51]

Muon isolation variables are also defined to separate non-isolated muons from semi-leptonic decays in jet and muons originating from W or Z bosons. The isolation depends on the tracks near the muon, and the amount of energy in the calorimeter near the muon, and can also be used to reject pions that mimic a muon signature.

Besides all the above mentioned criteria, a cosmic muon veto is applied, using information from the scintillator hit times and coincidence with the brunch crossing.

3.4 Jets

$p\bar{p}$ collisions at Tevatron produce a large amount of high energy quarks and gluons. The hadronization of quarks and gluons produces colour-neutral mesons and baryons. These highly-collimated collection of particles are called jets, which are then detected by the detector. Since these jets are associated with the original quark, they are important for the WH analysis.

In this section I will introduce the basic reconstruction and identification of jets at D0, then focus on my service work of jet identification certification on re-processed data at D0.

3.4.1 Jet Reconstruction

We use a cone algorithm at D0 for the Run II period data. Here cells are considered as massless objects with a 4-vector $P^{cell} = (E^{cell}, \vec{p}^{cell})$ [52], where E^{cell} is the energy deposited in the cell and $\vec{p}^{cell}$ is defined from the interaction point to the centre of the cell. After cells due to electronic noise are removed through the algorithm described in [52], the reconstructed towers are obtained from selected cells, and so the 4-vector could be calculated as:

$$\boldsymbol{P}^{tower} = (E^{tower}, \vec{p}^{tower}) = \sum_{i=cells\ in\ tower} (E_i, \vec{p}_i) \tag{3.1}$$

Now all the towers selected through the above steps are listed in decreasing p_T. A precluster is starting from the first tower I with $p_T^I > 500$ MeV, the next tower J will be combined to the formed precluster, if $\Delta R(precluster, J) < 0.3$ and $p_T^J > 1$ MeV. Such iteration will continued until no tower is left. Then all the preclusters getting from this step are used as the input to the D0 Run II cone algorithm.

The Run II cone algorithm uses the preclusters to form clusters, or called proto-jets, with $\Delta R_{cone} = 0.5$. It loops over p_T ordered preclusters: if a precluster P fulfils the condition $\Delta R(P, proto-jet) < \frac{R_{cone}}{2}$, it is added to the proto-jet, else it will be used to form a new proto-jet. This is repeated until a stable proto-jet is found, which means the position of the axis of the jet is not changed from one iteration to the next.

To deal with the possible overlap, we use the merging/splitting algorithm. If one jet shares one or more towers with another jet, these two jets are merged if the sum of p_T of the items shared exceeds 50% of the highest p_T jet, or these two are separated.

3.4.2 Jet Identification and Vertex Confirmation

To increase the purity of reconstructed jets, or say removing fake jets reconstructed not from partons of the hard scatter process, we process the jet identification and vertex confirmation.

From Run II, we apply the following cut on the jets, which are reconstructed as the previous section:

- $0.05 <$ Electromagnetic (EM) fraction $EMf < 0.95$, this is to reject electron or photon-like objects (< 0.95), and reject hardware noise (> 0.05)
- Coarse Hadronic (CH) fraction < 0.40, this is to reject fake jets with high energetic towers.
- Ratio of energy in hottest cell to that in the second hottest cell < 10,
- Minimal number of cells containing 90% of the cell's energy > 1.
- Level 1 trigger confirmation $\frac{E^{jet}_{\Delta R<0.5}}{E^{jet}_{reconstructed}} > 0.5$, where $E^{jet}_{\Delta R<0.5}$ denotes the energy of the trigger towers inside a $\Delta R < 0.5$ cone around the jet axis, $E^{jet}_{reconstructed}$ denotes the jet energy as reconstructed by the precision readout.

Jets with $|\eta| < 2.0$ were also required to have a (hard) track match with

- $p_T > 2.0$ GeV,
- $dR < 0.5$,
- $\chi^2 < 3$.

Similarly jets required to pass a track veto had no track match with

- $p_T > 0.5$ GeV,
- $dR < 0.5$,

- $\chi^2 < 3$.

We call a jet as "good jet" if it passes the above mentioned criteria.

Using the tag-and-probe method we study dijet events. We select a good track jet (tracks clustered into a cone jet) as tag jet, and then probe a good calorimeter jet (calorimetry energy clustered into a cone jet). We calculate the jet identification efficiency (passing "goodjet" criteria) as $JetIDEfficiency = \frac{N_{probe}}{N_{tag}}$. The JetID efficiencies are given by the D0 JetID group in [53], and such JetID efficiencies in Data and MC as well as the Data/MC Scale Factors (SF) are used by all the D0 physics groups. The followings are several example figures for the efficiencies and scale factors.

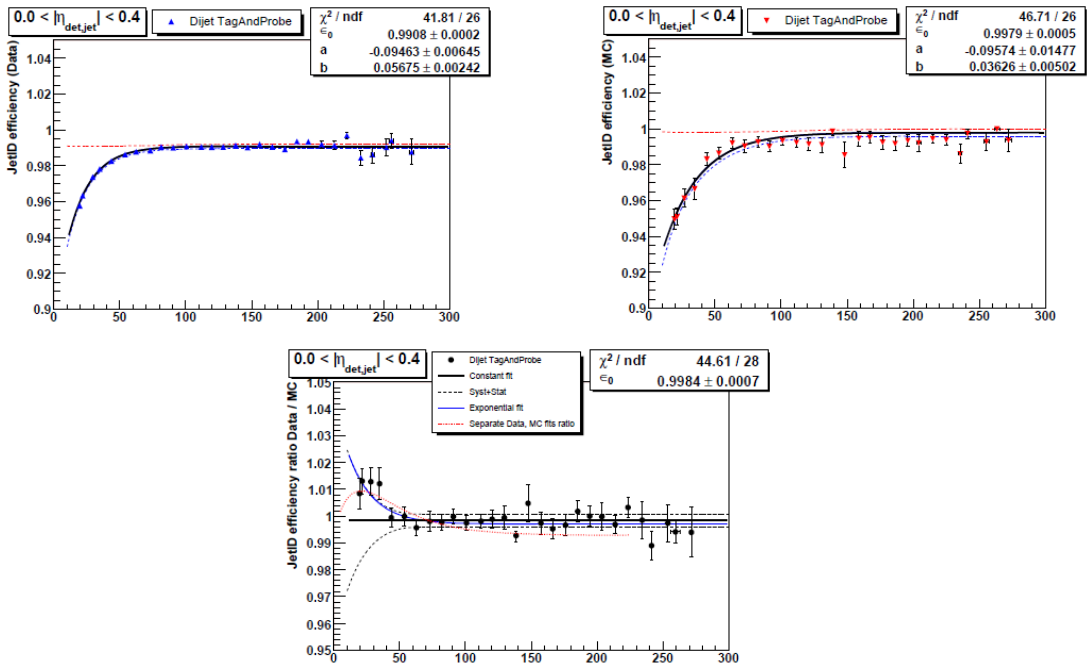


Figure 3.2: (a)JetID Efficiency in Data, (b)JetID Efficiency in MC, (c)JetID Data/MC Efficiency Scale factor. [53]

To further improve data/MC agreement in the Run IIb part of the analysis, we additionally require the confirmation that the selected jet is originating from the primary vertex. For this, we require the jet to have at least two tracks attached to the primary vertex. This gets rid of additional jets not belonging to the hard scatter. The criteria for a jet to be "vertex confirmed" are as the following:

- $p_T > 0.5$ GeV,
- $SMThits \geq 0$,
- $DCA_{xy} < 0.5$,
- $DCA_z < 1.0$,
- $\Delta z\ (vertex, track) < 2.0$ cm.

Jet Identification for Reprocessed IIb Data

Besides the originally-processed data and corresponding MC [54], the D0 collaboration reprocessed a selected sample of Run IIb raw data events. Such data reprocessing takes advantage in its reconstruction code and improved understanding of the detector. The Run IIb reprocessing contains updated track information, which could help reject multijet background. Also the updated track information influences taggability, and vertex reconstruction efficiency. Since the JetID is based on the track information, taggability and vertex, we need to remeasure the JetID for reprocessed IIb data. And such JetID efficiencies and MC/data corrections are provided to the whole D0 to use. The D0 $ZH \to \ell\ell b\bar{b}$ group redid the analysis for the reprocessed data, and got a slightly better sensitivity (few percent) than the original one, according to the ZH group report on the DZero Summer Workshop at NIU, in June 2013. The WH analysis was also on a similar track but there was not enough manpower to finish it up.

For my study of JetID, I use the reprocessed IIb1, IIb2, IIb3 and IIb4 data epochs, provided by the D0 Common Sample Group (CSG). The reprocessed datasets provided by the D0 CSG, and the D0 official di-jet reprocessed IIb1 and IIb3 MC samples have been used. PYTHIA [55] is the generator used to produce all the simulated samples.

I studied the JetID for the reprocessed Run IIb data to measure jet ID efficiency, the vertex confirmation efficiency and the combined vertex confirmation and taggability efficiency using the tag-and-probe method on dijet events, compared the reprocessed result to the previous result from default data, and provided the efficiencies and scale factors (SF) to all the D0 physics groups. These efficiencies are measured for data and MC separately, and parametrized as a function of jet p_T in different rapidity and primary vertex z bins. Here the JetID is defined above, the "vertex confirmed" means that jets must have at least two tracks attached to the primary vertex.

A jet is taggable if:

1. $E_T > 15$ GeV, $|\eta| < 2.5$, (i.e. Jet energy scale is defined)

2. If it contains a track-jet within $\Delta R < 0.5$

The track-jets, or track based jets are reconstructed as the following:

1. Pre-clustering: make precluster in z(along beam axis) of tracks that are nearby in z. Start from highest p_T tracks.

2. Track selection: associate each precluster to the closest Primary Vertex (PV), use tracks that have $p_T > 0.5$ GeV, ≥ 1 hit in the most precise section of the silicon.

3. From the preclusters, the tracks are clustered with a simple cone algorithm, with a track seed of $p_T > 1$ GeV.

The taggability is calculated as:

$$Taggability\ (E_T, \eta) = \frac{Number\ of\ taggable\ jets\ (E_T, \eta)}{Number\ of\ jets\ (E_T, \eta)}$$

As several quantities (object ID, reconstruction efficiencies) may depend on the luminosity, it is important to get in the simulation the same luminosity profile as in data. This is done by reweighting the instantaneous luminosity profile of the data overlayed to the

MC such that matches the data instantaneous luminosity profile. Only the shape of the luminosity profile is affected so this process does not affect the overall MC normalization. And it is recommended we use the centrally produced MC instantaneous luminosity profiles to reweight MC to Data luminosity. More details can be found through CSG documentation [54].

We use the tag-and-probe method to select back-to-back ($\Delta\phi > 3$) di-jet events from both data and MC [53, 56]. The tag jet must pass the jet ID selection and be vertex-confirmed, as well as have $p_T > 15$ GeV. For the jet ID selection efficiency, the denominator is the number of probe jets, the numerator is the good jets. The vertex confirmation efficiency is measured on top of the jet ID selection. The denominator is the number of good jets, the numerator is the one having at least 2 associated tracks from the primary vertex ($ntrkMultiplicity0() >= 2$). For the combined vertex confirmation and taggabiliy efficiency, the numerator is the number of events in which the vertex is confirmed and the jets are taggable.

To fully consider the various kinematic dependence, we split the events in 6 different z bins of primary vertex (PVZ) (-60 cm, -40 cm, -20 cm, 0 cm, 20 cm, 40 cm, 60 cm), and 22(14) different rapidity bins for ID selection (vertex confirmation or vertex confirmation + taggability), then measure the efficiencies as a function of jet p_T. The 22 rapidity bins used are (-3.2,-2.5,-2,-1.6,-1.4,-1.2,-1.,-0.8,-0.6,-0.4,-0.2,0,0.2,0.4,0.6,0.8,1.0,1.2,1.4,1.6,2,2.5,3.2). For vertex confirmation or vertex confirmation + taggability, we need to use a coarser binning, the 14 bins are (-3.2, -2.5, -2, -1.6, -1.2, -0.8, -0.4, 0, 0.4, 0.8, 1.2, 1.6, 2, 2.5, 3.2). The 1D efficiency as a function of rapidity and PVZ are shown in Figures 3.3 - 3.11.

From the Jet ID selection efficiency figures, we can see good agreement between the Run IIb1 data and the Run IIb1 MC, and also between the Run IIb1,2,3,4 data and the Run IIb3 MC, and the SF in these figures are equal to one. There is slight disagreement between Run IIb234 data and Run IIb3 MC, and the SF here is flat. So the reprocessed MC makes reasonable simulation to the reprocessed data, this suggests that the use of the Run IIb3 MC to simulate the whole dataset Run IIb1,2,3,4 data is better than the subsample Run IIb234 data.

From the vertex confirmation figures, and the vertex confirmation plus taggability figures, we can see good agreement between the data and MC at central η_{det} region, and central pvz region, corresponding to SF being near one. And at the high η_{det} region, and high pvz region, we see disagreement between the data and MC, and the SF we get from these regions are applied to correct the MC.

To constrain the statistical uncertainty, we use the exponential function

$$f(p_T) = a + b \cdot e^{-c \cdot p_T} \tag{3.2}$$

to fit the data and MC separately, and used the fitted results as the nominal efficiency, and fitted errors (see Equation 3.3, which includes parameters correlations) as the 1σ uncertainty. We calculate the scale factor based on the fitting parameters, and treat the uncertainties of the data and MC as uncorrelated and combine them in quadrature.

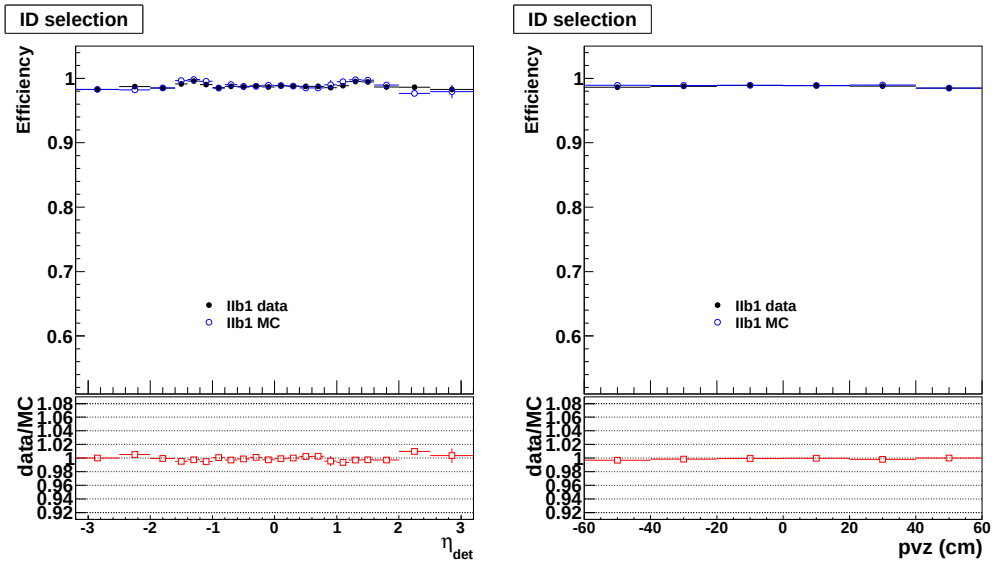


Figure 3.3: Jet ID selection efficiency as a function of η_{det} and primary vertex z from di-jet IIb1 data and IIb1 MC.

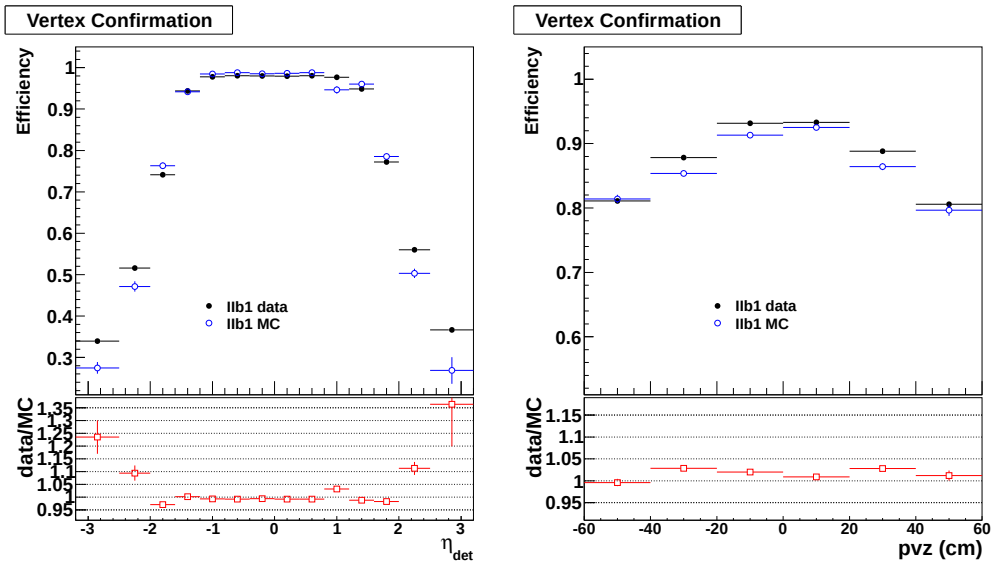


Figure 3.4: Vertex confirmation efficiency as a function of η_{det} and primary vertex z from di-jet IIb1 data and IIb1 MC.

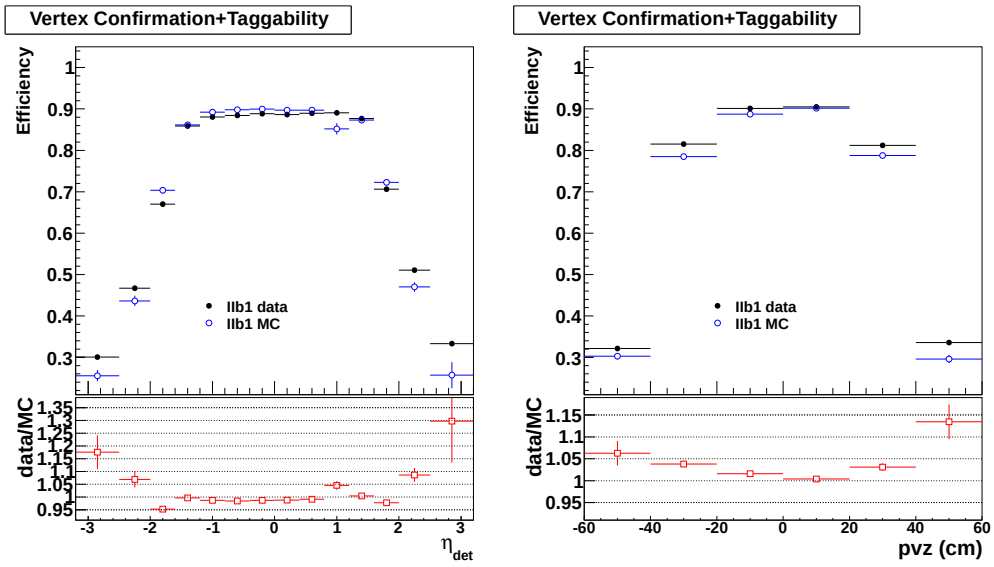


Figure 3.5: Vertex confirmation + taggability efficiency as a function of η_{det} and primary vertex z from di-jet IIb1 data and IIb1 MC.

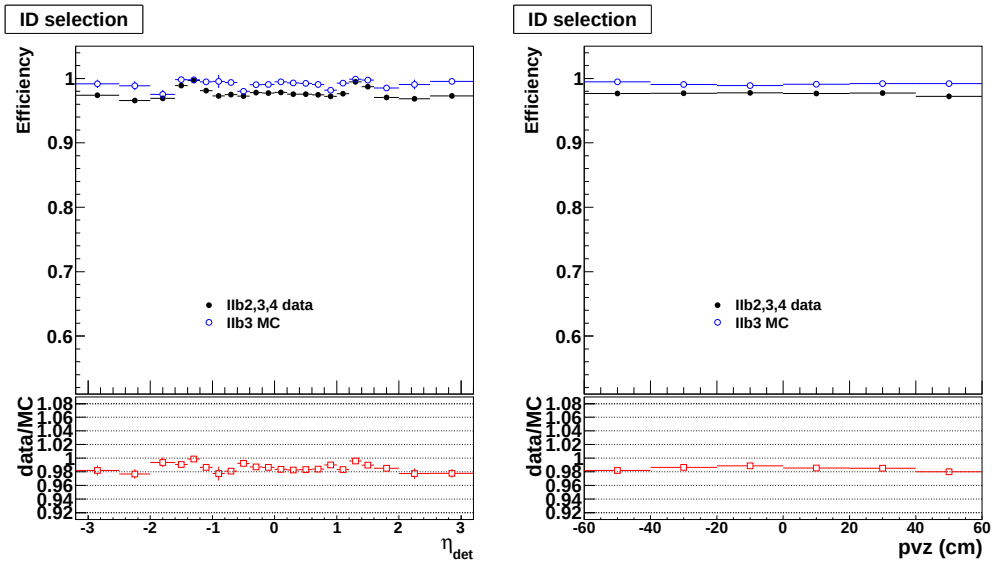


Figure 3.6: Jet ID selection efficiency as a function of η_{det} and primary vertex z from di-jet IIb2,3,4 data and IIb3 MC.

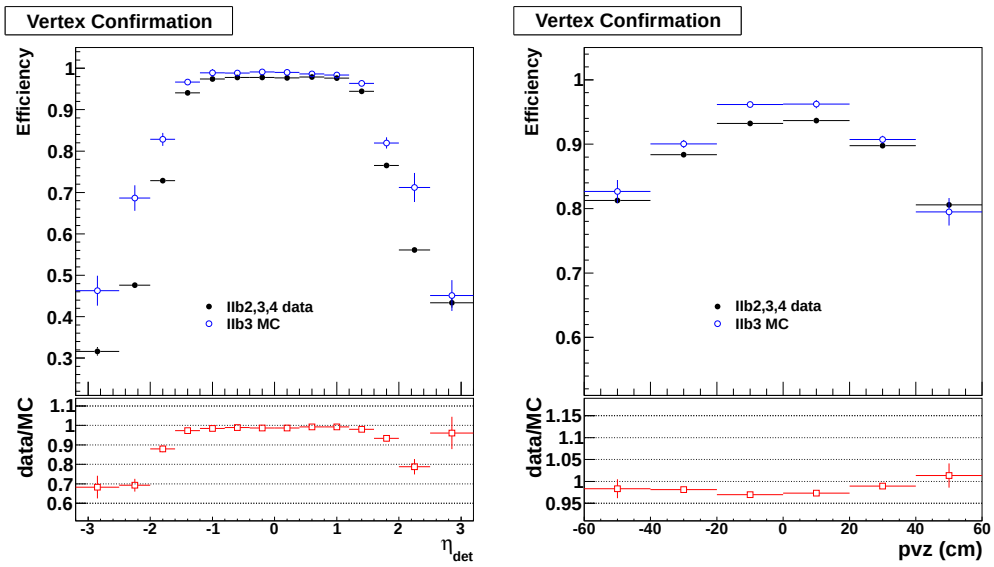


Figure 3.7: Vertex confirmation efficiency as a function of η_{det} and primary vertex z from di-jet IIb2,3,4 data and IIb3 MC.

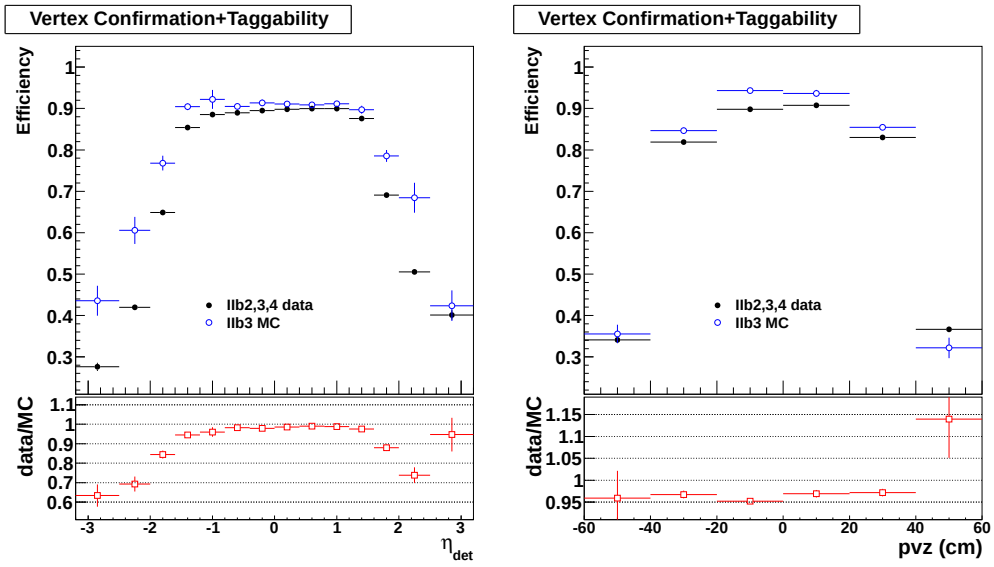


Figure 3.8: Vertex confirmation + taggability efficiency as a function of η_{det} and primary vertex z from di-jet IIb2,3,4 data and IIb3 MC.

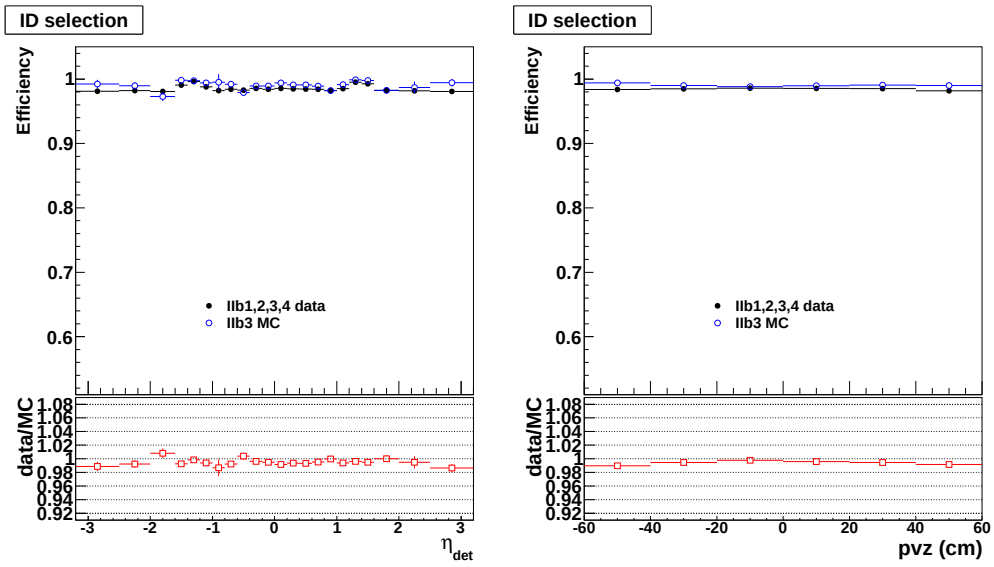


Figure 3.9: Jet ID selection efficiency as a function of η_{det} and primary vertex z from di-jet IIb1,2,3,4 data and IIb3 MC.

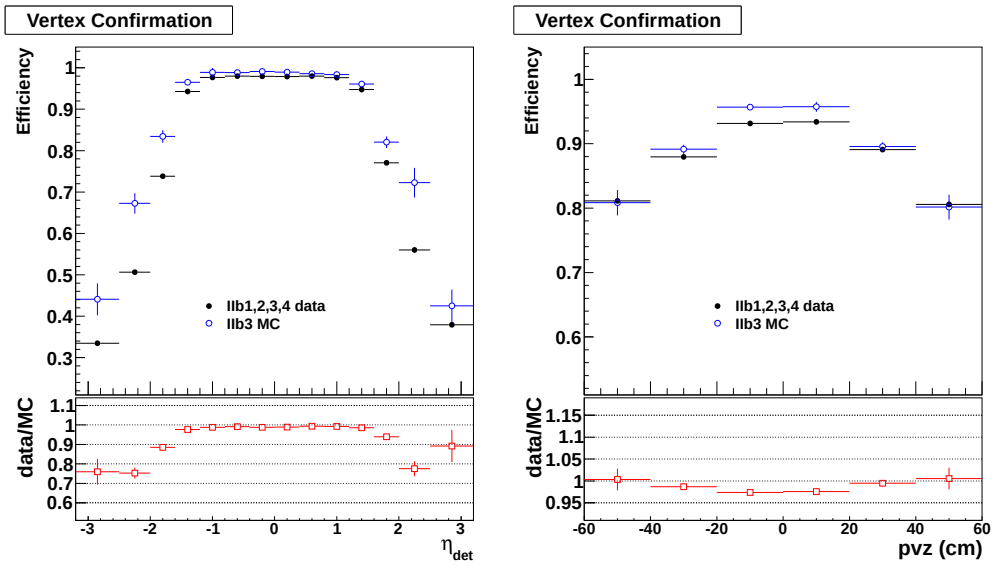


Figure 3.10: Vertex confirmation efficiency as a function of η_{det} and primary vertex z from di-jet IIb1,2,3,4 data and IIb3 MC.

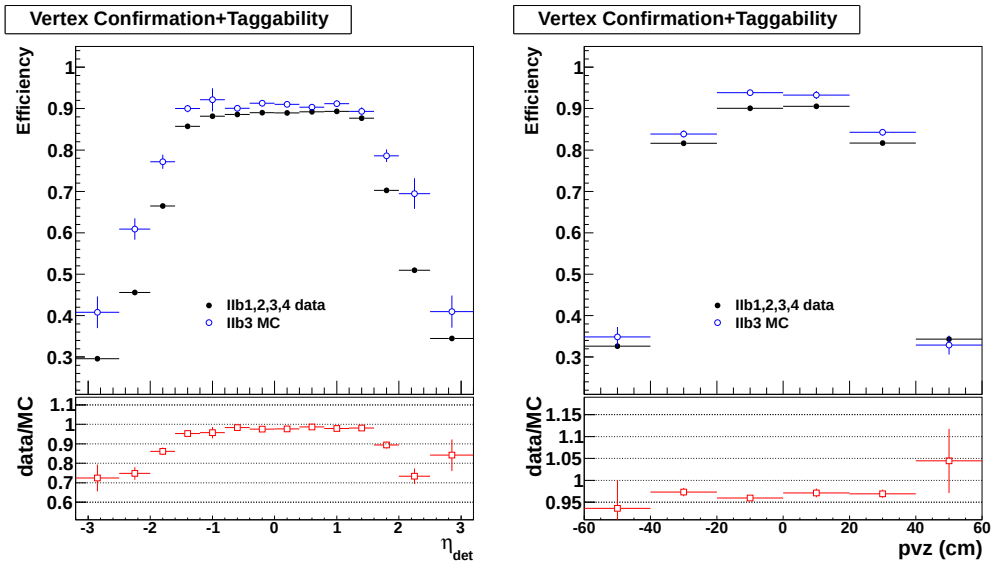


Figure 3.11: Vertex confirmation + taggability efficiency as a function of η_{det} and primary vertex z from di-jet IIb1,2,3,4 data and IIb3 MC.

$$\begin{aligned}\delta^2 f(p_T) =& (\frac{\partial f}{\partial a})^2 \cdot \delta a^2 + (\frac{\partial f}{\partial b})^2 \cdot \delta b^2 + (\frac{\partial f}{\partial c})^2 \cdot \delta c^2 \\ &+ 2 \cdot \frac{\partial^2 f}{\partial a \cdot \partial b} \cdot cov_{ab} \\ &+ 2 \cdot \frac{\partial^2 f}{\partial b \cdot \partial c} \cdot cov_{bc} \\ &+ 2 \cdot \frac{\partial^2 f}{\partial a \cdot \partial c} \cdot cov_{ac}\end{aligned} \tag{3.3}$$

where $\delta f(p_T)$ is the uncertainty for $f(p_T)$, δa, δb and δc are the uncertainties of the three fitting parameters, cov denotes the covariance between fitting parameters.

The jet ID selection efficiencies for IIb1,2,3,4 data and IIb3 MC in the central region are shown in Figures 3.12-3.15. The vertex confirmation efficiencies for IIb1,2,3,4 data and IIb3 MC in the central region are shown in Figures 3.16-3.17. The vertex confirmation + taggability efficiencies for IIb1,2,3,4 data and IIb3 MC in the central region are shown in Figures 3.18-3.19.

From the figures of the jet ID selection, vertex confirmation, and vertex confirmation + taggability efficiencies for the reprocessed data and MC shown here, we can see that the efficiency function behaves similarly to the previous result on default data and MC [53]. These p_T dependencies and the corresponding SF are provided to all the D0 groups.

When comparing to previous JetID result [53], we find that there is almost no change between the default data and the reprocessed data, for the running periods. And at high pvz, the primary vertex is different between the default and the reprocessed data.

Since the JetID is done after jet energy scale (JES, which will be desribed later in Section 3.4.3), and we find that the reprocessed data has fewer PV candidates, considering JES has offset correction that depends on number of PV candidates, we suggest that the usage of the reprocessed data require a new JES (at least for the offset term) and then a

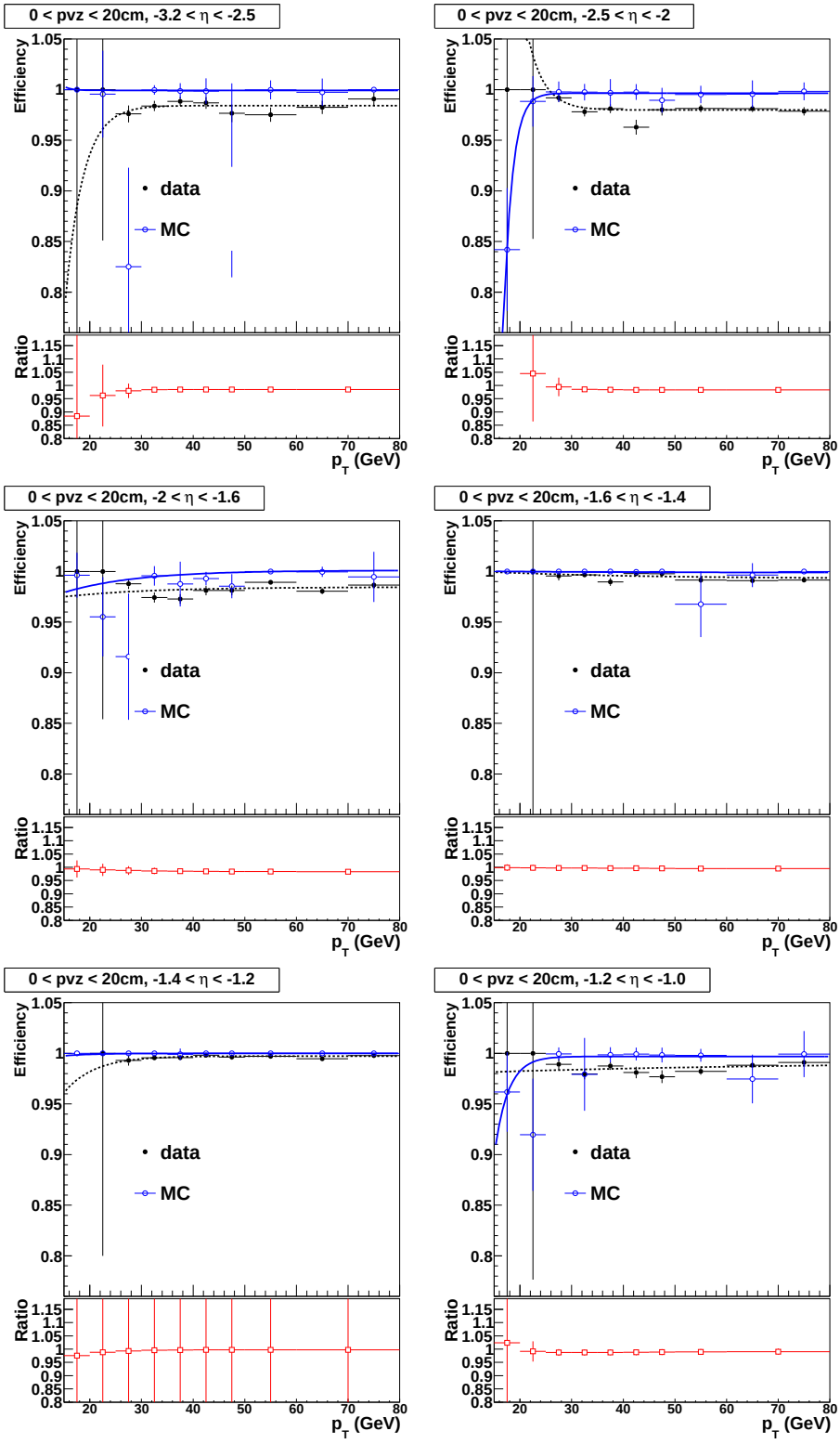


Figure 3.12: Jet ID selection efficiency as a function of jet p_T from di-jet IIb1,2,3,4 data and IIb3 MC.

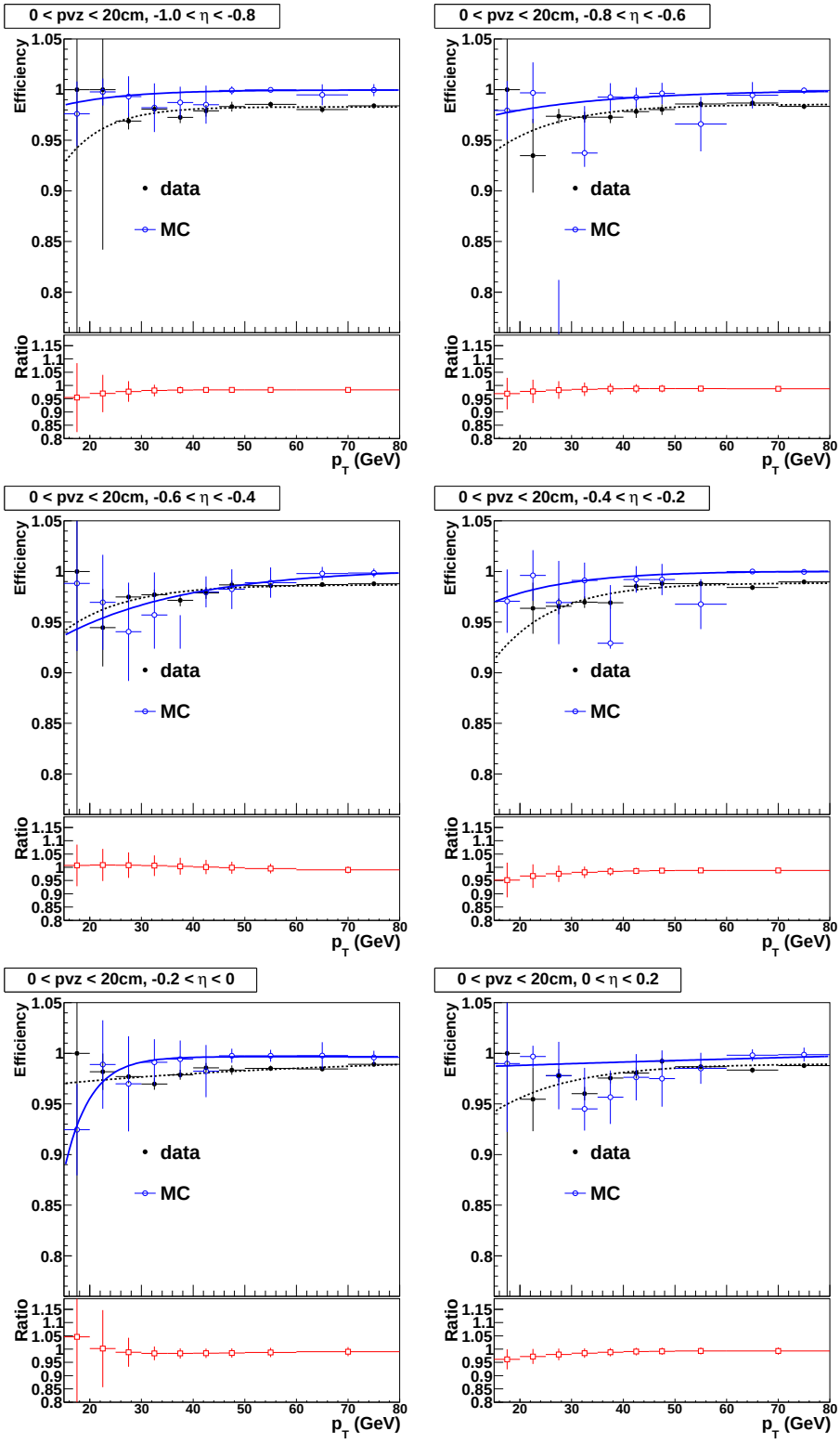


Figure 3.13: Jet ID selection efficiency as a function of jet p_T from di-jet IIb1,2,3,4 data and IIb3 MC.

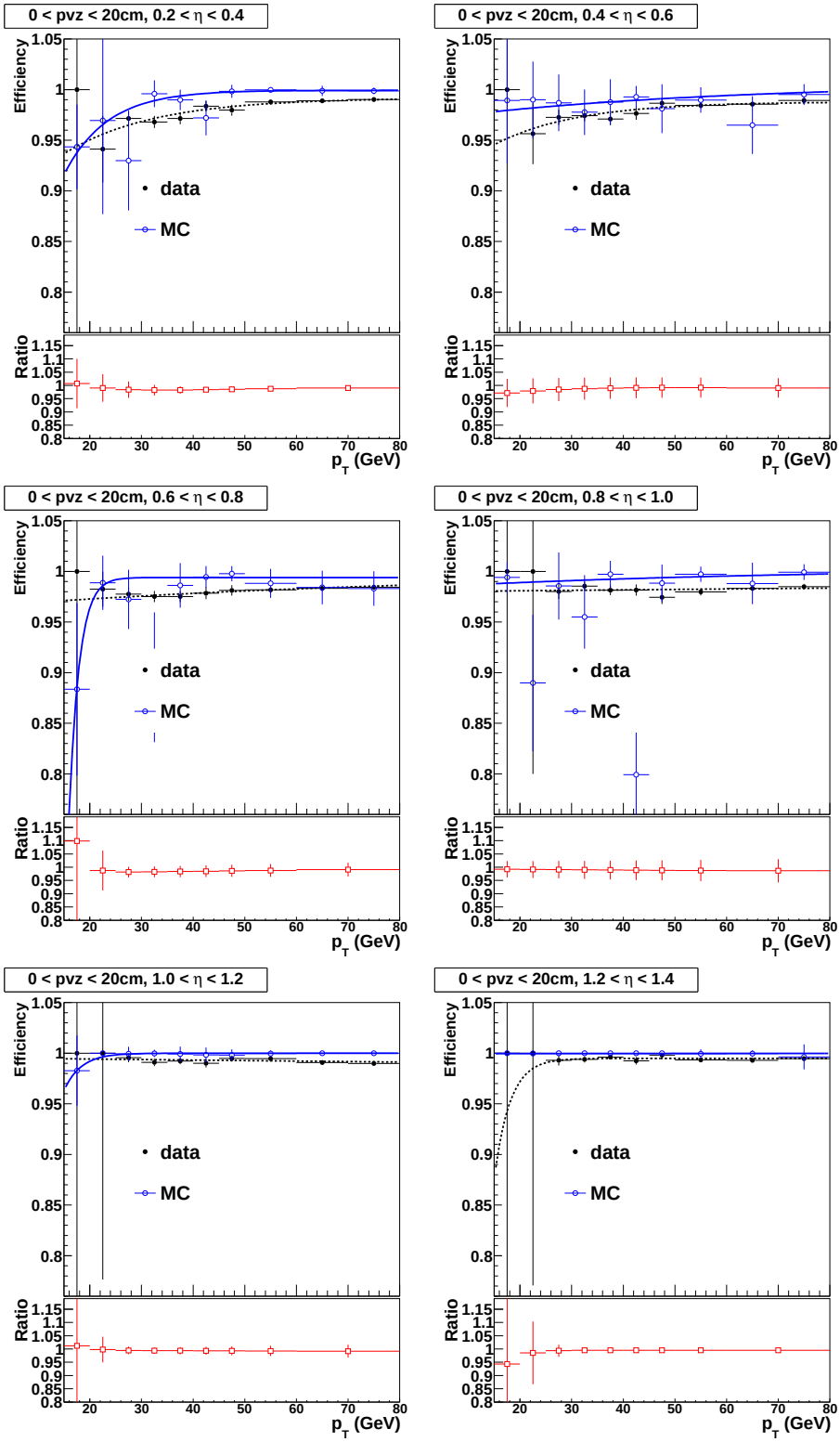


Figure 3.14: Jet ID selection efficiency as a function of jet p_T from di-jet IIb1,2,3,4 data and IIb3 MC.

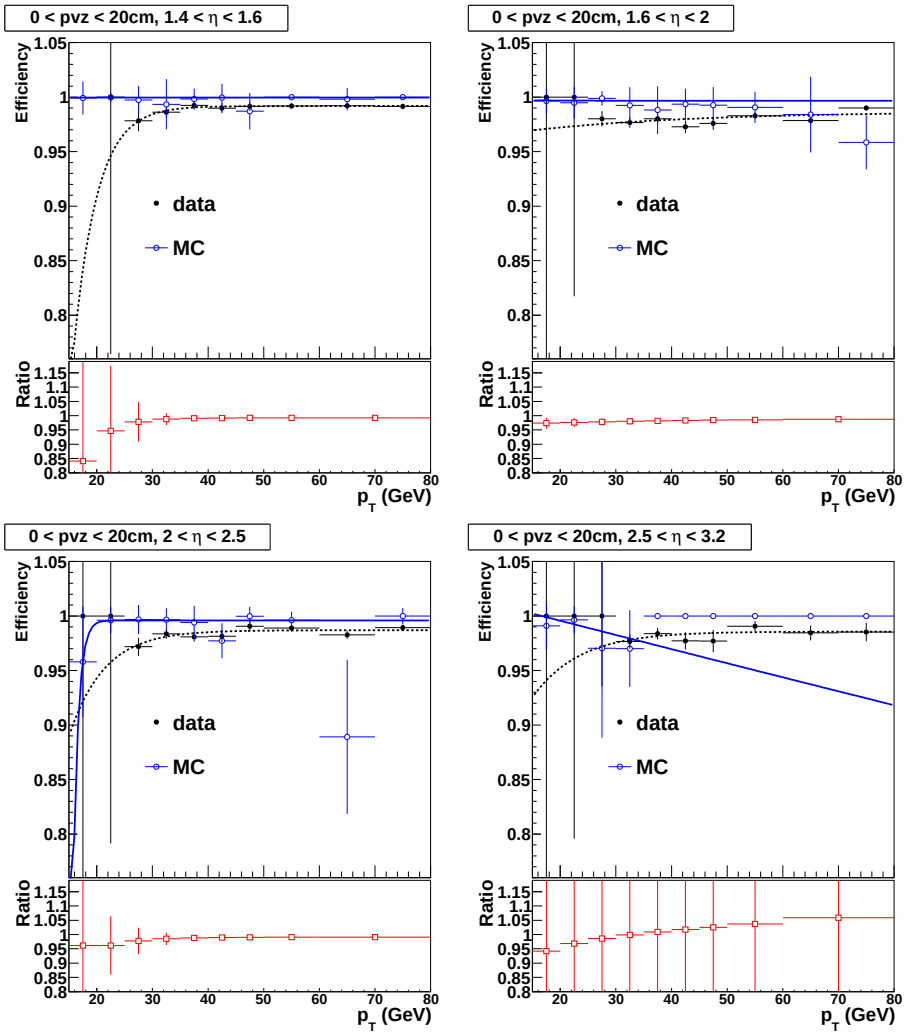


Figure 3.15: Jet ID selection efficiency as a function of jet p_T from di-jet IIb1,2,3,4 data and IIb3 MC.

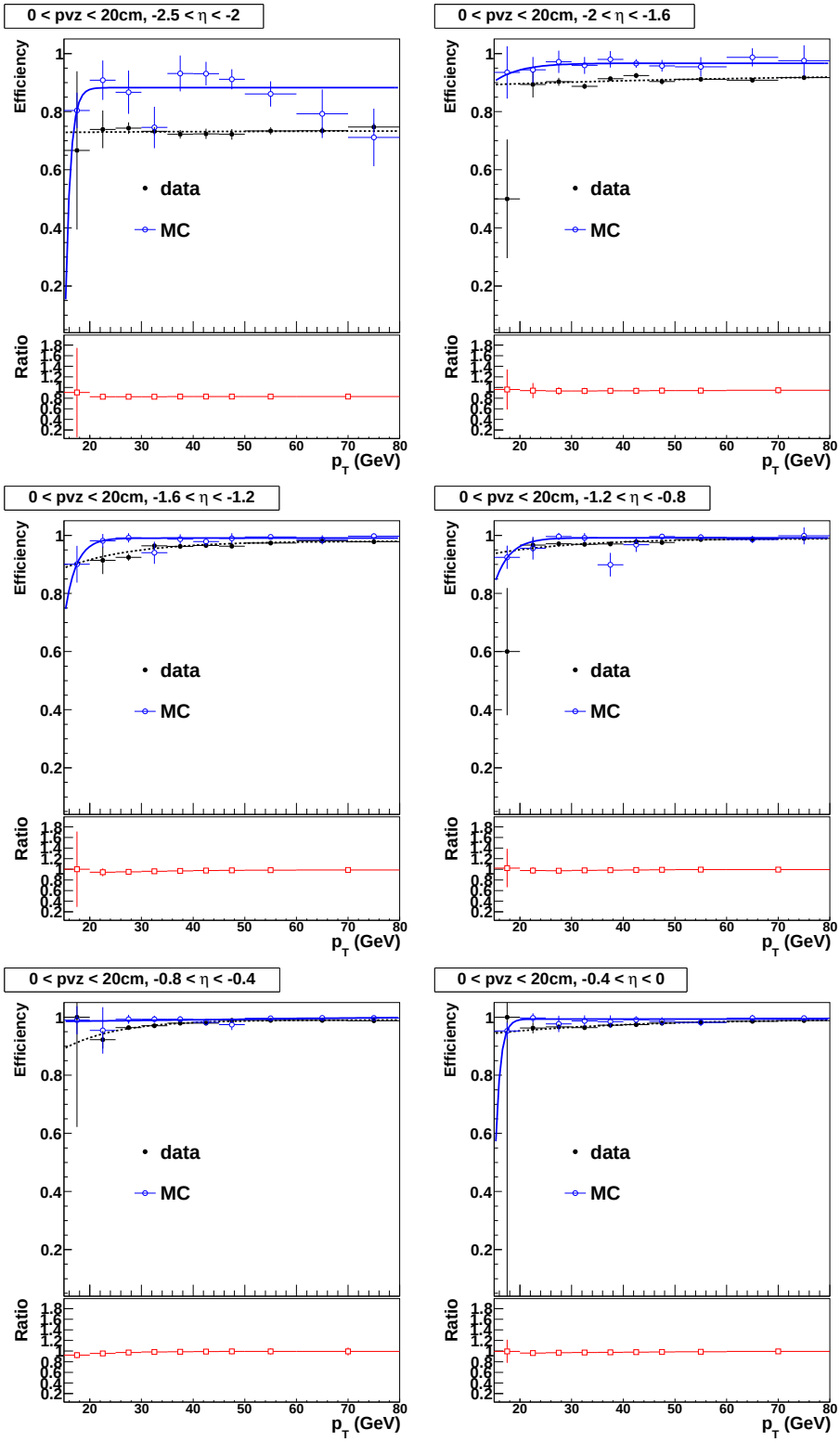


Figure 3.16: Vertex Confirmation efficiency as a function of jet p_T from di-jet IIb1,2,3,4 data and IIb3 MC.

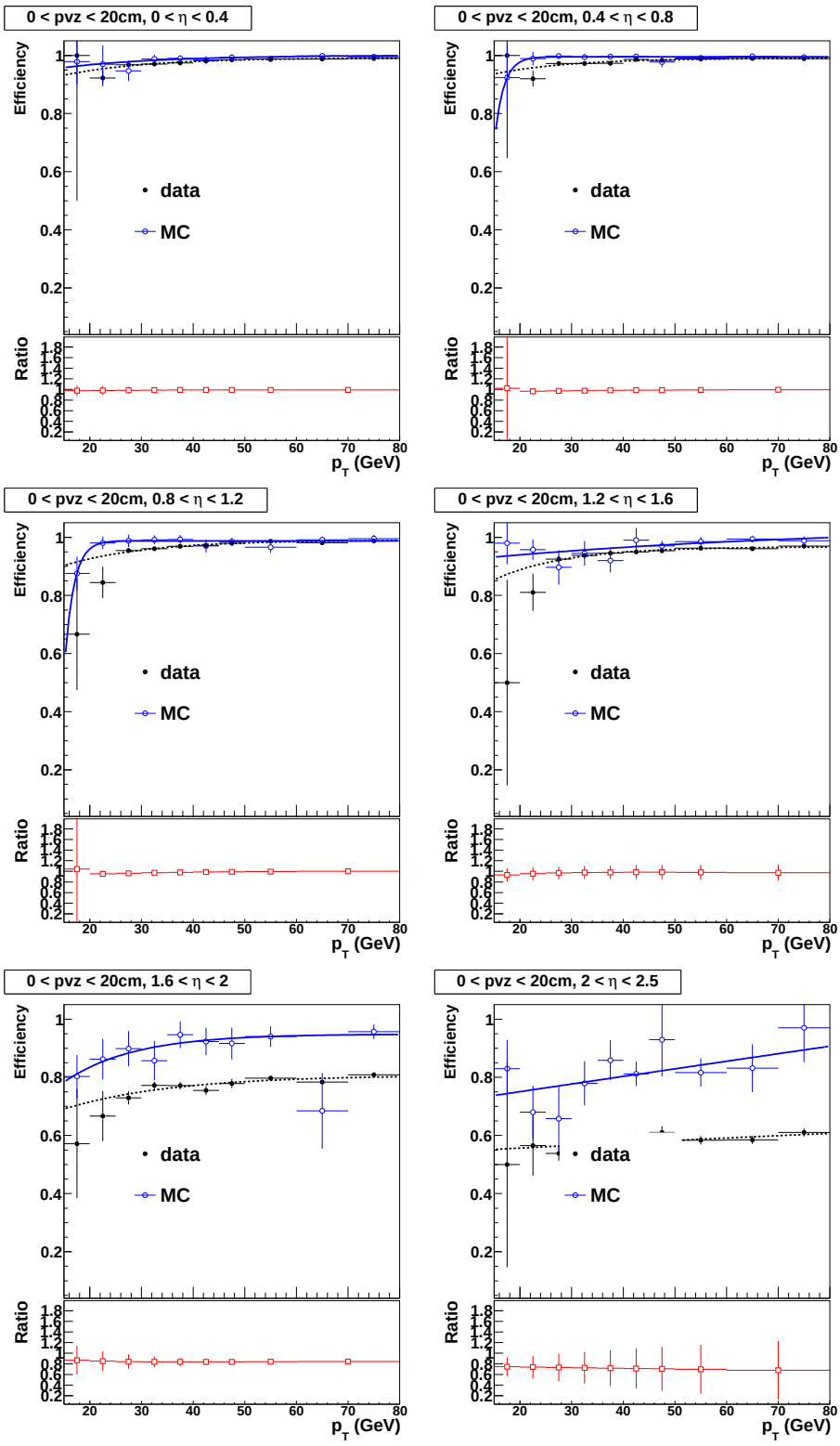


Figure 3.17: Vertex Confirmation efficiency as a function of jet p_T from di-jet IIb1,2,3,4 data and IIb3 MC.

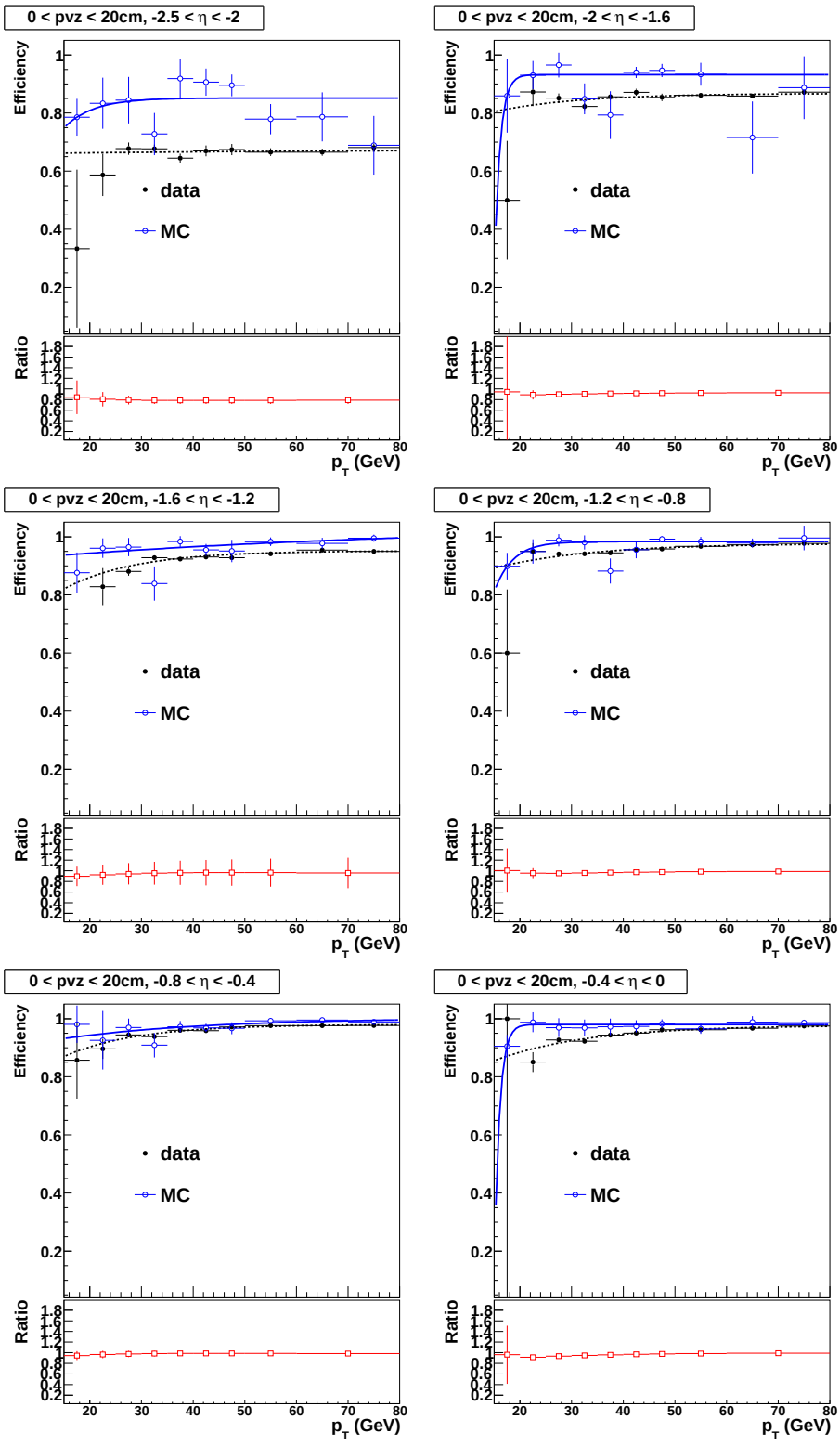


Figure 3.18: Vertex Confirmation + Taggability efficiency as a function of jet p_T from di-jet IIb1,2,3,4 data and IIb3 MC.

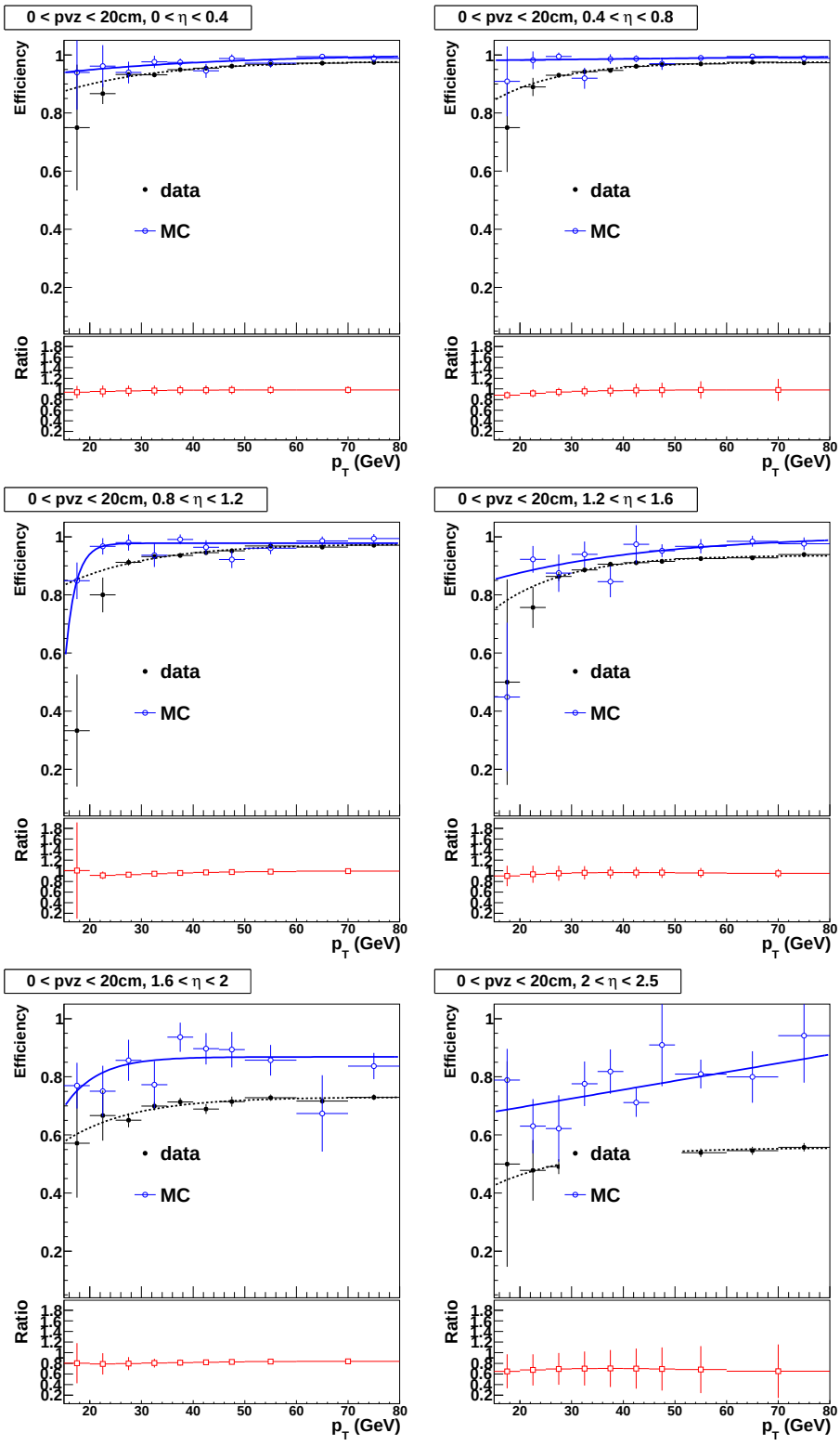


Figure 3.19: Vertex Confirmation + Taggability efficiency as a function of jet p_T from di-jet IIb1,2,3,4 data and IIb3 MC.

revised JetID.

All the jet ID selection efficiencies, the vertex confirmation efficiencies and the vertex confirmation + taggability efficiencies for IIb1 data and IIb1 MC, IIb2,3,4 data and IIb3 MC, IIb1,2,3,4 data and IIb3 MC can be found in our note [57]. And all these efficiencies and scale factors are provided to all the D0 physics groups to use.

3.4.3 Jet Energy Scale

A jet energy scale correction (JES) is used to correct the total energy of the jet measured with the detector. The JES considers various inaccuracies related to jet energy measurement [58]:

$$E_{jet} = \frac{E_{meas} - E_0}{R_{jet} S_{jet}} \tag{3.4}$$

where E_{meas} is the raw jet energy measured with the detector, E_0 is an offset energy (includes energy from minimum bias interactions, previous bunch crossings and noise in the electronics, et.), R_{jet} is the calorimeter energy response, and S_{jets} is for showers that have particles entering or leaving the jet cone.

The offset energy E_0 is measured in two parts: calorimeter response, and effects from additional interactions and energy from previous bunch crossings. The first part is measured by collecting data events at random and by measuring the average energy; the latter part is measured by taking a random sample of data events with an inelastic collision. The R_{jet} and S_{jet} can be measured using events where a photon and a jet are produced back to back. Since the photon energy can be measured to high precision, due to the precisely calibrated calorimeter response to electromagnetic particles for $Z \to ee$ events, we can derive the jet energy correction considering there should be no imbalance in transverse energy.

The MC studies provide showering corrections through comparing the inside and outside cone energy. And such corrections are used to fit shower templates to jet plus photon data to determine the ratio of energy inside the jet cone to the true energy within the cone. Due to semi-leptonic decays of B mesons, a separate correction (JESMU) is derived on jets containing muons to account for that.

3.4.4 Jet Shifting, Smearing, and Removal (JSSR)

To get a consistent agreement between MC and data, the JSSR is the algorithm recalibrating, smearing the jet resolution and possibly discarding simulated jets. Based on the γ+jet and Z+jet events, the p_T imbalance between the jet and other objects is calculated as

$$\Delta S = \frac{p_T^{jet} - p_T^{Z/\gamma}}{p_T^{Z/\gamma}} \tag{3.5}$$

The ΔS is fit in different bins of $p_T^{Z/\gamma}$ with the following function:

$$f(\Delta S) = N \times exp(-\frac{(\Delta S - < \Delta S >)^2}{\sigma^2}) \times (1 + erf(\frac{p_T^{jet} - \alpha}{\sqrt{2}\beta})) \tag{3.6}$$

where the second term is a turn on term used to model the jet reconstruction threshold.

The resolutions are fit in different bins of $p_T^{Z/\gamma}$:

$$\sigma(p_T) = \sqrt{\frac{a^2}{(p_T^Z)^2} + \frac{b^2}{p_T^Z} + c^2} \tag{3.7}$$

where a term describes the instrumental effect, and is important for low energy jets; the b term describes the fluctuations inherent in the development of the showers, the constant term describes calibration errors, dead material in front of the calorimeter, and non-uniformities in the calorimeter.

The imbalance term is fit as:

$$< \Delta S >= A + Bexp(-Cp_T^Z) + D(exp - E(p_T^Z)^2) \tag{3.8}$$

Simulated jets are smeared to make the resolution agrees with the data on average. The energy level of the jets (except b jet) is then shifted to account for the differences in the mean, $< \Delta S >$.

3.5 Missing Transverse Energy

Since we have zero momentum in the transverse plane, and the neutrino interacts weakly and deposits no energy in the detector, we can use the imbalance transverse energy to infer the existence of neutrinos.

This missing transverse energy (MET, or $\not\!\!E_T$) is calculated as the negative vector sum of transverse momenta of the calorimeter cells of the event:

$$\vec{\not\!\!E_T}^{raw} = -\sum^{cells} \vec{E_T} \tag{3.9}$$

The energy of the cells of the coarse hadronic layers of the calorimeter are not included in the sum since they tend to be noisy and reduce the resolution of the measurement. The $\not\!\!E_T$ is also corrected for the presence of any muons. All corrections to electrons, muons and jets are propagated into the $\not\!\!E_T$.

Since JES correction is applied, the energy of cells associated with a jet by the corrected energy should be replaced while calculated the sum of transverse momenta. And since muons deposit very little energy in the calorimeter, this part is subtracted during the final data analysis level.

Chapter 4

Tagging of *b*-quark Jets

In the WH process, the H decays into a pair of b quarks, it is thus crucial to identify ("tag") the final state jets originating from b quarks. Since the b quark has relatively heavy mass ($\sim$ 5 GeV), long lifetime ($\sim 10^{-12}$ s), and produces a displaced secondary decay vertex (several millimetres away from the primary interaction vertex), we can use such characteristics to distinguish b quarks from c and other light (u, s, d) quarks.

4.1 *b*-jets Properties

Unique signatures can be used to tag jets originating from b quarks. Firstly, the b quark is the heaviest quark with a mass around $\sim$ 5 GeV (except for the t quark which is heavier than the W or H bosons). Secondly, a relatively long lifetime of b hadrons (1.518 ± 0.007 ps for B^0 mesons [8]) implies that the hadron can fly a distance (up to few mm), and a secondary displaced decay vertex would be found several mm away from the primary interaction vertex. So we would consider the following items when tagging b quarks from light (u, d, s) or c quark:

- Secondary vertices, which are displaced away from the primary vertices, can be reconstructed from charged tracks.
- Particles with high impact parameters are identified
- The B meson will decay to an electron or muon 20% of the time, 11% directly and 9% after a cascade decay through an intermediate particle.

The latest MVA$_{bl}$ algorithm that we use to identify b jets is provided by the D0 b-ID group [59]. We will describe the input variables of the MVA$_{bl}$ and then the usage of the MVA$_{bl}$in 4.3.

4.2 *b* Jet Identification Algorithm Prerequisites

Jets have to fulfill certain criteria before being used as inputs for the identification algorithms. These include:

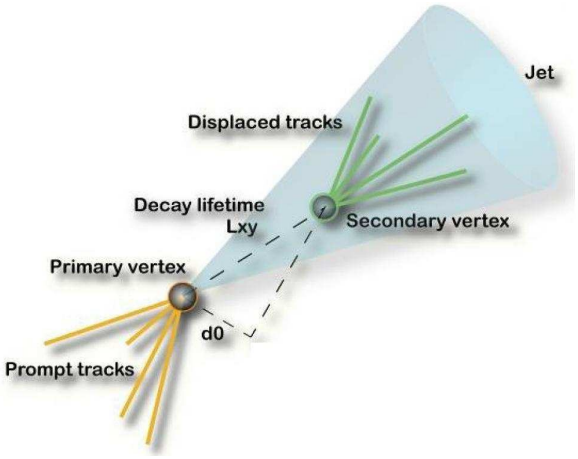


Figure 4.1: As b-hadrons have a long life-time ($\sim$ 1.5 ps), they can fly for a visible distance (20 GeV b-hadron decays after $\sim$ 2 mm) , denoted L_{xy} on this figure, before decaying. Where the decay occurs, a secondary vertex can be reconstructed from tracks produced by the decay products that are not attached to the primary vertex. A jet can be b-tagged if a secondary vertex is reconstructed or if the impact parameter d_0 can be measured (distance of closest approach of tracks to the primary vertex (DCA), defined as a straight line parallel to the z axis of the detector with (x,y) position as the one of the reconstructed primary vertex).

4.2.1 Taggability

Before applying the MVA$_{bl}$ algorithm, the jets are required to pass taggability requirement [59] for reducing the algorithm's dependence on the tracking system efficiency.

The taggability is defined as the ratio of the number of taggable jets over the number of calorimeter jets. The taggability dependency of p_T, η and z is observed and shown in Figure 4.2. The z' variable is used in order to account for correlation between $|z|$ and η and because tracking performance on the amount of material being crossed by charged particles [60]:

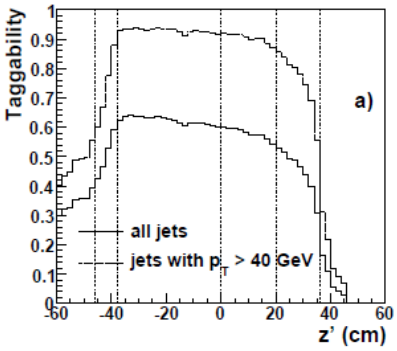
$$z' \equiv |z| \cdot sign(\eta \cdot z) \tag{4.1}$$

4.2.2 V^0 rejection

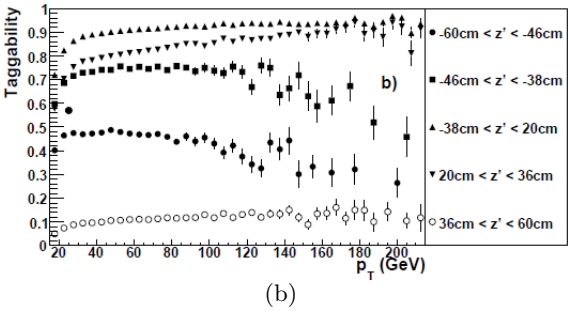
Since strange quarks ($V^0 s$) have long lifetimes, neutral hadrons containing $V^0 s$ have similar decay length signatures to b quarks. So we reject secondary vertices with two oppositely charged tracks which also satisfy the following criteria:

- The tracks associated to V^0 should have a $dca_{bid} < 200\mu$m, [1] this rejects the V^0 coming from the long lived neutral hadrons, but not from b hadron decays.
- To suppress mis-reconstructed tracks, we also apply cut on the z coordinates of the track, the z projection of each track must have a $dca_{bid} < 1$ cm.

[1] there are two dca_{bid} defined by the bID group, i.e. a distance of closest approach with respect to the the primary interaction vertex, 1. along the axis of the z beam, and 2. in the transverse plane with respect to the beam.



(a)



(b)

Figure 4.2: (a): Taggability as a function of $z' \equiv |z| \cdot sign(\eta \cdot z)$. The vertical lines denote the boundaries chosen for the parametrization in p_T and $|\eta|$. (b): taggability as a function of jet p_T , in different bins of z'. The curves for the two central bins are very similar and have been combined.

- $S_d = dca_{bid}/\sigma_{dca_{bid}}$ is the significance of the distance of closest approach of each track relative to the PV in the transverse plane, and we require $|S_d| > 3$.
- A selected V^0 candidate is rejected if its mass is compatible with the mass expected from K_S or Λ, 472 MeV $< m_{K_S} <$ 516 MeV and 1108 MeV $< m_\Lambda <$ 1122 MeV.

To reject photon conversion in the transverse plane, we use a nearly zero opening angle between the electron and positron. The tracks from the electron and positron are required to be less than 30μm apart at the point where their trajectories are parallel to each other. In addition, they must have an opposite electric charge, and their invariant mass must be less than 25 MeV.

4.3 MVA$_{bl}$ Algorithm

MVA$_{bl}$ uses six Random Forests (RF) as input variables to get a final single variable. Some inputs to the RFs were proposed by the D0 b-ID group prior to 2010 [60].

- Counting Signed Impact Parameters (CSIP): Based on the S_d of each track, CSIP counts the number of tracks to a jet. CSIP requires $N_{tracks} \leq 3, S_d > 2$ or $N_{tracks} \leq 2, S_d > 3$.
- Jet Lifetime Impact Parameter (JLIP): The algorithm provides JLIP probability using tracks information in a jet originating from the PV. These probabilities are combined among all the tracks matched to one jet, and are interpreted as the confidence level to determine whether it corresponds to a light or b quark. The probability relates to b quarks if it is close to 1, or else it is more likely to relate to light jets.
- Secondary Vertex Tagger (SVT): If a jet is close enough to a reconstructed secondary vertex, $\Delta R(SV, jet) < 0.5$, then the jet is tagged by this algorithm.

These three tools are combined using a neural network to construct the D0 NN algorithm [60], which was used in early Run IIa analysis. Outputs of these algorithms are now combined with other variables in the latest b-jet identification algorithm MVA$_{bl}$, used in all

current D0 analysis. This technique increases the tagging efficiency with respect to a single algorithm and provide the best discrimination between b and light jets.

The six RFs inputs to the final MVA_{bl} variable come from two categories: using impact parameter (IP) variables, and using secondary vertex (SVT) variables.

The input variables for the IP RF include the output of JLIP, CSIP, and seven other variables:

1. the reduced JLIP [60], which is computed by removing the track with the lowest probability of originating from the PV and then recalculating the JLIP;
2. the combined probability associated with the tracks with the highest and second highest probability of coming from the PV;
3. the largest distance in ΔR between any two tracks within a jet, or the total ΔR width of the tracks in the jet, $\max[\Delta R(\text{tracks})]$;
4. the sum of the ΔR distances between each track matched to the jet and the centre of the calorimeter jet, $\sum_{trk} \Delta R(trk, jet)$;
5. the p_T-weighted ΔR width of the tracks relative to the calorimeter jet, which is defined as:
$$\Theta = \frac{\sum_{trk} p_T^{trk} \times \Delta R(trk, jet)}{\sum_{trk} p_T^{trk}}$$
6. the total transverse momentum of all the selected tracks;
7. the total number of tracks matched to the jet.

Since various SVT configurations lead to different secondary vertices, the bID group train five SVT RFs, and each SVT RF combines 29 corresponding input variables:

1. the p_T fraction carried by the secondary vertex, p_T^{SVT}/p_T^{jet};
2. the track multiplicity originating from the secondary vertex;
3. the signed decay length significance in the plane transverse to the beam direction;
4. the JLIP probability of the tracks matched to the secondary vertex;
5. the $\chi^2/n.d.f.$ of the tracks matched to the secondary vertex;
6. the number of secondary vertices which can be reconstructed from the tracks matched to the jet;
7. the signed IP of the track with the highest momentum measured transverse to the direction of flight of the secondary vertex;
8. the number of tracks matched to the jets;
9. the pseudo-lifetime computed in the plane transverse to the beam;
10. the decay length measured in the plane transverse to the beam direction;
11. the decay length in the beam direction;

12. the p_T of the highest p_T track in the jet normalized to the p_T^{SVT}, p_T^1/p_T^{SVT};
13. the p_T of the second highest p_T track normalized to the secondary vertex p_T, p_T^2/p_T^{SVT};
14. the dca_{bid} of the secondary vertex to the PV in the plane transverse to the beam;
15. the dca_{bid} of the secondary vertex to the PV in the beam direction;
16. the mass of the secondary vertex (MSV), calculated by summing the invariant masses of all track four-momentum vectors assuming that all tracks originate from pions;
17. the p_T of the track which has the highest momentum measured relative to the direction of flight of the secondary vertex;
18. the momentum of the secondary vertex in the plane transverse to the calorimeter jet direction;
19. the p_T of the highest p_T track normalized to the total jet p_T , p_T^1/p_T^{jet};
20. the p_T of the second highest p_T track normalized to the total jet p_T , p_T^2/p_T^{jet};
21. the opening angle of the secondary vertex projected into the plane transverse to the beam direction;
22. the opening angle of the secondary vertex projected in the beam direction;
23. the Θ (as defined above) as measured for tracks matched to the secondary vertex;
24. max[ΔR(tracks)] of the tracks matched to the secondary vertex;
25. the p_T of the highest p_T track matched to the secondary vertex;
26. the p_T of the second highest p_T track matched to the secondary vertex;
27. the charge (q) of the jet, measured as $\sum_{trk} p_T^{trk} q^{trk}/p_T^{jet}$;
28. the signed decay length significance in the beam direction;
29. the ΔR size of the tracks matched to the secondary vertex.

So after using all the nine inputs to get the impact parameter RF, and after using the twenty-nine inputs to get the five secondary vertex RFs corresponding to five different secondary vertices, the b-ID group combines the IP RF and five SVT RFs as input to get the final MVA$_{bl}$ output.

When compared to the previous D0 b-ID NN algorithm [60], the MVA$_{bl}$ shows a significant improvement, which could be seen from Figure 4.3: at low value of misidentification rate, the MVA$_{bl}$ algorithm gets a higher b-jet identification efficiency, while at high values they tend to be similar.

For being used in all the D0 physics groups, the b-ID group define a set of operating points (OPs), and determine their efficiency and misidentification rates. For the MVA$_{bl}$ algorithm, these operating points are assigned the following names [59] as shown in Table 4.1:

These OPs are displayed in Figure 4.4 where the identification efficiency for b jets and the misidentification rate for light jets are shown as a function of the MVA$_{bl}$ output for simulated events.

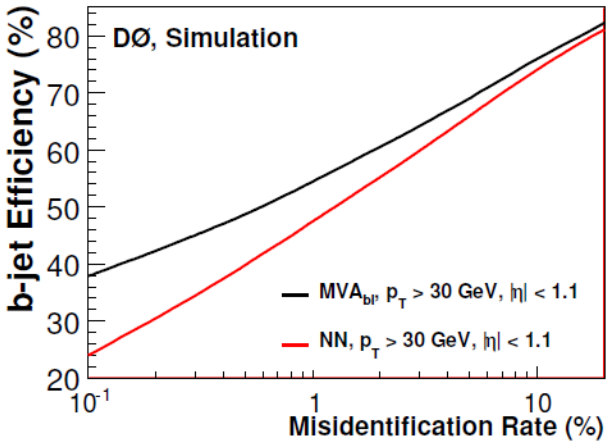


Figure 4.3: The efficiency for selecting a b jet and the misidentification rate as a function of the MVA_{bl} requirement as determined in simulations. The vertical lines correspond to the selected OPs, with (a) linear and (b) logarithmic scales. [59]

Operating Points name	MVA_{bl} criteria
L6	$MVA_{bl} > 0.02$;
L5	$MVA_{bl} > 0.025$;
L4	$MVA_{bl} > 0.035$;
L3	$MVA_{bl} > 0.042$;
L2	$MVA_{bl} > 0.05$;
Loose	$MVA_{bl} > 0.075$;
oldLoose	$MVA_{bl} > 0.1$;
Medium	$MVA_{bl} > 0.15$;
Tight	$MVA_{bl} > 0.225$;
VeryTight	$MVA_{bl} > 0.3$;
UltraTight	$MVA_{bl} > 0.4$;
MegaTight	$MVA_{bl} > 0.5$.

Table 4.1: Operating Points (OP) defined by the b ID group for the latest MVA_{bl} algorithm. [59]

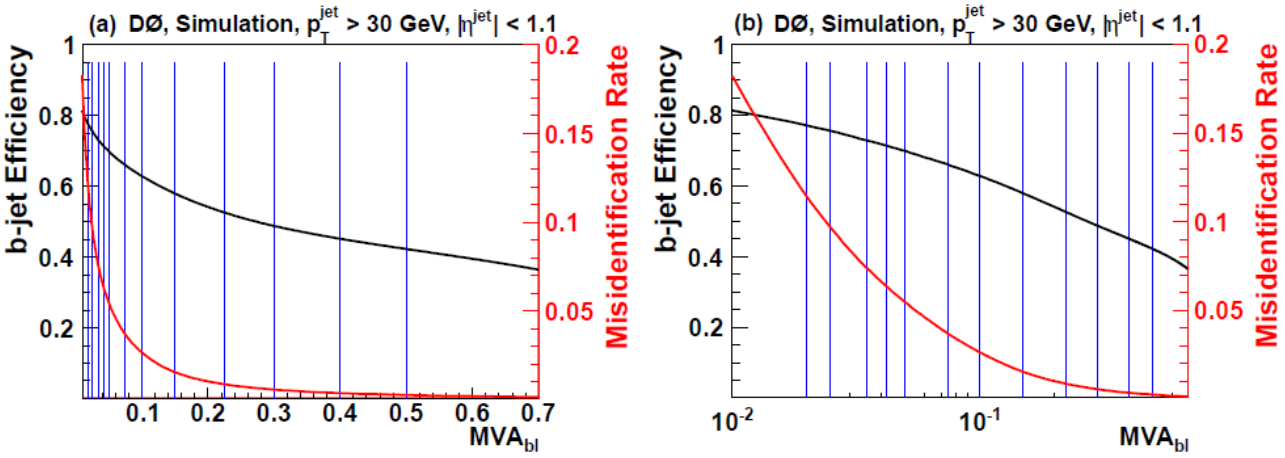


Figure 4.4: The efficiency for selecting a b jet and the misidentification rate as a function of the MVA$_{bl}$ requirement as determined in simulations. The vertical lines correspond to the selected OPs, with (a) linear and (b) logarithmic scales.

4.3.1 MVA$_{bl}$ Efficiency

Using the "System8" (or S8) methodology, described in [59], one can extract the data efficiencies for a set of specific MVA$_{bl}$ requirements. We will describe how to correct differences between MC and data in section 6.4.3. Here we just select two OPs, Loose and Tight for demonstration.

The operating point for the loose (medium, tight) threshold has an identification efficiency of 79% (57%, 47%) for individual b jets, averaged over selected jet p_T and η distributions, with a b-tagging misidentification rate of 11% (0.6%, 0.15%) for light-quark jets (lf), calculated by the method described in [60].

To use S8 method, we combine the c and light jet backgrounds into a single sample (cl jets), since the S8 method only accommodates a single background. And when considering the identification efficiency for c jets, it is assumed that b and c jets having the same data-to-simulation scale factor.

Figures 4.6 and 4.7 show the MC and data corrected efficiencies for b and c jets in dijet events, respectively. These efficiency curves are corrected with the parameterized correction factor derived in Fig 4.5. Figure 4.8 show the total systematic uncertainty for the S8 method, discussed in [60], parameterized as a function of jet p_T .

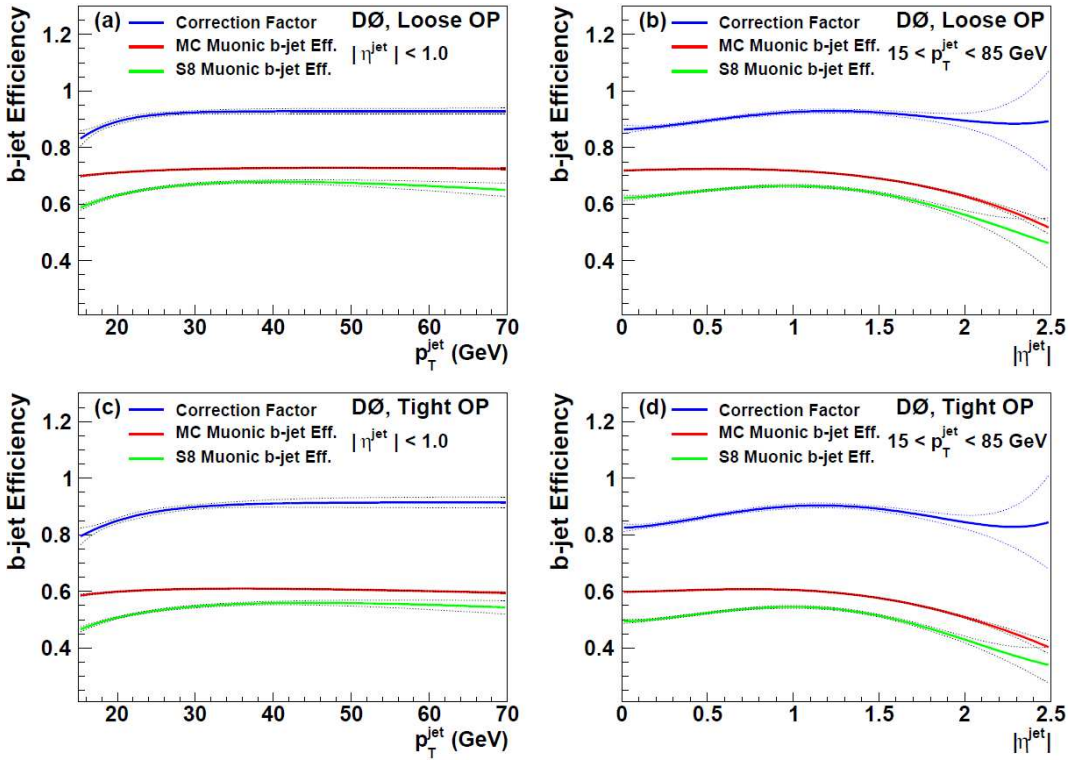


Figure 4.5: The efficiency for selecting a muonic b-jet in MC, shown in red, and data, in green, using the S8 method. The correction factor, SF, which will be used to model the algorithm's efficiency is shown in blue. Two OPs are shown (a,b) the Loose and (c,d) Tight . The efficiencies are parameterized as a function of (a,c) p_T, for central jets and versus (b,d) η. The band which surrounds the lines corresponds to $\pm 1\sigma$ total uncertainties.

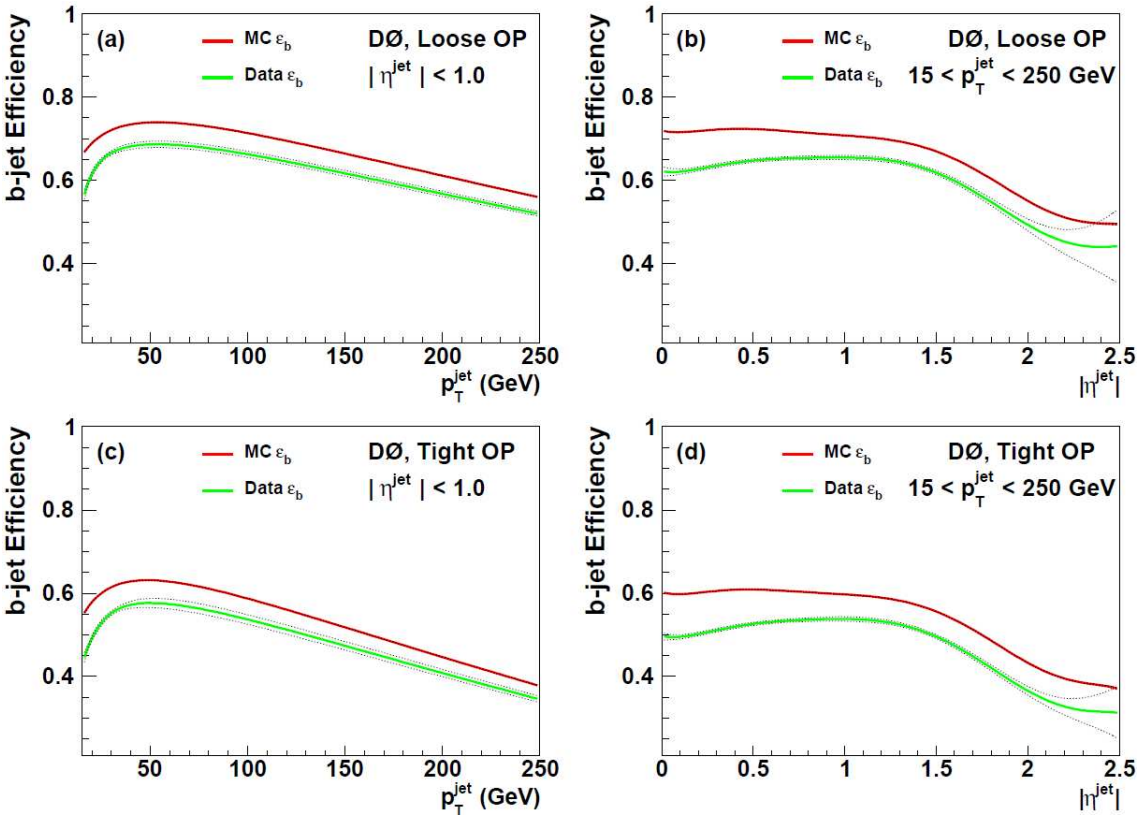


Figure 4.6: The MC b jet identification efficiency, as measured in dijet events, is shown in red along with the data b jet identification efficiency, in green. Two OPs are shown (a,b) the Loose and (c,d) Tight . The efficiencies are parameterized as a function of (a,c) p_T, for central jets and versus (b,d) η.

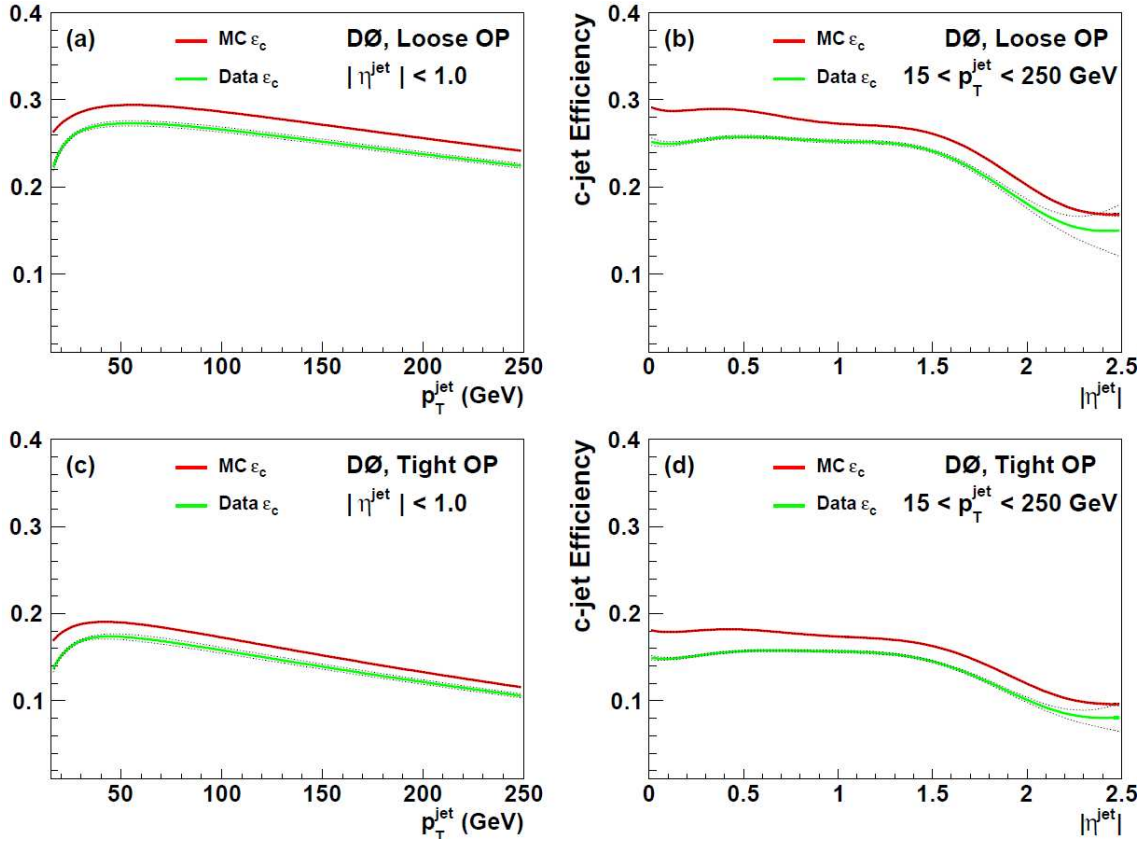


Figure 4.7: The MC c jet identification efficiency, as measured in dijet events, is shown in red along with the data b jet identification efficiency, in green. Two OPs are shown (a,b) the Loose and (c,d) Tight . The efficiencies are parameterized as a function of (a,c) p_T, for central jets and versus (b,d) η.

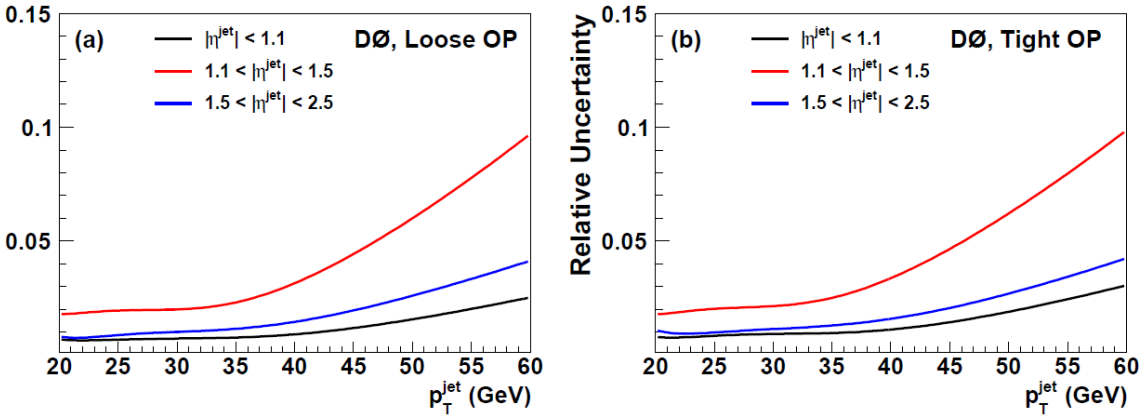


Figure 4.8: The total uncertainty as measured as a function of p_T for two choices of OPs (a) Loose and (b) Tight. Three different ranges of jet η were selected $|\eta| < 1.1$, in red, $1.1 < |\eta| < 1.5$, in green, and $1.5 < |\eta| < 2.5$, in blue.

The latest results of the misidentification rate for light jets extracted from our data are shown in Figure 4.9. Within three different η regions, these misidentification rates are parameterized in terms of p_T. We derive a correction factor through combining this data driven misidentification with the simulated one, and show it in Figure 4.10. These correction factors are then applied to the light jet simulations which passes the MVA_{bl} requirements.

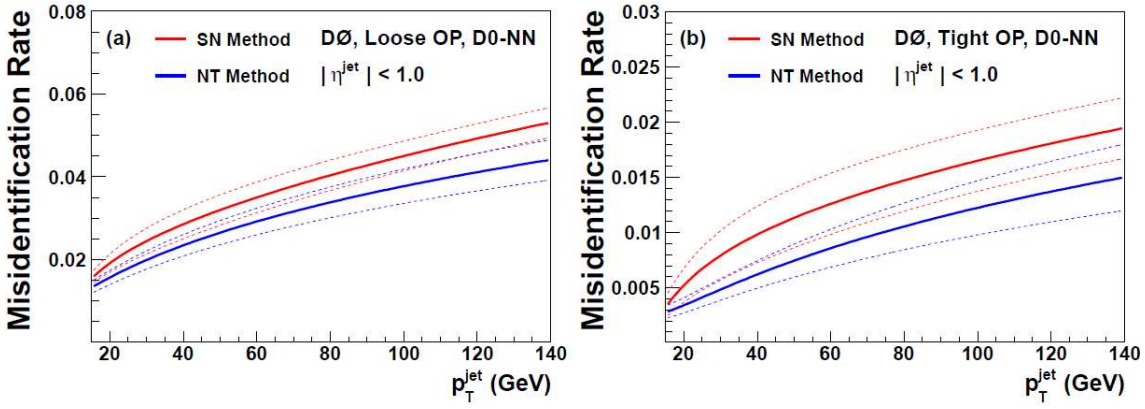


Figure 4.9: An example comparison between the misidentification rates of the D0-NN derived for two OP, (a) Loose and (b) Tight, for the new SN method [59] and the old method described in [60]. The colored bands which surround the values correspond to the total uncertainties.

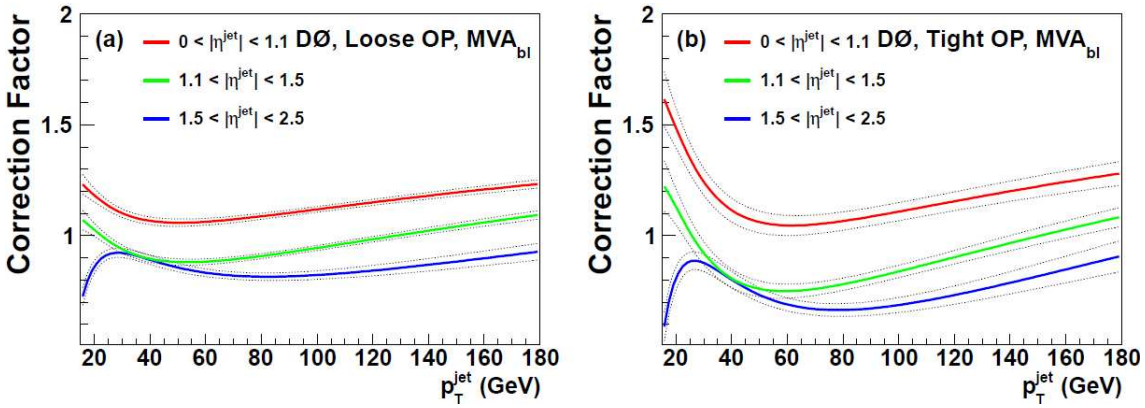


Figure 4.10: The correction factors for the light jet MC which are derived by taking the ratio of the data and MC misidentification rates. Two OPs are shown, (a) Loose and (b) Tight. These are further parameterized over jet p_T and for three different jet η intervals: $0 < |\eta| < 1.1$, $1.1 < |\eta| < 1.5$, and $1.5 < |\eta| < 2.5$. The black dotted lines represent the uncertainty on the fit.

Chapter 5

Monte Carlo Used in the WH Analysis and Event Preselection

As described in Section 1.6.2, the WH production is one of the most sensitive channels searching a low-mass Higgs boson, through looking for the final states with a lepton, a missing transverse energy and a pair of jets originating from b quarks. The leptonic branching ratio $W \to \ell\nu$ is three times lower than the dominant $W \to q\bar{q}'$, but backgrounds from QCD processes are significantly lowered. And also the high momentum lepton, and transverse missing energy is a much more obvious signature, compared to other decay modes, for instance the dominant production $gg \to b\bar{b}$ channel with only jets in final states.

This analysis has first been realized in 2004 and has been published, using 0.17 fb^{-1} of Tevatron RunII data [61]. Another result was published in 2008 using the full RunIIa dataset which represents 1.1 fb^{-1} of data [62]. A result with 5.3 fb^{-1} of Tevatron RunII data was published in 2011 [63] and 2012 [64]. The result which is presented in this document has been published in 2012 with 9.7 fb^{-1} [6] and a longer version of the article has been published recently in 2013 [7].

We cover the analysis strategy in this chapter. In Section 5.1 we introduce the analysis framework. Since the $p\bar{p}$ collision can produce many different particles decaying to different stable particles, constructing various final states, the final topology of an event can match the topology of several processes. It is crucial to differentiate processes with the same topology. To do this, we need firstly estimate the relative contribution from each process to our data, so in Section 5.2 we introduce how to use Monte Carlo simulating the possible outcomes for a $p\bar{p}$ collision and the detector response to the particles produced, and compare this to the data.

After a quick introduction on the analysis framework and introduction of Monte Carlo simulation to the data, we introduce the detailed preselection criteria corresponding to the WH events in Section 5.3.

5.1 Foreword on the Analysis Work Flow

An efficient data analysis framework for this search has been developed and maintained on the basis of the ROOT software, according to the final state topology we choose to search for the Higgs boson, i.e. $WH \to \ell\nu b\bar{b}$. In such a framework we could:

- select events in data that match the topology of our expected $WH \to \ell\nu b\bar{b}$ final state.

- estimate the contribution to our data from various processes that have the same final state topology as the Higgs boson signal, for instance: W+jets, Z+jets, diboson process.
- distinguish signal events from background events
- categorize our events based on b-tagging algorithms (since we are searching for a Higgs boson that decays to a pair of bottom quarks)
- distinguish signal and background using multivariate analysis (MVA) techniques (which are later detailed in Chapter 7)
- perform a statistical analysis to determine whether our data includes the Higgs boson signal, through the CL_S method in 9.1.
- validate our search methodology, through performing a cross check by measuring the cross section of $VZ \to \ell\nu b\bar{b}$

All figures shown in this document related to the WH analysis, published in 2012 [6] and 2013 [7], have been produced using this software. The colour legend for the different components is shown on Figure 5.1, except when the legend is displayed on the figures.

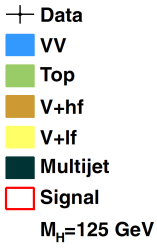


Figure 5.1: Colour legend for data to simulation comparison. Backgrounds denoted as VV (includes WZ, ZZ, and WW), Top (includes $t\bar{t}$ and single top), V+hf (includes $W + b\bar{b}$, $Z + b\bar{b}$), V+lf (includes $W + c\bar{c}$, $Z + c\bar{c}$), Multijet (i.e. QCD). More details on the simulated background samples used in the analysis are given in Section 5.2.2.

5.2 Dataset and Monte Carlo used in the WH analysis

5.2.1 Dataset

The D0 collaboration produces a common analysis format (CAF) of skimmed recorded and simulated data samples, which are used as the data samples in the WH analysis. The average run time to produce these skimmed files is about 24 hours. The output format is a light-weight ROOT tree for each sample with all relevant information to run the final selection code (final objects selection, b-tagging application, application of multivariate techniques and creation of input histograms for the final result extraction). This step is run in approximately an hour.

As introduced in 2.1, the WH analysis uses the data recorded from the D0 detector Run II period. After passing the good data quality criteria, which means all the detector parts working properly, the Run IIa data set has an integrated luminosity of 1.08 fb^{-1}. The Run IIb1 dataset has 1.22 fb^{-1}, the Run IIb2 dataset has 3.04 fb^{-1}, the Run IIb3 dataset has

1.99 fb^{-1}, and the Run IIb4 dataset has 2.40 fb^{-1} after requiring data quality and removing bad luminosity blocks from the luminosity system. The total integrated luminosity is 9.74 fb^{-1}.

5.2.2 Monte Carlo Samples and Generators

We account for all Higgs boson production and decay processes that can lead to a final state containing exactly one charged well isolated lepton, $\not\!\!E_T$, and two or more jets. The signal processes considered are:

- Associated production of a Higgs boson with a vector boson where the Higgs boson decays to $b\bar{b}$, $c\bar{c}$, $\tau\tau$, and the associated weak vector boson decays leptonically ($WH \to \ell\nu + c\bar{c}$, $WH \to \ell\nu + \tau\tau$, $ZH \to \ell\ell + b\bar{b}$, $ZH \to \ell\ell + c\bar{c}$, $ZH \to \ell\ell + \tau\tau$. Note that contributions from $Z(\to \ell\ell)H(\to b\bar{b})$ production arise from identifying only one charged lepton in the detector, with the other contributing to the $\not\!\!E_T$.)

- Associated production of a Higgs boson with a vector boson where the Higgs boson decays to VV, and the associated weak vector boson decays hadronically ($WH \to WWW \to \ell\nu jjjj$ $ZH \to ZWW \to \ell\ell jjjj$, we don't reconstruct all four jets.)

- Higgs boson production via gluon fusion with the subsequent decay $H \to VV$, where one weak vector boson decays leptonically (with exactly one identified lepton)($gg \to H \to WW \to \ell\nu jj$, $gg \to H \to ZZ \to \ell\ell jj$, where again one of the leptons is not reconstructed)

- Higgs boson production via vector boson fusion with the subsequent decay $H \to VV$, where one weak vector boson decays leptonically (with exactly one identified lepton)($H \to WW \to e\nu jj$, $H \to WW \to \mu\nu jj$, $H \to WW \to \tau\nu jj$, $H \to ZZ \to \ell\ell jj$)

The principle of events simulation is to expand the quantum field theory equations perturbatively, with a Monte-Carlo (MC) random generator. Due to numerical or theoretical challenges, such expansion usually be expanded to just leading order (LO) or next-to-leading order (NLO). At D0, we usually use the LO terms to simulate an object's kinematics, whereas we use the NLO calculations to normalize the cross section of a given process and correct the kinematics.

When performing the simulation, we need to consider the following four effects:

- the constituent quarks and gluons interact when $p\bar{p}$ collides, and carry some fraction of the proton or anti-proton momentum.

- the possibility that the interacting and/or resulting particles can radiate gluons or photons.

- the process of quarks forming hadrons (hadronization).

- the passage of particles through our detector.

PYTHIA is a Monte Carlo event generator calculating Feynman diagrams at the tree-level (or Leading Order perturbation theory (LO)), higher order loop calculations not being performed [55]. We produce the signal and diboson simulated samples withing PYTHIA [55] asking use of the leading order parton distribution functions CTEQ6L1 [65, 66].

ALPGEN is a tree-level matrix element (ME) calculator for a fixed number of partons in the final-state ($2 \to N$ processes). The $t\bar{t}$ and W+jets, Z+jets events have been generated with the matrix element generator ALPGEN [67] interfaced to PYTHIA for subsequent parton shower and hadronization.

The ALPGEN samples have been produced in exclusive bins of "light" (i.e. gluons or u, d, s quarks) parton multiplicity except for the "highest" bin obtained in an inclusive way, i.e. it includes higher multiplicities as well. ALPGEN uses the matching prescription as described in [68]. All ALPGEN W/Z+jets , $W + b\bar{b}$ and $W + c\bar{c}$ have undergone a process of heavy flavour (HF) skimming; that is, additional heavy-flavoured partons generated by PYTHIA have been removed so as not to double-count events. Specifically, $W + jj$ ($j = u$, d, s; "light jet") samples are skimmed to remove $W + c\bar{c}$ and $W + b\bar{b}$ events; the $W + c\bar{c}$ samples are not skimmed; and the $W + b\bar{b}$ samples are skimmed to remove $W + c\bar{c}$ events. The W+jets, signal, and $t\bar{t}$ samples include all three lepton flavours for the weak decays of the W boson. Other samples, such as $Z + HF$ and single-top, have been produced in bins of lepton flavor; they have been grouped in the tables for enhanced readability. The single-top sample was produced using CompHep [69].

All simulated events have been processed through the D0 detector simulation [70] based on the detailed detector material simulation package GEANT [71], the D0 detector electronics simulation (d0sim), and the reconstruction software (d0reco). Because of major upgrades, i.e installing Layer 0, between Run IIa and Run IIb, different sets of Monte Carlo samples are used in this analysis. In RunIIa, MC are simulated by version p17, and p20 is used in Run IIb MC samples. We additionally use different versions of the p20 MC for Run IIb1 (p20.09.03), Run IIb2 (p20.15.04), and Run IIb3 + Run IIb4 (p20.17.02) to account for changes to the primary vertex algorithm, changing tracking efficiency, and calorimeter calibrations between the different Run IIb datasets. The simulated events have been reweighted with the trigger efficiencies and with all the ratios of efficiencies of data to simulation (scale factors). Depending on the efficiency considered, these scale factors are constant or have a kinematic dependency which is taken into account.

A full list of MC samples used in this analysis can be found in [72].

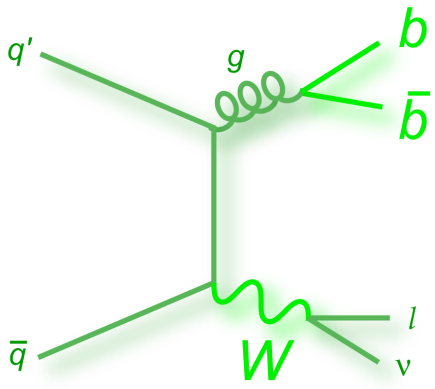
Feynman diagrams for the samples used are displayed in Figure 5.2.

This simulation process does not perfectly reproduce the data. We study these differences, and then apply corresponding corrections.

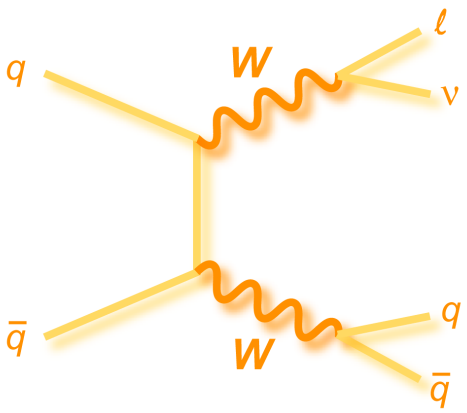
- Our simulation assumes only one proton interacting with one anti-proton in each event, in fact multiple interactions happen with higher luminosity events. So we overlay the simulation with data events from randomly chosen minimum bias (MB) events from different instantaneous luminosities.

- Our simulation assumes a Gaussian distribution with the primary interaction location along the z axis, in fact the distribution is not quite Gaussian. So the simulation is reweighted to match the distribution in data.

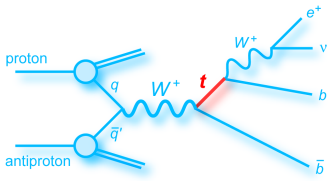
- When testing the Z boson p_T simulation through looking at $Z/\gamma^\star$ events decaying to two leptons, it was found that at low Z boson p_T, the simulation does not describe the data well in either ALPGEN or PYTHIA. A reweighting was derived such that simulated p_T spectrum matches that in data [73]. For the W boson p_T case, there is no measurement to derive an equivalent reweighting, so the W p_T is corrected using the Z p_T reweighting, adjusted by the ratio of the W to Z differential cross sections at next-to-next-to-leading-order (NNLO).



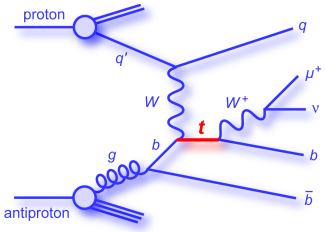
(a) Associated production of a W boson and a pair of b quarks from gluon-splitting.



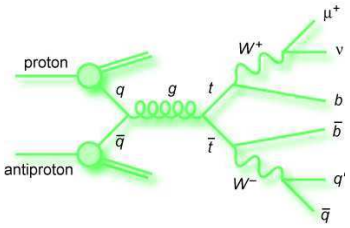
(b) Pair production of W bosons.



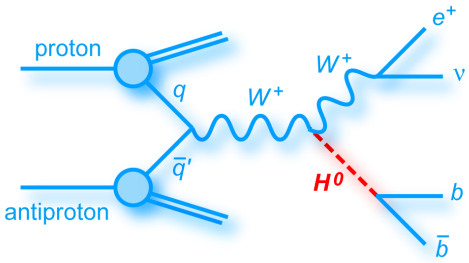
(c) Single top production in the s channel.



(d) Single top production in the t channel.



(e) $t\bar{t}$ production in the semileptonic final state.



(f) Associated production of a W and Higgs bosons.

Figure 5.2: Example Feynman diagrams of Monte Carlo processes used in the analysis.

Many thousands of events are generated for the simulation of each process. These must be normalized to the rates presented in our data. For classical particles with a density ρ, colliding with a speed v, the rate at which we expect a process with a cross section σ to occur is given by: $R^{classical} = v\rho\sigma$. We can express the analogous rate in quantum mechanics as: $R^{QM} = \mathcal{L}\sigma\eta$, where $\mathcal{L}$ is the luminosity of the colliding beams, σ is the cross section for the process under consideration, and η is the detector efficiency.

The cross section is calculated from theory, and the luminosity and efficiency are both measured from data, giving us everything we need to normalize our simulated samples. The signal cross sections and branching fractions are normalized to the SM predictions [74].

The cross sections for Higgs boson production used in this analysis are calculated at Next-to-Next-to-Leading order as recommended by the Tevatron Higgs Working Group [19], i.e. WH/ZH is calculated in [75], Vector Boson Fusion (VBF) in [76], and for gluon fusion, updated calculation from [77] is used. The cross sections for other SM processes used in the analysis are calculated at Next-to-Leading Order (NLO), (NNLO for $t\bar{t}$) [19]. The V+jets processes are normalized to data for each lepton flavour and jet multiplicity separately.

All the cross sections of the above signal processes are listed in Table 5.2. Cross section times branching ratios for the $WH \to \ell\nu + b\bar{b}$ is listed in Table 5.1. Cross section times branching ratios for other signal samples are listed in [72]. The cross sections for diboson simulated samples are given in Table 5.3.

$WH \to \ell\nu + bb$	
mass	$\sigma(\times\text{BR})$[pb]
$m_H = 90\,\text{GeV}$	0.10384
$m_H = 95\,\text{GeV}$	0.08651
$m_H = 100\,\text{GeV}$	0.07204
$m_H = 105\,\text{GeV}$	0.05978
$m_H = 110\,\text{GeV}$	0.04917
$m_H = 115\,\text{GeV}$	0.03986
$m_H = 120\,\text{GeV}$	0.03156
$m_H = 125\,\text{GeV}$	0.02425
$m_H = 130\,\text{GeV}$	0.01793
$m_H = 135\,\text{GeV}$	0.01272
$m_H = 140\,\text{GeV}$	0.00861
$m_H = 145\,\text{GeV}$	0.00552
$m_H = 150\,\text{GeV}$	0.00328

Table 5.1: *List of simulated $WH \to \ell\nu b\bar{b}$ signal samples. We use* `PYTHIA` *to generate signal events. Also shown are the cross section times branching ratio for one lepton flavour.*

5.3 Event Preselection

Since we are searching the $WH \to \ell\nu b\bar{b}$ signal, we look for exactly one high energy lepton (either an electron or muon), a large amount of $\not\!\!E_T$ from the neutrino in the W decay, and two or three jets. Consequently we apply the following criteria to these objects:

- lepton $p_T > 15$ GeV, to cut down on the number of multijet events without cutting out too many signal events from our sample.
- $|\eta| < 2$(muon); $|\eta| < 1.1$ or $1.5 < |\eta| < 2.5$(electron).

$m_H = 125\,\text{GeV}$	
signal	$\sigma(\times\text{BR})$[pb]
$WH \to \ell\nu + c\bar{c}$	0.00112
$WH \to \ell\nu + \tau\tau$	0.00267
$ZH \to \ell\ell + b\bar{b}$	0.00458
$ZH \to \ell\ell + c\bar{c}$	0.00021
$ZH \to \ell\ell + \tau\tau$	0.00050
$H \to WW \to e\nu jj$	0.02980
$H \to WW \to \mu\nu jj$	0.02930
$H \to WW \to \tau\nu jj$	0.03119
$H \to ZZ \to \ell\ell jj$	0.00358
$vbf \to H \to WW \to \ell\nu jj$	0.00636
$WH \to WWW \to \ell\nu jjjj$	0.02797
$ZH \to ZWW \to \ell\ell jjjj$	0.01696

Table 5.2: *List of simulated signal samples other than $WH \to \ell\nu b\bar{b}$. We use* `PYTHIA` *to generate signal events. Also shown are the cross section times branching ratio for one lepton flavour.*

Process	σ
WW inclusive	11.34
WZ inclusive	3.22
ZZ inclusive	1.20

Table 5.3: *List of simulated di-boson processes, and cross section. We generate all diboson MC events with* `PYTHIA`.

- $\not\!\!E_T > 15$ GeV(electron); $\not\!\!E_T > 20$ GeV(muon)($\not\!\!E_T$ is the missing transverse energy).
- jet $p_T > 20$ GeV.
- A primary vertex with at least three associated tracks, and $|z| < 60$ cm.
- The radial distance between the primary vertex and the interaction point be less than 1 cm.
- $M_T^W > 40$ GeV$-0.5 \times \not\!\!E_T$(M_T^W is the transverse mass of the reconstructed W boson).

The criteria will be discussed more in detail in the following sections.

5.3.1 Primary Vertex Selection

Since we are looking for the $WH \to \ell\nu b\bar{b}$ signal, i.e. the final objects including a charged lepton, and two charged b jets, we require that the event have a primary vertex (PV) with at least three associated tracks. And the PV position should be within the coverage of the SMT, i.e. $|z| < 60$ cm. Since the primary vertex should be matched to the interaction point, we require the radial distance between the primary vertex and the interaction point be less than 1 cm.

5.3.2 Lepton Selection

We want to cut down the multijet events without cutting out too many signal events, this could be done by applying the lepton $p_T > 15$ GeV cut, since real leptons decaying from W bosons tend to have larger p_T, compared to the misidentified leptons in the multijet sample.

To make sure the muon passes through the acceptance of the muon system, we require the muon $|\eta| < 2$. And for electron passing through the CC or EC, we require the electron $|\eta| < 1.1$ or $1.5 < |\eta| < 2.5$

Based on the lepton isolation, there are two categories of leptons: the "loose" and "tight" isolation leptons. The "loose" leptons are used to estimate the multijet background from data, and the "tight" leptons are used to perform the analysis. For electrons, the loose and tight isolation is defined by different requirements on the electron ID MVA output.

For a loose isolation muon, we only require it is outside of a reconstructed jet, $\Delta R(\mu, jet) > 0.5$. For a tight isolated muon, we expect that most of the momentum from charged particles near the muon should come from the muon itself, so we apply the cut $\frac{\sum_{\Delta R=0.5} p_T}{p_T^{muon}} < 0.4$, where $\sum_{\Delta R=0.5 tracks} p_T$ stands for the scalar sum of the p_T of the tracks within $\Delta R = 0.5$ of the muon. To reduce the probability that a neutral particle has deposited energy in the calorimeter near the muon, we apply the cut $\frac{\sum_{0.1<\Delta R<0.4} p_T}{p_T^{muon}} < 0.12$, where $\sum_{0.1<\Delta R<0.4 tracks} p_T$ stands for the transverse energy deposited in a hollow cone of $0.1 < \Delta R < 0.4$ around the muon.

Event Trigger

By triggering on a lepton we can greatly reduce the number of background event coming from multijet production, since leptons are produced less frequently than hadronic particles in a hadron collider.

Events containing an electron are selected by a logical OR of triggers that require an electromagnetic object, as described in [64]. The triggers' efficiencies are measured in data

and also applied to the MC simulation. These efficiencies are parametrized in electron η, azimuthal angle ϕ [78], and transverse momentum p_T. Depending on the trigger, and the electron's position within the detector, these efficiencies are $90 - 100\%$.

Events containing a muon are selected based on the logical OR of all available triggers, except those that contain lifetime based requirements that are used for b-jet identification. We determine the trigger efficiency in this inclusive trigger sample with the following two steps:

1. We consider a subset of triggers based on the logical OR of single muon and muon+jets triggers ($T_{\mu OR}$). This subset models the data well, and are about 70% efficient.

2. Compare the data in this subset to that in the inclusive trigger sample, and we can determine an inclusive trigger correction for the MC trigger efficiency, P_{corr}:

$$P_{corr} = \frac{(N_{Data} - N_{MJ})_{incl} - (N_{Data} - N_{MJ})_{T_{\mu OR}}}{N_{MC}} \tag{5.1}$$

   where the numerator is the difference (after subtracting off the multijet background component) between the number of data events in the inclusive and $T_{\mu OR}$ trigger samples, and the denominator is the total number of MC events.

3. The total trigger efficiency is estimated as $T_{\mu OR} + P_{corr}$, and is limited to be ≤ 1.

The correction P_{corr} is derived as a function of the sum of jet $p_T(H_T)$, and the $\not\!\!E_T$, in bins of muon η. In different regions of the detector, different triggers dominate chosen η bins. For $|\eta| < 1.0$, events are dominantly triggered by single muon triggers; while for $|\eta| > 1.6$, the muon+jets triggers are dominant; and for $1.0 < |\eta| < 1.6$, is a mixture of single muon and muon+jets triggers.

Detector supports allow only partial coverage of the muon detector in the $|\eta| < 1.0$ and $1.0 < |\eta| < 1.6$ region. Such a support structure impacts the trigger efficiency in these regions, so we derive separate corrections correspondingly. Examples of these corrections can be seen in Figure 5.3. This strategy results in a trigger efficiency of $80-100\%$, depending on the H_T and the muon's position in the detector.

5.3.3 Missing E_T Selection

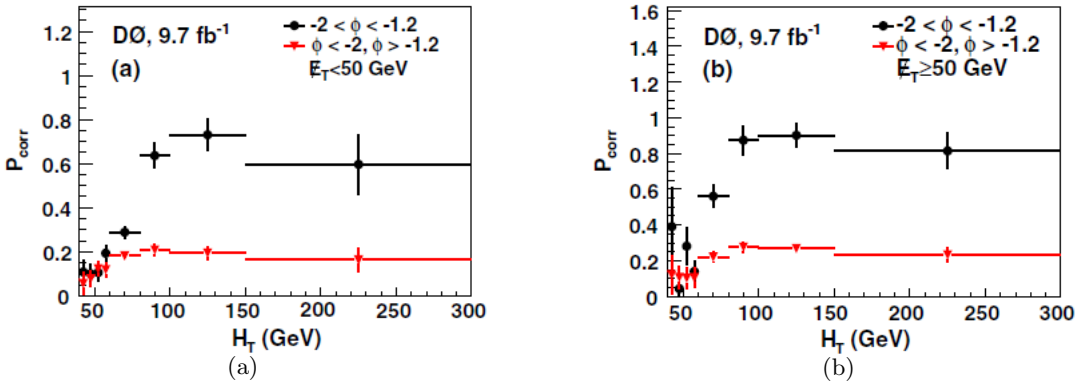
When the W boson decaying leptonically, one of the two daughter particles is a neutrino. This particle does not interact with the detector and thus its presence is characterized by missing energy in the transverse plane to the beam direction. As seen in Section 3.5, the missing transverse energy is recalculated after all EM objects and jets are corrected, and energy deposit from muons in the calorimeter is taken into account.

The requirement is set to $\not\!\!E_T > 15$ GeV for the electron channel to ensure orthogonality with the sample in which we estimate the multijet contribution to our data, and $\not\!\!E_T > 20$ GeV for the muon channel to further suppress the multijet background.

$\not\!\!E_T$ distributions for the the electron and muon channels are shown in Figure 6.7, 6.8 for the Run IIb period.

5.3.4 Jet Selection

The Standard Model predicts that the Higgs boson decays dominantly in a $b\bar{b}$ pair, in $M_H < 135$ GeV mass region. By analyzing the 2 jets channel, most of the sensitivity is



(a) (b)

Figure 5.3: Data-derived muon trigger correction to account for the resulting efficiency gain in moving from single muon and muon+jets triggers to inclusive triggers as a function of H_T for $|\eta| < 1.0$, shown (a) for events with $\not\!\!E_T < 50$ GeV and (b) for events with $\not\!\!E_T \geq 50$ GeV. The black circles show the correction when the muon is in the region of $\phi(-2 < \phi < -1.2)$ where there is a gap in the muon coverage for detector supports, and the red triangles show the correction elsewhere in ϕ.

obtained. But one can recover events which radiated an additional jet. Therefore, the analysis is split in two orthogonal channels (2 jets and 3 jets) and are analyzed separately and then combined at the limit calculation step. Since jets in our signal are coming from the decay of a Higgs, they will tend to have higher p_T than jets coming from gluon splitting as the jets in our $Wb\bar{b}$ background do. We therefore require that the jet p_T be larger than 20 GeV, after satisfying the identification criteria described in 3.4 and applying the corresponding corrections.

Figure 6.7, 6.8 shows the p_T, η and ϕ distributions for the leading p_T and next-to-leading p_T jets.

5.3.5 Triangular Cut

Since the multijet background consists of events with lepton and neutrino not coming from a real W boson, the transverse mass of the reconstructed W will tend to be low for these events. Since a straight cut on a single distribution is not the most efficient way to get rid of this background which is significant especially in the electron channel, events passing the following two dimensional cut are kept:

$$M_T^W > 40GeV - 0.5 \times \not\!\!E_T \tag{5.2}$$

This cut ensures the agreement between data and simulation is good for both low $\not\!\!E_T$ and low M_T^W regions. By applying the triangular cut, a large reduction of the multijet background is shown in Figure 5.4(b).

5.3.6 Vetoes

Since this result is a part of the D0 and Tevatron Standard Model Higgs combination, a particular attention has to be drawn to avoid biases in the process of combining several channels which can contain the same topology. It is thus important to avoid analyzing twice the same events in different analyses. Therefore a set of requirements are applied

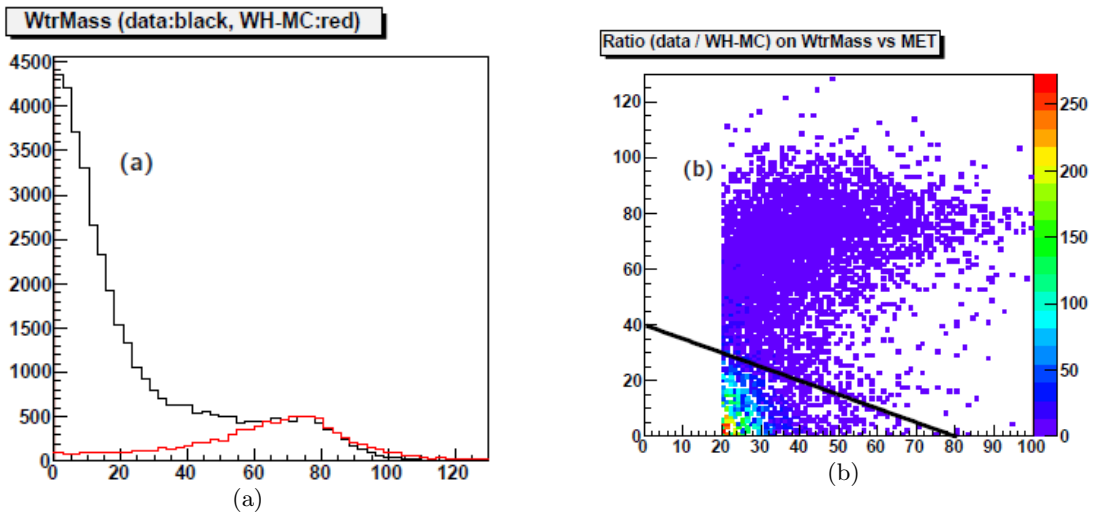


Figure 5.4: (a) M_T^W using loose electron criteria in data (black) as multijet sample, and WH MC (red). (b) 2D ratio plot of data/WH MC in the (M_T^W, $\not\!\!E_T$) plane. The multijet events are lying at low $\not\!\!E_T$ and low M_T^W and can therefore be eliminated by a triangle cut of $M_T^W > -0.5\not\!\!E_T + 40$ as shown by the black line. The WH MC is normalized on the W peak.

to avoid this potential bias when selecting events passing selection in another analysis. In the case of the WH analysis, orthogonality cuts are defined to remove events selected by the $ZH \to \ell\ell bb$ and $X + H \to \tau\tau jj$ analyses. The same object selection applied in these analyses is used to reject events that will fulfill such requirements.

In order to minimize backgrounds that contain two isolated leptons in the final state, such as those coming from Z and $t\bar{t}$ production, events are required to not contain any additional isolated lepton.

In the $WH \to e\nu b\bar{b}$ channel, events which additionally contains isolated muons[1] with $p_T > 15$ GeV, are rejected. Similarly, in the muon channel, events with an additional LOOSE electron are not considered.

Tau Veto

Hadronic tau candidates that passes the $X+H \to \tau\tau jj$ selection [79] are vetoed. Candidates can be any of the three D0 standard tau types [80], with the following requirements for each type:

Type1 $E_T > 10$ GeV, $p_T > 7$ GeV, $E_T/p_T > 0.5$, $NN_\tau > 0.9$

Type2 $E_T > 10$ GeV, $p_T > 5$ GeV, $NN_\tau > 0.9$

Type3 $E_T > 15$ GeV, $p_T > 7$ GeV, $\sum_{trk} p_T^{trk} > 15$ GeV, $NN_\tau > 0.95$

where p_T is the momentum of the associated track, NN_τ is the hadronic tau identification tool (analogous to the NN b-tagger) and $\sum_{trk} p_T^{trk}$ is the sum of tracks p_T associated to the τ candidate. This condition applied in both electron channel muon channels.There is a

[1]The definition of muons to be discarded in the analysis is: MediumNseg3 muon quality, Medium track quality, ΔR isolation requirements.

negligible loss in data and WH signal for both the electron and muon channels, while the largest difference can be seen in ZH and ZZ due to type 2 τs in the electron channel. The $t\bar{t}$ background is also reduced by 8% , mainly due to di-lepton final states where there can be a tau in the final state.

$ZH \to \ell\ell bb$ Veto

The $ZH \to \ell\ell bb$ [81] is split in different sub-channels. This search is performed on dilepton events with 2 or more jets. The dilepton pair can either be two muons or two electrons. A second lepton veto allows the WH analysis to be orthogonal to these two channels.

In order to ensure orthogonality with the $ZH \to \mu\mu b\bar{b}$ analysis, a veto is applied on events with two loose track-matched muons with $p_T > 10$ GeV, in the $60 < M_{\mu\mu} < 150$ GeV range. This allows as well to reject $Z \to \mu\mu$ and $t\bar{t}$ background events.

In the case of the $ZH \to eeb\bar{b}$ analysis, events are required to not contain additional electrons with $p_T > 15$ GeV that satisfy the loose electron criteria.

In order to increase their sensitivity, two alternate channels are considered in the analysis of final state with two leptons and two jets in order to increase the acceptance. The dilepton pair can also be a muon and an isolated track (to recover di-muon events with one of the lepton not having hits in the muon spectrometer) or an electron and a "ICR electron" (e_{ICR}).

Veto on $e + e_{ICR}$ Events

To reproduce the candidates of the $e + e_{ICR}jj$ analysis, the selection for the $e + e_{ICR}$ pair is done as following:

electron $p_T > 20$ GeV

ICR electron $p_T > 20$ GeV

Events with a reconstructed (electron,e_{ICR}) invariant mass greater than 5 GeV are vetoed. There is almost no data loss due to this veto. The largest differences can be seen in ZH and ZZ, which loose about 7%.

Veto on $\mu + \mu_{trk}$ Events

To reproduce the candidates of the $\mu + \mu_{trk}jj$ analysis, the selection for the μ+ track pair is done as the following:

muon $p_T > 10$ GeV

track $p_T > 20$ GeV

Events with a reconstructed (μ,trk) invariant mass between 60 GeV and 150 GeV are rejected.The largest differences can be seen in ZH and ZZ, where a loss of about 50% is observed, which indicates an effective $\mu\mu_{trk}$ selection.

Chapter 6

Treatment of the Background and Result of Event Selection

After selecting events, we will introduce the treatment of the various background in this Chapter. Since we are looking for a Higgs boson that decays to b-quarks, we employ b-tagging to categorize our candidate events in Section 6.4. Then we will show the event selection result in Section 6.5.

6.1 Reweighting of W+jets and Z+jets Samples

To correct the deficiency coming from certain poorly reproduced kinematic distributions by Alpgen, we derive Alpgen reweightings for our W/Z+jets samples. To minimize any possible signal contamination, we derive the reweightings before the application of b-tagging. There are three kinds of reweightings: the jet η, the lepton η, and the $p_T^W \sim \Delta R$ correlation reweighting.

- The jet η reweighting is to improve the description of jet angles for the two highest p_T jets. The correction has a fourth-order polynomial form, determined by fitting the ratio $\frac{N_{Data}-N_{QCD}-N_{SM-Vjets}}{N_{Vjets}}$.
- The lepton η reweighting is done in W+jet events by a second-order polynomial.
- The $p_T^W \sim \Delta R$ correlation reweighting is done for an observed discrepancy between our data and simulation that is correlated between the W boson $p_T(p_T^W)$ and $\Delta R(j1, j2)$. This two dimensional function is a product of a third-order polynomial in ΔR, and an error function plus a Gaussian plus a constant in p_T^W. The p_T^W reweighting is applied to W+jet events, and the ΔR correction is applied to both W and Z+jet events.

We always only use the two jets with highest p_T for the reweightings, even in three or more jet events. Although the reweightings indeed alter the shape of the distributions, we ensure that the functions preserve the total normalization of the `ALPGEN` samples. This is derived in the single muon OR sample and is applied to the electron channel. For the inclusive trigger muon sample, the same jet and lepton eta reweighting functions are used, but the W p_T and $\Delta R(jet1, jet2)$ are re-derived after the muon inclusive trigger correction is applied. These reweighting functions can be seen in Figure 6.1 and 6.2.

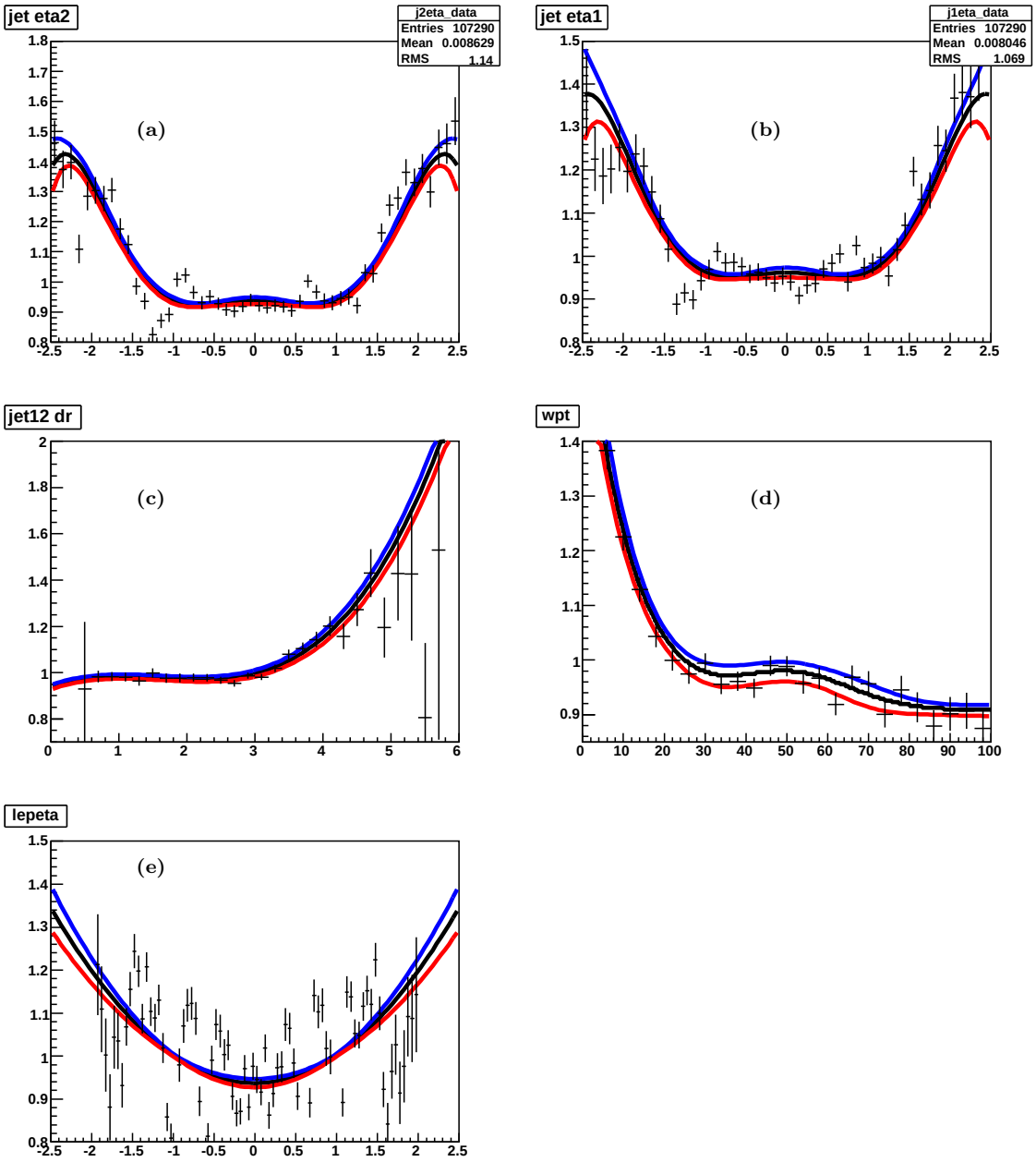


Figure 6.1: *Reweighting functions applied to correct the mis-modeling in* ALPGEN *V+jets MC samples of (a) η of the second leading jet, (b) η of the leading jet, (c) $\Delta R(jet1, jet2)$, (d) p_T of the W system, (e) η of the lepton. In all plots, the black curve is the nominal reweighting function, while the red and blue curves are uncertainties on the fit.*

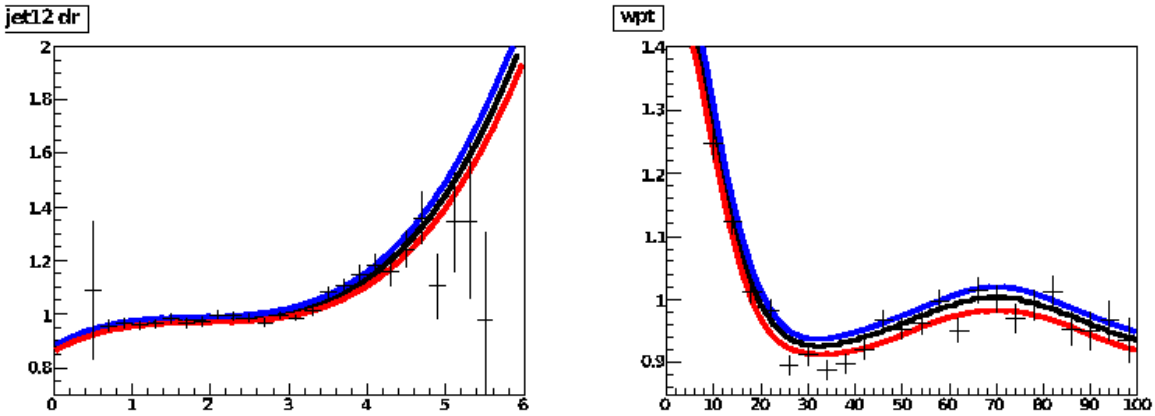


Figure 6.2: *Reweighting functions $\Delta R(jet1, jet2)$ (left) and p_T of the W (right) that are re-derived after applying the inclusive trigger correction and subsequently applied only to the inclusive muon sample. In all plots, the black curve is the nominal reweighting function, while the red and blue curves are the uncertainties on the fit.*

6.2 Multijet Background

A large amount of $p\bar{p}$ collisions will result in production of multiple jets. The lepton event trigger introduced in 5.3.2 can cut down on this background, however a jet is possibly misidentified as a lepton, especially if most of the jet's energy goes into a single pion upon hadronization. A charged pion can decay as $\pi^+ \to \mu + \nu_\mu$, which can lead to a misidentified muon. A neutral pion generally decays to a pair of photons, which can lead to a misidentified electron. It is also possible that the jet simply produces a shower shape similar to what expected for an electron, then we would mis-reconstruct an electron instead of a jet.

We estimate the multijet background from data. As we introduced in 5.3.2, based on the lepton isolation, there are two categories of leptons: the "Loose" and "Tight" isolation leptons. We define two orthogonal samples with similar kinematic distributions: events with a lepton that passes "Loose" isolation, and events with a lepton that passes "Tight" isolation. "Loose" events are used to estimate the multijet contribution, and "Tight" events are used in the analysis sample.

We apply a low $\not\!\!E_T$ cut ($5 < \not\!\!E_T < 15$ GeV) to estimate multijet contribution, since in this case we expect that usually the $\not\!\!E_T$ we measure to be from a mismeasurement instead of from a neutrino from a W boson decay. Then we reweight the events in data to properly estimate the number of events contributing to our analysis sample, through a template created using a modified version of the Matrix Method.

6.2.1 Multijet Background Modelling Strategy

To estimate multijet events, the matrix method works by forming the following pair of equations, which could be solved simultaneously:

$$N_L = N_\ell + N_{MJ}, \quad N_T = \varepsilon_\ell N_\ell + f_j N_{MJ} \tag{6.1}$$

where $N_L(N_T)$ is the number of events in data with a lepton that passes a loose (tight) isolation requirement, $N_\ell(N_{MJ})$ is the number of event in data with a real (misidentified)

lepton with the loose isolation requirements, ε_ℓ is the efficiency for a real lepton that passes the loose identification requirement to subsequently pass the tight identification requirement, and f_j is the rate at which a jet (that has been misidentified as a lepton) that passes the loose isolation requirement to subsequently passes the tight isolation requirement (f_j is so called the lepton fake rate).

Solving this system of equations for the number of multijet events in the sample with tight isolation requirements:

$$N_{MJ}^{T} = \frac{f_j}{\varepsilon_\ell - f_j}(\varepsilon_\ell N_L - N_T) \tag{6.2}$$

We can then weight each event in our templates by:

$$w = \frac{f_j}{\epsilon_\ell - f_j}(\epsilon_\ell - \Theta^T) \tag{6.3}$$

where $\Theta^T = 1$ if the event satisfies tight lepton identification criteria and 0 otherwise. The total multijet events in the tight sample is then the sum of weights over all events that pass the loose isolation requirements:

$$N_{MJ}^{T} = \sum_{i=1}^{N^L} w_i \tag{6.4}$$

Since this matrix method uses all events that pass the loose isolation requirements, events with a real lepton that passes both the loose and tight isolation requirements will contribute with a negative weight, and the multijet sample is statistically dependent from the data sample, which could potentially result in correlated fluctuations. To avoid this, we employ a slightly modified version of the matrix method where we only consider the events that pass the loose lepton isolation requirements, but fail the tight lepton isolation requirements (so-called loose-not-tight events). These events are weighted by:

$$w_i = \frac{f_j}{1 - f_j} \tag{6.5}$$

This will result in a multijet prediction of:

$$N_{MJ}^{T} = \sum_{i=1}^{N^{L-n-T}} w_i \tag{6.6}$$

where, N^{L-n-T} is the number of events in the loose-not-tight sample.

The real lepton efficiencies for electrons, ε_e, are parametrized as a function of detector $\eta(|\eta_{det}|)$, electron p_T, and instantaneous luminosity; while for muons, ε_μ is parametrized as a function of muon p_T.

6.2.2 Lepton Fake Rates

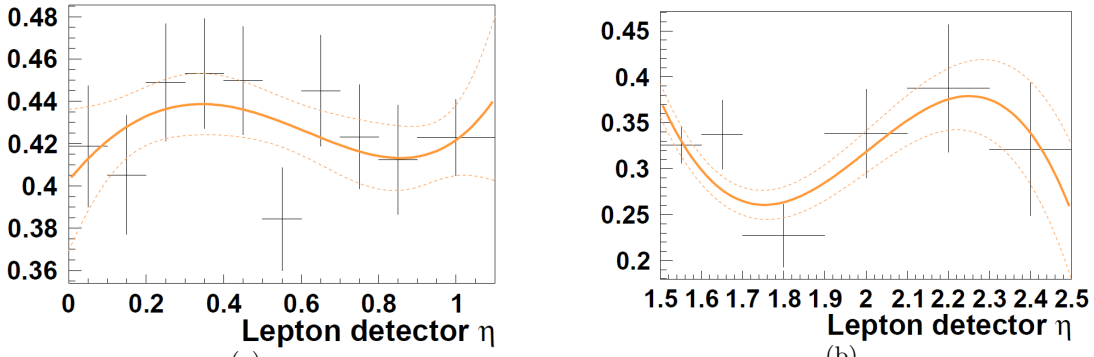
The lepton fake rate f_j, i.e. the efficiency with which jets (or photons in the electron channel) are misidentified as leptons. Electron fake rate, f_j^e is parametrized as a function of electron p_T , detector $\eta(|\eta_{det}|)$, and $\Delta\phi(\not\!\!E_T, e)$. Muon fake rate, f_j^μ is parametrized as a function of muon p_T, detector $\eta(|\eta_{det}|)$, and $\Delta\phi(\not\!\!E_T, \mu)$.

f_j is determined by studying the ratio of Loose to Tight MC-subtracted data events, with $5 < \not\!\!E_T < 15$ GeV cut and without the triangle cut:

$$f_{MJ} = \frac{N_T - MC_T}{N_L - MC_L} \tag{6.7}$$

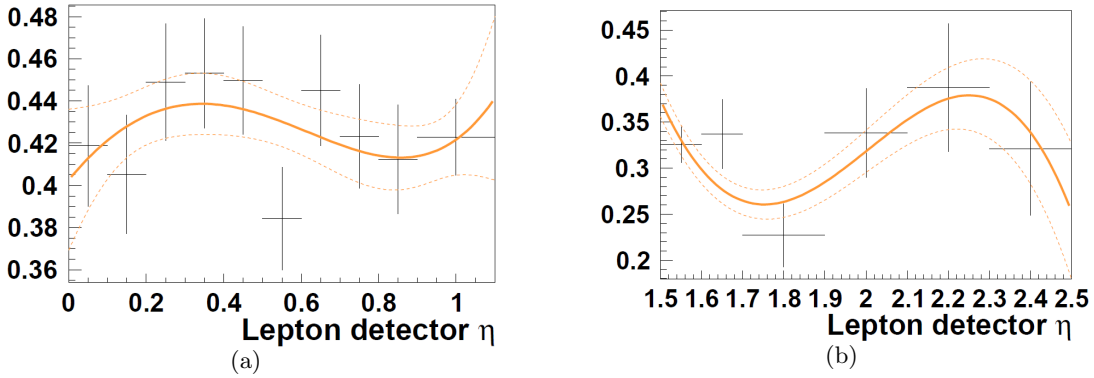
where $N_L(N_T)$ is the number of data events in the multijet estimation sample with a reconstructed Loose (Tight) lepton and $MC_L(MC_T)$ is the total MC estimation of the number of events with real Loose (Tight) leptons in them.

An example f_j^e is shown in Figure 6.3, and an example of f_j^μ is shown in Figure 6.4.



(a) (b)

Figure 6.3: Example of the probability for a jet that passes the loose lepton identification requirements to pass the tight lepton isolation requirements for the $15 < p_T < 17$ GeV range with $0.4 < min[\Delta\phi(\not\!\!E_T, jet)] < \pi$ as a function of the electron η in the (a) CC and (b) EC. The solid line is a fit to the data, and the dashed lines are the functions with the parameters shifted up and down by their uncertainties.



(a) (b)

Figure 6.4: Example of the probability for a jet that passes the loose lepton identification requirements to pass the tight lepton isolation requirements as a function of muon p_T for (a) $|\eta| < 1.0$ and (b) $1.0 < |\eta| < 1.6$, with $0 < \Delta\phi(\not\!\!E_T, \mu) < 1$. The solid line is a fit to the data, and the dashed lines are the functions with the parameters shifted up and down by their uncertainties.

6.3 Multijet and V+jet background normalization

Due to the large uncertainty in the W/Z+jets cross sections, we apply an additional scaling factor to both W+jets and Z+jets events based on a fit to data after subtracting the contributions from other known SM processes. The scaling factors are obtained after application of the triangle cut described in the next section and then remain fixed during the rest of the analysis. For both lepton channels the unbinned modified Matrix Method [82] using the fake rates described above is used build a template distribution of the W boson transverse mass (m_T^W) for the multijet events.

The MC(i.e $t\bar{t}$, single top, VV) subtracted data m_T^W(lepton p_T plus $\not\!\!E_T$) distribution is then fit with the multijet template and the V+jets ($V = W, Z$) MC samples in the electron and muon channels separately in order to extract their respective absolute normalization. The fit is performed with the ROOT TFractionFitter in order to properly account for the finite MC statistics. Separate normalization factors are determined for both the loose and tight lepton ID operating points for each jet multiplicity at the pre-tag level. We apply the same factor to all W+jets and Z+jets samples (they are treated as one shape in the fit).

The K^{exp} is defined as the factor by which TFractionFitter scales the V+jets MC in the fit. The values of the K^{exp} for each data set are summarized in Table 6.1.

Channel	II a	II b1	II b2	II b34
2jet				
Electron	1.00 ± 0.02	0.97 ± 0.02	0.98 ± 0.01	0.98 ± 0.01
Muon	1.11 ± 0.02	1.07 ± 0.03	1.11 ± 0.02	1.06 ± 0.02
3jet				
Electron	1.09 ± 0.06	1.13 ± 0.06	1.17 ± 0.04	1.11 ± 0.03
Muon	1.40 ± 0.06	1.29 ± 0.07	1.22 ± 0.06	1.29 ± 0.05

Table 6.1: The experimental K^{exp} factors for each lepton sample in different jet bins. Errors are statistical only. The total uncorrelated systematic uncertainty between the e and μ determinations is approximately 7-8%, based on trigger (3-4% e, 5% μ) and lepton ID (3% e, 2% μ) uncertainties.

6.4 b-tagging

The WH analysis searches for Higgs decays to $b\bar{b}$, so we employ b-tagging to categorize our candidate events based on the likelihood that the event contains either one or two b-tagged jets, using the MVA_{bl} b-tagging algorithm [59] described in Chapter 4. And the MVA_{bl} b-jet identification (bid) output is shown in Figure 6.5.

6.4.1 b-tagging optimization

Prior to Moriond 2012, we used the old criteria for classifying the events to different b-tagging categories, i.e. one tight, two loose, or two tight b-tagged jets. [64] Similarly to the CDF methodology at that time, we tried to split the two tagged b-tag categories, i.e. split the two loose into two loose$_{one}$, two loose$_{two}$, split the two tight into two tigh$_{one}$, two tight$_{two}$. To be specific, the previous 2 tag cuts are:

- $(MVA_{bl}^{jet1} + MVA_{bl}^{jet2})/2 > 0.3$ (VeryTight)

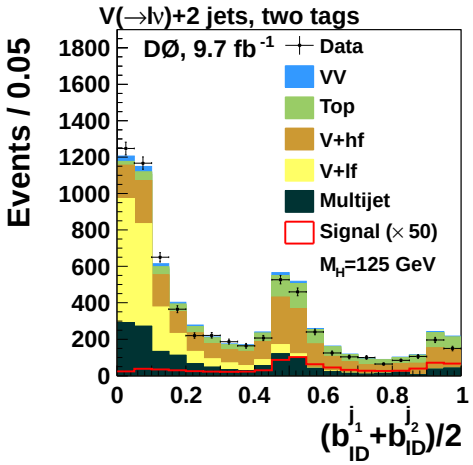


Figure 6.5: Average of the b-identification discriminant outputs of each jet in events with two jets. Discrepancies in data-MC agreement are within the systematic uncertainties described in Chapter 8.

- $0.02 < (MVA_{bl}^{jet1} + MVA_{bl}^{jet2})/2 < 0.3$ (L6)

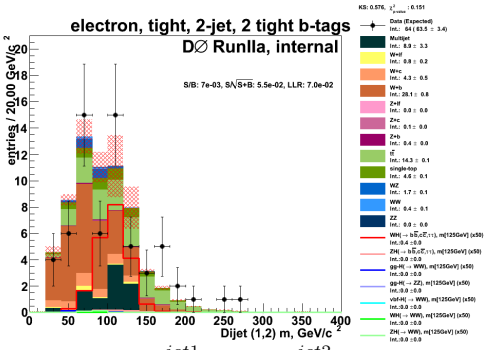
We tried to split those two categories into:

- $(MVA_{bl}^{jet1} + MVA_{bl}^{jet2})/2 > 0.5$ (MegaTight),
- $0.3 < (MVA_{bl}^{jet1} + MVA_{bl}^{jet2})/2 < 0.5$ (VeryTight to MegaTight)
- $0.15 < (MVA_{bl}^{jet1} + MVA_{bl}^{jet2})/2 < 0.3$ (Medium to VeryTight),
- $0.02 < (MVA_{bl}^{jet1} + MVA_{bl}^{jet2})/2 < 0.15$ (L6 to Medium)

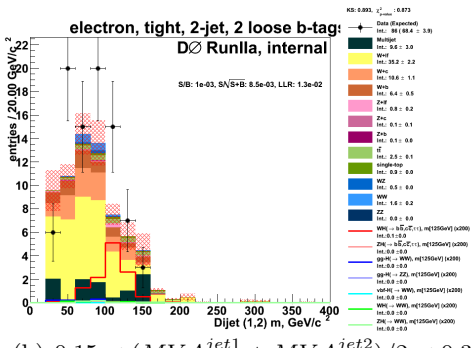
After applying such trial b-tagging, the dijet mass distributions are shown for example in Figure 6.6. $(MVA_{bl}^{jet1} + MVA_{bl}^{jet2})/2 > 0.5$ is the "tight-"tight category, and $0.3 < (MVA_{bl}^{jet1} + MVA_{bl}^{jet2})/2 < 0.5$ is the "loose-"tight category, these two are splitted from the default tight category $(MVA_{bl}^{jet1} + MVA_{bl}^{jet2})/2 > 0.3$. $0.15 < (MVA_{bl}^{jet1} + MVA_{bl}^{jet2})/2 < 0.3$ is the "tight-"loose category, $0.02 < (MVA_{bl}^{jet1} + MVA_{bl}^{jet2})/2 < 0.15$ is the "loose-"loose category, these two are splitted from the default loose category. From here we can see that the signals are more prominent in the splitted tight categories, whereas the improvement is not that significant in the splitted loose categories. Other distributions show similar behaviours. A simple limit calculation (without considering systematic, or other channels) also shows a similar conclusion, i.e. the trial b-tagging method gets a total 3.5% improvement: the splitted tight gets a 2.6% improvement, and the splitted loose gets a 20% improvement.

From these first optimization tests, we see that split the tight tag category would improve our analysis, so we want to split the tight region more properly. Considering Figure 6.5, there is a peak around 0.5, and the $V+lf$ background dominants the loose (low bID) region, and the $V+hf$ dominants the tight (high bID) region. After optimization studies, we set the cuts as:

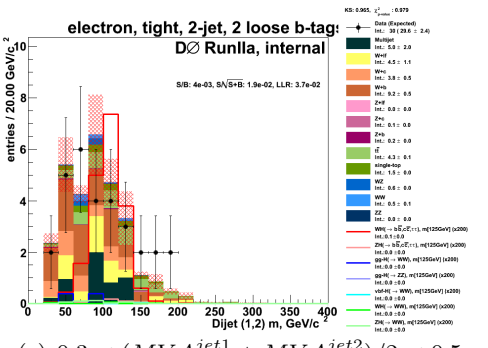
- $(MVA_{bl}^{jet1} + MVA_{bl}^{jet2})/2 > 0.55 \sim$ two tight



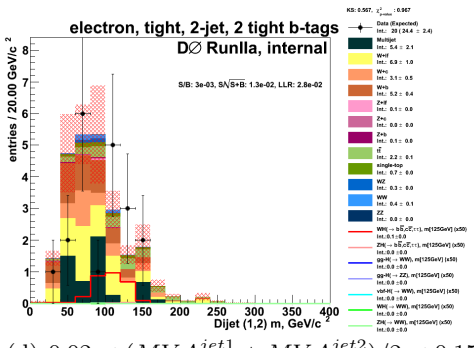
(a) $(MVA_{bl}^{jet1} + MVA_{bl}^{jet2})/2 > 0.5$



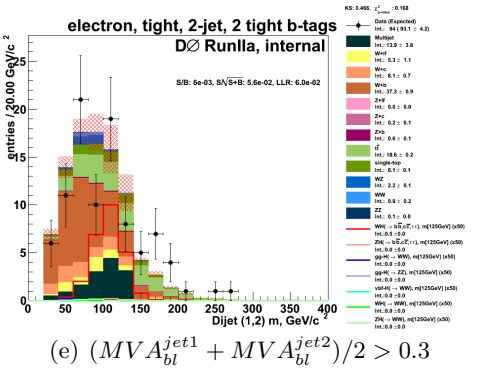
(b) $0.15 < (MVA_{bl}^{jet1} + MVA_{bl}^{jet2})/2 < 0.3$



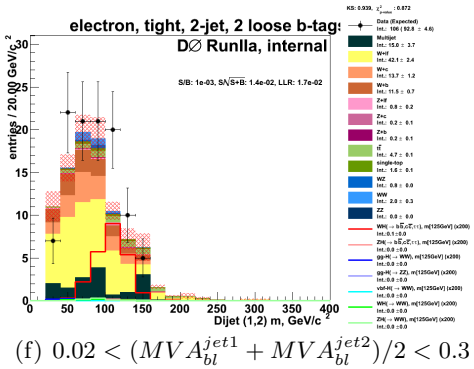
(c) $0.3 < (MVA_{bl}^{jet1} + MVA_{bl}^{jet2})/2 < 0.5$



(d) $0.02 < (MVA_{bl}^{jet1} + MVA_{bl}^{jet2})/2 < 0.15$



(e) $(MVA_{bl}^{jet1} + MVA_{bl}^{jet2})/2 > 0.3$



(f) $0.02 < (MVA_{bl}^{jet1} + MVA_{bl}^{jet2})/2 < 0.3$

Figure 6.6: Dijet mass distribution of different b-tagging category. The left three plots are "tight-"tight, "loose-"tight, and default tight categories, the right three plots are "tight-"loose, "loose-"loose, and default loose categories. The scale factors for the WH signal in each plots are respectively: 6.6(a): $\times 50$, 6.6(b): $\times 50$, 6.6(c): $\times 200$, 6.6(d): $\times 200$, 6.6(e): $\times 50$, 6.6(f): $\times 200$.

- $0.35 < (MVA_{bl}^{jet1} + MVA_{bl}^{jet2})/2 < 0.55 \sim$ two medium
- $0.02 < (MVA_{bl}^{jet1} + MVA_{bl}^{jet2})/2 < 0.35 \sim$ two loose

and we finally define four independent tagging samples with one tight, two loose, two medium, or two tight b-tagged jets. This analysis also considers an inclusive "pretag" sample, which contains all events that pass our selection requirements discussed in Section 5.3 before being classified into the different b-tagging categories. If exactly one jet is b-tagged, and the b-identification discriminant output for that jet, $b_{ID}^{j_i}$, satisfies the tight selection threshold ($b_{ID}^{j_i} > 0.15$), that event is considered part of the one tight b-tag sample. Events with two or more b-tagged jets are assigned to either the two loose b-tags, two medium b-tags, or two tight b-tags category, depending on the value of the average b-identification discriminant of the two jets with the highest discriminant values. And we choose the categories such that they each contain a different relative contribution of backgrounds, and a different ratio of signal to background events. In Figure 6.5, we see that below a bid output of 0.35, the background is dominated by V+light jet ($V + lf$) and multijet events with a small contribution from diboson (VV) events, and above 0.55, the dominant contribution is from V+heavy jet ($V + hf$) and top events. We also see that the signal is peaking near 1 and also near 0.5. I also tested the difference using $(b_{ID}^{j_1} + b_{ID}^{j_2})/2$ or $(b_{ID}^{j_1} \&\& b_{ID}^{j_2})$ as the cut criteria, where $(b_{ID}^{j_1} + b_{ID}^{j_2})/2$ means the average bid MVA output for the two jets passing the cut, and $(b_{ID}^{j_1} \&\& b_{ID}^{j_2})$ means both the bid MVA outputs for the two jets passing the cut. Finally we defined the three categories based on the average bid MVA output for the two jets: the double tight category is required to satisfy $(b_{ID}^{j_1} + b_{ID}^{j_2})/2 > 0.55$; the medium category is $0.35 < (b_{ID}^{j_1} + b_{ID}^{j_2})/2 \leq 0.55$; and the loose category is $0.02 < (b_{ID}^{j_1} + b_{ID}^{j_2})/2 \leq 0.35$.

If an event satisfies multiple tagging criteria, we place the event in the most stringent tag category. For example, events that pass the two tight b-tag requirements will also pass the two loose b-tag requirements, but are only counted once, in the two tight b-tag category.

N-Jets	1 tight tag / 2 loose/medium/tight tags
2	$WH \to \ell\nu b\bar{b}$
3	$WH \to \ell\nu b\bar{b}$

Table 6.2: Categories based on b-tagging and jet multiplicity.

6.4.2 b-tagging MC Corrections

To correct the tagging efficiency of the simulated events, data v.s. MC scale factors are applied (provided by the b-ID group) which depend on the kinematics of the jets.

In order to compute the total event weight, we apply the individual jet weight as follows:

$$CT_{weight}^{event} = \prod_j SF_j^{tagga} \cdot CT_j \tag{6.8}$$

where SF_j^{tagga} and CT_j are respectively taggability and continuous tag scale factors of the $j-th$ jet. The continuous tag scale factor is obtained from the data and Monte-Carlo tag rate functions (TRF) as follows: Consider a MC jet with a bid output value that lies between the $i-th$ and $i+1-th$ operating points. The fraction of all jets that fall in operating point i is: $f(i) = TRF_i - TRF_{i+1}$. To correct for the difference in $f(i)$ between data and MC, we

apply a weight: $w = \frac{f^{data}(i)}{f^{MC}(i)}$. If the bid output value is less than the loosest tag definition, we take $TRF_i^{data} = TRF_i^{MC} = SF = 1$. If the value is greater than the the tightest tag definition we take $TRF_{i+1}^{data} = TRF_{i+1}^{MC} = 0$. and $SF = 1$. Since $SF \cdot TRF^{MC} = TRF^{data}$, the continuous weight is computed as:

$$CT_j^i = \frac{TRF_{i+1}^{data} - TRF_i^{data}}{TRF_{i+1}^{MC} - TRF_i^{MC}} = SF_i SF_{i+1} \frac{TRF_{i+1}^{data} - TRF_i^{data}}{SF_i TRF_{i+1}^{data} - SF_{i+1} TRF_i^{data}} \tag{6.9}$$

where the jet j satisfies the b-id operating point i but fails the $i+1$ one. The systematic error on b-tagging is evaluated separately on the light flavour jet scale factor and the heavy flavour jet by varying the b-tagging scale factors by $\pm 1\sigma$. On average, the systematic uncertainty on bID is 5% per b-tagged jet.

6.4.3 b-tagged Event Distributions

Figures in [72] show several kinematic distributions of the $W + 2$ jet events in the electron and muon channel, respectively, for each of the 1 tight, 2 loose, 2 medium, and 2 tight tag samples. And I give examples in Figures 7.3 to 7.10. Figures 6.7 and 6.8 show some selected distributions of various kinematic at the pretag level.

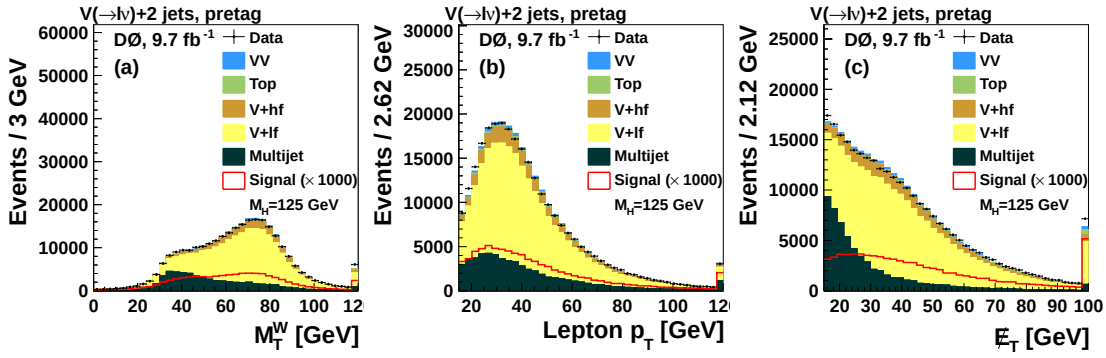
We see that our expected background events describe the data well. The observed agreement in all cases indicates that the simulation, which includes all the different Standard Model processes, describes the data well. The same plots for the $W + 3$ and jet channels are shown in [72].

The reconstructed dijet mass is shown in Figure 6.9 for the four different tag categories. As we require stricter b-tag requirements, we can see that the ratio of the number of signal events to the number of background events gets larger, and also that the relative background contributions changes. For example the two tight b-tag category is dominated by top and V+heavy jet production, whereas the two loose b-tag category has a larger contribution from V+light jet production.

6.5 Event Selection Result

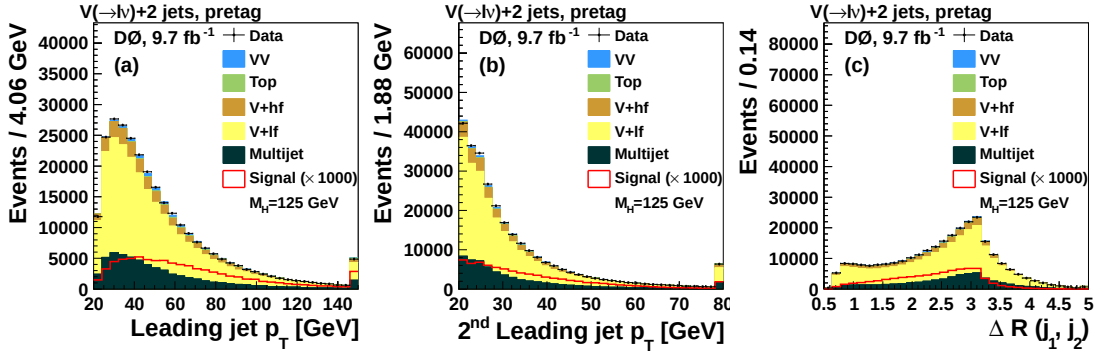
After passing all the selection criteria, treating the background, and categorizing based on the b-tagging, we will show the event selection result here. As described in Chapter 3, events are required to have one isolated charged lepton, large $\not\!\!E_T$, and two jets originating from a pair of b quarks.

The expected number of events from each signal and background category is compared to the observed data for each b-jet identification category for events with two jets, three jets in Tables 6.3 and 6.4 respectively. Selected kinematic distributions are shown for all selected events in Figures 6.7 and 6.8, and the dijet invariant mass for events with two jets is shown for four b-tag categories in Figures 6.9. In all plots, data points are shown with error bars that reflect the statistical uncertainty only. Discrepancies in data-MC agreement are within our systematic uncertainties described in Chapter 8.



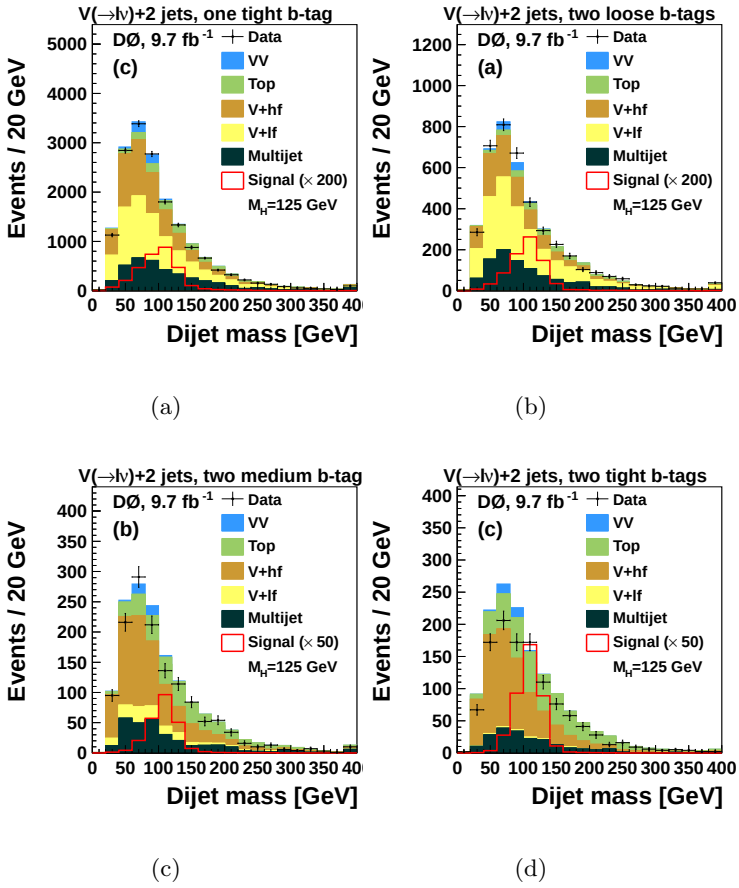
(a) (b) (c)

Figure 6.7: Distributions for all selected events with two jets of (a) transverse mass of the lepton-$\not\!\!E_T$system, (b) charged lepton pT, and (c) $\not\!\!E_T$. The signal is multiplied by 1000. Overflow events are added to the last bin.



(a) (b) (c)

Figure 6.8: Distributions for all selected events with two jets of (a) leading jet p_T, (b) second-leading jet p_T, and (c) ΔR between the leading and second-leading jets. The signal is multiplied by 1000. Overflow events are added to the last bin.



(a) (b)

(c) (d)

Figure 6.9: Invariant mass of the leading and second-leading jets in events with two jets and (a) one tight b tag, (b) two loose b tags, (c) two medium b tags, and (d) two tight b tags. The signal is multiplied by 200, 50, and 50, respectively. Overflow events are added to the last bin.

Table 6.3: Observed number of events in data and expected number of events from each signal and background source (where $V = W, Z$) for events with exactly two jets. The expected signal is quoted at $M_H = 125$ GeV. The total background uncertainty includes all sources of systematic uncertainty added in quadrature.

	Pretag	1 tight b-tag	2 loose b-tags	2 med. b-tags	2 tight b-tags
$VH \to \ell\nu bb$	37.3	11.6	3.2	4.6	7.7
$H \to VV \to \ell\nu jj$	24.7	1.8	0.3	0.07	0
$VH \to VVV \to \ell\nu jjjj$	13.0	1.2	0.3	0.04	0.01
Diboson	5686	535	109	42	38
$V + (g, u, d, s)$-jets	182 271	6174	1762	132	13
$V + (b\bar{b}/c\bar{c})$	27 443	5236	978	691	691
top ($t\bar{t}$ + single top)	3528	1289	247	333	462
Multijet	58 002	3700	946	298	195
Total expectation	276 930	16 935	4043	1496	1400
Total uncertainty	±14 998	± 1696	± 362	± 117	± 175
Observed events	276 929	16 406	4057	1358	1165

Table 6.4: Observed number of events in data and expected number of events from each signal and background source (where $V = W, Z$) for events with exactly three jets. The expected signal is quoted at $M_H = 125$ GeV. The total background uncertainty includes all sources of systematic uncertainty added in quadrature.

	Pretag	1 tight b-tag	2 loose b-tags	2 med. b-tags	2 tight b-tags
$VH \to \ell\nu b\bar{b}$	8.6	2.4	0.9	1.1	1.7
$H \to VV \to \ell\nu jj$	8.8	0.8	0.3	0.07	0.01
$VH \to VVV \to \ell\nu jjjj$	7.3	0.9	0.3	0.05	0.01
Diboson	1138	113	42	14	10
$V + (g, u, d, s)$-jets	24 086	976	582	34	3
$V + (b\bar{b}/c\bar{c})$	6625	1250	411	228	164
top ($t\bar{t}$ + single top)	3695	1123	365	460	570
Multijet	10 364	933	367	130	82
Total expectation	45 908	4395	1768	867	830
Total uncertainty	± 2582	± 528	± 209	± 118	± 113
Observed events	45 907	4278	1815	879	797

Chapter 7

Multivariate Signal Discriminants and Validation Through Diboson

Multivariate classifiers are commonly applied to maximize statistical sensitivities in high energy physics analysis. In this analysis we employ multivariate analysis (MVA) techniques to separate signal from background events. To separate signal from the MJ events, we use a boosted decision tree implemented with the TMVA package [83]. This multivariate analysis is described in Section 7.1. For the final multivariate analysis, we use a BDT in the one tight b-tag channel and all three two b-tag channels, and we use a random forest decision tree (RF) [84] implemented in the StatPatternRecognition (SPR) package [85, 86] for events in the zero and one loose b-tag channels.

7.1 Boosted Decision Trees

A decision tree is a binary tree structured classifier similar to the one sketched in Figure 7.1. The decision tree repeats a series of yes/no splits on classifiers to separate signal from background, until either a minimum number of events in a node is reached, or a node contains events that are pure signal or background.

Boosting is a way to enhance the MVA classification performance and stability. Boosting builds up a series of trees, where each tree is retrained, boosting the weights for events that were misclassified in previous trainings

In this analysis, we use the stochastic gradient boosting from the TMVA package [83], which combines a bagging and a boosting algorithm. The bagging aims at (re-)training individual trees with a random subset of the input variables and training events. The boosting technique boosts the events that are misclassified by assigning a weight to each intermediate tree. It is determined by minimizing a loss function $L(F, y)$(Equation 31 in TMVA user's guide [87]) [1] in a similar way as performed for neural networks training .

A more detailed description of these techniques can be found from the decision tree documentation of the TMVA package [83], we used the version 4.1.0.

We train a collection of decision trees built from replicas of the training data set. These replicas are made by sampling with replacement from the original data set, such that replicas

[1]The boosting procedure is now employed to adjust the parameters P such that the deviation between the model response $F(x)$ and the true value y obtained from the training sample is minimised. The deviation is measured by the so-called loss-function $L(F, y)$, a popular choice being squared error loss $L(F, y) = (F(x) - y)^2$.

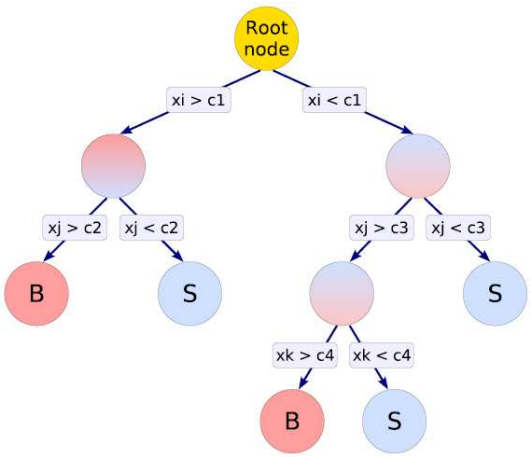


Figure 7.1: The decision tree uses classifiers to create a series of yes/no splits on events that are known to be classified as either signal or background. The splitting is done to maximally separate signal from background, and stops when either a node contains events that are pure signal or background, or when a minimum number of events in a node (leaf) is reached.

are the same size as the original but may contain duplicate events. This technique is commonly referred to as Bootstrap AGGregatING or "bagging". Discriminant variables are also sampled with replacement at each splitting. For a given test event the output of the forest is the combined average output of all trees. We use the bagged decision tree classifier (SprBaggerDecisionTree) from the package StatPatternRecognition in this analysis. A separate RF is trained for each mass point in each of the eight sub-channels. Details of the RFs are listed in Table 7.1.

Parameters used for Random Forest Training
Number of trees: 50
Number of input variables:
electron channel 2 jets excl (3 jets excl): 32 (50)
muon channel 2 jets excl (3 jets excl): 43 (50)
Criterion for decision tree optimization: gini index[2]
Minimal number of entries per tree leaf: 500
Max number of sampled features: 50

Table 7.1: Details of Random Forest construction.

7.1.1 Variable Selection

In general, the list of selected variables is built by applying the following steps:

1. starting from all available distributions, a good data / background agreement has to be observed. We require a Kolmogorov-Smirnov (KS) [88] probability threshold of 0.01 to build a list of "good" variables

[2]The gini index (GI) is defined as $GI = p(1-p)$ where the purity $p = s/\sqrt{s+b}$ where s and b correspond to the weighted sum of signal and background events

2. TMVA offers a ranking output which allows to measure the discriminating of each input variables after testing. In order to get the most simple structure of the decision trees, the less relevant variables and the ones which are similarly correlated to other ones for both signal and backgrounds events are removed if the performances of discriminant are not altered

It has been observed that using the list of inputs variables from the 2-jet channels for the 3-jet ones would not degrade the 3-jet channels' discriminating power. This is achieved by using "max/min"-like variables instead of individual ones, e.g the maximum $\Delta\eta$ between a jet and the lepton instead of the individual $\Delta\eta(jet_j, lepton)$ ones. Also, as described below, the Higgs mass is reconstructed from the di-jet invariant mass which drives most of the final sensitivity. It is formed with the pair of jets with the highest b-ID MVA BL outputs, it is thus the same variable used in both jet multiplicities. Table 7.2 list the MVA input variables for each jet multiplicity and tagging point.

The ranking of the BDT input variables is derived by counting how often the variables are used to split decision tree nodes, and by weighting each split occurrence by the separation gain-squared it has achieved and by the number of events in the node [89]. This measure of the variable ranking can be used for a single decision tree as well as for a forest [83]. Table 7.2 show the BDT variable rankings for the two tight b-tagged (2T), two medium b-tagged (2M), two loose b-tagged (2L), and one tight b-tagged (1T) channels.

The importance of variables used in the RF classifier are estimated using the validation sample. After training, validation events are run through the RF, once for each variable used. On each pass the class of each event is randomized whenever the variable under test is encountered and the change in the quadratic loss figure of merit is estimated. The class is defined to be 1/0 for signal/background. The quadratic loss is defined as:

$$F.O.M. = \frac{\sum_{i=1}^{numberofevents} wgt_i (event_i^{class} - RF(event_i))^2}{\sum_{i=1}^{numberofevents} wgt_i} \tag{7.1}$$

7.1.2 Training and Optimization

In order to optimize our sensitivity, the analysis samples have been split and subsequently trained separately for events with exactly 2 and 3 jets and for each tagging categories.

Along with tuning the individual Decision Tree parameters (e.g. number of trees, tree depths,. . .) several approaches have been tested in order to optimize the training procedure. We present here the ones which gave the best sensitivity. The following parameters were set:

shrinkage: this parameter constrains the learning rate of the algorithm by applying a weight to the individual trees. It ranges from 0 to 1. A value of 0.1 has been found useful to avoid overtraining.

number of trees: sets of 20, 50, 100, 200, 400 and 800 trees have been tested. A number of 200 trees was found to show the best performances

number of nodes: no variations on the discriminating power has been observed by adjusting this parameter, thus the default TMVA value (100000) has been set.

node split criteria: the gini index (GI) has been used. It is defined as $GI = p(1-p)$ where the purity $p = s/\sqrt{s+b}$ where s and b correspond to the weighted sum of signal and background events

node purity limit: in boosting, nodes with purity < 0.5 are signal, background otherwise

node cut: number of steps during node cut optimization

The training procedure follows two steps:

1. two sets of decision trees have been trained for each Higgs mass hypothesis against respectively $t\bar{t}$ and $W + b\bar{b}$, $c\bar{c}$ which correspond to the dominant backgrounds in the double-tag channel.

2. a final set of decision trees (MVA FINAL) are then trained using as input variables the union of the most powerful ones selected from the two above ones. This final discriminant is then applied to all non-trained events passing the complete selection cut-flow.

The training has been done separately for each Higgs mass point against all backgrounds. One third of events were respectively used for the training and testing steps. The last third is used for the final measurement. In order to maximize the granularity of the output MVA, we merged all epochs. A single set of trees has been thus been applied for both Run IIb1 and Run IIb2+3+4 datasets.

Figures 7.3 to 7.10 show the normalized input variables for the 1 tight tag, 2 loose tag, 2 medium and 2 tight channels, and 2 jet or 3 jet, respectively, and 7.13, 7.14 show the output of the final discriminant for both trained and test events for the signal and backgrounds events.

7.2 Rebinning

The MVA output has no protection against producing bins with insufficient statistics to properly evaluate a limit. In order to ensure that each bin in the final discriminant used to set our limits can be well described by our background models, we produce an initial MVA output distribution with 100 bin, then rebin to 25 bins in such a way that each bin contains a statistically significant estimate of the data.

The current rebinning algorithm combines bins from the highest MVA output down until these conditions are satisfied:

$$\begin{aligned} \mu_B &> 0.01 \\ \mu_{S+B} &> 0.1 \\ \sigma_B/\mu_B &< 0.28 \\ \sigma_{S+B}/\mu_{S+B} &< 0.2 \end{aligned} \tag{7.2}$$

where $\mu_B(\mu_{S+B})$ is the total expected background (signal + background) yield and $\sigma_B(\sigma_{S+B})$ is the statistical uncertainty of that background (signal + background) expectation, calculated as $\sqrt{\sum w_i^2}$ over the weights, w_i , for contributing events. The value of σ_B/μ_B $(\sigma_{S+B}\mu_{S+B})$ in Equation 7.2 can be understood as representing the inverse of the statistical significance of the background (signal plus background) away from zero, in standard deviations. The choice of B and S+B conditions in Equation7.2 are made to ensure that a single high-statistics background sample does not dominate the choice of binning by giving a high statistical significance to a very low expected yield. Separate conditions on the background-only and signal-plus-background samples allows for the possibility of having a

high signal-to-background ratio while maintaining a reasonable statistical understanding of the background.

After determining in which bin of the 100-bin initial distribution the thresholds are satisfied (hereafter the "cutoff bin"), we determine in which of the 25 bins of the final distribution the cutoff bin resides (the "target bin".) The bin content from all bins in the initial distribution above the cutoff bin is moved into the target bin of the final distribution, along with the content that comes from a simple 100-to-25 bin rebinning, taking care to avoid any double-counting. Bins in the final distribution below the target bin are simply the result of a straightforward 100-to-25 bin rebinning. Thus the target bin is the last bin with content. Bins above the target bin, which by definition contain no signal, background, or data, are ignored in the limit setting.

7.3 Performance

7.3.1 Multivariate Multijet Discriminators

To optimize without biasing the sample on which we evaluate the limit, our events are split into three statistically independent subsamples for training, testing, and performing the final statistical analysis published in 2013 [7]. Another optimization has been to add an input variable to the final MVA, compared to our previous analysis presented in Moriond 2012.

The new added input is trained using a BDT to separate the MJ background from the signal events: $VH(\rightarrow b\bar{b}, c\bar{c}, \tau\tau)$. The variables used in training these BDTs are chosen to exploit kinematic differences between the MJ and signal events, and are documented in [72]. To improve the training statistics, we combine signal events for M_H = 120, 125, and 130 GeV in the training. We find that a BDT trained on this combination of Higgs boson masses has a similar performance when applied to other masses, eliminating the need for a mass-dependent MJ discriminant. The BDT outputs $\mathrm{MVA}_{MJ}(VH)$ is shown in Figure 7.2, and Figures 7.11, 7.12. The $\mathrm{MVA}_{MJ}(VH)$ discriminant output is used as input variable to the final MVAs, as detailed in [72]. This is done separately between electron and muon, and I was in charge of the electron channel. We can see from the $\mathrm{MVA}_{MJ}(VH)$ figures that the multijet background is indeed pushed to the low $\mathrm{MVA}_{MJ}(VH)$ discriminant part, and separated from the signal.

7.3.2 Final $WH \rightarrow \ell\nu b\bar{b}$ MVA Analysis

In events with two or three jets and one tight b tag or two b tags, the $WH \rightarrow \ell\nu b\bar{b}$ process provides the dominant signal contribution. To separate signal from background, we train a BDT on the $WH \rightarrow \ell\nu b\bar{b}$ signal and all backgrounds. The lists of input variables to the MVA are described here below and the importance of the different BDT's is given in the Table 7.2. Figures 7.3 through 7.10 show examples of the most effective discriminating variables used in our BDTs for the two-jet and three-jet channels, respectively, in the one-tight-b-tag and all two-b-tag channels. Figures 7.13 and 7.14 show the BDT output for the two- and three-jet channels, respectively, in the one tight b-tag and all the two-b-tag channels.

The multivariate discriminants used in this search use various input variables listed in Table 7.2. Individual input variables are described in detail below. In the descriptions, ℓ refers to the electron or muon in a selected event, ν refers to the neutrino candidate, and j_n

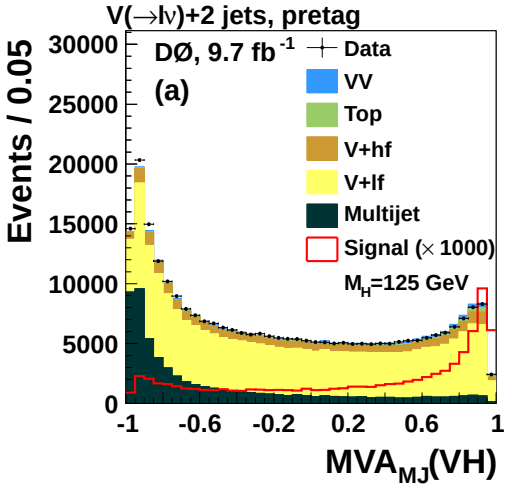


Figure 7.2: The multivariate discriminant output for $\mathrm{MVA}_{MJ}(VH)$, for all events. The signal is multiplied by 1000.

Table 7.2: Table of input variables for the final signal discriminant for the $WH \to \ell\nu b\bar{b}$ channel. Variables are listed by their rank of importance when used in the two tight b-tagged (2T), two medium b-tagged (2M), two loose b-tagged (2L), and one tight b-tagged (1T) categories. And the same inputs are used for the 3 jet categories.

Variable	2T	2M	2L	1T
$MVA_{MJ}(VH)$	1	1		
$m_{j_1 j_2}$	2	4	3	1
$p_T^W/(p_T^\ell + \not\!\!E_T)$	3	6	4	2
b_{ID}^{j12}	4	13	1	4
$cos(\chi^\star)$	5	3		
$max\lvert\Delta\eta(\ell, \{j_1 or j_2\})\rvert$	6	11	2	3
$q^\ell \times \eta^\ell$	7	2	6	6
$\Delta R(\ell, j_1)$	8	5		
$min[SIG(j_{12}, \{j_1 or j_2\})]$	9	15	9	5
$q^\ell \times \eta^{j_1}$	10	7	11	9
$\mathcal{V}(j_{12})$	11	12	7	11
$cos(\theta^\star)$	12	10		
$m_{\ell\nu j_2}$	13	16	12	13
m_T^{j12}	14	14		
$\mathcal{C}$	15	8	8	10
$\sum(p_T)^{VIS}$	16	9		
m^{Asym}			5	8
$\mathcal{A}$			10	12
$p_T^{j_2}$			13	7

refers to jets as ordered by p_T where j_1 is the jet with highest p_T. The p_z of the neutrino candidate is estimated by constraining the charged lepton and the neutrino system to the mass of the W boson and choosing the lowest magnitude solution. Figures 7.3 to 7.10 show all the input variables for each category (2 jet, 3 jet, one tight tag, two loose tag, two medium tag, two tight tag).

- MVA$_{MJ}(VH)$: Output of the multivariate discriminant trained to distinguish $WH \to \ell\nu b\bar{b}$ from the MJ background
- $m_{j_{12}}$: Invariant mass of the leading and second leading jets
- $p_T^W/(p_T^\ell + \not{E}_T)$: Ratio of the p_T^W to $p_T^\ell + \not{E}_T$
- $b_{ID}^{j_{12}}$: Averaged b-jet identification output for the highest energy b-tagged jets
- $\cos(\chi^\star)$: $\chi^\star = \angle(\ell, spin_W)$ in $\ell\nu$ system CM frame [90]
- $\max|\Delta\eta(\ell, \{j_1 or j_2\})|$: Maximum $\Delta\eta$ between the charged lepton and the leading or second leading jet
- $q^\ell \times \eta^\ell$: Product of the lepton charge and its pseudorapidity
- $\Delta R(\ell, j_1)$: ΔR between the charged lepton and the leading jet
- $\min[SIG(j_{12}, \{j_1 or j_2\})]$: Minimum SIG of the leading or second leading jet defined as $p_T^{min}(j_1, j_2)\Delta R(j_1, j_2)/\sum_{i=1}^{2} p_T^{j_i}$ with respect to the dijet system
- $q^\ell \times \eta^{j_1}$: Product of the the lepton charge and pseudorapidity of the leading jet
- $\mathcal{V}$: Velocity is defined for an object $C \to AB$ as $\ln\{1-\{1-4\times[(m_A^2+m_B^2)/m_C^2]^{1/2}\}^{1/2}\}$
- $\mathcal{V}(j_{12})$: Velocity of the dijet system
- $\cos(\theta^\star)$: $\theta^\star = \angle(W$, incoming u-type quark) in HCM frame [90]
- $m_{\ell\nu j_2}$: Invariant mass of the system consisting of the charged lepton, reconstructed neutrino (assuming $p_Z^{\ell 1}$, and second leading jet
- $m_T^{j_{12}}$: Transverse mass of the leading and second leading jets
- $\mathcal{C}$: Centrality is $(\sum_i p_T^i)/(\sum_i |\vec{p_i}|)$, where i runs over ℓ and all jets
- $\sum(p_T)^{VIS}$: Scalar sum of the p_T of the visible particles
- m^{Asym}: Mass asymmetry between $\ell\nu$ system and the dijet system: $(m_{\ell\nu}m_{j_{12}})/(m_{\ell\nu} + m_{j_{12}})$
- $\mathcal{A}$: Aplanarity is $3\lambda_3/2$ where λ_3 is the smallest eigenvalue of the normalized momentum tensor $S^{\alpha\beta} = (\sum_i p_i^\alpha p_i^\beta)/(\sum_i |\vec{p_i}|^2)$, where $\alpha, \beta = 1, 2, 3$ correspond to the x, y, z momentum components, and i runs over selected objects. Without arguments, it is calculated for all visible objects
- $p_T^{j_2}$: p_T of the second leading jet

We show the 10 variables, which are used for all the four b-tagging categories: $m_{j_1 j_2}$, $p_T^W/(p_T^\ell + \not{E}_T)$, $b_{ID}^{j_{12}}$, $max|\Delta\eta(\ell, \{j_1 or j_2\})|$, $q^\ell \times \eta^\ell$, $min[SIG(j_{12}, \{j_1 or j_2\})]$, $q^\ell \times \eta^{j_1}$, $\mathcal{V}(j_{12})$, $m_{\ell\nu j_2}$, $\mathcal{C}$. And we show m^{Asym}, $p_T^{j_2}$ for the 1 tight, 2 loose case; and $\Delta R(\ell, j_1)$, $m_T^{j_{12}}$ for the 2 medium, 2 tight case for their importance. We also show $MVA_{MJ}(VH)$ for all the four b-tagging categories in Figures 7.11 7.12, as it is the most significant variable for 2 medium and 2 tight, and the most important improvement when we updated the analysis after the Moriond 2012 results [91].

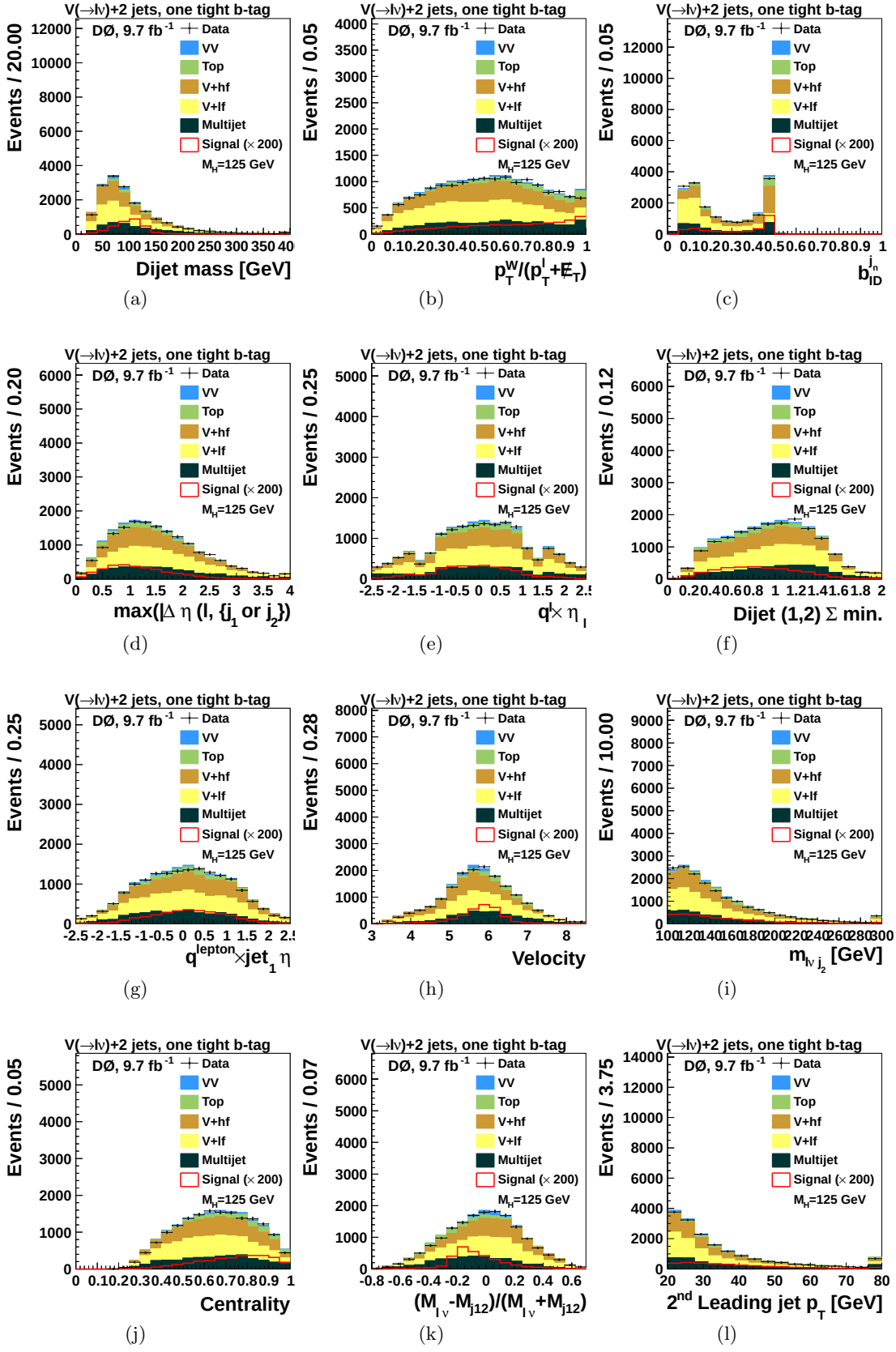


Figure 7.3: Distributions of some of the most significant inputs to the final discriminant in events with exactly two jets and one tight b tag.

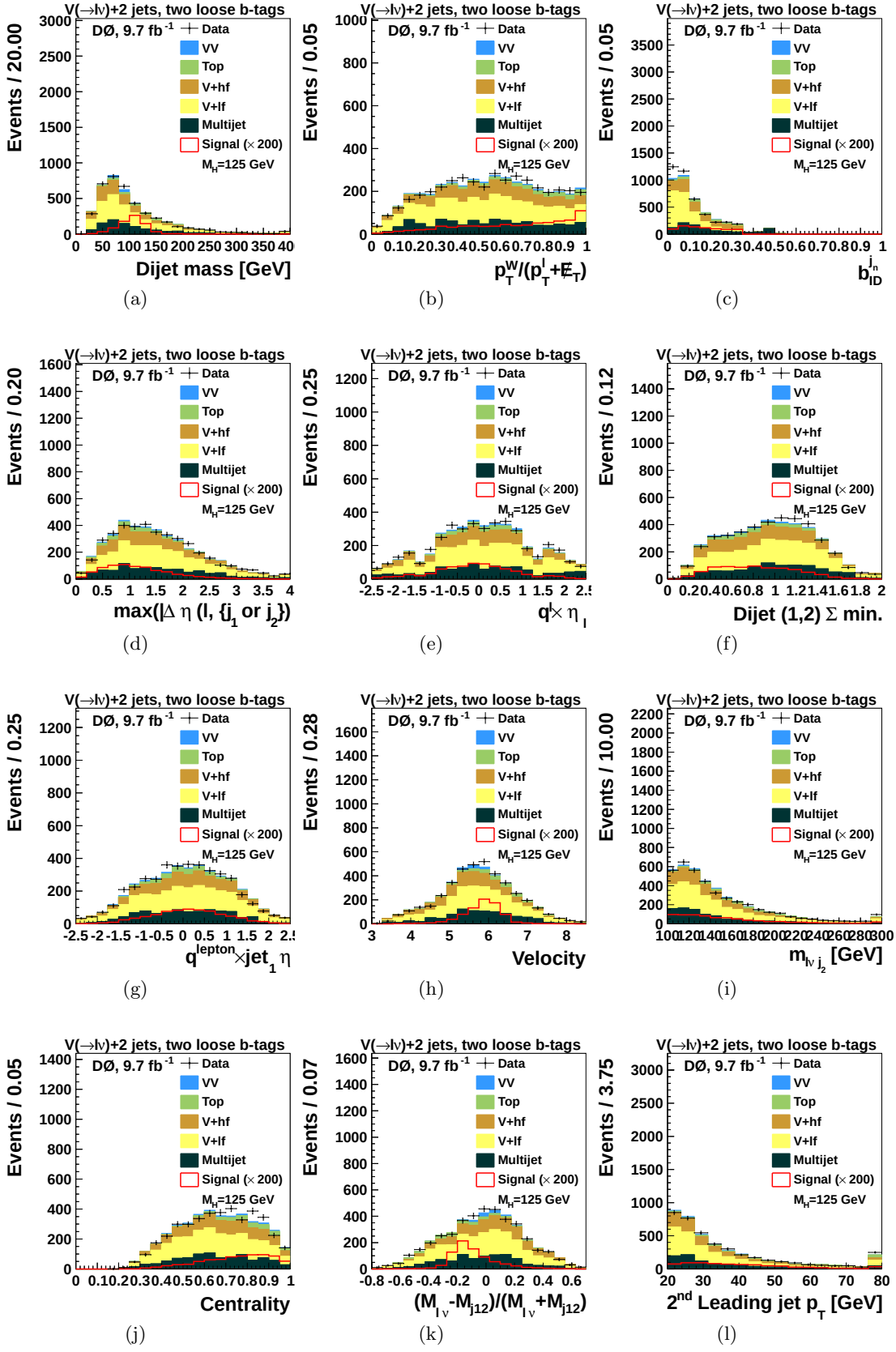


Figure 7.4: Distributions of some of the most significant inputs to the final discriminant in events with exactly two jets and two loose b tag.

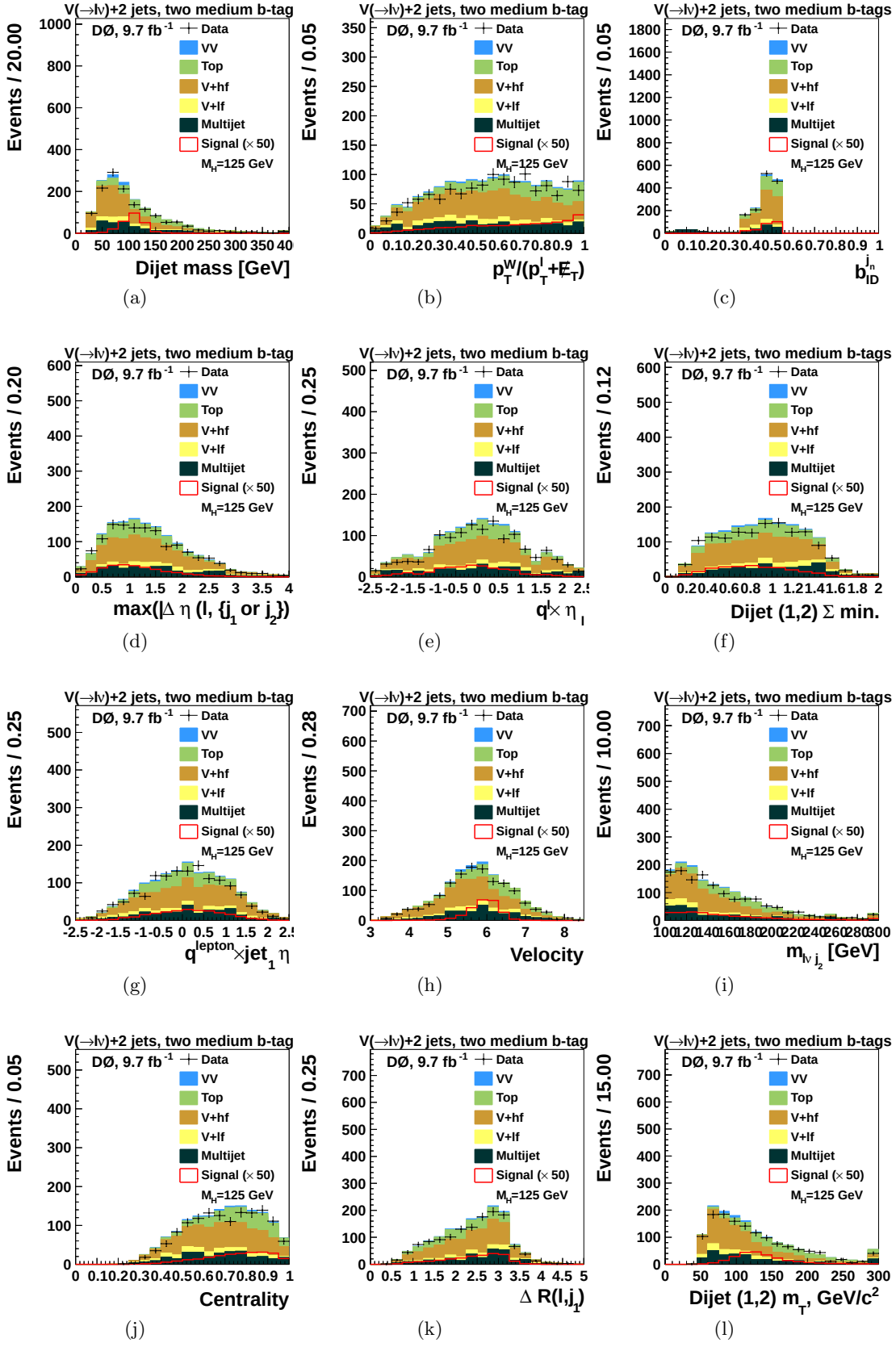


Figure 7.5: Distributions of some of the most significant inputs to the final discriminant in events with exactly two jets and two medium b tag.

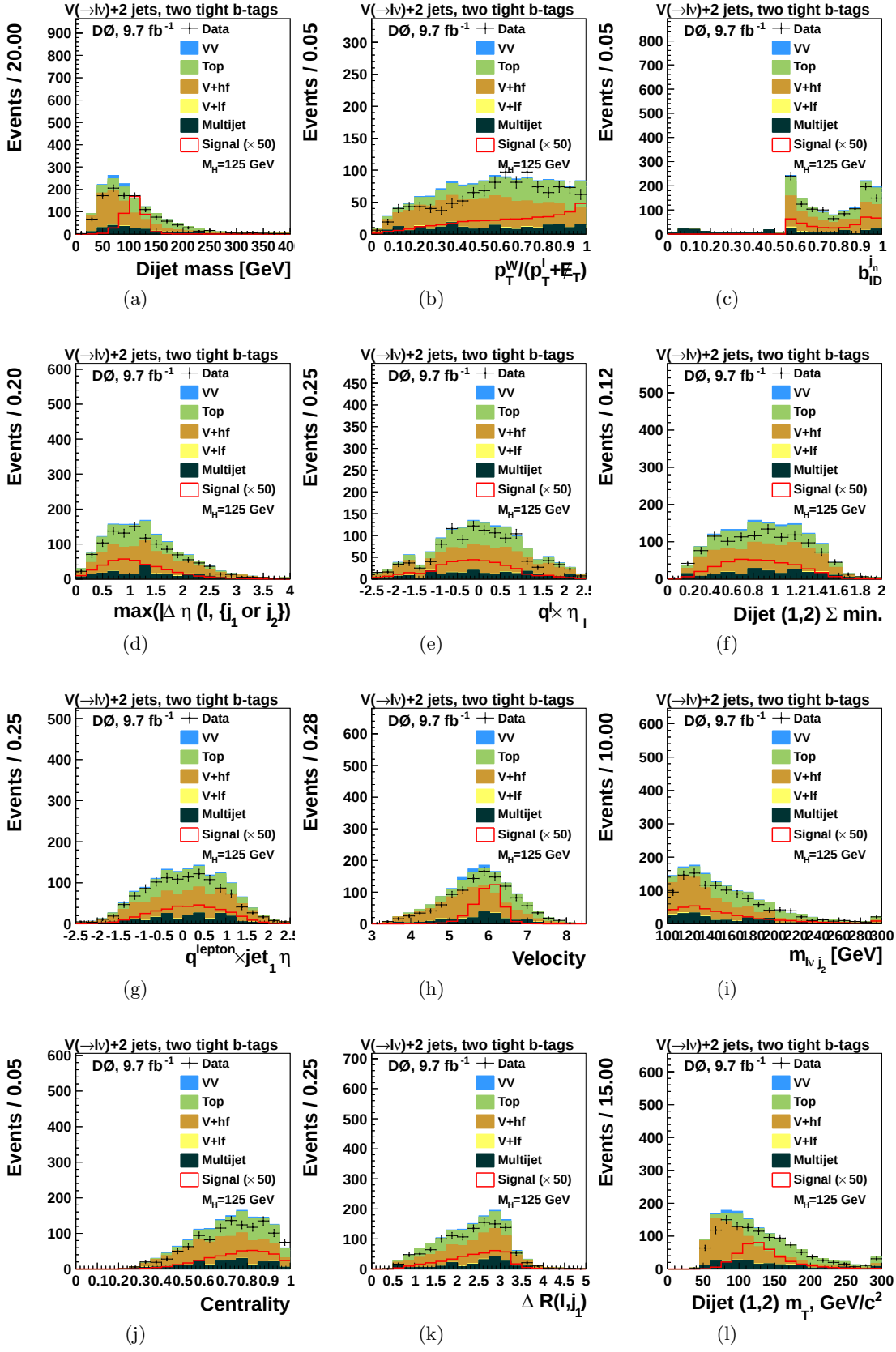


Figure 7.6: Distributions of some of the most significant inputs to the final discriminant in events with exactly two jets and two tight b tag.

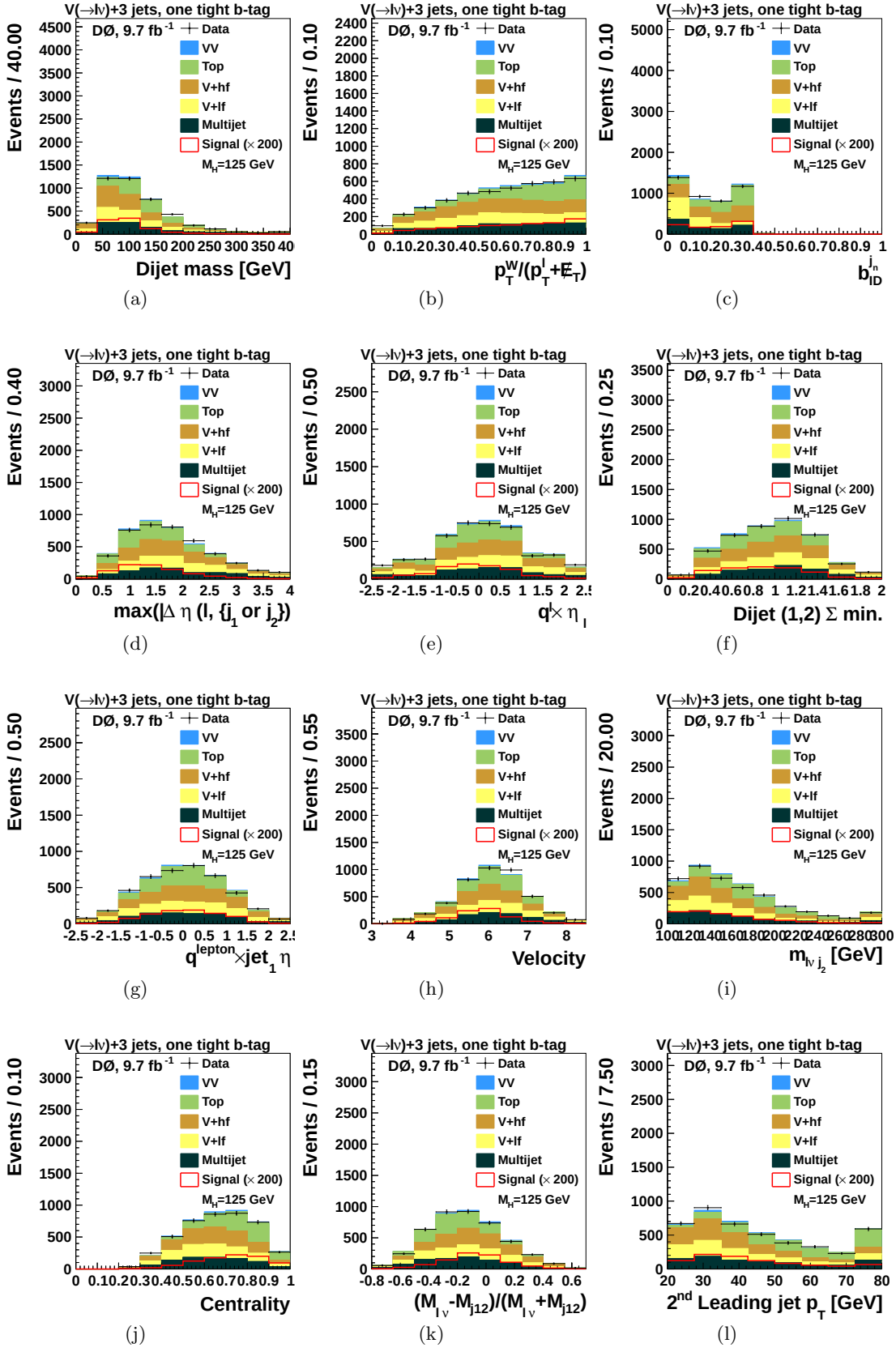


Figure 7.7: Distributions of some of the most significant inputs to the final discriminant in events with exactly three jets and one tight b tag.

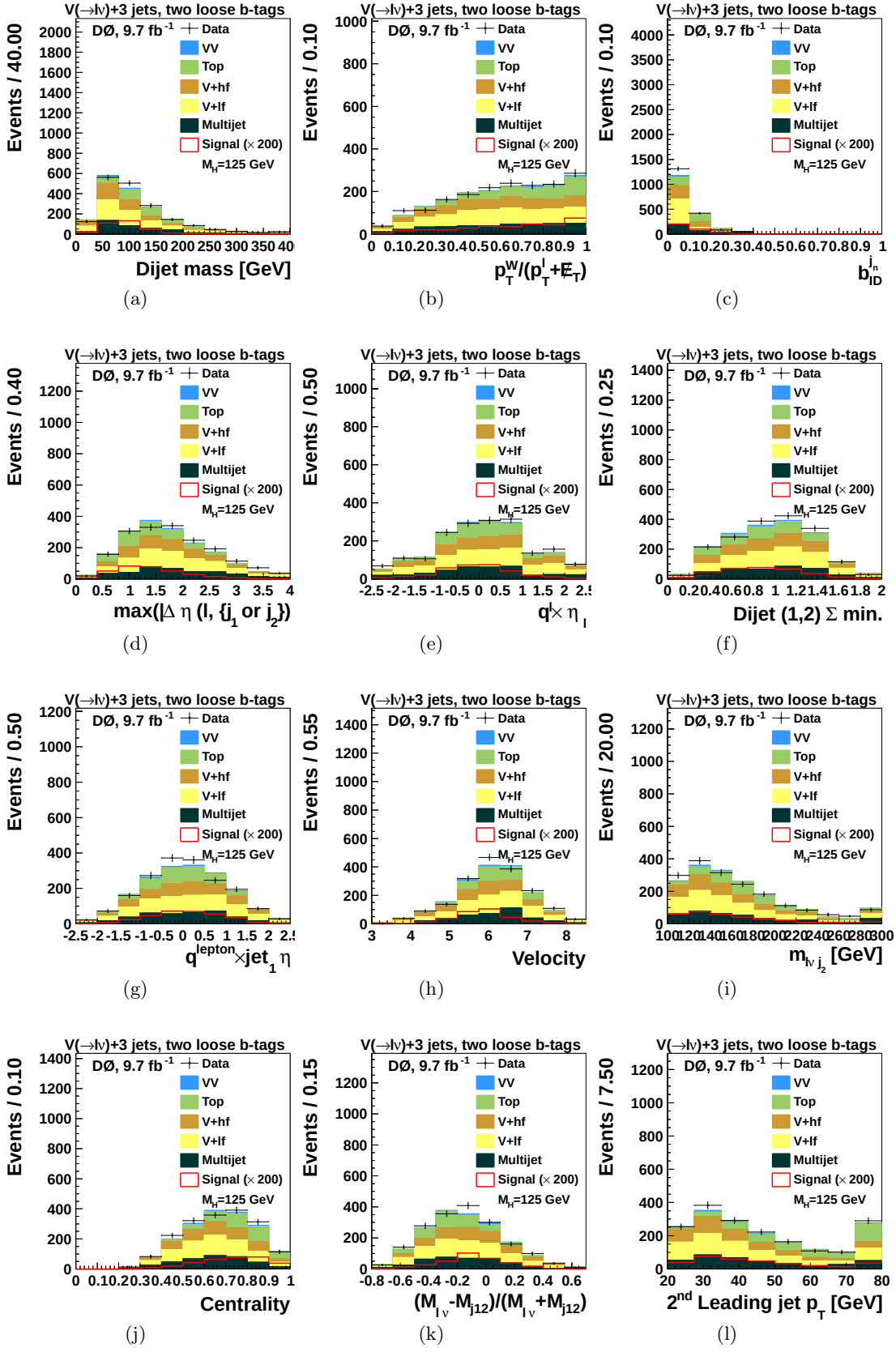


Figure 7.8: Distributions of some of the most significant inputs to the final discriminant in events with exactly three jets and two loose b tag.

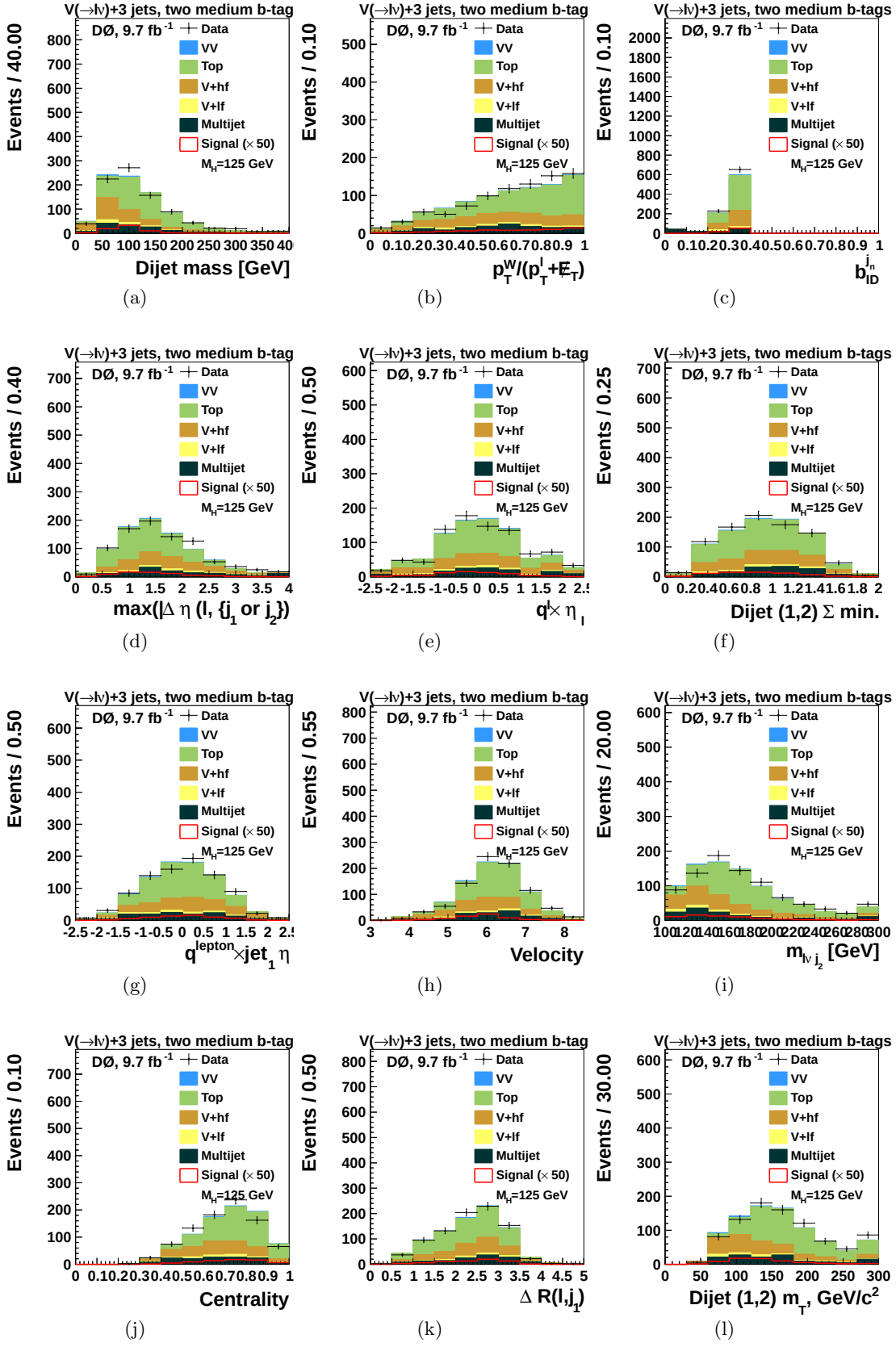


Figure 7.9: Distributions of some of the most significant inputs to the final discriminant in events with exactly three jets and two medium b tag.

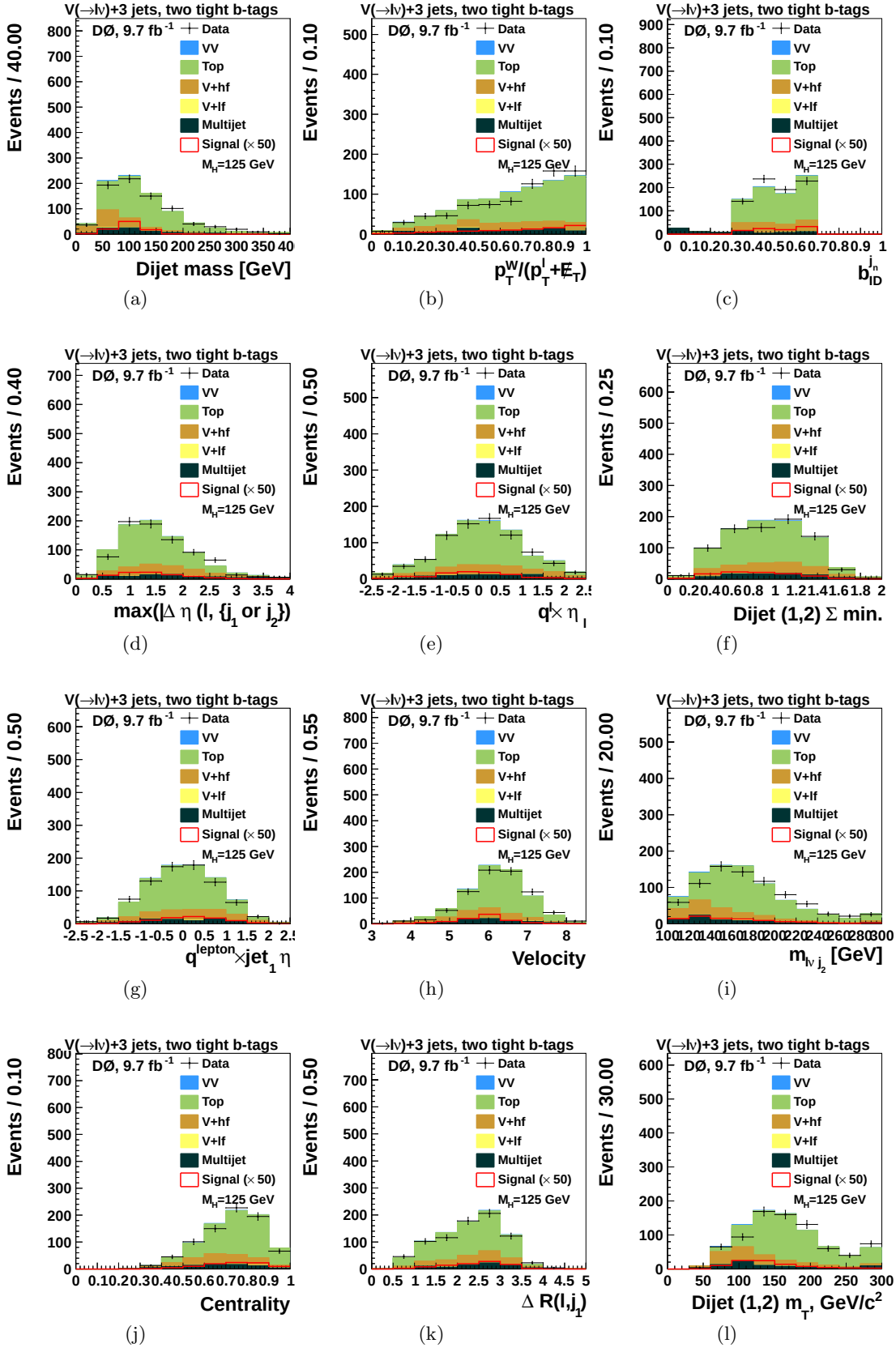


Figure 7.10: Distributions of some of the most significant inputs to the final discriminant in events with exactly three jets and two tight b tag.

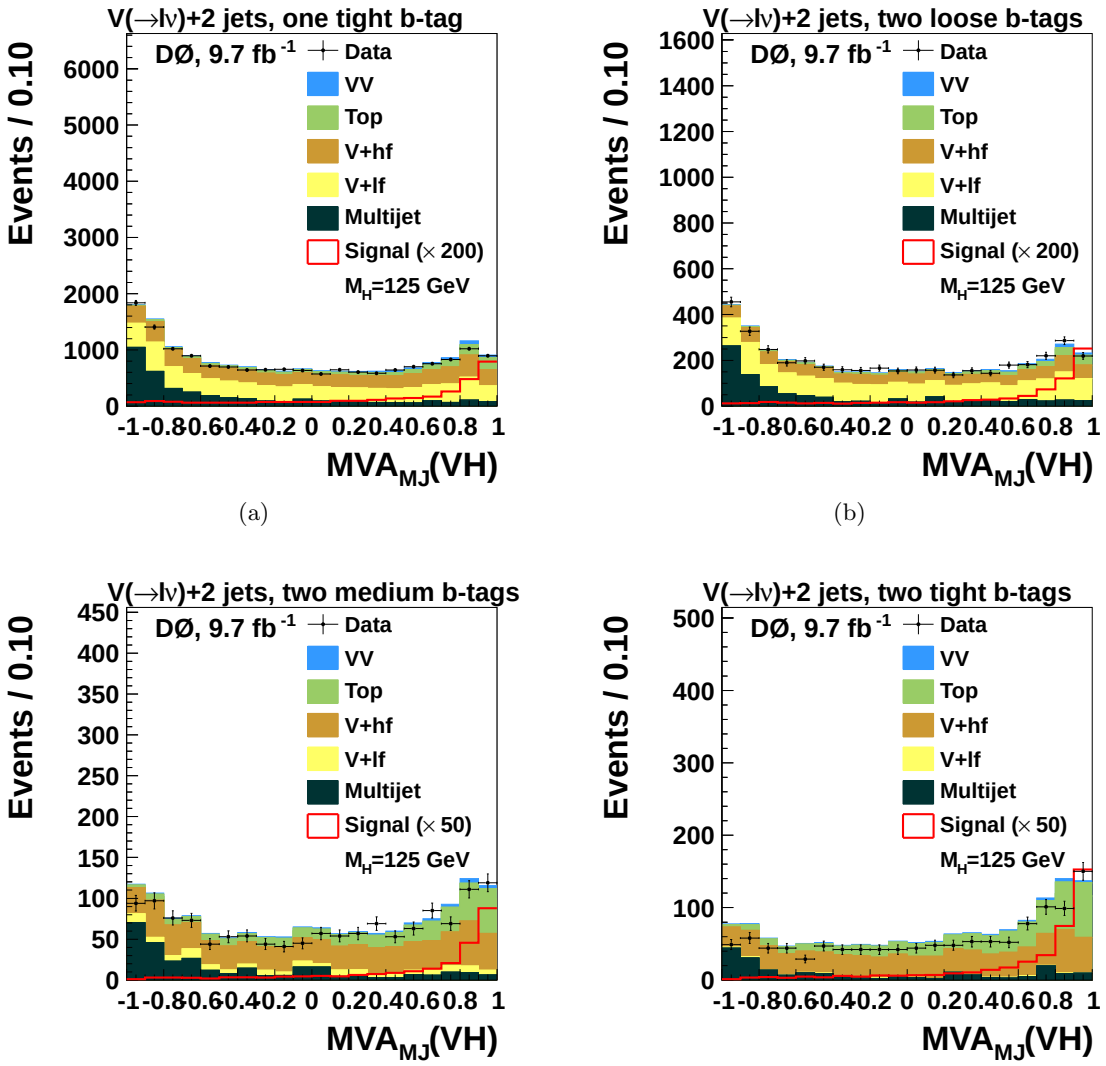


Figure 7.11: Distributions of some of the most significant inputs to the final discriminant in events with exactly two jets: (a) one loose (b) two tight (c)two medium (d)two tight. The signal is multiplied by 200, 200, 50, and 50, respectively. Overflow events are added to the last bin.

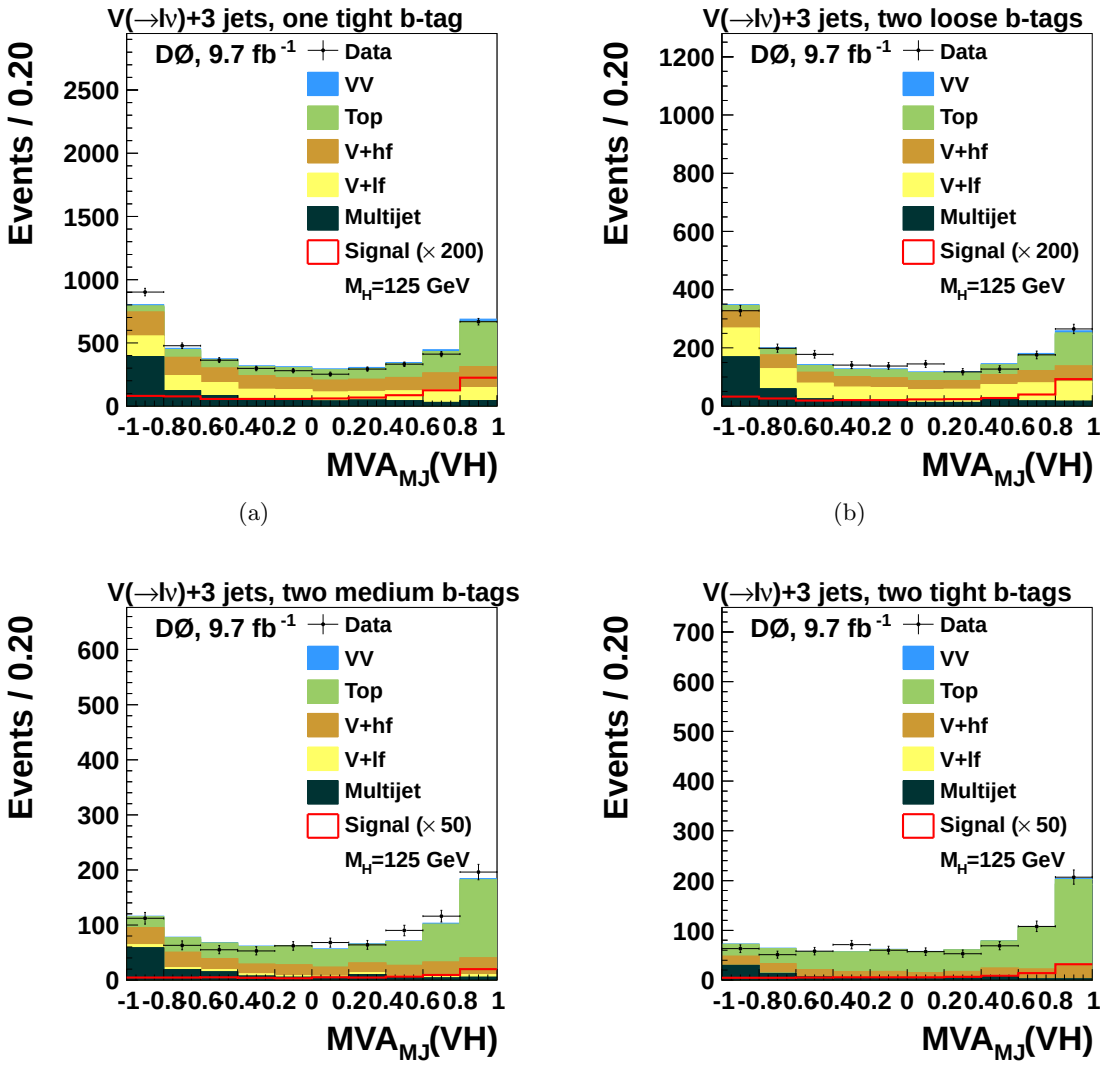


Figure 7.12: Distributions of some of the most significant inputs to the final discriminant in events with exactly three jets: (a) one loose (b) two tight (c)two medium (d)two tight. The signal is multiplied by 200, 200, 50, and 50, respectively. Overflow events are added to the last bin.

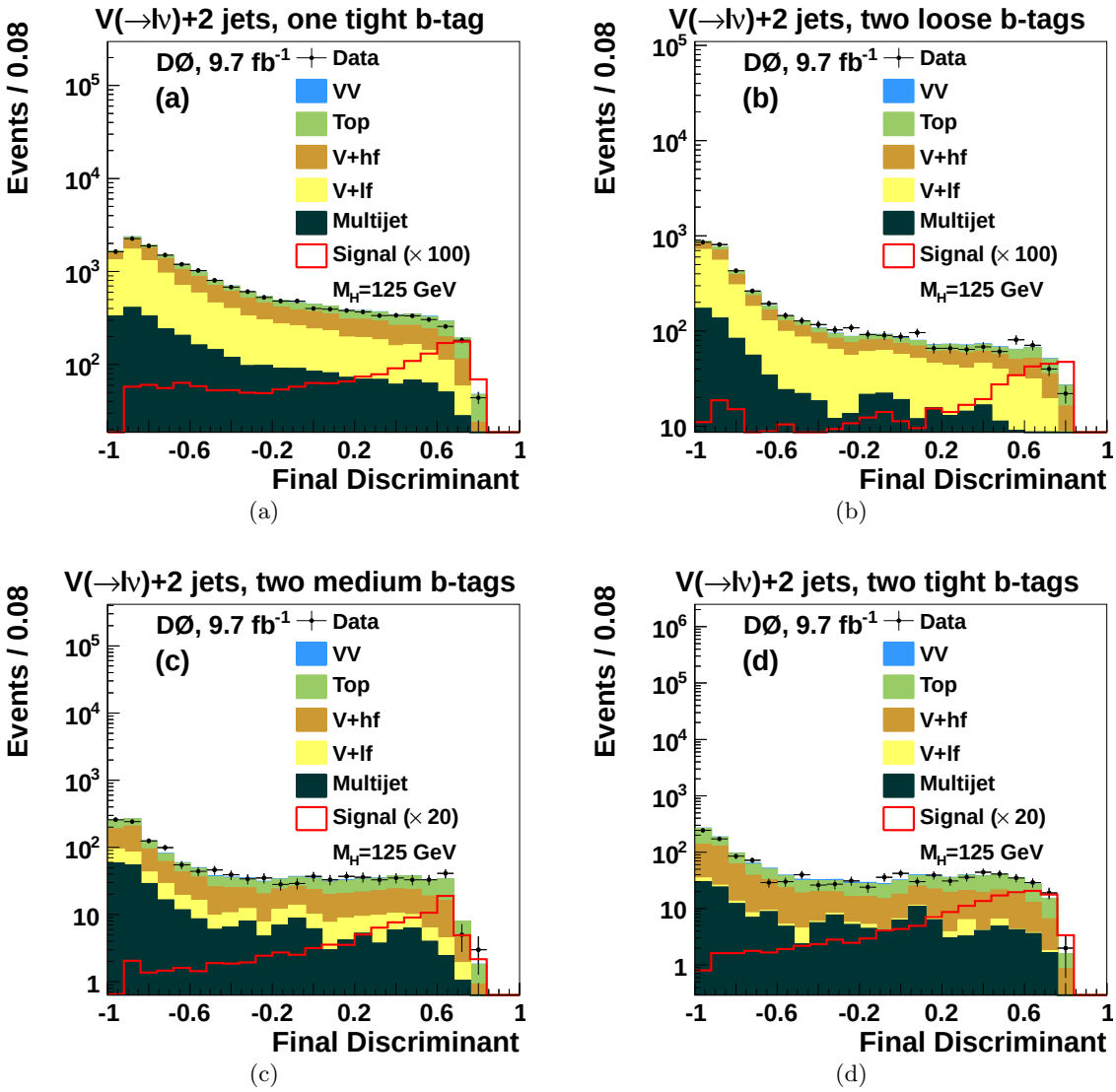


Figure 7.13: Distributions of the final discriminant output, after the maximum likelihood fit (described in Chapter 9), in events with exactly two jets and: (a) one tight b-tag, (b) two loose b-tags, (c) two medium b-tags, and (d) two tight b-tags. The signal is multiplied by 100, 100, 20, and 20, respectively.

7.4 WZ and ZZ Production with $Z \to b\bar{b}$

The SM processes $W(\to \ell\nu)Z(\to b\bar{b})$ and $Z(\to \ell\ell)Z(\to b\bar{b})$ where one of the leptons from the $Z \to \ell\ell$ decay is not reconstructed, result in the same final state signature as the Higgs boson in this search. Therefore, we search for these processes to validate our analysis methodology. The only change in the analysis is in the training of the final discriminant in events with two or three jets with one tight b-tag or two b-tags.

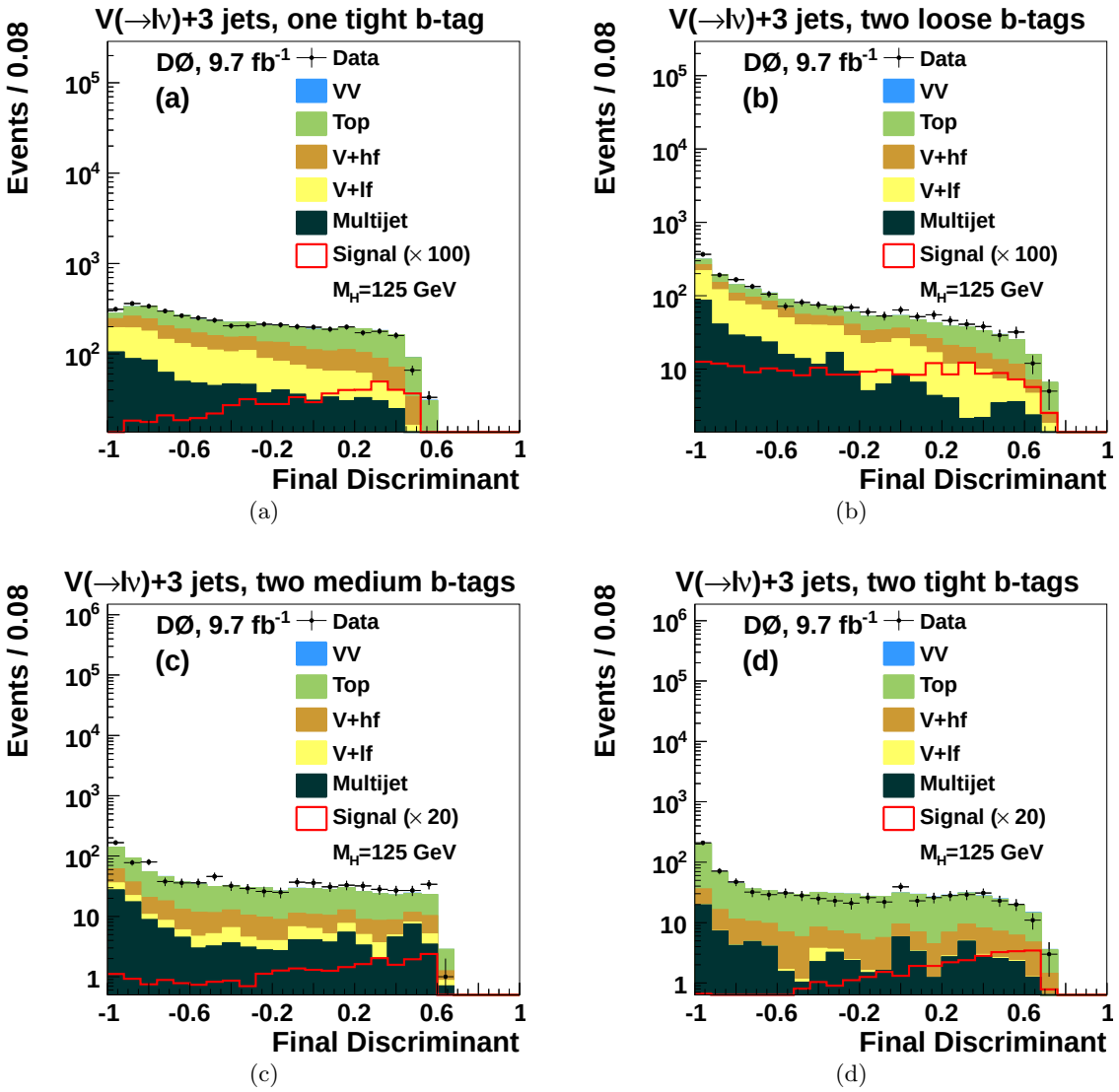


Figure 7.14: Distributions of the final discriminant output, after the maximum likelihood fit (described in Chapter 9), in events with exactly three jets and: (a) one tight b-tag, (b) two loose b-tags, (c) two medium b-tags, and (d) two tight b-tags. The signal is multiplied by 100, 100, 20, and 20, respectively.

7.4.1 Diboson MVA

In order to verify that we are sensitive to known SM processes with small cross sections, we use multivariate discriminants that utilize the same input variables as the discriminants used for the Higgs boson search, which means the data sample, reconstruction, process modelling, uncertainties, and sub-channel divisions are identical to those of the SM Higgs boson searches, but discriminant functions are trained to distinguish the contributions of SM diboson production from those of other backgrounds, and potential contributions from Higgs boson production are not considered. The modified $WH \to \ell\nu b\bar{b}$, $ZZ \to \ell\ell b\bar{b}$, and $ZZ \to \nu\nu b\bar{b}$ analyses (collectively called the VZ analyses) treat the WZ and ZZ processes as signal, and the WW process as a background.

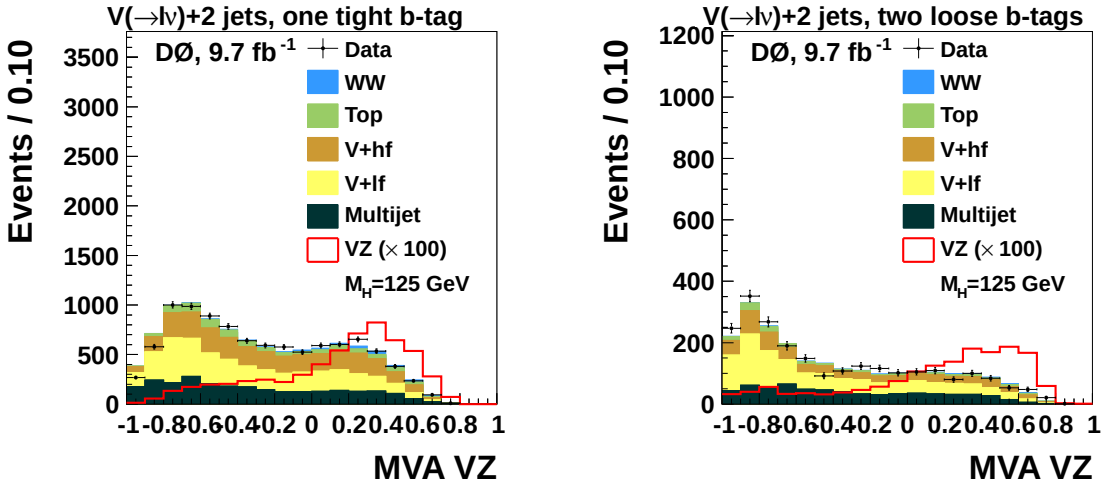
The input variables and training settings are exactly the same as the WH MVA. Figures 7.15 to 7.18 show the diboson MVA output distributions. We use them to estimate our sensitivity to the $WZ \to \ell\nu b\bar{b}$ and $ZZ \to \ell\ell b\bar{b}$ processes.

7.4.2 Result

The output of the diboson MVA discriminant, described in Section 7.4.1, is used to measure the combined WZ and ZZ limits and cross section by performing a maximum likelihood fit to data using signal plus background models, with maximization over the systematic uncertainties as described in detail in Chapter 8. Figure 7.19 shows the MVA discriminant output for the diboson cross section ($WZ + ZZ$) with background-subtracted data and signal scaled to the best fit value. The expected significance of the measurement using the MVA output is 1.8 standard deviations (s.d.). We measure a cross section of 0.50 ± 0.34 (stat.) ± 0.36 (syst.) times the expected SM cross section of 4.4 ± 0.3 pb.

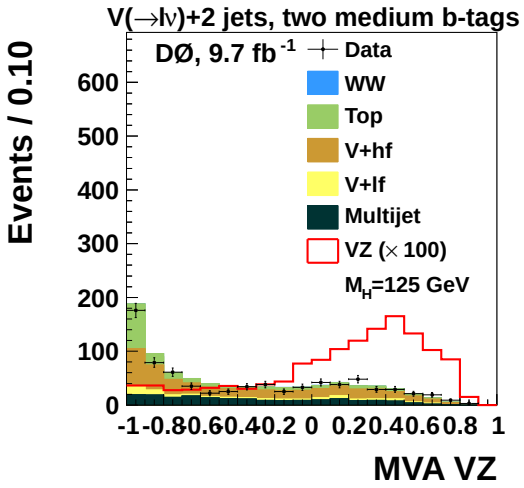
Within the D0 experiment, we also measure diboson production cross sections to validate our analyses techniques. Figure 7.20 shows the background-subtracted data for the dijet invariant mass in the VZ analyses. We observed a $WZ + ZZ$ production cross section of 0.73 ± 0.32 times the SM of 4.4 pb expected using MCFM. The non-zero significance is 2.4 s.d. with an expected significance of 3.4 s.d. We also observed a WW production cross of 1.01 ± 0.06 times the SM of 11.3 pb expected using MCFM. Since this VZ analysis have the same final states of our search for the Higgs boson, the ability to extract a small signal (VZ, i.e. $\ell\nu b\bar{b}$) from a large background validates our background modelling.

Among Tevatron combination, we also carry on the measurements of SM VZ production to validate our background modelling and methodology for WH search. WW production is considered as background. We observe a VZ cross section of $3.0 \pm 0.6(\text{stat}) \pm 0.7(\text{syst})$ pb whereas the expected SM is 4.4 ± 0.3 pb [92]. Figure 7.21 shows the combined background-subtracted data for dijet invariant mass in the VZ analysis. We fit the VZ signal and the background to the data, and then subtract the post-fit background. We also present the expected SM Higgs boson contribution.

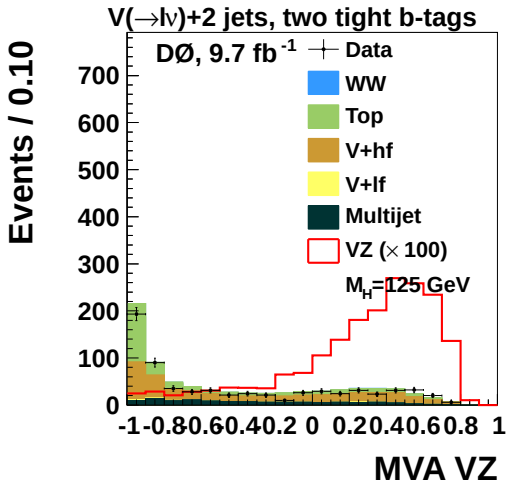


(a)

(b)



(c)



(d)

Figure 7.15: Distributions for the $W+2$ jet events (electron channel, Run II data). The data are compared to $Wb\bar{b}$, $t\bar{t}$, $W+$jets and other smaller expectations. The simulated processes are normalized to the integrated luminosity of the data sample using the expected cross sections (absolute normalization) except for the $W+2$ jets sample which is normalized to data on the "pre-tag sample", taking into account all the other backgrounds. a: 1 tight tag, b: 2 loose tag, c: 2 medium tag, d: 2 tight tag.

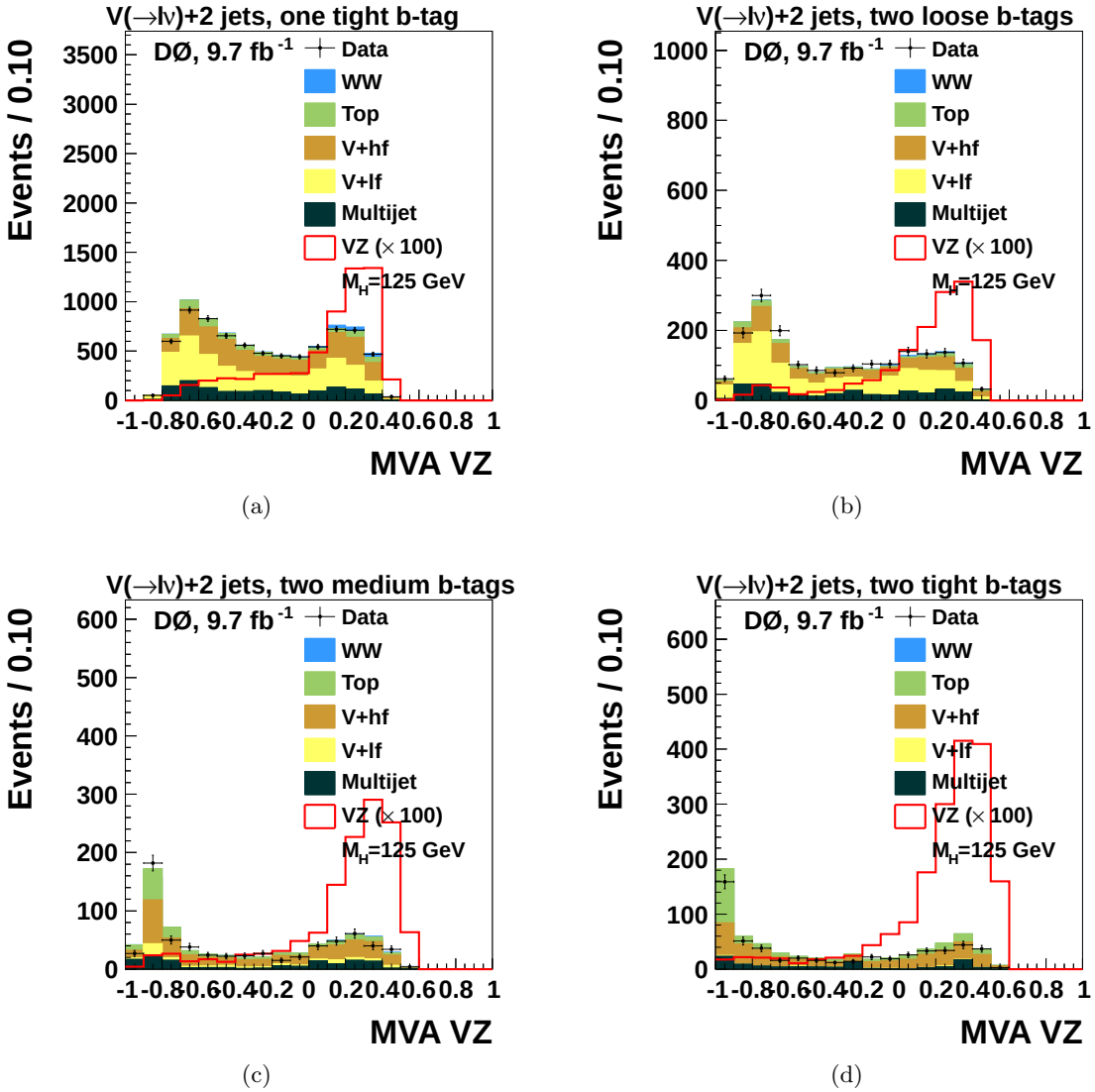
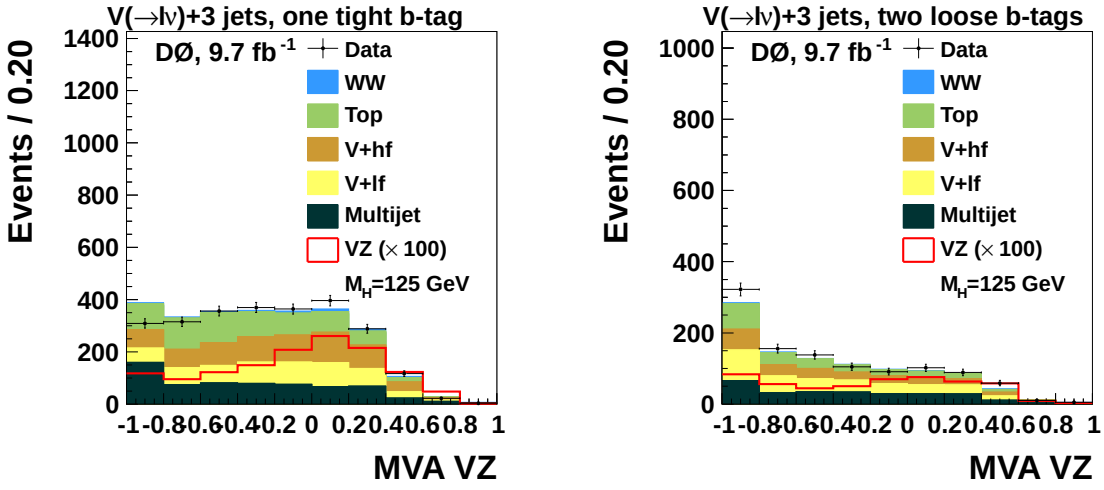
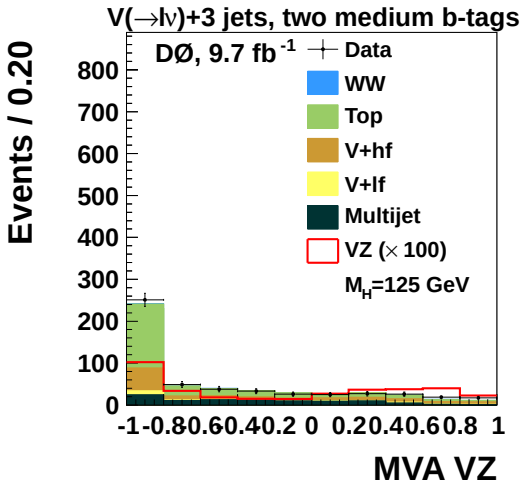


Figure 7.16: Distributions for the $W+2$ jet events (muon channel, Run II data). The data are compared to $Wb\bar{b}$, $t\bar{t}$, W+jets and other smaller expectations. The simulated processes are normalized to the integrated luminosity of the data sample using the expected cross sections (absolute normalization) except for the $W+2$ jets sample which is normalized to data on the "pre-tag sample", taking into account all the other backgrounds. a: 1 tight tag, b: 2 loose tag, c: 2 medium tag, d: 2 tight tag.

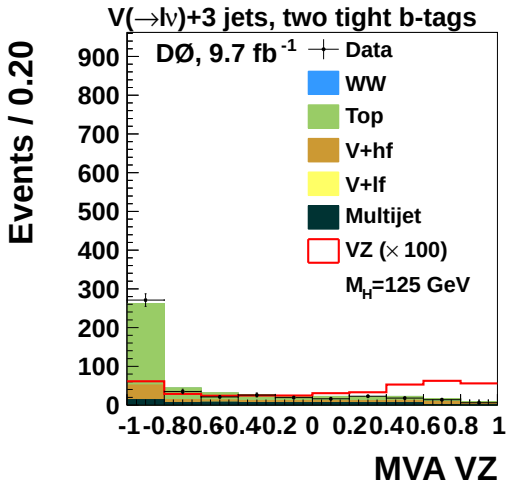


(a)

(b)

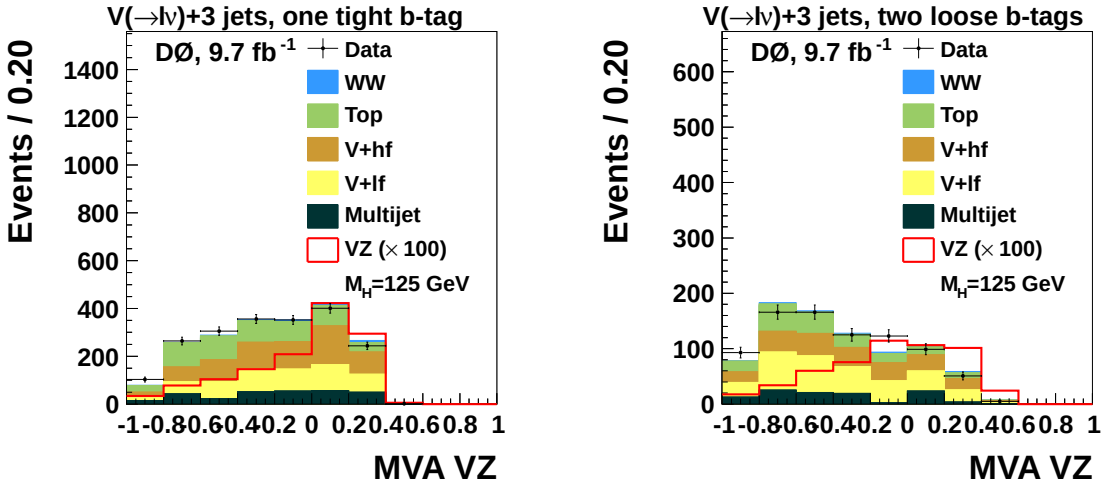


(c)



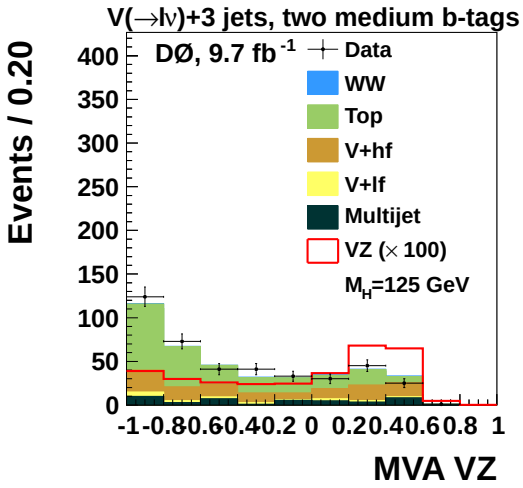
(d)

Figure 7.17: Distributions for the $W+3$ jet events (electron channel, Run II data). The data are compared to $Wb\bar{b}$, $t\bar{t}$, $W+$jets and other smaller expectations. The simulated processes are normalized to the integrated luminosity of the data sample using the expected cross sections (absolute normalization) except for the $W+2$ jets sample which is normalized to data on the "pre-tag sample", taking into account all the other backgrounds. a: 1 tight tag, b: 2 loose tag, c: 2 medium tag, d: 2 tight tag.

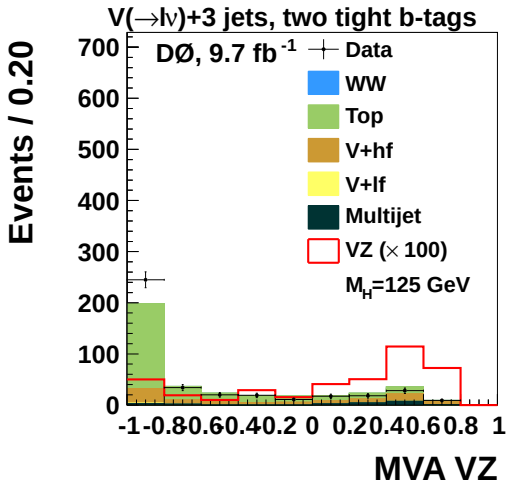


(a)

(b)



(c)



(d)

Figure 7.18: Distributions for the $W + 3$ jet events (muon channel, Run II data). The data are compared to $Wb\bar{b}$, $t\bar{t}$, W+jets and other smaller expectations. The simulated processes are normalized to the integrated luminosity of the data sample using the expected cross sections (absolute normalization) except for the $W + 2$ jets sample which is normalized to data on the "pre-tag sample", taking into account all the other backgrounds. a: 1 tight tag, b: 2 loose tag, c: 2 medium tag, d: 2 tight tag.

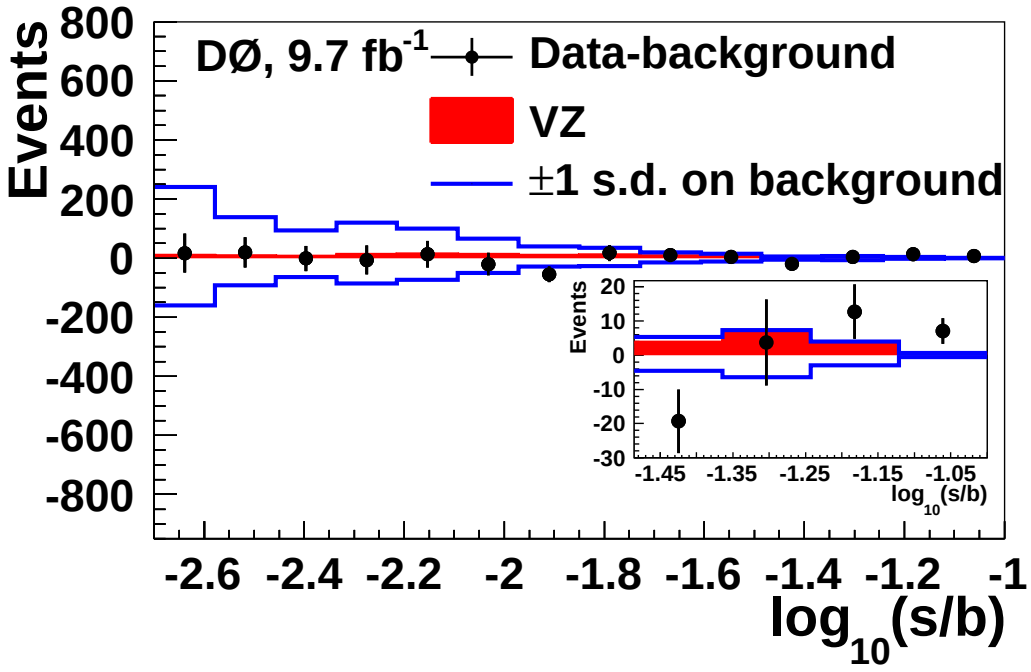


Figure 7.19: Final MVA discriminant output shown for the expected diboson signal and background-subtracted data rebinned as a function of $\log(S/B)$, after the maximum likelihood fit, summed over b-tag channels. The error bars on data points represent the statistical uncertainty only. The post-fit systematic uncertainties are represented by the solid lines. The signal expectation is shown scaled to the best fit value. The inset gives an expanded view of the high $\log(S/B)$ region.

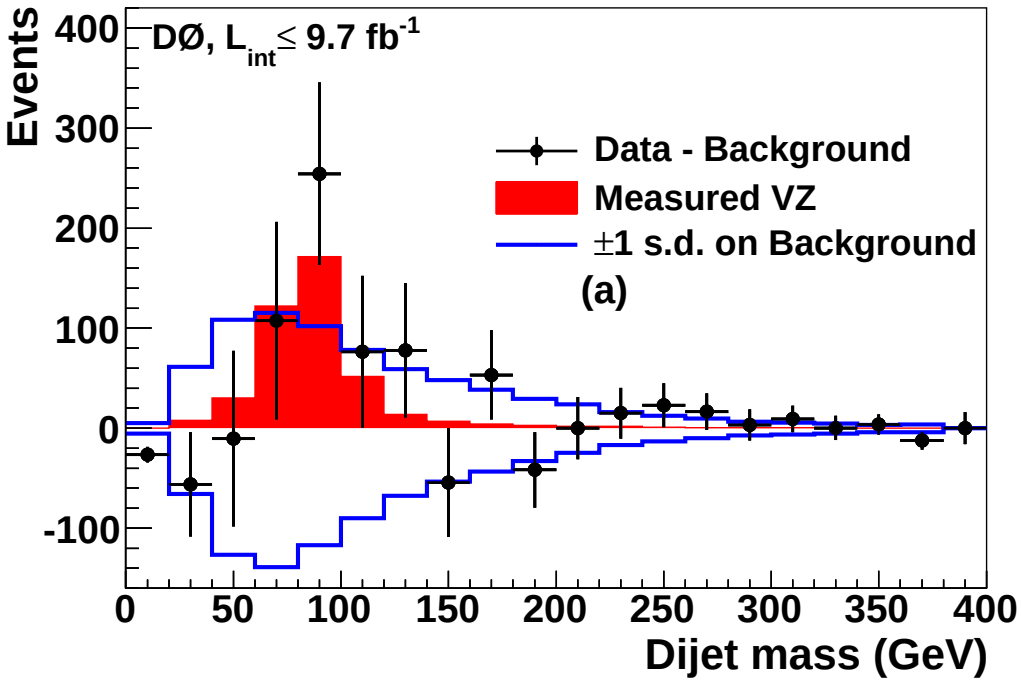


Figure 7.20: The dijet invariant mass in the combined $VZ \to Vb\bar{b}$ analyses. The black points show the background-subtracted data (with statistical errors), the red region shows the measured VZ signal, and the blue solid lines represent the ± 1 s.d. post-fit (after a fit to the $s + b$ hypothesis) systematic uncertainties constrained by data.

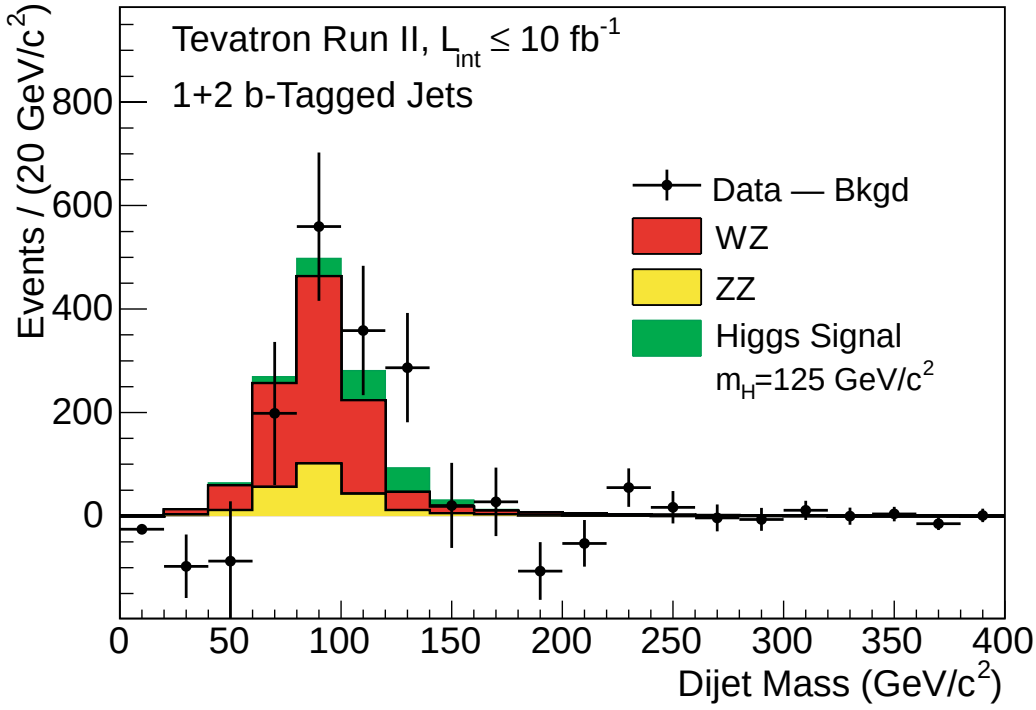


Figure 7.21: The dijet invariant mass in the combined $VZ \to Vb\bar{b}$ analyses on summed over CDF and D0's channels. The black points show the background-subtracted data of the reconstructed dijet mass (with statistical error bars). The red and yellow region shows the WZ and ZZ signal respectively. The green region shows the expected SM Higgs ($m_H = 125$ GeV) contributions.

Chapter 8

Systematic Uncertainties

We have obtained the final MVA discriminants in the previous Chapter 7, before we go to make the statistical statement about our data in the next Chapter 9, we must account for sources of uncertainty in our analysis. This includes both statistical and systematic uncertainties. The statistical uncertainties come from the number of data and MC events we have to describe kinematic distributions. The systematic uncertainties are uncertainties resulting from object reconstruction and calibration and MC generation or theory, and we are going to describe the various systematic uncertainties in this Chapter.

To determine the systematic uncertainties on signals and backgrounds, we repeat the full analysis with individually varying each source of systematic uncertainty by $\pm 1\sigma$ (where σ is the size of the uncertainty) for the parameter on which we are uncertain, and then taking the ratio of the obtained distribution to the nominal result (where no systematic variation has been applied).

For each lepton-type (e and μ) and for each jet multiplicity (2-jet, 3-jet) sample, we study all the sources of systematic uncertainty one by one, i.e. when we vary the luminosity systematics for instance, it's done for all categories of samples, and at the same time the other sources of systematics keep as the nominal. Each systematic variation is presented and applied for signal and for each of the main Standard Model backgrounds separately. Since the multijet (QCD) background is derived from data, we separately treat that. Systematic uncertainties are generally evaluated separately for each epoch, but then combined together in the limit setting.

Further details on the determination of the systematics (e.g the uncertainty arising from the uncertainty in the efficiency ratios between data and simulation, the uncertainties on the propagation of trigger, energy calibration, smearing etc.,) are described below.

Overall the total experimental systematic uncertainty for WH production is approximately 6%. We can divide the systematic uncertainties into two categories: uncertainties that affect only the rate of a signal or background process, and/or uncertainties that affect the shape of our MVA output distribution.

8.1 Uncertainties that affect only the rates of given processes

8.1.1 Cross Section Uncertainties

The theoretical uncertainties on the production cross sections times branching ratios of the background processes are 7% for $t\bar{t}$ production (as **Bkgd_Xsec_Top**), 7% for single-top

production, (as **Bkgd_Xsec_singletop**), 6% for WW, WZ and ZZ inclusive production (as **Bkgd_Xsec_EW**) [19].

We also apply uncertainties to the $K-$factors applied to the $W+$jets samples, which directly affect the cross section. We apply a 20% uncertainty on the $W/Z + HF$ $K-$factor (as **Bkgd_Xsec_W+HF**)(estimated from mcfm [93, 94]), and a 6% uncertainty on the $W/Z + LF$ $K-$factor (as **Bkgd_Xsec_W+LF**) [92], as well as an additional factor for 3-jet events (**Bkgd_Xsec_W+LF_3j**), since the $V+$jets experimental scaling factor for the three-jet channel is different from unity.

8.1.2 Parton Density Functions (PDF)

Cross sections uncertainties detailed earlier take into account the parton density functions (PDF) uncertainties. However the PDF uncertainty used in MC generation could impact both the cross section and the kinematic acceptance. PDF systematic retain only the changes due to kinematic acceptance. The normalization shouldn't be affected at the generator level. Hence, the overall yield is kept the same as the nominal case in order not to impact the cross section of a process. The resultant shaped-only systematic uncertainty is propagated to the analysis. After selection cuts applied, the residual normalization effect comes from the acceptance change. This procedure prevents us to not double count cross section uncertainties for limit settings. We assign a flat 2% uncertainty for PDFs.

8.1.3 Luminosity

Since there is uncertainty in the measurement of the luminosity, which affects the expected rates of signal and background processes, we treat the luminosity uncertainty separately (Split **DZeroLumi** of 4.6% and **Lumi** of 4.0% for combination).

8.1.4 Lepton-ID

Systematic errors for electron and muon identification are of different origins since their identification relies on different detector properties.

EM-ID

The EM-ID systematic for electron identification, reconstruction efficiency and energy smearing is obtained by varying by $\pm 1\sigma$ for the upward and downward systematic uncertainties. Details are available in [95]. We assign a 3% uncertainty on the efficiency of electron identification. We combine the EM Trigger systematic (2%) into the EM-ID uncertainty.

μ ID

The muon identification systematic error has three components: uncertainty on the muon identification efficiency, uncertainty on the track reconstruction efficiency, and the uncertainty on the muon isolation efficiency. Separate uncertainties in the Run IIa [96] and Run IIb [97] samples are applied.

- Identification Efficiency uncertainty: We apply a 0.8% uncertainty in Run IIa and a 1.2% uncertainty in Run IIb. For Run IIb we also apply an additional 2% systematic in events with muon $p_T < 20$ GeV as recommended by the Muon ID group.

- Track Reconstruction Efficiency Uncertainty: We apply a 2.3% uncertainty in Run IIa and a 1.4% uncertainty in Run IIb.

- Isolation Efficiency Uncertainty: We apply a 3.8% uncertainty in Run IIa and a 0.9% uncertainty in Run IIb.

8.1.5 Jet-ID

To evaluate the systematic error on the Jet ID, we shift the removal scale factor down 1σ (i.e. more jets will be removed than in the nominal case.) For the $+1\sigma$ Jet ID variation, we symmetrize the -1σ distribution. To reduce statistical fluctuations that could give rise to a bias in the result we rebin the JSSR systematic histograms to provide a smoother shape. For uniformity of result with the $ZH \to \nu\nu b\bar{b}$ and $ZH \to \ell\ell b\bar{b}$ analyses when we combine, we treat the Jet ID uncertainty as a flat uncertainty of 2%.

8.1.6 QCD

The systematic uncertainty on the QCD background is estimated by varying separately the lepton efficiency and the jet fake rate by $\pm 1\sigma$, as described in Section 6.2.

For the electron channel, the normalization changes due to the modelling imperfection is seen in the W transverse mass distribution which is used in the determination of V+Jets and Multi-jet scale factors. When we use a W transverse mass cut of 40 GeV in the SF determination, we observe a -5% change in the V+Jets and a $+15\%$ change in the MJ scale factor. We assign this changes as systematic uncertainty for Multi-jet QCD.

For the muon channel, we assign a 20% flat uncertainty on QCD. Since we fit the multijet and V+jet background rates to data simultaneously, we include an uncertainty on the V+jets rate of 2%, that is anti-correlated with the multijet rate.We treat the electron and muon QCD uncertainties as uncorrelated and also do not correlate between the two- and three-jet samples within a lepton channel as the fake rates and fits for each jet multiplicity are derived independently.

8.2 Uncertainties that affect the shape of our final MVA output include

8.2.1 Taggability (TAG)

As explained in 4.2.1, the uncertainty on the jet taggability ($\sim 3\%$ per jet) comes from the difference between the taggability scale factors being derived with one jet and being derived with two jets. The systematic uncertainty for taggability is evaluated by shifting the parameterizations of the taggability Scale Factor functions up and down by 1σ. Uncertainties from the fit of the parametrization are propagated to the final discriminant, giving a shaped uncertainty. These scale factors should not depend on jet multiplicity, so the uncertainty is taken to be large enough to cover this difference.

8.2.2 B-ID

The systematic uncertainty on b-ID is evaluated by using the $\pm 1\sigma$ deviation from the Tagging rate function (TRF). In this analysis, we are applying direct tagging to the MC, and apply a scale factor which is the ratio of the data TRF and the MC TRF. The uncertainty

on the b-tag efficiency is correlated between b- and c-jets, and not correlated between light and heavy jets, so we include uncertainties on the b-tagging efficiency by varying the heavy flavour tag rate functions up and down by one standard deviation in samples containing heavy quark jets, and we vary the light quark tag rate functions up and down by one standard deviation in samples devoid of heavy quark jets. We denote uncorrelated uncertainty between different MC epochs as bTag_LF for light flavour jets and bTag_HF for heavy flavour jets. The size of these uncertainties is $\sim 2-3\%$ per heavy quark jet, and $\sim 10\%$ for light quark jets.

Figure 8.1 is the systematic uncertainty b-ID for heavy flavour jets, which is a typical shape systematic, i.e. the scale of the uncertainty is different for each bin of the final distribution, therefore introducing a shape dependence. And all the systematic plots can be found in our analysis note [98]. In the Figure 8.1, the $+1\sigma$ ratio is shown as a red line, while the corresponding 1σ ratio is shown in blue. The nominal random forest output distributions are also included in the figures as dashed black lines. These distributions are arbitrarily scaled for presentation purposes and plotted from the line $y = y\_axis\_min$.

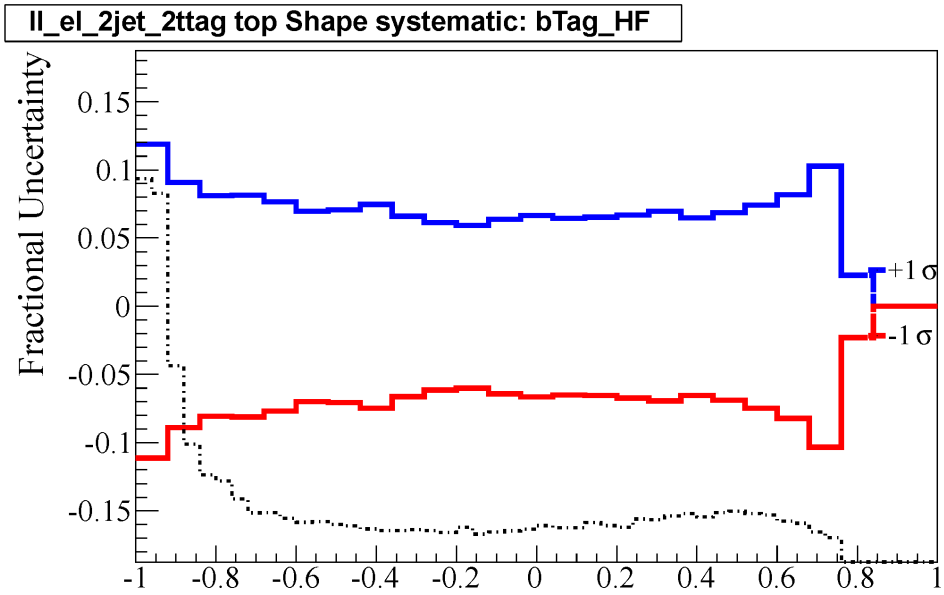


Figure 8.1: The systematic uncertainty on b-ID for heavy flavour jets, electron channel, two jet, two tight tag. The $+1\sigma$ variation is shown as a red line, while the corresponding -1σ variations is shown in blue. The shape of nominal random forest output distribution of the single top is also included in the figures as dashed black lines, with an arbitrary vertical scale.

8.2.3 Trigger

In the muon channel where all triggers are used (inclusive MU trigger, as mentioned in section 5.3.2) an additional systematic is applied to take into account the difference in the W_{p_T} and $\Delta R(j,j)$ alpgen reweightings used in the inclusive trigger sample compared to those used in the electron channel (i.e. the nominal trigger correction, which were derived in the super-OR muon sample, as mentioned in section 5.3.2). We first derive and apply an inclusive trigger correction based on the W_{p_T} and $\Delta R(j,j)$ that were derived in the super-OR sample. Then take difference between this and using the W_{p_T} and $\Delta R(j,j)$ reweightings derived on the inclusive trigger sample and apply it as the muon trigger systematic.

8.2.4 Jet Energy Scale (JES)

Due to low statistics in W+light jets and diboson MC samples after b-tagging, jet-related systematic errors are treated differently, namely Jet Energy Scale, jet resolution, jet identification and jet vertex confirmation.

The uncertainty on the jet energy scale(JES) is taken by shifting the JES parameters up and down by one standard deviation. Depending on whether the jets energy is shifted up or down, the jet acceptance is modified and events can either pass or not the jet cut. The final distribution shape can therefore be modified. Moreover, events can migrate between jet multiplicity channels. For instance, in an event with two jets passing the $p_T > 20$ GeV criteria and a third which does not fulfill this requirement, if the energy of the third jet is shifted, the event can potentially migrate from the 2 jets channel to the the three jets channel.

8.2.5 Jet Resolution (JSSR)

We evaluate the uncertainty on the jet energy resolution (JSSR) by shifting the JSSR parameters up and down by one standard deviation and we use the resulting jet instead of the nominal jet.

8.2.6 Vertex Confirmed Jet (VCJet)

The Vertex Confirmed Jets Scale Factor is applied as event weight, the systematic variation is also evaluated with event weight by shifting down or up by 1σ. We still use a separate taggability, therefore, we do not combine VC jet and taggable jet requirement. Because we do not use VC jets for Run IIa data, we apply this systematic only to the Run IIb MC.

8.2.7 ALPGEN

Several systematic uncertainties arise from ALPGEN inaccuracy to model the dominant V+jets backgrounds:

- A systematic uncertainty is evaluated for each kinematic reweighting for W/Z+jets samples only. A function is fitted and the uncertainty from the fit (using the correlation matrix from the fit) is taken as systematic uncertainty for the η reweightings of the leading and second leading jets as well as the reweighting of the ΔR between the two leading jets (for details on reweighting see Section 6.1). In addition, a shape uncertainty is assigned to the p_T^V distribution ($V = W, Z$).

- ALPGEN MLM matching parameters are varied. These changes impact the shape of the dijet invariant mass distribution and are taken as $\pm 1\sigma$ uncertainty bands. This shaped systematic error impacts only W/Z+light jets samples. $t\bar{t}$ samples, also produced with ALPGEN, are not affected by this uncertainty even though they are produced using the same generators. The reason is that jets from W decays in $t\bar{t}$+$0lp$[1] do not pass the process of matching since they do not originate from a radiation. However, $t\bar{t} + nlp$ $(n = 1, 2)$ could be affected by such systematic treatment, but the effect is expected to be small. More details on the determination of this uncertainty can be found in Reference [99].

- D0 analyses use the same tuning of ALPGEN parameters at the parton-level. Systematic uncertainties related to the ALPGEN event scale (k_T and Q^2) and the underlying event modelling are introduced. The scale uncertainty is evaluated separately for W/Z+light partons and W/Z+heavy partons. The underlying uncertainty is evaluated on W/Z+light partons, and applied to all W/Z+jets. More details on the determination of these uncertainties can be found in [99].

Previous studies of the ratio of the data-to-background expectation at the tagged stage, where we start to see W+heavy flavour events, showed a shape dependence in the single, double and combined tag samples, though within assigned systematic uncertainties.

8.2.8 Triangle cut

The uncertainty on the shape of our multijet estimate is determined by relaxing the requirement from Section 5.3 on M_T^W to $M_T^W > 30$ GeV $0.5 \times \not\!\!E_T$ and repeating the analysis with this selection in place. The positive and negative variations are taken to be symmetric. As with the uncertainty on the multijet rate, we do not correlate the multijet shape uncertainty across lepton, jet multiplicity, and b-tag categories.

8.3 A Summary of the Systematic Uncertainties

Distributions showing the $\pm 1\sigma$ variations are given in our analysis note [98] for shaped systematic errors. A summary of the systematic uncertainties, and their status as shape-only, normalization-only, or shape-and-normalization changing, appears in Tables 8.1 and 8.2.

[1] lp stands for light partons originating from additional radiations.

Collie name	Explanation, see following tables for the processes to which they apply
DZero_Lumi	Part of the Luminosity uncertainty coming from Dzero
Lumi	Part of the Luminosity uncertainty correlated with CDF
EMID	EM-id, -reconstruction, -scale, -smearing and -trigger (2%) added in quadrature
EMeff	systematic uncertainty for mis modelling on eta (north-south asymmetry)
MUID	muon-id , track-reconstruction, and -isolation efficiency
MUTrigger_WH	uncertainty on muon inclusive trigger correction.
JES	Jet Energy Scale uncertainties
JSSR	Jet Shifting/Smearing/Resolution uncertainties.
JetID	Jet-ID uncertainties
VCJet	Vertex confirmation scale factor uncertainties.
Tagga	Taggability uncertainties.
bTag_HF	Uncertainties on data/MC TRF ratios for heavy flavour jets
bTag_LF	Uncertainties on data/MC TRF ratios for light flavour jets
QCDev_2j	uncertainty on the electron QCD for 2jet sample, flat 15%
QCDev_3j	uncertainty on the electron QCD for 3jet sample, flat 15%
QCDev_sha	uncertainty on the electron QCD, shape
QCDmv_2j	uncertainty on the muon QCD for 2jet sample, flat 20%.
QCDmv_3j	uncertainty on the muon QCD for 3jet sample, flat 20%.
QCDmv_sha	uncertainty on the muon QCD, shape
PDF	Flat 2%.
ALPRW	uncertainty on the ALPGEN reweighting. Applied only to W/Z+jets.
ALP	ULE uncertainty on the underlying event. Applied only to W/Z+jets.
ALP_SCA	uncertainty on the ALPGEN scale parameter. Applied only to W/Z+jets.
ALP_MLM	uncertainty on the MLM matching parameter. Applied only to W/Z+light jets.
Bkgd_Xsec_EW	Diboson cross section uncertainty (6% .)
Bkgd_Xsec_Top	$t\bar{t}$ cross section uncertainty (7% .)
Bkgd_Xsec_singletop	single top cross section uncertainty (7% .)
Bkgd_Xsec_HF	uncertainty on the heavy flavour K-factor (20%.) Applied only to W/Z+$c\bar{c}$, $b\bar{b}$ samples.
Bkgd_Xsec_LF	uncertainty on the W/Z+light jet cross section (6% .) Applied only to both W/Z+light jet and W/Z+heavy jet samples.
Bkgd_Xsec_LF_3j	uncertainty on the W/Z+light jet cross section in 3jet sample(20% .) Applied only to both W/Z+light jet and W/Z+heavy jet samples only in 3jet.

Table 8.1: Explanations of the systematic error naming. The errors having a name starting with Bkgd apply only to the background, the others to Background and Signal, except for the QCD errors.

	electron	muon
DZero_Lumi	4.6%	
Lumi	4.0%	
EMID	SN	
MUID		3.0%
MUTrigger_WH		SN
JES	SN	
JSSR	SN	
JetID	2%	
VCJet	SN	
Tagga	SN	
bTag_HF	SN	
bTag_LF	SN	
QCDev_2j	15%	
QCDev_3j	15%	
QCDev_sha	S	
QCDmv_2j		20%
QCDmv_3j		20%
QCDmv_sha		S
PDF	2.0%	
ALPRW	SN	
ALP_ULE	SN	
ALP_SCA	SN	
ALP_MLM	SN	
Bkgd_Xsec_EW	6.0%	
Bkgd_Xsec_Top	7.0%	
Bkgd_Xsec_singletop	7.0%	
Bkgd_Xsec_HF	20.0%	
Bkgd_Xsec_LF	6.0%	
Bkgd_Xsec_LF_3j	20.0%	

Table 8.2: Correlation across the sixteen channels (2jet/3jet, electron/muon, 1tight/2loose/2medium/2tight tag) of the systematics uncertainties, see table 8.1 for explanations of these uncertainties. A given systematic is 100% correlated across all channels to which it is applied. A blank entry in the table means that that systematic is not applied in that channel. The errors having a name starting with Bkgd apply only to the background, the other to Background and Signal, except for the last error of the table. An S means that the systematic affects shape but not normalization, while SN means the systematic has both shape and normalization effects, and a percentage represents a flat systematic of the quoted size.

Chapter 9

Results on the Higgs Boson Search

Once we get the distribution of discriminant that distinguishes the signal from our expected background, as we do for the MVAs described in Chapter 7, we want to be able to make some statistical statement about our data, i.e. we want to know how much our data resembles a model that contains our signal. Since we do not see a significant excess in our data over our background expectation, we set an upper limit on the WH production cross section at a confidence level of 95% using the LLR as a test statistic as described in Section 9.1. The results obtained for the search of the $WH \to l\nu b\bar{b}$ signal will be presented in Section 9.2. After combining all the decay channels at D0, the data exhibit an excess above the background prediction of up to two standard deviations consistent with the presence of a 125 GeV SM Higgs boson, as described in 9.3. And for the combination of D0 and CDF, after combining all the decay channels, we observed an excess for m_H, which is described at Section 9.4.

9.1 The CL_S Method

9.1.1 Sensitivity Estimator

Considering the WH events (with small cross section for Higgs boson production) at D0 (with large number of collisions), these are kind of small probability events. To make a statistical statements about data, we start from describing the likelihood of observing d events given p predicted events by a Poisson parent distribution:

$$L(p,d) = \frac{p^d}{d!}e^{-p}$$

Here we use the Collie software package [100], based on a semi-frequentist CL_s method [101, 102], to compare two hypotheses: the Null (or background-only) hypothesis H0, and the Test (or signal+background) hypothesis H1. The likelihoods for each of these hypotheses are then:

$$L(b,d) = \frac{b^d}{d!}e^{-b}$$

$$L(s+b,d) = \frac{(s+b)^d}{d!}e^{-(s+b)}$$

where b and s correspond to the number of predicted background and signal events respectively. To evaluate and quantify which hypothesis is favoured for a given dataset, we take the likelihood ratio $Q = L(s+b)/L(s)$ as the sensitivity estimator, which is aimed to exploit as best as possible the sensitivity power of a given distribution (i.e. a final discriminant distribution). The likelihood ratio can be written as

$$Q(s,b,d) = \frac{L(d,s+b)}{L(d,b)} = \frac{e^{-(s+b)}(s+b)^d/d!}{e^{-b}b^d/d!} \tag{9.1}$$

This test statistic can be easily expanded to combine separate channels or bins of an histogram used as final variable of an analysis by multiplying each contribution:

$$\begin{aligned} Q' &= \prod_{i=1}^{N_{bins}} \prod_{j=1}^{N_{channels}} Q_{ij} \\ &= \prod_{i=1}^{N_{bins}} \prod_{j=1}^{N_{channels}} \frac{e^{-(s_{ij}+b_{ij})}(s_{ij}+b_{ij})^{d_{ij}}/d_{ij}!}{e^{-b_{ij}}b_{ij}^{d_{ij}}/d_{ij}!} \\ &= \prod_{i=1}^{N_{bins}} \prod_{j=1}^{N_{channels}} e^{-s_{ij}}(\frac{s_{ij}+b_{ij}}{b_{ij}})^{d_{ij}} \end{aligned} \tag{9.2}$$

where the sum over i is over the bins of the final MVA distribution, and the sum over j is over the different channels (lepton flavour, number of jets, and b-tag categories).

A more usual form can be used by transforming Q' to the negative log likelihood ratio (LLR):

$$\begin{aligned} LLR &= -2ln(Q') \\ &= 2\sum_{i=1}^{N_{bins}} \sum_{i=1}^{N_{channels}} (s_{ij} - d_{ij}ln(1+s_{ij}/b_{ij})) \end{aligned} \tag{9.3}$$

The sums come out because the probability for multiple independent bins is the product of the individual bin probabilities.

9.1.2 Handling of Systematic Uncertainties and Profile Likelihood Ratio

Before we go to quantify any excess that we observe above our background only prediction, by considering confidence levels and p-values, we should take into account systematic uncertainties on our predictions, since the predicted number of events in a given hypothesis is uncertain in reality.

Various parameters (have been described in Chapter 8) affect the b and $s+b$ predictions. Each of these parameters (referred to as nuisance parameters) has a range of possible values around some central value, and are considered as Gaussian distribution with the mean corresponding to the central value, and width corresponding to the uncertainty size. Thus for each hypothesis, we allow the number of predicted events to vary within the systematic uncertainties, and maximize the compatibility with our data. This method is called profiling.

According to detailed derivation in [103], the technique of χ^2 minimization has the benefit of reducing the degree to which the two model Gaussian probability distribution function (PDF) ($P(data|s+b)$ and $P(data|b)$) are broadened by the process of sampling

nuisance parameter values. And χ^2 can be written as:

$$\chi^2(\vec{R_k}) = 2\sum_{i=1}^{N_{bins}}\sum_{j=1}^{N_{channels}}(p(H)_i^{'} - d_{ij} - d_{ij}ln(\frac{p(H)_i^{'}}{d_{ij}})) + \sum_{k}^{N_n} R(H)_k^2 \tag{9.4}$$

where $p(H)_i^{'}$ is the predicted yield for a set of nuisance parameters, N_n is the total number of nuisance parameters (systematic uncertainties), and $R(H)_k$ is the deviation of the kth nuisance parameter from the central value in units of the Gaussian probability distribution width, σ_k:

$$R_k = \frac{\theta_k - \theta_k^0}{\sigma_k}$$

Now the log-likelihood ratio could be expressed using χ^2_{min}:

$$\begin{aligned} LLR &= -2ln(\frac{P(d|s+b,\theta_1)}{P(d|b,\theta_0)}) \\ &= \chi^2(b)_{min} - \chi^2(s+b)_{min} \end{aligned} \tag{9.5}$$

where θ_0 are the set of nuisance parameters that maximize the likelihood for b, and θ_1 are the set of nuisance parameters that maximize the likelihood for $s+b$.

9.1.3 Limit Calculation

The confidence level is a statement interpreted as the probability to observe a true value from a given interval. So a confidence interval with confidence level of $1-\alpha$ contains all values for which b is not rejected at a significance level α. The p-value is the probability that a hypothesis will fluctuate resulting in the observed data. Confidence levels for both hypotheses are related to the p-values as:

$$CL_B = 1 - PV_B = \int_{x_0}^{infty} H_{S+B}(LLR)dLLR$$

$$CL_{S+B} = PV_{S+B} = \int_{x_0}^{infty} H_B(LLR)dLLR$$

where x_0 is the reference LLR (for example the data LLR).

In order to prevent biases that could arise by choosing the CL_{S+B} to evaluate exclusion limits[1], the signal confidence level is estimated as:

$$CL_S = \frac{CL_{S+B}}{CL_B} \tag{9.6}$$

When $CL_S < \alpha$, we could exclude the s+b hypothesis at a confidence level of $1-\alpha$. An upper limit on the cross cross section can be set (with a C.L. of $1-\alpha$) by determining how much we would need to scale up the signal rate by until $CL_S = \alpha$.

For our analysis, to quote a "95% confidence level" (C.L.) limit, as it is usually given in high energy physics, the Collie software will iteratively scale the signal rate $s \propto \sigma_s L \epsilon_s$

[1]since using CL_{S+B} does not protect against making false exclusion statements from background-like fluctuations in our data

by increasing the signal SM cross section σ_s by steps to reach $\sigma_s^{95\%}$. Since the amount of signal impacts the CL_S value, this process will end when the condition

$$1 - CL_S \geq 95\% \tag{9.7}$$

is fulfilled, where CL_S can be either CL_S^{exp} (obtained by replacing the observed data with the median background expectation) or CL_S^{obs}.

9.2 Limits Obtained in the WH Analysis

In the WH analysis, using the LLR and CL method mentioned in the previous sections, we evaluate our data. As the branching fraction for the Higgs decay to b quarks is only significant below 150 GeV, at 13 discrete values within the Higgs boson mass range 90-150 GeV and spaced in increments of 5 GeV, we calculate the LLR for the full data combination including all the categories (electron/muon, two jets/three jets, one tight/two loose/two medium/two tight tagged b-jets, Run IIa/IIb1/IIb2/IIb34), and each of the two hypotheses (b/s+b) as a function of Higgs boson mass. The LLR for the $WH \to \ell\nu b\bar{b}$ search is shown in Figure 9.1. And we can see that our observed LLR agrees more with the background-only prediction at the mass region $m_H < 115$ GeV, and LLR_{obs} agrees more with the signal+background prediction while $m_H > 115$ GeV.

The MVA discriminant distributions, for the Higgs boson mass point MH = 125 GeV, after subtracting the total posterior background expectation are shown in Figure 9.2. The signal expectation is shown scaled to the observed upper limit (described later) and the uncertainties in the background after the constrained fit are shown by the solid lines.

Upper limits as the full combination are calculated by scaling the expected signal contribution to the value at which it can be excluded at the 95% C.L. ($\alpha = 0.05$). The expected limits are calculated from the background-only LLR distribution whereas the observed limits are quoted with respect to the LLR values measured in data. The upper limit on Higgs boson production cross section times branching ratio for each Higgs boson mass point considered as a ratio to the SM Higgs boson production cross section times branching ratio for the $WH \to \ell\nu b\bar{b}$ search in Figure 9.3 and Table 9.1. For a Higgs with $m_H = 125$ GeV, we set an observed (expected) limit of 4.8 (4.7) times the standard model cross section.

Table 9.1: The expected and observed 95% C.L. limits, as a function of the Higgs boson mass M_H, presented as ratios of production cross section times branching fraction to the expected SM prediction.

	WH 95% C.L. Limit $/\sigma_{SM}$												
M_H (GeV)	90	95	100	105	110	115	120	125	130	135	140	145	150
Expected	1.8	1.9	2.2	2.5	2.9	3.4	3.8	4.7	5.8	7.9	11.1	16.7	20.8
Observed	1.6	1.3	2.2	2.0	2.1	2.9	3.4	4.8	6.6	10.1	13.6	18.8	18.5

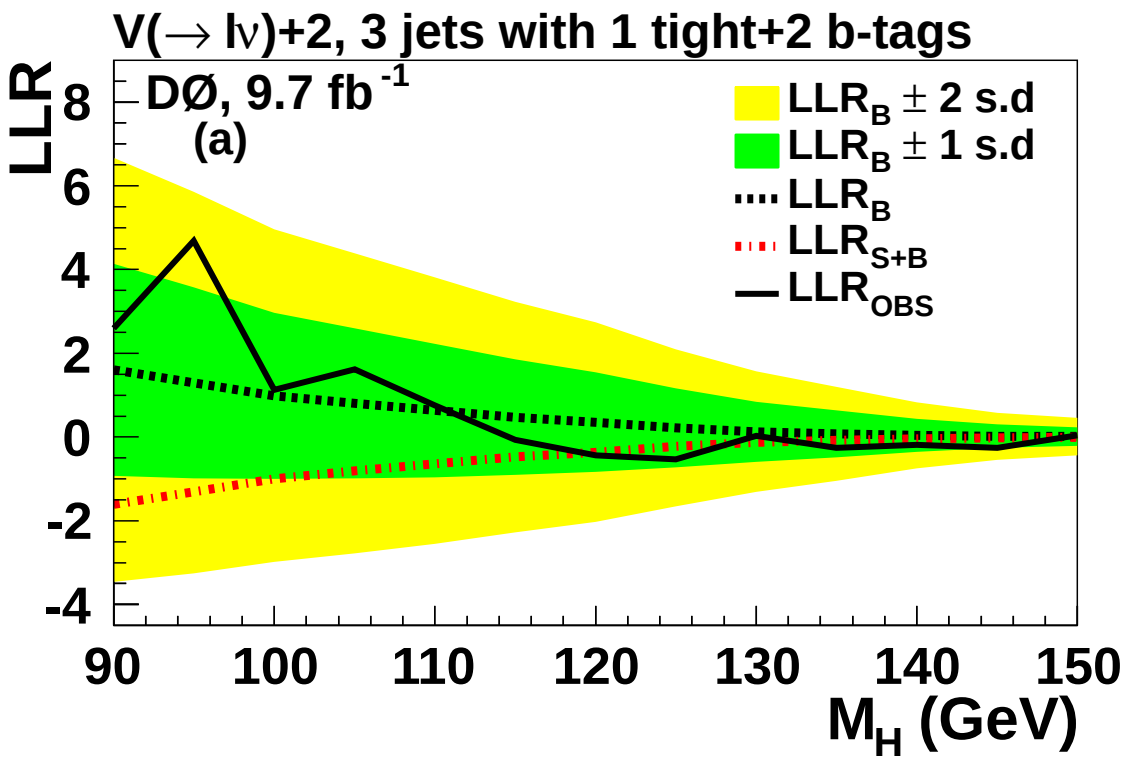


Figure 9.1: The expected and observed log-likelihood ratios as functions of the hypothesized Higgs boson mass m_H for the electron and muon, two- and three-jet, one tight and two-b-tag channels. The dashed red and black lines correspond to the median LLR of the signal + background and background-only hypotheses, respectively. The solid line corresponds to the LLR obtained from the data, and the shaded regions are the ± 1 s.d. and ± 2 s.d. values for the background-only hypothesis.

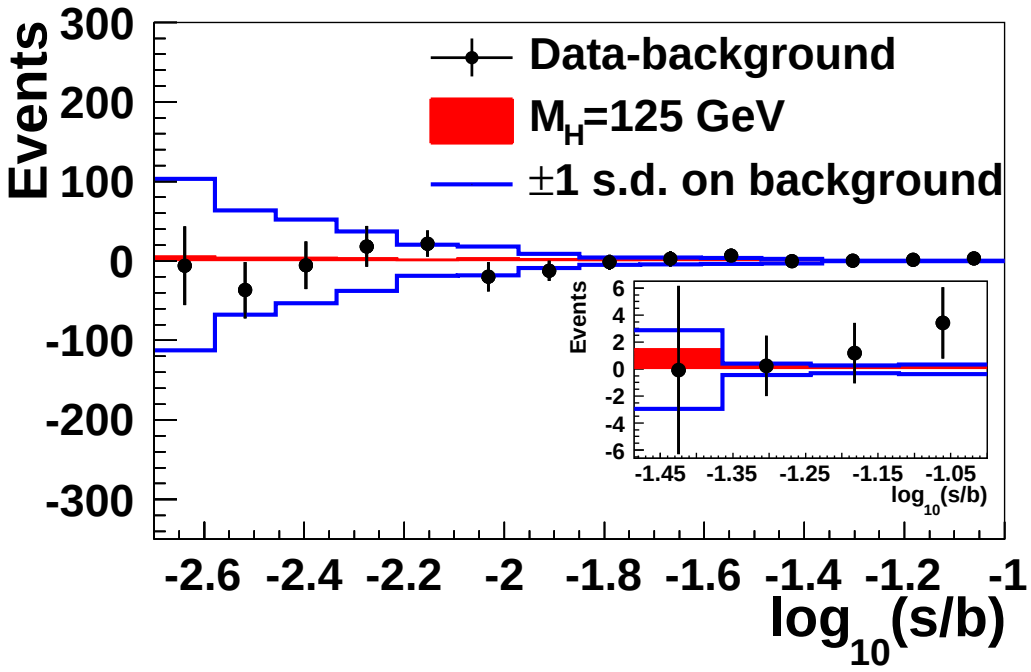


Figure 9.2: The MVA discriminant output distribution minus the total background expectation for $m_H = 125$ GeV rebinned as a function of $\log(S/B)$. The solid lines correspond to the post-fit (after the maximum likelihood fit) uncertainties. The signal expectation is shown scaled to the best-fit value. The inset gives an expanded view of the high- $\log(S/B)$ region.

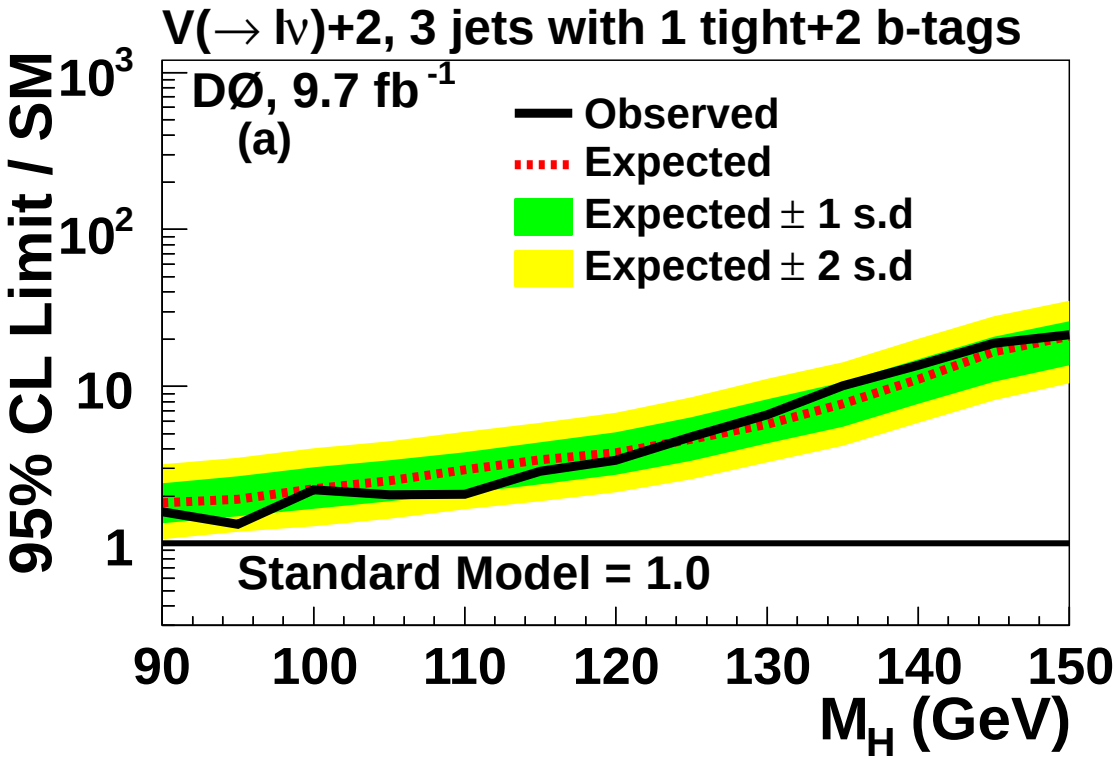


Figure 9.3: The expected and observed 95% C.L. upper limits on SM Higgs boson production for the electron and muon, two- and three-jet, one tight and two-b-tag channels. The limits are presented as ratios to the expected SM prediction. The dashed line corresponds to the expected limit, and the solid line corresponds to the limit observed in data. The shaded regions are the ± 1 s.d. and ± 2 s.d. values for the expected limit.

9.3 Combined search for the Higgs boson with the D0 experiment

While we do not see a significant excess in data over the background only prediction in the $WH \to \ell\nu b\bar{b}$ search on its own, we combine all Higgs boson searches at D0. Independently, the Higgs groups at D0 obtain limits on the product of the Higgs boson production cross section, σ_H, and branching fractions $\mathcal{B}(H \to b\bar{b}/W^+W^-/\tau^+\tau^-/\gamma\gamma)$ using individual channels [6, 7, 104], [105, 106, 79, 107, 108].

The LLR distribution for the D0 combination of all the channels is shown in Figure 9.4. This figure shows the median LLR values expected for the $s+b$ hypothesis (LLR_{s+b}), b hypothesis (LLR_b), and the results observed in data (LLR_{obs}). The shaded bands represent the ±1 and ±2 s.d. departures for LLR_b. From this figure we can see that the observed LLR is compatible with the $s+b$ hypothesis for $120 < M_H < 145$ GeV.

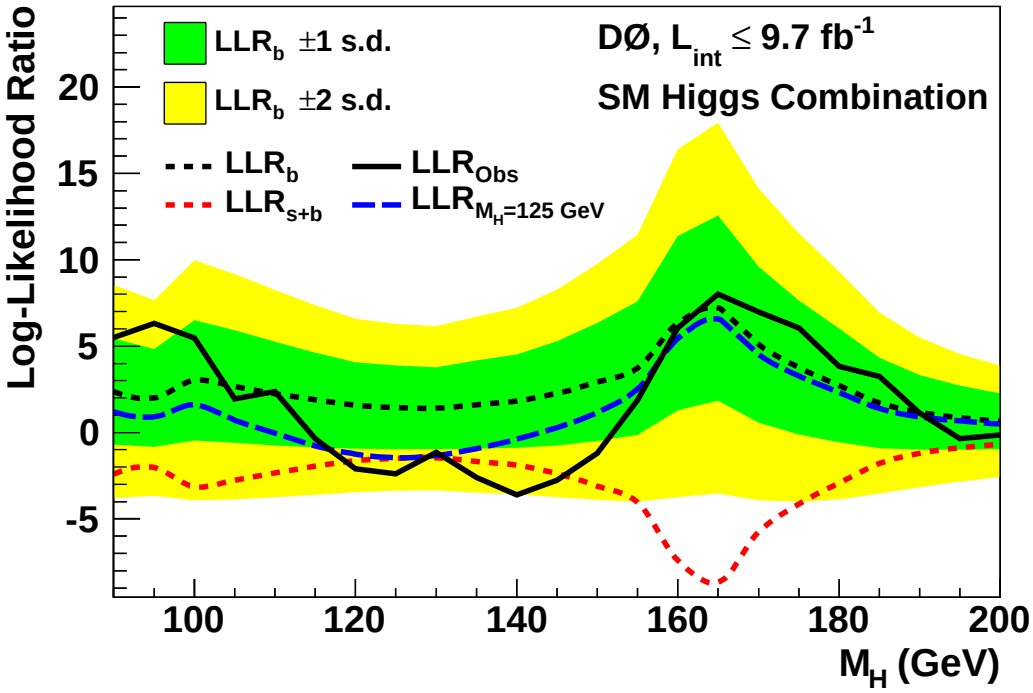


Figure 9.4: The observed (black solid line) and expected LLRs for the b (black short-dashed line) and s + b hypotheses (red/light short-dashed line), as well as the LLR expected in the presence of a SM Higgs boson with $m_H = 125$ GeV (blue long-dashed line) for all analyses combined for the range $90 \leq M_H \leq 200$ GeV. The shaded bands correspond, respectively, to the regions enclosing ±1 and ±2 s.d. fluctuations of the background.

Figure 9.5 shows the expected and observed upper limits on σ_H at 95% C.L. relative to the SM, for the mass region $90 \leq M_H \leq 200$ GeV, for all analyses combined. We exclude the SM Higgs boson at 95% C.L. in the mass ranges $90 < M_H < 101$ GeV and $157 < M_H < 178$ GeV. Our expected exclusion range is $155 < M_H < 175$ GeV.

As a further investigation of the excess, we calculate the best fit of the data to the ratio of σ_H to the SM prediction (σ_{Fit}/σ_{SM}). The result of this fit, shown in Figure 9.6, along with its band of ±1 s.d., yields a signal rate of approximately a factor of 1.4 larger than the SM cross section for m_H between 120 GeV and 145 GeV. For $m_H = 125$ GeV,

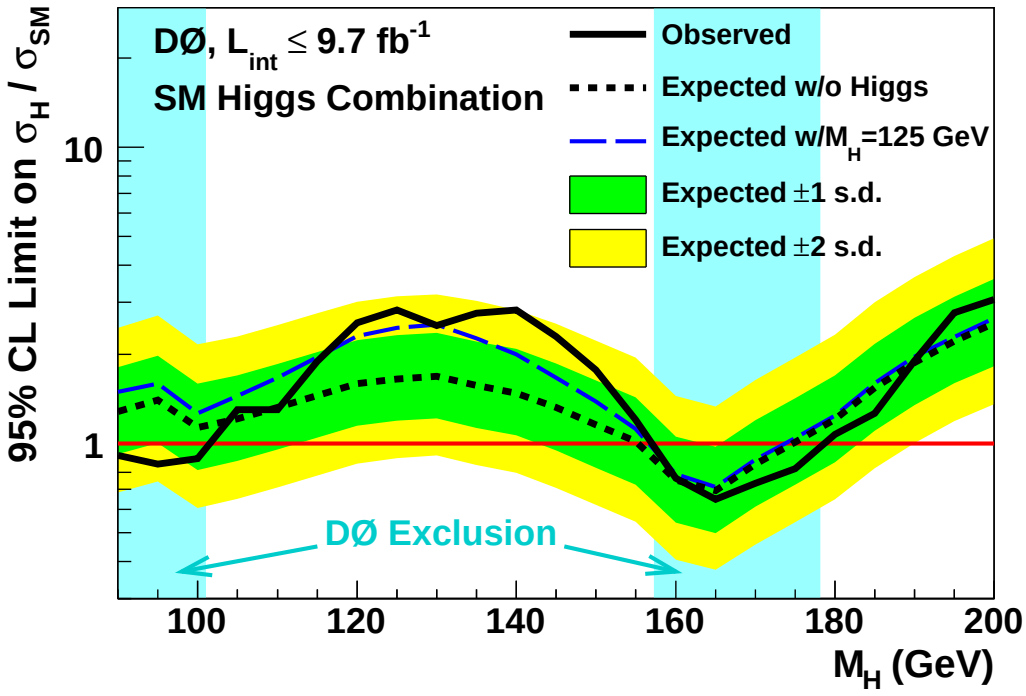


Figure 9.5: Expected (median) and observed ratios for the upper limits of the cross section σ_H at 95% C.L. relative to the SM values for all analyses combined for the range $90 \leq M_H \leq 200$ GeV. The shaded bands correspond to the regions enclosing ± 1 and ± 2 s.d. fluctuations of the background, respectively. The long-dashed line represents the expectation if a $m_H = 125$ GeV Higgs boson were present in the data with the SM cross section.

we obtain a ratio of 1.4±0.9. Compared to the narrow resonance observed at the LHC in the $H \to \gamma\gamma$ or $H \to ZZ \to 4\ell$ final states, at Tevatron we expect a Higgs boson signal to appear as a broad excess over background, due to the following dominated decay modes: the associated production with $H \to b\bar{b}$ decay and the $H \to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$, which have approximately 15% for the dijet invariant mass resolution ($H \to b\bar{b}$), or even poorer resolution ($H \to W^+W^-$).

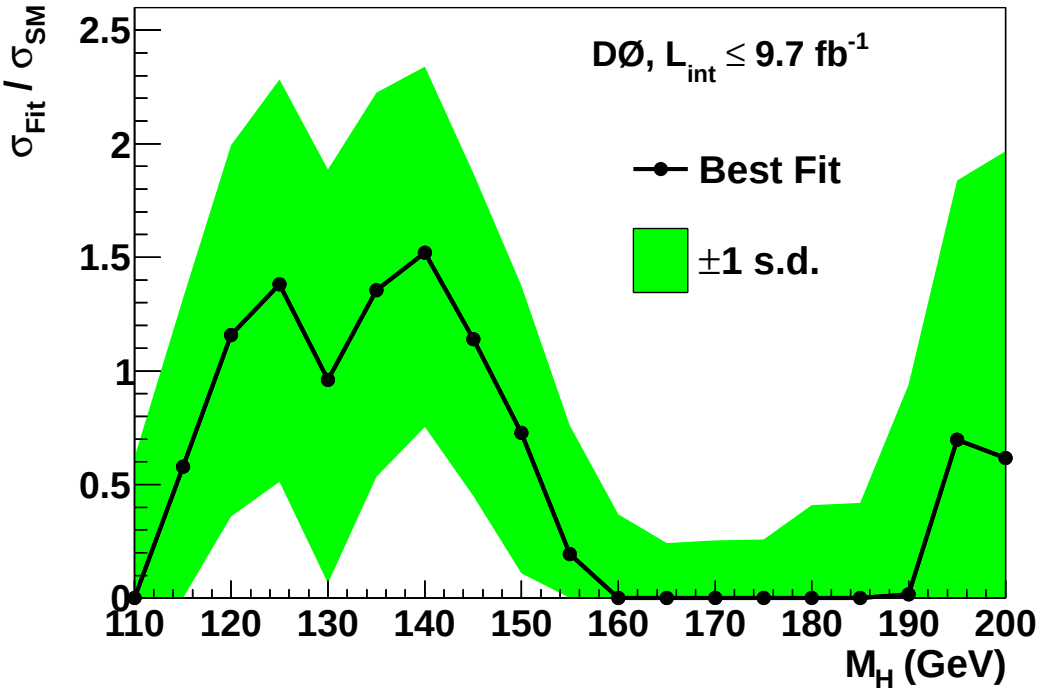


Figure 9.6: The best fit of the ratio $\sigma_H/(\sigma_H)_{SM}$ as a function of m_H for all analyses combined for the $110 \leq M_H \leq 200$ GeV. This indicates the values of the Higgs boson cross section that best match the observed data. The light shaded band indicates the ±1 s.d. region departure from the fit. The fit result is zero for masses below 110 GeV.

As the combination of searches for SM Higgs boson production with the D0 experiment using data corresponding to up to 9.7 fb^{-1} of $p\bar{p}$ collisions at $\sqrt{s} = 1.96$TeV, we set upper limits on the production cross section at 95% C.L. for Higgs boson masses of $90 < M_H < 200$GeV. The observed upper limits on SM Higgs boson production are $2.86(0.66)\times\sigma_{SM}$ at $m_H = 125(165)$ GeV, with an expected limit of $1.68\ (0.70)\times\sigma_{SM}$. We exclude the regions of $90 < M_H < 101$ GeV and $157 < M_H < 178$ GeV with an a priori expected exclusion of $155 < M_H < 175$ GeV. In the range of $m_H \approx 120-145$ GeV, the data exhibit an excess above the background prediction of up to two standard deviations consistent with the presence of a 125 GeV SM Higgs boson. The D0 analyses combined here also provide inputs to the overall Tevatron combination [5], which is described in Section 9.4

Since this thesis focuses on the $WH \to \ell\nu b\bar{b}$, we are interested in the $H \to b\bar{b}$ combination at D0. Figure 9.7 shows the LLR values from the combination of the results from the $ZH \to \ell\ell b\bar{b}$, $ZH \to \nu\nu b\bar{b}$ and $WH \to \ell\nu b\bar{b}$ searches, and illustrates a small excess that is compatible with the SM Higgs boson expected rate for $120 \leq M_H \leq 145$ GeV. Figure 9.8, as well as Tables 9.2, show the expected and observed 95% C.L. cross section limits in terms of ratio to the SM predictions for $H \to b\bar{b}$. Figure 9.9 shows the best fit of

the ratio $\sigma_H \cdot \mathcal{B}/(\sigma_H \cdot \mathcal{B})_{SM}$ for $m_H = 125$ GeV in each of the Higgs boson decay channels considered, as well as the central value for all analyses combined. These values are also given in Table 9.3. The D0 $H \to b\bar{b}$ analyses combined here also provide inputs to the overall Tevatron $H \to b\bar{b}$ combination [5], which will be described in Section 9.4

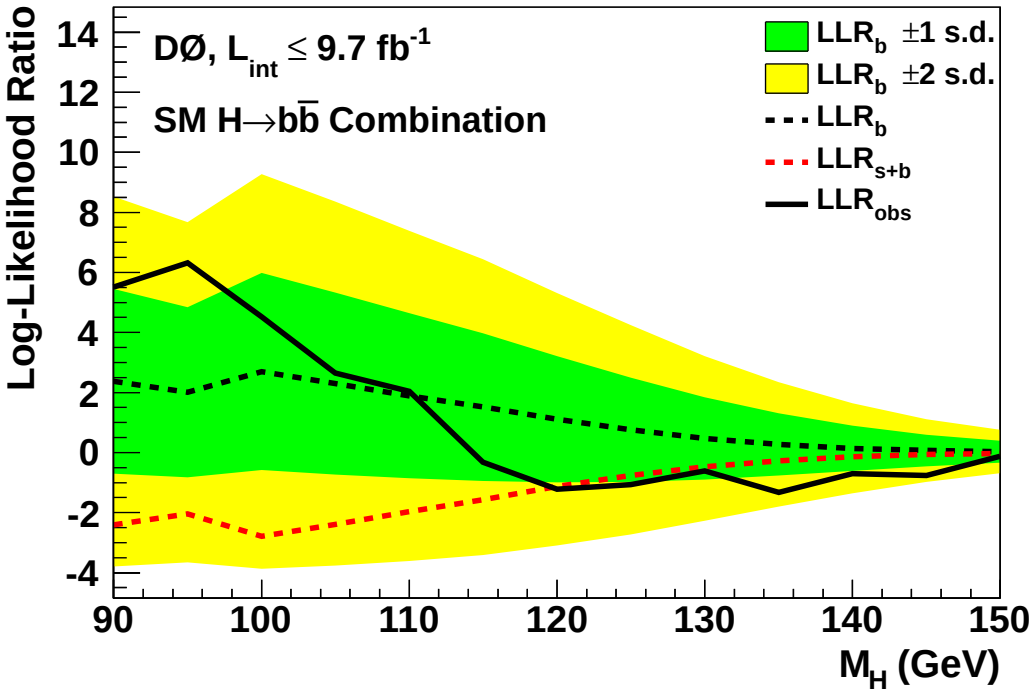


Figure 9.7: The observed (black solid line) and expected LLRs for the b (black short-dashed line) and $s+b$ hypotheses (red/light short-dashed line) for the combined $WH/ZH, H \to b\bar{b}$ analyses for the range $90 \leq M_H \leq 150$ GeV. The shaded bands correspond, respectively, to the regions enclosing ± 1 and ± 2 s.d. fluctuations of the background.

Table 9.2: Expected (median) and observed upper limits for $\sigma \times \mathcal{B}(H \to b\bar{b})$ relative to the SM at 95% C.L. for the combined $WH/ZH, H \to b\bar{b}$ analyses for the range $90 \leq M_H \leq 150$ GeV.

M_H (GeV)	90	95	100	105	110	115	120	125	130	135	140	145	150
Expected	1.29	1.40	1.21	1.31	1.45	1.63	1.92	2.33	2.99	3.96	5.52	7.91	11.35
Observed	0.96	0.89	1.05	1.33	1.51	2.25	2.96	3.49	4.29	6.92	8.65	13.85	13.90

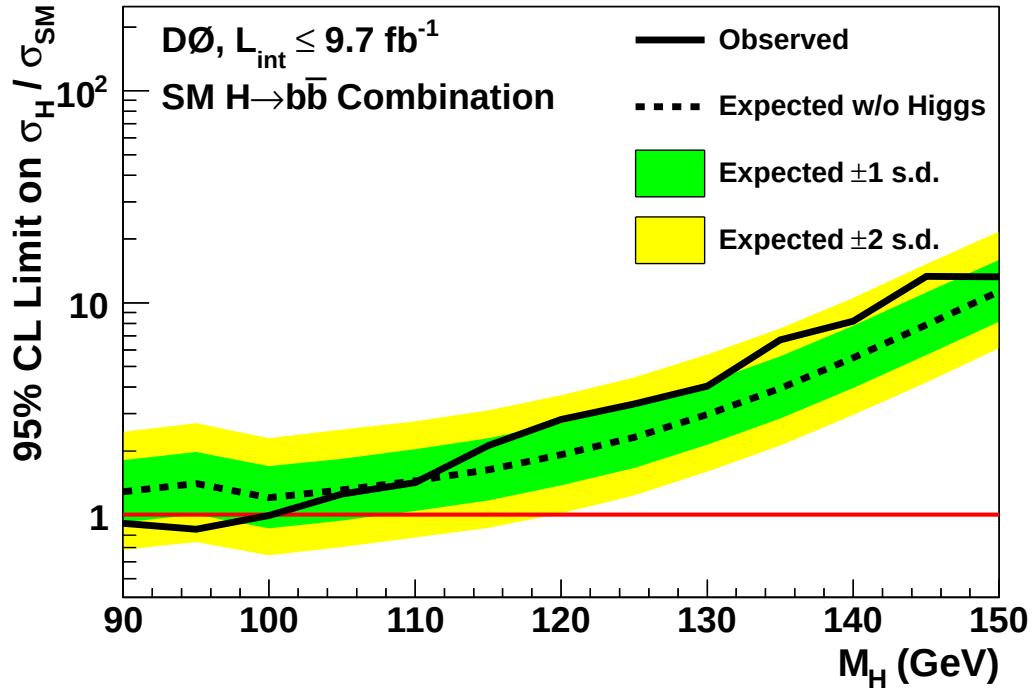


Figure 9.8: Expected (median) and observed ratios for the upper limits of the cross section H at 95% C.L. relative to the SM values for the combined $WH/ZH, H \to b\bar{b}$ analyses for the range $90 \leq M_H \leq 150$ GeV. The shaded bands correspond to the regions enclosing ± 1 and ± 2 s.d. fluctuations of the background, respectively.

Table 9.3: The best fit Higgs boson cross section times branching fraction as a ratio to the SM cross section times branching fraction for $m_H = 125$ GeV for the individual combinations according to Higgs boson decay mode, as well as the full combination.

Combined	$1.40^{+0.92}_{-0.88}$
$H \to \gamma\gamma$	$4.20^{+4.60}_{-4.20}$
$H \to W^+W^-$	$1.90^{+1.63}_{-1.52}$
$H \to \tau^+\tau^-$	$3.96^{+4.11}_{-3.38}$
$H \to b\bar{b}$	$1.23^{+1.24}_{-1.17}$

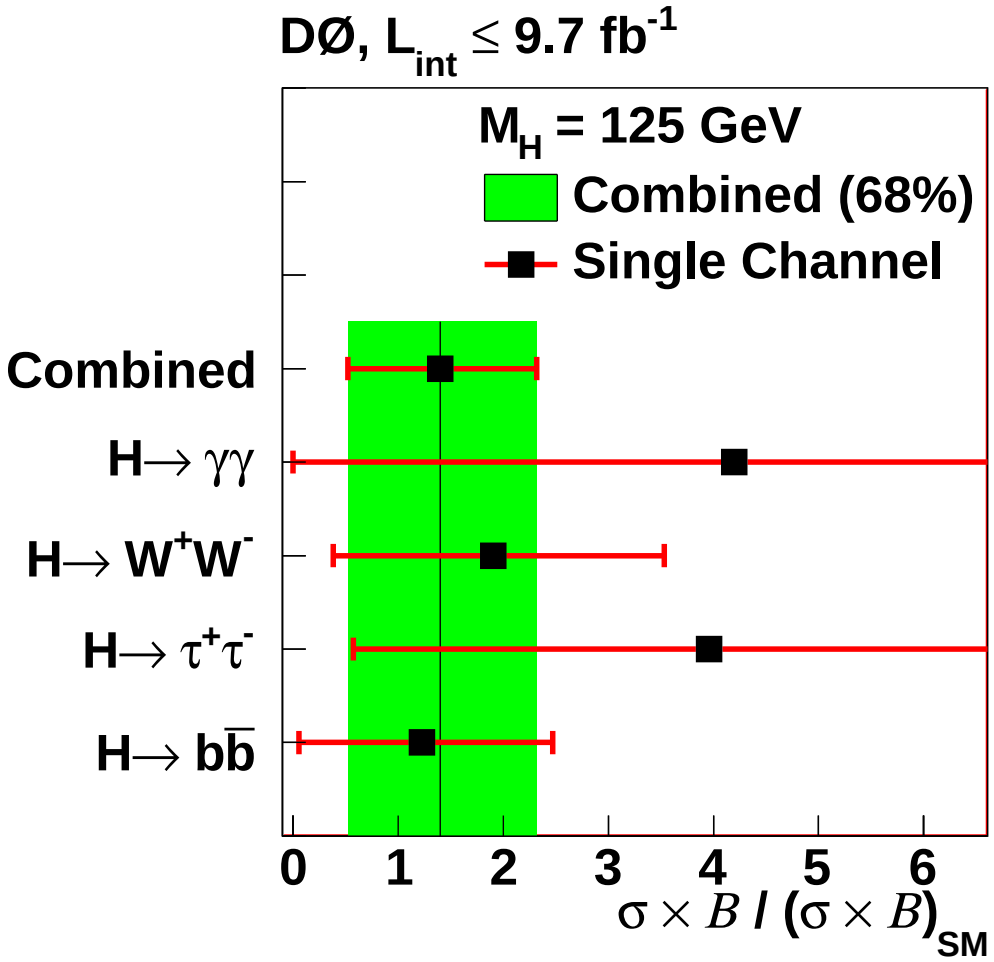


Figure 9.9: The best fit of $\sigma_H \cdot \mathcal{B}/(\sigma_H \cdot \mathcal{B})_{SM}$ for various Higgs boson decays for $m_H = 125$ GeV. The central value for all combined analyses is shown with its 1 s.d. band (shaded area).

9.4 Combined Higgs Boson Results from CDF and D0

We also combine the D0 results with those from CDF, i.e. all $H \to b\bar{b}$, $H \to W^+W^-$, $H \to \tau^+\tau^-$, $H \to \gamma\gamma$, and $H \to ZZ$ decay modes from both D0 and CDF. Figure 9.10 shows the LLR distributions for the combined analyses as functions of m_H in the range 90-200 GeV, along with the median of the LLR distributions for the background-only hypothesis (LLR_b), the signal-plus-background hypothesis (LLR_{s+b}), and the observed value for the data (LLR_{obs}). From the LLR figure we can see that the data are consistent with the background-only hypothesis (the black dashed line) at masses smaller than $\sim$ 110 GeV and above $\sim$ 145 GeV. For $115 < m_H < 140$ GeV, an excess well above two s.d. in the data with respect to the SM background expectation has an amplitude consistent with the expectation for a standard model Higgs boson (dashed red line).

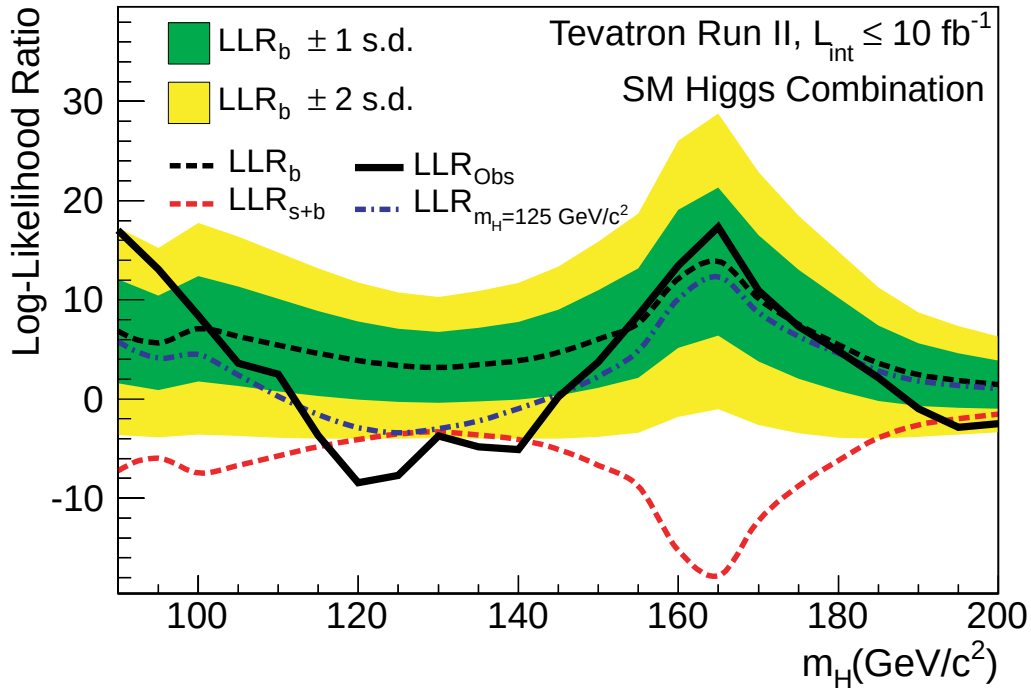


Figure 9.10: The log-likelihood ratio LLR as a function of Higgs boson mass for all of CDF and D0's SM Higgs boson searches in all decay modes combined. The solid line shows the observed LLR values, the dark long-dashed line shows the median expectation assuming no Higgs boson signal is present, and the dark- and light-shaded bands correspond, respectively, to the regions encompassing one and two s.d. fluctuations around the background-only expectation. The red long-dashed line shows the median expectation assuming a SM Higgs boson signal is present at each value of mH in turn. The blue short-dashed line shows the median expected LLR assuming the SM Higgs boson is present at $m_H = 125$ GeV.

At the Tevatron, the upper limit on SM Higgs boson production as a function of m_H is extracted in the range 90-200 GeV in terms of the ratio of the observed limit to the predicted SM rate. The ratios of the 95% C.L. expected and observed limit to the SM cross section using the Bayesian method [5] [2] are shown in Figure 9.11 for the combined CDF

[2]Bayesian method is a statistical technique, similar to the CL_S method we introduced in Section 9.1. In

and D0 analyses.

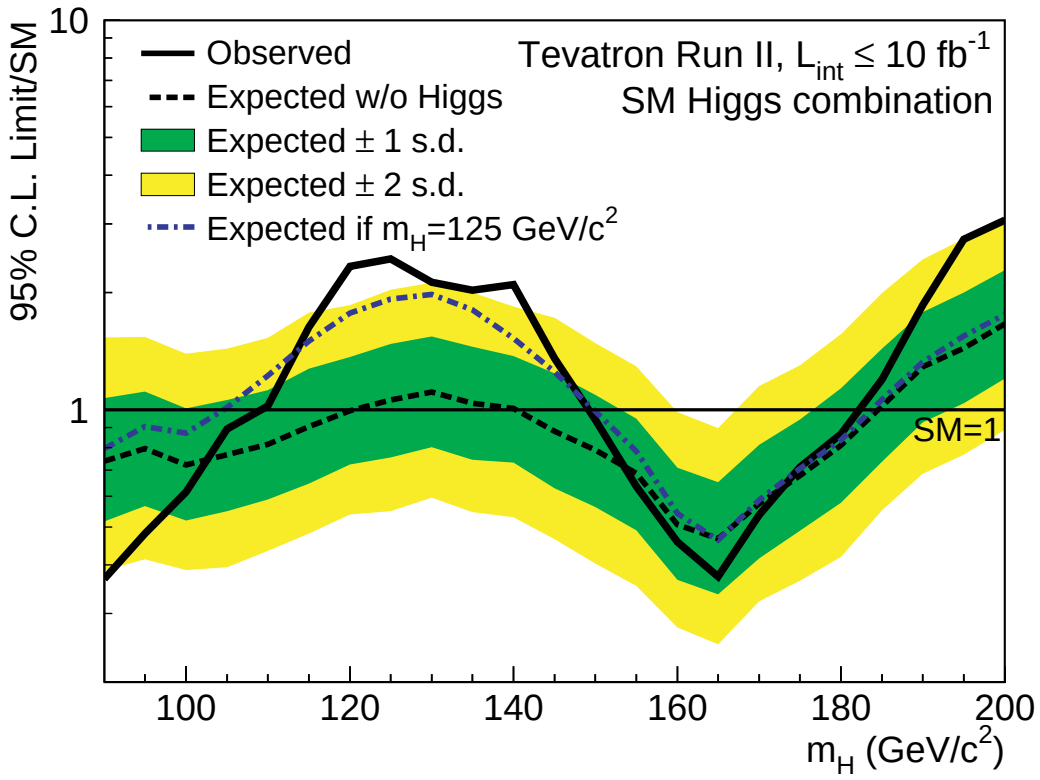


Figure 9.11: Observed and median expected (for the background-only hypothesis) 95% C.L. Bayesian upper production limits expressed as multiples of the SM cross section as a function of Higgs boson mass for the combined CDF and D0 searches in all decay modes. The dark- and light- shaded bands indicate, respectively, the one and two s.d probability regions in which the limits are expected to fluctuate in the absence of signal. The blue short-dashed line shows median expected limits assuming the SM Higgs boson is present at $m_H = 125$ GeV.

To quote ranges of Higgs boson masses that are excluded and that are expected to be excluded, we intersect linear interpolations of the observed and expected rate limits with the SM=1 line. The Higgs boson mass regions $90 < m_H < 109$ GeV and $149 < m_H < 182$ GeV are excluded at the 95% C.L.. The expected exclusion regions are $90 < m_H < 120$ GeV and $140 < m_H < 184$ GeV.

The observed excess for $115 < m_H < 140$ GeV is driven by an excess of data events with respect to the background predictions in the most sensitive bins of the discriminant distributions, favouring the hypothesis that a signal is present. The best-fit rate cross section, R^{fit}, is computed using the Bayesian calculation, and shown in Figure 9.12. The measured signal strength is within 1 s.d. of the expectation for a SM Higgs boson in the range $115 < m_H < 140$ GeV, with maximal strength between 120 GeV and 125 GeV. At 125 GeV, $R^{fit} = 1.44^{+0.49}_{-0.47}$ (stat)$^{+0.33}_{-0.31}$ (syst) ± 0.10 (theory).

the Bayesian method we assume a uniform prior in the signal yield, a Gaussian prior in the θ_k, truncated so that no prediction is negative. The posterior density function is then integrated over the θ_k (including correlations).

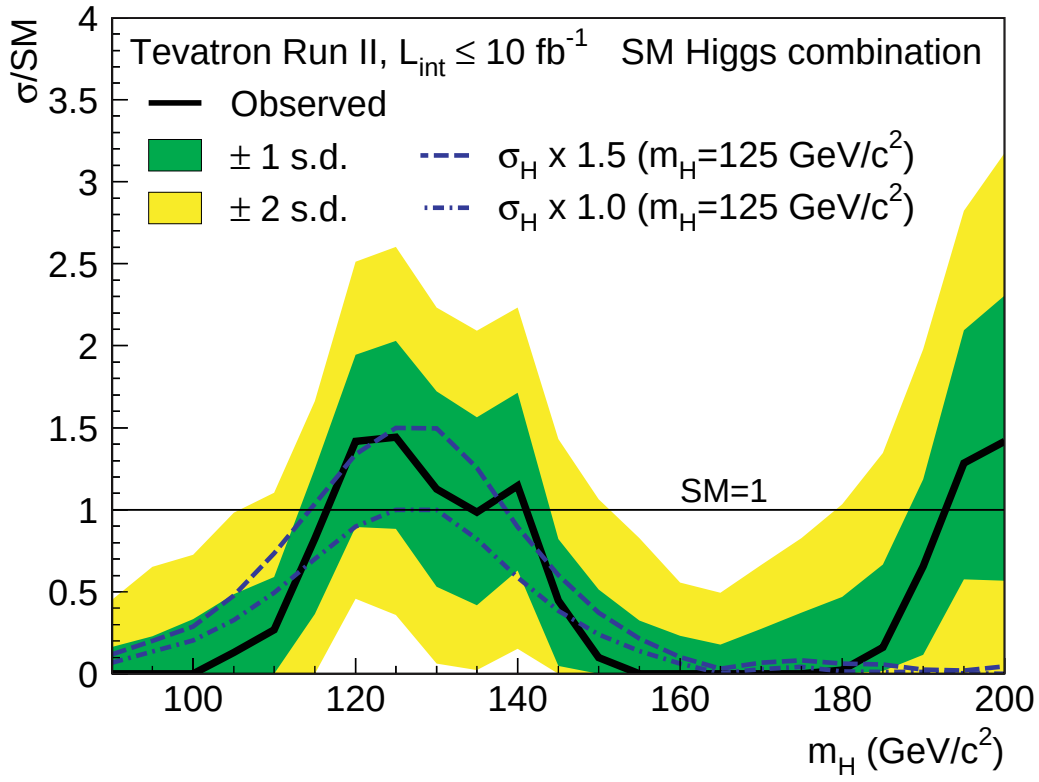


Figure 9.12: The best-fit signal cross section expressed as a ratio to the SM cross section as a function of Higgs boson mass for all of CDF and D0's SM Higgs boson searches in all decay modes combined. The dark- and light-shaded bands show the one and two s.d. uncertainty ranges on the fitted signal, respectively. Also shown with blue lines are the median fitted cross sections expected for a SM Higgs boson with $m_H = 125$ GeV at signal strengths of 1.0 times (short-dashed) and 1.5 times (long-dashed) the SM prediction.

We compute the significance of the excess in the data over the background prediction, at each hypothesized Higgs boson mass by calculating the local p-value under the background-only hypothesis using $R^{fit}_{profile}$, chosen a *priori*, as the test statistic. The observed local significance at $m_H = 125$ GeV corresponds to 3.0 standard deviations.

Since this thesis focus on the $WH \to \ell\nu b\bar{b}$, we are interested in the $H \to b\bar{b}$ combination at Tevatron. By combining the $H \to b\bar{b}$, $H \to W^+W^-$, $H \to \gamma\gamma$, and $H \to \tau^+\tau^-$ decay modes from both D0(described in previous Section) and CDF, the observed LLR distribution is shown in Figure 9.13, along with its expected values under the background-only and signal-plus-background hypotheses. The hypotheses that a SM Higgs boson is present with $m_H = 125$ GeV for signal strengths of 1.0 and 1.5 times the SM prediction are also given.

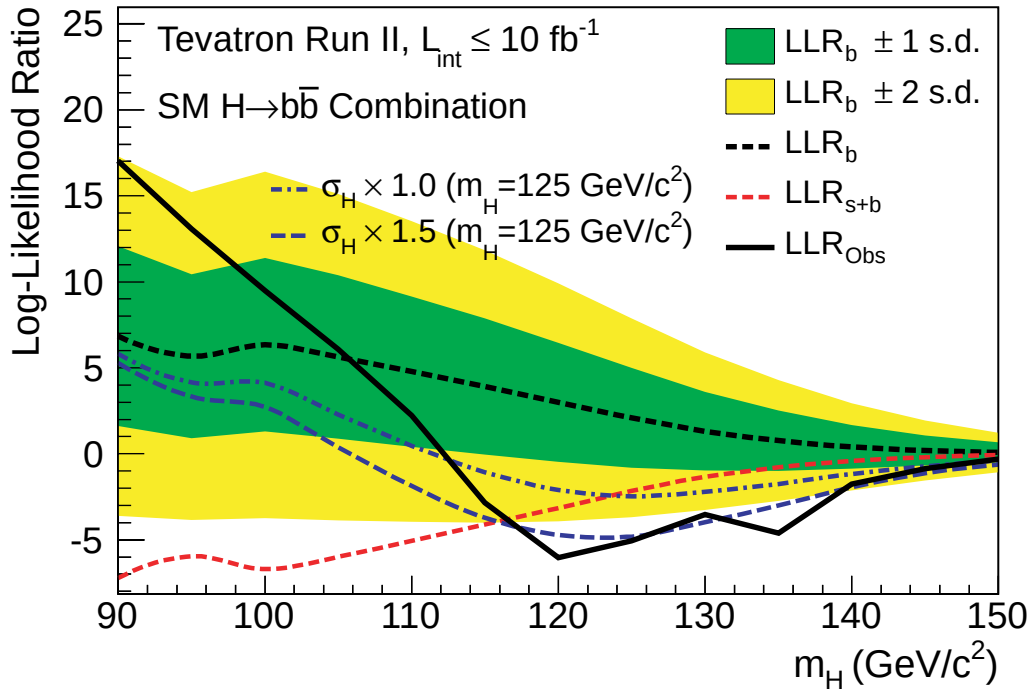


Figure 9.13: The log-likelihood ratio LLR as a function of Higgs boson mass from the combination of CDF and D0's SM Higgs boson searches focusing on the $H \to b\bar{b}$ decay mode. The solid line shows the observed LLR values, the dark short-dashed line shows the median expectation assuming no Higgs boson signal is present, and the dark- and light-shaded bands correspond, respectively, to the regions encompassing one and two s.d. fluctuations around the background-only expectation. The red long-dashed line shows the median expectation assuming a SM Higgs boson signal is present at each value of m_H in turn. The blue lines show the median expected LLR assuming the SM Higgs boson is present at $m_H = 125$ GeV with signal strengths of 1.0 times (shortdashed) and 1.5 times (long-dashed) the SM prediction.

We show the fitted $(\sigma_{WH} + \sigma_{WH}) \times \mathcal{B}(H \to b\bar{b})$ as a function of m_H, along with the SM prediction, in Figure 9.14. The figure also shows the expected cross section fits for each m_H, assuming that the SM Higgs boson with $m_H = 125$ GeV is present, both at the rate predicted by the SM, and also at a multiple of 1.5 times that of the SM. The best-fit rate corresponds to $(\sigma_{WH} + \sigma_{WH}) \times \mathcal{B}(H \to b\bar{b}) = 0.19^{+0.08}_{-0.09}$ (stat + syst) pb. For $m_H = 125$

GeV, the SM predicts $(\sigma_{WH} + \sigma_{WH}) \times \mathcal{B}(H \to b\bar{b}) = 0.12 \pm 0.01$ pb.

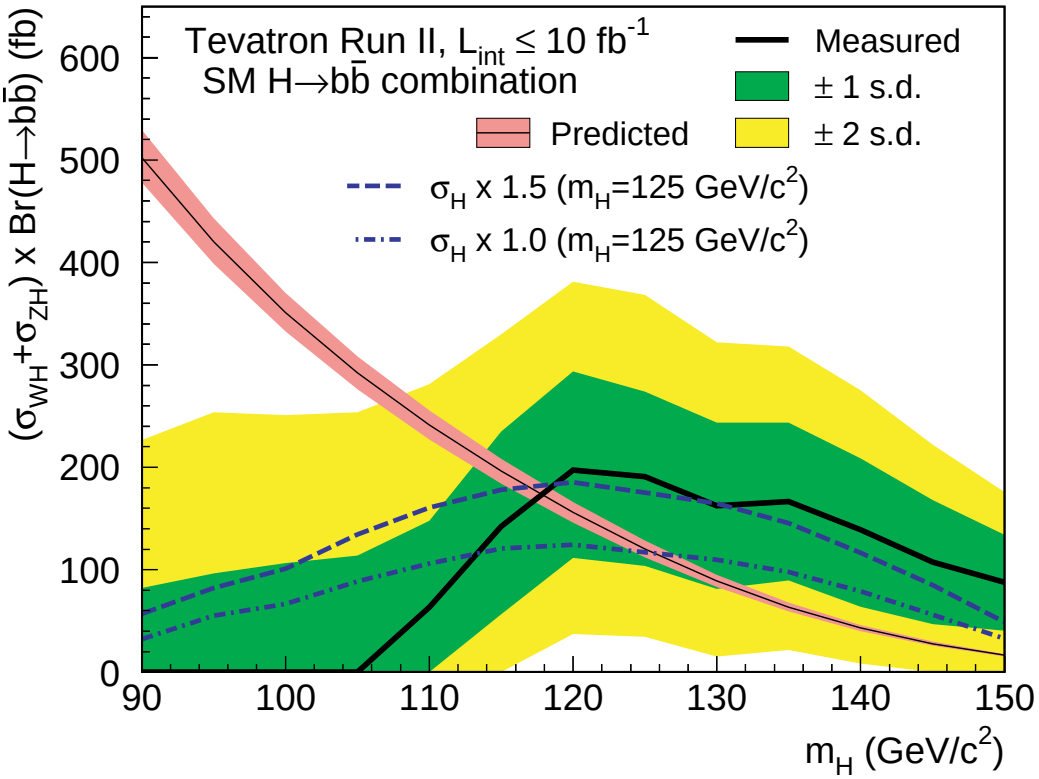


Figure 9.14: The best-fit signal cross section times branching ratio $(\sigma_{WH}+\sigma_{ZH})\times\mathcal{B}(H \to b\bar{b})$ as a function of Higgs boson mass from the combination of CDF and D0's SM Higgs boson searches focusing on the $H \to b\bar{b}$ decay mode. The dark- and light-shaded bands show the one and two s.d. uncertainty ranges on the fitted signal, respectively. Also shown with blue lines are the median fitted cross sections expected for a SM Higgs boson with $m_H = 125$ GeV at signal strengths of 1.0 times (short-dashed) and 1.5 times (long-dashed) the SM prediction. The SM prediction is shown as the smooth, falling curve where the narrow band indicates the theoretical uncertainty.

For $m_H = 125$ GeV, we obtain $R^{fit} = 1.44^{+0.59}_{-0.56}$ using all decay modes, and we obtain $1.59^{+0.69}_{-0.72}$ for the $H \to b\bar{b}$ decay mode.

Figure 9.15 shows the contribution of the four combinations for the different decay modes ($H \to W^+W^-$, $H \to b\bar{b}$, $H \to \gamma\gamma$, $H \to \tau^+\tau^-$) to the best-fit signal cross section for $m_H = 125$ GeV. The results are consistent with each other, with the full combination, and with the production of the SM Higgs boson at that mass.

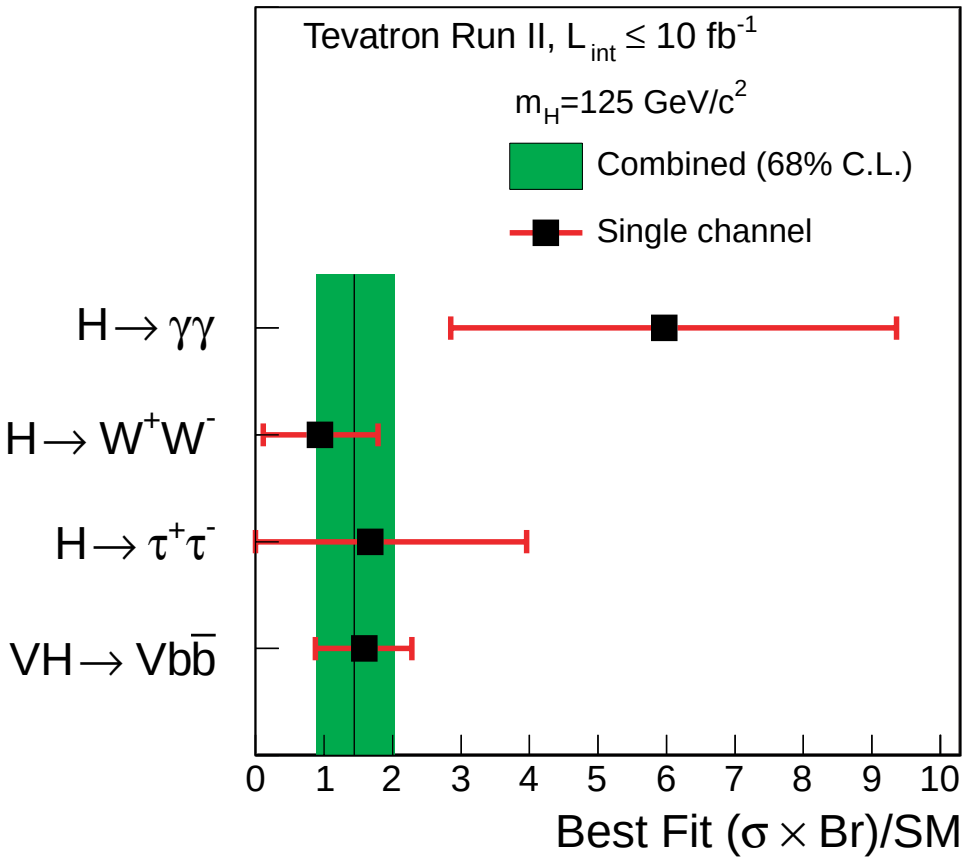


Figure 9.15: Best-fit values of $R = (\sigma \times \mathcal{B})/$ for the combinations of CDF and D0's Higgs boson search channels focusing on the $H \to W^+W$, $H \to b\bar{b}$, $H \to \gamma\gamma$, and $H \to \tau^+\tau^-$ decay modes for a Higgs boson mass of 125 GeV. The shaded band corresponds to the one s.d. uncertainty on the best-fit value of R for all SM Higgs boson decay modes combined.

Chapter 10

Summary

In this manuscript, I have presented the search for the Higgs boson in the $WH \to \ell\nu b\bar{b}$ at the D0 experiment, and my contributions to it. This is the most sensitive channel for low mass Higgs boson studies, and it is combined with ZH channels to search for $H \to b\bar{b}$ decays, first at D0, then, in combination with the equivalent channels analyzed by the CDF collaborations. These are then further combined with the other Higgs boson analysis channels of the Tevatron. Given their sensitivity, these results could also potentially be combined with the current ATLAS and CMS $H \to b\bar{b}$ analyses.

Most of the D0 detector parts are used to reconstruct WH events, since isolated electrons from W boson decay and the jets are reconstructed mainly from their characteristic signature in the calorimeter, isolated muons are detected in the muon spectrometer, and the neutrino from the W decay does not interact in the detector but its kinematics is reconstructed from the energy imbalance in the transverse plane. The tracking detectors are also heavily used to reconstruct precisely the momentum of the charged leptons and to identify b-jets. A good description of the event observables, hence of the background processes to this analysis, is achieved by correcting the objects from simulated events to account for reconstruction and identification differences in Data and Monte Carlo. I concentrated on different aspects of jet reconstruction, and in a future analyses one would try to improve the behaviour of jet reconstruction/identification at large pseudo-rapidity. The instrumental background (QCD or multijet) is derived from data, to accurately model the shapes and normalization of the different distributions, and I helped to test the optimization of the MVA_{MJ} for the electron channel.

To improve the sensitivity in the WH analysis, a major tool is the identification of jets coming from b quarks using displaced vertices consistent with the presence of b hadrons or multivariate analysis techniques to increase the discrimination between backgrounds and the WH signal. I improved the power of this algorithm by optimizing the set of b-tagging working points used in our analysis.

Last but not least, since the signal over background remained small after the above analyses steps, we concentrated on multivariate analysis (MVA) techniques to further improve our sensitivity. I thus improved our Boosted Decision Tree approach, making use of the best inputs our analysis group developed, in particular the output of a dedicated MVA to reject the instrumental background became one the most sensitive variables to improve our search.

Since a good agreement between data and simulation is obtained, and no signal excess is observed in data, upper observed (expected) limits (for $M_H = 125$ GeV) are set at 95%

confidence level on the ratio of the WH cross section multiplied by the branching ratio of $H \to b\bar{b}$ to its Standard Model prediction, which represent 4.8 (4.7) times the Standard Model expectation. This procedure is performed using a semi-frequentist approach in order to efficiently take into account systematic uncertainties, which are found to degrade the sensitivity by about 20% once taken into account and after minimizing their impact by constraining them to data.

The result presented in this thesis, based on an integrated luminosity of 9.7 fb^{-1} has been originally published in Physics Review Letter in September 2012 [6], and with some small modifications in Physics Review D in September 2013 [4]. It is included in the D0 and in the Tevatron combination.

The results of the $H \to b\bar{b}$ searches were validated through a measurement of the diboson ($WZ + ZZ$) production cross section using the same data samples and analysis techniques, treating those diboson processes as signal, and I produced the results of the electron channel for the WZ measurement. The resulting diboson cross-section measurement is in agreement with the SM prediction. The expected significance of the measurement is 1.8 standard deviations. We measure a cross section of 0.50 ± 0.34 (stat.) ± 0.36 (syst.) times the expected SM cross section of 4.4 ± 0.3 pb.

In the Higgs boson, search, when combining all CDF and D0 low mass channels, we observe a ~ 3 standard deviation excess of events in the mass range between 115 and 140 GeV. When adding all the channels, the local significance at $m_H = 125$ GeV corresponds to 3.0 standard deviations, with an expected significance of 1.9 standard deviations. The best-fit signal strength is $1.44^{+0.59}_{-0.56}$ times the SM expectation. The observed best-fit signal strengths obtained from each of the subprocess combinations are consistent with the expectations for a SM Higgs boson at $m_H = 125$ GeV.

Since five standard deviation significance is required for a discovery, the Tevatron could not achieve this goal for lack of statistics. Still, the Tevatron experiments achieved with this data sample a sensitivity comparable to early optimistic projections. [109] After ten years of gathering and analyzing data produced by the Tevatron collider, scientists from the CDF and D0 collaborations could announce evidence for fermionic decays simultaneously to the LHC Higgs boson discovery in the bosonic decay channels, while LHC experiments did not see any significant indication for fermionic decays until 2013. At the Tevatron, we saw the first evidence of associated production VH with Higgs boson decay in a pair of bottom quarks, a process which will remain difficult to observe at the future Run II of the LHC, even though it will have seven times the centre of mass energy of the Tevatron, and more than an order of magnitude higher integrated luminosity.

Bibliography

[1] G. Aad et al. (ATLAS Collaboration). *Phys. Lett. B*, 716:1, 2012.

[2] S. Chatrchyan et al. (CMS Collaboration). *Phys. Lett. B*, 716:30, 2012.

[3] T. Aaltonen et al. (CDF and D0 Collaborations). *Phys. Rev. Lett.*, 109:071804, 2012.

[4] V. M. Abazov et al. (D0 Collaboration). Combined search for the Higgs boson with the D0 experiment. *Phys. Rev. D*, 88:052011, 2013.

[5] D0 Collaboration) T. Aaltonen et al. (CDF Collaboration. Higgs boson studies at the Tevatron. *Phys. Rev. D*, 88:052014, 2013.

[6] V. M. Abazov et al. (D0 Collaboration). Search for the Standard Model Higgs Boson in Associated WH Production in 9.7fb^{-1} of $p\bar{p}$ Collisions with the D0 Detector. *Phys. Rev. Lett.*, 109:121804, 2012. http://prl.aps.org/abstract/PRL/v109/i12/e121804.

[7] V. M. Abazov et al. (D0 Collaboration). Search for the standard model Higgs boson in $\ell\nu$+jets final states in 9.7 fb^{-1} of $p\bar{p}$ collisions with the D0 detector. *Phys. Rev. D*, 88:052008, 2013. http://prd.aps.org/abstract/PRD/v88/i5/e052008.

[8] J. Beringer et al.(Particle Data Group). The Review of Particle Physics. *Phys. Rev. D*, 86(010001), 2012. http://pdg.lbl.gov/.

[9] Peter W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13:508509, Oct 1964. doi:10.1103/PhysRevLett.13.508.

[10] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.*, 13:321–323, Aug 1964. doi:10.1103/PhysRevLett.13.321.

[11] C. R. Hagen G. S. Guralnik and T. W. B. Kibble. Global conservation laws and massless particles. *Phys. Rev. Lett.*, 13:585–587, Nov 1964. doi:10.1103/PhysRevLett.13.585.

[12] http://www.nobelprize.org/nobel prizes/physics/.

[13] T. Aaltonen et al. (CDF and D0 Collaborations). 2012. arXiv:1204.0042.

[14] T. Aaltonen et al. (CDF and D0 Collaborations). *Phys. Rev. D*, 86:092003, 2012.

[15] The ALEPH, CDF, D0, DELPHI, L3, OPAL, and SLD Collaborations, the LEP Electroweak Working Group, the Tevatron Electroweak Working Group, and the SLD Electroweak and Heavy Flavour Working Groups, arXiv:1012.2367v2 (2011). The most recent values from March 2012, as quoted, are available from http://lepewwg.web.cern.ch/LEPEWWG/.

[16] L3 The ALEPH, DELPHI, OPAL Collaborations, and the LEP Working Group for Higgs Boson Searches. Search for the Standard Model Higgs boson at LEP. *Phys. Lett. B*, 565:61–75, 2003.

[17] G. Aad et al. (ATLAS Collaboration). *Phys. Lett. B*, 718:369, 2012.

[18] S. Chatrchyan et al. (CMS Collaboration). *Phys. Lett. B*, 710:284, 2012.

[19] The Tevatron New Phenomena and Higgs Working Group. 2012. CDF Note 10474.

[20] V.M. Abazov et al. Search for the Standard Model Higgs boson in the $ZH \to \nu\nu b\bar{b}$ channel in 5.2 fb^{-1} of $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ tev. *Phys. Rev. Lett.*, 104:071801, 2010. arXiv:0912.5285, doi:10.1103/PhysRevLett.104.071801.

[21] V.M. Abazov et al. Search for $ZH \to \ell\ell b\bar{b}$ production in 4.2 fb^{-1} of $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV. *Phys. Rev. Lett.*, 105:251801, 2010. arXiv:1008.3564,doi:10.1103/PhysRevLett.105.251801.

[22] Abdelhak Djouadi. The anatomy of electroweak symmetry breaking: Tome I: The Higgs boson in the Standard Model. *Physics Reports*, 457(1-4):1–216, February 2008.

[23] S. W. Herb et al. Observation of a dimuon resonance at 9.5 GeV in 400 GeV proton - nucleus collisions. *Phys. Rev. Lett.*, 39:252–255, 1977. doi:10.1103/PhysRevLett.39.252.

[24] F. Abe et al. Observation of top quark production in $p\bar{p}$ collisions. *Phys. Rev. Lett.*, 74:2626–2631, 1995. arXiv:hep-ex/9503002, doi:10.1103/PhysRevLett.74.2626.

[25] S. Abachi et al. Observation of the top quark. *Phys. Rev. Lett.*, 74:2632–2637, 1995. arXiv:hep-ex/9503003, doi:10.1103/PhysRevLett.74.2632.

[26] B. Baller et al. Direct observation of the tau neutrino. *Nucl. Phys. Proc. Suppl.*, 98:43–47, 2001. doi:10.1016/S0920-5632(01)01193-8.

[27] Accelerator - fermilab's tevatron. http://www.fnal.gov/pub/science/accelerator/.

[28] M. Abolins et al. Design and Implementation of the New D Level-1 Calorimeter Trigger. *Nucl. Instrum. and Methods, A*, 584.

[29] R. Angstadt et al. (D0 Collaboration). The Layer 0 Inner Silicon Detector of the D0 Experiment. *Nucl. Inst. Meth. A*, 622:298–310, 2010.

[30] http://www.fnal.gov/pub/today/archive/archive_2012/today12-08-21.html.

[31] C. M. Bhat and C. A. Rodriguez. Pbar Acceleration in the Main Injector for Run II: ESME Simulations. *Beams-doc-354-v1*, 2000.

[32] Available from: http://www-bd.fnal.gov/public/maininj.html.

[33] Gerald P. Jackson. Stochastic Cooling of the Recycler Antiproton Stack Momentum Spread Including Intrabeam Scattering and Barrier Buckets. *Beams-doc-172-v1*, 2002.

[34] S. Abachi et al. The D0 Detector. *Nucl. Instrum. Meth. A*, 338:185253, 1994. doi:10.1016/0168-9002(94)91312-9.

[35] V.M. Abazov et al. The Upgraded D0 detector. *Nucl. Instrum. Meth. A*, 565:463537, 2006. arXiv:physics/0507191, doi:10.1016/j.nima.2006.05.248.

[36] http://www-d0.fnal.gov/Run2Physics/cs/skimming/run2_overview.html.

[37] R. Angstadt, L. Bagby, A. Bean et al. The Layer 0 Inner Silicon Detector of the D0 Experiment. arXiv:0911.2522.

[38] M. Voutilainen. Jet p_T resolution for RunIIa final JES (v7.2) with dijet J4S jet corrections. D0 Note 5499.

[39] P. Gris. Electron Smearing Studies with RunIIa Data. D0 Note 5400.

[40] Hamamatsu Corporation.

[41] T. Edwards and others. The D0 Run IIb Luminosity Measurement. 2012. http://arxiv.org/pdf/1204.0461.pdf.

[42] Available from: http://www-d0.fnal.gov/Run2Physics/cs/caf.

[43] Alexander Khanov. Htf: histogramming method for finding tracks. the algorithm description. D0 Note 3778.

[44] G. Borissov. Ordering a Chaos or Technical Details of AA Tracking. Presentation at the All D0 Meeting, 28 February 2003.

[45] H. Greenlee. The D0 Kalman Track Fit. D0 Note 4303.

[46] R. Fruhwirth. Application of Kalman filtering to track and vertex fitting. *Nucl. Instrum. Meth. A*, 262:444450, 1987. doi:10.1016/0168-9002(87)90887-4.

[47] D. Bandurin et al. Electron and Photon Identification with p20 data. D0 Note 5761, 2008.

[48] X. Bu et al. A Multivariate based Electron ID. 2011. D0 note 6238.

[49] Masato Aoki. Electron Likelihood in p20. D0 Note 5675.

[50] D0 Collaboration. Electron and Photon Identification in the D0 Experiment. 2013. arXiv:1401.0029.

[51] D0 Collaboration. Muon reconstruction and identification with the Run II D0 detector. arXiv:1307.5202.

[52] E. Busato and B. Andrieu. Jet Algorithms in D0 RunII Software: Description and User's Guide. D0 Note 4457.

[53] Amnon Harel and Jiri Kvita. 2008. D0 Note 5634.

[54] http://www-d0.fnal.gov/Run2Physics/cs/MC/MC.html.

[55] T. Sjstrand et al. PYTHIA 6.3: Physics and manual. 2003. hep-ph/0308153, LU-TP-03-38, Lund.

[56] M. Rangel B. Casey, G. Facini. Jet ID/VC Efficiencies and Scale Factors using Dijet events. 2010. D0 Note 6058.

[57] Dikai Li Xuebing Bu and Jeongku Lim. Jet identification for Reprocessed IIb data. 2012. D0 note 6356.

[58] B. Abbott et al. Determination of the absolute jet energy scale in the D0 calorimeters. *Nucl. Instrum. Meth. A*, 424:352394, 1999. arXiv:hep-ex/9805009, doi:10.1016/S0168-9002(98)01368-0.

[59] D0 Collaboration. Improved b quark jet identification at the D0 experiment. 2013. arXiv:1312.7623v1.

[60] V. M. Abazov et al. *Nucl. Instrum. Methods Phys. Res., Sect. A*, 620:490, 2010.

[61] V.M. Abazov et al. *Phys. Rev. Lett.*, page 091802, 2005.

[62] V.M. Abazov et al. *Phys. Rev. Lett.*, page 051803, 2009.

[63] V. M. Abazov et al., (D0 Collaboration). Search for WH associated production in 5.3 fb^{-1} of $p\bar{p}$ collisions at the Fermilab Tevatron. *Physics Letters B*, 698:6–13, 2011.

[64] V. M. Abazov et al. (D0 Collaboration). Search for WH associated production in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$TeV. *Phys. Rev. D*, 86:032005, 2012. arXiv:1203.1082.

[65] H. L. Lai et al. *Phys. Rev. D*, 55:1280, 1997. hep-ph/9606399.

[66] J. Pumplin et al. New generation of parton distributions with uncertainties from global qcd analysis. *J. High Energy Phys.*, 07:012, 2002. arXiv:hep-ph/0201195v3.

[67] Fulvio Piccinini Roberto Pittau Michelangelo L. Mangano, Mauro Moretti and Antonio D. Polosa. ALPGEN, a generator for hard multiparton processes in hadronic collisions. *J. High Energy Phys.*, 07:001, 2003. arXiv:hep-ph/0206293.

[68] L. Christofek M. Begel and E. Shabalina. Determination of weighting factors for ALPGEN Monte Carlo signal and background samples. D0 Note 5016.

[69] E. Boos et al. CompHEP. *Nucl. Instrum. Methods Phys. Res. A*, 534:250.

[70] Y. Fisyak and J. Womersley. D0gstar D0 GEANT Simulation of the Total Apparatus Response. 1997. D0 Note 3191.

[71] R. Brun and F. Carminati. GEANT Detector Description and Simulation Tool, CERN Program Library Long Writeup W5013. unpublished, 1993.

[72] A. Abbinante et al. Search for Higgs Boson production in lepton plus jets final states with 9.7 fb^1 of RunII data. 2012. D0 Note 6334.

[73] (D0 Collaboration) V. M. Abazov et al. *Phys. Rev. D*, 76:012003, 2007.

[74] TEVNPH (Tevatron New Phenomena and Higgs Working Group). arXiv:1203.3774.

[75] J. Baglio and A. Djouadi. *J. High Energy Phys.*, 10:064, 2010. 1003.4266v2.

[76] S.-O. Moch P. Bolzoni, F. Maltoni and M. Zaro. *Phys. Rev. D*, 85:035002, 2012. 1109.3717.

[77] D. de Florian and M. Grazzini. *Phys. Lett. B*, 674:291, 2009. 0901.2427.

[78] The azimuthal angle, ϕ, is defined as the opening angle with respect to the x direction in a right-handed coordinate system defined by y as up and z as the protonbeam direction.

[79] V. M. Abazov et al. (D0 Collaboration). *Phys. Rev. D*, 88:052009, 2013.

[80] Romain Madar and Fabrice Couderc. Tau lepton identification using MVA optimizations, preshower and bID tools. D0 note 6061.

[81] V. M. Abazov et al. (D0 Collaboration). Search for $ZH \to \ell^+\ell^- b\bar{b}$ production in 9.7fb^{-1} of $p\bar{p}$ collisions with the D0 detector. *Phys. Rev. D*, 88:052010, 2013. arXiv:1303.3276.

[82] Aurelio Juste Alan Magerkurth Jianming Qian Jonas Strandberg Chun Xu Thomas Gadfort, Dale Johnston. Search for WH Production Using The Matrix Element Analysis Technique in 900 pb-1 of Data Collected with the D0 Detector. 2008. D0 Note 5365.

[83] Andreas Hoecker, Peter Speckmayer, Joerg Stelzer, Jan Therhaag, Eckhard von Toerne, and Helge Voss. TMVA: Toolkit for Multivariate Data Analysis. *PoS*, ACAT:040, 2007. We use version 4.1.0.

[84] L. Breiman. Random forests. *Mach. Learn.*, 45:5–32, 2001.

[85] I. Narsky. Optimization of Signal Significance by Bagging Decision Trees. arXiv:physics/0507157v1, 2005.

[86] I. Narsky. StatPatternRecognition: A C++ Package for Statistical Analysis of High Energy Physics Data. physics/0507143, 2005.

[87] http://tmva.sourceforge.net/docu/TMVAUsersGuide.pdf.

[88] *Kolmogorov-Smirnov Goodness-of-Fit Test.* http://www.itl.nist.gov/div898/handbook/eda/section3/

[89] R. Olshen L. Breiman, J. Friedman and C. Stone. *Classification and Regression Trees.* Wadsworth and Brooks/Cole Advanced Books and Software, Pacific Grove, CA, 1984.

[90] S. Parke and S. Veseli. *Phys. Rev. D*, 60:093003, 1999.

[91] D0 Collaboration. 2012. D0 Note 6309.

[92] J. M. Campbell and R. K. Ellis. *Phys. Rev. D*, 60:113006, 1999.

[93] J. M. Campbell. 2001. arXiv:hep-ph/0105226.

[94] R. K. Ellis J. M. Campbell and C. Williams. MCFM - Monte Carlo for FeMtobarn processes. http://mcfm.fnal.gov/.

[95] Konstantinos Petridis Maiko Takahashi Xuebing Bu, Tim Head and Yunhe Xie. Electron Identification for Summer 2010. D0 note 6116.

[96] G. Hesketh V. Lesne M. Owen R. Stroehmer V. Sharyy P. Calfayan, T. Gadford and B. Tuchming. Muon identification certification for p17 data. D0 Note 5157.

[97] O. Brandt et al. Muon Identication Certication for the Summer 2009 Extended Dataset (Run IIb-1 and -2). D0 Note 6025.

[98] J. Brown M. Cooke Y. Enari S. Greder K. Herner D. Li F. Miconi B. Penning G. Bernardi, D. Brown. D0 note 6213.

[99] J. Haley. D. Price W. Fisher. Studies of alpgen parameters, corrections and associated uncertainties. 2009. D0 note 5966.

[100] Wade Fisher. Collie: A Confidence Level Limit Evaluator. D0 Note 5595.

[101] A. L. Read. Presentation of search results: The CL(s) technique. *J. Phys., G*, 28:2693, 2002.

[102] T. Junk. *Nucl. Instrum. Methods Phys. Res., Sect. A*, 434:435, 1999.

[103] W. Fisher. Systematics and Limit Calculations. CollieChi2Formulism.pdf from the Collie Package. Report No. FERMILAB-TM-2386-E, 2007.

[104] [D0 Collaboration] V. M. Abazov et al. *Phys. Rev. Lett*, 109:121803, 2012.

[105] [D0 Collaboration] V. M. Abazov et al. *Phys. Lett. B*, 716:285, 2012.

[106] [D0 Collaboration] V. M. Abazov et al. *Phys. Lett. B*, 714:237, 2012.

[107] [D0 Collaboration] V. M. Abazov et al. *Phys. Rev. D*, 88:052005, 2013.

[108] [D0 Collaboration] V. M. Abazov et al. *Phys. Rev. D*, 88:052007, 2013.

[109] H. E. Haber J. D. Hobbs et al M. Carena, J. S. Conway. Report of the higgs working group of the tevatron run 2 susy/higgs workshop. arXiv:hep-ph/0010338.