UNIVERSITY OF CALIFORNIA
SANTA CRUZ

SEARCH FOR INVISIBLE DECAYS OF THE Υ(1S) MESON AT *BABAR*

A dissertation submitted in partial satisfaction of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

in

PHYSICS

by

Lucas O. Winstrom

June 2009

The Dissertation of Lucas O. Winstrom
is approved:

Professor Bruce Schumm, Chair

Professor Abraham Seiden

Professor Robert Johnson

Dean Lisa C. Sloan
Vice Provost and Dean of Graduate Studies

Table of Contents

List of Figures

List of Tables

Abstract

Search for Invisible Decays of the Υ(1S) Meson at *BABAR*

by

Lucas O. Winstrom

This details the search for invisible decays of the $\Upsilon(3S)$ meson at the *BABAR* experiment. This decay is potentially sensitive to new physics beyond the Standard Model. We use the *BABAR* $\Upsilon(3S)$ data set, analyzing 22.1 fb^{-1} of data, which corresponds to 91.4 million $\Upsilon(3S)$ mesons delivered by PEP-II. The dipion transition $\Upsilon(3S) \to \pi^+\pi^-\Upsilon(1S))$ is studied to identify a collection of $\Upsilon(1S)$ mesons upon which to perform this search. The measurement is $\mathcal{B}(\Upsilon(1S) \to \text{invisible}) = (-1.2 \pm 1.4(\text{stat.}) \pm 1.7(\text{syst.})) \times 10^{-4}$ which is consistent with the Standard Model prediction of $\sim 1 \times 10^{-5}$. This measurement sets an upper limit on the branching fraction of 2.9×10^{-4} at the 90% confidence level.

To the invisible particles . . .

where are you?

Acknowledgments

It is well known that a Ph.D. cannot be completed without the help and support of a small army, and I have been fortunate to have an extremely wonderful one helping me. Thank you to everyone who has been with me along the way.

My mother and father, Emily and Lee, have always been supportive of anything that I want to do, even something as harebrained as getting a degree in experimental particle physics. I know that my success began with their love and the home they provided for me. Thank you for starting me out and sticking with me the whole way. I also want to thank my brother, Zac, for all of his love and friendship over the years.

My advisor, Bruce Schumm, has been an incredible guide through the years leading up to this thesis. He has helped me maintain a larger view of the questions we are trying to answer, preventing me from being discouraged by details that could rapidly sap my resolve to continue asking. During this particular analysis, Stephen Sekula has been invaluable both as a mentor and a friend. I am very happy that I have been able to work closely with both of these scientists; it's been a lot of fun.

I would also like to express my appreciation to everyone involved in the Santa Cruz Institute of Particle Physics (SCIPP). SCIPP is a wonderful place to conduct scientific research. The friendly atmosphere and supportive people make graduate school much more enjoyable. I would especially like to thank Geogia Hamel and Vicki Johnson, who have made the myriad of administrative headaches available at UCSC vanish.

The *BABAR* collaboration has been a fantastic place to learn the details of experi-

mental particle physics research. It has been invaluable to be able to talk with other graduate students and senior physicists about common interests. Thank you to all the people who built and maintained the experiment, providing the data with which this measurement was made.

Friendships have made it possible to maintain my sanity when overwhelmed by frustrations with graduate school, and I am incredibly grateful to the people who have been part of my life. Dan and Shannon Westfahl, you have been constant sources of encouragement and love through high school, college, and into the present day. Matt White and Ryan Campbell, you are two of the most incredible human beings I have ever known. Thank you for helping me get through UW, getting me to do crazy things, and for being such fantastic friends.

Mike Griffo (and Gauge), thank you for freeing my heal and keeping me company through graduate school. I don't know what I would have done without someone who skis, makes such great pizza, and cheers me up during tough times. Chanda Prescod-Weinstein thank you for always being willing to listen and talk. Candace Church, you've become one of my closest friends and I'm sad that you left Santa Cruz for New Mexico. I miss you, and they don't do collider physics at Los Alamos.

Kate Rubin, spending time with you makes me happy, no matter what else is happening in my life. Thank you for giving me so much happiness.

There are a few people who made my time in college a wonderful experience, giving me the courage to continue along the academic track in the hope of finding more such friends. Thank you Pete Allen, Mark Blunk, Sam Coskey, Jane Dolliver, Ernie Esser, Jeff Giansiracusa, Matt Hillman, Bryan Sandoz, and Patrick Wagner. Thank you Joe Fitzgerald, Jeff Jones, and Derek Larson for being such friends.

Finally, I would like to thank the writers of *Futurama* for describing dark matter so well:

> You see, Vergon 6 was once filled with a super dense substance known as Dark Matter, each pound of which weighs over 10,000 pounds.
>
> – Professor Hubert J. Farnsworth, *Love's Labours Lost in Space*

Chapter 1

Motivation

Invisible decays of heavy particles can be used both as tests for Standard Model (SM) predictions and as probes for "new physics" phenomena not explained by the SM. Measurements of invisible Z^0 decays at LEP were able to precisely test the hypothesis of three generations of light neutrinos [1, 2, 3, 4]. A search for invisible decays of the B^0 meson [5], a channel with a SM branching fraction far below the sensitivity of the *BABAR* experiment, probed potential new physics that could enhance the decay rate.

In a similar vein, searches for the invisible decays of the $\Upsilon(1S)$ meson can be used to explore the possibility of physics beyond the standard model that might couple to the $b\bar{b}$ system.

1.1 $\Upsilon \rightarrow$ invisible Decays - Theory

Invisible upsilon decays can in principle be used both to test the Standard Model (SM), and as a search for new physics. The Standard Model process for invisible decays of the $\Upsilon(1S)$ meson proceeds by the transition of the b and $\bar{b}$ quarks to a pair of neutrinos mediated

by a Z^0 boson as show in Figure 1.1.

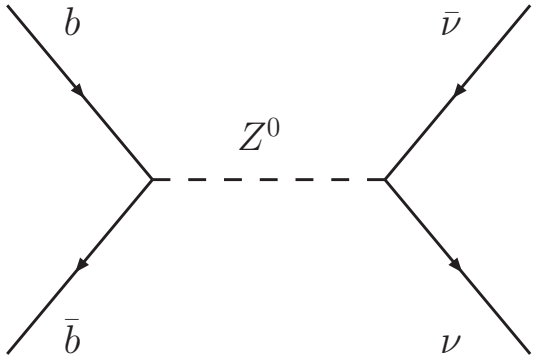


Figure 1.1: The process $\Upsilon \to \text{invisible}$ in the Standard Model.

A leading order calculation of this transition yields

$$\frac{\Gamma(\Upsilon \to \nu\overline{\nu})}{\Gamma(\Upsilon \to e^+e^-)} = \frac{27G^2M_\Upsilon^4}{64\pi^2\alpha^2}\left(-1+\frac{4}{3}\sin^2(\theta_W)\right)^2.$$

where neglected QCD corrections, electroweak radiative corrections, and corrections due to Higgs loops amount to $\sim\ 2-3\%$ [6]. This leads to a prediction for the Standard Model branching fraction of $\sim\ 1\times10^{-5}$. Since the *BABAR* data set can set an upper limit on this branching fraction at the 90% Confidence Level (CL) of $\sim 3\times10^{-4}$ (Chapter 6.5), this allows a search for potential new physics enhancing the invisible decay rate of the $\Upsilon(1S)$. A precision measurement of this decay will have to wait for the next generation of of $b\bar{b}$ factories.

There are a variety of hypothesized effects that can alter the Υ invisible branching fraction. Extensions to the Standard Model due to Super-Symmetric enhancements, which are explored in [6], can change this branching fraction, but not drastically enough for this search to differentiate them from a branching fraction of 0. These types of measurements will have to

wait for a data set capable of probing down to the 1×10^{-5} level or below.

However, light dark matter candidates which are not supersymmetric in origin could couple weakly to standard model particles to enhance this decay [7, 8]. This type of phenomenological coupling simply hypothesizes a new coupling between Υ states and Dark Matter particles below one half the mass of the Υ meson (in the case of this analysis, $\frac{1}{2}M_{\Upsilon(1S)} = 4.7\,\mathrm{GeV}/c^2$). Using relic density and the upper limit on dark matter candidate mass as constraints, this type of enhancement can lead to branching fractions of $\Upsilon \to \mathrm{invisible}$ of up to 1.8×10^{-3} [8]. This lies within reach of the *BABAR* data set and offers a compelling motivation for this search.

Unparticles [9], a hypothesized scale invariant sector of an effective field theory, can enhance this decay [10]. Scale invariant theories can arise as a low energy effective theory of high energy extensions to the Standard Model [9]. The unparticles of this theory do not have a definite mass, but rather a mass spectrum. The enhancement from this theory is sensitive to the dimensionality of the scale invariant sector and the coupling between the SM sector and this scale invariant sector. The current limits on these effects come from measurements of $\mathcal{B}(\Upsilon(1S) \to \mathrm{invisible})$ done elsewhere, and there are no other measurements that might be used to constrain effects from this framework to an unmeasurably small contribution. Although not as compelling as dark matter, unparticles provide additional motivation for this study.

1.2 $\Upsilon \to$ invisible Decays - Experiments

While all of these theories are potential explanations for the invisible decays of the $\Upsilon(1S)$ meson, they all have an identical signature: an enhanced rate of invisible decays. Should

this decay be measured in the future, other channels, such as $\Upsilon(1S) \to \gamma$ invisible, will have to be explored to differentiate between them.

Previous experimental searches for this decay have been done by the CLEO and Belle collaborations. Using 1.2 fb^{-1} of data collected at the $\Upsilon(3S)$ resonance, CLEO measured an upper limits at a 90% confidence level of $\mathcal{B}(\Upsilon(1S) \to \text{invisible}) < 3.9 \times 10^{-3}$ [11]. Using 2.9 fb^{-1} of data collected at the $\Upsilon(3S)$ resonance, Belle measured an upper limits at a 90% confidence level of $\mathcal{B}(\Upsilon(1S) \to \text{invisible}) < 2.5 \times 10^{-3}$ [12]. In this analysis, we use 22.1 fb^{-1} of data collected at the $\Upsilon(3S)$ resonance. While we cannot measure down to the Standard Model prediction for this decay, we do manage to set a limit of $\mathcal{B}(\Upsilon(1S) \to \text{invisible}) < 2.9 \times 10^{-4}$. We use a sophisticated multivariate event selection algorithm to create a sample of events with a minimum of background. We employ a statistical technique known as ${}_s\mathcal{P}lots$ [13] to calculate the expected number of events faking signal. Finally, we perform a Maximum Likelihood fit to extract the upper limit on the branching fraction.

Chapter 2

The *BABAR* Experiment

This thesis discusses a measurement which was performed using data collected by the *BABAR* [14] collaboration using the *BABAR* detector. The *BABAR* detector is a solenoidal, general purpose particle detector designed to observe position-electron collisions produced by the PEP-II (Positron Eelctron Project II) [15] accelerator at the SLAC National Accelerator Laboratory. A thorough description of this detector is published elsewhere [14]. This chapter will summarize the detector, the tracking and particle identification, and the event triggering detailed in that work and elsewhere. All figures showing the detector in this chapter come from [14].

Although *BABAR* was originally designed as a B meson factory observing the $B\bar{B}$ decays of $\Upsilon(4S)$ mesons, this analysis uses a special collection of $\Upsilon(3S)$ mesons produced between December 23, 2007 and March 1, 2008. This data set allows us to search for new physics not accessible in the $\Upsilon(4S)$ data, such as light Higgs and Dark Matter candidates, as well as pieces of the Standard Model particle spectrum which have not previously been

observed, such as the η_b[16]. The search for invisible decays of the $\Upsilon(1S)$ meson is a search for new physics, as motivated in Chapter 1.

2.1 Positron Electron Project II (PEP-II)

PEP-II is a positron electron collider originally designed to operate with a center-of-mass (CM) energy of 10.58 GeV, the mass of the $\Upsilon(4S)$ meson. The PEP-II team and the *BABAR* collaboration worked together to operate the collider at the mass of the $\Upsilon(3S)$ meson, 10.36 GeV, to collect the data that this analysis uses.

Positrons and electrons are accelerated into the two storage rings that make up PEP-II. These electrons have an energy of 9.0 GeV and the positrons have an energy of 3.0 GeV. The beams in these two counterrotating rings are tuned to the correct energies and brought into collision inside the *BABAR* detector. The different energies of the beams make this collision asymmetric, boosting the center of mass frame by $\beta\gamma = 0.58$ relative to the laboratory frame. These two beams are bunched to create dense, well spaced packets of particles before being brought into collision. The RMS spatial extent of these packets in the laboratory frame is $\sigma_{L_x} = 120\mu\mathrm{m}$, $\sigma_{L_y} = 5.6\mu\mathrm{m}$, and $\sigma_{L_z} = 9\mathrm{mm}$, where the coordinate frame of the experiment is defined as follows. The $\hat{z}$ direction is along the direction of electron beam flight, the $\hat{y}$ direction is upwards, and the $\hat{x}$ direction is defined such that $\hat{y} \times \hat{z} = \hat{x}$.

PEP-II has delivered a total luminosity at the $\Upsilon(3S)$ resonance of 30.2 fb^{-1}, corresponding to 122 million $\Upsilon(3S)$ mesons [17]. Through its lifetime, it delivered a total of 557 fb^{-1} of data, with a peak luminosity of $12.069 \times 10^{33}\mathrm{cm}^{-1}\mathrm{s}^{-1}$, more than four times the

original design specifications.

2.2 The *BABAR* Detector

BABAR is an azimuthally symmetric general-purpose particle detector, designed to work with PEP-II and detect the e^+e^- collisions generated by the collider. Diagrams of the detector and its subsystems are seen in Figures 2.1 and 2.2. Because, as discussed earlier, *BABAR* is designed to measure asymmetric collisions, there is a forward-backward asymmetry in the detector systems, giving a better center-of-mass (CM) frame fiducial coverage.

The detector consists of a number of subsystems, which proceed radially outwards from the beam pipe. Nearest the interaction point (IP) is the silicon vertex tracker (SVT) followed by the drift chamber (DCH). These two subsystems participate in particle identification (PID) and are the inputs for the tracking algorithms used at the *BABAR* collaboration. Outside of these lies the detector of internally reflected Cherekov radiation (DIRC), a key piece of the PID system and the electromagnetic calorimeter (EMC). These are surrounded by a superconducting solenoid which creates a 1.5T field. Finally, the instrumented flux return (IFR) surrounds the magnet and functions as a muon and neutral hadron detector.

This analysis particularly depends on the SVT, DCH, and EMC for measurement of the physics signal. The DIRC plays an important role in selection of control samples for the analysis, and the IFR is largely unused.

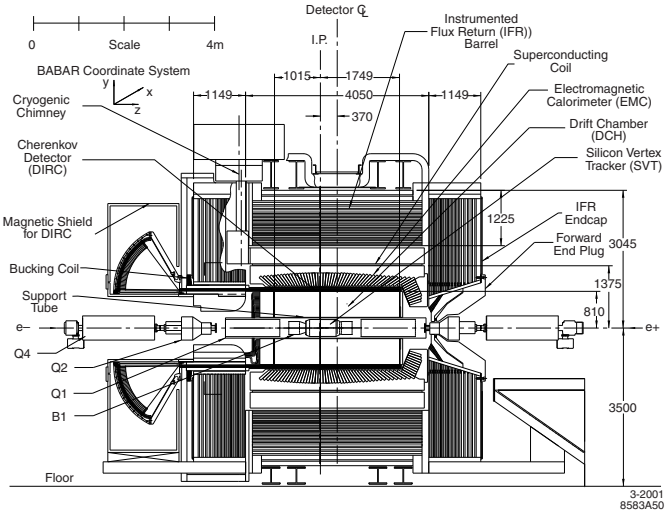


Figure 2.1: The *BABAR* detector longitudinal view.

2.2.1 Silicon Vertex Tracker (SVT)

The SVT is designed to measure angles and position of charged particles just outside the beam pipe. The SVT is composed of five layers of double-sided silicon strip detectors that are assembled from modules with readout at each end, thus reducing the inactive material in the acceptance volume. The inner three layers primarily provide position and angle information for the measurement of the vertex position, but also add momentum information to tracks that extend into them. They are mounted as close to the water-cooled beryllium beam pipe as practical, thus minimizing the impact of multiple scattering in the beam pipe on the extrapolation to the vertex. The outer two layers are at much larger radii, providing the coordinate and angle measurements needed for linking SVT and DCH tracks.

The principal use of the SVT in this analysis is the identification of tracks, and guaranteeing that tracks originate at the interaction point of the e^+e^- collision. Information on the

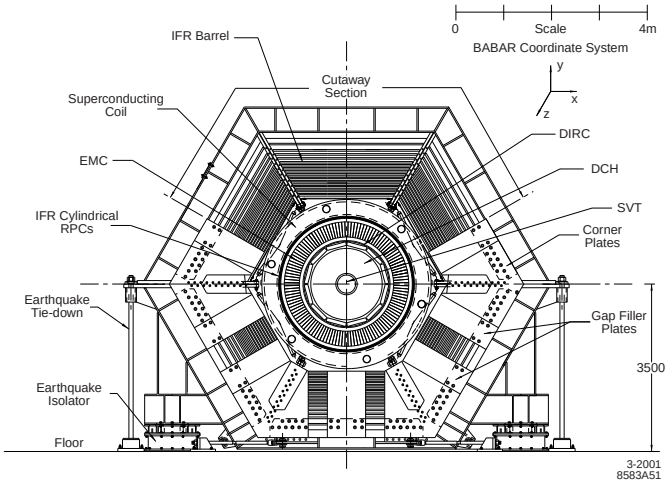


Figure 2.2: The *BABAR* detector transverse view.

origins of tracks is found by fitting track hits in both the SVT and the Drift Chamber (DCH) using a Kalman Filter. The resolution on distance from the interaction point is momentum dependent, and in the energy range of the tracks of interest to this analysis ($\sim 0.1 - 0.8\,\mathrm{GeV}$), the resolution is $\sim 0.1 - 0.3\mathrm{mm}$ for both the resolution in the z axis and the distance travelled in the $x - y$ plane. These are both far less than the $2.5\mathrm{cm}$ cut in z and $1.5\mathrm{cm}$ cut in the $x - y$ plane imposed on tracks used in this analysis. For higher energy tracks, the resolution becomes much better, with a resolution of $40\mu\mathrm{m}$ in the z axis and $25\mu\mathrm{m}$ in the $x - y$ plane. Information on the vertex probability of the two tracks is also used in this analysis, but the requirements on the vertexing are also very loose, much less than the $100\mu\mathrm{m}$ in z and $80\mu\mathrm{m}$ in the $x - y$ plane resolution that the SVT is capable of providing. Diagrams of the SVT can be seen in Figures 2.3 and 2.4.

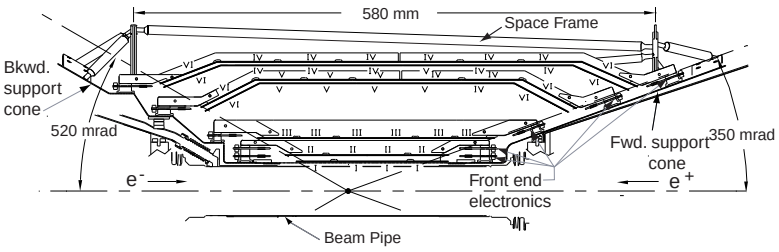


Figure 2.3: The SVT detector longitudinal view.

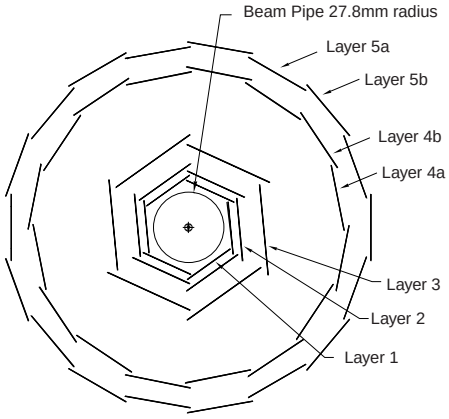


Figure 2.4: The SVT detector transverse view.

2.2.2 Drift Chamber (DCH)

The principal purpose of the DCH is the momentum measurement and tracking of charged particles. It also provides information for the charged particle trigger and a measurement of dE/dx (with a resolution of 6.9%) for particle identification. The DCH has 40 layers of approximated hexagonal cells that extend the length of the detector. These are divided into 10 superlayers. Low-mass wires and the gas mixture of helium-isobutane minimizes multiple scattering inside the DCH. The readout electronics are mounted on the backward endplate of

the chamber, minimizing the amount of material in front of the calorimeter endcap.

This analysis relies heavily on the momentum measurements of this subsystem, and its input to the particle identification (PID) algorithms. This system is the primary PID input for particles up to $0.7\,\mathrm{GeV}$, which includes almost all of the particles subject particle identification in signal sample of this study. When the DCH and SVT information is combined, the momentum resolution on a particle track is $\sigma_{p_T}/p_T = (0.13 \pm 0.01)\% \cdot p_T + (0.45 \pm 0.03)\%$. Figure 2.5 shows a tranverse view of the DCH, and Figure 2.6 shows a schematic of the first 4 superlayers of the DCH. Figure 2.7 shows the separation between K and π particles in dE/dx as a function of momentum. The separation power below $1\,\mathrm{GeV}/c$ can be seen very clearly.

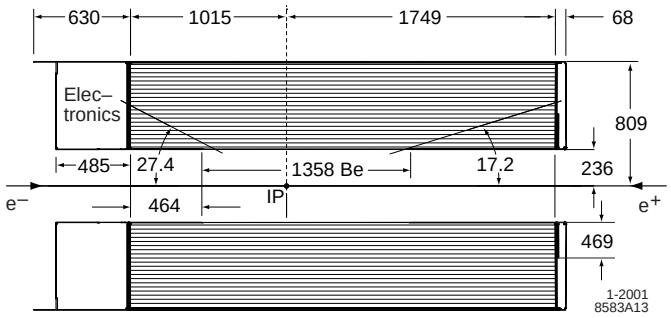


Figure 2.5: The DCH detector longitudinal view.

2.2.3 Detector of Internally Reflected Cherenkov Radiation (DIRC)

The DIRC is a novel device which provides separation information for pions and kaons from about $0.5\,\mathrm{GeV}$. Cherenkov light is produced in $4.9\mathrm{m}$ long bars of synthetic fused silica or rectangular cross section, $1.7\mathrm{cm} \times 3.5\mathrm{cm}$, and transported by total internal reflection, preserving the angle of emission, to an array of photomultiplier tubes. This array forms the

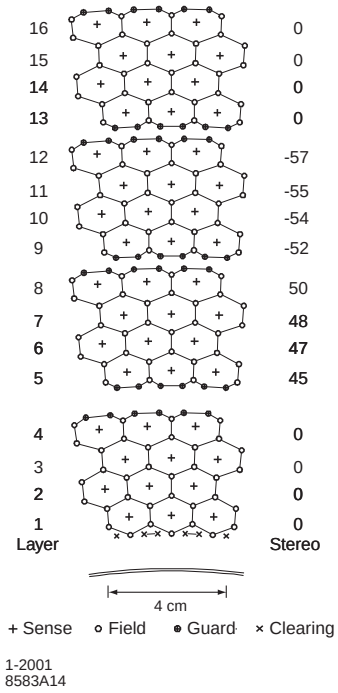


Figure 2.6: The first 4 superlayers of the DCH.

backward wall of a toroidal water tank that is located beyond the backward end of the magnet.

DIRC information plays some role in the PID information that this analysis uses, as the algorithms to discriminate pions and kaons use all information available to them. Our control samples benefit greatly from the ability to distinguish leptons from kaons at high energies. However, since the tracks in signal events that we study are very soft (maximally $0.8\,\mathrm{GeV}$), most of the identification information is delivered by the DCH. A schematic of the DIRC is seen in Figure 2.8.

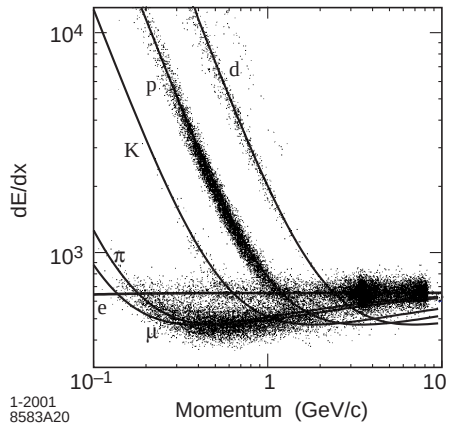


Figure 2.7: Measurement of dE/dx in the DCH as a function of track momenta.

2.2.4 Electromagnetic Calorimeter (EMC)

The EMC is designed to detect electromagnetic shower with excellent energy and angular resolution of the energy range of 20 MeV to 4 GeV. The EMC is a finely segmented array of projective geometry made of thallium doped cesium iodide (CsI(Tl)) crystals. The crystals are arranged in modules that are supported individually from an external support structure. This structure is built in two section, a barrel and a forward endcap. To obtain the best possible resolution, the amount of material in front of and in-between the crystals is held to a minimum. The energy resolution of this component is $\sigma_E/E = (2.32 \pm 0.30)/\sqrt[4]{E(\,\mathrm{GeV})}\% \oplus (1.85 \pm 0.12)\%$.

This analysis uses the EMC to guarantee that potential signal events are free of high energy particles not associated with the physics being studied. The requirements that an event be categorized as a signal event in this analysis requires that no single particle in the EMC be too energetic, and that the total energy deposited in the calorimeter be low. These statements are made precise later in the thesis, where these quantities are defined as inputs to a multivariate

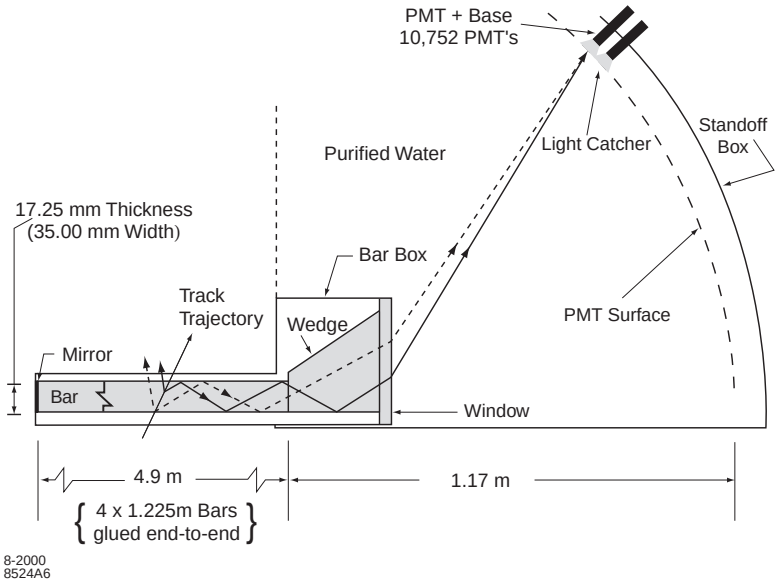


Figure 2.8: The DIRC detector.

selection algorithm. A longitudinal cross section of the EMC can be seen in Figure 2.9.

2.2.5 Instrumented Flux Return (IFR)

The IFR is designed to identify muons and to detect neutral hadrons. For this purpose, the magnet flux return steel in the barrel and the two end doors is segmented into layers. Between the steel absorbers, resistive plate capacitors and limited streamer tubes are inserted, which detect ionizing particles.

The IFR is not used in this analysis, as none of the high energy tracks, which could be either muons or electrons, used in control samples require particle identification. A drawing of the IFR can be seen in Figure 2.10.

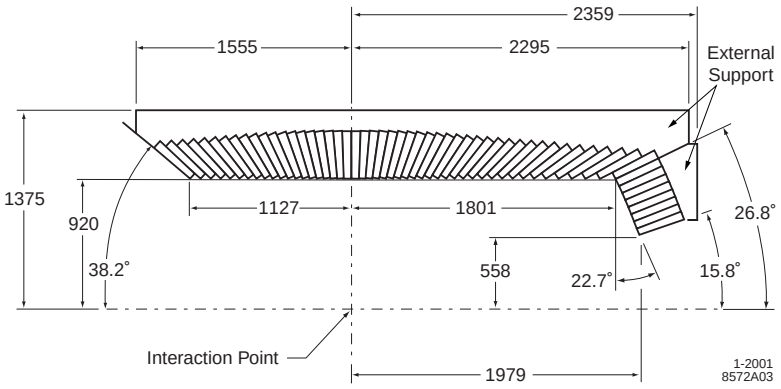


Figure 2.9: Longitudinal view of the top half of the EMC.

2.3 The *BABAR* Triggering System

The trigger system operates as a sequence of two independent stages, the second conditional upon the first. The Level 1 (L1) trigger is responsible for interpreting incoming detector signal, recognizing, and removing beam-induced background to a level acceptable for the Level 3 (L3) software trigger which runs on a farm of commercial processors.

L1 consists of pipelined hardware processors designed to provide an output rate of $\lesssim 2\mathrm{kHz}$. The L1 trigger selection is based on data from the DCH and EMC. Raw data from these components are delivered to the L1 processors and converted into trigger objects, some of which are defined as follows: Hits in the DCH that reach either superlayer 5 (classified as a short track) or superlayer 10 (classified as a long track). These tracks must have p_T at least 120 or 180 MeV/c to reach this far in the detector. Hi $p_T(> 120\,\mathrm{MeV}/c)$ tracks in the DCH. Energy deposits in small segments of the EMC which exceed 100 MeV. Other trigger objects can be defined to match different physics conditions (see e.g. Chapter 8.2.1). The maximum latency for a given collision is $12\mu\mathrm{s}$. L1 has 24 trigger lines which can be changed based on

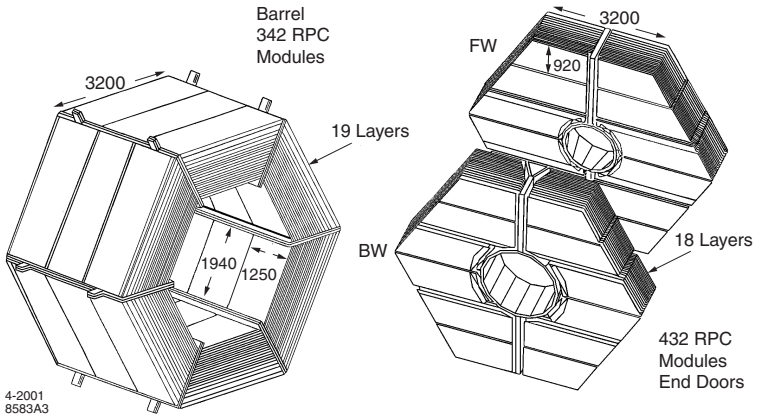


Figure 2.10: Diagram of the IFR, both barrel and end doors.

different physics conditions[18]. These lines can be set to trigger on the presence of different combinations of trigger objects (see e.g. Chapter 8.2.1). Once L1 passes an event to L3, it is analyzed and, if found to be interesting physics, passed on for permanent storage. A small subsample of events are passed without requiring L3 processing as a control sample check on this triggering system. A diagram of the data flow can be seen in Figure 2.11.

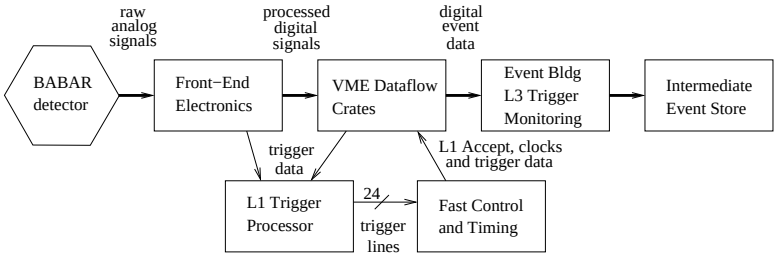


Figure 2.11: Schematic diagram of the data acquisition.

Chapter 3

Analysis Overview

This analysis seeks to measure or put an upper limit on the branching fraction of $\Upsilon(1S)$ mesons to invisible final states. The analysis takes place on data collected by the *BABAR* experiment from electron-positron collisions at the $\Upsilon(3S)$ resonance. In order to select events relevant to this search, the presence of two oppositely charged pions is required. This allows the presence of an invisible decaying $\Upsilon(1S)$ meson to be inferred from the mass the dipion system is recoiling against due to the dipion transition $\Upsilon(3S) \to \pi^+\pi^-\Upsilon(1S)$. There are two distinct backgrounds to the invisible signal:

- A non-peaking combinatoric background due to events containing containing two pions which are not due to the Υ transition, but that are kinematically similar to signal events
- A peaking background where the $\Upsilon(3S)$ undergoes the dipion transition to a non-invisibly decaying $\Upsilon(1S)$ but where the final decay products of the $\Upsilon(1S)$ escape detection

Methods are developed to suppress the combinatoric background and precisely understand the peaking background. Calculations are performed to understand the statistical sensitivity of this measurement. Finally, a maximum likelihood fit is performed to extract the yield of signal events. This yield is combined with the statistical sensitivity and systematic error calculations to quote an upper limit on the branching fraction $\mathcal{B}(\Upsilon(1S) \rightarrow \text{invisible})$. This chapter serves as an introduction to the analysis and a road map of the procedures that are described in greater detail in the following chapters.

3.1 Data and Monte Carlo Simulation

3.1.1 Data

From December 23, 2007 to March 1, 2008, PEP-II delivered beams at the $\Upsilon(3S)$ resonance to the *BABAR* detector with an integrated luminosity of 30 fb^{-1}. This analysis depends on the development of a special trigger configuration that was deployed in early January, and so uses 23.294 fb^{-1} of this data. This corresponds to 96.5×10^6 $\Upsilon(3S)$ mesons. Approximately 5% of the data is used for training, testing, and validating event selection methods which is then discarded, leaving 22.092 fb^{-1} or 91.4×10^6 $\Upsilon(3S)$ mesons for the final measurement.

3.1.2 Monte Carlo Simulation

A number of Monte Carlo samples are generated to model various components of this analysis:

- invisible $\Upsilon(1S)$ decays

- leptonic $\Upsilon(1S)$ decays
- generic $\Upsilon(1S)$ decays
- continuum processes
  - $\tau^+\tau^-$
  - light-quark production
  - radiative dilepton events

Monte Carlo simulations are run in large batches by the *BABAR* computing group. These simulations are discussed in more detail in Chapter 4.2.

3.2 Event Selection

The presence of the two pions from the transition $\Upsilon(3S) \to \pi^+\pi^-\Upsilon(1S)$ is required for identification of a $\Upsilon(1S)$ meson in an event. The formula for the mass that the two pions are recoiling against

$$m^2_{recoil} = s + m^2_{\pi\pi} - 2\sqrt{s}E_{\pi\pi} \tag{3.1}$$

determines whether or not a pair of pions might have originated from this transition. While events that contain only these two oppositely charged pions are used for the actual measurement, important control samples are also available within the data. The leptonic transition of the $\Upsilon(1S)$ can be detected and is used for two control samples: events with one lepton detected

and events with both leptons detected. We also use a sample with a single high energy neutral particle as a "nearly" invisible sample.

3.2.1 $\Upsilon(1S) \rightarrow$ invisible

This selection chooses events that have the two oppositely charged tracks (presumed to be from the dipion transition) with a recoil mass $\in [9.35, 9.57]\,\mathrm{GeV}/c^2$. There may be additional activity in the detector, as it is expected other activity from beams or noise will register during these events. This selection will contain a large amount of background from events that contain two pions but do not contain a $\Upsilon(1S)$. The region in m_{recoil} that is far away from the $\Upsilon(1S)$ mass can be used as a sideband region to study this background. This sideband is used for the measurement of, and to understand, the combinatoric background in the data.

3.2.2 $\Upsilon(1S) \rightarrow \ell^+\ell^-$

In addition to the dipion system in the $\Upsilon(1S) \rightarrow \mathrm{invisible}$ sample, this sample requires the presence of two high energy tracks from the leptonic decay of the $\Upsilon(1S)$. This sample is extremely pure, allowing the derivation of the m_{recoil} distribution shape from data. From this distribution, the sideband regions can be defined as being more than 5σ from the peak, and the signal region can be defined as being within 5σ of the peak. This sample is also used to calculate a scaling factor for the Monte Carlo simulation of leptonic $\Upsilon(1S)$ decays so the number of leptonic decays of the $\Upsilon(1S)$ escaping detection is accurately predicted.

3.2.3 $\Upsilon(1S) \to \ell$

We also collect events where there is one high energy track in addition to the dipion system. While this sample is not as pure as the two lepton sample, it occupies a different fiducial range of the detector. Its location in the back end of the detector is used to correct differences in event acceptance between the simulation and data.

3.2.4 "Nearly" Invisible Sample

This control sample is collected with the same dipion requirements as the $\Upsilon(1S) \to \mathrm{invisible}$, but requires the presence of an extra neutral cluster in the calorimeter to separate it from the invisible sample. We use this collection to compare the Monte Carlo predictions of generic $\Upsilon(1S)$ decays to those in actual data to constrain the number of peaking background events that might come from non-leptonic decays.

3.3 Non-Peaking Background Reduction

Since this analysis simply requires the presence of a dipion system with an invariant mass in a window, a large quantity of combinatoric background initially overwhelms the signal. In order to suppress this background, two methods are explored. A cut and count method is used as a baseline, and as a sanity check, and a sophisticated multivariate selection algorithm is trained and used for final event selection. For each method, a number of observables are explored for their power to discriminate between Monte Carlo signal events and sideband data from the $\Upsilon(1S) \to \mathrm{invisible}$ sample. Eventually the total signal yield will be determined using

a fit to m_{recoil}, so these variables must also be uncorrelated with m_{recoil}. A variety of variables are identified that can be used to discriminate between signal and background irrespective of recoil mass.

3.3.1 Cut and Count

Initially, a cut and count routine is performed to establish a baseline significance. This uses the PRIM algorithm [19] to find a multidimensional box containing the highest signal significance $S \equiv s/\sqrt{s+b}$. This gives a baseline significance that can be compared with a more sophisticated selection algorithm. While the cut and count routine is not used for the rest of the analysis, it is important to make sure that the multivariate algorithm works at least as well as the cut and count routine (preferably better), and that it does not give completely unbelievable gains over this routine, as that may point to errors elsewhere in the analysis.

3.3.2 Random Forest Decision Tree

The multivariate "Random Forest" algorithm (MVA) is based on a large number of "decision trees" trained with a set of random inputs. Decision tree outputs are weighted and summed to optimize the Gini index [20] of the Random Forest. Once the algorithm has been developed, a cut point must be determined. This is done by defining the procedure through which the signal yield will be measured and running several Monte Carlo experiments to understand the statistical sensitivity of the measurement. The results of these "toy" Monte Carlo experiments are used to minimize the 90% Confidence Limit on this measurement so that the best upper bound can be found.

3.4 Maximum Likelihood Fit

An extended maximum likelihood fit is performed on the data in the m_{recoil} to determine the yield of peaking and background components. We use the signal shape derived from the $\Upsilon(1S) \to \ell^+\ell^-$ control sample as a shape for both the peaking background and the signal, and we use a polynomial of order one as a shape for the non-peaking background.

3.5 Statistical Sensitivity Estimations

We create a Monte Carlo procedure that will generate distributions in m_{recoil} according to the signal shape for signal and background. The sideband data is used to estimate the shape of the non-peaking background, and the $\Upsilon(1S) \to \ell^+\ell^-$ control sample is used again to estimate the distribution of the peaking background and signal. These distributions are used to generate sample data sets which are added together to create a sample distribution of signal and background which can then be put through the fitting procedure described in section 3.4. Since in these toy experiments we know exactly the quantity of signal and background events, the fitted yield can be compared to the actual yield to determine the signal sensitivity of the fitting procedure.

By running these toy Monte Carlo experiments many times, we can create a distribution of yields based on various assumptions about values of signal yield. These distributions are used to determine confidence intervals around various branching fraction measurements. This includes both 68% and 90% confidence intervals. We use these confidence intervals to optimize the statistical signal sensitivity of the measurement as a function of Random Forest output.

3.6 Peaking Background Determination

When the $\Upsilon(1S)$ in the $\Upsilon(3S) \rightarrow \pi^+\pi^-\Upsilon(1S)$ transition decays into potentially detectable states, these products can sometimes escape detection. These decays result in a decay that is indistinguishable from a signal event, as only the dipion system can be detected. In order to estimate the number of decays of this sort, we use our Monte Carlo predictions and control samples to predict the number of undetectable events that occur. The most common type of undetectable decay is a leptonic transition of the $\Upsilon(1S)$ where the two final state leptons are produced outside the fiducial range of the detector. This accounts for approximately 99.8% of the peaking background after the MVA cut. We use the control samples to correct the Monte Carlo simulation of the $\Upsilon(1S)$ decay channels, and then use this corrected simulation to find the predicted number of peaking background events

Using the $\Upsilon(1S) \rightarrow \ell^+\ell^-$ control sample, we correct the scaling on the Monte Carlo simulation of the leptonic decays of the $\Upsilon(1S)$. We use the $\Upsilon(1S) \rightarrow \ell$ control sample to measure the backwards detector acceptance and correct the Monte Carlo simulation of the leptonic decays of the $\Upsilon(1S)$. We use the "near" invisible control sample to compare the prediction of non-leptonic decays of the $\Upsilon(1S)$ to the generic Monte Carlo prediction.

3.7 Systematic Errors and Results

Throughout the analysis, we operate on a set of data that is called "blinded". This simply means that we do not allow ourselves access to the invisible data in the signal region (defined in Chapter 3.2.2). The analysis is designed in this blind state to remove any potential

bias towards finding a signal. Once the MVA output cut has been fixed and the number of expected peaking events determined, the extended maximum likelihood fit is performed on the unblinded data. The systematic errors involved in this analysis are listed, added together, and combined with the statistical upper limit derived from this fitting (Chapter 3.4) using a Bayesian approach to calculate the upper limit on $\mathcal{B}(\Upsilon(1S) \rightarrow \mathrm{invisible})$.

Chapter 4

Data, Monte Carlo Simulation, and Preliminary Event Selection

This section describes the data used in this analysis, along with Monte Carlo (MC) samples which are intended to model the data. We explain the conditions of the data which we use in the analysis, as well as the luminosity of that data. We also document the assumed cross-sections for the various processes modeled in MC.

We describe the selection of several subsets of events from the entire data set. This type of selection is intended to be loose, efficiently selecting events of interest while rejecting the bulk of other kinds of events. These subsets are subsequently used to develop more sophisticated event selection techniques.

4.1 Data

The data used in this analysis were taken during a period between December 23, 2007 and March 1, 2008. During this period, data was taken both at and below the $\Upsilon(3S)$ resonance, but only data taken at the $\Upsilon(3S)$ resonance is used in this analysis. A special trigger configuration were developed to optimize the selection of the signal events, but this trigger was not deployed until early January. The new trigger provides a factor of seven gain in signal efficiency; we therefore only use data taken after this trigger was deployed.

The data are split into three subsamples. We use these subsamples to study selection criteria backgrounds that are not be possible to simulate, and then develop techniques to reject these backgrounds. Where the samples are used in such a way as to be unblinded (see Chapter 3.7) or otherwise potentially biased, they are removed from the rest of the analysis. These samples are nicknamed the "Low" and "High" , and "Medium" datasets. The "Low" and "High" sets are comprised of runs early in the $\Upsilon(3S)$ Run (Low) and later in the Run (High), representing a coarse sampling of conditions during data taking. The "Medium" dataset represents all of the runs which are not in either the "Low" or "High" samples. Taken together, these three subsamples consitute the available $\Upsilon(3S)$ data with the new trigger.

We use the "Low" data to train, test, and validate our event selection method (detailed in Chapter 5). Since this data represents a small fraction (5%) of the total dataset, we will exclude this data from the final signal selection. The final dataset will therefore contain only the "High" and "Medium" samples, totaling 22.092 fb^{-1}.

Dataset Name	Integrated Luminosity (fb^{-1})	$\Upsilon(3S)$ Count ($\times 10^6$)
Low	1.201	5.12
High	1.282	5.49
Medium	20.810	85.9
Total	23.294	96.5

Table 4.1: The dataset names and luminosities

4.2 Monte Carlo (MC) Samples

The *BABAR* computing group creates large batches of Monte Carlo simulation labelled by a Simulation Production (SP) number. This analysis takes place using the SP-10 series simulation. A number of MC samples were produced to model specific components of the data, including invisible $\Upsilon(1S)$ decay, leptonic $\Upsilon(1S)$ decay, generic $\Upsilon(3S)$ decay, and continuum processes (including $\tau^+ \tau^-$, light-quark production, and radiative dilepton events). Each of these is labelled with its own specific SP-mode number. These samples are detailed in Table 4.2. The simulation of the dipion transition $\Upsilon(3S) \rightarrow \pi^+\pi^-\Upsilon(1S)$ uses the matrix element analysis from [21] input to EvtGen [22]. The generic decays of the $\Upsilon(3S)$ and $\Upsilon(1S)$ and the continuum processes are generated by either *BABAR* specific event generators or PYTHIA/JETSET [23, 24]. The bhabha events are generated by BHWIDE [25] and the dimuon simulator is BKQED [26, 27, 28]

Once the events have been generated, the interaction with the detector are simulated by a detailed model of the *BABAR* detector based on GEANT 4 [29]

There are two signal MC samples. The "v01" sample resulted from the first full MC production for Run7 in SP10. However, due to a mistake MC were only generated in this

production to simulate conditions from a subset of the total $\Upsilon(3S)$ run, although the intention was to sample correctly over all conditions. The benefit of this sample is that we can use it to optimize cuts or train a multivariate discriminant and "throw it away" afterward, leaving us to use the full and correct simulation to compute efficiencies, etc. We will compare these two MC samples to show that there are no significant differences when using them in this way.

The Particle Data Group [30] branching fraction for the charged dipion transition is used, $\mathcal{B}(\Upsilon(3S) \to \pi^+\pi^-\Upsilon(1S)) = 0.0448$. We normalize generic $\Upsilon(3S)$ MC to the count of $\Upsilon(3S)$ mesons in each given conditions period of data-taking. Lepton universality is assumed for the leptonic decays of the $\Upsilon(1S)$ and the branching fraction for each leptonic final state is assumed to be 0.025 (the average value of leptonic branching fractions found in the PDG [30]). The cross-sections for light-quark continuum and $e^+e^- \to \tau^+\tau^-$ production are scaled from their values at the $\Upsilon(4S)$ assuming $1/s$ scaling: $\sigma_{uds} = 2.18\,\mathrm{nb}$, $\sigma_{c\bar{c}} = 1.36\,\mathrm{nb}$, and $\sigma_{\tau^+\tau^-} = 0.94\,\mathrm{nb}$. The cross-sections for radiative dilepton processes are taken from the generators for these processes, and are $\sigma_{\gamma e^+e^-} = 25.79\,\mathrm{nb}$ and $\sigma_{\gamma\mu^+\mu^-} = 1.1985\,\mathrm{nb}$.

4.3 Event Selection

4.3.1 Event Objects and Their Properties

In order to select particular events, it is convenient to define some objects that exist in the events. We need a way to define tracks in the detector, and a way to identify them. For our selections, the following object definitions are used:

Decay Mode	SP Mode Number	Generated Events
$\Upsilon(3S) \to \pi^+\pi^-\Upsilon(1S)(\to \text{invisible})$	8618	127000
$\Upsilon(3S) \to \pi^+\pi^-\Upsilon(1S)(\to \text{invisible})$ (SP10 "v01" Production)	8618	155000
$\Upsilon(3S) \to \pi^+\pi^-\Upsilon(1S)(\to \mu^+\mu^-)$	8780	1163000
$\Upsilon(3S) \to \pi^+\pi^-\Upsilon(1S)(\to e^+e^-)$	8781	1163000
$\Upsilon(3S) \to \pi^+\pi^-\Upsilon(1S)(\to \tau^+\tau^-)$	8782	993000
$\Upsilon(3S)$ Generic	8739	161704000
$u\overline{u}, d\overline{d}, s\overline{s}$	998	45360000
$c\overline{c}$	1005	53456000
$\tau^+\tau^-$	3429	34712000
Radiative Bhabha	2400	113368000
Radiative Dimuon	2981	33608000

Table 4.2: MC samples details, including SP mode number and number of generated events

- ChargedTracks are any track in the detector. These tracks are fit using a Kalman Filter (detailed in [14, 31, 32]) as helices in the detector.

- GoodTracksVeryLoose (GTVL) are Charged tracks that come from the interaction point (IP) of the experiment. In order to assure this, each must:

  – Have a maximum reconstructed momentum of 10 GeV

  – Have a distance of closest approach in the $x - y$ plane of no more than $1.5\mathrm{cm}$

  – Have a distance of closest approach on the z axis of no more than $2.5\mathrm{cm}$

For our GoodTracksVeryLoose (GTVL) we also use particle identification. Particle identification algorithms are developed by the PID group at the *BABAR* collaboration. The particle identification selectors used in this analysis employ a boosted decision tree algorithm [33] to determine the identity of a certain track. Boosted decision trees are used for particle identi-

fication in many particle physics experiments (e.g. [34]), and have been found to be extremely effective at *BABAR*. The output of a boosted decision tree algorithm is a real number between 0 and 1. In order to transform this real number into a binary selection (i.e. this track is or is not a muon), cut values on the decision tree output are set forth by the PID group. Each decision tree selector is divided into a hierarchy of cut values on the output which correspond to the false positive rate of that selector. There are typically 4 such levels named "VeryLoose", "Loose", "Tight", and "VeryTight", though there may be more or less depending on the particulars of the PID selector. The "Tight" values will have few false positives, and so collect a cleaner sample, but will also reject more true particles. In this analysis, our particle identification requirements are found to be fulfilled using loose particle selectors, and requiring tracks believed to be from pions to fail the selector. The names of the algorithms are: eKMVeryLoose, an electron selector; muBDTLooseLoP, a low momentum muon selectron; muBDTVeryLoose, a muon selector; and KKMVeryLoose, a kaon selector.

4.3.2 Data Samples

There are several different data samples in this analysis which are used in separate ways. We define how we select the events that are contained in each of these. We make use of one signal sample and three control samples.

4.3.2.1 Signal Sample Selection

The signal events in this analysis are expected to contain only the two pions from the dipion transition of the $\Upsilon(3S)$ to the $\Upsilon(1S)$ with no other detectable particles from the

primary physics processes involved. As with any other event in *BABAR*, we also expect there to be activity in detector systems due to noise or PEP-related activity

The signal selection is defined as follows. Events must contain exactly two GoodTracksVeryLoose (GTVL). No further restrictions are placed on, for instance, the number of ChargedTracks, which may exceed two. The tracks much have opposite charge, and have momenta in the center-of-mass (CM) frame satisfying $p^* < 0.8\,\mathrm{GeV}/c$. The two pions must have an invariant mass satisfying $m_{\pi\pi} \in [0.25, 0.95]\,\mathrm{GeV}/c^2$. The physical boundaries of the dipion mass are $2m_\pi = 0.28\,\mathrm{GeV}$ (minimum) and $M_{\Upsilon(3S)} - M_{\Upsilon(1S)} = 0.89\,\mathrm{GeV}$ (maximum). We widen the boundary around these physical limits to admit reconstruction resolution effects on the pions. Finally, we define the mass recoiling against the dipion system,

$$m^2_{recoil} = s + m^2_{\pi\pi} - 2\sqrt{s}E_{\pi\pi}, \tag{4.1}$$

where $\sqrt{s}$ is the collider CM energy (assumed to be $\sqrt{s} = M_{\Upsilon(3S)} = 10.36\,\mathrm{GeV}$) and $E_{\pi\pi}$ is the energy of the dipion system. This should be centered around the $\Upsilon(1S)$ mass ($M_\Upsilon(1S) = 9.46\,\mathrm{GeV}$) We require that $m_{recoil} \in [9.35, 9.57]\,\mathrm{GeV}/c^2$. This skim is found to be 64.4% efficient on signal MC events, where the bulk of the efficiency loss is simply due to the requirement that both pions be reconstructed.

4.3.2.2 Two Lepton Control Sample and One Lepton Control Sample Selection

The control sample of fully reconstructed visible leptonic $\Upsilon(1S)$ decays is constructed similarly to the signal sample. The only differences are that we require three or four GoodTracksVeryLoose (GTVL) (to select events where one or both of the final-state leptons are

reconstructed), and that these additional GTVL have CM momenta satisfying $p^* > 2.0\,\mathrm{GeV}/c$. We attempt to recover radiative energy loss by the lepton due to Bremsstrahlung radiation using nearby photons. No particle identification is required of these tracks, so while this is most effective for electrons, the procedure is carried out for both high energy tracks.

In the case where both leptons are reconstructed (i.e. there are exactly 4 GTVL), the $\Upsilon(1S)$ candidate is required to have a mass satisyfing $M_{\Upsilon(1S)} \in [9.0, 9.8]\,\mathrm{GeV}/c^2$. This skim is 60% efficiency on the Monte Carlo Simulation of the muon final state and 42% efficient on the Monte Carlo Simulation of the electron final state. A brief study of the electron final state using the MC truth reveals that the efficiency of the skim selection is the same between electron and muon until the $\Upsilon(1S)$ mass cut. Even after Bremsstrahlung recovery, radiation effects on the final-state electrons have a significant effect on the reconstructed mass.

Our sample of one-lepton final-state events is also constructed similarly to the two-lepton case, except we do not apply a $\Upsilon(1S)$ mass cut and we explicitly require three GTVL. These events will be discussed later, in section 8, when we describe our systematic errors.

The above selections reject all but 3.8% of the entire data triggered set.

4.3.2.3 "Near" Invisible Control Sample Selection

We create a control sample to study the data to Monte Carlo simulation agreement in a "near" invisible region. This selection is similar to the signal selection, but requires an extra high energy neutral object in the detector to isolate it from the signal sample. This selection is detailed in section 7.3.

4.3.3 $\Upsilon(1S) \to \ell^+\ell^-$ Sample Selection Refinements

The selection of the two lepton control sample is refined in this section in order to determine the signal and sideband regions for the analysis. Once the preliminary selections on two lepton events have beens made, we begin a more in depth evaluation. We wish to construct a sample of events which contain a dipion transition from $\Upsilon(3S)$ to $\Upsilon(1S)$ and have very little contamination from other sources. This will enable us to measure the spectrum of recoil mass from the data rather than the Monte Carlo simulation, and it will serve as a check on the $\Upsilon(1S) \to \ell^+\ell^-$ simulation.

Figure 4.1 shows the recoil mass of the events passing the preliminary two lepton sample selection. We see that the MC are generally in excess of the data everywhere, a significant shift of the MC relative to the data, and that there is a significant contribution (about 12%) from radiative dilepton events ($e^+e^- \to \ell^+\ell^-\gamma$), dominated by the radiative Bhabha. This latter effect is more obvious in Fig. 4.2.

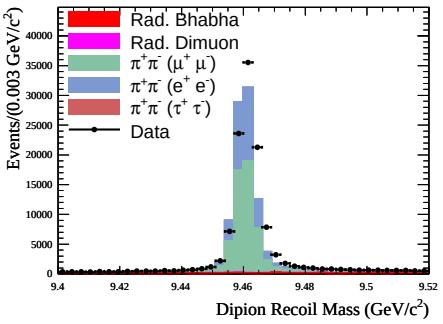


Figure 4.1: The recoil mass of the preliminary two lepton sample. The red and magenta sections are contamination from radiative dilepton events

We study these events and develop selection techniques to remove these backgrounds.

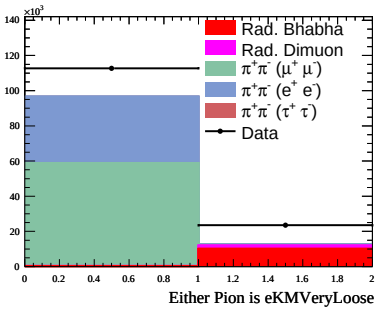
Since the radiative dilepton events should only have two hard tracks in them, the dipion system must be faked. The most likely candidate is photon pair-production in material, creating a pair of soft tracks. Since these are not true $\Upsilon(3S)$ events, the mass difference between the recontructed $\Upsilon(3S)$ and $\Upsilon(1S)$ must also be combinatoric. We also expected a small opening angle between the "pions", consistent with photon pair-production.

We plot these two variables ($\Delta M_{\Upsilon(3S),\Upsilon(1S)}$, and $\Delta\phi_{\pi\pi}$) in Fig. 4.2 and confirm that the majority of the radiative dilepton backgrounds have one or both "pions" identified as electrons (using eKMVeryLoose), a uniformly distributed value of the mass difference, and a narrow opening angle. We thus reject events where either "pion" is identified as an electron, and keep events where $\Delta M_{\Upsilon(3S),\Upsilon(1S)} \in [0.890, 0.920]\,\mathrm{GeV}/c^2$. We do not cut on the opening angle to avoid altering the kinematics of the dipion system in a significant way. This leaves us with a sample of events that is 99.8% pure $\Upsilon(3S) \to \pi^+\pi^-(\Upsilon(1S) \to \ell^+\ell^-)$ in the simulation. A comparison of two lepton sample m_{recoil} spectrum in data and Monte Carlo after this selection can be seen in Figure 4.3

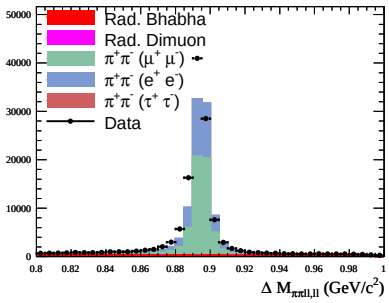
After these selection criteria area applied, we can fit the recoil mass distribution to obtain the yield of visible events, and to obtain a preliminary model of the recoil mass shape for signal events. This will allow us to define a "blinding region" in the recoil mass, $\pm 5\sigma$ around the peak position. We model the recoil mass distribution as the sum of a Cruijff function:

$$\mathcal{C}(m_{recoil}) = \frac{1}{N}\begin{cases} \exp[-(m_{recoil}-\mu)^2/(2\sigma_L^2 + \alpha_L(m_{recoil}-\mu)^2)], & m_{recoil} < \mu \\ \exp[-(m_{recoil}-\mu)^2/(2\sigma_R^2 + \alpha_R(m_{recoil}-\mu)^2)], & m_{recoil} > \mu \end{cases} \tag{4.2}$$

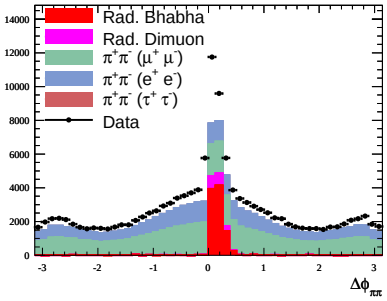
and a linear background. A fit of this model to the data yields the result shown in Figure 4.4.



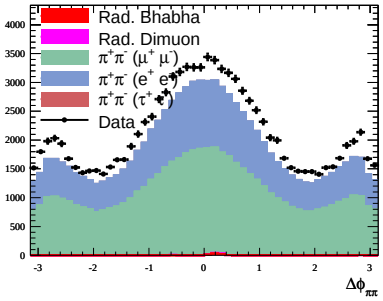
(a) A binary output, requiring either pion pass eKMVeryLoose. A majority of the radiative backgrounds have at least one "pion" identified as an electron, while the pions from $\Upsilon(3S) \rightarrow \pi^+\pi^-(\Upsilon(1S) \rightarrow \ell^+\ell^-)$ are almost completely identified as "not electrons"



(b) The mass difference between the reconstructed $\Upsilon(3S)$ and $\Upsilon(1S)$. The mass difference in the radiative dilepton samples is flat, while the distribution from $\Upsilon(1S) \rightarrow \ell^+\ell^-$ peaks at the true mass difference



(c) The opening angle between the pions in the plane transverse to the z-axis before performing any cuts to remove radiative dilepton events



(d) The opening angle between the pions in the plane transverse to the z-axis, after cuts on electron ID and the mass difference to remove radiative dilepton events

Figure 4.2: Variables under study to reject the radiative dilepton background

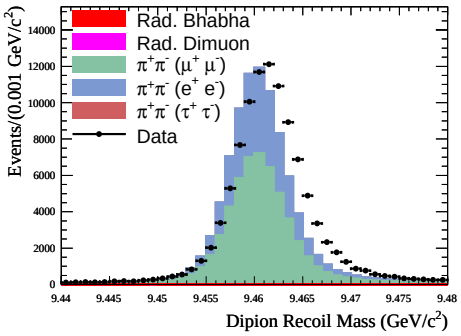


Figure 4.3: Comparing the recoil mass in data and Monte Carlo after cuts to reject radiative dilepton background.

The parameters of the fit are given in Table 4.3. Since this selection is a pure sample of dipion transition between the $\Upsilon(3S)$ and the $\Upsilon(1S)$ it is an ideal place to determine the shape of the recoil mass distribution of the dipion system. The differences between this sample and the invisible sample come from differences in triggering and are treated as systematic errors in Chapter 8.

We study the MC modeling of these events by fitting the sum of the muon and electron final-state Monte Carlo simulations with the same model. We use an unweighted number of events from these MC samples equivalent to an integrated luminosity of $30\,\mathrm{fb}^{-1}$. We allow for the linear component. We find that without this linear component, the quality of the fit to the MC is very poor ($\chi^2/\mathrm{ndof} \sim 10$ with 120 bins). This suggests that the linear component fit in the data is due to effects in the reconstructed signal events and not just due to an unmodeled flat background component. Including the linear component improves the fit quality, though it's still not perfect. The results of the fit are shown in Fig. 4.5 and in the right-most column of Table 4.3. We take the signal PDF model from the data, and use it in the rest of the analysis.

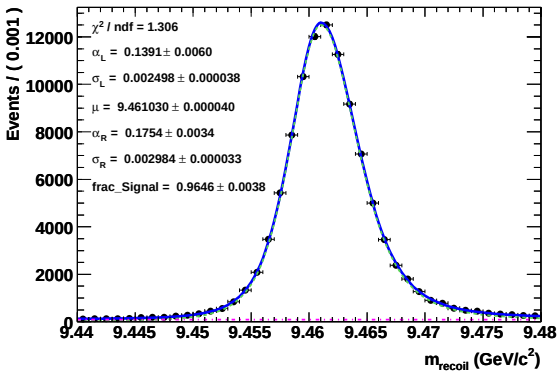


Figure 4.4: Fitting the recoil mass after rejecting radiative dilepton background.

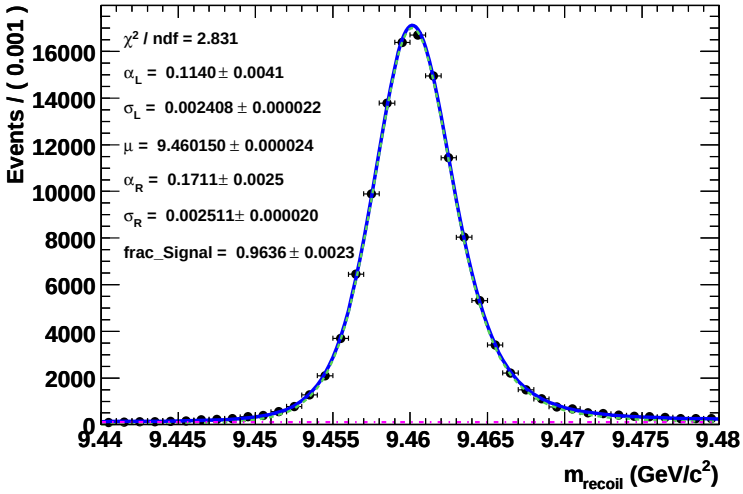


Figure 4.5: A fit for the PDF shape of the signal using visible final states signal MC simulations. The MC simulation is sampled at a rate consistent with the expected data luminosity, and is not weighted.

Cruijff Parameter	Data Value	MC Value
Peak Position (μ)	$(9.46103 \pm 0.00004)\,\mathrm{GeV}/c^2$	$(9.46015 \pm 0.00002)\,\mathrm{GeV}/c^2$
Left Width (σ_L)	$(2.50 \pm 0.04) \times 10^{-3}\,\mathrm{GeV}/c^2$	$(2.41 \pm 0.02) \times 10^{-3}$
Right Width (σ_R)	$(2.98 \pm 0.03) \times 10^{-3}\,\mathrm{GeV}/c^2$	$(2.51 \pm 0.02) \times 10^{-3}$
Left Tail (α_L)	0.139 ± 0.006	0.114 ± 0.004
Right Tail (α_R)	0.175 ± 0.003	0.171 ± 0.002
Cruijff (Peak) Fraction	$(96.5 \pm 0.4)\%$	$(96.4 \pm 0.2)\%$

Table 4.3: Fit results from the visible events in both data and MC

A few of the parameters are of immediate interest. First, the peak of the recoil mass distribution is found to be at $(9.46103 \pm 0.00004)\,\mathrm{GeV}/c^2$, while the PDG value of the $\Upsilon(1S)$ mass is $(9.46030 \pm 0.00026)\,\mathrm{GeV}/c^2$, which are not in great agreement with one another (the difference is about 2.8σ from zero). However, this analysis is not intended to perform a mass measurement. The width of the peak is asymmetric, which one would expect from the reconstruction uncertainty of the pions, and averages on the left and right to about $2.7\,\mathrm{MeV}/c^2$. An underlying flat part of the signal is just 3% of the total distribution, and is in good agreement with the number from the MC prediction.

We define the $\pm 5\sigma$ blinding region around the signal peak as $m_{recoil} \in$ [9.4485,9.4759] GeV/c^2. Data in our signal sample which has m_{recoil} in this window will be excluded from consideration when we develop algorithms to reject background. Data outside of this window will be referred to as "sideband" data.

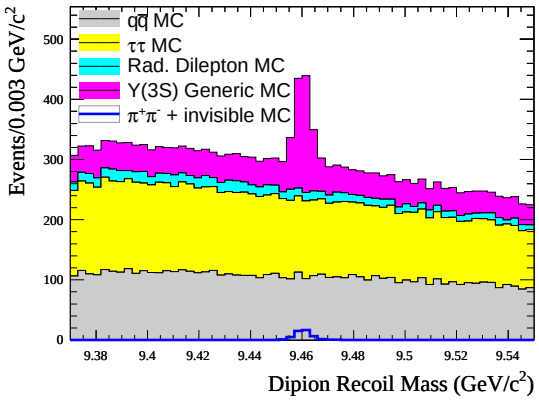
4.3.4 $\Upsilon(1S) \rightarrow$ invisible Sub-sample Selection Criteria

Now that we have defined the shape of the recoil mass in the events of interest, we examine the complete spectrum of recoil mass in our signal sample. While the peaking component is present, it is completely dwarfed by combinatoric backgrounds from other processes that generate two pions.

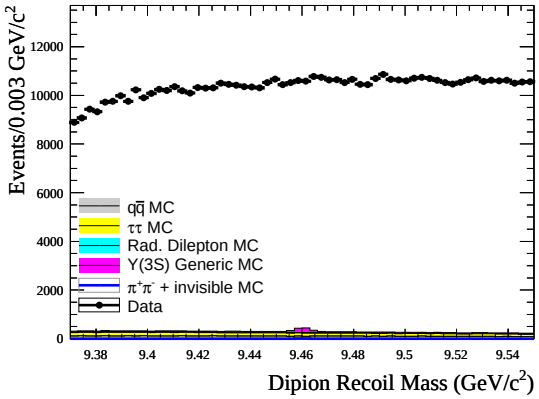
In Fig. 4.6, the left plot in the figure shows the stack of all MC backgrounds; radiative dilepton events contribute very little to this plot, and are hard to see between the $\tau^+\tau^-$ background and the generic $\Upsilon(3S)$ background. The right plot adds the data (the "Low" sample), and we see that the majority of the background events are actually not modeled at all by the MC.

Figure 4.6 illustrates the two major components of the background. The most significant is a non-peaking background comprised of light-quark continuum, $\tau^+\ \tau^-$, and a much larger unmodeled component. This unmodeled component comes primarily from two photon fusion processes that result in $\pi^+\pi^-$ pairs. The second, much smaller (but potentially irreducible) background is a peaking background from generic $\Upsilon(3S)$ decays. In Fig. 4.7, we plot the generic $\Upsilon(3S)$ broken down into events without a dipion transition to the $\Upsilon(1S)$, events with a dipion transition but with excluding leptonic decays of the $\Upsilon(1S)$, and finally the subsample containing only leptonic decays of the $\Upsilon(1S)$. At this early stage of the selection, the contribution of peaking backgrounds which are not from leptonic final states is about one-quarter of the total peaking background.

We will develop a multivariate selection algorithm to suppress the flat, combinatoric



(a) The MC prediction of the background (luminosity normalized to the "Low" sample). The peak in the generic $\Upsilon(3S)$ Monte Carlo illustrates the peaking background that comes from dipion transition between the $\Upsilon(3S)$ and the $\Upsilon(1S)$ with non-invisible final states which are simply not detected.



(b) Including data from the "Low" sample - notice that the MC prediction is still present but has been suppressed by the preponderance of unmodeled backgrounds.

Figure 4.6: The recoil mass in the invisible sample after all selection cuts

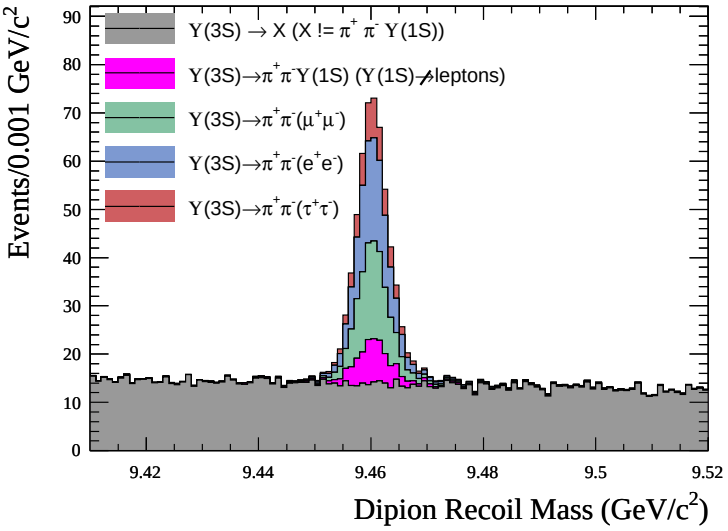


Figure 4.7: The recoil mass, in generic $\Upsilon(3S)$ MC. In order of stacking from bottom to top, the components are: events where the $\Upsilon(3S)$ decays into any final state *except* $\pi^+\pi^-\Upsilon(1S)$; events where the $\Upsilon(3S)$ decays *only* to $\pi^+\pi^-\Upsilon(1S)$, excluding the leptonic final states of the $\Upsilon(1S)$; events where $\Upsilon(3S) \to \pi^+\pi^-\Upsilon(1S)$ and $\Upsilon(1S)$ decays to muons, electrons, and taus, respectively. The MC is normalized to the "Low" data sample luminosity.

background using the signal Monte Carlo (MC) sample and the “Low” signal sample sideband. We will also conduct a detailed study of the peaking backgrounds and use this to predict the expected number of undetected dipion transitions in the signal sample.

Chapter 5

Background Suppression Techniques

In this chapter we describe the refinement of the selection of invisible decays of the $\Upsilon(1S)$, henceforth referred to as the "signal" channel. We describe the variables used to discriminate between signal and background processes, including both simple selection cuts and a more advanced multivariate approach to selecting these events.

5.1 Candidate Variables

There are a number of variables that may be potentially valuable for rejecting backgrounds. Backgrounds are expected to exhibit several features that distinguish them from the signal. First, the kinematics of the pions should differ from pions that come from a real $\Upsilon(3S) \to \pi^+\pi^-\Upsilon(1S)$ transition. Second, we expect backgrounds to typically have more extra detector activity, either in the form of extra tracks, neutrals, or IFR activity. We study these possibilities using signal MC and data from the recoil mass sidebands, which are defined in Chapter 4.3.3. Our first approach is designed to remove the dominant non-peaking background.

However, this process is also extremely effective at removing the peaking background from non-leptonic decays of the $\Upsilon(1S)$. The remaining peaking backgrounds will be discussed in Chapter 7. We require events in the following comparisons to satisfy $m_{recoil} \in [9.4, 9.52]\,\mathrm{GeV}/c^2$, which will serve as our fitting region at the end of this analysis (representing approximately 20σ to either side of the recoil mass peak).

Before proceeding to describe the variables, we make one important comment. We perform a pre-cut on the dipion transverse momentum at $p_T < 0.1\,\mathrm{GeV}/c$. Very low-p_T backgrounds overwhelm signal for p_T less than this, and they are removed later anyway by our cuts (even very loose cuts). We therefore show the raw p_T distribution (Figure 5.1), but require $p_T > 0.1\,\mathrm{GeV}/c$ for all other plots. For completeness, we show the same variables without the pre-cut in Appendix A.

The variables that we study are as follows:

- Dipion kinematics and properties
  - Dipion transverse momentum: we compute the magnitude of the dipion system's transverse momentum, p_T, in the laboratory frame. We expect backgrounds with small p_T, such as two-photon fusion processes, to appear quite distinctly in this variable (Figure 5.1). We find excellent separation between sideband data and signal MC.
  - Dipion polar angle: we use the absolute value of the cosine of the dipion system in the laboratory frame (Figure 5.2).
  - Dipion vertex probability: the dipions are required to meet at a common vertex,

with a beamspot constraint. We compare the vertex χ^2 probability [35, 36, 37] for signal and background (Figure 5.3). We find that the background tends to contain more failed vertexes (low probability) than signal.

  - We check whether either pion passes a particle identification (PID) selector level and define a variable for each pion that is "true" if it does pass the selector at the desired level. We identify the pions using the selectors eKMVeryLoose (electron ID), muBDTLooseLoP (muon ID), or KKMVeryLoose (kaon ID) (for a description of PID selectors, see Chapter 4.3.1). We then define a single boolean variable that is true if either of the pions passes this PID selector level. These single booleans for each of the three PID types are shown in Figures 5.4-5.6. We observe that the background exhibits a significantly larger electron and muon contamination than the signal, and a slightly larger kaon contamination.

- Extra Neutral/Charged Information

  - The number of ChargedTracks in an event, including the pions (Figure 5.7). We anticipate that while all events must contain just two GoodTracksVeryLoose (GTVL), backgrounds will tend to include several additional tracks that fail GTVL but populate the ChargedTracks list.

  - The total extra neutral energy: we compute the sum in the center of mass (CM) frame of all calorimeter energy deposits (Figure 5.8) that are not associated with the signal reconstruction (in the case of the signal events, there are no such neutrals; for the one and two lepton control samples, Bremsstrahlung neutrals are excluded from

this calculation).

- The CM energy of the highest-energy neutral cluster (Figure 5.9)
- The mass of the highest-momentum pair of neutral clusters (Figure 5.10). This will reveal background from π^0 and η decay, specifically, in the content of the extra neutrals. We see that there is a significant contribution from both sources.
- The cosine of the angle between the highest-energy neutral cluster and the normal to the plane of the dipion system (Fig 5.11). This variable reveals any correlation in production between the highest energy neutral and the dipion system. Events with no extra photons are defaulted to a value of 1.1, just above the physical region.
- We compare the number of "Tight" EMC K_L^0 candidates (Figs. 5.12). We find that the EMC K_L^0 multiplicity is a good discriminant between signal and background (though it's expected to be correlated with other neutral EMC cluster variables). EMC objects are identified as being K_L^0 candidates based on a boosted decision tree algorithm [38]

5.2 A Note on Data Subsamples for Training

In the following sections, we will discuss the use of the aforementioned variables in rejecting non-peaking background. Before doing this, we need to describe our procedure for splitting the data and MC into subsamples for training, testing, and validating any procedure for using these variables.

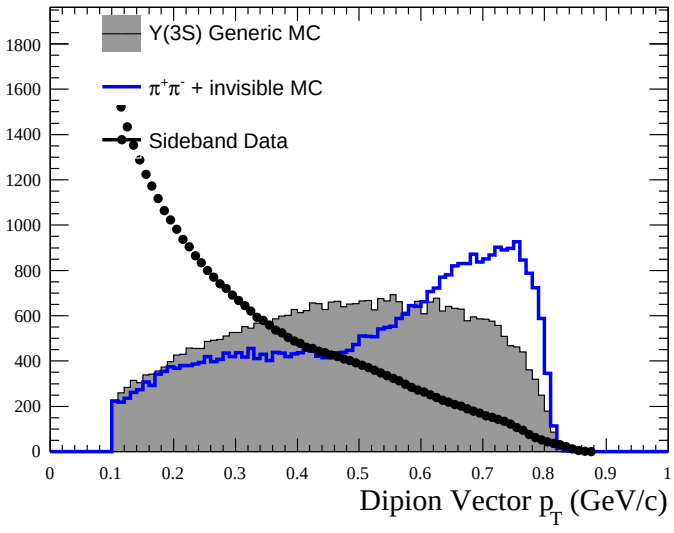


Figure 5.1: The dipion system p_T, compared between signal MC (blue, unshaded histogram), sideband data (black points), and $\Upsilon(3S)$ generic MC (grey, shaded histogram). All distributions are normalized to the yield in the sideband data, for a pure comparison of the shape.

We use the "Low" sample of data, and all signal MC from the signal v01 Monte Carlo (MC) dataset. After all selection cuts described in Chapter 4.3.2.1 and choosing events within the fit region ($m_{rec} \in [9.4, 9.52]$) of the recoil mass, we are left with 630k data sideband events and 70k signal events. We split these samples into three subsamples, two for training the background rejection algorithm and one for validating the performance of the algorithm. Two training samples are needed because the algorithm learns about signal and background from one, but its performance is measured (in order to correct the training) from the second. We will quote efficiencies and plot the output of multivariate discriminants using the validation sample, which is statistically independent of the other two samples and thus unbiased.

For the sake of simplicity, we exclude the "Low" sample from the final analysis of

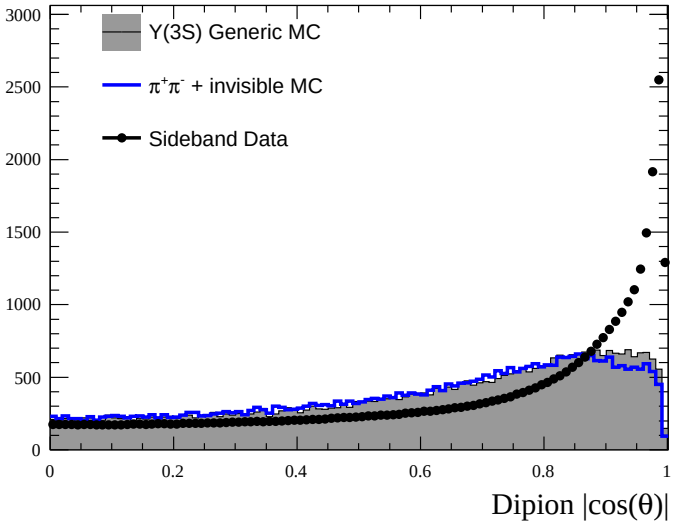
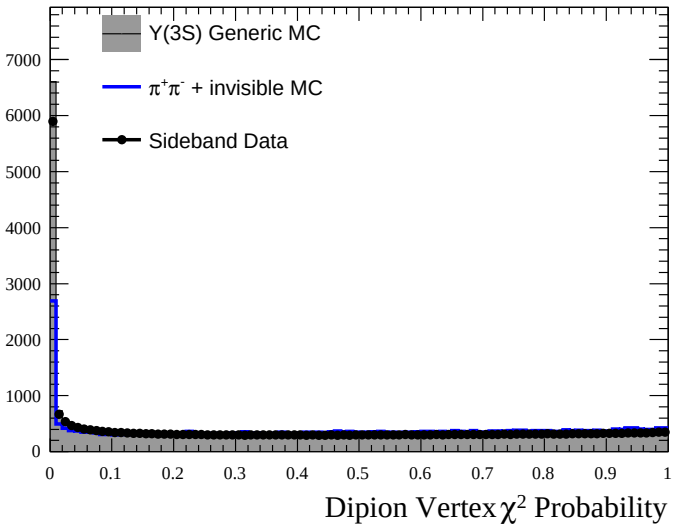
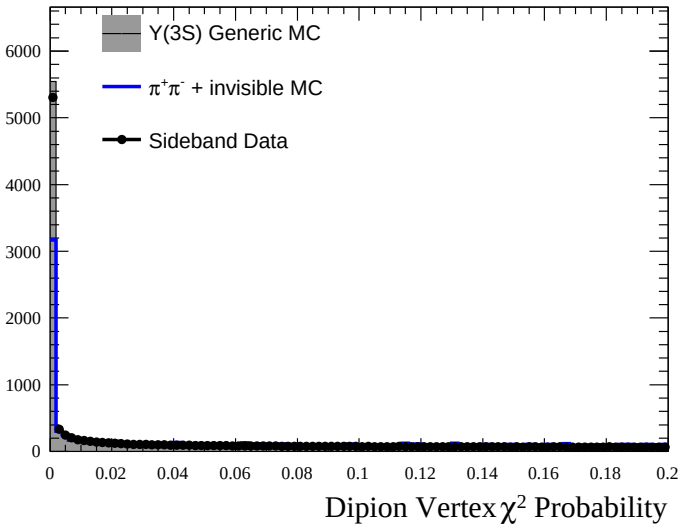


Figure 5.2: $|cos(\theta_{\pi\pi})|$, where $\theta_{\pi\pi}$ is the lab polar angle of the dipion system, compared between signal MC (blue, unshaded histogram), sideband data (black points), and $\Upsilon(3S)$ generic MC (grey, shaded histogram). All distributions are normalized to the yield in the sideband data, for a pure comparison of the shape.



(a) The whole probability range, 0.0-1.0



(b) A zoomed probability range, 0.0-0.2

Figure 5.3: The dipion vertex probability, compared between signal MC (blue, unshaded histogram), sideband data (black points), and $\Upsilon(3S)$ generic MC (grey, shaded histogram). All distributions are normalized to the yield in the sideband data, for a pure comparison of the shape.

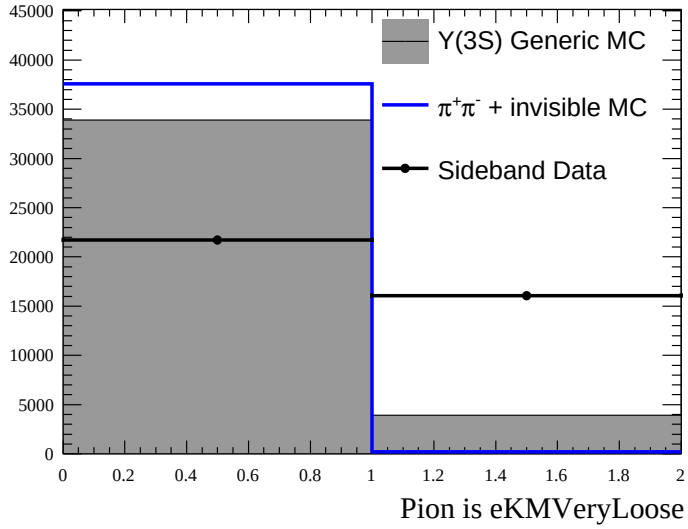


Figure 5.4: A boolean that is true if either pion passes eKMVeryLoose, compared between signal MC (blue, unshaded histogram), sideband data (black points), and $\Upsilon(3S)$ generic MC (grey, shaded histogram). All distributions are normalized to the yield in the sideband data, for a pure comparison of the shape.

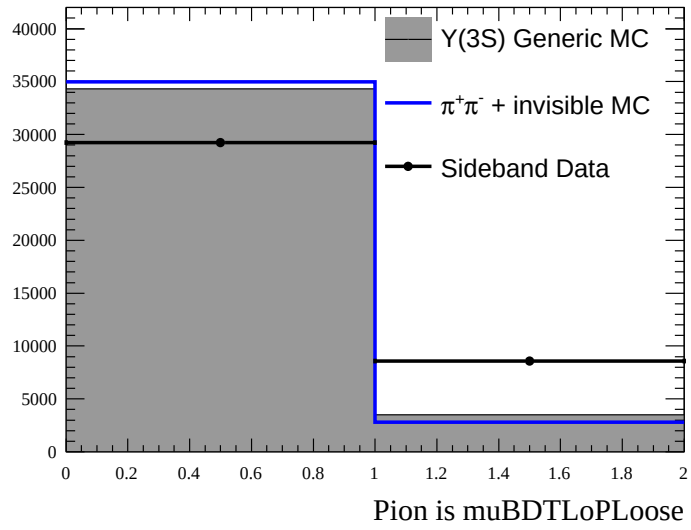


Figure 5.5: A boolean that is true if either pion passes muBDTLoPLoose, compared between signal MC (blue, unshaded histogram), sideband data (black points), and $\Upsilon(3S)$ generic MC (grey, shaded histogram). All distributions are normalized to the yield in the sideband data, for a pure comparison of the shape.

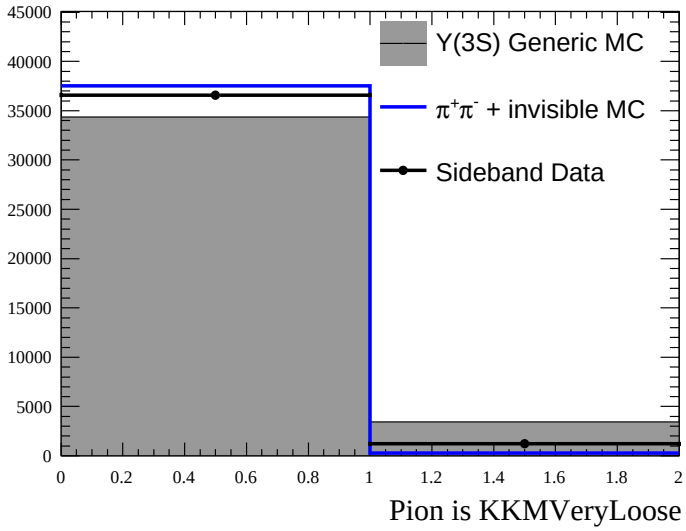


Figure 5.6: A boolean that is true if either pion passes KKMVeryLoose, compared between signal MC (blue, unshaded histogram), sideband data (black points), and $\Upsilon(3S)$ generic MC (grey, shaded histogram). All distributions are normalized to the yield in the sideband data, for a pure comparison of the shape.

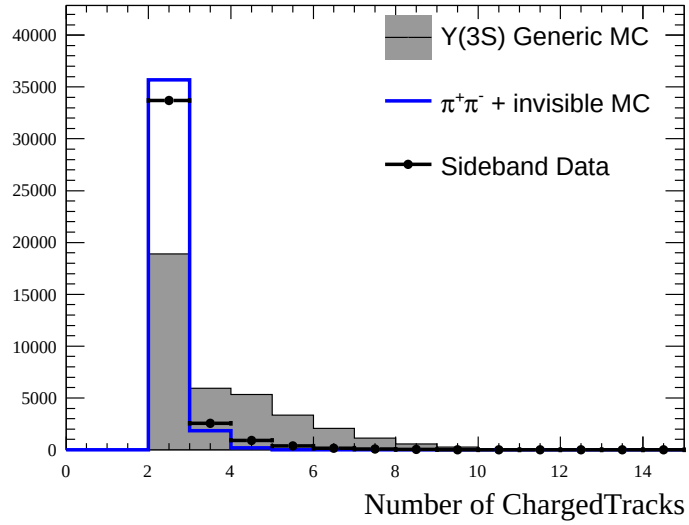


Figure 5.7: The number of ChargedTracks in the event, compared between signal MC (blue, unshaded histogram), sideband data (black points), and $\Upsilon(3S)$ generic MC (grey, shaded histogram). All distributions are normalized to the yield in the sideband data, for a pure comparison of the shape.

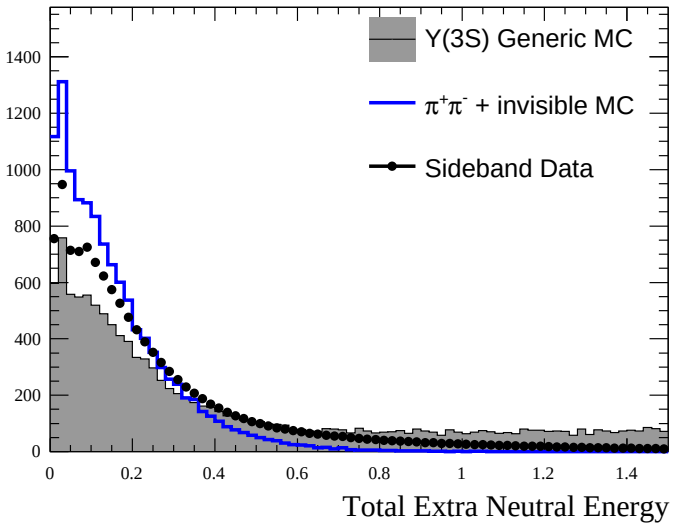


Figure 5.8: The extra neutral energy, compared between signal MC (blue, unshaded histogram), sideband data (black points), and $\Upsilon(3S)$ generic MC (grey, shaded histogram). All distributions are normalized to the yield in the sideband data, for a pure comparison of the shape.

the signal sample. This avoids having to disentangle the 1/3 of unbiased events from the biased 2/3 of the "Low" sample. Excluding the "Low" sample reduces the total $\Upsilon(3S)$ count from 96.5×10^6 to 91.4×10^6.

5.3 Candidate Variable Correlations

Since we will eventually be performing a fit to recoil mass to determine the total yield of invisible decays of the $\Upsilon(1S)$, it is imperative that the efficiency of the selection we develop be flat in m_{rec}. We first group variables into their two supercategories - dipion variables and extra energy variables - and plot their relative correlations and their correlations with the recoil mass. We have chosen the limited sample of available pion variables specifically to avoid strong

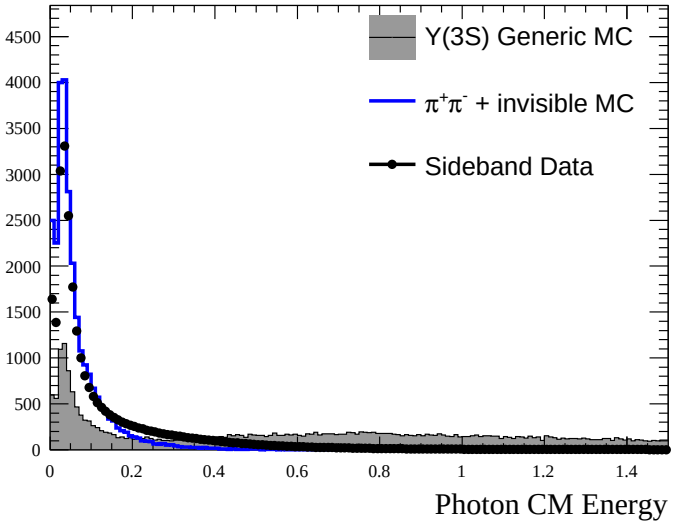
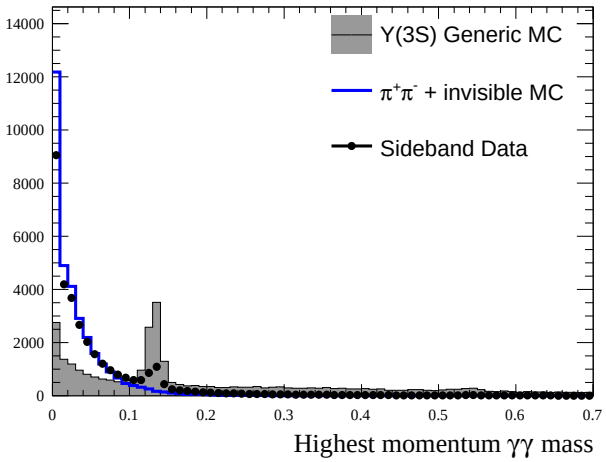
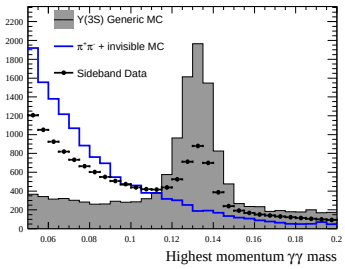


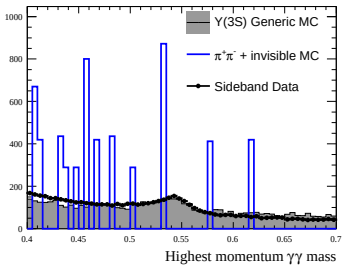
Figure 5.9: The center of mass (CM) energy of the highest-energy cluster, compared between signal MC (blue, unshaded histogram), sideband data (black points), and $\Upsilon(3S)$ generic MC (grey, shaded histogram). All distributions are normalized to the yield in the sideband data, for a pure comparison of the shape.



(a) A wide range of the two-neutral mass



(b) The π^0 mass region



(c) The η mass region

Figure 5.10: The mass of the highest-momentum (center of mass (CM) frame) neutral pair, compared between signal MC (blue, unshaded histogram), sideband data (black points), and $\Upsilon(3S)$ generic MC (grey, shaded histogram). All distributions are normalized to the yield in the sideband data, for a pure comparison of the shape.

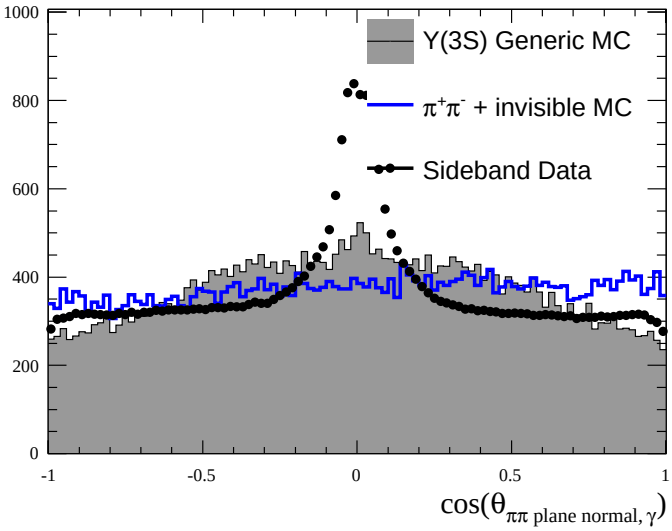


Figure 5.11: The cosine of the angle (center of mass (CM) frame) between the highest energy neutral and the normal to the plane of the dipion system, compared between signal MC (blue, unshaded histogram), sideband data (black points), and $\Upsilon(3S)$ generic MC (grey, shaded histogram). All distributions are normalized to the yield in the sideband data, for a pure comparison of the shape. Events that do not have any extra photons are defaulted to 1.1 and are not shown on this plot.

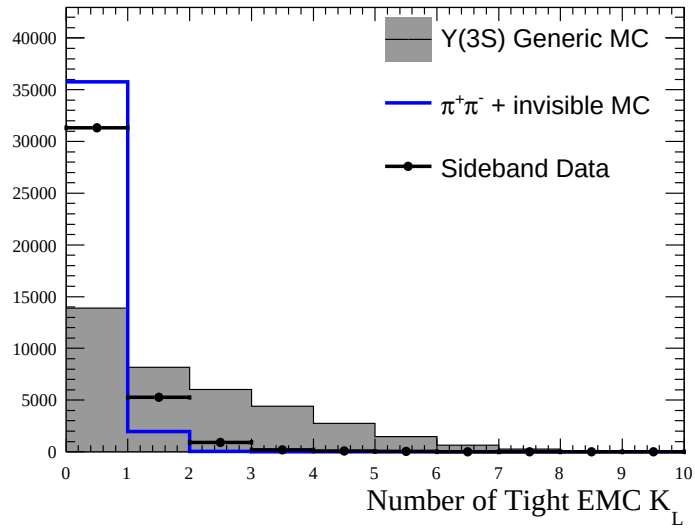


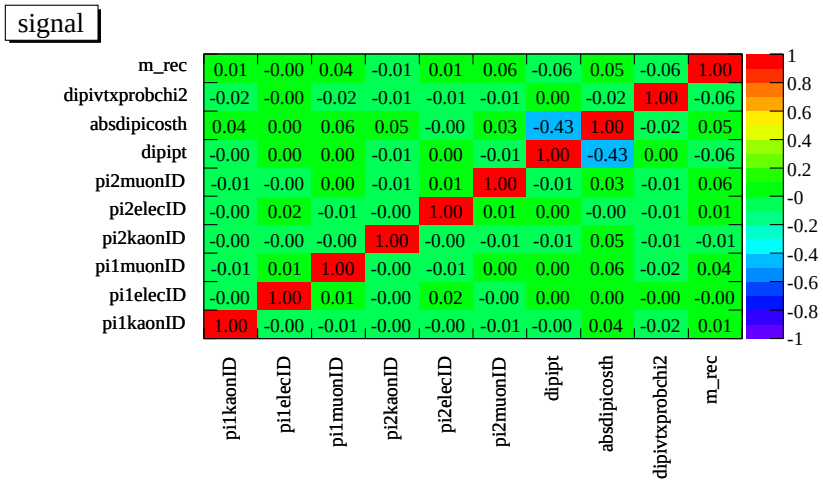
Figure 5.12: The multiplicity of EMC K_L^0 candidates, compared between signal MC (blue, unshaded histogram), sideband data (black points), and $\Upsilon(3S)$ generic MC (grey, shaded histogram). All distributions are normalized to the yield in the sideband data, for a pure comparison of the shape.

Ntuple Variable Name	Description
m_rec	Dipion recoil mass
dipivtxprobchi2	Dipion vertex probability
absdipicosth	$\lvert\cos(\theta_{\pi\pi})\rvert$
dipipt	Dipion p_T
pi[1,2]{kaon,elec,muon}ID	Pion PID
ntracks	ChargedTrack multiplicity, including the pions
nklemc	EMC K_L^0 multiplicity
g1pipicosthCMplane	cosine of the CM angle between the highest-momentum neutral cluster and the normal to the dipion plane
ggmass	Invariant mass of the highest-momentum pair of neutral clusters
eTotCalorNeutral	Total extra neutral energy in the CM frame
g1ECM	CM Energy of the highest-energy neutral cluster

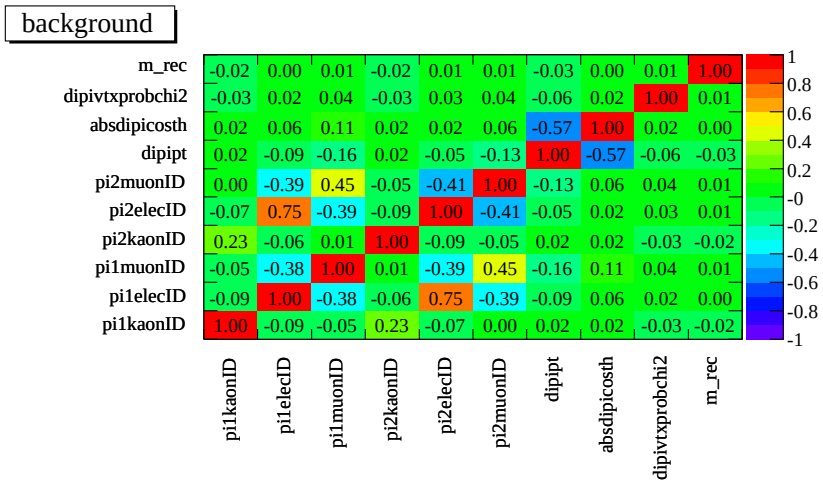
Table 5.1: The mapping between ntuple variable name and the variable description

correlations with the recoil mass, or a large number of weak ones that when taken together allow the background rejection algorithm to infer the value of the recoil mass in that event (and thus create a bias). The correlations of the PID, vertexing, and dipion kinematics are shown in Figure 5.13, while the correlations of the extra particles variables are shown in Figure 5.14. The key for deciphering the ntuple variable names is in Table 5.1.

The dipion variables are, for the most part, only weakly correlated with one another in the signal MC and sideband data (at the level of 6% or less), and also weakly correlated with the recoil mass. In the background, we see a pattern of correlations which are not present in the

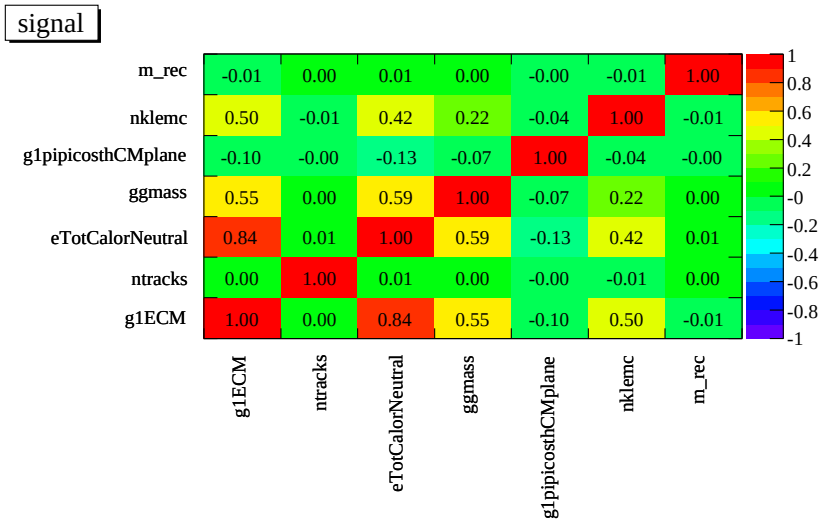


(a) Correlations in signal MC

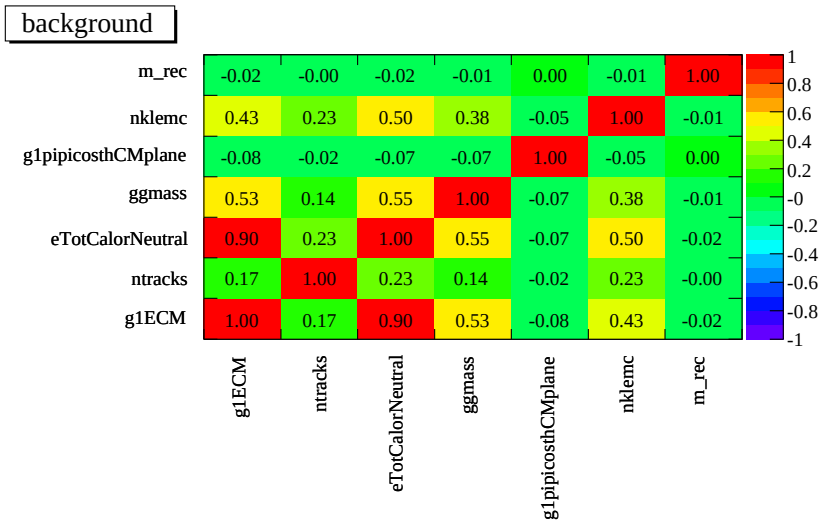


(b) Correlations in sideband data

Figure 5.13: The correlations between dipion variables in signal MC and sideband data. The plots are symmetric about the diagonal.



(a) Correlations in signal MC



(b) Correlations in sideband data

Figure 5.14: The correlations between extra energy variables in signal MC and sideband data. The plots are symmetric about the diagonal.

signal. The notable ones are:

- When one pion passes a certain PID selector, the other tends to as well
- A number of weak correlations, at the few percent level, across a wide range of variables that are not present at all in the MC signal sample – this should enable a multivariable algorithm to identify classes of background events.

Therefore, there is a lot of promise in these variables to reject backgrounds.

The extra particle variables are similarly very weakly or not at all correlated with recoil mass. However, in many cases they are strongly correlated with each other in both the signal and data, although from the plots above the nature of the correlation would tend to favor background rejection. We note that, as expected, the EMC K_L^0 multiplicity is strongly correlated with other neutral variables.

5.3.1 A Neural Net Study of Correlations

It is possible that while a group of variables may not be individually correlated with another variable, some function of those variable can be used to predict the variable in question. We are concerned that this may be the case with the input variables we choose for our analysis and m_{recoil}. In order to confirm that there is no way to reconstruct m_{recoil} from these inputs, we use a Neural Net algorithm trained on the same sideband data in the "Low" sample discussed earlier, with m_{recoil} as its output. We perform this training using several different sets of inputs to determine whether they can be used to reconstruct m_{recoil}. We then choose the maximal set that cannot reconstruct m_{recoil}. In Figure 5.15 the Neural Net has access to variables that allow

reconstruction of m_{recoil}. These extra variables are information about the individual pions in the events. Once information on the individual pions removed from the Neural Net (Figure 5.16), it cannot reconstruct m_{recoil}

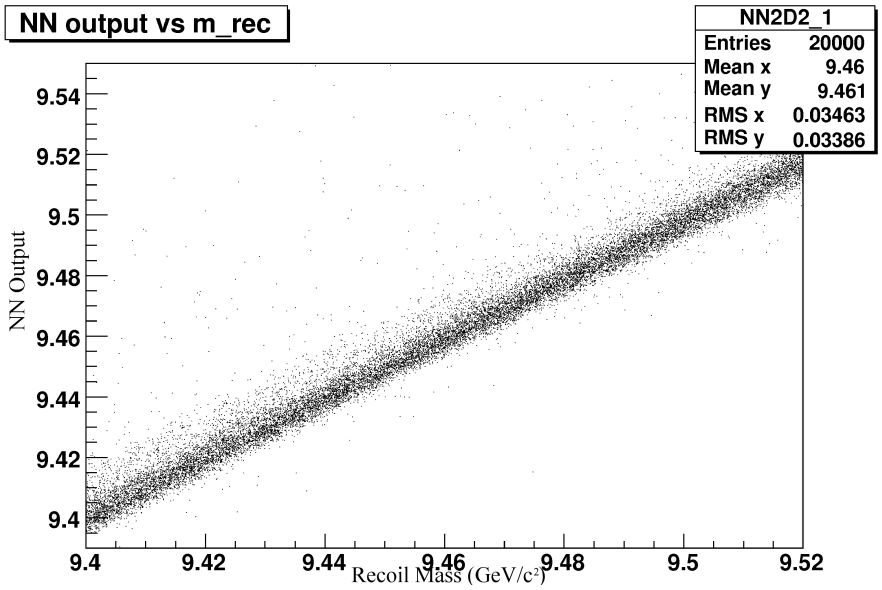


Figure 5.15: The output of a Neural Net algorithm trained to predict m_{recoil} that has access to variables allowing it to do so plotted vs the true m_{recoil}.

5.4 A Benchmark Cut-and-Count Analysis

We implement the PRIM algorithm [19] to create a preliminary cut-based selection routine. This algorithm makes multi-dimensional cuts on a sequence of variables in order to optimally reject background and select signal by searching for "bumps" in each variable that are signal-rich. We apply this algorithm to develop a baseline cut-and-count analysis on the aforementioned variables. This will serve as the standard against which we measure improvement in a more sophisticated multivariate approach. We want to make sure that there is a sensible

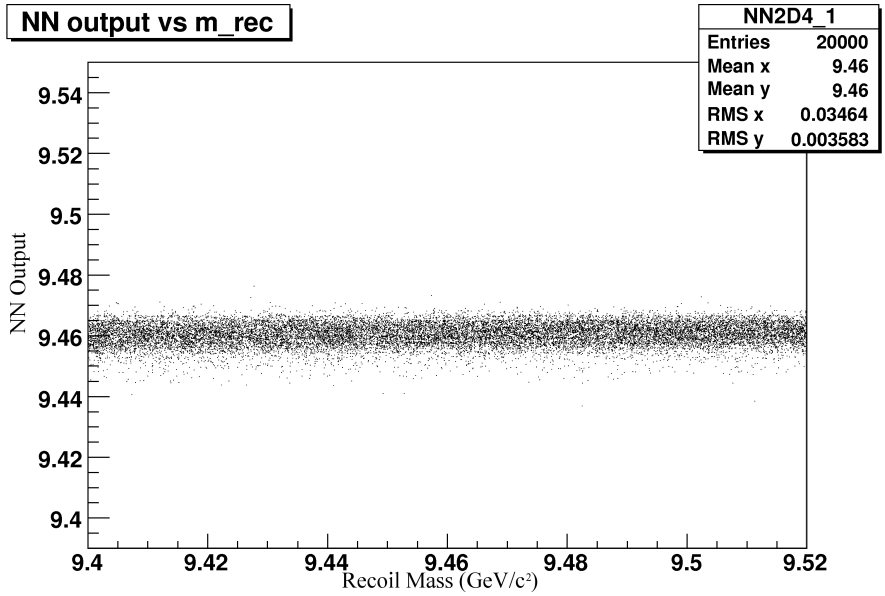


Figure 5.16: The output of a Neural Net algorithm trained to predict m_{recoil} that does not have access to variables allowing it to do so plotted vs the true m_{recoil}.

improvement from a multivariate approach.

We use the training and testing samples to train the algorithm, using the signal significance as the figure-of-merit (FOM) for the training. The signal is weighted to a target branching fraction of 5×10^{-4} for the optimization. There is one parameter that can be tuned to obtain the optimal FOM: the "peel parameter", which represents the maximum number of signal and background events that can be "peeled" away by a cut for a given cycle in the optimization. This is varied between 1% and 90%. The optimal peel parameter (maximizing the FOM) is found to be 50%. The cuts determined by the algorithm are shown in Table 5.2. The algorithm determines that with these cuts, it is not necessary to cut on the invariant mass of the highest-momentum pair of neutral clusters, the EMC K_L^0 multiplicity, or the angle between the highest-momentum neutral cluster and the normal to the dipion plane.

Variable	Selection
Kaon ID	fail
Electron ID	fail
Muon ID	fail
Dipion p_T	$[0.647, 0.821]\,\mathrm{GeV}/c$
$\lvert\cos(\theta_{\pi\pi})\rvert$	< 0.834
Dipion Vertex Probability	$> 6 \times 10^{-4}$
ChargedTrack Multiplicity	2
CM Energy of highest-energy photon	$< 0.170\,\mathrm{GeV}$
Total Extra Neutral Energy (CM)	$< 0.707\,\mathrm{GeV}$

Table 5.2: The selection criteria for optimal background rejection, as determined by the PRIM algorithm

We then apply these cuts to the validation sample and check the performance. We find a significance $S \equiv s/\sqrt{s+b}$ (where the s, the signal, is the signal Monte Carlo (MC), and b, the background, is the sideband data) of 0.205 in the validation sample, which represents a luminosity of just $0.4\,\mathrm{fb}^{-1}$. This represents 437 background and 4.31 signal events (assuming the previously mentioned signal branching fraction of 5×10^{-4}) in this integrated luminosity. These will be our benchmark number for a more complex multivariate algorithm. Also, note that we are computing this significance in a very wide region of the recoil mass ($m_{recoil} = [9.41, 9.52]\,\mathrm{GeV}/c^2$). This is the region in which we will later perform a maximum likelihood fit.

5.5 The Random Forest-based Multivariate Analysis (MVA)

We pursue a more advanced algorithm for rejecting background: the random forest. A "Random Forest" [20] is a method by which a number of decision trees [33] are trained, and the output of the algorithm is taken as the weighted vote of the output of each of the trees. The

training procedure sets the weights for the vote, to maximize the Gini index (the figure of merit (FOM) for this approach), which is related to the minimization of the loss of events from each category (signal categorized as signal, background as background). We use the same samples of events to train, test, and validate as were used for the cut-and-count approach. As an input to the Random Forest algorithm, we use the complete set of inputs defined in this chapter, and listed in Table 5.1. Each training cycle grows a decision tree from a random subset of the input variables - thus the name: "random forest".

We can control two parameters during the training process: the number of trees grown (training cycles) and the minimum number of events that are allowed to populate a terminal node of the tree (a node with no further splits). We fix the number of trees to 100 and try a variety of minimal events per terminal node, which we denote "l". Figure 5.17 shows the resulting training curves for the FOM vs. training cycle. We find the best performance (lowest FOM) for $l = 50$.

We then cross-check the performance of this algorithm against the cut-and-count analysis developed in the previous section. We fix the cut on the MVA output to 0.875 to obtain the same background yield and compare the signal yield and significance. We find that the MVA passes 3% more signal (4.43 events) for the same background, giving us a significance of 0.211 instead of 0.205.

We also cross-check the use of the "v01" signal MC to train and validate the algorithm. We compare the MVA output in the "v01" signal MC we used to train the MVA to the MVA output in the official signal MC (Figure 5.18). We find no significant differences between the two.

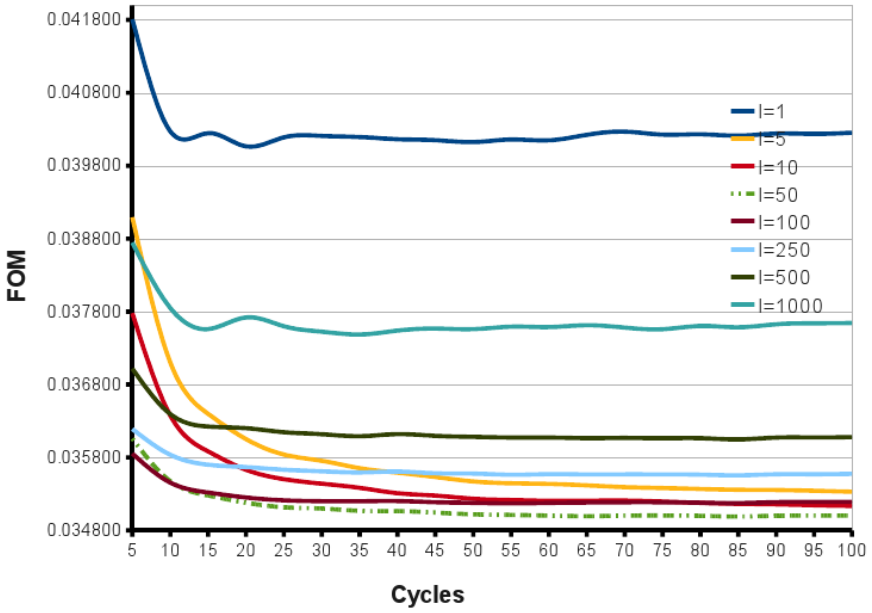


Figure 5.17: The figure of merit (FOM) of the random forest vs. the number of trees grown (training cycles), for a variety of minimal events allowed per terminal node ("l"). The optimal training configuration is shown as the dark green (dashed) line.

The output of the random forest algorithm, compared between signal MC and sideband data, is shown in Figure 5.19. We also compare the output of the algorithm just on the sideband data, for events in the upper and lower sidebands (Figure 5.20, and find no concerning differences between events from these two regions.

5.6 Signal Monte Carlo/Control Sample Comparison of MVA Output

The MVA can almost be applied to the control sample (two-lepton) events without changing its input configuration. We expect that most of the inputs will be very similar between signal MC and the control sample MC. However, this is not true for the track multiplicity vari-

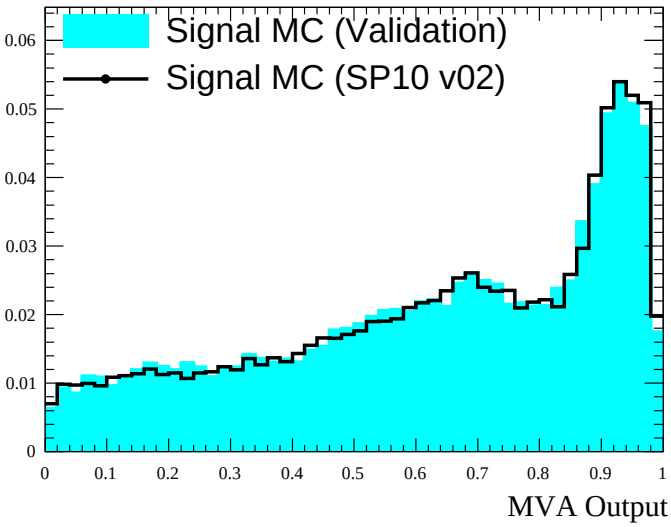
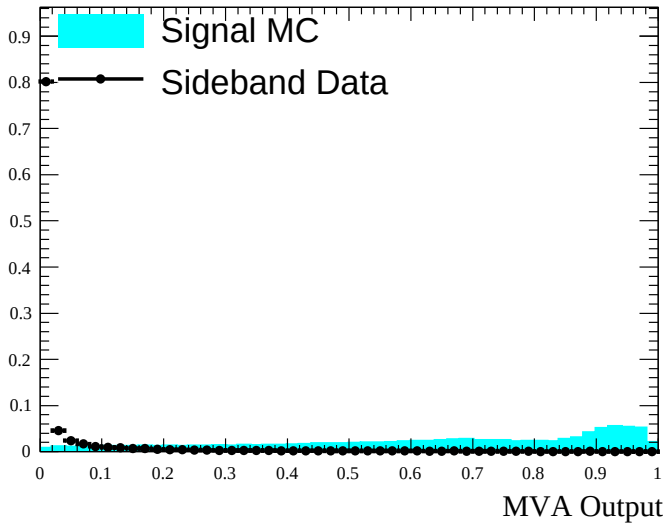


Figure 5.18: The output of the Random Forest (MVA) on signal MC from the "v01" validation sample (cyan, solid histogram) and the official signal MC (black, unfilled histogram). The plots are unit normalized over the MVA range of [0,1].

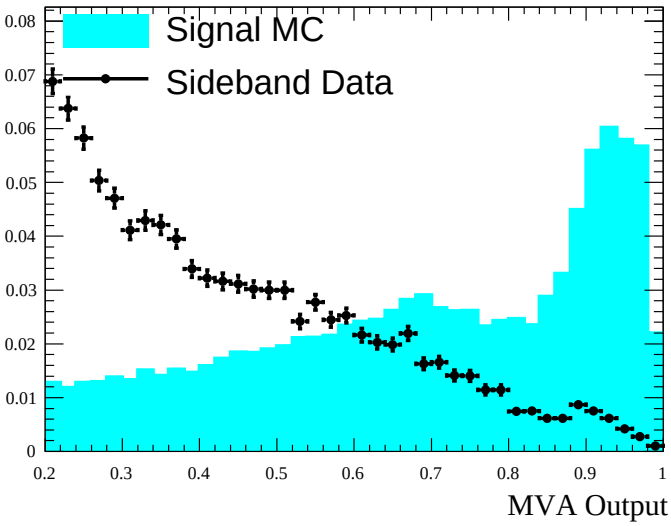
able. The track multiplicity contains the final-state leptons, which need to be subtracted. Therefore, for events reconstructed in the two-lepton final state we redefine $N_{tracks} \rightarrow N_{tracks} - 2$. We additionally ignore any photon used in Bremsstrahlung reconstruction so that this extra energy will not enter into the MVA inputs involving neutral objects. Later, when we discuss the one-lepton control sample, we will similarly modify this MVA input.

We can compare the overall shape of the MVA output between the signal MC and the two lepton control sample, and obtain a qualitative evaluation of the systematic difference between the MC and the data for a signal-like sample. The comparison of the MVA output is shown in Figure 5.21.

Qualitatively, the comparison of the distributions in this control sample is quite favor-

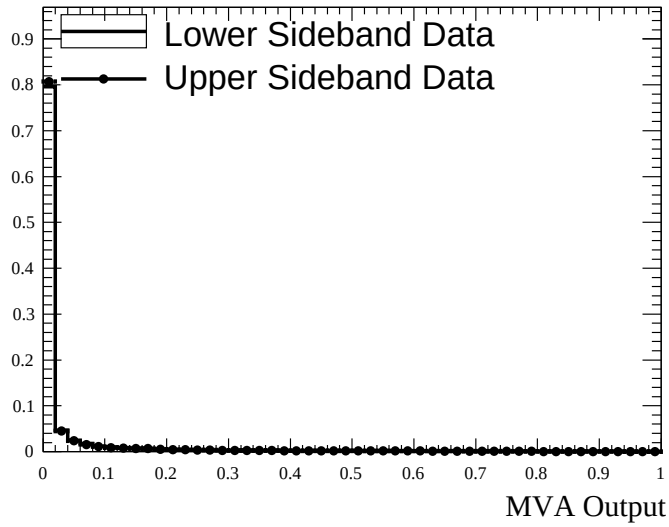


(a) Entire MVA output range

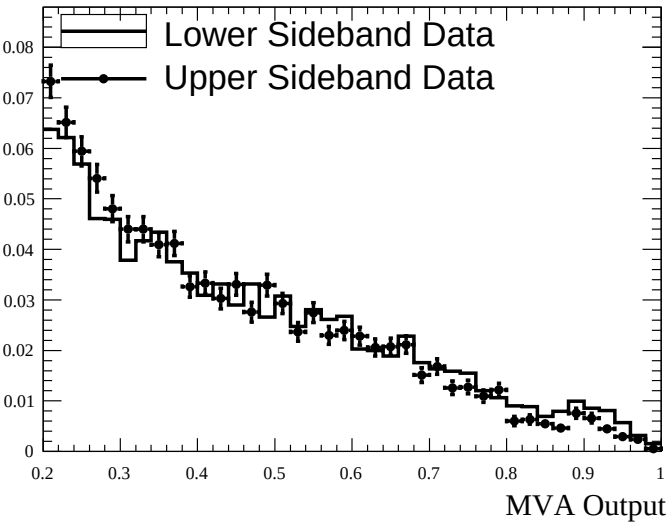


(b) MVA above 0.2

Figure 5.19: The output of the Random Forest (MVA) on signal MC and sideband data from the validation sample. The plots are unit normalized over the MVA range of [0,1].



(a) Entire MVA output range



(b) MVA above 0.2

Figure 5.20: The output of the Random Forest (MVA) on lower and upper sideband data from the validation sample. The plots are unit normalized over the MVA range of [0,1].

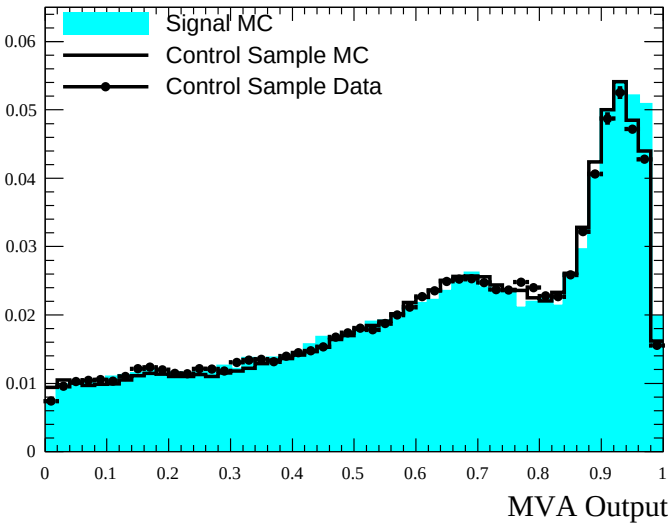


Figure 5.21: The output of the MVA compared between signal MC (cyan, filled histogram), control sample MC (black unfilled histogram), and the control sample data (black points). The plots are unit normalized over the MVA range of [0,1].

able. However, we observe clear shape differences above 0.6 between the signal MC and the two lepton control sample MC. While the control sample MC models the data almost perfectly, the signal MC and the control sample are not in perfect agreement. These differences are due to triggering effects and are investigated when we study systematic uncertainties in Chapter 8. There, we will use these samples to quantitatively calculate systematics errors in the use of this MVA.

5.7 Backgrounds After Loose Random Forest (MVA) Selection: Peaking Backgrounds

In Chapter 4.3.4, we studied the generic $\Upsilon(3S)$ MC and looked at the peaking and non-peaking background contributions from that source (Figure 4.7) prior to any additional cuts. We can now look at these events again, with the MVA applied to them. We are specifically interested in the contribution to the peaking background from the non-leptonic $\Upsilon(1S)$ decay. After the preliminary event selection, this source constituted 25% of the peaking background.

While there is a great deal of information on the $\Upsilon(1S) \rightarrow \ell^+\ell^-$ decays available, these decays only account for a small percentage of the total $\Upsilon(1S)$ branching fraction. There is a danger that the $\Upsilon(1S)$ may decay into much less well-understood non-leptonic states, and that these states will not be detected, contributing to our peaking background. In Figure 5.22 we plot the MVA output for various subsamples of the $\Upsilon(3S)$ generic MC, including both leptonic and non-leptonic $\Upsilon(1S)$ decays. We observe that most of the non-leptonic peaking background events also have a low-valued output from the MVA, and so we expect that once a cut has been applied, this background will be greatly suppressed. In order to make a preliminary estimate of the background we expect from this source, we apply a cut on the MVA output at 0.8 (which is determined to be the optimal cut in Chapter 6) to the generic $\Upsilon(3S)$ MC. We find that the leptonic final states dominate the peaking background sample, with less than 0.2 percent of the total contributed by non-leptonic sources after requiring $MVA > 0.8$, as shown in Figure 5.23.

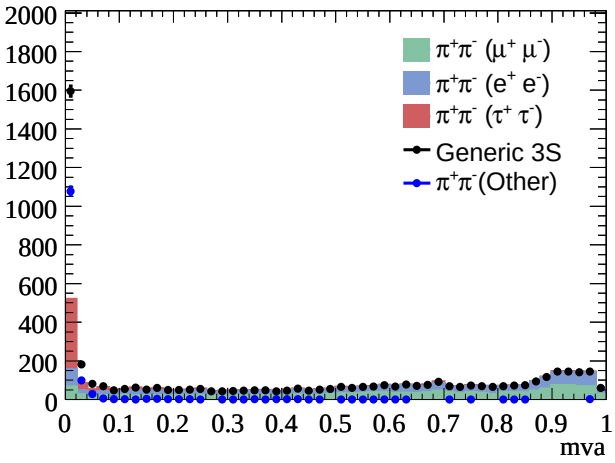
Since there are so few remaining non-leptonic Monte-Carlo events remaining, we list them here:

- $\Upsilon(1S) \to \pi^0 \bar{p} n K^+ \bar{K^0}$
- $\Upsilon(1S) \to K^{*+} K^- \gamma$
- $\Upsilon(1S) \to \omega \eta \pi^0$
- $\Upsilon(1S) \to \pi^- \bar{n} p \gamma$
- $\Upsilon(1S) \to \pi^0 \bar{n} n \gamma$

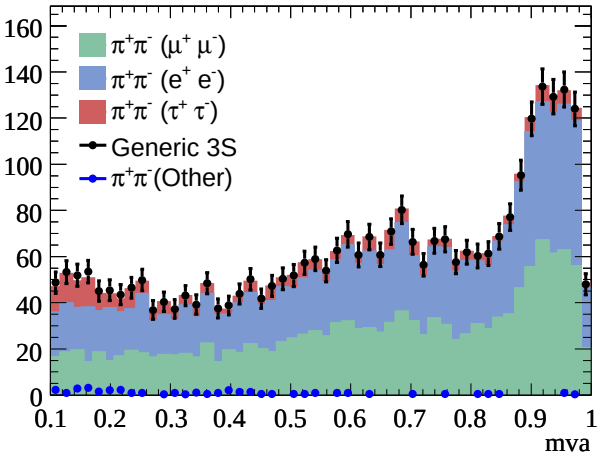
In Chapter 7 we use the one and two lepton control samples to calculate the expected peaking background from undetected leptonic decays of the $\Upsilon(1S)$. We also explore the uncertainties introduced into this measurement by the peaking backgrounds from both the $\Upsilon(1S) \to \ell^+\ell^-$ and the non-leptonic $\Upsilon(1S)$ decay modes using all three control samples: the leptonic control samples just mentioned, and the "near" invisible sample.

5.8 Backgrounds After Loose Random Forest (MVA) Selection: Other Backgrounds

We can now plot the sideband data and the sum of the Monte Carlo (MC) backgrounds after making a loose selection at MVA ¿ 0.5. The resulting distribution of the recoil mass is shown in Figure 5.24. The peaking component from the $\Upsilon(3S)$ MC contains all the subsamples described in the previous section, added together into a single distribution. Another example, with $MVA > 0.7$, is shown in Figure 5.25. For the cut at 0.5, we see the peaking background appears as a fairly prominent feature on top of the non-peaking background. However, non-peaking background is still clearly the dominant source of background events.



(a) All MVA values



(b) MVA above 0.1

Figure 5.22: Comparing the MVA output distributions for the leptonic and non-leptonic peaking backgrounds. We plot the entire generic $\Upsilon(3S)$ sample (black points), the subsample of events containing a dipion transition and a leptonic decay of the $\Upsilon(1S)$ (the stack of histograms), and overlay the contribution from non-leptonic $\Upsilon(1S)$ decays after a real dipion transition (the blue points). In the case where there are no events in the non-leptonic peaking background (blue points), there is no entry in the bin.

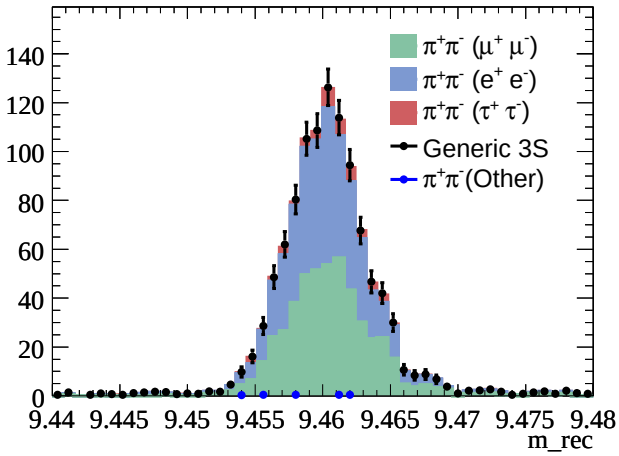


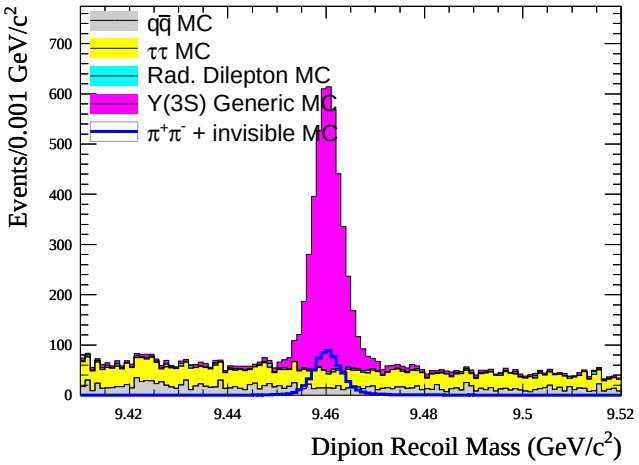
Figure 5.23: The recoil mass distribution of the peaking background components after a $MVA > 0.8$ cut. We plot the entire generic $\Upsilon(3S)$ sample (black points), the subsample of events containing a dipion transition and a leptonic decay of the $\Upsilon(1S)$ (the stack of histograms), and overlay the contribution from non-leptonic $\Upsilon(1S)$ decays after a real dipion transition (the blue points). In the case where there are no events in the non-leptonic peaking background (blue points), there is no entry in the bin.

For a tighter cut, the ratio of peaking and non-peaking background gets closer to 1.0, and the peaking background will be an even more prominent feature at this cut level. We will optimize the cut on the MVA using toy Monte Carlo experiments in Chapter 6. These studies will allow us to find the cut on the MVA output that maximizes our signal sensitivity.

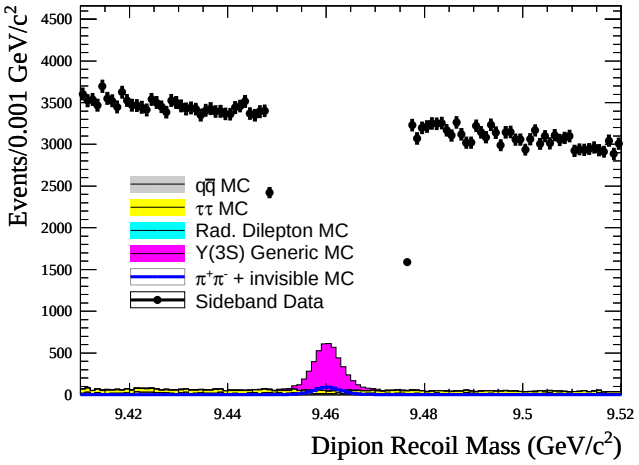
5.9 Behavior of MVA Input Variables After MVA cut

We plot the input variables to the Random Forest Algorithm (MVA) after a cut on the MVA output. Using the optimized value of the MVA cut determined in Chapter 6.5. In general, the MVA cut of 0.8 cuts quite hard on most of the input variables and shapes them to look signal-like for the sideband data. This suggests there is little extra gain to be had by any additional cuts on these input variables, and that the MVA has performed its intended function.

Since there non-peaking background has been suppressed as much as possible, we measure the remaining background using a fitting procedure (described in Chapter 6). This allows us to find the statistical error on the measurement. The systematic errors are calculated using the finalized selection routine. These errors are listed in Chapter 8

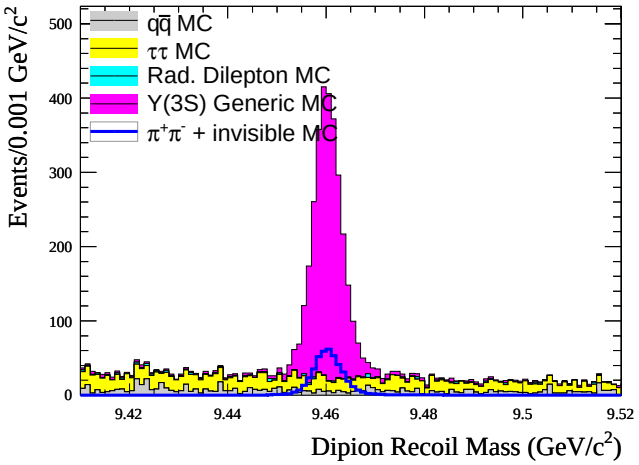


(a) Without sideband data

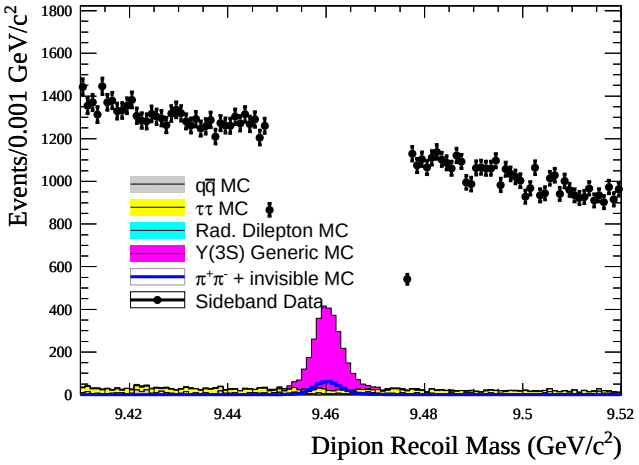


(b) Including sideband data

Figure 5.24: Recoil mass, after making an example selection on the MVA output of $MVA > 0.5$. Signal is normalized to a branching fraction of 5×10^{-4}.



(a) Without sideband data



(b) Including sideband data

Figure 5.25: Recoil mass, after making an example selection on the MVA output of $MVA > 0.7$. Signal is normalized to a branching fraction of 5×10^{-4}.

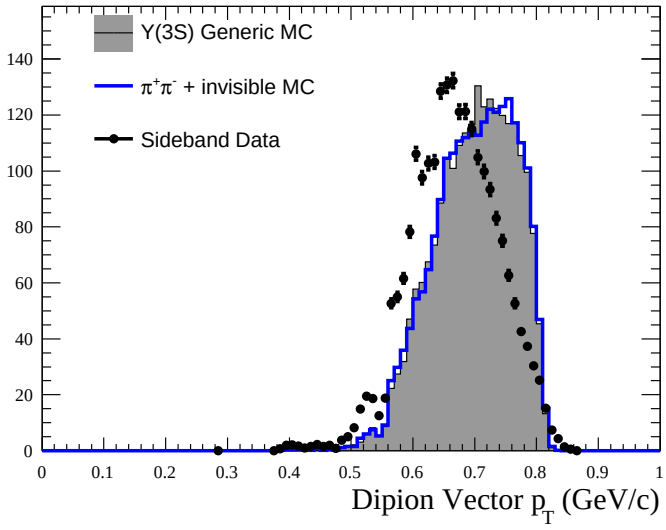


Figure 5.26: The dipion system p_T, compared between signal MC (blue, unshaded histogram), sideband data (black points), and $\Upsilon(3S)$ generic MC (grey, shaded histogram). All distributions are normalized to the yield in the sideband data, for a pure comparison of the shape.

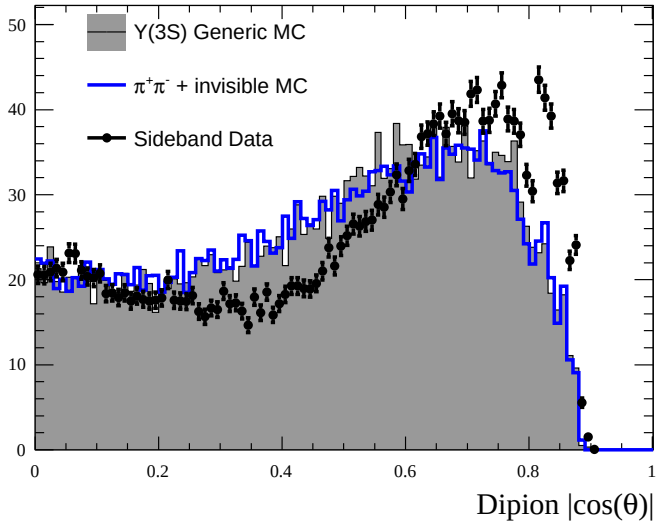


Figure 5.27: $|cos(\theta_{\pi\pi})|$, where $\theta_{\pi\pi}$ is the lab polar angle of the dipion system, compared between signal MC (blue, unshaded histogram), sideband data (black points), and $\Upsilon(3S)$ generic MC (grey, shaded histogram). All distributions are normalized to the yield in the sideband data, for a pure comparison of the shape.

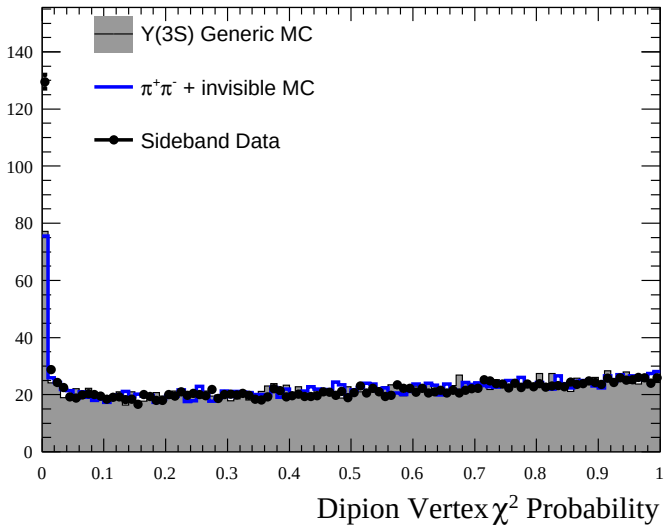


Figure 5.28: The dipion vertex probability, compared between signal MC (blue, unshaded histogram), sideband data (black points), and $\Upsilon(3S)$ generic MC (grey, shaded histogram). All distributions are normalized to the yield in the sideband data, for a pure comparison of the shape.

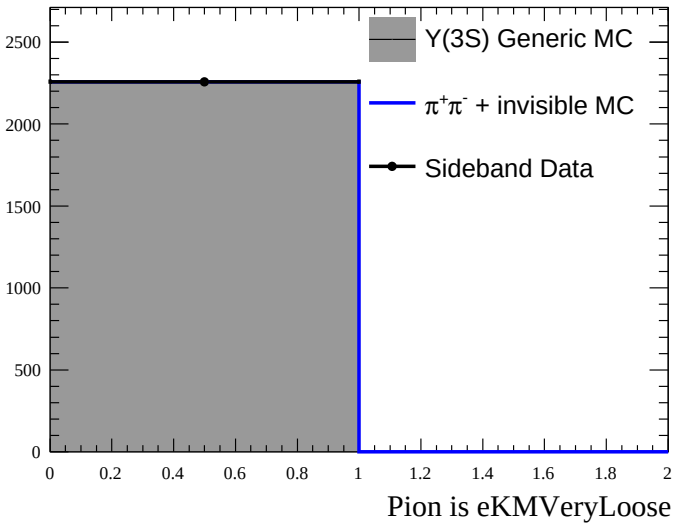


Figure 5.29: A boolean that is true if either pion passes eKMVeryLoose, compared between signal MC (blue, unshaded histogram), sideband data (black points), and $\Upsilon(3S)$ generic MC (grey, shaded histogram). All distributions are normalized to the yield in the sideband data, for a pure comparison of the shape.

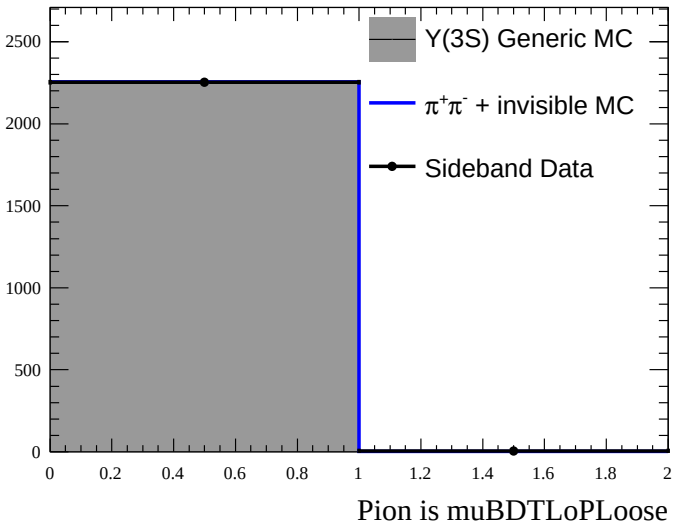


Figure 5.30: A boolean that is true if either pion passes muBDTLoPLoose, compared between signal MC (blue, unshaded histogram), sideband data (black points), and $\Upsilon(3S)$ generic MC (grey, shaded histogram). All distributions are normalized to the yield in the sideband data, for a pure comparison of the shape.

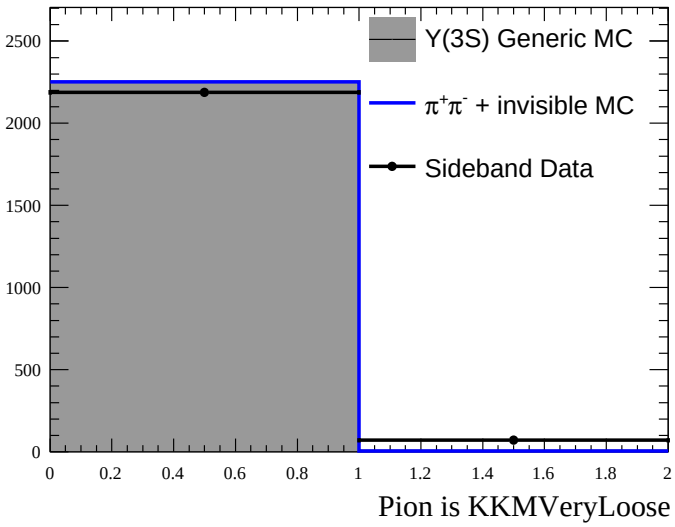


Figure 5.31: A boolean that is true if either pion passes KKMVeryLoose, compared between signal MC (blue, unshaded histogram), sideband data (black points), and $\Upsilon(3S)$ generic MC (grey, shaded histogram). All distributions are normalized to the yield in the sideband data, for a pure comparison of the shape.

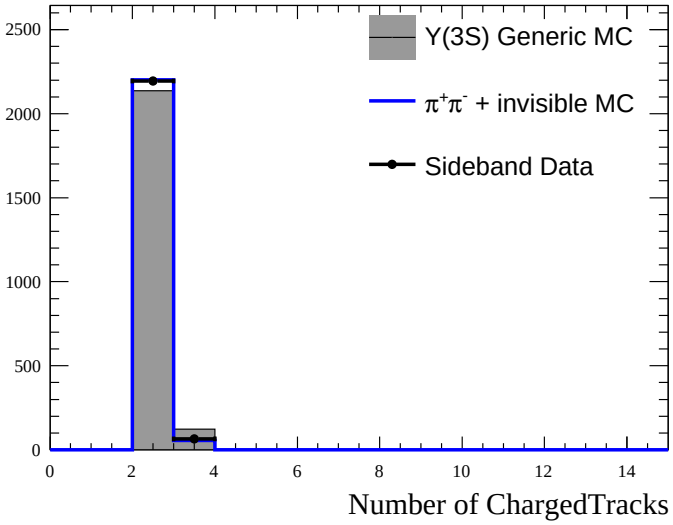


Figure 5.32: The number of ChargedTracks in the event, compared between signal MC (blue, unshaded histogram), sideband data (black points), and $\Upsilon(3S)$ generic MC (grey, shaded histogram). All distributions are normalized to the yield in the sideband data, for a pure comparison of the shape.

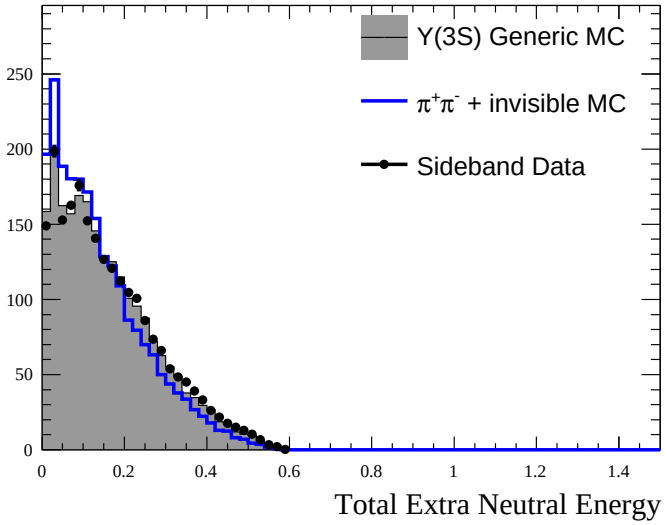


Figure 5.33: The extra neutral energy, compared between signal MC (blue, unshaded histogram), sideband data (black points), and $\Upsilon(3S)$ generic MC (grey, shaded histogram). All distributions are normalized to the yield in the sideband data, for a pure comparison of the shape.

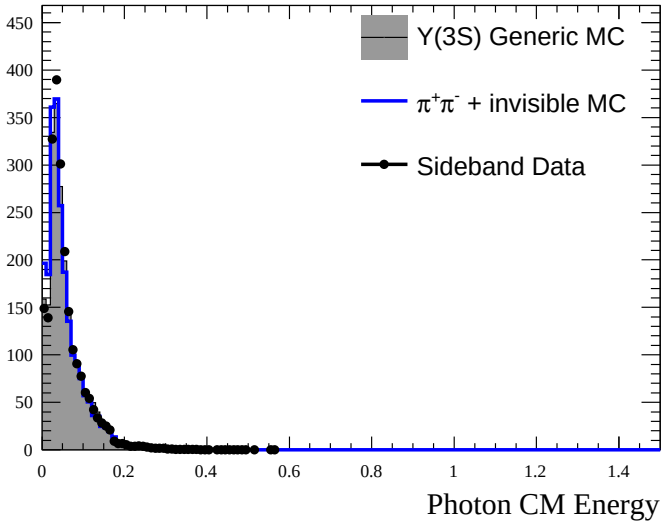
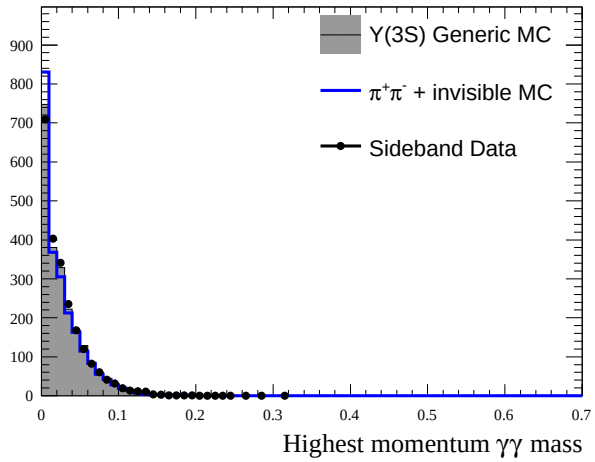
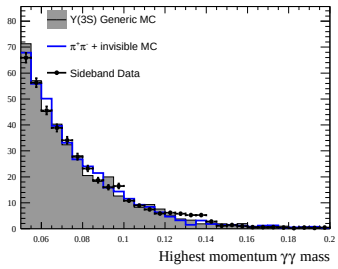


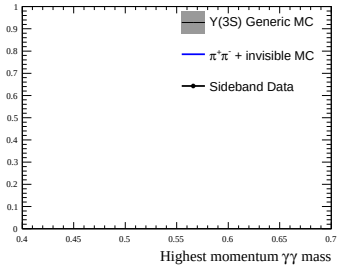
Figure 5.34: The CM energy of the highest-energy cluster, compared between signal MC (blue, unshaded histogram), sideband data (black points), and $\Upsilon(3S)$ generic MC (grey, shaded histogram). All distributions are normalized to the yield in the sideband data, for a pure comparison of the shape.



(a) A wide range of the two-neutral mass



(b) The π^0 mass region



(c) The η mass region

Figure 5.35: The mass of the highest-momentum (CM frame) neutral pair, compared between signal MC (blue, unshaded histogram), sideband data (black points), and $\Upsilon(3S)$ generic MC (grey, shaded histogram). All distributions are normalized to the yield in the sideband data, for a pure comparison of the shape.

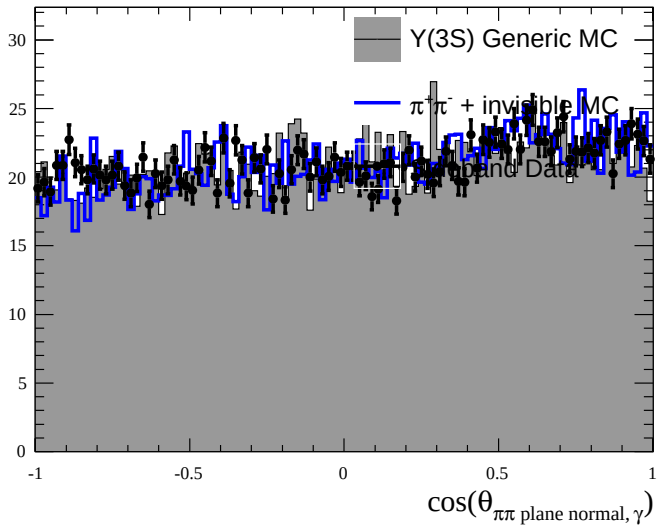


Figure 5.36: The cosine of the angle (CM frame) between the highest energy neutral and the normal to the plane of the dipion system, compared between signal MC (blue, unshaded histogram), sideband data (black points), and $\Upsilon(3S)$ generic MC (grey, shaded histogram). All distributions are normalized to the yield in the sideband data, for a pure comparison of the shape. Events that do not have any extra photons are defaulted to 1.1 and are not shown on this plot.

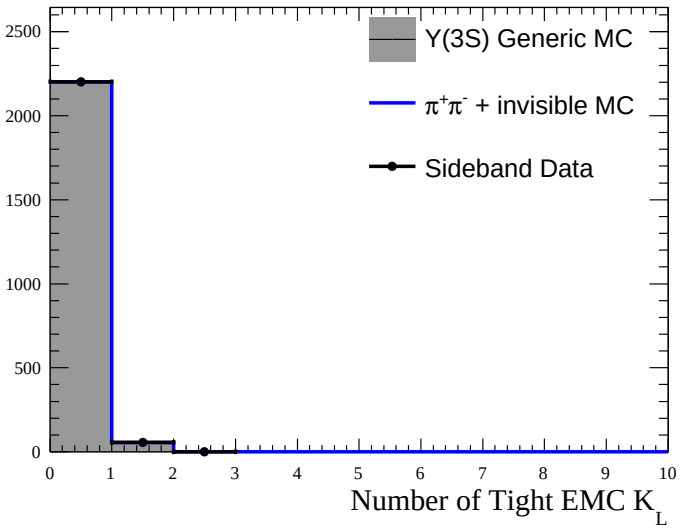


Figure 5.37: The multiplicity of EMC K_L^0 candidates, compared between signal MC (blue, unshaded histogram), sideband data (black points), and $\Upsilon(3S)$ generic MC (grey, shaded histogram). All distributions are normalized to the yield in the sideband data, for a pure comparison of the shape.

Chapter 6

Maximum Likelihood Fit of Recoil Mass, Statistical Signal Sensitivity Calculations, and Final Event Selection

We describe the maximum likelihood (ML) fit used to extract the signal events from the final signal selection in data. We develop the signal and peaking background models from data in the visible control sample, and the combinatoric background model from data in the recoil mass sideband. We use Monte Carlo experiments, called toys, to study the stability of the fit, and estimate the sensitivity of the final signal yield measurement.

6.1 Signal PDF

We construct the signal PDF by fitting the recoil mass in the two lepton control sample (as was done in Chapter 4.3.3) after making different levels of cuts on the Random Forest (MVA)

output. We scan over MVA cuts so that we can later optimize the cut on the output of the MVA using our full branching fraction extraction procedure. For each MVA cut, we fit the recoil mass and we compare the parameters of the model as a function of the cut (Table 6.1). We model the signal shape as the sum of a peaked shape - the Cruijff function - and a flat component; the fraction of the model occupied by the Cruijff is denoted $f_{peaking}$. The Cruijff function is given by,

$$\mathcal{C}(m_{recoil}) = \frac{1}{N} \begin{cases} \exp[-(m_{recoil}-\mu)^2/(2\sigma_L^2 + \alpha_L(m_{recoil}-\mu)^2)], & m_{recoil} < \mu \\ \exp[-(m_{recoil}-\mu)^2/(2\sigma_R^2 + \alpha_R(m_{recoil}-\mu)^2)], & m_{recoil} > \mu \end{cases} \tag{6.1}$$

There are slight changes in the parameterization of the Cruijff shape as a function of the cut on the MVA, while the fraction occupied by the Cruijff is relatively constant. An example fit, for $MVA > 05$, is shown in Fig. 6.1.

MVA Cut	μ	σ_L	σ_R	α_L	α_R	$f_{peaking}$
0.0	9.46101 ± 0.00004	0.00249 ± 0.00004	0.00301 ± 0.00003	0.14087 ± 0.00596	0.17412 ± 0.00346	0.96528 ± 0.00388
0.1	9.46099 ± 0.00004	0.00248 ± 0.00004	0.00301 ± 0.00003	0.14077 ± 0.00595	0.17136 ± 0.00356	0.96589 ± 0.00390
0.2	9.46099 ± 0.00004	0.00247 ± 0.00004	0.00301 ± 0.00003	0.14282 ± 0.00601	0.16945 ± 0.00370	0.96850 ± 0.00405
0.3	9.46099 ± 0.00004	0.00249 ± 0.00004	0.00300 ± 0.00003	0.14107 ± 0.00617	0.16796 ± 0.00377	0.96976 ± 0.00406
0.4	9.46098 ± 0.00004	0.00249 ± 0.00004	0.00300 ± 0.00003	0.14210 ± 0.00630	0.16540 ± 0.00398	0.97143 ± 0.00422
0.5	9.46096 ± 0.00004	0.00248 ± 0.00004	0.00301 ± 0.00004	0.14382 ± 0.00642	0.16197 ± 0.00429	0.97326 ± 0.00444
0.6	9.46093 ± 0.00004	0.00246 ± 0.00004	0.00301 ± 0.00004	0.14459 ± 0.00662	0.15869 ± 0.00467	0.97530 ± 0.00468
0.7	9.46091 ± 0.00004	0.00243 ± 0.00004	0.00299 ± 0.00004	0.14907 ± 0.00696	0.15577 ± 0.00536	0.97998 ± 0.00525
0.8	9.46090 ± 0.00005	0.00244 ± 0.00005	0.00296 ± 0.00005	0.14278 ± 0.00722	0.14363 ± 0.00613	0.98291 ± 0.00508
0.9	9.46079 ± 0.00006	0.00239 ± 0.00005	0.00299 ± 0.00006	0.13522 ± 0.00814	0.12094 ± 0.00833	0.98092 ± 0.00515

Table 6.1: Signal model parameters as a function of MVA output selection

The peaking background PDF is also taken to be this shape. We will fix the yield of the peaking background (calculated in Chapter 7) in the final maximum likelihood fit and fit for a signal excess on top of that component.

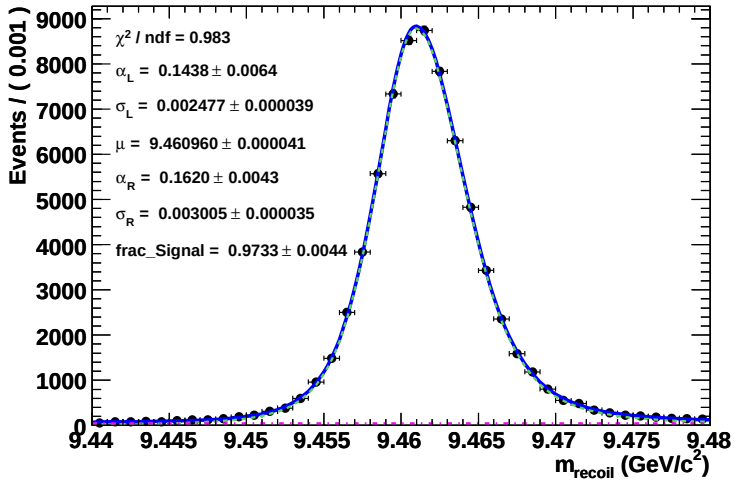


Figure 6.1: An example of fitting the recoil mass in the control sample for $MVA > 0.5$

6.2 Non-Peaking Background PDF

The non-peaking background shape is determined by a fit to the sideband data (the "Low" sample) using a linear function and interpolating through the blinding region. We fit the data using the function $P_1 + P_2 \times m_{recoil}$ using only events from the sideband. The linear shape, as a function of Random Forest (MVA) cut, is shown in Table 6.2. An example fit with $MVA > 0.5$ is shown in Fig. 6.2. We also compare the slope, as determined from the "Low" data sample, to that obtained by separately fitting the "High" sample. There is good agreement

MVA Cut	P_2 ("Low")	P_2 ("High")
0.0	0.03019 ± 0.00485	0.02199 ± 0.00034
0.1	0.01898 ± 0.00244	0.10354 ± 0.00654
0.2	0.00887 ± 0.07490	0.10845 ± 0.01506
0.3	-0.04109 ± 0.02817	-0.02572 ± 0.00950
0.4	-0.04651 ± 0.01134	-0.04299 ± 0.00736
0.5	-0.04315 ± 0.00534	-0.04878 ± 0.00517
0.6	-0.03484 ± 0.00275	-0.05067 ± 0.00359
0.7	-0.04700 ± 0.00232	-0.04444 ± 0.00203
0.8	-0.05093 ± 0.00181	-0.04766 ± 0.00093
0.9	-0.03338 ± 0.00122	-0.04599 ± 0.00205

Table 6.2: Non-peaking background model slope parameter as a function of MVA output selection

between these two, suggesting that any change in detector conditions is not strongly reflected in this procedure; and, this suggests that the MVA cut has the same effect on non-peaking background events taken earlier and later in the data set.

6.3 Signal Efficiency and Background Expectation

Signal efficiency, peaking background efficiencies, and bkg. yield are given in Table 6.3 as a function of the cut on the MVA output. The background yield is computed for a target luminosity, integrating the background model through the entire fit region for each MVA cut. The number of background at the target luminosity is extrapolated from the number of events in the "High" sample.

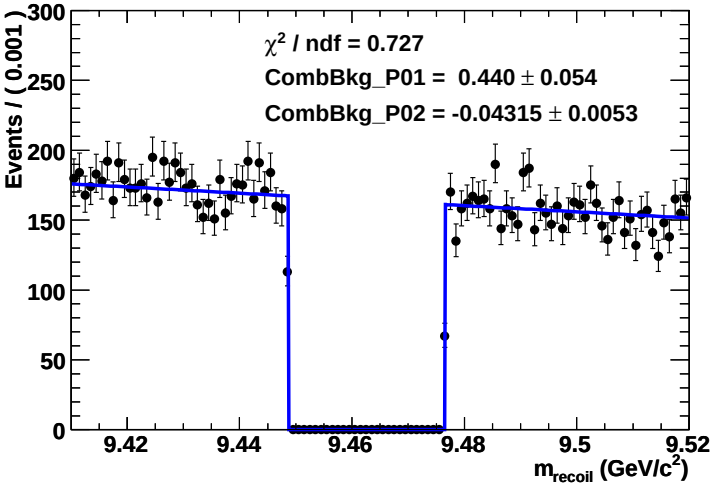


Figure 6.2: An example of fitting the recoil mass sideband data for the linear non-peaking background model, for $MVA > 0.5$

	Efficiencies				
MVA Cut	Signal	$\Upsilon(1S) \to e^+e^-$	$\Upsilon(1S) \to \mu^+\mu^-$	$\Upsilon(1S) \to \tau^+\tau^-$	Non-Peaking Bkg Yield (Projected from "High")
0.0	0.443	0.0301	0.0282	0.01182	7258226
0.1	0.423	0.0276	0.0263	0.00461	1576924
0.2	0.398	0.0257	0.0247	0.00366	1007835
0.3	0.373	0.0237	0.0229	0.00311	712514
0.4	0.344	0.0217	0.0210	0.00264	512332
0.5	0.309	0.0194	0.0187	0.00217	353689
0.6	0.267	0.0166	0.0161	0.00175	225725
0.7	0.214	0.0133	0.0130	0.00126	124047
0.8	0.164	0.0100	0.0098	0.00089	61512
0.9	0.102	0.0060	0.0060	0.00046	22471

Table 6.3: Signal and peaking background efficiencies, and the non-peaking background expectation, as a function of MVA cut

6.4 Example Uses of the Fit: $MVA > 0.5$

Before we optimize the cut, we explain how we conduct toy MC experiments and how we compute sensitivity using only statistical uncertainty. The following sections illustrate our techniques after making a cut of $MVA > 0.5$.

6.4.1 Toy Monte Carlo Studies

A "toy" Monte Carlo experiment is conducted using the signal and background PDFs described previously in this chapter. Random events are thrown from these distributions and added together to create a distribution that corresponds to a specific hypothesis about signal and background yields. This sample is then used to perform a maximum likelihood fit of the yields of the same PDF's used in its generation. This experiment is carried out many times to create a distribution of fitted signal yields for each signal hypothesis, which should be centered at the hypothetical yield. This enables us to determine what level of signal is required for us to detect any signal on top of a background.

We now study the behavior of the maximum likelihood function in the null signal hypothesis, to check for potential biases in the fit. We float the yield of signal events and non-peaking background events. The yield of peaking background is fixed for a given MVA cut using the following relationship:

$$N_{peaking} = N_{\Upsilon(3S)} \times \mathcal{B}(\Upsilon(3S) \to \pi^+\pi^-\Upsilon(1S)) \times \sum_{i=1}^{3} \mathcal{B}(\Upsilon(1S) \to \ell_i^+\ell_i^-) \times \varepsilon_i \qquad (6.2)$$

where the $N_{peaking}$ is the number of peaking background events, $N_{\Upsilon(3S)}$ is the number of $\Upsilon(3S)$ mesons, the sum over i is over the three $\Upsilon(1S) \to \ell^+\ell^-$ peaking background modes so

far under consideration, and ε_i is the inefficiency of the detector in each mode. For instance, for an MVA cut of 0.5, the number of leptonic peaking background events expected is about 4100.

We construct toy Monte Carlo experiments corresponding to the total luminosity available to the analysis. We generate and fit 1000 toy experiments, assuming the true branching fraction for the signal process is 0.0. We fit for the shape of the non-peaking background (floating the parameters of the shape), and the yields of the non-peaking background and signal. The peaking background is fixed to the nominal MC prediction. The pull distribution for the fitted signal yield is shown in Fig. 6.3. We find the pull distribution to be centered at 0.0 with a width statistically consistent with 1.0. We show the mean and width of the pull distribution as a function of input branching fraction hypothesis in Table 6.4. There is no trend suggesting obvious bias in the fit.

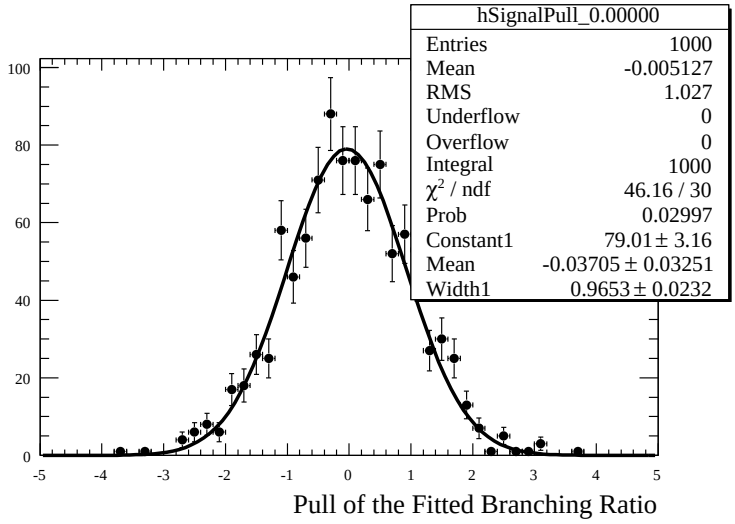


Figure 6.3: The pull distribution for 500 toy Monty Carlo experiments generated assuming the null signal hypothesis

	Pull Distribution	
Branching Fraction Hypothesis (Number of Signal)	Mean	Width
0.00000 (0)	-0.037 ± 0.033	0.965 ± 0.023
0.00010 (127)	0.004 ± 0.033	0.983 ± 0.026
0.00025 (317)	0.001 ± 0.032	0.973 ± 0.026
0.00050 (633)	-0.039 ± 0.032	0.995 ± 0.023
0.00075 (950)	0.030 ± 0.030	0.908 ± 0.025
0.00100 (1266)	0.012 ± 0.033	1.001 ± 0.025
0.00250 (3165)	0.024 ± 0.031	0.941 ± 0.023

Table 6.4: The variation of the mean and width of the pull distribution as a function of input branching fraction hypothesis (and, correspondingly, the input number of true signal events). The requirement on the MVA is $MVA > 0.5$.

6.4.2 Expected Sensitivity

We explore the potential sensitivity of this analysis in two scenarios: where we measure a branching fraction, and where we observe nothing and set an upper limit. We adopt the method of Cousins and Feldman [39] and construct confidence belts, where for a given input value of the true branching fraction we determine an allowed region of the fitted branching fraction. The branching fraction is expressed in terms of the signal efficiency, fitted signal yield, and the number of $\Upsilon(3S)$ as follows:

$$\mathcal{B}(\Upsilon(1S) \to \text{invisible}) = \frac{n_{\text{invisible}}}{N_{\Upsilon(3S)} \cdot \mathcal{B}(\Upsilon(3S) \to \pi^+\pi^-\Upsilon(1S)) \cdot \varepsilon_{\text{invisible}}} \tag{6.3}$$

where $n_{\text{invisible}}$ is the fitted yield of events with an invisible final state, $\varepsilon_{\text{invisible}}$ is the efficiency for reconstructing invisible final states, and $\mathcal{B}(\Upsilon(3S) \to \pi^+\pi^-\Upsilon(1S))$ is the branching fraction for the dipion transition to the $\Upsilon(1S)$.

We build confidence belts for each true branching fraction value by generating 500 toy experiments per hypothesis. We fit for the number of signal and fit the signal yield distribution using a single Gaussian. This approximates the fitted signal distribution using a smooth function (the Gaussian), which makes the next step easier. We then integrate the Gaussian from the left and the right and find the values of $\mathcal{B}_{fitted}(\Upsilon(1S) \to \text{invisible})$ that give us the probability of lying below (above) those values equal to $\alpha/2$, where the the Confidence Level $(CL) = 1 - \alpha$ defines the desired confidence level. If we want to measure the branching fraction, we construct 68% confidence belts; if we want an upper limit, we use 90% confidence belts. We construct these only assuming statistical uncertainty.

We choose a number of benchmark branching fraction hypotheses ((0.0, 1.0, 2.5, 5.0,

MVA Cut	Upper Limit ($\times 10^{-4}$)	Branching Fraction Error ($\times 10^{-4}$)
0.5	2.8	1.6
0.6	2.6	1.6
0.7	2.6	1.6
0.8	2.5	1.6
0.9	2.6	1.7

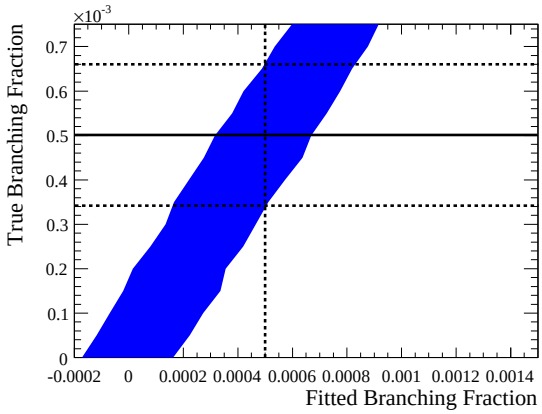
Table 6.5: The results of computing the upper limit or error on the branching fraction for each cut on the MVA output

7.5, 10.0, 25.0)$\times 10^{-4}$) and interpolate between the confidence belts for these points to get the confidence belts in-between. The 68% and 90% confidence belts are shown in Fig. 6.4. Based on these confidence belts, one expects to either measure a true branching fraction of 5×10^{-4} with an uncertainty of 1.6×10^{-4} (3.1σ statistical significance) or to set an upper limit in the absence of signal of $\mathcal{B}(\Upsilon(1S) \rightarrow \text{invisible}) < 2.8 \times 10^{-4}$.

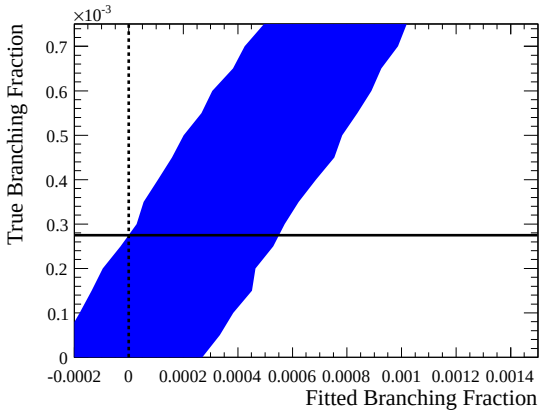
6.5 Optimization of the MVA Selection

We optimize the cut on the MVA output by scanning over cut values and performing the aforementioned toy MC study for each cut. We compute the expected upper limit and the error on the branching fraction (for $\mathcal{B} = 5 \times 10^{-4}$) for each cut, and compare them. We find (Table 6.5) that the error and the upper limit improve or remain the same until we move past $MVA > 0.8$. We therefore place the cut at $MVA > 0.8$ for the remainder of the analysis.

After performing the optimization, we look again for biases in the branching fraction measurement as a function of the input signal branching fraction. An example of a pull distribution for the null hypothesis is shown in Fig. 6.5. Table 6.6 shows the change in the mean and width of the pull distribution as a function of signal branching fraction.



(a) 68% confidence belts. The solid vertical line indicates the expected signal yield for a branching fraction of 5×10^{-4} and the horizontal lines the 1σ region on the branching fraction measurement.



(b) 90% confidence belts

Figure 6.4: Confidence belts constructed from toy Monte Carlo experiments for a number of true branching fraction hypotheses.

	Pull Distribution	
Branching Fraction Hypothesis	Mean	Width
0.00000	-0.014 ± 0.032	0.988 ± 0.024
0.00010	-0.048 ± 0.033	1.012 ± 0.026
0.00025	-0.010 ± 0.034	1.050 ± 0.028
0.00050	-0.039 ± 0.032	0.966 ± 0.022
0.00075	-0.015 ± 0.033	1.007 ± 0.024
0.00100	-0.018 ± 0.032	0.977 ± 0.022
0.00250	0.005 ± 0.033	0.998 ± 0.026

Table 6.6: The variation of the mean and width of the pull distribution as a function of input branching fraction hypothesis. The requirement on the MVA is $MVA > 0.8$.

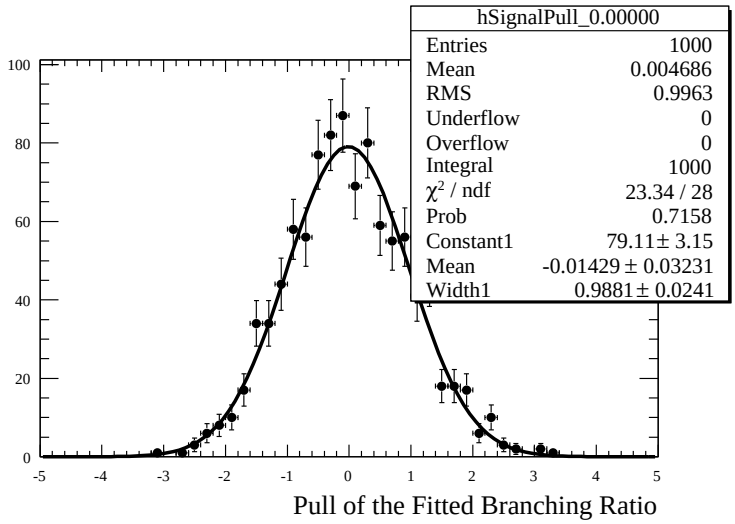


Figure 6.5: The pull distribution for 1000 toy Monte Carlo experiments generated assuming the null signal hypothesis, with $MVA > 0.8$

6.6 Limit of Sensitivity Due to Peaking Background Uncertainty

We here consider the physical limit of our sensitivity due to statistical uncertainty on the peaking background. Even if we could reject all non-peaking background, the peaking background would essentially be irreducible. Therefore, we can ask to what branching fraction our current analysis is limited before we can no longer improve the measurement. Given the efficencies for selecting peaking background at a cut of $MVA > 0.8$ (yielding 2119 MC-predicted leptonic peaking background events), we must observe a number of signal events, N_s, equal to a multiple of the error on the peaking background. If we consider the statistical error only and want to obtain a signal yield with a sigificance of $N_\sigma = N_s/\sqrt{N_s + N_{peaking}}$, then we find that the number of signal we observe must satisfy,

$$N_s = \frac{N_\sigma^2}{2}\left(1 + \sqrt{1 + 4N_{peaking}/N_\sigma^2}\right). \tag{6.4}$$

For the peaking background we estimated above, and for $N_\sigma = 3$, this this limits us to a branching fraction measurement that is no smaller than

$$\mathcal{B}(\Upsilon(1S) \to \text{invisible}) = \frac{n_{\text{invisible}}}{N_{\Upsilon(3S)} \cdot \mathcal{B}(\Upsilon(3S) \to \pi^+\pi^-\Upsilon(1S)) \cdot \varepsilon_{\text{invisible}}} = 2.1\times10^{-4}. \tag{6.5}$$

where $\varepsilon_m athrminvisible = 0.164$, $\mathcal{B}(\Upsilon(3S) \to \pi^+\pi^-\Upsilon(1S))$ = 0.0448, and $N_{\Upsilon(3S)} = 91.42 \times 10^6$. In other words, the statistical error on this analysis coming from the peaking background component is quite a bit smaller than the statistical error coming from the non-peaking peaking background.

However, the systematic uncertainty on the number of peaking background events is a more important problem. In the case of a systematic error the number of signal we have to

observe to achieve a given significance N_σ is given as $N_s = N_\sigma \times \sigma_{syst,}^{peaking} \times N_{peaking}$, where $\sigma_{syst,}^{peaking}$ is a percent systematic error. Even a 10% uncertainty on the peaking background limits the sensitivity of this analysis (again, at the 3σ level) to 9.5×10^{-4}. Controlling this systematic will be critical to interpreting the result.

6.7 Final Fit Configuration

The toy Monte Carlo (MC) approach has allowed us to determine the optimal configuration for our Random Forest. When we perform the final fit on the unblinded data, we will use the maximum likelihood fit to propagate the systematic errors through the yield calculation. Thus, once the systematic errors have been determined, they will be incorporated into the fit automatically, and the toy MC approach will not be needed to calculate the confidence intervals.

Chapter 7

Peaking Background

As discussed in Chapter 6.6, our understanding of the central value and uncertainty on the peaking background (coming from real dipion transistions ($\Upsilon(3S) \rightarrow \pi^+\pi^-\Upsilon(1S)$) where the $\Upsilon(1S)$ decays into visible final states that escape detection) is critical to the unbiased extraction of the yield of invisible signal events, as well as to the limitation of our sensitivity to a non-zero $\Upsilon(1S) \rightarrow \mathrm{invisible}$ branching fraction. In this section, we discuss our method for calculating our expected peaking background yield and determining a systematic uncertainty on that predicted background.

7.1 Naive Peaking Background Estimates

We can obtain a naive peaking background estimate by using $\Upsilon(3S)$ generic Monte Carlo (MC) simulation. After applying all selection criteria, the MC simulation predicts that we should expect peaking background from these sources following a real dipion transition: $\Upsilon(1S) \rightarrow e^+e^-$ (1019.5 events), $\Upsilon(1S) \rightarrow \mu^+\mu^-$ (1007.4 events), $\Upsilon(1S) \rightarrow \tau^+\tau^-$ (91.8

events), and $\Upsilon(1S) \rightarrow \mathrm{other}$ (2.9 ± 1.3 events). Thus, based on the assumption that the MC simulation is correct, we expect 2121.6 peaking background events, with the dominant contribution coming from electronic or muonic final states (48.1% and 47.5%, respectively), followed by the $\tau^+\tau^-$ final state (4.33%), and then other final states (0.14%).

However, we do not want to rely on the unchecked simulation for these predictions, especially when we have access to control samples from real data. There may be improperly-simulated detector or trigger effects, and there may be non-leptonic $\Upsilon(3S)$ decays which are not (or not properly) modelled by the generator. In principle, the latter effects could be quite substantial. We use the two lepton, one lepton, and "near" invisible control samples (defined in Chapter 4) in data to check the MC predictions.

7.2 Studies of the Leptonic Peaking Backgrounds

For a significant fraction of dilepton decays of $\Upsilon(1S)$, both leptons may escape detection, thus appearing just like an invisible decay. We can study this contribution by comparing data to Monte Carlo (MC) simulation for the large sample of visible events with both leptons detected, and also the substantial sample of events in which only one lepton is detected. The goal is to obtain a data/MC correction factor for the leptonic peaking background.

7.2.1 Control Samples for $\Upsilon(1S) \rightarrow \ell^+\ell^-$ Final States

We expect the dominant effects that cause an incorrect prediction of the peaking background to arise form two sources: an overall scale factor, based on inherent uncertainties in the

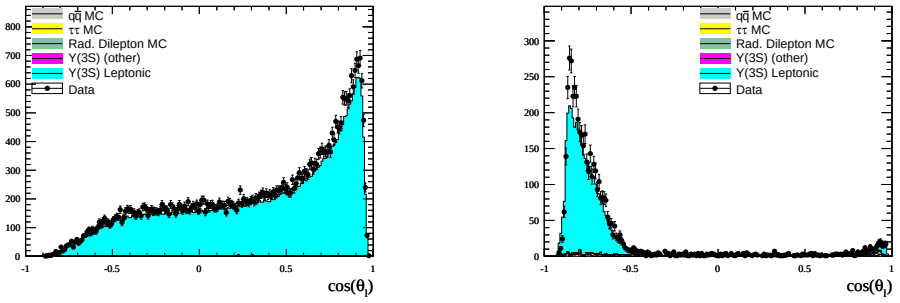
$\Upsilon(3S)$ count (about 1%), and branching fractions (at the level of about 5%), and detector acceptance mis-modeling. We can test both of these cases using control samples in which we reconstruct either two or one of the final-state leptons from decay of the $\Upsilon(1S)$.

In Chapter 4.3.2.2 we discussed the selection of events with one or two reconstructed hard tracks, assumed to be leptons from $\Upsilon(1S)$ decay . We refine the selection of these events by applying all cuts described in the discussion of the selection of invisible final states, including the cut on the modified Random Forest (MVA) output (as defined in Chapter 5.6) for these control sample events (Chapter 5.6).

7.2.2 Correcting for an Overall $\Upsilon(3S)$ Scaling Factor

After all selections are applied, we obtain 38833 events in the 2-lepton category and 4816 events in the 1-lepton category in data. We show the distribution of the cosine of the polar angle of the lepton (in the 2-lepton case, it's the polar angle of the positively charged lepton) in Figure 7.1. This figure shows that in both cases we have at least an overall normalization discrepancy between the MC prediction and the data in these events. We correct for this scaling factor using the 2-lepton sample. We determine the data and MC yield of events in the region $\cos\theta \in [-0.3, 0.3]$ (Figure 7.2), which is safely within the barrel of the detector and contains a large number of events. Tracking inefficiency of these tracks is negligible, as can be seen from the very small occupancy in the 1-lepton sample (Figure 7.1(b)) in this angular range. We find the Monte Carlo (MC) underpredicts the data yield, and needs to be scaled upward by 1.0882. Henceforth, we apply this scaling factor to $\Upsilon(3S)$ MC events. We will apply this scaling factor to the prediction of peaking background events, as well.

We show the lepton polar angle distribution after rescaling the MC simulation to the data in Figure 7.3. Overall, the MC models the data very well in shape. The modeling in the front-end is shown to be particularly good, suggesting that the acceptance modeling of the detector is excellent in the forward direction. The 2-lepton sample, however, has almost no statistics at the back-end of the detector. We check the modeling of the detector at the back-end in the next sample, using an independent 1-lepton control sample.



(a) 2-lepton control sample

(b) 1-lepton control sample

Figure 7.1: Cosine of the lepton polar angle in the lab frame in the 2-lepton and 1-lepton control samples before applying scaling correction.

7.2.3 Aside – The ${}_s\mathcal{P}lots$ Technique

We briefly pause here to discuss a method for computing the distributions contributed to a data set by different species of events. The ${}_s\mathcal{P}lots$ [13] method allows one to use a fit of signal and background distributions in one or more variables to project the signal and background distributions in another variable. Give a set of data which is a sum of different species of events (in the case of this analysis signal and background), an Extended Maximum Likelihood fit is performed on one or more of the observed variables (in the case of this analysis

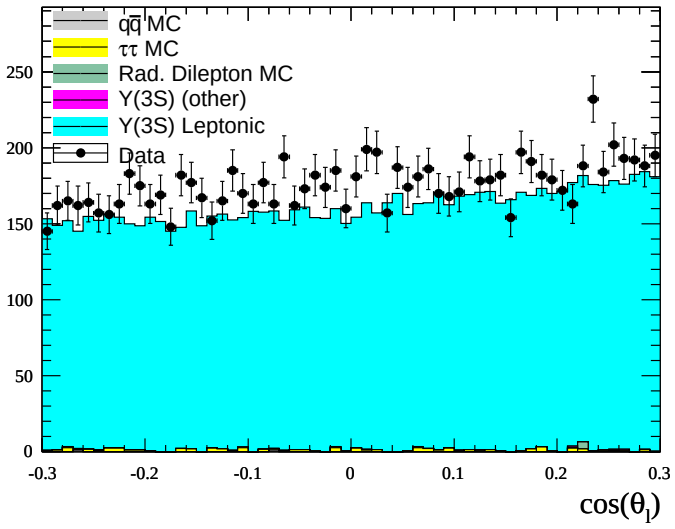
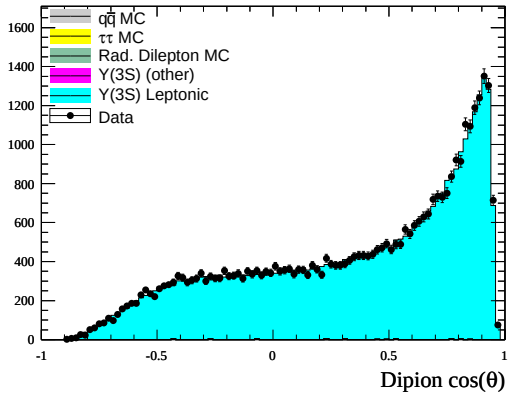


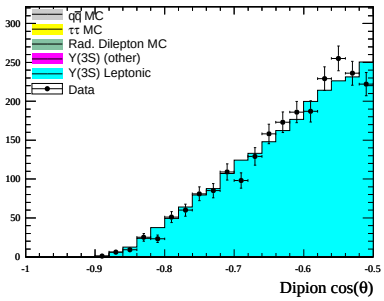
Figure 7.2: Cosine of the lepton polar angle in the lab frame in the 2-lepton control sample, shown in the center of the detector barrel and before applying scaling correction.

we perform a fit on m_{recoil}). The ${}_s\mathcal{P}lots$ technique takes this fit and the resulting correlation matrix, and computes a vector of weights for each event based on its position in the space of fitted observables. These ${}_s\mathcal{W}eights$ correspond to the species of events populating the sample. In the case of the 1-lepton sample about to be explored we have signal (from the process $\Upsilon(3S) \rightarrow \pi^+\pi^-(\Upsilon(1S) \rightarrow \ell^+\ell^-)$) ${}_s\mathcal{W}eights$ and background (from other processes) ${}_s\mathcal{W}eights$. Essentially, these ${}_s\mathcal{W}eights$ tell us how signal like or background like an event is. Using these weights, we can fill a histogram in a variable that is not used in the Likelihood fit with the ${}_s\mathcal{W}eights$ of the species we wish to study. In the case of the 1-lepton sample we will be creating a histogram of $\cos\theta$ filled with the signal ${}_s\mathcal{W}eights$ of events, creating a plot of 1-lepton $\cos\theta$ variable contributed only by the process $\Upsilon(3S) \rightarrow \pi^+\pi^-(\Upsilon(1S) \rightarrow \ell^+\ell^-)$.

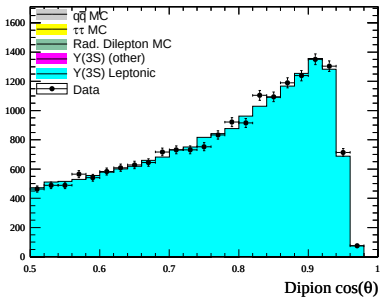
The ${}_s\mathcal{P}lots$ mathematics assumes that the variable(s) that are used in the Likelihood



(a) Positively charged lepton in 2-lepton events



(b) Back-end of the detector



(c) Front-end of the detector

Figure 7.3: Cosine of the positively charged lepton polar angle in the lab frame in the two-track control sample, after renormalizing MC to data in the region $\cos\theta = [-0.3, 0.3]$.

fit are uncorrelated with the variables being ${}_s\mathcal{P}lotted$. In Appendix B, we explore the effects of correlations on the effectiveness of the ${}_s\mathcal{P}lots$ technique. We find that for the use in the 1-lepton sample, this may introduce a small error which is calculated in Chapter 7.2.4.2, but that when used later for the full invisible fit, care must be taken to account for correlation effects.

7.2.4 Checking Detector Acceptance Using the 1 Lepton Control Sample

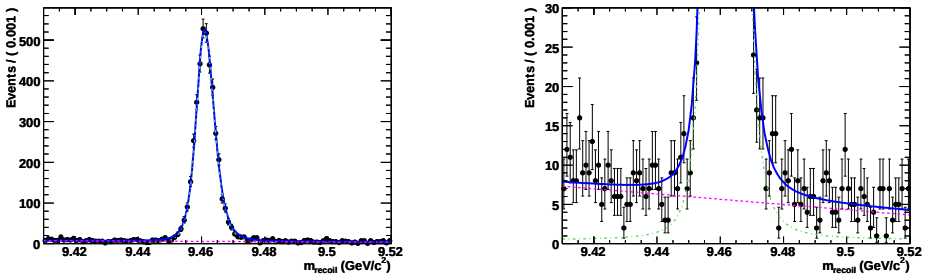
The primary concern with our estimate of the peaking background from electronic or muonic decays of the $\Upsilon(1S)$ is when one or both leptons are very close to the edge of the detector. If the detector edge is not well-simulated, we might be incorrectly predicting the rate at which final-state particles go from being detected to being undetected, and thus fake the invisible signal.

We study the detector edge using the 1-lepton events. As illustrated in Figure 7.1, the reconstructed one-lepton events occur primarily when the lepton is in the backward end of the detector. Due to the boost, even if the leptons are going back-to-back in the CM frame the forward-going lepton will tend to be boosted outside of the detector acceptance while the backward-going lepton will be boosted into the detector acceptance. The prevailance of backward-going 1-lepton events confirms this assumption.

We can see from the MC that the 1-lepton events contain a non-negligible fraction of peaking-background events, including a potentially unmodeled fraction in the far-backward direction, just before the efficiency drop for the 1-lepton events. In order to subtract these events from the polar angle distribution, we fit the recoil mass distribution of these events and unfold the polar angle distribution for peaking events using the ${}_s\mathcal{P}lots$ method [13]. The fit to the recoil

mass is shown in Figure 7.4 and the unfolded $\cos\theta$ distribution is shown in Figure 7.5 overlaid on $\Upsilon(3S)$ MC only. The recoil mass fit is performed using the signal PDF parameterization from the 2-lepton sample (Table 6.1) and a 1st-order polynomial background; the signal PDF parameters are fixed, the polynomial parameters are floated, and all yields are floated. The fit returns 4199 peaking and 617 non-peaking events; the non-peaking background is quite significant.

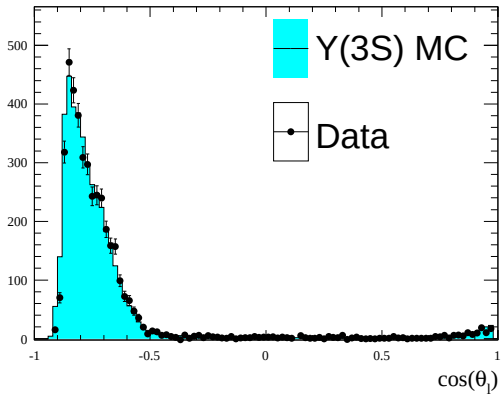
We see that the unfolded $\cos\theta$ distribution matches the re-scaled MC perfectly, except at the very back end of the detector; in the front end, it's unclear beyond the statistical uncertainty whether there is an effect. We concluded from the 2-lepton events that the front of the detector is well-modeled; given the statistics here, we have no strong reason to alter that assumption.



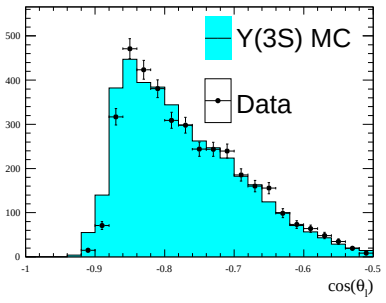
(a) Entire fit

(b) Zoomed view of the tail region

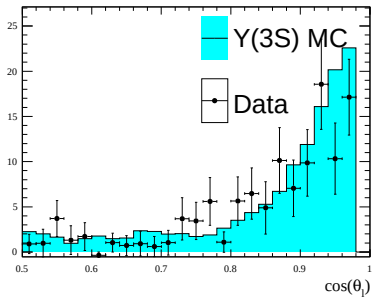
Figure 7.4: Fit for the peaking and non-peaking components of the 1-lepton recoil mass, showing (a) the whole fit and (b) a zoom of the region by the tails of the peak.



(a) Whole $\cos\theta$ region



(b) Backward region



(c) Forward region

Figure 7.5: Overlay of the unfolded 1-lepton sample $\cos\theta$ distribution in data (using ${}_s\mathcal{P}lots$) on the $\Upsilon(3S)$ MC for (a) the whole angular region, (b) the back end of the detector, and (c) the front end of the detector.

7.2.4.1 Applying a "Killing" Technique to 1-Lepton Events

We would like to account for the events in the Monte Carlo (MC) simulation that should have been classified as peaking background but were identified as single lepton events. We can determine this effect by applying a "killing" procedure to the MC simulation. We do this as follows:

- We bin the 1-lepton data and MC in $\cos\theta$ bins of size 0.02. We compute the ratio of data/MC in each bin.
- We apply the ratio in a random-number based lepton killing procedure as follows: For each bin, we define the killing factor as the data/MC ratio if it is between 0 and 1. If the ratio exceeds 1, we cannot add more MC events to that bin than already exist so we set the killing factor to 1.
- For each MC event, we determine in what bin the single lepton lies. If the killing factor for that bin is between 0 and 1, we throw a uniform random number for the lepton. If the value of that number exceeds the killing factor, we kill the lepton and treat the event as a 0-lepton event. Otherwise, we keep the event as a 1-lepton event.
- We only apply the killing procedure in the backward region defined earlier ($\cos\theta \in [-1.0, -0.84]$).

Applying this killing procedure and redefining the 1-lepton events leads to the $\cos\theta$ distributions in Figure 7.6, which are the post-killing analogues of Figure 7.5. We observe that the killing procedure appears to work very well; the MC now reproduces the data distribution in

the backward and forward regions of the detector. We then use the re-defined 0-lepton category to determine the peaking background prediction from this source, and we find 2470.1 predicted peaking leptonic background events.

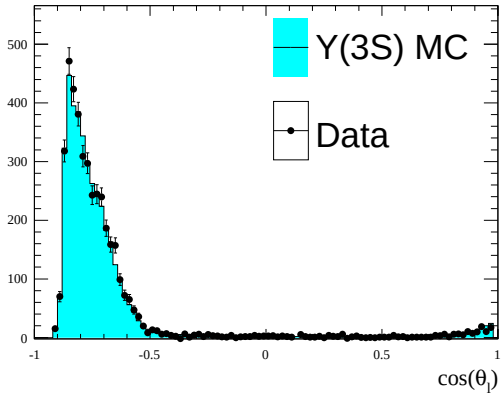
One more correction needs to be applied to the prediction from the killing procedure. This procedure cannot simulate the difference in trigger and selection efficiency for events that actually do not contain the reconstructed lepton. This effect could be quite significant, and we explore this possibility here.

We perform the following study using MC events. After all other cuts in the analysis, the signal selection described in Chapter 4.3.2.2 is 86% efficient on the 1-lepton events, and the trigger is 99% efficient. The total efficiency is therefore at least 85% when a lepton is present, in addition to the pions. However, once the lepton is lost the trigger and filter decisions are done entirely using the pions. The trigger is only 78% efficient on 0-lepton events that pass all other cuts, and the signal selection (Chapter 4.3.2.1) is 95% efficient. Therefore, once the lepton is lost the efficiency of selecting that event as a fake signal event is actually lower by $0.74/0.85 = 0.87$. This reduces the predicted peaking background from this killing procedure to $2301.1 + (2470.1 - 2301.1) \cdot 0.87 = 2448.1$.

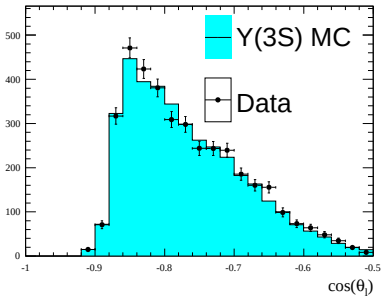
We take the difference between the unscaled and scaled numbers ($2470.1 - 2448.1 = 22.0$) as a systematic uncertainty on this procedure.

7.2.4.2 Uncertainty on the Leptonic Peaking Background Estimate

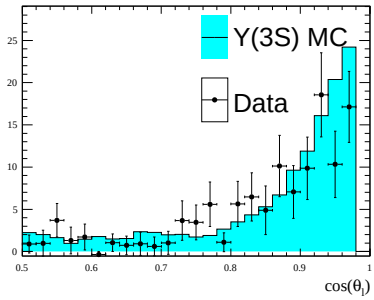
The uncertainty on the peaking background estimate comes from several sources. The first is MC statistics, and is negligible. The second has been discussed in the previous section,



(a) Whole $\cos\theta$ region



(b) Backward region



(c) Forward region

Figure 7.6: Overlay of the unfolded $\cos\theta$ distribution in data (using ${}_s\mathcal{P}lots$) on the $\Upsilon(3S)$ MC, after applying the killing procedure, for (a) the whole angular region, (b) the back-end of the detector, and (c) the front-end of the detector. The blue histogram is the sum of all $\Upsilon(1S)$ leptonic final state MC, after a dipion transition, and the magenta histogram contains all other $\Upsilon(3S)$ decays.

and is due to the differences in event selection efficiencies. Based on this, we assigned the difference between correcting and not correcting the background prediction for this effect as the systematic uncertainty due to this correction. This represents a 0.90% systematic uncertainty on the leptonic peaking background prediction.

There is also a systematic uncertainty on the number of events introduced by the ${}_s\mathcal{P}lots$ procedure This comes from the correlations between the angle of the single lepton and the recoil mass of the dipion system. We compare the number of events in the back of the detector in the pure MC sample and in the MC sample after being put through the same ${}_s\mathcal{W}eighting$ procedure as the 1 lepton data control sample (Figure 7.7) and find a total difference of $\sim$17 events. Since the data has a lower occupancy in these bins than the MC we scale this effect by the ratio of events in the data sample and those in the MC and find an uncertainty of 14 events. This gives an uncertaintainty of $14/2448.0 = 0.6\%$

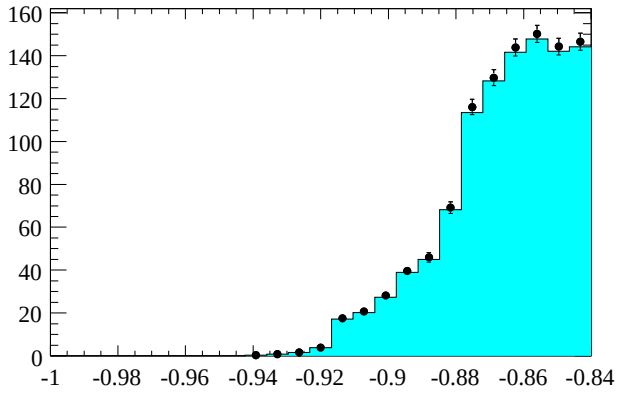


Figure 7.7: Plots of the one lepton Monte Carlo (MC) sample (Cyan Histogram) and the same MC sample after undergoing the same ${}_s\mathcal{W}eighting$ as the 1 lepton data control sample (black points).

Finally, we have the statistical uncertainty on the 1-lepton events. By scaling the data/MC ratios used for killing up or down by 1σ, we would reclassify more or less 1-lepton events as 0-lepton. We find that moving the killing factors up by 1σ yields 2440.8 events, and down by 1σ yields 2496.7. Taking half the difference between these two as the uncertainty yields 28.0 events. We then obtain a total uncertainty of $(22.0 + 28.0)/2448.0 = 2.0\%$

We add these effects in quadrature therefore take the total uncertainty on the leptonic peaking background to be 2.1%.

7.3 Studies of Non-leptonic Peaking Backgrounds

The Monte Carlo (MC) predictions for peaking background from $\Upsilon(1S)$ decays other than to dilepton states is very small. However there is no good reason to trust this prediction, so we need to use data to constrain such a contribution. We do this by comparing data to MC for "near" invisible $\Upsilon(1S)$ decays. In particular, we examine final states for which there are no extra tracks, but only one or more neutral calorimeter clusters, using variables such as total extra neutral energy or the center of mass (CM) energy of the highest-energy neutral.

We originally carried out this study on the Low + High data samples before unblinding the signal region. Thus we had to ensure that the range of values which could be studied for each variable lay outside the range accepted by the invisible-signal event selection. Events for the study were selected with a special skim, in which, based on figure 5.34, the highest-CM-energy photon in an event was required to be at least $0.25\,\mathrm{GeV}$. (Even with an MVA cut reduced to 0.5 from its normal value of 0.8, there are very few signal events predicted to have a larger

value.) Otherwise, the selection used the standard $\pi^+\pi^-$ selection. Finally, the cosine of the thrust angle was required to be below 0.85, to suppress an excess seen in sideband data but not in the generic $\Upsilon(3S)$ MC sample.

Fits to the recoil mass spectra for these pre-unblinding samples resulted in 217 ± 14 peaking events in the MC (after scaling by the nominal luminosity ratio of 0.35), and 260 ± 101 in the data, an encouraging similarity. We used ${}_s\mathcal{P}lots$ of maximum photon energy to show that in this near-invisible region the non-leptonic background for data is no more than five times the MC prediction, and we conservatively take this factor of five as the uncertainty for the non-leptonic peaking background in the invisible region.

However, rather than show these statistics-limited results here, we repeat the study post-unblinding using the full data sets. Figure 7.8 shows the recoil mass fits for generic MC (186.2 million $\Upsilon(3S)$ events) and all data (98.9 million, there being no need to exclude the Low+High data sets for this study), respectively. The relative luminosity of 0.5306 is nominal, not renormalized by the factor found from the two-lepton sample. The fits yield 3786 ± 100 peaking events in the MC (unscaled) and 3101 ± 308 in the data, for a data/scaled-MC ratio of 1.54 ± 0.16. Figure 7.9 compares ${}_s\mathcal{P}lots$ of this maximum photon energy for the MC and data fits, while figure 7.10 shows the data/MC ratio. The interesting region is toward the lowest energies, since that is closest to the region of interest for the invisible analysis. The plots do not show an increasing ratio as the energy decreases toward the inivisible region, and suggest that the limiting ratio is unlikely to exceed 4. Thus the factor-of-5 allowed for pre-unblinding is still safe as a systematic uncertainty.

Also post-unblinding, we have carried out a similar study for the total extra neutral

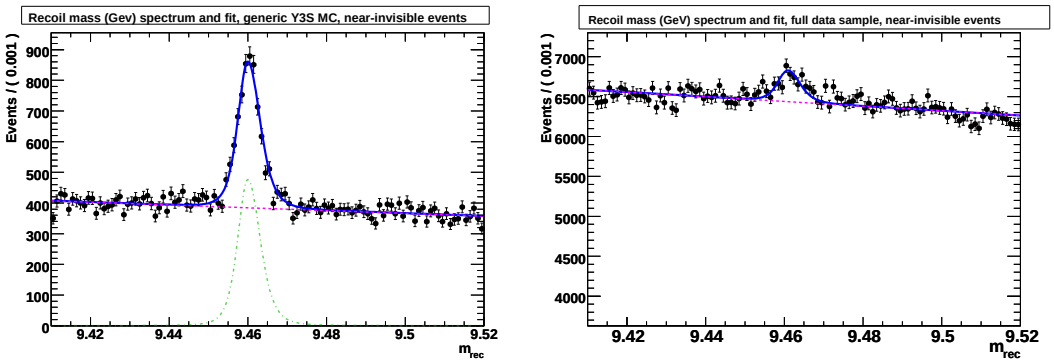


Figure 7.8: Fits of recoil mass spectra for a generic Monte Carlo (left) and all $\Upsilon(3S)$ data (right), for the "near-invisible" selection described in the text. The green dot-dashed curves show the peaking components of the fit (not visible on the zero-suppressed data plot), the red dashed curves show the linear non-peaking background, and the solid black curves show the total fits.

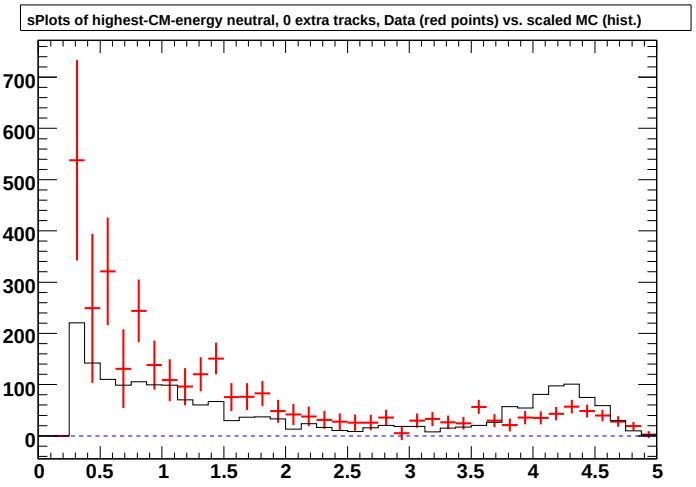


Figure 7.9: ${}_s\mathcal{P}lots$ of the maximum center of mass frame photon energy for the signal components of the fits shown in figure 7.8. The points with errors are from the full data sample, the histogram is the scaled generic Monte Carlo fit. The small MC uncertainties are not shown.

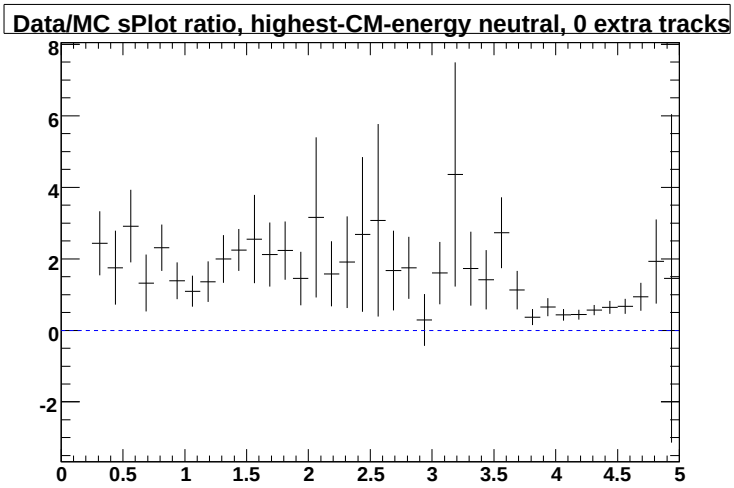


Figure 7.10: Ratio of ${}_s\mathcal{P}lots$ of the data to Monte Carlo simulation ${}_s\mathcal{P}lots$ shown in figure 7.9. Error bars include both data and MC uncertainties.

lab-frame energy, `eTot`. To avoid overlap with the first study, we required that the highest CM-frame photon energy be *below* $0.25\,\mathrm{GeV}$. The effect of the MVA cut is less sharp on `eTot`, with the efficiency tailing off between about 0.4 and $0.55\,\mathrm{GeV}$, but since we have already unblinded it is safe to look at the `eTot` spectrum down to $0.4\,\mathrm{GeV}$. Other cuts are as in the first "near" invisible selection. Figure 7.11 shows the recoil mass fits The fits yield 1112 ± 37 peaking events in the MC (unscaled) and 739 ± 208 in the data, for a data/scaled-MC ratio of 1.25 ± 0.36. Figure 7.12 compares ${}_s\mathcal{P}lots$ of this maximum photon energy for the MC and data fits, while figure 7.13 shows the data/MC ratio. It is most instructive to consider the first few bins above $0.6\,\mathrm{GeV}$, which are free of signal-like events. This study supports the upper limit for data of 4 to 5 times the MC prediction which we assigned based on the higher-statistics first study.

This study suggests that the MC prediction for the non-leptonic peaking background is appropriate, so we do not change that prediction of 2.8 events, and set the uncertainty on this

prediction at 14 events.

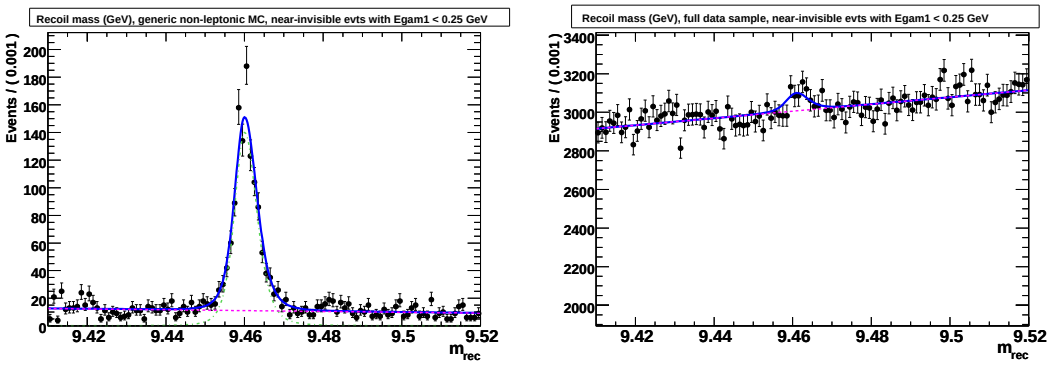


Figure 7.11: Fits of recoil mass spectra for a generic Monte Carlo (left) and all $\Upsilon(3S)$ data (right), for the second "near-invisible" selection (with maximum extra-photon center of mass frame energy below 0.25 GeV). The green dot-dashed curves show the peaking components of the fit (not visible on the zero-suppressed data plot), the red dashed curves show the linear non-peaking background, and the solid black curves show the total fits.

7.4 Final Peaking Background Estimate

Summing the prediction from leptonic and non-leptonic backgrounds, we obtain a final peaking background estimate of 2450.9 events, a 0.12% increase over the purely leptonic peaking background estimate. The systematic errors on this estimate, as well as all other systematics in the analysis are listed in Chapter 8

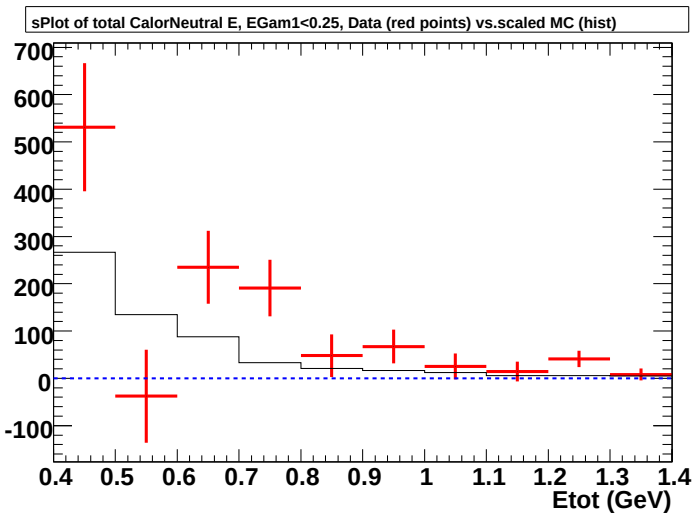


Figure 7.12: ${}_s\mathcal{P}lots$ of the maximum center of mass frame photon energy for the signal components of the fits shown in figure 7.11. The points with errors are from the full data sample, the histogram is the scaled generic Monte Carlo fit. The small MC uncertainties are not shown.

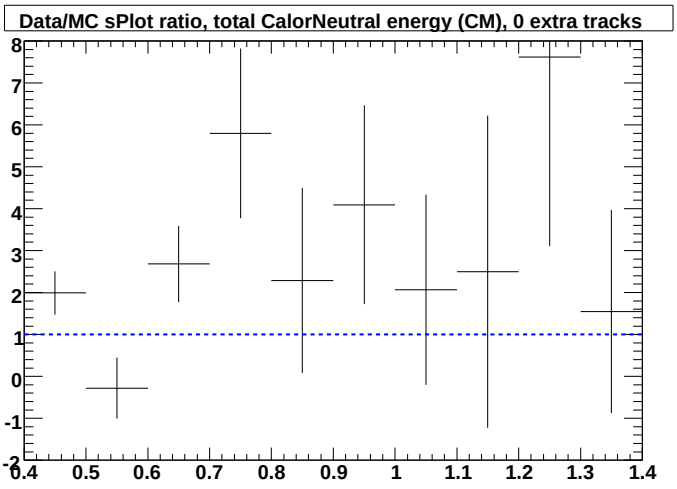


Figure 7.13: Ratio of ${}_s\mathcal{P}lots$ of the data to MC ${}_s\mathcal{P}lots$ shown in figure 7.12. Error bars include both data and MC uncertainties.

Chapter 8

Systematic Uncertainties

We describe the methods we use to assign systematic uncertainties to different parts of the analysis. We explain the way in which these uncertainties are combined to achieve a total systematic uncertainty on the signal yield and efficiency.

8.1 Scaling Correction

We discussed in Chapter 7.2.2 that we observe an overall normalization offset between MC and data in the two-lepton control sample. We take this correction (1.0882) as an overall correction to the normalization of the MC. Due to the statistics of the data used in determining this correction, there is a 1.1% uncertainty on this scaling factor. This scaling factor was determined after all selection criteria were applied to the two-lepton events.

We consider this as a measure of many effects at once, including:

- The branching fraction uncertainty and central value for $\Upsilon(3S) \to \pi^+\pi^-\Upsilon(1S)$, which

is reported by the Particle Data Group (PDG) [30] to be 4.48% with a 5% uncertainty.

- The branching fraction uncertainty on $\Upsilon(1S) \rightarrow \ell^+\ell^-$, where $\ell = e,\ \mu,\ \tau$, which the PDG reports as having a 5%/2%/4% uncertainty on the electron/muon/tau channel
- The tracking efficiency of the dipions, independent of trigger effects
- The tracking efficiency of the lepton tracks

Given the fact that systematic effects not applicable to the invisible final state (lepton tracking, leptonic branching fraction) contribute to in this correction, it is certainly conservative to apply this correction to the invisible final state (and therefore its efficiency).

8.2 Trigger Studies

BABAR triggering is discussed in Chapter 8.2.1. We consider both the hardware (Level 1) and software (Level 3) triggering systematic errors. The Level 1 triggering systematics use the 2 lepton control sample as an analog for signal events before triggering, and simulate the triggering in software. For the Level 3 triggering systematics we use a small sample of events which are passed through the Level 3 trigger expressly for this purpose.

8.2.1 Level 1 Trigger Systematics

We wish to compare the efficiencies of the hardware trigger on the Monte Carlo (MC) signal simulation and the invisible data. The objects that trigger the detector in these events are the two pions from the dipion transition $\Upsilon(3S) \rightarrow \pi^+\pi^-\Upsilon(1S)$. In the two lepton control

sample the two hard tracks from the $\Upsilon(1S) \rightarrow \ell^+\ell^-$ decay are responsible for triggering the detector, and so represent a sample of pions without trigger bias. We construct an offline algorithm to mimic the trigger and apply it to the two lepton sample in both data and MC simulation. The efficiency of this algorithm is used as the Level 1 trigger systematic.

We can use the final signal fit (discussed in Chapter 9.1) to create an ${}_s\mathcal{P}lot$ of the trigger lines fired by peaking events in data, and compare that to trigger lines fired in the peaking background MC. This is shown in Fig. 8.1. We find generally good agreement across the lines, with some cases where the data and the MC appear to disagree outside of the errors on each by several sigma.

We find that the trigger lines are highly correlated with one another and that the two-track triggers dominate the efficiency. We therefore concentrate on the study of just the two-track trigger lines. These lines are defined as follows:

- 2Zt & 1A & 1M
- 2A & 1Zk & 1M
- D2 & 1Zk & 1M
- D2 & 2M & 1Zt
- M*&1Z

The trigger objects involved are defined in Table 8.1.

There are a few steps that need to be taken to establish the effect of these trigger lines on invisible data:

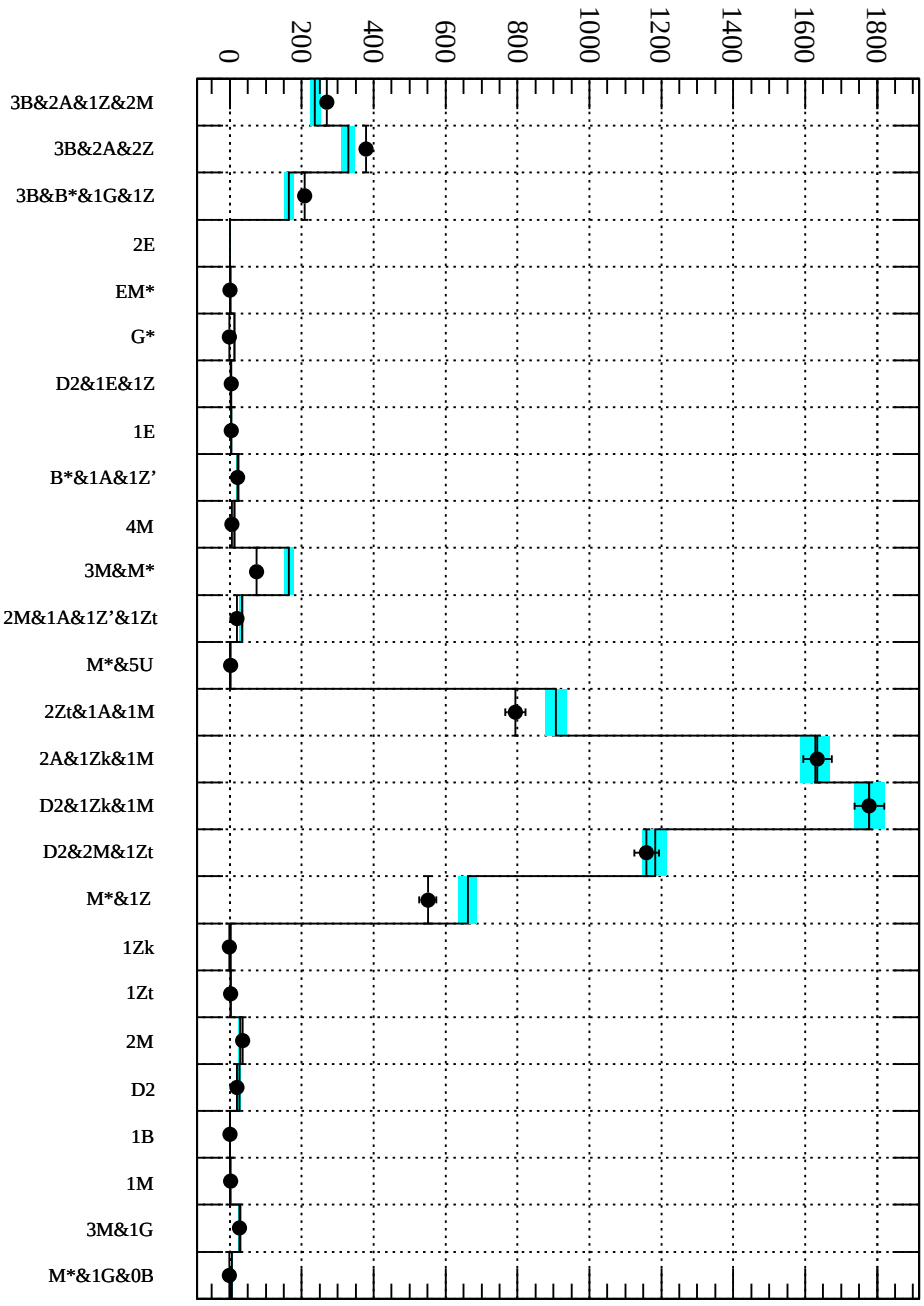


Figure 8.1: The number of events passing each Level 1 trigger line (inclusively) in the peaking component of data (black points with errors made with the ${}_s\mathcal{P}lot$ procedure) and in the peaking background MC (white with cyan error boxes).

Trigger Object	Cuts
A	a track reaching DCH superlayer (SL) 10, with $p_T > 0.150\,\mathrm{GeV}/c$
B	a track reaching SL5 with $p_T > 0.120\,\mathrm{GeV}/c$
D2	2B & 1A
Z	A track with $p_T > 0.2\,\mathrm{GeV}/c$ and $\lvert z\rvert < 12\,\mathrm{cm}$
Zt	a track reaching SL7, with $p_T > 0.2\,\mathrm{GeV}/c$ and $\lvert z\rvert < 10\,\mathrm{cm}$
Zk	a track reaching SL7, with a charge-asymmetric p_T cut. If positive, the track must satisfy $p_T > 0.8\,\mathrm{GeV}/c$ and if negative, the track must satisfy $p_T > 0.2\,\mathrm{GeV}/c$. In addition $\lvert z\rvert < 12\,\mathrm{cm}$
M	an EMC cluster with at least $0.13\,\mathrm{GeV}$ of energy
M*	two back-to-back M clusters with at least an opening angle of 117 degrees.

Table 8.1: Definitions of trigger objects using cuts on the off-line reconstructed quantities.

- The 2-lepton sample contains dipions that, in principle, are not affected significantly by the two-track triggers since they have likely been fired by the presence of the two high-momentum leptons, or the combination of one of the pions and one of the leptons. We need to test this assumption - that is, that the pion kinematics represent those in an "untriggered" signal sample

- Having established that the pions in the 2-lepton sample are analogous to untriggered pions in the signal MC, we then proceed to dissect the trigger lines and reproduce their effects using offline cuts (that is, cuts based on fully reconstructed objects instead of the more coarse trigger objects).

- After creating an algorithm that largely reproduces the effect of trigger cuts on signal using control sample MC, we apply those cuts to the data in the control sample and

measure the efficiency difference.

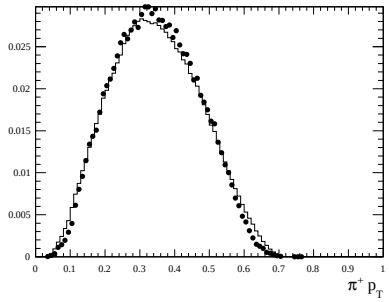
8.2.1.1 Pion Kinematics in the Two Lepton Control Sample

We begin by investigating whether the pions in the two lepton control sample Monte Carlo (MC) are a good analog of the pions in the untriggered signal MC events. We apply all cuts to the signal MC except the Level 1, Level 3, and invisible sample selection cuts. We apply to the control sample MC all cuts originally used to select those events, including trigger cuts. Again, the premise is that the trigger doesn't shape the pions in this sample like it does in the signal events. In addition, we have to apply all cuts used to define the 2-lepton sample if we are to reproduce that sample in data, where a trigger is already applied by definition.

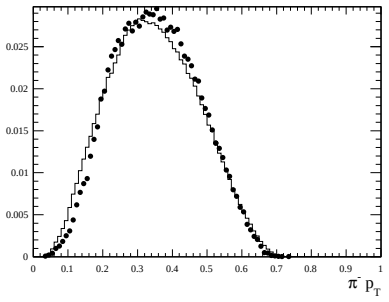
We compare the individual pion p_T and $\cos\theta$ distributions, along with the opening angle of the negative pion with respect to the positive pion, in Figures 8.2-8.4. These plots are all unit-normalized, so that the shape alone can be compared. We see small differences in the distributions but in general we find that the control sample MC is a very good analog of signal MC, prior to trigger cuts.

8.2.1.2 "Level 1 Trigger Inspired" Cuts

Before trying to reproduce the trigger lines with offline cuts, the basic ability to map drift chamber trigger (DCT) quantities (most likely to affect the efficiency) onto offline reconstruction quantities (e.g. pion p_T, dipion opening angle, radial extent of pion track) needs to be established. The most basic issue is whether the p_T cutoff is the same in the reconstructed object (track) as in the trigger object. For instance, signal Monte Carlo (MC) events passing a line

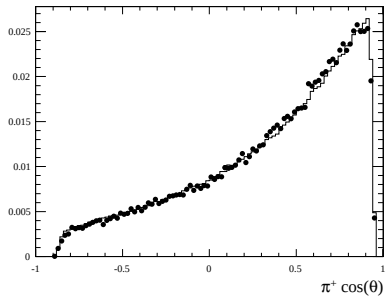


(a) $\pi^+ \, p_T$

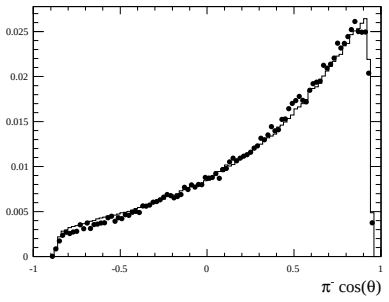


(b) $\pi^- \, p_T$

Figure 8.2: Individual reconstructed pion p_T distributions for signal MC (black markers, no errors shown) and the 2-lepton control sample MC (black histogram). Both distributions are unit normalized.



(a) $\pi^+ \cos(\theta)$



(b) $\pi^- \cos(\theta)$

Figure 8.3: Individual reconstructed pion $\cos\theta$ distributions for signal MC (black markers, no errors shown) and the 2-lepton control sample MC (black histogram). Both distributions are unit normalized.

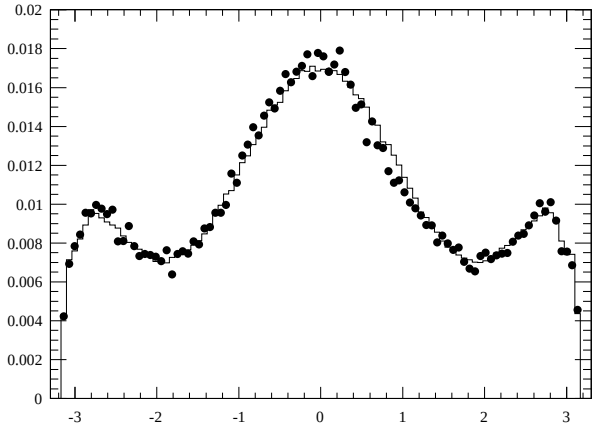


Figure 8.4: The dipion opening angle, relative to the π^+, for signal MC (black markers, no errors shown) and the 2-lepton control sample MC (black histogram). Both distributions are unit normalized.

containing 2Zt is expected to have neither track with $p_T < 0.2\,\mathrm{GeV}/c$. We find that fewer than 1% of signal events have a track with p_T below the expected threshold, and the distributions are suppressed rapidly below the cut threshold. While the cutoff is not perfectly sharp, it is very close to expectation.

We apply cuts to reproduce each line, treating the tracks like DCT objects. A few comments are required on our choices here. We use the p_T cuts outlined in the DCT definitions. We use the outermost drift chamber (DCH) layer hit by a reconstructed track to determine whether a track reaches the needed DCH superlayer. Since the tracks are GoodTracksVeryLoost (GTVL) (defined in Chapter 4.3.1), the $|z|$ cuts are already applied. We use the electromagnetic calorimeter (EMC) clusters associated with the tracks, and the EMC cluster associated with the highest energy photon, to define the M objects. The coarseness of the DCT (illustrated in

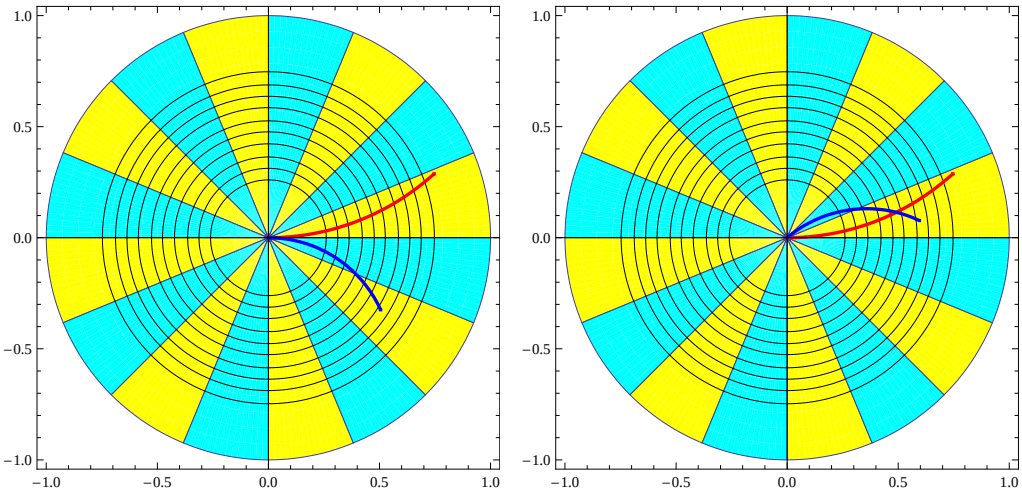
Figure 8.5) is the hardest effect to reproduce.

To reproduce the DCT coarseness, we cut on the opening angle between the pions in the plane transverse to the beam axis. There is a subtlety in defining the opening angle cut; it is not as simple as requiring that the opening angle be greater than the minimum ϕ bin width for a given object. Tracks bending in the magnetic field can either curve quickly away from one another, even for a small opening angle, or cross over each other, even for a larger opening angle. This type of effect can be seen in Figure 8.5, where we illustrate some possible dipion configurations. We rely on the signal MC to guide the cut on this opening angle. The opening angle distributions for each of the five trigger lines in question is shown in Figure 8.6. Figures 8.7–8.11 show the efficiency of each trigger line as a function of the opening angle. These plots make it much easier to observe where the effective veto cut on the opening angle needs to be placed.

The cuts that we will apply to the opening angle will be hard cuts, but it is clear that the cuts are by no means hard in the way they affect the reconstructed signal events. Therefore, we do not expect our cuts to reproduce the efficiency of the trigger in the control MC. We are primarily interested in how the cuts affect the shape of the dipion kinematic variables as they are presented to the Random Forest Algorithm (MVA).

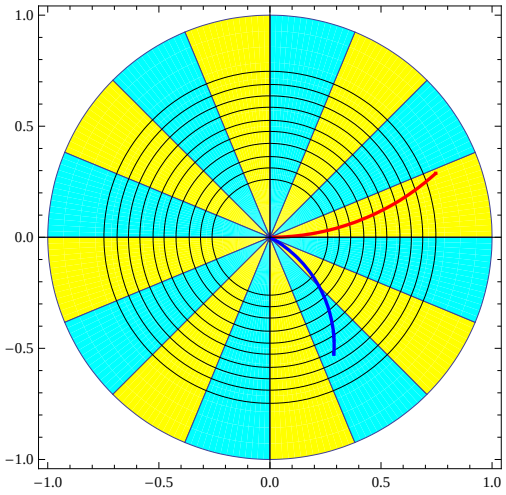
The opening angle cuts were determined by setting the cut boundaries to a point halfway down the slope of the efficiency on either side of the minimum in the efficiency plots (Figures 8.7–8.11). The regions we veto to approximate the effect of the trigger are detailed in Table 8.2.

The effect of trigger cuts is most pronounced on the dipion p_T, so we focus on that



(a) Angular Separation of 0.0 radians

(b) Angular Separation of 0.7 radians



(c) Angular Separation of -0.5 radians

Figure 8.5: Diagram of three different dipion system configurations. In each case the π^+ is in red and has $p_T = 0.5\,\mathrm{GeV}/c$ and the π^- is in blue and has $p_T = 0.25\,\mathrm{GeV}/c$. The angular separation in (a) is 0.0 radians, in (b) is 0.7 radians, and in (c) is -0.5 radians. The detector has been divided into 16 sections to illustrate the coarseness of the Level 1 Trigger, and the concentric black circles indicate the radius of the innermost sense wire in each superlayer.

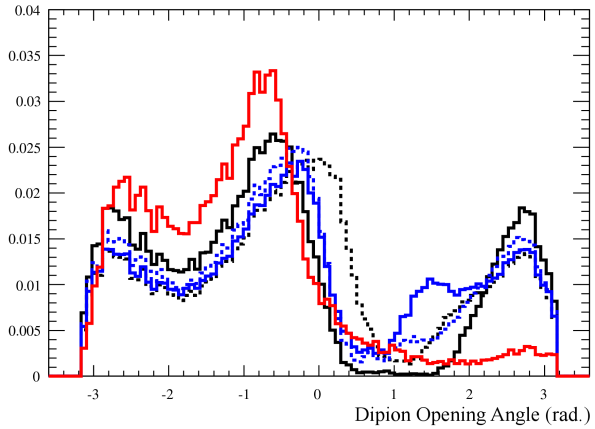


Figure 8.6: The dipion opening angle, relative to the π^+, for the application of different two-track triggers to the signal MC. All distributions are unit normalized. The distributions correspond to 2Zt & 1A & 1M (black, solid), 2A & 1Zk & 1M (black, dotted), D2 & 1Zk & 1M (blue, solid), D2 & 2M & 1Zt (blue, dotted), M* & 1Z (red, solid).

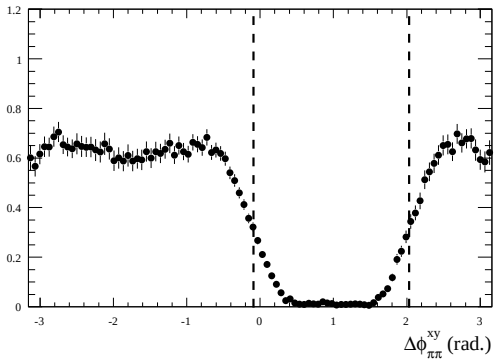


Figure 8.7: The efficiency of the 2Zt & 1A & 1M as a function of the dipion opening angle in signal MC. All cuts except the MVA cut, and cuts on other trigger lines, are already made prior to applying this trigger line cut. The dashed lines indicate where we define veto regions.

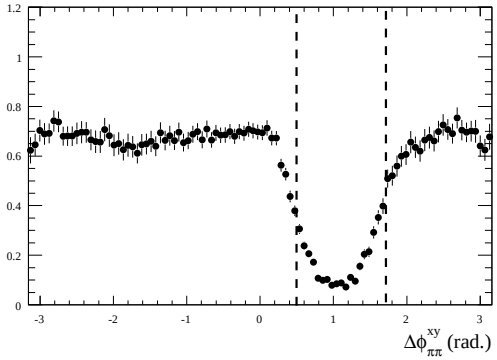


Figure 8.8: The efficiency of the 2A & 1Zk & 1M as a function of the dipion opening angle in signal MC. All cuts except the MVA cut, and cuts on other trigger lines, are already made prior to applying this trigger line cut. The dashed lines indicate where we define veto regions.

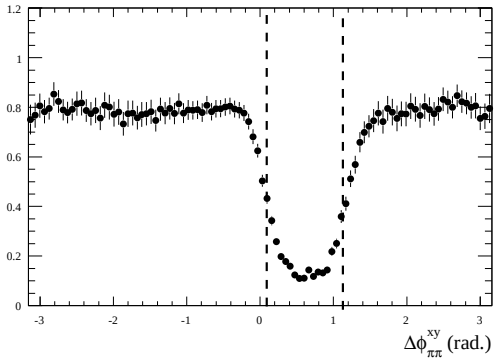


Figure 8.9: The efficiency of the D2 & 1Zk & 1M as a function of the dipion opening angle in signal MC. All cuts except the MVA cut, and cuts on other trigger lines, are already made prior to applying this trigger line cut. The dashed lines indicate where we define veto regions.

Trigger Object	Veto Region(s) in $\Delta\phi^{xy}_{\pi\pi}$
2Zt & 1A & 1M	$[-0.1, 2.0]$
2A & 1Zk & 1M	$[0.5, 1.7]$
D2 & 1Zk & 1M	$[0.1, 1.1]$
D2 & 2M & 1Zt	$[0.05, 1.7]$
M* & 1Z	$[-0.4, \pi]$ or $[-\pi, -2.8]$

Table 8.2: Opening angle veto regions for each two-track trigger line.

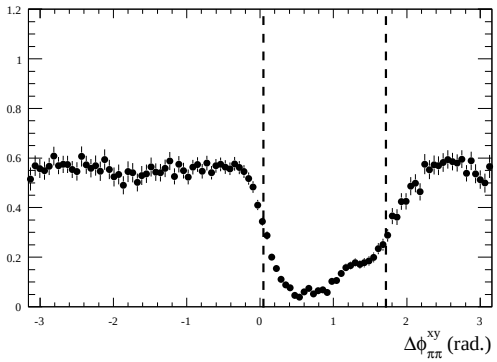


Figure 8.10: The efficiency of the D2 & 2M &1Zt as a function of the dipion opening angle in signal MC. All cuts except the MVA cut, and cuts on other trigger lines, are already made prior to applying this trigger line cut. The dashed lines indicate where we define veto regions.

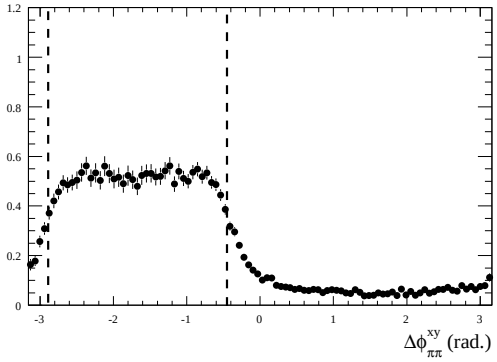


Figure 8.11: The efficiency of the M* & 1Z as a function of the dipion opening angle in signal MC. All cuts except the MVA cut, and cuts on other trigger lines, are already made prior to applying this trigger line cut. The dashed lines indicate where we define veto regions.

variable in our comparison. Figures 8.12 - 8.16 demonstrate the difference between not applying and applying these cuts to the control MC, compared to signal MC where the corresponding real trigger cut has been applied. Figure 8.17 shows the cumulative effect of applying all of these lines to the two lepton control sample MC events.

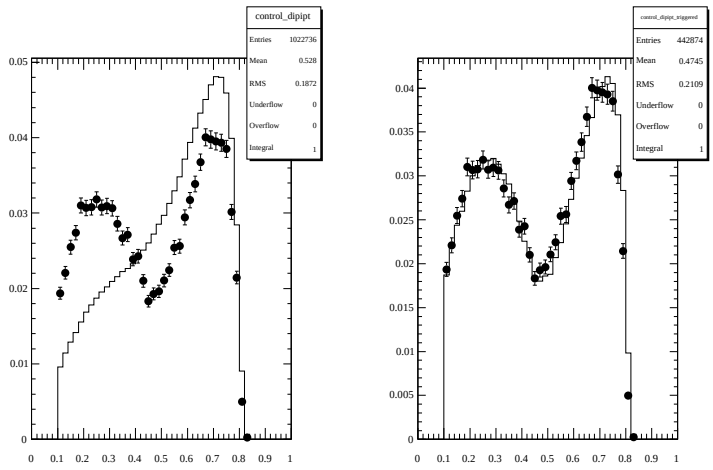


Figure 8.12: The dipion p_T for control MC (solid histogram) and signal MC (dot markers with errors). The left plot shows control MC before making the effective trigger cut (2Zt&1A&1M) and the right plot is after making the effective offline cut. The signal MC in both has all reconstruction cuts applied, and has the trigger line cut applied as well. All distributions are unit normalized.

We find the the cuts largely reproduce the shape of the dipion p_T. In addition, we look at two other dipion-kinematics-related variables used in the MVA - the dipion $\cos\theta$ and the cosine of the angle between the highest-energy photon and the normal to the dipion decay plane. These variables show significant differences between the control MC and signal MC until the application of our trigger cuts (Fig. 8.18). After the trigger cuts the agreement is excellent. The variable with the largest remaining discrepancy in shape between control and signal MC

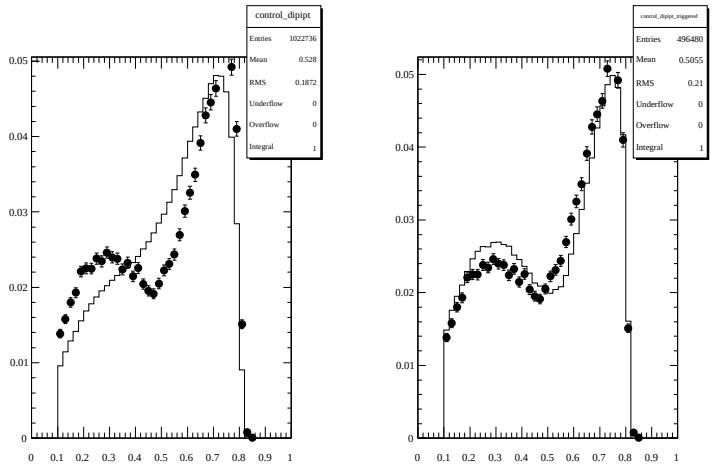


Figure 8.13: The dipion p_T for control MC (solid histogram) and signal MC (dot markers with errors). The left plot shows control MC before making the effective trigger cut (2A&1Zk&1M) and the right plot is after making the effective offline cut. The signal MC in both has all reconstruction cuts applied, and has the trigger line cut applied as well. All distributions are unit normalized.

is the dipion p_T, although this shape difference has been greatly reduced. We proceed with the selection and determine the net effect on MC-vs-MC agreement in the MVA output. We also apply the same "level 1 inspired" pseudo-trigger cuts to the signal MC before the real trigger cuts and observe the agreement between the signal and 2-lepton control MC samples under the same cuts.

We quantify the differences between signal and control MC when applying these cuts, and control data and MC, in Table 8.3. We find that the difference between the signal MC and control MC efficiencies is large, as expected - we did not anticipate that we could exactly reproduce efficiency numbers in the MC (the opening angle cut is clearly not a hard cut, as we modeled it). However, we find that application of the pseudo-trigger cuts to signal MC creates

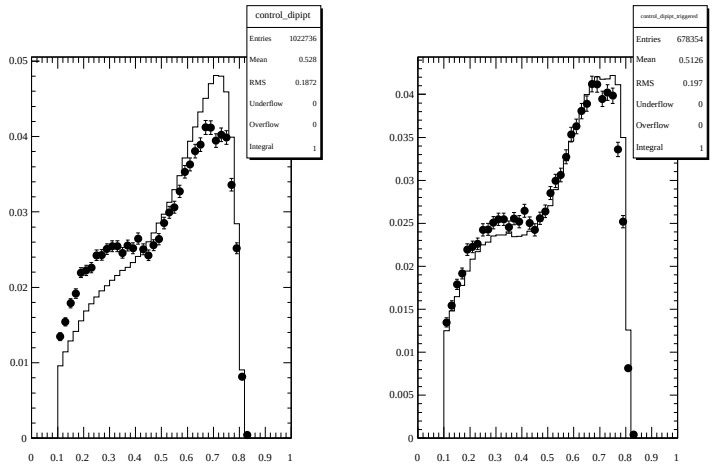


Figure 8.14: The dipion p_T for control MC (solid histogram) and signal MC (dot markers with errors). The left plot shows control MC before making the effective trigger cut (D2 & 1Zk & 1M) and the right plot is after making the effective offline cut. The signal MC in both has all reconstruction cuts applied, and has the trigger line cut applied as well. All distributions are unit normalized.

good agreement between the signal MC and the 2-lepton control sample MC. We also find that for those cuts the 2-lepton control MC and data agree very well. Relative to the control MC efficiency, the difference in the control data and MC efficiencies is a 2.1% effect. We take that as the systematic on the application of Level 1 trigger cuts to invisible signal.

Sample	**L1 Trigger Efficiency**
Signal MC (Trigger Simulation)	76.46%
Signal MC (Pseudo-Trigger Cuts)	70.58%
Control MC (Pseudo-Trigger Cuts)	68.13%
Control Data (Pseudo-Trigger Cuts)	66.69%

Table 8.3: Efficiency of Level 1 Trigger Cuts

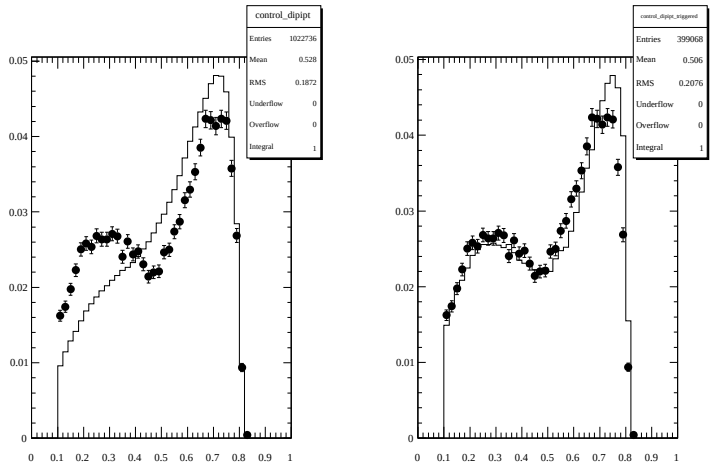


Figure 8.15: The dipion p_T for control MC (solid histogram) and signal MC (dot markers with errors). The left plot shows control MC before making the effective trigger cut (D2 & 2M & 1Zt) and the right plot is after making the effective offline cut. The signal MC in both has all reconstruction cuts applied, and has the trigger line cut applied as well. All distributions are unit normalized.

8.2.2 Level 3 Trigger Systematics

We study the ratio of Level 3 trigger efficiencies in data and Monte Carlo (MC) using events which pass Level 1, but are passed through Level 3 (`L3OutL1Open`). In other words, these events are not required to pass the software trigger, but are marked as passing if they do. These events are prescaled by a factor of 200 so as not to overwhelm us with statistics. We select events in data which pass all cuts, except the Level 3 trigger and in addition are part of this pass-through. We then apply the Level 3 trigger to these events, and compute the ratio of those which pass the Level 3 trigger to those which do not. In signal Monte Carlo, we perform a similar study, but are not limited by the pass-through statistics. Finally, we compute the ratio of data to MC and take that as a systematic correction with an uncertainty from the statistics of

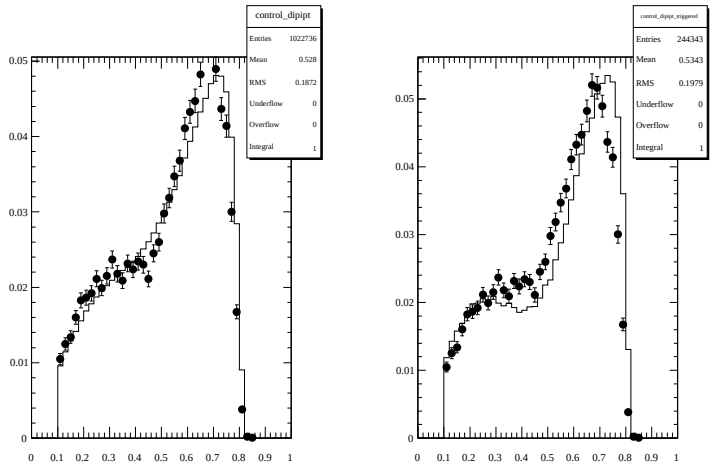


Figure 8.16: The dipion p_T for control MC (solid histogram) and signal MC (dot markers with errors). The left plot shows control MC before making the effective trigger cut (M* & 1Z) and the right plot is after making the effective offline cut. The signal MC in both has all reconstruction cuts applied, and has the trigger line cut applied as well. All distributions are unit normalized.

each sample of pass-through events.

The results are summarized in Table 8.4. We find the efficiencies in data and MC to be very similar, so that the data/MC ratio is essentially 1.0. The uncertainty on the ratio is dominated by the data statistics of events passing through Level 3.

8.2.3 Random Forest Systematics

The systematic uncertainty on the cut efficiency of the Random Forest MVA is potentially entangled with any issues that arise prior to the MVA. We find that it is essential to study the MVA output after applying Level 1 trigger inspired cuts to our 2-lepton events. The reason for needing to do this is two-fold: (1) the MVA was trained on sideband and signal MC

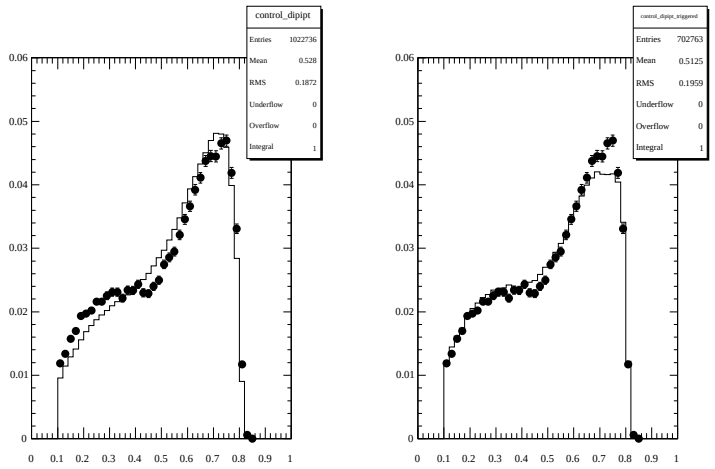
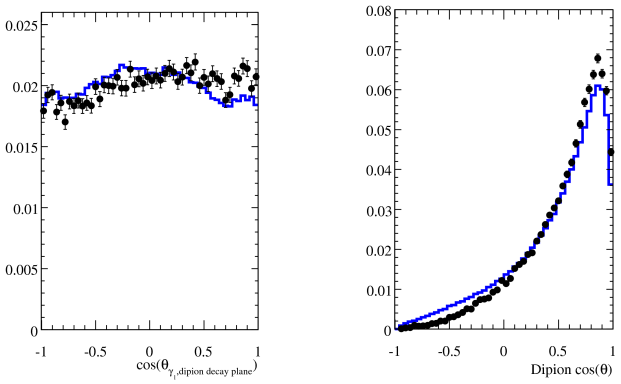


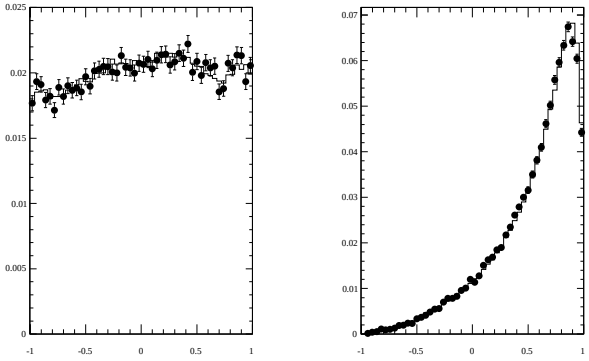
Figure 8.17: The dipion p_T for control MC (solid histogram) and signal MC (dot markers with errors). The left plot shows control MC before making the effective two-track trigger cuts and the right plot is after making the effective offline cuts. The signal MC in both has all reconstruction cuts applied, and has the trigger line cut applied as well. All distributions are unit normalized.

Selection	Data	Signal MC
All Cuts (except Level 3) and `L3OutL1Open`	324	57501
All Cuts and `L3OutL1Open`	316	56268
Efficiency	$(97.53 \pm 0.86)\%$	$(97.86 \pm 0.06)\%$
data/MC	0.9966 ± 0.0088	

Table 8.4: Results of the study of events passing `L3OutL1Open` in data and signal Monte Carlo



(a) $cos(\theta_{\gamma_{HE},\vec{p}_\pi \times \vec{p}_\pi})$ (b) dipion $\cos(\theta)$



(c) After Level 1 trigger cuts

Figure 8.18: The cosine of the angle between the highest-energy photon and the normal to the dipion decay plane and the dipion $\cos\theta$ in control MC (histogram) and signal MC (points with errors). The top pair of plots is shown for control MC before Level 1 two-track-trigger cuts and for signal with all cuts, including trigger cuts, applied. The bottom pair shows the same samples after Level 1 two-track triggers. All distributions are unit normalized.

events after trigger cuts, and those cuts cannot be removed from sideband data, and (2) the Level 1 trigger significantly shapes the inputs before they enter the MVA. It is critical to reproduce those shapes in the control sample before studying the MVA output.

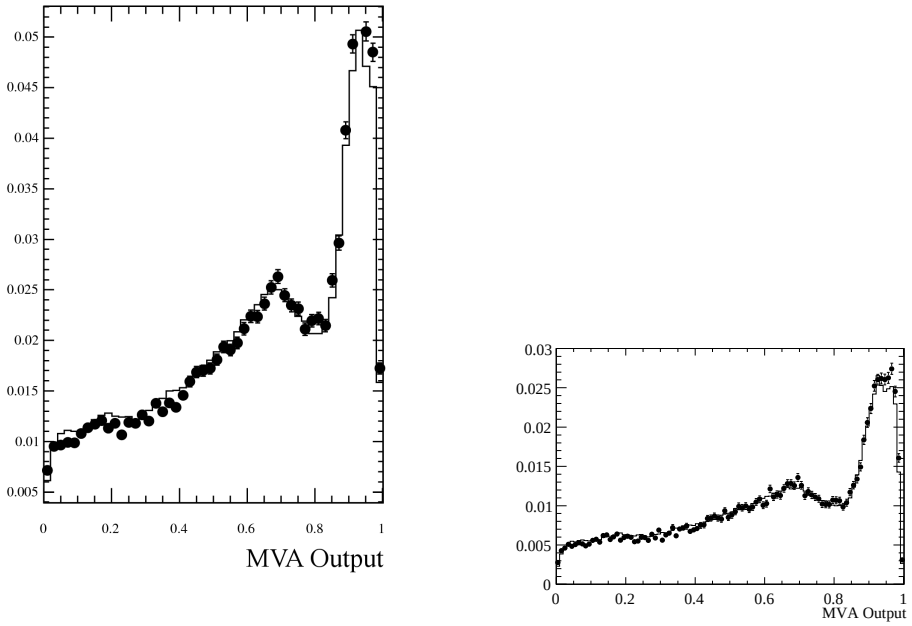
The distribution of the MVA output after trigger cuts is shown in Figure 8.19 for control MC and signal MC. We see that there are residual shape differences when the real trigger simulation is applied to the signal MC (Figure 8.19(a)). Those disagreements largely go away if we instead compare signal MC and 2-lepton control MC, both with pseudo-trigger cuts applied (Figure 8.19(b)). We also plot the distribution for data and MC in the 2-lepton control sample (Figure 8.20), and see small differences in shape there are well. We then apply our standard $MVA > 0.8$ cut and compare, quantitatively, the efficiency of making that cut. We find (Table 8.5) that the MC-to-MC difference is at the level of several percent. We find the data-to-MC difference is 4.0% relative to the control MC.

We move the cut on the MVA output in the 2-lepton control MC to achieve (almost) the same efficiency as in signal MC at the chosen cut of $MVA > 0.8$. A cut at $MVA > 0.787$ replicates this efficiency. For this cut (Table 8.5) we find the data/MC difference is slightly reduced to the level of 3.8%.

We use the data-MC difference for the nominal cut of $MVA > 0.8$ as the systematic uncertainty on the MVA.

8.2.4 Aside – ${}_s\mathcal{P}lots$ of MVA inputs

As a cross-check of our modeling of the data using the peaking background MC samples, which clearly dominate the peak in data, we use the ${}_s\mathcal{P}lot$ technique [13] to unfold



(a) Signal MC (Real Trigger Simulation Applied) vs. Control MC

(b) Signal MC (Pseudo-trigger cuts) vs. Control MC

Figure 8.19: The MVA output for 2-lepton control sample Monte Carlo (histogram) and signal sample Monte Carlo (points). Both distributions are unit-normalized.

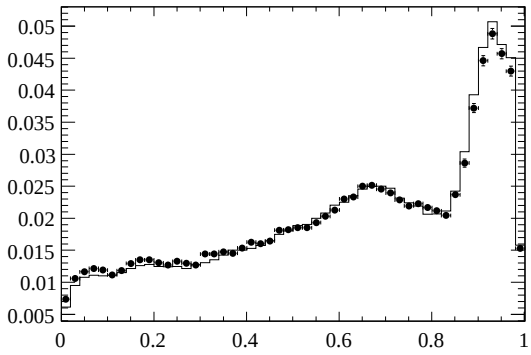


Figure 8.20: The MVA output for 2-lepton control sample Monte Carlo (histogram) and 2-lepton control sample data (points). Both distributions are unit-normalized.

Sample	$MVA > 0.8$ **Efficiency**
Signal MC (Trigger Simulation Applied)	35.91%
Signal MC (Pseudo-trigger Cuts Applied)	35.64%
Control MC (Pseudo-Trigger Cuts Applied)	34.50%
Control Data (Pseudo-Trigger Cuts Applied)	33.12%

Sample	**MVA Cut Achieving Signal MC Efficiency**
Signal MC (Trigger Simulation Applied)	35.91%
Signal MC (Pseudo-trigger Cuts Applied)	35.64%
Control MC (Pseudo-Trigger Cuts Applied)	35.91% ($MVA > 0.787$)
Control Data (Pseudo-Trigger Cuts Applied)	34.55% ($MVA > 0.787$)

Table 8.5: Efficiency of MVA cuts, both nominal (> 0.8) and when forcing the 2-lepton control MC to have the same efficiency as the signal MC.

the peaking component in variables other than the recoil mass, using our Maximum Likelihood fit to the recoil mass from Chapter **??**. The resulting plots are shown below, with the projected peaking component of data overlaid on the sum of peaking background MC samples. The most significant shape difference we observe is in the dipion $\cos(\theta)$, which appears more peaked in the forward direction in data than in MC. This is due to the correlations between the fitted variable m_{recoil} and the dipion $\cos(\theta)$, and is explained in detail in Appendix B.8

8.3 Peaking Background Systematic Uncertainties

Our prediction of the peaking background comes from leptonic and non-leptonic contributions. We discuss these contributions in Chapter 7. The systematic uncertainty on the peaking background prediction is found to be 2.1%.

The systematic uncertainty on the leptonic peaking background prediction discussed in Chapter 7.2.4.2 has several sources. We found the the total systematic uncertainty to be 2.0%.

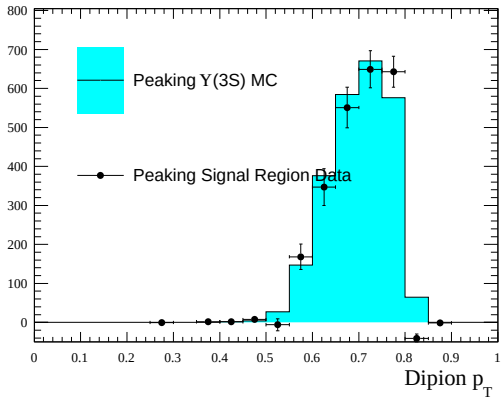


Figure 8.21: The dipion system p_T, compared between peaking background MC (cyan histogram) and the peaking component of the data projected from the recoil mass fit. MC is normalized to the $\Upsilon(3S)$ count, with corrections applied.

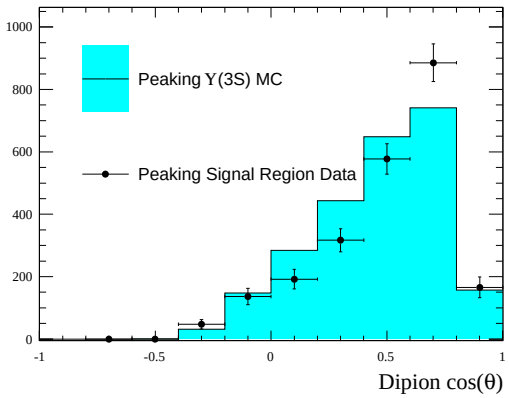


Figure 8.22: The dipion system $\cos(\theta)$, compared between peaking background MC (cyan histogram) and the peaking component of the data projected from the recoil mass fit. MC is normalized to the $\Upsilon(3S)$ count, with corrections applied.

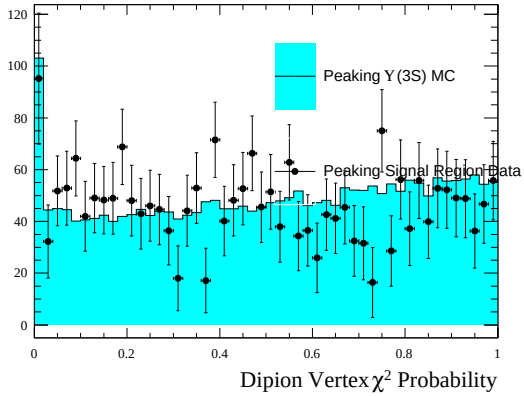


Figure 8.23: The dipion system vertex fit χ^2 probability, compared between peaking background MC (cyan histogram) and the peaking component of the data projected from the recoil mass fit. MC is normalized to the $\Upsilon(3S)$ count, with corrections applied.

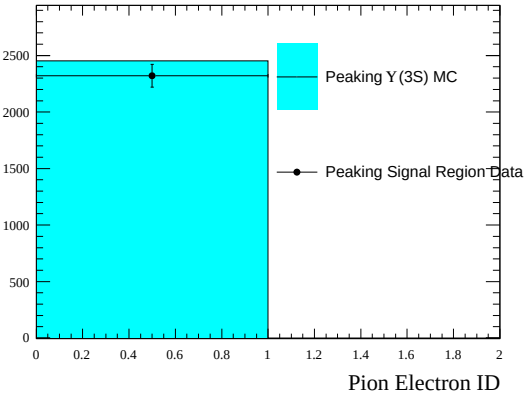


Figure 8.24: Whether either pion passes eKMVeryLoose, compared between peaking background MC (cyan histogram) and the peaking component of the data projected from the recoil mass fit. MC is normalized to the $\Upsilon(3S)$ count, with corrections applied.

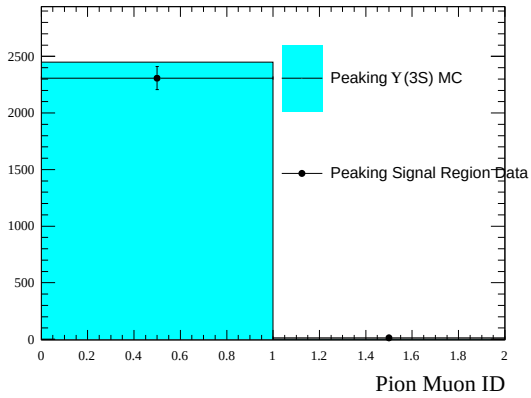


Figure 8.25: Whether either pion passes muBDTVeryLoose, compared between peaking background MC (cyan histogram) and the peaking component of the data projected from the recoil mass fit. MC is normalized to the $\Upsilon(3S)$ count, with corrections applied.

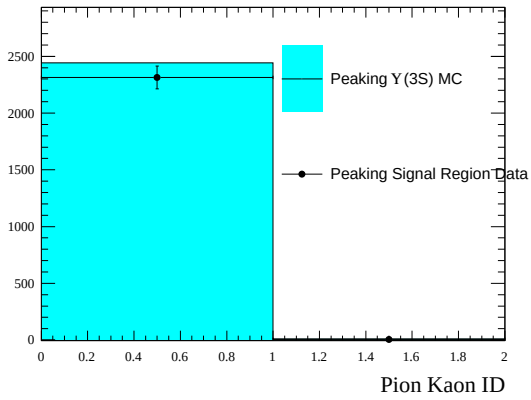


Figure 8.26: Whether either pion passes KKMVeryLoose, compared between peaking background MC (cyan histogram) and the peaking component of the data projected from the recoil mass fit. MC is normalized to the $\Upsilon(3S)$ count, with corrections applied.

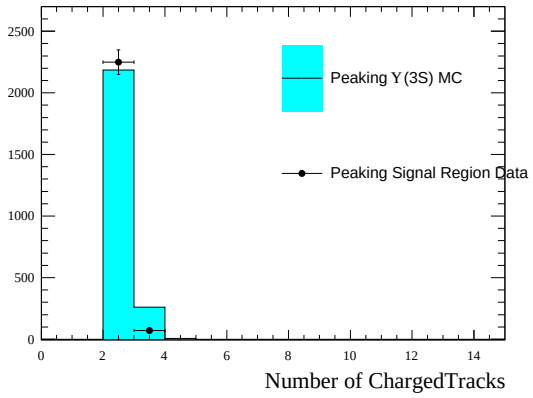


Figure 8.27: ChargedTrack multiplicity, compared between peaking background MC (cyan histogram) and the peaking component of the data projected from the recoil mass fit. MC is normalized to the $\Upsilon(3S)$ count, with corrections applied.

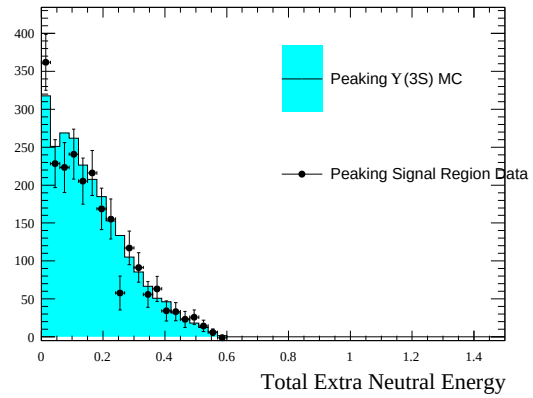


Figure 8.28: Total extra neutral energy (lab frame), compared between peaking background MC (cyan histogram) and the peaking component of the data projected from the recoil mass fit. MC is normalized to the $\Upsilon(3S)$ count, with corrections applied.

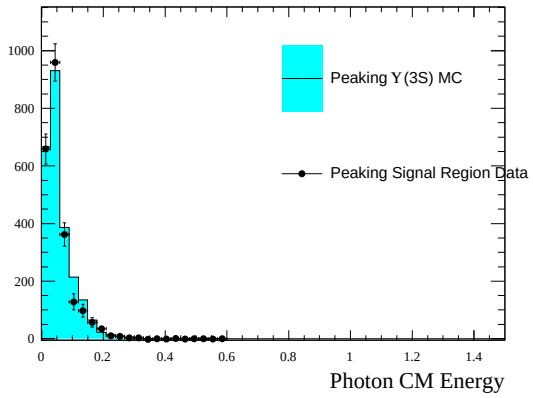


Figure 8.29: Energy of the highest-energy photon (CM frame), compared between peaking background MC (cyan histogram) and the peaking component of the data projected from the recoil mass fit. MC is normalized to the $\Upsilon(3S)$ count, with corrections applied.

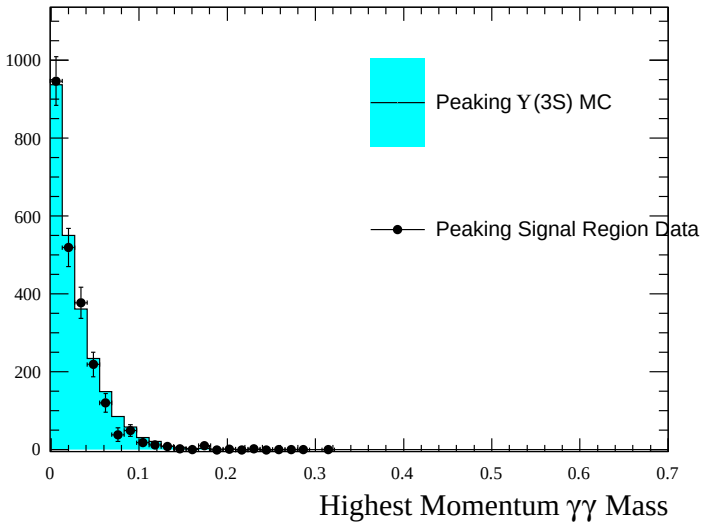


Figure 8.30: Highest lab-momentum photon pair-mass, compared between peaking background MC (cyan histogram) and the peaking component of the data projected from the recoil mass fit. MC is normalized to the $\Upsilon(3S)$ count, with corrections applied.

In section 7.3, we describe our approach to studying the non-leptonic peaking backgrounds and determine that the MC describes the yield of nearly invisible hadronic events to within a factor of 5. The non-leptonic peaking backgrounds represent 0.12% of the total peaking background. We therefore assign a 0.6% systematic uncertainty for this source, to be added in quadrature with the systematic uncertainty from the leptonic peaking background study.

8.4 Uncertainty on the $\Upsilon(3S)$ Count

This has been determined by the *BABAR* collaboration [17] to be 1.1%.

8.5 Systematic Uncertainties due to Signal Shape Parameters

We fix all of the parameters of the signal PDF in the fit to the data. We vary the parameters by their uncertainties determined from the 2-lepton control sample and observe a small variation on the branching fraction of just 0.033×10^{-4}. We include this as an additive systematic error on the branching fraction.

8.6 Total Systematic Uncertainty and Corrections

Table 8.6 summarizes the systematic uncertainties, their sources, and their applicability to the elements of the branching fraction extraction. Systematics which are correlated between signal efficiency and the peaking background estimate are marked by "[*]".

Source	**Correction**	**Uncertainty (%)**
Background Estimate		
2-lepton scaling correction [*]	1.088	1.1
1-lepton correction	1.000	2.1
hadronic peaking backgrounds	1.000	0.6
L1 Trigger [*]	1.000	2.1
L3 Trigger [*]	0.997	0.9
MVA [*]	1.000	4.0
Total (uncorrelated with Signal efficiency)	2.2	
Total (correlated with Signal efficiency) [*]	4.7	
Signal Efficiency		
2-lepton scaling correction [*]	1.088	1.1
L1 Trigger [*]	1.000	2.1
L3 Trigger [*]	0.997	0.9
MVA [*]	1.000	4.0
Total (uncorrelated with Peaking Bkg.)	0.0	
Total (correlated with Peaking Bkg.) [*]	4.7	
$\Upsilon(3S)$ Counting		
Counting	1.000	1.1
Additive Uncertainties (in BF units $\times 10^{-4}$)		
Signal Shape Parameters	N/A	0.033

Table 8.6: Systematic uncertainties and their applicability. The marker "[*]" indicates an uncertainty that is correlated between signal and peaking background.

Chapter 9

Results and Conclusions

We apply the extended maximum likelihood fit to the events in the data.

9.1 Unblinded Results

We unblind the fit region of the 91.42×10^6 $\Upsilon(3S)$ events (the sum of the "High" and "Medium" samples) and perform the final Extended Maximum Likelihood fit in m_{recoil}. For the fit, we:

- Fix the peaking background yield to 2357.8 events, which corresponds to the prediction from section 7.4 multiplied by the Level 3 trigger systematic correction from section 8.2.2. We fix the shape to that determined from the 2-lepton control sample
- Float the yield and shape of the non-peaking background
- Float the signal yield and fix the signal shape to that determined from the 2-lepton control sample

The remaining inputs to the unblinding are as follows:

- $\Upsilon(3S)$ Count: 91.42×10^6
- Signal efficiency: 16.4%
- Dipion transition BF: 4.48% [30]

We obtain a signal yield of

$$N_{signal} = -94.9 \pm 102.3 \tag{9.1}$$

The fit is shown in Fig. 9.1.

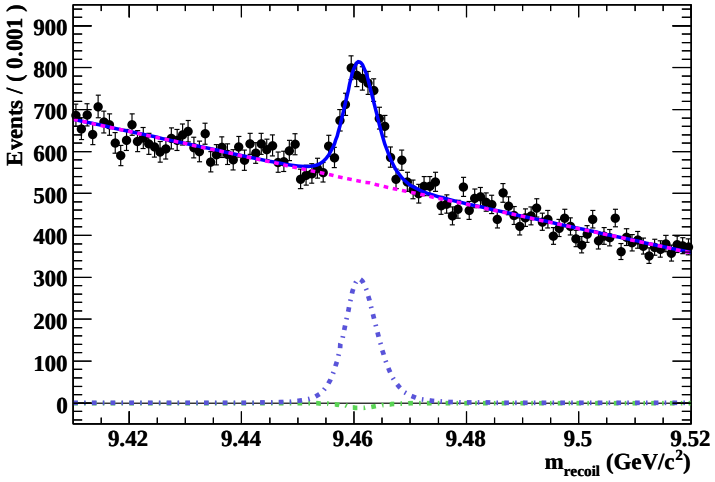


Figure 9.1: The result of the ML fit to the data sample. The total maximum likelihood fit is shown in solid blue; the non-peaking background component is shown in dashed magenta; the peaking background component is shown in blue dashed; the signal component is shown in green dashed.

We fit directly for the branching fraction $\mathcal{B}(\Upsilon(1S) \rightarrow \text{invisible})$ by writing the branching fraction in terms of the signal yield, as described at the beginning of Chapter 6.4.2.

$$\mathcal{B}(\Upsilon(1S) \rightarrow \text{invisible}) = \frac{n_{\text{invisible}}}{N_{\Upsilon(3S)} \cdot \mathcal{B}(\Upsilon(3S) \rightarrow \pi^+\pi^-\Upsilon(1S)) \cdot \varepsilon_{\text{invisible}}} \tag{9.2}$$

where $n_{\text{invisible}}$ is the fitted yield of events quoted above. This equation has several sources of uncertainty, detailed in Chapter 8. We write the peaking background in terms of correction factors ($\mathcal{C}$) determined to correct the MC prediction of this background:

$$N_{peaking} = N_{peaking}^{nominal} \cdot \prod_{i=0}^{n} \mathcal{C}_i. \tag{9.3}$$

We further write the efficiency of reconstructing signal and the efficiencies of mis-reconstructing $\Upsilon(1S) \rightarrow \ell^+\ell^-$ events as signal in terms of a component that includes correction factors, and their uncertainties. We expect the selection efficiencies to be the same for signal events and peaking background events, so these efficiencies are correlated between the signal and peaking background estimate. For instance,

$$\varepsilon_{signal} = \varepsilon_{signal}^{nominal} \cdot \prod_{i=0}^{n} \mathcal{C}_i. \tag{9.4}$$

This allows the fit to determine the systematic of varying the efficiencies (such as trigger efficiency) simultaneously for the peaking background and signal.

After incorporating the errors detailed in Chapter 8, we find a branching fraction of

$$\mathcal{B}(\Upsilon(1S) \rightarrow \text{invisible}) = (-2.4 \pm 1.4(\text{stat.}) \pm 1.7(\text{syst.})) \times 10^{-4} \tag{9.5}$$

where the systematic errors have been incorporated in the Likelihood function as Gaussians. Integrating the total Likelihood function upwards until it reaches 90% of its total area above 0,

we find that

$$\mathcal{B}(\Upsilon(1S) \to \text{invisible}) < 2.6 \times 10^{-4} \text{ at } 90\% \text{ CL} \tag{9.6}$$

This result is consistent with the Standard Model prediction (1×10^{-5}) for this decay, and improves upon the previous 90% confidence limit of 2.5×10^{-3} by almost an order of magnitude.

9.2 Conclusions

While this result probes into regions which could contain new physics, it does not approach the predictions of the Standard Model in sensitivity. Certainly there is substantial room between 1×10^{-5} and 2.9×10^{-4} in which light Dark Matter candidates or unparticles could be hiding. Additionally, should new physics be discovered elsewhere, this channel could be an effective precision probe of these physics effects if the differences between the Standard Model predictions and effects from these new physics effect the $\Upsilon(1S)$ system. This is the case in [6], where Super-Symmetric extensions to the standard model can suppress or enhance invisible Υ decays by a factor of 2.

This measurement on the *BABAR* data set is most limited by the systematic uncertainty on the expectation of the peaking background. Certainly the most direct way to improve this measurement will be by creating a more hermetic detector. A veto device close to the beam pipe to detect leptons from $\Upsilon(1S) \to \ell^+\ell^-$ decays would drastically reduce this problem. Trigger efficiency for low energy pions would also be extremely beneficial to this analysis, allowing a larger collection of $\Upsilon(1S)$ mesons from the same number of $\Upsilon(3S)$ decays. In addition to improving this analysis, though, that triggering ability would allow the measurements of other

invisible Υ decays through their dipion transitions. We cannot perform this analysis either on the transition $\Upsilon(3S) \to \pi^+\pi^-(\Upsilon(2S) \to \text{invisible})$ or $\Upsilon(2S) \to \pi^+\pi^-(\Upsilon(1S) \to \text{invisible})$ due to the low energies of the pions in those decays. As there are many interactions which have two low energy pions as a final state (as shown in the large non-peaking background to our analysis), an approach must be developed to to remove those events or the triggering rate will be much too large. For example a dependence on the p_T of the dipion system could be used to veto the vast majority of background coming from two photon fusion but allowing through almost all signal events. These consideration might be taken into account by those developing detectors for future high luminosity $b\bar{b}$ factories.

Appendix A

$\Upsilon(1S) \rightarrow$ invisible candidate variables, without a pre-cut on p_T

In section 5.1, we showed plots of candidate variables for the $\Upsilon(1S) \rightarrow \mathrm{invisible}$ analysis. These plots required events in the sideband data, signal MC, and generic $\Upsilon(3S)$ MC to pass the invisible signal selection, and have a pre-cut $p_T > 0.1\,\mathrm{GeV}/c$. We here show, for completeness, the variables without such a pre-cut.

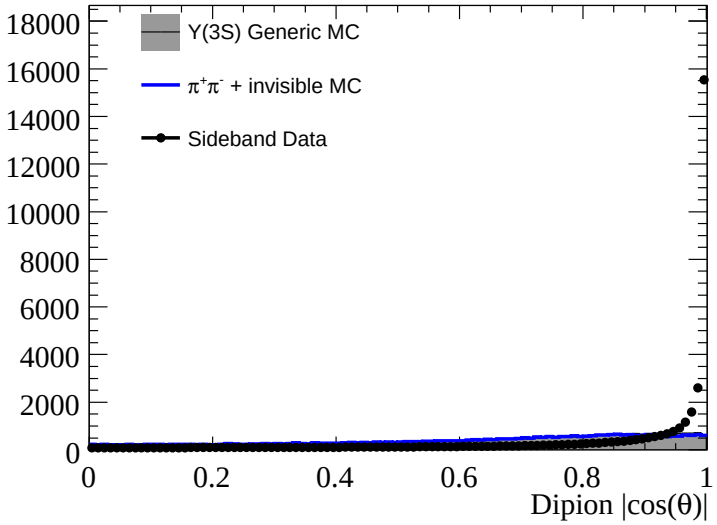
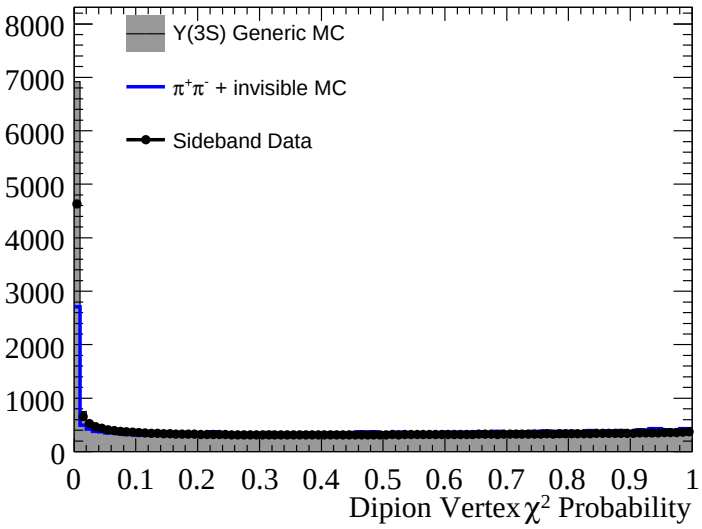
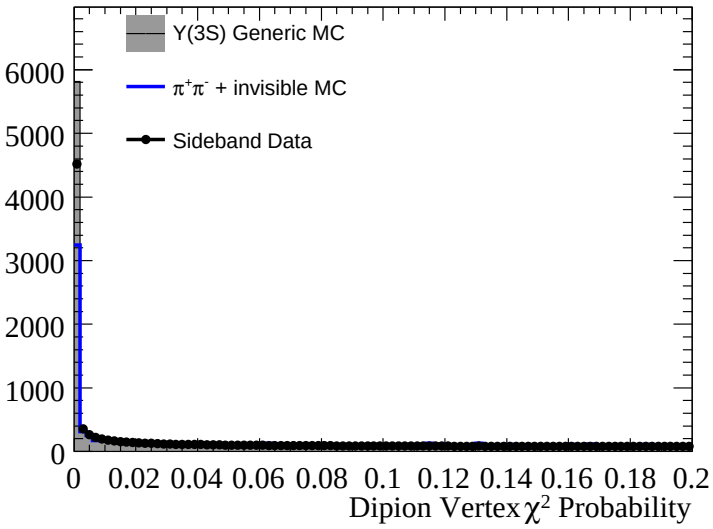


Figure A.1: $|cos(\theta_{\pi\pi})|$, where $\theta_{\pi\pi}$ is the lab polar angle of the dipion system, compared between signal MC (blue, unshaded histogram), sideband data (black points), and $\Upsilon(3S)$ generic MC (grey, shaded histogram). All distributions are normalized to the yield in the sideband data, for a pure comparison of the shape.



(a) The whole probability range, 0.0-1.0



(b) A zoomed probability range, 0.0-0.2

Figure A.2: The dipion vertex probability, compared between signal MC (blue, unshaded histogram), sideband data (black points), and $\Upsilon(3S)$ generic MC (grey, shaded histogram). All distributions are normalized to the yield in the sideband data, for a pure comparison of the shape.

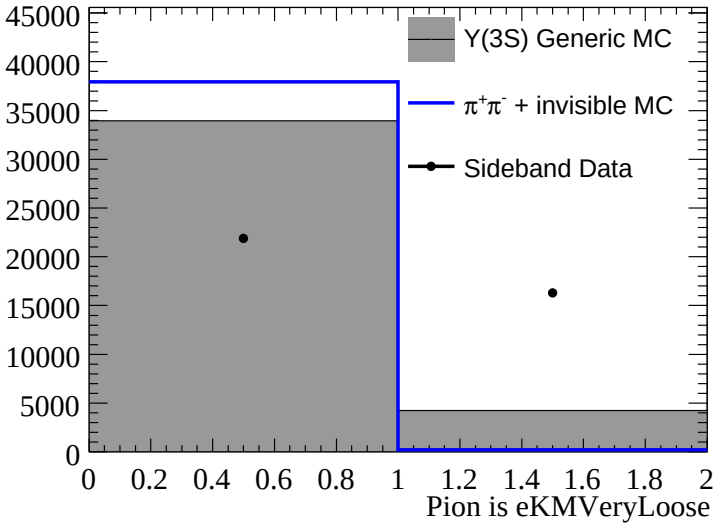


Figure A.3: A boolean that is true if either pion passes eKMVeryLoose, compared between signal MC (blue, unshaded histogram), sideband data (black points), and $\Upsilon(3S)$ generic MC (grey, shaded histogram). All distributions are normalized to the yield in the sideband data, for a pure comparison of the shape.

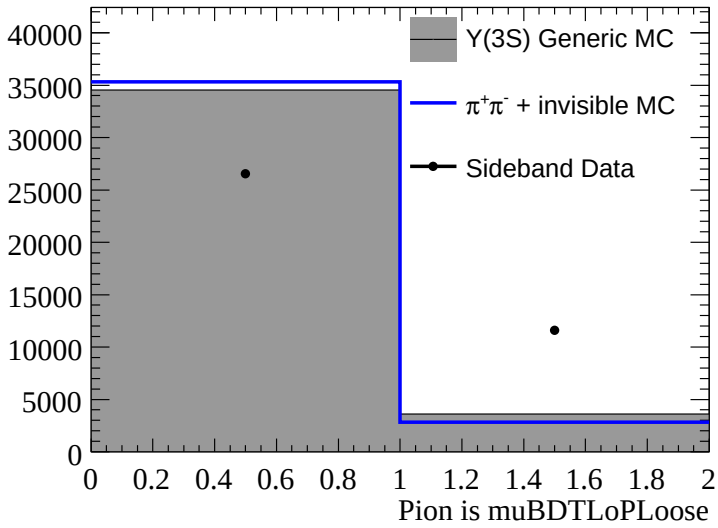


Figure A.4: A boolean that is true if either pion passes muBDTLoPLoose, compared between signal MC (blue, unshaded histogram), sideband data (black points), and $\Upsilon(3S)$ generic MC (grey, shaded histogram). All distributions are normalized to the yield in the sideband data, for a pure comparison of the shape.

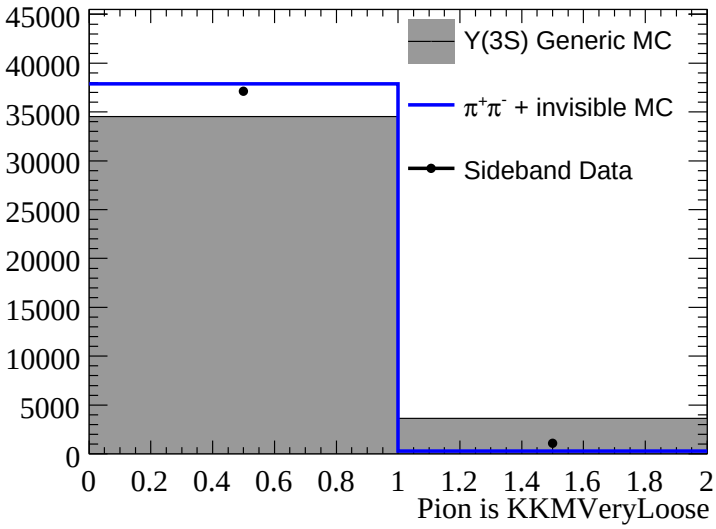


Figure A.5: A boolean that is true if either pion passes KKMVeryLoose, compared between signal MC (blue, unshaded histogram), sideband data (black points), and $\Upsilon(3S)$ generic MC (grey, shaded histogram). All distributions are normalized to the yield in the sideband data, for a pure comparison of the shape.

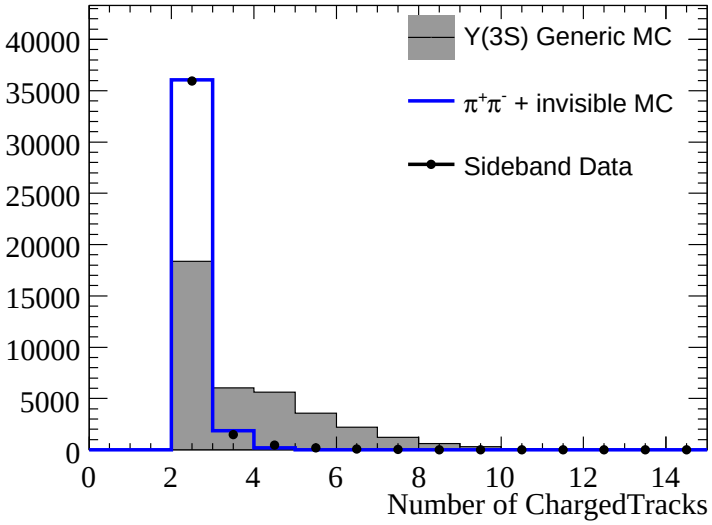


Figure A.6: The number of ChargedTracks in the event, compared between signal MC (blue, unshaded histogram), sideband data (black points), and $\Upsilon(3S)$ generic MC (grey, shaded histogram). All distributions are normalized to the yield in the sideband data, for a pure comparison of the shape.

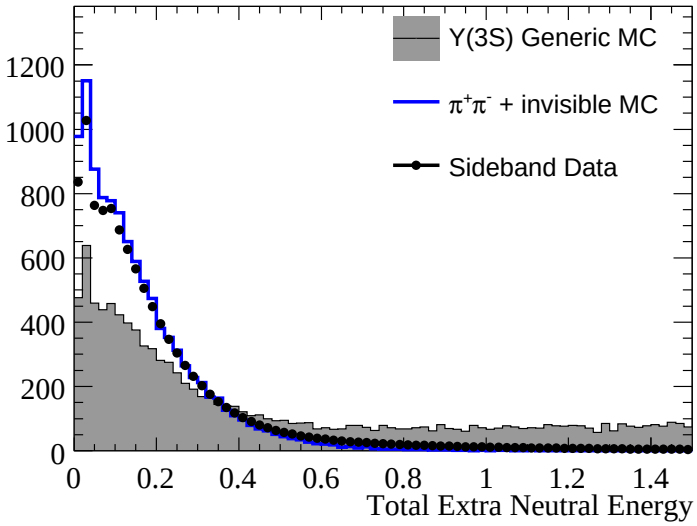


Figure A.7: The extra neutral energy, compared between signal MC (blue, unshaded histogram), sideband data (black points), and $\Upsilon(3S)$ generic MC (grey, shaded histogram). All distributions are normalized to the yield in the sideband data, for a pure comparison of the shape.

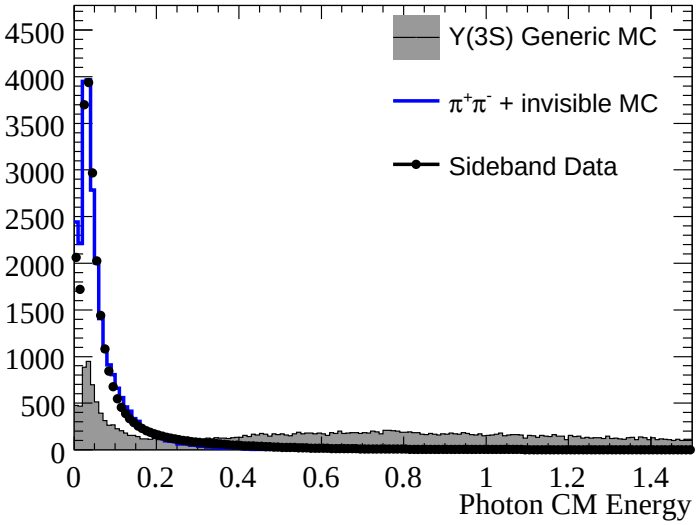
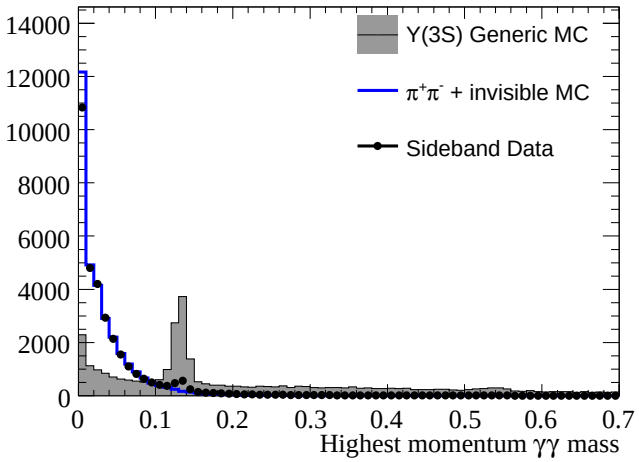
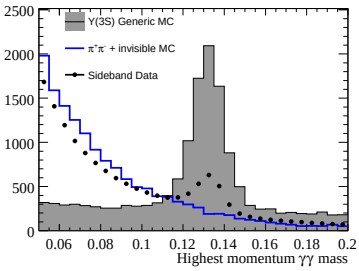


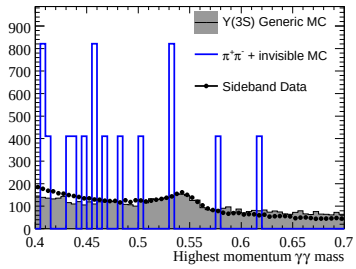
Figure A.8: The CM energy of the highest-energy cluster, compared between signal MC (blue, unshaded histogram), sideband data (black points), and $\Upsilon(3S)$ generic MC (grey, shaded histogram). All distributions are normalized to the yield in the sideband data, for a pure comparison of the shape.



(a) A wide range of the two-neutral mass



(b) The π^0 mass region



(c) The η mass region

Figure A.9: The mass of the highest-momentum (CM frame) neutral pair, compared between signal MC (blue, unshaded histogram), sideband data (black points), and $\Upsilon(3S)$ generic MC (grey, shaded histogram). All distributions are normalized to the yield in the sideband data, for a pure comparison of the shape.

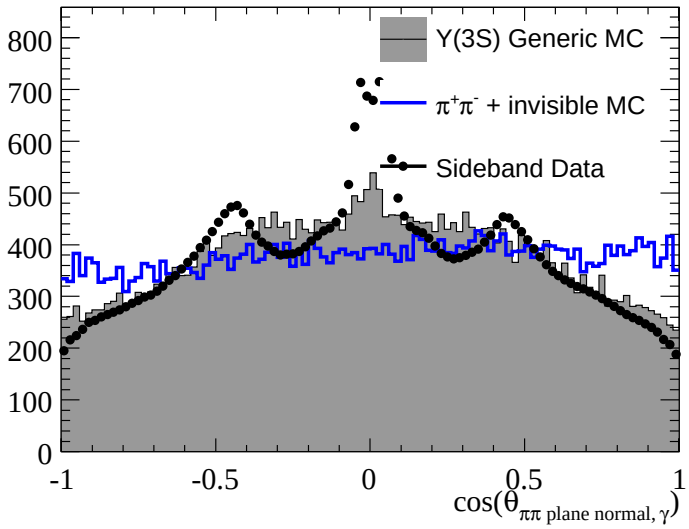


Figure A.10: The cosine of the angle (CM frame) between the highest energy neutral and the normal to the plane of the dipion system, compared between signal MC (blue, unshaded histogram), sideband data (black points), and $\Upsilon(3S)$ generic MC (grey, shaded histogram). All distributions are normalized to the yield in the sideband data, for a pure comparison of the shape. Events which do not have any extra photons are defaulted to 1.1 and are not shown on this plot.

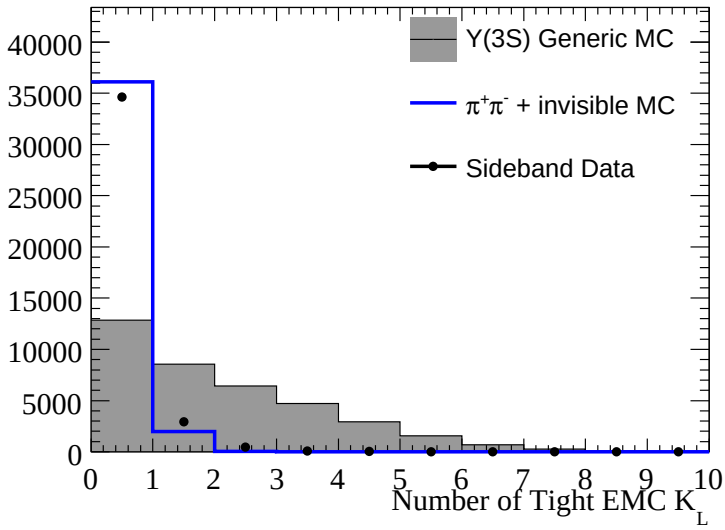


Figure A.11: The multiplicity of EMC K_L^0 candidates, compared between signal MC (blue, unshaded histogram), sideband data (black points), and $\Upsilon(3S)$ generic MC (grey, shaded histogram). All distributions are normalized to the yield in the sideband data, for a pure comparison of the shape.

Appendix B

Understanding Effects of Variable Correlations on ${}_s\mathcal{P}lots$

The following appendix deals with understanding the effects of correlations in data on ${}_s\mathcal{P}lots$ and how that effect can be removed from ${}_s\mathcal{P}lots$of MVA inputs.

B.1 Toy Monte Carlo used to understand the effect of correlations in data on ${}_s\mathcal{P}lots$

In our analysis, we fit in one variable and use ${}_s\mathcal{P}lots$ to project out the distribution in another variable. If there is a correlation in the data between the two variables, this can have an effect on the projected distribution that would not be expected if the correlation is not taken into account. This means that when we project out the signal portion of the data, there is a distortion being applied by the interaction of the ${}_s\mathcal{P}lot$ procedure and the correlations between the two

variables in the data.

In order to see if this effect can cause a significant problem, we create a toy Monte Carlo experiment with two variables (x and y) where there is a significant correlation in the background and no correlation in the signal. In this experiment, $x \in [0, 1]$ and $y \in [-1, 2.5]$ and the two PDF's are:

$$f_{signal} = e^{-\frac{(x-1/2)^2}{2(0.01)^2} - \frac{(y-1/2)^2}{2(0.1)^2}} \tag{B.1}$$

$$f_{background} = e^{-\frac{(y-(x+1/2))^2}{2(0.1)^2}} \tag{B.2}$$

A plot of the signal is shown in Fig. B.1 with a zoomed version in Fig. B.2. A plot of the background is shown in Fig. B.3. The correlation coefficient of Equation B.2 is 0.24.

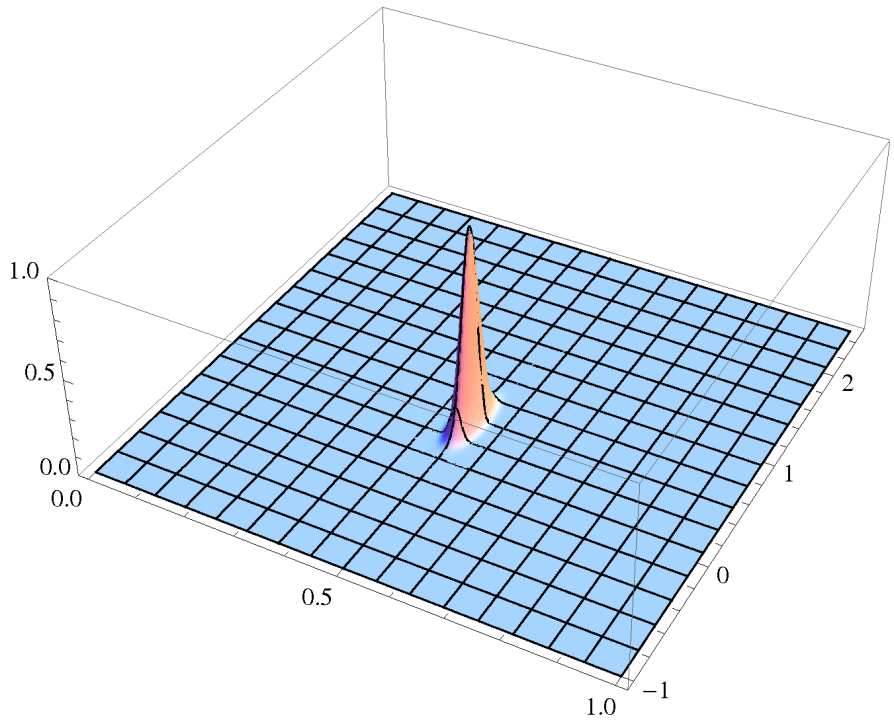


Figure B.1: Plot of eq. B.1, $x \in [0, 1]$ and $y \in [-1, 2.5]$

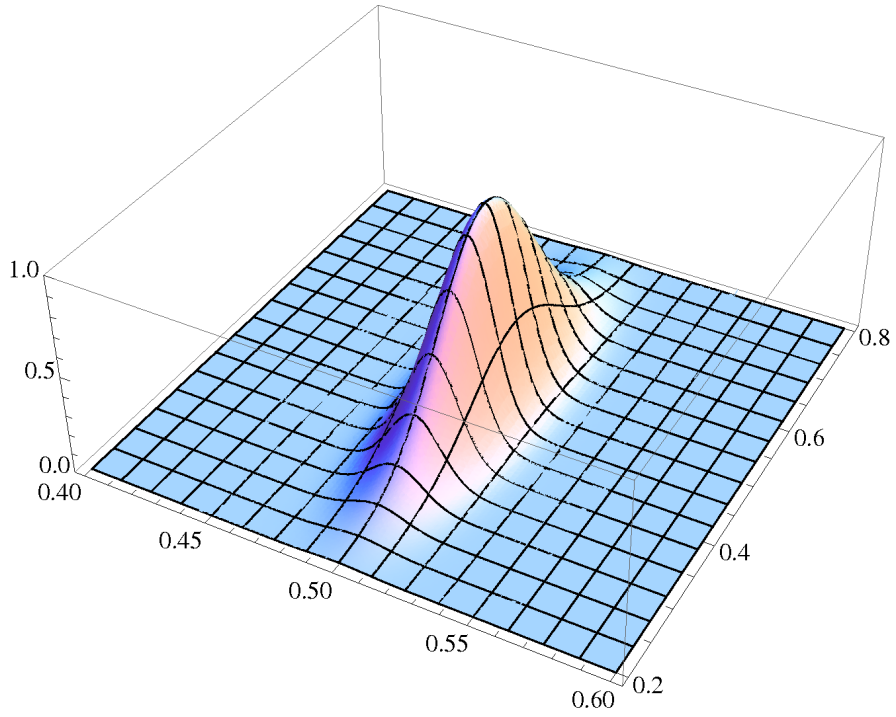


Figure B.2: Plot of eq. B.1, $x \in [0.4, 0.6]$ and $y \in [0.2, 0.8]$

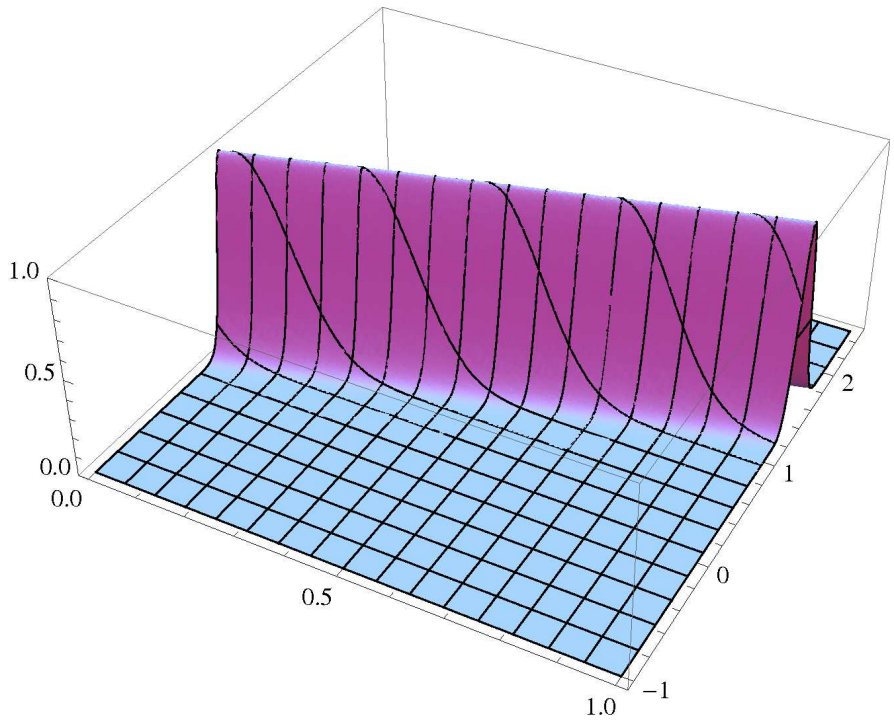


Figure B.3: Plot of eq. B.2, $x \in [0, 1]$ and $y \in [-1, 2.5]$

B.2 Toy Monte Carlo Experiment Results Part 1- Proof of Concept

The first run of this toy experiment has 1,000 signal events and 10,000 background events (experiment 1). The signal in x can be seen in fig. B.4, the background can be seen in fig. B.5, and their sum can be seen in B.6.

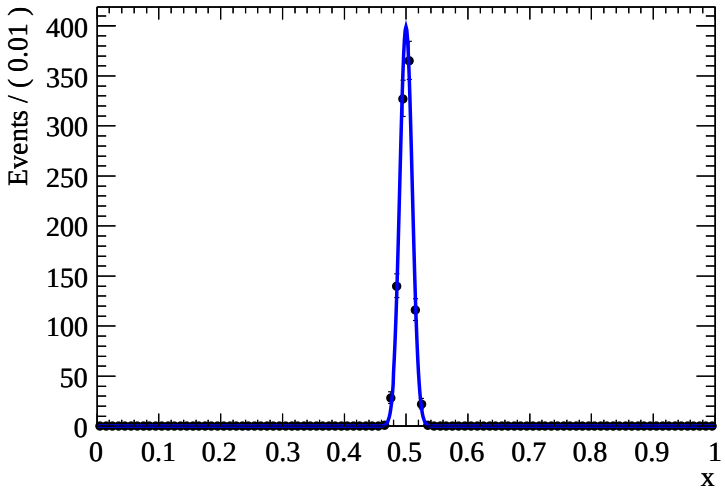


Figure B.4: Toy Monte Carlo signal simulation - experiment 1

We can then apply the ${}_s\mathcal{W}eighting$ procedure and derive a function (Fig. B.7) of the signal ${}_s\mathcal{W}eight$ in terms of x. Then we can look at the real and signal ${}_s\mathcal{W}eighted$ distributions in the y variable. The signal (Fig. B.8), background (Fig. B.9), and sum (Fig. B.10) can be seen to change as we expected, and introduce a large discrepancy from the true signal distribution in the sPlot.

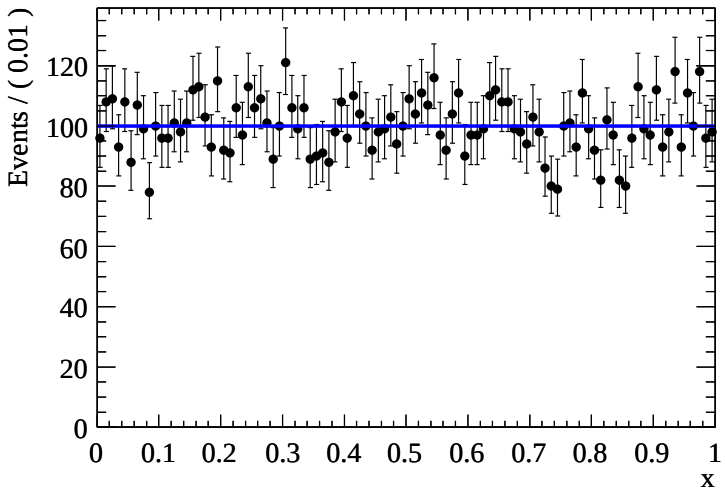


Figure B.5: Toy Monte Carlo background simulation - experiment 1

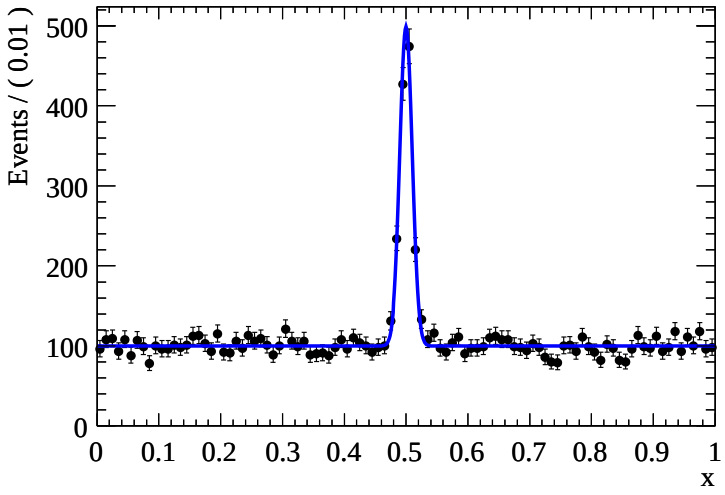


Figure B.6: Toy Monte Carlo combined simulation - experiment 1

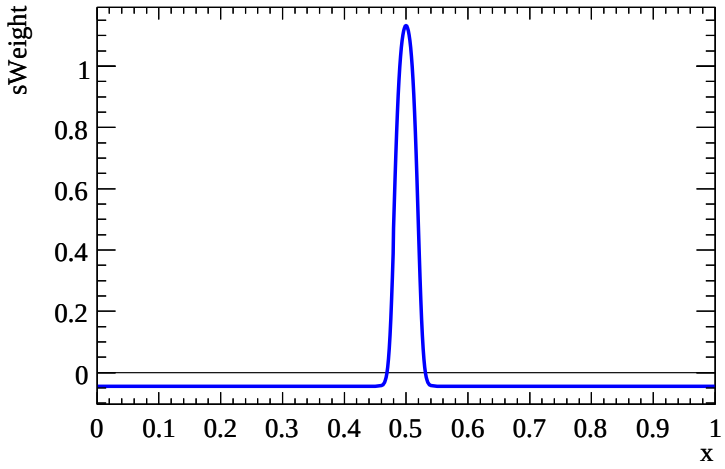


Figure B.7: Toy Monte Carlo signal ${}_s\mathcal{W}eight$ as a function of dipion x - experiment 1

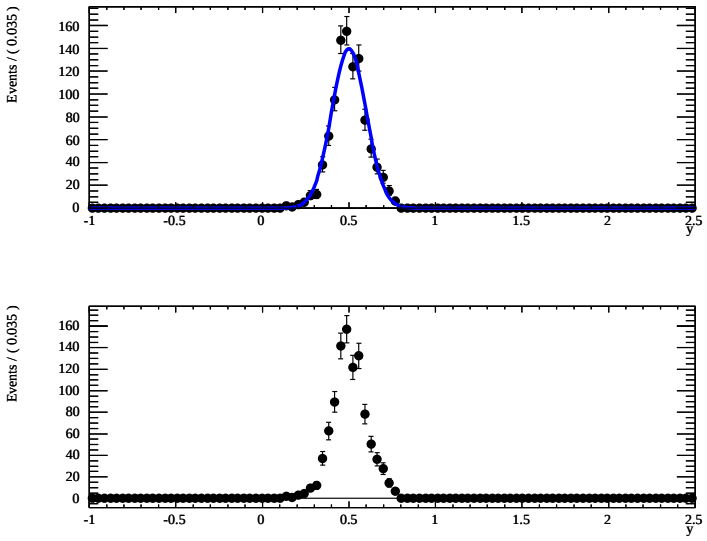


Figure B.8: Toy Monte Carlo signal simulation in y raw (above) and signal ${}_s\mathcal{P}lot$ (below) - experiment 1

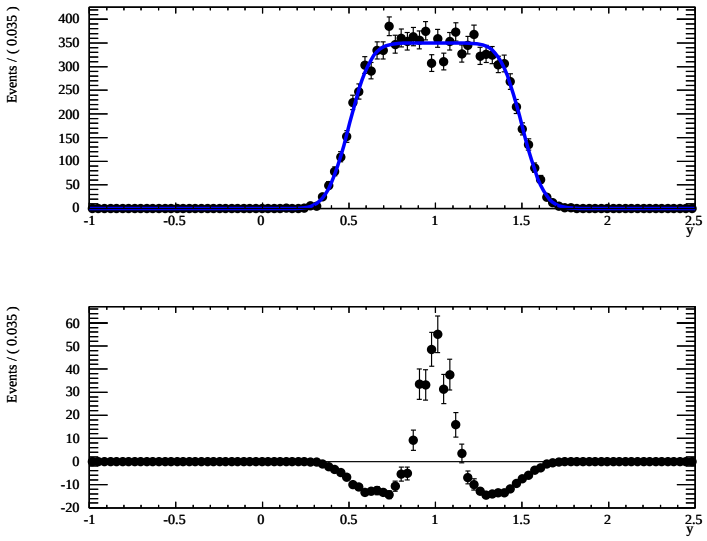


Figure B.9: Toy Monte Carlo background simulation in y raw (above) and signal ${}_s\mathcal{P}lot$ (below) - experiment 1

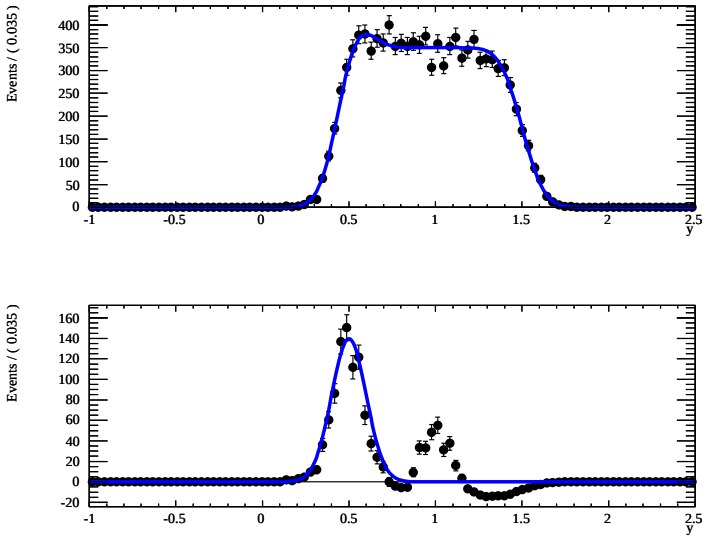


Figure B.10: Toy Monte Carlo combined simulation in y raw (above) and signal ${}_s\mathcal{P}lot$ (below) - experiment 1

B.3 Toy Monte Carlo Experiment Results Part II - Smaller Signal

We can run the same toy with 1,000 signal events and 100,000 background events (experiment 2). Plots similar to those in sec. B.2 can be seen in fig. B.11,B.12,B.13,B.14,B.15,B.16,B.17

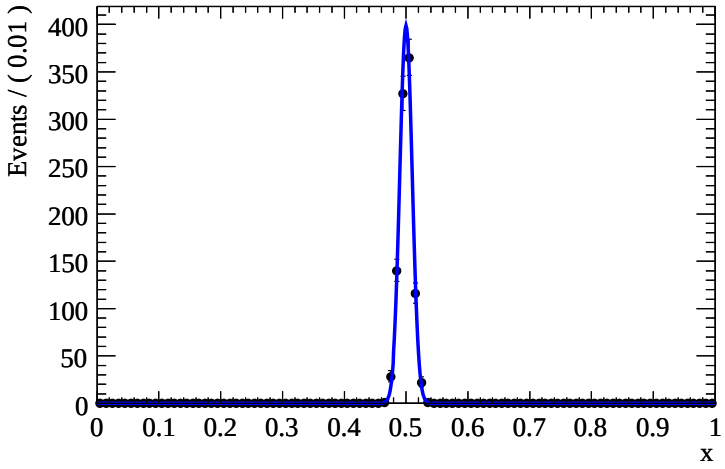


Figure B.11: Toy Monte Carlo signal simulation - experiment 2

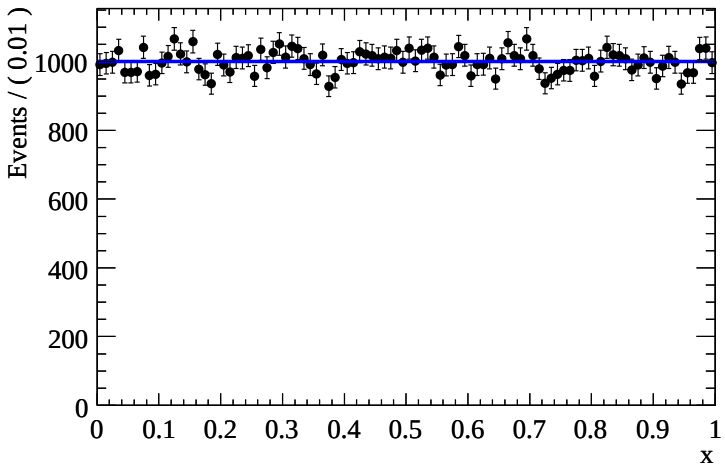


Figure B.12: Toy Monte Carlo background simulation - experiment 2

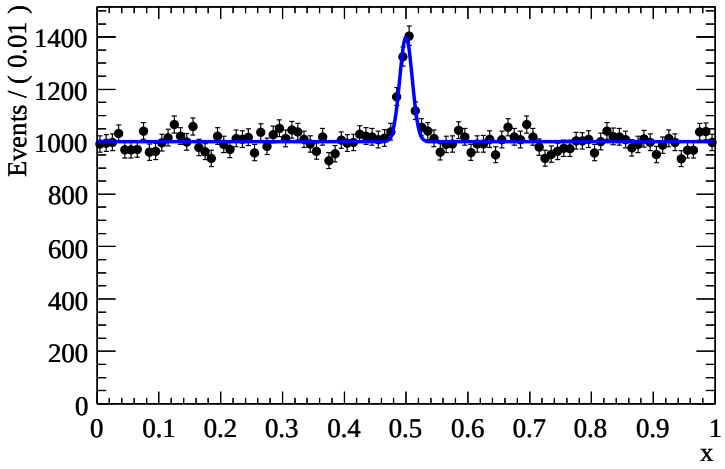


Figure B.13: Toy Monte Carlo combined simulation - experiment 2

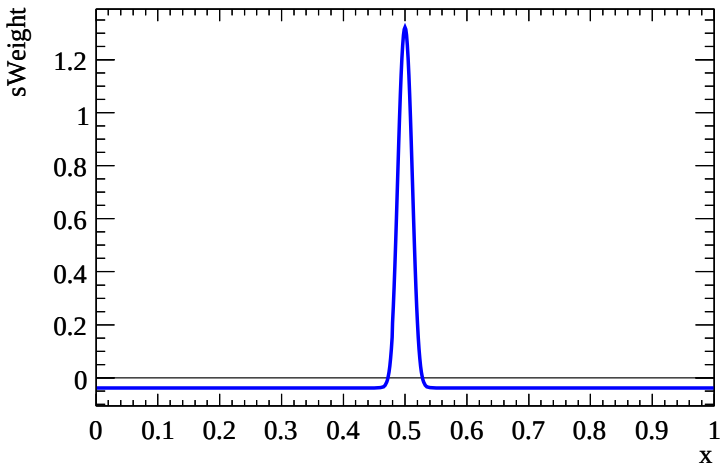


Figure B.14: Toy Monte Carlo signal ${}_s\mathcal{W}eight$ as a function of dipion x - experiment 2

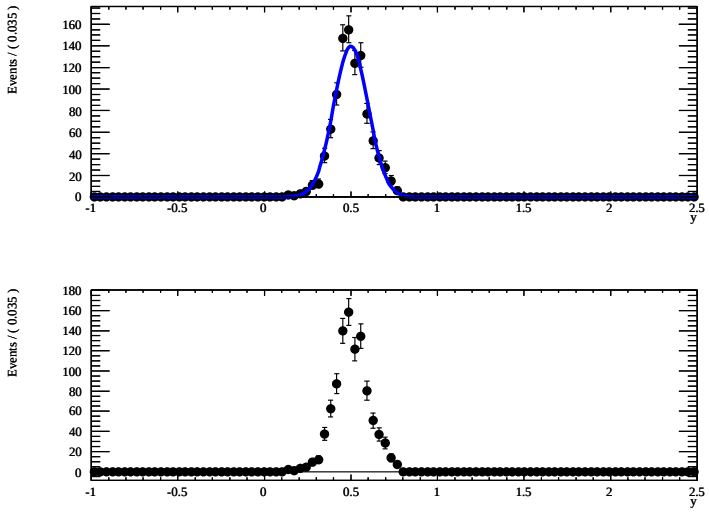


Figure B.15: Toy Monte Carlo signal simulation in y raw (above) and signal ${}_s\mathcal{P}lot$ (below) - experiment 2

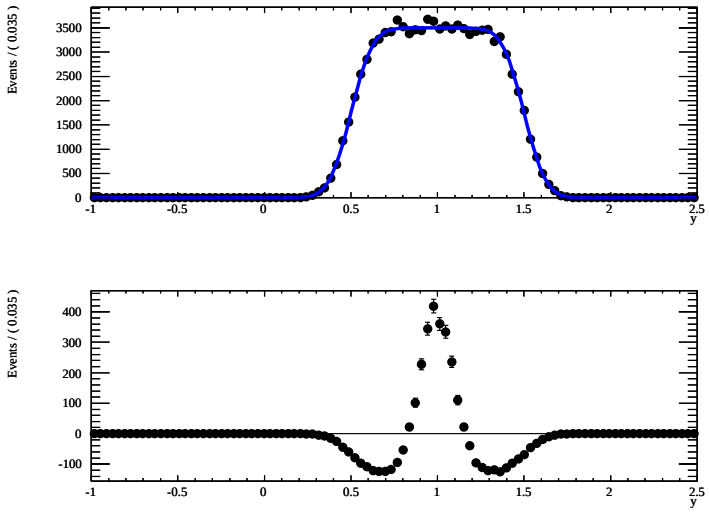


Figure B.16: Toy Monte Carlo background simulation in y raw (above) and signal ${}_s\mathcal{P}lot$ (below) - experiment 2

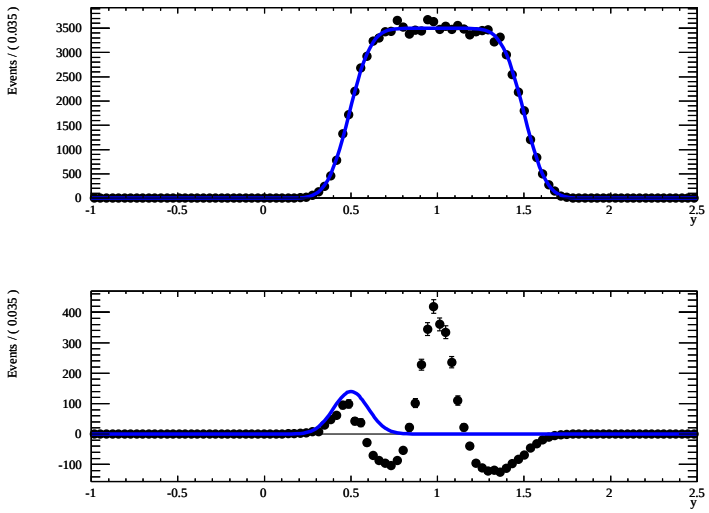


Figure B.17: Toy Monte Carlo combined simulation in y raw (above) and signal ${}_s\mathcal{P}lot$ (below) - experiment 2

B.4 Toy Monte Carlo Experiment Results Part III - Larger Signal

We can run the same toy with 50,000 signal events and 10,000 background events (experiment 3). This is similar to the case for our 1 lepton sample where there is very little background and we find good agreement between data and Monte Carlo. Plots similar to those in sec. B.2 can be seen in fig. B.18,B.19,B.20,B.21,B.22,B.23,B.24. The effect from the background is greatly reduced. In a simulation with 100,000 signal and 10,000 background events, the effect is invisible.

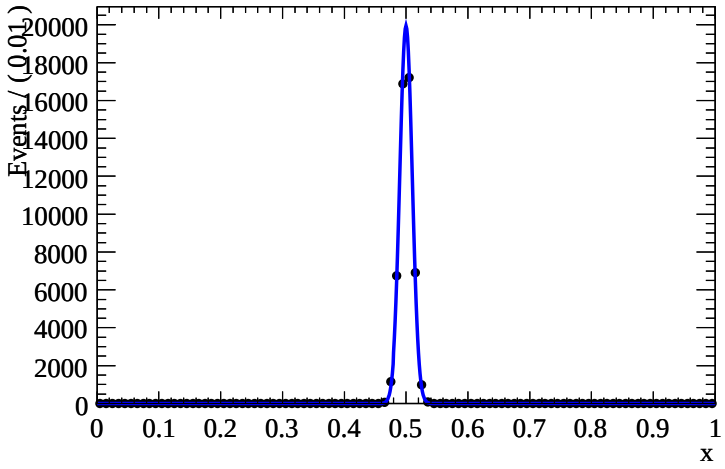


Figure B.18: Toy Monte Carlo signal simulation - experiment 3

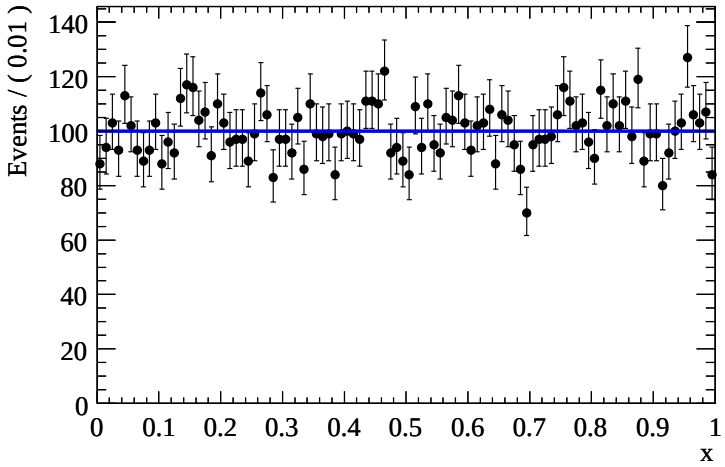


Figure B.19: Toy Monte Carlo background simulation - experiment 3

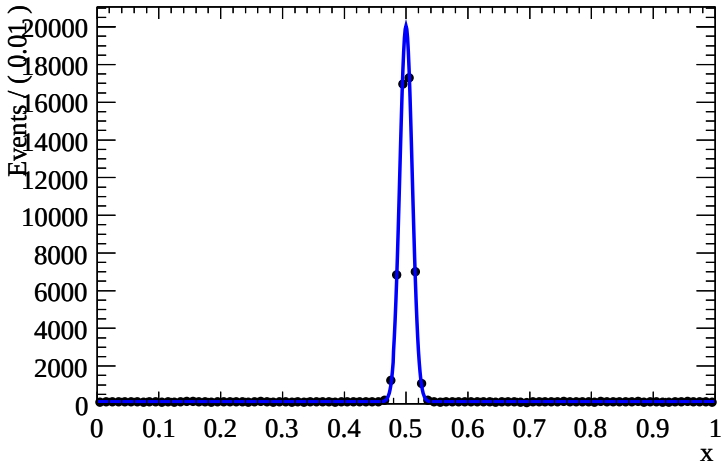


Figure B.20: Toy Monte Carlo combined simulation - experiment 3

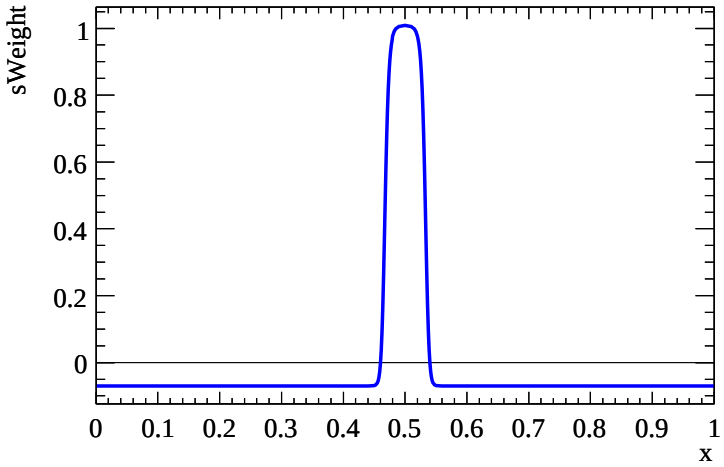


Figure B.21: Toy Monte Carlo signal ${}_s\mathcal{W}eight$ as a function of dipion x - experiment 3

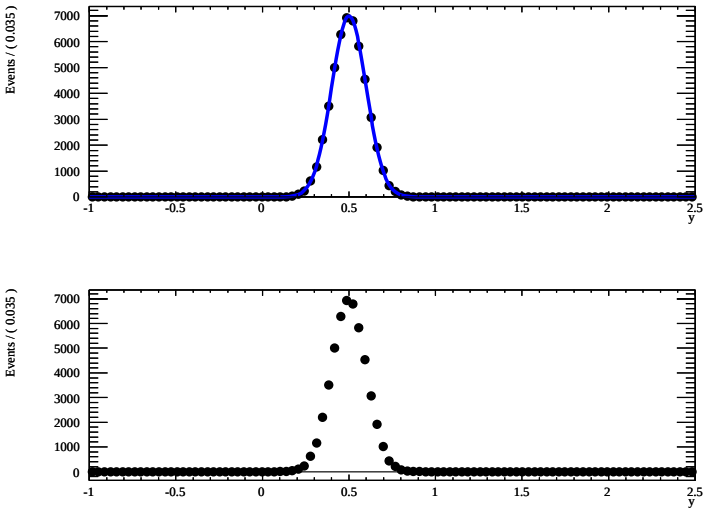


Figure B.22: Toy Monte Carlo signal simulation in y raw (above) and signal ${}_s\mathcal{P}lot$ (below) - experiment 3

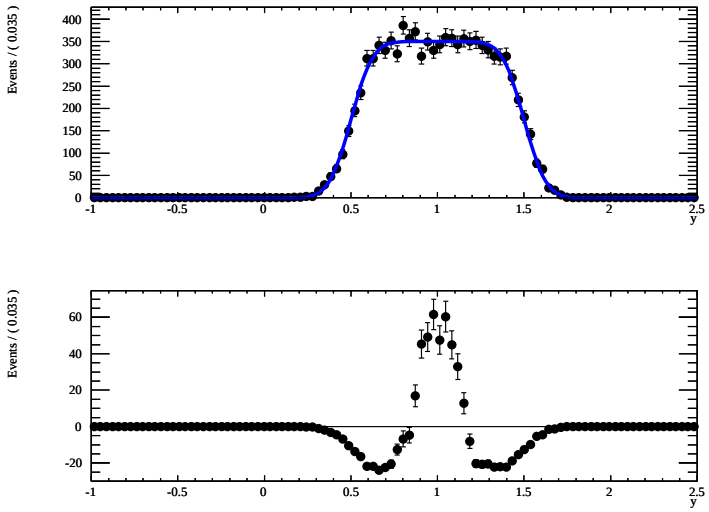


Figure B.23: Toy Monte Carlo background simulation in y raw (above) and signal ${}_s\mathcal{P}lot$ (below) - experiment 3

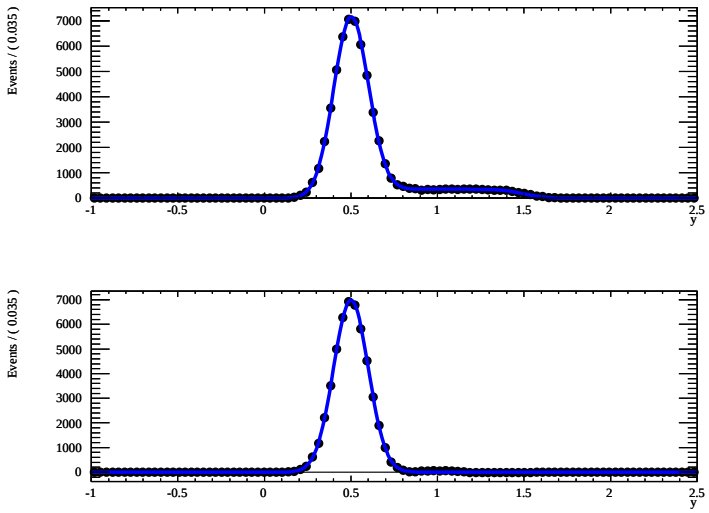


Figure B.24: Toy Monte Carlo combined simulation in y raw (above) and signal ${}_s\mathcal{P}lot$ (below) - experiment 3

B.5 Toy Monte Carlo Experiment Results Part IV - Less Correlated Background

We can change the toy experiment so that the correlation in the background is much weaker. We redefine the background function to be

$$f_{background} = e^{-\frac{(y-(0.05x+1))^2}{2(0.1)^2}} \tag{B.3}$$

The correlation coefficient of Equation B.3 is 0.036. A plot of this function can be seen in Fig. B.25.

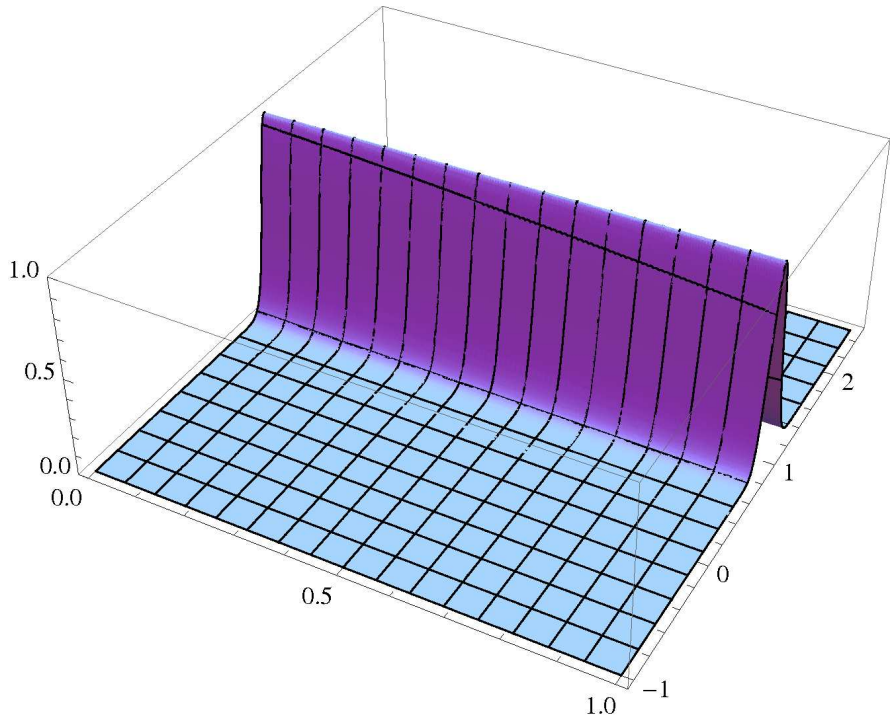


Figure B.25: Plot of eq. B.3, $x \in [0, 1]$ and $y \in [-1, 2.5]$

We generate 1,000 signal and 100,000 background events (experiment 4). We have

seen that in our real data, there is a correlation between the m_{recoil} and the dipion $\cos(\theta)$ angle, and this may be accountable for the change in projection shape. Plots similar to those in sec. B.2 can be seen in fig. B.26,B.27,B.28,B.29,B.30,B.31,B.32.

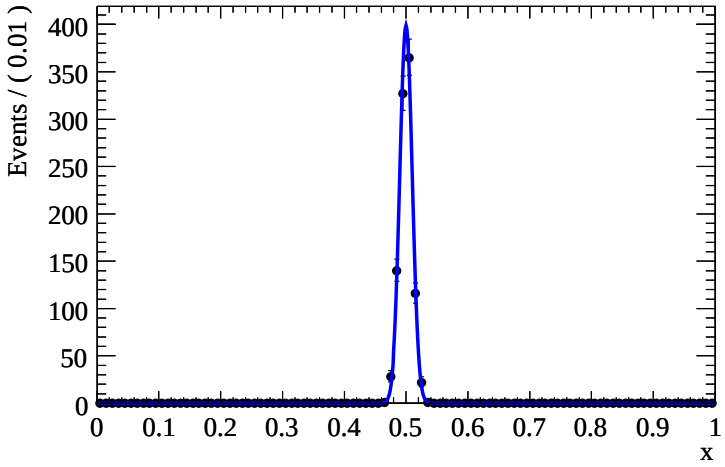


Figure B.26: Toy Monte Carlo signal simulation - experiment 4

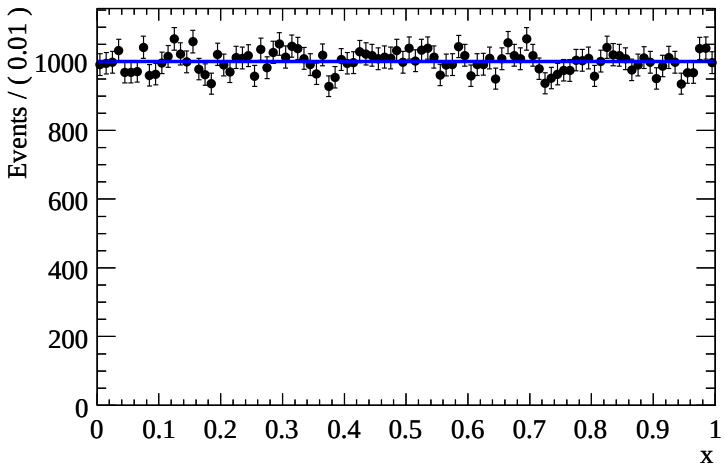


Figure B.27: Toy Monte Carlo background simulation - experiment 4

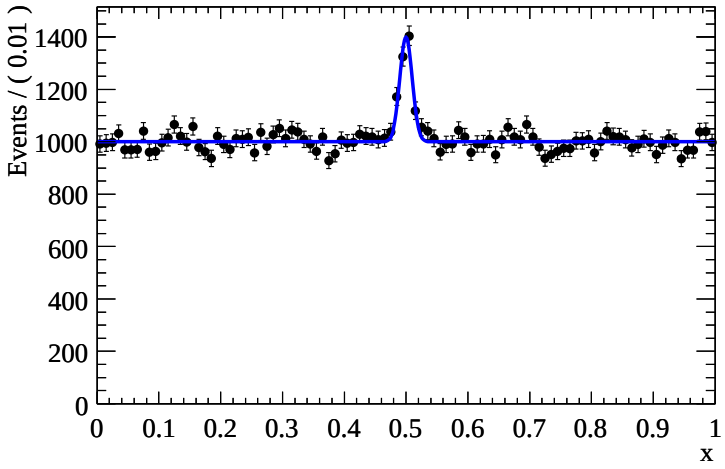


Figure B.28: Toy Monte Carlo combined simulation - experiment 4

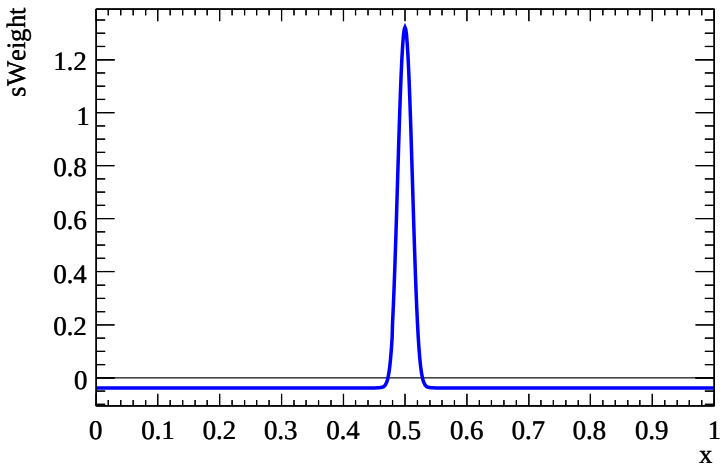


Figure B.29: Toy Monte Carlo signal ${}_s\mathcal{W}eight$ as a function of dipion x - experiment 4

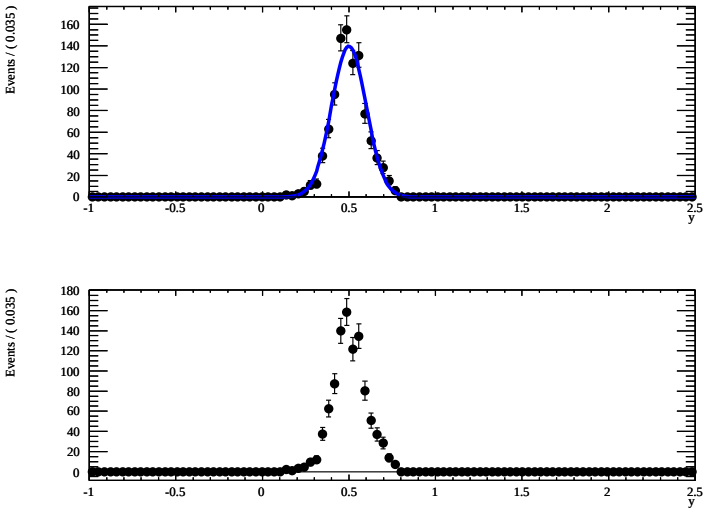


Figure B.30: Toy Monte Carlo signal simulation in y raw (above) and signal ${}_s\mathcal{P}lot$ (below) - experiment 4

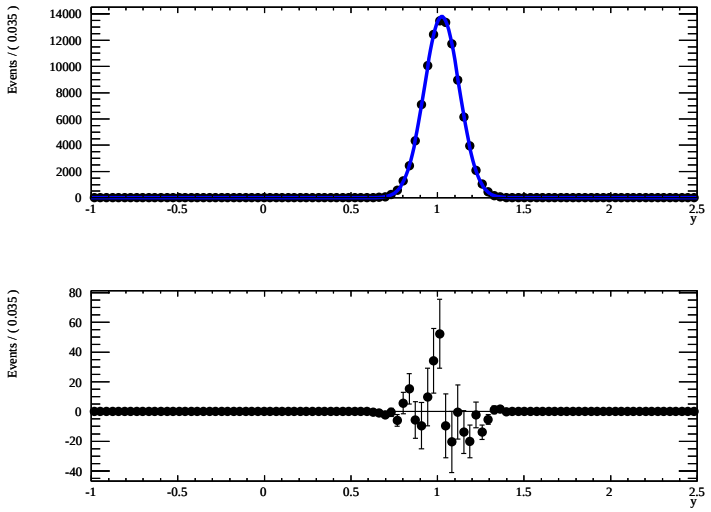


Figure B.31: Toy Monte Carlo background simulation in y raw (above) and signal ${}_s\mathcal{P}lot$ (below) - experiment 4

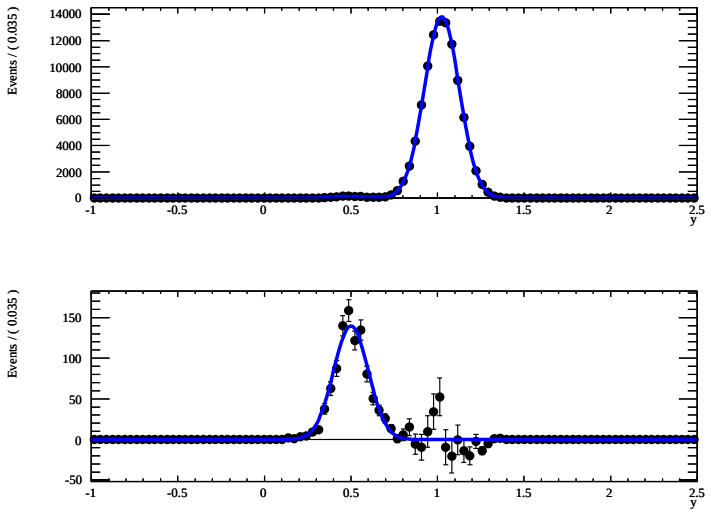


Figure B.32: Toy Monte Carlo combined simulation in y raw (above) and signal ${}_s\mathcal{P}lot$ (below) - experiment 4

B.6 Conclusion of ${}_s\mathcal{P}lots$ Toy Monte Carlo

Correlations in the variables involved in a ${}_s\mathcal{P}lot$ can have real effects on the distribution, and we should study this in out data sample.

B.7 Removing Correlation Corrections from Dipion $\cos(\theta)$

Since we have observed a correlation between m_{recoil} and the dipion $\cos(\theta)$ variables in our data, we wish to correct for this effect in our ${}_s\mathcal{P}lots$. While we have a simulated sample that predicts the peaking background distribution, we do not have a similar simulation for the non-peaking background. However, we do have the sideband data which we used to create our MVA and which is included in our maximum likelihood fit of the signal and background shape. We can take this sideband and use it to fit an estimate of the background distribution in the signal region.

After performing this fit, we use the extrapolated distribution to estimate what effect the correlation has on the ${}_s\mathcal{P}lot$ of peaking background. This distortion can then be removed from the ${}_s\mathcal{P}lot$ to compare directly with our simulation.

B.8 Fitting the Background Distribution

While it is difficult to create a two dimensional p.d.f of the m_{recoil} vs. $\cos(\theta)$ distribution, it is not difficult to define slices in $\cos(\theta)$ and fit the m_{recoil} distribution as a polynomial. We divide $\cos(\theta)$ into 10 bins, and perform an independent fit in each bin. This can be seen in Fig. B.33, B.34, B.35, B.36, B.37, B.38, B.39, B.40, B.41, and B.42.

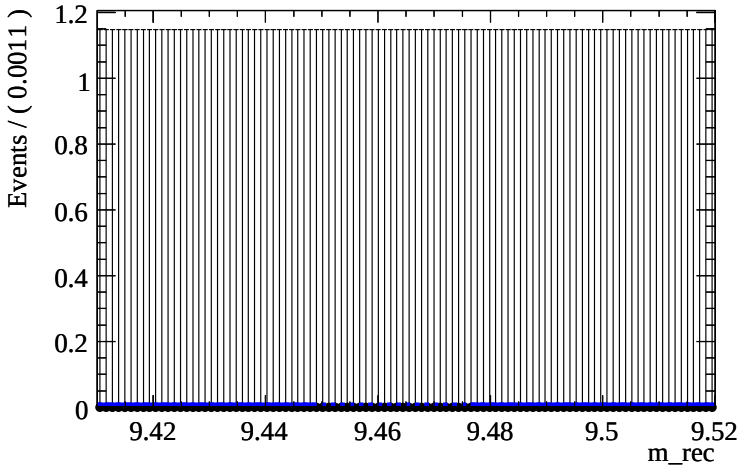


Figure B.33: m_{recoil} fit, $\cos(\theta) \in [-1.0, -0.8]$

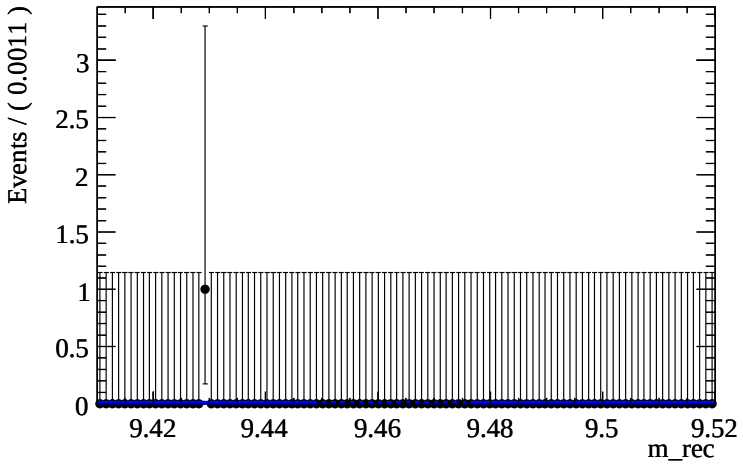


Figure B.34: m_{recoil} fit, $\cos(\theta) \in [-0.8, -0.6]$

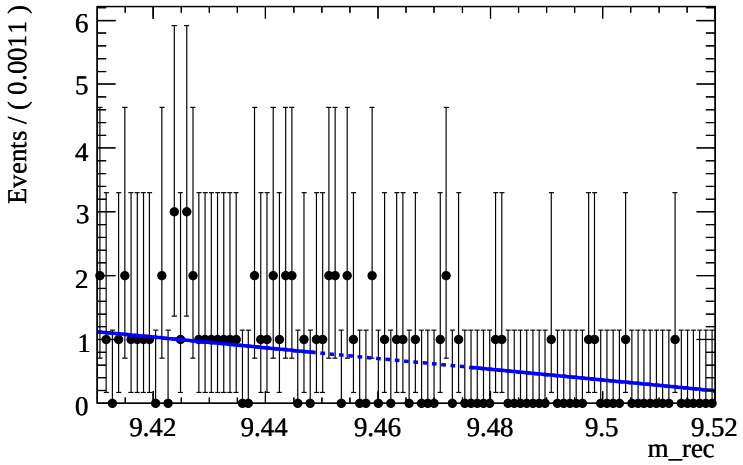


Figure B.35: m_{recoil} fit, $\cos(\theta) \in [-0.6, -0.4]$

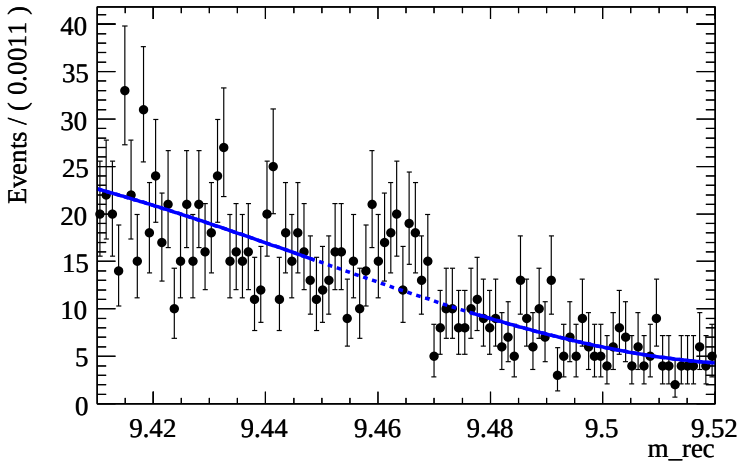


Figure B.36: m_{recoil} fit, $\cos(\theta) \in [-0.4, -0.2]$

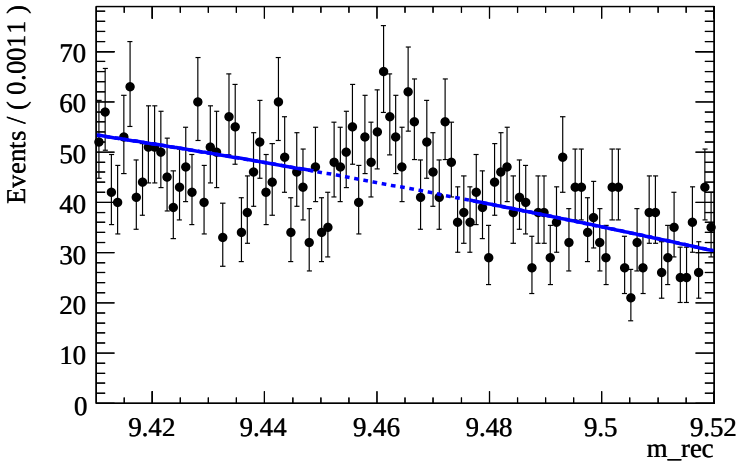


Figure B.37: m_{recoil} fit, $\cos(\theta) \in [-0.2, 0.0]$

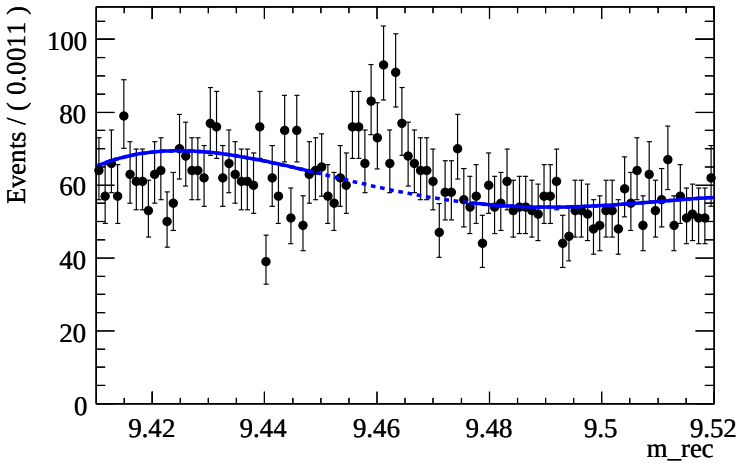


Figure B.38: m_{recoil} fit, $\cos(\theta) \in [0.0, 0.2]$

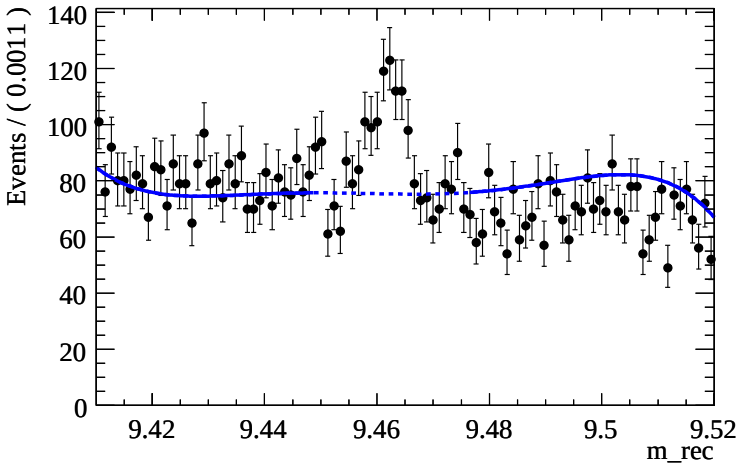


Figure B.39: m_{recoil} fit, $\cos(\theta) \in [0.2, 0.4]$

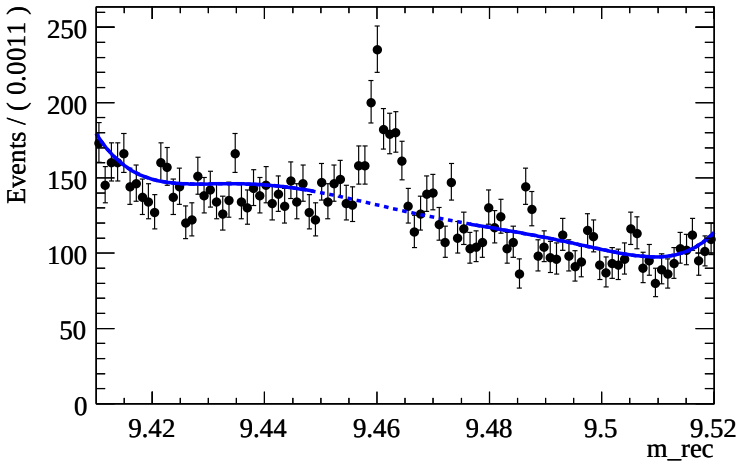


Figure B.40: m_{recoil} fit, $\cos(\theta) \in [0.4, 0.6]$

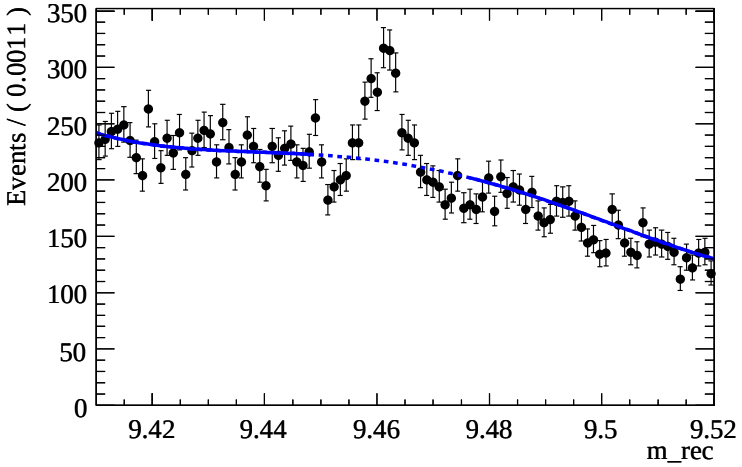


Figure B.41: m_{recoil} fit, $\cos(\theta) \in [0.6, 0.8]$

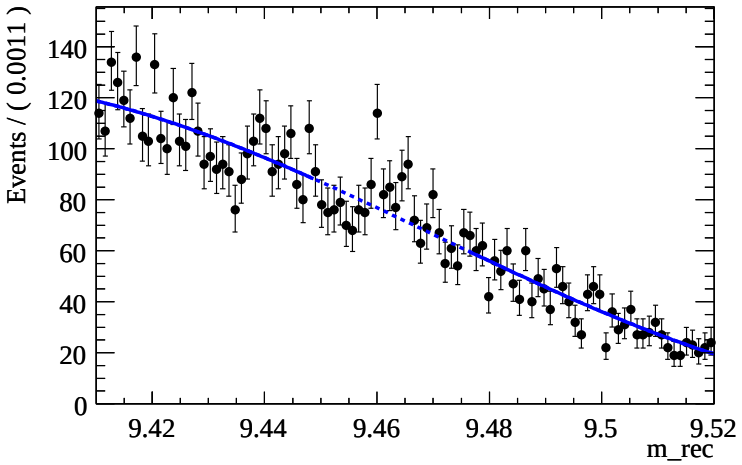


Figure B.42: m_{recoil} fit, $\cos(\theta) \in [0.8, 1.0]$

B.9 Adding ${}_s\mathcal{W}eighting$ to the Fitted Background p.d.f

From our complete maximum likelihood fit, we derive a function of signal ${}_s\mathcal{W}eight$ in terms of m_{recoil} (Fig. B.43). We can then take this function and multiply it by the p.d.f of the background that we have just derived (Sec. B.8) in the signal region for each bin in $\cos(\theta)$ to create a p.d.f of signal ${}_s\mathcal{W}eights$. This is multiplied by the number of non-peaking background events predicted to be in the signal region to find the total expected signal ${}_s\mathcal{W}eight$ contributed by the non-peaking background in the signal region for each bin of $\cos(\theta)$ (Fig. B.44). Similarly, the sum of the square of the signal ${}_s\mathcal{W}eights$ is calculated to compute the errors as $\sqrt{\sum({}_s\mathcal{W}eight)^2}$. These are conservative errors, as there is no consideration of how fitting through the signal region might effect the results.

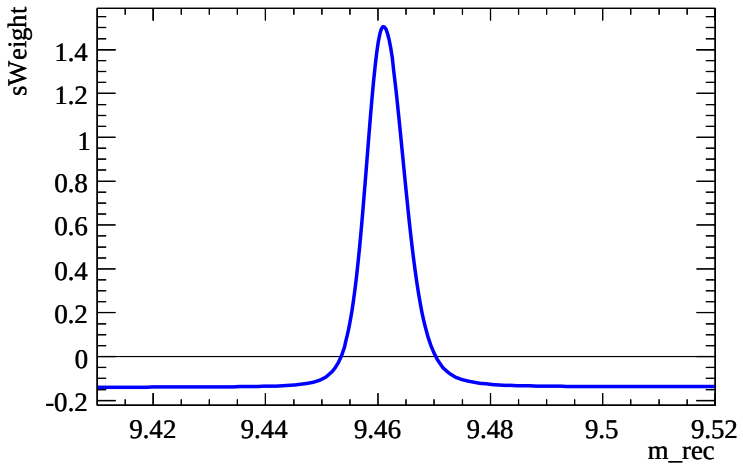


Figure B.43: Signal ${}_s\mathcal{W}eight$ as a function of m_{recoil}

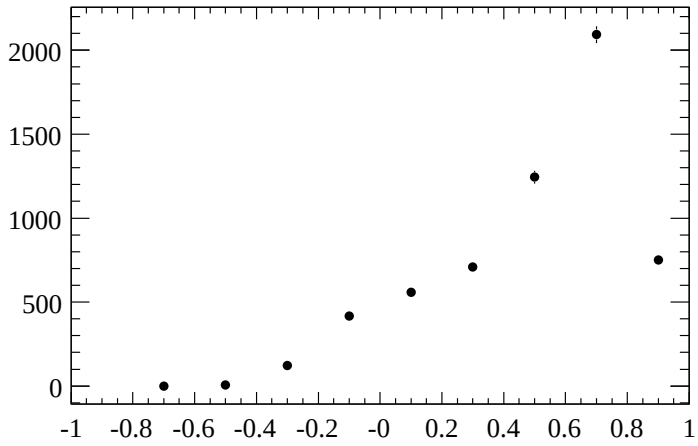


Figure B.44: Result of $\cos(\theta)$ slice fitting and signal ${}_s\mathcal{W}eighting$ with errors calculated as the square root of the sum of the signal ${}_s\mathcal{W}eights$. This represents the expected contribution to the ${}_s\mathcal{P}lot$ from the non-peaking background in the signal region.

B.10 Total Non-Peaking Background Distortion

Now that we have the distribution of the signal ${}_s\mathcal{W}eights$ in the signal region (Fig. B.44), we add the signal ${}_s\mathcal{W}eighted$ sideband data (Fig. B.45) and find the distortion introduced by the background correlation (Fig. B.46).

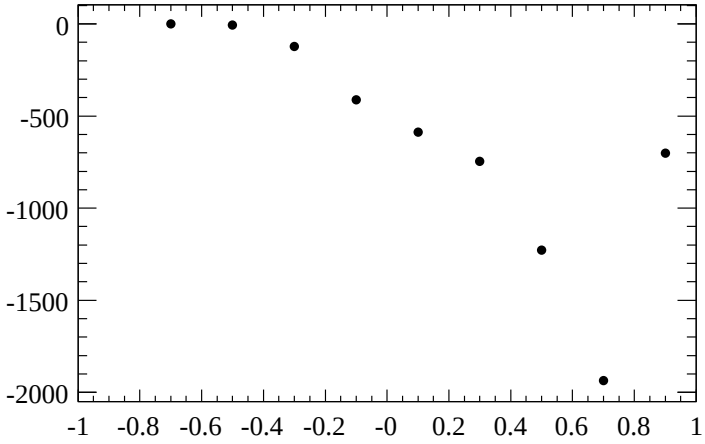


Figure B.45: Signal ${}_s\mathcal{P}lot$ of the sideband data. This is contribution to the ${}_s\mathcal{P}lot$ from the data in the sideband region

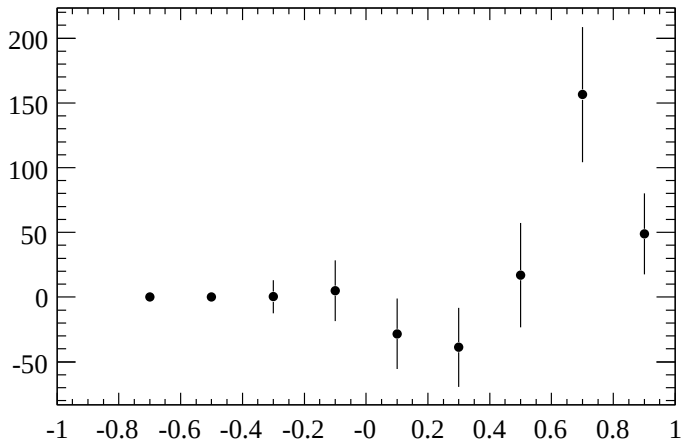


Figure B.46: Distortion to signal ${}_s\mathcal{P}lot$ from m_{recoil} and $\cos(\theta)$ correlation (sum of Fig. B.44 and B.45). This is the expected signal ${}_s\mathcal{P}lot$ from the non-peaking background distribution in both the sideband and signal region. It is subtracted from the signal ${}_s\mathcal{P}lot$ of the full data set to arrive at the signal ${}_s\mathcal{P}lot$ of the peaking background distribution (Fig. B.47)

B.11 Correcting the ${}_s\mathcal{P}lot$

We can now subtract the calculated distortion to the signal ${}_s\mathcal{P}lot$ from the signal ${}_s\mathcal{P}lot$ of the unblinded data and compare it to the signal ${}_s\mathcal{P}lot$ of the Monte Carlo simulation of the peaking background. This is displayed in Fig. B.47.

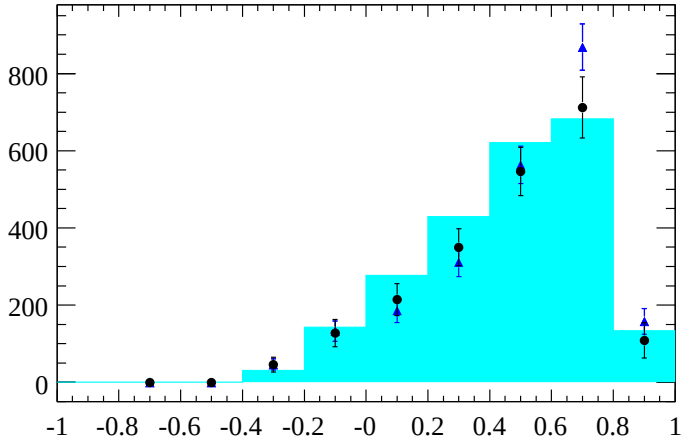


Figure B.47: Uncorrected data signal ${}_s\mathcal{P}lot$ in dark blue with triangular markers, corrected data signal ${}_s\mathcal{P}lot$ in black with circular markers, Monte Carlo simulation of non-peaking background signal ${}_s\mathcal{P}lot$ in Cyan histogram.

Bibliography

[1] **DELPHI** Collaboration, P. Abreu *et al.*, "Search for new phenomena using single photon events in the DELPHI detector at LEP," *Z. Phys.* **C74** (1997) 577–586.

[2] **L3** Collaboration, M. Acciarri *et al.*, "Determination of the number of light neutrino species from single photon production at LEP," *Phys. Lett.* **B431** (1998) 199–208.

[3] **OPAL** Collaboration, R. Akers *et al.*, "Measurement of single photon production in e+ e- collisions near the Z0 resonance," *Z. Phys.* **C65** (1995) 47–66.

[4] **ALEPH** Collaboration, D. Buskulic *et al.*, "A Direct measurement of the invisible width of the Z from single photon counting," *Phys. Lett.* **B313** (1993) 520–534.

[5] **BABAR** Collaboration, B. Aubert *et al.*, "Search for B^0 decays to invisible final states and to $\nu\bar{\nu}\gamma$," *Phys. Rev. Lett.* **93** (2004) 091802, `arXiv:hep-ex/0405071`.

[6] L. N. Chang, O. Lebedev, and J. N. Ng, "On the invisible decays of the Upsilon and J/psi resonances," *Phys. Lett.* **B441** (1998) 419–424, `arXiv:hep-ph/9806487`.

[7] B. McElrath, "Invisible quarkonium decays as a sensitive probe of dark matter," *Phys. Rev.* **D72** (2005) 103508, `arXiv:hep-ph/0506151`.

[8] B. McElrath, "Light Higgses and Dark Matter at Bottom and Charm Factories," arXiv:0712.0016 [hep-ph].

[9] H. Georgi, "Unparticle Physics," *Phys. Rev. Lett.* **98** (2007) 221601, arXiv:hep-ph/0703260.

[10] S.-L. Chen, X.-G. He, and H.-C. Tsai, "Constraints on Unparticle Interactions from Invisible Decays of Z, Quarkonia and Neutrinos," *JHEP* **11** (2007) 010, arXiv:0707.0187 [hep-ph].

[11] **CLEO** Collaboration, P. Rubin *et al.*, "Search for invisible decays of the Upsilon(1S) resonance," *Phys. Rev.* **D75** (2007) 031104, arXiv:hep-ex/0612051.

[12] **Belle** Collaboration, O. Tajima *et al.*, "Search for invisible decay of the Upsilon(1S)," *Phys. Rev. Lett.* **98** (2007) 132001, arXiv:hep-ex/0611041.

[13] M. Pivk and F. R. Le Diberder, "sPlot: a statistical tool to unfold data distributions," *Nucl. Instrum. Meth.* **A555** (2005) 356–369, arXiv:physics/0402083.

[14] **BABAR** Collaboration, B. Aubert *et al.*, "The BaBar detector," *Nucl. Instrum. Meth.* **A479** (2002) 1–116, arXiv:hep-ex/0105044.

[15] "PEP-II: An Asymmetric B Factory. Conceptual Design Report. June 1993,". SLAC-418.

[16] **BABAR** Collaboration, B. Aubert *et al.*, "Observation of the bottomonium ground state in the decay $\upsilon_{3S} \to \gamma eta_b$," *Phys. Rev. Lett.* **101** (2008) 071801, arXiv:0807.1086 [hep-ex].

[17] C. Hearty and G. McGregor, “Hadronic and Gamma-Pair Event Selection for Upsilon(3S) Counting,” *BABAR Analysis Document* (2008) .

[18] R. Bartoldus, A. Eisner, A. Ryd, and D. Su, “Trigger and filter documentation for the Run1 data,” *BABAR Analysis Document* (2000) .

[19] J. Friedman and N. Fisher, “Bump hunting in high dimensional data,” *Statistics and Computing* **9** (1999) 123–143.

[20] L. Breiman, “Random forests,” *Mach. Learn.* **45** (2001) no. 1, 5–32.

[21] **CLEO** Collaboration, D. Cronin-Hennessy *et al.*, “Study of Di-Pion Transitions Among Upsilon(3S), Upsilon(2S), and Upsilon(1S) States,” *Phys. Rev.* **D76** (2007) 072001, arXiv:0706.2317 [hep-ex].

[22] D. J. Lange and A. Ryd, “EvtGen A Monte Carlo Generator for B-Physics,” *BABAR Analysis Document* (2003) .

[23] T. Sjostrand, “High-energy physics event generation with PYTHIA 5.7 and JETSET 7.4,” *Comput. Phys. Commun.* **82** (1994) 74–90.

[24] T. Sjostrand *et al.*, “High-energy-physics event generation with PYTHIA 6.1,” *Comput. Phys. Commun.* **135** (2001) 238–259, arXiv:hep-ph/0010017.

[25] S. Jadach, W. Placzek, and B. F. L. Ward, “BHWIDE 1.00: O(alpha) YFS exponentiated Monte Carlo for Bhabha scattering at wide angles for LEP1/SLC and LEP2,” *Phys. Lett.* **B390** (1997) 298–308, arXiv:hep-ph/9608412.

[26] F. A. Berends and R. Kleiss, “Distributions in the Process e+ e- → mu+ mu- (gamma),” *Nucl. Phys.* **B177** (1981) 237.

[27] F. A. Berends, R. Kleiss, and S. Jadach, “Radiative Corrections to Muon Pair and Quark Pair Production in electron-Positron Collisions in the Z(0) Region,” *Nucl. Phys.* **B202** (1982) 63.

[28] F. A. Berends and R. Kleiss, “Distributions in the Process e+ e- → e+ e- (Gamma),” *Nucl. Phys.* **B228** (1983) 537.

[29] **GEANT4** Collaboration, S. Agostinelli *et al.*, “GEANT4: A simulation toolkit,” *Nucl. Instrum. Meth.* **A506** (2003) 250–303.

[30] **Particle Data Group** Collaboration, C. Amsler *et al.*, “Review of particle physics,” *Phys. Lett.* **B667** (2008) 1.

[31] D. N. Brown, “Kalman filter tracking algorithm,” in *CHEP ’97 Proceedings*. 1997. `http://www.ifh.de/CHEP97/paper/341.ps`.

[32] D. N. Brown, E. A. Charles, and D. A. Roberts, “The BaBar Track Fitting Algorithm,” `http://www.star.bnl.gov/public/comp/reco/babar.ps`.

[33] L. Breiman, J. Friedman, R. Olshen, and C. Stone, *Classification and Regression Trees*. Wadsworth and Brooks, Monterey, CA, 1984.

[34] H.-J. Yang, B. P. Roe, and J. Zhu, “Studies of boosted decision trees for Miniboone particle identification,” *Nucl. Instrum. Meth.* **A555** (2005) 370–385, `arXiv:physics/0508045`.

[35] M. Carpinelli and F. Martinez-Vidal, "Study of K_L^0 identification efficiency using the Emc," *BABAR Analysis Document* (2000) .

[36] P. Billoir, R. Fruhwirth, and M. Regler, "TRACK ELEMENT MERGING STRATEGY AND VERTEX FITTING IN COMPLEX MODULAR DETECTORS," *Nucl. Instrum. Meth.* **A241** (1985) 115–131.

[37] P. Billoir and S. Qian, "Fast vertex fitting with a local parametrization of tracks," *Nucl. Instrum. Meth.* **A311** (1992) 139–150.

[38] D. J. Lange and D. Wright, "Search for the forbidden decay $B^{\pm} \to K^{\pm}\tau\mu$," *BABAR Analysis Document* (2005) .

[39] G. J. Feldman and R. D. Cousins, "A Unified Approach to the Classical Statistical Analysis of Small Signals," *Phys. Rev.* **D57** (1998) 3873–3889, `arXiv:physics/9711021`.