MEASUREMENTS OF $\sigma(V + D^*)/\sigma(V)$ IN 9.7 FB^{-1} AT CDF RUN II

BY

KEITH MATERA

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Physics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2014

Urbana, Illinois

Doctoral Committee:

Associate Professor Mark Neubauer, Chair
Professor Kevin Pitts, Director of Research
Professor Scott Willenbrock
Professor Nadya Mason

Abstract

The Standard Model of particle physics has been remarkably successful, but the non-perturbative features of quantum chromodynamics must be tested and modeled with data. There have been many such tests, focused primarily on the use of jet-based probes of heavy flavor (bottom and charm quark) production at hadron colliders. In this thesis, we propose and test a strategy for identifying heavy flavor in events containing a W or Z vector boson (a V boson); this technique probes a much lower energy regime than can be explored by jet-based methods. In a sample of W and Z events skimmed from 9.7 fb^{-1} of high-p_T electron and muon data from CDF Run II $p\bar{p}$ collisions at center of mass energy $\sqrt{s} = 1.96$ GeV , we identify charm by fully reconstructing $D^*(2010) \to D^0(\to K\pi)\pi_s$ decays at the track level. Using a binned fit of $\Delta m = m(K\pi\pi_s) - m(K\pi)$ to count reconstructed D^* candidates, we then unfold these raw counts with acceptance values derived from Monte Carlo, and present measurements of $\sigma(W + D^*)/\sigma(W)$ and $\sigma(Z + D^*)/\sigma(Z)$ in the W/Z leptonic decay channels. All measurements are found to be in agreement with the predictions of Pythia 6.2 (PDF set CTEQ5L). These results include the first measurement of $W/Z + c$ production in events with zero jet objects at the Tevatron, and the first measurement of $W/Z + c$ production with $p_T(c) < 15$ GeV at the Tevatron.

Acknowledgments

Any acknowledgments must begin with Kevin Pitts, my advisor for the past three and a half years. He was responsible for the idea and direction of this work, but more importantly he taught me how to navigate the ever-heaving landscape of a live analysis. I owe most of my knowledge of particle physics to Kevin's ceaseless attempts to break colliders down into cartoon simplicity.

I owe the rest to CDF co-conveners Bo Jayatilaka and Christina Mesropian, and to CDF spokesperson Costas Vellidis. Our Thursday morning video conferences—undoubtedly as much a highlight of their week as it was mine—provided the full collaboration experience with only a fraction of the personnel. Their suggestions, advice, and guidance turned a few raw numbers into a legitimate physical study.

Going back a few years, I have to thank everyone who put in the time to help me out as an ACE during my data acquisition shifts. Without them, a plaque in the control room would forever read: "Most integrated luminosity lost: Keith "Runpocolypse" Matera, July 2011." A special thanks to John Freeman, for merging feng shui with the CDF trailer environment and putting the class in classical mechanics. Bo made an appearance in those trailers, too. He might actually be a ghost.

Olga Norniella for providing me with neural network code; Dale Stentz for jet treatment pointers; Jen-Chieh Peng for getting SU(N) to make sense in my head; the beleaguered Fermilab help desk for countless answered service tickets; Jessica Seckler, for some great work with Z bosons; Markus, Dave, and Greg, for the nights of problem sets and Minecraft; Mote and Frank, for occasionally humoring me when I wanted to discuss physics; Joe and Garrett—you two might end up providing something useful.

Finally, a special thanks to Bree, Kyser, and Riggs; you three provided me with rare worry-free moments during even the most stressful times of the past two years. One of these days, we'll actually have a weekend together that's not focused on particle physics or a home repair project. I love you all.

This thesis was supported in part by the University of Illinois, the Universities Research Association (award number DOE FERMI 11-S-009), and the United States Department of Energy (grant number DE-FG02-13ER42001).

Table of Contents

List of Tables

List of Figures

List of Abbreviations

SM	Standard Model
QED	Quantum Electrodynamics
QCD	Quantum Chromodynamics
EWK	Electroweak
DY	Drell-Yan
CDF	Collider Detector at Fermilab
MI	Main Injector
COT	Central Outer Tracker
L00	Layer 00 Silicon Tracker
SVXII	Silicon Vertex Tracker
ISL	Intermediate Silicon Layer Tracker
CEM	Central Electromagnetic Calorimeter
PEM	Plug Electromagnetic Calorimeter
CHA	Central Hadronic Calorimeter
WHA	Wall Hadronic Calorimeter
PHA	Plug Hadronic Calorimeter
CES	Central Electromagnetic ShowerMax Chamber
PES	Plug Electromagnetic ShowerMax Chamber
CMU	Central Muon Detector
CMP	Central Muon Upgrade Detector
CMX	Central Muon Extension Detector
BMU	Barrel Muon Detector
CLC	Cerenkov Luminosity Counters
XFT	eXtremely Fast Tracker
XTRP	Extrapolation Unit
SVT	Silicon Vertex Trigger

JES	Jet Energy Scale
MC	Monte Carlo Simulation
NLO	Next-to-Leading Order
NNLO	Next-to-Next-to-Leading Order
NNLL	Next-to-Next-to-Leading Log
ME	Matrix Element
LR	Likelihood Ratio
NN	Neural Network
CL	Confidence Level

List of Symbols

p_T	Transverse Momentum
η	Pseudorapidity
ϕ	Detector azimuthal angle
E_T	Transverse energy
$\not\!\!E_T$	Missing transverse energy
ΔR	Spatial separation in $\eta - \phi$ space
WC	$p\bar{p} \to W + c + X$ production events
WCC	$p\bar{p} \to W + g(\to c\bar{c}) + X$ production events
WBB	$p\bar{p} \to W + g(\to b\bar{b}) + X$ production events

Introduction

The Standard Model of particle physics has proven itself to be a tremendously successful description of the fundamental particles and their interactions. The last of its predicted particles, the Higgs boson, was identified in the summer of 2012, and while there exist tantalizing bits of evidence for new physics (non-zero neutrino masses, and the existence of dark matter), the standard model still lies at the heart of the field.

This makes it all the more interesting that standard model interactions are not perfectly well understood. Many studies in collider physics rely on theoretical models to predict the behavior of particle interactions. While electroweak interactions are well-described by perturbative expansions, the strong interactions of QCD are non-perturbative at low energies, and cannot be predicted from first principles. The best predictions are made by using models tuned to data, especially those involving more exotic states of matter that do not exist in our everyday lives—such as the "heavy flavor" bottom and charm quarks.

This thesis details a set of measurements of the behavior of charm quarks produced in association with vector bosons, using the CDF detector at the Tevatron collider. In Chapter 1, we introduce the theory of the standard model and quantum chromodynamics (QCD), and we describe gaps in our knowledge of QCD at low energy. A description of the Fermilab accelerator complex and CDF detector comprises Chapter 2.

We begin the analysis proper with Chapter 3, in which we discuss general procedures for reconstructing objects of interest to our study: objects such as leptons, missing energy, and jets. In Chapter 4, we focus on the selection criteria used to identify W and Z vector bosons (V bosons), and in Chapter 5 we focus on the process used to identify charmed D^* mesons. These methods are combined in Chapter 6 to measure the rate at which D^* are produced in associaton with V in $p\bar{p}$ collisions at CDF. We then focus—in Chapters 7, 8, and 9—on measuring this production rate as a function of D^* momentum; we also unfold observed rates to a ratio of physical cross-sections using acceptance values derived from simulated events. In Chapter 10, we finish our work by identifying what percentage of our charmed meson signal can be attributed to each of three possible production processes.

We conclude with a brief summary of our results and their meaning, and a discussion of future prospects for charmed meson reconstruction in hadron collider experiments.

Chapter 1

The Standard Model and Quantum Chromodynamics

Particle physics is the study of the basic building blocks of the universe. Our current understanding of things is that all matter is composed of a set of fundamental particles, and that all forces are mediated by a complimentary set. The full roster of these fundamental particles can be probed only by studying interactions at energy scales that are much higher than those which we experience in everyday life. To access these higher-energy states, physicists accelerate common subatomic particles to enormous speeds, and collide them in order to force a huge exchange of energy. The result is a spray of elementary particles—some common (such as electrons), and others much more exotic (such as the Higgs boson). In this chapter, we describe the full range of elementary particles that are currently known to populate our universe, and then describe the Standard Model that strives to unite them all behind one simple theoretical premise.

1.1 Matter

All matter is composed of bound states of point-like, half-integer spin particles called fermions. There are two types of these fermions, leptons and quarks, each of which interacts with a different subset of the four known forces (discussed in the next section). These fermions are further split into three generations of particles, loosely defined by tiers of mass, with the members of each successive generation acting essentially as heavier versions of the generation below it. A single generation consists of two leptons and two quarks, and members of each pair differ in charge by one unit in e (the magnitude of the electron charge, e=1.602$\times 10^{-19}$ C).

As an illustrative example, the first generation of leptons consists of the electron, e, (charge -1e) and the electron-neutrino, ν_e, (charge 0). The first generation of quarks consists of the up quark, u (charge $+2/3e$) and the down quark, d (charge -1/3e). As u and d quarks compose protons and neutrons, this first generation of particles makes up virtually all of the matter that we encounter on a daily basis. Table 1.1 lists these particles and their properties, as well as those of the remaining two generations. For each particle listed, there is also a corresponding anti-particle which has the same mass, but opposite charge and other quantum numbers.

	Leptons			Quarks		
Generation	**Flavor**	**Mass**	**Charge**	**Flavor**	**Mass**	**Charge**
1st	ν_e	< 2.2 eV/c^2	0	u	$1.7 - 3.1$ MeV/c^2	$+2/3e$
	e	0.511 MeV/c^2	$-e$	d	$4.1 - 5.7$ MeV/c^2	$-1/3e$
2nd	ν_μ	< 0.17 MeV/c^2	0	c	$1.290^{+0.05}_{-0.011}$ GeV/c^2	$+2/3e$
	μ	105.7 MeV/c^2	$-e$	s	100^{+30}_{-20} MeV/c^2	$-1/3e$
3rd	ν_τ	< 15.5 eV/c^2	0	t	$172.9^{+0.3}_{-0.9}$ GeV/c^2	$+2/3e$
	τ	1.777 GeV/c^2	$-e$	b	$4.190^{+0.18}_{-0.06}$ GeV/c^2	$-1/3e$

Table 1.1: A list of the fundamental particles that make up matter, divided by type (lepton or quark) and generation. For each particle listed here, there is also a corresponding anti-particle with the same mass, but opposite charge and other quantum numbers. Though not stated explicitly in this table, there are three possible color states (red, blue, and green) for each quark.

Force	**Boson**	**Mass**	**Fermions Affected**
Gravitational	Graviton	0	Quarks and leptons
Electromagnetic	Photon (γ)	0	Quarks and *charged* leptons
Weak	W^+, W^-	80.4 GeV/c^2	Quarks and leptons
	Z	91.2 GeV/c^2	
Strong	Gluon (g)	0	Quarks

Table 1.2: A list of the four fundamental forces, and the particles that mediate these forces. To the right, the fermions that are affected by each force. Though not stated explicitly in the table, there are eight possible color states for the gluon.

1.2 Forces

Interactions between the elementary fermions described in the previous section are governed by the four fundamental forces: gravity, electromagnetism, the weak force, and the strong force. Each of these is mediated by one or more integral-spin bosons: the photon (electromagnetism), the W^+, W^-, and Z bosons (weak force), the gluon (strong force), and the graviton (gravitation). The graviton has never been confirmed experimentally, but is believed to be massless and to carry spin 2. The remaining bosons are massless with spin 1, except for the W and Z bosons—these are given a non-zero mass through electroweak symmetry breaking, as discussed in the next chapter.

Not all fundamental fermions interact through all of these forces. First of all, in the types of interactions that are of interest in high energy physics, the gravitational force has a negligible effect when compared to any other force. We will, in fact, completely ignore gravity for the remainder of this thesis. The quarks interact through the strong, electromagnetic, and weak forces, but the leptons interact only through the electromagnetic and weak forces; a further restriction is that only charged particles may interact electromagnetically, which leaves neutrinos to interact only through the weak force. The four fundamental forces, and the bosons that mediate them, are listed in Table 1.2.

1.3 The Standard Model

The identification of the fundamental fermions which lie at the heart of matter, and of the bosons which mediate force between them, was a fantastic achievement. Perhaps most notable in this work was the creation of a general framework for describing the strong, weak, and electromagnetic interactions between quarks and leptons: the Standard Model (SM). The SM has been remarkably successful, correctly predicting the existence of the top quark and the Higgs boson, and—with the help of numerical simulations—even the complicated interactions between fundamental particles in such chaotic environments as hadron colliders. While the SM does not account for dark matter, or for the mass of neutrinos, it is extremely predictive in all other regimes, and it lies at the heart of all ongoing research in particle physics.

The SM is a quantum field theory, in which the fundamental fermions and the forces between them are waveforms that propagate through fields permeating all of space. Each particle is associated with its own field; a particle can be looked at as a 'bump' in its field, while its corresponding anti-particle is a 'hole', in a sense somewhat like electrons and holes in a semiconductor (an admittedly flawed analogy, given that it involves electrons, but hopefully an illustrative one). Mathematically, each fermion is described by a four-component Dirac spinor field, with the dynamics of these particles governed by a Lagrangian involving the fermion fields and their derivatives [1].

The form of the SM Lagrangian is determined by imposing natural symmetries; relativity, for example, requires Lorentz invariance. Gauge symmetries are perhaps the most important of these symmetries; essentially, these are statements that the phases of these fields are arbitrarily-defined and unobservable quantities. Therefore, a change in the definition of these phases ought to have no effect on physical observables. What is remarkable about these requirements, is just how far they get us. Local gauge symmetries—requiring invariance after shifting the phase θ by some space-time dependent continuous function $\theta(\overrightarrow{x}, t)$—require the introduction of integral-spin *gauge boson* fields; these, in turn, couple to the fermion fields and introduce interaction terms between the fermions. With the proper choice of gauge symmetry, these gauge fields can be identified with the fundamental forces of the SM [2, 3].

Explicitly, the gauge structure of the SM is said to be $\mathrm{SU(3)_{color}} \otimes \mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$. Enforcing gauge symmetry upon the $\mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$ piece results in a unified theory of the electromagnetic and weak interactions, known as electroweak theory. Under this (unmodified) theory, there are four massless gauge bosons—one for the $\mathrm{U(1)_Y}$ part, and three for the $\mathrm{SU(2)_L}$ part. Unfortunately, introducing mass terms for the quarks and leptons destroys the $\mathrm{SU(2)_L}$ symmetry, and we know from experiment that a successful electroweak theory must predict three *massive* gauge bosons (the $W^{\pm}$ and the Z). Both of these problems can be addressed by introducing an SU(2) scalar Higgs field to the SM Lagrangian. The correct choice

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

$$\begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} = \begin{pmatrix} 0.97427 \pm 0.00015 & 0.22534 \pm 0.00065 & 0.00351^{+0.00015}_{-0.00014} \\ 0.22520 \pm 0.00065 & 0.97344 \pm 0.00016 & 0.0412^{+0.0011}_{-0.005} \\ 0.00867^{+0.00029}_{-0.00031} & 0.0404^{+0.0011}_{-0.0005} & 0.999146^{+0.000021}_{-0.000046} \end{pmatrix}$$

Table 1.3: Top: the column to the left is filled by weak-interaction partners to the up-type quarks; to the right, a matrix describes the 'recipe of down-type quark mass eigenstate that composes each of the weak-interaction partner states. Bottom: the magntudes of the CKM mixing matrix elements. It is easy to see that couplings within the same generation of mass eigenstates are far larger than inter-generational couplings. While not shown here, a single complex phase δ in the CKM matrix allows for parity violation through the weak interaction.

of potential for this field breaks the $\mathrm{SU(2)_L}$ symmetry, simultaneously giving mass to the three $\mathrm{SU(2)_L}$ bosons, *and* introducing mass terms for the quarks and leptons. For this reason, the existence of a Higgs boson associated with this theoretically-successful Higgs field was long suspected, and finally confirmed in the summer of 2012 [4, 5].

The $\mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$ gauge symmetry and the associated Higgs field result in some interesting features. For one, the non-Abelian nature of the $\mathrm{SU(2)_L}$ symmetry results in the electroweak gauge bosons coupling to one another (*e.g.*, the charged $W^{\pm}$ coupling to photons). Perhaps more intriguing, the weak force eigenstates are not the same as the mass eigenstates generated by coupling to the Higgs (the eigenstates listed in Table 1.1). The weak interaction may thus transmute quarks and leptons of one flavor or generation into quarks and leptons of another flavor or generation. The couplings between mass eigenstates that are produced by the weak force are summarized in the Cabibbo-Kobayashi-Maskawa (CKM) mixing matrix (Table 1.3). A single complex phase in the CKM matrix allows for parity violation (in this context, parity violation manifests itself as unequal rates between interaction $a + b \to c$ and its anti-matter parallel $\bar{a} + \bar{b} \to \bar{c}$), which may be responsible for at least part of the observable universe's matter/anti-matter asymmetry.

While the electroweak interaction of the $\mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$ gauge symmetry is a rich theory, the $\mathrm{SU(3)_{color}}$ part of the SM's full symmetry may be considered richer still. This is the strong force piece of the SM, and it is referred to as quantum chromodynamics (QCD). The 'color' subscript and 'chromo' prefix both refer to the name of the of the strong force 'charge' —a three-valued property called color. A quark must have one color (red (R), green (G), or blue (B)), and an anti-quark one anti-color (anti-red ($\bar{R}$), anti-green ($\bar{G}$), or anti-blue ($\bar{B}$)). Quarks of one color are attracted to quarks of another color, and they bind together to form colorless states in combinations like RGB and $\bar{R}\bar{G}\bar{B}$, or $R\bar{R}$, $B\bar{B}$, and $G\bar{G}$. (This color analogy is best understood by recognizing that red, blue, and green light combine to produce white light, which is 'colorless'). Three-quark states are referred to as baryons, and quark-antiquark states as mesons.

Invariance under a $\mathrm{SU}(3)_{\mathrm{color}}$ gauge transformation requires the introduction of eight massless gauge bosons called gluons, each of which has its own color-plus-anti-color signature. The $\mathrm{SU}(3)_{\mathrm{color}}$ gauge symmetry is also non-Abelian, such that these gluons interact with one another. This has interesting effects on the coupling constant which describes the strength of QCD interactions: essentially, this coupling constant grows with the distance between two colored particles, such that they can never break free from one another and exist in an independent state. This is in marked contrast to the electromagnetic force, in which the strength of attraction between two charged objects falls off with distance.

This property of colored objects—their inability to exist in an independent state—is called confinement. To move two colored objects further away from one another requires an ever-increasing amount of energy; when the energy within the color field between the two objects reaches a critical level, a quark-antiquark pair is created from the field. These quarks are, in turn, now attracted—both to each-other, and to the two original objects. This process repeats, until all colored objects have bonded with other colored objects to create colorless, bound states. This is called *hadronization*, and it happens extremely quickly (on the order of 10^{-23} seconds). The result is that colored objects, such as free quarks or gluons, are never directly observed. What we have to work with are colored, bound states—baryons and mesons—from which we must piece together the properties of the original, colored objects.

For high-energy interactions, the effect of hadronization can be huge: a single quark-antiquark pair produced in a $p\bar{p}$ collision can hadronize into tens or hundreds of other particles. What simplifies the process somewhat is that hadronization must not violate energy and momentum conservation; if a quark hadronizes into dozens of other particles, the net momentum of this system must remain the same. As a result, particles that are sourced by a single progenitor quark often cluster together into *jets* which trace a cone-shaped path through space, moving outwards from a vertex centered upon the original quark's production point. Jets are our (admittedly, sometimes blurry) magnifying glass for studying QCD interactions, and we will describe them in more detail in Chapter 3.

1.4 Perturbation theory and heavy flavor

Quantum chromodynamics is a fascinating subject, and worthy of (several (hundred)) theses dedicated solely to the subject. Our analysis, however, is more concerned with how QCD is modeled. The complexity of QCD interactions, and especially of the process of hadronization, makes this a challenging question. Models of QCD are generally split into two factions: perturbative methods, and non-perturbative methods.

Perturbation theory is a technique for solving complex systems: an exact solution is found for a simplified case, and the complicated parts are treated as small perturbations to this exact solution. In models of the

SM, what we solve for is the amplitude of a given interaction—say, $u + \bar{d} \to W + g$. (Amplitude is a quantity that is related to probability, and we will elaborate on this in just a moment.) An exact solution is found by treating the system of particles as non-interacting; then, interaction terms are added in as a perturbation. The result is a power-law-like expansion in terms of a coupling constan, α, which describes the strength of that interaction. The value of α depends upon the type of interaction (strong, weak, electromagnetic), and also—perhaps most pertinently for the strong force—the interaction's energy and distance scale. As long as α is small, this perturbative approach holds. Perturbation theory is, however, perhaps easier discussed in terms of Feynman diagrams.

Feynman diagrams are cartoon representations of each term in a perturbation series. Each term is phrased as a particular path by which an interaction can take place; for our $u + \bar{d} \to W + g$ example, was the gluon emitted by the u or by the $\bar{d}$? Was a virtual gluon emitted and reabsorbed by the u prior to its fusion with the d? Each term represents one of these possibilities. Feynman diagrams are a way to visualize these options, and to determine which processes contribute to your interaction, without struggling through enormous sets of integrals. Just as important, there are a set of Feynman *rules* for translating these diagrams into mathematical terms; these can then be added together to provide the amplitude for your interaction. In sum, Feynman diagrams are an excellent shorthand for discussing interactions in particle physics. A few examples of these diagrams are shown in Figures 1.1 and 1.4. Once the Feynman rules have been applied to all Feynman diagrams representing a process (to some order), all terms are added together to get the amplitude for that process. A probability (more accurately, a cross-section, though the core meaning is very similar) can be obtained from the square of the amplitude, integrated over the phase space of interest.

Interactions between particles in a Feynman diagram take the form of vertices, and the Feynman rules require introducing a factor of α for each vertex; as said earlier, this value depends upon the interaction. For $\alpha \ll 1$, the lowest-order diagrams—those with the fewest vertices—will therefore have the most influence on your final result. This is extremely powerful, as higher-order terms are generally far more numerous than lower order terms—as the number of vertices involved increases, the number of ways in which the diagrams can be drawn increases even faster. As long as these higher-order terms do not propogate *too* quickly (we discuss one counterexample in the next section), we can achieve reasonably accurate predictions using only the lower-order terms. This is the essence of perturbation theory.

Perturbation theory has been used to great success in electroweak physics, for which $\alpha_{\text{EM}} \sim 1/137 \ll 1$. As a particularly stunning example, the anomolous magnetic moment of the electron as predicted by the SM has been experimentally confirmed to an accuracy of $1 : 10^9$. However, the coupling constant α_{QCD} can be quite large, and for low-energy interactions—specifically, for energy scales near $\Lambda_{\text{QCD}} \sim 200$ MeV—α_{QCD} can approach order 1. Here, perturbation theory fails completely. There are techniques for dealing with

these types of interactions (for example, lattice QCD calculations [6]); however, another way to circumvent this problem is to explore those certain regimes of QCD physics in which perturbation theory holds. This is where heavy flavor quarks (c, b, and t) become very useful.

The lightest of the heavy flavor quarks, the charm quark, has a mass of ~ 1.5 GeV. Pair production of $c\bar{c}$ therefore requires an interaction energy of at least 3 GeV $\gg \Lambda_{\rm QCD}$. Pair-production of bottom quarks ($m_b \sim 5$ GeV) lies even further beyond the $\Lambda_{\rm QCD}$ threshold, and the production of top quarks further still. Each of these heavy flavors has its own particular advantage.

The enormous top quark mass (~ 175 GeV) means that non-perturbative effects should have an extremely small effect on its production cross-section; in addition, the top quark decays so quickly (by the weak process $t \to W + b$) that it does not have time to hadronize. (Hadronization, a messy process with many low-energy quarks and gluons produced in close proximity to one another, is a prime example of non-perturbative physics.) However, the top quark production rate is much lower than that of bottom and charm quarks, which makes it difficult to obtain statistically-significant samples.

The bottom quark is produced much more frequently in hadron colliders than the top quark, and it decays much more slowly. This provides an important handle on jets of particles produced by a bottom quark. When a b hadronizes into a semi-stable meson or baryon, this bound state will in general travel some macroscopic distance through the detector volume before decaying. The result is a *displaced vertex*, in which several particles seem to originate from a point that is not centered upon the original collision point. With a high-resolution tracking chamber (such as that at Fermilab's CDF detector), it is possible to detect these displaced vertices and to use them to identify bottom quark production processes. Because of its relatively high abundance and this tell-tale displaced vertex, bottom quark production is an extremely common test of perturbative QCD.

Charm quark production lacks several of the 'advantages' possessed by its higher-mass, heavy flavor brethren. For one, charmed mesons decay relatively quickly, due in part to the much larger CKM mixing matrix element $|V_{cs}| \sim 0.97$ (to be compared with $|V_{cb}| \sim 0.04$). As a result, displaced vertices are not as strong of a handle for identifying charm production. In addition, the lower mass of the charm quark means that perturbation theory is less accurate of a technique for describing $c\bar{c}$ production, than it is for describing $b\bar{b}$ production. This can, however, also be looked at as an advantage—by straddling the boundary between perturbative and non-perturbative QCD, charm quark production allows for very stringent tests of perturbation theory, and may expose its weaknesses (or validate its strengths). Furthermore, the lighter mass of charm quarks means that they are produced in greater abundance than any of the other heavy flavor quarks. As such, while the lack of displaced vertices can make it difficult to identify charm, we also have greater statistics to work with.

1.5 Factorization, parton showers and NLO theory

We have discussed generally the usefulness of heavy flavor in testing the performance of techniques involving perturbative QCD. However, there are additional processes and properties that can be tested by probing the heavy flavor regime, all of which stem from observed variations from perturbative predictions. Specifically, the observed cross-section for $b\bar{b}$ and $c\bar{c}$ production is generally higher than that predicted by next-to-leading order (NLO) perturbative QCD. There are three major contenders for an explanation.

One, is that higher-order terms must be included in order to produce an accurate prediction of heavy flavor production. Despite the topically perturbative energy scale of, *e.g.*, $b\bar{b}$ production ($10\ \text{GeV} \gg \Lambda_{\text{QCD}}$), the leading-order calculation of this cross-section is two to three times smaller than the next-to-leading order cross-section [7]. The reason for this lies in the propagation of higher-order terms, as mentioned in the previous section. At lowest order (two vertices, α^2_{QCD}), $b\bar{b}/c\bar{c}$ production is brought about by quark-antiquark annihilation or gluon-gluon fusion. However, at the next highest order (α^3_{QCD}), the dominant form of $b\bar{b}$ production comes from appending a $g \to b\bar{b}/c\bar{c}$ vertex onto some lower-order QCD process (Figure 1.1). With the cross-section for $gg \to gg$ of order one hundred times larger than that for $gg \to b\bar{b}/c\bar{c}$, this means that even with an extra suppressing factor of α_{QCD}, the contribution of $gg \to gg$-type interactions to the heavy-flavor production cross-section is comparable to that of the leader-order term. By extension, it is not unreasonable to assume that yet higher-order terms might explain the observed discrepancy.

Another option is the possibility of unknown, beyond-the-standard-model (BSM) production processes which give rise to heavy flavor. Supersymmetry is one popular option. Supersymmetry is an extension to the standard model in which each fermion is paired with a supersymmetric partner boson, and each boson with a supersymmetric partner fermion. Quite unfortunately, these supersymmetric partners are labeled by prefacing the name of a SM fermion with the letter 's', and awkwardly forcing the suffix '-ino' into the names of SM bosons. As an example, the supersymmetric partner to the bottom quark is called the sbottom, and the partner of the gluon, the gluino. Under these theories, the pair-production of gluinos, each of which decays into a bottom and sbottom, may lead to an excess of bottom quark production [8]. Through bottom quarks decaying into charm (or by direct charm production through some other non-standard model process), an excess of charm could also be observed. Given the excesses that we do observe, it is possible that some new physics process (supersymmetric or otherwise) is responsible.

Finally, a third explanation points towards models of the hadronization process by which colored objects become colorless mesons and baryons. This process is also sometimes referred to as fragmentation. An improvement to what was (at the time) a leading model of fragmentation was used to improve the discrepancy between theory and experiment for b production from a factor of 3, to a factor of 1.7 [9]. This is a figure

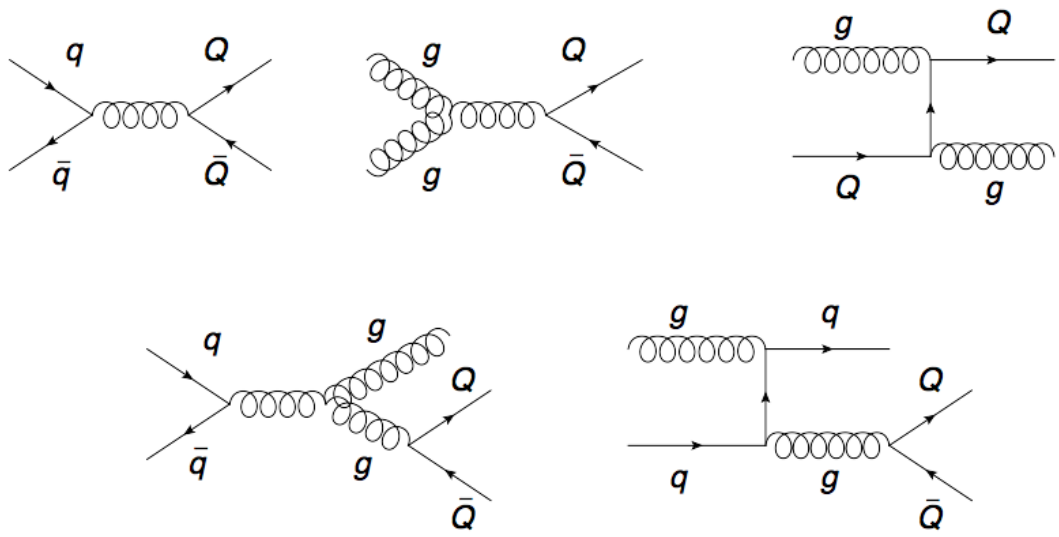


Figure 1.1: At lowest order, $b\bar{b}$ or $c\bar{c}$ production (labelled generally above as $Q\bar{Q}$) comes about through quark-antiquark, gluon-gluon, or quark-gluon scattering (top row). The latter process is the least common source of $Q\bar{Q}$ production in $p\bar{p}$ collisions, as the proton sea content contains very At next-to-leading order (NLO), $Q\bar{Q}$ production comes from pinning a $g \to Q\bar{Q}$ production vertex onto some lower-order QCD process (bottom row).

that holds to this day in such popular theoretical software as Pythia.

Fragmentation is a non-perturbative process that must be described phenomenologically. The standard scheme for doing so is called factorization: the problem is split into hard and soft processes, and the hard (high-energy) interaction is solved perturbatively, while the soft, non-perturbative piece is modeled off of experimental results (*e.g.* from the cleaner environment of e^+e^- colliders). The major assumption that lies behind factorization is simply that low-energy interactions do not have a significant effect on hard QCD interactions within the same collision.

A leading factorization-based approach is the parton shower model, which forms the basis for Monte Carlo programs such as Pythia [10] and Herwig [11]. The parton shower model begins by simulating the hard-scatter event of interest (*e.g.*, $u + \bar{d} \to W + g$), using perturbative QCD. Then, it adds initial state radiation (ISR) and final state radiation (FSR) to the interaction; this radiation is perhaps best described as gluons and photons which are radiated away by incoming and outgoing particles. Incoming and outgoing gluons and photons are also allowed to split into quark-antiquark, gluon-gluon, or lepton-antilepton pairs. This is all accomplshed using a probabilistic model, tuned to experiment. The result is a full simulation of high-energy collisions to extremely high order (with many final-state particles), without having to explicitly draw out each Feynman diagram and calculate each term.

An important final note about Monte Carlo parton shower models is that many (including Pythia) use

only leading-order predictions for the hard-scattering process in heavy flavor production. The reason for this is straightforward; the initial and final state radiation, by its nature, will simulate the NLO production processes that are accomplished by adding a $g \to b\bar{b}/c\bar{c}$ vertex onto a lower-order diagram. Explicitly adding in such production mechanisms through NLO production processes can result in double-counting. It is therefore important to take great care when mixing NLO perturbative calculations with the parton-shower framework. For the Pythia Monte Carlo models used in this thesis, we use only the predictions of lowest-order perturbative QCD paired with parton-showering.

1.6 Parton distribution functions

A final aspect of high energy collisions which we have not yet touched upon, is parton distribution functions (PDFs). The proton is nominally a bound state of two up quarks, and one down quark. However, the huge binding energy between these three quarks causes the color field around them to writhe with activity. The constant strong interactions between these three quarks are mediated by the exchange of gluons, which can split into temporary quark-antiquark pairs before recombining. The net result is that the proton is more accurately described as a bag of quarks and gluons: the "valence" quarks (the aforementioned u, u, and d) and the "sea" (everything else). When a proton and antiproton collide, the hard interaction may be between either two valence quarks, one valence quark and one sea parton, or two sea partons. The energy and momentum that is carried by each type of parton within the proton is described by the proton's PDF.

The PDF is a probabilistic distribution describing the fraction of energy, x, held by some parton, i, times that parton's average population, $x \cdot f_i(x)$. In essence, this quantity is the probability that in some given interaction with energy scale Q, a parton of type i interacts with energy $x \cdot Q$. The PDF is plotted as a function of x, and an example proton PDF is shown in Figure 1.2. At high x, the proton PDF is always largest for the valence quarks, as we would expect—these are the dominant 'source' of quarks in the proton, and are more likely to hold a large fraction of the proton's energy. However, the shape and size of the PDF for each parton is dependent upon the collision's center-of-mass energy scale, Q^2, and as Q^2 increases, the overwhelming source of hard interactions is that of sea-quarks and gluons with small x. This reflects the standard trend of ever-higher energy transfers probing ever-smaller regions of space, in which the sea quarks become far more prevalent. It is interesting to note that the gluon PDF is enormously larger than that of any quark flavor at low x; this is simply another reflection of the huge binding energy inherent in the strong force, here manifesting itself as a large gluon population.

There are many models for the proton PDF. Two of the more historically-dominant PDF sets are CTEQ [12] and MRST/MSTW [13], both of which are used in this thesis. To determine the proton PDF from

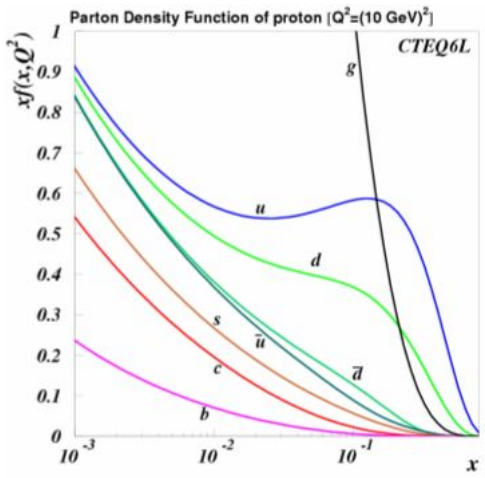


Figure 1.2: The proton's parton distribution function (PDF) at energy transfer scale $Q^2 = 10$ GeV^2, as predicted by PDF set CTEQ6L [12]. Notice that below $x \sim 0.1$, the PDF is dominated by gluons.

first principles is virtually impossible—this is an extremely non-perturbative regime, with huge numbers of low-energy, colored partons in close proximity to one another. Instead, these PDF models use experimental results taken at some particular Q^2 to determine specific pieces of the PDF, and then use perturbative techniques to extrapolate to new regimes.

There are four major types of experiment used to determine the proton PDF. One is deep-inelastic scattering processes (DIS), in which a lepton is scattered from a target containing protons. This is far cleaner than colliding two hadrons, and the final states produced provide insight into the nature of the proton. Lepton pair production also helps, in which a quark and its antiquark annihilate to produce a lepton-antilepton pair; in proton-proton collisions, for example, all such events must involve at least one sea quark.

Lepton charge asymmetry is a useful probe in $p\bar{p}$ collisions: the W bosons produced in the interaction, and their angular distribution, can provide insight into the nature of underlying quarks. Finally, high-p_T jet production can also contribute to our knowledge of the sea quarks: with each jet theoretically initiated by a single quark or gluon, we can trace jet structure back to the underlying hard-scatter of a hadron collision. For example, jets identified with strange quark production may provide insight into $gs \to gs$ scattering events, in which a strange sea quark is excited by collision with a gluon, and is ejected from the proton.

1.7 Vector bosons plus charm

In previous sections, we described the field of high-energy physics in great generality, and closed on three major concepts concerned with heavy flavor production. First, that heavy flavor production falls into the

regime of perturbative QCD predictions; second, that deviations from perturbative predictions may indicate a flawed model or new physics; and third, that heavy flavor production may also provide insight into the proton PDF. Identifying 'inclusive' heavy flavor production—charm, bottom, or top quark production by any means—is certainly a viable means for exploring these questions. However, several unique advantages can be gained by looking specifically at vector boson prodiction in association with charm quarks, $W/Z + c$.

First, the $W+c$ production process in $p\bar{p}$ collisions is sensitive to the magnitude of the Cabibbo-Kobayashi-Maskawa (CKM) matrix element V_{cs}, and also to the strange quark distribution function of the proton. To first order, $W + c$ production comes about through 'quark-gluon fusion', in which a strange or down quark absorbs energy from a gluon, and decays into a W boson and charm quark (Figure 1.4). While the strange quark PDF is smaller than the down quark PDF within the proton, CKM matrix element $|V_{cd}| \sim 0.225$ is sufficiently suppressed with respect to $|V_{cs}| \sim 0.973$ that most $s(d) + g \rightarrow W + c$ events come about through strange quark/gluon fusion (about 80%) [14, 15]. While the s-quark PDF at momentum transfer scale $Q^2 \sim 10$ GeV^2 and momentum fraction $x \sim 0.1$ has been determined by neutrino-nucleon deep inelastic scattering experiments [16, 17], the interpretation of this data is model-dependent. Some analyses suggest that the s-quark sea is suppressed with respect to the d-quark sea at all values of x [18, 19, 20], while others suggest that all light flavor sea quarks approach equal population as x decreases [21]. With $Q^2 \sim m_W^2$ for the production of a W boson in association with a charm quark, it is possible in $W + c$ events to probe this smaller x regime and to test these models in more detail.

A recent ATLAS search to identify $W+c$ production in pp collisions provided results with a preference for the symmetric s and d sea quark hypothesis [22]. In particular subsamples, this analysis was able to probe down to a charm quark momentum of 8 GeV. However, the uncertainty in this symmetry measurement was large ($\sim 30\%$), and corroborating or contradictory results would be well-received. The Tevatron has also investigated $W + c$ production. CDF measured the $W + c$ cross-section for ($p_T(c) > 20$ GeV as an indirect test of the value of $|V_{cs}|$, and also as a means of looking for any excess that might indicate a flaw in the SM prediction [23, 24]. No such excess was found, and all results (including the derived valued of $|V_{cs}|$) were found to agree with world best measurements.

$W/Z+g(\rightarrow c\bar{c})$ production is also an interesting venue for testing the properties of heavy flavor production. While searches for inclusive heavy flavor production (*e.g.*, $c\bar{c}$ or $b\bar{b}$ production with or without a W/Z) is certainly an interesting subject, first identifying a W or Z provides a strong handle on the momentum transfer involved in the collision. This, in turn, allows for more stringent tests of production rates and of perturbative QCD. As described in earlier in this chapter, the rates for inclusive $b\bar{b}$ and $c\bar{c}$ production are observed to be higher by a factor of 1.7 than NLO predictions with improved models of hadronization. This may be due to higher-order terms in the perturbative series, to new physics beyond the standard model,

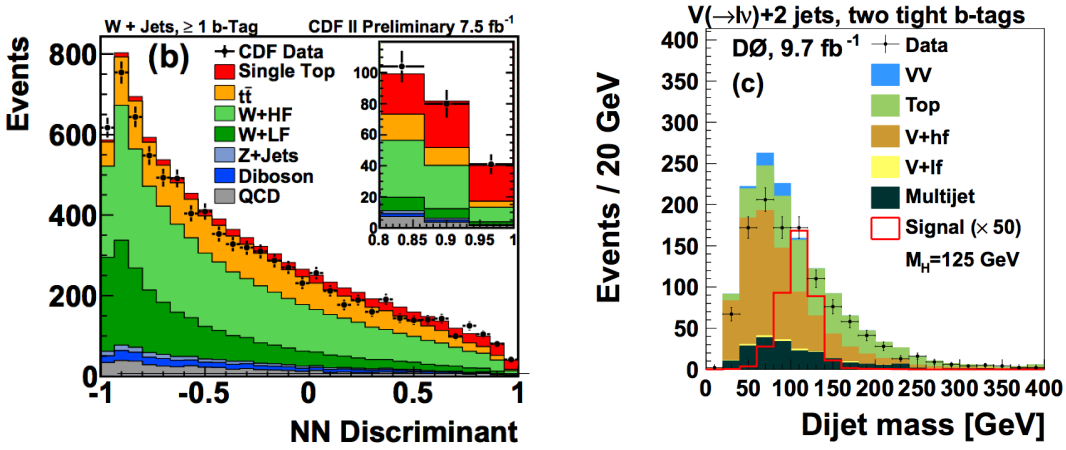


Figure 1.3: Vector boson plus heavy flavor as a source of background to Higgs and single top searches. To the left, a plot of a neural net discriminant used to search for single top decays in CDF events with at least one tagged bottom jet (search is for decay $t \to W(\to \ell\nu) + b$) [26]. To the right, a plot of invariant dijet mass for D0 events with two tightly-tagged bottom jets (search is for decay $p\bar{p} \to W(\to \ell\nu)H(\to b\bar{b})$). In both cases, the contribution of vector boson plus heavy flavor background (W + h.f. or V + h.f.) is seen to be quite sizeable. [27].

or to the need to further improve our models of hadronization. A recent analysis by the D0 experiment at Fermilab tested the (previously unmeasured) rate of charm quark production in association with Z bosons, and found it, too, to be compatible with the 1.7× rate enhancement of $g \to c\bar{c}$ splitting found in other analyses [25].

Finally, the signature of a vector boson plus charm event—one or more charged leptons and/or missing energy, plus heavy flavor jets—is shared by several other searches in high energy physics. Certain decay modes of the Higgs boson, top quark, dark matter candidates, and supersymmetric particles all look very similar to vector bosons plus heavy flavor, such that W/Z plus heavy flavor acts as background to these searches. Some examples are shown in Figure 1.3. To remove this background, these analyses often rely upon the predictions of theory to estimate how many of these background events will be in their signal sample. It is therefore very important, if we are to conduct accurate searches for rare processes, that these models be accurate. Only repeated experiment will tell us if they are.

In short, the applications of $W/Z + c$ physics are widespread. In this thesis, we do not seek to address all of the questions raised above, or even most of them. We are far from the first to recognize that these events are important, and there have been many studies of their properties; however, at the time that we began this work (and as recently as March of this year), there had been no attempt to probe $W/Z + c$ production in hadron colliders for charm momentum less than $15 - 20$ GeV. The reason for this was simple: the standard approach for identifying charm in an event is to search for charmed *jets* ([23, 24, 25]). Jet objects at hadron colliders are generally defined to have transverse momentum greater than $15 - 20$ GeV, which excludes charm

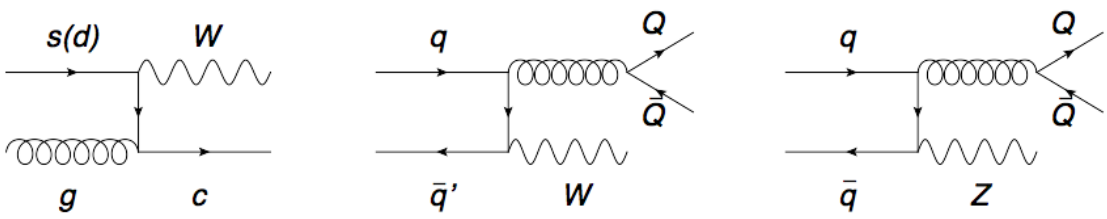


Figure 1.4: First order production processes for charm quark production in association with vector bosons. Production via $g \to Q\bar{Q}$ includes both direct charm production ($Q = c$), and bottom hadron decays (in which $Q = b$, and the b decays into a c).

production below this threshold.

However, if discrepancies between NLO QCD and observation lies in the failure of perturbative QCD, then we might expect for the discrepancy between theory and data to increase at lower charm momentum. In addition, it is possible that some new beyond-the-standard-model process is at work at this momentum scale, which could lead to an observable excess. We therefore propose a new approach for studying $W/Z + c$ production in hadron colliders. Instead of probing jet properties for evidence of a genitor charm quark, we search at the track level for a particular charmed meson decay: $D^*(2010) \to D^0(\to K\pi)\pi$. By fully reconstructing this decay, we find that we are able to identify D^* in W/Z events down to a transverse momentum of 3 GeV—nearly an order of magnitude lower than jet-based approaches (and a factor of 2.5$\times$ lower than a similar analysis presented by the ATLAS collaboration this March [22]).

In this thesis, we present our new approach in more detail, we explore its properties, and we identify and characterize backgrounds in the low-momentum regime that it has opened up for observation. In the process, we make the first observation of $W/Z + D^*$ production (abbreviated very slightly as $V + D^*$ production) at hadron colliders at low charm momentum. We show that full reconstruction allows us to identify up to three individual $W + D^*$ production processes, and that we are sensitive not only to $W/Z + c$ production, but also to $W + b\bar{b}$ production in which one of the resulting bottom hadrons decays into a D^*. Most importantly, we prove that is possible to probe this regime with good accuracy, and we provide a viable technique for future analyses to employ. We conclude by comparing what we have observed to theoretical predictions, in order to see how well theoretical models hold up in a low-momentum charm regime. We also suggest additional ways in which our technique might be used to answer other questions in hadron collider physics.

Chapter 2

The Tevatron Collider and CDF detector

This analysis uses data taken by the Run II Collider Detector at Fermilab (CDF), at the Fermilab Tevatron. The Tevatron was a 1 km radius synchrotron, acting as the final stage in an accelerator chain that brought protons and antiprotons together in collisions with a center of mass energy $\sqrt{s} = 1.96$ TeV. Of the six experimental halls spaced around the four mile accelerator ring, two featured collision points which were monitored by the CDF and D0 detectors during the Tevatron's Run II, from 2001 until 2011. Until the completion of the Large Hadron Collider in 2008, the Tevatron was the world's most powerful collider.

A schematic diagram of the Fermilab accelerator complex is pictured in Figure 2.1. As of the writing of this thesis, both the Tevatron and the CDF detector have been decomissioned. However, in all that follows, we choose to describe the accelerator, the collision process, and all related equipment in the present tense. This is simply to prevent unwieldy clashes with this author's tendency to write all scientific literature in the present-tense.

2.1 The Fermilab accelerator chain and the Tevatron

While nominally more recognizable than other parts of the accelerator chain, the Tevatron is only one part in a sequence of events leading to the production and collision of proton-antiproton beams at Fermilab. The proton beams described in the following sections, in particular, are not only used for $p\bar{p}$ collisions; they are also delivered to test beam facilities, to fixed target experiments, and to neutrino experiments at Fermilab.

2.1.1 Proton production and acceleration

The production of a $p\bar{p}$ collision at $\sqrt{s} = 1.96$ TeV at the Fermilab Tevatron begins with the production of a proton beam.

This beam originates from the Cockcroft-Walton pre-accelerator, which ionizes hydrogen gas (H_2) to create negatively-charged hydrogen ions (H^-). These ions are accelerated through a column that leads from a charged dome housing the hydrogren source, to a grounded wall; in the process, they gain an an energy of

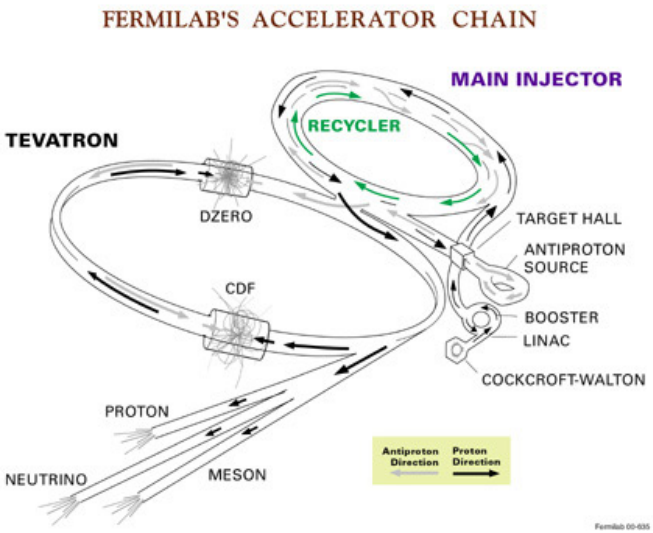


Figure 2.1: A schematic representation of the accelerator chain that is used to collide protons with antiprotons at a center of mass energy $\sqrt{s} = 1.96$ GeV. Not explicitly labeled are the Debuncher and Accumulator, both of which are located in the rounded triangular segment labeled Antiproton Source. Each piece of the chain is described in this section (section 2.1).

750 keV.

The ions are next fed through a transfer line into the linear accelerator (Linac). The Linac is a series of radio frequency (RF) cavities which accelerate the hydrogen ions to an energy of 400 MeV; this beam is focused by means of quadrupole magnets located both inside (the first five stations) and outside (the last seven stations) of the RF cavities.

A transfer line takes the resulting beam to the Booster. The Booster is the first synchrotron in the accelerator chain: a series of magnets arranged around a (75-meter radius) circle, with (19) RF cavities spaced throughout to provide the acceleration. These RF cavities operate at a frequency that is synchronized to the period of the circling protons, providing a constant boost—hence the Booster name. The Booster strips the electrons off of the hydrogen ions, leaving behind only protons, and brings these protons up to an energy of 8 GeV. Finally (at least until the Tevatron step), the Booster sends the beam of 8 GeV protons to the Main Injector (MI), another synchrotron with 18 accelerating cavities and a circumference seven times that of the booster. For beam that is destined for the Tevatron, the MI accelerates the beam to an energy 150 GeV.

Once the protons have reached 150 GeV of energy, they are sent to the Tevatron for the final step.

2.1.2 Antiproton production and acceleration

Antiproton production begins with a 120 GeV proton beam sent from the MI, as described in the previous section. This beam is sent to a target of nickel alloy, and collisions between the protons and the nickel produce a large number of secondary particles. Magnets and electric fields are used to filter off 8 GeV antiprotons from this spray, accelerate them, and send them to a synchrotron called the Debuncher.

The Debuncher is not a 'typical' accelerator, in that it does not increase the energy of the antiprotons. The antiprotons filtered off of collisions with the nickel target have a fairly large spread in momentum—including potentially large oscillations in the transverse plane—and this is not desirable for a final, well-behaved beam. To address this problem, the Debuncher uses stochastic cooling to reduce the transverse energy of the anti-protons, reducing this spread in momentum and producing a 'cooler', more well-behaved beam. This stochastic cooling works by detecting electrical signals caused by fluctuations in the momentum of groups of antiprotons, and providing a stimulus to move the momentum behavior of each group towards some average. This 'cools' the beam in the sense of reducing entropy, producing a more well-ordered state.

This cooled beam is then sent to another synchrotron, the Accumulator, which is housed in the same ring as the Debuncher. Here it is cooled further and temporarily stored. Antiprotons cannot be produced at nearly the same rate as high-momentum protons (in fact, it takes about one million protons hitting the nickel target to produce 20 antiprotons). Therefore, a stack of anti-protons must be built up over the course of many hours, before being injected into the Tevatron for collisions—in the meantime, they must be stored somewhere, and the Accumulator is used for part of this storage. The Accumulator is not the final location for storing the antiprotons, however, and they are next sent to the Recycler.

The Recycler, another synchrotron which is housed in the same ring as the MI, holds the antiprotons until the Tevatron is ready for a new beam store. In the meantime, it also cools the antiprotons yet further. It does this via stochastic cooling as previously described, but also through 'electron cooling', in which a cool, well-behaved (low transverse momentum) beam of electrons, at the same energy as the antiprotons, is overlaid with the antiproton beam. Interactions between the electrons and antiprotons bring the antiprotons even closer to a cooler, ordered state, allowing for a neater and more compact beam.

When the Tevatron is ready for a new beam store, antiprotons are sent from the Recycler to the Main Injector, where they are circulated in a direction that is opposite that of the proton beam. Here, the antiprotons are accelerated to 150 GeV of energy, and then sent to the Tevatron.

2.1.3 Tevatron

The Tevatron is the final destination for protons and antiprotons that are headed for collision. They are circulated in the Tevatron as beam 'stores', which consist of 36 'bunches' of each protons and antiprotons. One bunch contains about 3×10^{11} particles; each bunch is spaced 396 ns apart from the next, and the bunches are arranged into three groups of twelve, with 'abort gaps' between each group. These bunches are injected into the Tevatron by the MI, and ramped from 150 GeV up to 980 GeV, their final collision energy.

The performance of a particle collider is measured by instantaneous luminosity, defined as:

$$\mathcal{L} = \frac{fBNN_{\bar{p}}}{2\pi(\sigma_p^2 + \sigma_{\bar{p}}^2)}$$

Here, B is the number of bunches in each beam, $N_p(N_{\bar{p}})$ is the number of protons (antiprotons) in each bunch, $\sigma_p(\sigma_{\bar{p}})$ is the root-mean-square proton (antiproton) beam widths at the interaction point, and f is the bunch revolution frequency. Luminosity is usually defined in parallel with cross-section σ. For some given interaction $p\bar{p} \to X$, the cross-section for that interaction is related to luminosity as:

$$R(p\bar{p} \to X) = \sigma(p\bar{p} \to X) \times L$$

where R is the rate at which the interaction $p\bar{p} \to X$ takes place. The two quantities σ and L only have a quantifiable meaning when used in tandem. L is, roughly, a measurement of the number of $p\bar{p}$ collisions that take place per unit time, while integrated luminosity is a measure of the total number of $p\bar{p}$ collisions that have been produced over some finite range of time. Plots of the Tevatron peak and integrated luminosity records are shown in Figure 2.2.

These $p\bar{p}$ collisions take place at interaction points inside of the two detectors, CDF and D0, where proton and anti-proton bunches are made to intersect over a region of length on the order of one meter. Each crossing results in only a few collisions, and the bunches remain intact over the course of many crossings, with beams kept focused by magnets spaced around the Tevatron's circumference. As such, a single store can be circulated for many hours, either until the proton and antiproton bunches have been depleted to the point that a new store is desired (for maximum integrated luminosity), or until there is cause to dump (abort) the beam. This is done if the beam becomes unstable for some reason (*e.g.*, if there is a lightning strike somewhere in the vicinity of the Tevatron). If not aborted, an unstable beam could damage sensitive equipment in either the detectors, or the Tevatron ring itself. To prevent this, an unstable beam can be directed into a high-density target called the 'dump'. The previously-mentioned 'abort gaps' provide time for the Tevatron's magnets to set up a field that will reroute the beam into the dump.

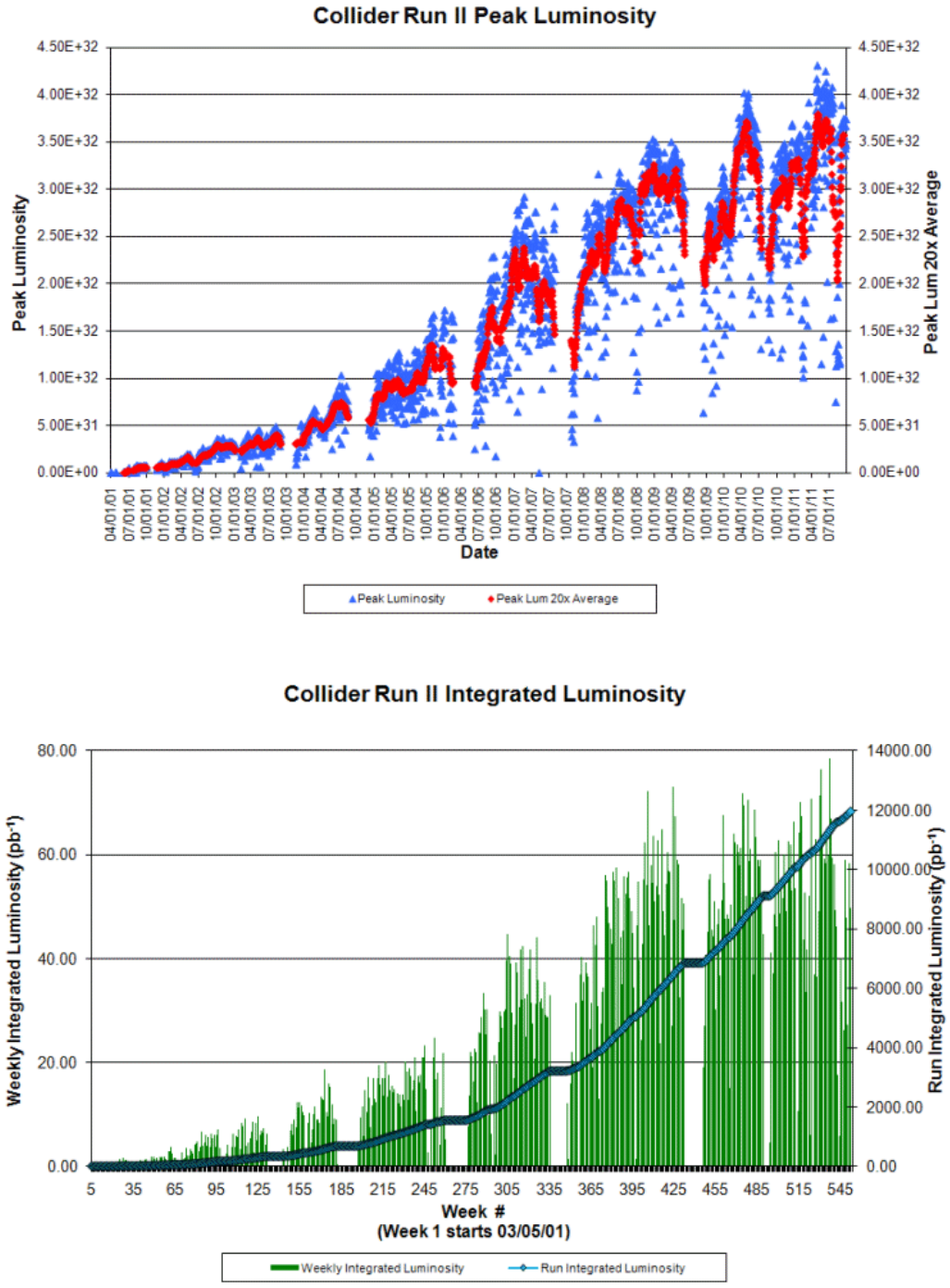


Figure 2.2: Peak (top) and integrated (bottom) luminosity at the Tevatron. The peak luminosity plot shows an increase with time that reflects continual improvements in the Tevatron's performance. The integrated luminosity plot also reflects this improvement, in the form of an accelerating rate of luminosity delivered to the two detectors, CDF and D0. Out of the ~ 12 fb^{-1} of integrated luminosity delivered to each detector, each recorded ~ 10 fb^{-1}—an efficiency of over 80%.

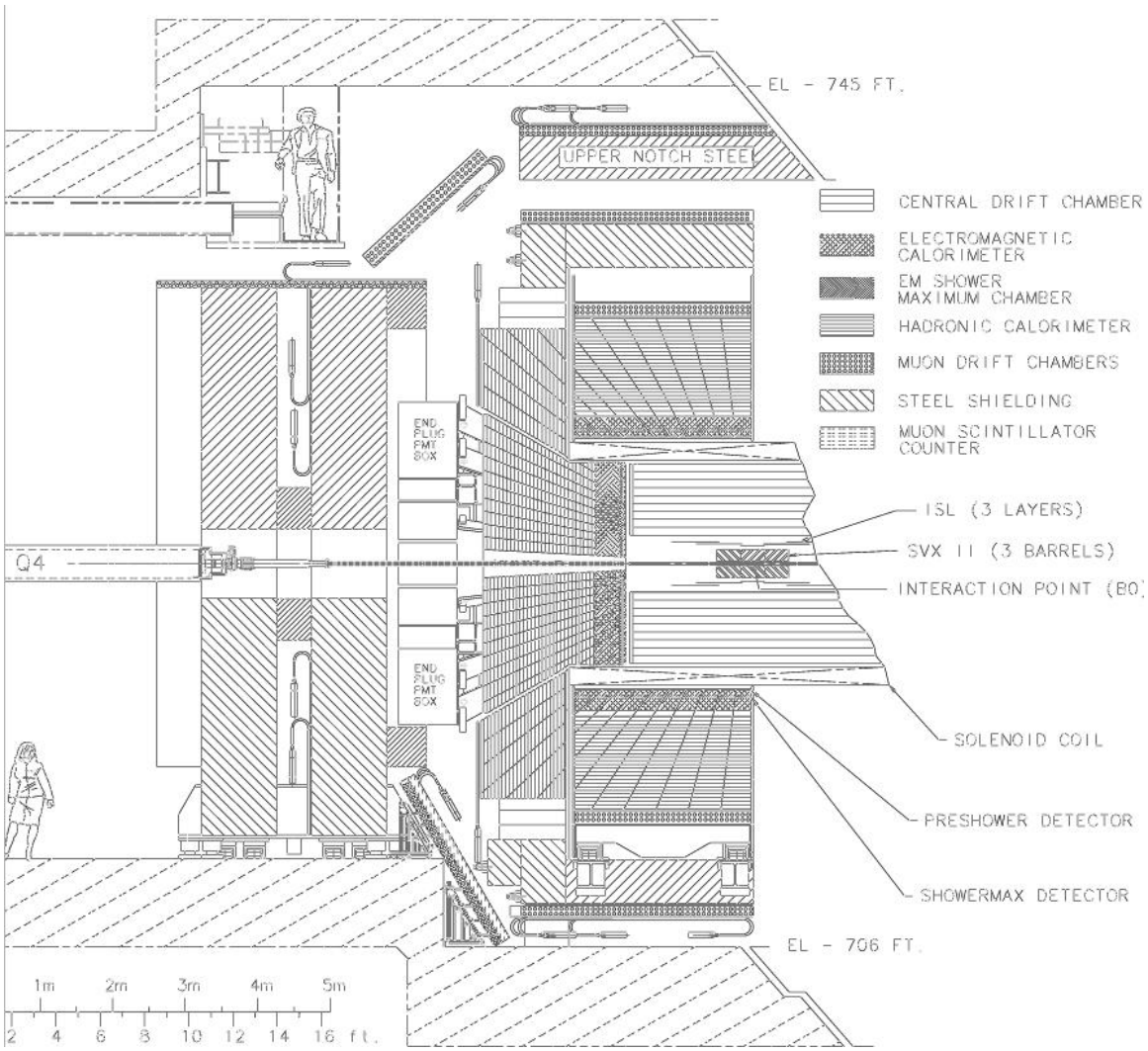


Figure 2.3: A longitudinal, top-down, schematic cross-section of the CDF detector for Run II,.

2.2 The CDF Detector in Run II

The CDF is a multipurpose detector, with tracking, calorimetry, and muon systems that make it possible to identify a vast range of collision by-products. We provide an overview of the detector in this section, with an emphasis on the subsystems that we use to identify objects in our analysis. A diagram of the CDF is displayed in Figure 2.3. A summary of detector signatures for some common physics objects is shown in Figure 2.4.

2.2.1 Coordinate system

The CDF detector is best described using different coordinate systems for different purposes. A standard Cartesian coordinate system is defined with $\hat{z}$ pointing in the direction of motion of the protons, $\hat{x}$ pointing

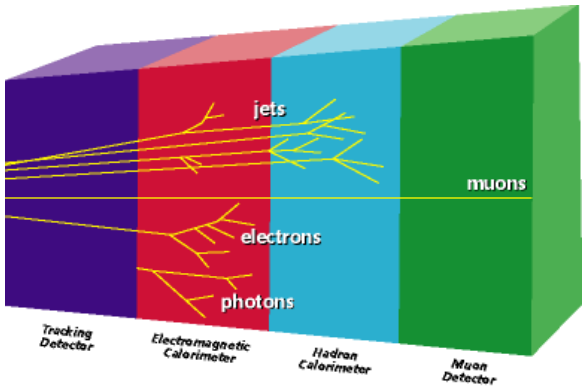


Figure 2.4: A cartoon displaying the signatures of four common colllision event by-products: muons, electrons, photons, and jets. Each type of object interacts with a different subset of the detector systems, allowing unique identification.

radially outwards from the center of the Tevatron ring, and $\hat{y}$ pointing vertically upwards up towards the sky.

In many applications, it is easier to describe the detector using a cylindrical (r, ϕ, z) or spherical (ρ, ϕ, θ) coordinate system, which better reflect the detector's (approximately) axial symmetry. In the cylindrical case, $\hat{z}$ is identified with the longitudinal direction, and the $x-y$ plane is used to describe transverse motion; azimuthal angle ϕ is measured with respect to $\hat{x}$. In spherical coordinates, ρ and ϕ describe the transverse plane; ϕ is once more measured with respect to $\hat{x}$), while polar angle θ is measured outwards from the $\hat{z}$ axis.

Another important coordinate is rapidity, y, defined as

$$y \equiv \frac{1}{2} \ln \left[\frac{E + |\overrightarrow{p}| \cos\theta}{E - |\overrightarrow{p}| \cos\theta} \right]$$

Rapidity is a Lorentz-invariant quantity which is constructed such that the number of particles produced per unit rapidity, dN/dy, is approximately constant. The number of particles per unit polar angle, $dN/d\theta$, is not constant due to variation in the $\hat{z}$ -momentum of incoming partons; while the protons (or anti-protons) in any given bunch each have the same average energy (980 GeV), there is variation at the parton level; even for two protons with identical energy and momentum, the momentum of an individual parton is a probabilistic fraction of the proton's total momentum. These probabilities are described by a proton's parton distribution functions (PDFs), as described in Chapter 1.

In practice, finding both E and $|\overrightarrow{p}|$ in order to calculate y requires knowledge of a particle's mass. To avoid this requirement, it is noted that in the relativistic limit $pc \gg mc^2$, rapidity can be reduced to the simpler 'pseudorapidity', η.

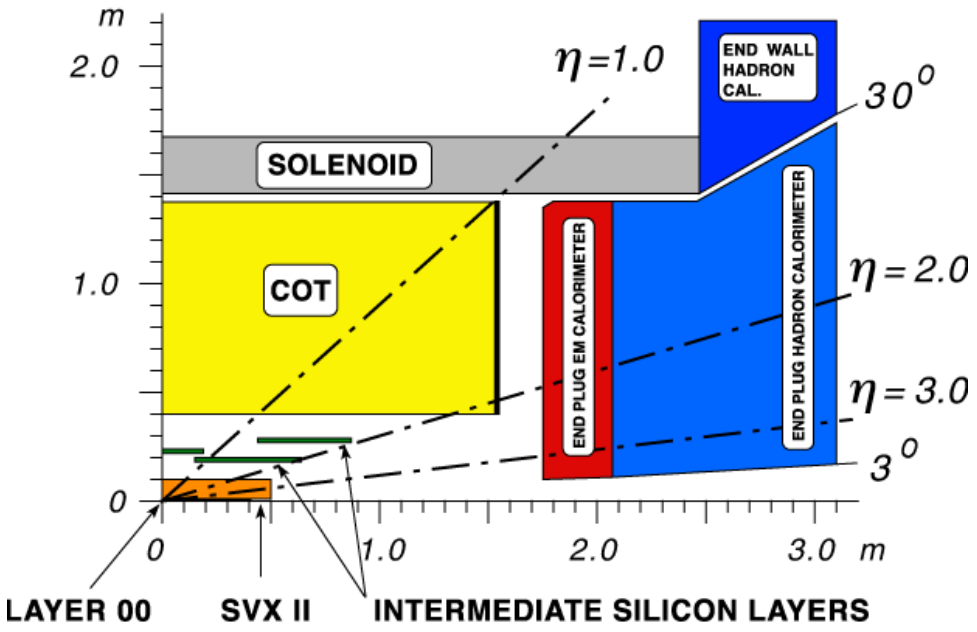


Figure 2.5: A reduced schematic of the CDF detector, emphasizing the tracking systems and a visual reference for coordinate η.

$$y(pc \gg mc^2) \simeq \eta \equiv -\ln\tan(\theta/2)$$

where $\tan\theta$ is defined in terms of the particle's transverse (p_T) and longitudinal (p_z) momentum, $\tan\theta = p_T/p_z$. Pseudorapidity is generally defined with respect to the collision point. However, when describing detector components, a variant definition η_{det} may be used; η_{det} is defined with respect to the center of the detector, $x = y = z = 0$. High $|\eta_{det}|$ parts of the detector are referred to as the *forward region,* while low $|\eta_{det}|$ parts are referred to as the *central region.* The detector is most sensitive in the central region, which is usually detects the by-products of the highest-energy collisions; incoming partons have $p_T = 0$, on average, such that a large value of p_T in a final state object requires a large transfer of energy and momentum.

2.2.2 Tracking

The CDF tracking system consists of two major parts: the smaller, three-part silicon tracking system near the beampipe—which provides high-resolution vertex reconstruction—and the larger central outer tracker (COT). These tracking subdetectors are immersed in a 1.4 T magnetic field running parallel to the beamline. This field is provided by a superconducting solenoid that surrounds the COT. The tracking subsystems are shown in the reduced schematic of the COT displayed in Figure 2.5.

The Silicon Tracking Systems

The three layers of the silicon system extend from the edge of the beampipe, at 1.35 cm, out to a radius of 29 cm. They consist of finely-segmented strips of silicon, with single hit resolution on the order of $6-10\ \mu m$. This resolution allows high precision reconstruction of collision vertices with a resolution $\sim 20\mu m$, which is useful for identifying displaced vertices from heavy flavor decays. The silicon detectors also have a low threshold for signal ($\sim$ 3 eV), compared to the gas ionization energies ($\sim$ 1015 eV) needed to produce a signal in the COT drift chamber.

The silicon system begins with Layer 00 (L00), a set of single-sided silicion detectors which are mounted directly on the beampipe, and which provide coverage for the region $|\eta| < 4.0$. Beyond this, from $r = 2.5-10.6$ cm, is the SVX II; this covers the region $|\eta| < 2.0$ with five double-sided layers of silicion microstrips, arranged on three cylindrical barrels. Regions with $2.0 < |\eta| < 4.0$ are partially covered by the SVX II (see Figure 2.6). The microstrips on the outer side of each layer are set at an angle with respect to the inner side, in order to provide stereo resolution. Finally, the Intermediate Silicon Layers (ISL), from $r = 20 - 29$ cm, cover the region $|\eta| < 2.0$. Like in the SVX II, each of the three layers comprising the ISL are double-sided, with the sensors on each side set at a different angle in order to provide stereo measurements [28].

Together, these three components provide high-resolution measurements of charged particle tracks near the collision point. This allows for the reconstruction of collision and decay vertices, which is particularly useful for identifying heavy flavor decays. We will take advantage of this in Chapter 5 when we reconstruct our $D^* \to D^0(K\pi)\pi$ decays in order to identify charm.

The Central Outer Tracker

The central outer tracker (COT) is an open-cell drift chamber: a cylindrical, 1.3 m shell with filled with an argon-ethane gas mixture, through which 30,240 sense wires are strung. Approximately half of these sense wires are axial (run parallel to the z axis), while the other half are "small-angle stereo", angled 2^o with respect to the z axis. These sense wires are arranged in wire planes, and within each wire plane the sense wires are alternated with potential wires that act as an anode. The wire planes are positioned opposite field sheets that act as a cathode. This produces a drift field through which charged particles can be attracted to the sense wires, registering a signal. These wire planes are arranged into 96 layers, which are further divided into 8 superlayers (4 axial, 4 stereo) [29].

When a charged particle moves through the COT, it ionizes the gas through which it passes. Electrons are drawn toward the sense wires, producing 'hits' that trace out the path taken by the charged particle. Track algorithms can later connect these hits and reconstruct the path. The path that a charged particle

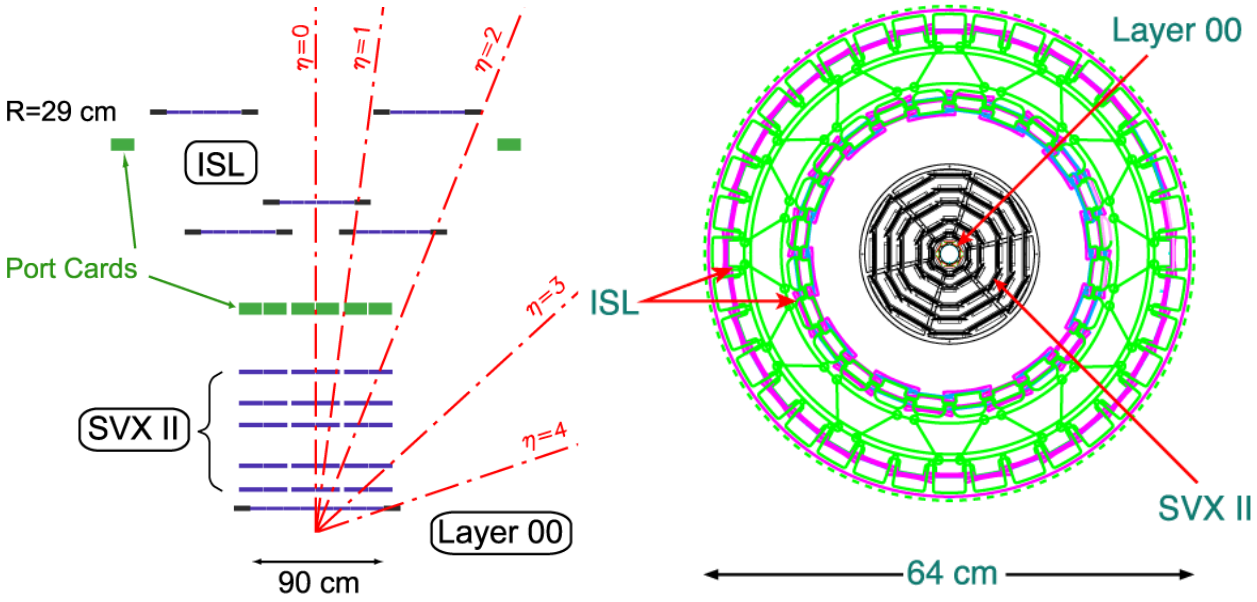


Figure 2.6: Cartoons representing longitudinal (left) and transverse (right) cross-sections of the silicon detector. The longitudinal view (presented in the $r-z$ plane) provides a nice reference for the coverage in η that is provided by each piece of the silicon system. The transverse view provides a 'beamline' view of the silicon systems, with L00 in red, SVX II in black, and ISL in pink; the remaining pieces are supports and infrastructure.

takes as it moves through the COT depends upon its charge and its momentum; as such, the COT tracking chamber allows for high-resolution determination of a particle's momentum and charge.

2.2.3 Calorimetry

Beyond the tracking volume and the solenoid lie the calorimeters. As the name would imply, the calorimeters measure the energies—and to some extent, the directions of motion— of particles produced in $p\bar{p}$ collisions. Unlike the tracking volume, the calorimeter system can register both charged and uncharged particles. There are many parts to the calorimeter system, but all parts work in approximately the same way.

General Calorimeter Function

In order to measure energy, each calorimeter consists of sheets of scintillators that are interspersed with thicker layers of high-density absorbing material. When a particle strikes the calorimeter, it interacts with the dense calorimeter material to produce a shower of lower-energy particles. As this shower radiates through the scintillator, any energy that is deposited in the scintillator volume is converted into light, which is then picked up and amplified by photomultiplier tubes to produce an electrical signal [30]. The number of photons produced in these scintillating layers is approximately proportionate to the energy of the initial particle, and

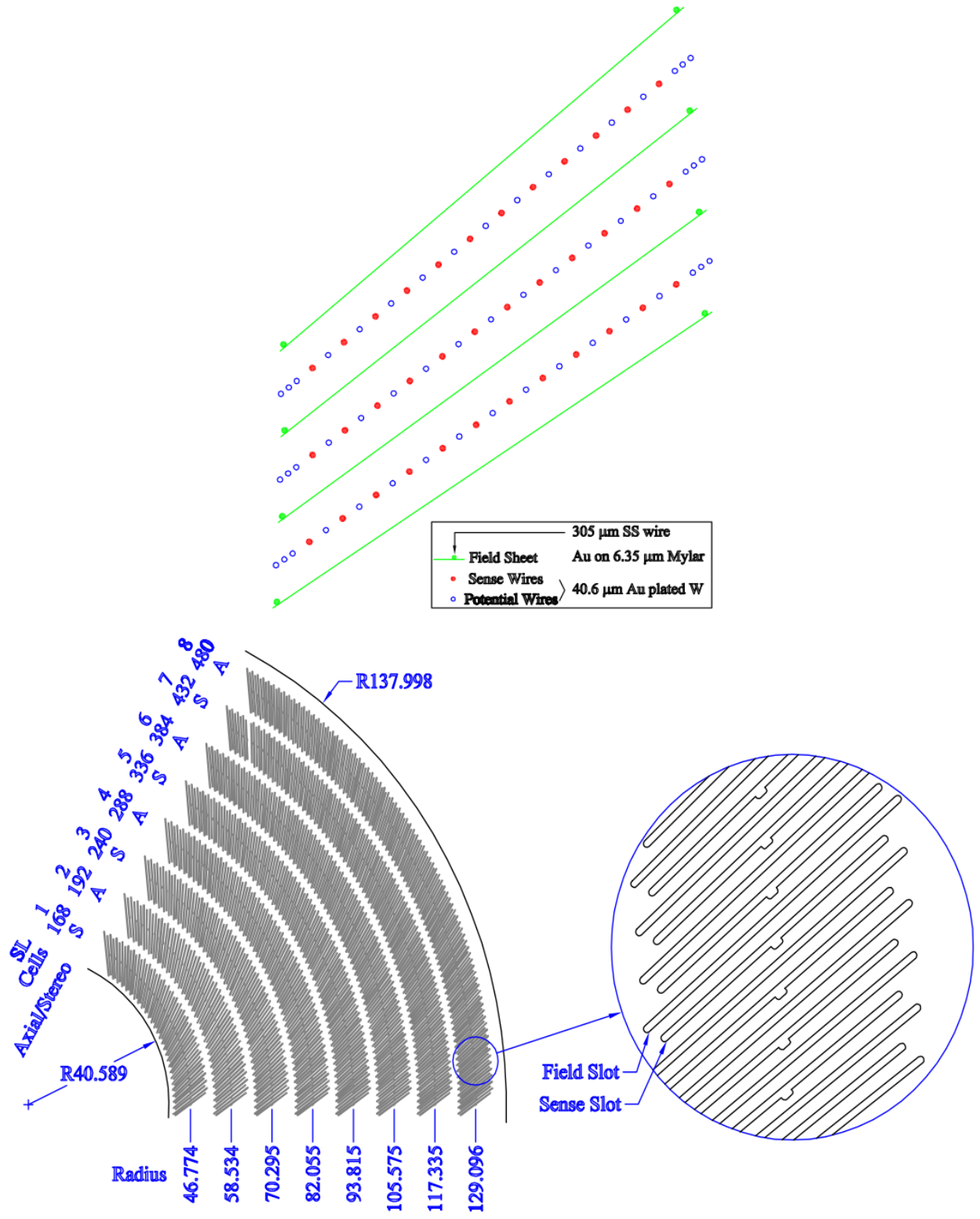


Figure 2.7: Top: cartoon representing the layout of sense and potential wires, arranged opposite field sheets, within the COT. The electric field generated between the anode potential wires and the cathode field sheet brings electrons towards the sense wires, causing 'hits' when a charged particle ionizes the gas near a sense wire. Bottom: the arrangement of wires and field sheets into eight superlayers within the COT; four superlayers are axial, parallel to the z-axis, and four superlayers are small-angle stereo (2^o) with respect to the z axis.

this relationship can be exploited to unfold measurements back to that initial particle energy.

The electromagnetic calorimeters use lead for the high-density absorbing layers, and are designed to measure the energy of electrons and photons. Photons are easily absorbed by these calorimeters, while electrons emit their energy very quickly through bremsstrahlung radiation (which is then absorbed); neither photons nor electrons can penetrate very far. As such, the electromagnetic parts of the calorimeters are located closest to the tracking volume. The hadronic calorimeters, which use iron for their absorbing layers, lie just outside of the electromagnetic calorimeters, and are substantially larger [30]. This reflects the fact that hadrons lose energy by ionization (a small effect) and nuclear interactions (which require collisions with nuclei)—in order for a hadron's energy to be absorbed by the calorimeter, it must travel through a large amount of material.

Calorimeter Specifics

CDF has electromagnetic and hadronic calorimeters—one 'central', and one 'plug'. The central calorimeter consists of the Central Electromagnetic (CEM), Central Hadronic (CHA), and Wall Hadronic (WHA) sections, while the plug calorimeter consists of the Plug Electromagnetic (PEM) and Plug Hadronic (PHA) sections. The CEM is further divided into 'towers' covering the range $|\eta| \leq 1.1$; each tower is an individual unit covering 15° in ϕ and 0.11 in η. The CHA matches the individual towers of the CEM, but only covers the range $|\eta| \leq 0.9$; however, it is supplemented by WHA coverage of the region $0.8 \leq |\eta| < 1.2$. Six radiation lengths into the CEM, a strip-and-wire-chamber Central Electromagnetic Shower (CES) detector provides additional measurements at the depth at which the electromagnetic shower of incoming particles is expected to deposit the largest amount of its energy. The CES can be used to better link particle tracks to calorimeter hits. The Central Pre-Radiate (CPR) wire chambers lie on the inner surface of the CEM, providing early measurements of showers—measuring this early behavior allows for better separation between electrons/photons and light hadrons [30].

The plug calorimeter set consists of the Plug Electromagnetic (PEM) and Plug Hadronic (PHA) sections, which cover the region $1.2 \leq |\eta| \leq 3.6$. As in the case of the central sections, these pieces are split into towers. These towers are clustered by η into 12 groups. The four groups at smaller $|\eta|$ match the central section with 15° divisions in ϕ, while the other eight towers use 8° divisions in ϕ. Again, in parallel to the corresponding central calorimeter piece, the Plug Electromagnetic Shower (PES) lies six radiation lengths into the PEM, and the Plug Pre-Radiate (PPR) likes on the inner surface of the PEM. Other than placement and size, these calorimeters function in essentially the same way as their central counterparts [30].

2.2.4 Muon sub-detectors

Beyond the calorimeters lies the outermost set of CDF sub-detectors: the muon sub-detectors. This is intentional; muons are highly-penetrating particles, and on average they do not lose much energy even to dense materials like lead and iron. They are unique in this sense, which provides a powerful identifier—to first order, any particle that makes it all the way through to the muon sub-detectors is a muon. The muon sub-detectors, then, need only to detect *a* charged particle that has made it through the calorimeter. In practice, there are also fake signatures (for example, high-energy pions can also make their way through the calorimeter), but we will consider those in section 3.4.

Each of the muon sub-detectors that we will describe here consist of groups of single-wire drift chambers—very small versions of the COT. As particles move through the sub-detectors, they leave behind small track signatures, or 'stubs', which can be matched to tracks in the COT in order to provide more information about the particle. The Central Muon (CMU) detector sits directly beyond the CHA, and covers the range $|\eta| \leq 0.6$. The Central Muon Upgrade (CMP) lies beyond both the CMU and a 61 cm layer of steel shielding (which prevents additional non-muon particles from hitting the CMP). It covers the same $|\eta| \leq 0.6$ range as the CMU, but breaks from the (until this point) approximately cylindrical symmetry of the CDF; the CMP takes the form of a rectangular box of wire chambers, which surrounds the rest of the CDF [31]. The CMP is supplemented by the Central Muon Extension (CMX), which covers the region $0.6 \leq |\eta| \leq 1.0$. Finally, the Intermediate/Barrel Muon Detector (I/BMU) covers the region $1.0 \leq |\eta| \leq 1.5$, positioned at an angle that such that it stretches radially from the plug calorimeters towards the CMX (see Figure 2.3) [31].

Together, the muon sub-detectors cover the region $|\eta| \leq 1.5$, and they are very useful for identifying both the leptonic decays of vector bosons, and the semi-leptonic decays of hadrons. As we will see, muon signatures are very clean, and they help to create signals that stand out from background. The coverage provided by each muon sub-detector is summarized in Figure 2.8.

2.2.5 Cerenkov Luminosity Counters (CLC)

Though we will not reference the Cerenkov Luminosity Counters (CLC) again in this thesis, they are an interesting part of the CDF detector. As the name would imply, the CLC is designed to provide a measure of the luminosity of $p\bar{p}$ interactions; as we saw in section 2.2.1, measuring the luminosity allows us to measure the cross-section of specific interaction processes.

In general, not all $p\bar{p}$ interactions result in a signature that is picked up by the detector systems described thus far. Inelastic collisions, in particular, often result in forward states which pass through uninstrumented regions of the detector. To count such events, the CLC occupies the region from $3.7 \leq |\eta| \leq 4.7$, and

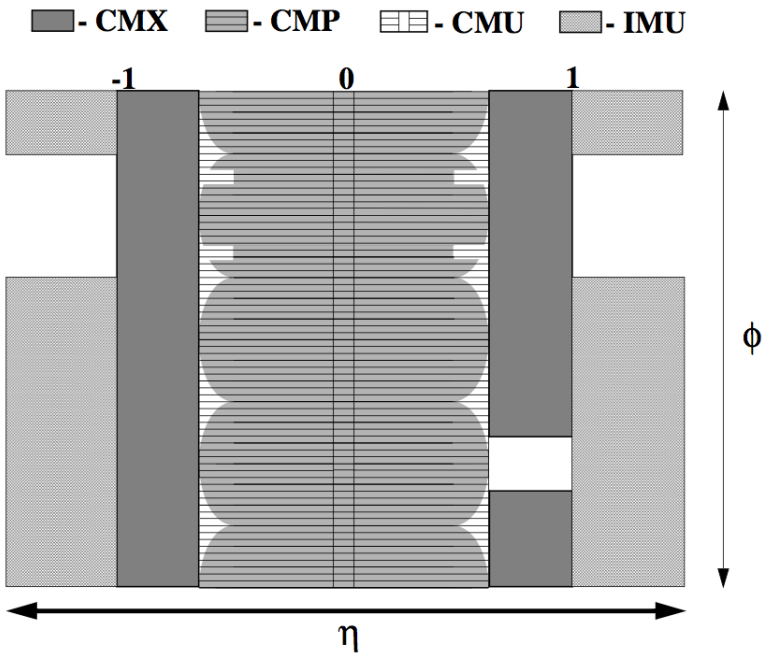


Figure 2.8: Summary of coverage by each of the muon sub-detectors in $\eta - \phi$ space.

it consists of long, conical tubes of isobutane gas which are angled to point towards the collision point [32]. Figure 2.9 displays the basic design of the CLC, as well as its placement within the detector. When forward particles interact with this gas, Cerenkov radiation is produced, and this radiation is amplified by photomultiplier tubes to produce a signal. Signal amplitude thresholds are employed to distinguish particles originating from the $p\bar{p}$ bunch crossings, from softer particles that originate in secondary interactions (*e.g.*, between prompt particles and the beampipe or detector volume) [32]. In all, this system allows for a better measurement of the total cross-section of $p\bar{p}$ interactions, in turn allowing for better resolution of cross-section measurements.

2.3 Trigger System and Data Acquisition (DAQ)

The rate of $p\bar{p}$ bunch crossings at the CDF is 1.7 MHz (one bunch crossing every 396 ns). This is meaningfully compared to the rate at which events can actually be written to tape and recorded, which has varied throughout the life of the experiment but is generally of order $\sim$100 Hz. Clearly, far more collisions are produced per second than can be recorded! The Trigger System is designed to sort through these collision events, and to select only those events that are most likely to be of interest to physics analyses. The Level 1 and Level 2 triggers are very fast hardware triggers that reject most of the events; the Level 3 triggers are software triggers that act upon partially-reconstructed events. We will briefly discuss each trigger layer

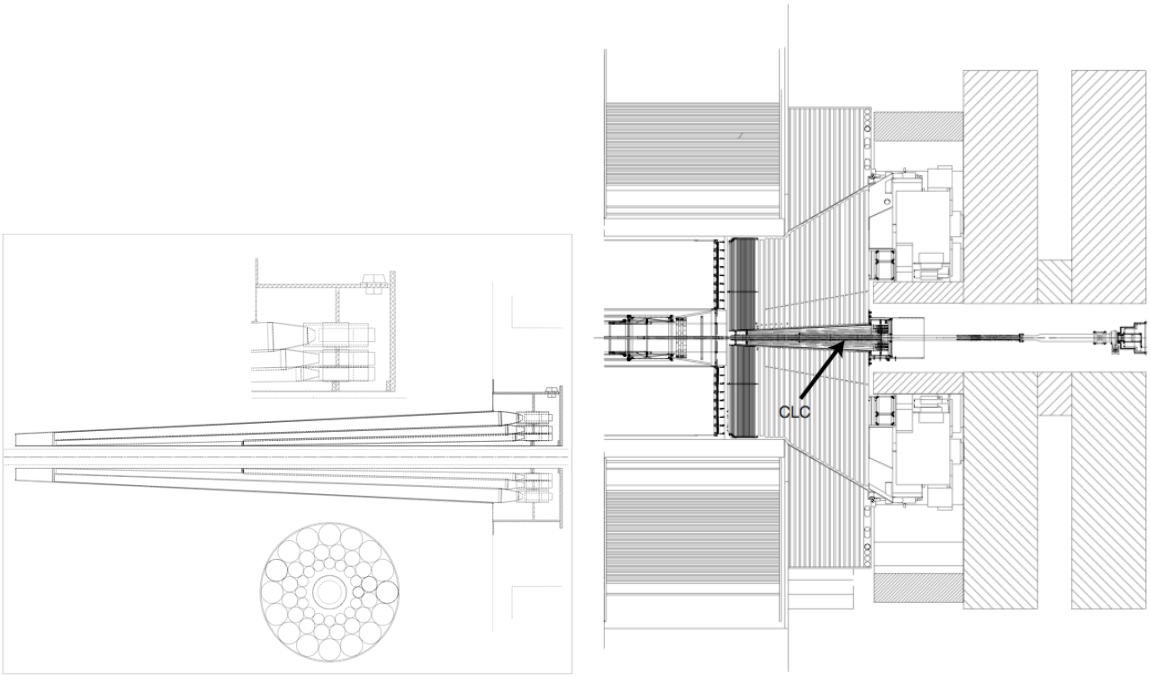


Figure 2.9: Left: the basic design of the CLC, consisting of long, thin, conical tubes angled towards the collision region. Right: the placement of the CLC within the forward detector volume (see Figure 2.3 for larger context).

here. A flow diagram of the trigger system is shown in Figure 2.10.

2.3.1 Level 1 trigger

At this stage, information about hits in the COT, the calorimeters, the muon system, and the luminosity system are available; they are used to reconstruct central tracks, and electron and muon signatures. There are a total of 64 different trigger conditions that can be satisfied at this stage, in order to get an event passed on to Level 2 [33]. Two of the more important subsystems at this level are the eXtremely Fast Tracker (XFT), which has the ability to find tracks with $p_T > 1.5$ GeV; and the Extrapolation Unit (XTRP), which is used to connect these tracks to stubs in the muon detectors (to construct primitive muon objects), and to the calorimeters (to construct primitive electron objects). Events with such muon and electron objects comprise the bulk of the data sample that we will discuss in this thesis. Other triggers at this stage act on primitive jet objects, photons, and the event missing energy $\not\!\!E_T$. Events pass the L1 trigger at a rate of $\sim 30,000$ Hz.

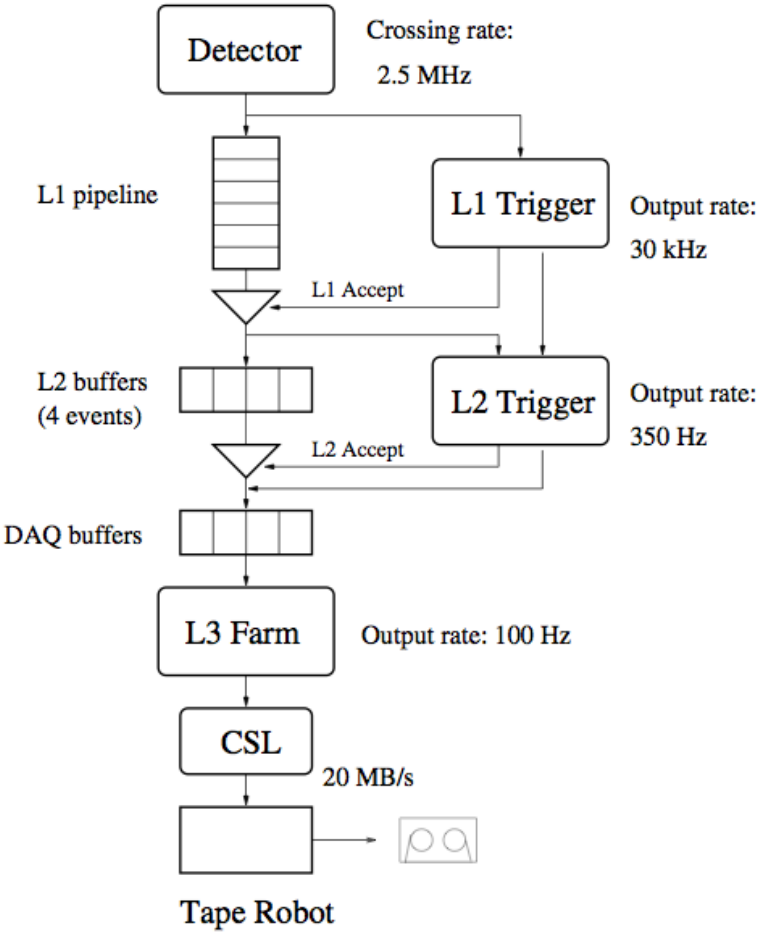


Figure 2.10: The basic flow structure of the CDF trigger system. The actual rate of events accepted and rejected at each step have varied as a function of time, but the basic structure has remained the same throughout CDF Run II. The Level 1 (L1) and Level 2 (L2) triggers are hardware based, and very quick; the Level 3 (L3) triggers are software-based and slower, but act on partially-reconstructed events to make high-level object decisions. Recorded events are stored on magnetic tape in large storage silos—an interesting, if antiquated feature.

2.3.2 Level 2 trigger

Any event that satisfies a Level 1 trigger is passed on to Level 2. Due to the reduced number of input events (reduced relative to the Level 1 input rate), the Level 2 trigger system has time to refine the primitive objects constructed by the Level 1 system, and to make finer cuts.

At this stage, information from the silicon detector is sent to the Silicon Vertex Trigger (SVT), which can identify the transverse impact parameter of tracks with respect to the primary vertex [34]. The impact parameter is, loosely speaking, a measure of how distinct a track's origin is from the primary vertex (we will discuss this further in section 3.3). Heavy flavor hadrons, such as charm and bottom mesons, often travel some finite distance from the primary vertex before decaying; this gives rise to tracks with a non-zero impact parameter. At Level 2, some triggers act to select events with large track impact parameters, in order to select events that are most likely to feature heavy flavor hadrons.

At this level, CES and PES information also becomes available, and can be used to further refine, *e.g.*, electron and photon objects. Any refined events which pass the Level 2 selection cuts are passed on to Level 3, at a rate of about 750 Hz.

2.3.3 Level 3 trigger

At Level 3, complete information about the event is read out to a PC farm running the Linux operating system. On one of these Linux nodes, each event is fully reconstructed with software algorithms. This provides improved momentum and energy resolution, full three-dimensional track reconstruction with matching to the muon stubs and to calorimeter signatures, and silicon-based secondary vertex reconstruction. This improved resolution allows high quality cuts, which select only the most interesting of the events that have made it through the Level 1 and Level 2 triggers. The rate at which events pass the Level 3 trigger cuts is about 100 Hz. These events are written to tape for offline analyses—analyses such as, for example, our search for $W/Z + D^*$ production.

Chapter 3

Object Identification

Our primary goal in this analysis is to identify a W/Z boson and D^* meson emanating from the same $p\bar{p}$ collision. However, we will find it useful to define other objects along the way. We collect these object definitions here, explaining the cuts and motivations behind each.

3.1 The data format

We begin by discussing the format in which our data is stored. When a CDF event is saved to tape, it is often reloaded, reconstructed, and stored in a data format which is more conducive to certain types of anlaysis. This might including constructing new objects that were not present in the original data format (such as secondary vertices), or even leaving out data that is not of interest (such as track uncertainties). This expedites data handling and processing for a given analysis type. For instance, the **BStntuple** format that is used by the CDF bottom group, is particularly well-suited for identifying events containing bottom hadrons and jets. The **TopStntuple** format is used by the CDF top group for top analyses.

For our analysis, we use the **Stntuple** data format. Stntuple stores the tracking information that we need in order to fully reconstruct $D^* \to D^0(K\pi)\pi_s$ decays in Chapter 5—a quality that is not shared by the more frequently-used **TopStntuple** and **BStntuple** formats. Objects stored in Stntuple form are usually prefaced with "TStn", *e.g.*, TStnElectron (electron objects stored in Stntuple format), or TStnJet (jet objects stored in Stntuple format). It is important to note that a TStnElectron will not always describe a true electron; it is often the case that some background event, such as a jet, can fake the signature needed for categorization as a TStnElectron object. I will often refer to such unproven objects as object 'candidates'. All objects of a certain type are stored in parent objects called "Blocks", *e.g.*, the TStnElectronBlock, which stores all TStnElectron objects in a given event.

3.2 Jets

3.2.1 Jet definitions

A jet in hadron collisions can be described at three levels: the parton level, the jet level, and the calorimeter level. All levels seek to describe a jet as one of the products of a hard-scatter event. Ideally, a single jet can be mapped back to a single parent parton (*e.g.*, a bottom quark), providing insight into the nature of partons despite our inability to observe them directly.

At the parton level, a jet is a quark or gluon produced by the hard-scatter process in a $p\bar{p}$ collision. As this quark or gluon moves through the detector (on the timescale of the strong interaction, 10^{-23} s), additional quarks and gluons may be produced by QCD radiation. This collection of partons is the parton-level jet, which shares the net momentum of the original parton.

The partons then hadronize, producing a spray of charged and neutral particles which move through the COT. For a parent parton with large initial momentum, these boosted particles trace a 'cone' of outgoing paths, with cone width dependent upon the momentum of the initial parton. This cone of outgoing particles, all of which originate from a single source, is a particle-level jet.

When these particles strike the calorimeter towers, they produce a pattern of energy signatures. These tower energies are combined by a jet clustering algorithm to produce the final, calorimeter-level jet definitions. The calorimeter-level jet definition is a staple in collider physics, providing a universal metric for characterizing diverse hard-scatter products—because neutral particles do not leave tracks, only the calorimeter has the ability to combine charged and uncharged particle signatures into a single measurement. Tracks within a cone of radius R about the jet center can then be associated wth the jet, resulting in an object which encodes a huge amount of information about the original parent parton. Therefore, when we refer to jets, we are usually talking about these highly-encompassing calorimeter-level jets.

3.2.2 The Midpoint and JetClu jet clustering algorithms

Calorimeter-level jets are constructed using jet clustering algorithms. The ideal algorithm gives a close relationship between jets at the three levels defined in the precious section. This thesis employs the so-called JetClu algorithm, but we also discuss here the popular Midpoint clustering algorithm.

Either jet clustering algorithm mentioned above begins by first creating a list of all calorimeter towers with $E_T > 1.0$ GeV, where

$$E_T = E_{\mathrm{EM}} \cdot \sin\theta_{\mathrm{EM}} + E_{\mathrm{HAD}} \cdot \sin\theta_{\mathrm{HAD}} \tag{3.1}$$

and θ is measured from the beamline with respect to the event's highest-p_T vertex. These towers are called seed towers. Picking the seed tower of highest E_T, the algorithm clumps together all seed towers within a cone of radius $R = 0.4$ about the first tower to create a precluster (or proto-jet). Removing preclustered towers from the list, the process is repeated until every seed tower has been assigned to exactly one precluster.

Each precluster's E_T-weighted center is found, and for the Midpoint algorithm, this center is used to create a list of midpoints (hence the name) between precluster centers. Midpoints that fall within the cones of two or more preclusters are then classified as the center of new proto-jets.

Preclusters are promoted to clusters by drawing a cone of radius R around each proto-jet's E_T-weighted center, and then grouping together all towers within the cone with $E_T > 100$ MeV. The cluster E_T-weighted center is then recalculated, a new cone is drawn, and the process is repeated until the list of included towers remains unchanged (or until the number of iterations reaches a maximum limit). In some algorithms (e.g., JetClu, which does *not* create midpoint preclusters), towers are only added to the cluster with each iteration—they are never removed. This is called ratcheting, and it can lead to very large clusters which contain towers that are no longer within the $R = 0.4$ cone. The Midpoint algorithm does not use ratcheting.

At this point, cluster overlaps are considered. If one cluster is completely contained within another, the smaller cluster (smaller E_T) is dropped. If two clusters partially overlap, the overlap fraction f (sum of E_T of common towers, divided by the E_T of the smaller cluster) is determined. If $f > 0.75$, the clusters are merged. If not, the shared towers are split between the two clusters, with each tower going to the cluster with the nearest E_T-weighted center. After each splitting/merging, cluster centers are redefined. At the end of this process, the remaining clusters are defined as our jets.

An interesting feature of the Midpoint algorithm is that, used as described above, it is possible for cluster cones to drift such that some seed towers do not fall into any jet. This can result in energy signatures that are effectively ignored. This problem is fixed by using a cone size of $R/2$ to find and define stable cones, and then using cone size R to determine jet properties at the last step (splitting/merging).

As said, we use the JetClu algorithm in this analysis—this is the default jet clustering algorithm used by the Stntuple data format. While the Midpoint algorithm may be considered more modern, we do not focus intensely on jets in this analysis—hence, we claim that JetClu is suffice.

3.2.3 Jet energy corrections

Jet E_T at the calorimeter level does not always well describe jet E_T at the particle level. There are many reasons for this, and there just as many solutions for correcting jet energies and strengthening the relationship between the levels of jet definition.

Absolute scale: Uninstrumented regions of the detector (*e.g.*, cracks between the calorimeters) lead to energy loss, and calorimeter reponse must be corrected for non-linearities in the relationship between signal strength and energy signature. This is done with a combination of MC simulations of the calorimeter response, and by measuring calorimeter response to single particles (*e.g.*, kaons and pions) and extrapolating to larger jets.

Relative scale: The calorimeter tower response also depends upon η, with plug calorimeter towers exhibiting different behavior than central calorimeter towers. Central towers are better understood, and can be better calibrated; in a $2 \to 2$ scattering process, the transverse energy of the two resulting jets should be equal. By measuring jet energy in such events in the well-instrumented region $0.2 < |\eta| < 0.6$ (which is also far away from cracks), a correction function is determined as a function of $|\eta|$ and p_T, and extrapolated to other regions.

Multiple interactions and the underlying event: Multiple $p\bar{p}$ collisions during a single bunch crossing can result in artificially high jet energies, with particles from one interaction falling into the jet comes of another. Particles produced in the underlying event (parton interactions in a $p\bar{p}$ collision that are not part of the hard scatter) can have the same effect. The multiple interaction correction is determined based upon the number of reconstructed primary vertices in the event, N_{vtx}—a fair measure of the number of $p\bar{p}$ collisions.

Out of Cone: Jet clustering algorithms will not always correctly measure jet energy—in particular, energy may 'leak' outside of the cone used to cluster jets ($R = 0.4$ for our analysis), resulting in a smaller measured jet energy than was present at the particle/parton level. This is corrected for by measuring energy differences between cones of varying size, and by comparing Monte Carlo predictions with data. These corrections are used to unfold calorimeter-level jet energy back to the parent parton energy.

In our analysis, we apply absolute scale, relative scale, and multiple interaction corrections, using the CDF Jet Corrections Group's jetCorr12 package, version 5. (In this package, underlying event corrections are actually part of the 'absolute scale' correction, which is why we do not state them explicitly.) This package returns a correction scale factor f_{scale} for each jet in a sample, for an input number of quality primary vertices N_{vtx} (vertex quality is determined by a method internal to the package). The corrected four-momentum of each jet is then found as

$$p^{(4)}_{\text{jet, corr}} = p^{(4)}_{\text{jet, raw}} \cdot f_{\text{scale}} \tag{3.2}$$

We only correct energy for jets with a minimum E_T of 8 GeV. We also do not correct for jets that are identified with an electron passing tight cuts defined by the TStnMuonID class (a set of cuts similar to,

but more relaxed than our tight electron cuts (Table 3.1)); electron energy is very well measured by the calorimeters, and does not need to be corrected. A jet is identified with an electron if it is the closest jet in ΔR to that electron (with max $\Delta R = 0.2$).

3.3 Electrons

We define three types of electrons: loose central electrons, tight central electrons, and plug electrons. Loose central electrons and plug electrons will be used to identify $Z \to ee$ events (section 4.2), while tight central electrons will be used to identify $W \to e\nu$ events (section 4.3). It is worth describing some of the quantities by which we define our electrons.

3.3.1 Electron classification

Pseudo-rapidity ($|\eta|$): Pseudo-rapidity has already been defined (section 2.2.1). However, we also use $|\eta|$ to define central ($|\eta_{\text{central}}| < 1.1$) versus plug ($1.2 < |\eta_{\text{plug}}| < 2.8$) electrons. The reason for this is straightforward: the incoming $p\bar{p}$ pair has essentially zero transverse momentum. Only in interactions with high energy/momentum transfer Q^2 will outgoing particles have large p_T and low $|\eta|$. By selecting only particles and tracks in the central ($|\eta| < 1.1$) region, an analysis can focus on high-energy interactions ,while neglecting background at high $|\eta|$. However, valuable information can be lost if the entire region $|\eta| > 1.1$ is neglected. For example, in $Z \to ee$ boson decays, it is often the case that decay products are split across the central and plug regions. By enforcing additional cuts on the electron pair, it is possible to make high-quality selection of electrons from Z events in the plug region—even if these plug electrons do not have as distinct of a signature as the electrons in the central region. We use both types of electrons in the course of this thesis.

Fiducial cuts (z_0, x_{CES}, z_{CES}): The detector is not hermitic, and there exist un- and under-instrumented gaps in its coverage. Lepton candidates with tracks that lead into these gaps may be cut, due to our truncated knowledge of their properties. Fiducial cuts often also include cuts on the z_0 position at which a lepton is traced back to the beamline, and on the x_{CES} and z_{CES} positions at which the electrons leave a signature in the electromagnetic shower (x_{CES} is measured with respect to the CES center for each tower cluster).

Central electron profile (χ^2_{CES}, L_{shr}, Δx_{CES}, and Δz_{CES}): As described in section 2.2.3, the electromagnetic calorimeter contains a region called the electromagnetic shower. Real electrons have an expected signature profile in this shower, described quantitatively by such variables as χ^2_{CES} and L_{shr}. Cuts on these values essentially require the electron signature to be 'electron like', compared to pre-determined profiles.

Cuts on $\Delta z_{\rm CES}$ and $Q \cdot \Delta x_{\rm CES}$ (where Q is track charge) require that the distance between the reconstructed electron track and the best matching CES cluster are small.

Plug electron profile ($\chi^2_{\rm PEM}$, PEM 3×3 fit tower, Pes5x9U(V), $\Delta x_{\rm PES}$, and $\Delta z_{\rm PES}$): As in the case of central electrons, plug electrons have an expected signature profile. These are described quantitatively by $\chi^2_{\rm PEM}$ and Pes5x9U(V), each of which compare observed energy distributions in some part of the detector with templates. PEM 3×3 fit tower compares the number of PEM towers used in the $\chi^2_{\rm PEM}$ fit, compared to the number that would be expected for a real electron. Cuts on $\Delta x_{\rm PES}/\Delta z_{\rm PES}$ parallel those on $\Delta x_{\rm CES}/\Delta z_{\rm CES}$ in the case of central electrons.

Beam constrained tracks: Z (and W) bosons are very short-lived particles, with a lifetime on the order of 10^{-25} s. They decay essentially immediately after production, such that the decay products appear to come from the collision point and within the crossing proton and anti-proton beams. By constraining the decay products to the beamline, lepton production from other sources (*e.g.* hadronic decays) are eliminated from our selection.

NCotAxSeg/**NCotStSeg**: As discussed in section 2.2.2, the COT is composed of four axial and four stereo superlayers. The accuracy of the information provided by a reconstructed track depends upon how many of these segments are used in reconstruction. In general, if there are too few layers involved, then the confidence with which we can claim that several hits have come from the same track will decline. A superlayer is considered involved if there are five or more layers within that superlayer that contain hits. A standard cut is that at least three axial and three stereo superlayers, must each have hits in five or more layers.

Track isolation (Iso1): Track isolation for an electron is defined as the ratio of the total transverse energy of all sources in a cone of $\Delta R = 0.4$ about the electron track (excluding the electron being considered), to the transverse energy of the electron,

$$\text{Iso1(electron)} = \Big(\sum_{\text{not electron}} E_T \Big) / E_T(\text{electron}) \tag{3.3}$$

In this sense, track isolation measures how 'jet like' an electron candidate is—a high isolation value means that there are many other particles in the vicinity of the electron candidate. In general, a lepton from a Z (or W) decay will have no tendency to align itself with other objects in the event (*e.g.*, with a particle jet). On the other hand, fake leptons are very frequently associated with jets: it is not uncommon for one pion

in a jet of particles to deposit most of its energy in the CEM. It is also possible for a hadron to decay into an electron and other particles, which would produce a real electron with a high isolation value—but for a vector boson analysis such as ours, this is not the kind of electron that we want to find. As such, cutting on track isolation helps to eliminate fake Z or W leptonic decay events.

Calorimeter energy ratio (Had/Em): This is simply the ratio of the elctron's energy deposited in the hadronic calorimeter (CHA or PHA), to the electron's energy deposited in the electromagnetic calorimeter (CEM or PEM). Real electrons will deposit most of their energy into the CEM/PEM, with very little leaking into the CHA/PHA, and so are expected to have small Had/Em.

Energy to momentum ratio (E/p): For quality electrons, this ratio is close to unity (at high CDF production energies, electron $p \simeq E$).

Transverse energy (E_T): Electrons from W or Z leptonic decays carry away roughly half of the boson mass as momentum. This means that the p_T and E_T of these electrons is generally much higher than it is for electrons produced by other processes. For electrons, E_T is better reconstructed than p_T: electrons moving in a magnetic field will radiate photons, which generally travel in the electron's initial direction of motion; the calorimeters will pick up this photon energy, but the reconstructed electron track may not properly reflect the initial electron p_T. Cuts requiring a high value of E_T are therefore standard in selecting electrons from leptonic W/Z boson decays.

3.3.2 Tight central electrons

Electrons from leptonic W decay have very high momentum, and should produce high quality tracks and signatures in the detector. We therefore place strict cuts on these electrons, in order to eliminate background. These cuts are featured in the 'Tight central electron' column of Table 3.1. In sum, we require a high transverse energy; an isolated central ($|\eta| < 1.1$) electron track that matches to the electromagnetic shower signature in η and ϕ; an energy signature that is mostly constrained to the electromagnetic calorimeter; and a value of E/p near unity.

3.3.3 Loose central electrons and plug electrons

Electrons from Z decay also have very high momentum (in fact, higher average momentum than in electrons from W decay: compare $M_W = 80.4$ GeV to $M_Z = 90.2$ GeV). However, we must identify two electrons for

Tight central electron	Loose central electron	Plug electron
$\|\eta_{\text{track}}\| < 1.1$	$\|\eta_{\text{track}}\| < 1.1$	$1.2 < \|\eta_{\text{PES}}\| < 2.8$
beam constrained $\|z_0\| < 60$ cm $\|x_{\text{CES}}\| < 21.5$ cm 9.0 cm $< \|z_{\text{CES}}\| < 230.0$ cm	beam constrained $\|z_0\| < 60$ cm $\|x_{\text{CES}}\| < 21.5$ cm 9.0 cm $< \|z_{\text{CES}}\| < 230.0$ cm	
$\chi^2_{\text{CES}} < 10.0$ -3.0 cm $< Q \cdot \Delta x_{\text{CES}} < 1.5$ cm $\|\Delta z_{\text{CES}}\| < 3.0$ cm $L_{\text{shr}} < 0.2$		$\chi^2_{\text{PEM}} < 10.0$ $0.65 <$ Pes5x9(U,V) < 1.0 PEM 3×3 fit tower=true
NCotAxSeg > 3 NCotStSeg > 3		
Iso1 (R=0.4) < 0.1	Iso1 (R=0.4) < 0.1	Iso1 (R=0.4) < 0.1
Had/Em $<0.05 \pm 0.0026$ $\cdot$Log(Em/100GeV)	Had/Em $<0.05 \pm 0.0026$ $\cdot$Log(Em/100GeV)	Had/Em $<0.05 \pm 0.0026$ $\cdot$Log(Em/100GeV)
0.5 GeV $< E/p <$ 2.0 GeV		
$E_T > 25$ GeV	$E_T > 25$ GeV	$E_T > 25$ GeV
beam constrained $p_T > 10$ GeV	beam constrained $p_T > 10$ GeV	

Table 3.1: The cuts used to define electrons in our analysis. For our $W \to e\nu$ events, we require that the electron pass the tight central cuts; for our $Z \to ee$ events, we require one electron to pass the loose central cuts, and the other to pass either the loose central cuts or plug cuts. In general, our cuts require electrons to have high E_T, an isolated track which leads into an instrumented region of the detector, and an 'electron-like' signature in the calorimeter (see section 3.3.1).

$Z \to ee$, versus one for $W \to e\nu$. If our electron cuts are too restrictive, then we will lose many Z candidates. In addition, we have an extra handle on electrons from Z, in that the two must form a common vertex with an invariant mass within the Z boson mass window. In sum, this means that we can be a bit more generous with our electron cuts for Z bosons.

For that reason, we define loose central electron cuts (a less restrictive vesion of the tight central electron cuts defined above); and plug electron cuts, for electrons in the region $1.2 < |\eta| < 2.8$. Both require an isolated track, cuts on E_T and $|\eta|$, and Had/Em values compatible with an electron; the plug cuts place additional requirements upon calorimeter profile. These cuts are also featured in Table 3.1.

3.4 Muons

We define two types of muons: loose central muons, and tight central muons. Loose central muons will be used to identify $Z \to \mu\mu$ events (section 4.2), while tight central electrons will be used to identify $W \to \mu\nu$

events (section 4.3). Again, it is worth describing some of the quantities by which we define our muons.

3.4.1 Muon signature properties

Many of the properties that we use to describe electrons, are also used to describe muons. We discuss here only those properties that are new, or that are different from their electron parallels in some important way.

Transverse momentum (p_T) **and energy** (E_T): A muon is a highly-penetrating particle, and leaves very little energy in the calorimeters: $E_{\rm Em} < 2.0$ GeV, $E_{\rm Had} < 6.0$ GeV are common muon selection criteria. On the other hand, the momentum of a high-energy muon can usually be reconstructed in the COT with excellent accuracy. With $p_T \simeq E_T$ for muons with high E_T and p_T, this makes p_T the better choice for characterizing the muon. For muons from W/Z boson decays, p_T is required to be large.

Pseudo-rapidity ($|\eta|$): As in the case of electrons, we use $|\eta|$ to define central ($|\eta_{\rm central}| < 1.1$) muons. Again, this is in order to select muons from high-energy intractions like those which produce W/Z bosons, over muons from lower-energy background events. While there exists instrumentation for identifying muons in the plug region (the BMU detector), we will not use plug muons in this analysis.

Muon detector stubs (CMUP or CMX stubs, $|\Delta x_{\rm CMU}|$, $|\Delta x_{\rm CMP}|$, $|\Delta x_{\rm CMX}|$): As described in section 2.2.4, the muon detector subsystems lie beyond the calorimeters and (for CMU/X) additional steel shielding. A highly-penetrating muon should leave a track in both the COT, and in one or more of the muon detector subsystems. As such, muon identification often requires that an extrapolated COT track be matched to either a track stub in the CMX detector, or to track stubs in both the CMU and CMP detectors. Matching takes the form of small differences in position between the extrapolated track, and a given stub in the muon detectors (cuts on variables $|\Delta x_{\rm CMU}|$, $|\Delta x_{\rm CMP}|$, and $|\Delta x_{\rm CMX}|$).

Iso1 (track isolation): Track isolation defined for muons in the same way that it is for electrons, but the denominator in Eq. 3.3 is changed to p_T of the muon track; this simply reflects that for muons, p_T is better measured than is E_T. Explicitly:

$$\text{Iso1(muon)} = \Big(\sum_{\text{not muon}} E_T \Big) / p_T(\text{muon}) \tag{3.4}$$

COT track fit quality ($\chi^2_{\rm COT}$): This is a measure of the quality of the muon track in the COT, as determined by track reconstruction algorithms. Decay-in-flight events (in which a charged particle decays into a muon), can produce 'kinked' tracks (one leg is the charged particle prior to the decay; the other is the muon after the decay) that connect to stubs in the muon sub detectors. This track quality cut eliminates such kinked tracks, and hence such decay-in-flight events.

Track impact parameter (d_0): The track impact parameter is found by extrapolating the track back towards the beamline, and then measuring the distance from the beamline to the track's point of closest approach. For a real W/Z decay, muons should originate from the beamline. Due to the finite resolution of the detector, d_0 will not always evaluate to a zero value. Requiring a small value of d_0, however, eliminates muons that are produced by the decay of secondary particles–in general, such muons will originate from a point that is transversely separated from the beamline, and tracing these muon tracks back to the beamline will produce non-zero d_0.

3.4.2 Tight central muons

Muons from $W \to \mu\nu$ decay are highly-penetrating particles with a large average transverse momentum: this results a unique signature involving the CDF muon detector subsystems. With only one muon in our decay, we place strict cuts in order to eliminate background. These cuts are featured in the 'Tight central muon' column of Table 3.2. In sum, we require a high transverse momentum; an isolated central ($|\eta| < 1.1$) muon track that can be matched to one or more stubs in the muon detector subsystems; small energy deposits in the calorimeter; and a COT track with highfit quality $\chi^2_{\rm COT}$ and small impact parameter d_0.

3.4.3 Loose central muons

As in the case of $Z \to ee$ decays, the two outgoing leptons in a $Z \to \mu\mu$ event allow us to be more relaxed with our cuts. We define loose central muon cuts, which differ from the tight cuts by relaxing track quality cuts, and by eliminating the $|\Delta x_{\rm CMU/P/X}|$ cuts. We still require an isolated track, stubs in the muon detector subsystem, cuts on p_T, small energy deposits in the calorimeter, and small impact parameter d_0. These cuts are featured in the 'Loose central muon' column of Table 3.2.

3.5 Cosmic Rays

Cosmic rays are muons produced by collisions between intersteller particles and the upper atmosphere. As stated before, muons are highly-penetrating particles, and these cosmic rays can make their way down to

Muon signature property	Tight central muon	Loose central muon
$\lvert\eta\rvert$	$\lvert\eta_{\text{track}}\rvert < 1.1$	$\lvert\eta_{\text{track}}\rvert < 1.1$
Muon subsystems	Has CMU and CMP stubs AND $\lvert\Delta x_{\text{CMU(CMP)}}\rvert < 3.0$ cm(5.0 cm) ——— OR ——— Has CMX stub AND $\lvert\Delta x_{\text{CMX}}\rvert < 6.0$ cm	Has CMU and CMP stubs ——— OR ——— Has CMX stub
NCotAx(St)Seg	NCotAxSeg > 3 NCotStSeg > 3	
Iso1 (R=0.4)	< 0.1	< 0.1
Had/Em ($p_T \leq 100$ GeV)	$E_{\text{Em}} < 2.0$ GeV $E_{\text{Had}} < 6.0$ GeV	$E_{\text{Em}} < 2.0$ GeV $E_{\text{Had}} < 6.0$ GeV
Had/Em ($p_T > 100$ GeV)	$E_{\text{Em}} = 0.0115 \cdot (p_T - 100)$ GeV $E_{\text{Had}} < 6.0 + 0.028 \cdot (p_T - 100)$ GeV	$E_{\text{Em}} = 0.0115 \cdot (p_T - 100)$ GeV $E_{\text{Had}} < 6.0 + 0.028 \cdot (p_T - 100)$ GeV
Track properties *(Both types also require track to fall into an instrumented region.)*	beam constrained $\lvert z_0\rvert < 60$ cm $\chi^2_{\text{COT}} < 4.0$ $d_0 < 0.3$ cm	beam constrained $\lvert z_0\rvert < 60$ cm $d_0 < 0.3$ cm
beam constrained p_T	> 20 GeV	> 20 GeV

Table 3.2: The cuts used to define muons in our analysis. For our $W \to \mu\nu$ events, we require that the muon pass the tight central cuts; for our $Z \to \mu\mu$ events, we require both muons to pass the loose central cuts. In general, our cuts require muons to have high p_T, an isolated track which leads into an instrumented region of the detector, small energy deposits in the calorimeter, and stubs in one or more of the muon detector subsystems. (see section 3.4.1).

the earth's surface. If a muon passes through the detector volume during a bunch crossing, and if this muon happens to pass close to the beamline, it can look very much as if two muons were produced back-to-back in a $p\bar{p}$ collision event (e.g., a $Z \to \mu\mu$ event). Alternately, if only one part of the cosmic ray is reconstructed, it may look as if a single high-p_T muon was produced (e.g., a $W \to \mu\nu$ event). To eliminate contamination from cosmic rays, we employ theTStnCosmicBlock $\to$ HasCosmicRay() accessor. The TStnCosmicBlock class is an Stntuple format version of the CosmicFinderModule class, which uses timing information, track quality, and impact parameters to identify cosmic ray events in CDF data[35]. For this thesis, we veto all events in our $W \to \mu\nu$ and $Z \to \mu\mu$ selection that are tagged as having a cosmic ray.

3.6 Missing transverse energy

Weakly-interacting particles that are produced in a collision (neutrinos, dark matter candidates) can pass through the detector without leaving any signal. However, we can still infer their presence by calculating an event's missing transverse energy ($\not\!\!E_T$). As earlier stated, the incoming partons have net $\overrightarrow{p_T} \sim 0$; therefore, the sum $\sum \overrightarrow{p_T}$ of all jets in an event ought to add up to zero. If there is a large discrepancy from zero, then this 'missing' energy may have been carried away by a weakly-interacting particle. The 'raw' value of $\not\!\!E_T$ is simply the negative of the sum $\sum \overrightarrow{p_T}$ of all jets in an event.

In our analysis, we take this value from the TStnMetBlock's index 1 entry—for high-p_T lepton events, this is $\not\!\!E_T$ calculated with respect to the high-p_T lepton's point of closest approach to the beamline (recall from section 3.2 that jet E_T is always defined with respect to some vertex). However, this value must be corrected for two major factors: muons, and jet corrections.

When a muon passes through the calorimeter, it deposits very little energy—this produces a 'fake' missing energy signature in the detector (fake insofar as that energy is not carried away by a weakly-interacting particle). We correct for this using the TStntuple::CorrectMetForMuons algorithm. This algorithm loops through all TStnMuon objects in the event's TStnMuonBlock, and selects those which pass a set of strict muon cuts (a slightly more relaxed version of our 'tight central muon' cuts (Table 3.2)). The difference between the muon's calorimeter energy and its track momentum is our false missing energy (recall $E \simeq p$ for high$-p$ muons). We define

$$\not\!\!E_{T,fake}(\text{muon}) = \overrightarrow{p_T} \cdot (1 - E/p) \tag{3.5}$$

for each tight muon, and subtract this term from our missing energy.

We must also account for jet corrections. $\not\!\!E_T$ as taken from the TStnMetBlock is calculated *before* any

jet energy corrections are applied. To modify $\not\!\!E_T$ to reflect these jet corrections, we define a difference vector

$$\Delta p^{(4)}_{\text{jet}} = p^{(4)}_{\text{jet, corr}} - p^{(4)}_{\text{jet, raw}} \tag{3.6}$$

for each jet, where $p^{(4)}_{\text{jet, raw}}$, $p^{(4)}_{\text{jet, corr}}$ are the jet four-momenta before and after corrections, respectively. We then add all of these $\Delta p^{(4)}_{\text{jet}}$ together to find $\Delta p^{(4)}_{total} = \sum_{\text{all jets}} \Delta p^{(4)}$. The negative of the transverse part of $\Delta p^{(4)}_{\text{total}}$ is our correction to $\not\!\!E_T$ for jets, and we add this to our muon-corrected $\not\!\!E_T$ value. At this point, we consider our $\not\!\!E_T$ to be fully corrected. In sum,

$$\not\!\!E_T \;=\; \not\!\!E_{T,\text{ raw}} \;-\; \sum_{\text{tight muons}} \overrightarrow{p_T} \cdot (1 - E/p) \;-\; \sum_{\text{all jets}} \left(p^{(4)}_{\text{jet, corr}} - p^{(4)}_{\text{jet, raw}}\right) \tag{3.7}$$

Chapter 4

Vector Boson Event Selection

Our analysis begins with the identification of Z and W vector bosons in the CDF high-p_T lepton datasets. We therefore begin this Chapter by describing the high-p_T lepton datasets, and the triggers which define them. We then discuss the cuts that we use to identify Z and W bosons, apply these cuts, and count the number of selected events. We conclude by determining what fraction of these events are fake Z or W events (other kinds of event, which simply mimic the Z or W signature).

4.1 The datasets

We look for our vector bosons in three high-p_T lepton Stntuple CDF datasets: HIGH_PT_MUON (abbreviated **bhmu**), HIGH_PT_ELECTRON (**bhel**), and PLUG_ELECTRON **(bpel**). Each of these datasets covers the full range of CDF Run II data taking (Periods 0 through 38), corresponding to 9.7 fb^{-1} of integrated luminosity.

A **bhmu** event is defined (at L3) by having a high-p_T central muon with $p_T > 18$, and either CMU and CMP stubs, or a CMX stub. The transverse distance between the extrapolated COT track and the associated CMU/P/X stubs must satisfy $|\Delta x| < 10$ cm. This dataset is a common candidate for vector boson searches; when a vector boson decays into two leptons, each lepton carries away approximately half of the Z or W boson's mass as momentum—this often leads to muons with very high transverse momentum, which is a clear signature to work with. Specifically, we use this dataset for our $W \to \mu\nu$ and $Z \to \mu\mu$ searches.

A **bhel** event is defined (at L3) by having a high-p_T central electron with $E_T > 18$ GeV and $p_T > 6$ GeV. The extrapolated electron track must match to a hit in the CES within $|\Delta x| < 3$ cm and $|\Delta x| < 5$ cm. Again, and for the same reasons, this dataset is a common candidate for vector boson searches. We use this dataset for our $W \to e\nu$ and $Z \to ee$ searches.

A **bpel** event is defined (at L3) by having a PEM cluster with $E > 20$ GeV, a matched electron track with Iso1 < 0.1, and missing energy $\not\!\!E_T > 15$ GeV. Note that the calorimeter energy restriction here is on E, and not on E_T; in the forward region, we will expect to see particles with lower transverse energy

and momentum, compared to the central region. So, we trigger on total E registered by the calorimeter, instead. A **bpel** event also requires that the ratio of energy deposited in the electromagnetic and hadronic calorimeters satisfies $E(\text{had})/E(\text{em}) < 0.075$. Due to the $\not\!\!E_T$ cut, this trigger is preferentially geared towards $W \to e\nu$ identification. However, in our analysis, we will use it to search for $Z \to ee$ events, only.

4.2 Identifying $Z \to \ell^+\ell^-$ events

We identify Z events only for those cases in which $Z \to ee$ or $Z \to \mu\mu$. These dilepton events produce very clean and recognizable signatures, making them an ideal pair of channels to work with. After finding two good electrons or muons, we calculate the dilepton invariant mass, and compare this to the known mass window of Z bosons. The dilepton invariant mass is then binned for all events near the Z boson mass window, and the resulting histogram is fit to a signal plus background hypothesis in order to count signal events.

4.2.1 Identifying $Z \to ee$ events

We begin our search for $Z \to ee$ decays by looking for two quality electrons in **bhel** and **bpel** events. For each event, we require that the (central) high p_T trigger electron pass our loose central electron cuts, and that a second (oppositely-signed) electron pass either our loose central electron cuts, or our plug electron cuts (Table 3.1). We then require that the dielectron invariant mass $M(e^+e^-)$ fall within the region 66 GeV $< M(e^+e^-) <$ 116 GeV. Finally, we check that the event does not pass our $W \to e\nu$ cuts (section 4.3); it is possible for an electron from $W \to e\nu$ decay to be paired with another (real or fake) electron, to produce a signature that is $Z \to ee$ like. Any event which passes all of these cuts, we tag as $Z \to ee$.

4.2.2 Identifying $Z \to \mu\mu$ events

Our procedure for identifying $Z \to \mu\mu$ events parallels that for identfying $Z \to ee$ events. We begin by searching for two quality muons in **bhmu** events. We require that the high-p_T trigger muon pass our loose central muon cuts, and that a second muon also pass those cuts (Table 3.2). The dimuon invariant mass must fall near the Z mass window (satisfying 66 GeV $< M(\mu^+\mu^-) <$ 116 GeV), and the event must not pass our $W \to \mu\nu$ cuts. The $Z \to \mu\mu$ selection differs from the $Z \to ee$ selection, in that we must also account for cosmic ray events; if a cosmic muon passes through the detector during a bunch crossing, it can produce a fake dimuon signature with an invariant mass that falls within our mass window. Therefore, we veto any events which are tagged as having a cosmic ray (section 3.5). Any event which passes these cuts, we tag as a $Z \to \mu\mu$ event.

$Z \to ee$ event selection cuts	$Z \to \mu\mu$ event selection cuts
High-p_T trigger electron passes loose central electron cuts	High-p_T trigger muon passes loose central muon cuts
A second electron passes loose central electron cuts ——— OR ——— A second electron passes plug electron cuts	A second muon passes loose muon electron cuts
Electrons are oppositely-signed	Muons are oppositely-signed
Dilepton invariant mass satisfies $66\ \text{GeV} < M(e^+e^-) < 116\ \text{GeV}$	Dilepton invariant mass satisfies $66\ \text{GeV} < M(\mu^+\mu^-) < 116\ \text{GeV}$
Event does not pass $W \to e\nu$ cuts	Event does not pass $W \to \mu\nu$ cuts
	Event is not tagged as a cosmic ray event

Table 4.1: The cuts that we use to identify $Z \to ee$ and $Z \to \mu\mu$ events in the high-p_T lepton dataset. The electron/muon cuts referenced above can be found in Tables 3.1, 3.2.

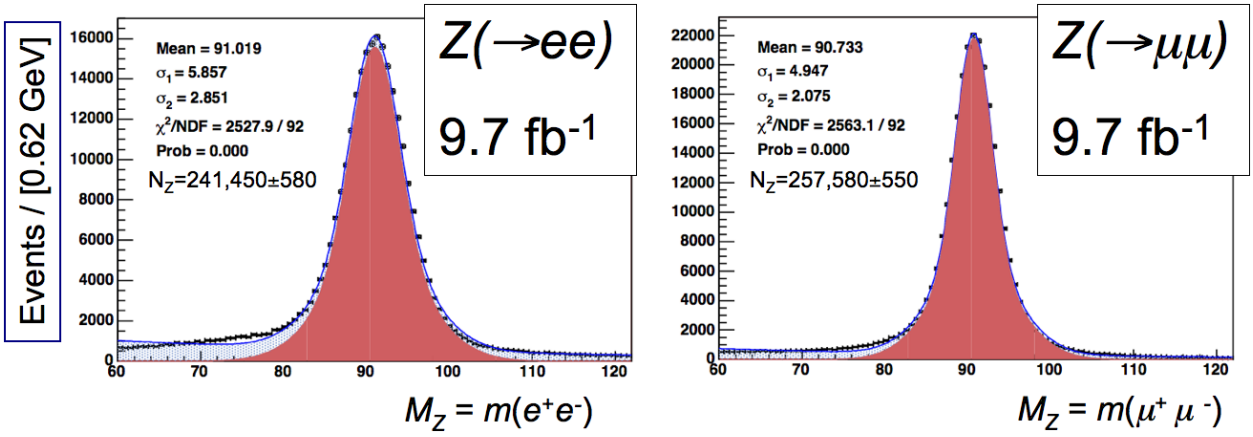


Figure 4.1: Plots of the invariant mass $M(\ell^+\ell^-)$ of all lepton pairs which pass our $Z \to ee/\mu\mu$ cuts. We fit the $M(\ell^+\ell^-)$ distribution with a double-Gaussian signal (widths (σ_1, σ_2) and mean shown on plots) plus an exponential background. The signal fit is shown in red, the combined fit in blue. By integrating under the signal function, we count $N_{Z\to ee} = 241,450 \pm 580$, and $N_{Z\to\mu\mu} = 269,451 \pm 560$. Note that only one lepton pair is accepted for each event, consisting of the high-p_T trigger lepton, and the first secondary lepton in the muon/electron block to satisfy all selection criteria.

4.2.3 Background to the $Z \to ee/\mu\mu$ selections

Our Z selection cuts will inevitably include some background events. Accounting for background in our Z selection is quite straightforward. We bin the invariant mass $M(\ell^+\ell^-)$ for all events which pass our $Z \to \ell^+\ell^-$ cuts. The Z resonance results in a clear peak near $M_Z = 91.2$ GeV (Figure 4.1). We then fit the $M(\ell^+\ell^-)$ histogram to a double-gaussian signal plus exponential background hypothesis, and integrate beneath the signal curve to count the number of true Z resonance events, $N_{Z\to\ell\ell}$. Doing this, we count $N_{Z\to ee} = 241,450 \pm 580$ and $N_{Z\to\mu\mu} = 257,600 \pm 550$.

4.2.4 The effects of binning and range

As can be seen in Figure 4.1, the choice of background shape does not perfectly describe the data. Because we are primarily interested in the size of the signal shape, this is a problem only if an improperly-shaped background ansatz affects our signal size. We can test this by varying the range and bin size of the $M(\ell^+\ell^-)$ histograms, and remeasuring the size of the signal peak. Different choices will result in fits of varying quality (as quantified in the χ^2/DOF value associated with each fit), and if large changes in the fit quality do not correspond to large changes in the signal size, then we can say that our signal count does not strongly depend upon our choice of background shape.

That being said, we check the effects of bin and range choice on the count of $Z \to \mu\mu$ and $Z \to ee$ events identified by our selection criteria. The results, and the conditions enforced to find each result, are summarized in Table 4.2. A few sample plots from these tests are shown in Figure 4.2. We find that changing the bin definitions has very little effect for a histogram range $M(\ell^+\ell^-) \in\sim [60.0, 122.0]$. However, the signal counts can be changed by $\sim 2\%$ when the $M(\ell^+\ell^-)$ range is reduced to $M(\ell^+\ell^-) \in\sim [70.0, 112.0]$. This 2% uncertainty is dwarfed by a much larger uncertainty of $\sim 30-50\%$ in our final measurements of the $Z + D^*$ production rate (as we'll see in chapter 9). Therefore, we consider the effect of rebinning—and by extension, the effect of our choice of background function—to be negligible in our final results. We will ignore this uncertainty henceforth.

4.2.5 Comparison to counts by other analyses

As a final check of the quality of our $Z \to \mu\mu/ee$ cuts, we also compare our $Z \to \mu\mu/ee$ signal counts against counts found by other CDF analyses. In particular, we look at two Fall 2013 analyses ([36, 37]) which measure $\sin^2\theta_W(M_W)$ using $\mu\mu/ee$ pairs in the Z boson region. Both of these analyses count $Z \to \mu\mu/ee$ events in the region $66 < M_Z/\text{GeV} < 116$, using the same high-$p_T$ lepton datasets that we do.

The first analysis [36] counted $51,951$ central-central $Z \to ee$ events in 2.1fb^{-1} of data, using cuts that

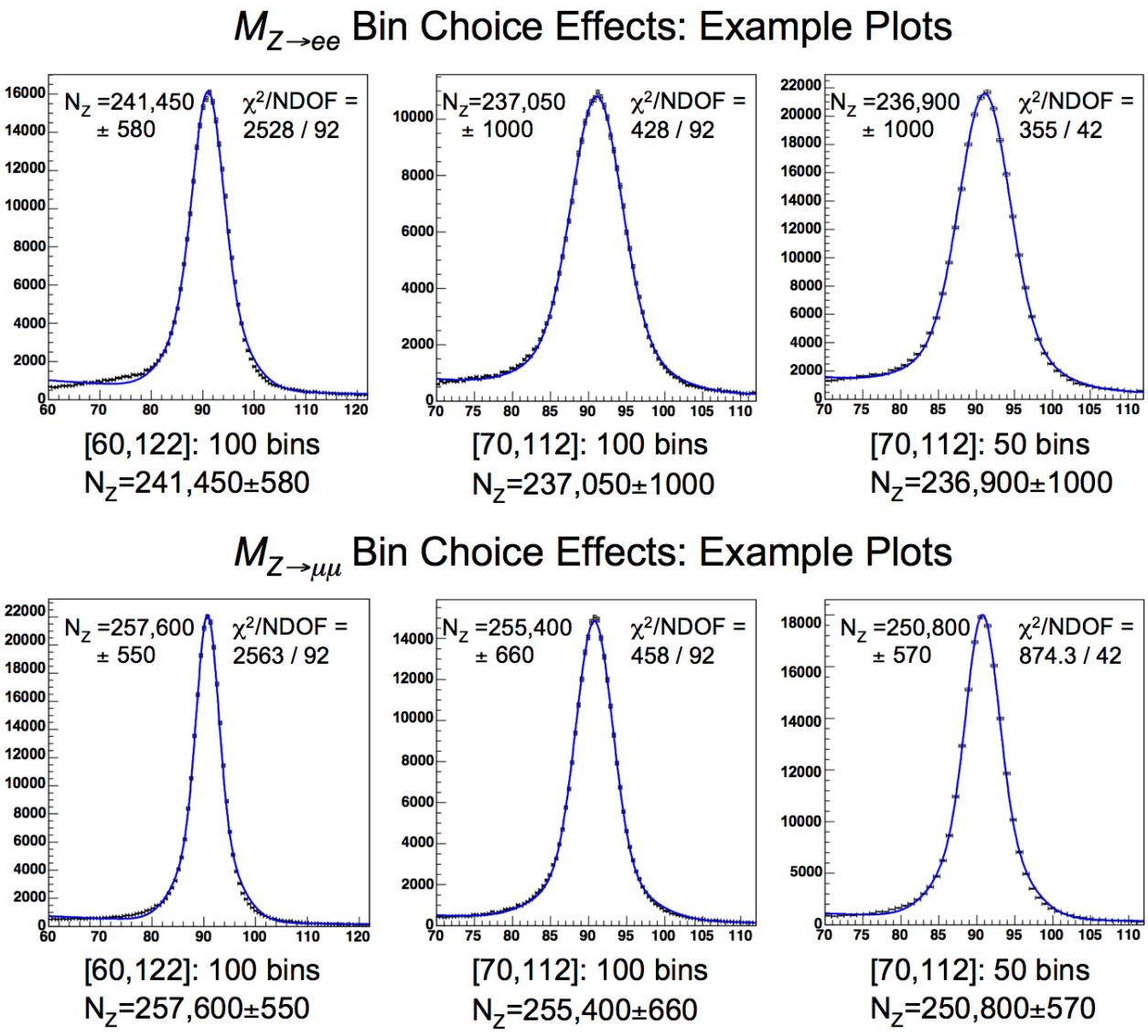


Figure 4.2: Sample plots of invariant mass $M(\ell^+\ell^-)$ for different bin and range choices. Plots for the $Z \to ee$ case are on the top row; for the $Z \to \mu\mu$ case, on the bottom row. For more details, and to see the results of other bin and range choices, see Table 4.2.

$M(\ell^+\ell^-)$ histogram range	# bins	$N_{Z\to\mu\mu}$	χ^2/NDOF	$N_{Z\to ee}$	χ^2/NDOF
[60.0, 122.0]	100	$257,600 \pm 550$	2563 / 92	$241,450 \pm 580$	2528 / 92
[60.0, 122.0]	50	$257,600 \pm 550$	2494 / 42	$241,530 \pm 580$	2452 / 42
[60.0, 122.0]+half bin shift	100	$257,500 \pm 550$	2539 / 92	$241,330 \pm 580$	2474 / 92
[60.0, 122.0]+half bin shift	50	$257,400 \pm 550$	2376 / 41	$240,980 \pm 580$	2335 / 41
[70.0, 112.0]	100	$255,400 \pm 660$	458 / 92	$237,050 \pm 1000$	428 / 92
[70.0, 112.0]	50	$250,800 \pm 570$	874.3 / 42	$236,900 \pm 1000$	355 / 42

Table 4.2: This table presents the number of Z events identified in our $Z \to \mu\mu$ and $Z \to ee$ samples, using different bin and range choices. While there is little variation as a function of bin width or placement in the range $M(\ell^+\ell^-) \in\sim [60.0, 122.0]$, we see that the signal count drops if histogram range is reduced to $M(\ell^+\ell^-) \in\sim [70.0, 112.0]$. This effect is small ($\sim 2\%$), and is most pronounced only for the last test (which has fewer bins in the detailed area about the M_Z peak, and thus poorer resolution). A 2% uncertainty is negligible compared to the $\sim 30-50\%$ uncertainty in our final measurements of the $Z + D^*$ production rate, and so we ignore this 'binning uncertainty' in the rest of this analysis.

were similar to our own. (Here, a 'central-central' event is one in which both electrons fall in the central region of the detector; a central-plug event, an event in which electron number one (two) falls in the central (plug) region.) Extrapolating to 9.7 fb^{-1}, we would estimate $\sim 239,960$ central-central $Z \to ee$ events in the full dataset, for this analysis. As our analysis counted very few central-plug $Z \to ee$ events (only ~ 120, compared to $\sim 240,000$ central-central candidates), we would expect our yield ($241,450 \pm 580$) to be similar to the yield of this analysis, which is indeed what we find. We could likely increase our $Z \to ee$ yield by adopting the looser cuts of [36], but there are two arguments against this: we already have well-established results using the central sample; and more forward $Z \to ee$ events are less likely to have D^* in the central region (our D^* selection method (chapter 5) requires central D^*).

The second analysis [37] counted $276,623$ $Z \to \mu\mu$ events in 9.2 fb^{-1} of data, again using cuts similar to (but looser than) our own. This is comparable to the $257,600 \pm 550$ events that we find in 9.7 fb^{-1} of data (taking into account the differences in selection criteria). We conclude that our counts of $Z \to \mu\mu$ and $Z \to ee$ are compatible with similar counts in other analyses; we are unlikely to be missing major useful cuts, or eliminating an unnecessarily large number of Z events due to excessively strict cuts.

Major selection criteria differences between the $\sin^2\theta_W(M_W)$ analysis and our own are noted in Table 4.3, along with the $Z \to \mu\mu/ee$ counts recorded by each.

4.3 Identifying $W \to \ell\nu$ events

We identify W events only for those cases in which $W \to e\nu$ or $W \to \mu\nu$. These events produce a very recognizable signature, consisting of a high-p_T lepton paired with large missing energy $\not\!\!E_T$ (energy that is carried away by the weakly-interacting neutrino). As we will see, however, these events are a bit more difficult to work with than the $Z \to ee/\mu\mu$ events; we do not have an invariant mass peak to work with, and

Z selection criteria	Our Analysis	$\sin^2\theta_W(M_W)$ Analysis
$Z \to ee$	$E_T > 25/25$ GeV (two central candidates) $E_T > 25/25$ GeV (one central, one plug)	$E_T > 25/15$ GeV (two central candidates) $E_T > 20/20$ GeV (one central, one plug)
$Z \to ee$	Central/Plug electrons must pass Loose/PlugIDWord cuts	Both electrons must pass 'standard' CDF electron selection cuts
$Z \to ee$ Count	$241,450 \pm 580$ mostly central-central events in 9.7 fb^{-1}	$51,951$ central-central events in 2.1 fb^{-1}
$Z \to \mu\mu$	Both muons: track matched to muon stub Both muons: $\lvert\eta\rvert < 1.1$	Muon 1: track matched to $\lvert\eta\rvert < 1$ muon stub Muon 2: any muon candidate
$Z \to \mu\mu$	No constraint on muon pair $\lvert y\rvert$	Muon pair: $\lvert y\rvert < 1$
$Z \to \mu\mu$ Count	$257,600 \pm 550$ in 9.7 fb^{-1}	$276,623$ in 9.2 fb^{-1}

Table 4.3: Differences between our $Z \to \mu\mu/ee$ selection criteria (Table 4.1), and those of the $\sin^2\theta_W(M_W)$ analyses [36, 37]. With these differences in mind, our $Z \to \mu\mu/ee$ counts are declared comparable with the counts of the $\sin^2\theta_W(M_W)$ analysis. The full sets of cuts used for our Z selection can be found in Tables 3.1, 3.2, 4.1.

so we must estimate the contributions of background by a more clever approach.

4.3.1 Identifying $W \to e\nu$

We start by skimming the **bhel** dataset for events with corrected $\not\!\!E_T > 25$ GeV. We also require that the (central) high p_T trigger electron pass our tight central electron cuts (Table 3.1). We then evaluate the transverse mass of this W candidate, defined as

$$M_T(W) = \sqrt{2.0 \cdot E_T(e) \cdot \not\!\!E_T \cdot \left(1 - \cos(\Delta\phi_{e\vec{\not\!\!E}_T})\right)} \tag{4.1}$$

In essence, $M_T(W)$ is the invariant mass of the electron and neutrino, $M(e\nu)$, but the z axis is ignored (along with any energy/momentum that is directed along it). The W transverse mass has a well-defined and well-studied shape, and, in particular, background from $W \to \tau\nu$ decays can be mostly removed by enforcing the cut $M_T(W) > 20$ GeV [38]. Finally, we reject the event if it passes our $Z \to ee$ cuts; it is possible for an electron from Z decays to be paired with real or fake missing energy, to produce a false W signature. We count $5,081,938$ $W \to e\nu$ candidates.

$W \to e\nu$ event selection cuts	$W \to \mu\nu$ event selection cuts
High-p_T trigger electron passes tight central electron cuts	High-p_T trigger muon passes tight central muon cuts
corrected $\not\!\!E_T > 25$ GeV	corrected $\not\!\!E_T > 20$ GeV
$M_T(W) > 20$ GeV	$M_T(W) > 25$ GeV
Event does not pass $Z \to ee$ cuts	Event does not pass $Z \to \mu\mu$ cuts
	Event is not tagged as a cosmic ray event

Table 4.4: The cuts that we use to identify $W \to e\nu$ and $W \to \mu\nu$ events in the high-p_T lepton dataset. The electron/muon cuts referenced above can be found in Tables 3.1 and 3.2.

4.3.2 Identifying $W \to \mu\nu$

Paralleling the $W \to e\nu$ case, we start by skimming **bhmu** for events with corrected $\not\!\!E_T > 20$ GeV. We then require the (central) high-p_T trigger muon pass our tight muon cuts (Table 3.2). We relax the $\not\!\!E_T$ and muon p_T cuts for $W \to \mu\nu$, compared to our cuts for $W \to e\nu$, due to a cleaner muon sample: in short, it is easier for a particle to fake an electron, than it is to fake a muon. The W transverse mass is also defined slightly differently in $W \to \mu\nu$ events: $E_T(e)$ is replaced by $p_T(\mu)$ in order to reflect more accurate measurement of muon p_T than muon E_T.

$$M_T(W) = \sqrt{2.0 \cdot p_T(\mu) \cdot \not\!\!E_T \cdot \left(1 - \cos(\Delta\phi_{\mu\vec{\not\!\!E}_T})\right)} \tag{4.2}$$

As in our $Z \to \mu\mu$ selection, we veto cosmic ray events: a cosmic muon can be paired with real or fake missing energy in order to produce a false $W \to \mu\nu$ signature. Finally, we reject the event if it passes our $Z \to \mu\mu$ cuts. We count $5,348,975$ $W \to \mu\nu$ candidates.

4.3.3 Background to the $W \to \ell\nu$ selections

Our $W \to \mu\nu/e\nu$ selection criteria ensure that the majority of tagged events are W bosons. However, some electroweak backgrounds—specifically, $Z \to \mu\mu/ee$ and $W \to \tau\nu$—may mimic our signal. We consider $W \to \tau\nu$ to be a 'background', because it may behave differently than our $W \to \mu\nu/e\nu$ signals (*e.g.* under

	$\not\!\!E_T$ boundary, $W \to \mu\nu$ ($e\nu$)	Charged lepton Iso1 boundary
Region **A**	$\not\!\!E_T$ $<$ 10 GeV	Iso1 $<$ 0.1
Region **B**	$\not\!\!E_T$ $<$ 10 GeV	Iso1 $>$ 0.3
Region **C**	$\not\!\!E_T$ $>$ 20 (25) GeV	Iso1 $>$ 0.3
Region **W**	$\not\!\!E_T$ $>$ 20 (25) GeV	Iso1 $<$ 0.1

Table 4.5: The boundaries used to define regions **A**, **B**, **C**, and **W** for W candidates in the $\not\!\!E_T$ / Iso1 plane. These are visualized in Figure 4.3. For regions **C** and **W**, the $\not\!\!E_T$ boundary depends upon whether we are consider the case of $W \to \mu\nu$ or $W \to e\nu$—the value for the $W \to e\nu$ case is written in parentheses.

the influence of a neural network, or in terms of overall acceptance rate).

In addition to these electroweak backgrounds, QCD jets may also mimic our W signature. As an example, a high-p_T charged pion may pass through the calorimeters and into the muon tracking chambers, mimicking a muon. At the same time, another jet with mismeasured energy may produce a large $\not\!\!E_T$ signature. The end result is what appears to be a $W \to \mu\nu$ signature. Only a small fraction of all jet events are expected to fake a W signature, but there are a huge number of jets produced in $p\bar{p}$ collisions. In sum, there will likely be some measureable QCD jet contamination in our W sample.

Estimating the background contribution of electroweak events is traditionally done by running Monte Carlo simulations. However, accurate estimates of QCD jet background requires a data-driven technique. We will in fact solve for both of these types of background contribution simultaneously, using a set of coupled equations. This custom technique eliminates many unknowns, and provides an easy determination of combined systematic uncertainty (see 6.4).

We begin by first removing the $\not\!\!E_T$ and lepton track isolation (Iso1) cuts from the $W \to \ell\nu$ selection criteria. We then define four regions as shown in Table 4.5 and Figure 4.3. To first order, regions **A**, **B**, and **C**, are composed entirely of QCD jet events [39]. We therefore approximate:

$$N_{QCD,W} \simeq \frac{N_{cand,A}}{N_{cand,B}} \cdot N_{cand,C} \tag{4.3}$$

where $N_{cand,X}$ is the number of candidate events in region **X**, and $N_{QCD,X}$ is the number of QCD jet events in region **X**. In reality, however, regions **A**, **B**, and **C** will contain at least some small amount of $Z \to \ell^+\ell^-$ and $W \to \tau\nu$ background, as well as $W \to \ell\nu$ signal events. We would like to correct (4.3), to account for this. We define:

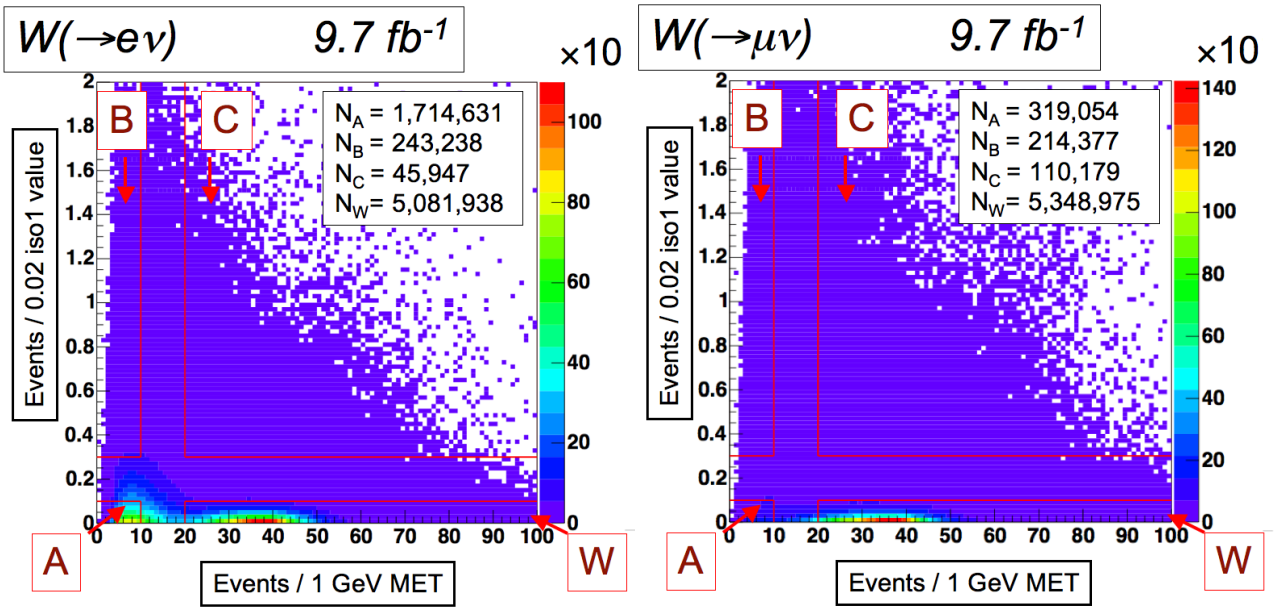


Figure 4.3: A two-dimensional distribution of all W candidates, with $\not\!\!E_T$ and charged lepton isolation cuts removed. We define four regions in this space, **A**, **B**, **C** and **W**, and use them to estimate the number of background events in signal region **W**.

$$
\begin{aligned}
X_{W\to\ell\nu} &\equiv \text{Number of } W\to\ell\nu\text{signal events in quadrant } \mathbf{X} \\
X_{W\to\tau\nu} &\equiv \text{Number of } W\to\tau\nu\text{background events in quadrant } \mathbf{X} \\
X_{Z\to\ell\ell} &\equiv \text{Number of } Z\to\ell^+\ell^-\text{background events in quadrant } \mathbf{X} \\
X_{\text{QCD}} &\equiv \text{Number of QCD jet background events in quadrant } \mathbf{X} \\
X_{\text{cand}} &\equiv \text{Number of total events in quadrant } \mathbf{X}
\end{aligned}
$$

We then expand the definition of 'QCD' background to mean any background that is not otherwise explicitly stated (the overwhelming majority of which is expected to be QCD jet events). This allows us to state:

$$X_{\text{cand}} = X_{W\to\ell\nu} + X_{Z\to\ell\ell} + X_{W\to\tau\nu} + X_{\text{QCD}} \tag{4.4}$$

We may then refine (4.3), writing

$$W_{\text{QCD}} = \frac{A_{\text{QCD}}}{B_{\text{QCD}}} \cdot C_{\text{QCD}} \tag{4.5}$$

Using (4.4), we rewrite (4.5) as

$$
\begin{aligned}
W_{\rm cand} - W_{W\to\ell\nu} - W_{W\to\tau\nu} - W_{Z\to\ell\ell} &= \frac{A_{\rm cand} - A_{W\to\ell\nu} - A_{W\to\tau\nu} - A_{Z\to\ell\ell}}{B_{\rm cand} - B_{W\to\ell\nu} - B_{W\to\tau\nu} - B_{Z\to\ell\ell}} \\
&\times \Big(C_{\rm cand} - C_{W\to\ell\nu} - C_{W\to\tau\nu} - C_{Z\to\ell\ell}\Big)
\end{aligned} \tag{4.6}
$$

Next, we use Monte Carlo simulations to scale values $X_{W\to\ell\nu}$, $X_{W\to\tau\nu}$, $X_{Z\to\ell\ell}$ to the (unknown) number of signal events in region **W**. Specifically, for electroweak process $Y \to ab$, we define a quantity $R^X_{Y\to ab}$ as:

$$
R^X_{Y\to ab} \equiv \frac{f^X_{Y\to ab}}{f^W_{W\to\mu\nu}} \cdot \frac{\sigma(Y\to ab)}{\sigma(W\to\mu\nu)} \tag{4.7}
$$

where $f^X_{Y\to ab}$ is the fraction of simulated signal events of type $Y \to ab$ which pass our $W \to \ell\nu$ selection cuts and fall in Region **X**, and $\sigma(Y \to ab)$ is the production cross section of events of type $Y \to ab$. For any electroweak background process, we may then connect $X_{Y\to ab}$ to the unknown $W_{W\to\ell\nu}$ as:

$$
X_{Y\to ab} = R^X_{Y\to ab} \cdot W_{W\to\ell\nu} \tag{4.8}
$$

This approach makes only one assumption: that the Monte Carlo simulation is good at predicting the properties of $W \to \ell\nu$ and $Z \to \ell^+\ell^-$decays.

Applying (4.8) to (4.5), we find

$$
\begin{aligned}
W_{\rm cand} - \Big(R^W_{W\to\mu\nu} + R^W_{W\to\tau\nu} + R^W_{Z\to\mu\mu}\Big)W_{W\to\mu\nu} &= \frac{A_{\rm cand} - \Big(R^A_{W\to\mu\nu} + R^A_{W\to\tau\nu} + R^A_{Z\to\mu\mu}\Big)W_{W\to\mu\nu}}{B_{\rm cand} - \Big(R^B_{W\to\mu\nu} + R^B_{W\to\tau\nu} + R^B_{Z\to\mu\mu}\Big)W_{W\to\mu\nu}} \\
&\times \Big(C_{\rm cand} - \Big(R^C_{W\to\mu\nu} + R^C_{W\to\tau\nu} + R^C_{Z\to\mu\mu}\Big)W_{W\to\mu\nu}\Big)
\end{aligned} \tag{4.9}
$$

Defining abbreviation

$$
R^X_{\rm sum} \equiv R^X_{W\to\mu\nu} + R^X_{W\to\tau\nu} + R^X_{Z\to\mu\mu}
$$

we rearrange (4.9) to find a quadratic equation for $W_{W\to\mu\nu}$,

$W \to \mu\nu$ selection	Raw Count	% $W \to \mu\nu$ ($\pm$ stat)	% $W \to \tau\nu$ ($\pm$ stat)	% $Z \to \mu\mu$ ($\pm$ stat)	% QCD jets ($\pm$ stat)
Reg. **A**	319,054	14.0 $\pm$ 0.1	1.71$\pm$ 0.04	18.29$\pm$ 0.05	**66.0$\pm$ 0.1**
Reg. **B**	214,377	.049 $\pm$ 0.08	.00004 $\pm$ .000022	.000019$\pm$ 0.00002	**99.928 $\pm$ 0.009**
Reg. **C**	110,179	20.8 $\pm$ 0.2	0.56 $\pm$ 0.04	0.60 $\pm$ 0.01	**78.0 $\pm$ 0.2**
Reg. **W**	5,348,975	**91.52 $\pm$ 0.02**	3.05 $\pm$ 0.01	3.853 $\pm$ 0.005	1.58 $\pm$ 0.01

$W \to e\nu$ selection	Raw Count	% $W \to e\nu$ ($\pm$ stat)	% $W \to \tau\nu$ ($\pm$ stat)	% $Z \to ee$ ($\pm$ stat)	% QCD jets ($\pm$ stat)
Reg. **A**	1,714,631	1.85 $\pm$ 0.02	0.179 $\pm$ 0.006	5.34 $\pm$ 0.01	**92.64 $\pm$ 0.02**
Reg. **B**	243,238	0.038 $\pm$ 0.007	0.096 $\pm$ .004	0.032 $\pm$ 0.002	**99.920 $\pm$ 0.008**
Reg. **C**	45,947	31.4 $\pm$ 0.5	0.93 $\pm$ 0.08	0.24 $\pm$ 0.01	**67.4 $\pm$ 0.5**
Reg. **W**	5,081,938	**93.62 $\pm$ 0.04**	1.91 $\pm$ 0.01	0.486 $\pm$ 0.002	3.98 $\pm$ 0.04

Table 4.6: The percent composition by production process of $W \to \mu\nu/e\nu$ candidates in regions **A**, **B**, **C**, and **W**. The largest percent contribution in each region is in bold. We see that the majority of candidate events in Regions **A**, **B**, and **C**, are QCD jet events (which, recall, is really a measure of all background events that are not electroweak—the vast majority of such events are expected to be QCD jet events). Uncertainties in the table above are statistical only; systematic uncertainties are taken into account in Chapter 7.

$$\begin{aligned} 0 \;=\; & \Big(W_{W\to\mu\nu}\Big)^2 \cdot \big(R^B_{\text{sum}} \cdot R^W_{\text{sum}} - R^A_{\text{sum}} \cdot R^C_{\text{sum}}\big) \\ & + \Big(W_{W\to\mu\nu}\Big) \cdot \big(R^A_{\text{sum}} \cdot C_{\text{cand}} + R^C_{\text{sum}} \cdot A_{\text{cand}} - R^B_{\text{sum}} \cdot W_{\text{cand}} - R^W_{\text{sum}} \cdot B_{\text{cand}}\big) \\ & + B_{\text{cand}} \cdot W_{\text{cand}} - A_{\text{cand}} \cdot C_{\text{cand}} \end{aligned} \tag{4.10}$$

Once we have solved for $W_{W\to\mu\nu}$, we can trace back with (4.8) to find the fraction of each type of background event in each quadrant. We do this, and put the results of our background analysis in Table 4.6. We see that, as assumed, the majority of events in Regions **A**, **B**, and **C** are QCD jet events (a term which, recall, is expanded here to mean all non-electroweak background events). The majority of events in signal region **W** are, as hoped, found to be signal events (92% for $W \to \mu\nu$, 94% for $W \to e\nu$).

Chapter 5

Charmed Meson (D^*) Selection

Now that we have identified our W candidates, we want to identify D^* mesons that have been produced in association with these candidates. In this chapter, we discuss our general procedure for tagging $D^* \to D^0(\to K\pi)\pi_s$ decays in a collision event. We count D^* produced in association with our W/Z selections, and identify sources of D^* background candidates. We train a neural network to discriminate between signal and background D^* candidates, and we then use this neural network to improve the statistical uncertainty of our D^* counts.

5.1 Identifying $D^* \to D^0(\to K\pi)\pi_s$ decays

We identify charm by looking for $D^*(2010) \to D^0(\to K\pi)\pi_s$ decays. For real D^*, the difference in invariant mass between the reconstructed D^* and D^0 vertices forms a sharp peak above background near $\Delta m_{peak} = 0.1456$ GeV; by measuring the magnitude of this peak, we may count the number of D^* decays in our sample. We break the $D^*(2010) \to D^0(\to K\pi)\pi_s$ identification process down into five steps, as described in the following five subsections.

5.1.1 Selecting D^* Decay Tracks

Let us assume that we have identified an event of interest: for example, if we are looking for $W/Z + c$ production, let us assume that we have identified a W or Z event. For $W \to \ell\nu$ decays, and for $Z(\ell^+\ell^-)$ decays in which only one lepton has a track, we begin by selecting all tracks within 2.0 cm of the high-p_T lepton's point of closest approach to the beamline. For Z decays in which both leptons have a track, we use the point of closest approach z_0 of the high-p_T trigger lepton (this preference should have no effect on our final results for real Z).

For each possible set of three tracks taken from our selection, we 'hypothesize' that one is the K, one the π, and one the π_s of a D^* decay. This includes assigning the appropriate K or π mass to each track (*e.g.*, when calculating invariant masses). We then make a number of checks against this hypothesis in order to

eliminate background.

5.1.2 Testing a set of tracks for the D^* decay signature

We first require that the K and π have opposite electric charge, as determined by the curvature of their tracks in the magnetic field. Then, as $D^0 \to K^-\pi^+$ is Cabbibo-favored versus $D^0 \to K^+\pi^-$ by a factor of 10^4, we require the two pions to have the same charge (note $D^{*+} \to D^0\pi^+$ and $D^{*-} \to \bar{D^0}\pi^-$). We also enforce a track quality cut, requiring each track to have at least 20 hits in the axial layers of the COT, at least 20 hits in the stereo layers of the COT, and $|\eta| < 1.1$. These hit requirements make it less likely that our 'track' is just a random assortment of hits falsely recognized as a true track, and the $|\eta|$ requirement confines our tracks to the fiducial region of the detector.

We then make additional kinematic selection cuts. The transverse momenta of the particles produced in this decay are expected to be higher on average than those of background tracks. Based on MC simulations, we enforce lower limits $p_T(K, \pi) > 400$ MeV, and $p_T(\pi_s) > 80$ MeV. Finally, due to the relativistic momentum of the decay products, we expect their opening angles to be relatively small. We therefore require $\Delta R < 1.1$ for each pair of tracks (where $\Delta R \equiv \sqrt{\Delta\phi^2 + \Delta\eta^2}$).

5.1.3 Fitting the tracks

With these preliminary cuts in place, we use the CDF TCtvmft fitting algorithm to reconstruct D^0 and D^* vertices from the K, π and π_s track candidates. This algorithm tests the validity of the hypothesis that two (or more) tracks have come from a common vertex, by moving each track about within the limits of uncertainty allowed by the resolution of the detector. Once a maximally-probable intersection point has been formed, the tracks are updated to reflect their new trajectories, and the probability of the fit is returned. A vertex can also be required to 'point' back towards another, to accomodate the following case: a particle is produced in a decay; it travels some finite distance away from the initial decay vertex; and it then decays itself, producing a secondary vertex.

In our case, we require that the K and π tracks form a D^0 production vertex, and that the momentum of this D^0 vertex points back towards and intersects with the π_s track, forming the D^* vertex. Because we are interested not only in direct charm, but also in charm from bottom decays, we do *not* limit our search to prompt D^* production: as such, we do not require the D^* to point back towards the beamline. (A D^* meson decays through the strong interaction, and as such it has a very short lifetime; it decays essentially immediately after production, and so a prompt D^* decay vertex would sit very close to the collision point.) We reject events only if the probability of the fit is 0 (*e.g.*, if the fit did not converge). For events with a

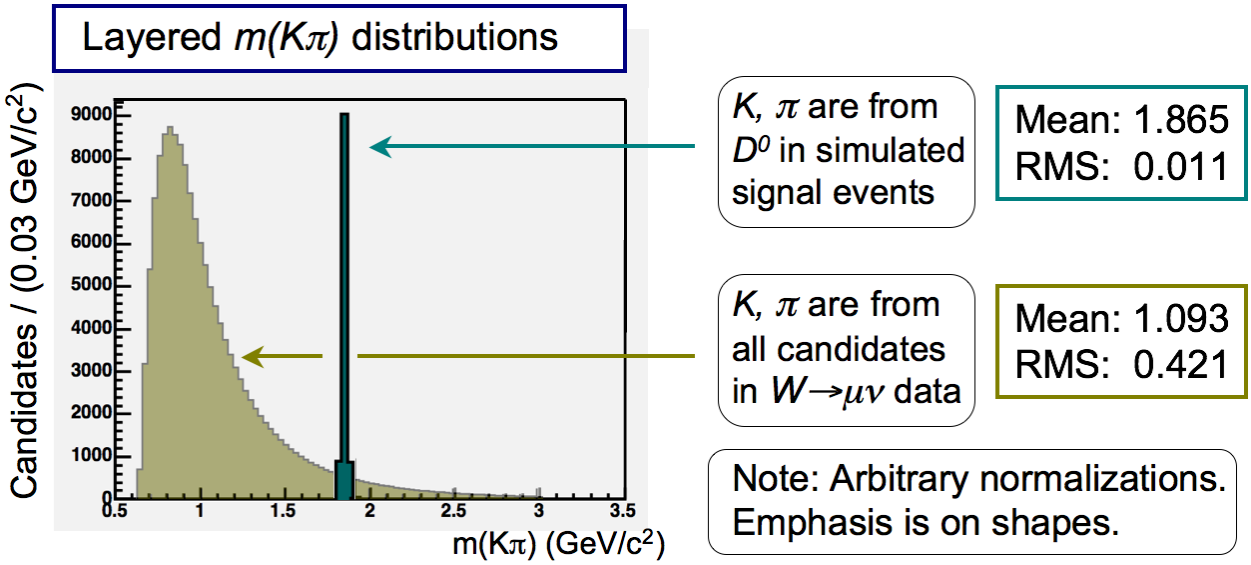


Figure 5.1: D^0 mass distributions for Monte Carlo $f + g \to W + c$ events (teal), and all D^* candidates in our $W \to \mu\nu$ selection that pass every cut up to (but not including) the D^0 mass cut.

non-zero fit probabilty, we store this probability and the χ^2/DOF of the fit for later analysis.

From these fitted tracks, we now construct the position vector $\mathbf{V}$ and four-momentum $p^{(4)} = (E, \mathbf{p})$ of each vertex. From the four-momenta, we can calculate the invariant mass of each vertex as $m = \sqrt{E^2 - \mathbf{p}^2}$. While we do nothing with the D^* mass at this point, we require the fitted D^0 mass to fall within $3\sigma = 0.03321$ GeV of the nominal D^0 mass peak, $m_{D^0} = 1.865$ GeV (see Fig. 5.1). This rejects over 99% of background, while eliminating only about $\sim 0.3\%$ of the signal.

We now use these reconstructed vertices to define a few quantities, that we will later use to describe our events with more precision. First, the position vector is used to determine the signed impact parameter $L_{xy} = \mathbf{p} \cdot \mathbf{V}/|\mathbf{p}|$ of each vertex. L_{xy} is positive if $\mathbf{p}$ points in the same direction as $\mathbf{V}$, and negative otherwise. For any particle which decays with a finite lifetime, we would expect $L_{xy} > 0$. However, the finite resolution of the detector allows real vertices with small $|\mathbf{V}|$ to give zero or negative L_{xy}. For this reason, we do not cut on L_{xy} of the D^* or D^0 vertex. We will later on use these L_{xy} values to separate signal events from background events, using a neural network (see 5.2). We also find L_{xy} of the reconstructed D^0 vertex with respect to the D^* vertex—in other words, $L_{xy} = \mathbf{p} \cdot \mathbf{V}/|\mathbf{p}|$ in which $\mathbf{V}$ is the displacement vector between the two reconstructed vertices, $\mathbf{V}$=$\mathbf{V}(D^0)$-$\mathbf{V}(D^*)$, and $\mathbf{p}$ is the momentum of the reconstructed D^0. We call this quantity Vtx Lxy.

Finally, we construct a set of signed impact parameters specific to each track. For b defined as the distance of closest approach between a track and the D^* or D^0 vertex, we define b_s as $b_s \equiv b \cdot \text{sign}(\cos(\phi))$, with ϕ as defined in Fig 5.2. We have six such parameters in all, one for each track paired with each vertex.

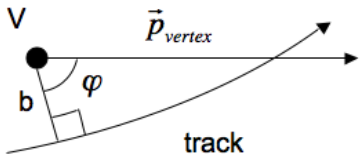


Figure 5.2: Sketch of a track and vertex for which b_s might be calculated. ϕ is the angle between the momentum of the reconstructed vertex, and the line connecting the track and vertex at the point of closest approach. We define b as the distance between the track and vertex at this point.

We write these as $b_s(\text{track}, \text{vertex})$, though the first argument can be replaced with another vertex if, *e.g.*, we are interested in the reconstructed D^0's point of closest approach to the beamline; in this case, instead of a track (the D^0 leaves no track in the COT), the reconstructed vertex momentum is traced backwards towards the target vertex.

These b_s parameters are complementary to the L_{xy} measurements, and provide a more detailed look into the kinematics of our decay near the production point. In particular, we expect tracks from the D^0 decay in signal events to produce b_s distributions with respect to the D^* vertex that are statistically separated from 0. This is in contrast to the combinatoric background, which has vanishing average impact parameters.

5.1.4 Binning Δm and counting signal

At this point we have enforced our cuts, performed the vertex fits, and rejected all events that failed these fits, eliminating a large amount of background. The last step towards counting D^* signal is to bin Δm for all remaining D^* candidates, and to fit the resulting histogram to a double-gaussian signal plus power-law background hypothesis, in order to count the number of tagged D^*. (We use templates for the signal and the background functions, and allow only the amplitude of each function to vary. For a description of these templates, and how we derived them, see Appendix A.) Before moving on to the next section, we tabulate all cuts used to eliminate background thus far, and we list all quantities calculated to describe our D^* candidates (Table 5.1).

5.1.5 Counting D^* in our W/Z selections

Having identified Z and W events in sections 4.2 and 4.3, we now use the machinery developed in this section to identify D^* produced in association with these vector boson events. We look separately at $Z \to \mu\mu$, $Z \to ee$, $W \to \mu\nu$ and $W \to e\nu$ events; we also combine decay modes to produce single $Z \to ee/\mu\mu$ and $W \to e\nu/\mu\nu$ samples, and then finally split the latter into two different subsets: events for which the W and tagged D^* have the same sign ($W + D^*_{SS}$ events), and events for which the W and tagged D^* have the

Variable	Signal Requirements	Variable	Signal Requirements
$p_T(K)$	400 MeV $< p_T$	$p_T(D^0)$	Not used for cuts
$p_T(\pi)$	400 MeV $< p_T$	$\lvert\mathbf{V}_{D^*} - \mathbf{V}_{D^0}\rvert$	Not used for cuts
$p_T(\pi_s)$	80 MeV $< p_T$	$b_s(K, D^*)$	Not used for cuts
$\Delta R(K, \pi)$	$\Delta R < 1.1$	$b_s(\pi, D^*)$	Not used for cuts
$\Delta R(K, \pi_s)$	$\Delta R < 1.1$	$b_s(\pi_s, D^*)$	Not used for cuts
$\Delta R(\pi, \pi_s)$	$\Delta R < 1.1$	$b_s(K, D^0)$	Not used for cuts
Fit χ^2/DOF	χ^2/DOF> 0	$b_s(\pi, D^0)$	Not used for cuts
Prob(χ^2/DOF)	Prob > 0	$b_s(\pi_s, D^0)$	Not used for cuts
$L_{xy}(D^*)$	Not used for cuts	$b_s(D^*,$ beam)	Not used for cuts
$L_{xy}(D^0)$	$L_{xy} > 0$	$b_s(D^0,$ beam)	Not used for cuts
$p_T(D^*)$	Not used for cuts	Vtx Lxy	Not used for cuts

Table 5.1: All properties used to distinguish between signal and background events. Requirements on each property—which must be satisfied in order for an event to be considered signal–are located under 'Signal Requirements'. Properties marked 'Not used for cuts ' have no absolute cuts, but may be used later on in other contexts (*e.g.*, as input to a neural network).

opposite sign ($W + D^*_{OS}$ events). We will later use these OS/SS counts to identify the fraction of our sample that comes from a particular production process (see 10.3). The relevant plots are shown in Figs. 5.3 and 5.4; signal counts are recorded in Table 5.2.

In the case of our $Z(\to \ell\ell) + D^*$ samples, we have a relatively small signal. In order to assess the strength of that signal (is it real signal, or is it a statistical fluctuation?), we can perform a significance test. In short, we fit the Δm distribution to our full signal plus background hypothesis, and then to a hypothesis of background only. We record the fit quality value χ^2/DOF for each (where DOF $\equiv$ number of degrees of freedom in the fit), and we look at the difference between these values. A value of $\chi^2/\text{DOF} \sim 1$ usually indicates a good fit, and if the signal plus background hypothesis is a much better fit than the hypothesis of background only, then we say that our signal is significant. Quantitatively [40], for a signal with gaussian-distributed uncertainty, and for a fit in which removing the signal decrements DOF by only 1, we may calculate

$$\text{Significance } / \ \sigma = \sqrt{\chi^2_{\text{sig+bkg}} - \chi^2_{\text{bkg only}}} \tag{5.1}$$

This quantity is only truly meaningful in the context of small signals, such as in the Δm distributions of our $Z + D^*$ samples. However, for the sake of comparison with our later, improved measurements (see 5.2), we measure this quantity for all of the $V + D^*$ signal types described above (Table 5.2). We see that our $Z + D^*$ signals have a significance that falls just below 3σ; this might be worrisome if we were looking to prove the existence of a new process or particle (*e.g.*, the Higgs boson). Such measurement usually requires

Category	N_{count} (Number D^* tagged)	$\chi^2_{\text{sig+bkg}}$ / DOF	$\chi^2_{\text{bkg only}}$ / DOF	Significance / σ
$W_{e\nu} + D^*$	367 ± 51	121.1 / 80	183.8 / 81	7.9
(with $\text{NN}_{cut} = 0.0$)	340 ± 30	100.8 / 79	292.7 / 80	13.9
$W_{\mu\nu} + D^*$	326 ± 43	90.4 / 79	158.5 / 80	8.3
(with $\text{NN}_{cut} = 0.0$)	294 ± 26	111.3 / 78	305.9 / 79	13.9
$W_{e\nu/\mu\nu} + D^*$	693 ± 66	121.3 / 79	247.5 / 80	11.2
(with $\text{NN}_{cut} = 0.0$)	634 ± 39	108.9 / 79	469.7 / 80	19.0
$W_{e\nu/\mu\nu} + D^*_{OS}$	383 ± 49	109.3 / 79	181.7 / 80	8.5
(with $\text{NN}_{cut} = 0.0$)	362 ± 30	114.6 / 78	328.9 / 79	14.6
$W_{e\nu/\mu\nu} + D^*_{SS}$	310 ± 45	115.2 / 80	175.5 / 81	7.8
(with $\text{NN}_{cut} = 0.0$)	272 ± 26	101.4 / 79	253.1 / 80	12.3
$Z_{ee} + D^*$	48 ± 17	112.8 / 76	120.4 / 77	2.8
(with $\text{NN}_{cut} = 0.0$)	17 ± 8	86.4 / 69	84.8 / 70	0.0
$Z_{\mu\mu} + D^*$	30 ± 12	139.2 / 74	145.8 / 75	2.6
(with $\text{NN}_{cut} = 0.0$)	310 ± 45	45.4 / 55	48.1 / 76	1.6
$Z_{ee/\mu\mu} + D^*$	78 ± 21	101.7 / 76	118.0 / 77	4.0
(with $\text{NN}_{cut} = 0.0$)	39 ± 11	113.1 / 73	123.5 / 74	3.2

Table 5.2: Tagged D^* counts as derived from our fitted plots of Δm for each type of signal event. We include counts found both with (top row of each result) and without (bottom row of each result) applying a neural network to reduce background. The neural network is discussed further in 5.2. V_{xy} is short for $V \to xy$. $V_{xy/ab}$ is short for a combined sample of $V \to xy$ and $V \to ab$ events.

a 3σ confidence level to be of interest, and a 5σ confidence level to be accepted as a real signal. However, when measuring the strength of a process which is already known to exist, lower significance values (such as those of our $Z + D^*$ signals) are acceptable for claiming observation. We have no such qualms with our $W + D^*$ signals: these are not only easily visible to the "naked" eye in our Δm distributions (Fig. 5.3), but are far beyond the 3-5σ significance threshold (Table 5.2).

5.2 Using a neural network to reduce background in Δm

In the previous section, we identified D^* candidates produced in association W and Z decays. We did this by binning Δm for all D^* candidates passing a set of cuts, and then fitting these Δm histograms to a signal plus background hypothesis. However, we would also like to look at differential measurements, *e.g.*, the number of D^* produced as a function of $p_T(D^*)$, or the number of D^* produced in events with N jets. By dividing up our D^* sample, we can provide more detailed final results; however, this will also reduce our signal strength, which in turn increases uncertainty in our measurements. In order to counteract this increased uncertainty and in order to provide the most meaningful final results, we want to eliminate as much background as possible in our Δm distributions. The less background there is, the stronger our signal will stand out over this background. This, then, is our next objective; to approach it, we turn to neural networks.

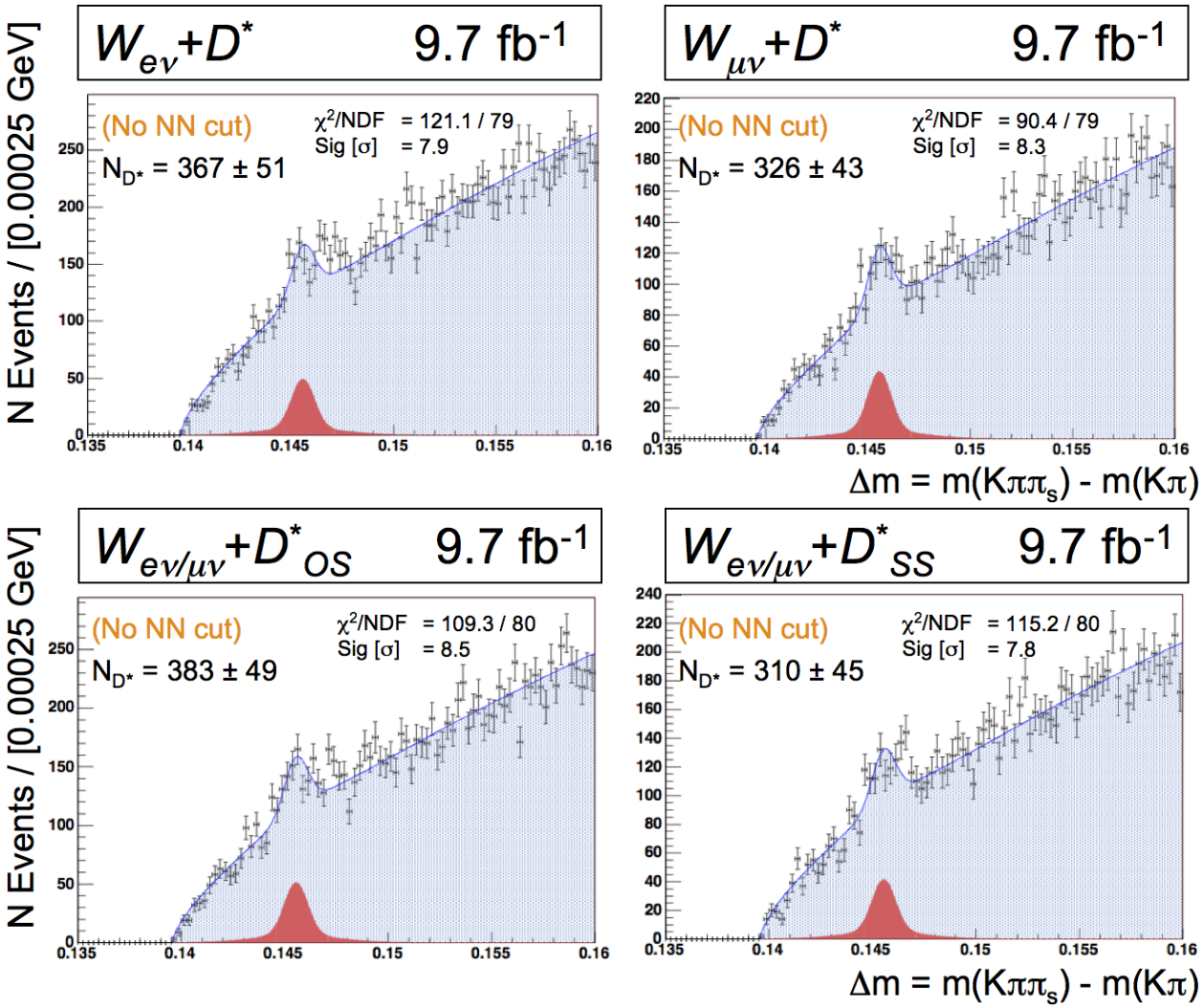


Figure 5.3: Fitted plots of Δm for $W(e\nu/\mu\nu) + D^*$ events.

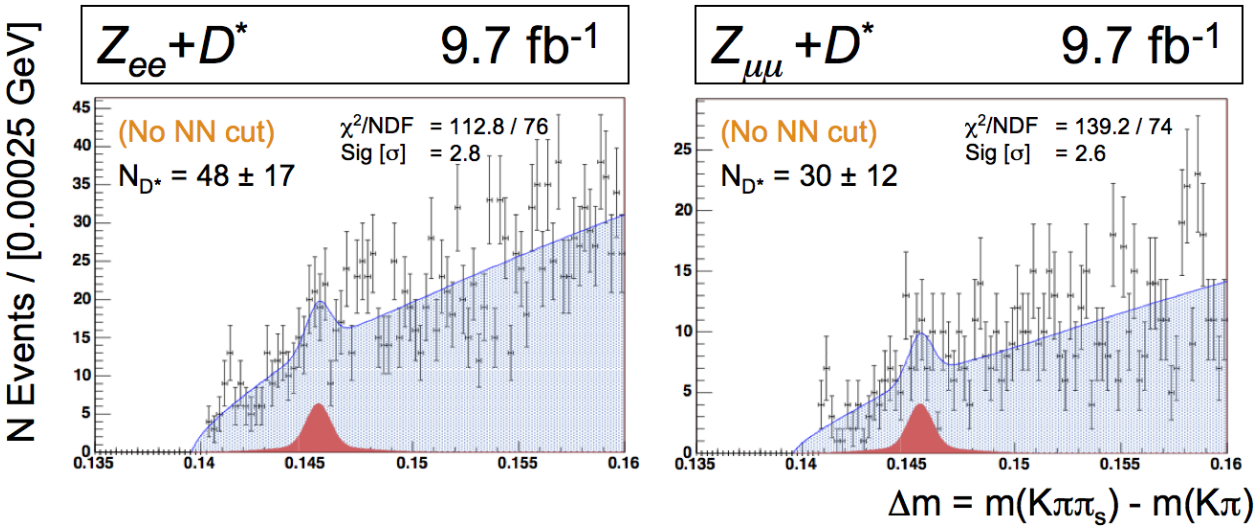


Figure 5.4: Fitted plots of Δm for $Z(ee/\mu\mu) + D^*$ events.

5.2.1 What is a neural network?

An artificial neural network (ANN) is a multivariate analysis tool. Unlike hard cuts, which set a clear line between what is considered background and what is considered signal (*e.g.*, our cuts on the reconstructed D^0 mass), a neural network cut allows these lines to be blurred by considering many properties in tandem. A D^* candidate with a large, negative L_{xy} value for example, might be D^*-like enough in all other ways that it will still pass our cuts. This kind of flexibility is very useful if your events are described by a large array of variables, none of which provide particularly clear boundaries between signal and background events.

A neural network gets its name by merit of its input and output structure, which can be said to resemble a web of interlocking neurons in the human brain. In the brain, a single neuron takes in inputs from many sources, through its dendrites, and produces a single output value; this value is carried along its axon, and then distributed as input to other neurons. The process is repeated many, many times, until some cell chooses to act (or not act) based upon the value delivered to it by some final neuron. The way that a brain learns to perform a certain to task, or to distinguish between a square and a circle, is by juggling these input and output values until a desired result is obtained.

An artificial neural network works in much the same way; the primary difference is that in an ANN, functions takes the place of chemical interactions. The network begins with a series of inputs, called simply the input layer. Each node in this input layer represents a single variable, which is provided by the user. These variables are sent to the nodes of a second, 'hidden' layer, each node of which combines these input values according to some function to produce a single output value. This value, in turn, is sent to the output layer, where the process is repeated; a neural network can have many hidden layers, and its final output layer can consist of one or more nodes. The primary purpose of each output node, however, is to represent in a single value how signal- or background-like an event is. In this way, many inputs can be consolidated into a single value representing the entire event; this value can be cut upon to provide a binomial yes-or-no decision. The general ANN structure is shown in Figure 5.5.

Like a biological neural network, a neural network must 'learn' to perform a task. This is accomplished by defining each function in the hidden layer with a large array of variable constants, which can be adjusted until the desired behavior is obtained. In general, a user provides the neural network with sample signal events and background events, which he wants to separate in other samples with unknown contents. The constants in the hidden layer are adjusted, until the final output value is a reliable measure of whether an event is signal- or background- like. This usually takes the form of an output value between -1 and 1, where the closer to -1 this value is, the more background-like the event. These adjustments are called 'training'.

We use the Neurobayes package in this analysis [41], a commercial software product used by the CDF

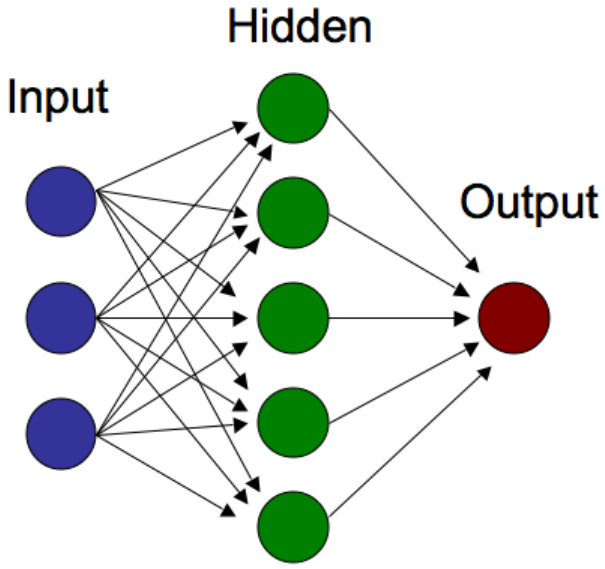


Figure 5.5: A schematic representing the structure of a neural network. The input layer is an array of variables, one per node, which describe an event. These inputs are each passed to the nodes of a second, 'hidden' layer, where they are processed to produce a second set of values. These values are then sent to an output layers (or to one or more additional hidden layers), where the process is repeated. At the end, each output node produces a single value representing how signal- or background- like an event is. Cuts on these final values are used to decide whether an event is signal- or background-like.

experiment. The hidden layers of this package are encrypted, and we are therefore unable to explain how they work in this thesis. Suffice it to say, it is very likely that they work exactly as described above. In the following subsections, we will describe the types of signal and background events that we wish to separate, and the variables (inputs) that we will use to create and train our neural network.

5.2.2 An interlude on production processes

Recall that to first order, the $W/Z + D^*$ signals measured in section 5.1 come from four different production processes: $s(d) + g \rightarrow W + c(\rightarrow D^*)$ and $p\bar{p} \rightarrow W/Z + c\bar{c}(\rightarrow D^*)$ (direct charm production); and $p\bar{p} \rightarrow W + B(\rightarrow D^*)$ (bottom hadron decays). For reference, these four production processes, and the abbreviations with which we refer to them, are re-listed in Table 5.3.

As we develop our neural network, we will need to keep these different production processes in mind. Let us assume, for example, that our trained neural network comes to identify the reconstructed D^* momentum as a critical discriminant for separating signal and background. While it is unlikely that D^* produced in *Wcc* vs *Zcc* events will have properties that are substantially different from one another, it *is* possible that the momentum distribution of D^* from *Wbb* events will be quite different than the momentum distribution of D^* from *Wcc* or *Wc* events. It is also possible that D^* from *Wcc* will have a momentum distribution

Production process	Abbreviation	Number of Monte Carlo signal events
$q + \bar{q}' \to W + g(\to b\bar{b}) \to W + D^* + X$	Wbb	2710
$q + \bar{q}' \to W + g(\to c\bar{c}) \to W + D^* + X$	Wcc	8437
$q + g \to W + c \to W + D^* + X$	Wc	7090
$q + \bar{q} \to Z + g(\to c\bar{c}) \to W + D^* + X$	Zcc	4995

Table 5.3: Our four production processes, along with the abbreviations used to refer to each, and the number of D^* candidates in each sample that fall within 3σ of the Δm peak.

different than that of D^* from Wc. While we do not expect the effect to be large, it is possible that this will cause, *e.g.,* NN_{Wcc} (a neural network trained to identify Wcc events), to preferentially select D^* from Wcc, over D^* from other production processes.

In practice, this neural network bias will not affect our final inclusive measurements (see the $A \cdot e$ 'unfolding' process in chapter 8). However, it provides some interesting insight into the nature of our signal. In fact, we will even take advantage of this bias and use it to separate Wcc and Wbb signal contributions from one another in 10.1. For now, we will simply keep track of the neural network acceptance rates for each of these four production processes.

Finally, we expect for most of our D^* signal to come from Wcc events. Therefore, in what follows, we will train our neural network to identify D^* from Wcc production. This choice does cause us to identify Wbb and Wc signal at a lower rate than Wcc signal, but the effect turns out to be small (see Figure 5.9), and may be easily accounted for.

5.2.3 Training the neural network

To create our neural network, we must first do two things: decide which variables to send to the ANN input layer; and collect sample signal and background events to train the neural network with. This may may at first seem contradictory; if we know which events are signal and which are background, then why train a neural network to begin with? The answer is that in the signal region (beneath the peak in our Δm plot), we have no way of knowing which events are background events, and which are signal. The signal peak is a statistical phenomena that rises above this uncertainty. We can, however, look at signal and background events in purer samples, and then use these to distinguish between signal and background in 'contested' regions.

First, we assemble a list of variables which can be used by the neural network to separate signal and background (Table 5.4). This list contains mostly kinematic variables (impact parameters, opening angles, and track/vertex momenta), as well as the vertex fit probability and χ^2/DOF value. In practice, the neural network does not use all of these variables (if changing some variable has only minimal impact on the final output value, a Neurobayes neural network will ignore it). Those that it does, we list in the order of

Rank	Variable
1	$L_{xy}(D^0)$
2	$p_T(K)$
3	$b_s(\pi_s, D^*)$
4	$\Delta R(K, \pi)$
5	$b_s(\pi, D^*)$
6	$b_s(K, D^0)$
7	$\Delta R(\pi, \pi_s)$
8	Fit χ^2/DOF
9	$p_T(\pi)$
10	$b_s(D^*,$ beam)

Rank	Variable
11	$L_{xy}(D^*)$
12	$p_T(D^*)$
13	$p_T(\pi_s)$
14	$\Delta R(K, \pi_s)$
15	$b_s(D^0,$ beam)
16	$\lvert\mathbf{V}_{D^*} - \mathbf{V}_{D^0}\rvert$
17	Vtx Lxy
18	$b_s(K, D^*)$
19	$b_s(\pi, D^0)$

Rank	Variable
Unused	$p_T(D^0)$
Unused	$b_s(\pi_s, D^0)$
Unused	Fit Prob

Table 5.4: The properties used to characterize our D^* candidates, both signal and background. It is these properties that the neural networks will use to separate signal and background. We also list the importance of each of these properties, as ranked by the neural network training program. Three of these variables ($p_T(D^0)$, $b_s(D^0, \pi_s)$ and Fit Prob) are determined by the training program to have negligible use in separating signal and background. These variables are, therefore, not used by the neural network. For more information about the variables listed above, please refer back to section 5.1.

importance assigned to them by the final, trained neural network.

5.2.3.1 Defining our signal sample

Now, we collect sample signal events for training. For this, we generate inclusive Pythia Monte Carlo $p\bar{p} \rightarrow W/Z$ events, skimming off only those events in which a D^* is produced in one of the four ways described aboove. We simulate these events in a detector environment using cdfSim, and then run them through our $W/Z + D^*$ tagging algorithms as if they were data. In general, Pythia Monte Carlo is very good at simulating the kinematic behavior of signal events; it is only the magnitude of these production processes that are considered unreliable until measured. For our purposes, then—training a neural network to recognize the kinematic properties of signal events—these simulated events will serve very well.

In the simulated samples described above, any D^* candidates that fall within 3σ of the Δm peak, $\Delta m_{peak} = 0.1455$, are considered signal events. While it is possible for some combinatoric background to fall into this region (tracks in signal events that do not come from a D^* decay may pass all of our cuts), this effect is considered to be negligible (contribution $\ll 1\%$). The properties of each type of signal event are stored in a data format called a ROOT Tree. These trees (Tree $_{Wbb}$, Tree $_{Wcc}$, etc) can then be used to train our neural network, NN_{Wcc}, and to estimate the efficiency with which this neural network identifies D^* from each production process.

5.2.3.2 Defining our background sample

We now need a sample of "D^*-like" background events. We expect that most of our background will be combinatoric background from underlying QCD processes; there are no known standard model processes that can mimic the Δm peak of the $D^* \to D^0(K\pi)\pi$ decay. While Monte Carlo simulations are quite good at predicting the kinematics of specific signal processes, they are not as reliable when it comes to predicting the behavior of general QCD background. Instead, we look to data. Specifically, we look in our selected W events for D^* candidates that pass all of our selection criteria, but do have not the charge assignments needed for a true signal event. (We choose to look in our W events, versus in our Z events, for two reasons: we have many more background candidates in our W sample than our Z sample; and we do not expect D^*-like background in Z events to be very kinematically different from D^*-like background in W events.)

We define two kinds of background according to the criteria above: "same-sign track" (SS) background, in which all tracks have the same sign; and "bad-sign π_s" (BSP) background, in which the soft pion has the opposite charge of the 'hard' pion. The SS background should be purely combinatoric. The BSP background is a combination of combinatoric background, real $D^0 \to K^+\pi^-$ decays paired with a random low-momentum track, and real signal events in which the sign of the soft pion is misreconstructed. We expect for the latter effect to be small; even if it is not, our $A \cdot \epsilon$ calculations (chapter 8) will take this into account when we report our final measurements.

Neural networks are optimally trained using the same number of signal and background events (in our largest Monte Carlo signal sample—the Wcc events—we count 8437 signal D^* candidates; we would also like to select 8437 D^*-like background events). However, in order to provide maximum discriminating power in the region beneath the Δm signal peak, we want our signal and background events to be as kinematically similar as possible. To fit both of these constraints, we combine our SS and BSP background samples, and count outwards from the Δm signal peak at $\Delta m_{peak} = 0.1455$. We select 8437 D^* candidates within 3.2σ of the peak, and define these as our background sample.

5.2.3.3 Bringing it all together

With our signal and background samples defined, we evaluate the variables used to describe each (Table 5.4), and send all of this information to the neural network NN_{Wcc} for training. The separating power of the final, trained neural network, in terms of the final output value distributions, is displayed in Figure 5.6. The distributions of each variable used to develop this neural network (for both signal and background events), are recorded in Appendix B. The correlation matrix for all input variables is also recorded in Appendix B.

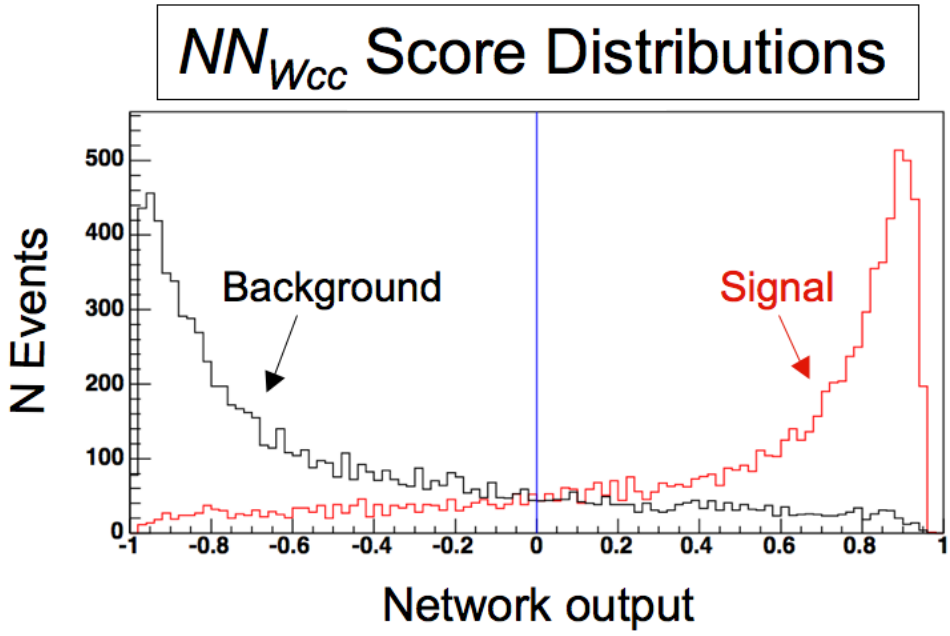


Figure 5.6: The neural network output score spectra for signal and background events, for trained neural network NN_{Wcc}. The signal event spectrum is drawn as a red histogram, while the background event spectrum is black.

5.2.4 Characterizing the neural network

As shown in Figure 5.6, our trained neural network is quite good at separating signal and background events from one another. The network output is close to 1 for the majority of signal events, and close to -1 for the majority of background events. However, in order to actually employ this neural network, we must cut upon its output score, and declare that all events with an output score greater than some value NN_{cut} are signal events.

Recall that our primary motivation in training this neural network is to reduce the number of D^*-like background events, thereby reducing the statistical uncertainty in our fits and our final measurements. To do this, we will want to choose an optimal cut value NN_{cut}. We determine that value by plotting a so-called figure of merit (FOM) as a function of neural network cut score, and looking for that neural net score which gives us the highest FOM. The FOM is, generally speaking, a measure of the 'cleanliness' of a signal amidst background—a high FOM value indicates less noise, hence less uncertainty when measuring the size of the signal peak. For our figure of merit, we choose standard expression

$$\text{FOM} = \frac{s}{\sqrt{s+b}} \tag{5.2}$$

Here, s is the number of signal events that pass a given NN cut value, and b the number of background

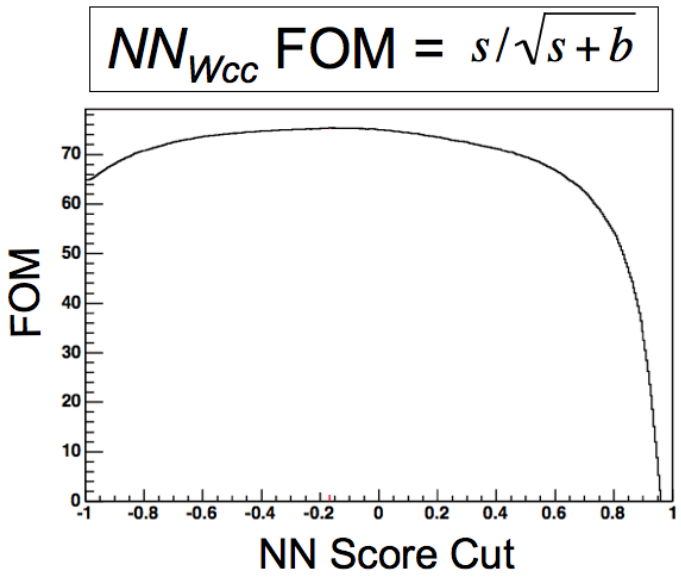


Figure 5.7: We plot FOM $\equiv s/\sqrt{s+b}$ as a function of the neural network score cut, NN_{cut}. A higher FOM score translates into a smaller uncertainty when fitting our discriminant (Δm) to a signal plus background hypothesis. As can be seen here, the 'peak' for the neural network is quite broad, within little change in the FOM value between 0.2 and -0.2. For the sake of simplicity, and for uniformity when using multiple neural networks (as we will in later sections), we choose to use $NN_{\text{cut}} = 0.0$ as our cut value between signal and background.

events that pass that cut value. Using the signal and background samples defined above to find s and b, we generate the FOM plot as shown in Figure 5.9. It is also interesting to look at at a plot of NN_{cut} versus the number of signal and background events that pass that cut, again using the signal and background samples defined above. This lets us predict about how much signal and background will be lost for, *e.g.*, a neural network cut of $NN_{\text{cut}} = 0.0$. We show these plots in Figure 5.8.

Recall that we would also like to see how our cut on the neural network output will affect different signal events from different production processes. For that purpose, we plot the fraction of each type of signal event that has a neural network output score greater than NN_{cut}, for each of our four signal processes (Figure 5.9). We see that for $NN_{\text{cut}} = 0.0$, our Wcc and Zcc samples have a very similar acceptance efficiency, but our Wc and Wbb samples have rather different efficiencies. This is to be expected—we will see later (chapter 8) that the rate at which we accept D^* candidates is directly dependent upon the transverse momentum of the D^* meson ($p_T(D^*)$). In single charm production (Wc), the W and c are produced back-to-back, and we would expect the D^* to have higher average momentum. In double charm production (Wcc), the momentum opposite to the W must be split between two charm quarks—this will reduce average $p_T(D^*)$. Finally, in double bottom production (Wbb), not only is the momentum opposite to the W split between two bottom quarks, but the D^* produced will only carry away a fraction of the momentum given to one bottom quark. As such, Wbb events will produce D^* with the lowest average $p_T(D^*)$. This is exactly the trend that we see

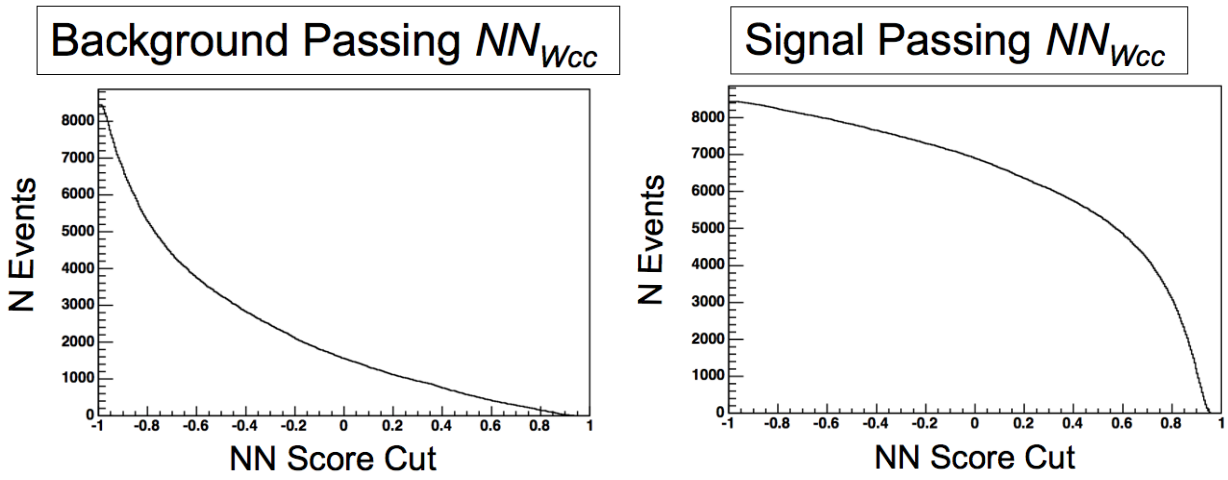


Figure 5.8: The number of signal and background events with NN_{cut} greater than some value, as a function of that value. To generate these plots, we employ the neural network and signal/background samples described earlier in this section (section 5.2). We see that for the choice $NN_{\text{cut}} = 0.0$, we would eliminate $\sim 80\%$ of background, while losing only $\sim 10\%$ of our signal.

in Figure 5.9.

5.3 Refined D^* counts in our W/Z selections

Choosing a cut on neural network score of $NN_{\text{cut}} = 0.0$, we run over all D^* candidates as selected in section 5.1. We then fit the resulting Δm plots. The results are plotted in Figures 5.10-5.12.

In the combined $W + D^*$ selection, we lose about 10% of our signal (as expected). In the combined $Z + D^*$ selection, however, we appear to lose significantly more signal (Figure 5.10). Looking at the $Z \to ee$ and $Z \to \mu\mu$ cases separately (compare Figures 5.4 and 5.12), we see that this loss is due primarily to reduction in the $Z(\to ee) + D^*$ signal. Because we do not expect for the neural network to affect Wcc and Zcc events very differently (recall Figure 5.9), and because the $Z + D^*$ fits involve low statistics, we attribute this signal 'loss' to be due to statistics only—the $Z + D^*$ pre-neural network signal count may have been driven artificially high by a statistical fluctuation (or the post-nerural network $Z + D^*$ signal count may have been driven artificially low).

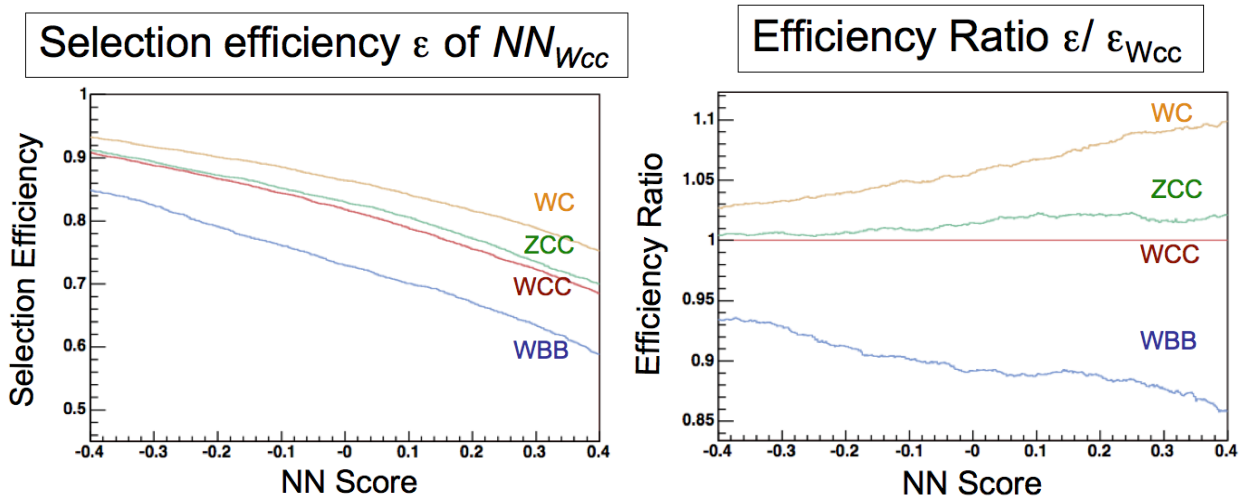


Figure 5.9: The rate at which our neural network accepts D^* candidates from each of four different production processes, plotted as a function of NN_{cut}.

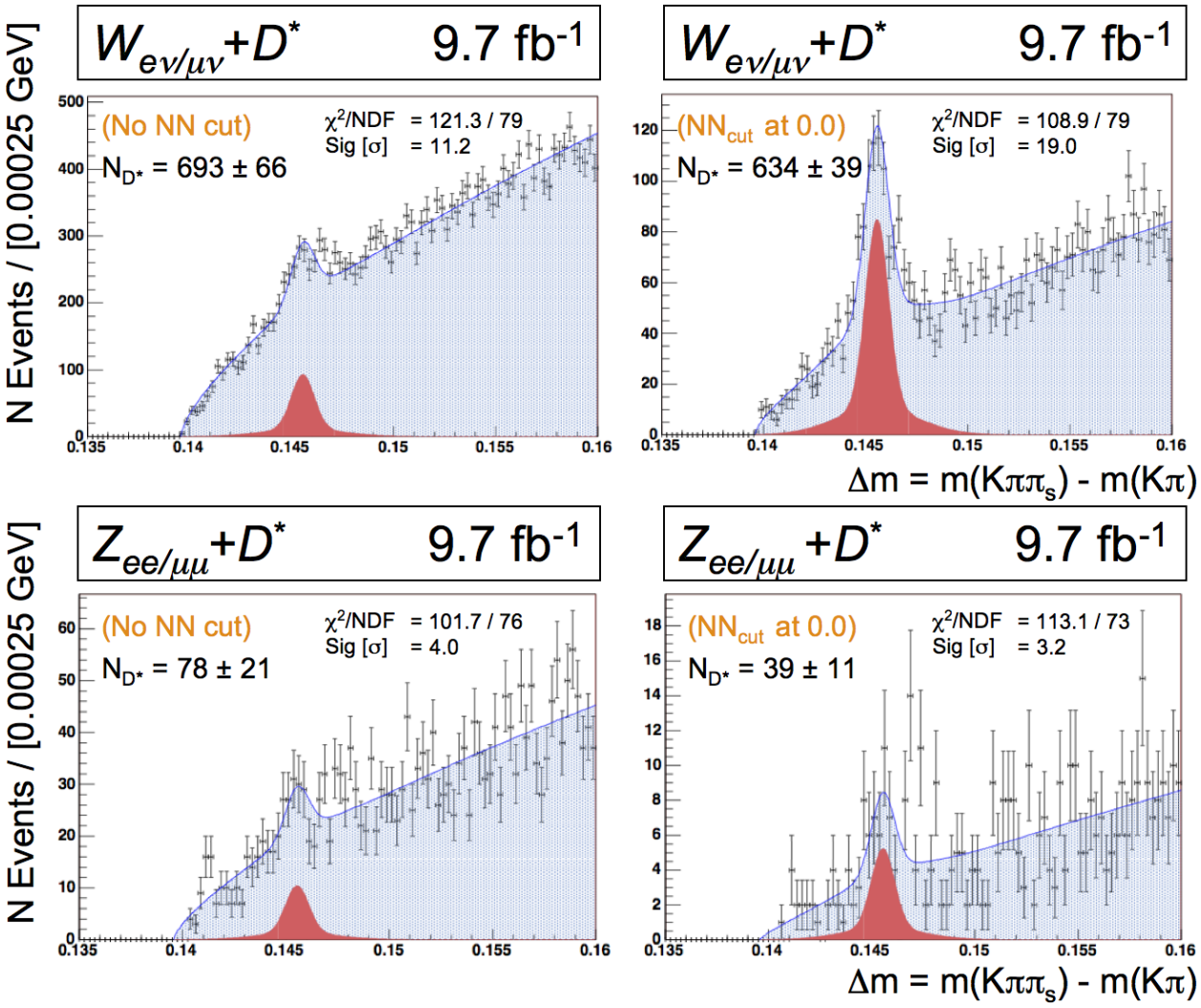


Figure 5.10: Fitted plots of D^* signal discriminant Δm for D^* candidates in both W and Z events. We plot Δm both before and after applying the NN cut of 0.0. We see marked improvement in the fit uncertainty of the $W + D^*$ signal, but less improvement of the $Z + D^*$ signal..

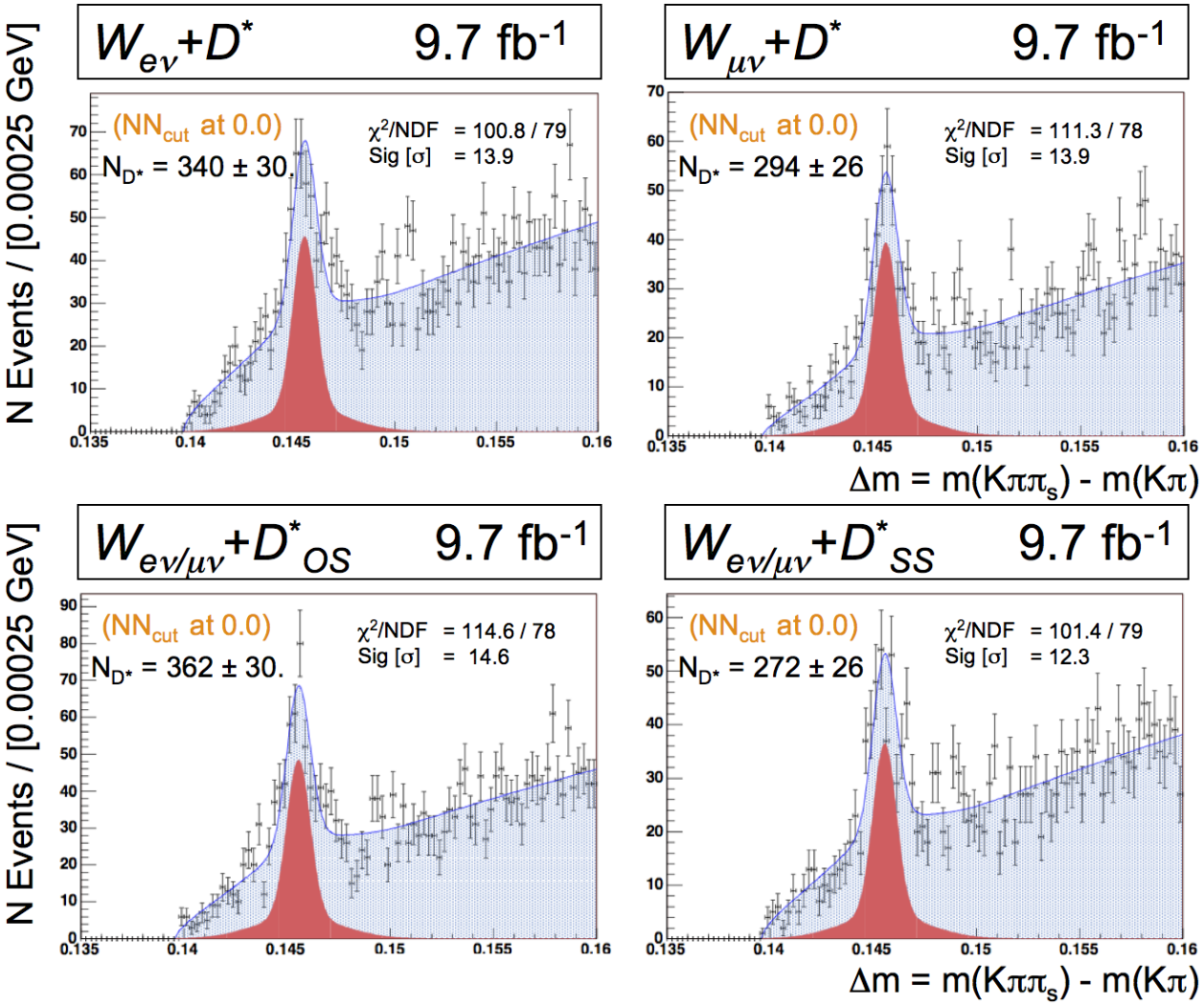


Figure 5.11: Fitted plots of Δm for $W(e\nu/\mu\nu) + c$ events.

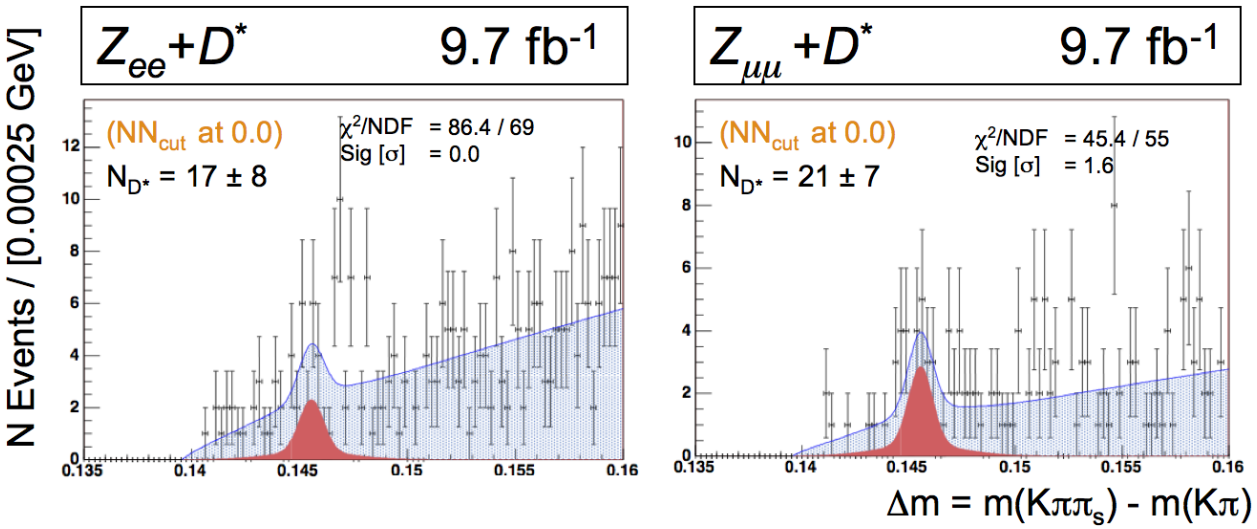


Figure 5.12: Fitted plots of Δm for $Z(ee/\mu\mu) + c$ events.

Chapter 6

Rate of D^* Production in Vector Boson Events

Following the techniques described in Chapter 5, we found low-uncertainty measurements of the number of $D^* \to D^0(\to K\pi)\pi_s$ decays in our selected W and Z events. Let us take a moment, however, to remember what it is that we ultimately wish to measure: a ratio of physical cross-sections $\sigma(V + c)/\sigma(V)$. On the surface, this might appear to be a very simple task. If, for example, every reconstructed W in our $W \to \mu\nu$ selection was a real W, then we could simply divide the number of tagged D^* in each sample, by the number of tagged $W \to \mu\nu$ events in that sample. After applying some 'unfolding' factors to this raw fraction, to take acceptance values and branching ratios into account, we would have our final measurement (chapter 8).

Unfortunately, it is unlikely that every event which passes our W/Z cuts are in fact W/Z events. For example, some "W" events may actually be misreconstructed Z decays, or may be QCD jets which fake the W signature. Further complicating this scenario, some of the D^* that we identified in our Δm plots may be associated with these *fake* W events. If the rate at which fake W/Z are produced in association with D^*, is similar to the rate at which *real* W/Z are produced in association with D^*, then the ratios $\sigma(W/Z + c)/\sigma(W/Z)$ will not be strongly affected. However, we cannot *a priori* assume this.

In this section, we work to remove the influence of background W/Z events from our D^* signal counts. In the process, we will determine the 'raw' fractions $f_{D^*}^{W/Z\ \mathrm{sig}}$ and $f_{D^*}^{W/Z\ \mathrm{bkg}}$, where $f_{D^*} \equiv N_{\mathrm{tagged}}^{W/Z+D^*}/N_{\mathrm{tagged}}^{W/Z}$ for each real (sig) and fake (bkg) W/Z events. As we will see, it is easiest to solve for both $f_{D^*}^{W/Z\ \mathrm{bkg}}$ and $f_{D^*}^{W/Z\ \mathrm{bkg}}$ simultaneously.

6.1 Solving for rate $f_{D^*}^{Z\ \mathbf{sig}}$

We will first determine the fraction of Z events (both real and fake) which are also tagged as having a $D^* \to D^0(\to K\pi)\pi_s$ decay. To do this, we define two regions along the $m(\ell^+\ell^-)$ axis for all Z candidates: the signal region is defined as $|m(\ell^+\ell^-) - 91\ \mathrm{GeV}| \leq 3\sigma_{M_Z}$, and the background region as $|m(\ell^+\ell^-) - 91\ \mathrm{GeV}| > 3\sigma_{M_Z}$. Here, σ_{M_Z} is the width of the Z boson signal peak; due to finite detector resolution that is generally worse for electrons than it is for muons, this value varies between $\sigma_{M_Z} = 3.0$ GeV for $Z \to e^+e^-$ events, and $\sigma_{M_Z} = 2.0$ GeV for $Z \to \mu^+\mu^-$ events.

We then fit the total $m(\ell^+\ell^-)$ distributions to an exponential background plus double-Gaussian signal hypothesis, as in section 4.2; integrate beneath the curves to count the number of signal and background Z candidates in each region, $N_Z^{\text{region=(sig,bkg)}}$ and $N_{Z,\text{ bkg}}^{\text{region=(sig,bkg)}}$; and bin Δm for all D^* candidates, fitting this distribution with our standard Δm signal plus background hypothesis. This gives us counts $N_{D^*}^{\text{region=sig}}$ and $N_{D^*}^{\text{region=bkg}}$. Finally, we construct two simple coupled equations to solve for $f_{D^*}^{Z,\text{bkg}}$ and $f_{D^*}^{Z,\text{sig}}$:

$$\begin{aligned} N_{D^*}^{\text{sig region}} &= f_{D^*}^{sig} \cdot N_Z^{\text{sig region}} + f_{D^*}^{bkg} \cdot N_{Bkg}^{\text{sig region}} \\ N_{D^*}^{\text{bkg region}} &= f_{D^*}^{sig} \cdot N_Z^{\text{bkg region}} + f_{D^*}^{bkg} \cdot N_{Bkg}^{\text{bkg region}} \end{aligned} \tag{6.1}$$

While we do not expect there to be many Z signal events in the background region, it is possible for the rate at which Z are produced with D^* to be much larger than the rate at which background events are produced with D^*. If that is the case, then all D^* in the background region may be from Z events; it is for this reason that we use the coupled set of equations above, instead of finding the background plus D^* rate in the background region, and then using that rate to estimate the number D^* from background in the signal region. The results of the approach described above are reported as "Test Number 4' 'in Table 6.1.

6.2 Systematic uncertainty in $f_{D^*}^{Z\ \text{sig}}$

The technique described in section 6.1 is expected to be accurate; there are no known sources of background which might peak in the signal region and cause a change in the behavior of background as a function of $m(\ell^+\ell^-)$. We test this assumption by looking at Z candidates in which both leptons have the same sign: the $m(\ell^\pm\ell^\pm)$ distribution is smooth across the signal region, with no peaks.

Regardless, we test the accuracy of our $Z + D^*$ rate-finding technique by performing the above analysis for several definitions of the $m(\ell^+\ell^-)$ signal and background regions. We collect the results of these tests in Table 6.1. We fit, with a constant value hypothesis, the full set of $f_{D^*}^{Z\ \text{sig}}$ and $f_{D^*}^{Z\ \text{bkg}}$ test values (Figure 6.1). The mean and uncertainty of these fits are taken to be the final fractions, $f_{D^*}^{Z\ \text{sig/bkg}}$, and their systematic uncertainties. These final results are as shown in Table 6.2.

6.3 Solving for rate $f_{D^*}^{W\ \text{sig}}$

We now wish to determine fractions $f_{D^*}^{W\ \text{sig}}$ and $f_{D^*}^{W\ \text{bkg}}$. The procedure for this is a bit more complicated than that used to find fractions $f_{D^*}^{Z\ \text{sig}}$ and $f_{D^*}^{Z\ \text{bkg}}$ in the previous section. This is due to the lack of a

Test No.	Sig Region $\lvert M_Z - 91.0\ \text{GeV}\rvert \leq$	Bkg Region $\lvert M_Z - 91.0\ \text{GeV}\rvert >$	$f_{D^*}^{Z\to\mu\mu}$ $(\times 10^{-5})$ $(\pm\text{stat})$ $(\sigma_{M_Z} = 2.0)$	$f_{D^*}^{Z\to ee}$ $(\times 10^{-5})$ $(\pm\text{stat})$ $(\sigma_{M_Z} = 3.0)$
1	2.0 σ_{M_Z}	2.0 σ_{M_Z}	11 ± 3	5.0 ± 3.1
2	2.0 σ_{M_Z}	3.0 σ_{M_Z}	11 ± 3	5.3 ± 3.1
3	2.0 σ_{M_Z}	5.0 σ_{M_Z}	11 ± 3	5.0 ± 3.0
4	3.0 σ_{M_Z}	3.0 σ_{M_Z}	9.3 ± 3.0	5.9 ± 3.0
5	3.0 σ_{M_Z}	5.0 σ_{M_Z}	9.3 ± 3.0	5.5 ± 3.0
6	3.0 σ_{M_Z}	7.0 σ_{M_Z}	9.2 ± 3.0	5.7 ± 3.1
7	5.0 σ_{M_Z}	5.0 σ_{M_Z}	8.2 ± 2.6	5.1 ± 3.3
8	5.0 σ_{M_Z}	7.0 σ_{M_Z}	8.1 ± 2.7	5.4 ± 3.6
9	2.0 σ_{M_Z}	7.0 σ_{M_Z}	11 ± 3	5.2 ± 3.1

Table 6.1: Tables containing the $f_{D^*}^{Z\ \text{sig}}$ values found by varying the definition of signal and background regions in our $Z \to \mu\mu$ and $Z \to ee$ selection samples. The signal (background) region is defined as all events for which M_Z is within (without) N σ_{M_Z} of the nominal peak at $M_Z = 91.0$ GeV. We approximate $\sigma_{M_Z} = 2.0$ for $Z \to \mu\mu$, $\sigma_{M_Z} = 3.0$ for $Z \to ee$. Statistical uncertainty in each test is dominated by uncertainty in the Δm fits.

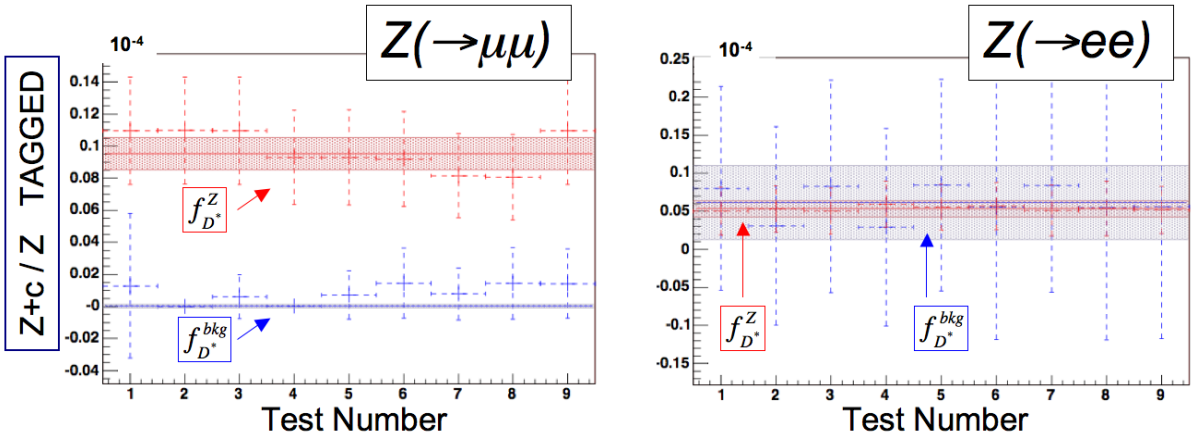


Figure 6.1: Plot of the $f_{D^*}^{Z\ \text{sig}}$ (red) and $f_{D^*}^{Z\ \text{bkg}}$ (blue) values found by varying the signal and background regions in our Z selection sample. We fit each set of values to a constant-value hypothesis (solid horizontal lines), and take the uncertainty in the best-fit value to be the systematic uncertainty of this method (shaded horizontal bars).

Signal Type	$f_{D^*}^{Z\ \text{bkg}}$ $(\times 10^{-5})$ $(\pm\ \text{stat} \pm \text{syst})$	$f_{D^*}^{Z\ \text{sig}}$ $(\times 10^{-5})$ $(\pm\ \text{stat} \pm \text{syst})$
$Z(\to \mu\mu) + D^*$	$0.0 \pm 0.2 \pm 0.1$	$9.6 \pm 2.9 \pm 1.0$
$Z(\to ee) + D^*$	$6 \pm 13 \pm 5$	$5.3 \pm 3.0 \pm 1.0$

Table 6.2: Final $f_{D^*}^{Z\ \text{sig/bkg}}$ fractions, with full statistical and systematic uncertainties.

clear signal discriminant such as M_Z. Instead, we break the process into two steps: first, we identify the fraction of selected W which are actually background events faking the W signature; then, we use this value to determine the rate at which real and fake W are produced in association with D^* ($f_{D^*}^{W\ \mathrm{sig}}$ and $f_{D^*}^{W\ \mathrm{bkg}}$). As we will see, there are many different sources of background to consider.

6.3.1 First considerations of EWK backgrounds to the $W + D^*$ selection

Recall the fraction of our W selection that is attributed to backgrounds, as recorded in Table 4.6. For both $W \to \mu\nu$ and $W \to e\nu$ events, we have percent-level non-signal events in region **W**, which is the region that corresponds to our signal selection. Any of these background events may be produced in association with a D^*. If not taken into account, this would contaminate our measurement of rates $f_{D^*}^{W\ \mathrm{sig}}$ and $f_{D^*}^{W\ \mathrm{bkg}}$.

We now assume that the $f_{D^*}^{Z\ \mathrm{sig}}$ rates, as found in the previous section, should be equal across all quadrants as defined in Table 4.5. Any $\not\!\!E_T$ in a $Z + D^*$ event should not be due to the underlying $Z + D^*$ production process, and any variation in the lepton isolation value should be dominated by a structureless background—the $W/Z + D^*$ production processes that we consider produce jets that are preferentially back-to-back with the W/Z decay products. We also assume that the $f_{D^*}^{W\ \mathrm{sig}}$ rates are equal across all regions. For a true $W \to \ell\nu$ event, a mismeasurement of $\not\!\!E_T$ should not strongly affect the odds of tagging an associated D^*. Once again, any variation in lepton isolation should be dominated by a structureless background.

Finally, we argue that $f_{D^*}^{W \to \tau\nu}$ should approximately equal $f_{D^*}^{W \to \mu\nu/e\nu}$. At the production level, these two rates should absolutely be equal—what happens to the W after production should not affect the properties of the associated D^*. However, we may tag D^* at a different rate in $W \to \tau\nu$ events than in $W \to \mu\nu/e\nu$ events: for example, the p_T distributions of tagged W events may differ across decay modes, which in turn would affect the p_T distributions of taggable D^*. For now, we claim that any variation from $f_{D^*}^{W \to \tau\nu} = f_{D^*}^{W \to \mu\nu/e\nu}$ is negligible—especially with the small number of $W \to \tau\nu$ events expected in region W (Table 4.6).

6.3.2 First considerations of QCD backgrounds to the $W + D^*$ selection

This leaves the question of D^* production in QCD-background events which fake the W signature. As we will see, these QCD-background events are the most important source of non-signal D^* in our sample. In general, our $f_{D^*}^{W\ \mathrm{sig}}$ rates are of order 10^{-5}, while our $f_{D^*}^{W\ \mathrm{bkg=QCD}}$ rates (henceforth $f_{D^*}^{\mathrm{QCD}}$, for brevity) can be as large as order 10^{-3}. With $f_{D^*}^{\mathrm{QCD}} \sim 100 \times f_{D^*}^{W\ \mathrm{sig}}$, the rate of D^* production in QCD background is much greater than the rate of D^* production in W signal events—even a few percent of QCD background events in region **W** can strongly affect our D^* count. As such, we will want to understand this rate in as much detail as possible.

There is no *a priori* reason to assume that $f_{D^*}^{\rm QCD}$ is constant across all four quadrants. Events with large lepton isolation (Iso1), may be more likely to come from heavy-flavor production; meanwhile, events with low $\not\!\!E_T$ may describe a different class of *QCD* background (*e.g.*, mismeasured jets), than events with large $\not\!\!E_T$ (*e.g.*, a heavy-flavor decay involving a lepton and neutrino). We therefore would like to measure the dependence of $f_{D^*}^{\rm QCD}$ on $\not\!\!E_T$ and Iso1.

To do this, we first assume that $f_{D^*}^{\rm QCD}$ can be written as a separable product of functions

$$f_{D^*}^{\rm QCD} = h(\not\!\!E_T) \times g(\text{Iso1}) \tag{6.2}$$

We then define $f_{D^*,X}^{\rm QCD}$ as the rate of QCD background plus D^* production in Region **X**. To first order, we may say that:

$$f_{D^*,A}^{\rm QCD} = h_{LO} \times g_{LO} \qquad f_{D^*,B}^{\rm QCD} = h_{LO} \times g_{HI} \qquad f_{D^*,C}^{\rm QCD} = h_{HI} \times g_{HI} \tag{6.3}$$

where h_{LO} is the average value of h in regions A and B (the low-$\not\!\!E_T$ regions), and h_{HI} is the average value of h in regions C and W (the high-$\not\!\!E_T$ regions); $g_{LO/HI}$ is defined similarly, for regions split by Iso1. This simplification allows us to say that:

$$f_{D^*,W}^{\rm QCD} = h_{HI} \times g_{LO} = \frac{f_{D^*,A}^{\rm QCD} \times f_{D^*,C}^{\rm QCD}}{f_{D^*,B}^{\rm QCD}} \tag{6.4}$$

6.3.3 Measuring rates $f_{D^*}^{W\ \textbf{sig}}$ and $f_{D^*}^{W\ \textbf{bkg}}$ (full approach)

By counting the number of D^* in each of the four $\not\!\!E_T$/Iso1 regions, we get four values which can be used to solve for four unknown rates. Using the $f_{D^*}^{Z\ \rm sig}$ rates as found in data, and assuming that $f_{D^*}^{Z\ \rm sig}$ and $f_{D^*}^{W\ \rm sig}$ are constant across all quadrants (as discussed earlier in this section), we may choose these four unknown rates to be $f_{D^*,A}^{\rm QCD}$, $f_{D^*,B}^{\rm QCD}$, $f_{D^*,C}^{\rm QCD}$, and $f_{D^*}^{W\ \rm sig}$. We will solve for these unknowns in the following iterative process.

First, assume that we have some estimate for $f_{D^*}^{W\ \rm sig}$. For each region **A**, **B**, and **C**, we estimate the number of true $W + D^*$ events by multiplying this fraction into the expected number of W events. We then subtract this value from the region's total D^* count. Similarly, we subtract the number of D^* expected to be from Z events. What's left in each quadrant is assumed to be the count of D^* from QCD background. Hence, dividing this remaining D^* count by the number of QCD events in each quadrant (Table 4.6), we find $f_{D^*,A}^{\rm QCD}$, $f_{D^*,B}^{\rm QCD}$, and $f_{D^*,C}^{\rm QCD}$.

We now calculate $f_{D^*,W}^{\rm QCD}$ using Eq 6.4. Using $f_{D^*,W}^{\rm QCD}$ and $f_{D^*}^{Z}$, and the number of events in region **W**

Tagged Rate ($\times 10^{-5}$)	... in $W \to \mu\nu$ events ($\pm$ stat)	... in $W \to e\nu$ events ($\pm$ stat)
$f^{\rm QCD}_{D^*,A}$	17 ± 5	23 ± 2
$f^{\rm QCD}_{D^*,B}$	152 ± 11	32 ± 7
$f^{\rm QCD}_{D^*,C}$	135 ± 17	20 ± 17
$f^{\rm QCD}_{D^*,W}$	16 ± 5	15 ± 2
$f^{W\ \rm sig}_{D^*}$	5.1 ± 0.5	5.3 ± 0.8

Table 6.3: Rates $f^{\rm QCD}_{D^*}$ and $f^{\rm W,\ sig}_{D^*}$ for both our $W \to \mu\nu$ and $W \to e\nu$ samples. These results suggest that $f^{\rm QCD}_{D^*}$ is strongly dependent on lepton isolation Iso1 (especially in the $W \to \mu\nu$ case), but only weakly dependent on $\not\!\!E_T$ (if at all: rate pairs $f^{\rm QCD}_{D^*,A}$ and $f^{\rm QCD}_{D^*,W}$, and $f^{\rm QCD}_{D^*,B}$ and $f^{\rm QCD}_{D^*,C}$, are each compatible within uncertainty—for both W decay modes).

that are expected to come from Z and QCD events, we can determine the number of D^* in region **W** that come from non-W sources. Subtracting out these D^* from the total D^* count in region **W**, we are left with the number of D^* produced in association with real W bosons. We divide this by the number of real W in region **W** to get a next-order approximation to $f^{W\ \rm sig}_{D^*}$.

With an initial guess of $f^{W\ \rm sig}_{D^*} = 0$, we iterate through this process until the value of $f^{W\ \rm sig}_{D^*}$ does not change by more than 0.1% of its value over two consecutive iterations. (As we will see, this is far lower than the uncertainty in our Δm signal fits—as such, it is more than sufficient that we iterate only to this level of accuracy).

Doing this for both our $W \to \mu\nu$ and $W \to e\nu$ samples, we find values for $f^{\rm QCD}_{D^*,A}$, $f^{\rm QCD}_{D^*,B}$, $f^{\rm QCD}_{D^*,C}$, $f^{\rm QCD}_{D^*,W}$ and $f^{W\ \rm sig}_{D^*}$ as recorded in Table 6.3.

6.3.4 Measuring rates $f^{W\ \mathbf{sig}}_{D^*}$ and $f^{W\ \mathbf{bkg}}_{D^*}$ (reduced uncertainty approach)

We see in Table 6.3 that $f^{\rm QCD}_{D^*}$ is strongly dependent on lepton isolation (especially in the $W \to \mu\nu$ case), but only weakly dependent on $\not\!\!E_T$. Specifically, rates $f^{\rm QCD}_{D^*,A}$ and $f^{\rm QCD}_{D^*,W}$ are compatible within uncertainty, as are rates $f^{\rm QCD}_{D^*,B}$ and $f^{\rm QCD}_{D^*,C}$. Unfortunately, for the case of $W \to e\nu$, a small population of D^* in region **C** leads to large uncertainties in $f^{\rm QCD}_{D^*,C}$, which in turn leads to large uncertainties in $f^{\rm QCD}_{D^*,W}$ as found by Eq 6.4. As such—and because $\not\!\!E_T$ dependence in $f^{\rm QCD}_{D^*}$ is shown to be weak—we choose instead to approximate rate $f^{\rm QCD}_{D^*,W}$ for both W samples as

$$f^{\rm QCD}_{D^*,W} \simeq f^{\rm QCD}_{D^*,A} \tag{6.5}$$

With this approximation made, we otherwise solve for $f^{\rm W}_{D^*}$ in the same way that we did in section 6.3.3. Doing so, we find (new) values for $f^{\rm W}_{D^*}$ as summarized in Table 6.4. We will use these values for the remainder of this analysis—any uncertainty in the procedures described here will be taken into account by our evaluation of systematic uncertainties in 6.4.

Signal Type	$f^{QCD}_{D^*,W}$ $(\times 10^{-5})$ ($\pm$ stat)	$f^{W}_{D^*}$ $(\times 10^{-5})$ ($\pm$ stat)
$W(\rightarrow \mu\nu) + D^*$	17 ± 5	5.1 ± 0.5
$W(\rightarrow e\nu) + D^*$	23 ± 2	5.0 ± 0.6

Table 6.4: Final values for the rate at which W bosons are produced in association with D^* mesons, found independently for decay modes $W \rightarrow \mu\nu$ and $W \rightarrow e\nu$.

We may also work backwards, and use final rates $f^{\mathrm{W}}_{D^*}$ and $f^{\mathrm{QCD}}_{D^*}$ to count the number of tagged D^* in Region W that are attributed to each signal and background process. Results are summarized in Table 6.5.

6.4 Systematic uncertainty in $f^{W\ \mathbf{sig}}_{D^*}$

The technique used above may not be completely accurate—any of our simplifying assumptions in section 6.3could be erroneous. We would like to estimate the quantitative effects of this uncertainty. Because the whole of our $f^{W\ \mathrm{sig}}_{D^*}$-finding technique depends upon the definition of $\not\!\!E_T$/Iso1 regions **A**, **B**, **C**, and **W**, this can be done by redefining those regions, and then redoing our measurements of $f^{W\ \mathrm{sig}}_{D^*}$, for each definition. The maximum variation in final results $f^{W\ \mathrm{sig}}_{D^*}$, taken across all tests, is identified as the systematic uncertainty associated with our technique.

Specifically, we first shift the upper $\not\!\!E_T$ boundary used to define regions **A** and **B**, while keeping all other region boundaries as defined in Table 4.5. This ensures that we do not vary our definition as to what constitutes a $W \rightarrow \ell\nu$ event, while also maintaining the accuracy of the important ratio defined in Equation 6.4. We then reestablish the original $\not\!\!E_T$ boundaries, and vary instead the lower *Iso1* boundary used to define regions **B** and **C**.

Unlike in the case of $f^{Z\ \mathrm{sig}}_{D^*}$, we do *not* average the results of all tests to find a central value; this would indirectly vary our definition of W events, and the original region **W** boundaries are widely accepted as the standard for W selection. Instead, we use the $f^{W\ \mathrm{sig}}_{D^*}$ rates reported in Table 6.4 as our central value, and these tests are used to find systematic uncertainty only. We tabulate the boundary definitions employed, and the $f^{W\ \mathrm{sig}}_{D^*}$ values found, by each test (Table 6.6). The results are also plotted, in Figure 6.2. Our final values for $f^{W\ \mathrm{sig}}_{D^*}$, including full statistical and systematic uncertainties, are summarized in Table 6.7.

$W \to \mu\nu$ selection $N_{\text{jets}} \geq 0$, $p_T \geq 0$	Fraction of all events in region **W** ($\pm$ stat)	Number of events in region **W** ($\pm$ stat)	D^* Production Rate $f_{D^*}^{\text{W}}$ ($\times 10^{-5}$) ($\pm$ stat)	$W + D^*$ Count ($\pm$ stat)
Total	1.0	$5,348,975$	5.5 ± 0.5	293 ± 26
$W \to \mu\nu$	0.9152 ± 0.0002	$4,895,240 \pm 880$	5.1 ± 0.5	250 ± 24
$W \to \tau\nu$	0.0305 ± 0.0001	$163,070 \pm 670$	5.1 ± 0.5	8.3 ± 0.8
$W \to \mu\nu$ & $\tau\nu$				258 ± 25
$Z \to \mu\mu$	0.03853 ± 0.00005	$206,110 \pm 270$	1.0 ± 0.3	2.1 ± 0.6
Jets Faking $W \to \mu\nu$	0.0158 ± 0.0001	$84,550 \pm 540$	17 ± 5	14 ± 4

$W \to e\nu$ selection $N_{\text{jets}} \geq 0$, $p_T \geq 0$	Fraction of all events in region **W** ($\pm$ stat)	Number of events in region **W** ($\pm$ stat)	D^* Production Rate $f_{D^*}^{\text{W}}$ ($\times 10^{-5}$) ($\pm$ stat)	$W + D^*$ Count ($\pm$ stat)
Total	1.0	$5,081,938$	5.7 ± 0.5	291 ± 27
$W \to e\nu$	0.9362 ± 0.0004	$4,757,570 \pm 2100$	5.0 ± 0.6	238 ± 29
$W \to \tau\nu$	0.01913 ± 0.0001	$97,193 \pm 560$	5.0 ± 0.6	4.9 ± 0.6
$W \to e\nu$ & $\tau\nu$				243 ± 29
$Z \to ee$	0.004863 ± 0.00001	$24,710 \pm 90$	5.3 ± 3.2	1.3 ± 0.8
Jets Faking $W \to e\nu$	0.03984 ± 0.0004	$202,500 \pm 2,100$	23 ± 2	47 ± 4

Table 6.5: The fractions of tagged W signal and background events that are also tagged as having a D^*. For easy reference, we also list the fraction of events in region **W** that are expected to come from each source, and the number of $W + D^*$ events that are attributed to each source. The number in the top right of each table is the total number of $W + D^*$ candidates in region **W** for that decay mode. Note that for $W + D^*$ signal counts, only the combined "$W \to \mu\nu$ & $\tau\nu$" and "$W \to e\nu$ & $\tau\nu$" values are calculated directly. The separated $W + D^*$ signal counts ($W \to e\nu$, $W \to \mu\nu$, and $W \to \tau\nu$) are derived from the combined results.

Test Name	Regions **A** and **B** $\not\!\!E_T$(GeV) $\leq$	Regions **B** and **C** Iso1 $\geq$	$f_{D^*}^{W(\to\mu\nu)\text{ sig}}$ ($\times 10^{-5}$)	$f_{D^*}^{W(\to e\nu)\text{ sig}}$ ($\times 10^{-5}$)
Iso Test 1	10.0	0.2	5.10 ± 0.53	5.1 ± 0.6
Iso Test 2	10.0	0.3	5.10 ± 0.53	5.0 ± 0.6
Iso Test 3	10.0	0.4	5.09 ± 0.53	4.9 ± 0.6
Iso Test 4	10.0	0.5	5.07 ± 0.54	4.8 ± 0.6
Iso Test 5	10.0	0.6	5.04 ± 0.54	4.6 ± 0.6
MET Test 1	8.0	0.3	5.05 ± 0.54	4.98 ± 0.55
MET Test 2	10.0	0.3	5.10 ± 0.53	5.02 ± 0.55
MET Test 3	12.0	0.3	5.10 ± 0.53	5.02 ± 0.55
MET Test 4	14.0	0.3	5.08 ± 0.53	5.02 ± 0.55
MET Test 5	16.0	0.3	5.11 ± 0.53	5.04 ± 0.55

Table 6.6: The charm fraction rates $f_{D^*}^{W\text{ sig}}$ found by varying the $\not\!\!E_T$ / Iso1 borders used to define regions **A**, **B**, **C**, and **W**. We take the maximum variation in $f_{D^*}^{W\text{ sig}}$ across all tests to be the systematic uncertainty associated with our "region-driven" technique for finding $f_{D^*}^{W\text{ sig}}$. We do this for both the $W \to \mu\nu$ and $W \to e\nu$ samples. Note that Iso1 Test 1 and $\not\!\!E_T$ Test 2 are the same—these tests both use the standard region definitions from Table 4.5. This does not affect the systematic uncertainty, as it does not affect the mininum/maximum $f_{D^*}^{W\text{ sig}}$ values.

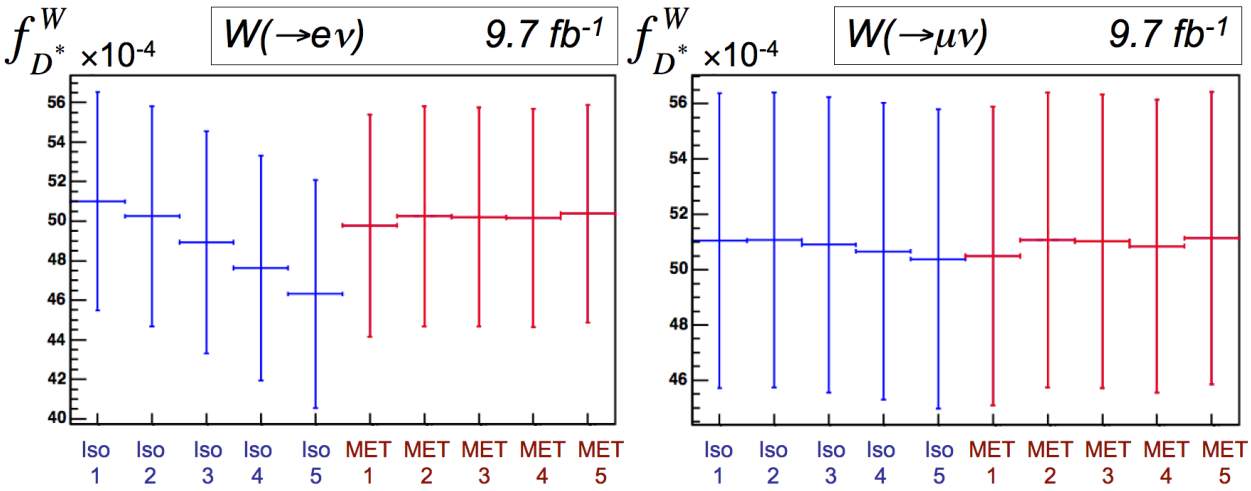


Figure 6.2: Plots of the values $f_{D^*}^{W\ \mathrm{sig}}$ found by varying the $\not\!\!E_T$ / Iso1 borders used to define regions **A**, **B**, **C**, and **W**. *Nota bene* that the error bars for each data point are statistical, and are highly-correlated across tests (the set of D^* candidates, which we look through to count D^* signal, changes only slightly as our $\not\!\!E_T$ / Iso1 borders are varied. Our systematic uncertainty is, in general, much smaller than this statistical uncertainty.

Signal Type	$f_{D^*}^{W\ \mathrm{sig}}$ $(\times 10^{-5})$ ($\pm$ stat $\pm$ syst)
$W(\to \mu\nu) + D^*$	5.1 ± 0.5 (stat) ± 0.1 (syst)
$W(\to e\nu) + D^*$	5.0 ± 0.6 (stat) ± 0.5 (syst)

Table 6.7: Final measurements of $f_{D^*}^{W\ \mathrm{sig}}$ for the $W \to e\nu$ and $W \to \mu\nu$ samples, including full statistical and systematic uncertainties.

Chapter 7

Kinematic Properties of $W/Z + D^*$ Events

Before getting started, let us simplify our terminology somewhat. Having completed our investigation of background, it is no longer necessary to specifically label $f_{D^*}^{W\ \mathrm{sig}}$ with the sig token. Henceforth, we refer to rates $f_{D^*}^{W}$ *only* in the context of D^* produced in association with real, "signal" W events. From this point forward, then, we will use $f_{D^*}^{W\ \mathrm{sig}}$ and $f_{D^*}^{W}$ interchangeably.

7.1 A preface

Thus far, we have found "inclusive" rates of D^* production, involving all identifiable $W + D^*$ events in our sample. We would also like to measure *specific* properties of our $W + D^*$ events. The particular strength of our fully-reconstructed D^* analysis, is our ability to detect low momentum D^*. This allows two novel measurements: a measurement of the rate at which D^* are produced in association with W bosons in *zero jet* events (jet objects are generally defined to have $E_T > 15, 20$ GeV); and a measurement of how $W + c$ production acts at low $p_T(c)$ (defined here as $p_T(c) < 15$ GeV). This is the first time that either of these measurements have been made at the Tevatron. We would also like to measure other properties of these events—how often the D^* is found within a jet, for example, and how often the D^* is produced back-to-back with the W boson. These are the measurements that we make in this chapter.

In general, we use exactly the same techniques developed in Chapters 4-6; the only difference are additional restrictions placed on which W and D^* are allowed to pass our cuts. We essentially run the analysis for several 'reduced' datasets (*e.g.*, D^* with p_T in ranges a, b or c, exclusively), and then merge these reduced analyses together to measure $f_{D^*}^{W\ \mathrm{sig}}$ as a function of some variable. In this analysis, we call this process 'splitting' our signal. The exact splits will be elaborated upon for each case.

7.2 Splitting by $p_T(D^*)$

We first look at the (normalized) $p_T(D^*)$ distributions of all D^* identified in our *Wcc*, *Wbb*, *Wc* and *Zcc* Monte Carlo samples, as defined in Table 5.3. These distributions are plotted in Figure 7.1. We do not split

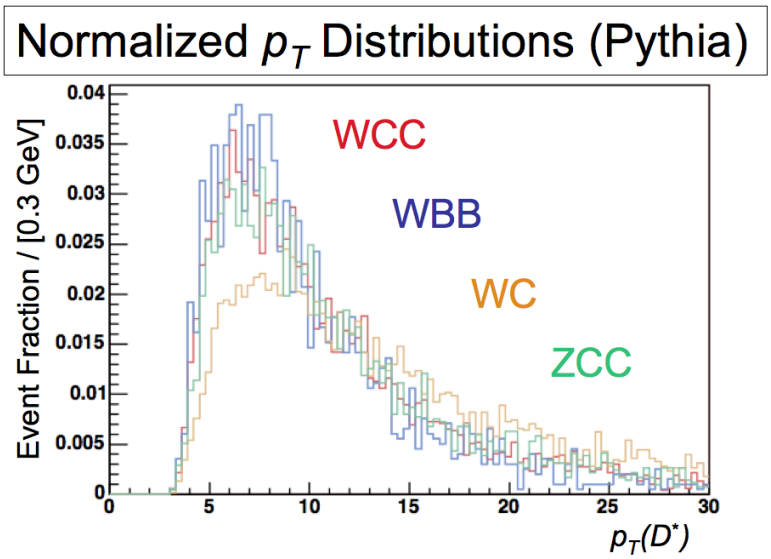


Figure 7.1: Normalized $p_T(D^*)$ distribution of all D^* tagged by our algorithm in fully-simulated $W/Z + D^*$ events generated by Pythia (and simulated with the cdfSim package). We include results from each of *Wcc*, *Wbb*, *Wc*, and *Zcc* Monte Carlo samples.

our *Zcc* sample by $p_T(D^*)$, but we add the *Zcc* distribution to Figure 7.1 for completeness and comparison.

As expected, we see that the Wc distribution is weighted towards higher $p_T(D^*)$ than the *Wcc* distribution is, and the *Wbb* distribution is weighted towards lower $p_T(D^*)$. In terms of $p_T(D^*)$, we would like to split our $W + D^*$ signal in such a way that the expected number of D^* in each region is approximately constant. This will assure a similar uncertainty in each region, rather than forcing huge uncertainties into 'tail' regions with low statistics. In chapter 10, we will show that our $W + D^*$ signal consists mostly of *Wcc* events, with some smaller fractions of *Wbb* and *Wc*. from *Wcc*. Therefore, we split our $W + D^*$ signal into eight regions, chosen such that number of predicted D^* from *Wcc* Monte Carlo is approximately the same in each region. These regions are defined as in Table 7.1. The highest-$p_T(D^*)$ region is capped at 30.0 GeV, for two reasons: first, this will allow us to measure the differential quantity $d\sigma(W + D^*)/\sigma(W) \cdot dp_T(D^*)$ without an infinite denominator in the last bin; second, beyond 30.0 GeV the number of D^* counted in our final $p_T(D^*)$ region does not change.

For each region we run our full $W + D^*$ analysis. Again, we compare our results to measurements of the same quantity in the inclusive Monte Carlo samples. All results are reported in Table 7.1. Once more, we find good agreement between simulation and data.

Region Number	Region Definition	$f^{D^*}_{W\to e\nu} \pm$ (stat) $\pm$ (syst) ($\times 10^{-6}$) $DATA$	$f^{D^*}_{W\to e\nu} \pm$ (stat) ($\times 10^{-6}$) MC
1	$[0.0, 3.0]$	$0.2 \pm 0.2 \pm 0.0$	0.0 ± 0.3
2	$[3.00, 6.50]$	$9.7 \pm 2.8 \pm 0.6$	19.0 ± 5.9
3	$[6.50, 7.75]$	$6.8 \pm 2.0 \pm 0.4$	5.5 ± 3.5
4	$[7.75, 9.00]$	$4.2 \pm 1.9 \pm 0.5$	6.6 ± 3.4
5	$[9.00, 11.25]$	$5.2 \pm 1.9 \pm 0.8$	2.7 ± 2.9
6	$[11.25, 13.5]$	$4.7 \pm 1.7 \pm 0.5$	4.3 ± 2.7
7	$[13.5, 22.0]$	$11.0 \pm 2.4 \pm 1.6$	11.1 ± 3.8
8	$[22.0, 30.0]$	$8.2 \pm 1.7 \pm 0.6$	3.8 ± 2.1

Region Number	Region Definition	$f^{D^*}_{W\to \mu\nu} \pm$ (stat) $\pm$ (syst) ($\times 10^{-6}$) $DATA$	$f^{D^*}_{W\to \mu\nu} \pm$ (stat) ($\times 10^{-6}$) MC
1	$[0.0, 3.0]$	$0.0 \pm 1.9 \pm 0.0$	0.0 ± 0.0
2	$[3.00, 6.50]$	$13.0 \pm 2.8 \pm 0.7$	11.1 ± 4.8
3	$[6.50, 7.75]$	$6.2 \pm 1.9 \pm 1.0$	6.4 ± 3.2
4	$[7.75, 9.00]$	$6.6 \pm 1.8 \pm 0.2$	6.4 ± 2.8
5	$[9.00, 11.25]$	$7.5 \pm 1.9 \pm 0.3$	7.1 ± 3.2
6	$[11.25, 13.5]$	$4.0 \pm 1.5 \pm 0.1$	7.0 ± 2.4
7	$[13.5, 22.0]$	$9.4 \pm 2.0 \pm 0.4$	13.7 ± 3.8
8	$[22.0, 30.0]$	$5.0 \pm 1.4 \pm 0.2$	4.0 ± 2.2

Table 7.1: $f^W_{D^*}$ as measured in different regions of $p_T(D^*)$. We split the $p_T(D^*)$ spectrum into eight regions, defined such that approximately the same number of observed D^* are expected to fall into each region. Again, we find good agreement between the inclusive MC sample and data. We see that region 1 contains a number of D^* compatible with zero. This is what we would expect, based on the $p_T(D^*)$ spectrum of accepted D^* from our $W + D^*$ Monte Carlo samples (Figure 7.1); our acceptance rate for all $W + D^*$ signal processes apparently drops to zero near $p_T(D^*) \sim 3.5 - 4$ GeV. We investigate this acceptance rate in more detail in Chapter 8.

N_{jets}	$f_{D^*}^{W(\to e\nu)} \pm$ (stat) $\pm$ (syst) $(\times 10^{-5})$ DATA	$f_{D^*}^{W(\to e\nu)} \pm$ (stat) $(\times 10^{-5})$ MC	$f_{D^*}^{W(\to \mu\nu)} \pm$ (stat) $\pm$ (syst) $(\times 10^{-5})$ DATA	$f_{D^*}^{W(\to \mu\nu)} \pm$ (stat) $(\times 10^{-5})$ MC
≥ 0	$5.0 \pm 0.6 \pm 0.3$	5.3 ± 1.0	$5.11 \pm 0.53 \pm 0.06$	5.6 ± 0.9
≥ 1	$18 \pm 2.8 \pm 0.6$	14 ± 4	$17.7 \pm 2.3 \pm 0.4$	19 ± 4
≥ 2	$37.7 \pm 9.9 \pm 1.2$	36 ± 13	$36 \pm 7 \pm 1$	$27 \pm 10.$
≥ 3	$102 \pm 36 \pm 23$	41 ± 36	$48 \pm 18 \pm 5$	0 ± 130
$= 0$	$2.4 \pm 0.4 \pm 0.2$	2.7 ± 0.6	$2.2 \pm 0.4 \pm 0.1$	2.1 ± 0.6
$= 1$	$12 \pm 2.3 \pm 0.5$	9.4 ± 3.6	$12 \pm 2 \pm 3$	17 ± 4
$= 2$	$24 \pm 9 \pm 2$	35 ± 14	$32 \pm 7 \pm 1$	32 ± 11
$= 3$	$32 \pm 30. \pm 7$	$35 \pm 40.$	$39 \pm 19 \pm 4$	7 ± 24

Table 7.2: $f_{D^*}^W$ rates measured as a function of N_{jets}, for jet $E_T > 15$ GeV, both for events in the inclusive high-p_T lepton datasets, and simulated events in the CDF EWK group's inclusive W datasets. We find good agreement in all cases save for $N_{\text{jets}} \geq 3$; however, within uncertainty, this disagreement is not large—it may be attributed to low statistics. All results above are plotted in Figure 9.3.

7.3 Splitting by N_{jets}

We define jets according to the midpoint algorithm with $\Delta R = 0.4$, as described in section 3.2, and we require that all jets have $E_T > 15$ GeV. From there, it is a simple matter of splitting our W event selection by the number of jets counted in that event. For each subset, we perform the complete D^*-counting and $f_{D^*}^{W\ \text{sig}}$-measuring analysis as described in Chapters 4-6. We do this both for the high-p_T lepton datasets, and for the CDF EWK group's inclusive Monte Carlo W samples (for comparison). All of these results are compiled in Table 7.2.

Chapter 8

Acceptance Rates and Unfolding

8.1 What we want, and what we have measured

We next turn to the question of how to unfold our tagged charm fractions, $f_{D^*}^{W/Z}$, into a ratio of physical cross-sections, $\sigma(W/Z + D^*)/\sigma(W/Z)$. To do this, we will have to determine our acceptance times efficiency rates, $A \cdot \epsilon$; these are a convolution of the detector-level fiducial and trigger acceptance rates, and the rate at which our tagging algorithm accepts those events that remain. In short, $A \cdot \epsilon$ is the rate at which we tag *all signal events of a given type.*

Recall that $f_{D^*}^{W/Z}$ is generally defined as

$$f_{D^*}^{W/Z} = \frac{\text{Number of } D^* \text{tagged in our } W/Z \text{ selection}}{\text{Number of } W/Z \text{ events in our selection}} \tag{8.1}$$

This quantity is related to the ratio of physical cross-sections, $\sigma(W/Z + D^*)/\sigma(W/Z)$, as

$$f_{D^*}^{W/Z} = \frac{\sigma\big(W/Z + D^*\big)}{\sigma\big(W/Z\big)} \cdot \frac{(A \cdot \epsilon)_{tag}^{W/Z,\text{ events with } D^*} \cdot (A \cdot \epsilon)_{tag}^{D^*,\text{ in } W/Z \text{ events}}}{(A \cdot \epsilon)_{tag}^{W/Z\text{ , inclusive}}} \times \text{Br}_{D^* \to D^0(\to K\pi)\pi_s} \tag{8.2}$$

where we define

$$\begin{aligned}
\text{Br}_{D^* \to D^0(\to K\pi)\pi_s} &\equiv \text{Branching ratio for } D^* \text{to decay as } D^* \to D^0(\to K\pi)\pi_s \\
(A \cdot \epsilon)^{W/Z\text{ , inclusive}} &\equiv \text{Rate at which we tag all } W/Z\text{, whether or not the event has a } D^* \\
(A \cdot \epsilon)^{W/Z\text{ , events with } D^*} &\equiv \text{Rate at which we tag } W/Z\text{, when the } W/Z \text{ is produced with a } D^* \\
(A \cdot \epsilon)^{D^*,\text{ in } W/Z \text{ events}} &\equiv \text{Rate at which we tag } D^*\text{, in an event containing a tagged } W/Z
\end{aligned} \tag{8.3}$$

There are a few important points to make about these quantities. For one, we cannot assume *a priori* that W/Z bosons that are produced alongside a D^* will be tagged at the same rate as a general sample

Parent process	Events total	Events tagged	$(A \cdot \epsilon)^{W/Z,\ \text{inclusive}}$
$p\bar{p} \to W(\to \mu\nu) + X$	7,871,895	1,721,976	21.9%
$p\bar{p} \to W(\to e\nu) + X$	7,858,019	1,532,542	19.5%
$p\bar{p} \to Z(\to \mu\mu) + g$	7,894,603	484,811	6.14%
$p\bar{p} \to Z(\to ee) + g$	7,835,135	378,008	4.82%

Table 8.1: W/Z tag rates, for the inclusive W/Z samples.

of W/Z bosons. It is possible, for example, that W/Z bosons that are produced alongside a D^* have a higher average p_T than those with no accompanying D^*. As such, in Equation 8.2, we cannot cancel out $(A \cdot \epsilon)^{W\to\mu\nu\ ,\ \text{events with}\ D^*}$ and $(A \cdot \epsilon)^{W\to\mu\nu\ ,\ \text{inclusive}}$. Additionally, we cannot assume that D^* tag rates will be the same across W/Z decay modes—we will either have to find the rate at which D^* are tagged in all four samples, or develop some invariant metric for the D^* tag rate. For these reasons, we combine the second two rates in equation 8.3 to create a combined rate $(A \cdot \epsilon)^{W/Z+D^*,\ \text{inclusive}}$:

$$(A \cdot \epsilon)^{W/Z+D^*,\ \text{inclusive}} = (A \cdot \epsilon)^{W/Z\ ,\ \text{events with}\ D^*} \times (A \cdot \epsilon)^{D^*,\ \text{in}\ W/Z\ \text{events}} \tag{8.4}$$

This rate (equation 8.4) is now just the total $A \cdot \epsilon$ for finding the W/Z *and* D^* in a sample of $W/Z + D^*$ events, using our selection criteria.

To determine physical cross-section ratios, then, we will need to determine both $(A \cdot \epsilon)^{W/Z\ ,\ \text{inclusive}}$ and $(A \cdot \epsilon)^{W/Z+D^*,\ \text{inclusive}}$, and we must do this for each of our selected W/Z samples. Determining these quantities, and using them to go from $f_{D^*}^{W/Z}$ to $\sigma\big(W/Z + D^*\big)/\sigma\big(W/Z\big)$, is the purpose of this chapter.

8.2 Finding inclusive rates $(A \cdot \epsilon)^{W/Z}$ and $(A \cdot \epsilon)^{W/Z+D^*}$

We will first determine the rates at which W/Z bosons are tagged, $(A \cdot \epsilon)^{W/Z\ ,\ \text{inclusive}}$. To do this, we run over the CDF EWK group's inclusive W/Z datasets with our W/Z tagging algorithm, and simply report the rate at which we accept events. These rates are summarized in Table 8.1. We see that we tag W events at approximately three times the rate at which we tag Z events; a larger W acceptance rate is to be expected, as we require only one lepton to fall into the central region $|\eta| < 1.1$ for $W \to \ell\nu$ decays; *both* leptons must fall into the central region for $Z \to \ell\ell$ decays, which is less likely. (Technically, we do allow for plug electrons in our Z candidates, but we found the number of accepted Z with at least one plug electron to be be negligible (section 4.2).)

To determine rates $(A \cdot \epsilon)^{W/Z+D^*,\ \text{inclusive}}$, we return to our Pythia-generated Monte Carlo samples from section 5.2. Recall that we constructed each sample by choosing a production process of interest, and then skimming off events which contained a $D^* \to D^0\pi_s \to (K\pi)\pi_s$ decay. In order to save time, we also made a

Parent process	:	count	Subprocess:	:	count	Subprocess Rate
$q + \bar{q}' \to W(\to \mu\nu) + g$	:	2,908,022,069	$g(\to c\bar{c})$	:	39,280,12	1.351 %
$q + \bar{q}' \to W(\to \mu\nu) + g$	:	4,004,507,486	$g(\to b\bar{b})$	:	12,255,600	0.306 %
$q + g \to W(\to \mu\nu) + \bar{q}'$		303,103,282	$\bar{q}' = c/\bar{c}$	:	28,196,185	9.303%
$q + \bar{q} \to Z(\to \mu\mu) + g$	:	5,507,241,075	$g(\to c\bar{c})$	:	51,136,473	0.929%

Table 8.2: A summary of the rates at which each first-order W/Z boson production process gives rise to our D^*-generating subprocesses. Subprocesses are defined relative to the parent process named on each line. These values are derived from Pythia simulation. Unertainty on these numbers is binomial, but negligible compared to all other uncertainties considered in this section.

check as to whether an event was 'fiducial' before simulating its interactions with the detector (full detector event simulation is extremely time intensive). In order to prevent any kind of interference with the production process, we did *not* place explicit restrictions on the production process' kinematics; we simply filtered off only those events that we would have a non-zero chance of tagging. Explicitly, the charged leptons from W/Z decays, and the K, π, and π_s from D^* decay, were all required to have $|\eta| < 1.2$. While these requirements allow for some non-fiducial events to get through (recall that our final cuts require $|\eta| < 1.1$), we wanted to make certain not to miss any events that we might be able to tag.

We record a number of statistics about these samples, in order to better identify the means by which D^* come to be associated with our W or Z events. Included among these statistics is the rate at which each 'parent process' gives rise to the subprocess that we are interested in—*e.g.*, what percentage of all $s(d) + g \to W + q$ events satisfy $q = c$, and what percentage of all $q + \bar{q}' \to W + g$ events are followed by $g \to c\bar{c}$ or $g \to b\bar{b}$. We present these statistics in Table 8.2. We also run over the simulated signal events with our $W/Z + D^*$ tagging algorithms, and count how many events pass each of our $W/Z + D^*$ cuts. We use these values to find combined rates These rates, and the counts used to find them, are recorded in Table 6.4. We see that in all cases, we are better able to identify $W/Z + D^*$ events if the W/Z decays to a muon(s)—this simply reflects the fact that our electron cuts are tighter than our muon cuts, as explained in sections 3.3-3.4.

8.3 Constructing p_T-dependent rates $(A \cdot \epsilon)^{W/Z+D^*}$

The tag rates found in the previous section are an important step, but overlook a rather important piece of the simulation/measurement relationship. Recall that while Pythia simulations are quite good at predicting the behavior of specific signal processes, they are not always great at determining the *rates* of production. This can have a profound impact on our tagging rate $(A \cdot \epsilon)^{W/Z}_{tag}$, if *e.g.* this tagging rate is strongly dependent upon momentum: if Pythia claims that half of all D^* in W/Z events have $p_T(D^*) < 3.0$ GeV, but reality dictates that 90% of all D^* in W/Z events have $p_T(D^*) < 3.0$ GeV, then our tag rates for inclusive D^* production, as determined by simulation, could be very far off!

Process (Event Count)	Number of D^* w/ decay $D^* \to D^0(\to K\pi)\pi_s$	Number of D^* with fiducial decay products	# events tagged for $W/Z + D^*$ (w/ NN)	Net tag rate $(A \cdot \epsilon)^{W/Z+D^*}$ (w/ NN)
$p\bar{p} \to W(\to e\nu) + g(\to c\bar{c})$ (36,790,414)	552,097	104,938	5,308	0.961%
$p\bar{p} \to W(\to \mu\nu) + g(\to c\bar{c})$ (39, 280, 812)	595,827	113,250	6,712	1.127%
$p\bar{p} \to W(\to e\nu) + c$ (28,196,185)	350,025	57,415	5,288	1.511%
$p\bar{p} \to W(\to \mu\nu) + c$ (38,510,409)	352,689	57,852	5,849	1.658%
$p\bar{p} \to Z(\to ee) + g(\to c\bar{c})$ (51,136,473)	426,723	43,460	2,509	0.5880%
$p\bar{p} \to Z(\to \mu\mu) + g(\to c\bar{c})$ (50,443,745)	414,882	42,254	3,230	0.779%

Table 8.3: A summary of the simulated Pythia Monte Carlo samples that we used to analyze the properties of $W/Z+D^*$ production. The ultimate goal of these samples, as used in this section, is to determine the rates at which we identify $W/Z + D^*$ events. The "w/ NN" token means that tag rates are for event acceptances after applying our neural network. **Important**: the final column above shows 'net' tag rate, which is the rate at which we tag all $W/Z + D^*$ events. This is *not* the same as the rate at which we tag events with $p_T(D^*) > 3$ GeV (see Table 8.5), which turns out to be a much more useful quantity. These values should be looked at only as a stepping stone towards our final results.

We do not have to worry about this for our W/Z boson acceptance; W/Z boson production simulations have been very well-tuned over the years, and are generally accepted to be quite accurate. And in general, Pythia is considered to be quite good even for specific QCD processes, so long as p_T is large. It is in the low-p_T regime that we do not have sufficient evidence for simulations being accurate (which is part of the reason for doing this analysis in the first place!).

We address this uncertainty by determining $A \cdot \epsilon$ as a function of $p_T(D^*)$. By doing this, we can ignore model-dependent assumptions about the true $p_T(D^*)$ distribution of inclusive $W/Z+D^*$ events. Specifically, we break the $p_T(D^*)$ axis into eight regions as defined in Table 7.1, and determine the rate at which we tag $W/Z + D^*$ events which fall into each region. These rates are recorded in Figure 8.1 and Table 8.4). As claimed earlier, our $A \cdot \epsilon$ values drop to zero below $p_T = 3$ GeV; this means that our analysis is completely insensitive to D^* content in that regime. With that in mind, we can only meaningfully unfold our results back to $p_T(D^*) \geq 3$ GeV; going back any further would be model-dependent extrapolation. This fits with what we observed in Table 7.1, in which the lowest bin (for $p_T \in [0, 3.0]$) is found to be empty both in data and in the inclusive Monte Carlo W samples. We also integrate to find $A \cdot \epsilon$ for all $W/Z + D^*$ events with $p_T(D^*) \geq 3$ GeV. These values are listed in Table 8.5.

$p_T(D^*)$ Range	$(A \cdot \epsilon)^{W(\to e\nu)+D^*}$	$(A \cdot \epsilon)^{W(\to \mu\nu)+D^*}$
$[0.0, 3.0]$	$(1.67 \pm 0.6) \times 10^{-5}$	$(6.7 \pm 3.9) \times 10^{-6}$
$[3.0, 6.50]$	0.0074 ± 0.0002	0.0084 ± 0.0002
$[6.50, 7.75]$	0.0258 ± 0.0008	0.0293 ± 0.0008
$[7.75, 9.00]$	0.034 ± 0.001	0.037 ± 0.001
$[9.00, 11.25]$	0.037 ± 0.001	0.044 ± 0.001
$[11.25, 13.5]$	0.041 ± 0.001	0.051 ± 0.001
$[13.5, 21.75]$	0.049 ± 0.002	0.056 ± 0.001
$[21.75, 30.0]$	0.056 ± 0.002	0.061 ± 0.002
$[3.0, \infty)$	0.021341 ± 0.000002	0.024171 ± 0.000002

Table 8.4: Values of $(A \cdot \epsilon)^{W+D^*}$ as a function of $p_T(D^*)$ (plotted in Figure 8.1). Uncertainties in this table are statistical and due to finite sample size. Other uncertainties (PDF) are introduced in the next chapter. Regions of $p_T(D^*)$ are chosen to match the splits described in section 7.2. We also include 'inclusive' $(A \cdot \epsilon)^{W+D^*}$ values for all events with $p_T(D^*) > 3$ GeV. We will use the full range of values shown in this table to unfold our $W + D^*$ samples as a function of p_T in chapter 9. It is important to note that we do not have enough statistics to unfold our $Z + D^*$ samples in the same way. We use only the inclusive ($p_T(D^*) > 3$ GeV) values of $(A \cdot \epsilon)^{Z+D^*}$ to unfold our $Z + D^*$ results in data.

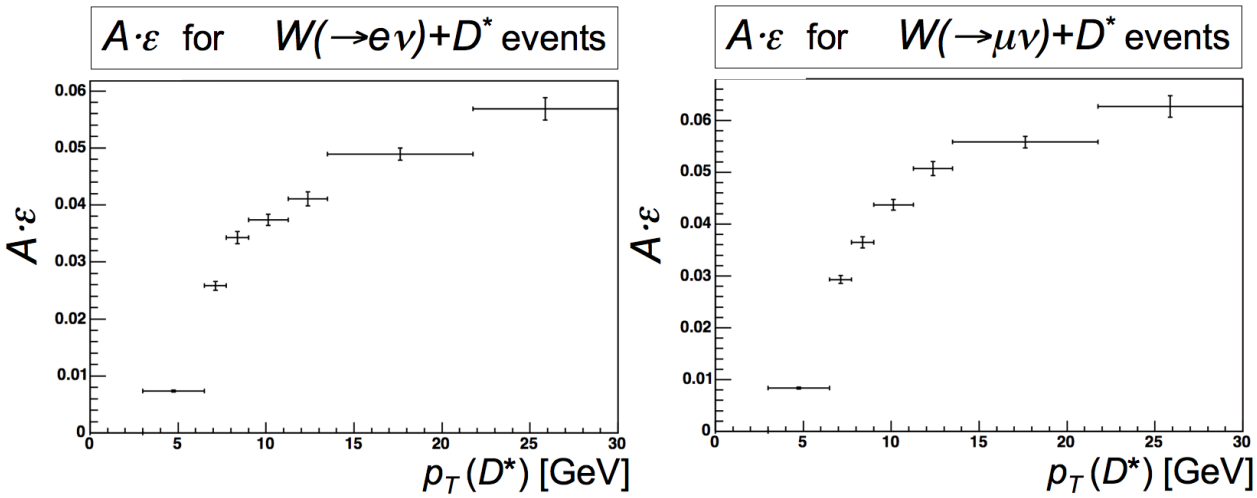


Figure 8.1: Plots of $(A \cdot \epsilon)^{W+D^*}$ as a function of $p_T(D^*)$ for both W selections considered by this analysis. Explicit values are recorded in Table 8.4. Bins are chosen to match the splits described in section 7.2. We see that the shapes of the $A \cdot \epsilon$ 'curves' are roughly similar, while the amplitudes vary. This suggests that these plots represent a single underlying $(A \cdot \epsilon)^{D*}$ distribution, independent of W decay mode, that is multiplied by some slightly-$D^*(p_T)$ dependent W tag rate. (Of course, this $p_T(D^*)$ dependence is only indirect, in the form of higher $p_T(D^*)$ being more commonly associated with higher p_T W bosons).

Important Values		
$\mathrm{Br}_{D^* \to D^0 \pi} = .677 \pm .005$	$\mathrm{Br}_{D^0 \to K\pi} = .0388 \pm .0005$	$\mathrm{Br}_{D^* \to D^0 (K\pi)\pi} = 0.0263 \pm 0.0004$

Process (with D^* understood to decay as $D^* \to D^0(\to K\pi)\pi$)	Inclusive W/Z tag rate	Inclusive $W/Z + D^*$ tag rate w/ NN ($p_T(D^*) > 3$ GeV)	Conversion factor
$p\bar{p} \to W(\to e\nu) + D^*$	0.195	0.021341	347 ± 5
$p\bar{p} \to W(\to \mu\nu) + D^*$	0.219	0.024171	344 ± 5
$p\bar{p} \to Z(\to ee) + D^*$	0.0482	0.009388	195 ± 3
$p\bar{p} \to Z(\to \mu\mu) + D^*$	0.0614	0.012385	189 ± 3

Table 8.5: Conversion factors, to go from a raw "tagged" fraction $f_{D^*}^{W/Z}$ to a physical fraction of cross-sections $\sigma(W/Z+c)/\sigma(W/Z)$. Uncertainly in the final column is dominated by uncertainty in $\mathrm{Br}_{D^* \to D^0 (K\pi)\pi}$. Statistical uncertainties are overwhelmed by uncertainty in the branching ratio, and are not shown. Other uncertainties (*e.g.*, PDF uncertainties) are considered and added to our final results in the next chapter.

8.4 From $A \cdot \epsilon$ to unfolding factors

Recall that our ultimate aim is to 'unfold' rates $f_{D^*}^{W/Z}$ into the ratio of physical cross-sections $\sigma(W+D^*)\sigma(W)$. We do this by creating an unfolding factor that performs this conversion. Explicitly, this factor is the inverse of a quantity found in Eq 8.2, and can be written as

$$\text{Unfolding Factor } (W/Z + D^*) = \frac{(A \cdot \epsilon)_{tag}^{W/Z \text{ , inclusive}}}{(A \cdot \epsilon)_{tag}^{W/Z+D^* \text{ , inclusive}} \times \mathrm{Br}_{D^* \to D^0 (\to K\pi)\pi_s}} \tag{8.5}$$

such that

$$\frac{\sigma\big(W/Z + D^*\big)}{\sigma\big(W/Z\big)} = f_{D^*}^{W/Z} \cdot \text{Unfolding Factor } (W/Z + D^*) \tag{8.6}$$

We evaluate these unfolding factors, using our values for $(A \cdot \epsilon)_{tag}^{W/Z \text{ , inclusive}}$ and $(A \cdot \epsilon)_{tag}^{W/Z+D^* \text{ , inclusive}}$ as reported in tables 8.1 and 8.3, and using the $D^*(2010) \to D^0(\to K\pi)\pi_s$ branching ratio $\mathrm{Br}_{D^* \to D^0 (\to K\pi)\pi_s} = 0.0263 \pm 0.0004$ as reported in the 2013 Particle Data Group (PDG) book [42]. In practice, these values are derived in real time by our program—however, to get a feel for the magnitude of these conversion factors, we report results for the inclusive case ($p_T(D^*) > 3$ GeV) in Table 8.5. We employ these unfolding factors in Chapter 9 to get our final results.

Chapter 9

Final Results, Full Systematics, and Comparisons to Theory

By applying the converison factors described in Chapter 8 to the raw fractions found in Chapter 6, we may finally report the ratio of physical cross-sections $\sigma(W/Z + D^*)/\sigma(W/Z)$; both for the inclusive sample $p_T(D^*) > 3$ GeV, and as a function of different variables. We will begin this chapter, however, by making a summary of all systematic uncertainties.

9.1 Summary of systematic uncertainties

The primary source of systematic uncertainty in Tables 9.3 and 9.4, is our choice of $\not\!\!E_T$ and Iso1 boundaries—what we will refer to here as our "method" uncertainty. This uncertainty was determined in section 6.4. There is, however, another source of systematic uncertainty that we must take into account—PDF uncertainty.

In determining our $A \cdot \epsilon$ values in section 8, we relied upon Monte Carlo predictions of the production processes. As explained earlier, while these simulations are usually quite good at predicting the kinematics of any given decay, the amplitude of each type of event is not as well modeled. One of the reasons for this is uncertainty in the parton distribution function (PDF) of the proton—the fraction of the proton's energy and momentum that is carried by its constituent partons. There are three major types of PDF uncertainty.

The first is the so-called 'eigenvector uncertainty'. In general, parton distribution functions are dependent upon several parameters which must be fit to data. Uncertainty in any one of these parameters leads to an uncertainty in the PDF. The CTEQ PDF sets which we use have twenty such parameters, such that any one PDF can be described in a parameter space which is spanned by 20 eigenvectors. The eigenvector uncertainty is determined by varying the value of each input parameter to its $\pm 90\%$ confidence level (CL) values, refitting the PDF, and then rerunning the PDF-dependent parts of the analysis [43].

In practice, rerunning an analysis forty one times (once for the central PDF, and twice for each eigenvector uncertainty-modified PDF) is time consuming and unnecessary. Instead, an analysis is run once (at the central PDF value), and the values measured are then reweighted to reflect the probability of each event occurring, for each modified PDF. These probabilities depend upon the fraction of energy and momentum carried by the two initial hard scatter particles (x_1 and x_2), and the hard-scatter momentum transfer Q^2.

Specifically, for each event, we define the relative weight of that event for PDF number i, $H_{PDF\ i}$, as,

$$H_{\mathrm{PDF}\ i} = \frac{\left(x_1 \cdot f(x_1)_{\mathrm{ptcl}\ 1} \cdot x_2 \cdot f(x_2)_{\mathrm{ptcl}\ 2}\right)_{\mathrm{PDF}\ i}}{\left(x_1 \cdot f(x_1)_{\mathrm{ptcl}\ 1} \cdot x_2 \cdot f(x_2)_{\mathrm{ptcl}\ 2}\right)_{\mathrm{PDF}\ 0}}$$

where PDF 0 is the 'central', unmodified PDF. Then, we define a modified value of any count x for PDF set i, x_i, as

$$x_i = x_0 \cdot H_{\mathrm{PDF}\ i}$$

Most commonly, x is a count of the number of events of some type which have passed your cuts; this is usually paired with some integrated count y of all events considered, in order to get an acceptance value $A \cdot \epsilon = x/y$ that has been reweighted to reflect the modified PDF distribution. Total eigenvector uncertainty (high and low) in $A \cdot \epsilon$ for a PDF with $2N$ eigenvector-driven variations is then calculated as

$$\begin{aligned}
\Delta(A \cdot \epsilon)^{+}_{max} &= \sqrt{\sum_{i=1}^{2N} \Big(max((A \cdot \epsilon)^{+}_{i} - (A \cdot \epsilon)_0, (A \cdot \epsilon)^{-}_{i} - (A \cdot \epsilon)_0, 0)\Big)^2} \\
\Delta(A \cdot \epsilon)^{-}_{max} &= \sqrt{\sum_{i=1}^{2N} \Big(max((A \cdot \epsilon)_0 - (A \cdot \epsilon)^{+}_{i}, (A \cdot \epsilon)_0 - (A \cdot \epsilon)^{-}_{i}, 0)\Big)^2}
\end{aligned}$$

There are two other types of PDF uncertainty to consider: the α_s uncertainty, $\Delta(A \cdot \epsilon)_{\alpha_s}$ (the amount by which $(A \cdot \epsilon)$ varies as strong coupling constant α_s is changed); and the model uncertainty, $\Delta(A \cdot \epsilon)_{\mathrm{model}}$ (the amount by which $(A \cdot \epsilon)$ varies if we move from one PDF set, say CTEQ5, to another, say MSTW2008). If $\Delta(A \cdot \epsilon)_{\mathrm{model}}$ is smaller than the eigenvector uncertainty described above, then it is ignored; otherwise, it is added in quadrature. $\Delta(A \cdot \epsilon)_{\alpha_s}$ is then added in quadrature to the resulting quantities, to give us our final PDF uncertainty [43].

Our values for $A \cdot \epsilon$ depend upon the proton PDF insofar as our ratio of acceptance rates may depend upon the PDF. We apply the procedure described above, and we find (both for the inclusive $p_T(D^*) > 3$ GeV sample, and for $(A \cdot \epsilon)$ as a function of $p_T(D^*)$), that our PDF uncertainty is quite small (Table 9.1). We also find the PDF uncertainty in the theoretical predictions for $\sigma(W + D^*)/\sigma(W)$, following the same procedure. These values are shown in Table 9.2.

Inclusive sample $p_T(D^*) \in [3.0, \infty)$	$\frac{\sigma(V+D^*)}{\sigma(V)}$ $(\times 10^{-2})$	% **stat unc.**	Bkgs. % syst unc.	Signal modeling % syst unc.	PDF % syst unc.	**Total** % **syst unc.**
$W(\to e\nu) + D^*$	1.74	**12**	10	$\ll 1$	2	**10**
$W(\to \mu\nu) + D^*$	1.75	**10**	2	$\ll 1$	2	**3**
$Z(\to ee) + D^*$	1.0	**57**	19	$\ll 1$	1	**19**
$Z(\to \mu\mu) + D^*$	1.8	**30**	11	$\ll 1$	1	**11**

$W(\to e\nu) + D^*$ $[p_T(D^*)$ range]	$\frac{d\sigma(V+D^*)}{\sigma(V) \cdot dp_T(D^*)}$ $(\times 10^{-3})$	% **stat unc.**	Bkgs. % syst unc.	Signal modeling % syst unc.	PDF % syst unc.	**Total** % **syst unc.**
$[3.0, 6.50]$	2.77	**29**	6	3	2	**7**
$[6.50, 7.75]$	1.60	**29**	6	3	2	**7**
$[7.75, 9.00]$	0.73	**45**	12	3	2	**13**
$[9.00, 11.25]$	0.44	**37**	15	3	3	**15**
$[11.25, 13.5]$	0.40	**36**	11	2	3	**12**
$[13.5, 21.75]$	0.21	**22**	15	4	3	**16**
$[21.75, 30.00]$	0.10	**21**	7	4	4	**9**

$W(\to \mu\nu) + D^*$ $[p_T(D^*)$ range]	$\frac{d\sigma(V+D^*)}{\sigma(V) \cdot dp_T(D^*)}$ $(\times 10^{-3})$	% **stat unc.**	Bkgs. % syst unc.	Signal modeling % syst unc.	PDF % syst unc.	**Total** % **syst unc.**
$[3.0, 6.50]$	3.69	**21**	5	2	2	**6**
$[6.50, 7.75]$	1.44	**31**	16	3	2	**16**
$[7.75, 9.00]$	1.21	**27**	3	3	2	**4**
$[9.00, 11.25]$	0.64	**25**	4	2	2	**5**
$[11.25, 13.5]$	0.29	**38**	3	2	2	**4**
$[13.5, 21.75]$	0.17	**21**	4	2	2	**5**
$[21.75, 30.00]$	0.08	**28**	4	3	3	**6**

Table 9.1: A summary of statistical and systematic uncertainties in our final results. Our statistical uncertainty is dominated by uncertainty from our fits of the Δm signal plus background plots, while our systematic uncertainty is dominated by the methods used to find our tagged fractions, labeled above as "Bkgs." (see 6.2, 6.4). Uncertainty due to finite sample size in our evaluation of $A \cdot \epsilon$ is identified as "signal modeling", while PDF uncertainties in $A \cdot \epsilon$ are labeled as such. Uncertainties not mentioned here are considered negligible. Many uncertainties (*e.g.*, trigger uncertainties) are almost entirely cancelled out by taking the ratio of cross-sections; this is one of the advantages to measuring this ratio, versus a total cross-section.

Signal Sample $p_T(D^*) \in [3.0, \infty)$	Pythia 6.2.16 (CTEQ5L) $\sigma(V + D^*)/\sigma(V)$ $(\times 10^{-2})$	% uncertainty (PDF)
$W \to e\nu/\mu\nu$	1.77	4.2
$Z \to ee/\mu\mu$	1.36	4.1

$W \to e\nu/\mu\nu$ $[p_T(D^*)$ range]	Pythia 6.2.16 (CTEQ5L) $\left[1/\sigma(V)\right] \times \left[d\sigma(V + D^*)\Big/ dp_T(D^*)\right]$ $(\times 10^{-3})$	% uncertainty (PDF)
$[3.0, 6.50]$	2.44	3.0
$[6.50, 7.75]$	2.01	5.0
$[7.75, 9.00]$	0.98	6.0
$[9.00, 11.25]$	0.73	7.2
$[11.25, 13.5]$	0.47	7.8
$[13.5, 21.75]$	0.20	8.8
$[21.75, 30.0]$	0.12	9.0

Table 9.2: The ratio of cross-sectons predicted by Pythia 6.2 (CTEQ5L), alongside the % uncertainty in this prediction due to PDF / modeling uncertainties (see 9.1, above).

Production process $(p_T(D^*) > 3$ GeV)	CDF Run II Preliminary $\sigma(V + D^*)/\sigma(V)$ (%) $\pm$(stat) $\pm$ (syst)	Pythia 6.2.16 (CTEQ5L) $\sigma(V + D^*)/\sigma(V)$ (%) $\pm$ (pdf unc)
$W(\to e\nu) + D^*$	$1.74 \pm 0.21 \pm 0.17$	1.77 ± 0.07
$W(\to \mu\nu) + D^*$	$1.75 \pm 0.17 \pm 0.05$	1.77 ± 0.07
Combined results: $W(\to e\nu/\mu\nu) + D^*$	$1.75 \pm 0.13 \pm 0.09$	1.77 ± 0.07
$Z(\to ee) + D^*$	$1.0 \pm 0.6 \pm 0.2$	1.36 ± 0.05
$Z(\to \mu\mu) + D^*$	$1.8 \pm 0.5 \pm 0.2$	1.36 ± 0.05
Combined results: $Z(\to ee/\mu\mu) + D^*$	$1.5 \pm 0.4 \pm 0.2$	1.36 ± 0.05

Table 9.3: The ratio of cross-sections $\sigma(W/Z + D^*)/\sigma(W/Z)$ for inclusive sample $p_T(D^*) > 3$ GeV, and the predictions of Pythia 6.2.16 simulation using PDF set CTEQ5L. The displayed restriction on $p_T(D^*)$ is the only kinematic restriction on these unfolded results. These are the final values, with full statistical and systematic uncertainties.

9.2 The inclusive sample, $p_T(D^*) > 3$ GeV

We record our final results in Table 9.3, for easy comparisons. For each process, we simply unfold the raw fraction to a rate of physical cross-sections using the acceptance values derived in the previous section, and the uncertainties summarized in Table 9.2 above. Results from the electron and muon decay channels are then combined using a best linear uncertainty estimate (BLUE), assuming that within each sample (W and Z), systematic uncertainties are fully correlated. For both our W samples and our Z samples, we find agreement within uncertainty between our unfolded results from data, and Pythia 6.2.16 predictions made with PDF set CTEQ5L. The Pythia predictions are made at the particle level (pre- CDF environment simulation).

$p_T(D^*)$ Range	CDF Run II Preliminary $\frac{d\sigma(W_{e\nu}+D^*)}{\sigma(W_{e\nu})\cdot dp_T(D^*)}$ $(\times 10^{-3})$ DATA $\pm$(stat) $\pm$ (syst)	CDF Run II Preliminary $\frac{d\sigma(W_{\mu\nu}+D^*)}{\sigma(W_{\mu\nu})\cdot dp_T(D^*)}$ $(\times 10^{-3})$ DATA $\pm$(stat) $\pm$ (syst)	CDF Run II Preliminary Combined Results: $\frac{d\sigma(W_{e\nu/\mu\nu}+D^*)}{\sigma(W_{e\nu/\mu\nu})\cdot dp_T(D^*)}$ $(\times 10^{-3})$ DATA $\pm$(stat) $\pm$ (syst)	Pythia 6.2.16 (CTEQ5L) $\frac{d\sigma(W_{e\nu/\mu\nu}+D^*)}{\sigma(W_{e\nu/\mu\nu})\cdot dp_T(D^*)}$ $(\times 10^{-3})$ MC $\pm$(stat) $\pm$ (syst)
$[3.0, 6.50]$	$2.8 \pm 0.9 \pm 0.2$	$3.7 \pm 0.8 \pm 0.2$	$3.3 \pm 0.6 \pm 0.2$	2.44 ± 0.07
$[6.50, 7.75]$	$1.60 \pm 0.48 \pm 0.08$	$1.44 \pm 0.48 \pm 0.24$	$1.53 \pm 0.34 \pm 0.15$	2.01 ± 0.09
$[7.75, 9.00]$	$0.73 \pm 0.32 \pm 0.08$	$1.21 \pm 0.32 \pm 0.07$	$0.97 \pm 0.23 \pm 0.08$	0.99 ± 0.04
$[9.00, 11.25]$	$0.44 \pm 0.18 \pm 0.09$	$0.64 \pm 0.16 \pm 0.04$	$0.56 \pm 0.12 \pm 0.06$	0.73 ± 0.04
$[11.25, 13.5]$	$0.40 \pm 0.13 \pm 0.09$	$0.29 \pm 0.11 \pm 0.02$	$0.32 \pm 0.09 \pm 0.04$	0.47 ± 0.04
$[13.5, 22.0]$	$0.21 \pm 0.05 \pm 0.02$	$0.17 \pm 0.04 \pm 0.01$	$0.17 \pm 0.03 \pm 0.01$	0.20 ± 0.02
$[22.0, 30.0]$	$0.13 \pm 0.03 \pm 0.01$	$0.083 \pm 0.024 \pm 0.007$	$0.102 \pm 0.018 \pm 0.008$	0.12 ± 0.01

Table 9.4: The ratio of cross-sections $\sigma(W + D^*)/\sigma(W)$ found in data as a function of $p_T(D^*)$, and the predictions of Pythia 6.2.16 using PDF set CTEQ5L. The displayed restrictions on $p_T(D^*)$ are the only kinematic restriction on these unfolded results. These are the final results, with full statistical and systematic uncertainties.

9.3 As a function of $p_T(D^*)$

We also report our results as a function of $p_T(D^*)$ (Table 9.4). These measurements are made by running our analysis exactly as described in the previous chapters, except for a restriction placed upon the momentum of prospective D^* candidates. Each momentum range is treated as an independent data sample. Once more, results from the electron and muon decay channels are combined using a best linear uncertainty estimate (BLUE), assuming that systematic uncertainties are fully correlated.These results are plotted in Figure 9.1.

9.4 As a function of $N_{\mathbf{jets}}(p_T > 15$ **GeV**$)$

It is also interesting to consider other properties of our $W + D^*$ events. For example, how many jets are in these events? This particular question, however, opens up many others. First, let us assume that we have measured the number of calorimeter-level jets in each $W + D^*$ event. It is unclear what the best level to 'unfold' back to, is—do we unfold from calorimeter-level jets to particle-level jets, or to parton-level jets? If the former choice is made, we must also decide how best to cluster jets at the particle-level; if the latter choice is made, how best do we map parton energy to jet energy?

In short, there are many possible options, and this question of 'unfolding' may be a question best left to future analyses. To completely, properly perform this unfolding, requires redoing our $A \cdot \epsilon$ calculations as a function of N_{jets}, and properly calibrating each of our unfolding values based on the mapping between calorimeter and particle (or parton) level jets. At the same time, this may not yield much more information

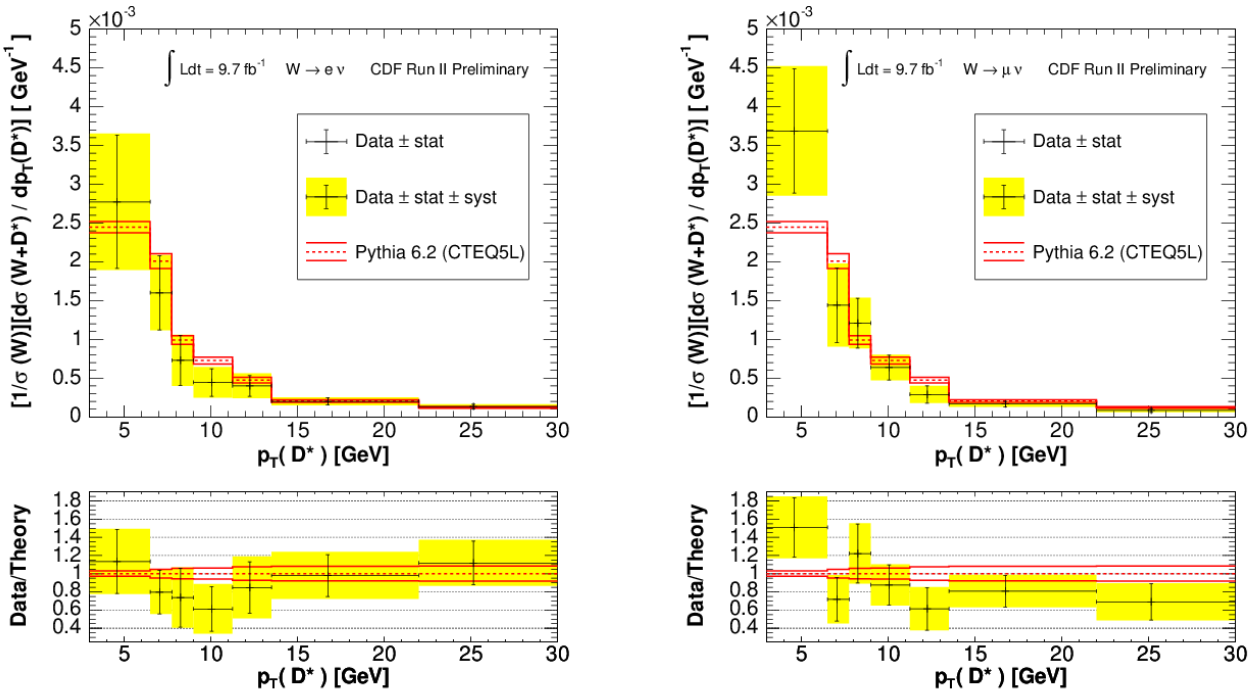


Figure 9.1: The ratio of physical cross-sections $\sigma(W + D^*)/\sigma(W)$ as a function of $p_T(D^*)$. The left plot is for the $W \to e\nu$ decay mode; the right is for the $W \to \mu\nu$ decay mode. The displayed restrictions on $p_T(D^*)$ are the only kinematic restriction on these results. In each case, error bars give the statistical uncertainty, while the sum in quadrature of the statistical and systematic errors is shown as a yellow error band. The dotted red line shows the prediction of the Pythia 6.2 obtained using the CTEQ5L PDF set, with solid red lines showing PDF uncertainty in this prediction. The ratio of the simulated distribution to data is shown in the lower panels.

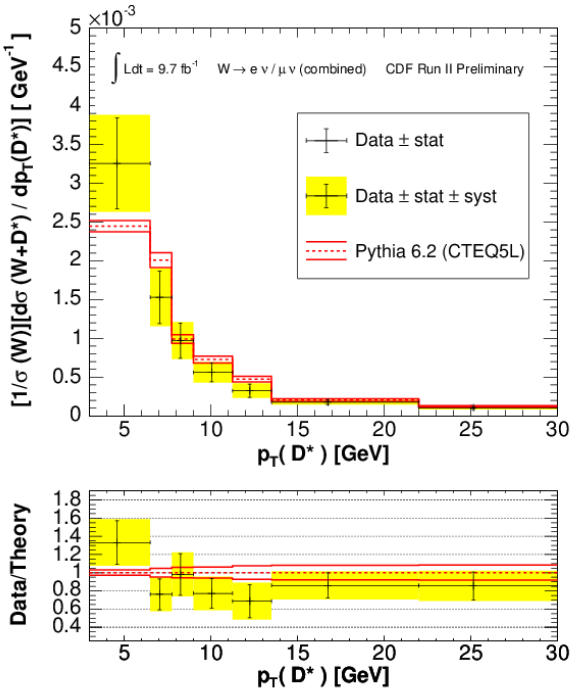


Figure 9.2: The ratio of physical cross-sections $\sigma(W + D^*)/\sigma(W)$ as a function of $p_T(D^*)$, for combined $W \to e\nu$ and $W \to \mu\nu$ results. Combination is made using a best linear uncertainty estimate with systematic uncertainties assumed to be fully correlated. The displayed restrictions on $p_T(D^*)$ are the only kinematic restriction on these results. In each case, error bars give the statistical uncertainty, while the sum in quadrature of the statistical and systematic errors is shown as a yellow error band. The dotted red line shows the prediction of the Pythia 6.2 obtained using the CTEQ5L PDF set, with solid red lines showing PDF uncertainty in this prediction. The ratio of the simulated distribution to data is shown in the lower panels.

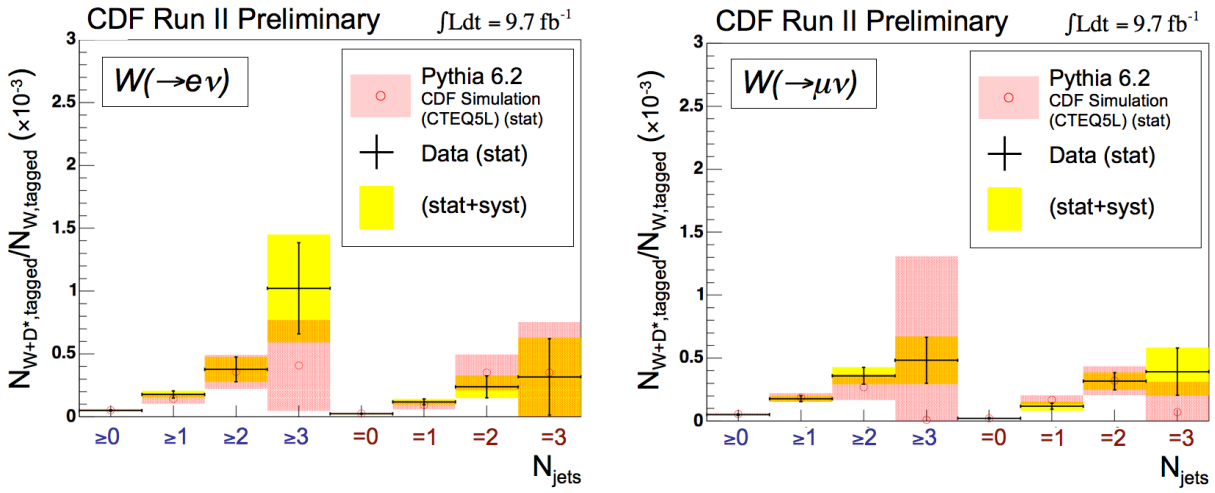


Figure 9.3: The ratio of tagged events $N_{\text{tagged}}(W + D^*)/N_{\text{tagged}}(W)$ as a function of N_{jets}. The left plot is for the $W \to e\nu$ decay mode; the right is for the $W \to \mu\nu$ decay mode. The error bars give the statistical uncertainty, while the sum in quadrature of the statistical and systematic errors is shown as a yellow error band. The red circle and red error band show the same quantity as measured in an inclusive W simulation sample, produced with Pythia 6.2 and using the CTEQ5L PDF set. The red error band corresponds to statistical uncertainties, due to a finite size simulation sample. Note that the uncertainty in these results does *not* take PDF uncertainties into account; this is simply an object-level comparison of MC simulation versus data. Note also that uncertainty in the simulated Pythia results for the $W \to \mu\nu$ case, is greater for the $N_{\text{jets}} \geq 3$ prediction than for the $N_{\text{jets}} = 3$ prediction; this is simply due to a larger number of D^* candidates in the former case—this drives up uncertainty in the fit of $\Delta(m)$ for that case.

than a simple comparison of calorimeter-level jet measurements between simulation and data

As a result, we decide to report simply the calorimeter-level 'tagged' fractions, $N_{\text{tagged}}(D^*)/N_{\text{tagged}}(W + D^*)$, as a function of $N_{\text{jets}}(p_T > 15$ GeV), as first recorded in Chapter 7. (Table 9.4). We plot these results in Figure 9.3. We see no significant disagreement between data and theory. (Speaking fancifully, the slight excess of events in data with $N_{\text{jets}} \geq 3$ may be due to top decays, $t \to W(\to \ell\nu) + b$, in which $b \to$ high p_T jets+D^*; however, we have no direct evidence for this.) With no significant discrepancies observed between theory and data, we do not believe that there is a good reason to unfold back to particle- or parton- level jets; we leave our jet measurements in 'tagged' fraction form. Note that the uncertainty in these results does *not* take PDF uncertainties into account; this is simply an object-level comparison of MC simulation versus data.

Chapter 10

Splitting Signal by Production Process

As explained in section 5.2.2, we expect to tag $W + D^*$ events from three different production processes. We recall the relevant production processes here:

$$\begin{aligned} q + \bar{q}' \to W + g(\to b\bar{b}) \to W + D^* + X &\equiv Wbb \\ q + \bar{q}' \to W + g(\to c\bar{c}) \to W + D^* + X &\equiv Wcc \\ q + g \to W + c \to W + D^* + X &\equiv Wc \end{aligned} \tag{10.1}$$

We would like to determine what percentage of our $W + D^*$ signal sample comes from each one of these production processes. This does not affect our unfolded cross-section ratios, as reported in Chapter 9. However, the ability to look at particular production sources is uncommon in jet-based techniques, making this the first real test of production process in $W/Z + c$ events. We accomplish this in three ways: by using neural networks; by using the signed impact parameter of the reconstructed D^0 vertex; and by looking at the sign relationship of the W and D^*.

10.1 Splitting with neural networks

As explored in section 5.2, a neural network will generally identify D^* from different sources (*Wcc*, *Wc*, and *Wbb)* with different efficiencies. We can take advantage of this to measure the fraction of our signal that comes from each production process. We first train three neural networks: NN_{Wcc}, NN_{Wc}, and NN_{Wbb}. Each network is trained to identify one type of signal event (labeled in the subscript) versus the *BSP* and *SS* background samples described in section 5.2.3. We run each of these three neural networks over each of our three simulated Monte Carlo signal samples (*Wcc*, *Wc,* and *Wbb*). This gives us the efficiencies with which each neural network is expected to identify each type of event. Choosing a neural network cut score of $NN_{\text{cut}} = 0.0$ for each case, we collect nine efficiencies as recorded in Table 10.1.

Next, we run our pre-neural network $W + D^*$ candidates (from data) through each neural network, ad

Production process	$\epsilon(NN_{Wcc})$ $NN_{cut} = 0.0$	$\epsilon(NN_{Wbb})$ $NN_{cut} = 0.0$	$\epsilon(NN_{Wc})$ $NN_{cut} = 0.0$
Wcc	0.818	0.595	0.759
Wbb	0.730	0.868	0.693
Wc	0.857	0.693	0.834

Table 10.1: The efficiencies with which our three basic neural networks identify D^* from each of the three first-order $W + D^*$ production processes; all D^* candidates must have a neural network score $NN_{\text{score}} > NN_{\text{cut}} = 0.0$ in order to pass, for all cases.

Production process	$\epsilon(NN_{Wcc} + NN_{Wcc \text{ vs } Wbb})$ $NN_{cut} = 0.0$ (both NNs)	$\epsilon(NN_{Wbb} + NN_{Wbb \text{ vs } Wcc})$ $NN_{cut} = 0.0$ (both NNs)	$\epsilon(NN_{Wc} + NN_{Wc \text{ vs } Wcc})$ $NN_{cut} = 0.0$ (both NNs)
Wcc	0.618	0.190	0.357
Wbb	0.213	0.677	0.295
Wc	0.652	0.195	0.535

Table 10.2: The efficiencies with which neural network pairs identifyD^* from each of the three first-order $W + D^*$ production processes. All D^* candidates must have neural network scores $NN_{\text{score}} > NN_{\text{cut}} = 0.0$ according to each neural network in a pair, in order to pass. This 'second-order' technique provides measurements of X_{D^*} with greatly reduced uncertainty, relative to the values found using unpaired neural networks.

require that each candidate have a neural network output score $NN_{\text{score}} > NN_{\text{cut}} = 0.0$ in order to pass. This results in three new Δm distributions; we fit these distributions to our standard signal plus background hypothesis, in order to make three different counts of $W + D^*$ signal events: $N_{D^*}^{Wcc}$, $N_{D^*}^{Wc}$, and $N_{D^*}^{Wbb}$. We now put these three counts, as well as our nine efficiency values, into equation 10.2. Solving, we find the fraction of pre-neural network $W + D^*$ signal events that fall into each production category; these fractions are labeled $X_{D^*}^{Wcc}$, $X_{D^*}^{Wc}$ and $X_{D^*}^{Wbb}$.

$$\begin{pmatrix} \epsilon(NN_{Wcc})_{Wcc,} & \epsilon(NN_{Wcc})_{Wc,} & \epsilon(NN_{Wcc})_{Wbb,} \\ \epsilon(NN_{Wc})_{Wcc,} & \epsilon(NN_{Wc})_{Wc,} & \epsilon(NN_{Wc})_{Wbb,} \\ \epsilon(NN_{Wbb})_{Wcc,} & \epsilon(NN_{Wbb})_{Wc,} & \epsilon(NN_{Wbb})_{Wbb,} \end{pmatrix} \begin{pmatrix} X_{D^*}^{Wcc} \cdot N_{D^*}^{TOT} \\ X_{D^*}^{Wc} \cdot N_{D^*}^{TOT} \\ X_{D^*}^{Wbb} \cdot N_{D^*}^{TOT} \end{pmatrix} = \begin{pmatrix} N_{D^*}^{NN_{Wcc}} \\ N_{D^*}^{NN_{Wc}} \\ N_{D^*}^{NN_{Wbb}} \end{pmatrix} \tag{10.2}$$

Doing this, we find fractions as reported in Table 10.3. Our uncertainty in these fractions is very large. We can reduce this uncertainty by adding a second layer of three new neural networks: $NN_{Wcc \text{ vs } Wbb}$, $NN_{Wbb \text{ vs } Wcc}$, and $NN_{Wc \text{ vs } Wcc}$. As the subscripts would imply, each new neural network is trained to identify one type of production process as signal, and another as background. This provides further separation between the three types of production process, which in turn allows finer measurement of the fractions X_{D^*}. We train this second tier of neural networks not only on the standard set of D^* properties (Table 5.4), but also on the score that each event is given by the three original neural networks, NN_{Wcc}, NN_{Wc}, and NN_{Wbb}.

Grouping each first-order neural network with one second order neural network as summarized in the headers of Table 10.2, and requiring that a D^* candidate must pass both neural networks in each set in

Production process $Pre-NN_{Wcc}$ *application*	Fraction found using single neural networks	Fraction found using neural network pairs
Wcc	0.10 ± 2.03	0.55 ± 0.30
Wc	0.80 ± 2.03	0.31 ± 0.28
Wbb	0.10 ± 0.16	0.14 ± 0.06

Correlation coefficient	For results using single neural networks	For results using neural network pairs
$\rho_{wcc,wc}$	-0.997	-0.982
$\rho_{wbb,wc}$	-0.022	0.177
$\rho_{wbb,wcc}$	0.056	-0.361

Table 10.3: The fraction of pre-neural network $W + D^*$ events that falls into each production category, as well as the correlation coefficients for these fractions. Note that 'Pre-NN_{Wcc} application' means that these fractions describe our $W + D^*$ signal *before* applying *any* neural network to reduce uncertainty. Once a neural network has been applied to our D^* candidates, the fractional content of the remaining signal shifts slightly, as described in Table 10.4.

order to be counted, we find nine new efficiencies (also in Table 10.2). We then run our pre-neural network $W + D^*$ selection through each neural network pair, to obtain three new counts N_{D^*}. Substituting these new efficiencies and counts into equation 10.2, we solve to find improved, 'second-order' fractions as reported in Table 10.3. On the whole, these second-order fractions have greatly reduced uncertainty with respect to the first-order fractions.

We are also interested in the correlation coefficients describing the relationship between the fraction uncertainties. For that reason, they are recorded in Table 10.3. We see that the *Wcc* and *Wc* fractions are the most strongly anti-correlated—essentially, the neural networks may confuse these two types of signal with one another, but much less frequently confuse them with *Wbb* events.

We would finally like to take the fractions reported in Table 10.3—which refer to the $W + D^*$ signal's fractional content *before* applying the neural network—and use them to determine the signal's fractional content *after* applying neural network NN_{Wcc} to reduce total uncertainty. This is done rather simply, using the NN_{Wcc} efficiencies in Table 10.1, and the correlation coefficients recorded in Table 10.3. Results are recorded in Table 10.4. However, we would like to see if we can back up these measurement using other methods. That is the purpose of the next two sub-sections.

10.2 Splitting with the D^0 signed impact parameter

We now consider a more traditional method for separating direct charm production (*Wcc* and *Wc*), from charm produced by the decay of bottom hadrons (*Wbb*): we look at the impact parameter (with respect to the beamline) of the reconstructed D^0 vertex. D^* from direct charm ought to have a smaller impact parameter, on average, than D^* from *Wbb*.

Production process $Post-NN_{Wcc}$ *application*	Fraction found working from single neural network results	Fraction found working from neural network pair results
Wcc	0.27 ± 1.90	0.56 ± 0.30
Wc	0.63 ± 1.91	0.32 ± 0.29
Wbb	0.09 ± 0.14	0.13 ± 0.05

Correlation coefficient	Correlation found working from single neural network results	Correlation found working from neural network pair results
$\rho_{wcc,wc}$	-0.997	-0.986
$\rho_{wbb,wc}$	-0.075	0.153
$\rho_{wbb,wcc}$	0.002	-0.312

Table 10.4: The fractional content of post-neural network $W + D^*$ signal, by production process. Also, the correlation coefficients for these post-neural network results. Note that 'Post-NN_{Wcc} application' means that these fractions describe our $W + D^*$ signal *after* applying a NN to reduce uncertainty. The NN that we applied to reduce uncertainty is not directly related to the NNs used to derive these fractions.

10.2.1 Development

In terms of impact parameter, we split events within our signal region into three groups: direct charm production, charm from bottom hadron decays, and fake D^* background. We didn't need to worry about background events in the previous section, as we were able to measure signal count directly—however, background events will affect the $b_s(D^0,\ \text{beam})$ distributions in this section, and must be considered.

Using the the Monte Carlo and background samples described in section 5.2.3, we develop template $b_s(D^0,\ \text{beam})$ distributions for each type of the three types of events (Table 10.5), . This is done by binning the (randomly) signed impact parameter 10.5 for all D^* candidates, and then fitting the resulting histogram with a gaussian plus symmetric-exponential ($e^{-|x-x_0|/\sigma}$) hypothesis. We sign the impact parameter *randomly* in order to make the fit easier—we find that better fits can be made using symmetric distributions. The gaussian piece of the fit is meant to model events near the beamline (where the distribution shape is dominated by vertex resolution); the exponential piece is meant to fit the tail, the shape of which is determined by the lifetime of the D^0 (and bottom hadron) decays. We generate $b_s(D^0,\ \text{beam})$ templates both with and without a cut on the NN_{Wcc} neural network value (Figure 10.1), to see what sort of effect the neural network might have. Generally, applying the neural net decreases the *Wbb* fraction relative to the *Wcc* and *Wc* content, and reduces the size of the $b_s(D^0,\ \text{beam})$ tails.

Fitting these impact parameter distributions is a slightly tricky process. We wish to include as much detail as possible about the gaussian core and the exponential tail of each distribution. However, selecting a fit region that is too small results in tails that are much longer than the data would suggest. For each case, then (with and without a cut on the neural network value), we choose a fit range such that the following conditions are satisfied: we see a (arbitrarily-defined) 'good' amount of the gaussian core of the distribution, occupying somewhere between one quarter and one-half of the visible range; and at the same time, we

Event Type	Sample used to generate template	# events used
Charm from bottom hadron decays	$q + \bar{q}' \to W + g(\to b\bar{b}) \to W + D^* + X$ (MC)	2710
Direct charm production	$q + \bar{q}' \to W + g(\to c\bar{c}) \to W + D^* + X$ (MC)	8437
Background	SS and BSP background	8437

Table 10.5: The three types of events that we seek to separate by impact parameter, and the samples used to represent each. While the direct charm events and D^* from bottom decay are modeled well by Monte Carlo (MC) samples, we use background samples from data (the samples described in section 5.2.3). Recall that SS background refers to candidates which pass all stages of our D^* selection criteria, except that all tracks have the same charge; in BSP background, the soft pion track has charge opposite that of the harder pion track.

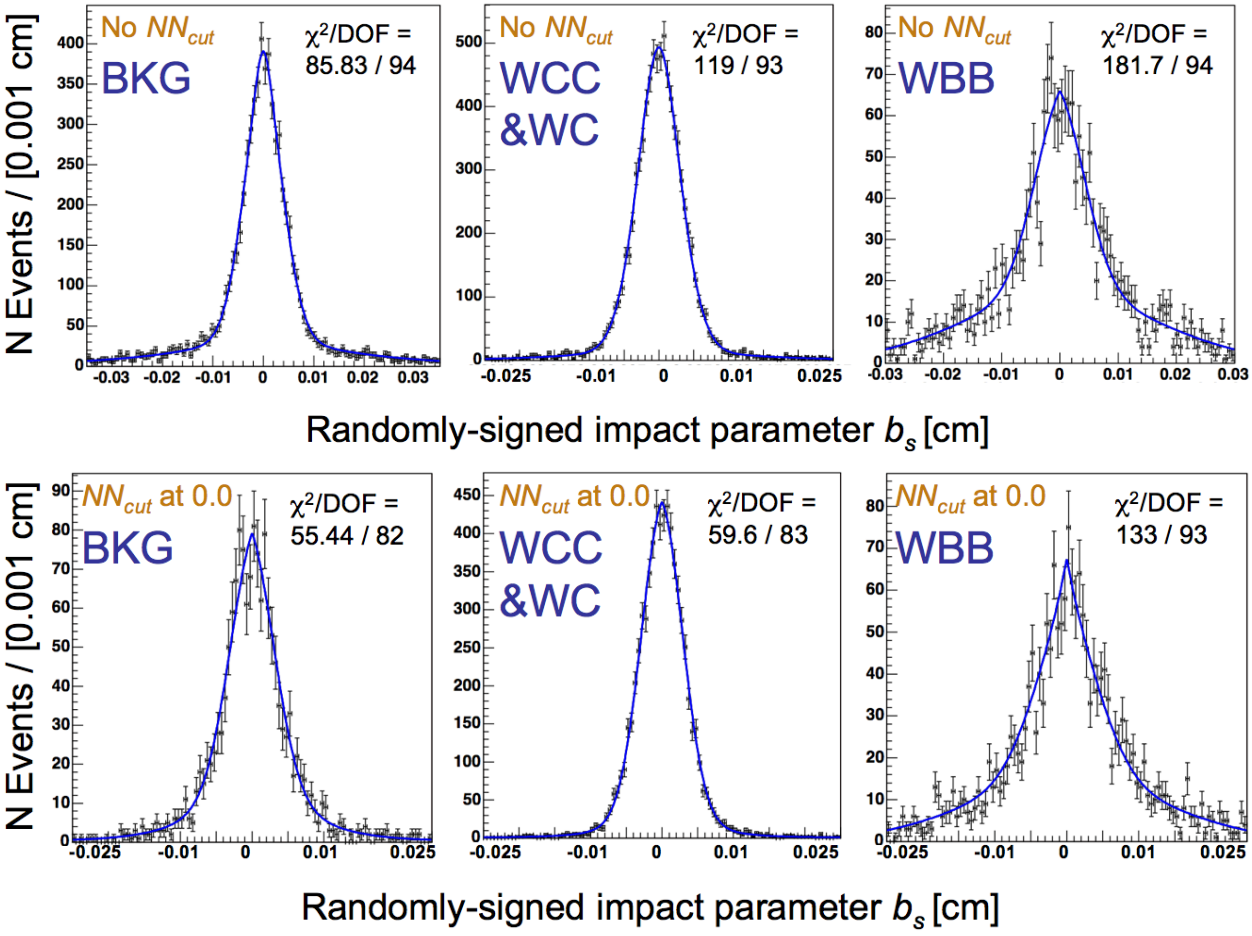


Figure 10.1: Example $b_s(D^0$, beam) distributions for the three types of event that we wish to distinguish between. Each distribution is fit with a gaussian plus symmetric-exponential function. The top row features distributions without any neural network cut; the bottom row shows distributions with a NN_{Wcc} neural network cut of 0.0. *Note bene* the difference in x-axis scale between the *Bkg* and *Wbb* distributions in the top row, and all other distributions.

require that the integral beneath this (normalized) distribution approaches unity as we extend the limits of integration to three times the range used to make the fits. This latter requirement comes from observations of the $b_s(D^0,\ \text{beam})$ distributions of our background and Monte Carlo signal samples, which fall to zero rather quickly.

10.2.2 Testing

With these templates defined, we can determine the D^* content of our signal region. We will do this by fitting the $b_s(D^0,\ \text{beam})$ distribution of all D^* candidates within 3σ of the Δm peak, with a combination of these three templates; the shapes of the templates are held fixed, and only the amplitudes are allowed to vary. Before doing this, however, we test the validity of this method with a series of 'toy' histograms.

For each of several values of $X_{D^*}^{Wbb}$, we generate a blank histogram. We then fill this histogram with the $b_s(D^0,\ \text{beam})$ values of 'simulated' background events, drawing entries at random from the *Bkg* $b_s(D^0,\ \text{beam})$ template (treating this template as a probability distribution). We then add a number of 'simulated' signal events in the same way, drawing from the *Wcc* and *Wbb* templates. For each test, the number of signal and background events added is based on the numbers found by fitting our Δm distribution in data, and then integrating the background and signal functions across the signal region (3σ from Δm_{peak}). Specifically, each test uses a value taken from gaussian distributions, which are centered upon the integrated signal and background values. The width of the gaussians is the uncertainty of the fit, propogated through to the integral. When all events have been added, we fit this distribution with a combination of the three templates, and see how well the input fractions $X_{D^*}^{Wbb}$ and $X_{D^*}^{Wcc/Wc}$ are recreated.

That is the general concept, at any rate. In practice, we fix the amplitude of the background curve to the total number of background events predicted by our Δm fit. We also fix the combined amplitudes of the charm and bottom templates to equal the total number of signal events predicted by the Δm fit. Only the fraction $X_{D^*}^{Wbb}$ is allowed to vary. Uncertainty in the meaured values of $X_{D^*}^{Wbb}$—as determined by these tests—should take into account the effect of fixing our total signal and background amplitudes.

It turns out that the effectiveness of this 'template' approach depends upon not only the number of signal and background events that populate the signal region, but also the fraction $X_{D^*}^{Wbb}$ being tested. We therefore test this technique for multiple values of X_b, ranging from 0.0 to 0.5 in units of 0.05. Creating and fitting $100,000$ toy histograms for each value of $X_{D^*}^{Wbb}$, we find good agreement between the input fraction and (average) ouput fraction (Figure 10.2).

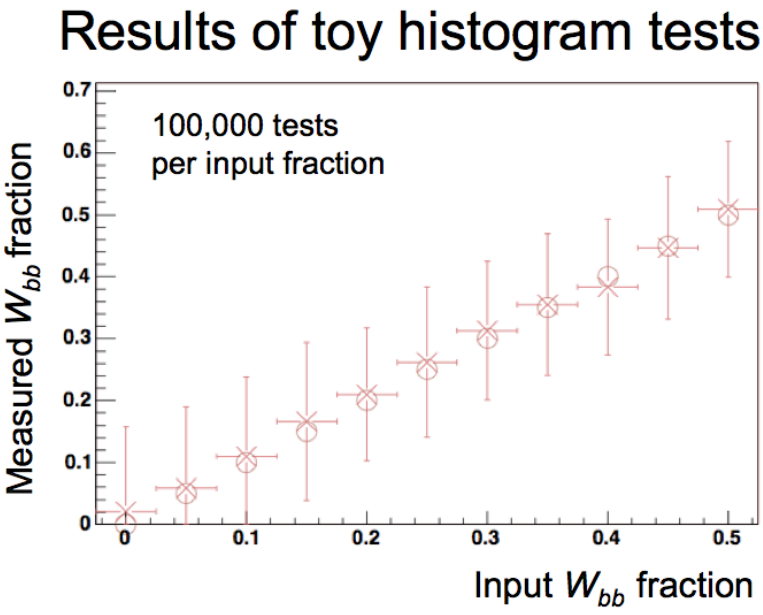


Figure 10.2: Measured values $X_{D^*}^{Wbb}$ obtained for several input values of $X_{D^*}^{Wbb}$. Each input value is tested by fitting 100,000 toy $b_s(D^0,\ \text{beam})$ histograms, each of which is generated with a number of signal and background events compatible with what we observe in data. We see that there is good agreement between the average measured value of $X_{D^*}^{Wbb}$, and the input value being tested. The tests shown are for post-neural network templates.

Sample Type	$X_{D^*}^{Wbb}$
No neural network cut	0.02 ± 0.12 (stat) ± 0.06 (syst)
Neural network cut at 0.0	0.08 ± 0.07 (stat) ± 0.03 (syst)

Table 10.6: Fractions $X_{D^*}^{Wbb}$, as found by fitting the $b_s(D^0,\ \text{beam})$ distribution of all events within 3σ of the Δm signal region, using a set of templates.

10.2.3 Results

With the effectiveness of this technique established, we can now apply it to the $b_s(D^0,\ \text{beam})$ distribution of all events within 3σ of the Δm signal peak, in our post-neural network $W + D^*$ data sample. We find fractions $X_{D^*}^{Wbb}$ as reported in Table 10.6.These results are compatible with those found in the previous section (Tables 10.3 and 10.4), but have larger uncertainty. As such, we will use the Wbb fractions found in the previous section (using the neural network-based technique) in our final results. It is nice, however, to see that our results are backed up by this more traditional approach.

.

10.3 Splitting with D^* and W signs

Recall our three $W + D^*$ production processes, as summarized in Eq. 10.1. In the case of Wc production, conservation of charge requires that the W and c be produced with opposite signs. When the W decays leptonically, the lepton will always have the same sign as the W. When the c hadronizes to a D^*, the c and D^* will have the same sign as one-another. In sum, this means that for a Wc event, the W and D^* that we tag ought to have opposite signs. We refer to such events as $(W + D^*)_{OS}$.

In the case of either Wcc or Wbb production, however, the W is produced along with a c and $\bar{c}$ (or b and $\bar{b}$), either of which could give rise to the D^* that we tag. This means that in a Wcc or Wbb event, we are equally likely to tag oppositely-signed W and D^*, $(W + D^*)_{OS}$, or same-signed W and D^*, $(W + D^*)_{SS}$.

By considering the difference between the number of $(W + D^*)_{OS}$ and $(W + D^*)_{SS}$ events tagged by our algorithm, we can estimate the contributions of Wc events versus Wcc/Wbb events. Quantitatively

$$N_{OS} = N_{Wc} + \frac{1}{2} N_{Wcc/Wbb}$$

$$N_{SS} = \frac{1}{2} N_{Wcc/Wbb}$$

and so

$$N_{Wc} = N_{OS} - N_{SS}$$

$$N_{Wcc/Wbb} = 2 \cdot N_{SS}$$

Measuring these quantities in our combined $W(\rightarrow \mu\nu/e\nu) + D^*$ sample (Figures 5.3,5.11, and Table 5.2), we find

$$\begin{aligned} X_{D^*}^{Wc}(\text{no neural network cut}) &= 0.107 \pm 0.097 \text{ (stat)} \\ X_{D^*}^{Wc}(\text{with neural network cut at 0.0}) &= 0.141 \pm 0.063 \text{ (stat)} \end{aligned} \tag{10.3}$$

Production process	Fraction of $W(\rightarrow \ell\nu) + D^*$ signal
$s(d) + g \rightarrow W + c$	$X_{D^*}^{Wc} = 0.14 \pm 0.06$
$q + \bar{q}' \rightarrow W + g(\rightarrow c\bar{c})$	$X_{D^*}^{Wcc} = 0.73 \pm 0.08$
$q + \bar{q}' \rightarrow W + g(\rightarrow b\bar{b})$	$X_{D^*}^{Wbb} = 0.13 \pm 0.05$

Table 10.7: The best measurements of the fraction of our $W + D^*$ signal that comes from processes Wbb and Wc. The best measurement for the Wbb fraction, $X_{D^*}^{Wbb}$, comes from our NN-based approach (section 10.2), while our best measurement for the Wc fraction, $X_{D^*}^{Wc}$, comes from the $OS - SS$ sign-based approach (section 10.2).

10.4 Combining results

We have determined what fraction of our sample comes from each of three production process, using several different techniques. Our best measurement for the fraction of Wc events, $X_{D^*}^{Wc}$, is found using the $OS - SS$ sign-based technique (section 10.3); our best measurement for the fraction of Wbb events after the neural network cut, $X_{D^*}^{Wbb}$, is found using the two-tiered neural-network technique (section 10.1). Because we arrived at these results using independent methods, we treat their uncertainties as uncorrelated. We report these best measurements in Table 10.7.

10.5 The effect of D^* from background

In splitting our signal into $Wcc/Wbb/Wc$ fractions, we assumed that the signal was composed entirely of D^* from these three signal processes. We know from Table 6.5 that this is not true—some small number of D^* in our signal comes from background events which fake the W signature. We assume that this effect is negligible within our (fairly large) uncertainties. This may be a question to revisit in the future.

Conclusions

This analysis was motivated by a gap in studies of perturbative QCD: while $W/Z + b$ production is well-tested, $W/Z + c$ production has been studied much less extensively. Never (before now) had we probed the regime $p_T(c) < 15$ GeV; with perturbative QCD relying on assumptions that are less accurate as $p_T \to 0$, this means that the $p_T(c) < 15$ GeV regime is a rigorous test of our models.

With this in mind, we created a new tool for identifying low-momentum charm by fully-reconstructing D^* mesons at the track level. We applied that tool to W/Z events in the CDF high-p_T lepton datasets, and tested this technique extensively. We characterized interesting backgrounds and systematic uncertainties with a high degree of precision, and we proved not only that this technique is viable, but that it is in fact quite successful: we were able to probe the full p_T spectrum of D^* produced in W events (for $p_T(D^*) > 3$ GeV). We found no discrepancy between data and Pythia MC predictions.

We also showed that this technique is capable of measuring much more than just inclusive cross-sections—working within the (relatively low statistics) regime of the CDF high-p_T lepton samples, we measured several other properties describing our events: how D^* production varies as a function of N_{jets}, and what percentage of our $W + D^*$ signal comes from each of three different production processes. These properties could, in the future, act as a starting point for other interesting analyses—not only at CDF (there are other charmed mesons that we could add to our sample), but also in high-statistics datasets at the Large Hadron Collider.

In short, the strength of this analysis lies primarily in its promise: low-momentum charm is infrequently considered a viable 'secondary tag' in events with other major players, like W/Z bosons or top quarks. More often than not, charm is considered an irreducible background, which must be emulated by (potentially faulty) Monte Carlo simulations. We have proven that this is not so—charmed mesons can take the role of a precise probe into the QCD processes that underlie the production of major event types. We have observed no discrepancies between data and Monte Carlo in W/Z production, but this is hardly the end of the road for this kind of work. Rather, the technique that we have developed in this thesis work is easily ported to new studies, and may open the way for a fascinating series of low-momentum measurements of QCD (see **The Future**, below).

The Future

After giving a joint CDF/D0 Tevatron talk at Moriond QCD 2014, this March, I was asked exactly zero questions about my $W/Z + D^*$ work. The reason for this soon became quite clear: among the CDF results that I discussed was a study of $\Upsilon + W/Z$ production. The first order production process for $\Upsilon + W/Z$ production, as advertised in the study's public paper, is simply $u + \bar{d} \to W/Z + g(\to b\bar{b})$, in which the $b\bar{b}$ pair condenses into a bound state. During the coffee break that proceeded my talk, it seemed that representatives from every major European country had become possessed with a monomanical desire to correct me: obviously, this cannot be the first order production process—the gluon is colored, and the Υ must be colorless. What did I really mean?

Whilst silently cursing the authors of that paper, I also became quite intrigued by this question. Kevin Pitts had suggested that we might use our D^* tagging technique to look for D^* produced in association with heavy quarkonium, and it was now easy to see why this would be fascinating. The $b\bar{b}$ pair obviously have to radiate off their color in some fashion in order to become an Υ; by tagging a particular charmed meson, we might be able to probe this process in considerable detail. What is the average angle of emission for this D^*, with respect to the Υ? How much energy or momentum is carried away? More importantly, how well do our models of perturbative QCD work in this extremely close-range, low-momentum menagerie of heavy flavor? If we are going to find significant discrepancies anywhere, it seems that this would be the ideal place for them to be hiding.

For this reason, I have been training a team of two undergraduates (and as of late May, one REU student) to modify my $W/Z + D^*$ tagging software for identifying $\Upsilon(\to \mu\mu) + D^*$ events. There are a colossal number of $\Upsilon + D^*$ candidates to work with in the CDF Run II dataset, compared to our ~ 600 or so $W + D^*$ candidates; from a very preliminary study, we have identified hundreds of thousands of $\Upsilon \to \mu\mu$ candidates, with thousands of D^* candidates split between the Υ events and their backgrounds. We will look at the production rate of D^* associated with each Υ state, and provide what may well be one of the first measurements of 'heavy quarkonium plus heavy flavor' production at low momentum.

This work obviously suggests a future study of $J/\psi + D^*$ production, or perhaps even of $\gamma + D^*$ production.

We might even imagine investigating the $X(3872)$ or $X(4260)$ particles, which some have suggested might be tetraquark states involving a bound $c\bar{c}$ pair [44, 45]. There are a huge number of options, and many opportunities for testing untested regimes of pertubative QCD. The code that we have have developed is very easily ported to other studies, which makes them quick and precedented projects, excellent for undergrads or graduate students, alike. In short, we have opened the door for a whole suite of interesting, low-momentum heavy flavor analyses. What the future may hold, in terms of our understanding of QCD, is unknown but very exciting—and I am eager to see our toolset pushed to potentially fascinating new limits.

Appendix A: Fitting Δm templates

In this appendix, we discuss the process by which we choose our signal and background template functions, for fitting Δm distributions.

Part I: Exploring signal and background shapes

To begin, we consider the *Wcc* Monte Carlo sample. This is our largest sample, and it will provide us with the greatest statistics. We fit the Δm distribution of this sample with two different shapes for signal, and two different shapes for background. The latter tests the stability of each signal shape. For signal, we try each a double gaussian

$$s_{\mathrm{DG}}(x) = N \cdot \Big(\frac{e^{-(x-x_0)^2/2\sigma_1^2}}{\sigma_1\sqrt{2\pi}} + R \cdot \frac{e^{-(x-x_0)^2/2\sigma_2^2}}{\sigma_2\sqrt{2\pi}} \Big) / (1+R) \tag{10.4}$$

$$\text{with } \sigma_2 \equiv R_\sigma \cdot \sigma_1$$

and a triple gaussian

$$s_{\mathrm{TG}}(x) = N \cdot \Big(\frac{e^{-(x-x_0)^2/2\sigma_1^2}}{\sigma_1\sqrt{2\pi}} + R \cdot \frac{e^{-(x-x_0)^2/2\sigma_2^2}}{\sigma_2\sqrt{2\pi}} + R_2 \cdot \frac{e^{-(x-x_0)^2/2\sigma_3^2}}{\sigma_3\sqrt{2\pi}} \Big) / (1+R+R_2) \tag{10.5}$$

$$\text{with } \sigma_2 \equiv R_\sigma \cdot \sigma_1 \text{ and } \sigma_3 \equiv R_{2,\sigma,} \cdot \sigma_1$$

The two signal functions have been normalized such that N is the integral of the function from $-\infty$ to ∞.

For background, we try a power-law

$W + cc$	x_0	σ_1	R	R_σ	R_2	$R_{\sigma,2}$	N	N_{Count}	χ^2/DOF
Pow. Law & DG	0.145546	0.00050	0.51	3.3			2.898	11593	281.6 / 74
Fit Uncertainty (±)	0.000007	0.00001	0.04	0.1			0.034	136	
Poly & DG	0.145547	0.000523	0.50	3.89			3.06	12250	358 / 72
Fit Uncertainty (±)	0.000006	0.000006	0.02	0.06			0.03	103	
Average:	**0.145547**	**0.000512**	**0.51**	**3.6**					
Pow. Law & TG	0.145542	0.00066	0.51	2.9	0.4	0.48	2.931	11723	226.1 / 72
Fit Uncertainty (±)	0.000007	0.00004	0.05	0.1	0.1	0.03	0.034	137	
Poly & TG	0.145543	0.000671	0.57	3.29	0.43	0.48	3.07	12294	274.1 / 70
Fit Uncertainty (±)	0.000005	0.000007	0.02	0.05	0.02	0.02	0.02	84	
Average:	**0.145543**	**0.00067**	**0.54**	**3.1**	**0.42**	**0.48**			

Table A.1: D^* counts in a signal plus background fit of Δm for Wcc MC events. Four permutations of signal and background function shape are tested.

$$b(x)_{\text{POW}} = A \cdot \left(x - m_{\pi^\pm}\right)^p \quad \text{with } m_{\pi^\pm} = 0.1396 \text{ GeV} \tag{10.6}$$

(where $m_{\pi^\pm}$ is the mass of the charged pion), and a third-order polynomial

$$b(x)_{\text{POLY}} = a + b \cdot x + c \cdot x^2 + d \cdot x^3 \tag{10.7}$$

For each combination of signal and background, we fit our Wcc Monte Carlo sample's Δm histogram. The fitted signal function's parameters are recorded in Table A.1. The fitted plots are recorded in Figure A.1.

While none of the permutations of signal and background provides a fit with low χ^2/DOF, we are not looking for an absolutely ideal fit. We are looking only for a reasonable fit to the shape of our MC sample, and the similarity in final signal counts between the four fits hints that we have accomplished just that.

Part II: Making a final decision

We do not expect for the shape of the Δm peak to vary significantly from one type of signal to the next. We will therefore use the results of our Wcc fits to generate a template that describes the shape of the Δm peak. In particular, we use the double and triple gaussian functions from earlier, but fix the R, R_σ, R_2 and $R_{2,\sigma}$ values to their averages across the two double gaussian / triple gaussian fits performed. We then define σ_0 as the average value of σ_1 across the two double gaussian / triple gaussian fits, and set $\sigma_1 = s \cdot \sigma_0$ for

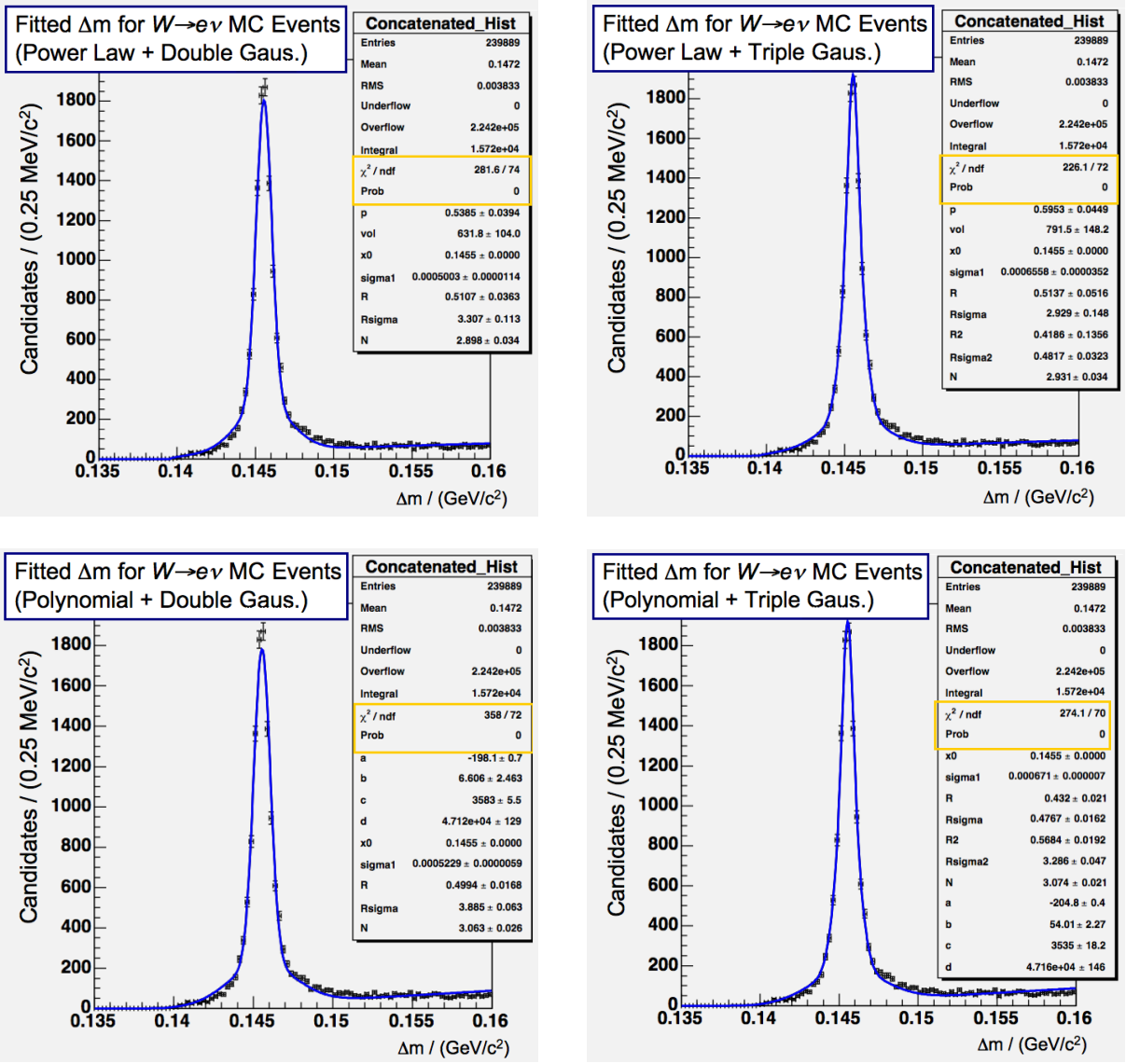


Figure A.1: Fitted Δm plots for Wcc MC events. Four permutations of signal and background function shape are tested.

some new parameter s, which describes by how much a given signal deviates from this template average.

In effect, this fixes the basic shape of the double/triple gaussian signals, allowing only the width and overall amplitude of the function to vary. We have now the double gaussian template:

$$s_{\mathrm{DG,TEMPLATE}}(x) = N \cdot \Big(\frac{e^{-(x-x_0)^2/2{\sigma_1}^2}}{\sigma_1\sqrt{2\pi}} + R \cdot \frac{e^{-(x-x_0)^2/2\sigma_2^2}}{\sigma_2\sqrt{2\pi}} \Big) / (1+R) \tag{10.8}$$

$$\text{with} \qquad \sigma_1 \;=\; s \cdot \sigma_0, \qquad \sigma_2 = R_\sigma \cdot \sigma_1,$$
$$R = 0.51, \qquad R_\sigma = 3.6, \qquad \sigma_0 = 0.000512$$

and the triple gaussian template

$$s_{\mathrm{TG}}(x) = N \cdot \Big(\frac{e^{-(x-x_0)^2/2\sigma_1^2}}{\sigma_1\sqrt{2\pi}} + R \cdot \frac{e^{-(x-x_0)^2/2\sigma_2^2}}{\sigma_2\sqrt{2\pi}} + R_2 \cdot \frac{e^{-(x-x_0)^2/2\sigma_3^2}}{\sigma_3\sqrt{2\pi}} \Big) / (1+R+R_2) \tag{10.9}$$

$$\text{with} \qquad \sigma_1 \;=\; s \cdot \sigma_0, \qquad \sigma_2 = R_\sigma \cdot \sigma_1, \qquad \sigma_3 = R_{\sigma,2} \cdot \sigma_1,$$
$$R = 0.54, \qquad R_2 = 0.42, \qquad R_\sigma = 3.1, \qquad R_{2,\sigma} = 0.48, \qquad \sigma_0 = 0.00067$$

For small signals (such as those that we will measure in data), the choice of double or triple gaussian signal has a negligible impact on the final measurement. We therefore choose the double-gaussian template (Eq 10.8) for fitting our Δm signal peak in data. We do not allow s to vary, in order to prevent statistics from forcing the fit into a sharp artificial peak. Instead, because signal widths in data are generally $10-15\%$ larger than in simulation, we fix $s = 1.1$, and allow only the amplitude to vary. This allows us to take advantage of our knowledge of the Δm signal shape, in order to get the most out of low-statistics samples.

Appendix B: Neural Network Qualities

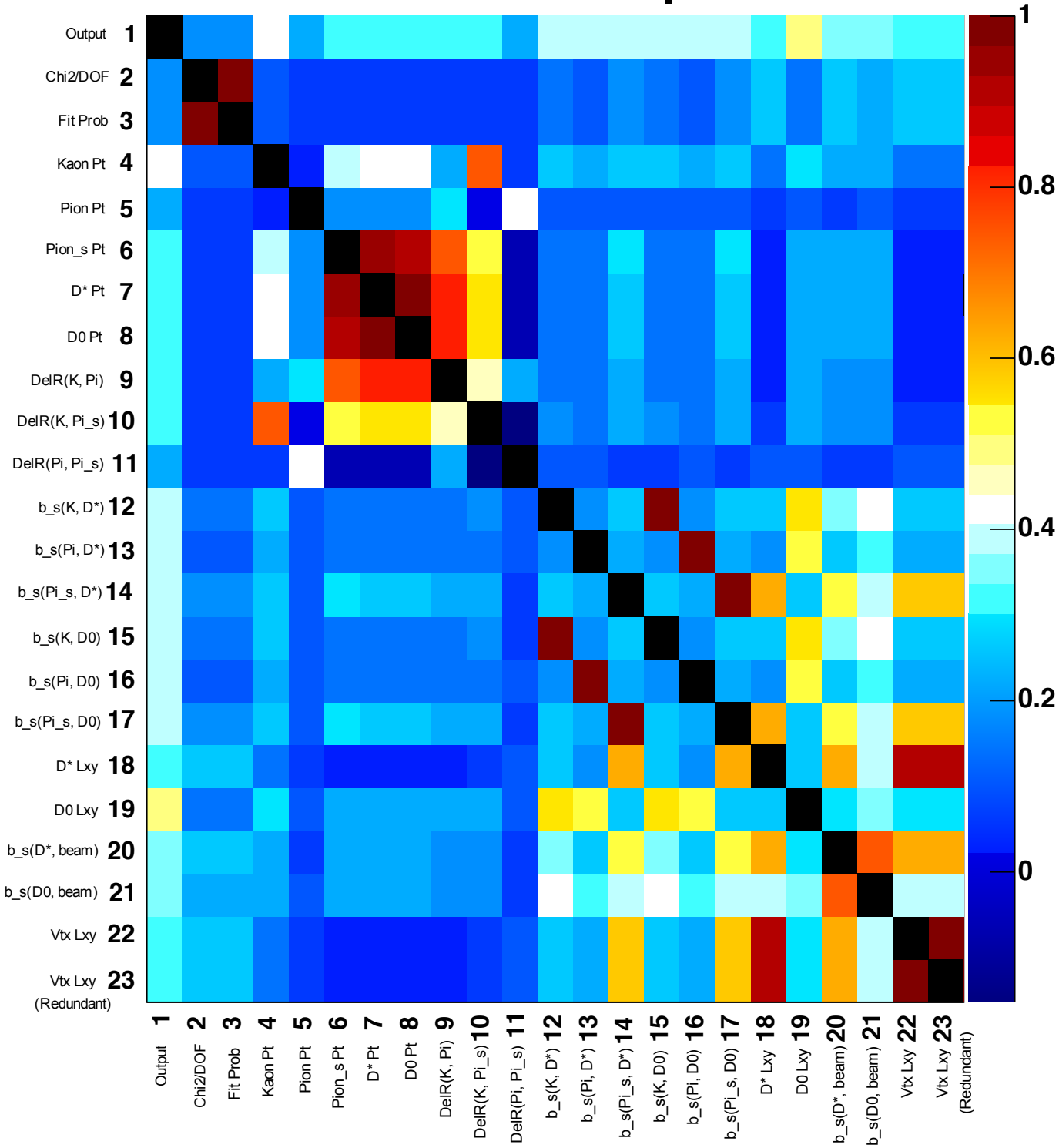
In this appendix, we include output from the training program that we used to develop our NN_{Wcc} neural network. The first page displays the correlation matrix for all variables used to train the network. The pages after this correlation matrix describe the signal (red) and background (black) distributions for each variable. Note that x-axis scales for each distribution *are not constant*. Finally, we include additional output from the training program detailing the signal purity and efficiency, the Gini index, and a visual diagram of how each node in the final neural network connects to each other node.

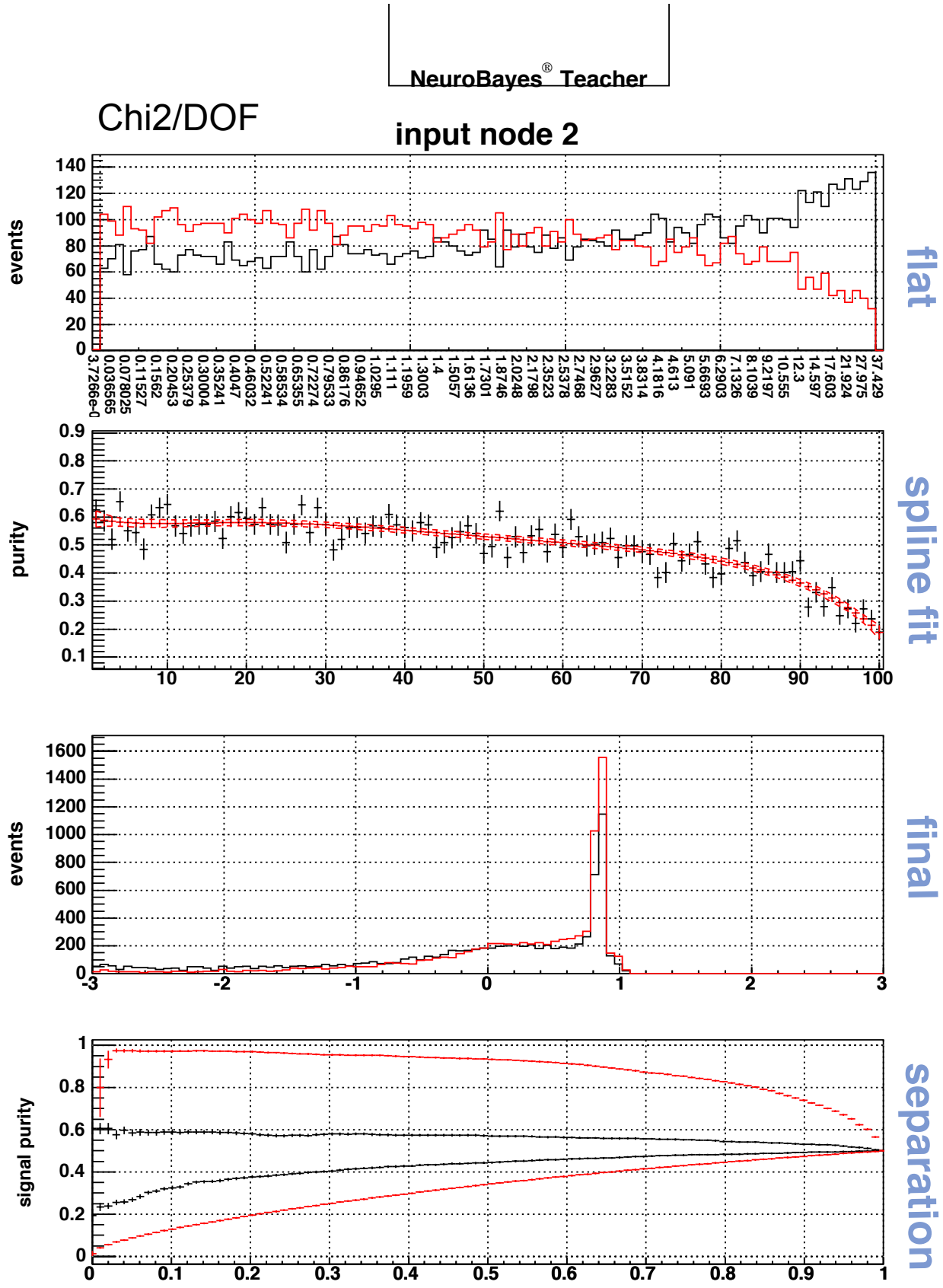
A note on Vtx Lxy and Vtx Lxy: Redundant

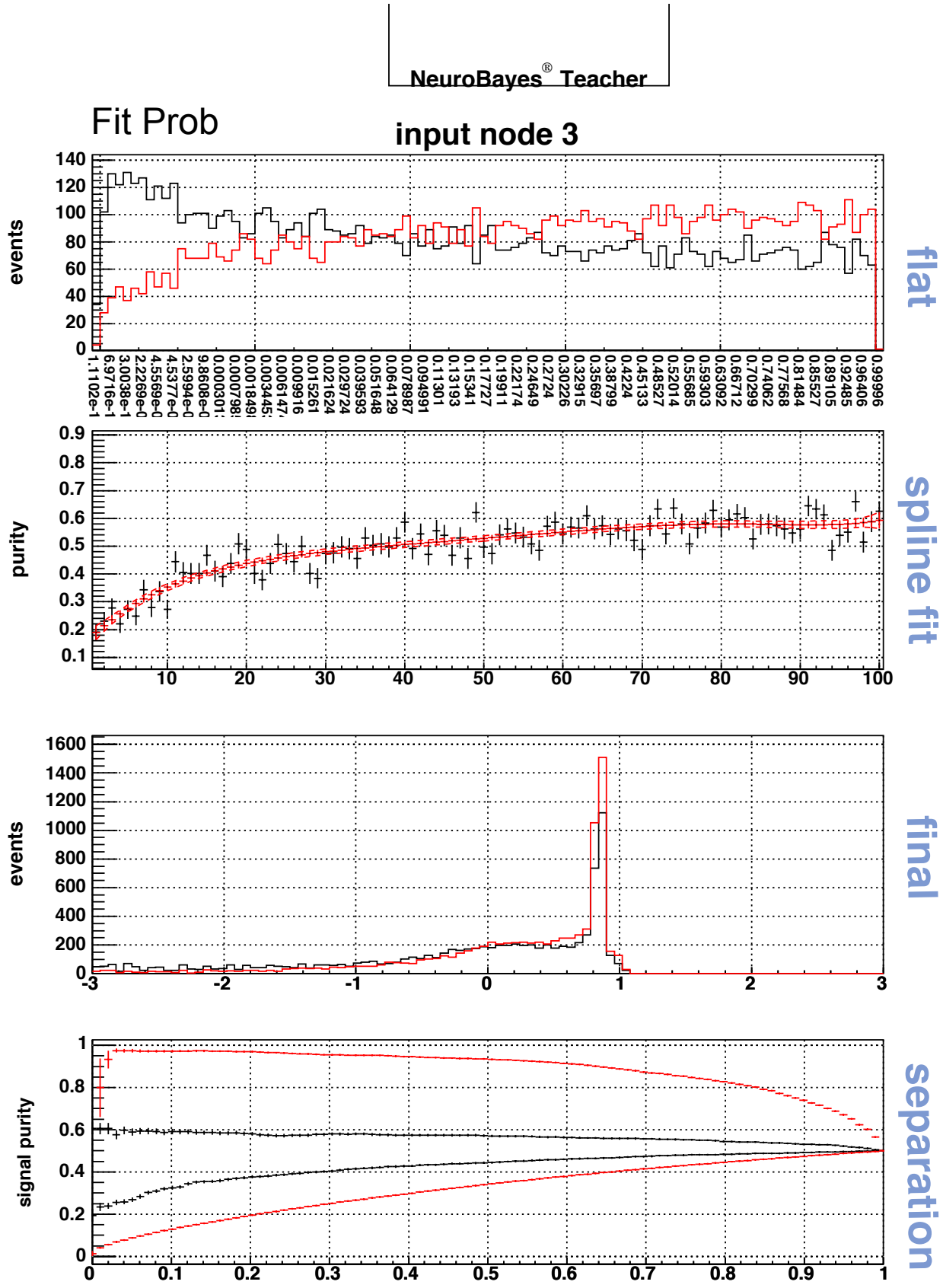
In the pages that follow, we have one variable named Vtx Lxy, and another named Vtx Lxy: Redundant. Both of these variables describe the signed impact parameter of the reconstructed D^* with respect to the D^0 vertex. The redundancy is due to having developed two definitions for this quantity, both of which turned out to be equivalent. This has no effect on the final neural network, as, due to this redundancy, the training program chose to ignore the second Vtx Lxy variable when constructing the neural network.

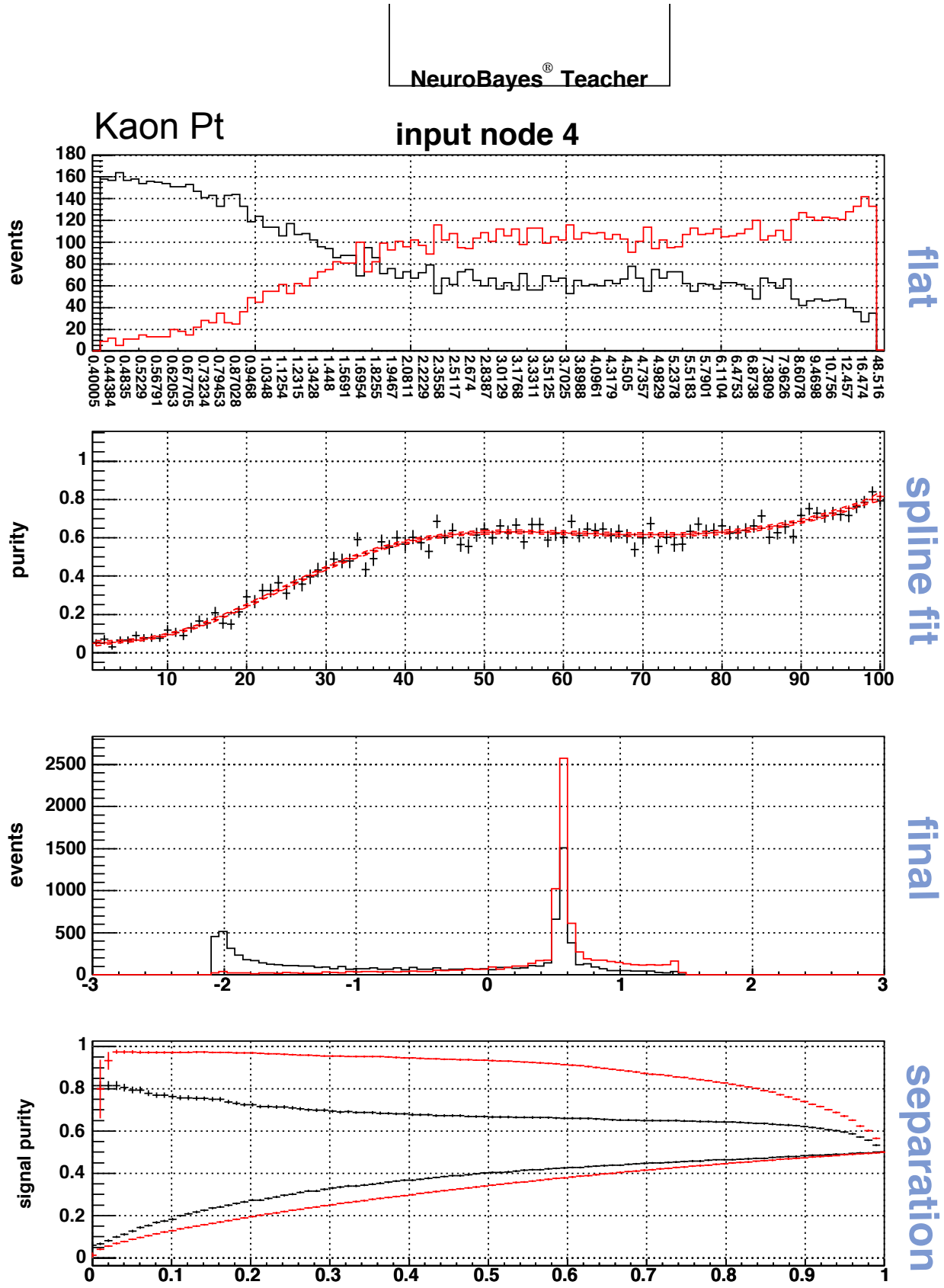
NeuroBayes® Teacher

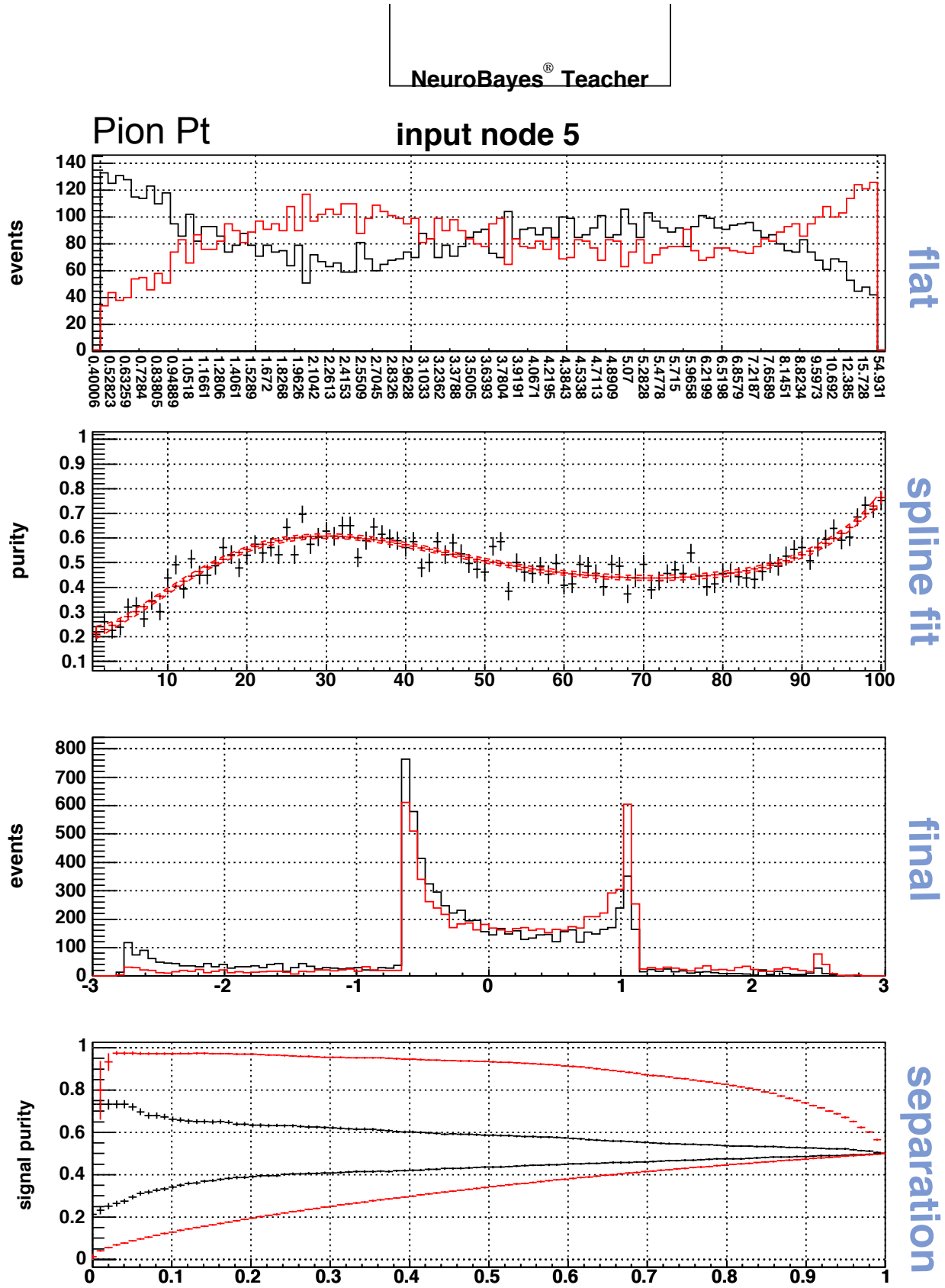
correlation matrix of input variables

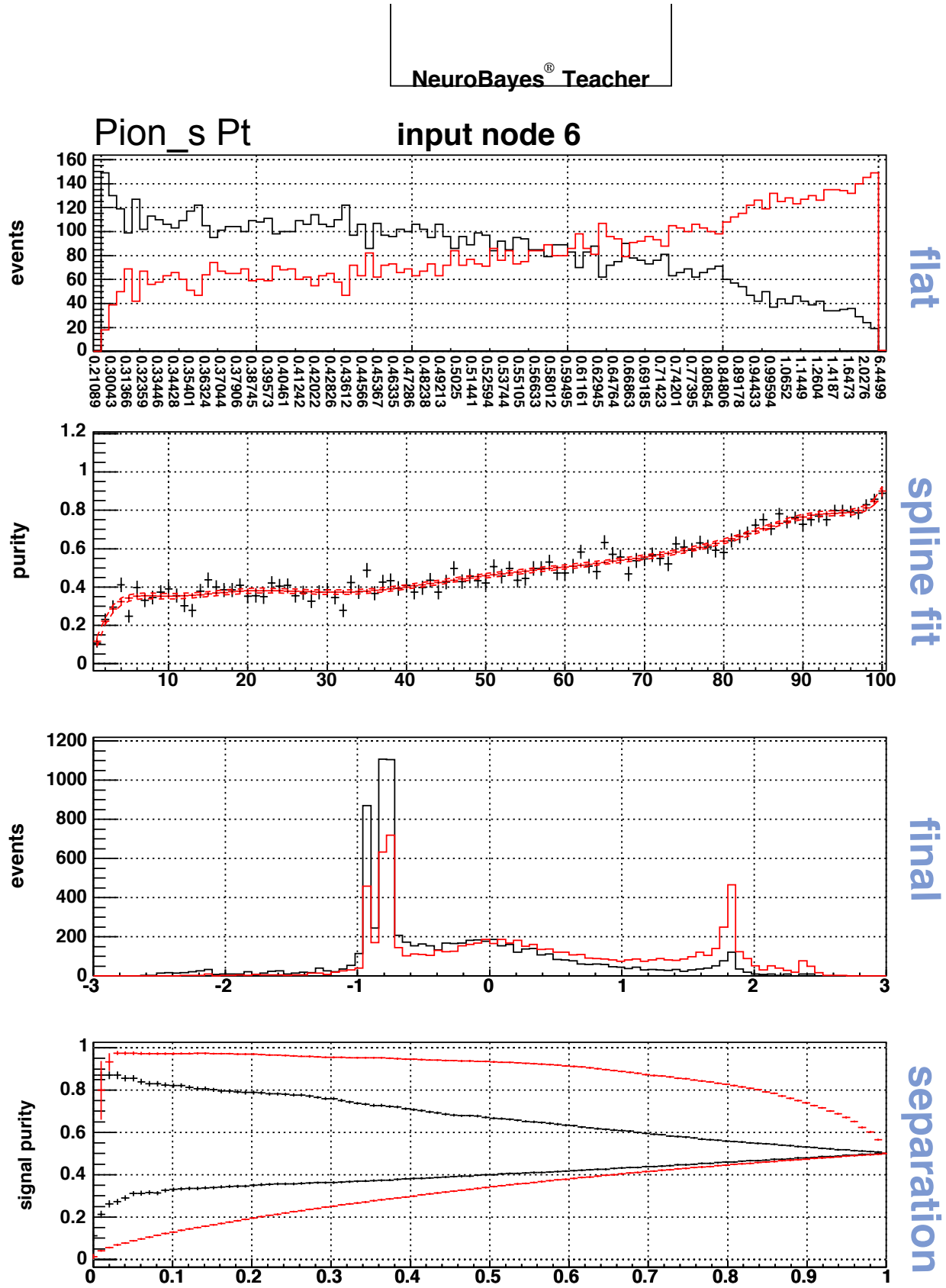


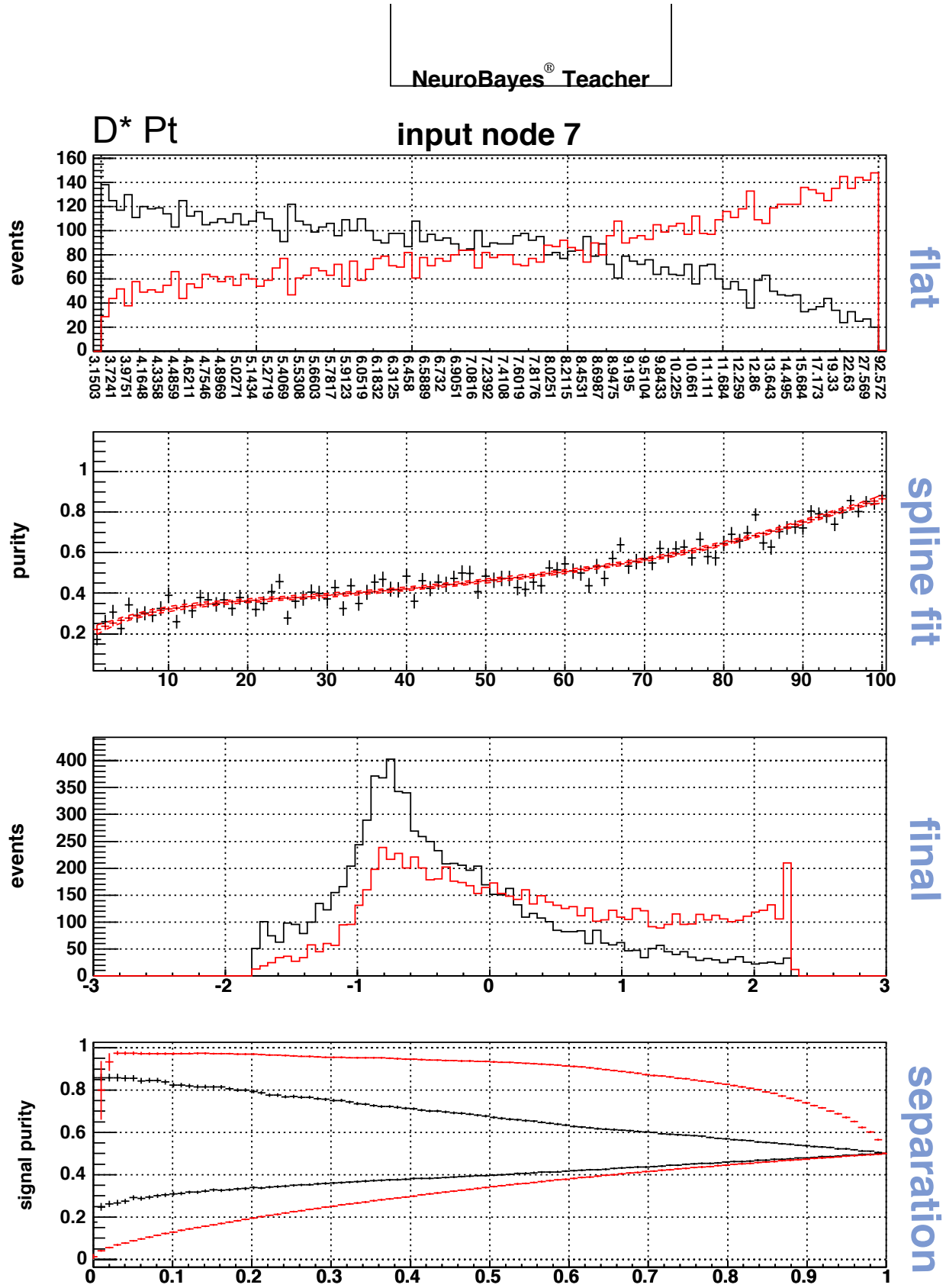

NeuroBayes® Teacher
Chi2/DOF
input node 2
events
flat
purity
spline fit
events
final
signal purity
separation

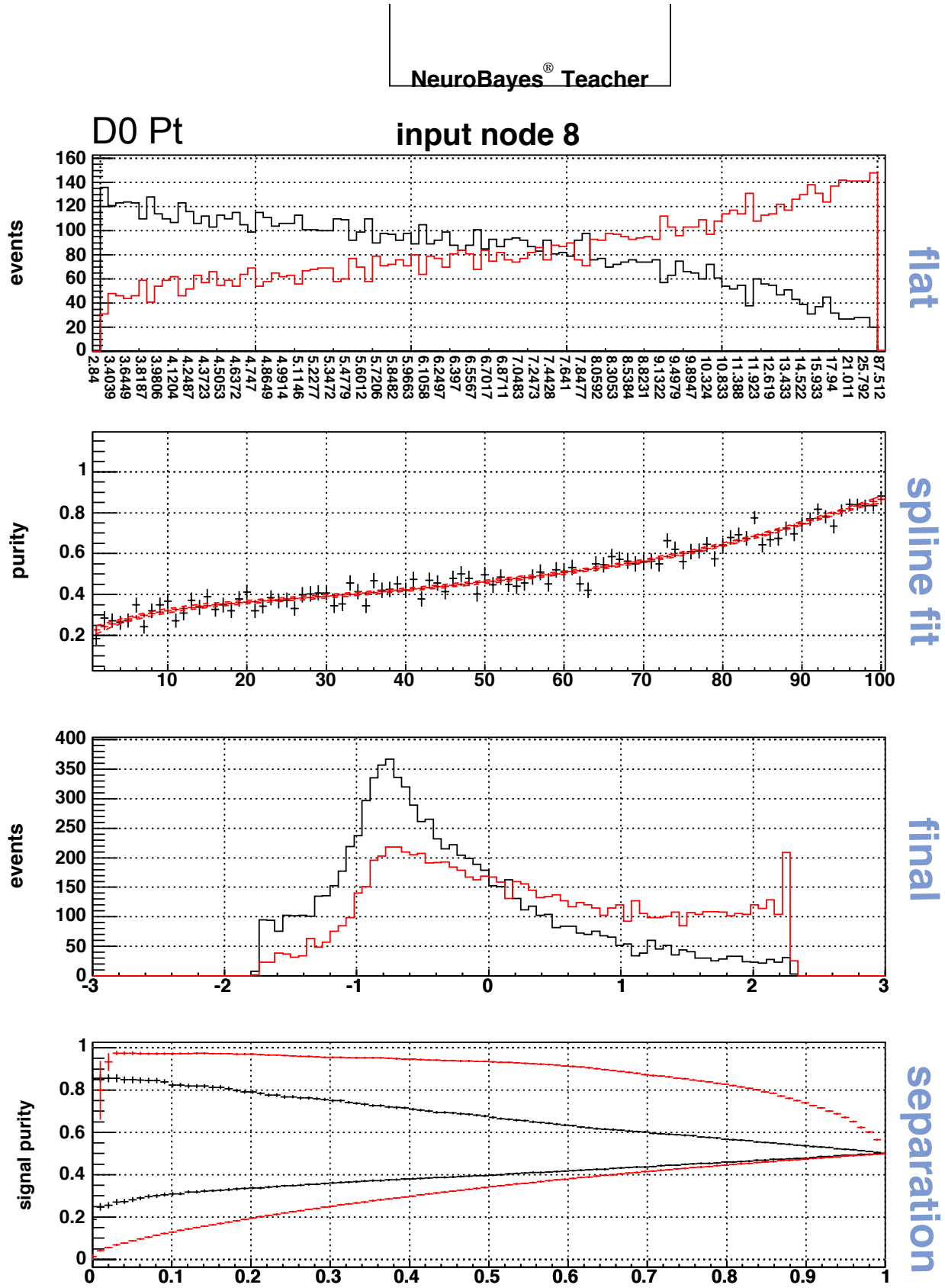

NeuroBayes® Teacher
Fit Prob
input node 3
events
flat
purity
spline fit
events
final
signal purity
separation

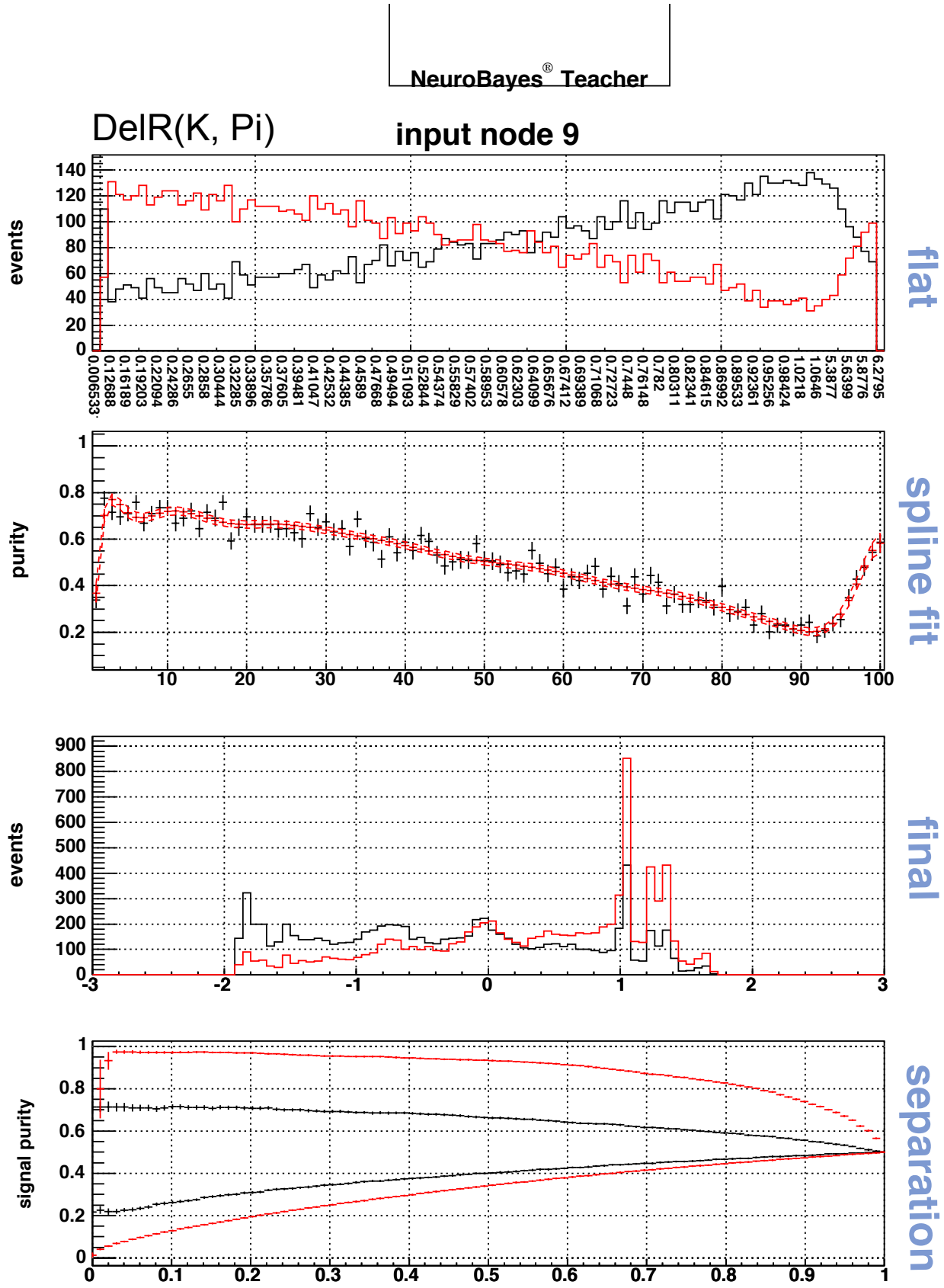

NeuroBayes® Teacher
Kaon Pt
input node 4
events
flat
purity
spline fit
events
final
signal purity
separation

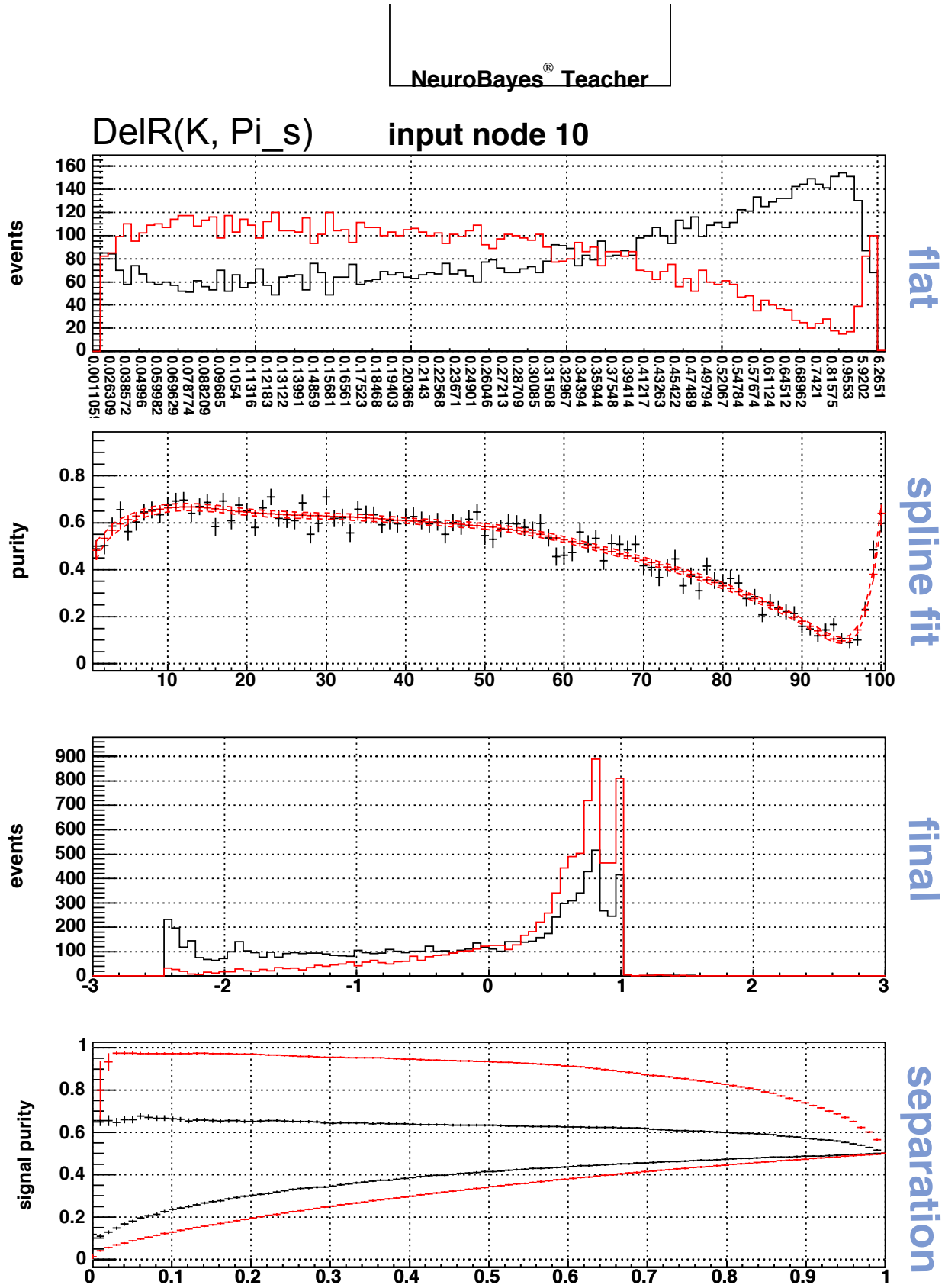

NeuroBayes® Teacher
Pion Pt
input node 5
events
flat
purity
spline fit
events
final
signal purity
separation

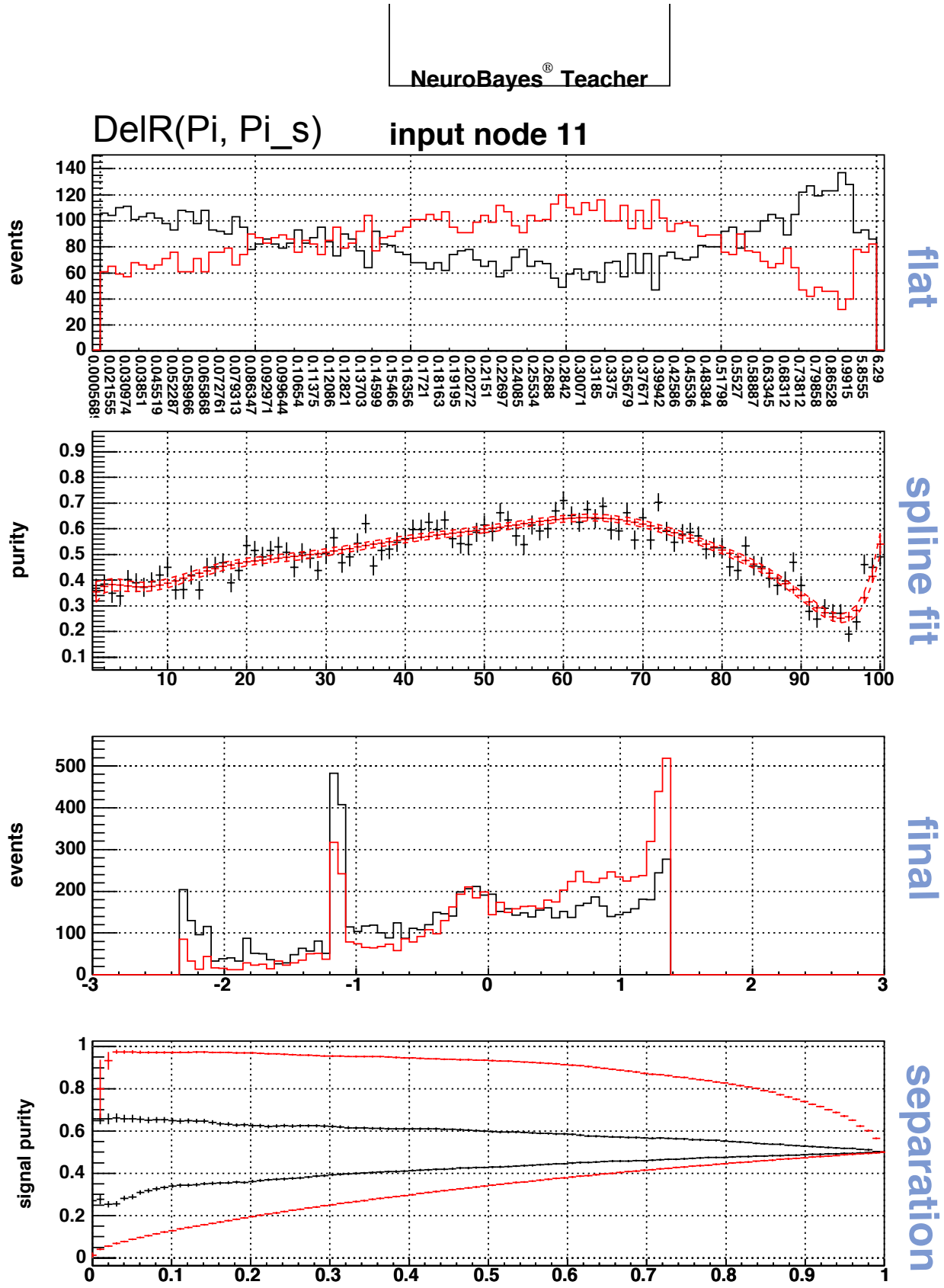

NeuroBayes® Teacher
Pion_s Pt
input node 6
events
flat
purity
spline fit
events
final
signal purity
separation

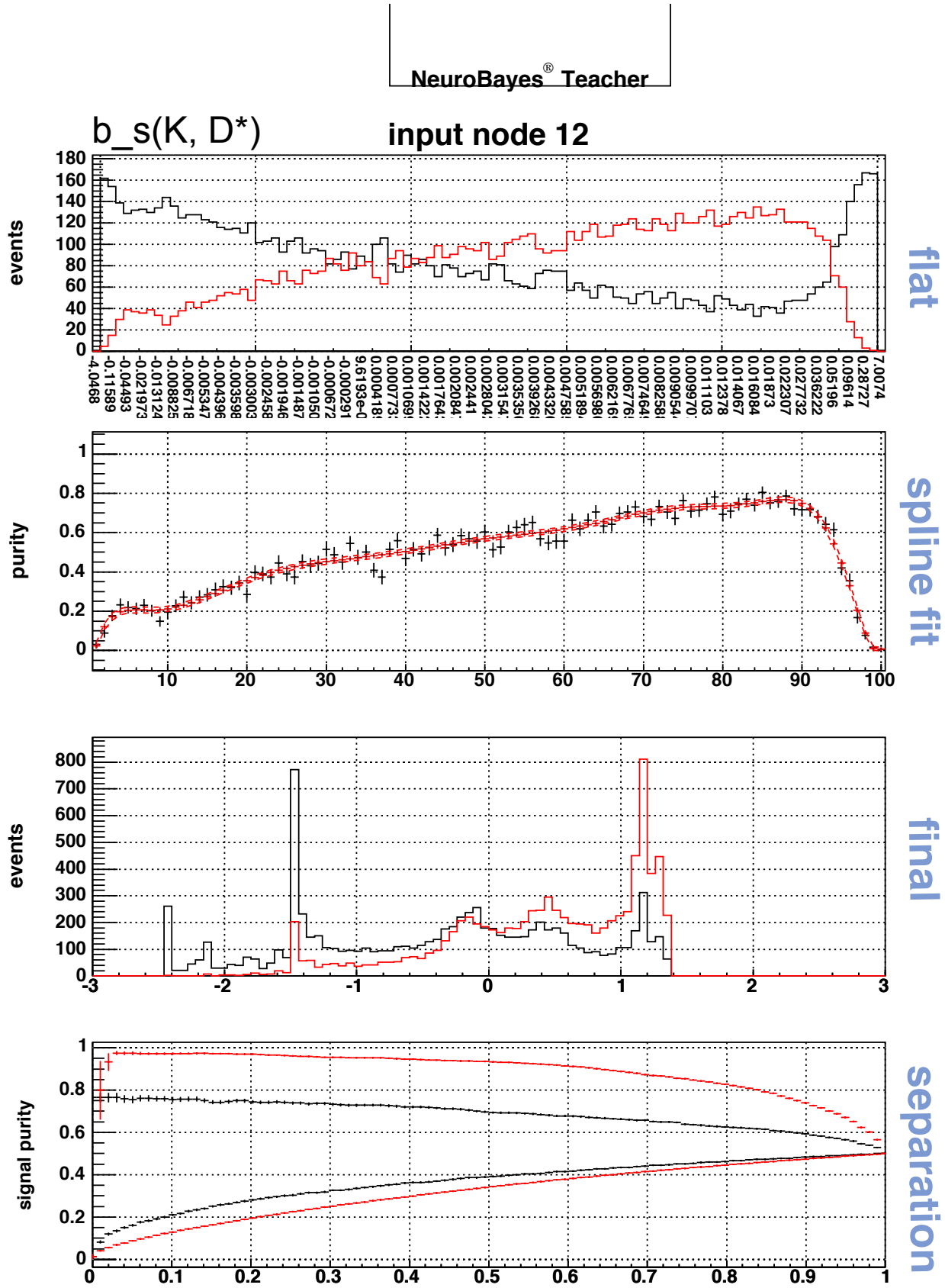

NeuroBayes® Teacher
D* Pt
input node 7
events
flat
purity
spline fit
events
final
signal purity
separation

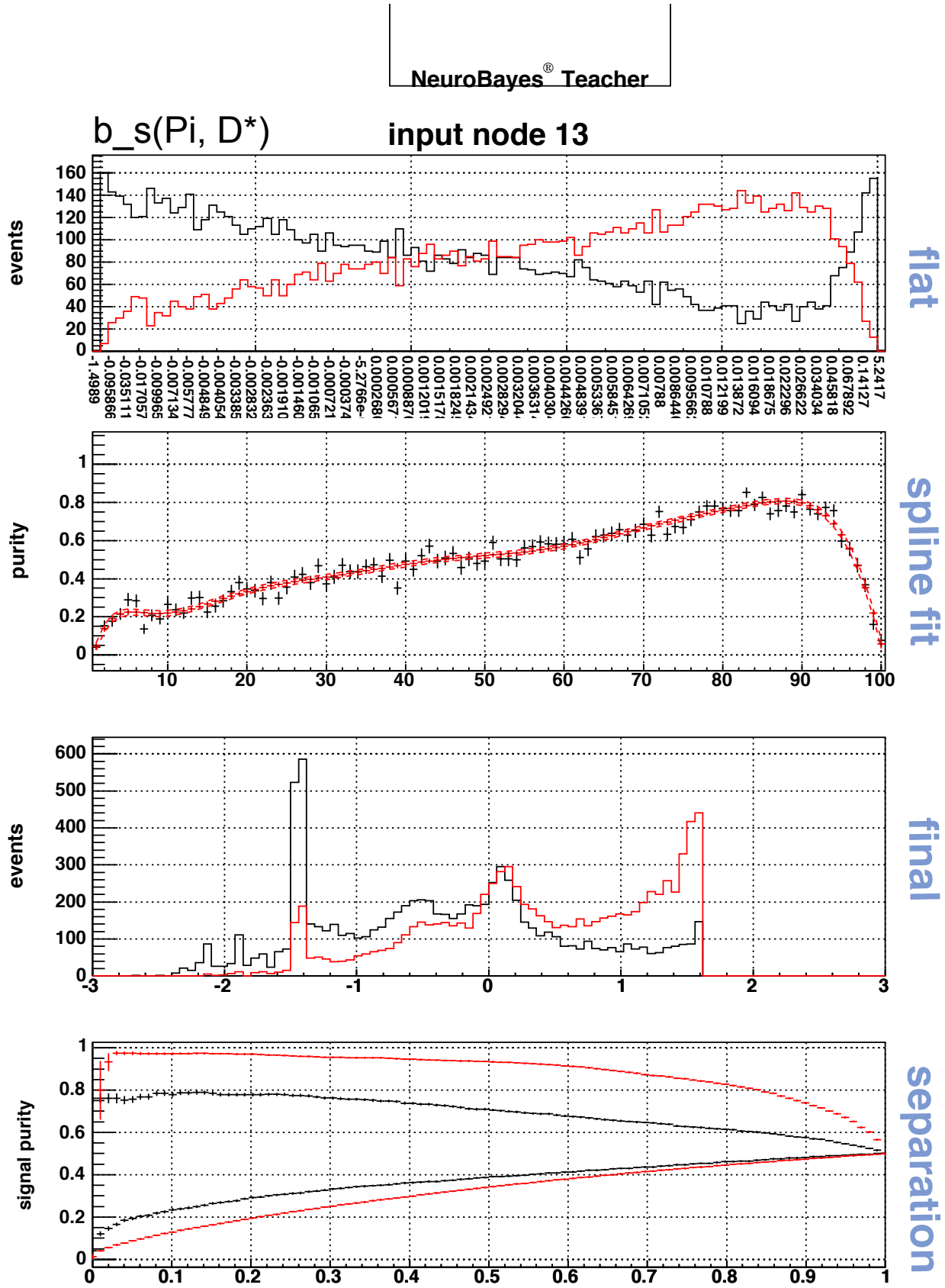

NeuroBayes® Teacher
D0 Pt
input node 8
events
flat
purity
spline fit
events
final
signal purity
separation

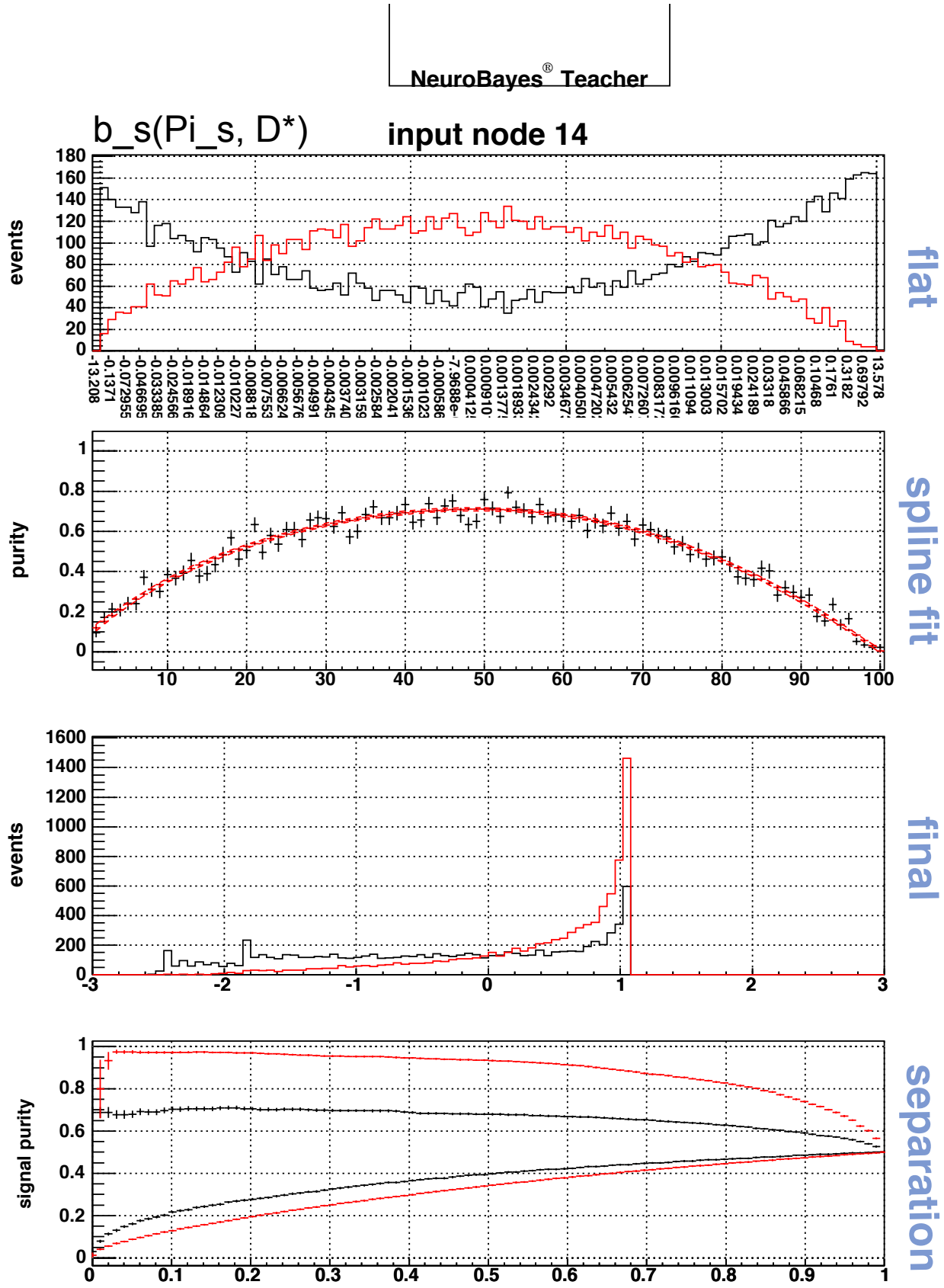

NeuroBayes® Teacher
DelR(K, Pi)
input node 9
events
flat
purity
spline fit
events
final
signal purity
separation

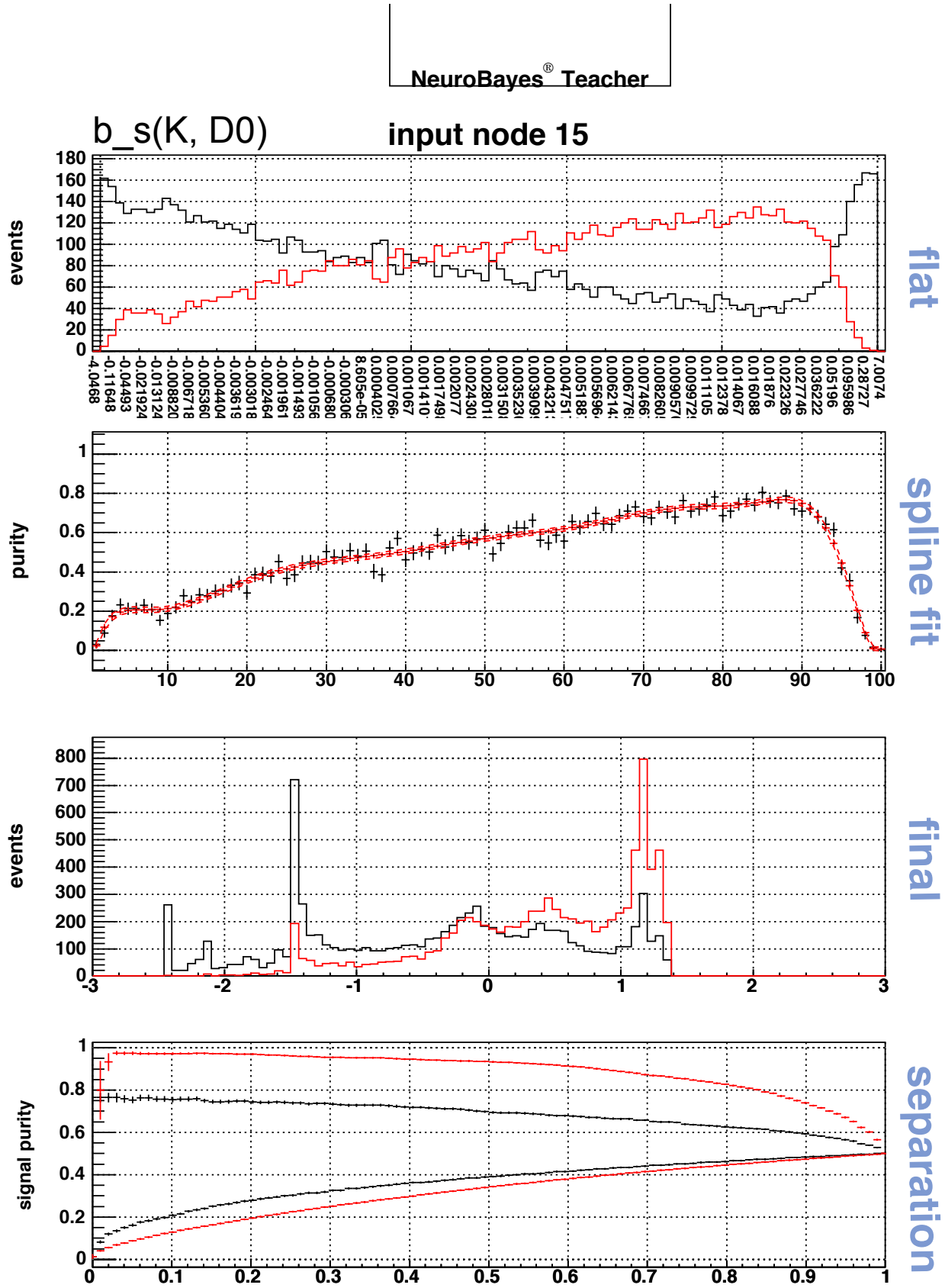

NeuroBayes® Teacher
DelR(K, Pi_s)
input node 10
events
flat
purity
spline fit
events
final
signal purity
separation

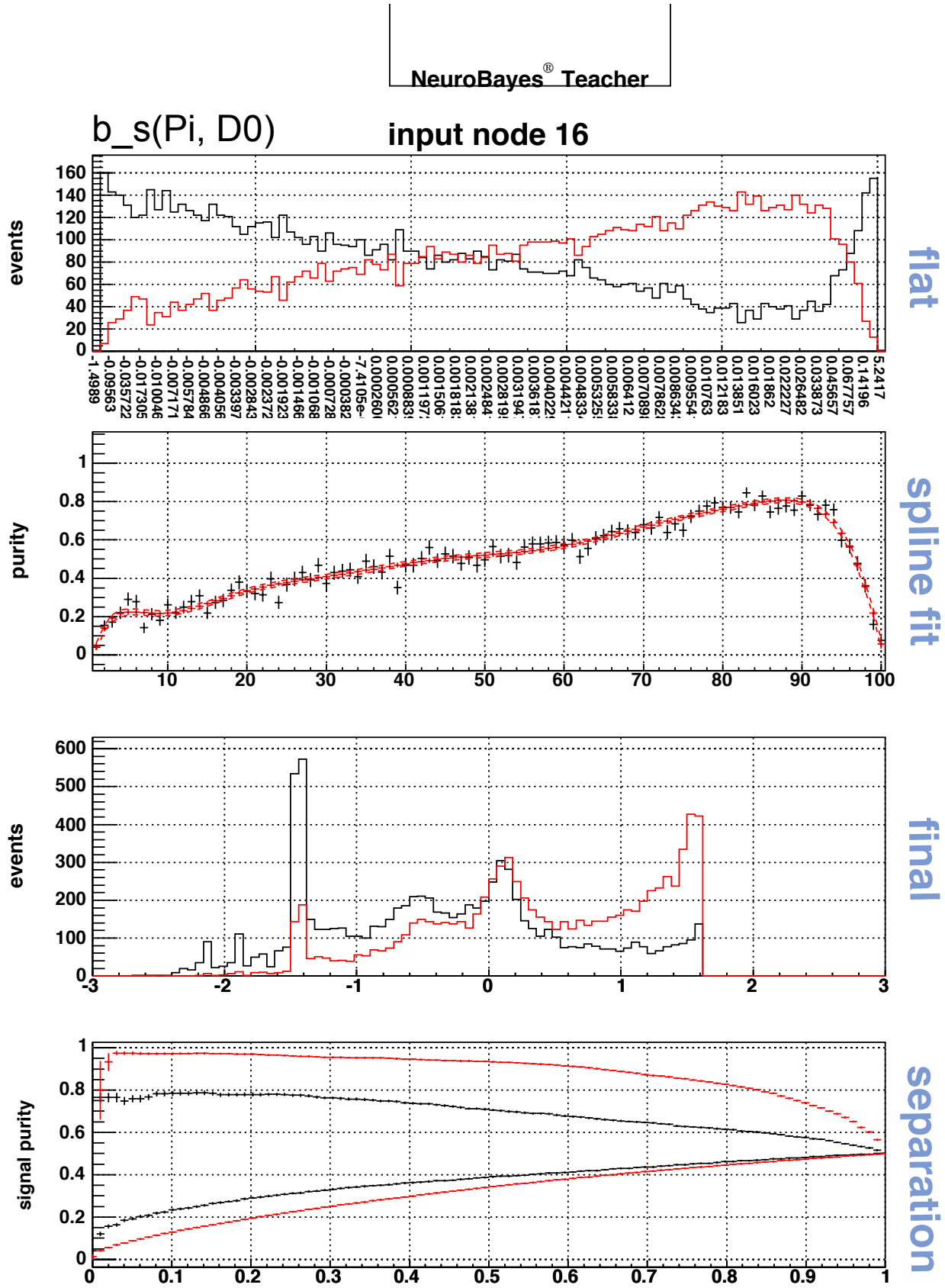

NeuroBayes® Teacher
DelR(Pi, Pi_s)
input node 11
events
flat
purity
spline fit
events
final
signal purity
separation

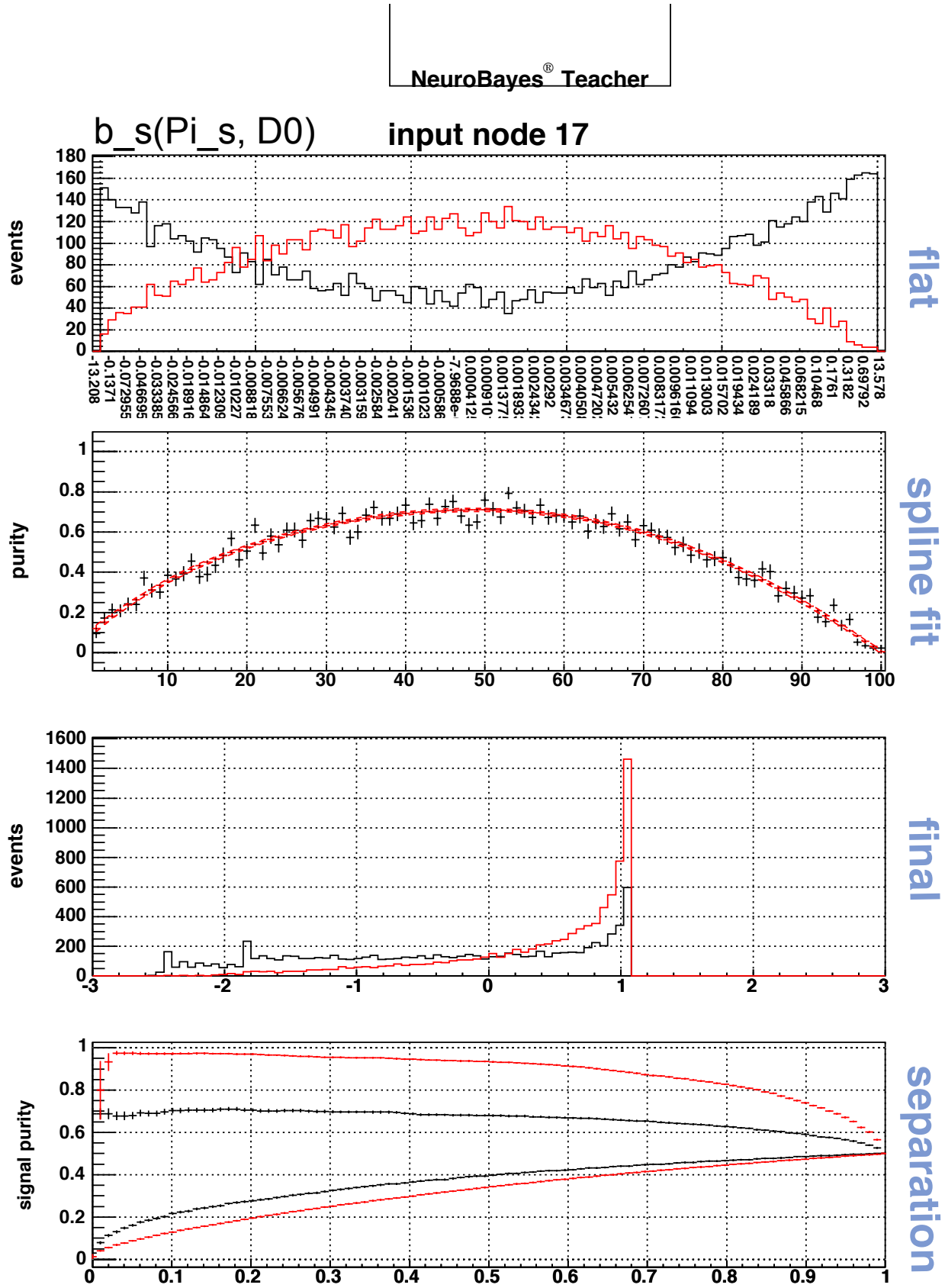

NeuroBayes® Teacher
b_s(K, D*)
input node 12
events
flat
purity
spline fit
events
final
signal purity
separation

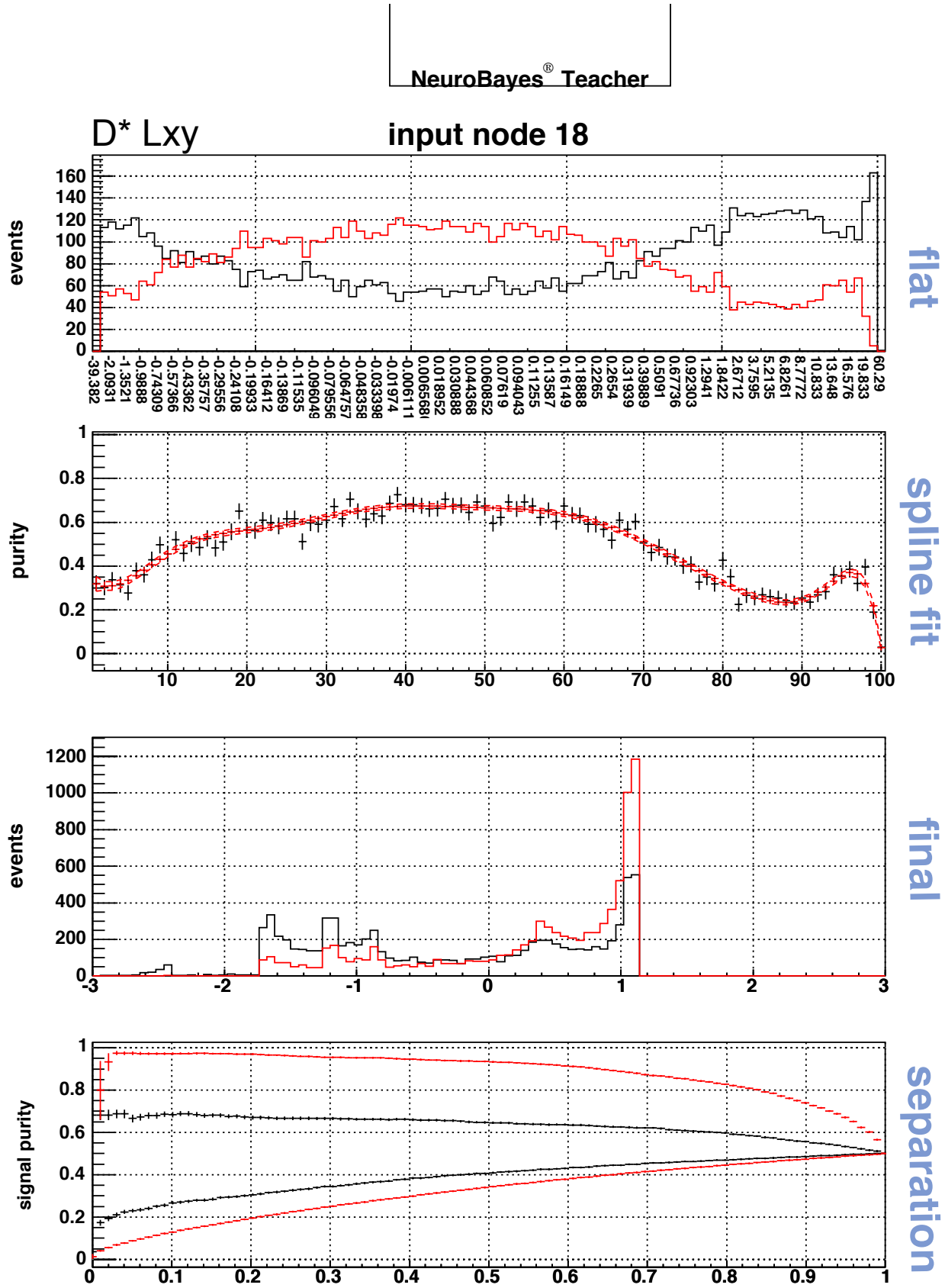

NeuroBayes® Teacher
b_s(Pi, D*)
input node 13
events
flat
purity
spline fit
events
final
signal purity
separation

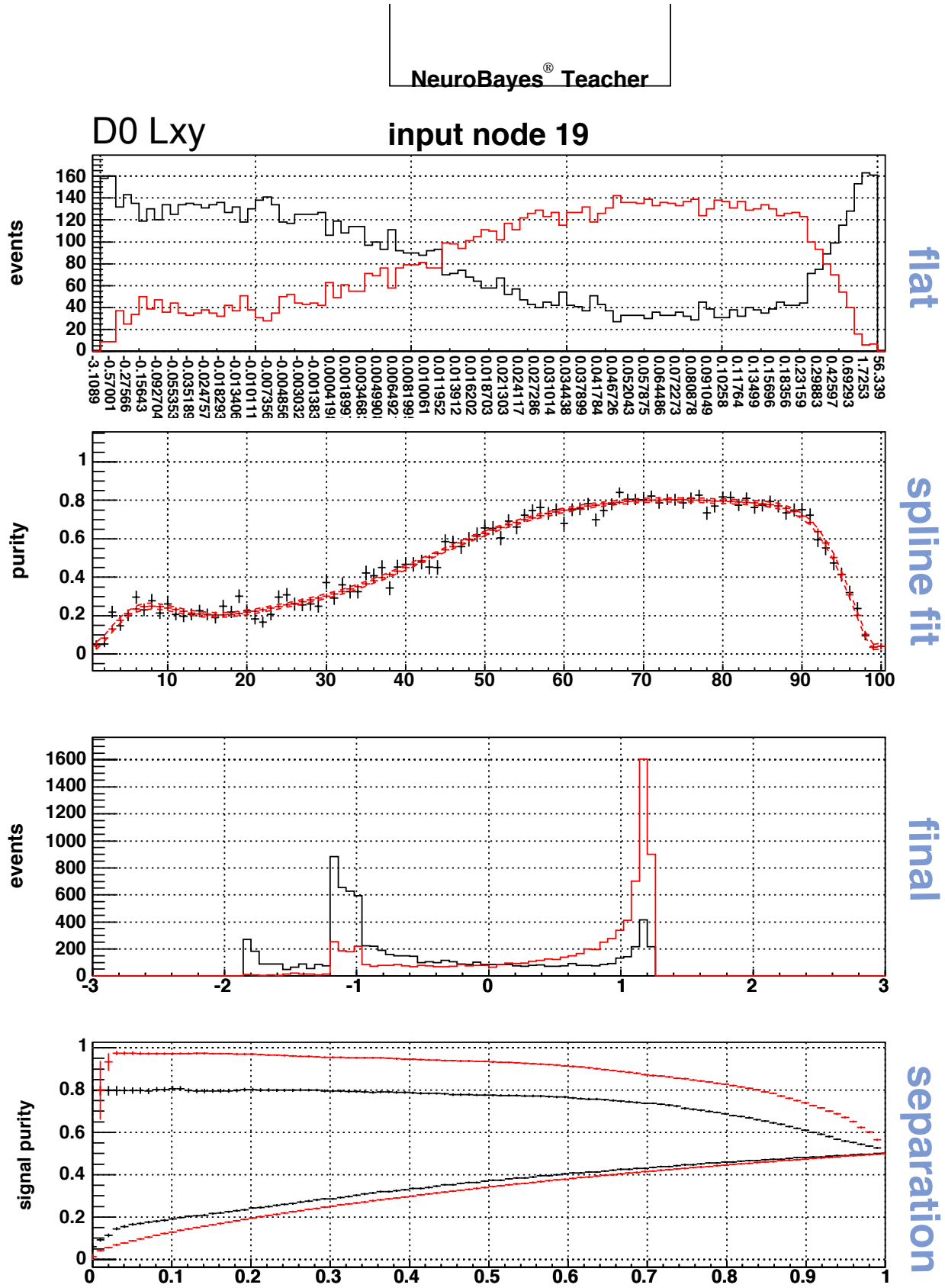

NeuroBayes® Teacher
b_s(Pi_s, D*)
input node 14
events
flat
purity
spline fit
events
final
signal purity
separation

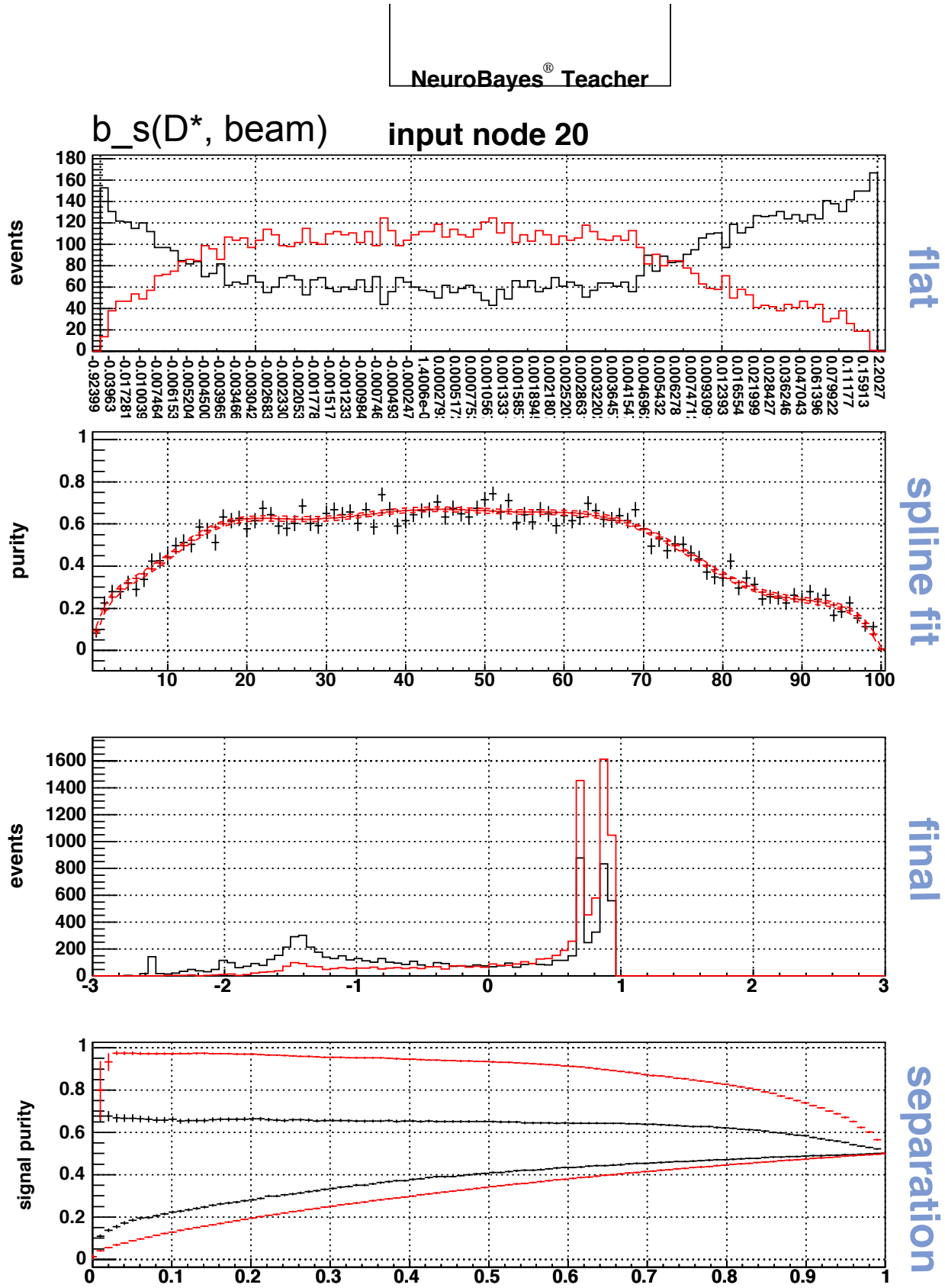

NeuroBayes® Teacher
b_s(K, D0)
input node 15
events
flat
purity
spline fit
events
final
signal purity
separation

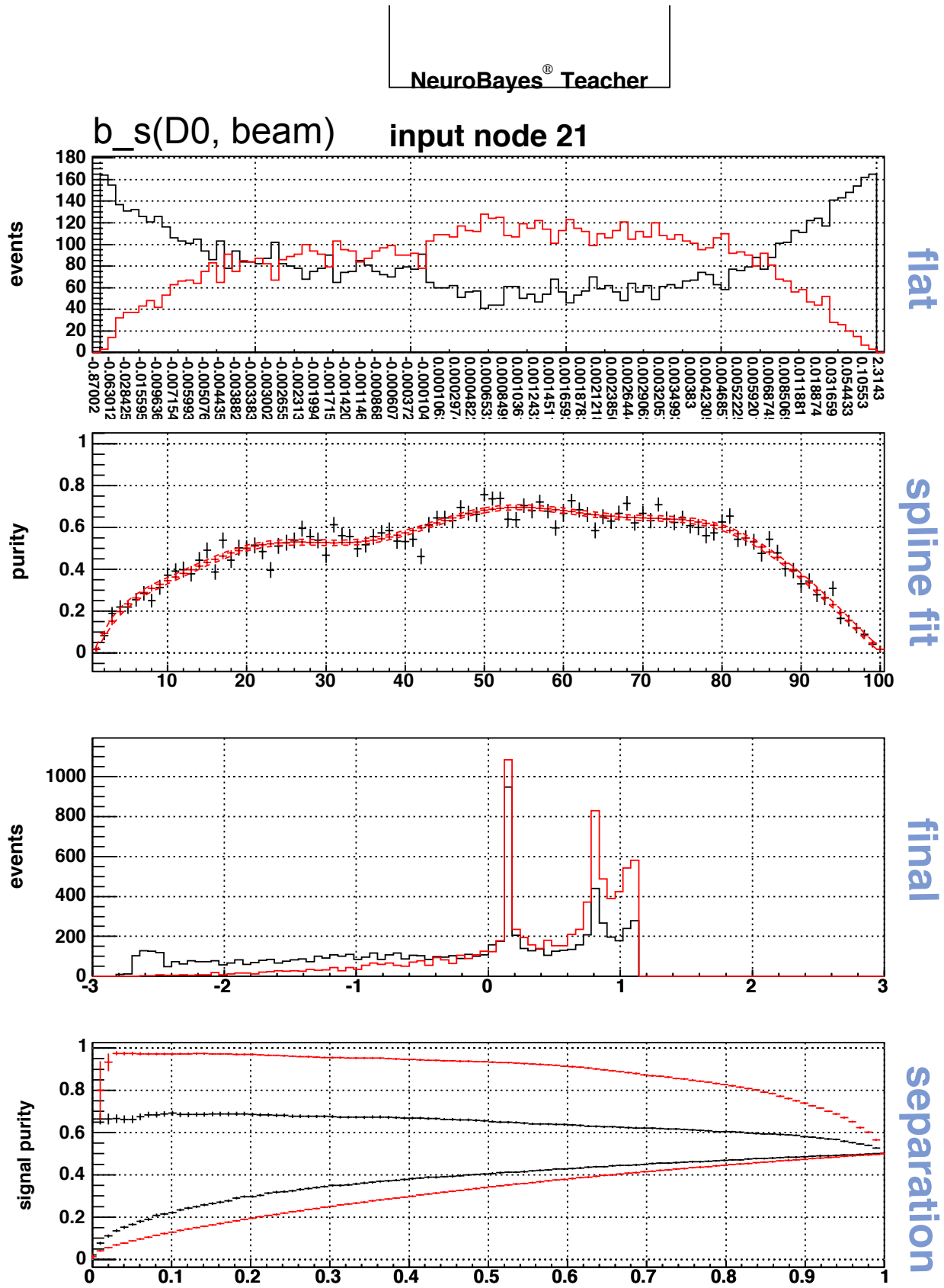

NeuroBayes® Teacher
b_s(Pi, D0)
input node 16
events
flat
purity
spline fit
events
final
signal purity
separation

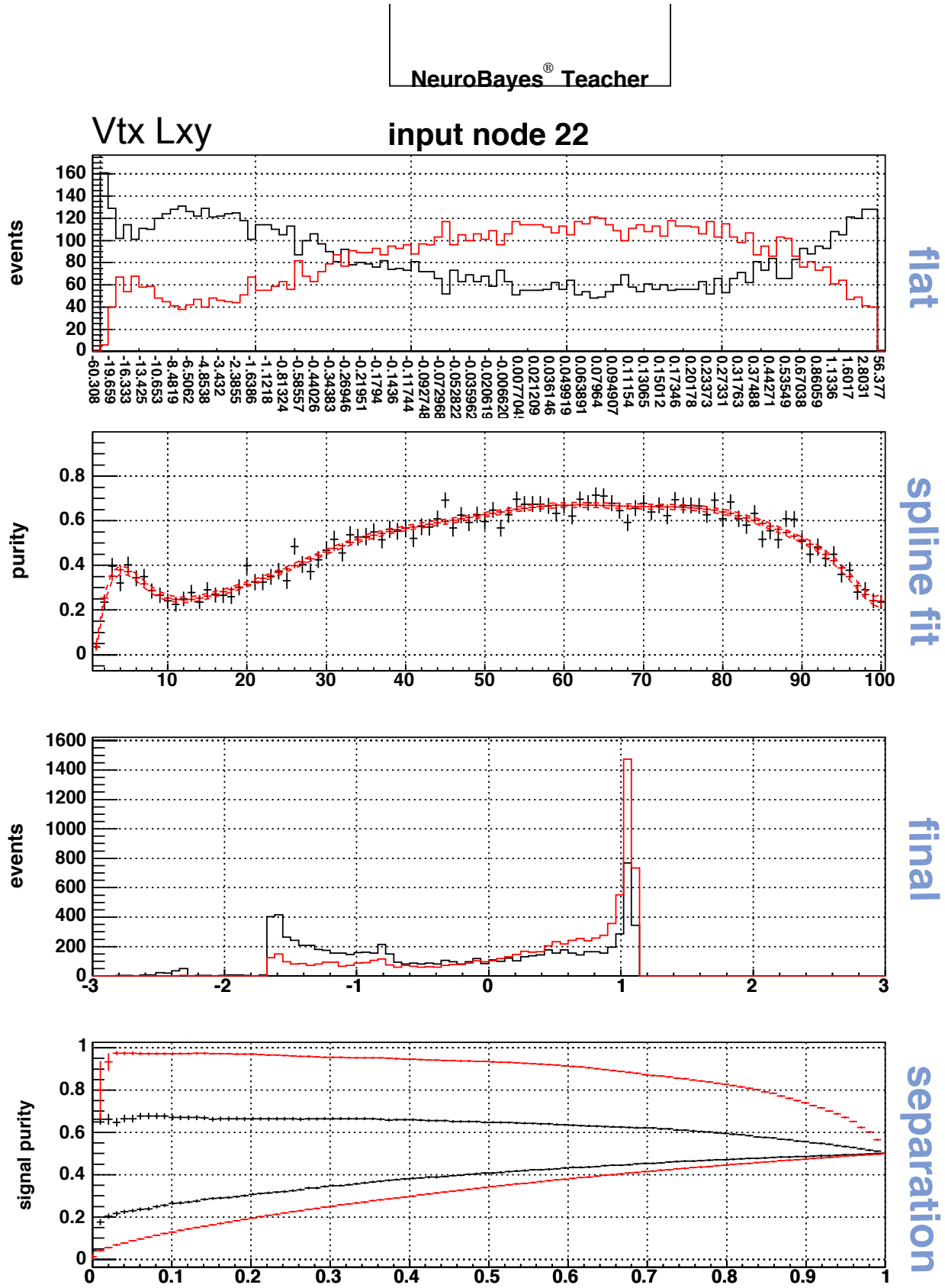

NeuroBayes® Teacher
b_s(Pi_s, D0)
input node 17
events
flat
purity
spline fit
events
final
signal purity
separation

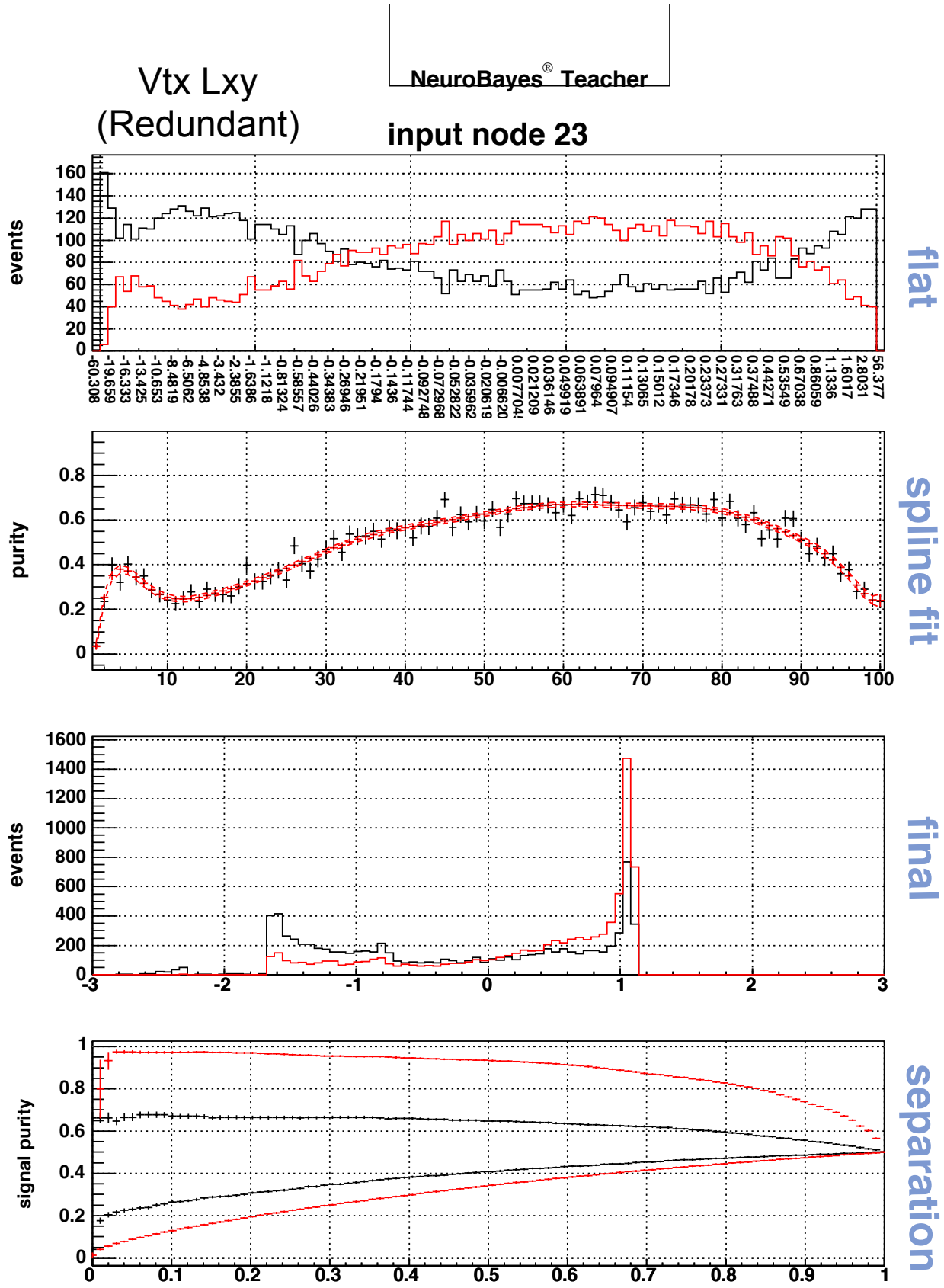

NeuroBayes® Teacher
D* Lxy
input node 18
events
flat
purity
spline fit
events
final
signal purity
separation


NeuroBayes® Teacher
D0 Lxy
input node 19
events
flat
purity
spline fit
events
final
signal purity
separation

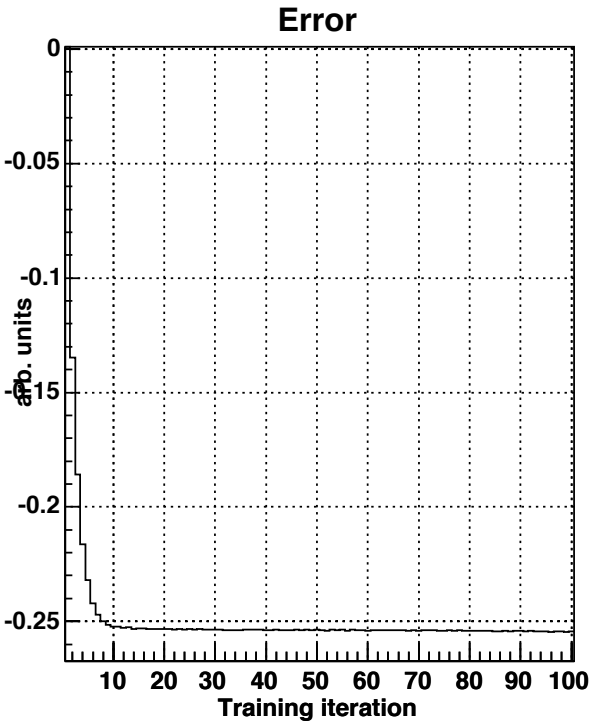

NeuroBayes® Teacher
b_s(D*, beam)
input node 20
events
flat
purity
spline fit
events
final
signal purity
separation

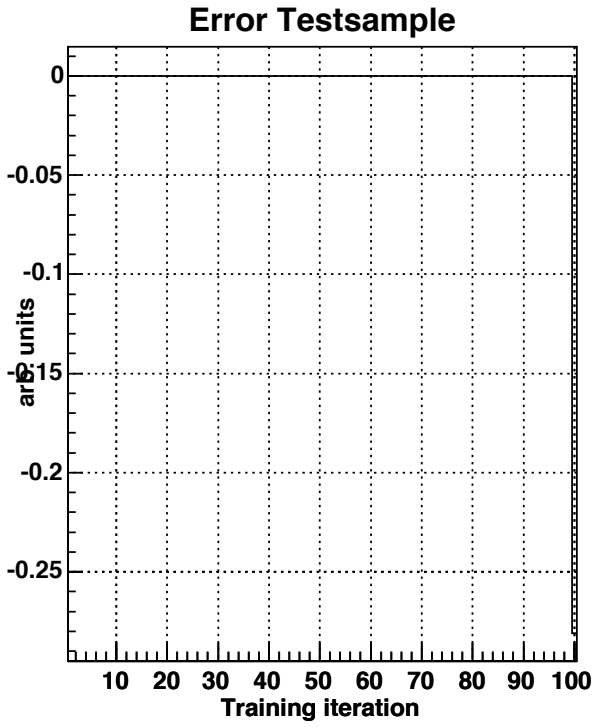

NeuroBayes® Teacher
b_s(D0, beam)
input node 21
events
flat
purity
spline fit
events
final
signal purity
separation

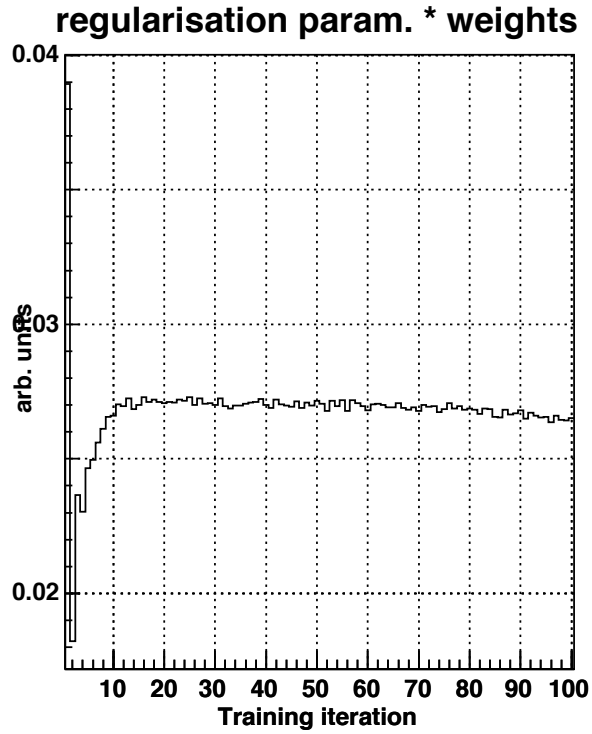

NeuroBayes® Teacher
Vtx Lxy
input node 22
events
flat
purity
spline fit
events
final
signal purity
separation

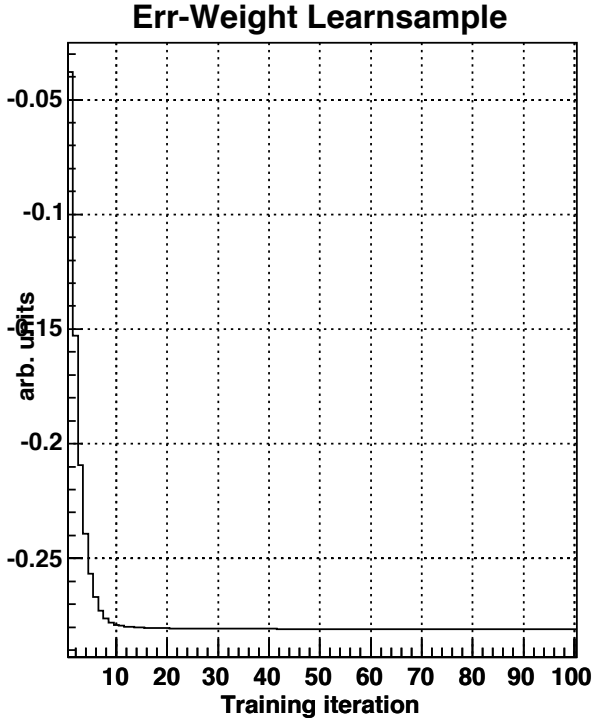

NeuroBayes® Teacher
Vtx Lxy
(Redundant)
input node 23
events
flat
purity
spline fit
events
final
signal purity
separation

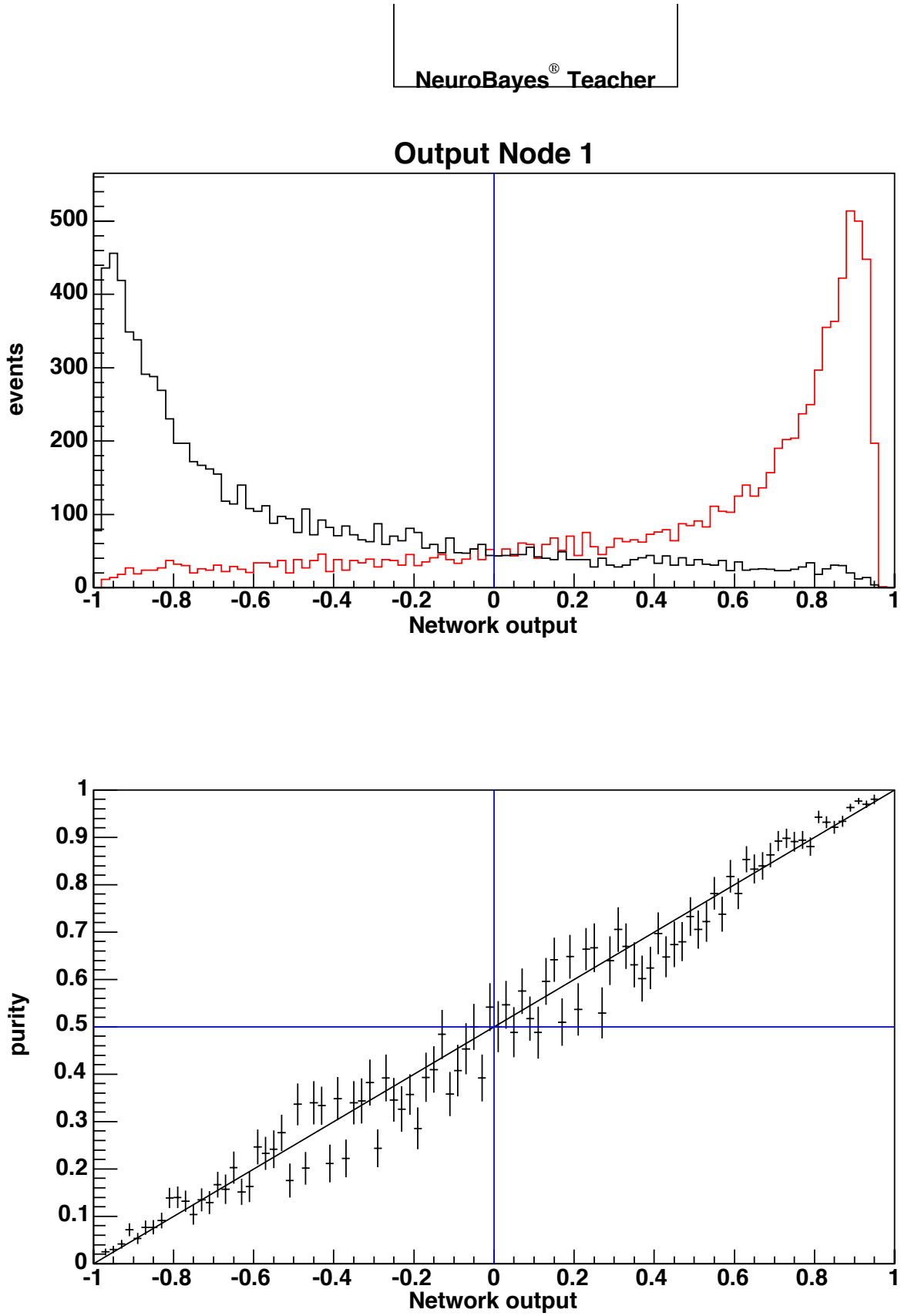
NeuroBayes® Teacher

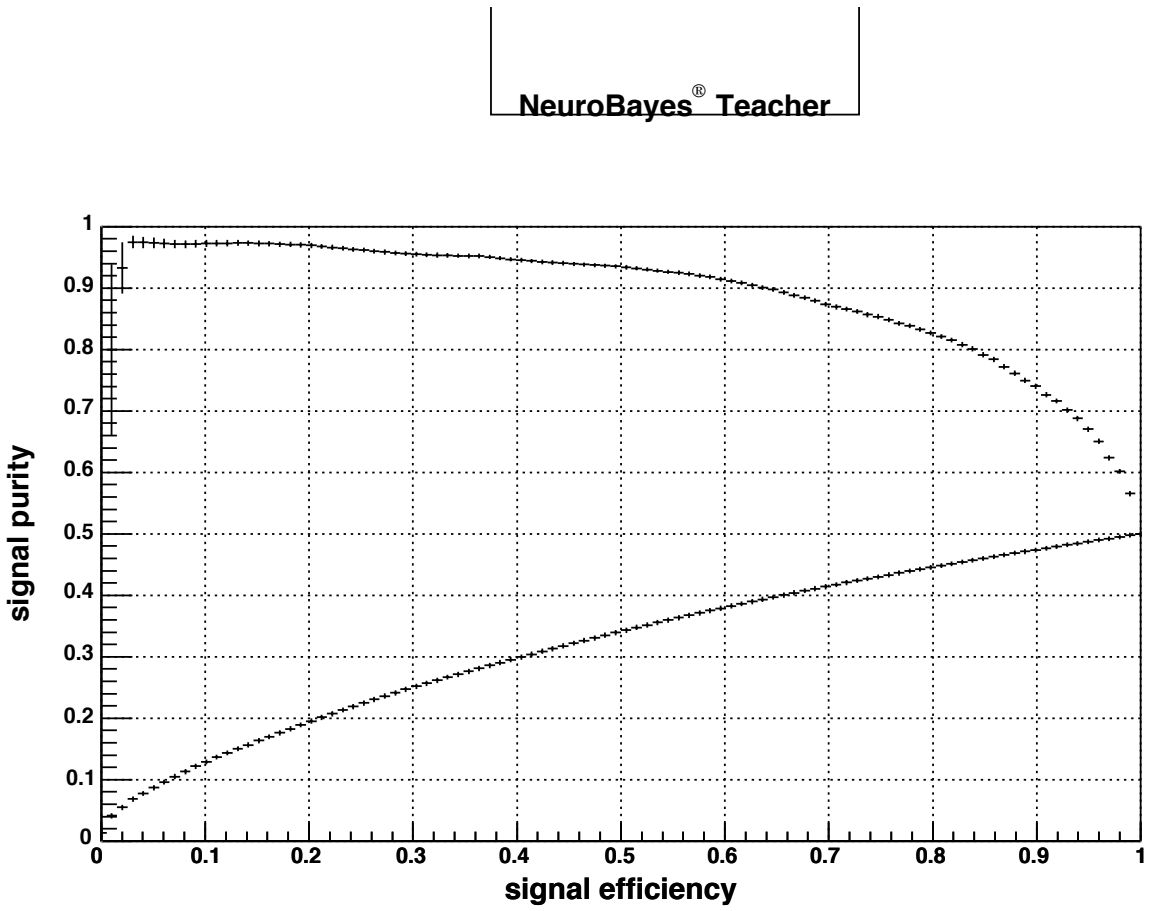


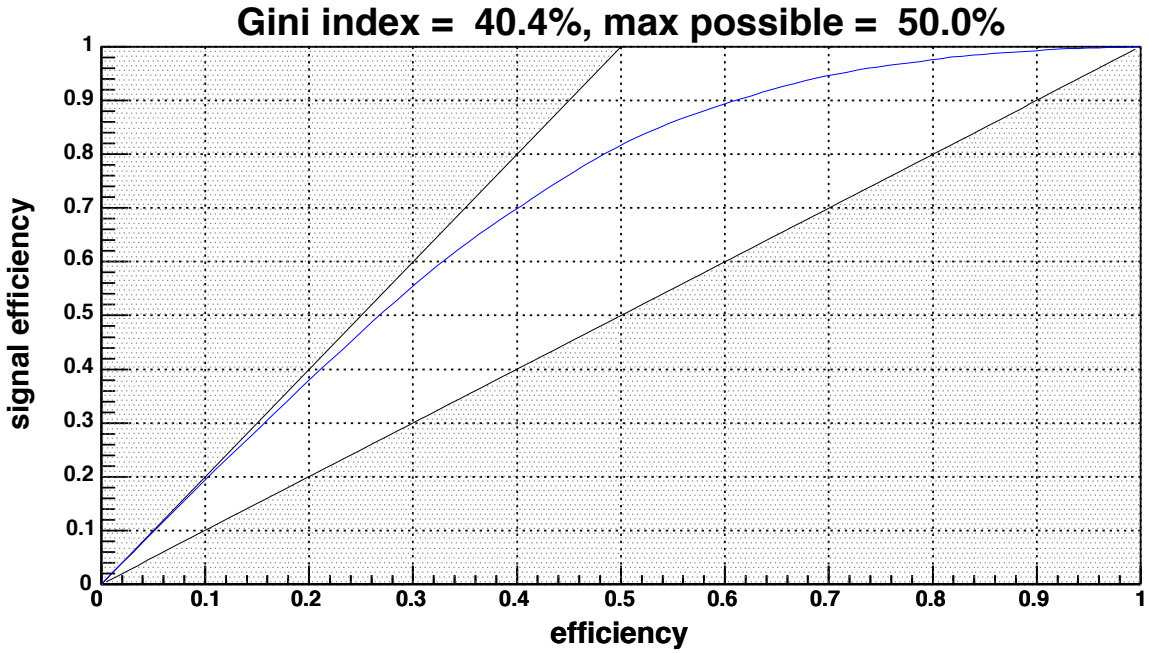




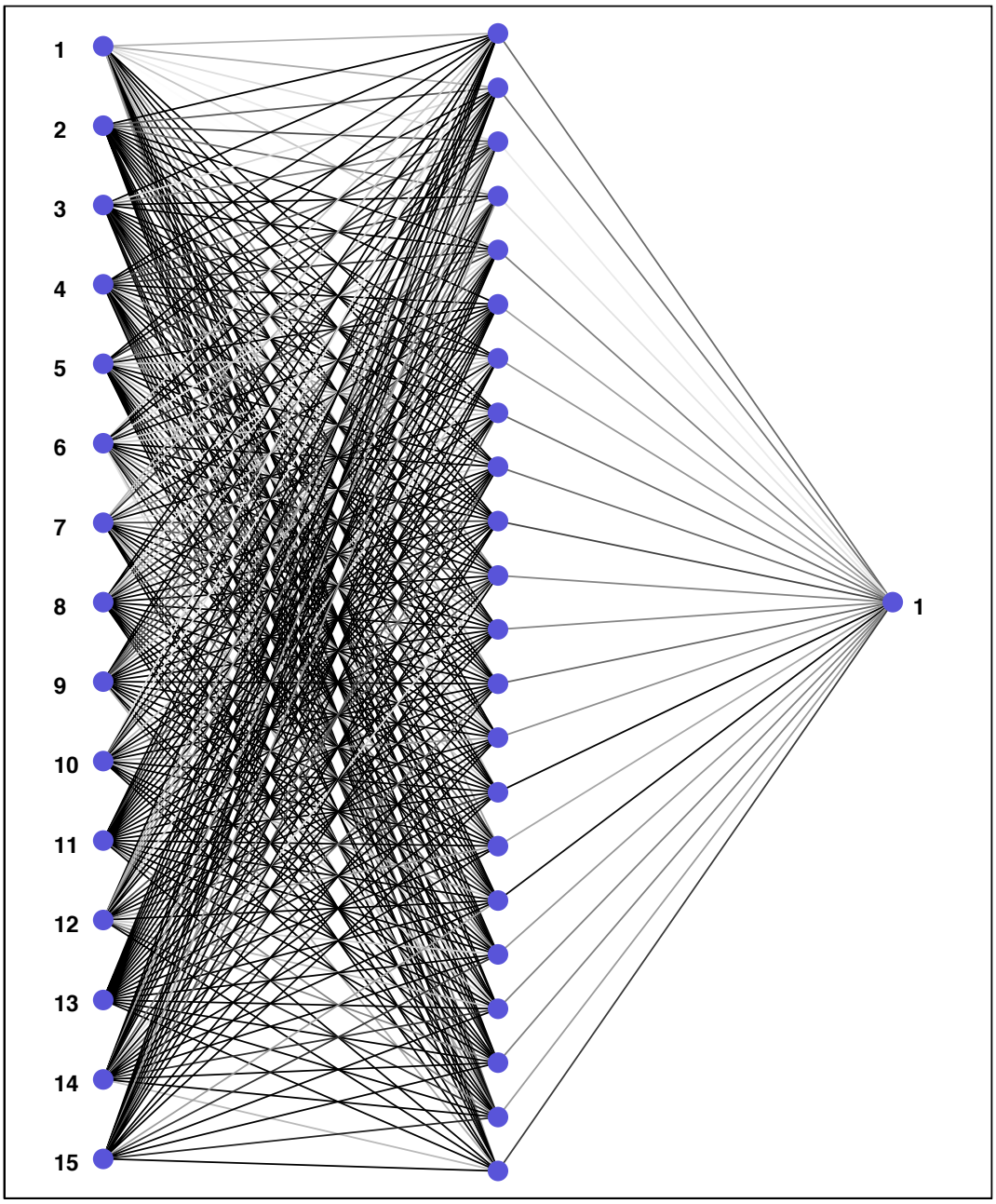



NeuroBayes® Teacher
Output Node 1
events
Network output
purity
Network output


NeuroBayes® Teacher
signal purity
signal efficiency
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1


Gini index = 40.4%, max possible = 50.0%
signal efficiency
efficiency
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

NeuroBayes® Teacher



References

[1] M.E. Peskin and D.V. Schroeder. *An Introduction to Quantum Field Theory.* Perseus Books, Reading, 1995.

[2] I. Aitchison and A. Hey. *Gauge Theories in Particle Physics, Volume I: From Relativistic Quantum Mechanics to QED.* Taylor and Francis Group, New York, 2003.

[3] I. Aitchison and A. Hey. *Gauge Theories in Particle Physics, Volume II: QCD and the Electroweak Theory.* Taylor and Francis Group, New York, 2004.

[4] G. Aad et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett. B*, 716, 2012.

[5] S. Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett. B*, 716, 2012.

[6] K. Wilson. Quarks, Gluon, and Lattices. *Phys. Rev. D*, 10(2445), 1974.

[7] S. Dawson P. Nason and R. Ellis. The One Particle Inclusive Differential Cross-Section for Heavy Quark Production in Hadronic Collisions. *Nucl. Phys. B*, 327(49), 1989.

[8] E.L. Berger. The puzzle of the bottom quark production cross section. *Int. J. Mod. Phys. A*, 18(1263), 2003.

[9] M. Cacciari and P. Nason. Is There a Significant Excess in Bottom Hadroproduction at the Tevatron . *Phys. Rev. Lett.*, 89(122003), 2002.

[10] S. Mrenna T. Sjostrand and P. Skands. Pythia 6.4 Physics and Manual. *JHEP*, 026(0605), 2006.

[11] G. Corcella et al. Herwig 6.5. *JHEP*, 010(0101), 2001.

[12] J. Pumplin et al. New generation of parton distributions with uncertainties from global QCD analysis. *JHEP*, 012(0207), 2002.

[13] A.D. Martin et al. Parton distributions for the LHC. *Eur. Phys. J. C*, 63(189-285), 2009.

[14] H.L. Lai et al. The Strange Parton Distribution of the Nucleon: Global Analysis and Applications. *JHEP*, 0704, 2007.

[15] U. Baur et al. The charm content of $W + 1$ jet events as a probe of the strange quark distribution function. *Phys. Lett.*, B318, 1993.

[16] D. Mason et al. Measurement of the Nucleon Strange-Antistrange Asymmetry at Next-to-Leading Order in QCD From NuTeV Dimuon Data. *Phys. Rev. Lett.*, 99(192001), 2007.

[17] M. Goncharov et al. Precise measurement of dimuon production cross-sections in muon neutrino Fe and anti-neutrino Fe deep inelastic scattering at the Tevatron. *Phys. Rev. D*, 64(112006), 2001.

[18] S. Klein S. Alekhin, J. Blumlein and S. Moch. The 3, 4, and 5-flavor NNLO Parton from Deep-Inelastic Scattering Data and at Hadron Colliders. *Phys. Rev. D*, 81(014032), 2010.

[19] R.S. Thorne A.D. Martin, W.J. Stirling and G. Watt. Parton distributions for the LHC. *Eur. Phys. J. C*, 63(189-285), 2009.

[20] R.D. Ball et al. A determination of parton distributions with faithful uncertainty estimates. *Nucl. Phys. B*, 809, 2009.

[21] H.L. Lai et al. New parton distributions for collider physics. *Phys. Rev. D*, 82(074024), 2010.

[22] G. Aad et al. Measurement of the production of a W boson in association with a charm quark in pp collisions at sqrt(s) = 7 TeV with the ATLAS detector . *JHEP (Submitted)*, 2014.

[23] T. Aaltonen et al. First measurement of the production of a W boson in association with a single charm quark in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV. *Phys. Rev. Lett.*, 100(091803), 2008.

[24] T. Aaltonen et al. Observation of the Production of a W Boson in Association with a Single Charm Quark. *Phys. Rev. Lett.*, 110(071801), 2013.

[25] V.M. Abazov et al. Measurement of associated production of Z bosons with charm quark jets in pp collisions at $\sqrt{s} = 1.96$ TeV .

[26] Proceedings for the Rencontres de Moriond EWK 2014. *Summary of single top quark production at the Tevatron*, 2014.

[27] V.M. Abazov et al. Search for the standard model Higgs boson in lv plus jets final states in 9.7 fb-1 of ppbar collisions with the D0 detector. *Phys. Rev. D*, 88(052008), 2013.

[28] A. Sill et al. CDF Run II silicon tracking projects. *Nucl. Instr. and Meth. A*, 447:1–8, 2000.

[29] T. Affolder et al. CDF Central Outer Tracker. *Nucl. Instr. and Meth. A*, 526:249–299, 2004.

[30] The CDF II Collaboration. CDF II Detector Technical Design Report. Technical Report FERMILAB-Pub-96/390-E, 1996.

[31] A. Artikov et al. Design and construction of new central and forward muon counters for CDF II. *Nucl. Instr. and Meth. A*, 538:358–371, 2004.

[32] Y.S. Chung et al. The Level-3 Trigger at the CDF Experiment at Tevatron Run II. Technical Report FERMILAB-Pub-01/031-E, 2001.

[33] International Conference on Computing in High Energy and Nuclear Physics (CHEP 2001). *The CDF Data Acquisition System for Tevatron Run II*, FERMILAB-CONF-01-242-E, 2001.

[34] B. Ashmanskas et al. The CDF Silicon Vertex Trigger. *Nucl. Instr. and Meth. A*, 518:532–536, 2004.

[35] A. Taffard. Run II Cosmic Ray Tagger, CDF Note 6100.

[36] T. Aaltonen et al. CDF Public Note 10952. June 2013.

[37] T. Aaltonen et al. CDF Public Note 11056. November 2013.

[38] T. Aaltonen et al. Precise Measurement of the W-Boson Mass with the CDF II Detector. *Phys. Rev. Lett.*, 108(151803), 2012.

[39] A Abulencia et al. Measurements of inclusive W and Z cross sections in pp collisions at $\sqrt{s} = 1.96$ TeV . *J. Phys. G: Nucl. Part. Phys.*, 34(12), 2007.

[40] E. Gross G. Cowan, K. Cranmer and O. Vitells. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J.*, C71(1554), 2010.

[41] M. Feindt and U. Kerzel. The NeuroBayes neural network package. *Nucl. Instr. and Meth. A*, 559, 2006.

[42] J. Beringer and others (Particle Data Group). 2012 Review of Particle Physics (and 2013 partial update for the 2014 edition). *Phys. Rev. D*, 86(010001), 2012.

[43] J.W. Huston J.M. Campbell and W.J. Sterling. Hard interactions of quarks and gluons: a primer for LHC physics. *Rep. Prog. Phys.*, 70(89), 2007.

[44] S. Dubnicka et al. Quark model description of the tetraquark state X(3872) in a relativistic constituent quark model with infrared confinement. *Phys.Rev.D*, 81(114007), 2010.

[45] The 6th International Workshop on Charm Physics (CHARM 2013). *X(4260) as a Mixed Charmonium-Tetraquark State*, 2013.