Belle Note 1522

Version 1.0
April 29, 2019

Search for decay $\Upsilon(5S) \to \gamma W_{bj}$

Nicholas Corrado, Vladimir Savinov

University of Pittsburgh, Pittsburgh, PA-15260, USA

Abstract

The recent discovery of the states Z_b and Z_b' implies the possible existence of a new family of hadronic resonances including molecular states dubbed W_{bJ}. We describe a search for W_{bJ} in the decay $\Upsilon(5S) \to \gamma W_{bJ}$ using 121.4 fb^{-1} of data collected at the Υ(5S) resonance with the Belle detector at the KEKB asymmetric-energy electron-positron collider. Using Monte Carlo simulation, we study Belle's sensitivity to the decay $\Upsilon(5S) \to \gamma W_{bJ}$, search for its presence in Belle data and describe the procedure we would use to establish an upper limit on the visible production cross section for these new states.

Contents

List of Figures

List of Tables

1 Introduction

1.1 Motivation

In this document, we describe a search for new hadronic states of matter – bottomonium-like particles dubbed W_{bJ} – in radiative decays of $\Upsilon(5S)$. These states are believed to be of molecular nature, where a pair of colored $B_{(s)}^{(*)}$ mesons, each containing a b or an anti-b quark, are held together by the strong interaction (in a way similar to single-pion exchange force mechanism in QCD-inspired low-energy models). As with conventional bottomonium, *i.e.* $b\bar{b}$ states, these molecular states exhibit their own spectroscopy. However, their masses and properties obviously could not be predicted using $q\bar{q}$ potential models. We are motivated by Belle's discoveries [1, 2, 3, 4] of the $Z_b(10610)$ and $Z_b(10650)$ states (referred to in the rest of this document as Z_b and Z_b' or just Z_b) and theoretical predictions which use the molecular picture to explain the nature of the Z_b and predict the existence of additional hadronic states. These predictions can be used to explain various long-standing puzzles in the (no longer pure) bottomonium at energies above the threshold for B meson pair production.

1.2 New Spectroscopy

Since the discovery of the Υ meson, the b quark, and B mesons [5], conventional bottomonium states have been a rich source of information about strong interaction dynamics in the approximately non-relativistic $b\bar{b}$ system. Vector bottomonium and bottomonium-like states ($\Upsilon(nS)$ mesons) can be produced directly in the e^+e^- annihilation. Three of these states – $\Upsilon(1S)$, $\Upsilon(2S)$, and $\Upsilon(3S)$ – have masses below the $B\bar{B}$ threshold [6]. These states are believed to be pure $b\bar{b}$, and their properties are relatively easy to understand using potential models. Such relativized models [7] predict 34 $b\bar{b}$ bound states below $\Upsilon(4S)$ energy, 15 of which have been observed. We show the predictions for the energy levels in the $b\bar{b}$ spectroscopy [8, 9] in Fig. 1.

Hadronic transitions (such as, *e.g.* $\Upsilon(3S) \to \pi^+\pi^-\Upsilon(1S)$) between bottomonium states provide an excellent opportunity to study QCD dynamics in non-perturbative regime by comparing the measured masses, widths, branching fractions, angular and invariant mass distributions with the theoretical predictions. For pure bottomonium states – $b\bar{b}$ resonances below $B\bar{B}$ threshold – the hadronic transitions proceed via radiating the strong field, *i.e.*, by emitting the gluons which convert into light hadrons. States above $B\bar{B}$ threshold, starting with $\Upsilon(4S)$, are significanly wider than the lower-mass states, and their hadronic transitions are known to exhibit certain properties that are unexpected for pure $b\bar{b}$ states. While the latter are well described from the perspective of Heavy Quark Spin Symmetry (HQSS) where transitions involving the spin of the heavy b quark are strongly suppressed, the former states, including the $\Upsilon(5S)$, require a different explanation [10].

The favored explanation for the properties of $\Upsilon(5S)$, including its decays to Z_b, is based on the molecular picture, where these vector bottomonium-like resonances are assumed to contain an admixture of pairs of colored heavy mesons. This hypothesis has been successfully employed [11] to explain the decays to and the existence of the six Z_b states. However, the details of the interaction responsible for these processes are not yet fully understood. Alternative explanations include a model with a diquark-antidiquark pair, where a pair of

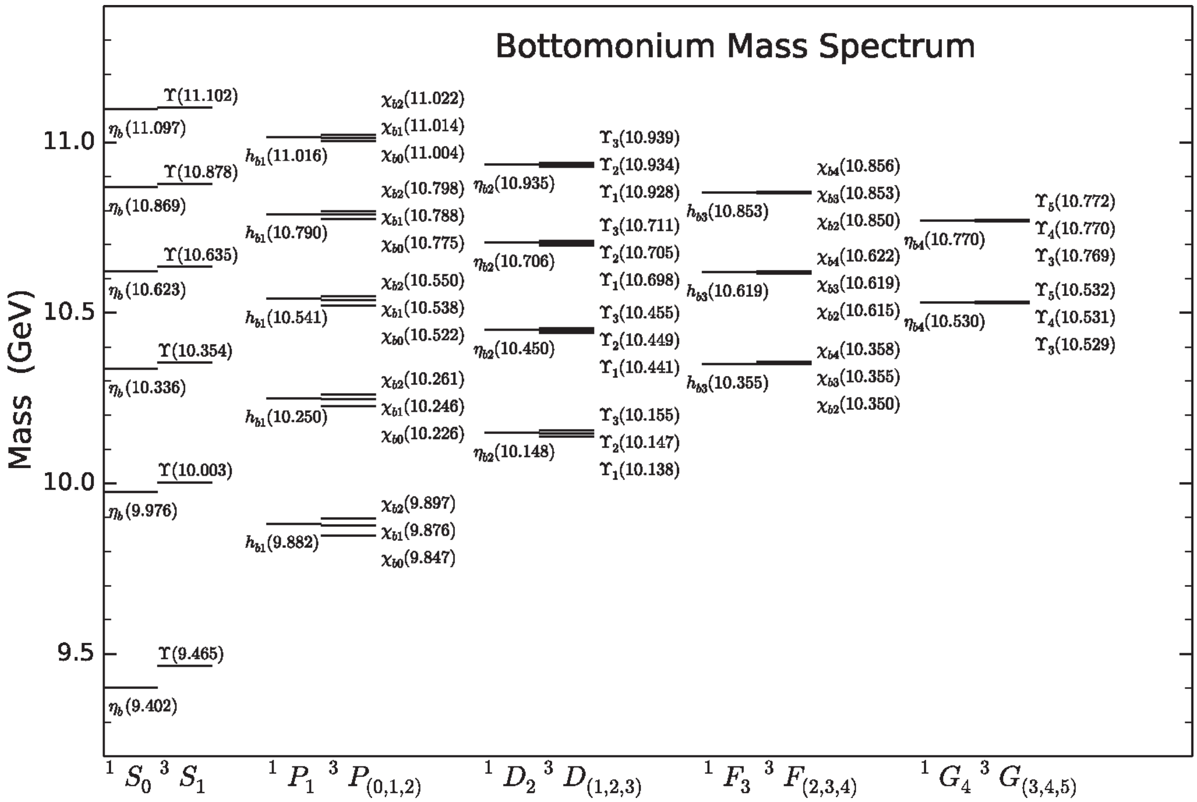


Figure 1: Pure (*i.e.* $b\bar{b}$) bottomonium mass spectrum [8] calculated using a relativized quark model [7].

quarks and a pair of antiquarks are each bound with a stronger force than the force holding diquark and antidiquark together. While the search described in this document is model-independent, our motivation is somewhat biased in favor of the molecular picture and has likely impacted our decisions about how to perform the analysis.

The main goal of out study is to test some of the predictions of the new spectroscopy [12] that predicts energy levels for the molecular bottomonium-like states depicted in Fig. 2, Namely, we describe a search for the partner states of Z_b, referred to as W_{bJ}, and we aim to obtain new information about hadronic dynamics in presence of the heavy b quarks. Improving the current understanding of such dynamics is of paramount importance for being able to use the hadronic decays of B mesons to extract possible contributions from the Beyond-the-Standard-Model (BSM) amplitudes, where the interplay between the strong interaction and the new BSM weak phases could not be reliably understood without the precise theoretical predictions for the QCD part.

1.3 Radiative Decays $\Upsilon(5S) \to \gamma W_{bJ}$

The Z_b states were discovered in single-pion transitions of $\Upsilon(5S)$ and $\Upsilon(6S)$, followed by another single-pion transition to the bottomonium states. According to molecular interpretation, $Z_b(10610)$ is primarily a $B\bar{B}^*$ state, while $Z_b(10650)$ (a.k.a. Z_b') is a $B^*\bar{B}^*$ state. Z_b are spin-1 isotriplets (both neutral and charged states were discovered in transitions

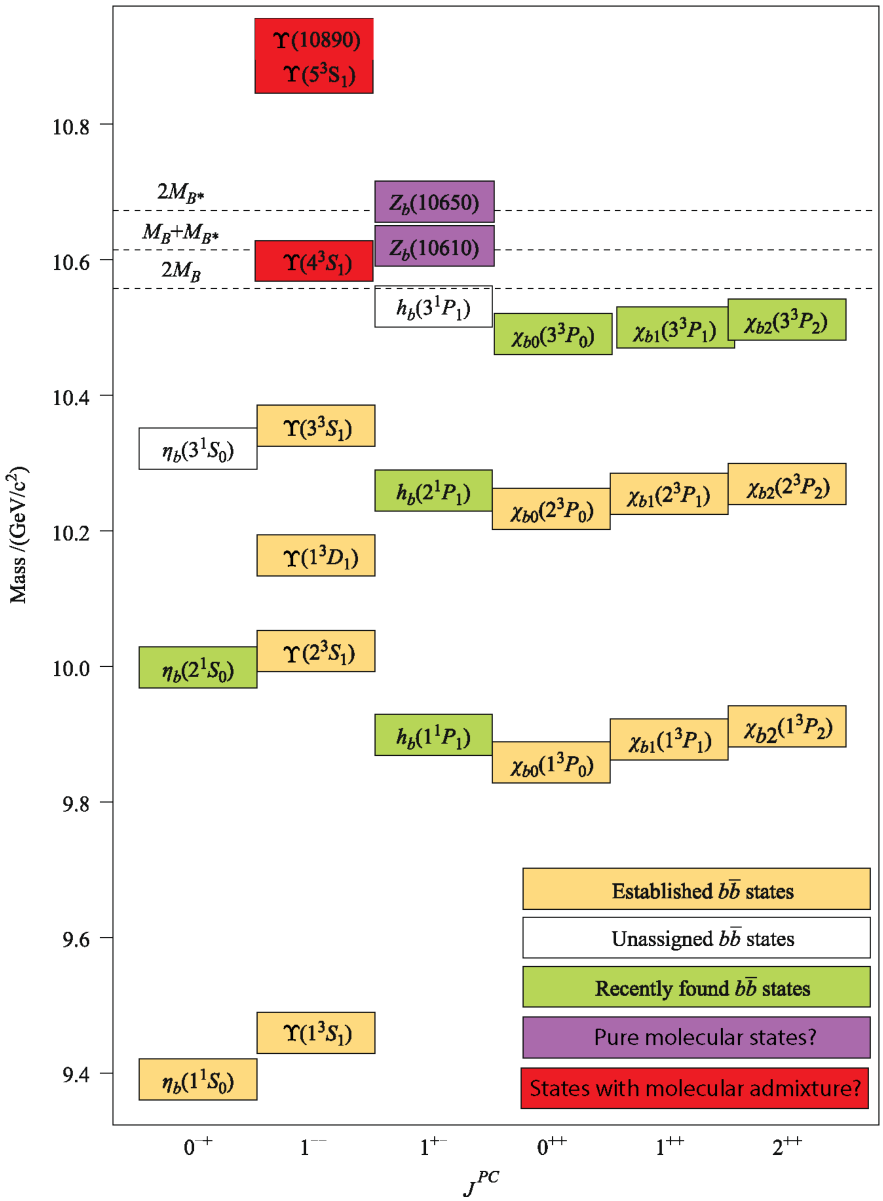


Figure 2: Most relevant (for our study) states in conventional bottomonium and bottomonium-like spectroscopies. We modified this figure from S. Olsen's review article [12]. Note that we took liberty to modify the original figure to better represent the contents of this Note, namely, we relabeled $\Upsilon(nS)$ $(n = 4, 5, 6)$ as "States with molecular admixture?" and Z_b states as "Pure molecular states?".

$\Upsilon(nS) \to \pi Z_b$ $(n = 5, 6)$. The hypothetical partners of positive G-parity states Z_b, *i.e.* the W_{bJ} states, would also be isotriplets but of negative G-parity (quantum numbers of the new molecular states are defined by quantum numbers of their partners in two-body decays of the $\Upsilon(5S)$ parent: while Z_b is accompanied by a pion, each W_{bj} is accompanied by a ρ meson (or a photon)). Therefore the W_{bJ} states are expected to appear in transitions $\Upsilon(nS) \to \rho W_{bJ}$. Conservation of angular momentum allows J in W_{bJ} to be 0, 1 or 2. Excited states such as W'_{b0} could exist as well. Quantum numbers assigned to Z_b and W_{bJ} states are summarized

$I^G(J^P)$	Name	Co-produced with (threshold, GeV/c^2)	Assumed composition	Decay channels
$1^+(1^+)$	$Z_b(10610)$	π (10.75)	$B\bar{B}^*$	$\Upsilon(nS)\pi, h_b(nP)\pi, \eta_b(nS)\rho$
$1^+(1^+)$	$Z_b'(10650)$	π (10.79)	$B^*\bar{B}^*$	$\Upsilon(nS)\pi, h_b(nP)\pi, \eta_b(nS)\rho$
$1^-(0^+)$	W_{b0}	ρ (11.34), γ (10.56)	$B\bar{B}$	$\Upsilon(nS)\rho, \eta_b(nS)\pi, \chi_b\pi$
$1^-(0^+)$	W_{b0}'	ρ (11.43), γ (10.65)	$B^*\bar{B}^*$	$\Upsilon(nS)\rho, \eta_b(nS)\pi, \chi_b\pi$
$1^-(1^+)$	W_{b1}	ρ (11.38), γ (10.61)	$B\bar{B}^*$	$\Upsilon(nS)\rho, \chi_b\pi$
$1^-(2^+)$	W_{b2}	ρ (11.43), γ (10.65)	$B^*\bar{B}^*$	$\Upsilon(nS)\rho, \chi_b\pi$

Table 1: Molecular isotriplet states which could be produced in the decays of $\Upsilon(5S)$ and $\Upsilon(6S)$ according to [10]. Note that the ρ could be replaced by a photon in the decays of $I_3 = 0$ states, but this would suppress the expected rate even more. Please see Fig. 3 as well.

in Table 1.

The $\Upsilon(5S)$ resonance does not have enough energy to allow the transition to W_{bJ} with sufficient amount of energy left for the two pions in the tail of the ρ invariant mass. In our analysis, instead of searching for decays with the ρ mesons, we have to allow for the $q\bar{q}$ annihilation and pay the price of approximately α_{em} in the branching fraction:

$$\frac{\Gamma(\Upsilon(5S) \to \gamma W_{bJ})}{\Gamma(\Upsilon(5S) \to Z_b\pi)} \sim \alpha_{\text{em}} \approx \frac{1}{137} \tag{1}$$

Therefore, we search for the transitions $\Upsilon(5S) \to \gamma W_{bJ}$. This indirect phase space limitation allows us to search only for the $I_3 = 0$ partners of the Z_b states, *i.e.* only the neutral component of each isotriplet can be found in such radiative transitions. We explain this strategy, suggested [13] by M.B. Voloshin, in Fig. 3.

To search for all new resonances expected in the new spectroscopy would require to collect a sizeable data sample at $\Upsilon(6S)$ or above its energy. Such possible future studies [14] at Belle II and many more interesting discussions (such as possible existence of isoscalar partners of Z_b and W_{bJ}) can be found elsewhere [10]. In the rest of this paper, we focus on the analysis of the full $\Upsilon(5S)$ data sample where we search for the decay $\Upsilon(5S) \to \gamma W_{bJ}$.

1.4 Expected Signal in Data

Belle previously reported [15] that charged Z_b states comprise approximately 2.54% of the 1819 $\Upsilon(1S)\pi^+\pi^-$ (followed by $\Upsilon(1S) \to \mu^+\mu^-$) events observed with the full data sample. The overall reconstruction efficiency in Z_b analysis was estimated to be around 46%. This allows us to estimate that, with an ideal, *i.e.* 100% efficient detector, we would expect to detect, approximately, 100 (charged) Z_b events.

While searching for W_{bj} events in radiative decays of $\Upsilon(5S)$, as elaborated in Section 1.3, we have to pay the price of α_{em}. Jumping a little bit ahead of ourselves, with our overall detection efficiency of 29%, we therefore expect to observe, on average, 0.2 W_{b0} events. This number, however, has a (hopefully very) large uncertainty, and, after all, we are (always!)

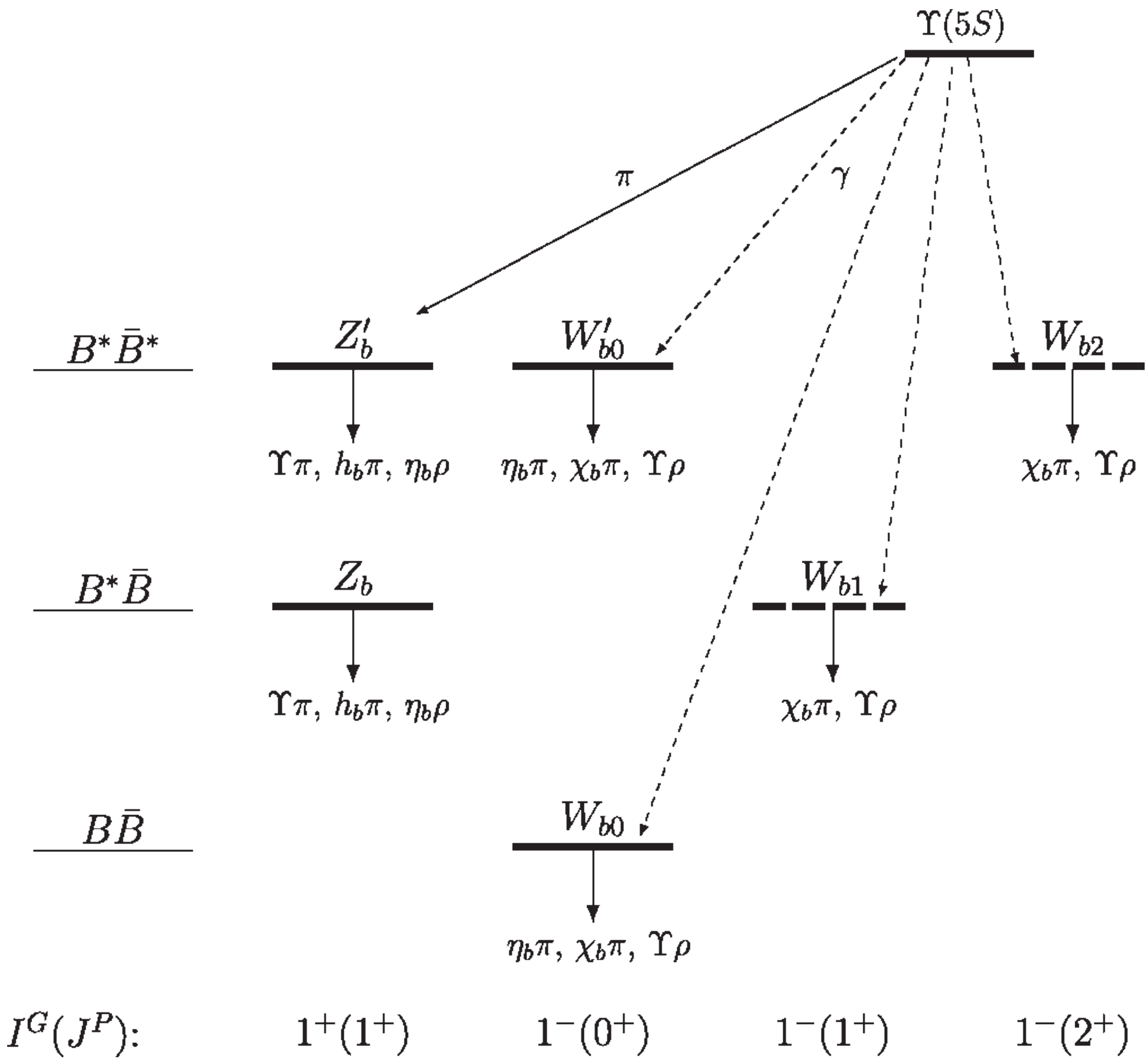


Figure 3: The expected family of isotriplet resonances from Ref. [13] (which the reader is advised to consult for relevant details). For $\Upsilon(6S)$ transitions, the photon is replaced by ρ. This would also allow to access charged W_{bJ} states. Also, please see Table 1.

driven by hope that nature might be kinder to us than we deserve. Also, tangentially, our LHC colleagues have been searching for signatures of SUSY for some time already, and, no matter how little has been observed so far, their noble quest will stop not. So why should we stop ours? On this philosophical note we conclude this discussion and proceed to describe our actual analysis.

2 Monte Carlo and Data Samples

To study the properties of signal events, we generate 100,000 Monte Carlo (MC) events for $\Upsilon(5S) \to \gamma W_{bJ}$ followed by $W_{bJ} \to \Upsilon(1S)\rho^0$, $\Upsilon(1S) \to \mu^+\mu^-$, $\rho^0 \to \pi^+\pi^-$ using MC generator EvtGen [16]. Detector response is simulated using GEANT4 [17]. W_{bJ} is generated with an intrinsic width of 15 MeV, similar to the widths of Z_b and Z_b'. Table 2 displays the decay models [18] used in MC simulation of signal processes. The PHOTOS package [19] is used to simulated final state radiation (FSR). To allow for softer FSR photons in simulation, we modified the PHOTOS package to lower the minimum energy of final state radiation. Please see Section 10.1 for details.

We use six streams of generic MC to study background events. Each stream is equivalent to a full Belle data sample of 121.4 fb^{-1} of $\Upsilon(5S)$ resonance data. We generate additional MC samples to study background events originating from $\Upsilon(5S) \to \Upsilon(1S)\pi^+\pi^- \to \mu^+\mu^-\pi^+\pi^-$ with initial state radiation (ISR) as well as events originating from $\Upsilon(5S) \to Z_b^{\pm}\pi^{\mp} \to \Upsilon(1S)\pi^{\pm}\pi^{\mp} \to \mu^+\mu^-\pi^{\pm}\pi^{\mp}$. We describe or studies of these processes in Section 6 and Section 7, respectively.

In this analysis, we use the full 121.4 fb^{-1} of on-resonance $\Upsilon(5S)$ data collected by the Belle detector at the KEKB collider from asymmetric energy e^+e^- collisions with $\sqrt{s} = 10.86$ GeV [20].

Decay Process	Decay Model used in Mote Carlo Simulation
$\Upsilon(5S) \to W_{bJ}\gamma$	VSP_PWAVE
$W_{bJ} \to \Upsilon(1S)\rho^0$	SVV_HELAMP
$\rho^0 \to \pi^+\pi^-$	VSS
$\Upsilon(1S) \to \mu^+\mu^-$	VLL
Final state radiation	PHOTOS (modified)

Table 2: EvtGen decay models used in Mote Carlo simulation of signal processes.

3 Selection Criteria

We reconstruct the decay mode $\Upsilon(5S) \to \gamma W_{bJ}$ followed by the decays $W_{bJ} \to \Upsilon(1S)\rho^0$, $\Upsilon(1S) \to \mu^+\mu^-$, $\rho^0 \to \pi^+\pi^-$. We select a fully-reconstructed final state particle combination consisting of $\pi^+\pi^-\mu^+\mu^-\gamma$.

3.1 Selection of Photon Candidates

We require reconstructed photons have energy between 100 and 600 MeV and polar angle between 17° and 150°. In the center of mass reference frame, the radiative photon is expected to be monochromatic with energy of approximately 300 MeV. To reject showers produced by neutral hadrons, we require $E_9/E_{25} > 0.75$, where the E_9/E_{25} ratio is defined as the energy summed in the 3 x 3 array of crystals surrounding the center of the shower (E_9) to that of the 5 x 5 array of crystals surrounding the center of the shower (E_{25}).

3.2 Selection of Pion and Muon Candidates

Pion candidates must satisfy $R_{K,\pi} < 0.9$, where $R_{K,\pi}$ is the "Kaon identification variable" defined as the likelihood ratio of the charged track to be due to a kaon versus a pion, and

Particle Candidate	Selection Criteria
γ	$100 <$ MeV $E(\gamma) < 600$ MeV
$\pi^{\pm}, \mu^{\pm}$	$dr < 0.3$ cm $\lvert dz \rvert < 2$ cm $p_T > 100$ MeV/c
$\pi^{\pm}$ PID	$R_{K,\pi} < 0.9$ $R_{e,hadron} < 0.9$
$\mu^{\pm}$	$R_{\mu} > 0.10$
ρ^0	0.420 GeV$/c^2 < M_{\pi^+\pi^-} < 1.020$ GeV$/c^2$
$\Upsilon(1S)$	9.3 GeV$/c^2 < M_{\mu^+\mu^-} < 9.6$ GeV$/c^2$
$\Upsilon(5S)$ (full event reconstruction)	10.2 GeV$/c^2 < M_{\pi^+\pi^-\mu^+\mu^-\gamma} < 11.5$ GeV$/c^2$ -0.05 GeV $< \Delta E < 0.03$ GeV Exactly four tracks: two muons and two pions

Table 3: Selection criteria for $\Upsilon(5S) \to \gamma W_{bJ}$

$R_{e,\mathrm{hadron}} < 0.9$, where $R_{e,\mathrm{hadron}}$ is the likelihood ratio of the charged track to be due to an electron versus a hadron. Similarly, muon candidates must satisfy $R_{\mu} > 0.1$, where R_{μ} is the likelihood ratio of the charged track to be due to a muon versus other particles detected by the KLM detector subsystem. After imposing the aforementioned requirements, we additionally require there to be four unique charged tracks – two pions and two muons. Events with more than four such tracks are rejected.

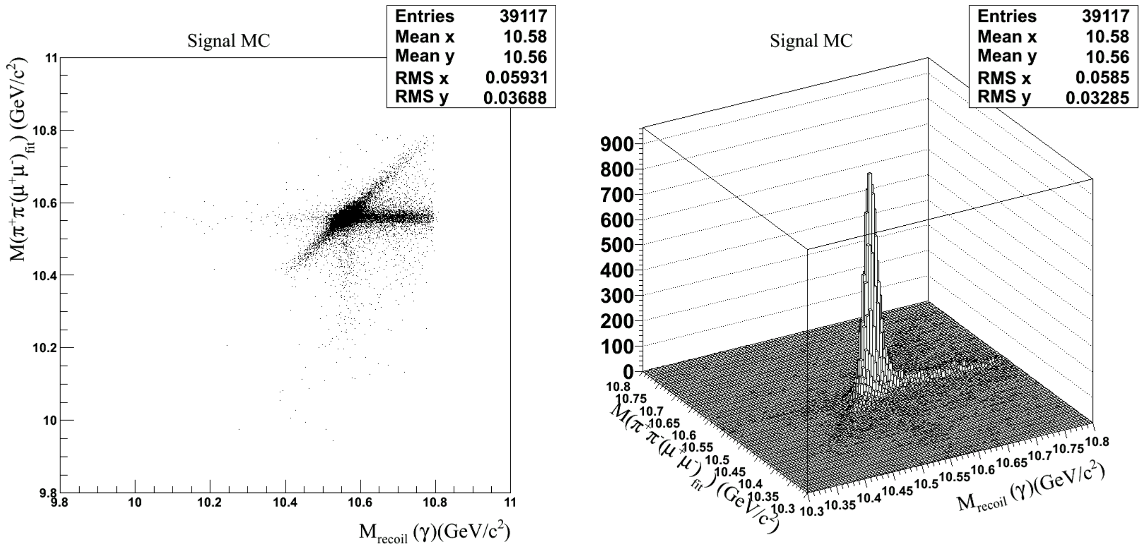
To select reconstructed track that originate near the interaction point, we require pion and muon candidates have $dr < 0.3$ cm and $|dz| < 2$ cm, where dr and dz are impact parameters in the radial and z directions, respectively. We also require pion and muon candidates to have transverse momenta $p_T > 100$ MeV. Candidate muon pairs must have an invariant mass between 9.3 GeV/c^2 and 9.6 GeV/c^2. Candidate pion pairs must have an invariant mass between 0.42 GeV/c^2 and 1.02 GeV/c^2.

3.3 Selection of $\Upsilon(5S)$ Candidates

$\Upsilon(5S)$ candidates are required to have an invariant mass between 10.2 GeV and 11.5 GeV. The muon pairs of selected $\Upsilon(5S)$ candidates are mass constrained to the nominal $\Upsilon(1S)$ invariant mass of 9.460 GeV/c^2. A summary of our selection criteria is shown in Table 3.

3.4 Best Candidate Selection

Approximately 32% of signal MC events satisfying our selection criteria have multiple signal candidates. This is exclusively due to relatively soft photons. In events with multiple signal candidates, we select the candidate that has an energy most consistent with the center of mass energy of the experimental run. The selected candidates are correctly MC-tagged to full



(a) (b)

Figure 4: $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ vs $M_{\rm rec}(\gamma)$ distribution for W_{B0} signal MC events. We show the lego plot in Fig. 4b to emphasize that the tail of $M_{\rm rec}(\gamma)$ is not as large as it appears in Fig. 4a. Note that Fig. 4b is plotted in a smaller range.

Quantity	Value
Intrinsic width of W_{bJ}	15 MeV/c^2
Charged track resolution	4 MeV
Photon energy resolution	8 MeV
Beam energy resolution	6 MeV

Table 4: Quantities contributing to widths of measured quantities

MC truth for signal 90% of the time. For fully reconstructed signal MC events with multiple candidates, our best candidate selection method selects a candidate correctly MC-tagged to full MC truth 88% of the time.

4 Signal Monte Carlo Studies

4.1 Signal Monte Carlo Distributions

To understand properties of signal events, we investigate two invariant mass variables, $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ and $M_{\rm rec}(\gamma)$, where subscript "fit" indicates that the muon pair is constrained to the nominal mass of $\Upsilon(1S)$. We define the invariant mass recoiling against X as

$$M_{\rm rec}(X) = \sqrt{(E_{\rm cm}(\exp) - E_{\rm cm}(X))^2 - |\vec{0} - \vec{p}_{\rm cm}(X)|^2} \qquad (2)$$

where $E_{\rm cm}(\exp)$ is the run's average energy, and $E_{\rm cm}(X)$ and $\vec{p}_{\rm cm}(X)$ are the energy and momentum of system X. Subscript "cm" is used for quantities evaluated in the center of mass reference frame of the experiment. For signal events, $M_{\rm recoil}(\gamma)$ and $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ are two independent ways to estimate the invariant mass of W_{bJ}. Fully reconstructed signal events fall along the main diagonal of the $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ vs $M_{\rm rec}(\gamma)$ plot shown in Fig. 4. We define energy balance ΔE as

$$\Delta E = E_{\rm cm}(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit}\gamma) - E_{\rm cm}(\exp). \tag{3}$$

ΔE is the most important variable we can use to select fully reconstructed signal event candidates.

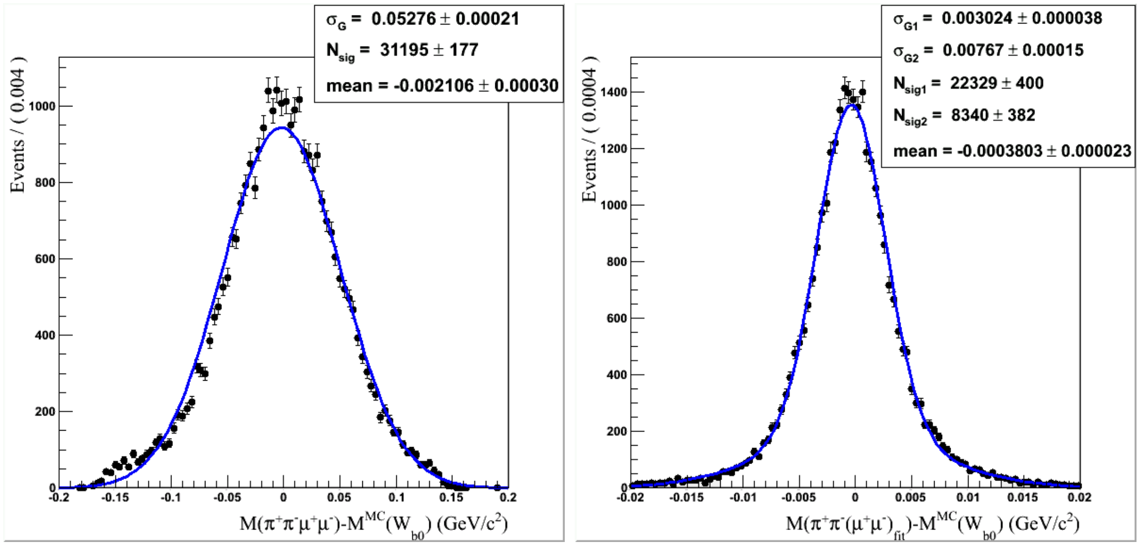
There are two effects contributing to the observed width of $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$: (1) the intrinsic width of W_{bJ}, and (2) the charged track reconstruction. Fig. 5 shows $M(\pi^+\pi^+\mu^+\mu^-)$ and $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ resolutions for signal events within the signal region and sideband regions (defined in Section 4.2). We model both resolutions as the sum of two Gaussians with the same mean and fit both resolutions. Contribution to $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ resolution from charged track reconstruction is primarily due to pions, since muon pairs are constrained to $\Upsilon(1S)$ invariant mass.

The distribution of $M_{\rm rec}(\gamma)$ has a long tail due to an underestimation of photon energy, causing an overestimation of $M_{\rm rec}(\gamma)$. Effects contributing to the observed width of $M_{\rm rec}(\gamma)$ include (1) intrinsic width of W_{bJ}, and (2) photon energy resolution. $M_{\rm rec}(\gamma)$ resolution is dominated by photon energy resolution.

Effects contributing to the observed shape of ΔE include (1) photon energy resolution, (2) charged track resolution, (3) beam energy resolution, and (4) the intrinsic width of W_{bJ}. ΔE resolution is dominated by photon energy resolution as well. The values of relevant widths are listed in Table 4. In signal MC we observe $\sigma_{\Delta E} \approx 12$ MeV, so we take advantage of this excellent energy resolution to select fully reconstructed events. Because the distribution of ΔE is asymmetric (primarily due to leakage from the calorimeter and relatively soft non-signal photons in signal events), we use an asymmetric selection and require $-0.05\,\text{GeV} \le \Delta E \le 0.03\,\text{GeV}$. This selection cuts out the long tail in the distriubtion of $M_{\rm rec}(\gamma)$ and reduces the efficiency by 20%. Note, however, that this selection primarily removes events where the signal photon is not reconstructed. After applying this selection on ΔE, signal reconstruction efficiency becomes approximately 31%. Fig. 6 displays ΔE resolution as well as quantities contributing to ΔE resolution.

4.2 Description of the Signal Region

Table 5 contains the definitions of four important regions in this analysis. Before investigating data, we blind the region where we expect to find signal. We refer to this region as the blinded region. The invariant masses of W_{b0}, W_{b1}, and W'_{b0} and W_{b2} are expected to be at the $B\overline{B}$, $B^*\overline{B}$, and $B^*\overline{B^*}$ thresholds, respectively. The blinded region is defined as the region between the $B\overline{B}$ and $B^*\overline{B^*}$ thresholds plus an additional margin of 70 MeV on either side. This corresponds to $10.49\,\text{GeV/c}^2 \le M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit}) \le 10.72\,\text{GeV/c}^2$. The boundary on the left side of the region is defined by the sloped line $M_{\rm rec}(\gamma) \ge M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit}) - 0.04$ GeV/c^2 which lies parallel to the main diagonal. Approximately 20% of signal events are



(a) $M(\pi^+\pi^+\mu^+\mu^-)$ resolution. Note that muons are not mass constrained.

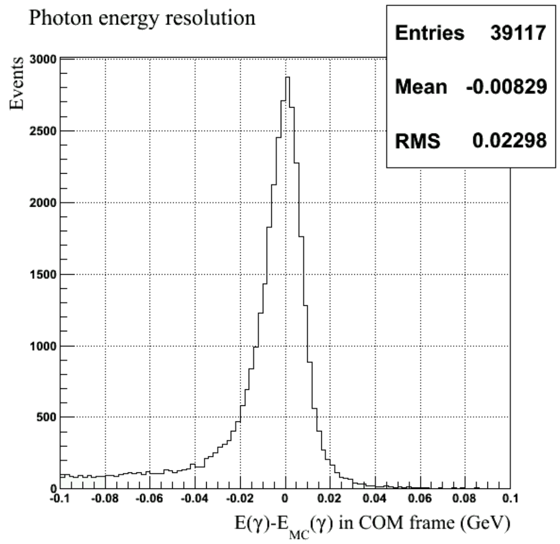
(b) $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ resolution (muons are mass constrained).

Figure 5: $M(\pi^+\pi^+\mu^+\mu^-)$ and $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ resolutions for signal events within the signal region and sideband regions (defined in Section 4.2). Note that the horizontal scales are different.

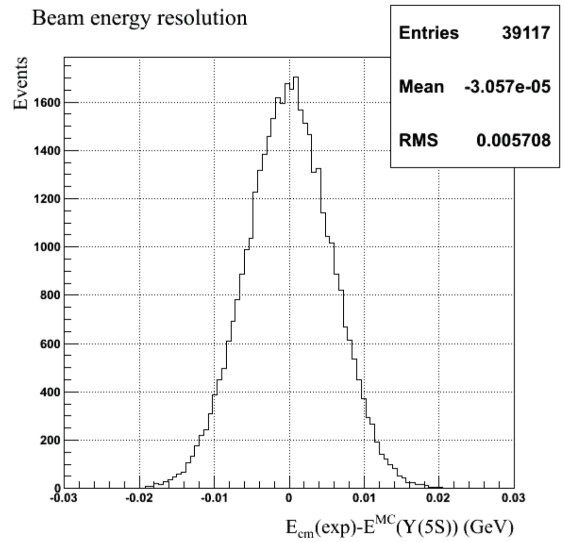
Region Name	Boundary Definitions
Blinded Region	$10.49\,\text{GeV/c}^2 \leq M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit}) \leq 10.72\,\text{GeV/c}^2$ $M_{\rm rec}(\gamma) \geq M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit}) - 0.04\ \text{GeV/c}^2$ $M_{\rm rec}(\gamma) \leq 10.8\ \text{GeV/c}^2$
Signal Region	$10.49\,\text{GeV/c}^2 \leq M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit}) \leq 10.72\,\text{GeV/c}^2$ $-0.05\,\text{GeV} \leq \Delta E \leq 0.03\,\text{GeV}$
Sideband Region	$10.38\,\text{GeV/c}^2 \leq M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit}) \leq 10.49\,\text{GeV/c}^2$ $10.72\,\text{GeV/c}^2 \leq M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit}) \leq 10.80\,\text{GeV/c}^2$ $-0.05\,\text{GeV} \leq \Delta E \leq 0.03\,\text{GeV}$.
Grand Sideband Region	$10.38\,\text{GeV/c}^2 \leq M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit}) \leq 10.80\,\text{GeV/c}^2$ $-0.20\,\text{GeV} \leq \Delta E \leq 0.20\,\text{GeV}$

Table 5: Definitions of the signal region and other important regions.

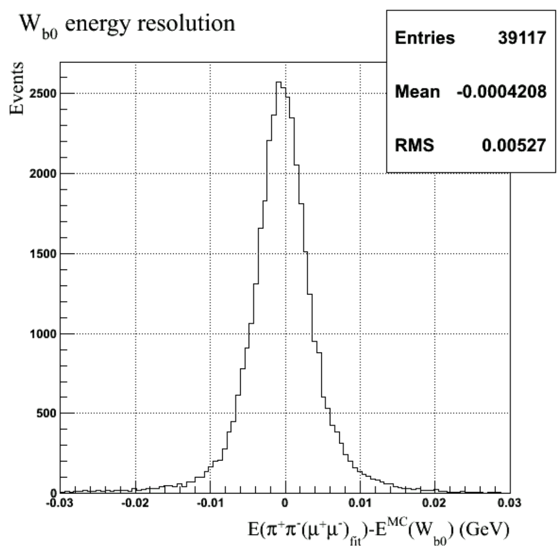
located in the long right tail of the distribution of $M_{\rm rec}(\gamma)$. A phase space boundary on the right side of the plot at $M_{\rm rec}(\gamma) \approx 10.75\ \text{GeV/c}^2$ forces this long tail of the $M_{\rm rec}(\gamma)$ distribution into a smaller region for the higher mass W_{bJ} states. Hence, we do not define a sloped boundary line as the right side of the signal region – a diagonal boundary would exclude more signal events for the lower mass states because of the aforementioned phase space boundary compressing the tail. Instead, we define the vertical line boundary $M_{\rm rec}(\gamma) \leq$



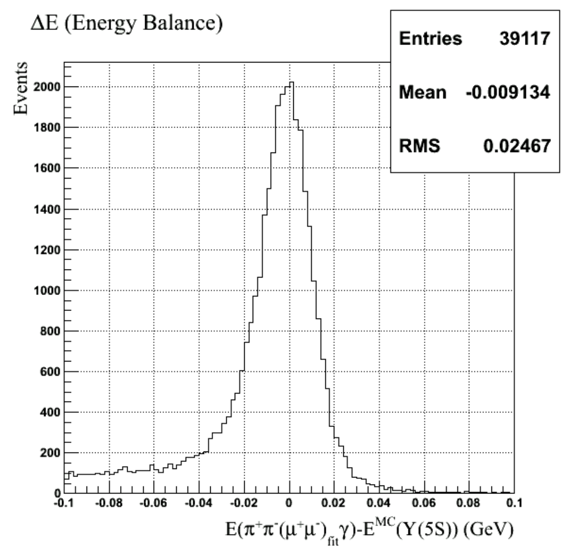
(a) Signal photon energy line shape in the COM reference frame.



(b) Beam energy resolution.



(c) $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ energy line shape (includes the effect of intrinsic W_{bJ} width and charged track reconstruction).



(d) Signal candidate energy line shape. Includes the effects of W_{bJ} intrinsic width and resolution.

Figure 6: ΔE resolution and quantities contributing to ΔE resolution.

$10.72\,\mathrm{GeV/c^2}$ which assures that approximately equal percentages of signal would be blinded for all masses of W_{bJ} states.

We define the signal region as the region contained within $10.49\,\mathrm{GeV/c^2} \leq M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit}) \leq 10.72\,\mathrm{GeV/c^2}$ satisfying $-0.05\,(\mathrm{GeV}) \leq \Delta E \leq 0.03\,\mathrm{GeV}$. The ΔE requirement selects only fully-reconstructed signal events, where signal is peaking.

The sideband region is essentially an extension of the signal region, defined as the regions within $10.38\,\mathrm{GeV/c^2} \leq M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit}) \leq 10.49\,\mathrm{GeV/c^2}$ and $10.72\,\mathrm{GeV/c^2} \leq$

$M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit}) \leq 10.80\,{\rm GeV/c^2}$ satisfying $-0.05\,({\rm GeV}) \leq \Delta E \leq 0.03\,{\rm GeV}$.

We additionally define the grand sideband region as the region within $10.38\,{\rm GeV/c^2} \leq M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit}) \leq 10.80\,{\rm GeV/c^2}$ satisfying $-0.20\,{\rm GeV} \leq \Delta E \leq 0.20\,{\rm GeV}$. This region is used when studying background in data.

Fig. 7 displays these four regions with our three signal MC samples. It is important to note that the blinded region is not completely contained within the grand sideband region and the signal region is not completely contained within the blinded region. This is due to historical reasons, as the blinded region was defined prior to the use of ΔE in this analysis.

4.3 Trigger Simulation

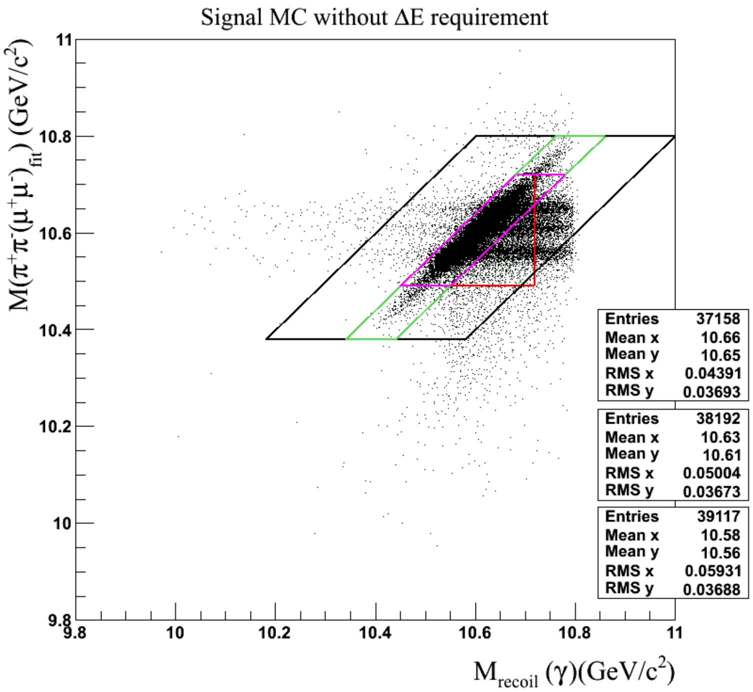
Relatively low final state particle multiplicity of our signal events requires us to investigate trigger efficiency. Trigger efficiency is simulated after full reconstruction. We find correlations between trigger efficiency and kinematics. Fig. 8 shows various 2-dimensional distributions of $\mu^+\cos(\theta)$ vs $\mu^+\cos(\theta)$, and we see that events failing to satisfy trigger are more likely to have one of the muons at a small angle with respect to the beam axis ($|\cos(\theta)| \geq 0.8$). Fig. 9 shows additional distributions of $\mu^+\cos(\theta)$ vs $\mu^+\cos(\theta)$ which we use to determine trigger efficiencies. When neither muon is at a small angle with respect to the beam axis, trigger efficiency is 96%. When one of the muons is at a small angle with respect to the beam axis, trigger efficiency drops to 89%. For all generated signal MC events, trigger efficiency is approximately 94%. After accounting for trigger efficiency, our overall efficiency drops from 31% to 29%.

5 Background Studies

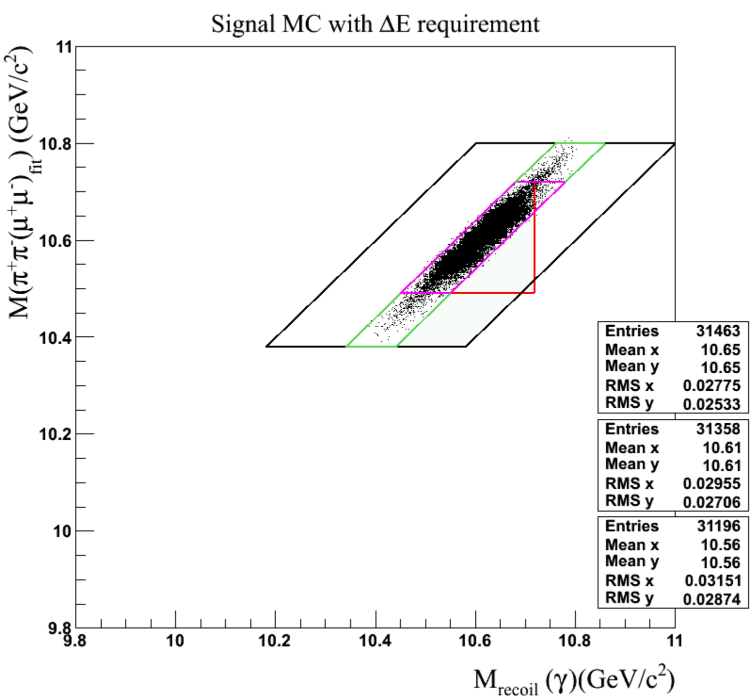
5.1 Generic Monte Carlo and Blinded Data

Fig. 10 shows the distribution of $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ vs $M_{\rm rec}(\gamma)$ for generic MC and blinded data events. Using MC truth, we identify the background decays in generic MC and blinded data and group them into eight categories which are defined in Table 6. No uds, charm, or B_sB_s generic MC events pass our selection criteria. A large number of non-B_sB_s events do satisfy our selection criteria, though they fall primarily outside the signal region. The ΔE requirement excludes most of these background events. The most prominent non-B_sB_s background sources are (cascade) dipion transitions to $\Upsilon(1S)$. We observe an enhancement in generic MC within the blinded region due to the decay $\Upsilon(5S) \to \Upsilon(2S)\pi^+\pi^-, \Upsilon(2S) \to \Upsilon(1S)\pi^+\pi^-$ where the selected signal pion candidates did not come from the same parent. The enhancement is removed when the ΔE constraint is applied, as such background events are not fully reconstructed.

We observe several regions where data events are clustering but generic MC events are not, and we have identified the likely origins of these events. The regions labeled X and Z in Fig. 10 are populated by events which are due to radiative returns to a lower mass $\Upsilon(nS)$ where the radiative photon is selected as our signal photon candidate. These events are fully reconstructed, and thus fall along the main diagonal of the plot. The region labeled Y includes processes involving radiative decays of $\chi_{bJ}(1P)$. These events have additional



(a) Does not include ΔE requirement.



(b) Includes ΔE requirement.

Figure 7: The blinded region (red), signal region (magenta), sideband region (green), and the grand sideband region (black). The plot in 7a includes the aforementioned ΔE requirement, while the plot in 7b does not. From top to bottom, the statistics boxes correspond to W'_{b0}, W_{b1}, and W_{b0} signal MC, respectively.

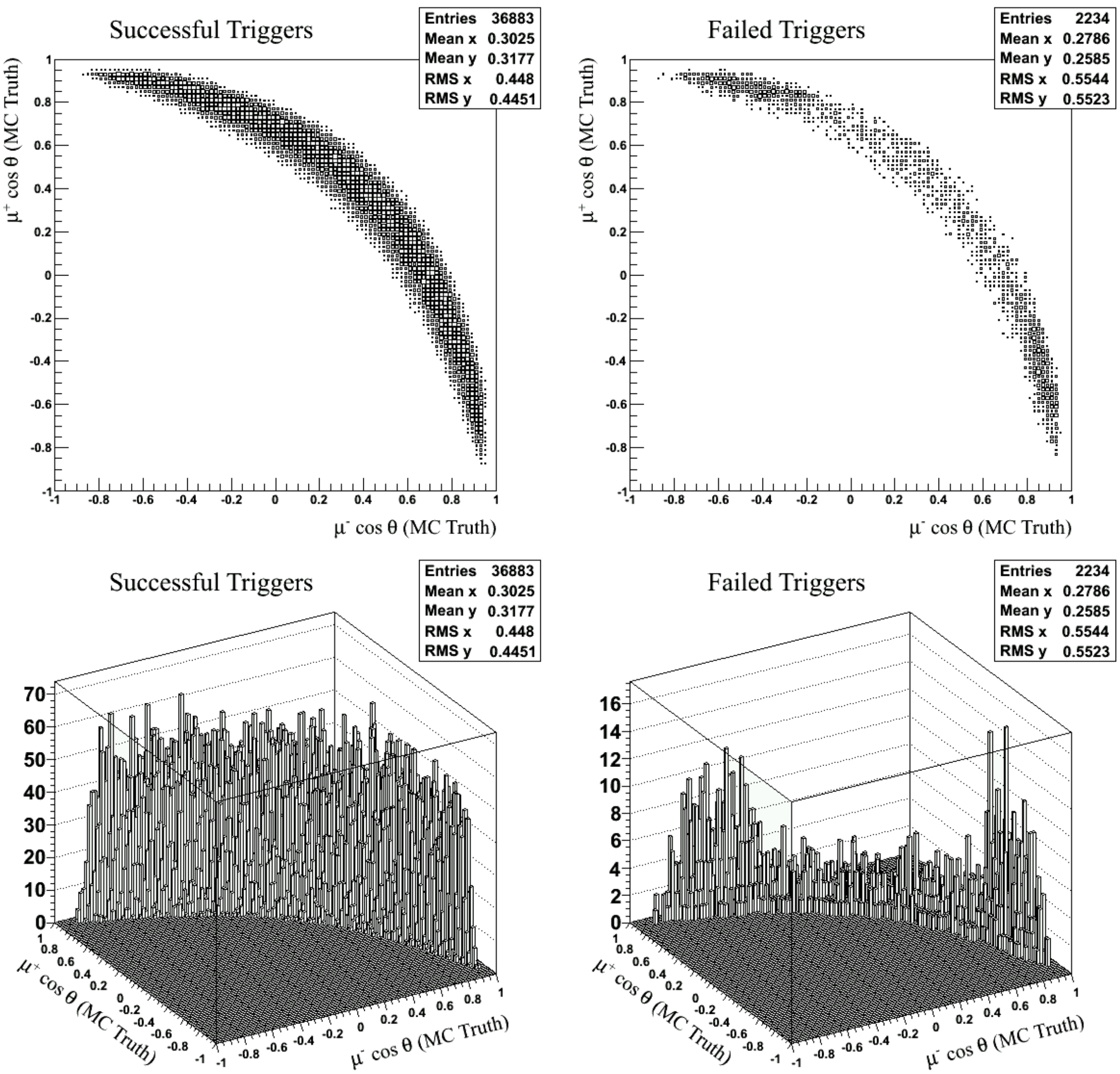


Figure 8: Reconstructed signal MC events that satisfy the offline trigger selection are plotted on the left, while events that fail the offline trigger selection are plotted on the right. We observe that events satisfying the trigger criteria are distributed more or less uniformly for kinematically allowed muons, but events failing to satisfy trigger are more likely to have one of the muons at a small angle with respect to the beam axis.

final state particles that are not reconstructed, and hence they fall below the main diagonal where $\Delta E < 0$. Events in categories X, Y, and Z are not of concern to us, since they are located far from the signal region.

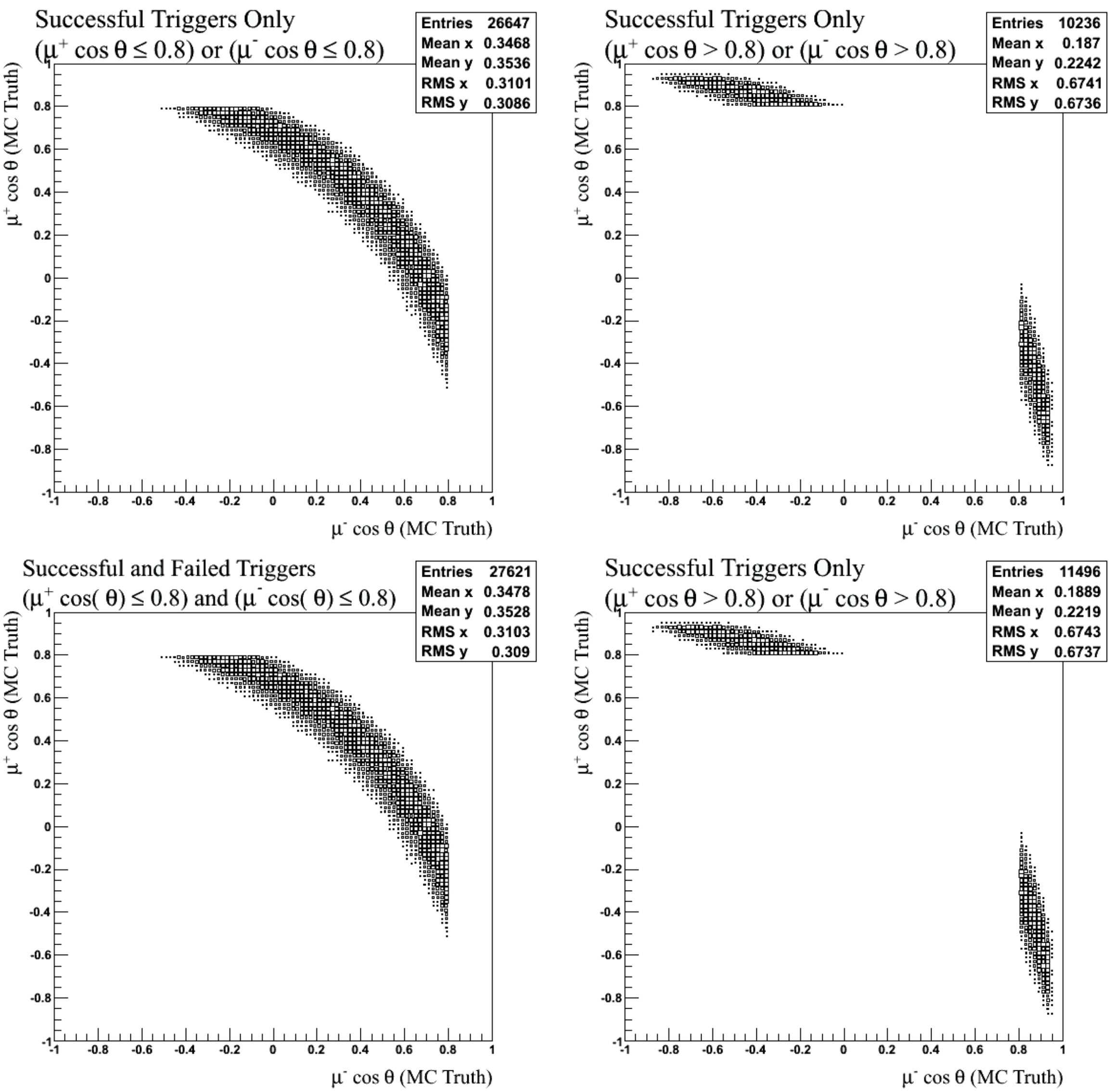


Figure 9: All reconstructed events in which both muons are generated with $|\cos(\theta)| < 0.8$ are plotted in the left two figures. Trigger efficiency for such events is approximately $(96 \pm 4)\%$. In the right two figures, we plot all reconstructed events where one of the muons is generated with $|\cos(\theta)| > 0.8$. Trigger efficiency for these events is reduced to about $(89 \pm 4)\%$.

6 Background from $\Upsilon(5S) \to \Upsilon(1S)\pi^+\pi^-$ with Initial State Radiation (ISR)

We find that dipion transitions to $\Upsilon(1S)$ (labeled 'A' in Fig. 10) have a much longer tail in data than in generic MC. This difference is shown in Fig. 11, and is due to initial state radiation (ISR). This tail contaminates the signal region, so we generate additional MC samples with ISR to study these backgrounds.

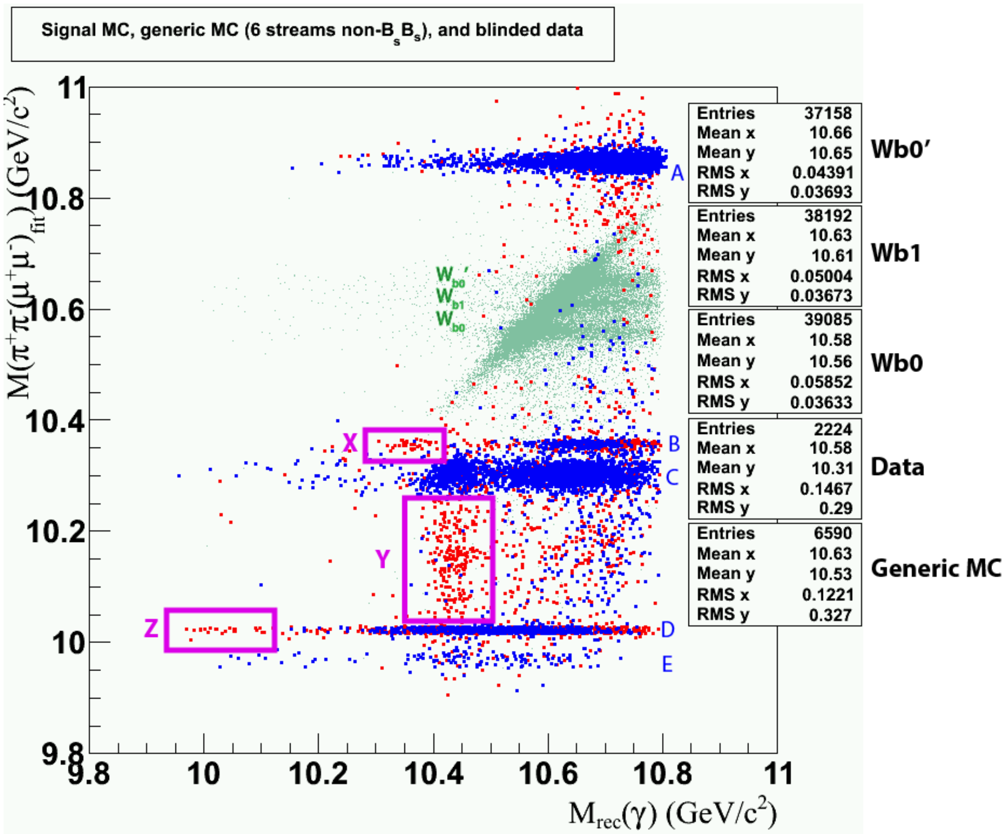


Figure 10: W_{b0}, W_{b1}, and W'_{b0} signal MC (light green), six streams of non-B_sB_s generic MC (blue), and data with the signal region blinded (red).

Label	Background
A	$\Upsilon(5S) \to \Upsilon(1S)\pi^+\pi^- \to \mu^+\mu^-\pi^+\pi^-$
B	$\Upsilon(5S) \to \Upsilon(3S)\pi^+\pi^- \to \Upsilon(1S)\pi^+\pi^-\pi^+\pi^- \to \mu^+\mu^-\pi^+\pi^-\pi^+\pi^-$ $\Upsilon(5S) \to \Upsilon(3S)\pi^0\pi^0 \to \Upsilon(1S)\pi^+\pi^-\pi^0\pi^0 \to \mu^+\mu^-\pi^+\pi^-\pi^0\pi^0$
C	$\Upsilon(5S) \to \Upsilon(2S)\pi^+\pi^- \to \Upsilon(1S)\pi^+\pi^-\pi^+\pi^- \to \mu^+\mu^-\pi^+\pi^-\pi^+\pi^-$ $\Upsilon(5S) \to \Upsilon(2S)\pi^+\pi^- \to \Upsilon(1S)\pi^0\pi^0\pi^+\pi^- \to \mu^+\mu^-\pi^0\pi^0\pi^+\pi^-$
D	$\Upsilon(5S) \to \Upsilon(2S)\pi^0\pi^0 \to \Upsilon(1S)\pi^+\pi^-\pi^0\pi^0 \to \mu^+\mu^-\pi^+\pi^-\pi^0\pi^0$
E	$\Upsilon(5S) \to \Upsilon(3S)\pi^+\pi^- \to \Upsilon(1S)\pi^0\pi^0\pi^+\pi^- \to \mu^+\mu^-\pi^0\pi^0\pi^+\pi^-$
X	$e^+e^- \to \Upsilon(3S)\gamma \to \Upsilon(1S)\pi^+\pi^-\gamma \to \mu^+\mu^-\pi^+\pi^-\gamma$
Y	Various processes involving $\chi_{bJ}(1P) \to \gamma\Upsilon(1S)$, *e.g.* $\Upsilon(5S) \to \Upsilon(1D)\pi^+\pi^-$, where $\Upsilon(1D) \to \gamma\chi_{bJ}(1P)$
Z	$e^+e^- \to \Upsilon(2S)\gamma \to \Upsilon(1S)\pi^+\pi^-\gamma \to \mu^+\mu^-\pi^+\pi^-\gamma$

Table 6: Backgrounds labeled in Fig. 10.

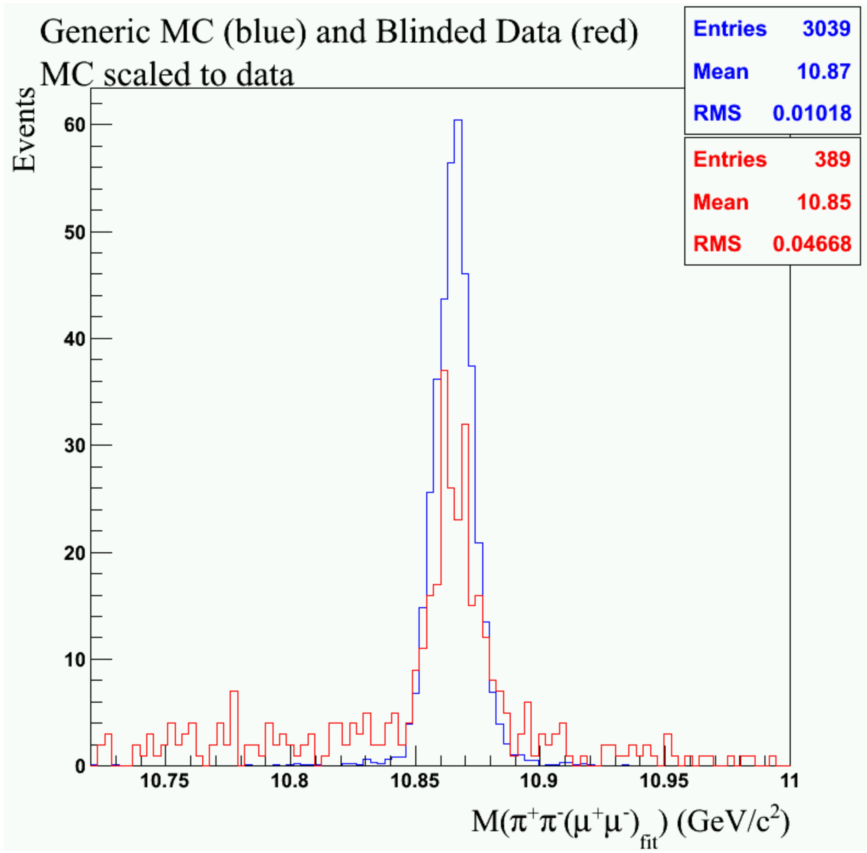


Figure 11: $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ distributions for $\Upsilon(5S) \to \Upsilon(1S)\pi^+\pi^-$ events (label 'A' in Table 6). Distributions for generic MC and blinded data are shown in blue and red, respectively. Generic MC does not include ISR and is normalized to the number of data events shown in the plotted range. We choose 10.72 GeV/c^2 as the lower limit of the range plotted, since lower masses would include the blinded region.

6.1 $\Upsilon(5S) \to \Upsilon(1S)\pi^+\pi^-$ ISR Monte Carlo Sample

The VectorISR model [18] is used to simulate ISR. We reweight the ISR photon energy spectrum according to the correct radiator function up to order α^2 [21] using a Monte Carlo method. After reweighting, there are approximately 110,000 events in our MC sample. A distribution of the reweighted ISR spectrum is shown in Fig. 12.

Fig. 13 shows the $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ vs $M_{\text{rec}}(\gamma)$ distribution for reweighted $\Upsilon(5S) \to \Upsilon(1S)\pi^+\pi^-$ events with ISR. Recall that the two plotted variables represent two independent ways to estimate the invariant mass of W_{bJ}, and therefore fully reconstructed events fall along the main diagonal of this plot. When the ISR photon of these backgrounds is selected as the signal photon candidate, these backgrounds are also fully reconstructed and fall along the main diagonal within the signal region. Approximately 3% of reconstructed events fall in the signal region. Fortunately, these backgrounds do not peak in the signal region in the distrubtion of $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$.

We simulate $\Upsilon(5S) \to \Upsilon(1S)\pi^+\pi^-$ with ISR using the models listed in Table 7. To determine if the choice of decay models affects the distribution shape of our signal variable

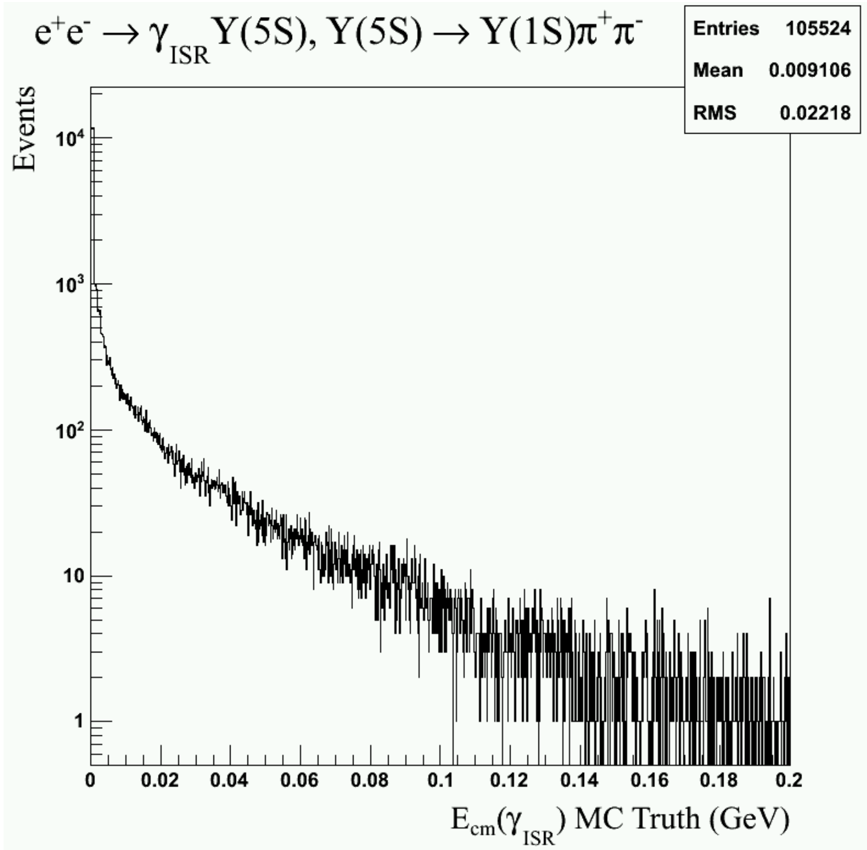


Figure 12: Reweighted ISR energy spectrum for $e^+e^- \to \gamma_{\rm ISR}\Upsilon(5S), \Upsilon(5S) \to \Upsilon(1S)\pi^+\pi^-$. Note that a log scale is used for the vertical axis.

$M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$, we generate additional samples using the VVPIPI decay [18] model for $\Upsilon(5S) \to \Upsilon(1S)\pi^+\pi^-$ and the VLL decay model [18] for $\Upsilon(1S) \to \mu^+\mu^-$. Fig. 14 shows the distribution of $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ for two different MC samples generated using different decay models.

We find that the choice of decay model has only a small effect on the shape of the $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ distribution. Furthermore, we plot the $\cos\theta$ of μ^+ in Fig. 15 and find that the presence of ISR has only a small effect on the the angular distriubtions of muons. To determine if ISR affects the width of the $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ distribution for signal processes $\Upsilon(5S) \to \gamma W_{bJ}$, we generate additional MC samples for the the signal process $\Upsilon(5S) \to \gamma W_{bJ}$ with ISR. We find that ISR has practically no effect on the width of the distribution of $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$.

6.2 Background Shape of $\Upsilon(5S) \to \Upsilon(1S)\pi^+\pi^-$ with ISR

It is likely that events due to $\Upsilon(5S) \to \Upsilon(1S)\pi^+\pi^-$ with ISR are a dominant source of backgrounds in the signal region. The rightmost plot in Fig. 16 shows the distribution of these events within the signal region for our reweighted MC. To see how the selection on ΔE affects the background shape, we loosen up the selection on ΔE in the left and middle plots in Fig. 16. Imposing a selection on ΔE has only a small effect on the shape of these

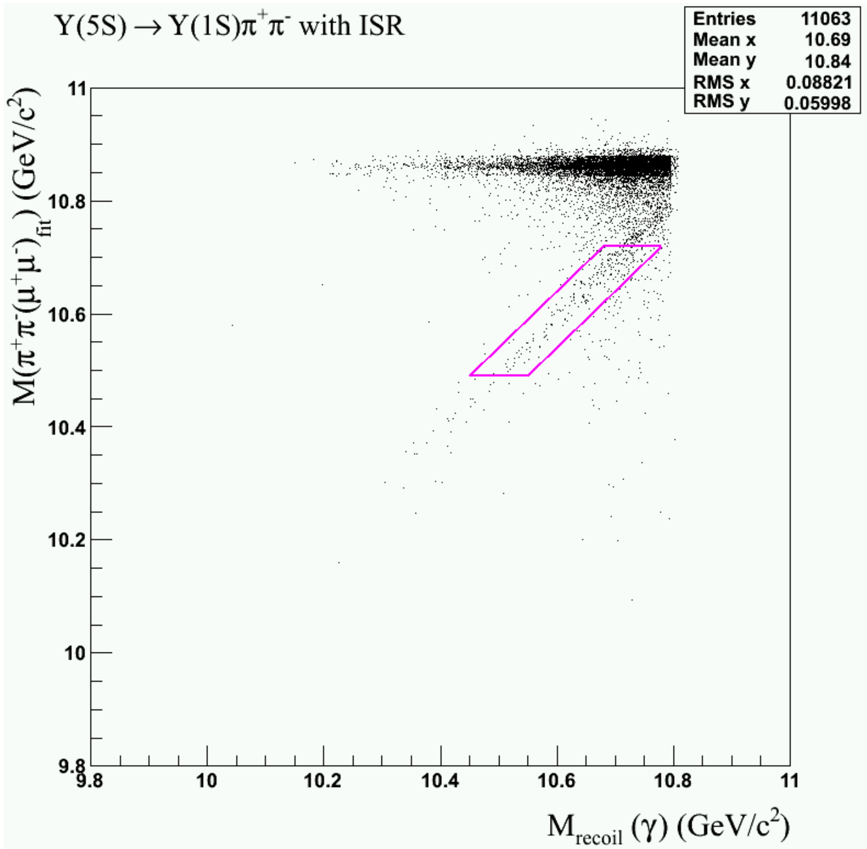


Figure 13: A 2-dimensional $M(\pi^+\pi^-(\mu^+\mu^-)_{\mathrm{fit}})$ vs $M_{\mathrm{rec}}(\gamma)$ distribution for $\Upsilon(5S) \to \Upsilon(1S)\pi^+\pi^-$ events with ISR (after reweighting). The signal region is outlined in magenta.

Decay Process	Decay Model used in Mote Carlo Simulation
$\Upsilon(5S) \to \Upsilon(1S)\pi^+\pi^-$	PHSP
$\Upsilon(1S) \to \mu^+\mu^-$	PHSP
Initial state radiation	VectorISR
Final state radiation	PHOTOS

Table 7: Decay models used in Mote Carlo simulation of $\Upsilon(5S) \to \Upsilon(1S)\pi^+\pi^-$ with ISR.

backgrounds in the signal region.

To determine if we can use this MC sample to estimate the number of background events in the signal region, we divide the grand sideband region shown in Fig. 17 into four smaller regions as defined in Table 8 and observe if the number of events in MC scales uniformly to data across all regions. Table 9 shows the number of ISR MC events and data events within the regions of interest. We see that ISR MC does not scale uniformly across all regions. While ISR studies improve the quality of our analysis and provide us with useful information about the shape of this background in the signal region, including ISR into our analysis does not sufficiently improve the scaling between data and MC in different regions

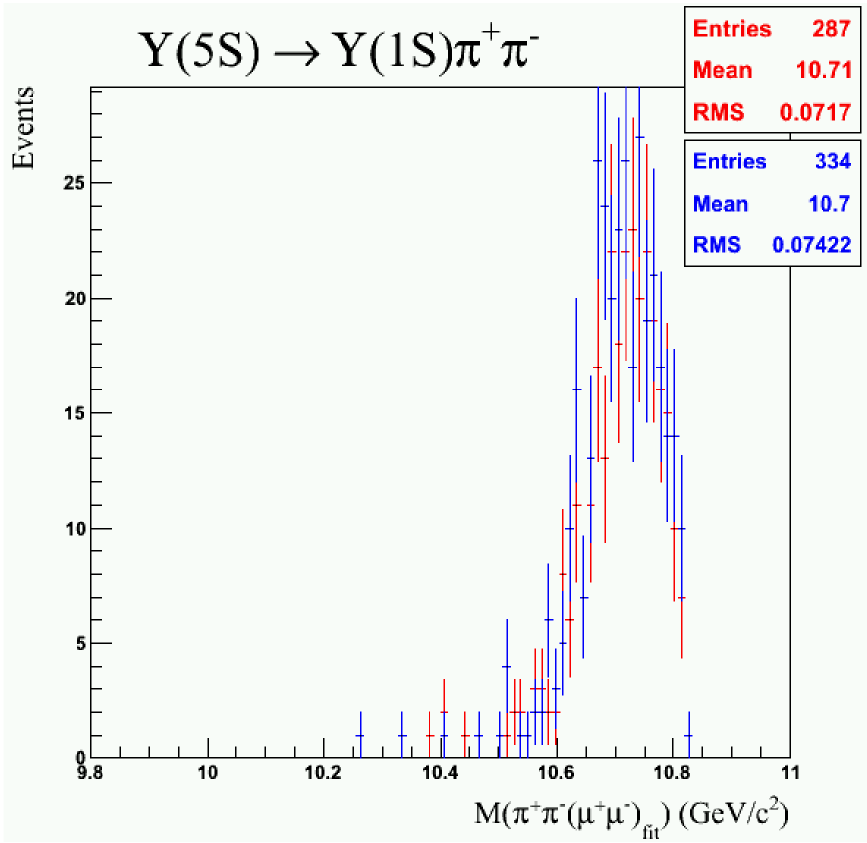


Figure 14: The distribution shown in blue is for events where $\Upsilon(5S) \to \Upsilon(1S)\pi^+\pi^-$ is generated using VVPIPI model [18] and $\Upsilon(1S) \to \mu^+\mu^-$ using VLL model [18]. The distribution shown in red is for events generated using PHSP model [18] for both processes. Neither samples contain ISR nor FSR, so they only differ by their decay models. The shapes of their $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ distributions are very similar. Note that although there is a difference in efficiency between the two samples, this is unimportant for our analysis, because we are only interested in possible difference between the shapes of these distributions.

of grand sideband.

7 Contribution from $\Upsilon(5S) \to Z_b^{(\prime)\pm}\pi^\mp$

Belle previously reported [15] that charged Z_b and Z_b' states comprise, respectively, approximately 2.54% and 1.04% of the 1819 $\Upsilon(1S)\pi^+\pi^-$ (followed by $\Upsilon(1S) \to \mu^+\mu^-$) events observed with the full data sample. The overall reconstruction efficiency in Z_b analysis was estimated to be around 46%. This allows us to estimate that, with an ideal, *i.e.* 100% efficient detector, we would expect to detect, approximately, 100 Z_b and 41 Z_b' events.

To estimate cross-feed between Z_b and W_{bj} analyses, we generated approximately 50,000 events for $\Upsilon(5S) \to Z_b^\pm\pi^\mp$ followed by $Z_b^\pm \to \Upsilon(1S)\pi^\mp$, $\Upsilon(1S) \to \mu^+\mu^-$. We also generated an additional 50,000 events for $\Upsilon(5S) \to Z_b'^\pm\pi^\mp$. These samples are 500 and 1000 larger than the numbers of such events which would be observed in data with an ideal detector.

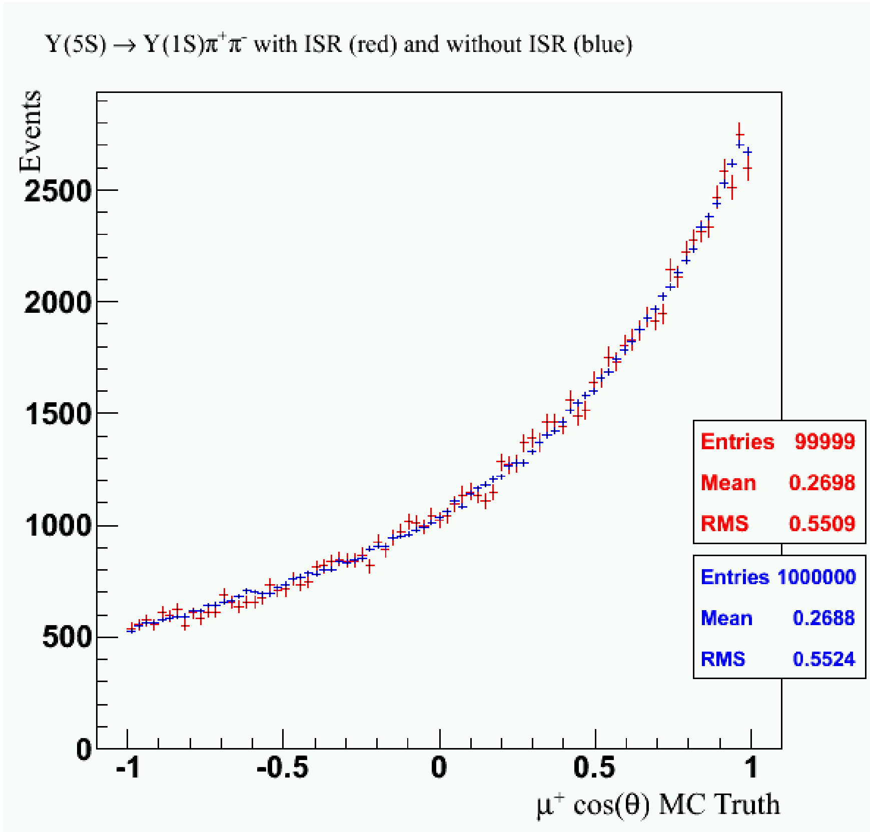


Figure 15: Distributions of $\cos\theta$ for μ^+ for $\Upsilon(5S) \to \Upsilon(1S)\pi^+\pi^-$ events. The distribution shown is red is for events generated with ISR while the distriubtion shown in blue is for events generated without ISR. Events in both distribtuions are generated using PHSP model for both $\Upsilon(5S) \to \Upsilon(1S)\pi^+\pi^-$ and $\Upsilon(1S) \to \mu^+\mu^-$. The blue distribution is normalized to the number of events in the red distribution.

Region Name	Boundary Definitions
Region 1	$10.72\ \text{GeV}/c^2 < M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}}) < 10.80\ \text{GeV}/c^2$ $-0.2\ \text{GeV} < \Delta E < 0.2\ \text{GeV}$
Region 2	$10.49\ \text{GeV}/c^2 < M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}}) < 10.72\ \text{GeV}/c^2$ $0.03\ \text{GeV} < \Delta E < 0.2\ \text{GeV}$
Region 3	$10.38\ \text{GeV}/c^2 < M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}}) < 10.49\ \text{GeV}/c^2$ $-0.2\ \text{GeV} < \Delta E < 0.2\ \text{GeV}$
Excluded Region	$10.49\ \text{GeV}/c^2 < M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}}) < 10.72\ \text{GeV}/c^2$ $-0.2\ \text{GeV} < \Delta E < 0.03\ \text{GeV}$

Table 8: Definitions of subdivisions of the grand sideband region. The Excluded Region is not considered in this analysis.

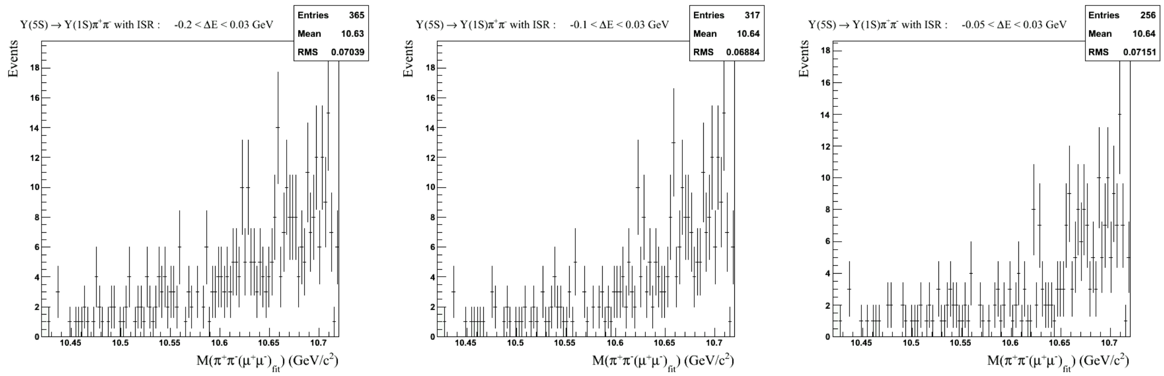


Figure 16: Distributions of $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ for $\Upsilon(5S) \to \Upsilon(1S)\pi^+\pi^-$ with ISR in the signal region for different ΔE requirements. The leftmost distribution requires -0.2 GeV $< \Delta E < 0.03$ GeV, the middle distribution requires -0.1 GeV $< \Delta E < 0.03$ GeV, and the rightmost distribution requires -0.05 GeV $< \Delta E < 0.03$ GeV. The upper bound of ΔE is kept at 0.03 GeV for all distributions, since very few signal events fall beyond $\Delta E > 0.03$ GeV.

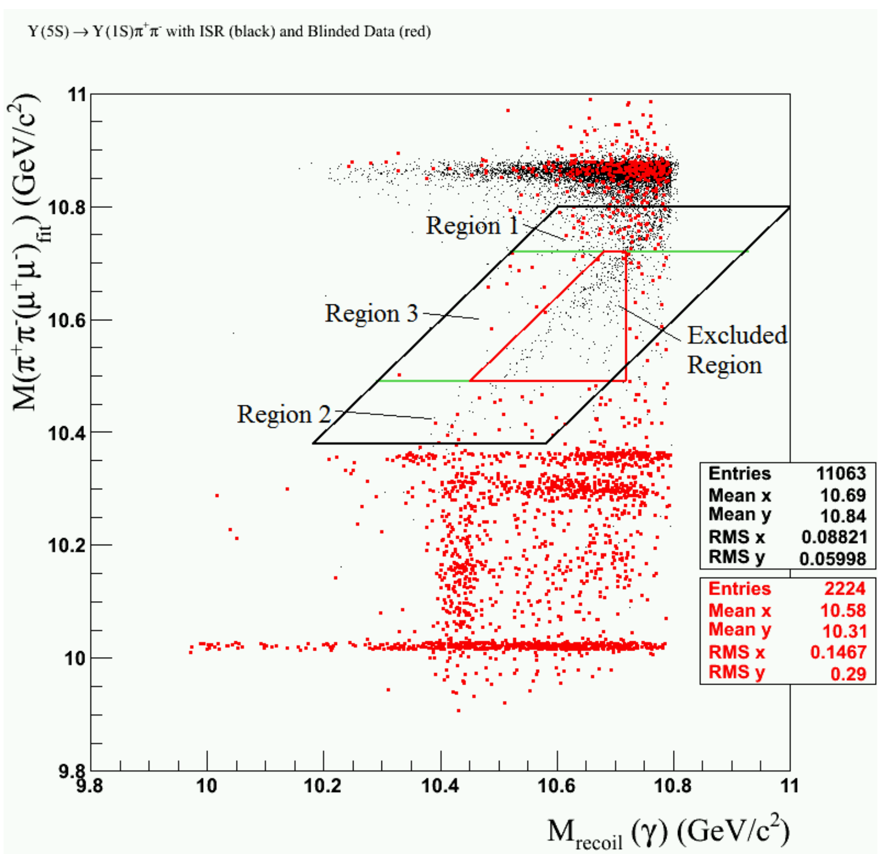
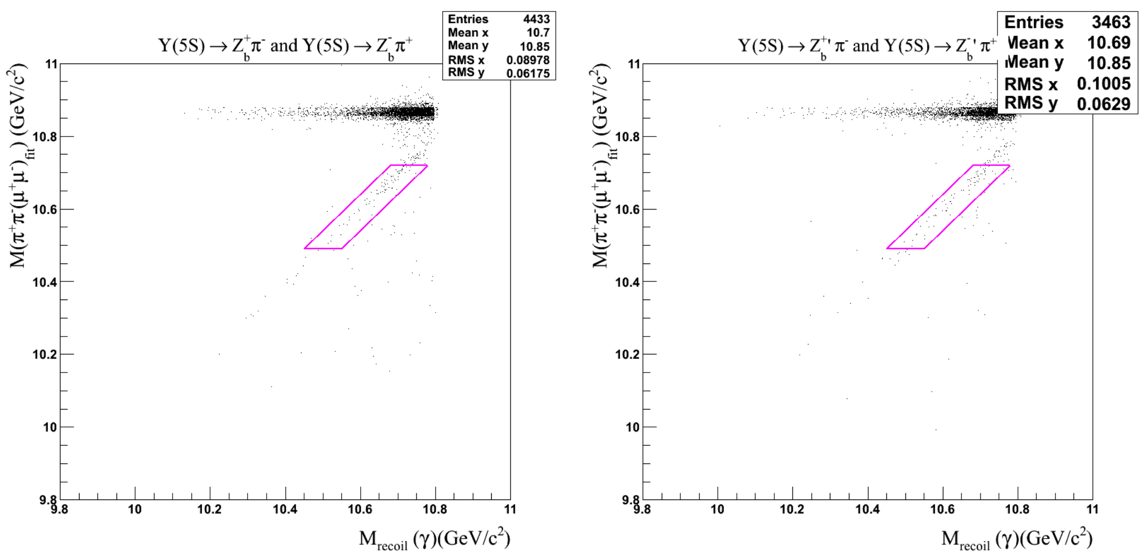


Figure 17: Subdivisions of the grand sideband region. The Excluded Region is not considered in this analysis.

Region	Number of events in ISR MC (N_{mc})	Number of events in blinded data (N_{data})	N_{mc}/N_{data}
Region 1	572	55	10.4
Region 2	28	23	1.2
Region 3	35	14	2.5

Table 9: Comparing the number of events in ISR MC and blinded data in the subdivided grand sideband Region

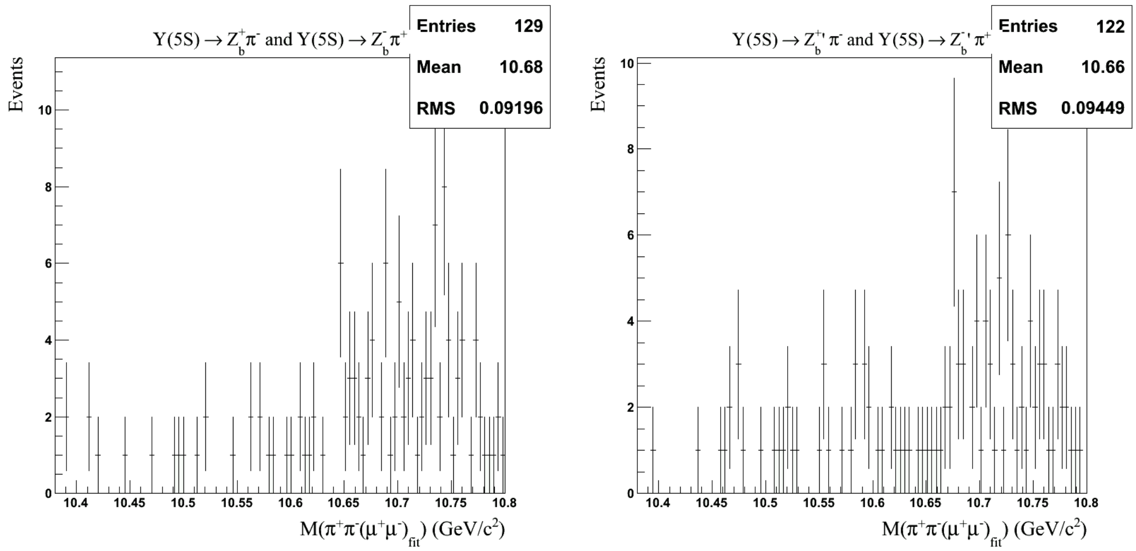


(a) $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ vs $M_{\rm rec}(\gamma)$ for $\Upsilon(5S) \to Z_b^{\pm}\pi^{\mp}$ MC

(b) $\Upsilon(5S) \to Z_b'^{\pm}\pi^{\pm}$ MC

Figure 18: The distribution of $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ vs $M_{\rm rec}(\gamma)$ for $\Upsilon(5S) \to Z_b^{(\prime)\pm}\pi^{\mp}$ MC.

The distribution of $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ vs $M_{\rm rec}(\gamma)$ is shown in Fig. 18 for both samples after applying our selection criteria for the W_{bj} analysis. Fig. 19 shows the distribution of $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ for events inside the signal and sideband region. It is important to note that, approximately, only 2% of events fall in the signal region for each of the two samples. Therefore, we expect less than 100 events from each of the two Z_b samples to be found in the signal region for the W_{bj} analysis. As explained earlier in this section, to predict the "contamination" of our signal region by Z_b events, this number has to be scaled down by the factors of 500 and 1000 for contributions from Z_b and Z_b', respectively. Therefore the process $\Upsilon(5S) \to Z_b^{(\prime)\pm}\pi^{\mp}$ in total, has negligible cross-feed contribution in the signal region and can be safely ignored.



(a) $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ for $\Upsilon(5S) \to Z_b^{\pm}\pi^{\mp}$ MC. (b) $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ for $\Upsilon(5S) \to Z_b\prime^{\pm}\pi^{\mp}$ MC.

Figure 19: The distribution of $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ for $\Upsilon(5S) \to Z_b^{(\prime)\pm}\pi^{\mp}$ MC for events inside the signal and sideband region.

8 Fitting

8.1 Signal and Background PDFs

To extract signal yield, we perform a one-dimensional extended unbinned ML fit to the variable $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ using RooFit [22]. We model the signal distribution of $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ as a Breit-Wigner convolved with the sum of two Guassians (to simulate effects of detector resolution as shown in Fig. 5). The observed width and shape of $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ distribution in signal MC remains practically the same after applying our ΔE requirement and after including ISR. Therefore, we fix the width of our signal PDF. We set the width of the Breit-Wigner to be $\sigma_{BW} = 15$ MeV/c^2 to match the intrinsic width of Z_b and Z_b'. The widths of the Gaussians used in colvolution are $\sigma_{G_1} \approx 3$ MeV/c^2 and $\sigma_{G_2} \approx 7.7$ MeV/c^2 to match the widths obtained from the fit to $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ resolution. We let mean of Breit-Wigner float within the fit, as W_{bJ} could be observed at different invariant masses for different spins J. Table 10 lists the values of parameters used in our signal PDF model.

We use an exponential $e^{\lambda x}$ to model background contributions due to ISR as well as possible non-resonant contribution from dimuon continuum events. Strictly speaking, the background distribution deviates from an exponential at $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit}) \approx 10.75$ GeV/c^2. because of the phase boundary at $M_{\rm rec}(\gamma) \approx 10.75$ GeV/c^2 seen in Fig. 4. This ever-present effect can be seen in figures showing the distribution of $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ for background events with our ΔE requirement (*e.g.* see Fig. 14, Fig. 20b, Fig. 20c). This shortcoming of our analysis will be taken care of in the next version of this Note. We would like to remark that the observed fall-off effect is easy to understand and describe in the model used for fitting, as it is exclusively due to the boundary of phase space.

To estimate the number of background events we expect in the signal region, we per-

Quanitity	Value Used in Signal PDF (MeV/c^2)
σ_{BW}	15
Mean of BW	floats betwen 10.38 and 10.80 GeV/c^2
σ_{G_1}	3.0 ± 0.1
σ_{G_2}	7.7 ± 0.2
Fraction of Gaussian 1	0.73 ± 0.01
Fraction of Gaussian 2	0.27 ± 0.01
Mean of both Gaussians	$(-3.8 \pm 0.2) \cdot 10^{-4}$

Table 10: Values of fixed quantities in the signal PDF model.

form an extended unbinned maximum likelihood fit to data only in the sideband regions. To account for uncertainty in the number of data events in the sideband region, we fit $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ within the range of 10.38 GeV/c^2 and 10.80 GeV/c^2 when extracting signal yield. This range corresponds to the signal region and sideband regions combined. From the fit, we obtain $\lambda = 3.7951$. We extract 59 ± 11 background events in the signal region and sideband regions combined. We expect 27 ± 5 of these background events to be in the signal region alone. Fits to W_{b0} signal MC, $\Upsilon(5S) \to \Upsilon(1S)\pi^+\pi^-$ MC with ISR MC, and data in the sidebands are shown in Fig. 20.

8.2 Confidence Belts

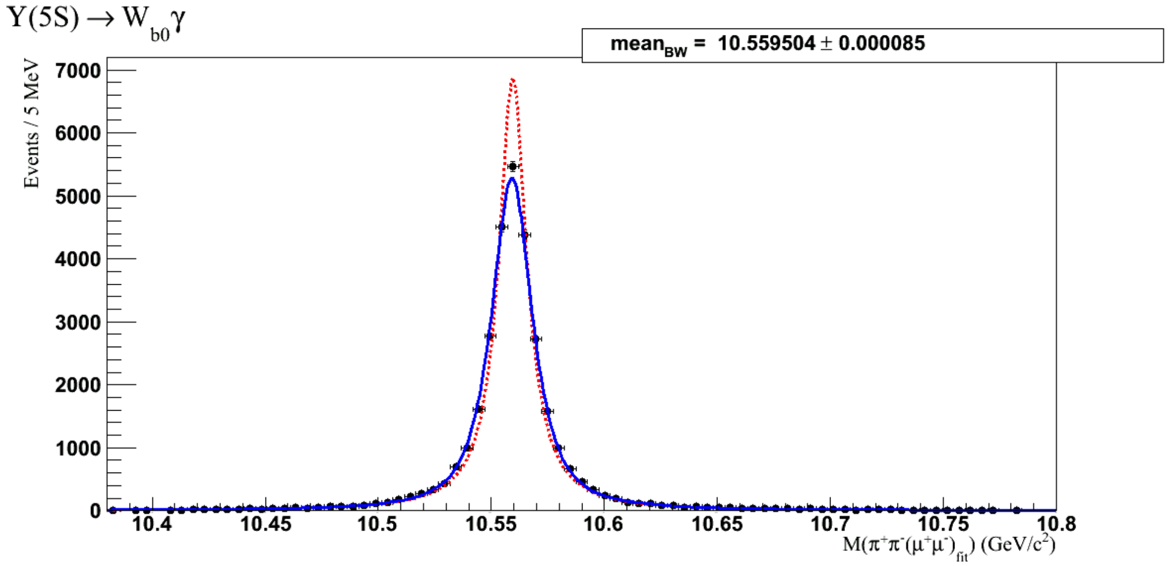
To construct a 90% confidence belt, we perform ensemble tests. Each ensemble test consists of 1000 toy MC experiments. In each toy MC experiment, we generate N_{sig} signal events and N_{bkg} background events according to their respective PDF lineshapes used for fitting signal and background. We then fit the generated events in the range 10.38 GeV/c^2 $<$ $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}}) < 10.80$ GeV/c^2 to our combined signal and background PDF to extract the fitted number of signal events $N_{\text{sig}}^{\text{fit}}$.

We construct our 90% confidence belt by performing ensemble tests with $N_{\text{bkg}}^{\text{gen}} = 59$ for values of $N_{\text{sig}}^{\text{gen}}$ from 0 to 70. We additionally construct a 90% confidence belt where we allow Poisson fluctuation in $N_{\text{bkg}}^{\text{gen}}$. These confidence belts are shown in Fig. 21.

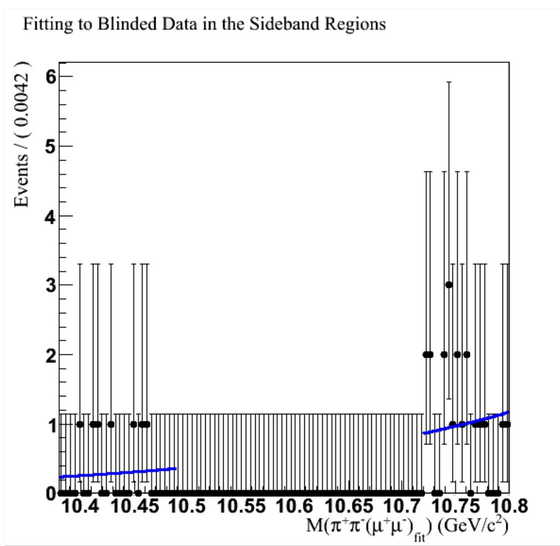
8.3 Linearity Study

To validate our fitting procedures, we perform a linearity study using ensemble tests. Ensemble tests are generated as described in Section 8.2. For each ensemble test of 1000 toy MC experiments, we calculate the average number of signal events from the fit and the error associated with the average. We vary $N_{\text{sig}}^{\text{gen}}$ from 0 to 10 in steps of 1 and from 10 to 50 in steps of 5 while fixing $N_{\text{bkg}} = 59$.

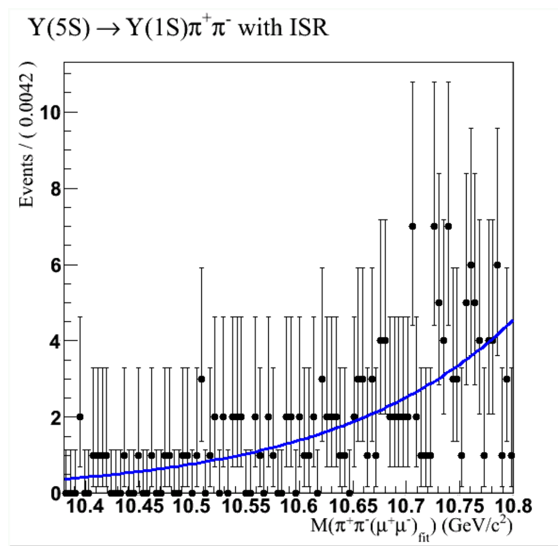
We plot the average number of signal events from the fit against $N_{\text{sig}}^{\text{gen}}$ as shown in Fig. 22. Fig. 23 displays distributions of $N_{\text{sig}}^{\text{fit}}$ for certain values of $N_{\text{sig}}^{\text{gen}}$. When $N_{\text{sig}}^{\text{gen}}$ is large, the distribution of $N_{\text{sig}}^{\text{fit}}$ is unbiased. For small $N_{\text{sig}}^{\text{gen}}$, however. we see an asymmetry in the



(a) Fit result for the distribution of $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ for signal MC in the signal and sideband region. The Breit-Wigner shape is shown in red. The blue distribution is the Breit-Wigner convolved with the sum of two Gaussians.



(b) Fit result for the distribution of $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ for data in the sideband region.



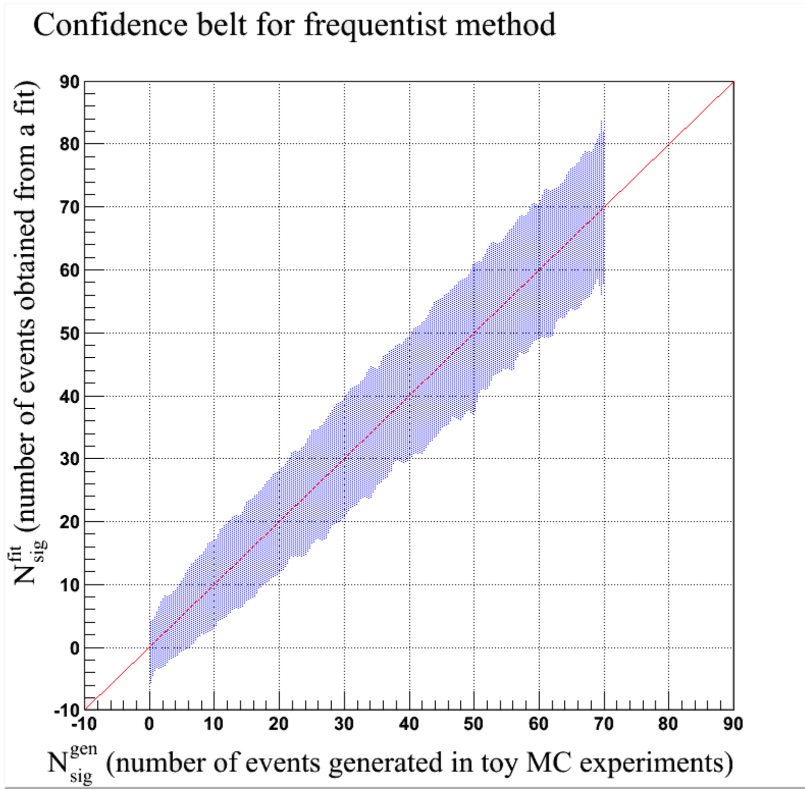
(c) Fit result for the distribution of $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ for $\Upsilon(5S) \to \Upsilon(1S)\pi^+\pi^-$ with ISR for events in the signal region.

Figure 20: Fitting background MC and data

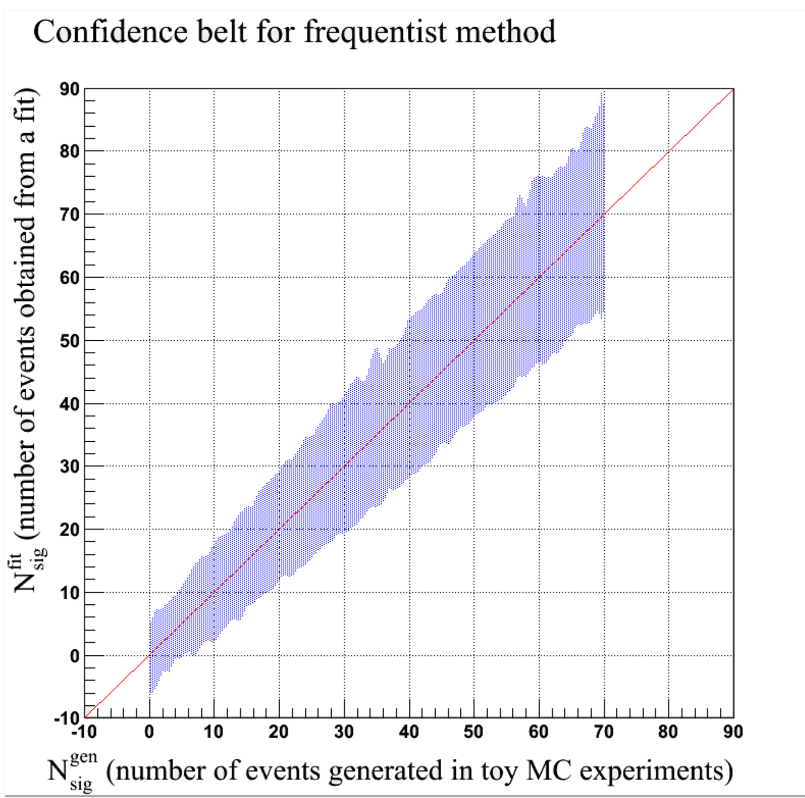
distribution of of $N_{\text{sig}}^{\text{fit}}$, indicating some bias. This effect is often observed for small statistics and is not unexpected.

8.4 Sensitivity Estimation

We estimate the upper limit on the branching fraction and visible cross section of $\Upsilon(5S) \to \gamma W_{bJ}$ in the absence of signal by performing an extended unbinned maximum likelihood fit



(a) Does not include Poisson fluctuations in $N_{\text{bkg}}^{\text{gen}}$



(b) Includes Poisson fluctuations in $N_{\text{bkg}}^{\text{gen}}$.

Figure 21: 90% confidence belts for frequentist method.

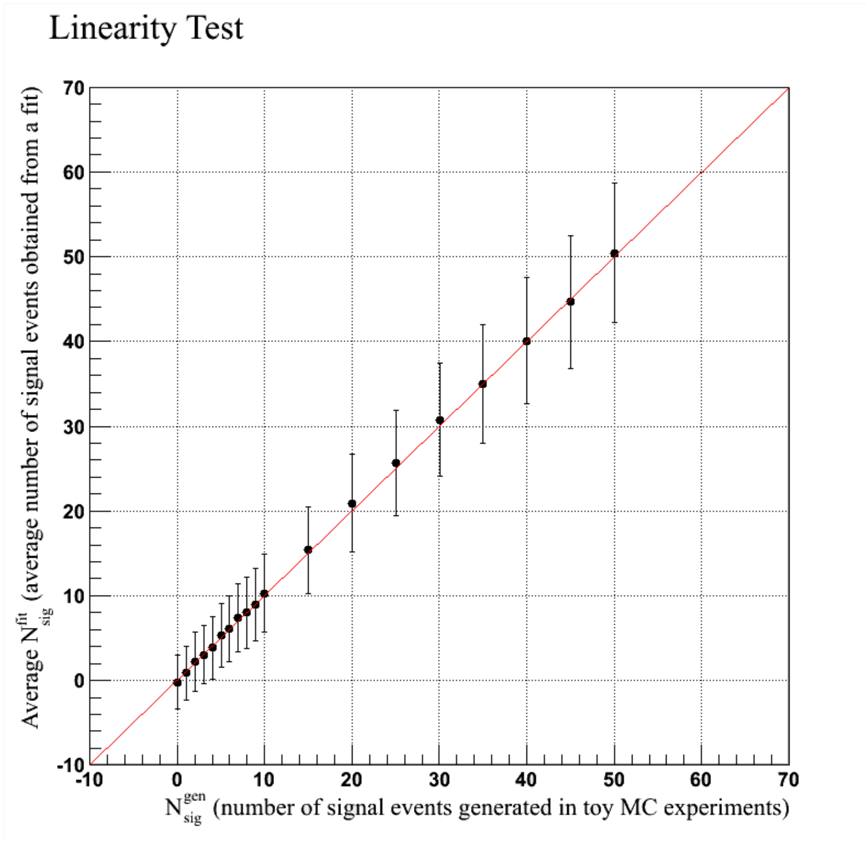


Figure 22: Average $N_{\text{sig}}^{\text{fit}}$ for varying values of $N_{\text{sig}}^{\text{gen}}$.

on toy MC generated according to the fit to the data sidebands. We generate 1000 toy MC samples with 59 background events, fit our combined signal and background shape to each sample, and then average the resulting signal yields. There is an average signal yield of -0.2 ± 3.2 events. Note that in Fig. 22, this average signal yield corresponds to the value plotted at $N_{\text{sig}}^{\text{gen}} = 0$. Using the confidence belt in Fig. 21, we determine the 95% confidence level upper limit on the number of signal events to be 10 events. We calculate the upper limit on the branching fraction in the absence of signal as follows:

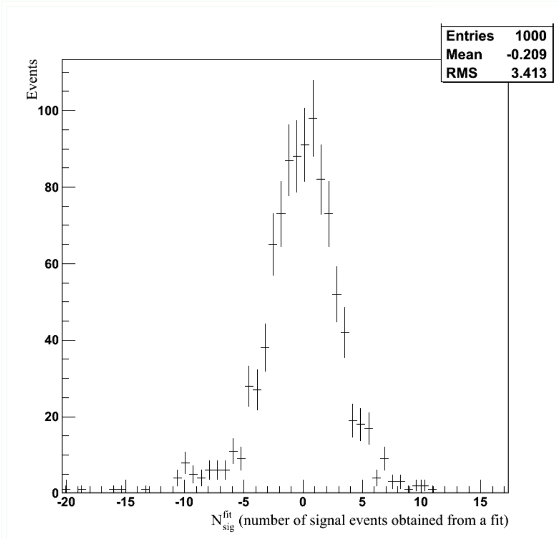
$$\mathcal{B}(\Upsilon(5S) \to \gamma W_{bJ}) \cdot \mathcal{B}(W_{bJ} \to \Upsilon(1S)\rho^0) = \frac{N_{\text{sig}}}{\epsilon \cdot N_{\Upsilon(5S)} \cdot \mathcal{B}(\Upsilon(1S) \to \mu^+\mu^-) \cdot \mathcal{B}(\rho^0 \to \pi^+\pi^-)} \tag{4}$$

where $N_{\Upsilon(5S)}$ is the number of $\Upsilon(5S)$ and ϵ is our reconstruction efficiency. Using Eq. 4, we determine the upper limit on the bracnhing fraction in the absence of signal to be 2.4×10^{-5}.

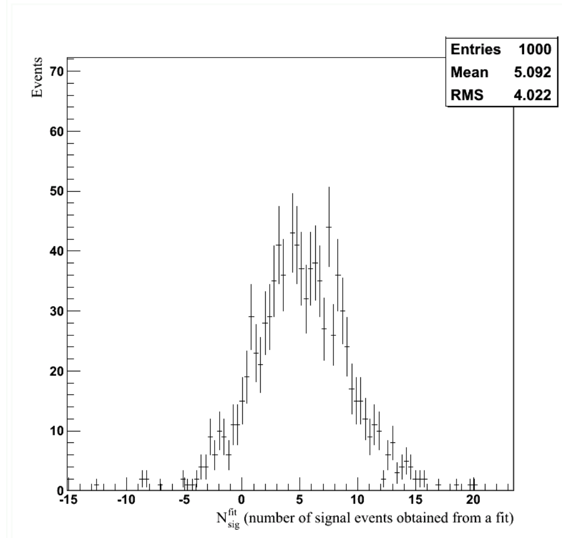
We calculate the visible cross section using

$$\sigma_{\text{vis}} = \frac{N_{\text{sig}}}{\epsilon \mathcal{B}(\Upsilon(1S) \to \mu^+\mu^-)\mathcal{B}(\rho^0 \to \pi^+\pi^-)\mathcal{L}} \tag{5}$$

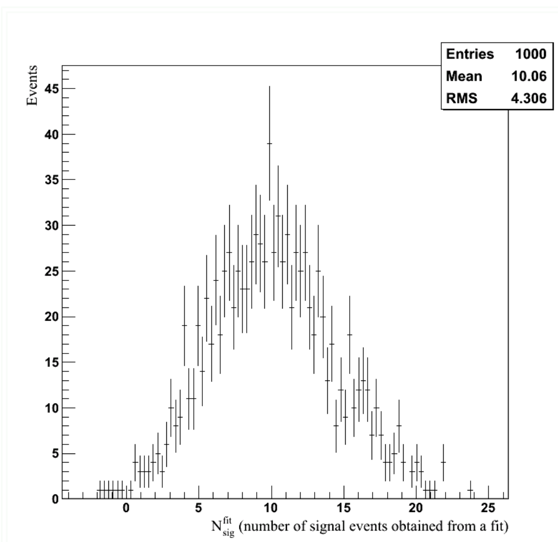
where $\mathcal{L}$ is the integrated luminosity. We find $\sigma_{\text{vis}} = (0.115 \pm 0.006)$ fb. All values used to calcuate the branching fraction and visible cross section are shown in Table 11.



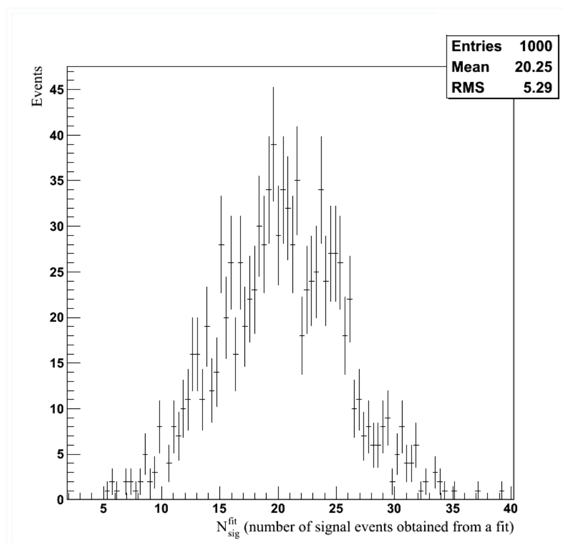
(a) Distribution of $N_{\text{sig}}^{\text{fit}}$ for an ensemble test with $N_{\text{sig}}^{\text{gen}} = 0$ and $N_{\text{bkg}}^{\text{gen}} = 59$.



(b) Distribution of $N_{\text{sig}}^{\text{fit}}$ for an ensemble test with $N_{\text{sig}}^{\text{gen}} = 5$ and $N_{\text{bkg}}^{\text{gen}} = 59$.



(c) Distribution of $N_{\text{sig}}^{\text{fit}}$ for an ensemble test with $N_{\text{sig}}^{\text{gen}} = 10$ and $N_{\text{bkg}}^{\text{gen}} = 59$.



(d) Distribution of $N_{\text{sig}}^{\text{fit}}$ for an ensemble test with $N_{\text{sig}}^{\text{gen}} = 20$ and $N_{\text{bkg}}^{\text{gen}} = 59$.

Figure 23: $N_{\text{sig}}^{\text{fit}}$ Distributions for ensemble tests with different $N_{\text{sig}}^{\text{gen}}$.

9 Summary

In this analysis, we describe a search for a new molecular state W_{bJ} which could be produced in the radiative transition $\Upsilon(5S) \to \gamma W_{bJ}$ followed by the decays $W_{bJ} \to \Upsilon(1S)\rho^0$, $\Upsilon(1S) \to \mu^+\mu^-$, $\rho^0 \to \pi^+\pi^-$ We fully reconstruct the signal final state consisting of two muons, two pions, and a photon. We perform a blind analysis by optimizing our selection criteria and analysis techniques using only MC samples before applying them to data. To search for the presence of W_{bJ} in Belle data, we propose to "unblind" 15% of the data in

Quantity	Value
N_{sig}	10
ϵ	$(29 \pm 0.17)\%$
$N_{\Upsilon(5S)}$	$(6.53 \pm 0.66) \cdot 10^6$
$\mathcal{B}(\Upsilon(1S) \to \mu^+\mu^-)$	$(2.48 \pm 0.05)\%$
$\mathcal{B}(\rho^0 \to \pi^+\pi^-)$	99.8%
$\mathcal{L}$	121.4 fb^{-1}

Table 11: Values of quantities used to calculate upper limits on visible cross section and the branching fraction. Uncertainty in $\mathcal{B}(\rho^0 \to \pi^+\pi^-)$ is negligible. Note that, for purposes of estimating upper limits, we use $N_{\text{sig}} = 10$, which is the 95% CL boundary of the 90% CL frequentist belt shown in Fig. 21 for $N_{\text{sig}}^{\text{fit}} = 3$, according to the result of the fit $N_{\text{sig}}^{\text{fit}} = -0.2 \pm 3.2$.

the signal region and then fit a one-dimensional distribution of $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ using the aforementioned models for signal and background shapes. We will use our confidence belt (Fig. 21) to either claim a discovery of W_{bJ} or establish an upper limit on the signal production rate (branching fraction) for the radiative decay $\Upsilon(5S) \to \gamma W_{bJ}$. The following sources of systematic uncertainties will be considered in our final estimate of the upper limit of the branching fraction of $\Upsilon(5S) \to \gamma W_{bJ}$:

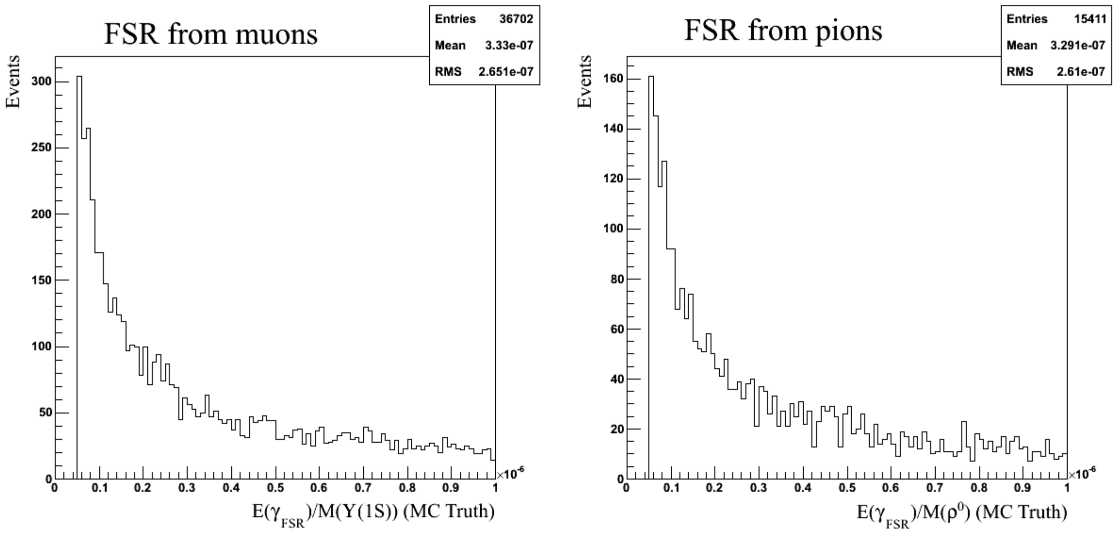
- Number of $\Upsilon(5S)$
- Signal Reconstruction Efficiency
- Daughter Branching Fractions
- MC statistics
- PDF parameterization
- Fit bias
- Trigger efficiency

10 Appendix

10.1 Final State Radiation

In the version of package PHOTOS used by Belle, the minimum FSR photon energy (evaluated in the center of mass frame of charged particle's parent) is calculated as follows:

$$E(\gamma_{FSR}) = (\text{XPHCUT}) \cdot 0.5 \cdot M(\text{parent}) \tag{6}$$



(a) FSR from muons.

(b) FSR from pions.

Figure 24: Final state radiation from charged tracks

where XPHCUT is a hardcoded constant set to 0.01. Hence, the minimum FSR energy is approximately 4 MeV for pions ($M(\rho^0) = 770$ MeV) and 50 MeV for muons ($M(\Upsilon(1S)) = 9.46$ GeV). The lower limit on FSR energy for muons is too high, so we lowered the value of XPHCUT to 10^{-7}. To accomplish this, we changed XPHCUT=0.01D0 to XPHCUT=0.0000001D0, recompiled the phocin.F source code and then rebuilt EvtGen with an updated PHOTOS library.

To verify that XPHCUT was successfully lowered to 10^{-7}, we plot the ratios $\frac{E(\gamma_{FSR})}{M(\Upsilon(1S))}$ and $\frac{E(\gamma^{\rho}_{FSR})}{M(\rho)}$ as generated in Fig. 24. Because these quantities are bounded from below by XPHCUT $\cdot$ 0.5, we prove that XPHCUT was successfully lowered.

References

[1] A. Bondar et al. Observation of two charged bottomonium-like resonances in Y(5S) decays. *Phys. Rev. Lett.*, 108:122001, 2012.

[2] I. Adachi et al. Evidence for a $Z_b^0(10610)$ in Dalitz analysis of $\Upsilon(5S) \to Y(nS)\pi^0\pi^0$. 2012.

[3] P. Krokovny et al. First observation of the $Z_b^0(10610)$ in a Dalitz analysis of $\Upsilon(10860) \to \Upsilon(\text{nS})\pi^0\pi^0$. *Phys. Rev.*, D88(5):052016, 2013.

[4] A. Garmash et al. Observation of $Z_b(10610)$ and $Z_b(10650)$ Decaying to B Mesons. *Phys. Rev. Lett.*, 116(21):212001, 2016.

[5] L. M. Lederman. The discovery of the Upsilon, bottom quark, and B mesons. In *The Rise of the standard model: Particle physics in the 1960s and 1970s. Proceedings, Conference, Stanford, USA, June 24-27, 1992*, pages 101–113, 1992.

[6] M. Tanabashi et al. Review of Particle Physics. *Phys. Rev.*, D98(3):030001, 2018.

[7] S. Godfrey and N. Isgur. Mesons in a Relativized Quark Model with Chromodynamics. *Phys. Rev.*, D32:189–231, 1985.

[8] S. Godfrey and K. Moats. Bottomonium Mesons and Strategies for their Observation. *Phys. Rev.*, D92(5):054034, 2015.

[9] S. Godfrey, K. Moats, and E. S. Swanson. B and B_s Meson Spectroscopy. *Phys. Rev.*, D94(5):054025, 2016.

[10] A. E. Bondar, R. V. Mizuk, and M. B. Voloshin. Bottomonium-like states: Physics case for energy scan above the $B\bar{B}$ threshold at Belle-II. *Mod. Phys. Lett.*, A32(04):1750025, 2017.

[11] A. E. Bondar, A. Garmash, A. I. Milstein, R. Mizuk, and M. B. Voloshin. Heavy quark spin structure in Z_b resonances. *Phys. Rev.*, D84:054010, 2011.

[12] Stephen Lars Olsen. A New Hadron Spectroscopy. *Front. Phys.(Beijing)*, 10(2):121–154, 2015.

[13] M. B. Voloshin. Radiative transitions from Upsilon(5S) to molecular bottomonium. *Phys. Rev.*, D84:031502, 2011.

[14] W. Altmannshofer et al. The Belle II Physics Book, arXiv:hep-ex/1808.10567, 2018.

[15] I. Adachi et al. Study of Three-Body Y(10860) Decays, arXiv:hep-ex/1209.6450, BELLE-CONF-1272, 2012.

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

[17] S. Agostinelli et al. GEANT4: A Simulation toolkit. *Nucl. Instrum. Meth.*, A506:250–303, 2003.

[18] A. Ryd, D. Lange, N. Kuznetsova, S. Versille, M. Rotondo, D. P. Kirkby, F. K. Wuerthwein, and A. Ishikawa. EvtGen: A Monte Carlo Generator for B-Physics. 2005.

[19] E. Barberio and Z. Was. PHOTOS: A Universal Monte Carlo for QED radiative corrections. Version 2.0. *Comput. Phys. Commun.*, 79:291–308, 1994.

[20] J. Brodzicka et al. Physics Achievements from the Belle Experiment. *PTEP*, 2012:04D001, 2012.

[21] M. Benayoun, S. I. Eidelman, V. N. Ivanchenko, and Z. K. Silagadze. Spectroscopy at B factories using hard photon emission. *Mod. Phys. Lett.*, A14:2605–2614, 1999. [Frascati Phys. Ser.15(1999)].

[22] Wouter Verkerke and David P. Kirkby. The RooFit toolkit for data modeling. *eConf*, C0303241:MOLT007, 2003. [186(2003)].