

Search for Di-Muon Decays of a Light Scalar Higgs Boson in Radiative Upsilon(1S) Decays

By Vindhya Prasad

August, 2012

SLAC National Accelerator Laboratory, Stanford University, Stanford CA 94309

Work funded in part by DOE Contract DE-AC02-76SF00515

Search for di-muon decays of a light scalar Higgs boson in radiative $\Upsilon(1S)$ decays

Vindhyawasini Prasad

A thesis
submitted for the degree of
Doctor of Philosophy



Department of Physics
Indian Institute of Technology Guwahati
Guwahati 781039, India

August 2012

Search for di-muon decays of a light scalar Higgs boson in radiative $\Upsilon(1S)$ decays

Vindhyawasini Prasad

A thesis
submitted for the degree of
Doctor of Philosophy

Supervisors:
Dr. Bipul Bhuyan and Dr. Poulouse Poulouse

Department of Physics
Indian Institute of Technology Guwahati
Guwahati 781039, India

August 2012

Certificate

It is certified that the work contained in the thesis entitled “*Search for di-muon decays of a light scalar Higgs boson in radiative $\Upsilon(1S)$ decays*” by Mr. Vindhya wasini Prasad, a Ph.D student of the Department of Physics, IIT Guwahati was carried out under our joint supervision and has not been submitted elsewhere for award of any degree.

Dr. Bipul Bhuyan

Dr. Poulouse Poulouse

*This thesis is dedicated to the memory of my late beloved
younger brother Sri Mata Prasad Mishra*

Acknowledgements

I would like to thank my thesis advisors, Dr. Bipul Bhuyan and Dr. Poulouse Poulouse for their guidance and constant support throughout my research work. They have motivated me to take this work as a challenge and without their help it would have been impossible to finish the same. My other doctoral committee members: Dr. G. S. Setlur, Dr. T. N. Dey and Dr. K. Kapoor also deserve special thanks for their frank comments and encouragements. I am thankful to all the faculty members and staff of the Department of Physics, Indian Institute of Technology Guwahati for their friendly behavior and help, whenever it was needed. My special thank extends to Mr. B. B. Purakayasthya and Mr. Pallav Dutta for helping me to solve the computer related problems on many occasions.

I wish to thank Prof. N. K. Mondal for providing me an opportunity to visit India Based Neutrino Observatory (INO) laboratory at Tata Institute of fundamental Research Center (TIFR), Mumbai, where I have learned different software techniques as well as worked on the LVDS-ECL-NIM translator for the INO data acquisition system (DAQ). These trainings gave me the first glimpse of a high energy physics experiment that kept me motivated throughout the period of this dissertation. I would also like to thank Dr. B. Satyanarayana and other INO colleagues for their kind help and suggestions during my visit of TIFR.

I wish to thank all the members of *BABAR* Tau/QED working group, past and present, especially Dr. Yury Kolomensky, Dr. Randal Sobie, Dr. Bertrand Echenard and Dr. Albarto Lusiani. Their insight and suggestions proved to be very important in finalizing this analysis. I would also like to extend my thanks to the review committee members: Dr. Bertrand Echenard, Dr. Andrew Manual Ruland, Dr. Randal Sobie, publication board chair, Dr. Bill Garry and the *BABAR* Physics Analysis Coordinator, Dr. Abner Sofer for their valued suggestions.

I am thankful to my fellow research group members: Satendra, Deepanwita, Biswajit, Kamal, Deepanjali, Biswaranjan and Nitin for making a pleasant work atmosphere and their assistance time to time. My sincere thank also extends to Meera, Biswanath, Poulami, Jahir, Soumen, Sunita, Sangeetha, Supriya, Jharnali, Manirupa, Parvendra, Arindam, Rahul, Vipin, Niraj, Himanshu and many others colleagues for making my experience somewhat more intellectual and mostly for lots of fun.

Last but not least, I am highly grateful to my parents, Smt. Savitri Mishra and Sri Gulab Dhar Mishra, my brothers, sisters, nephews, relatives and all my well wishers for believing in me, for their constant love and mental support, and inspiring me not only to pursue my Ph.D but also to dedicate myself and my works in the development of basic science throughout the rest of my life.

Abstract

We search for di-muon decays of a low-mass Higgs boson (A^0) in the fully reconstructed decay chain of $\Upsilon(2S, 3S) \rightarrow \pi^+\pi^-\Upsilon(1S)$, $\Upsilon(1S) \rightarrow \gamma A^0$, $A^0 \rightarrow \mu^+\mu^-$. The A^0 is predicted by several extensions of the Standard Model (SM), including the Next-to-Minimal Supersymmetric Standard Model (NMSSM). NMSSM introduces a CP -odd light Higgs boson whose mass could be less than 10 GeV/ c^2 . The data samples used in this analysis contain 92.8×10^6 $\Upsilon(2S)$ and 116.8×10^6 $\Upsilon(3S)$ events collected by the *BABAR* detector. The $\Upsilon(1S)$ sample is selected by tagging the pion pair in the $\Upsilon(2S, 3S) \rightarrow \pi^+\pi^-\Upsilon(1S)$ transitions. We find no evidence for A^0 production and set 90% confidence level (C.L.) upper limits on the product branching fraction $\mathcal{B}(\Upsilon(1S) \rightarrow \gamma A^0) \times \mathcal{B}(A^0 \rightarrow \mu^+\mu^-)$ in the range of $(0.28 - 9.7) \times 10^{-6}$ for $0.212 \leq m_{A^0} \leq 9.20$ GeV/ c^2 . We also combine our results with previous *BABAR* results of $\Upsilon(2S, 3S) \rightarrow \gamma A^0$, $A^0 \rightarrow \mu^+\mu^-$ to set limits on the effective coupling (f_γ) of the b -quark to the A^0 , $f_\gamma^2 \times \mathcal{B}(A^0 \rightarrow \mu^+\mu^-)$, at the level of $(0.29 - 40) \times 10^{-6}$ for $0.212 \leq m_{A^0} \leq 9.2$ GeV/ c^2 .

Contents

Acknowledgements	ix
Abstract	xi
Preface	xv
1 Theoretical & Phenomenological Framework	1
1.1 The Standard Model	1
1.1.1 Gauge Theories	2
1.1.1.1 Gauge theory of electromagnetic interaction	3
1.1.1.2 Gauge theory of strong interaction	3
1.1.1.3 Gauge theory of electroweak interaction	3
1.1.2 Electroweak symmetry breaking in the SM: The Higgs Mechanism	4
1.2 Drawback of the SM	7
1.2.1 Hierarchy problem of the SM	8
1.2.2 Unification	9
1.2.3 Dark Matter	9
1.3 The Minimal Supersymmetric Standard Model	10
1.3.1 The μ problem in MSSM	12
1.4 The Next-to-Minimal Supersymmetric Standard Model	12
1.5 Phenomenology of the light scalar states	14
2 The <i>BABAR</i> experiment	17
2.1 The PEP-II accelerator	18
2.2 The <i>BABAR</i> detector	20
2.2.1 Silicon Vertex Tracker (SVT)	20
2.2.2 Drift Chamber (DCH)	23
2.2.3 The Detector of Internally Reflected Cherenkov Light (DIRC)	26
2.2.4 The Electromagnetic Calorimeter	29
2.2.5 The Instrument Flux Return (IFR)	30
2.2.6 Trigger Selection	32
2.3 Chapter Summary	33

3	Event reconstruction and selection	35
3.1	Data Sets	35
3.2	Event Reconstruction and Event Pre-Selection	37
3.3	Event Selection	40
3.3.1	Pion selection variables	40
3.3.2	Muon selection variables	41
3.3.3	Track multiplicity and photon selection variables	41
3.3.4	Multivariate Analysis	47
3.3.4.1	Variable selection optimization using BumpHunter classifier	48
3.3.4.2	Variable selection optimization using Random Forest classifier	51
3.3.5	Final selection	53
3.4	Corrections of mean and width of m_{recoil}	55
3.5	Chapter Summary	56
4	Maximum Likelihood Fit	59
4.1	Theoretical overview of the ML fit	59
4.1.1	Extended ML Fit	61
4.2	Signal PDF	61
4.3	Background PDF	62
4.3.1	Fit Validation using a cocktail sample	66
4.3.2	Fit validation using Toy Monte-Carlo	69
4.4	Unblinding the $\Upsilon(2S, 3S)$ datasets.	70
4.5	Signal yield extraction using the 1d ML fit	75
4.6	Trial factor study: true significance observation	77
4.7	Chapter Summary	83
5	Systematic Uncertainties	85
5.1	PDF systematics	85
5.2	Fit Bias	86
5.3	Systematic uncertainty for Particle ID	86
5.4	Systematic uncertainty for the charged tracks	87
5.5	Systematic uncertainty for $\Upsilon(2S, 3S)$ kinematic fit χ^2	88
5.6	Systematic uncertainty for $\mathcal{B}(\Upsilon(2S, 3S) \rightarrow \pi^+\pi^-\Upsilon(1S))$	89
5.7	Systematic uncertainty for RF-selection	89
5.8	Systematic uncertainty due to photon selection	90
5.9	Systematic uncertainty for $\Upsilon(nS)$ counting	91
5.9.1	Final systematic uncertainties	91
5.10	Chapter Summary	92
6	Results and Conclusion	93

6.1	Upper-limit	93
6.2	Summary and Conclusion	95
A	Signal PDFs	99
A.1	Signal PDFs for $\mathcal{Y}(2S)$	99
A.2	Signal PDFs for $\mathcal{Y}(3S)$	102
B	Background PDFs	105
B.1	Background PDFs for $\mathcal{Y}(2S)$	105
B.2	Background PDFs for $\mathcal{Y}(3S)$	106
C	Toy Monte Carlo Results	107
C.1	For $\mathcal{Y}(2S)$	107
C.2	For $\mathcal{Y}(3S)$	110
	Bibliography	113

Preface

The Higgs boson is essential to explain the origin of mass of the elementary particles within the Standard Model (SM) via Higgs mechanism through spontaneous breaking of the electroweak symmetry. The Large Hadron Collider (LHC) experiment at CERN has found an evidence of a Higgs-like state which has a mass of $\approx 126 \text{ GeV}/c^2$. However, a light Higgs boson is also predicted by many extensions of the SM including the Next-to-Minimal Supersymmetric Standard Model (NMSSM). The Higgs sector of the NMSSM contains a total three CP -even, two CP -odd and two charged Higgs bosons. The lightest CP -odd Higgs boson (A^0) could have a mass below the $b\bar{b}$ production threshold, avoiding the constraints of Large Electron-Positron (LEP) Collider experiment. Such low-mass Higgs boson can be detected at the B -Factory via radiative $\Upsilon(nS) \rightarrow \gamma A^0$ ($n = 1, 2, 3$) decays. These Υ resonances have narrow width and are produced below the $B\bar{B}$ threshold, providing a clean environment for new physics searches.

In 2005, HyperCP experiment observed three anomalous events in the $\Sigma^+ \rightarrow p\mu^+\mu^-$ final state, that have been interpreted as candidates for CP -odd Higgs with the mass of $214.3 \pm 0.5 \text{ MeV}$ decaying into a pair of muons. In 2008, the CLEO experiment performed a search for A^0 production in the di-tau and di-muon in the final state in the radiative decays of $\Upsilon(1S)$ and ruled out the hyperCP prediction. Similar searches have been performed by $BABAR$ experiment in several final states, including $\Upsilon(2S, 3S) \rightarrow \gamma A^0$, $A^0 \rightarrow \mu^+\mu^-$, and more recently by BESIII experiment in $J/\psi \rightarrow \gamma A^0$, $A^0 \rightarrow \mu^+\mu^-$, and by CMS experiment in $pp \rightarrow A^0$, $A^0 \rightarrow \mu^+\mu^-$. These results have ruled out the hyperCP prediction as well as a substantial fraction of the NMSSM parameter space.

This thesis describes a search for the di-muon decays of the A^0 in the radiative decays of di-pion tagged $\Upsilon(1S)$ meson: $\Upsilon(2S, 3S) \rightarrow \pi^+\pi^-\Upsilon(1S)$, $\Upsilon(1S) \rightarrow \gamma A^0$, $A^0 \rightarrow \mu^+\mu^-$. The data samples used in this analysis were collected at $\Upsilon(2S)$ and $\Upsilon(3S)$ resonances by $BABAR$ detector

at the PEP-II asymmetric-energy e^+e^- collider located at SLAC National Accelerator Laboratory. A clean $\Upsilon(1S)$ sample is selected by tagging the di-pions in the $\Upsilon(2S, 3S) \rightarrow \pi^+\pi^-\Upsilon(1S)$ transition, resulting in a substantial background reduction compared to direct searches in $\Upsilon(2S, 3S) \rightarrow \gamma A^0$ decays. We find no evidence for the A^0 production in the $\Upsilon(2S, 3S)$ data samples, and set 90% C.L. upper limits on the $\mathcal{B}(\Upsilon(1S) \rightarrow \gamma A^0) \times \mathcal{B}(A^0 \rightarrow \mu^+\mu^-)$ for $\Upsilon(2S)$, $\Upsilon(3S)$ and combined data of $\Upsilon(2S, 3S)$ in the mass range of $0.212 \leq m_{A^0} \leq 9.20$ GeV/ c^2 . These results are combined with previous *BABAR* measurements of $\Upsilon(2S, 3S) \rightarrow \gamma A^0$, $A^0 \rightarrow \mu^+\mu^-$ to set limits on effective Yukawa coupling of bound b -quark to the A^0 . The results of this analysis have been published in *Phys. Rev. D* **87**, 031102 (R) (2013), [arXiv:1210.0287].

This thesis is organized in six chapters as discussed below:

Chapter 1 gives an overview of the SM and its limitations, and describes theoretically the most attractive replacement – Supersymmetry. The Minimal Supersymmetric Standard Model (MSSM) solves the hierarchy problem of the SM, but fails to explain why μ -parameter is of the order of electroweak scale which is so far from the next natural scale – the Planck scale. The NMSSM solves this problem while generating a μ -term and introduces an extra CP -even and CP -odd light Higgs bosons. Finally, this chapter reviews the phenomenology of the A^0 .

Chapter 2 provides a short description of the PEP-II electron-positron collider and the *BABAR* detector, which collected the $\Upsilon(2S, 3S)$ datasets for this analysis.

Chapter 3 describes the datasets used in this analysis, and the reconstruction of the $\Upsilon(2S, 3S)$ decay chains: $\Upsilon(2S, 3S) \rightarrow \pi^+\pi^-\Upsilon(1S)$, $\Upsilon(1S) \rightarrow \gamma A^0$, $A^0 \rightarrow \mu^+\mu^-$. It describes the discriminative variables used to separate the signal from background. Monte Carlo (MC) simulated events are used to study the detector acceptance and optimize the event selection criteria. A blind analysis technique is used in this work, where the full data samples are kept blind until all the selection criteria are finalized. A random forest (RF) classifier is used to improve the purity of $\Upsilon(1S)$ events from $\Upsilon(2S, 3S) \rightarrow \pi^+\pi^-\Upsilon(1S)$ transitions. Finally, It estimates the remaining backgrounds after applying all the selection criteria.

Chapter 4 discusses the signal and background probability density functions (PDFs), which are used to extract the signal from data. The fit procedure is validated by using a cocktail sample of $\Upsilon(2S, 3S)$ generic and 5% of $\Upsilon(2S, 3S)$ onpeak datasets, as well as a large number of Toy MC datasets. The full data sample of $\Upsilon(2S, 3S)$ are unblinded after finalizing

all the selection criteria and the ML fitting procedure. The signal yields are extracted using the unblinded $\Upsilon(2S, 3S)$ datasets. We also describe a trial factor study used to compute the true significance, i.e. the probability for pure background sample to fluctuate up to a given value of the signal yield.

Chapter 5 describes the possible systematic uncertainties and their sources for this analysis.

Chapter 6 presents the 90% confidence level (CL) Bayesian upper limits on the product branching fraction of $\mathcal{B}(\Upsilon(1S) \rightarrow \gamma A^0) \times \mathcal{B}(A^0 \rightarrow \mu^+ \mu^-)$ as a function of m_{A^0} , including the systematic uncertainties. The combined upper limits of this result with previous *BABAR* results of $\Upsilon(2S, 3S) \rightarrow \gamma A^0$, $A^0 \rightarrow \mu^+ \mu^-$ are also presented. Finally, we present a summary of the results and a brief conclusion.

Chapter 1

Theoretical & Phenomenological Framework

This chapter begins with an overview of the Standard Model (SM) of particle physics, including the Higgs mechanism which breaks the electroweak symmetry spontaneously in the model and provides masses to the W^\pm and Z^0 gauge bosons and the fermions. Section 1.2 reviews some limitations of this model and describes one of the possible theoretically attractive replacements – supersymmetry. The Minimal supersymmetric Standard Model (MSSM) solves the hierarchy problem of the SM, but fails to explain why the value of the μ -parameter is of the order of the electroweak scale, which is so far from the next natural scale – the Planck scale. The Next-to-Minimal Supersymmetric Standard Model (NMSSM) cures this problem and predicts a CP -odd light Higgs boson whose mass could be less than twice the mass of the b -quark. Finally, section 1.5 reviews some phenomenology related to the light scalar Higgs boson.

1.1 The Standard Model

The SM of Particle Physics describes all the known fundamental particles and their interactions [1–5]. It is a well-established theory, which has passed all the scrutiny by the high energy collider and precision experiments so far. Within this model, all the known matter is composed of spin-1/2 fermion constituents: the leptons and the quarks. There are six types

of lepton flavors forming three generations, which are called electron (e), muon (μ) and tau (τ) with electric charge $Q = -1$ (in the unit of the elementary charge of e), and the corresponding neutrinos ν_e , ν_μ and ν_τ with $Q = 0$. The quarks also comes in six different flavors: up (u), down (d), charm (c), strange (s), top (t) and bottom (b), and have fractional charges $Q = +\frac{2}{3}, -\frac{1}{3}, +\frac{2}{3}, -\frac{1}{3}, +\frac{2}{3}$ and $-\frac{1}{3}$, respectively. These fermions interact with each other via exchange of gauge bosons of integral spin-1. The gauge fields in the SM describe the three interactions: the electromagnetic interaction, the strong interaction and the weak interaction. The electromagnetic interaction is mediated by the photon (γ), the weak interaction is mediated by the weak vector bosons W^\pm and Z^0 , and strong interaction is mediated by the eight gluons (g_i). The gravity is not incorporated by the SM, because it is very weak compared to other interactions.

The fermions and the gauge bosons acquire mass via Higgs mechanism [6–9] through spontaneous breaking of electroweak symmetry, $SU(3)_C \times SU(2)_L \times U(1)_Y \rightarrow SU(3)_C \times U(1)_{em}$. In addition to providing the masses to the fermions and the W^\pm and Z^0 gauge bosons, the Higgs mechanism predicts an additional electrically neutral scalar Higgs boson. A Higgs like state has recently been discovered by the CMS and ATLAS experiments at CERN, and its mass is measured to be $126 \text{ GeV}/c^2$ [10].

1.1.1 Gauge Theories

The gauge theory is a special class of quantum field theory that introduces an invariance principle used to describe the interaction among all the fundamental constituents of matter. The interactions between the fundamental particles are dictated by symmetry principles, which are intimately connected with the ideas of conserved physical quantities. The connection between symmetries and conservation laws is described in the framework of Lagrangian field theory. The gauge symmetry of a physical system is realized through the invariance of the Lagrangian under gauge transformations, which are characterized by Lie group. Global-invariance (phase invariance) under gauge transformation leads to a conserved charge. The local gauge invariance (space-times dependence of parameter of the system) of the Lagrangian introduces a vector field, called gauge field, which governs the interaction. The quanta of the gauge fields are the gauge bosons mediating the interactions.

1.1.1.1 Gauge theory of electromagnetic interaction

The electromagnetic interaction is described by quantum electrodynamics [11]. The global invariance of $U(1)$ in the QED introduces the conservation of the electric charge (Q). The local gauge invariance of the gauge theory gives rise to the gauge field corresponding to a massless gauge boson (photon (γ)), which describes the interactions among the fundamental charged particles. The coupling constant (α) describes strength of the interaction between the photon and the fermions. However, α is a function of energy when quantum correction are considered. At low energy, the α is given by the fine structure constant, $\alpha = e^2/4\pi\hbar c = 1/137$. Due to the abelian nature of the $U(1)$ symmetry group, photon is charge-less, and do not interact with each other directly. The electromagnetic interaction is a long range interaction.

1.1.1.2 Gauge theory of strong interaction

The strong interaction is described by quantum chromodynamics (QCD) [12]. The symmetry group of QCD is $SU(3)_C$, where C refers to colour and 3 refers to the three possible colour states of the quarks, normally termed as red, green and blue. Colour symmetry is exact, so QCD calculations are independent of the colour of the quarks. For example, probability of a red quark scattering off a green quark is the same as the probability of a red quark scattering off a blue quark. The local gauge invariance of $SU(3)_C$ gives rise to eight types of the gluonic fields. QCD is a non-abelian theory, where the gluons carry both colour and anti-colour, in contrast to the photon in QED which does not carry the electric charge. Gluons interact with each other directly and as strongly as they do with quarks. Due to this gluon-gluon interaction, the strong force increases with distance resulting in confinement of quarks. This means, the quarks do not exist freely in Nature, but bind together by the strong force and form the mesons ($q\bar{q}$) and the baryons (qqq), where q stands for a quark and \bar{q} stands for an anti-quark.

1.1.1.3 Gauge theory of electroweak interaction

The electromagnetic and weak interactions are combined in an $SU(2)_L \times U(1)_Y$ gauge theory of electroweak interaction, developed by Glashow, Weinberg and Salam [1, 13]. The subscript “ L ” indicates that only the left-handed (right-handed) components of the fermion

(antifermion) fields take part in weak interactions. The fermions appear as left-handed doublets and right-handed singlets under the $SU(2)_L$. Global gauge invariance under the $SU(2)_L$ gauge transformation leads to the conservation of the weak-isospin, T . Requiring the local $SU(2)_L$ gauge invariance of the Lagrangian of the system introduces a weak-isospin triplet of the gauge fields, W_μ^i , $i = 1, 2, 3$. The $SU(2)_L$ is a non-abelian group which leads to the self-interactions of the gauge fields. The global gauge invariance under the $U(1)_Y$ transformation leads to the conservation of weak-hypercharge, Y . However, the local gauge invariance of $U(1)_Y$ introduces vector gauge field, B_μ . The weak-hypercharge, Y , third component of weak-isospin, T_3 , and electric charge, Q are related by the Gell-Mann-Nishijima relation:

$$Q = T_3 + \frac{1}{2}Y. \quad (1.1)$$

1.1.2 Electroweak symmetry breaking in the SM: The Higgs Mechanism

The gauge invariance of $SU(2)_L \times U(1)_Y$ requires massless gauge bosons, since the presence of a mass term for the gauge boson violates gauge invariance ($M^2 A_\mu A^\mu$ is not invariant under $A_\mu \rightarrow A_\mu - \partial_\mu \chi$, where χ is a function of position in space time. So M^2 must be zero in a gauge symmetric Lagrangian). This difficulty is circumvented by the Higgs mechanism through which electroweak symmetry breaking is achieved spontaneously [6]. The SM contains a weak-isospin $SU(2)_L$ doublet of complex scalar Higgs fields (with weak-hypercharge $Y=1$),

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \quad (1.2)$$

in the Lagrangian of the system. The most general renormalizable and $SU(2)_L \times U(1)_Y$ invariant Lagrangian allowed, involving only the gauge bosons and scalar fields is given by

$$\mathcal{L} = -\frac{1}{4}W_{\mu\nu}^i W^{\mu\nu i} - \frac{1}{4}B_{\mu\nu} B^{\mu\nu} + (D_\mu \Phi)^\dagger (D^\mu \Phi) - V(\Phi), \quad (1.3)$$

$$W_{\mu\nu}^i = \partial_\mu W_\nu^i - \partial_\nu W_\mu^i - g\epsilon^{ijk}W_\mu^j W_\nu^k,$$

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu,$$

$$D_\mu = \partial_\mu + \frac{1}{2}ig\tau^i W_\mu^i + \frac{1}{2}ig'YB_\mu, \quad (1.4)$$

where W_μ^i ($i = 1, 2, 3$) are the three massless $SU(2)_L$ gauge bosons, B_μ the massless $U(1)_Y$ gauge boson, and the scalar potential is given by

$$V(\Phi) = \mu^2|\Phi^\dagger\Phi| + \lambda|\Phi^\dagger\Phi|^2, \quad (1.5)$$

here g and g' are the gauge coupling constants of $SU(2)_L$ and $U(1)_Y$, respectively. For a choice of $\lambda > 0$ and $\mu^2 < 0$, the state of minimum energy for the potential V is not at zero, but at $|\Phi^\dagger\Phi| = -\mu^2/2\lambda \equiv v^2/2$ (Figure 1.1). The scalar field thus develops a non-vanishing vacuum expectation value (VEV), which is degenerate. A single value of the VEV can be chosen, which is essentially a choice of a preferred “direction” in the Higgs-doublet phase space. The usual choice is

$$\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}, \quad (1.6)$$

where $H(x)$ is a physical scalar field. The choice of this new ground state “spontaneously” breaks the $SU(2)_L \times U(1)_Y$ symmetries to $U(1)_{EM}$, while maintaining the renormalizability and unitarity of the theory. As the $U(1)_Y$ gauge symmetry remains unbroken in this transformation, the associated gauge boson, the photon, remains massless. However, three of the degrees of freedom of the scalar doublet (corresponding to Goldstone bosons) are “eaten by” or transformed into the longitudinal polarization components of the weak-isospin triplet of bosons, giving the W^\pm and Z^0 bosons their masses of $M_W = \frac{1}{2}vg$ and $M_{Z^0} = \frac{1}{2}v(g^2 + g'^2)^{1/2}$, respectively [1]. The mass eigenstates are expressed in terms of the gauge eigenstates as below:

$$\begin{aligned} W_\mu^\pm &= \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2), \\ Z_\mu^0 &= W_\mu^3 \cos\theta_W - B_\mu \sin\theta_W, \\ A_\mu &= W_\mu^3 \sin\theta_W + B_\mu \cos\theta_W, \end{aligned} \quad (1.7)$$

where A_μ is the gauge field of the electromagnetic interaction, and θ_W is the Weinberg mixing angle. The remaining degree of freedom corresponds to a massive neutral scalar particle, the Higgs boson, H^0 . The mass of this scalar is given by $m_{H^0}^2 = 2v^2\lambda$.

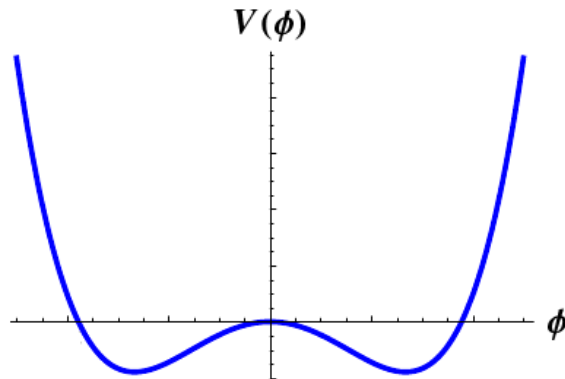


Figure 1.1: One dimensional projection of Higgs potential ($V(\phi)$) as a function of scalar field (ϕ). The ($V(\phi)$) develops a vacuum expectation value (VEV) at $\phi = 0$ when $\mu^2 < 0$.

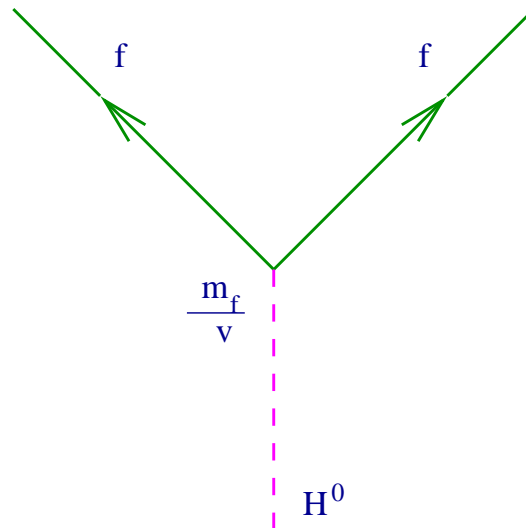


Figure 1.2: The Yukawa coupling of the Standard Model Higgs boson to the fermions.

The Higgs mechanism also provides masses to the quarks and leptons while including the following $SU(2)_L \times U(1)_Y$ gauge invariant terms for the first generation of leptons and quarks in the Lagrangian

$$\mathcal{L}_{\text{Lepton}} = -g_e \left[\begin{pmatrix} \bar{\nu}_e & \bar{e} \end{pmatrix}_L \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} e_R + \bar{e}_R \begin{pmatrix} \phi^- & \bar{\phi}^0 \end{pmatrix} \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L \right],$$

$$\mathcal{L}_{\text{Quark}} = -g_d \begin{pmatrix} \bar{u} & \bar{d} \end{pmatrix}_L \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} d_R - g_u \begin{pmatrix} \bar{u} & \bar{d} \end{pmatrix}_L \begin{pmatrix} -\phi^0 \\ \phi^- \end{pmatrix} u_R + \text{hermitian conjugate (h.c)}. \quad (1.8)$$

Here $\begin{pmatrix} -\phi^0 \\ \phi^- \end{pmatrix} = -i\tau_2 \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}^*$, where $\tau_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ is the isospin version of the Pauli matrix. Second and third generations of leptons and quarks have similar expressions. After breaking the symmetry spontaneously as discussed above, the Higgs scalar picks up a vacuum expectation value given by equation 1.6. This will generate the mass term of the fermion, and an interaction term with the Higgs particle

$$\mathcal{L}_{\text{Lepton}} = -m_e \bar{e}e - \frac{m_e}{v} \bar{e}eH^0,$$

$$\mathcal{L}_{\text{Quark}} = -m_d \bar{d}d - m_u \bar{u}u - \frac{m_d}{v} \bar{d}dH^0 - \frac{m_u}{v} \bar{u}uH^0, \quad (1.9)$$

with the identification $m_i = g_i v / \sqrt{2}$, where m_i is the mass of each fermion i , g_e , g_u and g_d are the Yukawa coupling constants (Figure 1.2). Thus, the strength of Higgs boson couplings to fermions is proportional to the corresponding particle masses.

1.2 Drawback of the SM

The SM is the result of many experimental observations and progress in the theoretical understanding of Nature. Most of the theoretical results of the SM agree with the experimental data. However, the SM can not be quantified as a “theory of everything”. There is no method to incorporate gravity which becomes important at energy scales approaching the Planck scale ($M_{\text{Planck}} = (8\pi G_N)^{-1/2} \sim 2.4 \times 10^{18} \text{ GeV}/c^2$) and so the SM must be considered as an effective theory at energies below this scale. Some of the important drawbacks of the SM and their possible solutions are described bellow:

1.2.1 Hierarchy problem of the SM

The mass of the SM Higgs boson is expected to be of the order of electroweak scale ($\sim O(M_W)$). The self-coupling effects in the scalar Higgs field involving higher-order fermionic loops are quadratically divergent (Figure 1.3(a)). A cut-off scale Λ_{cutoff} on the momentum integral can be introduced to prevent these radiative correction from going to infinity. The Higgs couples with fermion pair via an Yukawa interaction term of $-g_f \bar{H} f f$ in the Lagrangian. At one loop each fermion contributes a correction of mass term, which is [14]

$$\Delta m_{H^0,f}^2 = \frac{g_f^2}{8\pi^2} \left[-\Lambda_{cutoff}^2 + 6m_f^2 \ln\left(\frac{\Lambda_{cutoff}}{m_f}\right) \right]. \quad (1.10)$$

These corrections blow up as $\Lambda_{cutoff} \rightarrow \infty$. To explain the $m_{H^0}^2 \sim O(M_W)$ we need either $\Lambda_{cutoff} \lesssim 1$ TeV, or extreme fine tuning (adjusting the value of g_f accordingly) so that the correction is of the electroweak scale. This difficulty is known as hierarchy problem of the SM. Supersymmetric extension of the SM solves the hierarchy problem of the SM while introducing the superpartners of each fundamental particles that differ by half integral-spin [15]. The superpartners of the fermions also couple to the Higgs by a quartic interaction of the form $-g_S |H^0|^2 |S|^2$, and thereby contribute to the Higgs mass corrections through loops as shown in Figure 1.3(b). The loop correction contributes to the Higgs mass by:

$$\Delta m_{H^0,S}^2 = \frac{g_S}{16\pi^2} \left[\Lambda_{cutoff}^2 - 2m_S^2 \ln\left(\frac{\Lambda_{cutoff}}{m_S}\right) \right]. \quad (1.11)$$

It is seen from equation 1.10 and 1.11 that if every fermion is accompanied by a scalars with coupling $g_S = 2g_f^2$, the quadratic divergences cancel exactly. After adding the equation 1.10 and 1.11, the total correction is reduced to

$$\Delta m_{H^0,Tot}^2 \simeq \frac{g_f^2}{4\pi^2} (m_S^2 - m_f^2) \ln\left(\frac{\Lambda_{cutoff}}{m_S}\right). \quad (1.12)$$

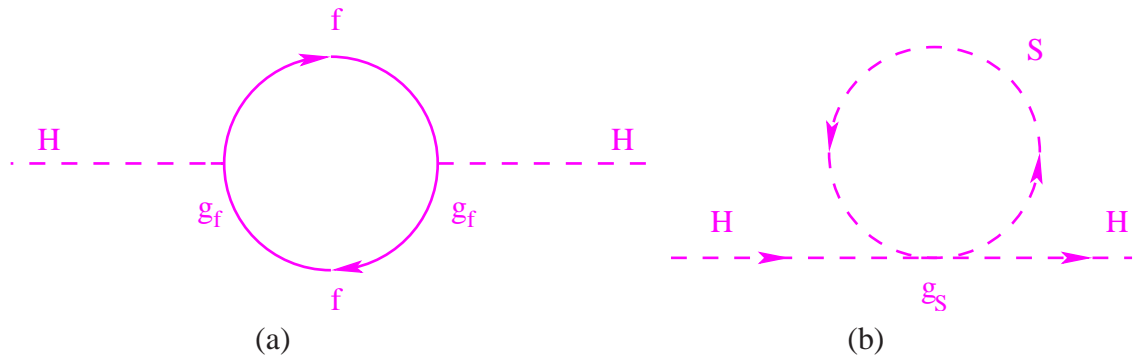


Figure 1.3: Loops affecting the squared Higgs mass from (a) fermions trilinear couplings and (b) scalar quartic couplings.

1.2.2 Unification

The SM unifies the electromagnetic and weak interactions at the electroweak scale. This motivation can also be extended to the grand unification scale [16, 17] where the electromagnetic, weak and strong interactions are unified together [18] through the supersymmetric extensions of the SM.

1.2.3 Dark Matter

There is ample evidence from observation like the rotation curve of galaxies that luminous matter in the universe accounts for only a small fraction of the total matter-energy density. The unknown matter content of the universe is called the dark matter (DM) [19–21]. It accounts for about 23% of the total matter density of the observable universe, while the ordinary matter accounts for only 4.6%, with the remainder being attributed to dark energy. The SM does not have viable candidate for DM particles. However, there are extension of SM including supersymmetric models which contain viable candidates for DM. We should mention that, such models do not explain the existence of dark energy, which contributes to about 73% of the total energy of the universe.

1.3 The Minimal Supersymmetric Standard Model

As we see in the last section that one of the best motivated extensions of the SM of particle physics is the introduction of Supersymmetry (SUSY) [14, 15, 22]. The SUSY solves the hierarchy problem of the SM and unifies the three forces of electromagnetic, weak and strong at the Grand unified scale. The Minimal Supersymmetric Standard Model (MSSM) is a minimal supersymmetric extension of the SM [23]. It is also based on the gauge symmetry of $SU(3)_C \times SU(2)_L \times U(1)_Y$. It transforms bosonic states into fermionic states and vice versa via an operator Q

$$Q|\text{Boson}\rangle = |\text{Fermion}\rangle \quad Q|\text{Fermion}\rangle = |\text{Boson}\rangle. \quad (1.13)$$

If the Q and its hermition conjugate Q^\dagger hold following commutation relations

$$\begin{aligned} \{Q, Q^\dagger\} &= P^\mu, \\ \{Q, Q\} &= \{Q^\dagger, Q^\dagger\} = 0, \\ [P^\mu, Q] &= [P^\mu, Q^\dagger] = 0, \end{aligned} \quad (1.14)$$

then the theory is able to describe the chiral fermions as they are observed in Nature [15]. Here, P^μ is the four-momentum, which is the generator of space-time translations. Irreducible representations of such types of algebra are called supermultiplets and describe the single particle state. A supermultiplet includes an equal number of fermionic (n_F) and bosonic (n_B) degrees of freedom, which means that every SM particle has their own superpartner, which has the same quantum numbers except their spin which differ by 1/2. The superpartner of fermions are scalar particles called sfermions, that of gauge boson are spin-1/2 particles called gauginos, and that of the Higgs bosons are spin-1/2 particles called Higgsinos.

A gauge or vector supermultiplet contains a massless vector boson ($n_B = 2$) and the superpartner of this boson, a spin-1/2 Weyl fermion ($n_F = 2$). The Weyl fermion does not have its own antiparticles in contrast to the Majorana fermion that is its own antiparticle. The known SM gauge bosons and the corresponding gauginos are contained in vector supermultiplet in the MSSM.

A chiral supermultiplet contains a spin-1/2 Weyl fermion ($n_F = 2$) and two real scalars (each $n_B = 1$, spin 0), which can be described by a complex scalar field. The Higgs bosons, Higgsinos, and spin-1/2 fermions and sfermions are part of such chiral superfields in the MSSM.

The MSSM postulates two complex $SU(2)_L$ doublet scalar superfields, denoted by \hat{H}_u and \hat{H}_d with weak-hypercharges $Y = \pm 1$:

$$\hat{H}_u = \begin{pmatrix} \hat{H}_u^+ \\ \hat{H}_u^0 \end{pmatrix}, \quad \hat{H}_d = \begin{pmatrix} \hat{H}_d^0 \\ \hat{H}_d^- \end{pmatrix}. \quad (1.15)$$

The superpotential of the MSSM involving the Higgs fields is given by [15]:

$$W = (g_u)^{ij} \hat{u}_i \hat{Q}_j \cdot \hat{H}_u - (g_d)^{ij} \hat{d}_i \hat{Q}_j \cdot \hat{H}_d - g_e^{ij} \hat{e}_i \hat{L}_j \cdot \hat{H}_d + \mu \hat{H}_u \cdot \hat{H}_d, \quad (1.16)$$

where the labels i, j are family indexes of quarks and leptons. The g_u, g_d and g_e are the Yukawa coupling constants of up-type quarks, down-type quarks and leptons, respectively. The μ -term mixes the two Higgs superfields.

The gauge-invariant Higgs scalar potential built from the two Higgs doublets in Equation 1.15 is consistent with the electroweak sector of the SM and spontaneously breaks $SU(2)_L \times U(1)_Y$ down to $U(1)_{EM}$. It is given by:

$$V = \frac{1}{8}(g^2 + g'^2)(|H_d|^2 - |H_u|^2)^2 + \frac{1}{2}g^2 |H_d^\dagger H_u|^2 + \mu^2(|H_d|^2 + |H_u|^2) + m_{H_d}^2 |H_d|^2 + m_{H_u}^2 |H_u|^2 + \mu B(H_u \cdot H_d + h.c.), \quad (1.17)$$

where $m_{H_d}^2, m_{H_u}^2$ and B are soft supersymmetry breaking parameters. This scalar potential is minimized by the vacuum expectation values (VEV's) of the Higgs fields

$$\langle H_d \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} v_1 \\ 0 \end{pmatrix}, \quad \langle H_u \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v_2 \end{pmatrix}, \quad (1.18)$$

which spontaneously breaks the electroweak symmetry, $SU(2)_L \times U(1)_Y \rightarrow U(1)_{EM}$. A conventional notation is used to relate the two VEVs by $\tan\beta = v_1/v_2$. The two VEVs can then be defined as $v_1 = \langle H_d \rangle = v \sin\beta$ and $v_2 = \langle H_u \rangle = v \cos\beta$ (where $v = \sqrt{v_1^2 + v_2^2} = 2m_W/g \simeq 246$

GeV). The physical MSSM Higgs sector consists of two neutral CP -even (H^0 and h^0), a neutral CP -odd (A) and a pair of charged (H^\pm) Higgs bosons. The MSSM also contains four neutralinos ($\tilde{\chi}_{1,2,3,4}^0$), among which the $\tilde{\chi}_1^0$ is the lightest supersymmetric particle (LSP) in the R-parity conserving model, and is a viable candidate of DM.

1.3.1 The μ problem in MSSM

The MSSM superpotential (Equation 1.16) contains a μ -term, which mixes the \hat{H}_u and \hat{H}_d chiral superfields, is the only dimensional coupling in the superpotential. The value of μ is expected to be of the order of electroweak scale, which is many orders of magnitude smaller than the next natural scale, the Planck scale. A possible solution for this problem can be found in the framework of the Next-to-Minimal Supersymmetric Standard Model (NMSSM).

1.4 The Next-to-Minimal Supersymmetric Standard Model

The NMSSM adds a singlet chiral superfield (\hat{N}) to the MSSM [24–26]. The superpotential of the NMSSM contains a trilinear term along with an \hat{N}^3 term instead of the μ -term of MSSM superpotential in Equation 1.16, basically defined as

$$W_{NMSSM} = (\dots) + \lambda \hat{N} \hat{H}_u \cdot \hat{H}_d + \frac{\kappa}{3} \hat{N}^3, \quad (1.19)$$

where λ and κ are dimensionless Yukawa couplings, and \hat{H}_u and \hat{H}_d are up and down types of Higgs superfields. The associated soft terms, which break the supersymmetry explicitly, are $\lambda A_\lambda N H_u H_d + \frac{1}{3} A_\kappa N^3$. In the presence of these soft supersymmetry breaking terms, a vacuum expectation value (VEV) of N , which is of the order of electroweak scale generates an effective μ -term with $\mu_{eff} = \lambda \langle N \rangle$, which solves the ‘ μ -problem’ of the MSSM [27]. As a result, the NMSSM Higgs sector contains a total of three CP -even, two CP -odd and two charged Higgs bosons. This model also contains a total of five neutral fermionic states, $\tilde{\chi}_{1,2,3,4,5}^0$, which are LSP (in the R-parity model) and viable candidates of DM. The Higgs sector of the NMSSM contains six independent parameters:

$$\lambda, \kappa, A_\lambda, A_\kappa, \tan\beta, \mu_{eff}, \quad (1.20)$$

where the sign conventions for the fields λ and $\tan\beta$ should be always positive, while κ , A_λ , A_κ and μ_{eff} may have either sign.

The mass of the lightest CP -odd Higgs boson (A^0) is controlled by the soft-trilinear coupling A_λ and A_κ and vanishes in the Peccei-Quinn symmetry limit, $\kappa \rightarrow 0$ [28], or a global $U(1)_R$ symmetry in the limit of vanishing soft term, $A_\lambda, A_\kappa \rightarrow 0$, which is spontaneously broken by the VEVs, resulting in a Nambu-Goldstone boson in the spectrum [29]. This symmetry is explicitly broken by the trilinear soft terms so that the A^0 is naturally small. In a generic case, the fermion coupling to the light pseudoscalar A^0 field can be defined by an interaction term:

$$\mathcal{L}^{f\bar{f}A^0} = -X_f \frac{m_f}{v} A^0 \bar{f}(i\gamma_5)f, \quad (1.21)$$

where X_f is the coupling constant, which depends on the type of fermion with a mass m_f [30, 31]. In the NMSSM, $X_d = \cos\theta_A \tan\beta$ for the down-type fermion pair and $X_u = \cos\theta_A \cot\beta$ for the up-type fermion pair, where θ_A is the mixing angle between the singlet (A_S) component and MSSM like doublet component (A_{MSSM}) of the A^0 . With this mixing angle, the lighter CP -odd state of the A^0 is defined as:

$$A^0 = \cos\theta_A A_{MSSM} + \sin\theta_A A_S. \quad (1.22)$$

Such light state of the A^0 is not excluded by the LEP constraints [32], where the CP -even Higgs boson, h , could decay dominantly into a pair of CP -odd scalars [25, 29, 33–37]. The LEP experiment has also excluded a SM-like h decaying to $b\bar{b}$ for $m_h < 114$ GeV/ c^2 and placed a strong constraints on $e^+e^- \rightarrow Zh \rightarrow Zb\bar{b}$ as well as the effective coupling of $C_{eff}^2 \equiv [g_{ZZh}^2/g_{ZZh_{SM}}^2] \mathcal{B}(h \rightarrow b\bar{b})$ [34]. The Large hadron collider (LHC) experiment will also not be able to discover such scalar states if h decays primarily into a pair of CP -odd scalars with m_A^0 bellow the $B\bar{B}$ threshold [33, 36, 37]. In this case, the A^0 can be accessible via the Υ decays [30, 31, 38–42] while using the large datasets of the current generation of B -Factories, such as *BABAR*, *CLEO* and *Belle* experiments.

1.5 Phenomenology of the light scalar states

The lightest state of the A^0 in the NMSSM is constrained to have the mass below the $b\bar{b}$ threshold, $2m_b$ [26], to avoid the detection at LEP. A pseudo-scalar axion having a mass around $360\text{--}800\text{ MeV}/c^2$ and decaying into a lepton pair with a Higgs-like coupling is also predicted by models motivated by astrophysical observations [43]. The low mass Higgs boson could explain the origin of mass of the light elementary particles, the mystery related to the Dark matter and Dark energy which contributes more than 90% matter density of the universe. In the framework of dark matter, the dark matter particles can annihilate into pairs of the dark photons, which subsequently decay to SM particles. In a minimal model [44], the dark photon mass is generated via the Higgs mechanism, adding a dark Higgs boson in the theory. The mass hierarchy between dark photon and dark Higgs boson is not constrained experimentally, so the dark Higgs boson could be light as well [45]. These light scalar states could be within the reach of present particle accelerators, such as the B -Factory at SLAC.

The branching fractions of $\mathcal{B}(\Upsilon(nS) \rightarrow \gamma A^0)$ ($n = 1, 2, 3$) are related to the effective Yukawa coupling (f_Υ) of the b -quark to the A^0 through [46–48]:

$$\frac{\mathcal{B}(\Upsilon(nS) \rightarrow \gamma A^0)}{\mathcal{B}(\Upsilon(nS) \rightarrow l^+ l^-)} = \frac{f_\Upsilon^2}{2\pi\alpha} \left(1 - \frac{m_{A^0}^2}{m_{\Upsilon(nS)}^2} \right), \quad (1.23)$$

where $l \equiv e$ or μ and α is the running fine structure constant. In the SM, the value of f_Υ is defined as:

$$f_{\Upsilon,SM}^2 = \sqrt{2} G_F m_b^2 C_{QCD} \approx (2 - 3) \times 10^{-4}, \quad (1.24)$$

where $C_{QCD} \approx 0.7 - 1.0$ [49] includes the QCD loop corrections and relativistic corrections to $\mathcal{B}(\Upsilon(nS) \rightarrow \gamma A^0)$ [48], as well as the leptonic width of $\Upsilon(nS) \rightarrow l^+ l^-$ [50]. However, the coupling of bound b -quark to the A^0 in the NMSSM is $f_{\Upsilon,NMSSM}^2 = \sqrt{2} G_F m_b^2 X_d^2 C_{QCD}$. The Yukawa coupling also depends upon the axion constant f_a in the axion model of Nomura and Thaler [43]. A study of the NMSSM parameter space predicts the branching fraction of $\Upsilon(1S) \rightarrow \gamma A^0$ to be in the range of $10^{-6} - 10^{-4}$ depending upon the A^0 mass, $\tan\beta$ and $\cos\theta_A$ [41].

In the SM, interactions between the leptons and gauge bosons are same for all the lepton flavors, and therefore the quantity $R_{ll'} = \Gamma_{\Upsilon(1S) \rightarrow ll} / \Gamma_{\Upsilon(1S) \rightarrow l'l'}$ with $l, l' = e, \mu, \tau$ and $l' \neq l$,

is expected to be close to one. In the NMSSM, any significant deviations of $R_{ll'}$ from unity would violate lepton universality, which may arise due to presence of the A^0 that couples to the $\Upsilon(1S)$. *BABAR* has measured the value of $R_{\tau\mu}(\Upsilon(1S)) = 1.005 \pm 0.013(\text{stat}) \pm 0.022(\text{syst})$ using a sample of $(121.8 \pm 1.2) \times 10^6$ $\Upsilon(3S)$ events, which shows no significant deviation from the expected SM value [51]. If the light CP -odd Higgs boson A^0 has a mass in the range of $9.2 < m_{A^0} < 12$ GeV/ c^2 , the NMSSM can account for the anomalous muon magnetic moment [52].

For large value of $\tan\beta$, the A^0 will primarily decay to heavier down-type fermion that is kinematically available. The branching fractions of $A^0 \rightarrow f\bar{f}$ as a function of $\tan\beta$ and m_{A^0} are summarized in [53]. The same reference [53] also summarizes the expected $\mathcal{B}(\Upsilon(3S) \rightarrow \gamma A^0)$ for various SUSY model parameters, with the constraint that the model does not require the ‘‘fine tuning’’ [34]. *BABAR* has previously searched for A^0 production in the radiative decays of $\Upsilon(nS) \rightarrow \gamma A^0$ with $n = 1, 2, 3$, where the A^0 decays to muons [54], taus [55], invisible [56, 57], or hadrons [58]. Similar searches have also been performed by CLEO in the di-muon and di-tau final states in radiative $\Upsilon(1S)$ decays [59], and more recently by BESIII in the decay chain of $J/\psi \rightarrow \gamma A^0$, $A^0 \rightarrow \mu^+\mu^-$ [60], and by CMS experiment in $pp \rightarrow A^0$, $A^0 \rightarrow \mu^+\mu^-$ [61]. *BABAR* results [54] for $A^0 \rightarrow \mu^+\mu^-$ decay rules out approximately 80% of the NMSSM parameter space in the $m_{A^0} < 2m_\tau$ range at $\tan\beta = 3$. Reference [62] interprets the *BABAR* [54] and CLEO [59] results in terms of the limit of X_d as a function of m_{A^0} and predicts that these results fit with an approximate limit of $X_d < 0.5$ for $\tan\beta = 5$.

This thesis describes a search for a di-muon resonance in the fully reconstructed decay chain of $\Upsilon(2S, 3S) \rightarrow \pi^+\pi^-\Upsilon(1S)$, $\Upsilon(1S) \rightarrow \gamma A^0$, $A^0 \rightarrow \mu^+\mu^-$. This search is based on a sample of $(92.8 \pm 0.8) \times 10^6$ $\Upsilon(2S)$ and $(116.8 \pm 1.0) \times 10^6$ $\Upsilon(3S)$ mesons collected with the *BABAR* detector at the PEP-II asymmetric-energy e^+e^- collider located at SLAC National Accelerator Laboratory. A sample of $\Upsilon(1S)$ mesons is selected by tagging the di-pion transition, which results in a substantial background reduction compared to direct searches of A^0 in $\Upsilon(2S, 3S) \rightarrow \gamma A^0$ decays. We assume that the A^0 is a scalar or pseudo-scalar particle with a negligible decay width compared to the experimental resolution.

Chapter 2

The *BABAR* experiment

BABAR is a high luminosity e^+e^- asymmetric energy collider experiment located at SLAC National Accelerator Laboratory, California, USA. It was primarily designed to study the CP -violation in B -meson decays, and therefore, for most of its run period, the experiment was operated at the e^+e^- center of mass (CM) energy corresponding to the $\Upsilon(4S)$ resonance, which is just above the $B\bar{B}$ threshold. This has allowed the *BABAR* to perform precision measurements of the B meson decays, probing deeply into the phenomena of CP -violation and thereby establishing the CKM formalism [63] of the SM. Despite its initial goal of the study of CP -violation in B -meson decays, the *BABAR* experiment has also carried out significant studies in many other fields of high energy physics such as: τ physics, physics of the heavy quarks, decays of the D-mesons and physics beyond SM such as low mass Higgs searches. To achieve the goal of some of these physics programs, the *BABAR* has also collected the data at the CM energy corresponding to the $\Upsilon(2S)$ and $\Upsilon(3S)$ resonances in the last phase of the data acquisition period in 2008.

This chapter outlines the design of the PEP-II B -Factory and the *BABAR* detector which enabled such a rich physics program from this experiment.

2.1 The PEP-II accelerator

The PEP-II is an asymmetric energy e^+e^- collider operating at the center-of-mass energy of $10.58 \text{ GeV}/c^2$ corresponding to the mass of the $\Upsilon(4S)$ resonance [64]. This resonance subsequently decays almost exclusively to both $B^0\bar{B}^0$ and B^+B^- pairs, which provide an ideal framework for studying the CP -violation in the B mesons decay. A schematic of the overall layout of the PEP-II collider is shown in Figure 2.1.

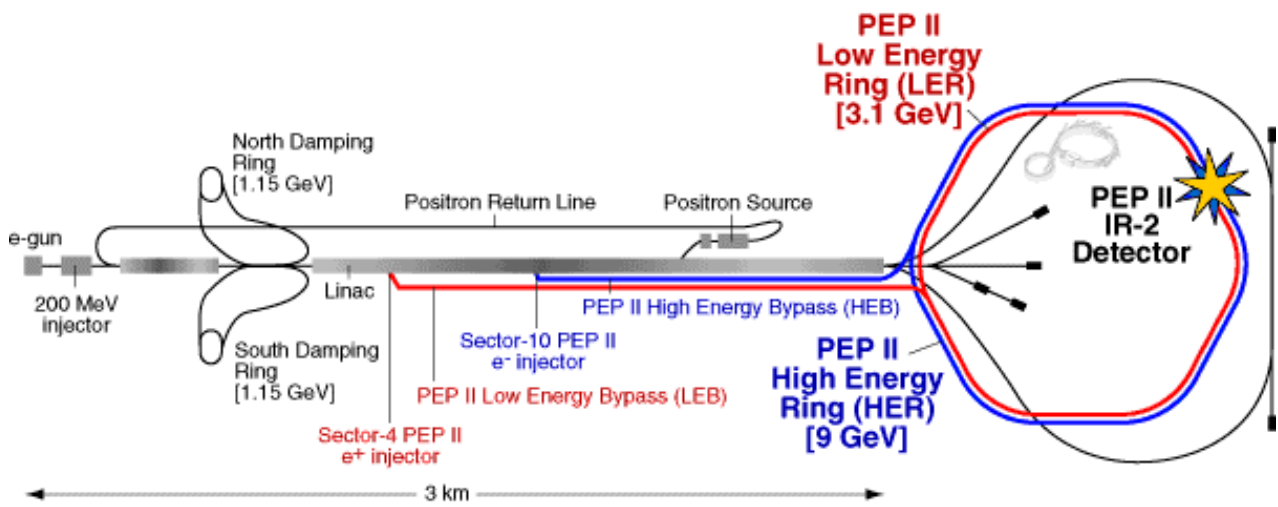


Figure 2.1: The diagram of the PEP-II Accelerator.

The *BABAR* experiment uses two accelerators: the SLAC linear accelerator (linac) and the PEP-II storage ring facility. The SLAC linac accelerates the electron and positron beams to the required high energies, and then it injects them into the PEP-II's storage rings. PEP-II consists of two storage rings, a high Energy Ring (HER) for the 9.0 GeV electron beam, and a low Energy Ring (LER) for the 3.1 GeV positron beam. The two beams move in opposite directions and collide at the interaction point, where the *BABAR* detector is located. The asymmetric beam energies cause the $\Upsilon(4S)$ system to be Lorentz-boosted by a factor of $\beta\gamma = 0.56$ in the laboratory frame, which is important for studying the CP -violation in the B -meson decays. This boost allows to reconstruct the decay vertices of the two B -mesons with enough accuracy to determine the relative decay time needed for time dependent CP -violation measurement.

PEP-II was operational from October 1999 to March 2008. During this period, the *BABAR* experiment has collected about 476 million of $\Upsilon(4S)$ events with an integrated luminosity of 433 fb^{-1} , 120 million of $\Upsilon(3S)$ events with an integrated luminosity of 28.05 fb^{-1} , and 98 million of $\Upsilon(2S)$ events with an integrated luminosity of 14.4 fb^{-1} . *BABAR* has also collected the data with an integrated luminosity of 53.85 fb^{-1} outside these resonances (off-resonance), which are mostly used for continuum background study. Figure 2.2 shows the integrated luminosity of the experiment throughout its running period.

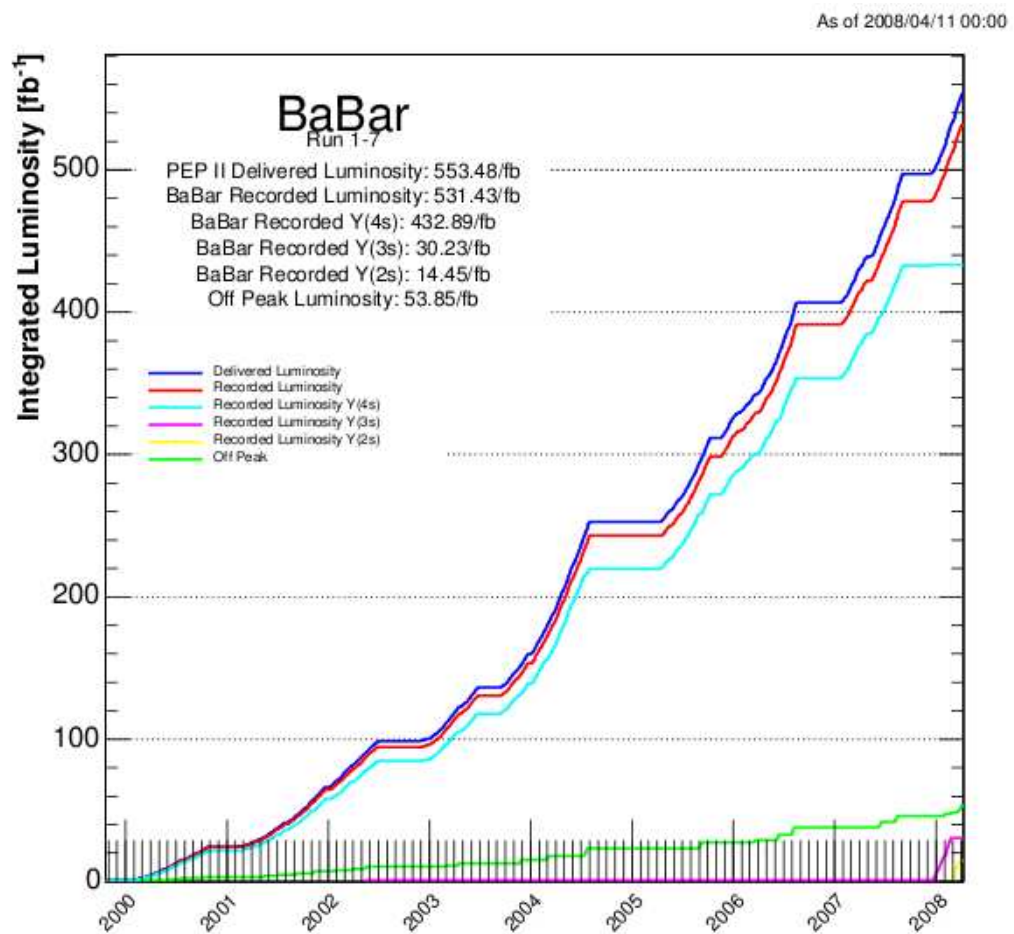


Figure 2.2: Integrated luminosity delivered by PEP-II to the *BABAR* experiment.

2.2 The BABAR detector

The BABAR detector is located at the collision point of the PEP-II accelerator [65]. To achieve the wide physics objective, it is necessary that the detector has a large acceptance, good vertexing, excellent reconstruction efficiencies for charged particles, good energy and momentum resolution, high lepton (particularly e and μ) and hadron identification efficiency and radiation hardness.

The BABAR detector consists of five sub-detectors: silicon vertex tracker (SVT) is positioned closest to the collision point and is responsible for measuring the decay vertices of the B -mesons, a drift chamber (DCH) for charged particle tracking and momentum measurement, a ring-imaging Cerenkov detector for particle identification, and an electromagnetic calorimeter (EMC) for measuring the electromagnetic showers from electrons and photons. These detector subsystems are contained within a large solenoidal magnet capable of generating a 1.5 T magnetic field, and for which the steel flux return is instrumented with a muon detection system. The BABAR detector is illustrated in Figure 2.3, and the following subsections describe these sub-detectors in more detail.

2.2.1 Silicon Vertex Tracker (SVT)

The Silicon Vertex Tracker (SVT) [66] is a semiconductor based tracking sub-detector of the BABAR experiment positioned very close to the collision point. It has been designed to provide the precise measurement of the reconstruction of the trajectory of the charged particles and decay vertices near the interaction region. It reconstructs the decay vertices of two primary B -mesons at the $\Upsilon(4S)$ resonance to determine the time difference between two B -mesons decay, which helps to study the time dependent CP -asymmetries. It is also capable of reconstructing the low momentum charged tracks below 120 MeV/ c that stop before reaching the DCH.

The SVT consists of five concentric cylindrical layers of double-sided AC-coupled silicon micro-strip sensors. The strips on the one side of each sensor are oriented parallel to the beam direction and used to measure the azimuthal angle (ϕ), while other side of the strips are perpendicular to the beam direction and used to measure the position of z . The inner 3 layers are barrel shaped and used to provide an accurate measurement of the impact parameters along z direction and in the $x - y$ plane. However, the outer two layers are arch shaped and used

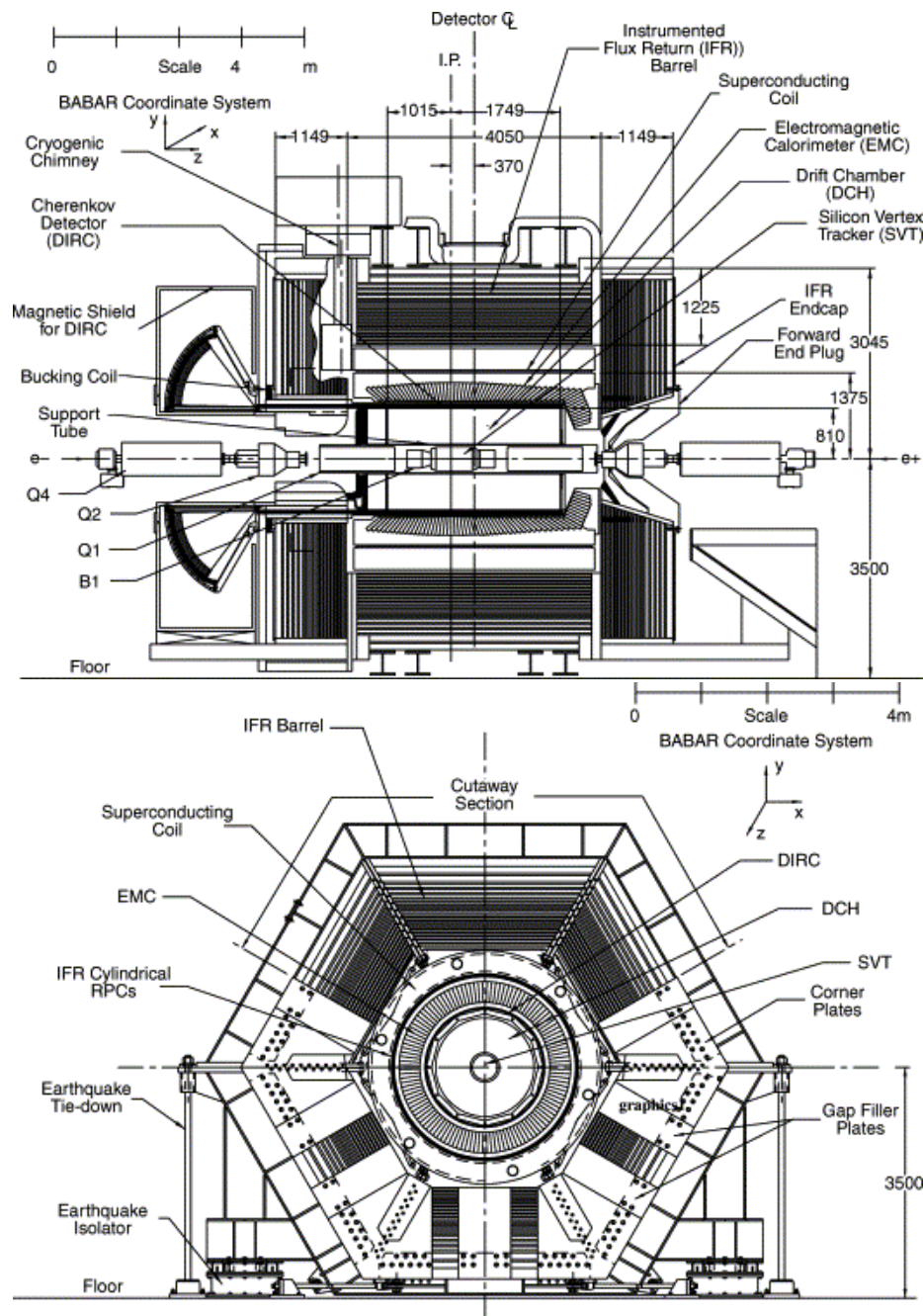


Figure 2.3: Longitudinal section and front end view of the *BABAR* detector [65].

to provide accurate polar angle measurement and can provide the standalone tracking for the low momentum particles that may not be capable of reaching the DCH. This arc design was chosen to minimize the amount of silicon required to cover the solid angle and to increase the crossing angle of the particles near the edges. These outer modules can not be tilted in ϕ like the inner modules because of their geometrical shape. To avoid the gap in the ϕ coordinate,

the two outer layers were divided into two sub-layers (4a, 4b, 5a, 5b as shown in Figure 2.4 (b)), and placed at slightly different radii. Figure 2.4 shows the fully assembled SVT with visible sensors of the outer layer and a transverse schematic view.

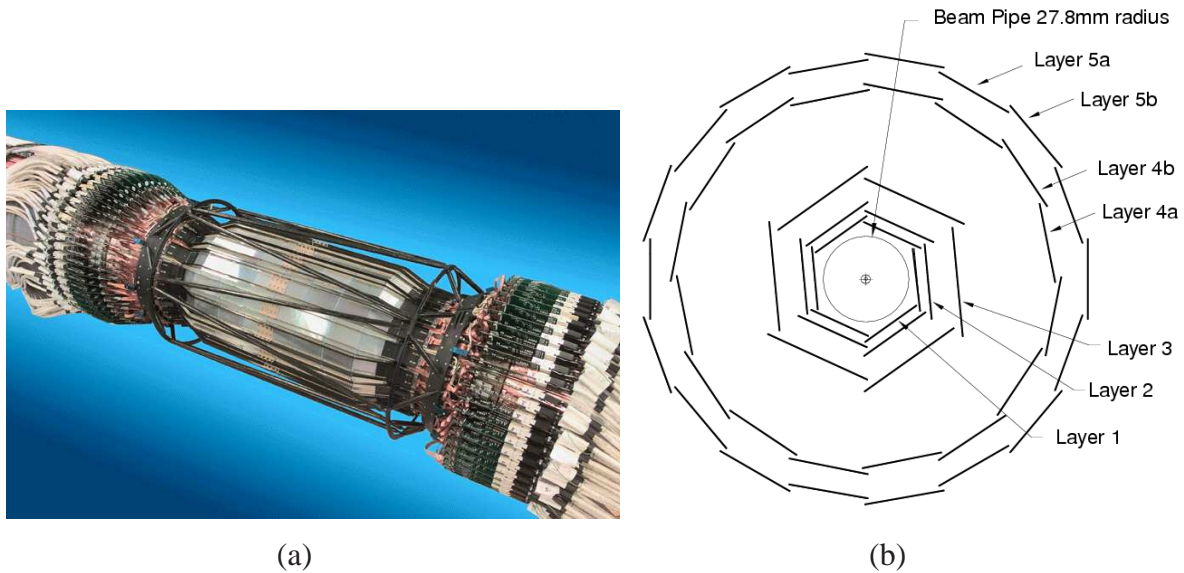


Figure 2.4: The Silicon Vertex Tracker (SVT) (a) fully assembled with visible outer layers and carbon fiber frame and (b) schematic view of the transverse section with the various layers around the beam pipe.

The SVT sensors are composed of a $300\ \mu\text{m}$ thick n-type bulk silicon substrate with p^+ and n^+ strips on opposite sides. These sensors work in the reverse bias mode and are held at a voltage of about 10 V above the depletion voltage, where the typical depletion voltages are 25–35 V. When a charged particles passes through the SVT sensors, it ionizes the materials creating the electron-hole pairs. The electron drifts to the n^+ strips and hole drifts to the p^+ strips. This results in an electrical signal which is read-out via capacitive couplings between the strips and the electronics.

The alignment of the SVT is performed in the following two steps: the local alignment to determine the relative position of all the silicon sensors and the global alignment to correct the movement of the SVT with respect to the rest of the other *BABAR* detectors. The local alignment of the 340 silicon sensors is performed by using a sample of $e^+e^- \rightarrow \mu^+\mu^-$ and the cosmic ray muons, and described by three translations, three rotations and a curvature. By using these parameters, it calculates the track residual using the SVT only hit and performs a χ^2 minimization to determine the best position for each sensor. The local alignment is stable

and performed only rarely. Once the local alignment is done, the SVT also requires to align globally with respect to the DCH since it is not supported structurally by the rest of the other *BABAR* detectors. The global alignment is performed by minimizing the difference between the track parameters fit with the SVT hit only as well as DCH hit only. Other monitoring systems such as temperature, humidity and electronic calibration are also used regularly to ensure the successful SVT operation.

The SVT also includes a radiation protection system consisting of a PIN and a diamond diode sensors located very near to the collision point. The PIN diode consists of p and n-types of semiconductors which are separated by an intrinsic semiconductor located between the regions of these two semiconductors. These radiation systems are used to protect the SVT by the colliding beams in the events of sudden high instantaneous or prolonged background levels that could damage the hardware components.

The SVT performs with an efficiency of 97%, which is calculated for each half-module by comparing the number of associated hit to the number of tracks crossing the active area of the half-module. The spatial resolution of the SVT ranges from 10–15 μm for the inner layers and 30–40 μm for the outer layers. The spatial resolution of the SVT is determined by measuring the distance between the track trajectory and the hit for the high momentum tracks in the two-track events. The SVT is also used to measure the energy loss (dE/dx) of the charged particles which passes through matter and deposit the energy in the sensor. The average dE/dx is used for the particle identification and gives a 2σ separation between kaons and pions up to momentum of 500 MeV/ c and between kaons and protons up to 1 GeV/ c

2.2.2 Drift Chamber (DCH)

The DCH is designed to measure the charged particle momentum with minimum transverse momentum of $p_T > 100$ MeV/ c and the angular distribution with a high precision. It is the main tracking device of the *BABAR* detector, and also enables the particle identification based on the dE/dx measurement for the low momentum of particles where the DIRC is not effective. It is also crucial to reconstruct the long lived particles such as K_s^0 , which often decays outside or on the edge of the SVT, so the chamber should be able to measure the longitudinal positions of a tracks with a resolution of ~ 1 mm. Combined with SVT, the *BABAR* tracking system provides excellent spatial and momentum resolution that enables the reconstruction of

the exclusive B and D -meson decays. The DCH complements the measurements of the impact parameter and the directions of the charged tracks provided by the SVT near the interaction point (IP), and it is also the key to the extrapolation of the charged tracks to the DIRC, EMC and IFR.

The DCH is a 280 cm long cylinder, with an inner radius of 23.6 cm and the outer radius of 81 cm (Figure 2.5). Since the $BABAR$ events are boosted in the forward direction, its design is therefore optimized to reduce the material in the forward end in front of the endcap calorimeter, and offset by 37 cm from the IP to give greater coverage in the forward region. The forward endplate is made thinner (12 mm) in the acceptance region of the detector compared to the rear endplate (24 mm), and all the electronics are mounted on the rear backward endplate. The inner cylinder is made of 1 mm beryllium corresponding to 0.28% of the radiation length (X_0), while the outer is made of 2 layers of carbon fiber of a honeycomb core corresponding to 1.5% of the X_0 .

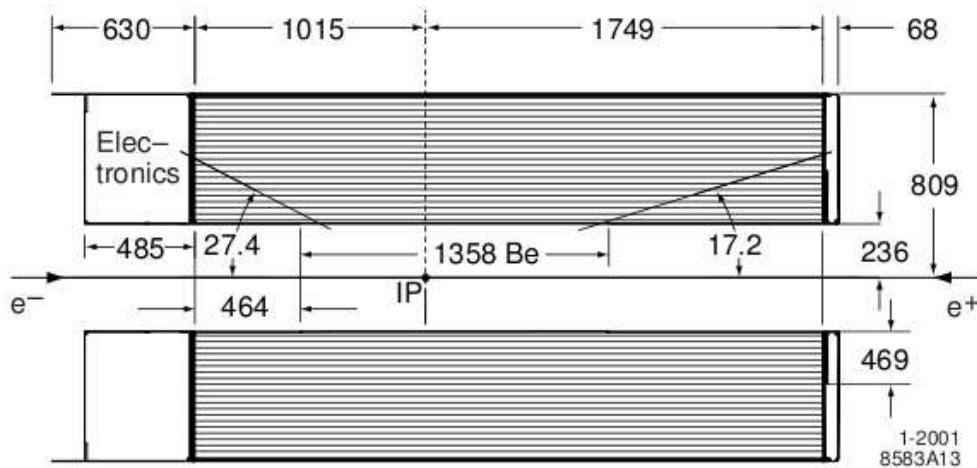


Figure 2.5: Longitudinal cross-section of the drift chamber.

The DCH consists of 7104 drift cells, which are arranged in 10 super-layers of 4 layers each, for a total of 40 layers. The stereo angles of the super-layers alternate between axial (A) and stereo (U,V) in following order: AUVAUVAUVA. The stereo angles increase from 45 mrad in the innermost super-layer to 76 mrad in the outermost super-layer. The chamber is filled with a 80:20 gas mixture of helium:isobutane to provide good spatial separation and resolution for the dE/dx measurement and reasonably short response time, where the helium is chosen to minimize the multiple scattering.

Figure 2.6 shows the design of the drift cells for the four innermost super-layers. The 7104 cells are hexagonal with a typical dimension of $1.2 \times 1.8 \text{ cm}^2$, to minimize the drift time. The sense wires is a $25 \text{ }\mu\text{m}$ gold-plated tungsten-rhenium wire, while the field wires are gold-plated aluminium with diameters of $120 \text{ }\mu\text{m}$ and $80 \text{ }\mu\text{m}$. A voltage of 1960 V is applied to the sense wires, while the field wires are held at ground.

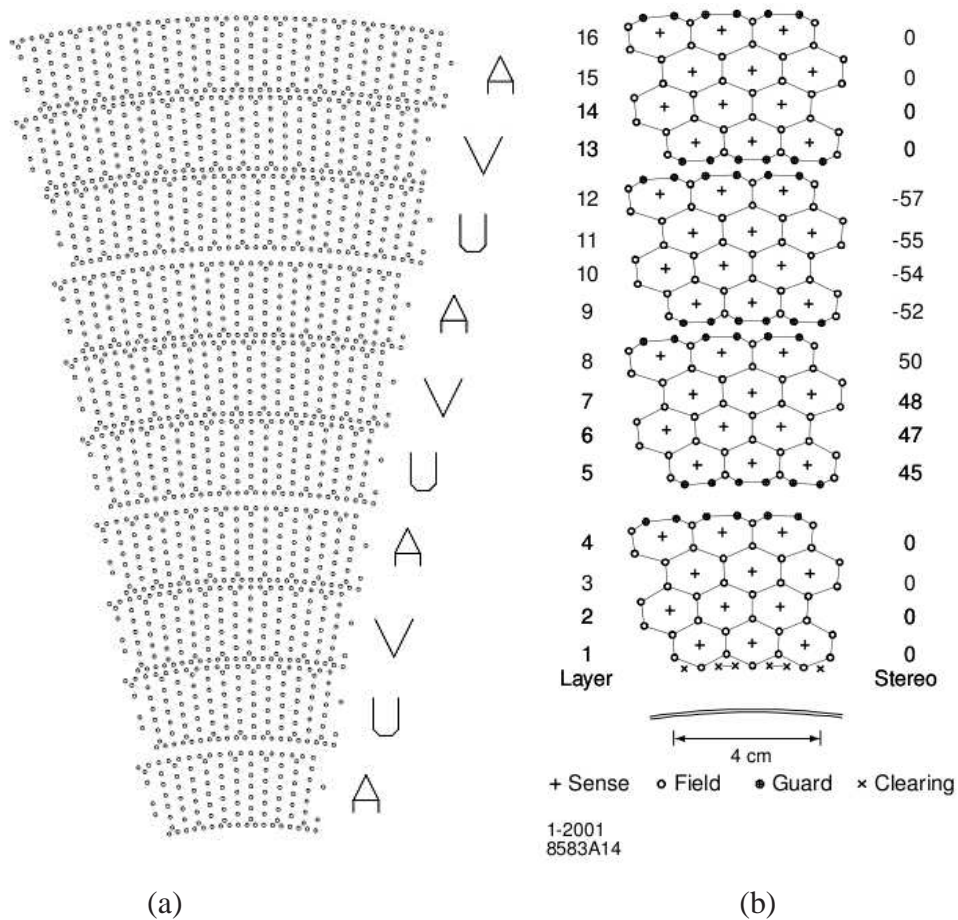


Figure 2.6: (a) Cell layout in the BABAR drift chamber. (b) DCH drift cell configuration for the four innermost super-layers. The numbers on the right give the stereo angles in mrad of the sense wires in each layer.

The ionized charged particles in the gas produces the free electrons that are accelerated towards the sense wires by the applied electromagnetic field. These accelerated electrons are further ionized and result in an avalanche of the electric charge near the wire. The avalanche accumulates at the sense wire producing a measurable electrical signal, which is amplified and read-out to the electronics. The integrated charge and drift time (time required for the ionized

electrons to reach the sense wire) provide the ionization energy-loss and position information of the charged particles, respectively.

The track of the charged particles is defined by five parameters $(d_0, \phi_0, \omega, z_0, \tan\lambda)$, which are measured at the point of closest approach to the z -axis, and their associated error matrix. The d_0 and z_0 represent the distance of a track from the origin of the coordinate system in the $x - y$ plane and along the z -axis; The ϕ_0 is the azimuthal angle of the track; λ is the dip angle relative to the transverse plane, and $\omega = 1/p_T$ is the curvature of the track. Based upon the full width half maxima, the distributions of these variables have the following resolution values: $\sigma_{d_0} = 23 \mu\text{m}$, $\sigma_{\phi_0} = 0.43 \text{ mrad}$, $\sigma_{z_0} = 29 \mu\text{m}$ and $\sigma_{\tan\lambda} = 0.53 \times 10^{-3}$. The DCH performs with a tracking efficiency of $(98 \pm 1)\%$ for $p_T > 200 \text{ MeV}/c$ and for polar angle $\theta > 500 \text{ mrad}$ at the voltage of 1960 V. The resolution of the measured p_T can be written as a linear function of $\sigma_{p_T}/p_T = (0.13 \pm 0.01)\% p_T + (0.45 \pm 0.03)\%$.

The specific energy loss per track is computed as a truncated mean from the lowest 80% of the individual dE/dx measurements. This value is computed after incorporating all the corrections. The corrections are needed to account for changes in gas pressure and mixture; differences in cell geometry and charge collection; signal saturation due to space charge buildup; non-linearities in the most probable energy loss at large track dip angles; and changes in cell charge collection as a function of track entrance angle. The corrections are all done once for a given High-Voltage (HV) setting and a given gas mixture whilst the gain corrections must be updated run by run. Corrections at the cell level can be large compared to the dE/dx resolution for a single cell, but have only a small impact on the average resolution of the ensemble of hits. The dE/dx as a function momentum is shown in Figure 2.7. The DCH achieves good separation between K and π upto 700 MeV/c .

2.2.3 The Detector of Internally Reflected Cherenkov Light (DIRC)

The DIRC is a new type of ring-imaging Cherenkov detector used for hadronic particle identification in the BABAR experiment. It provides the π/K separation of greater than 4σ for all tracks from pion Cherenkov threshold up to 4.2 GeV/c , and tags the flavor of a B meson via the cascade decay of $b \rightarrow c \rightarrow s$. Its imaging system is based upon the total internal reflection of Cherenkov photon produced in long quartz bar. When a particle passes through the

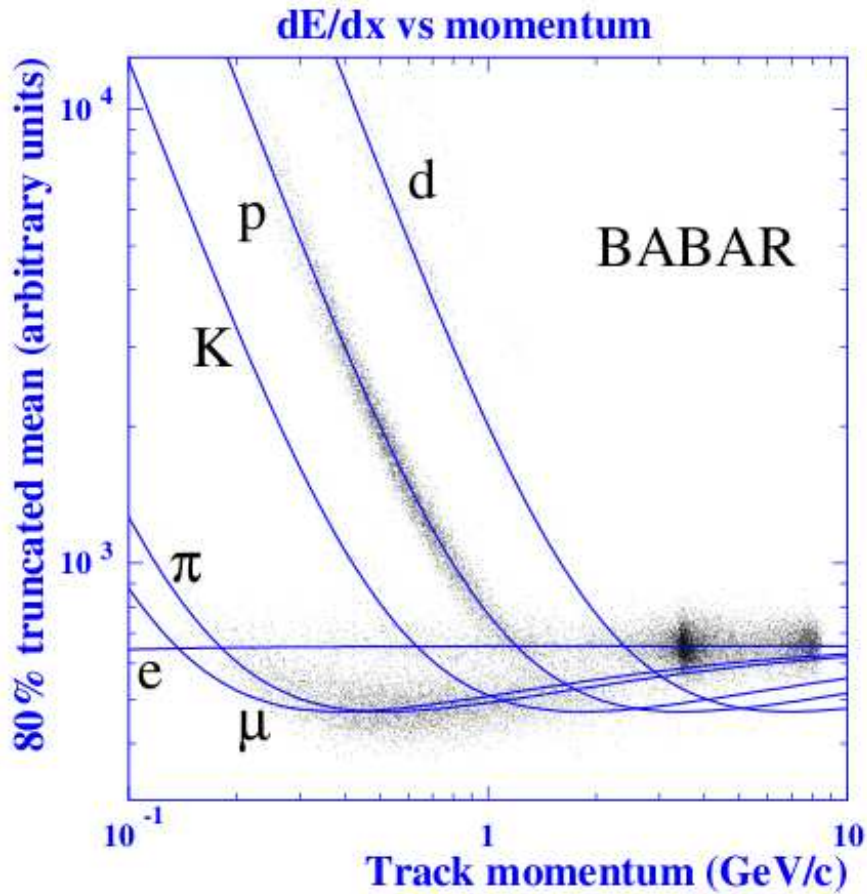


Figure 2.7: dE/dx in the DCH as a function of momentum for different particles.

medium with a velocity greater than the speed of light in that medium, it emits photons known as Cherenkov radiation. The angle of the Cherenkov radiation is defined as

$$\cos\theta_c = \frac{c}{nv}, \quad (2.1)$$

where c is the velocity of light, n is the refractive index of the medium and v is the speed of the particle.

The DIRC is a three-dimensional imaging device, used to sense the position and arrival time of the signal by using an array of densely packed photomultiplier tubes. It consists of 144 radiation-hard fused silica bars with a refractive index of $n = 1.473$. The bar serves

both as radiators and as light pipes for the light trapped in the radiator by total internal reflection. A schematic of the DIRC geometry illustrating the principle of light production, imaging and transportation is shown in Figure 2.8. Photons are generated by the particles above the Cherenkov threshold, trapped inside the bars and emerge into a water-filled expansion region, called a standoff box. A fused silica wedge is used to reflect photons at large angles to reduce the size of the required detection surface and hence recover those photons that would be lost due to internal reflection at the fused silica and water interface. Finally, the photo multiplier tubes (PMTs) detect the light and allow the Cherenkov angle and particle velocity to be measured. Once the velocity is known, the mass of the particle can be calculated using the momentum information from the DCH.

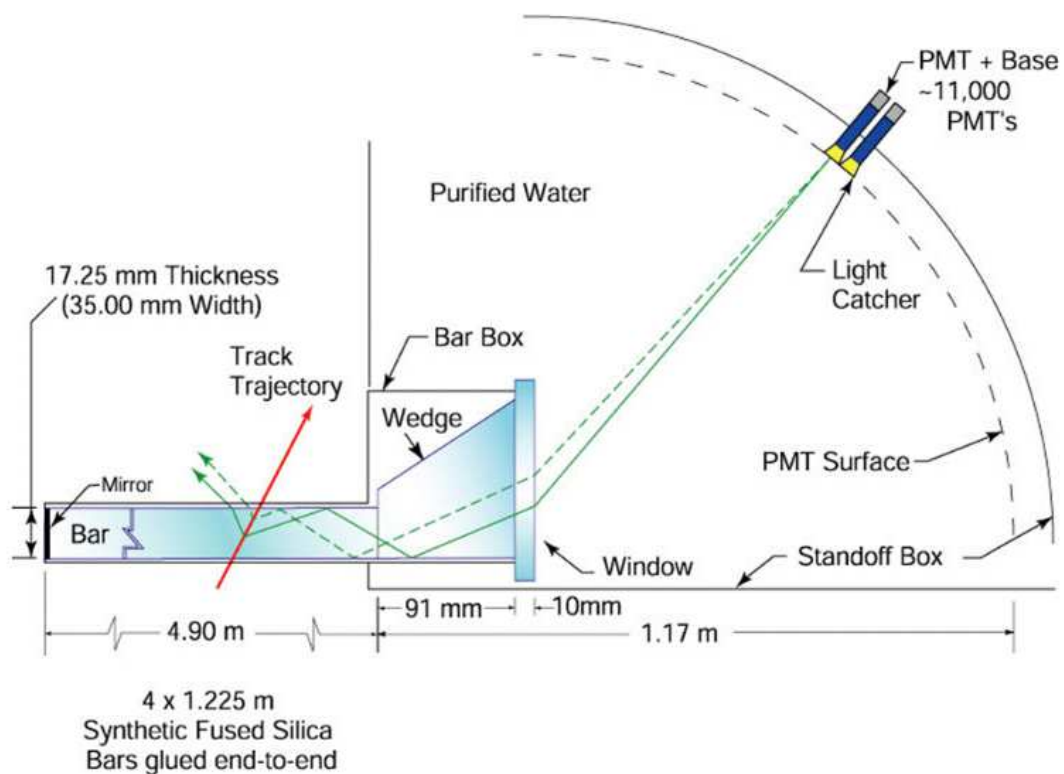


Figure 2.8: Schematic of the DIRC fused silica radiator bar and imaging region.

2.2.4 The Electromagnetic Calorimeter

The EMC is designed to measure the photon showers with excellent efficiency, energy and angular resolution over the energy range from 20 MeV to 9 GeV. The capability of the EMC allows the detection of photons from π^0 and η^0 as well as from electromagnetic and radiative processes. Most of the photons are produced by neutral pion decays with maximum energy of 200 MeV, hence the lower bound of energy is set to allow the reconstruction of B -meson decays containing multiple π^0 mesons. However, the upper bound of the energy range is set by calibrating and monitoring the luminosity of the photons produced via the QED processes, like $e^+e^- \rightarrow e^+e^-(\gamma)$. The EMC is also used to identify the electrons which allows to study of semi-leptonic and rare decays of B and D mesons, and τ leptons, and the reconstruction of vector mesons like J/ψ .

The EMC consists of a cylindrical barrel and a conical forward endcap. It has a full coverage in the azimuth and extends in polar angle from 15.8° to 141.8° corresponding to a solid-angle coverage of 90% in the CM system (Figure 2.9). The barrel contains 5,760 thallium-doped caesium iodide (CsI(Tl)) crystals arranged in 48 distinct rings with 120 identical crystals each. The endcap holds 820 crystals arranged in eight-rings, adding up to a total of 6,580 crystals. The crystals have a tapered trapezoidal cross-section and length of the crystals increases from 29.6 cm in the backward to 32.4 cm in the forward direction to limit the effects of shower leakage from increasingly higher energy particles. Two silicon PIN diodes mounted on the rear face of each crystal are used to readout the scintillation light.

A typical electromagnetic shower tends to spread over many adjacent crystals, forming a cluster of adjacent energy deposits. Pattern recognition algorithms are used to analyze the shower shape and to check whether cluster can be associated with the charged particles. Otherwise, the EMC cluster would be assumed to originate from a neutral particle. The energy resolution of a homogeneous crystal calorimeter is empirically described by

$$\frac{\sigma_E}{E} = \frac{a}{4\sqrt{E(\text{GeV})}} \oplus b, \quad (2.2)$$

where \oplus signifies addition in quadrature, and E and σ_E are the energy and rms value of a photon. The angular resolution is determined by the transverse crystal size and the distance

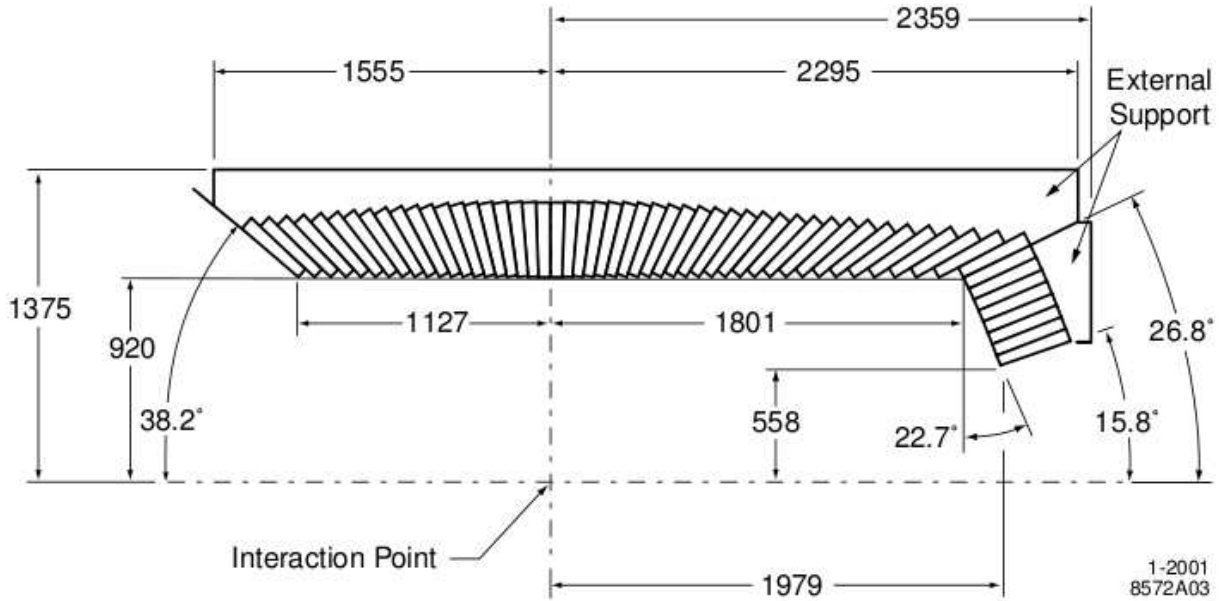


Figure 2.9: A longitudinal cross-section of the EMC (only the top half is shown) indicating the arrangement of 56 crystal rings. The detector is axially symmetric around the z -axis. All dimensions are given in mm.

from the interaction point, which is defined as

$$\sigma_{\theta} = \sigma_{\phi} = \frac{c}{\sqrt{E(\text{GeV})}} + d, \quad (2.3)$$

The energy dependent terms a and c are dominant at low energy and arise due to the fluctuations in photon statistics and electronic noise in the readout chain. Furthermore, beam-generated background will lead to a large numbers of additional photons that add additional noise. The constant terms b and d are dominant at higher energies (> 1 GeV) and arise due to non-uniformity in light collection and light absorption in the detector materials.

2.2.5 The Instrument Flux Return (IFR)

The IFR was designed to identify the muons with high efficiency and good impurity, and to detect neutral hadrons (primarily K_L^0 and neutrons) over a wide range of momenta and angles. Muons are important for tagging the B mesons via semileptonic decays, for the reconstruction

of the vector mesons, like the J/ψ , and for the study of semi-leptonic and rare decays of B and D mesons and τ leptons. K_L^0 detection is important to study the exclusive B decays, in particular CP eigenstates. The IFR also helps in vetoing charm decays and improving the reconstruction of neutrinos.

The IFR consists of one barrel and two endcap and uses the steel flux return of the magnet as a muon filter and hadron absorber. Single gap resistive plate chamber (RPC) with two-coordinate readout have also been chosen as an active detector. The IFR was originally equipped with 19 layers of RPC in the barrel and 18 in the endcaps. In addition, two layers of cylindrical RPCs are installed between the EMC and the magnet crystal to detect the particles existing the EMC. The RPC consists of two high resistivity Bakelite sheets coated with linseed oil separated by a 2 mm gap containing 56% argon, 38.8% Freon 134a, and 4.5% isobutane. The RPCs operate in the limited streamer mode at ~ 8 kV, and streamer signals readout by aluminum strips on the exterior of the plates. An illustration of the layout of the IFR is shown Figure 2.10.

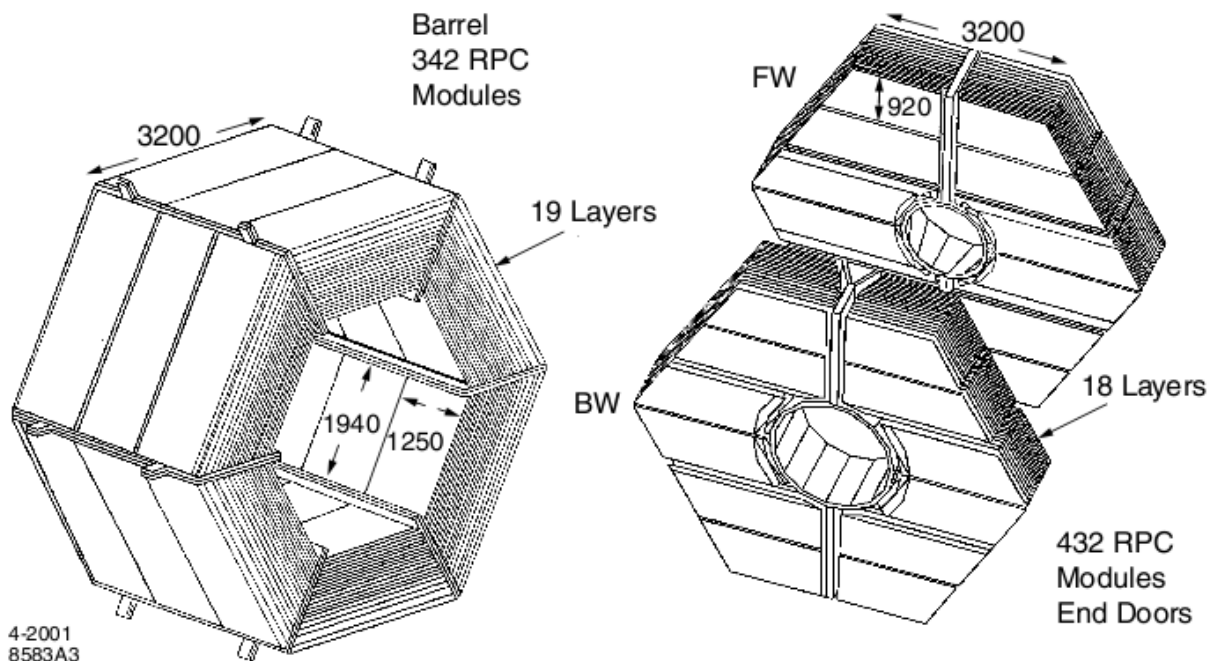


Figure 2.10: Layout of IFR barrel and endcaps. All the units are given in mm.

Unfortunately, it was found that the RPC degraded rapidly. Therefore, the muon detection

system was upgraded with a Limited Streamer Tubes (LST) [67, 68] during the detector shut-down periods from 2004-2006. The inner 18 layers of the RPC detector were replaced with 12 layers of LST detectors and 6 layers are filled with brass absorber, to improve muon-hadron separation. The LST detector consists of 7–8 cells with a dimension of $380 \times 15 \times 17 \text{ mm}^3$. The cells are composed of PVC plastic coated with a graphite paint, which is maintained at a ground potential, a central high voltage gold plated anode which is held in place by 6 wires holders, and are filled with a (89:8:3) gas mixture of CO_2 , ISO-butane and Argon. The LST also operates in the streamer mode, and the signals are collected by the external read-out strips.

2.2.6 Trigger Selection

The trigger system [69] was designed to select events of interest with a high, stable, and well-understood efficiency while rejecting background events and keeping the total event rate under 120 Hz. *BABAR* uses two types of the trigger systems: the hardware based level 1 Trigger (L1) and the software based Level 3 Trigger (L3). The details of L1 and L3 trigger systems are described below:

Level 1 Trigger system

The design of L1 trigger decision is based on the charged tracks in the DCH above a preset transverse momentum, showers in the EMC, and tracks detected in the IFR. The drift chamber trigger (DCT) processes the input data consisting of one bit from each of the 7104 cells to identify tracks. The Electromagnetic Trigger (EMT) receives input from the 280 towers in the EMC, and identifies the energy deposits in the EMC. The IFR is divided into ten sectors, namely the six barrel sextants and the four half end doors. The primary functions of the instrument flux return trigger (IFT) are to veto cosmic events and to identify muons from the interaction of $e^+e^- \rightarrow \mu^+\mu^-$, which can be used for measuring the detector parameters such as the luminosity. The output of the DCT, EMT and IFT are utilized to determine whether the signal event constraints a physics event by a Global Level Trigger. The frequency at which the events are accepted by the Level 1 Trigger is approximately 1 KHz.

Level 3 Trigger

The L3 trigger is an online application that acts primarily as an event filter. It implements in the framework of the Online Event Processing (OEP) and runs in parallel on a number of Unix processors. It is the first stage of the the DAQ system to select the events and is responsible to make a logging decision on the output of the L1 hardware trigger. Its design was required to reduce the L1 output of 1 kHz to a logging rate of about 120 Hz with a high efficiency in physics events of interest. It performs a partial reconstruction of the event based on the data from the drift chamber and drift chamber trigger as well as from the EMC. Its data are in part used by the offline reconstruction and its trigger decision records are input to the offline filters of DigiFilter and BGFilter.

The offline filters are based on prompt reconstruction (PR) used to provide a further selection of events before the full reconstruction. The selection is done using two levels of filters: the DigiFilter and BGFilter. The DigiFilter uses only information available from the L1 and L3 triggers to make the selection. It is primarily used to remove the calibrated events, such as radiative Babha events. The BigiFilter first runs as a part of the offline reconstruction to find drift chamber tracks and EMC clusters. Based on these tracks and clusters information, an event classification is done, where the events classified as multi-hadron, τ or two-prong etc are identified.

2.3 Chapter Summary

In this chapter, we describe the *BABAR* detector, PEP-II accelerator complex and the trigger system which are used to collect the dataset analyzed in this thesis.

Chapter 3

Event reconstruction and selection

This chapter describes the event reconstruction and the selection criteria applied to select signal-like events for the decay $\Upsilon(2S, 3S) \rightarrow \pi^+\pi^-\Upsilon(1S)$, $\Upsilon(1S) \rightarrow \gamma A^0$, $A^0 \rightarrow \mu^+\mu^-$. A blind analysis [70] technique is used, where the $\Upsilon(2S, 3S)$ datasets are blinded until all the selection criteria are finalized for an optimal value of signal-to-noise ratio. In this chapter, we describe the discriminative variables used to separate signal from background. A more advanced multivariate technique based BumpHunter and Random forest classifiers are also used to improve the purity of the $\Upsilon(1S)$ sample. We also discuss the $\Upsilon(2S)$ and $\Upsilon(3S)$ datasets used in this analysis along with the Monte-Carlo (MC) samples which are intended to model the data. The luminosities of these datasets are also documented.

3.1 Data Sets

The data sample used in this analysis was collected during Run 7, specifically during a period between December 2007 and April 2008 by the *BABAR* detector. The $\Upsilon(3S)$ dataset contains $(121.9 \pm 1.1) \times 10^6$ $\Upsilon(3S)$ events and the $\Upsilon(2S)$ dataset contains $(98.3 \pm 0.9) \times 10^6$ $\Upsilon(2S)$ events.

The $\Upsilon(3S)$ data set is divided into three sub samples: low, medium, and high which were collected in the beginning, middle and the end of Run7, respectively. The “Low” data set corresponds to about 4.2% of the total $\Upsilon(3S)$ on resonance data set and is used for checking

Dataset Name	Integrated Luminosity (fb^{-1})
For $\Upsilon(3S)$ dataset	
AllEvents-Run7-R24b $\Upsilon(3S)$ -OnPeak-Low	1.173
AllEvents-Run7-R24b $\Upsilon(3S)$ -OnPeak-Medium (So far blind)	25.594
AllEvents-Run7-R24b $\Upsilon(3S)$ -OnPeak-High (So far blind)	1.282
For $\Upsilon(2S)$ dataset	
AllEventsSkim-Run7- $\Upsilon(2S)$ -OnPeak-R24d-LowOnpeak	0.758
AllEventsSkim-Run7- $\Upsilon(2S)$ -OnPeak-R24d (So far blind)	13.56

Table 3.1: The luminosity of each data sample used in the analysis.

the agreement between data and MC and finally, for validating the analysis procedure. For the $\Upsilon(2S)$ analysis, a similar ‘‘Low’’ data set was generated which corresponds to 5.6% of the total $\Upsilon(2S)$ data set. The ‘‘Low’’ samples are kept blinded until all the selection criteria are finalized. We unblind these ‘‘Low’’ samples later to validate the fit procedure after applying all the optimal selection cuts. The luminosities of these samples are shown in Table 3.1. To avoid any bias, these samples are discarded from the final dataset.

MC simulated events are used to study the detector acceptance and optimize the event selection procedure. The EvtGen package [71] is used to simulate the $e^+e^- \rightarrow q\bar{q}$ ($q = u, d, s, c$) and generic $\Upsilon(2S, 3S)$ production, BHWIDE [72] to simulate the Bhabha scattering and KK2F [73] to simulate the decay processes of $e^+e^- \rightarrow (\gamma)\mu^+\mu^-$ (radiative di-muon) and $e^+e^- \rightarrow (\gamma)\tau^+\tau^-$. Signal events are generated using a phase-space (P-wave) model for the $A^0 \rightarrow \mu^+\mu^-$ ($\Upsilon(1S) \rightarrow \gamma A^0$) decay and the hadronic matrix elements measured by the CLEO experiment [74] are used for the $\Upsilon(2S, 3S) \rightarrow \pi^+\pi^-\Upsilon(1S)$ transition. The detector response is simulated by GEANT4 [75], and time-dependent detector effects are included in the simulation. The cross-sections for $e^+e^- \rightarrow q\bar{q}$ and lepton-pair productions are calculated from their values at the $\Upsilon(4S)$ assuming $1/s$ scaling, where \sqrt{s} is the e^+e^- CM energy at $\Upsilon(nS)$ ($n = 2, 3, 4$) resonances. Table 3.2 summarizes the number of generated signal MC events at different masses for the decay chains of $\Upsilon(2S, 3S) \rightarrow \pi^+\pi^-\Upsilon(1S)$, $\Upsilon(1S) \rightarrow \gamma A^0$, $A^0 \rightarrow \mu^+\mu^-$. The cross-sections and luminosities of these background decay processes are summarized in Table 3.3. We use these six types of background MCs and a signal MC sample in the mass range of 0.212 - 9.46 GeV/ c^2 to optimize the selection criteria.

Mass of A^0 GeV/ c^2	Number of events	Mass of A^0 GeV/ c^2	Number of events
For $\Upsilon(3S)$ dataset		For $\Upsilon(2S)$ dataset	
0.212	172k	0.212	126.2k
0.214	172k	0.214	126.2k
0.216	172k	0.216	126.2k
0.218	172k	0.218	126.2k
0.220	172k	0.220	126.2k
0.225	172k	0.225	126.2k
0.300	172k	0.500	126.2k
0.500	172k	0.750	126.2k
0.750	172k	1.0	87k
1.0	103k	1.5	126.2k
1.5	172k	2.0	87k
2.0	103k	3.0	87k
3.0	103k	4.0	87k
4.0	103k	5.0	87k
5.0	103k	6.0	87k
6.0	103k	6.7	87k
6.7	103k	7.0	87k
7.0	95k	7.5	87k
7.5	103k	8.0	87k
8.0	103k	8.25	126.2k
8.25	172k	8.5	126.2k
8.5	172k	8.75	126.2k
8.75	172k	9.0	87k
9.0	103k	9.10	126.2k
9.25	172k	9.20	126.2k
0.212 – 9.46	204k	9.25	126.2k
		0.212 – 9.46	174k

Table 3.2: The number of signal MC events generated at different masses for the decay chains of $\Upsilon(2S, 3S) \rightarrow \pi^+\pi^-\Upsilon(1S)$, $\Upsilon(1S) \rightarrow \gamma A^0$, $A^0 \rightarrow \mu^+\mu^-$.

3.2 Event Reconstruction and Event Pre-Selection

The events of interest are reconstructed using *BABAR* software packages designed for creating the lists of composite particles, automating the work of making combinations, performing the kinematic fits, making the pre-selection criteria and storing the events in an object-oriented based ROOT ntuple files [76]. To streamline the decay processes of $\Upsilon(2S, 3S) \rightarrow \pi^+\pi^-\Upsilon(1S)$, $\Upsilon(1S) \rightarrow \gamma A^0$, $A^0 \rightarrow \mu^+\mu^-$, the data and MC samples are filtered or “skimmed“. We select

Decay Mode	Generated Events	Cross-section (nb)	Luminosity (fb^{-1})
For $\Upsilon(3S)$ dataset			
$\Upsilon(3S) \rightarrow anything$	215456000	4.19	51.42
$e^+e^- \rightarrow q\bar{q}$ ($q = u, d, s$)	111576000	2.18	51.18
$e^+e^- \rightarrow c\bar{c}$	135224000	1.36	99.429
$e^+e^- \rightarrow \tau^+\tau^-$	47632000	0.94	50.672
$e^+e^- \rightarrow \gamma e^+e^-$	283856000	25.79	11.01
$e^+e^- \rightarrow \gamma\mu^+\mu^-$	68744000	1.1985	57.358
For $\Upsilon(2S)$ dataset			
$\Upsilon(2S) \rightarrow anything$	156400000	7.249	21.57
$e^+e^- \rightarrow q\bar{q}$ ($q = u, d, s$)	91025000	2.31	39.41
$e^+e^- \rightarrow c\bar{c}$	51420000	1.44	35.71
$e^+e^- \rightarrow \tau^+\tau^-$	20245000	1.04	19.47
$e^+e^- \rightarrow \gamma e^+e^-$	106268000	25.9	4.10
$e^+e^- \rightarrow \gamma\mu^+\mu^-$	26891000	1.30	20.69

Table 3.3: Background MC samples for different decay processes, which are used in this analysis.

events containing exactly four charged tracks and a single energetic photon with a center-of-mass (CM) energy greater than 200 MeV. The additional photons with CM energies below this threshold are also allowed to be present in the events. The two highest momentum tracks in the CM frame are required to have opposite charge, and are assumed to be muon candidates, combined to form the A^0 candidate. These tracks are required to have a distance of closest approach to the interaction point of less than 1.5 cm in the plane transverse to the beam and less than 10 cm along the beam-axis. The $\Upsilon(1S)$ candidate is reconstructed by combining the A^0 candidate with the energetic photon candidate and requiring the invariant mass of the $\Upsilon(1S)$ candidate to be between 9.0 and 9.8 GeV/ c^2 . The $\Upsilon(2S, 3S)$ candidates are formed by combining the $\Upsilon(1S)$ candidate with the two remaining tracks, assumed to be pions. The di-pion invariant mass must be in the range of $[2m_\pi, (m_{\Upsilon(2S,3S)} - m_{\Upsilon(1S)})]$, compatible with the kinematic boundaries of the $\Upsilon(2S, 3S) \rightarrow \pi^+\pi^-\Upsilon(1S)$ decay. Finally, we define the mass recoiling against the di-pion system to be:

$$m_{\text{recoil}}^2 = s + m_{\pi\pi}^2 - 2\sqrt{s}E_{\pi\pi}, \quad (3.1)$$

where \sqrt{s} is the collider CM energy (assumed to be $\sqrt{s} = M_{\Upsilon(3S,2S)}$) and $E_{\pi\pi}$ is the energy of the di-pion system. We require that m_{recoil} to be between 9.35 and 9.57 GeV/ c^2 . The m_{recoil}

is used to identify the $\Upsilon(3S, 2S) \rightarrow \pi^+\pi^-\Upsilon(1S)$ transitions and it should be peaked at $\Upsilon(1S)$ mass for signal like events. The entire decay chain is fit imposing a mass constraint on the $\Upsilon(1S)$ and $\Upsilon(2S, 3S)$ candidates, as well as requiring the energy of the $\Upsilon(2S, 3S)$ candidate to be consistent with the e^+e^- CM energy.

To distinguish the signal from backgrounds, we calculate the reduced mass [49] for an event which is defined as:

$$m_{\text{red}} = \sqrt{m_{\mu^+\mu^-}^2 - 4m_{\mu}^2}. \quad (3.2)$$

m_{red} is equal to twice the momentum of the muons in the rest frame of A^0 , and has a smooth distribution in the region of the kinematic threshold $m_{\mu^+\mu^-} \approx 2m_{\mu}$ ($m_{\text{red}} \approx 0$). It has a Gaussian-like distribution for signal and a flat distribution for background.

Further selection criteria are applied at the ntuple level. Events are required to satisfy L3 trigger (L3OutDch || L3OutEmc) and filter (RecoBGFilter && DigiFilter) flags. The trigger selection criteria reduces significant amount of combinatorial backgrounds while maintaining the signal selection efficiencies up to $\sim 99.95\%$ for both the $\Upsilon(2S, 3S)$ datasets. Further, we require that the momentum magnitude of most energetic charged particle to be less than $8.0 \text{ GeV}/c$. Figure 3.1 shows the m_{recoil} distribution for signal MC, combined background MC of $\Upsilon(3S, 2S)$ generic, radiative bhabha, radiative di-muon, $\tau^+\tau^-$, $c\bar{c}$ and uds.

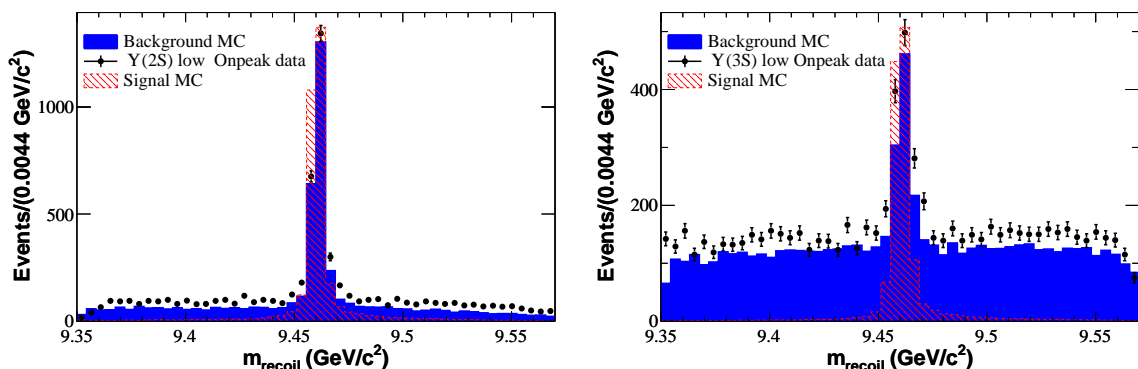


Figure 3.1: The distribution of the m_{recoil} for the low onpeak data sample (dots), together with the production of the background and signal MC for $\Upsilon(2S)$ (left) and $\Upsilon(3S)$ (right). The mean of the recoil mass in background MC has been corrected after comparing the recoil mass distributions in a control samples of data and MC, the details of which can be found in section 3.4. The background MC is normalized to the $\Upsilon(2S, 3S)$ low onpeak data samples.

3.3 Event Selection

This section describes the variables used to discriminate between signal and background processes. We also describe various multi-variate techniques used to discriminate signal events from the background events. The variables of interest can be split into three groups, which are pion, photon, and muon related variables. The pion related variables in the decay chains of $\Upsilon(3S, 2S) \rightarrow \pi^+\pi^-\Upsilon(1S)$ are identified by searching for two low momentum pions. The photon related variables in the decay chain of $\Upsilon(1S) \rightarrow \gamma A^0$ are identified by detecting a monochromatic photon. The muon related variables in the decay chain of $A^0 \rightarrow \mu^+\mu^-$ are identified by two high momentum muons. The kinematic variables related to these three groups are chosen as follows.

3.3.1 Pion selection variables

- **Costhpi_{pi}**: The cosine of the angle between two pions in the laboratory frame, shown in Figure 3.2(a) for $\Upsilon(2S)$ and Figure 3.2(b) for $\Upsilon(3S)$.
- **DiPip₃**: The transverse momentum of the di-pion system in the laboratory frame, shown in Figure 3.2(c) for $\Upsilon(2S)$ and Figure 3.2(d) for $\Upsilon(3S)$.
- **Pi2phi**: The azimuthal angle of each pion, shown in Figure 3.2(e) for $\Upsilon(2S)$ and Figure 3.2(f) for $\Upsilon(3S)$.
- **Pi2plab**: The transverse momentum of the pions, shown in Figure 3.3(a) for $\Upsilon(2S)$ and Figure 3.3(b) for $\Upsilon(3S)$.
- **DiPimass**: The di-pion invariant mass, shown in Figure 3.3(c) for $\Upsilon(2S)$ and Figure 3.3(d) for $\Upsilon(3S)$.
- **Costhetax**: The cosine of the angle formed between the π^+ in the di-pion frame and the direction of the di-pion in the $\Upsilon(2S, 3S)$ rest frame, shown in Figure 3.3(e) for $\Upsilon(2S)$ and Figure 3.3(f) for $\Upsilon(3S)$.
- **VDist**: The transverse position of the di-pion vertex, shown in Figure 3.4(a) for $\Upsilon(2S)$ and Figure 3.4(b) for $\Upsilon(3S)$.

- **RecoilMass:** the mass recoiling against the di-pion system, shown in Figure 3.4(c) for $\Upsilon(2S)$ and Figure 3.4(d) for $\Upsilon(3S)$.

The pion azimuthal angle in the radiative bhabha sample of both $\Upsilon(3S)$ and $\Upsilon(2S)$ datasets shows a multipeak structure, as shown in Figure 3.2(e) and 3.2(f). The peak structure is understood to be due to the random tracks which is removed after requiring that either one of the charged tracks must be identified as muon for A^0 reconstruction using muon Particle-ID (PID). Figure 3.5 shows the azimuthal angle of pion after applying the muon ID cut.

3.3.2 Muon selection variables

- **BDTMuon[1,2]IDFakeRate:** We require either one of the charged tracks for the A^0 reconstruction must be identified as muon by a standard Muon particle-ID algorithm, where the μ -to- π misidentification rate is about 3%. Figure 3.6 shows the muon PID Boolean distribution of the $\Upsilon(2S, 3S)$ datasets.

3.3.3 Track multiplicity and photon selection variables

- **nTracks:** We require that the number of charged tracks should be equal to four in the event.
- **xlmomgam:** The lateral moment [77] of a photon candidate in the electromagnetic calorimeter is defined as

$$\text{xlmomgam} = \frac{\sum_{i=3}^N E_i r_i^2}{\sum_{i=3}^N E_i r_i^2 + E_1 r_0^2 + E_2 r_0^2}, \quad (3.3)$$

where N is the number of crystals in the shower, E_i is the energy deposited in the i th crystal, r_i is the radius in the plane perpendicular to the line pointing from the interaction point to the shower center, and $r_0 = 5$ cm is the average distance between two crystals. The energies are ordered $E_1 > E_2 > \dots > E_N$. The xlmomgam quantity is used to differentiate the electromagnetic showers from the hadronic showers. The electromagnetic shower typically deposits a large fraction of their energy in one or two crystals, whereas the hadronic showers tend to be more spread out.

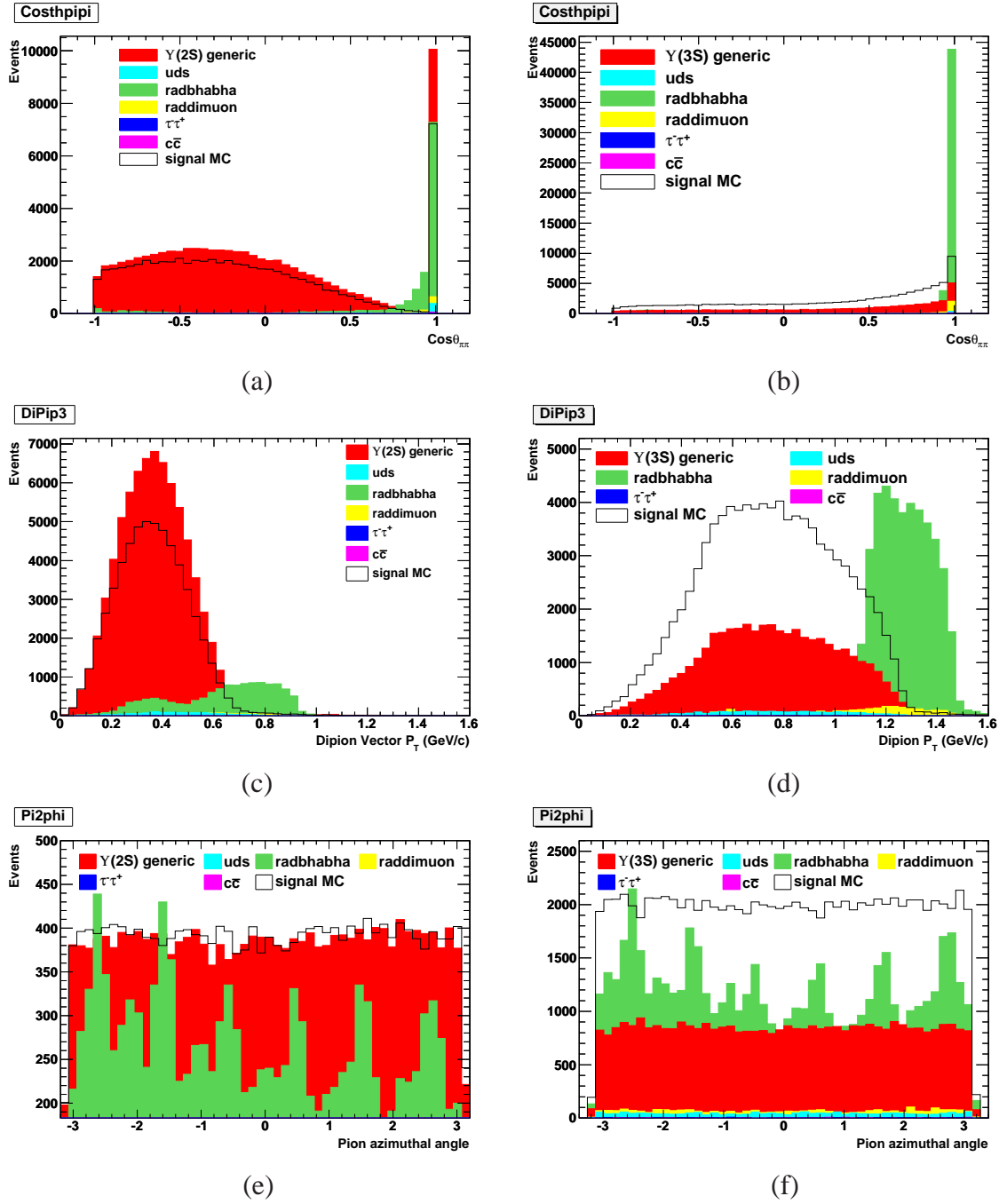


Figure 3.2: Di-pion related variables: (a, b) Cosine of angle between two pions in the laboratory frame (c, d) Transverse momentum of di-pion system in the laboratory frame and (e, f) Azimuthal angle of pion. Left plots are for $Y(2S)$ and right plots are for $Y(3S)$. All these variables are plotted after applying the pre-selection criteria.

- **Zmom42gam:** The Zernike A_{42} moment is defined as [78]:

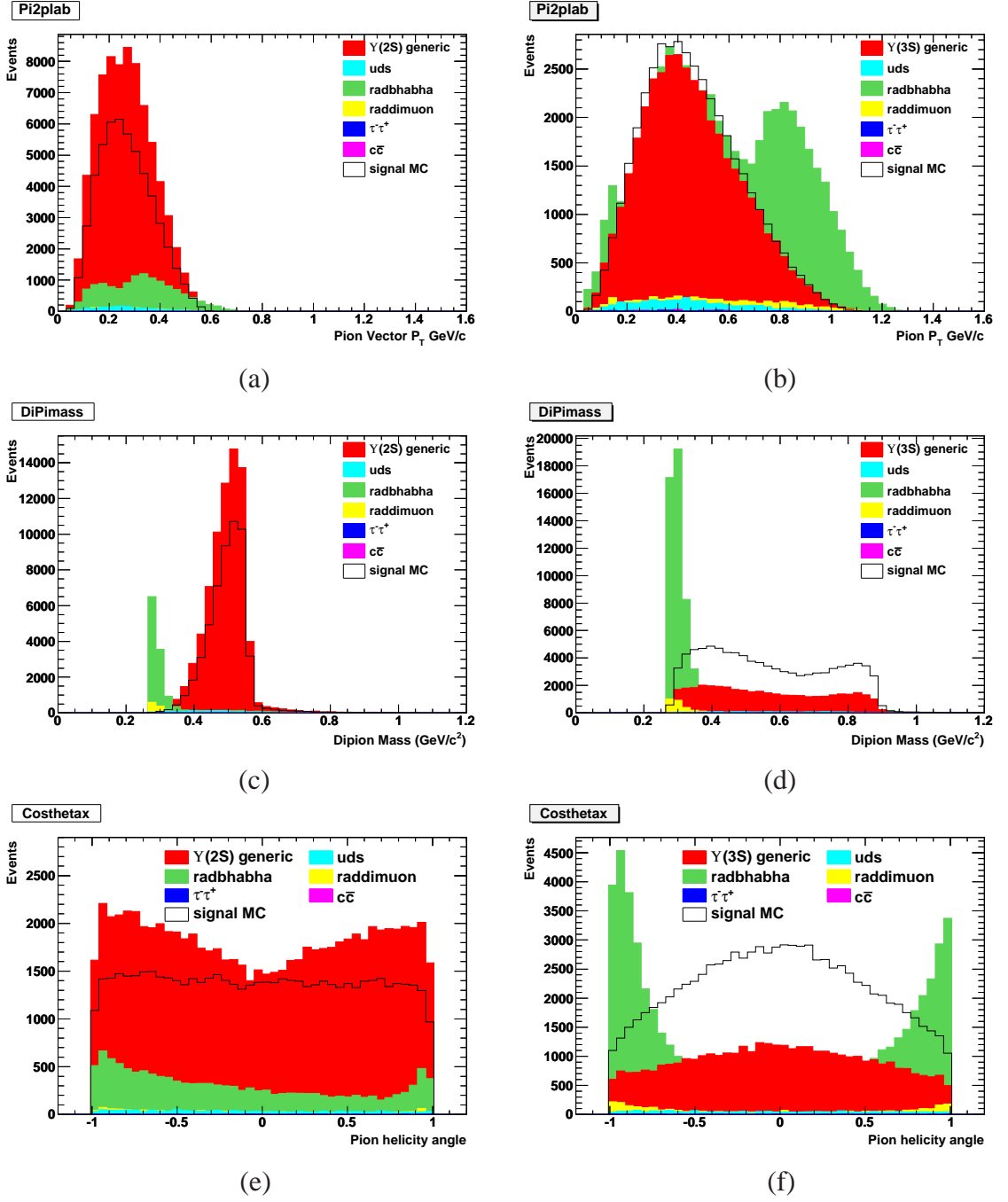


Figure 3.3: Di-pion related variables: (a, b) Transverse momentum of the pion (c, d) Di-pion invariant mass and (e, f) Cosine of pion helicity angle. Left plots are for $\Upsilon(2S)$ and right plots are for $\Upsilon(3S)$. All these variables are plotted after applying the pre-selection criteria.

$$A_{nm} = \sum_{k=1}^n (E_i/E) \cdot f_{nm}(\rho_i) e^{-im\phi_i}, \quad (3.4)$$

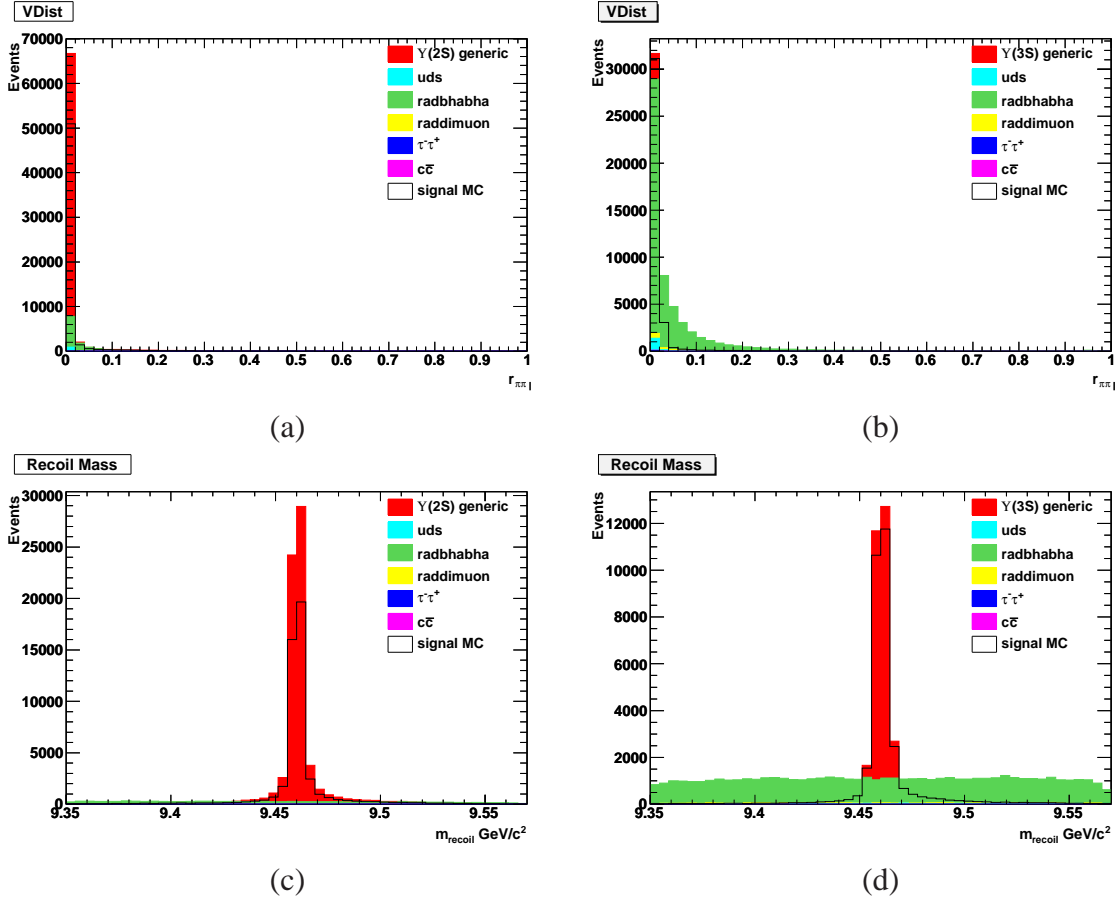


Figure 3.4: Di-pion related variables: (a, b) Transverse position of the di-pion vertex and (g, h) Mass recoiling against the di-pion system. Left plots are for $\Upsilon(2S)$ and right plots are for $\Upsilon(3S)$. All these variables are plotted after applying the pre-selection criteria.

where E_i is the energy deposited in the i^{th} crystal, E is the total energy deposited in the total crystals, f_{nm} are the polynomials of degree n and (ρ_i, ϕ_i) the location of the hit crystals in the EMC with respect to the center of the shower. The locations are defined in cylindrical coordinates with z -axis running from the beam spot to the centroid, with $\rho_i = r_i/R_0$ where $R_0 = 15$ cm. f_{nm} represents the Zernike function,

$$f_{nm}(\rho) = \sum_{s=0}^{(n-m)/2} \frac{(-1)^s (n-s)! \rho^n - 2s}{s! ((n+m)/2 - s)! ((n-m)/2 - s)!} \quad (3.5)$$

with $m \leq n$ and $(n-m)$ even. The Zmom42gam is used to characterize the azimuthal spread of the shower. It is also used to distinguish between electromagnetic showers

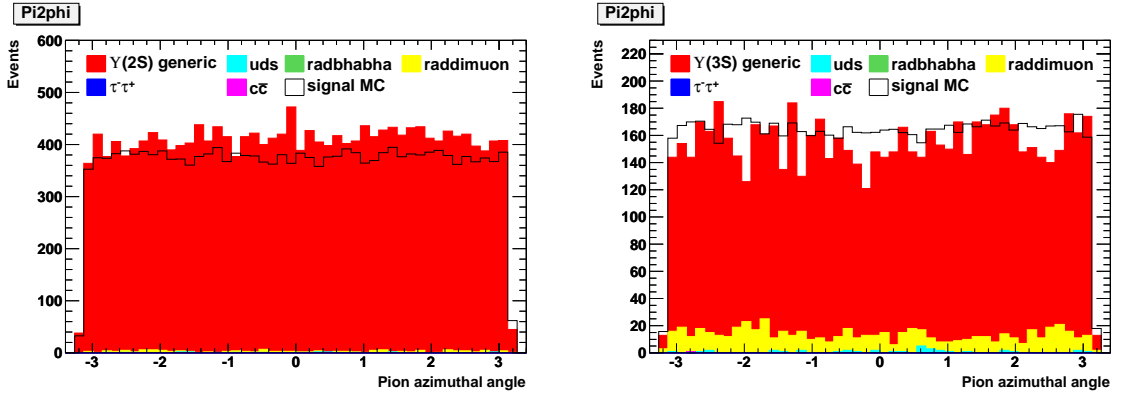


Figure 3.5: Azimuthal angle of pion for signal, $\Upsilon(3S, 2S)$ generic, uds, radiative bhabha, radiative di-muon, $\tau^+\tau^-$ and $c\bar{c}$ events for $\Upsilon(2S)$ (left) and $\Upsilon(3S)$ (right). These variables are plotted after applying the pre-selection criteria as well as requiring that either one of the tracks of the A^0 reconstruction using the muon PID must be identified as muon.

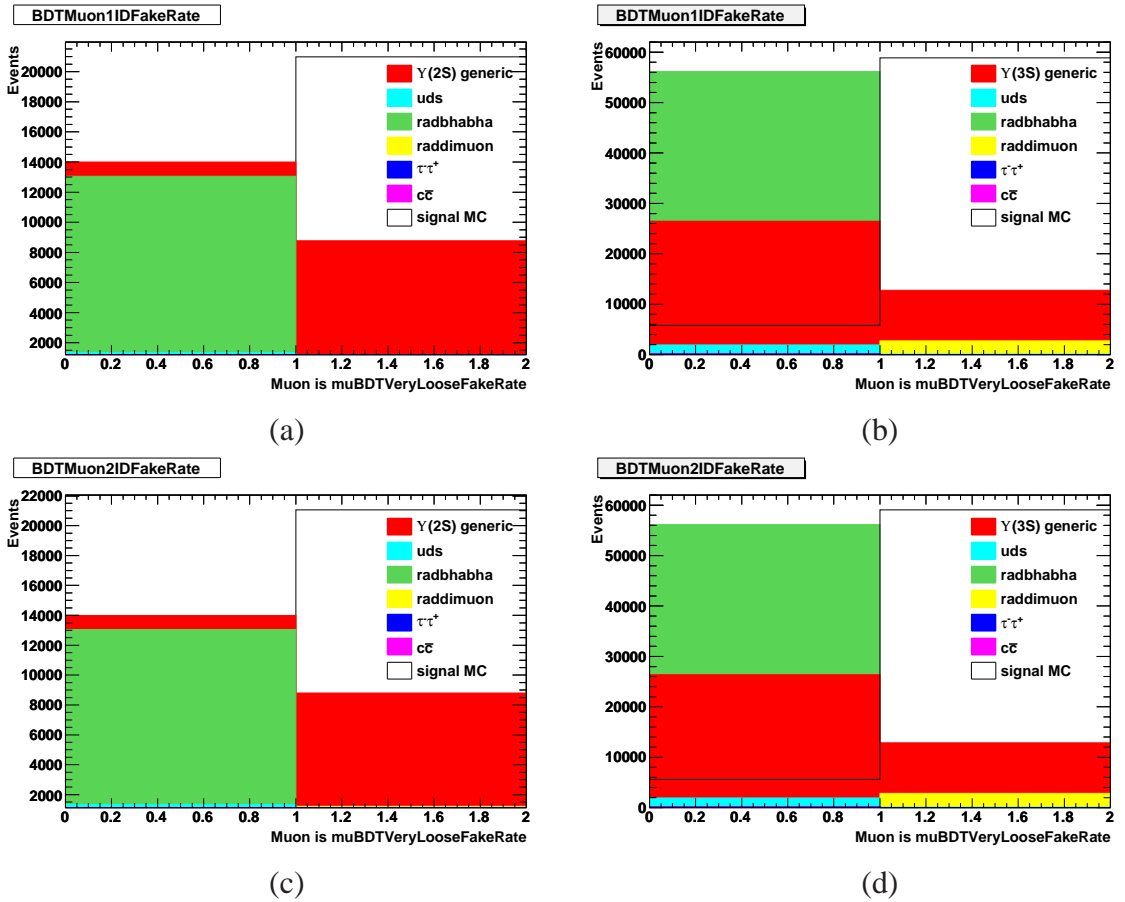


Figure 3.6: Muon particle-ID for $\Upsilon(3S, 2S)$ generic, uds, radiative bhabha, radiative di-muon, $\tau^+\tau^-$ and $c\bar{c}$ events for $\Upsilon(2S)$ (left) and $\Upsilon(3S)$ (right). We have plotted this variable at the pre-selection level.

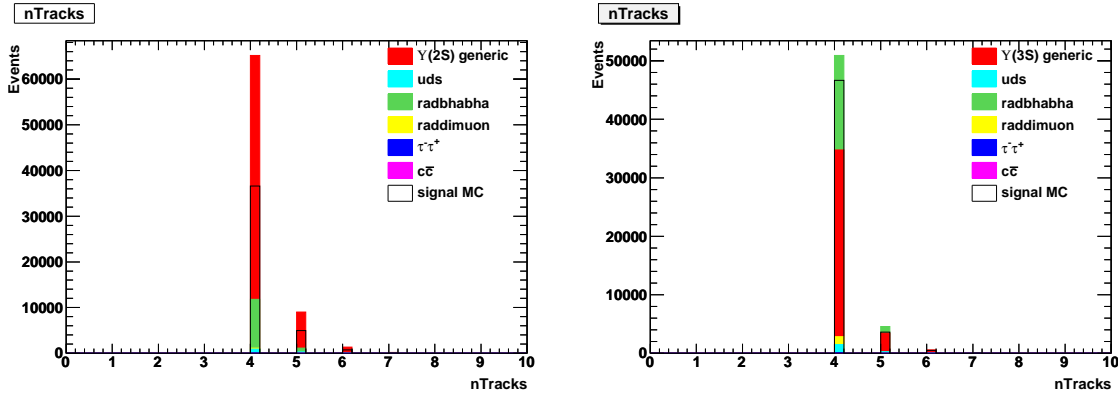


Figure 3.7: Number of ChargedTracks in signal, $\Upsilon(3S, 2S)$ generic, uds, radiative bhabha, radiative di-muon, $\tau^+\tau^-$ and $c\bar{c}$ events for $\Upsilon(2S)$ (left) and $\Upsilon(3S)$ (right). We have plotted this variable at the pre-selection level.

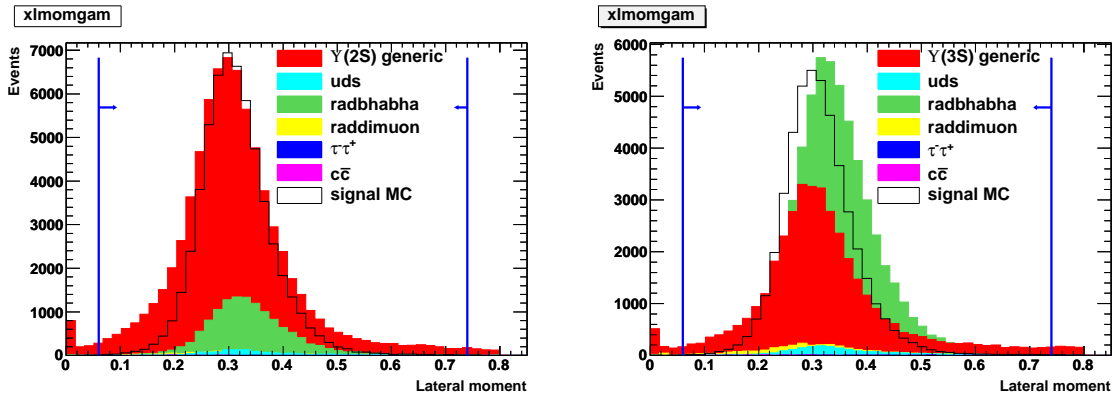


Figure 3.8: Lateral moment associated with photon from signal, $\Upsilon(3S, 2S)$ generic, uds, radiative bhabha, radiative di-muon, $\tau^+\tau^-$ and $c\bar{c}$ events for $\Upsilon(2S)$ (left) and $\Upsilon(3S)$ (right). We have plotted this variable at the pre-selection level.

and hadronic showers, because hadronic showers tend to be more irregular than electromagnetic shower.

We apply a loose selection cuts for the muon, track multiplicity and photon related variables. The selection criteria for the muon, track multiplicity and photon related variables for $\Upsilon(2S)$ and $\Upsilon(3S)$ datasets are summarized in Table 3.4. Figure 3.6, 3.7, 3.8 and 3.9 show the distributions of these variables.

We also apply a selection cut on the $\Upsilon(3S, 2S)$ kinematic fit χ^2 ($\chi^2_{\Upsilon(3S, 2S)} < 300$), which is calculated after fitting the entire decay chain using the CM beam energy constraints on the

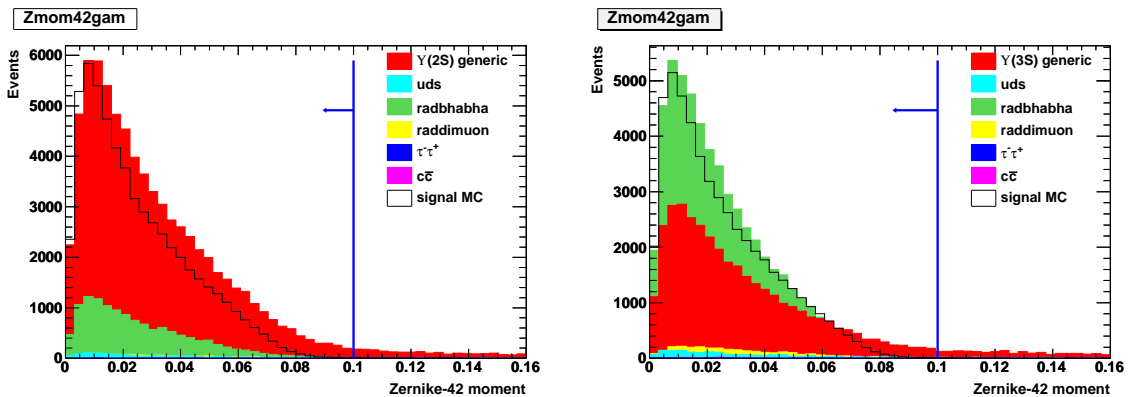


Figure 3.9: Zernike-42 moment associated with photon from signal, $\Upsilon(3S, 2S)$ generic, uds, radiative bhabha, radiative di-muon, $\tau^+\tau^-$ and $c\bar{c}$ events for $\Upsilon(2S)$ (left) and $\Upsilon(3S)$ (right). We have plotted this variable at the pre-selection level.

Variable name	Selection	
	$\Upsilon(3S)$	$\Upsilon(2S)$
Number of tracks	= 4	= 4
Lateral moment	[0.06, 0.75]	[0.06, 0.75]
Zernike-42 moment	< 0.1	< 0.1
Muon-ID	OR muon PID	OR muon PID

Table 3.4: Track multiplicity, photon and muon related selection variables.

$\Upsilon(3S, 2S)$ and mass constraints on the $\Upsilon(3S, 2S)$ and $\Upsilon(1S)$. Figure 3.10 shows the distribution of $\Upsilon(3S, 2S)$ kinematic fit χ^2 variable.

3.3.4 Multivariate Analysis

We use multivariate analysis (MVA) based BumpHunter algorithm and Random forest algorithm included in StatPatternRecognition [79] to optimize pions related variables. The full m_{red} range is used to optimize the pion related variables for both the datasets which are shown in Figure 3.11.

We split the data sample into 3 sub-samples, one for training set, one for validation set, and one for test set. The training and validation samples are used to train the MVAs. The test sample is used to check the performance of the MVAs after applying the selection criteria. Figure 3.12 and 3.13 show the correlation between the input variables for signal

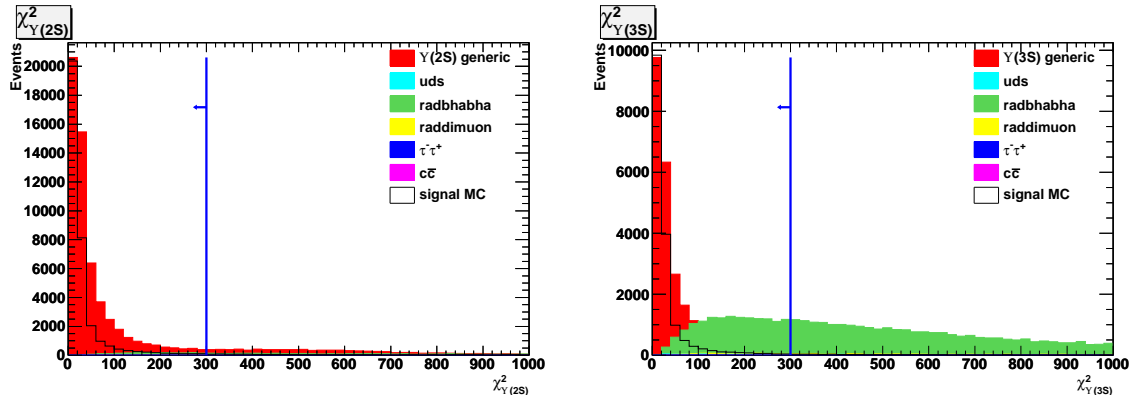


Figure 3.10: The $\Upsilon(2S, 3S)$ kinematic fit χ^2 for signal, $\Upsilon(3S, 2S)$ generic, uds, radiative bhabha, radiative di-muon, $\tau^+\tau^-$ and $c\bar{c}$ events for $\Upsilon(2S)$ (left) and $\Upsilon(3S)$ (right). We have plotted this variable at the pre-selection level.

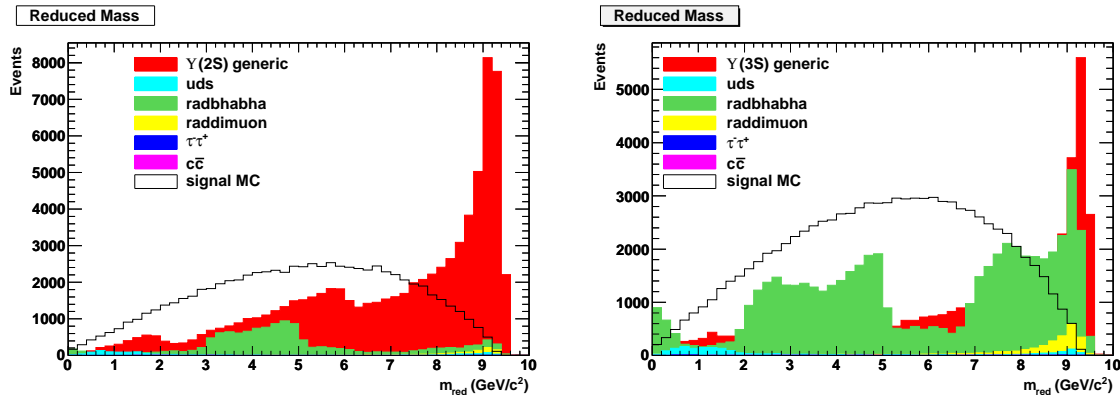


Figure 3.11: Reduced mass distribution for signal, $\Upsilon(3S, 2S)$ generic, uds, radiative bhabha, radiative di-muon, $\tau^+\tau^-$ and $c\bar{c}$ events for $\Upsilon(2S)$ (left) and $\Upsilon(3S)$ (right). We have plotted this variable at the pre-selection level.

and background in $\Upsilon(2S)$ and $\Upsilon(3S)$ datasets, respectively, which are used to train the BumpHunter and Random forest classifiers.

3.3.4.1 Variable selection optimization using BumpHunter classifier

The BumpHunter classifier is based on PRIM algorithm [80]. This classifier searches for a series of selection criteria that define an n-dimensional cuboid in n-dimensional variable space. Once a suitable region is found, the selection criteria are adjusted to optimize the figure of merit (FOM), such that the proportion of the number of events excluded by this adjustment does not exceed a fixed amount. This amount is known as the “peel” parameter.

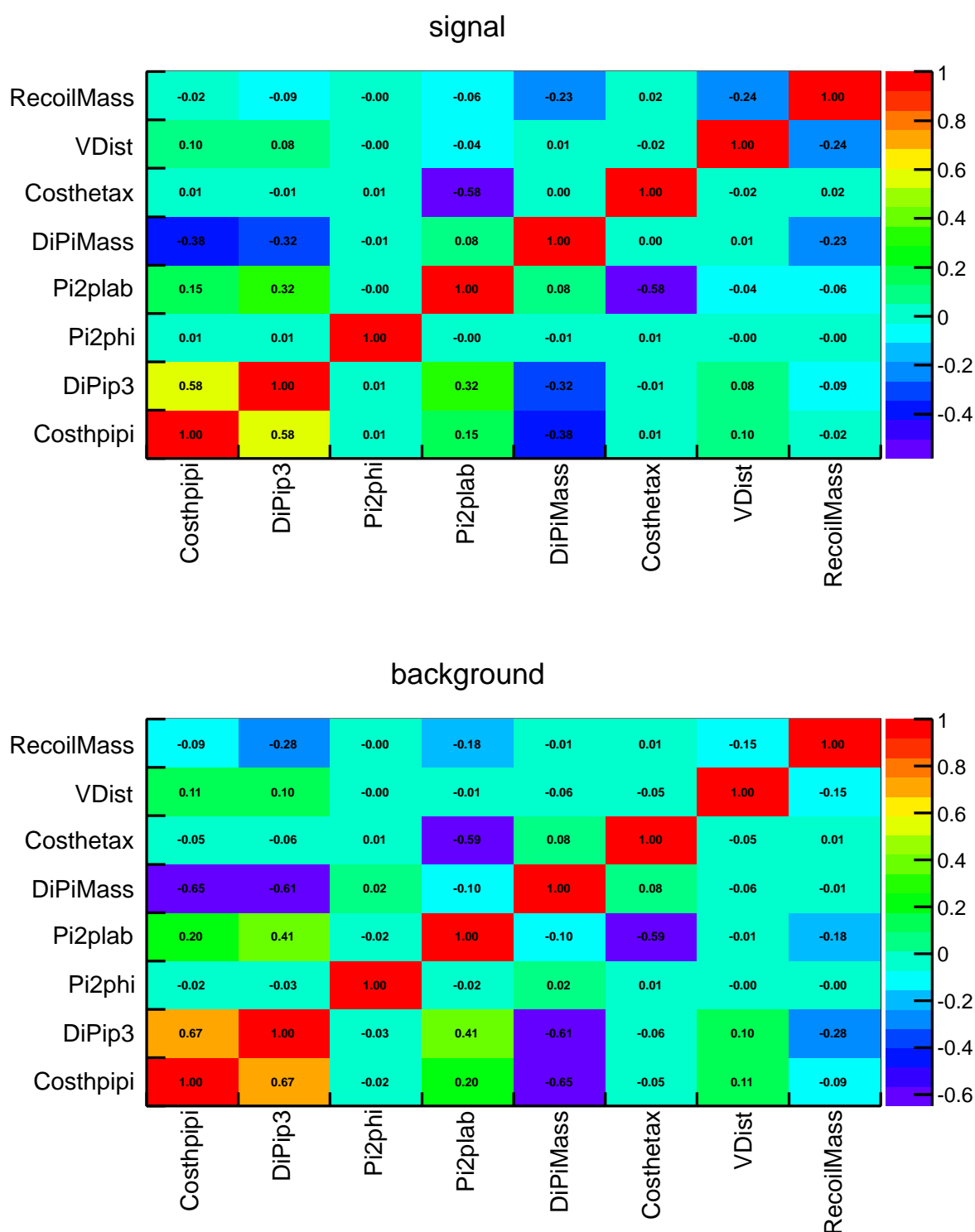


Figure 3.12: Correlation between MVA input variables in $\mathcal{V}(2S)$ for signal and Background MCs.

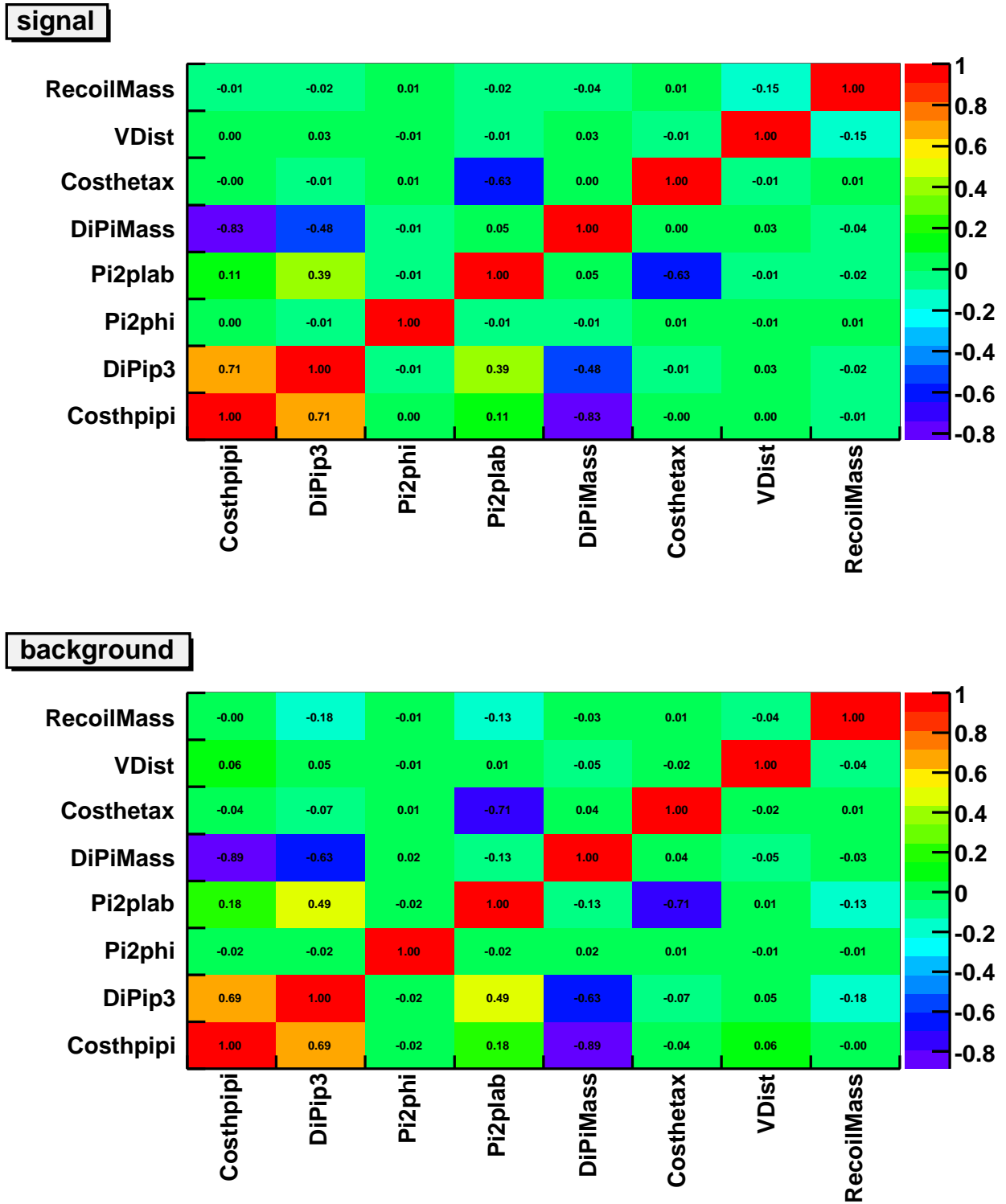


Figure 3.13: Correlation between MVA input variables in $\Upsilon(3S)$ for signal and Background MCs.

Variable name	Selection
Cosine of angle between two pions	> -0.999
Di-Pion transverse momentum	$< 1.239 \text{ GeV}/c$
Pion transverse momentum	$[0.070, 1.021] \text{ GeV}/c$
Pion helicity angle	$[-0.966, 0.947]$
Di-Pion mass	$[0.293, 0.894] \text{ GeV}/c^2$
Transverse position of di-pion vertex	$[5.50 \times 10^{-6}, 0.041]$
RecoilMass	$[9.451, 9.470] \text{ GeV}/c^2$

Table 3.5: Optimal set of cuts obtained from the BumpHunter with a peel parameter of 0.2 for $\Upsilon(3S)$.

The process is repeated until a cuboid is found which maximizes the FOM. In this analysis, we use Punzi FOM [81] for optimization which is defined as:

$$\frac{\epsilon}{0.5N_\sigma + \sqrt{B}}, \quad (3.6)$$

where N_σ is the number of standard deviations desired from the result, and ϵ and B are the average efficiency and background yield over a broad m_{A^0} range, respectively.

To train the BumpHunter classifier, we weigh the background MC by Run7 $\Upsilon(3S)$ onpeak luminosity (28.049 fb^{-1}) and weigh the signal MC by determining the number of expected signal event in our data while assuming a branching ratio of 10^{-6} . We train the BumpHunter MVA using training and validation sample of $\Upsilon(3S)$ to optimize the selection cuts. The peel parameter is varied between 1% and 95%. The optimal peel parameter (maximizing the FOM) is found to be 20%. The cuts determined by the algorithm are shown in Table 3.5.

We then apply these optimal cuts to the test sample and check the performance. We find 23925 signal MC events and 10009 background MC events for $\Upsilon(3S)$. This will be our benchmark numbers for a more complex multivariate analysis.

3.3.4.2 Variable selection optimization using Random Forest classifier

We use another advanced tool, the Random Forest (RF) classifier which was proposed by Breiman in 2001 [82]. RF is a method by which a number of decision trees are trained and

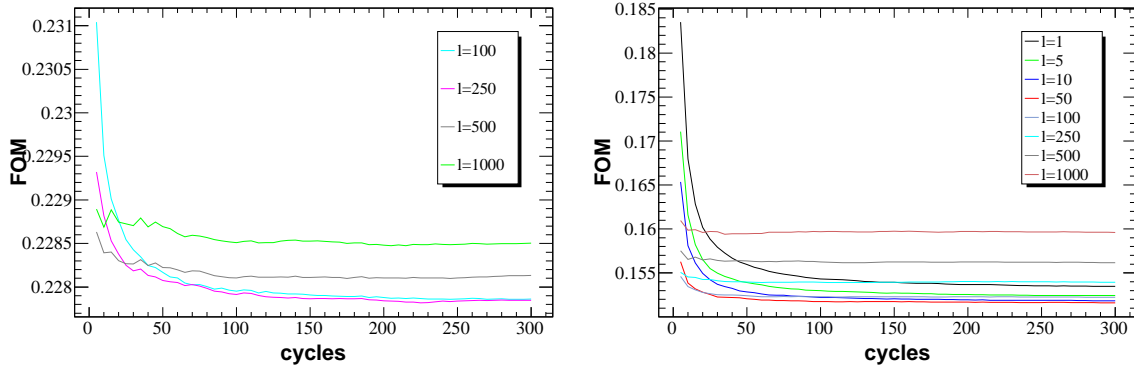


Figure 3.14: Figure of merit (FOM) of the RF vs. the number of trees grown (training cycles), for a variety of minimal events allowed per terminal node l for $\Upsilon(2S)$ (left) and $\Upsilon(3S)$ (right). We find the lowest FOM: $l=250$ for $\Upsilon(2S)$ and $l=50$ for $\Upsilon(3S)$.

the output of the algorithm is taken as the weighted vote of the output of each decision trees. Unlike the BumpHunter, a decision tree recursively splits training data into rectangular region (nodes). For each node, the tree examines all possible binary splits in each dimension and selects the one with the optimized FOM. In our case, the decision tree sets the weights for the vote, to maximize the Gini index (the FOM for this approach). StatPatternRecognition uses negative Gini-index ($= -2p.q$), where p and $q = 1 - p$ are fractions of correctly and incorrectly classified events in each node. The Gini index is related to the minimization of the loss of events from each category. Each training cycle grows a decision tree from a random set of 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 which are allowed to populate a terminal node of the tree (a node with no further splits). We fix the number of trees to 300 and try a variety of minimal events per terminal node, which we denote by “ l ”. Figure 3.14 shows the resulting training curves for the FOM vs. training cycle. We find the best performance (lowest FOM) for $l = 50$ for $\Upsilon(3S)$ and $l = 250$ for $\Upsilon(2S)$. The output of the RF, for both signal and combined background MC is shown in Figure 3.15. We use these RF outputs to calculate the survived signal and background events.

We cross-check the performance of RF algorithm against the BumpHunter algorithm for $\Upsilon(3S)$. We fix the cut on the RF output to a particular value to achieve the same background yield as BumpHunter and compare the signal MC yield and we find that the RF returns 5.88%

more signal MC events (25420 events) for the same background for $\Upsilon(3S)$. We shall use RF classifier for further analysis for both $\Upsilon(3S)$ and $\Upsilon(2S)$ datasets.

We optimize the cut on the RF discriminant using the Punzi figure of merit (FOM):

$$\frac{\epsilon}{0.5N_\sigma + \sqrt{\sum_i (B_i \times w_i)}}, \quad (3.7)$$

where $N_\sigma = 3$, ϵ is the average efficiency, B_i and w_i are the number of background events and background weights (for different $i = \Upsilon(3S, 2S)$ generic, uds, radiative bhabha, radiative di-muon, $\tau^+\tau^-$ and $c\bar{c}$) respectively. The weight of each data-set is defined as the ratio of two quantities a/b, where a is Run7 $\Upsilon(3S, 2S)$ onpeak luminosity and b is background sample luminosity. The RF output for signal and background MCs in the test sample is shown in Figure 3.15 for both $\Upsilon(3S)$ and $\Upsilon(2S)$ datasets. The optimized plot for the Punzi's FOM vs. RF discriminant is also shown in Figure 3.15(b) and Figure 3.15(d) for $\Upsilon(2S)$ and $\Upsilon(3S)$ sample, respectively.

3.3.5 Final selection

The final selection criteria for the $\Upsilon(3S, 2S) \rightarrow \pi^+\pi^-\Upsilon(1S)$; $\Upsilon(1S) \rightarrow \gamma A^0$; $A^0 \rightarrow \mu^+\mu^-$ analysis includes the following:

- Track multiplicity, photon and muon related cuts as described in Table 3.4.
- Pion related variables using RF classifier : $RF > 0.568$ for $\Upsilon(3S)$ and $RF > 0.388$ for $\Upsilon(2S)$.
- $\Upsilon(3S, 2S)$ kinematic fit $\chi^2 : \chi^2_{\Upsilon(3S, 2S)} < 300$.

We then apply the optimal selection cuts to the test samples for both signal and background MCs. The signal MC sample is used to compute the signal selection efficiency as a function of m_{A^0} after applying all the selection cuts. We use m_{red} distribution to perform the maximum likelihood (ML) fit for the signal yield extraction from data, the result of which will be presented in the next chapters. The signal efficiency varies between 38.3% (40.4%) and 31.7% (31.6%) for $\Upsilon(2S)$ ($\Upsilon(3S)$), and decreases monotonically with m_{A^0} . Figure 3.16

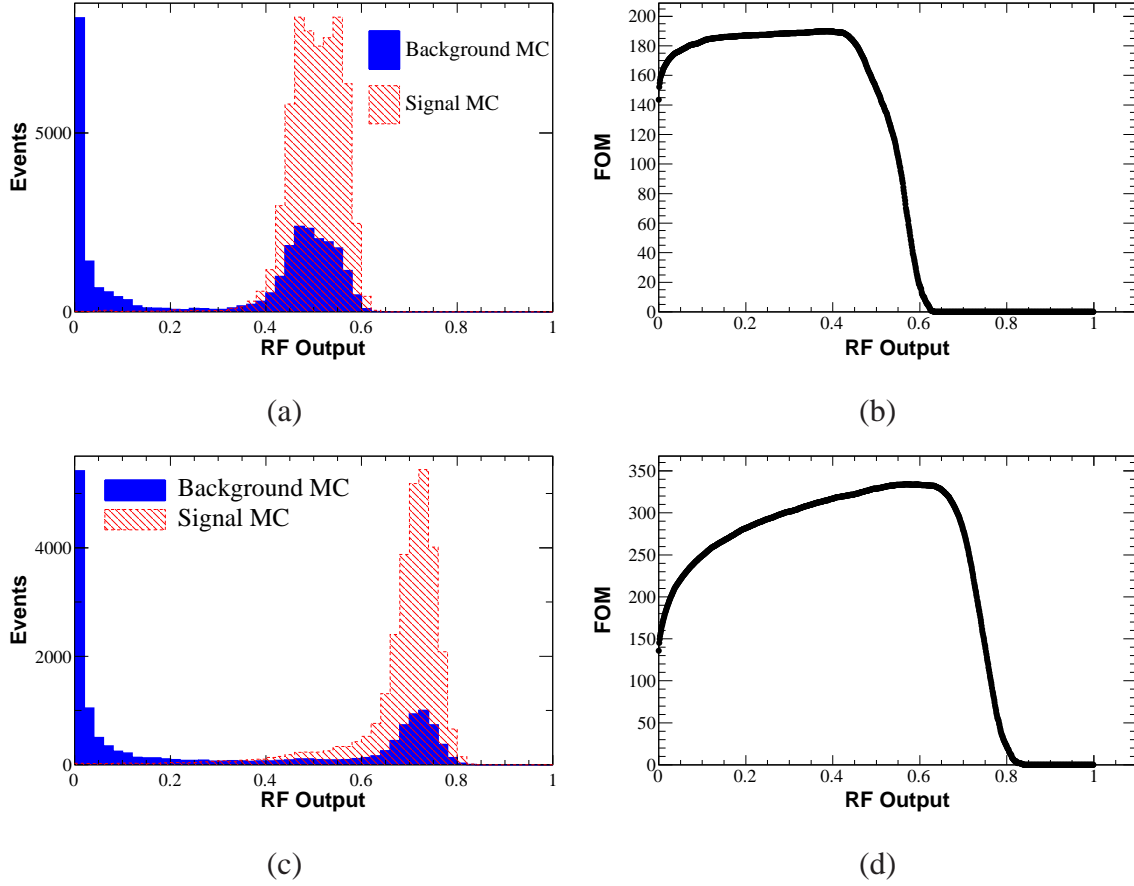


Figure 3.15: (a,c) The output of the RF for both signal and combined background MC for $\Upsilon(2S, 3S)$ and (b,c) Punzi's FOM as a function of RF cut. The backgrounds are normalized by $\Upsilon(3S, 2S)$ onpeak data sets. Top plots are for $\Upsilon(2S)$ and bottom plots are for $\Upsilon(3S)$. The optimized cut is $\text{RF} > 0.388$ (0.568) for the $\Upsilon(2S)$ ($\Upsilon(3S)$) dataset.

shows the remaining background events in the $\Upsilon(3S)$ and $\Upsilon(2S)$ data samples, after scaling up the number of events by 3 to represent the full sample. The background is dominated by $\Upsilon(2S, 3S)$ generic decays, rest of the other sources are negligible [83].

The m_{recoil} distributions of generic events and low onpeak datasets show that about 93% of the $\Upsilon(3S, 2S)$ generic events decay via $\Upsilon(3S, 2S) \rightarrow \pi^+\pi^-\Upsilon(1S), \Upsilon(1S) \rightarrow \text{anything}$ (Figure 3.17). Using MC-Truth information of the survived background events, it is found that about 99% of the events decay via $\Upsilon(1S) \rightarrow \gamma\mu^+\mu^-$. Figure 3.18 shows a MC-Truth Boolean distributions for $\Upsilon(1S) \rightarrow \gamma\mu^+\mu^-$ decays for both the datasets after applying all the selection cuts.

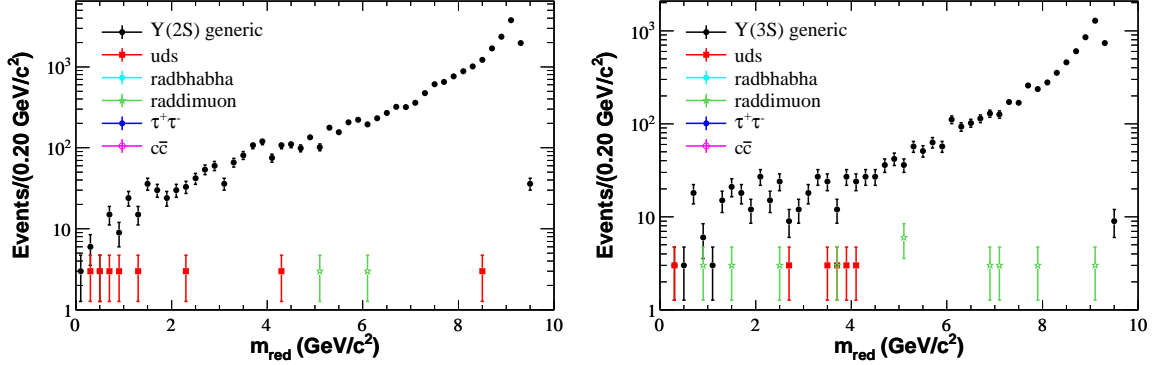


Figure 3.16: m_{recoil} distribution of the remaining background MCs events in the $\Upsilon(2S, 3S)$ datasets after applying all the optimal selection cuts. The left plot is for $\Upsilon(2S)$ and right plot is for $\Upsilon(3S)$. The test sample is scaled up by three to represent the full data samples in both $\Upsilon(2S)$ and $\Upsilon(3S)$. The most dominant remaining background is $\Upsilon(2S, 3S)$ generic decays in both the datasets. Contributions from other backgrounds are negligible.

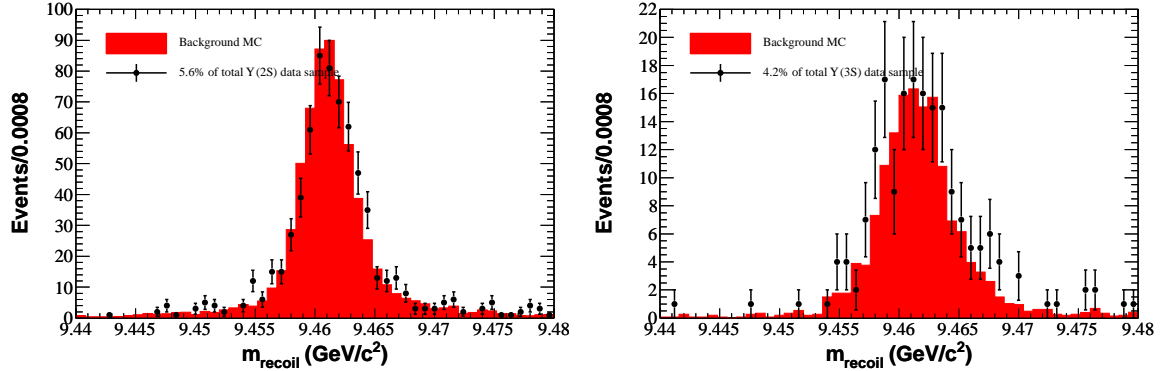


Figure 3.17: m_{recoil} distributions for generic and low onpeak datasets after applying all the selection criteria. Left plot is for $\Upsilon(2S)$ and right plot is for $\Upsilon(3S)$. The mean of the recoil mass in MC has been corrected after comparing the recoil mass distributions in a control samples of data and MC, the details of which can be found in section 3.4.

3.4 Corrections of mean and width of m_{recoil}

After the event reconstruction, it was observed that the mean and sigma of the m_{recoil} distribution is shifted by 1.0 MeV/ c^2 in MC, while compared to data. We use a control sample of $\Upsilon(2S, 3S) \rightarrow \pi^+\pi^-\Upsilon(1S)$, $\Upsilon(1S) \rightarrow \mu^+\mu^-$ in data and MC to study the mean and width value of m_{recoil} . We apply the following selection criteria to both data and MC after reconstructing the events:

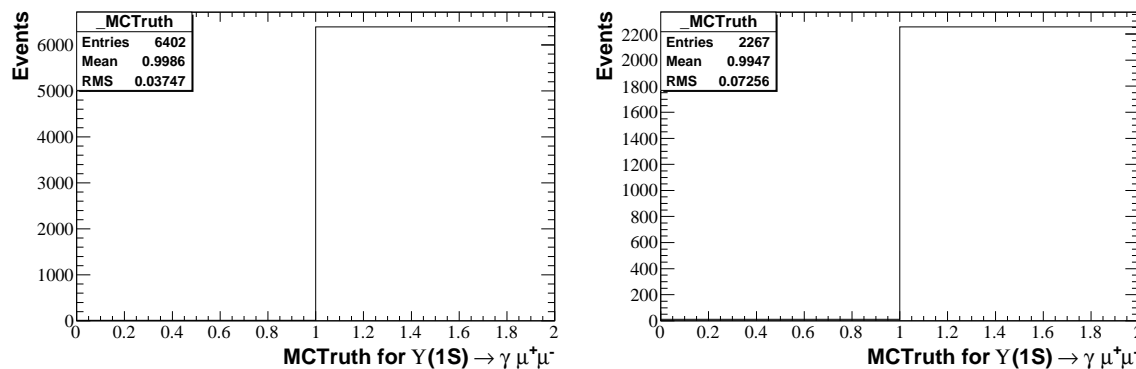


Figure 3.18: MC-Truth Boolean distributions for $\Upsilon(1S) \rightarrow \gamma\mu^+\mu^-$ decays. Left plot is for $\Upsilon(2S)$ and right plot is for $\Upsilon(3S)$.

- Two pions must not be misidentified as electron using a particle-ID algorithm where the π -to- e mis-identification rate is about 0.1%.
- Both leptons must be identified as muons by a muon particle-ID algorithms.
- CM energy and momentum are within $|\Delta E| < 0.2$ GeV and $|\Delta P| < 0.2$ GeV/c.
- The number of the charged tracks must be equal to four.
- RF selection cuts of the $\Upsilon(2S, 3S)$ datasets.

We use a sum of two Crystal Ball (CB) functions [84] with opposite side tails to model the m_{recoil} . The detail description about the CB function has been presented in the section 4.2 in chapter 4. The fit to the m_{recoil} distributions in both data and MC samples for both $\Upsilon(2S)$ and $\Upsilon(3S)$ datasets are shown in Figure 3.19 and 3.20, respectively. The mean of the recoil mass distribution in data appears to be shifted by less than 1 MeV/ c^2 and is also wider than MC, for both the $\Upsilon(2S, 3S)$ datasets. We correct the mean and width of the recoil mass distribution in MC by the observed difference in data and MC.

3.5 Chapter Summary

In this chapter, we have described the event reconstruction and the methods of event selection using the different multivariate techniques. The datasets used for this analysis are also

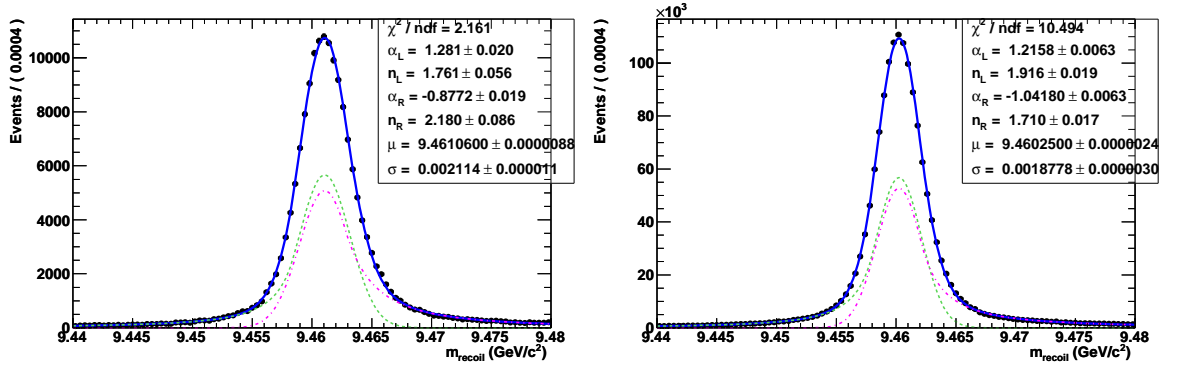


Figure 3.19: m_{recoil} distribution in $\Upsilon(2S) \rightarrow \pi^+\pi^-\Upsilon(1S)$, $\Upsilon(1S) \rightarrow \mu^+\mu^-$ events after applying all the selection cuts as mentioned in the section 3.4 including the RF selection cuts of $\Upsilon(2S)$. Left plot is for the data and right plot is for MC.

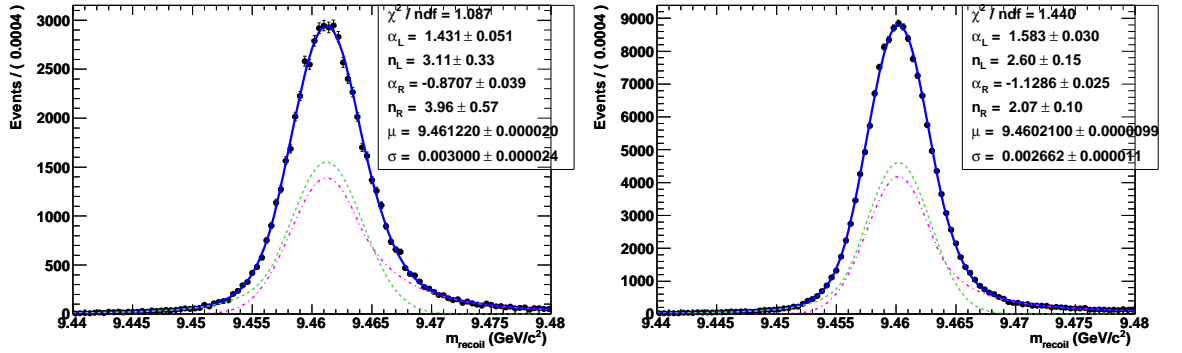


Figure 3.20: m_{recoil} distribution in $\Upsilon(3S) \rightarrow \pi^+\pi^-\Upsilon(1S)$, $\Upsilon(1S) \rightarrow \mu^+\mu^-$ events after applying all the selection cuts as mentioned in the section 3.4 including the RF selection cuts of $\Upsilon(3S)$. Left plot is for the data and right plot is for MC.

presented. Finally, we have discussed the remaining backgrounds after applying all the selection criteria. In the following chapter, we will describe the signal and background probability density functions (PDFs), which are used to extract the signal events from data.

Chapter 4

Maximum Likelihood Fit

This chapter begins with an overview of the maximum likelihood (ML) fit used to extract the signal events from the data [85, 86]. The RooFit [87] and RooRarFit [88] packages are used to conduct the 1d unbinned ML fit to the m_{red} distributions in the data samples. The signal and background probability density functions (PDFs) are developed using signal MC samples generated at 26 m_{A^0} points and the combined background MC, respectively. The fit validations are done using a cocktail samples of $\mathcal{Y}(2S, 3S)$ low onpeak datasets and $\mathcal{Y}(2S, 3S)$ generic MCs, as well as a large number of the Toy MC experiments with different embedded signal events at selected m_{A^0} points. The bias of the fit is considered as an additive systematic uncertainty. Finally, this chapter describes the trial factor study used to compute the true significance i.e., the probability for pure background event to fluctuate up to a given value of the signal yield.

4.1 Theoretical overview of the ML fit

The ML fit is a technique used to estimate the values of the parameters for a given finite sample of the data. Suppose a measurement of the random variable x is repeated several times for a finite values of x_1, \dots, x_n , where each x_i follow a probability density function (PDF) of $f(x_i; \theta)$ for a particular value of θ . Then the likelihood function in the interval of $[x_i, x_i + dx_i]$ is defined as:

$$\mathcal{L}(\theta) = \prod_{i=1}^n f(x_i; \theta). \quad (4.1)$$

The likelihood defined by equation 4.1 is called an unbinned likelihood, which is evaluated at each data point and no binning of the data is needed. In practice one often uses the negative log-likelihood (NLL)

$$-\log \mathcal{L}(\theta) = -\sum_{i=1}^n \log f(x_i; \theta), \quad (4.2)$$

that makes easier to estimate a parameter value while minimizing the NLL function. The unbinned ML estimator $\hat{\theta}$ for a parameter vector θ is defined as the value of θ for which the likelihood is maximal, or equivalently the negative log-likelihood is minimal.

The statistical uncertainty on a parameter θ is defined as the square-root of the variance. The ML estimator for the variance on θ is given by the second derivative of the log-likelihood at $\theta = \hat{\theta}$.

$$\hat{\sigma}(\theta)^2 = \hat{V}(\theta) = \left(\frac{d^2 \log(\mathcal{L}(\theta))}{d^2 \theta} \right)^{-1}, \quad (4.3)$$

In case there are multiple parameters, the variance of the ensemble of parameters is represented by the covariance matrix, which is defined as:

$$V(\theta, \theta') = \langle \theta \theta' \rangle - \langle \theta \rangle \langle \theta' \rangle = \left(\frac{\partial^2 \log(\mathcal{L}(\theta, \theta'))}{\partial \theta \partial \theta'} \right)^{-1}, \quad (4.4)$$

which can also be expressed in terms of variance and a correlation matrix

$$V(\theta, \theta') = \sqrt{V(\theta)V(\theta')} \cdot \rho(\theta, \theta'), \quad (4.5)$$

Here $\rho(\theta, \theta')$ expresses the correlation between the parameters of θ and θ' and have their values in the range of $[-1, 1]$.

4.1.1 Extended ML Fit

The extended ML function includes an extra factor for the probability of obtaining a sample of size N from a Poisson distribution of a mean ν

$$\mathcal{L}(\nu, \theta) = \prod_{i=1}^n f(x_i; \theta) \cdot e^{-\nu} \frac{\nu^N}{N!}, \quad (4.6)$$

where the ν describes the expected rate at which the total number of events are produced. The extended ML function is used to determine the number of signal and background events in a given data sample through a fit. The most straightforward approach to such an analysis is to define a composite probability density function (PDF) of $\mathcal{L}(x, \theta, \theta')$ as follows:

$$\mathcal{L}(x, \theta, \theta') = \frac{N_S}{N_S + N_B} \cdot S(x; \theta) + \frac{N_B}{N_S + N_B} \cdot B(x; \theta'), \quad (4.7)$$

where N_S and N_B are the number of signal and background events, respectively, $N = N_S + N_B$ the total number of events in the data sample, and $S(x; \theta)$ and $B(x; \theta')$ the PDFs of signal and background, respectively. A minimization of the extended ML fit estimates the yield of the N_S and N_B .

4.2 Signal PDF

In this analysis we perform an one-dimensional extended ML fit to the m_{red} distribution to extract the number of signal events. The m_{red} distributions of the signal are parametrized by a sum of two Crystal Ball (CB) [84] functions with opposite-side tails. The CB function is given by,

$$f(x|\mu, \sigma, \alpha, n) = C \cdot \begin{cases} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), & \frac{x-\mu}{\sigma} > -\alpha \\ \left(\frac{n}{|\alpha|}\right)^n \exp\left(-\frac{\alpha^2}{2}\right) \cdot \left(\frac{n}{|\alpha|} - |\alpha| + \frac{x-\mu}{\sigma}\right)^{-n}, & \frac{x-\mu}{\sigma} \leq -\alpha \end{cases} \quad (4.8)$$

where α determines where the usual Gaussian turns into a power function with the tail parameter n , and C is overall normalization. We constrain the mean (μ) parameters of the two CB functions to be the same, and for $m_{A^0} > 0.5 \text{ GeV}/c^2$ we also fix the relative weight of each

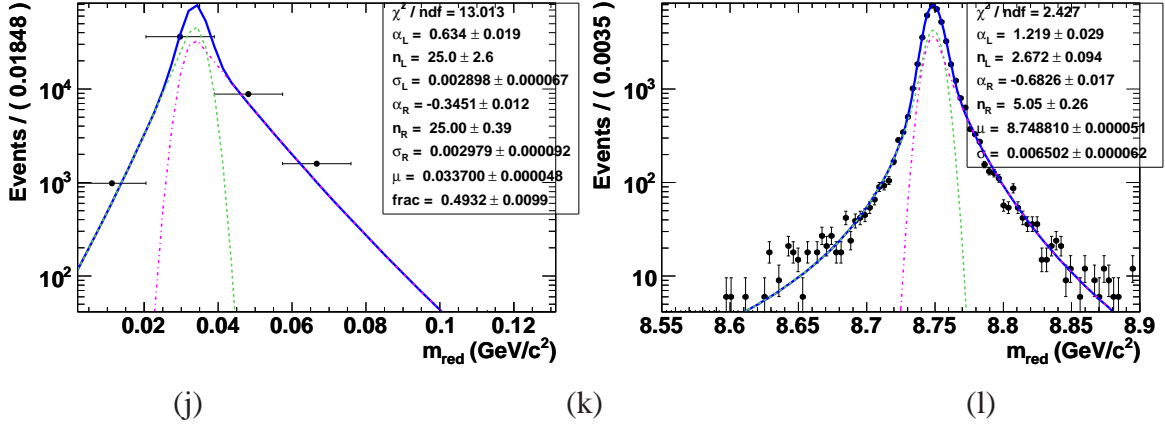


Figure 4.1: Signal PDFs for the Higgs mass of (a) $m_{A^0} = 0.214 \text{ GeV}/c^2$ and (b) $m_{A^0} = 8.75 \text{ GeV}/c^2$.

CB to $\text{frac} = 0.5$. For $m_{A^0} > 0.5 \text{ GeV}/c^2$ we also constrain the width (σ) parameters of the two CB functions to be same. Thus, in this mass range there are six floated parameters: mean (μ), sigma (σ), two tail cutoffs (α_L, α_R), and two powers (n_L, n_R). For $m_{A^0} \leq 0.5$, we float the two widths σ_L and σ_R separately, for a total of seven free parameters. We fit over fixed intervals in the mass regions: $0.002 \leq m_{\text{red}} \leq 1.85 \text{ GeV}/c^2$ for $0.212 \leq m_{A^0} \leq 1.50 \text{ GeV}/c^2$, $1.40 \leq m_{\text{red}} \leq 5.6 \text{ GeV}/c^2$ for $1.502 \leq m_{A^0} < 5.36 \text{ GeV}/c^2$ and $5.25 \leq m_{\text{red}} \leq 7.3 \text{ GeV}/c^2$ for $5.36 \leq m_{A^0} \leq 7.10 \text{ GeV}/c^2$. Above this range, we use sliding intervals $\mu - 0.2 < m_{\text{red}} < \mu + 0.15 \text{ GeV}/c^2$.

The fit to the m_{red} distributions for the signal MC for the selected mass points are shown in Figure 4.1. Rest of the other plots are shown in Appendix A in Figure A.1 – A.3 for $\Upsilon(2S)$ and in Figure A.4 – A.5 for $\Upsilon(3S)$ dataset. The summary of the PDF parameters for both datasets are shown in Figure 4.2 – 4.3. Figure 4.4 shows the signal selection efficiency as a function m_{A^0} for both the datasets. The PDF parameters of the signal are interpolated linearly from the known m_{A^0} points.

4.3 Background PDF

The background PDF in the range of $m_{A^0} \leq 1.50 \text{ GeV}/c^2$ is modelled using a MC sample of $\Upsilon(2S, 3S) \rightarrow \pi^+ \pi^- \Upsilon(1S)$, $\Upsilon(1S) \rightarrow (\gamma) \mu^+ \mu^-$ decays, which is described by a threshold

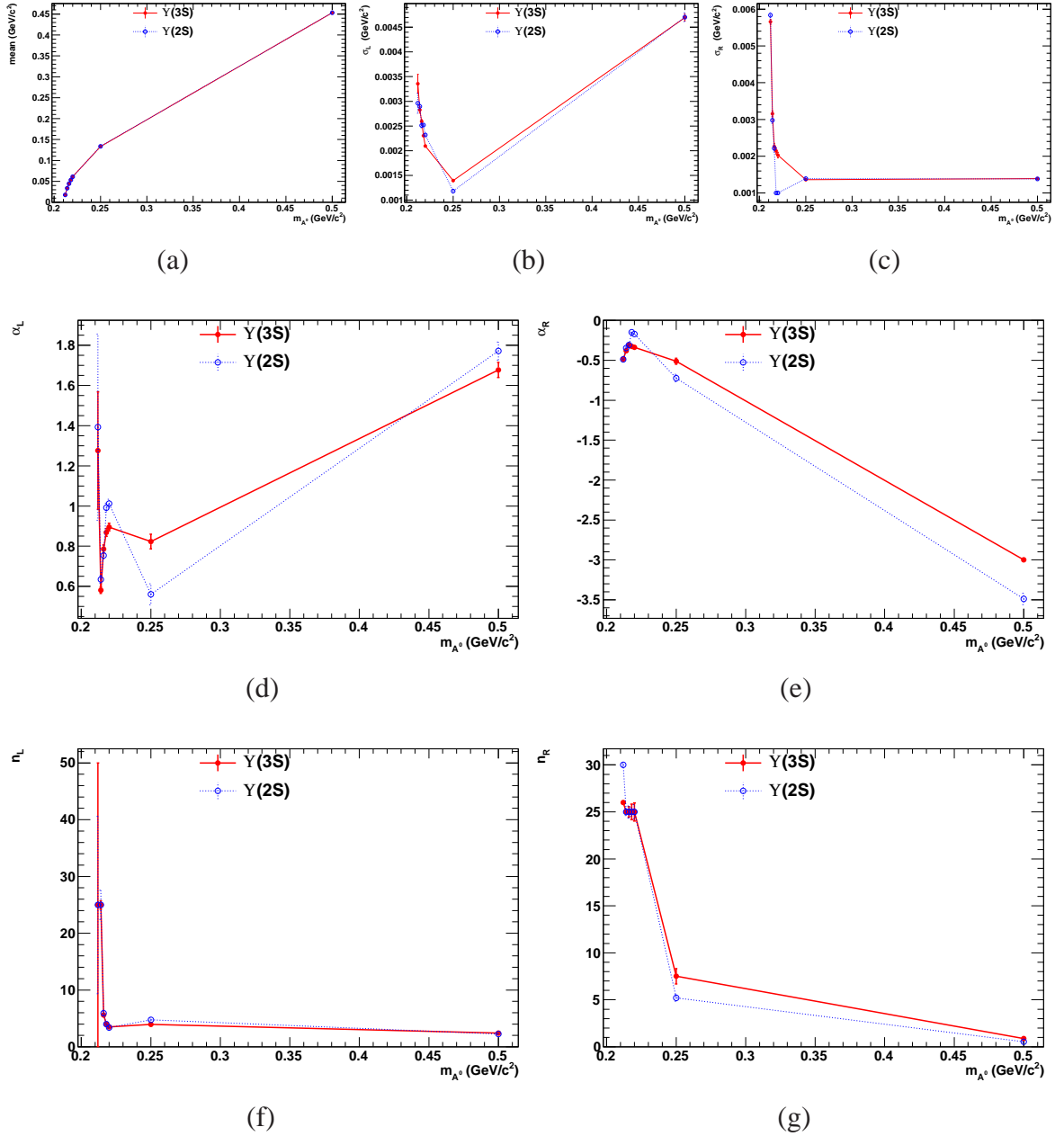


Figure 4.2: Parameters of the 1d ML fit to m_{red} distributions for signal MC for $m_{A^0} \leq 0.5$ GeV/c²: (a) mean of both CB functions (b) width of the “left” CB shapes (c) width of the “right” CB shapes (d) cutoff of the “left” CB (e) cutoff of the “right” CB (f) power of the “left” CB and (g) power of the “right” CB.

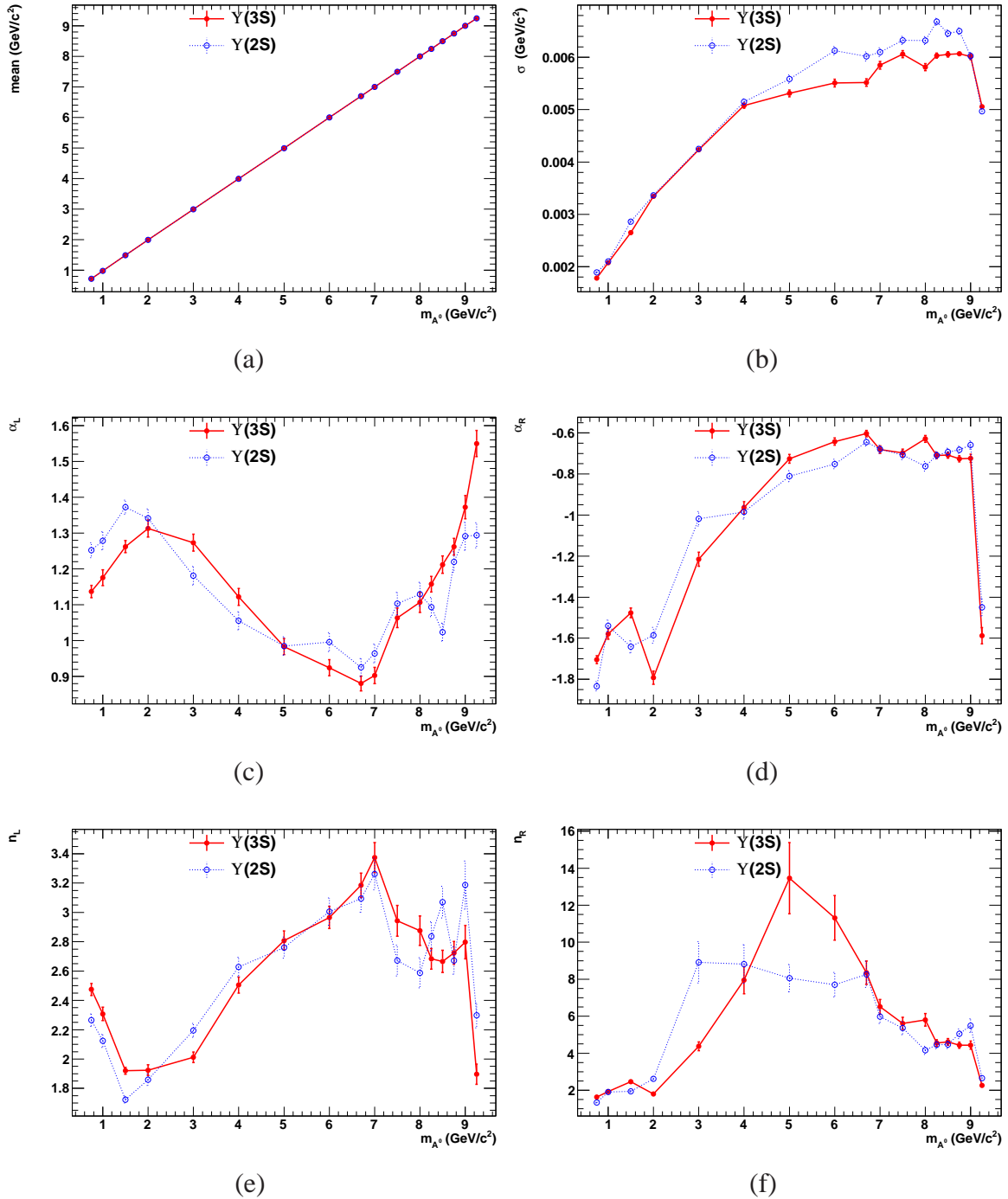


Figure 4.3: Parameters of the 1d ML fit to m_{red} distributions for signal MC for $m_{A^0} > 0.5$ GeV/c^2 : (a) mean of both CB functions, (b) width of both CB shapes, (c) cutoff of the “left” CB, (d) cutoff of the “right” CB, (e) power of the “left” CB, and (f) power of the “right” CB.

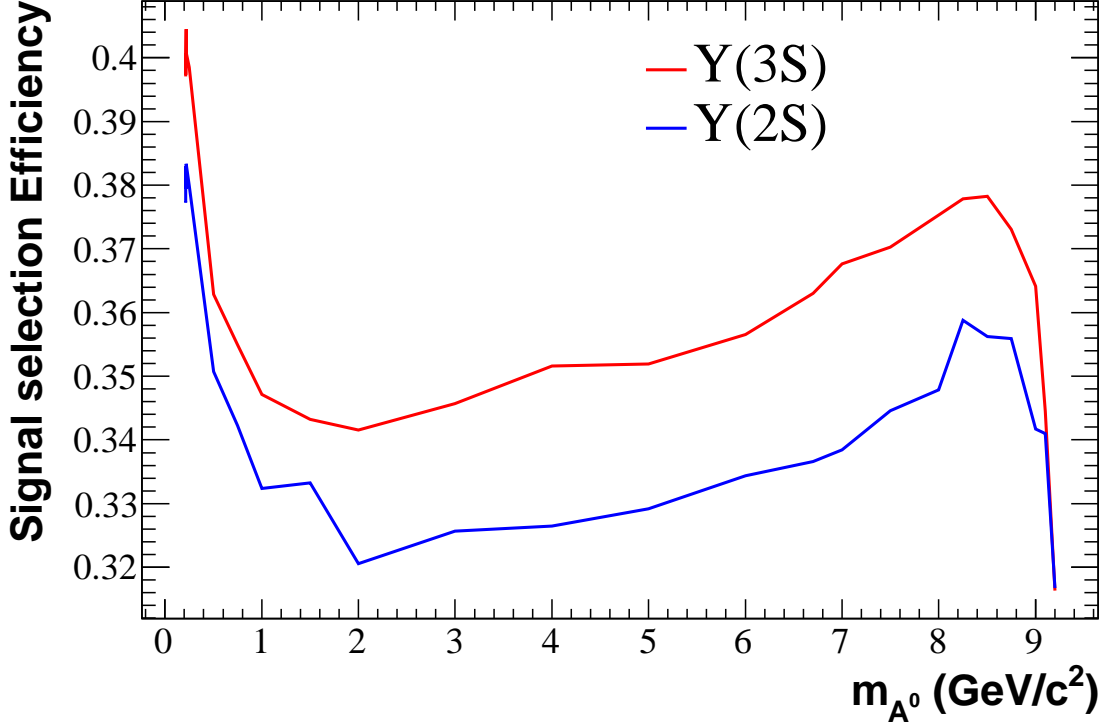


Figure 4.4: The signal selection efficiency as a function of m_{A^0} .

function

$$f(m_{\text{red}}) \propto [\text{Erf}(s(m_{\text{red}} - m_0)) + 1] + \exp\left(\sum_{\ell=0}^1 c_{\ell} m_{\text{red}}^{\ell}\right), \quad (4.9)$$

where s is a threshold parameter and m_0 is determined by the kinematic end point of the m_{red} distribution, and c_{ℓ} is the coefficient of ℓ^{th} order polynomial function. The background PDF is described by a second order Chebyshev polynomial in the range of $1.502 \leq m_{A^0} \leq 7.10$ GeV/c², and a first order Chebyshev polynomial for $m_{A^0} > 7.10$ GeV/c². The plots of background PDF near the threshold mass region are shown in Figure 4.5 for both $\Upsilon(2S, 3S)$ datasets. Rest of the other background PDFs are shown in the Appendix B in Figure B.1 and B.2 for $\Upsilon(2S)$ and $\Upsilon(3S)$, respectively.

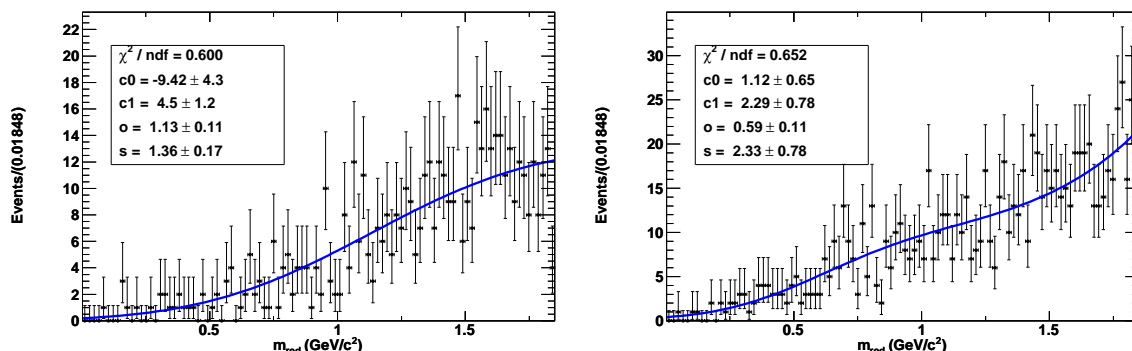


Figure 4.5: The background PDF in the threshold mass region in the m_{red} distribution. Left plot is for the $\Upsilon(2S)$ dataset and right plot is for the $\Upsilon(3S)$ dataset.

4.3.1 Fit Validation using a cocktail sample

The validation of the fit procedure is performed using a cocktail sample of the $\Upsilon(3S, 2S)$ low onpeak data-sample and 95% of $\Upsilon(3S, 2S)$ generic MC sample. The cocktail sample contains about 4522 events for $\Upsilon(3S)$ and about 12446 events for $\Upsilon(2S)$, as expected in the full data samples. Figure 4.6 shows the reduced mass distribution for $\Upsilon(3S, 2S)$ low onpeak and $\Upsilon(3S, 2S)$ generic samples after applying all the selection criteria. As seen in these figures the statistics is very limited in the low mass region in both the datasets. There are many regions in the m_{red} distribution where there are no events. The normal ML fit procedure gives large negative signal yield in a region of the m_{red} spectrum, where the statistics is limited. This problem can be avoided if we constraint the number of signal and background events to be greater or equal to zero. This constraint method works fine in the region of limited statistics and ignores the negative fluctuation in the datasets but introduces a bias, specially, where the statistics is little bit large, but not sufficient to use the normal fitting approach. To avoid these difficulties, we impose a lower cutoff to the signal yield to ensure that the total signal plus background PDF remains non-negative in the integration region [89].

We perform the scan for any possible peaks in the m_{red} distribution from $\Upsilon(3S, 2S)$ cocktail samples in the steps of half of m_{red} resolution, corresponding to 4585 points. The shape of the signal-PDF is fixed while the background-PDF shape, signal and background yields are allowed to float. The parameters of the signal PDF are interpolated between the known MC points. The representative plots of the 1d ML fit to the m_{red} distributions are shown in

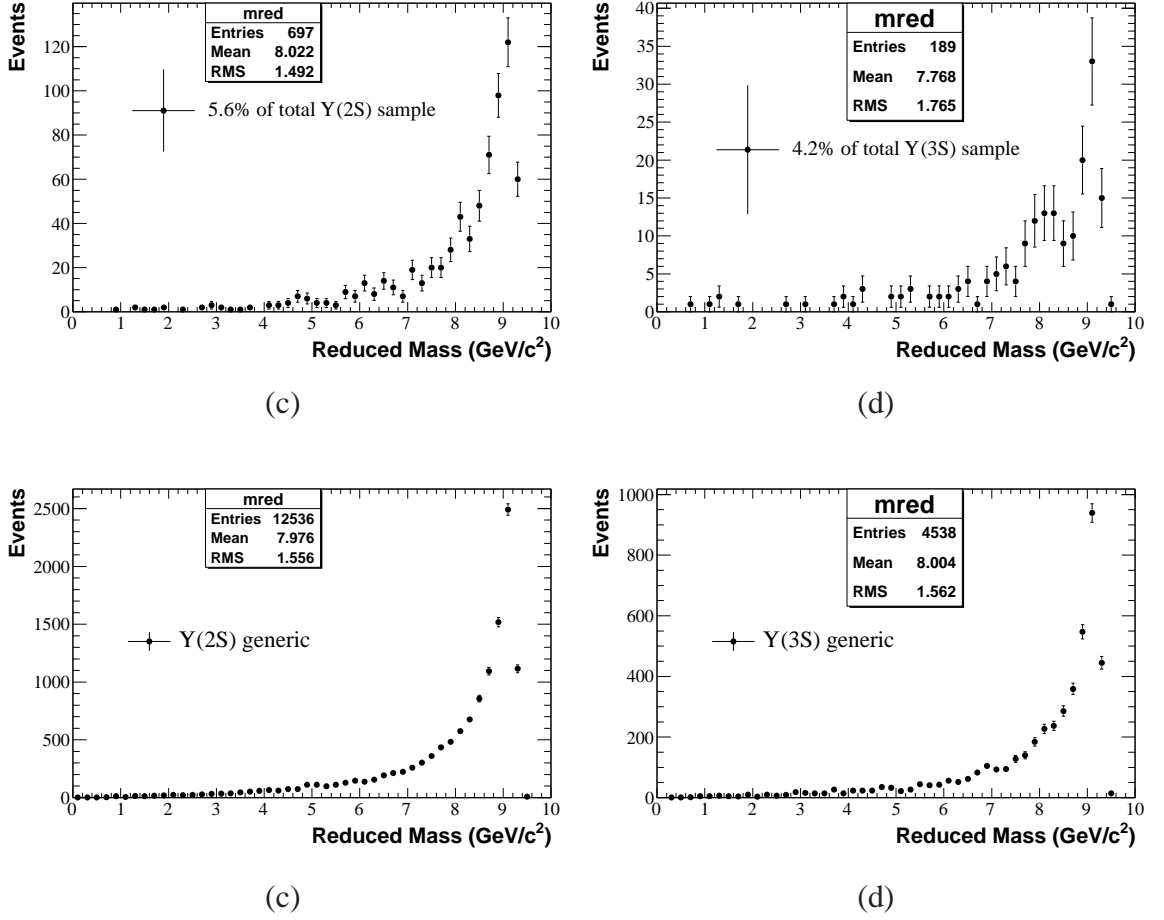


Figure 4.6: m_{red} distribution for $\Upsilon(3S, 2S)$ low onpeak and $\Upsilon(3S, 2S)$ generic samples. Left plots are for $\Upsilon(2S)$ right plots are for $\Upsilon(3S)$.

Figure 4.7 at some selected m_{A^0} points. The signal events (N_{sig}) as a function of m_{A^0} are shown in Figure 4.8. We also calculate a statistical significance (\mathcal{S}) which is defined as:

$$\mathcal{S} = \text{sign}(N_{sig}) \sqrt{-2 \ln(\mathcal{L}_0 / \mathcal{L}_{max})}, \quad (4.10)$$

where \mathcal{L}_{max} is the maximum likelihood value of a fit with a floating signal yield centered at m_{A^0} , and \mathcal{L}_0 is the likelihood value for the null hypothesis. Figure 4.9 shows the significance distributions for both the $\Upsilon(2S, 3S)$ cocktail datasets. The significance barely deviates more than 3σ for both the datasets. We also compute the combined significance of the $\Upsilon(2S, 3S)$ datasets, which is defined as:

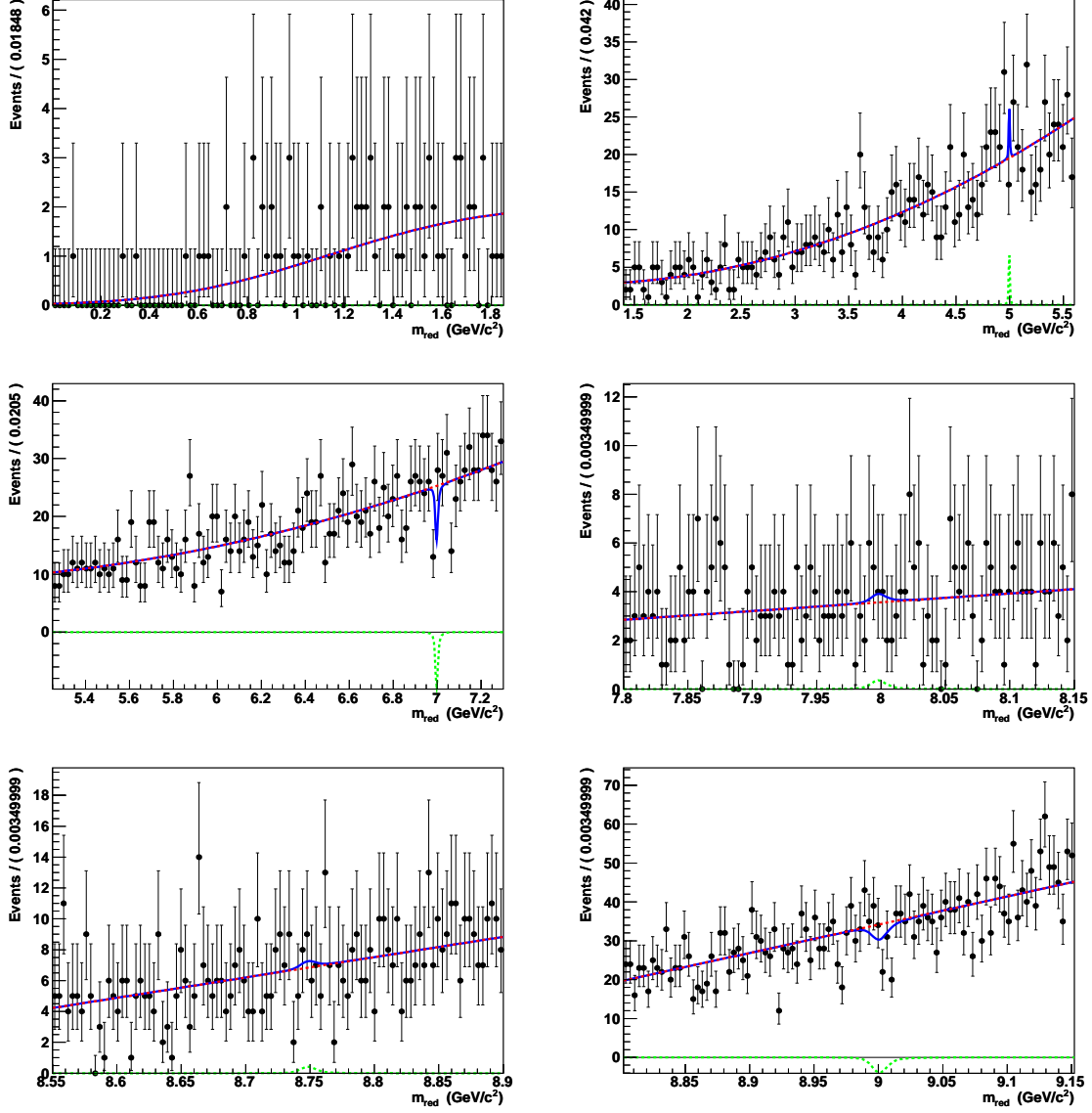


Figure 4.7: Projection plot onto m_{red} distributions at selected m_{A^0} points using the cocktail samples of $\Upsilon(2S, 3S)$. The total ML fit is shown in solid blue; the non-peaking background component is shown in dashed green; the signal component is shown in green dashed.

$$\mathcal{S} = \frac{w_{\Upsilon(2S)}\mathcal{S}_{\Upsilon(2S)} + w_{\Upsilon(3S)}\mathcal{S}_{\Upsilon(3S)}}{\sqrt{w_{\Upsilon(2S)}^2 + w_{\Upsilon(3S)}^2}}, \quad (4.11)$$

where $\mathcal{S}_{\Upsilon(2S,3S)}$ is the significance of the $\Upsilon(2S, 3S)$ data-sets, computed at each scanned m_{A^0} points and $w_{\Upsilon(2S,3S)}=1/\sigma_{N_{\text{sig}}}^2$ is the weight of the each data-sets.

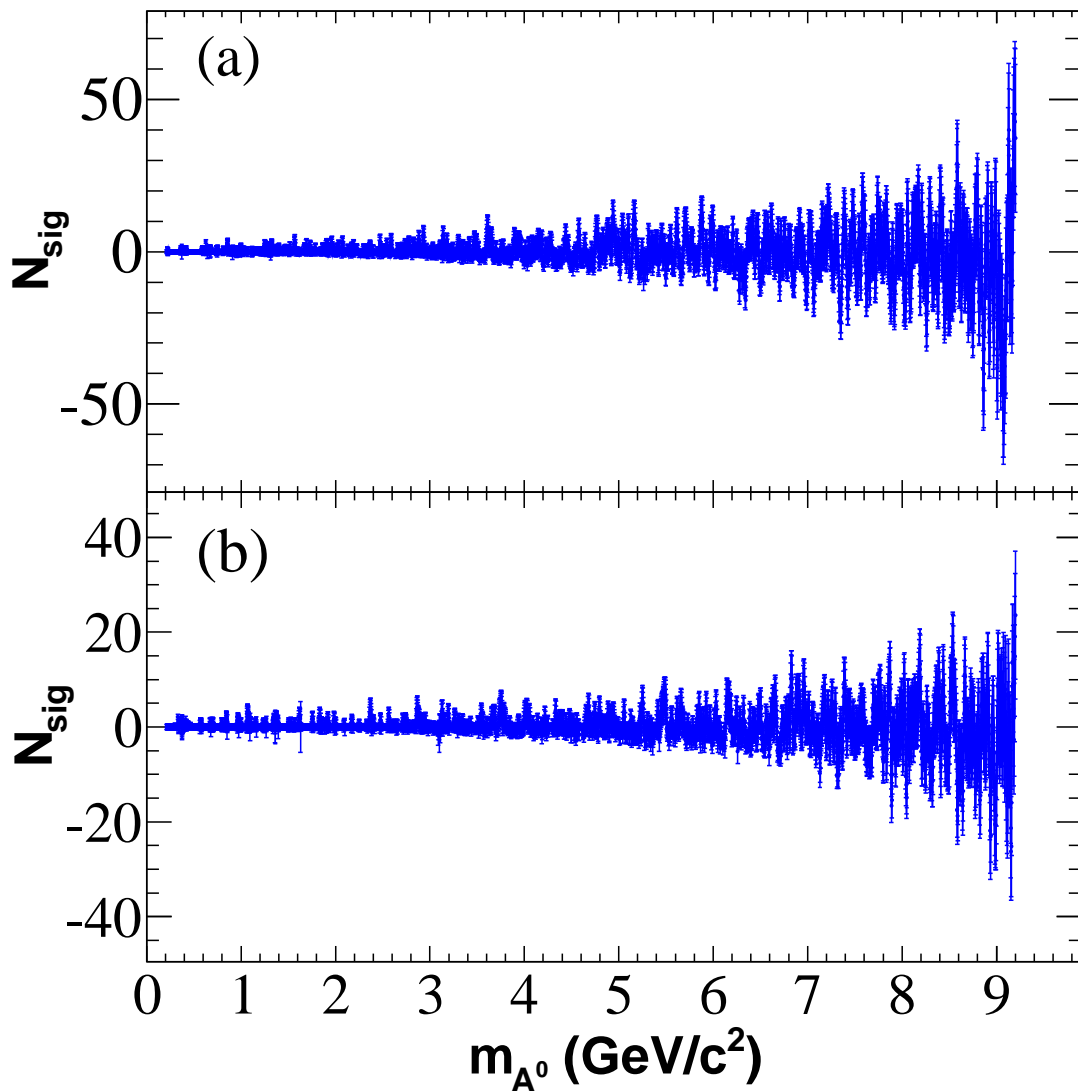


Figure 4.8: The number of signal events (N_{sig}) as a function of m_{A^0} for (a) the $\Upsilon(2S)$ dataset and (b) the $\Upsilon(3S)$ dataset. These plots are generated using the $\Upsilon(2S, 3S)$ cocktail samples.

4.3.2 Fit validation using Toy Monte-Carlo

We use a large number of toy Monte-Carlo experiments to validate the fit procedure further. We first fit the background PDF's to the $\Upsilon(3S, 2S)$ cocktail samples. Then, we generate the background events according to those PDFs, setting the background yields to the number expected in the Run7 $\Upsilon(3S, 2S)$ Onpeak datasets. The toy studies are done with different embedded signal events for each m_{A^0} points.

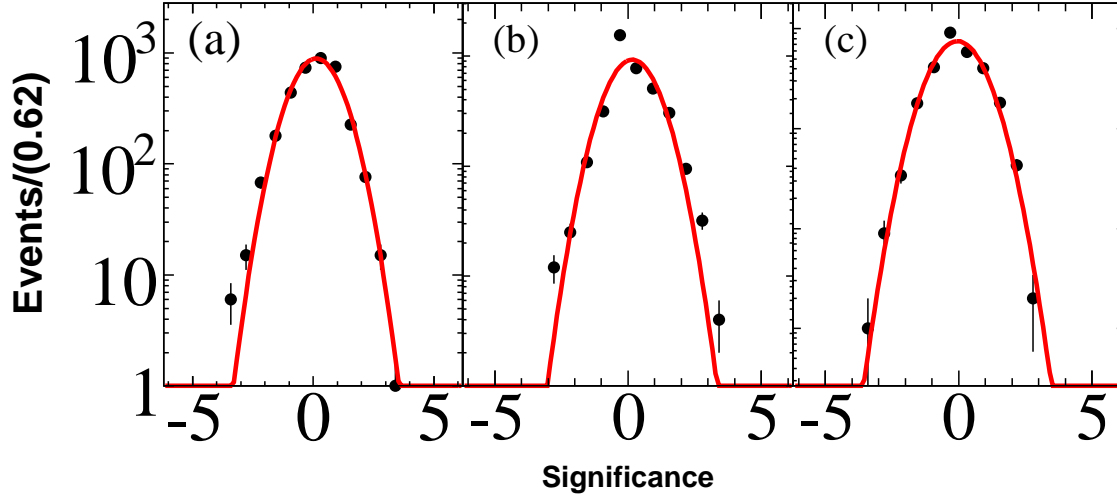


Figure 4.9: The signal significance (S) as a function of m_{A^0} for (a) the $\Upsilon(2S)$ dataset, (b) the $\Upsilon(3S)$ dataset and (c) the combined data of $\Upsilon(2S, 3S)$. These plots are generated using the $\Upsilon(2S, 3S)$ cocktail samples.

The average fit-residuals (the difference between the number of fitted and generated events) as a function of embedded signal events for each m_{A^0} are summarized in Appendix C in Figure C.1 – C.3 for $\Upsilon(2S)$ and in Figure C.4 – C.6 for $\Upsilon(3S)$. The fit-residual as a function of embedded signal event is fitted by a linear function. We accumulate the intercept value of the regression in a histogram for all the known mass points for both $\Upsilon(2S)$ and $\Upsilon(3S)$ [Figure 4.10]. Since we do not observe any significant bias in the fitting procedure, we assign the RMS value of the intercept of the regression as a systematic uncertainty. The RMS value of fit bias (ΔN_{sig}) is found to be 0.17 for $\Upsilon(3S)$ and 0.22 for $\Upsilon(2S)$, which will be considered as an additional source of systematic uncertainty for $\Upsilon(2S)$ and $\Upsilon(3S)$ datasets.

4.4 Unblinding the $\Upsilon(2S, 3S)$ datasets.

After finalizing all the selection criteria and the ML fitting procedure, including the validation of the analysis, we have unblinded the (116.8 ± 1.0) million $\Upsilon(3S)$ events (sum of the “High” and “Medium” samples) and (92.8 ± 0.8) million $\Upsilon(2S)$ events.

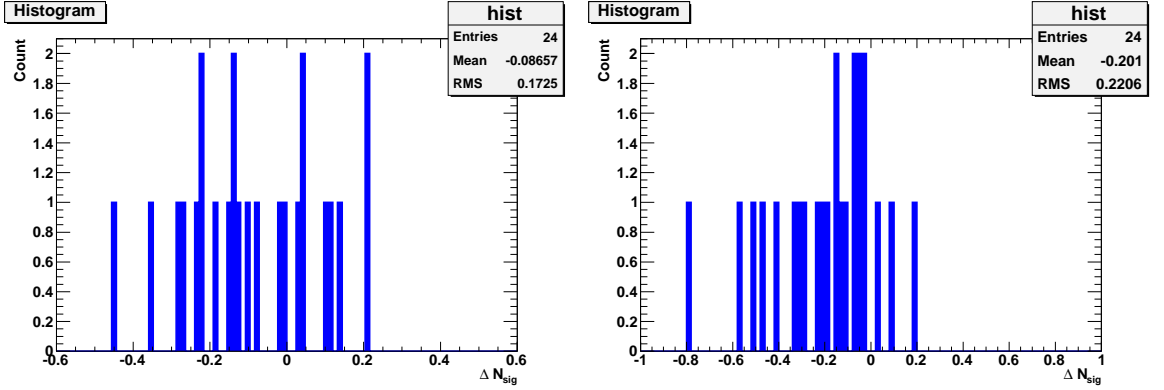


Figure 4.10: Histograms of the intercept of the regression for $\Upsilon(2S)$ (left) and $\Upsilon(3S)$ (right). The fit-residuals as a function of the embedded signal event are fitted by a linear function for each m_{A^0} points and intercept of the regression is accumulated in the Histogram.

A total of 11,136 $\Upsilon(2S)$ and 3,857 $\Upsilon(3S)$ candidates are selected by the selection criteria (mentioned in section 3.3.5) in the unblinded data samples of $\Upsilon(2S, 3S)$. Figure 4.11 and Figure 4.12 show the distributions of the m_{red} and m_{recoil} together with the remaining background MC samples of $\Upsilon(2S, 3S) \rightarrow \pi^+\pi^-\Upsilon(1S)$, $\Upsilon(1S) \rightarrow (\gamma)\mu^+\mu^-$ decays. The MCs are normalized to the data luminosity.

Two peaking components corresponding to ρ^0 and J/ψ mesons are observed in the $\Upsilon(3S)$ dataset. The ρ^0 -mesons are mainly produced in initial state radiation events, along with two or more pions, which disappears if we require both candidates to be identified as muons in the A^0 reconstruction (Figure 4.13) or apply a tighter (3σ) mass window cut on the m_{recoil} distribution. An enhancement of the ρ^0 background is observed outside the signal region of $[9.455, 9.48]$ GeV/c^2 in the m_{recoil} distribution of the $\Upsilon(3S)$ dataset (Figure 4.14). This data sample is used to model the ρ^0 background using a sum of a Gaussian and a constant linear function (Figure 4.14). The fixed PDF parameters of the Gaussian function are used to describe the ρ^0 peak in the final fit.

To understand the peaking component at the J/ψ mass position in the $\Upsilon(3S)$ dataset, we compute the mass of the system recoiling against the photon, which is defined as:

$$m_{\text{recoil}}^\gamma = s - 2 \cdot \sqrt{s} \cdot E_{CM}^\gamma, \quad (4.12)$$

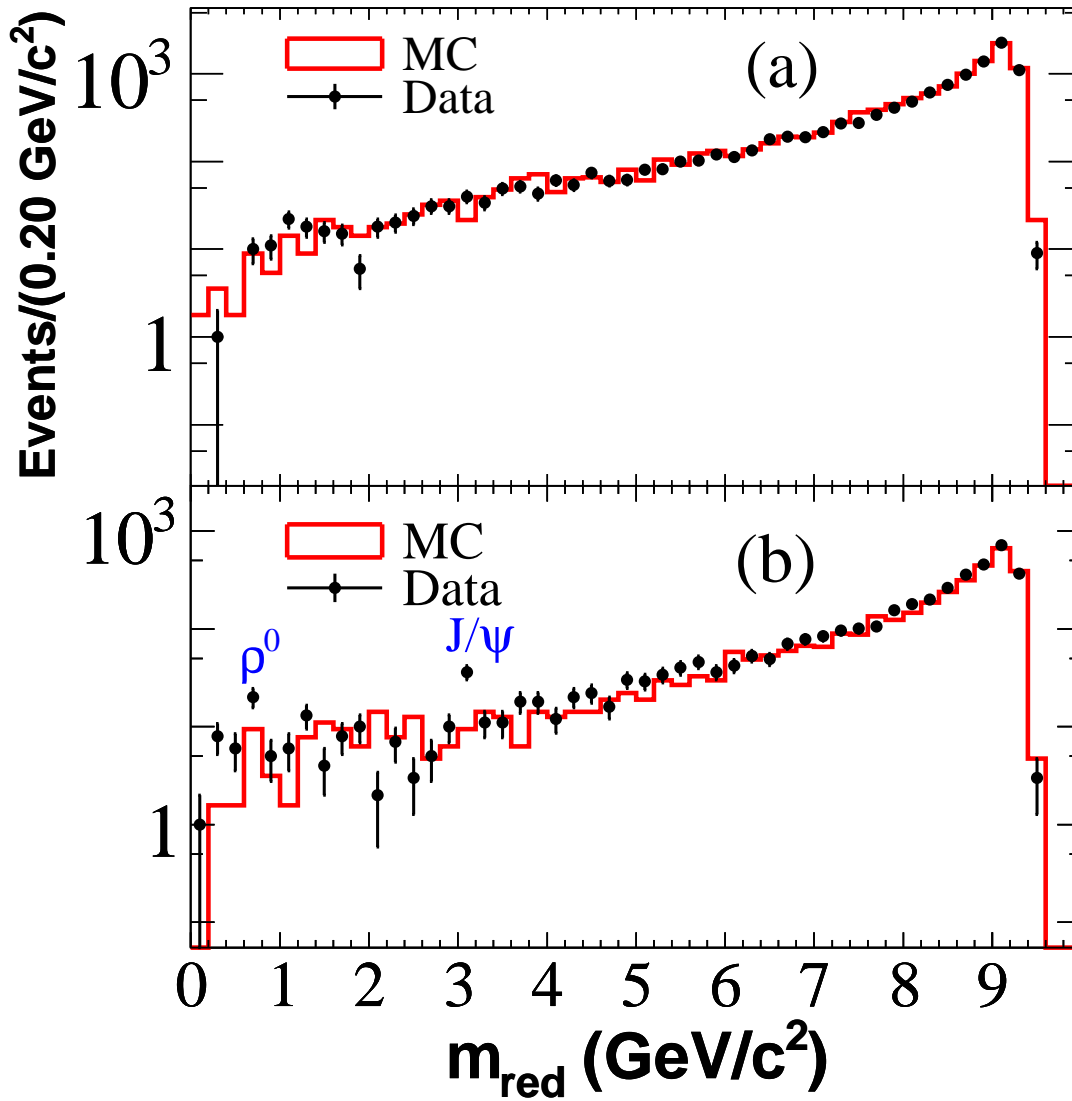


Figure 4.11: The distribution of m_{red} for (a) the $\Upsilon(2S)$ and (b) the $\Upsilon(3S)$ datasets, together with the remaining background Monte Carlo samples of $\Upsilon(2S,3S) \rightarrow \pi^+\pi^-\Upsilon(1S)$, $\Upsilon(1S) \rightarrow (\gamma)\mu^+\mu^-$ decays. The Monte Carlo are normalized to the data luminosity. Two peaking components corresponding to the ρ^0 and J/ψ mesons are observed in the $\Upsilon(3S)$ dataset.

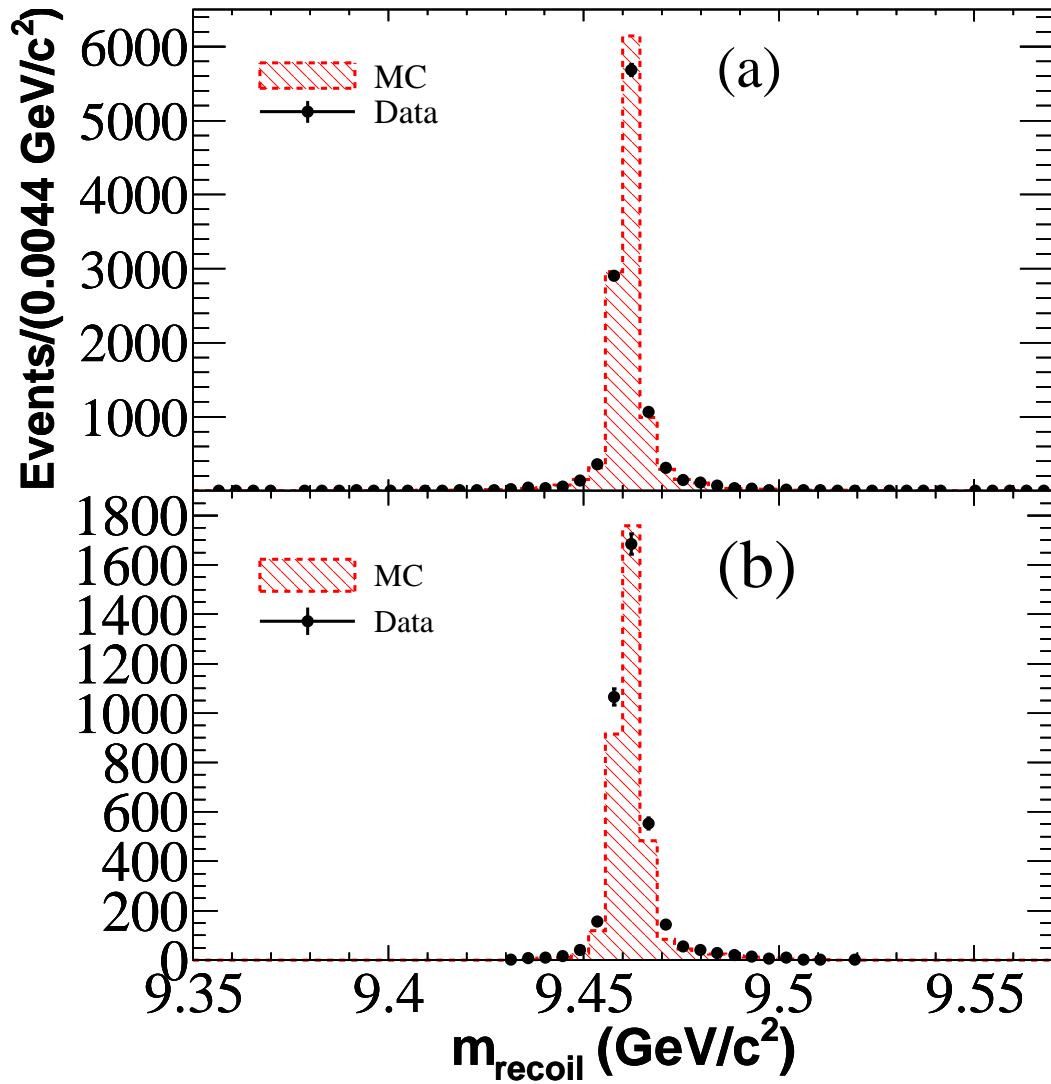


Figure 4.12: The distribution of m_{recoil} for (a) the $\Upsilon(2S)$ and (b) the $\Upsilon(3S)$ datasets, together with the remaining background Monte Carlo samples of $\Upsilon(2S, 3S) \rightarrow \pi^+\pi^-\Upsilon(1S)$, $\Upsilon(1S) \rightarrow (\gamma)\mu^+\mu^-$ decays. The Monte Carlo are normalized to the data luminosity.

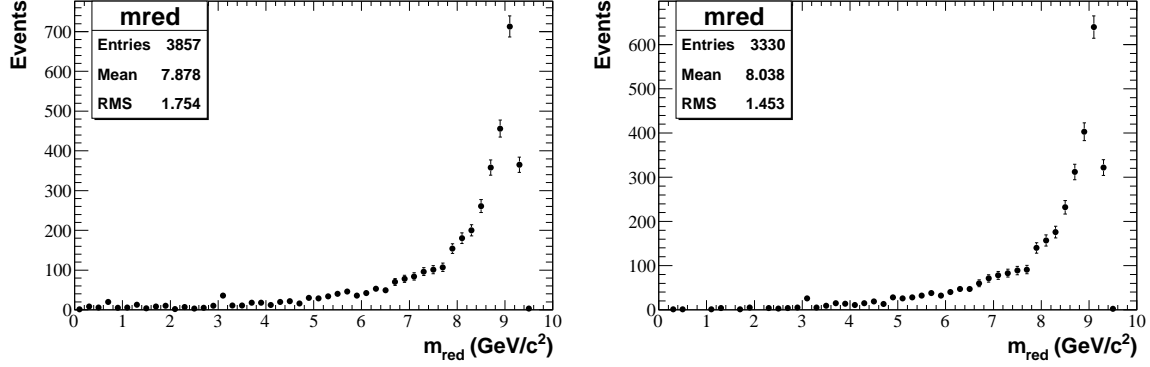


Figure 4.13: The m_{red} distribution for the unblinded $\Upsilon(3S)$ Onpeak data-set after applying all the selection criteria including the OR muon PID cut (left) and AND muon PID cut (right). The first peak disappears after applying the AND muon PID cut but the second peak does not.

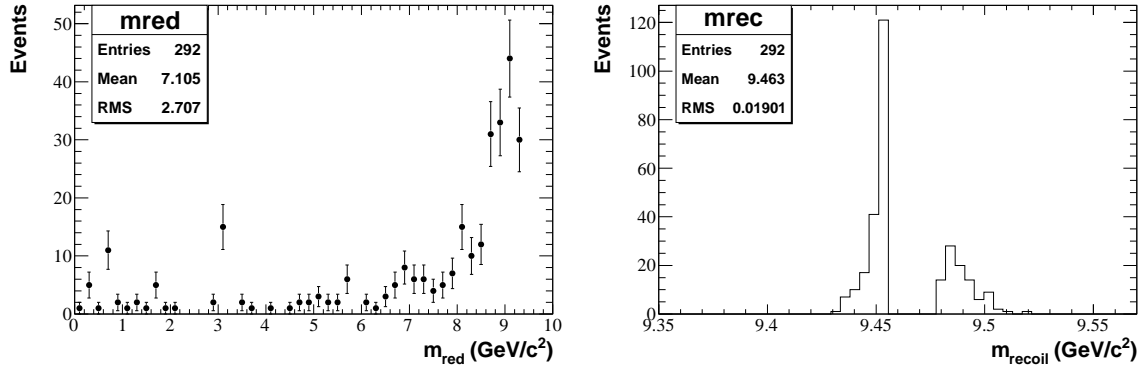


Figure 4.14: The m_{red} distribution (left) and m_{recoil} distribution (right) for the sideband $\Upsilon(3S)$ data of the m_{recoil} . We will use sideband of the m_{recoil} distribution in the $\Upsilon(3S)$ Onpeak dataset to model the ρ^0 background.

where \sqrt{s} is the CM energy of the e^+e^- system and E_{CM}^γ is the CM energy of the photon. The m_{recoil}^γ should peak at the mass position of the X resonance, in an ISR decay like $e^+e^- \rightarrow \gamma_{\text{ISR}}X$. Figure 4.16 shows the m_{recoil}^γ distribution in both m_{red} region of $[3.0 - 3.2]$ GeV/c^2 as well as the outside of this region using $\Upsilon(3S)$ onpeak dataset. It is clear that the m_{recoil}^γ distribution peaks at $\psi(2S)$ mass position for the m_{red} region of $3.0 - 3.2$ GeV/c^2 . We have also processed a sample of $e^+e^- \rightarrow \gamma_{\text{ISR}}\psi(nS)$ with generic decays of $\psi(nS)$. The m_{red} distribution of $\psi(nS)$ generic sample at J/ψ mass position is shown in Figure 4.17 (left). The m_{recoil}^γ distribution of $\psi(nS)$ generic sample is also shown in Figure 4.17 (right), which peaks at $\psi(2S)$ mass position. Using MC-Truth information of the survived MC events for $\psi(nS)$ generic decays, it is observed that about 95% of the events decay via $\psi(2S) \rightarrow \pi^+\pi^- J/\psi$ and about 99% of the

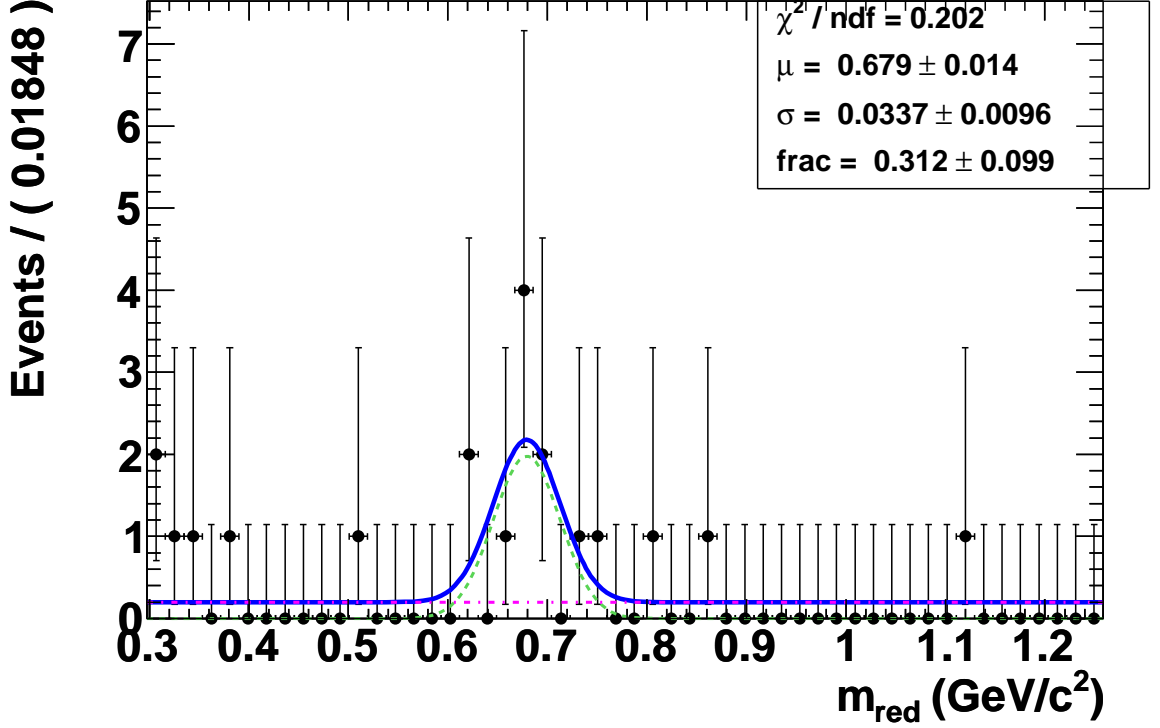


Figure 4.15: The peaking background PDF at the ρ^0 mass position. We use the sideband of the m_{recoil} distribution in the $\Upsilon(3S)$ onpeak dataset to model this background.

J/ψ events decay via $\mu^+\mu^-$ channel (Figure 4.18). We model the peaking component of the J/ψ background by a CB function using data sample of this $\psi(nS)$ generic decays (Figure 4.19). A high statistics data and MC samples of $e^+e^- \rightarrow \gamma_{ISR}\psi(2S)$, $\psi(2S) \rightarrow \pi^+\pi^-J/\psi$, $J/\psi \rightarrow \mu^+\mu^-$ have also been used to check the resolution of m_{recoil} distribution at the J/ψ mass peak position. We find a resolution of $(2.014 \pm 0.309) \times 10^{-3} \text{ GeV}/c^2$ in the data, compatible with the predictions of the MC of $(2.007 \pm 0.011) \times 10^{-3} \text{ GeV}/c^2$, which is obtained by applying the mass constraints on the $\psi(2S)$ to improve the resolution of the J/ψ . A similar exercise without the mass constraint results in an agreement between data and Monte Carlo as well. However, the resolution of these J/ψ event is not representative of that of the signal, because the kinematic is different.

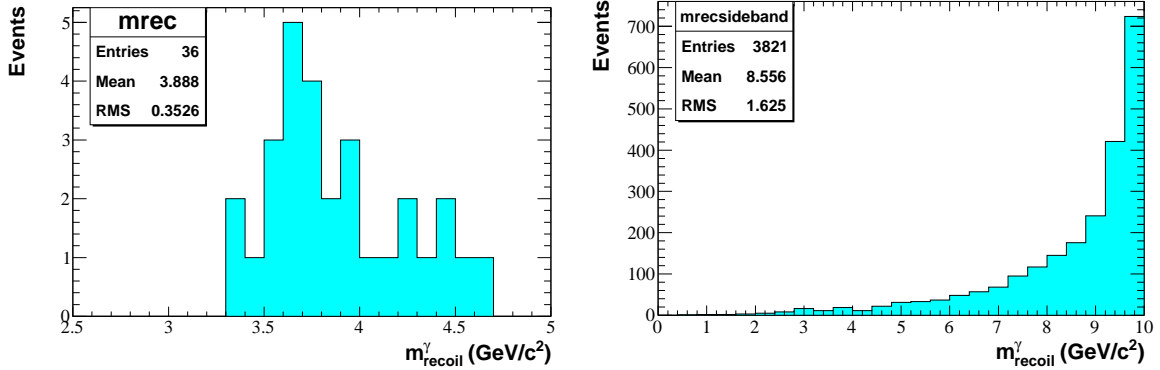


Figure 4.16: The m_{recoil}^γ distribution for the unblinded $\Upsilon(3S)$ Onpeak data-set after applying all the selection criteria. The left plot shows the m_{recoil}^γ distribution in the region of $m_{\text{red}} = [3.0, 3.2]$ GeV/c^2 and the right plot shows the same distribution in the region of $(m_{\text{red}} < 3.0) \parallel (m_{\text{red}} > 3.2)$ GeV/c^2 .

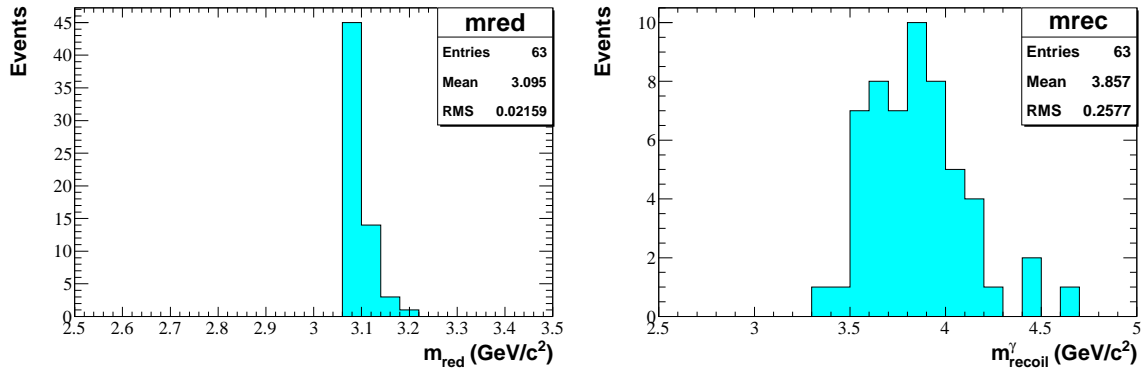


Figure 4.17: The distribution of m_{red} (left) and m_{recoil}^γ (right) in the $\psi(nS)$ generic decays sample after applying all the selection cuts.

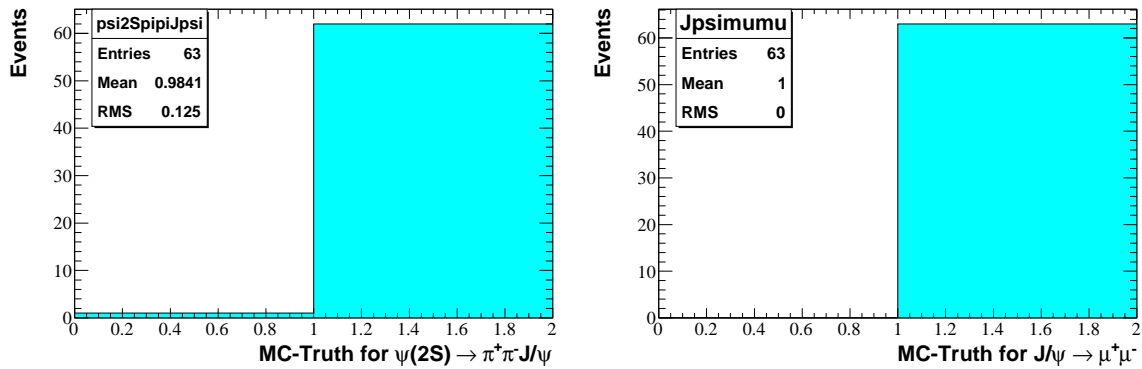


Figure 4.18: MC-Truth Boolean distribution for $\psi(2S) \rightarrow \pi^+\pi^-J/\psi$ decays (left) and for $J/\psi \rightarrow \mu^+\mu^-$ decays (right) in the $\psi(nS)$ generic decays sample.

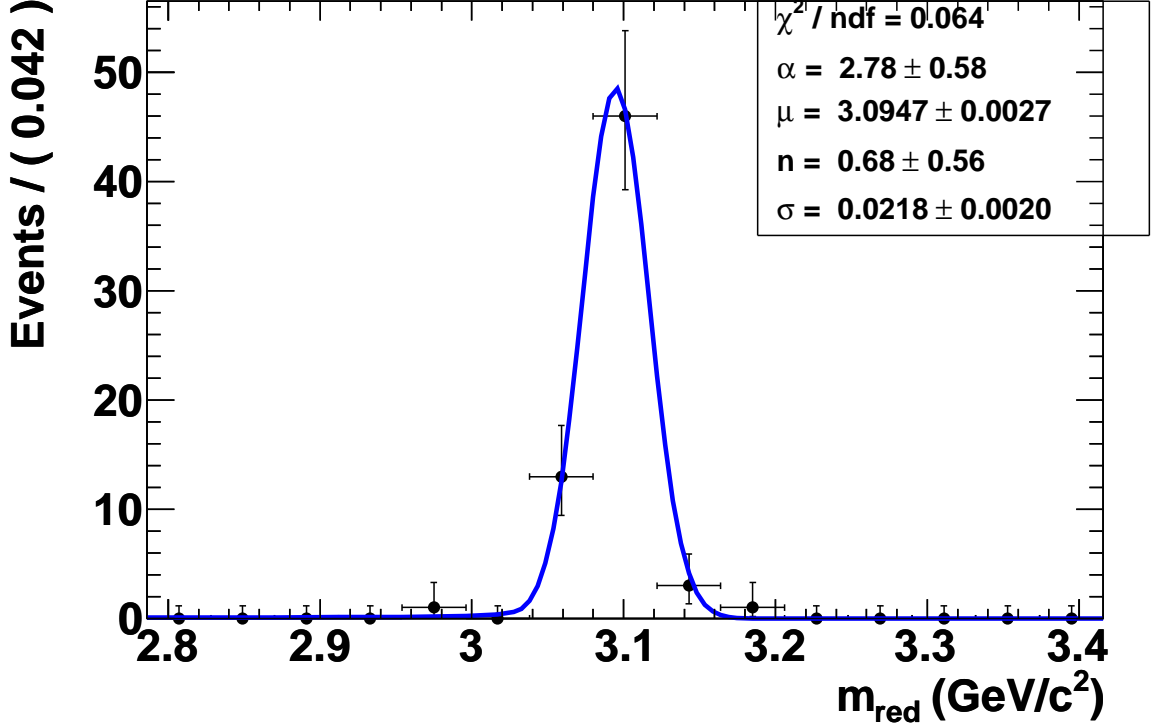


Figure 4.19: The background PDF for m_{red} at J/ψ mass position. We use a sample of $\psi(nS)$ generic decays to model this background.

4.5 Signal yield extraction using the 1d ML fit

We perform the likelihood scan for any possible peaks in the m_{red} distribution using the unblinded $\Upsilon(3S, 2S)$ onpeak data-set in the steps of half of m_{red} resolution, corresponding to 4585 m_{A^0} points. The J/ψ mass region in the $\Upsilon(3S)$ dataset, defined as $3.045 \leq m_{\text{red}} \leq 3.162$ GeV/c^2 , is excluded from the search due to large background from $J/\psi \rightarrow \mu^+\mu^-$ decays. The projection plots for selected mass points are shown in Figure 4.20 and 4.21. Figure 4.22 shows the number of signal events as well as signal significance for the $\Upsilon(3S, 2S) \rightarrow \pi^+\pi^-\Upsilon(1S)$; $\Upsilon(1S) \rightarrow \gamma A^0$; $A^0 \rightarrow \mu^+\mu^-$ decay as a function of m_{A^0} . Figure 4.23 shows the distribution of signal significance (S), where S is excluded in the range of $-0.04 < S < 0$. The significance is expected to follow a normal distribution with $\mu = 0$ and $\sigma = 1$ for the pure background hypothesis. The largest values of significance are found to be 3.62 (2.96) in the $\Upsilon(2S)$ ($\Upsilon(3S)$) dataset, and 3.24 for the combined $\Upsilon(2S, 3S)$ dataset.

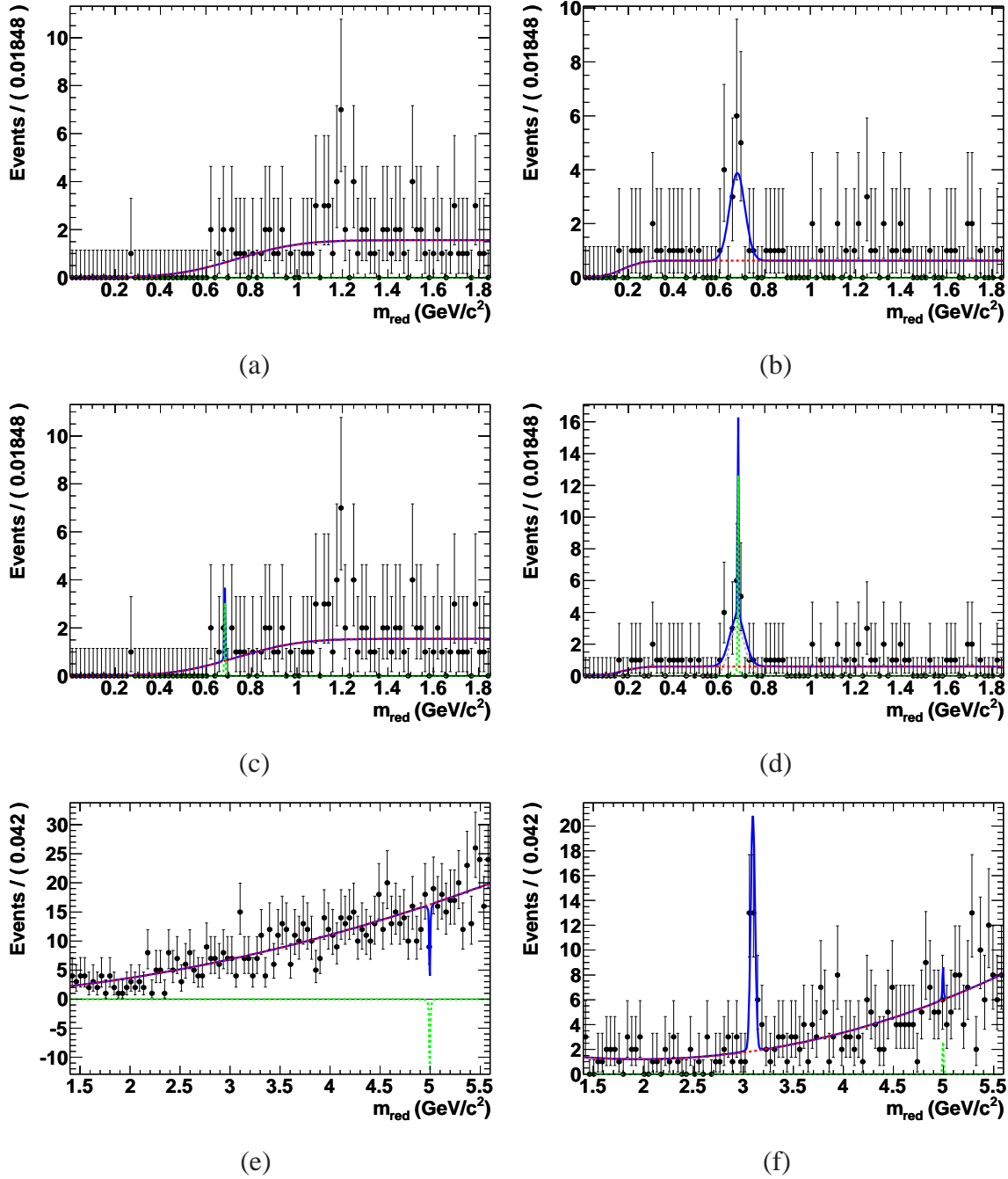


Figure 4.20: Result of the likelihood fit to the unblinded $\Upsilon(2S, 3S)$ onpeak data samples. Projection plot onto reduce mass distribution for the m_{A^0} of (a, b) $m_{A^0} = 0.212 \text{ GeV}/c^2$, (c, d) $m_{A^0} = 0.715 \text{ GeV}/c^2$ and (e, f) $m_{A^0} = 5.0 \text{ GeV}/c^2$. Left plots are for $\Upsilon(2S)$ and right plots are for $\Upsilon(3S)$ data sample. The total ML fit is shown in solid blue; the non-peaking background component is shown in dashed magenta; the signal component is shown in green dashed. The peaking components of ρ^0 and J/ψ resonances are modelled by a Gaussian and a CB function, respectively in the $\Upsilon(3S)$ data-set.

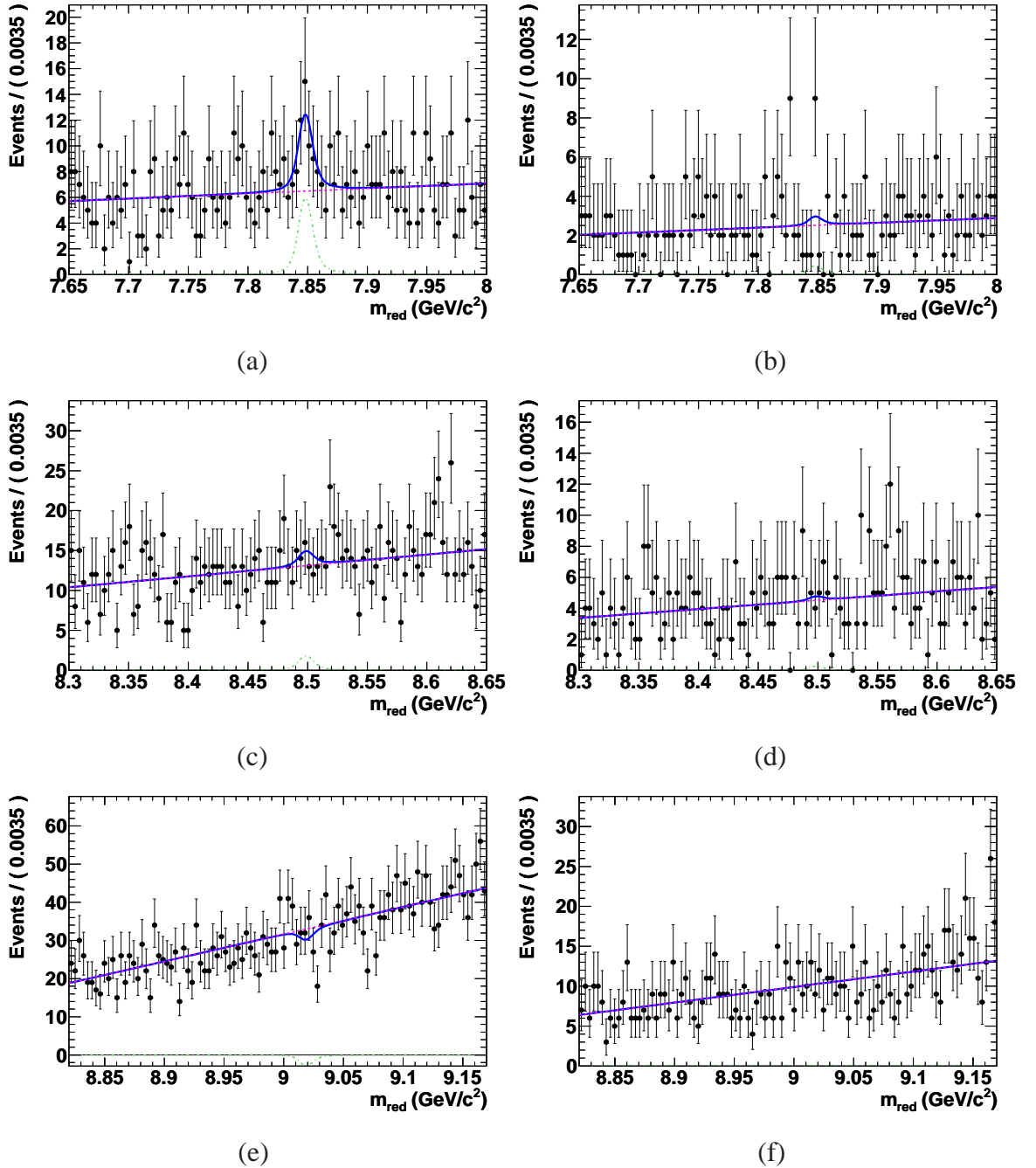


Figure 4.21: Result of the likelihood fit to the unblinded $\Upsilon(3S, 2S)$ onpeaks data samples. Projection plot onto reduce mass distribution for the m_{A^0} of (a, b) $m_{A^0} = 7.85 \text{ GeV}/c^2$ (c, d) $m_{A^0} = 8.5 \text{ GeV}/c^2$ and (e, f) $m_{A^0} = 9.02 \text{ GeV}/c^2$. Left plots are for $\Upsilon(2S)$ and right plots are for $\Upsilon(3S)$ data sample. The total ML fit is shown in solid blue; the non-peaking background component is shown in dashed magenta; the signal component is shown in green dashed.

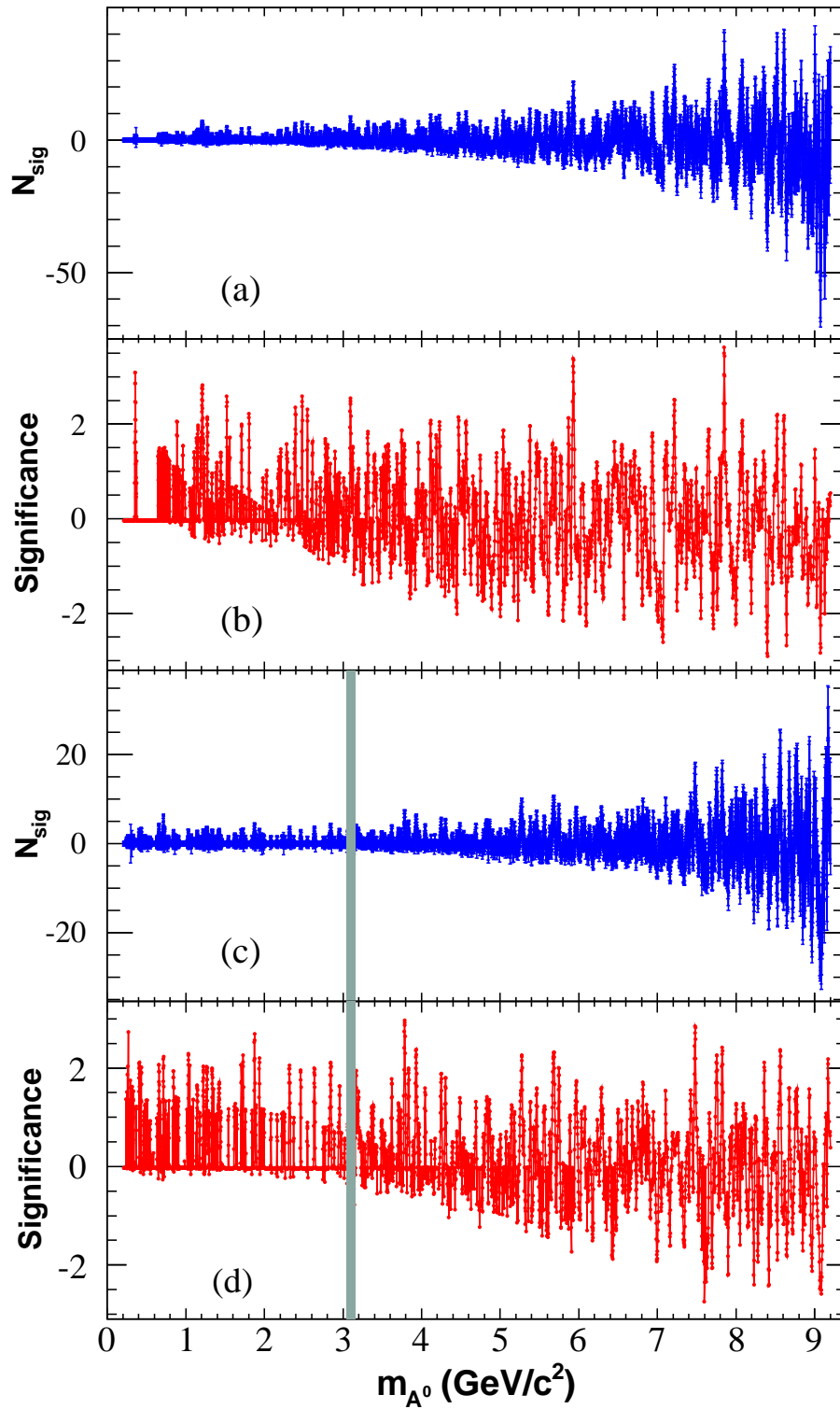


Figure 4.22: The number of signal events and signal significance as a function of m_{A^0} for (a,b) $\Upsilon(2S)$ and (c,d) for $\Upsilon(3S)$. The shaded area shows the region of the J/ψ resonance, excluded from the search in the $\Upsilon(3S)$ dataset.

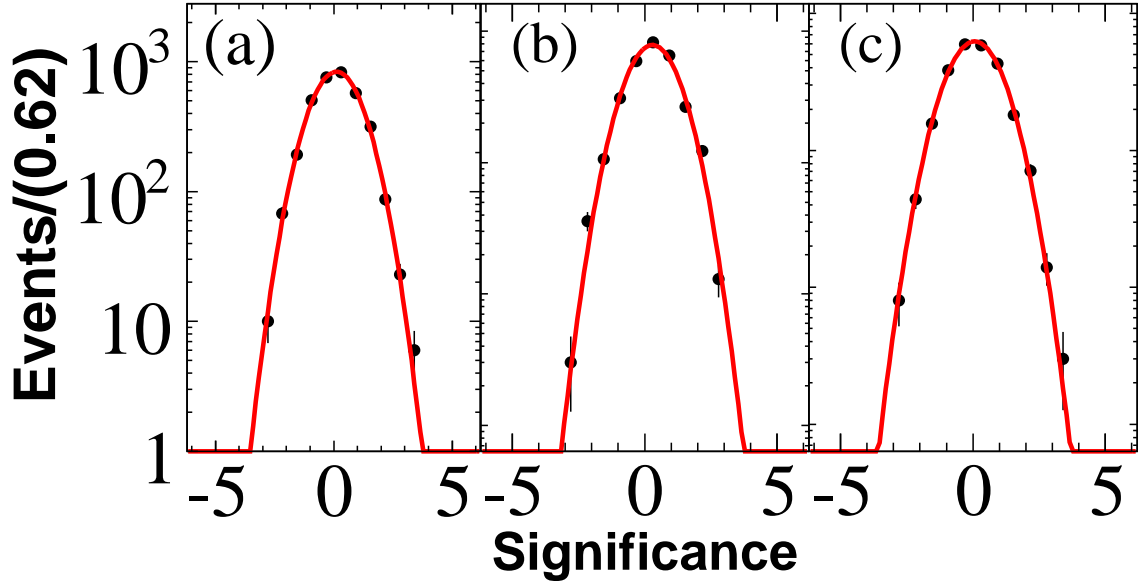


Figure 4.23: Histogram of the signal significance \mathcal{S} with statistical error for (a) $\Upsilon(2S)$ fit, (b) $\Upsilon(3S)$ fit and (c) the combined data of $\Upsilon(2S)$ and $\Upsilon(3S)$. The \mathcal{S} is excluded in the range of $-0.04 < \mathcal{S} < 0$. The overlaid curve shows the standard normal distribution expected in the absence of signal.

4.6 Trial factor study: true significance observation

When we search for a narrow resonance for A^0 at unknown mass points over a broad range of background, special care must be exercised in evaluating the true significance of observing a local excess of events. The log-likelihood ratio method is used to compute the significance of any positive signal observation. Since we need to scan the m_{red} distribution of the $\Upsilon(2S, 3S)$ onpeaks datasets at 4585 m_{A^0} points, we should expect at least a few statistical fluctuations at the level of $\mathcal{S} \approx 3$, even for the null hypothesis. Hence, we need to determine the probability for the background fluctuation to a particular value of \mathcal{S} anywhere in a given m_{A^0} range.

We generate toy Monte-Carlo data according to the PDFs using the background only hypothesis. We then scan the toy data in the same way as was done for the $\Upsilon(2S, 3S)$ onpeak data-sets for all 4585 m_{A^0} points, and pick up one of the maximum value of significance S_{max} from these 4585 m_{A^0} points. We repeat this process about 5000 times and accumulate the S_{max} value each time in a histogram. We also compute the $S_{\text{max}}^{\text{comb}} = (S_{\text{max}}^{\Upsilon(2S)} + S_{\text{max}}^{\Upsilon(3S)})/2$ for the combined $\Upsilon(2S, 3S)$ datasets.

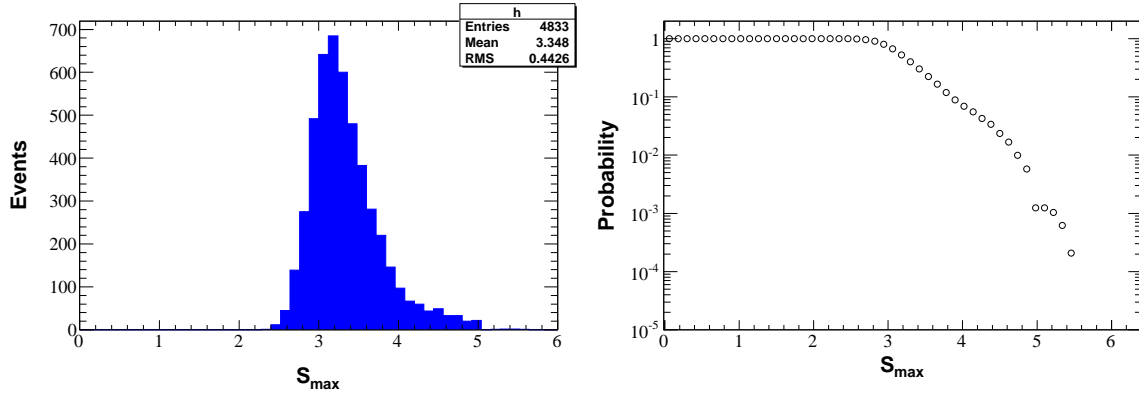


Figure 4.24: Histogram of the S_{max} (left) and its cumulative distribution (right) for the $\Upsilon(2S)$ dataset.

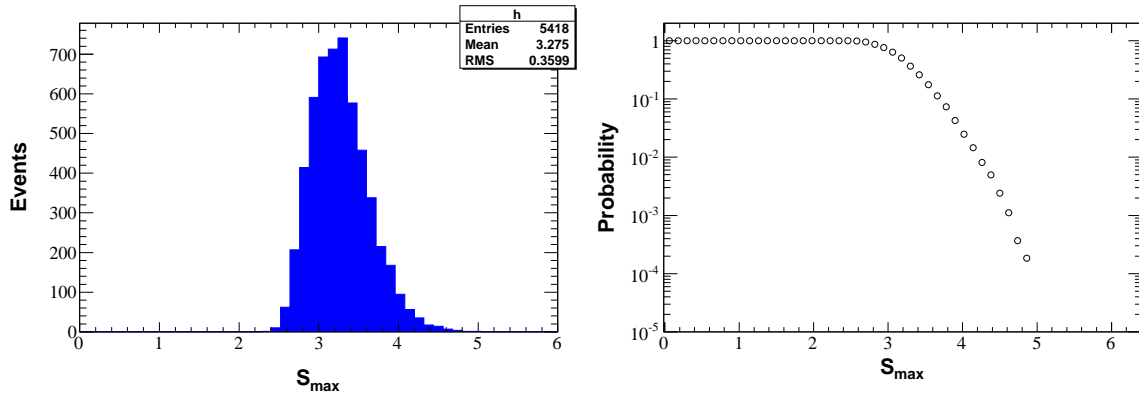


Figure 4.25: Histogram of the S_{max} (left) and its cumulative distribution (right) for the $\Upsilon(3S)$ dataset.

We compute the inverse cumulative distribution (also called p-value) of S_{max} while integrating the PDF of S_{max} from S_{max} to ∞ . The histograms of S_{max} and its inverse cumulative distribution for $\Upsilon(2S)$, $\Upsilon(3S)$ and combined $\Upsilon(2S, 3S)$ datasets are shown in Figure 4.24, 4.25 and 4.26, respectively. The p-value is the probability of a test statistics that describes the chance that a pure background would fluctuate to a signal peak with the significance S_{max} . If the null hypothesis is correct, the p-value is uniformly distributed between zero and one. To express a given value of probability in terms of standard deviations (σ), a convention is adopted for one sided Gaussian value of $p = 1.35 \times 10^{-3}$ for 3σ and $p = 2.865 \times 10^{-7}$ for 5σ .

We estimate the probability to observe a fluctuation of $S_{max} \geq 3.62$ (18.1%) in the $\Upsilon(2S)$ ($\Upsilon(3S)$) data-set to be 18.1% (66.2%), and $S_{max} \geq 3.24$ in the combined $\Upsilon(2S, 3S)$ data-set to be 46.5% based upon this trial factor study. Hence we interpret the observed local excess of events at several mass points in both the datasets as a mere background fluctuations.

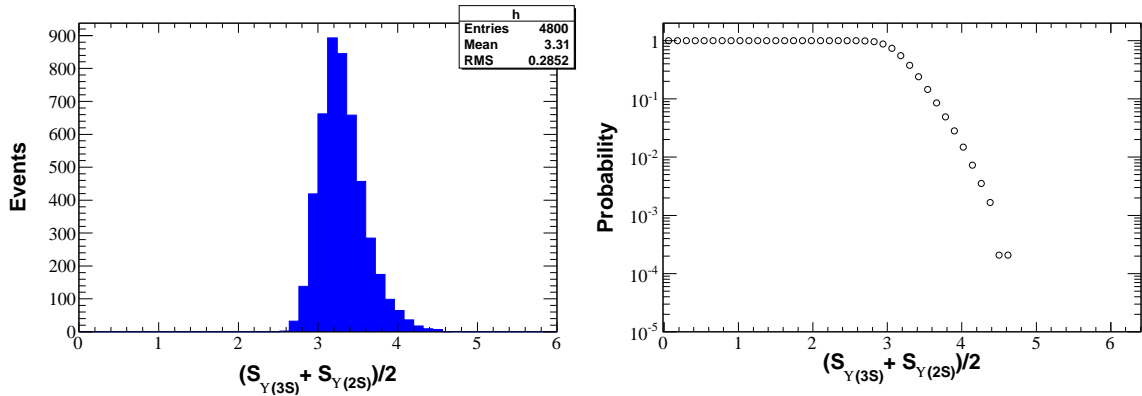


Figure 4.26: Histogram of the S_{max} (left) and its cumulative distribution (right) for the combined dataset of $\Upsilon(2S, 3S)$.

4.7 Chapter Summary

In this chapter we have described the ML fit procedures used to extract the signal yield from the data. We have developed the signal and background PDFs using the signal MC samples generated at 26 m_{A^0} points and the combined background MC. The fit validations are performed using a cocktail sample as well as a large number of Toy MC experiments with different embedded signal events at selected m_{A^0} points. The signal yields are extracted using the unblinded data of $\Upsilon(2S, 3S)$. A trial factor study is also performed, which shows that there is no evidence for the di-muon decay of the A^0 in the radiative decays of the $\Upsilon(1S)$ in the $\Upsilon(2S, 3S)$ data samples. The next chapter will describe the possible sources of systematic uncertainties for this analysis.

Chapter 5

Systematic Uncertainties

This chapter describes the sources of the systematic uncertainty which we consider in this analysis. Two kinds of systematic uncertainties are identified, which are additive and multiplicative systematics. The additive systematics reduce the significance of any observed peak and does not scale with the number of reconstructed events. It arises from the uncertainty on the PDF parameters and the fit bias. The multiplicative systematics do not change the significance of any observed peak and scales with the number of reconstructed events. The primary contributions to the multiplicative systematic uncertainties come from the RF classifier selection, muon-ID, photon-selection, tracking and $\Upsilon(2S, 3S)$ kinematic fit χ^2 .

5.1 PDF systematics

The dominant contribution to the additive systematic uncertainty comes from the uncertainties in the extracted signal yield (N_{sig}), which are primarily due to uncertainties in the PDF shapes. We evaluate the PDF systematic uncertainties after unblinding the Run7 $\Upsilon(3S, 2S)$ onpeak datasets by varying each parameter by its statistical error and observing the change in the fitted signal yield $\delta = \Delta N_{sig}$. The total systematic uncertainty in the signal yield is given by $\delta_{Tot} = \sqrt{\vec{\delta}^T C \vec{\delta}}$, where $\vec{\delta} = \langle \delta_1 \dots \delta_N \rangle$ and C is the parameter correlation matrix, giving a systematic uncertainty in the signal yield. The δ_{Tot} value is found to be very small for most of the m_{A^0} points and it varies from (0.00 – 0.62) events for the $\Upsilon(2S)$ dataset and (0.04 – 0.58) events for $\Upsilon(3S)$ dataset.

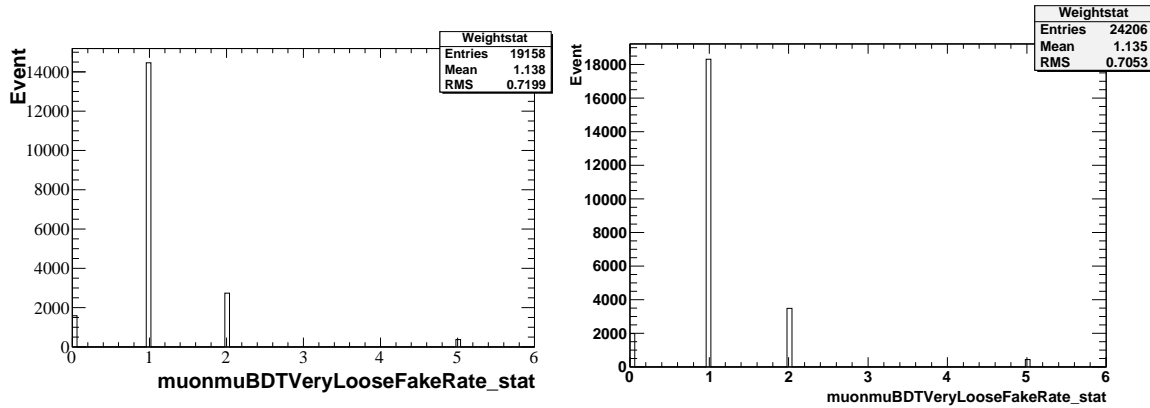


Figure 5.1: PID-weight status distribution for OR-muon PID selection cut for $\Upsilon(2S)$ (left) $\Upsilon(3S)$ (right).

5.2 Fit Bias

We perform a study of fit bias on the signal yield with a large number of Toy MC experiments as mentioned in section 4.3.2. The biases are consistent with zero and their average uncertainty is taken as a systematic uncertainty.

5.3 Systematic uncertainty for Particle ID

The systematic uncertainty for muon PID selection cuts is evaluated by a standard PID weighting recipe developed by the PID group in *BABAR* experiment [90]. This recipe creates a map that assigns a weight of each selected track, where weight is the ratio of efficiency in data and MC. Weight comes from the PID tables, which include the central value of weight, statistical uncertainty of weight, and status of given charged tracks. We first apply all the optimal selection cuts (excluding muon ID cut) to the signal MC sample in the mass range of 0.212 – 9.46 GeV/ c^2 . Then we check the status of charged tracks after applying the OR-muon PID (BDT-Muon1IDFakeRate || BDTMuon2IDFakeRate) selection cut. Figure 5.1 shows PID-weight status distribution for muon PID selection cut for both $\Upsilon(3S)$ and $\Upsilon(2S)$.

We construct a new table by generating 10000 Gaussian random number with mean = PID_weight, and sigma = PID_weighterr for the PIDWeight_status = 1, 2 and 3, where 1 means the PID efficiency of data and MC are well measured, 2 means the PID efficiency in

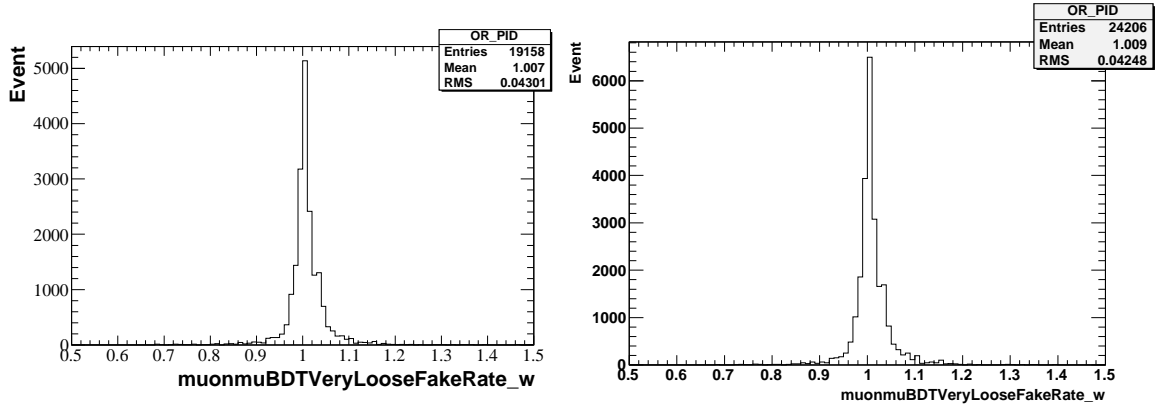


Figure 5.2: New generated PID_weight distribution (mean value of each Gaussian random number) for OR-muon PID selection cut for $\Upsilon(2S)$ (left) and $\Upsilon(3S)$ (right).

data and MC are poorly measured due to limited statistics, but still ok, and 3 means the PID efficiency in MC is zero, but upper limit in the weight is compatible with that of a lower momentum bin, from which the PID_weight and PID_weighterr have been taken. For other PIDWeight_status, we have used Gaussian mean = 1 and sigma = 0. Figure 5.2 shows the new generated PID_weight distribution (mean value of each Gaussian random number) for OR-muon PID selection cut for both $\Upsilon(3S)$ and $\Upsilon(2S)$. The systematic uncertainty on the muon PID efficiency is taken as the RMS value of the Gaussian. We find that the systematic uncertainty on the muon PID is 4.30% (4.25%) for $\Upsilon(2S)$ ($\Upsilon(3S)$).

5.4 Systematic uncertainty for the charged tracks

The systematic uncertainties for the four charged tracks are taken from [91], which results in a systematic uncertainty of 1.74% for the two highly energetic muon tracks. The systematic uncertainty for the pions with $p_T < 180$ MeV/c is taken from the soft-pion study and for tracks with $p_T > 180$ GeV/c, the systematic uncertainty is taken from the Tau31 study as discussed in [91]. Figure 5.3 shows that around (20.6%) (4%) signal MC events lie in the range of $p_T < 180$ MeV/c (for both pions) and 96% (79.4%) signal MC events lie in the range of $p_T > 180$ MeV/c for $\Upsilon(2S)$ ($\Upsilon(3S)$) dataset. So we evaluate the uncertainty due to the reconstruction of both pions to be 1.99% (1.76%) and the total systematic uncertainty for the four tracks to be 3.73% (3.5%) for $\Upsilon(2S)$ ($\Upsilon(3S)$).

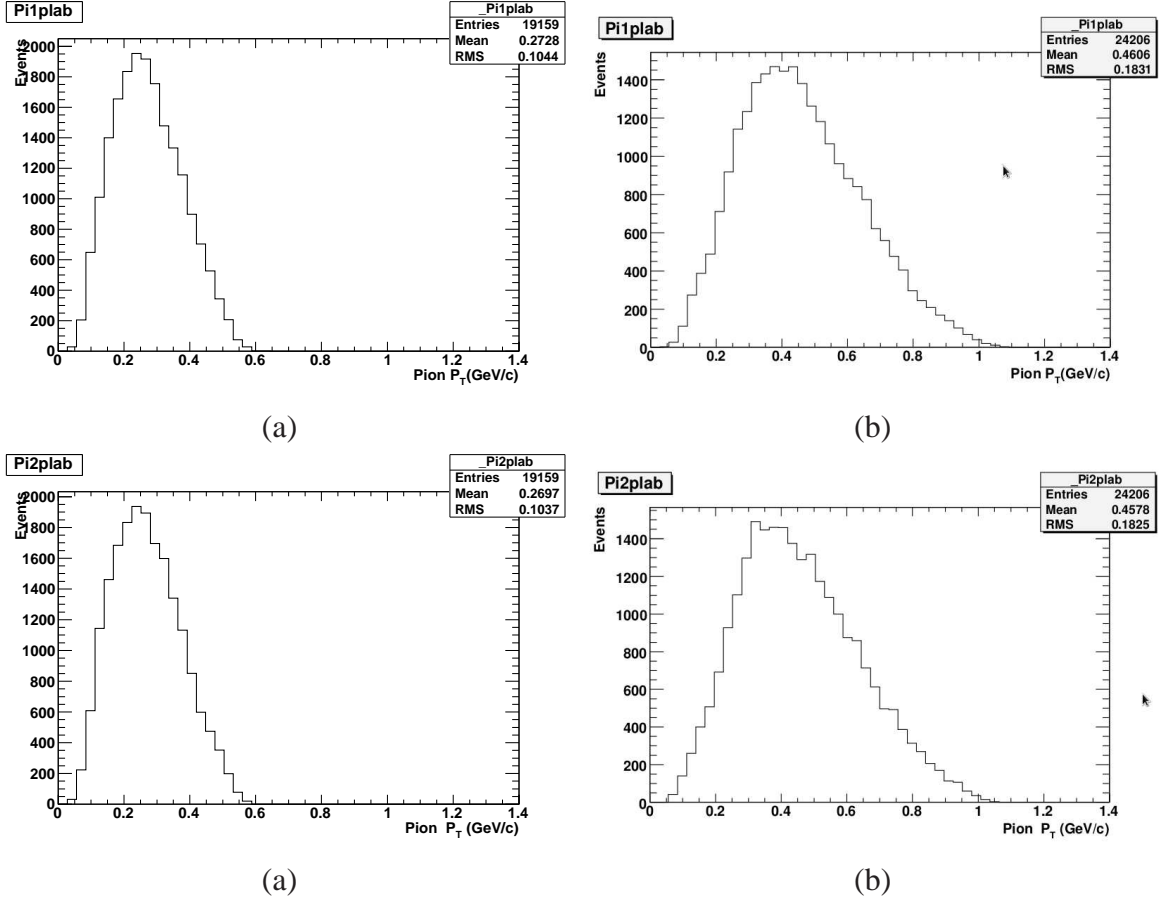


Figure 5.3: Pion transverse momentum distribution for both pions in signal MC. Left plots show for $\Upsilon(2S)$ and right plots show for $\Upsilon(3S)$.

5.5 Systematic uncertainty for $\Upsilon(2S, 3S)$ kinematic fit χ^2

We use the test sample of $\Upsilon(3S, 2S)$ generic MC and the $\Upsilon(3S, 2S)$ onpeak data samples to evaluate the systematic uncertainties for $\Upsilon(3S, 2S)$ kinematic fit χ^2 after unblinding the data samples. We first apply all the optimal selection cuts to the $\Upsilon(3S, 2S)$ generic MC and the $\Upsilon(3S, 2S)$ onpeak data samples except $\Upsilon(3S, 2S)$ kinematic fit χ^2 cut. We then apply $\Upsilon(3S, 2S)$ kinematic fit χ^2 to the both data and MC to calculate the systematic uncertainties. Figure 5.4 shows the $\Upsilon(3S)$ kinematic fit χ^2 distributions for both $\Upsilon(3S)$ and $\Upsilon(2S)$. The relative number of events for both data and MC after applying the kinematic fit χ^2 cut for both $\Upsilon(3S)$ and $\Upsilon(2S)$ are summarized in Table 5.1. The systematic uncertainty due to the $\Upsilon(3S, 2S)$ kinematic fit χ^2 are found to be 1.52% and 2.96% for the $\Upsilon(2S)$ and $\Upsilon(3S)$, respectively.

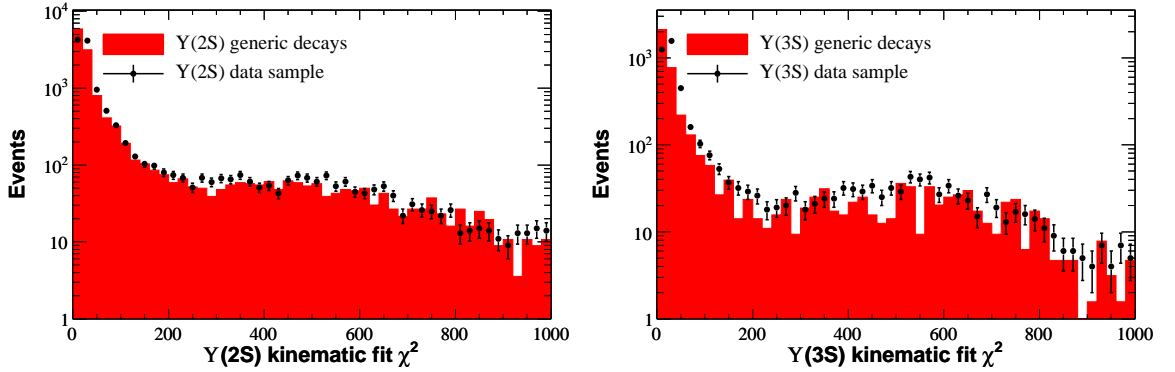


Figure 5.4: $\Upsilon(nS)$ kinematic fit χ^2 distributions after applying all the optimal selection cuts except $\Upsilon(nS)$ kinematic fit χ^2 cut for $\Upsilon(2S)$ (left) and $\Upsilon(3S)$ (right).

Selection cuts	$\Upsilon(2S)$		$\Upsilon(3S)$	
	Data	MC	Data	MC
Pre-selection cuts	13264	7508	4706	2682
$\chi^2_{\Upsilon(3S,2S)} < 300$	11136	6402	3857	2267
Efficiency	0.840 ± 0.0032	0.853 ± 0.0041	0.820 ± 0.0056	0.845 ± 0.007

Table 5.1: The relative number of events in data and MC after applying the $\Upsilon(3S, 2S)$ kinematic fit χ^2 cuts.

5.6 Systematic uncertainty for $\mathcal{B}(\Upsilon(2S, 3S) \rightarrow \pi^+\pi^-\Upsilon(1S))$

The uncertainties on the branching fractions $\mathcal{B}(\Upsilon(2S, 3S) \rightarrow \pi^+\pi^-\Upsilon(1S))$ are 2.2% and 2.3% for $\Upsilon(2S)$ and $\Upsilon(3S)$ datasets, respectively, which are taken from the PDG [92]

5.7 Systematic uncertainty for RF-selection

We study the systematic uncertainties for pion related variables using a control data and MC samples of $\Upsilon(3S, 2S) \rightarrow \pi^+\pi^-\Upsilon(1S)$; $\Upsilon(1S) \rightarrow \mu^+\mu^-$ as mentioned in section 3.4. Figure 5.5 shows the data and MC comparison of the output of RF for both $\Upsilon(2S)$ and $\Upsilon(3S)$ datasets and the relative number of events for both data and MC after applying RF cut are summarized in Table 5.2. Based on the relative difference in the efficiencies of the RF cut on the data and MC, we assign a systematic uncertainty of 2.21% for the $\Upsilon(2S)$ dataset and 2.16% for the $\Upsilon(3S)$ dataset.

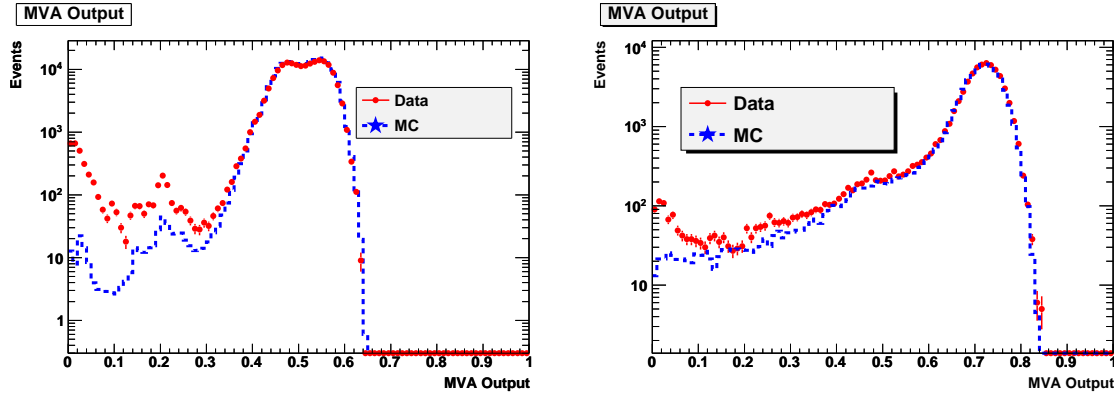


Figure 5.5: Comparison of dipion RF discriminant output between data and Monte-Carlo control samples for $\Upsilon(2S)$ (left) and $\Upsilon(3S)$ (right).

Selection cuts	$\Upsilon(2S)$		$\Upsilon(3S)$	
	Data	MC	Data	MC
Pre-selection cut	189644	1655701	66147	169913
RF > 0.388 (0.568) for $\Upsilon(2S)$ ($\Upsilon(3S)$)	183663	1642292	60116	157834
Efficiency	0.970 ± 0.0004	0.992 ± 0.00007	0.909 ± 0.0011	0.929 ± 0.0006

Table 5.2: The relative number of events in data and MC after applying the RF cut.

For $\Upsilon(3S)$ dataset		
Selection cuts	Efficiency % (Data)	Efficiency % (MC)
e2Mag < 0.2	95.42 ± 0.108	95.74 ± 0.063
Lateral moment [0.06, 0.74]	97.83 ± 0.075	99.13 ± 0.029
Zernika-42 moment < 0.1	98.14 ± 0.070	99.19 ± 0.028
Total Efficiency	92.78 ± 0.134	94.63 ± 0.070

Table 5.3: The relative efficiencies in data and MC after applying the photon related variables.

5.8 Systematic uncertainty due to photon selection

The systematic uncertainty related to the photon selection is measured using an $e^+e^- \rightarrow \gamma\gamma$ sample in which one of the photon converts into an e^+e^- pair in the detector material [57]. The relative selection efficiencies of the photon selection variables are summarized in Table 5.3 and we assign a systematic uncertainty of 1.96% for the photon related variables in the $\Upsilon(3S)$ dataset. Since the photon selection criteria are similar in both $\Upsilon(3S)$ and $\Upsilon(2S)$ datasets, we use the same systematic uncertainty value for $\Upsilon(2S)$.

5.9 Systematic uncertainty for $\Upsilon(nS)$ counting

The systematic uncertainty for $\Upsilon(3S)$ counting has been studied using the on-resonance and off-resonance samples of $\Upsilon(3S)$ data, and the MC samples [93]. The number of $\Upsilon(3S)$ events passing a set of selection criteria in an on-resonance sample is equal to the total number of hadronic events selected less the number of non- $\Upsilon(3S)$ events. The number of non- $\Upsilon(3S)$ events can be expressed in terms of the production cross section, the efficiency to pass the cuts, and the luminosity. The off-resonance sample is used to separate the number of non- $\Upsilon(3S)$ events. A sample of $e^+e^- \rightarrow \gamma\gamma$ events is also used to provide a relative luminosity normalization between between resonant and non-resonant $\Upsilon(3S)$ samples. This study quotes a systematic uncertainty of 0.86%. We also use this systematic uncertainty value for $\Upsilon(2S)$.

5.9.1 Final systematic uncertainties

Table 5.4 summarizes the final systematic uncertainties and their sources for both $\Upsilon(3S)$ and $\Upsilon(2S)$, which will be incorporated to evaluate the branching ratio or upper limit of B.R. in the analysis.

Source	Uncertainty	
	$\Upsilon(2S)$	$\Upsilon(3S)$
Additive systematic uncertainties (events)		
N_s PDF	(0.00 – 0.62)	(0.04 – 0.58)
Fit Bias	0.22	0.17
Total	(0.22 – 0.66)	(0.18 – 0.60)
Multiplicative systematic uncertainties (%)		
Muon-ID	4.30	4.25
Charged tracks	3.73	3.50
$\Upsilon(nS)$ kinematic fit χ^2	1.52	2.96
$\mathcal{B}(\Upsilon(nS) \rightarrow \pi^+\pi^-\Upsilon(1S))$	2.20	2.30
RF selection	2.21	2.16
Photon efficiency	1.96	1.96
$N_{\Upsilon(nS)}$	0.86	0.86
Total	7.00	7.32

Table 5.4: Systematic uncertainties and their sources.

5.10 Chapter Summary

In this chapter, we have summarized the possible sources of the systematic uncertainties for this analysis. These systematic uncertainties are included by convolving the likelihood curve with a Gaussian of width σ_{sys} , which is used to compute the 90% confidence level (C.L.) upper limit using the Bayesian approach with an uniform prior. The details are discussed in the next chapter.

Chapter 6

Results and Conclusion

This chapter presents the 90% C.L. Bayesian upper limits on the product branching fraction of $\mathcal{B}(\Upsilon(1S) \rightarrow \gamma A^0) \times \mathcal{B}(A^0 \rightarrow \mu^+ \mu^-)$ as well as the effective Yukawa coupling of the b-quark to the A^0 as a function of m_{A^0} , which are calculated in the absence of any signal events. Finally, we present the summary and conclusion of this dissertation.

6.1 Upper-limit

As discussed in section 4.6, the trial factor study shows that we find no evidence of signal for the di-muon decay of a light CP -odd scalar particle in the radiative $\Upsilon(1S)$ decays in the $\Upsilon(3S, 2S)$ samples. In the absence of any significant signal yield, we calculate the 90% C.L. upper limit on the product branching fraction $\mathcal{B}(\Upsilon(1S) \rightarrow \gamma A^0) \times \mathcal{B}(A^0 \rightarrow \mu^+ \mu^-)$ as a function of m_{A^0} , including the systematic uncertainties. The systematic uncertainty is included by convolving the likelihood curve with a Gaussian of width σ_{sys} . A convolution is an integral that blends one function with another producing new function that is typically viewed as modified version of the original functions. Mathematically, the convolution of the two functions f and g over an infinite range is given by:

$$h(x) = \int_{-\infty}^{\infty} f(x-y)g(y)dy \quad (6.1)$$

where $h(x)$ is the modified version of original functions f and g after the convolution. We plot the negative log likelihood (NLL) as a function of branching fraction (BF) and integrate it from zero upward until we find an integral which yields 90% of the total integral (above zero) under the likelihood curve to compute the 90% confidence level Bayesian upper limits. The BF is defined as:

$$BF = \frac{N_{sig}}{\epsilon \cdot \mathcal{B} \cdot N_{\Upsilon(nS)}} \quad (6.2)$$

where N_{sig} is the number of the fitted signal yield, ϵ is the signal selection efficiency, \mathcal{B} is the branching fraction of $\Upsilon(2S, 3S) \rightarrow \pi^+\pi^-\Upsilon(1S)$ transitions, and $N_{\Upsilon(nS)}$ is the number of $\Upsilon(2S, 3S)$ mesons used in this analysis. For combining the results of the $\Upsilon(2S, 3S)$ datasets, we add the log of the $\Upsilon(2S, 3S)$ likelihoods. Figure 6.1 shows the likelihood function as a function of BF at selected mass points for $\Upsilon(2S)$, $\Upsilon(3S)$ and combined data of $\Upsilon(2S, 3S)$. The correlated and uncorrelated systematic uncertainties are taken into account for combining the two datasets. The systematic uncertainties of $\Upsilon(nS)$ counting, photon efficiency, tracking and PID are considered as correlated systematic uncertainties and rest of the systematic uncertainties discussed Table 5.4 are considered as uncorrelated systematic uncertainties. The J/ψ mass region in the $\Upsilon(3S)$ dataset, defined as $3.045 \leq m_{red} \leq 3.162$ GeV/ c^2 , is excluded from the search due to a large background from $J/\psi \rightarrow \mu^+\mu^-$. Figure 6.2 shows the 90% C.L. upper limits on $\mathcal{B}(\Upsilon(1S) \rightarrow \gamma A^0) \times \mathcal{B}(A^0 \rightarrow \mu^+\mu^-)$ as a function of m_{A^0} . The limits vary between $(0.37 - 8.97) \times 10^{-6}$ for the $\Upsilon(2S)$ dataset, $(1.13 - 24.2) \times 10^{-6}$ for the $\Upsilon(3S)$ dataset, and $(0.28 - 9.7) \times 10^{-6}$ for the combined $\Upsilon(2S, 3S)$ dataset.

The branching fractions of $\mathcal{B}(\Upsilon(nS) \rightarrow \gamma A^0)$ ($n = 1, 2, 3$) are related to the effective Yukawa coupling (f_r) of the b -quark to the A^0 via Equation 1.23. The value of f_r incorporates the m_{A^0} dependent QCD and relativistic corrections to $\mathcal{B}(\Upsilon(nS) \rightarrow \gamma A^0)$ [48], as well as the leptonic width of $\Upsilon(nS) \rightarrow l^+l^-$ [50]. These corrections are as large as 30% to first order in strong coupling constant (α_s), but have comparable uncertainties [94]. The 90% C.L. upper limits on $f_r^2 \times \mathcal{B}(A^0 \rightarrow \mu^+\mu^-)$ for combined $\Upsilon(2S, 3S)$ datasets range from 0.54×10^{-6} to 2.99×10^{-4} depending upon the mass of A^0 , which is shown in Figure 6.4(a). For comparison, the results from previous *BABAR* measurements of $\Upsilon(2S, 3S) \rightarrow \gamma A^0$, $\mathcal{B}(A^0 \rightarrow \mu^+\mu^-)$ [54] are also shown. We combine our results with previous *BABAR* measurements [54], taking into account both correlated and uncorrelated uncertainties. Figure 6.3 shows the likelihood

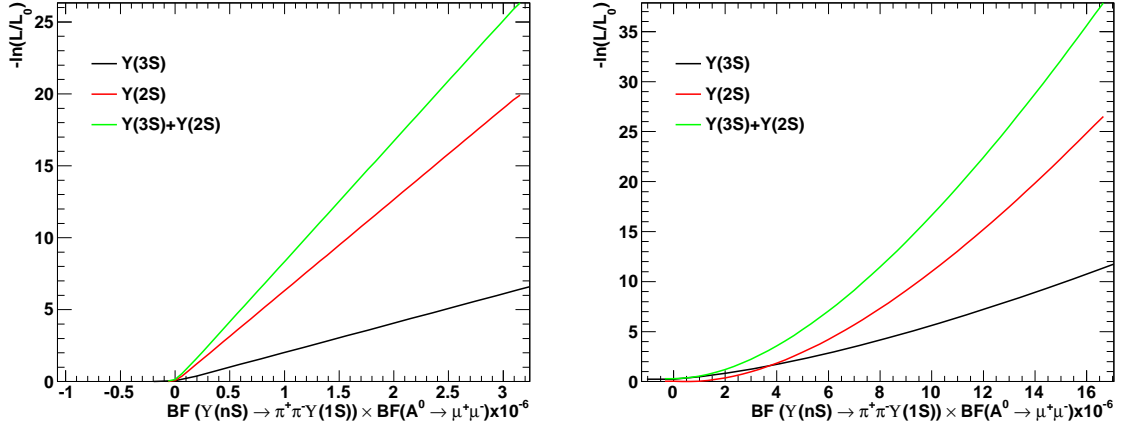


Figure 6.1: The likelihood function as a function of branching fraction (B.F.) for the Higgs mass of (a) $m_{A^0} = 0.212 \text{ GeV}/c^2$ and (b) $m_{A^0} = 8.21 \text{ GeV}/c^2$.

function as a function of f_T at selected m_{A^0} points for the combined data of $\Upsilon(2S, 3S)$, previous *BABAR* measurements [54] and combination of these two measurements. The combined upper limits on $f_T^2 \times \mathcal{B}(A^0 \rightarrow \mu^+ \mu^-)$ for these two measurements vary in the range of $(0.29-40) \times 10^{-6}$ for $m_{A^0} \leq 9.2 \text{ GeV}/c^2$ (Figure 6.4(b)).

6.2 Summary and Conclusion

This thesis describes a search for di-muon decays of a low-mass Higgs boson in the fully reconstructed decay chain of $\Upsilon(2S, 3S) \rightarrow \pi^+ \pi^- \Upsilon(1S)$, $\Upsilon(1S) \rightarrow \gamma A^0$, $A^0 \rightarrow \mu^+ \mu^-$. The $\Upsilon(1S)$ sample is selected by tagging the pion pair in the $\Upsilon(2S, 3S) \rightarrow \pi^+ \pi^- \Upsilon(1S)$ transitions, using a data sample of $(92.8 \pm 0.8) \times 10^6$ $\Upsilon(2S)$ and $(116.8 \pm 1.0) \times 10^6$ $\Upsilon(3S)$ mesons collected with the *BABAR* detector at the PEP-II asymmetric-energy $e^+ e^-$ collider located at SLAC National Accelerator Laboratory. The A^0 is assumed to be a scalar or pseudoscalar particle with a negligible decay width compared to the experimental resolution [30]. We find no evidence for A^0 production and set 90% confidence level (C.L.) upper limits on the product branching fraction $\mathcal{B}(\Upsilon(1S) \rightarrow \gamma A^0) \times \mathcal{B}(A^0 \rightarrow \mu^+ \mu^-)$ in the range of $(0.28 - 9.72) \times 10^{-6}$ for $0.212 \leq m_{A^0} \leq 9.20 \text{ GeV}/c^2$. These results improve the current best limits by a factor of 2–3 for $m_{A^0} < 1.2 \text{ GeV}/c^2$ and are comparable to the previous *BABAR* result [54] in the mass range of $1.20 < m_{A^0} < 3.6 \text{ GeV}/c^2$. Within this range, our limits rule out substantial amount of the parameter space allowed by the light Higgs [41] and axion [43] model. We

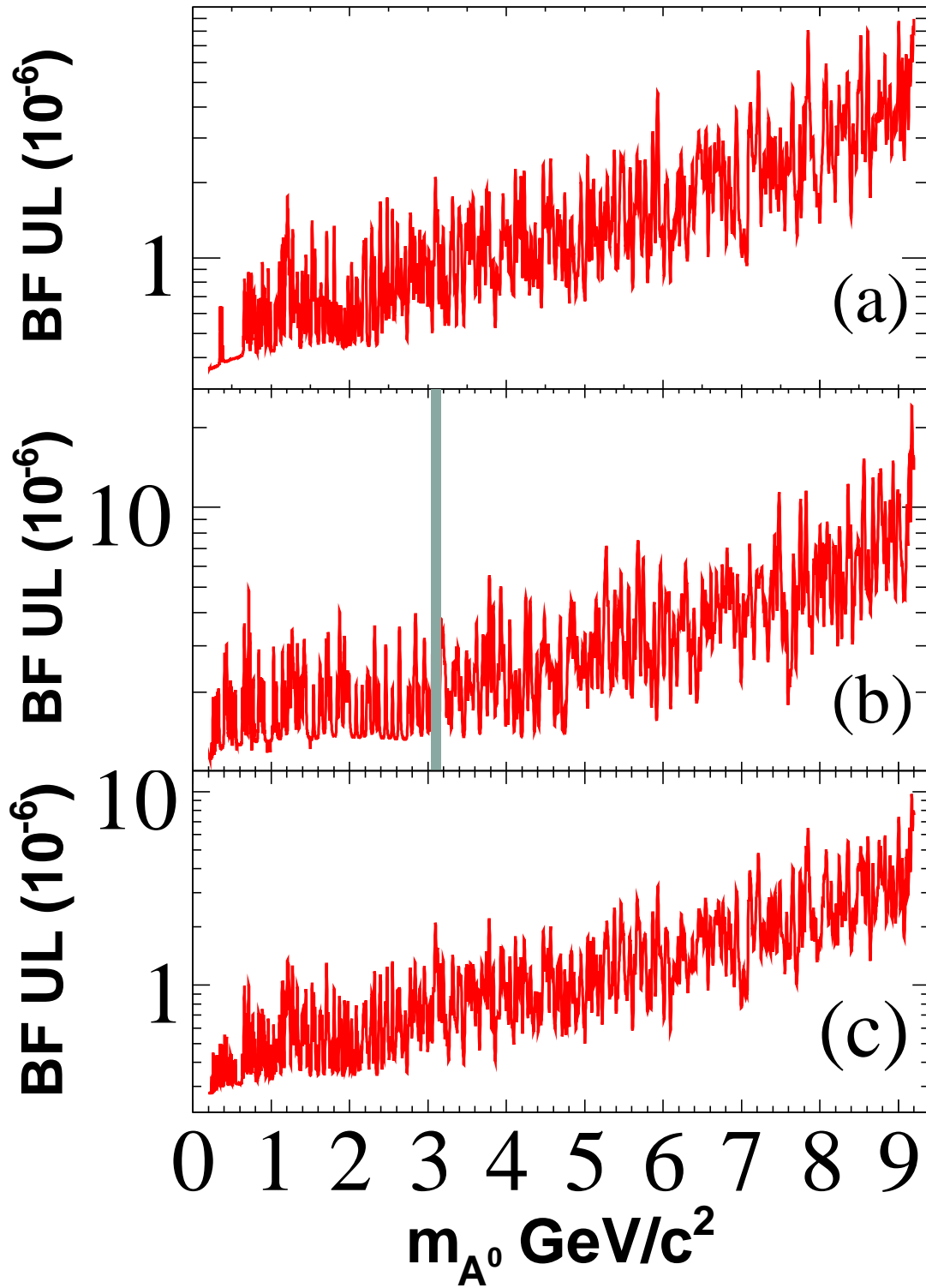


Figure 6.2: The 90% C.L. upper limit on the product of branching fractions $\mathcal{B}(\Upsilon(1S) \rightarrow \gamma A^0) \times \mathcal{B}(A^0 \rightarrow \mu^+ \mu^-)$ for (a) the $\Upsilon(2S)$ dataset, (b) the $\Upsilon(3S)$ dataset and (c) the combined $\Upsilon(2S, 3S)$ dataset. The shaded area shows the region of the J/ψ resonance, excluded from the search in the $\Upsilon(3S)$ dataset.

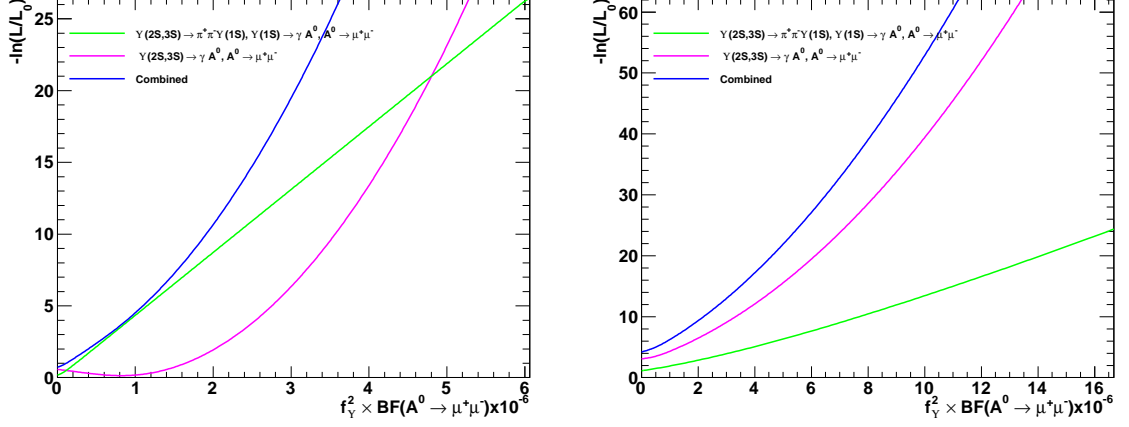


Figure 6.3: The likelihood function as a function of effective Yukawa coupling of b -quarks to the A^0 (f_γ) for the Higgs mass of (a) $m_{A^0} = 0.214 \text{ GeV}/c^2$ and (b) $m_{A^0} = 5.60 \text{ GeV}/c^2$.

also combine our results with previous *BABAR* results of $\Upsilon(2S, 3S) \rightarrow \gamma A^0, A^0 \rightarrow \mu^+ \mu^-$ to set limits on the effective coupling (f_γ) of the b -quarks to the A^0 , $f_\gamma^2 \times \mathcal{B}(A^0 \rightarrow \mu^+ \mu^-)$, at the level of $(0.29 - 40.18) \times 10^{-6}$ for $0.212 \leq m_{A^0} \leq 9.2 \text{ GeV}/c^2$. The combined limits on the product $f_\gamma^2 \times \mathcal{B}(A^0 \rightarrow \mu^+ \mu^-)$ are the most stringent to date, and significantly constrain the theoretical Models. A high luminosity $e^+ e^-$ asymmetric energy Super-B factory and International Linear Collider (ILC) experiments can significantly improve the searches of these low-mass scalar particles, difficult to explore by the LHC, and elucidate the structure of the new Physics.

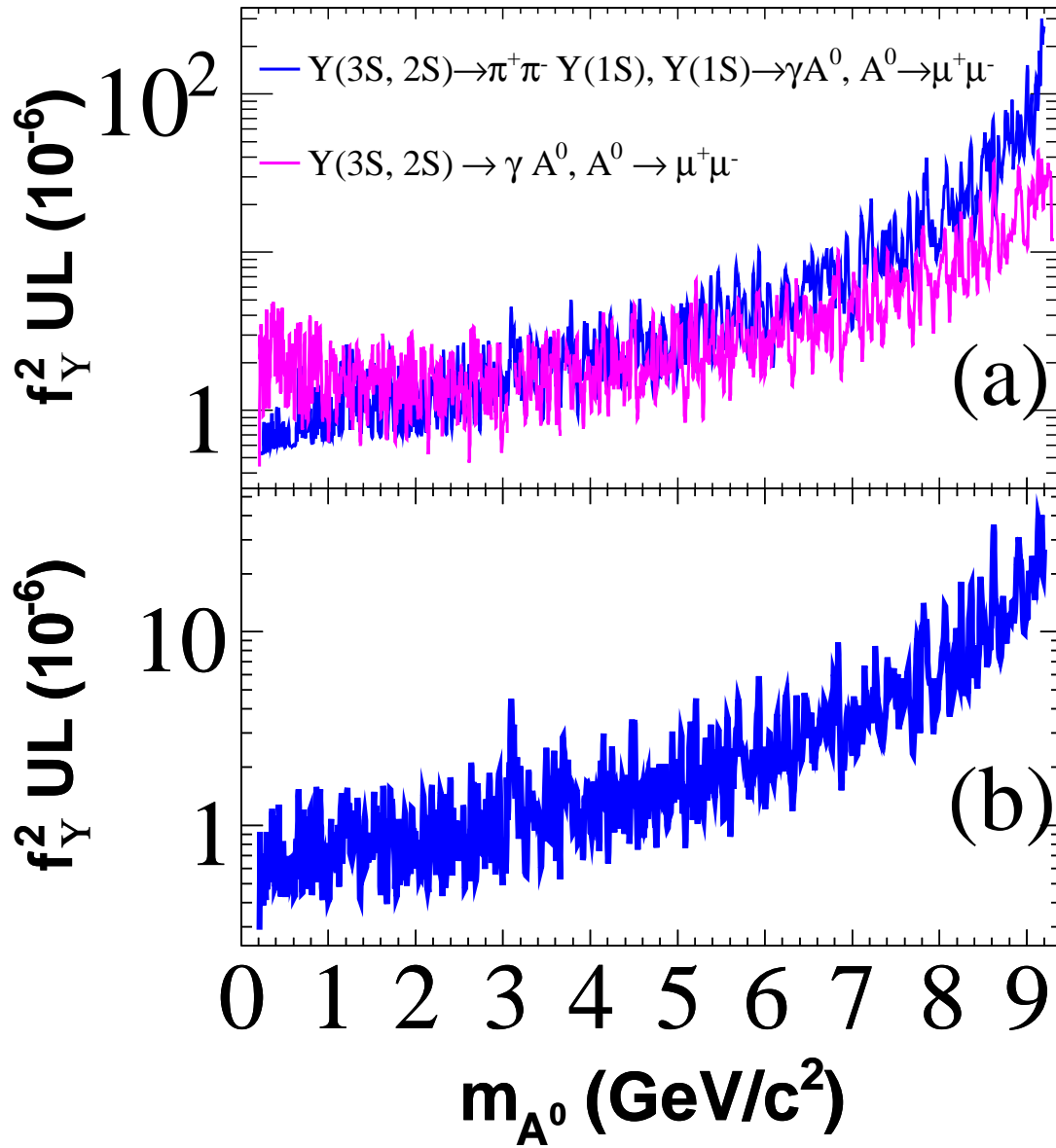


Figure 6.4: The 90% C.L. upper limit on the effective Yukawa coupling (f_Y) of bound b -quarks to the A^0 , $f_Y^2 \times \mathcal{B}(A^0 \rightarrow \mu^+ \mu^-)$, for (a) this and previous *BABAR* measurement of $Y(2S, 3S) \rightarrow \gamma A^0$, $A^0 \rightarrow \mu^+ \mu^-$ [54], and (b) the combined limit.

Appendix A

Signal PDFs

A.1 Signal PDFs for $\Upsilon(2S)$

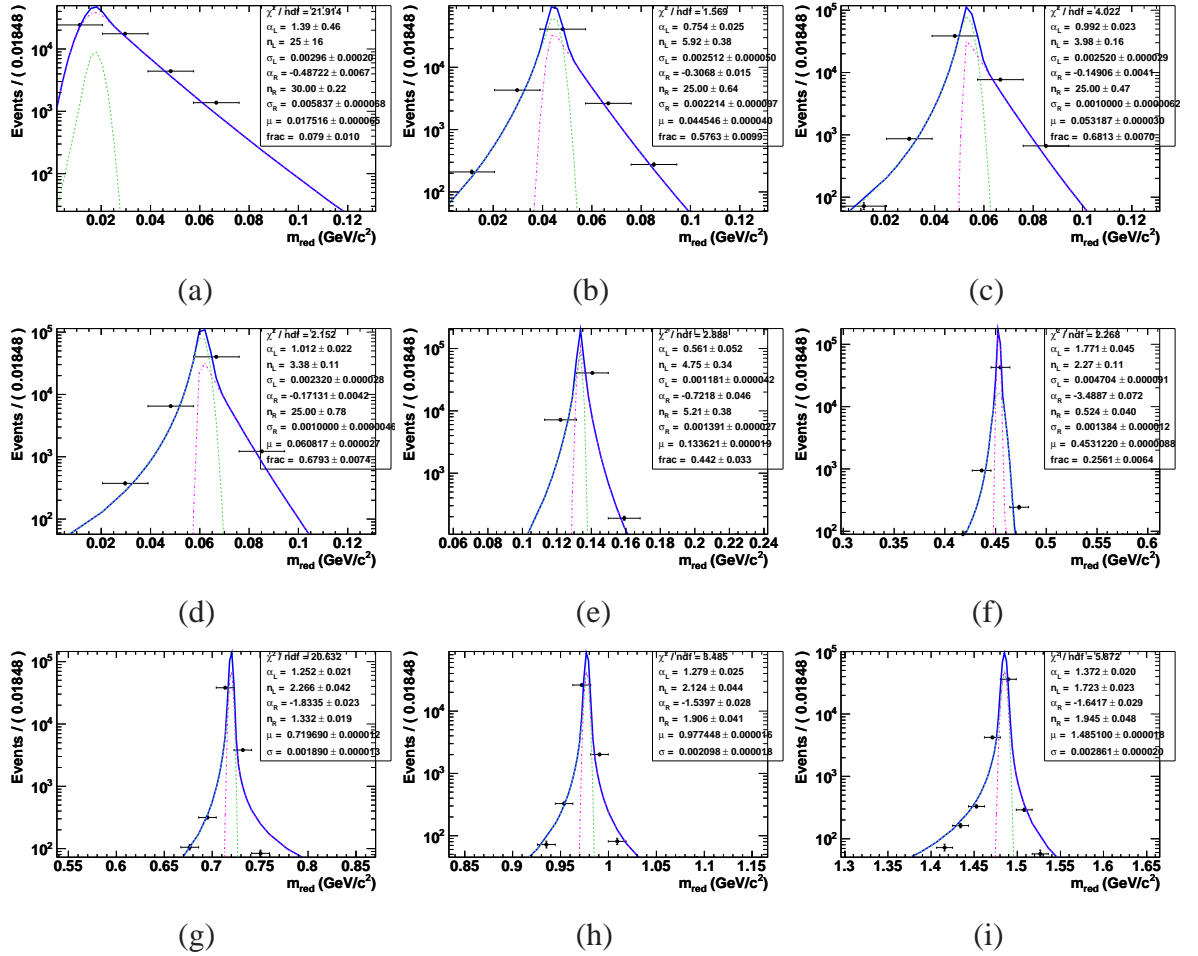


Figure A.1: Signal PDFs for the Higgs mass of (a) $m_{A^0} = 0.212 \text{ GeV}/c^2$ (b) $m_{A^0} = 0.216 \text{ GeV}/c^2$ (c) $m_{A^0} = 0.218 \text{ GeV}/c^2$ (d) $m_{A^0} = 0.220 \text{ GeV}/c^2$ (e) $m_{A^0} = 0.250 \text{ GeV}/c^2$ (f) $m_{A^0} = 0.500 \text{ GeV}/c^2$ (g) $m_{A^0} = 0.75 \text{ GeV}/c^2$ (h) $m_{A^0} = 1.0 \text{ GeV}/c^2$ and (i) $m_{A^0} = 1.5 \text{ GeV}/c^2$.

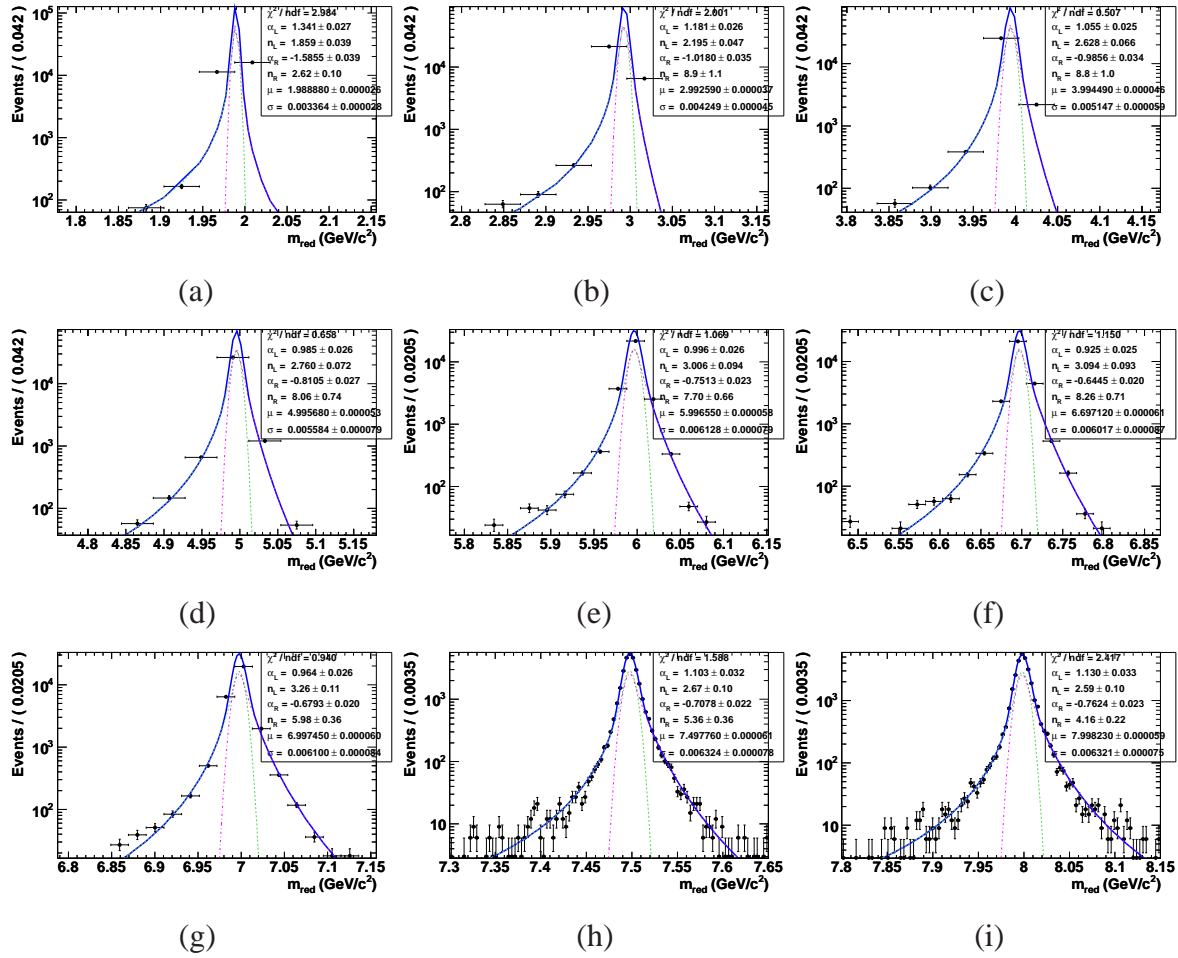


Figure A.2: Signal PDFs for the Higgs mass of (a) $m_{A^0} = 2.0$ GeV/c² (b) $m_{A^0} = 3.0$ GeV/c² (c) $m_{A^0} = 4.0$ GeV/c² (d) $m_{A^0} = 5.0$ GeV/c² (e) $m_{A^0} = 6.0$ GeV/c² (f) $m_{A^0} = 6.7$ GeV/c² (g) $m_{A^0} = 7.0$ GeV/c² (h) $m_{A^0} = 7.5$ GeV/c² and (i) $m_{A^0} = 8.0$ GeV/c².

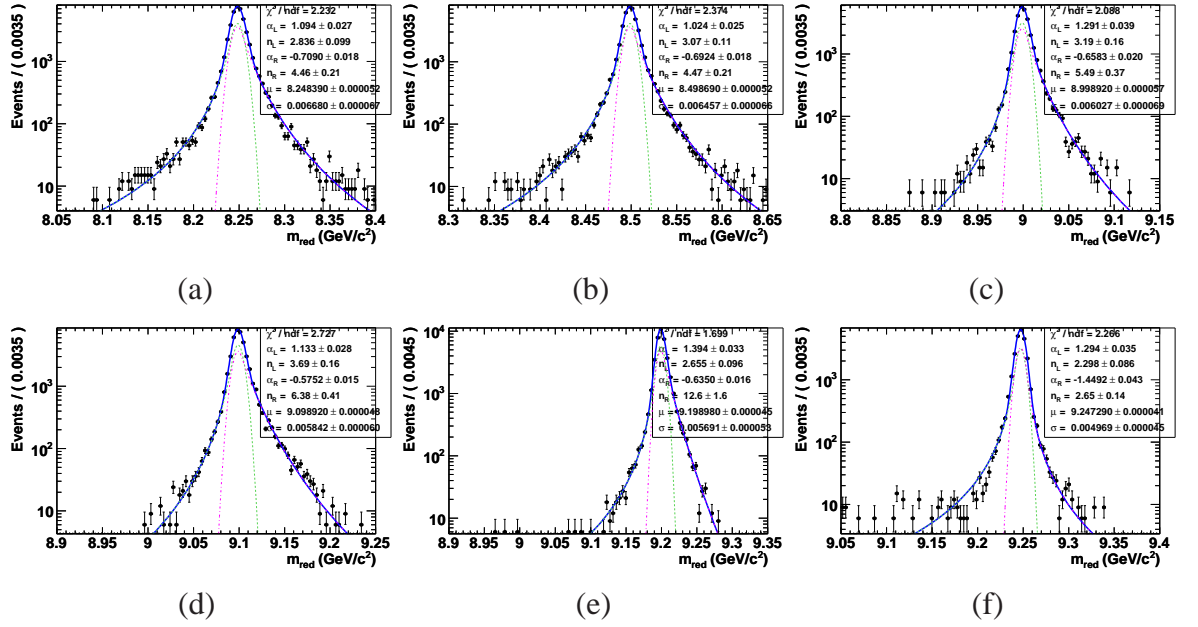


Figure A.3: Signal PDFs for the Higgs mass of (a) $m_{A^0} = 8.25$ GeV/c² (b) $m_{A^0} = 8.50$ GeV/c² (c) $m_{A^0} = 9.0$ GeV/c² (d) $m_{A^0} = 9.10$ GeV/c² (e) $m_{A^0} = 9.20$ GeV/c² and (f) $m_{A^0} = 9.25$ GeV/c².

A.2 Signal PDFs for $\Upsilon(3S)$

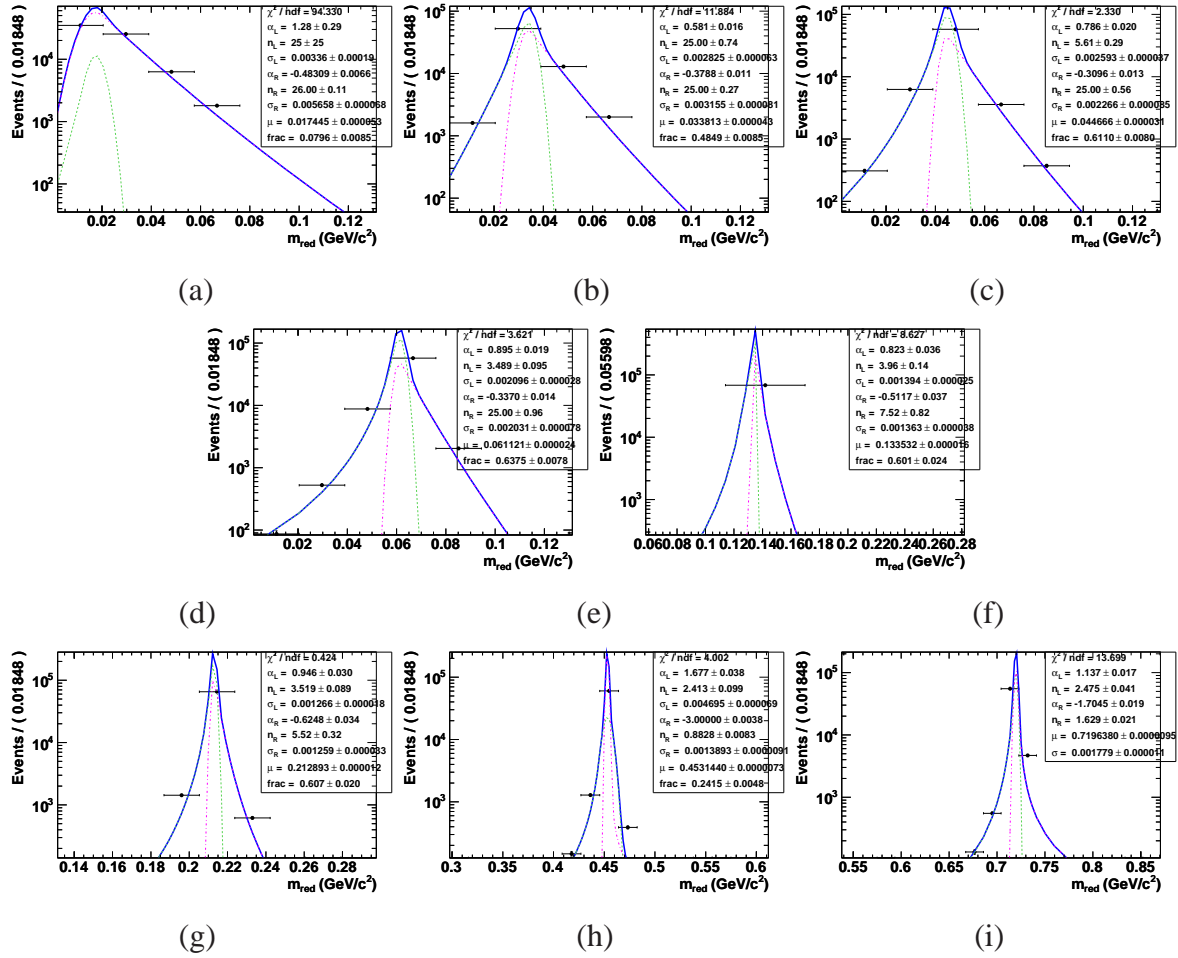


Figure A.4: Signal PDFs for the Higgs mass of (a) $m_{A^0} = 0.212 \text{ GeV}/c^2$ (b) $m_{A^0} = 0.214 \text{ GeV}/c^2$ (c) $m_{A^0} = 0.216 \text{ GeV}/c^2$ (d) $m_{A^0} = 0.218 \text{ GeV}/c^2$ (e) $m_{A^0} = 0.220 \text{ GeV}/c^2$ (f) $m_{A^0} = 0.250 \text{ GeV}/c^2$ (g) $m_{A^0} = 0.300 \text{ GeV}/c^2$ (h) $m_{A^0} = 0.50 \text{ GeV}/c^2$ and (i) $m_{A^0} = 0.75 \text{ GeV}/c^2$.

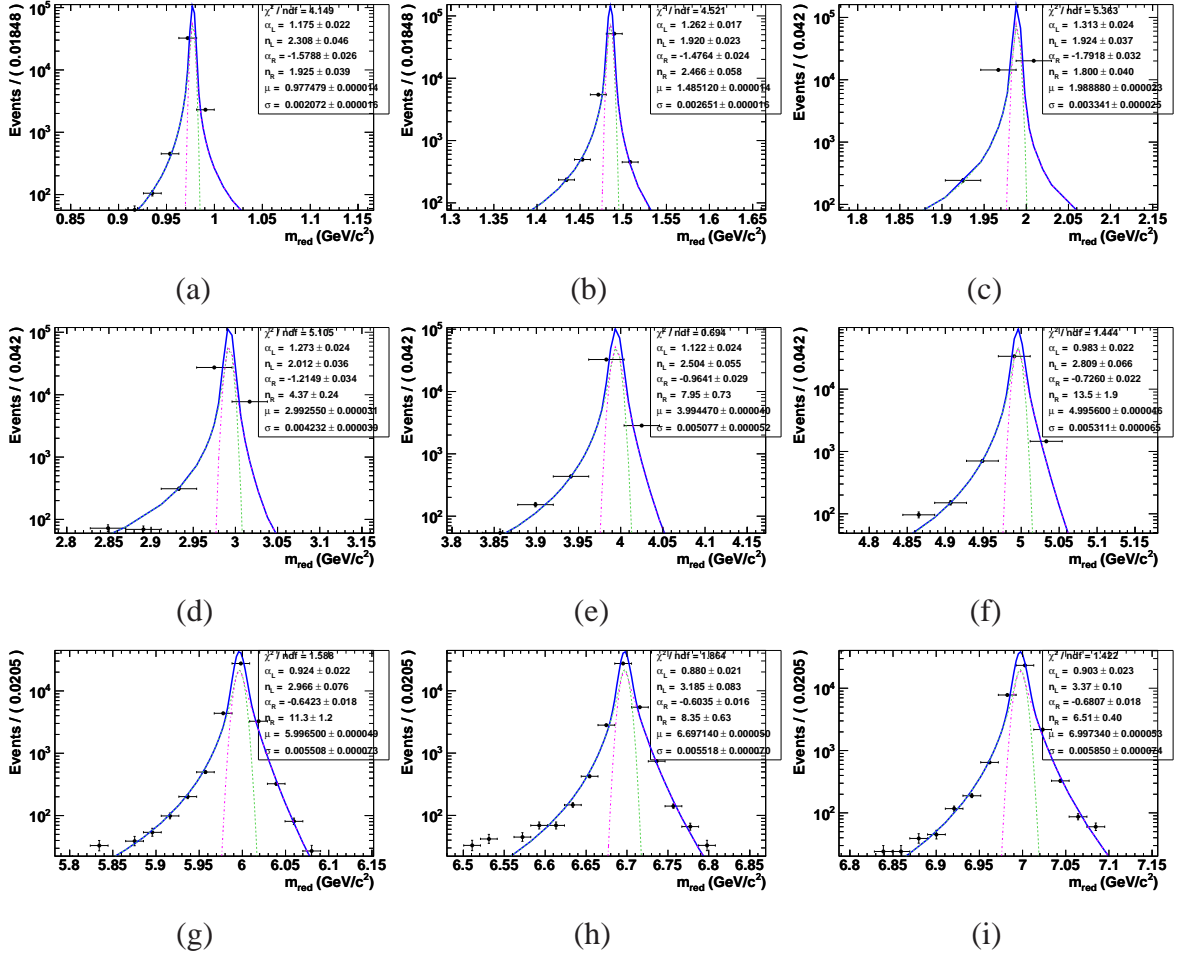


Figure A.5: Signal PDFs for the Higgs mass of (a) $m_{A^0} = 1.0 \text{ GeV}/c^2$, (b) $m_{A^0} = 1.5 \text{ GeV}/c^2$ (c) $m_{A^0} = 2.0 \text{ GeV}/c^2$ (d) $m_{A^0} = 3.0 \text{ GeV}/c^2$ (e) $m_{A^0} = 4.0 \text{ GeV}/c^2$ (f) $m_{A^0} = 5.0 \text{ GeV}/c^2$ (g) $m_{A^0} = 6.0 \text{ GeV}/c^2$ (h) $m_{A^0} = 6.7 \text{ GeV}/c^2$ and (i) $m_{A^0} = 7.0 \text{ GeV}/c^2$.

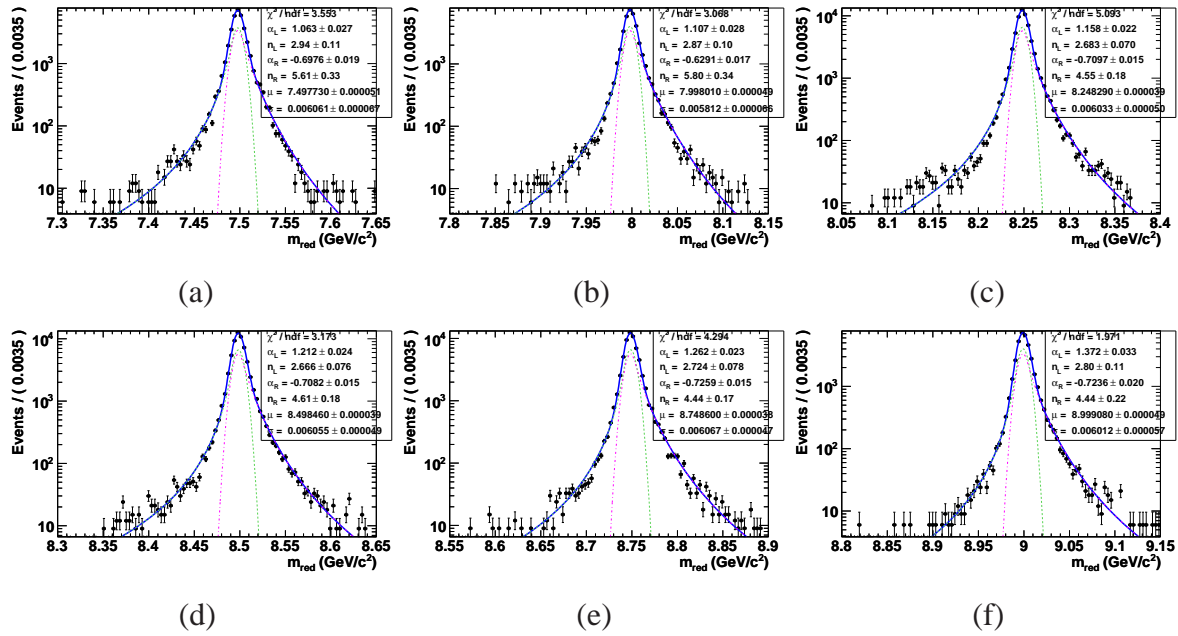


Figure A.6: Signal PDFs for reduced mass distribution for the Higgs mass of (a) $m_{A^0} = 7.5 \text{ GeV}/c^2$ (b) $m_{A^0} = 8.0 \text{ GeV}/c^2$ (c) $m_{A^0} = 8.25 \text{ GeV}/c^2$ (d) $m_{A^0} = 8.50 \text{ GeV}/c^2$ (e) $m_{A^0} = 8.75 \text{ GeV}/c^2$ and (f) $m_{A^0} = 9.0 \text{ GeV}/c^2$.

Appendix B

Background PDFs

B.1 Background PDFs for $\Upsilon(2S)$

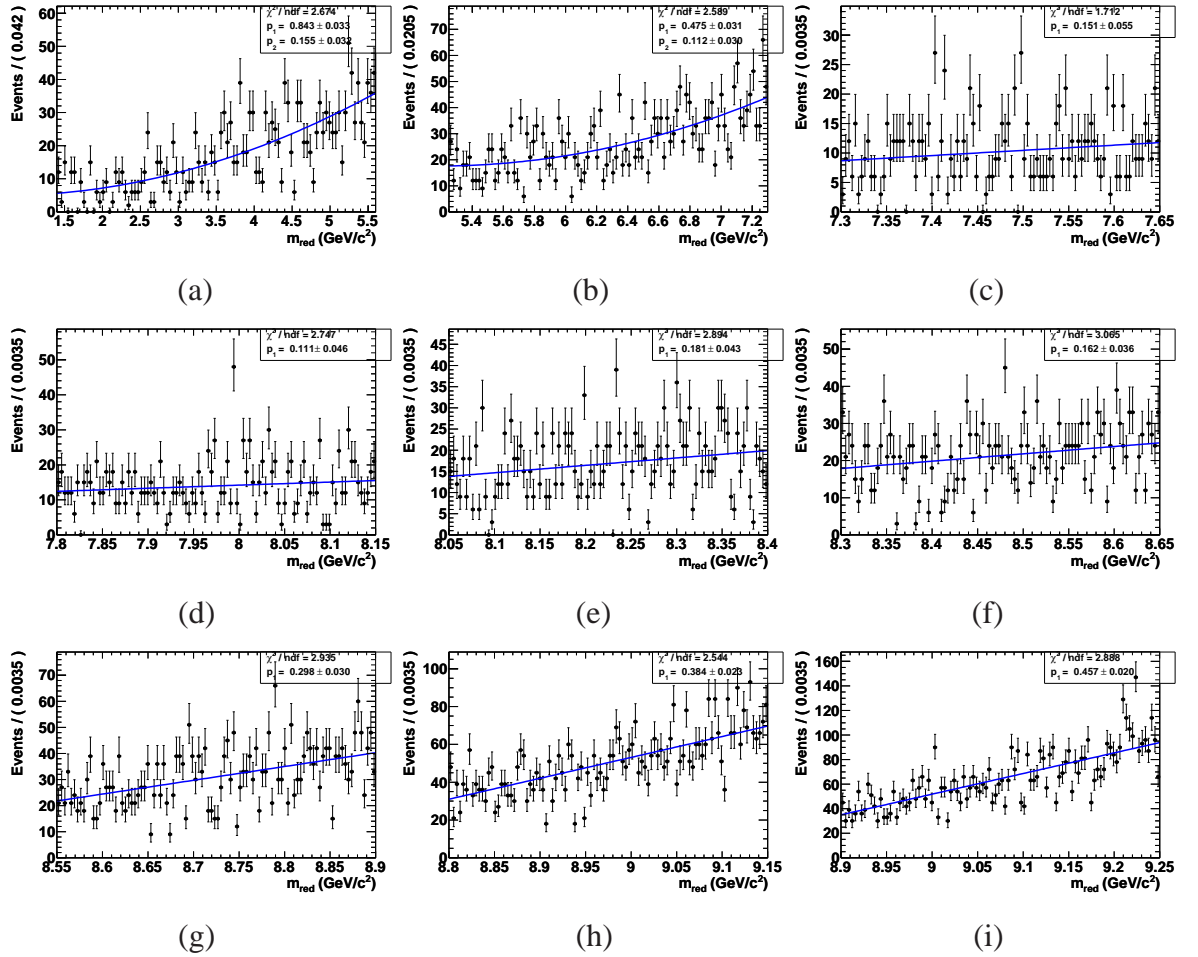


Figure B.1: Background PDFs for m_{red} distribution for the Higgs mass of (a) $1.502 \leq m_{A^0} \leq 5.50 \text{ GeV}/c^2$ (b) $5.25 \leq m_{A^0} \leq 7.3 \text{ GeV}/c^2$ (c) $m_{A^0} = 7.5 \text{ GeV}/c^2$ (d) $m_{A^0} = 8.0 \text{ GeV}/c^2$ (e) $m_{A^0} = 8.25 \text{ GeV}/c^2$ (f) $m_{A^0} = 8.5 \text{ GeV}/c^2$ (g) $m_{A^0} = 8.75 \text{ GeV}/c^2$ (h) $m_{A^0} = 9.0 \text{ GeV}/c^2$ and (i) $m_{A^0} = 9.1 \text{ GeV}/c^2$.

B.2 Background PDFs for $\gamma(3S)$

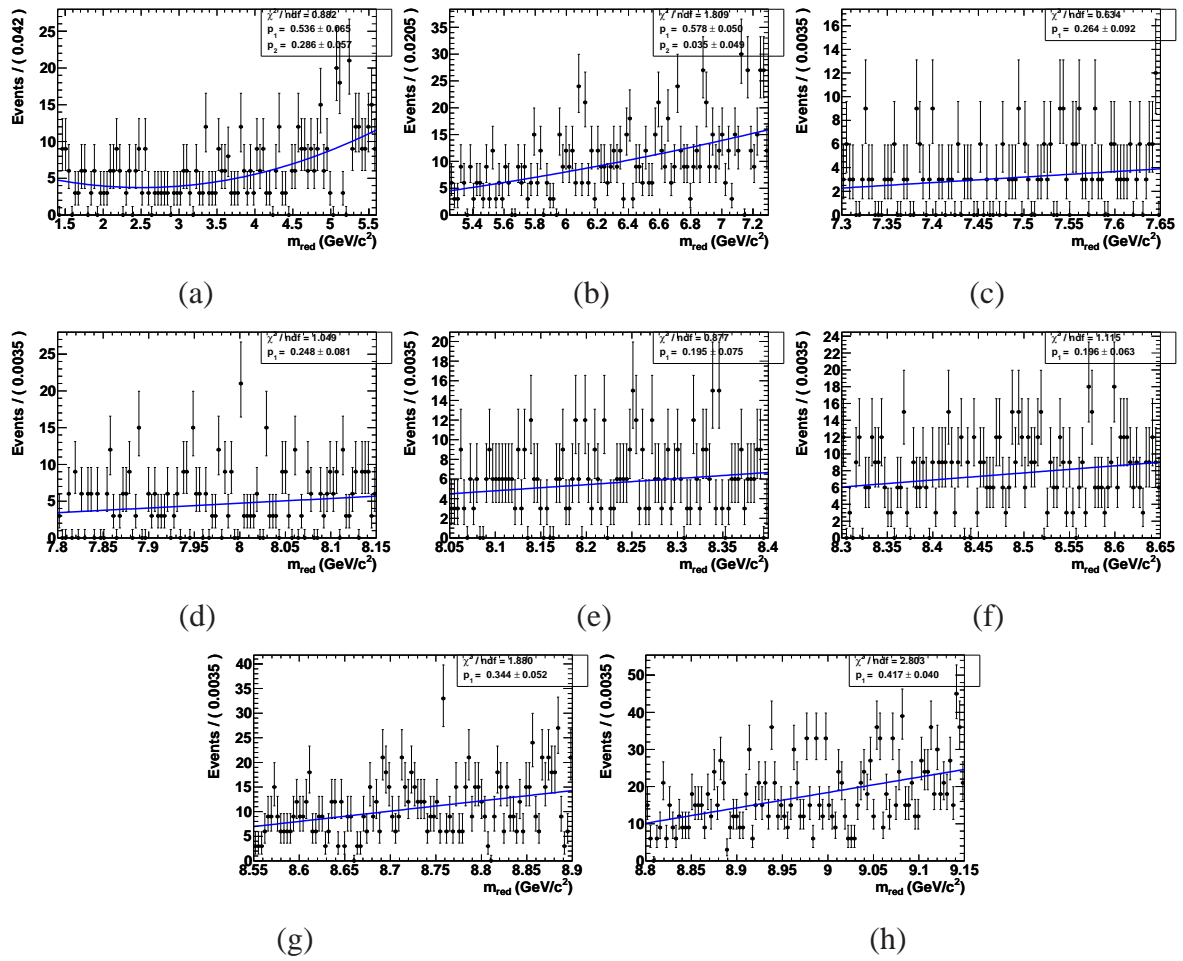


Figure B.2: Background PDFs for m_{red} distribution for the Higgs mass of (a) $1.502 \leq m_{A^0} \leq 5.50 \text{ GeV}/c^2$ (b) $5.25 \leq m_{A^0} \leq 7.3 \text{ GeV}/c^2$ (c) $m_{A^0} = 7.5 \text{ GeV}/c^2$ (d) $m_{A^0} = 8.0 \text{ GeV}/c^2$ (e) $m_{A^0} = 8.25 \text{ GeV}/c^2$ (f) $m_{A^0} = 8.5 \text{ GeV}/c^2$ (g) $m_{A^0} = 8.75 \text{ GeV}/c^2$ and (h) $m_{A^0} = 9.0 \text{ GeV}/c^2$.

Appendix C

Toy Monte Carlo Results

C.1 For $\gamma(2S)$

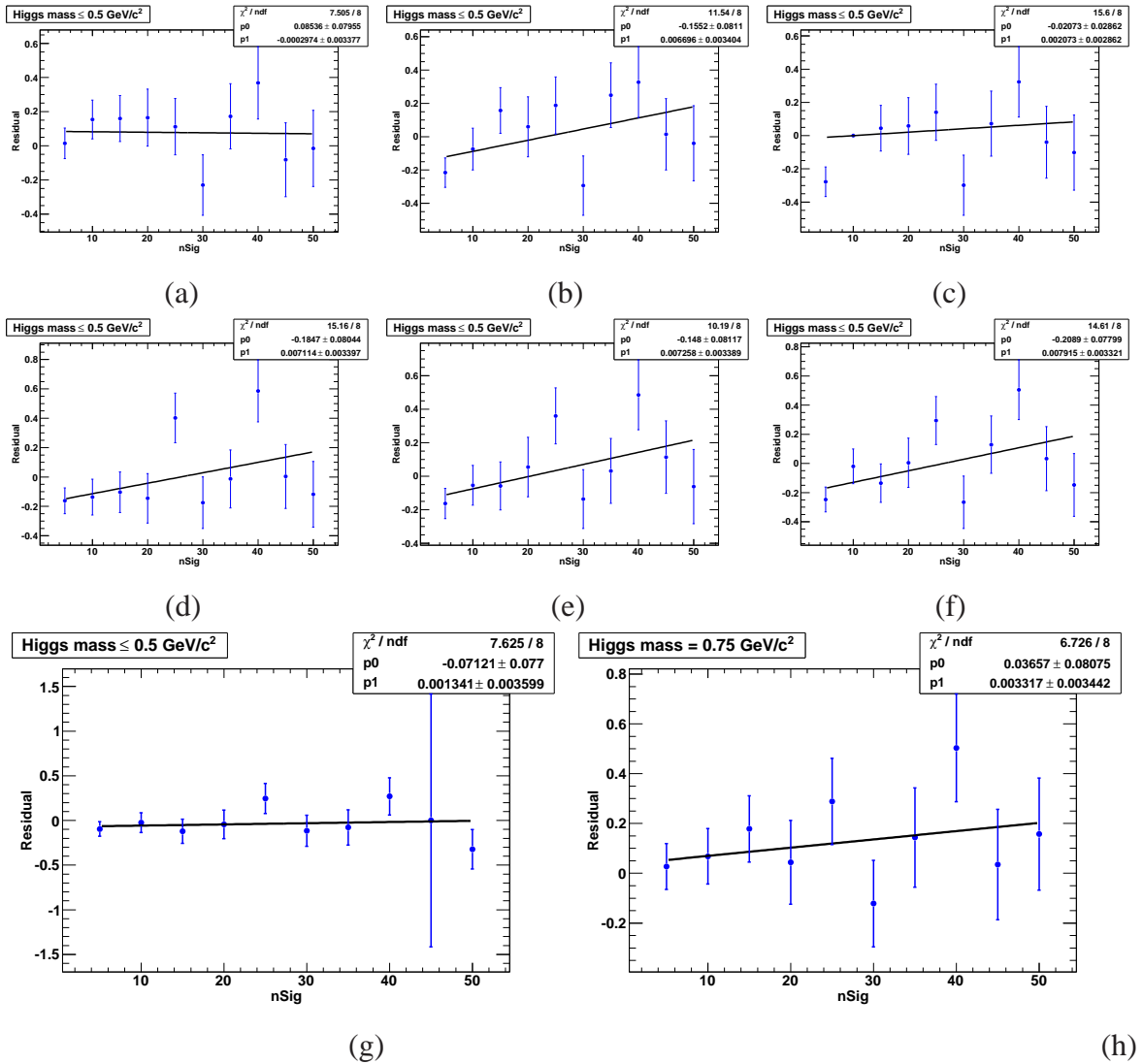


Figure C.1: Fit residuals for the Higgs mass of (a) $m_{A^0} = 0.212 \text{ GeV}/c^2$ (b) $m_{A^0} = 0.214 \text{ GeV}/c^2$ (c) $m_{A^0} = 0.216 \text{ GeV}/c^2$ (d) $m_{A^0} = 0.218 \text{ GeV}/c^2$ (e) $m_{A^0} = 0.220 \text{ GeV}/c^2$ (f) $m_{A^0} = 0.250 \text{ GeV}/c^2$ (g) $m_{A^0} = 0.50 \text{ GeV}/c^2$ and (h) $m_{A^0} = 0.75 \text{ GeV}/c^2$.

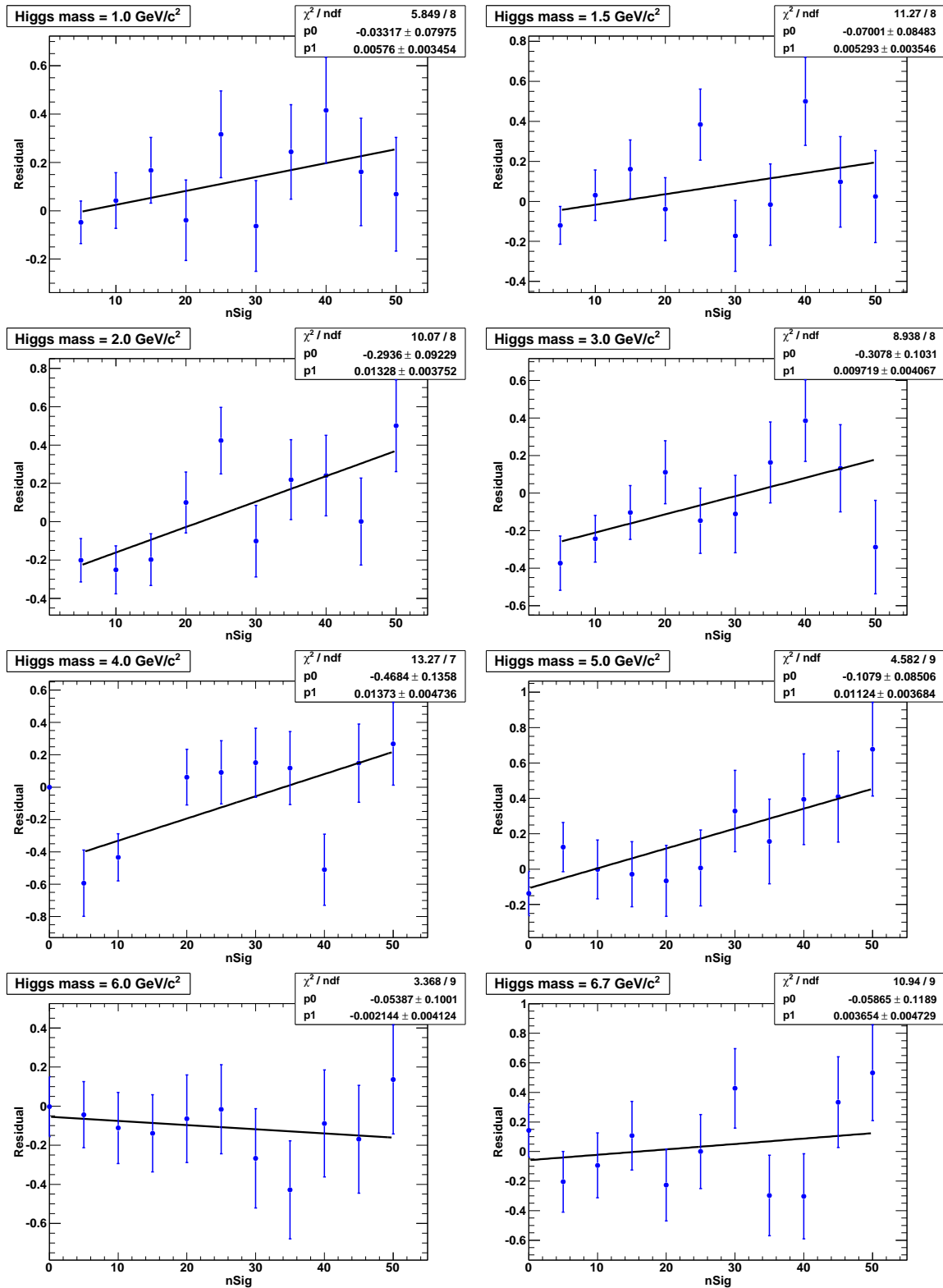


Figure C.2: Fit residuals for the number of signal events in the toy Monte Carlo experiments generated for each Higgs mass points.

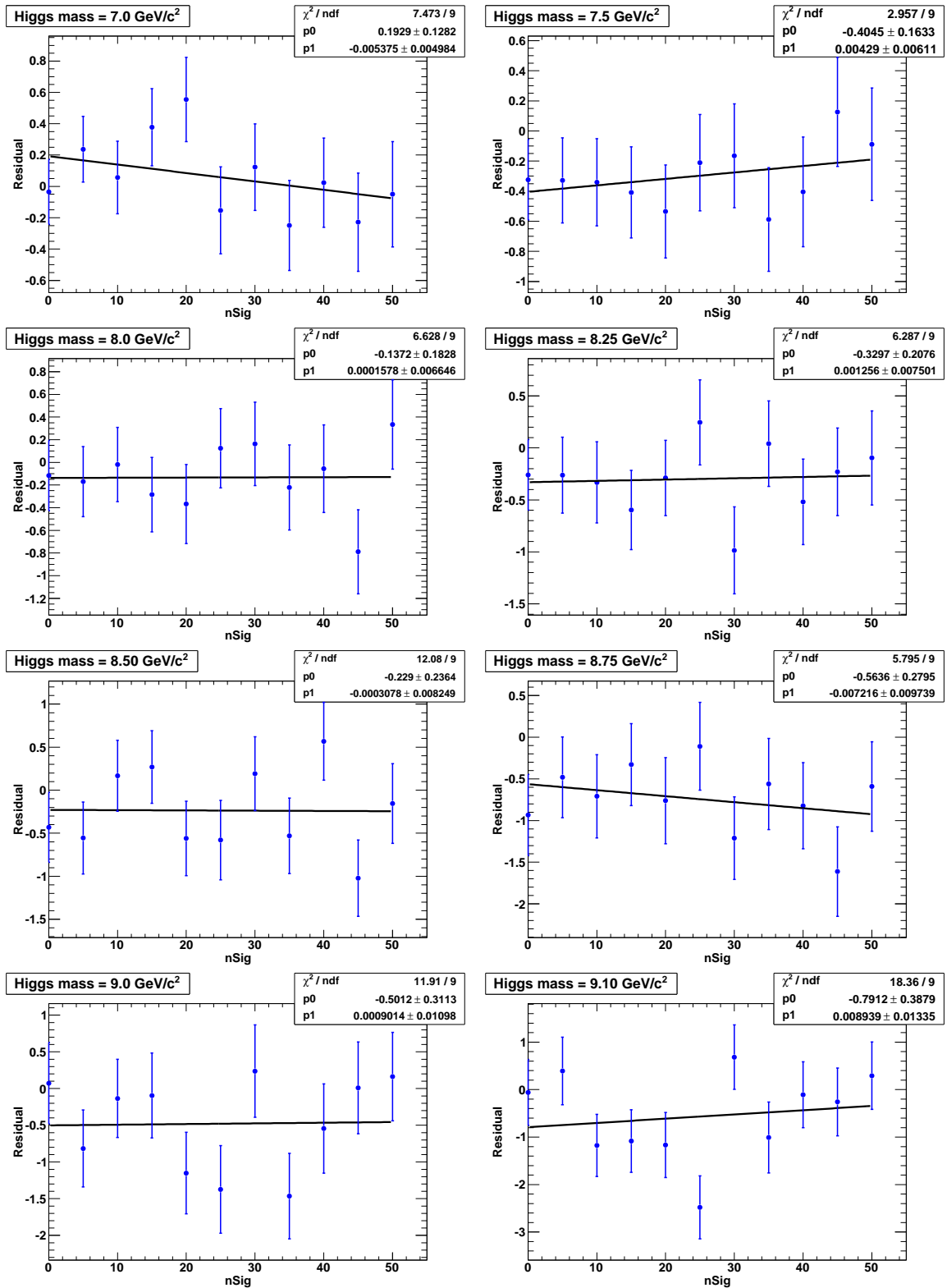


Figure C.3: Fit residuals for the number of signal events in the toy Monte Carlo experiments generated for each Higgs mass points.

C.2 For $\Upsilon(3S)$

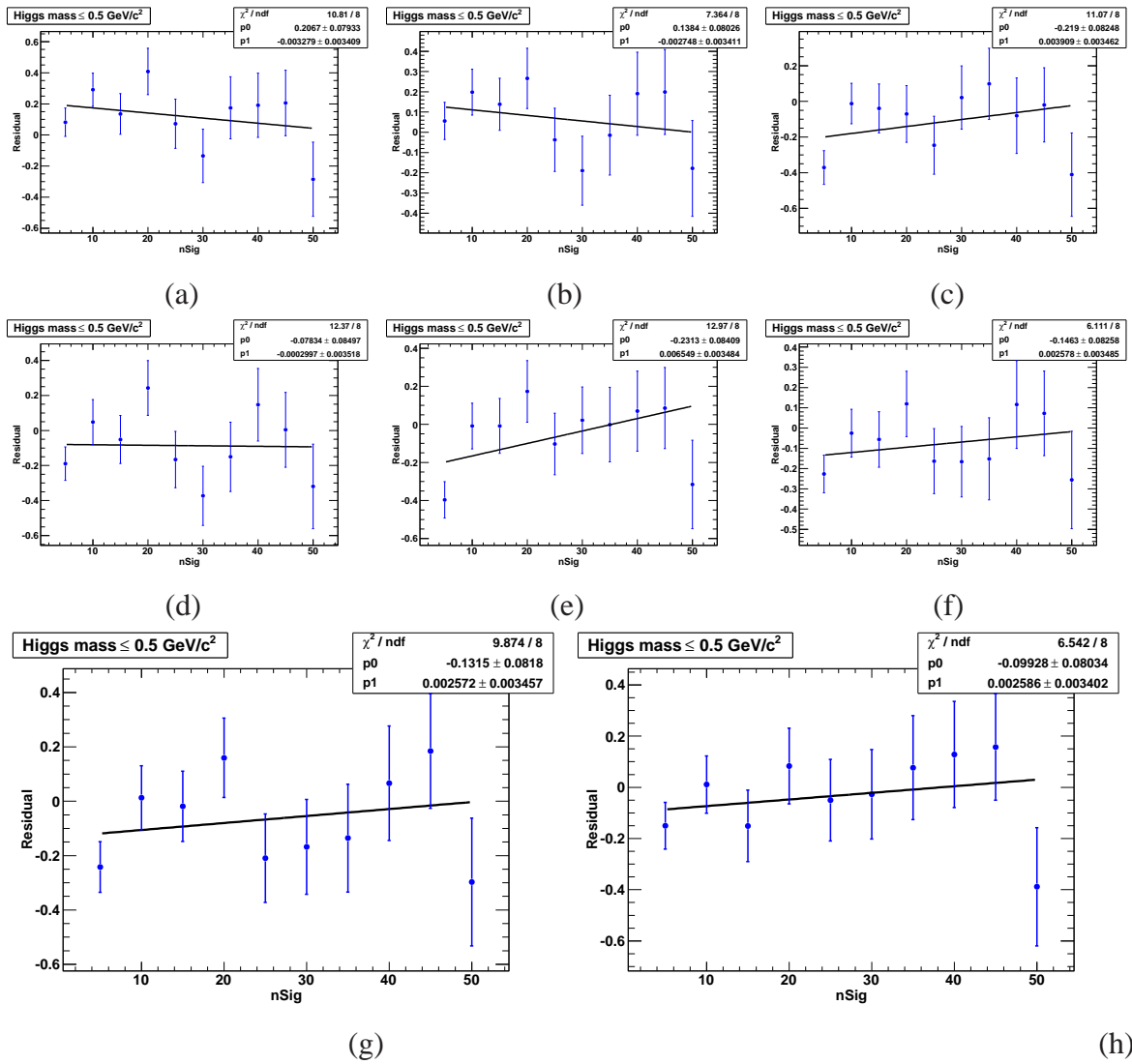


Figure C.4: Fit residuals for the Higgs mass of (a) $m_{A^0} = 0.212 \text{ GeV}/c^2$ (b) $m_{A^0} = 0.214 \text{ GeV}/c^2$ (c) $m_{A^0} = 0.216 \text{ GeV}/c^2$ (d) $m_{A^0} = 0.218 \text{ GeV}/c^2$ (e) $m_{A^0} = 0.220 \text{ GeV}/c^2$ (f) $m_{A^0} = 0.250 \text{ GeV}/c^2$ (g) $m_{A^0} = 0.300 \text{ GeV}/c^2$ and (h) $m_{A^0} = 0.50 \text{ GeV}/c^2$.

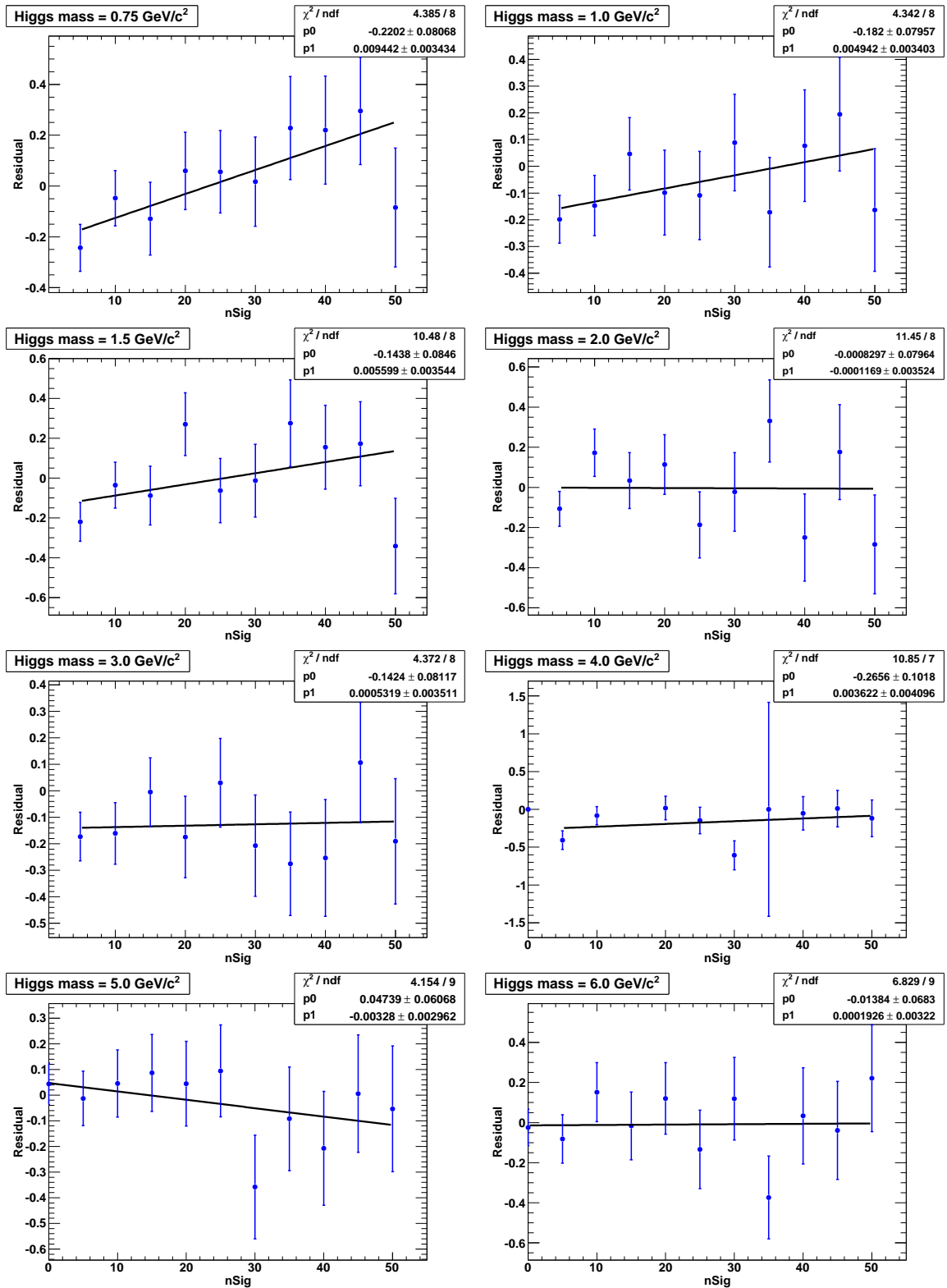


Figure C.5: Fit residuals for the number of signal events in the toy Monte Carlo experiments generated for each Higgs mass points.

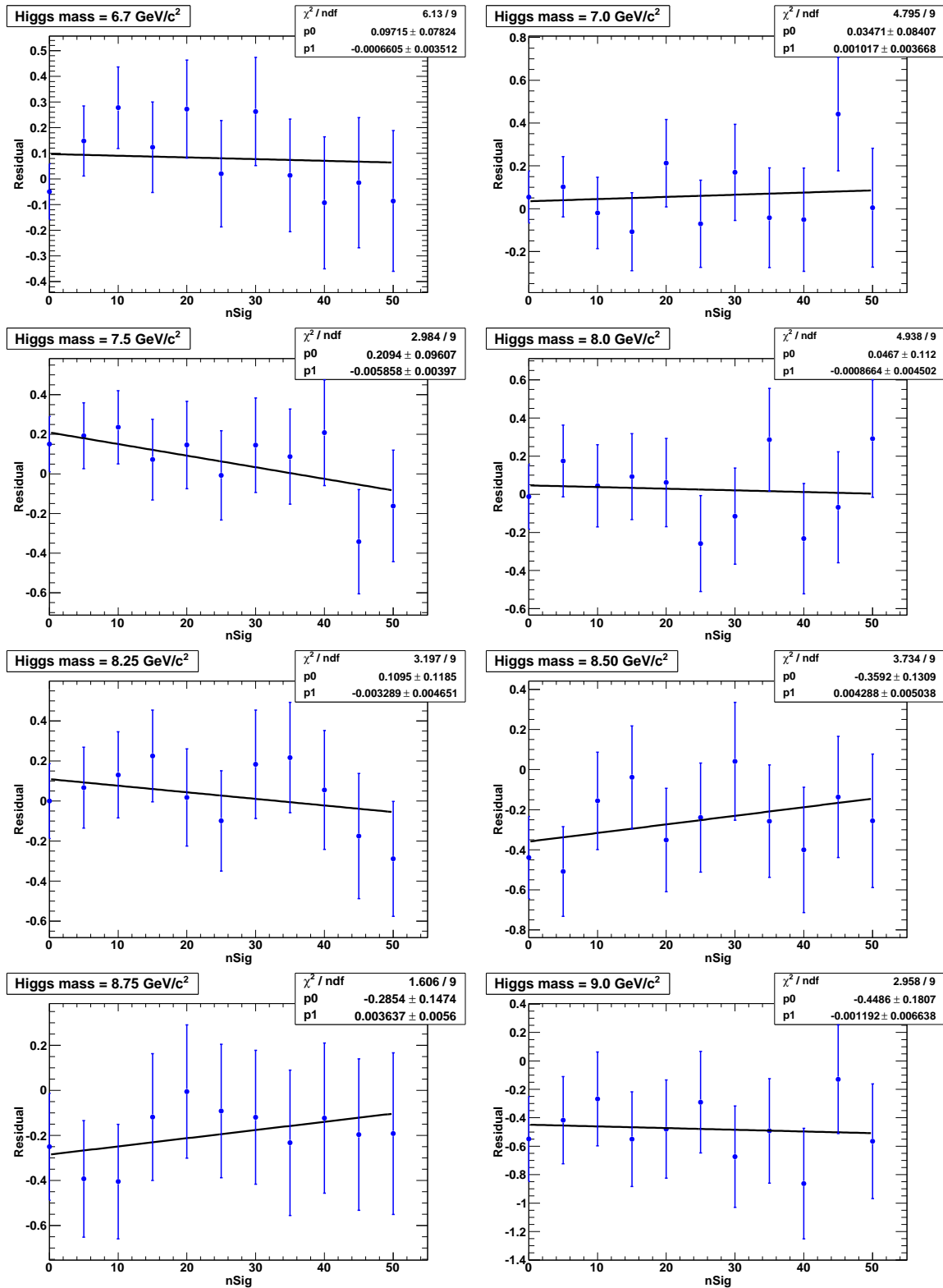


Figure C.6: Fit residuals for the number of signal events in the toy Monte Carlo experiments generated for each Higgs mass points.

Bibliography

- [1] S. Weinberg, *Phys. Rev. Lett.* **19**, 1264 (1967); A. Salam, in *Elementary Particle Theory*, edited by N. Svartholm (Almquist and Wiksells, Stockholm, 1969), p. 367; S. L. Glashow, *Nucl. Phys.* **22**, 579 (1961); S. L. Glashow, J. Iliopoulos, and L. Maiani, *Phys. Rev. D* **2**, 1285 (1970).
- [2] G. 't Hooft and M. Veltman, *Nucl. Phys.* **B44**, 189 (1972).
- [3] M. J. Herrero, arXiv:hep-ph/9812242 (1998).
- [4] ‘Quarks and Leptons’: An introductory course in Modern Particle Physics, F. Halzen and A. D. Martin, John Wiley & Sons, UK (1984); *Introduction to Elementary Particle Physics*, A. Bettini, Cambridge University Press, UK (2008); *Introduction to Elementary Particle*, D. Griffiths, John Wiley & Sons, InC (2004).
- [5] S. Dawson, arXiv:hep-ph/9901280 (1999).
- [6] P. W. Higgs, *Phys. Lett.* **12**, 132 (1964); P. W. Higgs, *Phys. Rev. Lett.* **13**, 508 (1964); P. W. Higgs, *Phys. Rev.* **145**, 1156 (1966).
- [7] F. Englert and R. Brout, *Phys. Rev. Lett.* **13**, 321 (1964).
- [8] G. S. Guralnik, C. R. Hagen and T. W. B. Kibble, *Phys. Rev. Lett.* **13**, 585 (1964).
- [9] T. W. B. Kibble, *Phys. Rev.* **155**, 1554 (1967).
- [10] G. Aad, *et al.* [Atlas Collaboration], *Phys. Lett.* **B716**, 1–29 (2012); S. Chatrchyan, *et al.* [CMS Collaboration], *Phys. Lett.* **B716**, 30–61 (2012).
- [11] R. P. Feynman, *Phys. Rev.* **76**, 769–789 (1949).

- [12] ‘QCD and collider physics’, R.K. Ellis, W. J. Stirling and B. R. Webber, Cambridge University Press (1996); ‘Foundation of Quantum Chromodynamic’, T. Muta, World Scientific, Singapur (1987); ‘The Theory of Quarks and Gluon Interaction’, F. J. Yndurain, Springer-Verlag 2nd. Edition (1993).
- [13] S.L. Glashow, Nucl. Phys. **22**, 579 (1961).
- [14] L. Pape and D. Treille, Rep. Prog. Phys. **69**, 2843–3067 (2006).
- [15] S.P. Martin, arXiv:hep-ph/9709356 (2011).
- [16] H. Georgi and S.L. Glashow, Phy. Rev. Lett. **32**, 438 (1974).
- [17] J. C. Pati and A. Salam, Phys. Rev. **D 10**, 275 (1974).
- [18] H. Georgi, H.R. Quinn, and S. Weinberg, Phys. Rev. Lett. **33** 451 (1974).
- [19] V. Trimble, Ann. Rev.of Astro and Astrophys. **25**, 425–472 (1987).
- [20] L. Bergstrom, Rept. Prog. Phys. **63**, 793 (2000).
- [21] R. W. Schnee, AIPConf. Proc. **903**, 8–15 (2007; arXiv:astro-ph/0612565 (2007)).
- [22] D. I. Kazakov, arXiv:hep-ph/0012288 (2001).
- [23] I. J. R. Aitchison, arXiv:0505105 (2005); H. E. Haber and G. L. Kane, Phys. Rep. **117**, 75 (1985).
- [24] U. Ellwanger, C. Hugonie and A.M. Teixeira, Phys. Reports **496** 1-77 (2010); J. Ellis, J. F. Gunion, H.E. Haber, L. Roszkowski and F. Zwirner, Phys. Rev. **D 39** 844 (1989).
- [25] R. Dermisek and J.F. Gunion, Phys. Rev. Lett. **95**, 041801 (2005);
- [26] G. Hiller, Phys. Rev. **D 70**, 034018 (2004)
- [27] J. E. Kim and H. P. Niles, Phys. Lett. **B 138** (1984) 150.
- [28] R. D. Peccei and H. R. Quinn, Phys. Rev. Lett. **38**, 1440 (1977).
- [29] B. A. Dobrescu, G. Landsberg and K. T. Matchev, Phys. Rev. **D 63**, 075003 (2001); B. A. Dobrescu and K. T. Matchev, J. High Energy Phys. **09**, 031 (2000).

- [30] E. Fullana and M.A. Sanchis-Lozano, Phys. Lett. **B 653**, 67 (2007).
- [31] F. Domingo, U. Ellwanger, E. Fullana, C. Hugonie and M.A. Sanchis-Lazano, JHEP **0901**, 061 (2009).
- [32] S. Schael et al. [ALEPH collaboration], Eur. Phys. J. **C 47**, 547 (2006).
- [33] R. Dermisek and J.F. Gunion, Phys. Rev. **D 73**, 111701 (2006).
- [34] R. Dermisek and J.F. Gunion, Phys. Rev. **D 77**, 015013 (2008).
- [35] R. Dermisek and J.F. Gunion, Phys. Rev. **D 75**, 075019 (2007).
- [36] S. Moretti, S. Munir and P. Poulose, Phys. Lett. **B 644**, 241 (2007); S. Chang, P. J. Fox and N. Weiner, Phys. Rev. Lett. **98** 111802 (2007); T. Stelzer, S. Wiesenfeldt and S. Willenbrock, Phys. Rev. **D 75**, 077701 (2007).
- [37] R. Dermisek and J.F. Gunion, Phys. Rev. **D 76**, 095006 (2007).
- [38] M. Drees and K. Hikasa, Phys. Rev. **D 41**, 1547 (2007).
- [39] M. A. Sanchis-Lazano, Mod. Phys. Lett. **A 17**, 2265 (2002); M. A. Sanchis-Lazano, Mod. Phys. Lett. **A 19**, 2183 (2002); M. A. Sanchis-Lazano, arXiv:hep-ph/0510374 (2006).
- [40] B. McElrath, Phys. Rev. **D 72** 103508 (2005).
- [41] R. Dermisek, J. F. Gunion and B. McElrath, Phys. Rev. **D 76**, 051105 (2007).
- [42] R. N. Hodgkinson, Phys. Lett. **B 665**, 219 (2008).
- [43] Y. Nomura and J. Thaler, Phys. Rev. **D 79** 075008 (2009).
- [44] B. Batell, M. Pospelov and A. Ritz, Phys. Rev. **D 79**, 115008 (2009).
- [45] J. P. Lees *et al.* [BABAR Collaboration], Phys. Rev. Lett. **108**, 211801 (2012).
- [46] F. Wilzek, Phys. Rev. Lett. **39**, 1304 (1977).
- [47] M. L. Mangano and P. Nason, Mod. Phys. Lett. **A 22**, 1373 (2007).
- [48] P. Nason, Phys. Lett. **B 175**, 223 (1986).

- [49] H. Kim, Y. Kolomensky and E. Petigura, et. al., *BABAR* Analysis Document #2107.
- [50] R. Barbieri *et al.*, Phys. Lett. **B 57**, 455 (1975).
- [51] P. del Amo Sanchez *et al.* [*BABAR* Collaboration], Phys. Rev. Lett. **104**, 191801 (2011).
- [52] J. F. Gunion, arXiv:0808.2509 (2008).
- [53] R. Dermisek and J.F. Gunion, Phys. Rev. **D 81**, 075003 (2010).
- [54] B. Aubert *et al.* [*BABAR* Collaboration], Phys. Rev. Lett. **103**, 081803 (2009).
- [55] B. Aubert *et al.* [*BABAR* Collaboration], Phys. Rev. Lett. **103**, 181801 (2009).
- [56] B. Aubert *et al.* [*BABAR* Collaboration], arXiv:0808.0017 (2008).
- [57] P. del Amo Sanchez *et al.* [*BABAR* Collaboration], Phys. Rev. Lett. **107**, 021804 (2011).
- [58] J. P. Lees *et al.* [*BABAR* Collaboration], Phys. Rev. Lett. **107**, 221803 (2011).
- [59] W. Love *et al.* [*CLEO* Collaboration], Phys. Rev. Lett. **101**, 151802 (2008).
- [60] M. Ablikim *et al.* [*BESIII* collaboration], arXiv:1111.2112 (2011).
- [61] S. Chatrchyan *et al.* [*CMS* collaboration], Phys. Rev. Lett. **109**, 121801 (2012).
- [62] F. Domingo, JHEP **1104**, 016 (2011).
- [63] N. Cabibbo, Phys. Rev. Lett. **10**, 531 (1963); M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 652 (1973).
- [64] W. Kozanecki, Nucl. Instrum. Meth., **A 446**, 59–64 (2000).
- [65] B. Aubert *et al.* [*BABAR* Collaboration], Nucl. Instrum. Meth., **A 479**, 1–116, (2002).
- [66] V. Re *et al.* [*BABAR* Collaboration], Nucl. Instrum. Meth., **A 511**, 1–5, (2003).
- [67] G. Benelli et al. [*BABAR* Collaboration], IEEE, **2**, 1145–1148 (2005); W. Menges et al. [*BABAR* Collaboration], IEEE, **3**, 1470–1474 (2005).
- [68] M. R. Convery *et al.* [*BABAR* Collaboration], Nucl. Instrum. Meth., **A 556**, 134–139, (2006).

- [69] R. Bartoldus et. al., *BABAR* Analysis Document #194.
- [70] P. F. Harison, *J. Phys.* **G 28**, 2679 (2002); A. Roodman, arXiv:physics/0312102 (2002).
- [71] D. J. Lange, *Nucl. Instrum. Meth.* **A 462**, 152 (2001).
- [72] S. Jadach, W. Placzek, B. F. L. Ward, *Phys. Lett.* **B 390** (1997) 298.
- [73] B. F. L. Ward, S. Jadach and Z. Was, *Nucl. Phys. Proc. Suppl.* **116**, 73 (2003).
- [74] D. Cronin-Hennessy *et al.* [CLEO Collaboration], *Phys. Rev.* **D 76** 072001 (2007).
- [75] S. Agostinelli *et al.* [GEANT4 Collaboration], *Nucl. Instrum. Methods Phys. Res., Sect. A* **506** 250 (2003).
- [76] <http://root.cern.ch/drupal/>.
- [77] A. Drescher et al. [ARGUS collaboration], *Nucl. Instrum. Meth* **237**, 464 (1985).
- [78] R. Sinkus and T. Voss, *Nucl. Instrum. Meth.* **A 391**, 360 (1997).
- [79] I. Narsky, arxiv:physics/0507143v1 (2005)
- [80] J. Friedman and N. Fisher, *Statistics and Computing* **9**, 123-143 (1999).
- [81] G. Punzi, arxiv:physics/0308063v2 (2003).
- [82] <http://www.stat.berkeley.edu/~breiman/RandomForests/>.
- [83] V. Prasad, B. Bhuyan and P. Poulou, *BABAR* Analysis Document #2304; J.P. Lees [BABAR collaboration], *Phys. Rev. D* **87**, 031102 (R) (2013), [arXiv:1210.0287].
- [84] M. J. Oreglia, Ph.D Thesis, report SLAC-236 (1980); J.E. Gaiser, Ph.D Thesis, report SLAC-255 (1982); T. Skwarnicki, Ph.D Thesis, report DESY F31-86-02 (1986).
- [85] R. Barlow, *Nucl. Instrum. Methods Phys. Res., Sect. A* **297** 496–506 (1990).
- [86] L. Lyons, W. W. M. Allison and P. Comellas, *Nucl. Instrum. Methods Phys. Res., Sect. A* **297** 496–506 (1990).
- [87] <http://roofit.sourceforge.net/>
- [88] <http://rarfit.sourceforge.net/>

-
- [89] B. Aubert *et al.* [BABAR Collaboration], Phys. Rev. Lett. **88**, 241801 (2002).
- [90] <http://www.slac.stanford.edu/BFROOT/www/Physics/Tools/Pid/PidOnMc/nutshell.html#weight>
- [91] T. Allmendinger *et al.*, arXiv:1207.2849.
- [92] J. Beringer *et al.* [Particle Data Group], Phys. Rev. **D 86**, 010001 (2012).
- [93] Grant MCGregor and Chris Hearty, *BABAR* Analysis Document #2069.
- [94] M. Beneke, A. Signer and V.A. Smirnov, Phys. Rev. Lett. **80**, 2535 (1998).