

Search for decay $\Upsilon(5S) \rightarrow \gamma W_{bJ}$

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 $\Upsilon(5S) \rightarrow \gamma W_{bJ}$ 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 $\Upsilon(5S) \rightarrow \gamma W_{bJ}$, 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.

16 Contents

17	1 Introduction	1
18	1.1 Motivation	1
19	1.2 New Spectroscopy	1
20	1.3 Radiative Decays $\Upsilon(5S) \rightarrow \gamma W_{bJ}$	3
21	1.4 Expected Signal in Data	4
22	2 Monte Carlo and Data Samples	6
23	3 Selection Criteria	6
24	3.1 Selection of Photon Candidates	7
25	3.2 Selection of Pion and Muon Candidates	8
26	3.3 Selection of $\Upsilon(5S)$ Candidates	8
27	3.4 Best Candidate Selection	8
28	4 Signal Monte Carlo Studies	8
29	4.1 Signal Monte Carlo Distributions	8
30	4.2 Description of the Signal Region	12
31	4.3 Trigger Simulation	15
32	5 Background Studies	18
33	5.1 Generic Monte Carlo and Blinded Data	18
34	6 Background from $\Upsilon(5S) \rightarrow \Upsilon(1S)\pi^+\pi^-$ with Initial State Radiation (ISR)	20
35	6.1 $\Upsilon(5S) \rightarrow \Upsilon(1S)\pi^+\pi^-$ ISR Monte Carlo Sample	20
36	6.2 Background Shape of $\Upsilon(5S) \rightarrow \Upsilon(1S)\pi^+\pi^-$ with ISR	22
37	7 Contribution from $\Upsilon(5S) \rightarrow Z_b^{(\prime)\pm}\pi^\mp$	26
38	8 Fitting	26
39	8.1 Signal and Background PDFs	26
40	8.2 Confidence Belts	28
41	8.3 Linearity Study	29
42	8.4 Sensitivity Estimation	31
43	9 Search Strategy Summary	33
44	10 Appendix	34
45	10.1 Final State Radiation	34
46	10.2 Changes in the Analysis between Note v1.5 and v2.0	35
47	10.2.1 Signal Photon Energy Conundrum	36
48	10.3 Fitting Strategy	41

49 List of Figures

50	1	Pure $b\bar{b}$ bottomonium mass spectrum for a relativized quark model	2
51	2	Spectrum of bottomonium and bottomonium-like mesons.	3
52	3	Expected family of molecular isotriplet resonances.	5
53	4	Various MC distributions which informed our selection criteria.	7
54	5	Signal MC distributions	10
55	6	$M(\pi^+\pi^+\mu^+\mu^-)$ and $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ resolutions for signal events	11
56	7	ΔE resolution and quantities contributing to ΔE resolution.	12
57	8	Definitions of important regions	14
58	9	Offline trigger selection for reconstructed signal MC events.	16
59	10	Offline trigger efficiency for reconstructed signal MC events.	17
60	11	Identifying background present in data but not in generic MC.	18
61	12	Motivation for ISR studies.	20
62	13	Reweighted ISR energy spectrum	21
63	14	Background due to ISR.	22
64	15	Effect of decay models on $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ distribution.	23
65	16	Effect of decay models on muon angular distribution.	24
66	17	Effect of ΔE cut on ISR background shape.	25
67	18	Definitions of subdivisions of the grand sideband region	25
68	19	Distribution of $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ vs $M_{\text{rec}}(\gamma)$ for $\Upsilon(5S) \rightarrow Z_b^{(\prime)\pm}\pi^\mp$ MC	27
69	20	Distribution of $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ for $\Upsilon(5S) \rightarrow Z_b^{(\prime)\pm}\pi^\mp$	27
70	21	Fitting background MC and data	29
71	22	90% confidence belts for frequentist method.	30
72	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 $f(x) = p_0 + p_1x$. The resulting fit parameters are shown in the box on the top right.	31
73			
74			
75	24	$N_{\text{sig}}^{\text{fit}}$ Distributions for ensemble tests with different $N_{\text{sig}}^{\text{gen}}$	32
76	25	Final state radiation from charged tracks	34
77	26	Signal invariant mass for a special MC sample with the original selection criteria.	38
78	27	Photon energy vs signal invariant mass for a special MC sample with the original selection criteria.	39
79			
80	28	Photon energy vs signal invariant mass for a special MC sample with relaxed selection criteria.	40
81			
82	29	Signal invariant mass for a special MC sample with relaxed selection criteria.	41
83	30	Signal invariant mass for a special MC sample with relaxed selection criteria (log scale).	42
84			
85	31	Photon energy vs signal invariant mass for blinded data with relaxed selection criteria.	43
86			
87	32	Signal invariant mass for blinded data with relaxed selection criteria.	44
88	33	Signal invariant mass for blinded data with relaxed selection criteria (log scale).	45
89	34	Signal invariant mass for blinded data with relaxed selection criteria (log scale, finer bins).	46
90			
91	35	Signal invariant mass for blinded data with relaxed selection criteria (log scale, fitting region).	47
92			

93	36	Photon energy vs signal invariant mass for ISR MC sample described in section 6 with relaxed selection criteria.	48
94			
95	37	Signal invariant mass for ISR MC sample described in section 6 with relaxed selection criteria.	49
96			
97	38	Signal invariant mass for ISR MC sample described in section 6 with relaxed selection criteria (log scale).	50
98			
99	39	Signal invariant mass for ISR MC sample described in section 6 with relaxed selection criteria (log scale, fitting region only).	51
100			
101	40	The results of the ML fit for ISR MC sample.	52
102	41	The results of the ML fit for ISR MC sample excluding the signal region.	52
103	42	The results of the ML fit for sidebands of the blinded data sample.	53
104	43	The results of the ML fit for ISR MC sample with background model and signal PDF shape.	53
105			
106	44	The results of the ML fit for ISR MC sample with background model and signal PDF shape (mass is a free parameter).	54
107			
108	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).	54
109			
110	46	The results of the scan for ISR MC sample with background model and signal PDF shape.	55
111			
112	47	The results of the scan for ISR MC sample + 50 toy MC signal events with background model and signal PDF shape.	55
113			
114	48	The results of the scan for ISR MC sample with background model and signal PDF shape.	56
115			
116	49	The results of the scan for ISR MC sample + 50 toy MC signal events with background model and signal PDF shape.	56
117			
118	50	The results of the scan for ISR MC sample + 50 signal MC signal events ($M(W_{bJ}) = 10.62 \text{ GeV}/c^2$) with background model and signal PDF shape.	57
119			

120 **List of Tables**

121	1	Possible molecular isotriplet states in decays of $\Upsilon(5S)$ and $\Upsilon(6S)$	4
122	2	EvtGen decay models used in Monte Carlo simulation of signal processes. . .	6
123	3	Selection criteria for $\Upsilon(5S) \rightarrow \gamma W_{bJ}$	9
124	4	Quantities contributing to widths of measured quantities	10
125	5	Definitions of the signal region and other important regions.	13
126	6	Backgrounds labeled in Fig. 11.	19
127	7	Decay models used in background ISR MC	22
128	8	Definitions of subdivisions of the grand sideband region	24
129	9	Comparing the number of events in ISR MC and blinded data in the subdivided grand sideband region	26
130			
131	10	Values of fixed quantities in the signal PDF model	28
132	11	Values of quantities used in upper limit calculation	33
133	12	The new (wider) blinded region in data. The important change is shown in	
134		red color. However, it is redundant and adds nothing new as compared to the	
135		second line in this table.	36

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

Hadronic transitions (such as, *e.g.* $\Upsilon(3S) \rightarrow \pi^+\pi^-\Upsilon(1S)$) 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

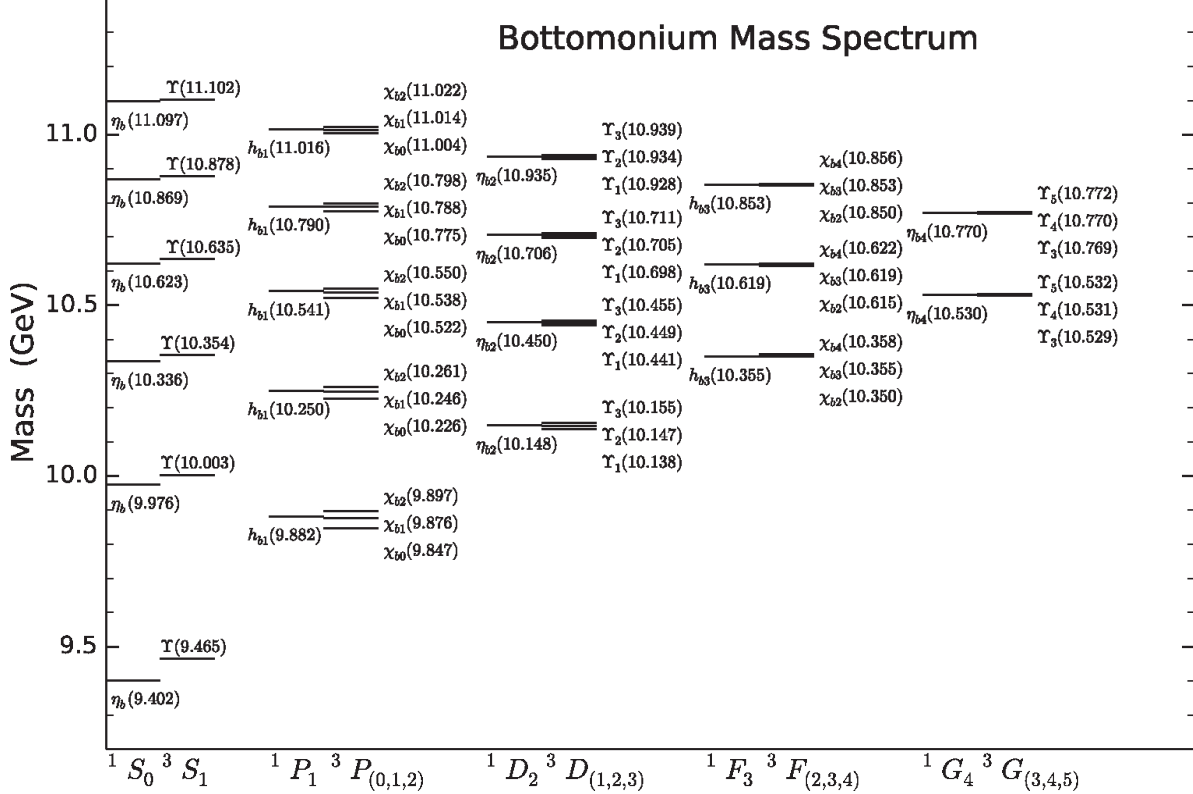


Figure 1: Pure (*i.e.* $b\bar{b}$) bottomonium mass spectrum [8] calculated using a relativized quark model [7].

176 transitions involving the spin of the heavy b quark are strongly suppressed, the former states,
 177 including the $\Upsilon(5S)$, require a different explanation [10].

178 The favored explanation for the properties of $\Upsilon(5S)$, including its decays to Z_b , is based
 179 on the molecular picture, where these vector bottomonium-like resonances are assumed to
 180 contain an admixture of pairs of colored heavy mesons. This hypothesis has been successfully
 181 employed [11] to explain the decays to and the existence of the six Z_b states. However,
 182 the details of the interaction responsible for these processes are not yet fully understood.
 183 Alternative explanations include a model with a diquark-antidiquark pair, where a pair of
 184 quarks and a pair of antiquarks are each bound with a stronger force than the force holding
 185 diquark and antidiquark together. While the search described in this document is model-
 186 independent, our motivation is somewhat biased in favor of the molecular picture and has
 187 likely impacted our decisions about how to perform the analysis.

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

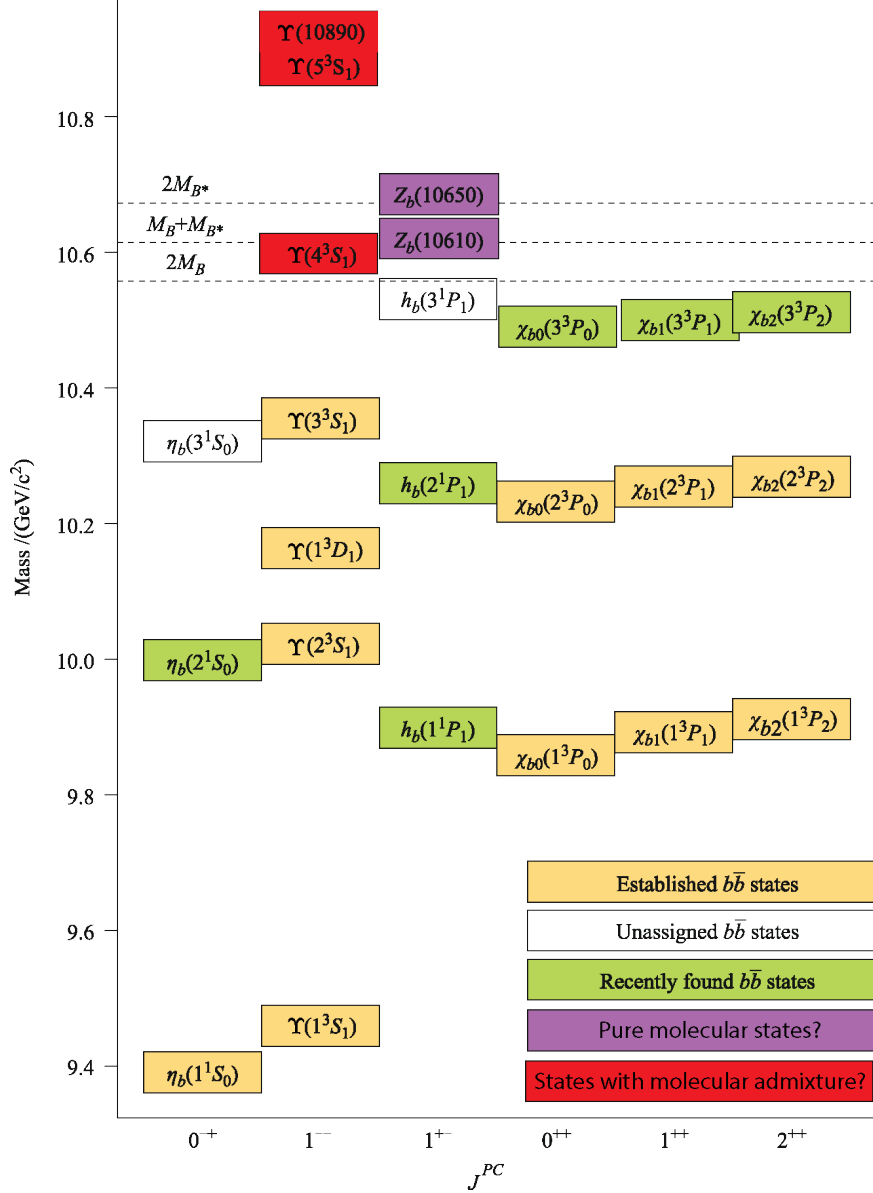


Figure 2: Most relevant (for our study) states in conventional bottomonium and bottomonium-like spectroscopies. We stole this figure from S. Olsen’s excellent review article [12]. Note that we took liberty to modify the original figure to better represent the contents of this Note, namely, we relabeled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

196 predictions for the QCD part.

197 1.3 Radiative Decays $\Upsilon(5S) \rightarrow \gamma W_{bJ}$

198 The Z_b states were discovered in single-pion transitions of $\Upsilon(5S)$ and $\Upsilon(6S)$, followed by
 199 another single-pion transition to the bottomonium states. According to molecular interpre-
 200 tation, $Z_b(10610)$ 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 ⁺)	$Z_b(10610)$	π (10.75)	$B\bar{B}^*$	$\Upsilon(nS)\pi, h_b(nP)\pi, \eta_b(nS)\rho$
1 ⁺ (1 ⁺)	$Z'_b(10650)$	π (10.79)	$B^*\bar{B}^*$	$\Upsilon(nS)\pi, h_b(nP)\pi, \eta_b(nS)\rho$
1 ⁻ (0 ⁺)	W_{b0}	ρ (11.34), γ (10.56)	$B\bar{B}$	$\Upsilon(nS)\rho, \eta_b(nS)\pi, \chi_{b\pi}$
1 ⁻ (0 ⁺)	W'_{b0}	ρ (11.43), γ (10.65)	$B^*\bar{B}^*$	$\Upsilon(nS)\rho, \eta_b(nS)\pi, \chi_{b\pi}$
1 ⁻ (1 ⁺)	W_{b1}	ρ (11.38), γ (10.61)	$B\bar{B}^*$	$\Upsilon(nS)\rho, \chi_{b\pi}$
1 ⁻ (2 ⁺)	W_{b2}	ρ (11.43), γ (10.65)	$B^*\bar{B}^*$	$\Upsilon(nS)\rho, \chi_{b\pi}$

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

201 Z_b are spin-1 isotriplets (both neutral and charged states were discovered in transitions
202 $\Upsilon(nS) \rightarrow \pi Z_b$ ($n = 5, 6$). The hypothetical partners of positive G -parity states Z_b , *i.e.* the
203 W_{bJ} states, would also be isotriplets but of negative G -parity (quantum numbers of the new
204 molecular states are defined by quantum numbers of their partners in two-body decays of the
205 $\Upsilon(5S)$ parent: while Z_b is accompanied by a pion, each W_{bJ} is accompanied by a ρ meson (or
206 a photon)). Therefore the W_{bJ} states are expected to appear in transitions $\Upsilon(nS) \rightarrow \rho W_{bJ}$.
207 Conservation of angular momentum allows J in W_{bJ} to be 0, 1 or 2. Excited states such as
208 W'_{b0} could exist as well. Quantum numbers assigned to Z_b and W_{bJ} states are summarized
209 in Table 1.

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

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

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

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

223 1.4 Expected Signal in Data

224 Belle previously reported [15] that charged Z_b states comprise approximately 2.54% of the
225 1819 $\Upsilon(1S)\pi^+\pi^-$ (followed by $\Upsilon(1S) \rightarrow \mu^+\mu^-$) events observed with the full data sample.

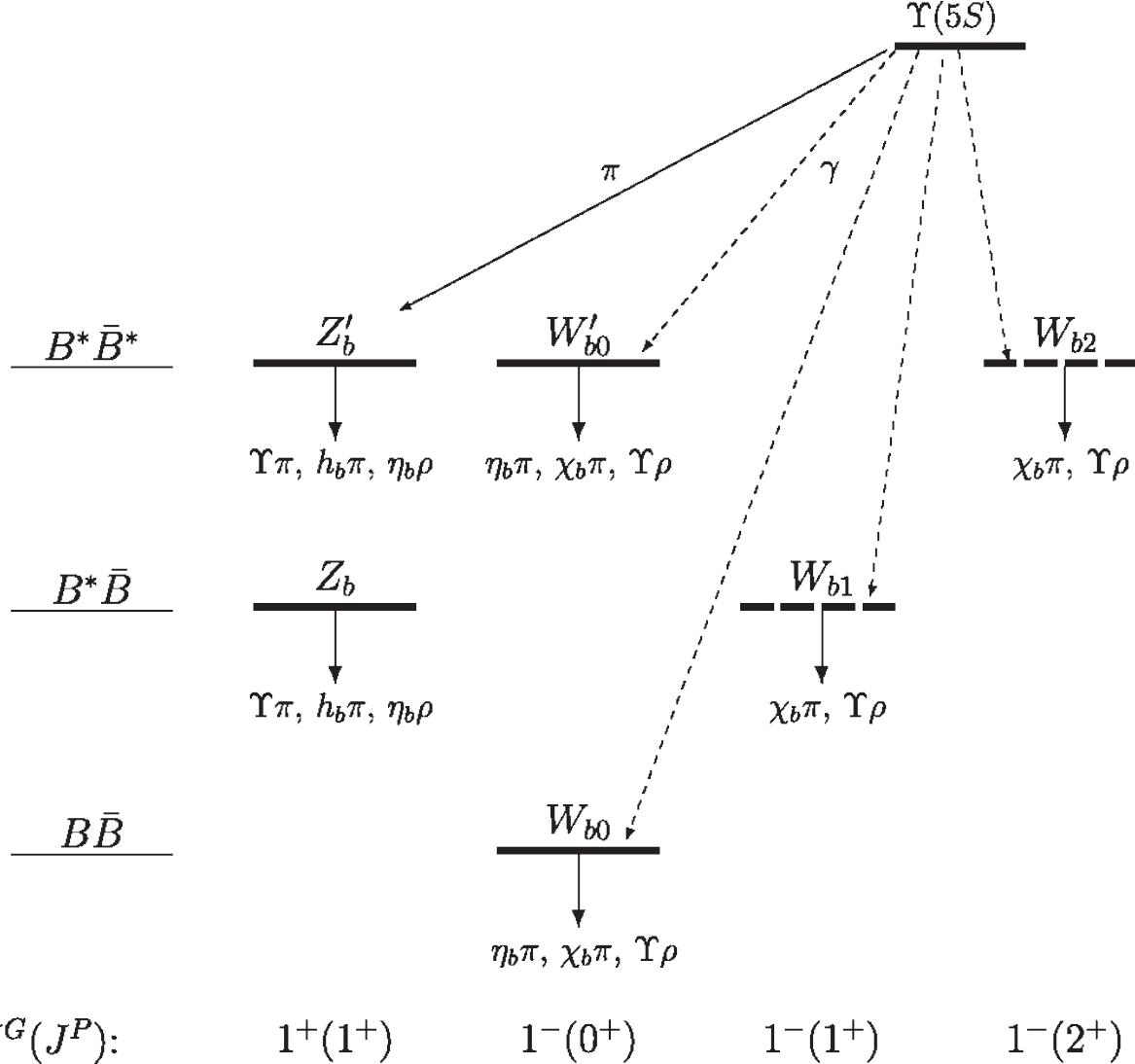


Figure 3: The expected family of isotriplet resonances from Ref. [13] (which the reader is advised to consult for relevant details). For $\Upsilon(6S)$ transitions, the photon is replaced by ρ . This would also allow to access charged W_{bJ} states. Also, please see Table 1.

226 The overall reconstruction efficiency in Z_b analysis was estimated to be around 46%. This
 227 allows us to estimate that, with an ideal, *i.e.* 100% efficient detector, we would expect to
 228 detect, approximately, 100 (charged) Z_b events.

229 While searching for W_{bj} events in radiative decays of $\Upsilon(5S)$, as elaborated in Section 1.3,
 230 we have to pay the price of α_{em} . Jumping a little bit ahead of ourselves, with our overall
 231 detection efficiency of 29%, we therefore expect to observe, on average, 0.2 W_{b0} events. This
 232 number, however, has a (hopefully very) large uncertainty, and, after all, we are (always!)
 233 driven by hope that nature might be kinder to us than we deserve. Also, tangentially, our
 234 LHC colleagues have been searching for signatures of SUSY for some time already, and, no
 235 matter how little has been observed so far, their noble quest will stop not. So why should we
 236 stop ours? On this philosophical note we conclude this discussion and proceed to describe

237 our actual analysis.

238 2 Monte Carlo and Data Samples

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

247 We use six streams of generic MC to study background events. Each stream is equivalent
 248 to a full Belle data sample of 121.4 fb^{-1} of $\Upsilon(5S)$ resonance data. We generate additional MC
 249 samples to study background events originating from $\Upsilon(5S) \rightarrow \Upsilon(1S)\pi^+\pi^- \rightarrow \mu^+\mu^-\pi^+\pi^-$
 250 with initial state radiation (ISR) as well as events originating from $\Upsilon(5S) \rightarrow Z_b^\pm\pi^\mp \rightarrow$
 251 $\Upsilon(1S)\pi^\pm\pi^\mp \rightarrow \mu^+\mu^-\pi^\pm\pi^\mp$. We describe our studies of these processes in Section 6 and
 252 Section 7, respectively.

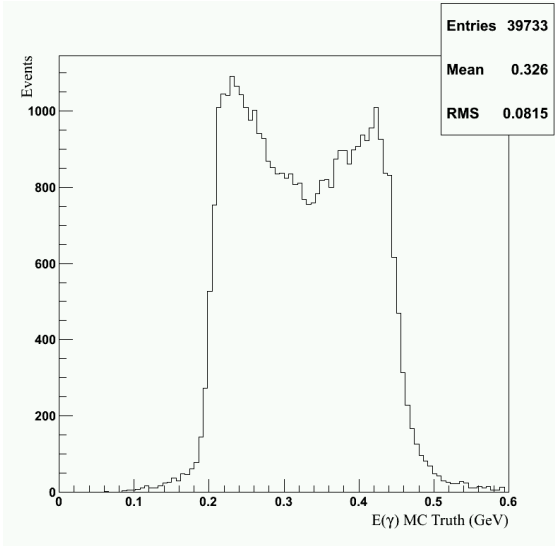
253 In this analysis, we use the full 121.4 fb^{-1} of on-resonance $\Upsilon(5S)$ data collected by the
 254 Belle detector at the KEKB collider from asymmetric energy e^+e^- collisions with $\sqrt{s} = 10.86$
 255 GeV [20].

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

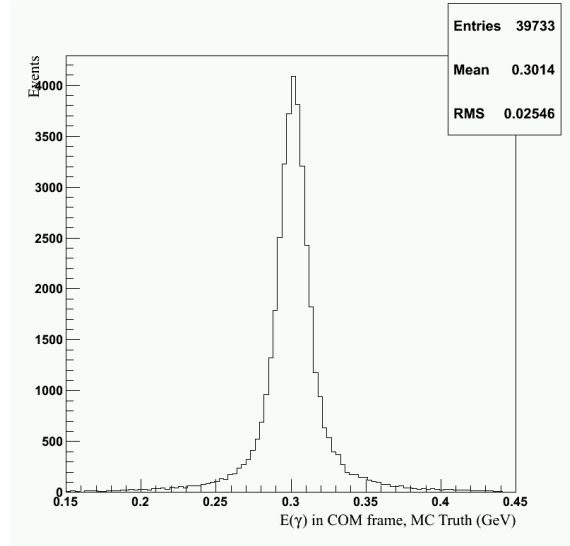
Table 2: EvtGen decay models used in Mote Carlo simulation of signal processes.

256 3 Selection Criteria

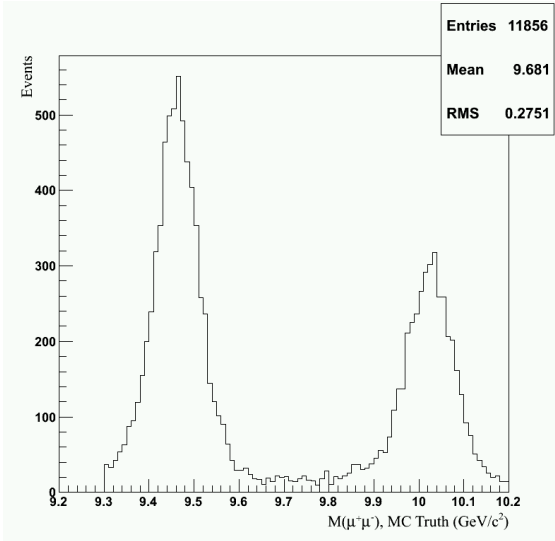
257 We reconstruct the decay mode $\Upsilon(5S) \rightarrow \gamma W_{bJ}$ followed by the decays $W_{bJ} \rightarrow \Upsilon(1S)\rho^0$,
 258 $\Upsilon(1S) \rightarrow \mu^+\mu^-$, $\rho^0 \rightarrow \pi^+\pi^-$. We select a fully-reconstructed final state particle combination
 259 consisting of $\pi^+\pi^-\mu^+\mu^-\gamma$. **The selection criteria that follow, though not systematically**
 260 **optimized, are based on MC truth distributions and typical choices made in previous Belle**
 261 **analyses.**



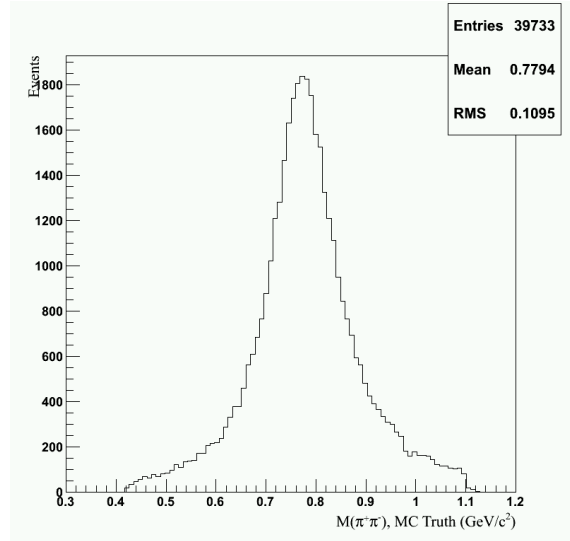
(a) Photon energy in the lab frame for events in signal MC.



(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.

262 3.1 Selection of Photon Candidates

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

268 shower (E_9) to that of the 5 x 5 array of crystals surrounding the center of the shower (E_{25}).
269 See Fig. 4a and Fig. 4b for relevant distributions.

270 3.2 Selection of Pion and Muon Candidates

271 Pion candidates must satisfy $R_{K,\pi} < 0.9$, where $R_{K,\pi}$ is the “Kaon identification variable”
272 defined as the likelihood ratio of the charged track to be due to a kaon versus a pion, and
273 $R_{e,\text{hadron}} < 0.9$, where $R_{e,\text{hadron}}$ is the likelihood ratio of the charged track to be due to
274 an electron versus a hadron. Similarly, muon candidates must satisfy $R_\mu > 0.1$, where
275 R_μ is the likelihood ratio of the charged track to be due to a muon versus other particles
276 detected by the KLM detector subsystem. After imposing the aforementioned requirements,
277 we additionally require there to be four unique charged tracks – two pions and two muons.
278 Events with more than four such tracks are rejected.

279 To select reconstructed tracks that originate near the interaction point, we require pion
280 and muon candidates have $dr < 0.3$ cm and $|dz| < 2$ cm, where dr and dz are impact
281 parameters in the radial and z directions, respectively. We also require pion and muon
282 candidates to have transverse momenta $p_T > 100$ MeV. Candidate muon pairs must have
283 an invariant mass between $9.3 \text{ GeV}/c^2$ and $9.6 \text{ GeV}/c^2$. Candidate pion pairs must have an
284 invariant mass between $0.42 \text{ GeV}/c^2$ and $1.02 \text{ GeV}/c^2$. See Fig. 4c and Fig. 4d for relevant
285 distributions.

286 3.3 Selection of $\Upsilon(5S)$ Candidates

287 $\Upsilon(5S)$ candidates are required to have an invariant mass between 10.2 GeV and 11.5 GeV.
288 The muon pairs of selected $\Upsilon(5S)$ candidates are mass constrained to the nominal $\Upsilon(1S)$
289 invariant mass of $9.460 \text{ GeV}/c^2$. A summary of our selection criteria is shown in Table 3.

290 3.4 Best Candidate Selection

291 Approximately 32% of signal MC events satisfying our selection criteria have multiple signal
292 candidates. This is exclusively due to relatively soft photons. In events with multiple signal
293 candidates, we select the candidate that has an energy most consistent with the center of
294 mass energy of the experimental run. The selected candidates are correctly MC-tagged to full
295 MC truth for signal 90% of the time. For fully reconstructed signal MC events with multiple
296 candidates, our best candidate selection method selects a candidate correctly MC-tagged to
297 full MC truth 88% of the time.

298 4 Signal Monte Carlo Studies

299 4.1 Signal Monte Carlo Distributions

300 To understand properties of signal events, we investigate two invariant mass variables,
301 $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ and $M_{\text{rec}}(\gamma)$, where subscript “fit” indicates that the muon pair is con-
302 strained to the nominal mass of $\Upsilon(1S)$. We define the invariant mass recoiling against X
303 as

Particle Candidate	Selection Criteria
γ	$100 \text{ MeV} \leq E(\gamma) \leq 600 \text{ MeV}$ $20 \text{ MeV} \leq E(\gamma) \leq 5000 \text{ MeV}$
π^\pm, μ^\pm	$dr < 0.3 \text{ cm}$ $dr < 0.4 \text{ cm}$ $dz < 2 \text{ cm}$ $ dz < 4 \text{ cm}$ $p_T > 100 \text{ MeV}/c$
π^\pm PID	$R_{K,\pi} < 0.9$ $R_{e,hadron} < 0.9$
μ^\pm	$R_\mu > 0.10$
ρ^0	$0.420 \text{ GeV}/c^2 < M_{\pi^+\pi^-} < 1.020 \text{ GeV}/c^2$
$\Upsilon(1S)$	$9.3 \text{ GeV}/c^2 < M_{\mu^+\mu^-} < 9.6 \text{ GeV}/c^2$
$\Upsilon(5S)$	$10.2 \text{ GeV}/c^2 < M_{\pi^+\pi^-\mu^+\mu^-\gamma} < 11.5 \text{ GeV}/c^2$ $-0.05 \text{ GeV} < \Delta E < 0.03 \text{ GeV}$
(full event reconstruction)	Exactly four tracks: two muons and two pions

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

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

304 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
305 momentum of system X . Subscript "cm" is used for quantities evaluated in the center of
306 mass reference frame of the experiment. For signal events, $M_{\text{rec}}(\gamma)$ and $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$
307 are two independent ways to estimate the invariant mass of W_{bJ} . Fully reconstructed signal
308 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.
309 We define energy balance ΔE as

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

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

312 There are two effects contributing to the observed width of $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$: (1) the
313 intrinsic width of W_{bJ} , and (2) the charged track reconstruction. Fig. 6 shows $M(\pi^+\pi^+\mu^+\mu^-)$
314 and $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ resolutions for signal events within the signal region and sideband
315 regions (defined in Section 4.2). We model both resolutions as the sum of two Gaussians
316 with the same mean and fit both resolutions. Contribution to $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ resolution
317 from charged track reconstruction is primarily due to pions, since muon pairs are constrained
318 to $\Upsilon(1S)$ invariant mass.

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

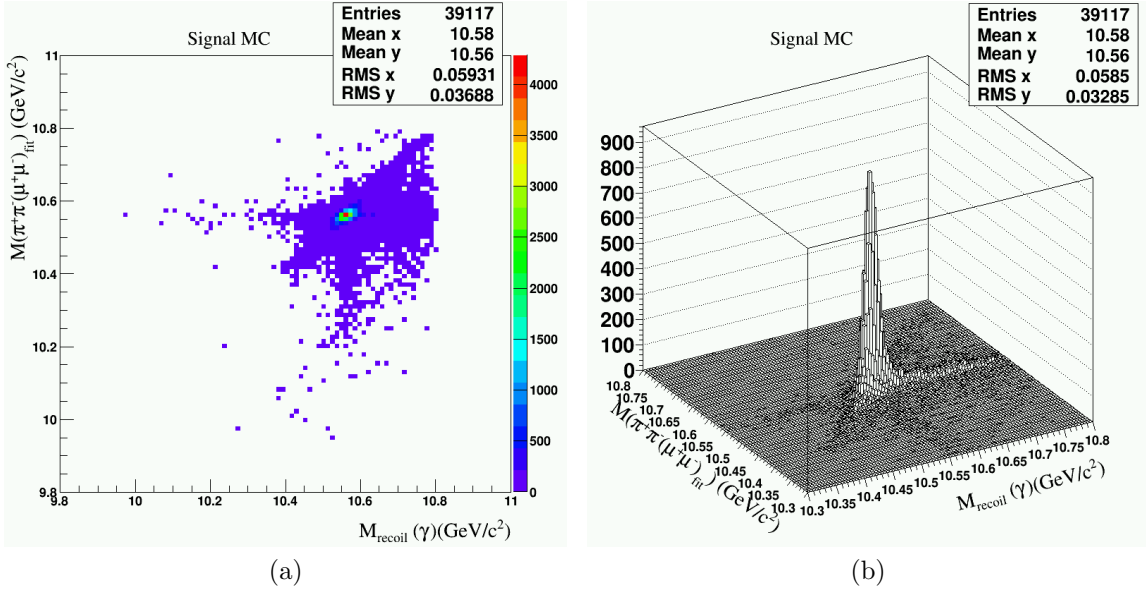
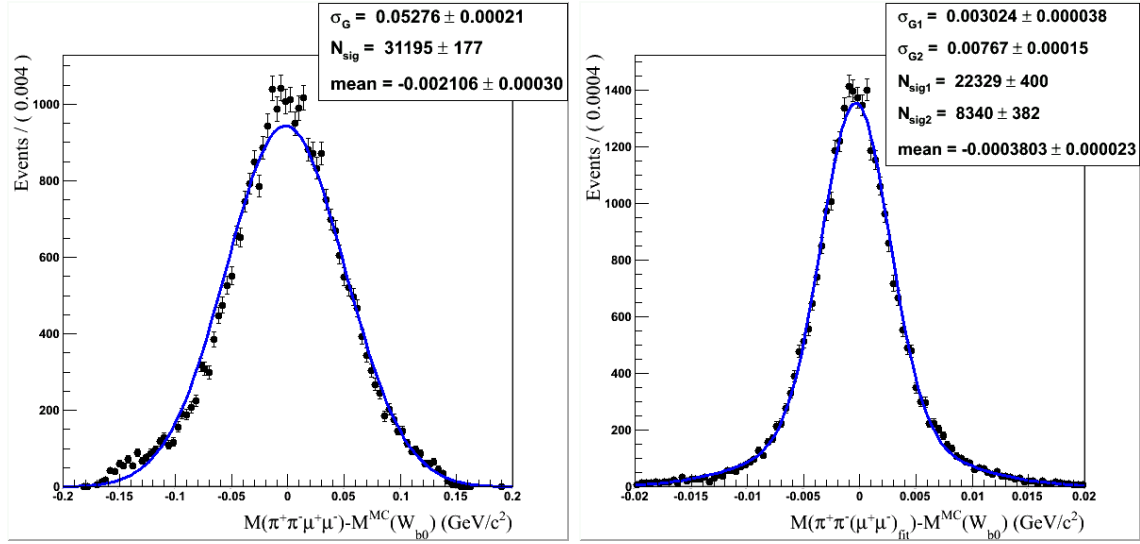


Figure 5: $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ vs $M_{\text{rec}}(\gamma)$ distribution for W_{B0} signal MC events. We show the lego plot in Fig. 5b to emphasize that the tail of $M_{\text{rec}}(\gamma)$ is not as large as it appears in Fig. 5a. Note that Fig. 5b is plotted in a smaller range.

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

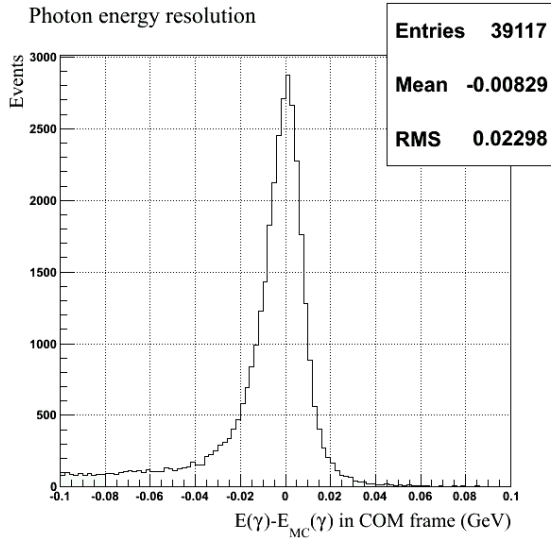
Table 4: Quantities contributing to widths of measured quantities

323 Effects contributing to the observed shape of ΔE include (1) photon energy resolution, (2)
324 charged track resolution, (3) beam energy resolution, and (4) the intrinsic width of W_{bJ} . ΔE
325 resolution is dominated by photon energy resolution as well. The values of relevant widths
326 are listed in Table 4. In signal MC we observe $\sigma_{\Delta E} \approx 12$ MeV, so we take advantage of this
327 excellent energy resolution to select fully reconstructed events. Because the distribution of
328 ΔE is asymmetric (primarily due to leakage from the calorimeter and relatively soft non-
329 signal photons in signal events), we use an asymmetric selection and require $-0.05 \text{ GeV} \leq$
330 $\Delta E \leq 0.03 \text{ GeV}$. This selection cuts out the long tail in the distribution of $M_{\text{rec}}(\gamma)$ and
331 reduces the efficiency by 20%. Note, however, that this selection primarily removes events
332 where the signal photon is not reconstructed. After applying this selection on ΔE , signal
333 reconstruction efficiency becomes approximately 31%. Fig. 7 displays ΔE resolution as well
334 as quantities contributing to ΔE resolution.

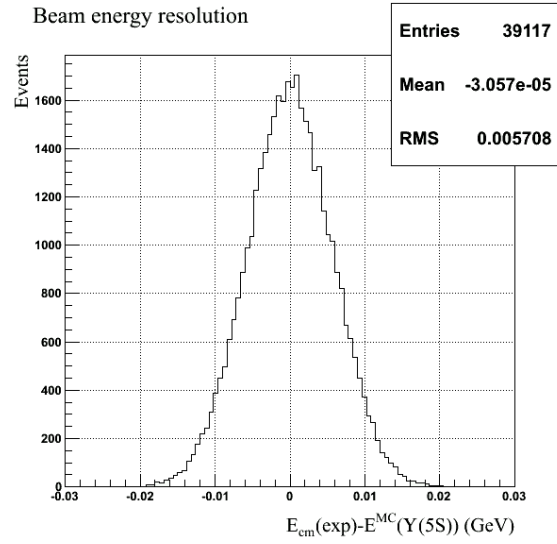


(a) $M(\pi^+\pi^+\mu^+\mu^-)$ resolution. Note that muons are not mass constrained. (b) $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ resolution (muons are mass constrained).

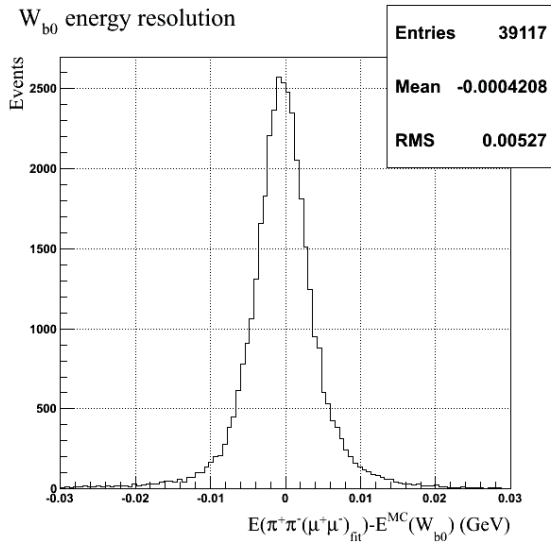
Figure 6: $M(\pi^+\pi^+\mu^+\mu^-)$ and $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ resolutions for signal events within the signal region and sideband regions (defined in Section 4.2). Note that the horizontal scales are different.



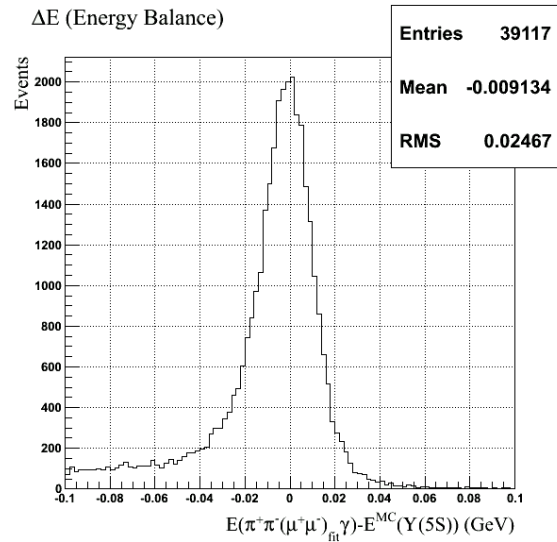
(a) Signal photon energy line shape in the COM reference frame.



(b) Beam energy resolution.



(c) $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ 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: ΔE resolution and quantities contributing to ΔE resolution.

335 4.2 Description of the Signal Region

336 Table 5 contains the definitions of four important regions in this analysis. Before investigating
 337 data, we blind the region where we expect to find signal. We refer to this region as the
 338 blinded region. The invariant masses of W_{b0} , W_{b1} , and W'_{b0} and W_{b2} are expected to be at
 339 the $B\bar{B}$, $B^*\bar{B}$, and $B^*\bar{B}^*$ thresholds, respectively. The blinded region is defined as the region
 340 between the $B\bar{B}$ and $B^*\bar{B}^*$ thresholds plus an additional margin of 70 MeV on either side.
 341 This corresponds to $10.49 \text{ GeV}/c^2 \leq M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}}) \leq 10.72 \text{ GeV}/c^2$. The boundary on

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

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

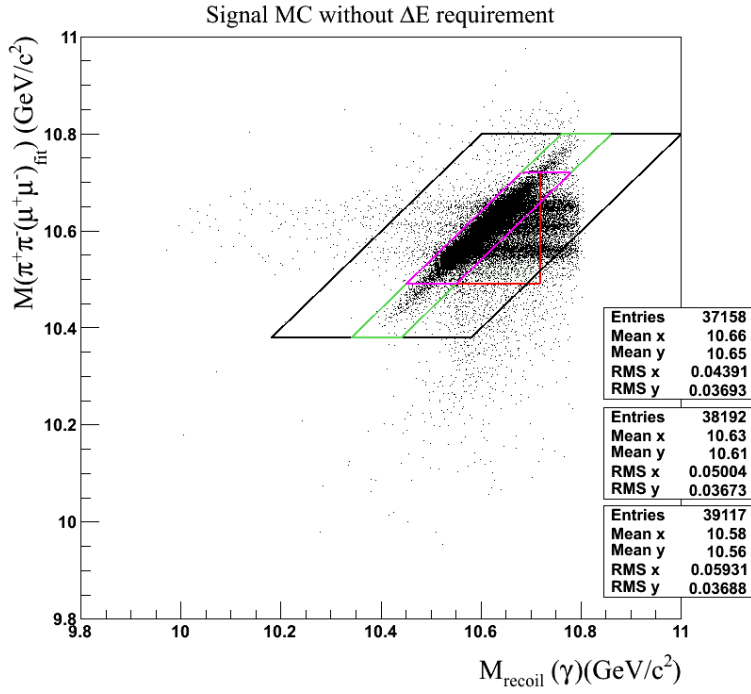
342 the left side of the region is defined by the sloped line $M_{\text{rec}}(\gamma) \geq M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}}) - 0.04$
343 GeV/c^2 which lies parallel to the main diagonal. Approximately 20% of signal events are
344 located in the long right tail of the distribution of $M_{\text{rec}}(\gamma)$. A phase space boundary on
345 the right side of the plot at $M_{\text{rec}}(\gamma) \approx 10.75 \text{ GeV}/c^2$ forces this long tail of the $M_{\text{rec}}(\gamma)$
346 distribution into a smaller region for the higher mass W_{bJ} states. Hence, we do not define
347 a sloped boundary line as the right side of the signal region – a diagonal boundary would
348 exclude more signal events for the lower mass states because of the aforementioned phase
349 space boundary compressing the tail. Instead, we define the vertical line boundary $M_{\text{rec}}(\gamma) \leq$
350 $10.72 \text{ GeV}/c^2$ which assures that approximately equal percentages of signal would be blinded
351 for all masses of W_{bJ} states.

352 We define the signal region as the region contained within $10.49 \text{ GeV}/c^2 \leq M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}}) \leq$
353 $10.72 \text{ GeV}/c^2$ satisfying $-0.05 \text{ (GeV)} \leq \Delta E \leq 0.03 \text{ GeV}$. The ΔE requirement selects only
354 fully-reconstructed signal events, where signal is peaking.

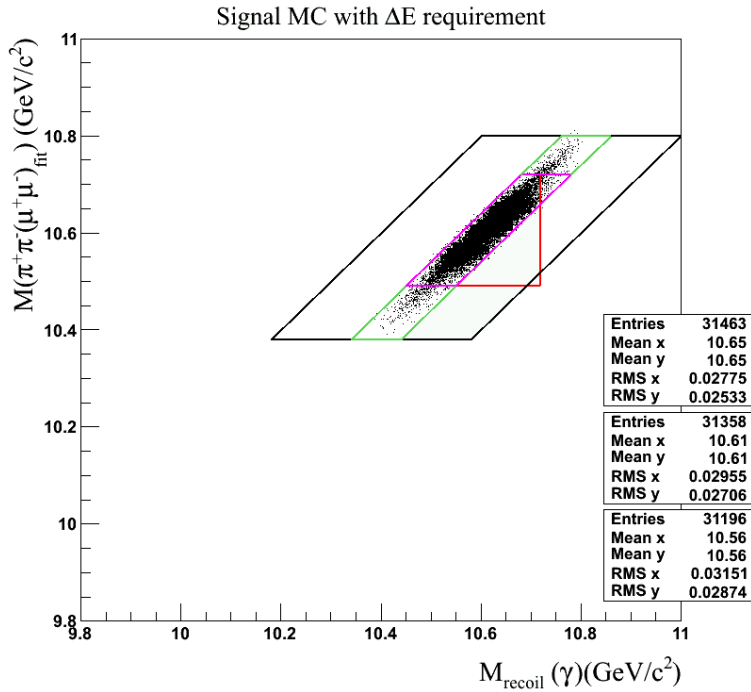
355 The sideband region is essentially an extension of the signal region, defined as the
356 regions within $10.38 \text{ GeV}/c^2 \leq M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}}) \leq 10.49 \text{ GeV}/c^2$ and $10.72 \text{ GeV}/c^2 \leq$
357 $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}}) \leq 10.80 \text{ GeV}/c^2$ satisfying $-0.05 \text{ (GeV)} \leq \Delta E \leq 0.03 \text{ GeV}$.

358 We additionally define the grand sideband region as the region within $10.38 \text{ GeV}/c^2 \leq$
359 $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}}) \leq 10.80 \text{ GeV}/c^2$ satisfying $-0.20 \text{ GeV} \leq \Delta E \leq 0.20 \text{ GeV}$. This region is
360 used when studying background in data.

361 Fig. 8 displays these four regions with our three signal MC samples. It is important to
362 note that the blinded region is not completely contained within the grand sideband region
363 and the signal region is not completely contained within the blinded region. This is due to
364 historical reasons, as the blinded region was defined prior to the use of ΔE in this analysis.



(a) Does not include ΔE requirement.



(b) Includes ΔE 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 ΔE 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.

365 4.3 Trigger Simulation

366 Relatively low final state particle multiplicity of our signal events requires us to investigate
367 trigger efficiency. Trigger efficiency is simulated after full reconstruction. We find correlations
368 between trigger efficiency and kinematics. Fig. 9 shows various 2-dimensional distributions
369 of $\mu^+ \cos(\theta)$ vs $\mu^- \cos(\theta)$, and we see that events failing to satisfy trigger are more likely
370 to have one of the muons at a small angle with respect to the beam axis ($|\cos(\theta)| \geq 0.8$).
371 Fig. 10 shows additional distributions of $\mu^+ \cos(\theta)$ vs $\mu^+ \cos(\theta)$ which we use to determine
372 trigger efficiencies. When neither muon is at a small angle with respect to the beam axis,
373 trigger efficiency is 96%. When one of the muons is at a small angle with respect to the beam
374 axis, trigger efficiency drops to 89%. For all generated signal MC events, trigger efficiency is
375 approximately 94%. After accounting for trigger efficiency, our overall efficiency drops from
376 31% to 29%.

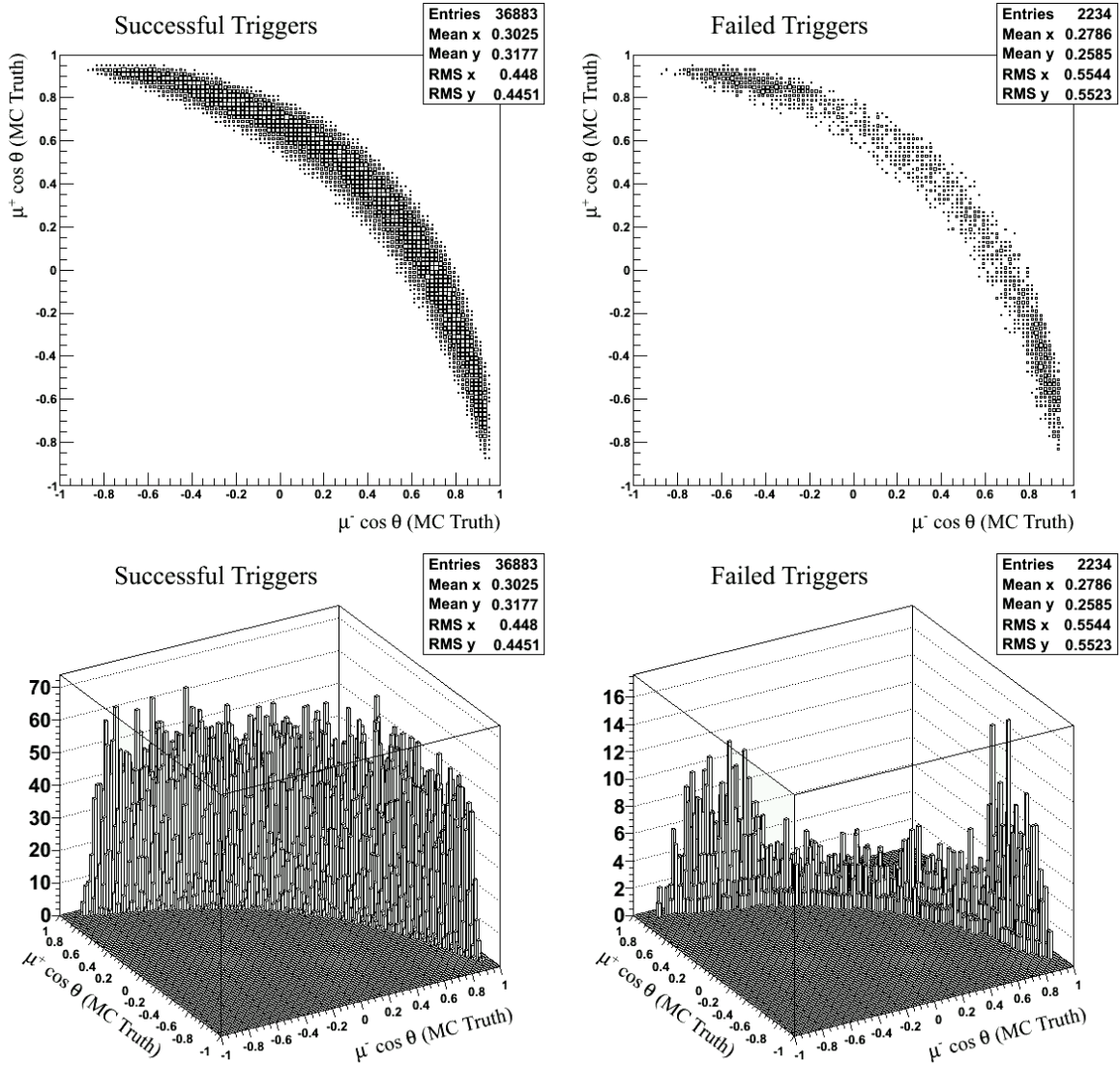


Figure 9: Reconstructed signal MC events that satisfy the offline trigger selection are plotted on the left, while events that fail the offline trigger selection are plotted on the right. We observe that events satisfying the trigger criteria are distributed more or less uniformly for kinematically allowed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

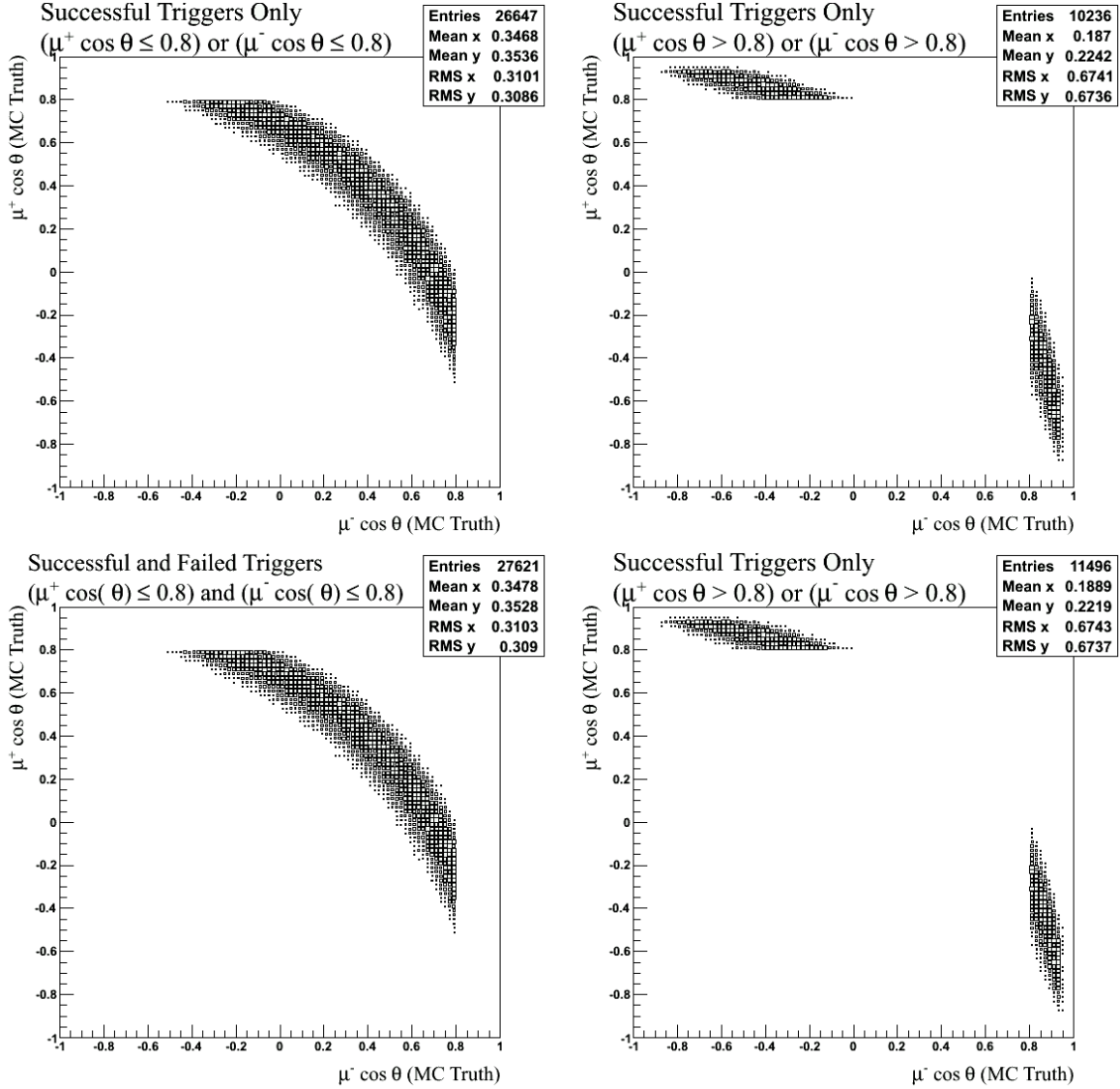


Figure 10: All reconstructed events in which both muons are generated with $|\cos(\theta)| < 0.8$ are plotted in the left two figures. Trigger efficiency for such events is approximately $(96 \pm 4)\%$. In the right two figures, we plot all reconstructed events where one of the muons is generated with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

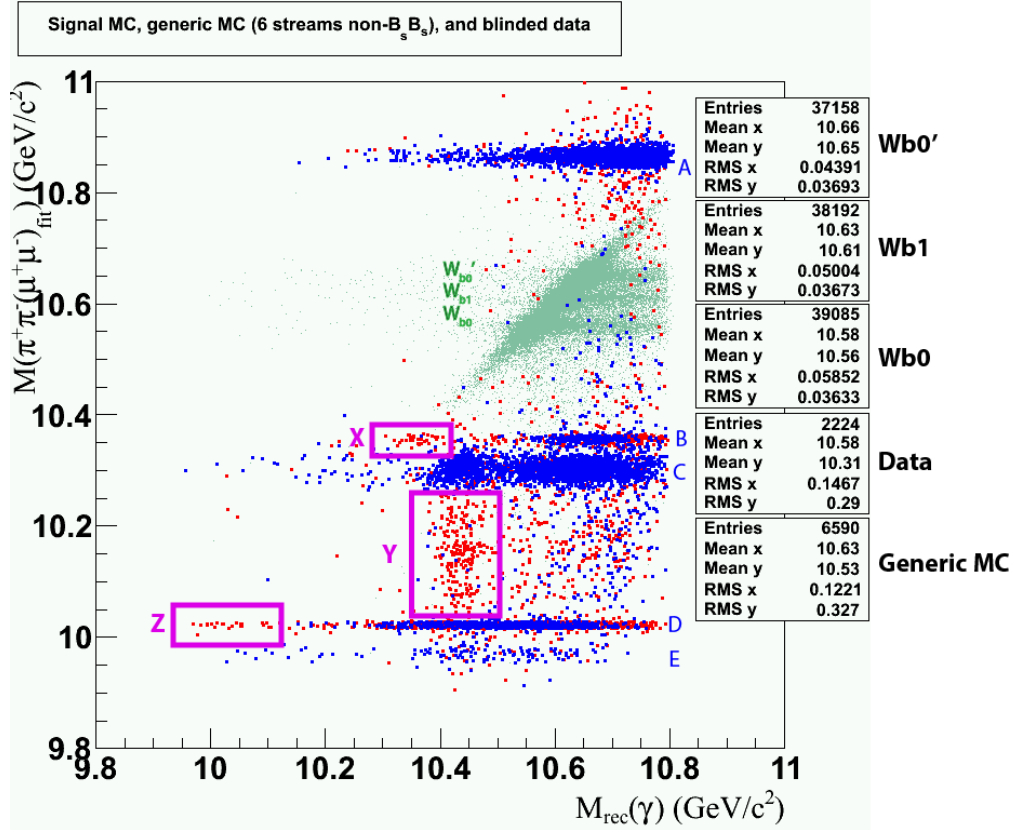


Figure 11: W_{b0} , W_{b1} , and W'_{b0} signal MC (light green), six streams of non- $B_s B_s$ generic MC (blue), and data with the signal region blinded (red).

5 Background Studies

5.1 Generic Monte Carlo and Blinded Data

Fig. 11 shows the distribution of $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ vs $M_{\text{rec}}(\gamma)$ 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 $B_s B_s$ generic MC events pass our selection criteria. A large number of non- $B_s B_s$ events do satisfy our selection criteria, though they fall primarily outside the signal region. The ΔE requirement excludes most of these background events. The most prominent non- $B_s B_s$ background sources are (cascade) dipion transitions to $\Upsilon(1S)$. We observe an enhancement in generic MC within the blinded region due to the decay $\Upsilon(5S) \rightarrow \Upsilon(2S)\pi^+\pi^-$, $\Upsilon(2S) \rightarrow \Upsilon(1S)\pi^+\pi^-$ where the selected signal pion candidates did not come from the same parent. The enhancement is removed when the ΔE constraint is applied, as such background events are not fully reconstructed.

We observe several regions where data events are clustering but generic MC events are not, and we have identified the likely origins of these events. The regions labeled X and Z in Fig. 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	$\Upsilon(5S) \rightarrow \Upsilon(1S)\pi^+\pi^- \rightarrow \mu^+\mu^-\pi^+\pi^-$
B	$\Upsilon(5S) \rightarrow \Upsilon(3S)\pi^+\pi^- \rightarrow \Upsilon(1S)\pi^+\pi^-\pi^+\pi^- \rightarrow \mu^+\mu^-\pi^+\pi^-\pi^+\pi^-$ $\Upsilon(5S) \rightarrow \Upsilon(3S)\pi^0\pi^0 \rightarrow \Upsilon(1S)\pi^+\pi^-\pi^0\pi^0 \rightarrow \mu^+\mu^-\pi^+\pi^-\pi^0\pi^0$
C	$\Upsilon(5S) \rightarrow \Upsilon(2S)\pi^+\pi^- \rightarrow \Upsilon(1S)\pi^+\pi^-\pi^+\pi^- \rightarrow \mu^+\mu^-\pi^+\pi^-\pi^+\pi^-$ $\Upsilon(5S) \rightarrow \Upsilon(2S)\pi^+\pi^- \rightarrow \Upsilon(1S)\pi^0\pi^0\pi^+\pi^- \rightarrow \mu^+\mu^-\pi^0\pi^0\pi^+\pi^-$
D	$\Upsilon(5S) \rightarrow \Upsilon(2S)\pi^0\pi^0 \rightarrow \Upsilon(1S)\pi^+\pi^-\pi^0\pi^0 \rightarrow \mu^+\mu^-\pi^+\pi^-\pi^0\pi^0$
E	$\Upsilon(5S) \rightarrow \Upsilon(3S)\pi^+\pi^- \rightarrow \Upsilon(1S)\pi^0\pi^0\pi^+\pi^- \rightarrow \mu^+\mu^-\pi^0\pi^0\pi^+\pi^-$
X	$e^+e^- \rightarrow \Upsilon(3S)\gamma \rightarrow \Upsilon(1S)\pi^+\pi^-\gamma \rightarrow \mu^+\mu^-\pi^+\pi^-\gamma$
Y	Various processes involving $\chi_{bJ}(1P) \rightarrow \gamma\Upsilon(1S)$, <i>e.g.</i> $\Upsilon(5S) \rightarrow \Upsilon(1D)\pi^+\pi^-$, where $\Upsilon(1D) \rightarrow \gamma\chi_{bJ}(1P)$
Z	$e^+e^- \rightarrow \Upsilon(2S)\gamma \rightarrow \Upsilon(1S)\pi^+\pi^-\gamma \rightarrow \mu^+\mu^-\pi^+\pi^-\gamma$

Table 6: Backgrounds labeled in Fig. 11.

394 fully reconstructed, and thus fall along the main diagonal of the plot. The region labeled
395 Y includes processes involving radiative decays of $\chi_{bJ}(1P)$. These events have additional
396 final state particles that are not reconstructed, and hence they fall below the main diagonal
397 where $\Delta E < 0$. Events in categories X, Y, and Z are not of concern to us, since they are
398 located far from the signal region.

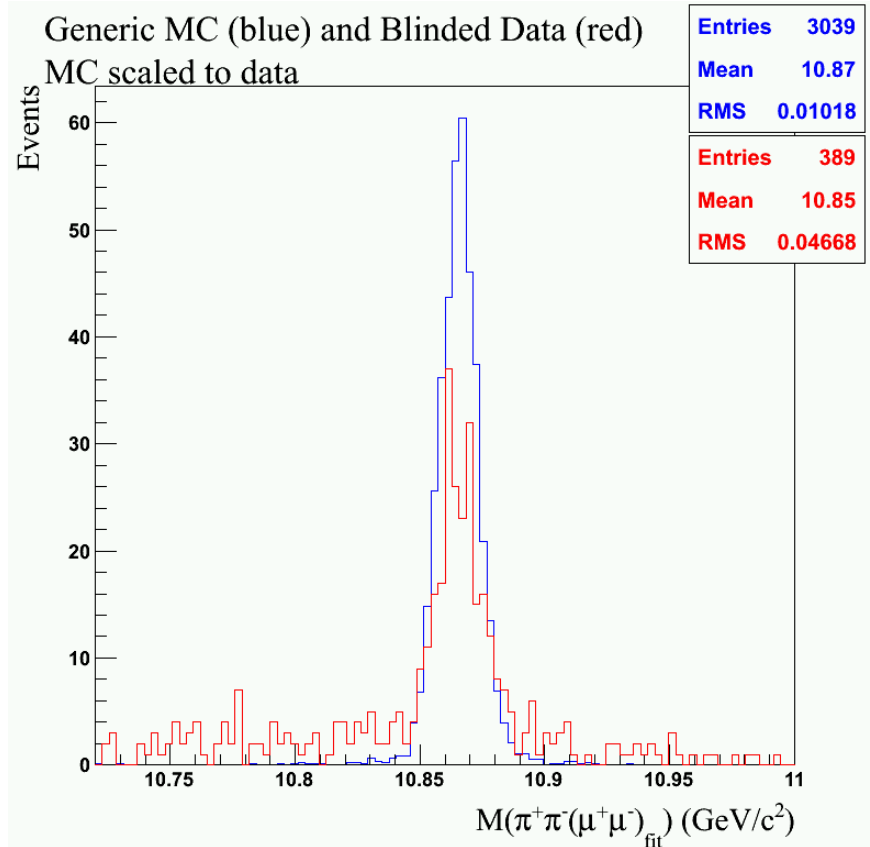


Figure 12: $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ distributions for $\Upsilon(5S) \rightarrow \Upsilon(1S)\pi^+\pi^-$ events (label 'A' in Table 6). Distributions for generic MC and blinded data are shown in blue and red, respectively. Generic MC does not include ISR and is normalized to the number of data events shown in the plotted range. We choose $10.72 \text{ GeV}/c^2$ as the lower limit of the range plotted, since lower masses would include the blinded region.

399 6 Background from $\Upsilon(5S) \rightarrow \Upsilon(1S)\pi^+\pi^-$ with Initial 400 State Radiation (ISR)

401 We find that dipion transitions to $\Upsilon(1S)$ (labeled 'A' in Fig. 11) have a much longer tail
402 in data than in generic MC. This difference is shown in Fig. 12, and is due to initial state
403 radiation (ISR). This tail contaminates the signal region, so we generate additional MC
404 samples with ISR to study these backgrounds.

405 6.1 $\Upsilon(5S) \rightarrow \Upsilon(1S)\pi^+\pi^-$ ISR Monte Carlo Sample

406 The VectorISR model [18] is used to simulate ISR. We reweight the ISR photon energy
407 spectrum according to the correct radiator function up to order α^2 [21] using a Monte Carlo
408 method. After reweighting, there are approximately 110,000 events in our MC sample. A
409 distribution of the reweighted ISR spectrum is shown in Fig. 13.

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

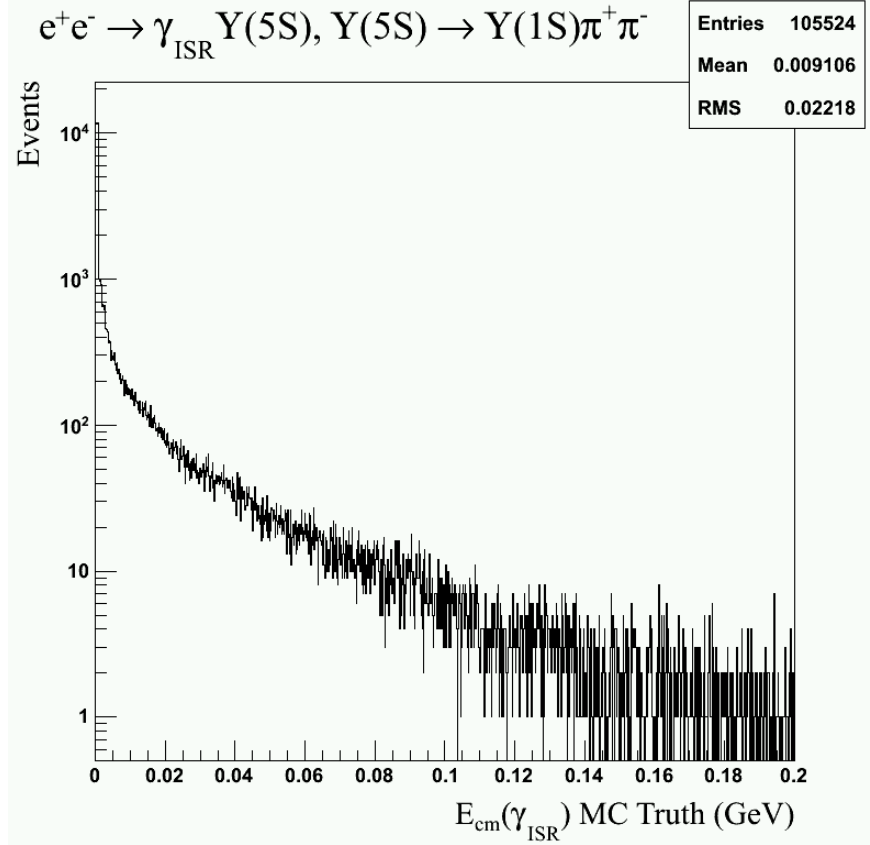


Figure 13: Reweighted ISR energy spectrum for $e^+e^- \rightarrow \gamma_{\text{ISR}}\Upsilon(5S), \Upsilon(5S) \rightarrow \Upsilon(1S)\pi^+\pi^-$. Note that a log scale is used for the vertical axis.

411 $\Upsilon(1S)\pi^+\pi^-$ events with ISR. Recall that the two plotted variables represent two independent
 412 ways to estimate the invariant mass of W_{bJ} , and therefore fully reconstructed events fall along
 413 the main diagonal of this plot. When the ISR photon of these backgrounds is selected as
 414 the signal photon candidate, these backgrounds are also fully reconstructed and fall along
 415 the main diagonal within the signal region. Approximately 3% of reconstructed events fall
 416 in the signal region. Fortunately, these backgrounds do not peak in the signal region in the
 417 distribution of $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$.

418 We simulate $\Upsilon(5S) \rightarrow \Upsilon(1S)\pi^+\pi^-$ with ISR using the models listed in Table 7. To
 419 determine if the choice of decay models affects the distribution shape of our signal variable
 420 $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$, we generate additional samples using the VVPIPI decay [18] model for
 421 $\Upsilon(5S) \rightarrow \Upsilon(1S)\pi^+\pi^-$ and the VLL decay model [18] for $\Upsilon(1S) \rightarrow \mu^+\mu^-$. Fig. 15 shows the
 422 distribution of $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ for two different MC samples generated using different
 423 decay models.

424 We find that the choice of decay model has only a small effect on the shape of the
 425 $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ distribution. Furthermore, we plot the $\cos\theta$ of μ^+ in Fig. 16 and find
 426 that the presence of ISR has only a small effect on the the angular distributions of muons. To
 427 determine if ISR affects the width of the $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ distribution for signal processes
 428 $\Upsilon(5S) \rightarrow \gamma W_{bJ}$, we generate additional MC samples for the the signal process $\Upsilon(5S) \rightarrow \gamma W_{bJ}$

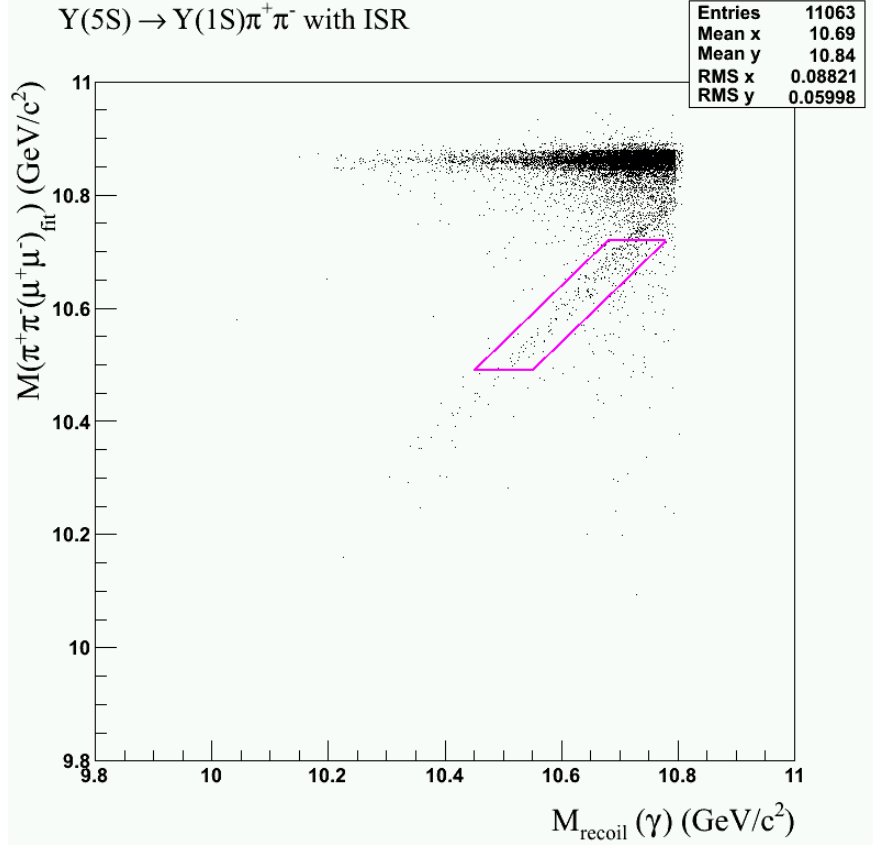


Figure 14: A 2-dimensional $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ vs $M_{\text{rec}}(\gamma)$ distribution for $\Upsilon(5S) \rightarrow \Upsilon(1S)\pi^+\pi^-$ events with ISR (after reweighting). The signal region is outlined in magenta.

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

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

429 with ISR. We find that ISR has practically no effect on the width of the distribution of
 430 $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$.

431 6.2 Background Shape of $\Upsilon(5S) \rightarrow \Upsilon(1S)\pi^+\pi^-$ with ISR

432 It is likely that events due to $\Upsilon(5S) \rightarrow \Upsilon(1S)\pi^+\pi^-$ with ISR are a dominant source of
 433 backgrounds in the signal region. The rightmost plot in Fig. 17 shows the distribution of
 434 these events within the signal region for our reweighted MC. To see how the selection on
 435 ΔE affects the background shape, we loosen up the selection on ΔE in the left and middle

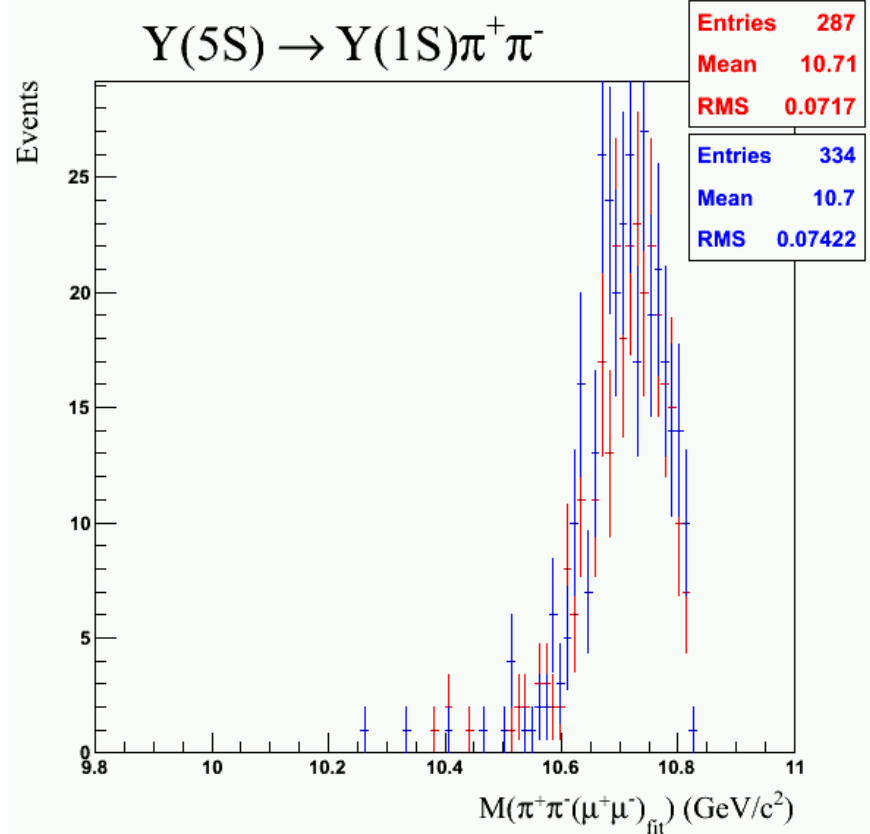


Figure 15: The distribution shown in blue is for events where $\Upsilon(5S) \rightarrow \Upsilon(1S)\pi^+\pi^-$ is generated using VVPIPI model [18] and $\Upsilon(1S) \rightarrow \mu^+\mu^-$ using VLL model [18]. The distribution shown in red is for events generated using PHSP model [18] for both processes. Neither samples contain ISR nor FSR, so they only differ by their decay models. The shapes of their $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ distributions are very similar. Note that although there is a difference in efficiency between the two samples, this is unimportant for our analysis, because we are only interested in possible difference between the shapes of these distributions.

436 plots in Fig. 17. Imposing a selection on ΔE has only a small effect on the shape of these
 437 backgrounds in the signal region.

438 To determine if we can use this MC sample to estimate the number of background events
 439 in the signal region, we divide the grand sideband region shown in Fig. 18 into four smaller
 440 regions as defined in Table 8 and observe if the number of events in MC scales uniformly
 441 to data across all regions. Table 9 shows the number of ISR MC events and data events
 442 within the regions of interest. We see that ISR MC does not scale uniformly across all
 443 regions. While ISR studies improve the quality of our analysis and provide us with useful
 444 information about the shape of this background in the signal region, including ISR into our
 445 analysis does not sufficiently improve the scaling between data and MC in different regions
 446 of grand sideband.

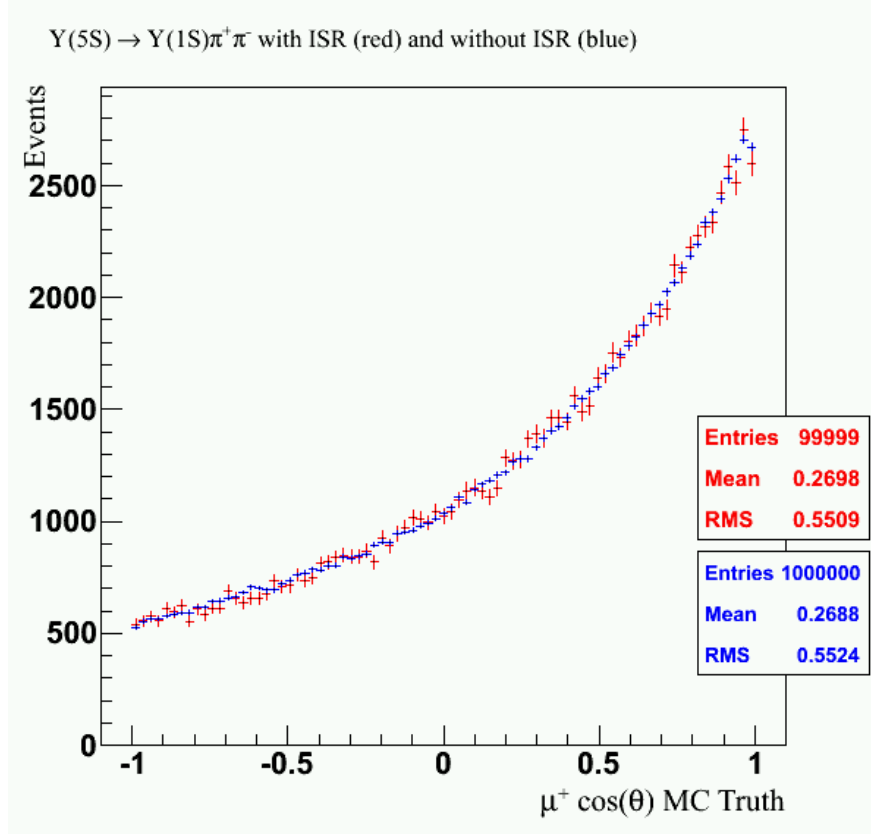


Figure 16: Distributions of $\cos \theta$ for μ^+ for $\Upsilon(5S) \rightarrow \Upsilon(1S)\pi^+\pi^-$ events. The distribution shown in red is for events generated with ISR while the distribution shown in blue is for events generated without ISR. Events in both distributions are generated using PHSP model for both $\Upsilon(5S) \rightarrow \Upsilon(1S)\pi^+\pi^-$ and $\Upsilon(1S) \rightarrow \mu^+\mu^-$. The blue distribution is normalized to the number of events in the red distribution.

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

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

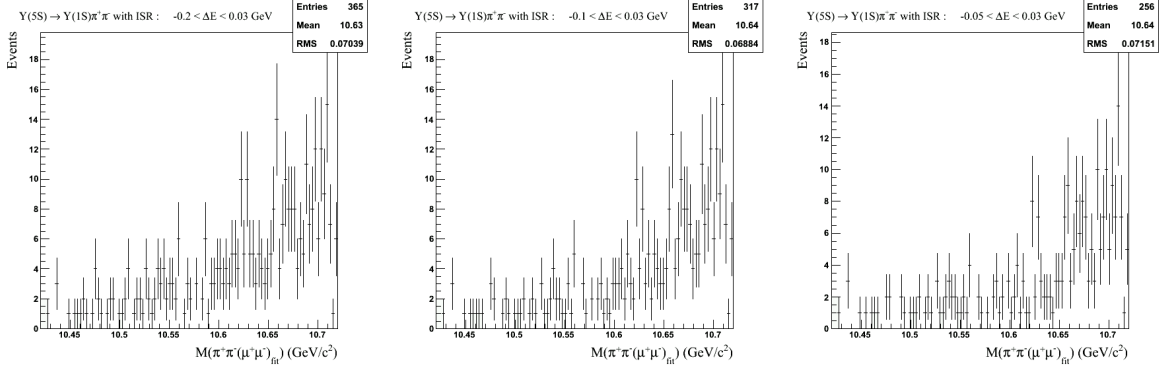


Figure 17: Distributions of $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ for $\Upsilon(5S) \rightarrow \Upsilon(1S)\pi^+\pi^-$ with ISR in the signal region for different ΔE requirements. The leftmost distribution requires $-0.2 \text{ GeV} < \Delta E < 0.03 \text{ GeV}$, the middle distribution requires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

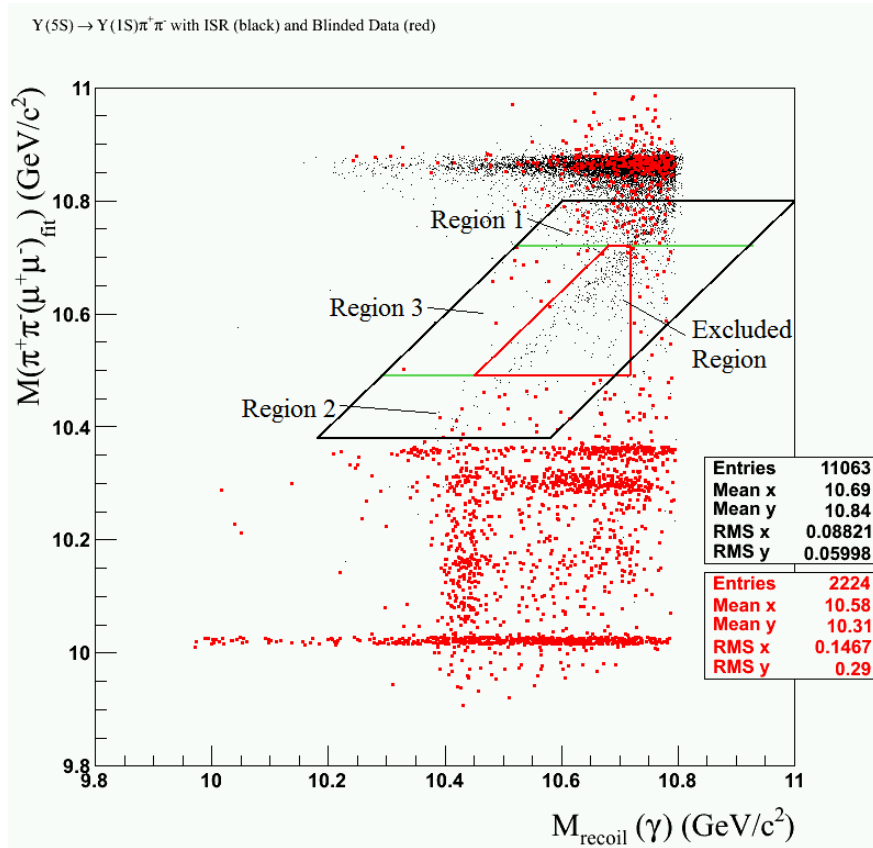


Figure 18: Subdivisions of the grand sideband region. The Excluded Region is not considered in this analysis.

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 $\Upsilon(5S) \rightarrow Z_b^{(\prime)\pm} \pi^\mp$

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 $\Upsilon(1S) \rightarrow \mu^+\mu^-$) 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 $\Upsilon(5S) \rightarrow Z_b^\pm \pi^\mp$ followed by $Z_b^\pm \rightarrow \Upsilon(1S)\pi^\mp$, $\Upsilon(1S) \rightarrow \mu^+\mu^-$. We also generated an additional 50,000 events for $\Upsilon(5S) \rightarrow Z_b'^\pm \pi^\mp$. 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 $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ vs $M_{\text{rec}}(\gamma)$ is shown in Fig. 19 for both samples after applying our selection criteria for the W_{bj} analysis. Fig. 20 shows the distribution of $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ 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 $\Upsilon(5S) \rightarrow Z_b^{(\prime)\pm} \pi^\mp$ 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 $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ using RooFit [22]. We model the signal distribution of $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ 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 $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ distribution in signal MC remains practically the same after applying our ΔE requirement and after including ISR. Therefore, we fix the width of our signal PDF. We set the width of the Breit-Wigner to be $\sigma_{BW} = 15 \text{ MeV}/c^2$ to match the intrinsic width of Z_b and Z_b' . The

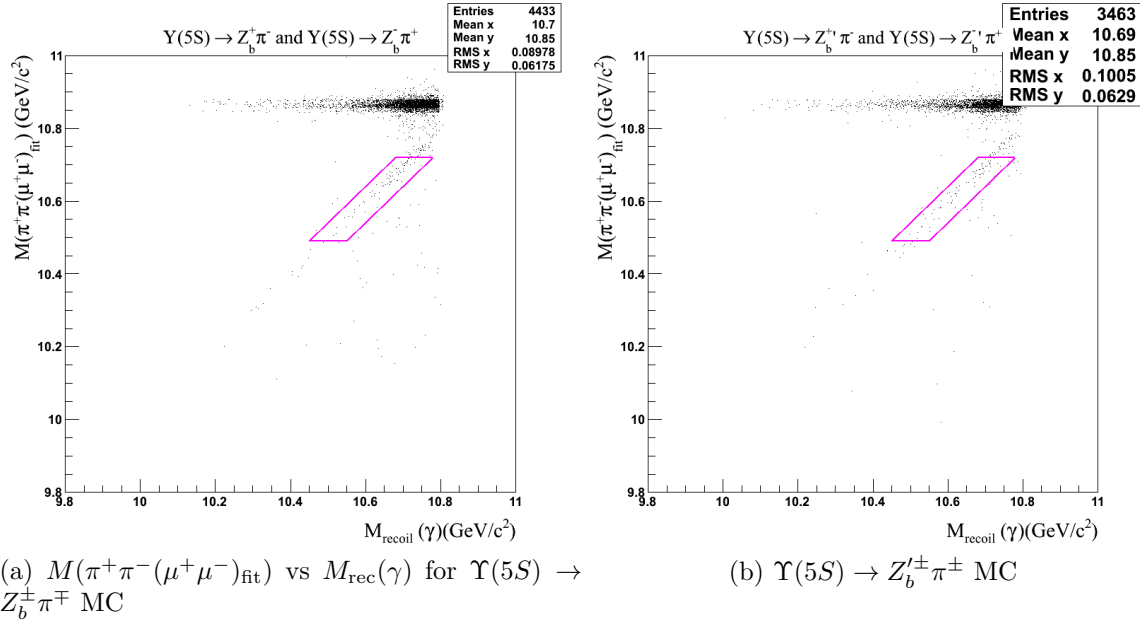
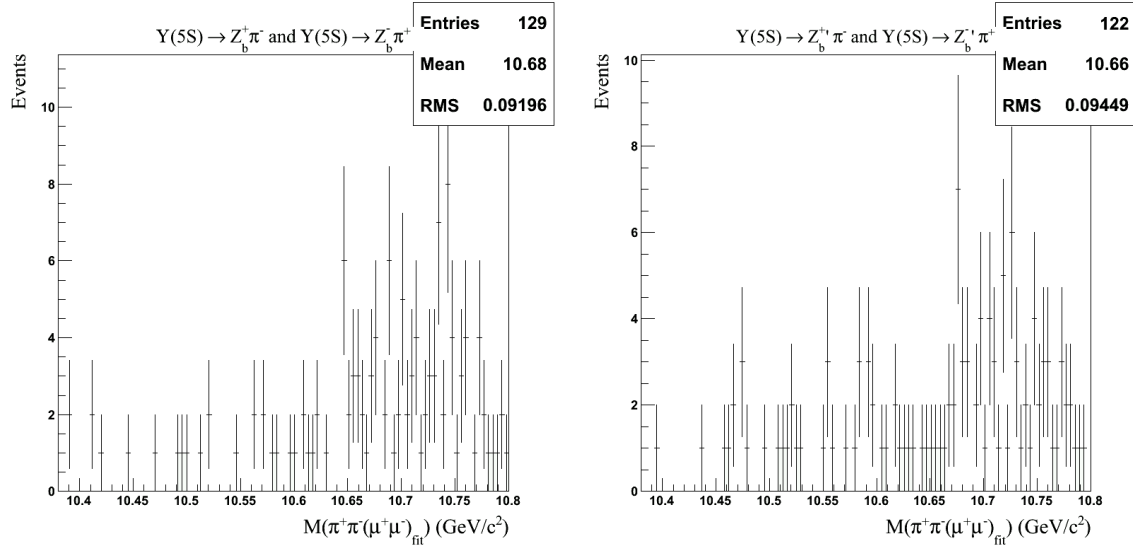


Figure 19: The distribution of $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ vs $M_{\text{recoil}}(\gamma)$ for $\Upsilon(5S) \rightarrow Z_b^{(\prime)\pm}\pi^\mp$ MC.



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

Figure 20: The distribution of $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ for $\Upsilon(5S) \rightarrow Z_b^{(\prime)\pm}\pi^\mp$ MC for events inside the signal and sideband region.

476 widths of the Gaussians used in convolution are $\sigma_{G_1} \approx 3$ MeV/c² and $\sigma_{G_2} \approx 7.7$ MeV/c² to
 477 match the widths obtained from the fit to $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ resolution. We let the mean
 478 of Breit-Wigner float within the fit, as W_{bJ} could be observed at different invariant masses
 479 for different spins J . Table 10 lists the values of parameters used in our signal PDF model.
 480 We use an exponential $e^{\lambda x}$ to model background contributions due to ISR as well as

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

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

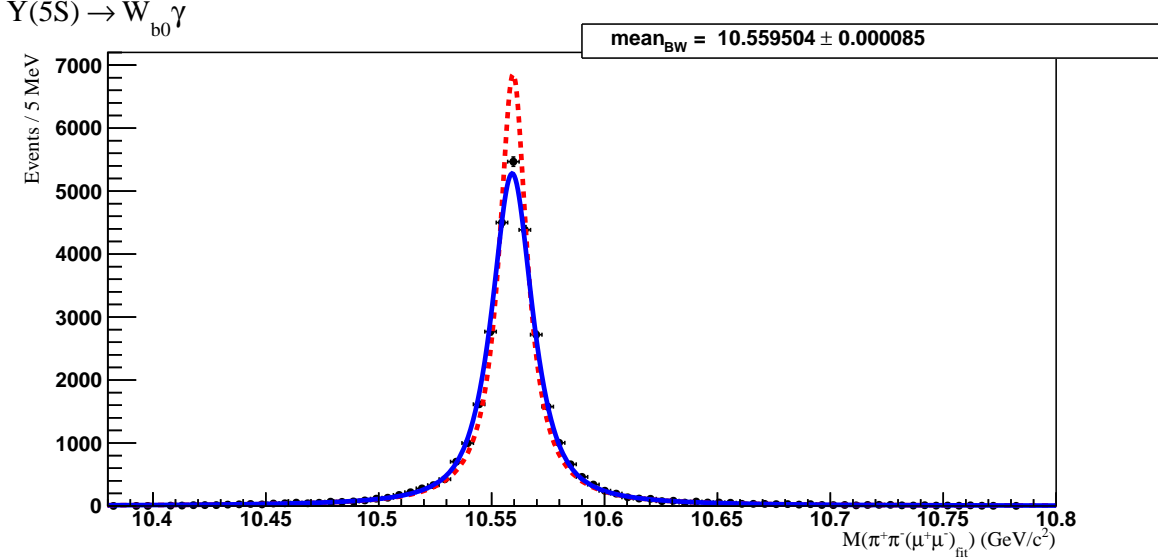
481 possible non-resonant contribution from dimuon continuum events. Strictly speaking, the
482 background distribution deviates from an exponential at $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}}) \approx 10.75$ GeV/c².
483 because of the phase space boundary at $M_{\text{rec}}(\gamma) \approx 10.75$ GeV/c² seen in Fig. 5. This
484 ever-present effect can be seen in figures showing the distribution of $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ for
485 background events with our ΔE requirement (*e.g.* see Fig. 15, Fig. 21b, Fig. 21c). This
486 shortcoming of our analysis will be taken care of in the next version of this Note. We would
487 like to remark that the observed fall-off effect is easy to understand and describe in the
488 model used for fitting, as it is exclusively due to the boundary of phase space.

489 To estimate the number of background events we expect in the signal region, we per-
490 form an extended unbinned maximum likelihood fit to data only in the sideband regions.
491 To account for uncertainty in the number of data events in the sideband region, we fit
492 $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ within the range of 10.38 GeV/c² and 10.80 GeV/c² when extracting
493 signal yield. This range corresponds to the signal region and sideband regions combined.
494 From the fit, we obtain $\lambda = 3.7951$. We extract 59 ± 11 background events in the signal
495 region and sideband regions combined. We expect 27 ± 5 of these background events to be
496 in the signal region alone. Fits to W_{b0} signal MC, $\Upsilon(5S) \rightarrow \Upsilon(1S)\pi^+\pi^-$ MC with ISR MC,
497 and data in the sidebands are shown in Fig. 21.

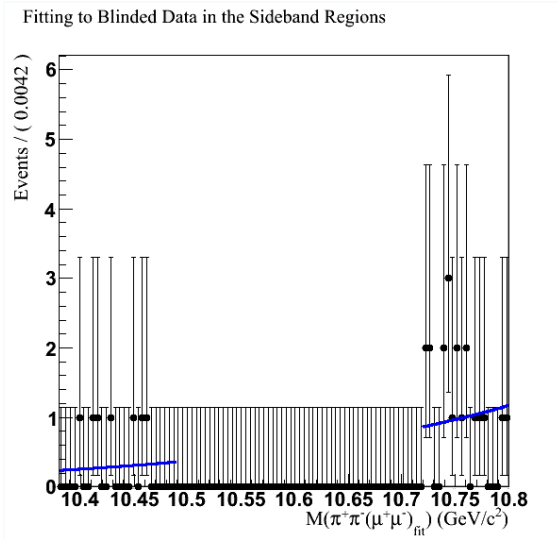
498 8.2 Confidence Belts

499 To construct a 90% confidence belt (with 5% on each side of the belt), we perform ensemble
500 tests. Each ensemble test consists of 1000 toy MC experiments. In each toy MC experiment,
501 we generate N_{sig} signal events and N_{bkg} background events according to their respective
502 PDF lineshapes used for fitting signal and background. We then fit the generated events in
503 the range $10.38 \text{ GeV}/c^2 < M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}}) < 10.80 \text{ GeV}/c^2$ to our combined signal and
504 background PDF to extract the fitted number of signal events $N_{\text{sig}}^{\text{fit}}$.

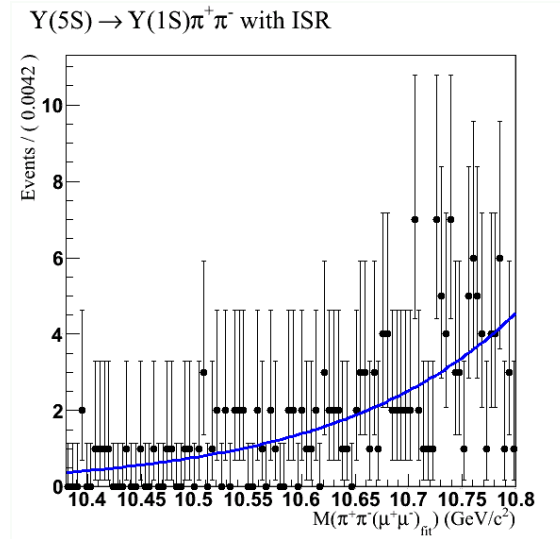
505 We construct our 90% confidence belt by performing ensemble tests with $N_{\text{bkg}}^{\text{gen}} = 59$ for
506 values of $N_{\text{sig}}^{\text{gen}}$ from 0 to 70. We additionally construct a 90% confidence belt where we allow
507 Poisson fluctuation in $N_{\text{bkg}}^{\text{gen}}$. These confidence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



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



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

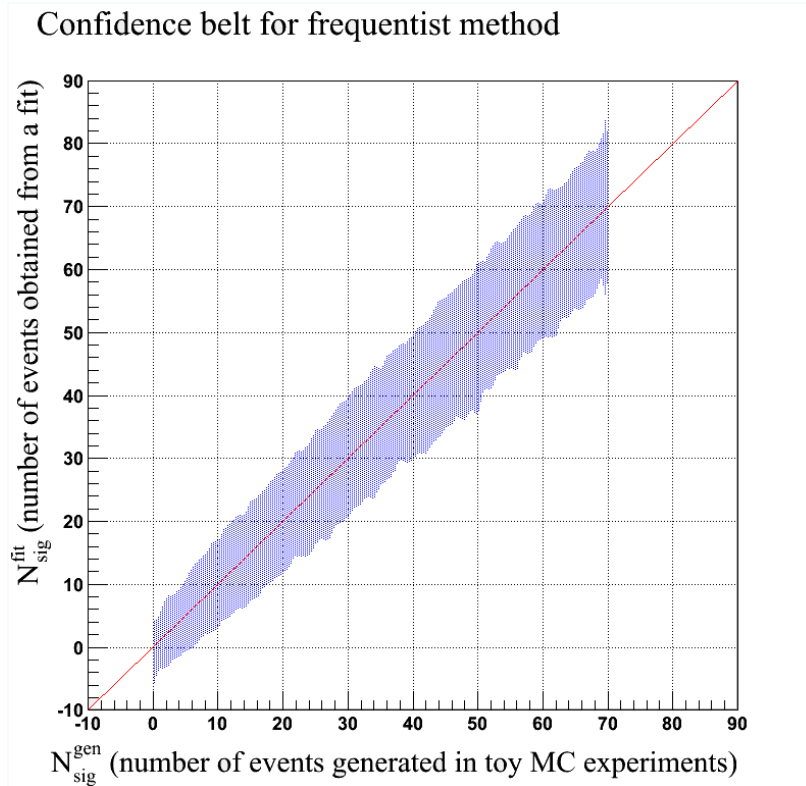


(c) Fit result for the distribution of $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ for $\Upsilon(5S) \rightarrow \Upsilon(1S)\pi^+\pi^-$ with ISR for events in the signal region.

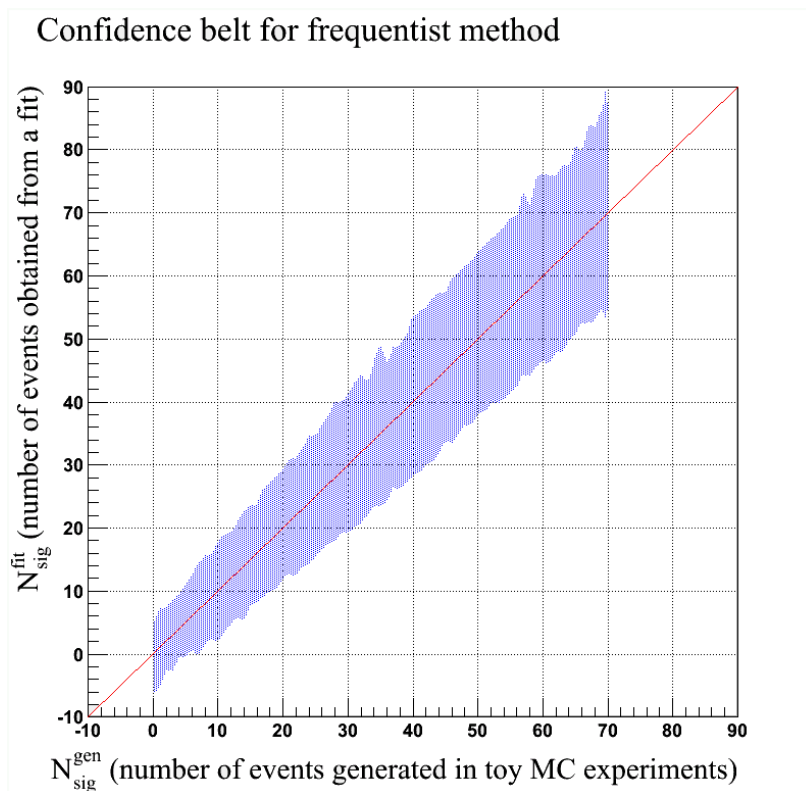
Figure 21: Fitting background MC and data

508 8.3 Linearity Study

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



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



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

Figure 22: 90% confidence belts for frequentist method.

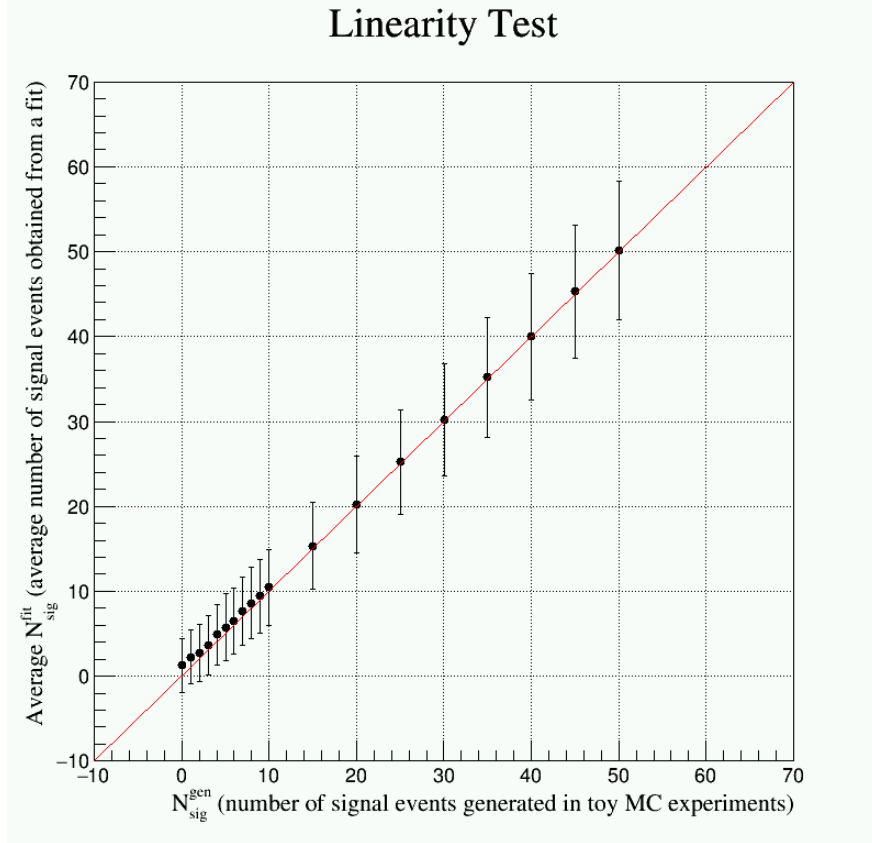
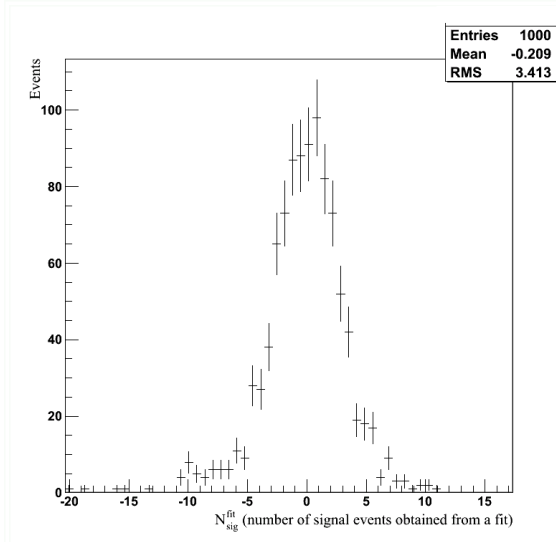


Figure 23: Average $N_{\text{sig}}^{\text{fit}}$ for varying values of $N_{\text{sig}}^{\text{gen}}$. The solid black line is the result of fitting these points to the linear function $f(x) = p_0 + p_1x$. The resulting fit parameters are shown in the box on the top right.

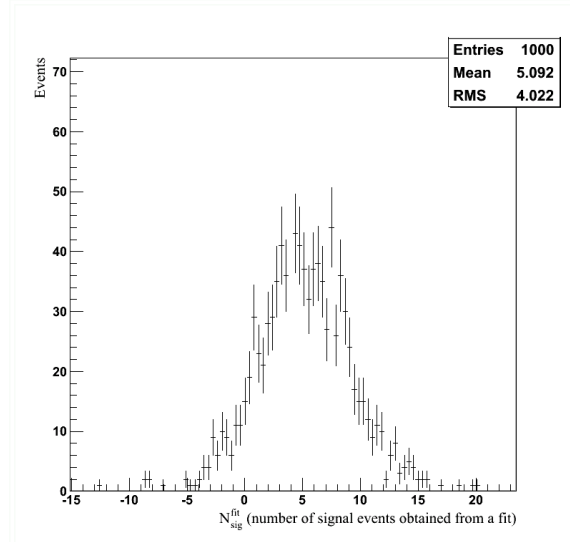
514 We plot the average number of signal events from the fit against $N_{\text{sig}}^{\text{gen}}$ and perform a fit
515 to a linear function $f(x) = p_0 + p_1x$. This plot and the results of the linear fit are shown
516 in Fig. 23. Fig. 24 displays distributions of $N_{\text{sig}}^{\text{fit}}$ for certain values of $N_{\text{sig}}^{\text{gen}}$. When $N_{\text{sig}}^{\text{gen}}$ is
517 large, the distribution of $N_{\text{sig}}^{\text{fit}}$ is unbiased. However, for small $N_{\text{sig}}^{\text{gen}}$, we see an asymmetry
518 in the distribution of $N_{\text{sig}}^{\text{fit}}$, indicating some bias. This effect is often observed for small
519 statistics and is not unexpected.

520 8.4 Sensitivity Estimation

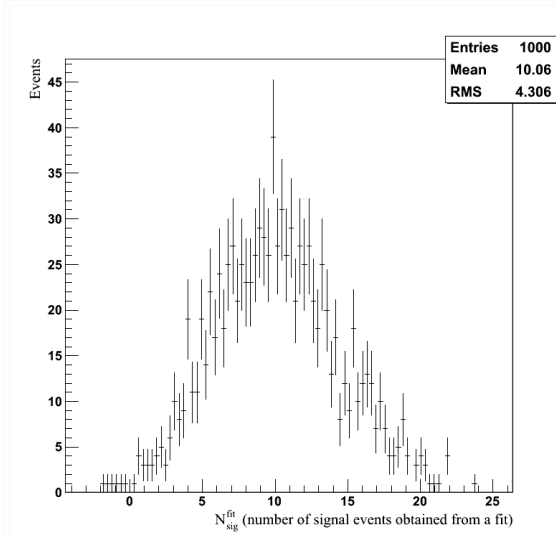
521 We estimate the upper limit on the branching fraction and visible cross section of $\Upsilon(5S) \rightarrow$
522 $\gamma W_{b,J}$ in the absence of signal by performing an extended unbinned maximum likelihood fit
523 on toy MC generated according to the fit to the data sidebands. We generate 1000 toy MC
524 samples with 59 background events, fit our combined signal and background shape to each
525 sample, and then average the resulting signal yields. There is an average signal yield of
526 -0.2 ± 3.2 events. Note that in Fig. 23, this average signal yield corresponds to the value
527 plotted at $N_{\text{sig}}^{\text{gen}} = 0$. Using the confidence belt in Fig. 22, we determine the 95% confidence
528 level upper limit on the number of signal events to be 10 events. We calculate the upper



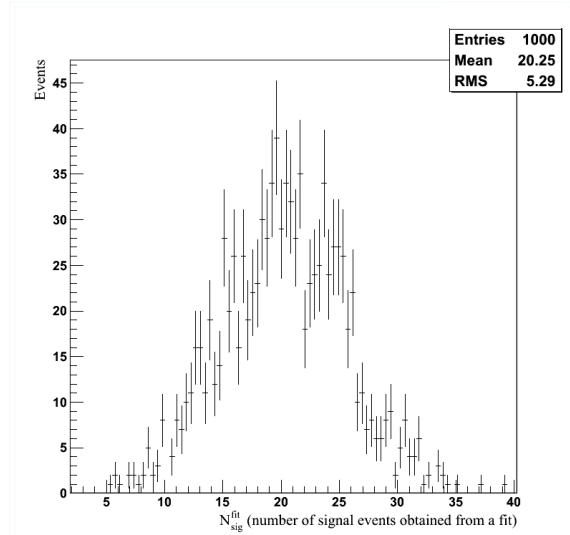
(a) Distribution of $N_{\text{sig}}^{\text{fit}}$ for an ensemble test with $N_{\text{sig}}^{\text{gen}} = 0$ and $N_{\text{bkg}}^{\text{gen}} = 59$.



(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 $N_{\text{sig}}^{\text{gen}}$.

529 limit on the branching fraction in the absence of signal as follows:

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

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

532 We calculate the visible cross section using

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

Table 11: Values of quantities used to calculate upper limits on visible cross section and the branching fraction. Uncertainty in $\mathcal{B}(\rho^0 \rightarrow \pi^+\pi^-)$ is negligible. Note that, for purposes of estimating upper limits, we use $N_{\text{sig}} = 10$, which is the 95% CL boundary of the 90% CL frequentist belt shown in Fig. 22 for $N_{\text{sig}}^{\text{fit}} = 3$, according to the result of the fit $N_{\text{sig}}^{\text{fit}} = -0.2 \pm 3.2$.

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

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

535 9 Search Strategy Summary

536 In this analysis, we describe a search for a new molecular state W_{bJ} which could be produced
537 in the radiative transition $\Upsilon(5S) \rightarrow \gamma W_{bJ}$ followed by the decays $W_{bJ} \rightarrow \Upsilon(1S)\rho^0$, $\Upsilon(1S) \rightarrow$
538 $\mu^+\mu^-$, $\rho^0 \rightarrow \pi^+\pi^-$. We fully reconstruct the signal final state consisting of two muons, two
539 pions, and a photon. We perform a blind analysis by optimizing our selection criteria and
540 analysis techniques using only MC samples before applying them to data.

541 To search for the presence of W_{bJ} in Belle data, we propose to "unblind" the data in
542 the signal region and then fit a one-dimensional distribution of $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ using
543 the aforementioned models for signal and background shapes. In the fit, we fix the width
544 of W_{bJ} to that of Z_b . Because we expect only one signal in our signal region, we plan to
545 scan the range of invariant masses of $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ and, for each assumed value of the
546 invariant mass, we perform a fit to data, where the background parameter λ is allowed to
547 float. If the fit returns a statistically significant result, we claim a discovery. We will then
548 produce a plot of the upper limit versus mass of W_{bJ} . This plot will be produced regardless
549 of whether or not the fit yields a significant result. Our confidence belt (Fig. 22) will be
550 used to either claim a discovery of W_{bJ} or establish an upper limit on the signal production
551 rate (branching fraction) for the radiative decay $\Upsilon(5S) \rightarrow \gamma W_{bJ}$. The following sources of
552 systematic uncertainties will be considered in our final estimate of the upper limit of the
553 branching fraction of $\Upsilon(5S) \rightarrow \gamma W_{bJ}$:

- 554 • Number of $\Upsilon(5S)$

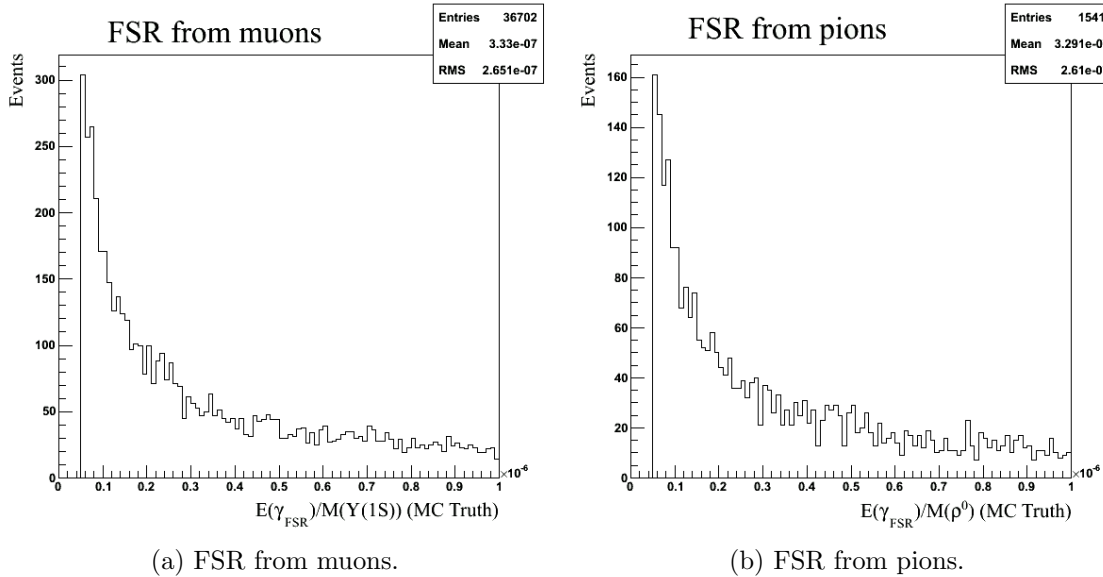


Figure 25: Final state radiation from charged tracks

- 555 • Signal Reconstruction Efficiency
- 556 • Daughter Branching Fractions
- 557 • MC statistics
- 558 • PDF parameterization
- 559 • Fit bias
- 560 • Trigger efficiency

561 10 Appendix

562 10.1 Final State Radiation

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

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

565 where XPHCUT is a hardcoded constant set to 0.01. Hence, the minimum FSR energy is
 566 approximately 4 MeV for pions ($M(\rho^0) = 770$ MeV) and 50 MeV for muons ($M(\Upsilon(1S)) =$
 567 9.46 GeV). The lower limit on FSR energy for muons is too high, so we lowered the
 568 value of XPHCUT to 10^{-7} . To accomplish this, we changed XPHCUT=0.01D0 to XPH-
 569 CUT=0.0000001D0, recompiled the phocin.F source code and then rebuilt EvtGen with an
 570 updated PHOTOS library.

571 To verify that XPHCUT was successfully lowered to 10^{-7} , we plot the ratios $\frac{E(\gamma_{FSR})}{M(\Upsilon(1S))}$
 572 and $\frac{E(\gamma_{FSR}^{\rho})}{M(\rho)}$ as generated in Fig. 25. Because these quantities are bounded from below by
 573 $XPHCUT \cdot 0.5$, we prove that XPHCUT was successfully lowered.

574 10.2 Changes in the Analysis between Note v1.5 and v2.0

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

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

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

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

600 • Signal Photon Energy

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

607 • Blinded Region in Data

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

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

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.

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

619 • **Impact Parameters for Charged Tracks**

620 The dz and dr selection criteria have been loosened to ≤ 4 cm and ≤ 0.5 cm, respec-
621 tively.

622 • **PID Requirements for Charged Pion Candidates**

623 These requirements have been eliminated.

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

626 **10.2.1 Signal Photon Energy Conundrum**

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

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

639 For as long as no bias is present in selecting signal event candidates, background events
640 of ISR origin are relatively well described (as you will see for yourself very soon) by the sum

641 of an exponential and a straight line of non-negative slope. Small non-peaking background is
 642 likely to be sufficiently-well approximated also by the same straight line (of zero or positive
 643 slope). However, our original selection criteria strongly suppressed ISR background at large
 644 values of the signal invariant mass $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$, making it very difficult to reliably
 645 obtain the shape of such biased background distribution using sidebands in data. We realized
 646 that, in order to significantly reduce such possible bias, we have to avoid suppressing ISR
 647 background in the sideband region.

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

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

671 In order to avoid the described bias in background photon energy spectrum and to be
 672 able to use the higher-mass sideband to perform a robust fit to such background, we release
 673 the cut on signal photon energy to the lowest possible value of 20 MeV (standard Belle
 674 reconstruction and MDST production do not go lower than that). Note that in our analysis
 675 we do not really care about possible energy dependence of photon reconstruction efficiency
 676 systematic uncertainty, because our signal is associated with photons of higher energy, but
 677 we need (even if smoothly suppressed) an exponential-like energy distribution of background
 678 photons to make extracting the signal from data reliable. In our approach we obtain the
 679 shape of background distribution from data also.

680 Interestingly, to improve our understanding of backgrounds, we also have to raise the
 681 cut on the other end of the photon energy region, though, in this case, for a different
 682 reason. As is explained in Table 6 and Fig. 10 of the original version of the note, there is
 683 a particular peaking background, namely, radiative (*i.e.* ISR) production of $\Upsilon(3S)$ followed
 684 by its dipion transition to $\Upsilon(1S)$, which is uncomfortably too close to the left side of our
 685 signal plus sideband region of $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$. When both charged pions and the photon

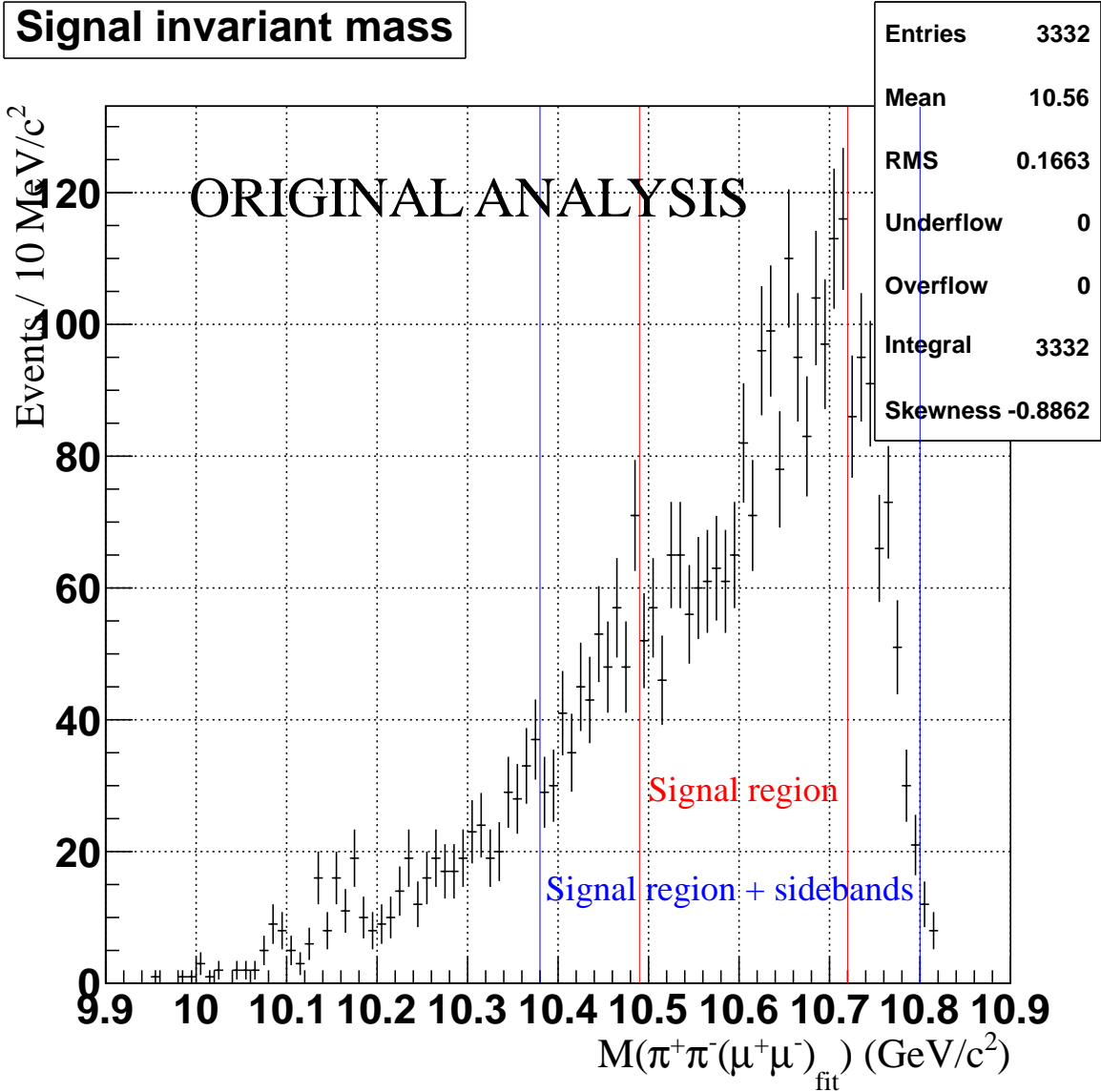


Figure 26: The reconstructed invariant mass $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ for a special ISR MC sample with the original selection criteria. Only best candidates are shown.

686 are misreconstructed, sometimes this unfortunate happenstance might shift some of such
 687 these events into the signal region. Therefore, it would be best to investigate this possible
 688 background using the data. This goal requires us to release the signal photon energy cut.
 689 For practical purposes, in the improved version of the analysis, we limit signal photon energy
 690 to 5 GeV. Note that our approach also facilitates possible measurement of ISR production
 691 of $\Upsilon(3S)$ and $\Upsilon(2S)$, which could be used to calibrate ISR MC.

692 After widening the signal photon candidate energy selection as described and explained,
 693 we plot the distributions of photon energy spectrum vs $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ in Fig. 28 and
 694 $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ in Fig. 29 (as always, for best candidates only) for MC events from our
 695 “ridiculous” background MC sample. We conclude that the current (*i.e.* new, relaxed) se-
 696 lection criteria allow us to perform a robust fit to the background using the sidebands of

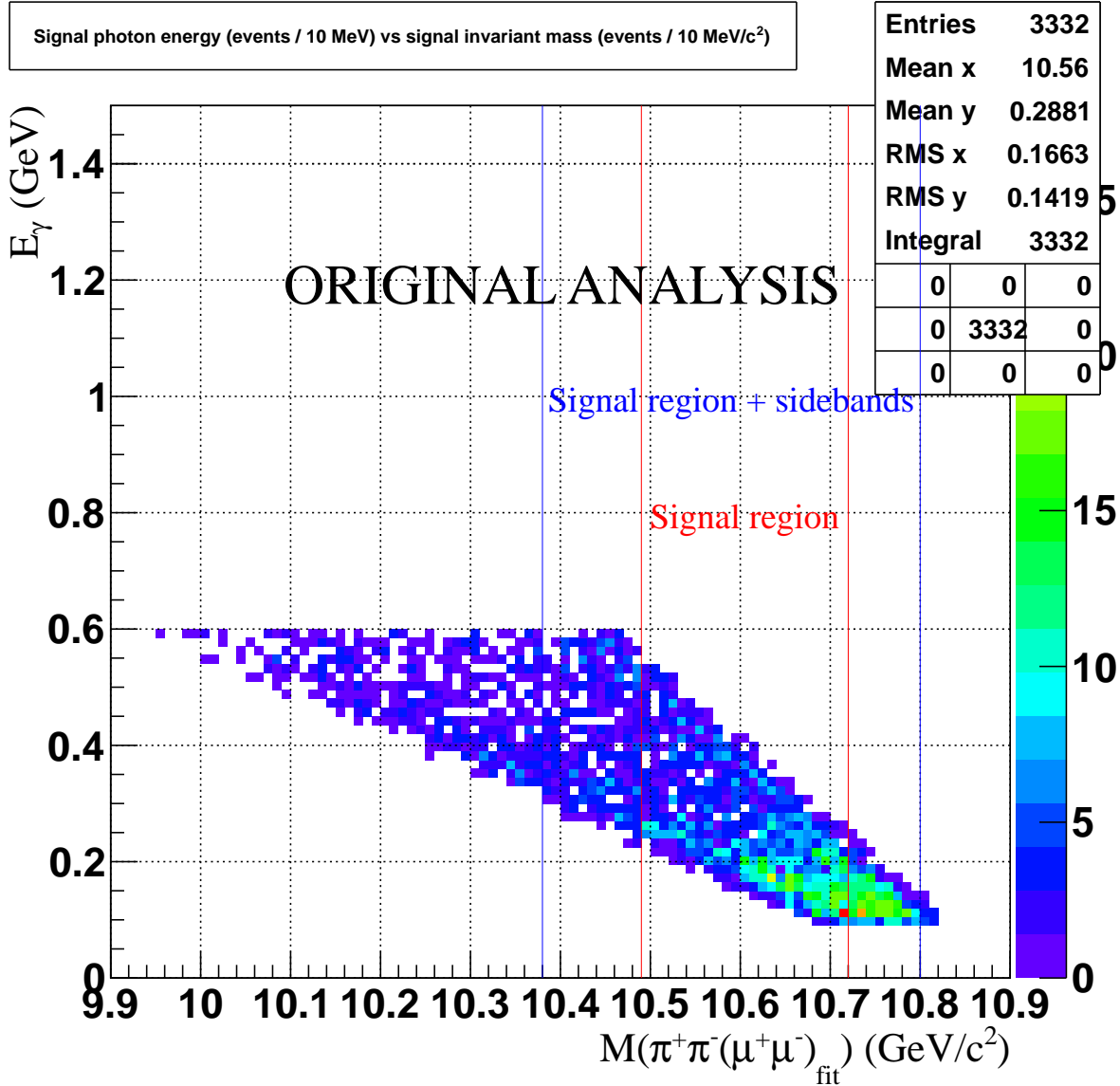


Figure 27: The reconstructed signal photon energy versus the invariant mass $wbjm$ for a special ISR MC sample with the original selection criteria. Only best candidates are shown.

697 $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ for the values of this variable up to 10.78 GeV/c². To further investi-
 698 gate the shape of $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ in this special MC sample, we plot the distribution of
 699 $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ using the logarithmic scale in Fig. 30. We observe that the distribution
 700 shown in the figure does not follow a simple exponential dependence on $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$
 701 in part because, as explained previously, the ISR spectrum in this MC sample is completely
 702 unreasonable.

703 We reskim the data and generic background MC, blind the signal region and present
 704 the data distributions of photon energy spectrum vs $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ in Fig. 31 and
 705 $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ in Fig. 32. We observe unambiguous signatures of $\Upsilon(2S)$ and $\Upsilon(3S)$
 706 ISR production. We also show the distribution of $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ for blinded data in
 707 Figures 33 and 34 using the logarithmic scale. Note that the last figure is plotted using

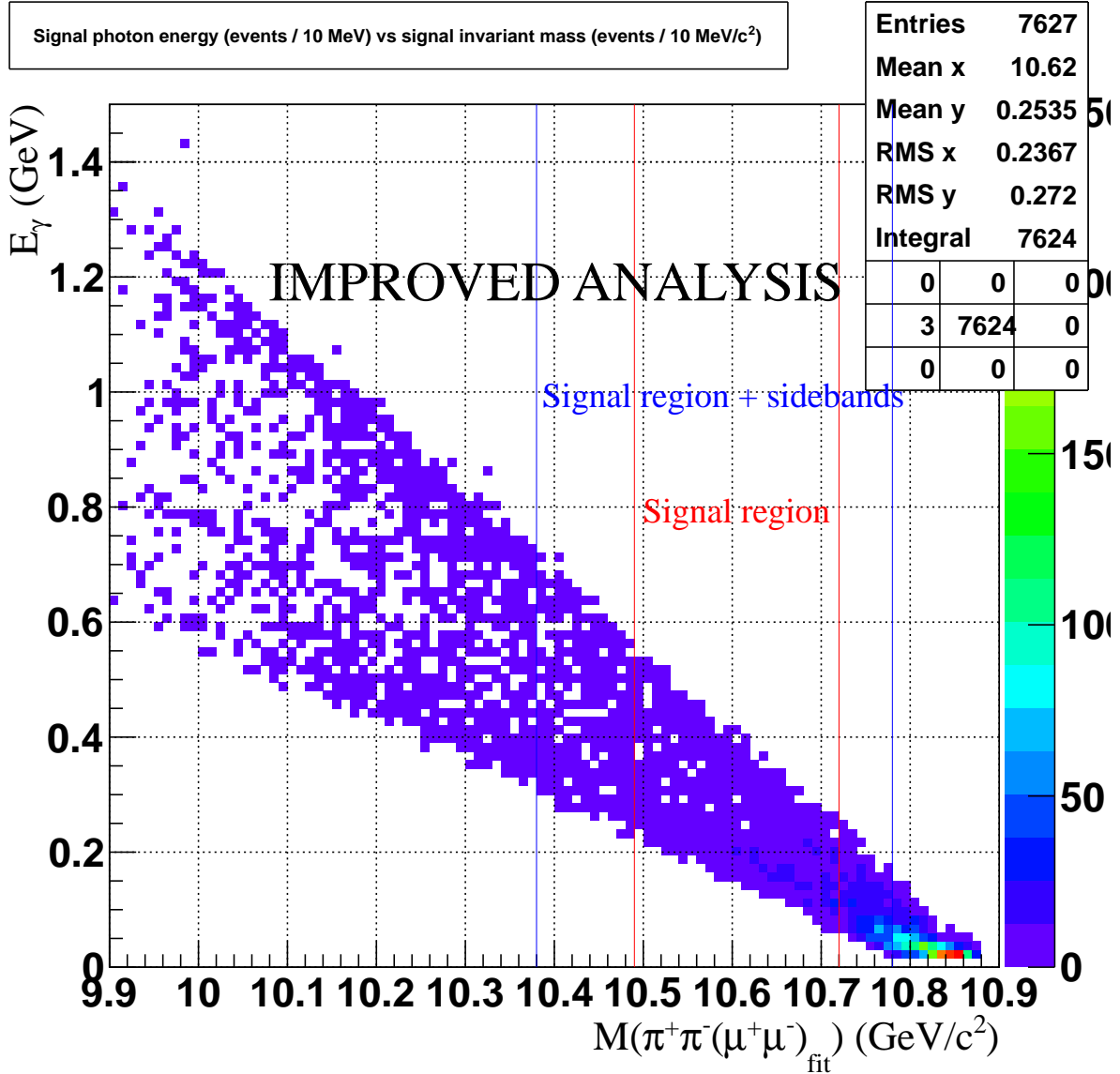


Figure 28: The reconstructed signal photon energy versus the invariant mass $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ for a special ISR MC sample with relaxed selection criteria. Only best candidates are shown.

708 a finer bin width of $2 \text{ MeV}/c^2$ in a narrower range of the invariant mass. Using the data,
 709 we perform a rough estimate of the width of the peak seen at the nominal mass of $\Upsilon(3S)$,
 710 $10.355 \text{ GeV}/c^2$, corresponding to events originating from radiative return to $\Upsilon(3S)$. Using
 711 our estimate of $5 \text{ MeV}/c^2$ (consistent with our MC-based understanding of resolution), we
 712 conclude that events in the left sideband of Fig. 34 are at least 5 width units away from
 713 this peak. Hence, it is unlikely that events in our left sideband are from radiative return to
 714 $\Upsilon(3S)$.

715 In the next section we explain our fitting strategy for extracting the signal from data
 716 (when the permission to unblind is granted). We fit the data in the range of signal invariant
 717 mass $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ between the two blue vertical lines shown in Figures 27–34, *i.e.*

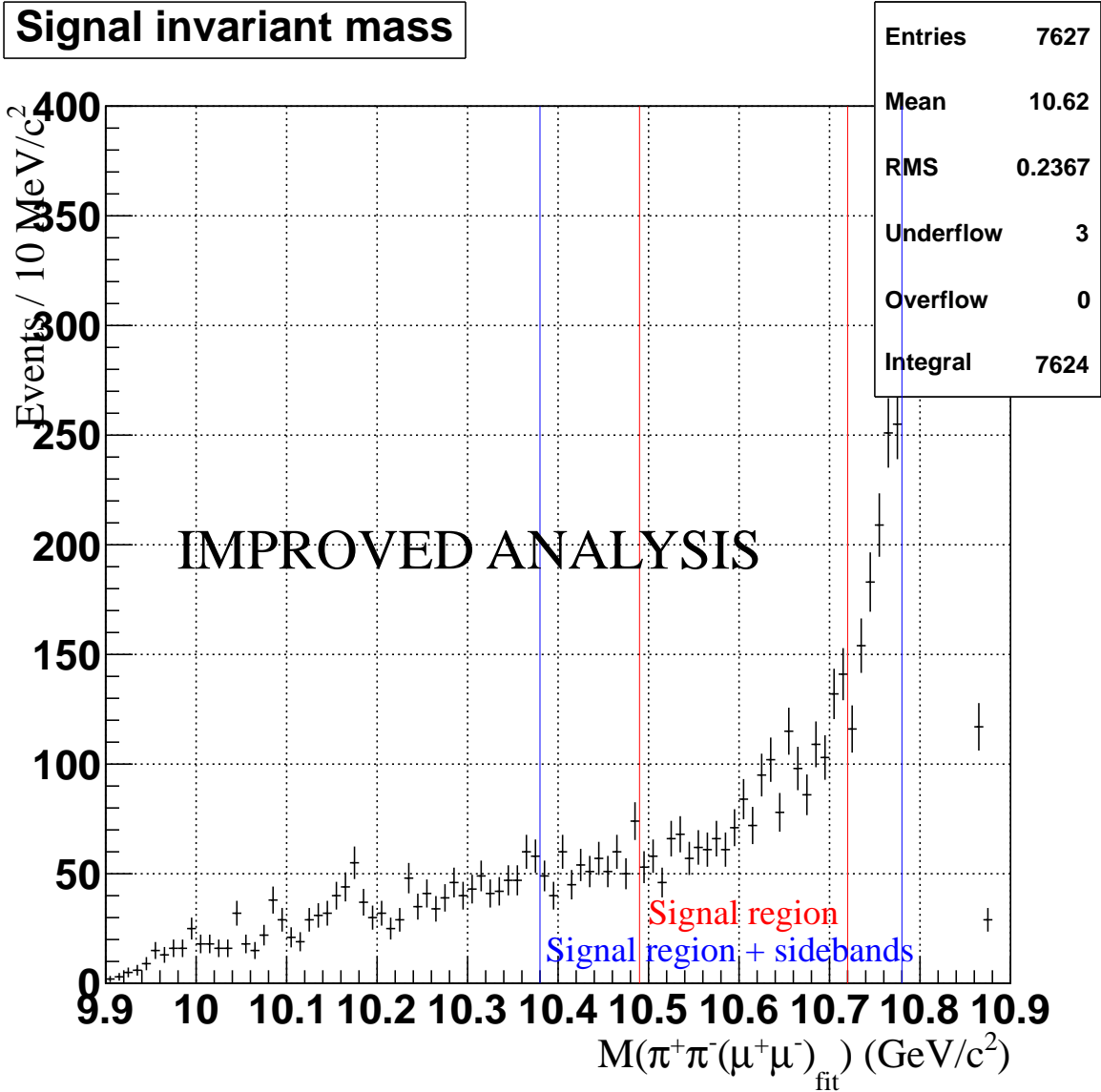


Figure 29: The reconstructed invariant mass $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ for a special ISR MC sample with relaxed selection criteria. Only best candidates are shown.

718 in the range $10.38 \text{ GeV}/c^2 \leq M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}}) \leq 10.78 \text{ GeV}/c^2$. This region of the
 719 invariant mass in blinded data is shown in Fig. 35.

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

724 10.3 Fitting Strategy

725 To extract the $W_{b,J}$ signal from the data or to estimate the upper limit on its production, we
 726 (plan to) fit the invariant mass distribution $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ in data with the sum of an

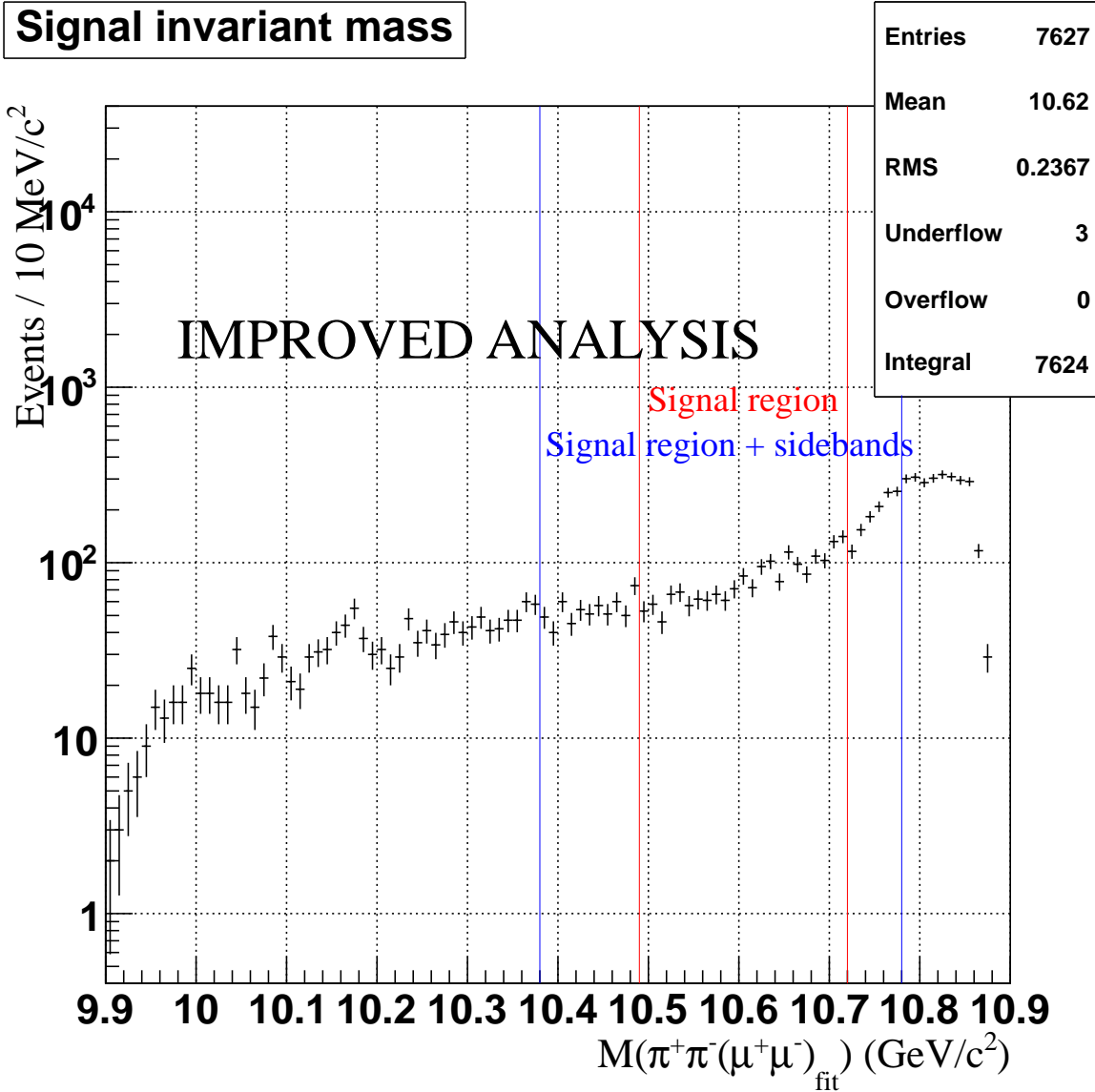


Figure 30: The reconstructed invariant mass $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ for a special ISR MC sample with relaxed selection criteria shown using the logarithmic scale. The ISR spectrum in this special MC sample has an unreasonable shape and could not be described by an exponential. Only best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

727 exponential, a straight line of zero or positive slope and the model for the signal shown in
 728 Fig. 21a. We plan to perform such fits for the values of (fixed) nominal mass of W_{bJ} between
 729 10.5 and 10.7 GeV/c² in steps of a few MeV/c². Important reference points here are provided
 730 by the invariant masses of Z_B and Z'_b which are, respectively, 10.610 and 10.650 GeV/c². We
 731 expect (or, rather, M. Voloshin expects) W_{bJ} to be roughly as wide (or narrow) as $Z_b^{(l)}$. This
 732 makes our life easier. Each fit will be performed independently. The shape of background
 733 distribution will be obtained from the data including the signal region. In our opinion, we
 734 can not obtain the shape of background exclusively from the sidebands because our sidebands

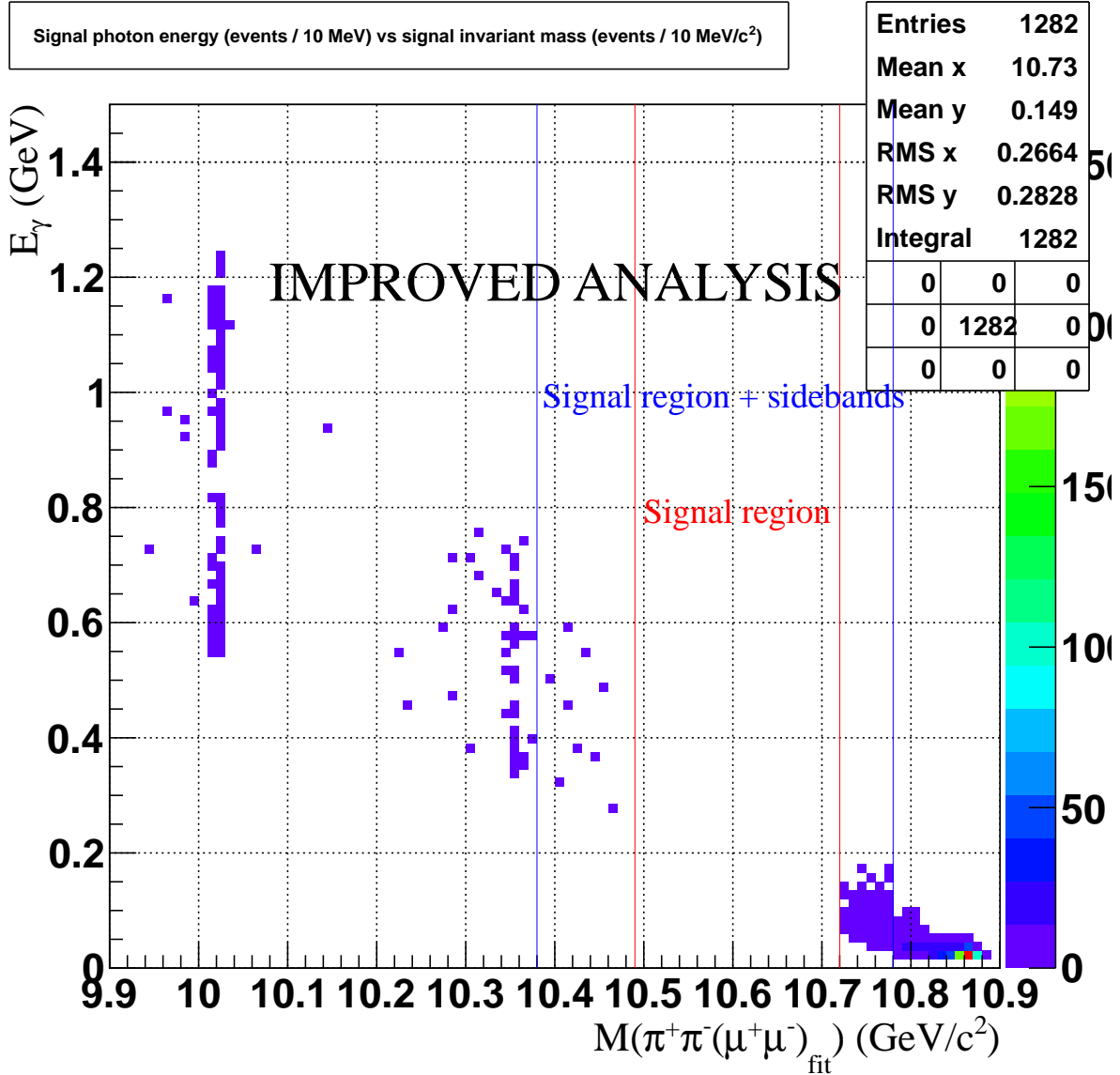


Figure 31: The reconstructed signal photon energy versus the invariant mass $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ for blinded data with relaxed selection criteria. Only best candidates are shown.

735 are relatively narrow and, also, if the shape of the background function is fixed using our
 736 sidebands, fitting with the model described in this section could easily introduce a significant
 737 bias in the results of the fit. Another limitation comes from the wide range of the invariant
 738 mass region where we are searching for the W_{bJ} . For each individual fit (with a particular
 739 hypothesis for W_{bJ} mass), the effective sideband region is going to be significantly wider than
 740 in our exercises discussed in this section. The key assumptions are: 1) there are no peaking
 741 backgrounds in the entire signal region, and 2) backgrounds can be modeled by the sum of
 742 an exponential and a straight line. Our confidence is based on MC studies using, first of all,
 743 our ISR MC samples.

744 We fit the invariant mass $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ in the range between $10.38 \text{ GeV}/c^2 \leq$

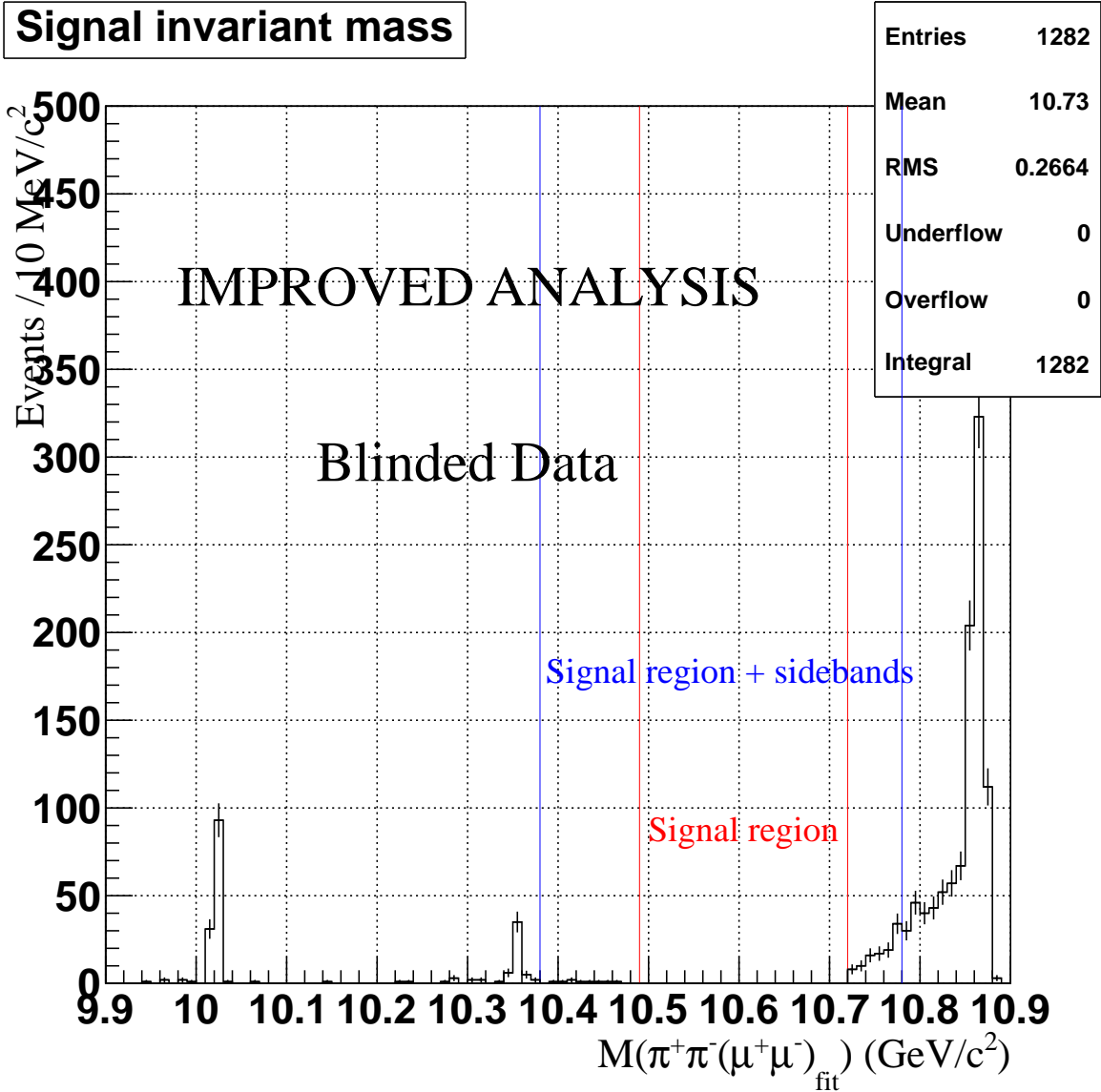


Figure 32: The reconstructed invariant mass $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ for blinded data with relaxed selection criteria. Only best candidates are shown.

745 $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}}) \leq 10.78 \text{ GeV}/c^2$. In principle, we can (significantly) extend the invariant
 746 mass included in the fit toward smaller values (therefore including the radiative production
 747 of $\Upsilon(3S)$ or even $\Upsilon(2S)$ in our fits), however, it is not clear to us if this would necessarily
 748 help us understand the shape of the background and to reduce the uncertainty in our model
 749 description of the data in the signal region.

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

753 We start our adventure by fitting the distribution of properly reweighted ISR MC sample
 754 for the decay $\Upsilon(5S) \rightarrow \Upsilon(1S)\pi^+\pi^-$ shown in Fig. 39. We fit this distribution using the
 755 unbinned extended maximum likelihood technique implemented in RooFit with the sum of

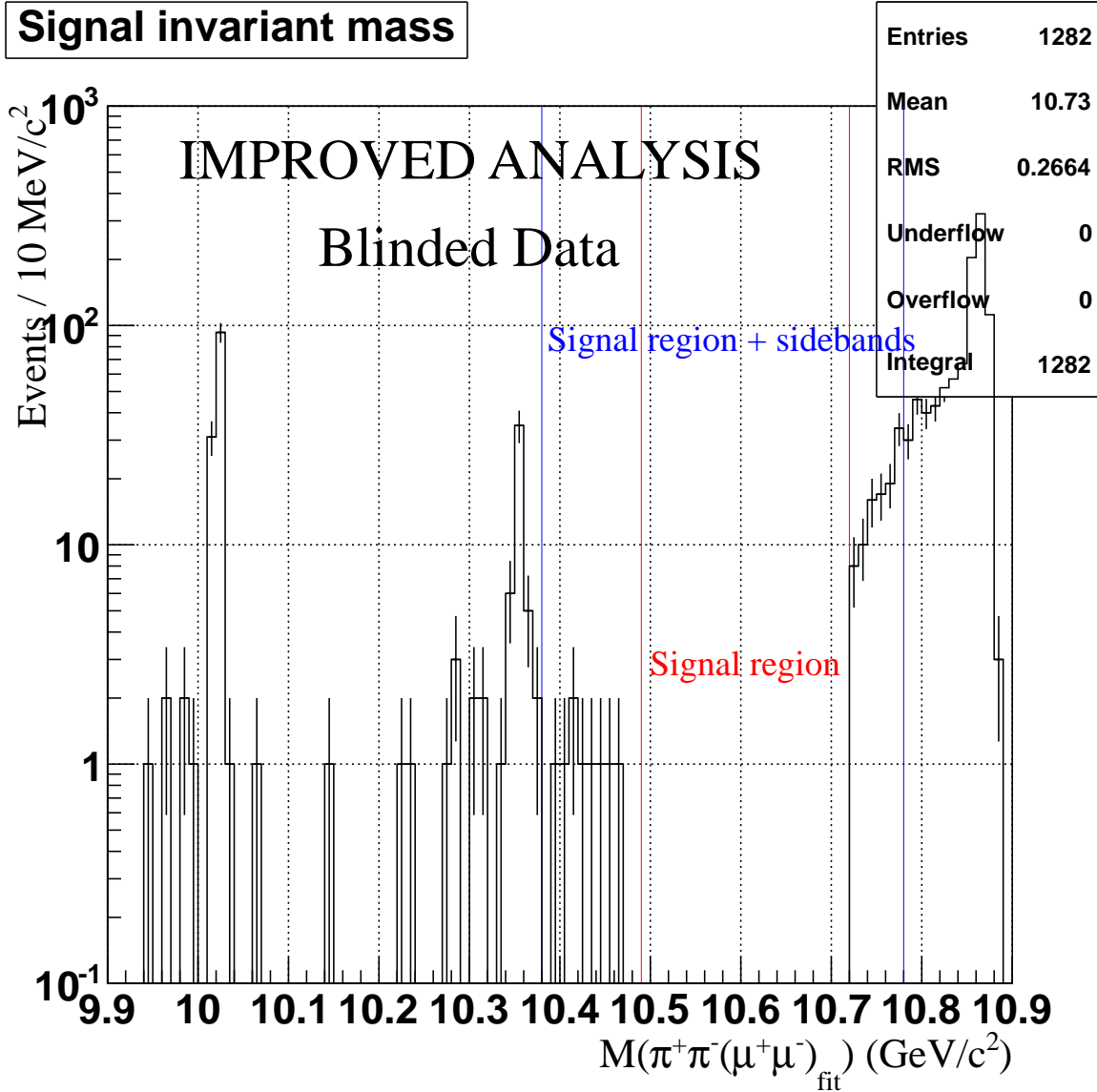


Figure 33: The reconstructed invariant mass $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ for blinded data with relaxed selection criteria plotted using the logarithmic scale. This is the same distribution as shown in Fig. 32. Only best candidates are shown.

756 an exponential and a straight line of zero or positive slope (if the curious reader really wants
 757 to know, we use `RooChebychev` for the latter). The results of this fit are shown in Fig. 40.
 758 We show the results of the fit using both the linear and the logarithmic scales because while
 759 one PDF is linear when plotted on log scale, the other PDF is, surprise, linear when plotted
 760 on linear scale because it is a line! Note that the fit has four parameters: α is the parameter
 761 of the exponent, `slope` is the slope of the straight line, $N1$ and $N2$ are the numbers of
 762 events obtained from the fit for background contributions parameterized by the exponent
 763 and the straight line, respectively. Note that we can not replace these two parameters by
 764 a single fraction parameter in an **extended** ML fit. When we fit real data, we plan to let
 765 the relative contributions from the two PDFs to be independently varying parameters in

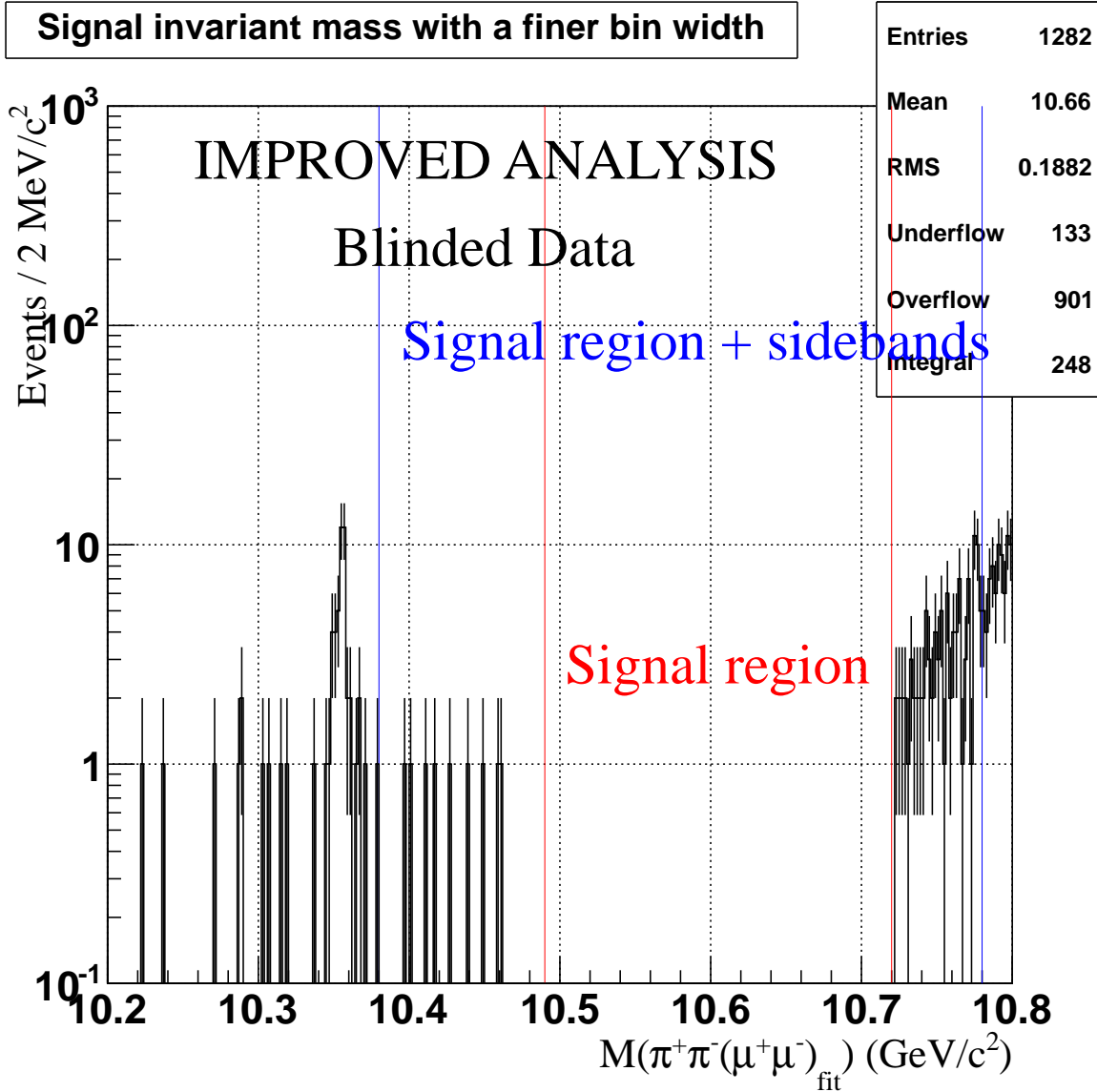


Figure 34: The reconstructed invariant mass $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ for blinded data with relaxed selection criteria plotted using the logarithmic scale for a narrower region of the invariant mass using a finer bin width than used for plots shown in Figures 32 and 33. Only best candidates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

766 the fit, same way it is the case in fits described here. This degree of freedom could be used
 767 to approximate (a small contribution, as we conclude from studying the sidebands) from
 768 non-peaking background possibly present in data in the signal region.

769 In the next step we exclude events in the signal region from the fit and repeat the
 770 described exercise for the same ISR MC sample. The results are shown in Fig. 41. One can
 771 easily notice that our sidebands are not sufficiently wide to use these to obtain the shape of
 772 background in the **entire** signal region. This is the reason why, when fitting the data, these
 773 two PDFs, an exponential and a straight line, will be combined with a signal PDF. In either
 774 case, one can see that presence of ISR definitely introduces a large systematic uncertainty in

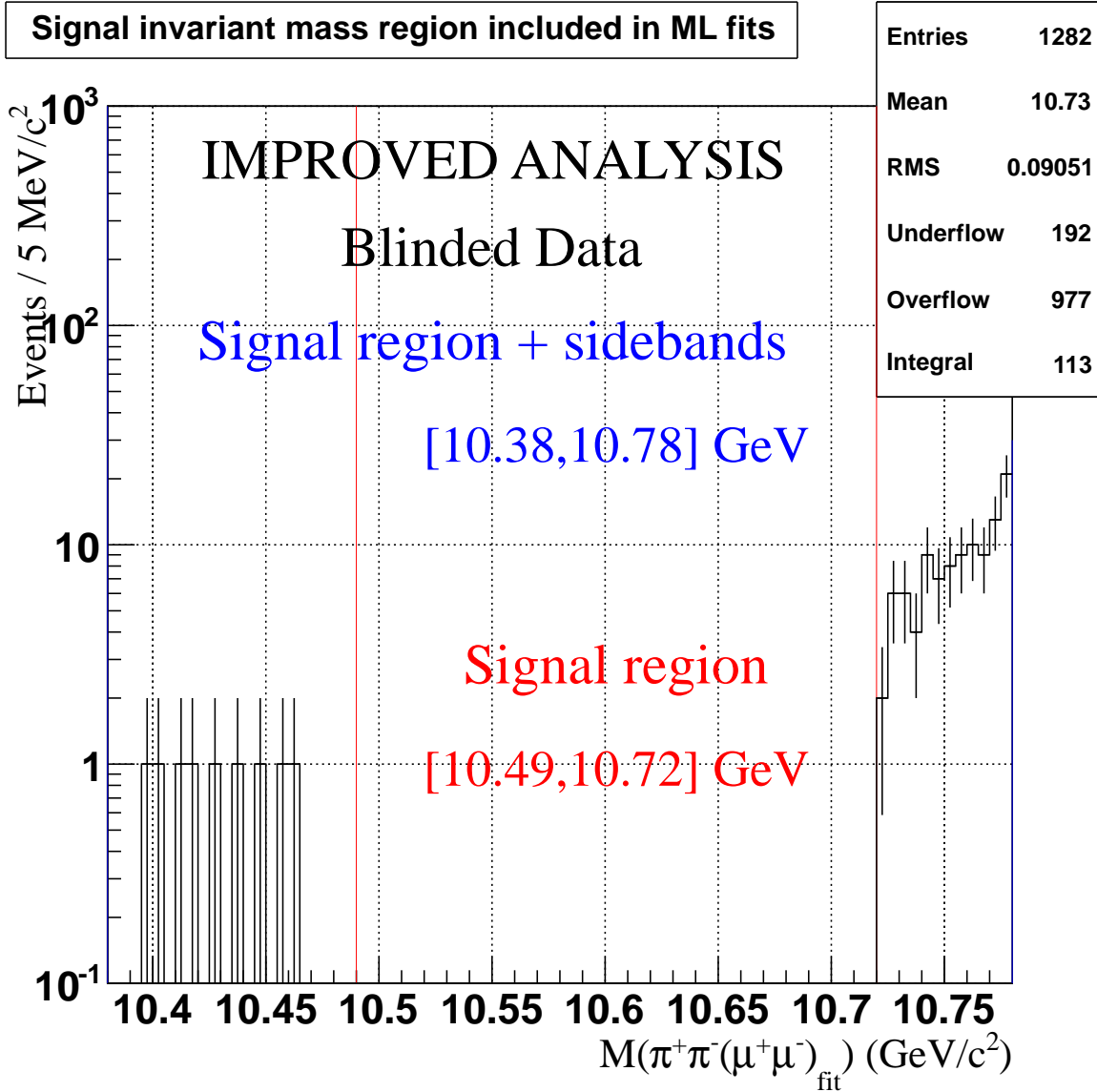


Figure 35: The reconstructed invariant mass $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ for blinded data with relaxed selection criteria plotted using the logarithmic scale for the region of the invariant mass included in the fits. Same events are shown as in Figures 32–34 but using a different bin width.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

775 our results for (relatively) larger invariant masses of W_{bJ} , in the region where the exponential
 776 contribution is rapidly increasing.

777 Now we try to fit the blinded data (just to see if the fit is going to converge at all).
 778 The results of the fit are shown in Fig. 42. Again, we observe that it would be unrealistic
 779 to expect our sidebands to predict the background in the entire, 400 MeV/c²-wide signal
 780 region.

781 Now we try to fit our precious ISR MC sample using three PDFs: the exponential
 782 and the straight line approximating background and the signal line shape shown in Fig. 21a
 783 (convolution of the signal Breit-Wigner with two Gaussians prepared using FFT plug-in in

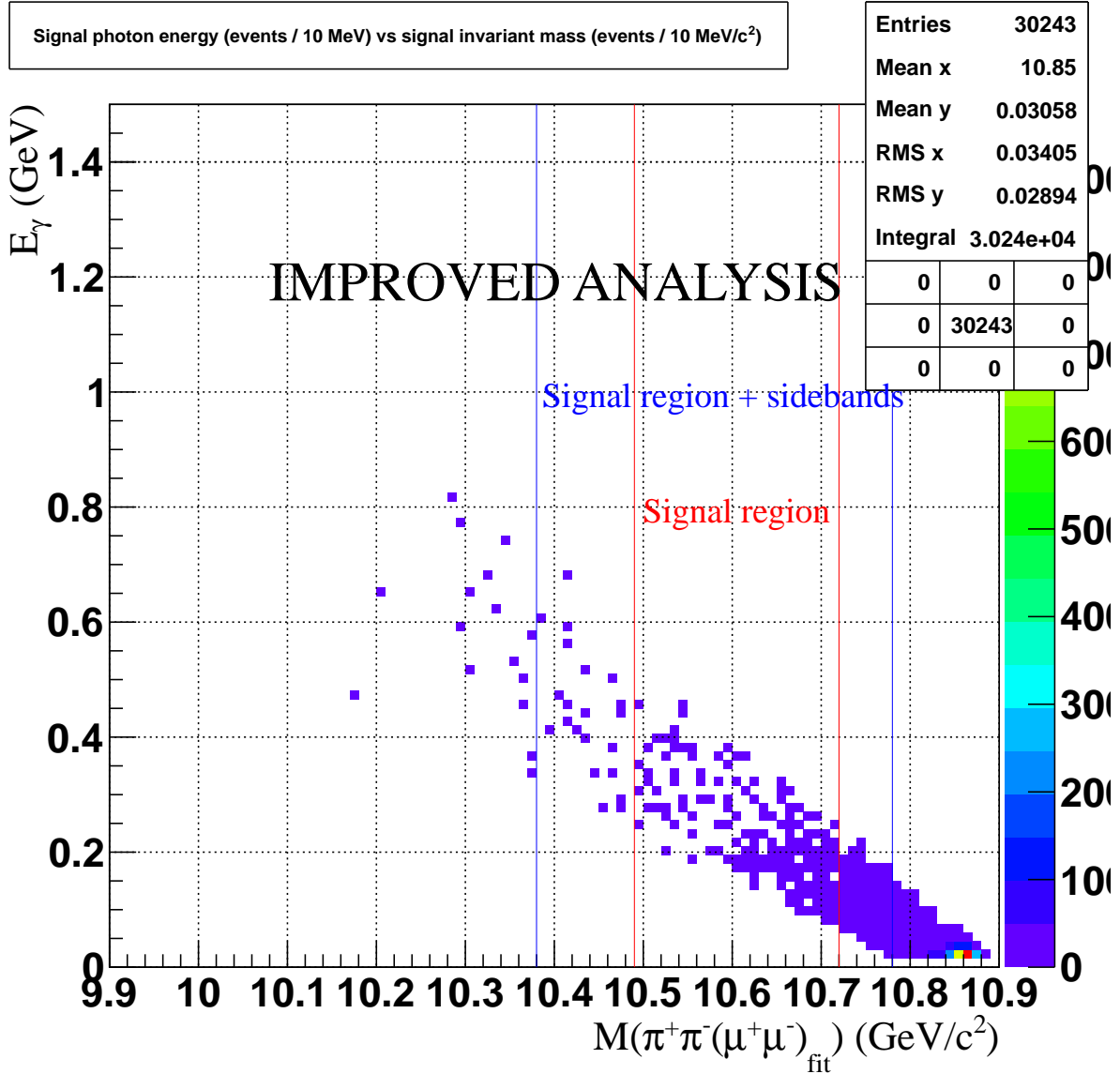


Figure 36: The reconstructed signal photon energy versus the invariant mass $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ for ISR MC sample described in section 6 with relaxed selection criteria. Only best candidates are shown.

784 ROOT). Note that no signal MC events have been added to the pure ISR MC sample yet.
 785 The number of signal events (NS) is an additional parameter in the fit, but the shape of
 786 the signal and its location (*i.e.* the invariant mass of W_{bJ} set at 10.6 GeV/c²) are fixed
 787 in this fit. As you can see in Fig. 43 the fit finds no statistically significant signal. Note
 788 that the solid green curve superimposed on the results of the fit shows how 50 signal events
 789 would look on average according to signal PDF description. The result of the fit for NS is
 790 a negative fluctuation.

791 Inspired by our success, we now ask the fitter to search for the signal (where there is
 792 none) in ISR MC sample. To do so we let the nominal mass of the W_{bJ} float in the range
 793 between 10.5 and 10.7 GeV/c². Note that in our future fits to data we plan to **scan** through

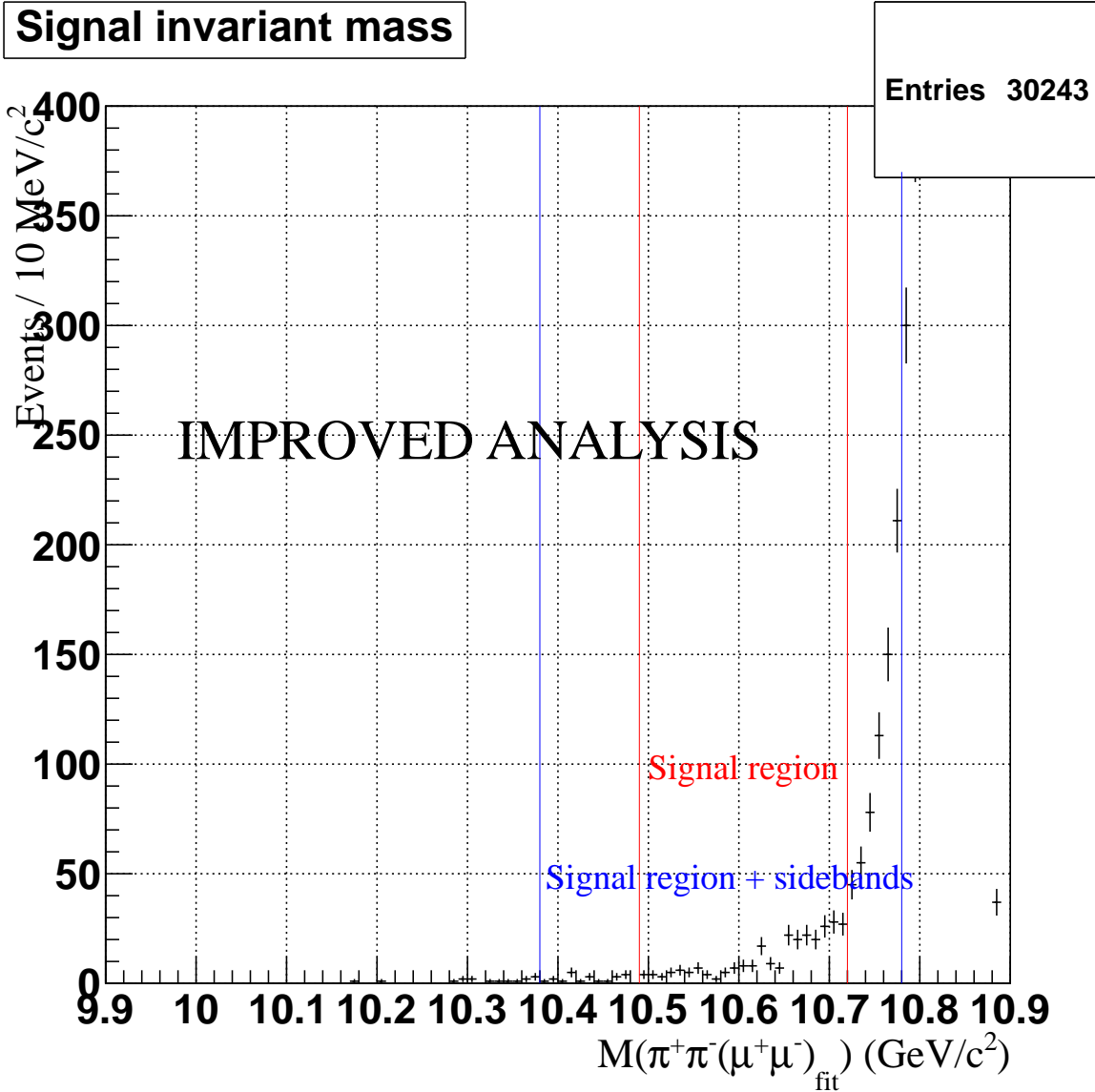


Figure 37: The reconstructed invariant mass $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ for ISR MC sample described in section 6 with relaxed selection criteria. Only best candidates are shown.

794 this interval of $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$, however, for the fit performed here we simply want to
 795 see the significance of the worst-case-scenario when the fit “discover” a signal where there
 796 is none. The results of this fit are shown in Fig. 44. Indeed, an obvious enhancement
 797 in the distribution of $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ is “picked-up” by the fitter as the most likely
 798 “signal”, however, as you can observe from the results of the fit, statistical significance of
 799 this “discovery” is consistent with a fluctuation. Such results are also likely to be obtained
 800 in the data, and, in case of low significance and no discovery, this would blow up the upper
 801 limit estimate.

802 Finally, being brave young pioneers, we decide to tackle a simulated data sample where
 803 50 events (with $W_{b,J}$ mass of $10.620 \text{ GeV}/c^2$) are randomly selected from one of our simulated
 804 signal MC sample and are added to the same ISR MC sample we are using for all our fits

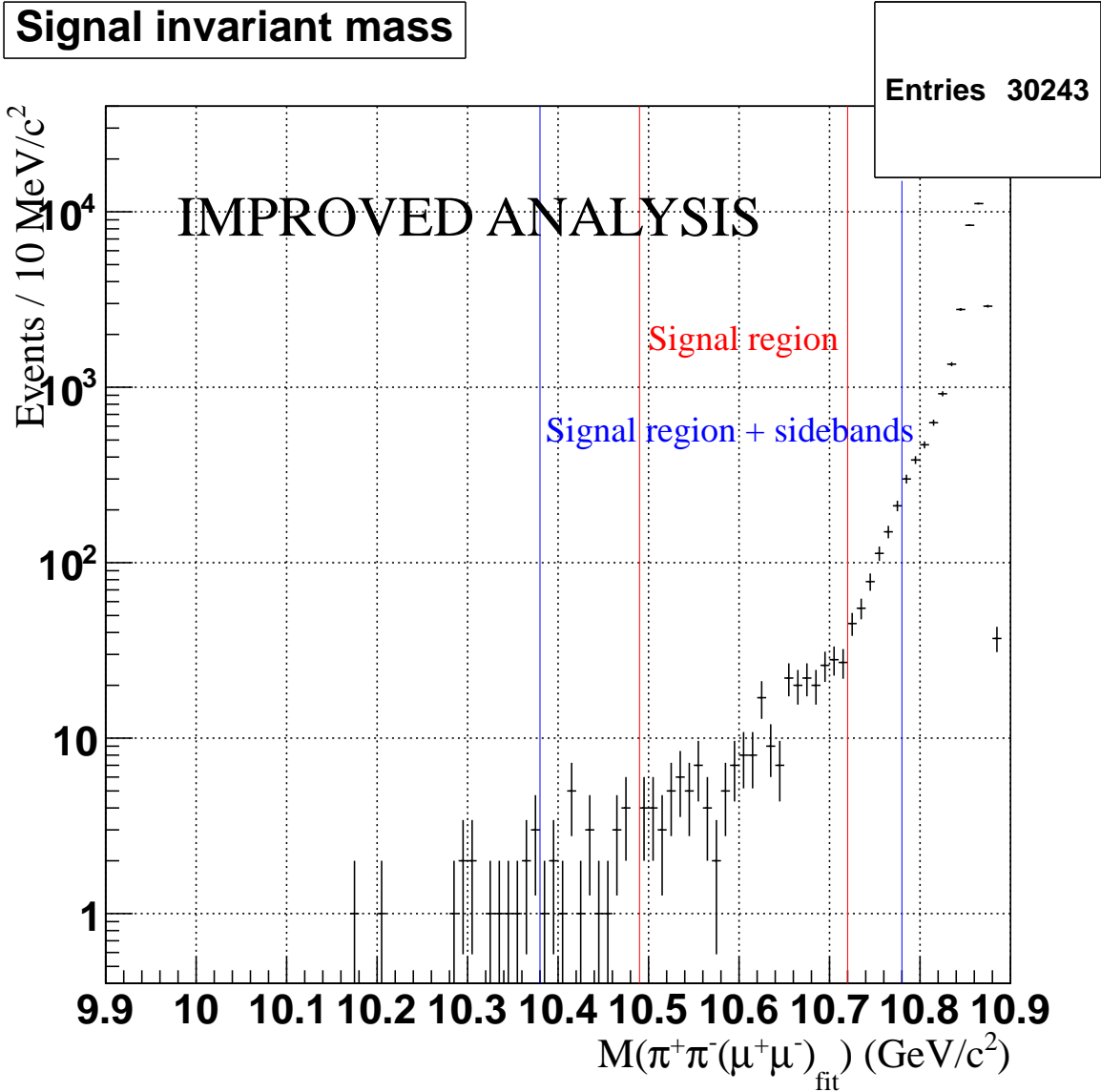


Figure 38: The reconstructed invariant mass $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ for ISR MC sample described in section 6 with relaxed selection criteria shown using the logarithmic scale. Only best candidates are shown.

805 described in this section. We let the fitter search for this signal and report the results in
 806 Fig. 45. The fitter finds the signal of the right magnitude. We conclude that our fitting
 807 procedure is working.

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

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

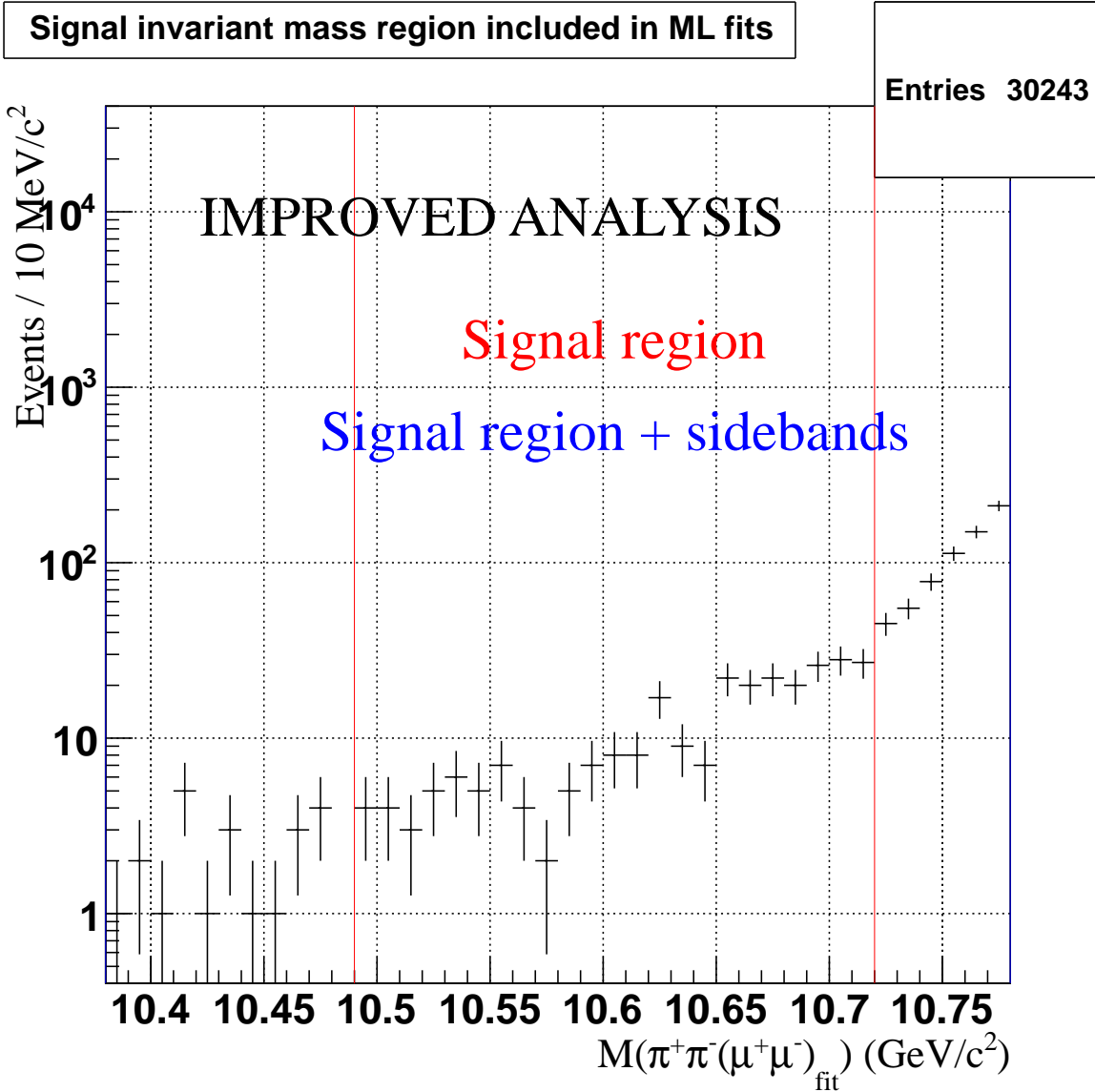


Figure 39: The reconstructed invariant mass $M(\pi^+\pi^-(\mu^+\mu^-)_{\text{fit}})$ for ISR MC sample described in section 6 with relaxed selection criteria shown using the logarithmic scale for the mass region used in the fits. Only best candidates are shown.

815 scan are shown in Fig. 47.

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

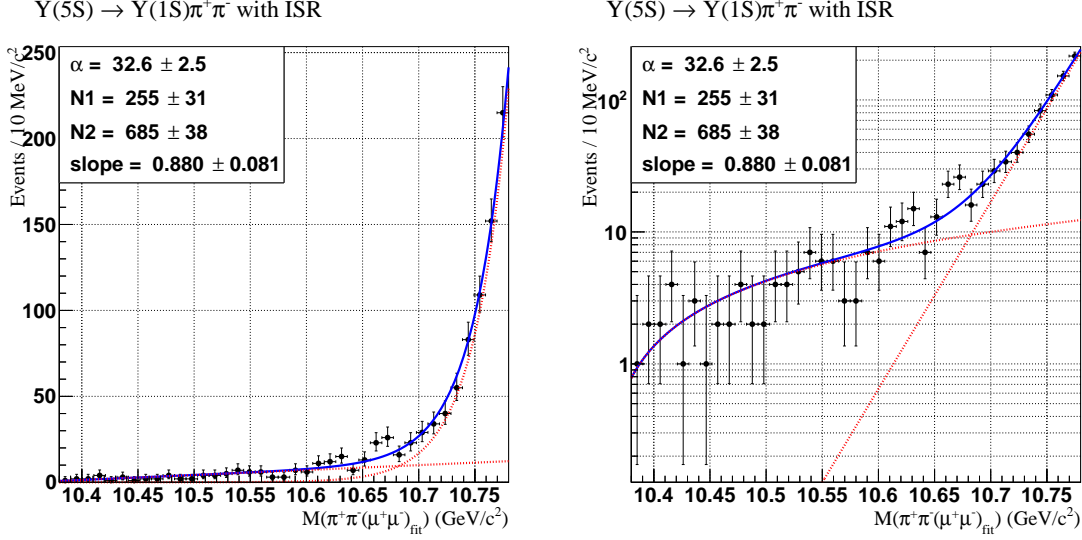


Figure 40: The results of the ML fit for ISR MC sample.

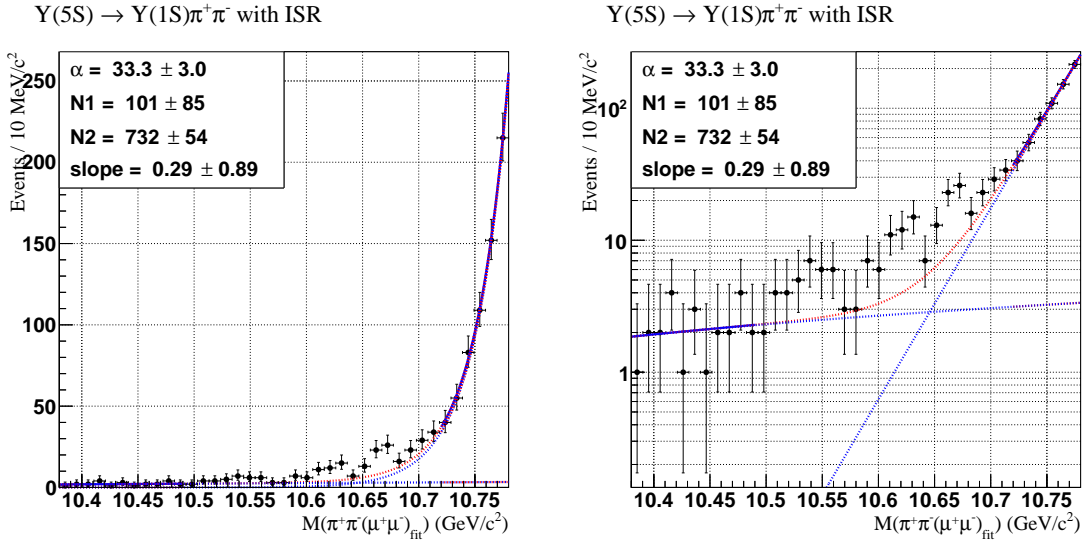


Figure 41: The results of the ML fit for ISR MC sample excluding the signal region.

825 Finally, to investigate the effect of possible presence of the signal W_{bJ} in data, we per-
 826 form a scan to ISR MC mixed with 50 signal MC events with nominal mass of W_{bJ} at
 827 $10.62 \text{ GeV}/c^2$. The results of the scan performed on this sample are shown in Fig. 50.

828 References

- 829 [1] A. Bondar et al. Observation of two charged bottomonium-like resonances in $Y(5S)$
 830 decays. *Phys. Rev. Lett.*, 108:122001, 2012.
- 831 [2] I. Adachi et al. Evidence for a $Z_b^0(10610)$ in Dalitz analysis of $Y(5S) \rightarrow Y(nS)\pi^0\pi^0$.
 832 2012.

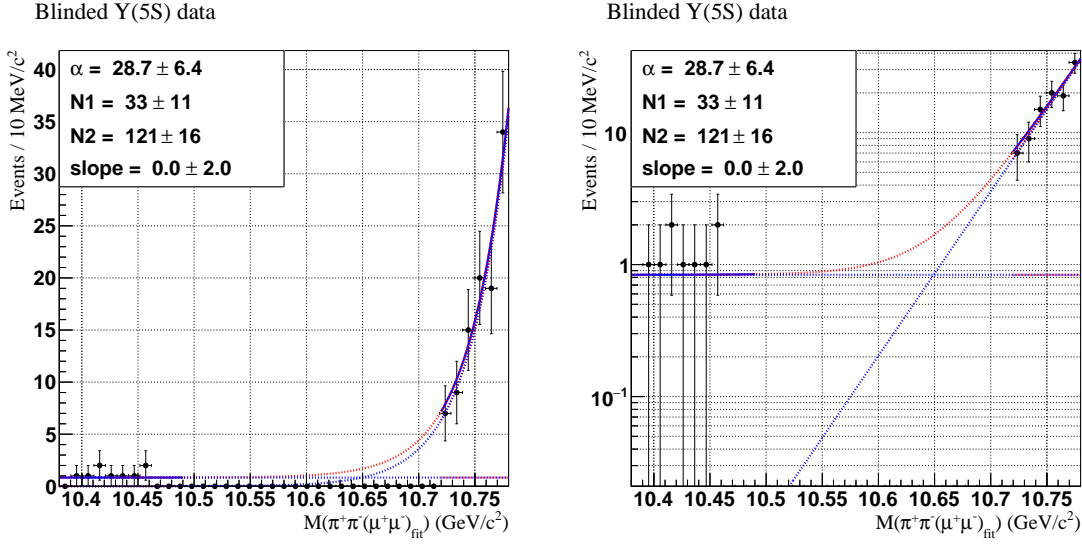


Figure 42: The results of the ML fit for sidebands of the blinded data sample.

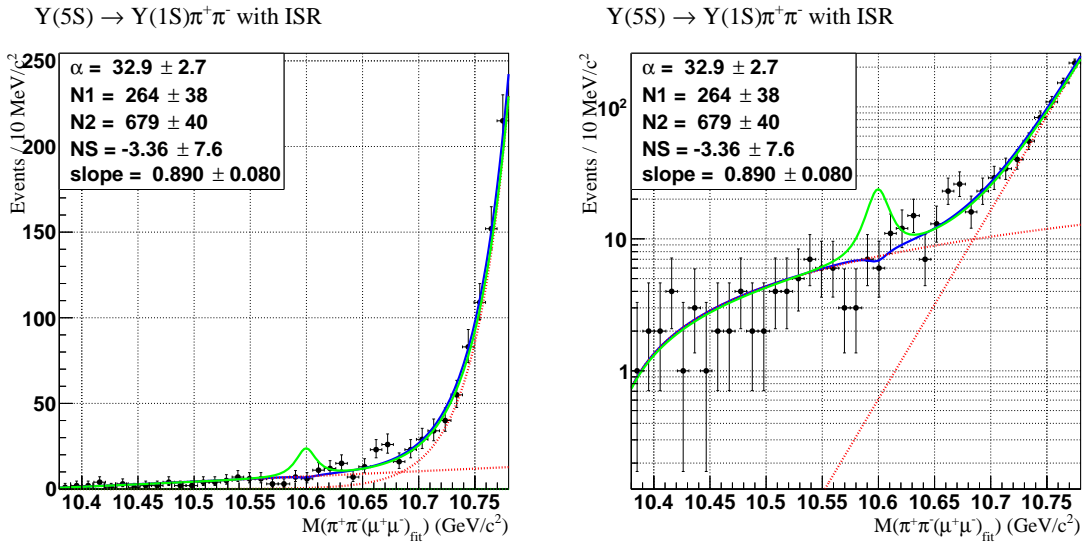


Figure 43: The results of the ML fit for ISR MC sample with background model and signal PDF shape.

- 833 [3] P. Krokovny et al. First observation of the $Z_b^0(10610)$ in a Dalitz analysis of $\Upsilon(10860)$
834 $\rightarrow \Upsilon(nS)\pi^0\pi^0$. *Phys. Rev.*, D88(5):052016, 2013.
- 835 [4] A. Garmash et al. Observation of $Z_b(10610)$ and $Z_b(10650)$ Decaying to B Mesons.
836 *Phys. Rev. Lett.*, 116(21):212001, 2016.
- 837 [5] L. M. Lederman. The discovery of the Upsilon, bottom quark, and B mesons. In
838 *The Rise of the standard model: Particle physics in the 1960s and 1970s. Proceedings,*
839 *Conference, Stanford, USA, June 24-27, 1992*, pages 101–113, 1992.
- 840 [6] M. Tanabashi et al. Review of Particle Physics. *Phys. Rev.*, D98(3):030001, 2018.

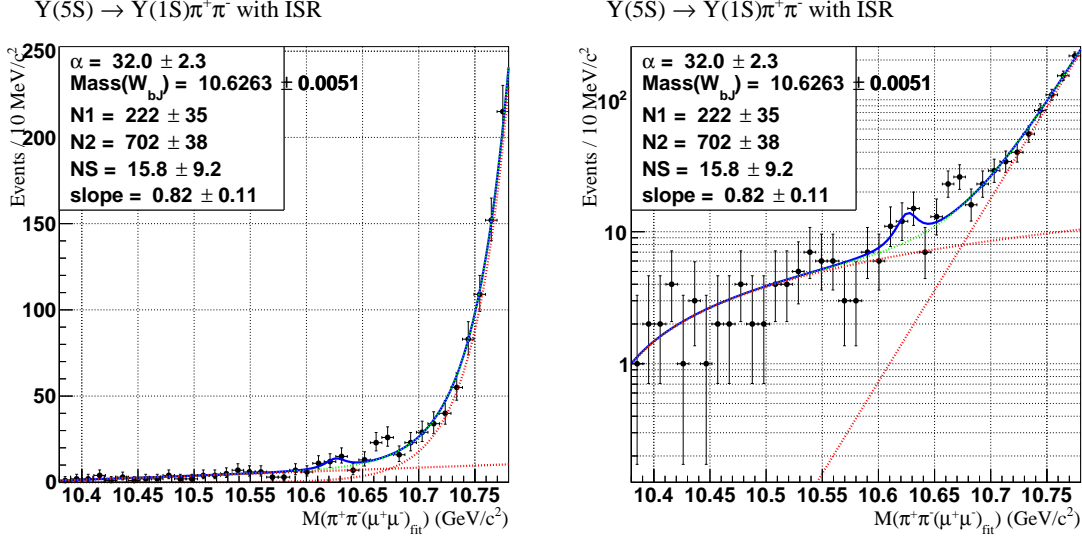


Figure 44: The results of the ML fit for ISR MC sample with background model and signal PDF shape (mass is a free parameter).

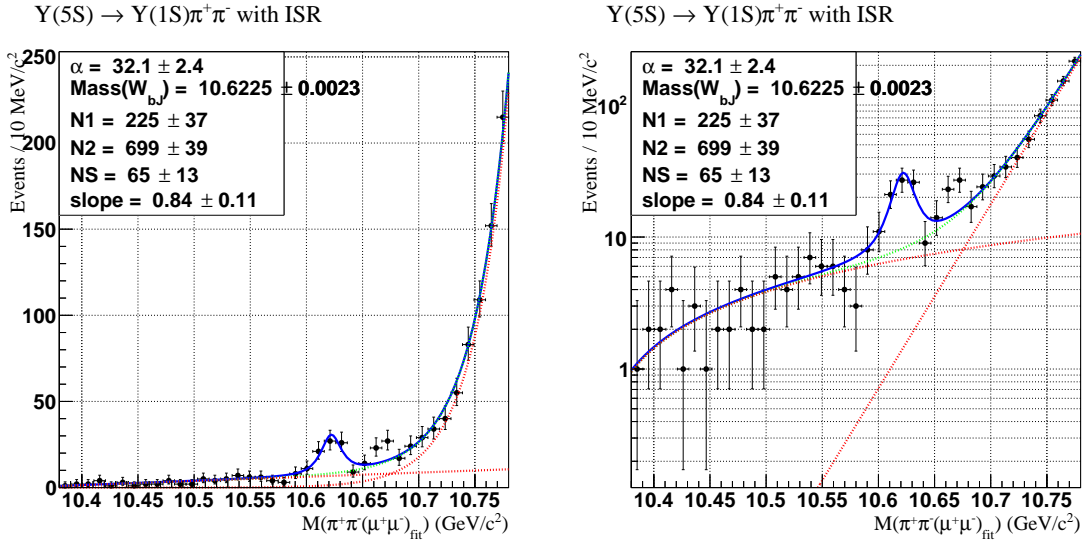


Figure 45: The results of the ML fit for ISR + signal (50 events) MC sample with background model and signal PDF shape (mass is a free parameter).

841 [7] S. Godfrey and N. Isgur. Mesons in a Relativized Quark Model with Chromodynamics.
 842 *Phys. Rev.*, D32:189–231, 1985.

843 [8] S. Godfrey and K. Moats. Bottomonium Mesons and Strategies for their Observation.
 844 *Phys. Rev.*, D92(5):054034, 2015.

845 [9] S. Godfrey, K. Moats, and E. S. Swanson. B and B_s Meson Spectroscopy. *Phys. Rev.*,
 846 D94(5):054025, 2016.

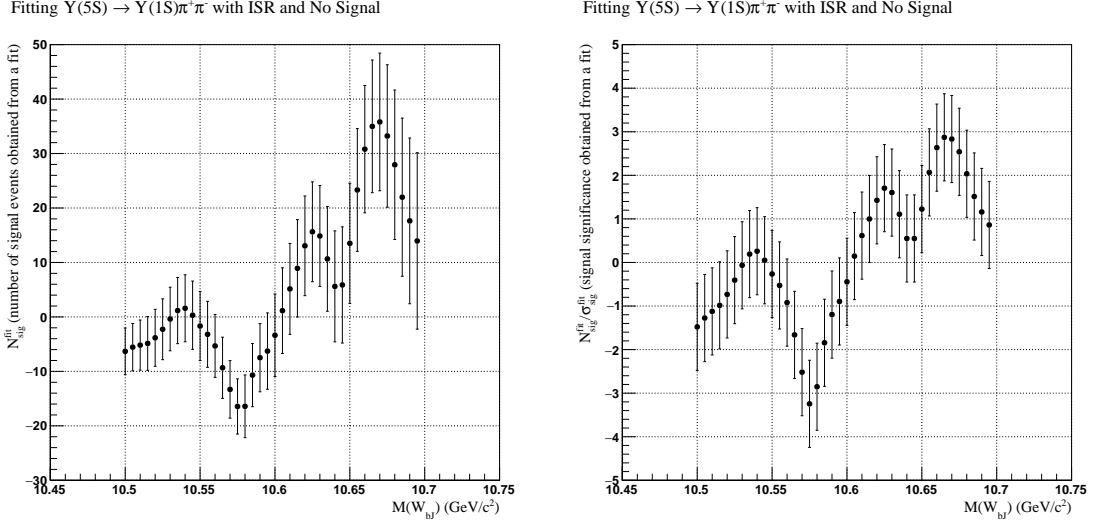


Figure 46: The results of the scan for ISR MC sample with background model and signal PDF shape.

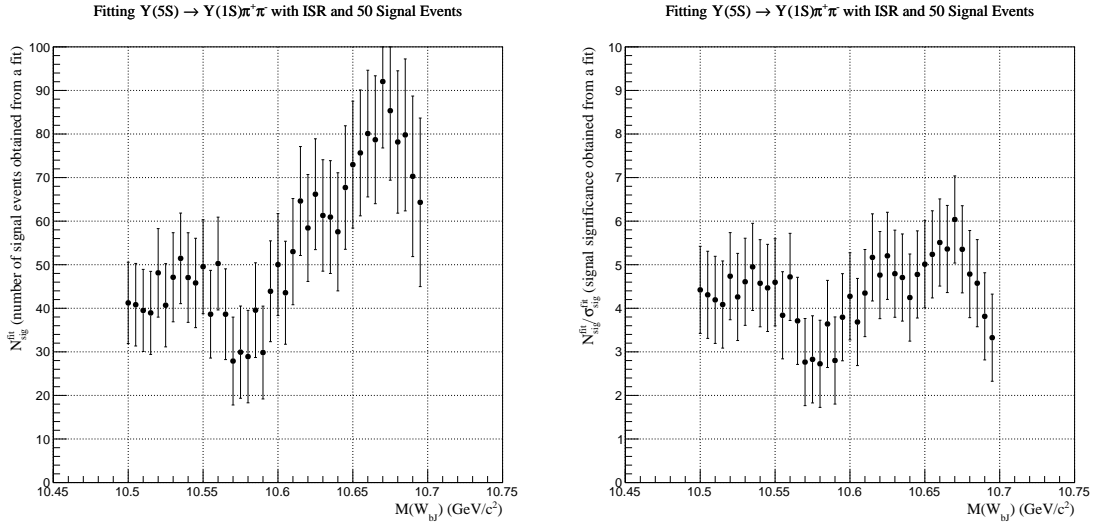


Figure 47: The results of the scan for ISR MC sample + 50 toy MC signal events with background model and signal PDF shape.

- 847 [10] A. E. Bondar, R. V. Mizuk, and M. B. Voloshin. Bottomonium-like states: Physics case
848 for energy scan above the $B\bar{B}$ threshold at Belle-II. *Mod. Phys. Lett.*, A32(04):1750025,
849 2017.
- 850 [11] A. E. Bondar, A. Garmash, A. I. Milstein, R. Mizuk, and M. B. Voloshin. Heavy quark
851 spin structure in Z_b resonances. *Phys. Rev.*, D84:054010, 2011.
- 852 [12] Stephen Lars Olsen. A New Hadron Spectroscopy. *Front. Phys.(Beijing)*, 10(2):121–154,
853 2015.

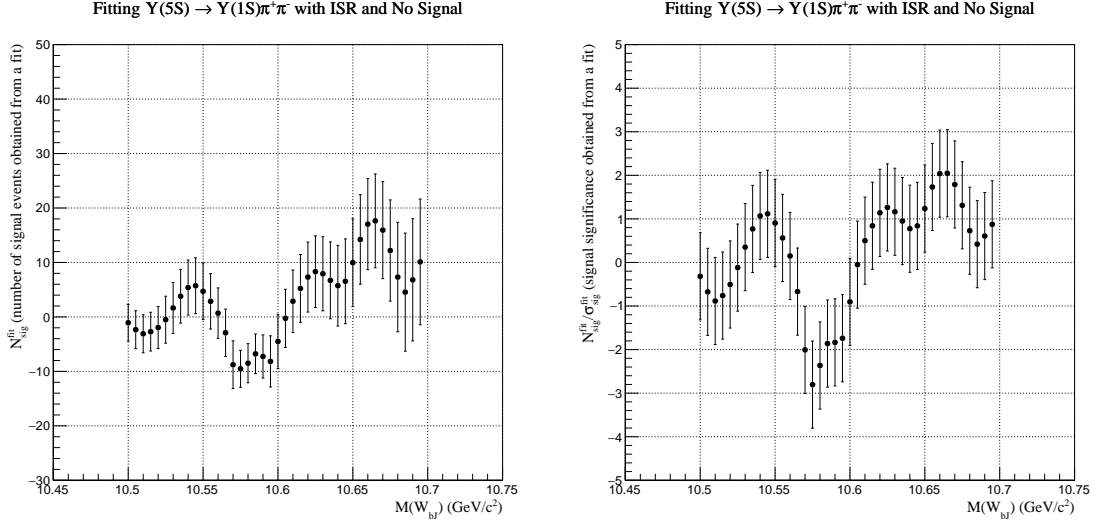


Figure 48: The results of the scan for ISR MC sample with background model and signal PDF shape.

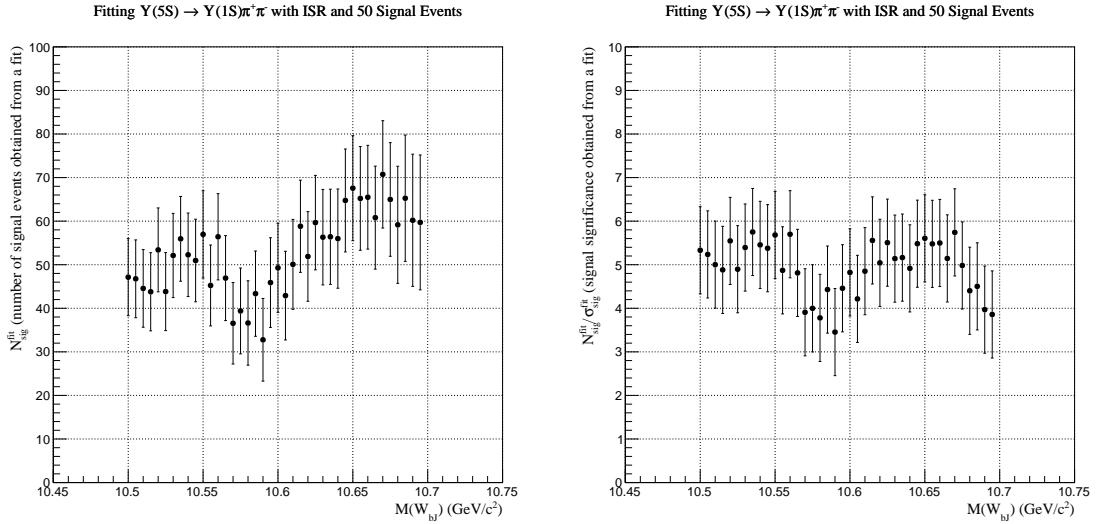


Figure 49: The results of the scan for ISR MC sample + 50 toy MC signal events with background model and signal PDF shape.

- 854 [13] M. B. Voloshin. Radiative transitions from Upsilon(5S) to molecular bottomonium.
855 *Phys. Rev.*, D84:031502, 2011.
- 856 [14] W. Altmannshofer et al. The Belle II Physics Book, arXiv:hep-ex/1808.10567, 2018.
- 857 [15] I. Adachi et al. Study of Three-Body Y(10860) Decays, arXiv:hep-ex/1209.6450,
858 BELLE-CONF-1272, 2012.
- 859 [16] D. J. Lange. The EvtGen particle decay simulation package. *Nucl. Instrum. Meth.*,
860 A462:152–155, 2001.

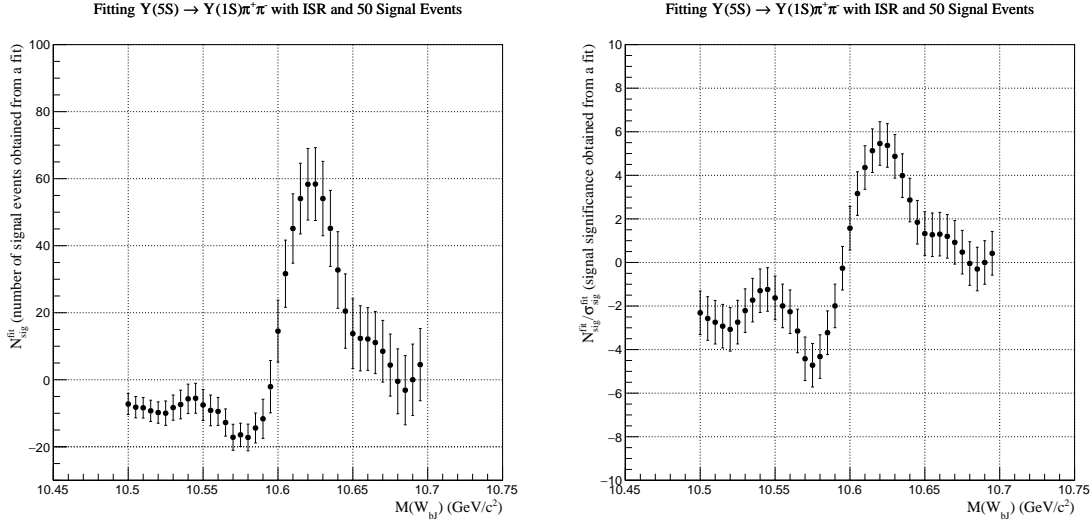


Figure 50: The results of the scan for ISR MC sample + 50 signal MC signal events ($M(W_{bJ}) = 10.62 \text{ GeV}/c^2$) with background model and signal PDF shape.

- 861 [17] S. Agostinelli et al. GEANT4: A Simulation toolkit. *Nucl. Instrum. Meth.*, A506:250–
 862 303, 2003.
- 863 [18] A. Ryd, D. Lange, N. Kuznetsova, S. Versille, M. Rotondo, D. P. Kirkby, F. K. Wuerth-
 864 wein, and A. Ishikawa. EvtGen: A Monte Carlo Generator for B -Physics. 2005.
- 865 [19] E. Barberio and Z. Was. PHOTOS: A Universal Monte Carlo for QED radiative cor-
 866 rections. Version 2.0. *Comput. Phys. Commun.*, 79:291–308, 1994.
- 867 [20] J. Brodzicka et al. Physics Achievements from the Belle Experiment. *PTEP*,
 868 2012:04D001, 2012.
- 869 [21] M. Benayoun, S. I. Eidelman, V. N. Ivanchenko, and Z. K. Silagadze. Spectroscopy at B
 870 factories using hard photon emission. *Mod. Phys. Lett.*, A14:2605–2614, 1999. [Frascati
 871 Phys. Ser.15(1999)].
- 872 [22] Wouter Verkerke and David P. Kirkby. The RooFit toolkit for data modeling. *eConf*,
 873 C0303241:MOLT007, 2003. [186(2003)].