Belle Note 1522

Version 1.9
August 2, 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 $\Upsilon$(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

Version 1.5 of this Note includes new plots, various corrections suggested and the answers to the questions asked by the referees. Also, this version includes two new appendices: section 10.2, "Changes in the Analysis between Note v1.5 and v2.0" and section 10.3, "Fitting Strategy". Note that the plots in the main part of the note (*i.e.* excluding the plots in the Appendix part) have been prepared using the original analysis. The changes outlined in the section 10.2 and used for preparing plots in sections 10.2 and 10.3, have not yet been applied to the main body of the text of this Note.

## 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 $\Upsilon$ 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 significantly 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

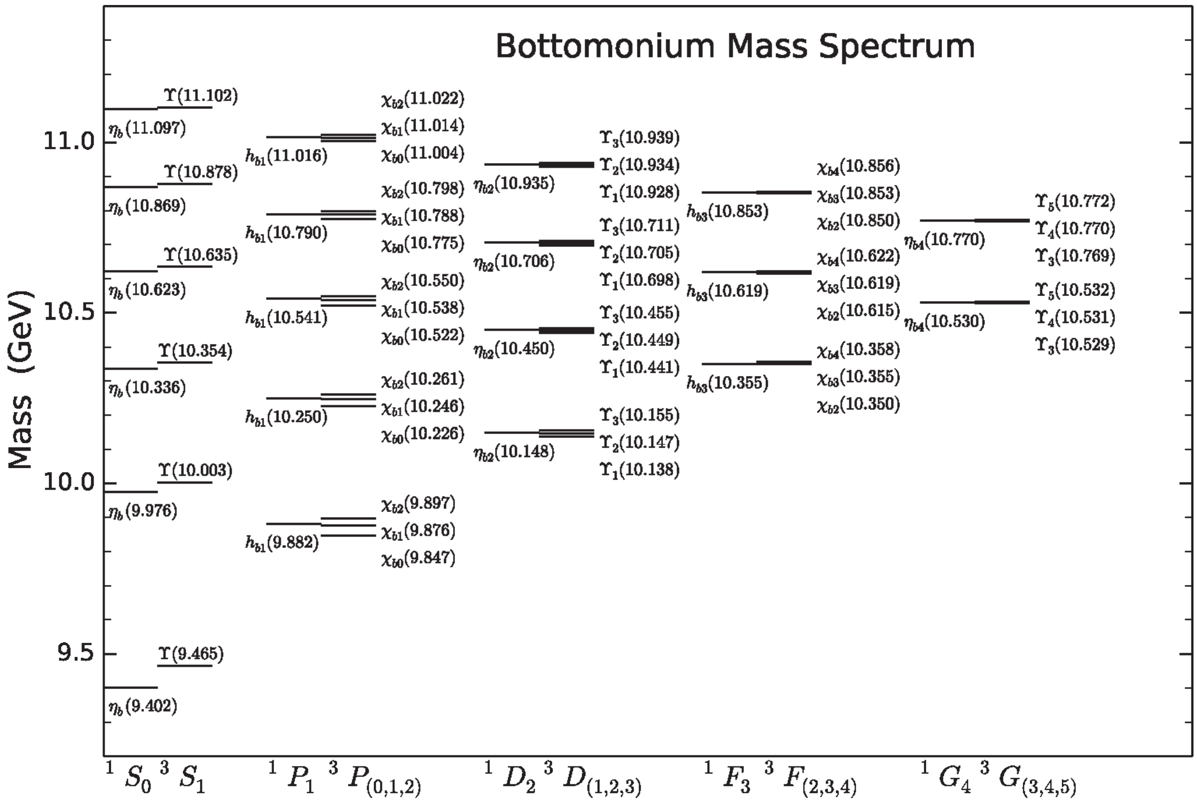



Figure 1: Pure (*i.e.* $b\bar{b}$) bottomonium mass spectrum [8] calculated using a relativized quark model [7].

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 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 our 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

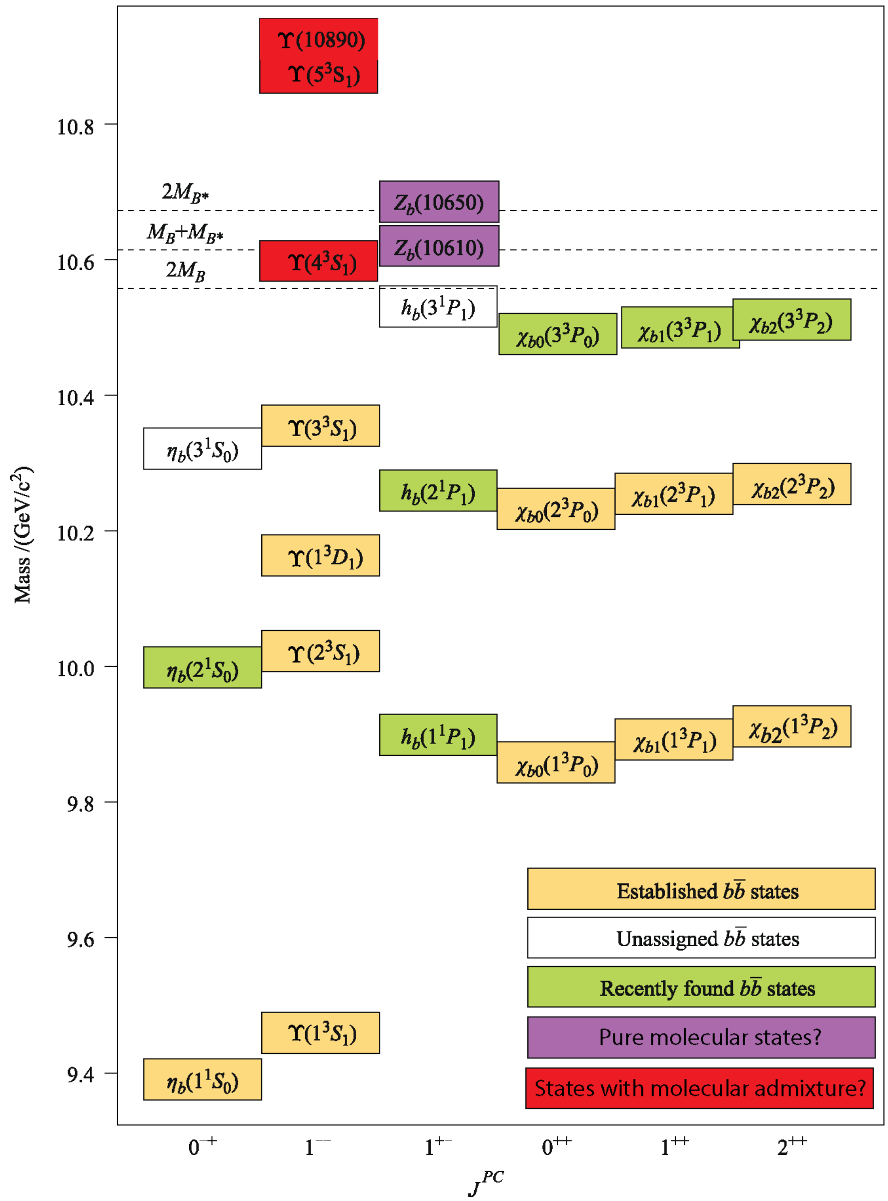



Figure 2: Most relevant (for our study) states in conventional bottomonium and bottomonium-like spectroscopies. We stole this figure from S. Olsen's excellent review article [12]. Note that we took liberty to modify the original figure to better represent the contents of this Note, namely, we relabeled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

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.

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

$Z_b$ are spin-1 isotriplets (both neutral and charged states were discovered in transitions $\Upsilon(nS) \to \pi Z_b$ $(n = 5, 6)$. 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 $\rho$ 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 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 $\rho$ invariant mass. In our analysis, instead of searching for decays with the $\rho$ mesons, we have to allow for the $q\bar{q}$ annihilation and pay the price of approximately $\alpha_{\rm em}$ in the branching fraction:

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

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.

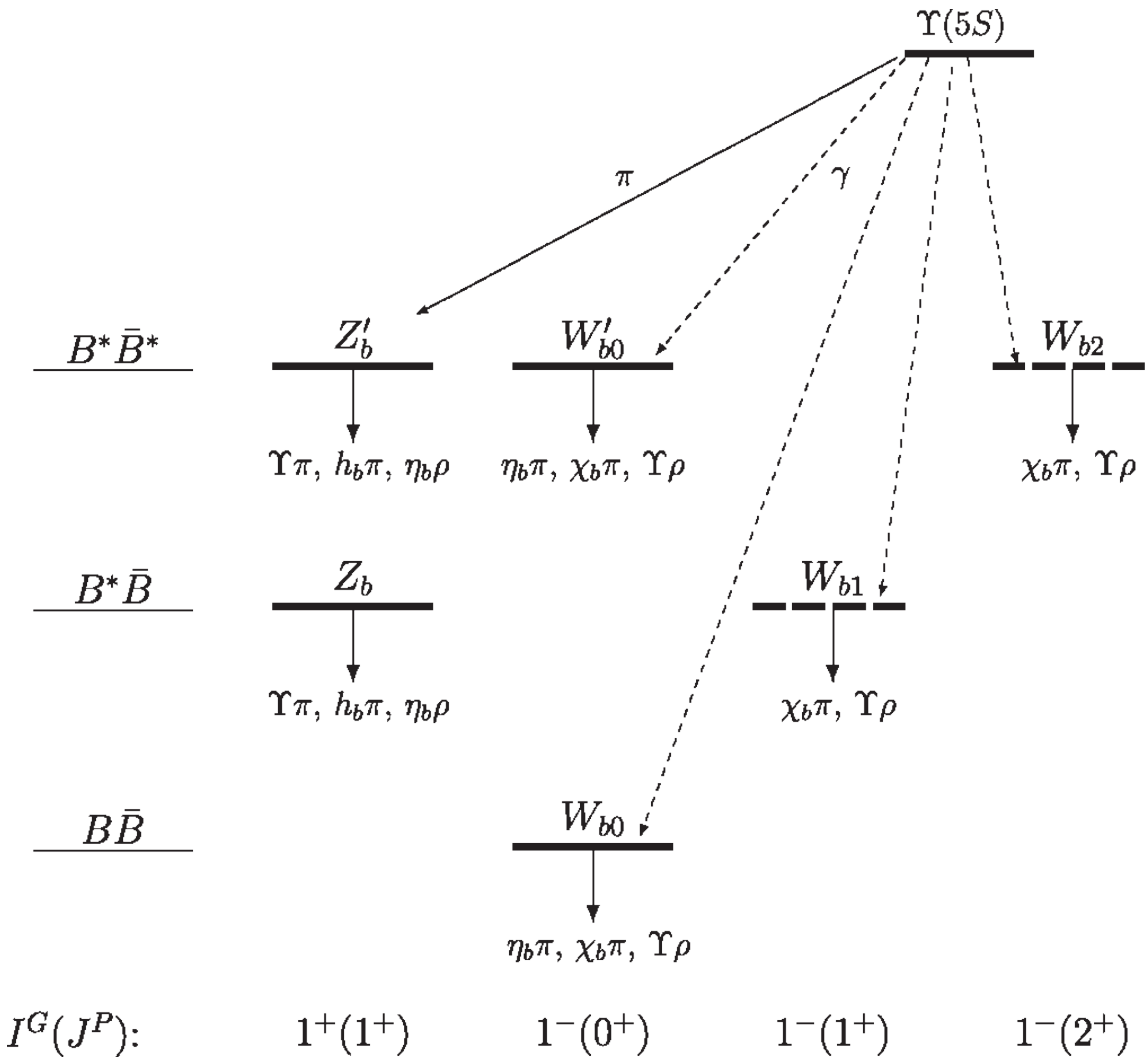



Figure 3: The expected family of isotriplet resonances from Ref. [13] (which the reader is advised to consult for relevant details). For $\Upsilon(6S)$ transitions, the photon is replaced by $\rho$. This would also allow to access charged $W_{bJ}$ states. Also, please see Table 1.

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 $\alpha_{\rm em}$. 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!) 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 our studies of these processes in Section 6 and Section 7, respectively.

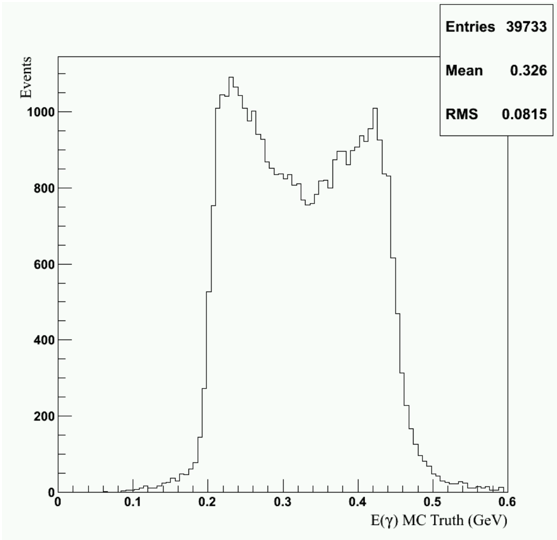
In this analysis, we use the full 121.4 fb$^{-1}$ of on-resonance $\Upsilon(5S)$ data collected by the Belle detector at the KEKB collider from asymmetric energy $e^+e^-$ collisions with $\sqrt{s} = 10.86$ GeV [20].

| 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) |

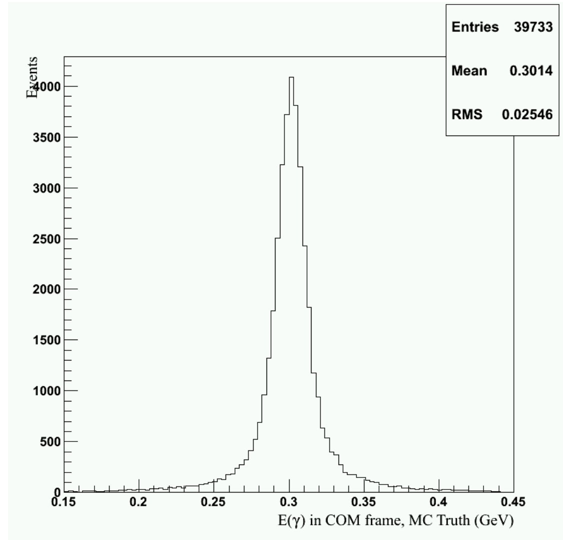
Table 2: EvtGen decay models used in Mote Carlo simulation of signal processes.

## 3 Selection Criteria

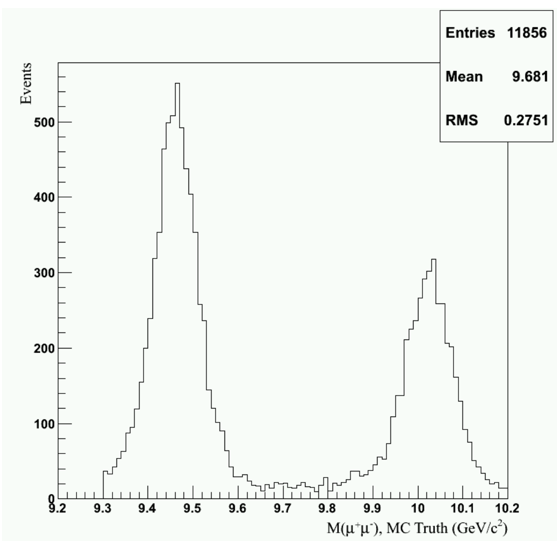
We reconstruct the decay mode $\Upsilon(5S) \to \gamma W_{bJ}$ followed by the decays $W_{bJ} \to \Upsilon(1S)\rho^0$, $\Upsilon(1S) \to \mu^+\mu^-$, $\rho^0 \to \pi^+\pi^-$. We select a fully-reconstructed final state particle combination consisting of $\pi^+\pi^-\mu^+\mu^-\gamma$. The selection criteria that follow, though not systematically optimized, are based on MC truth distributions and typical choices made in previous Belle analyses.



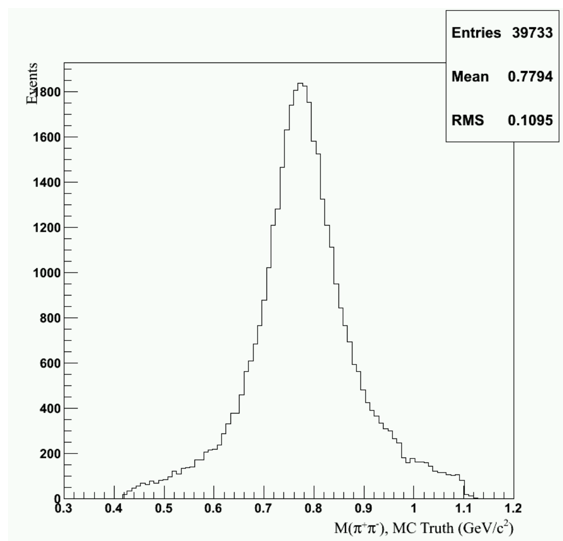

(a) Photon energy in the lab frame for events in signal MC.




(b) Photon energy in the COM frame for events in signal MC.




(c) ($M(\mu^+\mu^-)$ for events in generic MC. The left peak is $\Upsilon(1S)$ and the right peak is $\Upsilon(2S)$. Note that the right tail of $\Upsilon(1S)$ overlaps with the left tail of $\Upsilon(2S)$.




(d) $M(\pi^+\pi^-)$ for events in signal MC.

Figure 4: Various MC distributions which informed our selection criteria.

## 3.1 Selection of Photon Candidates

We require reconstructed photons have energy between 100 and 600 MeV (in the lab frame) 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 ($E_{25}$). See Fig. 4a and Fig. 4b for relevant distributions.

### 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 $R_{e,\text{hadron}} < 0.9$, where $R_{e,\text{hadron}}$ 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 $R_\mu$ 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 tracks 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$^2$. See Fig. 4c and Fig. 4d for relevant distributions.

### 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 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, $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ and $M_{\text{rec}}(\gamma)$, 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

| Particle Candidate | Selection Criteria |
|---|---|
| $\gamma$ | $\cancel{100\text{ MeV} \leq E(\gamma) \leq 600\text{ MeV}}$<br>$20\text{ MeV} \leq E(\gamma) \leq 5000\text{ MeV}$ |
| $\pi^{\pm}, \mu^{\pm}$ | $\cancel{dr < 0.3\text{ cm}}$ $dr < 0.5$ cm<br>$\cancel{\lvert dz \rvert < 2\text{ cm}}$ $\lvert dz \rvert < 4$ cm<br>$p_T > 100$ MeV/c |
| $\cancel{\pi^{\pm}\text{ PID}}$ | $\cancel{R_{K,\pi} < 0.9}$<br>$\cancel{R_{e,hadron} < 0.9}$ |
| $\mu^{\pm}$ | $R_{\mu} > 0.10$ |
| $\rho^0$ | $0.420\text{ GeV}/c^2 < M_{\pi^+\pi^-} < 1.020\text{ GeV}/c^2$ |
| $\Upsilon(1S)$ | $9.3\text{ GeV}/c^2 < M_{\mu^+\mu^-} < 9.6\text{ GeV}/c^2$ |
| $\Upsilon(5S)$<br><br>(full event reconstruction) | $10.2\text{ GeV}/c^2 < M_{\pi^+\pi^-\mu^+\mu^-\gamma} < 11.5\text{ GeV}/c^2$<br>$-0.05\text{ GeV} < \Delta E < 0.03\text{ GeV}$<br>Exactly four tracks: two muons and two pions |

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

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

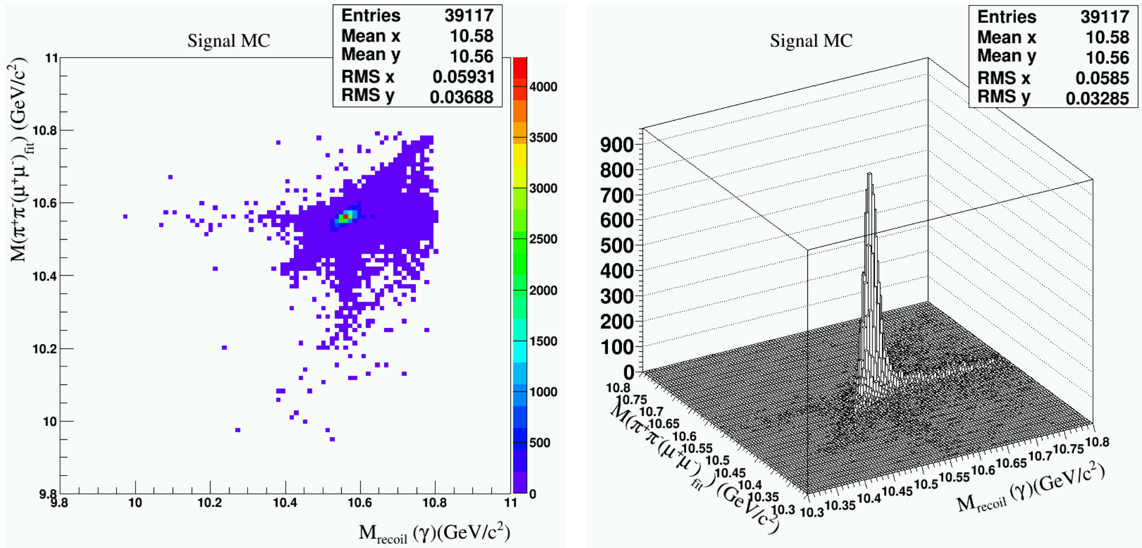
where $E_{\text{cm}}(\text{exp})$ is the run's average energy, and $E_{\text{cm}}(X)$ and $\vec{p}_{\text{cm}}(X)$ are the energy and momentum of system $X$. Subscript "cm" is used for quantities evaluated in the center of mass reference frame of the experiment. For signal events, $M_{\text{rec}}(\gamma)$ and $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ are two independent ways to estimate the invariant mass of $W_{bJ}$. Fully reconstructed signal events fall along the main diagonal of the $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ vs $M_{\text{rec}}(\gamma)$ plot shown in Fig. 5. We define energy balance $\Delta E$ as

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

$\Delta E$ 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 $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$: (1) the intrinsic width of $W_{bJ}$, and (2) the charged track reconstruction. Fig. 6 shows $M(\pi^+\pi^+\mu^+\mu^-)$ and $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ 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 $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ resolution from charged track reconstruction is primarily due to pions, since muon pairs are constrained to $\Upsilon(1S)$ invariant mass.

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



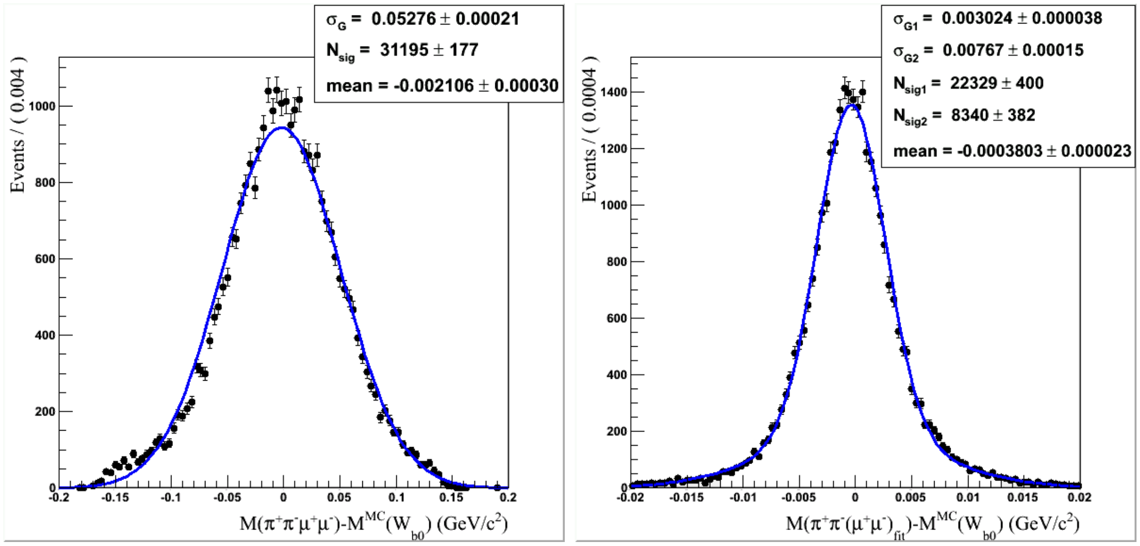

(a) (b)

Figure 5: $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ vs $M_{\rm rec}(\gamma)$ distribution for $W_{B0}$ signal MC events. We show the lego plot in Fig. 5b to emphasize that the tail of $M_{\rm rec}(\gamma)$ is not as large as it appears in Fig. 5a. Note that Fig. 5b 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

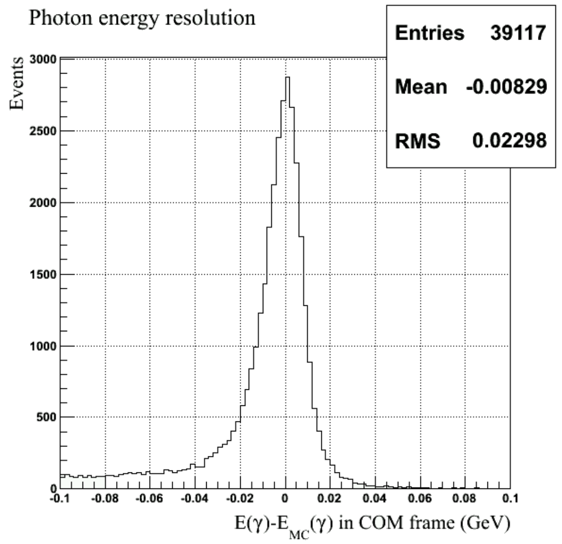
Effects contributing to the observed shape of $\Delta E$ include (1) photon energy resolution, (2) charged track resolution, (3) beam energy resolution, and (4) the intrinsic width of $W_{bJ}$. $\Delta E$ resolution is dominated by photon energy resolution as well. The values of relevant widths are listed 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 $\Delta E$ 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 $-0.05\,\text{GeV} \leq \Delta E \leq 0.03\,\text{GeV}$. This selection cuts out the long tail in the distribution 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 $\Delta E$, signal reconstruction efficiency becomes approximately 31%. Fig. 7 displays $\Delta E$ resolution as well as quantities contributing to $\Delta E$ resolution.



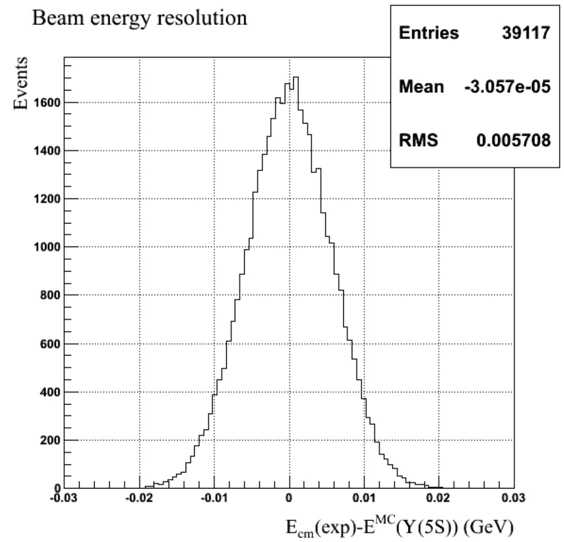

(a) $M(\pi^+\pi^+\mu^+\mu^-)$ resolution. Note that muons are not mass constrained.

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

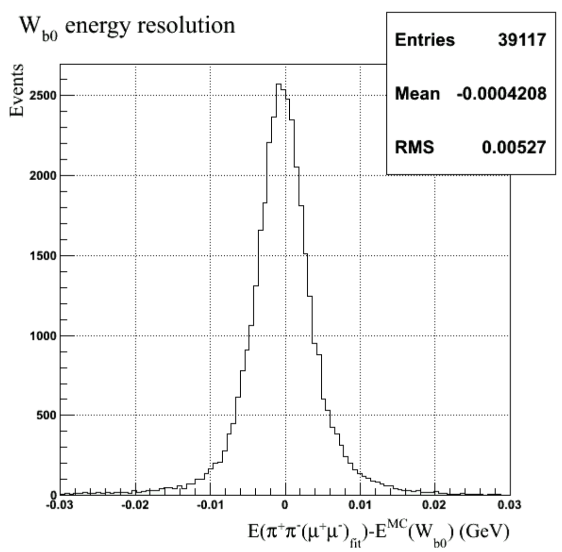
Figure 6: $M(\pi^+\pi^+\mu^+\mu^-)$ and $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ resolutions for signal events within the signal region and sideband regions (defined in Section 4.2). Note that the horizontal scales are different.



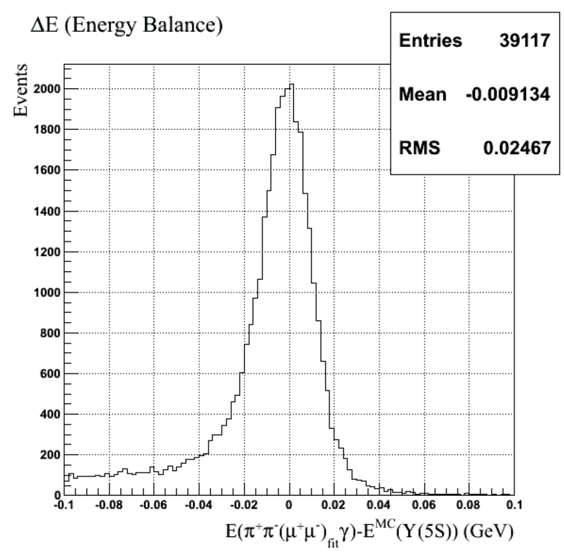

(a) Signal photon energy line shape in the COM reference frame.




(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 7: $\Delta E$ resolution and quantities contributing to $\Delta E$ 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 $10.49\,\mathrm{GeV/c^2} \leq M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit}) \leq 10.72\,\mathrm{GeV/c^2}$. The boundary on

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

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

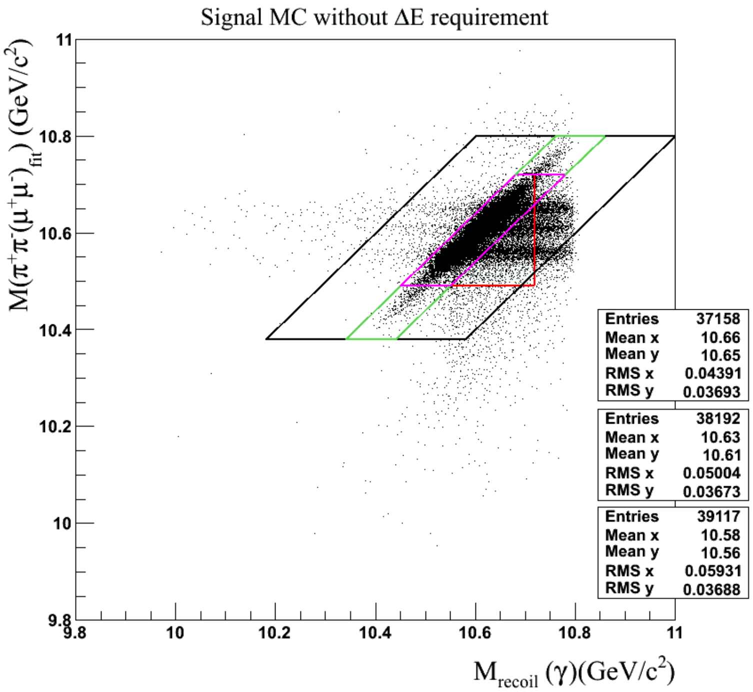
the left side of the region is defined by the sloped line $M_{\mathrm{rec}}(\gamma) \geq M(\pi^+\pi^-(\mu^+\mu^-)_{\mathrm{fit}}) - 0.04$ GeV/c$^2$ which lies parallel to the main diagonal. Approximately 20% of signal events are located in the long right tail of the distribution of $M_{\mathrm{rec}}(\gamma)$. A phase space boundary on the right side of the plot at $M_{\mathrm{rec}}(\gamma) \approx 10.75$ GeV/c$^2$ forces this long tail of the $M_{\mathrm{rec}}(\gamma)$ 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 $M_{\mathrm{rec}}(\gamma) \leq$ 10.72 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 $10.49\,\mathrm{GeV/c^2} \leq M(\pi^+\pi^-(\mu^+\mu^-)_{\mathrm{fit}}) \leq 10.72\,\mathrm{GeV/c^2}$ satisfying $-0.05\,(\mathrm{GeV}) \leq \Delta E \leq 0.03\,\mathrm{GeV}$. The $\Delta E$ requirement selects only fully-reconstructed signal events, where signal is peaking.

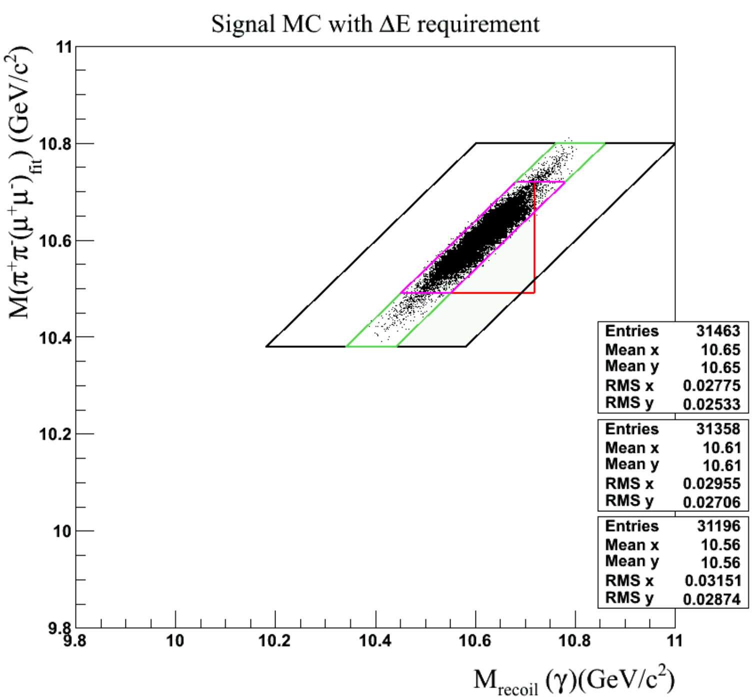
The sideband region is essentially an extension of the signal region, defined as the regions within $10.38\,\mathrm{GeV/c^2} \leq M(\pi^+\pi^-(\mu^+\mu^-)_{\mathrm{fit}}) \leq 10.49\,\mathrm{GeV/c^2}$ and $10.72\,\mathrm{GeV/c^2} \leq M(\pi^+\pi^-(\mu^+\mu^-)_{\mathrm{fit}}) \leq 10.80\,\mathrm{GeV/c^2}$ satisfying $-0.05\,(\mathrm{GeV}) \leq \Delta E \leq 0.03\,\mathrm{GeV}$.

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

Fig. 8 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 $\Delta E$ in this analysis.




(a) Does not include $\Delta E$ requirement.




(b) Includes $\Delta E$ requirement.

Figure 8: The blinded region (red), signal region (magenta), sideband region (green), and the grand sideband region (black). The plot in 8a includes the aforementioned $\Delta E$ requirement, while the plot in 8b does not. From top to bottom, the statistics boxes correspond to $W_{b0}', W_{b1}$, and $W_{b0}$ signal MC, respectively.

## 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. 9 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. 10 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%.

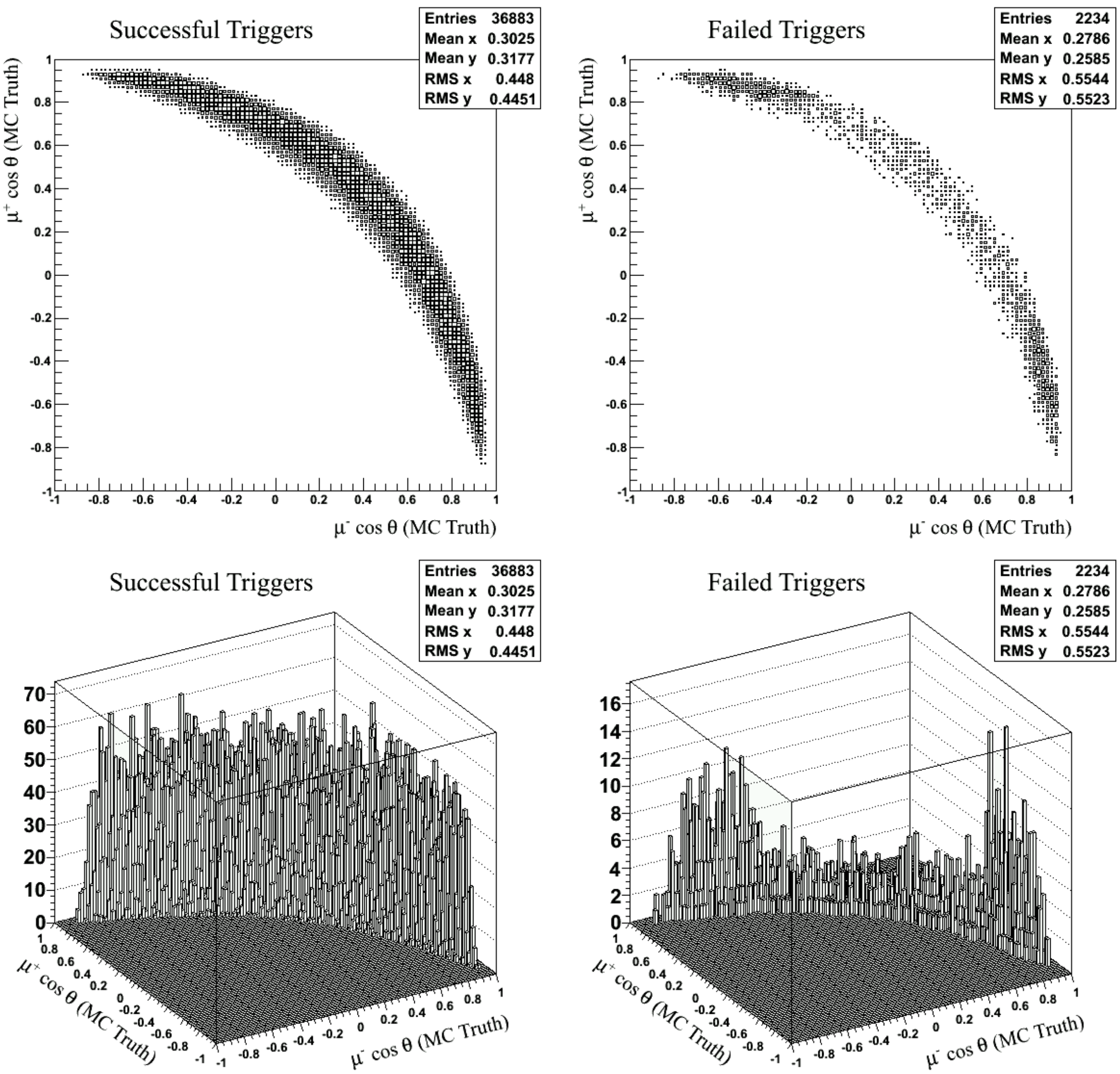



Figure 9: Reconstructed signal MC events that satisfy the offline trigger selection are plotted on the left, while events that fail the offline trigger selection are plotted on the right. We observe that events satisfying the trigger criteria are distributed more or less uniformly for kinematically allowed 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.

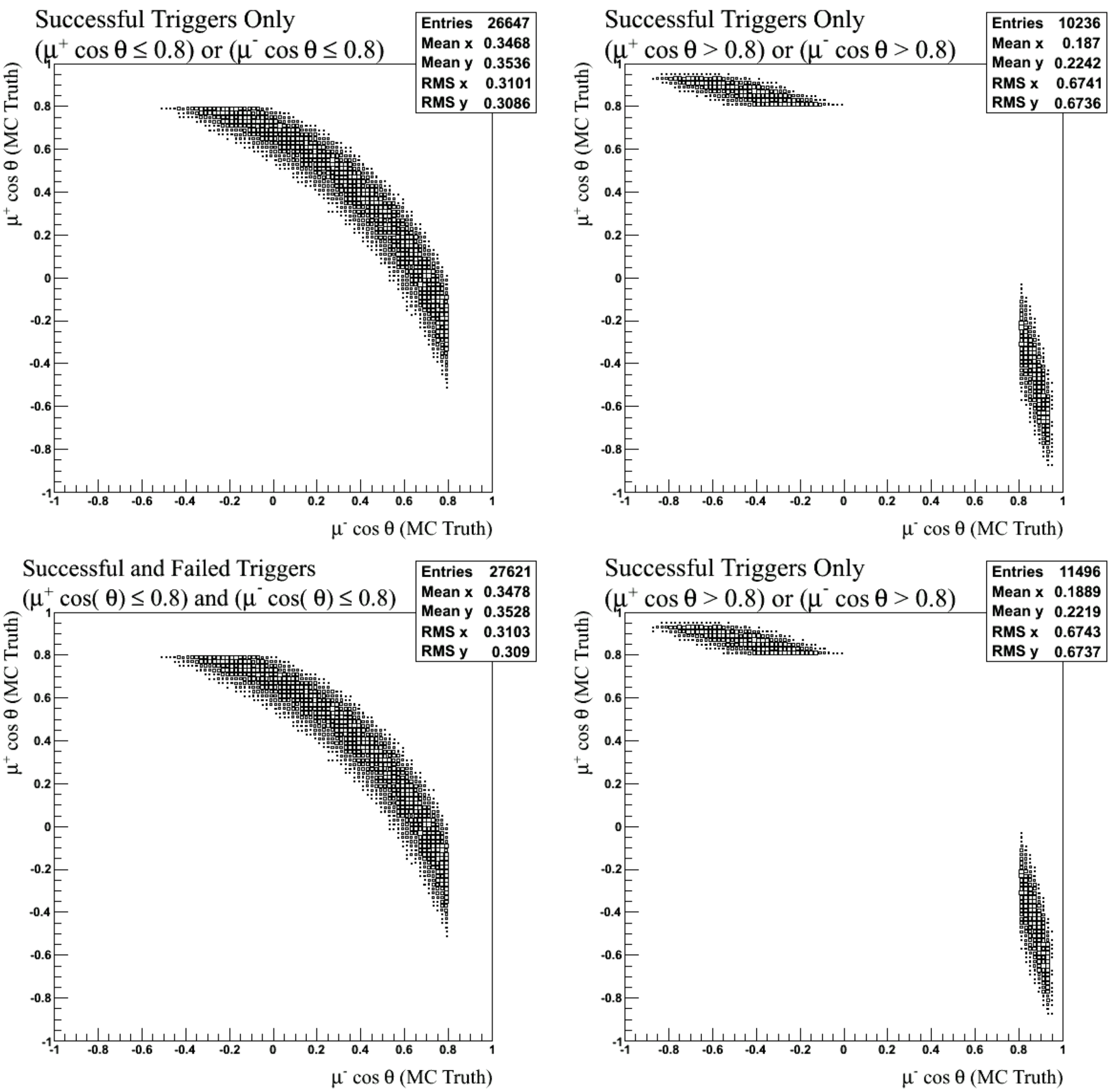



Figure 10: All reconstructed events in which both muons are generated with $|\cos(\theta)| < 0.8$ are plotted in the left two figures. Trigger efficiency for such events is approximately $(96 \pm 4)\%$. In the right two figures, we plot all reconstructed events where one of the muons is generated with $|\cos(\theta)| > 0.8$. Trigger efficiency for these events is reduced to about $(89 \pm 4)\%$.

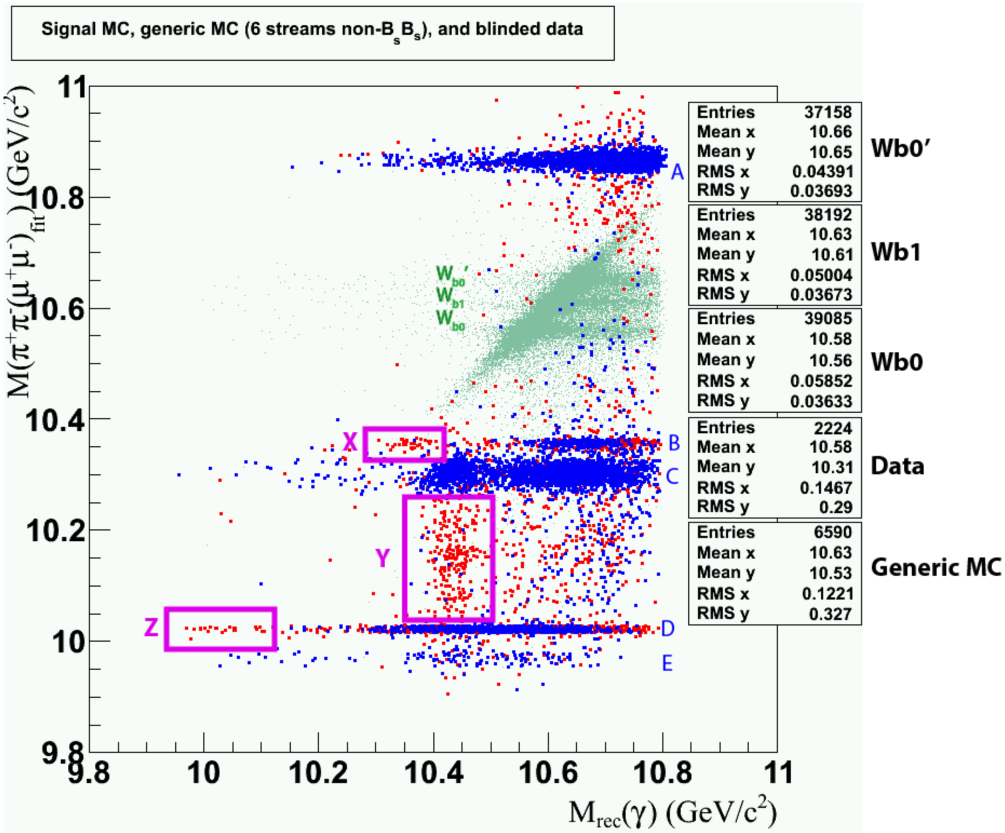



Figure 11: $W_{b0}$, $W_{b1}$, and $W'_{b0}$ signal MC (light green), six streams of non-$B_sB_s$ generic MC (blue), and data with the signal region blinded (red).

# 5 Background Studies

## 5.1 Generic Monte Carlo and Blinded Data

Fig. 11 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 $\Delta E$ 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 $\Delta E$ 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. 11 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

| Label | Background |
|---|---|
| A | $\Upsilon(5S) \to \Upsilon(1S)\pi^+\pi^- \to \mu^+\mu^-\pi^+\pi^-$ |
| B | $\Upsilon(5S) \to \Upsilon(3S)\pi^+\pi^- \to \Upsilon(1S)\pi^+\pi^-\pi^+\pi^- \to \mu^+\mu^-\pi^+\pi^-\pi^+\pi^-$<br>$\Upsilon(5S) \to \Upsilon(3S)\pi^0\pi^0 \to \Upsilon(1S)\pi^+\pi^-\pi^0\pi^0 \to \mu^+\mu^-\pi^+\pi^-\pi^0\pi^0$ |
| C | $\Upsilon(5S) \to \Upsilon(2S)\pi^+\pi^- \to \Upsilon(1S)\pi^+\pi^-\pi^+\pi^- \to \mu^+\mu^-\pi^+\pi^-\pi^+\pi^-$<br>$\Upsilon(5S) \to \Upsilon(2S)\pi^+\pi^- \to \Upsilon(1S)\pi^0\pi^0\pi^+\pi^- \to \mu^+\mu^-\pi^0\pi^0\pi^+\pi^-$ |
| 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 $\chi_{bJ}(1P) \to \gamma\Upsilon(1S)$,<br>*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. 11.

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 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.

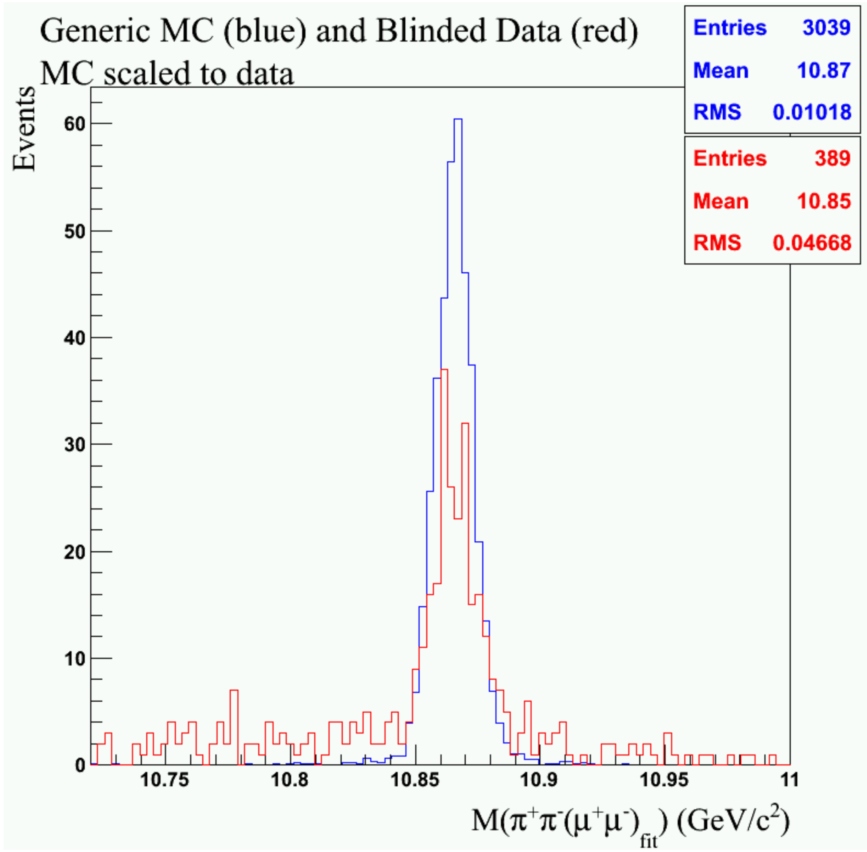



Figure 12: $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ distributions for $\Upsilon(5S) \to \Upsilon(1S)\pi^+\pi^-$ events (label 'A' in Table 6). Distributions for generic MC and blinded data are shown in blue and red, respectively. Generic MC does not include ISR and is normalized to the number of data events 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 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. 11) have a much longer tail in data than in generic MC. This difference is shown in Fig. 12, 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.

## 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 $\alpha^2$ [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. 13.

Fig. 14 shows the $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ vs $M_{\rm rec}(\gamma)$ distribution for reweighted $\Upsilon(5S) \to$

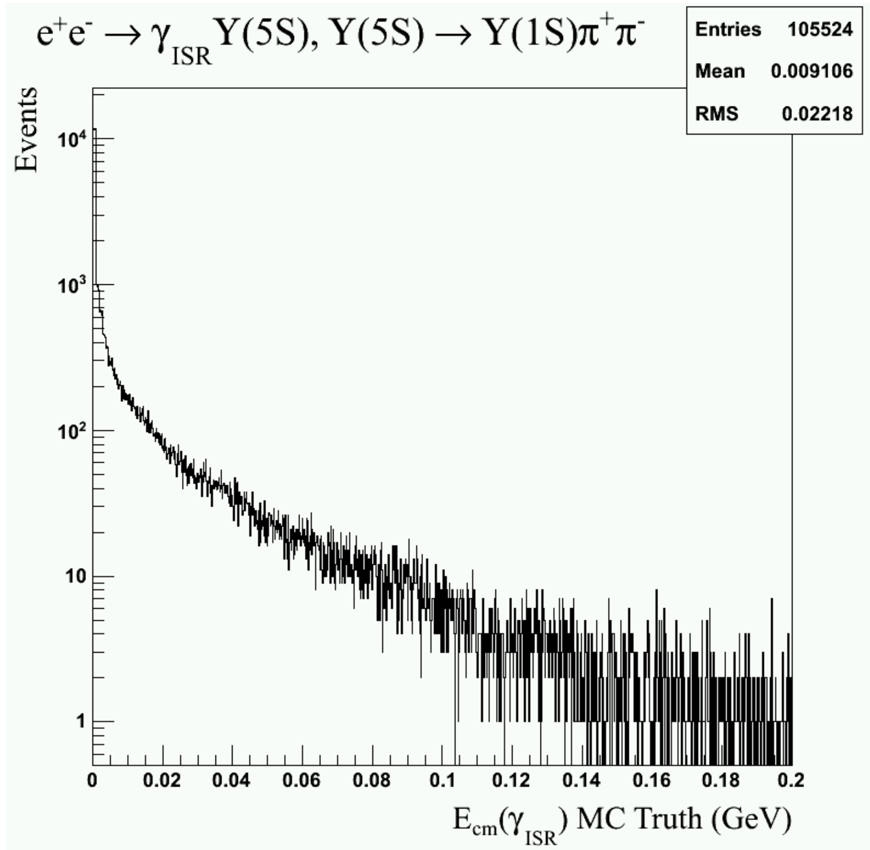



Figure 13: Reweighted ISR energy spectrum for $e^+e^- \to \gamma_{\rm ISR}\Upsilon(5S), \Upsilon(5S) \to \Upsilon(1S)\pi^+\pi^-$. Note that a log scale is used for the vertical axis.

$\Upsilon(1S)\pi^+\pi^-$ 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 distribution of $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$.

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 $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$, 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. 15 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 $\mu^+$ in Fig. 16 and find that the presence of ISR has only a small effect on the the angular distributions 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 $\Upsilon(5S) \to \gamma W_{bJ}$

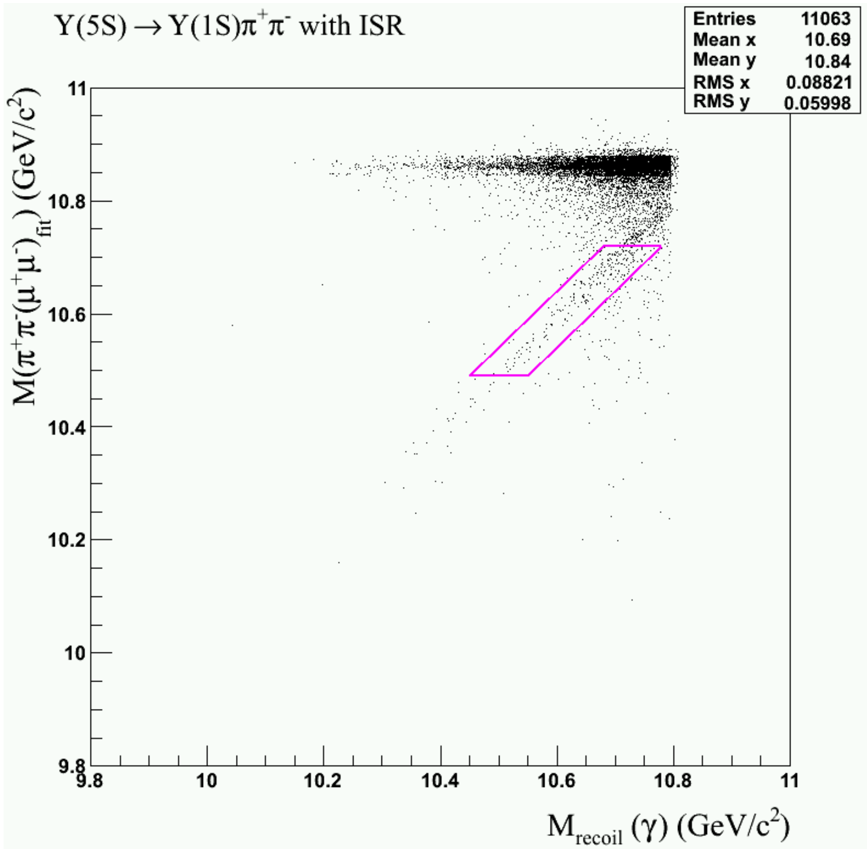



Figure 14: A 2-dimensional $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ vs $M_{\rm rec}(\gamma)$ distribution for $\Upsilon(5S) \to \Upsilon(1S)\pi^+\pi^-$ events with ISR (after reweighting). The signal region is outlined in magenta.

| 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.

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. 17 shows the distribution of these events within the signal region for our reweighted MC. To see how the selection on $\Delta E$ affects the background shape, we loosen up the selection on $\Delta E$ in the left and middle

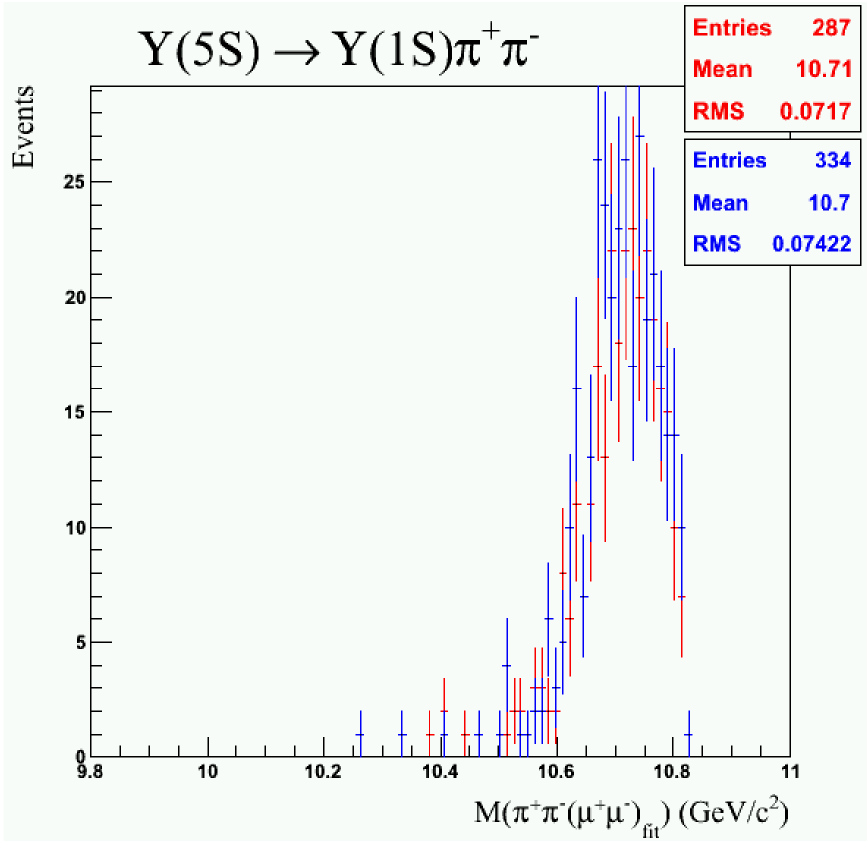



Figure 15: The distribution shown in blue is for events where $\Upsilon(5S) \to \Upsilon(1S)\pi^+\pi^-$ is generated using VVPIPI model [18] and $\Upsilon(1S) \to \mu^+\mu^-$ using VLL model [18]. The distribution shown in red is for events generated using PHSP model [18] for both processes. Neither samples contain ISR nor FSR, so they only differ by their decay models. The shapes of their $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ 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.

plots in Fig. 17. Imposing a selection on $\Delta E$ has only a small effect on the shape of these 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. 18 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 of grand sideband.

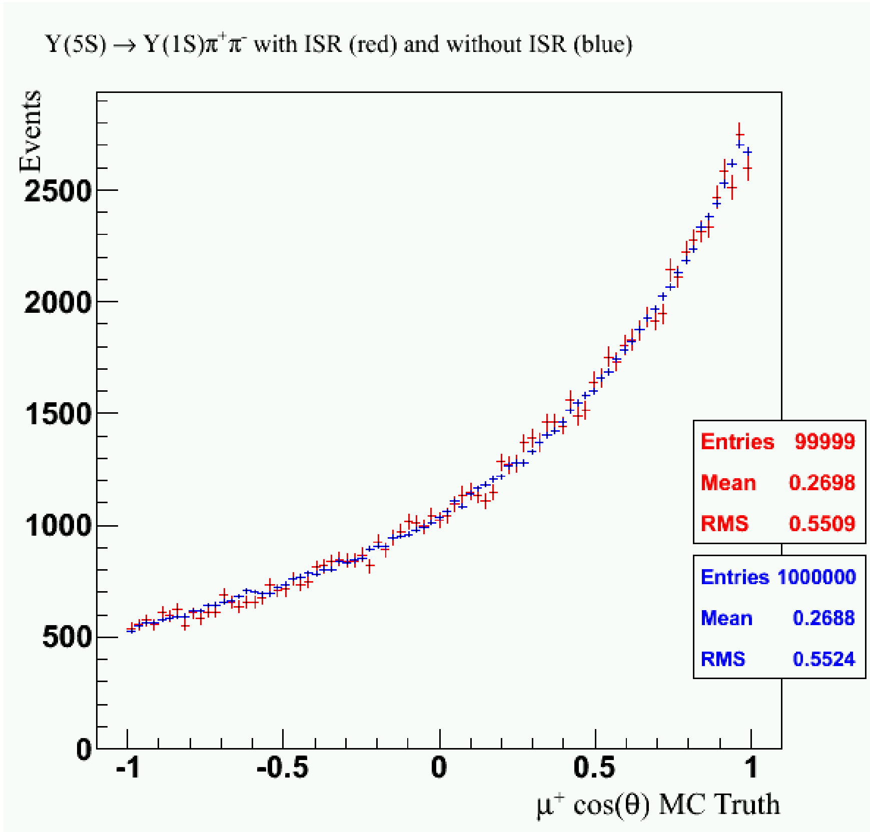



Figure 16: Distributions of $\cos\theta$ for $\mu^+$ for $\Upsilon(5S) \to \Upsilon(1S)\pi^+\pi^-$ events. The distribution shown is red is for events generated with ISR while the distribution shown in blue is for events generated without ISR. Events in both distributions are generated using PHSP model 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 | $10.72\ \mathrm{GeV}/c^2 < M(\pi^+\pi^-(\mu^+\mu^-)_{\mathrm{fit}}) < 10.80\ \mathrm{GeV}/c^2$<br>$-0.2\ \mathrm{GeV} < \Delta E < 0.2\ \mathrm{GeV}$ |
| Region 2 | $10.49\ \mathrm{GeV}/c^2 < M(\pi^+\pi^-(\mu^+\mu^-)_{\mathrm{fit}}) < 10.72\ \mathrm{GeV}/c^2$<br>$0.03\ \mathrm{GeV} < \Delta E < 0.2\ \mathrm{GeV}$ |
| Region 3 | $10.38\ \mathrm{GeV}/c^2 < M(\pi^+\pi^-(\mu^+\mu^-)_{\mathrm{fit}}) < 10.49\ \mathrm{GeV}/c^2$<br>$-0.2\ \mathrm{GeV} < \Delta E < 0.2\ \mathrm{GeV}$ |
| Excluded Region | $10.49\ \mathrm{GeV}/c^2 < M(\pi^+\pi^-(\mu^+\mu^-)_{\mathrm{fit}}) < 10.72\ \mathrm{GeV}/c^2$<br>$-0.2\ \mathrm{GeV} < \Delta E < 0.03\ \mathrm{GeV}$ |

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

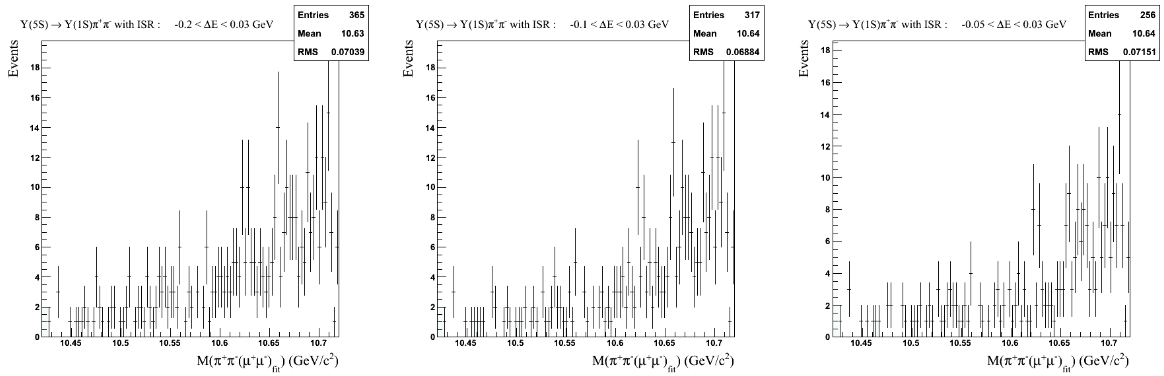



Figure 17: Distributions of $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ for $\Upsilon(5S) \to \Upsilon(1S)\pi^+\pi^-$ with ISR in the signal region for different $\Delta E$ requirements. The leftmost distribution requires $-0.2$ GeV $< \Delta E < 0.03$ GeV, the middle distribution requires $-0.1$ GeV $< \Delta E < 0.03$ GeV, and the rightmost distribution requires $-0.05$ GeV $< \Delta E < 0.03$ GeV. The upper bound of $\Delta E$ is kept at 0.03 GeV for all distributions, since very few signal events fall beyond $\Delta E > 0.03$ GeV.

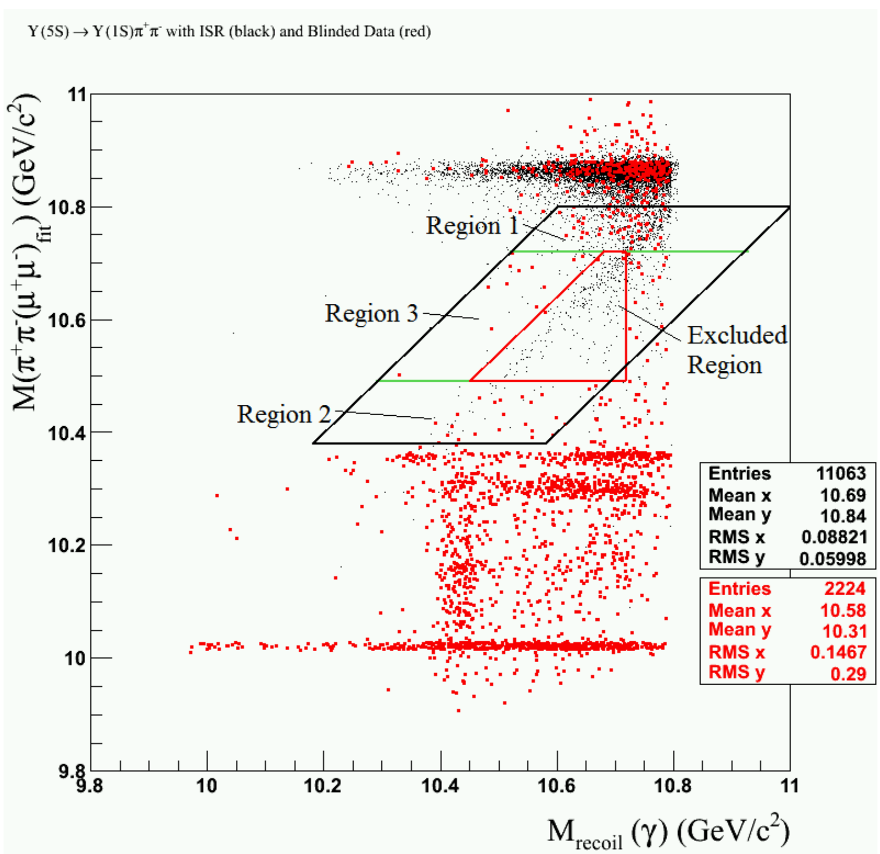



Figure 18: Subdivisions of the grand sideband region. The Excluded Region is not considered in this analysis.

| 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

# 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.

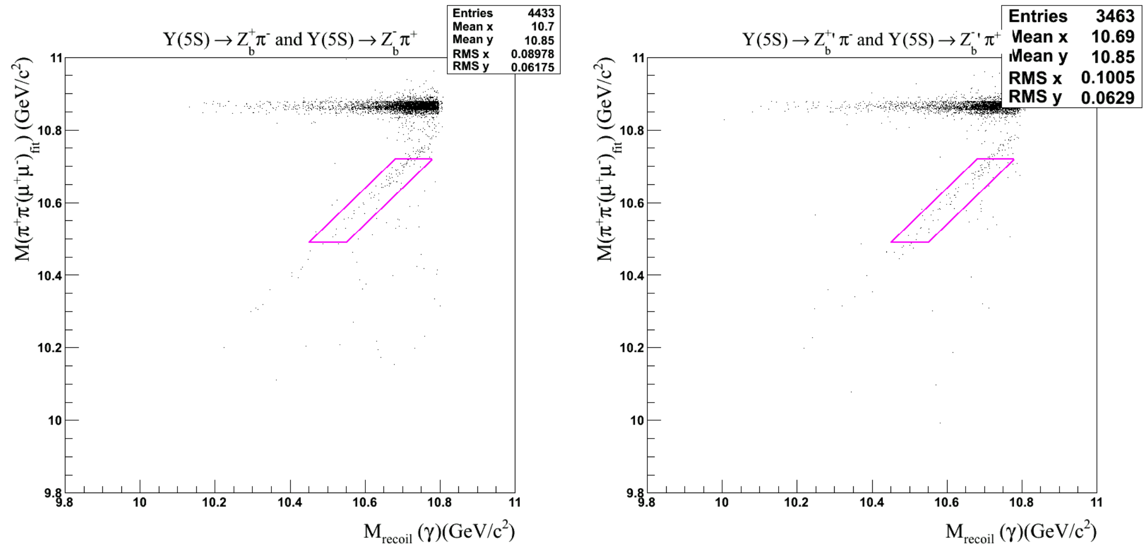
To estimate cross-feed between $Z_b$ and $W_{bj}$ analyses, we generated approximately 50,000 events for $\Upsilon(5S) \to Z_b^{\pm}\pi^{\mp}$ followed by $Z_b^{\pm} \to \Upsilon(1S)\pi^{\mp}$, $\Upsilon(1S) \to \mu^+\mu^-$. We also generated an additional 50,000 events for $\Upsilon(5S) \to Z_b'^{\pm}\pi^{\mp}$.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

The distribution of $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ vs $M_{\rm rec}(\gamma)$ is shown in Fig. 19 for both samples after applying our selection criteria for the $W_{bj}$ analysis. Fig. 20 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.

# 8 Fitting

## 8.1 Signal and Background PDFs

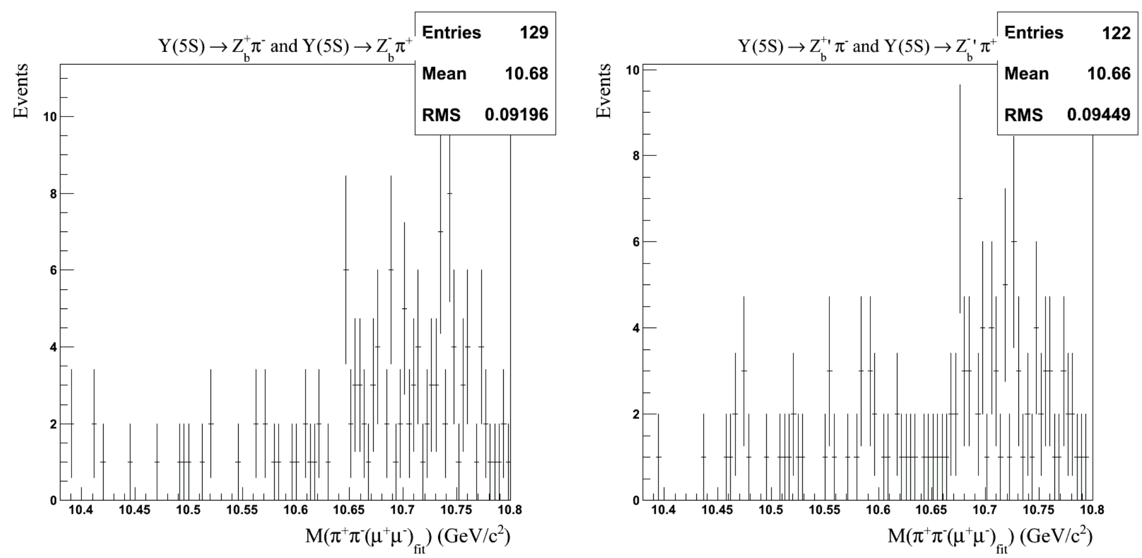
To extract signal yield, we perform a one-dimensional extended unbinned ML fit to the variable $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ using RooFit [22]. We model the signal distribution of $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ as a Breit-Wigner convolved with the sum of two Gaussians (to simulate effects of detector resolution as shown in Fig. 6). The observed width and shape of $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ distribution in signal MC remains practically the same after applying our $\Delta E$ 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




(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 19: 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.




(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 20: 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.

widths of the Gaussians used in convolution 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 the 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

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

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

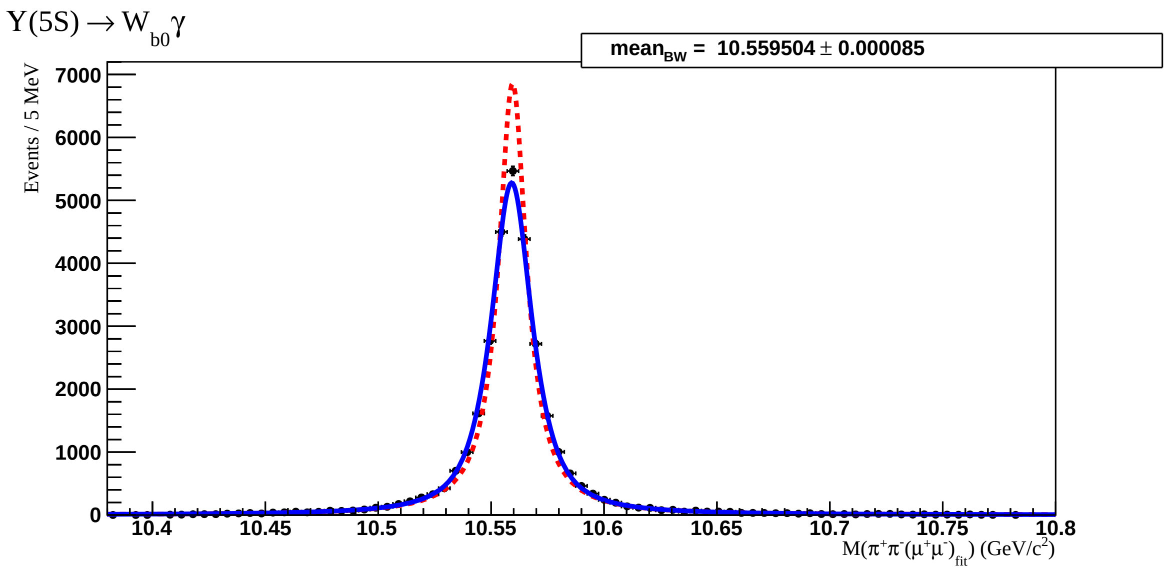
possible non-resonant contribution from dimuon continuum events. Strictly speaking, the background distribution deviates from an exponential at $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit}) \approx 10.75$ GeV/c$^2$. because of the phase space boundary at $M_{\rm rec}(\gamma) \approx 10.75$ GeV/c$^2$ seen in Fig. 5. 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 $\Delta E$ requirement (*e.g.* see Fig. 15, Fig. 21b, Fig. 21c). 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 perform 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 $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ 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 $59 \pm 11$ background events in the signal region and sideband regions combined. We expect $27 \pm 5$ 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. 21.

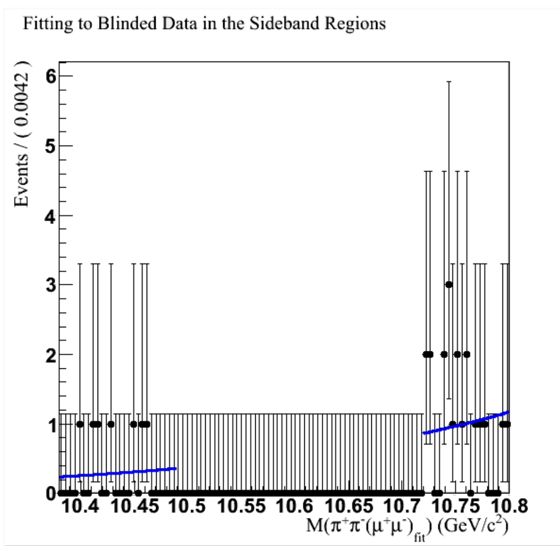
## 8.2 Confidence Belts

To construct a 90% confidence belt (with 5% on each side of the belt), we perform ensemble tests. Each ensemble test consists of 1000 toy MC experiments. In each toy MC experiment, we generate $N_{\rm sig}$ signal events and $N_{\rm bkg}$ 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$^2 < M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit}) <$ 10.80 GeV/c$^2$ to our combined signal and background PDF to extract the fitted number of signal events $N_{\rm sig}^{\rm fit}$.

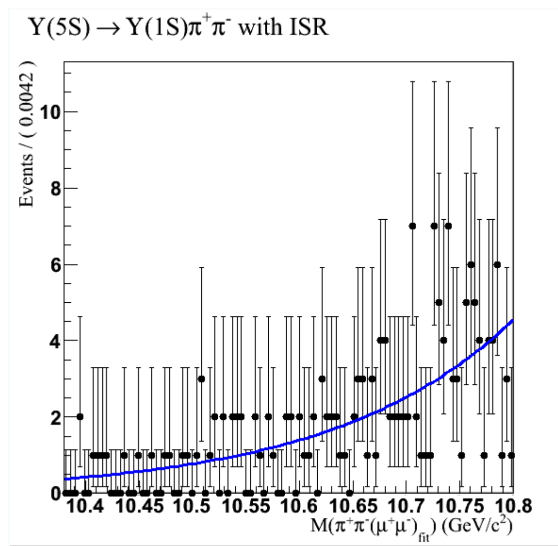
We construct our 90% confidence belt by performing ensemble tests with $N_{\rm bkg}^{\rm gen} = 59$ for values of $N_{\rm sig}^{\rm gen}$ from 0 to 70. We additionally construct a 90% confidence belt where we allow Poisson fluctuation in $N_{\rm bkg}^{\rm gen}$. These confidence belts are shown in Fig. 22.




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

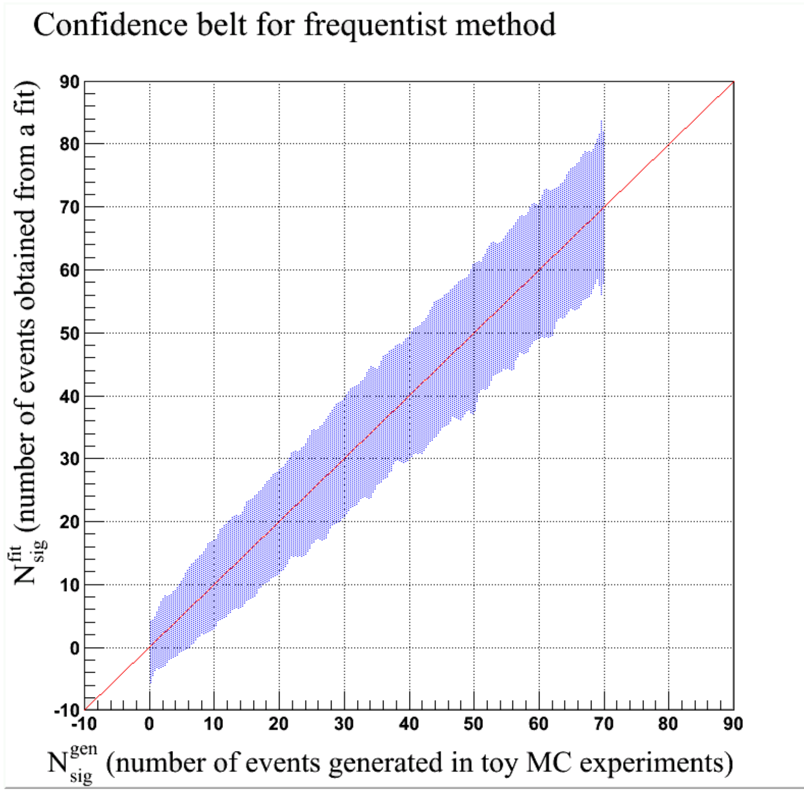



(c) Fit result for the distribution of $M(\pi^+\pi^-(\mu^+\mu^-)_{\mathrm{fit}})$ for $\Upsilon(5S) \to \Upsilon(1S)\pi^+\pi^-$ with ISR for events in the signal region.

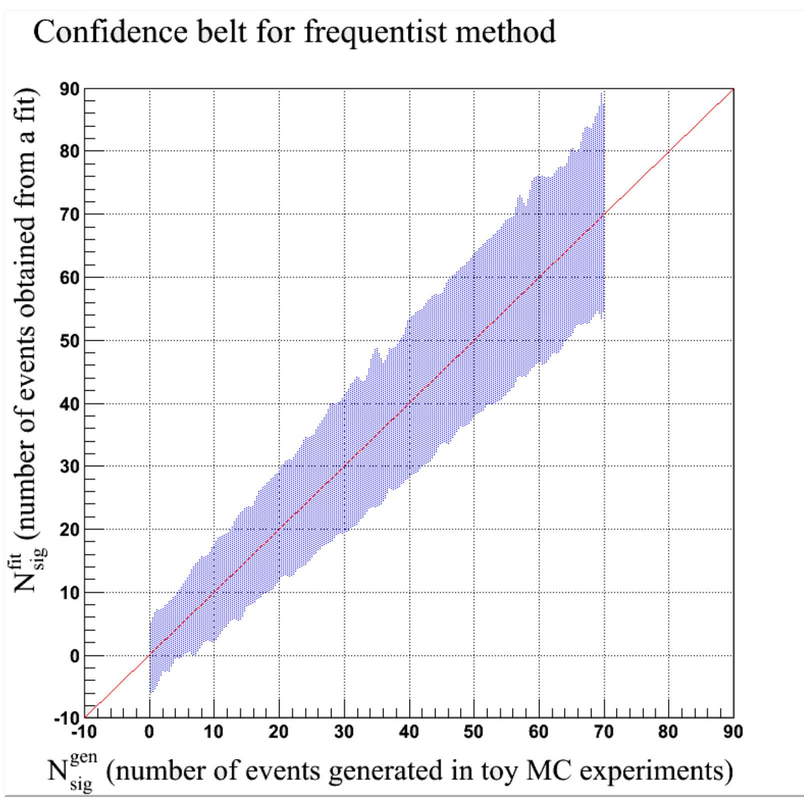
Figure 21: Fitting background MC and data

## 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 $N_{\mathrm{sig}}^{\mathrm{gen}}$ from 0 to 10 in steps of 1 and from 10 to 50 in steps of 5 while fixing $N_{\mathrm{bkg}} = 59$.




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




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

Figure 22: 90% confidence belts for frequentist method.

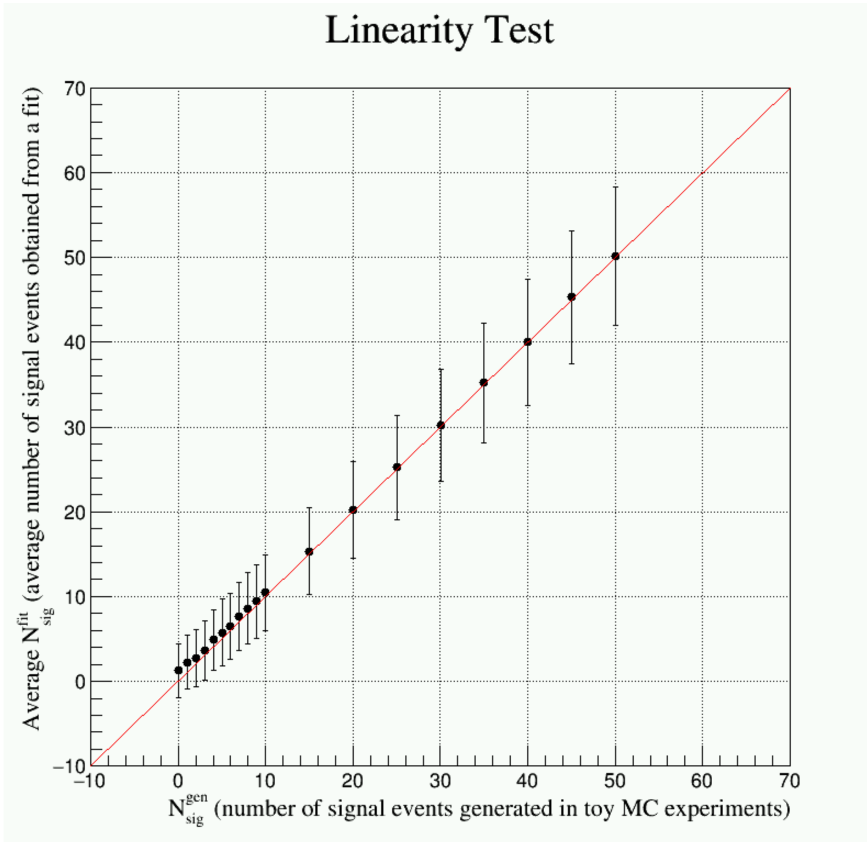
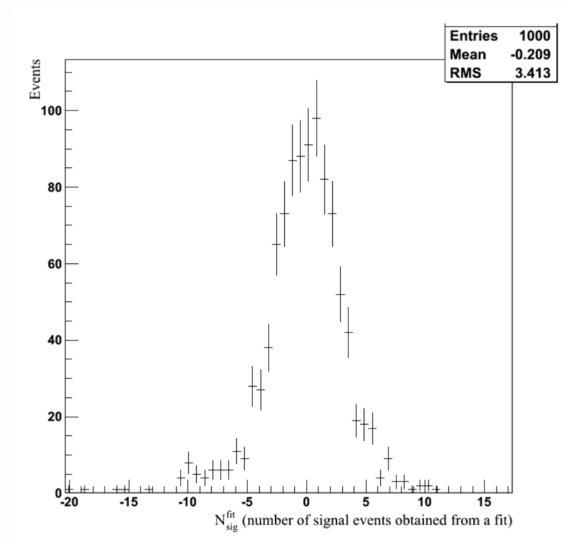



Figure 23: Average $N_{\text{sig}}^{\text{fit}}$ for varying values of $N_{\text{sig}}^{\text{gen}}$. The solid black line is the result of fitting these points to the linear function $f(x) = p_0 + p_1 x$. The resulting fit parameters are shown in the box on the top right.

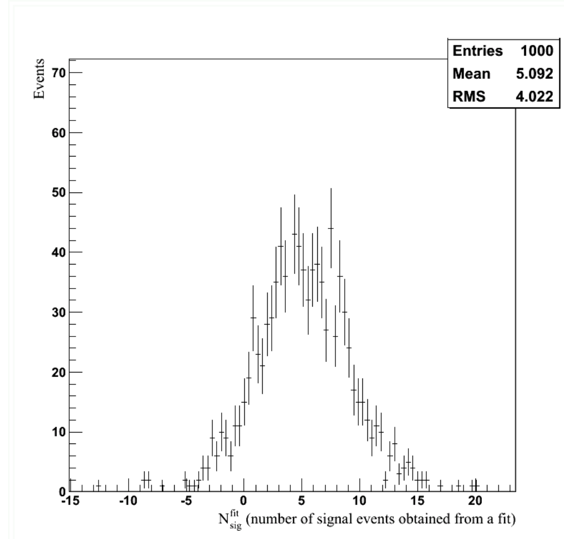
We plot the average number of signal events from the fit against $N_{\text{sig}}^{\text{gen}}$ and perform a fit to a linear function $f(x) = p_0 + p_1 x$. This plot and the results of the linear fit are shown in Fig. 23. Fig. 24 displays distributions of $N_{\text{sig}}^{\text{fit}}$ for certain values of $N_{\text{sig}}^{\text{gen}}$. When $N_{\text{sig}}^{\text{gen}}$ is large, the distribution of $N_{\text{sig}}^{\text{fit}}$ is unbiased. However, for small $N_{\text{sig}}^{\text{gen}}$, we see an asymmetry in the 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

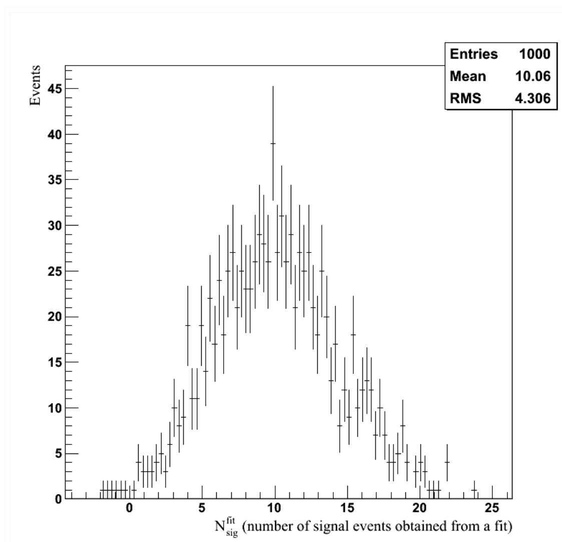
We estimate the upper limit on the branching fraction and visible cross section of $\Upsilon(5S) \to \gamma W_{bJ}$ in the absence of signal by performing an extended unbinned maximum likelihood fit on toy MC generated according to the fit to the data sidebands. We generate 1000 toy MC samples with 59 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 $-0.2 \pm 3.2$ events. Note that in Fig. 23, this average signal yield corresponds to the value plotted at $N_{\text{sig}}^{\text{gen}} = 0$. Using the confidence belt in Fig. 22, we determine the 95% confidence level upper limit on the number of signal events to be 10 events. We calculate the upper



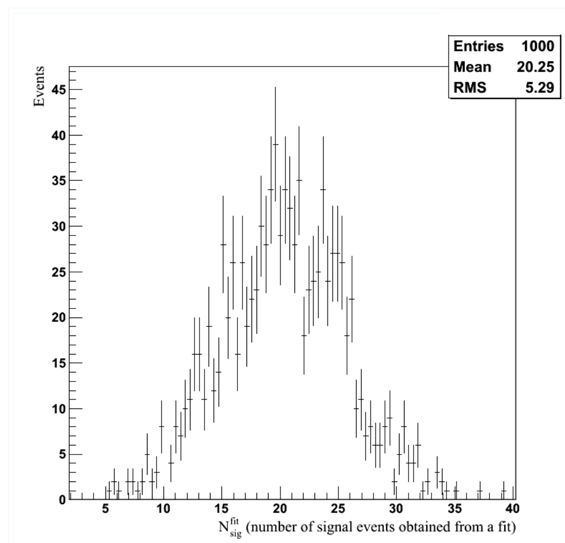

(a) Distribution of $N_{\text{sig}}^{\text{fit}}$ for an ensemble test with $N_{\text{sig}}^{\text{gen}} = 0$ and $N_{\text{bkg}}^{\text{gen}} = 59$.




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

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 $\epsilon$ is our reconstruction efficiency. Using Eq. 4, we determine the upper limit on the branching fraction in the absence of signal to be $2.4 \times 10^{-5}$.

We calculate the visible cross section using

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

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 $N_{\mathrm{sig}} = 10$, which is the 95% CL boundary of the 90% CL frequentist belt shown in Fig. 22 for $N_{\mathrm{sig}}^{\mathrm{fit}} = 3$, according to the result of the fit $N_{\mathrm{sig}}^{\mathrm{fit}} = -0.2 \pm 3.2$.

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

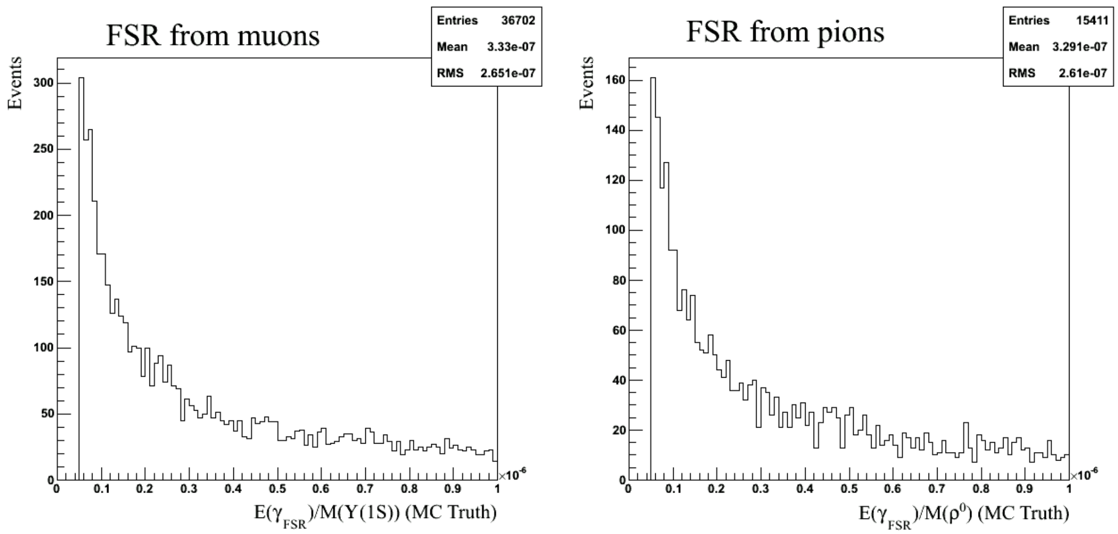
where $\mathcal{L}$ is the integrated luminosity. We find $\sigma_{\mathrm{vis}} = (0.115 \pm 0.006)$ fb. All values used to calculate the branching fraction and visible cross section are shown in Table 11.

# 9 Search Strategy 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' the data in the signal region and then fit a one-dimensional distribution of $M(\pi^+\pi^-(\mu^+\mu^-)_{\mathrm{fit}})$ using the aforementioned models for signal and background shapes. In the fit, we fix the width of $W_{bJ}$ to that of $Z_b$. Because we expect only one signal in our signal region, we plan to scan the range of invariant masses of $M(\pi^+\pi^-(\mu^+\mu^-)_{\mathrm{fit}})$ and, for each assumed value of the invariant mass, we perform a fit to data, where the background parameter $\lambda$ is allowed to float. If the fit returns a statistically significant result, we claim a discovery. We will then produce a plot of the upper limit versus mass of $W_bJ$. This plot will be produced regardless of whether or not the fit yields a significant result. Our confidence belt (Fig. 22) will be used 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 $\Upsilon(5S)$




(a) FSR from muons.

(b) FSR from pions.

Figure 25: Final state radiation from charged tracks

- 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:

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

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. 25. Because these quantities are bounded from below by XPHCUT $\cdot\, 0.5$, we prove that XPHCUT was successfully lowered.

## 10.2 Changes in the Analysis between Note v1.5 and v2.0

In this section we describe and explain the reasons for important changes we made in the analysis after Note v1.0 was released. These changes have not yet been applied to the main body of the text. All plots – except those in this and the next sections of this Note (v1.5) – have been made using the selection developed in the original analysis. The next version of this Note (v2.0) will have some of the critical plots and tables in the main section of the note updated to reflect for the changes described below.

Overall, there are four changes in the analysis related to (1) photon energy selection (extended), (2) the region blinded in data (increased), (3) selection criteria on $dr$ and $dz$ (relaxed) and (4) PID requirements for charged pions (removed).

The main improvement (extending signal photon energy spectrum) helps us to develop a robust and reliable approach to fitting the signal invariant mass spectrum (to be applied to data when the permission to open the signal region is secured). Extending the blinded region in data was done to avoid an annoying "undercoverage" demonstrated in Fig. 7 of Note v1.1, *i.e.* our decision to impose a requirement on $\Delta E$ was made after we had defined the blinded region, and hence, the top right corner of the signal region was not blinded originally. While no signal is expected in that corner, we would prefer to blind the entire signal region to simplify fitting. We deemed the changes in the selection applied to track impact parameters to be "right", so no additional systematics needs to be included in the result of the analysis. For the same reason we decided to remove PID cuts for charged pion candidates: any such selection criterion costs us some (even if very small) efficiency loss, and, more importantly, has some systematic uncertainty associated with it. We made the changes in PID and impact parameters selection just because we had to reskim the data and generic MC anyway.

The four changes are itemized (and elaborated more on) below, starting with the most important improvement in the analysis:

- **Signal Photon Energy**

  In the first version of this Note, as shown in Table 3, signal photon candidates were selected in the range between 100 and 600 MeV. As you can see in Fig. 4a of this version of the Note, this energy range is sufficient for signal photon selection, however, we found it to be very restrictive for purposes of fitting the signal invariant mass $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ including the sidebands. This effect is explained better in a dedicated subsection below.

- **Blinded Region in Data**

  As we already explained, our original blinding (which we decided about before we developed the fitting procedure) inadvertently exposed one corner of the signal region in data to possible inspection. This possibly introduces some bias, but more importantly

| Region Name | Boundary Definitions |
|---|---|
| The New Blinded Region | $10.49\,\mathrm{GeV/c^2} \leq M(\pi^+\pi^-(\mu^+\mu^-)_{\mathrm{fit}}) \leq 10.72\,\mathrm{GeV/c^2}$<br>$M_{\mathrm{rec}}(\gamma) \geq M(\pi^+\pi^-(\mu^+\mu^-)_{\mathrm{fit}}) - 0.04\ \mathrm{GeV/c^2}$<br>$M_{\mathrm{rec}}(\gamma) \leq 11.0\ \mathrm{GeV/c^2}$ |

Table 12: The new (wider) blinded region in data. The important change is shown in red color. However, it is redundant and adds nothing new as compared to the second line in this table.

makes the fitting of the $M(\pi^+\pi^-(\mu^+\mu^-)_{\mathrm{fit}})$ distribution slightly more difficult. As we had to reskim the data and repeat all analysis steps, we decided to extend the blinded region as shown in Table 12. Please compare this with Table 5 in the original version of the note. In one sentence, we blinded the entire right-side tail of the $M_{\mathrm{rec}}(\gamma)$ distribution. Note that the change shown in red color in this table is not even necessary: the second line of the boundary described in the table is sufficient to achieve our goal. We show the third line in the table only so the comparison with the previous version of this Note is easier to make.

- **Impact Parameters for Charged Tracks**

  The $dz$ and $dr$ selection criteria have been loosened to $\leq 4$ cm and $\leq 0.5$ cm, respectively.

- **PID Requirements for Charged Pion Candidates**

  These requirements have been eliminated.

As we already mentioned, these four changes in the selection criteria required us to reskim generic MC and data. That was easy.

### 10.2.1 Signal Photon Energy Conundrum

Well, retrospectively, extending signal photon energy selection was not really a very difficult decision, but it requires a thorough explanation. Below we try our best to walk the reader through the logic of our decision.

In our analysis we extract the signal yield by fitting the distribution of the signal invariant mass $M(\pi^+\pi^-(\mu^+\mu^-)_{\mathrm{fit}})$. We are confident (because we proved this) that we observe the ISR background (*i.e.* events where the production of $\Upsilon(5S)$ is accompanied by some initial state radiation). However, we suspect that there are other sources of non-peaking background, such as, *e.g.* poorly reconstructed events of all possible types, cosmic events overlapping with incompletely reconstructed collision events – you name it – present in data. The key part of our approach to fitting is that, on basis of our extensive and thorough studies of non-signal data and generic MC, we expect no peaking backgrounds to be present in the signal region.

For as long as no bias is present in selecting signal event candidates, background events of ISR origin are relatively well described (as you will see for yourself very soon) by the sum of an exponential and a straight line of non-negative slope. Small non-peaking background is likely to be sufficiently-well approximated also by the same straight line (of zero or positive slope). However, our original selection criteria strongly suppressed ISR background at large values of the signal invariant mass $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$, making it very difficult to reliably obtain the shape of such biased background distribution using sidebands in data. We realized that, in order to significantly reduce such possible bias, we have to avoid suppressing ISR background in the sideband region.

To demonstrate the effect we are trying to explain in this section, we generated a somewhat ridiculous MC sample, where an incredibly broad "structure" (an almost flat distribution of the invariant mass called, for purposes of MC production, "$\Upsilon(5S)$" (which it is most definitely not!)) was generated along with ISR in the $e^+e^-$ annihilation followed by this structure's "decay" to $\Upsilon(1S)\pi^+\pi^-$. Applying our selection criteria to such MC sample after detector simulation and reconstruction allows us to investigate the phase space of relevant kinematic parameters at sufficient level of precision to make meaningful conclusions. Note that we do not even try to reweight the ISR energy spectrum in this exercise, because all we need for our studies is a good coverage of phase space (which is already a good enough of a reason NOT to reweight such MC!).

We start by demonstrating, in Fig. 26, that, with the original signal photon energy selection, the reconstructed ISR background does not resemble an exponent in the signal plus sideband region of the invariant mass $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ (indicated by vertical lines in this figure). The explanation for the observed shape lies in the cut on signal photon energy requiring at least 100 MeV. Further, Fig. 27, where we show the 2D distribution of the reconstructed signal photon energy (in the lab) vs $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$, demonstrates that the range of the invariant mass $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ is actually biased on both ends of the spectrum – at higher masses because of the 100 MeV cut, at smaller masses because of the 600 MeV cut. Note that the relevant value of the invariant mass, where the ISR background is suppressed by the 100 MeV cut, is located at the intersection of the left-side of the "opening cone" of phase space with the horizontal line of the 100 MeV cut on the photon energy. Note that the opening angle of the cone describing the phase space is due to the $\Delta E$ selection, which we keep to be $-0.05\,\mathrm{GeV} \le \Delta E \le 0.03\,\mathrm{GeV}$.

In order to avoid the described bias in background photon energy spectrum and to be able to use the higher-mass sideband to perform a robust fit to such background, we release the cut on signal photon energy to the lowest possible value of 20 MeV (standard Belle reconstruction and MDST production do not go lower than that). Note that in our analysis we do not really care about the origin of such low-energy photons (and if they are really such) and possible energy dependence of photon reconstruction efficiency systematic uncertainty, because our signal is associated with photons of higher energy, but we need (even if smoothly suppressed) an exponential-like energy distribution of background photons to make extracting the signal from data reliable. Really low energy photon candidates could be even of instrumental origin, *e.g.* being due to noise in calorimeter electronics. This does not matter, as in our approach we obtain the shape of background distribution from data.

Interestingly, to improve our understanding of backgrounds, we also have to raise the cut on the other end of the photon energy region, though, in this case, for a different reason. As is explained in Table 6 and Fig. 10 of the original version of the note, there is a particular peaking background, namely, radiative (*i.e.* ISR) production of $\Upsilon(3S)$ followed

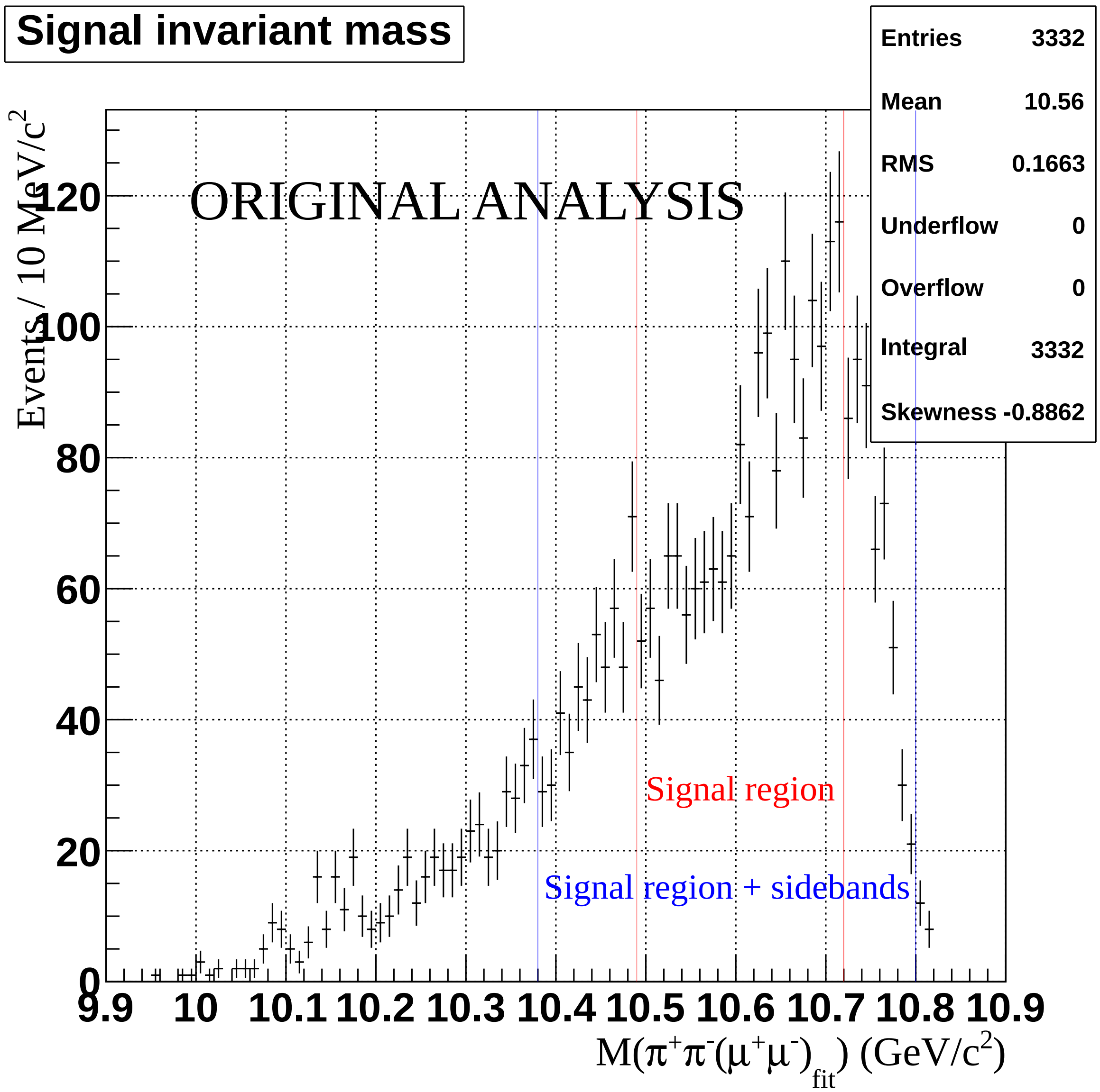



Figure 26: The reconstructed invariant mass $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ for a special ISR MC sample with the original selection criteria. Only best candidates are shown.

by its dipion transition to $\Upsilon(1S)$, which is uncomfortably too close to the left side of our signal plus sideband region of $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$. When both charged pions and the photon are misreconstructed, sometimes this unfortunate happenstance might shift some of such these events into the signal region. Therefore, it would be best to investigate this possible background using the data. This goal requires us to release the signal photon energy cut. For practical purposes, in the improved version of the analysis, we limit signal photon energy to 5 GeV. Note that our approach also facilitates possible measurement of ISR production of $\Upsilon(3S)$ and $\Upsilon(2S)$, which could be used to calibrate ISR MC.

After widening the signal photon candidate energy selection as described and explained, we plot the distributions of photon energy spectrum vs $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ in Fig. 28 and $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ in Fig. 29 (as always, for best candidates only) for MC events from our

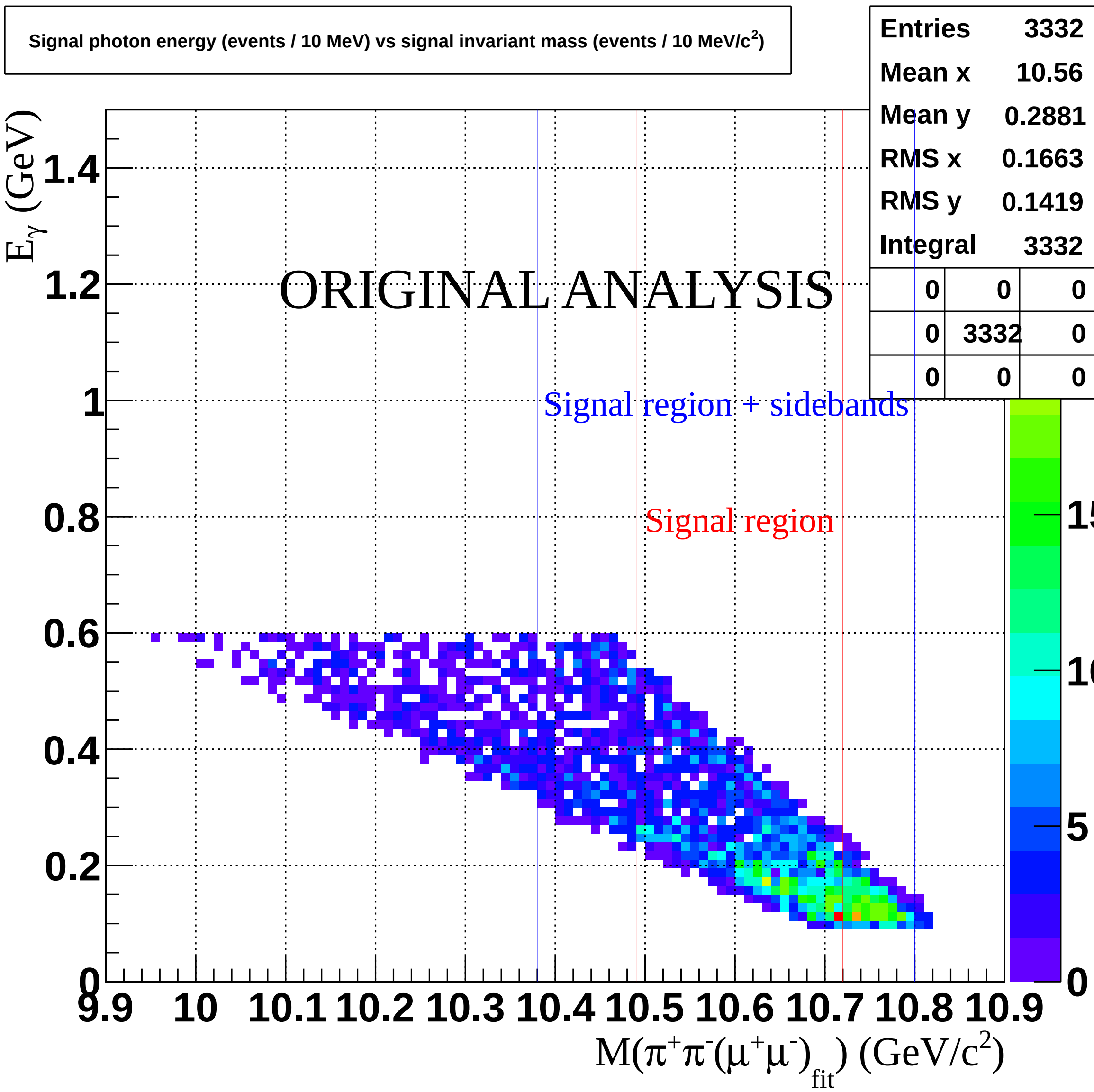



Figure 27: The reconstructed signal photon energy versus the invariant mass $wbjm$ for a special ISR MC sample with the original selection criteria. Only best candidates are shown.

"ridiculous" background MC sample. We conclude that the current (*i.e.* new, relaxed) selection criteria allow us to perform a robust fit to the background using the sidebands of $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ for the values of this variable up to 10.78 GeV/c$^2$. To further investigate the shape of $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ in this special MC sample, we plot the distribution of $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ using the logarithmic scale in Fig. 30. We observe that the distribution shown in the figure does not follow a simple exponential dependence on $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ in part because, as explained previously, the ISR spectrum in this MC sample is completely unreasonable.

We reskim the data and generic background MC, blind the signal region and present the data distributions of photon energy spectrum vs $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ in Fig. 31 and $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ in Fig. 32. We observe unambiguous signatures of $\Upsilon(2S)$ and $\Upsilon(3S)$

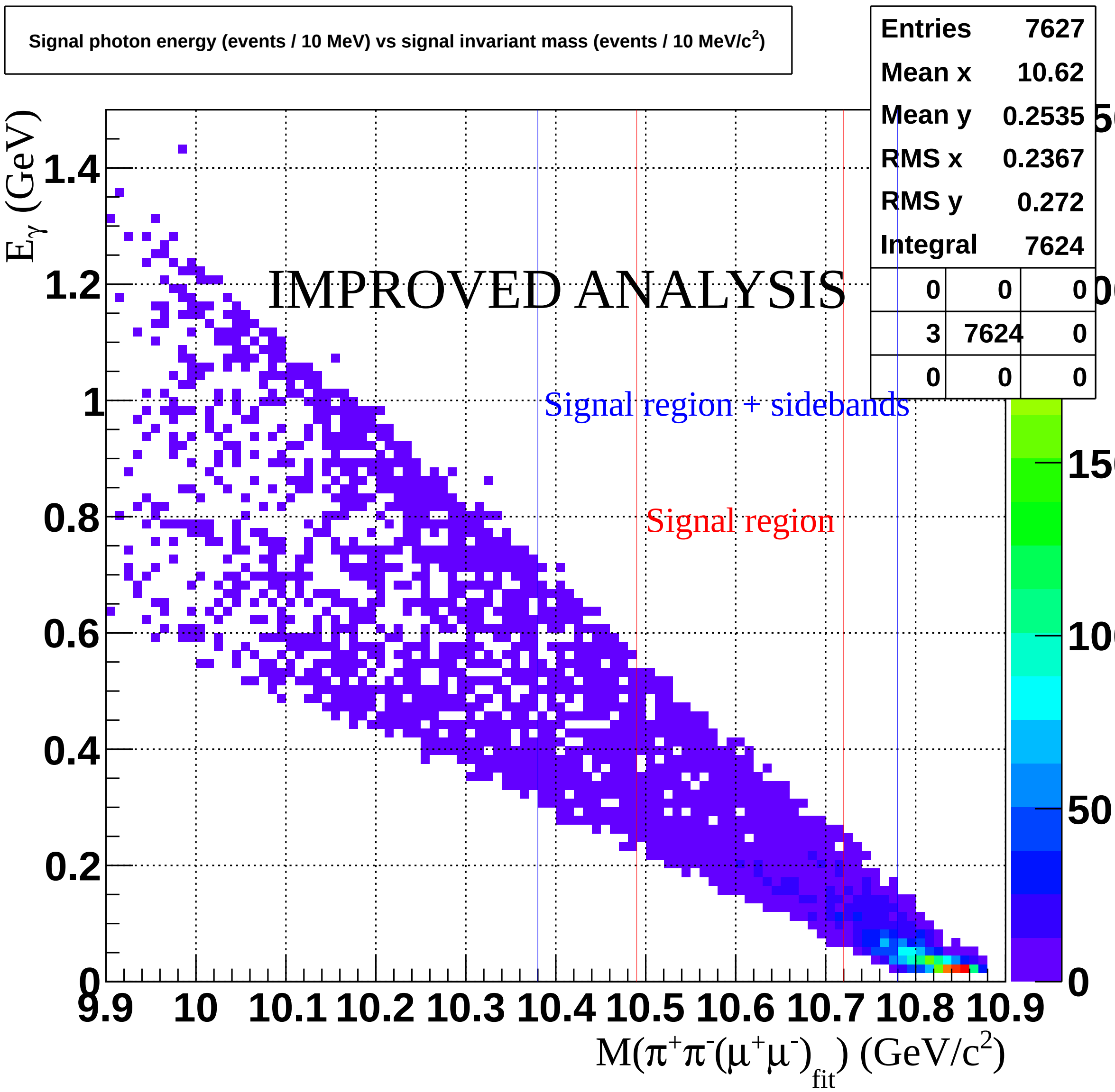



Figure 28: The reconstructed signal photon energy versus the invariant mass $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ for a special ISR MC sample with relaxed selection criteria. Only best candidates are shown.

ISR production. We also show the distribution of $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ for blinded data in Figures 33 and 34 using the logarithmic scale. Note that the last figure is plotted using a finer bin width of 2 MeV/c$^2$ in a narrower range of the invariant mass. Using the data, we perform a rough estimate of the width of the peak seen at the nominal mass of $\Upsilon(3S)$, 10.355 GeV/c$^2$, corresponding to events originating from radiative return to $\Upsilon(3S)$. Using our estimate of 5 MeV/c$^2$ (consistent with our MC-based understanding of resolution), we conclude that events in the left sideband of Fig. 34 are at least 5 width units away from this peak. Hence, it is unlikely that events in our left sideband are from radiative return to $\Upsilon(3S)$.

In the next section we explain our fitting strategy for extracting the signal from data

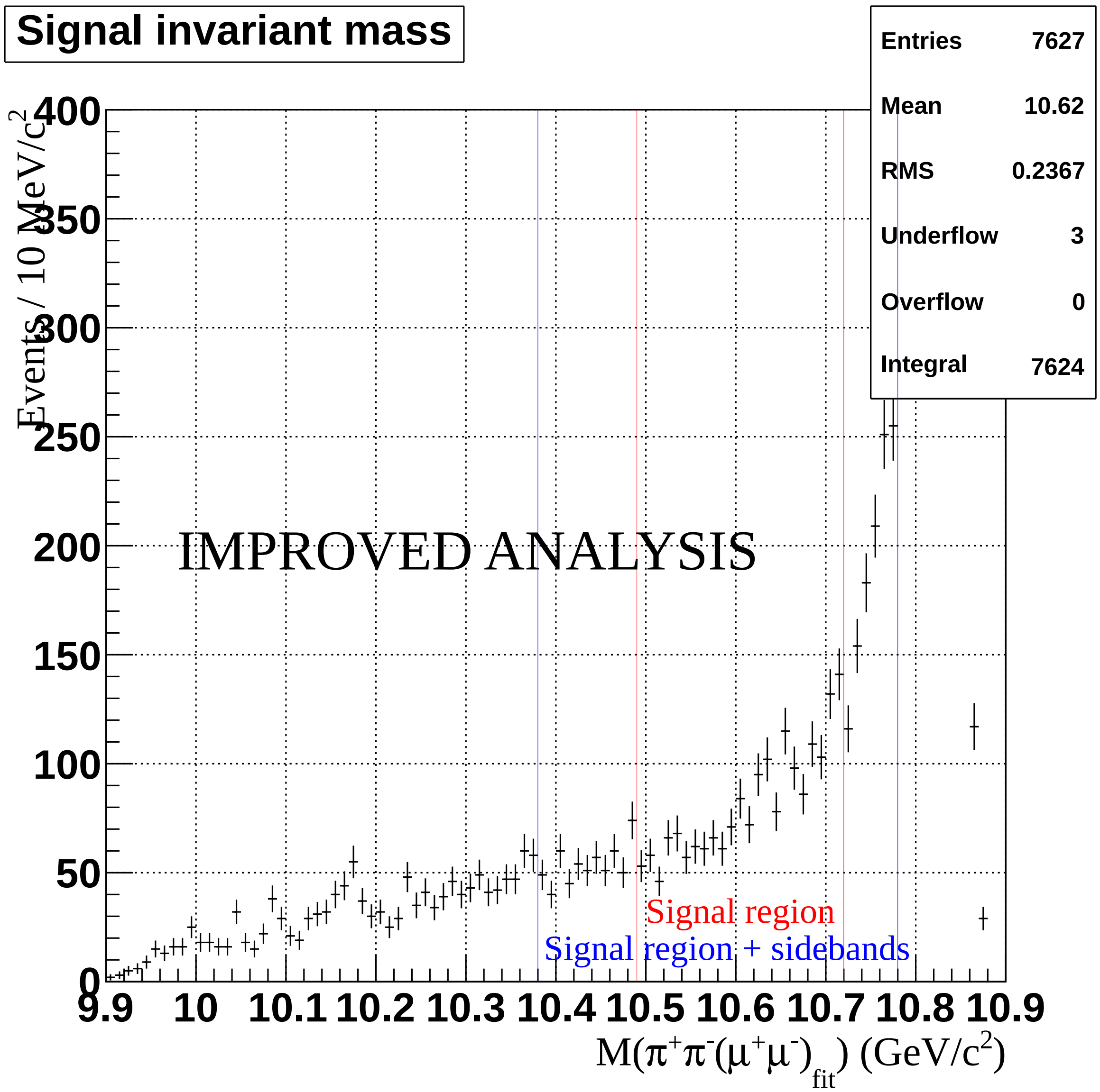



Figure 29: The reconstructed invariant mass $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ for a special ISR MC sample with relaxed selection criteria. Only best candidates are shown.

(when the permission to unblind is granted). We fit the data in the range of signal invariant mass $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ between the two blue vertical lines shown in Figures 27–34, *i.e.* in the range 10.38 GeV/c$^2$ $\leq$ $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ $\leq$ 10.78 GeV/c$^2$. This region of the invariant mass in blinded data is shown in Fig. 35. Systematic uncertainty due to the choice of the fitting region is investigated by varying the boundaries of this interval.

Finally, to conclude this section, we present Figures 36–39, where we show the distributions of the energy of the signal photoon candidate versus $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ and the projections onto the $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ invariant mass for our correctly reweighted ISR MC sample (described in section 6) for the decay $\Upsilon(5S) \to \Upsilon(1S)\pi^+\pi^-$.

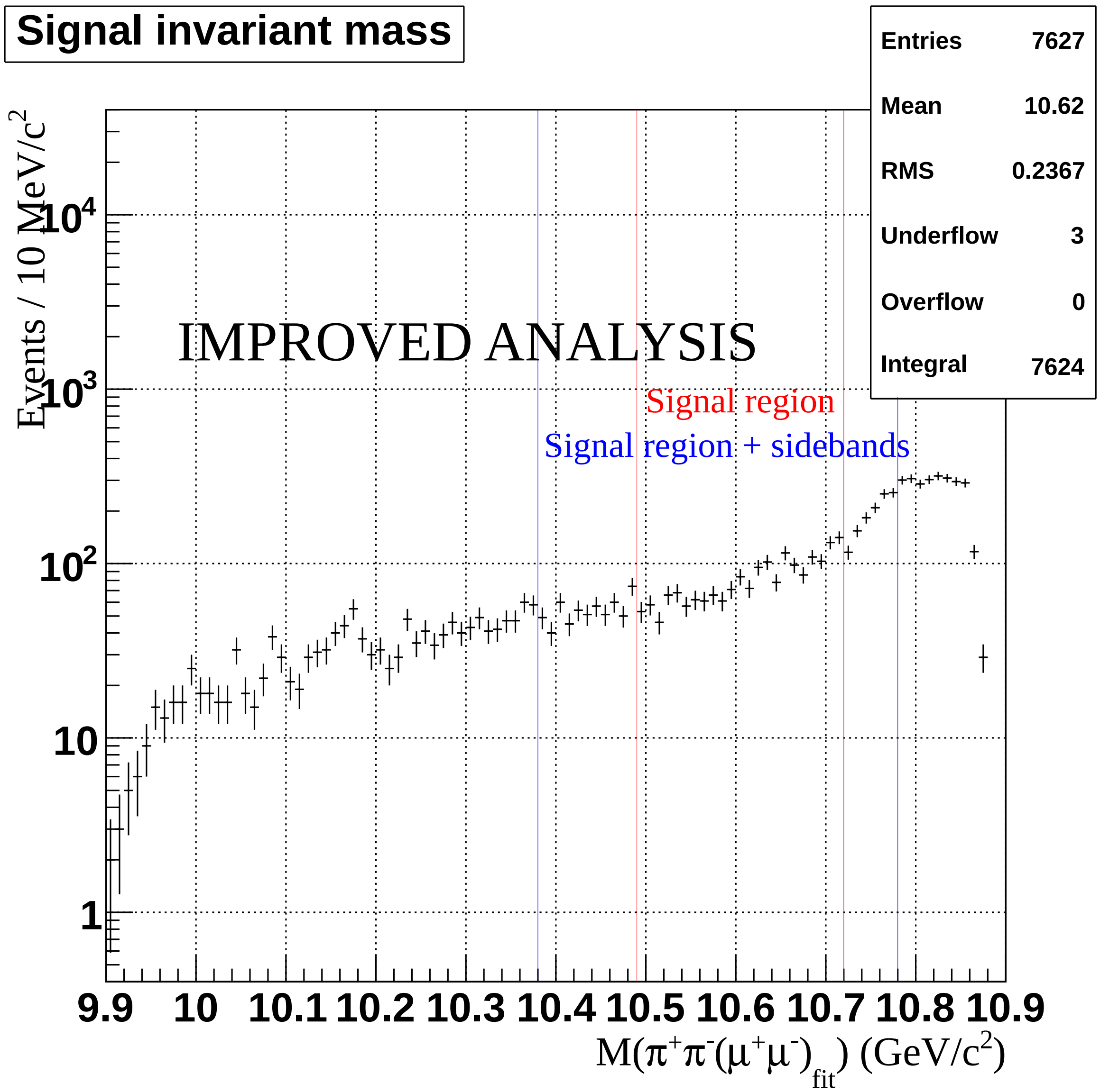



Figure 30: The reconstructed invariant mass $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ for a special ISR MC sample with relaxed selection criteria shown using the logarithmic scale. The ISR spectrum in this special MC sample has an unreasonable shape and could not be described by an exponential. Only best candidates are shown.

## 10.3 Fitting Strategy

To extract the $W_{bJ}$ signal from the data or to estimate the upper limit on its production, we (plan to) fit the invariant mass distribution $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ in data with the sum of an exponential, a straight line of zero or positive slope and the model for the signal shown in Fig. 21a. We plan to perform such fits for the values of (fixed) nominal mass of $W_{bJ}$ between 10.5 and 10.7 GeV/c$^2$ in steps of a few MeV/c$^2$. Important reference points here are provided by the invariant masses of $Z_B$ and $Z_b'$ which are, respectively, 10.610 and 10.650 GeV/c$^2$. We expect (or, rather, M. Voloshin expects) $W_{bJ}$ to be roughly as wide (or narrow) as $Z_b^{(\prime)}$. This

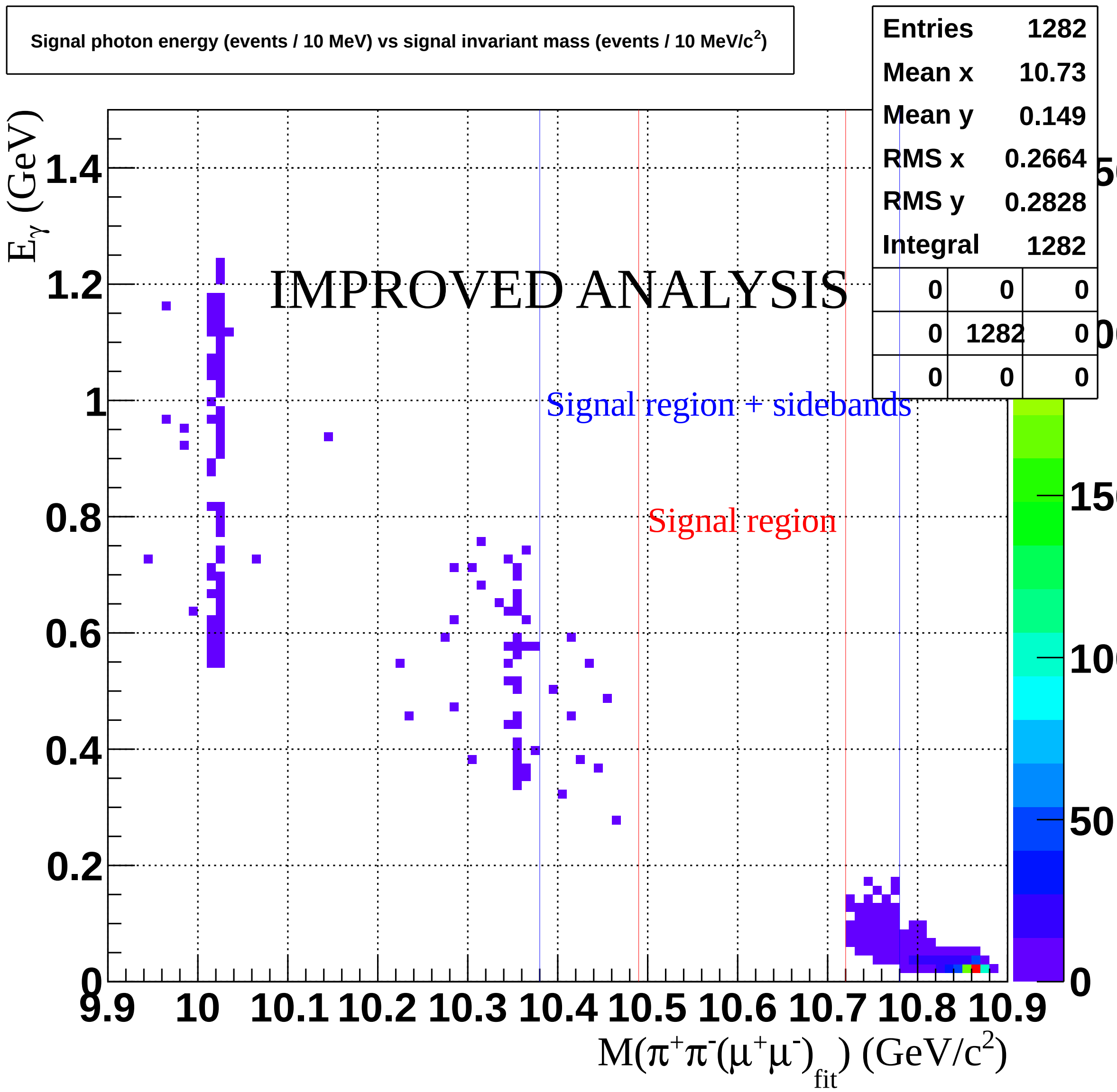



Figure 31: The reconstructed signal photon energy versus the invariant mass $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ for blinded data with relaxed selection criteria. Only best candidates are shown.

makes our life easier. Each fit will be performed independently. The shape of background distribution will be obtained from the data including the signal region. In our opinion, we can not obtain the shape of background exclusively from the sidebands because our sidebands are relatively narrow and, also, if the shape of the background function is fixed using our sidebands, fitting with the model described in this section could easily introduce a significant bias in the results of the fit. Another limitation comes from the wide range of the invariant mass region where we are searching for the $W_{bJ}$. For each individual fit (with a particular hypothesis for $W_{bJ}$ mass), the effective sideband region is going to be significantly wider than in our exercises discussed in this section. The key assumptions are: 1) there are no peaking backgrounds in the entire signal region, and 2) backgrounds can be modeled by the sum of

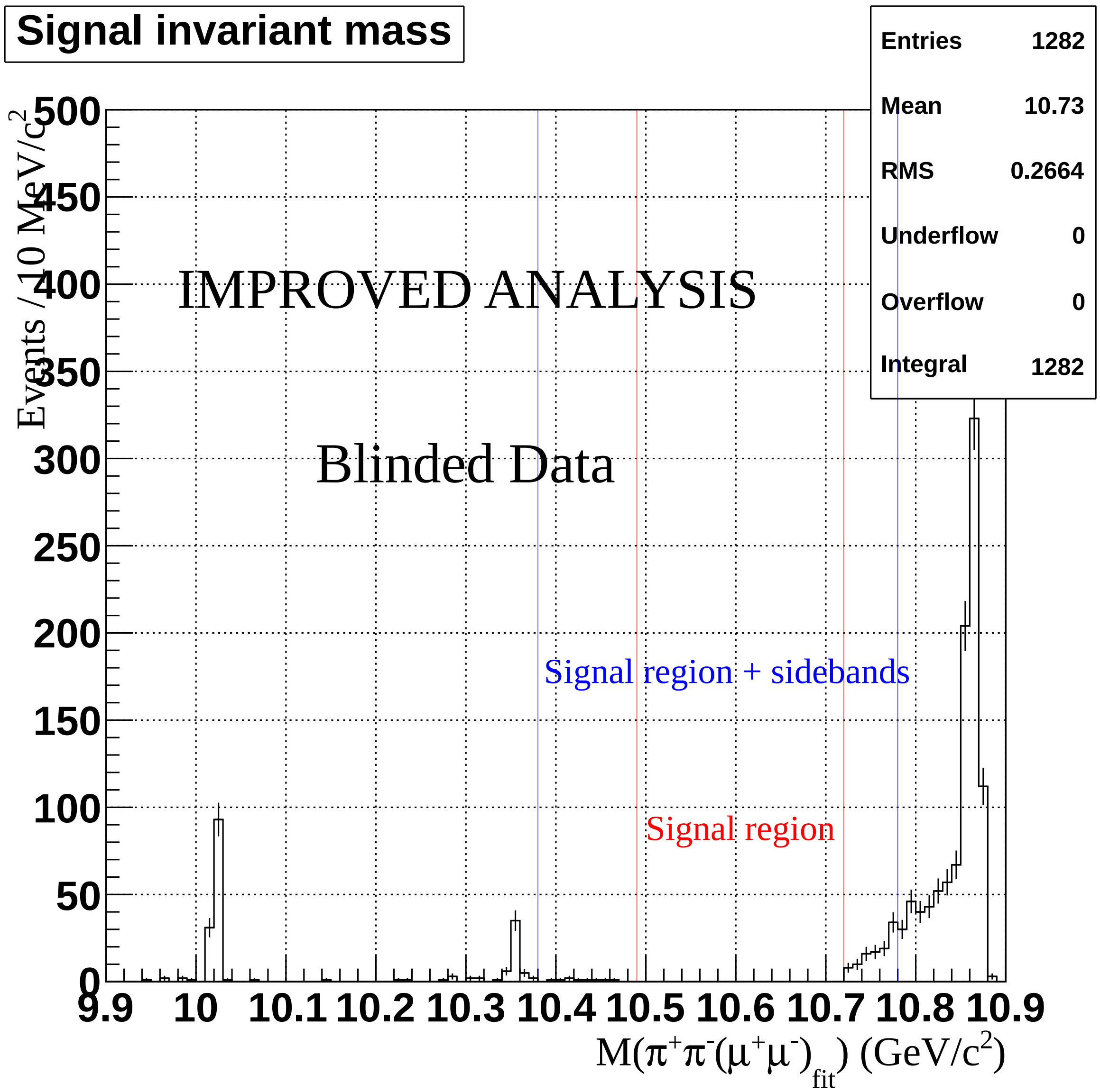



Figure 32: The reconstructed invariant mass $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ for blinded data with relaxed selection criteria. Only best candidates are shown.

an exponential and a straight line. Our confidence is based on MC studies using, first of all, our ISR MC samples.

We fit the invariant mass $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ in the range between 10.38 GeV/c$^2$ $\leq$ $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit}) \leq$ 10.78 GeV/c$^2$. In principle, we can (significantly) extend the invariant mass included in the fit toward smaller values (therefore including the radiative production of $\Upsilon(3S)$ or even $\Upsilon(2S)$ in our fits), however, it is not clear to us if this would necessarily help us understand the shape of the background and to reduce the uncertainty in our model description of the data in the signal region.

In this section we show some of the results of our unbinned extended maximum likelihood fits to $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ using a model implemented using RooFit for various MC samples under different conditions.

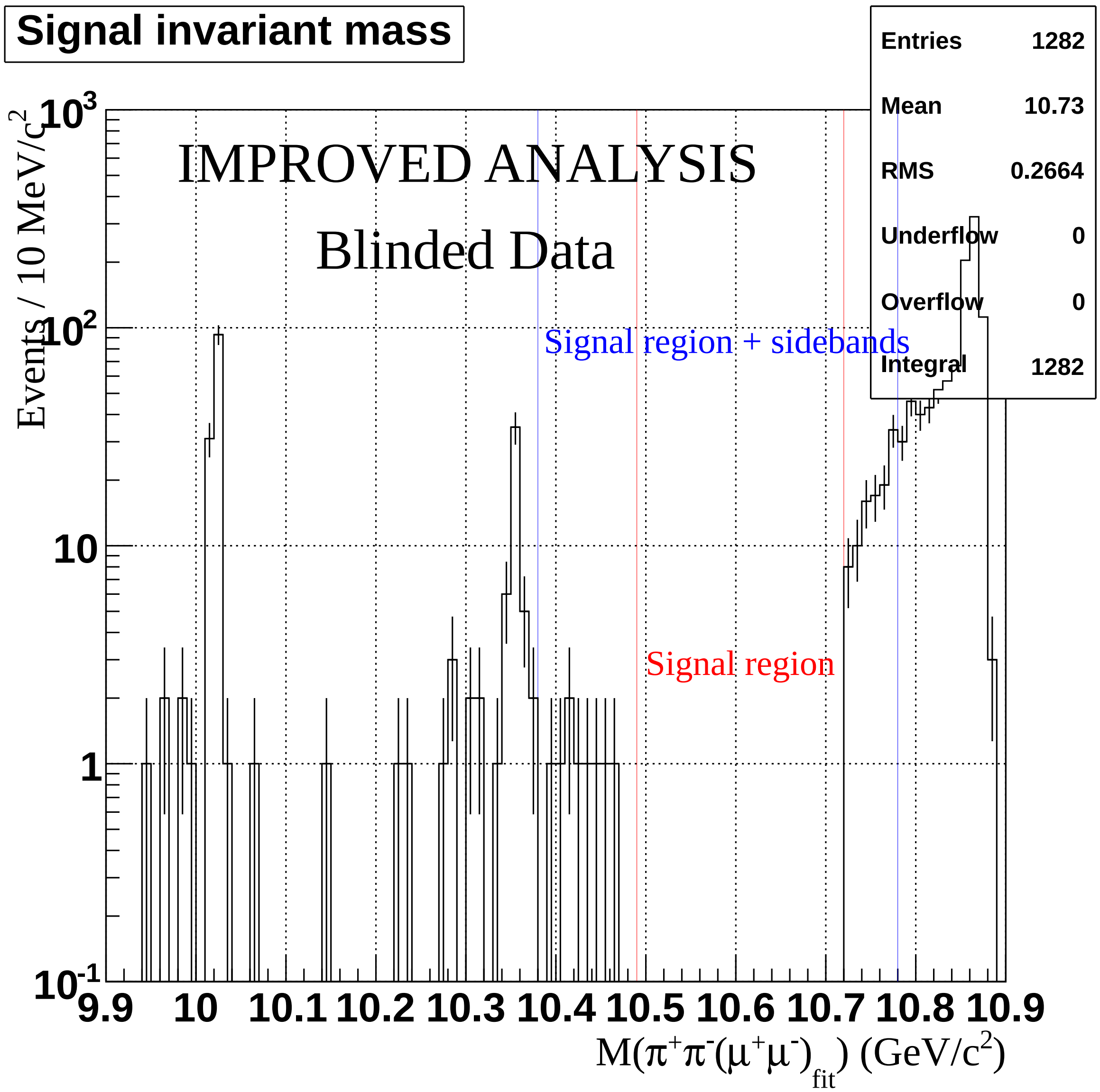



Figure 33: The reconstructed invariant mass $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ for blinded data with relaxed selection criteria plotted using the logarithmic scale. This is the same distribution as shown in Fig. 32. Only best candidates are shown.

We start our adventure by fitting the distribution of properly reweighted ISR MC sample for the decay $\Upsilon(5S) \to \Upsilon(1S)\pi^+\pi^-$ shown in Fig. 39. We fit this distribution using the unbinned extended maximum likelihood technique implemented in RooFit with the sum of an exponential and a straight line of zero or positive slope (if the curious reader really wants to know, we use RooChebychev for the latter). The results of this fit are shown in Fig. 40. We show the results of the fit using both the linear and the logarithmic scales because while one PDF is linear when plotted on log scale, the other PDF is, surprise, linear when plotted on linear scale because it is a line! Note that the fit has four parameters: $\alpha$ is the parameter of the exponent, slope is the slope of the straight line, $N1$ and $N2$ are the numbers of events obtained from the fit for background contributions parameterized by the exponent

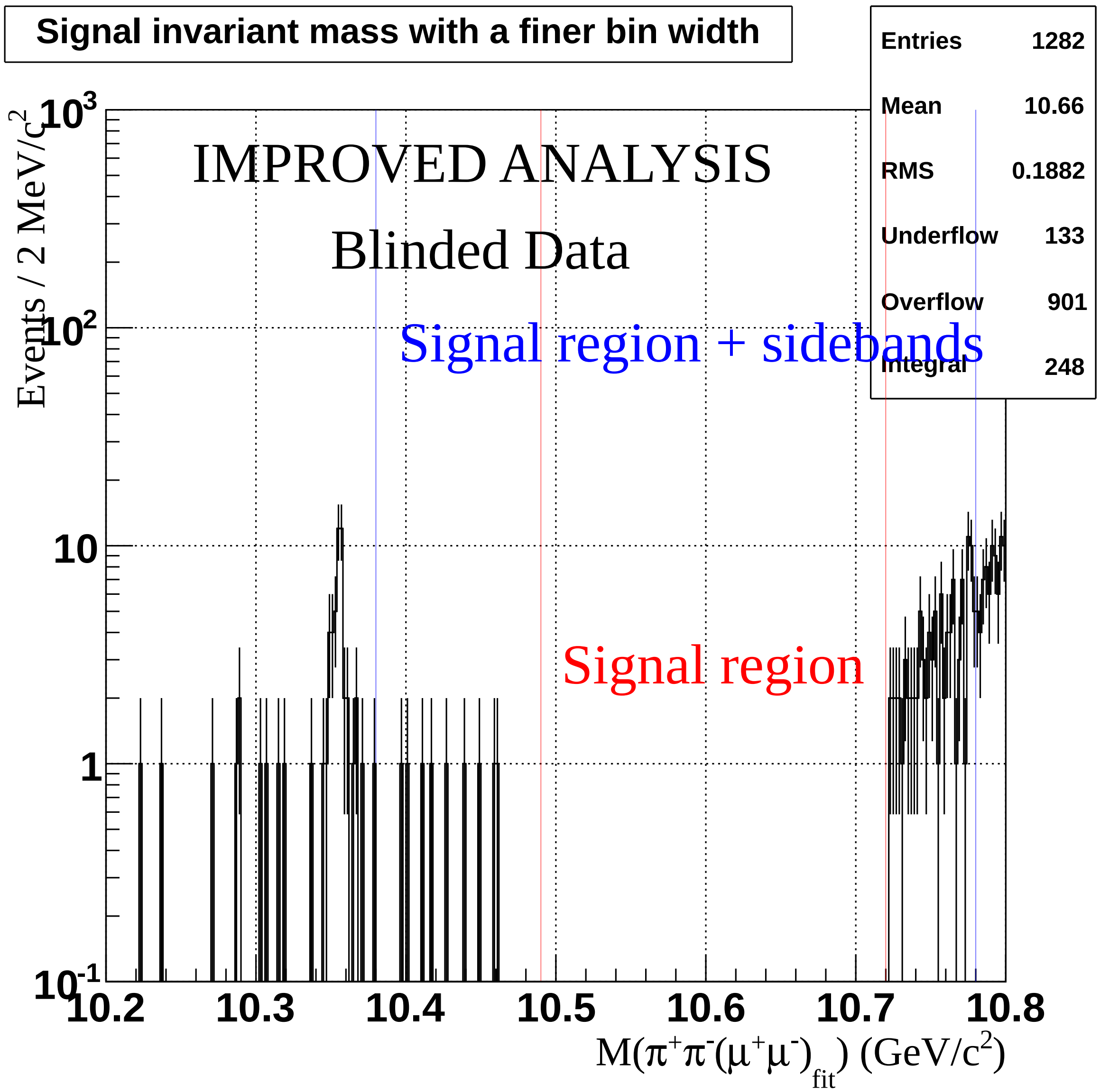



Figure 34: The reconstructed invariant mass $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ for blinded data with relaxed selection criteria plotted using the logarithmic scale for a narrower region of the invariant mass using a finer bin width than used for plots shown in Figures 32 and 33. Only best candidates are shown.

and the straight line, respectively. Note that we can not replace these two parameters by a single fraction parameter in an **extended** ML fit. When we fit real data, we plan to let the relative contributions from the two PDFs to be independently varying parameters in the fit, same way it is the case in fits described here. This degree of freedom could be used to approximate (a small contribution, as we conclude from studying the sidebands) from non-peaking background possibly present in data in the signal region.

In the next step we exclude events in the signal region from the fit and repeat the described exercise for the same ISR MC sample. The results are shown in Fig. 41. One can easily notice that our sidebands are not sufficiently wide to use these to obtain the shape of

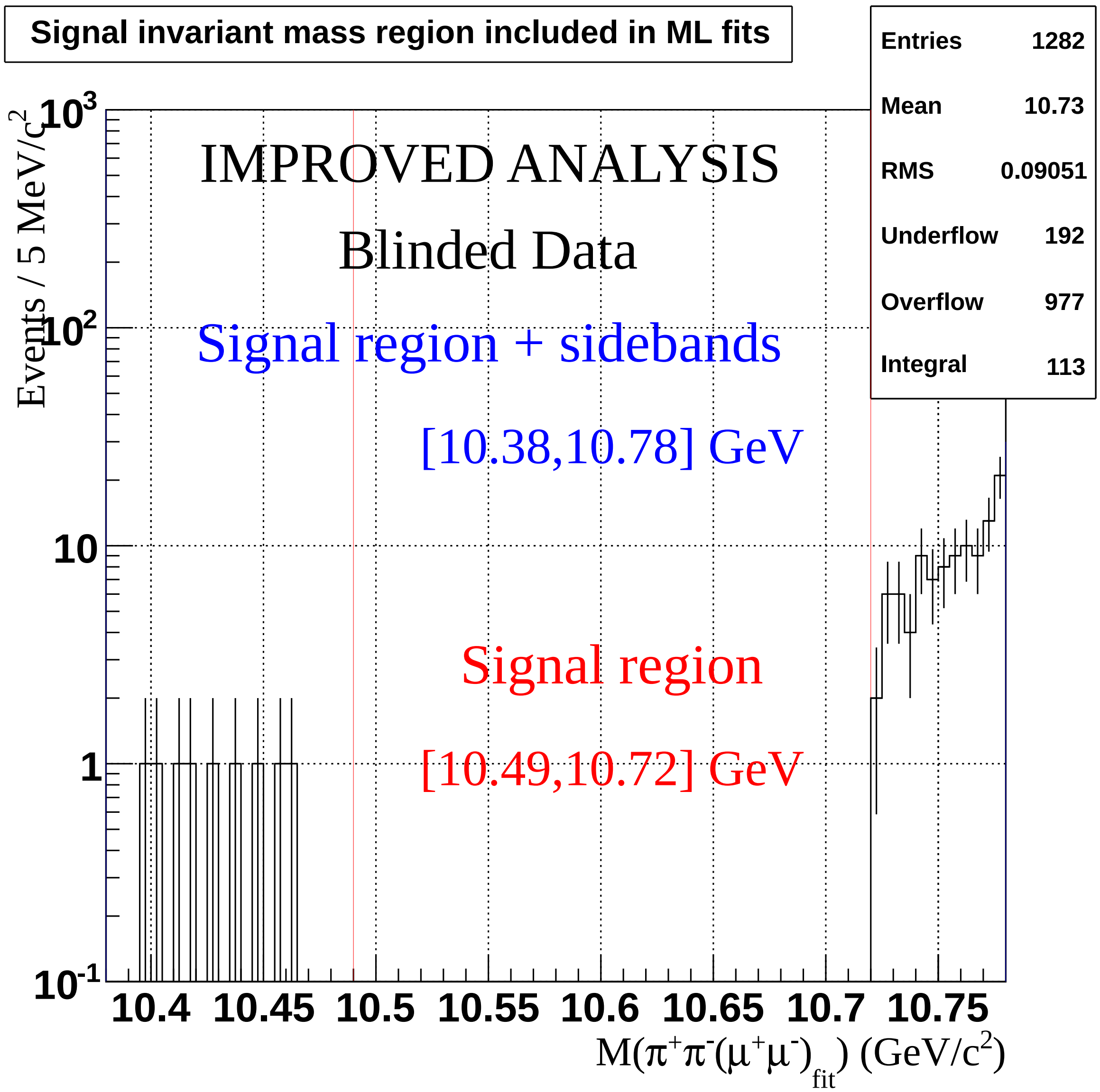



Figure 35: The reconstructed invariant mass $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ for blinded data with relaxed selection criteria plotted using the logarithmic scale for the region of the invariant mass included in the fits. Same events are shown as in Figures 32–34 but using a different bin width. Only best candidates are shown.

background in the **entire** signal region. This is the reason why, when fitting the data, these two PDFs, an exponential and a straight line, will be combined with a signal PDF. In either case, one can see that presence of ISR definitely introduces a large systematic uncertainty in our results for (relatively) larger invariant masses of $W_{bJ}$, in the region where the exponential contribution is rapidly increasing.

Now we try to fit the blinded data (just to see if the fit is going to converge at all). The results of the fit are shown in Fig. 42. Again, we observe that it would be unrealistic to expect our sidebands to predict the background in the entire, 400 MeV/c$^2$-wide signal region. Note that while our ISR MC does not predict the absolute amount of ISR we expect in data, it is

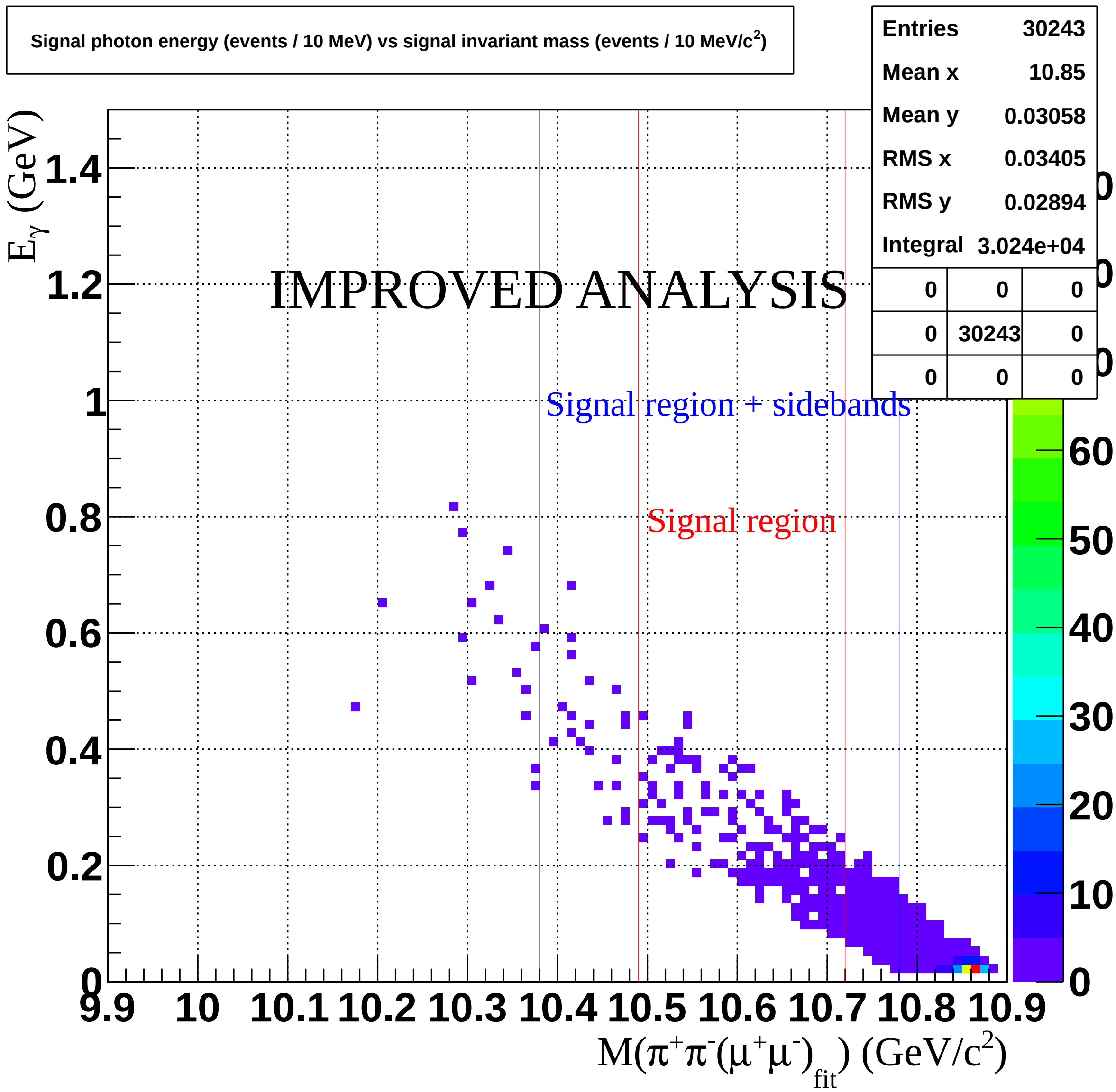



Figure 36: The reconstructed signal photon energy versus the invariant mass $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ for ISR MC sample described in section 6 with relaxed selection criteria. Only best candidates are shown.

interesting to compare the shape of the exponential component of the background obtained from the fits shown, respectively, in Fig. 41 and 42. The exponential part of the fit is driven almost exclusively by the right-side (*i.e.* by the heavier $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ mass) sideband of the $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ distribution. From our fits we obtain the values for the exponential parameter $\alpha$ for ISR MC $33.3 \pm 3.0$ and for blinded data $28.7 \pm 6.4$. It is good to see that these two values are consistent with each other.

Now we try to fit our precious ISR MC sample using three PDFs: the exponential and the straight line approximating backround and the signal line shape shown in Fig. 21a (convolution of the signal Breit-Wigner with two Gaussians prepared using FFT plug-in in ROOT). Note that no signal MC events have been added to the pure ISR MC sample yet.

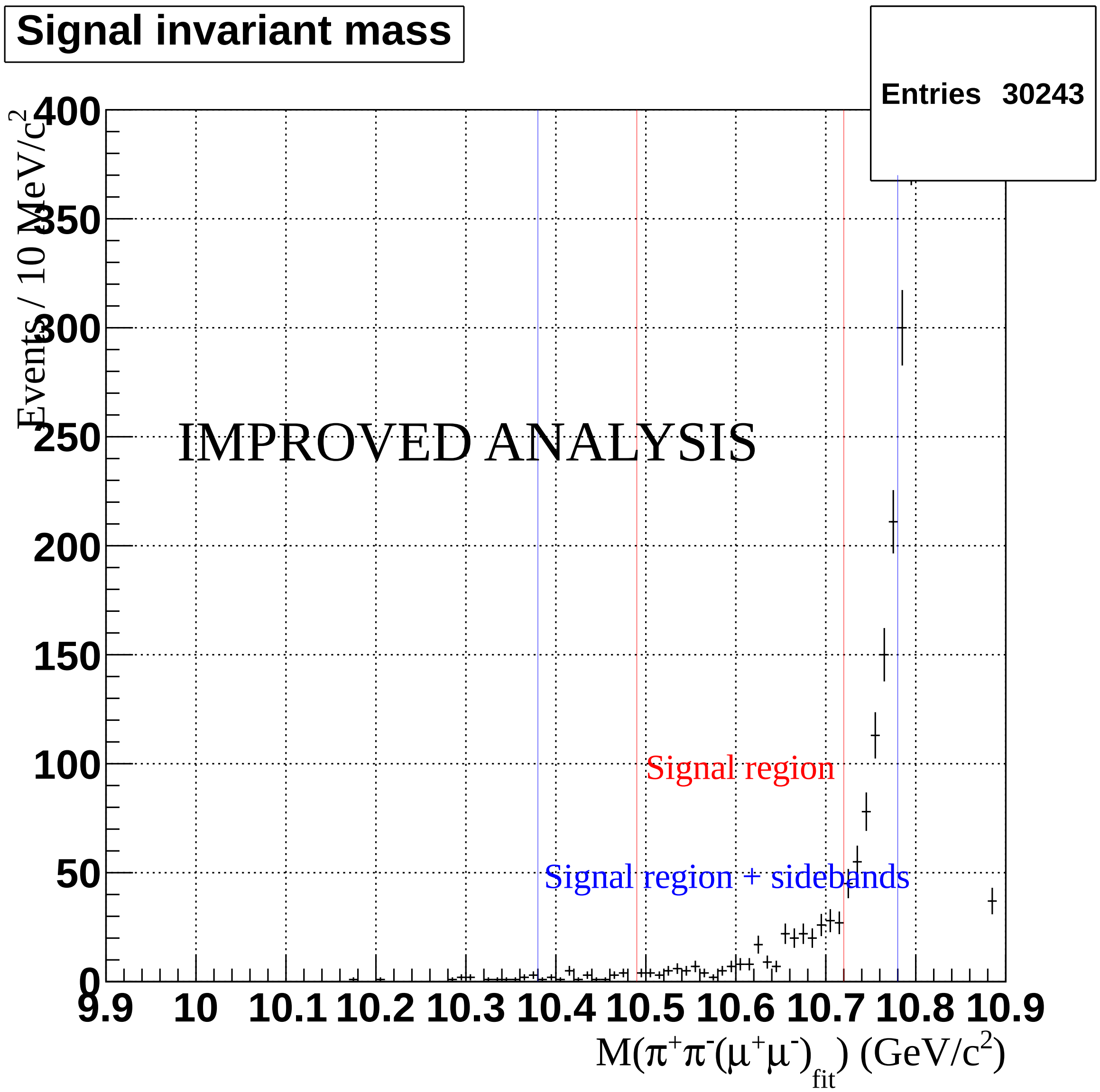



Figure 37: The reconstructed invariant mass $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ for ISR MC sample described in section 6 with relaxed selection criteria. Only best candidates are shown.

The number of signal events ($NS$) is an additional parameter in the fit, but the shape of the signal and its location (*i.e.* the invariant mass of $W_{bJ}$ set at 10.6 GeV/c$^2$) are fixed in this fit. As you can see in Fig. 43 the fit finds no statistically significant signal. Note that the solid green curve superimposed on the results of the fit shows how 50 signal events would look on average according to signal PDF description. The result of the fit for $NS$ is a negative fluctuation.

Inspired by our success, we now ask the fitter to search for the signal (where there is none) in ISR MC sample. To do so we let the nominal mass of the $W_{bJ}$ float in the range between 10.5 and 10.7 GeV/c$^2$. Note that in our future fits to data we plan to **scan** through this interval of $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$, however, for the fit performed here we simply want to see the significance of the worst-case-scenario when the fit “discover” a signal where there

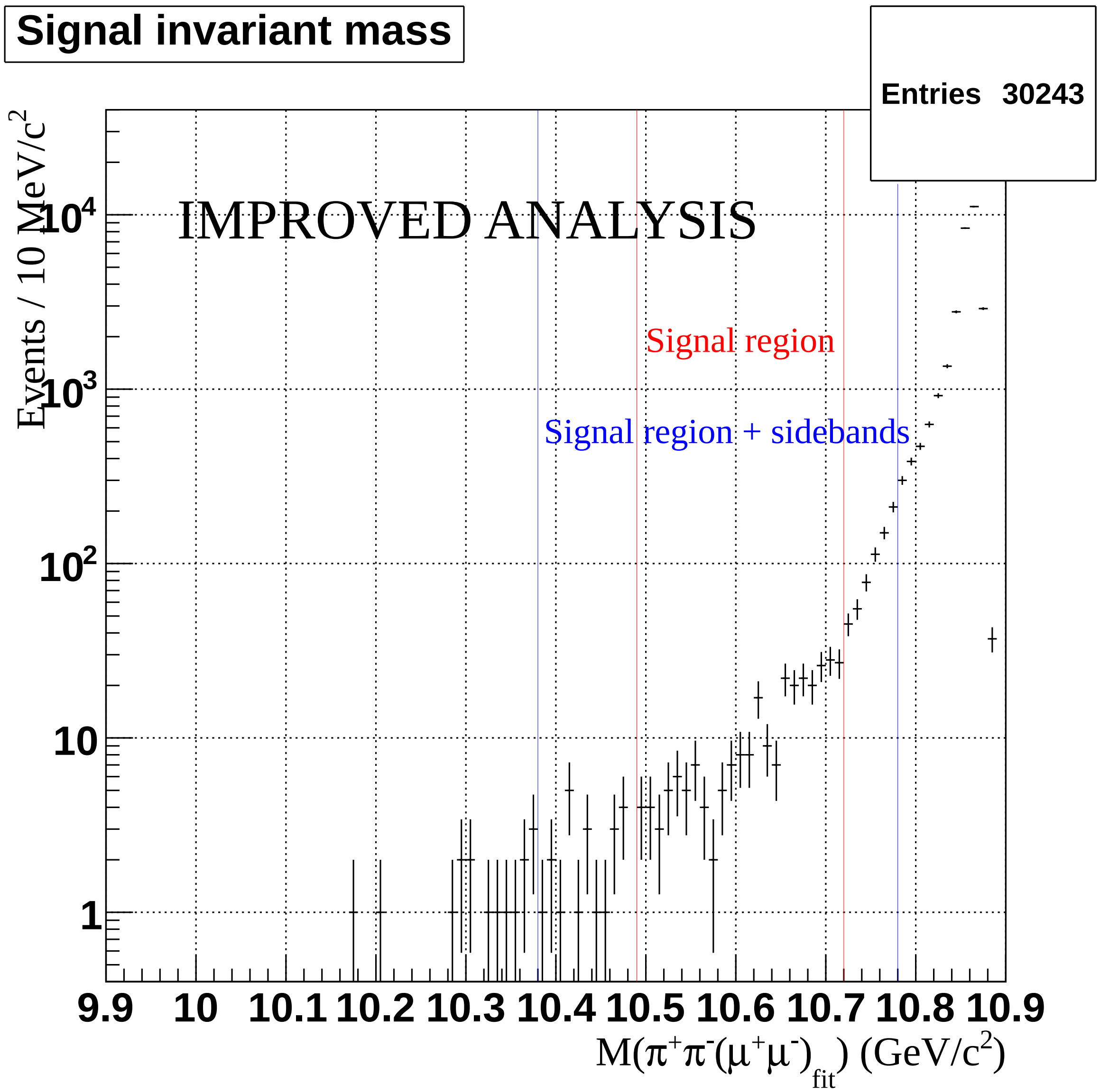



Figure 38: The reconstructed invariant mass $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ for ISR MC sample described in section 6 with relaxed selection criteria shown using the logarithmic scale. Only best candidates are shown.

is none. The results of this fit are shown in Fig. 44. Indeed, an obvious enhancement in the distribution of $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ is "picked-up" by the fitter as the most likely "signal", however, as you can observe from the results of the fit, statistical significance of this "discovery" is consistent with a fluctuation. Such results are also likely to be obtained in the data, and, in case of low significance and no discovery, this would blow up the upper limit estimate.

Finally, being brave young pioneers, we decide to tackle a simulated data sample where 50 events (with $W_{bJ}$ mass of 10.620 GeV/c$^2$) are randomly selected from one of our simulated signal MC sample and are added to the same ISR MC sample we are using for all our fits described in this section. We let the fitter search for this signal and report the results in

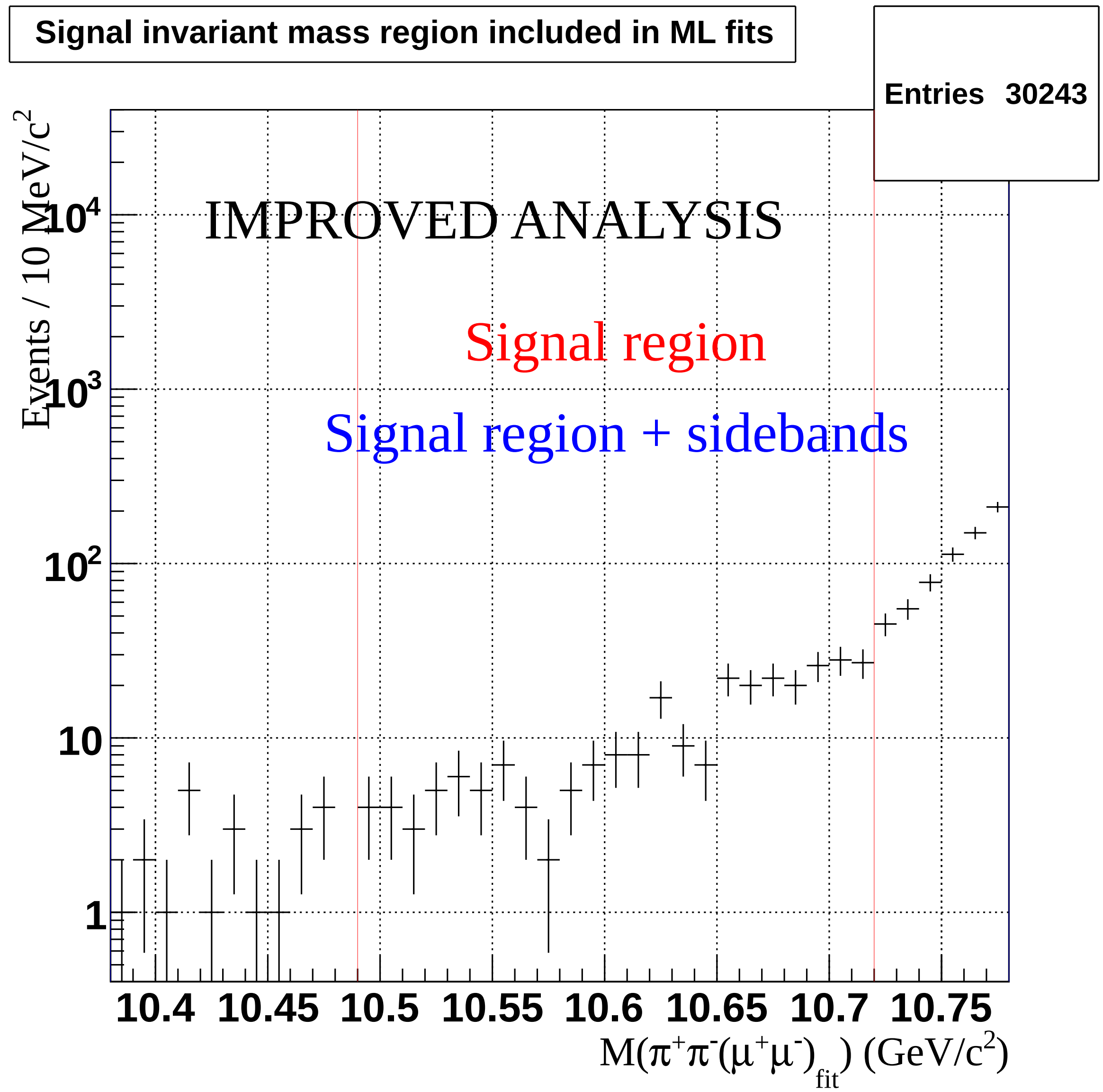



Figure 39: The reconstructed invariant mass $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ for ISR MC sample described in section 6 with relaxed selection criteria shown using the logarithmic scale for the mass region used in the fits. Only best candidates are shown.

Fig. 45. The fitter finds the signal of the right magnitude. We conclude that our fitting procedure is working.

Now knowing that our fitting procedure works, we perform a scan on our reweighted ISR MC sample for the decay $\Upsilon(5S) \to \Upsilon(1S)\pi^+\pi^-$, holding the nominal mass of $W_{bJ}$ fixed for values between 10.5 and 10.7 GeV/c$^2$ in steps of a 5 MeV/c$^2$. The results of the scan are shown in Fig. 46.

We perform another (identical) scan on our reweighted ISR MC sample, but this time, we include 50 toy signal MC events generated from our signal PDF (with the signal mass fixed at values between 10.5 and 10.7 GeV/c$^2$ in steps of a 5 MeV/c$^2$. The results of the scan are shown in Fig. 47.

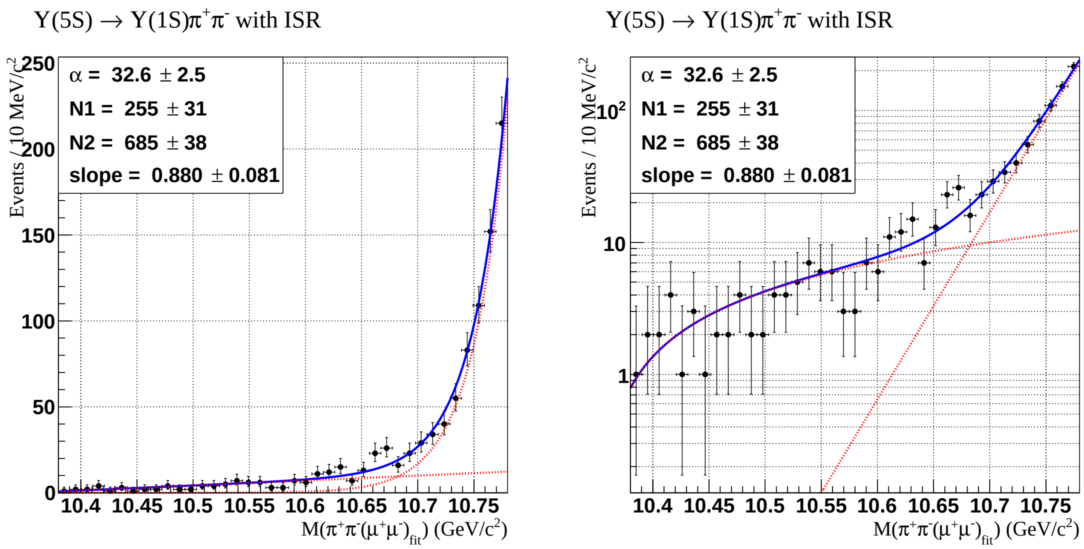



Figure 40: The results of the ML fit for ISR MC sample.

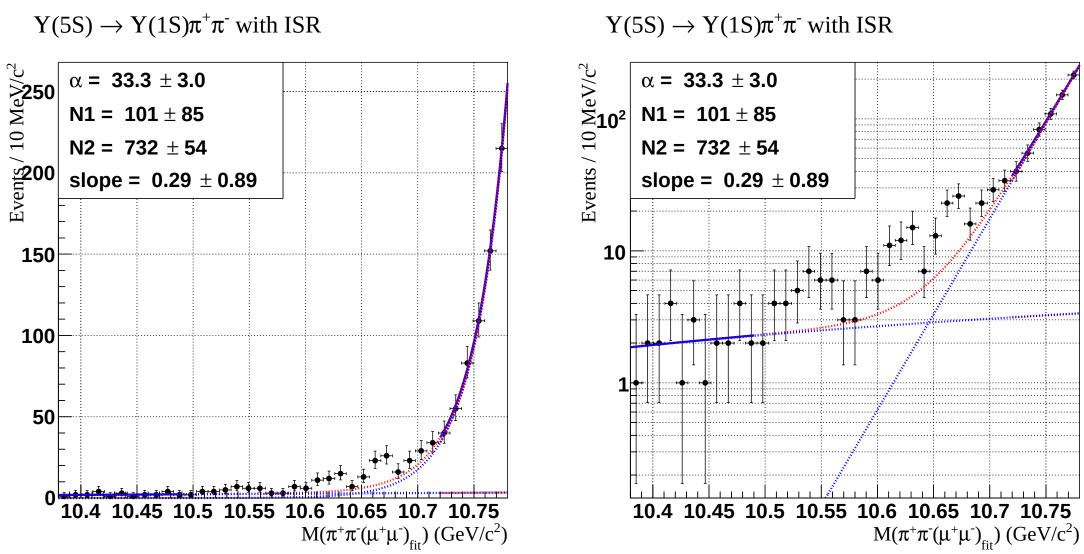



Figure 41: The results of the ML fit for ISR MC sample excluding the signal region.

Note that blind data contain 9 events in the lower-mass sideband and 104 events in the higher-mass sideband, while ISR MC has 22 events in the lower-mass sideband and 652 events in the higher-mass sideband. On basis of this comparison we conclude that our ISR MC sample is larger than the statistics in data. Therefore, future fits to data will likely yield results which are more competitive than the estimates shown in this section. To show an example of a possible improvement, we randomly reduce our ISR sample in the fitting region to half the size in the original ISR sample and repeat the scans to ISR-only sample and ISR sample mixed with 50 toy signal MC events. The results of these scans are shown in Figures 48 and 49.

To investigate the effect of possible presence of the signal $W_{bJ}$ in data, we perform a scan to ISR MC mixed with 50 signal MC events with nominal mass of $W_{bJ}$ at 10.62 GeV/$c^2$.

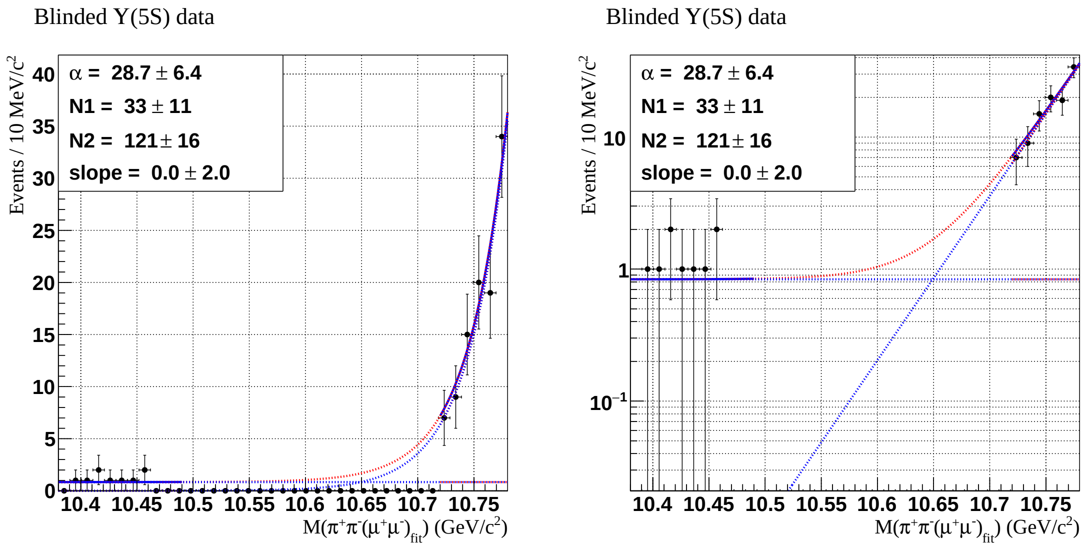



Figure 42: The results of the ML fit for sidebands of the blinded data sample.

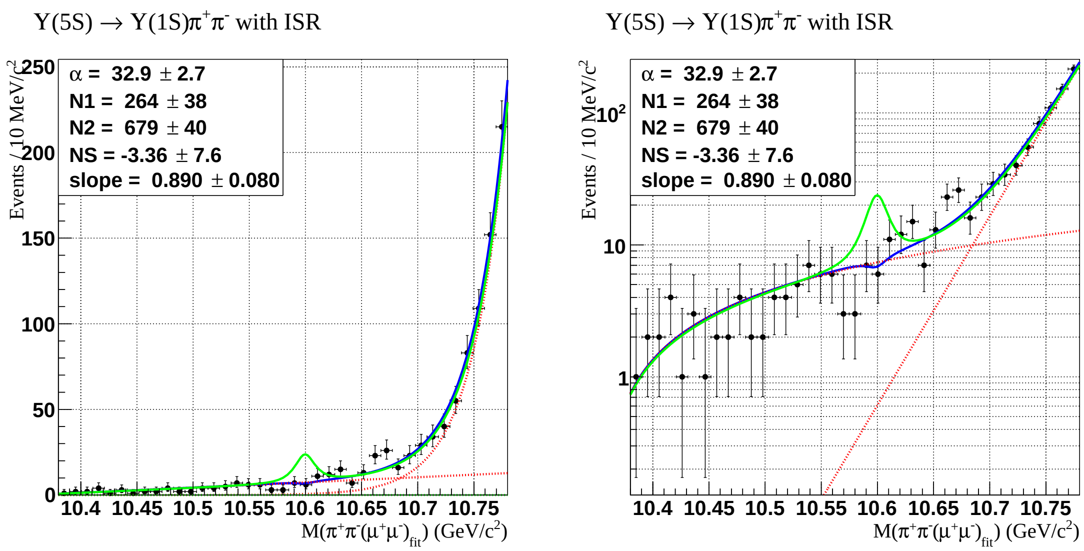



Figure 43: The results of the ML fit for ISR MC sample with background model and signal PDF shape.

The results of the scan performed on this sample are shown in Fig. 50.

All the fits discussed so far were performed in the range of $M(\pi^+\pi^-(\mu^+\mu^-)_{\rm fit})$ between 10.39 and 10.78 GeV/c$^2$. To understand how much this choice affects our results, we repeat the scan shown in Fig. 50 in a narrower range of the invariant masses, namely, between 10.42 and 10.74 GeV/c$^2$. The results of this scan are shown in Fig. 50.

To summarize, in fits shown in Figures 46 and 48 we fit ISR MC only. Fits shown in Figures 47 and 49 include 50 toy signal MC events, where for each point in the scan, these signal events were generated with the mass at that particular point. Fits shown in Figures 48, 49, 50 and 51 have less ISR MC than in fits with the results shown in Figures 46 and 47. Finally, scan fits shown in Figures 50 and 51 have 50 fully simulated (*i.e.* with the

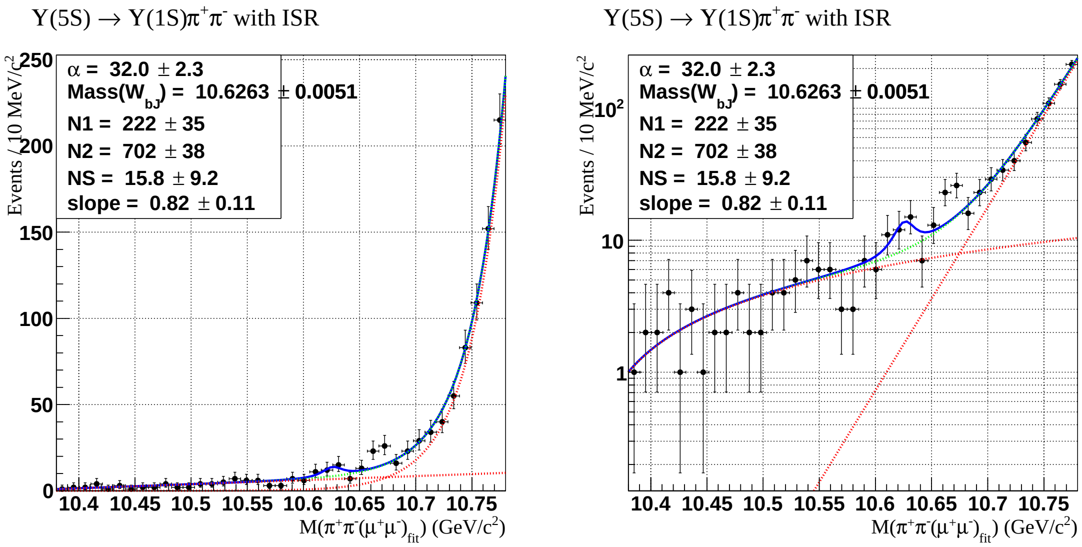



Figure 44: The results of the ML fit for ISR MC sample with background model and signal PDF shape (mass is a free parameter).

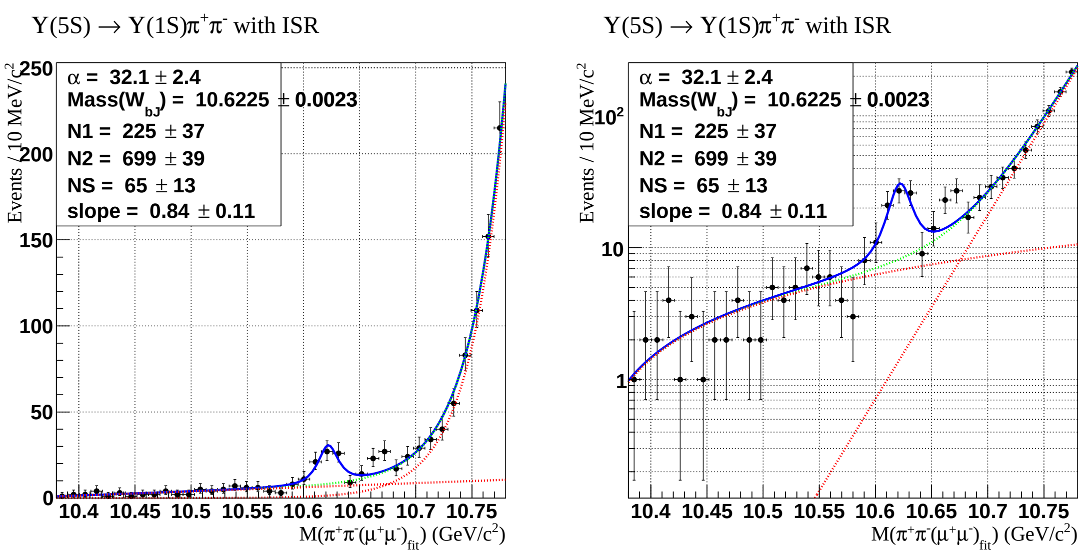



Figure 45: The results of the ML fit for ISR + signal (50 events) MC sample with background model and signal PDF shape (mass is a free parameter).

detector simulation) signal MC events at a particular $W_{bJ}$ mass (of 10.62 GeV/c$^2$).

# 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.

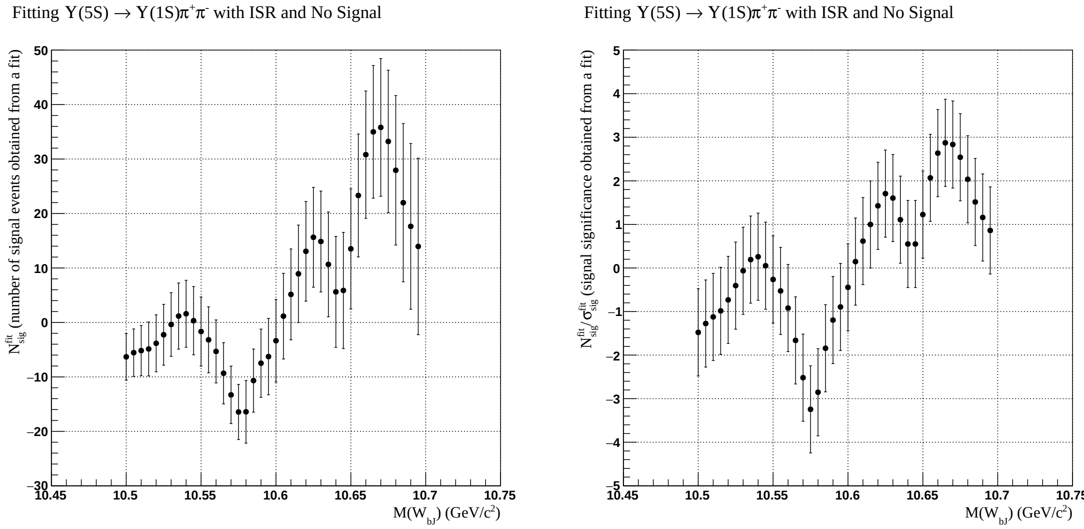



Figure 46: The results of the scan for ISR MC sample with background model and signal PDF shape.

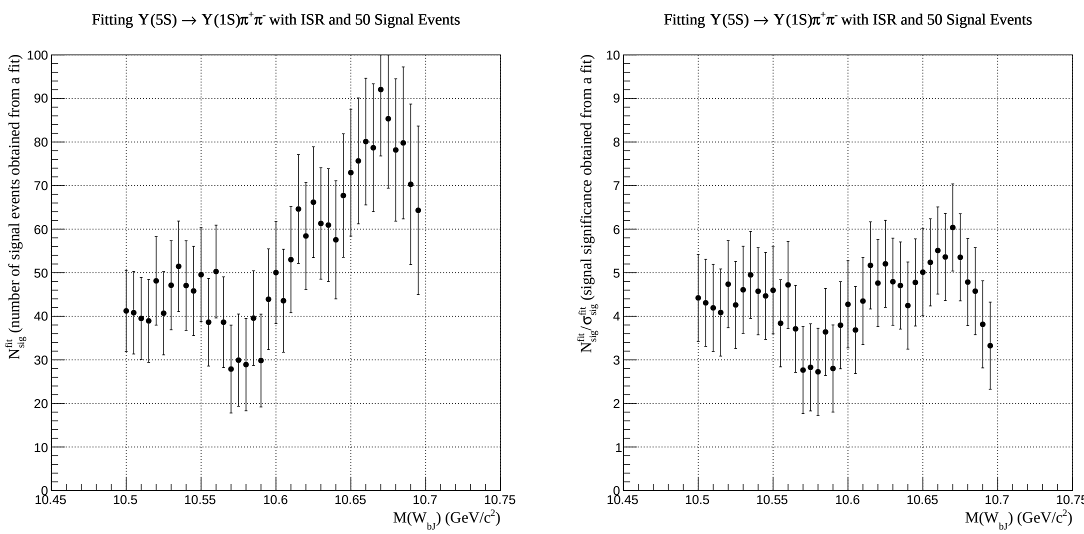



Figure 47: The results of the scan for ISR MC sample + 50 toy MC signal events (with $W_{bJ}$ mass in toy MC varied in the scan) with background model and signal PDF shape.

[3] P. Krokovny et al. First observation of the $Z_b^0(10610)$ in a Dalitz analysis of $\Upsilon(10860) \to \Upsilon(\mathrm{nS})\pi^0\pi^0$. *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.

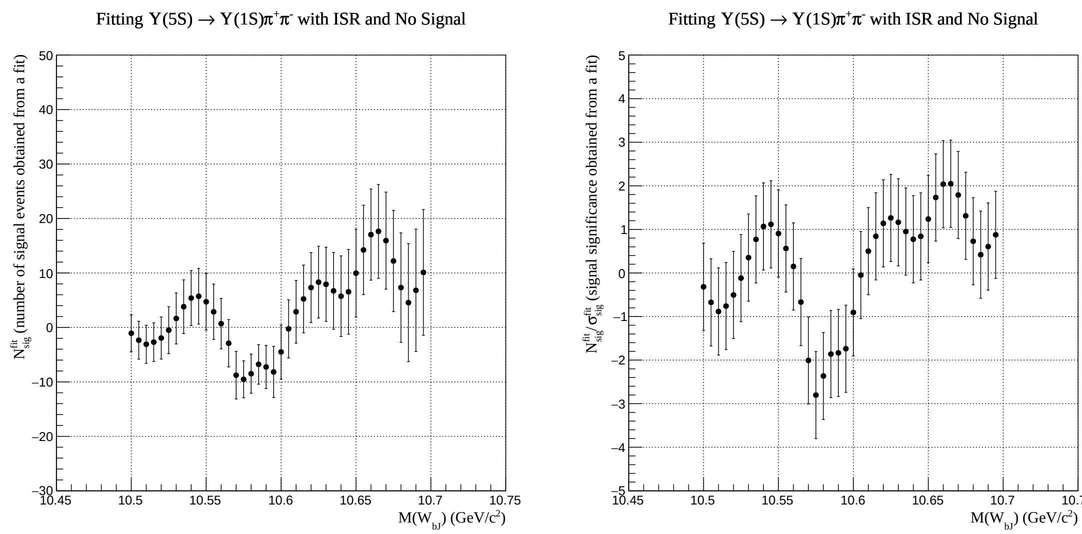



Figure 48: The results of the scan for ISR MC sample with background model and signal PDF shape.

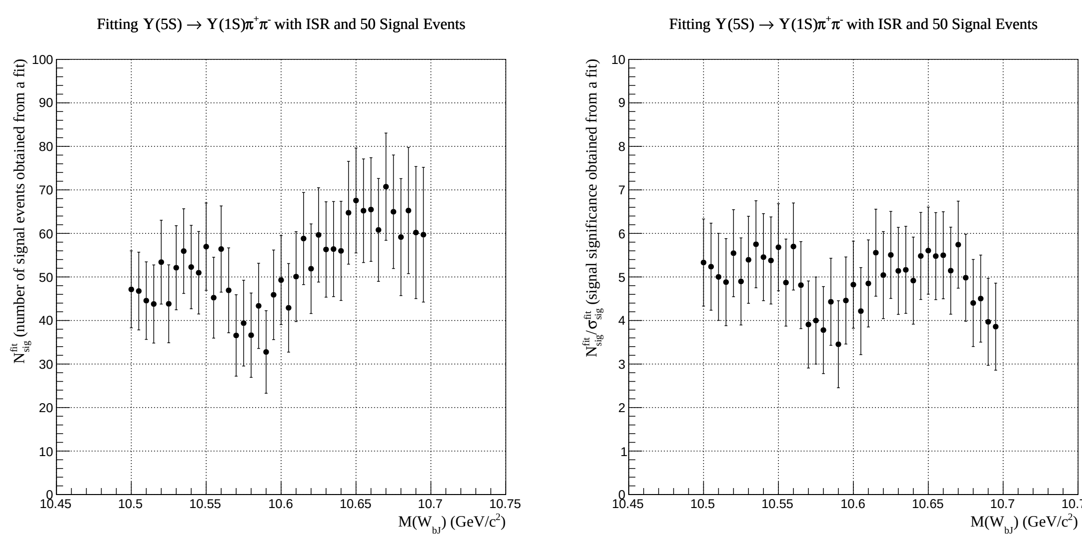



Figure 49: The results of the scan for ISR MC sample + 50 toy MC signal events (with $W_{bJ}$ mass in toy MC varied in the scan) with background model and signal PDF shape.

[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.

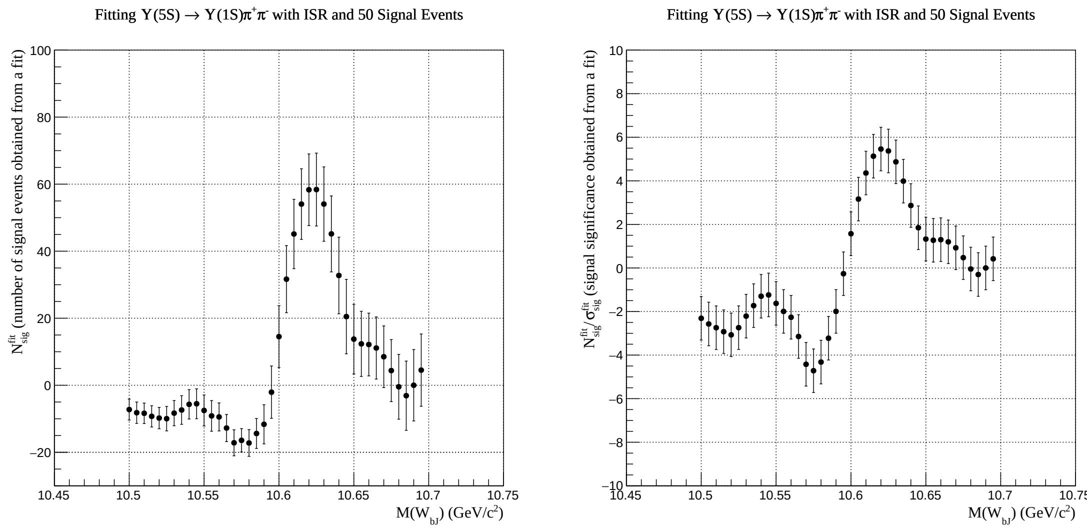



Figure 50: The results of the scan for ISR MC sample + 50 signal MC signal events ($M(W_{bJ}) = 10.62$ GeV/c$^2$) with background model and signal PDF shape.

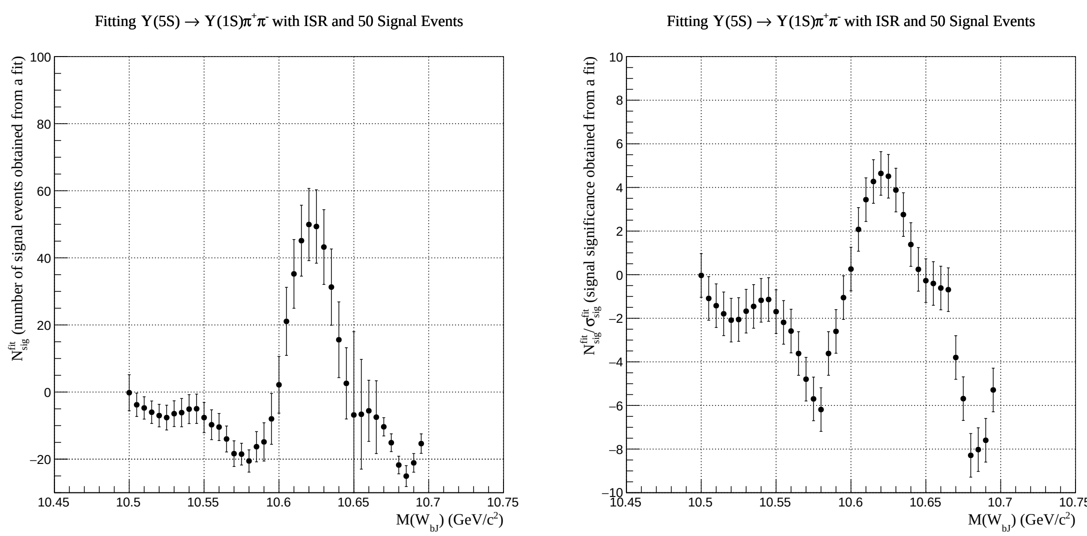



Figure 51: The results of the scan for ISR MC sample + 50 signal MC signal events ($M(W_{bJ}) = 10.62$ GeV/c$^2$) with background model and signal PDF shape in a narrower range of $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$.

[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)].