

STUDY OF THE BRANCHING RATIO FOR THE DECAY  
 $B^0 \rightarrow \pi^0 \pi^0 \pi^0$

CHRISTOPHER A. SERINO

Submitted to the faculty of the  
University of Massachusetts Amherst in partial fulfillment  
of the requirements for the degree of

BACHELOR OF SCIENCE with HONOURS

Department of Physics

Amherst, Massachusetts  
January 2007

STUDY OF THE BRANCHING RATIO FOR THE DECAY  
 $B^0 \rightarrow \pi^0 \pi^0 \pi^0$

A Thesis presented

by

CHRISTOPHER A. SERINO

Approved as to style and content by:

---

Carlo Dallapiccola, Chair

---

Stéphane Y. Willocq, Member

## Acknowledgments

There are, without exaggeration, a plethora of people without whom, this thesis simply could not exist. The sheer volume of involved persons makes me fearful I will forget to mention a fair few of them; as such, let me thank, before I go any further, anyone who has ever heard me utter the phrase *B meson*. That having been said, I believe the contributors to this thesis can be split into two groups: those who offered moral support, and those who offered scientific support. I shall begin with the former.

I offer a heart felt thank you to my family—Mom, Dad, Nicholas, and Robin—without whom it is unclear whether or not I would have even graduated. Similarly, I thank my grandparents, aunts, uncles, and cousins for all of their encouragement. Also, I would like to thank my friend Dave for all the hours we spent working together on problem sets, studying for exams, and discussing the fate of the Red Sox.

On the scientific front, I wish to thank the UMASS Amherst Particle Physics Group, especially Professor Blaylock, Professor Hertzbach, Professor Willocq, and Dr. Staengle for their expertise in the field and their advise there on. I also wish to thank Professor Guyer and Professor Kevrekidis for their guidance, support, and their efforts to teach scientific computing to me. I must also thank the Oxford Particle Physics Group for access to their computing facilities during my time at Trinity College, especially Professor Libby, as well as my college tutors: Professors Read and Wark. I would like to thank Professor Fermann of the UMASS Chemistry Department for first getting me involved in research and introducing me to computational based problem solving as well as for his continued support through the years. I would also like to thank the *BABAR* Collaboration, especially Dr. Back, Dr. Barlow, Dr. Pierini, and Dr. Wilson.

Finally, I would like to thank Professor Carlo Dallapiccola. It is impossible to summarize all of his contributions to this project and to my education, let alone the impact these contributions have had on me. Suffice it to say, Professor Dallapiccola has crafted me into a young curious physicist, eager to make him, and all of the others mentioned herein, proud of my contributions to the field. Hence, I present this thesis on behalf of everyone who has helped me along the way with the caveat that this is merely the beginning and that greater insight is, God willing, on the way.

# ABSTRACT

## Study of the Branching Ratio for the Decay

$$B^0 \rightarrow \pi^0\pi^0\pi^0$$

An upper limit at the 90% confidence level for the branching ratio of the rare three-body charmless decay  $B^0 \rightarrow \pi^0\pi^0\pi^0$  is presented. The analysis uses data collected from the more than three billion  $e^+ e^-$  collisions at SLAC's asymmetric  $B$  Factory that is operating at the  $\Upsilon(4S)$  resonance energy. The result of these collisions is data on over two-hundred million  $B^0\bar{B}^0$  decays. This data has been collected by the *BABAR* detector and has been initially analyzed using *BABAR* software. The value of  $\mathcal{B}(B^0 \rightarrow \pi^0\pi^0\pi^0)$ , previously unmeasured, is found to be less than  $1.56 \times 10^{-6}$  at the 90% confidence level.

# TABLE OF CONTENTS

|   | Page |
|---|------|
| ABSTRACT . . . . .  | iv   |
| LIST OF TABLES . . . . .  | vii  |
| LIST OF FIGURES . . . . .   | ix   |
| CHAPTER   |      |
| 1. <b>Introduction</b> . . . . .                                    | 1    |
| 1.1 The Standard Model . . . . .                                    | 1    |
| 1.2 An asymmetric $B$ Factory at SLAC . . . . .                     | 5    |
| 1.2.1 The Linear Accelerator . . . . .                              | 6    |
| 1.2.2 The PEP-II Positron-Electron Collider . . . . .               | 7    |
| 1.3 The $BABAR$ Detector . . . . .                                  | 8    |
| 1.3.1 Silicon Vertex Tracker . . . . .                              | 9    |
| 1.3.2 Drift Chamber . . . . .                                       | 10   |
| 1.3.3 Detector of Internally Reflected Cherenkov Light . . . . .    | 11   |
| 1.3.4 Electromagnetic Calorimeter . . . . .                         | 13   |
| 1.3.5 Muon and Neutral Hadron Detector . . . . .                    | 14   |
| 2. <b>Motivation</b> . . . . .                                      | 16   |
| 2.1 Eigenstates, Eigenfunctions, and Eigenvalue Equations . . . . . | 16   |
| 2.1.1 Eigenvalue Problems . . . . .                                 | 16   |
| 2.1.2 The $CP$ Operator . . . . .                                   | 17   |
| 2.1.3 $CP$ Eigenstates . . . . .                                    | 21   |
| 2.2 $CP$ Violation in a Matter Dominated Universe . . . . .         | 22   |
| 2.3 Theory . . . . .  | 26   |
| 2.3.1 Comparisons To $B^\pm$ Decays . . . . .                       | 27   |
| 2.3.2 Comparisons to Other $B^0$ Decay Modes . . . . .              | 30   |
| 3. <b>Monte Carlo Background Analysis</b> . . . . .                 | 34   |
| 3.1 Building the Monte Carlo Data Set . . . . .                     | 35   |
| 3.2 Types of Background . . . . .                                   | 37   |
| 3.2.1 Continuum Background ( $q\bar{q}$ ) . . . . .                 | 41   |

|       |  |     |
|-------|--|-----|
| 3.2.2 | $B^\pm$ Decays with Pion Containing Final States . . . . .   | 45  |
| 3.2.3 | $B^0$ Decays with Pion Containing Final States and $B^0$<br>$\rightarrow \pi^0\pi^0\pi^0$ Combinatorial Background . . . . . | 49  |
| 3.3   | Background Reduction . . . . .   | 52  |
| 3.3.1 | Skims . . . . .  | 52  |
| 3.3.2 | NonCharm3BodyUser: the Data Acquisition and Man-<br>agement Package . . . . .  | 55  |
| 3.3.3 | Cuts on Observables . . . . .  | 57  |
| 3.3.4 | The Optimized Data Set . . . . .   | 70  |
| 4.    | <b>Monte Carlo Signal Analysis</b> . . . . .   | 73  |
| 4.1   | Extracting $\mathcal{N}_{sig}$ and $\mathcal{N}_{bkg}$ from the $m_{ES}^{B^0}$ Distribution . . . . .                        | 74  |
| 4.1.1 | The Fitting Scheme: <code>Minuit</code> . . . . .  | 76  |
| 4.1.2 | Signal . . . . .   | 77  |
| 4.1.3 | Background . . . . .   | 79  |
| 4.1.4 | The Complete $m_{ES}^{B^0}$ Distribution . . . . .   | 81  |
| 4.2   | Calculating Confidence Levels for an Upper Limit Analysis . . . . .  | 82  |
| 4.2.1 | Theory . . . . .   | 84  |
| 4.2.2 | A Fictitious Example . . . . .   | 85  |
| 5.    | <b>Data Analysis</b> . . . . .   | 86  |
| 5.1   | The <i>BABAR</i> Data Set . . . . .  | 86  |
| 5.2   | The Skim-Reduced Data Set . . . . .  | 87  |
| 5.3   | The Cut-Reduced Data Set . . . . .   | 94  |
| 5.4   | An Upper Limit for $\mathcal{B}(B^0 \rightarrow \pi^0\pi^0\pi^0)$ . . . . .  | 95  |
| 6.    | <b>Conclusions</b> . . . . .   | 99  |
| 6.1   | Results . . . . .  | 99  |
| 6.2   | Summary of the Analysis . . . . .  | 100 |
| 6.3   | Difficulties with the Analysis . . . . .   | 102 |
| 6.4   | Future Work on $B^0 \rightarrow \pi^0\pi^0\pi^0$ . . . . .   | 103 |
|       | REFERENCES . . . . .   | 105 |

## LIST OF TABLES

| Table   | Page |
|---|------|
| 1. Qualitative Solutions to the $C$ Eigenvalue Problem . . . . .  | 19   |
| 2. Qualitative Solutions to the $P$ Eigenvalue Problem . . . . .  | 20   |
| 3. Qualitative Solutions to the $CP$ Eigenvalue Problem . . . . .   | 21   |
| 4. Theoretical Predictions for $\mathcal{B}(B^- \rightarrow \pi^- \pi^+ \pi^-)$ . . . . .                     | 27   |
| 5. Experimental Results on Measurements of $\mathcal{B}(B^\pm \rightarrow \pi^\pm \pi^\mp \pi^\pm)$ . . . . . | 28   |
| 6. Initial Size of the Monte Carlo Data Set . . . . .   | 36   |
| 7. Some $B^\pm$ Decays with Final States Containing At Least One $\pi^0$ . . . . .                            | 46   |
| 8. Number of $B^0$ Reconstructions for Events with $n$ Detected $\pi^0$ Mesons                                | 51   |
| 9. Expected Relative Decay Rates and the Skimmed MC Data . . . . .  | 53   |
| 10. Size of the Monte Carlo Data Set after Skimming the Data and Preserving the Decay Ratios . . . . .        | 53   |
| 11. Percentage of Data Remaining after Applying the Skims . . . . .   | 54   |
| 12. Average Number of Reconstructions per $e^+ e^-$ Collision . . . . .                                       | 54   |
| 13. Discretized Observable Domains for Cuts 1 . . . . .   | 63   |
| 14. Cuts 1 Results . . . . .  | 63   |
| 15. Size of the Monte Carlo Data Set after Applying Cuts 1 . . . . .  | 65   |
| 16. Percentage of Data Remaining after Applying Cuts 1 . . . . .  | 65   |
| 17. Discretized Observable Domains for Cuts 2 . . . . .   | 67   |
| 18. Cuts 2 Results . . . . .  | 68   |
| 19. Size of the Monte Carlo Data Set after Applying Cuts 2 . . . . .  | 69   |
| 20. Percentage of Data Remaining after Applying Cuts 2 . . . . .  | 69   |

|   |    |
|---|----|
| 21. Summery of Background Elimination . . . . .   | 70 |
| 22. Percentages of Background Eliminated during Various Background Re-<br>duction Stages . . . . .    | 71 |
| 23. Signal Efficiency . . . . .   | 73 |
| 24. A Fictitious Upper Limit Calculation . . . . .  | 85 |
| 25. A Summery of Runs 1–4 . . . . .   | 87 |
| 26. Number of Events in the Runs 1–4 Data Sets after Skims . . . . .                                  | 88 |
| 27. Number of Events in the Runs 1–4 Data Sets at Various Stages of<br>Background Reduction . . . . . | 89 |
| 28. The Effect of Cuts on Runs 1–4 . . . . .  | 94 |
| 29. Properties of the Two Signal Events . . . . .   | 95 |
| 30. Limit Calculator Inputs . . . . .   | 98 |

## LIST OF FIGURES

| Figure  | Page |
|---|------|
| 1.1. The Standard Model of Particle Physics . . . . .   | 2    |
| 1.2. The asymmetric $B$ Factory at SLAC . . . . .   | 5    |
| 1.3. The Linear Accelerator at SLAC . . . . .   | 6    |
| 1.4. The PEP-II Positron-Electron Collider at SLAC . . . . .  | 7    |
| 1.5. The <i>BABAR</i> Detector . . . . .  | 9    |
| 1.6. The Silicon Vertex Tracker . . . . .   | 9    |
| 1.7. The Drift Chamber . . . . .  | 10   |
| 1.8. The DIRC . . . . .   | 11   |
| 1.9. A photograph of the EMC . . . . .  | 13   |
| 1.10. A Reconstructed $\mu^-$ Track . . . . .   | 14   |
| 2.1. $CP$ Acts on an $e^-$ . . . . .  | 20   |
| 2.2. Feynman Diagrams for the Decays $B^0/\bar{B}^0 \rightarrow \pi^0\pi^0\pi^0$ . . . . .            | 26   |
| 2.3. Feynman Diagrams for the Decays $B^\pm \rightarrow \pi^\pm \pi^\mp \pi^\pm$ . . . . .            | 28   |
| 2.4. Feynman Diagram for the non-resonant Decay $B^0 \rightarrow \pi^+ D^-$ . . . . .                 | 31   |
| 3.1. Track Reconstruction and the Thrust Axis . . . . .   | 39   |
| 3.2. $\cos(\theta_{thrust})$ Distribution for $c\bar{c}$ Continuum Events . . . . .                   | 43   |
| 3.3. Fisher Discriminant Distribution for $c\bar{c}$ Continuum Events . . . . .                       | 43   |
| 3.4. $\cos(\theta_{thrust})$ Distribution for $u\bar{u}/d\bar{d}/s\bar{s}$ Continuum Events . . . . . | 44   |
| 3.5. Fisher Discriminant Distribution for $u\bar{u}/d\bar{d}/s\bar{s}$ Continuum Events . . . . .     | 45   |
| 3.6. $\cos(\theta_{thrust})$ Distribution for $B^\pm$ Background Events . . . . .                     | 47   |
| 3.7. $\Delta E$ Distribution for $B^\pm$ Background Events . . . . .                                  | 48   |

|       |  |     |
|-------|--|-----|
| 3.8.  | $\cos(\theta_{thrust})$ Distribution for $B^0$ Background Events . . . . .                         | 50  |
| 3.9.  | $\Delta E$ Distribution for $B^0$ Background Events . . . . .                                      | 51  |
| 3.10. | The Signal and Background Distributions for the Cut Observables . .                                | 58  |
| 3.11. | The Signal Distributions for the Cut Observables with a Cut on One<br>Observable Applied . . . . . | 60  |
| 3.12. | Observable Background Distributions with a Cut on One Observable<br>Applied . . . . .              | 61  |
| 3.13. | Baseline Cuts Scheme . . . . .   | 63  |
| 3.14. | Baseline Cuts . . . . .  | 64  |
| 3.15. | Tight Cuts Scheme . . . . .  | 68  |
| 3.16. | Tight Cuts . . . . .   | 69  |
| 3.17. | Monte Carlo $m_{ES}^{B^0}$ Background Distribution . . . . .                                       | 72  |
| 4.1.  | An Ideal $m_{ES}^{B^0}$ Signal Distribution . . . . .  | 78  |
| 4.2.  | Monte Carlo $m_{ES}^{B^0}$ Signal Distribution . . . . .   | 79  |
| 4.3.  | A Well-Behaved $m_{ES}^{B^0}$ Background Distribution . . . . .                                    | 80  |
| 4.4.  | Monte Carlo $m_{ES}^{B^0}$ Background Distribution . . . . .                                       | 81  |
| 4.5.  | An Ideal $m_{ES}^{B^0}$ Distribution . . . . .   | 82  |
| 4.6.  | The Confidence Level Calculator . . . . .  | 83  |
| 5.1.  | The Run 1-4 Distributions of the Cut Observables . . . . .   | 89  |
| 5.2.  | The Run 1 Distributions of the Cut Observables . . . . .   | 90  |
| 5.3.  | The Run 2 Distributions of the Cut Observables . . . . .   | 91  |
| 5.4.  | The Run 3 Distributions of the Cut Observables . . . . .   | 92  |
| 5.5.  | The Run 4 Distributions of the Cut Observables . . . . .   | 93  |
| 6.1.  | A Comparison of Monte Carlo and Run 1-4 Distributions of the Cut<br>Observables . . . . .          | 101 |

# CHAPTER 1

## Introduction

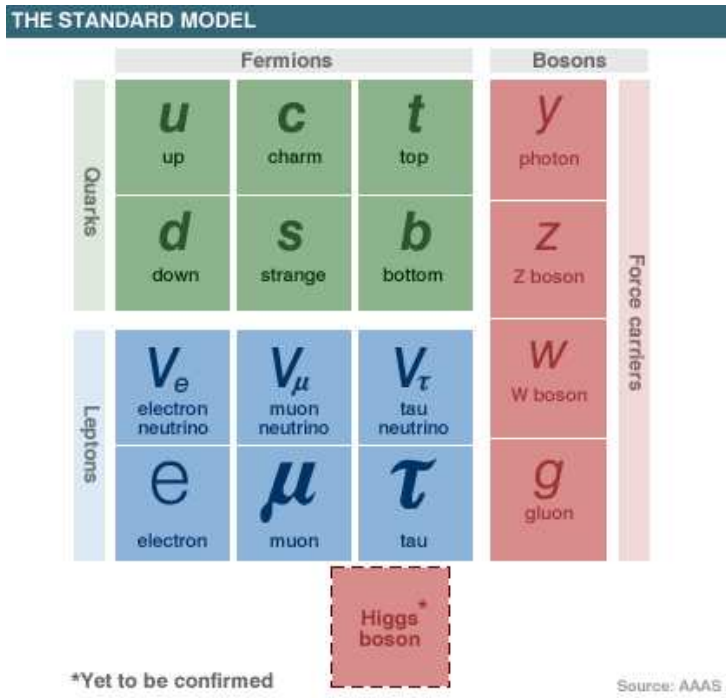
### 1.1 The Standard Model

The Standard Model of particle physics is a quantum field theory, consistent with both quantum mechanics and special relativity, which describes three of the the four fundamental forces of nature as well as the fundamental constituents of matter. Each of the three fundamental forces, the Strong Force and the Electroweak Force (which is the unification of the Weak Interaction and the Electromagnetic Force), are mediated by a force carrying particle or particles, shown in Figure 1.1. Further, when matter and antimatter are reduced down to their indivisible components; all that is left are the fundamental particles shown in Figure 1.1.

Each of the particles in the Standard Model can be classified as either *fermions* or *bosons*. Fermions are particles that posses odd-half integral spin. Any description of a group of fermions must be antisymmetric under interchange (see Equation (1.1)) and thus, these particles obey Fermi-Dirac Statistics. The fermions of the Standard Model are shown in Figure 1.1(b).

$$|1, 2, \dots, i, \dots, j, \dots, N\rangle = -|1, 2, \dots, j, \dots, i, \dots, N\rangle \quad (1.1)$$

Bosons, on the other hand, have integral or no spin. These particles are symmetric under interchange (see Equation (1.2)) and therefor obey Bose-Einstein Statistics.



(a) All Particle of the The Standard Model of Particle Physics

| <b>FERMIONS</b>                  |                         |                 | matter constituents<br>spin = 1/2, 3/2, 5/2, ... |                                 |                 |
|----------------------------------|-------------------------|-----------------|--|---------------------------------|-----------------|
| Leptons spin = 1/2               |                         |                 | Quarks spin = 1/2                                |                                 |                 |
| Flavor                           | Mass GeV/c <sup>2</sup> | Electric charge | Flavor   | Approx. Mass GeV/c <sup>2</sup> | Electric charge |
| ν <sub>e</sub> electron neutrino | <1×10 <sup>-8</sup>     | 0               | <b>u</b> up                                      | 0.003                           | 2/3             |
| <b>e</b> electron                | 0.000511                | -1              | <b>d</b> down                                    | 0.006                           | -1/3            |
| ν <sub>μ</sub> muon neutrino     | <0.0002                 | 0               | <b>c</b> charm                                   | 1.3                             | 2/3             |
| <b>μ</b> muon                    | 0.106                   | -1              | <b>s</b> strange                                 | 0.1                             | -1/3            |
| ν <sub>τ</sub> tau neutrino      | <0.02                   | 0               | <b>t</b> top                                     | 175                             | 2/3             |
| <b>τ</b> tau                     | 1.7771                  | -1              | <b>b</b> bottom                                  | 4.3                             | -1/3            |

(b) The fermions of The Standard Model

| <b>BOSONS</b>                |                         |                 | force carriers<br>spin = 0, 1, 2, ... |                         |                 |
|------------------------------|-------------------------|-----------------|---------------------------------------|-------------------------|-----------------|
| Unified Electroweak spin = 1 |                         |                 | Strong (color) spin = 1               |                         |                 |
| Name                         | Mass GeV/c <sup>2</sup> | Electric charge | Name                                  | Mass GeV/c <sup>2</sup> | Electric charge |
| <b>γ</b> photon              | 0                       | 0               | <b>g</b> gluon                        | 0                       | 0               |
| <b>W<sup>-</sup></b>         | 80.4                    | -1              |                                       |                         |                 |
| <b>W<sup>+</sup></b>         | 80.4                    | +1              |                                       |                         |                 |
| <b>Z<sup>0</sup></b>         | 91.187                  | 0               |                                       |                         |                 |

(c) The bosons of The Standard Model

Figure 1.1: The Standard Model of Particle Physics

The bosons of the Standard Model are shown in Figure 1.1(c).

$$|1, 2, \dots, i, \dots, j, \dots, N\rangle = +|1, 2, \dots, j, \dots, i, \dots, N\rangle \quad (1.2)$$

The fundamental fermions are the matter constituents and can be further divided into *leptons* and *quarks*. Together, the fermions of the Standard Model are organized into three *generations*, each generation consisting of a lepton, its antiparticle, the lepton's neutrino and anti-neutrino, a positively charged quark, a negatively charged quark, and their respective anti-quarks. The generations are organized by increasing mass and are shown in the columns of Figure 1.1(a) as well as the rows of Figure 1.1(b).

The lepton family consists of the electron ( $e^-$ ), the muon ( $\mu^-$ ), and the tau ( $\tau^-$ ), as well as their associated neutrinos ( $\nu_e$ ,  $\nu_\mu$ , and  $\nu_\tau$  respectively). Each of these six particles also has an antiparticle which is exactly the same in all observable ways except that it has opposite charge<sup>1</sup>. These particles are denoted as  $e^+$ ,  $\mu^+$ , and  $\tau^+$  and are associated with the anti-neutrinos  $\bar{\nu}_e$ ,  $\bar{\nu}_\mu$ , and  $\bar{\nu}_\tau$  respectively.

Quarks come in six flavors: up, down, charm, strange, top, and bottom (sometimes the latter being referred to as truth and beauty). They are denoted by  $u$ ,  $d$ ,  $c$ ,  $s$ ,  $t$ , and  $b$  respectively and each has its own antiparticle partner denoted by  $\bar{u}$ ,  $\bar{d}$ ,  $\bar{c}$ ,  $\bar{s}$ ,  $\bar{t}$ , and  $\bar{b}$  respectively. The positively charged quarks ( $u$ ,  $c$ , and  $t$  in order of increasing mass) carry  $\frac{2}{3}$  of the fundamental unit charge as determined in the now famous Millikan Oil Drop Experiment whereas the negatively charge quarks ( $d$ ,  $s$ , and  $b$  in order of increasing mass) carry  $-\frac{1}{3}$  of the fundamental unit charge.<sup>2</sup>

Why then, should the fundamental charge be taken as something other than the

---

<sup>1</sup>The relationship between neutrinos and their anti-neutrinos, both chargeless, is slightly more complicated than charge conjugation.

<sup>2</sup>Millikan's original data set contained one anomalous entry that lead him to calculate the  $e^-$  charge as  $\frac{1}{3}$  of the accepted value as established by the rest of his data. Ever since, this anomaly has inspired experimentalists to search for unbound quarks. One such experiment is being conducted at SLAC and its design is very similar to Millikan's, only at SLAC, all of the

charge carried by the negatively charged quarks? As it turns out, quarks, unlike leptons, do not exist on their own in nature as far as we know. That is to say, quarks are always observed in a bound state of multiple quarks and/or anti-quarks called *hadrons*.

Hadrons can be divided into two more groups called *mesons* and *baryons*. Mesons are bound states of quark anti-quark pairs (denoted  $q\bar{q}$ ) and have integral spin; thus, they are bosons. Baryons are bound states of three quarks (or anti-quarks) and contain odd-half integral spin which makes them fermions. The proton ( $u u d$ ) and the neutron ( $u d d$ ) are the two most well known baryons as, together, they make up the contents of the atomic nucleus. Some theorists have proposed bound states of five quarks, called penta-quarks, but they have yet to be observed.

Finally, quarks carry color charge which comes in three varieties: red, green, and blue. Color charge is to the strong force as electric charge is to the electromagnetic force. Since the strong force is not observed to interact with hadrons, for a bound state of quarks to be allowed, it must be “colorless”. Just as the electromagnetic force will not interact with electrically neutral particles, the strong force will not interact with colorless particles. For mesons, this can be accomplished by combining a color with its anti-color (e.g. a red quark with an anti-red quark) whereas neutral baryons must contain one quark of each color (thus, red plus green plus blue is equal to colorless).

The fundamental bosons are the force carriers. The massless, chargeless, photon ( $\gamma$ ) with unit spin mediates the electromagnetic force. The weak interaction is mediated by the  $Z^0$  gauge boson and the  $W^\pm$  gauge bosons. Each of these bosons possess unit spin and are the only fundamental bosons to have non-zero mass.

---

data collection and analysis is performed by computers with little to no human intervention. Nevertheless, Millikan’s anomalous data point remains the only “evidence” of an unbound quark.

The strong force, which is responsible for binding quarks together into hadrons, is mediated by eight massless, chargeless gluons all with unit spin. The gravitational force is not accounted for in the Standard Model, though in an effort to unify gravity with the Strong and Electroweak Forces, some theorists predict that the gravitational force is mediated by a single boson, named, though not observed, the graviton which possesses an intrinsic spin of two.

## 1.2 An asymmetric $B$ Factory at SLAC

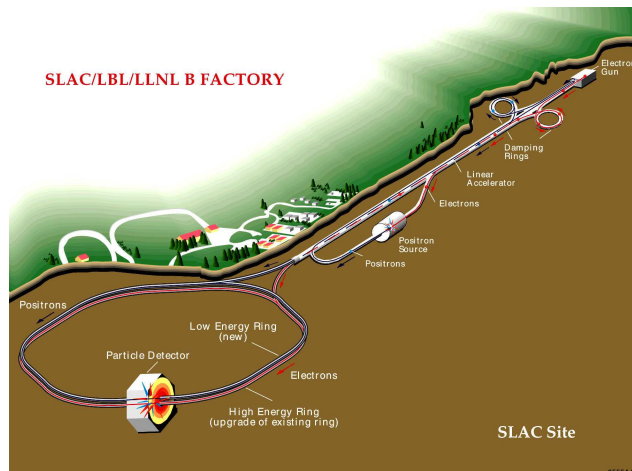


Figure 1.2: The asymmetric  $B$  Factory at SLAC

The Stanford Linear Accelerator Center (SLAC) in Menlo Park, California is the site of the *BABAR* Experiment. The accelerator at SLAC has three main components, the linear accelerator (linac), the PEP-II Positron-Electron Collider, and the *BABAR* detector. The facility was built in 1962 and the linear accelerator became operational in 1966. Construction on the PEP-II  $e^+e^-$  Collider began in 1994 and was up and running in July of 1998. In the same month, construction

on the *BABAR* detector was completed and the experiment began taking data. The PEP-II was designed in such a manner that high luminosities could be achieved at the *BABAR* detector (an important part of a *CP*-Violation experiment). The aforementioned three independently constructed apparatuses, now joint and working in unison at SLAC, make up the *B* Factory as is shown in Figure 1.2. The *B* Factory was designed in an asymmetric fashion so that Charge-Parity (*CP*) Violation experiments—the primary goal of the *BABAR* Experiment—could be conducted most easily.

### 1.2.1 The Linear Accelerator

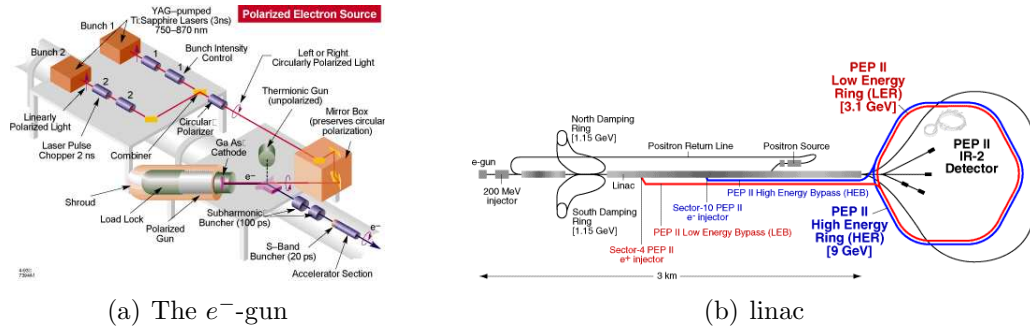


Figure 1.3: The Linear Accelerator at SLAC

The 3 km long linear accelerator at SLAC, shown in Figure 1.3(b), is the initial means of accelerating the electrons and positrons. Linac consists of a 200 MeV  $e^-$ -gun, shown in Figure 1.3(a), as well as two damping rings of 1.15 GeV each. Linac supplies high energy electrons and positrons for the PEP-II Collider which runs parallel to the linac for the rest of its 3 km stretch before it veers off into two separate storage rings, one for the  $e^+$ 's and one for the  $e^-$ 's, situated one on top of the other.

## 1.2.2 The PEP-II Positron-Electron Collider

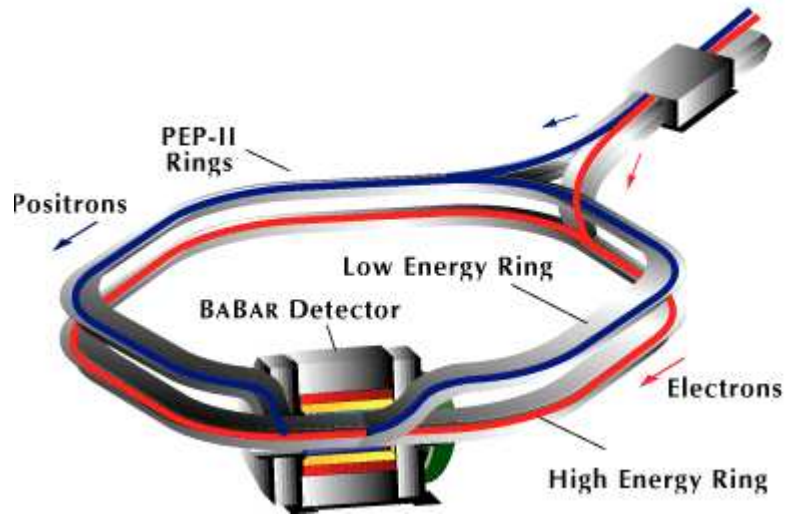


Figure 1.4: The PEP-II Positron-Electron Collider at SLAC

The PEP-II is an asymmetric  $e^+e^-$  collider and is pictured in Figure 1.4. By colliding leptons rather than hadrons, the initial state of the system is known with extremely high precision relative to the initial state of a head on hadron collision (such as those that will take place at CERN). This precision is useful in  $CP$  Violation studies. On the other hand, by colliding particles whose rest mass is much smaller than that of hadrons, it is more difficult to achieve high center of mass energies (hence CERN will be able to achieve the highest ever center of mass energy collisions). In the case of the asymmetric  $B$  Factory, maximizing the beam energy was not a concern as the objective was to operate at the  $\Upsilon(4S)$  resonance. The beam energy is tuned to the  $\Upsilon(4S)$  resonance because this energy is very nearly the same as twice the rest mass energy of the  $B^0$ . As a result, the  $\Upsilon(4S)$  always decays into a  $B\bar{B}$  pair (either a  $B^0\bar{B}^0$  pair or a  $B^+B^-$  pair). In order for the PEP-II to operate as an asymmetric collider, the  $e^+$  beam and the  $e^-$  beams are maintained

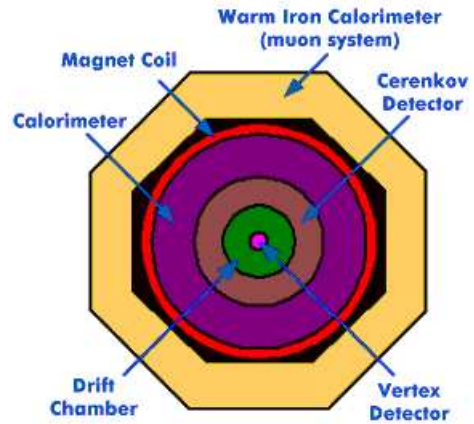
at different energies. The  $e^-$  beam is maintained at the higher energy of 9 GeV while the  $e^+$  beam is held at 3.1 GeV. The electrons and positrons collide head-on in the *BABAR* detector and the difference in beam energy creates an  $\Upsilon(4S)$  that has a significant momenta along the beam line in the laboratory frame while at rest, by definition, in  $e^+e^-$  center of mass frame. Since the  $\Upsilon(4S)$  rest mass energy is very nearly the same as twice the rest mass energy of the  $B^0$ , the center of mass of the  $B^0\bar{B}^0$  pair has a definite momenta along the beam line in the laboratory frame; however, in the center of mass frame, they too are very nearly at rest. The advantages to this methodology is two-fold. First, the lifetimes of the  $B^0$  and the  $\bar{B}^0$  can be measured accurately by measuring the distance they travel along the beam line. Secondly, reconstructing daughter particle tracks is made easier as they *all* must originate from a common origin in the plane orthogonal to the beam axis.

### 1.3 The *BABAR* Detector

The *BABAR* detector, shown in Figure 1.5, was designed with  $CP$ -Violation experiments in mind. As such, many of the design features were motivated by the necessary data and precisions needed to conduct a  $CP$ -Violation experiment. The detector contains five concentric cylindrical “sub-detectors” that work together to return as accurate and precise particle decay data as possible. For the purposes of the mode  $B^0 \rightarrow \pi^0\pi^0\pi^0$ , the most important decay to reconstruct is that of a neutral pion decaying into two photons ( $\pi^0 \rightarrow \gamma\gamma$ ). This is done primarily by using the data from the Electromagnetic Calorimeter “sub-detector” or the EMC. In addition to the “sub-detectors”, the *BABAR* detector contains a superconducting coil that creates a 1.5 Tesla solenoidal magnetic field that bends charged tracks



(a) An artist's interpretation of the *BABAR* detector.



(b) A cross-section of the *BABAR* detector

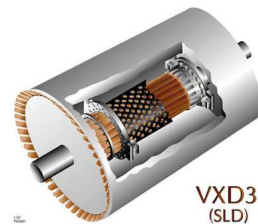
Figure 1.5: The *BABAR* Detector

according to their charge and momenta.[10]

### 1.3.1 Silicon Vertex Tracker



(a) A photograph of the SVT showing the array of silicon chips.



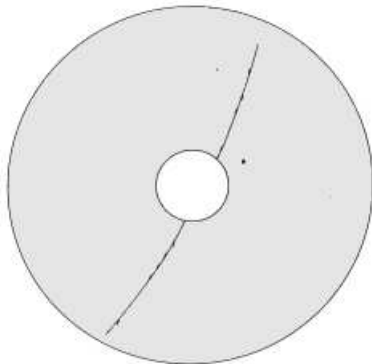
(b) An artist's interpretation of the SVT.

Figure 1.6: The Silicon Vertex Tracker

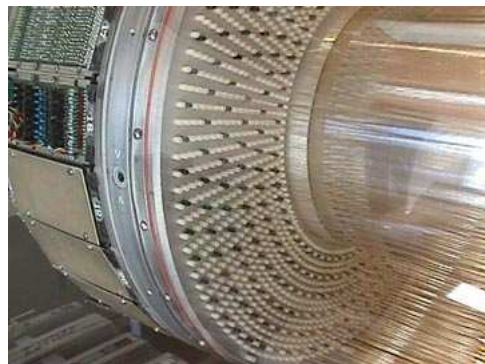
The inner most layer of the *BABAR* detector is the Silicon Vertex Tracker (SVT) whose main purpose is to reconstruct the decay vertex of the primary  $B$  decay. It is shown in Figure 1.6. The SVT also provides precision measurements of the

coordinates of charged tracks whose parents are particles other than the  $B$  or  $\bar{B}$ . While the Drift Chamber is primarily responsible for reconstructing charged tracks, particles whose momenta is below 100 MeV/ $c$  will not reach the Drift Chamber; further, the Drift Chamber reaches full efficiency only for tracks whose initial momenta is greater than 180 MeV/ $c$ . Thus, for tracks with momentum below this threshold, data from the SVT becomes important in efficient reconstructions of charged tracks. The SVT provides the most accurate angular coordinate information as multiple scatterings rarely occur within its volume. Finally, the SVT is the only source of data for particles that decay within its volume. The most notable of these particles is the  $K_s^0$ , a particle that is often in the final state of  $CP$  eigenstate decays (as the  $K_s^0$  is its own antiparticle).

### 1.3.2 Drift Chamber



(a) Two reconstructed tracks in the drift chamber.



(b) A photograph of the drift chamber and its precisely strung array of wires.

Figure 1.7: The Drift Chamber

The Drift Chamber (DCH), shown in Figure 1.7, is the second layer of the *BABAR* detector and is responsible for providing high precision measurements of

spacial coordinates while having a very short dead time. The chamber is filled with an inert, non-massive gas (a mixture of Helium and Isobutane in a 4:1 ratio) in order to minimize the chances of inducing a decay. In order to achieve the necessary high precisions—the DCH was designed to have a resolution better than  $140\ \mu\text{m}$ —the volume is filled with gold plated tungsten-rhenium wires of radius  $20\ \mu\text{m}$  and gold plated aluminum wires with radii of  $120\ \mu\text{m}$  and  $80\ \mu\text{m}$ . This complicated lattice of wiring (shown in Figure 1.7(b)) allows for up to forty independent measurements of the particle’s spacial coordinates. Multiple scatterings by low momentum particles limit the precision of the DCH, hence for it to operate at full efficiency those conditions discussed in Section 1.3.1 must be satisfied, else, the precision of the coordinate measurements is largely determined by data collected in the SVT.

### 1.3.3 Detector of Internally Reflected Cherenkov Light

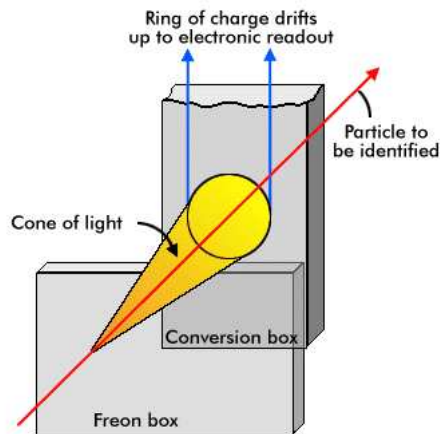


Figure 1.8: The DIRC operates by converting the emitted Cherenkov light into a ring of charge; the radius of which is a function of the particle’s speed.

The Detector of Internally Reflected Cherenkov Light, or the DIRC, uses a design that was first implemented in the *BABAR* detector. The DIRC's detection methods are based on the phenomenon of Cherenkov Light and it consists of 144 long, straight bars of synthetic quartz arranged in a 12-sided polygon barrel at the end of which is an array of photomultiplier tubes.[10] The detection method is shown in Figure 1.8. Cherenkov radiation is emitted by a particle when that particle's speed,  $v$ , satisfies the inequality given by Equation (1.3), where  $n$  is the refractive index of the material through which the particle is traveling.

$$v > \frac{c}{n} \tag{1.3}$$

In words, Cherenkov radiation is emitted when a particle's speed exceeds that of light within a given media. While the idea of using Cherenkov radiation in particle detection is not new, the way in which it is analyzed by the DIRC is. The DIRC inverts the traditional concept of ring-imaging Cherenkov counters (RICH counters) in that it relies on the detection of Cherenkov photons trapped in the radiator due to total internal reflection.[7, 13] The advantages of this new system include a reduction in the amount of material necessary to build the detector as well as a reduction in the volume of the detector which further allows for a relatively large Drift Chamber. The main purpose of the DIRC is particle identification and it is able to distinguish between pions and kaons with momenta up to 4.0 GeV/ $c$ . This identification process is useful for tagging purposes and particle identification in rare  $B$  decays.



Figure 1.9: A photograph of the EMC

### 1.3.4 Electromagnetic Calorimeter

Pictured in Figure 1.9 is the The Electromagnetic Calorimeter (EMC) whose design is based on quasi-projective CsI(Tl) crystals.[2] The crystals are aimed at the interaction point and are used to detect emitted photons. While the measurement of the energies of the  $\gamma$ 's is very good (see Equation 1.4) no information pertaining to the location of the decay vertex is available directly from the EMC.[10]

$$\frac{\sigma_E}{E} = \frac{1\%}{\sqrt[4]{E(\text{GeV})}} \oplus 1.2\% \quad (1.4)$$

To link photons to their decay vertices, information about “missing” energy and momenta must be used. As previously mentioned, the EMC is a key element in the decay mode  $B^0 \rightarrow \pi^0\pi^0\pi^0$  since  $\pi^0 \rightarrow \gamma\gamma$ . Thus the initial reconstruction of this  $B$  decay searches for events with at least six  $\gamma$ 's: pairs of which add to the total energy of one of the three  $\pi^0$  candidates. This will ultimately lead to a large amount of combinatorial background and is discussed in Section 3.2.3. Nevertheless, the high energy precision allows for an efficient  $\pi^0$  reconstruction with good mass resolution (as the resolution of the relativistic mass is closely related to the resolution of the

relativistic energy).<sup>3</sup>

### 1.3.5 Muon and Neutral Hadron Detector

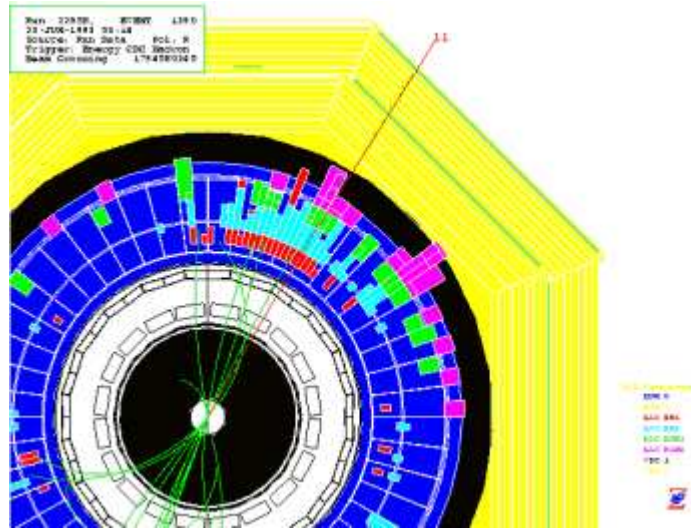


Figure 1.10: A  $\mu^-$  track in reconstructed using data collected in the Muon and Neutral Hadron Detector which is the yellow, outermost section pictured above.

The outermost layer of the *BABAR* detector is the Muon and Neutral Hadron Detector and is shown in Figure 1.10. While the  $\pi^0$  is a neutral hadron, it is created with insufficient momentum (given its life time) to travel all the way from the  $B^0$  decay vertex to the Muon and Neutral Hadron Detector and hence data from this sub-detector is unlikely to be used in this analysis.

Essentially, the job of the Muon and Neutral Hadron Detector is to identify all tracks have not decayed within the detector and have not yet been identified. The detection system is called the Instrumental Flux Return (IFR) and consists of

---

<sup>3</sup>For more information on  $\pi^0$  and  $\gamma$  identification see pages 144–145 of The *BABAR* Physics Book.[10]

a central barrel with two end caps (a soup can, more or less). As particles traverse this heavy iron casing, they leave energy deposits. By examining the positions of the energy deposits and the amount of energy deposited, the IFR can determine the identity of the particle. In terms of *CP*-Violation, its most important job is to detect  $\mu^-$ 's (as shown in Figure 1.10)—which are recognized as charged tracks penetrating deep into the iron core—and  $K_L^0$ 's.

# C H A P T E R 2

## Motivation

The study of the decay mode  $B^0 \rightarrow \pi^0\pi^0\pi^0$  was mainly inspired by the goal of *BABAR*: conducting *CP*-Violation experiments. The mode  $B^0 \rightarrow \pi^0\pi^0\pi^0$  is a *CP* eigenstate with an unmeasured branching ratio. The branching ratio (abbreviated  $\mathcal{B}(B^0 \rightarrow \pi^0\pi^0\pi^0)$ ) is defined as the ratio of the number of  $B^0 \rightarrow \pi^0\pi^0\pi^0$  decays to the total number of  $B^0$  decays as the total number of  $B^0$  decays becomes arbitrarily large, as is shown in Equation 2.1.

$$\mathcal{B}(B^0 \rightarrow \pi^0\pi^0\pi^0) \equiv \lim_{\mathcal{N}_{tot} \rightarrow \infty} \left( \frac{\mathcal{N}_{B^0 \rightarrow \pi^0\pi^0\pi^0}}{\mathcal{N}_{tot}} \right) \quad (2.1)$$

Hence, the motivation to measure the branching ratio is so that, in future, this mode may be used in *CP*-Violation experiments.

## 2.1 Eigenstates, Eigenfunctions, and Eigenvalue Equations

### 2.1.1 Eigenvalue Problems

Eigenvalue problems are common in physics and show up in many of the sub-disciplines. Eigenvalue problems are of the form:

$$\hat{\mathcal{O}}\Psi = \lambda\Psi \quad (2.2)$$

where  $\hat{O}$  is some operator that acts on the function  $\Psi$ , and  $\lambda$  can be either a real or complex constant, depending on the vector space. Any function,  $\Psi$ , that satisfies Equation 2.2 is said to be an *eigenfunction* of the operator  $\hat{O}$  with *eigenvalue*  $\lambda$ . The most obvious example of an eigenvalue problem in physics is the Time Independent Schrödinger Equation shown below.

$$\begin{aligned}
 \frac{p^2}{2m}\Psi + V\Psi &= E\Psi \\
 \frac{-\hbar^2}{2m}\nabla^2\Psi + V\Psi &= E\Psi \\
 \left(\frac{-\hbar^2}{2m}\nabla^2 + V\right)\Psi &= E\Psi \\
 \hat{H}\Psi &= E\Psi
 \end{aligned}
 \tag{2.3}$$

Often times, Equation 2.3 is written in terms of state kets rather than wavefunctions, in which case Equation 2.3 is written as below.

$$\hat{H}|\Psi\rangle = E|\Psi\rangle
 \tag{2.4}$$

When eigenvalue equations are written in this notation, the state ket  $|\Psi\rangle$  is said to be an *eigenstate* of  $\hat{H}$  with eigenvalue  $E$ .

### 2.1.2 The *CP* Operator

In physics, many of the operators seen in eigenvalue problems represent physical observables. For example, in Section 2.1.1 the Time Independent Schrödinger Equation is seen to contain an operator  $\hat{H}$ . This operator is the Hamiltonian and when it acts on a function or state ket, it returns the total energy of the system being described. This is represented by the eigenvalue  $E$  on the right hand side of Equation 2.3. Hence,  $\hat{H}$  is an operator that represents the total energy of a system.

Other times in physics, operators are not directly related to physical observables but are instead used to alter or perturb the state of the system in some way. For example, the creation and annihilation operators ( $a^\dagger$  and  $a$  respectively) for the quantum harmonic oscillator either bump the system up to the next excited state (creation) or lower it to the next less excited state (annihilation). Thus, the effects of  $a^\dagger$  and  $a$  are seen below.

$$a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle \quad (2.5)$$

$$a|n\rangle = \sqrt{n}|n-1\rangle \quad (2.6)$$

In Equation 2.5 and Equation 2.6 the *quantum number*  $n$  is used to denote the state of the system. Quantum numbers are used to describe certain properties of a system (such as energy or angular momentum). In Equation 2.5 and Equation 2.6,  $n$  is used to describe the excitation level of the system and hence, the system's energy is a function of  $n$ .<sup>1</sup> Quantum operators often return or change certain quantum numbers; for example, Equation 2.5 shows  $a^\dagger$  raising the quantum number  $n$  by one, Equation 2.7 shows  $J_z$  returning the quantum number  $m_j$ <sup>2</sup>, and Equation 2.8 shows  $J_+$  increasing the quantum number  $m_j$  by one.

$$J_z|j, m_j\rangle = m_j|j, m_j\rangle \quad (2.7)$$

$$J_+|j, m_j\rangle = \sqrt{j(j+1) - m_j(m_j+1)}|j, m_j+1\rangle \quad (2.8)$$

The  $CP$  operator is a combination of two operators: charge ( $C$ ) and parity ( $P$ ). The charge operator acting on some state  $|\Psi\rangle$  switches the electric charge of all the particles being described by  $|\Psi\rangle$ . Therefore,  $C$  acting on a particle  $k$  changes that particle into its antiparticle  $\bar{k}$  (see Figure 2.1). In terms of quantum numbers,

---

<sup>1</sup>In the case of the one-dimensional quantum harmonic oscillator:  $E(n) = (n + \frac{1}{2}) \hbar\omega$

<sup>2</sup> $m_j$  is a quantum number used to describe the projection of the total angular momentum onto a given axis (customarily, the axis is taken as the  $\hat{e}_z$  axis).

$C$  switches the sign of the charge quantum number. Further, the eigenvalue problem given by Equation 2.9 is satisfied by any group of particles, given by  $|\Psi\rangle$ , whose *potential* members  $k_i$ , satisfy either of the conditions given in Table 1.

$$C|\Psi\rangle = |\Psi\rangle \quad (2.9)$$

The parity operator,  $P$ , acting on some state  $|\Phi\rangle$  switches the parity of all the

| Condition Number | Description  |
|------------------|--|
| 1                | $k_i \in  \Psi\rangle$ if $k_i = k_i$ .  |
| 2                | $k_i \in  \Psi\rangle$ if $\forall k_i \exists k_j \in  \Psi\rangle \ni k_i = k_j$ . |

Table 1: Qualitative Solutions to the  $C$  Eigenvalue Problem

particles being described by  $|\Phi\rangle$  (see Figure 2.1). In other words, when  $P$  acts on a system  $|\Phi\rangle$ , the system returned,  $|\tilde{\Phi}\rangle$ , is a mirror image of the original system. In terms of quantum numbers,  $P$  switches the sign of the quantum number that describes the the projection of a particle's intrinsic angular momentum, or spin, onto a specific axis<sup>3</sup>. This quantum number is hereafter referred to as the spin component quantum number,  $m_s$ . The eigenvalue problem given by Equation 2.10 is satisfied by any group of particles, given by  $|\Phi\rangle$ , whose *potential* members  $|k, m_s\rangle$ , where  $k$  denotes the type of particle / all other quantum numbers, and  $m_s$  is the spin component quantum number, satisfy either of the conditions given in Table 2.

$$P|\Psi\rangle = |\Psi\rangle \quad (2.10)$$

---

<sup>3</sup>Again, this component is customarily taken to be the component lying along  $\hat{e}_z$ . Further, since the  $\hat{e}_z$  is free to be any direction, it is customary to let this vector lie parallel the particle's momentum vector thus the action of  $P$  on a system is easily defined in terms of quantum numbers.

| Condition Number | Description  |
|------------------|--|
| 1                | $ k, m_s\rangle \in  \Phi\rangle$ if $s = 0$ .   |
| 2                | $ k, m_s\rangle \in  \Phi\rangle$ if $\forall  k, m_s\rangle \exists  k, -m_s\rangle \in  \Phi\rangle$ . |

Table 2: Qualitative Solutions to the  $P$  Eigenvalue Problem

When charge and parity are combined to form a new operator,  $CP$ , the result is the sum of the individual components. Thus when the  $CP$  operator acts on some state  $|\Psi\rangle$ , both the electric charge and the parity of all the particles described are turned into their opposites. Thus, this process turns a left-handed particle  $k$  into a right-handed antiparticle  $\bar{k}$  (see Figure 2.1). In the language of quantum numbers,  $CP$  switches the signs of the charge quantum number and the spin component quantum number.

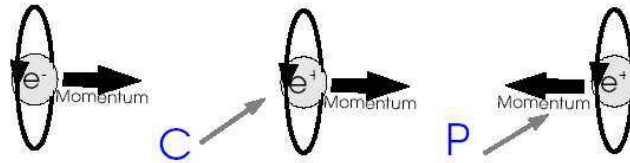


Figure 2.1: When  $C$  acts on an electron, the sign of the electric charge switches and the electron becomes a positron, the electron's antiparticle. When  $P$  acts on a positron its handedness is switched, the positron's spin, relative to its momentum (its  $\hat{e}_z$  component) is reversed.

### 2.1.3 $CP$ Eigenstates

In order to find  $CP$  eigenstates the eigenvalue problem given by Equation 2.11 must be solved.

$$CP|\Psi\rangle = |\Psi\rangle \quad (2.11)$$

Solutions to Equation 2.11, given by  $|\Psi\rangle$ , which describes a group of particles whose *potential* members are given by  $k_i$  whose state kets are  $|k, m_s\rangle$  where  $k$  denotes the type of particle / all other quantum numbers, and  $m_s$  is the spin component quantum number, must satisfy both condition (1) and condition (2) (using either (a) or (b)) from Table 3. Groups satisfying conditions labeled (a) must be made

| Condition Number |   | Description  |
|------------------|---|--|
| 1                | a | $k_i \in  \Psi\rangle$ if $k_i = k_i$ .  |
| 1                | b | $k_i \in  \Psi\rangle$ if $\forall k_i \exists k_j \in  \Psi\rangle \ni k_i = k_j$ .                     |
| 2                | a | $ k, m_s\rangle \in  \Phi\rangle$ if $m_s = 0$ .   |
| 2                | b | $ k, m_s\rangle \in  \Phi\rangle$ if $\forall  k, m_s\rangle \exists  k, -m_s\rangle \in  \Phi\rangle$ . |

Table 3: Qualitative Solutions to the  $CP$  Eigenvalue Problem

entirely of mesons since these particles must have spin zero<sup>4</sup> and must be their own antiparticles; whereas groups satisfying conditions labeled (b) may contain all types of hadrons and leptons. Both groups, however, have at least two common characteristics, that is, the net electric charge of the collection *must* be zero and the net spin component along  $\hat{e}_z$  *must* be zero.

---

<sup>4</sup>If a particle has a total spin given by quantum number  $s$ , then  $m_s$  may take the following values:  $\{-s, -s + 1, \dots, 0, \dots, s - 1, s\}$

## The State $\pi^0\pi^0\pi^0$

To determine whether or not  $\pi^0\pi^0\pi^0$  is a  $CP$  eigenstate, Equation 2.11 must be considered where  $|\Psi\rangle$  now describes the state  $\pi^0\pi^0\pi^0$ . Since this state does not *seem* to contain any  $\bar{\pi}^0$  particles, the only way  $\pi^0\pi^0\pi^0$  can be a  $CP$  eigenstate is if it satisfies condition (1a) from Table 3, namely,  $\pi^0$  is the same as  $\bar{\pi}^0$ . As for condition (2), the  $\pi^0$  is a spin zero ( $s = 0 \Rightarrow m_s = 0$ ) particle and thus satisfies condition (2a). To determine whether or not the  $\pi^0$  is its own antiparticle, the quark content of the  $\pi^0$  must be considered. The  $\pi^0$  is a superposition of  $u$ ,  $d$ ,  $\bar{u}$ , and  $\bar{d}$  quarks given by Equation 2.12.

$$\pi^0 = \frac{u\bar{u} + d\bar{d}}{\sqrt{2}} \quad (2.12)$$

Acting  $CP$  on Equation 2.12 yields:

$$\begin{aligned} CP(\pi^0) &= \bar{\pi}^0 \\ &= CP\left(\frac{u\bar{u} + d\bar{d}}{\sqrt{2}}\right) \\ &= \frac{\bar{u}u + \bar{d}d}{\sqrt{2}} \\ &= \pi^0 \end{aligned} \quad (2.13)$$

Thus, condition (1a) from Table 3 is satisfied since the  $\pi^0$  is its own antiparticle and thus the state  $\pi^0\pi^0\pi^0$  is a  $CP$  eigenstate.

## 2.2 $CP$ Violation in a Matter Dominated Universe

The universe contains a startling asymmetry in the amount of matter and antimatter that exists and yet, at its creation, it is believed an equal amount of both was formed. This observation has led theorists to believe some asymmetry must exist on the most fundamental level within the laws of nature. According to a

theorem proved in 1917 by the German mathematician Emmy Noether, every symmetry implies the existence of a conserved quantity.[12] For example, the rotational invariance of space-time leads to the conservation of angular momentum. Thus, by Noether's Theorem, the lack of symmetry between matter and antimatter should imply the existence of a quantity that is *not* conserved. That is to say, there should exist some *approximate* symmetry of nature, or a physically observable quantity that is *approximately* conserved. Experiments conducted in 1957 by Chien-Shiung Wu and her collaborators that showed only right-handed particles can decay via the weak force made charge parity a good candidate for this asymmetry; however, at the time, the theory of the weak force did not allow for  $CP$  violation. In 1964, suspicions that weak decays do not conserve  $CP$  were confirmed when researchers at Brookhaven National Laboratory observed  $CP$  violation in  $K_L^0$ 's. Nevertheless, the theory of the weak force still did not allow for  $CP$  violation; that is, until 1972 when Makoto Kobayashi and Toshihide Maskawa showed the weak force could violate  $CP$  *if* a third generation of quarks existed (we now know that a third generation does indeed exist).[12] This third generation would allow for elements of the Cabibbo-Kobayashi-Maskawa ( $CKM$ ) matrix (shown in Equation 2.14) to be complex and these complex entries would allow for quantum interference terms in the probability amplitudes of weak decays.[5] With this discovery, physicists began to ask whether or not the amount of  $CP$  violation allowed by the Standard Model was enough to account for the startling asymmetry observed.

$$\begin{aligned}
V &= \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \\
&= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} \\
&\quad c_{ij} \equiv \cos(\theta_{ij}), \quad s_{ij} \equiv \sin(\theta_{ij})
\end{aligned} \tag{2.14}$$

Three decades later, it is agreed the answer is no, not enough  $CP$  violation occurs within the Standard Model to allow for such a vast imbalance. As such, experimentalists continue to search for more  $CP$  violation while theorists have been proposing new theories with alternate sources of  $CP$  violation to account for the aforementioned imbalance (such as super-symmetry (SUSY) though  $CP$  violation is not the primary motivation for super-symmetric theories).

The un-conserved quantity related to  $CP$  violation in the electroweak force is  $CP$  number. The  $CP$  number of a system is the quantum number that describes the  $CP$  state of that system, just as the quantum number  $\ell$  is used to describe the orbital angular momentum of hydrogen-like systems. It has been observed and theoretically shown, albeit after the observations were made, weak decays do not conserve  $CP$  number, and so it is worth investigating what possible values a  $CP$  number can take. To do so, consider  $CP$  acting twice on some  $CP$  eigenstate,  $|\Phi\rangle$ ,

as in Equation 2.15.

$$\begin{aligned}
CPCP(|\Phi\rangle) &= CP^2(|\Phi\rangle) \\
&= CP(CP(|\Phi\rangle)) \\
&= CP\lambda(|\bar{\Phi}\rangle) \\
&= \lambda^2|\Phi\rangle
\end{aligned}
\tag{2.15}$$

Since  $CP$  acting on a  $CP$  eigenstate,  $|\Phi\rangle$ , takes that state into a state of its antiparticles,  $|\bar{\Phi}\rangle$ , then  $CP$  acting twice on a system,  $|\Phi\rangle$ , must return that system to its original state. Therefore,  $\lambda^2$  *must* equal unity and  $\lambda$  may take only the values  $\pm 1$ . Thus, weak decays can take systems of particles with  $CP$  number +1 into systems with  $CP$  number -1 and vice-versa. The only “problem” is that current theories predict that decays that do not conserve  $CP$  number should be quite rare, that is, not occurring often enough to account for a matter dominated universe.

The goal of  $CP$  experiments is to determine *the extent to which*  $CP$  violation occurs and whether or not that amount matches the predictions of the Standard Model. In performing these measurements, physicists look for asymmetries in the decay rates of  $CP$  eigenstates, that is, differences in the rate of decay for  $B^0 \rightarrow \pi^0\pi^0\pi^0$  and  $\bar{B}^0 \rightarrow \pi^0\pi^0\pi^0$ . This particular form of  $CP$  violation is known as *direct*  $CP$  violation and the difference in rate is predicted to be small as compared to either of the rates themselves. Experiments of this type of have already been conducted where  $B^0$  and  $\bar{B}^0$  decay to pions and kaons, and results have shown that 13% more  $B^0$  mesons decay into kaons and pions than  $\bar{B}^0$  mesons.[14] In order to measure the asymmetry in the rate, the branching ratio must first be measured, and thus is the motivation for measuring the  $B^0 \rightarrow \pi^0\pi^0\pi^0$  branching ratio: so that direct  $CP$  violation experimnts can be conducted using this mode.

## 2.3 Theory

The decay being considered,  $B^0 \rightarrow \pi^0\pi^0\pi^0$ , is a charmless three-body decay. That is to say, the final state contains three particles and the charm quark does not enter into the initial state, the final state, or any of the intermediate states of the decay. The primary (lowest order) Feynman Diagram for the decay is shown in Figure 2.2(a). The primary diagram for the decay  $\bar{B}^0 \rightarrow \pi^0\pi^0\pi^0$  is shown in Figure 2.2(b). Figure 2.2 shows that the  $B^0$  and the  $\bar{B}^0$  decay via the same diagram,

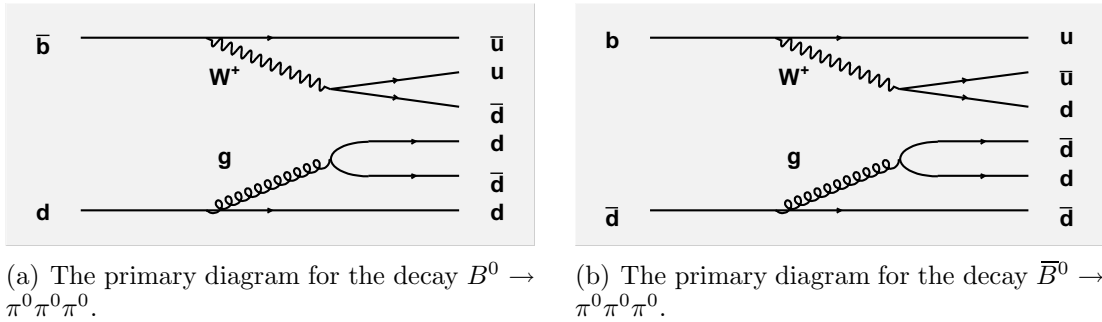


Figure 2.2: Feynman Diagrams for the Decays  $B^0/\bar{B}^0 \rightarrow \pi^0\pi^0\pi^0$

hence we should expect a symmetry in the branching ratios; however, as discussed in Section 2.2, weak decays (depicted as the emission of a virtual  $W^+$ ) do not conserve  $CP$  number and thus a small asymmetry is ultimately expected. In other words, the probability of the emission of a virtual<sup>5</sup>  $W^+$  differs depending on whether the  $W^+$  is emitted from a right-handed  $b$ , a left-handed  $b$ , a right-handed  $\bar{b}$ , or a left-handed  $\bar{b}$  quark, because the interaction between the  $W^\pm$  (a weak force mediating boson) and the  $b/\bar{b}$  quark is  $CP$  number dependent.

---

<sup>5</sup>The  $W^+$  emission is said to be virtual because the  $W^+$  boson cannot be detected. Nevertheless, to wind up with the correct final state, a  $W^+$  is need to decay into a  $u \bar{d}$  pair. Thus the  $W^+$  is said to be virtual as it is never detected but it must have existed, in some sense of the word, in order to get the  $u \bar{d}$  pair in the final state.

After reviewing the literature, no phenomenology papers have been written about this particular decay and there has been no experimental work done to date. Therefore, in order to get some idea of what type of results to expect and what to use for various parameters in the Monte Carlo studies, comparisons will be made between this decay and other similar decays.

### 2.3.1 Comparisons To $B^\pm$ Decays

The first decays that are considered are the decays  $B^\pm \rightarrow \pi^\pm \pi^\mp \pi^\pm$ . Both experimental and theoretical work has been done on this decay mode. Phenomenologists have considered the non-resonant mode  $B^- \rightarrow \pi^- \pi^+ \pi^-$  and have come up with two different predictions for this branching ratio based on two different treatments<sup>6</sup> of the value  $g_B B_s^* K(\pi)$ . [6] The results of this study are shown in Table 4. Experimentalists at *BABAR* have also considered this mode and have set an

| Mode   | $\mathcal{B}_{theory}^1$ | $\mathcal{B}_{theory}^2$ |
|--|--------------------------|--------------------------|
| $B^- \rightarrow \pi^- \pi^+ \pi^-$ (non-resonant) | $3.0 \times 10^{-6}$     | $1.7 \times 10^{-6}$     |

Table 4: Theoretical Predictions for  $\mathcal{B}(B^- \rightarrow \pi^- \pi^+ \pi^-)$

upper limit on the non-resonant branching ratio at a 90% confidence level as well as have determined a most probable value for  $\mathcal{B}(B^- \rightarrow \pi^- \pi^+ \pi^-)$ . [1] The results<sup>7</sup> are given in Table 5. All of the research agrees that the order of magnitude for the non-resonant decay is  $10^{-6}$ . The experimental best value lies right in between the

---

<sup>6</sup>In Table 4, the prediction  $\mathcal{B}_{theory}^1$  is made for  $g_B B_s^* K(\pi) = 2g/f_\pi \times (m_B m_{B_s^*})^{1/2}$  while  $\mathcal{B}_{theory}^2$  accounts for the off-shellness of the  $B_s^*$  by letting  $g_B B_s^* K(\pi) = 2g/f_\pi \times (m_B \sqrt{p_{B_s^*}^2})^{1/2}$ . [6]

<sup>7</sup>The first uncertainty is statistical, the second is systematic, while the third is model-dependent. [1]

| Mode   | $\mathcal{B}(B^\pm \rightarrow \text{Mode})(10^{-6})$ | 90% CL UL $\mathcal{B}(10^{-6})$ |
|--|---|----------------------------------|
| $B^\pm \rightarrow \pi^\pm \pi^\mp \pi^\pm$ (Total)        | $16.2 \pm 1.0 \pm 1.2 \pm 0.9$                        | —                                |
| $B^\pm \rightarrow \pi^\pm \pi^\mp \pi^\pm$ (non-resonant) | $2.3 \pm 0.9 \pm 0.3 \pm 0.4$                         | $< 4.6$                          |

Table 5: Experimental Results on Measurements of  $\mathcal{B}(B^\pm \rightarrow \pi^\pm \pi^\mp \pi^\pm)$

two theoretical predictions, and both of the theoretical predictions are consistent with the upper limit value.

The primary non-resonant diagrams for these decays are shown in Figure 2.3. These decays are also three-body charmless decays and the only difference between

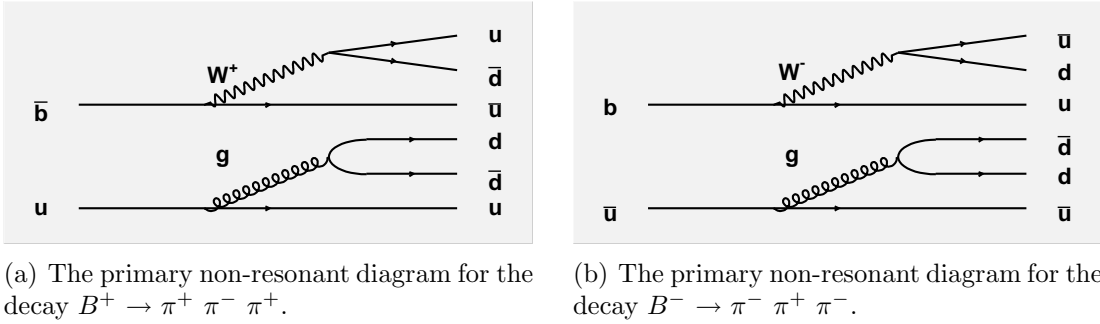


Figure 2.3: Feynman Diagrams for the Decays  $B^\pm \rightarrow \pi^\pm \pi^\mp \pi^\pm$

their diagrams and the primary diagram for  $B^0 \rightarrow \pi^0 \pi^0 \pi^0$  is the emission of an external  $W^\pm$  rather than an internal  $W^\pm$ . Therefore, to a lowest order approximation, it can be assumed that the difference in the theoretical calculation of these branching ratios and therefore the only difference in the values of the branching ratios is the result of the *type* of virtual  $W^\pm$  emission. The fact that the decay being studied requires an *internal*  $W^\pm$  decay leads to a phenomenon known as color suppression.

## Color Suppression

Color suppression is the result of a  $W^\pm$  decaying to a  $q\bar{q}$  pair that is split up and paired with other quarks in the final state of the decay. Consider Figure 2.2. The  $q\bar{q}$  pair created in the internal  $W^+$  decay is split and paired with a different anti-quark and quark in the final state. Quarks, as discussed in Section 1.1, carry color charge and interact via the Strong Force according to this charge. For a bound state of quarks to be allowed, the state must be colorless and thus the  $b$  and the  $\bar{d}$  in the initial state must carry opposite color charge (red and anti-red for example). The  $W^+$  is a gauge boson associated with weak decays and carries no information about color; therefore, when it decays into a  $q\bar{q}$  pair, while this pair is always colorless, it has equal probability of decaying into a red, anti-red pair a green, anti-green pair, or a blue, anti-blue pair. In other words, the probability of decaying into the same color and anti-color as the quarks bound as the initial  $B^0$  meson is one third. In calculating the branching ratio, the one third gets squared and enters as a factor of  $\frac{1}{9}$ . Since the  $d\bar{d}$  pair is created by a virtual gluon (a color-carrying particle that mediates the Strong Force) emitted from the  $\bar{d}$  quark, the  $d\bar{d}$  pair emitted from the gluon will carry the same color charge as the initial  $\bar{d}$  quark from the  $B^0$  bound state and thus this part of the decay is not color suppressed. Now consider the  $B^\pm$  decays in Figure 2.3. The  $q\bar{q}$  pair emitted by the  $W^\pm$  is paired with itself in the final state as a  $\pi^\pm$  and thus the color of the  $q$  is unimportant since the  $\bar{q}$  is guaranteed to carry the associated anti-color for the  $W^\pm$  decay to be allowed. Thus, if the ratio of the branching ratios of these two decays is considered, and it is assumed that the similar diagrams will lead to similar mathematical expressions that cancel against one another when the ratio is taken, only a ratio of terms describing the probabilities of emitting an external versus internal  $W^\pm$  will remain. This is seen

in Equation 2.16. Hence the result of color suppression is seen to be  $\frac{1}{9}$ —very nearly an order of magnitude.

$$\frac{\mathcal{B}(B^0 \rightarrow \pi^0\pi^0\pi^0)}{\mathcal{B}(B^+ \rightarrow \pi^+\pi^-\pi^+)} \approx \left( \frac{\mathcal{P}(W_{internal}^\pm)}{\mathcal{P}(W_{external}^\pm)} \right)^2 \approx \frac{1}{9} \quad (2.16)$$

Considering the results of the  $B^\pm \rightarrow \pi^\pm \pi^\mp \pi^\pm$  study from Section 2.3.1 and Equation 2.16, a branching ratio of  $2.3 \times 10^{-7}$  is expected for the decay  $B^0 \rightarrow \pi^0\pi^0\pi^0$ . Since this comparison technique is hardly rigorous, it is wise to consider only the order of magnitude of the prediction and thus,  $\mathcal{B}(B^0 \rightarrow \pi^0\pi^0\pi^0)$  is expected to be of the order  $10^{-7}$ .

### 2.3.2 Comparisons to Other $B^0$ Decay Modes

Now, in order to check the results of Section 2.3.1 (in a self-consistent manner) the decay  $B^0 \rightarrow D^- \pi^+$  is considered. Further, in considering this decay, it will be shown that  $B^0 \rightarrow \pi^0\pi^0\pi^0$  also experiences  $CKM$  suppression, a phenomenon discussed later in this section. The primary diagram for  $B^0 \rightarrow D^- \pi^+$  is shown in Figure 2.4. Similar to the mode  $B^\pm \rightarrow \pi^\pm \pi^\mp \pi^\pm$ , the mode  $B^0 \rightarrow D^- \pi^+$  involves the emission of an external  $W^+$ . Thus, the branching ratio of the mode  $B^0 \rightarrow \pi^0\pi^0\pi^0$  is suppressed by *at least* a factor of  $\frac{1}{9}$  relative to the mode  $B^0 \rightarrow D^- \pi^+$ . There are two other differences; however, between this mode and the mode  $B^0 \rightarrow \pi^0\pi^0\pi^0$ . First, this decay is a two-body decay rather than a three-body decay, and second, the final state contains a charm anti-quark (which is bound with a down quark in the final state to make the  $D^-$ ). While it is true, this mode does not emit a gluon from the  $d$  quark to create a  $q\bar{q}$  pair, this difference is less mathematically significant than the color suppression and the presence of the  $c$  quark and thus, for

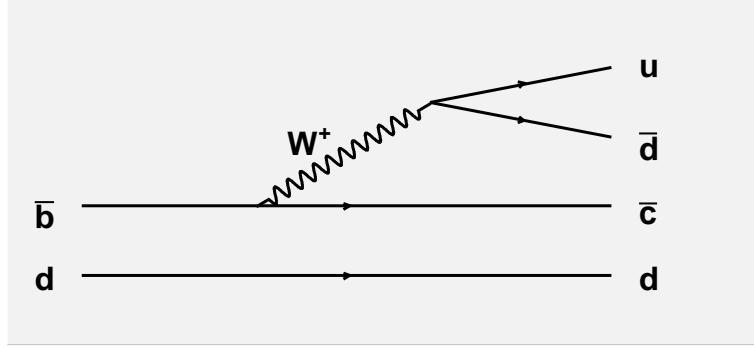


Figure 2.4: Feynman Diagram for the non-resonant Decay  $B^0 \rightarrow \pi^+ D^-$

the purposes of this naive estimate, that difference will be ignored. Additionally, the difference in the number of particles in the final state will be neglected as it, too, is expected to be less significant than color suppression and the presence of  $c$  quark. In Section 2.3.1, the probability of the  $\bar{b}$  quark decaying into a virtual  $W^\pm$  and a  $\bar{u}$  quark was not considered since both modes featured this decay; and thus, when the ratio of the branching ratios was taken, the contributions canceled against one and other.<sup>8</sup> However, in the decay  $B^0 \rightarrow D^- \pi^+$ , a  $\bar{b}$  quark decays into a  $W^+$  and a  $\bar{c}$ , whereas in the mode  $B^0 \rightarrow \pi^0 \pi^0 \pi^0$ , a  $\bar{b}$  quark decays into a  $W^+$  and a  $\bar{u}$ . The rates at which these different decays occur are functions of different  $CKM$  matrix elements (see Figure 2.14) and thus, the decays have different probability amplitudes. This difference leads to a phenomenon referred to as  $CKM$  suppression.

## $CKM$ Suppression

Consider the two Weak decays,  $\bar{b} \rightarrow W^+ \bar{u}$  and  $\bar{b} \rightarrow W^+ \bar{c}$ . The rates of these weak decays are *proportional* to the square of their associated  $CKM$  matrix

---

<sup>8</sup>Whether the mode involves a  $b$  decaying into a  $u$  or a  $\bar{b}$  decaying into a  $\bar{u}$  is, from a mathematical standpoint, irrelevant, as both decays are described using the same mathematical quantities

elements. That is, the former decay is proportional to  $|V_{ub}|^2$  and the latter is proportional to  $|V_{cb}|^2$ . Now, consider the ratio of the probabilities of each decay as in Equation 2.17 where it is assumed that the constants of proportionality (i.e. all of the factors affecting the decay rates other than the  $CKM$  matrix elements) are the same for both decays.

$$\frac{\mathcal{P}(\bar{b} \rightarrow W^+ \bar{u})}{\mathcal{P}(\bar{b} \rightarrow W^+ \bar{c})} = \left( \frac{|V_{ub}|}{|V_{cb}|} \right)^2 \approx 5 \times 10^{-3}. \quad (2.17)$$

Thus, from Equation 2.17 it is seen that the decay  $\bar{b} \rightarrow W^+ \bar{u}$  is suppressed by a factor of about 200 relative to the decay  $\bar{b} \rightarrow W^+ \bar{c}$ . Now, if the total suppression of  $B^0 \rightarrow \pi^0 \pi^0 \pi^0$  relative to  $B^0 \rightarrow D^- \pi^+$  is considered, the result is of the order  $10^{-4}$ , as seen in Equation 2.18.

$$\begin{aligned} \frac{\mathcal{B}(B^0 \rightarrow \pi^0 \pi^0 \pi^0)}{\mathcal{B}(B^0 \rightarrow D^- \pi^+)} &\approx \left( \frac{\mathcal{P}(W^+_{internal})}{\mathcal{P}(W^+_{external})} \right) \left( \frac{\mathcal{P}(\bar{b} \rightarrow W^+ u)}{\mathcal{P}(\bar{b} \rightarrow W^+ \bar{c})} \right) \\ &\approx \left( \frac{\mathcal{P}(W^+_{internal})}{\mathcal{P}(W^+_{external})} \right) \left( \frac{|V_{ub}|}{|V_{cb}|} \right)^2 \\ &\approx \left( \frac{1}{9} \right) (5 \times 10^{-3}) \\ &\approx 6 \times 10^{-4} \end{aligned} \quad (2.18)$$

The mode  $B^0 \rightarrow D^- \pi^+$  has been well studied and experiments have shown that  $\mathcal{B}(B^0 \rightarrow D^- \pi^+) = 3.7 \pm 1.5 \times 10^{-3}$ . [11] Thus, when the value of the decay rate for  $B^0 \rightarrow D^- \pi^+$  is considered with Equation 2.18, the expected branching ratio for  $B^0 \rightarrow \pi^0 \pi^0 \pi^0$  is of the order  $10^{-7}$ , which is consistent with the results obtained in Section 2.3.1.

Hence, by comparing the primary diagram for the mode  $B^0 \rightarrow \pi^0 \pi^0 \pi^0$  to similar decay diagrams with well known branching ratios,  $\mathcal{B}(B^0 \rightarrow \pi^0 \pi^0 \pi^0)$  is expected to be of the order  $10^{-7}$ . This estimate is self-consistent in that the same result was obtained investigating two independent decay modes. The first estimate

used both theoretical predictions and experimental data as well as the phenomenon of color suppression to arrive at a result while the latter used the world average for  $\mathcal{B}(B^0 \rightarrow D^- \pi^+)$  as well a *CKM* suppression argument. The order of magnitude estimates presented in Sections 2.3.1–2.3.2 will improve the accuracy of the Monte Carlo studies that are used in this analysis. These studies are discussed in Chapters 3–4.

## CHAPTER 3

### Monte Carlo Background Analysis

Before acquiring any of the *BABAR* data, let alone analyzing it, the analysis macros that will eventually examine the real data are tested and fine-tuned on *Monte Carlo* generated data sets (MC data). These data sets are produced using probability distribution functions (pdf's) and random number generators. The simulated data sets contain exactly the same information as the real data sets in addition to “truth” variables for all observables and a Monte Carlo particle ID for all tracks in a given event. The MC truth variables contain the true values of the observables that are calculated by the reconstruction algorithms while the MC particle ID's reveal the true identity of the particle (e.g.  $\pi^0$ ,  $\pi^+$ ,  $B^0$ , etc.). The Monte Carlo data is treated in the same way as *BABAR* data by the reconstruction algorithms, data acquisition modules, and analysis macros. By using data for which the “answer” is known, many different aspects of the analysis can be studied and the Monte Carlo truth variables can be used to estimate errors, and otherwise determine various sensitivities and efficiencies of the various analysis tools that will be used.

In this analysis, MC data is used in three main ways. First, the MC data is used to determine the optimal set of constraints to place on observables in or-

der to maximize the *significance* of the data set. The significance<sup>1</sup> is defined by Equation 3.1 and the aforementioned process is discussed in Section 3.3.3.

$$\mathcal{S} \equiv \frac{\mathcal{N}_{sig}}{\sqrt{\mathcal{N}_{bkg}}} \quad (3.1)$$

Second, by using the MC truth information and the particle ID information, it can be determined what percent of the total signal events remain after the cuts are applied to the data set. This percentage is called the signal efficiency. MC truth information further allows the significance of the data set to be determined during the optimization process. Finally, by analyzing the MC data set, a decision can be made whether an upper bound analysis or a direct measurement of the branching ratio should be performed using the *BABAR* data sets. MC data is also useful for debugging analysis macros, studying possible internal biases, and estimating how long the analysis macros will take to run over the data sets.

### 3.1 Building the Monte Carlo Data Set

While the beam energies at SLAC are tuned to the  $\Upsilon(4S)$  resonance so that the probability of producing a  $\Upsilon(4S) \rightarrow B\bar{B}$  decay is maximized, not every  $e^+ e^-$  collision results in a  $\Upsilon(4S)$ . As a matter of fact, only about one in every five  $e^+ e^-$  collisions result in a  $\Upsilon(4S)$ . When a  $\Upsilon(4S)$  is not produced, the decay  $e^+ e^- \rightarrow q\bar{q}$  ( $q\bar{q} \neq b\bar{b}$ )<sup>2</sup> is almost always seen. These types of decays are called *continuum* background events and are discussed in Section 3.2.1. Further, even if a  $\Upsilon(4S)$  is

---

<sup>1</sup>Signal is abbreviated as “sig” and background is abbreviated as “bkg”.

<sup>2</sup>Hereafter, it will be assumed, unless otherwise stated, that reactions labeled  $q\bar{q}$  are those with  $q \neq b$ .

produced, it does not always decay to a  $B^0\bar{B}^0$  pair; however since the  $\Upsilon(4S)$  rest mass is very nearly twice the  $B$  rest mass, it decays one-hundred percent of the time into a  $B^0\bar{B}^0$  pair *or* a  $B^+B^-$  pair, with *equal probability*. Therefore, in addition to losing about four out of every five events to continuum background, of the twenty percent remaining, half of the decays are  $B^+B^-$  decays. As a result, when the MC data set is constructed, it must contain proportionally accurate amounts of these types of decays (i.e. roughly one  $B$  event for every four  $q\bar{q}$  events and one  $B^+B^-$  event for every  $B^0\bar{B}^0$  event).

Before re-normalizing the data set to preserve these ratios, data for each type of decay is compiled into a master data set. Thus, data for  $e^+ e^- \rightarrow q\bar{q}$  decays,  $e^+ e^- \rightarrow \Upsilon(4S) \rightarrow B^0\bar{B}^0$  decays, and  $e^+ e^- \rightarrow \Upsilon(4S) \rightarrow B^+B^-$  decays are all necessary. The continuum background is separated into two data sets: one containing  $e^+ e^- \rightarrow c\bar{c}$  and the other containing  $e^+ e^- \rightarrow u\bar{u}/d\bar{d}/s\bar{s}$ . The  $c\bar{c}$  Monte Carlo data was acquired using the SP-1005 data set and the  $u\bar{u}/d\bar{d}/s\bar{s}$  data set was acquired using the SP-998 data set. Next, generic<sup>3</sup>  $B^+B^-$  decay data is needed. This is added to the collection using the SP-1235 data set. Finally, generic  $B^0\bar{B}^0$  data is needed. The SP-1237 data set is thus added to the rest of the collection to complete the Monte Carlo data set. Table 6 shows the contents of the Monte Carlo data set before any type of constraints or cuts are applied to it. It is noted that the ratios

|         | Parents        |             |             |                              |
|---------|----------------|-------------|-------------|------------------------------|
| Stage   | $B^0\bar{B}^0$ | $B^+B^-$    | $c\bar{c}$  | $u\bar{u}/d\bar{d}/s\bar{s}$ |
| Initial | 541 284 225    | 532 768 433 | 417 920 626 | 676 351 526                  |

Table 6: Initial Size of the Monte Carlo Data Set

---

<sup>3</sup>Here, generic means that the probability of decaying into a particular daughter state is set to the experimentally determined value or the theoretical prediction if no measurement of the decay mode has been made.

between the amounts of each type of data is incorrect; but, the data set is too large to handle at this stage anyway, thus, before the data set is renormalized, some loose cuts must be applied. These loose cuts, or *skims*, are discussed in Section 3.3.1. Even before these loose cuts are considered, the types of decays present in the data set and the expected properties of those decays are discussed in order to understand how best to apply background eliminating cuts to particular observables. That is, what observables can be constrained and what domain should the observable be constrained to, to best discriminate against background events whilst preserving as many signal events as possible.

## 3.2 Types of Background

Each type of background has its own “finger print”, so to speak. That is, the statistical distributions of each physical observable will take on a shape that is characteristic of the type of decay that was reconstructed to create the distributions. Since different decays display different distributions and, what is more, since signal distributions often peak in different regions than background distributions, by requiring the values of certain observables to lie within the region about the signal distribution’s maximum, much of the background can be eliminated from the data while preserving much of the signal. The observables that are used as cut variables in this analysis are:

- Cosine Theta-Thrust ( $\cos(\theta_{thrust})$ )
- Total Energy Difference ( $\Delta E$ )
- The Fisher Discriminant

- The  $\pi^0$  Invariant Mass ( $m_{\pi^0}$ )

These four observables were chosen for several reasons. First, the signal distribution is expected to differ significantly from at least one of the background distributions. Second, the observables are chosen such that when cuts are applied to all of the observables, all types of background will be discriminated against. Finally, these four variables are thought to be more or less independent of one and other. That is to say, cutting on all of the variables simultaneously will result in few redundancies and therefor, each restriction will remove a sizable and unique portion of the background contained within the data set.

$\cos(\theta_{thrust})$

The thrust axis of an event,  $\hat{T}$ , is defined to be the direction that maximizes the sum of the longitudinal momenta of the particles.[10] Geometrically, the thrust axis can be thought of as the best-fit linear regression drawn through the reconstructed coordinates of all of the particles for a given event, with the  $\chi^2$  of that fit weighted by the magnitude of the particle's momentum. The angle,  $\theta_{thrust}$ , is defined as the angle between the  $B^0$  thrust axis (i.e. the thrust axis calculated using tracks that are daughters, granddaughters, etc. of the  $B^0$  candidate) and the thrust axis for all other tracks in the event. Therefor, if  $\cos(\theta_{thrust})$  is close to  $\pm 1$ , the event's reconstruction lies mainly along a single axis. Figure 3.1 shows possible event reconstructions and discusses the properties of the thrust axis.

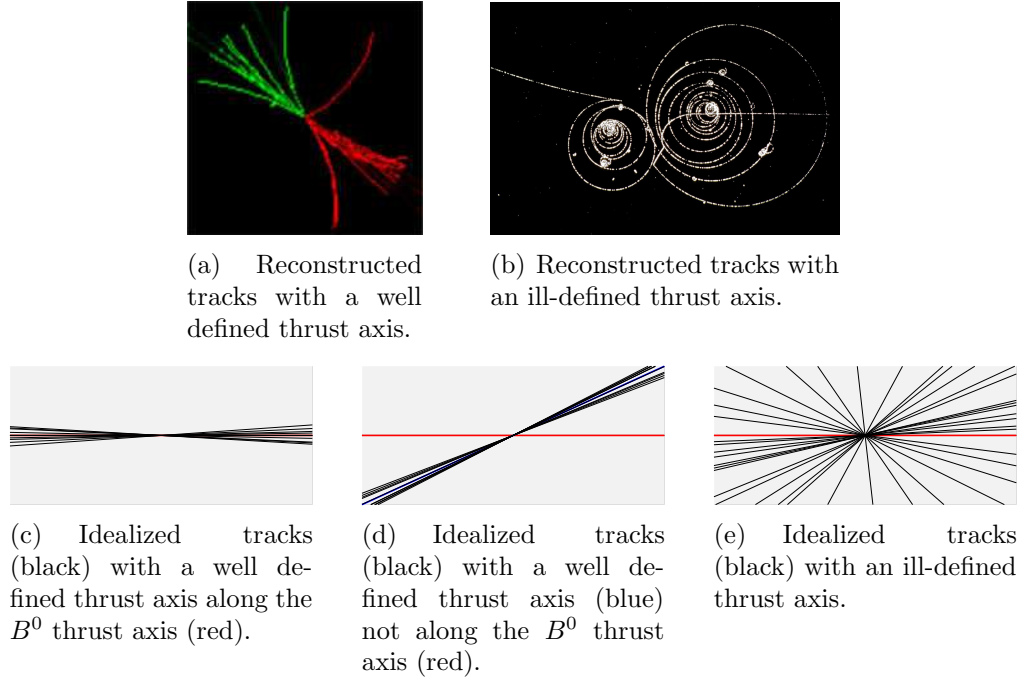


Figure 3.1: Track Reconstruction and the Thrust Axis

$\Delta E$

The observable  $\Delta E$  is the difference in energy between the  $B^0$  candidate and the beam energy in the  $e^+ e^-$  rest frame<sup>4</sup> (see Equation 3.2).

$$\Delta E = E_{B^0}^* - E_{beam}^* \quad (3.2)$$

Since energy must be conserved in all processes, if  $\Delta E$  takes values that lie away from zero, then it is quite likely the  $B^0$  candidate was reconstructed from the wrong daughter particles. Thus, by placing constraints on this variable, some of the combinatorial background from the  $\pi^0$  reconstruction should be eliminated.

---

<sup>4</sup>The (\*) superscript denotes the  $e^+ e^-$  rest frame.

## The Fisher Discriminant

A Fisher Discriminant<sup>5</sup> is defined as the ratio of the variance between a set of distributions to the variance within the distributions, where the variance is defined in the normal statistical sense. It is often interpreted as a signal to noise ratio. In particle physics, the Fisher Discriminate is formed as a linear combination of a set of measurements on observables. The coefficients of the Fisher Discriminant are determined in the standard way as to maximize the overall ability to distinguish between signal and background. For the purposes of a charmless  $B^0$  decay, the sets of measurements are those made on event shape or topological variables including the angles between candidate particles, and nine energy related topological variables collectively referred to as energy flow cones, and the ratio of the second to zeroth Fox-Wolfram moments. The energy flow cones are a set of nine concentric cones centered about the  $B^0$  momentum vector with their apex at the interaction point. The opening angles of these cones are distributed uniformly between  $20^\circ$  and  $180^\circ$ . The amount of energy possessed by tracks and neutral particles that lie between adjacent cones is summed and analyzed to determine whether the event looks most like particle jets, an isotropic  $B$  event, or something in between. The Fox-Wolfram moments ( $H_\ell$ ) are defined in Equation 3.4 where  $P_\ell$  is the  $\ell^{th}$  Legendre polynomial<sup>6</sup>,  $\mathbf{p}_i$  is the momentum of the  $i^{th}$  particle,  $\theta_{i,j}$  is the opening angle between particles

---

<sup>5</sup>For more information on Fisher Discriminant Theory, see pages 465–476 of R. A. Fisher’s *Contributions to Mathematical Statistics*. [8]

<sup>6</sup>The zeroth and fourth Legendre polynomials [15]:

$$\begin{aligned} P_0(x) &= 1 \\ P_2(x) &= \frac{1}{2}(3x^2 - 1) \end{aligned} \tag{3.3}$$

$i$  and  $j$ , and  $E_{vis}$  is the total visible energy of the event.[9]

$$H_\ell = \sum_{i,j} \frac{|\mathbf{p}_i| \cdot |\mathbf{p}_j|}{E_{vis}^2} P_\ell(\cos(\theta_{ij})) \quad (3.4)$$

Thus, the Fisher Discriminant makes use of topological information in order to maximize the ratio of signal to background.

## The $m_{\pi^0}$ Invariant Mass

The  $m_{\pi^0}$  invariant mass is literally just that: the relativistic invariant mass<sup>7</sup> of the  $\pi^0$  candidate, or, the mass of the pion as measured in its rest frame. This quantity is calculated using the energies of detected photons produced via the decay  $\pi^0 \rightarrow \gamma \gamma$ . The world average for the  $m_{\pi^0}$  is  $134.9739 \pm 0.0006 \text{ MeV}/c^2$ . [11] By restricting the  $m_{\pi^0}$  to lie within an interval<sup>8</sup> about the true  $\pi^0$  mass, false  $\pi^0 \rightarrow \gamma \gamma$  reconstructions can be cut out of the final data set.

### 3.2.1 Continuum Background ( $q\bar{q}$ )

The first type of background considered is the most abundant source of background: the continuum background or  $e^+ e^- \rightarrow q\bar{q}$ . The most powerful tool for continuum background discrimination exploits the differences in the characteristic topologies between continuum events and signal events ( $B^0 \rightarrow \pi^0 \pi^0 \pi^0$ ). In true signal events,  $B^0 \bar{B}^0$  pairs are produced via an  $\Upsilon(4S)$  decay. In the  $\Upsilon(4S)$  rest frame—the  $e^+ e^-$  rest frame—the  $B$  mesons have very low momenta and thus the  $B$  decays are nearly isotropic (See Figure 3.1(b) and Figure 3.1(e)). Further, there is little correlation between the unit normalized momentum vectors of the

---

<sup>7</sup> $m_{\pi^0} = \sqrt{E^2 - p^2}$  where  $c$ , the speed of light, has been set to one.

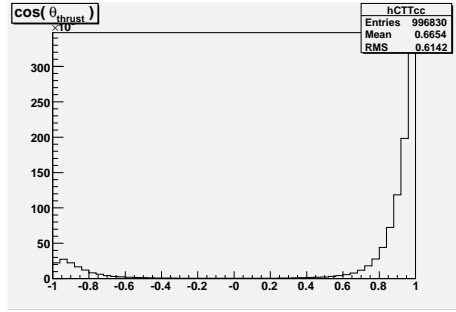
<sup>8</sup>Recall, from Section 1.3.4, that  $\pi^0$ 's are reconstructed with good mass resolution.

$B^0$  daughters, granddaughters, and so on, and the  $\bar{B}^0$  daughters, granddaughters, and so on. That is, they are not oriented “back to back” or anti-parallel to one and other. This event shape is in stark contrast to the shape of continuum events. In light quark ( $u, d, s$ ) continuum events, the event is characterized by a distinct two-jet shape, as in Figure 3.1(a), Figure 3.1(c), and Figure 3.1(d). That is to say, the event will have a definite preferred direction and each track with momentum  $\mathbf{p}$  will almost certainly have a “partner” track with momentum  $-\mathbf{p}$ . In terms of energy cones, it is expected that almost the entire event will fit within the cone with the smallest opening angle (and that cone’s mirror image). In the case of a  $c\bar{c}$  event, the jets are still present but are less pronounced due to the increased mass of the  $c$  quark relative to that of the  $u$ ,  $d$ , and  $s$  quarks.

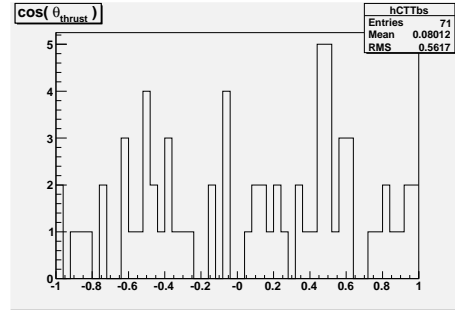
## The Charm Quark and Anti-quark ( $c\bar{c}$ )

The event  $e^+ e^- \rightarrow c\bar{c}$  is the most likely decay and, as such, makes up a great deal of the background. For every  $e^+ e^- \rightarrow \Upsilon(4S) \rightarrow B^0\bar{B}^0$  there are about three  $c\bar{c}$  continuum events. In order to reduce the amount of background, information pertaining to the topology of the event is used as discussed in Section 3.2.1. Therefore, the  $\cos(\theta_{thrust})$  and the Fisher Discriminant distributions will be examined.

Figure 3.2 compares the  $\cos(\theta_{thrust})$  distribution of the  $c\bar{c}$  continuum events with that of the signal events. The signal distribution is randomly distributed between plus and minus one whereas the  $c\bar{c}$  distribution peaks at plus one. A second, much less pronounced peak occurs at minus one. Thus, by restricting the value of  $\cos(\theta_{thrust})$  to regions with  $-\alpha \leq \cos(\theta_{thrust}) \leq \beta$ , where  $\alpha, \beta \approx 0.8$ , a great deal of  $c\bar{c}$  background can be eliminated. On the other hand, by placing these restraints on the data set, some signal will be lost as the signal is distributed



(a) The  $\cos(\theta_{thrust})$  distribution for the event  $e^+ e^- \rightarrow c\bar{c}$ .

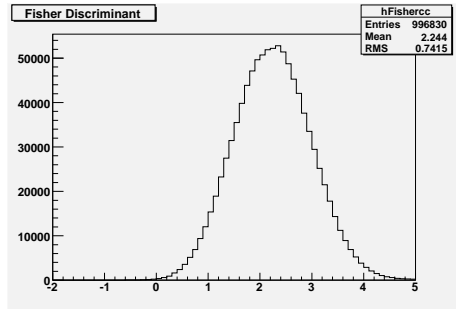


(b) The  $\cos(\theta_{thrust})$  distribution for the signal event.

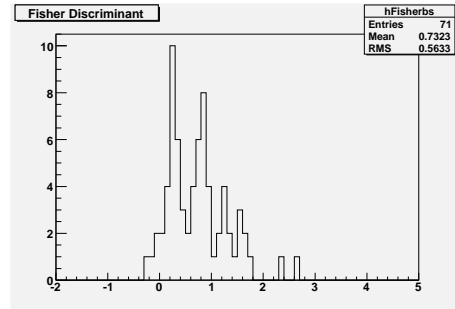
Figure 3.2:  $\cos(\theta_{thrust})$  Distribution for  $c\bar{c}$  Continuum Events

uniformly between  $\pm 1$ .

Figure 3.3 compares the distributions of the Fisher Discriminant for  $c\bar{c}$  events and signal events. Unlike the  $\cos(\theta_{thrust})$  signal distribution, the Fisher Discrimi-



(a) The Fisher Discriminant distribution for the event  $e^+ e^- \rightarrow c\bar{c}$ .



(b) The Fisher Discriminant distribution for the signal event.

Figure 3.3: Fisher Discriminant Distribution for  $c\bar{c}$  Continuum Events

nant signal distribution displays a non-random distribution and peaks near  $\frac{1}{2}$  with a decaying tail along the positive axis. The  $c\bar{c}$  distribution, on the other hand, has a broad maximum near two, and appears to fall off symmetrically about its maximum, rather than tail off in a preferred direction. So, while both distributions display peaks, the peaks are not concentric and by constraining the value of

the Fisher Discriminant about the signal distribution's peak,  $c\bar{c}$  background events should be eliminated; however, the relative broadness of the peaks will result in definite overlap between the two distributions and therefore, it is not immediately clear what the allowed domain should be for the Fisher Discriminant.

## The Up, Down, and Strange Quarks and Anti-quarks ( $u\bar{u}/d\bar{d}/s\bar{s}$ )

The light quark continuum background is made up of three decays:  $e^+ e^- \rightarrow u\bar{u}$ ,  $e^+ e^- \rightarrow d\bar{d}$ , and  $e^+ e^- \rightarrow s\bar{s}$ . The observables that are considered in this section are the same as in the previous one, since topological discrimination is the best way to distinguish between continuum background events and signal events.

While not quite as abundant as the  $c\bar{c}$  background, light quark continuum background makes up a substantial significant portion of the data set, with about two  $e^+ e^- \rightarrow u\bar{u}/d\bar{d}/s\bar{s}$  occurring for every  $e^+ e^- \rightarrow \mathcal{T}(4S)$ . Figure 3.4 compares the  $\cos(\theta_{thrust})$  distribution for the light quark continuum background events with that of the signal events and Figure 3.5 is a comparison of the Fisher Discriminant distributions for the same decays. Qualitatively, the light quark distributions in

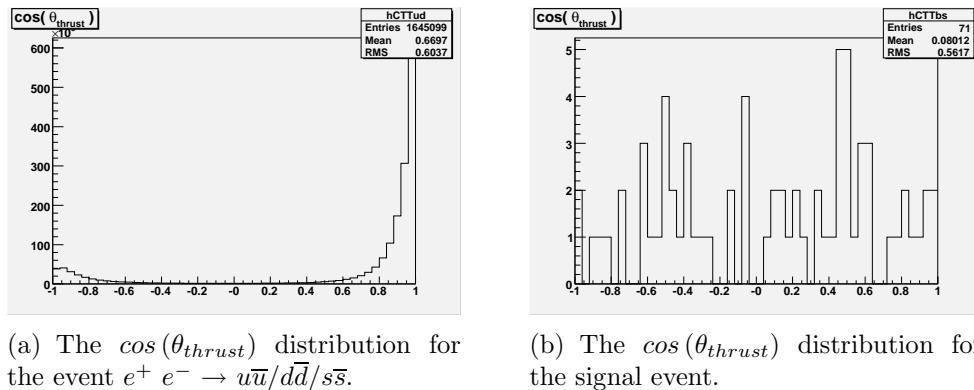
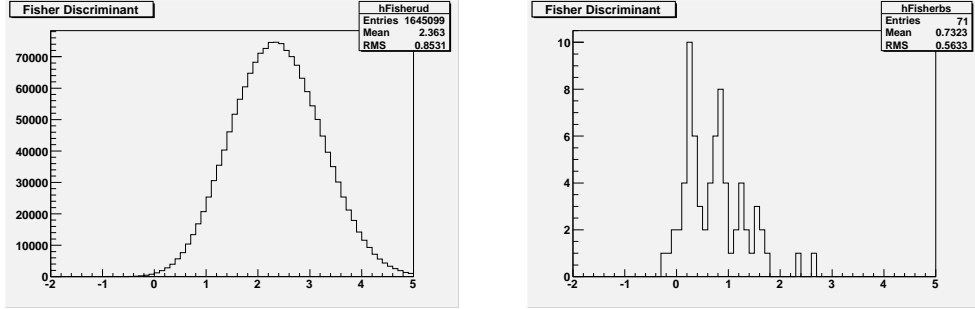


Figure 3.4:  $\cos(\theta_{thrust})$  Distribution for  $u\bar{u}/d\bar{d}/s\bar{s}$  Continuum Events



(a) The Fisher Discriminant distribution for the event  $e^+ e^- \rightarrow c\bar{c}$ .

(b) The Fisher Discriminant distribution for the signal event.

Figure 3.5: Fisher Discriminant Distribution for  $u\bar{u}/d\bar{d}/s\bar{s}$  Continuum Events

$\cos(\theta_{thrust})$  and the Fisher Discriminant look the same as the  $c\bar{c}$  distributions. As a result, the same conclusions are reached; namely, by placing constraints on the two event shape observables,  $\cos(\theta_{thrust})$  and the Fisher Discriminant, light quark continuum background will be reduced.

### 3.2.2 $B^\pm$ Decays with Pion Containing Final States

When  $e^+ e^-$  collisions result in a  $\Upsilon(4S)$ , half of the  $\Upsilon(4S)$ 's decay via  $\Upsilon(4S) \rightarrow B^+ B^-$ . Thus, if there are decay modes of the  $B^\pm$  mesons involving  $\pi^0$ 's, these  $\pi^0$ 's can be mistook for  $\pi^0$ 's from  $B^0 \rightarrow \pi^0 \pi^0 \pi^0$ . Table 7 shows some  $B^\pm$  decay modes with measured branching ratios that contain at least one  $\pi^0$  in the final state.[11] Thus, to eliminate  $B^\pm$  background, information must be used to discriminate against  $\pi^0$ 's originating from  $B^\pm$ 's rather than  $B^0$ 's. The  $m_{\pi^0}$  distribution cannot be used as the pion mass distribution should be centered on the world average (just as it will for signal events) and display roughly the same width as the signal distribution. Topologically, the  $B^\pm$  events and the signal events are expected

| $B^\pm \rightarrow$   | $\rightarrow \pi^0$ State |
|---|---------------------------|
| $D^{*0}$ (2010) $\pi^\pm \pi^\pm \pi^0$                                       | –                         |
| $\pi^\pm \pi^0$   | –                         |
| $\bar{D}^0 \pi^\pm$<br>$\bar{D}^0 \rho^\pm$                                   | $\bar{D}^0$               |
| $D^{*0}$ (2010) $\pi^\pm$   | $D^{*0}$ (2010)           |
| $\bar{D}^{*0}$ (2010) $\pi^\pm$   | $\bar{D}^{*0}$ (2010)     |
| $D^{*\mp}$ (2010) $\pi^\pm \pi^\pm$   | $D^{*\mp}$ (2010)         |
| $J/\psi K^\pm$<br>$J/\psi K^\pm \pi^\pm \pi^\mp$                              | $J/\psi$                  |
| $K^0 \pi^\pm$   | $K^0$                     |
| $K^\pm \pi^\pm \pi^\mp$<br>$K^\pm \rho^0$<br>$K^\pm \phi$<br>$K^\pm f^0(975)$ | $K^\pm$                   |

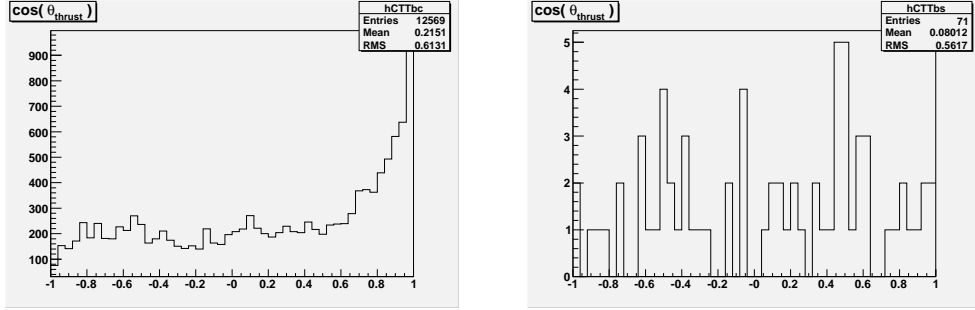
Table 7: Some  $B^\pm$  Decays with Final States Containing At Least One  $\pi^0$

to look the same,<sup>9</sup> thus  $\Delta E$  is the only observable left to consider. The definition of  $\Delta E$  (Equation 3.2) suggests that events with  $\pi^0$ 's that originate from  $B^\pm$  decays should take values less than zero whereas the signal distribution should be peaked at zero and decrease symmetrically about the origin. An interpretation of the Equipartition Theorem can be used to argue that  $\Delta E$  will peak in the negative region for  $B^\pm$  decays. This classical theorem states that the energy of a system will distribute itself evenly among its degree of freedom. Therefore, if a parent particle with energy  $E$  decays into  $n$  daughter particles, each daughter particle should have an energy close to  $\frac{E}{n}$ . With this in mind, two cases are considered.

The first case considers modes in which a  $B^\pm$  decays directly to a  $\pi^0$  daughter or daughters. The  $B^\pm$  mass is  $5277.6 \pm 1.4$  MeV/ $c^2$  and hence, in its rest frame it has an energy of 5277.6 MeV.[11] If the  $B^\pm$  decays into  $n$  daughters, then a given

---

<sup>9</sup>Actually, Monte Carlo analyses shows the  $B^+B^-$  background exhibits a peak in  $\cos(\theta_{thrust})$  at +1 (see Figure 3.6).



(a) The  $\cos(\theta_{thrust})$  distribution for  $B^\pm$  background events.

(b) The  $\cos(\theta_{thrust})$  distribution for the signal event.

Figure 3.6:  $\cos(\theta_{thrust})$  Distribution for  $B^\pm$  Background Events

$\pi^0$  daughter will have an energy around  $\frac{5277.6}{n}$  MeV<sup>10</sup>. If the other two  $\pi^0$  mesons in the reconstruction are the true  $\pi^0$  candidates (i.e.  $\pi^0$ 's from  $B^0 \rightarrow \pi^0\pi^0\pi^0$ ), then they will have a combined energy of about  $\frac{2m_{B^0}}{3}$ . Since the  $B^0$  mass is very close to the  $B^\pm$  mass, the “most confusing” value  $n$  can take, that is, the value of  $n$  that makes it most difficult to distinguish between signal and background, is  $n = 3$ . In this case, the value of  $E_{B^0}^*$  is given by Equation 3.5.

$$\begin{aligned}
 E_{B^0}^* &= \frac{2m_{B^0}}{3} + \frac{m_{B^\pm}}{3} \\
 &= \frac{1}{3}(2m_{B^0} + m_{B^\pm})
 \end{aligned}
 \tag{3.5}$$

The  $B^0$  mass, while nearly equal to the  $B^\pm$  mass, is actually  $5279.4 \pm 1.5$  MeV/ $c^2$  which is greater than the  $B^\pm$  mass ( $5277.6 \pm 1.4$  MeV/ $c^2$ ).<sup>[11]</sup> Thus when the event is reconstructed, it will appear to be missing energy.<sup>11</sup> This is seen in Equation 3.6.

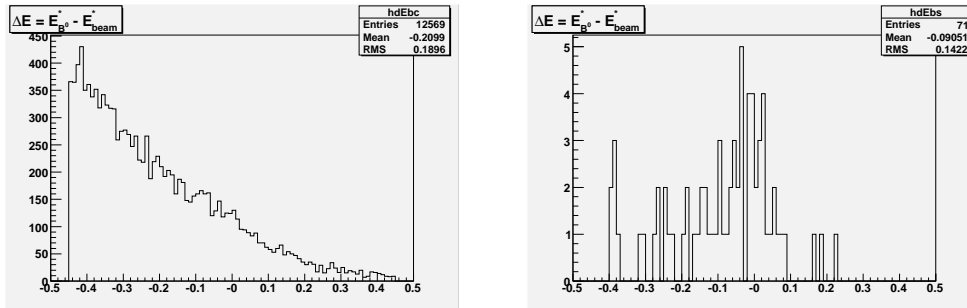
$$m_{B^\pm} < m_{B^0} \Rightarrow \frac{2m_{B^0}}{3} + \frac{m_{B^\pm}}{3} < E_{B^0, true}^* \Rightarrow \Delta E < 0
 \tag{3.6}$$

<sup>10</sup>Note that  $c$ , the speed of light, has been set to unity.

<sup>11</sup>Recall that if the  $B^0$  candidate is reconstructed correctly, then  $E_{B^0}^* \approx E_{beam}^*$ .

Since the masses of the  $B^0$  and the  $B^\pm$  mesons are very nearly the same, events for which  $n$  is different than three will have  $\Delta E$  values lying far from zero and, as such, should easily be recognized as background.

The second case to consider is when the  $\pi^0$  meson used to reconstruct the  $B^0$  candidate is a granddaughter (rather than daughter) of the  $B^\pm$  (e.g.  $B^\pm \rightarrow K^0\pi^\pm$  with  $K^0 \rightarrow \pi^0\pi^0$ ). In this case, the particle (e.g. a  $K^0$ ) that decays into a state containing a  $\pi^0$  will have an energy of about half the  $B^\pm$  rest mass energy. That is, in a quasi two-body  $B^\pm$  decay where the  $B^\pm$  decays directly to two daughter particles, both of the  $B^\pm$  daughter particles will have an energy of about  $\frac{m_{B^\pm}}{2}$ . Then, if one of the daughter particles undergoes a two-body decay (making the overall decay a quasi two-body decay) in which one of the daughters (and hence one of the  $B^\pm$  granddaughters) is a  $\pi^0$ , then that  $\pi^0$  will have an energy of about  $\frac{m_{B^\pm}}{4}$ . Therefore, the  $\frac{m_{B^\pm}}{3}$  term in Equation 3.6 is replaced by  $\frac{m_{B^\pm}}{4}$  and the event will, again, appear to be missing energy. Figure 3.7 is a comparison of the distributions of  $\Delta E$  for  $B^\pm$  background events and for signal events. The  $B^\pm$  distribution,



(a) The  $\Delta E$  distribution for  $B^\pm$  background events.

(b) The  $\Delta E$  distribution for the signal event.

Figure 3.7:  $\Delta E$  Distribution for  $B^\pm$  Background Events

Figure 3.7(a), appears to obey a power-law and is maximal at its left-hand limit (-0.5 GeV)—as expected from the discussion above—and minimal at its right-

hand limit (0.5 GeV); whereas the signal distribution is maximal about zero and tails off towards its left-hand limit.<sup>12</sup> Thus, by restricting  $\Delta E$  to take values in a neighborhood about zero, some of the  $B^\pm$  background should be eliminated.

### 3.2.3 $B^0$ Decays with Pion Containing Final States and $B^0 \rightarrow \pi^0\pi^0\pi^0$ Combinatorial Background

Finally, modes in which  $B^0$  mesons decay into states other than  $\pi^0\pi^0\pi^0$  and combinatorial background from actual  $B^0 \rightarrow \pi^0\pi^0\pi^0$  decays must be considered. Similar to the  $B^\pm$  modes, one source of background from  $B^0$  decays will be the reconstructed  $\pi^0$ 's that are produced by  $B^0$  mesons, but in decays that do not have  $\pi^0\pi^0\pi^0$  as the final state. These decays may involve modes such as  $B^0 \rightarrow D$  and  $B^0 \rightarrow K$ , where the  $D$  and  $K$  decay to states containing  $\pi^0$  mesons. There are also direct decay modes with final states that contain pions that will add to combinatorial background to such as  $B^0/\bar{B}^0 \rightarrow \pi^0\pi^0$ . Again,  $\Delta E$  will be the observable that, when constrained, should eliminate most of this background. Similar arguments as those presented in Section 3.2.2 can be made to show  $\Delta E$  is expected to peak at its left-hand limit for background decays. Also similar to the  $B^\pm$  background, while topological differences between  $B^0$  background decays and signal decays are expected to be insignificant, Monte Carlo studies suggest the background distribution in  $\cos(\theta_{thrust})$  tends towards +1 (see Figure 3.8) while the Fisher Discriminant distributions of the two data sets peak at the same value and behave too similarly

---

<sup>12</sup>In a perfect experiment, signal events would have  $\Delta E$  of exactly zero, or, at the very least, display a narrow maximum at zero. The tail seen in the distribution is a result of systematic error occurring in the EMC and is not, in a pure sense, physical. The EMC tends to systematically underestimate photon energies as some of the energy deposited is lost out the back of the CsI crystals since there are no photomultiplier tubes to detect electron showers in that region.

for the Fisher Discriminant to be a useful discriminating variable.

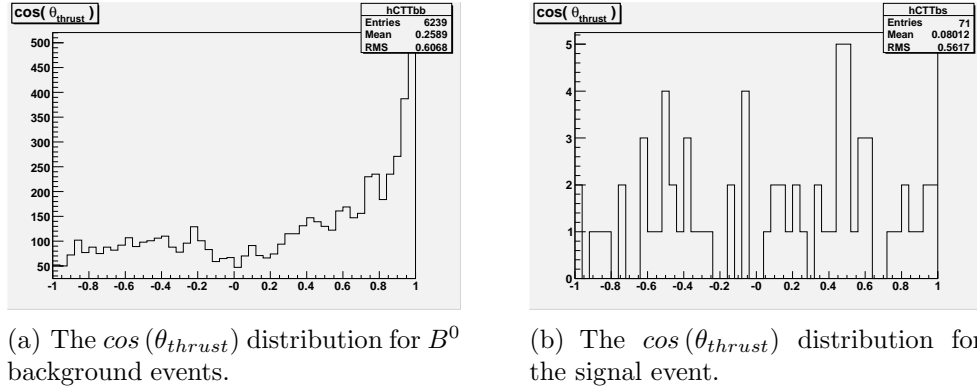


Figure 3.8:  $\cos(\theta_{thrust})$  Distribution for  $B^0$  Background Events

The final major source of background is combinatorial background. This type of background is the result of multiple reconstructions of the same event. Consider the decay  $\Upsilon(4S) \rightarrow B^0 \bar{B}^0$  where  $B^0 \rightarrow \pi^0 \pi^0 \pi^0$ . If the  $\bar{B}^0$  decay contains any  $\pi^0$ 's, there will be multiple ways of reconstructing the  $\pi^0 \pi^0 \pi^0$  final state. Equation 3.8 shows the number of unique combinations of  $\pi^0 \pi^0 \pi^0$  in events containing  $n$   $\pi^0$  mesons.

$$\mathcal{N}_{recon} = \binom{n}{3} = \frac{n!}{3!(n-3)!} \quad (3.7)$$

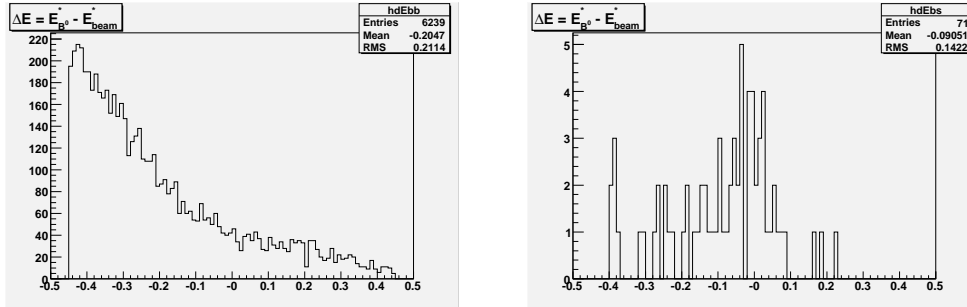
$$(3.8)$$

As can be seen by the presence of the factorial operator, the number of possible reconstructions grows fast, as is shown in Table 8. Combinatorial background can also be discriminated against using the  $\Delta E$  variable. Consider the case where two of the three  $\pi^0$  mesons are the correct  $\pi^0$ 's and the third is from the  $\bar{B}^0$  decay. Using the argument from Section 3.2.2, it can be shown that these types of

| $\mathcal{N}_{\pi^0}$ | $\mathcal{N}_{recon.}$ |
|-----------------------|------------------------|
| 3                     | 1                      |
| 4                     | 4                      |
| 5                     | 10                     |
| 6                     | 20                     |
| 7                     | 35                     |
| 8                     | 56                     |
| 9                     | 84                     |
| 10                    | 120                    |

Table 8: Number of  $B^0$  Reconstructions for Events with  $n$  Detected  $\pi^0$  Mesons

background events will peak in the negative region of the  $\Delta E$  distribution (unless, of course,  $\bar{B}^0 \rightarrow \pi^0\pi^0\pi^0$  or something very similar). Further, the absolute value of  $\Delta E$  should grow with the number of incorrect  $\pi^0$ 's used in the reconstruction. The  $\Delta E$  distributions are shown in Figure 3.9. Consistent with predictions, the



(a) The  $\Delta E$  distribution for  $B^0$  background events.

(b) The  $\Delta E$  distribution for the signal event.

Figure 3.9:  $\Delta E$  Distribution for  $B^0$  Background Events

$B^0$  background distribution is very similar to the  $B^\pm$   $\Delta E$  distribution in that it is maximal at its left-hand limit (-0.5 GeV), minimal at its right-hand limit (0.5 GeV), and appears to obey a power-law. Thus,  $\Delta E$  is a good discriminating variable against both types of  $B^0$  background.

### 3.3 Background Reduction

As discussed in Sections 3.2.1–3.2.3, constraints must be placed on certain observables in order to reduce the background present in the data set. Since  $B^0 \rightarrow \pi^0\pi^0\pi^0$  is a rare process, the amount of background to contend with will be larger, in a proportional sense, than it is for most analyses and hence, very tight or aggressive cuts will need to be made. The constraints placed on the observables will be referred to as *cuts* and the optimal set of cuts will be those which maximize the *significance*,  $\mathcal{S}$ , of the data set. However, the initial Monte Carlo data set contains far too much data for the analysis macros to run over (see Table 6), thus, *Skims* are used to considerably reduce the size of the data set.

#### 3.3.1 Skims

Skims are a very loose set of constraints imposed on the data set that significantly reduce the size of the data set while preserving 100%<sup>13</sup> of the signal originally present. The Release 16.1.3 BPi0Pi0Pi03body skims are applied to each data set during the data collection process.<sup>14</sup> After the skims are applied, the size of the data set is drastically reduced (see Table 10). This data set, however, cannot be analyzed as it is because the ratios between the different decay modes is not correct. The ratios must be of the form given in Table 9 ( $\frac{\mathcal{N}_{B^0\bar{B}^0}}{\mathcal{N}_{c\bar{c}}} = \frac{1}{3.98}$ ,  $\frac{\mathcal{N}_{u\bar{u}/d\bar{d}/s\bar{s}}}{\mathcal{N}_{c\bar{c}}} = \frac{2.36}{3.98}$ , etc.). Table 9 also shows that the limiting decay mode is the  $c\bar{c}$  mode. That is to say, if all of the  $B^0\bar{B}^0$  (or  $B^\pm$ , or  $u\bar{u}/d\bar{d}/s\bar{s}$ ) data were used, there would be an insufficient amount of  $c\bar{c}$  data to run on to preserve the proper ratios. Hence, in

---

<sup>13</sup>If not actually 100%, then 100% to a *very* good approximation.

<sup>14</sup>Release 16.1.3 is the most current Release for which the BPi0Pi0Pi03body skims exist.

|                              | $B^0\bar{B}^0$ | $B^+B^-$     | $u\bar{u}/d\bar{d}/s\bar{s}$ | $c\bar{c}$     |
|------------------------------|----------------|--------------|------------------------------|----------------|
| Initial                      | 541 284 225    | 532 768 433  | 417 920 626                  | 676 351 526    |
| Available                    | 11 532         | 21 480       | 305 491                      | 1 454 228      |
| Renorm.                      | 1.00           | 1.00         | 2.36                         | 3.98           |
| Keep All                     |                |              |                              |                |
| $B^0\bar{B}^0$               | 11 532         | 21 823       | 1 575 132                    | 2 881 837      |
| $B^+B^-$                     | 11 351         | 21 480       | 1 550 351                    | 2 836 499      |
| $c\bar{c}$                   | <b>2 237</b>   | <b>4 233</b> | <b>305 491</b>               | <b>558 922</b> |
| $u\bar{u}/d\bar{d}/s\bar{s}$ | 5 819          | 11 012       | 794 841                      | 1 454 228      |

Table 9: Expected Relative Decay Rates and the Skimmed MC Data

order to increase the number of events in the MC study, more  $c\bar{c}$  MC events that pass the BPi0Pi0Pi03body Skims must be created. Table 10 shows the amount of

| Stage         | Parents        |             |             |                                |
|---------------|----------------|-------------|-------------|--------------------------------|
|               | $B^0\bar{B}^0$ | $B^+B^-$    | $c\bar{c}$  | $u\bar{u}, d\bar{d}, s\bar{s}$ |
| Initial       | 541 284 225    | 532 768 433 | 417 920 626 | 676 351 526                    |
| After Skims   | 11 532         | 21 480      | 305 491     | 1 454 228                      |
| After Renorm. | 2 237          | 4 233       | 305 491     | 558 922                        |

Table 10: Size of the Monte Carlo Data Set after Skimming the Data and Preserving the Decay Ratios

data eliminated by the skims while Table 11 gives the same information, but now in terms of the percent of data remaining after the skims are applied. Note that re-normalizing the data set does not change the amount, percentage wise, of background present in the data set. This adjustment is made so that Monte Carlo data set closer resembles the data set nature will produce in the *BABAR* detector. Since this adjustment process will *not* be performed on the actual data set, the decrease in the number of events *cannot* enter into the calculation used to determine the percentage of background eliminated and the percentage of signal that is retained (i.e. the efficiency) by applying cuts to the data.

| Stage         | Parents                 |                         |                         |                                |
|---------------|-------------------------|-------------------------|-------------------------|--------------------------------|
|               | $B^0\bar{B}^0$          | $B^+B^-$                | $c\bar{c}$              | $u\bar{u}, d\bar{d}, s\bar{s}$ |
| Initial       | $1.00 \times 10^2\%$    | $1.00 \times 10^2\%$    | $1.00 \times 10^2\%$    | $1.00 \times 10^2\%$           |
| After Skims   | $2.13 \times 10^{-3}\%$ | $3.97 \times 10^{-3}\%$ | $5.64 \times 10^{-2}\%$ | $2.69 \times 10^{-1}\%$        |
| After Renorm. | $2.13 \times 10^{-3}\%$ | $3.97 \times 10^{-3}\%$ | $5.64 \times 10^{-2}\%$ | $2.69 \times 10^{-1}\%$        |

Table 11: Percentage of Data Remaining after Applying the Skims

There is still one final source of “background” to consider. When the *BABAR* algorithms reconstruct an  $e^+ e^-$  event, there are nearly always multiple reconstructions (see discussion in Section 3.2.3) of the same event. Thus, while the size of the data set is much smaller, Monte Carlo studies show that the data in Table 10 is slightly misleading. The results of this study reveal that there are about three reconstructions of every  $e^+ e^-$  collision and therefore; in a sense, the results in Table 10 are off by a factor of three. The results of the event reconstruction study are summarized in Table 12.

| $e^+ e^- \rightarrow$                | $\mathcal{N}_{recon.}$ | $\mathcal{N}_{events}$ | $\left\langle \frac{\mathcal{N}_{events}}{\mathcal{N}_{recon.}} \right\rangle$ |
|--------------------------------------|------------------------|------------------------|--|
| $c\bar{c}$                           | $9.968 \times 10^5$    | $2.811 \times 10^5$    | 3.5  |
| $u\bar{u}/d\bar{d}/s\bar{s}$         | $1.645 \times 10^6$    | $5.084 \times 10^5$    | 3.2  |
| $\Upsilon(4S) \rightarrow B^\pm$     | $1.257 \times 10^4$    | $3.922 \times 10^3$    | 3.2  |
| $\Upsilon(4S) \rightarrow B_{bkg}^0$ | $6.239 \times 10^3$    | $1.971 \times 10^3$    | 3.2  |
| Background                           | $2.661 \times 10^6$    | $7.954 \times 10^5$    | 3.3  |
| $\Upsilon(4S) \rightarrow B_{sig}^0$ | $7.100 \times 10^1$    | $6.800 \times 10^1$    | 1.0  |

Table 12: Average Number of Reconstructions per  $e^+ e^-$  Collision

### 3.3.2 NonCharm3BodyUser: the Data Acquisition and Management Package

Before restricting the range in which the observables can lie, the values of the observables must be calculated from the data collected by the *BABAR* detector and, what is more, these values must be stored in a way so that it is possible to cross-reference different observables for the same event (i.e. it must be possible to return the value of  $\cos(\theta_{thrust})$  for all events with  $\Delta E > 0$ ). The observables are calculated and their values stored using the executable created from the compilation of the `NonCharm3BodyUser` (NC3BU) package. Once all the necessary calculations have been made, the NC3BU applet arranges the data into an *ntuple*, or a multidimensional array, so that the data can be easily manipulated.

The `NonCharm3BodyUser` package contains files that calculate the values of observables and files that produce n-tuples for data storage for charmless three-body  $B$  decays, such as  $B^0 \rightarrow \pi^0 \pi^0 \pi^0$ . In this analysis, the Release 18.6.4 version of the package is used, the most current version of NC3BU to date. After compiling and linking the package, .tcl files are created and passed, as arguments, to the NC3BU executable that specify which data sets are to be analyzed. These data sets typically contain so many events that the CPU run time for `NonCharm3BodyUserApp` is greater than the maximum allowed CPU run time in the xlong queue on the Yakut cluster at SLAC. As such, it is first determined how many events can be run over in the allowed time and then the .tcl files are configured to split the data set into smaller data sets accordingly.

The `NonCharm3BodyUserApp` executable takes one argument and returns none. The argument it takes is the aforementioned .tcl file. An example .tcl file is contained within the NC3BU package and hence, all that is needed is to change the values

of certain “flags” so that the executable is properly configured. (N.B. Setting the values of these parameters is done at run time and therefore, recompilation of the package after changes are made is unnecessary.) The flags that are changed for this analysis are: `ConfigPatch`, `LevelofDetail`, `histFileName`, `Analysis`, `InputTclFile`, and `BetaMiniTuple`. The `ConfigPatch` flag may take the values “Run1”, “Run2”, or “MC”.<sup>15</sup> By choosing “MC”, the MC truth variables are recorded in the ntuple; whereas, these variables are not stored (as they are unavailable) for all other configuration values. The `LevelofDetail` may be set to “cache”, “micro”, “extend”, or “refit”. For this analysis, “cache” is used. The `histFileName` flag is a string which is the name of the output file (and the path to the directory in which the file is to be stored). This output file contains the ntuple which, in turn, contains the values of the observables. This file can be written to be compatible with `Paw` or `Root` by using the `BetaMiniTuple` flag. For this analysis, the output files are configured as `Root` files. The `Analysis` flag is the name of the skim used on the data sets. `NonCharm3BodyUser` is capable of handling twenty different types of analyses including `B0Pi0Pi0Pi03body`, the skim used for this analysis. It is important to set this flag correctly, else the MC truth information (i.e. the boolean that is set to true for events in which the decay is  $B^0 \rightarrow \pi^0\pi^0\pi^0$  and is set to false for all other decays) will be inaccurate which, in turn, will lead to all sorts of problems, not the least of which is confusing signal and background events during the calculation of  $\mathcal{S}$ , and the signal efficiency. Essentially, misconfiguring this flag leads to a worthless data set! Finally, there is an `InputTclFile` flag. This is set to the location and name of the .tcl file that contains the name(s) of the data set(s) that is(are) to be analyzed by `NonCharm3BodyUserApp`.

---

<sup>15</sup>For Run3 and Run4, `ConfigPatch` is set to “Run2”.

The output file created by `NonCharm3BodyUserApp` contains seven ntuples and seven one dimensional histograms. For the purposes of this analysis, all the data of interest is stored within the ntuple named `ntp3`. This ntuple contains all the observable values and all the MC truth information previously discussed<sup>16</sup>. With the necessary data stored in `Root` ntuples, it can now be analyzed.

### 3.3.3 Cuts on Observables

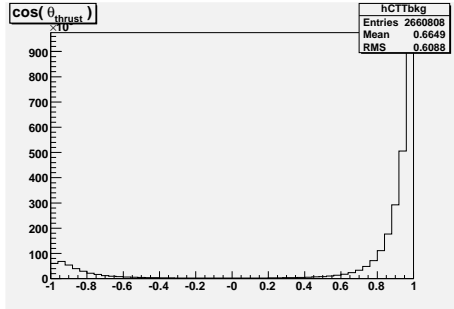
The last method of background discrimination is cutting on observables. This method takes advantage of the differences in the statistical distributions that different types of decays exhibit as discussed in Section 3.2. Further, it ensures that the reconstructed  $m_{\pi^0}$  lies within an acceptable region of uncertainty about its central value. Figure 3.10 shows the signal and background distributions of the four observables that will ultimately be constrained by cuts.

### Obtaining the Most Significant Cuts

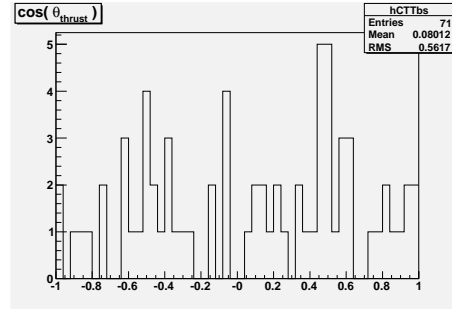
While the goal of the cuts is to eliminate as much background from the data set as possible, the cuts should also preserve as much signal as possible. In order to achieve this, the optimal set of cuts are those that maximize the significance,  $\mathcal{S}$ , of the data set. Before any cuts are applied, the significance of the data set is 0.042. The observables shown in Figure 3.10 are taken to be independent and thus, by constraining the range in which one observable can lie, data should be removed from the distributions of the other observables *at random* or *without bias*. To determine whether or not this is true, cuts are applied on a particular observable

---

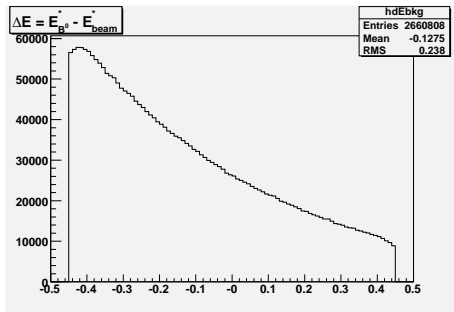
<sup>16</sup>For a full list of the information stored within this ntuple, see [www.slac.stanford.edu/~fwilson/3body\\_reson\\_new.shtml](http://www.slac.stanford.edu/~fwilson/3body_reson_new.shtml)



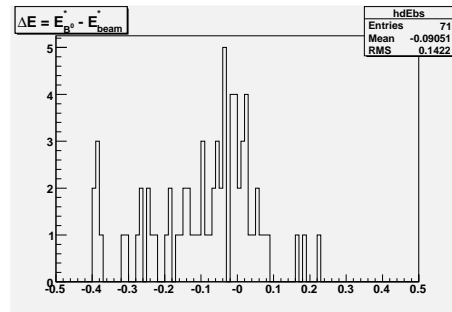
(a) The  $\cos(\theta_{thrust})$  distribution for all background events.



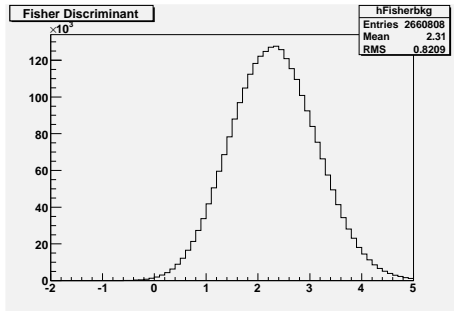
(b) The  $\cos(\theta_{thrust})$  distribution for all signal events.



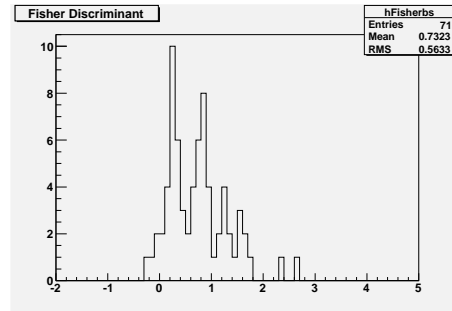
(c) The  $\Delta E$  distribution for all background events.



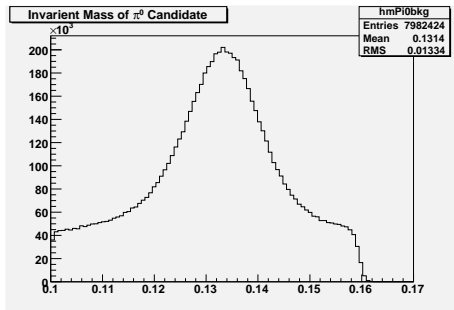
(d) The  $\Delta E$  distribution for all signal events.



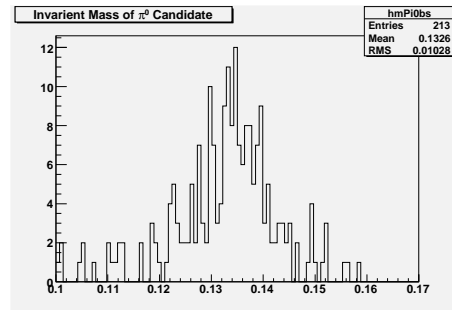
(e) The Fisher Discriminant distribution for all background events.



(f) The Fisher Discriminant distribution for all signal events.



(g) The  $m_{\pi^0}$  distribution for all background events.



(h) The  $m_{\pi^0}$  distribution for all signal events.

Figure 3.10: The Signal and Background Distributions for the Cut Observables

and the other three distributions are compared to their original distributions (i.e. the Fisher Discriminant,  $m_{\pi^0}$ , and  $\Delta E$  are plotted with cuts applied to  $\cos(\theta_{thrust})$ ). These plots are shown in Figure 3.11 and Figure 3.12. The plots suggest there exists a correlation between  $\cos(\theta_{thrust})$  and the Fisher Discriminant. As these are both topological or event shape variables, it is not entirely surprising that a correlation is observed. As a result of this correlation, it is important that the method by which  $\mathcal{S}$  is optimized is not reliant upon the assumption that the observables are independent quantities.

## Baseline Cuts

The baseline cuts (also referred to as Cuts 1) are used to determine the neighborhood in which the ideal cuts live, not necessarily determine their optimal values. Ultimately, the results of the baseline cuts will be used as the initial guess in a maximization scheme that is similar to a numerical “relaxation” schemes in that the results of the  $i^{th}$  iteration are used as starting “guesses” for the values in the  $i^{th} + 1$  iteration. The iterations continue until values calculated for the cuts “relax” to their optimal value. In other words, the procedure is carried out until the results of the  $j^{th}$  iteration are the same as the values calculated by the  $j^{th} + 1$  iteration. Before that iterative procedure can be carried out however, the baseline cuts must be established so that an initial best guess can be found. The baseline cuts are determined in the following way:

1. The full domain of the observable is discretize into  $N$  domains of size  $\Delta x$ .
2. The significance of the data set is calculated without any cuts applied to the observable.
3. A cut is placed at the left-hand limit of the domain plus  $\Delta x$ .

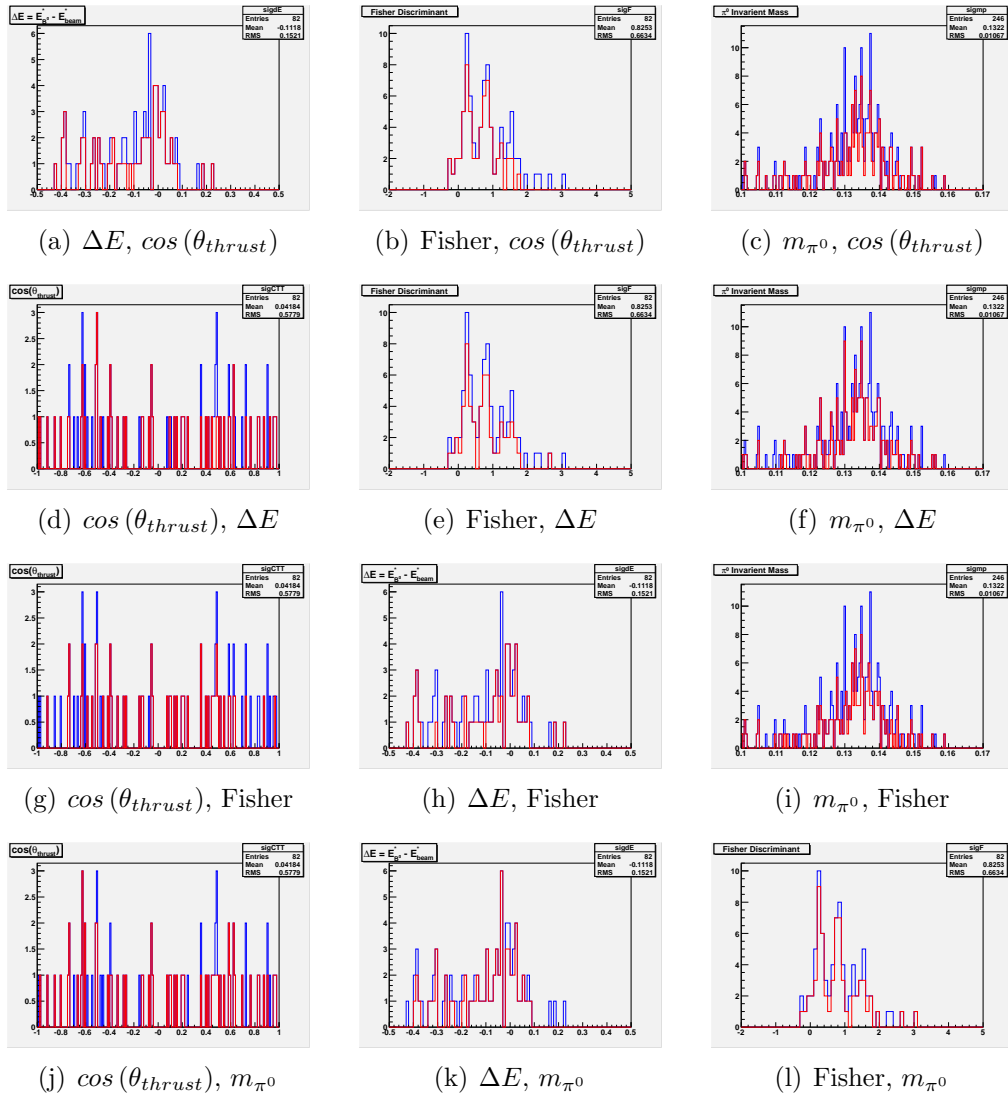


Figure 3.11: Plots labeled as, “a, b” are plots of the signal distribution of “a” (blue) with the cuts applied to “b”(red).

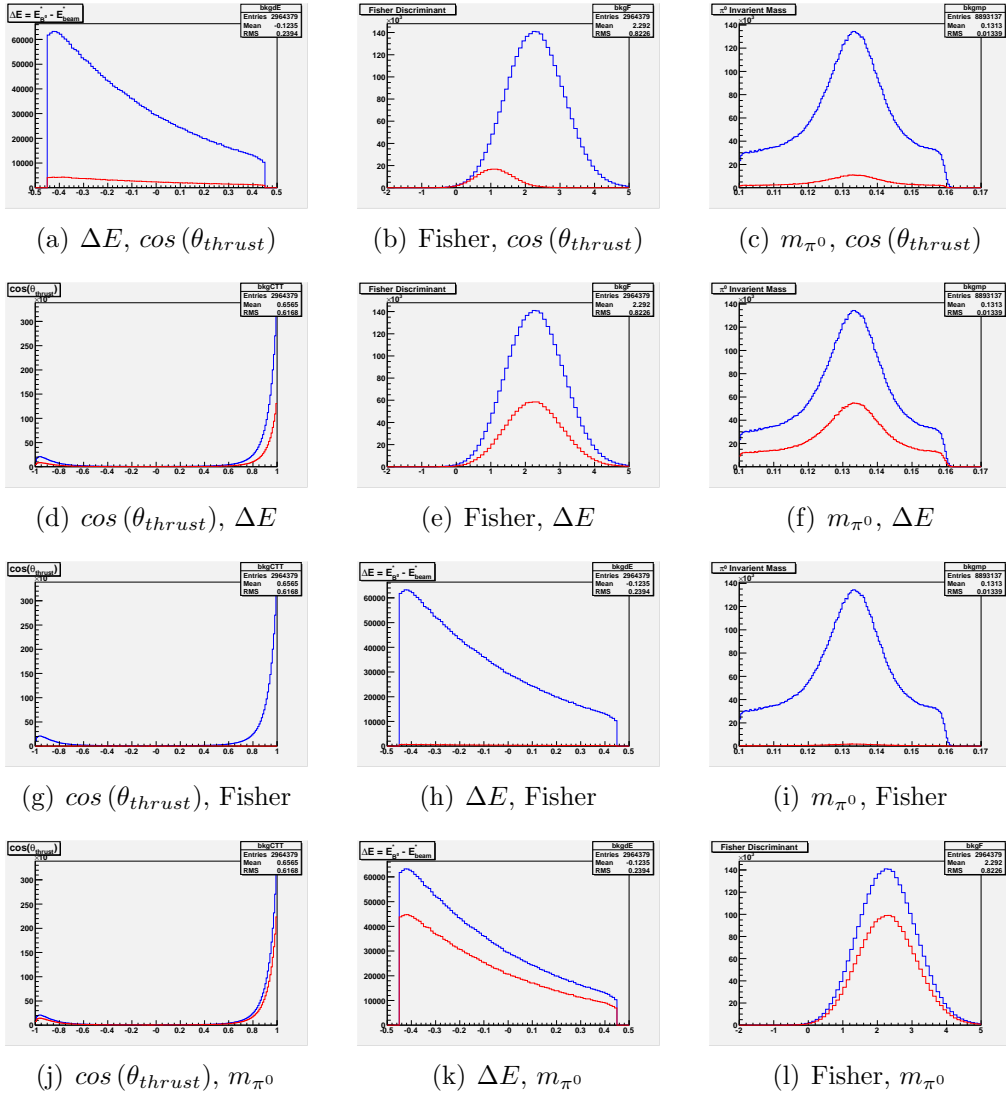


Figure 3.12: Plots labeled as, “a, b” are plots of the background distribution of “a” (blue) with the cuts 2 applied to “b” (red).

4. The significance of the data set is calculated and stored in memory.
5. The lower-bound cut is increased by  $\Delta x$ .
6. The significance of the data set is calculated and stored in memory.
7. Steps 5–6 are repeated until the left-hand cut reaches the point corresponding to the peak in the signal distribution.<sup>17</sup>
8. The cut with the best (i.e. maximum value) of  $\mathcal{S}$  is declared as the lower-bound cut value.
9. A cut is placed at the right-hand limit of the domain minus  $\Delta x$  (NB There is no longer any restraint on the lower bound value of the distribution.).
10. The significance of the data set is calculated and stored in memory.
11. The upper-bound cut is decreased by  $\Delta x$ .
12. The significance of the data set is calculated and stored in memory.
13. Steps 11–12 are repeated until the right-hand cut reaches the point corresponding to the peak in the signal distribution.
14. The cut with the best of  $\mathcal{S}$  is declared as the upper-bound cut value.

The baseline cut procedure is pictured in Figure 3.13. The domains of the various observables are digitized according to Table 13. The results of Cuts 1 are shown in Table 14 and Figure 3.14 shows the cuts applied to the distributions. Finally, Table 15 and Table 16 show the amount by which Cuts 1 reduce the background.

---

<sup>17</sup>In the case of  $\cos(\theta_{thrust})$  the signal distribution has no peak, thus the point  $x = 0$ , the middle of the full domain is used.

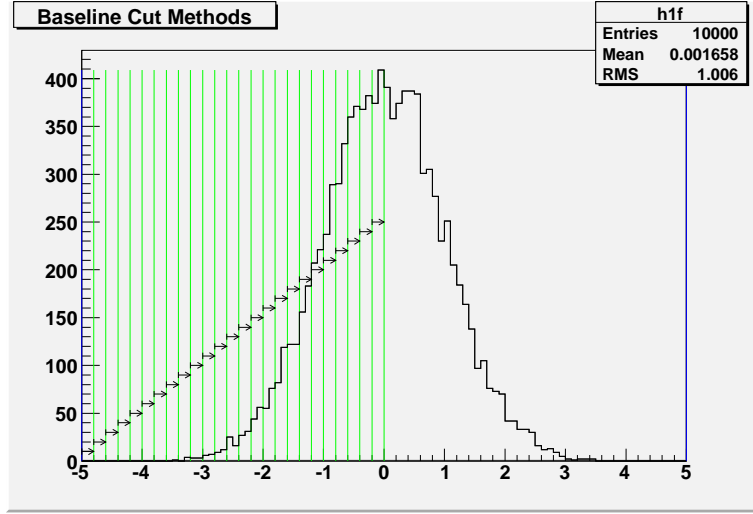


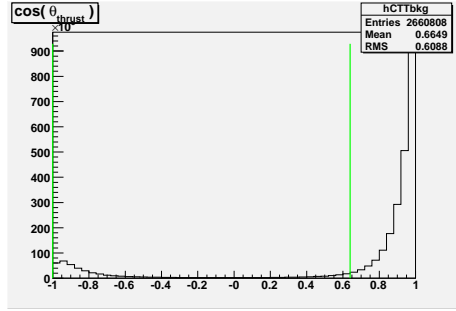
Figure 3.13: Baseline Cuts Scheme

| Observable                       | Left Hand Limit | Right Hand Limit | Peak | $\Delta x$ |
|----------------------------------|-----------------|------------------|------|------------|
| $\Delta E$ (MeV)                 | -0.50           | 0.50             | 0.00 | 0.020      |
| Fisher Discriminant              | -2.00           | 5.00             | 0.80 | 0.200      |
| $m_{\pi^0}$ ( $\text{GeV}/c^2$ ) | 0.10            | 0.17             | 0.13 | 0.005      |
| $\cos(\theta_{thrust})$          | -1.00           | 1.00             | 0.00 | 0.020      |

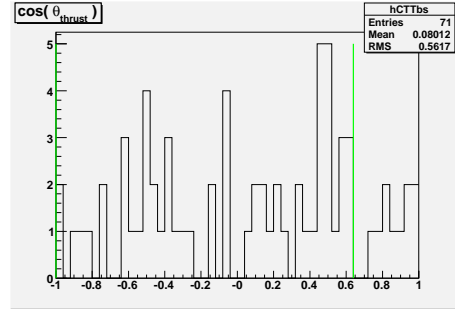
Table 13: Discretized Observable Domains for Cuts 1

| Observable              | Bound  | Value                 |
|-------------------------|--------|-----------------------|
| $\Delta E$              | $\geq$ | -0.28 MeV             |
|                         | $\leq$ | 0.10 MeV              |
| Fisher Discriminant     | $\geq$ | -0.4                  |
|                         | $\leq$ | 1.0                   |
| $M_{\pi^0}$             | $\geq$ | 0.12 $\text{MeV}/c^2$ |
|                         | $\leq$ | 0.15 $\text{MeV}/c^2$ |
| $\cos(\theta_{thrust})$ | $\geq$ | -1.0                  |
|                         | $\leq$ | 0.64                  |
| $\mathcal{S} = 0.231$   |        |                       |

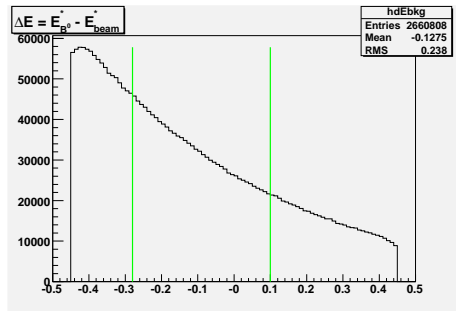
Table 14: Cuts 1 Results



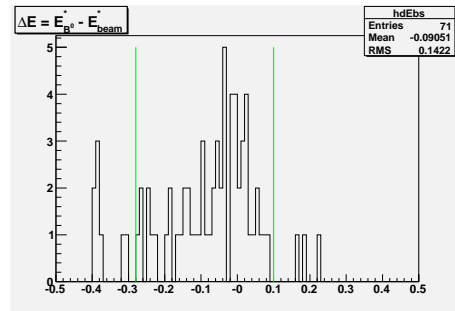
(a)  $\cos(\theta_{thrust})$  Background



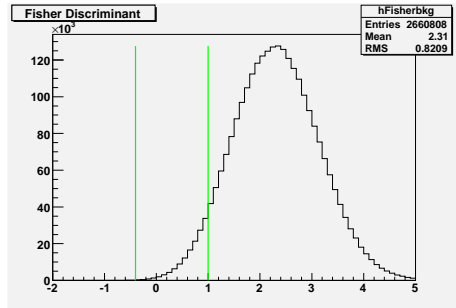
(b)  $\cos(\theta_{thrust})$  Signal



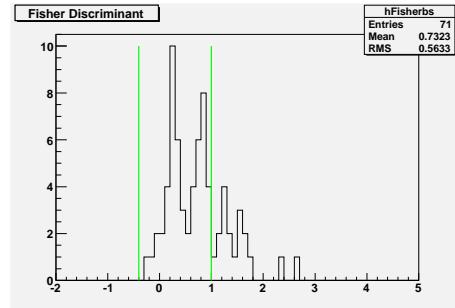
(c)  $\Delta E$  Background



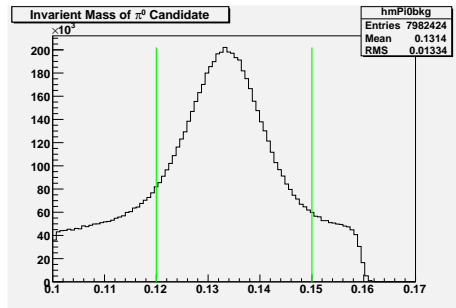
(d)  $\Delta E$  Signal



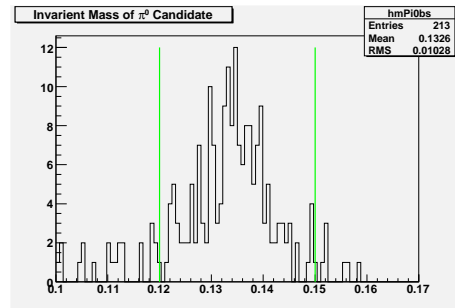
(e) Fisher Discriminant Background



(f) Fisher Discriminant Signal



(g)  $m_{\pi^0}$  Background



(h)  $m_{\pi^0}$  Signal

Figure 3.14: Baseline Cuts

|              | Parents        |          |            |                                |
|--------------|----------------|----------|------------|--------------------------------|
| Stage        | $B^0\bar{B}^0$ | $B^+B^-$ | $c\bar{c}$ | $u\bar{u}, d\bar{d}, s\bar{s}$ |
| After Cuts 1 | 275            | 615      | 3 144      | 7 521                          |

Table 15: Size of the Monte Carlo Data Set after Applying Cuts 1

|              | Parents                 |                         |                         |                                |
|--------------|-------------------------|-------------------------|-------------------------|--------------------------------|
| Stage        | $B^0\bar{B}^0$          | $B^+B^-$                | $c\bar{c}$              | $u\bar{u}, d\bar{d}, s\bar{s}$ |
| After Cuts 1 | $5.08 \times 10^{-5}\%$ | $1.14 \times 10^{-4}\%$ | $5.81 \times 10^{-4}\%$ | $1.39 \times 10^{-3}\%$        |

Table 16: Percentage of Data Remaining after Applying Cuts 1

The baseline cuts do eliminate a lot of background; however, it is unlikely they are the most significant cuts. The lower bound cut on  $\cos(\theta_{thrust})$  is good evidence that a more significant data set can be produced. Thus, a second optimization routine that uses the results of Cuts 1 is run on the data set.

## Tight Cuts Relaxation Procedure

While there is good reason to believe the significance has not yet been maximized, there is also good reason to believe the cuts that will maximize  $\mathcal{S}$  lie in a neighborhood about the current cut values. As such, the second optimization scheme searches for a better set of cuts by discretizing an interval about the cuts established by Cuts 1. By discretizing a smaller region, a smaller step size (i.e.  $\Delta x$ ) can be used; however, if it is unlikely that Cuts 1 are the optimal set of cuts, is any more likely the next set of cuts will return a data set with the best possible  $\mathcal{S}$ ? Further, how will it be determined when the cuts are maximized, that is, how can it be said for certain that there exist no other set of cuts that will increase the significance, or for that matter, can it be determined that there exists a unique

maximum in  $\mathcal{S}$ ?

The questions posed above have no obvious answers. Even if a set of cuts can be found that eliminates every background event can it be said for sure that there aren't another set of cuts that also eliminate all of the background but preserve one more signal event? In order to avoid having to make decisions about what value of  $\mathcal{S}$  is good enough, a scheme that is similar to a numerical "relaxation" procedure is used as discussed in the previous sub-section. The calculation of the best cut values works in the following way:

1. The domain is initially set with its left-limit equal to the lower-bound cut value (determined from the previous iteration) and its right-hand limit equal to its upper-bound cut value (also determined from the previous iteration).
2. The significance of the data set is calculated and stored in memory.
3. A domain given by  $x_{min}^{LB} = x_{LB}^{i-1} - \frac{x_{UB}^{i-1} - x_{LB}^{i-1}}{2}$  and  $x_{max}^{LB} = x_{LB}^{i-1} + \frac{x_{UB}^{i-1} - x_{LB}^{i-1}}{2}$  is discretize where  $x_{LB/UB}^{i-1}$  is initially set to the value established by Cuts 1, but, most generally, takes the value returned by the previous iteration.
4. A cut is placed at the left-hand limit of the domain defined in Step 3 and another cut is placed at  $x_{UB}^{i-1}$  as discussed in Step 1.
5. The significance of the data set is calculated and stored in memory.
6. The lower-bound cut *only* is increased by  $\Delta x$ .
7. The significance of the data set is calculated and stored in memory.
8. Steps 6–7 are repeated until the left-hand limit reaches the end of the domain defined in Step 3.

| Observable                | $\Delta x$ |
|---------------------------|------------|
| $\Delta E$ (MeV)          | 0.005      |
| Fisher Discriminant       | 0.025      |
| $m_{\pi^0}$ (GeV/ $c^2$ ) | 0.001      |
| $\cos(\theta_{thrust})$   | 0.005      |

Table 17: Discretized Observable Domains for Cuts 2

9. The cut with the best value of  $\mathcal{S}$  is declared as the lower-bound cut value.
10. A domain given by  $x_{min}^{UB} = x_{UB}^{i-1} - \frac{x_{UB}^{i-1} - x_{LB}^{i-1}}{2}$  and  $x_{max}^{UB} = x_{UB}^{i-1} + \frac{x_{UB}^{i-1} - x_{LB}^{i-1}}{2}$  is discretize.
11. A cut is placed at the left-hand limit of the domain defined in Step 10 and another cut is placed at  $x_{LB}^{i-1}$ .
12. The significance of the data set is calculated and stored in memory.
13. The upper-bound cut *only* is increased by  $\Delta x$ .
14. The significance of the data set is calculated and stored in memory.
15. Steps 13–14 are repeated until the left-hand limit reaches the end of the domain defined in Step 10.
16. The cut with the best value of  $\mathcal{S}$  is declared as the upper-bound cut value.
17. This entire procedure is repeated with the values  $x_{LB/UB}^{i-1}$  replaced by  $x_{LB/UB}^i$ .
18. Step 17 is repeated until  $\mathcal{S}$  no longer increases ( $\mathcal{S}^{i-1} \approx \mathcal{S}^i$ ) and/or the cut values stop changing ( $x_{LB/UB}^{i-1} \approx x_{LB/UB}^i$ ).

The tight cut process is pictured in Figure 3.15. The domains of the various observables were discretized according to Table 17. The results of the Tight Cuts

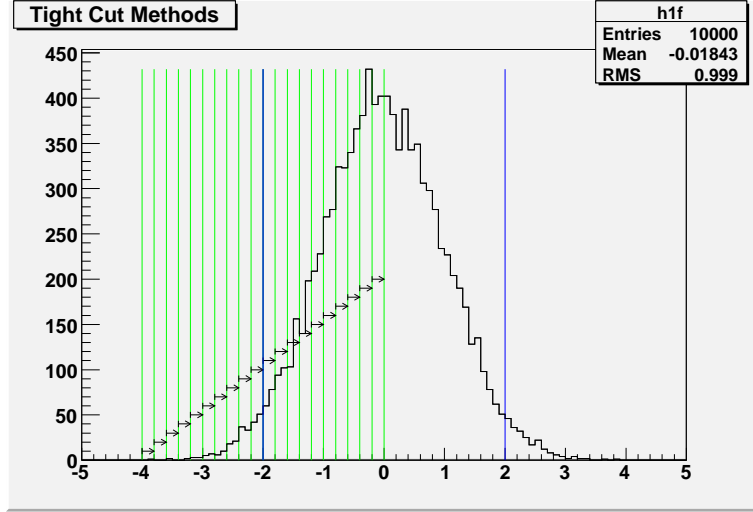


Figure 3.15: Tight Cuts Scheme

process (also referred to as Cuts 2) are shown in Table 18 and Figure 3.16 shows the cuts applied to the distributions. Table 19 and Table 20 show the amount by

| Observable                | Bound  | Iteration |         |         |         |        |        |
|---------------------------|--------|-----------|---------|---------|---------|--------|--------|
|                           |        | 1         | 2       | 3       | 4       | 5      | 6      |
| $\Delta E$ (MeV)          | $\geq$ | -0.095    | -0.1925 | -0.0875 | -0.015  | -0.015 | -0.005 |
|                           | $\leq$ | 0.12      | 0.0175  | 0.0675  | 0.015   | 0.015  | 0.015  |
| Fisher Discriminant       | $\geq$ | -0.275    | -0.2875 | -0.125  | -0.2875 | -0.05  | -0.05  |
|                           | $\leq$ | 0.45      | 0.3875  | 0.35    | 0.3875  | 0.35   | 0.15   |
| $M_{\pi^0}$ (MeV/ $c^2$ ) | $\geq$ | 0.122     | 0.127   | 0.1305  | 0.1265  | 0.13   | 0.1295 |
|                           | $\leq$ | 0.144     | 0.14    | 0.1425  | 0.1395  | 0.139  | 0.1395 |
| $\cos(\theta_{thrust})$   | $\geq$ | -0.64     | -0.52   | -0.52   | -0.53   | -0.525 | -0.42  |
|                           | $\leq$ | 0.52      | 0.1     | -0.04   | -0.04   | -0.035 | -0.04  |
| $\mathcal{S}$             | =      | 0.264     | 0.410   | 0.480   | 0.688   | 1.342  | –      |

Table 18: Cuts 2 Results

which Cuts 2 reduce the background. Cuts 2a refer to the the results of the iterative procedure after five iterations and Cuts 2b refer to the results of the iterative procedure after six iterations.

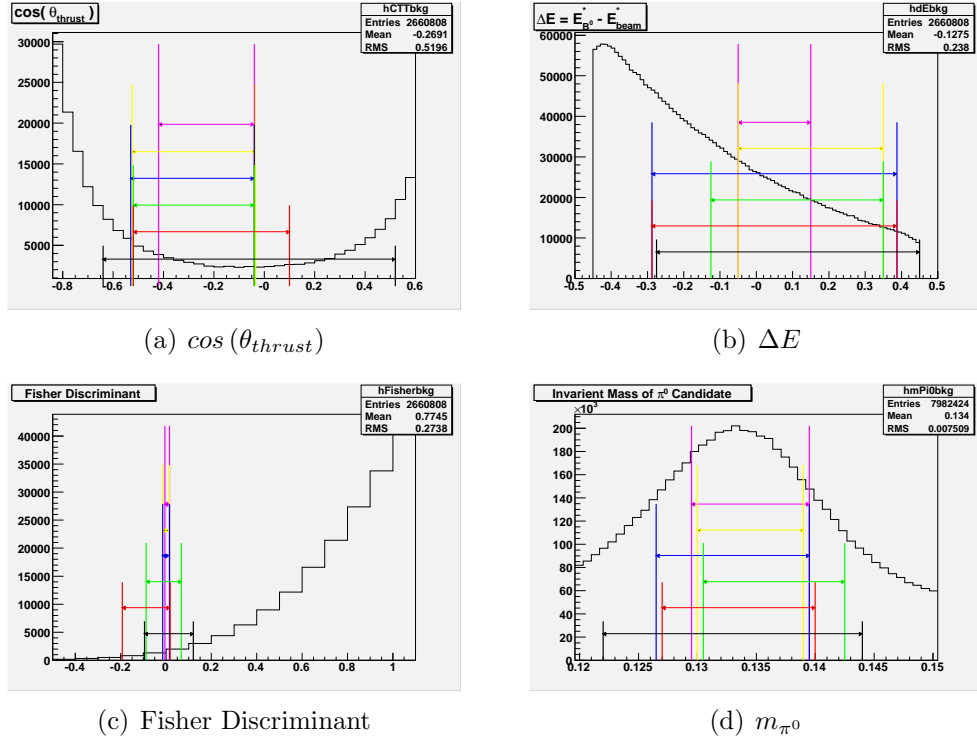


Figure 3.16: Tight Cuts

| Stage         | Parents        |          |            |                                |
|---------------|----------------|----------|------------|--------------------------------|
|               | $B^0\bar{B}^0$ | $B^+B^-$ | $c\bar{c}$ | $u\bar{u}, d\bar{d}, s\bar{s}$ |
| After Cuts 2a | 4              | 1        | 1          | 2                              |
| After Cuts 2b | 1              | 0        | 0          | 0                              |

Table 19: Size of the Monte Carlo Data Set after Applying Cuts 2

| Stage         | Parents                 |                         |                         |                                |
|---------------|-------------------------|-------------------------|-------------------------|--------------------------------|
|               | $B^0\bar{B}^0$          | $B^+B^-$                | $c\bar{c}$              | $u\bar{u}, d\bar{d}, s\bar{s}$ |
| After Cuts 2a | $7.39 \times 10^{-7}\%$ | $1.85 \times 10^{-7}\%$ | $1.85 \times 10^{-7}\%$ | $3.69 \times 10^{-7}\%$        |
| After Cuts 2b | $1.85 \times 10^{-7}\%$ | 0%                      | 0%                      | 0%                             |

Table 20: Percentage of Data Remaining after Applying Cuts 2

Cuts 2b yield zero background events and one signal event<sup>18</sup>, whereas Cuts 2a preserve four signal events but, the data set still contains four background events. There is no combination of cuts (in the neighborhood of the Cuts 2b) that will preserve more than one signal event and cut out all of the background, hence, the optimal set of cuts has been found. A summary of the entire background reduction, as well as a discussion of the properties of the optimized data set, is given in the next section.

### 3.3.4 The Optimized Data Set

The most significant data set has been obtained. The values of the cuts that optimize the data set are shown in Table 18. Cuts on topological variables proved to be the best discrimination method against continuum background, while cuts on kinematic variables are the best discrimination method against  $B^0\bar{B}^0$  and  $B^+B^-$  background. A summary of the background elimination process is shown in Table 21 and Table 22. Ultimately, the distribution of the energy substituted

| Stage         | Parents        |             |             |                                | $\mathcal{S}$ |
|---------------|----------------|-------------|-------------|--------------------------------|---------------|
|               | $B^0\bar{B}^0$ | $B^+B^-$    | $c\bar{c}$  | $u\bar{u}, d\bar{d}, s\bar{s}$ |               |
| Initial       | 541 284 225    | 532 768 433 | 417 920 626 | 676 351 526                    | –             |
| After Skims   | 11 532         | 21 480      | 305 491     | 1 454 228                      | –             |
| After Renorm. | 2 237          | 4 233       | 305 491     | 558 922                        | 0.042         |
| After Cuts 1  | 275            | 615         | 3 144       | 7 521                          | 0.231         |
| After Cuts 2a | 4              | 1           | 1           | 2                              | 1.342         |
| After Cuts 2a | 1              | 0           | 0           | 0                              | -             |

Table 21: Summary of Background Elimination

invariant mass, defined by Equation 3.9, of the  $B^0$  candidate will be fit in order

---

<sup>18</sup>MC Truth information reveals the one remaining  $B^0\bar{B}^0$  event is in fact a signal event.

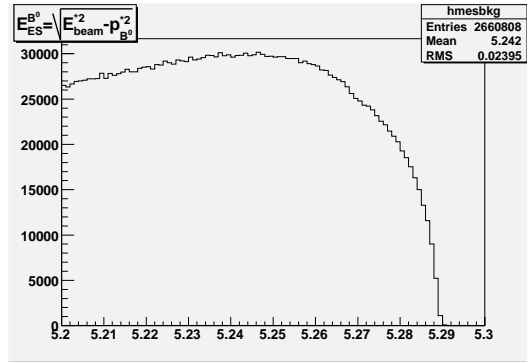
| Stage         | Parents                 |                         |                         |                                |
|---------------|-------------------------|-------------------------|-------------------------|--------------------------------|
|               | $B^0\bar{B}^0$          | $B^+B^-$                | $c\bar{c}$              | $u\bar{u}, d\bar{d}, s\bar{s}$ |
| Initial       | $1.00 \times 10^2\%$    | $1.00 \times 10^2\%$    | $1.00 \times 10^2\%$    | $1.00 \times 10^2\%$           |
| After Skims   | $2.13 \times 10^{-3}\%$ | $3.97 \times 10^{-3}\%$ | $5.64 \times 10^{-2}\%$ | $2.69 \times 10^{-1}\%$        |
| After Renorm. | $2.13 \times 10^{-3}\%$ | $3.97 \times 10^{-3}\%$ | $5.64 \times 10^{-2}\%$ | $2.69 \times 10^{-1}\%$        |
| After Cuts 1  | $5.08 \times 10^{-5}\%$ | $1.14 \times 10^{-4}\%$ | $5.81 \times 10^{-4}\%$ | $1.39 \times 10^{-3}\%$        |
| After Cuts 2a | $7.39 \times 10^{-7}\%$ | $1.85 \times 10^{-7}\%$ | $1.85 \times 10^{-7}\%$ | $3.69 \times 10^{-7}\%$        |
| After Cuts 2b | $1.85 \times 10^{-7}\%$ | 0.0%                    | 0.0%                    | 0.0%                           |

Table 22: Percentages of Background Eliminated during Various Background Reduction Stages

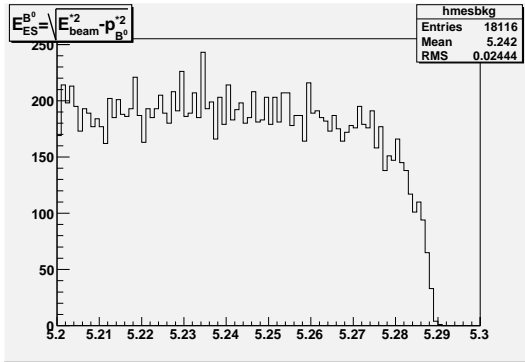
to determine the amount of signal present in the data set without using the MC Truth information.

$$m_{\text{ES}}^{B^0} = \sqrt{E_{\text{beam}}^{*2} - p_{B^0}^{*2}} \quad (3.9)$$

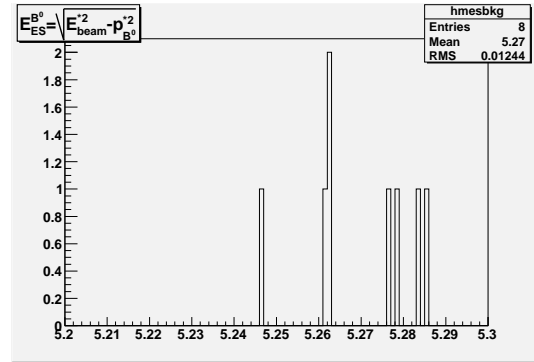
These distributions are shown in Figure 3.17. In the case of Figure 3.17(c) where only eight events are present in the entire distribution, a fit makes no sense. Similarly, when Cuts 2b (not pictured) are applied, only one event remains and again, a fit makes no sense. How to deal with these sparse distributions and how the necessary information can be extracted from the  $m_{\text{ES}}^{B^0}$  distribution is discussed in Chapter 4.



(a) The Monte Carlo  $m_{ES}^{D^0}$  background distribution after BPi0Pi0Pi03body Skims.



(b) The Monte Carlo  $m_{ES}^{D^0}$  background distribution with baseline cuts applied.



(c) The Monte Carlo  $m_{ES}^{D^0}$  background distribution with Cuts 2a applied.

Figure 3.17: Monte Carlo  $m_{ES}^{D^0}$  Background Distribution

# CHAPTER 4

## Monte Carlo Signal Analysis

Having optimized the significance of the Monte Carlo data set, the last task is to extract the necessary information in order to measure  $\mathcal{B}(B^0 \rightarrow \pi^0\pi^0\pi^0)$ . From definition of the branching ratio (see Equation 2.1), it is seen that the number of total  $B^0$  events and the number of signal ( $B^0 \rightarrow \pi^0\pi^0\pi^0$ ) events must be somehow extracted from the data. Clearly, this information must be extracted without using the MC Truth information; however, by using the MC Truth data, it can be determined how precisely this data can be extracted. Table 23 shows the amount of signal that remains after the cuts derived in Section 3.3.3 are applied. When

| Cuts    | Total $B^0\bar{B}^0$ Decays |                         | Signal Decays |            |
|---------|-----------------------------|-------------------------|---------------|------------|
|         | Number                      | Percentage              | Number        | Percentage |
| None    | 541 284 225                 | $1.00 \times 10^2\%$    | 351           | 100.00%    |
| Skims   | 11 181                      | $2.07 \times 10^{-3}\%$ | 351           | 100.00%    |
| Renorm. | 2 169                       | $2.07 \times 10^{-3}\%$ | 68            | 100.00%    |
| Cuts 1  | 244                         | $2.33 \times 10^{-4}\%$ | 31            | 45.59%     |
| Cuts 2a | 3                           | $2.86 \times 10^{-6}\%$ | 3             | 4.41%      |
| Cuts 2b | 1                           | $9.54 \times 10^{-7}\%$ | 1             | 1.47%      |

Table 23: Signal Efficiency

the value  $\mathcal{N}_{sig}$  is extracted from the real data, it will be important to remember

that the value obtained is only 1.47% of the total number of  $B^0 \rightarrow \pi^0\pi^0\pi^0$  events present in *total* the data set.

#### 4.1 Extracting $\mathcal{N}_{sig}$ and $\mathcal{N}_{bkg}$ from the $m_{ES}^{B^0}$ Distribution

To extract the number of signal events present in the data set, the  $m_{ES}^{B^0}$  distribution is fit with an empirically derived function. By integrating that function over the domain of the distribution and dividing that value by the histogram's bin width, the number of signal events present in the  $m_{ES}^{B^0}$  distribution is obtained. If that value is then divided by the signal efficiency (0.0147), the total number of signal events present in the entire data set is determined.

If the distribution of an observable,  $\xi$ , is given by  $f(\xi)$ , and the observable is plotted in a histogram with  $M$  bins of width  $\Delta\xi$  (defined in Equation 4.1) then the total number of events present in the distribution is given by Equation 4.2.

$$\Delta\xi \equiv \frac{\xi_{max} - \xi_{min}}{M} \quad (4.1)$$

$$\mathcal{N}_{sig} = \sum_{j=0}^M f(\xi + (\Delta\xi)j) \quad (4.2)$$

Now, consider the definition of the Riemann Integral given in Equation 4.3.

$$\begin{aligned} \int_{x_{min}}^{x_{max}} f(x) dx &= \lim_{M \rightarrow \infty} \sum_{j=0}^M f\left(x_{min} + \left(\frac{x_{max} - x_{min}}{M}\right)j\right) \left(\frac{x_{max} - x_{min}}{M}\right) \\ &= \lim_{\Delta x \rightarrow 0} \sum_{j=0}^{\infty} f(x_{min} + j\Delta x) \Delta x \end{aligned} \quad (4.3)$$

A simple, discrete approximation to Equation 4.3, such as those used to numerically evaluate definite integrals, is given in Equation 4.4.

$$\int_{x_{min}}^{x_{max}} f(x) dx \approx \sum_{j=0}^{\frac{x_{max}-x_{min}}{\Delta x}} f(x_{min} + j\Delta x) \Delta x \quad (4.4)$$

Thus, the relationship between the number of events in a distribution described by a function  $f$ , and the integral of that function is made clear when the right-hand side of Equation 4.2 is multiplied by  $\frac{\Delta\xi}{\Delta\xi}$ . Equation 4.5 shows this relationship.

$$\begin{aligned} \mathcal{N}_{sig} &= \sum_{j=0}^M f(\xi + (\Delta\xi)j) \\ &= \frac{\Delta\xi}{\Delta\xi} \sum_{j=0}^M f(\xi + (\Delta\xi)j) \\ &\approx \frac{1}{\Delta\xi} \int_{\xi_{min}}^{\xi_{max}} f(\xi) d\xi \end{aligned} \quad (4.5)$$

The signal and the background  $m_{ES}^{B^0}$  distributions have distinct enough shapes that the total fit function can be made from the sum of two functions: a function that describes the signal distribution and a function that describes the background distribution. This is extraordinarily convenient as it allows for the simple extraction of  $\mathcal{N}_{sig}$ , the number of signal events from a distribution containing both signal and background events. The signal distribution is discussed in Section 4.1.2 while the background distribution is discussed in Section 4.1.3. Keep in mind; however, these discussions are of the ideal case, when both distributions are well described by a function, both distributions are observable (i.e. the ratio of background events to signal events is not arbitrarily large), and both distributions have enough entries<sup>1</sup> so that each shape (i.e. the signal shape and background shape) is easily recognized—at least by the fitter if not the naked eye.

---

<sup>1</sup>For example, if fifty random events are generated according to a Gaussian distribution, the observed distribution may not be able to be fit very well with a Gaussian (i.e.  $\chi^2/ndf$  will be large). On the other hand, if one-hundred thousand random events are generated according to a Gaussian it is almost certain the resulting distribution will have a fit with a  $\chi^2/ndf$  approximately equal to unity.

In order to find good and accurate values for the various fit parameters (i.e. the free parameters in the fit functions) the signal events are temporarily separated from the background events using the MC Truth information and they are plotted in separate histograms. These distributions are fit with their respective fit functions and the values of the fit parameters are determined by a curve fitting algorithm. Once both distributions have been fit, the total distribution is plotted and fit with the total fit function—the sum of the background fit function and the signal fit function. During this final fit, only two parameters are left variable: the parameters that control the amplitude of each function. (NB The Monte Carlo is expected to simulate the actual data to such an extent that the only parameters that will need to be adjusted while fitting the real data is the amplitude of each fit function!) This final fitting process is discussed in Section 4.1.4.

#### 4.1.1 The Fitting Scheme: Minuit

In order to fit the  $m_{\text{ES}}^{B^0}$  distribution, the data is plotted in a histogram using `Root`. The minimization package `Root` uses to fit histograms is `Minuit`, a package originally written in FORTRAN by Fred James as a part of `PACKLIB`. It was converted to a C++ class by Rene Brun.[4] The minimization algorithm is `MIGRAD`, a variable-metric method with inexact line search, a stable metric updating scheme, and checks for positive-definiteness.[4] This analysis uses a Chi-squared minimization scheme with Poisson errors to determine the best fit. The `FUMLI` Minimization Package is used to minimize the Chi-square function.[4] For more information on the fitting scheme, refer to the `Root User's Guide`.

The “goodness” of the fit will be determined by the  $\chi^2/ndf$  value. If this ratio is close to unity, the fit is declared as good, whereas if this ratio is of order

10 or greater, the fit is decidedly poor.  $\chi^2$  is calculated via the usual statistical definition as is  $ndf^2$ . In order to perform an analysis which directly measures  $\mathcal{B}(B^0 \rightarrow \pi^0\pi^0\pi^0)$ , not only will the significance of the data set need to be around five or greater, the  $\chi^2/ndf$  will need to be close to unity. Otherwise, an upper-bound limit analysis will have to be performed. This type of analysis is described in Section 4.2.

### 4.1.2 Signal

The signal distribution is characterized by a Gaussian peak about the true  $B^0$  mass ( $5.2794 \pm 0.0015$  GeV/ $c^2$  [11]) with a tail that decays less rapidly than a Gaussian would to the left-hand side of the maximum. The empirically derived fit function was first used by the Crystal Ball collaboration and is hence named the crystalball function. It is defined in Equation 4.6.

$$f(x) = \begin{cases} N^{-1} \frac{\left(\frac{n}{|\alpha|}\right)^n e^{\frac{1}{2}\alpha^2}}{\left(\frac{n}{|\alpha|} - |\alpha| - x\right)^n} & \text{if } x \leq m_0 - \sigma\alpha \\ N^{-1} e^{-\frac{1}{2}\left(\frac{m_0 x}{\sigma}\right)^2} & \text{if } x > m_0 - \sigma\alpha \end{cases} \quad (4.6)$$

The crystalball function is piece-wise continuous and has a first derivative that does not necessarily exist about the point  $m = m_0 - \sigma\alpha$ . This discontinuity in the first derivative makes minimization schemes that rely on the first derivative (or any subsequent derivatives) difficult to implement; however, by fixing the parameter  $n$  and insisting it take only integer values, the fitter tends to converge. Nevertheless, with such a complicated function that has so many free parameters, good initial guesses for each parameter are needed in order for MIGRAD to converge rather than terminate

---

<sup>2</sup> $ndf = n - p$  where  $n$  is the number of points to be fitted and  $p$  is the number of free parameters in the fit function.

prematurely and/or return a nonpositive-definite error matrix. Ideally, the signal distribution will resemble the distribution shown in Figure 4.1. Unfortunately, the

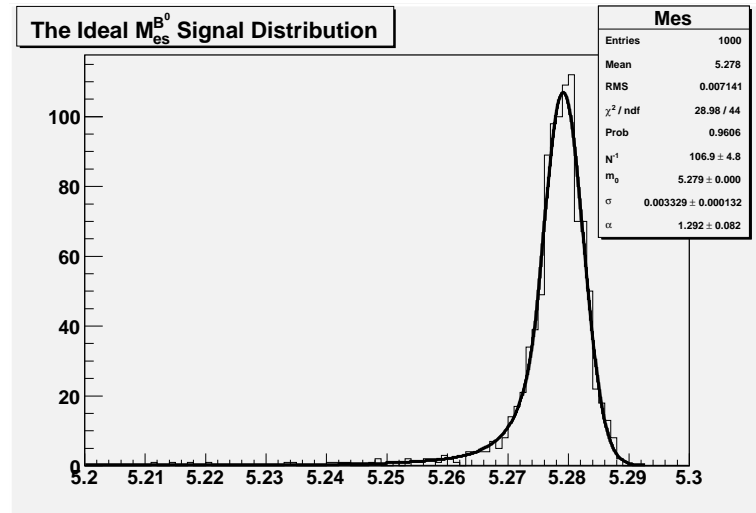
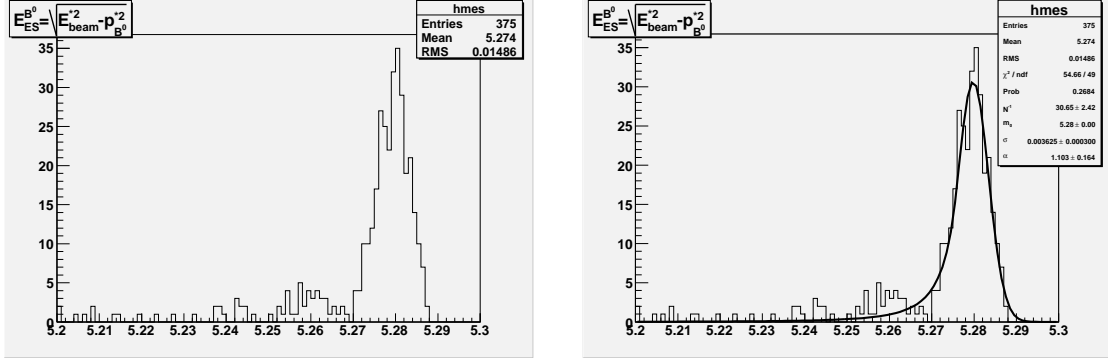


Figure 4.1: An Ideal  $m_{ES}^{B^0}$  Signal Distribution

relatively low branching ratio (recall the prediction of  $\mathcal{O}(10^{-7})$  from Section 2.3) results in low statistics for the Monte Carlo study. Further, the fact that this is a rare decay means that in order to observe any kind of signal distribution on top of the background distribution, aggressive cuts that drastically reduce the size of the data set must be made. In light of these two factors, the signal distribution is less than ideal. In fact, to call the final data set a distribution at all is misleading; rather, it is a histogram with two entries at  $5.24407 \text{ GeV}/c^2$ . This is the first hint that an upper limit analysis will be a more realistic goal than making a direct measurement, and it is a *very* suggestive hint! If renormalization is ignored, and all of the signal events are plotted, the distribution is indeed well defined by a crystalball function, and when fit, a  $\chi^2/ndf$  of 1.12 is achieved. These plots are shown in Figure 4.2. The fact remains however, that this distribution would be completely hidden within the background distribution and thus, observing the

signal distribution is impossible given the current size of the *BABAR* data set.



(a) The Monte Carlo  $m_{ES}^{B^0}$  signal distribution after the BPi0Pi0Pi03body Skims are applied.

(b) The Monte Carlo  $m_{ES}^{B^0}$  signal distribution fit with a crystalball function after the BPi0Pi0Pi03body Skims are applied.

Figure 4.2: Monte Carlo  $m_{ES}^{B^0}$  Signal Distribution

### 4.1.3 Background

The background distribution is characterized by a maxima about the true  $B^0$  mass to which the distribution slowly grows from its left-hand limit. As the right-hand limit is approached, the distribution falls off sharply to zero at some critical point<sup>3</sup>  $m_{ES}^{B^0} = m_0$ , and few, if any, entries are seen to the right of this critical value. The empirically derived fit function is known as the argus function and like the crystalball function, it is piece-wise continuous. Unlike the crystalball however, its first derivative never exists at the point  $m = m_0$ . The argus function is given by Equation 4.7.

$$f(x) = \begin{cases} N^{-1}x \left(1 - \left(\frac{x}{m_0}\right)^2\right)^p e^{c\left(1 - \left(\frac{x}{m}\right)^2\right)} & \text{if } x \leq m_0 \\ 0 & \text{if } x > m_0 \end{cases} \quad (4.7)$$

<sup>3</sup>This  $m_0$ , unlike the  $m_0$  parameter in the crystalball function, is *not* the  $B^0$  mass.

The potential problems caused by the discontinuity in the first derivative for the fit algorithm is avoided by fitting a *sub-section* of the distribution that is defined as the data with  $m < m_0$  with the non-zero part of the argus function. An ideal background distribution (i.e. a well behaved background distribution) will closely resemble Figure 4.3. As discussed in Section 4.1.2, the fact that  $B^0 \rightarrow \pi^0\pi^0\pi^0$  is a

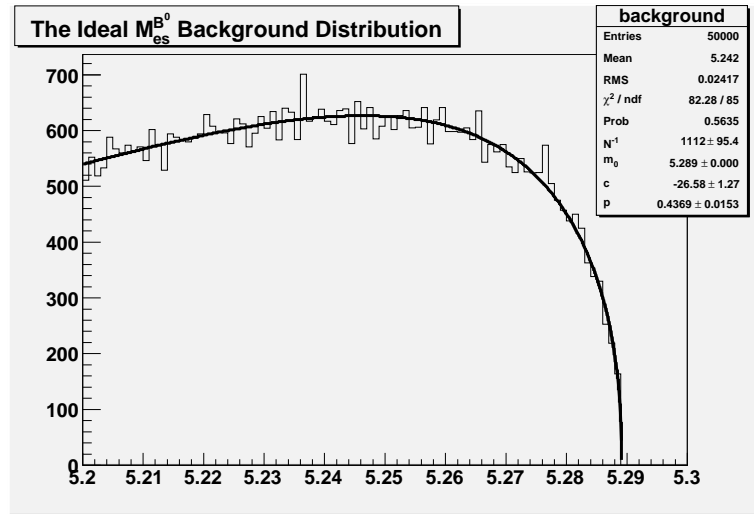


Figure 4.3: A Well-Behaved  $m_{ES}^{B^0}$  Background Distribution

rare process requires a very aggressive set of cuts to be applied. As such, *all* of the background is eliminated from the data set and so there is no distribution left to fit. However, if the  $m_{ES}^{B^0}$  background distribution is examined prior to making any cuts, a distribution that qualitatively seems well described by an argus function is observed as seen in Figure 4.4. Nevertheless, a quantitative analysis of the goodness of the fit reveals a  $\chi^2/ndf$  value of about 16. This suggests the distribution is *not* well described by an argus function or that the fluctuations in the distribution about its mean behavior (an argus function) are large. Large fluctuations are normally a sign of an insufficient amount of data. In the end, however, whether or not the fit is good is moot as the final set of cuts remove all of the background anyway. These

two points combined, further suggest a direct measurement must be abandoned for an upper-limit measurement.

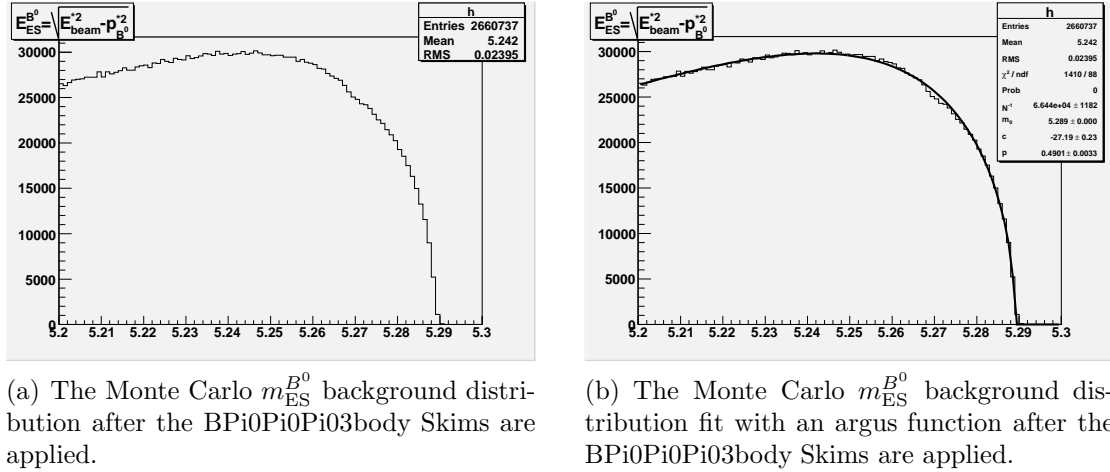


Figure 4.4: Monte Carlo  $m_{ES}^{B^0}$  Background Distribution

#### 4.1.4 The Complete $m_{ES}^{B^0}$ Distribution

At this stage, ideally, both the signal distribution and the background distribution of the  $m_{ES}^{B^0}$  observable are fit with their respective functions and the two histograms can be combined so that the total distribution may be fit with the total fit function—a linear combination of a crystalball function and an argus function. When the total distribution is fit, all of the parameters are set to the values established by the initial fits and are input into the new fit function as constants rather than parameters. The only two *free* parameters are  $N_{argus}^{-1}$  and  $N_{CB}^{-1}$ . These parameters control the amplitude of the argus and crystalball functions respectively. Ideally, the total  $m_{ES}^{B^0}$  distribution will resemble Figure 4.5. However, at this point, it is quite clear that the data set being dealt with is anything but ideal and thus, an

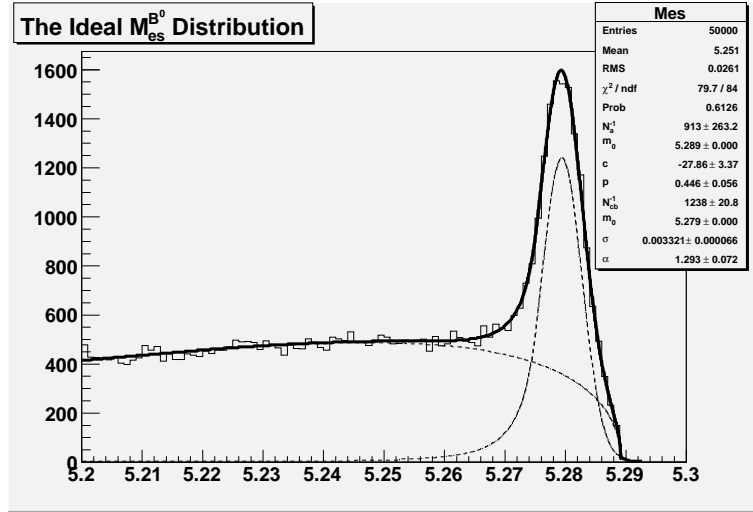


Figure 4.5: The ideal  $m_{\text{ES}}^{B^0}$  distribution is fit with a function (solid line) that is the sum of a crystalball function (dashed line) and an argus function (dotted line).

upper limit analysis is the only analysis that can be performed. Therefore, instead of measuring  $\mathcal{B}(B^0 \rightarrow \pi^0\pi^0\pi^0)$  to be some number  $\alpha \pm \beta \times 10^{-\gamma}$ , the goal is to say that  $\mathcal{B}(B^0 \rightarrow \pi^0\pi^0\pi^0)$  is *less than* some number  $a \times 10^{-b}$  with a confidence level of 90%. All that must be done now is figure out how to determine the value  $a \times 10^{-b}$  at a particular confidence level.

## 4.2 Calculating Confidence Levels for an Upper Limit Analysis

The *BABAR* Collaboration has developed an applet that can be used to calculate confidence limits on branching ratios for rare decays such as  $B^0 \rightarrow \pi^0\pi^0\pi^0$ . [3] The calculator, shown in Figure 4.2, takes seven arguments and returns three. It takes the following arguments: the number of observed signal events, the number of

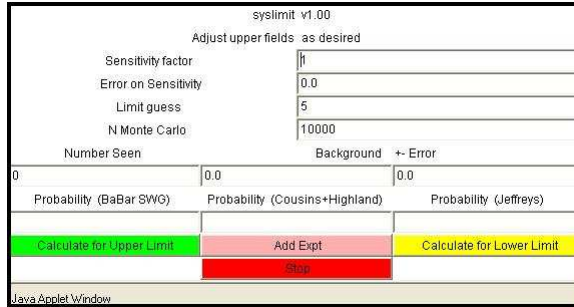


Figure 4.6: The Confidence Level Calculator

observed background events, the error (given as an absolute value) on the number of background events, the overall sensitivity and its error, the number of Monte Carlo events to be used in the evaluation (*unrelated* to the number of MC events used to optimize the cuts that will be placed on the data set), and a guess for the limit. The calculator then returns the probability that the number of signal events will be greater than the inputted guess for the limit. By adjusting the guess, a probability of 0.1 can be obtained. This guess value corresponds to a confidence level of 90%, that is to say, the value input as the guess is the upper limit for the number of signal events at the 90% confidence level. Finally, by dividing the limit guess by the efficiency with which the signal is preserved (in this case,  $\frac{1}{68}$ ), the upper limit for the total number of signal events present in the data set *before* any cuts are made is determined. Finally, division by the total number of  $B^0$  decays analyzed will yield an upper limit for the branching ratio. The three arguments returned by the calculator are three different probabilities each of which is calculated with a different algorithm. The three algorithms are the *BABAR* SWG, Cousins and Highland, and Jeffreys algorithms. This analysis uses the *BABAR* SWG probability.

### 4.2.1 Theory

Given a Poisson process<sup>4</sup> in which  $n$  events are observed and used to set limits on an observable  $R$ , the probability equalities are given by Equation 4.8 and Equation 4.9.

$$\sum_{r=0}^n P(r; \mu+) = a \quad (4.8)$$

$$\sum_{r=0}^{n-1} P(r; \mu-) = 1 - a \quad (4.9)$$

In the above equations,  $\mu\pm$  refers to the upper and lower limits on the mean ( $\mu$ ), given by  $RS$ <sup>5</sup> where  $S$  is the sensitivity, at the  $(1 - a)$  confidence level.[3] In the case where there are uncertainties on the quantities  $S$  and  $b$ , where  $b$  is the expected background, an exact calculation is not possible; however, a Monte Carlo technique can be used. The Monte Carlo process requires a trial value for the signal value. Then, a background drawn from a Gaussian of mean  $b$  and standard deviation  $\sigma_b$ , the error on the estimated background, is added to the the trial signal value. This value is then multiplied by a sensitivity drawn again from a Gaussian, this time with mean  $S$  and standard deviation  $\sigma_S$ , the error on the sensitivity. This value is then used as the mean of a Poisson distribution from which  $r$  events are generated. The probability  $a$  ( $1 - a$ ) is generated by counting the fractional number of times  $r$  is less than or equal to (greater than)  $n$ . [3]

---

<sup>4</sup>A Poisson process is one in which the probability of obtaining exactly  $n$  events in  $N$  trials is given by  $P_p(n|N) = \frac{N!}{n!(N-n)!} p^n (1-p)^{N-n}$  where  $p$  is the probability of an event occurring. As  $N \rightarrow \infty$ ,  $P(n) \rightarrow \frac{\nu e^{-\nu}}{n!}$  where  $\nu = Np$ . [16]

<sup>5</sup>Taking  $\mu = RS$  assumes that the factor for efficiency etc., collectively refers to as  $S$ , is known exactly (i.e.  $\sigma_S \equiv 0$ ). [3]

## 4.2.2 A Fictitious Example

Suppose 10,000 signal events of some rare process are observed and after applying a set of cuts, MC truth information reveals 50 signal events remain. The efficiency is calculated as  $\frac{50}{10000}$  or 0.5%. Further, the relative error on the efficiency is  $\frac{1}{\sqrt{10000}} = 0.01$  or 1% (of 0.05%).

In the “real data”, consisting of 5 million events, after the same set of cuts are applied, five events remain, four of which are thought to be signal (thus the error on the sensitivity is 1% of 4 or 0.04). Suppose further that a separate study has shown that the systematic error on the background should be taken as one half of one event. Then, the calculator takes the values shown in Table 24 and returns the value shown at the bottom of the same table. From the output of the

| Input                        |        |
|------------------------------|--------|
| $S$                          | 1.0    |
| $\sigma_S$                   | 0.04   |
| Limit Guess                  | 6      |
| $N$ Monte Carlo              | 50 000 |
| $\mathcal{N}_{sig}$          | 4      |
| $\mathcal{N}_{bkg}$          | 1      |
| $\sigma_{\mathcal{N}_{bkg}}$ | 0.5    |
| Output                       |        |
| $P$                          | 0.18   |

Table 24: A Fictitious Upper Limit Calculation

confidence limit calculator, it can be concluded that  $\frac{6}{0.005} = 1,200$  is the upper limit for the number of signal events originally present in the data set with a confidence level of 82%. Thus, the branching ratio for this rare decay is less than or equal to  $\frac{1200}{5000000} = 2.5 \times 10^{-4}$  at the 82% confidence level.

# CHAPTER 5

## Data Analysis

With the Monte Carlo analysis complete, the *BABAR* data set can be analyzed. First, the `NonCharm3BodyUserApp` will run over the data sets with the `B0Pi0Pi0Pi03body Skim` applied. Once `NonCharm3BodyUserApp` creates the `Root` files that contain all the observable values, the optimized cuts, determined in Section 3.3.3, can be applied to the data set. Next, by examining the entries left in the  $m_{\text{ES}}^{B^0}$  distribution, the number of signal events and background events left in the data set can be determined. Finally, the necessary information can be entered into the upper limit calculator and an upper limit for  $\mathcal{B}(B^0 \rightarrow \pi^0\pi^0\pi^0)$  will be determined.

### 5.1 The *BABAR* Data Set

The *BABAR* data set is separated into four<sup>1</sup> subsets of data: Runs 1–4. Each run corresponds to a year of data taking. For example, in the year 2006, Run 5 is being conducted. Each of these runs is divided into subsets: on peak and off peak data. On peak data corresponds to data collected by the *BABAR* detector while the

---

<sup>1</sup>To date, five sub sets of data exist: Runs 1–5; however, skims have yet to be applied to the Run 5 data.

$e^+ e^-$  beam energy is tuned to the  $\Upsilon(4S)$  resonance energy, while off peak data corresponds to data compiled at all other beam energies. Hence, for this study, only on peak data is analyzed.

The number of  $e^+ e^-$  collisions that occur during data taking is a well known number. The number of events can be calculated by knowing the beam’s luminosity and the time interval over which data is collected. However, recall from Section 3.1 that not every  $e^+ e^-$  collisions results in a  $\Upsilon(4S)$ , let alone a  $B^0\bar{B}^0$  pair. The number of  $B^0\bar{B}^0$  pairs produced is a very important quantity in nearly all analyses and hence this number must be calculated as accurately and as precisely as possible. The number of  $B^0\bar{B}^0$  pairs produced is known as  $B$  counting, and the  $B$  counting analysis was one of the first *BABAR* analyses performed. Now, the number of  $B^0\bar{B}^0$  events is documented on the *BABAR* website. The properties of Runs 1–4 are summarized in Table 25

| Data set    | Collections | Runs  | $B$ -counting | Luminosity ( $pb^{-1}$ ) |
|-------------|-------------|-------|---------------|--------------------------|
| Run1-OnPeak | 27          | 3 277 | 22 425 983.7  | 20 718.463               |
| Run2-OnPeak | 59          | 5 150 | 67 319 702.9  | 60 923.701               |
| Run3-OnPeak | 35          | 3 564 | 35 697 393.2  | 32 276.575               |
| Run4-OnPeak | 84          | 6 573 | 111 375 232.7 | 101 072.680              |
| Total       | –           | –     | 236 818 312.5 | –                        |

Table 25: A Summery of Runs 1–4

## 5.2 The Skim-Reduced Data Set

The *BABAR* data set contains data on more than two-hundred thirty million  $B^0\bar{B}^0$  decays. Therefore, before any analysis of the data set can be performed, skims must be applied to the data just as they were applied to the MC data. Again, the skim

used is the Release 16.1.3 B0Pi0Pi0Pi03body Skim. Table 26 shows the effect of applying the skims on the Runs 1–4 data sets.

|  | Run 1       | Run 2       | Run 3       | Run 4         |
|--|-------------|-------------|-------------|---------------|
| Stage  |             |             |             |               |
| Initial  | 537 174 990 | 866 976 254 | 447 426 127 | 3 025 423 292 |
| Skims  | 11 207      | 282 033     | 146 111     | 922 854       |
| Skims $\left\langle \frac{N_{events}}{N_{recon.}} \right\rangle$ | 2.94        | 3.06        | 3.03        | 3.16          |

Table 26: Number of Events in the Runs 1–4 Data Sets after Skims

The total data set (i.e. Runs 1–4 combined) is still very large; as such, it would take too long for `NonCharm3BodyUserApp` to run over the entire data set in one go. That is to say, the xlong queue at SLAC will not allow a job to run long enough for all of the data to be analyzed and stored in ntuples. As a result, the data set is trimmed into about fifteen subsets so that each job can run in the allowed time. Thus, fifteen `.tcl` files are prepared as arguments for `NonCharm3BodyUserApp` and the jobs are then submitted to the xlong queue at SLAC.

When the `Root` files are returned, they are merged to create four large ntuples, one for each run. Analysis macros are used to analyze the data by applying cuts, making histograms of interesting quantities, and counting the number of events and event reconstructions that are stored in the `Root` ntuples. One such set of histograms shows the distributions of the Cut variables. These histograms are shown in Figure 5.1. Histograms are also created to show the distributions of the cut variables for Runs 1–4 separately. These plots are shown in Figures 5.2–5.5. Table 27 shows the properties of the Runs 1–4 data sets at various stages during the analysis.

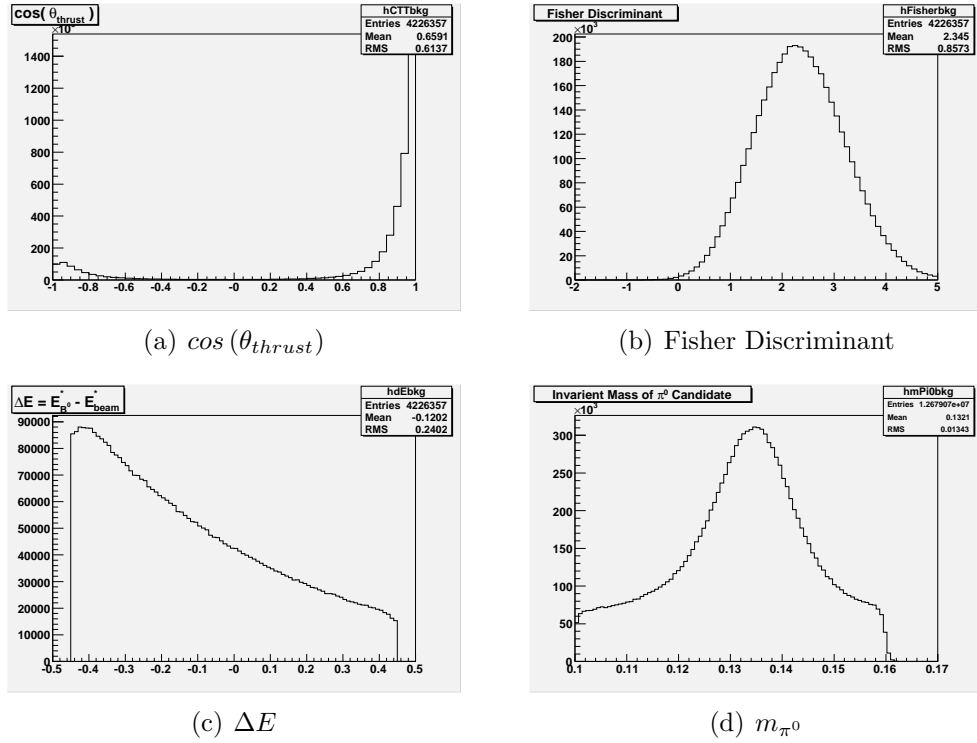
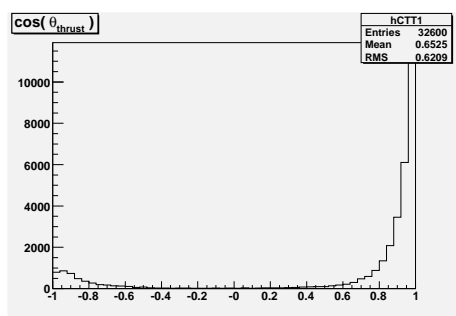


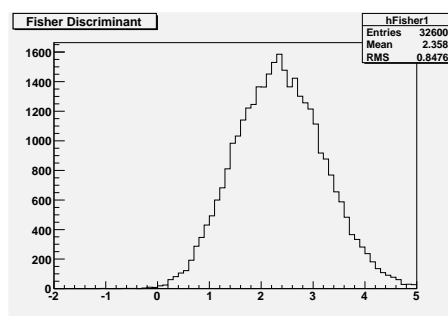
Figure 5.1: The Run 1-4 Distributions of the Cut Observables

|   | Run 1       | Run 2       | Run 3       | Run 4         |
|---|-------------|-------------|-------------|---------------|
| Stage   |             |             |             |               |
| Initial   | 537 174 990 | 866 976 254 | 447 426 127 | 3 025 423 292 |
| Skims   | 11 207      | 282 033     | 146 111     | 922 854       |
| Skims $\langle \frac{\mathcal{N}_{events}}{\mathcal{N}_{recon.}} \rangle$   | 2.94        | 3.06        | 3.03        | 3.16          |
| Cuts 2a   | 0           | 4           | 1           | 4             |
| Cuts 2a $\langle \frac{\mathcal{N}_{events}}{\mathcal{N}_{recon.}} \rangle$ | –           | 1           | 1           | 1             |
| Cuts 2b   | 0           | 0           | 0           | 2             |
| Cuts 2b $\langle \frac{\mathcal{N}_{events}}{\mathcal{N}_{recon.}} \rangle$ | –           | –           | –           | 1             |

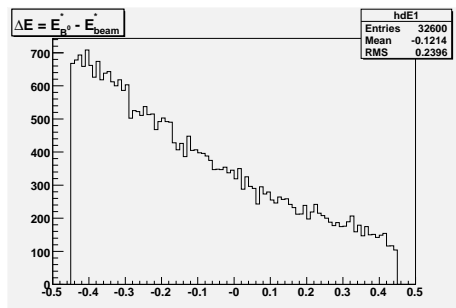
Table 27: Number of Events in the Runs 1–4 Data Sets at Various Stages of Background Reduction



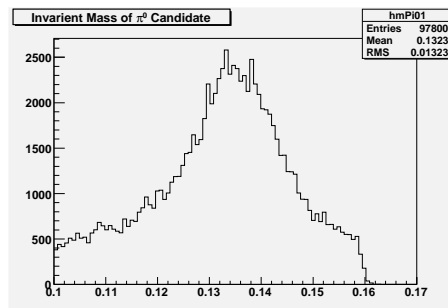
(a)  $\cos(\theta_{thrust})$



(b) Fisher Discriminant

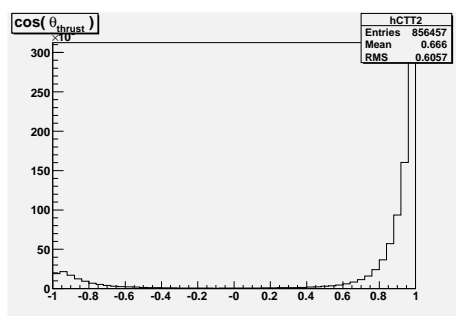


(c)  $\Delta E$

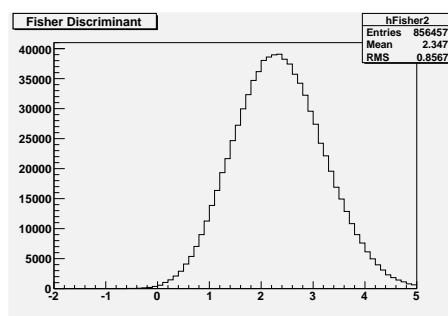


(d)  $m_{\pi^0}$

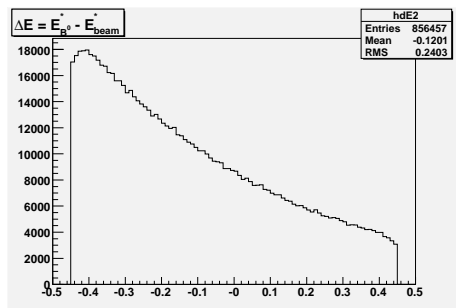
Figure 5.2: The Run 1 Distributions of the Cut Observables



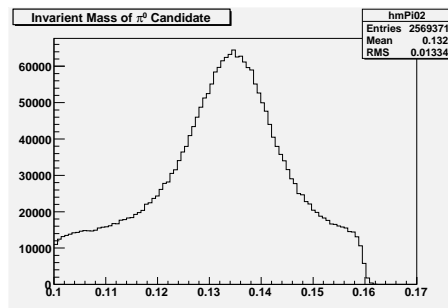
(a)  $\cos(\theta_{thrust})$



(b) Fisher Discriminant

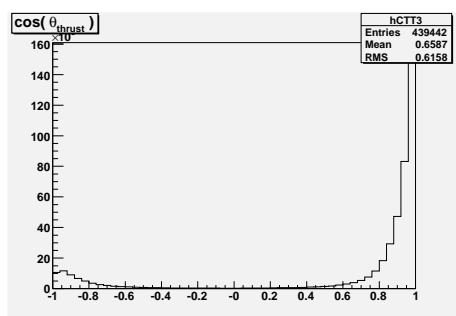


(c)  $\Delta E$

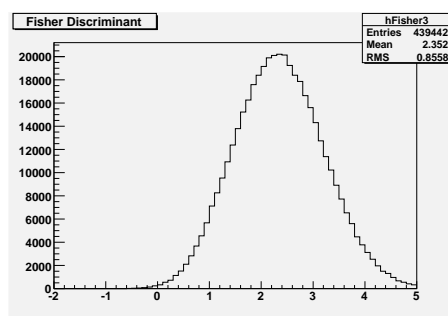


(d)  $m_{\pi^0}$

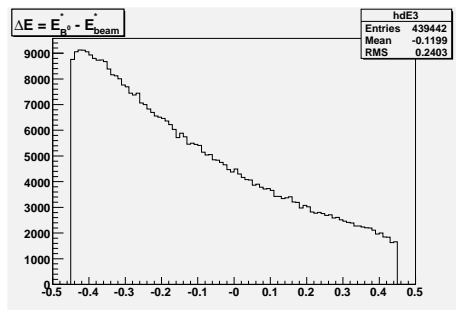
Figure 5.3: The Run 2 Distributions of the Cut Observables



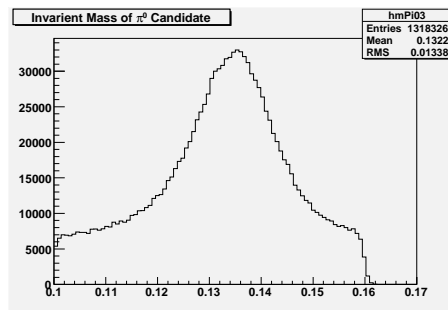
(a)  $\cos(\theta_{thrust})$



(b) Fisher Discriminant

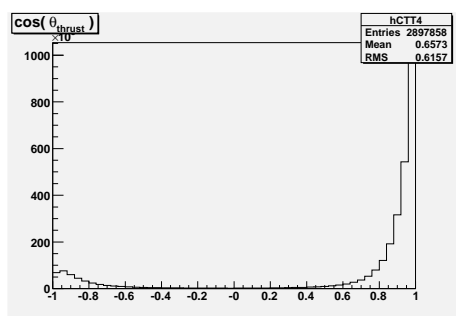


(c)  $\Delta E$

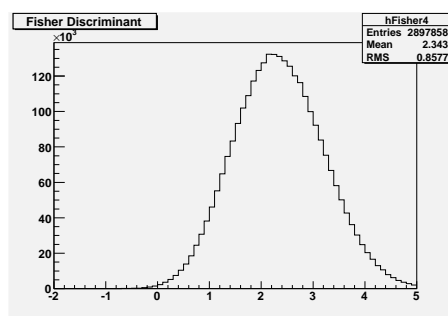


(d)  $m_{\pi^0}$

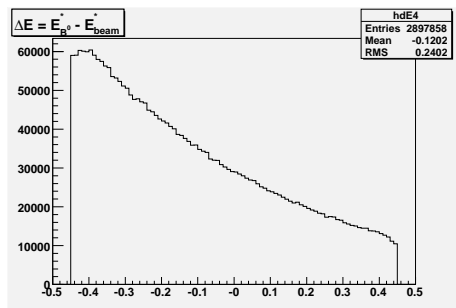
Figure 5.4: The Run 3 Distributions of the Cut Observables



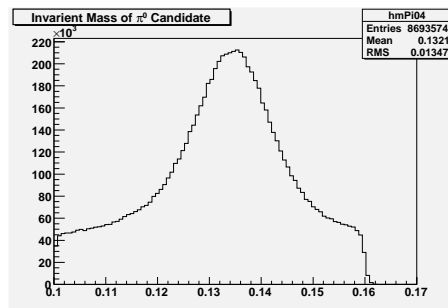
(a)  $\cos(\theta_{thrust})$



(b) Fisher Discriminant



(c)  $\Delta E$



(d)  $m_{\pi^0}$

Figure 5.5: The Run 4 Distributions of the Cut Observables

### 5.3 The Cut-Reduced Data Set

Before an upper limit is calculated, a new data set must be created, the contents of which is only the events that survive the cuts. Again, analysis macros are used to write all the information stored in the ntuples about events that pass the cuts into new Root files. When the optimal and tightest cuts are applied, Cuts 2b, all that survives from the initial 4.88 *billion* events analyzed are *two* events (see Table 27).

If the optimal set of cuts (i.e. Cuts 2b) are applied *individually* (i.e. applying only a cut on  $\cos(\theta_{thrust})$ , for example), the size of the data set is reduced to order  $10^3$  at best. While the constraints are never individually applied, it is interesting to see what the effect of doing so would be. Table 28 is a summary of a mini analysis that studies just that, the effect of a cut applied to a single variable on the data set. From Table 28, it is seen that the cut on the Fisher Discriminant removes the

|                         |   | Run 1  | Run 2   | Run 3   | Run 4   |
|-------------------------|---|--------|---------|---------|---------|
| Cut On                  | –   |        |         |         |         |
| –                       | $N_{events}$<br>$\left\langle \frac{N_{recon.}}{N_{sig}} \right\rangle$ | 11 207 | 282 033 | 146 111 | 922 854 |
| $\Delta E$              | $N_{events}$<br>$\left\langle \frac{N_{recon.}}{N_{sig}} \right\rangle$ | 594    | 15 050  | 7 631   | 49 996  |
| Fisher                  | $N_{events}$<br>$\left\langle \frac{N_{recon.}}{N_{sig}} \right\rangle$ | 15     | 774     | 428     | 4 115   |
| $m_{\pi^0}$             | $N_{events}$<br>$\left\langle \frac{N_{recon.}}{N_{sig}} \right\rangle$ | 914    | 23 439  | 11 995  | 75 582  |
| $\cos(\theta_{thrust})$ | $N_{events}$<br>$\left\langle \frac{N_{recon.}}{N_{sig}} \right\rangle$ | 127    | 3 954   | 2 035   | 12 712  |

Table 28: The Effect of Cuts on Runs 1–4

most background while the cut on  $m_{\pi^0}$  removes the fewest events from the data set. Kinematic variables do a good job reducing the number of reconstructions of

the same event, while the event shape variables do not.

Finally, before the upper limit is calculated, a study of the properties of the two remaining events is performed. The study obtains the values of various observables for the events that passed all the cuts as well as which Run the events came from. Table 29 summarizes the results of this study. The two events are

|                             | Event 1     | Event 2     |
|-----------------------------|-------------|-------------|
| Run Number                  | 4           | 4           |
| Event Number                | 218215-0    | 725498-0    |
| $\Delta E$ (GeV)            | -0.00462479 | -0.00462479 |
| Fisher                      | -0.0219509  | -0.0219509  |
| $m_{\pi^0_1}$ (MeV/ $c^2$ ) | 0.131409    | 0.131409    |
| $m_{\pi^0_2}$ (MeV/ $c^2$ ) | 0.133599    | 0.133599    |
| $m_{\pi^0_3}$ (MeV/ $c^2$ ) | 0.132792    | 0.132792    |
| $\cos(\theta_{thrust})$     | -0.0662942  | -0.0662942  |
| $m_{ES}^{B^0}$              | 5.24407     | 5.24407     |

Table 29: Properties of the Two Signal Events

essentially identical. All of the observables for both events, each coming from Run 4, take the same values, the only difference is their event number. The event number turns out to be the only way to know that these are indeed two unique events.

#### 5.4 An Upper Limit for $\mathcal{B}(B^0 \rightarrow \pi^0\pi^0\pi^0)$

With the most significant cuts applied to the *BABAR* data set, the upper limit for  $\mathcal{B}(B^0 \rightarrow \pi^0\pi^0\pi^0)$  can be calculated. Recall from Section 4.2 that the values needed to make this calculation are: the number of signal events observed, the number of background events observed, the error on the number of background events observed, the overall sensitivity and its error, the number of Monte Carlo events to be used in calculating the limit, and a guess for the signal event limit.

Two events remain in the data set after all of the cuts are applied. During the MC optimization process, the same cut values returned one signal event and zero background events. The study of the properties of the remaining events from the *BABAR* data set revealed the two events were not two different reconstructions of the same  $B^0$  decay (i.e. they were unique events), and hence, the number of signal events is taken as two and the number of background events is taken as zero.

While the number of background events observed is zero, there is still an uncertainty on this value. Strictly speaking, the uncertainty should be calculated by performing a very rigorous and formal analysis on the MC data; however, with no time left to perform this analysis (and it is expected to be a lengthy analysis) and estimate for the uncertainty is made. The error on the number of background events is typically taken as the systematic error in the reconstruction algorithm, the systematic error in data collection process, the errors as a result of imperfections in the *BABAR* particle detector, the uncertainty associated in the  $B$  counting, etc. All of these errors and uncertainties are *extremely* small relative to the number of events analyzed, and other studies in which serious background error analysis *is* performed typically use errors of about one half of one event. Hence, the error on the number of background events is taken as one half of one event.

For the overall sensitivity, one can choose between two values. The sensitivity can either be set to one, and an upper limit guesses of about five can be made, or, the sensitivity can be set to the signal efficiency derived during the MC analysis. For this analysis, the signal efficiency is  $\frac{1}{68} \approx 0.0147$ . If the sensitivity is taken as 0.047, then an upper limit guesses of order one hundred must be used. Either method will return exactly the same result. If the sensitivity is taken as one and the returned upper limit is  $L_1$ , then  $L_1$  is related to  $L_2$ , the upper limit returned

when the sensitivity is set to 0.047, through Equation 5.1.

$$\frac{L_1}{\frac{1}{68}} = 68L_1 = L_2 \quad (5.1)$$

The error on the sensitivity is calculated using the relative error determined from the Monte Carlo study. Recall that for a Poisson process, the relative error scales as  $\frac{1}{\sqrt{N}}$ , where  $N$  is the *total* number of signal events. In the Monte Carlo study, the total number of events,  $N$ , is 351. Hence, the relative error is  $\frac{1}{\sqrt{351}} \approx 0.0534$  or, 5.34%. To get the absolute error, the relative error is multiplied by the “best value” for the number of signal events, in this case, two. Hence, the error on the sensitivity is 5.34% of 2 or 0.107.

The number of Monte Carlo events used to calculate the limit is initially set to 10,000; however, as the 90% confidence level is approached, since the calculation is based on a statistical average, the number of MC events is increased to as many as ten million so that a more precise value for the upper limit is obtained. In other words, by using more Monte Carlo events to calculate the limit, the statistical fluctuations in the returned value decreases.

The upper limit guess is initially taken as five and it is adjusted until the calculator returns a *BABAR* AWG probability of 0.1. This corresponds to a 10% chance that the number of events in the data set is larger than the upper limit guess and hence, it is the 90% confidence limit for the upper bound value of the number of signal events originally present.

After adjusting the limit guess parameter a bit, the calculator was run three times with the inputs shown in Table 30. All three times the calculator was run, the same probability was returned: 0.1. Using Equation 5.1, the upper limit for the number of signal events originally present in the data set is found to be 369.92. Finally, using the results of  $B$  Counting, the total number of  $B^0$  decays are summed,

| Input               |            |
|---------------------|------------|
| $S$                 | 1.0        |
| $\sigma_S$          | 0.1067     |
| Limit Guess         | 5.44       |
| $N$ Monte Carlo     | 10 000 000 |
| $\mathcal{N}_{sig}$ | 2          |
| $\mathcal{N}_{bkg}$ | 0          |

Table 30: Limit Calculator Inputs

the ratio of  $B^0 \rightarrow \pi^0\pi^0\pi^0$  decays to total  $B^0$  decays is taken, and the upper limit for the branching ratio is found to be  $\frac{369.92}{236818312.5}$ , or  $1.56 \times 10^{-6}$ . Hence, the upper limit for  $\mathcal{B}(B^0 \rightarrow \pi^0\pi^0\pi^0)$  is  $1.56 \times 10^{-6}$  at the 90% confidence level.

# CHAPTER 6

## Conclusions

### 6.1 Results

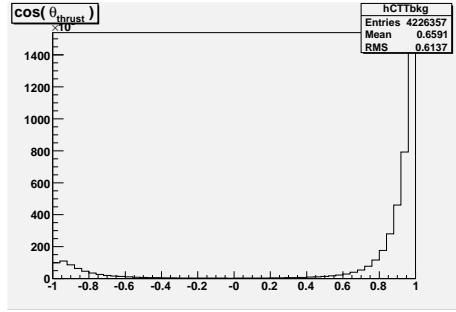
The upper limit for  $\mathcal{B}(B^0 \rightarrow \pi^0\pi^0\pi^0)$  is found to be  $1.56 \times 10^{-6}$  at the 90% confidence level. This value is consistent with the naive estimates made in Section 2.3.2, where the order of magnitude for this branching ratio was determined to be  $10^{-7}$ . This calculation was based on the data collected from the more than three billion  $e^+ e^-$  collisions taking place at SLAC's asymmetric  $B$  Factory. The billions of collisions, taking place at the  $\Upsilon(4S)$  resonance energy, yield a data set with information on over two-hundred million  $B^0\bar{B}^0$  decays. Data from these  $B$  meson decays is collected by the *BABAR* detector and is reconstructed using algorithms written by the *BABAR* collaboration.

A great deal of emphasis was placed on the Monte Carlo data analysis with which the optimal set of cuts on the total energy ( $\Delta E$ ), the Fisher Discriminant, the neutral pion invariant mass ( $m_{\pi^0}$ ), and the cosine of theta thrust ( $\cos(\theta_{thrust})$ ) were determined. While much discussion of the MC data takes place, there is no mention of "how good" this data is or, how well it represents nature. That is to say, are the probability distribution functions that are used to generate the data accurate models of what is seen in nature? Figure 6.1 compares the MC distribution

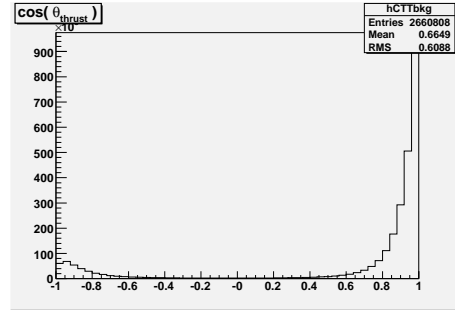
in the aforementioned observables to the distributions produced from the Runs 1–4 data. This figure suggests that the pdf’s do indeed do a good job of modeling nature; however, the size of the MC data set is small when compared to the size of the *BABAR* data set. The size of the MC data set is limited by the number of  $c\bar{c}$  events that pass the B0Pi0Pi0Pi03body Skims.

## 6.2 Summery of the Analysis

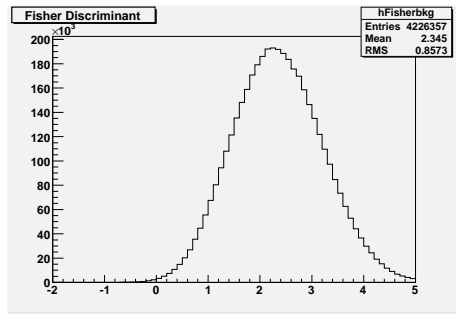
This analysis was quite involved and included a great many steps in order to set an upper limit on  $\mathcal{B}(B^0 \rightarrow \pi^0\pi^0\pi^0)$ . First, the proper packages needed to extract, analyze, and sort the data into ntuples were sought. The package `NonCharm3BodyUser` was chosen for the analysis and it was compiled and linked. Next, MC data was obtained. Four types of MC data were compiled into a master MC data set. The four types were  $e^+ e^- \rightarrow c\bar{c}$ ,  $e^+ e^- \rightarrow u\bar{u}/d\bar{d}/s\bar{s}$ ,  $e^+ e^- \rightarrow B^\pm$ , and  $e^+ e^- \rightarrow B^0\bar{B}^0$ . Before the data could be analyzed, the data needed to be skimmed to reduce the overwhelming size of the data set. The B0Pi0Pi0Pi03body Skims were used. Next, the data set had to be renormalized so that the ratios between the decay types were the same as those observed in nature. Once the data sets were renormalized, the `Root` ntuples produced by `NC3BU` could be analyzed in order to find the cuts on  $\Delta E$ , the Fisher Discriminant,  $m_{\pi^0}$ , and  $\cos(\theta_{thrust})$  that maximized the significance of the data set. This was a two step process that involved a set of “loose” cuts and then an iterative procedure that used the results from  $i^{th} - 1$  iteration as the seed or best guess for the  $i^{th}$  iteration. When the optimal cuts were determined, the real data sets (Runs 1–4) were skimmed using the B0Pi0Pi0Pi03body Skims, analyzed, and sorted into ntuples with `NC3BU`. After the data was stored into `Root` compatible files, analysis macros created a data



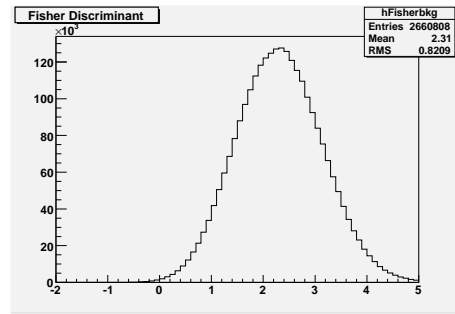
(a)  $\cos(\theta_{thrust})$  Run 1-4



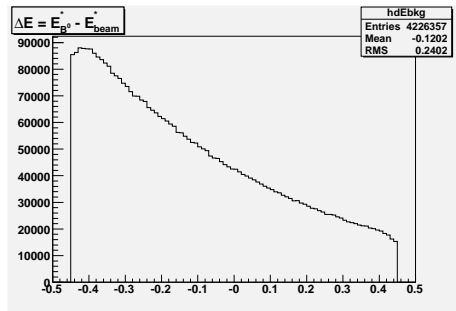
(b)  $\cos(\theta_{thrust})$  MC Background



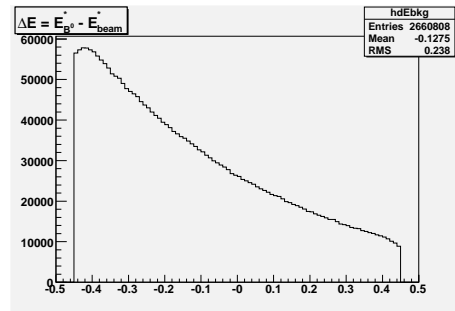
(c) Fisher Discriminant Run 1-4



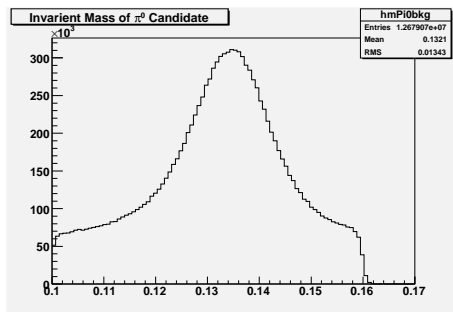
(d) Fisher Discriminant MC Background



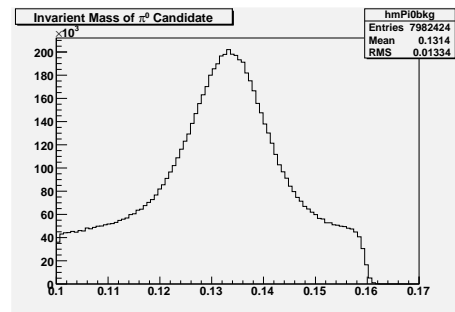
(e)  $\Delta E$  Run 1-4



(f)  $\Delta E$  MC Background



(g)  $m_{\pi^0}$  Run 1-4



(h)  $m_{\pi^0}$  MC Background

Figure 6.1: A Comparison of Monte Carlo and Run 1-4 Distributions of the Cut Observables

set containing only those decays that passed the optimal cuts previously derived. That data set had two entries. Using an upper limit calculator program based on the statistics of Poisson processes, an upper limit for the number of signal events originally present in the data set was determined at the 90% confidence level. Finally, with the help of  $B$  counting, the number of signal events was divided by the total number of  $B^0\bar{B}^0$  pairs analyzed and an upper limit for  $\mathcal{B}(B^0 \rightarrow \pi^0\pi^0\pi^0)$  was determined. It was found that  $\mathcal{B}(B^0 \rightarrow \pi^0\pi^0\pi^0)$  is less than  $1.56 \times 10^{-6}$  at the 90% confidence level.

### 6.3 Difficulties with the Analysis

The decay mode  $B^0 \rightarrow \pi^0\pi^0\pi^0$  is a rare decay that suffers from two different types of QCD suppression. First, the decay is color suppressed as a result of an internal virtual  $W^+$  emission. Second, the decay is  $CKM$  suppressed in that the mode's primary diagram calls for the inclusion of the term  $|V_{ub}|$  in the calculation of the branching ratio. Thus, to begin with, the branching ratio is expected to be small,  $\mathcal{O}(10^{-7})$ . What makes the measurement even more difficult is the neutral final state that decays directly to photons does not leave very much information about itself in the *BABAR* detector. Some rare decays, such as  $B^0 \rightarrow K_s^0 K_s^0 K_s^0$  have charged tracks in their final state ( $K_s^0 \rightarrow \pi^+ \pi^-$ ) which allows information from the DIRC and the DCH to be used during the analysis; however,  $B^0 \rightarrow \pi^0\pi^0\pi^0$  is reconstructed from the detection of six photons detected by the EMC which provides no direct coordinate information about the neutral pion. This leads to an abnormally large background distribution and less observables to cut on. For example, in the decay  $B^0 \rightarrow K_s^0 K_s^0 K_s^0$ , each of the six pion masses can be cut on, as can each of the three kaon masses. Also, combinatorial background can be reduced

by demanding that the net charge of the two pions being reconstructed into a  $K_s^0$  is zero, and that they originated from the same place in space-time, which is possible using data taken by the DHC. As a result of all these variables to cut on, plus the ones cut on in this analysis, a great deal of background can be removed from the data set while preserving a larger percentage of the signal.

#### 6.4 Future Work on $B^0 \rightarrow \pi^0\pi^0\pi^0$

Now that an upper limit has been set, the future of  $B^0 \rightarrow \pi^0\pi^0\pi^0$  has many possibilities. The next analysis on this mode should certainly perform a rigorous analysis on the background error and uncertainty. This will allow for a more accurate value of the upper limit presented herein to be reported. Another possibility is a direct measurement of  $\mathcal{B}(B^0 \rightarrow \pi^0\pi^0\pi^0)$  rather than setting an upper limit. In order for this to be achieved, a more sophisticated analysis would need to be performed, perhaps one involving multi-binned fits. Further, a larger MC data set will probably be needed for this analysis. This will involve creating more  $c\bar{c}$  data that passes the B0Pi0Pi0Pi03body Skims. For a direct measurement analysis to advance, more observables will likely need to be involved in the analysis, and the *BABAR* data set will need to have increased substantially.

This analysis leads nicely into another possibility for the future of  $B^0 \rightarrow \pi^0\pi^0\pi^0$ : a resonance analysis. Consider the decay  $B^0 \rightarrow A\pi^0$  with  $A \rightarrow \pi^0\pi^0$ , where  $A$  is some unknown hadron. This analysis can be performed almost immediately using the data that was collected for this analysis. By calculating the invariant mass of two of the three pions, and plotting the distribution, resonances, should they exist, will appear as “spikes” in the distribution. Then, by fitting the distribution, or at the very least, the sharp peak(s) in the distribution, the location(s) of

the center(s) of the peak(s) can be determined and using a table of the properties of various hadrons, the identity of the particle(s) can be determined. Some resonances involving “common” hadrons can be ruled out immediately because they do not simultaneously conserve angular momentum and sign under interchange. One candidate for a resonance that is not immediately ruled out is  $f^0$ .

Finally, since  $B^0 \rightarrow \pi^0\pi^0\pi^0$  is a  $CP$  eigenstate, if the branching ratio is measured, then the amount of  $CP$  violation that occurs for this decay could be studied. This, however, is far in the future and may be beyond the scope of *BABAR*. In fact, it may require a super  $B$  Factory such as the one being planned for in Japan.

Nevertheless, now that the branching ratio has been determined to be less than  $1.56 \times 10^{-6}$  at the 90% confidence level, a benchmark has been set, and the mode  $B^0 \rightarrow \pi^0\pi^0\pi^0$  is ready for further analyses that will hopefully provide even more insight into the behavior of the most elementary particles in this universe.

## REFERENCES

- [1] Amplitude Analysis of the Decays  $B^\pm \rightarrow \pi^\pm \pi^\mp \pi^\pm$ . *BABAR-PUB-05/028*, *SLAC-PUB-11332*. (2005).
- [2] Technical design report. (*BABAR* Collaboration), SLAC-REP-372 (1995).
- [3] Roger Barlow. A Calculator for Confidence Intervals. *Computer Physics Communications*, 149:97, 2002.
- [4] ed. Brun, Rene and ed. Rademakers, Fons. *Root* User's Guide v. 5.12. July 2006, <http://root.cern.ch/root/doc/RootDoc.html>.
- [5] L. L. Chau and W. Y. Keung. *Physical Review Letters*. **53**, 1802(1984).
- [6] Hai-Yang Cheng and Kwei-Chou Yang. Nonresident three-body decays of  $D$  and  $B$  mesons. *Physical Review D*. **66**, 054015(2002).
- [7] P. Coyle and *et al.* Nuclear Instrumentation Methods. **A343**, page 292, 1994.
- [8] R. A. Fisher. *Contributions to Mathematical Statistics*. John Wiley, 1950.
- [9] G. C. Fox and S. Wolfram. *Nuclear Physics*. **B149**, 413 (1979).
- [10] ed. Harrison, P. F. and ed. Quinn, Helen R. The *BABAR* Physics Book: Physics at an asymmetric  $B$  Factory Papers from Workshop on Physics at an asymmetric  $B$  Factory (*BABAR* Collaboration Meeting), Rome, Italy, 11-14 Nov 1996, Princeton, NJ, 17-20 Mar 1997, Orsay, France, 16-19 Jun 1997 and Pasadena, CA, 22-24 Sep 1997.
- [11] ed. Lide, David R. *CRC Handbook of Chemistry and Physics*. Chemical Rubber Publishing Company, 73<sup>rd</sup> edition, 1992.
- [12] Helen R. Quinn and Michael S. Witherell. The Asymmetry between Matter and Antimatter. *Scientific American*, 279:76–82, October 1998.
- [13] B. N. Ratcliff. *SLAC-PUB-5946 (1992)*. The  $B$  Factory Detector for PEP-II: A Status Report.
- [14] Peter Weiss. Antimatter Loses Again. *Science News*, 166:126, August 2004.

- [15] Eric W. Weisstein. Legendre polynomial. From *MathWorld*—A Wolfram Web Resource. <http://mathworld.wolfram.com/LegendrePolynomial.html>.
- [16] Eric W. Weisstein. Poisson distribution. From *MathWorld*—A Wolfram Web Resource. <http://mathworld.wolfram.com/PoissonDistribution.html>.