Study of Charged Particle Species Produced in Association with $\bar{B}^0$, B^-, and $\bar{B}_s^0$ Mesons in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV.

Denys Usynin

A DISSERTATION
in
Physics and Astronomy

Presented to the Faculties of the University of Pennsylvania in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

2005

Joseph Kroll, Advisor
Supervisor of Dissertation

Andrea Liu
Graduate Group Chairperson

Dedication

This dissertation is dedicated to little Maria, who is no trouble at all.

Acknowledgments

There are many people who have directly contributed to this dissertation, and many more who contributed indirectly by helping me or influencing me in one way or another. Thanking everyone would be the proper and impossible thing to do, but some people have more than just deserved thanks, they shaped me as a scientist and a person. First of all, my adviser, Joseph Kroll, to whom I am grateful for teaching me to seek the meaning of all things, and Matthew Jones, with whom we worked together for many years and who inspired me to never stop at "good enough" but strive for "excellent". They both were more than just my guides and collegues in Physics, but have in many ways been my guides in life during the years of graduate school.

I am very grateful to MIT group at CDF, especially Guillelmo Gomez-Ceballos, Christoph Paus, Stephanie Menzemer and Ilya Kravchenko, for their help and advice. So many results of their work were used either in my research that it would not be fair to mention their contribution only in the Bibliography.

Barry Wicklund has always been a source of interesting ideas and insights, and

a lot of the research that I have done been a direct result of his suggestions.

I thank Gino Segre and Randall Kamien who taught me graduate courses at the University and managed to not only teach me their subject matter but changed the entire way I think and look at myself as a scientist.

The community at CDF has always been there for me to provide true scientific environment, whatever that means. I am however mostly grateful to them for making the years spent at Fermilab pleasant, which is mostly due to them being very nice, friendly and fun to be around people. We had many scientific and non-scientific discussions with (in no particular order) Chris Neu, Andrew Kovalev, Aart Heijboer, Boris Iyutin, Konstantin Anikeev, Shin-Shan (Eiko) Yu, Joel Heinrich, and others. Our professors Nigel Lockyer, Brig Williams, Evelyn Thompson and Joseph Kroll have contributed tremendously to the feeling of healthy community within the University of Pennsylvania group at Fermilab.

I would like to thank the University of Pennsylvania and United States Department of Energy for providing the funding for this work. I also thank the Fermilab stuff, the CDF Collaboration, and everyone else who contributed to building the laboratory and the CDF experiment - without their effort this dissertation would never come to life.

I remember with gratitude my high school teachers, especially Elena Elnikova, who showed me that studying can be fun and not a chore, and that Physics is mysterious and intriguing, which I still feel is true.

I thank my parents for bringing me up the way I am, for always believing in me and for always encouraging me to do what I feel is right and study what I think is interesting. I thank my sister for always supporting me; she may not realize how much her support meant to me, but if all good things come back to us she will live long and happy.

Finally, I wholeheartedly thank my wife Valida for being with me through all the difficult times and turning them into much easier ones by the sheer force of her love. And last but not least of all, I thank our little daughter Maria who has been very, very patient with us all this time.

ABSTRACT

Study of Charged Particle Species Produced in Association with $\bar{B}^0$, B^-, and $\bar{B}_s^0$ Mesons in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96\,\text{TeV}$.

Denys Usynin

Joseph Kroll, Advisor

We study the yields of charged kaons, charged pions, and protons produced in a association with B mesons produced in proton-antiproton collisions at center of mass energy 1960 GeV using 355 pb^{-1} of data collected with the CDF detector at the Fermilab Tevatron. This is the first reported measurements of these yields at a hadron collider. The B mesons are reconstructed using their semileptonic decays:

$B^0 \to l^+ D^- X, D^- \to K^+\pi^-\pi^-$;

$B^0 \to l^+ D^{*-} X, D^{*-} \to \pi^- \bar{D}^0, \bar{D}^0 \to K^+\pi^-$;

$B^+ \to l^+ \bar{D}^0 X, \bar{D}^0 \to K^+\pi^-$;

$B_s \to l^+ D_s^- X, D_s^- \to \pi^-\phi, \phi \to K^+K^-$.

The K, π, and p are identified using the Time of Flight detector (TOF), the CDF spectrometer, and the specific ionization (dE/dx) measured in the central drift chamber (COT). The fraction of charged kaons produced in association with $\bar{B}_s^0$ mesons is found to be larger than the fraction produced in association with the

$\bar{B}^0$ and B^- mesons, as expected from naive models of heavy quark hadronization to mesons. The particle species yields are found to be in qualitative agreement with simulation of B meson production in hadron collisions from the `PYTHIA` Monte Carlo, although the yield of kaons around $\bar{B}_s^0$ mesons is found to be larger in the simulation when compared to the data. These studies are important for understanding methods of identifying the flavor of $\bar{B}_s^0$ mesons in measurements of $\bar{B}_s^0$ flavor oscillations and charge conjugation-parity (CP) violation in $\bar{B}_s^0$ meson decays.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

This dissertation presents a study of the production of the charged particles associated with B meson formation. B mesons are a colorless bound state of a b quark and a light anti-quark. The standard convention for naming the weakly decaying B mesons is $\bar{B^0} = b\bar{d}, B^- = b\bar{u}, \bar{B_s^0} = b\bar{s}$. We measure the fractional composition of the charged particles produced in association with $\bar{B^0}, B^-, \bar{B_s^0}$ mesons in terms of the particle species π, K, p, e, and also the yields at which these particles are produced. The yield of a particle of given species is defined as the average number of particles of that type found within a certain kinematic region around a reconstructed B meson. We measure both the absolute yields of all particles of given type within the detectable momentum range and the momentum-dependent yields of these particles. The fractional composition is defined as a ratio of the yield of given particle type to the combined yield of all charged particles.

The first part of this Introduction is devoted to giving the motivation for performing the analysis presented here. It briefly touches on the Physics concepts (and arising complications) involved in these studies and thus introduces the reader to Chapter 2 where the relevant Physics concepts are described in more detail.

The second part of the Introduction gives a structural overview of the entire analysis making it easier for the reader to navigate the document.

1.1 Motivation

While the measurements of particle species production in association with B mesons formation are quite interesting by themselves, being the first of their kind, the main motivation for performing this research at CDF is to provide information and help development of the tools for studying the mixing of the $\bar{B}_s^0$ meson. The mixing is a process by which a particle can undergo transitions into its own anti-particle and back. It was first discovered in the neutral kaon system, and more recently in the $\bar{B}^0$ and $\bar{B}_s^0$ systems. Measuring the mixing parameters of the $\bar{B}_s^0$ mesons is one of the most important current goals of particle physics. Knowing the mixing frequency of $\bar{B}_d^0$ and $\bar{B}_s^0$ mesons would allow us to improve our knowledge of the weak interactions by measuring the elements of Cabbibo-Kobayashi-Maskawa (CKM) matrix, and would facilitate a test of CKM matrix unitarity and potentially probe for new physics beyond the Standard Model. A more detailed review of these topics and the reasons for the intense interest in them is given in Chapter 2.

Observing the mixing of B mesons requires being able to determine the flavor of the B meson at the point of production as well as the point of decay. The latter is usually known after reconstructing the decay, and the former has to be inferred by one of the indirect techniques collectively called tagging methods (algorithms). One of the techniques for extracting the B production flavor is called same-side tagging (SST), and, in the case of $\bar{B}_s^0$ meson, this technique is frequently referred to as same-side kaon tagging (SSKT) for reasons which are explained below.

The basic idea behind the SST is based on our understanding of the process of the B meson formation. A naive illustration of this process is presented in Figure 1.1.

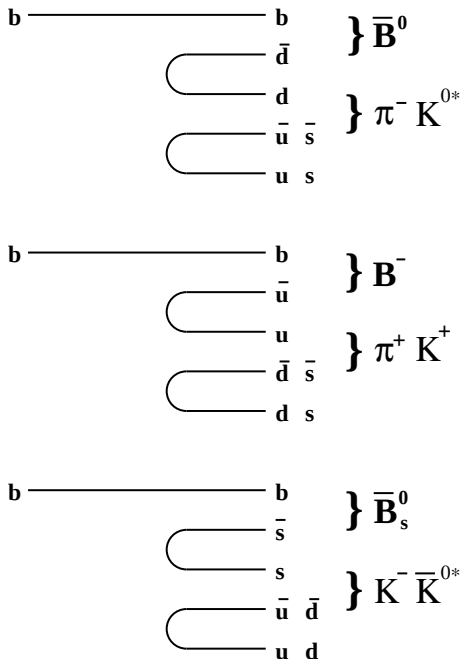


Figure 1.1: Naive description of the formation of B mesons.

The b quarks are produced in a $p - \bar{p}$ collision and then form a B meson by coupling with one of the lighter quark species. This process is called fragmentation and is believed to be dominated by the forces of strong interactions, which conserve the quark flavor. Therefore the quarks are created as quark-antiquark pairs during the fragmentation process, with lighter quarks more likely to be produced than heavier quarks since it requires less energy to create them.

Based on this qualitative description of the b fragmentation, one can assume that if, for example, a $\bar{B}_s^0$ meson was found in an event, then an $s\bar{s}$ pair was created during fragmentation. One of these s quarks combines with a b quark to form a B meson and the other may form a K meson (giving rise to the name of the kaon tag). Similarly, pions are more likely to have been produced in an event if a $\bar{B}^0$ or B^- meson is found, since in these events a light quark remains from B hadronization and it is more likely to couple to another light quark than a heavier s quark. In addition, the B^- meson will often be charge-correlated with nearby fragmentation products, as shown in Figure 1.1.

Same-side tagging algorithms attempt to identify the particles associated with b fragmentation and infer the B meson flavor from them. It is usually assumed that the directions of the momenta of the particles created from the same quark pair are correlated, so that the particles are selected in a cone[1] around the B meson being tagged (same-side B meson). This assumption is supported by predictions of some

[1]This is not a usual geometric cone but rather a kinematic selection that is invariant under Lorentz boost along the detector axis in a highly relativistic limit.

fragmentation models as well as experimental evidence.

Applying an SST algorithm to a reconstructed B^- candidate to extract its flavor at the creation point one can immediately test for correlations between the charge of the particles associated with B production and the b quark flavor, since the B^- does not mix and the decay flavor is the same as the production flavor. The mixing parameters of the $\bar{B}^0$ meson are well known and the correlations of the SST in this case could also be established experimentally from the observation of the flavor mixing itself.

Another complication that one has to deal with when studying the same-side tagging of B mesons is that a b quark may first form an excited B^{**} state which then strongly decays into a B particle that is reconstructed. In case of the semileptonic decays the B mesons themselves may also have unreconstructed intermediate excited D states. This means that particles selected around B meson come from the B^{**} and $D^{*(*)}$ decays as well as from quark fragmentation. There is ample evidence that decay products of B^{**} and $D^{*(*)}$ mesons make large contribution to the same-side tags. The analysis presented here is also complicated by the presence of these secondary particles.

The measurements of the SST performance done for one B meson species do not apply to any other type of B mesons due to the differences in fragmentation and decay signatures of the B^{**} mesons. The challenge therefore lies with studying the tagging of the $\bar{B}_s^0$ mesons, which oscillate very rapidly, preventing a direct

measurement of the mixing frequency so far. The two possibilities for studying and developing the SSKT are either to observe its correlation with the other tagging techniques[2], or to rely on a Monte-Carlo simulation of the quark fragmentation process. The former is very difficult due to limited statistical power, both because the sample of $\bar{B}_s^0$ mesons collected by CDF so far is relatively small, and because the other tagging algorithms have limited power. It currently does not seem possible that enough data will be gathered by CDF in the foreseeable future to provide sufficient data for these studies. Therefore, unless we are able to observe the flavor oscillations of $\bar{B}_s^0$ directly, we have to rely on the Monte-Carlo techniques to study the SST algorithms at CDF.

The Monte-Carlo approach however also faces a difficulty, which is that it's not known how well the fragmentation is simulated with respect to the types of the particles produced. The Monte-Carlo simulation of the b quark production and creation of the particles associated with that process was tuned on parameters such as particle yields and momentum spectra. This analysis will help fill a gap in our experimental knowledge by identifying charged particles produced in association with the B mesons. We measure the yields and fractions of the individual particle species in a cone around B decay. We identify the particle type by using timing informa-

[2]The other main type of tagging is called opposite side tagging (OST). Since b quarks are produced as $b\bar{b}$ pairs in strong interactions, OST tags attempt to identify the flavor of the other b quark and assume that it is opposite to the flavor of the b quark of interest. A detailed discussion of OST and other tagging methods is outside the scope of this document.

tion from the Time-of-Flight (TOF) detector and particle energy loss information (dE/dx) from the Central Outer Tracker (COT) detector.

In order to achieve these goals the analysis has to deal with many of the issues outlined above. Because of the limited size of the $\bar{B}_s^0$ samples collected at CDF so far, we perform the analysis on a sample of semileptonic B decays, which provide the highest number of reconstructed $\bar{B}_s^0$ mesons. We also try to minimize the effect of the secondary particles (such as $D^{*(*)}$ decay products) on our results.

We compare our measurements with the prediction of the `Pythia` [46] Monte-Carlo generator. This will help us decide whether we can rely on Monte-Carlo for predicting the performance of the SSKT. These measurements may also help estimate how well the models used by Monte-Carlo generators predict the relative abundance of the particles produced in association with the B mesons.

Finally, we compare the kaon production around different B meson species. As described above, we expect more kaons to be produced around $\bar{B}_s^0$ mesons than B^- or $\bar{B}^0$ mesons. However, quantifying this statement and the comparison itself is not a trivial matter, and is discussed further in this dissertation.

1.2 Overview of the analysis procedure

This analysis would not be possible without the experimental apparatus, the CDF II detector. The design, construction and operation of the CDF II detector should be rightfully thought of as the first stage of this analysis. Many people, including

the author of this dissertation, have committed countless hours to these activities allowing many interesting analyses to be produced. Chapter 3 gives an overview of the CDF II detector and some of the data analysis techniques that have become standard for all the analyses produced at CDF, including this one.

For the studies described in this dissertation there are several clearly identifiable stages of data analysis which can be summarized as follows:

- **Reconstruct the B decay candidates.**

  We use partially reconstructed semileptonic decays of B mesons. We concentrate on the four semileptonic decays providing the highest statistics:

  $B^0 \to l^+ D^- X, D^- \to K^+ \pi^- \pi^-$;

  $B^0 \to l^+ D^{*-} X, D^{*-} \to \pi^- \bar{D}^0, \bar{D}^0 \to K^+ \pi^-$;

  $B^+ \to l^+ \bar{D}^0 X, \bar{D}^0 \to K^+ \pi^-$;

  $B_s \to l^+ D_s^- X, D_s^- \to \pi^- \phi, \phi \to K^+ K^-$.

  Unless stated otherwise, the notations throughout this dissertation assume that charge conjugate particles and decay modes are also used. The choice of the data sample and the decays is dictated by the limitations in statistics of the $\bar{B}_s^0$ sample. The fully reconstructed hadronic decays have lower statistics and simply do not provide enough events to make a meaningful measurement. This choice however does add the difficulty of dealing with the sample composition issues. Chapter 4.1 gives the details of the B candidate reconstruction and selection, while Chapter 5 discusses the techniques for measuring and

correcting for the sample composition.

- **Measure particle content around B decays.**

  We restrict this analysis to studying the charged particles found around the B decays because finding neutral particles such as K^0 suffers from larger reconstruction inefficiencies and does not provide comparable statistics. We combine TOF, dE/dx and invariant mass information in an unbinned likelihood fit. Since the particle content changes rapidly with the track transverse momentum p_T the fit is performed in several ranges of p_T. Particles are selected in a cone around the lD reconstructed momentum. We reject particles that do not come from the same primary vertex as the B meson. Chapter 6 describes the procedure we use to extract the particle composition information from the samples of selected tracks.

  - **Develop and calibrate the tools for particle identification.**

    Part of this analysis stage is developing the necessary particle ID tools. In particular, we calibrate the TOF performance in bins of particle p_T since the fits are performed on track samples restricted to certain p_T regions. These items are discussed in Appendices.

- **Correct the measurements for sample composition.**

  The reconstructed lD signals in the semileptonic samples do not represent the pure B meson signals. The backgrounds from other B mesons are present at

the level of 10-20% and a correction has to be applied to the final result.

We also attempt to correct the measurements for the secondary particles coming from unreconstructed intermediate D states. Chapter 5 provides the details.

- **Correct the measurement for the detector efficiency.**

  In order to compare the data with the generator level Monte-Carlo simulation we correct our measurements by the detector efficiencies to estimate the yields of the charged particles[3] in a cone around B signals.

- **Compare the measurements with the Monte-Carlo prediction.**

  We compare the measured fractions and yields with the prediction of the `Pythia` generator. Generation was performed in `msel=1` mode (which means the quark creation is simulated using flavor creation, flavor oscillation and flavor excitation modes, as explained in Section 2.2), and it uses the Peterson fragmentation function [41] to describe the energy-momentum transfer from b quark to B meson. Section 6.5 explains the generation of the Monte-Carlo sample.

Finally, Chapter 7 discusses the results of the measurements, the interpretation of these results and the Monte-Carlo comparison.

[3] What we measure directly is the yields of the tracks satisfying certain selection criteria, and only a fraction of charged particles will produce such tracks.

Chapter 2

Physics Overview

2.1 The Standard Model

The goal of particle physics is to understand the behavior of the elementary particles. The theory in particle physics that quantitatively describes the behavior of all known particles is called the Standard Model. It encompasses all known particles and interactions, with the exception of gravity.

In the Standard Model, all matter consists of 12 fundamental fermions (half integer spin particles), six quarks and six leptons subdivided into three generations of two quarks and two leptons. These particles can interact through four different types of forces: strong, weak, electromagnetic and gravitational. Leptons participate in the weak interactions, while quarks can also interact via the strong forces, governed by a quantum property called color. The interactions between these fun-

damental particles are described by a set of integer spin particles, or bosons, that are often called the force carriers. The carriers of the strong force are called gluons, the carriers of the weak force are W and Z bosons, and photons are the carriers of the electromagnetic force.

While the Standard Model has been incredibly successful in explaining a host of experimental observations, it is by no means complete and many activities focus on searching for phenomena that cannot be explained by it. In this Chapter we provide an overview of the physics processes underlying the phenomena studied in our analysis, how they are understood in terms of the Standard Model and how they are relevant for the tests of the Standard Model's validity.

2.2 b quark production at Tevatron

As stated before, this analysis studies the b quark properties associated with their production in a high energy $p\bar{p}$ collision. The b quark is a third generation quark. It is the heaviest quark that can couple to other quarks to form hadrons, and the second heaviest quark after top quark. Chapter 3 describes in more detail how protons and anti-protons are produced and made to collide. During this collision the constituent particles of the $p\bar{p}$ pair (quarks and gluons) interact, and, as a result, a variety of secondary products may be created.

Quantitatively, the Standard Model describes the strong interaction with a theory called "Quantum Chromodynamics" (QCD). According to this theory, the force

potential between the quarks grows with the distance between them which prevents quarks from existing in a free state. The quarks therefore have to form colorless composite particles, such as mesons and baryons, and can only be observed indirectly. Mesons are the two-quark states where color neutrality is achieved by coupling a quark-antiquark pair, for example a red and quark an red anti-quark. The baryons are three-quark states where quarks of three different colors (red, blue, green) form a color-neutral (also called white) state.

The strong interactions between quarks conserve quark flavor, which means that when heavy quarks are created during the $p\bar{p}$ collision, they are created in pairs of a quark and corresponding anti-quark. The leading order QCD processes resulting in heavy quark creation involve two-to-two QCD transitions as shown in Figure 2.1, via either the annihilation of two lighter quarks or fusion of two gluons. Since the rest frame of the $p\bar{p}$ system is close to being at rest in the detector frame of reference, one can naively expect that in the leading order processes, the heavy quarks will be created with approximately opposite momenta. This leads to the rise of the terms "same-side" and "opposite-side" b quark when referring to the quark pairs.

The 2-2 transition resulting in $b\bar{b}$ pair production is often referred to as the flavor creation process. The lowest order diagrams in Figure 2.1 contribute to only a part of the total b quark production cross-section. Figure 2.2 shows additional diagrams that when included make a non-trivial contribution to the flavor creation process. These diagrams correspond to the next-to-leading order (NLO) processes

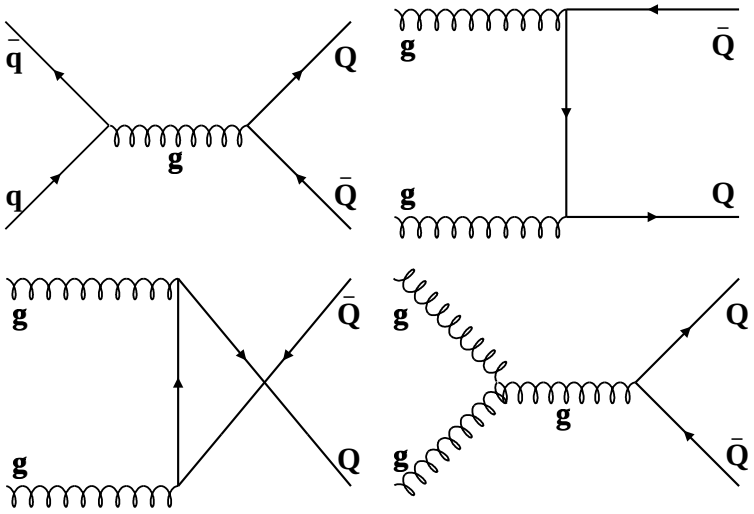


Figure 2.1: The Feynman diagrams showing the flavor creation type transitions leading to heavy quark Q creation at the Tevatron. These diagrams show the leading order QCD processes.

called initial and final state gluon emission. Finally, Figure 2.3 shows additional NLO processes that are called flavor excitation and gluon splitting. They contribute significantly to the b quark production at Tevatron. Note that in these additional NLO processes one cannot make the assumption about the relative momenta of the b quarks and anti-quarks being roughly opposite. In fact, the study of the momentum correlations of the same side and opposide side b quarks is one of the major sources of information we have for measuring the relative contributions of the flavor creation versus the NLO processes.

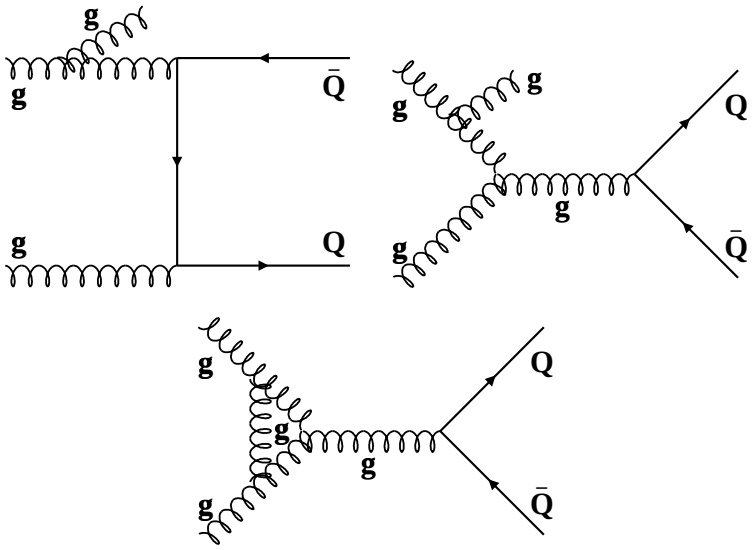


Figure 2.2: The Feynman diagrams showing the flavor creation type transitions leading to heavy quark Q creation at the Tevatron. These diagrams correspond to the next to leading order QCD processes. The two diagrams on top are called real gluon emission and the bottom is called the virtual gluon emission diagram.

2.3 Fragmentation models

After a pair of quarks is created in the $p\bar{p}$ collision they undergo a process of forming composite hadron particles. This process is called *fragmentation*. The term *hadronization* is used interchangibly sometimes. Fragmentation is a long distance process with small momentum transfer, so that perturbative QCD techniques cannot be used for an analytical description of this process. Instead, a number of models have been proposed to describe quark fragmentation. We will briefly describe two of these models, cluster fragmentation and string fragmentation. The simplified drawings illustrating these models are shown in Figure 2.4 and are explained in the following paragraphs.

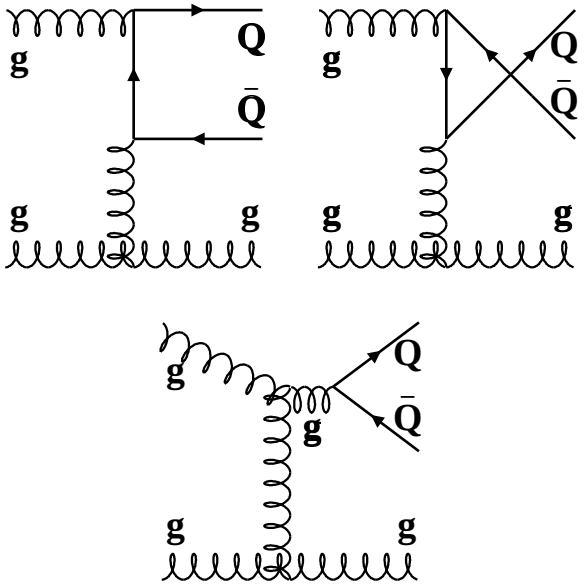


Figure 2.3: The Feynman diagrams illustrating the flavor excitation(top) and gluon splitting (bottom) processes leading to heavy quark Q creation. These are the next to leading order QCD processes making significant contribution to the b quark cross-section at the Tevatron.

In the cluster model the quarks lose part of their momenta through gluon emission, and the resulting gluons are split non-perturbatively into pairs of light quarks: $g \to q\bar{q}$. These quarks are then broken into colorless clusters which undergo simple decays into pairs of hadrons chosen according to the density of the clusters and their quantum numbers. This model's strong features are the small number of parameters and a simple mechanism for the suppression of heavy particle production and generation of the momentum spectra. However it has encountered significant difficulties in explaining heavy quark fragmentation[1] and the suppression of baryon

[1]It is very difficult to describe both light quark and heavy quark fragmentation within the bounds of the same phenomenological model such as the cluster fragmentation model. The fragmentation of heavy quarks is expected to differ significantly from that of light quarks due to very

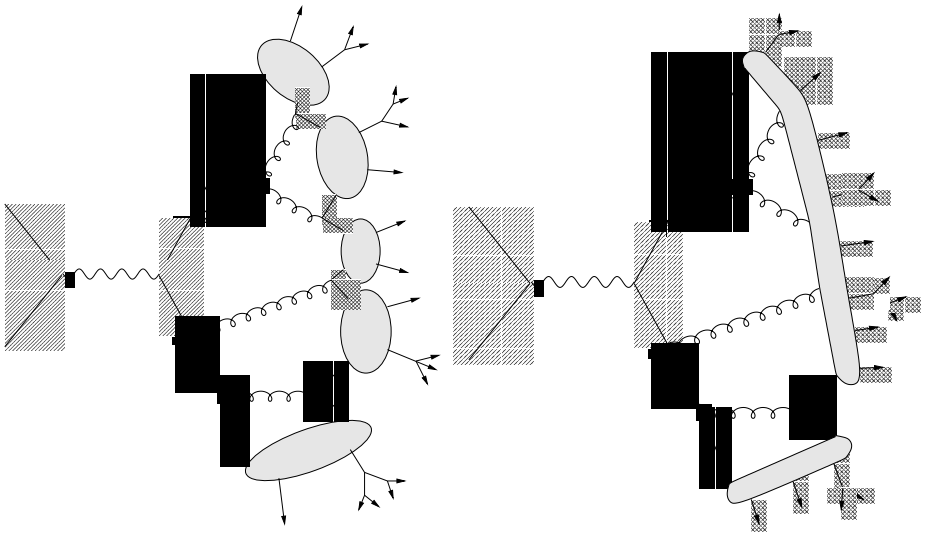


Figure 2.4: Simplified diagrams of the cluster (left) and string (right) fragmentation models.

production.

The string fragmentation model has been the most successful in quantitatively describing the fragmentation process and is the one most often used in Monte-Carlo generators such as `Pythia` and `Herwig`. This model posits that a string exists between the quarks produced in a $p\bar{p}$ collision and the energy of this string represents the quark-quark potential and grows with the string size. This leads to the fragmentation of the string *i.e.*, when the distance between the quarks reaches a certain limit, the string breaks into two or more strings by creating quark-antiquark pairs, thus moving into a lower energy state. Depending on the initial energy of the

diffirent kinematics. A heavy quark plays the leading role in determining the kinematics of the fragmentation process and for example will often transfer most of its momentum to the resulting hadron. This is not true for the light quark hadronization.

$b\bar{b}$ pair this division process may be repeated one or more times. The b quarks then couple with other quarks to form composite particles such as mesons and baryons. A number of other particles may also be created. The strings were dubbed so because they are described in QCD as clouds of gluons forming string-like spatial structures, or color fields.

To explain the breaking of the string there are several phenomenological models within the string fragmentation model. The number of these models has grown significantly over the years. Here we will discuss only one such model, the Peterson quark fragmentation model [41]. It has traditionally been the most popular choice for simulations, although it has been found that other fragmentation models describe the experimental results better in some cases. For a recent example and a review of other functions see [21].

The fundamental property of the fragmentation process that all fragmentation functions aim to describe is the momentum-energy transfer from the original quark to the resulting hadron. The Peterson fragmentation function is derived from the assumption that the amplitude of the hadronization process is inversely proportional to the energy difference between the initial and the final states of the fragmentation process. If a heavy quark Q with initial momentum P fragments into a hadron H of momentum zP by coupling with a lighter quark $\bar{q}$ (so that a quark q remains in the string), then the energy difference can be written as

$$\Delta E = E_Q - E_H - E_q = \sqrt{m_Q^2 + P^2} - \sqrt{m_H^2 + (zP)^2} - \sqrt{m_q^2 + (1-z)^2 P^2}.$$

Making an assumption of a heavy quark fragmenting at high momentum, so that $m_Q \approx m_H$ and $m_Q/P << 1$, and expanding in terms of $(m_Q/P)^2$, one can derive

$$\Delta E = \frac{m_Q^2}{2P}\left(1 - \frac{1}{z} - \frac{\epsilon_Q}{1-z}\right),$$

with the notation $\epsilon_Q = (m_{\bar{q}}/m_Q)^2$. By adding the normalization and the phase space factors, the Peterson fragmentation function relates the transition amplitude to this energy transfer as

$$|\mathcal{M}(Q \to Q\bar{q})|^2 \propto \frac{1}{\Delta E^2},$$

from which we can finally derive a probability $D(z)$ of finding an H hadron of momentum zP among the debris of the original heavy quark Q:

$$D(z) = \frac{N}{z}\frac{1}{\left(1 - \frac{1}{z} - \frac{\epsilon_Q}{1-z}\right)^2},$$

where N is the normalization constant. The parameter ϵ_Q is usually considered a free parameter of the model. It is important to understand that this model is purely phenomenological, and even though it can be tuned to describe many practical situations, it breaks down in many scenarios.

A typical example of the Peterson fragmentation function used in the Monte-Carlo generators is shown in Figure 2.5. Note that the momentum fraction transferred to the hadron is larger for a heavier b quark than for the c quark.

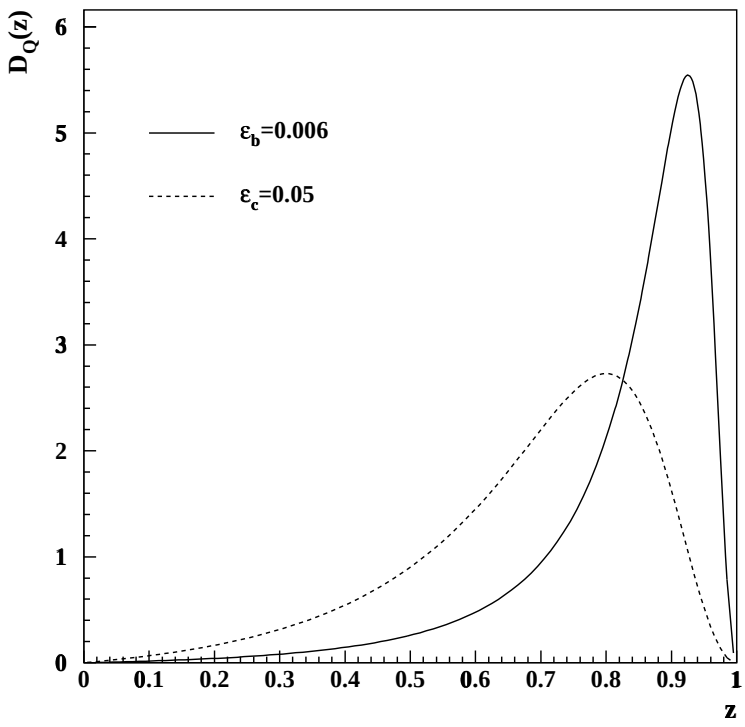


Figure 2.5: An example of the Peterson fragmentation function for b (solid curve) and c (dashed) quarks commonly used in the simulations of the fragmentation process.

2.4 The weak interactions and CKM formalism

The theory of weak interactions is one of the most interesting topics of modren physics. The weak interactions do not conserve quark flavor. They also do not obey charge-parity (CP) conservation laws and therefore may hold a key to explaining why the Universe consists primarily of matter and not of a equal mixture of matter and anti-matter.

This section will describe the current theory of the weak interactions accepted in the Standard Model. It should be noted however that the amount of CP vi-

olation present in the Standard Model is not enough to account for the observed matter-antimatter asymmetry and many physicists believe that our view of the weak interactions will change dramatically in the future. Our current inability to explain the matter-antimatter asymmetry of the Universe is one of the reasons why many experiments are testing the Standard Model's description of the weak force. The next sections will show how the B mixing studies at CDF contribute to these efforts.

In the Standard Model, the weak interactions are described by assuming that the weak interaction eigenstates and mass eigenstates of the quarks are not the same and instead a weak eigenstate is a mixture of the mass eigenstates of different quarks. This is quantitatively described by a Cabbibo-Kobayashi-Maskawa (CKM) [33, 25] matrix which connects the weak eigenstates (d', s', b') and the corresponding mass eigenstates d, s, b as

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \equiv \hat{V}_{\mathrm{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{2.1}$$

The CKM matrix is a unitary matrix which forbids the flavor-changing transitions involving neutral force-carriers (also called neutral current interactions) which are not observed in nature. The elements of the CKM matrix can be complex numbers and that allows for the CP violation in the Standard Model.

The CKM matrix can be parametrized in a variety of ways [16, 29] to emphasize

some of its properties. Two of the most common parameterizations used are the standard parameterization, recommended by the Particle Data Group [24], and the Wolfenstein parameterization [52].

The standard parameterization uses the angles $\theta_{i,j}(i,j=1,2,3)$ and complex phase δ as follows:

$$\hat{V}_{\rm CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -s_{23}c_{12} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \qquad (2.2)$$

where $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$. c_{ij} and s_{ij} can all be chosen to be positive and δ may vary in the range $0 \leq \delta \leq 2\pi$.[2] The benefit of this parameterization is that it exposes the phase δ which is responsible for CP violation. This CP violating phase is multiplied by a very small parameter which shows clearly the suppression of the CP violating mechanism. This parameterization also describes the CKM matrix precisely.

The Wolfenstein parameterization uses the fact that the flavor changing elements of the CKM matrix are small, and expands the CKM matrix elements in terms of a parameter $\lambda = |V_{us}| = 0.22$, so that

$$\hat{V} = \begin{pmatrix} 1 - \frac{\lambda 2}{2} & \lambda & A\lambda 3(\varrho - i\eta) \\ -\lambda & 1 - \frac{\lambda 2}{2} & A\lambda 2 \\ A\lambda 3(1 - \varrho - i\eta) & -A\lambda 2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4)\,. \qquad (2.3)$$

[2]The measurements of CP violation in the kaon decays show that δ has to be in the range $0 < \delta < \pi$.

While this parameterization is not exact, it makes the relative size of different elements apparent and transparently illustrates the CP-violating mechanism.

The unitarity of the CKM matrix implies there are six independent relations involving its elements which have to hold. For B physics the most interesting of these is

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0, \tag{2.4}$$

because it involves the V_{ub}, V_{cb} and V_{td} matrix elements at the same time, and these elements are important in B hadron decays. Using the Wolfenstein parameterization this relation can be written in a more intuitive way. The product $V_{cd}V_{cb}^* = -A\lambda^3 + \mathcal{O}(\lambda^7)$ is almost real. Using this number as a scale, we expand the other two members of the relation 2.4 as

$$\frac{1}{A\lambda^3}V_{ud}V_{ub}^* = \bar{\varrho} + i\bar{\eta},$$

and

$$\frac{1}{A\lambda^3}V_{td}V_{tb}^* = 1 - (\bar{\varrho} + i\bar{\eta}),$$

where $\bar{\varrho} = \varrho(1 - \frac{\lambda^2}{2})$ and $\bar{\eta} = \eta(1 - \frac{\lambda^2}{2})$. Written in this way, the relation 2.4 describes a triangle in the complex plane $(\bar{\varrho}, \bar{\eta})$ as shown in Figure 2.6. In fact, this triangle, together with the elements $|V_{us}|$ and $|V_{cb}|$ contains all the information about the CKM matrix.

The sides CA and BA in the unitarity triangle are often denoted by R_b and R_t,

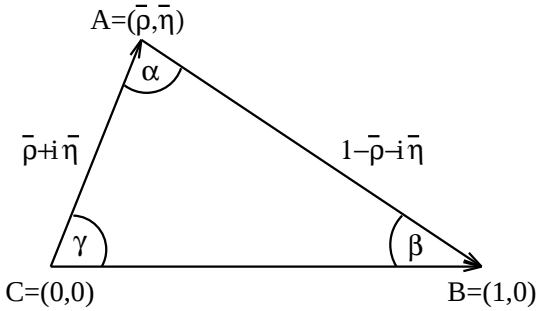


Figure 2.6: The triangle illustrating the unitarity of the CKM matrix.

and can be written as

$$R_b \equiv \frac{|V_{ud}V_{ub}^*|}{|V_{cd}V_{cb}^*|} = \sqrt{\bar{\varrho}^2 + \bar{\eta}^2} = (1 - \frac{\lambda^2}{2})\frac{1}{\lambda}\left|\frac{V_{ub}}{V_{cb}}\right|, \tag{2.5}$$

$$R_t \equiv \frac{|V_{td}V_{tb}^*|}{|V_{cd}V_{cb}^*|} = \sqrt{(1 - \bar{\varrho})^2 + \bar{\eta}^2} = \frac{1}{\lambda}\left|\frac{V_{td}}{V_{cb}}\right|. \tag{2.6}$$

The angles β and γ of the unitarity triangle are related directly to the complex phases of the CKM-elements V_{td} and V_{ub}, respectively, through $V_{td} = |V_{td}|e^{-i\beta}$, $V_{ub} = |V_{ub}|e^{-i\gamma}$.

The angles and sides of the unitarity triangle have been the subject of active investigation in the last decade as they provide a direct test of the CKM unitarity and this test may give evidence for new physics.

2.5 Mixing of B mesons and CKM

As explained in the Introduction, one of the primary goals of our analysis is to contribute to the measurement of $\bar{B}_s^0$ mixing at CDF by helping the development of the flavor tagging tools. This section will describe the physics of B meson mixing

and how the studies of B mesons at CDF can be used to constraint the CKM matrix elements.

2.5.1 B meson mixing

Here we a present a brief overview of B meson mixing. The full formalism of this phenomena is covered extensively in the literature [10, 12, 45] and we therefore will only highlight the most relevant points.

The fact that the quarks can undergo both flavor-conserving strong interactions and flavor-changing weak interaction, coupled with the fact that B_q^0 and $\bar{B}_q^0$ mesons have the same quantum numbers allows for transition of these particles into their respective antiparticles in the Standard Model. This is the process called flavor oscillation or flavor mixing. There are two neutral B meson systems ($B_d^0 - \bar{B}_d^0$ and $B_s^0 - \bar{B}_s^0$, here denoted as $B_q^0 - \bar{B}_q^0$, $q = s, d$) which exhibit particle-antiparticle mixing. The leading order Feynman diagrams of this process are shown in Figure 2.7.

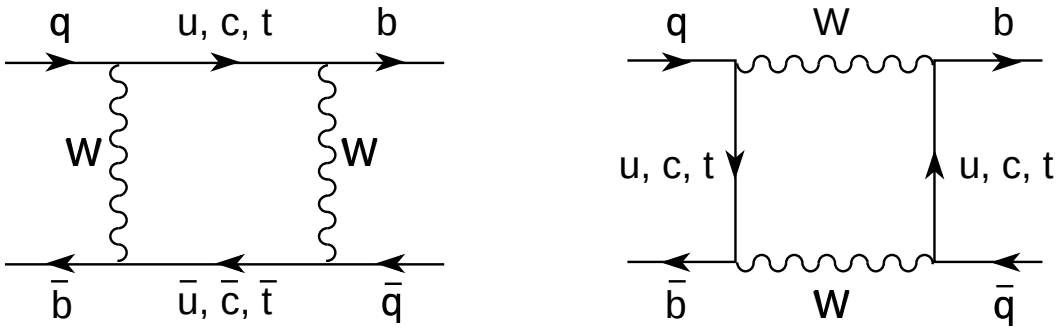


Figure 2.7: The Feynman diagrams of the leading order mixing transitions.

In each $B_q^0 - \bar{B}_q^0$ system, the light (L) and heavy (H) mass eigenstates can be represented as a superposition of the strong interaction eigenstates with the weights p and q:

$$|B_{\mathrm{L,H}}\rangle = p|B_q^0\rangle \pm q|\overline{B}_q^0\rangle \,.$$

The mass eigenstates have a mass difference $\Delta m_q = m_{\mathrm{H}} - m_{\mathrm{L}} > 0$ and a total decay width difference $\Delta\Gamma_q = \Gamma_{\mathrm{H}} - \Gamma_{\mathrm{L}}$.

If we introduce the notations $g_+(t)$ and $g_-(t)$ to write the time-dependent evolution of a $B_q^0 - \bar{B}_q^0$ system that starts in a pure B_q^0 state as

$$B_q^0(t) = \begin{pmatrix} g_+(t) & g_-(t) \end{pmatrix} \begin{pmatrix} B_q^0 \\ \\ \bar{B}_q^0 \end{pmatrix}, \tag{2.7}$$

then the Shrodinger equation describing this state is

$$i\frac{d}{dt}\begin{pmatrix} g_+(t) \\ \\ g_-(t) \end{pmatrix} = \begin{pmatrix} m - \frac{1}{2}i\Gamma & m_{12} - \frac{1}{2}i\Gamma_{12} \\ \\ m_{12}^* - \frac{1}{2}i\Gamma_{12}^* & m - \frac{1}{2}i\Gamma \end{pmatrix} \cdot \begin{pmatrix} g_+(t) \\ \\ g_-(t) \end{pmatrix}, \tag{2.8}$$

where m is the mass of the weak eigenstates and Γ is the decay width. The parameters m_{12} and Γ_{12} are determined from theory by evaluating the Feynman diagrams of the mixing transitions. This equation is often equivalently written as

$$i\frac{d}{dt}\begin{pmatrix} g_+(t) \\ \\ g_-(t) \end{pmatrix} = (\mathcal{M} + \Gamma) \cdot \begin{pmatrix} g_+(t) \\ \\ g_-(t) \end{pmatrix}, \tag{2.9}$$

where $\mathcal{M}$ is called the mass matrix and contains only the m elements and Γ is called the decay matrix and contains only the Γ elements.

Diagonalizing the Hamiltonian in the equation 2.8 to obtain the CP eigenstates in terms of flavor eigenstates one can derive

$$q = p = \frac{1}{\sqrt{2}}$$

The time-evolution of a pure B_q^0 or $\bar{B}_q^0$ state from the time $t = 0$ is given by a solution to 2.8:

$$|B_q^0(t)\rangle = g_+(t)\,|B_q^0\rangle + \frac{q}{p}\,g_-(t)\,|\overline{B}_q^0\rangle\,, \tag{2.10}$$

$$|\overline{B}_q^0(t)\rangle = g_+(t)\,|\overline{B}_q^0\rangle + \frac{p}{q}g_-(t)\,|B_q^0\rangle\,, \tag{2.11}$$

which means that the flavor of the system either does (+) or does not (-) change at time t with the probabilities proportional to

$$|g_\pm(t)|^2 = \frac{e^{-\Gamma_q t}}{2}\left[\cosh\left(\frac{\Delta\Gamma_q}{2}\,t\right) \pm \cos(\Delta m_q\,t)\right]\,, \tag{2.12}$$

where $\Gamma_q = (\Gamma_{\rm H} + \Gamma_{\rm L})/2$.

This relation means that by measuring the time dependence of the $B_q^0 \to \bar{B}_q^0$ transition one can extract the value of Δm_q. In practice this is done by the experiments by reconstructing the B_q^0 decays (which provides information about the flavor of the final state) and using the flavor tagging techniques (as described in the Introduction) to deduce the flavor of the initial state. If the decay time can also be determined by the experimental apparatus, then this information is sufficient to detect the time-dependant $B_q^0 - \bar{B}_q^0$ oscillations and measure Δm_q.

On the other hand, by explicit evaluation of the box diagrams of the mixing process one can make a theoretical prediction for the values of m_q (which are the

diagonal elements of the mass matrices). Specifically, we have

$$M_{12} = -\frac{G_F^2 m_W^2 \eta_B m_{B_q} B_{B_q} f_{B_q}^2}{12\pi^2} S_0(m_t^2/m_W^2) \left(V_{tq}^* V_{tb}\right)^2 , \tag{2.13}$$

where G_F is the Fermi constant, m_W is the W boson mass and m_t the mass of top quark, m_{B_q}, f_{B_q} and B_{B_q} are the B_q^0 mass, weak decay constant and bag parameter correspondingly. $S_0(x_t)$ is a known function that can be approximated by $0.784\, x_t^{0.76}$ and V_{ij} are the elements of the CKM matrix. The QCD correction η_B is of order unity. Equation 2.13 shows a direct connection between the mixing oscillation frequency and the values of the CKM matrix elements.

2.5.2 Using mixing measurement to constraint CKM

The presence of mixing has been observed in both $\bar{B_s^0}$ and $\bar{B_d^0}$ system. The mixing frequency in B_d system has been known with some precision for a decade, while the first measurement of the $\bar{B_s^0}$ mixing frequency is expected to be made at Tevatron.

The importance of the mixing measurements stems from the fact that the ratio of their mixing parameters can be expressed as

$$\frac{\Delta m_s}{\Delta m_d} = \frac{m_{B_s}}{m_{B_d}} \frac{|V_{ts}|^2}{|V_{td}|^2} \frac{F_{B_s}^2 B_{B_s}}{F_{B_d}^2 B_{B_d}}. \tag{2.14}$$

The measurement of this ratio would help enormously in the determination of the right side of the unitarity triangle. The first measurement of Δm_s, along with the established value of Δm_d is expected to constraint the right side of the unitary triangle to better than 6% since the theoretical uncertainties on the hadronic matrix

elements partially cancel in the ratio

$$\xi = F_{B_s}\sqrt{B_{B_s}}/F_{B_d}\sqrt{B_{B_d}} = 1.15 \pm 0.05^{+0.12}_{-0.00},$$

which is a flavor-symmetry breaking factor obtained from lattice QCD calculations [51].

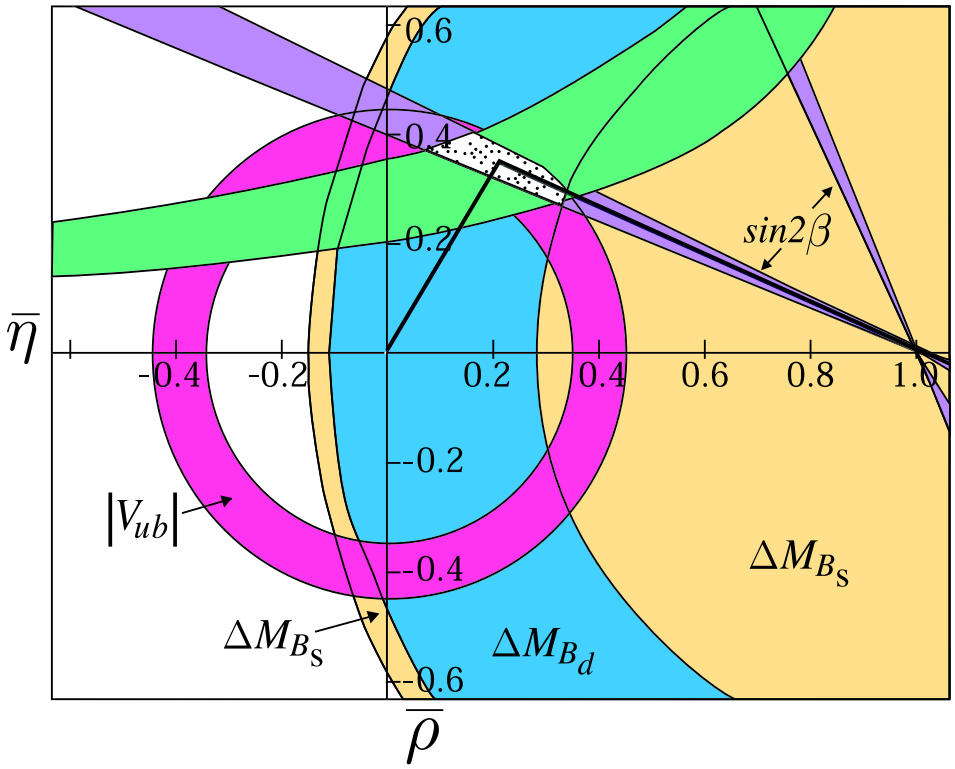


Figure 2.8: The current constraints on the unitarity triangle parameters.

Figure 2.8 shows the current constraints [34, 24] on the parameters of the unitarity triangle. The green band corresponds to the measurements of the CP-violating parameter from mixing in the neutral kaon system; the blue band represents the constraint from the measurements of mixing of the $\bar{B}^0$ meson. The magenta region

is the result of the measurement of the $sin(2\beta)$ CKM angle performed by measuring the CP-violation in the B decays (CDF made a first measurement of this quantity). The red circles outline the constraints on the size of the $|V_{ub}|/|V_{cb}|$ which can be measured directly in B hadron decays.

The yellow area is the constraint from the $\bar{B}_s^0$ mixing parameters, on which there is currently no experimental measurement. The world-best experimental limit on the value of Δm_s is $\Delta m_s > 14.4\text{ps}^{-1}$[24]. An upper limit can also be obtained, but only from the theoretical calculations using the Standard Model parameters. This theoretical limit is currently $\Delta m_s < 27.5\text{ps}^{-1}$ [15], and is not included in Figure 2.8. Measuring the value of Δm_s will allow to combine the Δm_s and Δm_d regions into one narrow band strongly constraining the unitarity relation. Doing this would be a great achievement; for that reason the $\bar{B}_s^0$ mixing measurement is one of the most anticipated results in the particle physics today.

Chapter 3

Detector Overview

The experimental apparatus used in our research is the Collider Detector at Fermilab, or CDF detector[1], where the results of the $p\bar{p}$ collisions are observed. The protons and anti-protons are delivered to the detector by an accelerator complex. Both the accelerator and the detector are hosted at Fermi National Accelerator Laboratory[2], Illinois, USA.

This Chapter will give an overview of the accelerator complex and the CDF II detector to the extent that the reader may understand their role in our analysis.

Apart from using the hardware, every analysis at CDF relies on a number of low level procedures that need to be performed for the data to become easily usable. Examples of such procedures are using the filtering software to reject unneeded

[1] Also called CDF II after a serious redesign that took place in 1996-2000. The current set of activities at CDF detector is also called RUN II.

[2] Also referred to as Fermilab or FNAL.

events or using reconstruction software to convert the raw hardware response of the detector into information about particle trajectories and so on. These common techniques used in data analysis and processing are also a subject of this Chapter.

3.1 The accelerator complex

The accelerator complex at Fermilab produces the protons and anti-protons, accelerates them to the energy of 980 GeV and delivers them to the experiments where the $p\bar{p}$ collisions can occur at effective center of mass energy $\sqrt{s} = 1.96$ TeV. The entire complex is often collectively called the Tevatron, although technically it is the name of the last stage of the complex. An accelerator complex consists of several parts because the protons cannot be created and accelerated from rest to 980 GeV in the same device. The schematic drawing of the entire complex is shown in Figure 3.1. In the following text we will describe the functioning of the accelerator and refer to this Figure.

3.1.1 Proton production

Protons are produced in the Cockroft-Walton chamber from hydrogen gas. The gas is put in a chamber with large electric field and the hydrogen ions are produced by the means of electric discharges. The ions are accelerated by the electric field and then a magnetic field is used to filter out the positively charged particles to obtain a beam of H^- ions. This system can really be thought of as the first accelerator in

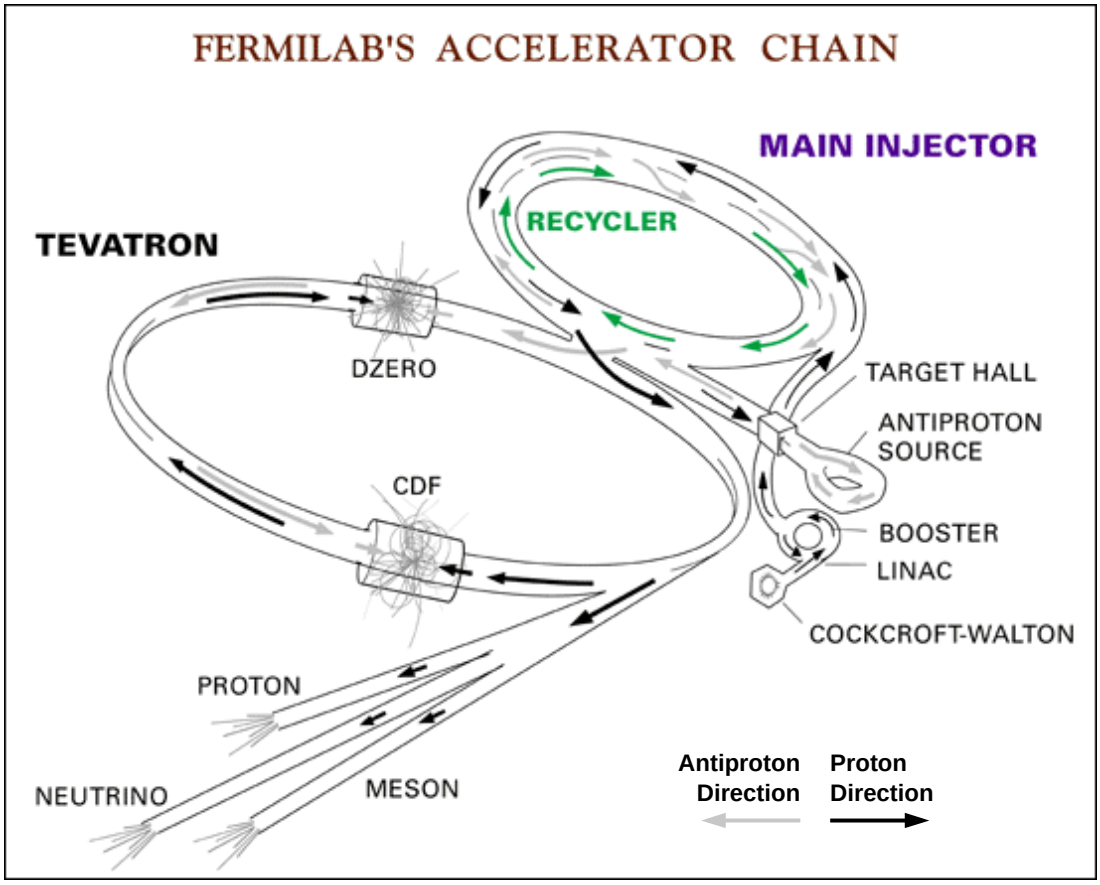


Figure 3.1: Simplified overview of the Fermilab's Tevatron accelerator complex.

the complex and is often called a Pre-accelerator or Preacc.

The average ion energy in the beam after Preacc is 750 KeV and the beam is fed into a linear accelerator (Linac) through a transfer pipe. The particles are accelerated in the Linac to an energy of 400 MeV. The Linac beam is also used in a cancer treatment facility at Fermilab which is an example of the variety of the projects that the accelerator complex serves.

The next stage is the Booster. It the first circular accelerator (or *synchrotron*) in the chain and has a diameter of 150 meters. The negative H^- ions from the

Linac are injected into the Booster and the electrons are stripped from them by sending the beam through a carbon foil. This leaves only the proton beam in the Booster. The protons are accelerated to an energy of 8 GeV in the Booster before being transfered to the next stage of acceleration, the Main Injector.

3.1.2 Anti-Proton production

Once a moderately energetic beam of protons is available in the accelerator [3] the anti-protons are produced by directing that beam on a metal target [4]. Many nuclear reactions take place when a beam of protons interacts with the target, and the products of these reactions are selected based on their momentum and charge using a system of magnets. This allows to collect and beam of approximately 8 GeV anti-protons from a spray of these secondary particles.

This beam of anti-protons is transfered to a synchrotron called the *Debuncher*. The debuncher has a radius of 90 meters and does not accelerated the beam further. The anti-protons collected after the proton beam hits the target naturally have a large momentum spread and the purpose of the Debuncher is to efficiently capture this beam and make it much more manageable using techniques such as stochastic cooling.

[3] In fact, a proton beam from the Main Injector is used.

[4] Potentially many materials could be used for the target. Nickel is the metal currently used; copper was the metal used for some time but the heat generated when protons strike the target was a problem when using copper.

The anti-protons from the Debuncher are sent to the *Accumulator* which is a 75 meter radius synchrotron housed in the same tunnel as the Debuncher. The purpose of this machine is to store the 8 GeV $\bar{p}$ beam until it has to be sent to the Main Injector.

3.1.3 The Main Injector and Tevatron

The last two stages of the accelerator complex are the Main Injector and the Tevatron. The Main Injector is a synchrotron over 500 meters in radius that is capable of accepting 8 GeV protons from the Booster or 8 GeV anti-protons from the Accumulator and then accelerating them to 150 GeV. It can then send the accelerated beam to the fixed target location, or to the Tevatron for further acceleration, or to the *NuMi* neutrino experiment. The Main Injector can accelerate the beam every 2.2 seconds.

The last stage of acceleration is the Tevatron, a synchrotron more than 4 miles in circumference and the largest Fermilab accelerator. At the time of the writing of this thesis it is also the highest energy accelerator in the world, to be superseded by LHC in a few years. It is also the only accelerator at Fermilab that uses cryogenic cooling, allowing the use of the super-conducting magnets [5].

The primary purpose of the Tevatron is to store the protons and anti-protons circulating inside at the energy of 980 GeV so that they can collide and produce

[5]Using super-conducting materials for the magnets allows achieve high currents and magnetic fields without the danger of overheating the materials.

interesting particles. The protons and anti-protons are sent to the Tevatron from the Main Injector and accelerated from 150 GeV to 980 GeV. At this point they can be stored in the Tevatron for long time colliding inside the CDF and D0[6] collision halls. This stable running condition is called a *store* and can sometimes go on for a few days.

The protons and anti-protons inside the Tevatron are stored in 36 bunches. For the actual collisions to occur, the Tevatron directs the bunches of protons and anti-protons at each other, so that their trajectories intersect at some point (collision point) and this process is called *bunch crossing*. During the bunch crossing one or more of the $p\bar{p}$ collisions may occur, creating interesting secondary particles that the detectors can reconstruct.

3.1.4 Summary

The Tevatron accelerator complex produces, accelerates and delivers to CDF the protons and anti-protons that collide inside the detector. The products of these collisions can be detected and reconstructed by the detector, and are a subject of the studies of physicists.

[6]D0 is the other collider detector at Fermilab.

3.2 The CDF II detector

The Collider Detector at Fermilab (CDF) is a general purpose solenoidal detector located around the point of the $p\bar{p}$ bunch crossing. The collision point is the geometrical center of the detector, with the detector itself being approximately symmetrical both around the horizontal axis (z) and the vertical plane intersecting at the interaction point. An overview of the detector is shown in Figure 3.2. The proton and anti-proton beams travel towards each other along the horizontal axis called *beam axis* or *beam line*. Any plane perpendicular to the beam line is called a transverse plane, and the intersection point of the beam line and the transverse plane is referred to as the *beam spot*.

Close to the beam line there are the tracking[7] systems of the detector (L00, COT, SVX and ISL) which are discussed later. A super-conducting magnet forms CDF's solenoid. The magnet and its cooling system (cryostat) are located outside the CDF tracking systems. The magnet creates a uniform horizontal magnetic field of 1.4 T along the z axis inside the tracking volume. Charged particles inside the tracking volume have curved trajectories; we measure the particle momenta based on the observed curvature.

Outside the tracking volume there are the systems that measure the time of flight of the particles (TOF), identify photons and electrons (calorimeters), detect muons (muon chambers). All these detector systems are described briefly in the

[7] Tracking is a process of reconstructing the trajectories of charged particles in the detector.

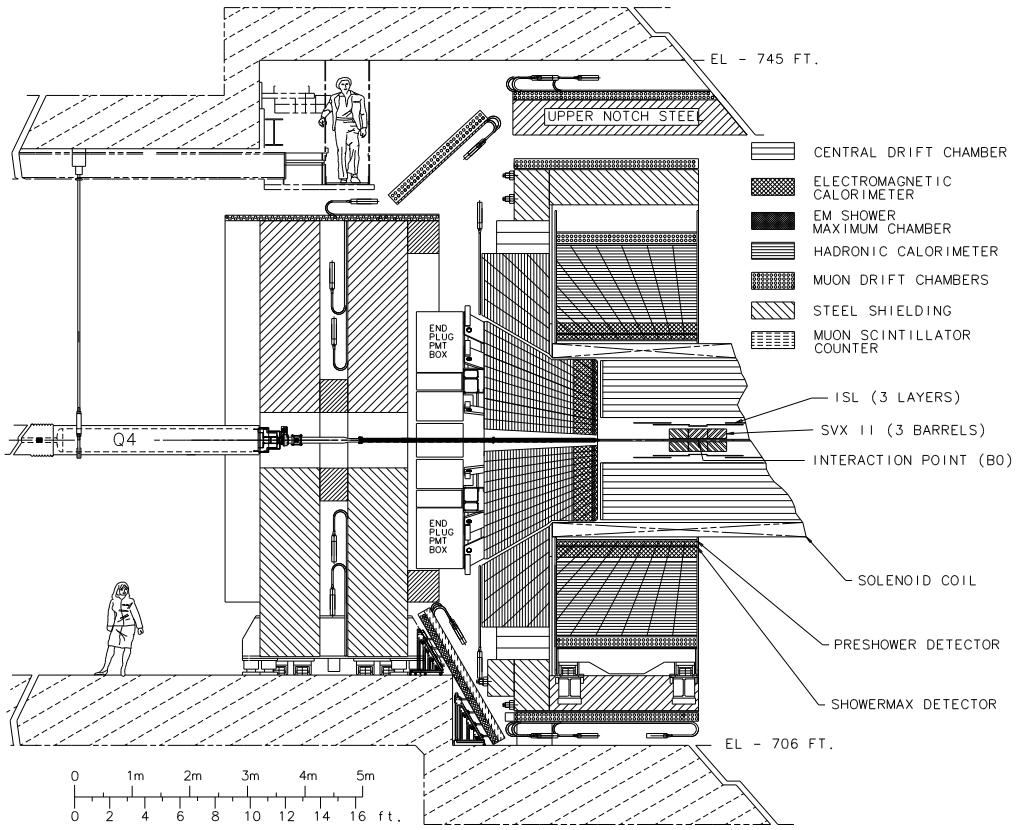


Figure 3.2: A partial cross-section view of the CDF detector.

following text.

3.2.1 The CDF coordinate system

The coordinate systems used at CDF reflect the detector symmetries. The horizontal axis along which the $p\bar{p}$ beams travel determines the z axis; positive z direction is the direction of the proton beam. The Cartesian system is chosen with the y axis pointing up and the x axis complements this frame. Since the detector has cylindrical symmetries we often use polar or cylindrical coordinates. The ϕ polar

angle is defined as the angle from the positive direction of the x axis, and the θ polar angle is counted from the positive direction of the z axis.

Another very frequently used variable when talking about the detector geometry is pseudo-rapidity η. This variable is defined as

$$\eta = -log(tan(\frac{\theta}{2})).$$

It is convenient because it is a purely geometrical quantity related to the polar angle θ, and, at the same time, it can be shown that this variable is Lorentz invariant under a boost in z direction in ultra-relativistic limit. Since a lot of the physics at CDF can be approximated by ultra-relativistic, or massless, limit this variable is useful for describing the geometry of the decays.

3.2.2 Tracking at CDF

As mentioned above, tracking is a process of reconstructing the trajectories of the particles in the detector. These reconstructed trajectories are called *tracks*. When charged particles travel in the detector they interact with the detector material which results in the energy loss during material ionization. The tracking systems work by detecting the ionization from tracks and based on this information mapping out the particle trajectory. Tracking is of course central to our analysis, as we try to measure the yields of the charged particles around the B signals, and most of the information about charged particles at CDF comes from tracking, such as charge, magnitude and direction of the momentum, and even particle velocity information

from the energy loss measurements.

Figure 3.3 shows the cross-section view of the tracking systems at CDF. In the following sections we will review these systems.

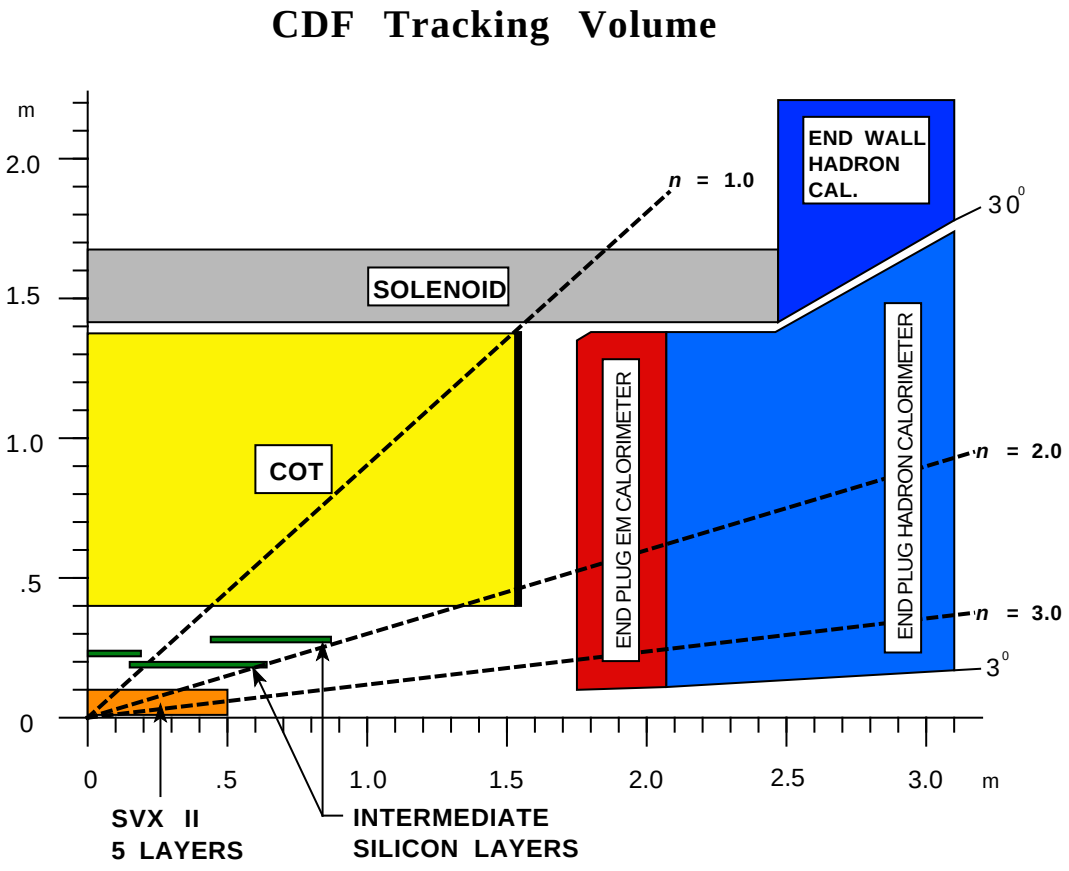


Figure 3.3: An overview of the tracking volume of the CDF detector.

3.2.3 Central-Outer Tracker (COT)

The COT detector is the main tracking device at CDF. It provides the tracking of charged particles at large radii in the pseudo-rapidity region $|\eta| < 1$. The COT

is an open cell cylindrical drift chamber located between the silicon detector and the time-of-flight detector. It covers the radial region from 43.4 cm to 132 cm and spans 150 cm in z direction. It provides coverage in the entire ϕ region. For charged particles traveling inside the COT volume, the detector provides an accurate measurement of transverse momentum p_T (momentum component perpendicular to the beam axis) and a less precise measurement of η.

The COT chamber contains 30,240 sense wires that span the entire length of the detector in z. Roughly half of the wires are aligned in the z direction (called axial wires) and half run at a 2 degree angle to the z axis (called small angle stereo wires) which facilitates the measurement of η. COT can also provide imprecise measurements of the particle impact parameter d_0 (the distance of the closest approach of particle trajectory to the interaction point) and the angle ϕ_0 defined as momentum direction at the point of closest approach to the interaction point.

During the operation the volume of COT is filled with a gas mixture consisting mainly of argon and ethane. Charged particles traversing COT ionize the gas producing electrons, which then travel to the nearest sense wire (voltage is applied to the wires). An avalanche multiplication of charge happens inside the high electric field region in the close vicinity of the wire through the electron-atom collisions. The resulting charge reaches the wire and this so-called 'hit' is read out by the electronics. The COT electronics is capable of reading out several simultaneous hits on the same sense wire at different z positions. The location of the wire itself

is not sufficient to get good precision on the hit position because it takes non-negligible time (called *drift time*) for the ionization charge to reach the wire. After the correction is applied for the drift time, the hit position can be measured with approximately 200 micron precision by the COT.

The tracking reconstruction algorithms find (reconstruct) particle trajectories (which in the presence of a magnetic field are helixes) that best correspond to the observed sense wire hits. Reconstructed trajectories are referred to as tracks. Since an average particle has hits on about 60 wires the resulting precision of measuring the track curvature (and hence momentum) is quite impressive and on most tracks is $\delta p_T/p_T^2 < 0.02$.

The tracks with available COT information are important for several reasons: they form the basis of the TOF reconstruction to provide particle identification information for track parent particle; they can be used in the silicon reconstruction to match the hits in the SVX detector to the COT track trajectory; finally they themselves contain information about particle velocity through the measured energy loss in the COT detector volume. All the tracks that we use to measure the particle production around B in this analysis are required to have the COT information.

Another extremely important application of the COT is particle identification. Since the amount of ionization charge can be measured by the COT one can deduce the amount of energy a particle has lost while traversing the COT volume. Since the particle energy loss in the material dE/dx is related to its velocity the COT can

be used for identifying the charged particle species by extracting the particle mass from velocity and momentum measurements. This feature is very important for our analysis since the dE/dx measurement complements the reduced TOF particle identification power at higher p_T (see Figure 3.8).

A more technical description of the COT detector and tracking algorithms is available in References [3, 49].

3.2.4 Silicon Detectors

The silicon strip detectors at CDF [30] provide a precise determination of the particle trajectory close to the beam line. This is important in our analysis, not only for improving the B meson reconstruction, but also for determining whether the tracks found in the event are consistent with the point of B decay.

The primary silicon detector system is the Silicon Vertex Detector (SVX) which is a cylindrical detector 3×29 cm long located between the radii of 2.5 cm and 10.7 cm. Two additional detectors, the Intermediate Silicon Layer (ISL) [2] and Layer 00 (L00) extend the silicon tracking to larger and smaller radii correspondingly. The tracking reconstruction used by our analysis makes use of SVX and L00 systems only.

A silicon micro-strip detector is consists of wafers of semi-conducting silicon crystals, 300 micron thick, supported on structures called ladders. Each crystal has a thin layer of p-type silicon strips on one side and n-type silicon stripe on the

other side of the silicon substrate (see Figure 3.4). Voltage is applied to the silicon substrate creating an electric field between the n and p layers. When a charged particle traverses the silicon material it produces ionization in the form of electron-hole pairs that drift in the electric field creating charge signals on both the n and p micro-strips, as illustrated in Figure 3.4. Due to the small size of these finely

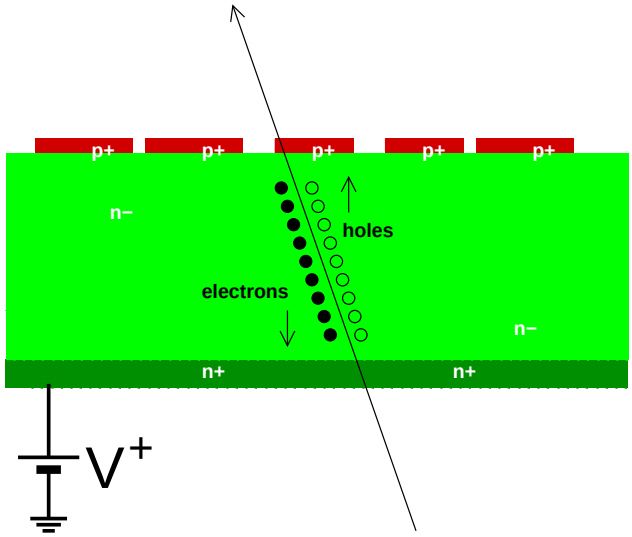


Figure 3.4: An illustration of the basic design and particle detection in a typical micro-strip silicon detector.

spaced strips, the positions of strips where the signal (or hit) is detected provide an extremely precise measurement ($\approx 12\mu m$) of the hit position. By making n and p strips run in different (e.g., orthogonal) directions, simultaneous measurements of the positions of the hits are made both in transverse plane and along the z axis.

The mechanical assembly of the SVX detector consists of 5 layers made up of 12 wafers each, as illustrated in Figure 3.5. The layers are positioned at increasing

radii. The ladders are mounted between two beryllium bulkheads to form a single SVX barrel, with the entire SVX system consisting of 6 barrels located co-linearly at different z positions. The signals from the ladders are read out at both ends for a total number of 405,504 channels, which creates a significant challenge for the data acquisition system. The Layer 00 silicon system operates on the same principles as the SVX system. Introduction of the L00 amounts to adding the additional silicon wafers at a very small distance to the beam pipe (about 1.5 cm) helping further improve the resolution of the track impact parameter.

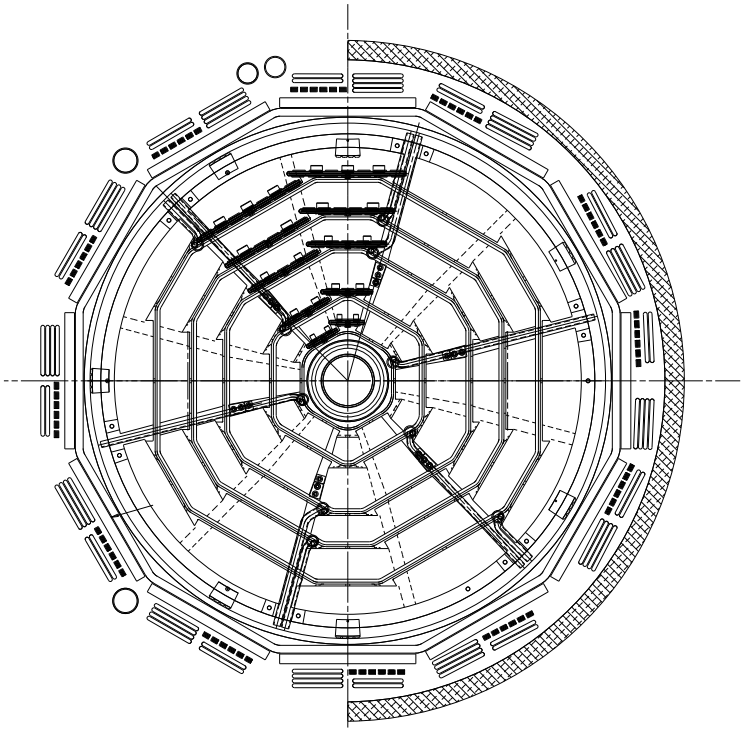


Figure 3.5: An $r - \phi$ plane cross-section view of one SVX barrel.

The presence of several tracking systems at CDF allows for a variety of ap-

proaches to determining track trajectories given a set of hits found in the COT, SVX, ISL and L00 detectors. Indeed, many tracking algorithms are used at CDF with the main differences being the order in which hits of the different types are processed. In this analysis we only use tracks that have COT information. All such tracks are created with reconstruction algorithms starting with the set of $r-\phi$ COT hits to form the first approximation for the trajectory fit (called seed) in the transverse plane. This seed is then used to attach other types of hits to the fitted track trajectory, for example $r-\phi$ SVX hits or small angle stereo COT and SVX hits.

3.2.5 Muon Chambers and Calorimeters

There are several detector systems at CDF dedicated to the lepton detection. While our analysis would not be possible without lepton identification, only some of these systems are used to aid in B meson reconstruction. We rely on a standardized set of procedures to achieve that. Therefore we will only briefly describe these systems here.

Mmuon detection at CDF is provided by four systems:

- The Central Muon Detector (CMU) provides coverage at $\eta < 0.6$ at $p_T >$ 1 GeV/c.

- The Central Muon Upgrade (CMP) provides roughly same $\phi - eta$ coverage as the CMU but is located further in r (behind CMU) and thus allows for

cleaner muon detection. The drawback is that muons have to have p_T above ≈ 3 GeV/c to reach the CMP system.

- The Central Muon Extension (CMX) and
- Intermediate Muon Detector (IMU) allows for muon detection at higher pseudo-rapidity regions (up to ≈ 0.9 and ≈ 1.5 correspondingly) and are not used in our analysis.

The muon detector systems are shielded by large amounts of detector material, which means that muons are much more likely to reach these systems than any other particle of similar momentum. The muon detectors are simple tracking chambers which only provide a few hits for any single particle. The hits left in the muon chambers can then be matched to the extrapolated COT track trajectory. In case of a match a muon object is formed from a track and a muon chamber stub (a set of muon chamber hits matched to the same track). The main disadvantage of the muon identification at CDF are the relatively large fake muon rates (compared to other similar experiments) that are due mostly to the punch-throughs (hadrons that reach the muon systems) and decays in flight[8]. For more information on the muon systems at CDF see [7, 6].

Electron detection at CDF is provided by a set of calorimeter systems. The central region (entire ϕ region and η up to ≈ 1.1) is covered by the Central Electro-

[8] (Muons that are the decay product of long-lived particles. The muons can be detected and accidentally matched to a COT track

Magnetic (CEM), Central Hadron (CHA) and Wall Hadron (WHA) calorimeters. In the higher η regions the Plug Electro-Magnetic (PEM) and Plug Hadron (PHA) calorimeter systems extend the coverage up to $\eta \approx 3.5$.

The operational volume of the calorimeters is made up of several layers of the scintillator material separated by the metal sheets. The energy deposited in the scintillator is measured using the photo-multiplier tubes.

3.2.6 Time-of-Flight Detector

The Time-of-Flight system [1] is the primary detector system used for K/π particle separation at low transverse momenta. The TOF detector measures the time it takes for the particles to travel from the collision point to the TOF system. Combining this information with the measured quantities from COT (flight length L and particle momentum p) one can estimate the particle mass as

$$m = \frac{p}{c}\sqrt{\left(\frac{ct}{L}\right)^2 - 1}.$$

The TOF detector consists of 216 bars of scintillator material of dimensions $4 \times 4 \times 279$ cm forming a thin cylinder in the space between the COT detector and the solenoid's cryostat system, as shown in Figure 3.6. Bars are located at roughly equal distances in ϕ at the radius of 140 cm and the full system covers the region $|\eta| < 1$ and the entire ϕ region.

Interaction of a charged particle with the detector material is called a hit. When a charged particle traverses the scintillator material, photons are emitted, and these

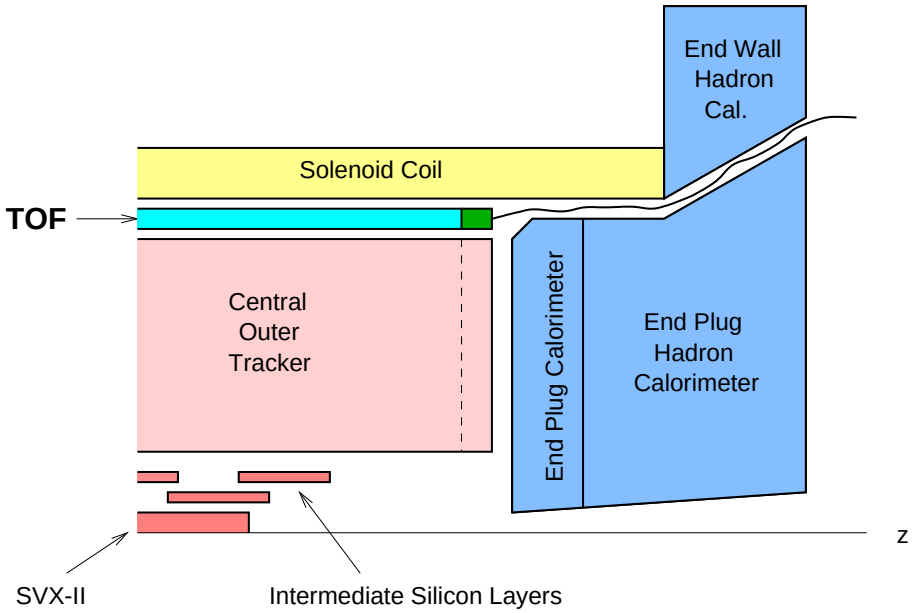


Figure 3.6: Location of TOF system in the CDF detector.

photons propagate towards the ends of the scintillator bar. The photo-multiplier tubes (PMTs) at the both ends of the bar collect the light and the front-end electronics measures the arrival time of the PMT pulse and the amount of charge collected (which is related to the number of photons produced). The physical location of the bar where the hit was registered provides the ϕ location of the hit; comparing the arrival times of the pulses at the two ends of the bar determines the z position of the hit. The COT tracks are then extrapolated to the TOF system location to find the matches between the track trajectories and the hits in the TOF system. This associates the timing information from TOF with the tracking information from COT.

Figure 3.7 shows the overview of the electronics read-out system for one TOF

channel. It starts with the photo-multiplier tube where the light is collected and converted to a charge pulse. The pre-amplifier receives a nearly differential pulse from the PMT base and drives it to the discriminator. The Pre-amplifier is designed to have bilinear gain *i.e.*, the gain is reduced for bigger pulses which increases the dynamic range and makes it possible to measure the charges from very large pulses produced by heavy slow particles. The discriminator, which has an adjustable

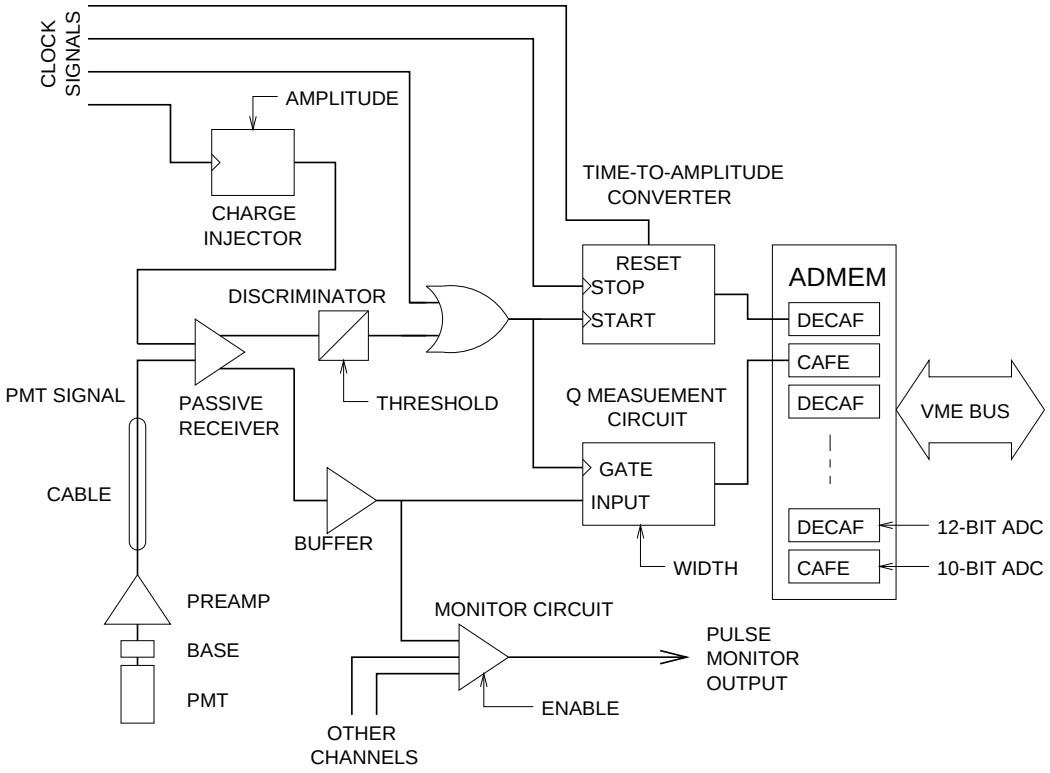


Figure 3.7: Electronic readout chain of one TOF channel.

threshold, selects the signals to be processed by effectively filtering out the noise. After the discriminator the signal path is split and the signal is sent to the time and charge measurement circuits.

The discriminator provides a start signal to the Time-to-Amplitude Converter, or TAC, which is later stopped by the CDF common stop signal. The TAC readout is converted into a digitized time reading by a 12-bit Analog-to-Digital Converter chip (ADC). The digitized time is buffered in the VME module called an ADMEM for the ultimate readout by the data acquisition system. The charge measurement circuit converts the received signal into a current that is passed to a charge sensitive ADC that is also located on the ADMEM module. The primary purpose of measuring the pulse charge is to perform the correction for the variation in the timing measurement of the pulses with varying amplitudes (so called time-slewing correction).

The TOF electronics also provides the operations for configuration, monitoring, calibration and testing of the system. A more comprehensive description of the TOF front-end electronics can be found in [17]. The timing resolution of the electronic readout itself is less than 20 ps and contributes very little to the overall timing resolution of the TOF detector.

The TOF electronics measures the time of the PMT pulse with respect to the CDF common stop signal. In order to calculate the actual flight time of the particle the direct measurement has to be corrected by the time it takes the photons to reach the PMTs, by the time-slewing correction, time it takes the signals to travel the length of the cables in the data aquisition system, and so on. Main contributions to the TOF timing resolution include physics effects such as the random variations in

the number of photoelectrons in the photo-multipliers and variations in the electron transit time inside the PMT. In addition, the variations in the shapes of the PMT pulses make a precise determination of the pulse arrival time based on the threshold crossing time impossible. Situations where more than one particle produces a hit in the scinitillator bar cannot always be detected and contribute significantly to the TOF resolution. Extrapolation of the measured particle trajectories to the TOF detector is itself imprecise and also makes a contribution to the timing uncertainty. In the end, we achieve an overall time of flight timing resolution of 100-150 ps for most particles. The separation power for the various particle species that is achieved with TOF alone is shown in Figure 3.8. For comparison, the K/π separation from the COT dE/dx measurement is also shown to illustrate the complementary power of COT with respect to the TOF particle identification.

Measuring the arrival times of charged particles with respect to the CDF common stop signal is not sufficient to deduce particle masses because the time of the $p\bar{p}$ interaction t_0 with respect to the common stop signal can vary significantly from event to event[9]. However the combined TOF timing measurements for all the particles in the event can be used to estimate the value of t_0 by assuming that these particles consist of known mixture of pions, kaons and protons and estimating the value of t_0 that is most consistent with all observed particles. This is illustrated in Figure 3.9 where an event is shown where 3 $p\bar{p}$ interactions occur during a bunch

[9]This is mostly due to the relatively large size of the proton and antiproton bunches resulting in a long time window during which a $p\bar{p}$ interaction can occur.

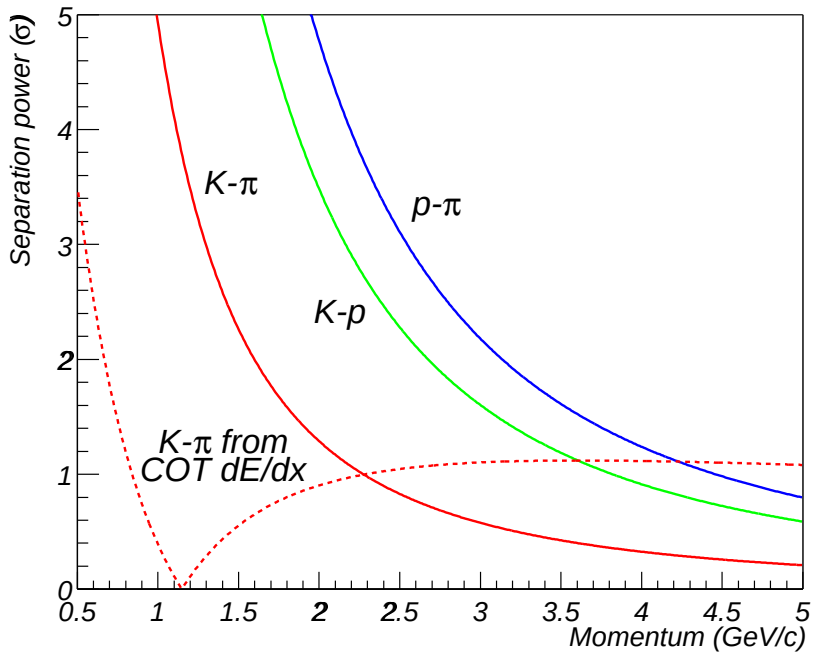


Figure 3.8: Separation power of the TOF particle identification.

crossing. The event is reconstructed in TOF and collision time t_0 is calculated for every track assuming pion hypothesis (the majority of the tracks in CDF are from the charged pion particles). The three peaks corresponding to each of the interactions can be used to determine the most likely value of t_0 for every interaction. In practice tracks are associated with one of the interaction points by finding the vertexes corresponding to the $p\bar{p}$ collision point and extrapolating the tracks to those vertexes, and TOF information is rarely needed to resolve multiple interactions. In most events the t_0 can be determined with a resolution of roughly 50 ps which is significantly smaller than the intrinsic timing resolution of the TOF measurement.

In conclusion, the combination of particle identification information from the COT and TOF detectors allows us to separate the particles based on their mass in practically all the p_T regions accessible at CDF.

3.3 CDF trigger system

The trigger system at CDF has several purposes: to provide event filtering for the reduction of the size of the incoming data; perform early data classification into several streams based on the event type; reconstruct the events online enabling high-level decision making and data monitoring.

Some experiments, especially those operating on e^+e^- machines, have trigger systems which specialize in tagging the events meaning that every event is saved to storage and the trigger system just provides a way to separate the data stream into several categories. Such an approach would be unfeasible at CDF due to the high event rate and much lower fraction of the events with interesting interactions.

The proton-antiproton collisions in the CDF detector occur at a rate of about 3MHz, with almost every bunch crossing having a $p\bar{p}$ collision. With event sizes exceeding 100 kilobytes, saving every event would present an enormous technical challenge in terms of both the bandwidth and storage capacity. Also, doing so would not be prudent because only a small fraction of the collisions produce events which are likely to be interesting in physics analyses. Fortunately, a set of signatures can be identified that physics events of interest are likely to produce. For example, a

semileptonic B meson decay would likely produce an event with a lepton and a track not coming from the primary interaction point. The strategy then is to identify events containing signatures hinting at the presence of the desired processes and save only those events, rejecting the rest.

To identify the events of interest the CDF II detector has an online trigger system built with a three-layer architecture, each layer providing different degree of sophistication in decision making and event reconstruction. Since the time to make a decision may vary significantly from event to event, all the trigger systems are pipelined meaning that while the current event is processed additional events are stored in the system waiting for their turn.

Below we give a brief overview of each trigger system. The schematic diagram of the first two level systems is shown in Figure 3.10 and is explained in the following text.

3.3.1 Level 1

The first, lowest level trigger system at CDF II is Level 1 trigger. This trigger system is designed to process the events extremely fast and provides the highest rate reduction (it accepts events at about 10-20kHz rate). It is implemented entirely in a hardware system consisting of custom design printed circuit boards (PCBs). Due to event filtering time constraints the level of event reconstruction taking place at this stage is the least sophisticated of the three levels in the trigger. Level 1

only uses information from the central tracking chamber, muon chambers and the calorimeter, and forms simple objects (called *primitives*) based on that information. The trigger can then make simple decisions, usually based on counting the numbers of objects of certain type present in the event, such as number of muons or tracks above given p_T threshold.

Two hardware systems are available in Level 1 to facilitate the formation of the primitives. A system called Extremely Fast Tracker (XFT) is capable of quickly identifying 2-D track candidates of particles with p_T above $\approx$ 1.5 GeV/c by using the information from the four axial layers of the Central Outer Tracker (COT) detector. The algorithm reconstructs the track segments in the COT layers by searching for a matching pattern of hits in a lookup table. This approach allows for very fast track reconstruction.

The XFT tracks are then used in an Extrapolator Unit (XTRP) system to extrapolate the tracks to the other sub-detectors such as muon chambers or calorimeter, allowing for reconstruction of leptons with known p_T.

3.3.2 Level 2

Every event that is accepted by the Level 1 trigger is passed on to the Level 2 trigger system. The Level 2 trigger further reduces the event rate to about 200Hz and can analyze the events with a higher degree of sophistication. At this stage information is added from the shower maximum strip (CES) detectors and from

the SVX detector. There are three additional hardware systems at this level for building primitives: L2CAL, XSEC and SVT. The L2CAL system builds energy clusters from the Level 1 trigger towers in the calorimeter. The XSEC system extrapolates XFT tracks to the CES detector creating electron candidates.

The SVT system is the most important for B Physics as it allows measurement of the track impact parameter d_0 which is sensitive to the lifetime of the mother particles in case of the secondary tracks.[10] SVT works by extrapolating the XFT tracks into the volume of the SVX detector and matching them to the $r - \phi$ hits in the SVX. This improves the p_T and ϕ measurement on the track and allows a d_0 measurement. A selection based on the large impact parameter of at least one track in an event is a basis of many heavy flavor triggers at CDF and allows fairly clean selection without requiring high p_T tracks. More information on the SVT system is available in [8].

3.3.3 Level 3

The last stage of the trigger system is Level 3. After a Level 2 accept the entire detector is read out and all the information is assembled into a data structure known as the *event record* by a system called the `EventBuilder`. The event record is then passed on to the Level 3 system. The design and implementation of the EventBuilder-Level 3 stage is presented in [27].

[10] Mother and daughter particles are the terms often used to describe a decaying particle and its decay products.

Level 3 trigger performs a nearly complete reconstruction of the event in the detector, building high level objects corresponding to leptons, tracks, etc. using the full resolution of the detector. At this point the event is analyzed and final decisions are made, reducing the event rate to around 70Hz.

In order to achieve these computation intensive goals the Level 3 trigger system is implemented as a farm of several hundred commodity computers running Linux, also called "processor" nodes. Each processor node is dedicated to reconstructing and analyzing one event at a time. A few nodes are designated for orchestrating the coherent work of the entire Level 3 system and sending the accepted events to the data handling system. Every event passed by Level 3 system is stored to tape.

3.4 Semileptonic B triggers at CDF

As mentioned in the Introduction, our analysis uses semileptonic decays of the B mesons. The events we use are selected by two of the semileptonic triggers at CDF to which we will refer as eSVT and μSVT for triggers selecting semileptonic B decays that have correspondingly an electron or a muon in the final state. We also collectively refer to these triggers as lSVT.

Due to their relatively long lifetime weakly decaying B mesons can travel as far as several millimeters in the detector before they decay. This means that the tracks of their decay products may not be consistent with these tracks originating from the $p\bar{p}$ interaction point (primary vertex) resulting in these tracks having large

impact parameters. Therefore the *l*SVT triggers require a presence of an SVT track with large impact parameter to further reduce the trigger rates. Below we explain how this is implemented within the bounds of the CDF trigger system, which also serves as a good illustration of the function of the various trigger systems described in Section 3.3.

3.4.1 *e*SVT trigger

At Level 1 only the most simple signatures can be identified due to the limited reconstruction capabilities in the limited amount of decision time. Therefore at Level 1 an event is only required to have one primitive corresponding to an electron seen in the central region with $p_T > 4\text{GeV/c}$ and ratio of hadronic to electromagnetic energies in the calorimeter $E_{HAD}/E_{EM} < 0.125$.

Such an event is passed on to Level 2 where SVX information is available. The event is then required to have at least one SVT track reconstructed in addition to the electron, and the electron and SVT track have to satisfy the following set of requirements in addition to Level 1 ones:

- transverse energy of the electron in the calorimeter $E_T > 4\text{GeV/c}$;
- p_T of the SVT track $p_T > 2$ GeV/c;
- impact parameter of the SVT track $0.120 \leq d_0 \leq 1\text{mm}$;
- χ^2 of the fit for the SVT track $\chi^2 > 15$;

- the difference in ϕ_0 between the SVT track and the electron track $2 < \phi_0 < 90$ degrees.

The p_T requirements are partially due to the limited reconstruction capabilities but are mainly present to reduce the rate and the backgrounds. Tracks with very large d_0 are rejected to reduce backgrounds from very long lived particles such as Λ, K^0 and photon conversions.

Finally, in Level 3 every event that has passed L1/L2 semileptonic triggers is analyzed again after the full detector reconstruction. All the requirements of Level 1 and Level 2 systems are applied again to the more precisely reconstructed objects. In addition, Level 3 imposes the following requirements:

- SVT track selected in Level 2 is required to have a matching track in COT, where matching is based on ϕ_0 and p_T parameters;
- the invariant mass of the electron and SVT track combined has to be less than 5 GeV/c (SVT track is assumed to be a pion; the value is chosen based on the Monte-Carlo simulation);
- at least 3 towers with electro-magnetic energy found for an electron;
- $E_{HAD}/E_{EM} < 0.080$;
- $L_{shr} < 0.2$;
- CES ${\chi_x}^2 < 15$;

- CES ${\chi_z}^2 < 10$;
- CES $\Delta x < 2$ cm;
- CES $\Delta z < 3$ cm.

The last 5 requirements are based on the extrapolation of the COT track to the CES detector, x and z are the coordinates in the CES coordinate systems and the extrapolation procedure performs a fit based on the distance between the extrapolated track position and the hits in the CES strips - selection is then based on this distance and its χ^2 in both dimensions.

3.4.2 μSVT trigger

The logic of the μSVT trigger is the same as for the eSVT trigger, but due to the differences in the hardware of these systems the kinematic cuts were not always identical.

At Level 1 this trigger requires a presence of the central muon primitive with p_T of the associated XFT track of at least 4GeV/c. At Level 2 a displaced SVT track is added with a requirement $0 < \phi_0 < 90$ - analogous but not identical to eSVT trigger. The Level 3 system confirms L1/L2 decision and imposes the following selection criteria on the muon:

- trigger muon has hits in CMU and CMP detectors;
- trigger muon's best matching track has $p_T > 4$ GeV/c;

- CMU $\Delta x < 15$ cm;
- CMP $\Delta x < 40$ cm.

All the Level 3 requirements concerning lepton-SVT track pair are the same as in the eSVT trigger. The quantities Δx in CMP and CMU are defined similarly to the CES quantities described above and represent the distance between the hits in the muon chamber and extrapolated position of the matched track.

3.5 Data selection

3.5.1 Standard CDF data processing

The events accepted by the trigger system and stored to tape are not optimal for the physics analyses because the up-to-date calibrations of the detectors are usually not available for online event reconstruction. A standard software package `Production` exists at CDF for final quality reconstruction of the data offline. It not only builds the highest precision physics objects in the event record (based on the best available detector calibrations, beam line position measurements and so on) but also separates the data into different datasets based on the triggers that accepted each and every event. In our case, events that passed the eSVT trigger were added to the `xbel0d` dataset and events accepted by μSVT trigger are part of the `xbmu0d` dataset. The datasets created by `Production` are what physicists typically use to perform their analyses. The version of the `Production` package

used for reconstruction of the data we analyze is 5.3.3.

3.5.2 B group data processing standards

The procedures outlined above are standard for anyone performing analyses at CDF. The B physics group has its own set of guidelines that have been applied to this analysis.

B group has a standard set of data periods during which the data quality is considered acceptable for physics analyses. Such data periods have to be identified by each physics group and a single set for the whole collaboration would not be appropriate. For example, if the SVX detector were not operational during a period of time (due to beam conditions), some top quark analyses could still be performed successfully while most of the B meson sensitive triggers would not function. On the other hand if the calorimeter system were non-functional some of the hadronic B analyses would not suffer.

The data used in this analysis were selected according to those guidelines. These data were taken at CDF from April 2001 to August 2004 and they correspond to a total integrated offline luminosity of approximately $355 pb^{-1}$.

The semileptonic B triggers were described here as they were implemented at the time of writing this dissertation, but historically some of the trigger selection criteria have changed. Most of the effects introduced by these historic differences are taken out by using a standard selection tool `LeptonSvtSel`. This insures data uniformity

and makes the detailed description of the most current trigger selection relevant. Using `LeptonSvtSel` is a recommended procedure for treating the semileptonic dataset for B analyses. The software package `LeptonSvtSel` confirms the trigger decision offline. First the trigger requirements are applied to online objects (SVT tracks) to make sure the event would be accepted by the most current version of the trigger. Then the same requirements are applied to the reconstructed offline objects (i.e. SVX tracks are used that have a matching SVT track used in making the trigger decision) to make sure the trigger requirements are still met.

Events accepted at this last stage of standard data processing are the data sample used in the analysis described here. This concludes the description of the experimental apparatus and the data processing techniques that are common for most analyses performed at CDF.

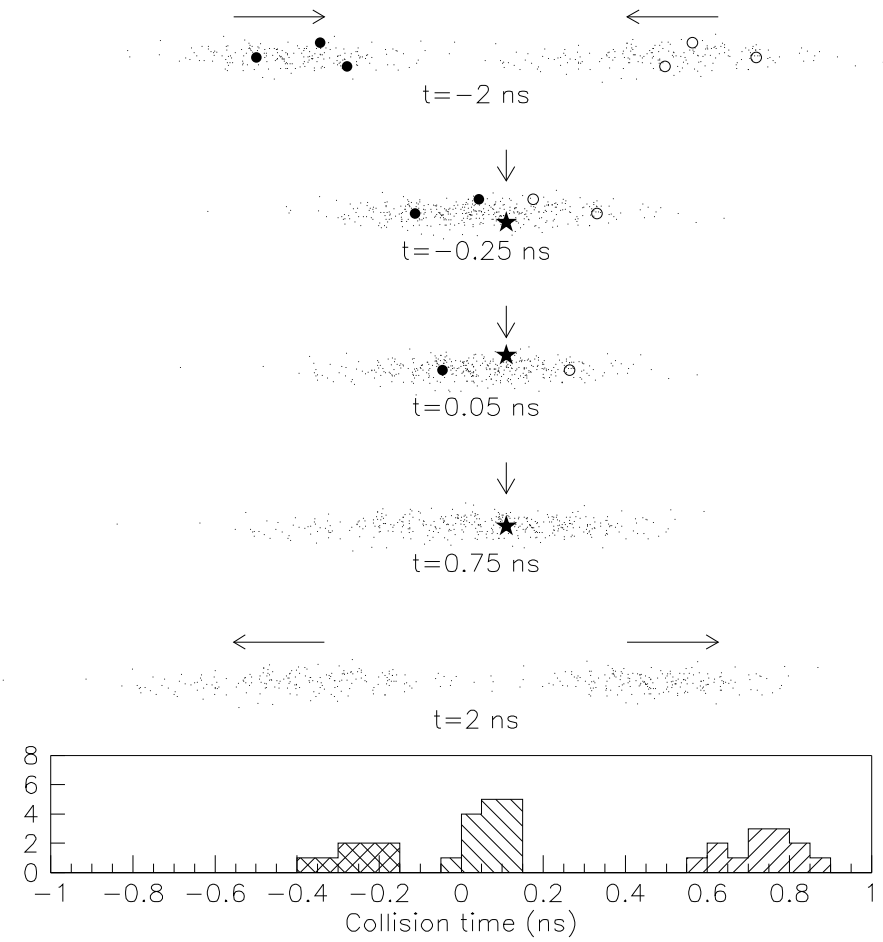


Figure 3.9: An illustration of event t_0 calculation using TOF. Closed and open dots represent the protons and antiprotons in the Tevantron bunches, each star indicates a $p\bar{p}$ collision occuring at some time during bunch crossing. Each interaction corresponds to a peak in a histogram of t_0 times calculated using TOF information.

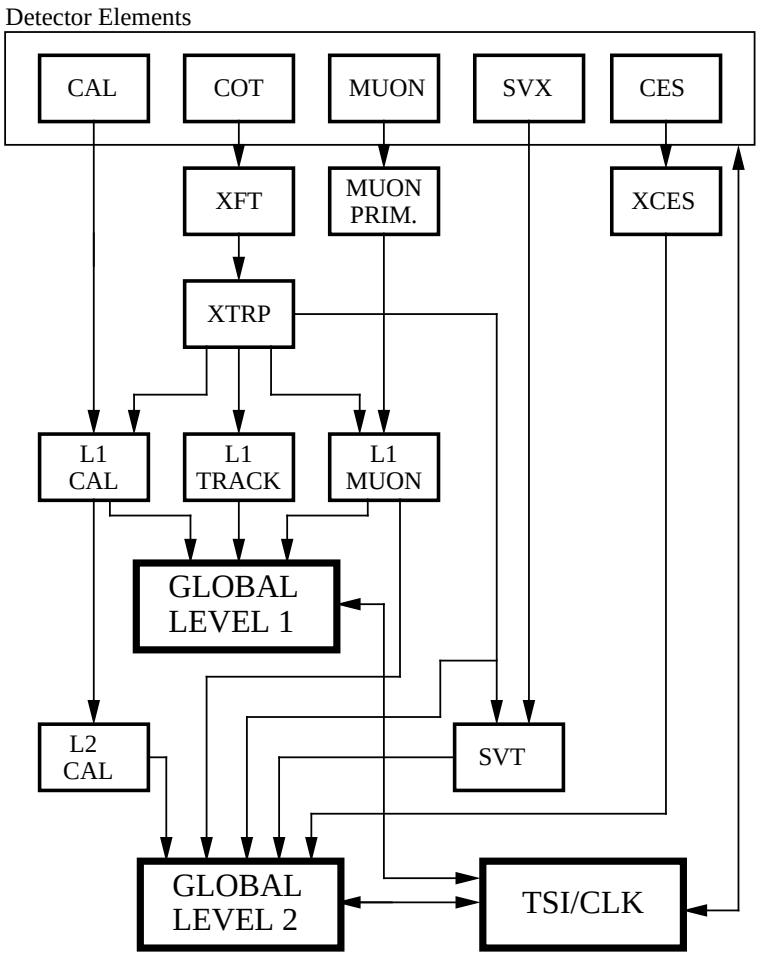


Figure 3.10: Structural diagram of the L1/L2 trigger system at CDF II.
PJW 9/23/96

Chapter 4

Event reconstruction and selection

This Chapter focuses on the aspects of the data analysis and processing that are specific to this dissertation research. We describe the selection and reconstruction of the B meson candidates and the procedure used to collect a sample of tracks that are associated with the production of the B mesons.

4.1 B candidates selection

As already mentioned, in this analysis we use the samples of partially reconstructed semileptonic B decays because they provide the highest statistics. Four modes were selected which provide the highest event yield:

$B^0 \to l^+ D^- X, D^- \to K^+ \pi^- \pi^-$;

$B^0 \to l^+ D^{*-} X, D^{*-} \to \pi^- \bar{D}^0, \bar{D}^0 \to K^+ \pi^-$;

$B^+ \to l^+ \bar{D}^0 X, \bar{D}^0 \to K^+ \pi^-$;

$B_s \to l^+ D_s^- X, D_s^- \to \pi^- \phi, \phi \to K^+ K^-$.

Since each mode has to be analyzed individually, and because the particle fraction fits can only be performed reliably on samples with at least few hundred signal events (the reasons will become clear in Chapter 6), there is no benefit from using additional lower statistics decay modes. The purpose of this Chapter is to describe the procedure used to find and select the lD candidates and associated tracks for further analysis.

The reconstruction criteria used in this analysis are the same as employed in [19] and [20], mainly to make sure that the sample composition calculations and event yields can be standardized and easily cross-checked.

An event topology of a typical event containing a semileptonic B decay is shown in Figure 4.1. A point of $p\bar{p}$ interaction is the primary vertex of the event and is the point of origin of the B meson. B meson travels a short (several millimeters) distance in the detector before weakly decaying into a lepton and D meson. D meson then undergoes a weak decay itself. The vector between the primary vertex and B decay vertex (length of B flight path) is a useful quantity related to the B lifetime. The projection of this vector on the B meson transverse momentum vector is commonly denoted L_{xy} and it assigned a positive sign if this projection points in the same direction as the B momentum. A related quantity is the proper decay time ct which is the travel time of B meson from the primary vertex to its decay point. It is assigned the same sign as L_{xy}.

In this analysis we only attempt to reconstruct hadronic decays of the D mesons with the decay products of the D meson being kaons and pions. These D meson decay products will often have large impact parameter due to long combined lifetime of the B and D mesons. The semileptonic triggers are designed to identify a lepton from the B decay and an SVT track from the D decay.

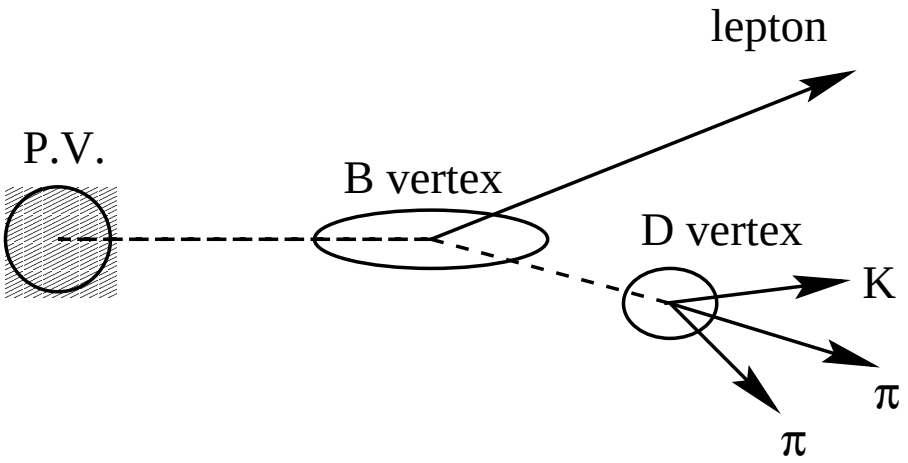


Figure 4.1: Decay topology of a typical $B \to lD$ event. B meson originates at the event primary vertex (P.V.) and decays into a lepton, a neutrino (unreconstructed and therefore not shown) and a D meson. D meson travels a short distance before decaying into a number of secondary particles (usually kaons and pions).

A software package `BottomMods`, which is commonly used in the B physics group at CDF, has been used to select lD candidates by applying very loose reconstruction criteria. The procedure used by this software can be briefly outlined as follows: first a set of tracks is found that are candidates for the decay products of the secondary decays (usually D meson decay) by taking all combinations of tracks and calculating whether they are compatible with a single point of origin [1] called a *vertex*. If a

[1]This is called vertex fitting and a `FORTRAN` collection of routines is used called `CTVMFT` to

vertex is found and the invariant mass of the track combination (assuming some particle mass assignment) is within reasonable bounds of the rest mass of the decay particle(D meson) then a candidate particle is created.

A similar procedure is then performed to find a (possible) single point of origin of the trigger lepton and the trajectory of the reconstructed D candidate. After both decay vertexes are found, the selection of the candidates is usually performed on quality criteria such as vertex fit χ^2[2], invariant mass window of the track combinations (requirements are imposed on masses of both B and D decay products), and reconstructed flight path of the decay particles. As mentioned above, the events accepted by the `BottomMods` package as the ones containing lD candidates of interest have very loose selection criteria, and additional filtering is done at the analysis level.

When a particle travels through the volume of the tracking chamber it loses some energy through material interaction, therefore its momentum changes and the trajectory is not a perfect helix. Since energy loss is particle type dependent[3] the tracking algorithm assumes a pion hypothesis when fitting the track trajectory and extracting particle momentum. During the decay reconstruction procedure, every

perform the fit itself.

[2]To form a χ^2 of the vertex fit one has to use the uncertainties on the track parameters calculated using one of the tracking algorithms and the distance of closest approach of each track to the fitted vertex position.

[3]Strictly speaking, energy loss of a particle, or dE/dx, is determined only by its velocity and the type of the material.

decay track is associated with a particle species, i.e. if for example a $D^0 \to K\pi$ decay is reconstructed for a real D^0 meson then kaon and pion are correctly identified. When it is possible to assign a particle type in this way, an energy loss correction is applied to determine the correct values of momentum and other track parameters. This allows for more precise reconstruction of the lD decays.

Table 4.1 gives an overview of our final selection requirements. The Pre-selection requirements that are not listed in the table are that at least 3 $r\phi$ SVX hits are found on all the decay tracks(not applied to the soft pion in lD^{*-} reconstruction) and that an SVT trigger track found by the L2/L3 trigger is one of the D meson decay products. This helps remove bad quality tracks and ensures that trigger decision is made on the B candidate[4].

These selection requirements were optimized to produce the best ratio of $\frac{S}{\sqrt{S+B}}$ although this particular choice of the optimization criteria is not important for this analysis.

The selection is performed on the variables such as:

- the χ^2 and the probability of the vertex fit.
- The distance between the primary vertex and D decay point in the $r-\phi$ plane, denoted L_{xy}. Most backgrounds are short-lived and produce combinatorial candidates with very small values of L_{xy} that are symmetric about 0, so

[4]This is important for example for generating Monte-Carlo samples which emulate the trigger behavior.

Cut	lD^0	lD^+	lD^{*+}	lD_s
B vertex prob. $>$	10^{-6}	10^{-5}	10^{-6}	10^{-7}
$\chi^2_{xy}(D) <$	10	10	10	10
$L_{xy}/\sigma L_{xy}(D \to PV) >$	6	11	6	5
$L_{xy}(D \to PV)$(cm) $>$	0.02	0.04	0.02	0.02
$p_T(K) >$(GeV/c)	0.4	0.7	0.4	0.6
$\sigma_{ct}(B)$(cm) $<$	0.04	0.04	0.04	0.04
$m(lD)$(GeV/c^2) $>$ to $<$	2.3 to 5.0	2.3 to 5.0	2.3 to 5.0	2.3 to 5.0
$ct(D)$(cm) $>$ to $<$	-0.015 to 0.1	-0.015 to 0.2	-0.015 to 0.1	-0.015 to 0.1
$p_T(\pi_*)$(GeV/c) $>$	-	-	0.2	-
$(m(D^*) - m(D))$(GeV/c^2) $>$ to $<$	-	-	0.1435 to 0.1475	-
$\lvert m(KK) - m(\phi_{PDG})\rvert$(GeV/$c^2$) $<$	-	-	-	0.0075
$\chi^2_{r\phi}(\phi) <$	-	-	-	30
$\lvert cos(\psi)\rvert >$	-	-	-	0.2

Table 4.1: Selection criteria on the lD samples.

requiring positive value of L_{xy} has a lot of power in background rejection.

- Significance of the above variable $L_{xy}/\sigma L_{xy}$ is strongly correlated with the L_{xy} requirement but makes it possible to improve selection on very well measured vertexes.

- The measured proper decay time of the D meson ct_D. Again, this is correlated with the two above requirements.

- Uncertainty on the measured proper decay time of the B meson $\sigma(ct)$. This helps remove poorly reconstructed candidates.

- Invariant mass of the combined tracks in the decay. Obviously it should not exceed the mass of the B meson. Monte-Carlo also shows that with our trigger configuration and kinematic selection criteria the distribution of this mass lies mainly above 2.5 GeV/c. This motivates the requirement $2.3 < M(lD) < 5.0$ GeV/c.

Other variables are decay specific. For lD^{*-} they are the minimum transverse momentum $p_T(\pi_*)$ of the soft pion π_* from the D^{*-} decay, and the difference in invariant masses of the tracks forming D^0 and D^{*-} decays. The latter is a very powerful selection criteria and explains why lD^{*-} decays have much less background contamination.

The additional selections in the lD_s mode are made on the quality of the reconstructed $\phi \to KK$ decay where the requirements are made on the χ^2 of the vertex

fit and on the invariant mass of the KK pair (it has to fall into a narrow window around the PDG [24] value of the ϕ mass). The helicity angle in the $D_s \to \phi\pi$ decay is another powerful selection criteria and is used in the lD_s candidate selection. It is defined as the angle between the D_s momentum direction and the direction of the K mesons from the ϕ decay, calculated in the rest frame of the ϕ particle.

Since the semileptonic decays of the B mesons include a neutrino as one of the decay products, and these neutrinos are practically undetectable in CDF, the rest mass of the B meson is impossible to reconstruct completely. Therefore we use the distribution of the reconstructed invariant mass of the D mesons to compute the signal yields and subtract the combinatorial background effects. The invariant mass peaks corresponding to the reconstructed decays are shown in Figures 4.2 and 4.3, and Table 4.2 shows the number of signal events reconstructed in μSVT and eSVT samples and their signal to background ratios. These numbers are given here only

	μ SVT		e SVT	
	Yields	S/B	Yields	S/B
lD^-	30100 ± 293	1.6	21520 ± 257	1.8
lD^{*-}	12880 ± 127	18.1	9800 ± 113	16.4
lD^0	55780 ± 374	3.0	43080 ± 356	2.9
lD_s	2540 ± 72	2.1	1730 ± 57	2.4

Table 4.2: Number of signal events and S/B for the selected lD samples.

to illustrate the quality and size of our data sample.

In order to calculate these quantities the invariant mass distribution of the D mesons was fit with two Gaussian distributions centered at the same mean value

for signal events and a first degree polynomial for the background. The region in which S/B is calculated is centered around the mean value of the Gaussian fits and its width is three times the weighted average of the standard deviations of the two fitted Gaussian curves. This region is indicated in Figures 4.2 and 4.3 by two vertical lines.

Finally, every reconstructed candidate has been assigned a weight by which it contributes to the measurements described in this document, so that the measurements are effectively performed on the samples of B mesons all having the same p_T spectra. The discussion of the reasons for this is more appropriate elsewhere, and Section 7.3 expands on this topic in detail. Figure 7.1 shows the template of the B meson p_T that was used in this analysis.

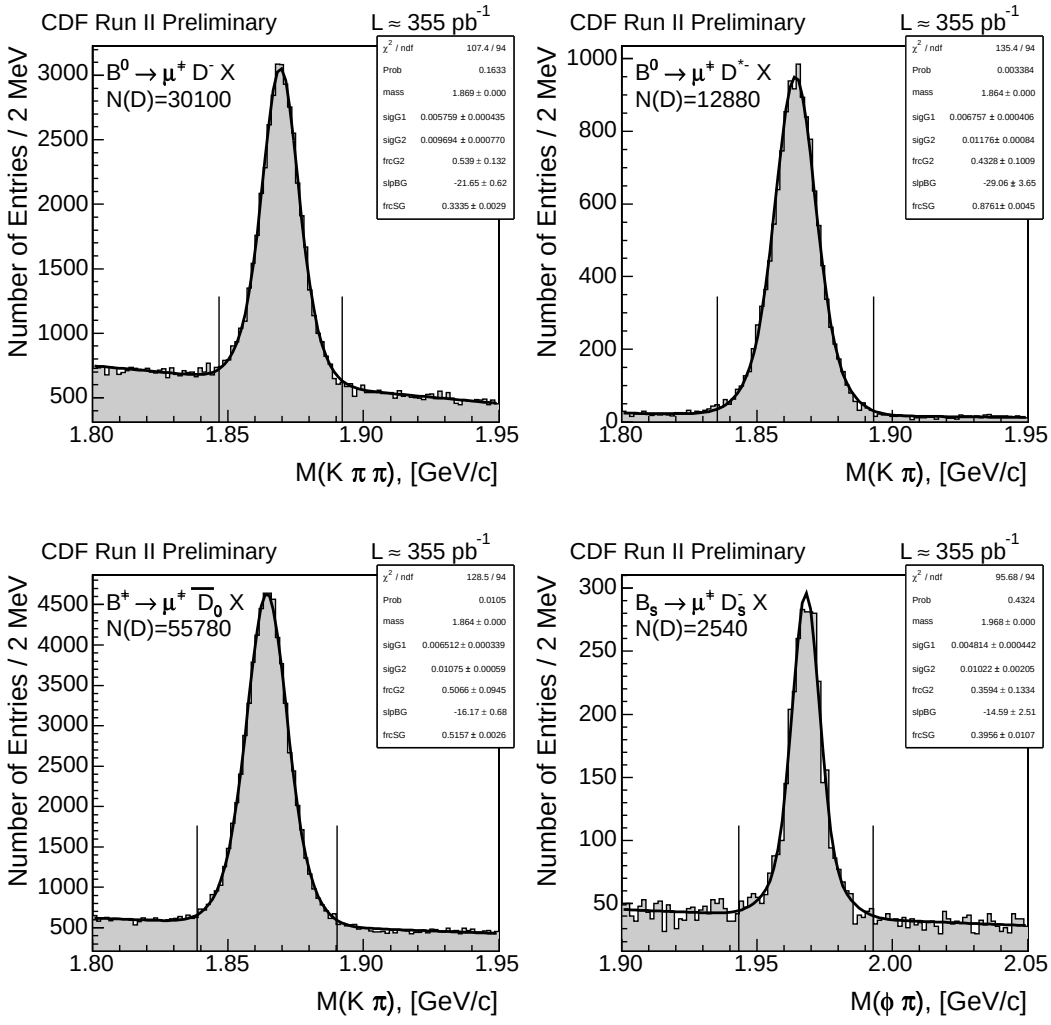


Figure 4.2: Invariant mass distributions for reconstructed D states in the partially reconstructed semileptonic signals from μSVT trigger sample.

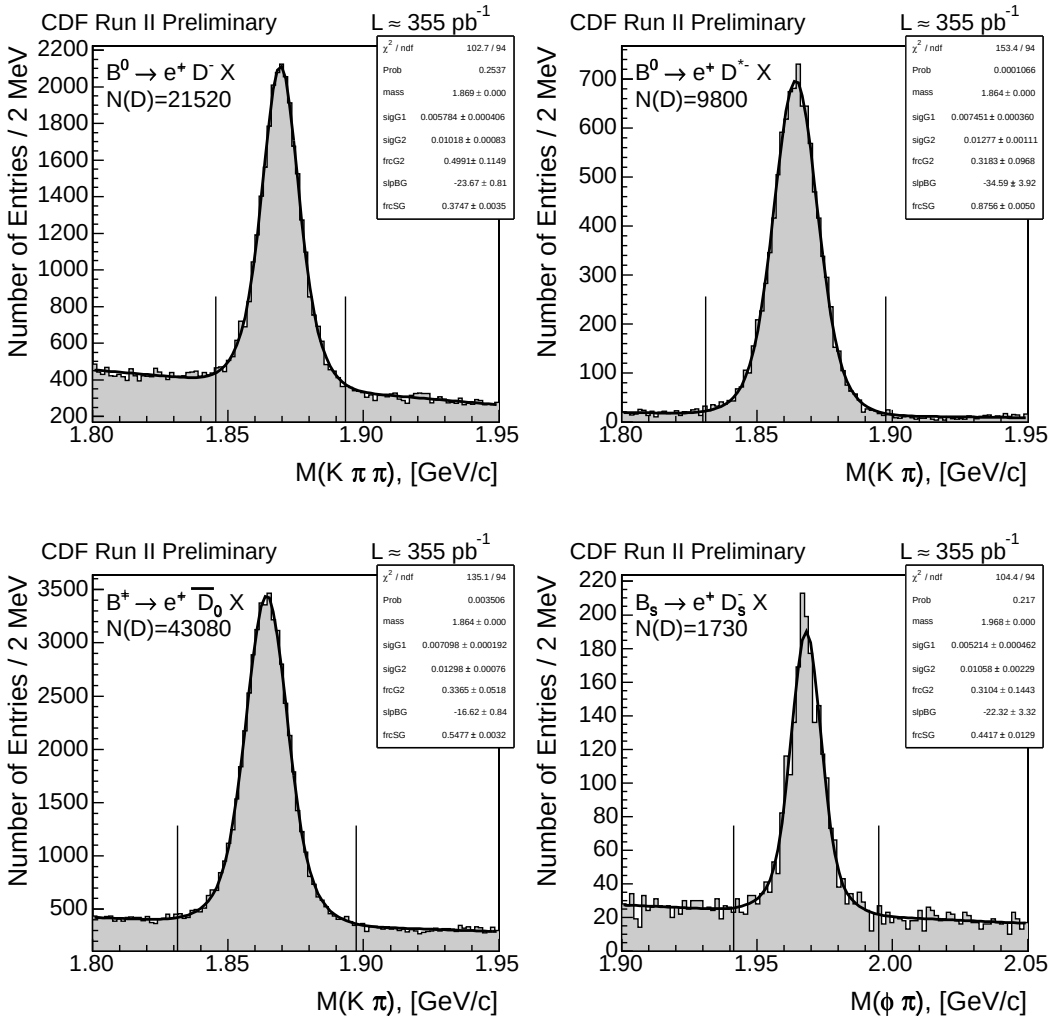


Figure 4.3: Invariant mass peaks for reconstructed D states in the partially reconstructed semileptonic signals from eSVT trigger sample.

4.2 Selection of particles associated with B formation

In order to select a set of tracks associated with B meson formation we first reconstruct the semileptonic B modes as described in Section 4.1 and analyze all the candidates in the event. Of our primary interest are the particles associated with b quark fragmentation and the decay products of the strong decays of excited B^{**} mesons, everything else is considered a background in our sample. The tracks associated with these particles are characterized by small impact parameter (as they should originate from the same primary vertex as the B meson itself) and this is the basis of our selection criteria.

In summary, after a B decay candidate has been selected we look for all the tracks in the event that satisfy the following requirements:

1. The track has not been identified as a decay product from a B decay.

2. The track is within the cone $\Delta R < 0.7$ of the direction of lD momentum. The track ϕ_0 is taken at the point of closest approach of the track's helix to the event primary vertex to calculate the value of $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$. The motivation for the ΔR requirement is the assumption that if one u, d or s quark from a quark pair formed a B meson then the particle formed by the other quark would have a correlated direction of the momentum (and thus be on the *same side* of the event). The value of the ΔR requirement

is motivated by the studies[37] done at CDF showing that b fragmentation products are contained in a cone 0.6 around B meson. In these studies the distribution of $\Delta\phi$ between the fully reconstructed decay and the other tracks in the event was compared for decays of W bosons and B mesons. In the first case the distribution was flat (it included no fragmentation tracks, underlying event only) and in the second case the distribution peaked around B meson direction. The peak was contained in a region $|\Delta\phi| < 0.6$ therefore we choose a value of 0.7 for our ΔR requirement. This is illustrated in Figure 4.4 where the distribution of $\Delta\phi$ between the lD candidate momentum and some other particle momentum is shown for all particles found in an event.

3. The track has at least 3 $R-\phi$ hits in the SVX detector. This requirement helps ensure good track quality, as well as to reject particles not associated with the event primary vertex. Many secondary particles from long lived decays would originate outside or in the middle of the SVX detector volume and have much lower probability of passing this selection requirement.

4. The z_0 of the track is within 3cm of the trigger lepton z_0. This requirement is necessary to ensure that the track comes from the same primary interaction point as the B meson. More than one $p\bar{p}$ collision can occur in a sinle bunch crossing. In addition to the event primary vertex where the B candidate originates, additional vertex(es) may be present. This requirements helps reduce backgrounds from these additional $p\bar{p}$ interactions.

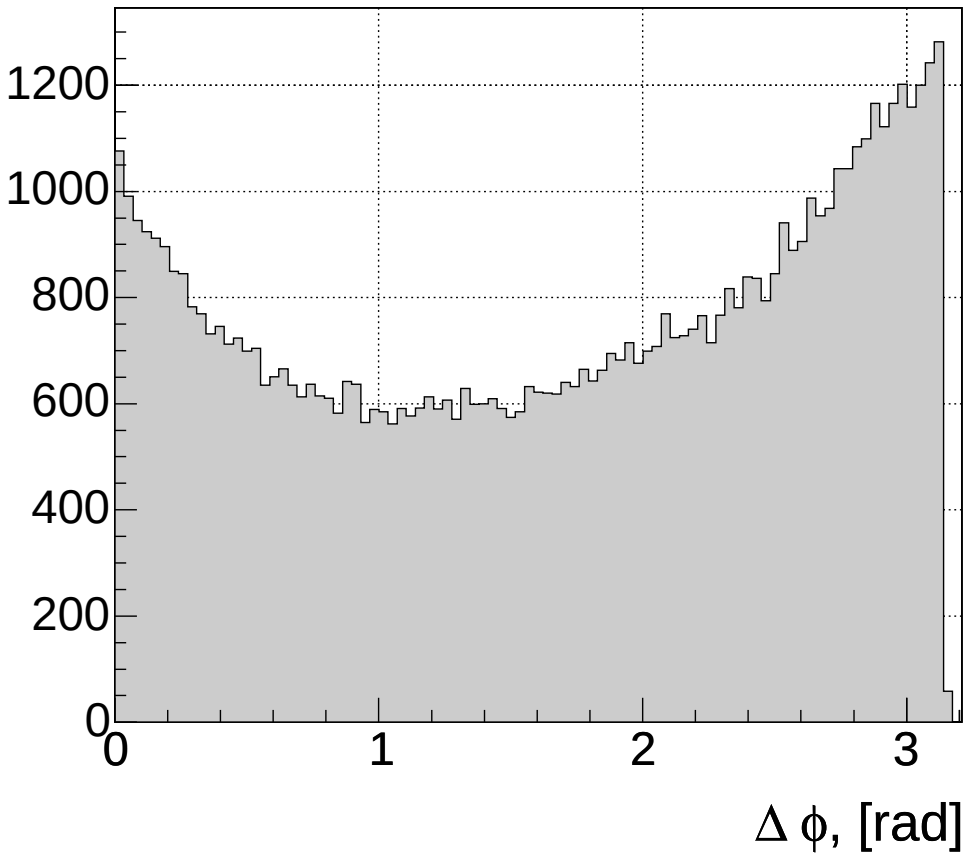


Figure 4.4: The distribution of the angle $\Delta\phi$ between the lD candidate momentum and the momentum of all particles found in an event.

5. The impact parameter significance for the track satisfies $|\sigma d_0/d_0| < 4$. This is used to reject secondary particles. The value was chosen after fitting the track impact parameter significance with two Gaussians and roughly corresponds to 3σ region of the narrow Gaussian. This is illustrated by Figure 4.5. The

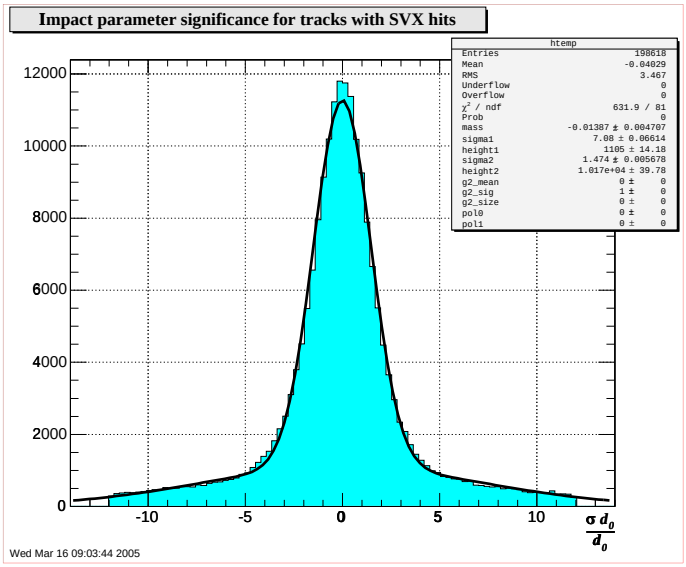


Figure 4.5: Impact parameter significance for the tracks around μD^- signal that were not identified as B decay products. Tracks are required to have at least 3 $r\phi$ SVX hits. The histogram is fitted with two Gaussian distributions. Based on the fit we chose to select the tracks associated with B formation by imposing a requirement of $\frac{\sigma d_0}{d_0} < 4$.

narrow peak represents the prompt tracks that we are interested in, and the wider distribution underneath the narrow peak represents secondary particles (decay products of relatively long lived particles) and possibly particles from $p\bar{p}$ interactions other than the one from which B meson originates.

6. TOF, dE/dx information is available for the track. This is necessary for particle identification and implies that the track has also passed the loose quality criteria imposed by the TOF reconstruction.

7. Track transverse momentum is above 400 MeV/c. This is the lowest momentum to which TOF and dE/dx were calibrated.

The sample of tracks selected according to the above criteria for a particular type of lD decay is expected to be enriched in tracks produced in association with the B formation. This is the sample that is used in the analysis to measure the particle content.

Chapter 5

Sample composition of the reconstructed lD candidates

5.1 Introduction

So far we described how the lepton-charm candidates are reconstructed and selected. The selected sample includes a big contribution from combinatorial background, which consists of random track combinations mimicking the signatures of the B decays. These have an invariant mass distribution that is very different from that of the true B decays. These backgrounds therefore can easily be dealt with by using statistical methods such as sideband subtraction. The technique we use to remove the effect of the combinatorial background is described in Chapter 6.

This Chapter is devoted to a discussion of a different issue. When dealing with

semileptonic decays one cannot fully reconstruct the invariant B mass because the neutrino for all practical purposes cannot be detected. Therefore if any other decay products are also not included in the reconstructed decay chain we do not know the missing decay products are there. A decay of a particular B meson will not always produce the same reconstructed D meson associated with a lepton because of the possibility that the B decays through the various unreconstructed excited D meson states. For example, the following two decay chains:

$$\begin{aligned} B^+ &\rightarrow \nu\ell^+\overline{D}^0 \\ B^0 &\rightarrow \nu\ell^+ D^{**-}, \quad D^{**-} \rightarrow \overline{D}^0\pi^-_{**} \end{aligned}$$

could both be reconstructed as valid $l\bar{D}^0$ candidates, if the soft pion π^-_{**} in the second decay is lost in the reconstruction [1]. What this means is that our partially reconstructed lD decays do not correspond to pure B meson signals but rather represent a mixture of the decays of B meson species.

Obviously this effect will have to be taken into account to correct the measurements performed in the selected data sample, if we want these measurements to represent pure B meson samples. This Chapter will explain the general approach to estimating the sample composition and then describe a technique of using our knowledge of B sample composition to convert the measurements of particle production around lD signals to the ones corresponding to the B signals.

[1] And in fact it will always be lost, since we do not even attempt to look for it due to relatively low efficiency of finding them. We however attempt to find the soft pion from D^{*-} decay and the efficiency of finding it is denoted $\epsilon(\pi_*)$.

This analysis is by no means the first to deal with such issues. The techniques that we use to estimate the composition of the samples of reconstructed lD candidates were first used in [37] during CDF Run-I and recently also employed by CDF Run-II analysis [19]. We base our approach to the sample composition issue on the procedures outlined in these analyses and extend it to suit our needs.

5.2 Overview of the sample composition analysis

Figure 5.1 shows the decay paths through which the $\bar{B}^0, B^-$ mesons can decay into the final D^0, D^+ states. If one could somehow determine all the branching ratios of the decays involved in these transitions, and the trigger and reconstruction efficiencies of all the decay paths, then the composition of the lD samples could be computed analytically by estimating the contributions of every decay path in the reconstructed signature. This is the essence of the procedure we use in this analysis. The parameters necessarily involved in such a calculation are

1. the relative production of the $\bar{B}^0, B^-$ and $\bar{B}^0_s$ mesons in the $p\bar{p}$ collisions. The current world's best measurements [24] show that $\bar{B}^0, B^-$ are produced at the same rate (with a precision of $\approx 3\%$) and the $\bar{B}^0_s$ mesons are produced at a rate of about 0.269 ± 0.035 of the rate of the $\bar{B}^0, B^-$ meson production.

2. The relative branching ratios for B mesons to decay semileptonically into lD^{**}, lD^*, or lD state, which are denoted f^{**}, f^* and f correspondingly. Since

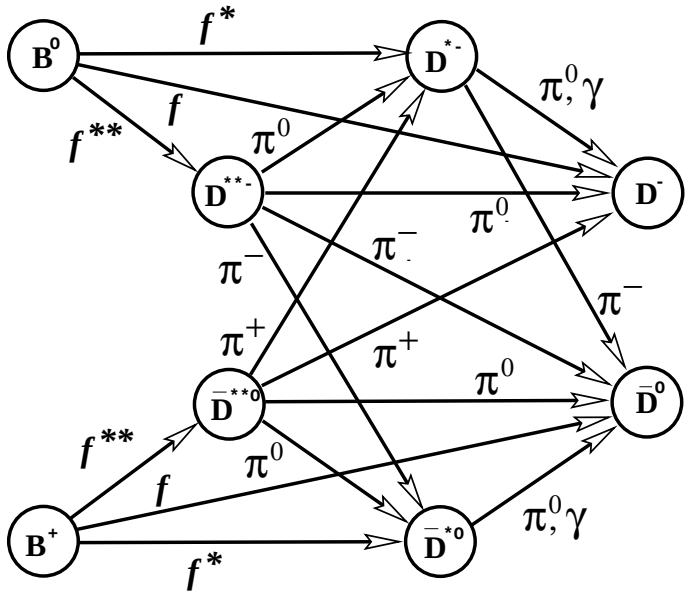


Figure 5.1: The diagram illustrating the possible decay paths of $\bar{B^0}, B^-$ mesons leading to final D^0, D^+ states in the semileptonic channels. The decay of the B meson contains a lepton and a neutrino not shown here.

excited D mesons always decay strongly they all ultimately lead to one of the lD candidate signatures and therefore $f^{**} + f^* + f = 1$. The parameters $R_f = f^*/f$ and f^{**} are often used as two independent variables. The current best estimates for the values of these quantities are $R_f = 2.14 \pm 0.14, f^{**} = 0.31 \pm 0.005$ (see discussion in the rest of this Chapter).

3. The probability for D^{**} mesons to decay via D^* (as opposed to D) state, denoted P_V. This parameter is poorly known, and current world's best measurement is $P_V = 0.627 \pm 0.14$.

4. The ratio of the semileptonic branching ratios of $\bar{B^0}, B^-$ mesons (it can be shown that this ratio is equal to the ratio of the lifetimes of these mesons $\tau^+/\tau^0 = 1.086 \pm 0.017$).

5. The branching ratios of the decays of the various D^* mesons. Most of these are known with the absolute uncertainty of about 0.01.

6. The efficiency of finding the soft pion from the $D^{*-} \to \bar{D^0}\pi^-_*$ decay, here used as $\epsilon(\pi^-_*) = 0.579 \pm 0.007$. This parameter is directly related to the migration of the candidates between the lD^0 and lD^{*-} samples. This efficiency is measured in the data.

7. The trigger and reconstruction efficiencies for all the decays paths involved, these were measured in realistic [2] Monte-Carlo. Monte-Carlo samples were generated using the `BGen` [5] generator and decayed using `EvtGen` [35], with each decay path having a dedicated sample generated.

Except for the last two pieces of information in the list above, all parameters were derived from the current world-best measurements of the branching ratios of the meson decays [24]. This is just a short overview of the sample composition parameters and an in-depth description of them is provided later in the Chapter.

[2] Realistic in this context means that the detector behavior was simulated at the lowest possible level. Usually it means that the Monte-Carlo first simulates the interactions of the particles with the detector material and then estimates the detector response based on that information, typically using the standard data reconstruction procedures.

The reader may have noticed that the $\bar{B}_s^0$ decays were absent from the above discussion. They are somewhat decoupled from this description of the sample composition. First, the background from $\bar{B}^0, B^-$ mesons in the lD_s mode is dominated by the decays of $B \to D^{(*)}D_s^{(*)}$ with the subsequent semileptonic decay of the $D^{(*)}$ meson; these are fundamentally different in the sense that they are not due to the unreconstructed $D^{*(*)}$ states within the semileptonic B decay chain. These backgrounds can be fully determined from the respective branching ratios of these decays and their reconstruction efficiencies and are roughly equal for $\bar{B}^0$ and B^- cases.

In the other direction, the background from $\bar{B}_s^0$ into the $\bar{B}^0, B^-$ modes occurs through the mode $\bar{B}_s^0 \to lD_s^{**}\nu$ with subsequent $D_s^{**} \to D^0K^-$ and $D_s^{**} \to D^{-(*)}K^0$ decays. Monte-Carlo studies show these backgrounds to be very small (of an order of a percent or less), partly because of the lower production rate of the $\bar{B}_s^0$ mesons. For this reason these two modes are not shown in the Figure 5.1.

Since the $\bar{B}_s^0$ backgrounds in the $\bar{B}^0, B^-$ modes are small, they can be ignored in the error propagation calculations (Section 5.6) while the same isn't true for the parameters of the $\bar{B}^0, B^-$ crosstalk[3].

For the above reasons, it was found easier for the $\bar{B}_s^0$ associated backgrounds to estimate them in the inclusive semileptonic B Monte-Carlo sample where the appropriate branching ratios were set to the world best values [24] in `EvtGen` [35].

[3]It will become clear by the end of the Chapter, but in essence the analytical dependence of the $\bar{B}^0, B^-$ sample composition from the input parameters is needed to properly estimate the error analysis and corrections of the results.

In conclusion, having the values of all the above parameters one can list all the decay chains of the B mesons and calculate their contributions to each of the lD modes. Table 5.1 shows these contributions in terms of the branching fractions and soft pion reconstruction efficiency (this parameter is necessary there because it directly affects the migration of lD^0 candidates into the lD^{*-} mode). The contribution of every decay path is calculated simply as the product of the relative branching ratios of all decays involved in said path, times the efficiency $\epsilon(\pi_*^-)$ where needed.

Table 5.1 illustrates the analysis components specific to the decay path. Every decay path contribution also needs to be multiplied by the semileptonic branching ratio of the corresponding B decay, since these are not the same for $\bar{B}^0$ and B^- mesons. This and other branching ratio inputs are discussed in Section 5.4. The decay path contribution also depends on the efficiency for this particular decay path to be found in the event, i.e. the contributions of the paths in Table 5.1 need one additional factor, and that is combined detector/reconstruction efficiency for a path. These efficiencies are discussed in Section 5.3.

5.3 Detector and reconstruction efficiencies used in the sample composition analysis

We have already introduced the notion of a *direct decay* chain which is by definition a B decay path which is reconstructed in the lD candidate as fully as possible, i.e.

B^0 Decay Chains	Decay Signatures ℓD^{*-}	ℓD^-	ℓD^0
$\to D^{**-}\ell^+\nu$			
$D^{**-} \to \overline{D}^{*0}\pi^-_{**}$			
$\overline{D}^{*0} \to \overline{D}^0\pi^0_*$	—	—	$f^{**}P_V(2/3)\mathcal{B}^*(D^0\pi^0_*)$
$\overline{D}^{*0} \to \overline{D}^0\gamma$	—	—	$f^{**}P_V(2/3)\,\mathcal{B}^*(\overline{D}^0\gamma)$
$D^{**-} \to D^{*-}\pi^0_{**}$			
$D^{*-} \to \overline{D}^0\pi^-_*$	$f^{**}P_V(1/3)\mathcal{B}^*(\overline{D}^0\pi^-_*)\epsilon(\pi_*)$	—	$f^{**}P_V(1/3)\,\mathcal{B}^*(\overline{D}^0\pi^-_*)(1-\epsilon(\pi_*))$
$D^{*-} \to D^-\pi^0_*$	—	$f^{**}P_V(1/3)\,\mathcal{B}^*(D^-\pi^0_*)$	—
$D^{*-} \to D^-\gamma$	—	$f^{**}P_V(1/3)\,\mathcal{B}^*(D^-\gamma)$	—
$D^{**-} \to \overline{D}^0\pi^-_{**}$	—	—	$f^{**}(1-P_V)(2/3)$
$D^{**-} \to D^-\pi^0_{**}$	—	$f^{**}(1-P_V)(1/3)$	—
$\to D^{*-}\ell^+\nu$			
$D^{*-} \to \overline{D}^0\pi^-_*$	$f^*\mathcal{B}^*(\overline{D}^0\pi^-_*)\epsilon(\pi_*)$	—	$f^*\mathcal{B}^*(\overline{D}^0\pi^-_*)(1-\epsilon(\pi_*))$
$D^{*-} \to D^-\pi^0_*$	—	$f^*\mathcal{B}^*(D^-\pi^0_*)$	—
$D^{*-} \to D^-\gamma$	—	$f^*\mathcal{B}^*(D^-\gamma)$	—
$\to D^-\ell^+\nu$	—	f	—
B^+ Decay Chains			
$\to \overline{D}^{**0}\ell^+\nu$			
$\overline{D}^{**0} \to D^{*-}\pi^+_{**}$			
$D^{*-} \to \overline{D}^0\pi^-_*$	$f^{**}P_V(2/3)\mathcal{B}^*(\overline{D}^0\pi^-_*)\epsilon(\pi_*)$	—	$f^{**}P_V(2/3)\mathcal{B}^*(\overline{D}^0\pi^-_*)(1-\epsilon(\pi_*))$
$D^{*-} \to D^-\pi^0_*$	—	$f^{**}P_V(2/3)\mathcal{B}^*(D^-\pi^0_*)$	—
$D^{*-} \to \overline{D}^-\gamma$	—	$f^{**}P_V(2/3)\mathcal{B}^*(D^-\gamma)$	—
$\overline{D}^{**0} \to \overline{D}^{*0}\pi^0_{**}$			
$\overline{D}^{*0} \to \overline{D}^0\pi^0_*$	—	—	$f^{**}P_V(1/3)\mathcal{B}^*(\overline{D}^0\pi^0_*)$
$\overline{D}^{*0} \to \overline{D}^0\gamma$	—	—	$f^{**}P_V(1/3)\mathcal{B}^*(\overline{D}^0\gamma)$
$\overline{D}^{**0} \to D^-\pi^+_{**}$	—	$f^{**}(1-P_V)(2/3)$	—
$\overline{D}^{**0} \to \overline{D}^0\pi^0_{**}$	—	—	$f^{**}(1-P_V)(1/3)$
$\to \overline{D}^{*0}\ell^+\nu$			
$\overline{D}^{*0} \to \overline{D}^0\pi^0_*$	—	—	$f^*\mathcal{B}^*(\overline{D}^0\pi^0_*)$
$\overline{D}^{*0} \to \overline{D}^0\gamma$	—	—	$f^*\mathcal{B}^*(\overline{D}^0\gamma)$
$\to \overline{D}^0\ell^+\nu$	—	—	f

Table 5.1: The contributions of the various B decay chains into reconstructed lD modes. Dashes indicate that a particular decay chain makes no contribution. The total contribution to a given sample is simply the sum over the entries in a vertical column. The branching ratio of $D^* \to XY$ is denoted by "$\mathcal{B}^*(XY)$." All $\pi^0_{(*)}$'s, π_{**}'s, and γ's are lost from the reconstruction, and π^-_*'s are identified with efficiency $\epsilon(\pi_*)$. See the text for a discussion of the other parameters.

the only missing particle is a neutrino. When we calculate the sample composition we add up the contributions F_i^0 from all $\bar{B}^0$ decays paths and the contributions F_i^+ from all B^- decay path, and similarly for $\bar{B}_s^0$, to calculate the relative fractions of the meson species in the lD sample, so that

$$f(\bar{B}^0) = \frac{\sum_i F_i^0}{\sum_i F_i^0 + \sum_j F_j^+ + \sum_k F_k^s} \tag{5.1}$$

would be fraction of $\bar{B}^0$ and so on.

As discussed, contribution of every decay path includes the factors associated with B species production, B species semileptonic branching rations, and decay path efficiencies in addition to the factors identified in Table 5.1. From Equation 5.1 it follows that only the relative values of the above quantities will affect the final result, which simplifies some of the calculations and in some cases leads to a decrease in uncertainties (by cancelling out common poorly known factors).

For the decay paths contributing to the $\bar{B}^0$ and B^- modes separate Monte-Carlo samples were generated using the `BGen` [5] b quark/B meson generator and `EvtGen` [35] B decay generator. The trigger and reconstruction efficiencies were then measured relative to the efficiencies of the direct decay paths, as just discussed. The trigger efficiencies are listed in Table 5.2. The efficiencies of the lD candidate selection for all decay paths are listed in Table 5.3. By definition, these efficiencies normalize to unity for the direct decay paths in every reconstruction mode.

The efficiencies for trigger and reconstruction quoted above were measured in the in μSVT Monte-Carlo samples. They were found to agree within 0.5% with the

efficiencies in eSVT samples and therefore the results of the sample composition studies apply to both samples. In our analysis, we correct the results for the sample composition in the combined eSVT+μSVT sample only.

5.4 Values of the branching ratios used in sample composition analysis

As mentioned above, for disentangling the composition of the semileptonic samples we need the values of poorly known relative branching ratios R_f and f^{**}:

$$R_f \equiv \frac{f^*}{f} = \frac{B(B \to \nu l^+ D^*)}{B(B \to \nu l^+ D)}$$

$$f^{**} \equiv \frac{B(B \to \nu l^+ D^{**})}{B(B \to \nu l^+ DX)} \equiv 1 - f - f^*.$$

In this section we explain how the values of these ratios were compiled from the available experimental results. The techniques used here are based on the 1991 CLEO analysis [18] which reports the measurements of the values of R_f and f^{**}. Some of the CLEO measurements have not been repeated since then, however the original paper derives the values of R_f and f^{**} using as input the values of the branching ratios of the decays $D^0 \to K^-\pi^+$ and $D^+ \to K^-\pi^+\pi^+$. Since that time, the precision of the measurements of these branching fractions has improved. In the recent B_d mixing analysis at CDF [19], the CLEO calculations were repeated to re-derive the values of R_f and f^{**} from the raw measurements performed in 1991 using the current PDG[24] values of the branching ratios where possible. This

Signature	Code	Decay channel					Relative Efficiency
ℓD^-	100.002	$B^0 \to$			$D^- \to$	$K\pi\pi$	1.00 ± 0.02
	100.202	$B^0 \to$		$D^{*-} \to$	$\gamma D^- \to$	$K\pi\pi$	1.20 ± 0.02
	100.212	$B^0 \to$		$D^{*-} \to$	$\pi^0_* D^- \to$	$K\pi\pi$	1.19 ± 0.02
	121.002	$B^0 \to$	$D^{**-} \to$	π^0_{**}	$D^- \to$	$K\pi\pi$	0.85 ± 0.01
	121.202	$B^0 \to$	$D^{**-} \to$	$\pi^0_{**} D^{*-} \to$	$\gamma D^- \to$	$K\pi\pi$	0.78 ± 0.01
	121.212	$B^0 \to$	$D^{**-} \to$	$\pi^0_{**} D^{*-} \to$	$\pi^0_* D^- \to$	$K\pi\pi$	0.78 ± 0.01
	212.002	$B^+ \to$	$\overline{D}^{**0} \to$	π^+_{**}	$D^- \to$	$K\pi\pi$	1.06 ± 0.02
	212.202	$B^+ \to$	$\overline{D}^{**0} \to$	$\pi^+_{**} D^{*-} \to$	$\gamma D^- \to$	$K\pi\pi$	0.91 ± 0.01
	212.212	$B^+ \to$	$\overline{D}^{**0} \to$	$\pi^+_{**} D^{*-} \to$	$\pi^0_* D^- \to$	$K\pi\pi$	0.91 ± 0.01
ℓD^0	100.221	$B^0 \to$		$D^{*-} \to$	$\pi^-_* \overline{D}^0 \to$	$K\pi$	1.22 ± 0.02
	121.221	$B^0 \to$	$D^{**-} \to$	$\pi^0_{**} D^{*-} \to$	$\pi^-_* \overline{D}^0 \to$	$K\pi$	0.85 ± 0.01
	122.001	$B^0 \to$	$D^{**-} \to$	π^-_{**}	$\overline{D}^0 \to$	$K\pi$	0.95 ± 0.01
	122.101	$B^0 \to$	$D^{**-} \to$	$\pi^-_{**} D^{*0} \to$	$\gamma \overline{D}^0 \to$	$K\pi$	0.83 ± 0.01
	122.111	$B^0 \to$	$D^{**-} \to$	$\pi^-_{**} D^{*0} \to$	$\pi^0_* \overline{D}^0 \to$	$K\pi$	0.84 ± 0.01
	200.001	$B^+ \to$			$\overline{D}^0 \to$	$K\pi$	1.00 ± 0.01
	200.101	$B^+ \to$		$D^{*0} \to$	$\gamma \overline{D}^0 \to$	$K\pi$	1.32 ± 0.02
	200.111	$B^+ \to$		$D^{*0} \to$	$\pi^0_* \overline{D}^0 \to$	$K\pi$	1.32 ± 0.02
	211.001	$B^+ \to$	$\overline{D}^{**0} \to$	π^0_{**}	$\overline{D}^0 \to$	$K\pi$	0.83 ± 0.01
	211.101	$B^+ \to$	$\overline{D}^{**0} \to$	$\pi^0_{**} \overline{D}^{*0} \to$	$\gamma \overline{D}^0 \to$	$K\pi$	0.94 ± 0.01
	211.111	$B^+ \to$	$\overline{D}^{**0} \to$	$\pi^0_{**} \overline{D}^{*0} \to$	$\pi^0_* \overline{D}^0 \to$	$K\pi$	0.84 ± 0.01
	212.221	$B^+ \to$	$\overline{D}^{**0} \to$	$\pi^+_{**} D^{*-} \to$	$\pi^-_* \overline{D}^0 \to$	$K\pi$	0.83 ± 0.01
ℓD^{*-}	100.221-s	$B^0 \to$		$D^{*-} \to$	$\pi^-_* \overline{D}^0 \to$	$K\pi$	1.00 ± 0.02
	121.221-s	$B^0 \to$	$D^{**-} \to$	$\pi^0_{**} D^{*-} \to$	$\pi^-_* \overline{D}^0 \to$	$K\pi$	0.68 ± 0.01
	212.221-s	$B^+ \to$	$\overline{D}^{**0} \to$	$\pi^+_{**} D^{*-} \to$	$\pi^-_* \overline{D}^0 \to$	$K\pi$	0.66 ± 0.01

Table 5.2: Trigger efficiencies of the various B decay paths measured in the 21 B signal samples and normalized to the efficiency measured for the direct decay path. The Monte Carlo statistical errors are also shown. The decay channel "code" is a numbering scheme used for internal identification.

Signature	Code	Decay channel					Relative Efficiency
ℓD^-	100.002	$B^0 \to$			$D^- \to$	$K\pi\pi$	1.00 ± 0.01
	100.202	$B^0 \to$		$D^{*-} \to$	$\gamma D^- \to$	$K\pi\pi$	0.96 ± 0.01
	100.212	$B^0 \to$		$D^{*-} \to$	$\pi_*^0 D^- \to$	$K\pi\pi$	0.96 ± 0.01
	121.002	$B^0 \to$	$D^{**-} \to$	π_{**}^0	$D^- \to$	$K\pi\pi$	0.95 ± 0.01
	121.202	$B^0 \to$	$D^{**-} \to$	$\pi_{**}^0 D^{*-} \to$	$\gamma D^- \to$	$K\pi\pi$	0.97 ± 0.01
	121.212	$B^0 \to$	$D^{**-} \to$	$\pi_{**}^0 D^{*-} \to$	$\pi_*^0 D^- \to$	$K\pi\pi$	0.95 ± 0.01
	212.002	$B^+ \to$	$\overline{D}^{**0} \to$	π_{**}^+	$D^- \to$	$K\pi\pi$	0.77 ± 0.01
	212.202	$B^+ \to$	$\overline{D}^{**0} \to$	$\pi_{**}^+ D^{*-} \to$	$\gamma D^- \to$	$K\pi\pi$	0.82 ± 0.01
	212.212	$B^+ \to$	$\overline{D}^{**0} \to$	$\pi_{**}^+ D^{*-} \to$	$\pi_*^0 D^- \to$	$K\pi\pi$	0.81 ± 0.01
ℓD^0	100.221	$B^0 \to$		$D^{*-} \to$	$\pi_*^- \overline{D}^0 \to$	$K\pi$	0.90 ± 0.01
	121.221	$B^0 \to$	$D^{**-} \to$	$\pi_{**}^0 D^{*-} \to$	$\pi_*^- \overline{D}^0 \to$	$K\pi$	0.92 ± 0.01
	122.001	$B^0 \to$	$D^{**-} \to$	π_{**}^-	$\overline{D}^0 \to$	$K\pi$	0.91 ± 0.01
	122.101	$B^0 \to$	$D^{**-} \to$	$\pi_{**}^- D^{*0} \to$	$\gamma \overline{D}^0 \to$	$K\pi$	0.92 ± 0.01
	122.111	$B^0 \to$	$D^{**-} \to$	$\pi_{**}^- D^{*0} \to$	$\pi_*^0 \overline{D}^0 \to$	$K\pi$	0.91 ± 0.01
	200.001	$B^+ \to$			$\overline{D}^0 \to$	$K\pi$	1.00 ± 0.01
	200.101	$B^+ \to$		$D^{*0} \to$	$\gamma \overline{D}^0 \to$	$K\pi$	0.93 ± 0.01
	200.111	$B^+ \to$		$D^{*0} \to$	$\pi_*^0 \overline{D}^0 \to$	$K\pi$	0.93 ± 0.01
	211.001	$B^+ \to$	$\overline{D}^{**0} \to$	π_{**}^0	$\overline{D}^0 \to$	$K\pi$	0.94 ± 0.01
	211.101	$B^+ \to$	$\overline{D}^{**0} \to$	$\pi_{**}^0 \overline{D}^{*0} \to$	$\gamma \overline{D}^0 \to$	$K\pi$	0.95 ± 0.01
	211.111	$B^+ \to$	$\overline{D}^{**0} \to$	$\pi_{**}^0 \overline{D}^{*0} \to$	$\pi_*^0 \overline{D}^0 \to$	$K\pi$	0.93 ± 0.01
	212.221	$B^+ \to$	$\overline{D}^{**0} \to$	$\pi_{**}^+ D^{*-} \to$	$\pi_*^- \overline{D}^0 \to$	$K\pi$	0.91 ± 0.01
ℓD^{*-}	100.221-s	$B^0 \to$		$D^{*-} \to$	$\pi_*^- \overline{D}^0 \to$	$K\pi$	1.00 ± 0.02
	121.221-s	$B^0 \to$	$D^{**-} \to$	$\pi_{**}^0 D^{*-} \to$	$\pi_*^- \overline{D}^0 \to$	$K\pi$	1.05 ± 0.02
	212.221-s	$B^+ \to$	$\overline{D}^{**0} \to$	$\pi_{**}^+ D^{*-} \to$	$\pi_*^- \overline{D}^0 \to$	$K\pi$	0.99 ± 0.02

Table 5.3: Reconstruction efficiencies of the various B decay paths after the lD candidate selection. The efficiencies are normalized with respect to the direct decay path efficiency. The Monte-Carlo statistical errors are also shown. The decay channel "code" is a numbering scheme used for internal identification.

recalculation was described in great detail in a CDF internal publication [9] but since this is not easily available outside the collaboration, we reproduce the basic assumptions and the results of that study here.

The main quantity that is still available only from [18] is the measurement of

$$R_{Dl} \equiv \frac{B(B \to \nu l^+ DX)}{B(B \to \nu l^+ X)}.$$

The original CLEO result is $R^{91}_{D^0 l} = 0.67 \pm 0.09 \pm 0.10$ and $R^{91}_{D^+ l} = 0.26 \pm 0.07 \pm 0.04$ yielding a combined quantity

$$R^{91}_{Dl} \equiv R^{91}_{D^0 l} + R^{91}_{D^+ l} = 0.93 \pm 0.11 \pm 0.11,$$

where the first quoted uncertainty is statistical and the second is the systematic uncertainty due to imprecisely known D branching ratios. The original paper [18] quotes the raw measurements and the technique of deriving R_{Dl} which allowed CDF [19] to recalculate these values with the current world best measurements of D branching ratios shown in Table 5.4. This recalculation produced a value

	B_{91}, used in [18] %	B_{02}, PDG 2002 %
$D^0 \to K^- \pi^+$	$4.2 \pm 0.4 \pm 0.4$	3.80 ± 0.09
$D^+ \to K^- \pi^+ \pi^+$	$9.0 \pm 1.3 \pm 0.4$	9.0 ± 0.6

Table 5.4: The values of the D mesons branching ratios used by the CLEO 1991 analysis and the current world best measurements used to re-derive the quantities R_f and f^{**}. The measurements of these ratios that were used in our analysis were compiled in PDG 2002 Review of Particle Physics.

$R^{02}_{Dl} = 1.0056 \pm 0.1275 \pm 0.0483$ which was saturated (and of course unphysical) so

it was transformed into the final form

$$R_{Dl}^{02} = 1.00^{+0.00+0.00}_{-0.12-0.04}.$$

Additional branching ratios needed for the calculation of f^{**} are listed in Table 5.5. The third row of that table represents the sum of the D, D^* ratios combined into a species averaged value, which is the single composite quantity we need for f^{**} computation, as will be shown momentarily. Combining all these input parameters

	B^0, %	B^-, %
$B(B \to \nu l^+ D)$	$2.11 \pm 0.10 \pm 0.14$	$2.15 \pm 0.21 \pm 0.05$
$B(B \to \nu l^+ D^*)$	$4.60 \pm 0.18 \pm 0.11$	$5.45 \pm 0.78 \pm 0.13$
$< B(B \to \nu l^+ D^{(*)}) >$	$7.16 \pm 0.42 \pm 0.21$	

Table 5.5: The values of the B mesons branching ratios used to derive the quantities R_f and f^{**}. The third row represents the sum of the D, D^* ratios combined into a species average, which is the composite quantity we need for the f^{**} computation. The current best measurements of these ratios were compiled in PDG 2002 Review of Particle Physics. The errors quoted by PDG are here split into statistical and systematic by separating the (assumed) contribution due to the uncertainties on the D meson branching ratios. This systematic is assumed 100% correlated between the modes.

with the average total leptonic branching ratio $< B(B \to \nu l^+ X) >= 10.38 \pm 0.32$ % taken from 2002 PDG[28] we can calculate the value of f^{**} from the relation

$$\begin{aligned} 1 - f^{**} &= f + f^* = \frac{B(B \to \nu l^+ D)}{B(B \to \nu l^+ DX)} + \frac{B(B \to \nu l^+ D^*)}{B(B \to \nu l^+ DX)} \\ &= \frac{< B(B \to \nu l^+ D^{(*)}) >}{B(B \to \nu l^+ DX)} = \frac{< B(B \to \nu l^+ D^{(*)}) >}{< B(B \to \nu l^+ X) > \cdot R_{Dl}} \\ &= 0.6895^{+0.0455+0.0015}_{-0.0456-0.0016}. \end{aligned}$$

Based on this calculation, in our sample composition analysis we use a value

$$f^{**} = 0.31 \pm 0.05.$$

For the calculation of R_f we use the values from Table 5.5 to form the error-weighted species averaged branching ratios of the D, D^* decays and derive

$$\begin{aligned} R_f &\equiv \frac{f^*}{f} = \frac{<B(B\to\nu l^+ D^*)>}{<B(B\to\nu l^+ D)>} \\ &= \frac{4.63\pm0.18\pm0.11}{2.12\pm0.10\pm0.11} \\ &= 2.181 \pm 0.1299 \pm 0.0599. \end{aligned}$$

In our analysis we used a value

$$R_f = 2.18 \pm 0.14.$$

5.5 Results of the sample composition measurements

The results of the sample composition calculations are most conveniently expressed in a matrix form where the number of events in each of the lD samples is presented as a linear superposition of the fractional contributions of the three B mesons species:

$$\begin{pmatrix} lD^- \\ lD^{*-} \\ lD^0 \\ lD_s \end{pmatrix} = \begin{pmatrix} 0.854 & 0.139 & 0.007 \\ 0.899 & 0.098 & 0.003 \\ 0.261 & 0.716 & 0.023 \\ 0.063 & 0.063 & 0.874 \end{pmatrix} \times \begin{pmatrix} B^0 \\ B^+ \\ B_s^0 \end{pmatrix}. \quad (5.2)$$

By construction, the sum of the elements in every row of the sample composition matrix is 1. We do not quote the errors on the sample composition parameters because they are not used directly in our calculations. The error estimation procedure is explained in detail in the next Section.

5.6 Using the sample composition parameters to correct the measurements of particle content

Once the sample composition is known, it is possible convert the yields of particles found around B decays into the particle yields around lD decays (which can be measured in data) by adding these B yields weighted by the sample composition parameters. In other words, the measured fractions of B mesons in the sample are exactly the parameters we need to convert the yields around B signals to yields around lD signals.

Therefore we can now rewrite Equation 5.2 as

$$\begin{pmatrix} m_d \\ m_d^* \\ m_u \\ m_s \end{pmatrix} = \begin{pmatrix} 0.854 & 0.139 & 0.007 \\ 0.899 & 0.098 & 0.003 \\ 0.261 & 0.716 & 0.023 \\ 0.063 & 0.063 & 0.874 \end{pmatrix} \times \begin{pmatrix} M_d \\ M_u \\ M_s \end{pmatrix}. \tag{5.3}$$

We used m and M to denote the yields of the same particle type around lD and B signals correspondingly. The subscript refers to the quark coupled with the b quark in the B meson corresponding to the direct decay path, and the $*$ superscript denotes the lD^{*-} mode.

However this relation is "backwards", i.e. we measure $m_i^{(*)}$ in the data, but need to know the values of $M_i^{(*)}$. One possible solution is to use the above equation to estimate the most likely values of M_i by finding the values of M_i which minimize

the sum

$$\chi^2 = \sum_i \left(\frac{m_{fit} - m_{measured}}{\sigma(m)} \right)^2 , \tag{5.4}$$

where the m_{fit} are calculated from the above matrix for the estimated values of M_i. This however would only be a correct procedure if the sample composition parameters were known precisely. Since some of the branching ratios used in the calculation of the sample composition parameters are known very poorly, the uncertainties on the sample composition parameters are significant. The proper procedure then is somewhat more involved.

We use the following notation in the rest of this section to describe the cross-talk parameters and the values of the yields:

$$\mathbf{C} = \begin{pmatrix} 1 - c_1 - s_1 & c_1 & s_1 \\ 1 - c_1^* - s_2 & c_1^* & s_2 \\ c_2 & 1 - c_2 - s_3 & s_3 \\ s & s & 1 - 2s \end{pmatrix} \tag{5.5}$$

$$\mathbf{x} = \begin{pmatrix} m_d \\ m_d^* \\ m_u \\ m_s \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} M_d \\ M_u \\ M_s \end{pmatrix} \tag{5.6}$$

Note that the equation 5.4 can be rewritten as

$$\chi^2(\mathbf{X}) = (\mathbf{x} - \hat{\mathbf{x}}(\mathbf{X}))^T \mathbf{W}^{-1} (\mathbf{x} - \hat{\mathbf{x}}(\mathbf{X})), \tag{5.7}$$

where we used the notation $\hat{\mathbf{x}}(\mathbf{X}) = \mathbf{C} \cdot \mathbf{X}$ and $\mathbf{W}$ is the covariance matrix of the measured yields:

$$\mathbf{W} = \begin{pmatrix} \sigma^2_{m_d} & 0 & 0 & 0 \\ 0 & \sigma^2_{m^*_d} & 0 & 0 \\ 0 & 0 & \sigma^2_{m_u} & 0 \\ 0 & 0 & 0 & \sigma^2_{m_s} \end{pmatrix} \tag{5.8}$$

In presence of the uncertain sample composition parameters, the covariance matrix will have to include the effect of these additional uncertainties. We assume that the matrix $\mathbf{C}$ is a function of a set of parameters $\mathbf{y}$ which have a covariance matrix $\mathbf{V}$. Then, given a set of fit parameters $\mathbf{X}$, the covariance matrix for the expected parameters $\hat{\mathbf{x}}$ is given by

$$\mathrm{Cov}(\hat{\mathbf{x}}) = \left(\frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{y}}\right)^T \mathbf{V} \left(\frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{y}}\right), \tag{5.9}$$

where

$$\frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{y}} = \frac{\partial(\mathbf{C} \cdot \mathbf{X})}{\partial \mathbf{y}} \equiv \mathbf{T} \tag{5.10}$$

Then the proper expression to minimize is

$$\chi^2(\mathbf{X}) = (\mathbf{x} - \hat{\mathbf{x}}(\mathbf{X}))^T (\mathbf{W} + \mathbf{T}^\mathbf{T}\mathbf{V}\mathbf{T})^{-1} (\mathbf{x} - \hat{\mathbf{x}}(\mathbf{X})), \tag{5.11}$$

The challenge therefore lies in estimating the matrices $\mathbf{T}$ and $\mathbf{V}$. Since there is very little contribution of the $\bar{B}^0_s$ decays in modes other than lD_s we choose to ignore the uncertainties on the composition parameters s_i, and choose four parameters in

conversion matrix $\mathbf{C}$ that have non-negligible uncertainties:

$$\mathbf{c} = \begin{pmatrix} c_1 \\ c_1^* \\ c_2 \\ s \end{pmatrix} \tag{5.12}$$

Similarly, most of the components in the calculation of the sample composition described above are known well or very well, and their uncertainties can be ignored. For example, all the trigger and reconstruction efficiencies are measured to a precision of about 1%. There are four parameters that have the largest uncertainties, with all other parameters being known with precisions an order of magnitude better. This is the summary of these parameters and the values we used in the sample composition calculations:

$$\mathbf{y} = \begin{pmatrix} f^{**} \\ R_f \\ P_V \\ s \end{pmatrix} = \begin{pmatrix} 0.31 & \pm & 0.05 \\ 2.14 & \pm & 0.14 \\ 0.627 & \pm & 0.13 \\ 0.063 & \pm & 0.025 \end{pmatrix} \tag{5.13}$$

Here we introduced a variable $R_f \equiv f^*/f$. We also assume that these parameters are uncorrelated, which is a reasonable assumption given that two of them were estimated from non-related branching ratios, and other two were measured independently by several experiments. We therefore form $\mathbf{V}$ as a diagonal covariance matrix based on the uncertainties shown in equation 5.13.

The matrix of $\frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{y}}$ was calculated as follows. We notice that because of the way matrix $\mathbf{C}$ is written we can separate the differentiation of the parameters in matrix $\mathbf{C}$ from the fitted values $\mathbf{X}$ so that we have

$$\mathbf{T}_{ij} \equiv \left(\frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{y}}\right)_{ij} = \sum_k \frac{\partial \mathbf{c}_{ik}}{\partial \mathbf{y}_j} \mathbf{X}_k \equiv \frac{\partial \mathbf{c}_i}{\partial \mathbf{y}_j} \cdot \hat{\mathbf{X}}_i, \tag{5.14}$$

where

$$\hat{\mathbf{X}} = \begin{pmatrix} -1 & 1 & 0 \\ -1 & 1 & 0 \\ 1 & -1 & 0 \\ 1 & 1 & -2 \end{pmatrix} \cdot \mathbf{X} \tag{5.15}$$

The functions for analytic calculations of the parameters $\mathbf{c}_i$ from the values of $\mathbf{y_i}$ were then numerically differentiated with the respect to each of the parameters $\mathbf{y_i}$ to obtain the matrix $\frac{\partial \mathbf{c}}{\partial \mathbf{y}}$:

$$\frac{\partial \mathbf{c}}{\partial \mathbf{y}} = \begin{pmatrix} 0.6358 & 0.02097 & -0.159 & 0 \\ 0.4114 & -0.04423 & 0.1407 & 0 \\ 0.4166 & 0.02532 & -0.008686 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{5.16}$$

The covariance matrix of the sample composition parameters is also interesting and can be calculated as

$$\mathrm{Cov}(\mathbf{c}) = \left(\frac{\partial \mathbf{c}}{\partial \mathbf{y}}\right)^T \mathbf{V} \left(\frac{\partial \mathbf{c}}{\partial \mathbf{y}}\right) \tag{5.17}$$

$$\mathrm{Cov}(\mathbf{c}) = \begin{pmatrix} 0.007729 & -0.0001166 & 0.0008114 & 0 \\ -0.0001166 & 5.201e-05 & -0.0001347 & 0 \\ 0.0008114 & -0.0001347 & 0.0004529 & 0 \\ 0 & 0 & 0 & 0.000625 \end{pmatrix}. \quad (5.18)$$

At this point we have everything in hand to perform the minimization of the expression in equation 5.11 given the measured values of the yields and their statistical uncertainties. This was done using the `MINUIT` [32] minimization package for every p_T range as well as for the integrated measurements.

Using this minimization procedure with and without the additional covariance matrix introduced by the sample composition uncertainties allows one to separate the statistical and systematic uncertainties on the measurements. After obtaining the result and the associated `MINUIT` error σ_t using the outlined procedure one can repeat the procedure by using just the statistical uncertainties (*i.e.* set $\mathbf{V} = 0$) and assign the error reported by `MINUIT` to the statistical uncertainty σ_{stat}, with systematic uncertainty being $\sigma_{syst} = \sqrt{\sigma_t^2 - \sigma_{stat}^2}$.

5.7 Backgrounds in the sample of the tracks reconstructed in a cone $\Delta R < 0.7$ around lD signals

Measuring the track yields around B signals has one additional complication due to the B sample composition. Namely, various decay chains leading to the same partially reconstructed lD signal have a number of unreconstructed decay products. These particles constitute a background in our track sample, which means that the issues of the composition of B sample and track sample are related.

A decay process that makes significant contribution to the sample of tracks we select is the decay products of the excited B states (B^{**} mesons). Since B^{**} states undergo very fast strong decays $B^{**}_x \to B_x X$ into lower B states their decay products will be found as tracks coming from the same primary vertex as the reconstructed B meson. They are therefore indistinguishable from the particles produced during b quark fragmentation. Also, there is very little knowledge about B^{**} production and decays, and correction of our results by Monte-Carlo prediction of these backgrounds would be unrealistic. Therefore we consider these tracks to be a part of our sample of particles associated with B meson production.

The main source of physics background is therefore the tracks from the intermediate $D^{*(*)}$ states. To reduce the effect of these secondary particles on our results we use a realistic Monte-Carlo to predict the amount of this background. We found

that a vast majority of the tracks coming from the excited D states were pions. After the measured particle yields in the data are corrected for the detector reconstruction efficiencies, we subtract the expected $D^{*,**}$ pion contribution from the resulting pion yields. Table 5.6 lists the yields used for the correction. We found that the size of this correction is around 6% in $\bar{B}^0, B^-$ modes.

p_T	$\mu^+ D^-$	$\mu^+ D^{*-}$	$\mu^+ \bar{D}_0$	$\mu^+ D_s^-$
0.4-0.7	0.011	0.018	0.033	0.0056
0.7-1	0.0067	0.0088	0.015	0
1-1.5	0.011	0.0091	0.012	0.00059
1.5-2.5	0.014	0.0085	0.011	0.003
2.5-5	0.0092	0.0052	0.0073	0.003

Table 5.6: Yields of pions from excited D states that were selected as fragmentation tracks.

In the lD_s mode, the background decays $B_x \to D_x^{(*)} D_s$ (where x subscript can mean either u or d quarks) can cause some tracks from these decays to be included in the track sample selection. We used a realistic Monte-Carlo sample based on the `BGen` [5] generator to estimate this contribution. We found that the total yield of these tracks is 0.005 ± 0.003, and they consist of roughly 60% kaons and 40% pions. The theoretical systematic uncertainty on the size of this correction is quite large, about 30%. In addition, due to the small amount of this background measuring it in several p_T regions would require large amounts of Monte-Carlo. Since this background is also small compared to the errors on other corrections we do not take it into account.

Chapter 6

Technique for measuring the particle content

After the sample of tracks has been selected we measure the particle content. We do this by estimating the fractions of different particle species in the sample, and also by measuring the yields of each of the particles species. The majority of the particles that are reconstructed as the tracks in the detector will belong to one of the five species: π, K, p, e, μ. Our particle identification capabilities allow us to distinguish between all these types with the exception of muons being practically identical to pions in both the TOF and COT detectors.

In this Chapter we describe the technical implementation of the procedure for extraction the particle species content of the sample of tracks associated with the formation of various B mesons.

6.1 Description of the fitting algorithm

This section describes the technique used for measuring the particle content of our data sample. We measure the numbers of each particle species in our sample of tracks by fitting the distributions of the track parameters that are sensitive to particle mass with the distributions of these parameters expected for each particle species.

These track parameters are the particle arrival time as measured in the TOF detector and particle energy loss in the COT detector. In order to correct the measurements for the combinatorial background in the reconstructed B sample we use the invariant mass distribution of the D meson M_D. The three measurements (time of flight, dE/dx, M_D) are combined in an unbinned extended likelihood fit as described in the following text.

If we were to know the particle type beforehand then we could compare the predicted production time of the particle t_{pred} (TOF measured arrival time minus its time of flight predicted from mass, momentum and track trajectory) with the production time measured by the detector (event t_0)[1]. The distribution of $t_{pred} - t_0$ represents the resolution of the TOF detector. An incorrect particle species hypothesis has an effect of shifting the resolution function in the time axis by

[1]Event t_0 is estimated from the TOF detector measurements by taking all the tracks in the event with TOF information and calculating the most likely average production time assuming some average mixture of particle species in the event.

amount depending on particle momentum and mass. Therefore if you select tracks with similar momenta then the distribution of $t_{pred}(\pi) - t0$ would consist of several peaks each representing a contribution from one of the particle species. Each peak is described by the TOF resolution function centered at a value of $t_{pred}(\pi) - t_{pred}(i)$ where i refers to the particle type in question. Note that $t_{pred}(\pi) - t_{pred}(i)$ depends on track momentum for $i \neq \pi$, which means that as the range of momenta of the selected particles becomes wider the peaks will start to wash out and eventually merge into a single smooth distribution.

When a particular track is found in an event then the probability that it has a TOF measured arrival time t_{meas} and a COT measured energy loss dE/dx can be written as

$$\begin{aligned} P(M_D, t_{meas}, dE/dx) = \sum_{c=S,B} P_c(M_D) \times \\ \times \sum_{i=\pi,K,p,e} f_c^i \cdot R_{TOF}^i(t_{meas} - t_{exp}^i) \cdot R_{dE/dx}^i(Z^i). \end{aligned} \tag{6.1}$$

Here f_c^i is the probability that the track from either signal or background [2] component was produced by a particle species "i", P_c are the probabilities that given event is a signal or background event, and $R_{TOF}, R_{dE/dx}$ are the resolution functions for TOF and dE/dx measurements which may depend on track and event parameters as well as particle type. Note that we require all f^i to add up to 1 in both signal and background region, and also $P_S + P_B = 1$.

The resolution functions are measured in the calibration samples as described in

[2] c stands for component, either signal or background.

Appendix A. In order to determine P_S and P_B we fit the invariant mass distribution of the D mesons from the partially reconstructed lD decays. The fitted histogram has one entry for every track used in a current fit (see Figure 6.1, note that the fits are done in bins of fragmentation track p_T). signal component fit m_S is a double Gaussian distribution and background fit m_B is a first order polynomial. We then determine the probabilities as $P_S = \frac{m_S(M)}{m_S(M)+m_B(M)}$ and $P_B = 1 - P_S$ for an event where D meson invariant mass is M.

Given the knowledge of the resolution functions and P_S and P_B we can write the extended log-likelihood expression for the sample of particles:

$$log(L) = Nlog(n) - n + \sum logP(M_D, t_{meas}, dE/dx), \tag{6.2}$$

where

$$n = \sum_{c=S,B} \sum_{i=\pi,K,p,e} n_c^i \tag{6.3}$$

denotes the sum of the fitted numbers of all particle species, and N is the total (known) size of the track sample. The sum is taken over all the tracks in the sample. A simultaneous fit to the TOF, dE/dx and M_D distributions is performed. We minimize $-log(L)$ using the `MINUIT` package to calculate the fractions of particle species in a sample of tracks.

The particle fractions are determined as

$$f_c^i = \frac{n_c^i}{\sum_{i=\pi,K,p,e} n_c^i} \tag{6.4}$$

We use an extended likelihood function instead of normal likelihood to ensure that

the measured uncertainties on the numbers of particles of each type are calculated in accordance to the Poisson statistics. The extended log-likelihood [22] is derived from the standard one by adding the first two terms in equation 6.2. This means that the number of tracks in the fit n becomes an additional free parameter constrained by the observed number of tracks N to which it eventually converges. Using extended likelihood fit in our case does not change any results but changes the uncertainties estimated by the `MINUIT` by correctly taking into account the Poisson statistical nature of the track sample.

In this description of the fitting procedure the parameters one could fit for are the particle fractions f^i and the resolution function parameters. In practice when measuring the resolution function parameters this procedure was applied to a clean sample of pions [3] and the particle fractions were fixed so that $f_S^\pi = 1$, $f_{bg}^\pi = 0.998$ and $f_{bg}^K = 0.002$ [4]. The resolution function parameters were then measured by the fit.

Conversely, when fitting for the particle fractions we fixed the resolution function parameters to the values we previously measured and only the parameters related to the particle content were fit for.

[3] We used the soft pions from the lD^{*-} sample, same sample as used for measuring the fractions.

[4] These fractions were first determined by measuring the particle content of the sample with imperfect resolution functions and then used as fixed parameters for the final determination of the TOF resolution. It is enough to perform this iterative procedure once to achieve good precision.

6.2 Dependence of fraction on track p_T

In this description of the likelihood function, we assumed that the particle fractions f^i are constants. However we do know that they depend on various particle and event properties, most prominently particle transverse momentum p_T. For this reason the fits for the particle fractions are performed in several restricted ranges in p_T. If the fractions change significantly within a bin, then our measurement of of the fractions will have a bias.

We improve the fit procedure by parametrizing the particle fractions with a first degree polynomial function in particle p_T to better match the behavior of the data. We tried fitting for the gradients as additional parameters and using an iterative method. We found that the best results[5] were achieved by first measuring the fractions assuming no p_T dependence, and then using these measurements in different p_T ranges to estimate the dependence of the fractions on p_T and use these estimates to fix the gradients in the fit. This procedure is iterative in nature and can be performed several times for the best precision. The pion fraction is always constrained from the remaining particle fractions so that the sum of all fractions is 1.

When performing the fit this way, we recalculate the $f^i(p_T)$ for every track from the measured fit parameters $f^j = \frac{n^j}{\sum_i n^i}$ and the fraction gradients $\frac{df^i}{dp_T}$. For example,

[5] There is not always enough statistics within a p_T range to reliably fit for the particle fraction dependence on p_T.

the fraction at the low edge of the bin is given by expression

$$f^j_{low} = \frac{n^j}{\sum_i n^i} - \frac{df^j}{dp_T} \cdot \frac{\sum(p_T - p_{T,low})}{\sum_i n^i}, \tag{6.5}$$

where n^j is the number of tracks of a particle type j, and the sum in the denominator is performed over all particle types to give a total number of tracks in the sample, and the sum in the numerator is performed over all the tracks in the event[6]. The actual fit parameters in the fitting framework were $f^j = \frac{n^j}{\sum_i n^i}$ and the total number of particles n (which normally converges to the number of tracks in the sample N), as shown in equation 6.2.

The full likelihood fit procedure was tested with Monte-Carlo techniques with and without this additional parameterization. This test validates the fitting procedure, measures the size of the bias introduced by assuming the fractions are constant in the p_T range, and checks whether the improved procedure reduces this bias substantially.

In this Monte-Carlo study we created a data sample from a sample of tracks around μD^- meson signals that satisfy the usual analysis criteria from Section 4.2. P_{sg} and P_{bg} were also calculated for every event as described above, and an event was randomly classified as a signal or a background event according to these probabilities. Then we simulated the measured TOF and dE/dx values by choosing

[6] One can derive this formula by writing down a number of tracks of certain type in the sample as $n^j = \sum(f_0 + f' \cdot \Delta p) = f_0 N + f' \sum \Delta p$ with N being the size of the track sample and the sum is performed over all tracks.

a particle species and setting them to the predicted values for that particle type, adding random smearing according to the measured resolution functions.

The particle type was chosen randomly depending on track p_T. We measured the particle fractions in data and fit them with smooth functions. These curves were then used as the input for simulating the particle fractions in the Monte-Carlo data sample so that we mimic the behavior of the true data sample as closely as possible.

We then performed the likelihood fits in the same p_T ranges as in the fits on the actual data, and we compared the results of the fit with the true simulated particle fractions. A bias is present in most of the measurements when we assume flat particle fraction distributions. The bias however is small compared to the statistical error (and for a smaller sample of tracks around $\bar{B}_s^0$ signals the statistical error grows while the bias remains constant).

Then we allowed the fit to parameterize the fractions with linear functions and performed the fits again, by fixing the gradients to their simulated values at the low edge of the p_T bin. The results of this study are shown in Figure 6.2 for the signal and background events. 3150 Monte-Carlo data sets were created and evaluated. We see no evidence of a bias in the likelihood fit within the statistical error and conclude that our procedure is safe in that respect.

When fitting the actual data, the fit was performed as an iterative procedure. First the fits were done assuming constant fractions within the p_T bins. Then the

fitted fractions were used to estimate the dependence of f^i on p_T and to fix the gradients in the fitting framework. Then this procedure was performed once again to get the final result.

6.3 Detector efficiencies

When presenting the measured yields of each particle species in a sample we attempt to give the numbers describing the actual particles produced in b fragmentation. Therefore we correct the measured track yields by detector efficiencies to arrive to the final particle yields.

To measure the detector efficiency we first measure the COT acceptance, *i.e.* the fraction of particles produced in a cone $\Delta R < 0.7$ around lD signals that are reconstructed as COT tracks with at least 20 position measurements. We use realistic Monte-Carlo sample for this purpose and find these efficiencies to be 0.978, 0.969, 0.976, 0.959 respectively for μ^+D^-, μ^+D^{*-}, $\mu^+\bar{D}_0$, $\mu^+D_s^-$ modes. The statistical uncertainty of this measurement is negligible compared to other sources of errors.

Then we measure the efficiency for these COT tracks to have at least 3 SVX $R-\phi$ hits, TOF and dE/dx information. This is done on the real μSVT data from the detector. The results are summarized in Table 6.1. The errors on the measured efficiencies are added to the statistical errors on the measured yields.

The procedure was performed in these two steps because SVX, TOF and dE/dx efficiencies are not reproduced by the Monte-Carlo simulation sufficiently well. On

Combined TOF/dEdX/SVX efficiencies.				
p_T	$\mu^+ D^-$	$\mu^+ D^{*-}$	$\mu^+ \bar{D}_0$	$\mu^+ D_s^-$
0.4-0.7	0.5263 ± 0.0031	0.5460 ± 0.0049	0.5320 ± 0.0024	0.506 ± 0.011
0.7-1	0.5388 ± 0.0040	0.546 ± 0.0064	0.5312 ± 0.0030	0.485 ± 0.014
1-1.5	0.5156 ± 0.0041	0.543 ± 0.0066	0.5408 ± 0.0032	0.497 ± 0.016
1.5-2.5	0.5259 ± 0.0046	0.533 ± 0.0077	0.5397 ± 0.0037	0.502 ± 0.017
2.5-5	0.529 ± 0.0062	0.525 ± 0.011	0.540 ± 0.0056	0.586 ± 0.023

Table 6.1: The efficiencies (as measured in the μSVT sample) for fragmentation tracks with at least 20 COT hits to have TOF/dE/dx information and at least 3 $R-\phi$ SVX hits. The uncertainties on these efficiencies are added to the statistical errors.

the other hand, these efficiencies are easily accessible from the data, and simulation of the COT efficiency and acceptance in the Monte-Carlo is much more reliable.

To account for a possible remaining systematic uncertainty due to imperfect COT simulation in the Monte-Carlo, we assign an additional systematic uncertainty to our correction. This uncertainty is 1% in all p_T region except the lowest one, where it is 5% because the differences might be more significant.[7] This should allow for a conservative estimate of this systematics that is still very small compared to other sources of uncertainties.

[7]There have been studies performed at CDF showing significant COT inefficiency in data compared to Monte-Carlo in the low p_T region (below 550 MeV/c). However our studies did not show any additional inefficiencies and we chose not to assign additional corrections. Instead we assign a larger systematic uncertainty in the lowest p_T region, since none of the quoted studies can be considered complete and conclusive.

6.4 Kaon decays in flight

Charged kaons traveling in the detector sometimes decay before traversing the CDF tracking system. Some of these kaons will either not be reconstructed or will be reconstructed incorrectly. We need to correct for this effect in the data since we would like to compare with the fragmentation level Monte-Carlo.

If a kaon of transverse momentum p_T travels from the beam axis to the TOF detector (positioned at radius R) then the probability that it will decay before reaching the TOF is

$$P(K^+ \to X) = 1 - exp\left(-arcsin\left(\frac{RC}{2p_T}\right) \cdot \frac{2mc}{c\tau C}\right), \tag{6.6}$$

where $C = \frac{p_T}{R_{curvature}}$ is a constant relating track's p_T and curvature, m is the kaon rest mass, τ is the kaon lifetime, and c is the speed of light.

Figure 6.3 shows this probability as a function of kaon p_T in the p_T region used in our analysis. LKon decay is a significant effect only in the lower p_T range.

In the kaon's rest frame, the charged decay product of kaon decay (either π or μ) has a momentum around 300 MeV/c distributed isotropically. The momentum vector is boosted by the original kaon velocity to give the four-momentum in the detector's frame of reference. Given that the p_T of the kaon is not large (say below 1-1.5 GeV/c), the momentum of the charged decay product differs from the original kaon momentum by at least 10-15%. This difference is randomly distributed among all the components of the 3-momentum vector. Three possibilities exist for how the kaon is reconstructed in the detector:

1. The kaon decays inside the SVX detector or in the first layer of the COT. Its decay product is reconstructed as a COT track. However the efficiency for this track to have SVX hits and a small impact parameter is low. Nevertheless, if it is selected, it contributes to the measured pion fraction, since the actual track is produced by either pion or muon particle.

2. The kaon decays inside of the COT detector. These kaons will most likely not be reconstructed - either because they fail the fit for the helix in the COT, or because they fail the match of the TOF/SVX hits onto one track trajectory. If such a kaon is reconstructed, it may contribute to either pion or kaon measured fraction depending on where the decay occured in the detector.

3. The kaon decays inside the last layer of COT. These kaons are likely to be reconstructed and contribute to the measured kaon fraction.

The higher p_T kaons will produce decay products of almost identical momentum to the parent kaon and therefore are much more likely to be reconstructed.

There are therefore two effects that have to be taken into account. The first is the efficiency for the kaons that decay to be reconstructed in the detector and to pass our selection criteria. The second is the contribution of some of the decaying kaons to the measured pion fraction. Table 6.2 gives a summary of our estimates of the size of these effects.

We measure the first effect by calculating the efficiency for the particles found in the MC truth information to be reconstructed as a CDF track that satisfies

p_T, GeV/c	K reaching TOF, %	K found as a track, %	K id-ed as K, %	K id-ed as π, %
0.4-0.7	71.6	82.9 ± 1.5	79.0 ± 1.5	3.90 ± 1.5
0.7-1.0	79.7	91.7 ± 1.4	88.4 ± 1.4	3.30 ± 1.5
1.0-1.5	84.8	98.3 ± 1.2	95.4 ± 1.7	3.95 ± 1.4
1.5-2.5	91.8	99.6 ± 1.1	95.6 ± 2.3	3.98 ± 2.0
2.5-5.0	93.8	$100. \pm 1.2$	97.0 ± 1.5	3.00 ± 1.5

Table 6.2: For different p_T ranges, shown are the probability that kaon reaches TOF system before it decays, the probability that kaon will be reconstructed as a CDF track satisfying our selection criteria, and the probability that it will contribute to kaon or pion measured fraction (so that numbers in the third column equal the sum of the numbers in the 4th and 5th columns).

our selection criteria. This efficiency is first calculated for true pions and then for true kaons. The ratio of these two efficiencies gives us the efficiency for kaon reconstruction that is due to the decays in flight. The HEPG particles are selected in a cone around B decay and within detector acceptance.

To measure the number of kaons that contribute to the pion fraction, we measure the particle fractions in a sample of true kaons in Monte-Carlo. Selecting tracks matched to the kaons, we measure the particle fraction using TOF measurements, and we find that a small fraction of the kaons are indeed identified as pions. The relative size of these pion and kaon contributions in the Monte-Carlo allows us to estimate the fraction of kaons that undergo decay in flight and later contribute to the pion component of the sample.

Unfortunately measuring this second correction can only be done reliably in the first 3 p_T ranges. The dE/dx measurement is not simulated well in Monte-Carlo, and the COT hit efficiency for kaons undergoing decay in flight may also be simulated incorrectly. This becomes important at higher momentum because particle identification relies heavily on the dE/dx measurement in higher p_T regions where the separation power of TOF is reduced. However in the higher p_T ranges the decay in flight also has much smaller effect. Our approach there is to assume that the kaons that decay in flight contribute equally to the pion and kaon fractions. We assume 50% systematic error on the pion contribution, which should safely cover the uncertainty.

6.5 Monte-Carlo Generation

One of the primary motivation of this analysis is to compare the particle content seen around the B signals in the data with the predictions of the Monte-Carlo generators. As discussed in the Introduction, we rely on the fragmentation models used in the Monte-Carlo generators such as `Pythia` to predict the performance of the same-side tagging algorithms used on $\bar{B}_s^0$ decays. Seeing how well the Monte-Carlo reproduces the production of various particle species seen around B decays is an important piece of information when deciding whether the Monte-Carlo simulation can be trusted.

In this section we describe the procedure we used to generate a sample of simulated events and how we use them to compare our measurements in the data to the predictions of the `Pythia` [46] generator.

We start by generating the `Pythia/Jetset` events that simulate the collisions of $p\bar{p}$ in the Tevatron and the following fragmentation process, as well as the underlying event. The b quark production at Tevatron is dominated by the processes of flavor creation, flavor oscillation and gluon splitting. The Monte-Carlo must simulate all three processes. We run the `Pythia` generator with the parameter `msel=1` which means all these quark creation processes are simulated. This however also means that all quark flavors have to be produced in the simulation, and we filter and discard all events that do not contain a b quark. Using this generation procedure increases the amount of the computing power necessary to generate the sample by

orders of magnitude but this increase cannot be avoided.

After `Pythia` proceeds to simulate the fragmentation and forms the composite particles the simulated event is passed on the `EvtGen` [35] decay package, which simulates the decays of all the unstable particles. We force the decays of some of the particles. Specifically, for B mesons containing a b quark, we simulate only the semileptonic decays containing muons in the final state. The decays of the particles containing the anti-quark $\bar{b}$ are left unbiased. The decays of the D mesons that are the decay products of the B mesons containing a b quark are also forced so that the lowest-level D meson states are allowed to decay only in the channels that we reconstruct. For example $D^0 \to K\pi$ is the only allowed decay for D^0 meson. Again, this was only done for the decays following the path of the b quark and does not bias the decay of the other B hadron in the event.

After these decays are simulated the event is processed by detector response simulation package based on GEANT [14], [13]. This simulates the low-level response of all detector systems, allowing for a reliable estimate of trigger, detector and reconstruction effects. We use the standard CDF trigger simulation package `TRIGsim` to simulate trigger decisions of the Level 1 and 2 trigger systems based on the raw data banks. The events are filtered based on the `TRIGsim` decision. Finally, the standard CDF reconstruction is performed on the events to create the final sample. The complete Monte-Carlo sample is then processed and analyzed in the same way as the data sample from the detector.

The number of the partially reconstructed B mesons in our Monte-Carlo sample is roughly equal to the same number found in the data. However access to the truth information (information about which particles were simulated) in Monte-Carlo allows us to avoid the reduction in statistical power due to the detector inefficiencies and having to use imprecise TOF/dE/dx measurements. This allows for much smaller uncertainties on the measured particle content in the Monte-Carlo sample since we can just count the particles in the truth information record.

From the description of the simulation process it follows that the relative sample composition of the $\bar{B}^0$ and B^- decays is reproduced by the Monte-Carlo with the best precision allowed by the current knowledge of the charmed and strange decays. However, since the semimuonic decays of the B mesons are selected, the backgrounds in the main $\bar{B}^0_s$ mode from the $\bar{B}^0, B^-$ mesons [8] are not present in the Monte-Carlo sample. For this reason, direct comparison between the data and Monte-Carlo is not performed based on the results obtained on the samples of tracks found around lD signals. The final comparisons are done only for the yields and fractions calculated for the tracks associated with the B meson signals, as opposed to the lD signals. In other words, the data to Monte-Carlo comparisons that we present in Section 7 are based on the results that were obtained in the data by correcting for the sample composition as described in Section 5, and in Monte-Carlo by selecting the events coming from the appropriate B meson based on the truth information.

[8] These backgrounds are the result the hadronic decays of the B followed by the semileptonic decays of the daughter D mesons.

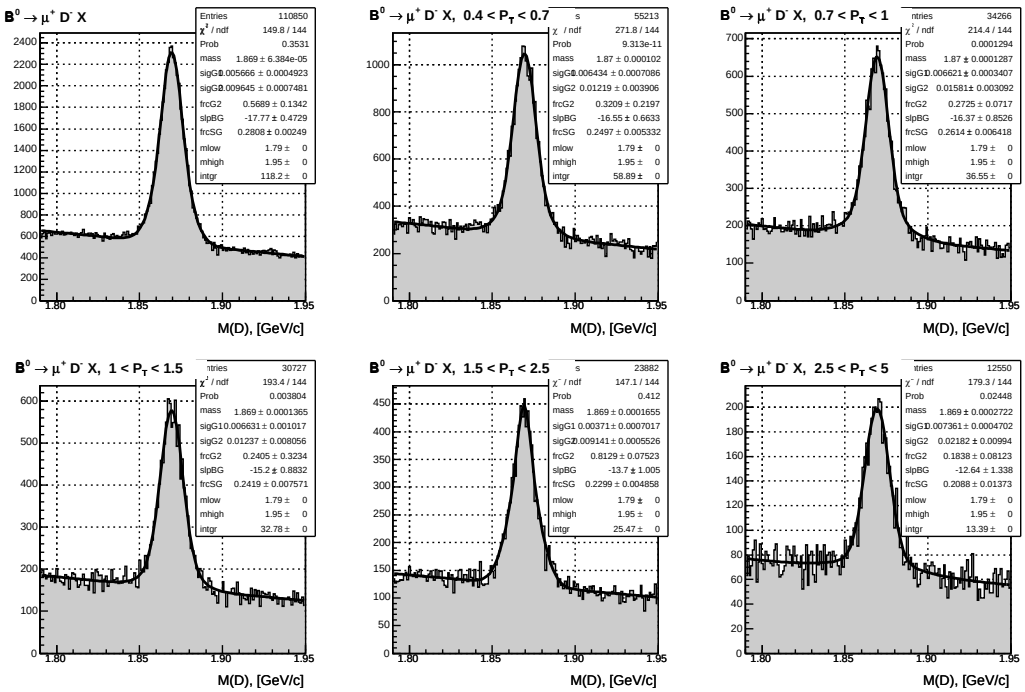


Figure 6.1: Examples of mass fits used to set signal and background probabilities. For each fit iteration a histogram of D invariant mass is made with one entry per track in the sample.

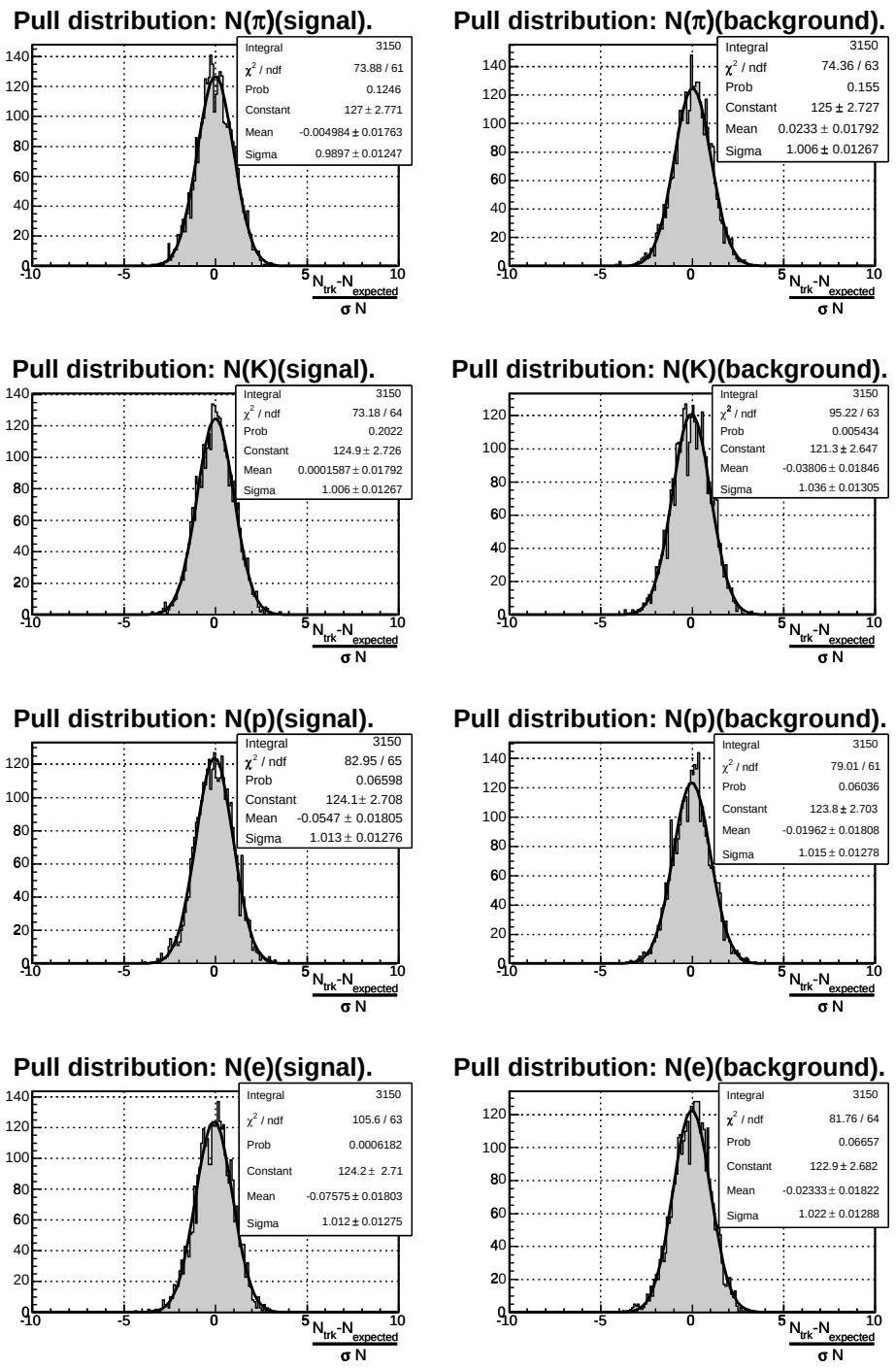


Figure 6.2: The pull distributions of the numbers of particles measured in the fit validation tests performed on μD^- sample, as explained in the text.

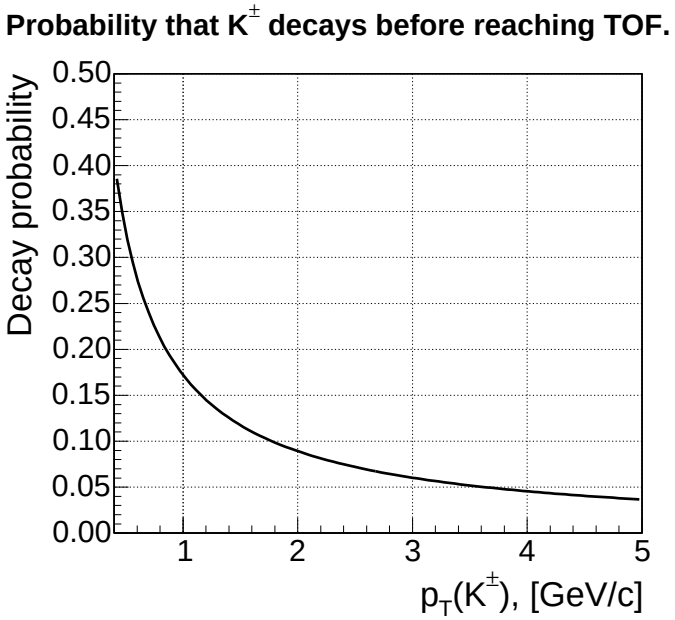


Figure 6.3: Probability that a $K^{\pm}$ produced at the beamline decays in flight before reaching the TOF system.

Chapter 7

Results and Discussion

At this point we have described the tools and techniques developed for this analysis. In this Chapter we give an overview of our results, and present the conclusions and the interpretations of our measurements. Our analysis has two goals. The first goal is to measure the charged particle production in association with B meson formation. This is the first such measurement at a hadron collider and it can be used to tune Monte-Carlo generators or fragmentation models. Second, we compare our measurement to the prediction of the current `Pythia` generator to provide information on using Monte-Carlo simulation to predict the same-side tagging performance.

Keeping in mind those goals, we first present our "raw" uncorrected measurements and then proceed to show the results corrected for all the effects mentioned in the preceding discussion, such as applying sample composition corrections and taking into account the efficiencies. Finally, the discussion of the data follows,

concentrating on comparing the data and the Monte-Carlo, and examining the differences between the B mesons.

The majority of the particles that are reconstructed as the tracks in the detector belong to one of the five species: π, K, p, e, μ. Our particle identification capabilities allow us to distinguish between all these types with the exception of muons and pions that have similar masses and therefore are practically identical in both the TOF and the COT detectors. Therefore in all the results quoted in this dissertation the measurements quoted for the pions should be understood as "including the muon contribution".

7.1 Summary of the measurements

Tables 7.2, 7.3 and 7.4 show the measured fractions and yields of the charged tracks reconstructed around the lD signals in each of our trigger samples as well as the combined measurement. The columns are the decay type around which the tracks were selected, the boundaries of the p_T ranges, the fractions of π, K, p in percent and their yields, number of tracks in the bin and finally total number of events in the sample. The yield is calculated as $M(i) = N_{TRK} * f_i / N_{EVT}$ for each particle type (i) in each p_T bin.

These tables show the actual fit results with no corrections applied. The errors quoted in these tables are the statistical errors from the errors on the fit parameters only. As mentioned in the Section 6.5 our Monte-Carlo sample does not have all the

components of the data sample, therefore we do not make any direct comparison between data and Monte-Carlo for measurements corresponding to the lD signals.

Special care is required to combine eSVT and μSVT trigger samples. Electrons cannot be reliably simulated by the Monte-Carlo, so we only make comparisons to the μSVT Monte-Carlo sample. We measure the detector acceptance corrections to the measured yields only for the μSVT sample. The difference in the measured yields in the eSVT and μSVT samples is expected to be significant because trigger electrons have larger acceptance than trigger muons in the CDF detector ($|\eta| < 1$ for central electrons versus $|\eta| < 0.6$ for central muons) so the tracks associated with lD signals also have different acceptance. On the other hand, measured *fractions* of particles associated with B meson production are expected to be identical in both trigger samples (and they indeed agree quite well). Therefore our approach to combining the data samples is to run the full analysis on the combined lSVT sample to measure the particle fractions, but the measured yields are rescaled to the values measured in the μSVT sample only.

To conclude the discussion of the raw measurements done in the lD samples, we present Figures 7.2, 7.3 and 7.4 which show the examples of likelihood fits in lD^0 channel. The quantities in the histograms are $log(dE/dx)$ and $t_\pi - t_0$ as defined in previous discussion. The comparison between data and the fit is made by calculating a probability density function for every track in the fitted sample and drawing a sum of such functions over the histogram. The total probability functions for K, π, p, e

hypotheses are also shown in blue, red, magenta and brown correspondingly.

Figures 7.5 through 7.13 show the yields and fractions of the charged particle species seen in the cone $\Delta R < 0.7$ around the partially reconstructed B decays. At this point all the corrections are included. The error bars include the systematic uncertainties. The blue curves show the results from the data, and the red curves correspond to the Monte-Carlo predictions. These plots present the measurements of the particle content as a function of track p_T. The yields in these Figures were normalized to the widths of the p_T bins in which the fits were performed. For reference and convenience the data presented in these graphs are also available in Tables 7.6, 7.7 and 7.8.

The fractions of the particle species in the full sample of the tracks (*i.e.*, over the entire p_T range from 400 MeV/c to 5 GeV/c) after all the corrections are shown in Table 7.1 for B mesons. The errors shown are first the statistical error, then the systematic uncertainty that is strongly correlated between all included B decay modes, and finally the uncorrelated systematic uncertainty. These systematic errors are summarized in Section 7.5.

7.2 Comparing data and Monte-Carlo

The figures and tables mentioned in the previous section show all the relevant measurements between data and Monte-Carlo. In general, the Monte-Carlo seems to reproduce the data well in the higher p_T region, with discrepancies growing as

Total fractions around B mesons, data Value $\pm$ (stat.) $\pm$(common syst.) $\pm$ (syst.)			
	π,%	K,%	p,%
B^0	$81.4 \pm 0.10 \pm 1.6 \pm 0.33$	$11.6 \pm 0.25 \pm 0.34 \pm 0.085$	$5.05 \pm 0.15 \pm 0.12 \pm 0.072$
B^+	$78.9 \pm 0.12 \pm 3.1 \pm 0.35$	$12.1 \pm 0.28 \pm 0.57 \pm 0.093$	$6.70 \pm 0.17 \pm 0.22 \pm 0.077$
B_s	$71.0 \pm 0.73 \pm 1.3 \pm 1.8$	$20.2 \pm 1.2 \pm 0.59 \pm 0.29$	$6.62 \pm 0.74 \pm 0.17 \pm 0.055$
Total fractions around B mesons, Monte-Carlo			
	π,%	K,%	p,%
B^0	84.357 ± 0.074	10.3 ± 0.15	5.05 ± 0.11
B^+	82.039 ± 0.077	11.2 ± 0.14	6.40 ± 0.11
B_s	70.24 ± 0.54	25.0 ± 0.82	4.40 ± 0.44

Table 7.1: Total fractions of particles seen in a cone around partially reconstructed B signals.

p_T becomes lower. This is hardly surprising as historically most studies and tuning efforts concentrated on the higher momentum region.

The production rate of kaons around $\bar{B}_s^0$ mesons is consistently lower in the data compared to the Monte-Carlo. This may be an indication that Monte-Carlo prediction of the same-side kaon tagging power may be an overestimate. One interesting feature appearing in the plots of particle yields versus track p_T is that production of protons and kaons seems to be slightly suppressed at lower p_T, probably due to the heavier mass of these particles. This effect is reproduced in Monte-Carlo reasonably well for protons, but is missing for kaons.

Another small difference between data and Monte-Carlo is the presence of electrons at a small level in the data sample, which is mostly due to remaining secondary particles (such as conversions) that were not removed by our selection requirements. These are a source of background in our sample and could be removed by sideband

subtraction. We did not make this subtraction for two reasons: first because the statistics is insufficient in the sideband region to reliably measure the particle content, and second because the particle content may depend on the particle impact parameter[1].

7.3 Comparing measurements in $\bar{B}_s^0, \bar{B}^0$ and B^- samples.

The difference in particle production between the different B meson species may be indicative of the differences in the fragmentation processes leading to their creation. We observe that kaons are more likely to be produced around $\bar{B}_s^0$ mesons than around the other two species. The statistical significance of this observation is 5.3σ when the kaon fraction is compared.

However, when comparing particle production around different B modes one can choose to compare either the fractions or yields. We explore this issue by looking at the fractions and yields around the lD signals in subsamples of the lD signals with different p_T. Figures 7.14, 7.15 and 7.16 show how the kaon and proton fraction and the total particle yield change with the p_T of the lD signal. These results are not corrected for the sample composition since selecting lD candidates based on their p_T

[1]Different types of secondary particles all have different lifetimes and decay product composition, while the tracks from secondary $p\bar{p}$ interactions produce uniformly distributed minimum bias events overlaying the main event.

changes the sample content. Also, since the Monte-Carlo does not reproduce all the sample composition components one should not directly compare data with Monte-Carlo measurements in these plots, but compare the trends visible in these samples. The yield shows significant dependence on the B meson p_T while the fractions seem to remain constant (we did observe some correlation between the p_T spectra of the B mesons and the associated tracks, so the fractions can not be exactly constant, but this dependence is smaller than the uncertainties of our measurements).

Because of this strong dependence of multiplicity on $p_T(B)$, we re-weighted all the lD modes in this analysis to the same B meson p_T spectrum using the average K-factors[2] measured in Monte-Carlo (0.8255, 0.8595, 0.8310, 0.8362 for μ^+D^-, μ^+D^{*-}, $\mu^+\bar{D}_0$, $\mu^+D_s^-$ modes respectively). All the results shown in this dissertation were obtained using these re-weighted p_T spectra. Still, the yields around different B modes are not the same, and comparing the relative production of particles (fractions) may be more justified.

One could argue that re-weighting the events based on the B meson transverse momentum is not correct for a meaningful comparison between the B meson species, since fragmentation depends on the initial properties of the quarks and gluons involved. For example, if one could calculate the energy remaining in the string after the B meson was formed, then re-weighting the events based on that variable would

[2] K-factor is a quantity commonly used to describe the semileptonic decays and is defined as the ratio of the combined momentum of all reconstructed decay products to the momentum of the decaying B meson.

allow for a different type of comparison of particle production around $\bar{B}_s^0, \bar{B}^0, B^-$. However even in this case the heavier kaon particles would have a different p_T spectrum from the pions, and, therefore, different efficiency to be reconstructed in the detector. This difference implies that a statement such as "more kaons are expected to be produced around $\bar{B}_s^0$ mesons compared to $\bar{B}^0, B^-$" would still be very difficult to unambiguously check experimentally. More importantly, such a procedure would mean that the analysis would be tied to one of the string fragmentation models in that it would rely on its calculation of the energy of the string, and its prediction of the detector acceptance for fragmentation particles. Our analysis as presented here avoids any dependence on the theoretical fragmentation models.

In conclusion, it is very difficult (if not impossible) to compare conclusively and quantitatively the particle production around different B meson species. However, some evidence of higher kaon production in association with $\bar{B}_s^0$ meson formation is seen compared to the $\bar{B}^0, B^-$ mesons. Such an observation is expected from the quark model[3].

[3]Interestingly, if quark model did not exist the results of our analysis could be used to suggest that such model needs to be created. The "mysterious" connection between $\bar{B}_s^0$ mesons and $K^{\pm}$ mesons when produced in the $p\bar{p}$ collisions is naturally explained by the presence of the same type of constituent in both particles.

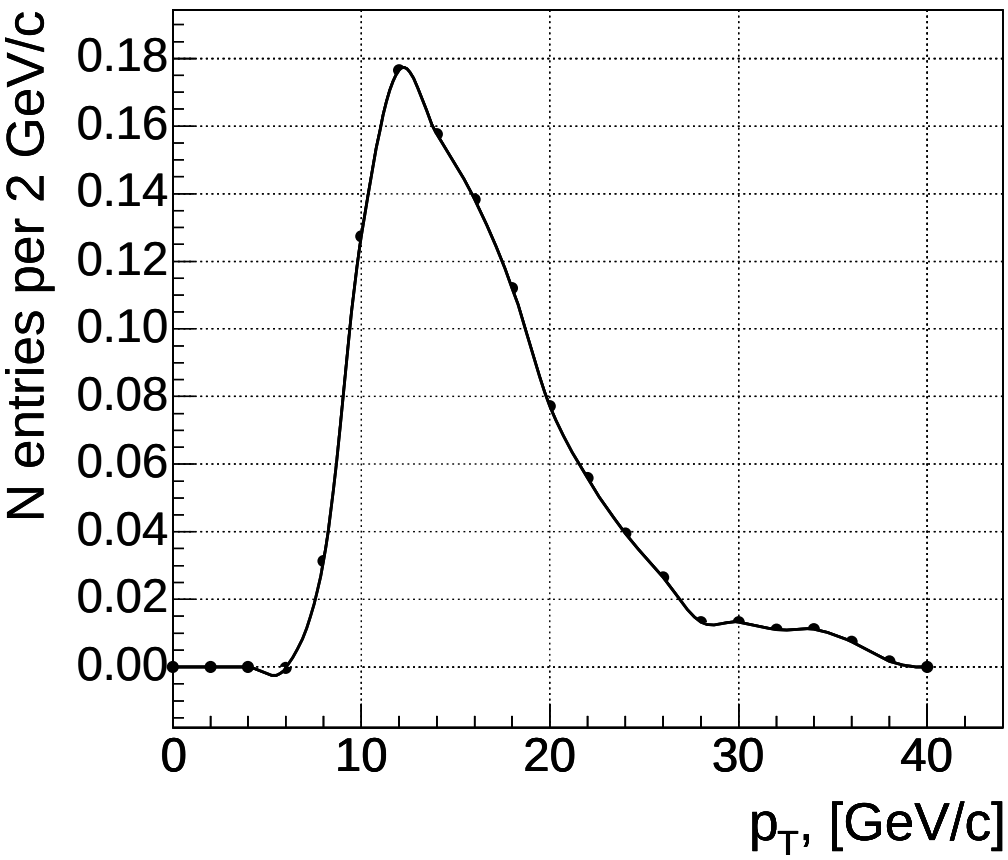


Figure 7.1: Spectrum of the B meson transverse momentum to which all the reconstructed decays were rescaled using average K-factors.

Signal Fit - μ SVT							
Type	p_T	π,%	K,%	p,%	e, %	N_{TRK}	N_{EVT}
μ^+D^-	0.4-0.7	88.89 ± 0.13	6.23 ± 0.31	1.95 ± 0.18	2.93 ± 0.31	10649(0.356)	29898
μ^+D^-	0.7-1	84.73 ± 0.23	9.56 ± 0.50	4.49 ± 0.31	1.22 ± 0.29	6699(0.224)	29898
μ^+D^-	1-1.5	77.61 ± 0.36	13.5 ± 0.70	7.72 ± 0.43	1.21 ± 0.32	5860(0.196)	29898
μ^+D^-	1.5-2.5	75.64 ± 0.50	14.7 ± 1.1	8.23 ± 0.66	1.45 ± 0.44	4441(0.149)	29898
μ^+D^-	2.5-5	70.8 ± 0.97	20.6 ± 2.3	6.34 ± 1.7	2.28 ± 1.0	2051(0.0686)	29898
$\mu^+D^-,\sum$	0.4-5	82.50 ± 0.13	10.7 ± 0.33	4.90 ± 0.20	1.94 ± 0.17	29702(0.993)	29898
μ^+D^{*-}	0.4-0.7	88.92 ± 0.15	6.23 ± 0.35	1.99 ± 0.18	2.85 ± 0.35	4316(0.339)	12722
μ^+D^{*-}	0.7-1	81.70 ± 0.34	12.1 ± 0.61	4.66 ± 0.41	1.58 ± 0.37	2660(0.209)	12722
μ^+D^{*-}	1-1.5	77.18 ± 0.45	13.4 ± 0.85	7.67 ± 0.55	1.73 ± 0.37	2532(0.199)	12722
μ^+D^{*-}	1.5-2.5	77.18 ± 0.57	12.6 ± 1.3	8.96 ± 0.88	1.22 ± 0.49	1774(0.139)	12722
μ^+D^{*-}	2.5-5	69.5 ± 1.2	13.8 ± 3.1	13.0 ± 2.2	3.71 ± 1.3	851(0.0669)	12722
$\mu^+D^{*-},\sum$	0.4-5	81.85 ± 0.16	10.5 ± 0.39	5.53 ± 0.26	2.16 ± 0.20	12135(0.954)	12722
$\mu^+\bar{D}_0$	0.4-0.7	88.96 ± 0.091	6.25 ± 0.20	1.83 ± 0.090	2.95 ± 0.19	18539(0.353)	52555
$\mu^+\bar{D}_0$	0.7-1	81.86 ± 0.19	10.9 ± 0.35	4.90 ± 0.22	2.32 ± 0.21	11698(0.223)	52555
$\mu^+\bar{D}_0$	1-1.5	76.95 ± 0.25	13.5 ± 0.46	8.17 ± 0.30	1.36 ± 0.22	11218(0.213)	52555
$\mu^+\bar{D}_0$	1.5-2.5	70.63 ± 0.40	16.1 ± 0.74	11.8 ± 0.46	1.48 ± 0.30	8325(0.158)	52555
$\mu^+\bar{D}_0$	2.5-5	65.2 ± 0.79	19.5 ± 1.7	12.5 ± 1.3	2.77 ± 0.72	3776(0.0719)	52555
$\mu^+\bar{D}_0,\sum$	0.4-5	80.38 ± 0.098	11.3 ± 0.23	6.13 ± 0.15	2.24 ± 0.12	53557(1.02)	52555
$\mu^+D_s^-$	0.4-0.7	85.79 ± 0.62	8.61 ± 1.2	2.46 ± 0.61	3.14 ± 1.1	700(0.278)	2519
$\mu^+D_s^-$	0.7-1	73.5 ± 1.5	17.9 ± 2.1	6.19 ± 1.3	2.40 ± 1.2	414(0.165)	2519
$\mu^+D_s^-$	1-1.5	69.1 ± 2.0	22.1 ± 2.8	7.36 ± 1.7	1.46 ± 1.4	372(0.148)	2519
$\mu^+D_s^-$	1.5-2.5	63.9 ± 2.9	20.3 ± 4.4	14.2 ± 2.8	1.62 ± 2.1	289(0.115)	2519
$\mu^+D_s^-$	2.5-5	44.5 ± 8.7	$18.0 \pm 13.$	35.3 ± 7.4	2.19 ± 6.5	133(0.0529)	2519
$\mu^+D_s^-,\sum$	0.4-5	73.8 ± 0.75	15.7 ± 1.4	8.22 ± 0.97	2.36 ± 0.77	1910(0.758)	2519

Table 7.2: The fit results for the signal component in μSVT trigger sample. The measured fractions are shown for all track p_T ranges, as well as for the entire sample. The rightmost two columns show the number of tracks found in the p_T ranges (and also this number normalized to the number of the reconstructed lD signal) and the number of the reconstructed lD signal events.

Signal Fit - e SVT							
Type	p_T	π,%	K,%	p,%	e, %	N_{TRK}	N_{EVT}
e^+D^-	0.4-0.7	89.61 ± 0.15	5.76 ± 0.35	1.91 ± 0.18	2.72 ± 0.34	6489(0.3)	21607
e^+D^-	0.7-1	83.51 ± 0.30	10.5 ± 0.62	4.26 ± 0.37	1.69 ± 0.37	4232(0.196)	21607
e^+D^-	1-1.5	78.65 ± 0.43	13.9 ± 0.85	5.96 ± 0.51	1.49 ± 0.39	3668(0.17)	21607
e^+D^-	1.5-2.5	72.85 ± 0.69	15.7 ± 1.4	10.3 ± 0.82	1.16 ± 0.52	2692(0.125)	21607
e^+D^-	2.5-5	77.7 ± 0.96	15.1 ± 3.1	9.96 ± 2.2	-2.72 ± 1.1	1126(0.0522)	21607
$e^+D^-, \sum$	0.4-5	82.78 ± 0.16	10.6 ± 0.39	5.00 ± 0.24	1.67 ± 0.20	18210(0.843)	21607
e^+D^{*-}	0.4-0.7	86.39 ± 0.26	6.58 ± 0.52	4.14 ± 0.34	2.89 ± 0.49	2795(0.284)	9858
e^+D^{*-}	0.7-1	82.66 ± 0.39	9.49 ± 0.75	5.69 ± 0.55	2.16 ± 0.47	1873(0.19)	9858
e^+D^{*-}	1-1.5	78.07 ± 0.53	13.2 ± 0.89	8.22 ± 0.71	0.511 ± 0.35	1653(0.168)	9858
e^+D^{*-}	1.5-2.5	77.93 ± 0.70	13.2 ± 1.7	9.64 ± 1.1	-0.753 ± 0.31	1141(0.116)	9858
e^+D^{*-}	2.5-5	71.4 ± 1.6	15.4 ± 4.1	10.2 ± 2.8	2.96 ± 1.6	449(0.0456)	9858
$e^+D^{*-}, \sum$	0.4-5	81.79 ± 0.21	10.0 ± 0.46	6.45 ± 0.32	1.72 ± 0.24	7913(0.803)	9858
$e^+\bar{D}_0$	0.4-0.7	88.35 ± 0.11	6.61 ± 0.24	2.00 ± 0.12	3.04 ± 0.23	12760(0.324)	39441
$e^+\bar{D}_0$	0.7-1	82.78 ± 0.20	10.7 ± 0.39	5.02 ± 0.27	1.52 ± 0.23	8099(0.205)	39441
$e^+\bar{D}_0$	1-1.5	76.61 ± 0.30	13.9 ± 0.56	7.85 ± 0.36	1.69 ± 0.25	7166(0.182)	39441
$e^+\bar{D}_0$	1.5-2.5	71.67 ± 0.45	14.2 ± 0.88	13.2 ± 0.56	0.888 ± 0.31	5392(0.137)	39441
$e^+\bar{D}_0$	2.5-5	66.9 ± 0.92	10.6 ± 2.3	19.9 ± 1.5	2.56 ± 0.81	2236(0.0567)	39441
$e^+\bar{D}_0, \sum$	0.4-5	80.87 ± 0.11	10.4 ± 0.26	6.68 ± 0.18	2.07 ± 0.13	35657(0.904)	39441
$e^+D_s^-$	0.4-0.7	88.85 ± 0.62	6.81 ± 1.3	1.73 ± 0.74	2.61 ± 1.4	452(0.258)	1756
$e^+D_s^-$	0.7-1	75.1 ± 1.8	17.0 ± 3.0	5.96 ± 1.8	1.94 ± 1.9	253(0.144)	1756
$e^+D_s^-$	1-1.5	71.3 ± 2.1	20.7 ± 3.4	6.97 ± 2.1	1.00 ± 1.7	316(0.18)	1756
$e^+D_s^-$	1.5-2.5	58.5 ± 4.1	35.6 ± 5.1	4.70 ± 3.1	1.17 ± 1.5	180(0.102)	1756
$e^+D_s^-$	2.5-5	Insufficient statistics.					
$e^+D_s^-, \sum$	0.4-2.5	77.2 ± 0.82	16.6 ± 1.5	4.38 ± 0.86	1.85 ± 0.85	1203(0.685)	1756

Table 7.3: The fit results for the signal component in eSVT trigger sample. The measured fractions are shown for all track p_T ranges, as well as for the entire sample. The rightmost two columns show the number of tracks found in the p_T ranges (and also this number normalized to the number of the reconstructed lD signal) and the number of the reconstructed lD signal events.

Signal Fit - l SVT										
Type	p_T	π,%	K,%	p,%	e, %	$M(\pi)$	$M(K)$	$M(p)$	N_{TRK}	N_{EVT}
l^+D^-	0.4-0.7	89.18 ± 0.10	6.05 ± 0.23	1.94 ± 0.11	2.83 ± 0.22	0.294 ± 0.0031	$((2.00 \pm 0.080)10^{-2}$	$((6.40 \pm 0.38)10^{-3}$	17002(0.33)	51488
l^+D^-	0.7-1	84.16 ± 0.18	10.0 ± 0.39	4.39 ± 0.24	1.40 ± 0.22	0.178 ± 0.0025	$((2.13 \pm 0.092)10^{-2}$	$((9.31 \pm 0.53)10^{-3}$	10913(0.212)	51488
l^+D^-	1-1.5	77.75 ± 0.28	13.8 ± 0.54	7.08 ± 0.33	1.33 ± 0.25	0.144 ± 0.0023	$((2.56 \pm 0.12)10^{-2}$	$((1.31 \pm 0.066)10^{-2}$	9513(0.185)	51488
l^+D^-	1.5-2.5	74.34 ± 0.40	15.2 ± 0.86	9.04 ± 0.51	1.41 ± 0.34	0.102 ± 0.0021	$((2.09 \pm 0.14)10^{-2}$	$((1.24 \pm 0.077)10^{-2}$	7076(0.137)	51488
l^+D^-	2.5-5	72.43 ± 0.72	19.2 ± 1.9	7.35 ± 1.4	0.975 ± 0.76	$((4.44 \pm 0.16)10^{-2}$	$((1.18 \pm 0.14)10^{-2}$	$((4.50 \pm 0.91)10^{-3}$	3158(0.0613)	51488
$l^+D^-,\sum$	0.4-5	82.44 ± 0.10	10.8 ± 0.25	4.94 ± 0.16	1.87 ± 0.13	0.7627 ± 0.0053	$((9.95 \pm 0.26)10^{-2}$	$((4.57 \pm 0.15)10^{-2}$	47666(0.926)	51488
l^+D^{*-}	0.4-0.7	88.74 ± 0.13	6.28 ± 0.28	2.07 ± 0.13	2.91 ± 0.27	0.277 ± 0.0037	$((1.96 \pm 0.095)10^{-2}$	$((6.47 \pm 0.42)10^{-3}$	7068(0.314)	22485
l^+D^{*-}	0.7-1	82.17 ± 0.26	11.0 ± 0.50	5.09 ± 0.33	1.79 ± 0.28	0.163 ± 0.0029	$((2.17 \pm 0.11)10^{-2}$	$((1.01 \pm 0.069)10^{-2}$	4497(0.2)	22485
l^+D^{*-}	1-1.5	77.44 ± 0.34	13.3 ± 0.63	7.88 ± 0.48	1.39 ± 0.27	0.141 ± 0.0027	$((2.43 \pm 0.13)10^{-2}$	$((1.44 \pm 0.095)10^{-2}$	4144(0.184)	22485
l^+D^{*-}	1.5-2.5	77.36 ± 0.42	12.8 ± 1.0	9.31 ± 0.65	0.581 ± 0.32	$((9.82 \pm 0.24)10^{-2}$	$((1.62 \pm 0.15)10^{-2}$	$((1.18 \pm 0.091)10^{-2}$	2888(0.128)	22485
l^+D^{*-}	2.5-5	70.3 ± 0.97	14.3 ± 2.4	11.8 ± 1.7	3.52 ± 1.0	$((3.91 \pm 0.18)10^{-2}$	$((7.98 \pm 1.6)10^{-3}$	$((6.59 \pm 1.1)10^{-3}$	1288(0.0573)	22485
$l^+D^{*-},\sum$	0.4-5	82.08 ± 0.13	10.2 ± 0.30	5.63 ± 0.20	2.04 ± 0.15	0.7186 ± 0.0062	$((8.97 \pm 0.29)10^{-2}$	$((4.93 \pm 0.19)10^{-2}$	19887(0.884)	22485
$l^+\bar{D}_0$	0.4-0.7	88.720 ± 0.065	6.38 ± 0.15	1.90 ± 0.072	2.99 ± 0.15	0.3018 ± 0.0020	$((2.17 \pm 0.056)10^{-2}$	$((6.47 \pm 0.25)10^{-3}$	31289(0.34)	91948
$l^+\bar{D}_0$	0.7-1	82.19 ± 0.14	10.8 ± 0.27	4.97 ± 0.17	2.02 ± 0.16	0.1764 ± 0.0017	$((2.32 \pm 0.064)10^{-2}$	$((1.07 \pm 0.039)10^{-2}$	19748(0.215)	91948
$l^+\bar{D}_0$	1-1.5	76.83 ± 0.19	13.6 ± 0.36	8.04 ± 0.23	1.50 ± 0.16	0.153 ± 0.0016	$((2.71 \pm 0.083)10^{-2}$	$((1.60 \pm 0.050)10^{-2}$	18318(0.199)	91948
$l^+\bar{D}_0$	1.5-2.5	71.03 ± 0.30	15.4 ± 0.57	12.3 ± 0.36	1.22 ± 0.22	0.106 ± 0.0015	$((2.30 \pm 0.100)10^{-2}$	$((1.83 \pm 0.060)10^{-2}$	13720(0.149)	91948
$l^+\bar{D}_0$	2.5-5	65.97 ± 0.60	16.5 ± 1.4	14.9 ± 0.98	2.64 ± 0.55	$((4.31 \pm 0.12)10^{-2}$	$((1.08 \pm 0.11)10^{-2}$	$((9.73 \pm 0.75)10^{-3}$	6026(0.0655)	91948
$l^+\bar{D}_0,\sum$	0.4-5	80.570 ± 0.073	10.9 ± 0.17	6.32 ± 0.11	2.18 ± 0.087	0.7800 ± 0.0036	0.106 ± 0.0019	$((6.12 \pm 0.12)10^{-2}$	89104(0.969)	91948
$l^+D_s^-$	0.4-0.7	87.03 ± 0.44	8.15 ± 0.85	2.12 ± 0.42	2.70 ± 0.82	0.235 ± 0.0092	$((2.20 \pm 0.25)10^{-2}$	$((5.73 \pm 1.2)10^{-3}$	1215(0.274)	4431
$l^+D_s^-$	0.7-1	72.5 ± 1.2	18.4 ± 1.7	6.14 ± 1.0	2.93 ± 1.0	0.113 ± 0.0066	$((2.87 \pm 0.32)10^{-2}$	$((9.58 \pm 1.7)10^{-3}$	708(0.16)	4431
$l^+D_s^-$	1-1.5	70.2 ± 1.4	21.4 ± 2.1	7.52 ± 1.3	0.932 ± 0.96	0.104 ± 0.0070	$((3.16 \pm 0.40)10^{-2}$	$((1.11 \pm 0.21)10^{-2}$	672(0.152)	4431
$l^+D_s^-$	1.5-2.5	60.5 ± 2.3	26.8 ± 3.2	10.5 ± 2.0	2.17 ± 1.5	$((6.66 \pm 0.64)10^{-2}$	$((2.94 \pm 0.48)10^{-2}$	$((1.16 \pm 0.25)10^{-2}$	502(0.113)	4431
$l^+D_s^-$	2.5-5	35.0 ± 7.4	36.5 ± 8.0	27.4 ± 6.4	1.04 ± 4.2	$((1.41 \pm 0.46)10^{-2}$	$((1.47 \pm 0.51)10^{-2}$	$((1.10 \pm 0.35)10^{-2}$	194(0.044)	4431
$l^+D_s^-,\sum$	0.4-5	73.55 ± 0.57	17.5 ± 1.0	6.77 ± 0.67	2.22 ± 0.54	0.533 ± 0.015	0.126 ± 0.0090	$((4.90 \pm 0.52)10^{-2}$	3293(0.743)	4431

Table 7.4: The fit results for the signal component in combined μSVT and eSVT trigger samples. The measured fractions are shown for all track p_T ranges, as well as for the entire sample. The rightmost two columns show the number of tracks found in the p_T ranges (and also this number normalized to the number of the reconstructed lD signal) and the number of the reconstructed lD signal events.

Signal Fit - MC										
Type	p_T	π,%	K,%	p,%	e, %	$M(\pi)$	$M(K)$	$M(p)$	N_{TRK}	N_{EVT}
$\mu^+ D^-$	0.4-0.7	89.71 ± 0.14	7.83 ± 0.37	1.90 ± 0.20	0.553 ± 0.11	0.582 ± 0.0089	$((5.09 \pm 0.26)10^{-2}$	$((1.23 \pm 0.13)10^{-2}$	4794(0.649)	7385
$\mu^+ D^-$	0.7-1	85.14 ± 0.25	10.6 ± 0.53	4.03 ± 0.35	0.276 ± 0.096	0.343 ± 0.0068	$((4.26 \pm 0.24)10^{-2}$	$((1.62 \pm 0.15)10^{-2}$	2977(0.403)	7385
$\mu^+ D^-$	1-1.5	80.04 ± 0.33	13.2 ± 0.58	6.31 ± 0.43	0.449 ± 0.12	0.319 ± 0.0066	$((5.27 \pm 0.27)10^{-2}$	$((2.52 \pm 0.18)10^{-2}$	2946(0.399)	7385
$\mu^+ D^-$	1.5-2.5	78.57 ± 0.41	11.7 ± 0.66	9.38 ± 0.60	0.361 ± 0.13	0.225 ± 0.0055	$((3.35 \pm 0.21)10^{-2}$	$((2.69 \pm 0.19)10^{-2}$	2118(0.287)	7385
$\mu^+ D^-$	2.5-5	74.0 ± 0.76	14.1 ± 1.1	10.9 ± 1.00	0.950 ± 0.33	$((8.78 \pm 0.34)10^{-2}$	$((1.67 \pm 0.15)10^{-2}$	$((1.30 \pm 0.13)10^{-2}$	876(0.119)	7385
$\mu^+ D^-, \sum$	0.4-5	83.92 ± 0.13	10.6 ± 0.25	5.04 ± 0.18	0.466 ± 0.058	1.558 ± 0.015	0.196 ± 0.0052	$((9.36 \pm 0.36)10^{-2}$	13713(1.86)	7385
$\mu^+ D^{*-}$	0.4-0.7	89.76 ± 0.13	7.79 ± 0.34	1.85 ± 0.18	0.592 ± 0.10	0.574 ± 0.0081	$((4.98 \pm 0.24)10^{-2}$	$((1.19 \pm 0.12)10^{-2}$	5571(0.639)	8719
$\mu^+ D^{*-}$	0.7-1	85.18 ± 0.23	10.4 ± 0.49	4.14 ± 0.33	0.302 ± 0.094	0.334 ± 0.0062	$((4.07 \pm 0.22)10^{-2}$	$((1.62 \pm 0.14)10^{-2}$	3414(0.392)	8719
$\mu^+ D^{*-}$	1-1.5	80.23 ± 0.31	12.2 ± 0.54	7.24 ± 0.44	0.307 ± 0.097	0.297 ± 0.0058	$((4.53 \pm 0.23)10^{-2}$	$((2.68 \pm 0.18)10^{-2}$	3229(0.37)	8719
$\mu^+ D^{*-}$	1.5-2.5	77.64 ± 0.40	12.8 ± 0.64	9.31 ± 0.57	0.235 ± 0.099	0.212 ± 0.0049	$((3.50 \pm 0.20)10^{-2}$	$((2.54 \pm 0.17)10^{-2}$	2379(0.273)	8719
$\mu^+ D^{*-}$	2.5-5	76.26 ± 0.66	13.6 ± 1.0	9.58 ± 0.89	0.581 ± 0.24	$((8.71 \pm 0.32)10^{-2}$	$((1.55 \pm 0.13)10^{-2}$	$((1.09 \pm 0.11)10^{-2}$	996(0.114)	8719
$\mu^+ D^{*-}, \sum$	0.4-5	84.07 ± 0.12	10.4 ± 0.23	5.10 ± 0.17	0.414 ± 0.051	1.503 ± 0.013	0.186 ± 0.0046	$((9.12 \pm 0.32)10^{-2}$	15590(1.79)	8719
$\mu^+ \bar{D}_0$	0.4-0.7	90.021 ± 0.065	7.47 ± 0.17	1.94 ± 0.094	0.572 ± 0.052	0.6098 ± 0.0044	$((5.06 \pm 0.13)10^{-2}$	$((1.31 \pm 0.065)10^{-2}$	20903(0.677)	30857
$\mu^+ \bar{D}_0$	0.7-1	84.61 ± 0.13	10.6 ± 0.26	4.43 ± 0.18	0.385 ± 0.055	0.3483 ± 0.0034	$((4.35 \pm 0.12)10^{-2}$	$((1.82 \pm 0.077)10^{-2}$	12700(0.412)	30857
$\mu^+ \bar{D}_0$	1-1.5	79.43 ± 0.17	13.3 ± 0.29	7.00 ± 0.22	0.296 ± 0.049	0.309 ± 0.0032	$((5.16 \pm 0.13)10^{-2}$	$((2.72 \pm 0.094)10^{-2}$	12008(0.389)	30857
$\mu^+ \bar{D}_0$	1.5-2.5	72.26 ± 0.25	15.0 ± 0.34	12.3 ± 0.32	0.426 ± 0.068	0.214 ± 0.0026	$((4.44 \pm 0.12)10^{-2}$	$((3.66 \pm 0.11)10^{-2}$	9153(0.297)	30857
$\mu^+ \bar{D}_0$	2.5-5	70.38 ± 0.40	15.9 ± 0.54	13.0 ± 0.50	0.709 ± 0.13	$((8.90 \pm 0.17)10^{-2}$	$((2.01 \pm 0.081)10^{-2}$	$((1.65 \pm 0.073)10^{-2}$	3902(0.126)	30857
$\mu^+ \bar{D}_0, \sum$	0.4-5	82.606 ± 0.065	11.1 ± 0.12	5.87 ± 0.094	0.461 ± 0.028	1.571 ± 0.0071	0.210 ± 0.0026	0.112 ± 0.0019	58668(1.9)	30857
$\mu^+ D_s^-$	0.4-0.7	82.39 ± 0.57	15.1 ± 1.2	1.79 ± 0.47	0.699 ± 0.30	0.479 ± 0.019	$((8.79 \pm 0.80)10^{-2}$	$((1.04 \pm 0.28)10^{-2}$	791(0.581)	1361
$\mu^+ D_s^-$	0.7-1	74.6 ± 1.00	21.6 ± 1.7	3.60 ± 0.83	0.166 ± 0.18	0.266 ± 0.014	$((7.69 \pm 0.75)10^{-2}$	$((1.28 \pm 0.31)10^{-2}$	484(0.356)	1361
$\mu^+ D_s^-$	1-1.5	63.6 ± 1.4	31.0 ± 1.9	5.31 ± 1.1	0 ± 0	0.193 ± 0.012	$((9.43 \pm 0.83)10^{-2}$	$((1.61 \pm 0.34)10^{-2}$	413(0.304)	1361
$\mu^+ D_s^-$	1.5-2.5	54.0 ± 1.9	38.1 ± 2.2	7.84 ± 1.5	0 ± 0	0.122 ± 0.0095	$((8.63 \pm 0.80)10^{-2}$	$((1.78 \pm 0.36)10^{-2}$	307(0.226)	1361
$\mu^+ D_s^-$	2.5-5	42.8 ± 3.2	45.5 ± 3.1	11.1 ± 2.5	0.680 ± 0.70	$((4.36 \pm 0.57)10^{-2}$	$((4.63 \pm 0.58)10^{-2}$	$((1.13 \pm 0.29)10^{-2}$	138(0.102)	1361
$\mu^+ D_s^-, \sum$	0.4-5	70.34 ± 0.54	25.0 ± 0.81	4.36 ± 0.43	0.341 ± 0.13	1.10 ± 0.028	0.392 ± 0.017	$((6.84 \pm 0.71)10^{-2}$	2136(1.57)	1361

Table 7.5: The fit results for the signal component in μSVT Monte-Carlo sample. The measured fractions are shown for all track p_T ranges, as well as for the entire sample. The rightmost two columns show the number of tracks found in the p_T ranges (and also this number normalized to the number of the reconstructed lD signal) and the number of the reconstructed lD signal events.

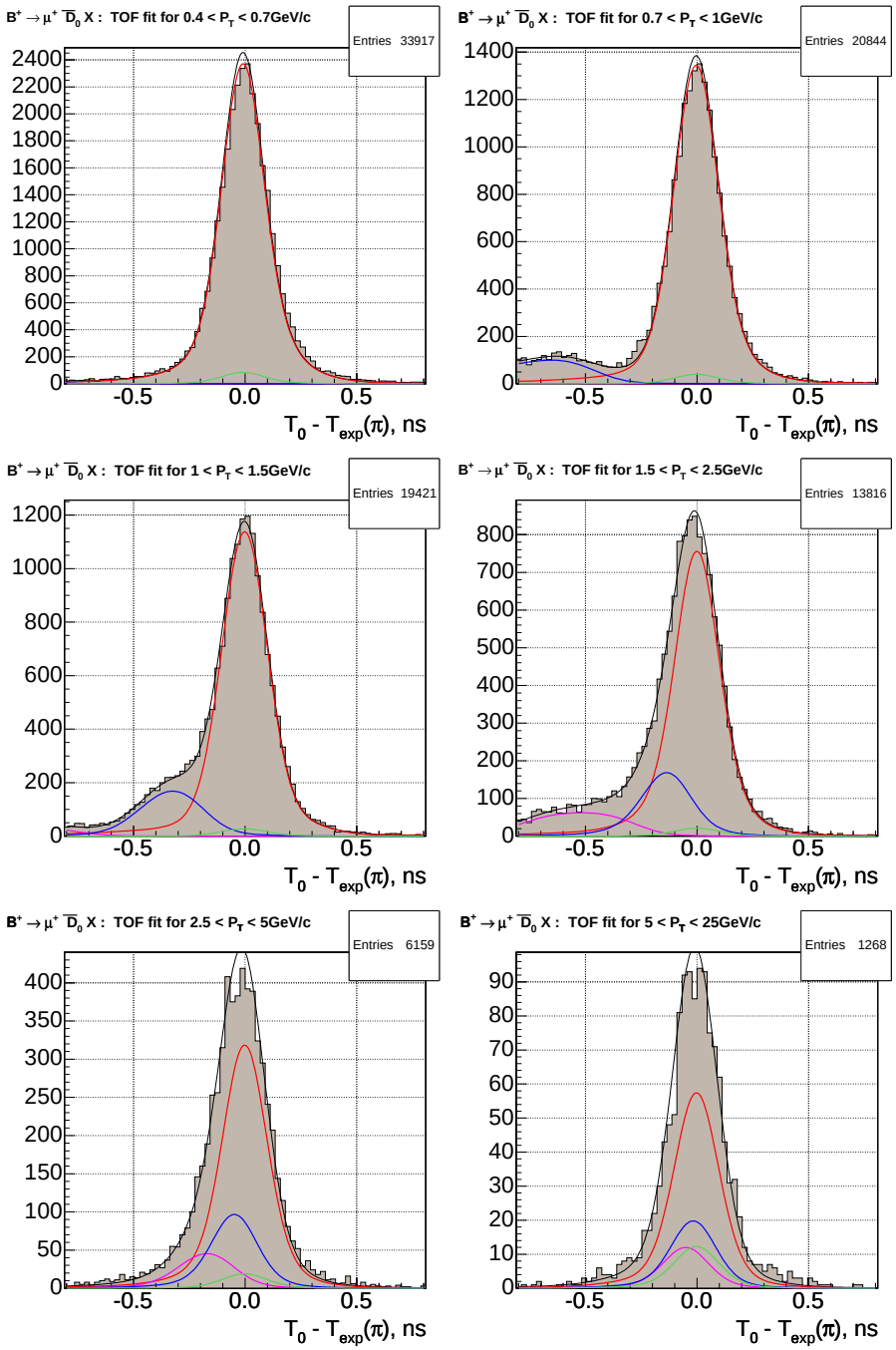


Figure 7.2: Illustration of the likelihood fits for the particle fractions in the lD^0 channels. Shaded histogram is a distribution of the measured time of flight of the particles compared to the time of flight expected for the particles of pion mass. The curves correspond to the same distribution expected to be produced by the numbers of particles of each type that were measured by the particle content fit.

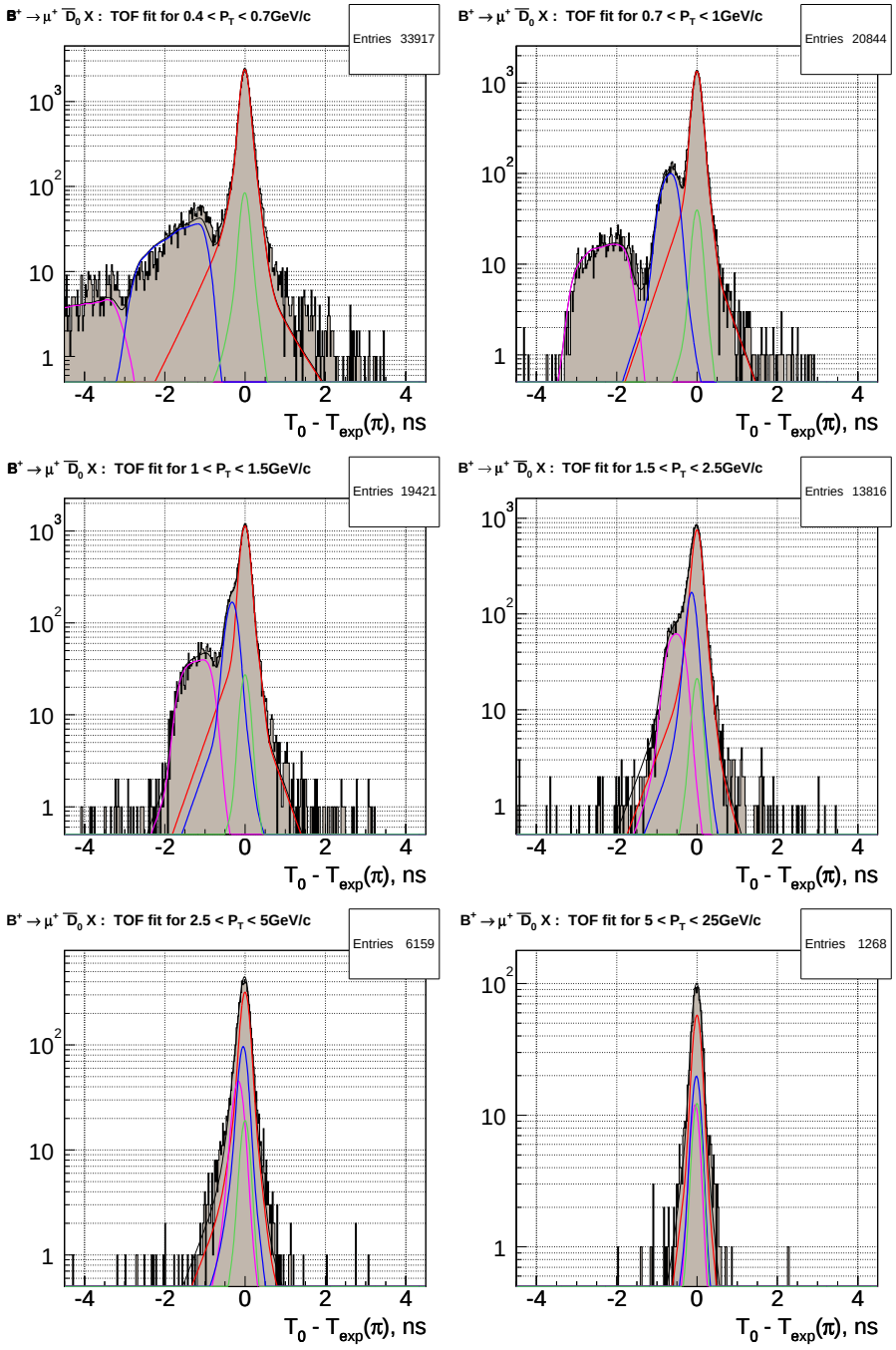


Figure 7.3: Illustration of the likelihood fits for the particle fractions in the lD^0 channels. Shaded histogram is a distribution of the measured time of flight of the particles compared to the time of flight expected for the particles of pion mass. The curves correspond to the same distribution expected to be produced by the numbers of particles of each type that were measured by the particle content fit. The distributions are shown on logarithmic scale.

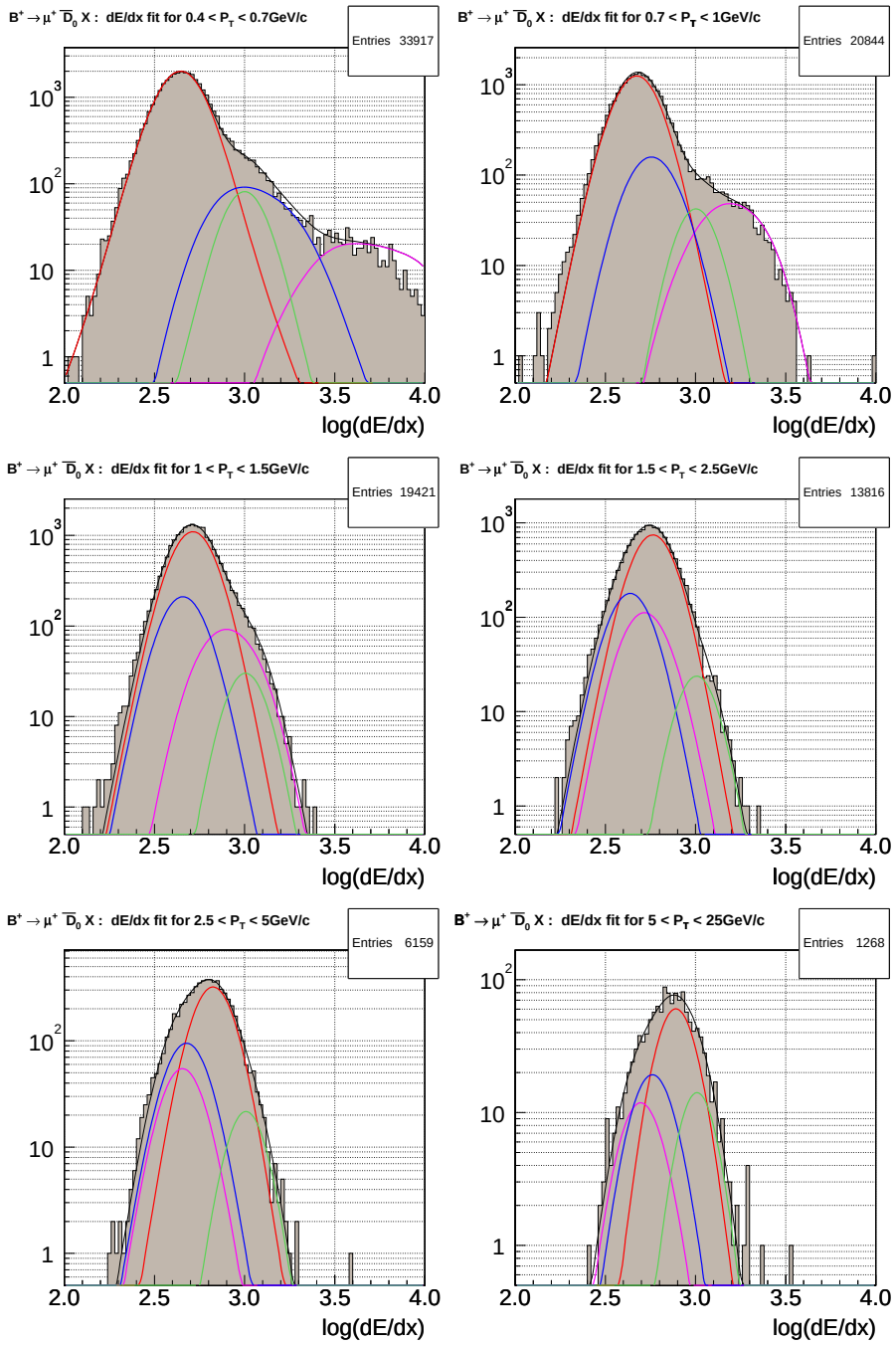


Figure 7.4: Illustration of the likelihood fits for the particle fractions in the lD^0 channels. Shaded histogram is a distribution of the logarithm of the measured energy loss of the particles in the COT detector. The curves correspond to the same distribution expected to be produced by the numbers of particles of each type that were measured by the particle content fit.

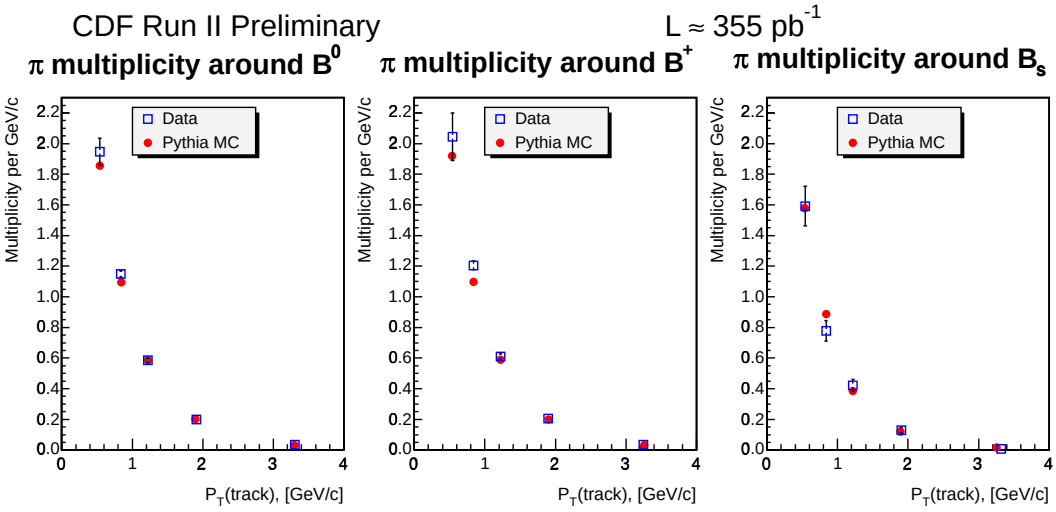


Figure 7.5: The measured pion yield in a sample of the tracks associated with B meson production shown as a function of the transverse momenta of these tracks. The error bars include the systematic uncertainties. The blue squares show the results from the data, and the red points correspond to the Monte-Carlo measurements.

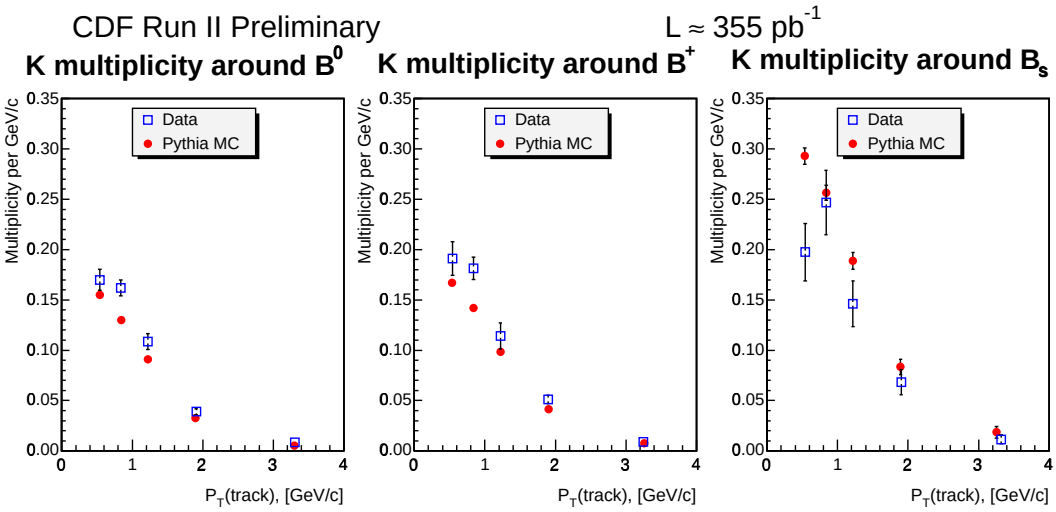


Figure 7.6: The measured kaon yield in a sample of the tracks associated with B meson production shown as a function of the transverse momenta of these tracks. The error bars include the systematic uncertainties. The blue squares show the results from the data, and the red points correspond to the Monte-Carlo measurements.

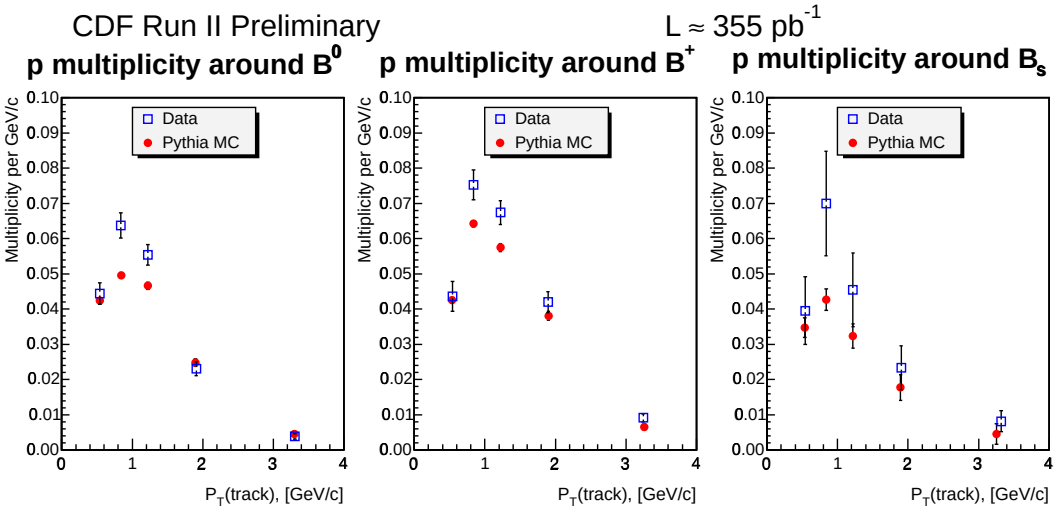


Figure 7.7: The measured proton yield in a sample of the tracks associated with B meson production shown as a function of the transverse momenta of these tracks. The error bars include the systematic uncertainties. The blue squares show the results from the data, and the red points correspond to the Monte-Carlo measurements.

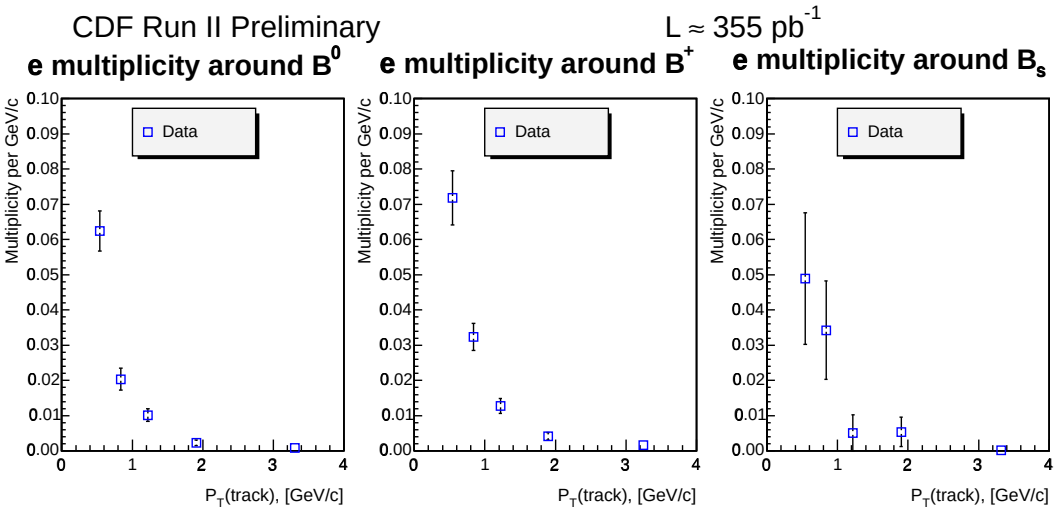


Figure 7.8: The measured electron yield in a sample of the tracks associated with B meson production shown as a function of the transverse momenta of these tracks. The error bars include the systematic uncertainties. The blue squares show the results from the data, and the red points correspond to the Monte-Carlo measurements.

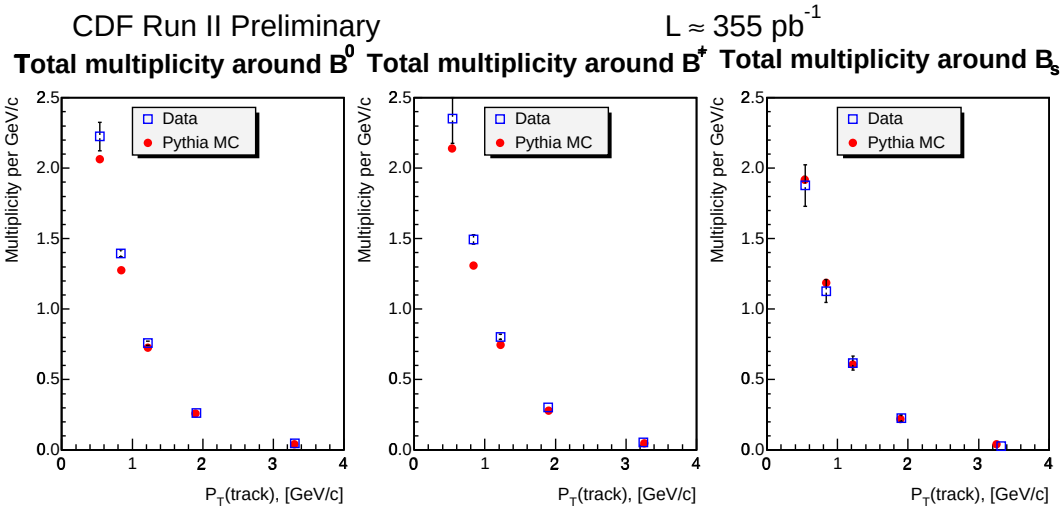


Figure 7.9: The measured combined particle yield in a sample of the tracks associated with B meson production shown as a function of the transverse momenta of these tracks. The error bars include the systematic uncertainties. The blue squares show the results from the data, and the red points correspond to the Monte-Carlo measurements.

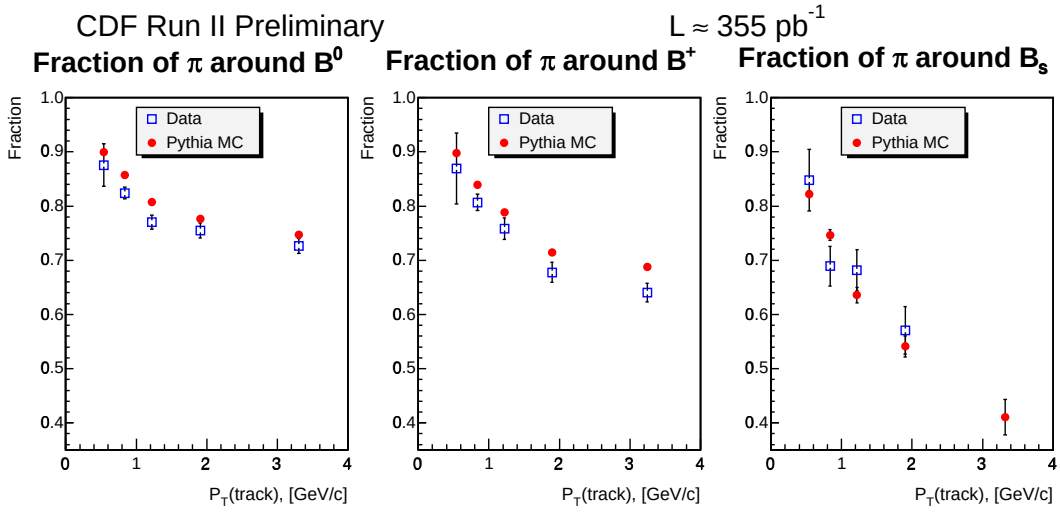


Figure 7.10: The measured fraction of the pions in a sample of the tracks associated with B meson production shown as a function of the transverse momenta of these tracks. The error bars include the systematic uncertainties. The blue squares show the results from the data, and the red points correspond to the Monte-Carlo measurements.

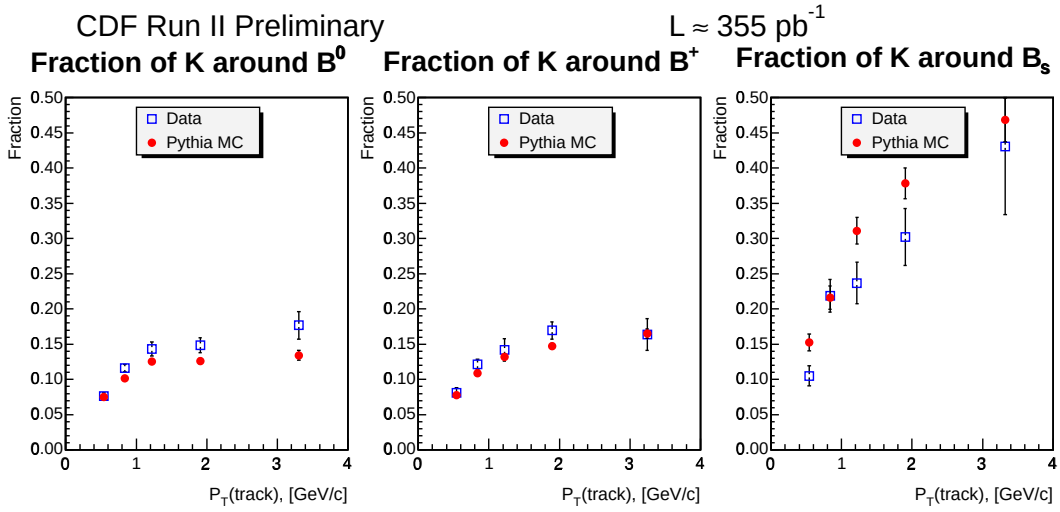


Figure 7.11: The measured fraction of the kaons in a sample of the tracks associated with B meson production shown as a function of the transverse momenta of these tracks. The error bars include the systematic uncertainties. The blue squares show the results from the data, and the red points correspond to the Monte-Carlo measurements.

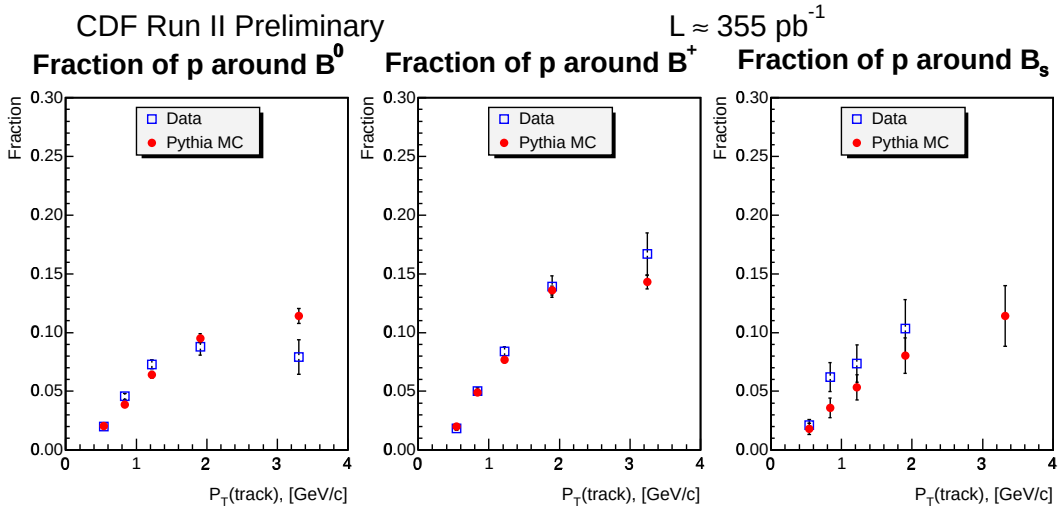


Figure 7.12: The measured fraction of the protons in a sample of the tracks associated with B meson production shown as a function of the transverse momenta of these tracks. The error bars include the systematic uncertainties. The blue squares show the results from the data, and the red points correspond to the Monte-Carlo measurements.

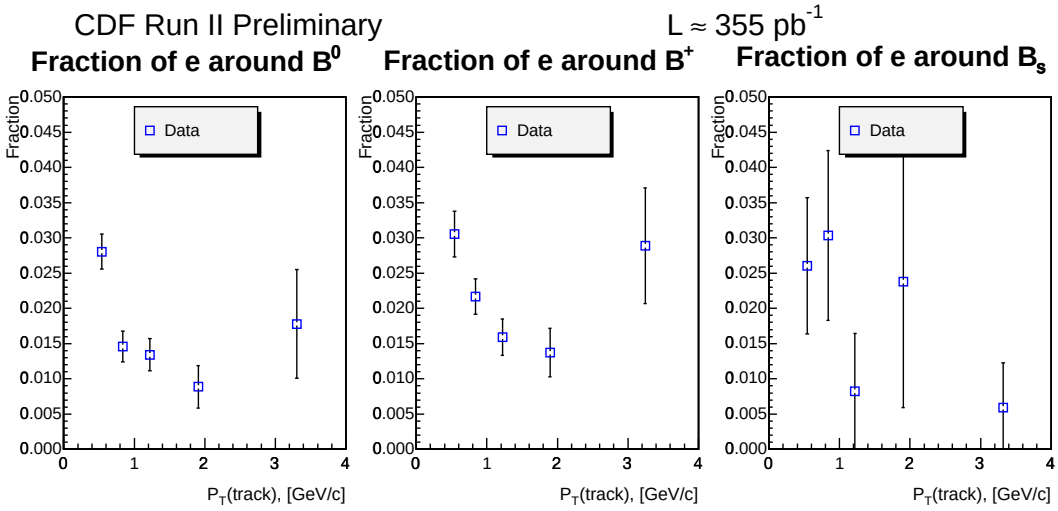


Figure 7.13: The measured fraction of the electrons in a sample of the tracks associated with B meson production shown as a function of the transverse momenta of these tracks. The error bars include the systematic uncertainties. The blue squares show the results from the data, and the red points correspond to the Monte-Carlo measurements.

Particle yields around B mesons, data. Value $\pm$ (stat.) $\pm$(common syst.) $\pm$ (syst.)					
	p_T	π	K	p	total
B^0	0.4-0.7	$1.948 \pm 0.020 \pm 0.086 \pm 0.0080$	$0.1700 \pm 0.0065 \pm 0.0080 \pm 0.00091$	$0.04438 \pm 0.0023 \pm 0.0019 \pm 2.0e-05$	$2.225 \pm 0.022 \pm 0.097 \pm 0.0092$
B^0	0.70-1.0	$1.150 \pm 0.015 \pm 0.014 \pm 0.0049$	$0.1619 \pm 0.0066 \pm 0.0043 \pm 0.00063$	$0.06370 \pm 0.0035 \pm 0.00078 \pm 0.00031$	$1.396 \pm 0.017 \pm 0.017 \pm 0.0053$
B^0	1.0-1.5	$0.5844 \pm 0.0090 \pm 0.0096 \pm 0.0016$	$0.1087 \pm 0.0047 \pm 0.0065 \pm 0.00041$	$0.05539 \pm 0.0028 \pm 0.00082 \pm 0.00037$	$0.7586 \pm 0.011 \pm 0.0096 \pm 0.0020$
B^0	1.5-2.5	$0.1976 \pm 0.0041 \pm 0.0035 \pm 0.00030$	$0.03888 \pm 0.0027 \pm 0.0014 \pm 0.00065$	$0.02306 \pm 0.0015 \pm 0.0011 \pm 0.00070$	$0.2619 \pm 0.0051 \pm 0.0037 \pm 0.0017$
B^0	2.5-5.0	$0.03539 \pm 0.0012 \pm 0.00058 \pm 1e-4$	$0.008615 \pm 0.0011 \pm 0.00021 \pm 1e-4$	$0.003857 \pm 0.00072 \pm 0.00025 \pm 0.00015$	$0.04873 \pm 0.0018 \pm 0.00081 \pm 0.00023$
B^0	$\sum_{p_T}$	$1.504 \pm 0.010 \pm 0.030 \pm 0.0061$	$0.2144 \pm 0.0053 \pm 0.0063 \pm 0.0016$	$0.09317 \pm 0.0030 \pm 0.0022 \pm 0.0013$	$1.846 \pm 0.012 \pm 0.021 \pm 0.0091$
B^+	0.40-0.70	$2.045 \pm 0.021 \pm 0.15 \pm 0.0088$	$0.1912 \pm 0.0074 \pm 0.015 \pm 0.0010$	$0.04359 \pm 0.0026 \pm 0.0033 \pm 2.5e-05$	$2.351 \pm 0.023 \pm 0.17 \pm 0.014$
B^+	0.70-1.0	$1.205 \pm 0.018 \pm 0.021 \pm 0.0057$	$0.1813 \pm 0.0077 \pm 0.0080 \pm 0.00074$	$0.07525 \pm 0.0041 \pm 0.0013 \pm 0.00037$	$1.494 \pm 0.020 \pm 0.026 \pm 0.0069$
B^+	1.0-1.5	$0.6093 \pm 0.010 \pm 0.016 \pm 0.0019$	$0.1142 \pm 0.0054 \pm 0.012 \pm 0.00049$	$0.06740 \pm 0.0031 \pm 0.0013 \pm 0.00043$	$0.8037 \pm 0.012 \pm 0.014 \pm 0.0022$
B^+	1.5-2.5	$0.2043 \pm 0.0048 \pm 0.0054 \pm 0.00037$	$0.05108 \pm 0.0032 \pm 0.0024 \pm 0.00080$	$0.04197 \pm 0.0019 \pm 0.0021 \pm 0.00081$	$0.3015 \pm 0.0062 \pm 0.0056 \pm 0.0018$
B^+	2.5-5.0	$0.03502 \pm 0.0014 \pm 0.00079 \pm 0.00012$	$0.008961 \pm 0.0014 \pm 0.00043 \pm 0.00012$	$0.009122 \pm 0.00092 \pm 0.00060 \pm 0.00017$	$0.05469 \pm 0.0022 \pm 0.0012 \pm 0.00025$
B^+	$\sum_{p_T}$	$1.576 \pm 0.011 \pm 0.063 \pm 0.0071$	$0.2424 \pm 0.0063 \pm 0.011 \pm 0.0019$	$0.1339 \pm 0.0037 \pm 0.0044 \pm 0.0015$	$1.997 \pm 0.014 \pm 0.042 \pm 0.012$
B_s	0.40-0.70	$1.592 \pm 0.074 \pm 0.10 \pm 0.023$	$0.1975 \pm 0.026 \pm 0.013 \pm 0.00098$	$0.03957 \pm 0.0093 \pm 0.0025 \pm 0.00025$	$1.878 \pm 0.081 \pm 0.12 \pm 0.022$
B_s	0.70-1.0	$0.7774 \pm 0.055 \pm 0.030 \pm 0.023$	$0.2467 \pm 0.030 \pm 0.010 \pm 0.0043$	$0.06993 \pm 0.015 \pm 0.0025 \pm 3.0e-05$	$1.128 \pm 0.066 \pm 0.042 \pm 0.017$
B_s	1.0-1.5	$0.4207 \pm 0.034 \pm 0.018 \pm 0.010$	$0.1462 \pm 0.021 \pm 0.0088 \pm 0.0020$	$0.04548 \pm 0.010 \pm 0.0020 \pm 0.00092$	$0.6174 \pm 0.042 \pm 0.025 \pm 0.0084$
B_s	1.5-2.5	$0.1288 \pm 0.016 \pm 0.0059 \pm 0.0041$	$0.06824 \pm 0.012 \pm 0.0031 \pm 0.0014$	$0.02333 \pm 0.0060 \pm 0.0015 \pm 0.00053$	$0.2258 \pm 0.021 \pm 0.0096 \pm 0.0029$
B_s	2.5-5.0	$0.006668 \pm 0.0038 \pm 0.00053 \pm 0.0016$	$0.01135 \pm 0.0044 \pm 0.00055 \pm 0.00015$	$0.008177 \pm 0.0029 \pm 0.00043 \pm 9.2e-05$	$0.02635 \pm 0.0065 \pm 0.0015 \pm 0.0013$
B_s	$\sum_{p_T}$	$1.066 \pm 0.037 \pm 0.020 \pm 0.027$	$0.3029 \pm 0.023 \pm 0.0088 \pm 0.0043$	$0.09936 \pm 0.012 \pm 0.0026 \pm 0.00083$	$1.502 \pm 0.046 \pm 0.026 \pm 0.022$

Table 7.6: The yields of charged particles seen around B mesons in data. The yield is defined as the number of tracks in a cone $\Delta R < 0.7$ around the lD momentum normalized to the bin width. Total yield is is not divided by the width of the p_T region.

Particle fractions around B mesons, data Value $\pm$ (stat.) $\pm$(common syst.) $\pm$ (syst.)					
	p_T,[GeV/c]	π,%	K,%	p,%	e,%
B^0	0.40-0.70	$87.6 \pm 0.11 \pm 3.9 \pm 0.36$	$7.64 \pm 0.27 \pm 0.36 \pm 0.041$	$1.99 \pm 0.10 \pm 0.087 \pm 0.00090$	$2.81 \pm 0.21 \pm 0.13 \pm 0.018$
B^0	0.70-1.0	$82.4 \pm 0.20 \pm 1.0 \pm 0.35$	$11.6 \pm 0.42 \pm 0.31 \pm 0.045$	$4.56 \pm 0.24 \pm 0.056 \pm 0.022$	$1.46 \pm 0.21 \pm 0.040 \pm 0.025$
B^0	1.0-1.5	$77.0 \pm 0.27 \pm 1.3 \pm 0.21$	$14.3 \pm 0.53 \pm 0.86 \pm 0.054$	$7.30 \pm 0.34 \pm 0.11 \pm 0.049$	$1.34 \pm 0.23 \pm 0.035 \pm 0.012$
B^0	1.5-2.5	$75.5 \pm 0.38 \pm 1.3 \pm 0.11$	$14.8 \pm 0.87 \pm 0.53 \pm 0.25$	$8.80 \pm 0.52 \pm 0.44 \pm 0.27$	$0.886 \pm 0.29 \pm 0.033 \pm 0.040$
B^0	2.5-5.0	$72.6 \pm 0.69 \pm 1.2 \pm 0.17$	$17.7 \pm 1.9 \pm 0.44 \pm 0.19$	$7.92 \pm 1.4 \pm 0.52 \pm 0.31$	$1.78 \pm 0.76 \pm 0.099 \pm 0.016$
B^0	$\sum_{p_T}$	$81.4 \pm 0.10 \pm 1.6 \pm 0.33$	$11.6 \pm 0.25 \pm 0.34 \pm 0.085$	$5.05 \pm 0.15 \pm 0.12 \pm 0.072$	$1.90 \pm 0.12 \pm 0.047 \pm 0.019$
B^+	0.40-0.70	$87.0 \pm 0.12 \pm 6.5 \pm 0.38$	$8.13 \pm 0.29 \pm 0.63 \pm 0.044$	$1.85 \pm 0.11 \pm 0.14 \pm 0.0010$	$3.05 \pm 0.22 \pm 0.23 \pm 0.020$
B^+	0.70-1.0	$80.7 \pm 0.23 \pm 1.4 \pm 0.38$	$12.1 \pm 0.45 \pm 0.54 \pm 0.049$	$5.04 \pm 0.26 \pm 0.084 \pm 0.025$	$2.17 \pm 0.24 \pm 0.073 \pm 0.027$
B^+	1.0-1.5	$75.8 \pm 0.31 \pm 1.9 \pm 0.23$	$14.2 \pm 0.58 \pm 1.5 \pm 0.061$	$8.39 \pm 0.35 \pm 0.17 \pm 0.054$	$1.59 \pm 0.25 \pm 0.059 \pm 0.014$
B^+	1.5-2.5	$67.8 \pm 0.51 \pm 1.8 \pm 0.12$	$16.9 \pm 0.89 \pm 0.81 \pm 0.27$	$13.9 \pm 0.53 \pm 0.71 \pm 0.27$	$1.37 \pm 0.33 \pm 0.064 \pm 0.050$
B^+	2.5-5.0	$64.0 \pm 0.94 \pm 1.4 \pm 0.21$	$16.4 \pm 2.1 \pm 0.79 \pm 0.21$	$16.7 \pm 1.4 \pm 1.1 \pm 0.32$	$2.89 \pm 0.81 \pm 0.14 \pm 0.040$
B^+	$\sum_{p_T}$	$78.9 \pm 0.12 \pm 3.1 \pm 0.35$	$12.1 \pm 0.28 \pm 0.57 \pm 0.093$	$6.70 \pm 0.17 \pm 0.22 \pm 0.077$	$2.28 \pm 0.13 \pm 0.079 \pm 0.021$
B_s	0.40-0.70	$84.8 \pm 0.62 \pm 5.5 \pm 1.2$	$10.5 \pm 1.2 \pm 0.69 \pm 0.052$	$2.11 \pm 0.48 \pm 0.13 \pm 0.013$	$2.60 \pm 0.95 \pm 0.19 \pm 0.055$
B_s	0.70-1.0	$68.9 \pm 1.5 \pm 2.6 \pm 2.0$	$21.9 \pm 2.1 \pm 0.91 \pm 0.38$	$6.20 \pm 1.2 \pm 0.22 \pm 0.0027$	$3.04 \pm 1.2 \pm 0.12 \pm 0.039$
B_s	1.0-1.5	$68.1 \pm 1.8 \pm 2.9 \pm 1.6$	$23.7 \pm 2.6 \pm 1.4 \pm 0.33$	$7.37 \pm 1.5 \pm 0.33 \pm 0.15$	$0.823 \pm 0.82 \pm 0.054 \pm 0.059$
B_s	1.5-2.5	$57.1 \pm 3.0 \pm 2.6 \pm 1.8$	$30.2 \pm 3.8 \pm 1.4 \pm 0.60$	$10.3 \pm 2.4 \pm 0.65 \pm 0.23$	$2.38 \pm 1.8 \pm 0.11 \pm 0.061$
B_s	2.5-5.0	$25.3 \pm 11. \pm 2.0 \pm 6.2$	$43.1 \pm 9.4 \pm 2.1 \pm 0.59$	$31.0 \pm 7.7 \pm 1.7 \pm 0.35$	$0.589 \pm 0.59 \pm 0.091 \pm 0.24$
B_s	$\sum_{p_T}$	$71.0 \pm 0.73 \pm 1.3 \pm 1.8$	$20.2 \pm 1.2 \pm 0.59 \pm 0.29$	$6.62 \pm 0.74 \pm 0.17 \pm 0.055$	$2.21 \pm 0.62 \pm 0.067 \pm 0.027$

Table 7.7: The fractions (in percent) of charged particles seen around B mesons in data.

Particle content around B mesons in Monte-Carlo.					
	p_T,[GeV/c]	π,%	K,%	p,%	M_{TRK}
B^0	0.40-0.70	89.984 ± 0.082	7.51 ± 0.22	2.06 ± 0.12	2.064 ± 0.018
B^0	0.70-1.0	85.74 ± 0.14	10.2 ± 0.31	3.88 ± 0.21	1.28 ± 0.014
B^0	1.0-1.5	80.76 ± 0.19	12.5 ± 0.35	6.41 ± 0.27	0.727 ± 0.0081
B^0	1.5-2.5	77.67 ± 0.26	12.6 ± 0.41	9.53 ± 0.37	0.260 ± 0.0034
B^0	2.5-5.0	74.68 ± 0.47	13.4 ± 0.68	11.4 ± 0.64	0.0402 ± 0.00086
B^0	$\sum_{p_T}$	84.357 ± 0.074	10.3 ± 0.15	5.05 ± 0.11	1.726 ± 0.0089
B^+	0.40-0.70	89.742 ± 0.077	7.80 ± 0.21	1.99 ± 0.11	2.139 ± 0.017
B^+	0.70-1.0	83.90 ± 0.15	10.8 ± 0.30	4.91 ± 0.21	1.31 ± 0.013
B^+	1.0-1.5	78.87 ± 0.20	13.2 ± 0.33	7.69 ± 0.27	0.747 ± 0.0078
B^+	1.5-2.5	71.47 ± 0.29	14.7 ± 0.39	13.6 ± 0.38	0.280 ± 0.0034
B^+	2.5-5.0	68.80 ± 0.49	16.5 ± 0.64	14.3 ± 0.61	0.0461 ± 0.00087
B^+	$\sum_{p_T}$	82.039 ± 0.077	11.2 ± 0.14	6.40 ± 0.11	1.803 ± 0.0086
B_s	0.40-0.70	82.22 ± 0.58	15.3 ± 1.2	1.81 ± 0.47	1.92 ± 0.069
B_s	0.70-1.0	74.6 ± 1.00	21.6 ± 1.7	3.60 ± 0.83	1.19 ± 0.054
B_s	1.0-1.5	63.6 ± 1.4	31.1 ± 1.9	5.32 ± 1.1	0.607 ± 0.030
B_s	1.5-2.5	54.1 ± 1.9	37.8 ± 2.2	8.05 ± 1.5	0.220 ± 0.013
B_s	2.5-5.0	41.0 ± 3.3	46.8 ± 3.1	11.4 ± 2.6	0.0395 ± 0.0034
B_s	$\sum_{p_T}$	70.24 ± 0.54	25.0 ± 0.82	4.40 ± 0.44	1.55 ± 0.034

Table 7.8: The fractions (in percent) of the charged particles seen around B mesons in the Monte-Carlo, as well as the total yield of all tracks.

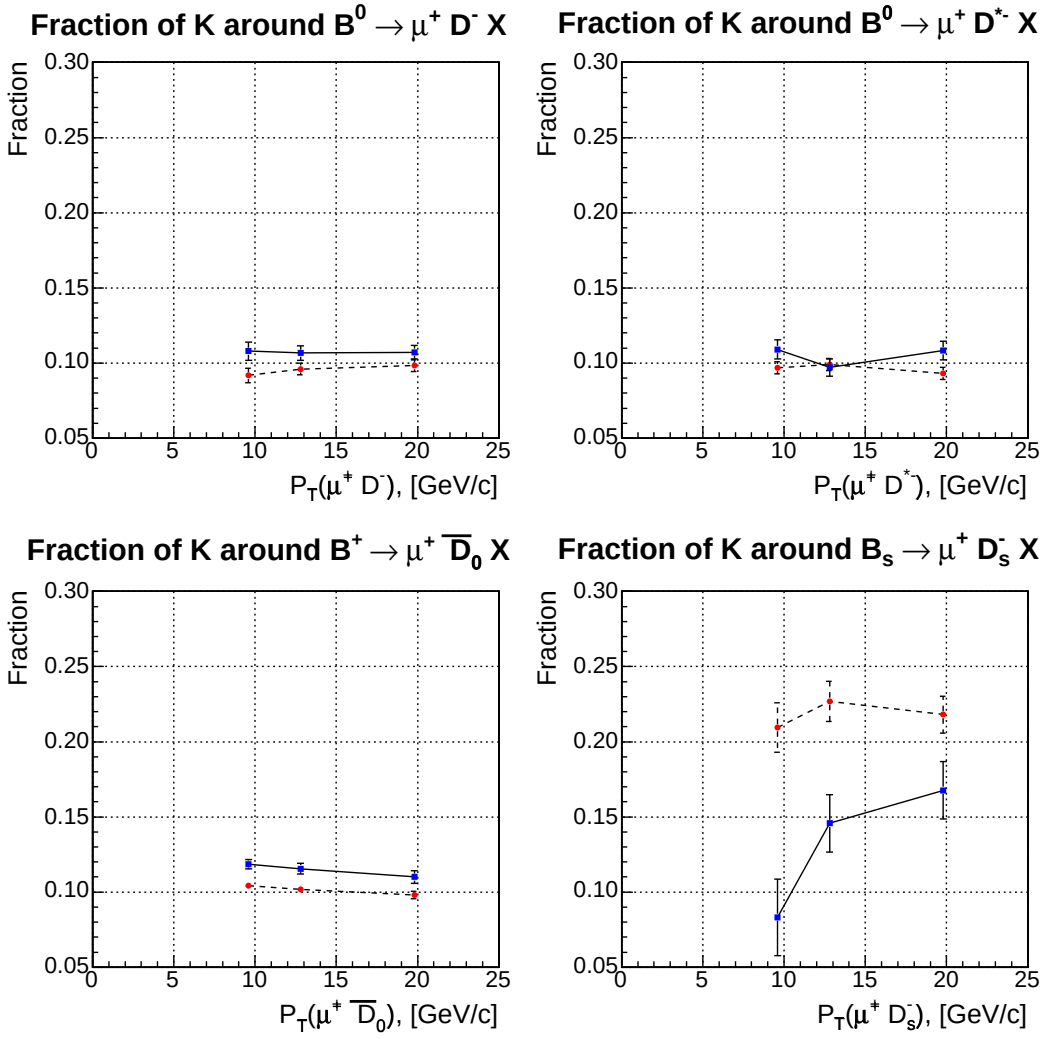


Figure 7.14: The measured kaon fraction as a function of $p_T(lD)$. The blue dots and solid line show the results for data, and red dots and dashed line show the results for Monte-Carlo.

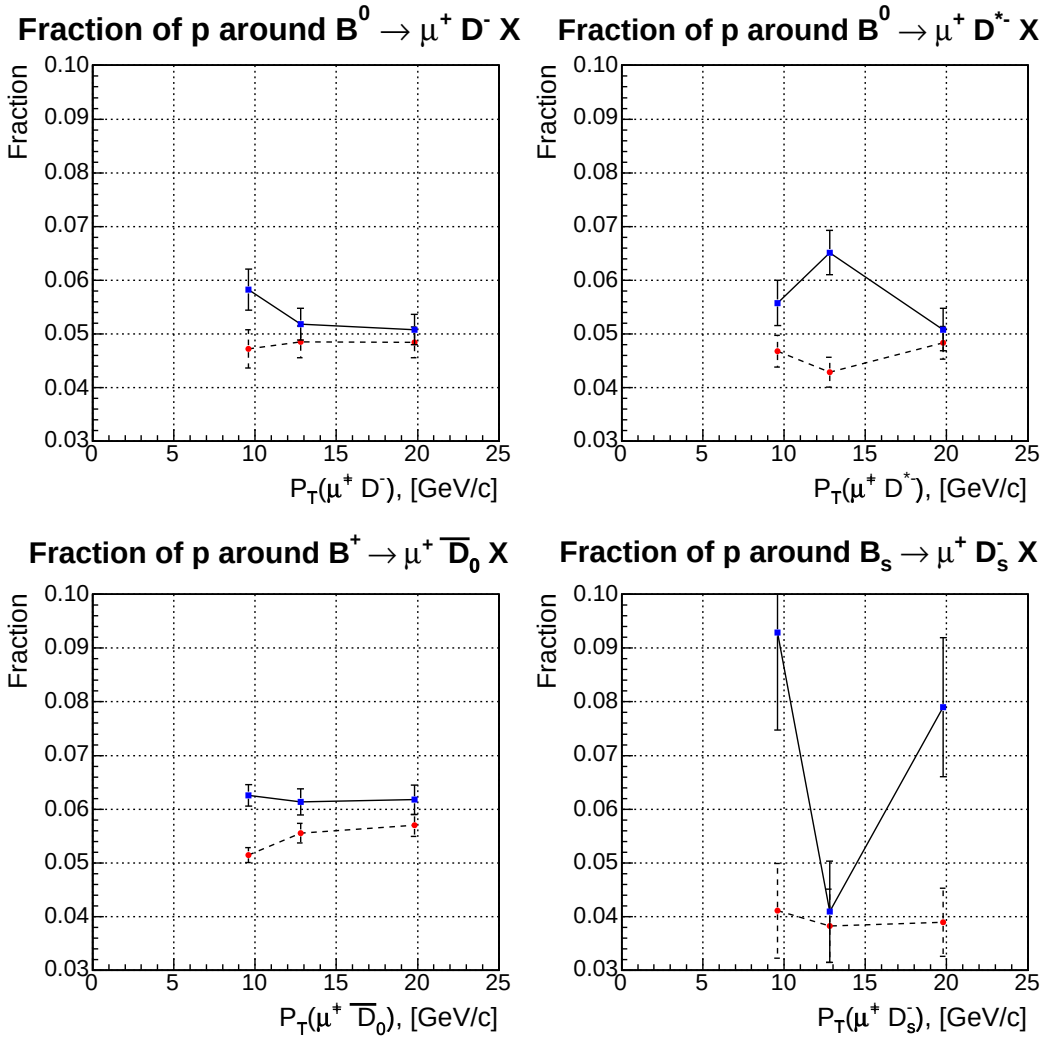


Figure 7.15: The measured proton fraction as a function of $p_T(lD)$. The blue dots and solid line show the results for data, and red dots and dashed line show the results for Monte-Carlo.

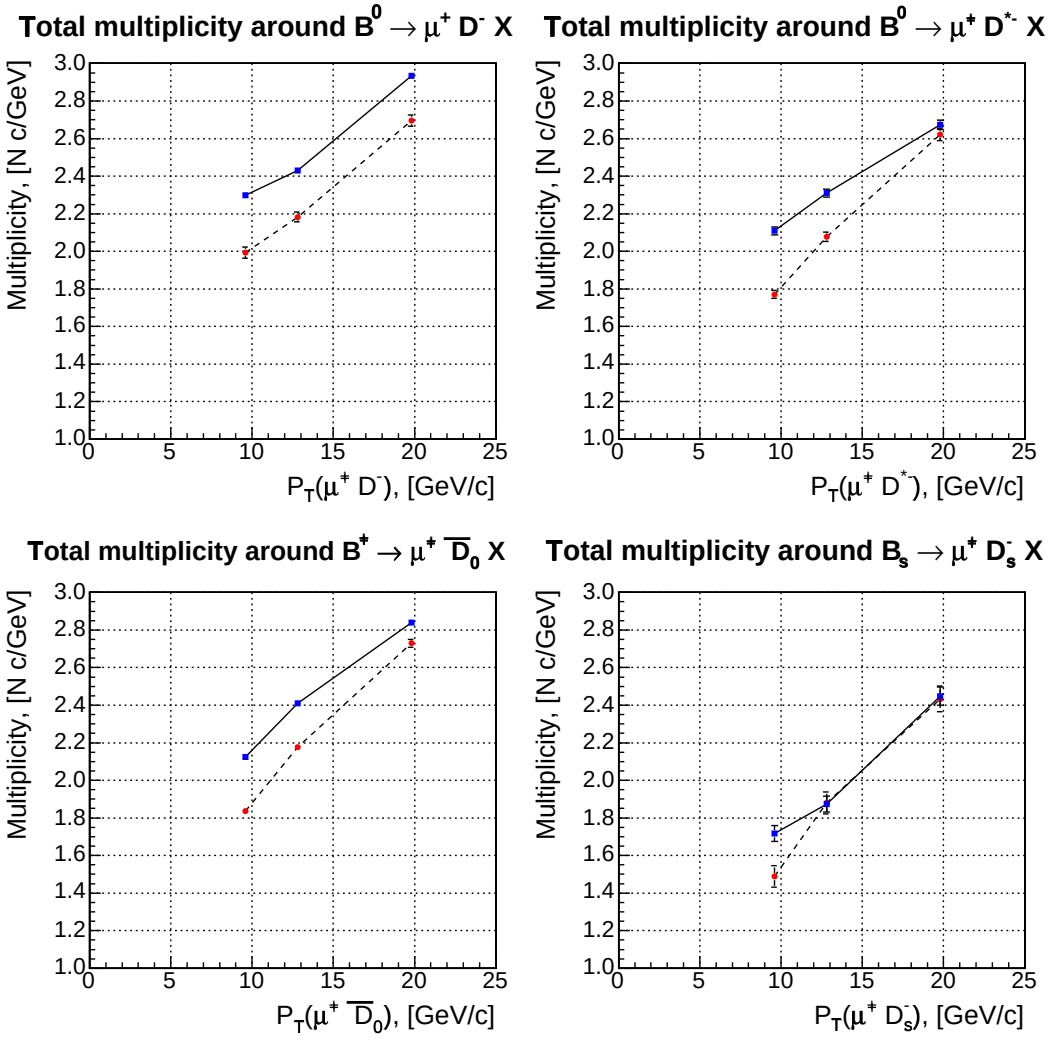


Figure 7.16: Combined track yield as a function of $p_T(lD)$. The blue dots and solid line show the results for data, and red dots and dashed line show the results for Monte-Carlo.

7.4 Particle production in different ΔR regions

Another study we performed concerns the change in the particle content of the tracks associated with B formation with the cone size ΔR. We divided the sample of tracks in three categories based on the distance $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ to the lD momentum direction. The regions chosen were $[0; 0.3], [0.3; 0.7], [0.7; 1.0]$.

Tables 7.9 and 7.10 summarize the results of this study. As expected from the most fragmentation models, the leading particle associated with B meson formation has momentum direction strongly correlated with that of the B meson. This effect is especially strong in the case of $\bar{B}_s^0$ meson and the associated kaon. The results in

Total fractions around B mesons, tracks with $\Delta R < 0.3$, data Value $\pm$ (stat.) $\pm$(common syst.) $\pm$ (syst.)			
	π,%	K,%	p,%
B^0	$81.0 \pm 0.20 \pm 1.4 \pm 0.21$	$9.33 \pm 0.47 \pm 0.26 \pm 0.12$	$4.83 \pm 0.30 \pm 0.15 \pm 0.11$
B^+	$77.8 \pm 0.25 \pm 3.0 \pm 0.24$	$11.6 \pm 0.51 \pm 0.42 \pm 0.14$	$7.58 \pm 0.34 \pm 0.31 \pm 0.12$
B_s	$58.2 \pm 1.9 \pm 0.90 \pm 2.4$	$21.5 \pm 2.2 \pm 0.34 \pm 0.47$	$7.26 \pm 1.6 \pm 0.22 \pm 0.039$
Total fractions around B mesons, tracks with $0.3 < \Delta R < 0.7$, data			
	π,%	K,%	p,%
B^0	$81.0 \pm 0.14 \pm 1.4 \pm 0.42$	$10.9 \pm 0.27 \pm 0.24 \pm 0.057$	$5.09 \pm 0.18 \pm 0.077 \pm 0.056$
B^+	$81.5 \pm 0.14 \pm 2.8 \pm 0.45$	$10.8 \pm 0.30 \pm 0.41 \pm 0.063$	$6.46 \pm 0.20 \pm 0.14 \pm 0.060$
B_s	$69.4 \pm 0.94 \pm 0.77 \pm 1.8$	$15.8 \pm 1.3 \pm 0.31 \pm 0.13$	$6.05 \pm 0.80 \pm 0.10 \pm 0.070$
Total fractions around B mesons, tracks with $0.7 < \Delta R < 1.0$, data			
	π,%	K,%	p,%
B^0	$79.4 \pm 0.17 \pm 1.3 \pm 0.39$	$11.4 \pm 0.31 \pm 0.20 \pm 0.100$	$5.98 \pm 0.21 \pm 0.072 \pm 0.029$
B^+	$79.9 \pm 0.17 \pm 2.8 \pm 0.41$	$12.6 \pm 0.32 \pm 0.36 \pm 0.11$	$6.07 \pm 0.21 \pm 0.13 \pm 0.030$
B_s	$76.2 \pm 0.76 \pm 0.73 \pm 1.4$	$10.7 \pm 1.2 \pm 0.29 \pm 0.25$	$4.89 \pm 0.80 \pm 0.072 \pm 0.15$

Table 7.9: The measured total particle fractions in data in the samples of tracks found in different ΔR regions.

the data also support the conjecture that the b fragmentation tracks are contained in the cone $\Delta R \approx 0.7$ around the lD meson direction.

Total fractions around B mesons, tracks with $\Delta R < 0.3$, Monte-Carlo Value $\pm$ (stat.) $\pm$(common syst.) $\pm$ (syst.)			
	π,%	K,%	p,%
B^0	84.94 ± 0.13	10.1 ± 0.27	4.57 ± 0.20
B^+	81.59 ± 0.14	11.8 ± 0.26	6.25 ± 0.20
B_s	61.8 ± 1.2	34.1 ± 1.5	3.76 ± 0.75
Total fractions around B mesons, tracks with $0.3 < \Delta R < 0.7$, Monte-Carlo			
	π,%	K,%	p,%
B^0	84.12 ± 0.089	10.3 ± 0.18	5.24 ± 0.13
B^+	82.23 ± 0.091	11.0 ± 0.17	6.46 ± 0.13
B_s	73.73 ± 0.58	21.3 ± 0.94	4.67 ± 0.53
Total fractions around B mesons, tracks with $0.7 < \Delta R < 1.0$, Monte-Carlo			
	π,%	K,%	p,%
B^0	82.44 ± 0.10	11.0 ± 0.19	6.18 ± 0.15
B^+	81.45 ± 0.10	11.4 ± 0.18	6.69 ± 0.15
B_s	77.18 ± 0.53	16.6 ± 0.89	5.88 ± 0.60

Table 7.10: The measured total particle fractions in Monte-Carlo in the samples of tracks found in different ΔR regions.

The Monte-Carlo reproduces the general trends seen in the data but significantly overestimates the relative kaon production in the $\bar{B}_s^0$ mode. It also appears that the cone containing the fragmentation tracks is wider in Monte-Carlo.

7.5 Sources of uncertainties

The treatment of most of the uncertainties arising in this analysis is discussed in the Chapters explaining the technical aspects and procedures of our measurement. Here we list these uncertainties to remind the reader what they are, and how they are used.

The statistical uncertainties quoted in this note are the errors returned by the likelihood fit (and they are obtained assuming Poisson statistics on the number of

tracks of each type) and the errors on the measured detector acceptance.

The systematic uncertainty that is due to the imperfect knowledge of the resolution function was estimated by allowing some of the resolution function parameters to change by one standard deviation. The analysis was then performed with these new resolution function parameters and the change in the fitted yields and fractions was used to estimate the systematic uncertainty. The parameters we allow to change are the width of the TOF resolution, the size of the negative tail of the TOF resolution function, and the correction on the dE/dx value. These parameters all have different importance in different p_T bins and should be weakly correlated. Table 7.11 shows the change in fitted fractions after the resolution function parameters were changed. The numbers are signed to show that the changes were strongly correlated across the different lD modes. We therefore can assume that for comparing different B modes between themselves this systematic uncertainty may be negligibly small. For this reason this systematic uncertainty is quoted separately from the rest of the systematic errors.

Additional detector-related systematic uncertainties come from our estimation of the COT efficiency and the effects of the kaon decays in flight. These corrections affect all the B decay modes in a similar way and therefore the associated uncertainties are added to the common systematic error.

Another source of systematic error comes from the uncertainty on the sample composition parameters. They originate from the experimental errors on the

Systematic errors on fitted fractions.					
Type	p_T	π,%	K,%	p,%	e, %
l^+D^-	0.4-0.7	0.0373	0.0125	0.00022	-0.0501
l^+D^-	0.7-1	-0.155	0.195	0.0022	-0.043
l^+D^-	1-1.5	-0.946	0.918	0.0622	-0.034
l^+D^-	1.5-2.5	-0.887	0.468	0.455	-0.0359
l^+D^-	2.5-5	-0.207	-0.31	0.605	-0.0879
l^+D^-,$\sum$	0.4-5	-0.356	0.281	0.121	-0.0456
l^+D^{*-}	0.4-0.7	0.0474	0.0115	-1e-05	-0.059
l^+D^{*-}	0.7-1	-0.186	0.221	0.00202	-0.0365
l^+D^{*-}	1-1.5	-0.872	0.847	0.0594	-0.0338
l^+D^{*-}	1.5-2.5	-0.863	0.438	0.457	-0.0319
l^+D^{*-}	2.5-5	-0.19	-0.174	0.493	-0.129
l^+D^{*-},$\sum$	0.4-5	-0.344	0.283	0.11	-0.0491
$l^+\bar{D}_0$	0.4-0.7	0.038	0.0107	0.00061	-0.0494
$l^+\bar{D}_0$	0.7-1	-0.196	0.237	0.00349	-0.0444
$l^+\bar{D}_0$	1-1.5	-0.943	0.919	0.0605	-0.0361
$l^+\bar{D}_0$	1.5-2.5	-0.826	0.412	0.454	-0.0408
$l^+\bar{D}_0$	2.5-5	-0.255	-0.398	0.746	-0.0924
$l^+\bar{D}_0$,$\sum$	0.4-5	-0.368	0.282	0.134	-0.0471
$l^+D_s^-$	0.4-0.7	0.0618	0.0109	0.00128	-0.074
$l^+D_s^-$	0.7-1	-0.177	0.225	0.00209	-0.0502
$l^+D_s^-$	1-1.5	-0.932	0.837	0.129	-0.0341
$l^+D_s^-$	1.5-2.5	-0.758	0.397	0.411	-0.0503
$l^+D_s^-$	2.5-5	-0.641	-0.109	0.693	0.0567
$l^+D_s^-$,$\sum$	0.4-5	-0.355	0.277	0.127	-0.0498

Table 7.11: The change in the fitted particle fractions after some of the TOF/dEdX resolution function parameters were allowed to change by one standard deviation. These numbers were used to estimate the systematic uncertainty due to the imperfect knowledge of the resolution functions.

branching ratios of the decays of B and D mesons (mostly excited states). The way we treat the sample composition parameters is described in detail in Section 5. The errors resulting from the presence of the sample composition uncertainties are included in the yields around the B signals.

Conclusion and Outlook

We studied the production of the charged particles associated with B meson formation by examining the particle content around $B^-, \bar{B}^0$ and $\bar{B}^0_s$ mesons decaying semileptonically. We use an unbinned likelihood fit to measure the particle fractions and yields in the samples of tracks selected around the partially reconstructed B signals. Calibrated TOF and dE/dx information is used for the particle identification.

From the quark model a significant difference in kaon production in association with $\bar{B}^0_s$ meson formation is expected when compared to the $\bar{B}^0, B^-$ mesons. Our measurements strongly confirm this expectation and present the first direct observation of this effect at a hadron collider. The excess of kaons around $\bar{B}^0_s$ mesons versus $B_{d,u}$ mesons can only be attributed to the b quark fragmentation and the difference in kaon production is therefore directly sensitive to the underlying fragmentation process.

We compare the results obtained in CDF with the prediction of the `Pythia` [46] generator using the Peterson fragmentation function and observe small (of order

2σ significance) discrepancy in the production of various particle species associated with B meson formation. In particular, the observed kaon production around $\bar{B}_s^0$ mesons is somewhat lower than the `Pythia` prediction. The Monte-Carlo correctly reproduces all the general trends seen in our data.

The results of these analysis can be used to improve the performance of the simulation generators, to help develop the B flavor tagging techniques at CDF and, by these means, to contribute to the first measurement of the $\bar{B}_s^0$ mixing frequency.

Appendix A

TOF and dE/dx performance

The precision with which TOF detector can measure the particle flight time has to be taken into account when measuring the particle content, as explained in Section 6.1.

It should be noted that this is not the first CDF analysis to develop a particle identification technique using TOF and dE/dx information. One such technique is described in [38]. The reason we use our own parameterization of TOF resolution is primarily because the parameterization derived in [38] assumes p_T independence of the TOF resolution function. Since in our analysis the measurements are done in ranges of track p_T it is crucial to take out the systematic effects that would appear from treating all the p_T regions uniformly. Also, the main motivation for technique described in [38] is separating pions and kaons on track by track basis, whereas in our analysis the correct statistical separation is important. That means that we

should put more emphasis on simulating the tails in the TOF resolution correctly, which is not a serious concern for track by track separation power.

A description of the TOF detector calibration and reconstruction procedures are given in references [26], [31] and [36]. The front-end electronics is discussed in detail in [17]. We will not therefore dwell on these topics and in this section we describe the treatment of the TOF resolution as relevant for this analysis.

If we were to know the particle type beforehand then we could compare the predicted production time of the particle t_{pred} (TOF measured arrival time minus its time of flight predicted from mass, momentum and track trajectory) with the production time measured by the detector (event t_0). The distribution of $t_{pred} - t0$ would represent the resolution of the TOF detector.

We measure the TOF resolution on a sample of pions selected from D^{*-} decays from the decay chain $B^0 \to l^+ D^{*-} X, D^{*-} \to \pi^- \bar{D^0}, \bar{D^0} \to K^+ \pi^-$. We take both pions in this channel and require $0.4 < p_T < 5.0$ GeV/c and $|M(D^0) - 1.864| < 0.015$ GeV/c, where $M(D^0)$ is the mass of the reconstructed D^0 meson. The signal to background ratio for this selection is 31, and this sample is estimated to contain less than 0.5% of kaons (kaon contamination was measured using imperfect TOF resolution function).

There are two main contributions to the TOF resolution function. First is from the cases when the track was correctly reconstructed and matched in TOF. This is the dominant narrow part of the distribution and we describe it with a double

Gaussian fit. Second is the scenario where the reconstruction might have not been entirely successful due to either additional unmatched hits on the bar, incorrect mass hypothesis on one or more tracks during event t_0 calculation or for a variety of other reasons. We found that this part of the resolution function can be sufficiently well described by a positive and negative exponential tails that are brought smoothly to zero at $t = 0$.

Properly describing the tails is highly important for our analysis. As described in Section 6.1, in lower p_T regions the contribution from kaons in the $t_{pred}(\pi) - t0$ distribution lands on the negative tail of the pion peak. Therefore any error on the size of the negative tail of the TOF resolution function directly translates into an error on the kaon fraction. Since the kaon fraction is low at low p_T (typically a few percent) the resulting relative change can be quite significant. This argument was the main motivation for measuring the resolution function in ranges of p_T, since we did see a noticeable change in the size of the tails with the p_T of the track (negative tail fraction varies within 2 - 6% and positive from 1 to 2%).

Our parameterization of the TOF resolution function was chosen as

$$\begin{aligned} R_{TOF} &= f_{G1} G_1(\Delta t, \sigma_{G1}, \Delta T) \\ &+ f_{G2} G_1(\Delta t, \sigma_{G2}, \Delta T_{G2}) \\ &+ f^+ E^+(\Delta t - \Delta T) + f^- E^-(\Delta t - \Delta T), \end{aligned} \tag{A.1}$$

where $G(x, \sigma, \Delta T)$ is a Gaussian distribution with mean ΔT and width σ and $E(x)$ describes the tails of the distribution. The tail part of the distribution is empirically

parametrized as

$$E(x) = \frac{2\tau(A^2+\tau^2)}{A^2(1+e^{-\pi\tau/A})} \times \begin{cases} e^{-\tau x}\frac{(1-cos(Ax))}{2}, & 0 < x < \frac{\pi}{A} \\ e^{-\tau x}, & x > \frac{\pi}{A} \end{cases} . \qquad (A.2)$$

Note that $\int_0^\infty E dx = 1$. This function is defined for $x > 0$ and behaves like an exponential for $x > \frac{\pi}{A}$ and below that value slowly goes to 0. It is also smooth for all values of x. Figure A.1 shows a graph of this function for a typical set of parameters. The positive exponential tail is then defined as $E^+(x) = E(\tau^+, x), x \geq 0$. Since $E(x)$ is defined only for $x > 0$ the negative tail for $x \leq 0$ is defined as $E^-(x) = E(\tau^-, -x)$.

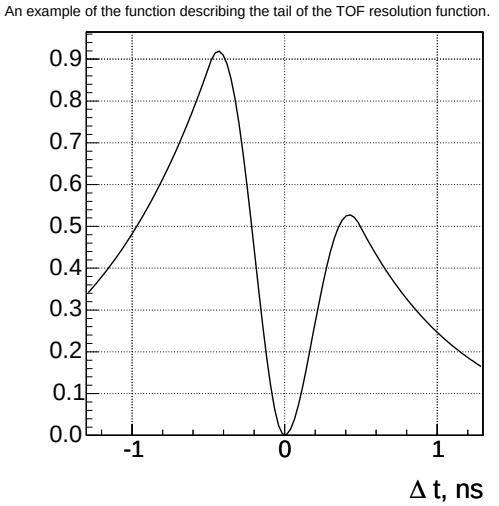


Figure A.1: An example of the function describing the long tails in the TOF resolution function.

The width of the resolution function strongly depends on the z position of the hit in the bar. The most simple parameterization for this dependence is linear:

$$\frac{1}{\sigma^2} = \frac{1}{(\sigma_E + \alpha_E(L/2 - z))^2} + \frac{1}{(\sigma_W + \alpha_W(L/2 + z))^2}, \qquad (A.3)$$

where the subscripts E and W refer to the east and west PMT, σ is the channel resolution for a hit in front of the PMT, α is the parameter describing the dependence of the channel resolution on the hit position and L is the bar length. Both σ and α eventually will be available from the calibration database, however at the time this analysis was performed they were not. Therefore when we fit for the resolution function we also fit for these parameters. Due to the limited size of the calibration sample in this analysis we assume that all detector channels are similar (so that also $\sigma_E = \sigma_W$ and $\alpha_W = \alpha_E$). Both Gaussians in equation A.1 are parametrized in this way so that $\sigma_{G1} = \sigma_{G1}(\sigma_1, \alpha_1)$ and $\sigma_{G2} = \sigma_{G2}(\sigma_2, \alpha_2)$. Finally, parameter ΔT represents the shift of the resolution function from zero and parameter ΔT_{G2} is the shift of the second Gaussian distribution with respect to the first.

The resolution of the TOF detector has also been found to depend on the p_T of the tracks. Since in our analysis we measure the particle fractions in the bins of track p_T, we also measured the TOF resolution in several regions of p_T. The graphical representations of the fits for the resolution parameters are shown in Figures A.3 and A.4. Table A.1 lists the values of the fitted parameters and the associated uncertainties. These values are given here for reference only. It should be mentioned that the parameterization used here is purely empirical and the parameters do not represent physics processes taking place inside the detector. Also, the values of some of the parameters are strongly correlated which is not reflected in the Table A.1.

The particle energy loss when passing through the detector material dE/dx is

measured by the Central-Outer Tracker (COT) detector. The dE/dx measurement was calibrated (see [44] and [23]) so that the resolution function is a Gaussian distribution with a mean and width that vary with particle species.

However we have found that there is a bias in the predicted mean value of the dE/dx measurement. We have measured the correction in the sample of pions selected from D^{*-} decays from the decay chain $B^0 \to l^+ D^{*-} X$, $D^{*-} \to \pi^- \bar{D}^0$, $\bar{D}^0 \to K^+\pi^-$. We found that the pull of $Z = log(dE/dx) - log(dE/dx_{expected})$ is shifted in the positive direction by 0.095 ± 0.009 of the predicted width for the pion hypothesis. This is illustrated in Figure A.2. We apply this correction to our data sample in this analysis.

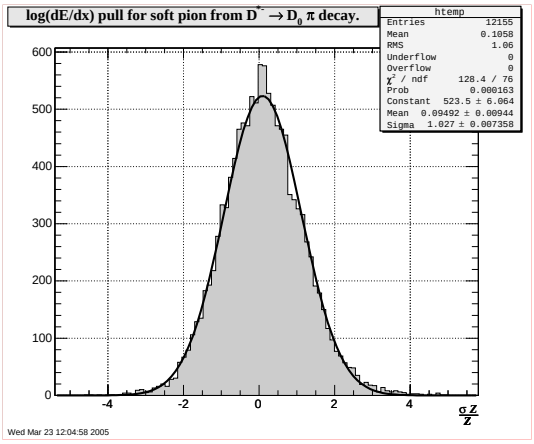


Figure A.2: Pull distribution of $Z = log(dE/dx)$ for the soft pion from $D^{*-} \to \pi^- \bar{D}^0$ decay from μSVT and eSVT samples. The distribution is fit with a Gaussian distribution. We measure the mean to be shifted by 0.095 ± 0.009. The pull is consistent with unity.

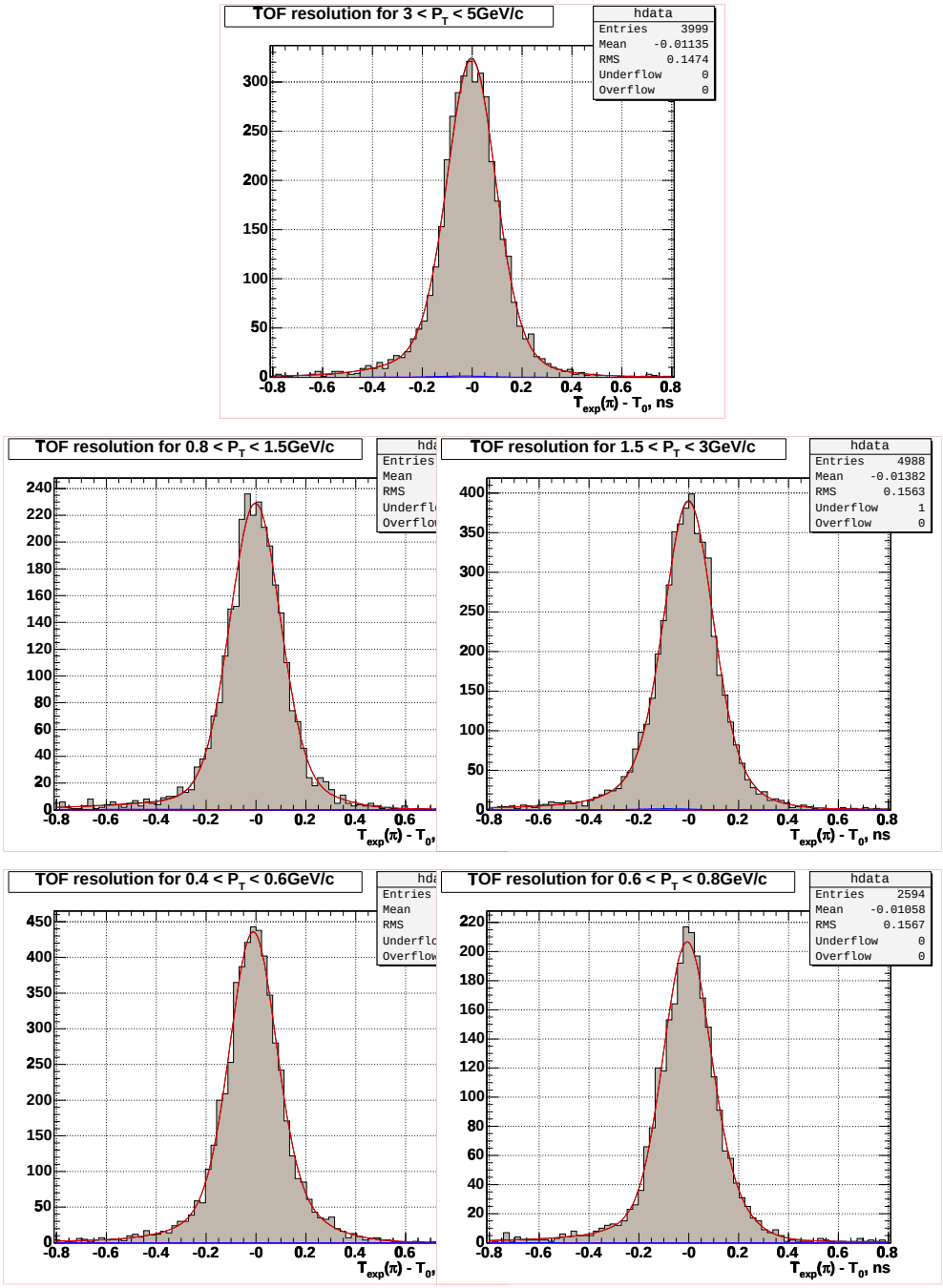
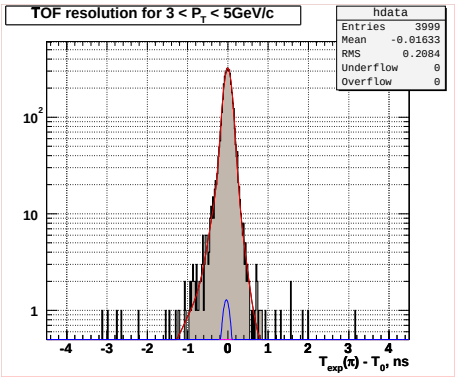
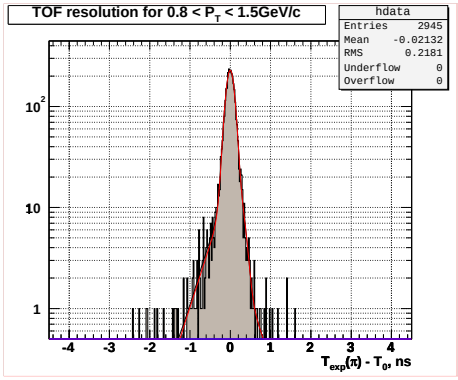
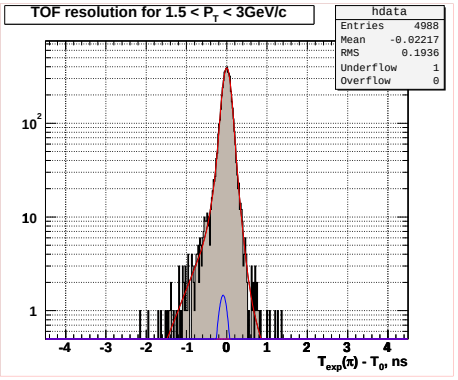
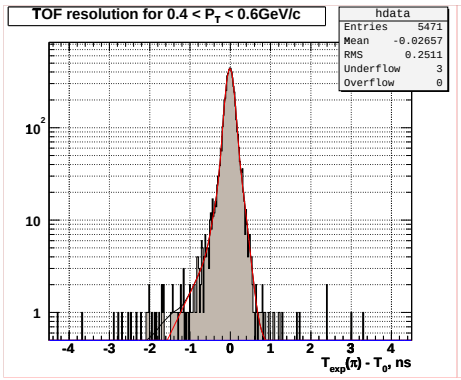


Figure A.3: Likelihood fits for the TOF resolution function parameters.









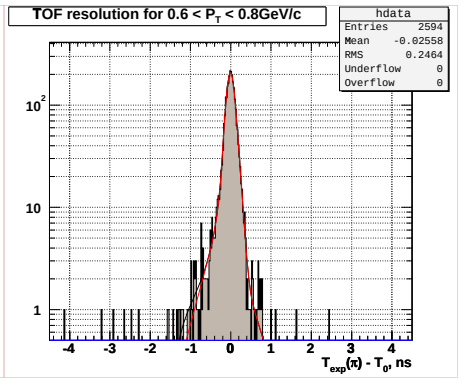


Figure A.4: Likelihood fits for the TOF resolution function parameters shown on the logarithmic scale.

	unit	0.4-0.6 GeV/c	0.6-0.8 GeV/c	0.8-1.5 GeV/c	1.5-3.0 GeV/c	3.0-5.0 GeV/c
σ_1	ps	53.6 ± 4.9	$110. \pm 63.$	65.1 ± 6.1	53.4 ± 4.6	51.2 ± 4.9
A	ns^{-1}	6.71 ± 2.5	8.76 ± 1.7	8.63 ± 2.2	8.81 ± 1.4	7.46 ± 0.93
τ^+	ns^{-1}	1.41 ± 0.32	2.90 ± 0.56	2.80 ± 0.82	3.38 ± 0.75	1.69 ± 0.43
f^+	%	0.713 ± 0.23	2.02 ± 0.54	1.66 ± 0.86	1.33 ± 0.50	0.836 ± 0.26
τ^-	ns^{-1}	2.08 ± 0.35	3.37 ± 0.38	2.86 ± 0.30	2.53 ± 0.26	2.01 ± 0.46
f^-	%	3.14 ± 0.84	5.67 ± 0.85	5.79 ± 0.84	4.86 ± 0.85	2.33 ± 0.87
σ_2	ps	$154. \pm 27.$	$61.2 \pm 14.$	$157. \pm 170$	$130. \pm 26.$	$153. \pm 58.$
f_{G2}	%	23.9 ± 4.1	44.1 ± 9.9	2.49 ± 1.6	19.8 ± 5.6	15.6 ± 4.0
ΔT	ps	-11.6 ± 2.4	-17.0 ± 8.0	-2.42 ± 3.0	-0.743 ± 2.6	-2.38 ± 2.1
ΔT_{G2}	ps	$24.9 \pm 16.$	$39.1 \pm 17.$	$285. \pm 60.$	$2.54 \pm 19.$	$-25.1 \pm 29.$
α_1	ns/cm	0.581 ± 0.056	$((1.87 \pm X)10^{-4}$	0.601 ± 0.064	0.624 ± 0.039	0.654 ± 0.067
α_2	ns/cm	1.16 ± 0.20	1.08 ± 0.19	$((1.03 \pm X)10^{-4}$	1.09 ± 0.27	1.39 ± 0.32
$w_{h/2}$	ps	128	167	149	130	114

Table A.1: Fitted TOF resolution function parameters. The last row in this table shows the width of the TOF resolution at half height assuming a track hitting the bar in front of the PMT and perfect knowledge of event t_0 (and is not a fit parameter).

Appendix B

Validation of fitter in samples of known content

We study the performance of our likelihood fit and check for possible systematic effects on control samples with known composition. For this purpose we take the same reconstructed semileptonic decays that are used for the rest of this analysis and perform the particle content measurements on the sample of the decay products of D decays. We use D^0 and D^+ modes which provide the highest statistics for such comparison.

Assuming that the TOF efficiency doesn't depend on particle species or p_T, we expect to measure 66.6% of pions and 33.3% of kaons in a sample of all tracks from D^- decays and 50%/50% in tracks from D^0 decays. Obviously, these ratios may be different when only a subset of tracks is selected in a given p_T region. Table B.1

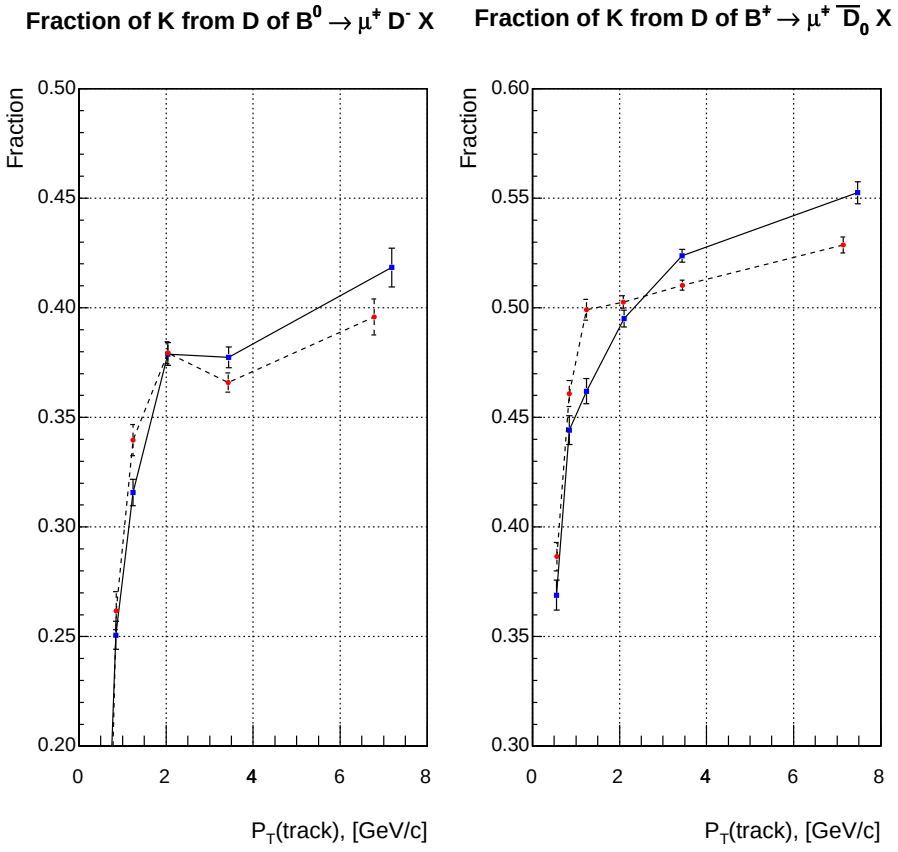


Figure B.1: The measured K fractions in the samples of the D meson decay tracks in $\mu^+ D^-$ and $\mu^+ \bar{D}_0$ decays.

shows the results of the fits performed on a sample of kaon and pion tracks from these decays. The total fraction is very close to the expected result. We also try to determine whether the fits in every p_T bin are correct by comparing these measurements to Monte-Carlo prediction. This is also illustrated graphically by Figure B.1.

Unfortunately this method is limited in its use, first because the statistical errors on the measurements within the p_T bins are rather large, and second because the

Signal Fit - Data							
Type	p_T	π,%	K,%	p,%	e, %	N_{TRK}	N_{EVT}
μ^+D^-	0.4-0.7	99.8022 ± 0.0041	0.198 ± 0.27	0 ± 0	0 ± 0	3523(0.116)	30282
μ^+D^-	0.7-1	74.94 ± 0.41	25.1 ± 0.64	0 ± 0	0 ± 0	4832(0.16)	30282
μ^+D^-	1-1.5	68.43 ± 0.42	31.6 ± 0.61	0 ± 0	0 ± 0	7615(0.251)	30282
μ^+D^-	1.5-2.5	62.11 ± 0.38	37.9 ± 0.52	0 ± 0	0 ± 0	14246(0.47)	30282
μ^+D^-	2.5-5	62.26 ± 0.34	37.7 ± 0.48	0 ± 0	0 ± 0	18459(0.61)	30282
μ^+D^-	5-25	58.2 ± 0.70	41.8 ± 0.89	0 ± 0	0 ± 0	6048(0.2)	30282
$\mu^+D^-,\sum$	0.4-25	66.17 ± 0.17	33.83 ± 0.27	0 ± 0	0 ± 0	54726(1.81)	30282
$\mu^+\bar{D}_0$	0.4-0.7	63.1 ± 0.70	36.9 ± 0.69	0 ± 0	0 ± 0	3145(0.0565)	55629
$\mu^+\bar{D}_0$	0.7-1	55.6 ± 0.68	44.4 ± 0.65	0 ± 0	0 ± 0	3870(0.0696)	55629
$\mu^+\bar{D}_0$	1-1.5	53.8 ± 0.57	46.2 ± 0.57	0 ± 0	0 ± 0	6001(0.108)	55629
$\mu^+\bar{D}_0$	1.5-2.5	50.49 ± 0.41	49.51 ± 0.39	0 ± 0	0 ± 0	16837(0.303)	55629
$\mu^+\bar{D}_0$	2.5-5	47.63 ± 0.33	52.37 ± 0.30	0 ± 0	0 ± 0	27729(0.498)	55629
$\mu^+\bar{D}_0$	5-25	44.8 ± 0.61	55.25 ± 0.50	0 ± 0	0 ± 0	10149(0.182)	55629
$\mu^+\bar{D}_0,\sum$	0.4-25	49.62 ± 0.20	50.38 ± 0.19	0 ± 0	0 ± 0	67733(1.22)	55629
Signal Fit - MC							
μ^+D^-	0.4-0.7	100 ± 0	0 ± 0	0 ± 0	0 ± 0	1243(0.169)	7359
μ^+D^-	0.7-1	73.82 ± 0.52	26.2 ± 0.87	0 ± 0	0 ± 0	1884(0.256)	7359
μ^+D^-	1-1.5	66.04 ± 0.51	34.0 ± 0.71	0 ± 0	0 ± 0	2925(0.398)	7359
μ^+D^-	1.5-2.5	62.06 ± 0.39	37.9 ± 0.50	0 ± 0	0 ± 0	5778(0.785)	7359
μ^+D^-	2.5-5	63.39 ± 0.34	36.6 ± 0.44	0 ± 0	$((1.25 \pm 1.3)10^{-2}$	7540(1.02)	7359
μ^+D^-	5-25	60.4 ± 0.67	39.6 ± 0.82	0 ± 0	0 ± 0	2128(0.289)	7359
$\mu^+D^-,\sum$	0.4-25	66.13 ± 0.19	33.86 ± 0.26	0 ± 0	$((4.38 \pm 4.5)10^{-3}$	21501(2.92)	7359
$\mu^+\bar{D}_0$	0.4-0.7	61.35 ± 0.51	38.7 ± 0.65	0 ± 0	0 ± 0	3468(0.115)	30044
$\mu^+\bar{D}_0$	0.7-1	53.9 ± 0.55	46.1 ± 0.59	0 ± 0	0 ± 0	3795(0.126)	30044
$\mu^+\bar{D}_0$	1-1.5	50.10 ± 0.47	49.90 ± 0.48	0 ± 0	0 ± 0	5535(0.184)	30044
$\mu^+\bar{D}_0$	1.5-2.5	49.74 ± 0.29	50.26 ± 0.29	0 ± 0	0 ± 0	14939(0.497)	30044
$\mu^+\bar{D}_0$	2.5-5	48.97 ± 0.23	51.03 ± 0.23	0 ± 0	$((4.58 \pm 4.4)10^{-3}$	23566(0.784)	30044
$\mu^+\bar{D}_0$	5-25	47.14 ± 0.39	52.86 ± 0.37	0 ± 0	0 ± 0	8783(0.292)	30044
$\mu^+\bar{D}_0,\sum$	0.4-25	50.02 ± 0.14	49.97 ± 0.14	0 ± 0	$((1.80 \pm 1.7)10^{-3}$	60088(2)	30044

Table B.1: The fitted particle fractions in the samples of the D meson decay tracks in μ^+D^- and $\mu^+\bar{D}_0$ decays. The signal component fit results are shown for both data and MC samples.

measured fractions within the bins depend very strongly on the D meson spectrum (which means the comparison itself suffers from large systematics). For example, the results shown here were obtained by re-weighting data and Monte-Carlo $p_T(D)$ spectrum to the same distribution. If we instead re-weight to the same $p_T(lD)$ spectra the results in Monte-Carlo change by several standard deviations.

Therefore while there is noticeable disagreement between data and Monte-Carlo in certain p_T regions, this does not indicate a problem with the technique we use. The fact that the integrated fractions agree with the expected values quite well is encouraging.

So even though we can not assign any systematic uncertainty based on this comparison, it provides a useful cross-check, and no obvious problems with our procedure are discovered.

Appendix C

Probing for the effects of other backgrounds

The studies of the B meson lifetimes performed in the semileptonic samples have encountered major difficulties in properly subtracting the effects of the background. One of the possible explanations was presence of unknown short-lived background in the semileptonic sample that has real D meson component in it but no B mesons. We test for the possible presence and effect of such backgrounds on our results by adding an additional selection requirement $L_{xy}(B) > 0.1$ cm to our B meson reconstruction. This reduces the data size by about one third, but should nearly eliminate any short-lived background events. We then perform the full analysis on this data sample and check for any change in the results. This study is summarized in Table C.1 where the first part of the Table represents the results we obtain by

performing our analysis following the standard prescription, and the second part shows the results obtained after imposing the additional L_{xy} requirement. The dif-

Total fractions around B mesons, standard data selection Value ± (stat.) ±(common syst.) ± (syst.)			
	π,%	K,%	p,%
B^0	$81.4 \pm 0.10 \pm 1.6 \pm 0.33$	$11.6 \pm 0.25 \pm 0.34 \pm 0.085$	$5.05 \pm 0.15 \pm 0.12 \pm 0.072$
B^+	$78.9 \pm 0.12 \pm 3.1 \pm 0.35$	$12.1 \pm 0.28 \pm 0.57 \pm 0.093$	$6.70 \pm 0.17 \pm 0.22 \pm 0.077$
B_s	$71.0 \pm 0.73 \pm 1.3 \pm 1.8$	$20.2 \pm 1.2 \pm 0.59 \pm 0.29$	$6.62 \pm 0.74 \pm 0.17 \pm 0.055$
Total fractions around B meson decays with $L_{xy}(B) > 0.1$ cm			
	π,%	K,%	p,%
B^0	$81.9 \pm 0.13 \pm 1.6 \pm 0.32$	$11.1 \pm 0.33 \pm 0.34 \pm 0.096$	$5.10 \pm 0.20 \pm 0.11 \pm 0.081$
B^+	$78.7 \pm 0.15 \pm 3.2 \pm 0.33$	$12.3 \pm 0.33 \pm 0.58 \pm 0.10$	$6.81 \pm 0.21 \pm 0.22 \pm 0.087$
B_s	$70.0 \pm 0.94 \pm 1.4 \pm 2.0$	$20.8 \pm 1.5 \pm 0.66 \pm 0.32$	$6.35 \pm 0.96 \pm 0.17 \pm 0.085$

Table C.1: Comparison of measured particle fractions in tracks around decays with standard selection criteria and decays with additional $L_{xy}(B) > 0.1$ cm requirement.

ference in measured particle fractions is well within the measurement uncertainties. The sample composition parameters have not been re-evaluated with the new selection, but the changes in the sample composition parameters are expected to be negligible. Therefore we conclude that any additional short-lived backgrounds that may be present in our samples do not have a significant effect on our results. The change in measured yields is not shown here but was also smaller than the statistical error.

Bibliography

[1] D. Acosta et al. A time-of-flight detector in cdf-ii. *Nucl. Instrum. Meth.*, A518:605–608, 2004.

[2] A. Affolder et al. Intermediate silicon layers detector for the cdf experiment. *Nucl. Instrum. Meth.*, A453:84–88, 2000.

[3] T. Affolder et al. Cdf central outer tracker. *Nucl. Instrum. Meth.*, A526:249–299, 2004.

[4] K. Anikeev. *Measurement of the Lifetimes of B Meson Mass Eigenstates*. PhD Dissertation, MIT, 2004.

[5] Konstantin Anikeev, Pasha Murat, and Christoph Paus. Description of bgenerator. *CDF Internal Publication*, 5092, 1999.

[6] A. Artikov et al. Design and construction of new central and forward muon counters for cdf ii. *Nucl. Instrum. Meth.*, A538:358–371, 2005.

[7] G. Ascoli et al. Cdf central muon detector. *Nucl. Instrum. Meth.*, A268:33, 1988.

[8] Bill Ashmanskas et al. The cdf silicon vertex trigger. *Nucl. Instrum. Meth.*, A518:532–536, 2004.

[9] G. Bauer et al. Measurement of b^0 oscillations using same-side tagging in semileptonic b decays. *CDF Internal Publication*, 7011, 2004.

[10] I.I. Bigi and A.I. Sanda. *CP violation.* Cambridge University Press, 2000.

[11] D. Bortoletto. B physics at hadron colliders. *hep-ph/0212022*, 2002.

[12] G.C. Branco, L. Lavoura, and J.P. Silva. *CP violation.* Clarendon Press Oxford, 1999.

[13] R. Brun, F. Bruyant, M. Maire, A. C. McPherson, and P. Zanarini. Geant3. *CERN-DD/EE/84-1*, 1984.

[14] R. Brun, R. Hagelberg, M. Hansroul, and J. C. Lassalle. Geant: Simulation program for particle physics experiments. user guide and reference manual. *CERN-DD-78-2-REV*, 1978.

[15] Andrzej J. Buras. Flavour dynamics: Cp violation and rare decays. *hep-ph/0101336*, 2001.

[16] Ling-Lie Chau and Wai-Yee Keung. Comments on the parametrization of the kobayashi-maskawa matrix. *Phys. Rev. Lett.*, 53:1802, 1984.

[17] C. Chen, M. Jones, et al. Front end electronics for the cdf-ii time-of-flight system. *IEEE Trans. Nucl. Sci.*, 50(6), 2003.

[18] CLEO Collaboration. Exclusive and inclusive semileptonic decays of b mesons to d mesons. *Phys. Rev. D*, 43:651, 1991.

[19] The CDF Collaboration. Measurement of b^0 oscillations and calibration of flavor tagging in in semileptonic decays. *CDF Publication*, 7549, 2005.

[20] The CDF Collaboration. Study of b_s^0 oscillations using semileptonic $b_s^0 \to d_s^- l^+ \nu$ decays. *CDF Publication*, 7542, 2005.

[21] The SLD Collaboration. Measurement of the b-quark fragmentation function in z^0 decays. *hep-ph/0202031*, 2002.

[22] Glen Cowan. *Statistical data analysis*. Oxford University Press Inc., New York, 1998.

[23] d'Auria S. et al. Track-based calibration of the cot specific ionization. *CDF Internal Publication*, 6932, 2004.

[24] S. Eidelman et al. Review of particle physics. *Phys. Lett.*, B592:1, 2004.

[25] H. Fritzsch and J. Plankl. The mixing of quark flavors. *Phys. Rev. D*, 35:1732, 1987.

[26] Gomez G., Jones M., et al. Online calibration of the cdf-ii time-of-flight detector. *CDF Internal Publication*, 6050, 2002.

[27] G. Gomez-Ceballos et al. Event builder and level 3 at the cdf experiment. *Nucl. Instrum. Meth.*, A518:522–524, 2004.

[28] K. Hagiwara et al. Review of particle physics. *Phys. Rev.*, D66:1, 2002.

[29] Haim Harari and Miriam Leurer. Recommending a standard choice of cabibbo angles and ckm phases for any number of generations. *Phys. Lett.*, B181:123, 1986.

[30] Christopher S. Hill. Operational experience and performance of the cdfii silicon detector. *Nucl. Instrum. Meth.*, A530:1–6, 2004.

[31] Fernandez J. et al. The time-of-flight off-line calibrations. *CDF Internal Publication*, 6990, 2004.

[32] F. James and M. Roos. 'minuit' a system for function minimization and analysis of the parameter errors and correlations. *Comput. Phys. Commun.*, 10:343–367, 1975.

[33] M. Kobayashi and T. Maskawa. Cp violation in the renormalizable theory of weak interaction. *Prog. Theor. Phys.*, 49:652–657, 1973.

[34] R.V. Kowalewski. B physics and cp violation. *hep-ph/0305024*, 2003.

[35] D. J. Lange. The evtgen particle decay simulation package. *Nucl. Instrum. Meth.*, A462:152–155, 2001.

[36] Jones M. Tof calibration studies. *CDF Internal Publication*, 6948, 2004.

[37] Petar Maksimovic. *Observation of $\pi - B$ meson Charge-flavor Correlations and Measurement of Time Dependent $B^0\bar{B}^0$ Mixing in $p\bar{p}$ Collisions.* PhD Dissertation, MIT, 1997.

[38] Squillacioti P. et al. Particle identification by combining tof and dedx information. *CDF Internal Publication*, 7488, 2005.

[39] Donals Perkins. *Introduction to High Energy Physics.* Cambridge University Press; 4th edition, 2000.

[40] Michael E. Peskin and D. V. Schroeder. *An Introduction to quantum field theory.* Reading, USA: Addison-Wesley, 1995.

[41] C. Peterson, D. Shlatter, I. Schmitt, and P.M. Zerwas. Scaling violations in inclusive e+e- annihilation spectra. *Phys. Rev. D*, 27:105, 1983.

[42] A. Pich. Weak decays, quark mixing and cp violation: Theory overview. *hep-ph/9709441*, 1997.

[43] Alexandre P. Pronko. *Fragmentation of quark and gluon jets in proton-antiproton collisions at $\sqrt{s} = 1.8$ TeV.* PhD Dissertation, University of Florida, 2005.

[44] Yu S. et al. Cot de/dx measurement and corrections. *CDF Internal Publication*, 6361, 2004.

[45] O. Schneider. $b^0 - \bar{b}^0$ mixing. *hep-ph/0405012*, 2003.

[46] Torbjorn Sjostrand, Leif Lonnblad, and Stephen Mrenna. Pythia 6.2: Physics and manual. *hep-ph/0108264*, 2001.

[47] E. J. Thomson et al. Online track processor for the cdf upgrade. *IEEE Trans. Nucl. Sci.*, 49:1063–1070, 2002.

[48] Tatjana Unverhau. *A Measurement of the Lifetime of the λ_b Baryon with the CDF detector at the Tevatron Run II.* PhD Dissertation, University of Glasgow, 2004.

[49] R. Veenhof. Garfield, recent developments. *Nucl. Instrum. Meth.*, A419:726–730, 1998.

[50] B.R. Webber. Fragmentation and hadronization. *hep-ph/9912292*, 1991.

[51] Hartmut Wittig. Status of lattice calculations of b-meson decays and mixing. *hep-ph/0310329*, 2003.

[52] L. Wolfenstein. Parametrization of the kobayashi-maskawa matrix. *Phys. Rev. Lett.*, 51:1945, 1983.

[53] Shin-Shan Yu. *First Measurement of the Ratio of Branching Fractions $\mathcal{B}(\Lambda_b \to \Lambda_c^+ \mu^- \bar{\nu}_\mu)/\mathcal{B}(\Lambda_b \to \Lambda_c^+ \pi^-$ at CDF II.* PhD Dissertation, University of Pennsylvania, 2005.