# STUDY OF SINGLE TOP QUARK PRODUCTION USING BAYESIAN NEURAL NETWORKS WITH DØ DETECTOR AT THE TEVATRON

**A THESIS**

**Submitted to the**

**FACULTY OF SCIENCE**

**PANJAB UNIVERSITY, CHANDIGARH**

**for the degree of**

DOCTOR OF PHILOSOPHY

**2012**


JYOTI JOSHI

DEPARTMENT OF PHYSICS

CENTRE OF ADVANCED STUDY IN PHYSICS

PANJAB UNIVERSITY, CHANDIGARH

INDIA

.....

*Dedicated to Bhagwaan ji and my family*

.....

# Acknowledgments

*Finally, the time has come when I can thank and acknowledge all the people who helped me to make this thesis possible. First and foremost, I am deeply grateful to my supervisor, Prof. Suman B. Beri for her patient guidance, encouragement and advice. I consider myself very fortunate for being able to work with a very considerate and encouraging professor like her. Without her enlightening suggestions and offering to accomplish this research, I would not have been able to finish my research both at Fermilab and Panjab University. I was continually amazed by her willingness to proof read my thesis numerous times. A thesis of this nature which was based on two continents, is only possible with the help of her enormous and careful editing. During my initial stay at Fermilab, her invaluable cooking techniques in addition to the multivariate physics analysis techniques gave me a life time experience.*

*I am thankful to Chairman, Department of Physics, Panjab University for providing adequate facilities to work in the department. Many thanks and regards go in particular to Prof. J. M. Kohli, Prof. J. B. Singh, Prof. M. Kaur and Dr. V. Bhatnagar for their continuous support and efforts to provide us adequate facility in PUHEP lab. I would also like to thank the technical, computing and purchase section staff at Panjab University.*

*Special thanks go to the previous and current conveners of single top working group at DØ: Prof. Cecilia Gerber, for her continuous guidance, support*

*and caring while being at Fermilab and helping me to shape up my thesis; Dr. Reinhard Schwienhorst, for helping in all the stastical issues of the analysis and always being ready to answer my questions; Dr. Ann Heinson, for bringing experience to the group and for taking care of the details during the analysis; Dr. Victor Bazterra and Dr. Liang Li for all the encouragement and work that has lead this analysis possible. I am thankful to my fellow grad students Nathan, Mark, Weigang and Yun-Tse for always being more than willing to help and the hard work all of you have put into the group. Many thanks to Prof. Harrison B. Prosper and Dr. Supriya Jain for helping me to understand the details of bayesian neural networks and always helping me out when I was stuck. Working with all of you is the most learning and enjoyable experience of my life. Also, I owe my thanks to DØ spokespersons specially to Dr. Dmirti Denisov and all the members of DØ collaboration including physicists, engineers and technicians who worked hard, day and night, to make it successful.*

*I am deeply thankful to my friends and fellow grad students at Panjab University; Nishu, Archana di, Manuk, Monika, Manie, Suneel, Sunil, Anil, Lovedeep, Bhawan, Kanishka, Sandy, Sumali, Prabhdeep, Inderpal, Nitish, Ritu and Ruchi who were always present for helping me out from all kind of problems with their suggestions, discussions and creating enjoyable environment in the lab. I would especially like to thank all my Indian friends whom I have met at Fermilab during my two and a half year of stay, Sonam, Richa, Ruchika, Dileep, Pratima di, Suvadeep, Shilpee and Amit. Special thanks go to Tim Uncle for all the spiritual, emotional and practical support. Friends are a great asset to life and this becomes true while speaking about all of you.*

*Last but not the least, I would like to thank Almighty God for blessing me with very understanding and caring family. Thank you Mom and Dad,*

*for all the unconditional love and encouragement. This thesis would have been simply impossible without your support. My heartiest regards to my Grandma for showering her warm blessings on me. I simply thank my siblings Ritu, Varinder and Karan for their constant support in all phases of my life. Special thanks go to Tayaji and Tayiji for their enormous love and continuous support.*

*This research work has been supported by Department of Science and Technology (DST) and the support from Fermilab. I am thankful to all of them for their confidence in me.*

Dated: Jyoti Joshi

Chandigarh

# Abstract

Top quark, the heaviest and most intriguing among the six known quarks, can be created via two independent production mechanisms in $p\bar{p}$ collisions. The primary mode, strong $t\bar{t}$ pair production from a $gtt$ vertex, was used by the DØ and CDF collaborations to establish the existence of the top quark in March 1995. The second mode is the electroweak production of a single top quark or antiquark, which has been observed recently in March 2009. Since single top quarks are produced at hadron colliders through a $Wtb$ vertex, thereby provide a direct probe of the nature of $Wtb$ coupling and of the Cabibbo-Kobayashi-Maskawa matrix element, $V_{tb}$. So this mechanism provides a sensitive probe for several, standard model and beyond standard model, parameters such as anomalous $Wtb$ couplings.

In this thesis, we measure the cross section of the electroweak produced top quark in three different production modes, $s+t$, $s$ and $t$-channels using a technique based on the Bayesian neural networks. This technique is applied for analysis of the 5.4 $fb^{-1}$ of data collected by the DØ detector. From a comparison of the Bayesian neural networks discriminants between data and the signal-background model using Bayesian statistics, the cross sections of the top quark produced through the electroweak mechanism have been measured as:

$$\sigma(p\bar{p} \to tb + X, tqb + X) = 3.11^{+0.77}_{-0.71} \text{ pb}$$

$$\sigma(p\bar{p} \to tb + X) = 0.72^{+0.44}_{-0.43} \text{ pb}$$

$$\sigma(p\bar{p} \to tqb + X) = 2.92^{+0.87}_{-0.73} \text{ pb}$$

The $s$+$t$-channel has a gaussian significance of $4.7\sigma$, the $s$-channel $0.9\sigma$ and the $t$-channel $4.7\sigma$. The results are consistent with the standard model predictions within one standard deviation. By combining these results with the results for two other analyses (using different MVA techniques) improved results

$$\sigma(p\bar{p} \to tb + X, tqb + X) = 3.43^{+0.73}_{-0.74} \text{ pb}$$

$$\sigma(p\bar{p} \to tb + X) = 0.68^{+0.38}_{-0.35} \text{ pb}$$

$$\sigma(p\bar{p} \to tqb + X) = 2.86^{+0.69}_{-0.63} \text{ pb}$$

were obtained with a significance of $5.4\sigma$, $1.8\sigma$ and $5.0\sigma$ respectively for $s$+$t$, $s$ and $t$-channels. Using this measured cross section and constraining $0 \leq |V_{tb}|^2 \leq 1$, the lower limit has been calculated to be $|V_{tb}| > 0.79$ with 95% confidence level (C.L.). Another measurement of $tqb$ production cross section is done using the same dataset and discriminant but without any assumption on the $tb$ production rate. From this measurement, we obtain a cross section of $2.90^{+0.59}_{-0.59}$ pb for $t$-channel and corresponding significance of $5.5\sigma$.

In addition to the above mentioned work, a search is made for the anomalous $Wtb$ couplings in single top quark production. Within the Standard Model, the $Wtb$ vertex is purely left-handed, and its amplitude is given by the $V_{tb}$, related to weak interaction between a top and a $b$-quark. In a more general way, additional anomalous couplings such as right-handed vectorial couplings and left and right-handed tensorial couplings can also be considered. An analysis based on the Bayesian neural networks method is used to separate the signal from expected backgrounds. We find no evidence for anomalous couplings and set 95% C.L. limits on these couplings as $|V_{tb} \cdot f_{L_T}|^2 < 0.06$, $|V_{tb} \cdot f_{R_V}|^2 < 0.93$ and $|V_{tb} \cdot f_{R_T}|^2 < 0.13$. This result represents the most stringent direct constraints on anomalous $Wtb$ interactions.

This work has been done in collaboration with DØ experiment but the analyses and results presented in this thesis are my contribution.

# Contents

# List of Tables

# List of Figures

# Chapter 1

# Introduction

***Reason, Observation, and Experience - the Holy Trinity of Science.***

- Robert G. Ingersoll.

Science is the systematic study of the nature and behavior of the material and physical universe, based on observation, experiment, and measurement, and the formulation of laws to describe these facts in general terms. At the most basic level known, where things cannot be further subdivided, the science is particle physics. High energy particle physics concerns itself with the basic constituents of matter at the most fundamental level - the subatomic particles from which all matter in the known universe is made. The laws governing these particles are gathered in a Quantum Field Theory (QFT) known as the Standard Model (SM) of particles and fields. These laws may not be the final statement and there may be physics beyond the SM, which is expected to be observed at the very highest energies. The subject of particle physics deals with the detection and measurement of the properties and behavior of these fundamental particles, along with the four apparently distinct forces (electromagnetic, strong, weak, and gravity) by which they are known to interact. Out of these four interactions, gravity is ignored in particle physics.

To study these fundamental particles and forces, huge particle accelerators (machines capable of accelerating subatomic particles and colliding them at very high energies) are

required. The principle method of study is to cause particles to collide at the highest possible energies using accelerators and to observe the particle interactions with special detectors. With higher energies, smaller and smaller length scales are revealed. The *Tevatron* collider, which came online in 1983, accelerates particles to near light speed on a six-kilometer racetrack before steering them into head-on collisions. At those energies, the debris from collisions can contain fleeting particles never before observed by physicists. In its 28 years of operation the Tevatron made countless contributions to particle physics. Its importance is most tangibly demonstrated by the 1995 discovery of the top quark, one of the last missing pieces of the SM. The top quark is the heaviest fundamental particle ever observed. It is 100,00 times heavier than the lightest quark  the up quark  and has about the same mass as a gold atom, though it is much smaller. Its mass is the reason of why it was the last quark to be discovered: the more massive a particle, the more energy is required to create it in a particle accelerator. Before the Large Hadron Collider, the Tevatron was the only accelerator in the world capable of reaching energies high enough to make the top quark.

This thesis provides an experimental test to the SM, namely through the top quark, and also sets constraints on theories beyond the SM. At the Tevatron the top quark is primarily studied when a top ($t$) and antitop ($\bar{t}$) quark are produced together; this process occurs for roughly one out of every five billion $p\bar{p}$ collisions. The SM also predicts that the top quark can be produced singly at the Tevatron; however, it is expected to occur at half the rate of top-antitop production, or in nearly one out of every ten billion collisions. The present study of this thesis deals in particular with (i) the electroweak production of top quarks, called single top quarks, where one can test theoretical assumptions directly such as that the SM has three generations and that the top quark can be produced through the electroweak force much like the other quarks and (ii) the top quark produced through this channel can be used as a probe into physics beyond the SM, particularly anomalous couplings, if any discrepancies arise from the experimental verification of this SM cross section. So these two aspects: (i) measurement of the single top quark production cross section and (ii) search

for anomalous couplings in single top quark at Tevatron using DØ have been studied and presented in this thesis.

The outline if the thesis is as follows. Chapter 2 of this thesis gives an introduction to the Standard Model of particle physics. Also described in this chapter is the motivation for studying single top quark production and the production of the various background contributions. Chapter 3 describes the experimental facilities of the Tevatron and the *DØ* detector. It is not possible to directly measure quantities that are physically significant. Instead, there are large arrays of digitized signals which must be converted into the physically meaningful quantities. This is called event reconstruction and object identification which will be discussed in Chapter 4. The modeling of single top quark production and the background components, the event selection applied and the systematic uncertainties of the background estimation are discussed in Chapter 5. Single top quark production is a rare process relative to its backgrounds. This analysis uses Bayesian Neural Networks (BNN) as a multivariate method to separate single top quark events from the background events which will be described in Chapter 6. The following two chapters 7 and 8 describes the two analyses performed: measurement of single top quark production cross section and the search for anomalous $Wtb$ couplings. Chapter 7 discusses how the bayesian neural networks are created and applied to the dataset, and how the cross section and the signal significance are determined using the BNN output. Chapter 8 gives the theoretical background of the anomalous $Wtb$ couplings. It also gives a simple overview of how the monte carlo samples are generated for different anomalous couplings and looks at the multivariate technique of BNN that is used to derive limits on the anomalous couplings. Finally, Chapter 9 summarizes both the analyses and conclusions of the analyses.

# Chapter 2

# Top Quarks within the Standard Model

***Our goal in physics is to see things simply, to understand a great many complicated phenomena in a unified way, in terms of a few simple principles.*** - Steven Weinberg.

## 2.1 The Standard Model

The current model of sub-atomic structure used to understand matter is known as the Standard Model of particles and fields (SM) [1, 2, 3, 4, 5, 6, 7, 8]. Development of this model began in the late 1960s, and has continued through today with contributions from many scientists across the world. Although the SM itself is a very complicated theory, the basic structure of the model is fairly straightforward. The SM attempts to explain everything in the universe in terms of fundamental particles. A fundamental particle is one which cannot be broken down into anything else. These fundamental particles are the building blocks of matter, and the things which hold matter together.

As almost everybody knows, matter is made from atoms, which stick together to form molecules. Simple substances such as air and water are made from simple molecules con-

taining only two or three atoms. More complicated substances, such as mammoths and scientists, are made from more complicated molecules such as proteins and DNA, and contain millions of atoms, which stick together to make cells, tissue, fur and brains. There are around a hundred different types of atoms (known as elements), from hydrogen to uranium, catalogued by chemists in the periodic table.

But what are atoms made from ? Although for a long time thought to be fundamental, indivisible particles, early 20th Century studies showed that atoms were made up of smaller particles: protons, neutrons and electrons. Protons and neutrons stick together in a tiny point called the nucleus, which is orbited by electrons. The hydrogen atom is a single proton orbited by a single electron, whereas a uranium atom has nearly a hundred protons, neutrons and electrons.

Electrons readily jump from atom to atom, allowing all sorts of interesting chemistry and electronics. However, the protons and neutrons remain stuck together in the nucleus except in nuclear reactions. Nuclear fission occurs when an unstable heavy nucleus (such as uranium) splits into two; this process drives nuclear power stations and atomic weapons. Nuclear fusion occurs when two light nuclei fuse together to form a heavier atom; this energy fuels stars. Unstable or radioactive nuclei may also decay and emit particles due to various processes in the nucleus.

**But what are protons, neutrons and electrons made from ?**

To answer this question we leave the world of biologists, chemists, and atomic and nuclear physicists, and enter the realm of particle physicists and the standard model of particle physics. From the work of many, a brief review of the Standard Model is given here.

### 2.1.1 The Zoo of Particles

There are two types of elementary particles. The first are the basic building blocks of matter and the second are the particles which generate interactions. The "matter" particles are fermions and fall into two classes: quarks and leptons. Both quarks and leptons are spin 1/2

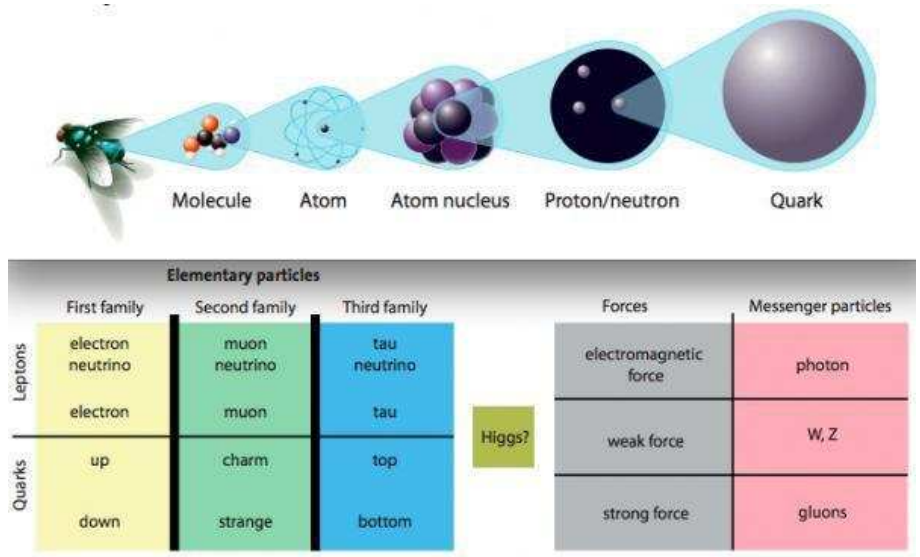


Figure 2.1: Electrons and quarks are the smallest building blocks of all matter. The Standard Model comprises the fundamental building blocks of matter and three of the four fundamental forces. While all known matter is built with particles from the first family, the other particles exists but only for extremely short time periods. To complete the Model a new particle is needed – the Higgs particle.

particles, and therefore obey Fermi-Dirac statistics. As indicated in Figure 2.1, quark and leptons are each arranged in three generations, containing particles of similar properties but differing in mass. For each particle there exists an associated anti-particles.

There are six different flavors of quarks, labeled (in order of increasing mass) *up*, *down*, *strange*, *charm*, *bottom* and *top* as given in Table 2.1 [9] with their basic properties. Quarks carry fractional electrical charges of + 2/3e or - 1/3e, with "e" the charge of the electron. Quarks are never observed as single particles but form bound states called *hadrons* of either three quarks to form *baryons*, or by pairing a quark with an antiquark to form *mesons*. Protons (made up of two up-quarks and one down-quark) and neutrons (made up of two down-quarks and one up-quark) are the most common examples of baryons. Pions ($\pi^0$, $\pi^{\pm}$) and kaons ($K^{\pm}$, $K^0$, $\bar{K}^0$) are the most common types of mesons.

There are three different flavors of charged leptons carrying a charge of - 1: *electron* ($e^-$), *muon* ($\mu^-$) and *tau* ($\tau^-$). While electrons exists in all atoms, muons and taus can only be observed in energetic processes like cosmic ray showers, or in high energy particle collisions.

There are three neutral leptons, called *neutrinos* ($\nu$), each associated with a charged lepton: $\nu_e$, $\nu_\mu$ and $\nu_\tau$. Neutrinos interact extraordinarily weakly with matter and their masses are negligibly small.

| Gen | Leptons (spin=$\frac{1}{2}$ ) | | | Quarks (spin=$\frac{1}{2}$ ) | | |
|---|---|---|---|---|---|---|
| | Flavors | Charge | Mass (MeV) | Flavors | Charge | Mass (MeV) |
| 1 | electron $e$ | $-1$ | 0.511 | up $u$ | $+2/3$ | $1.7 - 3.1$ |
| | $e$ neutrino $\nu_e$ | 0 | $< 2 \times 10^{-6}$ | down $d$ | $-1/3$ | $4.1 - 5.7$ |
| 2 | muon $\mu$ | $-1$ | 105.7 | charm $c$ | $+2/3$ | $1.29 \times 10^3$ |
| | $\mu$ neutrino $\nu_\mu$ | 0 | $< 0.19$ | strange $s$ | $-1/3$ | 100 |
| 3 | tau $\tau$ | $-1$ | 1777 | top $t$ | $+2/3$ | $(173.3 \pm 1.1) \times 10^3$ |
| | $\tau$ neutrino $\nu_\tau$ | 0 | $< 18.2$ | bottom $b$ | $-1/3$ | $4.19 \times 10^3$ |

Table 2.1: The fundamental particles and their basic properties.

### 2.1.2 The Particle Interactions

The great success of the SM is that it explains how fundamental particles interact with each other. Particles interact by four fundamental interactions: gravitational, electromagnetic, strong and weak. All the interactions work by exchanging an additional force particle between the particles which experience the force. For example, the electromagnetic force is mediated by the photon, and an electron can emit a photon which is then absorbed by another electron. Figure 2.2 shows the drawing of four SM interactions and their mediating particles.

The *gravitational interaction*, is the familiar attractive force between all particles having mass. This is the interaction that makes objects fall to the ground but also governs the movements of planets and galaxies. It is mediated by the graviton which is a massless boson with spin 2. Although gravity is the most obvious force in everyday life, it is not explained by the standard model. This is not a problem when studying microscopic particles, however, as it is incredibly weak compared to the other forces.

The *electromagnetic interaction* occurs between all charged particles with electric charge, and is mediated by the massless photon. The electromagnetic interaction is responsible

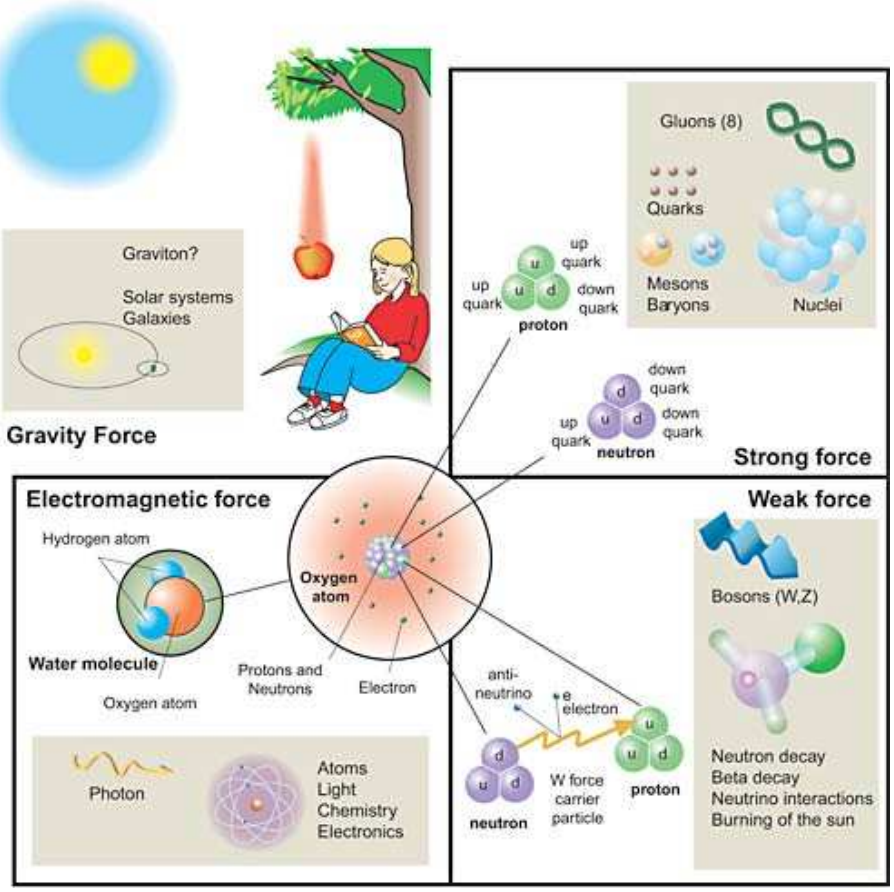


Figure 2.2: The four forces (or interactions) of Nature, their force carrying particles and the phenomena or particles affected by them. The three interactions that govern the microcosmos are all much stronger than gravity and have been unified through the Standard Model.

for a number of common phenomena in the world that surrounds us, such as friction and magnetism. It causes electrons to stick to atoms, and determines the behavior of atoms and molecules. Although electromagnetism has an infinite range and is much stronger than gravity, it plays a smaller role at large scales as it can be both attractive and repulsive.

The *strong interaction* is what holds the nucleus of an atom together. It is mediated by gluons and allows color-charged particles to interact with each other. As the nucleus contains only protons (with positive charge) and neutrons (with zero charge), the electromagnetic force would cause it to fly apart. However, the much stronger strong force holds them together. Unlike gravity and electromagnetism, the strong force has a very short range, so it plays no role in larger scale objects. Another unique property of the strong force between two particles is that its strength increases as they move apart (while they are still within its range). In contrast the strength of the electromagnetic or gravitational force decreases with distance. This explains one of the properties of quarks: they are only ever seen within another particle (eg, within protons, neutrons or pions). In fact, nobody has ever seen a lone quark or managed to isolate one. This is because as we try to separate two quarks, the force pulling them together increases, so an ever-increasing amount of energy is needed to continue pulling them apart. Eventually there is enough energy present to create a quark-antiquark pair out of the vacuum, and the two original quarks separate - now as quark-antiquark particles.

The *weak interaction* [5] is carried by the bosons, $W^{\pm}$ and $Z^0$, particles that, unlike the photon and the graviton, have very large masses (approximately 100 proton masses!) [10]. That is why the interaction has a short range. It acts on both quarks and leptons and is responsible for some radioactive decays. The weak force allows interactions amongst all the particles except gluons. It is closely related to the electromagnetic interaction and the two interactions are said to be united in the electroweak interaction, which was elucidated in the 1970s. Gerardus t Hooft and Martinus Veltman received the 1999 Nobel Prize for the final formulation of this theory.

Table 2.2 [9] shows in detail the force mediators and their basic properties.

| Force | Gauge Boson | Charge | Spin | Mass (GeV/c$^2$) | Range | Rel. Strength |
|---|---|---|---|---|---|---|
| Strong | Gluon (g) | 0 | 1 | 0 | $10^{-15}m$ | 1 |
| EM | Photon($\gamma$) | 0 | 1 | 0 | $\infty$ | 1/137 |
| Weak | $W^{\pm}$ | $\pm1$ | 1 | $80.423 \pm 0.039$ | $10^{-18}m$ | $10^{-5}$ |
| | $Z^0$ | 0 | 1 | $91.188 \pm 0.002$ | | |
| Gravity | Graviton (G) | 0 | 2 | 0 | $\infty$ | $10^{-38}$ |

Table 2.2: The force mediators and their properties.

### 2.1.3 The Gauge Theories

Gauge theories describe physical systems based on the idea that there is a symmetry governing their actions, and that transformations by this symmetry can be performed both locally and globally. These symmetries, such as charge, parity, or time, in combination with quantization of a theory, have led to the development of quantum field theories (QFT) [11, 12].

In the late 1940's the first field theory, *Quantum Electrodynamics* (QED) [13], was being developed by Richard P. Feynman, Julian S. Schwinger, and Tomonaga Shinichiro. One of the most elegant and precise theories to date (correctly predicting the anomalous magnetic moment of the electron to more than 10 decimal places [13]), it describes the interactions of all particles with electric charge. With this triumph, the theory was further extended by Sheldon Glashow, Abdus Salam, and Steven Weinberg in the 1960s and the electromagnetic and weak forces were united into one force, the electroweak force.

The electroweak gauge theory is described by a $SU(2)_L \otimes U(1)_Y$ group ( $SU(2)_L$ refers to left-handed doublets, while $U(1)_Y$ refers to the weak hypercharge, $Y^{weak} = 2(Q - I_3^{weak})$, where Q is the electric charge, and $I_3^{weak}$ the weak isospin.); it predicted four force mediators, two charged, and two neutral. With the weak force being experimentally short ranged, it became necessary for some of these mediators to gain mass. This process, known as spontaneous symmetry breaking (SSB), correctly predicted the massive $W^{\pm}$ and $Z^0$ bosons,

kept the massless photon, and retained normalizability of the theory.

With such success, the model for strong interactions was based on the QED and electroweak theories, and the theory of *Quantum Chromodynamics* (QCD) [11] was created. QCD is the QFT describing the strong interactions. It is based on an SU(3) gauge field, which leads to 8 mediating massless gauge bosons called *gluons*. As the electric charge is the fundamental parameter of QED, a "color" charge was created for QCD. This charge comes in three flavors, red (r), green (g), and blue (b), and their anti-colors. The theory requires all unbound strongly interacting particles to exist in colorless states, which is possible only by combining all colors, rbg, all anti-colors, $\bar{r}\bar{b}\bar{g}$, or a color and its anti-color, e.g. $g\bar{g}$. Quarks possess a single color, and the theory requires eight gluons, each possessing a color and a differing anti-color. These guidelines explain why all mesons are made of two quarks, like the $\pi^+(u_b\bar{d}_{\bar{b}})$, and all baryons are made up of three quarks, like the proton $(u_r u_g d_b)$ as discussed in Section 2.1.1.

One interesting feature of QCD is that the strength of the coupling increases with decreasing energy scale, i.e. at low energies and long distances the interaction becomes too strong to be treated within the framework of perturbation theory. This leads to *confinement*, which implies that all objects carrying color can never be found as free particles in nature because they are confined in color-neutral hadrons. The quarks that combine into baryons or mesons are referred to as *valence quarks*, and they constantly interact with each other by exchanging gluons. Since gluons can couple to each other, they can emit more gluons that can further split into virtual quark-antiquark pairs called *sea quarks*.

Experimentally, quarks and gluons are manifested as *jets* of color-neutral hadrons. This means that if a single parton (quarks and gluons) emerges from a particle collision, gluons will be radiated which subsequently produce quark-antiquark pairs to form a *parton shower*. Ultimately the partons combine into a jet of hadrons moving in the direction close to that of the original parton. This final step is called *hadronization*.

The strong coupling constant, $\alpha_s$, can be expressed to leading-log in $Q^2$ [11] as:

$$\alpha_s(Q^2) = \frac{12\pi}{(11c - 2n_f)\log(Q^2/\Lambda^2)} \tag{2.1}$$

where $Q$ expresses the magnitude of the momentum transferred in the interaction, $n_f$ indicates the number of quark flavors (6 in the SM), and c is the number of quark colors (3 in the SM). $\Lambda$ is the QCD scale parameter, defined as:

$$\Lambda^2 = \mu_R^2 \exp\frac{-12\pi}{(11c - 2n_f)\alpha_s(\mu_R^2)} \tag{2.2}$$

The parameter $\mu_R$ is an arbitrary renormalization scale introduced to regulate divergences in the perturbative calculation of $\alpha_s$. Eq. 2.1 shows that the strength of the coupling decreases with increasing momentum transfer $Q^2$. Therefore, quarks and gluons are said to be *asymptotically free* when probed at high energies. Theoretical work on asymptotic freedom by Gross, Politzer, and Wilczek was rewarded with the 2004 Nobel Prize. On the other hand, as $Q^2$ approaches $\Lambda$, the coupling becomes large and perturbative calculations are no longer possible.

### 2.1.4 The Electroweak Symmetry Breaking (EWSB) and Standard Model Higgs

The electroweak theory is based on the realization that the quantum of light, the photon, and the quanta of $\beta$ decay, the $W^{\pm}$ bosons, are intimately related. Just as isospin, a symmetry of strong interactions, identifies the neutron and proton as partners, a new symmetry, weak isospin, identifies an electron and its neutrino as partners. The electroweak symmetry is far from exact. The $W$ and $Z$ bosons are among the heaviest known elementary particles, while the photon is the lightest, though they are related by this symmetry. Similarly, the neutrino and the electron can hardly be confused, even though they are partners.

*How is the electroweak symmetry broken ?* A magnetic field, for example, applied to

an atom breaks its rotational symmetry, or heating up a crystal until it melts breaks the discrete symmetry of the lattice. The theory requires that electroweak symmetry be broken in a similar, though more intricate manner. Without the electroweak symmetry breaking (EWSB), the $W^{\pm}$ and $Z$, and all the quarks and leptons would be massless. If any progress is to be made in understanding these masses, the source of EWSB must be discovered. It is one of the core questions in high-energy physics.

The simplest model of EWSB is realized with a doublet of complex scalar fields that introduce four new degrees of freedom of which a single neutral scalar particle, the *Higgs boson*, remains after symmetry breaking and remaining three degrees of freedom are absorbed by the $W$ and $Z$ bosons thus giving them mass.

At hadron colliders, Higgs bosons can be produced via four different production mechanisms:

- gluon fusion, $gg \to H$, which is mediated at lowest order by a heavy quark loop;
- vector boson fusion (VBF), $qq \to qqH$;
- associated production of a Higgs boson with weak gauge bosons, $qq \to W/ZH$ (Higgs Strahlung, Drell-Yan like production);
- associated Higgs boson production with heavy quarks, $gg, qq \to ttH$, $gg, qq \to bbH$ (and $gb \to bH$).

The lowest order production cross sections for the four different processes are shown in the Figure 2.3 for the Tevatron collider as a function of the Higgs boson mass. The dominant production mode is the gluon-fusion process. In the low mass region it amounts at leading order to about 20% of the gluon-fusion cross section, whereas it reaches the same level for masses around 800 $GeV/c^2$. At the Tevatron $p\bar{p}$ collider, the contribution of the associated W/ZH production mode is also important as it provides a handle to pull this signal out of the QCD background and Higgs boson searches exploit this production mode.

The LEP experiments set a lower limit of 114.4 $GeV$ at 95% C.L. [14] using 2.461 $fb^{-1}$ of data collected from $e^+e^-$ collisions at center of mass energies ranging from 189-209 GeV. On

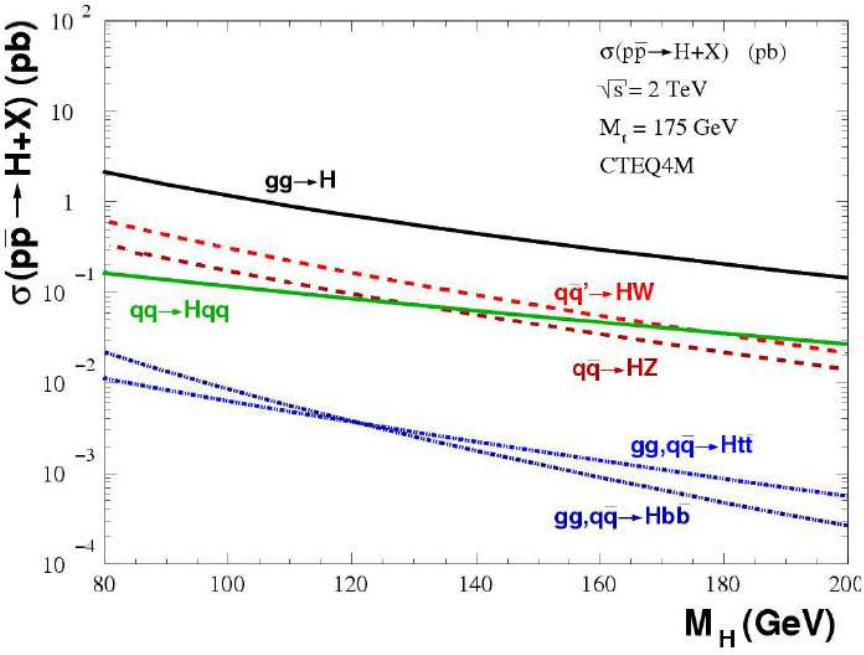


Figure 2.3: Leading Order (LO) production cross section for a Standard Model Higgs boson as a function of the Higgs boson mass at the Tevatron $p\bar{p}$ collider. In the cross section calculation the CTEQ4M parton distribution function parametrization has been used.

the basis of present theoretical knowledge, the Higgs sector in the Standard Model remains largely unconstrained. While there is no direct prediction for the mass of the Higgs boson, an upper limit of $\approx 1\ TeV/c^2$ can be inferred from unitary arguments [15]. Further constraints can be derived under the assumption that the Standard Model is valid only up to a cutoff energy scale $\Lambda$, beyond which new physics becomes relevant. Requiring that the electroweak vacuum be stable and that the Standard Model remains perturbative allows one to set upper and lower bounds on the Higgs boson mass [16, 17]. For a cutoff scale of the order of the Planck mass, the Higgs boson mass is required to be in the range $130 < m_h < 190\ GeV/c^2$. If new physics appears at lower mass scales, the bound becomes weaker, e.g., for $\Lambda = 1\ TeV/c^2$ the Higgs boson mass is constrained to be in the range $85 < m_h < 420\ GeV/c^2$. These constraints are well motivated, but not very powerful.

At the Tevatron, very dedicated program on Higgs boson searches is being carried out. The latest results from the DØ and CDF experiments for full mass range based on up to

8.6 $fb^{-1}$ of data, excluded the SM Higgs with mass $m_h$ at 95% C.L. between 156 and 177 GeV/c$^2$, and between 100 and 108 GeV/c$^2$ [18].

## 2.2 The Top Quark

With the discovery of bottom quark (also known as beauty quark) by the E288 experiment at Fermilab in 1977 [19], search for its left-handed isospin, third generation partner quark started extensively, since in the SM all quarks exist in left-handed isospin doublets. Finally in 1995, the long predicted top quark was discovered in $p\bar{p}$ collisions at the Tevatron by the DØ and CDF collaborations [20, 21]. Our group at Panjab University was also a part of this discovery.

Most of the properties of top quark are similar to that of an up-quark, interacting via the weak and strong forces, having a charge of +2/3e, and being a spin 1/2 fermion. On the other hand, it also exhibits some unique properties. The top quark is the heaviest of the six quarks, nearly as heavy as gold nucleus. Current world average of the top quark mass given by $D\emptyset$ and CDF collaborations is 173.3 $\pm$ 0.6 (stat) $\pm$ 0.9 (syst) GeV/$c^2$ (Figure: 2.4) [22]. This mass is of the order of the electroweak symmetry breaking scale $\upsilon = (\sqrt{2}G_F)^{-1/2} =$ 246 GeV, (where $G_F$ is the Fermi constant), and the top Yukawa coupling $\lambda_t = 2^{3/4}G_F^{1/2}m_t$, (where $m_t$ is the top quark mass) is numerically very close to unity and one expects that a study of top quark physics might reveal details of the electroweak symmetry breaking mechanism [23].

The large top mass further leads to an important property of the top quark of having a relatively large decay width. The top quark decay width, $\Gamma_t$, has been theoretically calculated in the Standard Model to second order in QCD [24] and to first order in electroweak [25] (EW) corrections, and is found to be $\sim$ 2 GeV. This means that the top decay lifetime, $\tau_t$ ($\approx 0.4 \times 10^{-24}$ s) is much smaller than the typical time for formation of QCD bound states, $\tau_{QCD}$ ($\approx 3 \times 10^{-24}$ s). This property of the top quark makes it a very clean source of

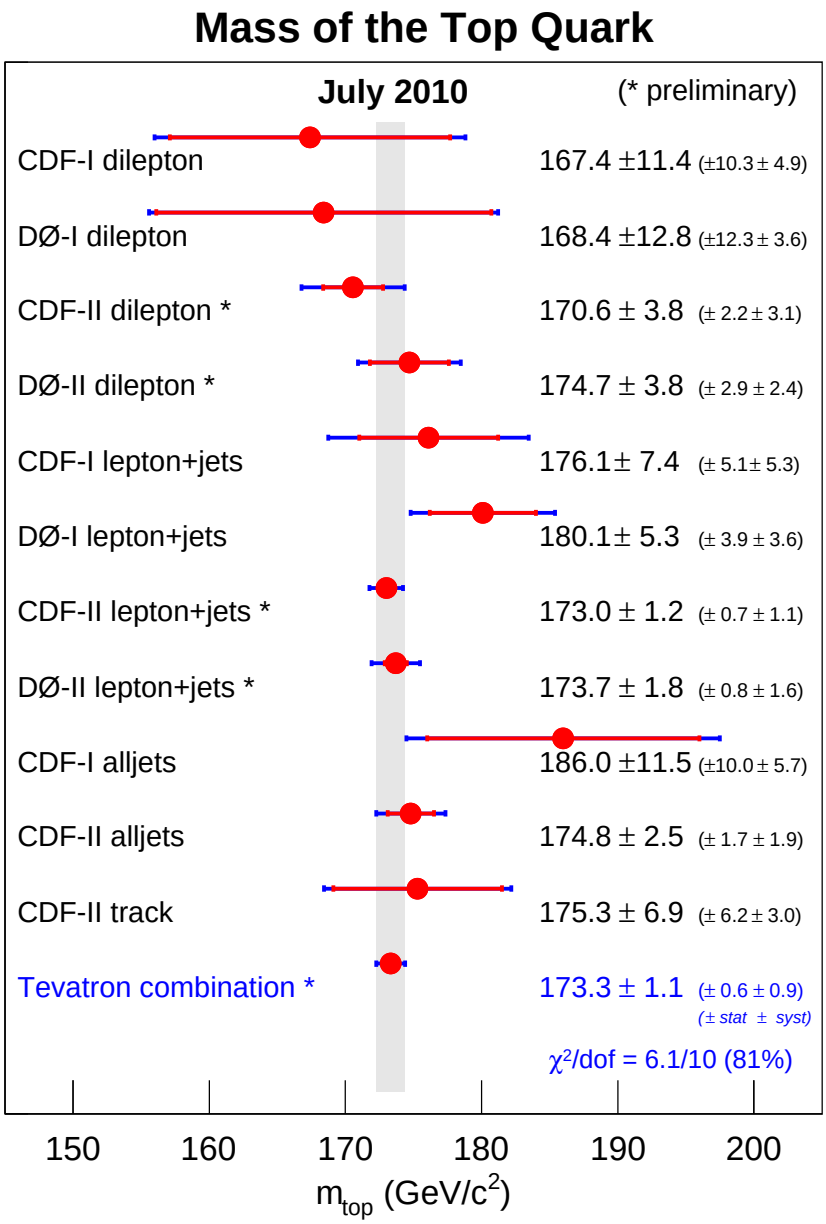


Figure 2.4: Summary of the input measurements and resulting Tevatron average mass of the top-quark.

fundamental information. In particular, the momentum and spin information carried by the top quark is expected to be passed on to its decay products directly, without being distorted by the hadronization process. For example, angular distributions of the top quark decay products are mainly determined by the momentum and spin state of the top quark itself and are not smeared out by hadronization effects [26].

The top quark couples to both the strong and weak fields and allows study via two production modes: pair production (Section 2.2.1) and single top quark production (Section 2.3). The work presented in the thesis deals with the single top quark production.

### 2.2.1 Top Quark Pair Production and Decay

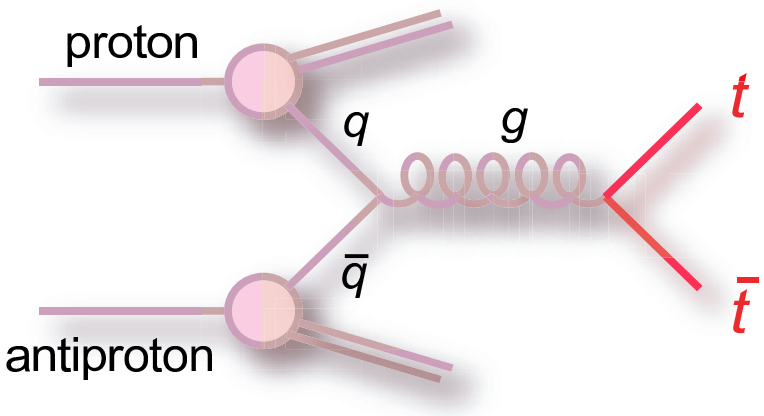


Figure 2.5: Leading Order Feynman diagram for the quark-antiquark annihilation process in the production of top quark pairs.

As discussed in the previous section, the top quark was first observed at Fermilab through pair production via the strong interaction [20, 21]. At a $p\bar{p}$ collider, both the proton valence quarks, and the gluons and quarks from the parton sea can contribute to the production. A large fraction of the parton's momenta is needed for $t\bar{t}$ production - threshold production of $t\bar{t}$ pairs requires $\sim 18\%$ of the total center of mass energy at $\sqrt{s} = 1.96$ TeV. For this reason, about 85% of the $t\bar{t}$ pairs production comes from the quark process $q\bar{q} \rightarrow t\bar{t}$ (Figure: 2.5), and 15% from the gluon process $gg \rightarrow t\bar{t}$ (Figure: 2.6) [27, 28]. At the LHC, the gluon-gluon

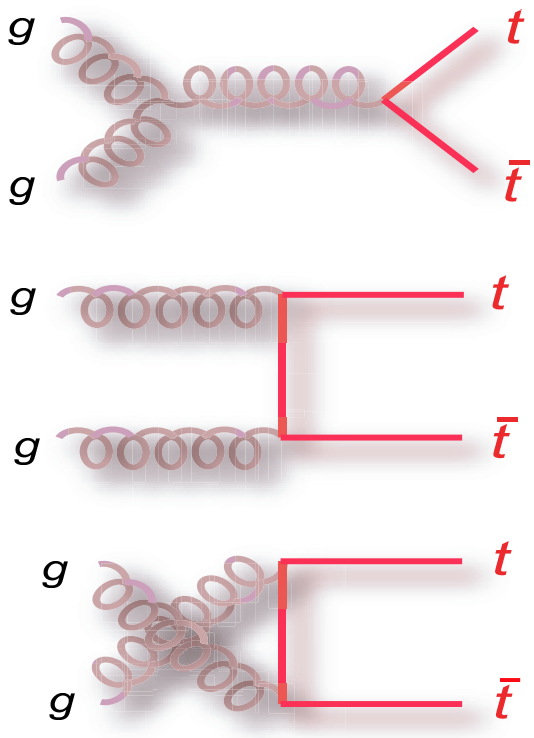


Figure 2.6: Leading Order Feynman diagram for the gluon fusion processes in the production of top quark pairs.

fusion dominates with about 90%. DØ has measured the $t\bar{t}$ cross section to be $7.78^{+0.77}_{-0.64}$ $pb$ for top quark mass of 172.5 GeV [29].

Within the Standard Model the top quark decays via the weak interaction to a $W$ boson and a down-type quark $q$ ($q = d, s, b$). The rate of the decay is proportional to the CKM matrix element $|V_{tq}|^2$. Under the assumption of three fermion families and a unitary $3 \times 3$ CKM matrix the elements $|V_{tq}|$ are severely constrained. The decay $t \to W^+b$ and its charged conjugate occurs to almost 100%, with $|V_{tb}| = 0.999100^{+0.000034}_{-0.000004}$[30].

However, in several extensions of the Standard Model the $3 \times 3$ CKM is a sub-matrix of a larger matrix and would not appear unitary. In this case the $|V_{tq}|$ elements can significantly deviate from their Standard Model values. Experimentally this can be checked with the single top quark production mode, which depends directly on $|V_{tb}|^2$, or with the ratio of branching fractions,

$$R = \frac{B(t \to Wb)}{B(t \to Wq)} \tag{2.3}$$

$R$ can be expressed in terms of the CKM matrix elements as

$$R = \frac{|V_{tb}|^2}{|V_{td}|^2 + |V_{ts}|^2 + |V_{tb}|^2} \tag{2.4}$$

In case of the existence of a fourth quark generation, the denominator does not change. The limits on the mass of a fourth generation quark are already larger than the top quark mass [30], therefore the top quark decays only into the quarks of the three known generations.

The $t\bar{t}$ decay channels are classified according to the decay of the two $W$ bosons from the top decays. The two $W$ bosons from the $t$ and $\bar{t}$ decays then decay either leptonically or hadronically. The branching ratio for $W$ decay to a lepton and neutrino is $B(W \to l\nu_l) = 1/3$, with equal probabilities for $l = e, \mu, \tau$. Hadronically, $W$ decays to a quark-antiquark pair with $B(W \to q_1\bar{q_2}) = 2/3$. The top quark and subsequent $W$ decay are shown diagrammatically in Figure 2.7.

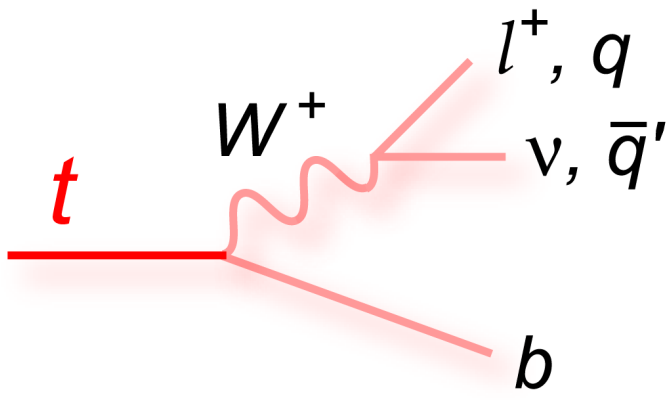


Figure 2.7: Feynman diagram for top quark decay to final state particles.

Since W-bosons can decay either leptonically or hadronically, the $t\bar{t}$ decay channels are classified according the $W$-decay. The three channels are *dilepton*, *lepton+jets*, and *all-jets*.

**All hadronic/all-jets channel** The branching ratio for $W$-boson to decay hadronically is about 46%. This is clearly the best channel with respect to statistics, and all particles in the final state can be detected as jets in the detector. This channel suffers the most, however, from uncertainties in jet energy scale, and also from very large background.

**Lepton+jets channel** Lepton+jets events provide the best compromise between the various factors of statistics, signal-to-background ratio, and resolution of measurements for final-state particles. In the lepton+jets channel one $W$ boson decays hadronically and the other one leptonically into an electron or muon and its corresponding neutrino. Decays of the $W$ boson into a $\tau$ with leptonic decay of the $\tau$ are included in the lepton+jets channel. The total branching fraction of this channel is about 35%.

**Dilepton channel** The decay channel in which both $W$ bosons decay leptonically is known as the dilepton ($ll$) channel. The branching ratio of this channel is about 6%. The dilepton channel has less hadronic activity, since there are only 2 $b$-quarks in the final state. This reduces somewhat the systematic uncertainty due to uncertainties in the jet energy scale. This advantage is offset, however, by the presence of two neutrinos in the final state for which momenta can not be directly measured, and also by the poor statistics relative to

the other two decay channels.

In Figure 2.8 all decay channels and the corresponding branching ratios are summarized.

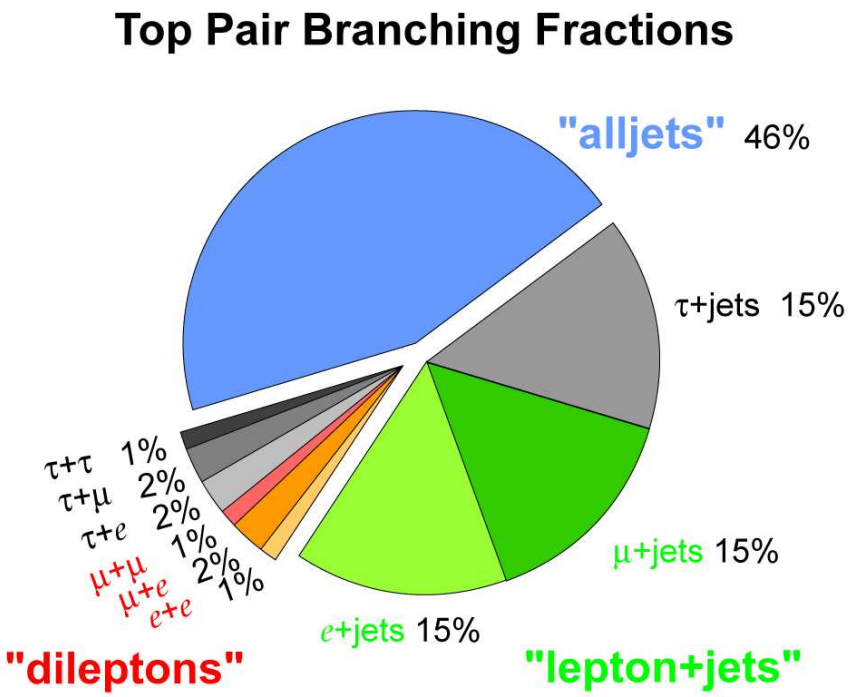


Figure 2.8: $t\bar{t}$ decay channels and the corresponding branching ratios. The branching ratios correspond to the theoretical predictions at tree level.

## 2.3 Electroweak Production of the Top quark : Single Top

Besides the production of top quark pairs involving strong interactions, the production of single top quarks via electroweak interaction is predicted by the standard model. Top quarks produced via an electroweak interaction commonly called single top because only one top quark is produced in the event. The very fact that electroweak interactions are involved, makes single top study very interesting because it provides a direct window to the electroweak symmetry breaking sector of the Standard Model. Single top production mechanisms and the related physics of top quark properties have been the subject of many studies [31, 32,

33, 34, 35, 36, 37, 38].

At hadron colliders, there are three production modes that contribute and each is defined by the virtuality ($Q^2_W$), or four-momentum squared, of the $W$ boson mediating the production. All the three modes are described below.

- **$s$-channel production**

The $s$-channel mode of production shown in Figure. 2.9 [39, 40] is mediated by a time-like $W$ boson ($Q^2_W > (m_{top} + m_b)^2$) and only includes a $2 \rightarrow 2$ process at leading order. This process occurs when a quark and an anti-quark fuse into a virtual $W$ boson, which then splits into a $t$ and $\bar{b}$ quark. We refer to the $s$-channel process as "tb," which includes both $t\bar{b}$ and $\bar{t}b$.

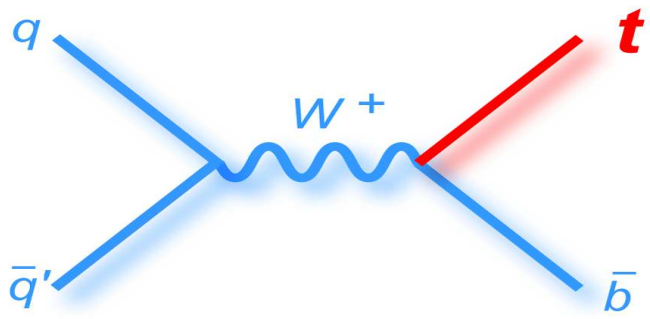


Figure 2.9: The leading order Feynman diagram for the $s$-channel single top production.

Figure. 2.10 (left) shows the theoretical Next Next to Leading Order (NNLO) cross section for this process at the Tevatron as a function of top quark mass. This analysis assumes a top mass of 172.5 GeV which results in a cross section of 1.04 pb [41]. At LHC, where gluon initial states dominates, this channel is highly suppressed as the initial state of this process is two quarks. Due to this, the Tevatron will have the best chance of observing s-channel single top quarks until the LHC has been running for many years.

This sub-process is an attractive mode for studying top quark properties, because the initial state partons are quarks with relatively large momentum fraction $x$, and thus the

parton densities are well understood. In fact, this process may provide the best measurement of the Cabbibo-Kobayashi-Maskawa matrix element, $V_{tb}$ [42]. On the other hand, this mode suffers from a smaller cross section than the $t$-channel mode, and a larger dependance on the mass of the top quark [43].

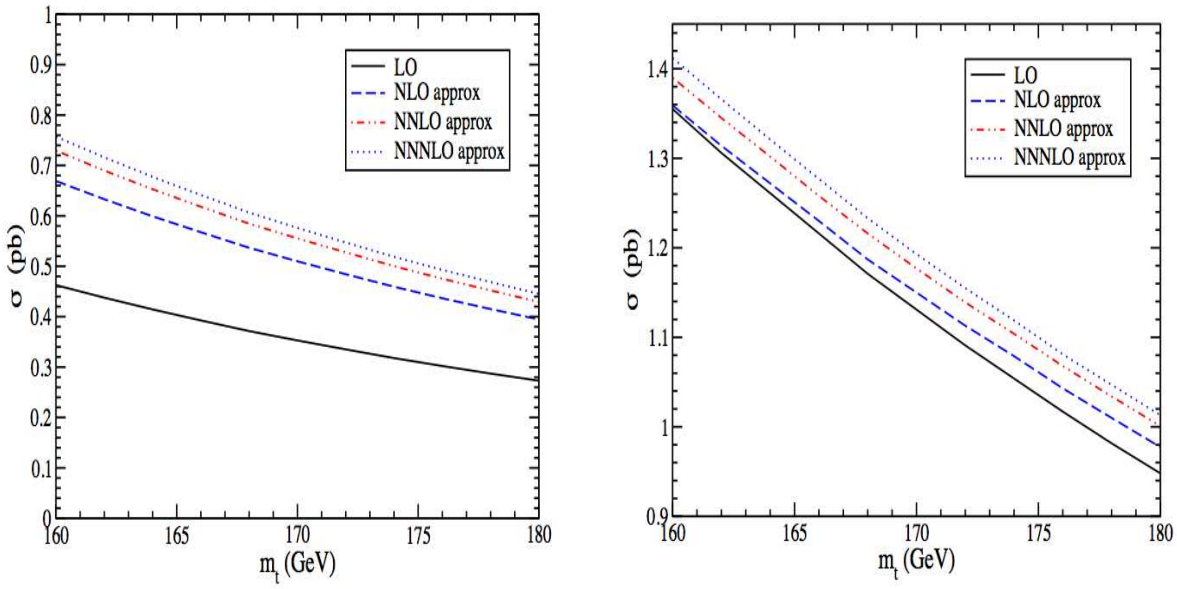


Figure 2.10: Single top quark cross section for $s$-channel (left) and $t$-channel (right) as a function of top quark mass. The plots show the cross sections for top quark production. The cross section for anti-top quark production is identical at the Tevatron. Cross section values quoted throughout the text are for top and anti-top quark production combined.

- **$t$-channel production**

The dominant production mode at the Tevatron, the $t$-channel is mediated by a space-like $W$ boson ($Q^2_W < 0$) as shown in Figure 2.11 [44, 45, 46]. In this case a $b$ quark, coming from gluon splitting, fuses with a $W^+$ boson, producing a top quark. This process is also referred to as $W$-gluon fusion, because the $b$ quark arises from a gluon splitting to $b\bar{b}$. The presence of gluon in the $t$-channel makes things relatively complicated because the gluon distribution function is not well-known.

This process includes a $2 \to 2$ part with a $b$ quark from the proton sea in the initial state as shown in Figure 2.11 (left), and a dominant $2 \to 3$ part, where an extra soft $\bar{b}$ anti-quark appears in the final state explicitly as shown in Figure 2.11 (right). We also refer to the

$t$-channel process as $tqb$, which includes $tq\bar{b}$, $\bar{t}\bar{q}b$, $tq$ and $\bar{t}\bar{q}$.

Figure 2.10 (right) shows the theoretical NNLO cross section for this process. Assuming a top quark mass of 172.5 GeV, the cross section is predicted to be 2.26 $pb$ at the Tevatron [41].

The $t$-channel production mode has the advantage of a larger cross section and a smaller percentage dependance on top mass than the $s$-channel process. This mode is also of interest because within the Standard Model, it provides a way to directly probe the partial width of the top quark, $\Gamma(t \rightarrow W^+b)$, through the effective-$W$ approximation [42]. The major drawback of the $W$-gluon fusion mode is that it suffers from a larger theoretical uncertainty due to the uncertainty in the $b$ quark parton density.

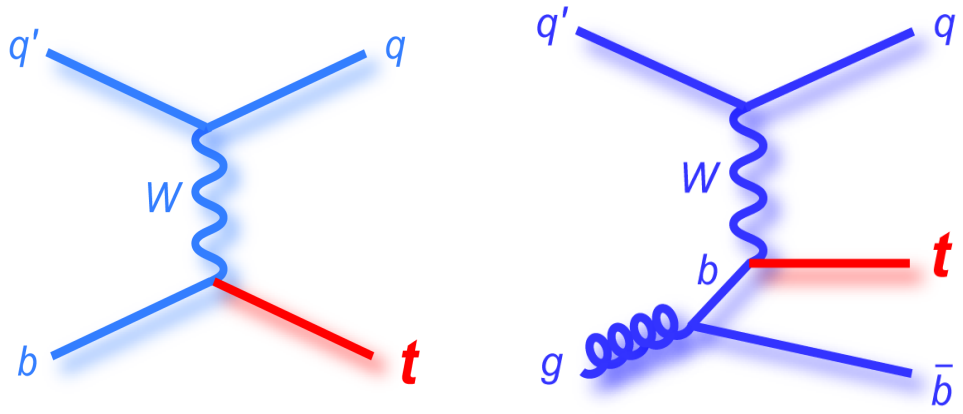


Figure 2.11: The leading order $2 \rightarrow 2$ (left) and $2 \rightarrow 3$ process (right) Feynman diagrams for t-channel single top quark production. The left diagram is a subset of the right.

- **$tW$ production**

The associated $tW$ production mode involves an on-shell $W$ boson ($Q^2_W = M^2_W$) and this process occurs when a $b$ quark radiates a $W$, as shown in the Figure. 2.12. This process involves an incoming $b$ quark from either the proton or anti-proton sea, and a gluon from the other sea. The products are a real $W$ boson and a top quark. This mode has a very small cross section (0.28 $pb$) at the Tevatron [41], because of the small gluon parton density and due to presence of a massive $W$ boson and top quark in the final state. Comparatively, at the LHC this production mode becomes important with a cross section of 66.5 $pb$ at a center

of mass energy of 14 TeV [47] (this compares to an $s$-channel cross section of $\sim$ 11 $pb$ and a $t$-channel cross section of $\sim$ 247 $pb$ [48]). Due to the small cross section of this channel at the Tevatron, this process is assumed to be negligible and is not studied for this analysis.

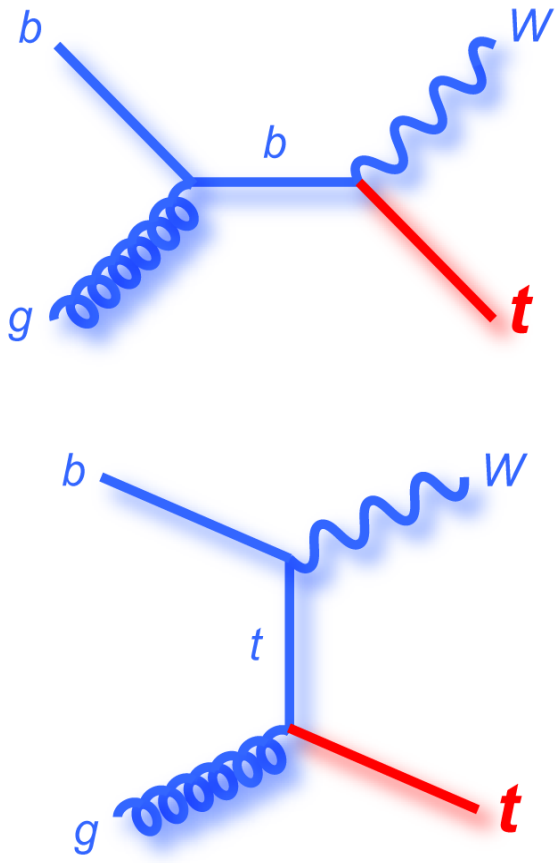


Figure 2.12: The leading order Feynman diagrams for the tW associated single top production.

Present work deals with the study of single top quark production cross section in $s$-channel, $t$-channel and also both channels combined. Also, physics beyond the SM is explored in single top quark production.

### 2.3.1 Why Study Single Top ?

Even though the cross section for single top production is smaller than that for top pair production, studying single top quark production at hadron colliders provides us access to

the properties of the top quarks which are not accessible through $t\bar{t}$ production studies alone.

First, a measurement of the production cross section provides the only direct measurement of the total top quark decay width and the CKM matrix element $|V_{tb}|^2$ [49], without having to assume three quark generations or CKM matrix unitarity. Second, measuring the spin polarization of single top quarks that can be used to test the V-A structure of the top quark electroweak charged current interaction. Third, the presence of various new SM and non-SM phenomena may be inferred by observing deviations from the predicted rate of the single top signal and by comparing different production modes. Fourth, the single top quark final state presents an irreducible background to several searches for SM or non-SM signals, for example Higgs boson searches in the associated production channel. Below, the motivations which led us to study single top production are discussed in detail.

**CKM Matrix Element $V_{tb}$ Measurement**

Top quark decays into a $W$ boson and a down-type quark. The flavor of the down-type quark is determined by the Cabbibo-Kobayashi-Maskawa (CKM) matrix. The CKM quark mixing matrix describes the relationship between the quark mass eigenstates $(d, s, b)$ and the weak eigenstates $(d^{'}, s^{'}, b^{'})$ during charge current interactions.

$$\begin{bmatrix} d^{'} \\ s^{'} \\ b^{'} \end{bmatrix} = \begin{bmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{bmatrix} \begin{bmatrix} d \\ s \\ b \end{bmatrix} \tag{2.5}$$

The $V_{qq'}$ terms represents the strength of the mixing of any up-type quark to any down-type quark. If we assume that there are only three quark generations, then unitarity of the CKM matrix and current measurements of several elements in the matrix tightly restrict the

values for the remaining elements, including $V_{tb}$ as shown in the Equation 2.6 [50].

$$\begin{bmatrix} 0.97428 \pm 0.00015 & 0.2253 \pm 0.0007 & 0.00347^{+0.00016}_{-0.00012} \\ 0.2252 \pm 0.0007 & 0.97345^{+0.00015}_{-0.00016} & 0.0410^{+0.0011}_{-0.0007} \\ 0.00862^{+0.00026}_{-0.00020} & 0.0403^{+0.011}_{-0.0007} & 0.999152^{+0.000030}_{-0.000045} \end{bmatrix} \tag{2.6}$$

In Equation 2.6, the values for the three generation CKM matrix are shown and the value of $V_{tb}$ is relatively fixed. If we relax the assumption of three quark generations, then the value of the matrix elements become virtually unconstrained and particularly $V_{tb}$ is allowed in the following range [38]:

$$0.08 < |V_{tb}| < 0.9993 \tag{2.7}$$

Therefore, comparing the range of $|V_{tb}|$ as allowed by unconstrained matrix (Eq. 2.7) with that of the constrained matrix (Eq. 2.6), could shed light on the total number of quark generations. Now lets discuss how to do the direct measurement of $V_{tb}$ element from single top production cross section.

The Single top quark production involves the $Wtb$ vertex which provides a factor of:

$$\frac{-ig_w}{2\sqrt{2}} V_{tb} \gamma^u (1 - \gamma^5) \tag{2.8}$$

to the single top production matrix element. Hence the single top quark cross section is proportional to $|V_{tb}|^2$ which allows for a direct measurement of $|V_{tb}|$ without the assumption of the unitarity of the CKM matrix. From a measurement of the single top quark production cross section, a measurement of $|V_{tb}|$ can be therefore extracted. A measurement that differs significantly from the range specified in Eq. 2.6 would be clear evidence for physics beyond the standard model, and could possibly indicate the existence of a fourth generation of quarks.

**Measurement of Spin Polarization**

In the Standard Model, all the single top quarks are produced through the left-handed $(V-A)$ electroweak interaction and therefore expected to be highly polarized. As already discussed in Section 2.2, the top decay lifetime is much smaller than the time required to form the QCD bound states, hence no hadronization occurs for the top quark and thus spin correlations are directly passed on to the final decay products. So, measuring the spin polarization of single top quarks can be used to test the $V-A$ structure of the top quark electroweak charged current interactions [51, 52, 53, 38].

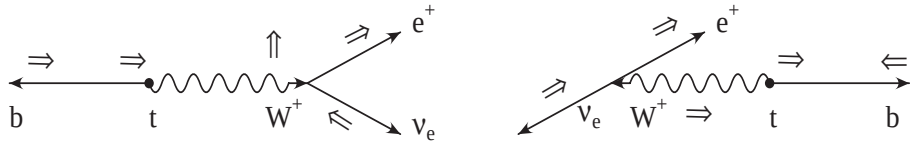


Figure 2.13: The polarization of the top quark and its decay products for longitudinal (left) and left-handed (right) scenarios. The larger arrow above each particle represents the preferred direction of the polarization, and the skinnier arrows represents the preferred direction of momentum in the top quark rest frame [38]. For $t$ decays, the $e^+$ prefers to travel along the polarization of the $t$ and lepton polarization tends to align with the top quark polarization.

In the top quark rest frame, the left-handed nature of the weak interaction leaves a left-handed b quark and either a longitudinally polarized $W$ boson, or a left-handed $W$ boson. In the longitudinal case (Figure 2.13 (left)), the $Ws$ momentum is preferentially aligned with the top polarization. The leptonic decay products tend to align with the $W$ polarization direction, but inheriting its momentum, they are also aligned with the top polarization. In the left-handed case (Figure 2.13 (right)), the $b$ quark is forced to travel along the direction of the top polarization, and the $W$ thus recoils in the opposite direction. In the ensuing $W$ leptonic decay, the charged lepton is forced to be right-handed and thus prefers to travel in a direction opposite to the $W$ parent. The lepton is preferably emitted in the same direction as the top quark spin. The distribution of the angle $\theta_l$ between the lepton momentum in the

top rest frame, and the top polarization vector is given by [38] :

$$F(\theta_l) = \frac{1}{2}(1 + \cos\theta_l) \tag{2.9}$$

Hence by measuring the degree to which top quarks are polarized, one can test the left-handed structure of the $Wtb$ vertex. In this analysis, the top spin polarization information is extensively used to identify the single top quark events.

**Beyond Standard Model : New Physics**

Single top study serves as a proto-type for new physics searches. New physics can influence the single top quark production by inducing non-standard weak interactions via loop effects, or by providing new sources of single top quark events [54, 55, 56, 57, 58]. The two prominent modes of single top production respond quite differently to different realizations of physics beyond Standard Model [38].

One possible form of new physics in single top quark production is a resonance coupled to the top quark. A heavy vector boson, $W^{'}$, can affect the rate of single top production by contributing additional diagrams in which the $W^{'}$ is exchanged as shown in Figure 2.14 (left). Depending upon the relative sign of the couplings between heavy boson and fermions, it can either raise or lower the single top cross section as the initial and final states are the same for both the $W$ exchange and $W^{'}$ exchange. The $s$-channel mode is more sensitive to an exotic charged boson which couples to top and bottom [59]. Because the exchanged particle is time-like, there is the possibility that it can be produced on-shell, resulting in a large enhancement of the cross section. On the other hand, the $t$-channel exchange results in a space-like momentum, which can never go on-shell, and thus the amplitude for the heavy particle is always suppressed by the mass of the heavy boson, $1/M^2_{W'}$.

Another possible sign of the beyond SM physics in the single top quark production can be searched via the anomalous $Wtb$ couplings. The large mass of the top quark suggests that it has large couplings to the electroweak symmetry breaking sector of the SM and may have

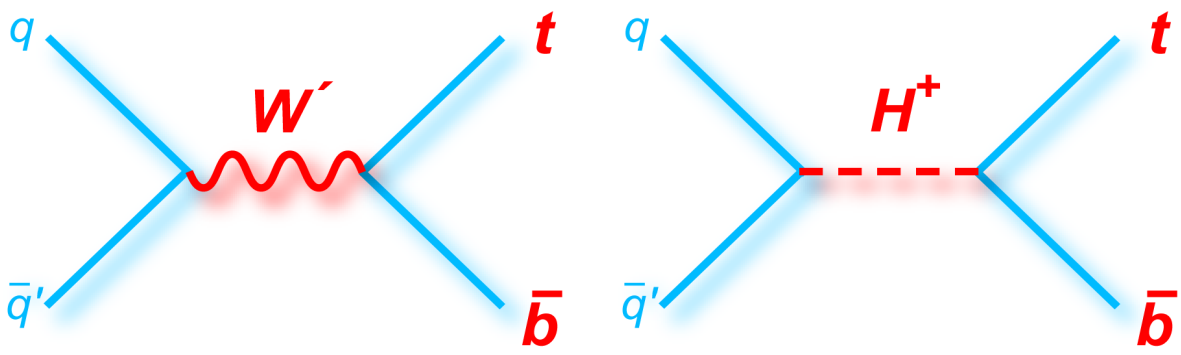


Figure 2.14: Feynman diagram showing how an additional heavy charged vector particle ($W'$) (left) or charged higgs (right) can contribute to the $s$-channel process.

non-standard interactions with weak gauge bosons. The measurement of single top quark production provides a unique probe to study the interactions of top quarks with $W$ bosons. Work done on the search for anomalous $Wtb$ couplings in single top quark production is described in detail in Chapter 8.

Other types of heavy resonances can also be added to the Standard Model, such as those found in supersymmetry (SUSY) e.g. a charged Higgs (Figure. 2.14 (right)). Since Higgs couples very weakly with fermions, this effect can largely modify the top decay width and branching ratios instead of single top production cross section.

Flavor-changing-neutral-currents (FCNC) interactions can also have a noticeable effect through single top quark production [38] as it is possible that the top quark couples differently to light particles from what is predicted by Standard Model. Such an interaction would have varying effects on the different single top processes, and by measuring each of these individual cross sections, one could compare the relative changes to provide evidence for this new physics. The FCNC processes can have a drastic effect on the $t$-channel mode because they involve new interactions between the top quark, a boson ($\gamma, Z, g, or H$), and one of the light quarks, ($c$ $or$ $u$), the $t$-channel mode can be enhanced. For example, in the case of a $Z-t-c$ interaction there is the process $qc \rightarrow qt$ with a $Z$ exchanged. The fact that high energy proton collisions contain more $c$ quarks than $b$ quarks further enhances the new physics contribution compared to the SM. Whereas, for the $s$-channel process it is

almost impossible to measure a FCNC effect due to the low cross section and the difficulty of extracting this channel from the backgrounds.

### 2.3.2 Single Top Event Signature

According to the SM, the top quark is expected to decay almost exclusively into a $b$-quark and a $W$ boson as other decays are suppressed by the small values of $|V_{ts}|$ and $|V_{td}|$ matrix elements. Now $W$ boson can either decay hadronically or leptonically. The signatures of single top events with hadronic $W$ decays are final states containing only jets. This channel is called the *all-jets* or *all-hadronic* channel. Although one or two of the jets are $b$-jets, the signal is swamped by QCD multijet background. Whereas, when $W$ boson decays leptonically, the presence of a lepton and missing transverse energy in final states reduces the multijets background to a manageable amount. The final states considered in this analysis contain an electron or a muon (or their antiparticles) and missing transverse energy accounting for the neutrino. These are called the *lepton+jets* channels (electron and muon).

Figure 2.15 (left) shows the representative Leading Order $s$-channel diagram with the decay explicitly shown. The main components of the final state are the $\bar{b}$ that was created along with the top quark, the $b$-quark from the top quark decay, the charged lepton that comes from the $W$ boson the top quark decays into, and the neutrino from the same $W$ boson. The neutrino is evident only as the missing transverse energy ($\not{E}_{\mathrm{T}}$) in the detector. Figure 2.15 (right) shows the representative Leading Order $t$-channel diagram with the decay explicitly shown. The main components of the final state are the $b$-quark from the top quark decay, the charged lepton that comes from the $W$ boson the top quark decays into, the neutrino from the same $W$ boson, and the forward light "spectator" quark. Occasionally in the $t$-channel, a second $b$ jet is produced from the splitting of the initial-state gluon with enough transverse momentum $p_T$ to be reconstructed.

In spite of the small cross section and huge amount of backgrounds present, single top quark events have some unique kinematic features which helps to discriminate from the

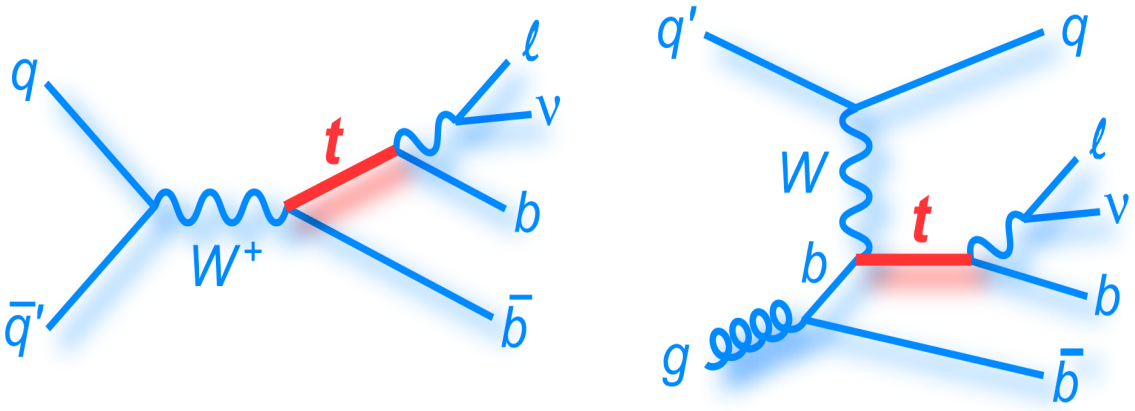


Figure 2.15: Leading Order Feynman diagram for single top production via the $s$-channel (left) and $t$-channel (right), followed by a leptonic decay of the top quark.

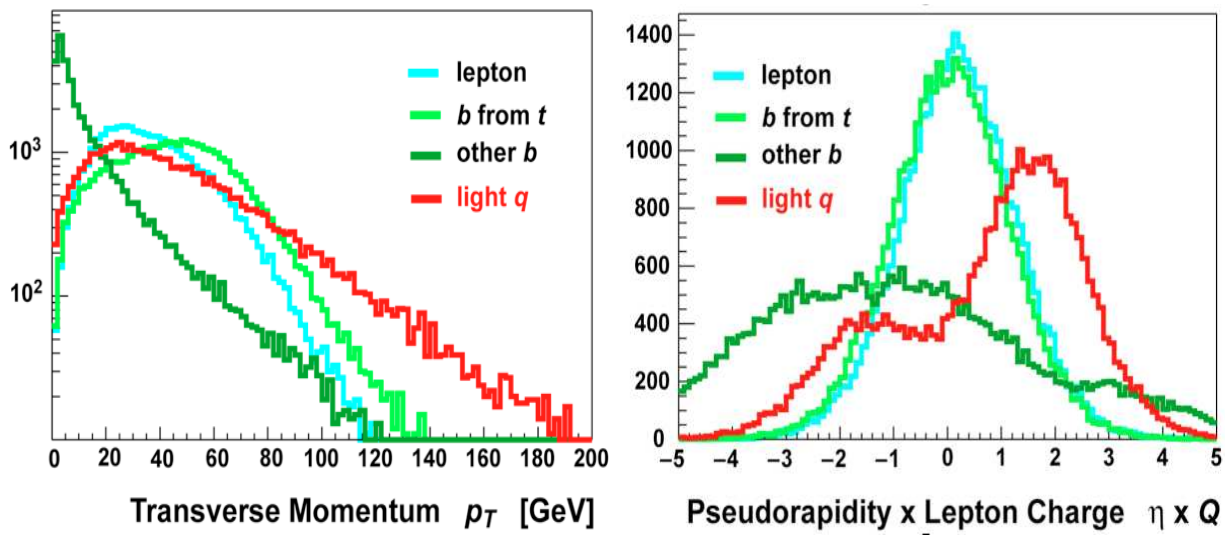


Figure 2.16: Parton-level kinematic distributions for single top $t$-channel from the single top Monte Carlo samples. The $p_T$ spectrum for each final state particle is shown on the left, the $Q(l) \times \eta$ spectra are shown on the right. These distributions were generated after parton showering was applied.

background. Some characteristics kinematic features of single top quark production can be seen in Figure. 2.16. The $b$ quark emitted from the top quark decay tends to be central and has large transverse momentum $p_T$, and the lepton originating from the $W$ boson has a softer $p_T$ spectrum. This occurs because the preferred direction of the lepton is anti-aligned with the top quark direction due to the $V - A$ nature of the weak force, as already discussed in Section. 2.3.1. The $b$ quark produced in association with the top quark in $t$-channel single top production tends to have high rapidity and low momentum and is often not reconstructed in the analysis. The light quark produced in the $t$-channel has reasonably large $p_T$, but its most distinguishing feature is the asymmetric $Q(l) \times \eta$ distribution shown in Figure 2.16 (right), where $Q(l)$ is the charge of the lepton in the event. This asymmetry arises since the final state light quark produced during single $t$ ($\bar{t}$) production most often is a $d$ ($\bar{d}$) quark that moves in the same direction as the proton (antiproton) [38]. The light quark $\eta$ will hence tend to have the same sign as the charge of the lepton from the top decay.

### 2.3.3 Background Processes

The processes which share approximately the same final states as single top quark events or in other sense mimic the signal events are referred to as the background processes. The main background sources taken into account for this analysis are described below.

#### $W$ + jets

As described in Section 2.3.2 the top quark in single top quark events will decay to a $W$ boson and $b$ quark, where the $W$ boson is only considered to decay to a lepton and neutrino. With an additional $b$ quark or light quark this makes the signature of single top quark events as: one high $p_T$ lepton, large missing $\not{E}_{\mathrm{T}}$ , and two or more jets. The similar event signature can be generated by $W$ boson produced together with two or more jets. This background is referred to as "$W$ + jets". This is one of the dominant backgrounds to the single top and mainly for single top events with lower jet multiplicity. Figure 2.17 shows example diagrams

of $W$ boson plus jets production with $W + b\bar{b}$ on left and $W$ + light jets on right.

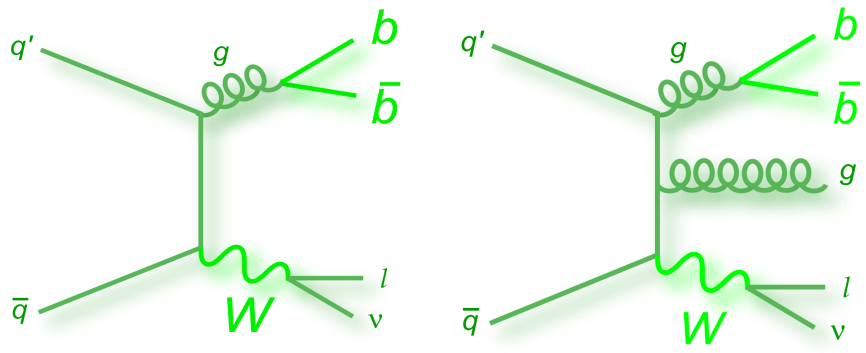


Figure 2.17: Example leading order Feynman diagrams for a $W$ +jets event.

### $t\bar{t}$ pair production

Another large background present in the dataset are events originating from top pair production. The top pair production background, referred to as $t\bar{t}$, is defined by the decay of the two $W$ bosons, from the decay of the two top quarks. This event resembles a single top event with slightly higher jet multiplicity. The first case when one of the $W$ bosons decays to two quarks and other decays to a lepton and neutrino is referred to as *"lepton+jets"* ($t\bar{t} \rightarrow l + jets$) because the final state in the event is one lepton, one neutrino, and four quarks. The other way in which a $t\bar{t}$ event can enter the data sample is when both $W$ bosons decay to leptons and neutrinos. In this case, there are two quarks, two leptons, and two neutrinos. If one of the leptons is mistakenly identified as a jet, this event mimics a single top event and passes the selection criteria. These events are referred to as *"dilepton"* ($t\bar{t} \rightarrow ll$) events. An example Feynman diagram for the $t\bar{t} \rightarrow l + jets$ process is shown in Fig 2.18.

### Multijets

The third largest background present in the dataset is multijet events produced by the strong interaction. The background processes responsible for these events in the dataset are quite different for electron events and muon events. In electron events one of the reconstructed

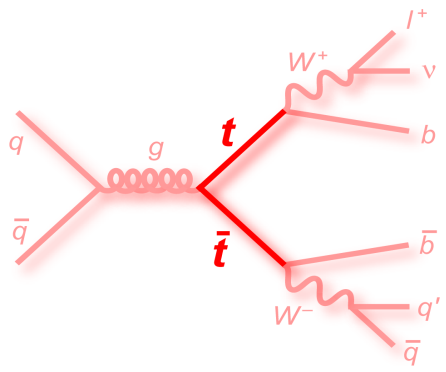


Figure 2.18: Feynman Diagram showing leading order top quark pair production and decay in $t\bar{t} \rightarrow l + jets$ channel.

jets will have a large electromagnetic fraction causing it to be mis-identified as an electron. In muon events a gluon will decay to a $b\bar{b}$ pair and one of the $B$ mesons will undergo a semi-leptonic decay and produce a muon. In both cases, another jet may not be properly reconstructed leading to a large amount of missing energy in the event, thus mimicking the single top quark event signature. An example Feynman diagram for a multijet process producing a lepton, missing $\not{E}_{\mathrm{T}}$, and jets is shown in Figure 2.19.

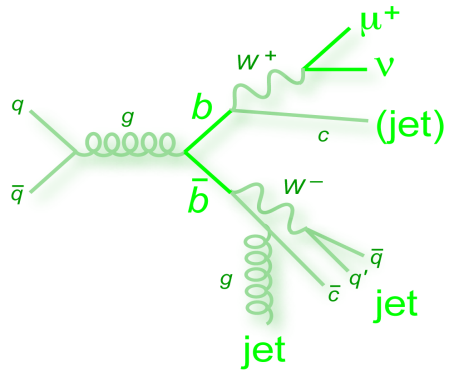


Figure 2.19: Example Feynman diagram for a multijet event.

### $Z$ + jets

The production of $Z$ boson plus jets is shown in Figure 2.20. This background can mimic single top event signature in two ways. One when a $Z$ boson decays to two isolated leptons (electrons or muons) and if one of the leptons is not measured and mis-identified as missing

transverse energy and second way is when there are two muons in the final state and one accidentally overlaps a jet and is thereby identified as a tagging muon from a semi-leptonic $b$ decay. This background is less prominent as compared to $t\bar{t}$ and $W$ + jets backgrounds because to pass the event selection, along with the lepton mis-identification, a $b$-jet mis-identification is also required at the same time.

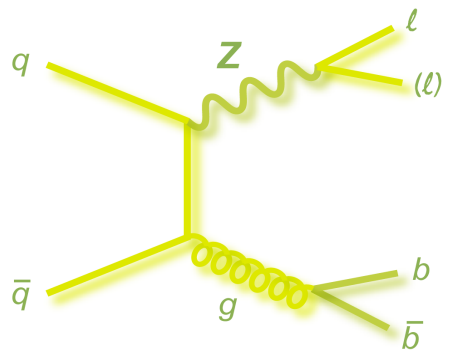


Figure 2.20: Example Feynman diagram of $Z + jets$ production.

**Dibosons**

Another less prominent background to single top production, due to its small cross section, is the diboson production. Figure 2.21 shows the example $WW$ production. Also there can be $ZZ$ and $WZ$ productions. These events enter into dataset when one boson undergoes leptonic decay and other hadronic decay, and some of the jets are classified as containing $b$-quarks.

### 2.3.4 Observation of Single Top Quark and Current Status

During Run I (1992-1996), at the Tevatron with center of mass energy of $\sqrt{s}$ = 1.8 TeV, DØ [60, 61] and CDF [62, 63] experiments published first limits on the cross sections using about 90 $pb^{-1}$ of data. These searches first used the simple kinematic event selection and then followed by the use of multivariate analysis technique, neural networks to set upper

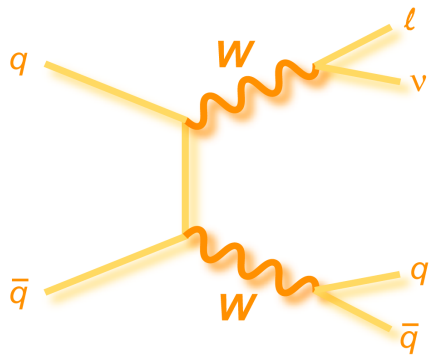


Figure 2.21: Example Feynman diagram for diboson production showing $WW$ production.

limits on the cross sections. The limits on the cross sections for $s$-channel and $t$-channel production were about 10 - 20 times greater than the predicted values.

The Tevatron collision energy was increased to 1.96 TeV in 2001 and the DØ and CDF detectors were significantly upgraded including the improvements to central solenoid system and silicon tracking system. Both experiments set improved limits on single top cross section. DØ analyzed 230 $pb^{-1}$ of data using neural networks (NN) for signal-background separation and a Bayesian binned likelihood calculation using the NN output distributions, and set 95% confidence level (C.L.) upper limits of 6.4 $pb$ in the $s$-channel and 5.0 $pb$ in the $t$-channel in 2005 [64, 65]. Whereas CDF analyzed 160 $pb^{-1}$ of Run II data using a cut-based selection and a maximum-likelihood fit and set 95% C.L. upper limits of 13.6 $pb$ on $s$-channel production and 10.1 $pb$ on $t$-channel production in 2005 [66].

The next step in the search led to the major improvement and hence evidence of the single top quark. DØ performed analysis on 0.9 $fb^{-1}$ of data and combined the $s$- and $t$-channels (assuming the SM ratio of the two parts), and applied three multivariate methods to separate signal from background to reach 3.4 standard deviation ($\sigma$) significance for a single top quark signal. The measured cross section for $tb + tqb$ production combined was $4.9 \pm 1.4$ $pb$ [32]. The measurement significance represents a probability of 0.035% for the background to have fluctuated up and given a false measurement of signal with a cross section of at least 4.9 $pb$. A significance greater than $3\sigma$ is considered in the high energy physics community not to be

sufficient for a claim of discovery or first observation (which is set at $5\sigma$), but is high enough to indicate the evidence for the process. The CDF collaboration performed a similar analysis on 2.2 $fb^{-1}$ of data and reached a significance for single top quark signal of $3.7\sigma$, published in 2008 [33]. They measured the cross section for $tb + tqb$ production to be $2.2 \pm 0.7$ $pb$.

Evidence of single top quark was the first milestone achieved by DØ and CDF collaborations. After evidence, quest for single top quark started at more fast pace and both collaborations started tuning their methods to reach the second milestone, which is the Observation of single top quark. Finally after $\sim$ 14 years of discovery of top quark by pair production, DØ [67] and CDF [68] collaborations published the first observation of single top quark production in 2009. DØ result is based on the 2.3 $fb^{-1}$ of data and it measures a cross section of $3.94 \pm 0.88$ $pb$ for $tb + tqb$ which corresponds to a $5.0\sigma$ significance for observation. Similarly, CDF performs analysis on 3.2 $fb^{-1}$ of data and measures a cross section of $2.3^{+0.5}_{-0.6}$ $pb$ for $tb + tqb$ which also corresponds to a $5.0\sigma$ significance. After the observation of single top quark production with combined $s$- and $t$-channels, DØ also measured the cross section of $t$-channel alone using same 2.3 $fb^{-1}$ dataset and published the direct evidence for electroweak production of single top quarks through the $t$-channel exchange of a virtual W boson [69]. Also, along with SM cross section measurements, DØ has performed searches to see the sensitivity of single top quark production to non-SM contributions. DØ has performed analyses for FCNC [70] and $W^{'}$ [71] searches and set corresponding limits.

The analysis presented in this thesis is based on the 5.4 $fb^{-1}$ data collected by the DØ detector and mainly concentrates on the cross section measurement of $s$- and $t$-channels independently and also provide a precise cross section measurement of $s + t$-channel using Bayesian Neural Networks. Also, this thesis provides a search for the anomalous couplings in single top quark production using the same dataset.

# Chapter 3

# Experimental Setup: The Tevatron Collider and DØ Detector

> ***The strongest arguments prove nothing so long as the conclusions are not verified by experience. Experimental science is the queen of sciences and the goal of all speculation.*** - Roger Bacon.

The data collected for the analysis originates from proton-antiproton ($p\bar{p}$) collisions produced by the Tevatron accelerator at Fermi National Accelerator Laboratory which is located 30 miles west of Chicago, Illinois, USA and is spread over 6800 acres of land [72, 73, 74, 75, 76]. Proton and antiprotons travel around the roughly four mile Tevatron ring in about 21 microseconds ($\mu s$) and collide at a $\sqrt{s} = 1.96$ TeV center-of-mass energy.

Both protons and antiprotons travel in orbits within the ring designed such that they meet only twice upon completing an entire circuit, at BØ and DØ, which are home to the two multipurpose detectors - the collider detector at Fermilab (CDF) and the DØ detector. Three super-bunches of particles travel around the ring with a spacing of 2.6 $\mu s$ in between them. Inside each super-bunch, there are twelve individual bunches, separated by 396 nanoseconds ($ns$). Each bunch initially contains up to $\sim$ 300 billion protons and $\sim$ 100 billion antiprotons, where the limiting factor is the number of antiprotons available for use. Collisions occur when

these bunches meet at their crossing points along the ring and the individual protons and antiprotons interact.

The information on a specific interaction, produced when particle beams collide, is contained in a number called the *cross section* $\sigma$. Cross sections can be measured experimentally and calculated from theory. It is, basically, the probability that a particular fundamental interaction will take place in the collision. The number of events of a particular kind expected to be produced per second is given by:

$$\text{Number of events per second} = \mathcal{L} \times \sigma, \tag{3.1}$$

where $\mathcal{L}$ is the luminosity of the colliding beam. Luminosity is a performance measure of the colliding beams that is independent of any specific fundamental particle interaction. It contains the information of the incoming particle flux, the number of incoming particles per second, and the number of target particles per unit area, which is the colliding, or target beam.

$$\mathcal{L} = \frac{BN_pN_{\bar{p}}}{\beta\epsilon} \tag{3.2}$$

where B is the number of bunches per beam, $N_p(N_{\bar{p}})$ is the number of protons (antiprotons) in a bunch, $\epsilon$ is the emittance (95% of the beam area), and $\beta$ is the effective beam overlap area of the two bunches when they collide.

A commonly used number is *integrated luminosity* which is given by

$$\textit{Integrated Luminosity} = \int \mathcal{L}\, dt, \tag{3.3}$$

and is the luminosity, integrated over the total time a high energy particle physics experiment has run. Integrated Luminosity, when multiplied by the cross section, gives the number of events of the kind that were produced, which means, the physics productivity of an accelerator depends on the maximum luminosity that it can give and how long it can run at

that luminosity.

The Tevatron Collider physics program is comprised of two stages : Run I and Run II. The Run I (1992-1996) collider program, operating at energy of $\sqrt{s} = 1.8$ TeV, delivered an integrated luminosity of about 130 $pb^{-1}$ to both CDF and DØ experiments. Run I produced a large number of exciting physics results. The main highlight of these results is the discovery of top quark by CDF and DØ experiments in 1995 [20, 21]. Top Quark discovery became the first milestone of the Tevatron. After Run I phase, there was six year upgrade period for the collider and both experiments. In order to continue the discovery of new particles and to increase the sensitivity for new physics, many efforts were taken to increase the center of mass energy from 1.8 TeV to 1.96 TeV and to deliver more than a 20-fold increase in the number of particle collisions observed and recorded by the two experiments. The DØ experiment utilized the first year of collisions to commission the detector, trigger and electronics. The last major element to be completed - the central tracker - which is crucial for particle position and charge determination, was fully instrumented in April 2002, marking the beginning of DØ Run II physics program. Tevatron Run II physics program ended on September 30, 2011. This excellent machine has delivered $\sim 11.9$ $fb^{-1}$ of data in its full RunII operation. Figure 3.1 shows the complete RunII Integrated luminosity graph for delivered and recorded data.

The data analyzed in the present work for this thesis was recorded by the DØ detector during Run II of the Tevatron operation in the years 2002-2009. This chapter briefly describes the technical details of the Tevatron Collider and the DØ detector and its subsystems during the Run II.

## 3.1 The Fermilab Accelerator Chain

The Tevatron is the final stage in a sequence of seven accelerators [77, 78]. A Cockcroft-Walton pre-accelerator, a linear accelerator (Linac) and a synchrotron (Booster) provide a

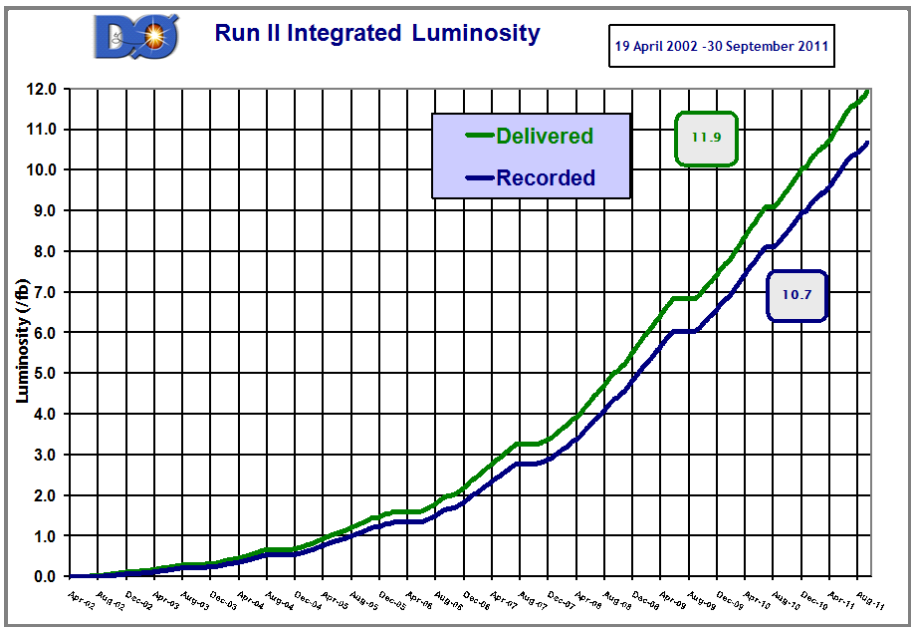


Figure 3.1: Tevatron RunII Integrated Luminosity.

source of 8 GeV protons. The antiproton Debuncher and Accumulator are two components of the antiproton source. The Main Injector serves as the final boosting stage before injecting protons and antiprotons into the Tevatron. It also provides the necessary source of energetic protons which are needed in the antiproton source. Figure 3.2 gives an overview of the Fermilab accelerator complex.

### 3.1.1 Cockroft-Walton pre-accelerator

The purpose of the pre-accelerator is to produce negatively charged hydrogen ions ($H^-$) with an energy of 750 keV, which are then transferred into the Linac. Hydrogen gas ($H_2$) enters a magnetron [79] surface-plasma source as shown in Figure 3.3. Due to the electric field between the anode (negatively charged) and cathode (positively charged), the electrons are stripped away from the hydrogen atoms to create a plasma. The positively charged hydrogen ions then strike the surface of the cathode to collect extra electrons and thereby form negatively charged hydrogen ions. The $H^-$ ions are extracted through the anode aperture with an electric field of 18 kV applied by the extractor plate (Figure 3.3).

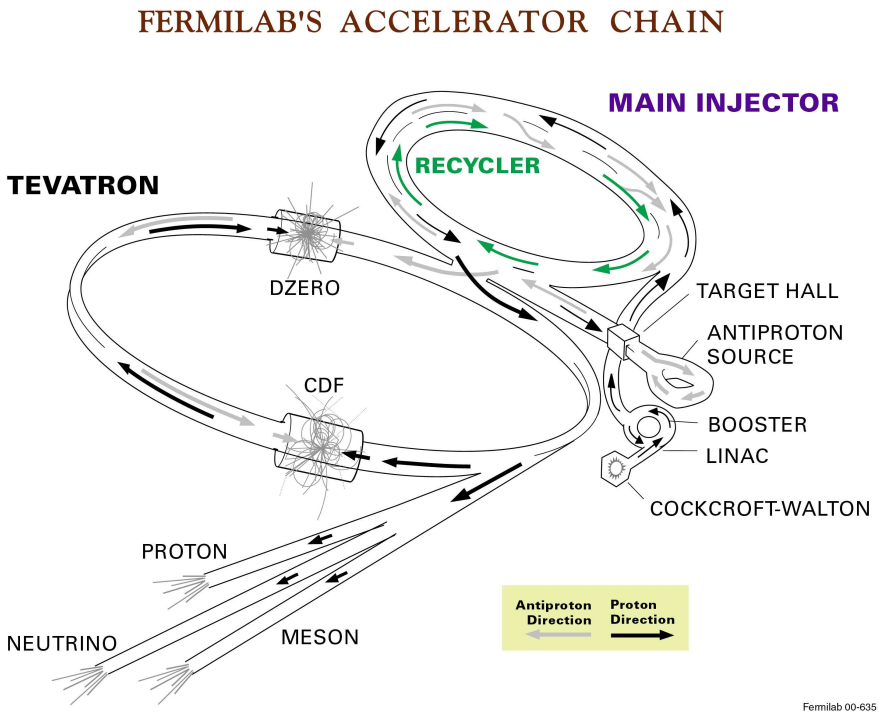


Figure 3.2: Schematic view of the Fermilab accelerator chain.

A commercial Cockcroft-Walton generator as shown in Figure 3.4 produces a 750 kV potential difference by charging capacitors in parallel from an AC voltage source and discharging them in series, via diodes. The Cockcroft-Walton generator is used to further accelerate the $H^-$ ions to an energy of 750 keV. After exiting the Cockcroft-Walton device, the $H^-$ ions travel through a transfer line. Before entering the Linac the continuous stream of $H^-$ ions passes through a single gap radio frequency (RF) cavity which bunches the beam at the RF frequency of the Linac (201.24 MHz).

### 3.1.2 The LINAC

The Linac [80] receives bunches of 750 keV $H^-$ ions from the pre-accelerator and accelerates them further to an energy of 400 MeV using RF cavities. The RF cavities are contained within a collection of steel tanks which hold a sequence of drift tubes separated from each other by gaps as shown in Figure 3.5. In order to accelerate $H^-$ ions, the cavities are designed in such a way that particles traveling in the gaps experience an acceleration, while

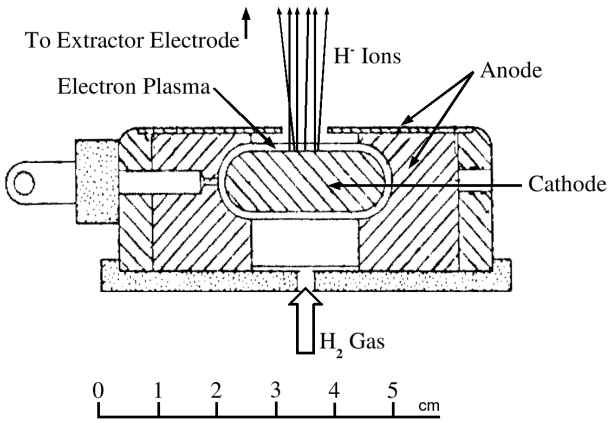


Figure 3.3: Schematic view of magnetron operation for the hydrogen ion source.

Figure 3.4: View of Cockcroft-Walton pre-accelerator at Fermilab.

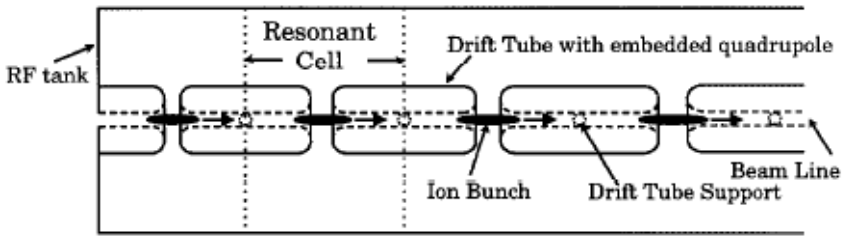


Figure 3.5: Simplified diagram of drift tube accelerator.

particles traveling in the drift tubes are shielded from the RF. After passing through the Linac, bunches of 400 MeV $H^-$ ions are transferred into the Booster.

### 3.1.3 The Booster Synchrotron

The Booster [81] is the first synchrotron in the chain of accelerators. It consists of a sequence of dipole and quadrupole magnets and 17 RF cavities arranged in a circle with a diameter of 151 m, and accelerates protons to an energy of 8 GeV. Negatively charged $H^-$ ions coming from the Linac are merged with protons ($H^+$ ions) circulating in the Booster with the help of dipole magnets. The electrons are subsequently stripped from the $H^-$ ions by letting the combined beam pass through a carbon foil.

Once the Booster is filled with proton bunches, the RF cavities provide an acceleration up to 8 GeV. At the same time, the field strength in the dipole magnets is adjusted accordingly in order to maintain a constant radius for the circulating particles. Once the protons have reached an energy of 8 GeV, they are transferred into the Main Injector.

### 3.1.4 The Main Injector

The next stage of the acceleration is the Main Injector [82]. The Main Injector is circular synchrotron, 3 km in circumference, and located tangentially next to the Tevatron ring at the $F\emptyset$ straight section. This ring is a new addition for the Run II upgrade and it replaces the Main Ring, which operated during Run I. The Main Injector is capable of delivering up

to 3 times as many protons as the Main Ring. Being located outside of the enclosure of the Tevatron ring (in contrast the Run I Main Ring was inside the same tunnel), it reduces beam halos and backgrounds seen in the colliding detectors during Run I. There are two functions that the Main Injector performs. The first is to coalesce the proton bunches injected from the Booster into single high-density bunch of $\sim 5 \times 10^{22}$ protons and further boost their energy to 150 GeV, before the proton beam is delivered to the Tevatron. The second function is to extract proton bunches at 120 GeV and then deliver them to the anti-proton facility.

### 3.1.5 The Anti-proton Source

The Antiproton Source consists of three major components: the *Target Station*, the *Debuncher*, and the *Accumulator*. In the first step, the Target Station receives 120 GeV protons from the Main Injector and diverts them onto a Nickel Target. This produces a shower of secondary particles (including antiprotons) at many different angles and with a large spread in particle momentum. A Lithium lens and bending magnets are used to focus the beam and remove positively charged particles. A process called *stochastic cooling* [83] is used in both the Debuncher and the Accumulator to reduce the spread in momentum and position of the antiprotons, thereby "cooling" them.

Both the Debuncher and Accumulator are located in a rounded-triangle shaped tunnel with a circumference of about 51 m. Antiprotons coming from the Target Station are transferred into the Debuncher where the momentum spread of the particles is reduced. It is technically very challenging to accumulate a large quantity of antiprotons. On average, for every 1 million protons that hit the Nickel target, only about 20 antiprotons can be gathered. Therefore, the Accumulator stores antiprotons until a sufficient amount has been generated to be transferred into the Main Injector. The Accumulator is capable of storing antiprotons over many hours.

### 3.1.6 The Recycler

The 8 GeV Recycler ring is located in the same tunnel as the Main Injector, directly above the Main Injector beam line. The Recycler provides the following functionality to increase the number of protons, and thus increase the luminosity:
(a) A high reliability storage ring for antiprotons. It can store $2.5 \times 10^{12}$ antiprotons.
(b) A post-accumulator ring. By emptying the antiprotons from the Accumulator into the Recycler periodically, the Accumulator can operate in its optimum intensity regime.
(c) A receptacle for antiprotons left over (about 80 %) at the end of the Tevatron stores. By re-using these antiprotons, the luminosity is more than doubled.

### 3.1.7 The Tevatron Ring Synchrotron

The Tevatron is the final stage in the sequence of proton and antiproton acceleration. It has a diameter of 2 km ($\sim$ 6 km circumference) and uses superconducting magnets which operate at liquid helium temperature providing magnetic fields of up to 4 Tesla. Protons and antiprotons are accelerated to 980 GeV, leading to a center-of-mass collision energy of 1.96 TeV.

Protons and antiprotons travel in groups of particles (*bunches*) in opposite directions while sharing the same beam pipe. A full revolution (*turn*) takes $\sim$ 21 $\mu$s. The Tevatron injects 36 bunches of both protons and antiprotons for each store. A three fold symmetry is imposed by separating the 36 bunches into three super-bunches. Overall, this leads to a time structure where bunches of protons and antiprotons (*live bunch crossings* or *zero bias events*) collide at 1.7 MHz.

## 3.2 DØ Detector

The DØ detector [84, 85, 86, 87], one of the two multipurpose detector designed to study the proton-antiproton collisions originating from the Tevatron Collider, was proposed in 1983

and is in operation since 1992 when Run I started. It went through subsequent upgrades. The DØ detector is a collection of smaller sub-detectors working in tandem to detect and measure all particles produced from the hard scatter collision. The inner-most detectors near the beam pipe are the tracking detectors, described in Section 3.2.2, which record the paths of charged particles as they enter and leave the detector. Section 3.2.3 describes the preshower detectors placed after the tracker. The next layer of the detector is the calorimeter, described in Section 3.2.4. The calorimeter measures the energy of the lightest electromagnetically interacting particles, such as the electrons and the photons, and strongly interacting particles, such as pions or neutrons. Another sub-detector, called the luminosity monitor, described in Section 3.2.5, is designed to record the presence of an inelastic $p\bar{p}$ collision in the bunch crossing. This information is used in the analysis to normalize backgrounds and expected signal yields. The outer-most layer of the DØ detector is the muon detector, described in Section 3.2.6. Once the collision has been measured a complex set of trigger decisions, described in Section 3.2.7, must be satisfied before the event is recorded to tape for later analysis. Figure 3.6 shows a schematic of the detector and the relative location of each of these subsystems.

### 3.2.1 DØ Coordinate System

Before discussing about the parts of the DØ detector in detail, first a brief discussion of the coordinate system which is used to describe the detector and data analysis is given here. The Cartesian right-handed coordinate system is used to describe the four-momenta of particles arising from $p\bar{p}$ collisions. The $z$-axis is aligned along the beam axis, with the positive $z$-direction along the proton beam. The $y$-axis points vertically upward, and the $x$-axis points horizontally toward the center of the Tevatron ring. The $\phi$ and $\theta$ angles describe, respectively, the azimuthal and polar angles, with $\theta$=0 along the beam pipe. When polar

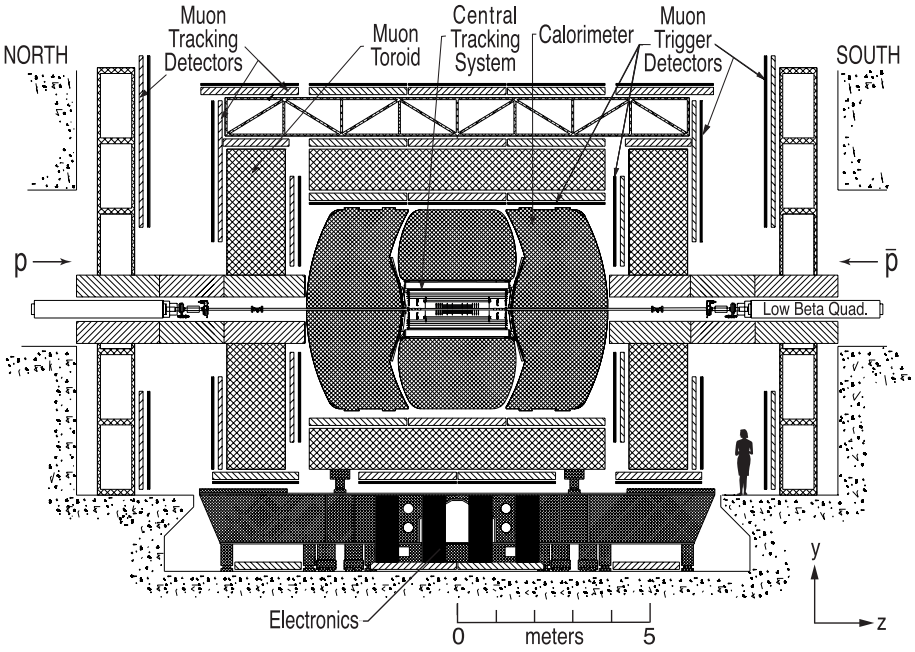


Figure 3.6: Schematic of the DØ detector.

coordinates are used, $r$ is the transverse distance, defined as

$$r = \sqrt{x^2 + y^2}, \tag{3.4}$$

A useful parameter for physics analyses is the pseudorapidity, $\eta$, defined as

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right), \tag{3.5}$$

In the relativistic limit ($E >> m$), $\eta$ is a good approximation of the rapidity of the particle

$$y = \frac{1}{2}\ln\left(\frac{E + p_z}{E - p_z}\right), \tag{3.6}$$

This is a convenient parameter for hadron colliders because the multiplicity of high energy particles is approximately constant in rapidity. Rapidity distributions are also invariant under Lorentz boosts in the $z$-direction.

In $p\bar{p}$ collisions, longitudinal momentum of the individual partons is not known. For this reason, analyses usually work with transverse momenta. Transverse momentum is defined

to be

$$p_T = p \sin\theta, \tag{3.7}$$

Similarly, transverse energy is defined as

$$E_T = E \sin\theta \tag{3.8}$$

It is assumed that the transverse momentum of the $p\bar{p}$ system is initially zero.

The beam bunch length is about 30 $cm$ in $z$, so collisions do not necessarily occur at the center of the detector. It is therefore necessary to distinguish between detector $\eta$, which is measured assuming a particle trajectory from the center of the detector, and physics $\eta$, measured with respect to the point along the $z$-axis where the collision occurred. Unless otherwise noted, all $\eta$ values should be taken to be detector $\eta$.

### 3.2.2 Central Tracking System

The inner-most layer of the detector, constructed directly outside of the Tevatron beam pipe serves as tracking system which measures the momentum, electric charge, impact parameter, position and trajectory of the charged particles produced in a collision. The central tracking system locates the primary interaction vertex (PV) with a resolution of 35 $\mu$m along the beam-line and the impact parameter (IP) with resolution of better than 15 $\mu$m in $r - \phi$ for particles with $p_T > 10$ GeV at $|\eta| = 0$. A solenoid provides a nearly uniform 2T magnetic field parallel to the beam axis. Charged particles leave a pattern of hits in the layers of the tracking detectors, and these hits are used to reconstruct a curved trajectory in 3-dimensional space. The curvature of the track gives the momentum and the sign of electric charge. The components of the central tracking system are:

1. the silicon microstrip tracker,
2. the central fiber tracker, and

3. the solenoid

A cross-section of the central tracking region is shown in Figure 3.7

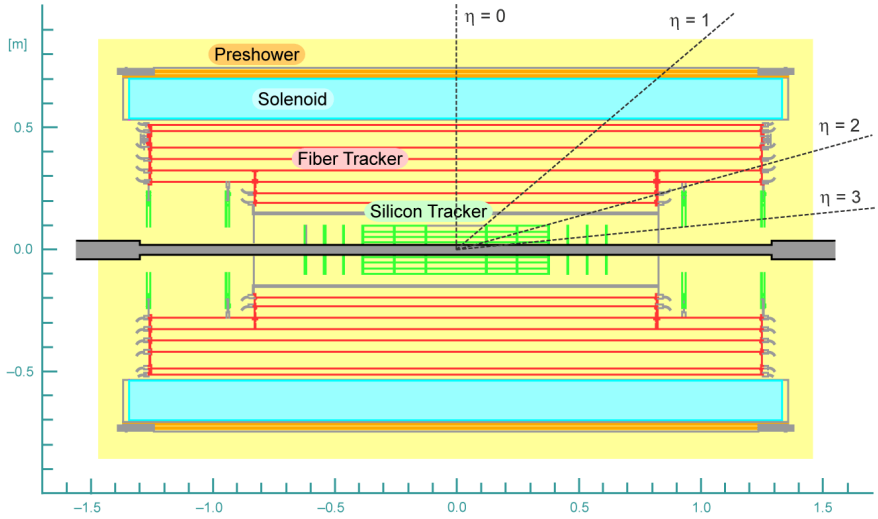


Figure 3.7: Cross-sectional view of the DØ Central Tracking Systems.

### Silicon Microstrip Tracker

The detector which is located immediately outside the Tevatron beam pipe and is designed to provide high resolution position measurements of charged particles that is used in tracking and vertexing over nearly the full pseudorapidity ($\eta$) coverage of the calorimeter and muon systems is the Silicon Microstrip Tracker (SMT) [88, 89]. The large $z$ distribution of the $p\bar{p}$ interaction region ($\sigma_z \sim 28$ cm) provides a challenge for designing a detector in which tracks are predominantly perpendicular to detector surfaces for all $\eta$. This feature led to a hybrid system shown in Figure 3.8 with barrel detectors measuring primarily the $r-\phi$ coordinate and disk detectors which measure $r-z$ as well as $r-\phi$. Thus vertices for high $|\eta|$ particles are reconstructed in three dimensions by the disks, and vertices of particles at small values of $|\eta|$ are measured in the barrels.

The fundamental detecting unit is a silicon wafer with hundreds of parallel microstrips

across its surface. The strips are held at a bias voltage, and as charged particles pass through, they ionize the silicon and produce current flows measured by the wire strips known as "hits". These silicon wafers are arranged in six barrels concentric with the beam pipe, 12 F-disks perpendicular to it, and 4 larger H-disks (Figure 3.8). The barrels are composed of rectangular wafers, called ladders, arranged in four concentric layers around $z$-axis. There is additional layer, Layer 0, the innermost layer, that was installed before RunIIb to maintain high quality tracking and pattern recognition in spite of detector aging and the higher instantaneous luminosity.

Each of the 6 barrels is 12 cm long and has 72 ladders arranged in four layers, with each layer having two sub-layers at slightly different radii to provide azimuthal overlap, as illustrated in Figure 3.9. A combination of double sided and single sided layers are used. The double sided wafers provide coordinates in the $r-\phi$ plane from the axial side, while the other side, by having an angle of 2°or 90° with respect to the beam line, gives a measurement in the $r-z$ plane. The single sided layers have an axial orientation. Each barrel is capped with a disk of wedge detectors, called the F-disks. These disks comprise twelve wedges made of double sided silicon wafers with trapezoidal shapes with each disk rotated by 7.5° with respect to neighboring disk. In both the far-forward and far-backward regions, two large diameter disks, called the H-disks are installed to improve momentum resolution for tracks up to $|\eta| = 3$.

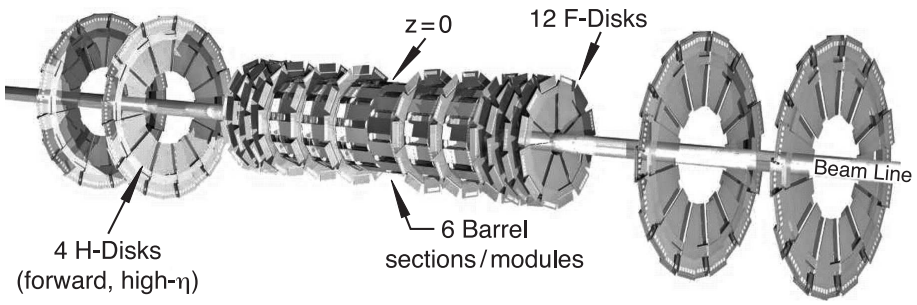


Figure 3.8: A schematic of the silicon microstrip tracker.

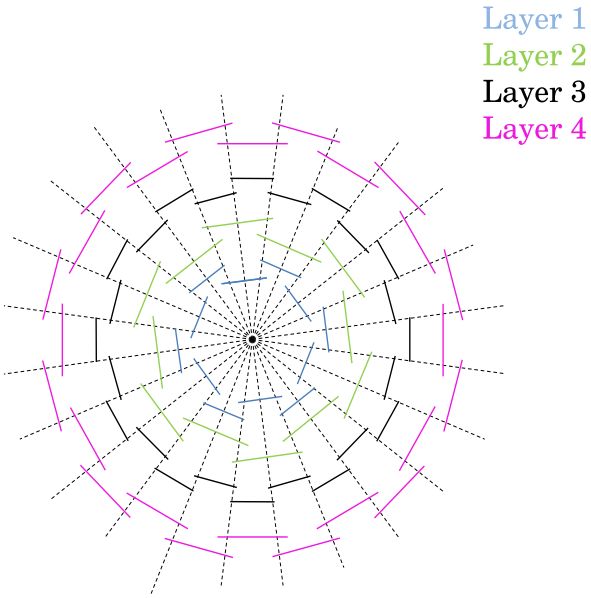


Figure 3.9: Arrangement of barrel modules in $r - \phi$.

### Central Fiber Tracker

The scintillating central fiber tracker (CFT) [90] surrounds the SMT and provide tracking coverage up to $|\eta| < 1.6$. CFT is comprised of concentric layers of scintillating fibers constructed from polystyrene doped with an organic fluorescent dye, paraterphenyl (Figure 3.10) to detect the passage of charged particles. As particles travel through the fibers, a small fraction of their energy is deposited and converted into light. This light travels through a wave guide and is collected by a "visible light photon counter" (VLPC) outside the detector, signaling a hit in the fiber tracker. The 835 $\mu$m fibers are arranged in eight concentric cylinders about the beam pipe, with the innermost at a radius of 20 cm and the outermost at 52 cm. Each cylinder is composed of a double layer of fibers, with one layer parallel to the $z$ axis and the other layer $\pm 3^{o}$(alternating between cylinders). The location of particle in a layer is identified by the intersection of any two fibers that received a signal.

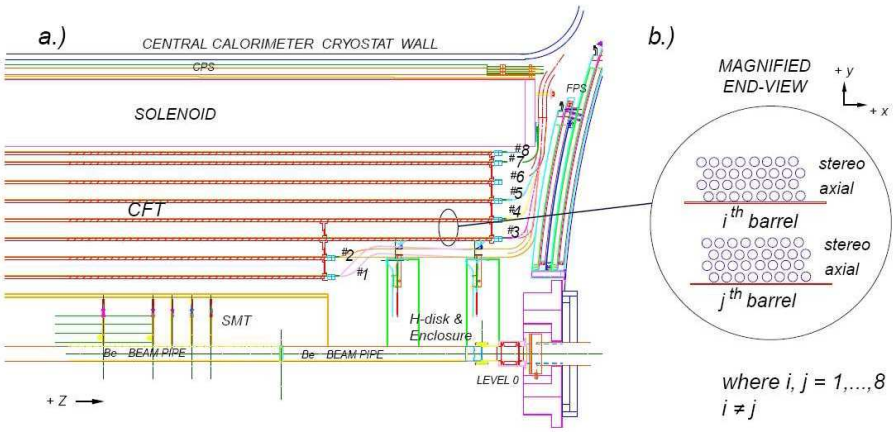


Figure 3.10: Layout of the CFT illustrating the layers of scintillating fibers.

### Solenoid Magnet

The superconducting solenoidal magnet was designed [91] and installed between RunI and RunII, to optimize the momentum resolution $\delta p_T/p_T$, and tracking pattern recognition. The solenoidal magnet is 2.73 m long with an outer radius of 71.0 cm and is made of niobium-titanium wire that is cooled to less than 4.7 K using liquid helium. In spite of the size constraints of the surrounding calorimeter, it provides a 2T magnetic field parallel to the $z$-axis with two possible polarity configuration (Figure 3.11) that causes charged particles to curve as they travel radially out from the interaction point. The radius of curvature in meters is given by,

$$R = \frac{p_T}{0.3qB}; \tag{3.9}$$

where $p_T$ is the transverse momentum in units of GeV, $q$ is the charge of the particle in units of electron change ($e$), and $B$ is the magnetic field in Tesla; thus providing a measurement of the momenta of charged particles passing through the SMT and CFT.

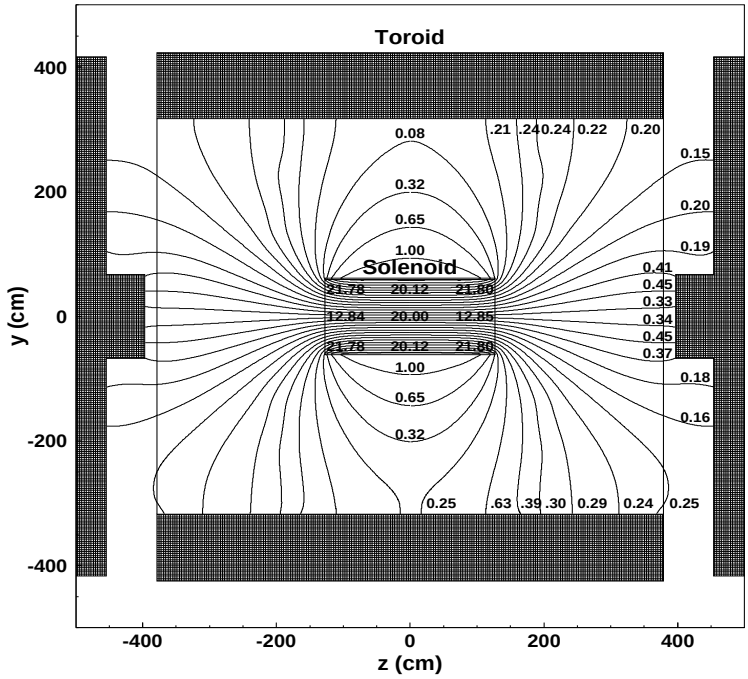


Figure 3.11: Magnetic field lines in the DØ detector with both the solenoidal and toroidal magnets at full current.

### 3.2.3 Preshower Detectors

To increase the resolution of the calorimeter after traversing the material in the solenoid, the remaining 5 cm gap between the solenoid and calorimeter is instrumented with another detector - the preshower. There are two preshower detectors located just before the calorimeters : Central Preshower (CPS) covering $|\eta| < 1.2$ and a forward preshower covering $1.4 < |\eta| < 2.5$ (shown in Figure 3.7). These preshower detectors play an important role in improving calorimetry measurements and are also sensitive enough to aid the tracking measurements, thus leading to enhance electron and photon identification. The preshower detectors use scintillating fibers with a triangular cross section arranged into layers as shown in Figure 3.12.

The central preshower (CPS) [92] consists of three layers of scintillating fibers, triangular in cross section, arranged at parallel, +23.77°, and -24.02° angles to the $z$ axis. There is also a thin (7/32") radiative lead plate encased by two stainless steel plates (1/32" each) between the solenoid and the CPS. The plates add about two radiation lengths of material

for particles normal to the $z$ axis, and up to four at larger $\eta$.

The forward preshower (FPS) [93] rests in the thin cavity between central and end calorimeters. It consists of two layers of double layered scintillators, separated by an 11 mm lead-steel-lead plate. The first layer is known as the minimum ionizing particle or MIP layer, and the second layer following the lead plate is known as the absorber layer. All charged particles will leave a hit in the MIP layer, and electrons and photons will begin to shower in the absorber layer, producing a cluster of hits. Each layer is made of 48 wedges, and each wedge is double layered. Scintillating fibers in the layers are placed at 22.5° with respect to one another. The layers are slightly offset such that particles do not cross more than one set of uninstrumented cracks between the wedges.

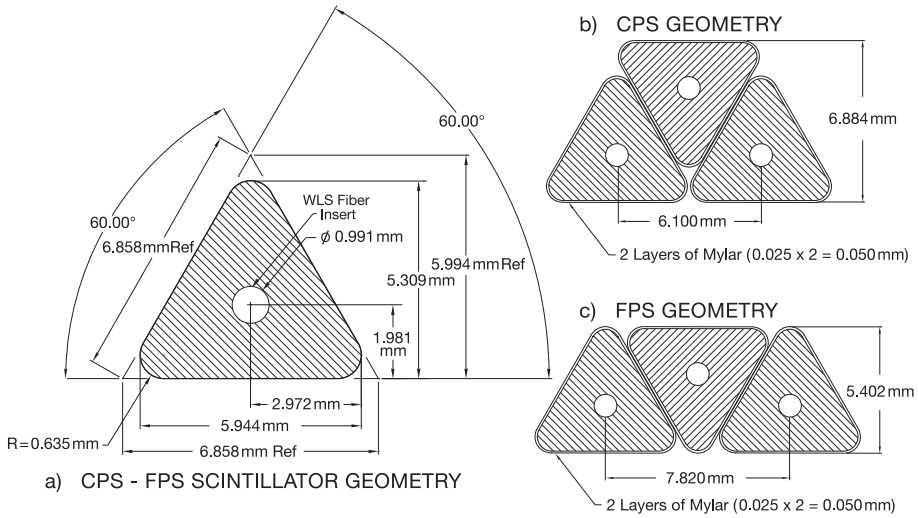


Figure 3.12: (a) Cross section diagram of a scintillating fiber used in the preshower detectors. (b) Diagram showing how fibers are stacked in each layer of the CPS. (c) Diagram showing how fibers are stacked in each layer of the FPS.

### 3.2.4 Calorimeter

Till now all the detectors studied keep track of the particles produced in a given collision, without disturbing them in any way (except, of course, bending them in the magnetic field). These detectors were designed to trace the trajectories of the incoming particles, to measure their momentum and charge, and to reconstruct their vertices. The other important infor-

mation is the energy of these particles. Also, since the tracker is good for detecting only charged particles, another type of detector is required which can detect neutral particles too.

The DØ calorimeter [87] is used for the identification of electrons, photons, jets and muons and to establish the transverse energy balance in an event. The main parts of the DØ calorimeter are shown in Figure 3.13. The central calorimeter (CC) extends in pseudorapidity to roughly $|\eta| = 1.1$, and the two end (forward) calorimeters, ECN (north) and ECS (south), extend coverage to $|\eta| \approx 4$. The active medium for all the calorimeters is liquid argon and each of the three calorimeters (CC, ECN and ECS) is located within a cryostat that maintains the temperature at approximately 90 K.

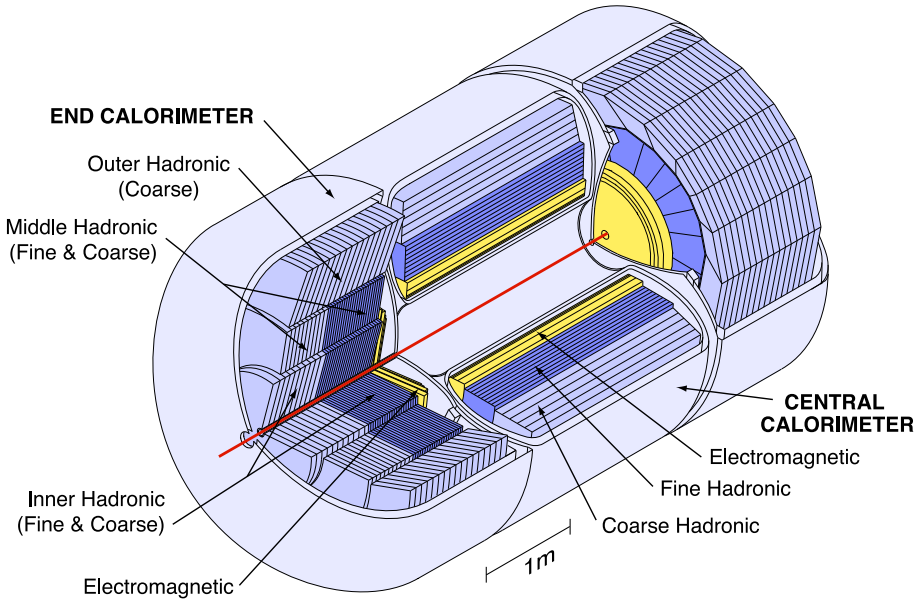


Figure 3.13: Isometric view of the DØ central and two end calorimeters.

A unit cell of the DØ Uranium/Liquid-Argon sampling calorimeter, shown in Figure 3.14, consists of a grounded metal absorber plate and a signal board with the resistive surfaces connected to a high voltage source of typically 2.0 to 2.5 kV. Incoming high energy particles will react via the electromagnetic or strong force with the absorber plates resulting in showers of secondary particles. Liquid argon, being the active medium, is ionized by these secondary particles and the resultant electrons collect on the signal board where the magnitude of the charge is read out. Electron drift time across the 2.3 mm gap is 450 ns, longer than the

time between two consecutive bunch crossings. The calorimeter signal processing hardware, through fast signal shaping, is aptly capable of distinguishing and correctly disentangling these two signals.

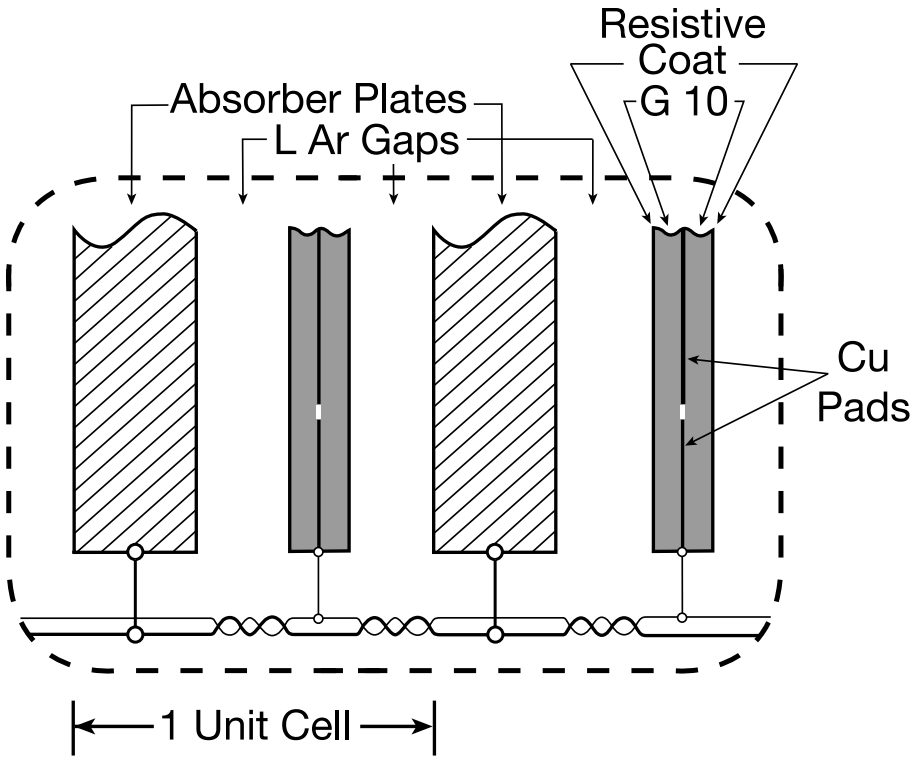


Figure 3.14: Schematic view of the unit cell for the calorimeter.

The calorimeter modules themselves are further segmented into three distinct sections to adequately sample the secondary particle shower shape and size, for a range of particles and range of energies. In order of increasing radius, these are: electromagnetic (EM) section with relatively thin uranium absorber plates, fine-hadronic (FH) with thick uranium plates and the coarse-hadronic (CH) with thick copper or stainless steel plates.

As the EM objects tend to decay over a shorter distance than hadrons, the innermost layers of both the CC and EC are the electromagnetic layers and the outer are hadronic calorimeter. The EM calorimeter is constructed of alternating layers of depleted Uranium, which acts as the shower inducing material, and the liquid Argon, which acts as the active medium. The depleted Uranium plates are 3 mm thick in the central region and 4 mm thick in

the forward end-cap region while the liquid Argon active region is 2.3 mm thick as also shown in Figure 3.14. The EM calorimeter has four layers of cells representing nearly 21 radiation lengths. The hadronic calorimeter is actually two detectors as said before: one called the fine hadronic calorimeter which employs 6 mm thick Ur-Ni alloy as the shower inducing material and the coarse hadronic calorimeter which used 46.5 mm thick plates of copper in the central region and stainless steel in the forward region. The hadronic calorimeter also uses liquid Argon as the active material. The combination of the fine and coarse hadronic calorimeters provides an additional 7 radiation lengths to the detector. The numerous radiation lengths are important to ensure that a particle deposits nearly all of its energy in the detector.

The DØ calorimeter also has fine segmentation (i.e. radial size of the cells), which allows for excellent energy and position measurement of particles as they shower in the detector. The segmentation of the EM calorimeter in $\delta\eta \times \delta\phi$ is $0.1 \times 0.1$ for all layers except the third layer, where the segmentation is $0.05 \times 0.05$. The fine segmentation in the third layer is because the electromagnetic shower is expected to reach a maximum in this layer, and the finer segmentation thus improves the energy resolution. The fine hadronic layers of the calorimeter also have a segmentation of $0.1 \times 0.1$, while the segmentation in the coarse hadronic calorimeter is $0.2 \times 0.2$. An octant of the DØ calorimeter including segmentation can be seen in Figure 3.15.

From $1.1 < |\eta| < 1.4$, because of the multiple cryostat design, inter-cryostat detectors (ICD) have been added to sample the shower energy that is lost by particles that transverse the module endcaps and cryostat walls. A particle passes through this region of less active material before reaching the end of the calorimeter. However there is a sufficient amount of passive material for showering to occur. To sample the showers in this region scintillating tile detectors of size in $0.1 \times 0.1$ in $\delta\eta \times \delta\phi$ were installed on the external walls of the endcap calorimeters.

There are about 55,000 read out channels. The readout is accomplished in three steps.

1. In the first step, the signals from the uranium-liquid argon detector are transported to

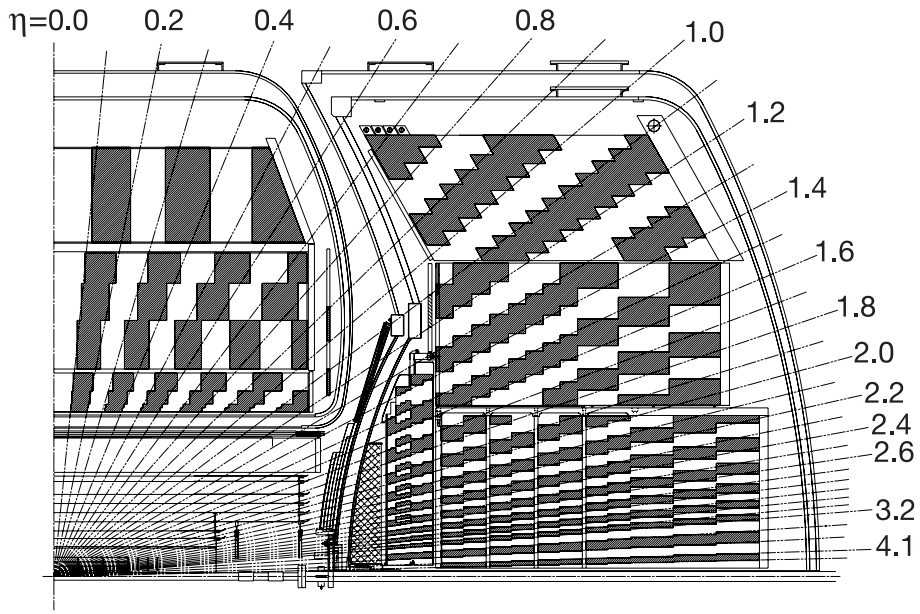


Figure 3.15: Schematic view of a portion of the DØ calorimeters showing the transverse and longitudinal segmentation pattern. The shading pattern indicates groups of cells ganged together for signal readout. The rays indicates pseudorapidity intervals from the central of the detector.

charge per-amplifiers located on the cryostats.

2. In the second step, the signals from the pre-amplifiers are transported to signal shaping and analog storage circuits (baseline subtractor boards or BLSs) located underneath the cryostats. The BLSs hold the signal for about 4 $\mu$s until the trigger is available, and provide baseline subtraction to remove any low frequency noise or pileup present in the signal. In addition, faster shaped analog sums of the signals are picked off to provide prompt inputs to the calorimeter trigger for both Level 1 and Level 2 trigger decisions.

3. The precision signals from the BLSs are sent over 130 m to analog-to-digital converters (ADCs), and then enter the data acquisition system for the Level 3 trigger decision and storage to tape.

### 3.2.5 Luminosity Monitor

The primary purpose of the luminosity monitor (LM) is to determine the Tevatron luminosity at the DØ interaction region. This is a very important task as the precision of a cross section determination depends on the precision of the luminosity measurement. This is accomplished by detecting inelastic $p\bar{p}$ collisions with a dedicated detector. The LM also serves to measure beam halo rates and to make a fast measurement of the $z$ coordinate of the interaction vertex.

The LM detector consists of two arrays of 24 plastic scintillation counters with photomultiplier tube (PMT) readout, located at $z = \pm 140$ cm as shown in Figure 3.16. The arrays are located in front of the end calorimeters and occupy the radial region between the beam pipe and forward preshower detectors. The counters are 15 cm long and cover the pseudorapidity range, $2.7 < |\eta| < 4.4$.

Collision products will arrive at each set of scintillators roughly in coincidence, while beam halo products passing through the detector will appear distinctly separated.

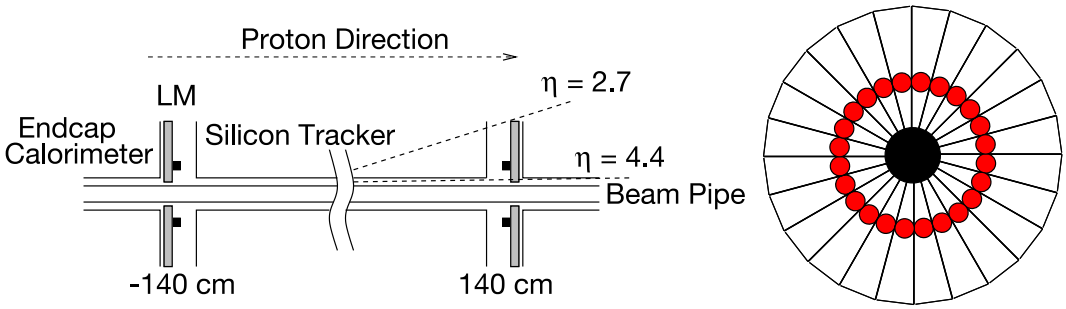


Figure 3.16: Schematic drawing showing LM. Left: the location of the LM detectors; Right: the geometry of the LM counters and the locations of the PMTs (solid red dots).

The luminosity is determined from:

$$\mathcal{L} = \frac{fN}{\sigma} \tag{3.10}$$

where, $f$ is the beam crossing frequency, $\sigma$ is the effective cross section for the luminosity monitor, and $N$ is the average number of inelastic collisions per beam crossing measured by

the luminosity monitor. Since $N$ is typically greater than one, it is important to account for multiple $p\bar{p}$ collisions in a single beam crossing. This is done by counting the fraction of the beam crossings with no collisions and using Poisson statistics to determine $N$.

### 3.2.6 Muon System

The outer most layer of the DØ detector is the muon system (Figure 3.6) [87]. A special detector is required to measure muons because muons do not undergo hadronic interactions but only loose energy by ionization, as they are the only charged particles that can pass through the several meters of highly dense material present in the calorimeter. The muon detector has two active regions called the central region for $|\eta| < 1$ and the forward region for $1 < |\eta| < 2$. The system also employs a 2T toroid iron magnet to bend the muons from their original paths. The addition of the magnetic field helps to provide a local momentum measurement in the event in which momentum can not be determined from the tracking detector. Additional shielding surrounding the beam pipe near the forward muon detector is designed to reduce the spurious beam effects and dramatically reduces the amount of radiation to which detector is exposed. A schematic of the muon system and the beam shielding can be seen in Figure 3.17.

The muon system at DØ is a three layer (A, B and C layers with A being the innermost layer) detector, both in the central and forward regions, consisting of drift chambers for precise position measurement and scintillator counters for muon identification and fast triggering as shown in Figure 3.18. The scintillator counters produce light when the muon passes through the detector which is then collected by a photo-multiplier tube. The drift chambers have a central wire held at a large voltage surrounded by an inert gas. As the muon enters the chamber it will ionize the gaseous organic compound mixture and the resulting free charges will drift towards the wire. The position of the muon is found by analyzing the current profile in the wire. In the central region, the drift chambers are called PDTs (proportional drift tubes) and are rather large with typical areas of $2.8 \times 5.6\ m^2$. The forward

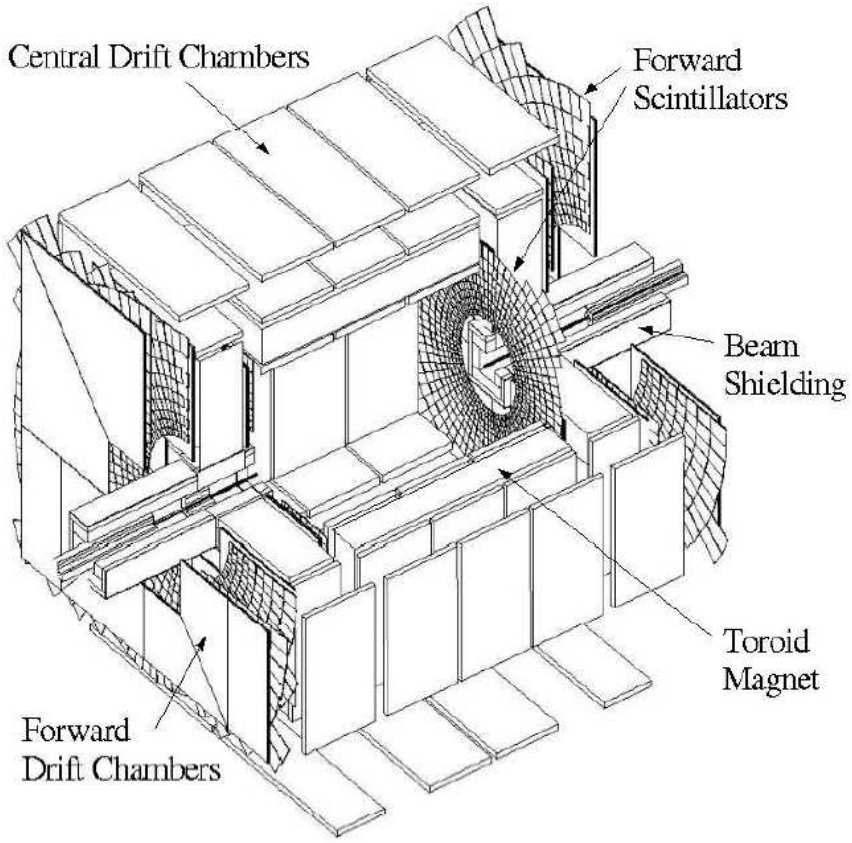


Figure 3.17: 3D view of the DØ muon detector.

region uses smaller drift chambers called MDTs (mini drift tubes), which are a collection of eight cells of size $9.4 \times 9.4\ mm^2$. The position resolution of the drift chambers is $\sim 1$ mm.

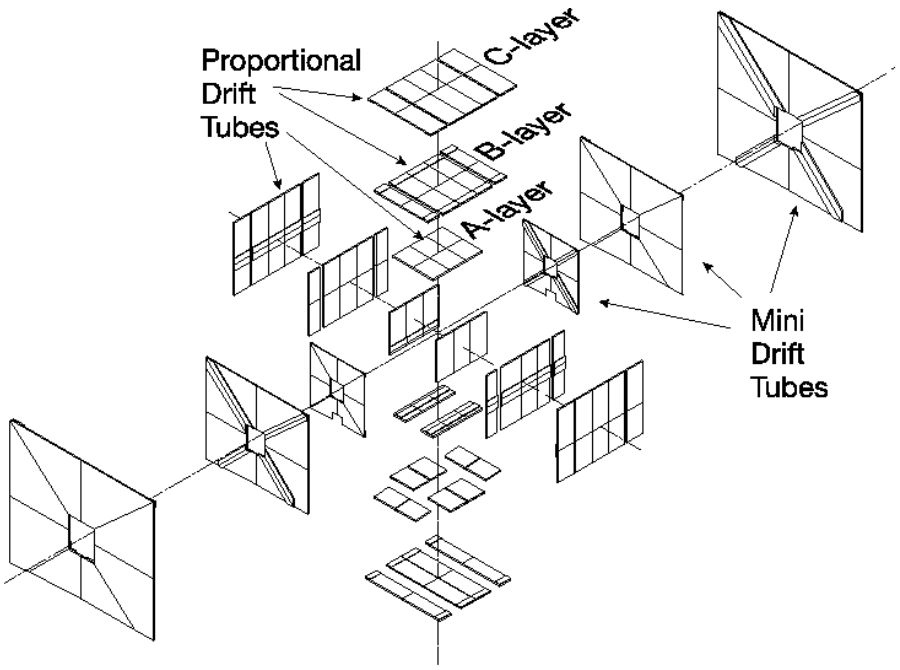


Figure 3.18: An exploded view of the DØ muon drift tube system. From the interaction point, muons traverse the A, B and C layers.

### 3.2.7 Trigger and Data Acquisition System

The proton-antiproton beams produce about 1.7 million collisions per second at the center of the DØ detector. The information collected for each collision is called an event. As most of the collisions do not produce interesting physics, not every event needs to be saved to the tape. Also, due to the constraints of cost for data storage and computing power, only selected events will make it to permanent data storage to be used in an analysis. All these points lead to the necessity of having an efficient trigger system. The simple layout of the DØ trigger system [94, 95] consists of three levels shown in Figure 3.19 with each succeeding level examining fewer events but in greater detail and with more complexity.

The first level consists of hardware components, the second level uses both software and hardware, and the third level, which does a full event reconstruction, is based on software

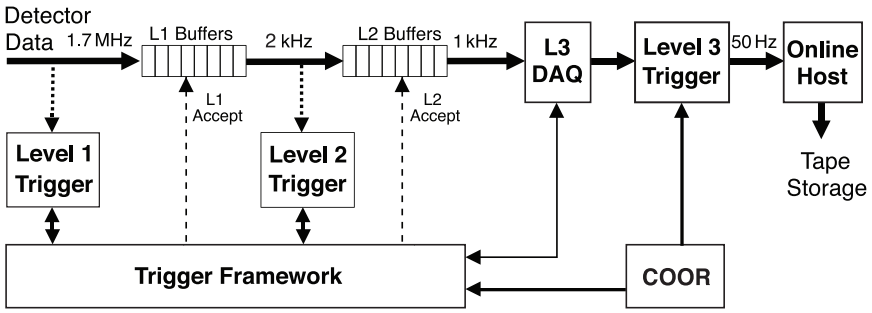


Figure 3.19: Overview of the DØ trigger and data acquisition systems.

only.

### The Level 1 trigger

Level 1 uses hardware to find tracks in the CFT/PS (called CTT trigger), transverse energy deposits and global missing energy in the calorimeter, and muon tracks in the muon system. The muon tracks can be required to be matched with CTT tracks, or trigger on the tracks independently. The level 1 trigger has only 4.2 $\mu$s to make a decision and the rate should be reduced to 2 kHz, from an incoming rate of about 1.7 MHz.

### The Level 2 trigger

The Level 2 system is comprised of two stages, a preprocessor stage and a global trigger stage. The preprocessor identifies the objects such as tracks, electrons, jets and muons. The global stage allows the first opportunity to examine the correlation between objects, such as tracks and leptons. The Level 2 trigger has a time budget of 100 $\mu$s and must reduce the rate to 1 kHz.

### The Level 3 trigger

The Level 3 trigger/data acquisition system is implemented entirely in software which performs an approximate reconstruction of the event and makes a trigger decision using the full event information. The Level 3 trigger has a time budget of 100 ms and must reduce the

readout rate to 50 Hz.

The trigger system is closely integrated with the read out of data. Each event that satisfies the successive L1 and L2 triggers is fully digitized, and all of the data blocks for the event are transferred to a single commodity processor in the L3 farm. The L1 and L2 buffers play an important role in minimizing the experiment's deadtime by providing FIFO (First In First Out) storage to hold event data awaiting a Level 1 decision or awaiting transfer to Level 3.

# Chapter 4

# Event Reconstruction and Object Identification

***Everything should be made as simple as possible, but not simpler.***

- Albert Einstein.

The DØ detector described in Chapter 3, is designed to distinguish and define fundamental objects such as tracks, electrons, muons, and jets. However, the raw data from the detector are just digitized pulse height and time information. And hence a very large amount of information is recorded for every event that passes the final level of the triggering system. These data must be processed by event reconstruction algorithms to produce physics objects and their properties, representing the particles that originated from the $p\bar{p}$ collisions. The collection of complex reconstruction algorithms is referred to as DØ Offline Reconstruction Program (*d0reco*) and this process is called object identification. The algorithms process either collider events recorded during online data taking or simulated events produced with the DØ Monte Carlo (MC) program. This is a step-by-step procedure which involves (a) decoding of the raw information of individual detectors to associate electronics channels with physical detector elements, including the application of detector specific calibration constants; (b) This decoding information is used to reconstruct the clusters of energy eg.

from calorimeter and hits from the tracking detectors; (c) The next step is to build global tracks out of the hits in the SMT and CFT detectors and then to search for primary and secondary vertex candidates indicating the locations of $p\bar{p}$ interactions; (d) The results of all the above algorithms are then used to finally identify the standard physics candidates (photons,electrons, muons, neutrinos and jet candidates). This chapter describes the algorithms and procedures used to reconstruct and identify the physics objects used throughout this thesis.

## 4.1 Particle Tracks

As charged particles traverse through the tracking system (SMT and CFT), their paths are curved by the magnetic field of the solenoid and along these paths, these particles deposit a small amount of energy in many layers of material in the tracker, called "hits." The collection of hits from a single particle form a particle track. The DØ tracking algorithms reconstruct a particle track from the collection of hits in any given event. It is very difficult task to recognize which hit is associated which track in any given event as there are tracks from the secondary collisions and random electric noise.

To perform this difficult task two different algorithms are used to find tracks from the detector hits: The Histogram Track Finder (HTF) [96], The Alternative Algorithm (AA) [97], and a combination of the HTF and AA called Global Track Reconstruction (GTR).

**HTF**: The Histogram Track Finder algorithm [96] takes 2D hits and uses them to find the most likely values of curvature and trajectory angles for track candidates. For each hit, the algorithm calculates a value of the curvature, $\rho$, and the direction of the track at the point of closest approach to the beam spot, $\phi$, for a track constrained to pass through the beam spot. The curvature is calculated with

$$\rho = \frac{qB}{p_T} \tag{4.1}$$

where $q$ is the electric charge and $B$ is the magnetic field. A 2D histogram is constructed with $\rho$ on one axis and $\phi$ on the other, and the bin corresponding to the calculated ($\rho$,$\phi$) coordinate is incremented. This is done for all hits in the event. Hits belonging to the same track will contribute to the same peak, and hits belonging to different tracks will give a randomly distributed background. The Hough transform improves the fit by taking errors in measured values into account, giving ranges of $\phi$ for each hypothetical $\rho$ value considered for each hit. This is described in detail in Ref. [96]. The result is a series of lines that intersect at a particular ($\rho$,$\phi$) value. The histogram bins with too few hits are discarded, as are bins for which all hits are contained in neighboring bins. The bins that remain are used to form candidate tracks.

**AA**: The Alternative Algorithm begins by generating a pool of track candidates using the hits in the SMT. The algorithm selects all sets of three hits which lie along a path originating from the beam spot. It then extrapolates the path of the track outward to either the next layer of the SMT or to the CFT to calculate the point where the track should have crossed the next layer. This algorithm checks whether there is a hit near that location, and then extrapolates to the next layer, and repeats the procedure. At each layer a $\chi^2$ of hits with respect to the track is calculated and the hit becomes associated with the track if its $\chi^2$ is less than a certain value. If there is no hit in the layer, the algorithm continues and records a "miss" for this track. At the end of this procedure, a list of tracks is produced along with hits, misses, and $\chi^2$-s. If the number of hits is less than 3 in the SMT, the AA finds "CFT only" tracks.

**GTR**: In Global Track Reconstruction, the reconstructed tracks from both the HTF and AA algorithms are used and the list of tracks is passed through a track fitter based on the Kalman filter algorithm [98, 99] to obtain the final set of tracks in the event.

## 4.2 Primary Vertices

The point in space at which the proton and antiproton collision of the hard scattering occur is referred to as the "Primary Vertex" (PV). The hard scatter interaction vertex is very important to locate to allow discrimination of physics objects resulting from the $p\bar{p}$ collision and objects created from noise in the detector or other low energy $p\bar{p}$ inelastic collisions.

The location of the PV is reconstructed by means of an adaptive primary vertex algorithm [100]. This algorithm mainly consists of three steps: *track selection*, *vertex fitting* and *vertex selection*.

**Track Selection** removes poorly measured and fake tracks by requiring tracks to have $p_T > 0.5\ GeV$ and two or more hits registered in SMT, and assign the tracks to a vertex to which the extrapolated paths of the tracks point.

**Vertex Fitting** is a three-step process. In the first step, all selected tracks within each $z$-cluster are fitted into a common vertex using the Kalman Filter vertex fitting algorithm [99], to get an estimate of the beam position and width. In this fit, the track(s) with the highest $\chi^2$ contribution to the vertex are removed in turn, until the total vertex $\chi^2$ per degree of freedom is smaller than 10. In the second step, in each of the z-clusters, tracks with a distance of closest approach (dca) to the previously determined beam spot smaller than $5\sigma$ ($dca/\sigma(dca) < 5$), are preselected. Finally, the preselected tracks are fitted into a common vertex using the Adaptive vertex fitter algorithm. In this case each track receives a weight, shown in Eq. 4.2, that includes the $\chi^2$ of the previous track fit.

$$\omega_i = \frac{1}{1 + \exp^{(\chi_i^2 - \chi_{cutoff}^2)/2T}} \tag{4.2}$$

where the values for $\chi^2_{cutoff}$ and T are 16 and 4, respectively. The vertex fitting procedure is repeated until the difference of weights from the previous iteration for each track is less than $10^{-4}$.

**Vertex Selection** consists in the identification of the hard-scatter and additional mini-

mum bias (MB) vertices in the event. The hard-scatter vertex is selected using a minimum bias[1] probability selection algorithm [101], where each vertex is assigned a probability to be consistent with a MB interaction. The vertex with the smallest MB probability is defined as the hard-scatter PV. The algorithm is based on the fact that tracks from hard interactions have a harder $p_T$ spectrum than tracks from MB interactions. The first step of the algorithm is to cluster all reconstructed vertices in the $z$ direction within 2 cm of each other, and select the highest multiplicity vertex in each cluster, so that split primary vertices [2] are removed. For every selected vertex, all tracks within some distance around the vertex are used to compute the MB probability.

## 4.3 Electromagnetic Clusters

Electromagnetic (EM) clusters are localized deposits of energy in the calorimeter consistent with a purely electromagnetic interaction. Prior to identifying calorimeter objects such as electrons, jets, and $\not\!\!E_T$ , it is necessary to apply algorithms to remove undesired cells i.e, to suppress noise. The first of these algorithms is the NADA algorithm [102], which is designed to remove calorimeter "hot cells". Hot cells are cells which contain spurious or excessive energies due to detector problems such as hardware failure, electronic noise, uranium noise, or argon contamination, or physics processes such as cosmic ray showers or backscattering of beam particles interacting outside the interaction region. The NADA algorithm removes cells with high $p_T$ if neighboring cells have energy below a threshold $p_T$.

Another algorithm called T42 [103] is also used to reduce calorimeter noise. Calorimeter noise, somewhat more subtle than hot cells, is caused by readout fluctuations and energy deposition from previous beam crossings (known as pile-up). The T42 algorithm uses thresholds which are multiples of the RMS of the noise distribution, $\sigma$, to determine whether to keep or reject calorimeter cells. Cells with energy greater than $2.5\sigma$ are kept if neighboring

[1] A minimum bias vertex is a vertex from an inelastic $p\bar{p}$ collision.

[2] These are low multiplicity vertices, close to the real high multiplicity PV, formed from poorly reconstructed tracks that do not get attached to (although they may come from) the real PV.

cells have energies expected to be from signal rather than noise. Cells are believed to have energies from signal if above +4 $\sigma$.

## 4.4 Electrons

An electron produced from some physics process traverses through the tracking system and stops in the ElectroMagnetic (EM) calorimeter. Electrons being charged particles, leave traces in the SMT and CFT while passing through the tracking system and these traces can be reconstructed as the track of the electron. In addition, when an electron enters the calorimeter, its energy is deposited in the calorimeter and it is fully absorbed. Track matching with energy cluster in the calorimeter distinguishes electron from photon since photon which is neutral does not leave hits in the tracking system. These two features can be used to reconstruct electrons. The energy deposits of electrons in the calorimeter are reconstructed as towers in the $\eta \times \phi$ space. These towers are sorted by $E_T$, with the highest $E_T$ tower defined as a seed. Then, all towers within a cone of $\Delta$R ($= \sqrt{\Delta\eta^2 + \Delta\phi^2}$) $< 0.2$ around a seed are added and become an EM cluster. Parameters for each EM cluster derived with calorimeter information and track matching are used to identify electrons.

1. **EM fraction ($f_{em}$)**: This is defined as the energy ratio of the cluster in the EM calorimeter to the total energy deposited in the calorimeter.

$$f_{em} = \frac{E_{em}(\Delta\mathrm{R} < 0.2)}{E_{tot}(\Delta\mathrm{R} < 0.2)} \tag{4.3}$$

As a true EM object is expected to deposit most of its energy in the first few EM layers of the calorimeter. So, it is expected that electrons have large value of $f_{em}$. For a real electron, its value is close to one.

2. **Isolation**: EM objects should be isolated in $\eta \times \phi$ space from the surrounding clusters. A second cone with radius $\Delta\mathrm{R} < 0.4$ is defined and only 15% of the total energy within the

cone of $\Delta R < 0.4$ is allowed to be outside the $\Delta R < 0.2$ cone or to be non-Em energy:

$$f_{iso} = \frac{E_{tot}(\Delta R < 0.4) - E_{em}(\Delta R < 0.2)}{E_{em}(\Delta R < 0.4)} < 0.15 \tag{4.4}$$

where $E_{em}(\Delta R < 0.2)$ is the energy of EM object within $\Delta R < 0.2$ and $E_{tot}(\Delta R < 0.4)$ is the total energy deposited in the calorimeter within $\Delta R < 0.4$.

3. **The $\chi^2$ of H-Matrix**: The shower development of an electron in the calorimeter is different from hadronic particles. A $7 \times 7$ H-matrix quantifies how the shower development is similar to an electron. It is computed using seven correlated variables: the fraction of energy in each of the four EM layers, the shower width in $\Delta R$, the log of the total shower energy, the $z$ position of the primary vertex. The $\chi^2$ of the H-matrix ($\chi^2_{HM}$) is required to be smaller than 50.

4. **Track Match Probability**: Once the EM cluster in calorimeter is well identified, we enhance its identification by using information from the tracking system. The $\chi^2$ is calculated for the closest track to the electron cone and this $\chi^2$ is then converted into the probability that the track is associated with the given electromagnetic cluster.

5. **EM Likelihood**: Finally, a likelihood discriminant [104] is built in order to differentiate a cluster coming from an Electromagnetic object from one of an Hadronic object with a large EM fraction. A number of variables (noted as the vector $x$) are put together to define a likelihood that a track-matched electron object is really an electron:

$$\mathcal{L} = \frac{P_{sig}(x)}{P_{sig}(x) + P_{bkg}(x)} \tag{4.5}$$

where $P_{sig}(x) = \prod_i P_{sig,i}(x_i)$ and $P_{bkg}(x) = \prod_i P_{bkg,i}(x_i)$, that is, the probabilities for signal and background are the product of the probabilities for the individual variables. The associated track is the one with the highest $P(\chi^2_{spatial})$ that has $\delta\phi < 0.05$ and $\delta\eta < 0.05$. The version of the likelihood that we use is based on seven variables: EM fraction, the fraction of the transverse energy of the cluster in the calorimeter over the transverse momentum of the

matched track, H-matrix, the $\chi^2$ probability of track matching, distance of closest approach (dca), the total number of track within a cone of $\Delta R < 0.05$, and sum of $p_T$ of all tracks.

In the analysis presented in this thesis, the following electron definitions are used :

**Ultra-loose electron** This electron definition is used to model the multijet background in order to increase the acceptance. An ultra-loose electron is required to have $f_{em} > 0.9$, $\chi^2_{HM} < 50$, $f_{iso} < 0.2$ and $p_T > 15GeV$. There is no requirement for a matching track.

**Loose isolated electron** This electron definition is used to veto events with more than one isolated lepton. In addition to the ultra-loose requirements, the energy deposition in the calorimeter must be matched to a charged particle track from the tracking detectors with $p_T > 5GeV$ and $z(track, PV) < 1\ cm$. Also the isolation criteria is tightened to 0.15 i.e, $f_{iso} < 0.15$.

**Tight isolated electron** This is the main electron definition used in the analysis to select the electrons. A tight isolated electron must pass all the requirements of loose isolated electron and also have a value of EM-likelihood, $\mathcal{L}_{em} > 0.85$.

## 4.5 Muons

Muon, being minimum ionizing particles (MIP), do not produce showers like electrons or hadrons in the calorimeter. Therefore, muons are reconstructed and categorized based on the pattern of hits in the muon system as well as by having a spatially matched track in the central tracking system [105]. The muon system delivers the unambiguous muon identification and the tracking system provides precise momentum resolution and a high efficiency of finding tracks in the entire angular acceptance region of the muon system. Muons found only in the muon system are labeled as “local” and muons with an associated track in the central tracking are called “global”.

Muons are categorized depending on the location and types of hits in the muon system. These classifications are referred to as “tight”, “medium”, and “loose”. Tight muons require

drift tube and scintillator hits both inside and outside the toroid, and result from a local muon track match. Medium muons require drift tube and scintillator hits outside the toroid. In the bottom region, where muon detector coverage is reduced, this requirement is relaxed. Loose muons require only one reconstructed segment. Muons are further subdivided by the number of segments ($nseg$) or layers that a hit was found in: $nseg = 1$ requires at least one A layer (inside the toroid) hit, $nseg = 2$ requires at least one B or C layer (outside the toroid) hit, and $nseg = 3$ includes at least one hit in the A layer and one in the B or C layer. From these categories, many types and qualities of muons can be defined, and the accepted ones are determined by the needs of a given analysis.

In this analysis, following muon definitions are used :

**Loose isolated muon** A loose isolated muon used for vetoing events containing more than one isolated lepton is defined as:

• The quality of the muon must be medium $|nseg| = 3$;

• To veto against cosmic ray muons, a timing requirement is added requiring the scintillator hits to be within 10 $ns$ of the expected time for muons from the collisions to reach the scintillators.

• Muons are also required to pass medium track quality definitions. This requires a track match with the central tracker of $\chi^2/dof < 4$ and distance of closest approach to the primary vertex $< 0.2\ cm$ for a CFT track or $< 0.02\ cm$ for a SMT track.

• The $p_T(\mu)$ of the muon must be greater than 15 GeV; and

• $\Delta\mathrm{R}(\mu, jet) > 0.5$.

**Tight isolated muon** Tight isolated muons are the loose muons that must pass the additional isolation criterion called "TopScaledLoose," which means that the sum of the momenta of all other tracks in a cone of $\Delta\mathrm{R} < 0.5$ around the muon and the energy deposited in a cone of $0.1 < \Delta\mathrm{R} < 0.4$ around the muon must both be less than 20% of the muon $p_T$. The tight isolated muon definition is used as the standard muon definition throughout this analysis unless otherwise specified.

## 4.6 Jets

The strongly interacting particles such as quarks and gluons cannot exist alone as a free particles and undergo hadronization into colorless particles due to color confinement. These particles are detected as showers within a narrow cone in the calorimeter, called jets [107]. A jet will shower in the electromagnetic and hadronic calorimeters and its energy is measured by sampling this shower in the many layers of the DØ calorimeter (Figure 4.1). A proper measurement of the jet energy and direction is needed to determine the original quark or gluon energy and momentum.

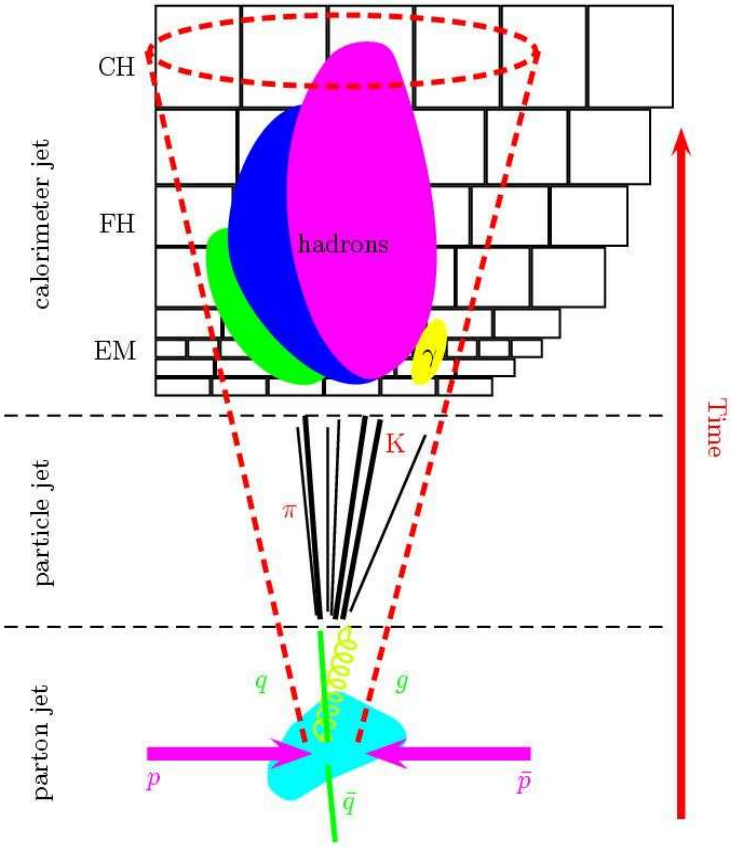


Figure 4.1: Diagram illustrating a reconstructed jet at the parton, particle and detector (calorimeter) levels.

To reconstruct the final jets some jet algorithms are used. The purpose of a jet algorithm is to reconstruct these jets from the energy deposits in the calorimeter towers in such a way

that the kinematic properties can be related to those of the outgoing quark or gluon. An ideal jet algorithm needs to fulfill certain requirements [107]. It needs to be Infrared (IR) and collinear safe. The jet algorithm must not be sensitive in its behavior to soft radiation. The ideal jet algorithm also needs to be collinear safe, which means that it should not be sensitive to collinear radiation. For example, if a jet algorithm depends on the energy deposit in just one tower as a seed, collinear radiation going into another cell could cause the energy to not be enough to form a seed. The algorithm should be invariant to boost in the longitudinal direction, which is not controlled in a hadron collider. The algorithm should be insensitive to the details of the final state (such as extra radiation). And of course, the algorithm should be experimentally well behaved: straightforward and efficient to implement, stable with luminosity, efficient, precise and unbiased.

At DØ, jet reconstruction is done with the "Improved Legacy Cone Algorithm" (ILCA, or Run II Cone Algorithm) [107, 108]. The cone algorithm in general is based on the idea of associating all particles that lie within a cone of radius $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ in $\eta \times \phi$ space. In order to reject calorimeter noisy cells an algorithm called T42 [103] is used as discussed in detail in Section. 4.3. The ILCA selects calorimeter towers with transverse energy $> 0.5$ GeV as seeds, and collects all calorimeter towers within a cone of radius $\Delta R = 0.5$ around the seed. If the cone has transverse energy greater than 1 GeV, it is defined as a jet candidate. The central axis of the jet candidate is an $E_T$ weighted midpoint of the towers. Next, a new cone is drawn around the new direction and the towers within the new cone are used to calculate the new central axis of the jet candidate, which is compared to the central axis of the old jet candidate. This process continues until the jet axis does not change appreciably or the maximum number of iterations (50) is reached. In the final step, the jet finding algorithm decides whether to merge or split jets that share energy. If the shared energy is less than 50% of the individual jet energies, jets are split into two distinct ones, otherwise they are merged.

After the reconstruction of jets, certain quality cuts are applied in order to distinguish

fake jets due to calorimeter noise from physical jets as described below:

• A cut on the fraction of the total calorimeter energy in the jet cone contained in the EM region is applied at $0.05 < f_{em} < 0.95$ (where $f_{em} = E_{em}/E_{tot}$), to remove electromagnetic particles from jets.

• Since the noise level is higher in the coarse hadronic calorimeter, a cut on the fraction of the jet energy deposited in the coarse hadronic calorimeter is applied at $f_{ch} < 0.4$ (where $f_{ch} = E_{ch}/E_{tot}$). This cut is designed to remove fake jets which are clustered around noise.

• A cut on the ratio of the energy in the most energetic cell of a jet to that of the second most energetic cell ($f_{hot}$) is applied at $f_{hot} < 10$ in order to remove jets clustered partially from noisy hot cells.

• $n90$ is defined as the number of calorimeter towers in a jet that contain 90% of the total energy of the jet. To remove jets clustered from a single hot tower, n90 is required to be greater than 1.

• A jet is required to be isolated from all electromagnetic clusters, $\Delta R > 0.5$.

**Jet Energy Scale**

After ensuring the quality of the reconstructed jets, we have to make sure that these jets do represents the real hard scattered quarks, which we are looking for. After scattering, these quarks undergo hadronization and form a jet of particles, and only after showering through the detector these form calorimeter jets as described in Figure 4.1. Usually, the jet energy measured in the calorimeter is not equal to the energy of the parton from which the jet is assumed to be arisen. This can be due to the non-linearities, dead material and showering effects in the calorimeter. The purpose of the DØ jet energy scale (JES) correction is to correct the calorimeter jet energy back to the stable-particle jet level before interaction with the detector. JES attempts to correct the reconstructed jet energy, $E_{jet}^{raw}$, back to the particle level energy that is the jet energy before the interaction with the calorimeter, $E_{jet}^{corr}$ [109].

The correction can be written as

$$E_{jet}^{corr} = \frac{E_{jet}^{raw} - O}{R_{jet} F_{\eta} S} \tag{4.6}$$

The different components in Equation. 4.6 are described below:

• *Offset Energy*, $O$ Energy in the clustered cells from electronic noise, pile-up, the underlying event, multiple interactions, and noise from radioactive decay of the uranium in the calorimeter. The offset correction is measured by averaging over minimum bias data samples and averaging over $\phi$, and is given as a function of detector $\eta$ and instantaneous luminosity.

• *Calorimeter Response*, $R$ is a measure of the calorimeter response to a jet. It is less than one because of the energy loss in the detector before the calorimeter and the poorer response of calorimeters to hadrons than electrons. In addition, the measured jet energy can be distorted due to the non-linear response of the calorimeter to the particle energies, a different response of the calorimeter to different particles, and un-instrumented regions of the detector or dead material. This response ($R$) is determined using the transverse energy balance in back-to-back photon+jet events. The transverse energy of the photon is measured very precisely and provides the target transverse energy for the jet. This correction is derived as a function of the jet energy and detector $\eta$.

• *Inter-calibration*, $F_{\eta}$ is a calibration factor applied to make the response uniform as a function of jet $\eta$ across the central and end-cap calorimeters, and the inner cryostat regions. The size of this correction for a typical jet in this analysis ia around 5%

• *Showering Correction*, $S$ is a measure of the energy that radiates outside of the cone during the shower development in the calorimeter. Furthermore, the solenoid field can change a particle's trajectory. Therefore, the showering correction, $S$ corrects for the net energy difference due to such showering effects.

Further, in order to take into account the differences between Monte Carlo and data for jet resolution, jet reconstruction efficiencies and identification efficiencies, a method called "Jet shifting, smearing and removal" (JSSR) is introduced [110, 111]. JSSR does a re-calibration,

smearing and discarding of simulated jets, after which the behavior is similar to jets in data.

## 4.7 $b$ Jets

Identifying jets arising from the hadronization of $b$ quarks is very important for this analysis as we have two $b$ quarks in the final state. Thus, an efficient $b$ quark jets selection will suppress the backgrounds significantly and enable us to make purer single top events samples. After hadronization, $b$ quarks form a $B$ hadron which is a bound state of a $b$ quark and one or two light quarks. Due to slightly longer lifetimes than lighter hadrons, $B$ hadrons can travel a few millimeters in the detector before decaying, resulting in a decay vertex usually displaced from the primary interaction point that can be reconstructed as a secondary vertex. Another peculiar property is that about 20% of $b$ jets contain a muon inside the jet cone. These distinguishing features and other kinematic properties can be used to identify $b$ jets from other light quark jets, the process of identifying $b$ jets is called *b-tagging*.

A neural network (NN) $b$-jet tagger developed by the DØ $B-ID$ group is used to identify $b$-jets [112]. First, the jets are required to be "taggable" and then the jets are "tagged". Some quality requirements, such as jets must have at least two good quality tracks with SMT hits pointing to a common origin, are first imposed to reject different tracking efficiencies, badly reconstructed jets and detector effects, in general. This quality selection is called "taggability". The NN $b$-tagging algorithm uses seven variables, ranked in order of separation power, to discriminate $b$ jets from other jets: (a) decay length significance of the secondary vertex; (b) weighted combination of the track's impact parameter significance; (c) probability that the jet originates from the primary vertex; (d) $\chi^2$ per degree of freedom of the secondary vertex; (e) number of tracks used to reconstruct the secondary vertex; (f) mass of the secondary vertex; and (g) number of secondary vertex found in the jet.

There are in total 12 different operating points provided by the $b$-tagger, each defined by a cut on the output of the NN tagger. For the analysis presented in this thesis the following

operating points were used:

• Events that contain exactly two taggable jets passing the OLDLOOSE working point ($NNoutput > 0.5$) are considered to be "Two Tag" events.

• For "One Tag" events, exactly one jet passes the TIGHT ($NNoutput > 0.775$) $b$-tagging cut and no other jet passes the OLDLOOSE cut ($NNoutput > 0.5$)

The veto of a second OLDLOOSE candidate in the one tag definition ensures that there is no overlap between the one tag and two tags samples.

## 4.8 Missing Transverse Energy, $\not\!\!E_T$

Neutrinos are weakly interacting particles which traverse the detector without giving rise to any electronic signal. So, neutrinos are not detected at all in the DØ detector, and their presence must be inferred from energy imbalance in the transverse plane. Before the collision, the proton and antiproton only have momentum in the $z$ direction, so the momentum in the $x-y$ plane, the transverse momentum, is *zero*. Conservation of momentum guarantees that to be the case after the collision as well. This fact can be used to infer the presence of particles that escape detection, with neutrinos being a prime example. The missing transverse energy, $\not\!\!E_T$, is the negative of the vectorial sum of the transverse momenta of particles observed in the detector. In practice, we compute the $\not\!\!E_T$ by adding up vectorially the transverse energies in all cells of the EM and FH layers of the calorimeter [106]. There are some corrections to be applied to the missing transverse energy for the following cases:

• The missing transverse energy needs to be corrected if there are reconstructed muons in the event. Since muon deposit a small amount of its energy in the calorimeter, which must be subtracted.

• Similarly, the jet energy scale correction changes the balance in the transverse plane. The momentum component added due to the jet energy scale for each jet needs to be subtracted from the raw $\not\!\!E_T$.

• The $\not\!\!E_T$ is corrected for the coarse hadronic calorimeter energy belonging to the jets in the event. Since due to large noise presence, the energy in the coarse hadronic calorimeter is not taken into account.

• There are also small corrections needed if there are electrons and photons in the event due to the electron and photon energy scales.

# Chapter 5

# Samples, Event Selection and Uncertainties

***Probable impossibilities are to be preferred to improbable possibilities.***

- Aristotle.

This chapter describes the data and monte carlo (MC) samples used in this analysis. The MC generation and different MC corrections needed to properly model the DØ data is discussed in Section 5.2. This section also describes the set of selection cuts which are applied to remove mis-measured events or events which are unlikely single top quark candidates and at the same time trying to minimize physics processes that can mimic the signal event signature. In general, the cuts are designed to select events with one high $p_T$ lepton from the $W$ boson decay, large missing $\not\!\!E_T$ indicating a neutrino in the final state, and two to four jets. Kinematic distributions between data and the sum of all background models are compared after the event selection and all the corrections to validate the modeling. This chapter also gives details about the systematic uncertainties used in the analysis.

## 5.1 Data Sample

The data sample analyzed in this thesis was collected from Tevatron $p\bar{p}$ collisions at center-of-mass energy $\sqrt{s} = 1.96$ TeV using the DØ detector. Only a fraction of the data delivered by the Tevatron is recorded by the DØ detector. Out of this recorded data only the fraction for which all detector systems are functioning well is used for the data analysis which is called "good data". The selection of good data i.e, data quality monitoring is performed on two levels, *online* and *offline*. It is crucial for a high data taking efficiency to catch the malfunction of detector components, of the readout or the triggering as early as possible. Online data quality monitoring guarantees an immediate reaction to a problem, thus maximizing the good data that gets recorded. However, there are data quality issues which are not recognized online. The remaining deficient data is eliminated by offline data quality monitoring. Finally, only the fraction of the data which is reconstructed is actually used for the analysis presented. This amounts for approximately 80% of the data delivered by the Tevatron.

The data sample used in the analysis comprises of data collected in two run periods. Run IIa (also known as "p17" from the reconstruction version used) consists of data taken between August 2002 and February 2006 (Run numbers 151831-215670) and Run IIb data (also known as "p20" from the reconstruction version used) was taken for the period from June 2006 until June 2009, which in total corresponds to $\sim 5.4\ fb^{-1}$ of good data. The Run IIb data was recorded at higher instantaneous luminosities, and with the upgraded detector and hence requires the subsequent trigger changes. The integrated luminosity, run periods, and trigger versions used for data in this analysis are shown in Table. 5.1.

### 5.1.1 Triggers

To increase the significant signal acceptance gain, data used in the analysis is required to pass through any "reasonable" trigger where a reasonable trigger is any trigger except $b$-tagging, gap, and forward proton triggers and in the case of the muon channel, an EM

| Data sample | Trigger Version | Integrated Luminosity [pb$^{-1}$] |
|---|---|---|
| Run IIa | v8 – v14 | 1078.81 |
| Run IIb | v15a | 534.44 |
| | v15b | 688.02 |
| | v15c | 397.31 |
| | v16 | 2661.90 |
| Total | v8-16 | 5360.48 |

Table 5.1: Integrated luminosities for the datasets used in this analysis.

trigger and in the case of the electron channel, a muon trigger. 814 triggers are used for the Run IIa electron channel, 413 triggers for the Run IIa muon channel, 647 triggers for the Run IIb electron channel, and 490 triggers for the Run IIb muon channel and the ORing of all these triggers is known as the "Mega-OR" where the Single-Lepton-OR is a subset of the Mega-OR. The Single-Lepton-OR is defined as any trigger that finds either a single muon or electron in the event and is modeled using turn-on curves that give the efficiency of an event to pass the trigger requirements at all 3 levels parameterized in $p_T$ , $\eta$, or $\phi$. Since a large number of triggers were used, the trigger efficiency was assumed to be 100% for single top signal events that passed the offline selection. This was verified by comparing two ratios: first, the ratio of Single-Lepton- OR selected data to Mega-OR selected data, and second, the ratio of MC simulation that has the Single-Lepton-OR turn-on curves applied to MC simulation that is 100% efficient. If the ratio of the two ratios is around one, then assuming that the Single-Lepton-OR is modeled correctly, the Mega-OR trigger efficiency is 100%. Figure 5.1 shows an example of these ratios as a function of a topological variable sensitive to the trigger selection: Jet1 pT which is the leading jet transverse momentum.

## 5.2 Monte Carlo Modeling

In order to understand the experimental data and for the cross section measurement, it is very important to identify the background processes involved and to properly model them.

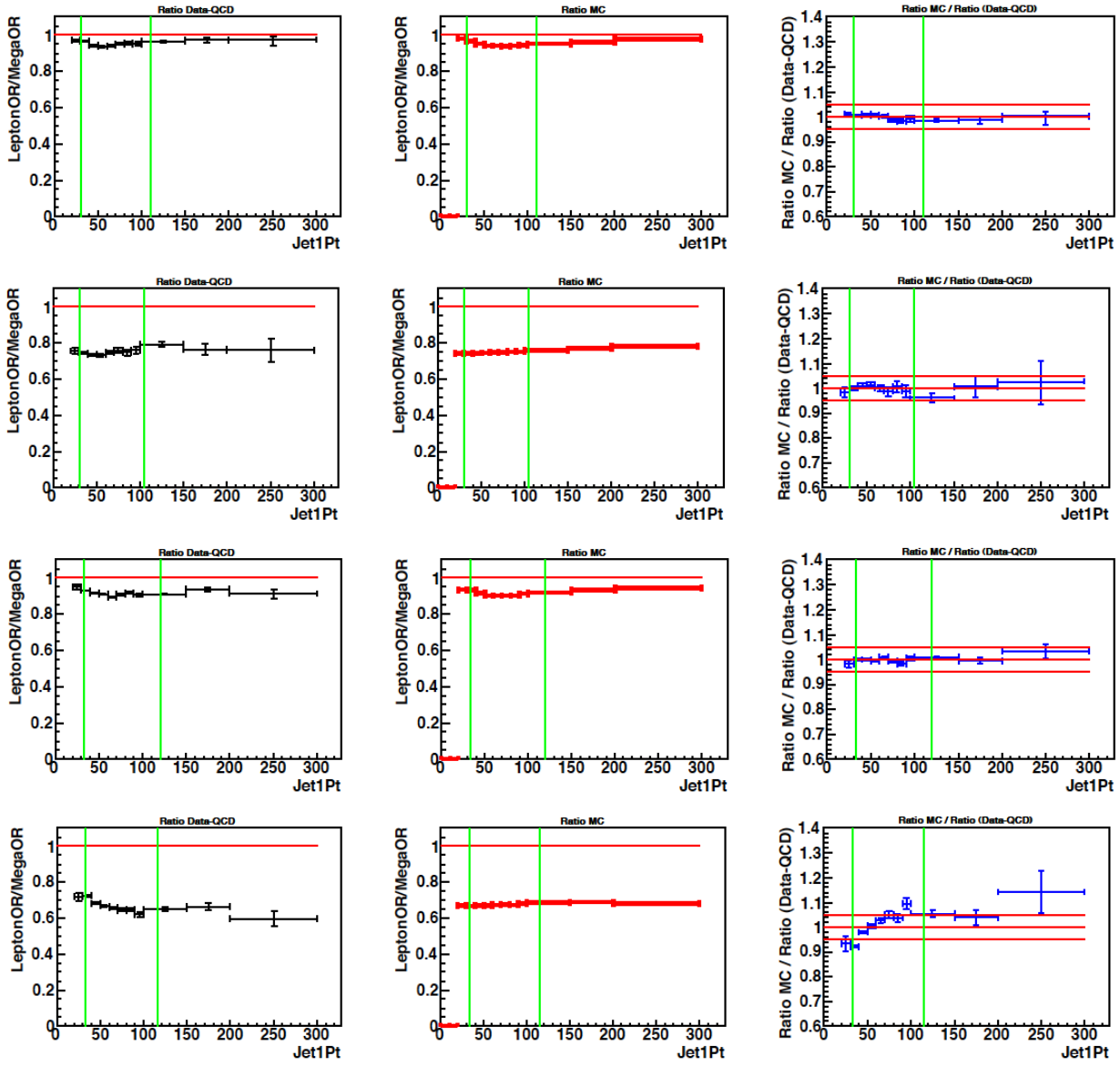


Figure 5.1: Ratio of Single-Lepton-OR to Mega-OR in Data-minus-QCD (left column); ratio of Single-Lepton-OR corrected MC yields to the total sum of MC yields (center column); ratio of the previous two ratios, namely the Mega-OR trigger efficiency (right column) as a function of Jet1 $p_T$ . The rows correspond to Run IIa electrons, Run IIa muons, Run IIb electrons, and Run IIb muons.

As discussed in Section 2.3.3 the main backgrounds for the single top signal events are $W$+jets, $t\bar{t}$, $Z$+jets, dibosons and multijets. Also it is important to simulate the signal events. All the background and signal samples are simulated with Monte Carlo (MC) event generators. The steps involved in the MC event simulation are shown in the Figure 5.2. The first step involved in the MC generation, is the creation of the hard scattered parton $4-vectors$ ($E_x$,$p_x$,$p_y$,$p_z$). This is accomplished by various event generator packages (for example, PYTHIA, shown in Figure 5.2) available depending upon the calculations of the matrix elements for a process, and using the proton parton distribution functions (PDF) generated by the CTEQ group [113]. Then, all stable particles produced are passed through a full detector simulation that models the interactions between the particles and the material in the detector using GEANT [114]. The showering of each hadron through all detector material, including the many liquid argon-uranium layers of the calorimeter, is modeled, as well as the energy deposition at each step. A simulation program known as DØGSTAR [115] accomplishes this task, and another package known as DØSim [116] simulates the detector electronics response to the deposited energy. In addition to the deposited energy, electronic noise is added for all detector systems. Further, the TrigSim package models the DØ trigger system. The trigger electronics and the effects of the trigger on the data selection is applied for L1 and the filtering code which is used at L2 and L3. The output which is obtained from this stage is similar as from DØ data acquisition system. And hence the final step in the Monte Carlo generation process is to reconstruct the event in the same way as a real data event is reconstructed (see Chapter 4). The event generators and the settings used for all signal and background processes modeled by MC are stated below.

### 5.2.1 Monte Carlo Event Samples

#### Signal MC Events Generation

The CompHEP-SingleTop Monte Carlo event generator [117] is used to generate the single top signal events and it is modeled in a way to reproduce next-to-leading order (NLO)

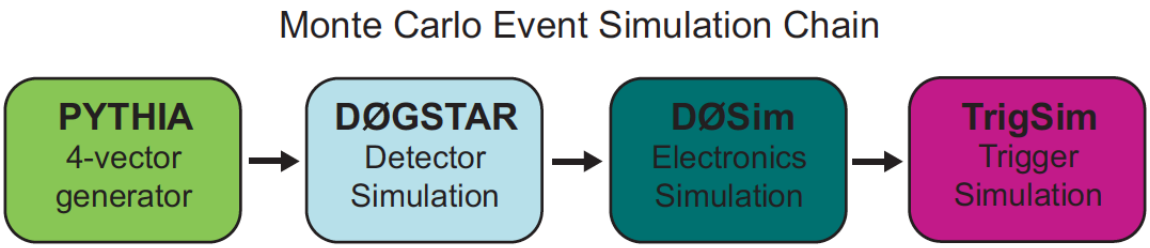


Figure 5.2: The chain of packages used to generate Monte Carlo events. PYTHIA is one example of many 4-vector generators, and in practice, TrigSim is replaced by a trigger efficiency calculation.

kinematics [48] using modified leading order (LO) generators. For the $s$-channel simulation at LO, reproduces NLO kinematics without any changes [48] but in case of the $t$-channel some special modeling is required. Simulated events from the $2 \to 2$ calculations are kept if $p_T(\bar{b}) \leq 10\ GeV$ and ones from the $2 \to 3$ process are used if $p_T(\bar{b}) > 10\ GeV$. The $2 \to 2$ process is scaled by a $K$ factor to make the rates at the cut-off point match: $K = 1.21$. [117]. Finite widths for the top quark ($\sim$ 1.5 GeV) and $W$ boson ($\sim$ 2.0 GeV) and a top quark mass of 172.5 GeV is used for signal simulation. In all signal models, the top quarks and their daughter $W$ bosons are decayed at the time of production, before later processing with PYTHIA [118], so that all spin properties of the top quarks are preserved in the angular correlations of the final decay products.

For modeling the parton kinematics in the protons and antiprotons, CTEQ6M next-to-leading-order parton distribution functions (PDF) [113] and factorization scale of $m_{top}^2$ for $s$-channel and $(m_{top}/2)^2$ for $t$-channel is set. To add the underlying event from the $p\bar{p}$ interaction, the initial-state and final-state radiation, and to hadronize and fragment the final state quarks and gluons into jets, PYTHIA is used. They both use TAUOLA [119] to decay tau leptons and EvtGen [120] is used to model the decay of $b$-hadrons.

**Background MC Events Generation**

All background samples ($W$+jets, $Z$+jets, $t\bar{t}$ and dibosons) except multijets are simulated using Monte Carlo models. A leading order matrix element event generator, ALPGEN [121] version 2.11 is used to model $W$+jets, $Z$+jets and $t\bar{t}$ backgrounds. The version of ALPGENused includes parton-jet matching [122] to avoid double counting some regions of jet kinematics. The samples are generated in the following sets ($lp$ = light partons): $W + 0lp$, $W+1lp$, $W+2lp$, $W+3lp$, $W+4lp$, $W+ \geq 5lp$ (this set includes $W$+single massless charm); $Wc\bar{c} + 0lp$, $Wc\bar{c} + 1lp$, $Wc\bar{c} + 2lp$, $Wc\bar{c}+ \geq 3lp$; and $Wb\bar{b} + 0lp$, $Wb\bar{b} + 1lp$, $Wb\bar{b} + 2lp$, $Wb\bar{b}+ \geq 3lp$, which are summed weighted by the ALPGENleading log (LL) average cross section for each subset. Further, for $W$+jets, CTEQ6L1 [113] PDFs were used with a factorization scale at $m_W^2 + \sum_{alljets} m_T^2$ where $m_T$, the transverse mass, is defined as $m_T^2 = m^2 + p_T^2$ and extends to all final state partons excluding the $W$ decay products. Furthermore, the $W$+light-parton ($Wlp$) jet samples, as defined above, have additional parton-level cuts on the light partons, such as $p_T(lp) > 8\ GeV$ and $\Delta\mathrm{R}(lp; lp) > 0.4$ for all massless partons. This is done to avoid divergences in the cross section. Events with two $c$ quarks after parton showering are removed from the $Wlp$ and $Wbb$ samples, as well as events with $b$ quarks in the $Wlp$ sample, such that there is no phase-space overlap between the samples [123].

The $Z$+jets samples are generated similarly to the $W$+jets samples. The $Z$ bosons are set to decay leptonically, and the factorization scale used is $m_Z^2 + \sum m_T^2$. Separate samples for the $Zlp$, $Zbb$ and $Zcc$ processes are generated with up to four partons in the final state.

The $t\bar{t}$ samples either have one of the $W$ bosons decaying to $l\nu$ while the other decays to two quarks ($l + jets$), or both $W$ bosons decay leptonically ($dilepton$). Matrix elements for $t\bar{t}$ production with 0 to 2 additional light partons are used. The top quark mass is set to 172.5 GeV (just as for the signal sample), and the factorization scale to $m_T^2 + \sum p_T^2(jets)$.

The diboson samples were generated with PYTHIA with inclusive decays. The samples generated are $WW$, $WZ$, and $ZZ$. All these MC samples have the duplicate events removed. Table 5.2 shows the number of generated MC events and the cross section for each process.

| | The Monte Carlo Event Sets | | | |
|---|---|---|---|---|
| Event Type | Cross Section [pb] | Branching Fraction | No. of p17 Events | No. of p20 Events |
| Signals | | | | |
| $tb \to \ell$+jets | $1.04^{+0.04}_{-0.04}$ | $0.3240 \pm 0.0032$ | 0.6M | 0.5M |
| $tqb \to \ell$+jets | $2.26^{+0.12}_{-0.12}$ | $0.3240 \pm 0.0032$ | 0.6M | 0.5M |
| **Signal total** | $\mathbf{3.30^{+0.16}_{-0.16}}$ | $\mathbf{0.3240 \pm 0.0032}$ | **1.2M** | **1.0M** |
| Backgrounds | | | | |
| $t\bar{t} \to \ell$+jets | $7.46^{+0.48}_{-0.67}$ | $0.4380 \pm 0.0044$ | 1.5M | 1.3M |
| $t\bar{t} \to \ell\ell$ | $7.46^{+0.48}_{-0.67}$ | $0.1050 \pm 0.0010$ | 1.5M | 1.3M |
| **Top pairs total** | $\mathbf{7.46^{+0.48}_{-0.67}}$ | $\mathbf{0.5430 \pm 0.0054}$ | **3.0M** | **2.6M** |
| $Wb\bar{b} \to \ell\nu bb$ | 90.5 | $0.3240 \pm 0.0032$ | 2.7M | 3.0M |
| $Wc\bar{c} \to \ell\nu cc$ | 260 | $0.3240 \pm 0.0032$ | 2.7M | 3.0M |
| $Wjj \to \ell\nu jj$ | 23,831 | $0.3240 \pm 0.0032$ | 55M | 97M |
| **W+jets total** | **24,182** | $\mathbf{0.3240 \pm 0.0032}$ | **60.4M** | **103M** |
| $Zb\bar{b} \to \ell\ell bb$ | 38.7 | $0.03366 \pm 0.00002$ | 0.7M | 0.7M |
| $Zc\bar{c} \to \ell\ell cc$ | 106 | $0.03366 \pm 0.00002$ | 0.7M | 0.7M |
| $Zjj \to \ell\ell jj$ | 7,032 | $0.03366 \pm 0.00002$ | 14M | 4.0M |
| **Z+jets total** | **7,177** | $\mathbf{0.03366 \pm 0.00002}$ | **15.4M** | **5.4M** |
| $WW \to$ anything | $11.6 \pm 0.4$ | $1.0 \pm 0.0$ | 2.0M | 0.7M |
| $WZ \to$ anything | $3.25 \pm 0.11$ | $1.0 \pm 0.0$ | 1.0M | 0.6M |
| $ZZ \to$ anything | $1.33 \pm 0.04$ | $1.0 \pm 0.0$ | 1.0M | 0.5M |
| **Diboson total** | $\mathbf{16.2 \pm 0.6}$ | $\mathbf{1.0 \pm 0.0}$ | **4.0M** | **1.8M** |

Table 5.2: The cross sections, branching fractions, and initial numbers of events in the Monte Carlo event samples. The symbol $\ell$ stands for electron plus muon plus tau decays.

### 5.2.2 Event Selection Cuts

A loose event selection is applied to find $W$-like events containing an isolated lepton, missing transverse energy, and two to four jets with high transverse momentum. A series of topological cuts are imposed to reduce the multijet background in each channel to an acceptable level.

The samples after this selection, which is called "pre-tagged" are dominated by $W$+jets events, with some $t\bar{t}$ contribution that becomes more significant for higher jet multiplicities, and smaller contributions from multijets, $Z$+jets and diboson events. To be able to extract single top events, the presence of one or two $b$-tagged jets is also required.

**Common selection for both electron and muon channels**

- Good quality (for data).
- Instantaneous luminosity $> 0$.
- Pass MegaOR trigger requirement as discussed previously in Section. 5.1.1.
- Good primary vertex: $|z_{\rm PV}| < 60$ cm with at least three tracks attached.
- Two, three, or four good jets with $p_T > 15$ GeV and $|\eta^{\rm det}| < 3.4$. For Run IIb, jets are required to be vertex-confirmed, with $CPF0 > 0.85$ or $CPF0 = -1$, where $CPF0$ is the charge particle energy fraction for the primary vertex. Vertex confirmation requires that if a jet contains tracks, those tracks pointing back to the primary vertex must contain at least 85% of the energy of the jet.
- The leading jet $p_T > 25$ GeV.
- For Run IIb only, the second leading jet $p_T$ cut is raised to 20 GeV in events with exactly three good jets.
- For Run IIb only, the second leading jet $p_T$ cut is raised to 25 GeV in events with exactly four good jets.
- Jet triangle cut of $|\Delta\phi(\text{leading jet}, \not\!\!E_{\rm T})|$ vs. $\not\!\!E_T$:

- $|\Delta\phi|$ from 1.5 to $\pi$ rad when $\not\!\!E_T$ = 0 GeV, and $\not\!\!E_T$ from 0 to 35 GeV when $|\Delta\phi| = \pi$ rad

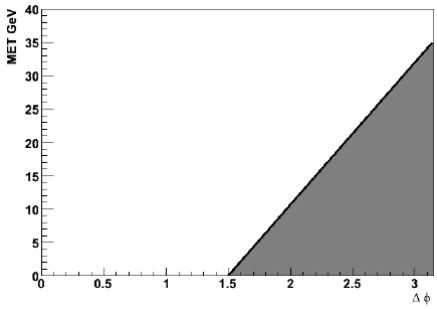


Figure 5.3: Figure showing the triangle cut between $\not\!\!E_T$ and $\Delta\phi$.

- Missing transverse energy
  - $20 < \not\!\!E_T < 200$ GeV in events with exactly two good jets.
  - $25 < \not\!\!E_T < 200$ GeV in events with three or more good jets.
- Scalar sum of the transverse energies of all good jets
  - $H_T$(alljets) $>$ 50 (55) GeV in events with exactly two good jets for Run IIa (Run IIb) period.
  - $H_T$(alljets) $>$ 75 (80) GeV in events with exactly three good jets for Run IIa (Run IIb) period.
  - $H_T$(alljets) $>$ 100 (110) GeV in events with exactly four good jets for Run IIa (Run IIb) period.
- Scalar sum of the transverse energies of all good jets, the lepton transverse momentum, and the missing transverse energy
  - $H_T(\text{alljets}, \text{lepton}, \not\!\!E_T) > 120$ GeV in events with exactly two good jets.
  - $H_T(\text{alljets}, \text{lepton}, \not\!\!E_T) > 140$ GeV in events with exactly three good jets.
  - $H_T(\text{alljets}, \text{lepton}, \not\!\!E_T) > 160$ GeV in events with exactly four good jets.

- Jet triangle cut of second leading jet $p_T$ vs. $H_T$(alljets) for Run IIb in events with exactly two good jets:
  - second leading jet $p_T$ from 0 to 27.5 GeV when $H_T$(alljets) = 0 GeV, and $H_T$(alljets) from 0 to 165 GeV when second leading jet$p_T$ = 0 GeV.

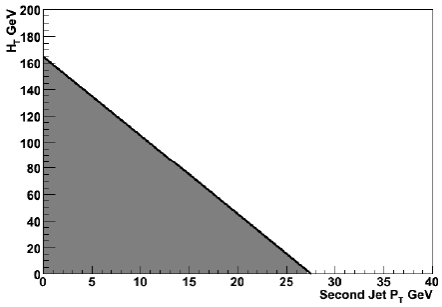


Figure 5.4: Figure showing the triangle cut between $H_T$ and *Second Jet* $p_T$.

**Electron channel selection**

- Only one tight electron with $|\eta^{\text{det}}| < 1.1$ and $p_T > 15$ (20) GeV in events with 2 (3 or more) good jets.
- No additional loose electron with $p_T > 15$ GeV.
- No tight isolated muon with $p_T > 15$ GeV and within $|\eta^{\text{det}}| < 2.0$.
- Electron coming from the primary vertex: $|\Delta z(e, \text{PV})| < 1$ cm.
- Electron triangle cuts of $|\Delta\phi(e, \not\!\!E_T)|$ vs. $\not\!\!E_T$
  - $|\Delta\phi|$ from 2 to 0 rad when $\not\!\!E_T = 0$ GeV, and $\not\!\!E_T$ from 0 to 40 GeV when $|\Delta\phi| =$ 0 rad.
  - $|\Delta\phi|$ from 1.5 to 0 rad when $\not\!\!E_T = 0$ GeV, and $\not\!\!E_T$ from 0 to 50 GeV when $|\Delta\phi| = 0$ rad.
  - $|\Delta\phi|$ from 2 to $\pi$ rad when $\not\!\!E_T = 0$ GeV, and $\not\!\!E_T$ from 0 to 24 GeV when $|\Delta\phi| = \pi$ rad.

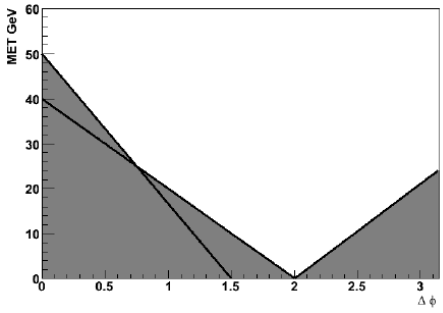


Figure 5.5: Figure showing the triangle cut between $\Delta\phi$ and $\not\!\!E_T$ for electrons.

**Muon channel selection**

- Only one tight muon with $p_T > 15$ GeV and $|\eta^{\text{det}}| < 2.0$.
- No additional loose muons with $p_T > 4$ GeV.
- No loose electron with $p_T > 15$ GeV and within $|\eta^{\text{det}}| < 2.5$.
- Muon coming from the primary vertex: $|\Delta z(\mu, \text{PV})| < 1$ cm.
- Muon triangle cuts of $|\Delta\phi(\mu, \not\!\!E_{\text{T}})|$ vs. $\not\!\!E_T$.
  - $|\Delta\phi|$ from 1.2 to 0 rad when $\not\!\!E_T = 0$ GeV, and $\not\!\!E_T$ from 0 to 85 GeV when $|\Delta\phi| = 0$ rad.
  - $|\Delta\phi|$ from 2.5 to $\pi$ rad when $\not\!\!E_T = 0$ GeV, and $\not\!\!E_T$ from 0 to 30 GeV when $|\Delta\phi| = \pi$ rad.

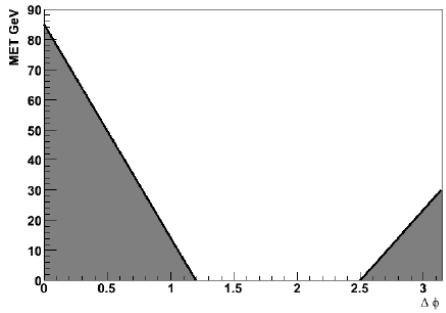


Figure 5.6: Figure showing the triangle cut between $\Delta\phi$ and $\not\!\!E_T$ for muons.

- Muon track curvature significance cuts of |TrackCurvSig| vs. $|\Delta\phi(\mu, \not\!\!E_T)|$, where |TrackCurvSig| = $|\frac{q/p_T}{\sigma(1/p_T)}|$, and $q$ and $p_T$ are the charge and transverse momentum of the charged track associated with the muon
  - $|\Delta\phi|$ from $0.875\pi$ to $\pi$ rad when |TrackCurvSig| = 0, and |TrackCurvSig| from 0 to 4 (6) when $|\Delta\phi| = \pi$ rad for Run IIa (Run IIb) period.
  - $|\Delta\phi|$ from 2 to $\pi$ rad when |TrackCurvSig| = 0, and |TrackCurvSig| from 0 to 2 (3) when $|\Delta\phi| = \pi$ rad for Run IIa (Run IIb) period.

  These cuts are needed to reject events with poorly measured muons that cause an excess in data over background model in the $\Delta\phi$ distributions.
- Transverse momentum of the leading jet within the ICD region of the detector: leading jet $p_T >$ 30 GeV when $1.0 <$ |leading jet $\eta^{\text{det}}| < 1.5$.
- Transverse momentum of the second leading jet within the ICD region of the detector: leading jet $p_T > 25$ GeV when $1.0 <$ |leading jet $\eta^{\text{det}}| < 1.5$ only for Run IIb period with exactly three good jets.

### 5.2.3 Multijets Modeling

The multijet (QCD) background is modeled by using data with much looser lepton selection than used for signal selection. The purpose of using looser selection is to increase the sample statistics by ten-fold and to ensure sufficient statistics after $b$-tagging to make a proper measurement of this background. An orthogonal dataset is created where all the selection cuts stated above are imposed, except the lepton requirements. The lepton requirements for multijet background for electron and muon channels are described below:

**Electron channel**

- Only one ultraloose electron that fails the seven-variable EM-likelihood cut (i.e. it satisfies $\mathcal{L} < 0.85$) as described in Section. 4.4 and also is not required to have a matching track.

- No second ultraloose electron is allowed in any $|\eta^{\text{det}}|$ region. (This requirement rejects all events with a tight isolated electron, thus keeping this sample orthogonal to the main one.)

By using the looser lepton selection, differences in shape for some variable distributions were introduced [124]. These kinematic biases can be removed by reweighting the events with a correction factor calculated from the matrix method equations [125]. The kinematic correction factor QCDWeight $= \varepsilon_{\text{fake}-e}/(1-\varepsilon_{\text{fake}-e})$ depends on $\varepsilon_{\text{fake}-e}$, which is the efficiency for a fake ultraloose electron to pass the tight electron criteria.

$\varepsilon_{\text{fake}-e}$ can be measured as the ratio of the number of events with a *tight* electron over the number of events with an *ultraloose* electron in the $\not\!\!E_T < 10$ GeV region. Even though the low- $\not\!\!E_T$ region is dominated by multijet events, there is a non-negligible contamination from real electrons, in particular for the tight sample. This *signal contamination* originates mostly from $Z$+jets events and to some extend also from $W$+jets events. The expected contributions to the ultraloose and tight samples originating from $W$+jets and $Z$+jets events are estimated from MC and subtracted from the number of events in data before computing $\varepsilon_{\text{fake}-e}$. As shown from Figure 5.7 $\varepsilon_{\text{fake}-e}$ has a relative kinematic dependance on lepton $p_T$. No significant difference has been observed in the fake rate for different jet multiplicities and $\varepsilon_{\text{fake}-e}$ is thus measured in events with at least two jets. Figure 5.7 shows the effect of correcting for the signal contamination in $\varepsilon_{\text{fake}-e}$ and in the kinematic bias correction factor.

**Muon channel**

- The $\Delta R(\mu, \text{jet}) > 0.5$ isolation requirement is dropped.
- Events with a tight isolated muon are rejected. This ensures orthogonality.
- Jets near the muon are removed and the $\not\!\!E_T$ of the event is corrected accordingly in order to make the jets reproduce those in the signal data.

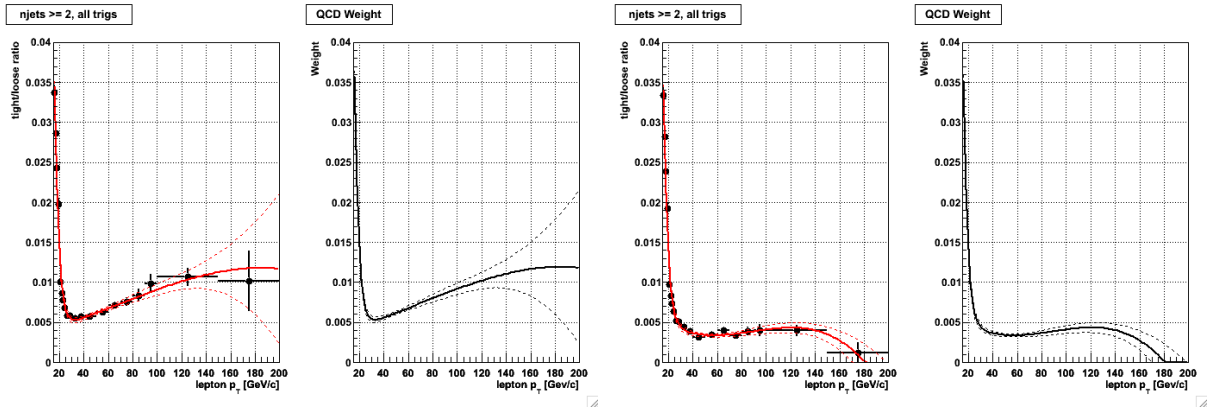


Figure 5.7: $\varepsilon_{\text{fake}-e}$ and QCDWeight $= \varepsilon_{\text{fake}-e}/(1-\varepsilon_{\text{fake}-e})$ as a function of electron $p_T$ before (left) and after (right) the signal contamination correction described in the text.

### 5.2.4 Monte Carlo Corrections

After the full MC simulations as described in Section 5.2 and event selection, some correction factors for MC reconstruction and selection efficiencies need to be applied in order to mimic the true data events completely. Since during the MC simulations and applying the detector response to MC events, the wear-and-tear of the detector, for example, debris build-up and aging effects, are not considered and hence reconstruction efficiencies for electrons, muons and jets tend to be overestimated in the simulations. To account for these effects, scale factors and smearing factors are measured from data and applied to the MC Samples. The corrections that are applied to the simulated samples in order to reach agreement with data in this analysis are described below.

- **Primary vertex position reweighting**

The distribution of the $z$ position of the primary interaction point tends to be wider in data than it is in the simulation. A weight (correction factor) is applied to MC so that the distribution of the primary vertex $z$ location matches data [126].

- **Instantaneous luminosity reweighting**

The instantaneous luminosity for a MC event is determined from the corresponding value for the overlaid zero bias data event. And as zero bias data is collected at different run periods and hence the instantaneous luminosity distribution of the MC is reweighted by applying a

scale factor to events so that these distributions agree.

• $Z$ $p_T$ **Reweighing**

As described in Section 5.2.1, $Z + jets$ sample is generated by ALPGEN, and the $Z$ $p_T$ modeled from this sample does not match the data. To account for this effect, a scale factor depending on the true $Z$ $p_T$ and jet multiplicity is used to make the sample agree with Next Leading Order (NLO) calculations [127].

• **Electron Identification Efficiencies**

The efficiency to identify an electron candidate depends on the factors like EM clusters properties, its isolation, a track match, and a likelihood. MC efficiencies exceed the data in all respects, and are therefore scaled accordingly. $Z \to ee$ events are used to derive the efficiencies through a standard tag-and-probe method. The ratios of the data and Monte Carlo efficiencies are used to scale the Monte Carlo and is applied to tight electrons. For the ultraloose electrons used in multijets, this correction factor is consistent with one [128, 129, 130, 131]. The correction factor is defined as:

$$\varepsilon_{e-ID} = \frac{\varepsilon^{Data}_{Presel}}{\varepsilon^{MC}_{Persel}} \times \frac{\varepsilon^{Data}_{PostPersel}}{\varepsilon^{MC}_{PostPersel}} \tag{5.1}$$

where $\varepsilon$ is the efficiency which can be derived for the data or MC sample and for the preselection or post-preselection scale factors.

• **Muon Efficiency Correction**

In the MC samples, the detector resolution of muon events is overestimated and to account for such overestimation and to match the resolution with data, the reconstructed Monte Carlo muons need to be smeared to ensure that the resolution matches that of real muons. The momentum smearing can be expressed as

$$\left(\frac{q}{p_T}\right)' \to \frac{q}{p_T} + \left(A + \frac{B}{p_T}\right) \times G, \tag{5.2}$$

where the parameter $G$ is a random number generated from a Gaussian distribution centered

at 0 and a width of 1. The parameters $A$ and $B$ are measured for muons with an SMT track hit in two regions ($\eta < 1.6$ and $\eta > 1.6$) and for muons without an SMT hit. The muon track is defined by the charge and radius of curvature, which is proportional to $q/p_T$ , thus the natural quantity to smear is $q/p_T$.

After the smearing is applied, the tag-and-probe method [132, 133] is used to measure the muon correction factor which is defined as the product of three independent factors for reconstruction, track matching, and isolation:

$$\varepsilon_{\mu-ID} = \frac{\varepsilon^{Data}_{Reco}}{\varepsilon^{MC}_{Reco}} \times \frac{\varepsilon^{Data}_{Track|Reco}}{\varepsilon^{MC}_{Track|Reco}} \times \frac{\varepsilon^{Data}_{Isolation|Track}}{\varepsilon^{MC}_{Isolation|Track}} \tag{5.3}$$

• **Jet Corrections**

MC jets have a better energy resolution, a higher reconstruction efficiency, and sometimes a higher average jet energy than what is observed in data. A procedure called JSSR (Jet Smearing Shifting and Removal) is applied at DØ [134] to correct for this effect. This algorithm is used for the $W + jets$, $Z + jets$, and diboson samples while leaving it off for the top pair and single top samples as recommended by the DØ JSSR group.

• **$V$+jets eta Reweighting**

The $V$+jets ($W$ and $Z$+jets) MC samples used in this analysis are generated with ALPGEN. The leading log ALPGEN provides a good modeling to these background samples in general, but some discrepancies, specifically in jet distributions at large $\eta$ are seen when compared with data. The reason of these discrepancies can be due to intrinsic limitations of the leading log approximation and/or potentially any bias introduced by zero bais overlay in MC or difficulties in simulation due to some poorly instrumented detector regions such as ICR. To deal with this, the $V$+jets samples are reweighted before $b$-tagging to reach agreement with the jet $\eta$ distributions observed in data. The reweighting is derived by comparing the $V$+jets sample to the data after subtraction of all other backgrounds. Reweighting functions are derived for the following variables sequentially: leading jet $\eta$, second leading jet $\eta$, $\Delta R$ between two leading jets, and thereafter the third and fourth jet $\eta$ when applicable [129].

These reweighting functions are derived such that the overall normalization stays the same. Only the kinematic shape of the samples are affected.

• $b$ **Jet Identification and Corrections**

Identification of $b$-Jets is a very important tool in the analysis which reduces the signal to background ratio even more after event selection cuts since the single top event signature requires at least one $b$-tagged jet. For considering jets to be $b$-tagged, first jets are required to be "*taggable*". A taggable jet is a calorimeter jet matched within $\Delta R < 0.5$ to a track jet. The calorimeter jet must therefore be within the region covered by the tracking system. This is imposed by requiring $\eta_{jet} < 2.5$. The track jet must consist of at least two tracks, with $\Delta R < 0.5$ between them. Each track in the track jet must have at least one SMT hit, and at least one of the tracks must have $p_T > 1$ GeV/c. The Neural Network (NN) algorithm [112] developed by DØ b-id group is used to identify jets originating from $b$-hadrons. As defined in Section 4.7, events are considered to have two tagged jets if exactly two jets have a $b$-tag $NNoutput > 0.5$. Events which do not pass the two tag criteria are considered to be one tagged events if there is exactly one jet with a $b$-tag $NNoutput > 0.775$.

For $W$+light jets MC sample, a different tagging technique is used. Since this sample contains, by definition, no heavy flavor jets and therefore the only events which pass the tagging criteria are those light jets which fake a tagged jet. As this fake rate is only a few percent, most of the $W + light\ jet$ sample is thrown away resulting in low statistics in the selected sample. In order to increase the statistics of the sample, the method of random tagging is devised [135, 129]. This algorithm assigns every jet a 50% probability to be tagged and then adjusts the weights of these events to match the expected yields after tagging.

The NN tagger is applied directly to select $b$ jets in data samples, whereas in MC events parametrizations are used to define probabilities for a jet to be tagged. The following subsections describe how these probabilities are derived and how they are used.

**Taggability Scale Factors**

The ratio of the number of jets passing the taggability requirement to the total number of jets is called taggability efficiency. The taggability efficiency depends on geometric features of the detector and the jet to be reconstructed and hence is parameterized in jet $p_T$, jet $\eta$ and the primary vertex $z$ coordinate ($PV_z$) for the event. Taggability requires jets to have well defined tracks. In general, the taggability increases as the jet traverses more layers of the tracking system and tends to zero when $\eta_{jet}$ is high such that jets can not reach the SMT. On the other hand, the taggability increases with the $p_T$ of the jet, because the particle content increases and therefore the number of tracks within it, as well as tracks with higher $p_T$ curve less in the magnetic field and are more likely to reach the tracking detector.

A taggability scale factor for each jet ($SF_T$) can be derived by comparing the taggability measured in data and MC. A per-jet scale factor, $T(\eta, p_T, PV_z)$ is defined to be equal to the ratio of taggability in data to that of MC. A scale factor, per event, can by constructed by taking the product of scale factors for taggable and untaggable jets as follows:

$$SF_T(p_T, \eta_{det}, PV_z) = \frac{T^{\text{data}}(p_T, \eta_{det}, PV_z)}{T^{\text{MC}}(p_T, \eta_{det}, PV_z)} \tag{5.4}$$

A similar scale factor ($SF_{\bar{T}}$) is used for each untaggable jet:

$$SF_{\bar{T}}(p_T, \eta_{det}, PV_z) = \frac{1 - T^{\text{data}}(p_T, \eta_{det}, PV_z)}{1 - T^{\text{MC}}(p_T, \eta_{det}, PV_z)}. \tag{5.5}$$

An event weight to correct efficiency differences between data and MC due to taggability can be obtained by taking the product of all of these scale factors assuming the taggability correction is weakly correlated

$$w_T = \prod_{i=1}^{N_T} SF_T(p_{T,i}, \eta_{det,i}, PV_{z,i}) \times \prod_{j=1}^{N_{\bar{T}}} SF_{\bar{T}}(p_{T,j}, \eta_{det,j}, PV_{z,j}), \tag{5.6}$$

where $N_T$ is the number of taggable jets and $N_{\bar{T}}$ is the number of untaggable jets per event. To account for some of the missing correlations between jets by measuring a tag-

gable/untaggable jet multiplicity scale factor defined as:

$$SF(N_T, N_{\bar{T}}) = \frac{E^{\text{data}}(N_T, N_{\bar{T}})}{E^{\text{MC}}(N_T, N_{\bar{T}})} \tag{5.7}$$

where $E^{\text{data}}(N_T, N_{\bar{T}})$ and $E^{\text{MC}}(N_T, N_{\bar{T}})$ is the number of events in data and MC after been corrected by Eq. 5.6, with the corresponding jet multiplicities $N_T$ and $N_{\bar{T}}$ for taggable and untaggable jets respectively. The total global event weight is then obtained by multiplying this last scale factor to Eq. 5.6 and is applied to every MC event [129].

$$w_T = SF(N_T, N_{\bar{T}}) \times \prod_{i=1}^{N_T} SF_T(p_{T,i}, \eta_{det,i}, PV_{z,i}) \times \prod_{j=1}^{N_{\bar{T}}} SF_{\bar{T}}(p_{T,j}, \eta_{det,j}, PV_{z,j}). \tag{5.8}$$

**Tagging Scale Factor**

Similar as taggability, the efficiency for a jet to be tagged differs in MC than data. To correct this effect efficiency for a jet to be tagged in both data and MC is measured [136]. This efficiency is parametrized in jet $\eta$ and jet $p_T$ for each working point of the NN tagger. The scale factor for any jet is then calculated as:

$$SF_{tag}(p_T, \eta, NN_{out}) = \frac{\epsilon^{data}(p_T, \eta, NN_{out})}{\epsilon^{MC}(p_T, \eta, NN_{out})} \tag{5.9}$$

where $\epsilon$ is the tagging efficiency in data or MC. An event scale factor is constructed by taking the product of these scale factors over all taggable jets.

$$\prod_{i=1}^{N_T} SF_{tag}(p_{T,i}, \eta_i, NN_{out,i}) \tag{5.10}$$

### 5.2.5 Monte Carlo Sample Normalization

The $t\bar{t}$, $Z$+jets, dibosons, and single top samples are normalized to the integrated luminosity 5.4 $fb^{-1}$ of the dataset using the cross sections and branching fractions listed in Table 5.2. The $W$+jets background is corrected in the same way as the other MC samples, but here

further corrections are needed. First, the sample is normalized to the ALPGEN leading log cross sections listed in Table 5.2, but these cross section calculations for $Wb\bar{b}$, $Wc\bar{c}$, and $Wjj$ ($Wjj$ includes $W$+$c$+light partons) have sizable uncertainties and are very sensitive to the renormalization and factorization scales. From comparisons of LL and NLO calculations, $K'$ and $K'_{HF}$ scale factors have been derived [137] with which to scale the $Wb\bar{b}$, $Wc\bar{c}$ and $Wjj$ samples; these are shown in Table 5.3. Since this analysis also includes the $Z$+jets backgrounds, the $K'$ and $K'_{HF}$ factors used to scale these are also shown.

| | NLO/LL $K'$ and $K'_{HF}$ Factors for $V$+Jets Cross Sections | | | | | | |
|---|---|---|---|---|---|---|---|
| | $Wjj$ | $Wcj$ | $Wc\bar{c}$ | $Wb\bar{b}$ | $Zjj$ | $Zc\bar{c}$ | $Zb\bar{b}$ |
| Run IIa | 1.246 | 1.648 | $1.246 \times 1.47$ | $1.246 \times 1.47$ | 1.253 | $1.253 \times 1.67$ | $1.253 \times 1.52$ |
| Run IIb | 1.249 | 1.648 | $1.249 \times 1.47$ | $1.249 \times 1.47$ | 1.239 | $1.239 \times 1.67$ | $1.239 \times 1.52$ |

Table 5.3: Theoretical $K'$-factors used to scale the boson+jets backgrounds. When a product of two numbers is shown, the first one, $K'$, is for the ratio of NLO to LL for all events and the second one is the heavy-flavor scale factor $K'_{HF}$. For $W$+jets the numbers are the same for the two data-taking periods because the version of ALPGEN used was the same for each background model.

An additional heavy-flavor scale factor correction is derived for $W + jets$ normalization from comparison with data and is applied to the $Wb\bar{b}$, $Wc\bar{c}$ MC samples. These heavy flavor scale factors can be measured by comparing the samples with zero, one, and two tags [129]. The measured heavy flavor scale factor is consistent with one.

### 5.2.6 Multijets and $W$+jets and Normalization

As described in Section 5.2.5, all background samples except $W$+jets and mutijets are normalized using theory cross sections. The sum of $W$+jets and multijets backgrounds is normalized to pre-taged data i.e, two normalization scale factors, $\lambda_{wjets}$ and $\lambda_{multijets}$, are derived such that the total number of predicted events match data before any $b$-tagging selection and is applied to the $W + jets$ and multijet samples respectively. These two scale factors

$\lambda_{\text{wjets}}$ and $\lambda_{\text{multijets}}$ are anti-correlated, as shown in the following equation:

$$N = \lambda_{\text{wjets}} N_{\text{wjets}} + \lambda_{\text{multijets}} N_{\text{multijets}}, \tag{5.11}$$

where $N = N_{\text{data}} - N_{\text{non-wjets-MC}}$ and $N_{\text{data}}$, $N_{\text{non-wjets-MC}}$, $N_{\text{wjets}}$ and $N_{\text{multijets}}$ are the event yield in data, non-$W + jets$ MC, $W + jets$ and multijet samples, respectively.

The $\lambda_{\text{wjets}}$ and $\lambda_{\text{multijets}}$ are determined by comparing the lepton $p_T$ , $\not\!\!E_T$ , and $M_T(W)$ distributions between data and background, which all have significantly different shapes in $W + jets$ and multijet backgrounds. The procedure according to which $\lambda_{\text{wjets}}$ and $\lambda_{\text{multijets}}$ are calculated is the following:

1. Select the set of kinematic variables to be used (in this case lepton $p_T$, $\not\!\!E_T$, and $M_T(W)$).

2. Start with $\lambda_{\text{wjets}} = 0$ and calculate the corresponding $\lambda_{\text{multijets}}$ using Eq. 5.11.

3. Do a Kolmogorov-Smirnov test (KS-test) between data and background distributions for the three kinematic variables and record their KS-test values.

4. Increase $\lambda_{\text{wjets}}$ by a given step size (by default 0.001).

5. Repeat the above three steps until $\lambda_{\text{wjets}}$ reaches 4.0 or when $\lambda_{\text{multijets}}$ becomes negative. The reason of choosing range of 0-4 is as it was found to be sufficiently large to cover the range of values.

The shapes of the three variables in the data are compared to the MC in each case and the values of $\lambda_{\text{wjets}}$ and $\lambda_{\text{multijets}}$ are chosen to be the values which give the best overall agreement. The fit is performed for each jet multiplicity, lepton type, and run period. The derived values are shown in Table 5.4.

| | IKS normalization scales | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | $\lambda_{\text{wjets}}$ | | | | $\lambda_{\text{multijets}}$ | | | |
| | Run IIa (p17) | | Run IIb (p20) | | Run IIa (p17) | | Run IIb (p20) | |
| | $e$ | $\mu$ | $e$ | $\mu$ | $e$ | $\mu$ | $e$ | $\mu$ |
| 2 jets | 1.071 | 1.042 | 0.971 | 0.954 | 0.444 | 0.014 | 0.388 | 0.034 |
| 3 jets | 1.326 | 1.288 | 1.082 | 1.105 | 0.351 | 0.025 | 0.343 | 0.045 |
| 4 jets | 1.371 | 1.553 | 1.070 | 1.294 | 0.309 | 0.033 | 0.383 | 0.024 |

Table 5.4: $W$+jets and multijets scales from the iterative KS normalization method.

## 5.3 Event Yields

The term "yield" is defined as the predicted number of events of the signal and background in the nearly 5.4 $fb^{-1}$ of data. As the $W + jets$ and multijets backgrounds are normalized to data before $b$-tagging, the sum of the backgrounds and signal is defined to equal the number of events observed in the data. The yields after the selection described in Section 5.2.2 and before $b$-tagging is shown in Table 5.5. After $b$-tagging is applied, the sample is divided into events containing exactly one and exactly two $b$-tagged jets. The event yield for One tagged and Two tagged are shown in Tables 5.6 and 5.7.

| | Pretagged event yields with channels combined | | | |
|---|---|---|---|---|
| | Electron+Muon, Run IIa+Run IIb | | | |
| | 2 jets | 3 jets | 4 jets | All Channels |
| Signal | | | | |
| $tb$ | 167 ± 0.69 | 71 ± 0.45 | 22 ± 0.25 | 259 ± 0.86 |
| $tqb$ | 319 ± 1.2 | 148 ± 0.85 | 50 ± 0.51 | 517 ± 1.6 |
| $tb$+$tqb$ | 487 ± 1.4 | 219 ± 0.96 | 71 ± 0.57 | 777 ± 1.8 |
| Backgrounds | | | | |
| $t\bar{t}\to ll$ | 414 ± 1.2 | 292 ± 0.96 | 95 ± 0.52 | 801 ± 1.6 |
| $t\bar{t}\to l$+jets | 384 ± 2.5 | 1,078 ± 3.9 | 1,242 ± 4.1 | 2,703 ± 6.2 |
| $Wb\bar{b}$ | 2,977 ± 15 | 951 ± 7.6 | 242 ± 3.4 | 4,171 ± 17 |
| $Wc\bar{c}$ | 6,709 ± 30 | 2,187 ± 16 | 550 ± 7.1 | 9,446 ± 35 |
| $Wcj$ | 6,684 ± 36 | 1,417 ± 16 | 251 ± 5.5 | 8,352 ± 40 |
| $Wjj$ | 79,561 ± 116 | 16,696 ± 49 | 3,190 ± 19 | 99,448 ± 127 |
| $Zb\bar{b}$ | 238 ± 2.8 | 80 ± 1.7 | 22 ± 0.89 | 340 ± 3.4 |
| $Zc\bar{c}$ | 506 ± 5.6 | 161 ± 3.2 | 44 ± 1.6 | 711 ± 6.7 |
| $Zjj$ | 6,018 ± 41 | 1,239 ± 16 | 246 ± 6.6 | 7,503 ± 44 |
| Dibosons | 2,631 ± 12 | 673 ± 5.9 | 145 ± 2.7 | 3,450 ± 14 |
| Multijets | 5,029 ± 8.8 | 1,440 ± 5.9 | 404 ± 2.4 | 6,874 ± 11 |
| Background Sum | 111,151 ± 134 | 26,213 ± 57 | 6,432 ± 23 | 143,796 ± 147 |
| Background + Signal | 111,638 ± 134 | 26,432 ± 57 | 6,503 ± 23 | 144,573 ± 147 |
| Data | 111,635 ± 334 | 26,432 ± 163 | 6,503 ± 81 | 144,570 ± 380 |
| $S:B$ | 1:228 | 1:120 | 1:90 | 1:185 |

Table 5.5: Pretagged event yields with statistical uncertainty for each jet multiplicity and for all analysis channels combined.

| Once tagged event yields with channels combined | | | | |
|---|---|---|---|---|
| | Electron+Muon, Run IIa+Run IIb | | | |
| | 2 jets | 3 jets | 4 jets | All Channels |
| Signal | | | | |
| $tb$ | 66 ± 0.46 | 27 ± 0.30 | 7.7 ± 0.17 | 100 ± 0.57 |
| $tqb$ | 133 ± 0.78 | 58 ± 0.55 | 18 ± 0.32 | 209 ± 1.0 |
| $tb$+$tqb$ | 198 ± 0.91 | 85 ± 0.63 | 26 ± 0.36 | 309 ± 1.2 |
| Backgrounds | | | | |
| $t\bar{t}\rightarrow ll$ | 168 ± 0.84 | 121 ± 0.75 | 37 ± 0.38 | 326 ± 1.2 |
| $t\bar{t}\rightarrow l$+jets | 144 ± 1.6 | 424 ± 2.7 | 462 ± 2.8 | 1,030 ± 4.2 |
| $Wb\bar{b}$ | 989 ± 9.4 | 315 ± 4.6 | 84 ± 2.2 | 1,388 ± 11 |
| $Wc\bar{c}$ | 554 ± 9.5 | 208 ± 4.9 | 56 ± 2.3 | 818 ± 11 |
| $Wcj$ | 612 ± 15 | 137 ± 6.9 | 24 ± 2.2 | 774 ± 16 |
| $Wjj$ | 1,066 ± 4.1 | 301 ± 2.2 | 74 ± 1.1 | 1,441 ± 4.8 |
| $Zb\bar{b}$ | 78 ± 1.6 | 28 ± 1.2 | 7.3 ± 0.53 | 113 ± 2.1 |
| $Zc\bar{c}$ | 39 ± 1.4 | 17 ± 1.2 | 4.5 ± 0.51 | 60 ± 1.9 |
| $Zjj$ | 82 ± 5.7 | 31 ± 3.6 | 5.3 ± 1.0 | 118 ± 6.8 |
| Dibosons | 158 ± 3.1 | 48 ± 1.7 | 12 ± 1.1 | 218 ± 3.7 |
| Multijets | 258 ± 2.2 | 117 ± 1.9 | 37 ± 0.83 | 412 ± 3.1 |
| Background Sum | 4,150 ± 22 | 1,745 ± 11 | 804 ± 5.2 | 6,699 ± 25 |
| Background + Signal | 4,348 ± 22 | 1,830 ± 11 | 830 ± 5.3 | 7,008 ± 25 |
| Data | 4,284 ± 65 | 1,772 ± 42 | 851 ± 29 | 6,907 ± 83 |
| $S:B$ | 1:21 | 1:21 | 1:31 | 1:22 |

Table 5.6: Once tagged event yields with statistical uncertainty for each jet multiplicity and for all analysis channels combined.

| Twice tagged event yields with channels combined | | | | |
|---|---|---|---|---|
| | Electron+Muon, Run IIa+Run IIb | | | |
| | 2 jets | 3 jets | 4 jets | All Channels |
| Signal | | | | |
| $tb$ | 38 ± 0.31 | 17 ± 0.21 | 5.2 ± 0.12 | 60 ± 0.39 |
| $tqb$ | 7.8 ± 0.21 | 14 ± 0.27 | 7.3 ± 0.19 | 29 ± 0.39 |
| $tb$+$tqb$ | 46 ± 0.37 | 31 ± 0.34 | 13 ± 0.22 | 90 ± 0.55 |
| Backgrounds | | | | |
| $t\bar{t}\to ll$ | 90 ± 0.56 | 83 ± 0.52 | 27 ± 0.28 | 199 ± 0.82 |
| $t\bar{t}\to l$+jets | 31 ± 0.80 | 203 ± 1.7 | 334 ± 2.1 | 568 ± 2.8 |
| $Wb\bar{b}$ | 252 ± 4.5 | 90 ± 2.6 | 26 ± 1.2 | 367 ± 5.3 |
| $Wc\bar{c}$ | 41 ± 2.3 | 25 ± 1.8 | 11 ± 1.2 | 77 ± 3.2 |
| $Wcj$ | 16 ± 0.65 | 7.3 ± 0.55 | 1.6 ± 0.16 | 24 ± 0.87 |
| $Wjj$ | 30 ± 0.35 | 17 ± 0.28 | 6.0 ± 0.15 | 53 ± 0.47 |
| $Zb\bar{b}$ | 16 ± 0.64 | 8.5 ± 0.49 | 2.6 ± 0.33 | 27 ± 0.87 |
| $Zc\bar{c}$ | 3.2 ± 0.39 | 1.4 ± 0.27 | 0.56 ± 0.15 | 5.1 ± 0.50 |
| $Zjj$ | 2.9 ± 1.3 | 1.8 ± 0.81 | 0.64 ± 0.28 | 5.4 ± 1.6 |
| Dibosons | 20 ± 0.81 | 6.6 ± 0.49 | 2.0 ± 0.27 | 29 ± 0.98 |
| Multijets | 19 ± 0.62 | 13 ± 0.64 | 6.3 ± 0.33 | 39 ± 0.95 |
| Background Sum | 520 ± 5.5 | 456 ± 3.9 | 418 ± 2.8 | 1,394 ± 7.3 |
| Background + Signal | 566 ± 5.5 | 487 ± 3.9 | 431 ± 2.8 | 1,484 ± 7.3 |
| Data | 597 ± 24 | 535 ± 23 | 432 ± 21 | 1,564 ± 40 |
| $S:B$ | 1:11 | 1:15 | 1:33 | 1:16 |

Table 5.7: Twice tagged event yields with statistical uncertainty for each jet multiplicity and for all analysis channels combined.

## 5.4 Data-Background Model Comparison

This section shows the plots where the data are compared with the total background and signal predictions in order to ensure that the background is well modeled. The agreement between the data and the signal-background samples is studied for a long list of variables, both for each channel individually, and for various combinations of channels. Figure 5.8 shows the color scheme used to label the data and the different signal and background components in these plots.

After applying all of the event selection criteria, correction factors and $b$-tagging, the entire dataset and MC sample is divided into six subsamples. This division is performed by splitting the sample by the number of jets and tagged jets. All events contain either exactly one or two tagged jets and either two, three, or four jets, creating six subsamples. Figures 5.9 and 5.10 show some of the important kinematic distributions of the variables studied in the analysis for all six channels combined.

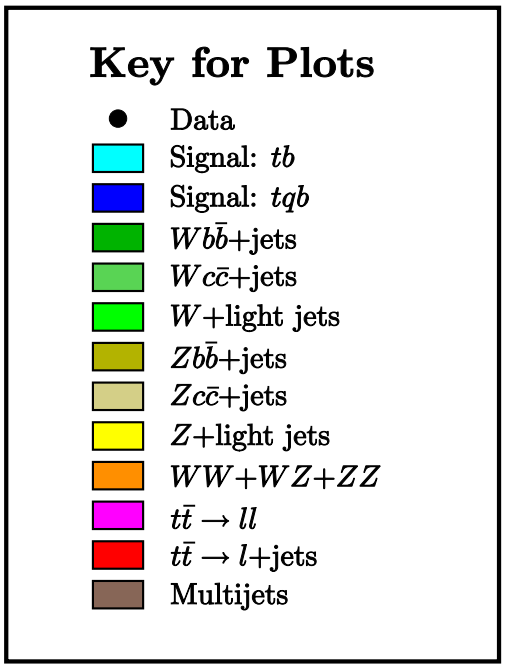


Figure 5.8: Illustration of the color scheme used in plots of signal and backgrounds in the analysis.

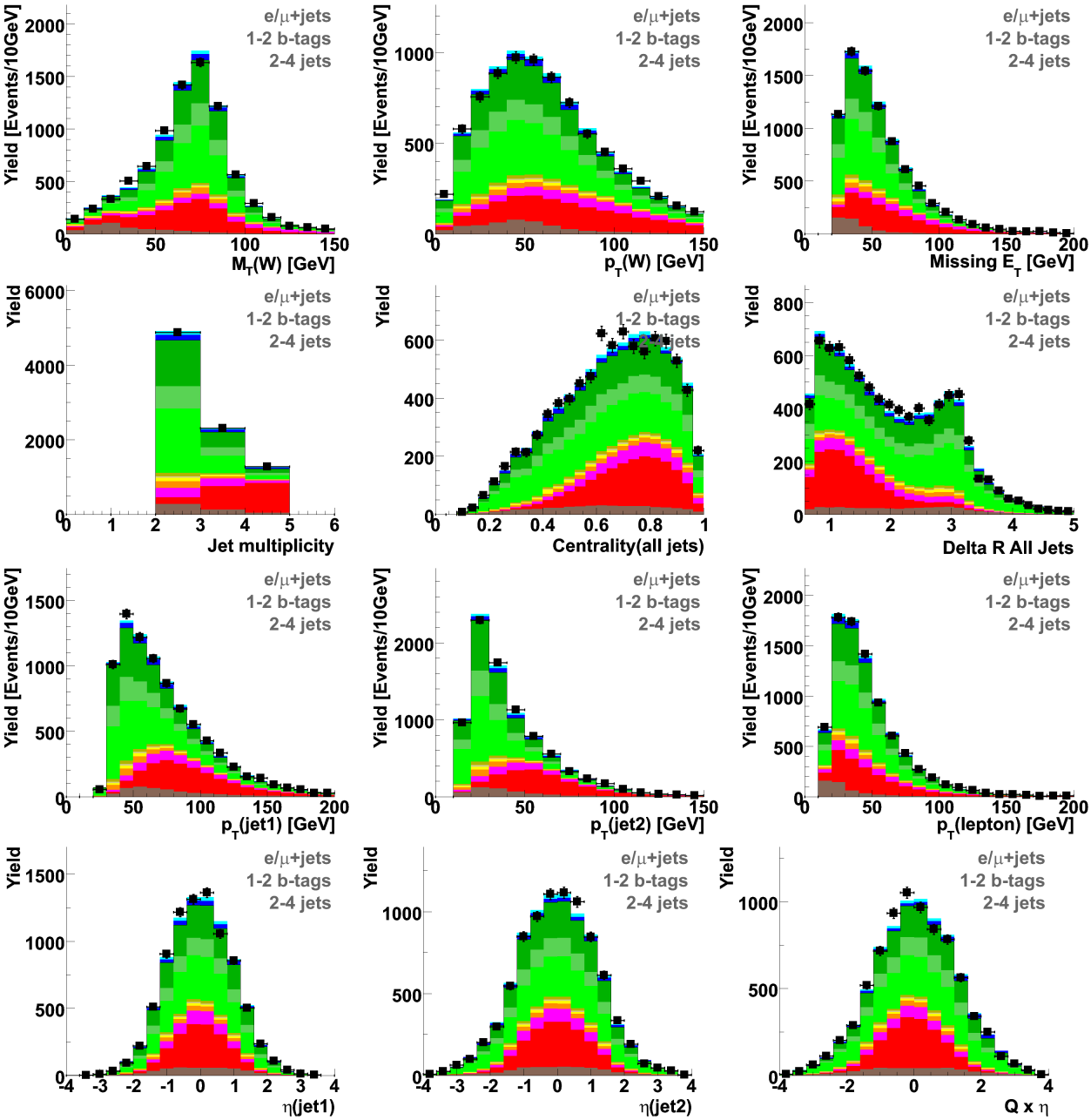


Figure 5.9: Data-background agreement of some variables after $b$ tagging has been applied (all 6 channels combined).

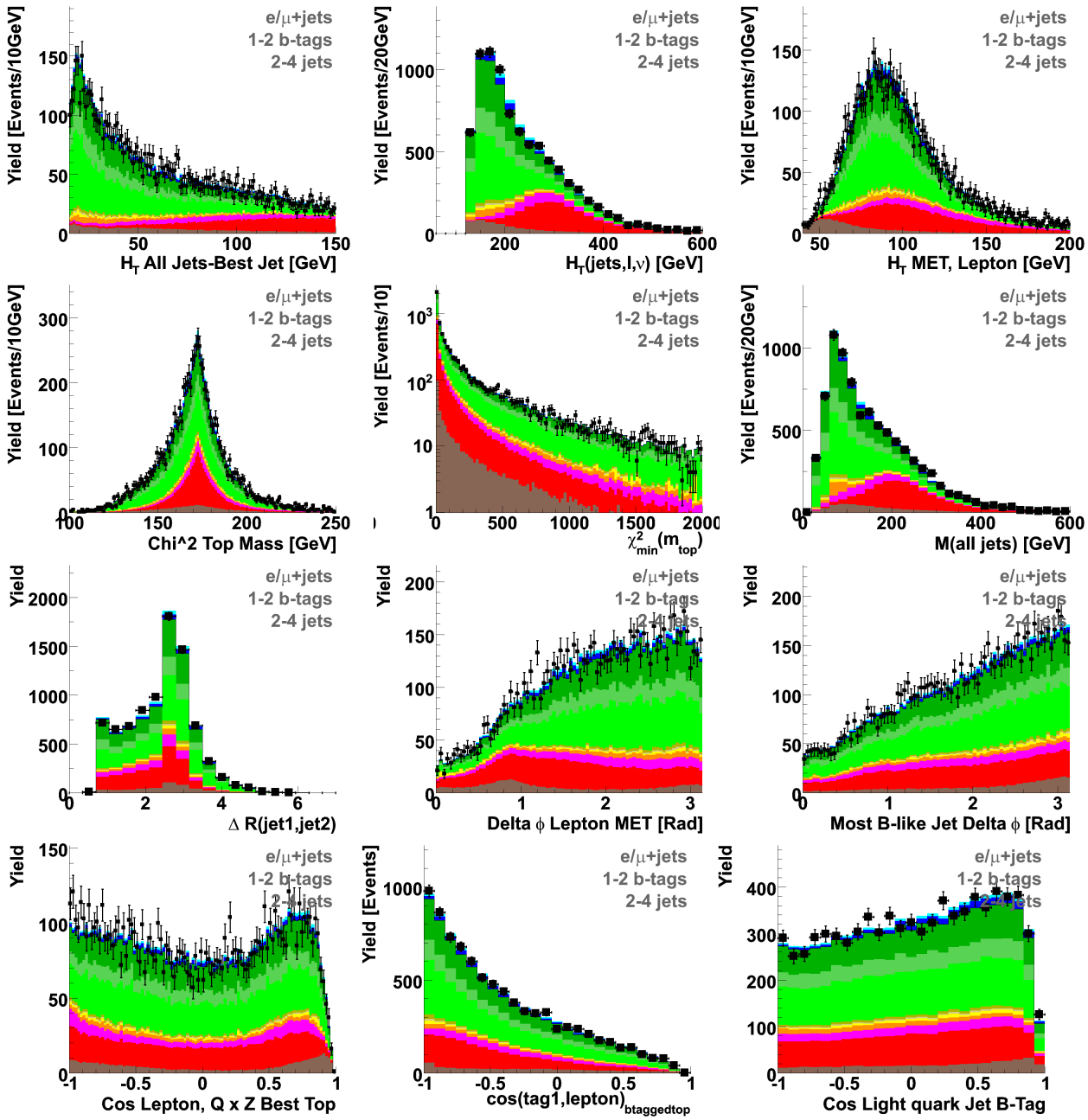


Figure 5.10: Data-background agreement of some more variables after $b$ tagging has been applied (all 6 channels combined).

## 5.5 Cross Check Samples

To test whether the background model reproduces the data in regions dominated by one type of background, two cross-check samples are selected for each of the electron and muon channels. The selection criteria are the same as for the main event samples as described in Section 5.2.2, but then the following additional requirements are imposed:

**"$W$+jets" sample**

- Exactly two jets
- $H_T$(lepton,$\not\!\!E_T$,alljets) $< 175$ GeV
- One $b$-tagged jet

**"$t\bar{t}$" sample**

- Exactly four jets
- $H_T$(lepton,$\not\!\!E_T$,alljets) $> 300$ GeV
- One or two $b$-tagged jets

Table 5.8 shows the yields of two cross-check samples with backgrounds and channels combined. For the $W$+jets samples, $W$+jets events form 82% of the samples, and the $t\bar{t}$ component forms less than 2% of the samples. For the $t\bar{t}$ samples, $t\bar{t}$ events form 84% of the samples, and the $W$+jets events form only 12% of the samples. The sum of the backgrounds agrees reasonably well with the observed data in each channel.

Figure 5.11 shows the $W$ transverse mass distribution for the cross-check samples, for all channels combined (electron and muon, Run IIa and Run IIb).

| | Yields for the Cross-Check Samples | | | |
|---|---|---|---|---|
| | Electron+Muon, Run IIa+Run IIb | | | |
| | $W$+Jets Samples | | $t\bar{t}$ Samples | |
| Signals | | | | |
| $tb$+$tqb$ | 60 ± 0.49 | | 13 ± 0.25 | |
| Backgrounds | | | | |
| $t\bar{t}$ | 29 ± 0.51 | (1.4%) | 573 ± 2.9 | (84%) |
| $W$+jets | 1,660 ± 15 | (82%) | 86 ± 2.0 | (13%) |
| $Z$+jets & dibosons | 168 ± 4.0 | (8.4%) | 11 ± 0.87 | (1.6%) |
| Multijets | 155 ± 1.7 | (7.7%) | 11 ± 0.46 | (1.6%) |
| Background Sum | 2,013 ± 16 | | 682 ± 3.7 | |
| Backgrounds+Signals | 2,073 ± 16 | | 694 ± 3.7 | |
| Data | 2,035 ± 45 | | 754 ± 27 | |
| $S : B$ | 1:34 | | 1:54 | |

Table 5.8: Yields after selection in the cross-check samples, for channels and backgrounds combined. The numbers in parentheses are the percentages of the total background+signal for each source.

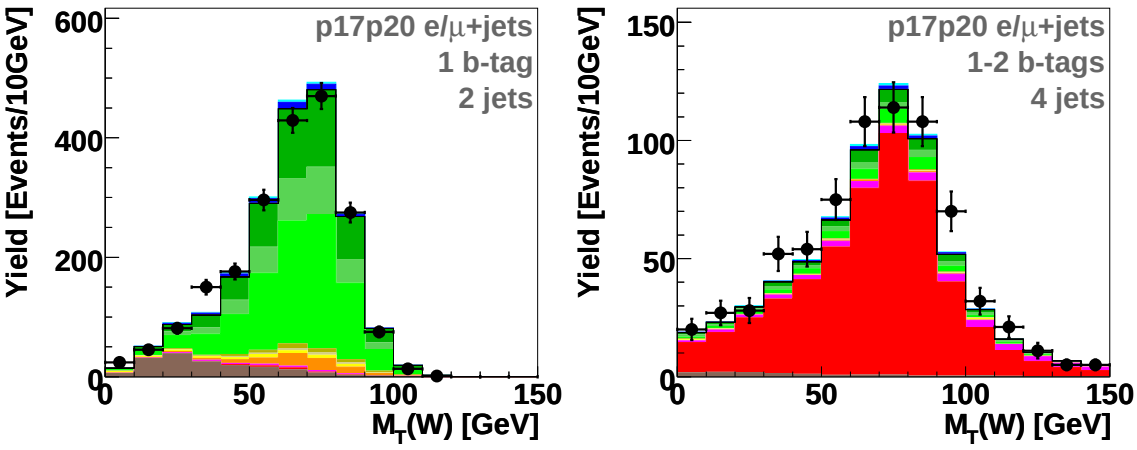


Figure 5.11: The $W$ transverse mass distribution for the "$W$+jets" and "$t\bar{t}$" cross-check samples, for all channels combined.

## 5.6 Systematic Uncertainties

There are two types of systematic uncertainties that can affect this analysis: 1) uncertainties on the normalization of the background samples, 2) uncertainties which effects the shape of the distributions of the background samples and of the expected signal. This section describes all the uncertainties considered in this analysis and the relative uncertainties on each of the sources is summarized in Table 5.9. For all the six analysis channels separately, detailed tables of uncertainties are listed in Appendix A.

UNCERTAINTIES AFFECTING NORMALIZATION

- **Integrated luminosity**

  The 6.1% uncertainty on the luminosity estimate affects the signal, $t\bar{t}$, $Z$+jets, and diboson yields [129].

- **Theory cross sections**

  The uncertainties on the single top and $t\bar{t}$ cross sections for top mass 172.5 GeV are $\pm 3.8\%$ for s-channel ($tb$), $\pm 5.3\%$ for t-channel ($tqb$), and $+6.4\%, -9.0\%$ for $t\bar{t}$. And for $tb$+$tqb$ combined, the uncertainty is $\pm 4.8\%$ [41, 138]. The uncertainties on the $Z$+jets cross section is set to 3.3% as per Ref. [139] and on the diboson cross sections is set to 7% [140].

- **Branching fractions**

  The branching fractions for a $W$ boson to decay to an electron, muon, or tau lepton, have an average uncertainty of 1.5% [9].

- **Parton distribution functions**

  The effect of changing the parton distribution functions on the signals is evaluated by reweighting single top MC events according to the 40 sets of CTEQerror PDFs and measuring the signal acceptance for each of them. A systematic uncertainty of 3% is assigned to the signal acceptances from this source. The PDF uncertainty on

the $t\bar{t}$, $Z$+jets, and diboson background yields is included in the theory cross section uncertainties listed above.

- **Trigger efficiency**

  This analysis uses an OR of many trigger conditions which gives a trigger efficiency close to 100% as described in Section 5.1.1. The uncertainty on the trigger efficiency is taken to be 5% in all channels. This measured efficiency is always equal to or smaller than this value in each channel [129].

- **Instantaneous luminosity reweighting**

  The instantaneous luminosity distributions of all MC samples are reweighted to make them match Run IIa or Run IIb data distributions as appropriate. The initial distributions are from the zero bais data overlaid on the MC events to simulate multiple interactions, and are generally at too low values for later data-taking conditions. The uncertainty on this reweighting is 1.0%.

- **Primary vertex modeling and selection**

  The distribution of the z position of the primary vertex is reweighted in MC to match that in data [126]. The uncertainty due to this reweighting is 0.05% (negligible). The uncertainty on the difference in primary vertex selection efficiency between data and MC is 1.4%.

- **Color Reconnection**

  The systematics due to color reconnections is estimated by taking the difference in signal efficiency between a sample generated with PYTHIA TuneAPro and PYTHIA TuneACPro [141]. The relative difference on preselection and b-tagging efficiency between both simulations is found to be 1.0% and is assigned as an additional uncertainty to the central efficiencies for $t\bar{t}$ background and signal MC.

- **Relative b/light jet response**

  The uncertainty on the difference in inclusive jet response for b-hadrons and normal jets is calculated in Ref. [142] and a difference of 1.8% is measured.

- **Electron reconstruction and identification efficiency**

  The uncertainty on the MC scale factor for electron reconstruction and identification includes the dependence of the electron identification scale factor on the variables ignored in the parametrization (jet multiplicity, track match, likelihood, electron $p_T$ and $\phi$ coordinate), and the limited statistics in each bin of the parametrization. The assigned total uncertainty is 4.8% [131, 143].

- **Muon reconstruction and identification efficiency**

  The MC scale factor uncertainties for muon reconstruction and identification, including isolation requirements, are estimated by the muon ID group [133] as coming from the tag and probe method, background subtraction, and limited statistics in the parametrization. The assigned total uncertainty is 2.1%.

- **Jet Fragmentation and higher-order effects**

  The acceptance of $t\bar{t}$ events generated with ALPGEN+PYTHIA (as used in the analysis) was compared with ALPGEN+HERWIG and MC@NLO+HERWIG. The difference in acceptance is used to estimate the systematic uncertainty from both jet fragmentation and higher order effects. The uncertainty is dependent on jet multiplicity and ranges from 0.7% to 7.8%. The jet fragmentation systematic effect is applied only to the $Z$+jets, single top, and diboson samples [129]. The uncertainty on higher order effects added is applied to the $t\bar{t}$ sample.

- **Initial-state and final-state radiation**

  This uncertainty is evaluated in $t\bar{t}$ samples generated with different amounts of initial and final-state radiation and comparing the acceptance to the nominal settings. The

overall measured uncertainty varies between 0.8% and 10.9% depending on the sample and jet multiplicity [129].

- **b-jet fragmentation**

  The size of the uncertainty from the $b$-jet modeling was evaluated in the $t\bar{t}$ pairs cross section analysis following the method described in Ref. [144]. The uncertainty arises from the difference between the fragmentation parametrizations preferred by SLD vs. LEP data. A 2.0% value is measured and applied to signal, $t\bar{t}$ and $Z$bb samples.

- **Taggability**

  Taggability efficiencies are calculated with one sigma up and down variations of taggability scale factor values. Contributions come from the statistical uncertainty on the measurement and on the $\eta$, $p_T$ , $PV_z$ parametrization. These are then used to measure the uncertainty on the selection efficiency of each sample ranging from 3.5-17% [129].

- **W+jets heavy-flavor scale factor correction**

  The heavy-flavor scale factor correction, $\lambda_{HF}$, for $Wb\bar{b}$ and $Wc\bar{c}$ is measured in a sample that is orthogonal to the one in the analysis and found to be consistent with 1 as described in Section 5.2.5. The uncertainty on the $\lambda_{HF}$ scale factor is 12% and the uncertainty on the $Wb\bar{b}$ and $Wc\bar{c}$ scale factor is 5% [129].

- **Z+jets heavy-flavor scale factor correction**

  The uncertainty on the heavy-flavor scale factor for $Zb\bar{b}$ and $Zc\bar{c}$ is taken from NLO calculations. The uncertainty on $\lambda_{HF}$ for $Z$+jets is considered to be the same as measured in the $W$+jets sample (12%).

- **W+jets and multijets normalization**

  The uncertainty from the $W$+jets and QCD multijets normalization are determined from a fit to the pre-tagged data, as described in Section 5.2.6, is estimated at 30-40% for the QCD multijets sample and 1.8% for the $W$+jets sample [129].

- **Sample statistics**

  The MC and data samples used to estimate the signal and background shapes are limited in size. In particular the multijet background samples have low statistics after $b$-tagging, as seen in Tables 5.6 and 5.7. The statistical uncertainty on the background is taken into account for each sample in each bin of the final discriminant distribution.

UNCERTAINTIES AFFECTING NORMALIZATION AND SHAPE

- **Jet reconstruction and identification**

  To estimate the uncertainty associated with the efficiency to reconstruct and identify jets, a sample is created by turning off the correction factor to the jet ID. And then this sample is symmeterized by taking half the difference, in the MVA output, for the nominal and the systematic. This is then added or subtracted to the nominal in order to create symmetric JetID- up and down samples. The normalization uncertainty of this systematic ranges from 0.04% to 3.7%.

- **Jet energy resolution**

  The JER systematics samples used to estimate the shape changing uncertainty are created by fluctuating JER correction by one sigma up and down. The normalization of the samples changes from between 0.2% and 11.6% for all signal and background samples.

- **Jet energy scale**

  The JES correction is raised and lowered by one standard deviation on each MC sample and the whole analysis repeated, which produces a shape-changing uncertainty, and an overall normalization uncertainty. The normalization part ranges from 0.3% to 14.6% for all MC samples.

- **Vertex confirmation**

  The uncertainty in the efficiency correction for MC vertex confirmation is estimated

by creating two samples, one with a one sigma higher correction, and one with a one sigma lower correction. In addition to the shape changing effects, the normalization changes between 0.1% and 9.6%.

- **$b$-tagging**

  The uncertainty associated with the $b$-tagging correction factor for MC is fluctuated one sigma up and down to create the $b$-tagging shape changing systematic samples. A normalization difference of 4.3-15% is measured in the single tag sample and from 5.8-11.2% in the two tag sample.

- **$V$+jets angular corrections**

  The angular corrections are defined such that they preserve normalization in the pretag sample. The reweighting procedure does not change the normalization of the $W$ and $Z$+jets samples before $b$-tagging. Although after $b$-tagging, small normalization variations ($\approx$ 0.3%) are introduced by taggability and $b$-tagging scale factors. To estimate the shape differences, two additional samples with the angular corrections fluctuated up and down by one sigma are created.

| Relative Systematic Uncertainties | |
|---|---|
| **Components for Normalization** | |
| Integrated luminosity | 6.1% |
| $t\bar{t}$ cross section | 9.0% |
| $Z$+jets cross section | 3.3% |
| Diboson cross sections | 7.0% |
| Branching fractions | 1.5% |
| Parton distribution functions (signal acceptances only) | 2.0% |
| Triggers | 5.0% |
| Instantaneous luminosity reweighting | 1.0% |
| Primary vertex selection | 1.4% |
| Color reconnection | 1.0% |
| b/light jet response | (0.3-1.0)% |
| Electron identification | (2.8-3.8)% |
| Muon identification | 2.1% |
| Jet fragmentation and higher order effects | (0.7-7.0)% |
| Initial-and final-state radiation | (0.8-10.9)% |
| $b$-jet fragmentation | 2.0% |
| Taggability | (3.1-21.5)% |
| $W$+jets heavy-flavor correction | 12.0% |
| $Z$+jets heavy-flavor correction | 12.0% |
| $W$+jets normalization to data | 1.8% |
| Multijets normalization to data | (30-40)% |
| MC and multijets statistics | (0.2-16)% |
| **Components for Normalization and Shape** | |
| Jet reconstruction and identification | (0.04-3.7)% |
| Jet energy resolution | (0.2-11.6)% |
| Jet energy scale | (0.3-14.6)% |
| Vertex confirmation | (0.1-9.6)% |
| $b$ tagging, single-tagged | (4.3-14.0)% |
| $b$ tagging, double-tagged | (5.8-11.2)% |
| Angular correction | 0.3% |

Table 5.9: A summary of the relative systematic uncertainties for each of the correction factors or normalizations. The uncertainty shown is the error on the correction or the efficiency, before it has been applied to the MC or data samples.

# Chapter 6

# Bayesian Neural Networks

***The important thing in science is not so much to obtain new facts as to discover new ways of thinking about them.*** - Sir William Bragg.

The single top analysis is very challenging in the sense that not only the signal production cross section is very small but also the background is huge and is very difficult to separate from the signal. After applying the event selection and corrections described in Chapter 5, the signal to background ratio is very small. As seen from Tables 5.6 and 5.7, signal to background ratio (S:B) is 1:22 and 1:16 for once tagged and twice tagged events respectively. Hence after this low S:B ratio, a simple cut-based analysis is still not sensitive enough to measure the cross section of a top quark generated via the electroweak channel. And expected single top signal is still smaller than the uncertainty on the background. Traditionally, this kind of problem is dealt by applying further tight selection criteria (cuts) on discriminating variables and select a subset of the original sample with an enhanced signal to background ratio. But in order to maximize our signal acceptance, imposing any further cuts is avoided and instead a multivariate approach is chosen that combines the discriminating power of several variables, while still retaining 100% of signal sample after event selection. This approach manifests in the form of using Bayesian Neural Networks (BNN) in this thesis. To give the main idea, the technique requires the selection of several discriminating variables,

a sample of monte carlo signal events, and a sample of background events. The network studies the input samples, and produces a function that guides background events towards zero and signal events towards one. This Chapter basically deals with the theory involved and mathematical structure of BNN. The application of BNN to measure the single top production cross section and to compute limits on non-SM couplings in single top is discussed in next two Chapters 7 and 8, respectively.

## 6.1 Basic Concepts of Neural Networks

A neural network (NN) [145] is a powerful data modeling tool that is able to capture and represent complex input/output relationships. NN uses the inductive strategy to estimate models from data and this estimation process is referred to as "learning" or "learning from data". By learning one means obtaining representation of data or finding patterns in the data which can be used for several purposes such as predicting future or classification.

The most commonly used NNs are Multilayer Perceptrons (MLP), also known as feed-forward neural networks [146, 147]. A schematic of feed-forward NN is shown in Figure 6.1. An MLP consists of an interconnected group of neurons or nodes arranged in layers; each node processes information received by it with an activation (or transformation) function, and passes on the result to the next layer of nodes. The first layer, called the *input layer*, receives the feature variables, followed by one or more *hidden layers* of nodes and the last *output layer* gives the final response of the network. Each of the interconnections is characterized by a weight, and each of the processing nodes may have a bias or a threshold. The weights and thresholds are the network parameters, often collectively referred to as weights, whose values are learned during the training phase. The activation function is generally a non-linear function that allows for flexible modeling. The output of the network with one hidden

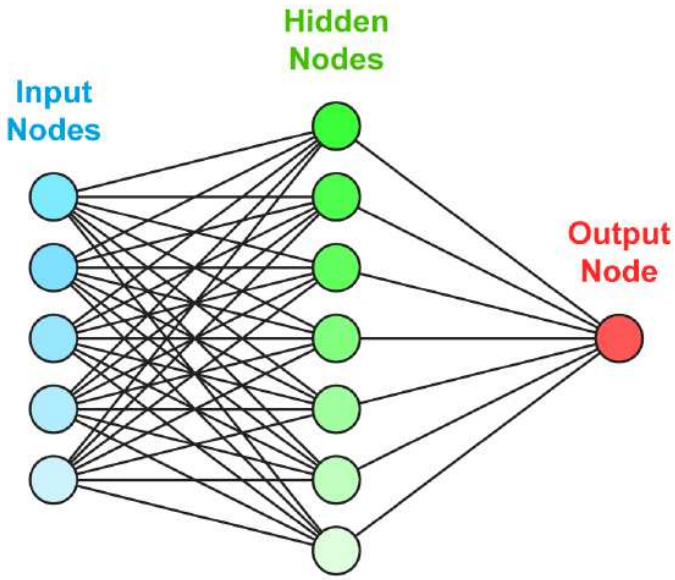


Figure 6.1: Illustration of the structure of a neural network with three layers: one layer with five input nodes, one layer with seven hidden nodes, and one layer with one output node.

layer of nodes and a data set with $d$ input variables $x = x_1, x_2, ...x_d$, is

$$O(x) = g(\sum_j w_j h_j + \theta), \tag{6.1}$$

where $h_j$ is the output form the hidden nodes,

$$h_j = g(\sum_j w_{ij} x_i + \theta). \tag{6.2}$$

Here, $i$ runs over the inputs $x$ and $j$ runs over the hidden units. $w_{ij}$ is the weight on the connection from hidden unit $j$ to input unit. The non-linear activation function $g$ is commonly taken as a sigmoid

$$g(a) = \frac{1}{1 + e^{-a}}. \tag{6.3}$$

If $g(a) \sim a$, the outputs $h_j$ at the hidden layer would be linear combinations of the inputs and the network with a single layer of adaptive weights would be a linear model. The

sigmoid function is linear close to $a \sim 0$ , nonlinear for higher values of $a$, and saturates for large values. Therefore, a network with a sigmoidal activation function contains a linear model as a special case. The network parameters are determined by minimizing an empirical risk function, usually the mean square error between the actual output $O_p$ and the desired (target) output $t_p$,

$$E = \frac{1}{N_p} \sum_{p=1}^{N} (O_p - t_p)^2 \tag{6.4}$$

where "$p$" denotes a feature vector or pattern.

There are some limitations of conventional NN. The critical issue in developing a neural network is generalization: how well the network make predictions for cases that are not in the training set ? NNs can suffer from either underfitting or overfitting. A network that is not sufficiently complex can fail to detect fully the signal in a complicated data set, leading to underfitting. A network that is too complex may fit the noise, not just the signal, leading to overfitting. Overfitting is especially dangerous because it can easily lead to predictions that are far beyond the range of the training data with many of the common types of NNs. Overfitting can also produce wild predictions in multilayer perceptrons even with noise-free data. We shall see in the next sections that by adopting Bayesian learning approach we can minimize the problems of overfitting.

## 6.2 Overview - Bayes' Theorem

A Bayesian network is a graphical model that encodes probabilistic relationships among variables of interest. When used in conjunction with statistical techniques, the graphical model has several advantages for data analysis. One, because the model encodes dependencies among all variables, it readily handles situations where some data entries are missing. Two, a Bayesian network can be used to learn causal relationships, and hence can be used to gain understanding about a problem domain and to predict the consequences of interven-

tion. Three, because the model has both a causal and probabilistic semantics, it is an ideal representation for combining prior knowledge (which often comes in causal form) and data. Four, Bayesian statistical methods in conjunction with Bayesian networks offer an efficient and principled approach for avoiding the over-fitting of data.

Therefore, the Bayesian approach to statistical analysis is that of inductive inference. It uses prior knowledge or new information or data to update probabilities. Hence it is a natural paradigm for learning from data. It is an intuitive and rigorous framework for handling classification and parameter estimation problems. At the heart of Bayesian inference [148] is Bayes theorem, written in general form as,

$$p(B|A) = \frac{p(A|B)p(B)}{p(A)} \tag{6.5}$$

where the conditional probabilities $p(B|A)$ and $p(A|B)$ (i.e, probability of B, given A and probability of A, given B; respectively) are referred to as the posterior probability and likelihood, respectively, $p(B)$ is the prior probability of B (i.e, probability of B irrespective of A), and the denominator is simply the total probability of A, $p(A) = \int p(A|B)p(B)dB$. If B is discrete, the integral is replaced by sum.

Let us consider a binary classification problem in which an event must be classified either as due to a signal process $s$, or as due to a background process $b$ . This is achieved by placing a cut on the ratio of the probabilities for the two classes,

$$r(x) = \frac{p(s|x)}{p(b|x)} = \frac{p(x|s)p(s)}{p(x|b)p(b)}, \tag{6.6}$$

where $p(x|s)$ and $p(x|b)$ are class conditional densities or likelihoods of the data for signal and background classes, respectively; $p(s)$ and $p(b)$ are prior probabilities. The discriminant 'r' is called the Bayes discriminant, where $r(x)$ is constant corresponds to a decision (separation) boundary in the feature space. The Bayes rule is to assign an input vector, that is, a data point, to signal class if $p(s|x) > p(b|x)$. This rule minimizes the probability

of misclassification. And the classifier which minimizes the misclassification rate is said to have reached the *Bayes limit*. The problem of discrimination, then mathematically reduces to that of calculating the Bayes discriminant $r(x)$ or any one-to-one function of it.

The posterior probability for the desired class $s$, becomes,

$$p(s|x) = \frac{p(x|s)p(s)}{p(x|s)p(s) + p(x|b)p(b)} = \frac{r}{1+r} \tag{6.7}$$

There are parametric and non-parametric methods to estimate $p(x|s)$ and $p(x|b)$. And Neural Networks being one of the method to approximate the posterior probability. Since $p(s)$ and $p(b)$ are not always known, then one can calculate the discriminant function as,

$$D(x) = \frac{s(x)}{s(x) + b(x)} \tag{6.8}$$

where $s(x) = p(x|s)$ and $b(x) = p(x|b)$. Hence the Bayes posterior probability for a signal class is related to $D(x)$ by,

$$p(s|x) = \frac{D(x)}{[D(x) + (1 - D(x))/k]}, \tag{6.9}$$

where $k = p(s)/p(b)$. The discriminant D(x) is often referred to as the likelihood discriminant.

Hence posterior probability plays a central role in pattern recognition, and Bayes' theorem allows it to re-expressed in terms of quantities which may be more easy to calculate. In the next section, we shall see that the problem of learning in neural networks from a Bayesian perspective can be treated simply by application of the rules of probability described above.

## 6.3 BNN Mathematics

The training of NNs can be viewed as a problem of inference, which can be addressed from a Bayesian viewpoint as discussed above. This perspective leads to a method called Bayesian

neural networks (BNN).

If a network is trained with events, described by a vector of variables $x$, such that signal events are labeled by $t = 1$ and background events by $t = 0$, then the network output $y$ approximates the posterior probability

$$y \approx Prob(t = 1|x) = \frac{p(x|1)p(1)}{p(x|1)p(1) + p(x|0)p(0)}, \tag{6.10}$$

that is, the probability that an event defined by the variables $x$ belongs to the signal class $t = 1$. $p(x|1)$ and $p(x|0)$ are the probability density functions for the signal and background classes respectively, and $p(1)$ and $p(0)$ are the corresponding class prior probabilities. Typically, one trains with equal numbers of signal and background events, in which case the priors cancel out. The label $t$ is referred to as the *target*.

The idea behind Bayesian neural networks (BNN) is to cast the task of training a network as a problem of inference, which is solved using Bayes' theorem. The latter is used to assign a probability density to each point $w$ in the parameter space of the neural network. Each point w corresponds to a network defined by a specific set of parameter values. In the conventional methods of training NNs, one attempts to find a single "best" network, i.e, a single "best" set of network parameters (weights). Bayesian training provides a posterior density for the network weights, $p(w|training\ data)$. The idea behind BNN is to assign a probability density to each point $w$ in the parameter space of the neural network. Then, one performs a weighted average over all points, that is, over all possible networks. Given the training data $(t_1, x_1), ..., (t_N, x_N)$, where $t_i$ is the known label associated with data $x_i$, the probability density assigned to point $w$, that is, to a network, is given by Bayes' theorem

(Section 6.2)

$$\begin{aligned} p(w|t,x) &= \frac{p(t,x|w)p(w)}{p(t,x)}, \\ &= \frac{p(t|x,w)p(x|w)p(w)}{p(t|x)p(x)}, \\ &= \frac{p(t|x,w)p(w)}{p(t|x)} \end{aligned} \tag{6.11}$$

where we have assumed that data $x$ do not depend on $w$, in which case $p(x|w) = p(x)$. Thus, in order to assign a probability density to a network, defined by the point $w$, we need the likelihood $p(t|x,w)$ and the prior density $p(w)$.

Consider a class of neural networks defined by the functional form

$$y(x,w) = \frac{1}{1+\exp[-f(x,w)]}, \tag{6.12}$$

where

$$f(x,w) = b + \sum_{j=1}^{H} v_j \tanh(a_j + \sum_{i=1}^{P} u_{ij}x_i), \tag{6.13}$$

having $P$ inputs, a *single* hidden layer of $H$ nodes and a single output. The parameters $u_{ij}$ and $v_j$ are called weights and $a_j$ and $b$ are called biases. Both sets of parameters are usually referred to collectively as weights, $w$.

Since, for a correctly trained network, the probability that $t = 1$ is $y(x,w)$, and (1-$y$) for $t = 0$, the probability of the set of targets $t = (t_1, t_2, ..., t_N)$, given the data $x = (x_1, x_2, ..., x_N)$, is

$$p(t|x,w) = \prod_{i=1}^{N} y^{t_i}(1-y)^{1-t_i}, \tag{6.14}$$

in which the events are assumed to be independent. Given an event with data $x^{'}$, a reasonable estimate of the probability that it belongs to the signal class (assuming $p(0) = p(1)$) is given

by the weighted average

$$\bar{y}(x^{'}) = \bar{y}(x^{'}|t, x) = \int y(x^{'}, w)p(w|t, x), \tag{6.15}$$

where the posterior density $p(w|t, x)$ is given by Eq. (6.11) is computed using the likelihood, Eq. (6.14) and prior $p(w)$, which is to be discussed in next Section 6.5. The only practical and feasible way to perform the high-dimensional integral in Eq. 6.15 is to sample the density $p(w|t, x)$, in some appropriate way, and to approximate the integral using the average

$$\bar{y}(x^{'}) = \bar{y}(x^{'}|t, x) \approx \frac{1}{K}\sum_{k=1}^{K} y(x^{'}, w_k), \tag{6.16}$$

where $K$ is the number of points $w$ sampled from $p(w|t, x)$. Again, since each point $w$ corresponds to a different neural network function in the class of networks with $P$ inputs and $H$ hidden nodes. The average is therefore an average over *networks*.

As discussed in the Section 6.2, there are several advantages offered by BNN over conventional feed-forward NN [149, 150]. Each point $w$ correspond to networks that are tightly fit to the training data. Such networks will typically perform poorly on an independent set of events. But in case of BNN, it is an average over many networks, one expects to produce an estimate of the signal class probability, $y = p(1|x)$, that is less likely to be affected by "over training." Moreover, in the Bayesian approach, there is less need to limit, severely, the number of hidden nodes because a low probability density will be assigned to points $w$ that correspond to unnecessarily large networks, in effect, pruning them away. The network can be as large as is computationally feasible so that the class of functions defined by the network parameter space includes a subset with good approximation to the true mapping.

## 6.4 Computing the Posterior Probability Density

To compute the average in Eq. 6.16, it is necessary to generate a sample of points $w$ from the posterior density, Eq. 6.11. Unfortunately, sampling from the posterior density is not feasible using simple numerical methods. In practice, a sample is generated using Markov Chain Monte Carlo (MCMC) methods [149, 151, 152]. In the MCMC method, one steps through a parameter space in such a way that points are visited with a probability which is proportional to the density $p(w|t, x)$. Points where $p(w|t, x)$ is large will be visited more often than points where $p(w|t, x)$ is small. The methods of choice for sampling complex densities, such as $p(w|t, x)$, originate in the field of computational statistical physics. The problem of moving through the network parameter space is re-cast as a problem of statistical mechanics, specifically, of a single particle moving through a (rather complicated) potential.

The posterior density is written as

$$p(w|t, x) = \exp[-V(q)], \tag{6.17}$$

where $V(q) = -\ln p(w|t, x)$ (with $q \equiv w$) is interpreted as a spatially varying "potential" through which the "particle" moves. One adds a "kinetic energy" term $T(p) = 1/2\ p^2$ where $p$ is a vector of dimensionality equal to that of the network parameter space. The "mass" of the "particle" can be taken to be unity by appropriate re-scaling. The motion of the particle is governed by its "Hamiltonian", $H = T + V$. For a Hamiltonian system, the particle will, eventually, visit every phase space point $(q, p)$ arbitrarily closely in such a way that the density of points in phase space is proportional to $\exp(-H)$. By randomly (and appropriately) injecting or removing "energy" from the system, different constant energy regions of phase space $(p, q)$ can be explored. A Markov chain $q_1, q_2, ..., q_N$ is thereby created, which converges (eventually) to a sequence of points that constitute a sample from the density $p(w|t, x)$. Since the correlation between adjacent points is very high, typically 0.9 or higher, one usually saves a point, that is, a network, after every $L$ steps, to lessen the correlation

between the saved points. Usually, the initial part of the Markov chain is discarded because, in general, it will not be a faithful sample of the required density.

## 6.5 The BNN Prior

In order to complete the Bayesian calculation, it is necessary to specify a prior density over the network space. In the Flexible Bayesian Modeling (FBM) package of Radford Neal [149], used in the analysis, a Gaussian prior centered at zero is specified for each weight. Smaller weights yields smoother sigmoid functions with a consequent better approximation to the discriminant. However, the variance for the weights belonging to a given group (either *input-to-hidden weights* ($u_{ij}$), *hidden-biases* ($a_j$), *hidden-to-output weights* ($v_j$) or *output-bias* ($b$)) is chosen to be the same: $\sigma_u^2$, $\sigma_a^2$, $\sigma_v^2$, or $\sigma_b^2$, respectively. However, since we do not know, a *priori*, what these variances should be, their values are allowed to vary over a large range, while favoring small variances. This is done by assigning each variance a gamma prior

$$p(z) = (\frac{\alpha}{\mu})^{\alpha} \frac{z^{\alpha-1} e^{-z\frac{\alpha}{\mu}}}{\Gamma(\alpha)}, \tag{6.18}$$

where $z = \sigma^{-2}$, and with the mean $\mu$, and shape parameter $\alpha$ set to some fixed plausible values. The inverse of variance $z = \sigma^{-2}$ is sometimes referred to as the *precision*. The gamma prior is referred to as a *hyperprior* and the parameter (here the precision) for which it is a prior is called a *hyperparameter*.

Measurement of single top production cross section and analysis done for present work using BNN is given in the next chapter.

# Chapter 7

# Analysis 1

# Cross Section Measurement

> ***Measure what can be measured, and make measurable what cannot be measured.*** - Galileo Galilei.

This chapter describes the application of Bayesian neural networks (BNN) to the single top quark search. We will see how BNN helps to separate the single top quark signal events from background events, and how the single top quark production cross section is measured using the BNN output distributions. Measurements of the signal significance and the CKM matrix element $|V_{tb}|$ are also presented, as well as the cross checks done for the measurements.

## 7.1 BNN Analysis

In the BNN analysis, a one-to-one function of the probability say, $\Pr(S|x)$ is approximated such that an event characterized by the variable $x$ belongs to the signal class $S$. This probability can be written using Bayes Theorem as described in Section 6.2, in terms of the probability density functions for the signal and backgrounds, $f(x|S)$ and $f(x|B)$, and the

corresponding class prior probabilities, $prior(S)$ and $prior(B)$:

$$\Pr(S|x) = \frac{f(x|S)\,\mathrm{prior}(S)}{f(x|S)\,\mathrm{prior}(S) + f(x|B)\,\mathrm{prior}(B)}. \tag{7.1}$$

Either by applying a cut on $\Pr(S|x)$, or by weighting events by it [153], can minimize the probability to misclassify events. An approximation to the discriminant

$$D(x) = \frac{f(x|S)}{f(x|S) + f(x|B)}, \tag{7.2}$$

is built using equal numbers of signal and background events.

**Road-Map of the Analysis**

Following are the steps involved in the BNN analysis for the single top production cross section measurement :

- The first step in the analysis is to prepare the samples with appropriate splitting which are to be used for training, testing and measurement purposes.

- The next important step in the analysis is the selection of input variables. Input variables are selected for all six analysis channels, these selected input variables should have very good discrimination between signal and background in order to build a strong discriminant.

- Having selected the input variables for different analysis channels, training of the networks is performed to build a discriminant $D(x)$. Performance of the BNN is also checked.

- The next step is to compute the posterior density of the single top cross section using binned likelihoods formed from the BNN output distributions by combining different tag and jet multiplicity channels, and optimize the bin and channel combination using a SM signal+background ensemble. Cross checks are done in order to study the

possible bias of the method. And finally the expected and observed posterior densities are computed for each of the six analysis channels as well as the optimum channel combination and cross section is determined.

- Last step in the measurement is to compute a $p$-value that quantifies the significance of the result using a background-only ensemble.

The detailed procedure of all these steps is discussed in the next sections of this chapter.

### 7.1.1 Input Samples

The single top signal samples consists of $s$ and $t$ channel MC events given by SM prediction. The background consists of $W$+jets, $t\bar{t}$, $Z$+jets, dibosons MC events, plus multijet events derived from data. Each of these signal and background samples, created as described in Section 5.2 are divided into three independent subsets which are further used in the analysis. The division of samples is done to reduce the bias which can be introduced if the same events are used for both the training and measurement. So, in order to keep the MC events used in the cross section measurement independent from the training sample, the original samples (signal and background) are divided into three sub-samples as follows :

- the first subset called "training", contains the one fourth of the samples and is used to train the BNN.

- the second subset called "testing" contains one fourth of the samples and is used to train the combination BNN as will be discussed in Section 7.2.6.

- the last subset known as "yield" contains the remaining half of the samples and is used to verify the BNNs convergence and to do the final measurement and all the plots.

The original samples are split by putting the first event in the testing, the second event in the training and the 3rd and 4th event in the yield sample, and repeating the process

until the end of the sample was reached. All three resulting sub-samples are normalized back to the full dataset: i.e, testing/training weights are multiplied by 4.0 and yield weights are multiplied by 2.0, where the actual factors are calculated as yields in the original over yield in the sub-sample.

### 7.1.2 Variables Selection

Selection of variables, which offers good discrimination between signal and background, is one of the most important steps in the BNN analysis. Usually, for most of the multivariate methods, the computational burden increases rapidly with the dimensionality of the input data. Moreover, keeping the dimensionality low relative to the size of the training data sample reduces the tendency to over-fit a discriminant to the training data. In this particular analysis, presented in this thesis, we have made a step forward regarding variable optimization. Comparing to the previous single top cross section measurement analysis using BNN [154], this analysis [155] optimized the variables selection and used less number of variables, in order to decrease the complexity of the analysis.

The selection of variables for the BNN analysis is guided by the following premise: for the purposes of signal/background discrimination, the information contained in variables that are functions of the fundamental degrees of freedom of the events, principally the 4-vectors of identified objects [150], is no greater than that contained in the fundamental degrees of freedom. This mathematical fact is the basis of the matrix element (ME) method [156], which computes a semi-analytical approximation to the discriminant using the 4-vectors only. The virtue of searching for derived variables is that one may find a few that capture most of the relevant information in an event. Ideally, this will permit a reduction in the dimensionality of the problem and therefore a simplification of the analysis.

The virtue of an *ab initio* attempt to use the 4-vectors directly is that one knows a priori that this set, along with the charges and identity of the objects, comprises all the relevant physics information available in an event. The only possible drawback of such an approach

is that the computational method used may not be able to model the 4-vector distributions with sufficient accuracy nor account fully for the fact that, in practice, the identity of objects is ambiguous to differing degrees. In this analysis we choose to build the BNN using the lepton and jet 4-vectors, the missing transverse energy 2-vector and variables that include charge and b-tagging information.

The 4-vector for a particle can be taken to be either: $p_x$, $p_y$, $p_z$, and $E$, or $p_T$, $\eta$, $\phi$, and $m$. We use the latter set since this set is what we actually measure. In our case, the mass ($m$) of each final-state object, lepton, neutrino, or jet, is negligible, and therefore considered to be 0, leaving the variables $p_T$, $\eta$ and $\phi$. For the neutrino, we have only $p_T$ and $\phi$, where $p_T$ is approximated by the missing transverse energy (METPt). However, since we expect $\phi$-symmetry of the final-state objects for both signal and background, we can remove one degree of freedom. We remove it by breaking the $\phi$-symmetry: the $\phi$ of each object is measured with respect to that of the lepton. Since the lepton's $\phi$ is now always 0, it can be dropped from the variable set.

Each channel has 2 lepton variables, 2 neutrino variables, and 4 variables for each jet, resulting in $2 + 2 + 4 * (n_{\text{jets}}) = 12, 16$, and 20 variables for events with 2, 3, and 4 jets, respectively. The 4th jet variable is the b-tag weight. To add the charge information, we included the variable **QTimesEta** ($Q \times \eta$ of the leading light jet) in the set, except for the 2Jet 2Tag channel since, for this particular channel, this variable is not well modeled. In addition to these variables, we also included **WTransverseMass** (Transverse mass of the reconstructed $W$ boson) for all the 6 analysis channels resulting in a reduction of the total error for s+t expected cross section of $\sim 5\%$. The same lists of input variables are used for training $s + t$, $s$, and $t$ - modes, for all the 6 analysis channels.

Also, the variables selected for the analysis had underwent the Kolmogorov-Smirnov (KS) test to check the quantification of the data and MC agreement. This KS test takes as input the data and background model distribution for a given variable and returns a value between 0 and 1. If data and MC observables were derived from the same underlying function, the

resulting KS distribution for those observables will be uniform between 0 and 1. We have checked that the variables, which are selected should have KS value greater that 0.10. The list of selected variables for all the six analysis channels are described in Tables B.1 to B.6 separately, of Appendix B.

### 7.1.3 Training and Validation

#### Training

For each of the 6 analysis channels, a training set $T$ is built by combining 10,000 signal and 10,000 background unweighted events, which are sampled randomly from the signal and background weighted events in the training samples, and with each event having a probability of being picked by the sampling proportional to it weight. For this analysis, three separate training sets are created. One training set is for dedicated $t$-channel ($tqb$) which takes $t$-channel as signal and backgrounds include $s$-channel, $W$+jets, $t\bar{t}$, $Z$+jets, dibosons and multijets. Second training is for dedicated $s$-channel ($tb$) where $s$-channel as signal and backgrounds contain $t$-channel, $W$+jets, $t\bar{t}$, $Z$+jets, dibosons and multijets. Third training set is for $s+t$-channel ($tbtqb$) together and hence the signal for this case is $s+t$-channel and background contains $W$+jets, $t\bar{t}$, $Z$+jets, dibosons and multijets. For all the 6 channels in each training set, the structure of each neural network is fixed: each contains $N_{var}$ inputs which are listed in Tables B.1 to B.6 of Appendix B, and $H = 20$ hidden nodes. The numbers $N_{var}$ and $H$ define a $1 + (N_{var} + 2)H$-dimensional parameter space each point $w$ of which corresponds to an instance of a neural network function. From each training set $T$, a posterior density $p(w|T)$ over the network parameter space is constructed.

A Markov Chain Monte Carlo (MCMC) technique [151, 152] is used to draw a sample of networks out of the posterior density $p(w|t, x)$ as explained in Section 6.4 of Chapter 6. A complete cycle of running through the entire training is called an "epoch". Each iteration of the MCMC consists of 20 epochs and the result of each iteration is the average of the outputs of the 20 training epochs. The MCMC is run with 300 iterations and a sample of

100 networks pertaining to the last 100 networks is used to approximate the discriminant $D(x)$, (Eq. 7.2), via a Bayesian neural network, $\bar{y}(x)$, defined in Eq. 6.16 as

$$\bar{y}(x) \approx \frac{1}{100} \sum_{k=201}^{K=300} y(x, w_k), \tag{7.3}$$

where $x$ represents a vector with values of the input variables and $w_k$ are the adjustable parameters of the neural networks. The measure of the BNN performance is given by the area below the signal vs. background efficiency curve, which will be discussed later this section.

**Treatment of Noise in the Training Data**

The Bayesian neural network is sensitive to the presence of "noise" in the training data, which degrades its performance. By noise it is meant MC events with large weights. When training on such samples, the prior densities may become excessively broad and the MCMC sampling yields networks whose parameters have values that scatter over a large range. This leads to discriminant distributions that are excessively jagged and therefore bad approximations to the discriminant $D(x)$, Eq. 7.2. To overcome this problem, the network parameter prior widths are held fixed instead of letting them adapt to the training data. This limits the range of the network parameter values thereby favoring smoother approximations to the discriminant $D(x)$. A single neural network is trained for each of the six analysis channels separately, using the back-propagation algorithm as implemented in the JETNET program [157], which aims to minimize the error given in Eq. 6.4. The structure of the JETNET-trained networks is identical to that used for the BNNs, as are the training data. For each of the six JETNET-trained networks (one per channel) per each training set, the root-mean-square of the network parameters in each of the three groups (input-to-hidden weights $\sigma_u$, hidden-to-output weights $\sigma_a$, and hidden biases $\sigma_v$) is calculated. These root-mean-squares serve then as the standard deviations of the zero mean Gaussian priors in the BNN. The three standard deviations used in each of the six channels for all the three training

sets are shown in Table 7.1. The width $\sigma_b$ is set to a large value (= 100) so that the output bias is only minimally constrained. With this choice of prior densities, the BNN parameters are no longer driven to large values, resulting in smooth discriminants $D(x)$.

| | Values of the Gaussian Prior Widths used in the BNN | | | | | |
|---|---|---|---|---|---|---|
| | 1 $b$tag 2jet | 1 $b$tag 3jet | 1 $b$tag 4jet | 2 $b$tag 2jet | 2 $b$tag 3jet | 2 $b$tag 4jet |
| | | | $s+t$-channel ($tbtqb$) | | | |
| $\sigma_u$ | 0.24 | 0.17 | 0.18 | 0.22 | 0.20 | 0.20 |
| $\sigma_a$ | 0.38 | 0.36 | 0.44 | 0.31 | 0.30 | 0.49 |
| $\sigma_v$ | 0.68 | 0.59 | 0.68 | 0.66 | 0.69 | 0.74 |
| | | | $s$-channel ($tb$) | | | |
| $\sigma_u$ | 0.18 | 0.17 | 0.19 | 0.24 | 0.18 | 0.21 |
| $\sigma_a$ | 0.28 | 0.33 | 0.26 | 0.33 | 0.29 | 0.36 |
| $\sigma_v$ | 0.56 | 0.56 | 0.69 | 0.71 | 0.59 | 0.78 |
| | | | $t$-channel ($tqb$) | | | |
| $\sigma_u$ | 0.31 | 0.21 | 0.20 | 0.22 | 0.20 | 0.21 |
| $\sigma_a$ | 0.37 | 0.38 | 0.42 | 0.34 | 0.44 | 0.38 |
| $\sigma_v$ | 0.66 | 0.69 | 0.78 | 0.58 | 0.69 | 0.63 |

Table 7.1: Standard deviations for the gaussian priors for the input-to-hidden ($\sigma_u$), hidden-to-output weights ($\sigma_a$) and hidden biases ($\sigma_v$), obtained from a single neural network trained separately for each of the 6 channels and for three different training sets, using JETNET.

**Verification**

To check the convergence of the Markov Chain which is used to sample the $\bar{y}(x)$ from the posterior density $p(w|t,x)$ as discussed above, following diagnostic method is used.

Three ensembles, out of which two are independent ensembles of signal events, $E_{sig}$ and $E'_{sig}$, with each half of the statistics in the signal yield samples and a third ensemble, $E_{bkg}$, containing all the statistics in the background yield sample are created. All the three ensembles are normalized to the same number, $N$. The events in $E_{sig}$ and $E_{bkg}$ are weighted by their corresponding BNN output values. Mathematically, this is equivalent to multiplying

the densities $f_{E_{sig}}(x|S)$ and $f_{E_{bkg}}(x|B)$ by the BNN function $\bar{y}(x)$ :

$$f_n(x|S) = f_{E_{sig}}(x|S)\,\bar{y}(x), \tag{7.4}$$

$$f_n(x|B) = f_{E_{bkg}}(x|B)\,\bar{y}(x). \tag{7.5}$$

Since the number of events in the ensemble are large, the densities $f_{E_{sig}}(x|S)$ and $f_{E_{bkg}}(x|B)$ can be replaced by $Nf(x|S)$ and $Nf(x|B)$, since their shape is given by $f(x|S)$ and $f(x|B)$ and their normalization by $N$. Thus we can re-write above equations as:

$$f_n(x|S) = Nf(x|S)\,\bar{y}(x), \tag{7.6}$$

$$f_n(x|B) = Nf(x|B)\,\bar{y}(x). \tag{7.7}$$

And $g(x)$, the sum of the two weighted densities is defined as:

$$\begin{aligned} g(x) &= f_n(x|S) + f_n(x|B), \\ &= N\,[f(x|S) + f(x|B)]\,\bar{y}(x). \end{aligned} \tag{7.8}$$

As the Markov chain proceeds, we get

$$\bar{y}(x) \to D(x) \Longrightarrow g(x) \to Nf(x|S) \quad \text{(from Eq. 7.2)}$$

that is, one should recover the signal density. Thus, by comparing $g(x)$ extracted from $E_{sig} + E_{bkg}$, with the distributions from the independently signal-only sample $E'_{sig}$, it is possible to verify the convergence of the Markov Chain. In particular, one should recover all possible marginal densities, that is, projections to one-dimension:

$$\begin{aligned} \int_{x\notin z} g(z)\,dz &= \int_{x\notin z} f_n(z|S)\,dz + \int_{x\notin z} f_n(z|B)\,dz, \\ &= m_n(x|S) + m_n(x|B) \to N\,m(x|S). \end{aligned} \tag{7.9}$$

First row of Figures 7.1 to 7.6 shows the verification plots for three of the variables used in the $s+t$-channel training set for all the 6 analysis channels. For plotting simplicity, the values of each of these variables are shifted and scaled in an event-by-event basis such that the resulting distributions have zero mean and unit variance.

The black dots represent the signal distribution $N\,m(x|S)$ for variable $x$ extracted from ensemble $E'_{sig}$. The blue and green histograms correspond to the distributions $m_n(x|S)$ and $m_n(x|B)$, which are the one-dimensional projections of the densities $f_n(x|S)$ and $f_n(x|B)$ in the variable $x$, and the red histogram (sum of the blue and green ones) is the projection of the sum $g(x)$. Thus, according to what has been just explained (and summarized in Eqs. 7.6 to 7.9), the convergence should be judged by the agreement between the black dots and the red histogram. The convergence was checked in all the input variables used in each channel for all the training sets, and good agreement was found between the black dots and red histogram, indicating that the Markov chains have converged.

An equivalent verification test can be performed directly on the BNN output. The second row of Figures 7.1 to 7.6 shows the additional plots using BNN output. The left plot shows the BNN output for signal (blue) and background (green) normalized to unity. Denoting by $S$ and $B$ the amount of signal and background that are left after cutting on the BNN output, the plot at the center shows $S/\sqrt{B}$ and $S/\sqrt{S+B}$ versus the BNN output cut. It is important to recall that this is only a performance test and that no cut on the BNN output is ultimately done; the cross section is measured taking into account the whole range of BNN discriminant. Finally, the right plot shows the signal efficiency ($\epsilon_S$) versus the background efficiency ($\epsilon_B$) when cutting on different values of the BNN output. In this last plot, the area above the curve is shown as a way of quantifying the separation power of the BNN between signal and background. The lower this number, the better the separation.

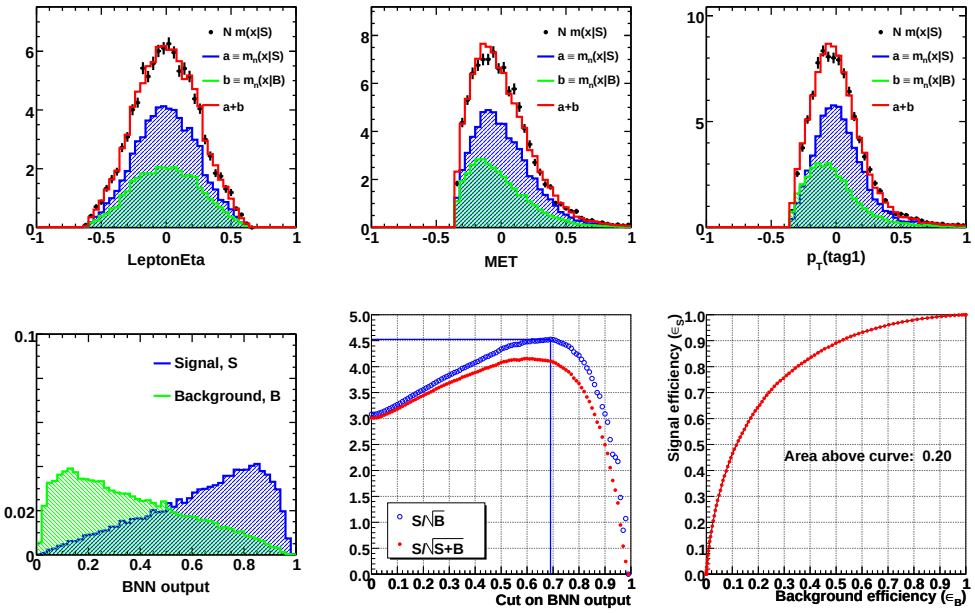


Figure 7.1: Top Row: verification plots, Bottom Row: BNN outputs normalized to unity (green: background, blue: signal), and the $S/\sqrt{B}$, $S/\sqrt{S+B}$, and signal efficiency ($\epsilon_S$) versus background efficiency ($\epsilon_B$) curves, for the =1tag, =2jet of s+t-channel.

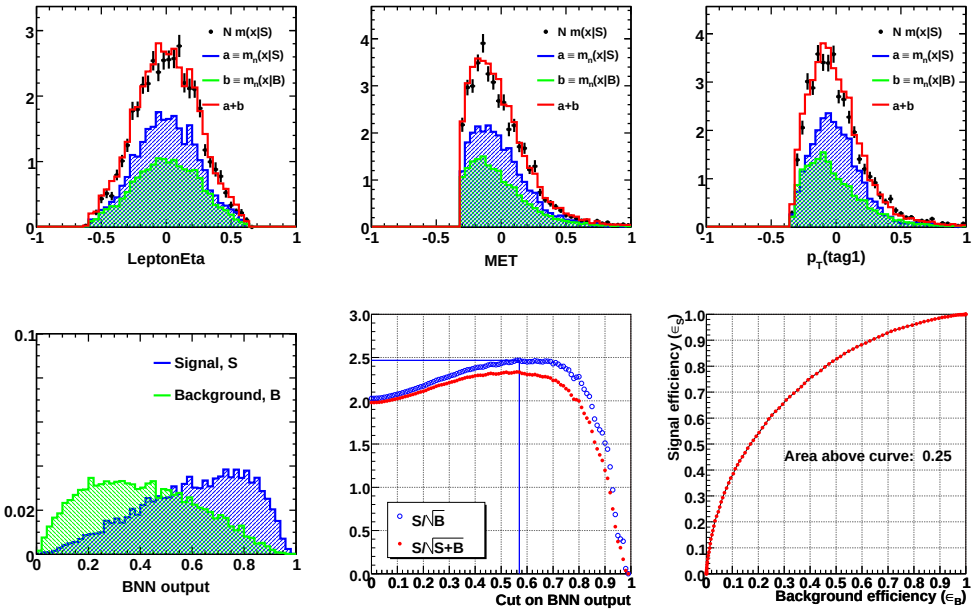


Figure 7.2: Top Row: verification plots, Bottom Row: BNN outputs normalized to unity (green: background, blue: signal), and the $S/\sqrt{B}$, $S/\sqrt{S+B}$, and signal efficiency ($\epsilon_S$) versus background efficiency ($\epsilon_B$) curves, for the =1tag, =3jet of s+t-channel.

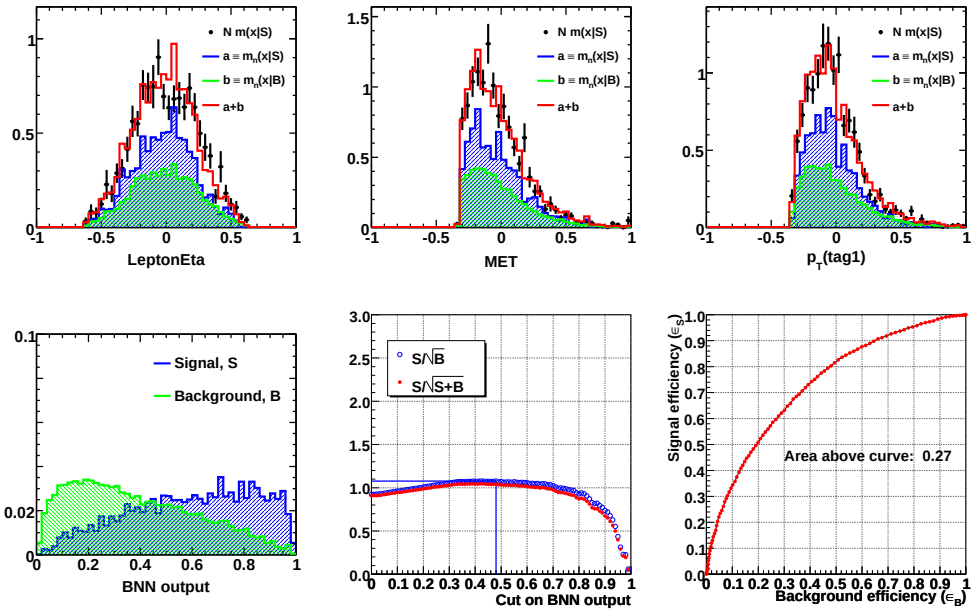


Figure 7.3: Top Row: verification plots, Bottom Row: BNN outputs normalized to unity (green: background, blue: signal), and the $S/\sqrt{B}$, $S/\sqrt{S+B}$, and signal efficiency ($\epsilon_S$) versus background efficiency ($\epsilon_B$) curves, for the =1tag, =4jet of s+t-channel.

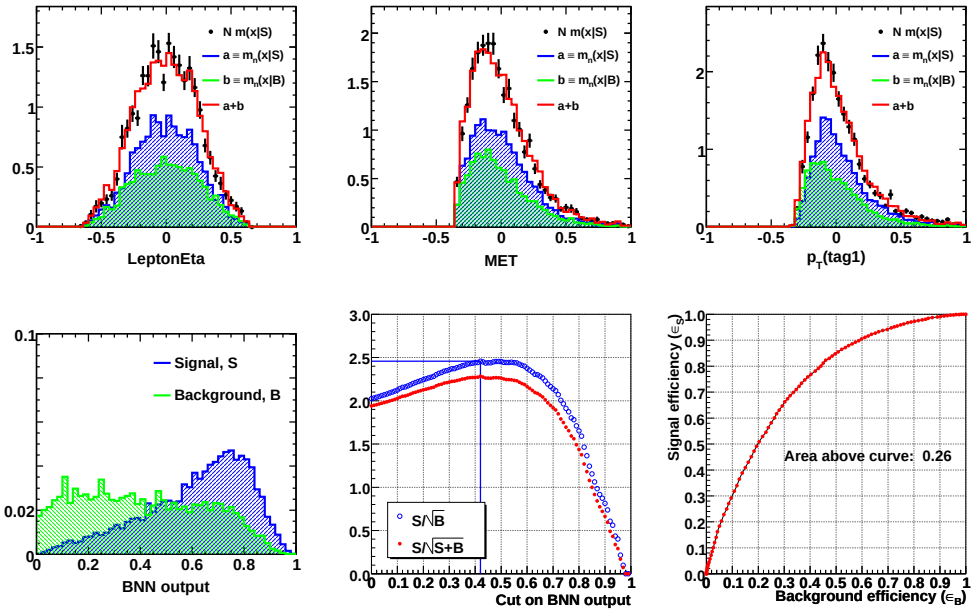


Figure 7.4: Top Row: verification plots, Bottom Row: BNN outputs normalized to unity (green: background, blue: signal), and the $S/\sqrt{B}$, $S/\sqrt{S+B}$, and signal efficiency ($\epsilon_S$) versus background efficiency ($\epsilon_B$) curves, for the =2tag, =2jet if s+t-channel.

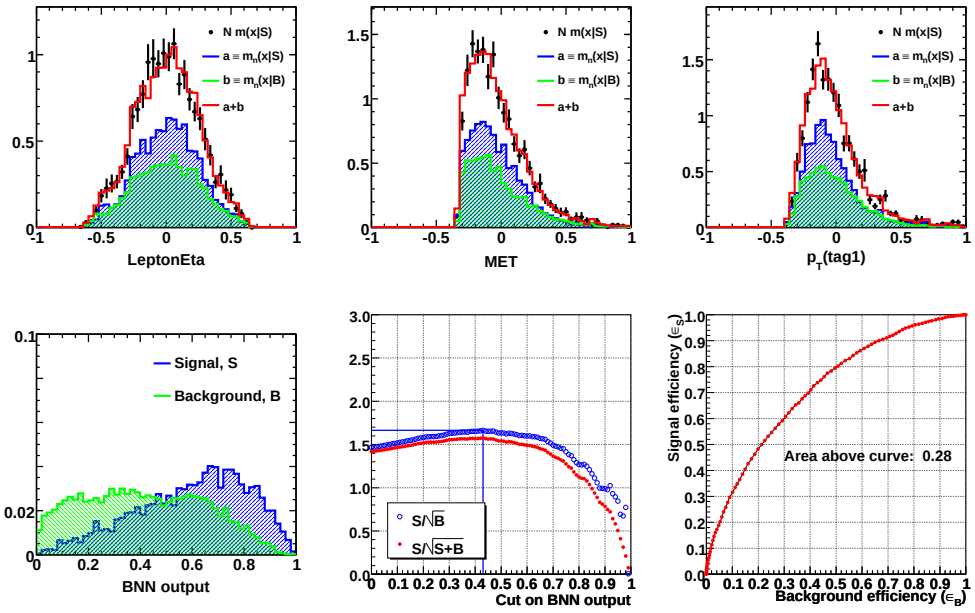


Figure 7.5: Top Row: verification plots, Bottom Row: BNN outputs normalized to unity (green: background, blue: signal), and the $S/\sqrt{B}$, $S/\sqrt{S+B}$, and signal efficiency ($\epsilon_S$) versus background efficiency ($\epsilon_B$) curves, for the =2tag, =3jet of s+t-channel.

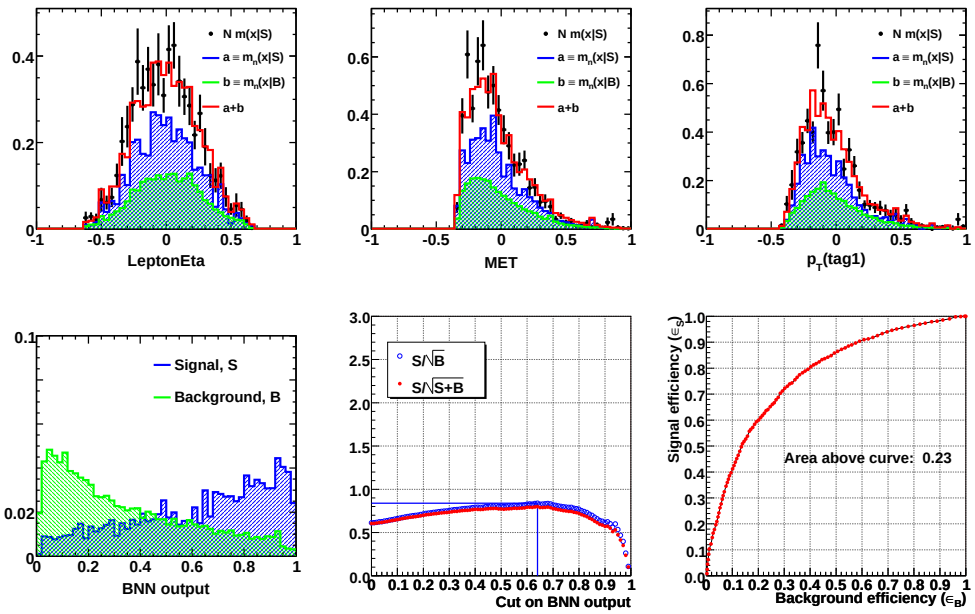


Figure 7.6: Top Row: verification plots, Bottom Row: BNN outputs normalized to unity (green: background, blue: signal), and the $S/\sqrt{B}$, $S/\sqrt{S+B}$, and signal efficiency ($\epsilon_S$) versus background efficiency ($\epsilon_B$) curves, for the =2tag, =4jet of s+t-channel.

### 7.1.4 BNN Output Binning Transformation

After the training is done and after checking the convergence of BNN, the next step is to apply BNN filter given by Eq. 7.3 to the signal and background events of the yield sample. The resulting BNN discrimination output is a number between 0 and 1, with background-like (signal-like) events populating mostly the low (high) discriminant region as can be seen from top plot in Figure 7.7. Due to this kind of distribution, there may be some bins in which there are some signal but no background events, specially in high discriminant region. This can create problems with the stability in the cross section calculation and hence results in large MC statistical uncertainty. To overcome this problem, a binning transformation is applied to BNN discriminant outputs which ensures that there is a minimum amount of effective background events in each bin.

For each of the six analysis channels of all the training sets, a monotonic re-binning scheme is used corresponding to the output distributions for the backgrounds. First of all, the background sum is very finely separated in 10,000 uniform bins and normalized to unity. This 10,000-bin input distribution is then re-arranged such that the transformed discriminant output for the background sum follows a $1/x$-curve between 0.05 and 0.8 (normalized to 1 over this range) and a linear slope from the intersept of the $1/x$ graph at 0.8 down to zero at 0.95. There are no shape constraints between 0.95 and 1.0, but requiring at least 4 effective background events in each of the output bins of width 0.0025 (i.e. 400 bins between 0 and 1). The transformation is then constrained so that when put into a histogram with 25 bins, each bin with 0.04 bin-width will contain at least 64 effective background events. Figure 7.7 shows the transformation function (bottom left), and the signal and background BNN output distributions before (top) and after the binning transformation (bottom right). This condition ensures that the MC statistical uncertainty stays reasonably small. This binning transformation increases the resolution in the signal region without introducing statistical fluctuations.

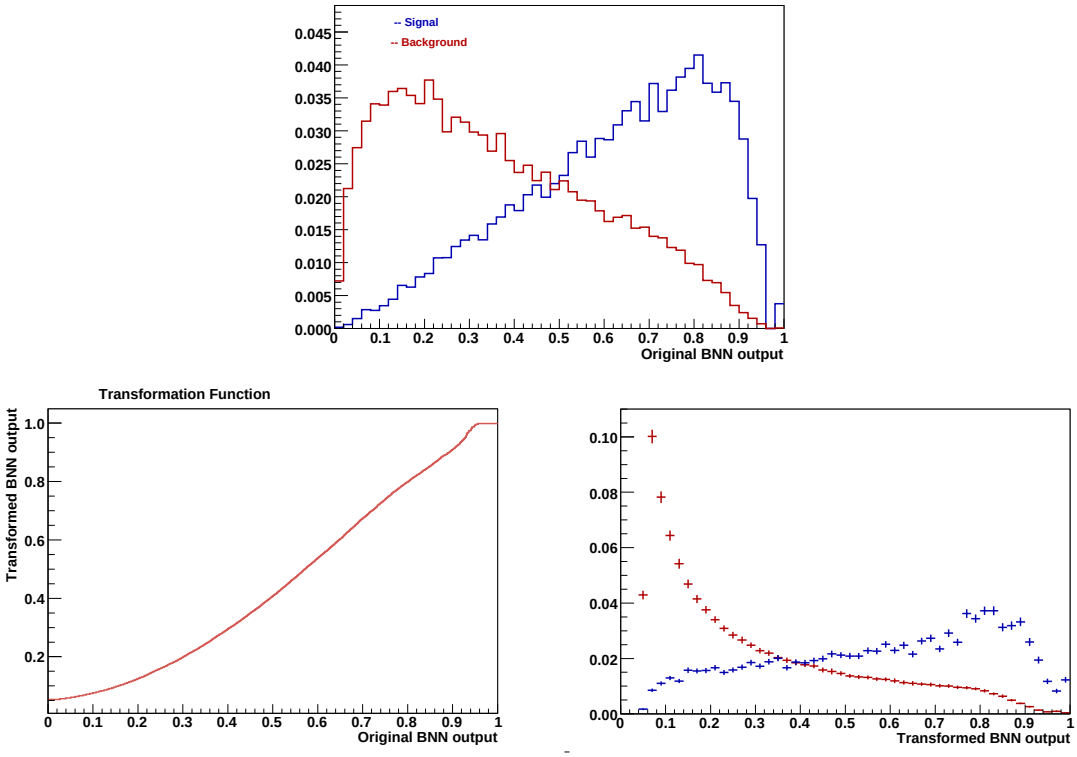


Figure 7.7: (Top) The BNN discriminant output for tb+tqb signal (blue) and all backgrounds combined (red) before binning transformation. (Bottom Left) The derived monotone transformation function. (Bottom Right): Same discriminant distributions as in the Top plot, but after the monotone transformation.

### 7.1.5 BNN Discriminant Output

This section shows the resulting BNN output distributions after applying BNN filter to signal and background yield sample and after proper binning transformation. Figure 7.8 shows the discriminant plots for $s+t$-channel (labelled as "*tbtqb*") for all the six analysis channels in the first two rows and all channels combined $s+t$-channel discriminant at bottom. Similarly, Figures 7.9 and 7.10 show the discriminant outputs for $s$ and $t$-channels, (labeled as "*tb*" and "*tqb*"') respectively, following the same schema. The color scheme used in these plots to label the data and the different signal and background components is shown in Figure 5.8 of Chapter 5.

BNN OUTPUT FOR $s+t$ CHANNEL

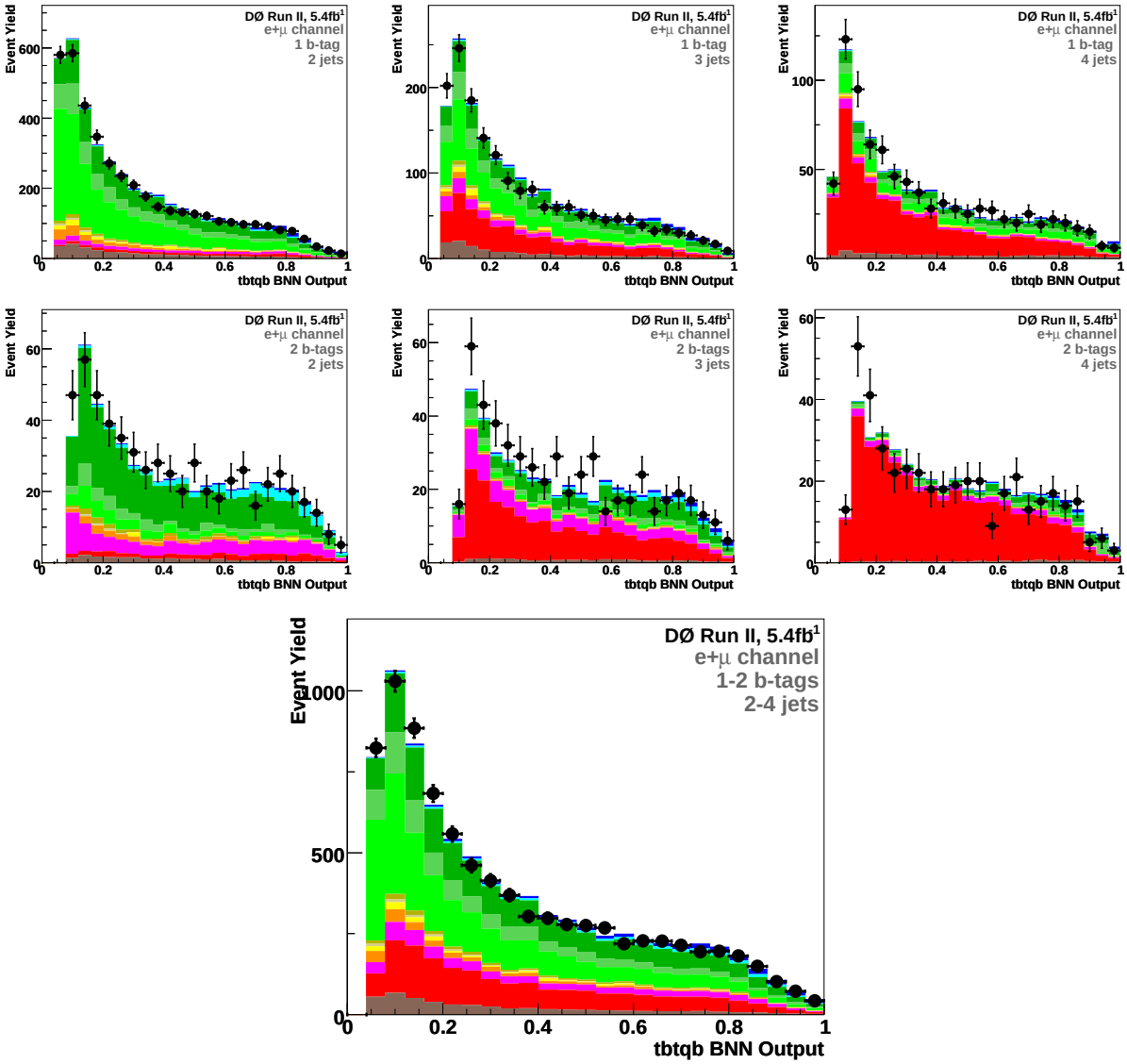


Figure 7.8: BNN output for $s+t$-channel. First two rows show the results for the six individual analysis channels: 2 jets (left), 3 jets (middle), 4 jets (right); single tag (top row), double tag (second row). The bottom plot shows the discriminant output for the six analysis channels combined.

BNN OUTPUT FOR $s$ CHANNEL

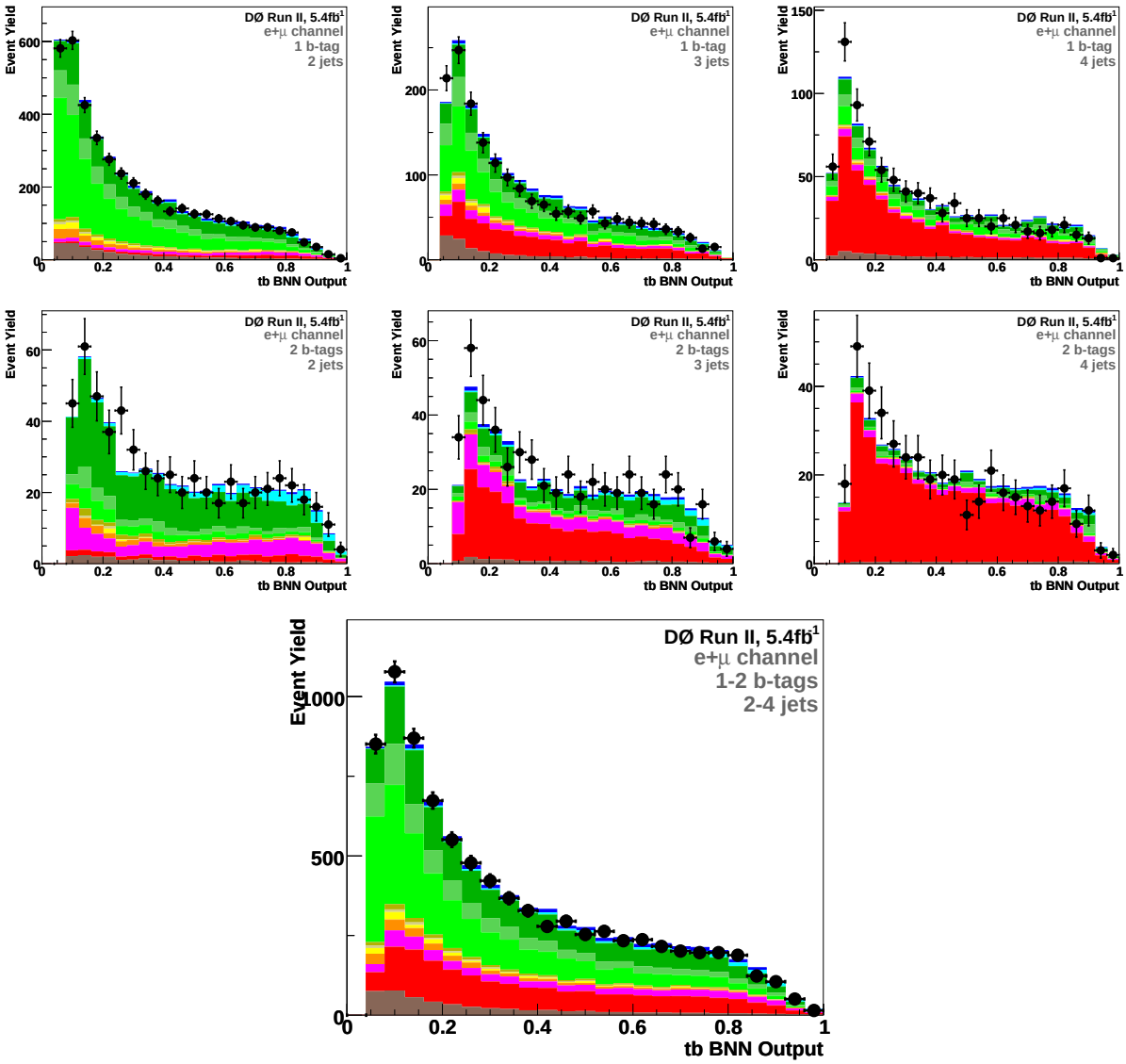


Figure 7.9: BNN output for $s$-channel. First two rows show the results for the six individual analysis channels: 2 jets (left), 3 jets (middle), 4 jets (right); single tag (top row), double tag (second row). The bottom plot shows the discriminant output for the six analysis channels combined.

BNN OUTPUT FOR $t$ CHANNEL

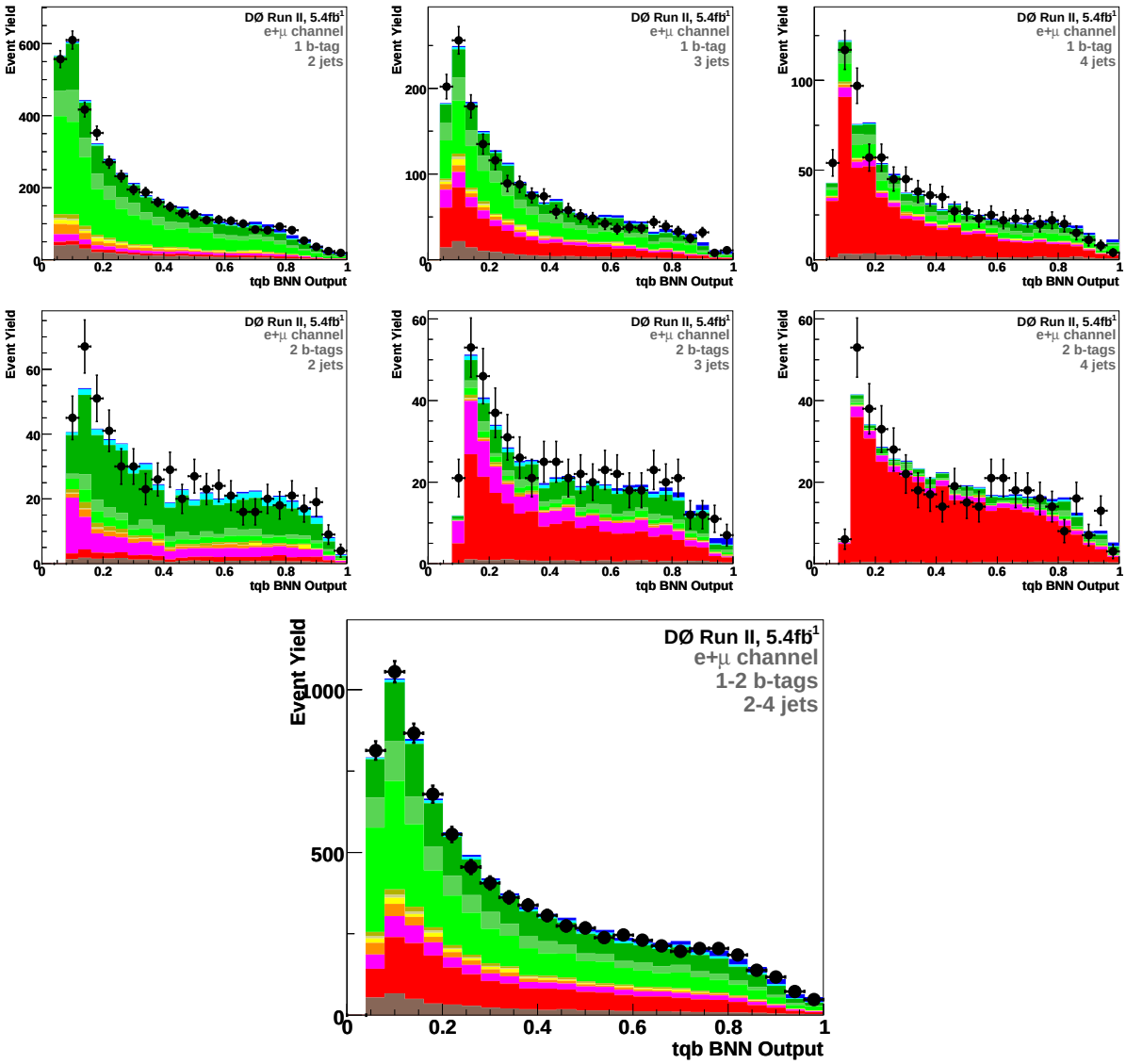


Figure 7.10: BNN output for $t$-channel. First two rows show the results for the six individual analysis channels: 2 jets (left), 3 jets (middle), 4 jets (right); single tag (top row), double tag (second row). The bottom plot shows the discriminant output for the six analysis channels combined.

### 7.1.6 BNN Cross Checks

In order to validate every step of the BNN analysis, another cross-check has been done. As discussed in Section 5.5 of Chapter 5, two cross check samples are created to test whether the background model reproduces the data in regions dominated by one type of background. The selection cuts used to create these samples are described in Section 5.5. This section shows the BNN output distributions for these cross-check samples. Figure 7.11, 7.12 and 7.13 show the BNN output distributions for these cross-check samples for $s+t$, $s$ and $t$-channels, respectively. Good agreement is seen between data and MC background models for both of the two main backgrounds in the analysis. Thus, leads to the conclusion that the background model describes the data well within uncertainties.

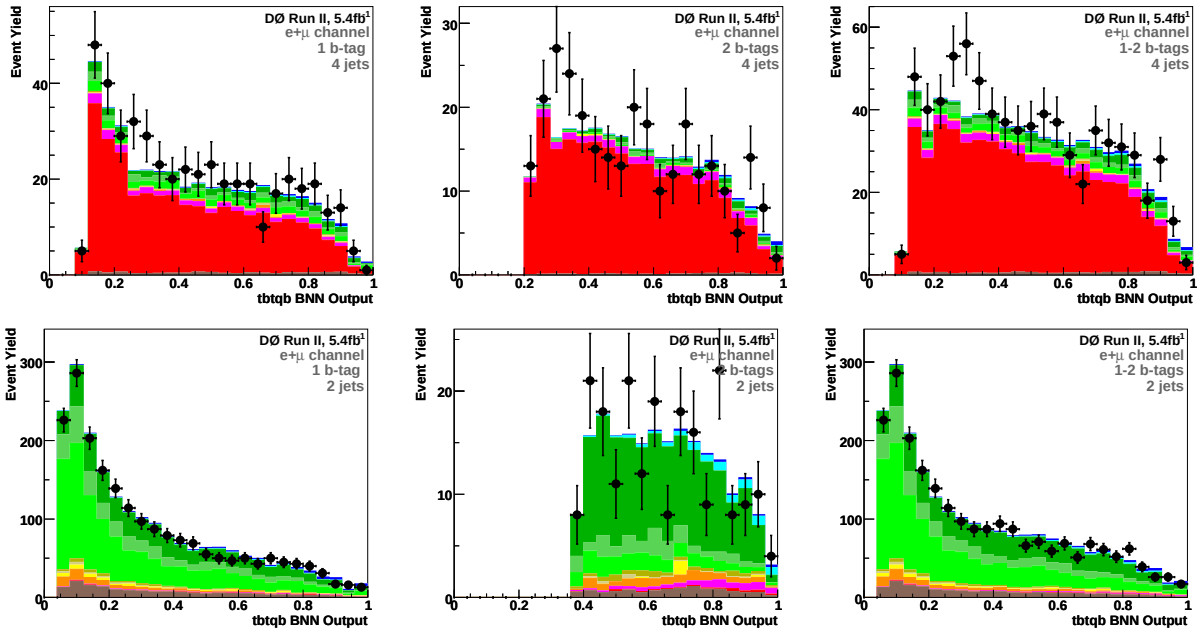


Figure 7.11: BNN cross check samples for $s+t$ channel. First row: $t\bar{t}$ cross check samples, Second row: $W$+jets cross check samples.

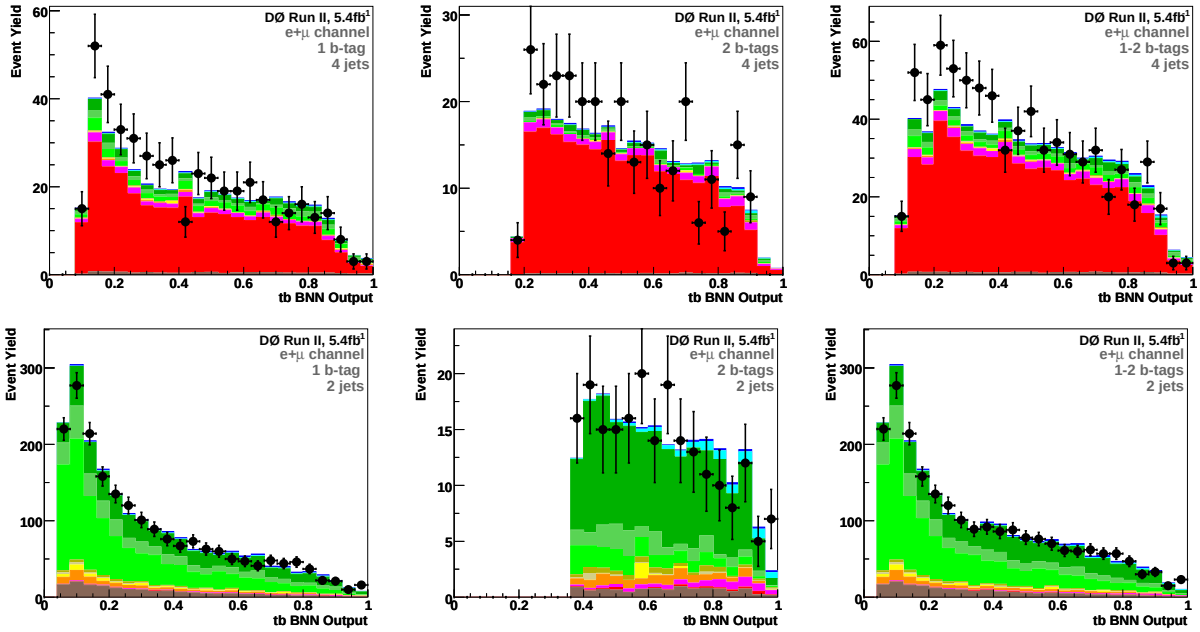


Figure 7.12: BNN cross check samples for $s$ channel. First row: $t\bar{t}$ cross check samples, Second row: $W$+jets cross check samples.

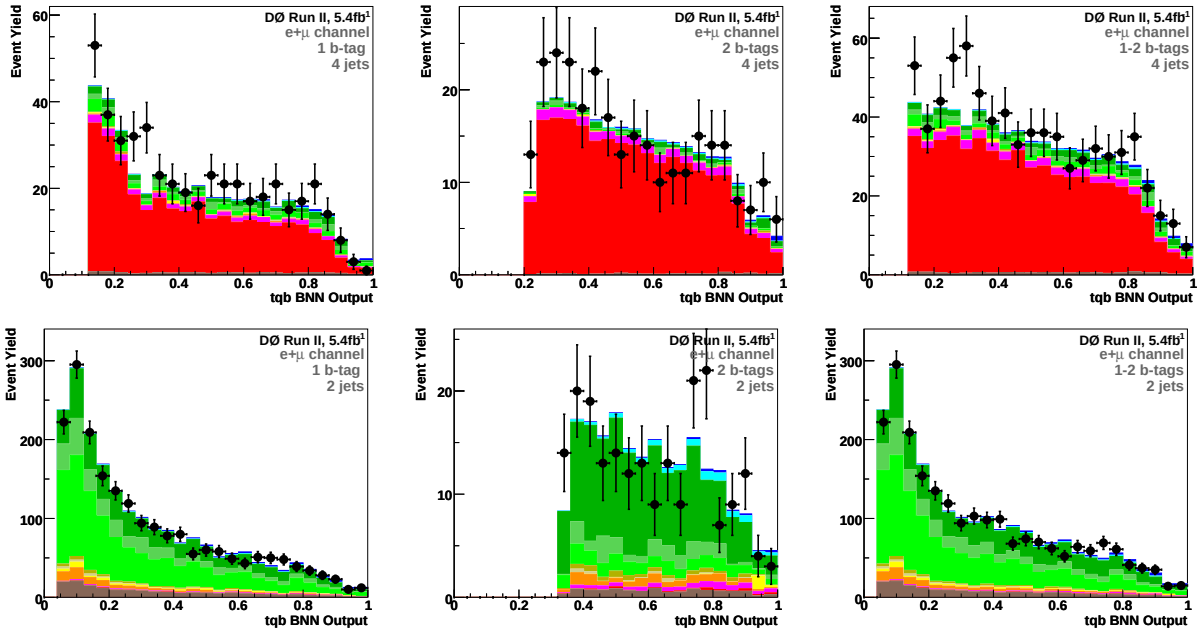


Figure 7.13: BNN cross check samples for $t$ channel. First row: $t\bar{t}$ cross check samples, Second row: $W$+jets cross check samples.

## 7.2 Measurement of Single Top Cross Section

This section describes how the cross section measurement is done using the final BNN output distributions shown in Figures 7.8 to 7.10 by the application of Bayesian approach and how this measurement is validated and calibrated by creating ensembles of pseudo-datasets [158, 159]. In this section expected and observed cross section measurements for single top $s+t$, $s$ and $t$-channels, separately, are shown. And finally, how the result obtained from BNN is combined with other multivariate methods and hence the measurement of CKM matrix element $|V_{tb}|$ is described in next two Sections.

### 7.2.1 Bayesian Statistical Approach

The probability (proportional to likelihood) to observe a count $D$, if the mean count is $d$, is given by the Poisson distribution:

$$P(D|d) = \frac{e^{-d}\ d^{D}}{\Gamma(D+1)} \tag{7.10}$$

where $\Gamma$ is the gamma function. The mean count $d$ is a sum of the predicted contributions from the signal $s$ and $N$ sources of background $b$:

$$\begin{aligned} d &= s + \sum_{i=1}^{N} b_i = \sigma(\alpha l) + \sum_{i=1}^{N} b_i, \\ &= \sigma a + \sum_{i=1}^{N} b_i \end{aligned} \tag{7.11}$$

where $\alpha$ is the signal acceptance, $l$ is the integrated luminosity; $a \equiv \alpha l$ is the effective luminosity for the signal. A combined likelihood can be constructed as the product of the single-bin likelihoods for dealing with many bins from a single or several histograms [32, 160]:

$$L(\mathbf{D}|\mathbf{d}) \equiv L(\mathbf{D}|\sigma, \mathbf{a}, \mathbf{b}) = \prod_{j=1}^{N_{bins}} P(D_j|d_j) \tag{7.12}$$

where $\mathbf{D}$ and $\mathbf{d}$ are vectors of the observed and predicted number of events in each bin, and $\mathbf{a}$ and $\mathbf{b}$ are vectors of effective luminosity and background yields.

The likelihood function $L(\mathbf{D}|\mathbf{d}) \propto P(\mathbf{D}|\mathbf{d})$. Bayes' theorem as discussed in Section 6.2 can be used to invert the likelihood function, and obtain:

$$P(\mathbf{d}|\mathbf{D}) \equiv P(\sigma, \mathbf{a}, \mathbf{b}|\mathbf{D}) = \frac{L(\mathbf{D}|\mathbf{d})\pi(\mathbf{d})}{\mathcal{N}}, \tag{7.13}$$

where $\mathcal{N}$ is an overall normalization and is obtained from the requirement $\int P(\sigma|\mathbf{D})d\sigma = 1$. Further, this can be integrated over $\mathbf{a}$ and $\mathbf{b}$ termed as the *nuisance* parameters, to obtain the posterior probability density for the signal cross section, given the observed count $\mathbf{D}$:

$$P(\sigma|\mathbf{D}) = \frac{1}{\mathcal{N}} \iint L(\mathbf{D}|\sigma, \mathbf{a}, \mathbf{b})\pi(\sigma, \mathbf{a}, \mathbf{b})d\mathbf{a}d\mathbf{b}. \tag{7.14}$$

The quantity $\pi(\sigma, \mathbf{a}, \mathbf{b})$ is the prior probability that encodes what is known about the parameters, $\sigma$, $\mathbf{a}$, $\mathbf{b}$. Assuming that any prior knowledge of $\mathbf{a}$ and $\mathbf{b}$ is independent of the signal cross section $\sigma$, and therefore factorize the prior probability as follows:

$$\pi(\sigma, \mathbf{a}, \mathbf{b}) = \pi(\sigma)\pi(\mathbf{a}, \mathbf{b}). \tag{7.15}$$

The prior probability density $\pi(\mathbf{a}, \mathbf{b})$ encodes all knowledge of the effective signal luminosity and background yields, including all systematic uncertainties and their correlations. A flat prior for $\sigma$ is used such that:

$$\begin{aligned} \pi(\sigma) &= \frac{1}{\sigma_{max}}, 0 < \sigma < \sigma_{max} \\ &= 0, \qquad \text{otherwise,} \end{aligned} \tag{7.16}$$

where $\sigma_{max}$ is the sufficiently high upper bound on the cross section such that the posterior probability for $\sigma > \sigma_{max}$ is negligible. This choice of prior implies a maximal ignorance of any preferred value for the signal cross section, and is best regarded as a *convention*. The

posterior probability density for the signal cross section is therefore:

$$P(\sigma|\mathbf{D}) = \frac{1}{\mathcal{N}\sigma_{max}} \iint L(\mathbf{D}|\sigma, \mathbf{a}, \mathbf{b})\pi(\mathbf{a}, \mathbf{b})d\mathbf{a}d\mathbf{b}. \tag{7.17}$$

The resulting peak position of the posterior distribution $P(\sigma|\mathbf{D})$ is regarded as the cross section measurement, and the 68.27% interval about the peak as an estimate of its uncertainty, as shown in Figure 7.14. In the case the posterior is too broad or the peak is too close to zero, the lower bound may reach and stay at zero and the upper bound is further moved until 68.27% of the area is reached.

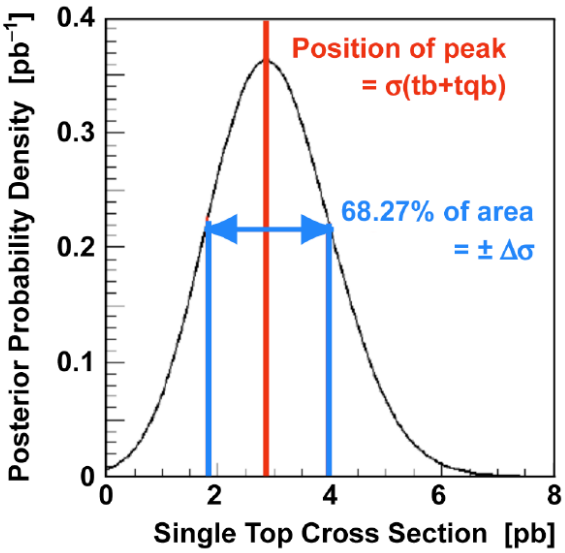


Figure 7.14: Illustration of the posterior density $P(\sigma|\mathbf{D})$. The measured cross section is the peak position $\sigma_{peak}$, and the uncertainty of the measurement is the interval $\Delta\sigma$ covering 68.27% of the posterior as indicated in the plot.

### Prior Modeling - Modeling the effects of Systematic Uncertainties

The systematic uncertainties enter in the modeling of the prior density $\pi(\mathbf{a}, \mathbf{b})$. The integral in Eq. 7.17 is done numerically using Monte Carlo importance sampling. A large number, $\sim$ 1M ($N_{sample}$) of points $(\mathbf{a_n}, \mathbf{b_n})$ are generated randomly by sampling the prior density

$\pi(\mathbf{a}, \mathbf{b})$. The posterior density is then estimated as:

$$P(\sigma|\mathbf{D}) \approx \frac{1}{\mathcal{N}\sigma_{max}N_{sample}} \sum_{n=1}^{N_{sample}} L(\mathbf{D}|\sigma, \mathbf{a_n}, \mathbf{b_n}) \tag{7.18}$$

Uncertainties that affect the normalization only for example, integrated luminosity, lepton-ID corrections etc. are modeled as the widths of Gaussian distributions with means set to the expected yields. The shape-changing uncertainties are modeled bin by bin by evaluating the effect of shifting the uncertainty up and down by one standard deviation.

### 7.2.2 Expected Cross Section

In this section, expected cross section measurement for single top $s + t$, $s$ and $t$-channels is presented. The cross section is measured using the Bayesian technique described above. Before looking into the real data and getting the observed measurement, it is always good and advisable to run the analysis with only MC values and get the *expected* measurement of the cross section. By seeing the expected result, we can safely tune up the parameters like the upper bound on the prior ($\sigma_{max}$), number of integration steps etc. to increase the sensitivity of the result which can not be done while having observed result first in order not to bias the measurement.

The cross section is measured by forming a binned likelihood as a product over all six channels (2, 3 or 4 jets with 1 or 2 $b$-tags) and bins using the full discriminant outputs shown in Figures 7.8 to 7.10 without any cut applied, but by setting the number of data events in each analysis channel equal to the expected number of background events plus the number of signal events predicted when using the SM cross sections of 1.04 pb for $s$-channel and 2.26 pb for $t$-channel (for the top mass of 172.5 GeV) and hence called *expected* cross section. The measurement is done separately for $s + t$-channel (considering $s + t$ as signal in $s$:$t$ proportions given by the SM and all other contributions as backgrounds), $s$-channel (considering $s$ as signal and adding the $t$ contribution to the backgrounds as a Gaussian

prior centered on the SM prediction) and for $t$-channel (taking $t$ as signal and adding the $s$ contribution to the backgrounds as a Gaussian prior centered on the SM prediction).

Figures 7.15 to 7.17 shows the Bayesian posterior probability density distribution for expected results, for the individual six analysis channels and for all the six channels combined, and including all the systematics. The vertical solid line shows the position of the peak and the vertical dashed lines show the bounds of the 68.3%, 95.4%, 99.7%, 99.994% and 99.99994% two-sided confidence intervals.

Table 7.2 summarizes the expected cross sections with all systematic uncertainties included for the six analysis and for all channels combined, for all the three $s + t$, $s$ and $t$-modes. The expected values is found to be in good agreement with the SM values of 3.30 $pb$ ($s + t$-channel), 1.04 $pb$ ($s$-channel) and 2.26 $pb$ ($t$-channel). Also from the Table 7.2, it is clear that the resulting uncertainty on the cross section is reduced by combining all the six channels. Table 7.3 shows the "peak over half-width" defined as the ratio of the peak position of the posterior over the size of the interval given by the peak position and the position of the lower bound of the 68.27% two-sided confidence interval. This provides an approximation to the sensitivity of the result.

Expected Cross Section FOR $s+t$ CHANNEL

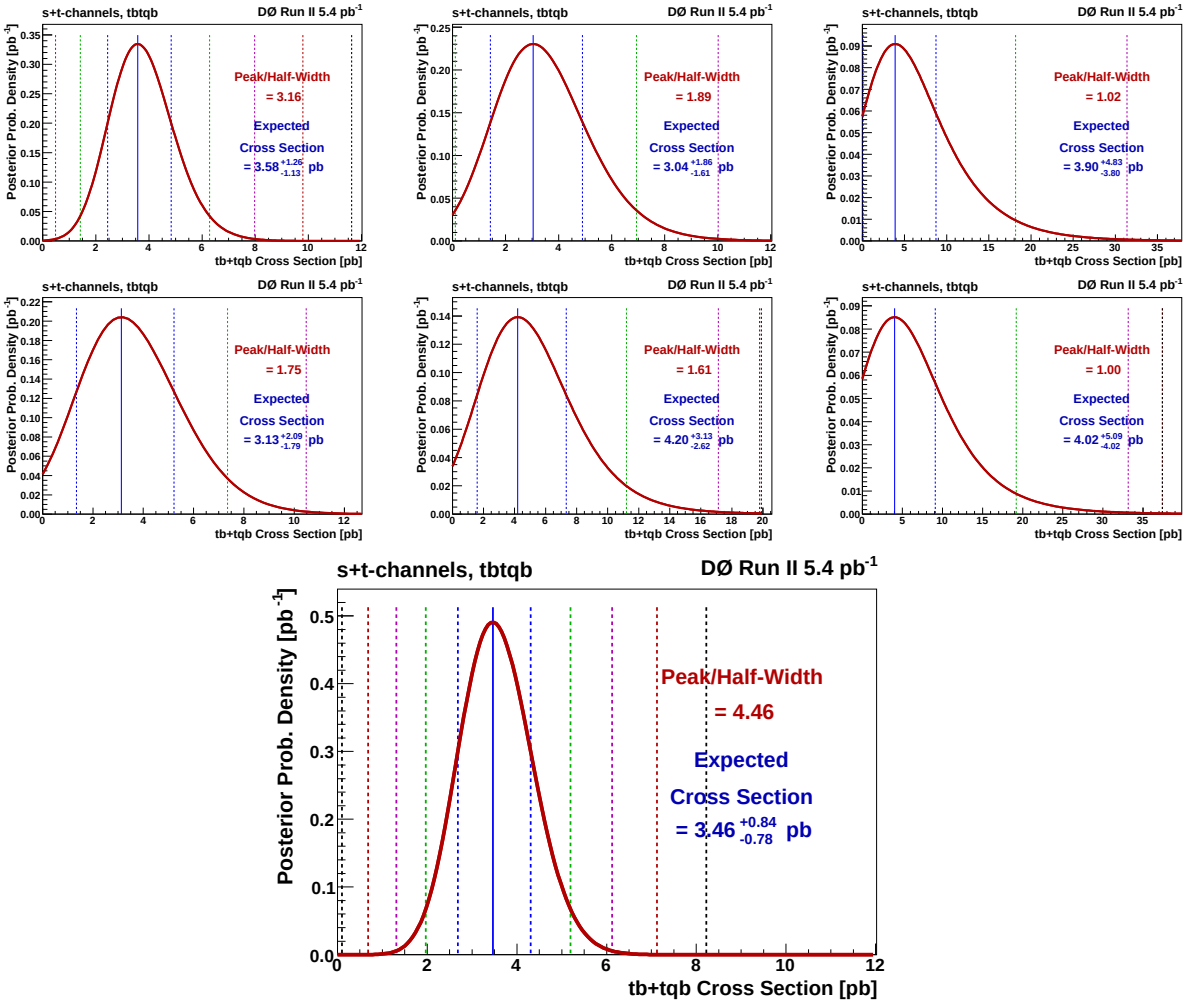


Figure 7.15: Expected posterior density for $s+t$-channel. First two rows show the results for the six individual analysis channels: 2 jets (left), 3 jets (middle), 4 jets (right); single tag (top row), double tag (second row). The bottom plot shows the expected posterior density for all the six analysis channels combined.

Expected Cross Section FOR $s$ CHANNEL

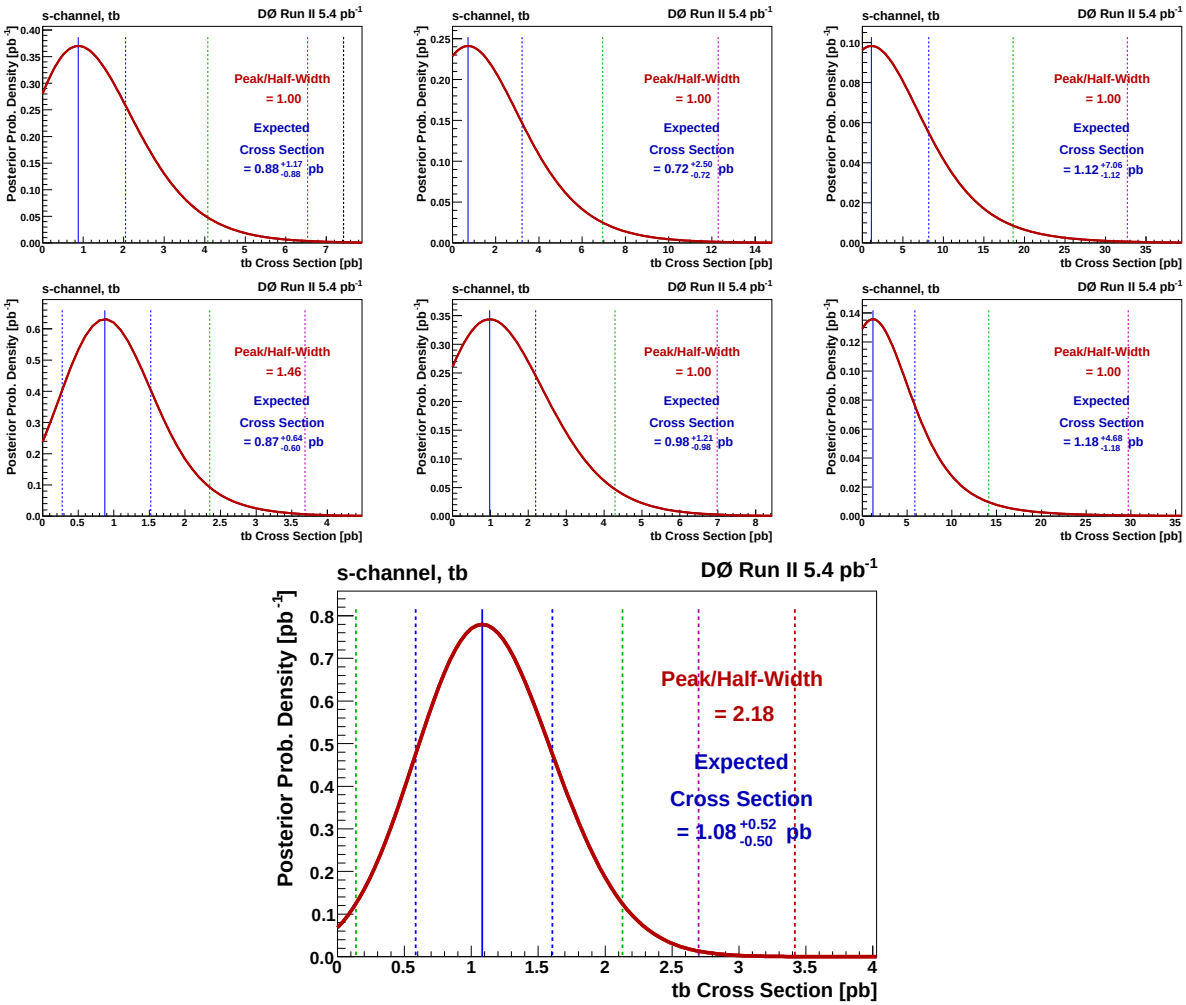


Figure 7.16: Expected posterior density for $s$-channel. First two rows show the results for the six individual analysis channels: 2 jets (left), 3 jets (middle), 4 jets (right); single tag (top row), double tag (second row). The bottom plot shows the expected posterior density for all the six analysis channels combined.

Expected Cross Section FOR $t$ CHANNEL

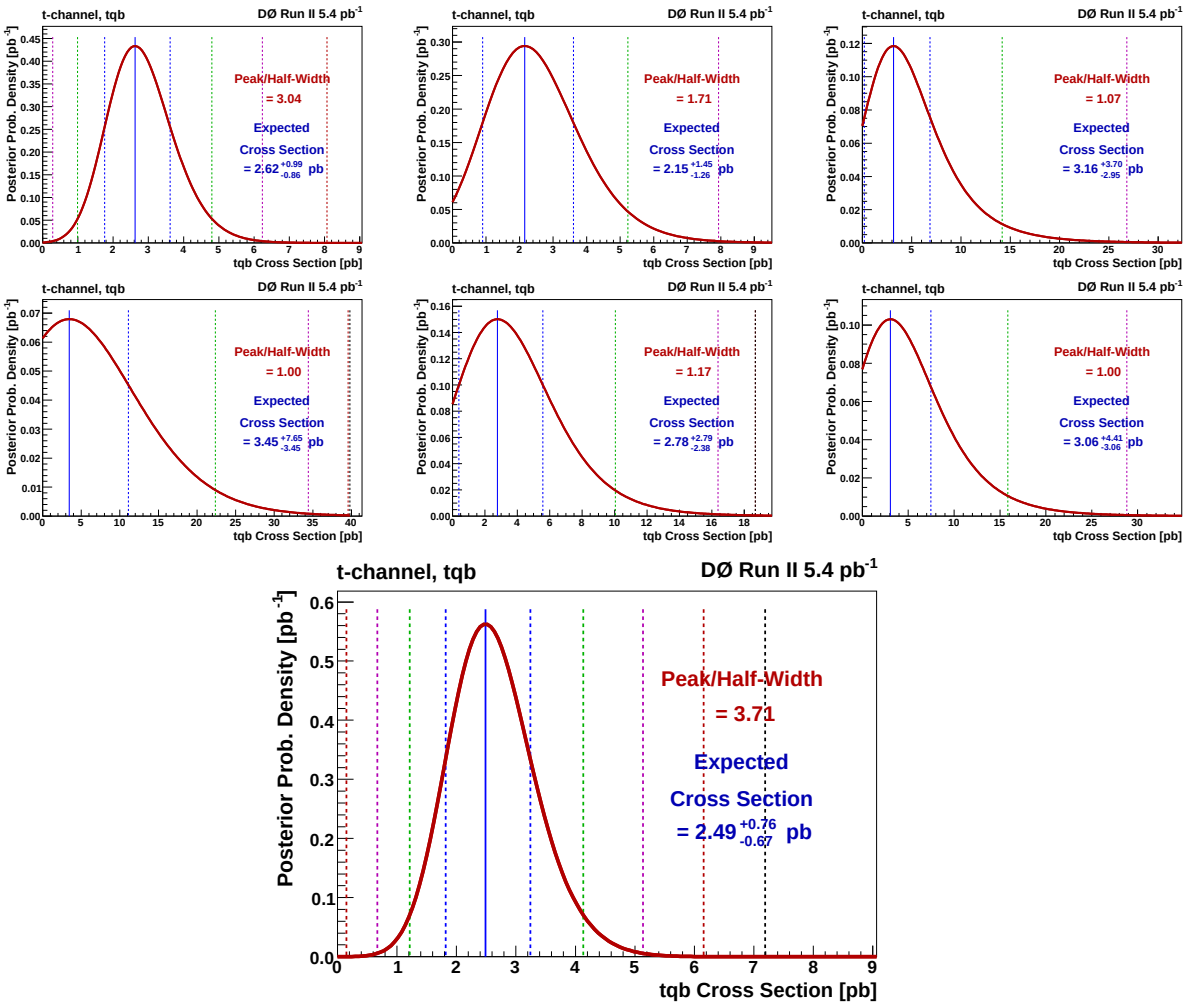


Figure 7.17: Expected posterior density for $t$-channel. First two rows show the results for the six individual analysis channels: 2 jets (left), 3 jets (middle), 4 jets (right); single tag (top row), double tag (second row). The bottom plot shows the expected posterior density for all the six analysis channels combined.

| Channel | Expected Cross Sections $(tb+tqb)$ $s+t$-mode | $(tb)$ $s$-mode | $(tqb)$ $t$-mode |
|---|---|---|---|
| 1tag / 2jets | $3.58^{+1.26}_{-1.13}$ | $0.88^{+1.17}_{-0.88}$ | $2.62^{+0.99}_{-0.86}$ |
| 1tag / 3jets | $3.04^{+1.86}_{-1.61}$ | $0.72^{+2.50}_{-0.72}$ | $2.15^{+1.45}_{-1.26}$ |
| 1tag / 4jets | $3.90^{+4.83}_{-3.80}$ | $1.12^{+7.06}_{-1.12}$ | $3.16^{+3.70}_{-2.95}$ |
| 2tag / 2jets | $3.13^{+2.09}_{-1.79}$ | $0.87^{+0.64}_{-0.60}$ | $3.45^{+7.65}_{-3.45}$ |
| 2tag / 3jets | $4.20^{+3.13}_{-2.62}$ | $0.98^{+1.21}_{-0.98}$ | $2.78^{+2.79}_{-2.38}$ |
| 2tag / 4jets | $4.02^{+5.09}_{-4.02}$ | $1.18^{+4.68}_{-1.18}$ | $3.06^{+4.41}_{-3.06}$ |
| 1-2tags / 2-4jets | $\mathbf{3.46^{+0.84}_{-0.78}}$ | $\mathbf{1.08^{+0.52}_{-0.50}}$ | $\mathbf{2.49^{+0.76}_{-0.67}}$ |

Table 7.2: Expected cross section, with all systematic uncertainties taken into account, for six analysis channels and for all the three modes. The values from all channels combined are shown in bold type.

| Channel | Expected Posterior Peak Over Half-Width $(tb+tqb)$ $s+t$-mode | $(tb)$ $s$-mode | $(tqb)$ $t$-mode |
|---|---|---|---|
| 1tag / 2jets | 3.16 | 1.00 | 3.04 |
| 1tag / 3jets | 1.89 | 1.00 | 1.71 |
| 1tag / 4jets | 1.02 | 1.00 | 1.07 |
| 2tag / 2jets | 1.75 | 1.46 | 1.00 |
| 2tag / 3jets | 1.61 | 1.00 | 1.17 |
| 2tag / 4jets | 1.00 | 1.00 | 1.00 |
| 1-2tags / 2-4jets | **4.46** | **2.18** | **3.71** |

Table 7.3: Expected posterior peak over half-width, with all systematic uncertainties taken into account, for six analysis channels and for all the three modes. The values from all channels combined are shown in bold type.

### 7.2.3 Ensembles - Linearity Test

A linearity test is performed, to ensure that there is no bias in the cross section measurement. A set of ensembles of pseudo-datasets are generated from the total pool of background model events using different values for the signal cross sections. All the systematic uncertainties are taken into account during the ensemble generation for each pseudo-dataset individually. The probability to pick any given event is proportional to the event weight modified by the unique systematic shift for the pseudo-dataset in question [159].

Five ensembles have been generated, each with at least 9,800 entries (pseudo-experiments) for all the three single top modes, and each with a different amount of single top. For $s+t$-mode, ensembles are generated with cross sections: 2 pb, 3.30 pb (SM), 5 pb, 7 pb and 10 pb with SM $s$ : $t$ ratio; for $s$-mode cross sections used are: 0.5 pb, 1.04 pb (SM), 1.5 pb, 2.5 pb and 3.5 pb; and for $t$-mode cross section values used are: 1.5 pb, 2.26 pb (SM), 3.5 pb, 5 pb and 7 pb. In each such ensemble, the cross section is measured treating each pseudo-dataset as real data. The measured cross section distribution in each ensemble for $s+t$, $s$ and $t$-modes can be seen in Figures 7.18, 7.19 and 7.20, respectively. A Gaussian fit around the mean of the distribution is performed and the resulting Gaussian mean (and its error) used as the measured cross section in the ensemble and a straight-line fit is done to the measured single top cross section as a function of the input value of the single top cross section used in the ensemble generation.

The bias in the cross section measurement is determined from a straight-line fit to the mean of the distributions vs. the input cross sections used to generate each ensemble type as shown in the bottom right plots of Figures 7.18, 7.19 and 7.20 for $s+t$, $s$ and $t$-modes, respectively. A linear response is obtained for all the three single top modes with slope of the line is about unity with an offset close to zero. Thus, it can be concluded that the bias in our cross section measurement is negligible and can be ignored henceforth.

LINEARITY TEST FOR $s+t$ CHANNEL

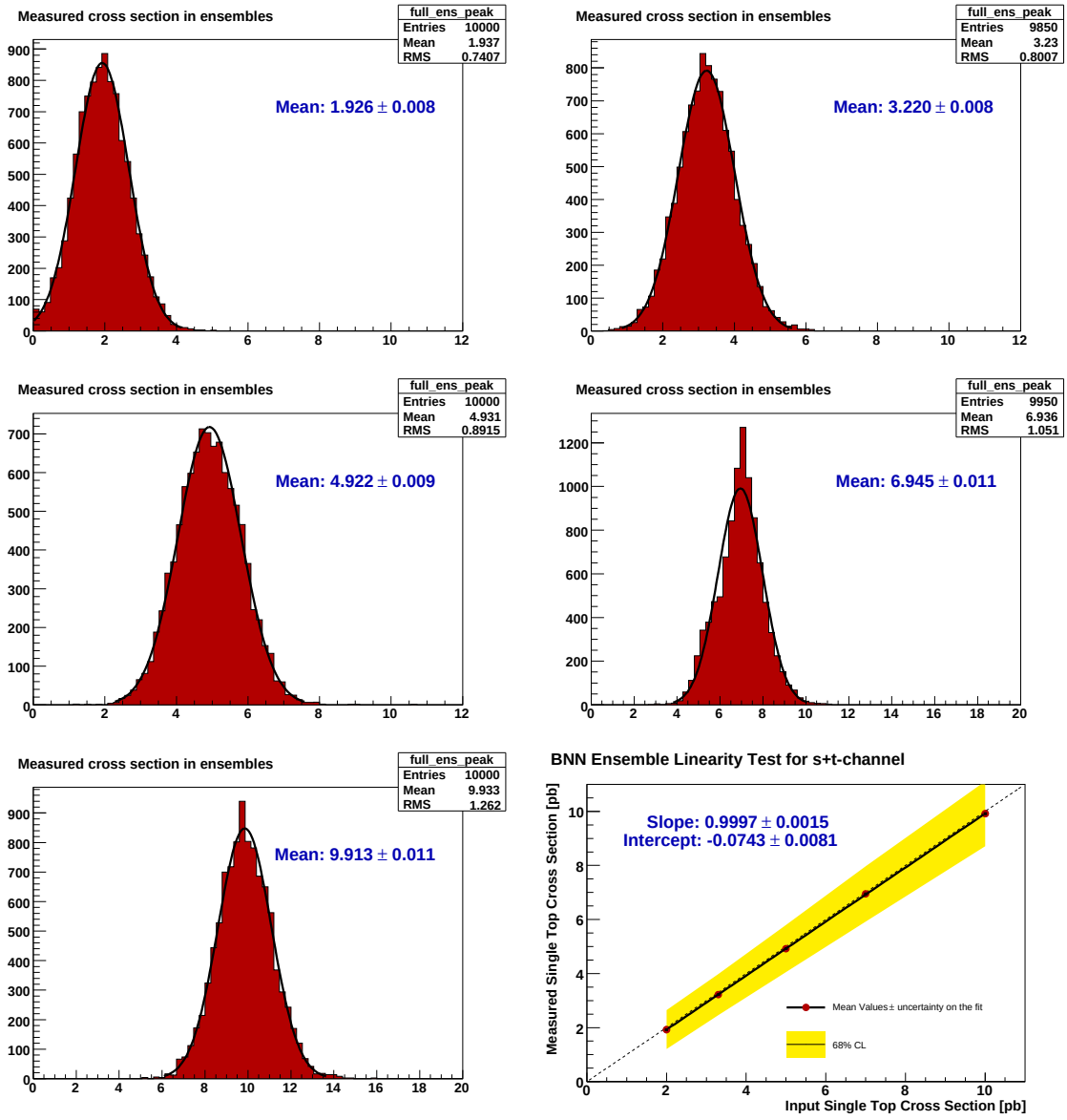


Figure 7.18: Measured single top cross sections in ensembles with various amounts of single top for $s+t$-channel. The upper row shows the results from the 2 pb and the SM (3.30 pb) ensembles, the second row shows the results from the 5 pb and the 7 pb ensembles, the third row shows the results from the 10 pb ensembles and Linear fit through the means from the Gaussian fits.

LINEARITY TEST FOR $s$ CHANNEL

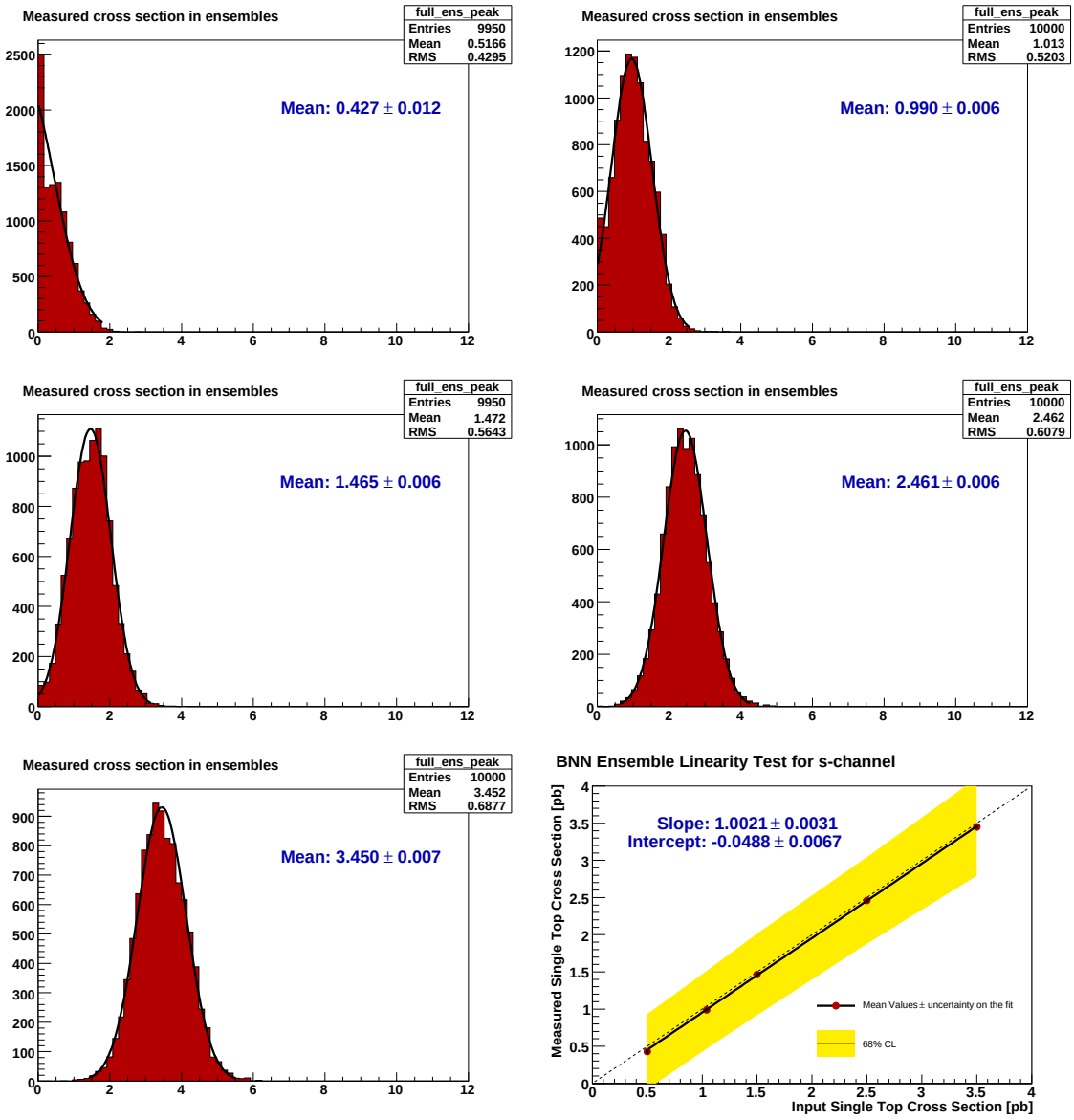


Figure 7.19: Measured single top cross sections in ensembles with various amounts of single top for $s$-channel. The upper row shows the results from the 0.5 pb and the SM (1.04 pb) ensembles, the second row shows the results from the 1.5 pb and the 2.5 pb ensembles and the third row shows the results from the 3.5 pb ensembles and Linear fit through the means from the Gaussian fits.

LINEARITY TEST FOR $t$ CHANNEL

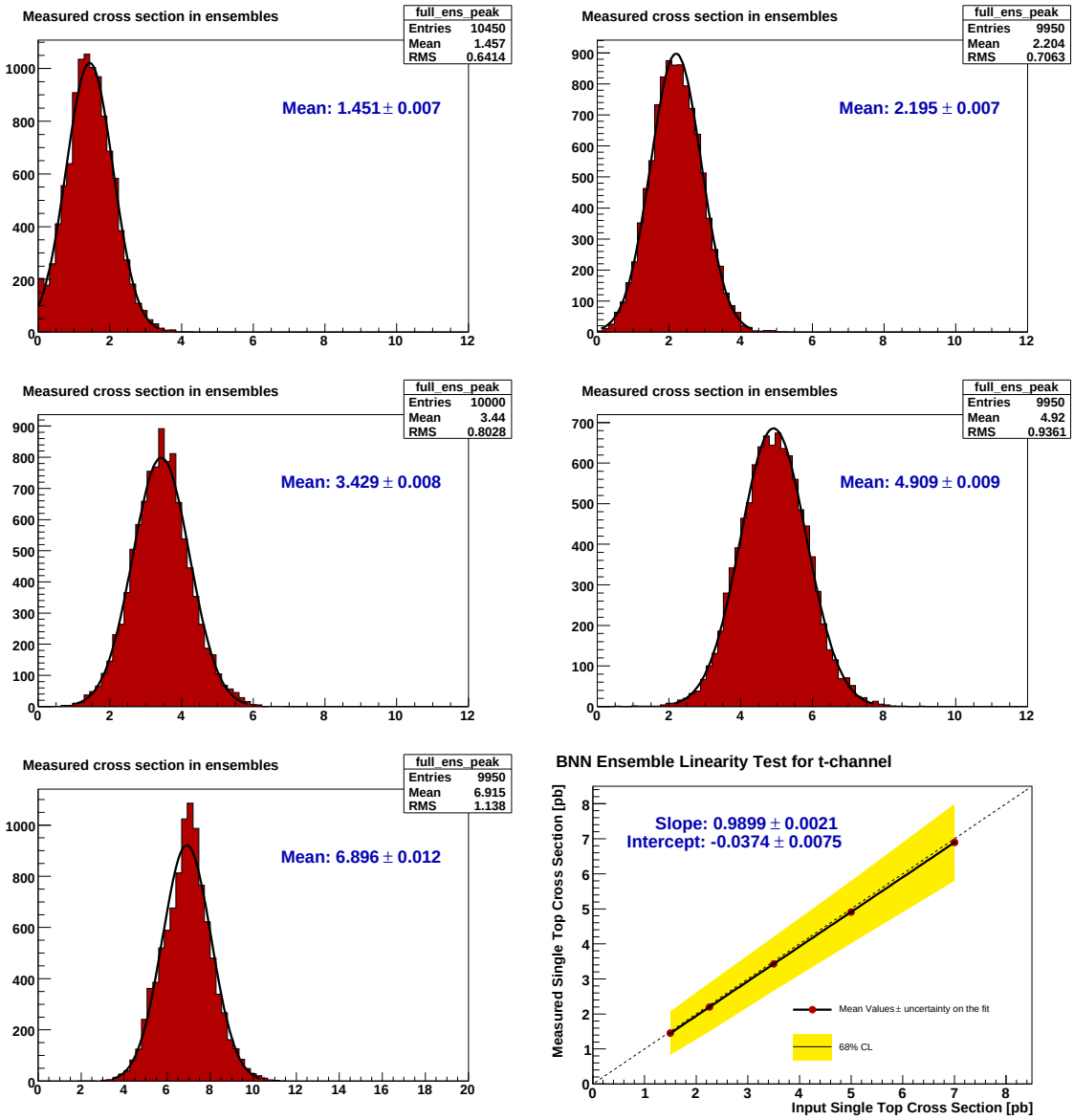


Figure 7.20: Measured single top cross sections in ensembles with various amounts of single top for $t$-channel. The upper row shows the results from the 1.5 pb and the SM (2.26 pb) ensembles, the second row shows the results from the 3.5 pb and the 5 pb ensembles and the third row shows the results from the 7 pb ensembles and Linear fit through the means from the Gaussian fits.

### 7.2.4 Observed Cross Section

Finally, after all the cross checks and ruling out for the possibility of any significant bias in the analysis, the measurement of the cross section has been performed using the actual data set. Figure 7.21 shows the final BNN output discriminant for all the six channels combined in log scale across the entire range of the output values for $s+t$, $s$ and $t$-modes, respectively. The BNN output for all the channels and for all the modes are shown in Figures 7.8 to 7.10.

Figures 7.22 to 7.24 show the Bayesian posterior probability density distribution for observed results, for the individual six analysis channels and for all the six channels combined, and including all the systematics. Table 7.4 and 7.5 summarizes the observed cross sections and peak over half-width with all the systematic uncertainties included for the six analysis and for all channels combined.

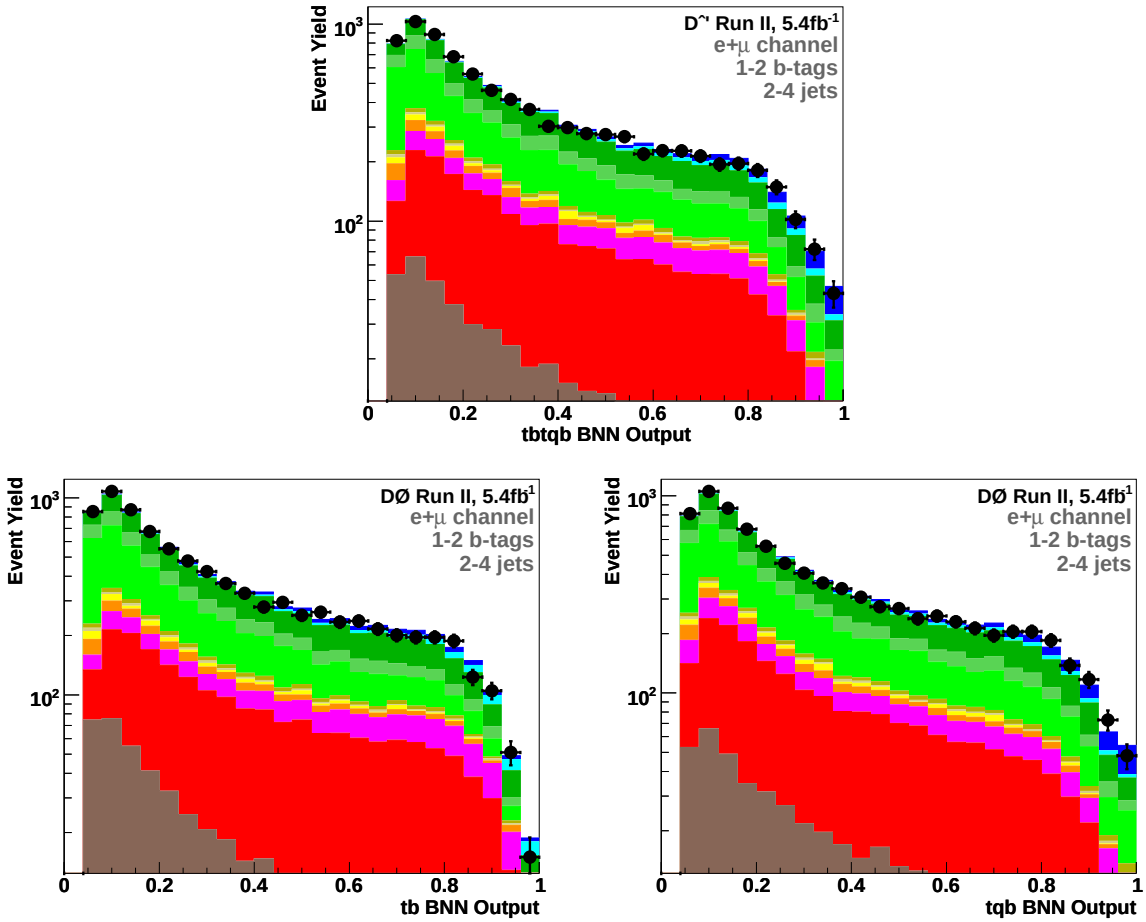


Figure 7.21: Bayesian neural network discriminant output for all the six channels combined for $s+t$, $s$ and $t$-modes in log scale. The single top contribution in this plot is normalized to the measured cross section.

Observed Cross Section FOR $s+t$ CHANNEL

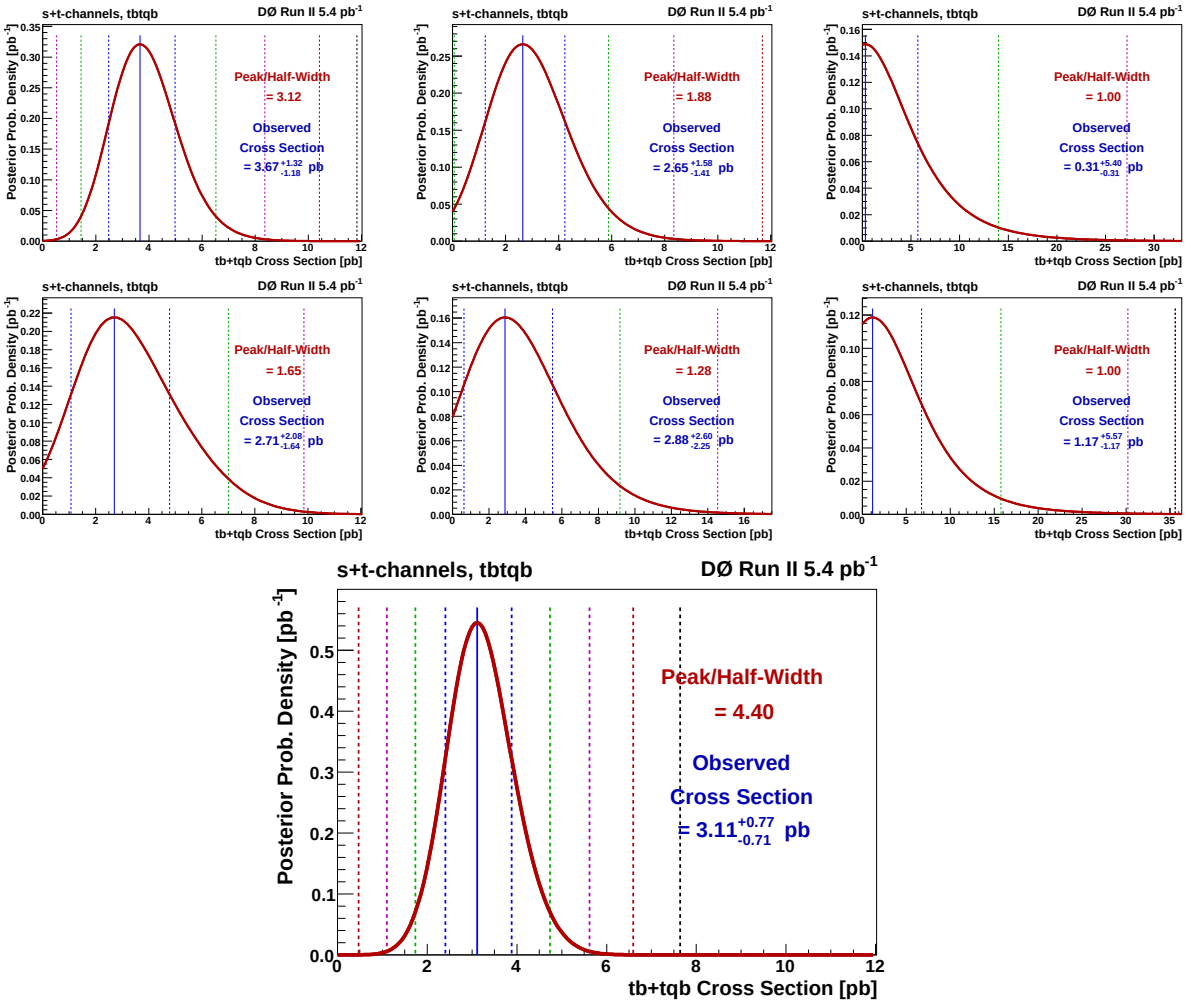


Figure 7.22: Observed posterior density for $s+t$-channel. First two rows show the results for the six individual analysis channels: 2 jets (left), 3 jets (middle), 4 jets (right); single tag (top row), double tag (second row). The bottom plot shows the observed posterior density for all the six analysis channels combined.

Observed Cross Section FOR $s$ CHANNEL

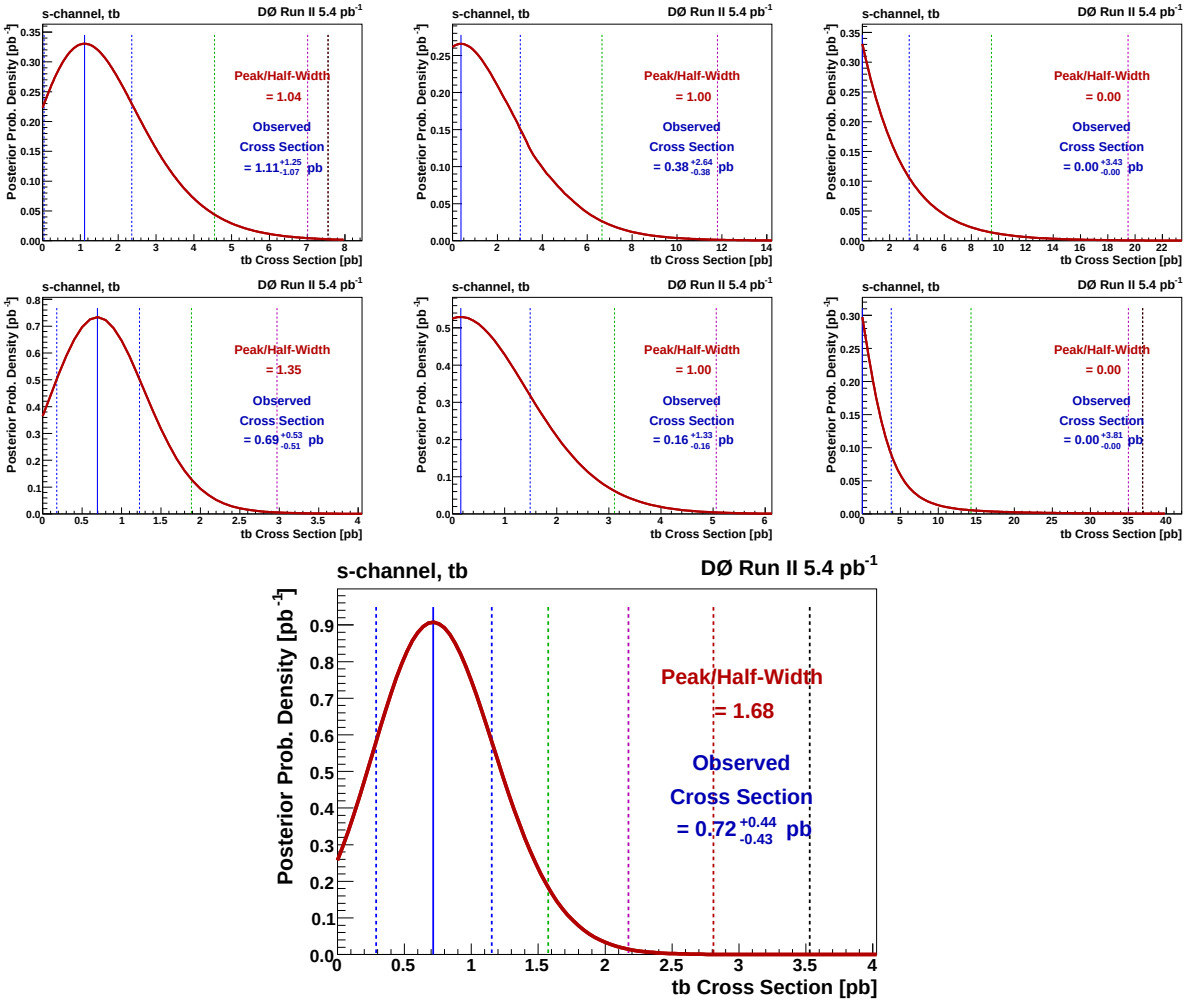


Figure 7.23: Observed posterior density for $s$-channel. First two rows show the results for the six individual analysis channels: 2 jets (left), 3 jets (middle), 4 jets (right); single tag (top row), double tag (second row). The bottom plot shows the observed posterior density for all the six analysis channels combined.

Observed Cross Section FOR $t$ CHANNEL

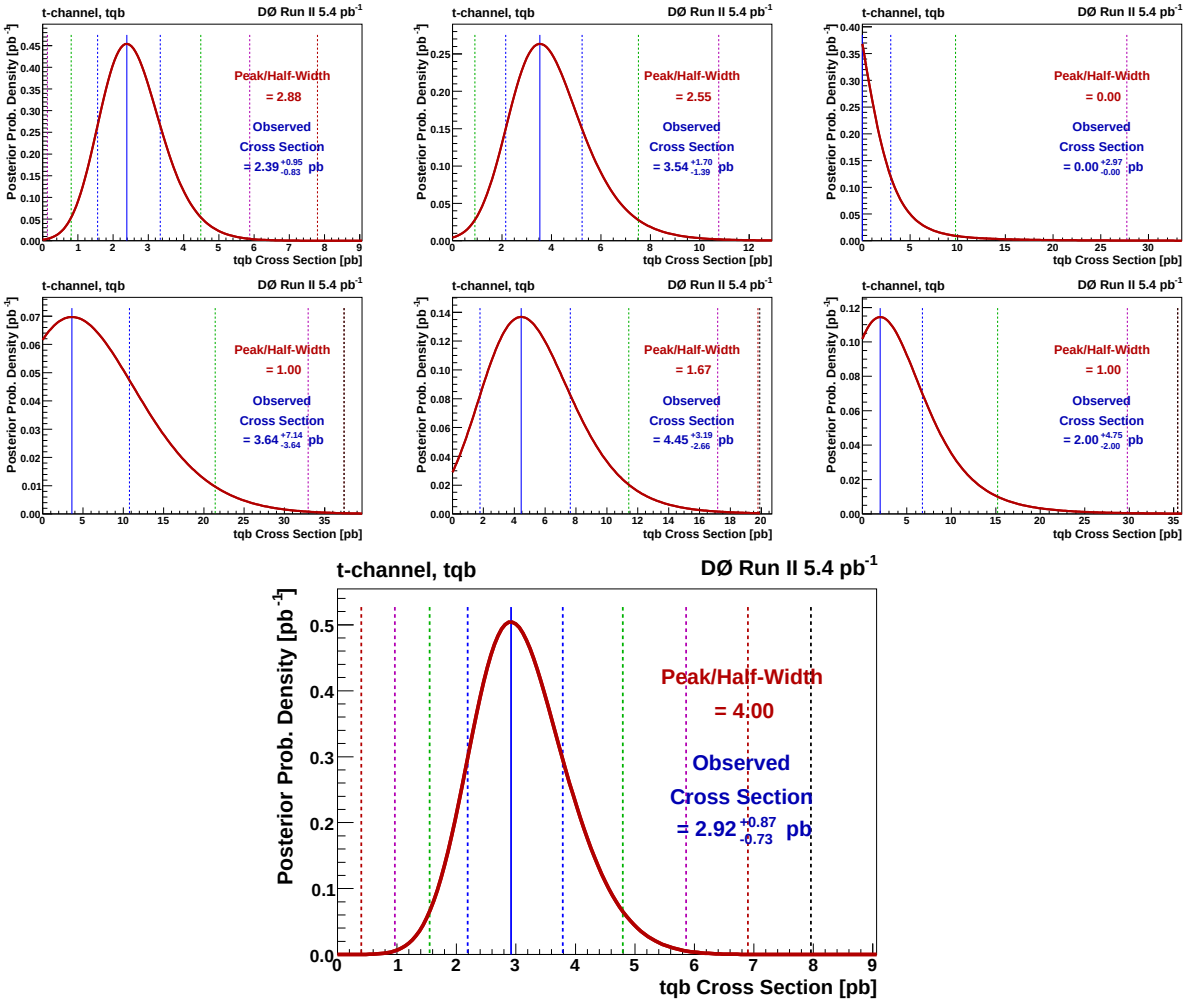


Figure 7.24: Observed posterior density for $t$-channel. First two rows show the results for the six individual analysis channels: 2 jets (left), 3 jets (middle), 4 jets (right); single tag (top row), double tag (second row). The bottom plot shows the observed posterior density for all the six analysis channels combined.

| Observed Cross Sections | | | |
|---|---|---|---|
| Channel | $(tb+tqb)$ $s+t$-mode | $(tb)$ $s$-mode | $(tqb)$ $t$-mode |
| 1tag / 2jets | $3.67^{+1.32}_{-1.18}$ | $1.11^{+1.25}_{-1.07}$ | $2.39^{+0.95}_{-0.83}$ |
| 1tag / 3jets | $2.65^{+1.58}_{-1.41}$ | $0.38^{+2.68}_{-0.38}$ | $3.54^{+1.70}_{-1.39}$ |
| 1tag / 4jets | $0.31^{+5.40}_{-0.31}$ | $0.00^{+3.43}_{-0.00}$ | $0.00^{+2.97}_{-0.00}$ |
| 2tag / 2jets | $2.71^{+2.08}_{-1.64}$ | $0.69^{+0.53}_{-0.51}$ | $3.64^{+7.14}_{-3.64}$ |
| 2tag / 3jets | $2.88^{+2.60}_{-2.25}$ | $0.16^{+1.33}_{-0.16}$ | $4.45^{+3.19}_{-2.66}$ |
| 2tag / 4jets | $1.17^{+5.57}_{-1.17}$ | $0.00^{+3.81}_{-0.00}$ | $2.00^{+4.75}_{-2.00}$ |
| 1-2tags / 2-4jets | $\mathbf{3.11^{+0.77}_{-0.71}}$ | $\mathbf{0.72^{+0.44}_{-0.43}}$ | $\mathbf{2.92^{+0.87}_{-0.73}}$ |

Table 7.4: Observed cross section, with all systematic uncertainties taken into account, for six analysis channels and for all the three modes. The values from all channels combined are shown in bold type.

| Observed Posterior Peak Over Half-Width | | | |
|---|---|---|---|
| Channel | $(tb+tqb)$ $s+t$-mode | $(tb)$ $s$-mode | $(tqb)$ $t$-mode |
| 1tag / 2jets | 3.12 | 1.04 | 2.88 |
| 1tag / 3jets | 1.88 | 1.00 | 2.55 |
| 1tag / 4jets | 1.00 | 0.00 | 0.00 |
| 2tag / 2jets | 1.65 | 1.35 | 1.00 |
| 2tag / 3jets | 1.28 | 1.00 | 1.67 |
| 2tag / 4jets | 1.00 | 0.00 | 1.00 |
| 1-2tags / 2-4jets | **4.40** | **1.68** | **4.00** |

Table 7.5: Observed posterior peak over half-width, with all systematic uncertainties taken into account, for six analysis channels and for all the three modes. The values from all channels combined are shown in bold type.

### 7.2.5 Significance Measurement

The conventional method of measuring significance of the excess over background using a very large ensembles of pseudo-datasets generated with background only events as used in the single top observation analysis [67], requires a large amount of computing time. Since now the significance of the signal is even bigger than the previous analysis, a new technique called asymptotic approximation of the log-likelihood ratio (AALLR) [68, 161, 162, 163] is used to calculate the significance of the measurement. After checking that the two approaches are consistent [155], using AALLR method seems more beneficial and time saving.

In case of the conventional method of ensembles, the significance is measured as one minus the probability of making the measurement assuming null hypothesis is true. In this case, null hypothesis is that there are no actual single top events in the dataset. From the measured cross section in the ensemble, the probability that data containing no single top quark events could fluctuate up to or above the measured cross section in the real data set, is calculated dividing the number of pseudo-datasets above the measured cross section by the total number of pseudo-datasets. This probability is known as the"p-value" ($\alpha$). The significance of the measurement is then estimated as the quantile function (the inverse of the normal cumulative distribution function) at probability 1 - $\alpha$; so, it is measured in units of the standard deviation ($\sigma$) of a Gaussian distribution:

$$\sigma = \sqrt{2}\ erf^{-1}(1 - 2\alpha) \tag{7.19}$$

In case of AALLR method, the log-likelihood ratio is defined as:

$$Q = -2 \ln \frac{L(\mu = 1)}{L(\mu = 0)} \tag{7.20}$$

where $L(\mu)$ is the likelihood after integrating over all the systematics for a cross section $\sigma = \mu\sigma_0$ where $\mu$ is the strength parameter and $\sigma_0$ is the theoretical cross section for signal. This ratio tests the compatibility of the data with two hypotheses: a null hypothesis where

there is only background ($\mu = 0$) and an alternative hypothesis with background plus signal, where the number of signal events corresponds to the theoretical cross section ($\mu = 1$).

If the LLR is evaluated in an ensemble of events $A$ with a signal strength $\mu'$, the LLR will follow a Gaussian distribution with mean $\alpha^2(1 - 2\mu')$ and variance $4\alpha^2$ [68], where $\alpha$ is given as:

$$\alpha = \frac{\sqrt{-2 \ln \frac{L(\mu)}{L(\hat{\mu})}}}{|\mu - \mu'|} \tag{7.21}$$

where $\hat{\mu}$ is the maximum likelihood value for the signal strength. Under the asymptotic assumption the value $\alpha$ is independent of the particular choice of strength value $\mu$ in Eq. 7.21. In case of $\mu^{'} = 0$ i.e, ensemble with only background events (B) and $\mu = \mu_{expected}$, then the asymptotic probability density function $p_B(Q)$ is a Gaussian with mean $\alpha_B^2$ and variance $4\alpha_B^2$. And in case of an ensemble with signal plus background events i.e, $\mu^{'} = 1$, the asymptotic probability density function $p_{S+B}(Q)$ is a Gaussian with mean $-\alpha_{S+B}^2$ and variance $4\alpha_{S+B}^2$.

The p-value is defined as the probability that the LLR value is smaller or equal to the observed result $Q_{\text{obs}}$ in the case of background fluctuations (no signal):

$$p = \int_{-\infty}^{Q_{\text{obs}}} \rho_B(q)dq = \Phi\left(\frac{Q_{\text{obs}} - \alpha_B^2}{2\alpha_B}\right) \tag{7.22}$$

where $\Phi$ is the cumulative distribution function for a Gaussian distribution centered in 0 with a variance equal to 1. To calculate the expected "p-value", we set $Q_{\text{obs}}$ equal to the mean of the distribution $\rho_{S+B}$, where in the asymptotic approach this is equivalent to $Q_{\text{obs}} = -\alpha_{S+B}^2$.

Using the above described AALLR method, expected and observed significance is calculated for $s+t$, $s$ and $t$-channels as shown in Figure 7.25 with the parameters obtained given in Table 7.6. The green curve shows the probability that the background only hypothesis will yield a given log-likelihood ratio. The blue curve shows the probability that the background plus SM signal will yield a given log-likelihood ratio. Both curves are normalized to unit area. The "p-value" at a specific log-likelihood point is then calculated by integrating the

area under the green (background only) curve to the left of that point on the x-axis. In order to calculate a significance, one point on this axis must then be chosen. For the expected significance, the point chosen is the mode of the blue (signal plus background) curve which is represented by the dashed line. The grey band corresponds to the one sigma uncertainty on this value. In order to calculate the observed significance, the log-likelihood ratio value of the observed result is calculated. This point is shown by the large black arrow.

| Channels | $\alpha_B$ | $\alpha_{S+B}$ | $Q_{\text{obs}}$ | $\rho_{exp}$ | $\rho_{obs}$ |
|---|---|---|---|---|---|
| $s+t$ | 4.10 | 4.87 | $-22.25$ | $4.9\sigma$ | $4.7\sigma$ |
| $s$ | 1.92 | 2.19 | $-0.20$ | $2.2\sigma$ | $0.9\sigma$ |
| $t$ | 3.07 | 4.21 | $-19.48$ | $4.4\sigma$ | $4.7\sigma$ |

Table 7.6: Parameters of the likelihood which are used in the calculation of the significances using the AAALR approach.

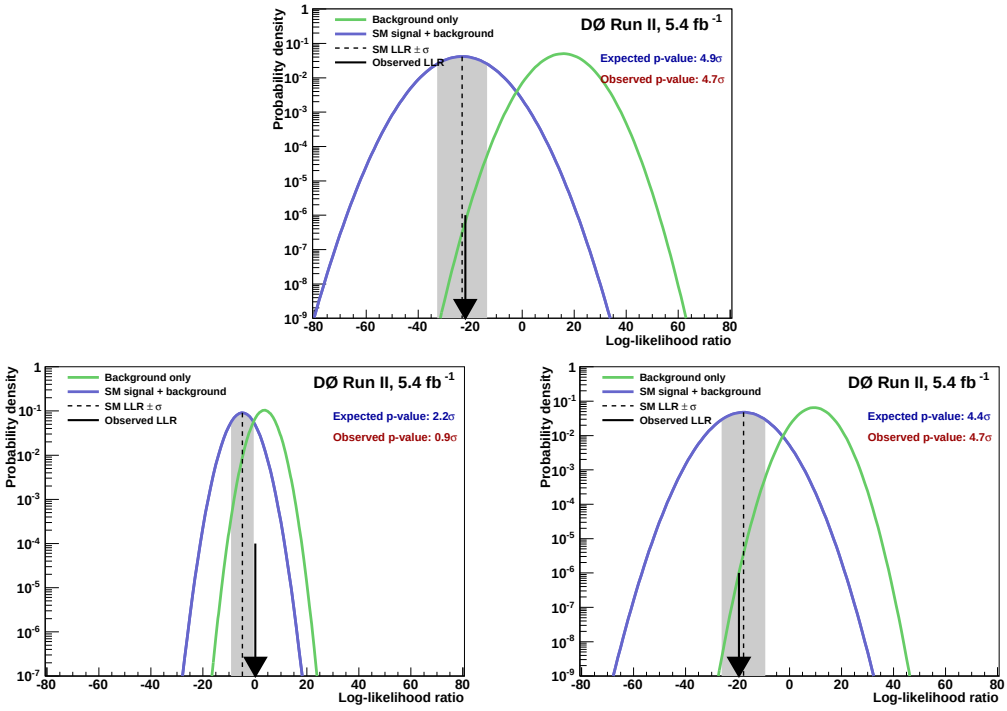


Figure 7.25: Significance of the BNN $s+t$, $s$, $t$-channels cross section measurement.The green curve is the probability density of the background only hypothesis, and the blue curve is the probability density of the standard model hypothesis. Both curves are normalized to unit area. The large arrow represents the observed value.

### 7.2.6 Combination Analysis

Two other multivariate analysis methods: Boosted Decision Trees (BDT) [164] and Neuroevolution of Augmenting Topologies (NEAT) [165] are also performed in parallel, in addition to the BNN analysis. These analyses use the same dataset, background modeling, event selection and consider the same sources of systematic uncertainties as described in Chapter 5. All the three methods differ in the choice of input variables used for training their discriminants in order to have less correlation between methods [155, 166, 167]. These two methods also have three training sets corresponding to $s+t$, $s$ and $t$-modes as described in Section 7.1.3 and hence at the end we have three independent discriminants for each of the three modes ($s+t$, $s$ and $t$-modes) for all six analysis channels combined coming from BNN, BDT and NEAT. The single top cross section and the significance are measured individually for each analysis with the same techniques described for BNN analysis. These measurements are highly correlated as the same dataset is used. But since the machinery of all three methods is different and due to the different choice of input variables, the methods are not entirely correlated with each other, additional performance can be gained by combining the information from all three methods.

In order to check the correlation present between the three MVA methods, there are two methods. In the first approach, the transformed output for each of the three methods (BNN output, DT output, NEAT output) for a given event can be plotted against each other. Figure 7.26 shows three scatter plots formed by taking all combination of pairs between the three values. As all three methods use the binning transformation as described in Section 7.1.4, they all have very similar shapes for the background distributions. This allows the direct comparison of the outputs, as the distributions are background dominated.

The second method of comparing the correlation between methods is to create many ensembles of MC events. Sampling events in such a way as to include correlations between systematic uncertainties and correlations between methods [159]. 5000 cross section measurements with full systematic uncertainties, are done with each of the three methods. Scatter

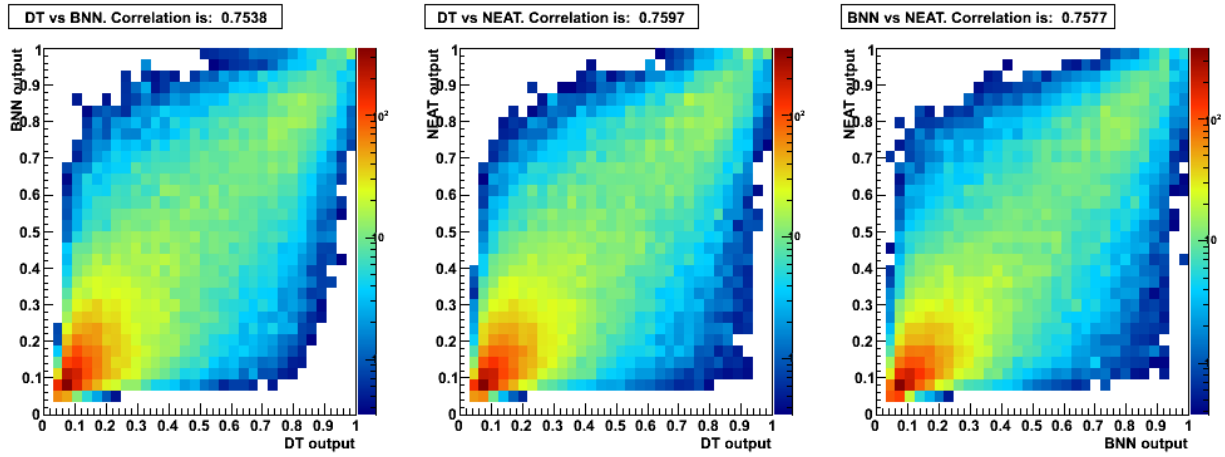


Figure 7.26: Correlations between MVA outputs

plots showing the correlation between cross section measurements are shown in Figure 7.27

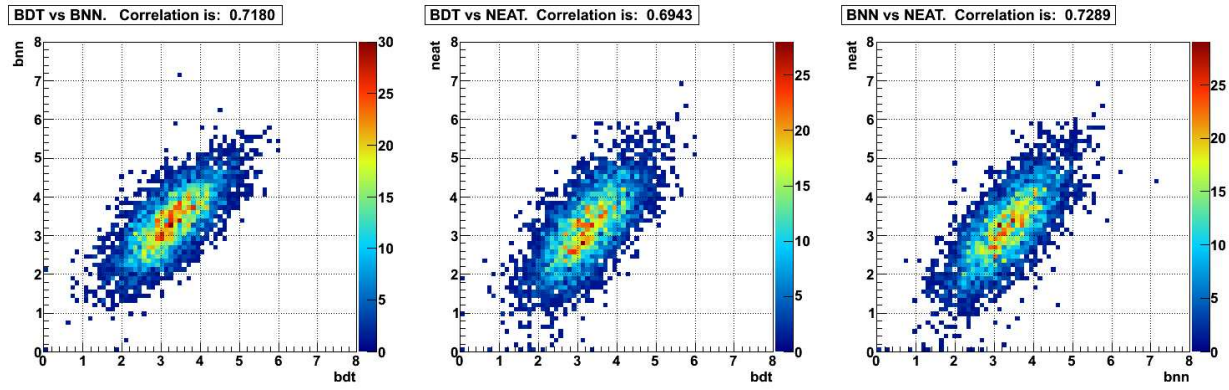


Figure 7.27: Cross section correlations between methods

As seen from both the methods, all the three MVA methods are correlated with each other to some extent but still not 100% correlated and hence there is possibility of increasing the sensitivity of measurement by combining all the information from three MVA methods. The combination of three methods is done using another Bayesian neural network, which takes as inputs the outputs of the three individual methods, and provides its own discriminant output [155]. This combination BNN is trained on a independent subset of the dataset denoted as "testing sample" as described in Section 7.1.1 which is different from both the samples used to train the individual methods and the MC used to make the cross section measurement. Three different combination BNN discriminants are constructed for $s+t$, $s$

and $t$-channels. The combination discriminants (BNNComb) for $s$ ($tb$) and $t$ ($tqb$)-channels take as inputs the three discriminant outputs of BNN, BDT and NEAT, and these are trained by assuming $tb$ or $tqb$ as signals, respectively. The combined $s+t$-channel ($tbtqb$) discriminant takes as input the six discriminant outputs of BNN, BDT, and NEAT that are trained separately for the $tb$ and $tqb$ signal. The training for $s+t$-channel discriminant treats the combined $tb+tqb$ contribution as signal with relative production rates predicted by SM. The combination BNN discriminant outputs for all six channels combined for all the three cases is shown in Figure 7.28.

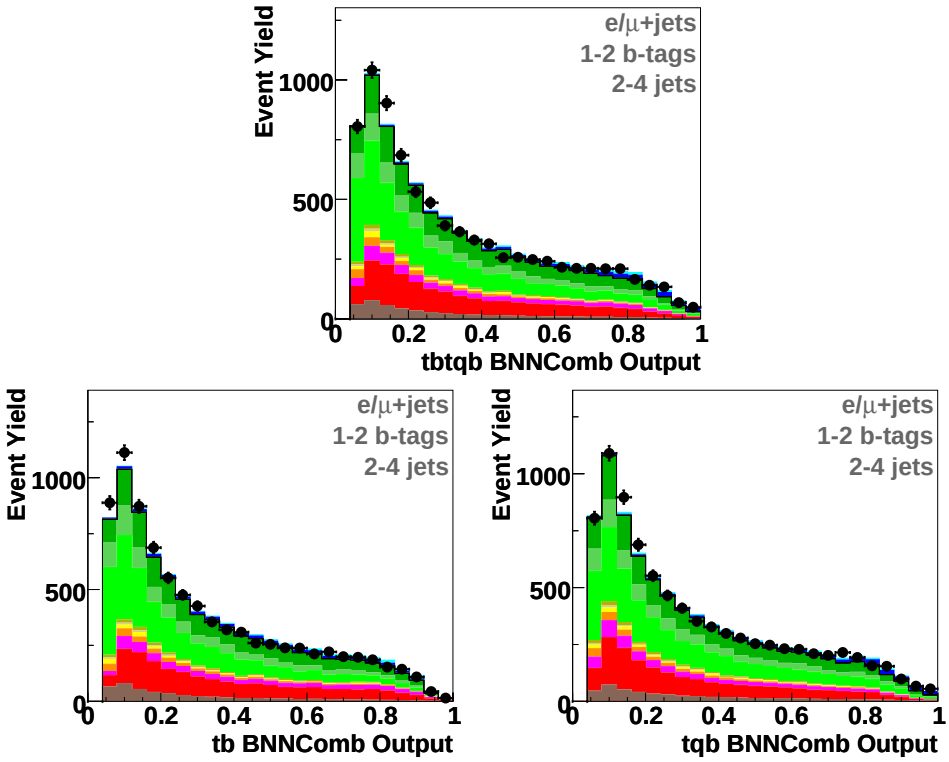


Figure 7.28: Combination BNN output for (First row) $s+t$-channel and (second row) $s$ and $t$-channels, for all the six analysis channels combined.

To the output of the combination, same Bayesian statistics approach is used to measure the cross section for $s+t$, $s$ and $t$-channels combination BNNs. Table 7.7 lists the expected and observed cross section values obtained for different combination BNNs and Figure 7.29

shows the resulting expected and observed posterior probability density curves [166].

| MVA Method | Expected Cross Section (pb) | Observed Cross Section (pb) |
|---|---|---|
| $s+t$-channel BNN combination | $3.49^{+0.77}_{-0.71}$ | $3.43^{+0.73}_{-0.74}$ |
| $s$-channel BNN combination | $1.12^{+0.45}_{-0.43}$ | $0.68^{+0.38}_{-0.35}$ |
| $t$-channel BNN combination | $2.43^{+0.67}_{-0.61}$ | $2.86^{+0.69}_{-0.63}$ |

Table 7.7: Expected and observed cross sections in $pb$ for $s+t$, $s$, $t$-channels BNN combination for all the six analysis channels combined.

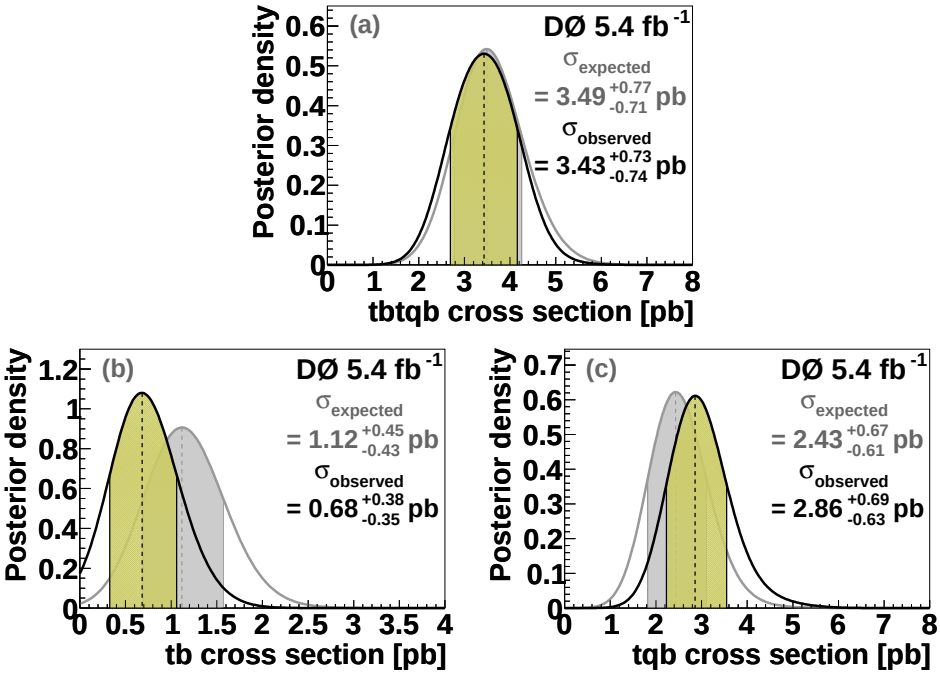


Figure 7.29: The expected (back) and observed (front) posterior probability density distributions for (a) $tbtqb$, (b) $tb$, and (c) $tqb$ production for all the six analysis channels combined. The shaded bands indicate the 68% C.L.s from the peak values.

Finally, the significance of the measurement for combination BNNs is calculated by the AALLR method explained in Section 7.2.5 as used for significance calculation for BNN analysis. Table 7.8 lists the expected and observed significance for BNN combination $s+t$, $s$

and $t$-channels and Figure 7.30 shows the corresponding plots for significance calculation. We have observed the $5.0\sigma$ significance for the $t$-channel and hence claiming the first observation of the single top $t$-channel.

| MVA Method | Expected Significance | Observed Significance |
|---|---|---|
| $s+t$-channel BNN combination | $5.5\sigma$ | $5.4\sigma$ |
| $s$-channel BNN combination | $2.7\sigma$ | $1.8\sigma$ |
| $t$-channel BNN combination | $4.3\sigma$ | $5.0\sigma$ |

Table 7.8: Expected and observed significance for $s+t$, $s$ and $t$-channels combination BNNs.

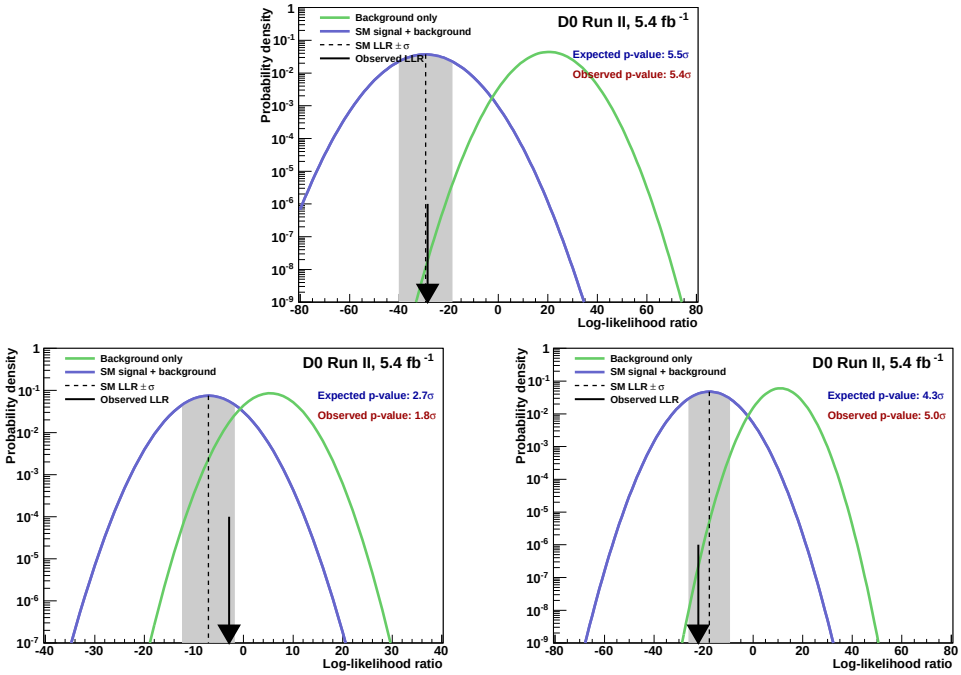


Figure 7.30: Significance of the combination BNN $s+t$, $s$, $t$-channels cross section measurement.The green curve is the probability density of the background only hypothesis, and the blue curve is the probability density of the standard model hypothesis. Both curves are normalized to unit area. The large arrow represents the observed value. The observed p-value is calculated by taking the integral of area under the green curve to the left of the arrow. The expected p-value is calculated by taking the integral of the area under the green curve to the left of the dashed line (at the mode of the blue curve).

## 7.3 Measurement of $|V_{tb}|$

As discussed in Section 2.3.1, the single top quark cross section is proportional to the square of the CKM matrix element, $V_{tb}$ [49] i.e, $\sigma \propto |V_{tb}|^2$. Single top quark production thus provides a sensitive and direct measurement of $|V_{tb}|$ without any assumption on the number of quark families or the unitarity of the CKM matrix [168, 155]. The assumptions made for this measurement are that the only existing production mechanism of electroweak produced top quarks are by a W boson; that $|V_{td}|^2 + |V_{ts}|^2 << |V_{tb}|^2$, implying that the branching fraction of $W \to tb \sim 100\%$ and that the electroweak produced top quark is dominated by this interaction; and that the $Wtb$ interaction is CP-conserving and of the $V-A$ type. The $Wtb$ interaction can be anomalous since $|V_{tb}f_1^L|$ can be greater than 1 where $f_1^L$ is the left-handed $Wtb$ coupling. The cross section for an electroweak produced top quark is thus directly proportional to the square of the effective $tbW$ coupling or in other words $V_{tb}$. Limiting the measurement to the range [0,1] implies the additional assumption that $f_1^L = 1$. The case of anomalous $Wtb$ couplings in single top production is studied in the next Chapter.

Bayesian posterior [158, 159] is formed using the BNN cross section analysis and a flat prior in $|V_{tb}f_1^L|^2$ is chosen. The prior is restricted to the region [0,1] in order to obtain the posterior for $|V_{tb}|^2$. In order to extract $|V_{tb}|$ from the measured cross section, additional theoretical uncertainties need to be considered and applied to the $s$-channel and $t$-channel samples separately in order to take the correlations into account properly. The resulting values for the uncertainties used for top mass 172.5 GeV are $\pm 3.8\%$ for s-channel $tb$, $\pm 5.3\%$ for t-channel $tqb$ [41].

The Bayesian posterior density for $|V_{tb}f_1^L|^2$ and the posterior for $|V_{tb}|^2$ obtained after restricting the prior to be non-zero only in the range [0,1] corresponding to setting $f_1^L = 1$ is shown in the Figure 7.31. The most probable value of $|V_{tb}f_1^L|^2$ is given by the peak of the posterior (a) of Figure 7.31 and is $|V_{tb}f_1^L|^2 = 1.03^{+0.22}_{-0.22}$ and hence $|V_{tb}f_1^L| = 1.02^{+0.10}_{-0.11}$. And from the posterior density (b) of Figure 7.31, $|V_{tb}|^2 = 0.99^{+0.1}_{-0.19}$ and hence $|V_{tb}| = 0.99^{+0.01}_{-0.10}$. From this posterior, a 95% confidence lower limit yields $|V_{tb}| > 0.79$ [166].

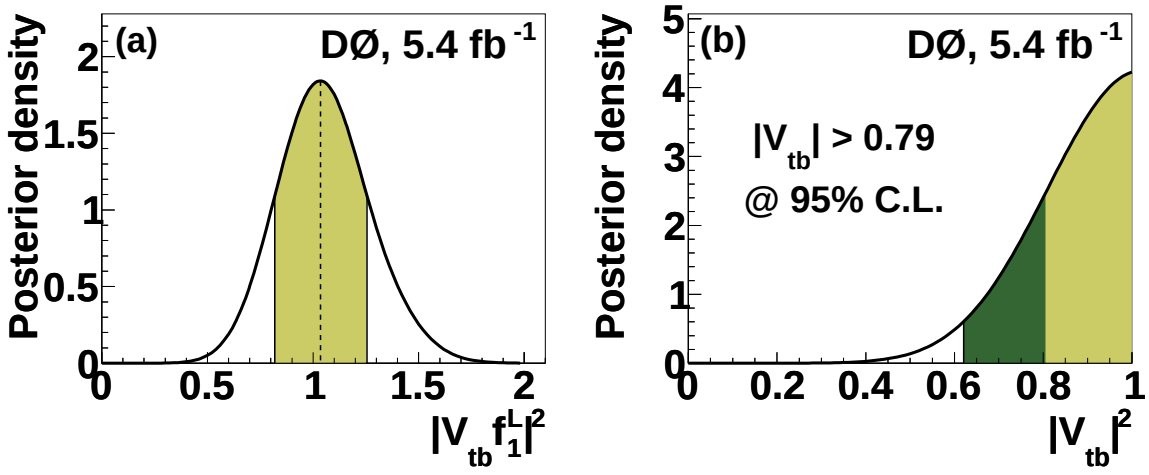


Figure 7.31: The posterior density functions for (a) $|V_{tb}f_1^L|^2$ and (b) $|V_{tb}|^2$. The shaded (dark shaded) band indicate regions of 68% (95%) C.L. relative to the peak values.

## 7.4 Model-Independent Measurement of $t$-channel cross section

A second approach to measure the $t$-channel cross section is used by following the same procedure which led to the evidence for the single top $t$-channel production [163, 162]. In this measurement, the same combination BNN discriminant as in Figure. 7.28 (upper right) is used which is specifically trained to isolate $t$-channel ($tqb$) by taking $s$-channel ($tb$) as an additional background contribution with its normalization set to the SM prediction. A two- dimensional (2D) posterior probability density is constructed as a function of the cross sections for the $tqb$ and $tb$ processes. The output discriminants for the signals, backgrounds, and data are used to form a binned likelihood as a product over all six analysis channels and all bins. No constraint is imposed on the relative rates of $tb$ and $tqb$ production. A Poisson distribution for the observed number of events and uniform prior probabilities with positive values for the two signal cross sections is assumed . Then integration is done over the systematic uncertainties which are described by Gaussian priors that preserve the correlations

between bins and channels. The $tqb$ cross section is then extracted from a one-dimensional posterior probability density obtained from this 2D posterior by integrating over the $tb$ axis, thus not making any assumptions about the value of the $s$-channel cross section. Similarly, the $tb$ cross section is obtained by integrating over the $tqb$ axis.

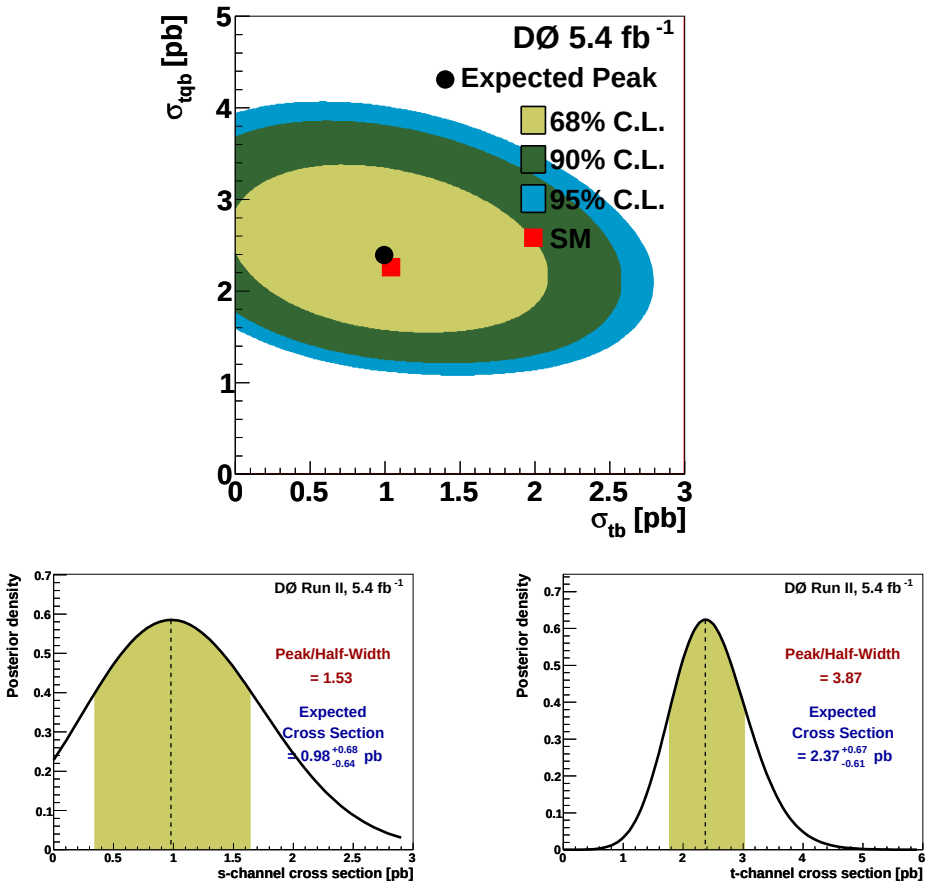


Figure 7.32: Expected 2d posterior density distribution for BNNcomb for t-channel calculated for the six analysis channels combined (2-4 jets and 1 or 2 tags), including all systematic uncertainties. The resulting one dimensional posterior densities for s-(t-)mode, obtained after integrating over the t-(s-)mode, are shown below.

Figure. 7.32 and 7.33 show the expected and observed 2D posterior density distributions and projected 1D posterior density when integrating over the $s$-channel axis. For completion, the 1D posterior density obtained when integrating over the $t$-channel axis is also included. Table 7.9 summarizes the expected and observed values for $t$-channel, and also includes the

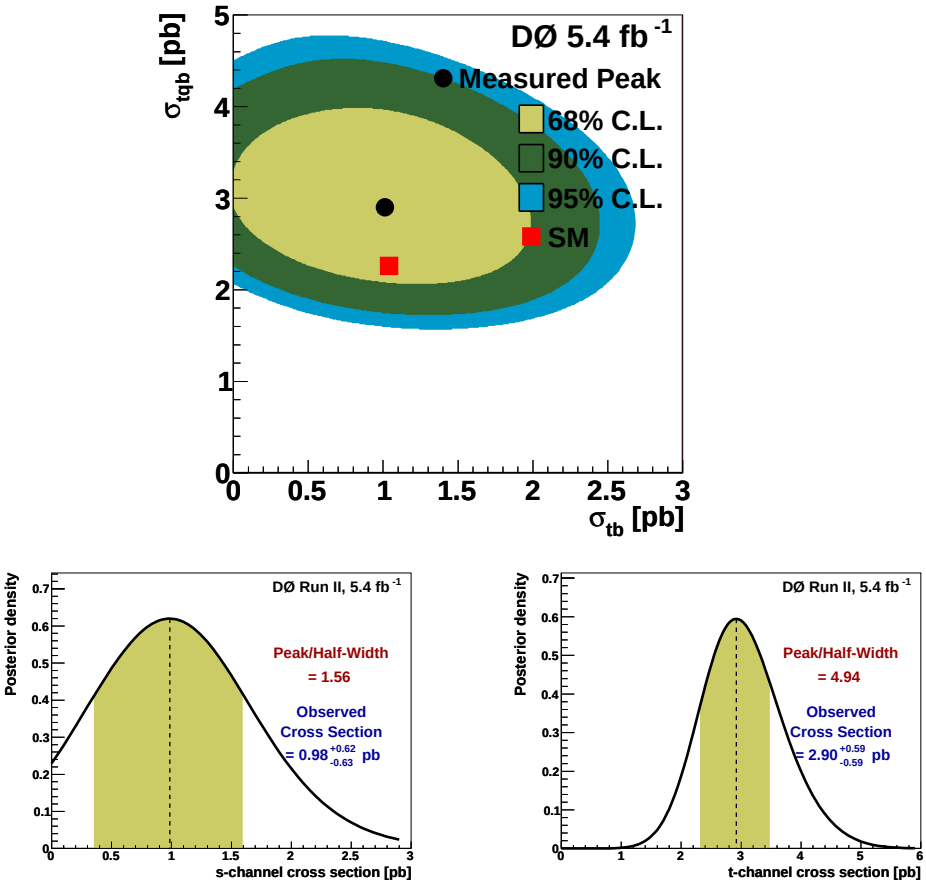


Figure 7.33: Observed 2d posterior density distribution for BNNcomb for t-channel calculated for the six analysis channels combined (2-4 jets and 1 or 2 tags), including all systematic uncertainties. The resulting one dimensional posterior densities for s-(t-)mode, obtained after integrating over the t-(s-)mode, are shown below.

$s$-channel values for completion.

Hence, using this method, the cross section of $t$-channel, with full systematics included, comes out to be equal to $2.90^{+0.59}_{-0.59}$ pb and the corresponding observed significance of this measurement with AALLR method comes out to be equal to $5.5\sigma$ [167], hence the first observation of the single top $t$-channel. The corresponding significance plot for this measurement is shown in Figure. 7.34.

| | s-channel | | t-channel | |
|---|---|---|---|---|
| Systematics | Expected | Observed | Expected | Observed |
| Stat | $1.03^{+0.48}_{-0.47}$ | $1.08^{+0.48}_{-0.47}$ | $2.25^{+0.51}_{-0.50}$ | $2.22^{+0.50}_{-0.48}$ |
| Full | $0.98^{+0.68}_{-0.64}$ | $0.98^{+0.62}_{-0.63}$ | $2.37^{+0.67}_{-0.61}$ | $2.90^{+0.59}_{-0.59}$ |

Table 7.9: Expected and observed cross section for t-channel obtained from the 2d posterior constructed from the BNNcomb for t-channel discriminant trained for t-channel.

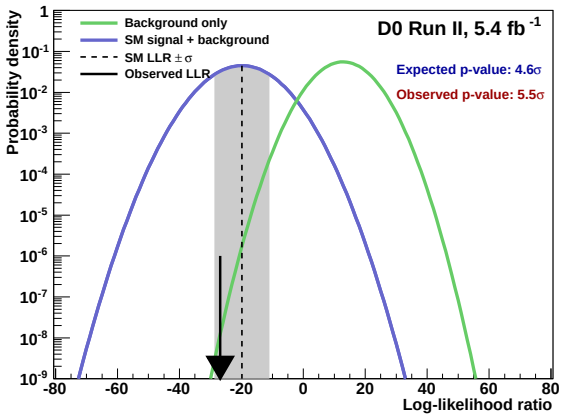


Figure 7.34: Significance of the $t$-channel cross section measured using 2D posterior approach and BNNComb discriminant trained with $t$-channel.

## 7.5 Results

This first analysis presents the measurement of single top quark production cross section using 5.4 $fb^{-1}$ of data collected in the DØ detector at Fermilab. The measurement is done for $s+t$, $s$ and $t$-channels separately. Bayesian neural networks are used to separate expected single top quark signals from background and the Bayesian neural network output distributions across the six independent channels used in the analysis are combined using a binned likelihood and the single top cross section in all the three channels is measured using a Bayesian method. For the top quark mass of 172.5 GeV, using bayesian neural network analysis, we measure the cross sections for $s+t$ ($tb+tqb$), $s$ ($tb$) and $t$ ($tqb$) production to be

$$\sigma(p\bar{p} \rightarrow tb+X, tqb+X) = 3.11^{+0.77}_{-0.71} \text{ pb}$$

$$\sigma(p\bar{p} \rightarrow tb+X) = 0.72^{+0.44}_{-0.43} \text{ pb}$$

$$\sigma(p\bar{p} \rightarrow tqb+X) = 2.92^{+0.87}_{-0.73} \text{ pb}$$

The $s+t$, $s$ and $t$ channels have a significance of $4.7\sigma$, $0.9\sigma$ and $4.7\sigma$, respectively. The corresponding theoretical predictions are $3.30 \pm 0.13$ pb, $1.04 \pm 0.04$ pb and $2.26 \pm 0.12$ pb.

After combining these results from Bayesian neural networks analysis with other two parallel multivariate analyses, the final combined measurement [166] for $s+t$, $s$ and $t$-channels is

$$\sigma(p\bar{p} \rightarrow tb+X, tqb+X) = 3.43^{+0.73}_{-0.74} \text{ pb}$$

$$\sigma(p\bar{p} \rightarrow tb+X) = 0.68^{+0.38}_{-0.35} \text{ pb}$$

$$\sigma(p\bar{p} \rightarrow tqb+X) = 2.86^{+0.69}_{-0.63} \text{ pb}$$

with a significance of $5.4\sigma$, $1.8\sigma$ and $5.0\sigma$ respectively for $s+t$, $s$ and $t$-channels. Hence the

first observation of single top $t$-channel. Measurement of $tb+tqb$ production assumes the SM ratio between $tb$ and $tqb$ production; and of $tb$ and $tqb$ measurement assumes respectively, $tqb$ and $tb$ production rates as predicted by the SM. Another measurement of $tqb$ production cross section is done using the same dataset and discriminant but without any assumption on the $tb$ production rate [167] which is described in Section 7.4. From this measurement, we obtain a cross section of $2.90^{+0.59}_{-0.59}$ pb for $t$-channel and corresponding significance of $5.5\sigma$. All the results are in good agreement with theoretical predictions and with the results from this analysis. Finally, a direct limit on the CKM matrix element $|V_{tb}| > 0.79$ at 95% assuming a flat prior with $0 \leq |V_{tb}|^2 \leq 1$ is derived and the results so obtained have been published in Ref. [166].

# Chapter 8

# Analysis 2

# Anomalous $Wtb$ Couplings

***You never know what is enough unless you know what is more than enough.*** - William Blake.

As seen from the previous chapter, the precise measurement of single top quark production cross section for $s+t$, $s$ and $t$-channel is done using 5.4 $fb^{-1}$ of data with multivariate techniques that can extract a small signal from an extremely large background. The single top quark study can be extended to search for the EWSB sector as described in the Section 2.3.1. The large mass of the top quark implies that it has large couplings to the electroweak symmetry breaking sector of the SM and may have non-standard interactions with the weak gauge bosons. Single top quark production provides a unique probe to study the interactions of the top quark with the $W$ boson. This chapter focuses on the theoretical background and methodology used in present study to search for the anomalous $Wtb$ couplings in single top quark production and see if some signature of beyond SM physics can be obtained.

## 8.1 Theory

As we have discussed in Chapter 2, the SM has been extraordinarily successful in describing the data taken at the energies of present colliders. This is true for both the strong and electroweak interactions. However, the SM does not explain everything and especially the mass generation sector of standard model gives rise to many unanswered questions, making a strong case for new physics beyond SM.

New physics can manifest itself either in terms of new particles or as new couplings changing the cross sections and angular distributions of existing processes. The couplings of weak gauge bosons to light quarks are very well known from precision measurements at LEP and low energy experiments [169], but this is not the case for the top quark, whose couplings to the weak gauge bosons are not well known. The large top quark mass, comparable to the EWSB scale, offers a possibility that the top quark plays a key role in the mechanism of EWSB or opens a window of sensitivity to new particles related to EWSB that strongly couple to the top quark. In such cases, modifications to top quark interactions, in particular with weak gauge bosons (as the mediators of the EW interaction which acquire mass via EWSB), could yield the first signs of new physics. The Tevatron can sensitively probe the top quark interaction to the $W$ boson via measurements of single top quark production and top quark decays in $t\bar{t}$ production, each yielding complementary information. For instance, in the case of single top quark production and assuming that the production mechanism involves only a $W$ boson exchange, the cross section is proportional to the square of the effective coupling between the top quark and the $W$ boson, and can be used to study the $Wtb$ coupling [34, 55, 170, 171].

Within the standard model with three generations of quarks, the charged current interactions of the top quark are (at lowest order) of the type $V$-$A$, and involve a $W$ boson and

a down-type quark $q$ ($q = d, s, b$):

$$\Gamma^{\mu}_{tqW} = -\frac{g}{\sqrt{2}} V_{tq} \bar{u}(p_q) \gamma^{\mu} P_L u(p_t), \tag{8.1}$$

where $|V_{tq}|$ is one of the element of the 3×3 unitary the Cabibbo-Kobayashi-Maskawa (CKM) matrix [172, 173], and $P_L = (1 - \gamma_5)/2$ is the left-handed projection operator.

In general, the $Wtb$ coupling could be more complicated than predicted by the standard model, involving anomalous couplings. The most general, lowest dimension $Wtb$ vertex is given by [174]:

$$\mathcal{L} = -\frac{g}{\sqrt{2}} \bar{b} \gamma^{\mu} (L_V P_L + R_V P_R) t W^-_{\mu} - \frac{g}{\sqrt{2}} \bar{b} \frac{i \sigma^{\mu\nu} q_{\nu}}{M_W} (L_T P_L + R_T P_R) t W^-_{\mu} + h.c., \tag{8.2}$$

where $q_{\nu}$ is the $W$ boson four-momentum and the $L_{V,T} = V_{tb}.f_{L_{V,T}}$ and $R_{V,T} = V_{tb}.f_{R_{V,T}}$ couplings can a-priori be CP-violating. The form factor $f_{L_V}(f_{L_T})$ represents the left-handed vector (tensor) coupling, $f_{R_V}(f_{R_T})$ represents the right-handed vector (tensor) coupling, and $V_{tb}$ is the Cabibbo-Kobayashi-Maskawa matrix element [49]. Within the standard model, CP is conserved in the $Wtb$ vertex and $L_V \equiv V_{tb} \simeq 1$ and $R_V = L_T = R_T = 0$, at tree level. Measurements of top quark decays in $t\bar{t}$ production, e.g. $W$ boson polarization, are insensitive to $|V_{tb}|$, but allow to constrain the Lorentz structure of the $Wtb$ vertex. The single top quark cross section is directly proportional to the square of the effective $Wtb$ coupling, and can be different from the standard model prediction either because $|V_{tb}| < 1$, because of anomalous couplings, or both.

The size of the right-handed vector coupling, $R_V$, is indirectly constrained by the measured branching fraction for the $b \to s\gamma$ process [175] and has an upper bound of $|R_V| \leq 0.04$ [176]. At low energies, the contribution from the left- and right-handed tensor coupling is suppressed by the term $q_{\nu}/M_W$ but this contribution can become significant at higher energies. The structure of the $Wtb$ vertex has also been studied directly by CDF and D0 ex-

periments at the Tevatron in top quark decays in the $t\bar{t}$ pair production mode by measuring the $W$ boson helicity [177, 178].

For single top quark production, the cross section is directly proportional to the $Wtb$ coupling but kinematics and angular distributions depend on the structure of the $Wtb$ vertex. The effect of anomalous couplings on angular distributions and event kinematics is shown in parton level plots in Figure 8.1, which shows $p_T$ distributions of lepton and neutrino from the $W$ boson (decaying from top quark), and the distribution of $cos\theta$, where $\theta$ is the angle between the direction of the lepton and the top quark direction, in the top quark reference frame.

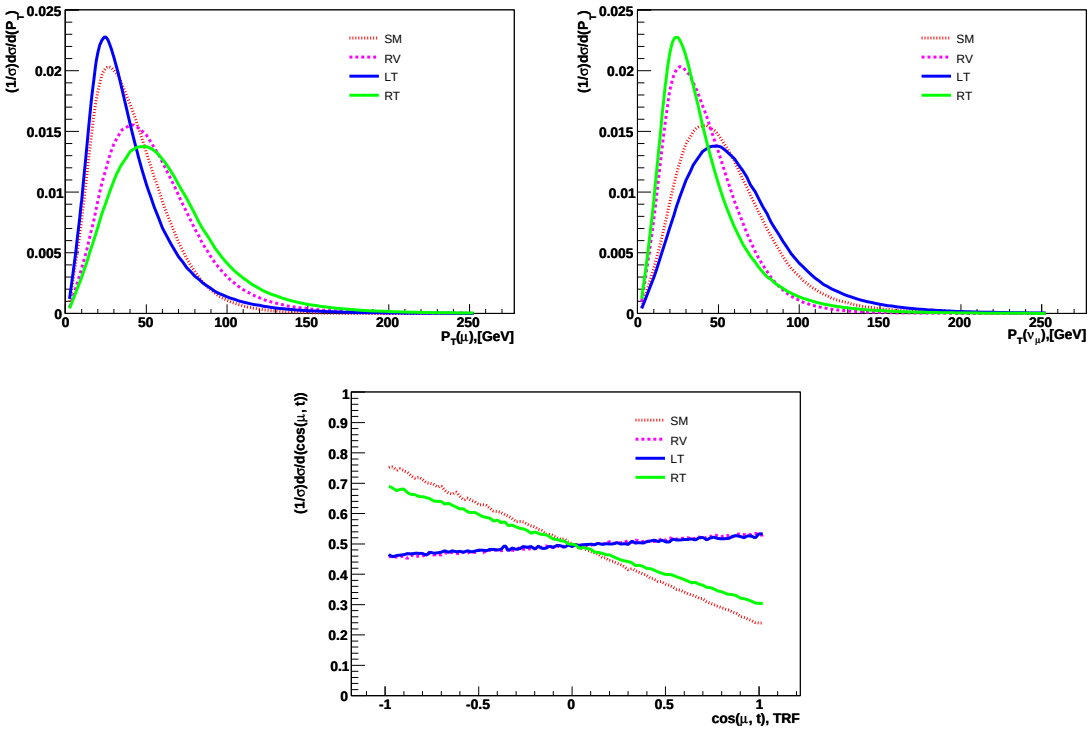


Figure 8.1: Lepton $p_T$ (top quark left) and neutrino $p_T$ (top right) distributions for all four couplings. The lower plot shows angular distribution $cos\theta$ where $\theta$ is the angle between the direction of the lepton and the top quark direction, in the top quark reference frame.

Hence the measurement of anomalous vector left-handed and right-handed coupling, and tensor left-handed and right-handed coupling, in the single top quark production can be done and limits can be set on four anomalous couplings [179, 180]. This measurement requires

some assumptions to be made. First assumption, is to ignore the flavor-changing neutral current interactions or heavy scalar or vector boson exchange i.e, the only existing production mechanism of single top quarks involves the interaction with a $W$ boson. Second is to assume that $|V_{ts}|$ and $|V_{td}|$ are negligible compared to $|V_{tb}|$ i.e, $|V_{td}|^2 + |V_{ts}|^2 \ll |V_{tb}|^2$. And the third and final assumption is that the $Wtb$ vertex is charge-parity ($CP$) conserving.

## 8.2 Signal and Background Modeling

In this analysis we used the same 5.4 $fb^{-1}$ dataset and triggers as described in Section 5.1 for the cross section measurement [166, 167]. The Lagrangian given in Eq. 8.2 has been incorporated into the CompHEP package and has been included into the SingleTop [117] generator for event generation [120] for a top quark mass $m_t = 172.5\ GeV$ using the CTEQ6M [113] parton distribution functions. The anomalous $Wtb$ couplings are taken into account in both production and decay in the generated samples. The event kinematics for both $s$-channel and $t$-channel processes reproduce distributions from next-to-leading-order calculations [48]. The decay of the top quark and the resulting $W$ boson are carried out in the singletop [117] generator in order to preserve information about the spin of the particles. PYTHIA [118] was used to add the underlying event and initial- and final-state radiation. TAUOLA [119] was used to decay tau leptons, and EvtGen [120] to decay $b$ hadrons. Five single top signal samples ($L_V$ (SM) sample generated with $f_{L_V} = 1$; $R_V$ sample generated with $f_{R_V} = 1$; $L_T$ sample generated with $f_{L_T} = 1$; $R_T$ sample generated with $f_{R_T} = 1$; and $L_V R_T$ sample generated with $f_{L_V}, f_{R_T} = 1$) are generated according to the coupling values lists in Table 8.1.

| Sample | $L_V$ | $L_T$ | $R_V$ | $R_T$ |
|---|---|---|---|---|
| 1 | 1 | 0 | 0 | 0 |
| 2 | 0 | 1 | 0 | 0 |
| 3 | 0 | 0 | 1 | 0 |
| 4 | 0 | 0 | 0 | 1 |
| 5 | 1 | 0 | 0 | 1 |

Table 8.1: Parameters used in the generation of the single top MC samples.

Out of the four couplings ($L_V$ , $L_T$ , $R_V$ , $R_T$ ), three coupling scenarios are considered: ($L_V, L_T$ ), ($L_V, R_V$ ) and ($L_V, R_T$ ), and for each case, assuming the other two non-SM couplings are negligible. In the first scenario i.e, ($L_V, L_T$ ) the signal is represented by a superposition of samples 1 and 2,

$$s = L_V^2 s_1 + L_T^2 s_2 \tag{8.3}$$

where $s_1$ and $s_2$ represents the signal sample 1 and 2 from Table 8.1, that means a sample with pure left-handed vector $s_1$ and tensor $s_2$ couplings. or in terms of process cross sections

$$s = \sigma_{L_V} a_1 + \sigma_{L_T} a_2 \tag{8.4}$$

where $a_i = s_i/\sigma_i$ and $\sigma_i$ are the effective acceptance and cross sections used in the normalization of the samples. For instance, $a_2$ and $\sigma_{L_T}$ are the effective acceptance and cross section resulting from left-handed tensor ($L_T$) contribution to single top production. Similarly the second scenario, ($L_V, R_V$ ), uses samples 1 and 3, resulting in the following superpositions

$$s = L_V^2 s_1 + R_V^2 s_3 \tag{8.5}$$

and

$$s = \sigma_{L_V} a_1 + \sigma_{R_V} a_3. \tag{8.6}$$

The third scenario, ($L_V, R_T$ ), is more complicated because of the interference of the two amplitudes. Here, samples 1, 4, and 5 are superimposed as follow

$$s = L_V^2 s_1 + R_T^2 s_4 + L_V R_T (s_5 - s_1 - s_4). \tag{8.7}$$

where third term is the interference contribution to the signal yield. In terms of process

cross sections we obtain

$$s = \sigma_{L_V} a_1 + \sigma_{R_T} a_4 + \left( \frac{\sigma_{L_V} \sigma_{R_T}}{\sigma_1 \sigma_4} \right)^{\frac{1}{2}} (\sigma_5 a_5 - \sigma_1 a_1 - \sigma_4 a_4). \tag{8.8}$$

The main background contributions are same as those from the cross section measurement analysis described in Section 5.2.1 i.e, $W$ bosons produced in association with jets ($W$+jets), $t\bar{t}$ production, and multijet production in which a jet with high electromagnetic content mimics an electron, or a muon contained within a jet originating from the decay of a heavy-flavor quark ($b$ or $c$ quark) appears to be isolated. Diboson ($WW, WZ, ZZ$) and $Z$+jets processes add small contributions to the background. Table 8.2 shows the cross sections, branching fractions, and initial number of events of the MC samples.

### 8.2.1 Event Yields and Signal Acceptances

Exactly same event selection criteria as described in Section 5.2.2 is applied to all the different coupling signals and background samples. Also, all the samples are modeled with all the corrections described in Section 5.2.4. After all the event selection and corrections, the resulting event yield for signal samples is listed in Table 8.3, where "yield" refers to the number of events of the signal predicted to be in the nearly 5.4 $fb^{-1}$ of data analyzed. The signal acceptance is defined as:

$$\mathcal{A} = \frac{\mathcal{B}}{\mathcal{N}_{\text{initial}}} \sum_{\mathcal{N}_{\text{selected}}} \varepsilon_{\text{corrections}}, \tag{8.9}$$

where $\mathcal{B}$ is the branching fraction and $\mathcal{N}_{\text{initial}}$ is the initial number of events, for each MC sample as listed in Table 8.2, $\mathcal{N}_{\text{selected}}$ is the number of MC events remaining after selection as listed in Table 8.3, and $\varepsilon_{\text{corrections}}$ are efficiency correction factors to account for the differences between data and MC from particle ID, vertex confirmation, taggability and $b$-tagging.

Table 8.4 shows the percentage of each single top quark signal that remains after selection

| The Monte Carlo Event Sets | | | | |
|---|---|---|---|---|
| Event Type | Cross Section [pb] | Branching Fraction | No. of p17 Events | No. of p20 Events |
| Signals $L_V$ (LH vector(SM)) | | | | |
| $tb \to e$+jets | $1.04 \pm 0.14$ | $0.1080 \pm 0.0032$ | 291,094 | 248,415 |
| $tb \to \mu$+jets | $1.04 \pm 0.14$ | $0.1080 \pm 0.0032$ | 289,002 | 263,437 |
| $tb \to \tau$+jets | $1.04 \pm 0.14$ | $0.1080 \pm 0.0032$ | 288,909 | 249,947 |
| $tqb \to e$+jets | $2.26 \pm 0.30$ | $0.1080 \pm 0.0032$ | 290,065 | 273,518 |
| $tqb \to \mu$+jets | $2.26 \pm 0.30$ | $0.1080 \pm 0.0032$ | 289,311 | 274,418 |
| $tqb \to \tau$+jets | $2.26 \pm 0.30$ | $0.1080 \pm 0.0032$ | 290,066 | 247,397 |
| Signals $L_T$ (LH Tensor) | | | | |
| $tb \to e$+jets | $7.33 \pm 1.42$ | $0.1080 \pm 0.0032$ | 284,693 | 267,008 |
| $tb \to \mu$+jets | $7.33 \pm 1.42$ | $0.1080 \pm 0.0032$ | 279,958 | 270,316 |
| $tb \to \tau$+jets | $7.33 \pm 1.42$ | $0.1080 \pm 0.0032$ | 289,259 | 274,869 |
| $tqb \to e$+jets | $1.66 \pm 0.24$ | $0.1080 \pm 0.0032$ | 289,645 | 272,728 |
| $tqb \to \mu$+jets | $1.66 \pm 0.24$ | $0.1080 \pm 0.0032$ | 290,750 | 266,941 |
| $tqb \to \tau$+jets | $1.66 \pm 0.24$ | $0.1080 \pm 0.0032$ | 290,412 | 269,522 |
| Signals $R_V$ (RH vector) | | | | |
| $tb \to e$+jets | $1.07 \pm 0.14$ | $0.1080 \pm 0.0032$ | 289,445 | 272,779 |
| $tb \to \mu$+jets | $1.07 \pm 0.14$ | $0.1080 \pm 0.0032$ | 286,486 | 269,681 |
| $tb \to \tau$+jets | $1.07 \pm 0.14$ | $0.1080 \pm 0.0032$ | 289,943 | 275,415 |
| $tqb \to e$+jets | $2.00 \pm 0.30$ | $0.1080 \pm 0.0032$ | 289,509 | 271,410 |
| $tqb \to \mu$+jets | $2.00 \pm 0.30$ | $0.1080 \pm 0.0032$ | 289,791 | 251,188 |
| $tqb \to \tau$+jets | $2.00 \pm 0.30$ | $0.1080 \pm 0.0032$ | 291,146 | 319,602 |
| Signals $R_T$ (LH Tensor) | | | | |
| $tb \to e$+jets | $7.33 \pm 1.42$ | $0.1080 \pm 0.0032$ | 289,487 | 268,962 |
| $tb \to \mu$+jets | $7.33 \pm 1.42$ | $0.1080 \pm 0.0032$ | 275,773 | 271,280 |
| $tb \to \tau$+jets | $7.33 \pm 1.42$ | $0.1080 \pm 0.0032$ | 288,649 | 270,446 |
| $tqb \to e$+jets | $2.09 \pm 0.22$ | $0.1080 \pm 0.0032$ | 289,064 | 269,924 |
| $tqb \to \mu$+jets | $2.09 \pm 0.22$ | $0.1080 \pm 0.0032$ | 288,618 | 273,234 |
| $tqb \to \tau$+jets | $2.09 \pm 0.22$ | $0.1080 \pm 0.0032$ | 289,946 | 268,642 |
| Signals $L_V + R_T$ | | | | |
| $tb \to e$+jets | $6.11 \pm 0.89$ | $0.1080 \pm 0.0032$ | 289,109 | 271,319 |
| $tb \to \mu$+jets | $6.11 \pm 0.89$ | $0.1080 \pm 0.0032$ | 279,929 | 271,433 |
| $tb \to \tau$+jets | $6.11 \pm 0.89$ | $0.1080 \pm 0.0032$ | 289,326 | 269,098 |
| $tqb \to e$+jets | $4.45 \pm 0.75$ | $0.1080 \pm 0.0032$ | 289,876 | 270,135 |
| $tqb \to \mu$+jets | $4.45 \pm 0.75$ | $0.1080 \pm 0.0032$ | 290,978 | 270,581 |
| $tqb \to \tau$+jets | $4.45 \pm 0.75$ | $0.1080 \pm 0.0032$ | 289,170 | 269,401 |
| Backgrounds | | | | |
| $t\bar{t}$ total | $7.46^{+0.48}_{-0.67}$ | $0.5340 \pm 0.0054$ | 3,060,960 | 3,052,800 |
| $W + jets$ total | $24,182$ | $0.3240 \pm 0.0032$ | 60,346,211 | 72,176,448 |
| $Z + jets$ total | $7,177$ | $0.03366 \pm 0.00002$ | 23,042,922 | 8,888,380 |
| *dibosons* total | $16.2 \pm 0.6$ | $1.0 \pm 0.0$ | 3,937,211 | 1,889,856 |

Table 8.2: The cross sections, branching fractions, initial numbers of events, and integrated luminosities of the Monte Carlo event samples.

| Tagged event yields with channels combined for single top samples | | | | |
|---|---|---|---|---|
| | Electron+Muon, Run IIa+Run IIb | | | |
| | 2 jets | 3 jets | 4 jets | All Channels |
| $LH$ Vector ($L_V$) | | | | |
| $(tb)_{L_V}$ | 104 ± 0.56 | 44 ± 0.36 | 13 ± 0.21 | 160 ± 0.69 |
| $(tqb)_{L_V}$ | 140 ± 0.81 | 72 ± 0.61 | 26 ± 0.37 | 239 ± 1.1 |
| $(tb+tqb)_{L_V}$ | 244 ± 0.98 | 116 ± 0.71 | 39 ± 0.42 | 399 ± 1.3 |
| ($LH$ Tensor($L_T$)) | | | | |
| $(tb)_{R_T}$ | 756 ± 5.6 | 344 ± 3.8 | 103 ± 2.1 | 1,202 ± 7.1 |
| $(tqb)_{R_T}$ | 103 ± 0.86 | 67 ± 0.72 | 28 ± 0.49 | 198 ± 1.2 |
| $(tb+tqb)_{R_T}$ | 858 ± 5.7 | 410 ± 3.9 | 132 ± 2.2 | 1,400 ± 7.2 |
| $RH$ Vector($R_V$) | | | | |
| $(tb)_{R_V}$ | 105 ± 0.75 | 43 ± 0.49 | 12 ± 0.27 | 160 ± 0.93 |
| $(tqb)_{R_V}$ | 122 ± 1.0 | 61 ± 0.74 | 22 ± 0.48 | 206 ± 1.3 |
| $(tb+tqb)_{R_V}$ | 227 ± 1.3 | 104 ± 0.89 | 35 ± 0.55 | 366 ± 1.6 |
| $RH$ Tensor($R_T$) | | | | |
| $(tb)_{L_T}$ | 730 ± 5.4 | 316 ± 3.5 | 92 ± 2.0 | 1,139 ± 6.7 |
| $(tqb)_{L_T}$ | 117 ± 1.0 | 86 ± 0.96 | 40 ± 0.66 | 242 ± 1.6 |
| $(tb+tqb)_{L_T}$ | 847 ± 5.5 | 402 ± 3.6 | 132 ± 2.1 | 1,381 ± 6.9 |
| $L_V + R_T$ | | | | |
| $(tb)_{L_V+R_T}$ | 607 ± 4.5 | 284 ± 3.2 | 86 ± 1.8 | 976 ± 5.8 |
| $(tqb)_{L_V+R_T}$ | 268 ± 2.3 | 167 ± 1.9 | 67 ± 1.2 | 502 ± 3.2 |
| $(tb+tqb)_{L_V+R_T}$ | 874 ± 5.0 | 451 ± 3.7 | 152 ± 2.2 | 1,478 ± 6.6 |

Table 8.3: Tagged event yields for single top samples with statistical uncertainty for each jet multiplicity and for all analysis channels combined.

in each channel after $b$-tagged jets have been required. The entries in this table have been calculated using the yields shown in Table 8.3 and the following equation:

$$\mathcal{A} = \frac{\text{Yield}}{\sigma \times \mathcal{L}} \tag{8.10}$$

in which $\sigma$ is the predicted cross section (for $tb$ or $tqb$) and $\mathcal{L}$ is the integrated luminosity for each sample.

| Combined Signal Acceptance | | | | |
|---|---|---|---|---|
| | Electron+Muon, Run IIa+Run IIb | | | |
| | 2 jets | 3 jets | 4 jets | All Channels |
| $LH$ Vector ($L_V$) | | | | |
| $(tb)_{L_V}$ | 1.9% | 0.78% | 0.23% | 2.9% |
| $(tqb)_{L_V}$ | 1.2% | 0.60% | 0.21% | 2.0% |
| $(tb+tqb)_{L_V}$ | 1.4% | 0.65% | 0.22% | 2.3% |
| $LH$ Tensor($L_T$) | | | | |
| $(tb)_{R_T}$ | 1.9% | 0.87% | 0.26% | 3.0% |
| $(tqb)_{R_T}$ | 1.1% | 0.75% | 0.31% | 2.2% |
| $(tb+tqb)_{R_T}$ | 1.8% | 0.84% | 0.27% | 2.9% |
| $RH$ Vector($R_V$) | | | | |
| $(tb)_{R_V}$ | 1.8% | 0.74% | 0.20% | 2.7% |
| $(tqb)_{R_V}$ | 1.1% | 0.56% | 0.20% | 1.9% |
| $(tb+tqb)_{R_V}$ | 1.3% | 0.63% | 0.21% | 2.1% |
| $RH$ Tensor($R_T$) | | | | |
| $(tb)_{L_T}$ | 1.8% | 0.79% | 0.23% | 2.8% |
| $(tqb)_{L_T}$ | 1.0% | 0.76% | 0.35% | 2.1% |
| $(tb+tqb)_{L_T}$ | 1.6% | 0.79% | 0.25% | 2.6% |
| $L_V + R_T$ | | | | |
| $(tb)_{L_V+R_T}$ | 1.8% | 0.86% | 0.26% | 2.9% |
| $(tqb)_{L_V+R_T}$ | 1.1% | 0.69% | 0.27% | 2.1% |
| $(tb+tqb)_{L_V+R_T}$ | 1.5% | 0.79% | 0.26% | 2.6% |

Table 8.4: Combined signal acceptances for single top samples for each jet multiplicity and for all analysis channels combined.

## 8.3 Systematic Uncertainties for Anomalous Couplings

All the systematic uncertainties accounted in this analysis are the same as in the cross section measurement analysis described in Section 5.6 and details are given in our DØ notes [129, 181]. The systematic uncertainty tables are shown in Appendix A. There is one additional uncertainty considered in this analysis due to the signal mis-modeling effect. As stated in Section 8.2, anomalous single top samples are generated using the same coupling in the top production and top quark decay, e.g. $R_V$ (0100) coupling sample has right-handed coupling both in top quark production and in the top quark decay. But, since top quark has mass, it may be possible that the single top quark is produced through a left-handed coupling and decay through a right-handed coupling (and vice versa). A study is done to properly take this effect into account using the parton level MC samples of different couplings.

A sample $L_V R_V$ (1100) is generated, which allows for top quark to be produced through one coupling and decay through the other coupling. Then already trained scenario $(L_V, R_V)$ BNN filter is applied to this sample to produce the discriminant output and compare this new output to the sum of $L_V$ (1000) and $R_V$ (0100) discriminant outputs.

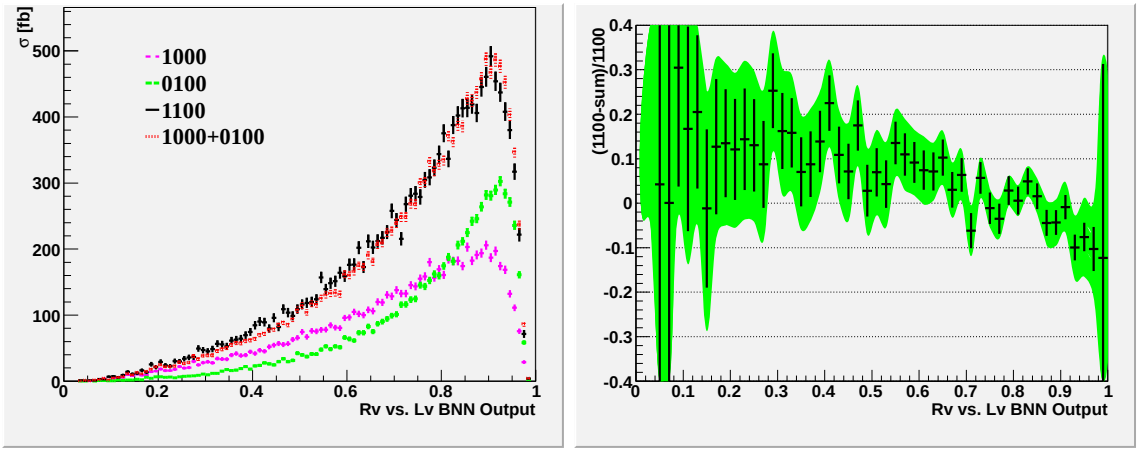


Figure 8.2: Left: Plot showing the overlay of two inputs (1000 and 0100), the 1100 sample, and the simple sum of 1000 and 0100. Right: Plot showing the difference between 1100 and the simple sum model (1000+0100), relative to 1100. The green error band shows the 15% anti-correlated uncertainty on 1000 and 0100.

Figure 8.2 (left) shows the comparison of BNN discriminant outputs of two inputs (1000, 0100), the 1100 sample, and the simple sum of 1000 and 0100. The agreement between simple sum model (1000 + 0100) and the 1100 sample is in general quite good but some differences can be seen in the high discriminant range. To better illustrate the difference, each bin of the 1100 discriminant distribution is taken, subtracted the simple sum, and then result is divided by 1100, shown in Figure 8.2 (right). Figure 8.2 (right) also shows a green error band which reflects a 15% anti-correlated uncertainty on 1000 and 0100. With this additional uncertainty the green band encompasses all the difference. Thus a normalization systematic uncertainty of 15% is imposed to account for any signal mis-modeling effect as reported in our analysis note [181]. This study is performed only in the $(L_V, R_V)$ scenario, but being conservative the same uncertainty is applied in the $(L_V, R_T)$ and $(L_V, L_T)$ scenarios as well.

## 8.4 BNN Analysis

For this analysis, again a multivariate technique, Bayesian Neural Networks as described in Section 7.1 is used to separate the signal from the background. All the steps involved for the BNN analysis are the same as with cross section measurement but for this analysis there are additional single top samples with different couplings. These additional samples are also divided into three sub-samples ("training", "testing" and "yield" samples) after going through all the selections and corrections in a similar way as done in Section 7.1.1.

### Variable Selection

The variables used for the analysis are the same as used in the cross section measurement analysis explained in Section 7.1.2. In addition to these variables, four more angular variables are added based on the top quark spin and $W$ helicity information which provides more sensitivity to the discriminant because of the expected difference in kinematics of the events

with different couplings.

The list of selected variables along with their description, for all the six analysis channels ranked according to their KS-values are described in Tables B.7 to B.12 separately of Appendix B. Same variables are used for all the training scenarios. Figures 8.3 and 8.4 shows some of the sensitive variables used for training.

**Training Scenarios**

As discussed before, in this analysis there are four single top samples: left-handed vector coupling, $L_V$ (SM); right-handed vector coupling $R_V$; left-handed tensor coupling, $L_T$ and right-handed tensor coupling, $R_T$. Our aim is to study the anomalous $Wtb$ couplings and obtain substantial improvements on the limits of these couplings. For that, out of the four couplings, we consider three cases of pairing the left-handed vector coupling with each of the other three non-SM couplings: $(L_V, R_V\ )$, $(L_V, L_T\ )$ and $(L_V, R_T\ )$, and for each case we assume the other two non-SM couplings are negligible. This pairing allows us to limit the complexity of the analysis, while increasing the statistical power and sensitivity for the anomalous coupling under study.

Hence, there are three training scenarios considered for this analysis: first one is $L_V - L_T$ scenario, and the signal in this case is single top ($tb+tqb$) with anomalous left-handed tensor coupling; for the second $L_V - R_V$ scenario, the signal is single top ($tb+tqb$) with anomalous right-handed vector coupling; and for the $L_V - R_T$ scenario, the signal is single top ($tb+tqb$) with anomalous right-handed tensor coupling. In each scenario, signal is trained against backgrounds. The background includes Monte Carlo events for single top ($tb + tqb$) with anomalous left-handed vector coupling, $t\bar{t} \to \ell$+jets, $t\bar{t} \to \ell\ell$+jets, dibosons, $W$+jets with separate components for $Wb\bar{b}$, $Wc\bar{c}$ and $W$+light partons, $Z$+jets with separate components for $Zb\bar{b}$, $Zc\bar{c}$ and $z$+light partons. Each background component is represented in proportion to its expected fraction in the background model.

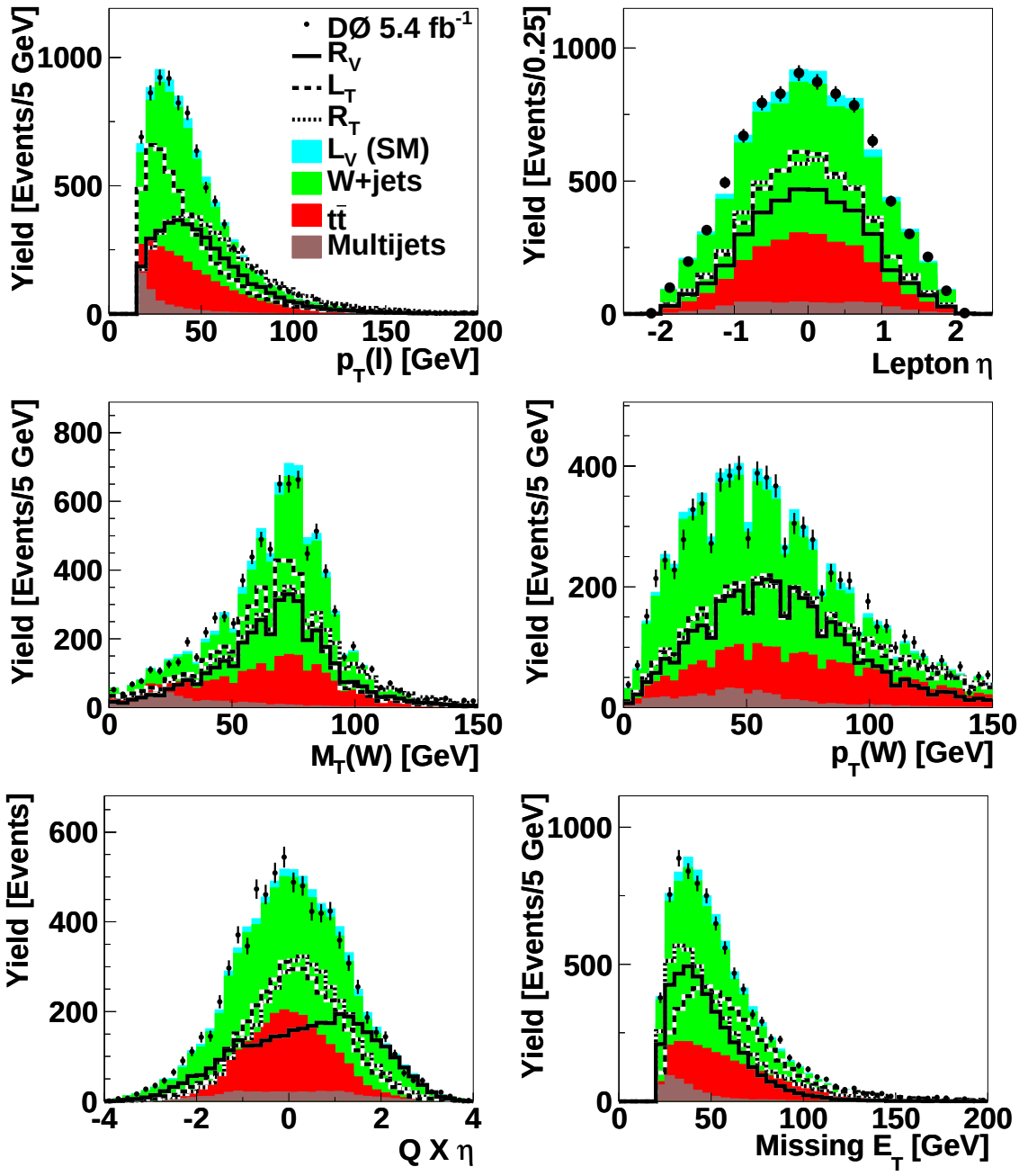


Figure 8.3: Comparison of the SM backgrounds and data for selected discriminating variables with all channels combined. Superimposed are the distributions from single top quark production ($tb + tqb$) with one non-vanishing non-SM coupling (all other couplings set to zero) normalized to 10 times the SM single top quark cross section. The $W$+jets contributions include the smaller backgrounds from $Z$+jets and dibosons.

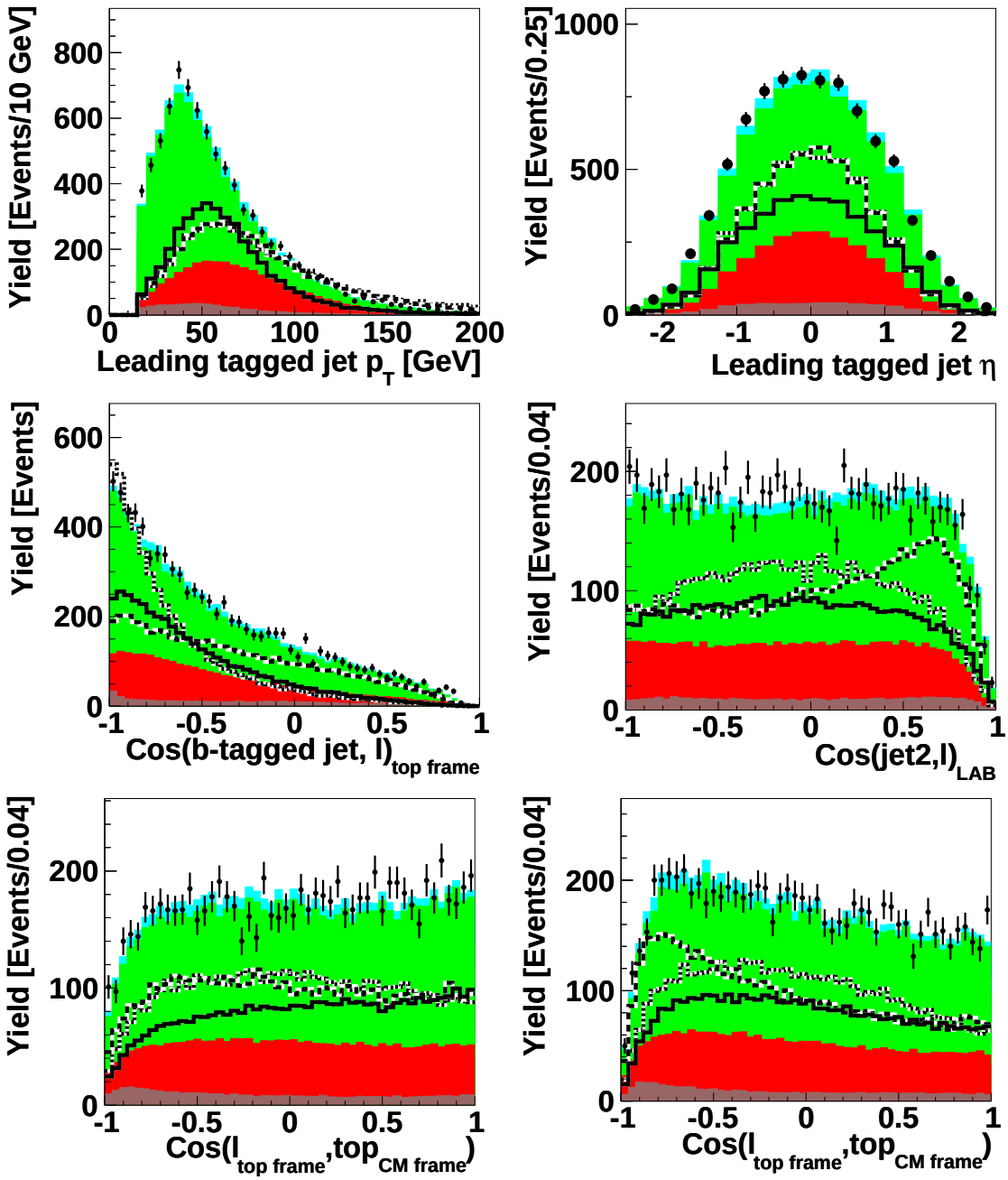


Figure 8.4: Comparison of the SM backgrounds and data for selected discriminating variables with all channels combined. Superimposed are the distributions from single top quark production ($tb + tqb$) with one non-vanishing non-SM coupling (all other couplings set to zero) normalized to 10 times the SM single top quark cross section. The $W$+jets contributions include the smaller backgrounds from $Z$+jets and dibosons.

### 8.4.1 BNN Output

The BNN output discriminant distributions for the six analysis channels considered are plotted using the "yields" sample after the binning transformation as described in Section 7.1.4. Distributions are shown separately for the different scenarios : $L_V - L_T$, $L_V - R_V$, $L_V - R_T$, respectively. Figures 8.5, 8.6, and 8.7 show the BNN output distributions for the 6 analysis channels for the three different scenarios ($L_V - L_T$,$L_V - R_V$,$L_V - R_T$).

BNN OUTPUT FOR $L_V - L_T$ CHANNEL

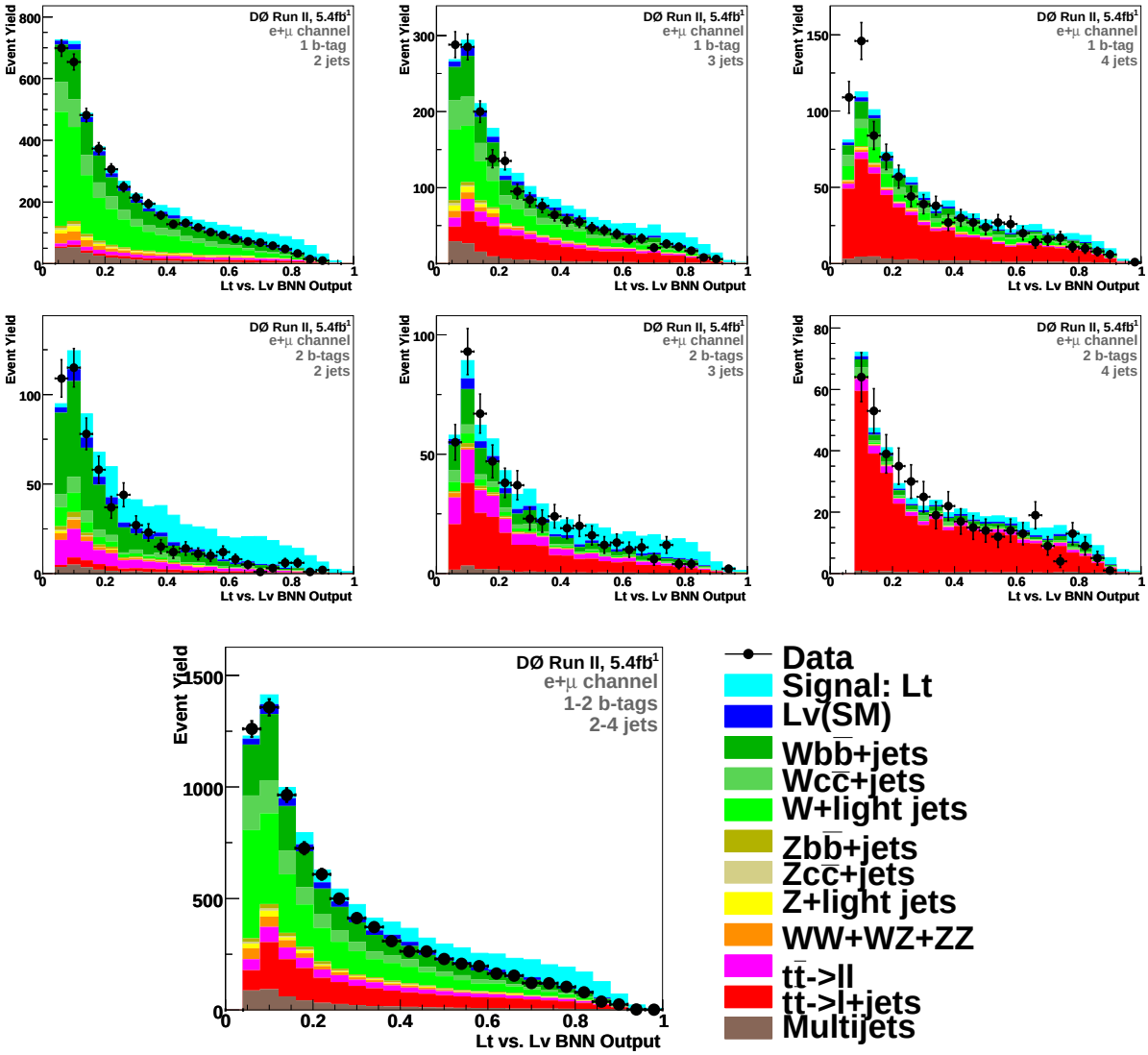


Figure 8.5: BNN output for $L_V - L_T$ Scenario. First two rows show the results for the six individual analysis channels: 2 jets (left), 3 jets (middle), 4 jets (right); single tag (top row), double tag (second row). The bottom left plot shows the discriminant output for the six analysis channels combined and bottom right plots shows the color scheme used in plots of signals and backgrounds in the $L_V - L_T$ Scenario..

BNN OUTPUT FOR $L_V - R_V$ Scenario

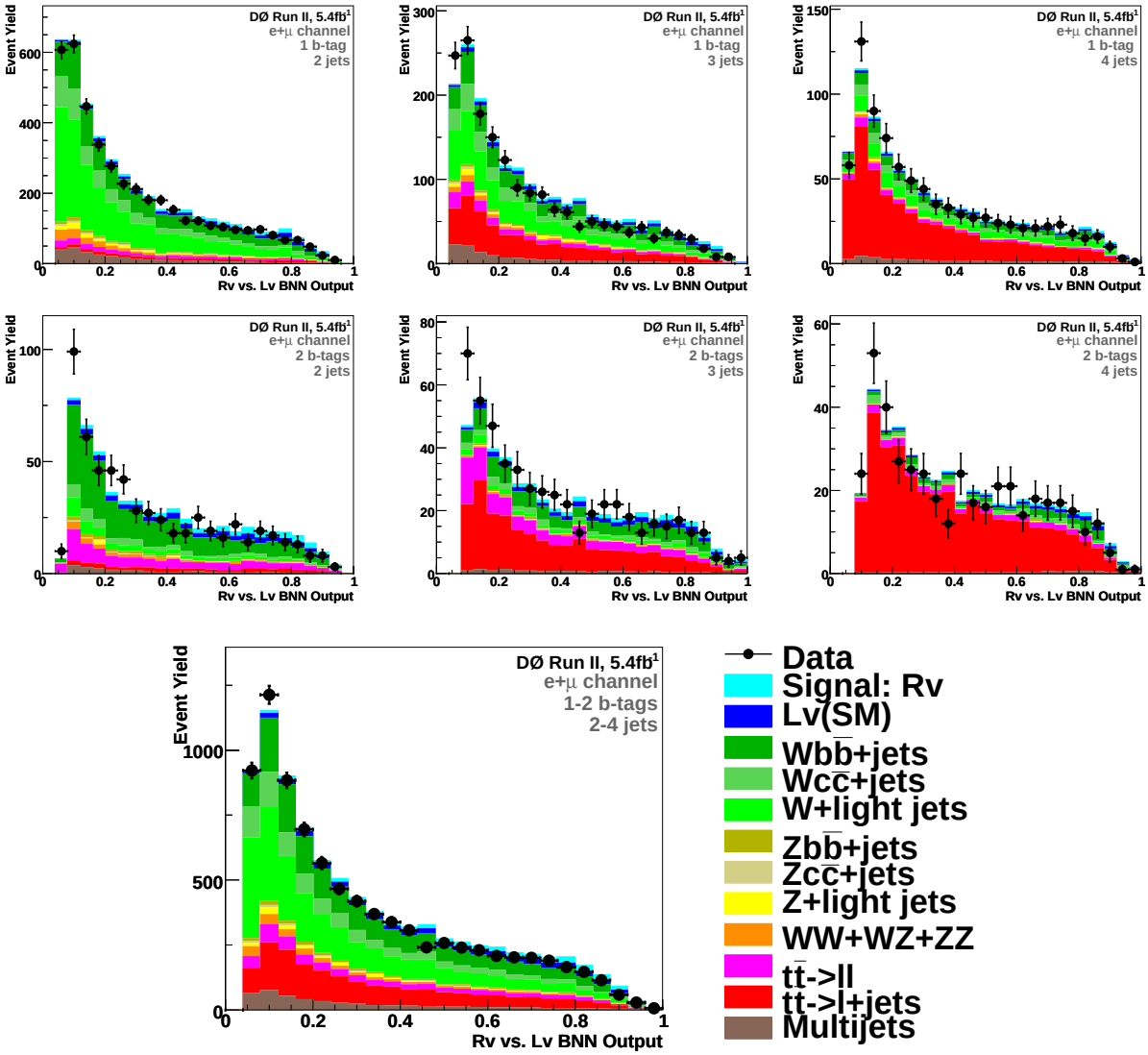


Figure 8.6: BNN output for $L_V - R_V$ Scenario. First two rows show the results for the six individual analysis channels: 2 jets (left), 3 jets (middle), 4 jets (right); single tag (top row), double tag (second row). The bottom left plot shows the discriminant output for the six analysis channels combined and bottom right plots shows the color scheme used in plots of signals and backgrounds in the $L_V - R_V$ Scenario.

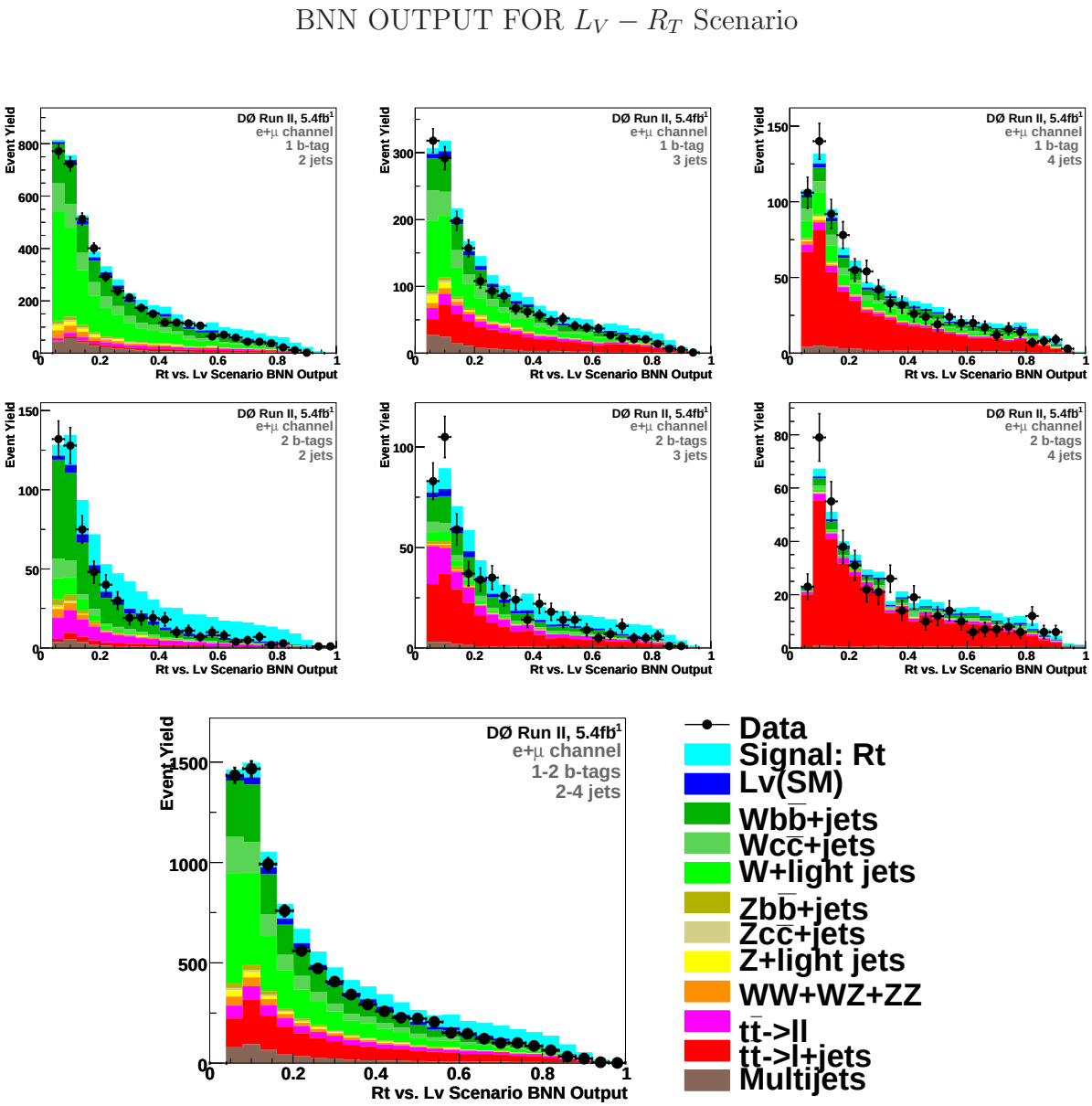


Figure 8.7: BNN output for $L_V - R_T$ Scenario. First two rows show the results for the six individual analysis channels: 2 jets (left), 3 jets (middle), 4 jets (right); single tag (top row), double tag (second row). The bottom left plot shows the discriminant output for the six analysis channels combined and bottom right plots shows the color scheme used in plots of signals and backgrounds in the $L_V - R_T$ Scenario.

### 8.4.2 BNN Cross Checks

The cross check samples are generated for single top samples with anomalous couplings as described in Section 5.5 and BNN output discriminant distributions are plotted after running filters on these samples. This allows us to test whether or not each part of the background model is adequately described in regions where the $W$+jets or the $t\bar{t}$ background dominates. Figure 8.8 shows the BNN output distributions for these cross-check samples for $L_V - L_T$ Scenario. Likewise, Figures 8.9 and 8.10 show the distributions for $L_V - R_V$ and $L_V - R_T$ Scenarios, respectively. Each figure is organized as follows: the two columns in first two rows correspond to events with 1 $b$-tag (left), 2 $b$-tags (center) and 1-2 $b$-tags combined (right). First row corresponds to the $t\bar{t}$, and the bottom one to the $W$+jets cross-check sample. The color scheme used is same as used in Figures 8.5 to 8.7. A good agreement is seen between the predicted background and the observed data in all instances.

CROSS CHECK SAMPLES FOR $L_V - L_T$ Scenario

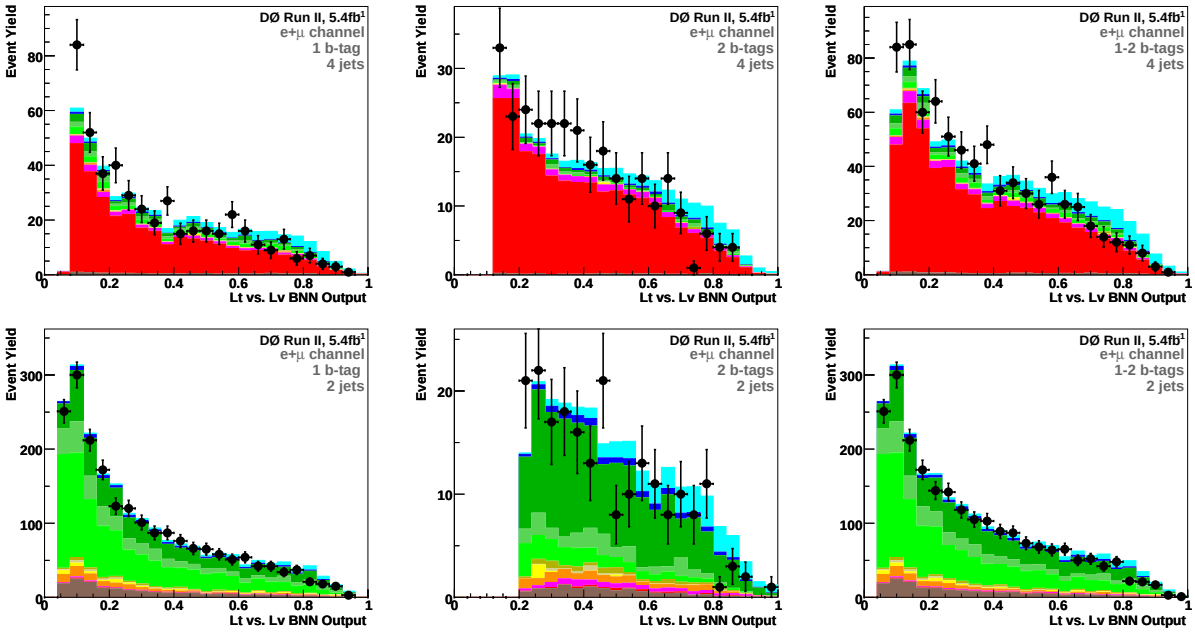


Figure 8.8: BNN cross check samples for $L_V - L_T$ Scenario.

CROSS CHECK SAMPLES FOR $L_V - R_V$ and $L_V - R_T$ Scenarios

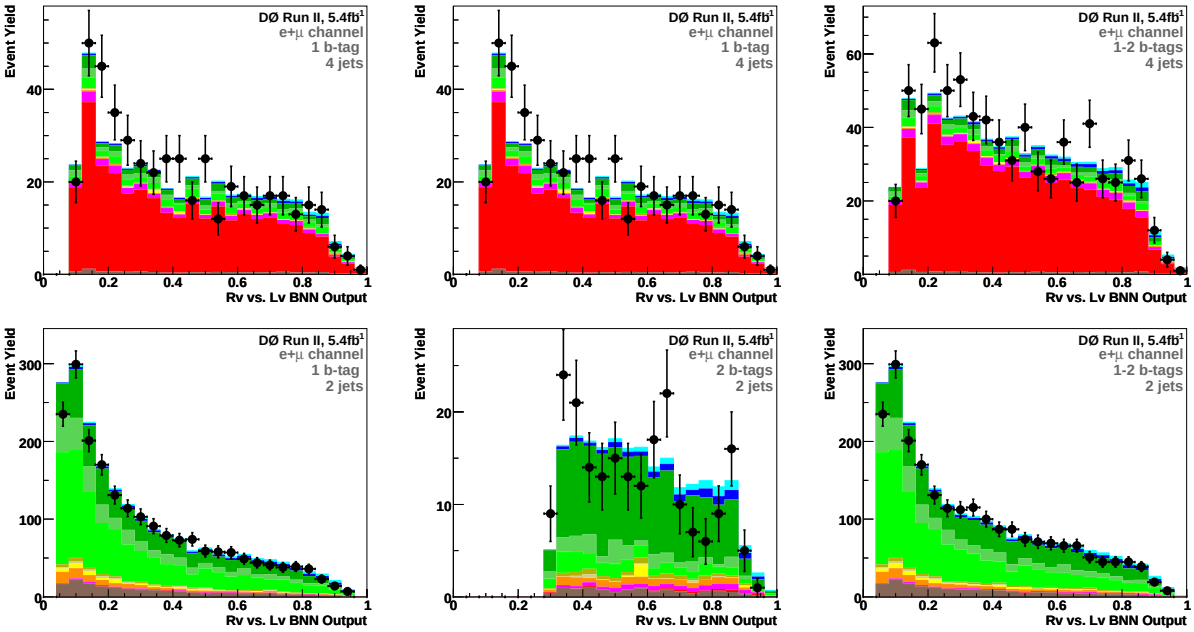


Figure 8.9: BNN cross check samples for $L_V - R_V$ Scenario.

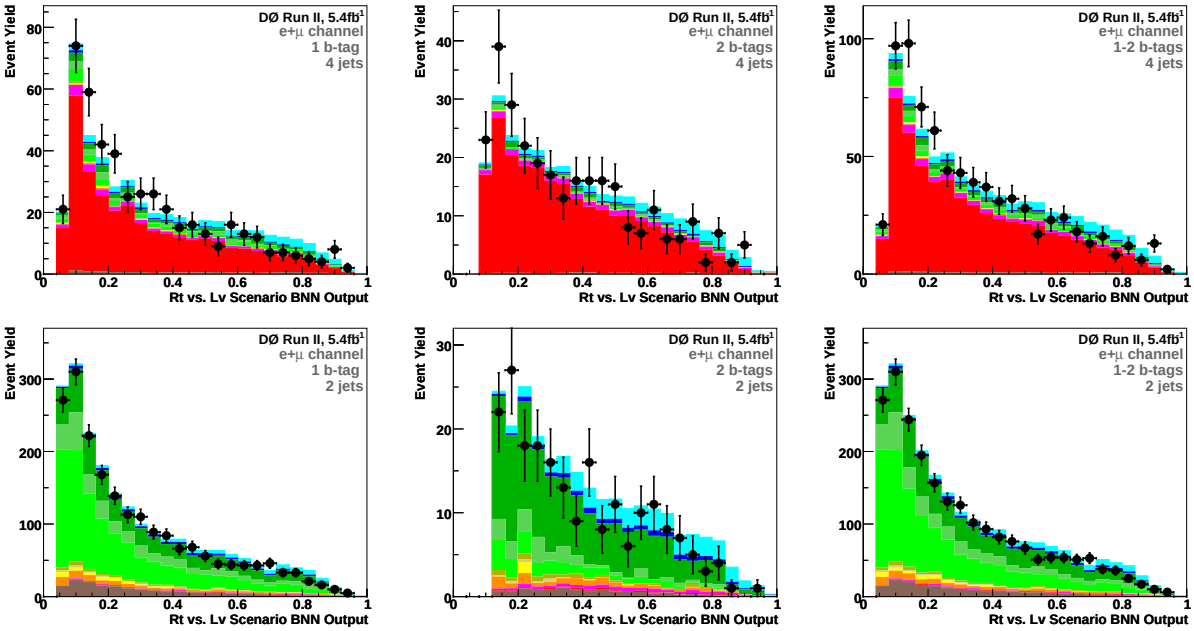


Figure 8.10: BNN cross check samples for $L_V - R_T$ Scenario.

## 8.5 Expected Limits

A Bayesian statistical approach [158, 159, 177, 180] is used to compare data to the signal predictions given by different anomalous couplings using BNN discriminant output distributions. A two-dimensional (2D) posterior probability is computed as a function of $|V_{tb} \cdot f_{L_V}|^2$ and $|V_{tb} \cdot f_X|^2$, where $V_{tb} \cdot f_X$ is one of the two non-SM couplings $X = \{L_T, R_V\}$. For these two cases the single top quark contribution is represented by a superposition of two samples as also described in Section 8.2:

$$s = |V_{tb} \cdot f_{L_V}|^2 s_{L_V} + |V_{tb} \cdot f_X|^2 s_X, \tag{8.11}$$

where $s_{L_V}$ ($s_X$) are the mean expected count of single top quarks for the assumptions $f_{L_V} = 1$ ($f_X = 1$) and the other couplings are set to zero. In the $(L_V, R_T)$ scenario, the two couplings interfere, and to account for the effect of the interference, the single top quark contribution is represented by the superposition of three samples:

$$\begin{aligned} s &= |V_{tb} \cdot f_{L_V}|^2 s_{L_V} + |V_{tb} \cdot f_{R_T}|^2 s_{R_T} + \\ &+ |V_{tb} \cdot f_{L_V}||V_{tb} \cdot f_{R_T}|(s_{L_V R_T} - s_{L_V} - s_{R_T}), \end{aligned} \tag{8.12}$$

where $s_{R_T}$ is the mean count assuming a right-handed tensor coupling only $f_{R_T} = 1$, and $s_{L_V R_T}$ is the one where both couplings $f_{L_V} = 1$ and $f_{R_T} = 1$. The last sample is indicated as "$L_V + R_T$". A Poisson distribution for data counts and uniform prior probability for nonnegative values of the SM and non-SM couplings is assumed. The output discriminants for the signal, backgrounds, and data as shown in Figures 8.5 to 8.7 are used to form a binned likelihood as a product over all six analysis channels and all bins, taking into account all systematic uncertainties and their correlations. The expected posterior probabilities are obtained by setting the number of data counts to be equal to the predicted sum of the signal and backgrounds. The expected two dimensional posterior probability distributions with

68%, 90% and 95% confidence level contours are shown in the Figures 8.11 to 8.19. The two dimensional probabilities are projected onto the two axes and if the projection has a local maximum, its position is quoted as the measured value of the coupling and if it does not have a local maximum, a 95% C.L. upper limit is quoted. Figures 8.11 to 8.19 show the (first row ) expected 2D posterior density distributions for cross section and resulting 1D posterior probabilities for cross section after integrating over the two axes of 2D distribution and (second row) expected 2D posterior density distributions for $|V_{tb} \cdot f|^2$ and resulting 1D posterior probabilities for $|V_{tb} \cdot f|^2$ after integrating over the two axes of 2D distribution. For all the three scenarios, three expected cases are studied by either setting one of the coupling to "0" or setting both equals to "1". All the cases, gives the reasonable expected result. Table 8.5 summarizes the expected result obtained [181].

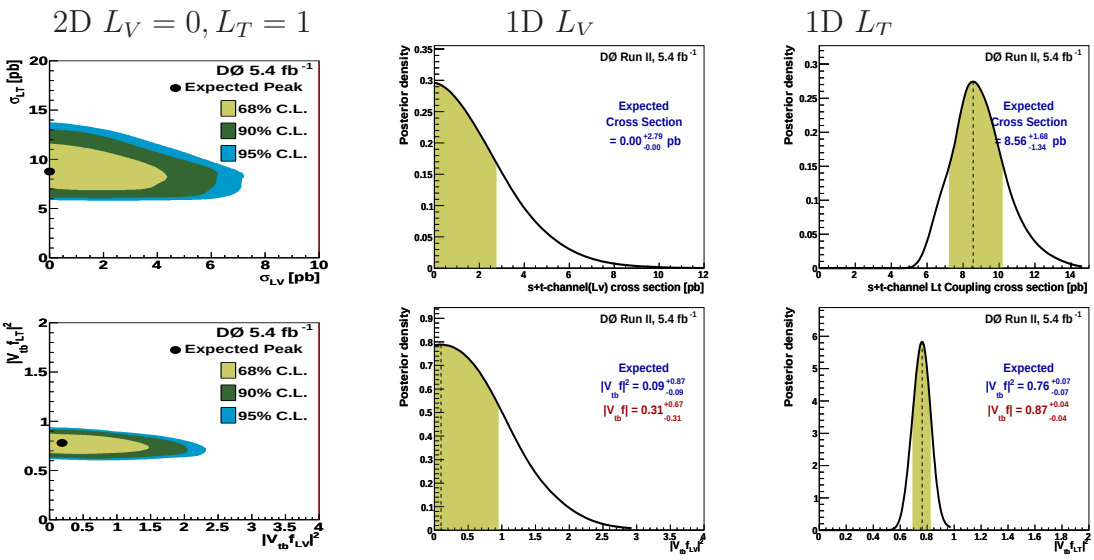


Figure 8.11: $L_V - L_T$ scenario ($L_V = 0, L_T = 1$): first row, first plot is expected 2D posterior density distribution for cross section, second(third) plot is the resulting 1D posterior densities for $L_V(L_T)$ after integrating over $L_T(L_V)$; second row: first plot is expected 2D posterior density distribution for $|V_{tb}.f|^2$ measurement, second(third) plot is the resulting 1D posterior densities for $L_V(L_T)$ after integrating over $L_T(L_V)$.

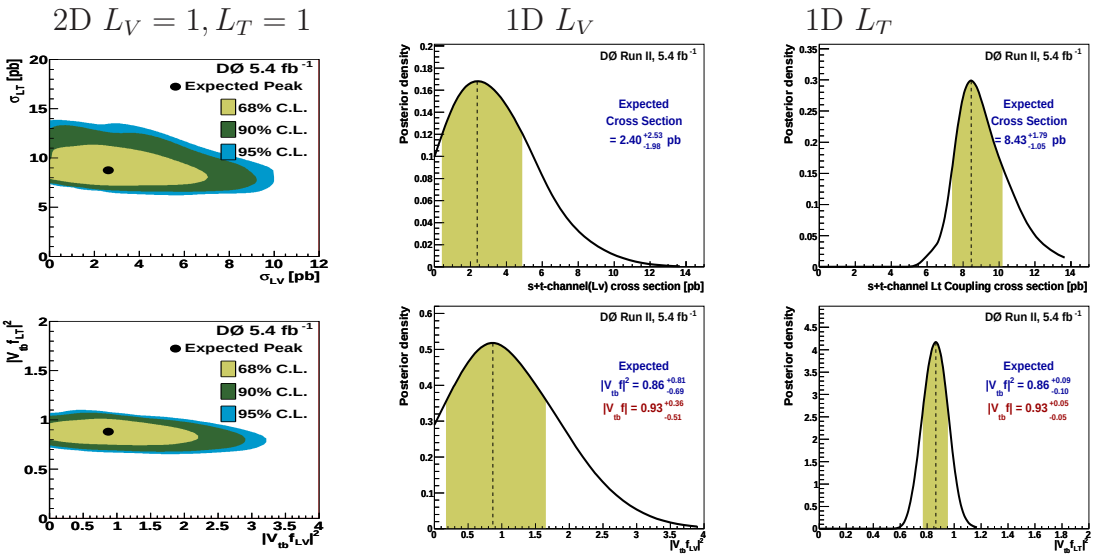


Figure 8.12: $L_V - L_T$ scenario ($L_V = 1, L_T = 1$): first row, first plot is expected 2D posterior density distribution for cross section, second(third) plot is the resulting 1D posterior densities for $L_V(L_T)$ after integrating over $L_T(L_V)$; second row: first plot is expected 2D posterior density distribution for $|V_{tb}.f|^2$ measurement, second(third) plot is the resulting 1D posterior densities for $L_V(L_T)$ after integrating over $L_T(L_V)$.

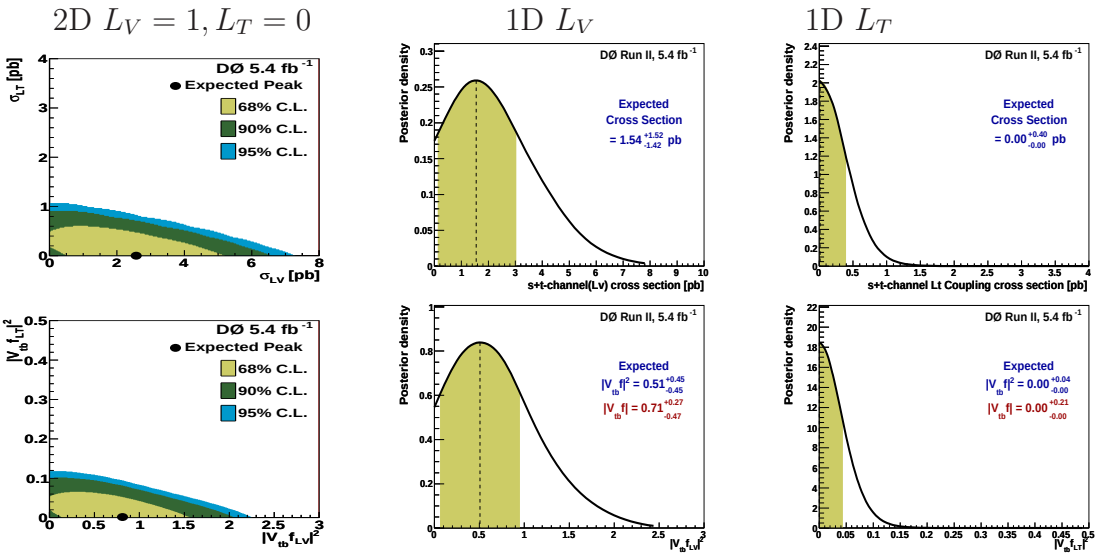


Figure 8.13: $L_V - L_T$ scenario ($L_V = 1, L_T = 0$): first row, first plot is expected 2D posterior density distribution for cross section, second(third) plot is the resulting 1D posterior densities for $L_V(L_T)$ after integrating over $L_T(L_V)$; second row: first plot is expected 2D posterior density distribution for $|V_{tb}.f|^2$ measurement, second(third) plot is the resulting 1D posterior densities for $L_V(L_T)$ after integrating over $L_T(L_V)$.

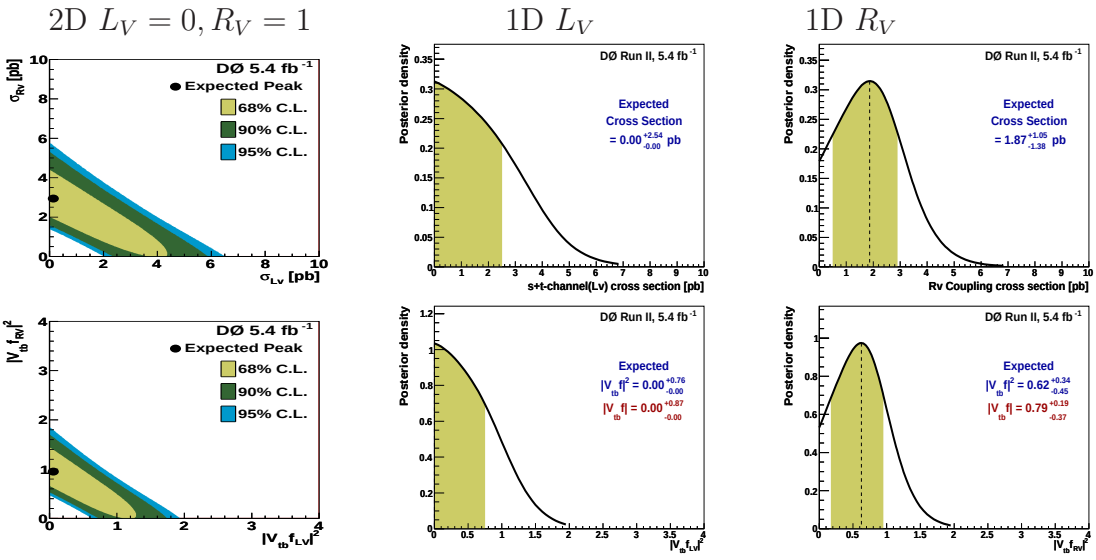


Figure 8.14: $L_V - R_V$ scenario ($L_V = 0, R_V = 1$): first row, first plot is expected 2D posterior density distribution for cross section, second(third) plot is the resulting 1D posterior densities for $L_V(R_V)$ after integrating over $R_V(L_V)$; second row: first plot is expected 2D posterior density distribution for $|V_{tb}.f|^2$ measurement, second(third) plot is the resulting 1D posterior densities for $L_V(R_V)$ after integrating over $R_V(L_V)$.

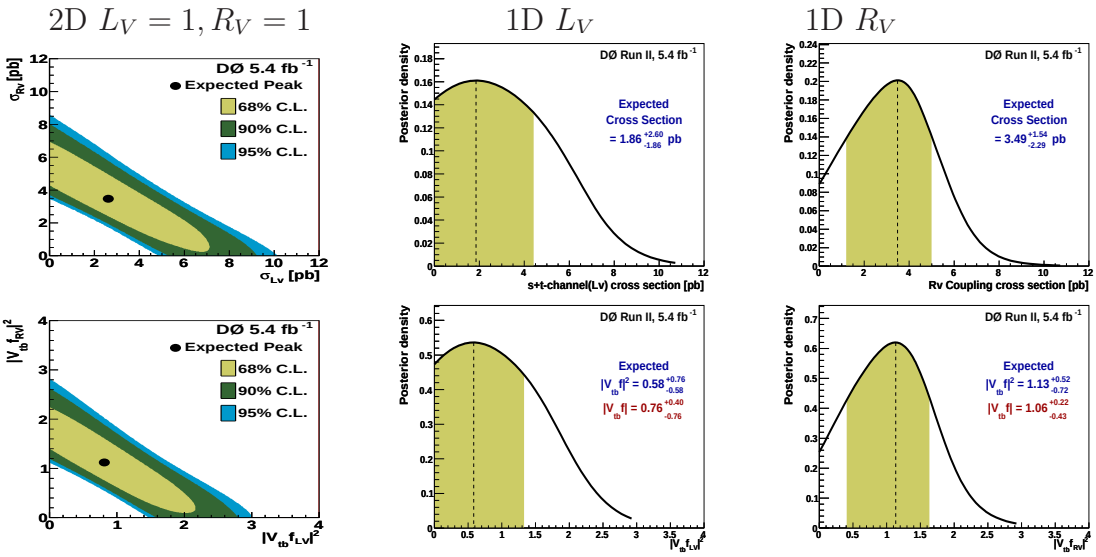


Figure 8.15: $L_V - R_V$ scenario ($L_V = 1, R_V = 1$): first row, first plot is expected 2D posterior density distribution for cross section, second(third) plot is the resulting 1D posterior densities for $L_V(R_V)$ after integrating over $R_V(L_V)$; second row: first plot is expected 2D posterior density distribution for $|V_{tb}.f|^2$ measurement, second(third) plot is the resulting 1D posterior densities for $L_V(R_V)$ after integrating over $R_V(L_V)$.

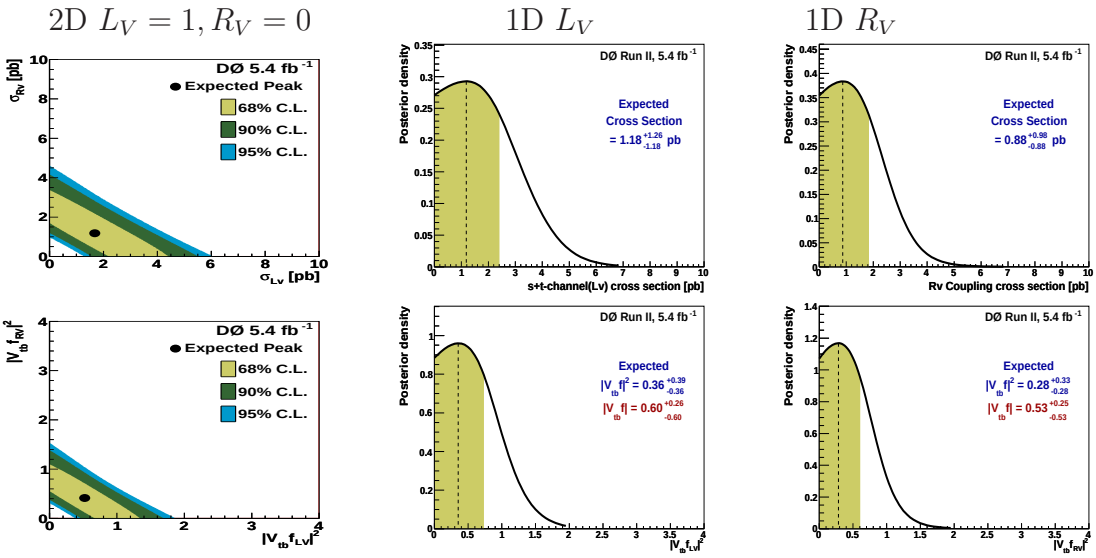


Figure 8.16: $L_V - R_V$ scenario ($L_V = 1, R_V = 0$): first row, first plot is expected 2D posterior density distribution for cross section, second(third) plot is the resulting 1D posterior densities for $L_V(R_V)$ after integrating over $R_V(L_V)$; second row: first plot is expected 2D posterior density distribution for $|V_{tb}.f|^2$ measurement, second(third) plot is the resulting 1D posterior densities for $L_V(R_V)$ after integrating over $R_V(L_V)$.

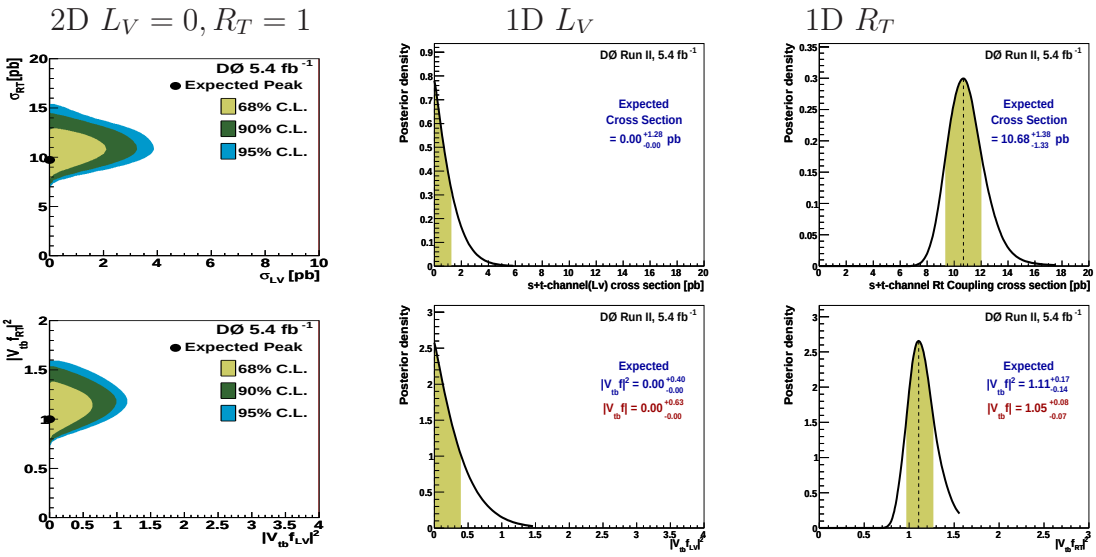


Figure 8.17: $L_V - R_T$ scenario ($L_V = 0, R_T = 1$): first row, first plot is expected 2D posterior density distribution for cross section, second(third) plot is the resulting 1D posterior densities for $L_V(R_T)$ after integrating over $R_T(L_V)$; second row: first plot is expected 2D posterior density distribution for $|V_{tb}.f|^2$ measurement, second(third) plot is the resulting 1D posterior densities for $L_V(R_T)$ after integrating over $R_T(L_V)$.

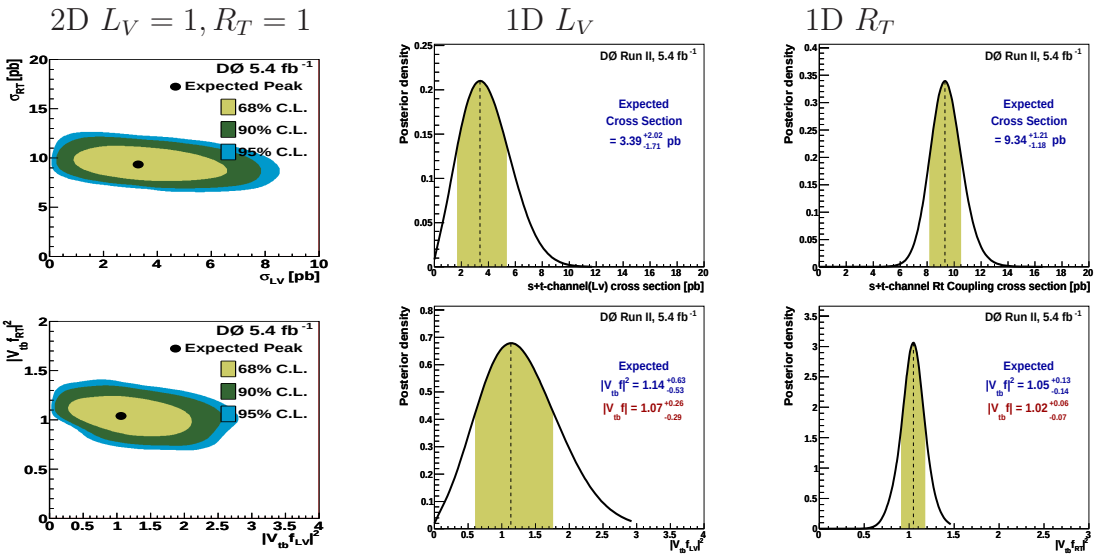


Figure 8.18: $L_V - R_T$ scenario ($L_V = 1, R_T = 1$): first row, first plot is expected 2D posterior density distribution for cross section, second(third) plot is the resulting 1D posterior densities for $L_V(R_T)$ after integrating over $R_T(L_V)$; second row: first plot is expected 2D posterior density distribution for $|V_{tb}.f|^2$ measurement, second(third) plot is the resulting 1D posterior densities for $L_V(R_T)$ after integrating over $R_T(L_V)$.

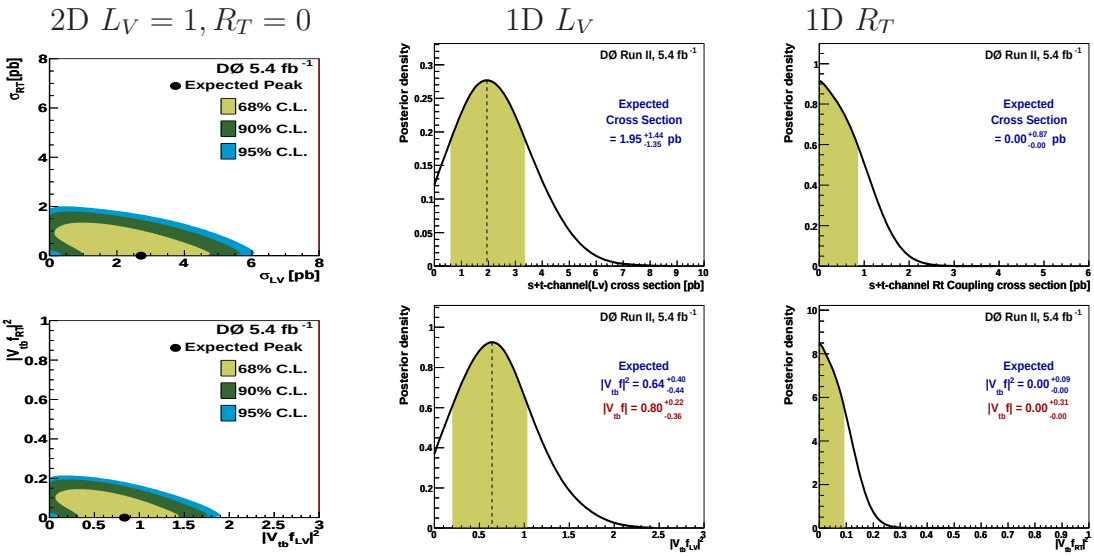


Figure 8.19: $L_V - R_T$ scenario ($L_V = 1, R_T = 0$): first row, first plot is expected 2D posterior density distribution for cross section, second(third) plot is the resulting 1D posterior densities for $L_V(R_T)$ after integrating over $R_T(L_V)$; second row: first plot is expected 2D posterior density distribution for $|V_{tb}.f|^2$ measurement, second(third) plot is the resulting 1D posterior densities for $L_V(R_T)$ after integrating over $R_T(L_V)$.

| | Expected Cross Section | Expected $\lvert V_{tb}.f\rvert^2$ |
|---|---|---|
| $\underline{L_V\text{-}L_T}$ | | |
| $L_V$=0,$L_T$=1 | | |
| $L_V$(SM) | $0.00^{+2.79}_{-0.00}$ pb | $0.09^{+0.87}_{-0.09}$ |
| $L_T$ | $8.56^{+0.68}_{-0.34}$ pb | $0.76^{+0.07}_{-0.07}$ |
| $L_V$=1,$L_T$=1 | | |
| $L_V$(SM) | $2.40^{+2.53}_{-1.98}$ pb | $0.86^{+0.81}_{-0.69}$ |
| $L_T$ | $8.43^{+1.79}_{-1.05}$ pb | $0.86^{+0.09}_{-0.10}$ |
| $L_V$=1,$L_T$=0 | | |
| $L_V$(SM) | $1.54^{+1.52}_{-1.42}$ pb | $0.51^{+0.45}_{-0.45}$ |
| $L_T$ | $0.00^{+0.40}_{-0.00}$ pb | $0.00^{0.04}_{-0.00}$ |
| $\underline{L_V\text{-}R_V}$ | | |
| $L_V$=0,$R_V$=1 | | |
| $L_V$(SM) | $0.00^{+2.54}_{-0.00}$ pb | $0.00^{+0.76}_{-0.00}$ |
| $R_V$ | $1.87^{+1.05}_{-1.38}$ pb | $0.62^{+0.34}_{-0.35}$ |
| $L_V$=1,$R_V$=1 | | |
| $L_V$(SM) | $1.86^{+2.60}_{-1.86}$ pb | $0.58^{+0.76}_{-0.58}$ |
| $R_V$ | $3.49^{+1.54}_{-2.29}$ pb | $1.13^{+0.52}_{-0.72}$ |
| $L_V$=1,$R_V$=0 | | |
| $L_V$(SM) | $1.18^{+1.26}_{-1.18}$ pb | $0.36^{+0.39}_{-0.36}$ |
| $R_V$ | $0.88^{+0.98}_{-0.88}$ pb | $0.28^{+0.33}_{-0.28}$ |
| $\underline{L_V\text{-}R_T}$ | | |
| $L_V$=0,$R_T$=1 | | |
| $L_V$(SM) | $0.00^{+1.28}_{-0.00}$ pb | $0.00^{+0.40}_{-0.00}$ |
| $R_T$ | $10.68^{+1.38}_{-1.33}$ pb | $1.11^{+0.17}_{-0.14}$ |
| $L_V$=1,$R_T$=1 | | |
| $L_V$(SM) | $3.39^{+2.02}_{-1.71}$ pb | $1.14^{+0.63}_{-0.53}$ |
| $R_T$ | $9.34^{+1.21}_{-1.18}$ pb | $1.05^{+0.13}_{-0.14}$ |
| $L_V$=1,$R_T$=0 | | |
| $L_V$(SM) | $1.95^{+1.44}_{-1.35}$ pb | $0.64^{+0.40}_{-0.44}$ |
| $R_T$ | $0.00^{+0.87}_{-0.00}$ pb | $0.00^{+0.09}_{-0.00}$ |

Table 8.5: Expected cross section and $\lvert V_{tb}.f\rvert^2$ with systematic uncertainties for three different Scenarios.

## 8.6 Observed Limits

After obtaining the reasonable expected result, the measurement is done with actual data using the same technique. The observed 2D posterior probability density distributions for all the three different scenarios $L_V - L_T$, $L_V - R_V$ and $L_V - R_T$ are shown in Figure 8.20. We do not observe significant deviations from the SM expectations and therefore compute 95% C.L. upper limits on the anomalous couplings by integrating out the left-handed vector coupling to get a one-dimensional posterior probability density. Table 8.6 summarizes the observed cross section and coupling values with all the systematic uncertainties included and this result have been published in Ref. [182].

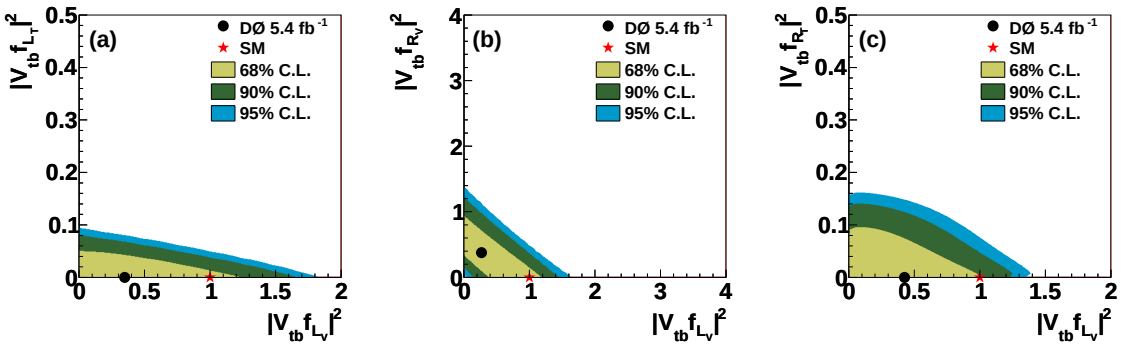


Figure 8.20: Two-dimensional posterior probability density distributions for the anomalous couplings. The left row (a) shows the distribution for the ($L_V$, $L_T$) scenario, the middle row (b) for the ($L_V$, $R_V$) scenario, and the right row (c) for the ($L_V$, $R_T$) scenario. The dots represent the peak posterior from our data in comparison with the SM predictions.

| Scenario | Observed Cross Section | Observed Coupling |
|---|---|---|
| ($L_V$-$L_T$) | $< 0.60$ pb | $\|V_{tb} \cdot f_{L_T}\|^2 < 0.06$ |
| ($L_V$-$R_V$) | $< 2.81$ pb | $\|V_{tb} \cdot f_{R_V}\|^2 < 0.93$ |
| ($L_V$-$R_T$) | $< 1.21$ pb | $\|V_{tb} \cdot f_{R_T}\|^2 < 0.13$ |

Table 8.6: Observed One-dimensional upper limits at 95% C.L. for anomalous $Wtb$ couplings in the three scenarios..

## 8.7 Results

This analysis searches for the anomalous $Wtb$ couplings using 5.4 $fb^{-1}$ of DØ data in the single top quark final state and present new direct constraints on a general $Wtb$ interaction. The standard model provides a purely left-handed vector coupling at the $Wtb$ vertex, while the most general, lowest dimension Lagrangian allows right-handed vector and left- or right-handed tensor couplings as well. We obtain precise limits on these anomalous couplings by comparing the data to the expectations from different assumptions on the $Wtb$ coupling. We find no evidence for anomalous couplings and set 95% C.L. limits on these couplings as $|V_{tb} \cdot f_{L_T}|^2 < 0.06$, $|V_{tb} \cdot f_{R_V}|^2 < 0.93$ and $|V_{tb} \cdot f_{R_T}|^2 < 0.13$. The results so obtained have been published in Ref. [182]. These represent improvements in the limits by factors of 2.6 to 5.0 in terms of couplings squared compared to the previous results [180] while a factor of approximately 2.5 is expected from the increase in integrated luminosity. This result represents the most stringent direct constraints on anomalous $Wtb$ interactions.

# Chapter 9

# Conclusions

This thesis presents a measurement of single top quark production cross section and searches for anomalous $Wtb$ couplings in single top quark production using 5.4 $fb^{-1}$ of data collected in the DØ detector at Fermilab. The events selected must have one isolated lepton, large $\not\!\!E_T$, at least one $b$-tagged jet, and at least one more light jet. A multivariate analysis, the Bayesian Neural Networks, is performed to separate out the small signal from a large background where the uncertainties in the background modeling are of the same order as the signal yield. The Bayesian neural network output distributions across the six independent channels used in the analysis, are combined using a binned likelihood to measure the $s+t$ ($tb+tqb$), $s$ ($tb$) and $t$ ($tqb$)-channel single top cross section and to measure the precise limits on the anomalous couplings by comparing the data to the expectations from different assumptions on the $Wtb$ coupling using a Bayesian method.

For the top quark mass of 172.5 GeV, using Bayesian neural network analysis, we measure the cross sections for $s+t$ ($tb+tqb$), $s$ ($tb$) and $t$ ($tqb$) production to be

$$\sigma(p\bar{p} \rightarrow tb + X, tqb + X) = 3.11^{+0.77}_{-0.71} \text{ pb}$$

$$\sigma(p\bar{p} \rightarrow tb + X) = 0.72^{+0.44}_{-0.43} \text{ pb}$$

$$\sigma(p\bar{p} \to tqb + X) = 2.92^{+0.87}_{-0.73} \text{ pb}$$

with a significance of $4.7\sigma$, $0.9\sigma$ and $4.7\sigma$, respectively, for $s+t$, $s$ and $t$-channels. By combining these results from Bayesian neural networks analysis with other two parallel multivariate analyses, the final combined measurement for $s+t$, $s$ and $t$-channels is

$$\sigma(p\bar{p} \to tb + X, tqb + X) = 3.43^{+0.73}_{-0.74} \text{ pb}$$

$$\sigma(p\bar{p} \to tb + X) = 0.68^{+0.38}_{-0.35} \text{ pb}$$

$$\sigma(p\bar{p} \to tqb + X) = 2.86^{+0.69}_{-0.63} \text{ pb}$$

with a significance of $5.4\sigma$, $1.8\sigma$ and $5.0\sigma$ respectively for $s+t$, $s$ and $t$-channels [166]. Hence the first observation of single top $t$-channel and the most accurate measurements of both $s+t$-channel and $s$-channel single top have been presented. This measurement of $tb + tqb$ production assumes the SM ratio between $tb$ and $tqb$ production; and of $tb$ and $tqb$ measurement assumes $tqb$ and $tb$ production rates respectively, as predicted by the SM. Another measurement of $tqb$ production cross section is done using the same dataset and discriminant but without any assumption on the $tb$ production rate. From this measurement, we obtain a cross section of $2.90^{+0.59}_{-0.59}$ pb for $t$-channel which corresponds to a significance of $5.5\sigma$, and these result have been published [167]. Also, a direct limit on the CKM matrix element $|V_{tb}| > 0.79$ at 95% C.L. assuming a flat prior with $0 \leq |V_{tb}|^2 \leq 1$ is derived and the results so obtained have been published in Ref. [166].

We find no evidence for anomalous couplings and set 95% C.L. limits on these couplings as $|V_{tb} \cdot f_{L_T}|^2 < 0.06$, $|V_{tb} \cdot f_{R_V}|^2 < 0.93$ and $|V_{tb} \cdot f_{R_T}|^2 < 0.13$. The results so obtained have been published in Ref. [182]. These represent improvements in the limits by factors of 2.6 to 5.0 in terms of couplings squared compared to the previous results while a factor of approximately 2.5 is expected from the increase in integrated luminosity. This result represents the most stringent direct constraints on anomalous $Wtb$ interactions.

# Appendix A

# Systematic Uncertainty Tables

The following tables A.1- A.6 in this appendix provides the full list of all the systematics uncertainties considered in this analysis for each six analysis channels separately.

Since before $b$-tagging, $W + jets$ and multijets backgrounds are normalized to data, hence the simulated $W + jets$ components are not affected by most of the systematic uncertainties. However, there are uncertainties on the relative compositions of the $W + jets$ components, and due to the $W + jets$ and multijets normalization. These IKS systematic uncertainties are anti-correlated due to the constraint to match data before $b$-tagging, which is indicated by a minus sign for one of the values. Also, the ISR/FSR systematics are treated as anti-correlated across the different jet multiplicities.

UNCERTAINTIES FOR CHANNELS WITH TWO JETS

SINGLE TAG

| | Percentage Errors | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $t\bar{t}ll$ | $t\bar{t}lj$ | $Wbb$ | $Wcc$ | $Wlp$ | $Zbb$ | $Zcc$ | $Zlp$ | dibosons | multijet | $tb$ | $tqb$ | $tb+tqb$ |
| Luminosity | 6.1 | 6.1 | — | — | — | 6.1 | 6.1 | 6.1 | 6.1 | — | 6.1 | 6.1 | 6.1 |
| Xsect. | 9.0 | 9.0 | — | — | — | 3.3 | 3.3 | 3.3 | 7.0 | — | 3.8 | 5.3 | 4.8 |
| Branching frac. | 1.5 | 1.5 | — | — | — | — | — | — | — | — | 1.5 | 1.5 | 1.5 |
| PDF | — | — | — | — | — | — | — | — | — | — | 2.0 | 2.0 | 2.0 |
| Triggers | 5.0 | 5.0 | — | — | — | 5.0 | 5.0 | 5.0 | 5.0 | — | 5.0 | 5.0 | 5.0 |
| Lumi. rewtg. | 1.0 | 1.0 | — | — | — | 1.0 | 1.0 | 1.0 | 1.0 | — | 1.0 | 1.0 | 1.0 |
| Prim. vertex | 1.4 | 1.4 | — | — | — | 1.4 | 1.4 | 1.4 | 1.4 | — | 1.4 | 1.4 | 1.4 |
| Color Reconnection | 1.0 | 1.0 | — | — | — | — | — | — | — | — | 1.0 | 1.0 | 1.0 |
| b/light Jet Response | −0.4 | −0.4 | — | — | — | 1.0 | — | — | — | — | 0.5 | 0.5 | 0.5 |
| Lepton ID | 2.9 | 2.9 | — | — | — | 2.4 | 2.4 | 2.6 | 2.7 | — | 2.8 | 2.8 | 2.8 |
| Jet frag. and higher order | 1.6 | 1.6 | — | — | — | — | — | — | — | — | 0.7 | 0.7 | 0.7 |
| ISR/FSR | 4.3 | 4.3 | — | — | — | — | — | — | — | — | 0.8 | 0.8 | 0.8 |
| b-jet frag. | 2.0 | 2.0 | — | — | — | 2.0 | — | — | — | — | 2.0 | 2.0 | 2.0 |
| Taggability | 7.1 | 5.6 | 5.6 | 5.1 | 6.2 | 6.2 | 6.4 | 5.8 | 5.7 | — | 6.6 | 5.9 | 6.1 |
| $\lambda^W_{\mathrm{HF}}$ | — | — | 12.0 | 12.0 | — | — | — | — | — | — | — | — | — |
| $\lambda^Z_{\mathrm{HF}}$ | — | — | — | — | — | 12.0 | 12.0 | — | — | — | — | — | — |
| IKS | — | — | 1.8 | 1.8 | 1.8 | — | — | — | — | −34.6 | — | — | — |

Table A.1: Uncertainties requiring exactly one tag and two jets.

DOUBLE TAG

| | Percentage Errors | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $t\bar{t}ll$ | $t\bar{t}lj$ | $Wbb$ | $Wcc$ | $Wlp$ | $Zbb$ | $Zcc$ | $Zlp$ | dibosons | multijet | $tb$ | $tqb$ | $tb+tqb$ |
| Luminosity | 6.1 | 6.1 | — | — | — | 6.1 | 6.1 | 6.1 | 6.1 | — | 6.1 | 6.1 | 6.1 |
| Xsect. | 9.0 | 9.0 | — | — | — | 3.3 | 3.3 | 3.3 | 7.0 | — | 3.8 | 5.3 | 4.8 |
| Branching frac. | 1.5 | 1.5 | — | — | — | — | — | — | — | — | 1.5 | 1.5 | 1.5 |
| PDF | — | — | — | — | — | — | — | — | — | — | 2.0 | 2.0 | 2.0 |
| Triggers | 5.0 | 5.0 | — | — | — | 5.0 | 5.0 | 5.0 | 5.0 | — | 5.0 | 5.0 | 5.0 |
| Lumi. rewtg. | 1.0 | 1.0 | — | — | — | 1.0 | 1.0 | 1.0 | 1.0 | — | 1.0 | 1.0 | 1.0 |
| Prim. vertex | 1.4 | 1.4 | — | — | — | 1.4 | 1.4 | 1.4 | 1.4 | — | 1.4 | 1.4 | 1.4 |
| Color Reconnection | 1.0 | 1.0 | — | — | — | — | — | — | — | — | 1.0 | 1.0 | 1.0 |
| b/light Jet Response | −0.4 | −0.4 | — | — | — | 1.0 | — | — | — | — | 0.5 | 0.5 | 0.5 |
| Lepton ID | 2.9 | 2.8 | — | — | — | 2.3 | 2.4 | 3.2 | 2.7 | — | 2.8 | 2.8 | 2.8 |
| Jet frag. and higher order | 1.6 | 1.6 | — | — | — | — | — | — | — | — | 0.7 | 0.7 | 0.7 |
| ISR/FSR | 4.3 | 4.3 | — | — | — | — | — | — | — | — | 0.8 | 0.8 | 0.8 |
| b-jet frag. | 2.0 | 2.0 | — | — | — | 2.0 | — | — | — | — | 2.0 | 2.0 | 2.0 |
| Taggability | 3.5 | 3.2 | 3.8 | 3.7 | 4.4 | 3.8 | 3.4 | 3.6 | 3.4 | — | 3.5 | 3.9 | 3.6 |
| $\lambda^W_{\mathrm{HF}}$ | — | — | 12.0 | 12.0 | — | — | — | — | — | — | — | — | — |
| $\lambda^Z_{\mathrm{HF}}$ | — | — | — | — | — | 12.0 | 12.0 | — | — | — | — | — | — |
| IKS | — | — | 1.8 | 1.8 | 1.8 | — | — | — | — | −34.9 | — | — | — |

Table A.2: Uncertainties requiring exactly two tags and two jets.

UNCERTAINTIES FOR CHANNELS WITH THREE JETS

SINGLE TAG

| | Percentage Errors | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $t\bar{t}ll$ | $t\bar{t}lj$ | $Wbb$ | $Wcc$ | $Wlp$ | $Zbb$ | $Zcc$ | $Zlp$ | dibosons | multijet | $tb$ | $tqb$ | $tb+tq$ |
| Luminosity | 6.1 | 6.1 | — | — | — | 6.1 | 6.1 | 6.1 | 6.1 | — | 6.1 | 6.1 | 6.1 |
| Xsect. | 9.0 | 9.0 | — | — | — | 3.3 | 3.3 | 3.3 | 7.0 | — | 3.8 | 5.3 | 4.8 |
| Branching frac. | 1.5 | 1.5 | — | — | — | — | — | — | — | — | 1.5 | 1.5 | 1.5 |
| PDF | — | — | — | — | — | — | — | — | — | — | 2.0 | 2.0 | 2.0 |
| Triggers | 5.0 | 5.0 | — | — | — | 5.0 | 5.0 | 5.0 | 5.0 | — | 5.0 | 5.0 | 5.0 |
| Lumi. rewtg. | 1.0 | 1.0 | — | — | — | 1.0 | 1.0 | 1.0 | 1.0 | — | 1.0 | 1.0 | 1.0 |
| Prim. vertex | 1.4 | 1.4 | — | — | — | 1.4 | 1.4 | 1.4 | 1.4 | — | 1.4 | 1.4 | 1.4 |
| Color Reconnection | 1.0 | 1.0 | — | — | — | — | — | — | — | — | 1.0 | 1.0 | 1.0 |
| b/light Jet Response | 0.0 | 0.0 | — | — | — | 0.6 | — | — | — | — | 0.8 | 0.8 | 0.8 |
| Lepton ID | 2.8 | 2.8 | — | — | — | 2.5 | 2.5 | 2.7 | 2.7 | — | 2.7 | 2.8 | 2.8 |
| Jet frag. and higher order | 1.7 | 1.7 | — | — | — | — | — | — | — | — | 3.7 | 3.7 | 3.7 |
| ISR/FSR | 2.2 | 2.2 | — | — | — | — | — | — | — | — | −5.9 | −5.9 | −5.9 |
| b-jet frag. | 2.0 | 2.0 | — | — | — | 2.0 | — | — | — | — | 2.0 | 2.0 | 2.0 |
| Taggability | 13.3 | 8.6 | 7.9 | 8.0 | 8.8 | 9.4 | 7.7 | 14.8 | 8.7 | — | 9.3 | 8.6 | 8.9 |
| $\lambda^W_{\mathrm{HF}}$ | — | — | 12.0 | 12.0 | — | — | — | — | — | — | — | — | — |
| $\lambda^Z_{\mathrm{HF}}$ | — | — | — | — | — | 12.0 | 12.0 | — | — | — | — | — | — |
| IKS | — | — | 1.8 | 1.8 | 1.8 | — | — | — | — | −36.2 | — | — | — |

Table A.3: Uncertainties requiring exactly one tag and three jets.

DOUBLE TAG

| | Percentage Errors | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $t\bar{t}ll$ | $t\bar{t}lj$ | $Wbb$ | $Wcc$ | $Wlp$ | $Zbb$ | $Zcc$ | $Zlp$ | dibosons | multijet | $tb$ | $tqb$ | $tb+tqb$ |
| Luminosity | 6.1 | 6.1 | — | — | — | 6.1 | 6.1 | 6.1 | 6.1 | — | 6.1 | 6.1 | 6.1 |
| Xsect. | 9.0 | 9.0 | — | — | — | 3.3 | 3.3 | 3.3 | 7.0 | — | 3.8 | 5.3 | 4.8 |
| Branching frac. | 1.5 | 1.5 | — | — | — | — | — | — | — | — | 1.5 | 1.5 | 1.5 |
| PDF | — | — | — | — | — | — | — | — | — | — | 2.0 | 2.0 | 2.0 |
| Triggers | 5.0 | 5.0 | — | — | — | 5.0 | 5.0 | 5.0 | 5.0 | — | 5.0 | 5.0 | 5.0 |
| Lumi. rewtg. | 1.0 | 1.0 | — | — | — | 1.0 | 1.0 | 1.0 | 1.0 | — | 1.0 | 1.0 | 1.0 |
| Prim. vertex | 1.4 | 1.4 | — | — | — | 1.4 | 1.4 | 1.4 | 1.4 | — | 1.4 | 1.4 | 1.4 |
| Color Reconnection | 1.0 | 1.0 | — | — | — | — | — | — | — | — | 1.0 | 1.0 | 1.0 |
| b/light Jet Response | 0.0 | 0.0 | — | — | — | 0.6 | — | — | — | — | 0.8 | 0.8 | 0.8 |
| Lepton ID | 2.8 | 2.8 | — | — | — | 2.5 | 2.4 | 2.2 | 2.7 | — | 2.8 | 2.8 | 2.8 |
| Jet frag. and higher order | 1.7 | 1.7 | — | — | — | — | — | — | — | — | 3.7 | 3.7 | 3.7 |
| ISR/FSR | 2.2 | 2.2 | — | — | — | — | — | — | — | — | −5.9 | −5.9 | −5.9 |
| b-jet frag. | 2.0 | 2.0 | — | — | — | 2.0 | — | — | — | — | 2.0 | 2.0 | 2.0 |
| Taggability | 9.3 | 5.9 | 6.8 | 6.6 | 6.7 | 6.9 | 7.2 | 6.3 | 8.4 | — | 6.3 | 6.4 | 6.4 |
| $\lambda^W_{\mathrm{HF}}$ | — | — | 12.0 | 12.0 | — | — | — | — | — | — | — | — | — |
| $\lambda^Z_{\mathrm{HF}}$ | — | — | — | — | — | 12.0 | 12.0 | — | — | — | — | — | — |
| IKS | — | — | 1.8 | 1.8 | 1.8 | — | — | — | — | −36.0 | — | — | — |

Table A.4: Uncertainties requiring exactly two tag and three jets.

UNCERTAINTIES FOR CHANNELS WITH FOUR JETS

SINGLE TAG

| | Percentage Errors | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $t\bar{t}ll$ | $t\bar{t}lj$ | $Wbb$ | $Wcc$ | $Wlp$ | $Zbb$ | $Zcc$ | $Zlp$ | dibosons | multijet | $tb$ | $tqb$ | $tb+tqb$ |
| Luminosity | 6.1 | 6.1 | — | — | — | 6.1 | 6.1 | 6.1 | 6.1 | — | 6.1 | 6.1 | 6.1 |
| Xsect. | 9.0 | 9.0 | — | — | — | 3.3 | 3.3 | 3.3 | 7.0 | — | 3.8 | 5.3 | 4.8 |
| Branching frac. | 1.5 | 1.5 | — | — | — | — | — | — | — | — | 1.5 | 1.5 | 1.5 |
| PDF | — | — | — | — | — | — | — | — | — | — | 2.0 | 2.0 | 2.0 |
| Triggers | 5.0 | 5.0 | — | — | — | 5.0 | 5.0 | 5.0 | 5.0 | — | 5.0 | 5.0 | 5.0 |
| Lumi. rewtg. | 1.0 | 1.0 | — | — | — | 1.0 | 1.0 | 1.0 | 1.0 | — | 1.0 | 1.0 | 1.0 |
| Prim. vertex | 1.4 | 1.4 | — | — | — | 1.4 | 1.4 | 1.4 | 1.4 | — | 1.4 | 1.4 | 1.4 |
| Color Reconnection | 1.0 | 1.0 | — | — | — | — | — | — | — | — | 1.0 | 1.0 | 1.0 |
| b/light Jet Response | 0.3 | 0.3 | — | — | — | 0.8 | — | — | — | — | 0.9 | 0.9 | 0.9 |
| Lepton ID | 2.8 | 2.8 | — | — | — | 2.6 | 2.5 | 2.4 | 2.7 | — | 2.7 | 2.8 | 2.8 |
| Jet frag. and higher order | −7.0 | −7.0 | — | — | — | — | — | — | — | — | 4.7 | 4.7 | 4.7 |
| ISR/FSR | 0.8 | 0.8 | — | — | — | — | — | — | — | — | −10.9 | −10.9 | −10.9 |
| b-jet frag. | 2.0 | 2.0 | — | — | — | 2.0 | — | — | — | — | 2.0 | 2.0 | 2.0 |
| Taggability | 15.9 | 11.8 | 14.9 | 9.2 | 14.4 | 10.6 | 13.4 | 10.7 | 9.5 | — | 12.6 | 12.1 | 12.2 |
| $\lambda^W_{\mathrm{HF}}$ | — | — | 12.0 | 12.0 | — | — | — | — | — | — | — | — | — |
| $\lambda^Z_{\mathrm{HF}}$ | — | — | — | — | — | 12.0 | 12.0 | — | — | — | — | — | — |
| IKS | — | — | 1.8 | 1.8 | 1.8 | — | — | — | — | −33.8 | — | — | — |

Table A.5: Uncertainties requiring exactly one tag and four jets.

DOUBLE TAG

| | Percentage Errors | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $t\bar{t}ll$ | $t\bar{t}lj$ | $Wbb$ | $Wcc$ | $Wlp$ | $Zbb$ | $Zcc$ | $Zlp$ | dibosons | multijet | $tb$ | $tqb$ | $tb+tqb$ |
| Luminosity | 6.1 | 6.1 | — | — | — | 6.1 | 6.1 | 6.1 | 6.1 | — | 6.1 | 6.1 | 6.1 |
| Xsect. | 9.0 | 9.0 | — | — | — | 3.3 | 3.3 | 3.3 | 7.0 | — | 3.8 | 5.3 | 4.8 |
| Branching frac. | 1.5 | 1.5 | — | — | — | — | — | — | — | — | 1.5 | 1.5 | 1.5 |
| PDF | — | — | — | — | — | — | — | — | — | — | 2.0 | 2.0 | 2.0 |
| Triggers | 5.0 | 5.0 | — | — | — | 5.0 | 5.0 | 5.0 | 5.0 | — | 5.0 | 5.0 | 5.0 |
| Lumi. rewtg. | 1.0 | 1.0 | — | — | — | 1.0 | 1.0 | 1.0 | 1.0 | — | 1.0 | 1.0 | 1.0 |
| Prim. vertex | 1.4 | 1.4 | — | — | — | 1.4 | 1.4 | 1.4 | 1.4 | — | 1.4 | 1.4 | 1.4 |
| Color Reconnection | 1.0 | 1.0 | — | — | — | — | — | — | — | — | 1.0 | 1.0 | 1.0 |
| b/light Jet Response | 0.3 | 0.3 | — | — | — | 0.8 | — | — | — | — | 0.9 | 0.9 | 0.9 |
| Lepton ID | 2.8 | 2.8 | — | — | — | 2.6 | 2.6 | 2.4 | 2.7 | — | 2.8 | 2.8 | 2.8 |
| Jet frag. and higher order | −7.0 | −7.0 | — | — | — | — | — | — | — | — | 4.7 | 4.7 | 4.7 |
| ISR/FSR | 0.8 | 0.8 | — | — | — | — | — | — | — | — | −10.9 | −10.9 | −10.9 |
| b-jet frag. | 2.0 | 2.0 | — | — | — | 2.0 | — | — | — | — | 2.0 | 2.0 | 2.0 |
| Taggability | 11.7 | 8.2 | 9.8 | 6.4 | 8.8 | 9.7 | 18.4 | 21.1 | 7.7 | — | 8.6 | 14.1 | 11.7 |
| $\lambda^W_{\mathrm{HF}}$ | — | — | 12.0 | 12.0 | — | — | — | — | — | — | — | — | — |
| $\lambda^Z_{\mathrm{HF}}$ | — | — | — | — | — | 12.0 | 12.0 | — | — | — | — | — | — |
| IKS | — | — | 1.8 | 1.8 | 1.8 | — | — | — | — | −34.3 | — | — | — |

Table A.6: Uncertainties requiring exactly two tag and four jets.

# Appendix B

# Selected BNN Input Variables

**Selected BNN Input Variables for Cross Section Measurement Analysis**

| Analysis Channel: 1 b-tag 2 jets | | |
|---|---|---|
| Rank | Variable | K-S |
| 1 | LeadingLightQuarkJetBTagNN | 1.000 |
| 2 | LeadingLightQuarkJetEta | 0.987 |
| 3 | LeptonEta | 0.984 |
| 4 | LeadingBTaggedJetLeptonDeltaPhi | 0.974 |
| 5 | LeadingBTaggedJetPt | 0.969 |
| 6 | METPt | 0.941 |
| 7 | LeptonPt | 0.849 |
| 8 | LeadingLightQuarkJetLeptonDeltaPhi | 0.693 |
| 9 | WTransverseMass | 0.558 |
| 10 | QTimesEta | 0.520 |
| 11 | DeltaPhiLeptonMET | 0.488 |
| 12 | LeadingBTaggedJetBTagNN | 0.440 |
| 13 | LeadingBTaggedJetEta | 0.378 |
| 14 | LeadingLightQuarkJetPt | 0.176 |

Table B.1: Set of selected variables, ordered by their KS values, for the 1 b-tag 2 jets channel.

| Analysis Channel: 1 b-tag 3 jets | | |
|---|---|---|
| Rank | Variable | K-S |
| 1 | SecondLightQuarkJetBTagNN | 1.0000 |
| 2 | LeptonPt | 0.990 |
| 3 | LeadingLightQuarkJetBTagNN | 0.939 |
| 4 | LeptonEta | 0.886 |
| 5 | LeadingLightQuarkJetLeptonDeltaPhi | 0.865 |
| 6 | LeadingBTaggedJetPt | 0.799 |
| 7 | DeltaPhiLeptonMET | 0.797 |
| 8 | LeadingLightQuarkJetPt | 0.787 |
| 9 | LeadingBTaggedJetLeptonDeltaPhi | 0.702 |
| 10 | LeadingBTaggedJetEta | 0.647 |
| 11 | METPt | 0.585 |
| 12 | SecondLightQuarkJetEta | 0.554 |
| 13 | SecondLightQuarkJetLeptonDeltaPhi | 0.524 |
| 14 | WTransverseMass | 0.441 |
| 15 | SecondLightQuarkJetPt | 0.417 |
| 16 | LeadingLightQuarkJetEta | 0.416 |
| 17 | QTimesEta | 0.143 |
| 18 | LeadingBTaggedJetBTagNN | 0.139 |

Table B.2: Set of selected variables, ordered by their KS values, for the 1 b-tag 3 jets channel.

| Analysis Channel: 1 b-tag 4 jets | | |
|---|---|---|
| Rank | Variable | K-S |
| 1 | LeadingLightQuarkJetBTagNN | 1.000 |
| 2 | LeadingBTaggedJetEta | 0.995 |
| 3 | LeadingBTaggedJetPt | 0.991 |
| 4 | Jet4BTagNN | 0.990 |
| 5 | WTransverseMass | 0.940 |
| 6 | LeadingLightQuarkJetEta | 0.939 |
| 7 | SecondLightQuarkJetLeptonDeltaPhi | 0.915 |
| 8 | LeptonPt | 0.837 |
| 9 | LeadingBTaggedJetBTagNN | 0.822 |
| 10 | SecondLightQuarkJetBTagNN | 0.820 |
| 11 | SecondLightQuarkJetPt | 0.789 |
| 12 | LeadingLightQuarkJetPt | 0.778 |
| 13 | DeltaPhiLeptonMET | 0.759 |
| 14 | LeadingLightQuarkJetLeptonDeltaPhi | 0.712 |
| 15 | Jet4Eta | 0.615 |
| 16 | SecondLightQuarkJetEta | 0.592 |
| 17 | Jet4Phi | 0.559 |
| 18 | LeadingBTaggedJetLeptonDeltaPhi | 0.317 |
| 19 | QTimesEta | 0.259 |
| 20 | METPt | 0.215 |
| 21 | LeptonEta | 0.168 |
| 22 | Jet4Pt | 0.105 |

Table B.3: Set of selected variables, ordered by their KS values, for the 1 b-tag 4 jets channel.

| Analysis Channel: 2 b-tag 2 jets | | |
|---|---|---|
| Rank | Variable | K-S |
| 1 | LeptonPt | 0.994 |
| 2 | SecondBTaggedJetEta | 0.992 |
| 3 | SecondBTaggedJetPt | 0.981 |
| 4 | LeadingBTaggedJetLeptonDeltaPhi | 0.890 |
| 5 | DeltaPhiLeptonMET | 0.842 |
| 6 | LeadingBTaggedJetEta | 0.839 |
| 7 | SecondBBTaggedJetLeptonDeltaPhi | 0.805 |
| 8 | METPt | 0.792 |
| 9 | LeadingBTaggedJetPt | 0.739 |
| 10 | LeptonEta | 0.645 |
| 11 | WTransverseMass | 0.501 |
| 12 | LeadingBTaggedJetBTagNN | 0.461 |
| 13 | SecondBTaggedJetBTagNN | 0.296 |

Table B.4: Set of selected variables, ordered by their KS values, for the 2 b-tag 2 jets channel.

| Analysis Channel: 2 b-tag 3 jets | | |
|---|---|---|
| Rank | Variable | K-S |
| 1 | SecondBTaggedJetLeptonDeltaPhi | 1.000 |
| 2 | LeadingLightQuarkJetEta | 1.000 |
| 3 | LeadingLightQuarkJetBTagNN | 1.000 |
| 4 | QTimesEta | 0.999 |
| 5 | LeadingLightQuarkJetLeptonDeltaPhi | 0.950 |
| 6 | LeptonEta | 0.938 |
| 7 | LeadingLightQuarkJetPt | 0.938 |
| 8 | SecondBTaggedJetEta | 0.908 |
| 9 | SecondBTaggedJetBTagNN | 0.888 |
| 10 | WTransverseMass | 0.880 |
| 11 | LeadingBTaggedJetEta | 0.862 |
| 12 | SecondBTaggedJetPt | 0.844 |
| 13 | LeadingBTaggedJetPt | 0.714 |
| 14 | LeptonPt | 0.505 |
| 15 | DeltaPhiLeptonMET | 0.438 |
| 16 | METPt | 0.405 |
| 17 | LeadingBTaggedJetBTagNN | 0.343 |
| 18 | LeadingBTaggedJetLeptonDeltaPhi | 0.164 |

Table B.5: Set of selected variables, ordered by their KS values, for the 2 b-tag 3 jets channel.

| Analysis Channel: 2 b-tag 4 jets | | |
|---|---|---|
| Rank | Variable | K-S |
| 1 | SecondBTaggedJetLeptonDeltaPhi | 1.000 |
| 2 | SecondBTaggedJetBTagNN | 1.000 |
| 3 | LeadingLightQuarkJetEta | 1.000 |
| 4 | LeadingLightQuarkJetBTagNN | 1.000 |
| 5 | LeadingBTaggedJetLeptonDeltaPhi | 0.999 |
| 6 | QTimesEta | 0.998 |
| 7 | SecondLightQuarkJetBTagNN | 0.996 |
| 8 | LeadingBTaggedJetPt | 0.996 |
| 9 | LeadingBTaggedJetEta | 0.981 |
| 10 | LeptonPt | 0.976 |
| 11 | SecondLightQuarkJetLeptonDeltaPhi | 0.937 |
| 12 | LeadingBTaggedJetBTagNN | 0.920 |
| 13 | METPt | 0.910 |
| 14 | DeltaPhiLeptonMET | 0.819 |
| 15 | LeadingLightQuarkJetPt | 0.732 |
| 16 | SecondBTaggedJetPt | 0.709 |
| 17 | SecondBTaggedJetEta | 0.704 |
| 18 | LeadingLightQuarkJetLeptonDeltaPhi | 0.644 |
| 19 | SecondLightQuarkJetEta | 0.505 |
| 20 | WTransverseMass | 0.369 |
| 21 | LeptonEta | 0.319 |
| 22 | SecondLightQuarkJetPt | 0.184 |

Table B.6: Set of selected variables, ordered by their KS values, for the 2 b-tag 4 jets channel.

**Selected BNN Input Variables for Anomalous $Wtb$ Coupling Analysis**

| Analysis Channel: 1 b-tag 2 jets | | |
|---|---|---|
| Rank | Variable | K-S |
| 1 | LeadingLightQuarkJetBTagNN | 1.000 |
| 2 | LeadingLightQuarkJetEta | 0.987 |
| 3 | LeptonEta | 0.984 |
| 4 | LeadingBTaggedJetLeptonDeltaPhi | 0.974 |
| 5 | LeadingBTaggedJetPt | 0.969 |
| 6 | CosBTaggedJetLeptonBTaggedTop | 0.960 |
| 7 | METPt | 0.941 |
| 8 | LeptonPt | 0.849 |
| 9 | LeadingLightQuarkJetLeptonDeltaPhi | 0.693 |
| 10 | CosLeptonQZBestTop | 0.659 |
| 11 | WTransverseMass | 0.558 |
| 12 | QTimesEta | 0.520 |
| 13 | DeltaPhiLeptonMET | 0.488 |
| 14 | LeadingBTaggedJetBTagNN | 0.440 |
| 15 | LeadingBTaggedJetEta | 0.378 |
| 16 | CosLeptonBTaggedTopFrameBTaggedTopCMFrame | 0.178 |
| 17 | CosLightQuarkJetLeptonBTaggedTop | 0.177 |
| 18 | LeadingLightQuarkJetPt | 0.176 |

Table B.7: Set of selected variables, ordered by their KS values, for the 1 b-tag 2 jets channel.

| Analysis Channel: 1 b-tag 3 jets | | |
|---|---|---|
| Rank | Variable | K-S |
| 1 | SecondLightQuarkJetBTagNN | 1.0000 |
| 2 | LeptonPt | 0.990 |
| 3 | CosLeptonBTaggedTopFrameBTaggedTopCMFrame | 0.986 |
| 4 | CosBTaggedJetLeptonBTaggedTop | 0.941 |
| 5 | LeadingLightQuarkJetBTagNN | 0.939 |
| 6 | LeptonEta | 0.886 |
| 7 | LeadingLightQuarkJetLeptonDeltaPhi | 0.865 |
| 8 | LeadingBTaggedJetPt | 0.799 |
| 9 | DeltaPhiLeptonMET | 0.797 |
| 10 | LeadingLightQuarkJetPt | 0.787 |
| 11 | LeadingBTaggedJetLeptonDeltaPhi | 0.702 |
| 12 | LeadingBTaggedJetEta | 0.647 |
| 13 | METPt | 0.585 |
| 14 | CosLightQuarkJetLeptonBTaggedTop | 0.555 |
| 15 | SecondLightQuarkJetEta | 0.554 |
| 16 | SecondLightQuarkJetLeptonDeltaPhi | 0.524 |
| 17 | WTransverseMass | 0.441 |
| 18 | CosLeptonQZBestTop | 0.268 |
| 19 | SecondLightQuarkJetPt | 0.417 |
| 20 | LeadingLightQuarkJetEta | 0.416 |
| 21 | QTimesEta | 0.143 |
| 22 | LeadingBTaggedJetBTagNN | 0.139 |

Table B.8: Set of selected variables, ordered by their KS values, for the 1 b-tag 3 jets channel.

| Analysis Channel: 1 b-tag 4 jets | | |
|---|---|---|
| Rank | Variable | K-S |
| 1 | LeadingLightQuarkJetBTagNN | 1.000 |
| 2 | LeadingBTaggedJetEta | 0.995 |
| 3 | LeadingBTaggedJetPt | 0.991 |
| 4 | Jet4BTagNN | 0.990 |
| 5 | CosLeptonQZBestTop | 0.963 |
| 6 | WTransverseMass | 0.940 |
| 7 | LeadingLightQuarkJetEta | 0.939 |
| 8 | CosLightQuarkJetLeptonBTaggedTop | 0.934 |
| 9 | SecondLightQuarkJetLeptonDeltaPhi | 0.915 |
| 10 | LeptonPt | 0.837 |
| 11 | LeadingBTaggedJetBTagNN | 0.822 |
| 12 | SecondLightQuarkJetBTagNN | 0.820 |
| 13 | SecondLightQuarkJetPt | 0.789 |
| 14 | LeadingLightQuarkJetPt | 0.778 |
| 15 | DeltaPhiLeptonMET | 0.759 |
| 16 | CosLeptonBTaggedTopFrameBTaggedTopCMFrame | 0.729 |
| 17 | LeadingLightQuarkJetLeptonDeltaPhi | 0.712 |
| 18 | Jet4Eta | 0.615 |
| 19 | SecondLightQuarkJetEta | 0.592 |
| 20 | Jet4Phi | 0.559 |
| 21 | CosBTaggedJetLeptonBTaggedTop | 0.328 |
| 22 | LeadingBTaggedJetLeptonDeltaPhi | 0.317 |
| 23 | QTimesEta | 0.259 |
| 24 | METPt | 0.215 |
| 25 | LeptonEta | 0.168 |
| 26 | Jet4Pt | 0.105 |

Table B.9: Set of selected variables, ordered by their KS values, for the 1 b-tag 4 jets channel.

| Analysis Channel: 2 b-tag 2 jets | | |
|---|---|---|
| Rank | Variable | K-S |
| 1 | LeptonPt | 0.994 |
| 2 | SecondBTaggedJetEta | 0.992 |
| 3 | SecondBTaggedJetPt | 0.981 |
| 4 | LeadingBTaggedJetLeptonDeltaPhi | 0.890 |
| 5 | DeltaPhiLeptonMET | 0.842 |
| 6 | LeadingBTaggedJetEta | 0.839 |
| 7 | SecondBBTaggedJetLeptonDeltaPhi | 0.805 |
| 8 | CosBTaggedJetLeptonBTaggedTop | 0.798 |
| 9 | METPt | 0.792 |
| 10 | LeadingBTaggedJetPt | 0.739 |
| 11 | CosLeptonBTaggedTopFrameBTaggedTopCMFrame | 0.725 |
| 12 | LeptonEta | 0.645 |
| 13 | WTransverseMass | 0.501 |
| 14 | LeadingBTaggedJetBTagNN | 0.461 |
| 15 | SecondBTaggedJetBTagNN | 0.296 |
| 16 | CosLeptonQZBestTop | 0.249 |

Table B.10: Set of selected variables, ordered by their KS values, for the 2 b-tag 2 jets channel.

| Analysis Channel: 2 b-tag 3 jets | | |
|---|---|---|
| Rank | Variable | K-S |
| 1 | SecondBTaggedJetLeptonDeltaPhi | 1.000 |
| 2 | LeadingLightQuarkJetEta | 1.000 |
| 3 | LeadingLightQuarkJetBTagNN | 1.000 |
| 4 | QTimesEta | 0.999 |
| 5 | CosLightQuarkJetLeptonBTaggedTop | 0.965 |
| 6 | LeadingLightQuarkJetLeptonDeltaPhi | 0.950 |
| 7 | CosLeptonBTaggedTopFrameBTaggedTopCMFrame | 0.938 |
| 8 | LeptonEta | 0.938 |
| 9 | LeadingLightQuarkJetPt | 0.938 |
| 10 | SecondBTaggedJetEta | 0.908 |
| 11 | SecondBTaggedJetBTagNN | 0.888 |
| 12 | WTransverseMass | 0.880 |
| 13 | LeadingBTaggedJetEta | 0.862 |
| 14 | SecondBTaggedJetPt | 0.844 |
| 15 | CosBTaggedJetLeptonBTaggedTop | 0.759 |
| 16 | LeadingBTaggedJetPt | 0.714 |
| 17 | CosLeptonQZBestTop | 0.559 |
| 18 | LeptonPt | 0.505 |
| 19 | DeltaPhiLeptonMET | 0.438 |
| 20 | METPt | 0.405 |
| 21 | LeadingBTaggedJetBTagNN | 0.343 |
| 22 | LeadingBTaggedJetLeptonDeltaPhi | 0.164 |

Table B.11: Set of selected variables, ordered by their KS values, for the 2 b-tag 3 jets channel.

| Analysis Channel: 2 b-tag 4 jets | | |
|---|---|---|
| Rank | Variable | K-S |
| 1 | SecondBTaggedJetLeptonDeltaPhi | 1.000 |
| 2 | SecondBTaggedJetBTagNN | 1.000 |
| 3 | LeadingLightQuarkJetEta | 1.000 |
| 4 | LeadingLightQuarkJetBTagNN | 1.000 |
| 5 | CosLeptonQZBestTop | 1.000 |
| 6 | LeadingBTaggedJetLeptonDeltaPhi | 0.999 |
| 7 | QTimesEta | 0.998 |
| 8 | SecondLightQuarkJetBTagNN | 0.996 |
| 9 | LeadingBTaggedJetPt | 0.996 |
| 10 | CosBTaggedJetLeptonBTaggedTop | 0.989 |
| 11 | LeadingBTaggedJetEta | 0.981 |
| 12 | LeptonPt | 0.976 |
| 13 | SecondLightQuarkJetLeptonDeltaPhi | 0.937 |
| 14 | LeadingBTaggedJetBTagNN | 0.920 |
| 15 | METPt | 0.910 |
| 16 | CosLeptonBTaggedTopFrameBTaggedTopCMFrame | 0.879 |
| 17 | DeltaPhiLeptonMET | 0.819 |
| 18 | LeadingLightQuarkJetPt | 0.732 |
| 19 | CosLightQuarkJetLeptonBTaggedTop | 0.727 |
| 20 | SecondBTaggedJetPt | 0.709 |
| 21 | SecondBTaggedJetEta | 0.704 |
| 22 | LeadingLightQuarkJetLeptonDeltaPhi | 0.644 |
| 23 | SecondLightQuarkJetEta | 0.505 |
| 24 | WTransverseMass | 0.369 |
| 25 | LeptonEta | 0.319 |
| 26 | SecondLightQuarkJetPt | 0.184 |

Table B.12: Set of selected variables, ordered by their KS values, for the 2 b-tag 4 jets channel.

### Description of the Input Variables

The naming convention used for the variables in both the analyses is as follows. Jets are ordered in $p_T$ such that "jet1" is the jet with the largest $p_T$ , "jet2" has the second largest and so forth. In the samples with exactly one tag, the "leading tagged jet" refers to the only jet passing the $b$-tagging criteria. In the samples with exactly two tags, the "leading tagged jet" is the jet passing the $b$-tagging criteria which has the largest $p_T$ , the "second tagged jet" is the other tagged jet. The "leading light quark jet" and "second light quark jet" are the untagged jets ranked in $p_T$ order.

The definition of four additional angular variables used in the anomalous $Wtb$ coupling analysis is :

**Cos BTaggedJetLepton BTaggedTop**: Cosine between leading tagged jet and the lepton in the tagged top frame.

**Cos LeptonQZ BestTop**: Cosine between the lepton and the (z-axis)*(lepton charge) in the best top frame.

**Cos LeptonBTaggedTopFrame BTaggedTopCMFrame**: Cosine between lepton (in tagged top frame) and the tagged top (in the CM frame) where the top quark is constructed from the leading tagged jet and $W$.

**Cos LightQuarkJetLepton BTaggedTop**: Cosine between leading light quark jet and the lepton in the tagged top frame.

# Bibliography


[1] M. Veltman, "Facts and Mysteries in Elementary Particle Physics", World Scientific (2003). 2.1

[2] D. Griffiths, "Introduction to Elementary Particles", John Wiley Sons (1987). 2.1

[3] D. H. Perkins, "Introduction to High Energy Physics," Addison-Wesley Publishing Company (1972). 2.1

[4] S. B. Beri, "Concepts of High Energy and Nuclear Physics," Elite Publishing House (2007). 2.1

[5] F. Halzen and A. Martin, "Quarks and Leptons", John Wiley Sons (1984). 2.1, 2.1.2

[6] C. Quigg, "Gauge Theories of the Strong, Weak, and Electromagnetic Interactions", Westview Press (1983). 2.1

[7] R. Ellis *et al.*, "QCD and Collider Physics", Cambridge University Press (1996). 2.1

[8] H. B. Prosper and M. Danilov, "Techniques and Concepts of High-Energy Physics XII", NATO Science Series (2003). 2.1

[9] K. Nakamura *et al.*, [Particle Data Group] J. Phys. G 37, 075021 (2010). 2.1.1, 2.1.2, 5.6

[10] S. Eidelman *et al.*, "Review of Particle Physics," Phys. Lett. B 592, 1 (2004). 2.1.2

[11] M. Peskin, D. Schroeder, "An Introduction to Quantum Field Theory", Westview Press (1995). 2.1.3

[12] P. A. M. Dirac, "The quantum theory of dispersion", Proc. R. Soc., London A114, 243, 710 (1927). 2.1.3

[13] R. P. Feynman, "QED : The Strange Theory of Light and matter", Princeton University Press, Princeton, New Jersey (1985). 2.1.3

[14] R. Barate *et al.*, "LEP Working Group for Higgs boson searches", Phys. Lett. B 565, 61 (2003). 2.1.4

[15] B. W. Lee *et al.*, "Strength of Weak Interactions at Very High Energies and the Higgs Boson Mass", Phys. Rev. Lett. 38 883 (1977);
M. Quiros, "Constraints on the Higgs boson properties from the effective potential", [hep-ph/9703412] (1997);
A. Ghinculov and T. Binoth, "Loops and legs beyond perturbation theory", Acta Phys. Polon. B30 99 (1999). 2.1.4

[16] L. Maiani, G. Parisi and R. Petronzio, Nucl. Phys. B 136 115 (1979);
N. Cabibbo *et al.*, Nucl. Phys. B 158 295 (1979);
R. Dashen and H. Neunberger, Phys. Rev. Lett. 50 1897 (1983);
D. J. E. Callaway, Nucl. Phys. B 233 189 (1984);
M. A. Beg *et al.*, Phys. Rev. Lett. 52 883 (1984);
M. Lindner, Z. Phys. C31 295 (1986). 2.1.4

[17] G. Altarelli and G. Isidori, Phys. Lett. B 337 141 (1994);
J. A. Casas, J. R. Espinosa and M. Quiros, Phys. Lett. B 342 171 (1995), Phys. Lett. B 383 374 (1996);
B. Grzadkowski and M. Lindner, Phys. Lett. B 178 81 (1986);
T. Hambye and K. Riesselmann, Phys. Rev. D 55 7255 (1997). 2.1.4

[18] The CDF, DØ Collaborations, the TEVNPHWG Working Group, "Combined CDF and DØ Upper Limits on Standard Model Higgs Boson Production with up to 8.6 $fb^{-1}$ of Data", FERMILAB-CONF-11-354-E, arXiv:1107.5518v2 [hep-ex] (2011). 2.1.4

[19] 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 (1977). 2.2

[20] S. Abachi, *et al.*, "Observation of the Top Quark," Phys. Rev. Lett. 74, 2632-2637 (1995). 2.2, 2.2.1, 3

[21] F. Abe *et al.*, (CDF Collaboration), "Observation of Top Quark Production in $p\bar{p}$ Collisions with the Collider Detector at Fermilab," Phys. Rev. Lett. 74, 2626 (1995). 2.2, 2.2.1, 3

[22] The CDF, DØ Collaborations, the TEVEWWG Working Group, "Combination of CDF and DØ results on the mass of the top quark using up to 5.6 $fb^{-1}$ of data," FERMILAB-TM-2466-E, arXiv:1007.3178v1 [hep-ex] (2010). 2.2

[23] R. D. Pecci and X. Zhang, "Dynamical Symmetry Breaking and Universality Breakdown," Nucl. Phys. B 337, 269 (1990). 2.2

[24] K. G. Chetyrkin, R. Harlander, T. Seidensticker, and M. Steinhauser, "Second Order QCD Corrections to $\Gamma(t \to Wb)$," Phys. Rev. D 60, 114015 (1999). 2.2

[25] A. Denner and T. Sack, "The Top Width," Nucl. Phys. B 238, 46 (1991). 2.2

[26] I. I. Bigi, Y. L. Dokshitzer, V. Khoze, J. kuehn and P. Zerwas, Phys. Lett. B 181, 157 (1986). 2.2

[27] R. Bonciani, S. Catani, M. Mangano, and P. Nason, "NLL Resummation of the Heavy-Quark Hadroproduction Cross-Section," Nucl. Phys. B 529, 424 (1998). 2.2.1

[28] N. Kidonakis and R. Vogt, "Next-to-Next-to-Leading Order Soft-Gluon Corrections in Top Quark Hadroproduction," Phys. Rev. D 68, 114014 (2003). 2.2.1

[29] V. Avazov *et al.* (DØ Collaboration), "Measurement of the top quark pair production cross section in the lepton+jets channel in proton-antiproton collisions at $\sqrt{s}$=1.96 TeV," Phys. Rev. D 84, 012008 (2011). 2.2.1

[30] W. M. Yao *et al.*, [Particle Data Group] "Review of particle physics," J. Phys. G 33, 1 (2006). 2.2.1, 2.2.1

[31] V. M. Abazov *et al.* (DØ Collaboration), "Evidence for Production of Single Top Quarks and First Direct Measurement of $|V_{tb}|$", Phys. Rev. Lett. 98, 181802 (2007). 2.3

[32] V. M. Abazov *et al.* (DØ Collaboration), "Evidence for production of single top quarks," Phys. Rev. D 78, 012005 (2008). 2.3, 2.3.4, 7.2.1

[33] T. Aaltonen *et al.* (CDF Collaboration), "Measurement of the Single-Top-Quark Production Cross Section at CDF," Phys. Rev. Lett. 101, 252001 (2008). 2.3, 2.3.4

[34] A. Heinson, *et al.*, "Single top quarks at the Fermilab Tevatron," Phys. Rev. D 56, 3114-3128 (1997). 2.3, 8.1

[35] B. Abbott *et al.* (DØ Collaboration), "Search for Electroweak Production of Single Top Quarks in $p\bar{p}$ Collisions," Phys. Rev. D 63, 031101 (2001);
V. M. Abazov *et al.* (DØ Collaboration), "Search for Single Top Quark Production at DØ Using Neural Networks," Phys. Lett. B 517, 282 (2001). 2.3

[36] D. Acosta *et al.* (CDF Collaboration), "Optimized search for single top quark production at the Fermilab Tevatron," Phys. Rev. D 69, 052003 (2004);
D. Acosta *et al.* (CDF Collaboration), "Search for Electroweak Single-Top-Quark Production in $p\bar{p}$ Collisions at $\sqrt{s}$ = 1.96 TeV," Phys. Rev. D 71, 012005 (2005). 2.3

[37] M. T. Bowen, *et al.*, "In search of lonely top quarks at the Tevatron," Phys. Rev. D 72, 074016 (2005). 2.3

[38] T. M. P. Tait and C. P. Yuan, "Single top quark production as a window to physics beyond the standard model," Phys. Rev. D 63, 1014018 (2000). (document), 2.3, 2.3.1, 2.3.1, 2.13, 2.3.1, 2.3.1, 2.3.1, 2.3.2

[39] S. Cortese and R. Petronzio, Phys. Lett. B 253, 494 (1991). 2.3

[40] T. Stelzer and S. Willenbrock, Phys. Lett. B 357, 125 (1995). 2.3

[41] N. Kidonakis, "Single top quark production at the Fermilab Tevatron: Threshold resummation and finite-order soft gluon corrections," Phys. Rev. D. 74, 114012 (2006). 2.3, 2.3, 2.3, 5.6, 7.3

[42] T. Tait and C. -P. Yuan, "The Phenomenology of Single Top Quark Production at the Fermilab Tevatron," [hep-ph/9710372] (1997). 2.3, 2.3

[43] S. Mrenna and C. -P.Yuan, "Effects of QCD Resummation on $W^+h$ and $t\bar{b}$ Production at the Tevatron," Phys. Lett. B 416, 200-207 (1998). 2.3

[44] S. S. D. Willenbrock and D. A. Dicus, "Production of Heavy Quarks from W-Gluon Fusion," Phys. Rev. D 34, 155 (1986). 2.3

[45] C. -P. Yuan, "A New Method to Detect a Heavy Top Quark at the Fermilab Tevatron," Phys. Rev. D 41, 42 (1990). 2.3

[46] R. K. Ellis and S. J. Parke, "Top quark production by W gluon fusion," Phys. Rev. D 46, 3785 (1992). 2.3

[47] J. Campbell and F. Tramontano, "Next-to-leading order corrections to $Wt$ production and decay," Nucl. Phys B 726 (2005). 2.3

[48] Z. Sullivan, "Understanding single-top-quark production and jets at hadron colliders," Phys. Rev. D 70, 114012 (2004). 2.3, 5.2.1, 8.2

[49] N. Cabibbo, Phys. Rev. Lett. 10, 531 (1961);
M. Kobayashi and T. Maskawa, Prog. Theor. Phys. 49, 652 (1973). 2.3.1, 7.3, 8.1

[50] K. Nakamura *et al.*, [Particle Data Group] "The Review of Patricle Physics," Chapter 11, THE CKM QUARK-MIXING MATRIX, J. Phys. G 37, 075021 (2010). 2.3.1

[51] G. Mahlon and S. Parke, "Improved Spin Basis for Angular Correlation Studies in Single Top Quark Production at the Fermilab Tevatron," Phys. Rev. D 55, 7249 (1997). 2.3.1

[52] G. Mahlon and S. Parke, "Observing Spin Correlations in Single Top Production and Decay," Phys. Lett. B 476, 323 (2000). 2.3.1

[53] E. E. Boos and A. V. Sherstnev,"Spin Effects in Processes of Single Top Quark Production at Hadron Colliders," Phys. Lett. B 534, 97 (2002). 2.3.1

[54] D. O. Carlson and C. -P. Yuan, Phys. Lett. B 306, 386 (1993). 2.3.1

[55] D. O. Carlson, E. Malkawi and C. P. Yuan, "Probing the couplings of the Top Quark to Gauge Bosons," Phys. Lett. B 337, 145 (1994). 2.3.1, 8.1

[56] E. Boos, L. Dudko and T. Ohl, Eur. Phys. J. C 11, 473 (1999). 2.3.1

[57] C. -R. Chen, F. Larios and C. P. Yuan, Phys. Lett. B 631, 126 (2005). 2.3.1

[58] A. Datta, P. J. ODonnell, Z. H. Lin, X. Zhang and T. Huang, Phys. Lett. B 483, 203 (2000). 2.3.1

[59] V. M. Abazov *et al.* (DØ Collaboration),"Search for $W^{'}$ boson resonances decaying to a top quark and a bottom quark," Phys. Rev. Lett. 100, 211803 (2008). 2.3.1

[60] B. Abbott *et al.* (DØ Collaboration), "Search for Electroweak Production of Single Top Quarks in $p\bar{p}$ Collisions," Phys. Rev. D 63, 031101 (2001). 2.3.4

[61] V. M. Abazov *et al.* (DØ Collaboration), "Search for Single Top Quark Production at DØ Using Neural Netwoks," Phys. Lett. B 517, 282 (2001). 2.3.4

[62] D. Acosta *et al.* (CDF Collaboration), "Search for Single Top Quark Production in $p\bar{p}$ Collisions at $sqrts = 1.86TeV$," Phys. Rev. D 65, 091102 (2002). 2.3.4

[63] D. Acosta *et al.* (CDF Collaboration), "Optimized search for single-top-quark production at the Fermilab Tevatron," Phys. Rev. D 69, 052003 (2004). 2.3.4

[64] V. M. Abazov *et al.* (DØ Collaboration), "Search for single top quark production in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV," Phys. Lett. B 622, 265 (2005). 2.3.4

[65] V. M. Abazov *et al.* (DØ Collaboration), "Multivariate searches for single top quark production with the D0 detector.," Phys. Rev. D 75, 092007 (2007). 2.3.4

[66] D. Acosta *et al.* (CDF Collaboration), "Search for electroweak single-top-quark production in $p\bar{p}$ collisions at $sqrts$=1.96TeV," Phys. Rev. D 71, 012005 (2005). 2.3.4

[67] V. M. Abazov *et al.* (DØ Collaboration), "Observation of Single Top Quark Production," Phys. Rev. Lett. 103, 092001 (2009). 2.3.4, 7.2.5

[68] T. Aaltonen *et al.* (CDF Collaboration), "Observation of Electroweak Single Top-Quark Production," Phys. Rev. Lett. 103, 092002 (2009). 2.3.4, 7.2.5, 7.2.5

[69] V. M. Abazov *et al.* (DØ Collaboration), "Measurement of the $t$-channel single top quark production cross section," Phys. Rev. Lett. B 682, 363-369 (2010). 2.3.4

[70] V. M. Abazov *et al.* (DØ Collaboration), "Search for flavor changing neutral currents via quark-gluon couplings in single top quark production using 2.3 $fb^{-1}$ of $p\bar{p}$ collisions," Phys. Lett. B 693, 81-87 (2010). 2.3.4

[71] V. M. Abazov *et al.* (DØ Collaboration), "Search for $W^{'} \rightarrow tb$ resonances with left- and right-handed couplings to fermions," Phys. Lett. B 699, 145-150 (2011). 2.3.4

[72] L. M. Lederman, "The Tevatron," Scientific American, 264(3), 48 (1991). 3

[73] H. T. Edwards, "The Tevatron Energy Doubler: A Superconducting Accelerator," Ann. Rev. Nucl. Part. Sci. 35, 605 (1985). 3

[74] TeVI Group, "Design Report Tevatron I projects," Fermilab Technical Report FERMILAB-DESIGN-1984-01 (1984). 3

[75] F. T. Cole *et al.*, "A Report on the Design of the Fermi National Accelerator Laboratory Superconducting Accelerator," Beams Document 1888-v1 (1979). 3

[76] J. Thompson, "Introduction to Colliding Beams at Fermilab," Fermilab Technical Report FermiLab TM-1909 (1994). 3

[77] http://www-ad.fnal.gov/runII/index.html. 3.1

[78] http://www-bd.fnal.gov/public/chain.html. 3.1

[79] C. W. Schmidt and C. D. Curtis, "A 50 mA Negative Hydrogen-Ion Source," IEEE Transactions on Nuclear Science, NS-26, 4120 (1979). 3.1.1

[80] C. D. Curtis *et al.*, "Linac H-Beam Operation and Uses at FermiLab," IEEE Transactions on Nuclear Science, NS-26, 3760 (1979). 3.1.2

[81] E. L. Hubbard *et al.*, "Booster synchrotron," Fermilab Technical Report Fermilab TM-405 (1973). 3.1.3

[82] C. S. Mishra, "The Fermilab Main Injector," Fermilab Technical Report Fermilab Conf-92/372 (1992). 3.1.4

[83] S. Nagaitsev *et al.*, "Antiproton Cooling in the Fermilab Recycler Ring," Fermilab Technical Report Fermilab-Conf-05-550-AD (2005). 3.1.5

[84] T. LeCompte, H. T. Diehl, "The CDF and DØ Upgrades for Run II," Annu. Rev. Nucl. Part. Sci. 50, 71-117, 2000. 3.2

[85] S. Abachi *et al.* (DØ Collaboration), "The DØ Detector," Nucl. Inst. Methods A338, 185 (1994). 3.2

[86] S. Abachi *et al.* (DØ Collaboration), "The DØ Upgrade: The Detector and its Physics," Fermilab Technical Report Fermilab-Pub-96/357-E (1996). 3.2

[87] V. M. Abazov *et al.* (DØ Collaboration), "The Upgraded DØ Detector," Nucl. Inst. Methods A565, 463 (2006). 3.2, 3.2.4, 3.2.6

[88] R. Angstadt *et al.*, "The Layer 0 Inner Silicon Detector of the DØ Experiment," Nucl. Inst. and Methods A 622, 298-310 (2010). 3.2.2

[89] S. N. Ahmed *et al.*, "The DØ Silicon Microstrip Tracker," Nucl. Inst. and Methods A 634, 8-46 (2011). 3.2.2

[90] D. Adams *et al.*, "The DØ Upgrade: Central Fiber Tracker, Technical Design Report," DØ Note 4164 (2003). 3.2.2

[91] J. Brzezniak (DØ Collaboration), "Conceptual Design of a 2 Tesla Superconducting Solenoid for the Fermilab DØ Detector Upgrade," Fermilab Technical Report FERMILAB-TM-1886 (1994). 3.2.2

[92] M. Adams *et al.*, "Design Report of the Central Preshower Detector for the DØ Upgrade," DØ Note 3014 (1996). 3.2.3

[93] The DØ Collaboration, "The DØ Upgrade : Forward Preshower, Muon System and Level2 Trigger," DØ Note 2894 (1996). 3.2.3

[94] M. Abolins *et al.*, "The Run IIb Trigger Upgrade for the DØ Experiment," IEEE Trans. Nucl. Sci. 51, 340344 (2004). 3.2.7

[95] V. M. Abazov *et al.* (DØ Collaboration), "Run IIb Upgrade Technical Design Report," Fermilab Pub-02-327-E (2002). 3.2.7

[96] A. Khanov, "HTF: Histogramming Method for Finding Tracks. The Algorithm Description.," DØ Note 3778 (2000). 4.1, 4.1

[97] G. Borissov, "Ordering a Chaos or... Technical Details of AA Tracking," Available from http://www-d0.fnal.gov/global-tracking/talks/20030228/talk-adm-030228.ps (2003). 4.1

[98] R. Fruhwirth, "Application of Kalman Filtering to Track and Vertex Fitting," Nucl. Instrum. Methods A 262 444 (1987). 4.1

[99] H. Greenlee, "The DØ Kalman Track Fit", DØ Note 4303 (2004). 4.1, 4.2

[100] C. Tully and A. Schwartzman, "Primary vertex reconstruction by means of adaptive vertex fitting," DØ Note 4918 (2005). 4.2

[101] A. Schwartzman and M. Narain, "Probabilistic Primary Vertex Selection", DØ Note 4042 (2002). 4.2

[102] G. Bernardi, *et al.*, "Improvement of the NADA Algorithm: Hot Cell Killing in DØ Run II Data", DØ Note 4057 (2002). 4.3

[103] J. -R. Vlimant, *et al.*, "Technical description of the T42 Algorithm for the calorimeter noise suppression", DØ Note 4146 (2003). 4.3, 4.6

[104] A. Kumar *et al.*, "Electron Likelihood Study," DØ Note 4769 (2005). 4.4

[105] P. Calfayan *et al.*, "Muon Identification Certification for p17 data," DØ Note 5157 (2006). 4.5

[106] S. Trincaz-Duvoid and P. Verdier, "Missing $\not\!\!E_T$ reconstruction in p17," DØ Note 4474 (2004). 4.8

[107] G. C. Blazey *et al.*, "Run II Jet Physics," DØ Note 3750 (2000). 4.6, 4.6

[108] E. Busato and B. Andrieu, "Jet algorithms in the DØ Run II software: Description and users guide," DØ Note 4457 (2004). 4.6

[109] K. DeVaughan *et al.*, "Jet Energy Scale Determination for DØ RunIIb (final p20 version)," DØ Note 5801 (2008). 4.6

[110] N. Makovec, J. -F. Grivaz, "Shifting, Smearing and Removing Simulated Jets," DØ Note 4914 (2005). 4.6

[111] C. Orchando, J. -F. Grivaz, "SSR for p17," DØ Note 5609 (2008). 4.6

[112] T. Scanlon, "Development and optimization on MC of a Neural Network tool to identify *b*-jets," DØ Note 4889 (2005). 4.7, 5.2.4

[113] J. Pumplin *et al.*, "New Generation of Parton Distributions with Uncertainties from Global QCD Analysis," J. High Energy Phys. 0207, 012 (2002). Used versions CTEQ6M (signals) and CTEQ6L1 (backgrounds). 5.2, 5.2.1, 5.2.1, 8.2

[114] R. Brun and F. Carminati, "GEANT: Detector Description and Simulation Tool," CERN Program Library Long Writeup, Report No. W5013 (1993). 5.2

[115] Y. Fisyak, J. Womersley, "D0gstar DØ GEANT Simulation of the Total Apparatus Response," DØ Note 3191 (1991). 5.2

[116] DØSIM. http://www-d0.fnal.gov/computing/MonteCarlo/simulation/d0sim.html. 5.2

[117] E. E. Boos *et al.*, "Method for Simulating Electroweak Top-Quark Production Events in the NLO Approximation: SingleTop Generator," Phys. Atom. Nucl. 69, 1317 (2006). 5.2.1, 8.2

[118] T. Sjostrand, S. Mrenna, and P. Skands, "PYTHIA 6.4 Physics and Manual," J. High Energy Phys. 0608, 026 (2006). Used PYTHIA version 6.409. 5.2.1, 8.2

[119] S. Jadach, Z. Was, R. Decker and J. H. Kuhn, Comput. Phys. Commun. 76, 361 (1993). Used TAUOLA version v.2.5. 5.2.1, 8.2

[120] D. J. Lange,"The EvtGen Particle Decay Simulation Package," Nucl. Instrum. Meth. A 462, 152 (2001). 5.2.1, 8.2

[121] M. L. Mangano *et al.*, "ALPGEN, a Generator for Hard Multiparton Processes in Hadronic Collisions," J. High Energy Phys. 0307, 001 (2003). Used ALPGEN version 2.11. 5.2.1

[122] S. Hoeche *et al.*, "Matching Parton Showers and Matrix Elements," [hep-ph/0602031] (2006). 5.2.1

[123] D. Gillberg, "Heavy Flavour Removal and Determination of Weighting Factors for ALPGEN $W + jets$ Monte Carlo," DØ Note 5129 (2006). 5.2.1

[124] C. E. Gerber *et al.*, "Studies on the Instrumental Background in electron+jets+MET Final State Selections using p17 Monte Carlo," DØ Note 5955 (2009). 5.2.3

[125] E. Barberis *et al.*, "The Matrix Method and its Error Calculation," DØ Note 4564 (2004);

R. Schwienhorst and E. Perez, "Uncertainty Calculation for Combined W+jets and QCD Backgrounds in Analyses with W+Jets Signatures," DØ Note 4597 (2004). 5.2.3

[126] H. Schellman, "The Longitudinal Shape of the Luminous Region at DØ," DØ Note 5142 (2006);

H. Schellman, "Run IIb Longitudinal Beam Shape," DØ Note 5540 (2007). 5.2.4, 5.6

[127] B. Tiller, T. Nunneman, "Measurement of the differential $Z0$-boson production cross-section as a function of transverse momentum," DØ Note 4660 (2004). 5.2.4

[128] E. Aguilo *et al.*,"Single Top Quark Production in 2.3 $fb^{-1}$ of Data - Signal and Background Modeling and Event Selection," DØ Note 5810 (2008). 5.2.4

[129] S. Bala *et al.*, "Single Top Quark Production in 5.4 $fb^{-1}$ of Data Signal and Background Modeling and Event Selection," DØ Note 6099 (2010). 5.2.4, 5.2.4, 5.2.4, 5.2.5, 5.6, 8.3

[130] J. Hays *et al.*, "Single Electron Efficiencies in p17 Data and Monte-Carlo using p18.05.00 d0correct," DØ Note 5105 (2006). 5.2.4

[131] O. Atramentov *et al.*, "Electron and Photon Identification with p20 data," DØ Note 5761 (2008). 5.2.4, 5.6

[132] T. Gadfort *et al.*, "Muon Identification Certification for p17 Data," DØ Note 5157 (2006). 5.2.4

[133] O. Brandt *et al.*, "Muon Identification Certification for the Summer 2009 Extended Dataset (Run IIb-1 and -2)," DØ Note 6109 (2010). 5.2.4, 5.6

[134] G. Blazey, "Run II Jet Physics," DØ Note 3750 (2000). 5.2.4

[135] J. BackusMayes, "*b*-Tagging Studies in $ZH \to \mu\mu bb$" Presentation in b-ID Meeting DØ Collaboration Week 12 May (2010). 5.2.4

[136] T. Gadfort *et al.,* "Performance of the DØ NN b-tagging Tool on p20 Data," DØ Note 5554 (2007). 5.2.4

[137] S. Muanza,"A Compilation of MCFM v5.1 Cross Section," DØ Note 5300 (2006). 5.2.5

[138] S. Moch and P. Uwer, "Theoretical status and prospects for top-quark pair production at hadron colliders," Phys. Rev. D 78, 034003 (2008). 5.6

[139] F. Deliot *et al.,* "Z/$\gamma^\star$ cross section computation at NNLO using the FEWZ code," DØ Note 6050 (2010). 5.6

[140] M. Buehler *et al.,* "Search for standard-model Higgs production in the lepton + jets final state for $H \to WW^* \to lnujj$ decay in 2.58 $fb^{-1}$ of Run II Data," DØ Note 5851 (2009). 5.6

[141] F. Deliot *et al.,* "Systematic Uncertainties in Top Quark Measurements," DØ Note 6024 (2010). 5.6

[142] A. Harel, "An estimate of the data over MC, b over light jet response uncertainty for RunIIa JES," DØ Note 5654 (2008). 5.6

[143] Y. Peters *et al.,* "Measurement of the $t\bar{t}$Cross Section using Kinematic and b-tagging Information in the Lepton+Jets Channel with 5.3 fb$^{-1}$," DØ Note 6025 (2010). 5.6

[144] Y. Peters *et al.,* "Reweighting of the Fragmentation Function for the DØ Monte Carlo," DØ Note 5325 (2007). 5.6

[145] C. M. Bishop, "Neural Networks for Pattern Recognition," Clarendon Press, Oxford (1998);
R. Beale and T. Jackson, "Neural Computing: An Introduction," Adam Hilger, New York (1991). 6.1

[146] E. K. Blum and L. K. Li, "Approximation theory and feedforward networks," Neural Networks, 4, 511-515 (1991). 6.1

[147] D W. Ruck *et al.*, "The multilayer perceptron as an approximation to a Bayes optimal discriminant function," IEEE Trans. Neural Networks 1 (4), 296-298 (1990);
E. A. Wan, "Neural network classification: a Bayesian interpretation," IEEE Trans. Neural Networks 1 (4), 303-305 (1990). 6.1

[148] A. Kendall, "Advance Theory of Statistics: Bayesian Inference" Volume 2B, Oxford University Press, New York (2002) 6.2

[149] R. M. Neal, "Bayesian Learning for Neural Networks", Springer Verlag, New York (1996). 6.3, 6.4, 6.5

[150] S. Bala *et al.*, "Bayesian Neural Networks," DØ Note 4846 (2005). 6.3, 7.1.2

[151] B. A. Berg, "Markov Chain Monte Carlo Simulations and Their Statistical Analysis", World Scientific, Singapore (2004). 6.4, 7.1.3

[152] S. Duane *et al.*, "Hybrid Markov Chain Monte Carlo", Phys. Lett. B 195, 216-222 (1987). 6.4, 7.1.3

[153] R. Barlow, "Event Classification Using Weighting Methods," J. Comp. Phys. 72, 202 (1987). 7.1

[154] E. Aguilo *et al.*,"Observation of Single Top Quark Production in 2.3 $fb^{-1}$ of Data using Bayesian Neural Networks," DØ Note 5812 (2008). 7.1.2

[155] S. Bala *et al.*, "Measurement of the Single Top Quark Production Cross Section in 5.4 $fb^{-1}$ of Data," DØ Note 6128 (2010). 7.1.2, 7.2.5, 7.2.6, 7.2.6, 7.3

[156] E. Aguilo *et al.*,"Study of Single Top Qaurk Production in 2.3 $fb^{-1}$ of Data using Matrix Elements," DØ Note 5813 (2008). 7.1.2

[157] C. Peterson, T. Rognvaldsson, and L. Lonnblad, "JETNET 3.0-A versatile artificial neural network package", Computer Physics Communications 81, 185-220 (1994). 7.1.3

[158] S. Jain *et al.*, "Computing Limits Using a Bayesian Approach in the Package `top_statistics`," DØ Note 5123 (2006). 7.2, 7.3, 8.5

[159] S. Jain *et al.*, "Statistical methods implemented in the Package `top_statistics`," DØ Note 5817 (2009). 7.2, 7.2.3, 7.2.6, 7.3, 8.5

[160] E. T. Jaynes and L. Bretthorst, "Probability Theory: The Logic of Science", Cambridge University Press, England (2003). 7.2.1

[161] G. Jain *et al.*, "Asymptotic formulae for likelihood-based tests of new physics," arXiv:1007.1727v2 [hep-ex] (2010). 7.2.5

[162] V. M. Abazov *et al.* (DØ Collaboration), "Measurement of the t-channel single top quark production cross section," Phys. Lett. B 682, 363 (2010). 7.2.5, 7.4

[163] E. Aguilo *et al.*, "Model-independent measurement of t-channel single top quark production," DØ Note 5954 (2009). 7.2.5, 7.4

[164] L. Breiman *et al.*, "Classification and Regression Trees," Wadsworth, Stanford (1984). 7.2.6

[165] K. O. Stanley and R. Miikkulainen, Evolutionary Computation 10, 99 (2002). 7.2.6

[166] V. M. Abazov *et al.* (DØ Collaboration), "Measurements of single top quark production cross sections and $|V_{tb}|$ in $p\bar{p}$ collisions at $\sqrt{s} = 1.96\ TeV$," Phys. Rev. D 84, 112001 (2011). 7.2.6, 7.2.6, 7.3, 7.5, 8.2, 9

[167] V. M. Abazov *et al.* (DØ Collaboration), "Model-Independent Measurement of $t$-Channel Single Top Quark Production in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96\,TeV$," Phys. Lett. B 705, 313 (2011). 7.2.6, 7.4, 7.5, 8.2, 9

[168] E. Aguilo *et al.*, "Measurement of $|V_{tb}|$ Using the Single Top Quark Observation Analyses from 2.3 $fb^{-1}$ of data," DØ Note 5816 (2009). 7.3

[169] The LEP Collaboration and the LEP Electroweak Working Group, "A Combination of preliminary electroweak measurements and constraints on the standard model," LEPEWWG-2007-01, CERN-PH-EP-2006-042 and arXiv:hep-ex/0612034 (2007). 8.1

[170] E. Boos, L. Dudko, and T. Ohl, "Complete Calculations of $Wb\bar{b}$ and $Wb\bar{b}$+jet Production at Tevatron and LHC: Probing Anomalous $Wtb$ Couplings in Single-Top Production," Eur. Phys. J. C 11, 473 (1999). 8.1

[171] E. Malkawi and C.-P.Yuan,"Global Analysis of the Top Quark couplings to Gauge Bosons," Phys. Rev. D 50, 4462 (1994). 8.1

[172] N. Cabibbo,"Unitarity Symmetry and Leptonic Decays," Phys. Rev. Lett. 10, 531 (1963). 8.1

[173] M. Kobayashi and K. Maskawa,"CP Violations in the Renormalizable Theory of Weak Interaction," Prog. Theor. Phys. 49, 652 (1973). 8.1

[174] G. L. Kane *et al.,* Phys. Rev. D 45, 124 (1992). J. A. Aguilar-Saavedra, Nucl. Phys. B 812, 181 (2009). 8.1

[175] S. Eidelman *et al.,* [Particle Data Group] Phys. Lett. B 592, 1 (2004). 8.1

[176] F. Larios, M. A. Perez and C. P. Yuan, Phys. Lett. B 457, 334 (1999). 8.1

[177] V. M. Abazov *et al.* (DØ Collaboration), Phys. Rev. Lett. 100, 062004 (2008). 8.1, 8.5

[178] CDF Collaboration, CDF Note 8280 (2006). 8.1

[179] E. E. Boos *et al.*, "Search for Anomalous $Wtb$ couplings with 0.9 $fb^{-1}$ of DØ Data," DØ Note 5649 (2008). 8.1

[180] V. M. Abazov *et al.* (DØ Collaboration), "Search for anomalous $Wtb$ couplings in single top quark production," Phys. Rev. Lett. 101, 221801 (2008). 8.1, 8.5, 8.7

[181] S. Bala *et al.*, "Search for Anomalous $Wtb$ couplings in single top production with 5.4 $fb^{-1}$ of DØ Data," DØ Note 6206 (2011). 8.3, 8.3, 8.5

[182] V. M. Abazov *et al.* (DØ Collaboration), "Search for anomalous $Wtb$ couplings in single top quark production in $p\bar{p}$ collisions at $\sqrt{s} = 1.96\ TeV$," Phys. Lett. B 708, 21 (2012). 8.6, 8.7, 9

# List of Publications

## Publications and DØ Internal Notes as Primary **Author** (as an outcome of work done)

1. V. M. Abazov *et al.* (D0 Collaboration), "Model-independent measurement of $t$-channel single top quark production in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV," Phys. Lett. B **705**, 313 (2011) [arXiv:1105.2788 [hep-ex]].

2. V. M. Abazov *et al.* (D0 Collaboration), "Measurements of single top quark production cross sections and $|V_{tb}|$ in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV," Phys. Rev. D **84**, 112001 (2011) [arXiv:1108.3091 [hep-ex]].

3. V. M. Abazov *et al.* (D0 Collaboration), "Search for anomalous $Wtb$ couplings in single top quark production in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV," Phys. Lett. B **708**, 21 (2012) [arXiv:1110.4592 [hep-ex]].

4. J. Joshi *et al.* (D0 Collaboration), "Observation of single top quark production at D0 using Bayesian neural networks," Conf. Proc. C **0908171**, 431 (2009).

5. S. Bala *et al.*, "Single Top Quark Production in 5.4 $fb^{-1}$ of Data  Signal and Background Modeling and Event Selection," DØ Note 6099, (2010).

6. S. Bala *et al.*, "Measurement of the Single Top Quark Production Cross Section in 5.4 $fb^{-1}$ of Data," DØ Note 6128, (2010).

7. S. Bala *et al.*, "Search for Anomalous $Wtb$ couplings in single top production with 5.4 $fb^{-1}$ of DØ Data," DØ Note 6206, (2011).

8. S. Bala *et al.*, "Search of CP Violation in Single Top Quark Production in 5.4 $fb^{-1}$ of Data." DØ Note 6205 (2011).

# Full List of Papers (Published/Accepted)

1. V. M. Abazov *et al.* (D0 Collaboration), "Search for Higgs bosons of the minimal supersymmetric standard model in $p\bar{p}$ collisions at $\sqrt(s) = 1.96$ TeV," arXiv:1112.5431 [hep-ex].

2. V. M. Abazov *et al.* (D0 Collaboration), "Search for universal extra dimensions in ppbar collisions," arXiv:1112.4092 [hep-ex].

3. V. M. Abazov *et al.* (D0 Collaboration), "Measurements of $WW$ and $WZ$ production in $W$ + jets final states in $p\bar{p}$ collisions," arXiv:1112.0536 [hep-ex].

4. V. M. Abazov *et al.* (D0 Collaboration), "$Z\gamma$ production and limits on anomalous $ZZ\gamma$ and $Z\gamma\gamma$ couplings in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV," arXiv:1111.3684 [hep-ex].

5. V. M. Abazov *et al.* (D0 Collaboration), "Search for a Narrow $t\bar{t}$ Resonance in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV," arXiv:1111.1271 [hep-ex].

6. V. M. Abazov *et al.* (D0 Collaboration), "Measurement of the relative branching ratio of $B_s^0 to J/\psi f_0(980) \rightarrow B_s^0 \rightarrow J/\psi\phi$," arXiv:1110.4272 [hep-ex].

7. V. M. Abazov *et al.* (D0 Collaboration), "Evidence for spin correlation in $t\bar{t}$ production," arXiv:1110.4194 [hep-ex].

8. V. M. Abazov *et al.* (D0 Collaboration), "Measurement of the inclusive jet cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV," arXiv:1110.3771 [hep-ex].

9. V. M. Abazov *et al.* (D0 Collaboration), "A Search for charged massive long-lived particles," [arXiv:1110.3302 [hep-ex]].

10. V. M. Abazov *et al.* (D0 Collaboration), "$W\gamma$ production and limits on anomalous $WW\gamma$ couplings in $p\bar{p}$ collisions," Phys. Rev. Lett. **107**, 241803 (2011) [arXiv:1109.4432 [hep-ex]].

11. V. M. Abazov *et al.* (D0 Collaboration), "Measurement of the CP-violating phase $\phi_s^{J/\psi\phi}$ using the flavor-tagged decay $B_s^0 \to J/\psi\phi$ in 8 fb$^{-1}$ of $p\bar{p}$ collisions," arXiv:1109.3166 [hep-ex].

12. V. M. Abazov *et al.* (D0 Collaboration), "Model independent search for new phenomena in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV," arXiv:1108.5362 [hep-ex].

13. V. M. Abazov *et al.* (D0 Collaboration), "Forward-backward asymmetry in top quark-antiquark production," Phys. Rev. D **84**, 112005 (2011) [arXiv:1107.4995 [hep-ex]].

14. V. M. Abazov *et al.* (D0 Collaboration), "Search for the standard model and a fermiophobic Higgs boson in diphoton final states," Phys. Rev. Lett. **107**, 151801 (2011) [arXiv:1107.4587 [hep-ex]].

15. V. M. Abazov *et al.* (D0 Collaboration), "Search for first generation leptoquark pair production in the electron + missing energy + jets final state," Phys. Rev. D **84**, 071104 (2011) [arXiv:1107.1849 [hep-ex]].

16. V. M. Abazov *et al.* (D0 Collaboration), "Search for associated Higgs boson production using like charge dilepton events in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV," Phys. Rev. D **84**, 092002 (2011) [arXiv:1107.1268 [hep-ex]].

17. V. M. Abazov *et al.* (D0 Collaboration), "Measurement of the anomalous like-sign dimuon charge asymmetry with 9 $fb^{-1}$ of p pbar collisions," Phys. Rev. D **84**, 052007 (2011) [arXiv:1106.6308 [hep-ex]].

18. V. M. Abazov *et al.* (D0 Collaboration), "Precision measurement of the ratio B($t \to Wb$)/B($t \to Wq$) and Extraction of $V_{tb}$," Phys. Rev. Lett. **107**, 121802 (2011) [arXiv:1106.5436 [hep-ex]].

19. V. M. Abazov *et al.* (D0 Collaboration), "Search for neutral Minimal Supersymmetric Standard Model Higgs bosons decaying to tau pairs produced in association with $b$ quarks in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV," Phys. Rev. Lett. **107**, 121801 (2011) [arXiv:1106.4885 [hep-ex]].

20. V. M. Abazov *et al.* [D0 Collaboration], "Search for Higgs bosons decaying to $\tau\tau$ pairs in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV," Phys. Lett. B **707**, 323 (2012) [arXiv:1106.4555 [hep-ex]].

21. V. M. Abazov *et al.* (D0 Collaboration), "Search for doubly-charged Higgs boson pair production in ppbar collisions at sqrts = 1.96 TeV," arXiv:1106.4250 [hep-ex].

22. V. M. Abazov *et al.* (D0 Collaboration), "Bounds on an anomalous dijet resonance in $W$+jets production in ppbar collisions at $\sqrt{s} = 1.96$ TeV," Phys. Rev. Lett. **107**, 011804 (2011) [arXiv:1106.1921 [hep-ex]].

23. V. M. Abazov *et al.* (D0 Collaboration), "Direct measurement of the mass difference between top and antitop quarks," Phys. Rev. D **84**, 052005 (2011) [arXiv:1106.2063 [hep-ex]].

24. V. M. Abazov *et al.* (D0 Collaboration), "Measurements of inclusive $W$+jets production rates as a function of jet transverse momentum in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV," Phys. Lett. B **705**, 200 (2011) [arXiv:1106.1457 [hep-ex]].

25. V. M. Abazov *et al.* (D0 Collaboration), "Precise measurement of the top-quark mass from lepton+jets events at D0," Phys. Rev. D **84**, 032004 (2011) [arXiv:1105.6287 [hep-ex]].

26. V. M. Abazov *et al.* (D0 Collaboration), "Measurement of the $t\bar{t}$ production cross section using dilepton events in $p\bar{p}$ collisions," Phys. Lett. B **704**, 403 (2011) [arXiv:1105.5384 [hep-ex]].

27. V. M. Abazov *et al.* (D0 Collaboration), "Measurement of the production fraction times branching fraction $\boldsymbol{f(b \to \Lambda_b) \cdot \mathcal{B}(\Lambda_b \to J/\psi\Lambda)}$," Phys. Rev. D **84**, 031102 (2011) [arXiv:1105.0690 [hep-ex]].

28. V. M. Abazov *et al.* (D0 Collaboration), "Precise measurement of the top quark mass in the dilepton channel at D0," Phys. Rev. Lett. **107**, 082004 (2011) [arXiv:1105.0320 [hep-ex]].

29. V. M. Abazov *et al.* (D0 Collaboration), "Measurement of spin correlation in $t\bar{t}$ production using a matrix element approach," Phys. Rev. Lett. **107**, 032001 (2011) [arXiv:1104.5194 [hep-ex]].

30. V. M. Abazov *et al.* (D0 Collaboration), "Search for a fourth generation $t'$ quark in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV," arXiv:1104.4522 [hep-ex].

31. V. M. Abazov *et al.* (D0 Collaboration), "Measurement of $\sin^2\theta^{\ell}_{\rm eff}$ and $Z$-light quark couplings using the forward-backward charge asymmetry in $p\bar{p} \to Z/\gamma^* \to e^+e^-$ events with $\mathcal{L} = 5.0\,{\rm fb}^{-1}$ at $\sqrt{s} = 1.96$ TeV," Phys. Rev. D **84**, 012007 (2011) [arXiv:1104.4590 [hep-ex]].

32. V. M. Abazov *et al.* (D0 Collaboration), "Measurement of the $ZZ$ production cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV," Phys. Rev. D **84**, 011103 (2011) [arXiv:1104.3078 [hep-ex]].

33. V. M. Abazov *et al.* (D0 Collaboration), "Determination of the pole and MSbar masses of the top quark from the $t\bar{t}$ cross section," Phys. Lett. B **703**, 422 (2011) [arXiv:1104.2887 [hep-ex]].

34. V. M. Abazov *et al.* (D0 Collaboration), "Measurement of three-jet differential cross sections $d\sigma_{3\mathrm{jet}}/dM_{3\mathrm{jet}}$ in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV," Phys. Lett. B **704**, 434 (2011) [arXiv:1104.1986 [hep-ex]].

35. V. M. Abazov *et al.* (D0 Collaboration), "Search for flavor changing neutral currents in decays of top quarks," Phys. Lett. B **701**, 313 (2011) [arXiv:1103.4574 [hep-ex]].

36. V. M. Abazov *et al.* (D0 Collaboration), "Measurement of spin correlation in $t\bar{t}$ production using dilepton final states," Phys. Lett. B **702**, 16 (2011) [arXiv:1103.1871 [hep-ex]].

37. V. M. Abazov *et al.* (D0 Collaboration), "Search for the Standard Model Higgs Boson in the $H \to WW \to \ell\nu q'\bar{q}$ Decay Channel," Phys. Rev. Lett. **106**, 171802 (2011) [arXiv:1101.6079 [hep-ex]].

38. V. M. Abazov *et al.* (D0 Collaboration), "Azimuthal decorrelations and multiple parton interactions in photon+2 jet and photon+3 jet events in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV," Phys. Rev. D **83**, 052008 (2011) [arXiv:1101.1509 [hep-ex]].

39. V. M. Abazov *et al.* (D0 Collaboration), "Measurement of the top quark pair production cross section in the lepton+jets channel in proton-antiproton collisions at $\sqrt{s}$=1.96 TeV," Phys. Rev. D **84**, 012008 (2011) [arXiv:1101.0124 [hep-ex]].

40. V. M. Abazov *et al.* (D0 Collaboration), "Measurement of color flow in $\mathbf{t\bar{t}}$ events from $\mathbf{p\bar{p}}$ collisions at $\sqrt{\mathbf{s}} = \mathbf{1.96}$ TeV," Phys. Rev. D **83**, 092002 (2011) [arXiv:1101.0648 [hep-ex]].

41. V. M. Abazov *et al.* (D0 Collaboration), "Search for $W' \to tb$ resonances with left- and right-handed couplings to fermions," Phys. Lett. B **699**, 145 (2011) [arXiv:1101.0806 [hep-ex]].

42. 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," Phys. Lett. B **698**, 6 (2011) [arXiv:1012.0874 [hep-ex]].

43. V. M. Abazov *et al.* (D0 Collaboration), "Search for resonant WW and WZ production in ppbar collisions at $\sqrt{\mathbf{s}} = \mathbf{1.96}$ TeV," Phys. Rev. Lett. **107**, 011801 (2011) [arXiv:1011.6278 [hep-ex]].

44. V. M. Abazov *et al.* (D0 Collaboration), "Measurement of the W boson helicity in top quark decays using 5.4 fb$^{-\mathbf{1}}$ of $\boldsymbol{p\bar{p}}$ collision data," Phys. Rev. D **83**, 032009 (2011) [arXiv:1011.6549 [hep-ex]].

45. V. M. Abazov *et al.* (D0 Collaboration), "Search for neutral Higgs bosons in the multi-$b$-jet topology in 5.2fb$^{-1}$ of $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV," Phys. Lett. B **698**, 97 (2011) [arXiv:1011.1931 [hep-ex]].

46. V. M. Abazov *et al.* (D0 Collaboration), "A measurement of the ratio of inclusive cross sections $\sigma(p\bar{p} \to Z + b\,\text{jet})/\sigma(p\bar{p} \to Z + \text{jet})$ at $\sqrt{s} = 1.96$ TeV," Phys. Rev. D **83**, 031105 (2011) [arXiv:1010.6203 [hep-ex]].

47. V. M. Abazov *et al.* (D0 Collaboration), "Search for single vector-like quarks in ppbar collisions at $\sqrt{s} = 1.96$ TeV," Phys. Rev. Lett. **106**, 081801 (2011) [arXiv:1010.1466 [hep-ex]].

48. V. M. Abazov *et al.* (D0 Collaboration), "Precise study of the $Z/\gamma^*$ boson transverse momentum distribution in $p\bar{p}$ collisions using a novel technique," Phys. Rev. Lett. **106**, 122001 (2011) [arXiv:1010.0262 [hep-ex]].

49. V. M. Abazov *et al.* (D0 Collaboration), "Determination of the width of the top quark," Phys. Rev. Lett. **106**, 022001 (2011) [arXiv:1009.5686 [hep-ex]].

50. V. M. Abazov *et al.* (D0 Collaboration), "Search for pair production of the scalar top quark in the electron+muon final state," Phys. Lett. B **696**, 321 (2011) [arXiv:1009.5950 [hep-ex]].

51. V. M. Abazov *et al.* (D0 Collaboration), "High mass exclusive diffractive dijet production in $\mathbf{p\bar{p}}$ collisions at $\sqrt{\mathbf{s}} = 1.96$ TeV," Phys. Lett. B **705**, 193 (2011) [arXiv:1009.2444 [hep-ex]].

52. V. M. Abazov *et al.* (D0 Collaboration), "Measurement of $t\bar{t}$ production in the tau + jets topology using $p\bar{p}$ collisions at sqrts = 1.96 TeV," Phys. Rev. D **82**, 071102 (2010) [arXiv:1008.4284 [hep-ex]].

53. V. M. Abazov *et al.* (D0 Collaboration), "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 [hep-ex]].

54. V. M. Abazov *et al.* (D0 Collaboration), "Search for New Fermions ('Quirks') at the Fermilab Tevatron Collider," Phys. Rev. Lett. **105**, 211803 (2010) [arXiv:1008.3547 [hep-ex]].

55. V. M. Abazov *et al.* (D0 Collaboration), "Search for events with leptonic jets and missing transverse energy in $\mathbf{p\bar{p}}$ collisions at $\sqrt{\mathbf{s}} = \mathbf{1.96}$ TeV," Phys. Rev. Lett. **105**, 211802 (2010) [arXiv:1008.3356 [hep-ex]].

56. V. M. Abazov *et al.* (D0 Collaboration), "Search for a heavy neutral gauge boson in the dielectron channel with 5.4 $fb^{-1}$ of ppbar collisions at $\sqrt{\mathbf{s}} = \mathbf{1.96}$ TeV," Phys. Lett. B **695**, 88 (2011) [arXiv:1008.2023 [hep-ex]].

57. V. M. Abazov *et al.* (D0 Collaboration), "Search for diphoton events with large missing transverse energy in 6.3 fb$^{-1}$ of $\mathbf{p\bar{p}}$ collisions at $\sqrt{\mathbf{s}} = \mathbf{1.96}$ TeV," Phys. Rev. Lett. **105**, 221802 (2010) [arXiv:1008.2133 [hep-ex]].

58. V. M. Abazov *et al.* (D0 Collaboration), "Search for sneutrino production in emu final states in 5.3 $fb^{-1}$ of ppbar collisions at $\sqrt{\mathbf{s}} = \mathbf{1.96}$ TeV," Phys. Rev. Lett. **105**, 191802 (2010) [arXiv:1007.4835 [hep-ex]].

59. V. M. Abazov *et al.* (D0 Collaboration), "Search for flavor changing neutral currents via quark-gluon couplings in single top quark production using 2.3 $fb^{-1}$ of ppbar collisions," Phys. Lett. B **693**, 81 (2010) [arXiv:1006.3575 [hep-ex]].

60. V. M. Abazov *et al.* (D0 Collaboration), "Search for the rare decay $B_s^0 \to mu^+mu^-$," Phys. Lett. B **693**, 539 (2010) [arXiv:1006.3469 [hep-ex]].

61. V. M. Abazov *et al.* (D0 Collaboration), "Measurement of the $WZ \to \ell\nu\ell\ell$ cross section and limits on anomalous triple gauge couplings in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV," Phys. Lett. B **695**, 67 (2011) [arXiv:1006.0761 [hep-ex]].

62. V. M. Abazov *et al.* (D0 Collaboration), "Measurement of the normalized $Z/\gamma^* - > \mu^+\mu^-$ transverse momentum distribution in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV," Phys. Lett. B **693**, 522 (2010) [arXiv:1006.0618 [hep-ex]].

# Conferences/Schools Attended

1. "**Root @ Heavy Ions**", 12-16 February 2008, University of Jammu, India.

2. "**Chandigarh Science Congress**", 14-15 March 2008, Panjab University, India.

3. "**International Conference on Early Physics at LHC with CMS Detector (LHC08)**", 26-29 March 2008, TIFR, India.

4. "**XXIV International Symposium on Lepton Photon Interactions at High Energies (Lepton Photon 09)**", 17-22 August 2009, Hamburg, Germany.

5. "**2010 CERN-Fermilab Hadron Collider Physics Summer School**", 16-27 August 2010, Fermilab, USA.

6. "**Phenomenology 2011 Symposium**", 9-11 May 2011, University of Wisconsin,Madison, USA.

7. "**New Perspectives 2011 - 44th Fermilab Users Meeting**", 31 May 2011, Fermilab, USA.

8. "**XXV International Symposium on Lepton Photon Interactions at High Energies (Lepton Photon 2011)**", 22-27 August 2011, TIFR, India.