SEARCH FOR $B_S \to \eta'\eta$ IN BELLE DATA

by

Anthony Zummo

Submitted to the Graduate Faculty of

the Kenneth P. Dietrich School of Arts and Sciences in partial

fulfillment

of the requirements for the degree of

Bachelor of Philosophy

University of Pittsburgh

2017

UNIVERSITY OF PITTSBURGH

DEPARTMENT OF PHYSICS AND ASTRONOMY

This thesis was presented

by

Anthony Zummo

It was defended on

April 14th 2017

and approved by

Dr. Vladimir Savinov, Department of Physics and Astronomy

Dr. Russell Clark, Department of Physics and Astronomy

Dr. Tae Min Hong, Department of Physics and Astronomy

Dr. Roy Briere, Department of Physics, Carnegie Mellon University

Thesis Advisor: Dr. Vladimir Savinov, Department of Physics and Astronomy

SEARCH FOR $B_S \to \eta'\eta$ IN BELLE DATA

Anthony Zummo, BPhil

University of Pittsburgh, 2017

We search for the decay $B_s \to \eta'\eta$ using 121.4 fb^{-1} of data collected at the $\Upsilon(5S)$ resonance with the Belle detector at the KEKB asymmetric-energy electron-positron collider. This decay is suppressed in the Standard Model of particle physics and proceeds through transitions sensitive to new physics. The expected branching fraction for $B_s \to \eta'\eta$ in the Standard Model is 33.5×10^{-6}. This decay has not been observed yet. We use Monte Carlo simulation to optimize our selection criteria for signal events and a Neural Network to separate signal from background. Our study maximizes discovery potential for this process.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1.0 INTRODUCTION

1.1 THEORETICAL BACKGROUND

1.1.1 The Standard Model of Particle Physics

The Standard Model of particle physics was formulated in the 1960's and 1970's to classify all known subatomic particles and describe their interactions. The Standard Model classifies fundamental particles into two groups based on the laws of statistics these particles obey. Particles with integer spin obey Bose-Einstein statistics and are referred to as "bosons", while particles with half-integer (*e.g.* 1/2, 3/2, 5/2, ...) spin obey Fermi-Dirac statistics and are referred to as "fermions". The primary difference between the Bose-Einstein statistics that describe bosons and the Fermi-Dirac statistics that describe fermions is that identical fermions cannot occupy the same quantum state simultaneously while bosons can. This phenomenon is called the "Pauli exclusion principle". Included in the Standard Model's fundamental particles are five bosons that mediate the different fundamental interactions between particles. These force-carrying particles are the photon which mediates electromagnetic interactions, the gluon which mediates strong interactions, the charged W and neutral Z bosons which mediate weak interactions, and the Higgs boson. In addition, the Standard Model includes three "generations" of fermions, further divided into two categories of quarks and leptons. There are six "flavors" of quarks within the Standard Model three of which have electric charge $q = +2/3$ (up, charm, and top quarks) and three of which have electric charge $q = -1/3$ (down, strange, and bottom quarks). These particles interact through the strong, weak, and electromagnetic forces and bind together in different combinations to make up composite particles. For example, protons and neutrons are baryons – particles

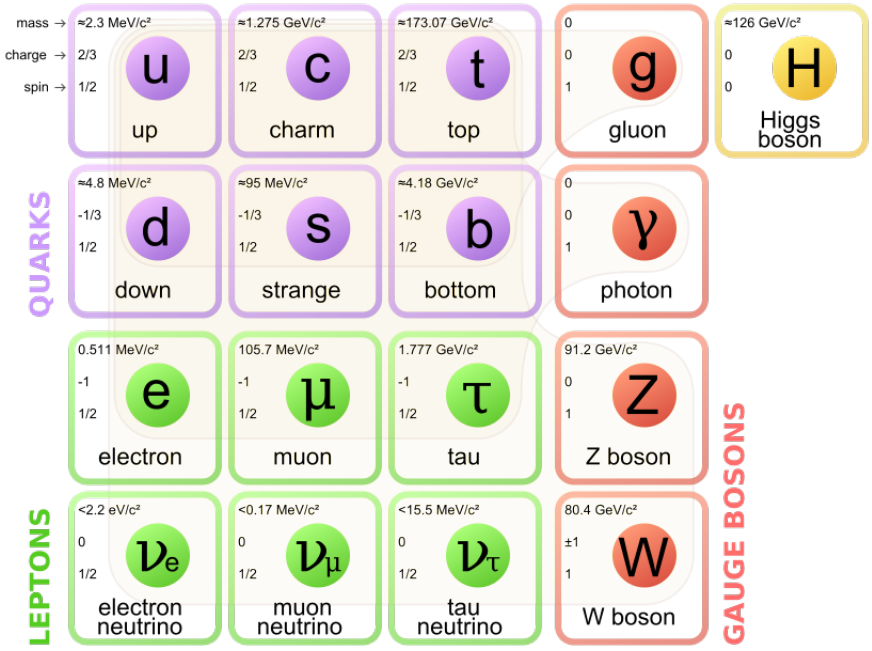

Figure 1: Fundamental particles of the Standard Model [1]

composed of three quarks – and pions are mesons – particles composed of a quark-antiquark pair. There are also six flavors of leptons within the Standard Model. Charged leptons (electrons, muons, and taus) have an electric charge of $q = -1$ and interact through the weak and electromagnetic forces. Neutrinos (ν_e, ν_μ, and ν_τ) have no electric charge and therefore only interact through the weak force. All fundamental particles also have corresponding antiparticles [1] which have the same properties of the particles except with the opposite quantum numbers (such as their electric and color charges). Antiparticles are denoted either through the explicit conjugation of charge (*e.g.* the antiparticle of e^- is e^+) or through the use of an overbar for uncharged particles (*e.g.* the antiparticle of ν_e is $\bar{\nu_e}$). The categorization of the fundamental particles in the Standard Model as well as their mass, charge and spin are summarized in Fig. 1.

The Standard Model not only classifies the fundamental particles of the universe, but also describes the interactions between these particles. The Standard Model has been extremely successful in explaining experimental observations and making reliable predictions. However,

[1] In some cases (γ, Z^0), particles are their own antiparticles.

there are several important experimental facts that the Standard Model does not address. For example, the universe contains much more matter than antimatter, neutrinos have masses, and astrophysical data suggest there is a large amount of dark matter in the universe. None of these observations can be explained by the Standard Model. In addition the Standard Model does not include any description of gravity, one of the four fundamental forces. Because of this, searching for "New Physics" (NP) or physics "Beyond the Standard Model" (BSM) is one of the main goals of particle physics. Searches for BSM physics are conducted in several different areas, including the "cosmic frontier"- which includes experiments that search for new physics using cosmological data, the "energy frontier"- which includes experiments that search for new physics at the highest possible energies, and the "intensity frontier"- which includes experiments that search for indirect signs of new physics by precisely measuring and comparing rare processes to Standard Model predictions. The Belle experiment is one such experiment searching for BSM physics at the intensity frontier. The Belle experiment was designed to allow the properties of particles called B and B_s mesons to be studied with unprecedented precision.

1.1.2 B mesons

B mesons are particles consisting of a b antiquark or quark bound to any lighter quark or antiquark (*i.e.* u, d, c, or s) via the strong force. Because the b quark or antiquark can be bound to several different flavors of quarks, there are several different types of B mesons: B^0, $\bar{B}^0$, B^+, B^-, B_s, $\bar{B}_s$, B_c^-, and B_c^+. Similar to most subatomic particles B mesons are unstable and on average will decay into "daughter particles" after approximately 10^{-12} seconds. B mesons have many interesting properties which has led several experiments to study these particles in great depth. One example of these properties is that the neutral B meson states (B^0 and $\bar{B}^0$, and B_s and $\bar{B}_s$) are able to oscillate between themselves and their corresponding antiparticle. This process can occur through the exchange of two W bosons and is shown for B_s–$\bar{B}_s$ oscillations in Fig. 2.

Another interesting property of B mesons is that Charge Parity (CP) symmetry is violated in their decays. In fact, the primary goal of the Belle experiment was the measurement

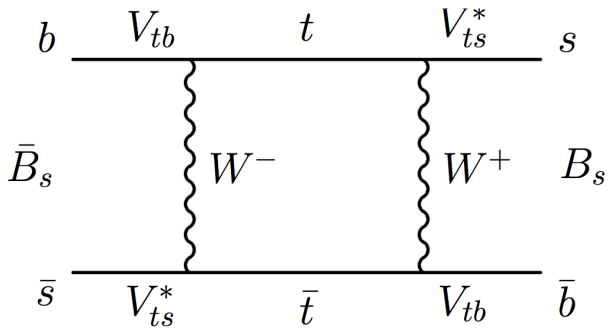

Figure 2: Feynman diagram of $B_s\bar{B}_s$ mixing

of CP violations in B meson decays. Charge conjugation is an operation that transforms all particles into their antiparticles given by:

$$\hat{C}\,|p\rangle = |\bar{p}\rangle \tag{1.1}$$

where $\hat{C}$ is the charge conjugation operator and $|p\rangle$ is an arbitrary particle. Under charge conjugation symmetry, this operation would leave physics unchanged. However, this symmetry is broken in nature. In addition, a parity transformation is an operation which flips the sign of all of the spatial coordinates given by:

$$\hat{P}\,|\vec{x}\rangle = \pm\,|\vec{x}\rangle \tag{1.2}$$

Under parity transformation symmetry this operation would leave physics unchanged, however, this symmetry is also broken in nature. A charge parity transformation is simply the combined transformations of charge and parity and this potential symmetry is also found to be broken in nature. These broken symmetries give insight into the differences between matter and antimatter, and can also give insight into the matter antimatter asymmetry in the universe. For this reason, discovering sources of CP violation is one of the most important goals in experimental particle physics.

1.1.3 Rare and Penguin Decays

One type of process that can be studied at intensity frontier experiments are rare decays of B and B_s mesons. While there is no definition for what is considered "rare", there are many decays that only occur on the order of once per million B mesons. The probability of these decays occuring, called the "branching fraction", is thus $\mathcal{B} = 10^{-6}$. This probability can be calculated theoretically for many decays using the framework of the Standard Model. By measuring these branching fractions experimentally, we are able to test the Standard Model predictions.

One type of decay that is of particular interest are the so-called "penguin" decays. These decays occur through transitions in which a particle is emitted and reabsorbed by the same particle. An example of this type of transition and its resemblance to a penguin is shown in Fig. 3. The Feynman diagrams of penguin decays thus have a characteristic "loop". This is opposed to simpler diagrams called "tree" diagrams. These penguin decays are suppressed in the Standard Model and occur through the transition of a b quark to a d, or s quark ($b \to d, s$). These transitions are interesting because the respective "tree" and "penguin" amplitudes of these decays could be of similar magnitude, so the interference effects between these amplitudes could be sizeable even if one of the amplitudes is relatively small. In addition, new particles not described by the Standard Model can enter the loop and contribute new interesting effects. The presence of new particles could thus manifest itself in several different ways, such as a deviation of the branching fraction from the Standard Model prediction, CP violating asymmetries, or allowing decays forbidden in the Standard Model. By searching for new rare decays of B_s mesons that are sensitive to NP, we are able to either provide evidence for NP or constrain existing models.

1.2 THE BELLE EXPERIMENT

The Belle experiment collected data from 1999-2010 and was conducted by an international collaboration of over 400 physicists and was located at the High Energy Accelerator Re-

Figure 3: Feynman diagram of a penguin transition

search Organization (KEK) in Tsukuba, Japan. This experiment operated by accelerating electrons and positrons to have energies of 8 GeV and 3.5 GeV, respectively, before colliding them at the interaction point within the Belle detector. These high energy collisions can result in electron-positron annihilation, creating many unique and interesting particles and resonances. In particular, the Belle experiment was designed to produce many B and B_s mesons, allowing their properties to be studied in great depth. This type of experiment is called a "B-factory" and is acheived by colliding electrons and positrons at the energy of the Upsilon (Υ) resonances which commonly decay to pairs of B mesons. The Belle experiment has collected large data samples of the Υ resonances at the (1S) through (5S) energies [2]. The Belle experiment primarily operated at the $\Upsilon(4S)$ resonance ($\sqrt{s} = 10.58$) and collected a data sample of 703 fb^{-1} corresponding to 772 million $B\bar{B}$ pairs. The Belle experiment also collected a large data sample at the $\Upsilon(5S)$ resonance ($\sqrt{s} = 10.86$). These $\Upsilon(5S)$ resonances commonly decay into a pair of quantum mechanically entangled B_s mesons. 121.4 fb^{-1} of data was collected at the $\Upsilon(5S)$ resonance corresponding to a sample of 6.5 million $B_s\bar{B}_s$ pairs.

[2]These are the first through fifth radially excited states of the Υ resonance

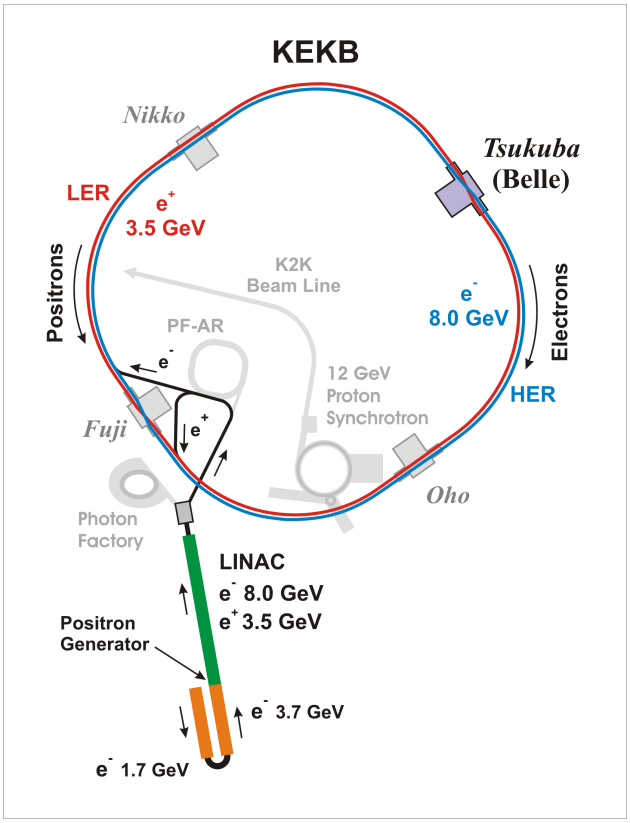

Figure 4: Schematic diagram of the KEKB accelerator [2]

1.2.1 KEKB

KEKB is the asymmetric collider that accelerated and guided the electrons and positrons to the interaction region. It consisted of two storage rings that are 3 km in circumference. The first of these rings is the high energy ring (HER) which contains an electron beam with energy $E = 8.0$ GeV and the second is a low energy ring (LER) which contains a positron beam with energy $E = 3.5$ GeV. The electrons and positrons are accelerated using a linear accelerator before being injected into their respective storage rings. A diagram of the accelerator is shown in Fig. 4. The KEKB accelerator is described in detail in Ref. [2].

1.2.2 The Belle Detector

The Belle detector is a multilayer particle detector located at the interaction point of KEKB. Because the center-of-mass system of e^+e^- collisions is not at rest in the detector's frame

Table 1: Subdetectors of the Belle detector

Subdetectors
Silicon Vertex Detector (SVD)
Central Drift Chamber (CDC)
Time-of-Flight Detector (TOF)
Aerogel Cherenkov Counter (ACC)
Electromagnetic Calorimeter (ECL)
K_L and Muon Detector (KLM)
Extreme Forward Calorimeter (EFC)

of reference, the detector itself was designed to be asymmetric. The detector is permeated by a uniform 1.5 T magnetic field in the direction of the beam axis causing the trajectories of charged particles to be helical. Several subdetectors make up the Belle detector. These subdetectors are listed in order from closest to furthest from the interaction point in Table 1.

These detectors have several functions allowing for the analysis of data from e^+e^- collisions. The subdetectors reconstruct the trajectories of particles as they traverse the detector. They also perform measurements to identify the types of the detected particles. The Belle detector and its subdetectors are described in detail in Ref. [3] and a diagram of the detector is shown in Fig. 5.

1.3 EXPERIMENTAL TECHNIQUES

1.3.1 Particle Reconstruction

Because B_s mesons have an extremely short lifetime, they decay before they are able to be detected directly by the innermost subdetectors of the Belle detector. Therefore, in order to indirectly detect B_s mesons, we must first directly detect the secondary particles

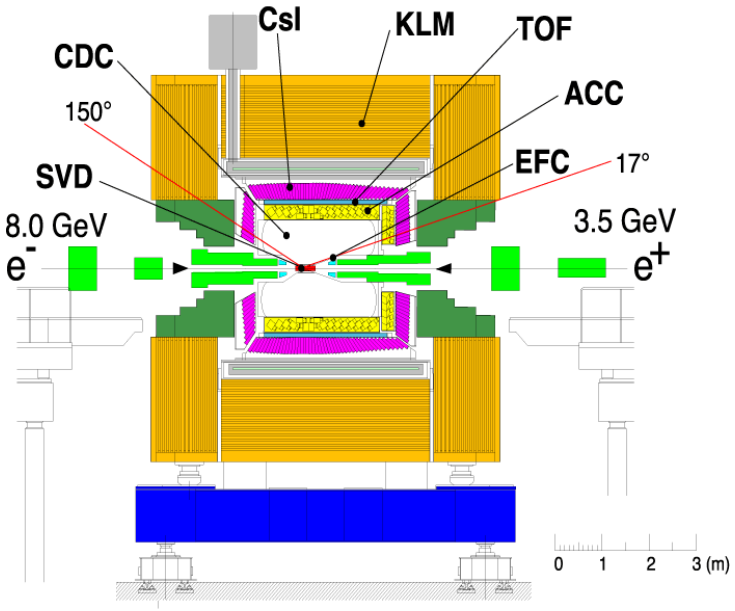

Figure 5: Schematic diagram of the Belle Detector [3]

created in their decays. After these secondary particles are detected, we are able to infer the properties of the candidate B_s meson from its daughters. This process is called particle reconstruction. Because energy and momentum are conserved in particle interactions, we are able to measure the energy and momentum of a B_s meson candidate by simply adding the energy and momenta of its daughter particles. In order to determine the mass of the B_s candidate we can use the relativistic equation (in natural units where $c = 1$)

$$m^2 = E^2 - p^2 \tag{1.3}$$

where E, m, and p are the reconstructed energy, mass, and momentum, respectively, of the B_s candidate. In order to search for B_s mesons decaying to certain daughter particles, we must find all instances of the daughter particles of this decay and reconstruct all possible candidates for the parent particle. Plotting the mass (or other properties) of the B_s candidates shows that candidates that do not come from actual B_s meson decays are randomly distributed, while candidates that do come for B_s meson decays are distributed closer to the actual mass of a B_s meson. We exploit these differences in the distributions to statistically estimate the number of actual B_s mesons in a sample.

1.3.2 Monte Carlo Simulation

In order to develop and validate the analysis procedures used by the Belle collaboration, samples of Monte Carlo (MC) data are generated. These MC data samples are obtained from simulations of the events that occur in the e^+e^- collisions of KEKB and the detector response of Belle. The simulations use pseudorandom number generators to determine which physical processes occur in each e^+e^- collision and the efficiency and accuracy of the measurements performed by the Belle detector. The likelihood of processes being chosen to occur in the simulations are determined by previous measurements and theoretical predictions of the probabilities for these processes. In addition, the probabilities for these processes can be controlled by the user to ensure certain events occur in these simulations. Analyses commonly use two types of simulated MC data samples. In background MC, the probabilities of processes occuring are determined by previous measurements and theoretical predictions. In signal MC, the probabilities of simulated processes are set to ensure that the events being studied are generated. These two simulated data samples are used to quantify the differences between signal and background distributions and to optimize our analysis techniques.

1.4 THE DECAY

Because the final state products of $B_s \to \eta'\eta$ are hadrons containing no charm quarks, $B_s \to \eta'\eta$ is a charmless hadronic decay. Charmless hadronic decays of B mesons are important both to understanding CP-violation and flavor mixing in the Standard Model (SM), and to searching for physics beyond the Standard Model (BSM) [4]. $B_s \to \eta'\eta$ proceeds primarily through $b \to u$ transitions, or $b \to s$ transitions. This can be seen in the three representative Feynman diagrams of the decays $B_s \to \eta'\eta$ and $B_s \to \eta'\pi^0$ in Fig. 6. Penguin decays are sensitive to BSM scenarios including a fourth fermion generation, supersymmetry with broken R-parity, and a two-Higgs doublet model with flavor changing neutral currents. These BSM scenarios can affect the branching fraction (and CP asymmetries) of these decays.

The branching fraction of this decay has been predicted using several different methods

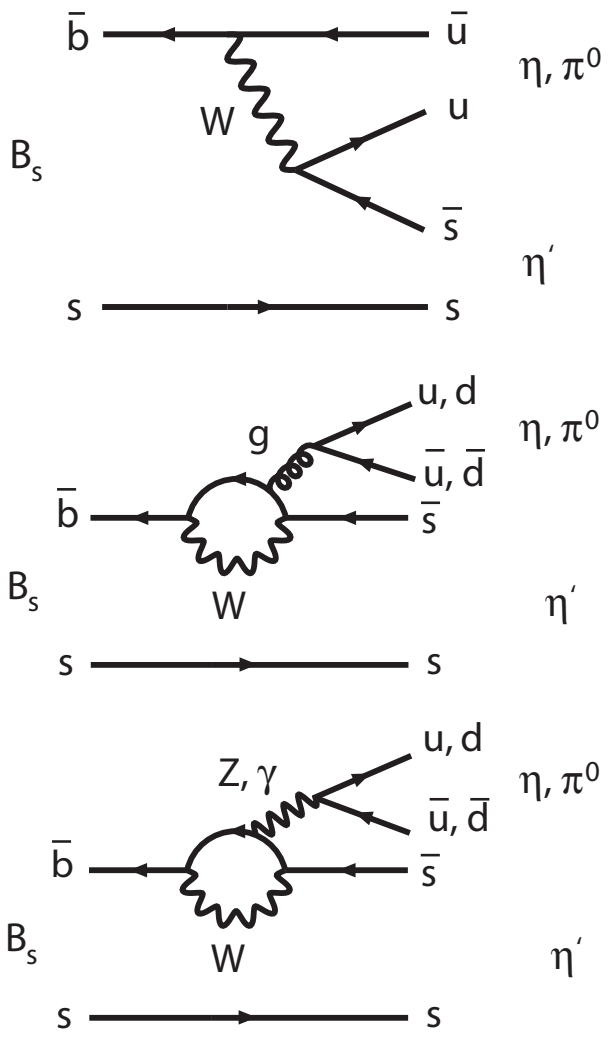

Figure 6: Tree level, gluonic penguin, and radiative penguin Feynman diagrams of $B_s \to \eta'\eta$ and $B_s \to \eta'\pi^0$ decays.

including soft collinear effective theory (SCET) [5], perturbative QCD (pQCD) [6], QCD factorization (QCDF) [7], flavor SU(3) symmetry [8], and flavor U(3) symmetry [9]. These predictions range from $(21.0-41.2)\times 10^{-6}$. This decay has not yet been observed, but could be within our sensitivity.

In addition, data on $B_{d,s} \to \eta\eta, \eta\eta', \eta'\eta'$ branching fractions can be used to extract CP-violating parameters [9]. Of these six decays channels, measurements of the branching fractions of four of them are needed to make any conclusions. Only the branching fraction of $B_s \to \eta'\eta'$ has been measured so far [10].

2.0 ANALYSIS STRATEGY SUMMARY

In this analysis, we search for the decay $B_s \to \eta'\eta$ using the entire Belle data sample of $\Upsilon(5S)$ on-resonance data. We perform a blind analysis by optimizing our analysis techniques using only MC samples of signal and background events before applying these techniques to the real data. We begin by generating a large sample of simulated $B_s \to \eta'\eta$ signal events. Official MC simulations of Belle background processes are used as our background simulations. We use these samples to first identify several variables that exhibit a separation between signal and background distributions. These variables are used either in our selection criteria, background suppression strategy, or final maximum likelihood fit to extract the signal yield in data. We then optimize our selection criteria for $B_s \to \eta'\eta$ signal events. Background events are then further suppressed using a two step process. First, we optimize a Fisher discriminant [11] using the software package RooKSFW [12]. RooKSFW combines several variables that provide separation between signal and background into a single discriminating variable. This variable is then used as one of several input variables to NeuroBayes [13], a neural network, which also calculates a single discriminating variable used in our selection criteria. We also use a transformed version of this variable in our maximum likelihood fit. After further suppressing backgrounds, we use the software package RooFit [14] to model the distributions of our fitting variables for the signal MC sample and the sidebands of the background MC sample. These models are then combined and applied to the full background MC sample to validate our fitting methods and ensure we can use the same method for data. We then perform ensemble tests to check for biases in our fitting functions. Finally, we perform a maximum likelihood fit to the data to extract the signal yield.

After measuring the signal yield, the branching fraction is calculated as:

$$\mathcal{B}(B_s \to \eta'\eta) = \frac{N_{sig}}{2 \cdot N_{B_s^* \bar{B}_s^*} \cdot \epsilon_{B_s^* \bar{B}_s^*} \cdot \mathcal{B}_{daughters}} \tag{2.1}$$

where $N_{B_s^* \bar{B}_s^*}$ is the number of $B_s^* \bar{B}_s^*$ pairs, the factor of 2 is from the two B_s^* mesons in each pair, $\epsilon_{B_s^* \bar{B}_s^*}$ is our reconstruction efficiency, and $\mathcal{B}_{daughters}$ is the product of the relevant branching fractions:

$$\mathcal{B}_{daughters} = \mathcal{B}(\eta' \to \rho^0\gamma) \cdot \mathcal{B}(\rho^0 \to \pi^+\pi^-) \cdot \mathcal{B}(\eta \to \gamma\gamma) \tag{2.2}$$

3.0 MONTE CARLO AND DATA SAMPLES

3.1 SIGNAL MONTE CARLO SAMPLE

We have generated a sample of 100,000 $B_s \to \eta'\eta$ events in proportion to the cross sections for $B_s^*\bar{B}_s^*$, $B_s^*\bar{B}_s$, and $B_s\bar{B}_s$ production at $\Upsilon(5S)$ and distributed over Belle data samples 53, 67, 69, and 71[1] according to data statistics. We have also generated samples of 100,000 $B_s \to \eta'\eta$ events for each of $B_s^*\bar{B}_s^*$, $B_s^*\bar{B}_s$, and $B_s\bar{B}_s$ individually. We use EvtGen [15] to generate events and GEANT4 [16] to simulate the detector response.

3.2 BACKGROUND MONTE CARLO SAMPLE

We use 4 streams of generic $\Upsilon(5S)$ MC to study background events. Each stream corresponds to an integrated luminosity of 121.4 fb^{-1} of $\Upsilon(5S)$ on-resonance data. This is equivalent to the full Belle data sample. The dominant source of background in these decay channels is light-quark continuum events ($u\bar{u}, d\bar{d}, s\bar{s}$, and $c\bar{c}$).

3.3 DATA SAMPLE

In this analysis, we use the entire 121.4 fb^{-1} of $\Upsilon(5S)$ on-resonance data. This dataset corresponds to $(6.53 \pm 0.66) \times 10^6$ $B_s\bar{B}_s$ pairs. It was collected by the Belle detector at the

[1]The Belle data sample consists of 73 experiments, each run under slightly different conditions. These 4 experiments were taken at the Υ(5S) resonance.

KEKB collider from asymmetric energy e^+e^- collisions with $\sqrt{s} = 10.86$ GeV.

4.0 SELECTION CRITERIA

4.1 PHOTON (γ) SELECTION

Photons deposit energy in the ECL by first undergoing a conversion into an electron-positron pair (*i.e.* pair production process) followed by repeated bremsstrahlung and pair production until the entire energy of the photon is absorbed. These events lead to a "shower" of electrons, positrons, and photons in the detectors until the particles have low enough energy to excite the calorimeter material. These energy deposits in the crystals of the ECL are used to reconstruct photons in data. All photons are required to have energy greater than 100 MeV.

4.2 PION ($\pi^{\pm}$) SELECTION

The drift chamber is filled with a gas mixture with so-called field and sensitive wires strung through its volume. Groups of wires are arranged in cells with large electric field gradient. Charged particles traversing the CDC ionize the gas, the ionized electrons are drawn to sensitive wires (located at the center of the cells). When the ionized electrons reach the wires they create a current which can be measured. Digitized response of the detector to this current is generally reffered to as a hit. As the particle continues through the drift chamber, it creates a trail of ionization which is responsible for many hits in the drift chamber cells. These hits can be used to reconstruct the trajectory taken by the particle. Because charged particles in a magnetic field follow helical trajectories, measuring the curvature of the tracks allows the momentum of the particle to be estimated.

We require pion candidates to satisfy $dr < 0.2$ cm and $|dz| < 4$ cm, where dr and dz are defined as the distances from the interaction point to the point of closest approach in the radial and beam axis directions respectively. We also require pion candidates to have transverse momentum $p_T > 100$ MeV. In addition, pion candidates must satisfy $R_{K,\pi} < 0.6$, where $R_{K,\pi}$ is the "Kaon identification variable" defined as the likelihood ratio of the charged track being produced by a kaon vs a pion.

4.3 RHO (ρ^0) SELECTION

We reconstruct ρ^0 meson candidates using the decay channel $\rho^0 \to \pi^+\pi^-$. This decay has a branching fraction of $\mathcal{B}(\rho^0 \to \pi^+\pi^-) = 0.998$. These ρ^0 meson candidates are required to have an invariant mass between 0.435 and 1.005 GeV.

4.4 ETA (η) SELECTION

We reconstruct η meson candidates using the decay channel $\eta \to \gamma\gamma$. This decay has a branching fraction of $\mathcal{B}(\eta \to \gamma\gamma) = 0.394$. These η meson candidates are required to have an invariant mass between 0.47 GeV and 0.57 GeV. After reconstruction, we perform a mass fit to our η candidates, constraining them to their nominal mass of 0.548 GeV.

4.5 ETA PRIME (η') SELECTION

We reconstruct η' meson candidates using the decay channel $\eta' \to \rho^0\gamma$. This decay has a branching fraction of $\mathcal{B}(\eta' \to \rho^0\gamma) = 0.291$. These η' meson candidates are required to have an invariant mass between 0.92 GeV and 0.98 GeV. The distribution of η' candidate invariant masses in both signal and background MC are shown in Fig. 7. All plots with distributions

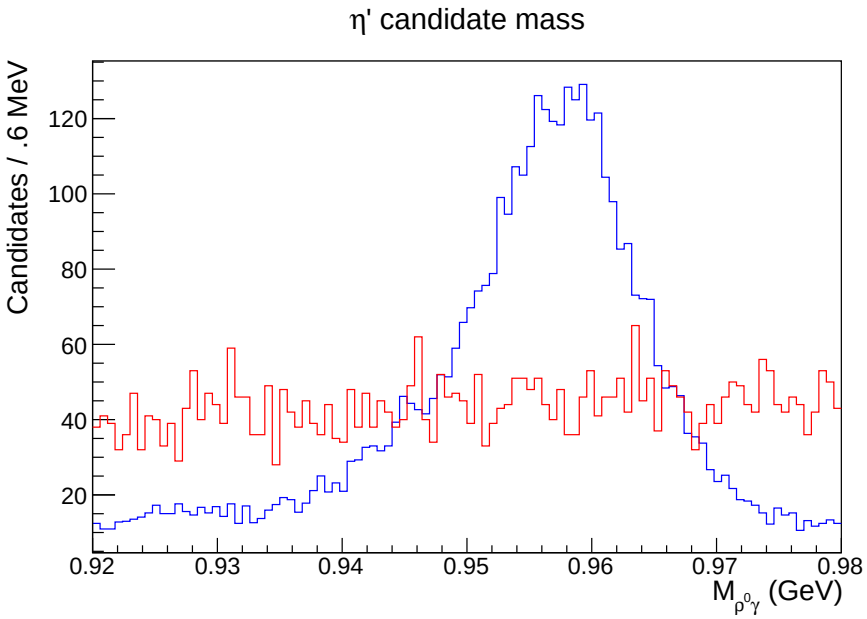

Figure 7: Distributions of $M_{\rho^0\gamma}$ for signal (blue) and background (red) MC

of signal and background on the same axes are shown with the number of events normalized to the number of events in background MC.

4.6 B_S MESON SELECTION

We reconstruct B_s meson candidates using the decay channel $B_s \to \eta'\eta$. We define two variables – the beam-constrained mass (M_{bc}) and the energy difference (ΔE) – to select our B_s meson candidates. The beam constrained mass is the mass of the B_s meson calculated by replacing the energy of the B_s meson with that of the beam:

$$M_{bc} = \sqrt{E_{beam}^2 - p_{B_s}^2} \tag{4.1}$$

The energy difference is the difference between the energy of the beam and the B_s meson:

$$\Delta E = E_{B_s} - E_{beam} \tag{4.2}$$

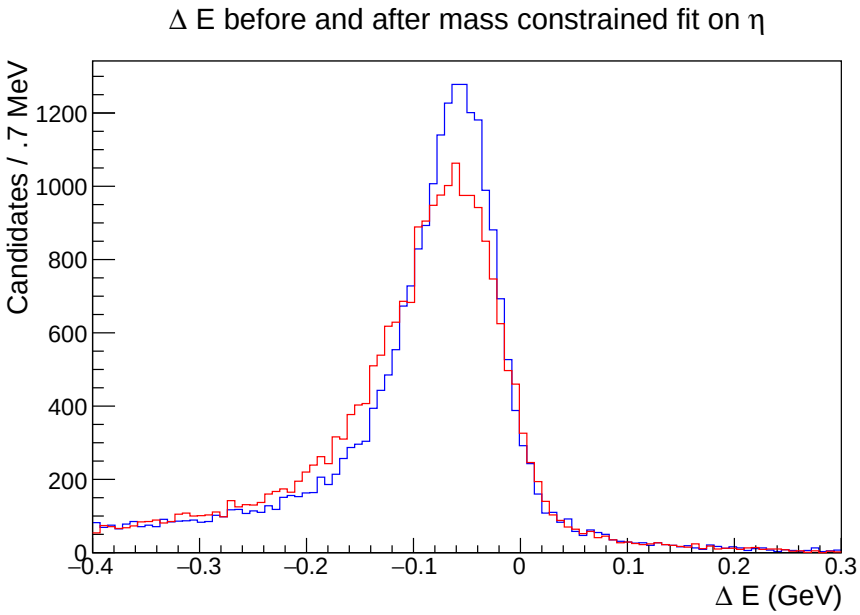

Figure 8: Distribution of ΔE before (blue) and after (red) mass constraint of η

In these definitions, E_{beam} is defined as the center of mass energy of the e^+e^- system divided by two the beam energy, E_{B_s} is the energy of the B_s candidate, and p_{B_s} is the momentum of the B_s candidate. All quantities are calculated in the center of mass rest frame of the $\Upsilon(5S)$. These variables are used due to their improved resolution compared to the mass and energy of the B_s meson candidate. This is because the energy of the beam can be more accurately measured than the energy of the B_s meson candidate. The effect of our mass constraint of η candidates on ΔE is shown in Fig. 8. The distributions of M_{bc} and ΔE for B_s candidates in both signal and background MC are shown in Fig. 10. The selection criteria used for choosing B_s candidates are summarized in Table 2.

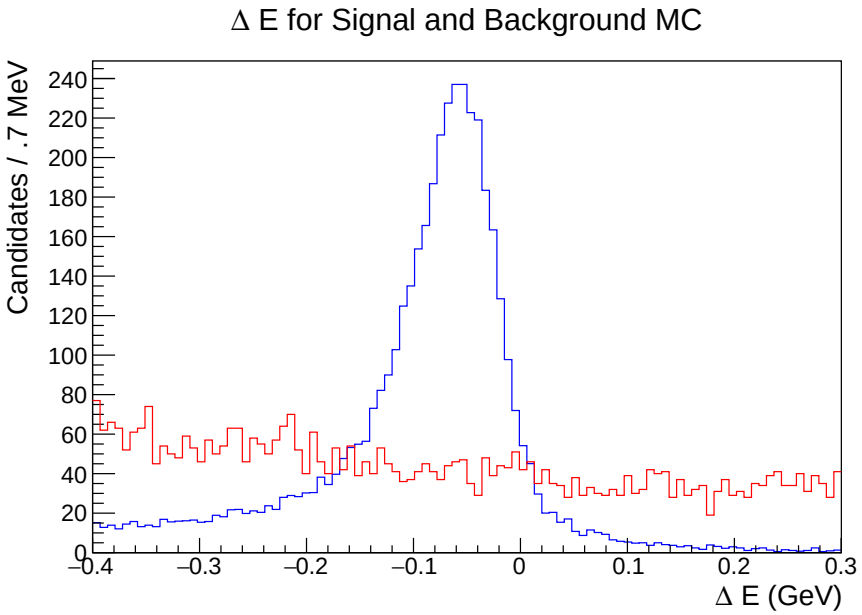

Figure 9: Distribution of ΔE for signal (blue) and background (red) MC

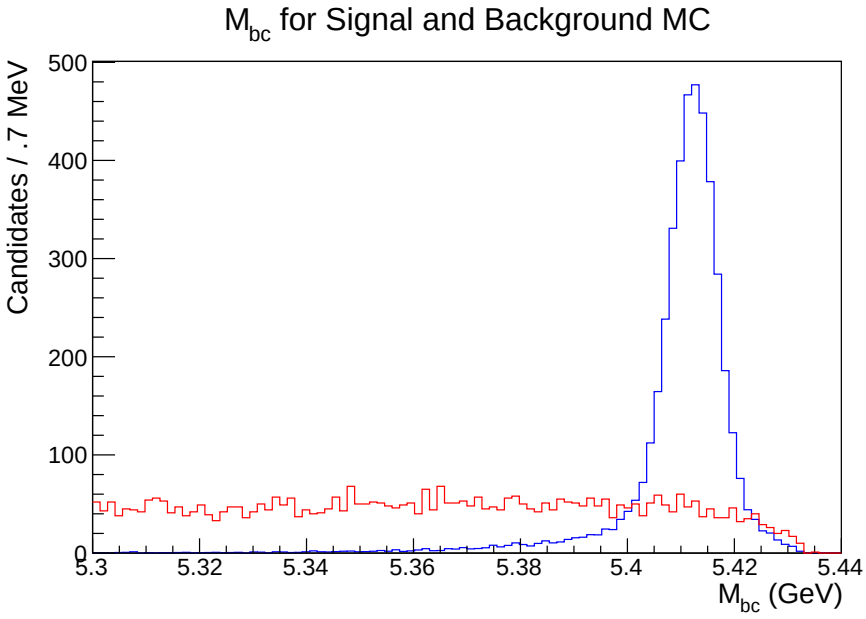

Figure 10: Distribution of M_{bc} for signal (blue) and background (red) MC

Table 2: Selection criteria for $B_s \to \eta'\eta$

Particle	Criteria
γ	$E > 100$ MeV
$\pi^{\pm}$	$dr < 0.2$ cm $\lvert dz \rvert < 4$ cm $p_T > 100$ MeV/c $R_{K,\pi} < 0.6$
ρ^0	$0.435\ \text{GeV}/c^2 < M_{\pi^+\pi^-} < 1.005\ \text{GeV}/c^2$
η	$0.47\ \text{GeV}/c^2 < M_{\gamma\gamma} < 0.57\ \text{GeV}/c^2$
η'	$0.92\ \text{GeV}/c^2 < M_{\rho^0\gamma} < 0.98\ \text{GeV}/c^2$
B_s	$M_{bc} > 5.3\ \text{GeV}/c^2$ -0.4 GeV $< \Delta E <$ 0.3 GeV

5.0 BACKGROUND SUPPRESSION

The primary source of background events for charmless hadronic decays such as $B_s \to \eta'\eta$ is continuum background events from light quark production ($e^+e^- \to q\bar{q}$ where $q = u, d, s, c$). The light quarks in continuum events have large momenta, causing these events to have a "jet-like" shape, while $B_s\bar{B}_s$ events are distributed isotropically. This difference is illustrated in Fig. 11. Because of this, we are able to separate $q\bar{q}$ events from $B_s\bar{B}_s$ events using variables that describe the shape of the event. These variables are described in the next section.

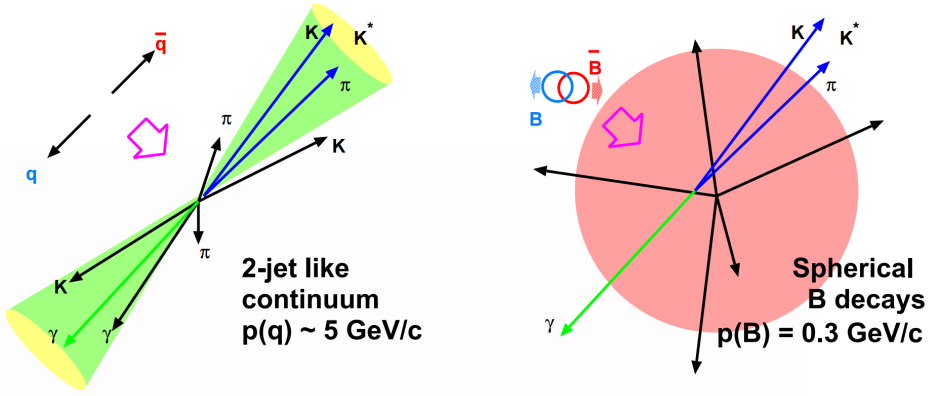

Figure 11: Diagrams showing the topological differences between continuum background (left) and signal B (right) events

5.1 BACKGROUND SUPPRESSION VARIABLES

1. $|\cos(\theta_{thrust})|$: θ_{thrust} is the angle between the thrust axis of the detected particles from the B_s decay and the thrust axis of all other particles detected in the event. The thrust axis is defined as the axis on which the projections of the momenta of the particles has the greatest magnitude. Because of the spherical shape of $B_s\bar{B}_s$ events, these axes will be randomly distributed and $|\cos(\theta_{thrust})|$ should have a uniform distribution, while continuum events will peak near $|\cos(\theta_{thrust})| = 1$.
2. $|\cos(\theta_{bt})|$: θ_{bt} is the angle between the thrust axis of the detected particles from the B_s decay and the beam axis. Because of the spins and large initial momenta of the quarks in continuum events, continuum events are more likely to be aligned with the beam axis.
3. R_2: The l^{th} order Fox-Wolfram moment [17] is defined as

$$H_l = \sum_{i,j} \frac{|\mathbf{p}_i|\,|\mathbf{p}_j|}{E_{\text{vis}}^2} P_l(\cos\theta_{ij}) \ , \tag{5.1}$$

   where θ_{ij} is the opening angle between particles i and j, E_{vis} is the total visible energy of the event, and P_l are the Legendre Polynomials.

   R_2 is defined as the ratio of the second to the zeroeth order Fox-Wolfram moments:

$$R_2 = \frac{H_2}{H_0} \tag{5.2}$$

   This variable describes the shape of the events and has a range from 0 to 1. Events with more spherical shapes have a value of R_2 closer to 0, while events with a "jet-like" shape have a value of R_2 closer to 1. This allows us to separate spherical $B_s\bar{B}_s$ events from jet-like continuum events.
4. Kakuno-Super-Fox-Wolfram (KSFW) moments [12]:

   KSFW moments are modified versions of the Fox-Wolfram moments that utilize thrust axes and the separate tracks of the signal and other B meson.

   - R_l^{so}

The R_l^{so}'s are moments associated with the signal ("s") B_s meson and the other ("o") B_s meson in the event and are defined as:

$$R_l^{so} \equiv \frac{(\alpha_c)_l (H_{\text{charged}})_l^{so} + (\alpha_n)_l (H_{\text{neutral}})_l^{so} + (\alpha_c)_l (H_{\text{missing}})_l^{so}}{E_{\text{beam}} - \Delta E} \tag{5.3}$$

R_l^{so}'s are composed of three components, "charged," "neutral," and "missing energy".

For $l = 1$ and 3,

$$(H_{\text{charged}})_l^{so} \equiv \sum_i \sum_j \beta_l^{so} Q_i Q_j |\vec{p}_j| P_l(\cos\theta_{ij}), \tag{5.4}$$

$$(H_{\text{neutral}})_l^{so} = H_{\text{missing}} = 0 \tag{5.5}$$

For $l = 0$, 2, and 4,

$$(H_{\text{charged,neutral,missing}})_l^{so} \equiv \sum_i \sum_j \beta_l^{so} |\vec{p}_j| P_l(\cos\theta_{ij}) \tag{5.6}$$

The index i iterates over the tracks of the signal candidate and the index j iterates over the same category (charged, neutral, or missing) of tracks of the other B_s meson. Q_i is the charge of particle i. In these equations, α and β are Fisher coefficients. This gives 11 R_l^{so} parameters.

- R_l^{oo}

  The R_l^{oo}'s are moments associated with both particles coming from the other B_s meson in the event.

  For $l = 1, 3$

$$R_l^{oo} \equiv \frac{\sum_j \sum_k \beta_l^{oo} Q_j Q_k |p_j||p_k| P_l(\cos\theta_{jk})}{(E_{\text{beam}} - \Delta E)^2} \tag{5.7}$$

  For $l = 0, 2, 4$

$$R_l^{oo} \equiv \frac{\sum_j \sum_k \beta_l^{oo} |p_j||p_k| P_l(\cos\theta_{jk})}{(E_{\text{beam}} - \Delta E)^2} \tag{5.8}$$

  In these equations, the indices j and k iterate over the tracks not associated with the B_s signal candidate. This gives five R_l^{oo} parameters

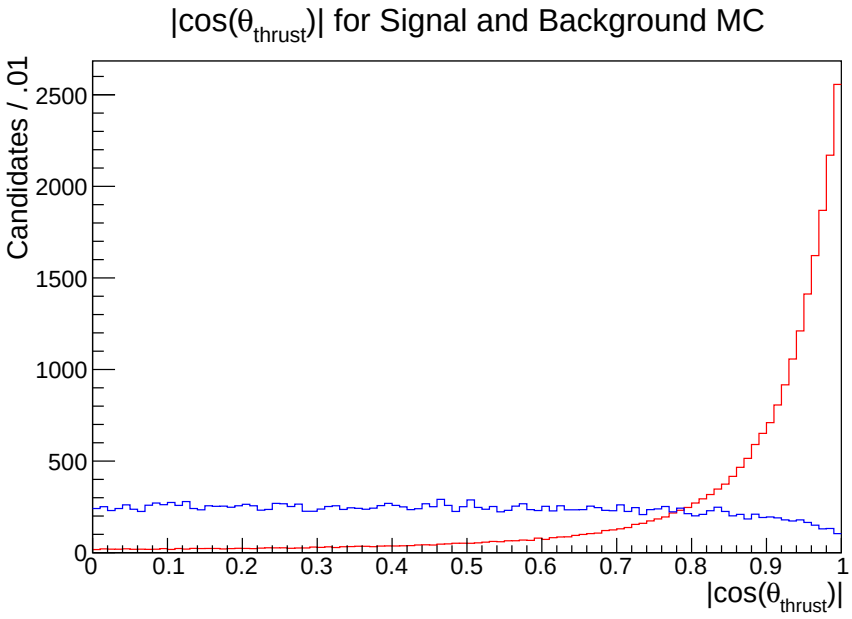

Figure 12: Distribution of $|\cos(\theta_{thrust})|$ for signal (blue) and background (red) MC

- $\sum_{n=1}^{N_t} |p_{t,n}|$

  This is the sum of transverse momenta of all visible particles in an event.

- mm^2

  mm^2 is the missing mass squared defined as:

$$mm^2 \equiv (E_{\Upsilon(5S)} - \sum_i E_i)^2 - \sum_i |p_i|^2 \tag{5.9}$$

This method has a total of 18 parameters.

The distributions of these background suppressing variables are shown in Fig. 12, 13, and 14 with signal distributions in blue and background distributions in red. Again, all plots with distributions of signal and background on the same axes are shown with the number of events normalized to the number of events in background MC. All distributions exhibit some separation between signal and background events.

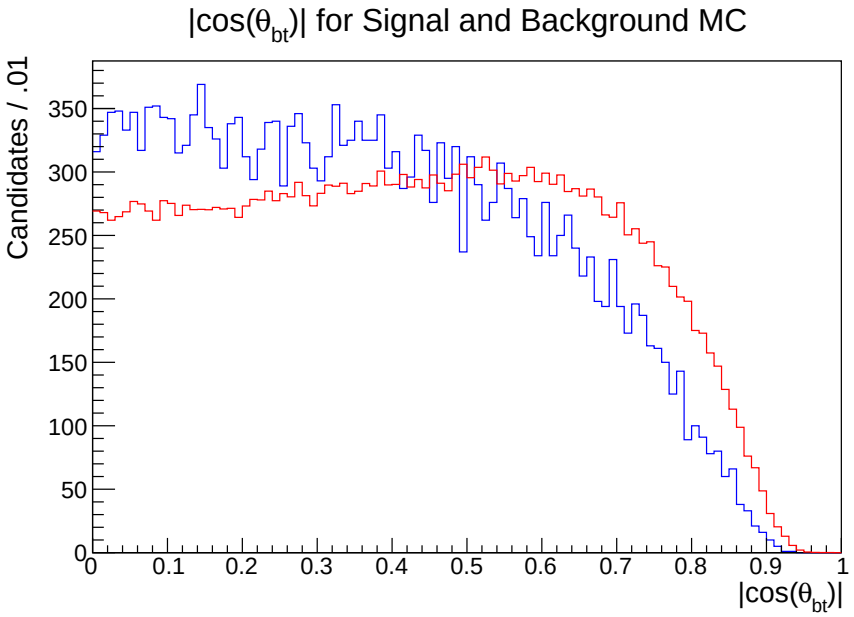

Figure 13: Distribution of $|\cos(\theta_{bt})|$ for signal (blue) and background (red) MC

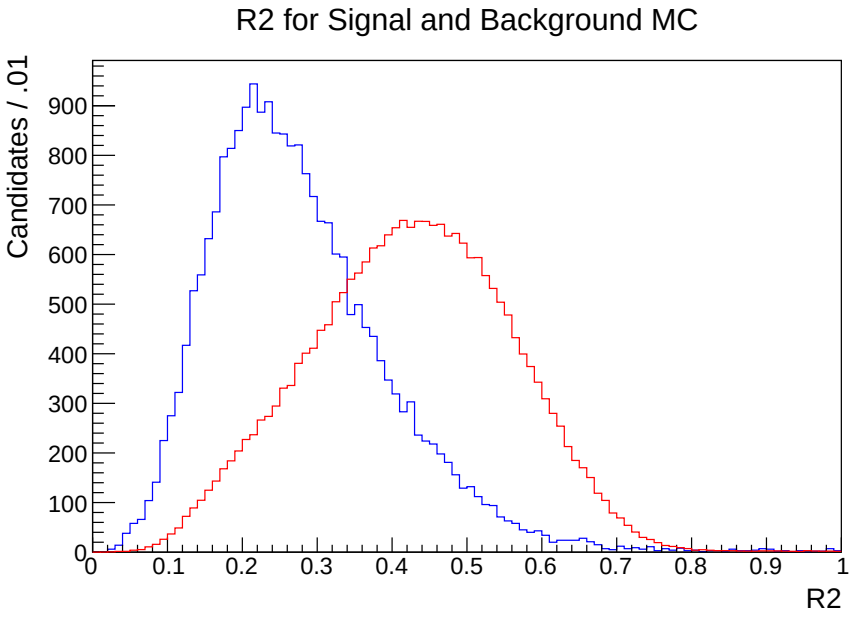

Figure 14: Distribution of R2 for signal (blue) and background (red) MC

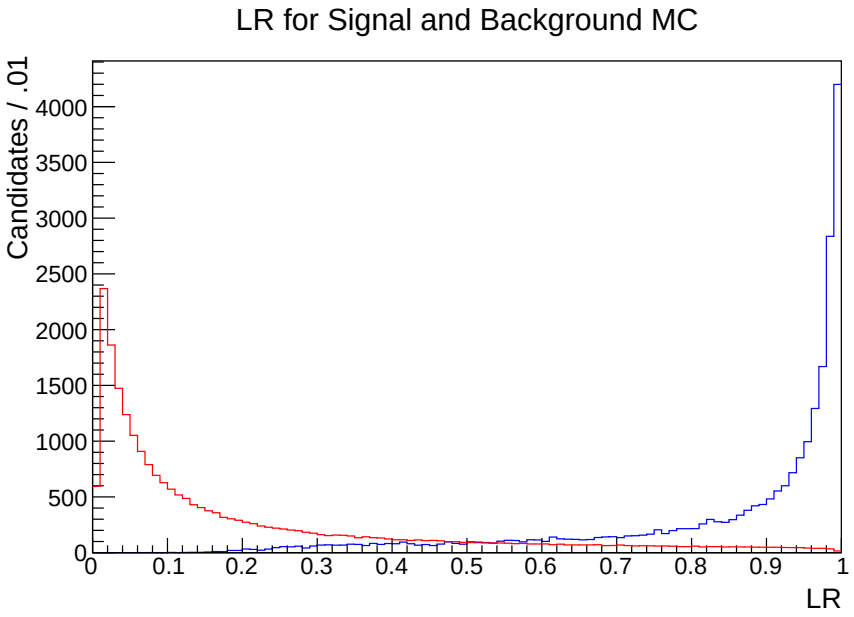

Figure 15: Distribution of $\mathcal{LR}$ for signal (blue) and background (red) MC

5.2 BACKGROUND SUPPRESSION STRATEGY

5.2.1 Likelihood Ratio and RooKSFW

In order to suppress continuum backgrounds, we first compute a likelihood ratio using the program RooKSFW. RooKSFW optimizes a Fischer discriminant using the 18 KSFW Moments described in the previous section as discriminating variables. This program then creates a likelihood ratio defined as $\mathcal{LR} = \frac{P_s}{P_s+P_b}$ where P_s is the probability of the event being a signal event and P_b is the probability of the event being a background event. This ratio peaks at 1 for signal events and 0 for background events. One stream of generic MC background and our MC sample of $B_s \to \eta'\eta$ events are used to calculate the likelihood ratio. Fig. 15 shows the distribution of $\mathcal{LR}$, which is used as an input to a neural network described in the following section.

5.2.2 NeuroBayes

We then use NeuroBayes to further suppress background events. NeuroBayes is a neural network package designed for classification problems, which uses a multivariate method based on Bayes' Theorem. NeuroBayes takes input variables and provides separation between signal and background events. By taking all input variables and their correlations into account for each event, NeuroBayes computes a single discriminating variable (NN) with a value between -1 and 1. Events with an NN value closer to -1 are more likely to be background events while events with an NN value closer to 1 are more likely to be signal events.

Events which pass our selection criteria described in Table 2 are input to NeuroBayes. We use four input variables: $\mathcal{LR}$, $|\cos(\theta_{thrust})|$, $|\cos(\theta_{bt})|$, and R_2. The use of NeuroBayes is separated into two steps. First, the network is trained using signal and background samples. Because this optimization requires similar numbers of signal and background events, we use our MC sample of $B_s \to \eta'\eta$ signal events and 1/10 of the events in one stream of generic MC background. We then validate the training of our neural net using independent samples of both signal and background MC. Within statistical uncertainties, these independent samples give the same results as our training sample and indicate that the neural net has not been overtrained. The results of these training and validation steps are shown in Fig. 16.

The distribution of NN is shown in Fig. 16. We add an additional selection criterion by requiring $NN_{min} > 0.6$. Fig. 17 shows the signal purity and background rejection for several values of NN_{min}. By requiring $NN_{min} > 0.6$, we are able to remove 96% of background events while retaining 76% of signal events. Because of the large peaks at -1 and 1, this distribution is difficult to fit using simple functions. In order to include this variable in our multidimensional fit, we define a new variable NB':

$$NB' = \log \frac{NN - NN_{min}}{NN_{max} - NN} \tag{5.10}$$

where we set $NN_{min} = 0.6$ and $NN_{max} = 0.998$. $NN_{min} = 0.6$ is given by our selection criteria on NN and is used to reject continuum background events. $NN_{max} = 0.998$ is given by the maximum value of NN obtained from our sample of signal MC decays and is used to ensure the distribution is symmetric (*i.e.* at larger values of NN_{max} the distribution becomes

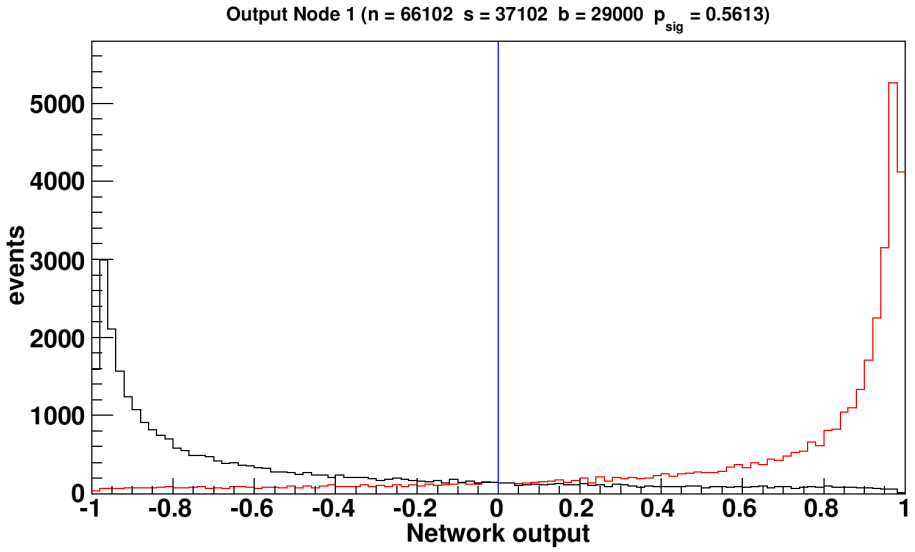

Figure 16: Distribution of NN for signal (blue) and background (red) MC

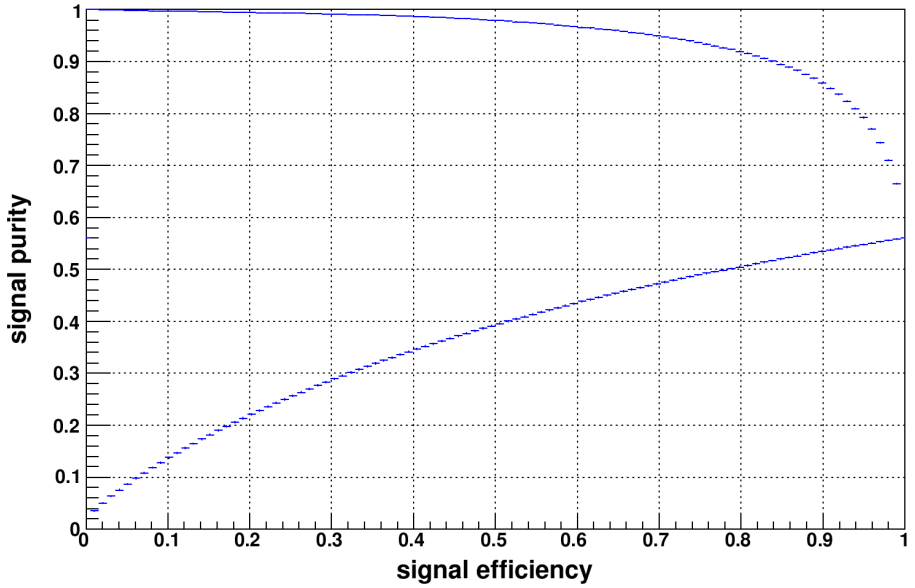

Figure 17: Plot of the signal purity vs signal efficiency for several values of NN_{min}

asymmetric). This function transforms our NN variable for both signal and background events to a more Gaussian-like shape shown in Fig. 18.

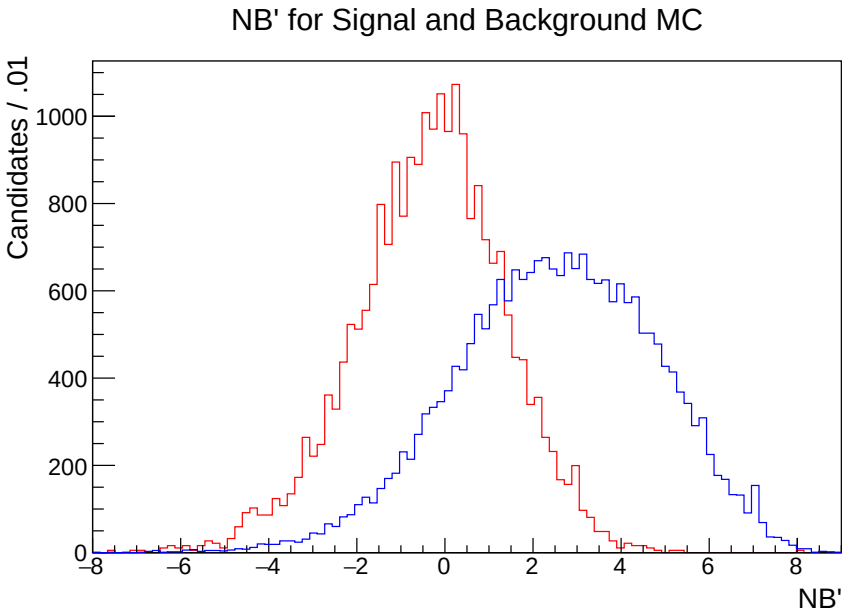

Figure 18: Distributions of NB′ for signal (blue) and background (red) MC

6.0 BEST CANDIDATE SELECTION

After applying all selection criteria and background suppression, 15% of signal events contain multiple B_s candidates. The number of candidates per event before best candidate selection is shown in Fig. 19. For these events, we define a variable $\chi^2 = \chi^2_\eta + \chi^2_{\pi^+\pi^-}$ where χ^2_η and $\chi^2_{\pi^+\pi^-}$ are the χ^2 of the mass of the η candidate defined as:

$$\chi^2_\eta = \left(\frac{m_{\gamma\gamma} - m_\eta}{\sigma_{\gamma\gamma}}\right)^2 \tag{6.1}$$

and the χ^2 of the vertices of the $\pi^\pm$ candidates which is the distance between the tracks of the pions and the presumed vertex in units of standard deviations. However, using this method, we find that in a significant number of events, multiple candidates have equal χ^2.

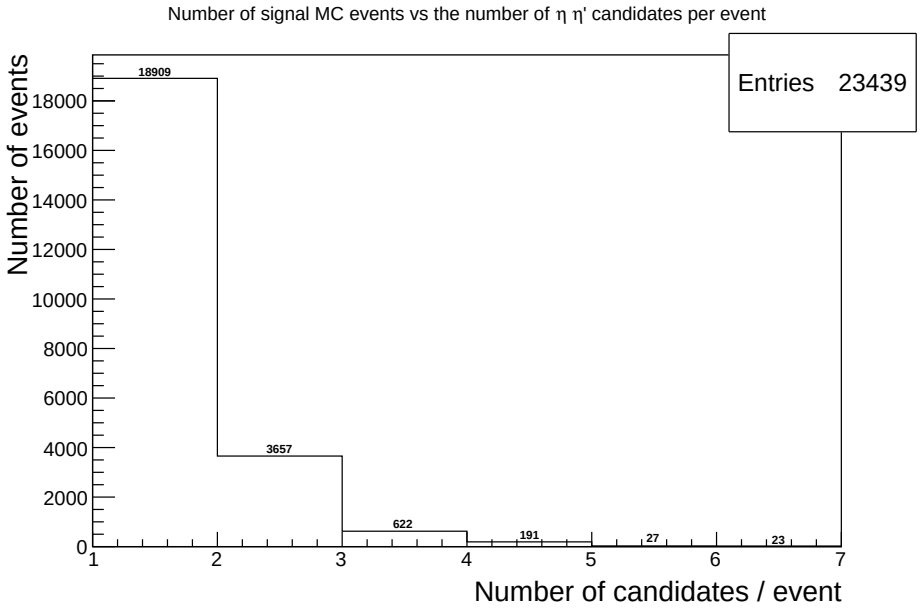

Figure 19: Number of candidates per event before best candidate selection

This occurs because multiple candidates have daughter η candidates reconstructed from the same photons and ρ^0 candidates reconstructed from the same pions. In these cases, the B_s candidates only differ by the photon coming from the η' decay $\eta' \to \rho^0\gamma$. Out of these candidates, we choose the candidate whose photon from this decay has the greatest energy. Based on MC simulations, this best candidate selection method chooses the correct B_s candidate 96% of the time.

7.0 FITTING

In order to extract the signal yield, we use RooFit to perform a four-dimensional (4D) extended unbinned maximum likelihood fit to the variables M_{bc}, ΔE, NB', and $M_{\eta'}$. The likelihood is defined as

$$\mathcal{L} = e^{N_{sig}+N_{bkg}} \prod_{i=1}^{N} \left(N_{sig}\mathcal{P}_{sig}(M_{bc}^i, \Delta E^i, NB'^i, M_{\eta'}^i) + N_{bkg}\mathcal{P}_{bkg}(M_{bc}^i, \Delta E^i, NB'^i, M_{\eta'}^i) \right) \quad (7.1)$$

where N is the total number of events, N_{sig} is the total number of signal events, and N_{bkg} is the total number of background events. We allow N, N_{sig}, and N_{bkg} to float in our fit while holding all other fitting parameters constant. $\mathcal{P}_{sig}$ and $\mathcal{P}_{bkg}$ are the probability density functions (PDF) of the signal and background components of our fitting functions. The PDFs are factorized as

$$\mathcal{P}_{sig,bkg}(M_{bc}, \Delta E, NB', M_{\eta'}) = \mathcal{P}_{sig,bkg}(M_{bc}) \cdot \mathcal{P}_{sig,bkg}(\Delta E) \cdot \mathcal{P}_{sig,bkg}(NB') \cdot \mathcal{P}_{sig,bkg}(M_{\eta'}) \quad (7.2)$$

7.1 CORRELATIONS BETWEEN FITTING VARIABLES

In order to factorize our PDF as shown in Equation 7.2 our variables must be uncorrelated. Two dimensional distributions of all combinations of our fitting variables are shown in Fig. 20 for signal MC and in Fig. 21 for background MC. These distributions show no significant correlations between our fitting variables. Possible small correlations ignored in our approximations of $\mathcal{P}_{sig,bkg}(M_{bc}, \Delta E, NB', M_{\eta'})$ will be included in the systematic uncertainties.

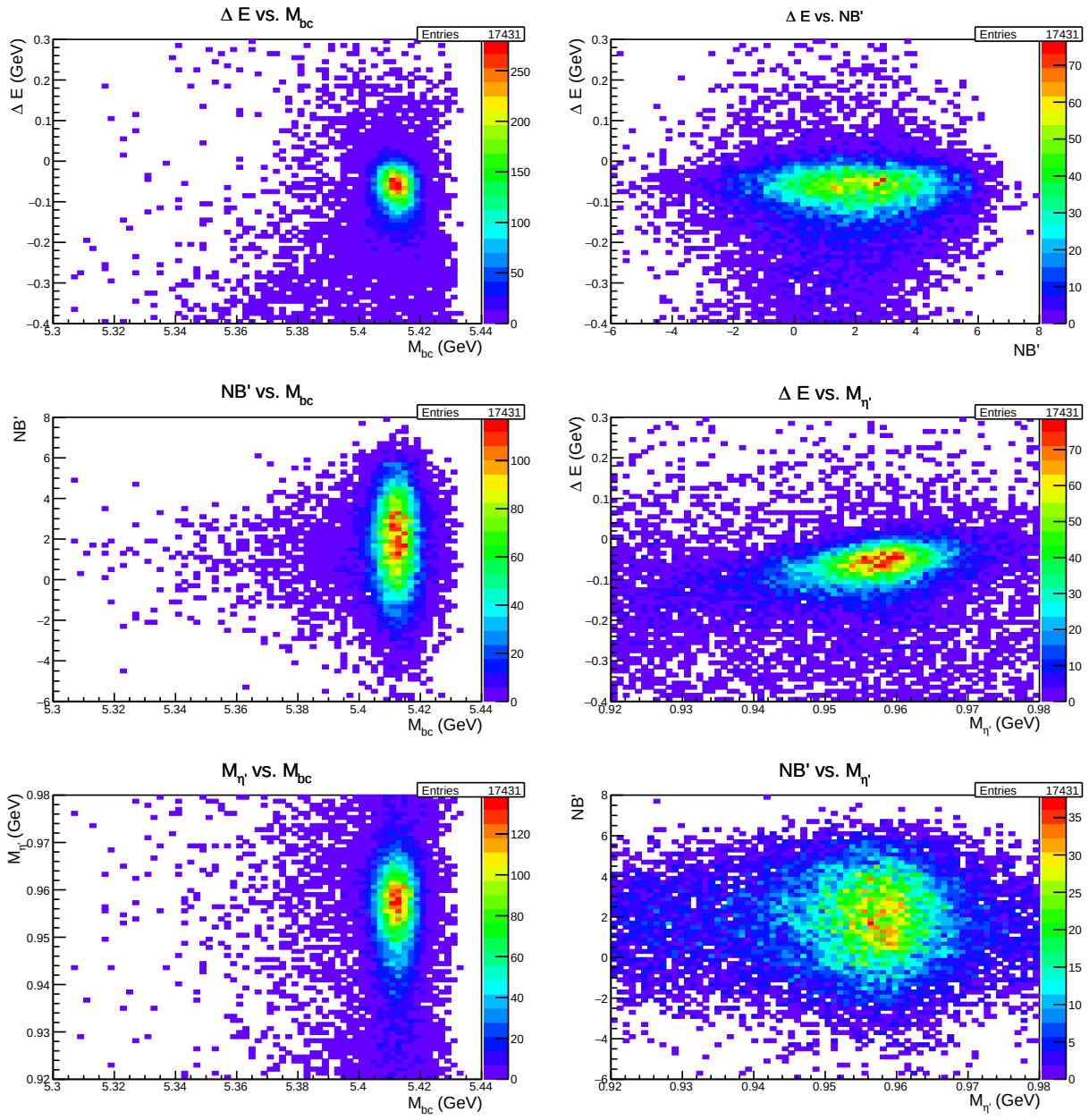

Figure 20: 2D distributions showing no significant correlations between our fitting variables in signal MC.

7.2 SIGNAL PDFS

We fit the signal distributions of four variables – M_{bc}, ΔE , NB', and $M_{\eta'}$ – in our signal MC sample. The distributions of M_{bc}, ΔE, and NB', in signal MC are described by the

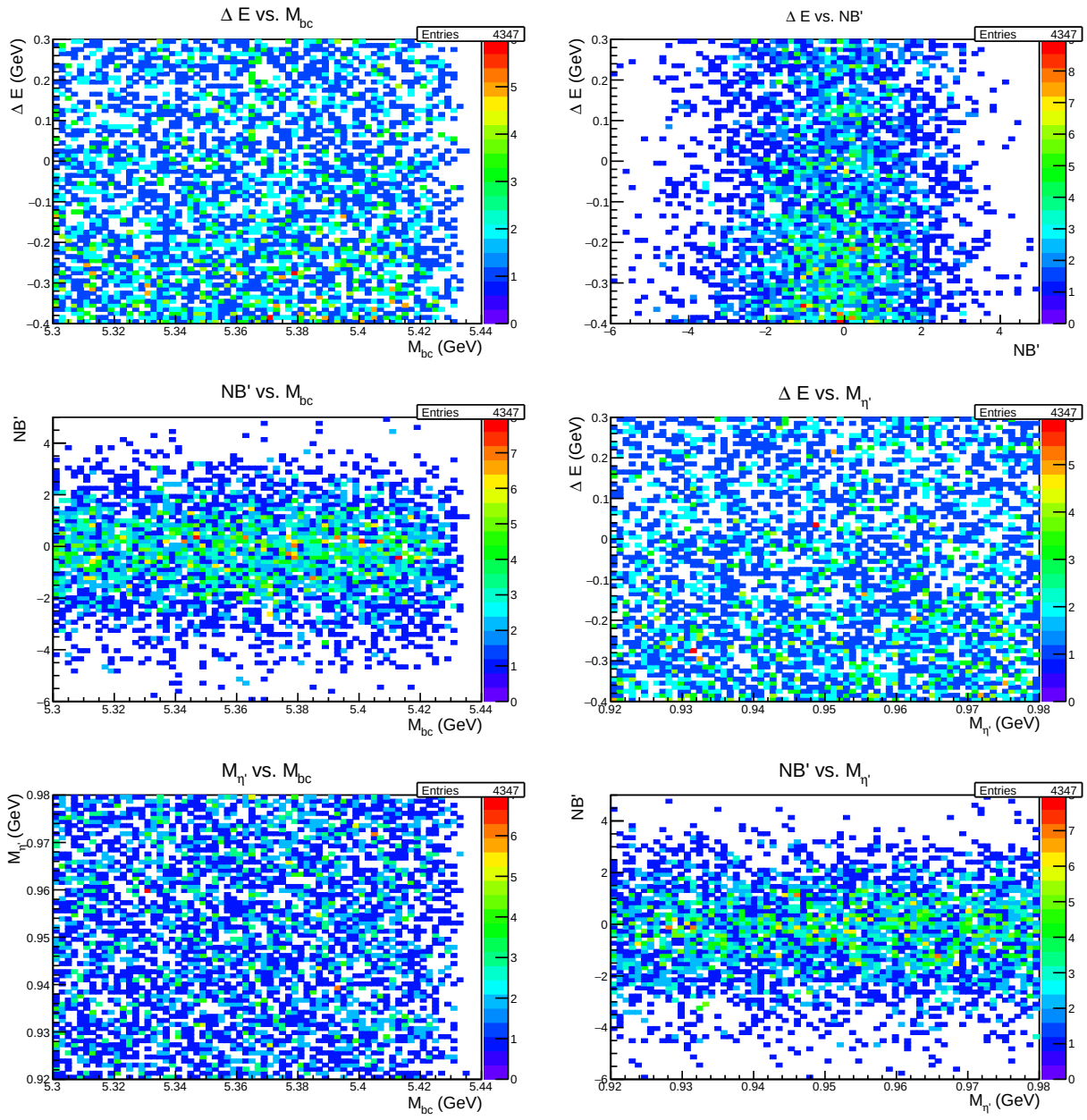

Figure 21: 2D distributions showing no significant correlations between our fitting variables in background MC.

sum of a Gaussian and a Crystal Ball function. The Crystal Ball function is a piecewise function which consists of a Gaussian above a certain threshold and a power law below the same threshold. The distribution of $M_{\eta'}$ in signal MC is described by the sum of a Crystal Ball function and a first order Chebychev polynomial (a straight line). These PDFs are

Table 3: Fitting functions for signal distributions

Fitting Variable	Signal Fitting Functions
M_{bc}	Gaussian
	Crystal Ball
ΔE	Gaussian
	Crystal Ball
NB'	Gaussian
	Crystal Ball
$M_{\eta'}$	Chebychev Polynomial (First Order)
	Crystal Ball

summarized in Table 3. When fitting, the means of the Gaussians and Crystal Balls for the signal distributions of M_{bc}, ΔE, and NB' are required to be the same value. In addition, when fitting the signal distribution for $M_{\eta'}$, we fix the mean to the nominal mass of η' (0.958 GeV). All other parameters of the signal distributions are allowed to float in the fit. When performing the maximum likelihood fit to the real data, the paramaters of our fitting functions are fixed to the values obtained in these fits. The distributions and fitted functions are plotted and shown in Fig. 22.

7.3 BACKGROUND PDFS

We fit the background distributions of four variables – M_{bc} ,ΔE, NB', and $M_{\eta'}$ – first using the sidebands of our background MC sample. The distribution of M_{bc} in background MC is

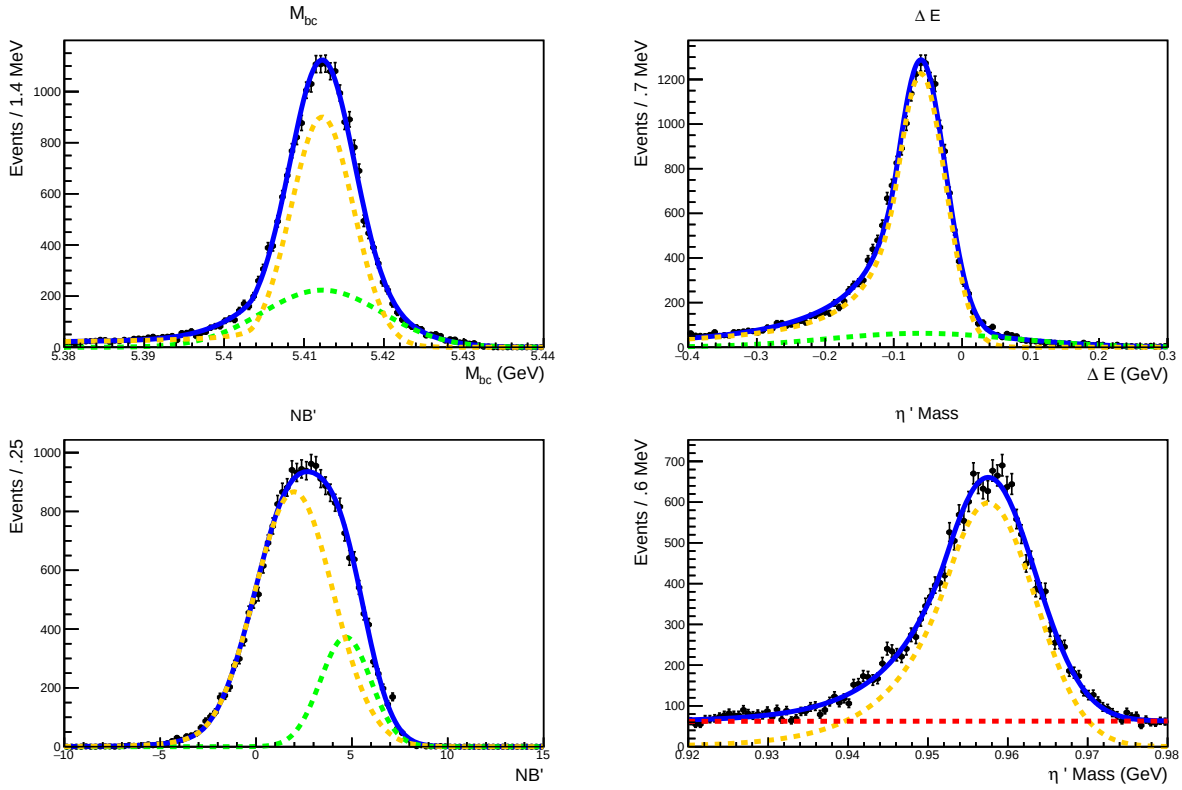

Figure 22: Projections of 4D maximum likelihood fit to the signal MC sample

described using an ARGUS function defined as:

$$\mathcal{P}_{\text{ARGUS}}\left(M_{bc};\alpha,E_{beam}\right)=M_{bc}\sqrt{1-\left(\frac{M_{bc}}{E_{beam}}\right)^{2}}e^{-\alpha\left(1-(M_{bc}/E_{beam})^{2}\right)} \tag{7.3}$$

where the parameter α defines the shape of the distribution and E_{beam} is the cutoff. The distribution of ΔE is described using a second order Chebychev polynomial (a parabola). The distribution of NB' is described using a Gaussian. The distribution of $M_{\eta'}$ in background MC is described by a first order Chebychev polynomial. These PDFs are summarized in Table 4. All parameters of the background distributions are allowed to float in the fit. We then fix all parameters to the values obtained in the fit to the sidebands of our background MC sample and perform a fit to our entire background MC sample including the signal region. This fit is shown in Fig. 23.

Because our fit to sideband MC can be applied to the full MC sample as shown above, we use the same procedure when performing a maximum likelihood fit to the Belle data sample.

Table 4: Fitting functions for background distributions

Fitting Variable	Signal Fitting Functions
M_{bc}	ARGUS
ΔE	Chebychev Polynomial (Second Order)
NB'	Gaussian
$M_{\eta'}$	Chebychev Polynomial (First Order)

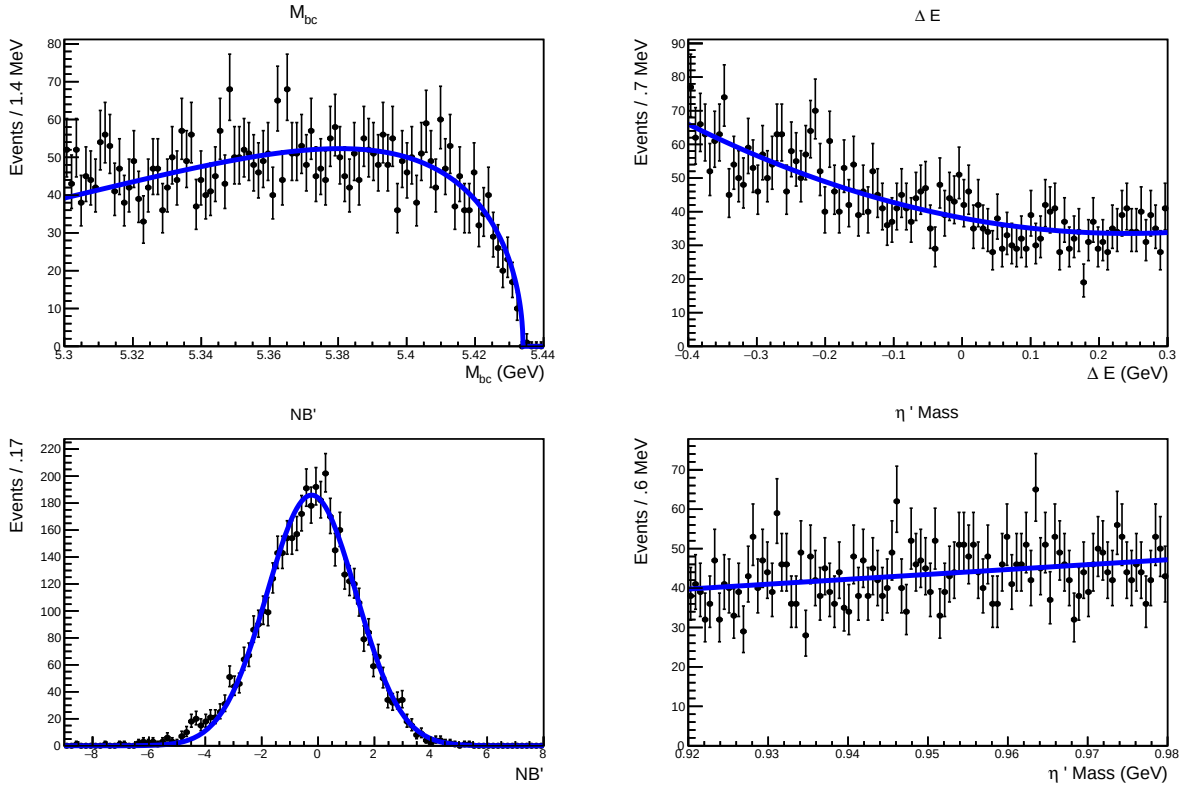

Figure 23: Projections of 4D maximum likelihood fit to the full background MC sample

The results of our maximum likelihood fit to sideband data are shown in Fig. 24. Again all parameters are floated in this fit and the values obtained for the parameters will be fixed in the fit to the full data sample.

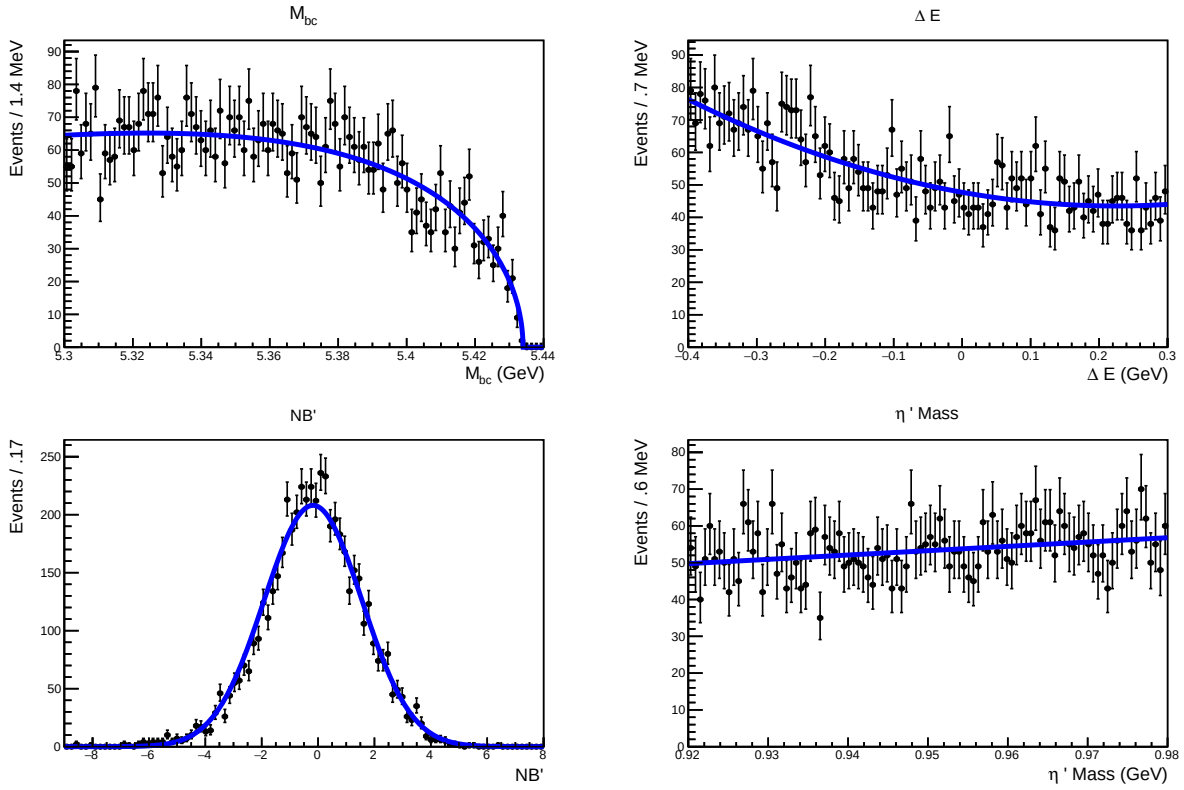

Figure 24: Projections of 4D maximum likelihood fit to the sidebands of the Belle data sample

7.4 SIGNAL EFFICIENCY AND EXPECTED BACKGROUND

By performing maximum likelihood fits to our signal MC distributions we are able to extract the signal yield from our signal MC sample. This MC sample consisted of 100,000 signal MC events and our 4D fit returns a value of $N_{sig} = 23,439$. This indicates an efficiency of $\epsilon = 23.4\%$ for our signal MC events.

By performing maximum likelihood fits to our background MC distributions we are able to extract the number of expected background events in our data sample. We are able to improve this approximation by comparing the number of background events in our MC sidebands to the full MC sample. This ratio is then used to estimate the number of events in the full Belle data sample from its sidebands. Using this approximation we expect 5,500 background events in the full Belle data sample.

7.5 ENSEMBLE TESTS

We perform so-called "ensemble tests" in order to test our fitting model for possible biases and estimate its associated systematic errors. These ensemble tests use generated toy MC experiments to test the validity of the model. In each toy MC experiment, we generate a number of signal events, N_{sig}^{gen}, and a number of background events, N_{bkg}^{gen}. These events are generated based only on the PDF lineshapes of our fitting functions for signal and background. In order to accurately estimate systematic errors, these values must be statistically similar to the numbers of events expected in the real data. The generated events are then fit using a sum of our multidimensional PDFs for signal and background events multiplied by N_{sig}^{fit} and N_{bkg}^{fit}:

$$\mathcal{D}_{both}(M_{bc}, \Delta E, NB', M_{\eta'}) = N_{sig}^{fit}\mathcal{P}_{sig}(M_{bc}, \Delta E, NB', M_{\eta'}) + N_{bkg}^{fit}\mathcal{P}_{bkg}(M_{bc}, \Delta E, NB', M_{\eta'}) \tag{7.4}$$

This process is repeated 1,000 times for varying values of N_{sig}^{gen} and a constant value of N_{bkg}^{gen} equal to the expected number of background events in the real data sample (5,500). For each ensemble test consisting of 1,000 toy MC studies, we plot histograms of the extracted number of signal events (N_{sig}^{fit}), the error on the extracted number of signal events (σ_{sig}^{fit}), and the pull (P). Where the pull is defined as the number of standard deviations the fitted number of events is from the generated number of events:

$$P = \frac{N_{sig}^{fit} - N_{sig}^{gen}}{\sigma_{sig}^{fit}} \tag{7.5}$$

In the case where our fitting parameter (N_{sig}^{fit}) is unbiased, we expect the pull distribution to be a Gaussian with a mean of zero and a standard deviation of one. We also plot the the negative log likelihood, $-\ln\mathcal{L}$, of toy MC experiments. Several ensemble tests are performed and the distributions described above are plotted in Fig. 25. These distributions show unbiased fits with a large number of signal events, however, when N_{sig} is small, we see an asymmetry in the pull distributions indicating a bias. This bias is expected due to

the small number of events in the signal region. This bias occurs only with a small value of N_{sig}. Therefore, we use a frequentist method to determine an upper limit on the branching fraction in the case where our fit extracts a small value for N_{sig}. A frequentist approach using confidence intervals accounts for this bias.

7.6 CONFIDENCE INTERVALS AND SENSITIVITY ESTIMATE

We perform ensemble tests to create the 90% confidence belt for a frequentist approach. We use similar number of N_{bkg}^{gen} as expected in data and vary N_{sig}^{gen} from 0 to 70 events. The lower bounds of the 90% confidence belt are given by the fitted number of signal events N_{sig}^{fit} for which 5% of the results are less than this value. The upper bounds of the 90% confidence belt are given by the fitted number of signal events N_{sig}^{fit} for which 5% of the results are greater than this value. This confidence belt is shown in Fig. 26. After performing a fit to the full data sample, this confidence belt will be used to either set an upper limit on the branching fraction or to claim a discovery of the decay.

After creating this confidence belt, we estimate the upper limit on the branching fraction of this decay in the absence of signal. We perform a maximum likelihood fit to all four streams of background MC and average the result to correspond to the statistics in data. There is an average signal yield of 5.4 events in the four streams of background MC. Using our confidence belt in Fig. 26, we determine the 95% upper limit on the number of signal events is 26 events. Using Equation 2.1, we calculate the upper limit on the branching fraction in the absence of signal to be $\mathcal{B}(B_s \to \eta'\eta) = 7.6 \times 10^{-5}$.

7.7 PRELIMINARY FIT TO 15% OF THE DATA SAMPLE

We use RooFit to perform a 4D extended unbinned maximum likelihood fit to the data recorded by Belle experiment 53. This sample corresponds to 15% of the entire Belle data sample. In order to perform this fit, we first perform a 4D fit to the sidebands of the data

sample using only background PDFs and floating all parameters. The results of this fit is shown in Fig. 27. After performing this fit, we fix all PDF parameters to the values obtained in the fits except for N_{sig} and N_{bkg} which are allowed to float. We then fit the partial data sample including the signal region and obtain a signal yield of $N_{sig} = 2.6 \pm 4.0$ events where 4.0 is the statistical error on our measurement. The results of this fit projected to all four fitting variables are shown in Fig. 28. In this projection, we require all events to be in the signal region of the other three fitting variables. This greatly reduces the number of background events as most background comes from the contributions of sidebands.

7.8 UPPER LIMIT ESTIMATION FOR 15% OF THE DATA SAMPLE

Using the methods described in Section 7.6, we construct a confidence belt for the experiment 53 data sample. After creating this confidence belt, we estimate the upper limit on the branching fraction of this decay based on the results for the experiment 53 data sample. The signal yield is 2.6 events and by using the confidence belt in Fig. 29, we determine the 95% upper limit on the number of signal events to be 16 events. Using Equation 2.1, we calculate the upper limit on the branching fraction from experiment 53 data to be $\mathcal{B}(B_s \to \eta'\eta) = 2.7 \times 10^{-4}$. This result is consistent with the estimate of the branching fraction upper limit of 7.6×10^{-5} from the full background MC sample.

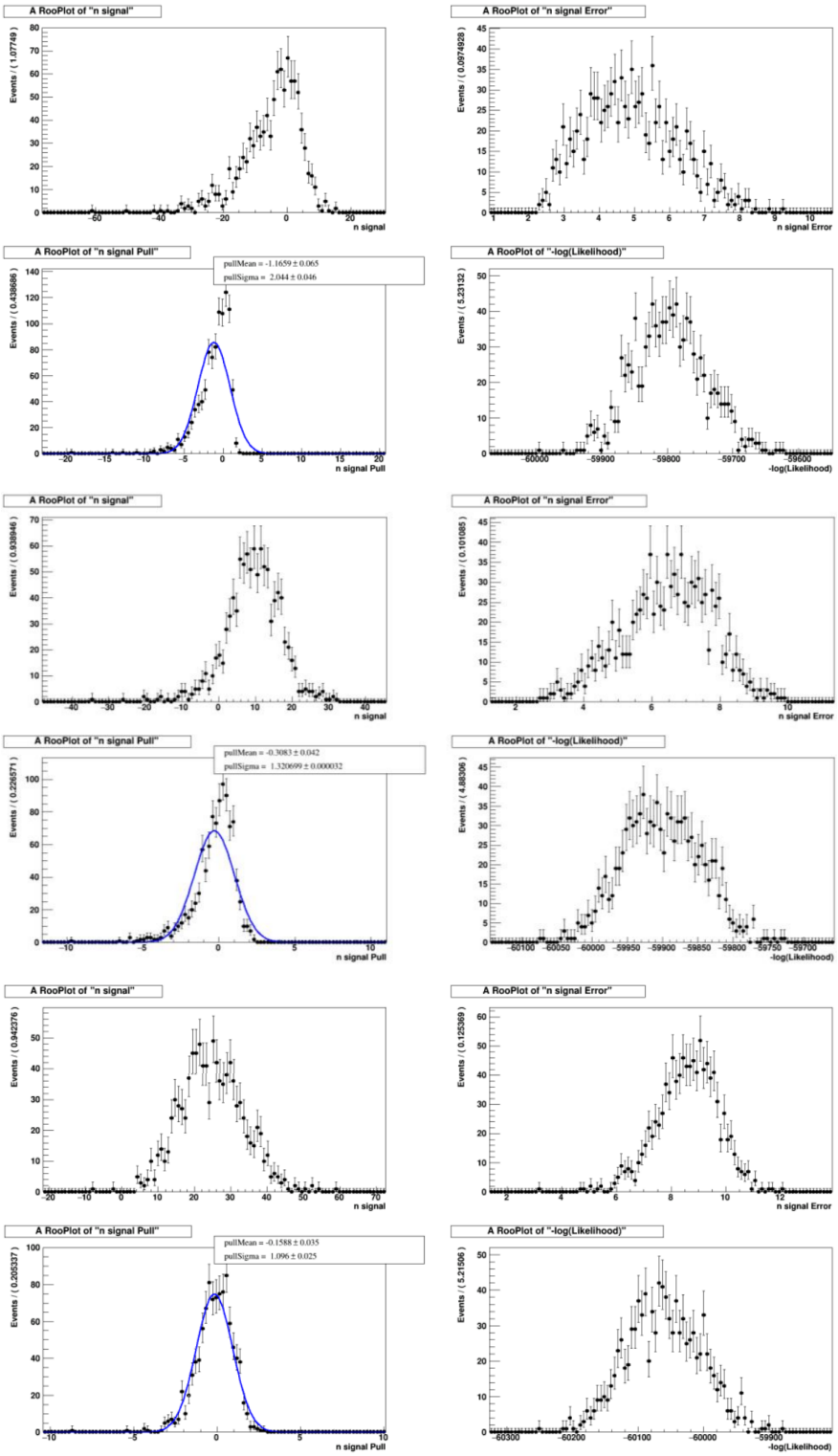

Figure 25: Results of ensemble tests with $N_{sig} = 0$ (top), $N_{sig} = 10$ (middle) and $N_{sig} = 25$ (bottom)

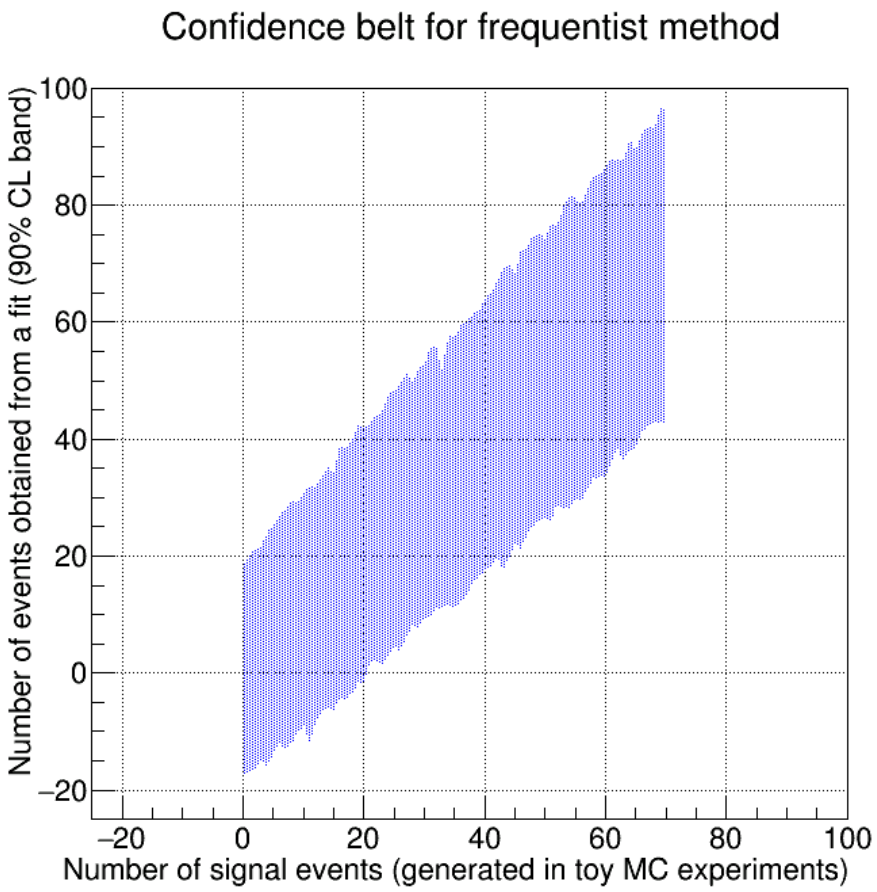

Figure 26: 95% Confidence belt for the full data sample

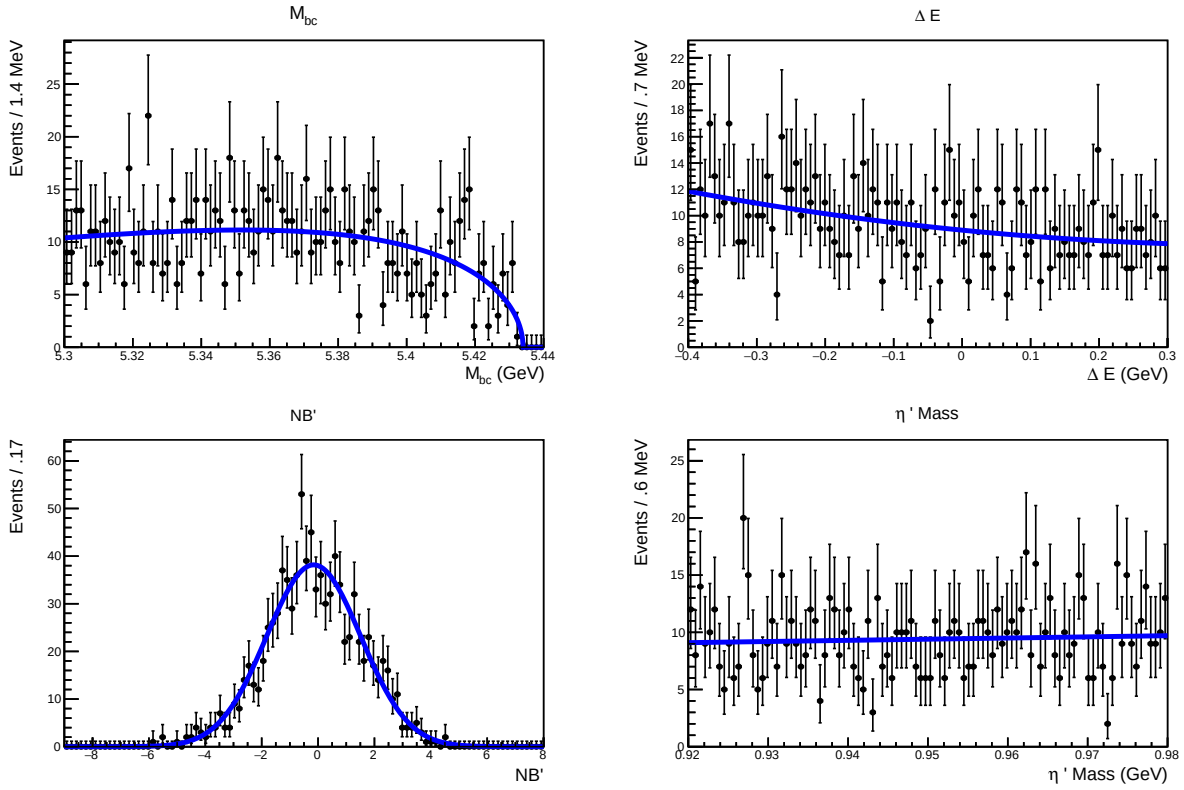

Figure 27: Projections of the 4D maximum likelihood fit to the sidebands of 15% of the Belle experiment 53 data sample

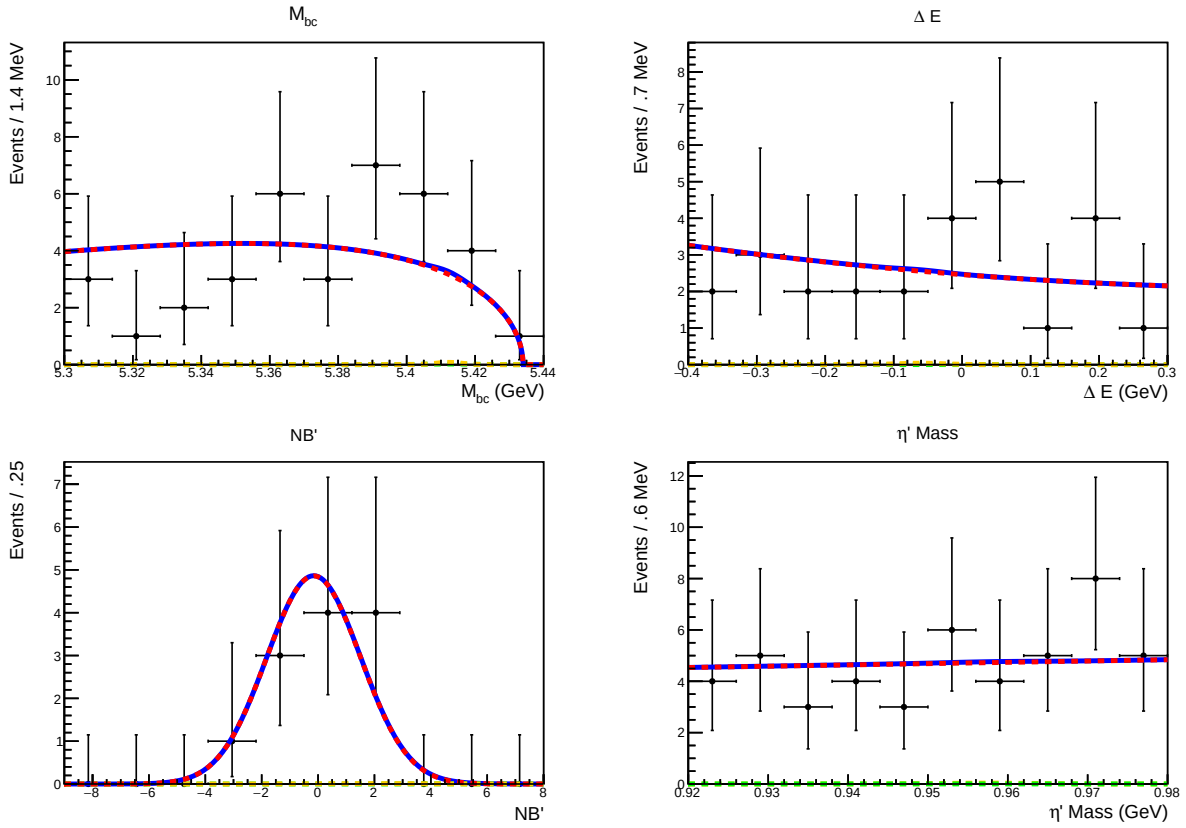

Figure 28: Projections of the 4D maximum likelihood fit to 15% of the Belle experiment 53 data sample

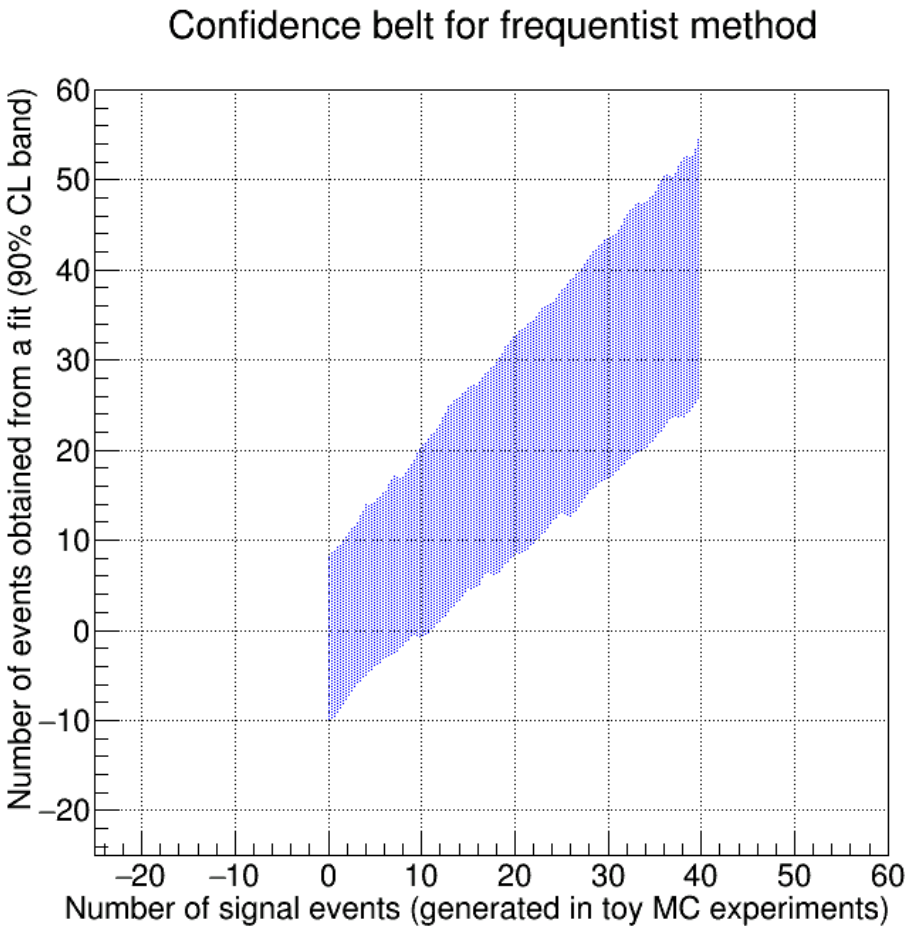

Figure 29: 95% Confidence belt for the experiment 53 data sample

8.0 OUTLOOK AND CONCLUSIONS

In this analysis we search for the decay $B_s \to \eta'\eta$. By reconstructing B_s candidates using the decay channels $\eta' \to \rho^0\gamma$ and $\eta \to \gamma\gamma$, optimizing our selection criteria and background suppression techniques on MC, and performing a 4D maximum likelihood fit to 15% of the Belle data sample, we extract a signal yield of $N_{sig} = 2.6 \pm 4.0$ events. This fit indicates no evidence of $B_s \to \eta'\eta$ decays in our data sample.

In the near future, we will perform a fit to the full Belle data sample to estimate the signal yield. We will use this measurement to either measure or set an upper limit on the branching fraction of $B_s \to \eta'\eta$. Studies of systematic uncertainties must then be performed to determine the consistency of the measured branching fraction with its theoretical prediction. If this measurement disagrees with the Standard Model prediction, this result would suggest BSM physics. However, if this measurement agrees with the Standard Model prediction, it could still be used to constrain the possible effects of NP or to rule out specific models.

In order to improve our analysis, we could reconstruct multiple channels of η' and η decays. While $\eta' \to \rho^0\gamma$ and $\eta \to \gamma\gamma$ have the largest branching fraction of $B_s \to \eta'\eta$ decays at $\mathcal{B} = 0.114$, there are five more decay channels of η and η' with $\mathcal{B} > 0.05$. Including these decay channels in our analysis will improve our sensitivity to this decay.

In addition, because $B_s \to \eta'\pi^0$ can have the same final state, we could extend this analysis to also include a search for this decay. However, the predicted branching fraction in the Standard Model is significantly smaller and a discovery would be much less likely.

BIBLIOGRAPHY

[1] Wikimedia Commons. Standard model of elementary particles, 2006.

[2] S. Kurokawa and E. Kikutani. Overview of the KEKB accelerators. *Nuclear Instruments and Methods A*, 499(1):1, 2003.

[3] A. Abashian et. al. (Belle Collaboration). The Belle Detector. *Nuclear Instruments and Methods A*, 479(1):117, 2002.

[4] M. Artuso et al. B, D and K Decays. *Eur. Phys. J.*, C57:309, 2008.

[5] Alexander R. Williamson and Jure Zupan. Two body B decays with isosinglet final states in soft collinear effective theory. *Phys. Rev. D*, 74:014003, 2006.

[6] Ahmed Ali, Gustav Kramer, Ying Li, Cai-Dian Lü, Yue-Long Shen, Wei Wang, and Yu-Ming Wang. Charmless nonleptonic B_s decays to PP, PV, and VV final states in the perturbative QCD approach. *Phys. Rev. D*, 76:074018, 2007.

[7] Hai-Yang Cheng and Chun-Khiang Chua. QCD factorization for charmless hadronic B_s decays revisited. *Phys. Rev. D*, 80:114026, 2009.

[8] Hai-Yang Cheng, Cheng-Wei Chiang, and An-Li Kuo. Updating $B \to PP, VP$ decays in the framework of flavor symmetry. *Phys. Rev. D*, 91:014011, 2015.

[9] Yu-Kuo Hsiao, Chia-Feng Chang, and Xiao-Gang He. Global $SU(3)/U(3)$ flavor symmetry analysis for $B \to PP$ with $\eta - \eta'$ mixing. *Phys. Rev. D*, 93:114002, 2016.

[10] R. Aaij et. al. (LHCb Collaboration). Observation of the $B_s^0 \to \eta'\eta'$ decay. *Phys. Rev. Lett.*, 115(5):051801, 2015.

[11] R. A. Fisher. The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, 7(2):179, 1936.

[12] Mikihiko Nakao. Continuum suppression. Belle Analysis School, 2009. Unpublished.

[13] M. Feindt, F. Keller, M. Kreps, T. Kuhr, S. Neubauer, D. Zander, and A. Zupanc. A hierarchical NeuroBayes-based algorithm for full reconstruction of B mesons at B factories. *Nuclear Instruments and Methods A*, 654(1):432, 2011.

[14] Wouter Verkerke and David P. Kirkby. The RooFit toolkit for data modeling. 2003.

[15] David J. Lange. The EvtGen particle decay simulation package. *Nuclear Instruments and Methods A*, 462(12):152, 2001.

[16] S. Agostinelli et al. Geant4 simulation toolkit. *Nuclear Instruments and Methods A*, 506(3):250, 2003.

[17] Geoffrey C. Fox and Stephen Wolfram. Observables for the analysis of event shapes in e^+e^- annihilation and other processes. *Phys. Rev. Lett.*, 41:1581, 1978.