SLAC-R-776

Study of the Rare Decay B Mesons Decaying to X Mesons Positive And Negative Leptons at BABAR

By Ventzislav B. Koptchev

Ph.D. Thesis

Stanford Linear Accelerator Center, Stanford University, Stanford, CA 94309

Work supported by Department of Energy contract DE–AC02–76SF00515.

STUDY OF THE RARE DECAY $B \to X_s \, \ell^+ \ell^-$ AT *BABAR*

A Dissertation Presented

by

VENTZISLAV B. KOPTCHEV

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

June 2004

Department of Physics

STUDY OF THE RARE DECAY $B \to X_s \, \ell^+ \ell^-$ AT *BABAR*

A Dissertation Presented

by

VENTZISLAV B. KOPTCHEV

Approved as to style and content by:

Stéphane Y. Willocq, Chair

Carlo Dallapiccola, Member

Eugene Golowich, Member

Grant W. Wilson, Member

Jonathan L. Machta, Department Head
Physics

ACKNOWLEDGEMENTS

I have learned a great deal from all I have worked with over the years and gratefully acknowledge my debt to them.

I would like to thank the Experimental High Energy Physics group at the University of Massachusetts - Amherst, for giving me the opportunity to work at a modern high energy physics fascility and for their support throughout my graduate studies. My special thanks go to my adviser, Prof. Stéphane Willocq, for his invaluable help during the initial frustrating period, and expert advice on so many issues related to the analysis and my education in general. I am grateful to Hermann Staengle and Toshinori Abe for their crucial contributions in the areas of signal extraction, background estimates and systematic error evaluation, without their help the timely completion of the analysis would not have been possible.

In the course of my graduate studies I have benefited from numerous lectures I have taken at UMass. I would especially like to thank Prof. Eugene Golowich, whose unique lecturing style has brought back memories of the excitement felt at the time of my first encounters with the wonderful world of Physics.

I would like to thank my primary school physics teacher Asen Tasev for kindling the spark and all the authors, including Paolo Maffei and the late Carl Sagan, who kept it growing.

Finally I would like to thank my parents for the sacrifices and unwavering support that allowed me to pursue my interests.

ABSTRACT

Flavor-changing neutral current transitions are forbidden at tree level in the Standard Model and can only occur via higher order diagrams. Since the amplitudes for such loops are dominated by the heaviest known particles, and non-SM effects are expected to contribute at the same order as the SM, such processes are an ideal place to look for new physics.

We present a measurement of the inclusive branching fraction for the flavor-changing neutral current process $B \to X_s \ell^+ \ell^-$ with a sample of 81.9 fb^{-1}, collected with the *BABAR* detector at the Stanford Linear Accelerator Center. The final state is reconstructed from e^+e^- or $\mu^+\mu^-$ pairs and a hadronic system consisting of one $K^\pm$ or K_s and up to two pions, with at most one π^0. We observe a signal of 40 ± 10(stat) ± 2(syst) events and extract a branching fraction $\mathcal{B}\,(B \to X_s\,\ell^+\ell^-) = (5.6 \pm 1.5(\text{stat}) \pm 0.6(\text{exp. syst}) \pm 1.1(\text{model syst})) \times 10^{-6}$ for $m_{ll} > 0.2$ GeV.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Theory

1.1 Introduction

The *Standard Model* (SM) combines the *Electroweak theory* of Weinberg [4] and Salam [5] with the theory of strong interactions – *Quantum Chromodynamics* (QCD)– to produce the contemporary theory of elementary particles. The model is based on relativistic quantum field theory, where the Lagrangian formalism is the natural choice, since Lorentz invariance is built in by construction. The Lagrangian function $\mathcal{L}$ is a real, relativistically invariant function depending on the fields and their first derivatives

$$\mathcal{L}(\psi_i(x), \partial_\mu \psi_i(x)). \tag{1.1}$$

In general, $\mathcal{L}$ contains kinetic terms for each field. In addition, it may contain interaction terms – products of the interacting fields multiplied by a coupling constant, which specifies the strength of the interaction. Mass terms are also possible and are quadratic with respect to the massive field. A field's kinetic term depends on $\partial_\mu \psi_i(x)$, the exact form being constrained by the requirement of Lorentz invariance and the field's transformation properties under the Lorentz group. Thus, a scalar field representing a spin-0 particle is a singlet under the Lorentz group and

the kinetic term T for such a field is

$$T = (\partial^{\mu}\phi(x))(\partial_{\mu}\phi(x)). \tag{1.2}$$

A spinor representing a spin 1/2 particle, transforms as

$$\psi'(x') = exp(i\epsilon_{\mu\nu}\sigma^{\mu\nu})\psi(x), \tag{1.3}$$

where

$$\sigma^{\mu\nu} = \frac{i}{2}[\gamma^{\mu}, \gamma^{\nu}], \tag{1.4}$$

resulting in a kinetic term

$$T = \bar{\psi}(x)\gamma^{\mu}\partial_{\mu}\psi(x), \quad \text{with } \bar{\psi}(x) \equiv \psi^{\dagger}(x)\gamma_0. \tag{1.5}$$

A spin-1 particle field A_{μ}, transforms as a Lorentz vector and the kinetic term is introduced via the field strength tensor F

$$T = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}, \;\; F_{\mu\nu} \equiv \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu}. \tag{1.6}$$

In the Standard Model, force mediators appear in the form of spin-1 fields, referred to as *gauge fields*, which arise when local gauge invariance is imposed on the Lagrangian. Under local transformations represented by a group G of dimension N, the *matter fields* transform as

$$\psi(x) \rightarrow U\psi(x), \quad \text{with } U(x) \equiv e^{-i\alpha_a(x)T^a}, \tag{1.7}$$

where T_a form an $N \times N$ representation of the group generators and $\alpha_a(x)$ are arbitrary real functions. The requirement of local gauge invariance dictates that all field derivatives in $\mathcal{L}$ must enter as

$$D_{\mu}\psi(x) \equiv \partial_{\mu}\psi(x) + igT^aW^a_{\mu}(x)\psi(x), \tag{1.8}$$

where g is a coupling constant and W^a are a set of N Lorentz vectors - the gauge fields. The gauge invariant Lagrangian will thus be of the form

$$\mathcal{L}(\psi_i, (\partial_\mu + igT^a W^a_\mu)\psi_i) \tag{1.9}$$

and will contain interaction terms between the gauge and the matter fields. The derivative defined in Eq.(1.8) must be covariant under the group transformations

$$D_\mu \psi \to U D_\mu \psi, \tag{1.10}$$

which imposes the following infinitesimal transformation rule for the gauge fields:

$$\delta W^k_\mu = c_{lmk} W^m_\mu \alpha_l(x) + \frac{1}{g}\partial_\mu \alpha_k(x), \tag{1.11}$$

where c_{lmk} are the group's structure constants

$$[T_j, T_k] = i c_{jkl} T_l. \tag{1.12}$$

The gauge invariant kinetic term associated with the gauge fields W^a is then

$$\mathcal{L}_0 = -\frac{1}{4} F^j_{\mu\nu} F^{j,\mu\nu}, \tag{1.13}$$

where F is the field strength tensor

$$F^j_{\mu\nu} \equiv \partial_\mu W^j_\nu - \partial_\nu W^j_\mu + g c_{jkl} W^k_\mu W^l_\nu, \tag{1.14}$$

which transforms under the adjoint representation of G

$$\delta F^j_{\mu\nu} = c_{jkl} \alpha_k F^l_{\mu\nu}. \tag{1.15}$$

It follows from Eqs.(1.13 and 1.14), that in a non-Abelian case ($c_{jkl} \neq 0$), the gauge fields will be self-coupled. A mass term in the form $W_\mu W^\mu$, however, cannot be introduced since, according to Eq.(1.11), it would not be gauge invariant.

In the Standard Model, particle masses are generated through the Higgs Mechanism (spontaneous symmetry breaking in a gauge theory). To employ this mechanism, a number of scalar fields with degenerate ground state are introduced in the Lagrangian, all ground states being related by gauge transformations. The coupling between the gauge bosons and the scalar fields is ensured by the local gauge invariance requirement, through the covariant derivative. The choice of a particular ground state is arbitrary due to the gauge symmetry, however the symmetry is then broken as that state is not invariant under G. The state may be invariant under a sub-group of G - G' of dimension N', the difference N-N' indicating the number of broken generators. The physical fields corresponding to interaction mediators are obtained through an appropriate unitary transformation of the initial gauge fields. The ones multiplied by a G' generator in the Lagrangian remain massless, the other N-N' acquire mass. Expansion of the scalar fields around the chosen ground state, gives rise to N-N' massless scalar fields - *Goldstone bosons*, which are eliminated through an appropriate gauge choice, effectively transferring their degree of freedom to a longitudinal component of the now massive vector bosons. The remaining scalar fields appear as massive *Higgs bosons*. In Nature, there are 9 unbroken symmetry generators - those associated with electric charge and the strong interaction colors. Hence there are 9 massless mediators, the photon and the 8 strong interaction carriers, the gluons.

1.2 Model Construction

When building a Lagrangian for a physical interaction the choice of a gauge symmetry group is motivated by the number of force mediators [6]. For the electromagnetic interaction there is one carrier - the photon and correspondingly a

group with one generator $U(1)$. The weak interaction is conveyed by three vector bosons, $W^{\pm}$ and Z^0, and the simplest group with three generators is $SU(2)$. The eight gluon carriers of the strong interaction are represented by $SU(3)$. The matter fields in the Lagrangian represent the fundamental constituents - quarks and leptons to which at least 4 scalar fields have to be added to invoke the Higgs mechanism, since three of the carriers – the weak interaction bosons – are massive. The transformation properties of the matter fields under the chosen gauge group determine their coupling to the gauge bosons. The simplest way to obtain the observed weak-interaction coupling structure is to arrange the leptons and quarks in left-handed $SU(2)$ doublets and right-handed $SU(2)$ singlets, and the 4 scalar fields in a complex $SU(2)$ doublet. Quark and lepton masses are generated by introducing *Yukawa terms* to the Lagrangian, through which quarks and leptons are coupled to the scalars in a gauge invariant manner. It is regarded that the $U(1)$ and $SU(2)$ transformations are generated by *weak hypercharge* Y_W and *weak isospin* I_W, related to the electric charge Q, the only unbroken $U(1) \times SU(2)$ generator, by the Gell-Mann - Nishijima relation Eq.(1.16)

$$Y_W/2 = Q - I_{3W}. \tag{1.16}$$

The electroweak interaction properties of the matter fields are summarized in Table 1.

As far as the strong interaction is concerned the quarks are arranged in an $SU(3)$ triplet, the transformations being generated by the quark color. As colorless, all other particles are represented by $SU(3)$ singlets.

The overall gauge group is thus $U(1)_{weak\ hypercharge} \times SU(2)_{weak\ isospin} \times SU(3)_{color}$, which we refer to as $U(1)_Y \times SU(2)_I \times SU(3)_C$ in the following. It undergoes spontaneous breaking to $U(1)_{el.\ charge} \times SU(3)_C$, resulting in one massless gauge boson,

Table 1: Electroweak properties of the fundamental matter constituents.

Particle	Y_W	I_{3W}	Q
$\binom{\nu_e}{e^-}_L \binom{\nu_\mu}{\mu^-}_L \binom{\nu_\tau}{\tau^-}_L$	-1	$\binom{1/2}{-1/2}$	$\binom{0}{-1}$
$e_R\ \mu_R\ \tau_R$	-2	0	-1
$\binom{u}{d}_L \binom{c}{s}_L \binom{t}{b}_L$	1/3	$\binom{1/2}{-1/2}$	$\binom{2/3}{-1/3}$
$u_R\ c_R\ t_R$	4/3	0	2/3
$d_R\ s_R\ b_R$	-2/3	0	-1/3
$\binom{\phi^+}{\phi^0}$	1	$\binom{1/2}{-1/2}$	$\binom{1}{0}$

the photon A_μ, three massive ones, the W_μ^+, W_μ^- and Z_μ^0 bosons, and a massive scalar field, the Higgs field. The initial $SU(3)$ gauge symmetry remains unbroken and its gauge bosons unrelated to the electroweak carriers, keeping Quantum Chromodynamics a separate theory.

The initial gauge-invariant electro-weak Lagrangian is

$$\mathcal{L}_{EW} = \mathcal{L}_{scalar} + \mathcal{L}_{gauge} + \mathcal{L}_{matter} + \mathcal{L}_{YL} + \mathcal{L}_{YQ}. \tag{1.17}$$

The scalar part is

$$\mathcal{L}_{scalar} = (D^\mu \phi)^\dagger (D_\mu \phi) - V(\phi^\dagger \phi), \tag{1.18}$$

where D_μ is the $U(1)_Y \times SU(2)_I$ covariant derivative

$$D_\mu = \partial_\mu + \frac{ig\prime}{2} Y_W B_\mu + \frac{ig}{2} \overrightarrow{\tau} \cdot \overrightarrow{W_\mu}. \tag{1.19}$$

The 2-dimensional $SU(2)$ generators are the Pauli matrices $\overrightarrow{\tau}$, g' and g are the $U(1)_Y$ and $SU(2)_I$ coupling constants, respectively, B_μ and $\overrightarrow{W_\mu}$ are the $U(1)_Y$ and $SU(2)_I$ gauge bosons, and ϕ is the Higgs complex doublet

$$\phi = \begin{pmatrix} \phi^+ \\ \\ \phi^0 \end{pmatrix}. \tag{1.20}$$

The potential in Eq.(1.18) is

$$V(\phi^\dagger\phi) = \mu^2(\phi^\dagger\phi) + \lambda(\phi^\dagger\phi)^2 \tag{1.21}$$

with $\lambda > 0$ and $\mu^2 < 0$ so that there is a degenerate vacuum state with expectation value $v \equiv \sqrt{\frac{-\mu^2}{\lambda}}$. Choosing a vacuum state of the scalar field

$$\phi_0 = \begin{pmatrix} 0 \\ v/\sqrt{2} \end{pmatrix} \tag{1.22}$$

breaks both $U(1)_Y$ and $SU(2)_I$ symmetries, but preserves an invariance under the $U(1)_Q$ generated by the electric charge operator

$$e^{i\alpha Q}\phi_0 = \phi_0, \tag{1.23}$$

since

$$Q\phi_0 = \frac{1}{2}\left(\tau_3 + Y\right)\phi_0 = 0. \tag{1.24}$$

Expanding ϕ about the minimum of the potential gives

$$\phi = exp\left(\frac{i\zeta\cdot\tau}{2v}\right)\begin{pmatrix} 0 \\ \left(v+H\right)/\sqrt{2} \end{pmatrix}, \tag{1.25}$$

where ζ are the Goldstone boson fields and H is the Higgs scalar field. The exponential factor can be eliminated with a gauge transformation leaving

$$\phi = \begin{pmatrix} 0 \\ \left(v+H\right)/\sqrt{2} \end{pmatrix}. \tag{1.26}$$

The kinetic term for the gauge fields is

$$\mathcal{L}_{gauge} = -\frac{1}{4}F^l_{\mu\nu}F^{l\mu\nu} - \frac{1}{4}f_{\mu\nu}f^{\mu\nu} \tag{1.27}$$

and the field strength tensors are

$$F^l_{\mu\nu} = \partial_\mu W^l_\nu - \partial_\nu W^l_\mu + g\epsilon_{jkl}W^j_\mu W^k_\nu \tag{1.28}$$

for the $SU(2)_I$ gauge fields and

$$f_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \tag{1.29}$$

for the $U(1)_Y$ gauge field. The matter term is

$$\mathcal{L}_{matter} = \bar{R} i\gamma^\mu (\partial_\mu + \frac{ig\prime}{2} B_\mu Y_W) R + \bar{L} i\gamma^\mu D_\mu L, \tag{1.30}$$

where L denotes a left-handed $SU(2)_I$ doublet

$$L \equiv Q_{iL}, L_{iL} \tag{1.31}$$

and R denotes a right-handed $SU(2)_I$ singlet

$$R \equiv U_{iR}, D_{iR}, L_{iR}, \tag{1.32}$$

with

$$Q_{iL} \equiv \binom{u}{d}_L, \binom{c}{s}_L, \binom{t}{b}_L, \tag{1.33}$$

$$L_{iL} \equiv \binom{\nu_e}{e}_L, \binom{\nu_\mu}{\mu}_L, \binom{\nu_\tau}{\tau}_L, \tag{1.34}$$

$$U_{iR} \equiv u_R, c_R, t_R, \tag{1.35}$$

$$D_{iR} \equiv d_R, s_R, b_R, \tag{1.36}$$

$$L_{iR} \equiv e_R, \mu_R, \tau_R, \tag{1.37}$$

ν_R terms are missing since neutrinos are treated as massless. The interaction part of $\mathcal{L}_{matter}$ can be summarized in

$$\mathcal{L}_{INT} = g' J_B^\mu B_\mu + g \overrightarrow{J}^\mu \cdot \overrightarrow{W}_\mu, \tag{1.38}$$

where the currents are

$$J_B^\mu = -\frac{1}{2} \bar{R} \gamma^\mu Y_W R - \frac{1}{2} \bar{L} \gamma^\mu Y_W L \tag{1.39}$$

and

$$J_i^\mu = -\frac{1}{2}\bar{L}\gamma^\mu\tau_i L. \tag{1.40}$$

The physical fields W_+^μ, W_-^μ, Z^μ and A^μ corresponding to W^+, W^-, Z and the photon are linear combinations of the initial gauge fields

$$\begin{aligned} A^\mu &= \sin\theta_W W_3^\mu + \cos\theta_W B^\mu, \\ Z^\mu &= \cos\theta_W W_3^\mu - \sin\theta_W B^\mu, \\ W_\pm^\mu &= \frac{1}{\sqrt{2}}\left(W_1^\mu \mp i W_2^\mu\right), \end{aligned} \tag{1.41}$$

where θ_W is the *weak mixing angle*. In terms of them the interactions are

$$\begin{aligned} \mathcal{L}_{INT} &= \frac{g}{2\sqrt{2}}[J_-^\mu W_{\mu+} + J_+^\mu W_{\mu-}] \\ &+ [(g\cos\theta_W + g'\sin\theta_W)J_3^\mu - g'\sin\theta_W J_{em}^\mu]Z_\mu \\ &+ [g'\cos\theta_W J_{em}^\mu + (g'\cos\theta_W - g\sin\theta_W)J_3^\mu]A_\mu, \end{aligned} \tag{1.42}$$

where the *charged current* and the *electromagnetic current* are defined as

$$J_+^\mu = 2\left(J_1^\mu - iJ_2^\mu\right) = 2\bar{U}_{iL}\gamma^\mu D_{iL}, \tag{1.43}$$

$$J_-^\mu = 2\left(J_1^\mu + iJ_2^\mu\right) = 2\bar{D}_{iL}\gamma^\mu U_{iL}, \tag{1.44}$$

$$\begin{aligned} J_{em}^\mu &= J_3^\mu + J_B^\mu = -\frac{1}{2}\left[\bar{L}\gamma^\mu \begin{pmatrix} 1+Y_W & 0 \\ 0 & -1+Y_W \end{pmatrix} L + \bar{R}\gamma^\mu Y_W R\right] \\ &= -Q\bar{f}\gamma^\mu f, \end{aligned} \tag{1.45}$$

where $U_{iL} = (u, c, t, \nu_e, \nu_\mu, \nu_\tau)$, $D_{iL} = (d, s, b, e, \mu, \tau)$, $Q = I_3 + Y_W/2$ and f is a fermion field. In order to identify the last term in Eq.(1.42) with the electromagnetic interaction

$$\mathcal{L}_{INT}^{em} = eJ_{em}^\mu A_\mu, \tag{1.46}$$

we must impose the relation $e = g'\cos\theta_W = g\sin\theta_W$, where e is the electron charge.

Thus the interaction Lagrangian takes the form

$$\begin{aligned}\mathcal{L}_{INT} &= \frac{e}{2\sqrt{2}\sin\theta_W}(W^{\mu}_{+}J_{\mu-} + W^{\mu}_{-}J_{\mu+}) \\ &+ \frac{e}{2\cos\theta_W\sin\theta_W}Z^{\mu}J_{\mu\ NC} + eA^{\mu}J_{\mu\ em}, \end{aligned} \tag{1.47}$$

where the *weak neutral current* is defined as

$$\begin{aligned} J^{\mu}_{NC} &= 2(J^{\mu}_{3} - \sin^2\theta_W J^{\mu}_{em}) = -\bar{L}\gamma^{\mu}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} L + 2\sin^2\theta_W Q\bar{f}\gamma^{\mu}f \\ &= \bar{f}\gamma^{\mu}(2Q\sin^2\theta_W - I_3 + I_3\gamma^5)f. \end{aligned} \tag{1.48}$$

The mass Lagrangian for the gauge bosons results from the covariant derivatives in $\mathcal{L}_{scalar}$, where after substituting the physical gauge fields and the expression Eq.(1.26) for ϕ, one gets

$$\begin{aligned} M_{W^{\pm}} &= gv/2, \\ M_Z &= M_{W^{\pm}}/\cos\theta_W. \end{aligned} \tag{1.49}$$

The scalar Lagrangian also contains a mass term for the Higgs field

$$m_H = \sqrt{-2\mu^2} \tag{1.50}$$

and interaction terms with the gauge bosons.

The lepton fields acquire mass through the Yukawa term

$$\mathcal{L}_{YL} = -G_i(\bar{L}_{iL}\phi L_{iR} + \bar{L}_{iR}\phi^{\dagger}L_{iL}), \tag{1.51}$$

where G_i are unknown coupling constants. After substituting Eq.(1.26) $\mathcal{L}_{YL}$ becomes

$$\mathcal{L}_{YL} = -G_i(\bar{\ell}_{iR}\ell_{iL} + \bar{\ell}_{iL}\ell_{iR})\frac{v+H}{\sqrt{2}} = \frac{-G_i v}{\sqrt{2}}\bar{\ell}_i\ell_i - \frac{-G_i}{\sqrt{2}}\bar{\ell}_i\ell_i H, \tag{1.52}$$

where ℓ_i is e, μ or τ and the leptons have acquired mass $m_i = G_i v/\sqrt{2}$. The Yukawa term for the quarks is

$$\mathcal{L}_{YQ} = -G^u_{ij}\bar{Q}_{iL}\bar{\phi}U_{jR} - G^d_{ij}\bar{Q}_{iL}\phi D_{jR} + h.c., \tag{1.53}$$

where

$$\bar{\phi} = i\tau_2\phi^* = \begin{pmatrix} \bar{\phi}^0 \\ -\phi^- \end{pmatrix}. \tag{1.54}$$

Local gauge invariance does not require the matrices $G^{u/d}_{ij}$ to be diagonal. In fact, to obtain the observed quark mixing they cannot be. Therefore, after the symmetry breaking, non diagonal mass matrices arise

$$M^{u/d}_{ij} = \frac{v}{\sqrt{2}} G^{u/d}_{ij}. \tag{1.55}$$

To determine the quark mass eigenstates, the mass matrices must be diagonalized

$$(\mathrm{U}^{u/d}_L)^\dagger M^{u/d} \mathrm{U}^{u/d}_R = M^{u/d}_{diag}, \tag{1.56}$$

where $\mathrm{U}^{u/d}_L$ and $\mathrm{U}^{u/d}_R$ are unitary matrices linked to the basis change

$$\begin{pmatrix} u_{L/R} \\ c_{L/R} \\ t_{L/R} \end{pmatrix} \to \mathrm{U}^u_{L/R} \begin{pmatrix} u_{L/R} \\ c_{L/R} \\ t_{L/R} \end{pmatrix}, \quad \begin{pmatrix} d_{L/R} \\ s_{L/R} \\ b_{L/R} \end{pmatrix} \to \mathrm{U}^d_{L/R} \begin{pmatrix} d_{L/R} \\ s_{L/R} \\ b_{L/R} \end{pmatrix}. \tag{1.57}$$

Since $M^{u/d}$ are hermitian by construction, $\mathrm{U}^{u/d}_L = \mathrm{U}^{u/d}_R$. It is worth noting that J^μ_{NC} (1.48) and J^μ_{em} (1.45) are flavor diagonal in both the *weak* and the *mass* basis. This is the essence of the GIM mechanism and implies absence of flavor-changing neutral current (FCNC) transitions at the tree level. Thus the neutral currents are not affected by the basis change. However, for the hadronic charged currents this leads to

$$J^\mu_{+\ HAD} = 2(\bar{u}_L, \bar{c}_L, \bar{t}_L)\gamma^\mu (\mathrm{U}^u_L)^\dagger \mathrm{U}^d_L \begin{pmatrix} d_L \\ s_L \\ b_L \end{pmatrix}, \tag{1.58}$$

$$J^{\mu}_{-\ HAD} = 2(\bar{d}_L, \bar{s}_L, \bar{b}_L)(\mathrm{U}_L^d)^{\dagger}\mathrm{U}_L^u\gamma^{\mu}\begin{pmatrix} u_L \\ c_L \\ t_L \end{pmatrix}. \tag{1.59}$$

The unitary matrix

$$V_{CKM} \equiv (\mathrm{U}_L^u)^{\dagger}\mathrm{U}_L^d \tag{1.60}$$

is known as the Cabibbo-Kobayashi-Maskawa (CKM) matrix [7] [8]. Because of the unitarity and the 5 arbitrary phases between the six quarks, there are only four independent parameters in the CKM matrix. They account for 4 of the 18 Standard Model parameters to be determined by experiment, the other 14 being the 9 fermion masses, 3 coupling constants, the Higgs mass and Higgs vacuum expectation value. A convenient representation, in which the matrix elements are expressed as a power series of a small parameter λ, was introduced by Wolfenstein (1.61).

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} 1-\frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + O(\lambda^4). \tag{1.61}$$

So far $|V_{ud}|$ and $|V_{us}|$ have been determined precisely in nuclear beta decay and $K \to \pi\ell\nu$, respectively and $|V_{cb}|$ has been determined from semileptonic B decays. These give values for the parameters $\lambda = 0.22$ and $A = 0.84$ accurate to about 1%. The other two parameters however, are still poorly determined.

1.3 Search for New Physics

1.3.1 CP violation

In addition to the continuous transformations of the Poincare group, there are three independent discrete transformations which preserve the space-time interval $t^2 - \overrightarrow{x}^2$. These are the *Charge conjugation*, the *Parity* and the *Time reversal*. The charge conjugation operator C transforms a particle into its anti-particle, the parity operator P flips the sign of the space components in a 4-vector and the time reversal operator T changes the 4-vector's time component sign. All three together (i.e. CPT) represent a fundamental symmetry of Nature. However, they are individually violated in the weak interactions. Parity and Charge conjugation violation were discovered by Wu et al. [9] and Goldhaber et al. [10], respectively, and are accommodated in the SM with the choice of gauge transformation properties of the left and right-handed matter fields. This results in the charged currents having the form $\bar{f_L}\gamma^\mu f_L = 1/2\bar{f}\gamma^\mu(1-\gamma^5)f$, which has a *vector - axial vector* (V-A) structure. CP (and therefore T) violation was first observed in the K system by Christenson et al. [11] and recently confirmed by the *BABAR* and $BELLE$ collaborations in the B system [12] and [13]. The phenomenon is built in the model through the presence of a non-zero complex phase in the CKM matrix as the weak Hamiltonian then does not commute with the operator T. In 1967 Sakharov suggested that CPV is necessary to generate the matter-antimatter asymmetry in the Universe [14]. However, despite the huge success of the Standard Model in recent years, this sole source of CP violation in the theory is not enough to explain the amount of the observed asymmetry.

The phase information in the CKM matrix can be displayed using the so-called

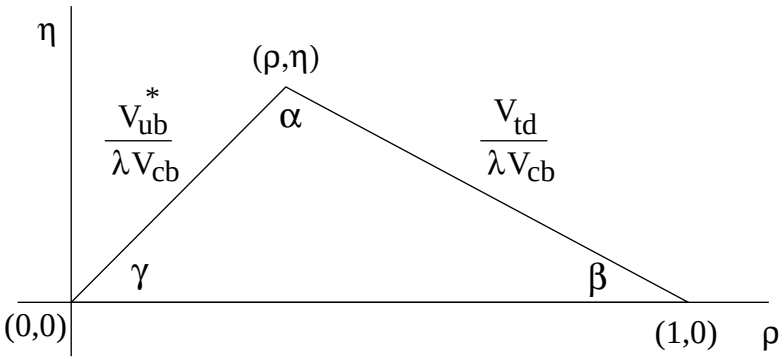

Figure 1: The unitarity triangle.

unitarity triangle (Fig. 1), which follows from the orthogonality of the first and third columns

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0, \tag{1.62}$$

rescaled by $|V_{cd}V_{cb}^*| = A\lambda^3 = \lambda|V_{cb}|$. The non-zero imaginary component $i\eta$ in V_{ub} and V_{td} see Eq.(1.61) results in a non-zero area of the triangle. The angles α, β and γ can be measured via CP violation in B meson decays. The sides of the triangle can also be measured independently. Any deviation from $\alpha + \beta + \gamma = \pi$ or any inconsistency with the independently measured sides would indicate physics beyond the Standard Model. Another place to look for new physics are rare decays which could provide complementary information to the CP asymmetry measurements and thus help identifying the source of any new physics.

1.3.2 New Physics in Rare B Decays

Although FCNC transitions are forbidden at tree level in the Standard Model, due to the different masses of the same-charge quarks, such processes can occur via one-loop diagrams (see Fig.2), albeit at a very low rate [15][16][17][18]. Due to the

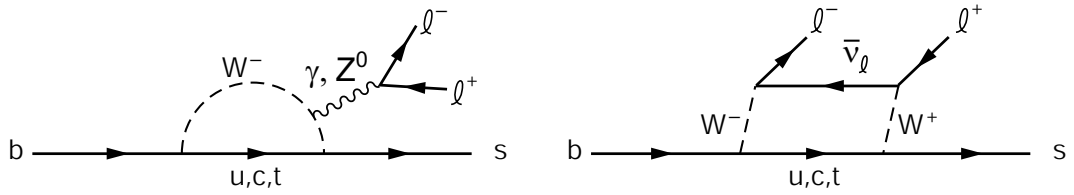

Figure 2: *Penguin* and *Box* Feynman diagrams for the transition $b \to s\ell^+\ell^-$ are shown on the left and right, respectively.

combinations of CKM matrix elements involved, such loops are quite sensitive to new physics. One-loop processes are possible in c and t decays, but these particles have the CKM-favored decays $c \to s$ and $t \to b$ accessible to them. On the other hand since the s and b quarks have no kinematically-allowed CKM-favored decay, the relative importance of the loop decays is greater. Both s and b loop decay amplitudes are dominated by t exchange due to its much higher mass with respect to u and c. Since the heaviest known particles (t, W, Z) dominate the amplitudes, rates for these processes are very sensitive to non-SM extensions with for example, heavy charged Higgs or supersymmetric particles. This makes the FCNC processes the most sensitive low energy probes for such extensions to the Standard Model. Searching for new physics in b rather than s decays has some definite advantages, however. Since the s quark mass is relatively small, $s \to d$ loop diagrams are typically dominated by large non-perturbative effects. In the case of $b \to s(d)$, due to $|V_{tb}|$ being close to unity, the loop decays are less suppressed.

At very short distance (high energy) scales $O(M_{W,Z,t})$, quark decays are represented by Feynman diagrams (Fig. 2) with W, Z and quark propagators. A picture of a decaying hadron with a mass on the order of $O(M_{b,c,s})$ can be described by a set of effective point-like vertices (Fig. 3), represented by local operators Q_i. The coupling constants associated with those operators are known as the Wilson coef-

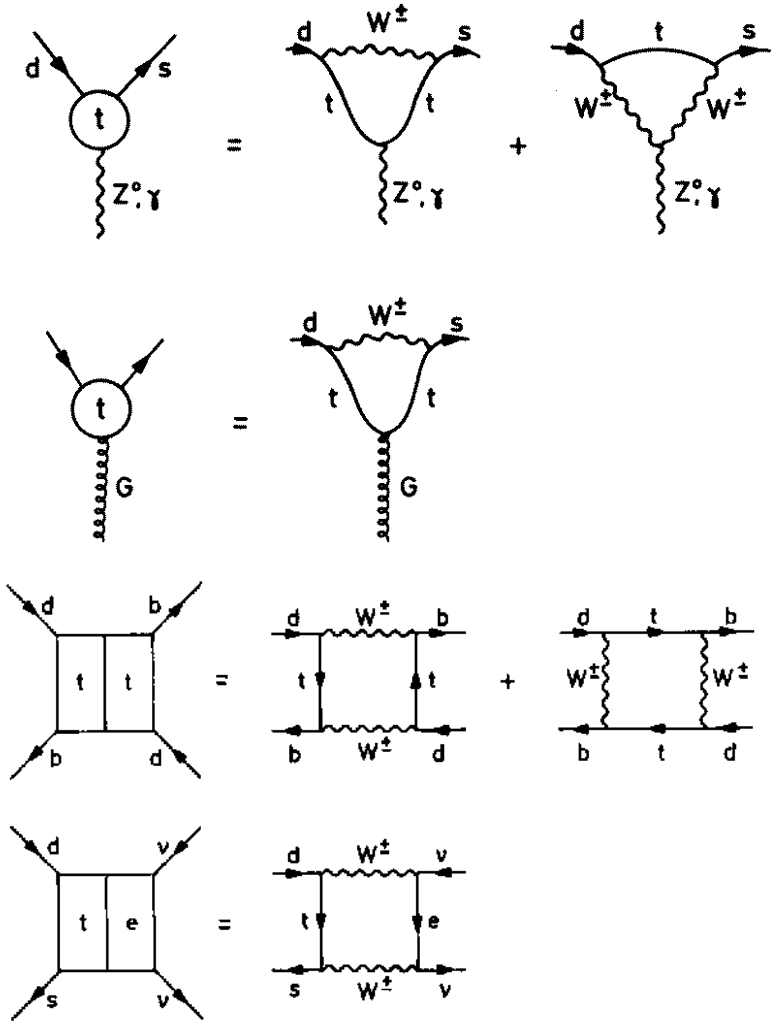

Figure 3: Effective "penguin" and "box" vertices in terms of basic vertices.

ficients C_i. The effective Hamiltonian describing weak hadron decays can thus be written in the form

$$H_{eff} = \frac{G_F}{\sqrt{2}} \sum_i V^i_{CKM} C_i(\mu) Q_i(\mu), \tag{1.63}$$

where G_F is the Fermi constant, μ is a separation energy scale and Q_i are the relevant local operators which govern the decays in question. The factors V_{CKM} and the Wilson coefficients C_i describe the strength with which a given operator

enters the Hamiltonian. In the case of beta decay, H_{eff} is

$$H^{\beta}_{eff} = \frac{G_F}{\sqrt{2}} \cos\theta_c [\bar{u}\gamma_\mu(1-\gamma_5)d\ \bar{e}\gamma^\mu(1-\gamma_5)\nu_e], \tag{1.64}$$

which corresponds to the Fermi theory of point-like 4-fermion interactions to which the known u-d mixing has been incorporated via the Cabibbo angle θ_c. In this particular case the Wilson coefficient is equal to unity and the local operator is given by a product of two currents.

Thus, to obtain an effective low energy theory relevant for scales $O(\mu \sim M_b \sim 5GeV)$, the heavy degrees of freedom – W, Z^0 and t – must be integrated out, after which they do not appear explicitly in the theory, but their effects are hidden in the Wilson coefficients. The Wilson coefficients $C_i(\mu)$ represent the short distance(energy scale higher than μ) electroweak interactions and the perturbative (due to asymptotic freedom) strong interactions. To calculate them, they are first obtained at high scale by evaluating the Feynman diagrams with full W, Z, t and new particle exchanges and include short-distance QCD effects. They are then evolved to a low-energy scale using renormalization group equations. This evolution mixes the operators

$$C_i(\mu) = \sum_j U_{ij}(\mu, M_W) C_j(M_W). \tag{1.65}$$

The renormalization guarantees that the μ dependence of C_i is canceled by the μ dependence of Q_i. Since extensions of the Standard Model contribute additional diagrams at the high-energy scale, they modify the values of the Wilson coefficients in the effective low-energy theory. The operator elements $< X|Q_i(\mu)|B >$ summarize the long-distance (energy scale lower than μ) non-perturbative strong interactions (here $|B>$ represents the B meson and $|X>$ the final state it decays into). Unlike the Wilson coefficients, the operator elements cannot be obtained perturbatively due to the confining nature of strong interactions at large

distances. Non-perturbative methods such as lattice calculations, QCD sum rules, Heavy Quark Effective Theory(HQET) and so on are used in this case, but they all have limitations. Consequently the dominant theoretical uncertainties in the decay amplitudes reside in the matrix elements $< Q_i(\mu) >$. However, when $|X >$ represents an inclusive final state, to leading order $< X_s|Q_i|B > \sim < s|Q_i|b >$, where $< s|Q_i|b >$ is an operator element for free quarks which can be calculated perturbatively if no resonant intermediate states are involved. The first nonperturbative corrections are of second order $O(1/M_b^2)$, and are small, thanks to the large b quark mass.

The subject of my study is the inclusive decay $B \to X_s\, \ell^+\ell^-$, for which the effective Hamiltonian reads [1]

$$H_{eff} = -\frac{4G_F}{\sqrt{2}} V_{ts}^{\star} V_{tb} \sum_{i=1}^{10} C_i(\mu) Q_i(\mu). \tag{1.66}$$

The first six coefficients $C_i(\mu)$, $i = 1, \ldots, 6$, determine the non-leptonic B-decay rates and the B-hadron lifetimes. Since data on B decays and the results obtained in the SM from Eq.(1.66), including QCD effects, are in good agreement with each other (20%), there is not much room left for the first six coefficients involving the four-quark operators to deviate from their SM values. The coefficients of interest are $C_7(\mu)$, ...,$C_{10}(\mu)$, since they are generated at scale $\mu = M_W$ by electroweak loops (penguins and boxes). Once extracted from experiment, these can be compared to the theoretical predictions. If there is a deviation from the SM values, new models accounting for the discrepancy can be looked for. If no deviation is observed, the results can be used to set bounds on new physics. In order to measure the values of these coefficients, several observables can be used [19]:

- Invariant dilepton mass distributions $d\mathcal{B}(B \to X_s\, \ell^+\ell^-)/d\hat{s}$, where

$$\hat{s} = \tfrac{s}{M_b^2}, \;\; s = (p_{\ell^+} + p_{\ell^-})^2,$$

- Forward-backward (FB) charge asymmetry $\mathcal{A}(\hat{s})$ in $B \to X_s\, \ell^+\ell^-$, defined with respect to the angular variable $z \equiv \cos\theta$, where θ is the angle of the ℓ^+ with respect to the b-quark direction in the center-of-mass system of the dilepton pair.

These quantities can be expressed in terms of the Wilson coefficients

$$\begin{aligned}\frac{d\mathcal{B}(B \to X_s\, \ell^+\ell^-)}{d\hat{s}} &= K(1-\hat{s})^2\{(|C_9^{eff}(\mu)|^2 + |C_{10}(\mu)|^2)(1+2\hat{s}) \\ &\quad +|C_7(\mu)|^2\frac{4}{\hat{s}}(2+\hat{s}) + 12\ \mathrm{Re}\ (C_7(\mu)C_9^{eff}(\mu))\}, \quad (1.67)\end{aligned}$$

$$\mathcal{A}(\hat{s}) = -\frac{3}{2}K(1-\hat{s})^2 C_{10}(\mu)\{C_9^{eff}(\mu)\hat{s} + 4C_7(\mu)\}, \quad (1.68)$$

where K is a constant and $C_9^{eff}(\mu) = C_9(\mu) + f(C_1(\mu), ..., C_6(\mu), \hat{s})$. Figures 4 and 5 show dilepton invariant mass and FB asymmetry distributions for several SM extensions as well as the SM predictions.

Recent calculations of the branching fractions are provided in Table 2. In the electron channel, the branching fraction is predicted to be $\mathcal{B}(B \to X_s\, e^+e^-) = (4.2 \pm 0.7) \times 10^{-6}$, for $m(e^+e^-) > 0.2$ GeV.

Table 2: Branching fraction predictions from Ali et al. [1]

Decay	Branching fraction
$B \to K e^+e^-$	$(0.35 \pm 0.12) \times 10^{-6}$
$B \to K^* e^+e^-$	$(1.58 \pm 0.49) \times 10^{-6}$
$B \to X_s\, e^+e^-$	$(6.89 \pm 1.01) \times 10^{-6}$
$B \to K \mu^+\mu^-$	$(0.35 \pm 0.12) \times 10^{-6}$
$B \to K^* \mu^+\mu^-$	$(1.19 \pm 0.39) \times 10^{-6}$
$B \to X_s\, \mu^+\mu^-$	$(4.15 \pm 0.70) \times 10^{-6}$

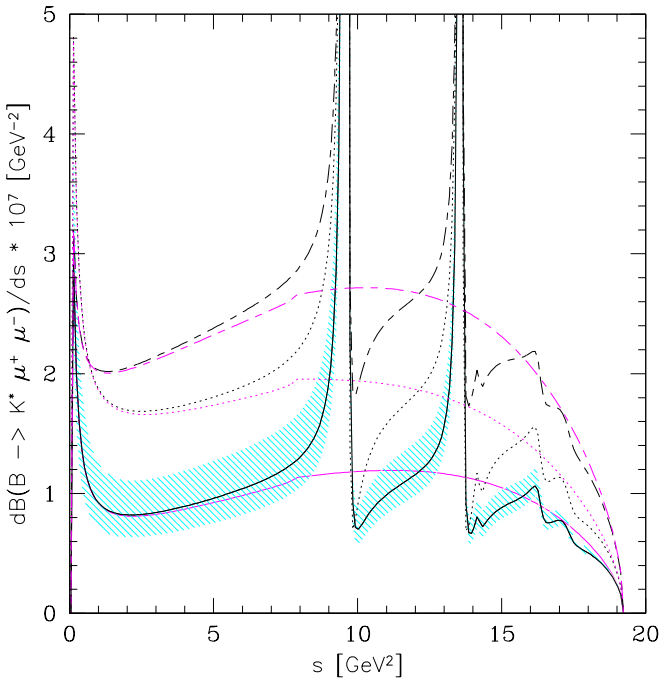

Figure 4: The dilepton invariant mass distribution in $B \to K^* \mu^+ \mu^-$ decays [3]. The solid line represents the SM and the shaded area depicts the form factor-related uncertainties. The dotted line corresponds to the SUGRA model with $R_7 = -1.2$, $R_9 = 1.03$ and $R_{10} = 1$. The long-short dashed lines correspond to an allowed point in the parameter space of the MIA-SUSY model, given by $R_7 = -0.83$, $R_9 = 0.92$ and $R_{10} = 1.61$. The corresponding pure SD spectra are shown in the lower part of the plot.

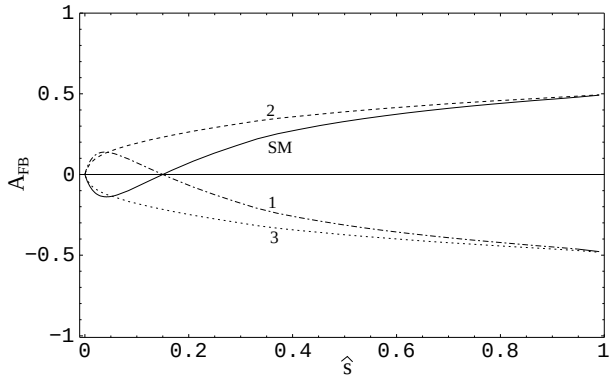

Figure 5: Differential Forward–Backward asymmetry for the decay $B \to X_s \ell^+ \ell^-$ [1]. The curves correspond to the points indicated in Fig.6.

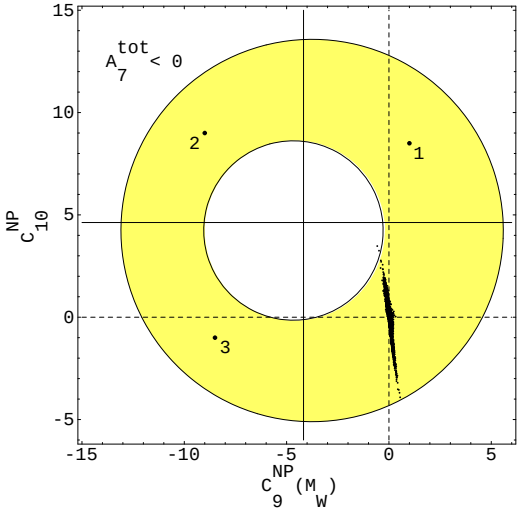

Figure 6: Superposition of all constraints for C9 and C10 [1].

CHAPTER 2

The *BABAR* experiment

The goal of the *BABAR* [20] experiment is to study B meson decays produced by the PEP-II asymmetric electron-positron collider at the Stanford Linear Accelerator Center (SLAC). Although the detector design was prompted by the primary

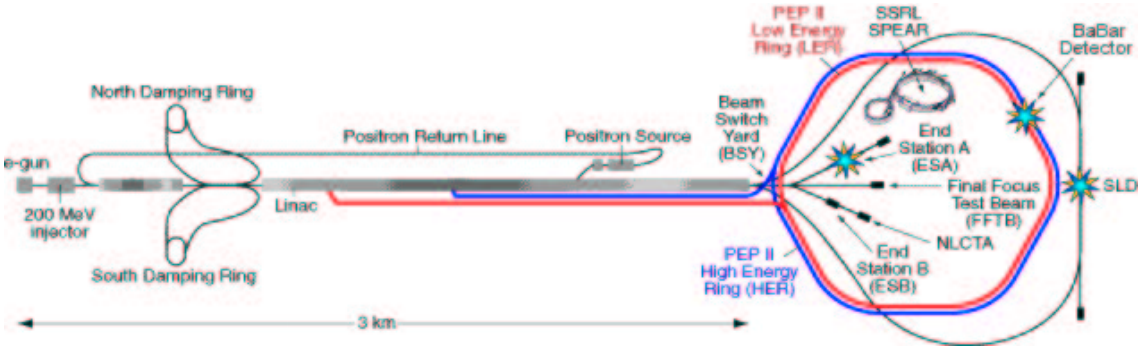

Figure 7: The SLAC Linear Collider, PEP-II and *BABAR*.

purpose of the experiment – the measurement of CP violation in the B system – it is versatile enough to make a wide range of B physics studies possible.

2.1 The PEP-II B Factory

The PEP-II B Factory is designed to operate at a luminosity of 3×10^{33} $\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and above. It consists of two storage rings for the 9.0 GeV electrons and 3.1 GeV positrons located one on top of the other. The electrons go clockwise around the old lower ring, while the positrons go anticlockwise around the newly built upper ring. Being less energetic, the positron beam is diverted into the interaction region. The beam energies are selected to produce the $\Upsilon(4S)$ resonance with a boost of

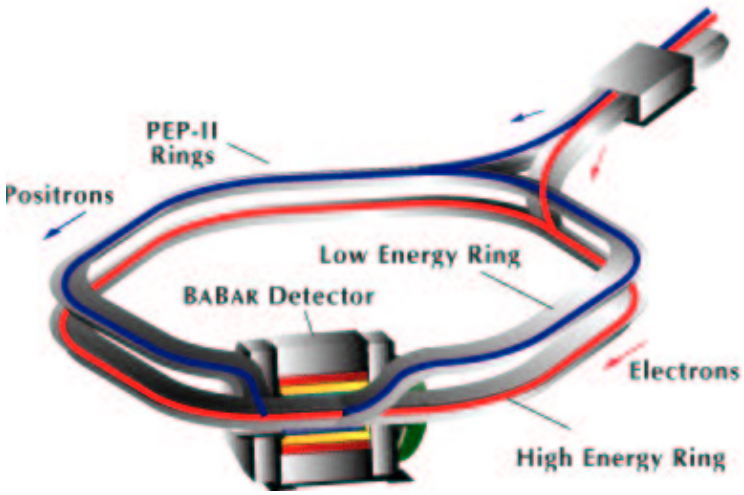

Figure 8: PEP-II

$\beta\gamma = 0.56$. This resonance decays exclusively to B^0 $\overline{B}^0$ and B^+B^- pairs and thus provides an ideal laboratory for the study of B mesons. The boost makes it possible to reconstruct the decay vertices of the two B mesons, determine their relative decay times, and thus measure the time dependence of their decay rates. Until recently PEP-II operated on a 40-50 minute cycle with the beams being replenished at the end of each fill. Due to higher backgrounds during injection data taking was halted and some detector components switched off. Since March 2004 a new *trickle*

injection method was implemented. With it, new bunches are injected continuously at a rate of up to 10 per second. The average $b\bar{b}$ production rate is thus increased by up to 50%. *BABAR* divides the data into runs, defined as periods of three hour duration or less during which beam and detector conditions are judged to be stable. While most of the data are recorded at the peak of the $\Upsilon(4S)$ resonance, about 12% are taken at a center of mass (c.m.) energy 40 MeV lower to allow for studies of non-resonant (*continuum*) background. This background is due to light quark (udsc) production the cross section for which, at the resonance, is about three times larger than the one for $\Upsilon(4S)$ (1.05 nb).

The *BABAR* coordinate system is right-handed with the z-axis coinciding with the detector principle axis and the positive direction along the electron beam. This axis is offset relative to the beam axis by about 20 mrad in the horizontal plane. The positive y-axis points upward and the positive x-axis points away from the center of the storage rings.

Several beam parameters are crucial for *BABAR* physics analyses. The mean beam energies are determined from the total magnetic bending strength. They enter in the calculation of $\Delta E = E^*_B - E^*_{beam}$ and the *energy-substituted mass* $m_{\rm ES} = \sqrt{E^{*2}_{beam} - p^{*2}_B}$, where E^*_B and p^*_B are the energy and 3-momentum of the reconstructed B candidate and the $*$ indicates that these quantities are measured in the c.m. frame. These two kinematic variables are largely uncorrelated and are widely used to separate signal from background in analyses involving exclusive B-meson decays. The use of E_{beam} rather than E_B in the calculation of $m_{\rm ES}$ results in better resolution. The absolute luminosity is derived from $e^+e^- \to e^+e^-$ and $e^+e^- \to \mu^+\mu^-$ processes. The position of the luminous region is important for decay-time-dependent analyses. In the transverse plane it is determined from the distribution of the distance of closest approach to the z-axis of tracks in well mea-

sured two-track events as a function of the azimuth ϕ. The longitudinal position is derived from the longitudinal vertex distribution of the two tracks.

2.2 The *BABAR* Detector

The *BABAR* detector was built by a large international collaboration of scientists. A longitudinal section through the detector center and an end view are shown on Figures 9 and 10, respectively. Because of the non-zero boost and in order to maximize acceptance, it is mounted asymmetrically with respect to the interaction point (IP). The detector incorporates five sub-systems, each providing complementary information. The charged particle tracking is done with a silicon vertex tracker (SVT) and a drift chamber (DCH). The SVT measures positions and angles of charged particles near the IP, while the main purpose of the DCH is momentum measurement of charged particles and providing additional particle identification information via dE/dx. The third detector component is the ring-imaging Cherenkov detector (DIRC) and is used for charged particle identification. Photons are detected in the electromagnetic calorimeter (EMC), which is designed to provide excellent energy and angular resolution. The muon and neutral hadron part of the detector (IFR) is an integral part of the magnet flux return. More detailed information about each of these subsystems is provided in the following subsections.

In order to achieve the desired charged particle momentum resolution, the detector is immersed in a magnetic field of 1.5T. The system consists of a superconducting solenoid, a flux return, a conventional water cooled field compensating coil, to reduce magnetic flux leakage at the backward end to levels acceptable for the PEP-II components and the DIRC photo-multipliers, and a cryostat for the super-

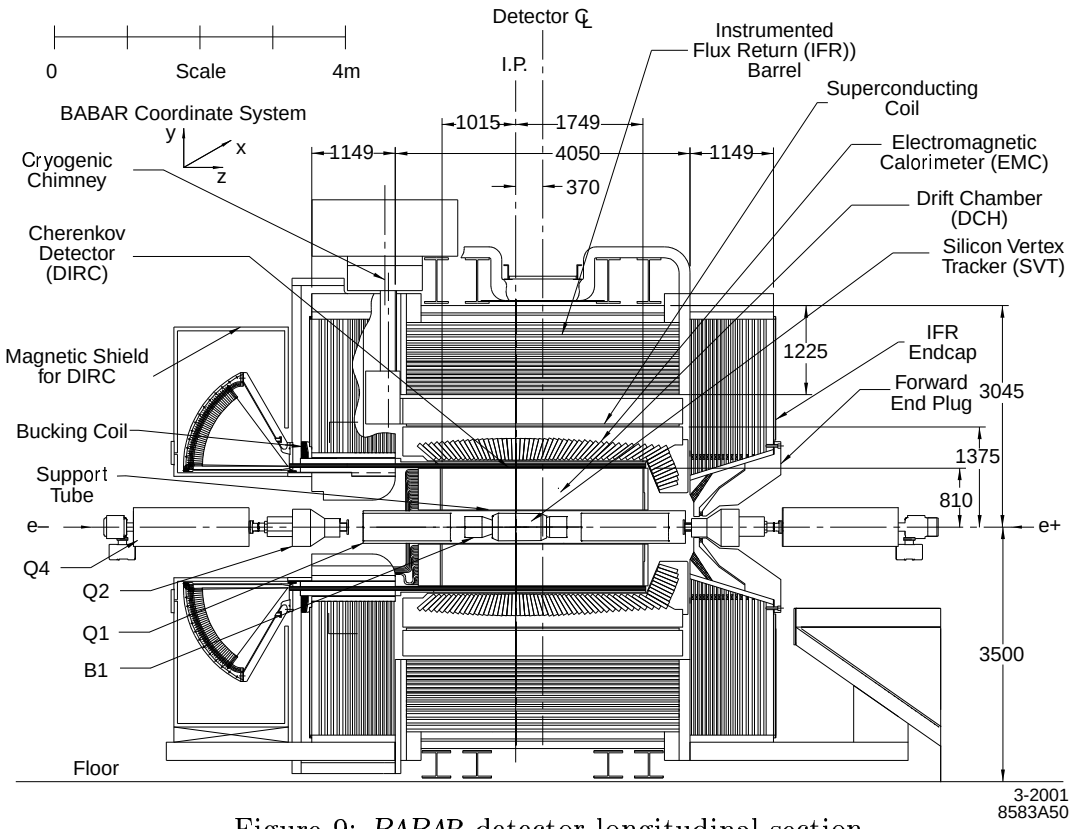

Figure 9: *BABAR* detector longitudinal section.

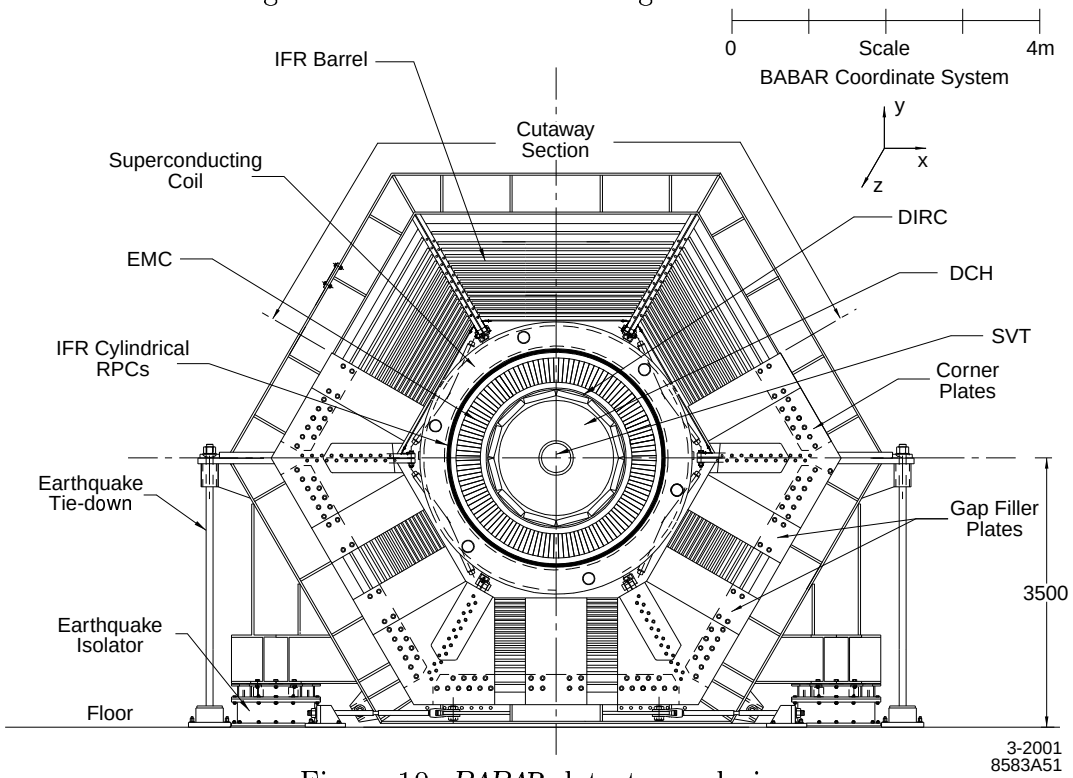

Figure 10: *BABAR* detector end view.

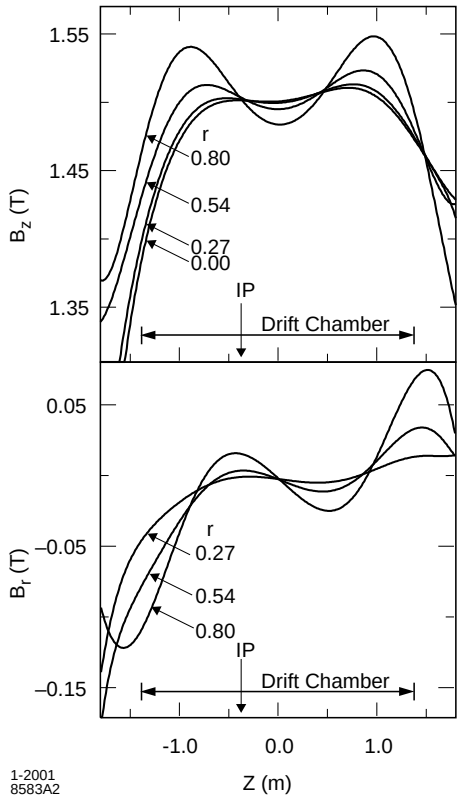

Figure 11: The magnetic field components B_z and B_r as a function of z for various radial distances r (in m).

conducting coil. The direction of the field is along the detector's z axis, which is also the approximate direction of the electron beam. To facilitate track fitting, the field is uniform to within a few percent in the tracking volume, see Fig. 11. The cryostat uses helium to indirectly cool the superconducting solenoid to an operating temperature of 4.5K.

2.2.1 SVT

The main goal of the experiment – measurement of time-dependent CPV – depends greatly on the SVT and imposes a stringent requirement on the mean vertex resolution of a fully reconstructed B candidate to be better than 80 μm along the z-axis. The resolution in the x-y plane needs to be approximately 100 μm to reconstruct $B^0 \to D^+D^-$ decays for example. In addition many B decay products have low transverse momenta $p_t < 120 MeV/c$, the minimum that can be measured reliably in the DCH alone. In particular for the identification of slow pions from D^* decays, a tracking efficiency of more than 70% is desirable for tracks with transverse momenta ranging from 50 to $120 MeV/c$. Additional requirements on reliability are imposed due to the inaccessibility and the large radiation dosage the SVT must withstand. In order to meet these objectives the SVT incorporates five layers of double-sided silicon strip detectors, Figures 12 and 13. The first three layers are mounted close to the beam pipe and provide angle and vertex position information. The outer two layers are at larger radii and are used to link the SVT and DCH tracks as well as for standalone tracking for particles with $p_t < 120 MeV/c$. The strips on the opposite sides of each sensor are orthogonal to

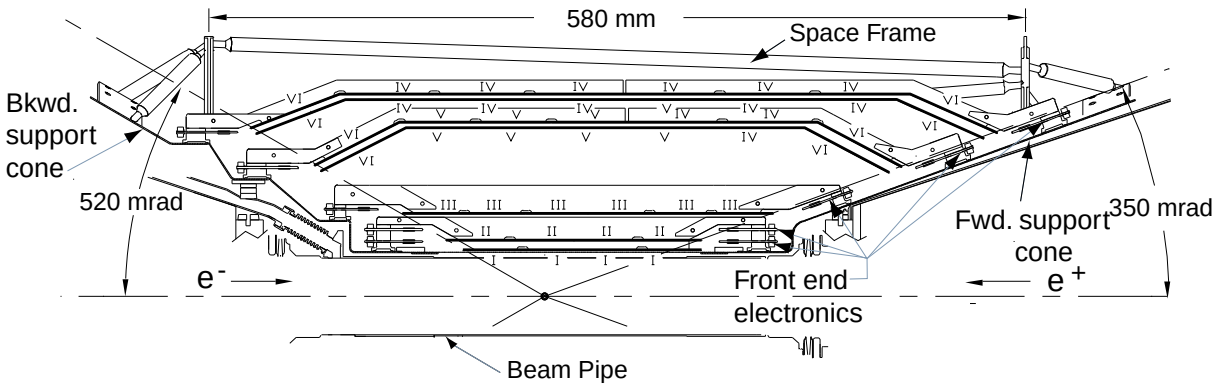

Figure 12: Schematic view of SVT: longitudinal section.

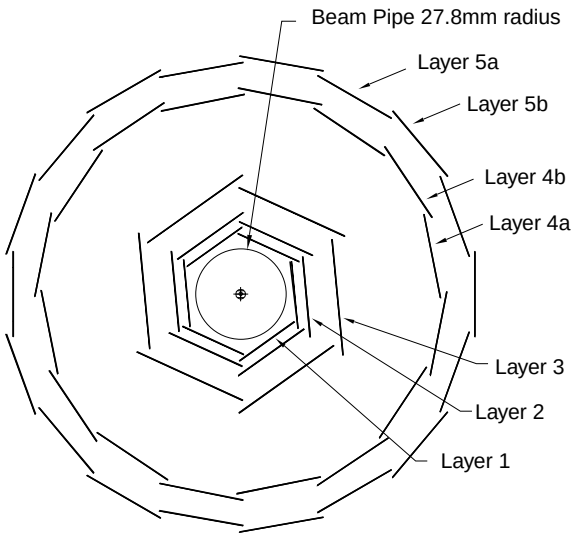

Figure 13: Schematic view of SVT: transverse section.

each other, the ϕ measuring strips run parallel to the beam and the z measuring strips are transverse to the beam axis. To maximize acceptance while minimizing detector material the outer two layers have an arch design. The efficiency of each layer can be measured by comparing the number of associated hits to the number of tracks crossing the active area of the layer and is 97%. The hit resolution is determined from high-momentum tracks in two prong events by measuring the distance, in the plane of the sensor, between the track trajectory and the hit, see Fig. 14.

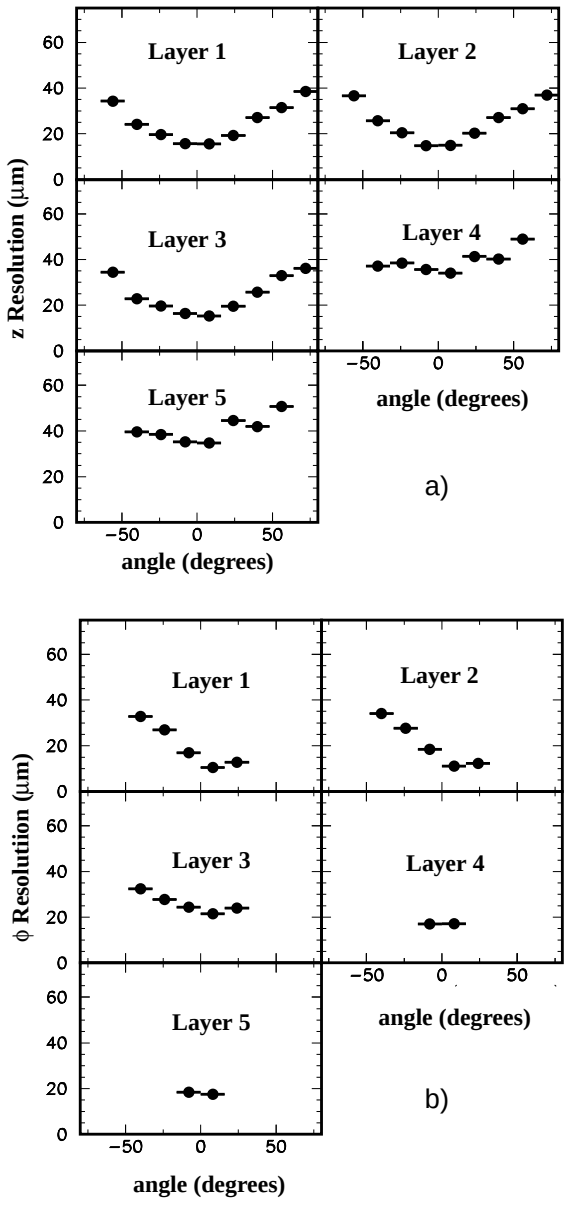

Figure 14: SVT hit resolution in the a) z and b) ϕ coordinate in microns, plotted as a function of track incident angle in degrees.

2.2.2 DCH

The DCH detects charged particles and measures their momenta with high precision. In addition it helps π/K separation by measuring the ionization loss dE/dx. In the barrel this information is complementary to the one from the DIRC, but beyond the DIRC acceptance the DCH is the only device providing some discrimination. The reconstruction of decay vertices outside of the SVT relies solely on the DCH as well. For this purpose, in addition to transverse momenta and positions, the chamber should be capable of measuring longitudinal positions with a resolution of 1mm. The DCH consists of 40 layers of cells, filled with a helium-based gas mixture, see Fig. 15. The cells are close to hexagonal in shape and consist of one sense wire surrounded by six field wires. A positive voltage of $1960V$ is applied to the sense wires while the field wires are at ground potential. In order to obtain longitudinal information, the wires in 24 of the layers are at small angles with respect to the z-axis. Multiple scattering is minimized by the choice of gas and low-mass aluminum wires. The mixture of helium and isobutane is provided at a stable 80/20 ratio by a gas recirculation system at a constant pressure. During normal operation a complete gas recirculation is done in six hours, and a full volume of fresh gas is added every 36 hours. In order to prevent electrical discharge a relatively high level of humidity is maintained by passing a controlled fraction of the gas through a watter bubbler. Oxygen is removed from the gas mixture with a palladium catalytic filter. The readout electronics are mounted on the back plate of the chamber to minimize material in front of the calorimeter endcap.

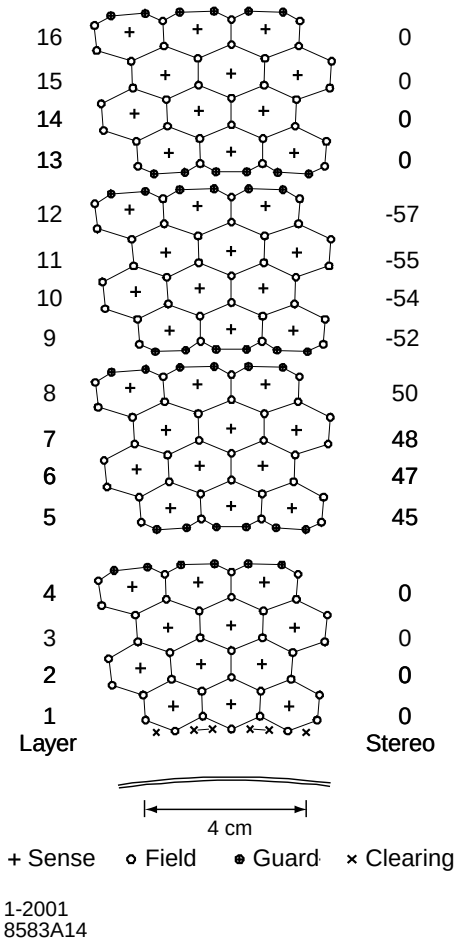

Figure 15: Schematic layout of drift cells for the four innermost superlayers. Lines have been added between field wires to aid in visualization of the cell boundaries. The numbers on the right side give the stereo angles (mrad) of sense wires in each layer.

2.2.3 DIRC

The ring-imaging Cherenkov detector utilizes the property of charged particles moving faster than the speed of light in a medium, to emit Cherenkov light. The light is emitted in a cone around the particle direction and the cone angle with respect to the axis, known as Cherenkov angle θ_c, depends on the particle velocity v and the medium index of refraction n, $\cos\theta_c = 1/n\beta$ ($\beta = v/c$, c = velocity of light in vacuum). The radiator material of the DIRC is synthetic, fused silica in the

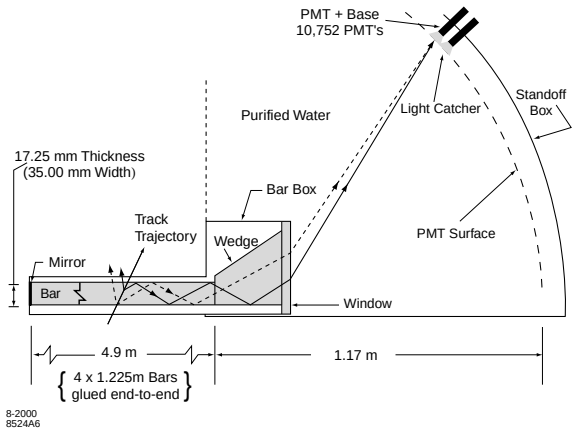

Figure 16: Schematics of the DIRC.

form of 4.9-m long, thin bars with rectangular cross section. The bars are optically isolated with air gaps from one-another and also serve as light pipes transporting the light by total internal reflection, while preserving the Cherenkov angle, to an array of photo-multiplier tubes (PMTs), where the angle is measured, see Fig. 16. Along with the DCH information of the track's angle and momentum, the Cherenkov angle allows to determine the particle velocity and hence identity. It is imperative for the total internal reflection that no condensation occurs on the bars. This is achieved by maintaining a flow of filtered nitrogen gas. The choice of fused silica as the bar material was based on a combination of properties among which resistance to ionizing radiation, large index of refraction, and because it allows excellent optical finish on the surfaces. To avoid instrumenting both ends of the bars with photon detectors and minimize interference with other detector components in the forward region, where particles are preferentially produced, mirrors are placed at the forward ends, while at the back the bars merge into a water-filled tank, called the *standoff box*. The photons are detected by PMTs arranged radially on the rear tank surface, at about 1.2m from the bar end. Purified water is used in the standoff box due to its index of refraction being close to that of fused silica. The distance

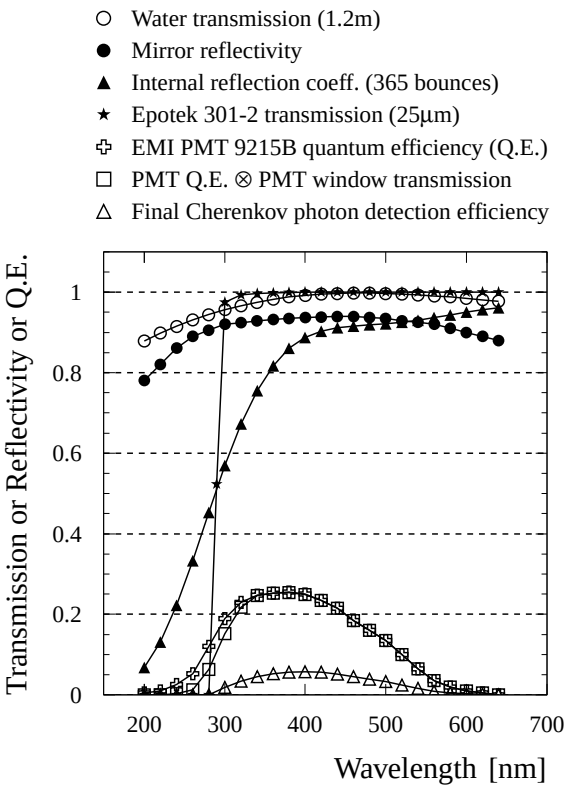

Figure 17: Transmission, reflectivity and quantum efficiency for various components of the DIRC as a function of wavelength for a $\beta = 1$ particle at normal incidence to the center of a bar

between the bars and the PMTs along with their sizes gives a Cherenkov angle resolution of about 7 mrad. With transmission dispersions included, the overall single photon resolution is estimated to be $\sim$ 10 mrad. The photon detection efficiency of the DIRC depends on the optical properties of the components as well as the PMTs specifics. The wavelength dependent efficiency of some of the components is shown on Fig. 17, where the data points are for particles entering the center of the bar at 90°. In order to maintain good water transparency at wavelengths down to 300 nm, the water must be ultra-pure, de-ionized, de-gassed and free of bacteria. For this purpose several systems are used including filters, a reverse osmosis unit and a UV lamp. Water quality is continuously monitored by checking its resistivity, pH-value, temperature and transparency. For the latter,

three lasers with wavelengths 266nm, 325nm and 442nm are used. The entire volume can be recirculated up to four times a day. Potential water seal failure is detected by a dedicated system with a capability of draining the water in 12 minutes.

2.2.4 EMC

The electromagnetic calorimeter (EMC) is used for energy and momentum reconstruction of photons as well as electron identification. It must provide excellent energy and angular resolution over a wide energy range. Photons with energies up to 9 GeV need to be measured from QED processes, like $e^+e^- \to e^+e^-(\gamma)$ and $e^+e^- \to \gamma\gamma$, for calibration and luminosity determination. On the other hand photons with energies as low as 20 MeV are needed in order to reconstruct B-meson decays containing low-energy π^0s and η^0s. The requirement on the energy resolution comes from rare B decays containing π^0s and is on the order of 1-2%, since at energies lower than 2GeV it dominates the mass resolution of the reconstructed π^0s. At higher energies, however, the angular resolution plays a bigger role, which imposes a requirement for the latter to be roughly few mrads. To meet these objectives the EMC consists of appropriately sized thallium-doped cesium iodide CsI crystals arranged projectively with respect to the IP, see Fig. 18. To maximize acceptance in the forward region, a conical endcap is added to the cylindrical barrel extending the polar angle coverage from 15.8° to 141.8°. The crystals have trapezoidal cross section and increase in length from 29.6cm in the backward to 32.4cm in the forward direction to fully contain the electromagnetic showers from increasingly higher energy particles. The energy of the electromagnetic showers is measured by reading the scintillation light. For that purpose two silicon PIN diodes

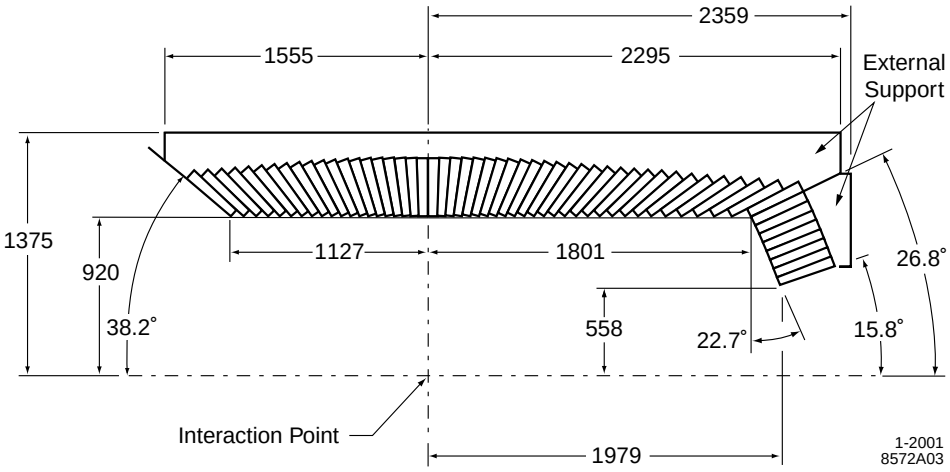

Figure 18: A longitudinal cross section of the EMC (top half). The detector is axially symmetric around the z-axis. All dimensions are given in mm.

are glued to the rear face of the crystals and the surrounding area is covered with white reflective paint. The surfaces are polished to facilitate internal reflection as the crystals act not only as a scintillation medium but also as light guides. Transmitted light is recovered in part by wrapping the crystals in defuse white reflectors. In order to achieve uniform light yield along the crystal each was individually measured by placing a highly collimated radioactive source at 20 points along its length. Fine adjustments were made by selectively roughening or polishing the surface to reduce or increase its reflectivity. The crystals were then wrapped in aluminum foil which was grounded to provide Faraday shield and covered with a layer of mylar for electrical isolation. Since the photodiode leakage current depends strongly on the temperature and the diode-crystal joints could experience stress under temperature change due to different termal expansion, an effective cooling system is imperative. The system consists of Fluorinert (polychlorotrifluoro-ethylene) and water chillers for the preamplifiers and the digitizing electronics. The light yield

of the individual crystals varies significantly and also changes with time under the impact of beam radiation. The front of the crystal is most strongly affected resulting in increased attenuation of the transmitted scintillation light. Individual crystal calibration must therefore be performed at different energies, corresponding to different shower penetration, to compensate the effects of the radiation damage. At 6.12 MeV radioactive photon source provides absolute calibration at lower energies, while at higher energies (3-9GeV) the relation between polar angle and energy of $e^{\pm}$ from Bhabha events is used. The information obtained from the EMC is processed by reconstruction algorithms which separate the electromagnetic shower, that typically spreads over many adjacent crystals, into *clusters* and *bumps* which are associated with a particle. Other algorithms are responsible for reconstructing the angle by employing various weighting schemes on the crystals based on the deposited energy.

2.2.5 IFR

The IFR is designed for muon and neutral hadron identification over a large range of momenta and angles. The steel flux return is used as a muon filter and hadron absorber. It is segmented into layers ranging from 2 to 10cm in thickness with resistive plate chambers (RPCs) inserted in the gaps between the layers. The RPCs detect streamers from minimum-ionizing particles via two-coordinate capacitance readout strips, see Fig. 19. There are 19 RPC layers in the barrel and 18 in the endcaps. An RPC schematic is shown in Fig. 20. It consists of two bakelite sheets separated by a 2mm gap. The external bakelite surfaces are coated with graphite and an 8kV voltage is applied between them. The signals are red capacitively, on both sides of the gap, by X and Y aluminum strips. The

strips are separated by the ground aluminum plane with a foam sheets. The RPCs operate with a non-flammable gas mixture of argon, Freon and isobutane. The gas is distributed among the chambers at a constant pressure of 6.5 Torr through a parallel manifold system. The IR experimental hall does not have temperature regulation and during the first summer of operation the maximum hall temperature frequently exceeded 31° C causing the dark current in many RPC modules to exceed the capabilities of the high-voltage system. Subsequently water cooling was installed stabilizing the temperature at 20-21° C in the barrel, 22° C in the backward and 24° C in the forward endcaps. During operation at high temperatures many RPCs ($> 50\%$) showed some efficiency reduction compared to earlier measurements. After the instalation of the cooling system some of them continued to deteriorate, while others remained stable. The cause of the efficiency loss remains under investigation. Recently 24 modules in the endcaps have been replaced by new RPCs. After evaluation of the results with these new RPCs, plans on further IFR improvement can be made.

2.3 Detector Readout

The detector readout is achieved with a number of tightly coupled hardware and software systems, see Fig. 21. The front-end electronics for all detector subsystems is mounted on the detector to minimize noise pickup. It performs initial signal amplification and digitization as well as data transfer via optical fibers to the data acquisition system. A two-level trigger is designed to reject beam-induced background. The first level (L1) is implemented in hardware and relies on inputs from the DCH, EMC and IFR. The level 3 (L3) trigger is implemented in software and selects events to be stored for further processing. A provisional second level

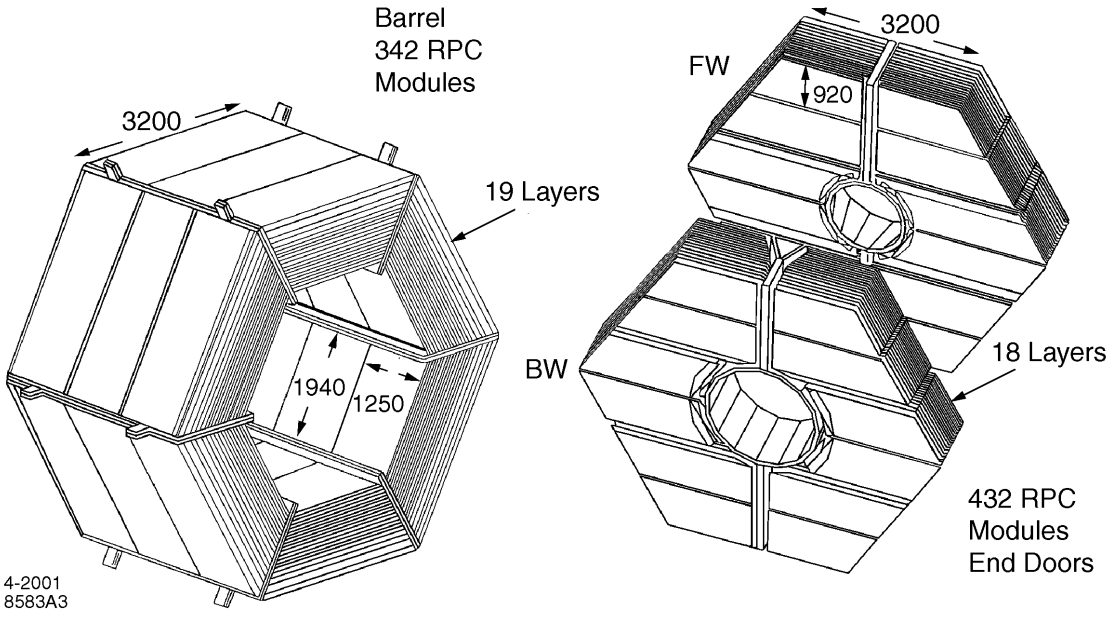

Figure 19: Overview of the IFR: Barrel sectors and forward (FW) and backward (BW) end doors; the shape of the RPC modules and their dimensions are indicated.

trigger is possible should severe conditions require additional sophistication. The online prompt reconstruction (OPR) reads the raw data after L3, selects physics events and writes the results to the event store.

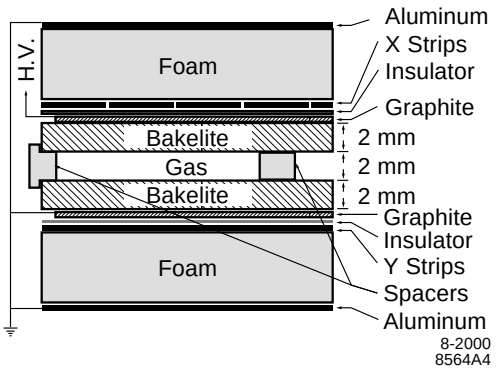

Figure 20: RPC cross section with the schematics of the high voltage connection.

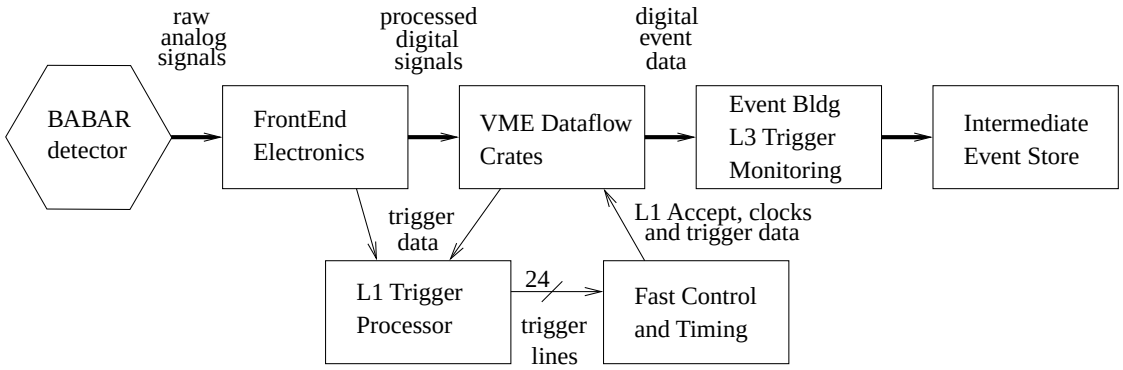

Figure 21: Schematic diagram of the data acquisition.

CHAPTER 3

Particle Identification

In the inclusive $B \to X_s \ell^+\ell^-$ analysis we adopt a sum of exclusive modes technique. B candidates are reconstructed from an oposite charge e or μ pair, one $K^\pm$ or K_s and up to three pions, of which at most one π^0. Thus all particle identification (PID) systems are involved in the analysis and due to the rareness of the process, outstanding performance is necessary. The PID criteria used are outlined in the following sections.

3.1 Electron Id

For the electrons we use the *Likelihood Selector*. A standard *very loose* set of cuts is applied to all charged tracks in the event:

- $0.5 < E/p < 5$, where E and p are the EMC-measured energy and the DCH-measured momentum of the associated track, respectively.
- Number of crystals associated with the track > 3
- $500 < dE/dx$ (the ionizing energy loss in DCH) < 1000

A likelihood L for each particle hypothesis is then calculated based on the variables:

- E/p
- Lateral moment of the EMC cluster: LAT $= \frac{RM}{RM+25(E_1+E_2)}$. Here E_1 and E_2 are the energies of the two most energetic crystals and $RM = \sum(r_i/1.08)^2 E_i$, where the sum goes over all but the two most energetic crystals and r is the distance between a point located 12.5 cm from the crystal surface along its axis and the line connecting the cluster centroid (also 12.5 cm from the surface) and the origin of the coordinate system. For clusters with less than three crystals LAT is 0.
- $\triangle\phi =$ charge$(\phi_{AtEMC} - \phi_{cluster})$, where ϕ_{AtEMC} is the ϕ coordinate of the track intersection point with the EMC and $\phi_{cluster}$ is the ϕ coordinate of the cluster centroid projected to the crystal face.
- dE/dx (DCH)
- The Cherenkov angle from DIRC: θ_C

A likelihood fraction is defined as

$$f = \frac{P_e L(e)}{P_e L(e) + P_\pi L(\pi) + P_K L(K) + P_p L(p)}, \tag{3.1}$$

where P_e and P_K are equal to unity, $P_\pi = 5$ and $P_p = 0.2$. These values are prompted by the average multiplicity of these particles in an event. Electron candidates are selected by using the default *tight* selection corresponding to a cut of $f > 0.95$. Plots of the efficiency and the pion missid rate as a function of the momentum are shown on Figs. 22 and 23.

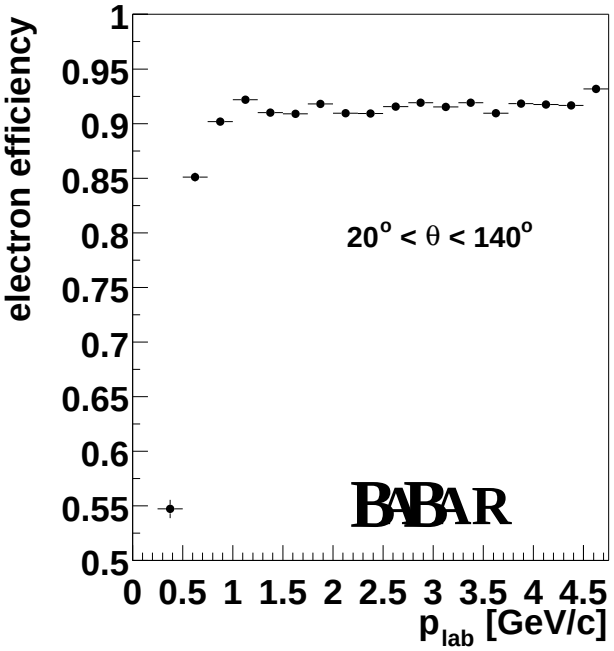

Figure 22: Electron Id efficiency vs momentum.

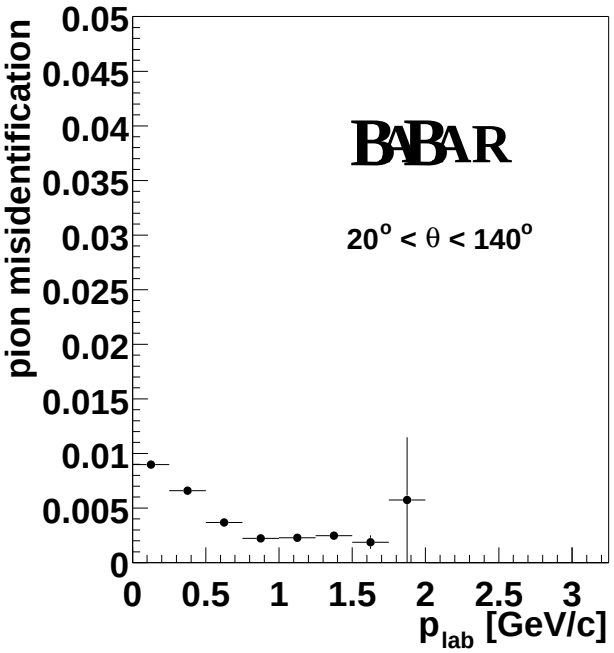

Figure 23: Pion miss-Id vs momentum.

3.2 Muon Id

The muons are identified with the following variables:

1. Calorimeter energy: E_{cal}
2. Number of IFR layers hit: N_L
3. A boolean variable which is true when a cluster has a hit in the inner RPC: I_h
4. The first IFR hit layer in the cluster: F_h
5. The last IFR hit layer in the cluster: L_h
6. The number of interaction lengths traversed by the track in the detector: λ
7. The number of interaction lengths in the detector which a muon with the same momentum is expected to traverse: λ_{exp}
8. The χ^2 of the IFR cluster hits with the extrapolated track from DCH: χ^2_{trk}
9. The χ^2 of the IFR track fit: χ^2_{fit}
10. The average multiplicity of hit strips per layer: $\bar{m}$
11. The standard deviation of $\bar{m}$: $\sigma_{\bar{m}}$

We use the *very tight selector* corresponding to cuts:

- $0.05 < E_{cal} < 0.4$
- $N_L \geq 2$
- $\triangle\lambda = \lambda_{exp} - \lambda < 0.8$

- $\lambda > 2.2$
- $\chi^2_{trk} < 5$
- $\chi^2_{fit} < 3$
- $T_c = \left[\begin{array}{l} \frac{N_L}{L_h - F_h + 1} \quad \text{if } \mathrm{I_h} = \text{false} \\ \frac{N_L}{L_h - F_h} \quad \text{if } \mathrm{I_h} = \text{true} \end{array} \right] > 0.34$
- $\bar{m} < 8$
- $\sigma_{\bar{m}} < 4$

Tha plateau efficiency, starting from around 2 GeV, for the very tight selector is about 70.3% with missidentification rate of 2.3%.

3.3 Charged Kaon Id

Charged kaons are identified with the *PidKaonMicroSelector*. The selector uses a neural network, taking inputs from the SVT, the DCH and the DIRC. These inputs are:

- The momentum of the candidate: p
- $\frac{L_K^{SVT}}{L_K^{SVT} + L_\pi^{SVT}}$
- $\frac{L_K^{DCH}}{L_K^{DCH} + L_\pi^{DCH}}$
- $\frac{L_K^{DIRC}}{L_K^{DIRC} + L_\pi^{DIRC}}$

where L_K^{system} and L_π^{system} are likelihoods for K and π hypothesis, respectively, for the corresponding detector system. In this analysis we use the *KMicroLoose* selection, corresponding to a cut of 0.5 on the neural network output. The efficiency

and pion miss-Id for the KMicroLoose selector are shown in Figs. 24 and 25, the red and blue dots correspond to an older and newer code release version.

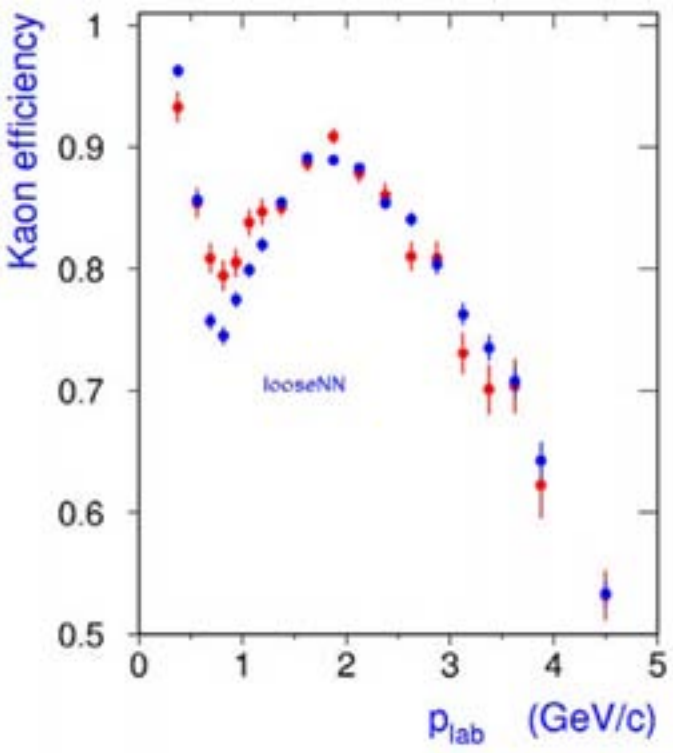

Figure 24: Kaon Id efficiency vs momentum.

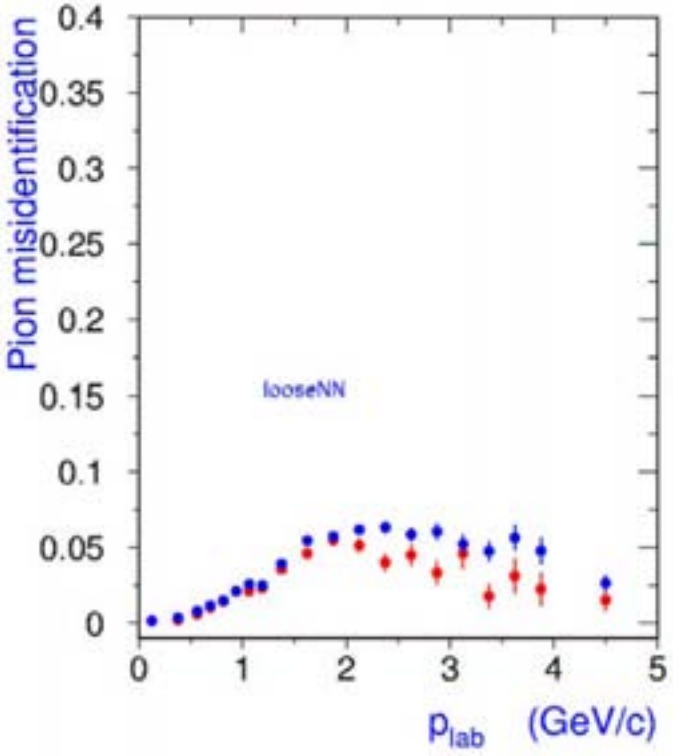

Figure 25: Pion miss-Id vs momentum.

3.4 Neutral Kaon Id

We only consider Ks decaying into two charged pions. The candidates are taken from the *KsDefault* list and are made out of oppositely charged tracks with fitted mass within 0.025 GeV of the PDG Ks mass, 497.7 MeV. In order to purify the sample three additional cuts are applied:

- A tighter cut on the reconstructed Ks mass: $|m(\pi\pi) - m(Ks)| < 11.2$ MeV
- $L_{dec} > 2$ mm, where L_{dec} is the decay length calculated as the distance between the Ks decay point and the event primary vertex
- $\cos\alpha > 0.99$, where α is the angle between the Ks momentum and the primary vertex – decay point line

3.5 Charged Pion Id

For the charged pions we use particles from the *GoodTracksLoose* list which are not consistent with being *KMicroVeryTight* and *eMicroVeryTight*. The pion candidates in the GoodTracksLoose list satisfy the following requirements:

- Min transverse momentum: 0.1 GeV
- Max momentum: 10 GeV
- Min number of DCH hits: 12
- Max distance of closest approach (DOCA) in XY plane: 1.5 cm
- Min Z DOCA: -10 cm

- Max Z DOCA: 10 cm

The kaon candidates in the KMicroVeryTight list correspond to a cut of 0.68 on the neural network output described in Sec. 3.3. The eMicroVeryTight lists consists of candidates which satisfy the requirements:

- $540 < dE/dx < 860$
- EMC shower of at least 3 crystals
- $0.89 < E/p < 1.2$
- $0.1 < LAT < 0.6$
- $-10 < A_{42} < 0.11$, where A_{42} is the modulus of the Zernike moment of order (4,2) [21].
- At least three photons detected in the DIRC and Cherenkov angle consistent with electron hypothesis within 3 σ
- DCH track matched to the EMC cluster

3.6 Neutral Pion Id

The π^0 particles are composed of photon candidates identified by EMC bumps unmatched to any track. The bumps must have $LAT < 0.8$ and energy $E > 0.05$ GeV. In order to compute the photons 3-momenta their origin is assumed to be at the primary event vertex. In addition we apply cuts:

- $E(\pi^0) > 0.4$ GeV
- $|m(\gamma\gamma) - m(\pi 0)| < 10$ MeV

CHAPTER 4

Analysis overview

Our aim is to measure the overall inclusive branching fraction for the $B \to X_s\, \ell^+\ell^-$ decay, considering electron and muon modes, as well as the branching fractions in several dilepton and hadronic mass regions. The inclusive branching fraction measurement for the electron modes is done for $m_{e^+e^-} > 0.2$ GeV. The results we present are based on a sample of 81.9 fb^{-1} of on-resonance data. We also used 9.6 fb^{-1} of off-resonance data for continuum background studies.

The sum of exclusive modes technique we have adopted allows us to use the strong background suppression power of the two kinematic variables ΔE and $m_{\rm ES}$ described in Section 2.1. However, it also introduces a significant hadronization model dependence, since the signal reconstruction efficiency varies with the particle multiplicity of the hadronic system (X_s) in the final state.

4.1 Monte Carlo generator

We use Monte Carlo (MC) generated signal and background events to estimate the final signal reconstruction efficiency, tune the background suppression cuts and asses the systematic errors. The overall signal model is a combination of three separate Monte Carlo generators: $B \to K\ell^+\ell^-$, $B \to K^*\ell^+\ell^-$ for $m_{Xs} < 1.1\ GeV$

and $b \to s\ell^+\ell^-$ for $m_{Xs} \geq 1.1\ GeV$. The exclusive $B \to K\ell^+\ell^-$ and $B \to K^*\ell^+\ell^-$ decays are generated according to the model by Ali, Ball, Handoki and Hiller [3], where the required form factors are computed using light cone QCD sum rules. For the inclusive $b \to s\ell^+\ell^-$ decays we set the spectator-quark mass m_q to zero and the Fermi motion parameter p_F to 410 MeV, as suggested by the CLEO $B \to X_s\gamma$ data [22]. All three generators use NNLO Wilson coefficients, following Ali et al. [1]. The three pieces are combined according to the branching fraction predictions for $m_{ll} > 0.2$ GeV, see end of Sec. 1.3.2, multiplied by a factor reflecting the difference between the lifetimes of the B^+ and B^0 mesons: $\tau(B^+)/\tau(B^0) = 1.083 \pm 0.017$ [2]. Since the exclusive predictions in Table 2 are given for B^0 decays, we multiply the corresponding branching fractions for B^+ decays by 1.083. The predictions for the inclusive decays, however, are averaged over charged and neutral B mesons, thus we multiply the branching fraction for $b \to s\ell^+\ell^-$ by 1.04 for charged B mesons and 0.96 for neutral B mesons.

Figures 26 through 29 show the hadronic and dilepton mass distributions for the combined MC signal model in the electron and muon channels. The fractions of final states in different hadronic topologies are given in Table 3.

The various signal and background MC samples used in this analysis are summarized in Table 4.

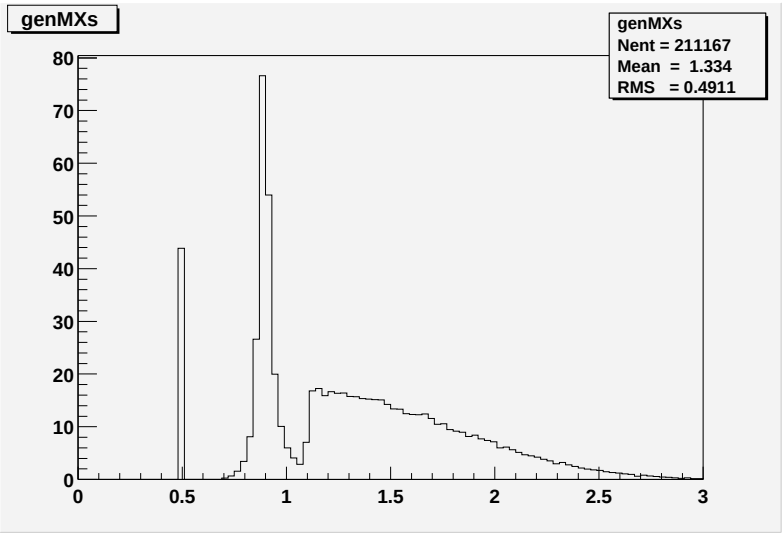

Figure 26: Hadronic mass distribution of the signal model in the electron channel. The entries are normalized to the expected number in 81.9 fb^{-1}.

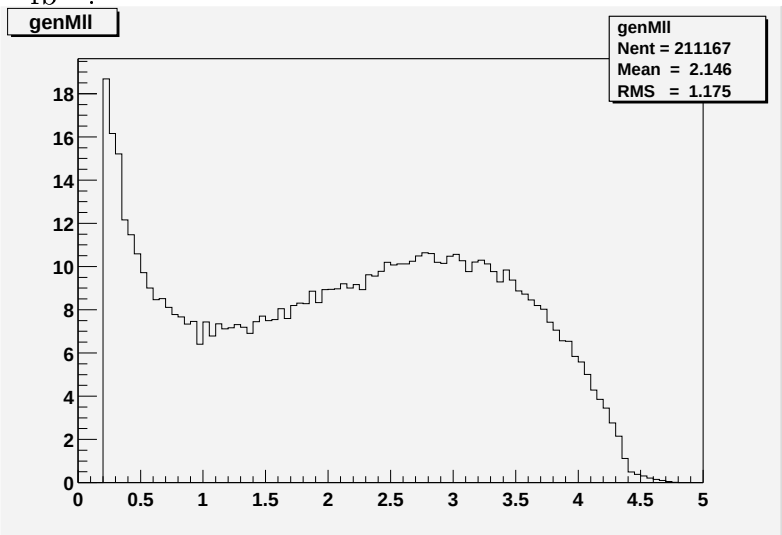

Figure 27: Dilepton mass distribution of the signal model in the electron channel for $m_{ll} > 0.2$ GeV. The entries are normalized to the expected number in 81.9 fb^{-1}.

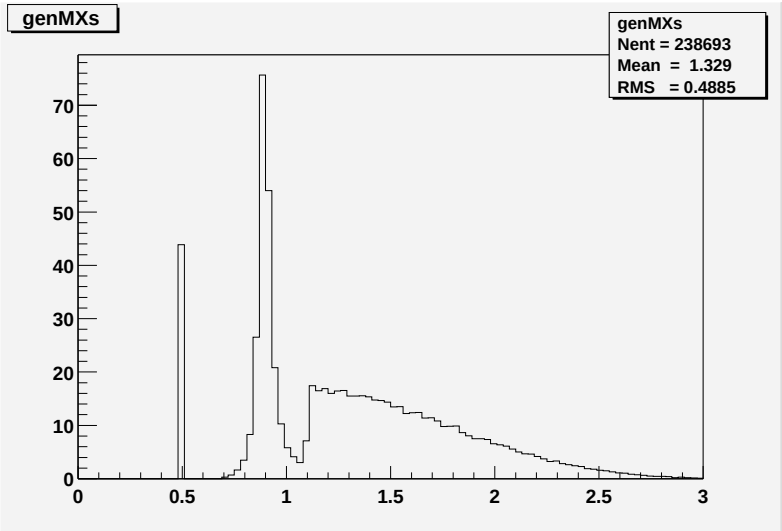

Figure 28: Hadronic mass distribution of the signal model in the muon channel. The entries are normalized to the expected number in 81.9 fb^{-1}.

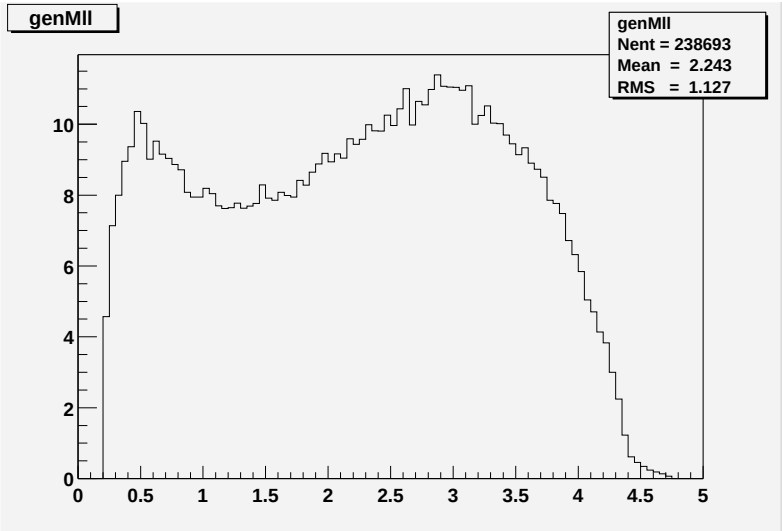

Figure 29: Dilepton mass distribution of the signal model in the muon channel. The entries are normalized to the expected number in 81.9 fb^{-1}.

Table 3: Fractions of B decays in different hadronic final states for the signal model.

X_s topology	$B \to X_s e^+ e^-$(with $m_{ll} > 0.2$ GeV)	$B \to X_s \mu^+ \mu^-$
K	0.044	0.044
$K\pi^0$	0.071	0.072
$K\pi$	0.125	0.126
$K\pi\pi^0$	0.056	0.053
$K\pi\pi$	0.049	0.051
$K\pi\pi\pi^0$	0.036	0.036
$K\pi\pi\pi$	0.010	0.010
K_s	0.021	0.021
$K_s\pi^0$	0.032	0.032
$K_s\pi$	0.069	0.069
$K_s\pi\pi^0$	0.031	0.030
$K_s\pi\pi$	0.023	0.022
$K_s\pi\pi\pi^0$	0.015	0.016
$K_s\pi\pi\pi$	0.006	0.006

4.2 Analysis roadmap

The analysis consists of several stages, which are outlined below. The procedure is motivated by the fact that the signal yield is extremely small, which necessitates background suppression better than $1/10^7$ in order to obtain a final signal-to-background ratio of about 1/1.

1. Events are required to have a good primary vertex and pass a loose *skim* (see Sec. 5.1) intended to select hadronic events with at least two charged leptons.

2. In each event B candidates are formed by first picking a lepton pair, and then adding one $K^\pm$ or K_s and up to three pions, of which at most one can be a π^0. In this way fourteen different modes are considered: K, Kπ^0, Kπ, K$\pi\pi^0$, K$\pi\pi$, K$\pi\pi\pi^0$, K$\pi\pi\pi$, K_s, $K_s\pi^0$, $K_s\pi$, $K_s\pi\pi^0$, $K_s\pi\pi$, $K_s\pi\pi\pi^0$, $K_s\pi\pi\pi$.

3. The number of B candidates per event is reduced by applying a set of loose

Table 4: Monte Carlo samples used in the analysis.

Sample	# events
$B \to K^+ e^+ e^-$	54000
$B \to K_s e^+ e^-$	51000
$B \to K^* e^+ e^-$ with $m_{Xs} < 1.1$ GeV	46828
$B \to X_s e^+ e^-$ with $m_{Xs} \geq 1.1$ GeV	59658
$B \to K^+ \mu^+ \mu^-$	51000
$B \to K_s \mu^+ \mu^-$	52000
$B \to K^* \mu^+ \mu^-$ with $m_{Xs} < 1.1$ GeV	49703
$B \to X_s \mu^+ \mu^-$ with $m_{Xs} \geq 1.1$ GeV	52597
$B \to J/\psi X$	390000
$B \to \psi(2S) X$	137000
Generic uds	96.3×10^6
Generic $c\overline{c}$	60.7×10^6
Generic $B^+ B^-$	75.2×10^6
Generic $B^0 \overline{B}^0$	75.1×10^6

preselection cuts. A signal likelihood is then calculated for each candidate passing the preselection to pick the best candidate in the event.

4. The best candidate is subjected to a series of background rejection *postselection* cuts. These include cuts on the dilepton invariant mass to suppress backgrounds from $B \to J/\psi X_s$, $B \to \psi(2S) X_s$ and $B \to X_s \gamma$ with $\gamma \to e^+ e^-$, which are peaking in m_{ES}, as well as cuts reducing the combinatorial background from $B\bar{B}$ and continuum events.

5. The signal yield is extracted with an extended unbinned maximum likelihood fit to the m_{ES} distribution.

CHAPTER 5

Signal reconstruction and background suppression

5.1 Event selection

In the first selection stage, events are required to have a good primary vertex (converging fit) and pass the following skim criteria, generally satisfied by multihadron, but not by continuum events:

1. The number of charged tracks is greater than three.

2. The ratio between the second and zeroth order Fox-Wolfram moments [23], R_2, calculated using charged tracks and neutral calorimeter clusters must be less than 0.5. The value of R_2 ranges from 0 to 1 and is indicative of the jettiness of the event. Values close to 1 indicate a jetty event, while values closer to 0 correspond to a more spherical event.

3. Either the DCH or the EMC L3 trigger fired to guarantee an event triggered by an electron positron collision rather than a cosmic event.

4. The event contains at least two leptons, which could be either electrons or muons satisfying loose particle identification criteria and lab frame momenta greater than 0.5 GeV and 0.8 GeV, respectively.

The efficiency for the above requirements is 95% for both $B \to X_s e^+ e^-$ and $B \to X_s \mu^+ \mu^-$ signal MC.

5.2 B candidate selection

In the second selection stage, the e^+e^- or $\mu^+\mu^-$ pair, if any, with the highest lab frame energy and converging vertex is picked. The signal efficiency for this stage is 93.5% for the electron and 84.8% for the muon modes. Because of their relatively low mass, the electrons are likely to loose significant amount of their energy via Bremsstrahlung. We perform a Bremsstrahlung recovery by picking up to three photons lying within a small angular region around the electron direction to calculate the electron energy. For vertexing we use the original non-recovered electrons. Using only the highest energy dilepton, numerous $B \to X_s \ell^+ \ell^-$ candidates are formed in every event by adding a K or a K_s, and up to three pions (of which at most one π^0). A candidate is rejected if it fails any of the following preselection requirements:

1. The dilepton vertex fit probability ($P_{\ell\ell vtx}$) satisfies $\log(P_{\ell\ell vtx}) > -10$.

2. The B vertex fit probability (P_{Bvtx}), using only the charged particles, satisfies $log(P_{Bvtx}) > -10$.

3. The invariant mass of the hadronic system $m_{Xs} < 2.5$ GeV.

4. The beam energy-substituted mass, see Sec. 2.1, satisfies $5.00 < m_{ES} < 5.29$ GeV.

5. The difference between the energy of the B candidate and the beam energy in the CM frame $|\Delta E| < 0.3$ GeV.

These criteria retain 77% and 90% of the electron and muon MC signal decays, respectively. The best candidate in the event is selected among the ones passing the preselection. In an average MC signal event there are about 5 candidates to choose from with some events having up to 30, see Figs. 30 and 31, the figures also show distributions of the number of combinations per event for $B\bar{B}$, continuum MC and data. In events with two or more candidates, the one with the highest signal likelihood is selected. The likelihood function is based on distributions of the variables ΔE, P_{Bvtx} and $\cos\theta_B$, where θ_B is the angle between the momentum of the B candidate and the beam axis in the CM frame. Figures 32 and 33 show the distributions of these variables for truth-matched (correctly reconstructed) signal decays and all other decays from the signal MC sample, hereby referred to as *cross-feed* decays. The ΔE distributions for the truth-matched signal exhibit a radiative tail, which is more pronounced for the electron modes. Since the initial electron and positron are relativistic, the $\Upsilon(4S)$ is created with its spin aligned parallel to the z direction. Thus the two scalar B mesons are preferentially produced in perpendicular directions to the z axis in the CM frame to conserve the angular momentum, resulting in the characteristic $\cos\theta_B$ distribution for the truth-matched signal. A cross-feed event would typically miss a signal decay product and/or include a particle from the other B. Such events tend to have flatter ΔE and $\cos\theta_B$ distributions and a lower vertex quality (if a charged particle from the other B was included).

The probability density functions (PDFs) for ΔE, $log(P_{Bvtx})$ and $\cos\theta_B$ are the sum of a Novosibirsk and a Gaussian function, the sum of a first-order polynomial and an exponential, and a second order polynomial, respectively. The signal likelihood is the product of these one-dimensional PDFs. The Novosibirsk function is

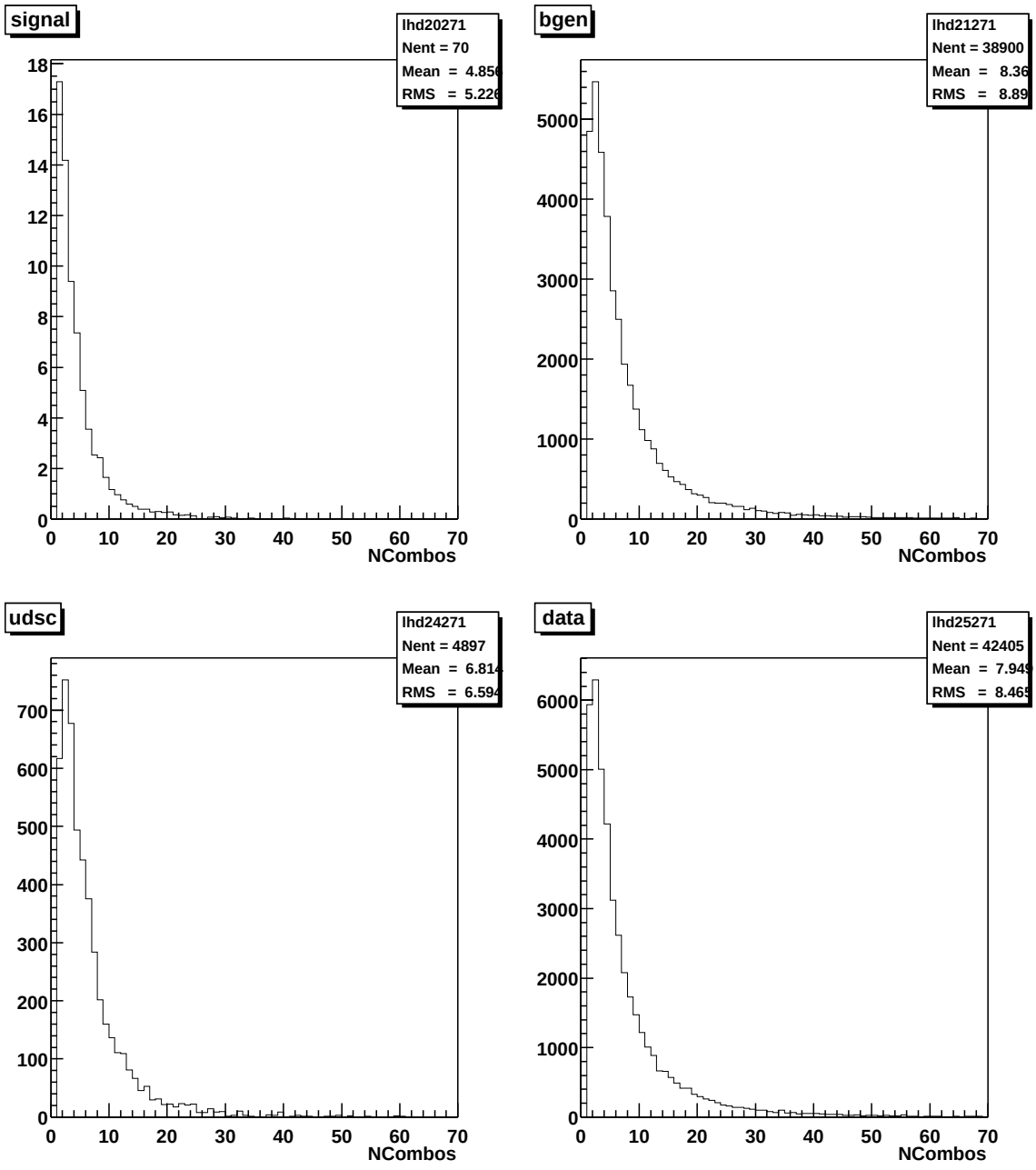

Figure 30: Distributions of the number of $B \to X_s e^+ e^-$ candidates per event after preselection for $B \to X_s \ell^+ \ell^-$ (top left), $B\bar{B}$ (top right), continuum MC (bottom left), and data events (bottom right).

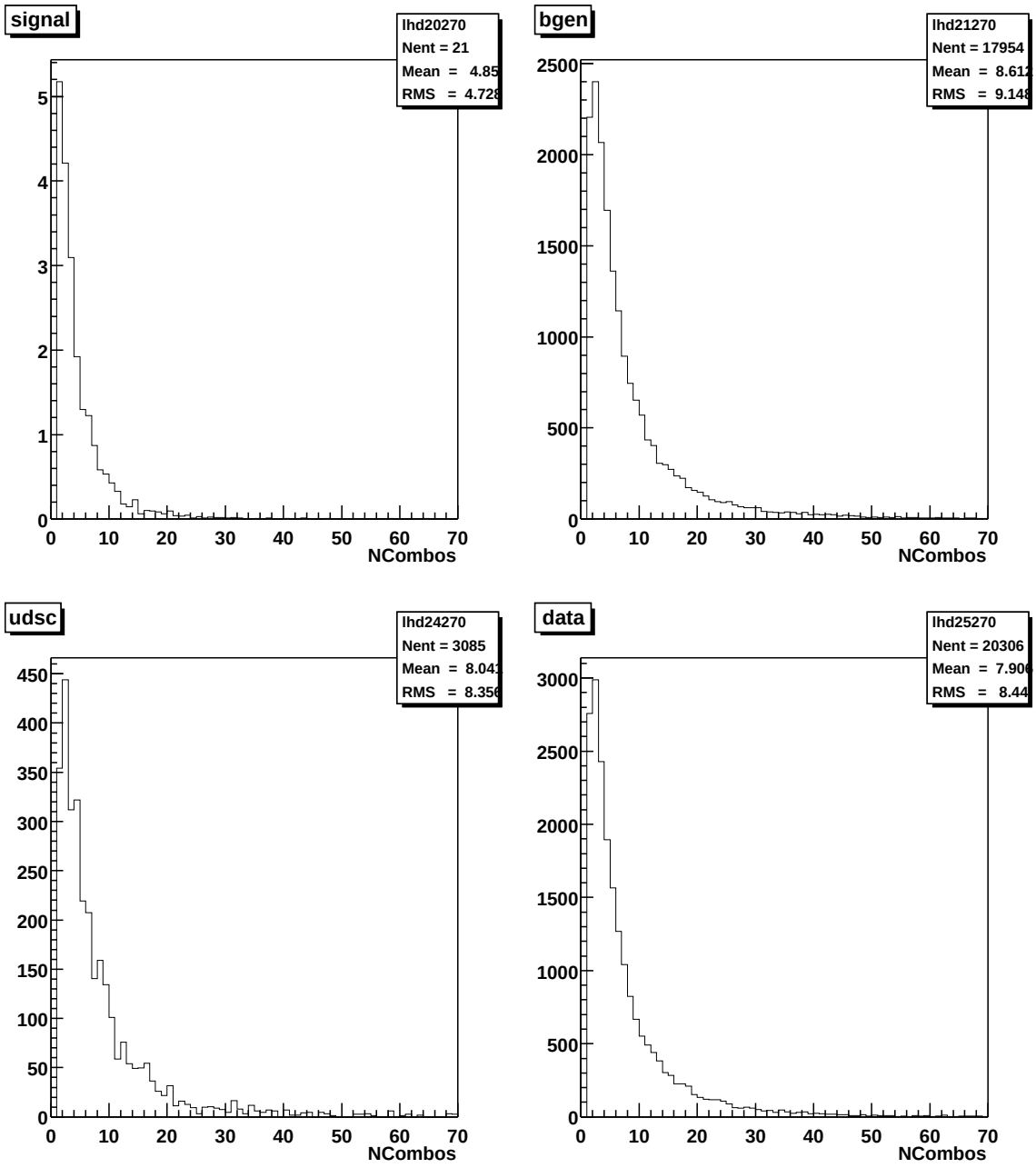

Figure 31: Distributions of the number of $B \to X_s \mu^+ \mu^-$ candidates per event after preselection for $B \to X_s \ell^+ \ell^-$ (top left), $B\bar{B}$ (top right), continuum MC (bottom left), and data events (bottom right).

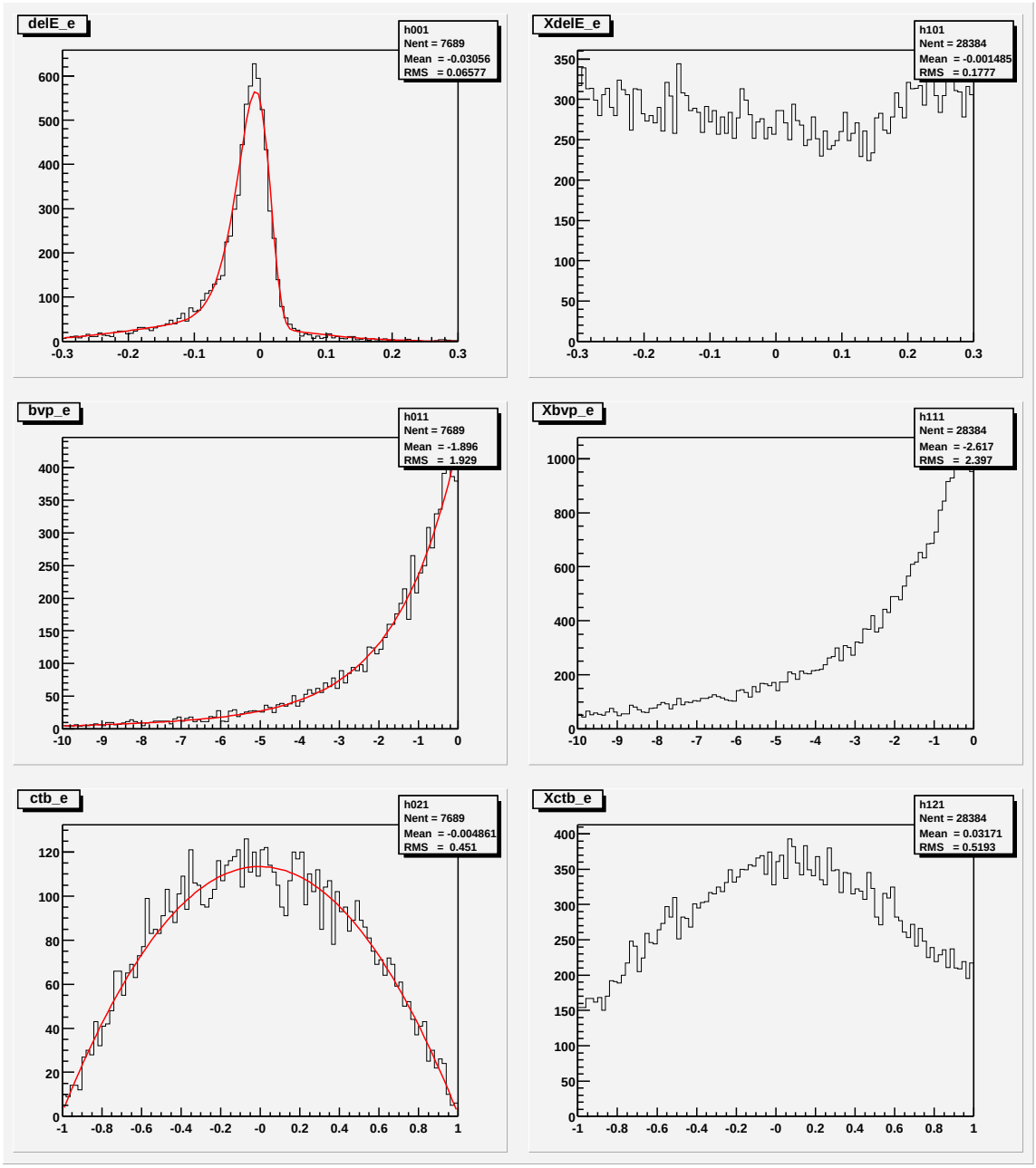

Figure 32: Distributions of ΔE (top), $\log(P_{Bvtx})$ (middle), and $\cos\theta_B$ (bottom) for truth-matched $B \to X_s e^+ e^-$ decays (left column), and all other decays (right column), a.k.a. cross-feed, in the signal $B \to X_s \ell^+ \ell^-$ MC.

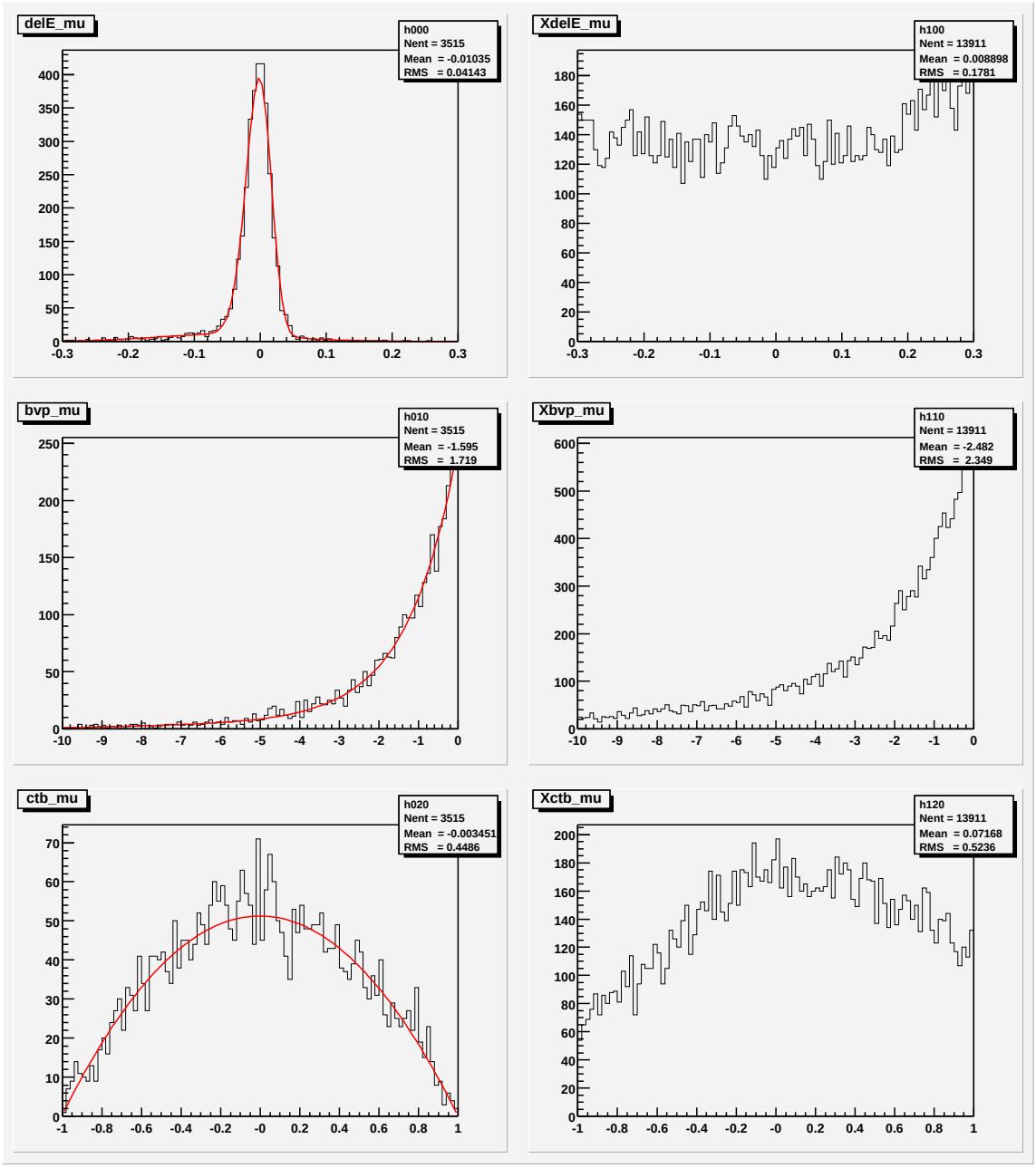

Figure 33: Distributions of ΔE (top), $\log(P_{Bvtx})$ (middle), and $\cos\theta_B$ (bottom) for truth-matched $B \to X_s\mu^+\mu^-$ decays (left column), and all other decays (right column), a.k.a. cross-feed, in the signal $B \to X_s\ell^+\ell^-$ MC.

effectively a Gaussian with an asymmetric tail

$$f_{Novo}(x,\mu,\sigma,\tau) = N \exp\left(\frac{-\log^2\left(1+\tau\frac{\sinh(\tau\sqrt{\log 4})}{\tau\sqrt{\log 4}}\frac{x-\mu}{\sigma}\right)}{2\tau^2} - \frac{\tau^2}{2}\right), \tag{5.1}$$

where μ is the x value at which the function peaks, σ is the Gaussian width, and τ is the tail parameter. The signal efficiencies for the fourteen considered hadron topologies for the best candidate selection step are 76.5% and 85% for the electron and muon modes, respectively.

5.3 Background suppression

In the third selection stage, events from various background sources are suppressed. Backgrounds that peak in the $m_{\rm ES}$ distribution, and thus mimic the $B \to X_s\, \ell^+\ell^-$ signal are:

1. Charmonium $B \to J/\psi X_s$ and $B \to \psi(2S) X_s$ decays, with J/ψ or $\psi(2S) \to e^+e^-$ or $\mu^+\mu^-$, constitute a potentially large background since these decays have a signature identical to the signal and a branching fraction higher by an order of magnitude. They are effectively removed by cutting on the dilepton mass around the J/ψ and $\psi(2S)$ masses, and provide an excellent control sample for the analysis.

2. $B \to X_s\gamma$ decays followed by a gamma conversion in the material $\gamma \to e^+e^-$ are a modest source of peaking background in the electron channel. This background is removed with a minimum cut on the dilepton mass.

3. Hadronic $B \to D^{(*)}\pi$ decays can also mimic signal decays if a $\pi^+\pi^-$ pair is misidentified as either an e^+e^- or a $\mu^+\mu^-$ pair. The latter is considered in this analysis due to the relatively high $\pi \to \mu$ misidentification rate.

Purely combinatorial background originates from both $B\bar{B}$ and continuum events. In the case of the $B\bar{B}$ background, the most common source are events in which each B meson decays semileptonically and the final state hadrons combine to form the $B \to X_s\, \ell^+\ell^-$ candidate.

An initial suppression of the combinatorial backgrounds is achieved by tightening some of the preselection requirements (first 3 cuts listed below). A cut on the maximum value of m_{Xs} is particularly useful as backgrounds rise with increasing m_{Xs}, whereas the signal decreases. Similarly, background events populate $m_{\rm ES}$

evenly, while the signal is concentrated near the B mass. These and additional cuts employed to reduce the backgrounds are:

1. $m_{Xs} < 1.8$ GeV.
2. $5.20 < m_{\rm ES} < 5.29$ GeV.
3. $-0.2 < \Delta E < 0.1$ GeV.
4. The z-separation between the two lepton at their point of closest approach to the beam spot in the xy plane satisfies $|\Delta z| < 0.15$ cm.
5. The difference between the energy of the rest of the event and the beam energy in the CM frame $-5.0 < \Delta E^{ROE} < 2.0$ GeV, where the rest of the event is formed out of all charged tracks and neutral EMC clusters not included in the B candidate.
6. The beam energy-substituted mass for the rest of the event $m_{ES}^{ROE} > 4.9$ GeV.
7. The lepton pair is required not to contain any electron candidate consistent with a possible gamma conversion.

The above criteria constitute the *postselection* of the analysis. After the postselection, charmonium background is suppressed by removing candidates with dilepton mass in the following ranges:

$$\begin{aligned} 2.70 < m(e^+e^-) < 3.25 \ \ \text{GeV}, \\ 2.80 < m(\mu^+\mu^-) < 3.20 \ \ \text{GeV}, \\ 3.45 < m(e^+e^-) < 3.80 \ \ \text{GeV}, \\ 3.55 < m(\mu^+\mu^-) < 3.80 \ \ \text{GeV}. \end{aligned} \tag{5.2}$$

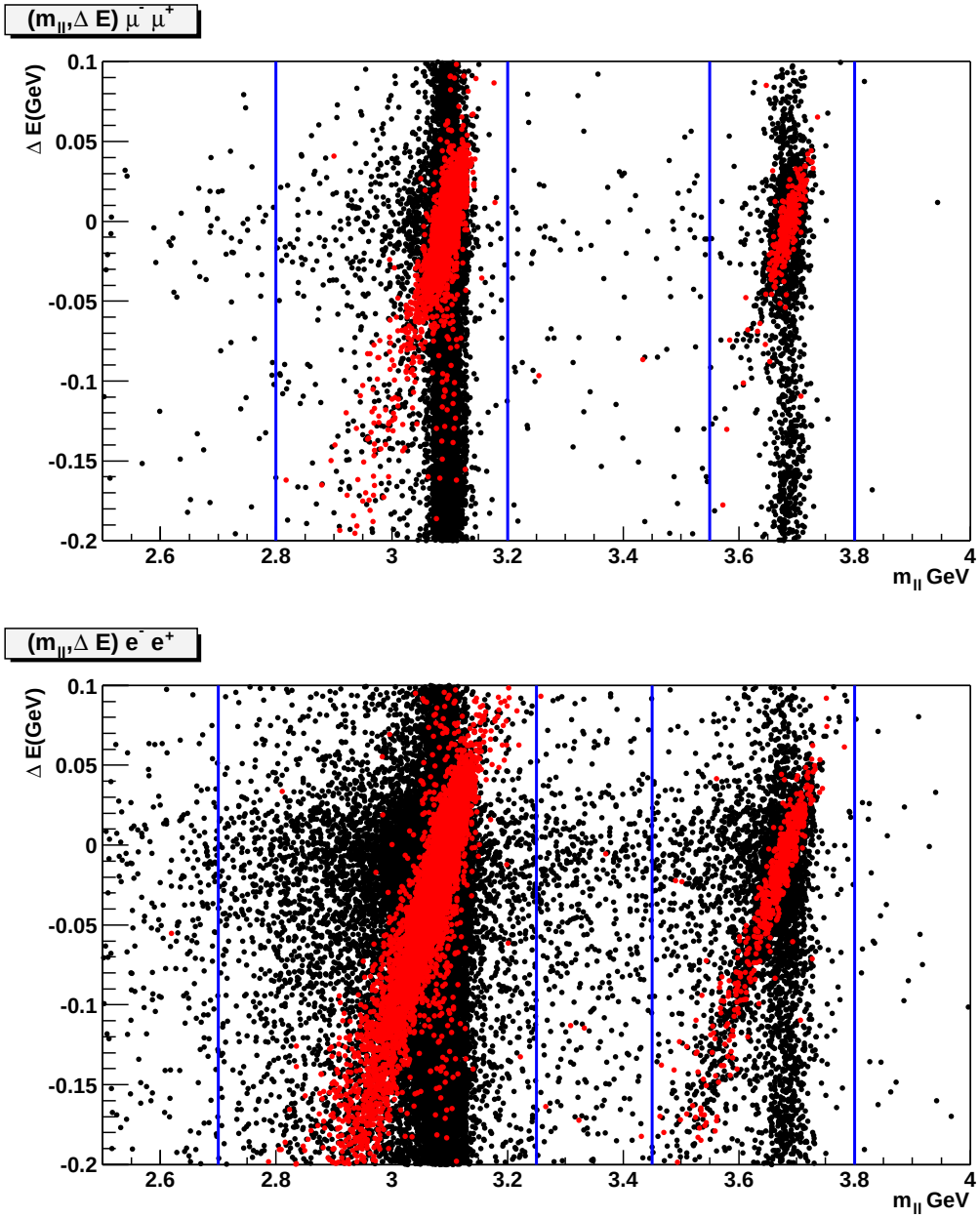

Figure 34: Scatter plots of ΔE vs. dilepton mass in $B\bar{B}$ MC events for which the leptons of the reconstructed $B \to X_s \ell^+ \ell^-$ candidate are either a J/ψ or $\psi(2S)$ decay products. The muon channel is shown on top and the electron channel is at the bottom. The vertical bands represent the veto regions. The red points indicate fully reconstructed charmonium B decays, while the black ones correspond to charmonium B decays with miss-reconstructed hadronic part.

Figure 34 shows scatter plots of ΔE (for the B decay candidate) as a function of dilepton mass in simulated $B \to J/\psi X_s$ and $B \to \psi(2S)X_s$ events. These *charmonium veto* cuts use the Brem-recovered dilepton mass and remove 98.50% (99.85%) of this background in the electron (muon) channel. However, we found that a small peaking component still remained in the electron channel and was traced to events in which one or both the electrons erroneously pick up a photon in the Bremsstrahlung recovery process. This may cause a sufficient increase of the dilepton mass to evade the charmonium veto. To reduce the effect of such events, the veto is applied to the dilepton mass before *and* after the Bremsstrahlung recovery. This additional requirement results in a loss of about 3% in signal efficiency. Single vs. double charmonium veto comparisons are shown on Fig. 35. Figure 36 illustrates the effect of the charmonium veto on the $m_{\rm ES}$ distributionfor $B \to X_s\ell^+\ell^-$ candidates in $B\bar{B}$ MC events for which the leptons of the reconstructed candidate are either a J/ψ or $\psi(2S)$ decay products.

Another source of peaking background (in the electron channel only) comes from $B \to X_s\gamma$ followed by a conversion of the photon into an e^+e^- pair in the detector material. Figure 37 shows $m_{\rm ES}$ and dilepton mass distributions for data and MC in the sideband region $m_{\rm ES} < 5.27$ GeV. The distributions are provided before and after the likelihood ratio cut (to be described later). A common feature is the excess of data in the very low dilepton mass region, where gamma conversions are expected to contribute. Because of this discrepancy between data and MC in the low m_{ll} region and because this region is not of theoretical interest, we require the dilepton mass to be greater than 0.2 GeV. If the already chosen pair fails this requirement the event is rejected.

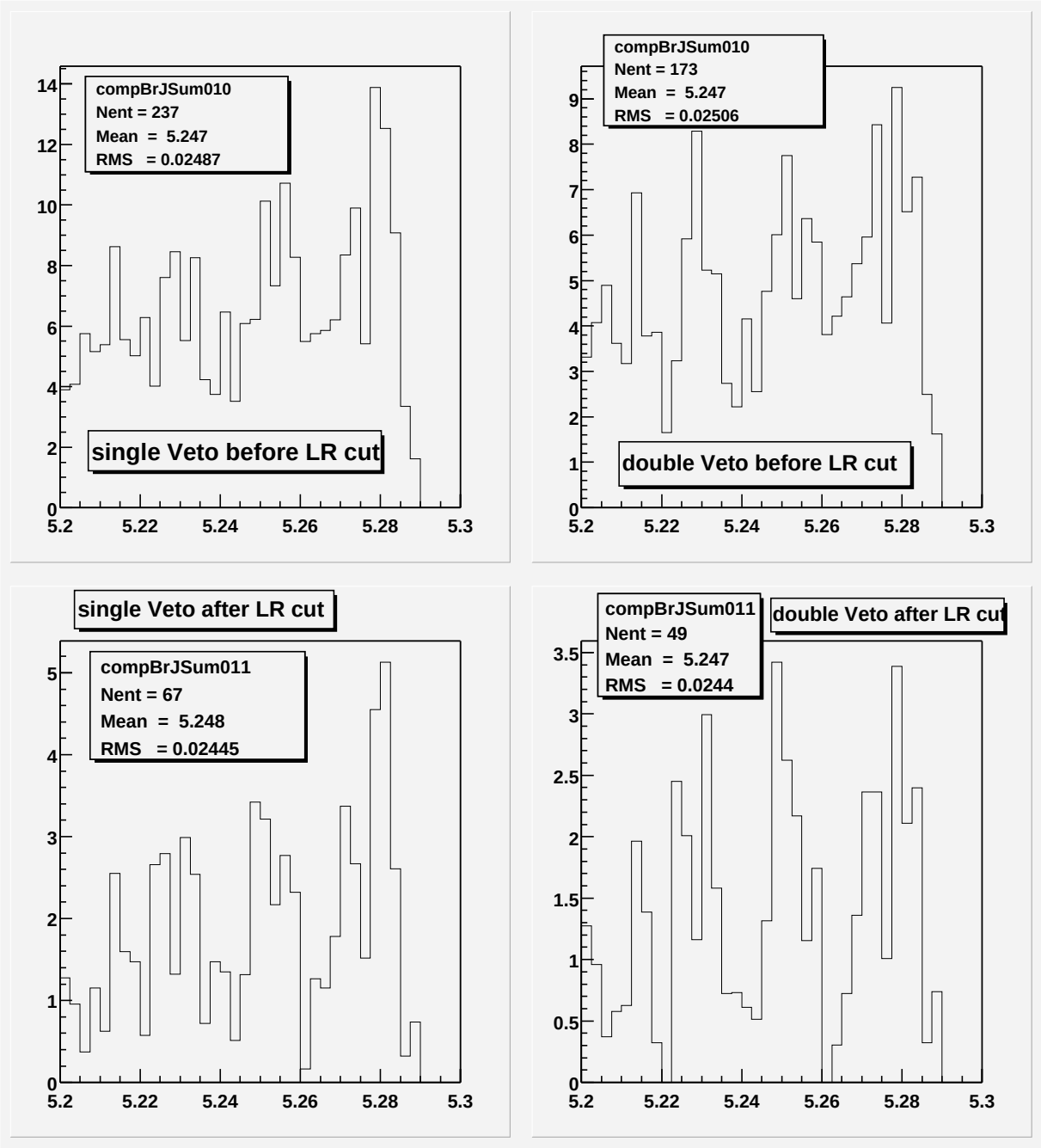

Figure 35: Distributions of m_{ES} for $B \to X_s e^+ e^-$ candidates in $B\overline{B}$ MC events for which the dilepton originates from either a J/ψ or $\psi(2S)$ decay, before likelihood ratio cut (top) and after (bottom). The left-hand plots correspond to applying a single charmonium veto and those on the right correspond to applying a double veto.

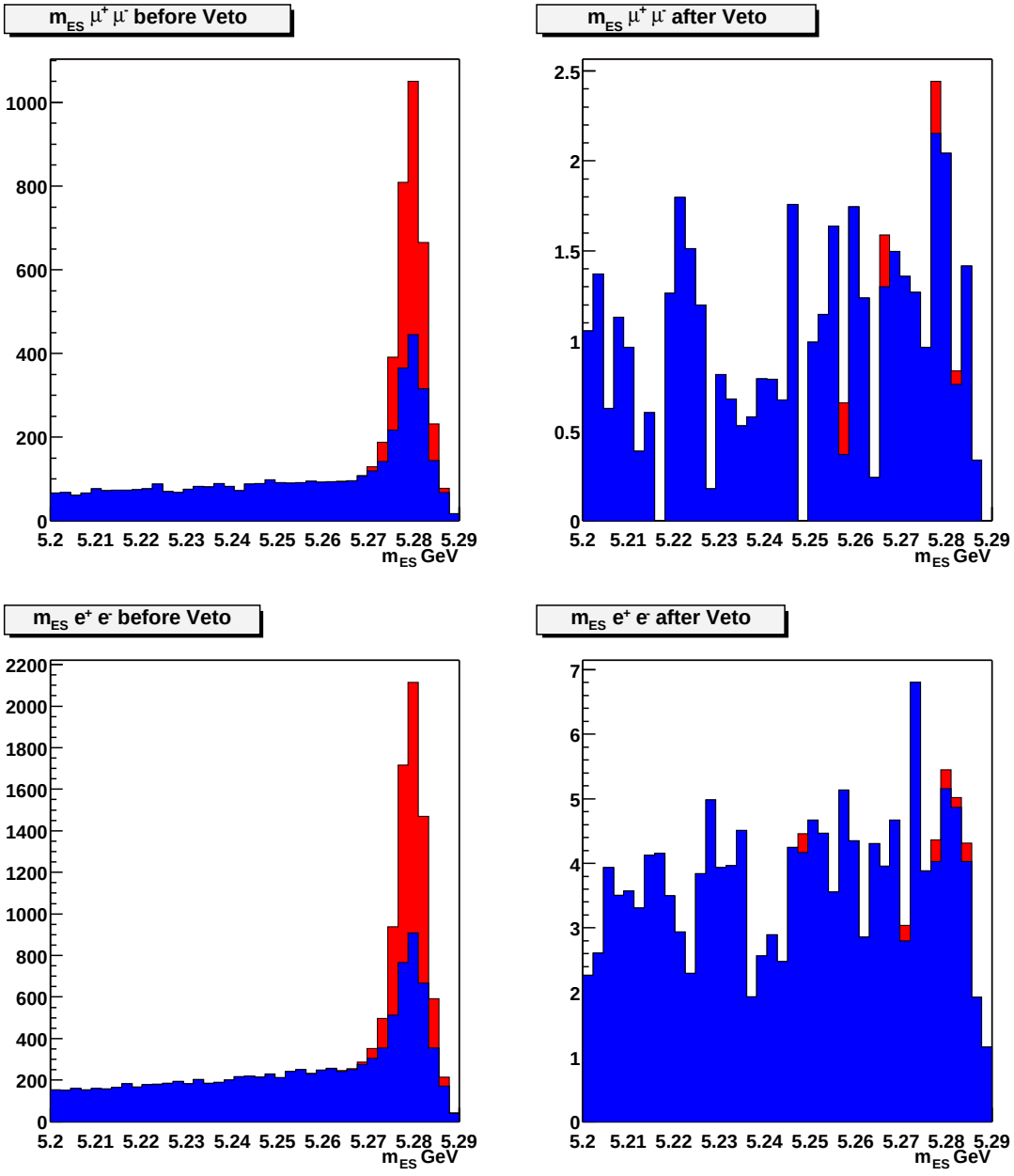

Figure 36: Distributions of $m_{\rm ES}$ for $B \to X_s \ell^+ \ell^-$ candidates in $B\bar{B}$ MC events for which the leptons of the reconstructed candidate are either a J/ψ or $\psi(2S)$ decay products. The muon channel is shown on top and the electron channel is at the bottom. All cuts prior to the charmonium veto are applied on the left and the charmonium veto is included on the right. The red area represents fully reconstructed charmonium B decays and the blue corresponds to charmonium B decays with miss-reconstructed hadronic part.

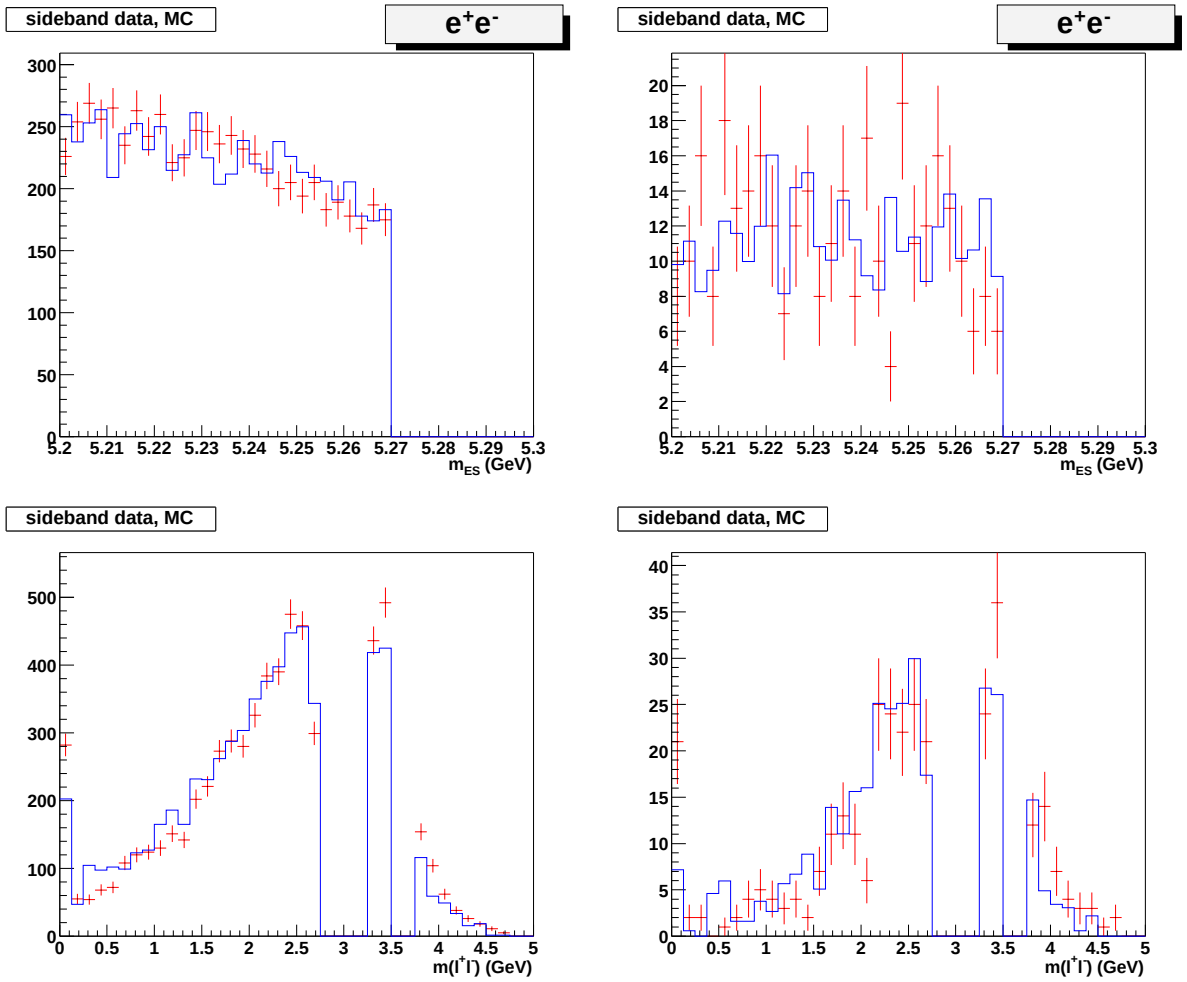

Figure 37: Distributions of m_{ES} (top) and dilepton mass (bottom) for $B \to X_s e^+e^-$ candidates in $m_{\rm ES} < 5.27$ GeV sideband on-resonance data (points) and MC (histograms). The plots to the right include the likelihood ratio cut. MC distributions have been normalized to correspond to 81.9 fb^{-1}.

Further suppression of the combinatorial background is achieved by cutting on a likelihood ratio defined as

$$\mathcal{L}_R = \frac{\mathcal{L}^{signal}}{\mathcal{L}^{signal} + \mathcal{L}^{B\bar{B}} + \mathcal{L}^{cont}}. \tag{5.3}$$

The likelihood functions for each sample are products of nine PDFs – P_i – for that sample, $\mathcal{L}^{sample} = \prod_i P_i^{sample}$, each PDF providing some background separation. The variables we use for the PDFs are:

- Kinematical

  - ΔE

  - ΔE^{ROE}

  - $\mathrm{m}_{\mathrm{ES}}^{\mathrm{ROE}}$

  - $\cos\theta_{miss}$, where θ_{miss} is the angle between the missing momentum in the event and the z axis in the CM frame

- Topological

  - Δz

  - $\log(P_{Bvtx})$

  - $|\cos\theta_T|$, where θ_T is the angle between the thrust axes of the B candidate and the rest of the event in the CM frame

  - R_2, using only charged tracks

- Angular momentum

  - $\cos\theta_B$

Table 5: Functional forms used to fit the nine components of the $\mathcal{L}_R$ likelihood.

Variable	PDF form
ΔE	Sum of 2 Gaussians
ΔE^{ROE}	Novosibirsk + Gaussian
$\mathrm{m}_{\mathrm{ES}}^{\mathrm{ROE}}$	Novosibirsk + Gaussian
Δz	Sum of 2 Gaussians
$\log(P_{Bvtx})$	First order polynomial + exponential
$\cos\theta_{miss}$	Second order polynomial + sum of 2 exponentials
$\cos\theta_B$	Second order polynomial
$\lvert\cos\theta_T\rvert$	First order polynomial + exponential
R_2	Novosibirsk

The variables ΔE, $\log(P_{Bvtx})$ and $\cos\theta_B$ are the same as those already used for picking the best candidate. The *rest of the event* variables – ΔE^{ROE} and $\mathrm{m}_{\mathrm{ES}}^{\mathrm{ROE}}$ – are useful against combinatorial $B\bar{B}$ background, since these events are characterized by significant amount of missing energy due to neutrinos from both B decays. The topological variables rely either on the fact that the charged tracks in a signal event must be consistent with originating from a point – Δz and $\log(P_{Bvtx})$ – or on the overall event shape – $\lvert\cos\theta_T\rvert$ and R_2 – the latter being useful primarily against continuum. Plots of the nine variables, showing signal and background components along with the fits used to parametrize them as PDFs are shown in Figs. 78 to 94. Table 5 lists the functional form of each PDF. The ΔE, ΔE^{ROE} and $\mathrm{m}_{\mathrm{ES}}^{\mathrm{ROE}}$ variables are particularly efficient at rejecting $B\bar{B}$ background, especially for events with two semileptonic decays which are characterized by larger missing energy than signal events. For continuum suppression, the event shape variables $\lvert\cos\theta_T\rvert$ and R_2 are most useful.

5.4 Likelihood ratio cut optimization

The optimal value of the likelihood ratio cut is determined by maximizing the statistical significance of the signal yield N_{sig}. To this end the cut value is varied, and the corresponding signal yield is extracted from a binned maximum likelihood fit to the $m_{\rm ES}$ distribution in the MC, see for example the bottom left plot in Fig. 38. The signal significance is defined as N_{sig}/σ_{sig}, where σ_{sig} is the error on N_{sig} as determined from the fit. The fit function is a sum of an Argus-shape function [24] for the background events and a Gaussian for the truth-matched signal events. The shape parameters of these two functions are determined from separate fits to the background-only and signal-only $m_{\rm ES}$ distributions, see Fig. 38. The small cross-feed component is included in the background for this optimization method.

As mentioned earlier the background increases rapidly with the increase of multiplicity (or hadronic mass), while at the same time the signal decreases. This brings a dramatic drop of the signal-to-background ratio at high hadronic masses and necessitates a separate likelihood ratio cut optimization in different hadronic mass regions. The optimizations were performed separately for electron and muon modes in three m_{Xs} ranges following the method described above. For the m_{Xs} ranges we chose: $m_{Xs} < 0.6$ GeV, $0.6 < m_{Xs} < 1.1$ GeV and $m_{Xs} > 1.1$ GeV, the values roughly correspond to the boundaries between the three MC signal components. The likelihood ratio distributions in these three regions are shown in Figs. 39 through 44.

Given the small expected signal yield and the substantial background for candidates with final state topologies containing 3 pions, when optimizing we also considered a case where all 3-pion modes are excluded. In addition we explored the effect of defining the likelihood ratio in such a way as to account for the difference

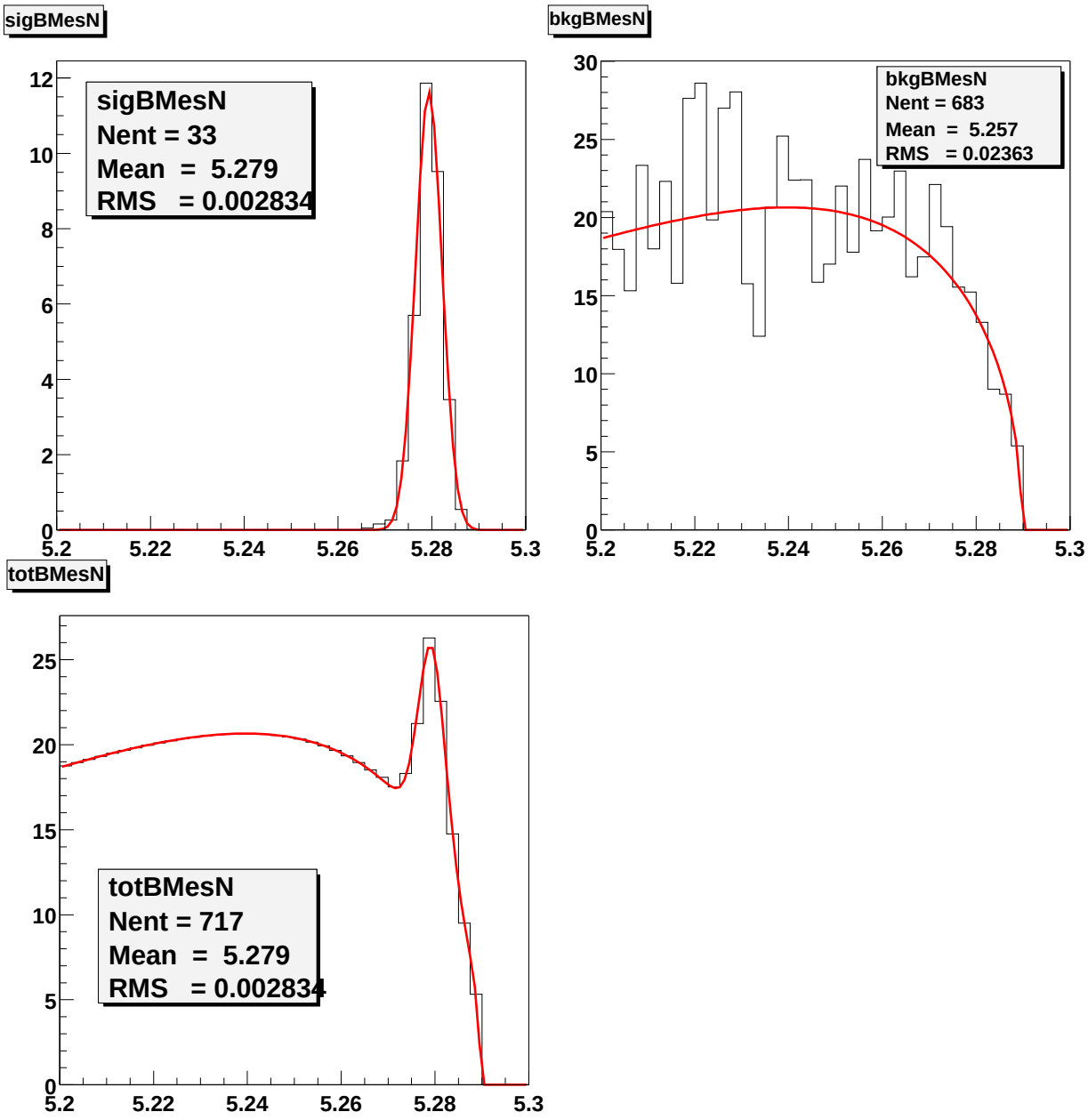

Figure 38: Distribution of m_{ES} in the electron channel with a particular set of likelihood ratio cuts for the following components in the MC: truth-matched $B \to X_s e^+ e^-$ signal (top left), total background including cross-feed (top right), and sum of signal and smoothed background (bottom left). The lines represent the result of a binned maximum likelihood fit to each distribution separately. Entries have been normalized to correspond to 81.9 fb^{-1}.

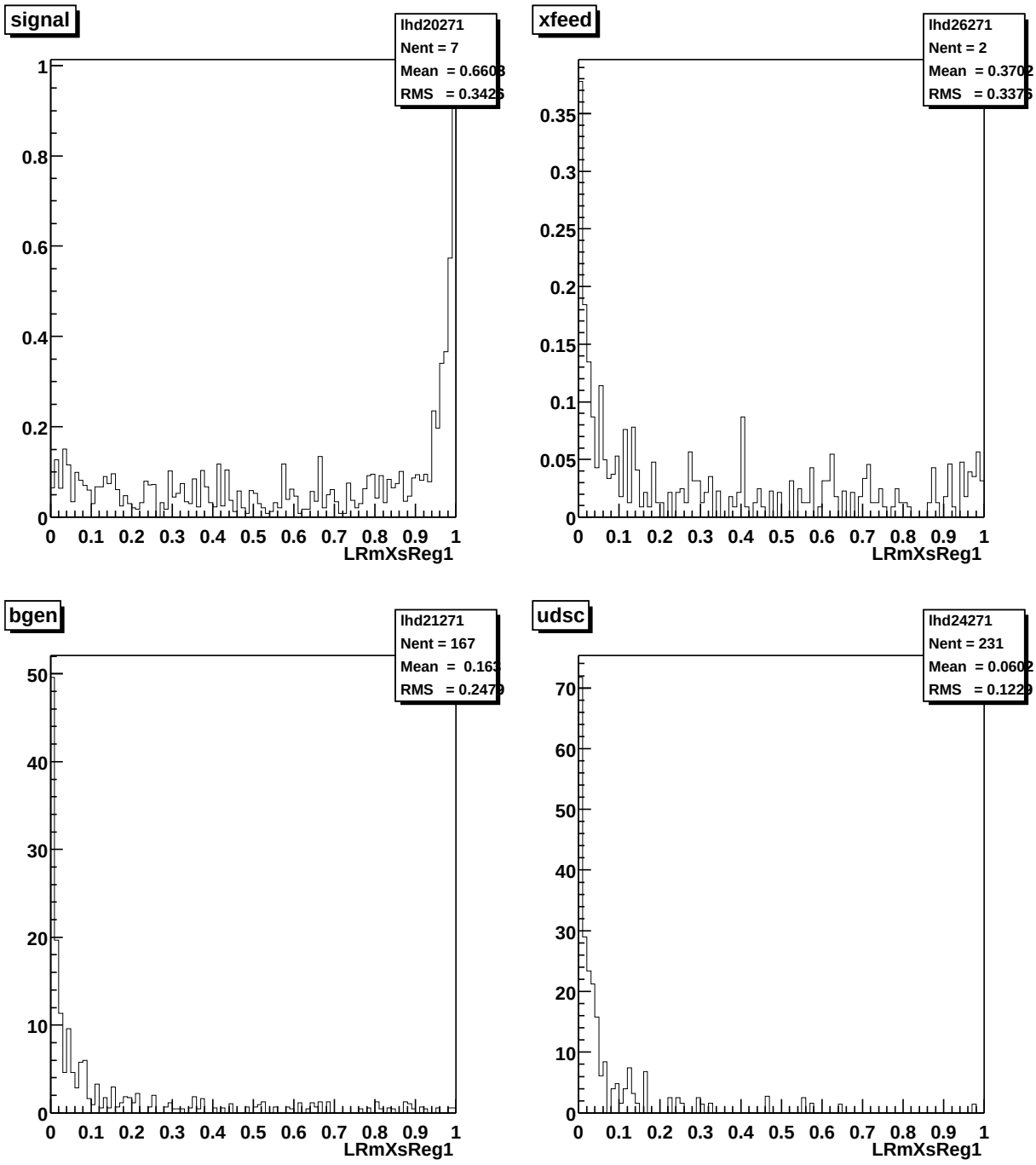

Figure 39: Distributions of the likelihood ratio for $B \to X_s\, e^+e^-$ candidates with $M_{X_s} < 0.6$ GeV in truth-matched signal (top left), cross-feed (top right), $B\overline{B}$ (bottom left) and continuum (bottom right) events.

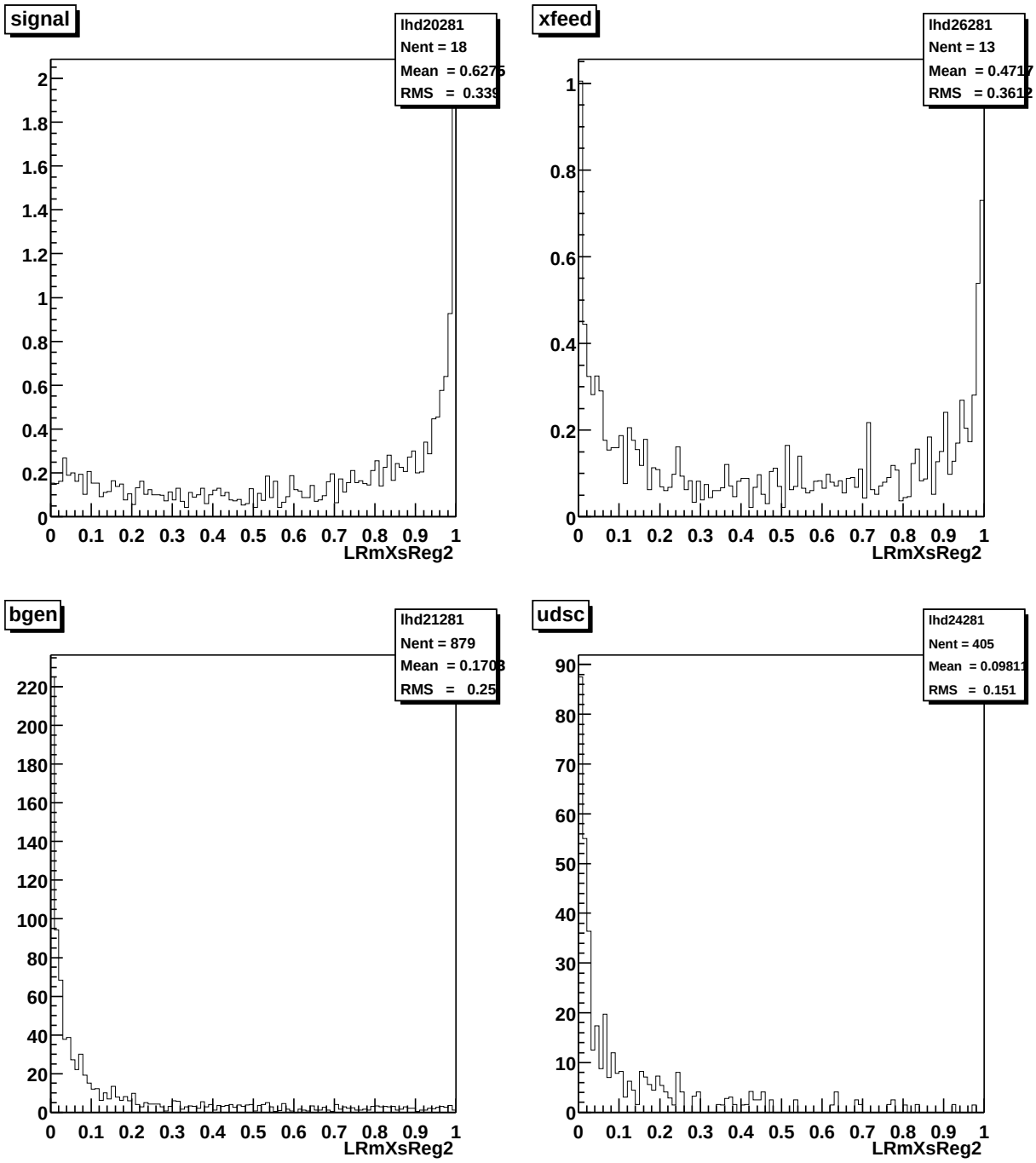

Figure 40: Distributions of the likelihood ratio for $B \to X_s\, e^+e^-$ candidates with $0.6 < M_{X_s} < 1.1$ GeV in truth-matched signal (top left), cross-feed (top right), $B\overline{B}$ (bottom left) and continuum (bottom right) events.

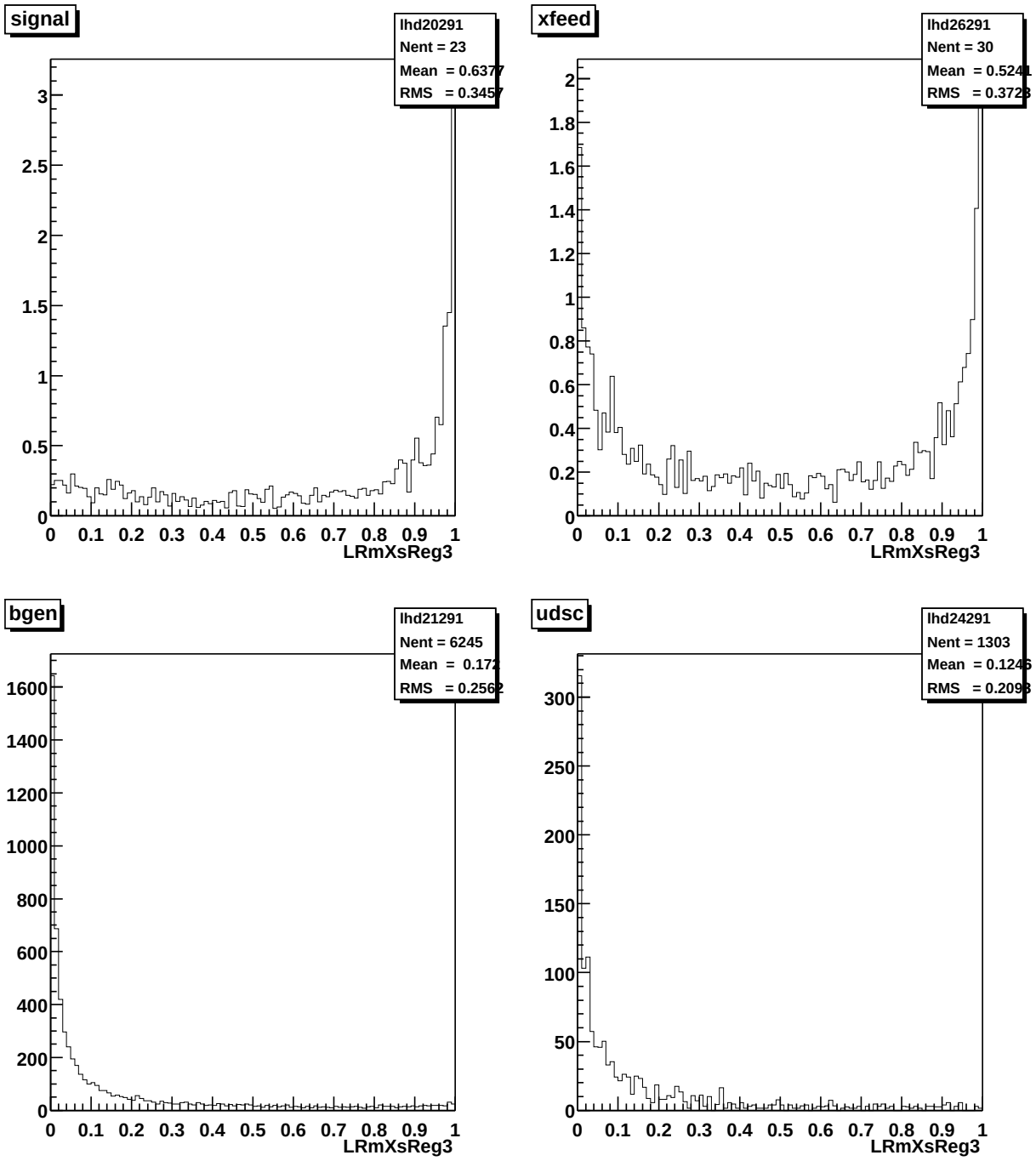

Figure 41: Distributions of the likelihood ratio for $B \to X_s\, e^+e^-$ candidates with $1.1 < M_{X_s} < 1.8$ GeV in truth-matched signal (top left), cross-feed (top right), $B\overline{B}$ (bottom left) and continuum (bottom right) events.

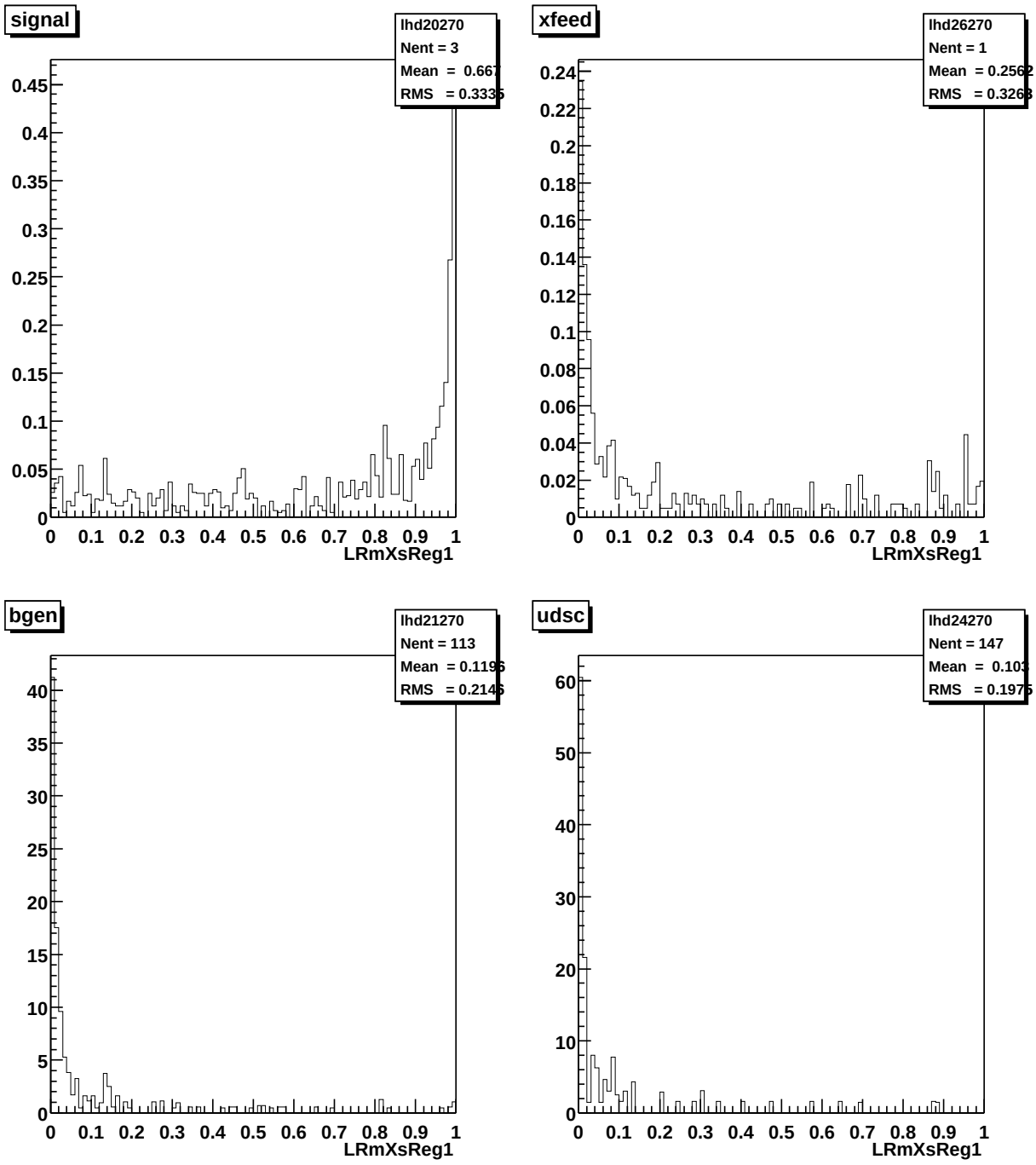

Figure 42: Distributions of the likelihood ratio for $B \to X_s\, \mu^+\mu^-$ candidates with $M_{X_s} < 0.6$ GeV in truth-matched signal (top left), cross-feed (top right), $B\overline{B}$ (bottom left) and continuum (bottom right) events.

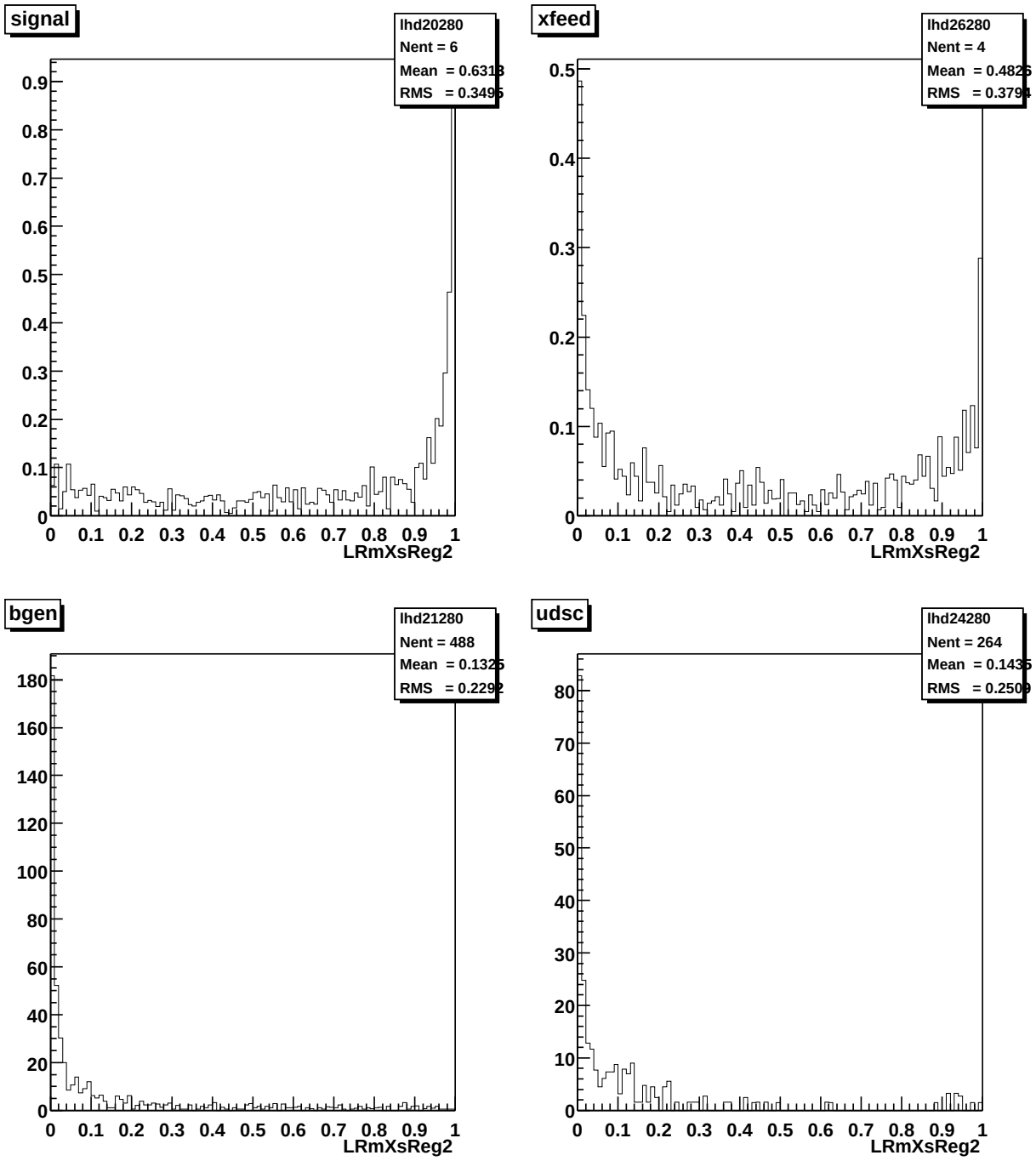

Figure 43: Distributions of the likelihood ratio for $B \to X_s\, \mu^+\mu^-$ candidates with $0.6 < M_{X_s} < 1.1$ GeV in truth-matched signal (top left), cross-feed (top right), $B\overline{B}$ (bottom left) and continuum (bottom right) events.

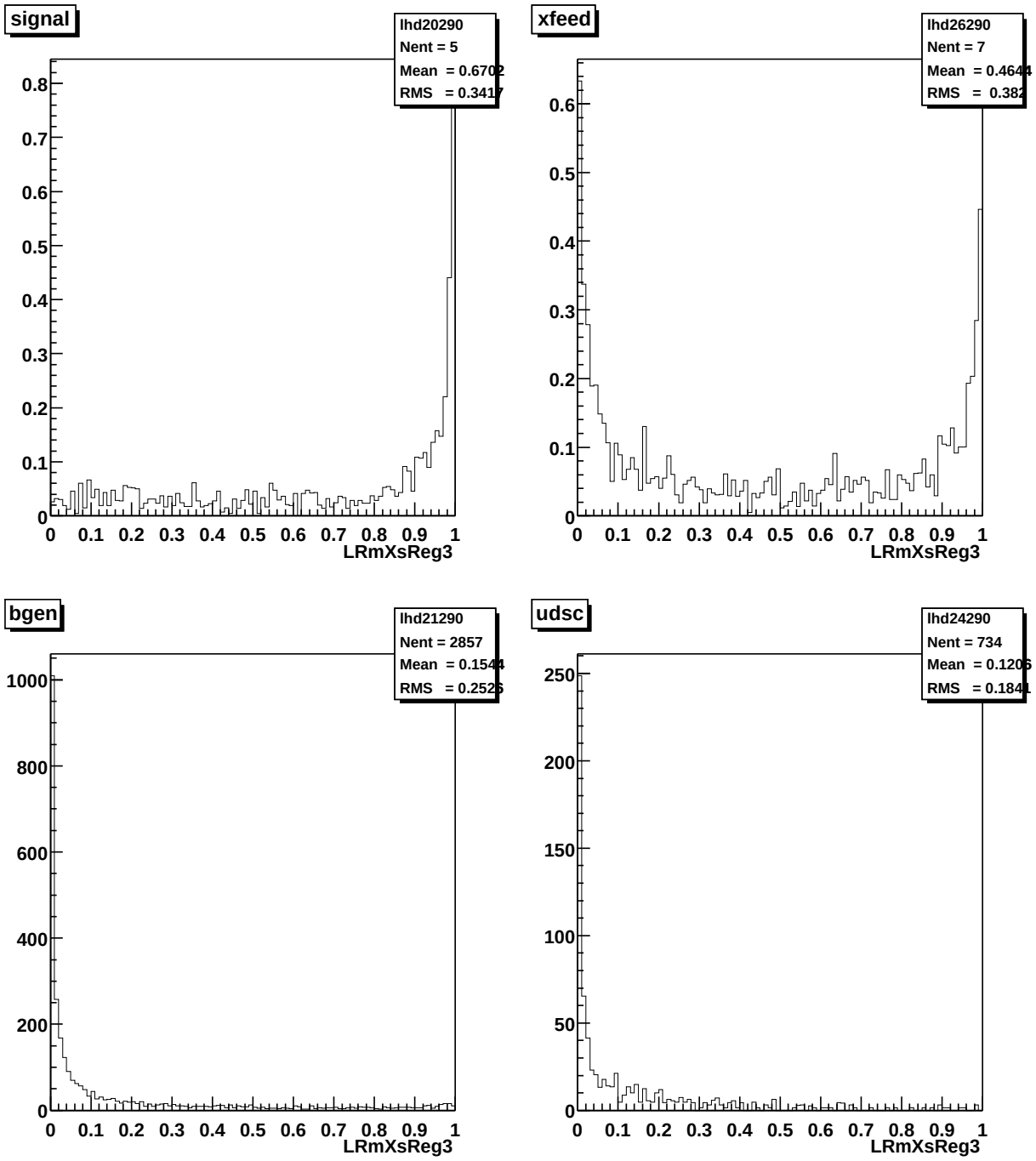

Figure 44: Distributions of the likelihood ratio for $B \to X_s\, \mu^+\mu^-$ candidates with $1.1 < M_{X_s} < 1.8$ GeV in truth-matched signal (top left), cross-feed (top right), $B\overline{B}$ (bottom left) and continuum (bottom right) events.

in the number of candidates from $B\bar{B}$ and continuum events

$$\mathcal{L}'_R = \frac{\mathcal{L}^{signal}}{\mathcal{L}^{signal} + f_{B\bar{B}}\mathcal{L}^{B\bar{B}} + f_{cont}\mathcal{L}^{cont}}, \tag{5.4}$$

where $f_{B\bar{B}}$ and f_{cont} represent the fraction of the background due to $B\bar{B}$ and continuum events: $f_{B\bar{B}} = 79\%(75\%)$ and $f_{cont} = 21\%(25\%)$ in the electron (muon) channel. The results of the different optimization methods and the optimal cut values for each case are presented in Table 5.1. Based on these results we apply the following set of cuts as the final stage of the event selection:

1. The 3 pion modes are excluded. As a result, the final sample contains a total of 10 different modes. Given the final states produced in our particular MC model (see Table 3) and the fraction of $K\ell^+\ell^-$ modes (see Table 2), this set of 10 modes corresponds to approximately 50% of all possible final states. If one assumes that the missing K_L component is equal to the K_S component, this fraction becomes 70%.

2. Likelihood ratio cuts for $B \to X_s e^+ e^-$:

$$\begin{aligned} \mathcal{L}_R &> 0.3 \text{ for } m_{Xs} < 0.6 \text{ GeV}, \\ \mathcal{L}_R &> 0.4 \text{ for } 0.6 < m_{Xs} < 1.1 \text{ GeV}, \\ \mathcal{L}_R &> 0.9 \text{ for } m_{Xs} > 1.1 \text{ GeV}. \end{aligned} \tag{5.5}$$

3. Likelihood ratio cuts for $B \to X_s \mu^+ \mu^-$:

$$\begin{aligned} \mathcal{L}_R &> 0.2 \text{ for } m_{Xs} < 0.6 \text{ GeV}, \\ \mathcal{L}_R &> 0.6 \text{ for } 0.6 < m_{Xs} < 1.1 \text{ GeV}, \\ \mathcal{L}_R &> 0.9 \text{ for } m_{Xs} > 1.1 \text{ GeV}. \end{aligned} \tag{5.6}$$

	$\mathcal{L}_R$ cut								Signal		Bkg		Significance	
	e			μ					e	μ	e	μ	e	μ
Single $\mathcal{L}_R$ cut	0.7			0.9					24.0	5.7	596	146	2.71	1.25
m_{Xs} range (GeV)	< 1.1	$1.1 - 1.4$	> 1.4	< 1.1	$1.1 - 1.4$	> 1.4								
3 $\mathcal{L}_R$ cuts (with 3 π)	0.7	0.2	0.9	0.9	0.2	0.9	22.6	8.4	417	226	2.99	1.72		
3 $\mathcal{L}_R$ cuts (no 3 π)	0.7	0.4	0.9	0.9	0.2	0.9	21.8	8.3	349	200	3.08	1.77		
3 $\mathcal{L}'_R$ cuts (no 3 π)	0.7	0.7	0.9	0.9	0.4	0.9	23.4	9.0	487	289	2.84	1.58		

Table 5.1: Signal and background yields for different likelihood ratio cut optimization approaches. The first row shows the results obtained with a single cut, the following rows correspond to separate cuts in different m_{Xs} regions, with and without the 3-pion modes. The last row corresponds to the likelihood ratio $\mathcal{L}'_R$ defined in Eq. 5.4. The likelihood ratio cut values are given on the left-hand side in the order of the m_{Xs} ranges.

An improvement of the signal significance by about 0.3 and 0.5 σ is achieved in the electron and muon channels compared to the single $\mathcal{L}_R$ cut case. Figures 45 and 46 show the final m_{Xs} spectrum of reconstructed truth-matched MC signal in the electron and muon chanels.

The MC-based signal reconstruction efficiencies at the different stages of the analysis for the 14 final state topologies are presented in Tables 28 and 29. By far, the largest source of inefficiency is due to the requirement that all final state particles are reconstructed and correctly identified (the column "in lists" in the tables). Similarly, Tables 30 and 31 show efficiencies for MC background events in the electron and muon channels.

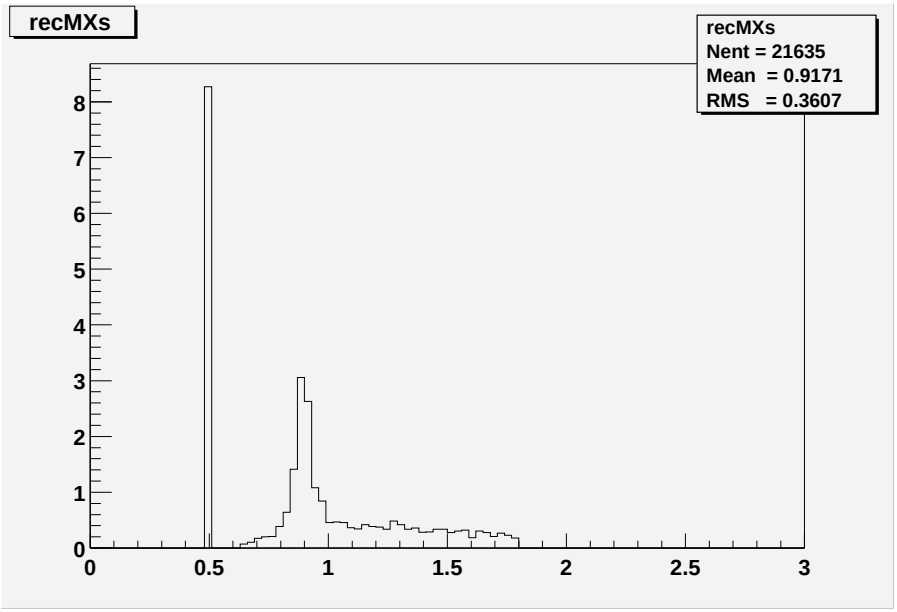

Figure 45: Distributions of hadronic mass for correctly reconstructed $B \to X_s\, e^+e^-$ signal after the nominal likelihood ratio cuts are applied. The selection also excludes final states containing 3 pions. The Belle model is used for the signal MC.

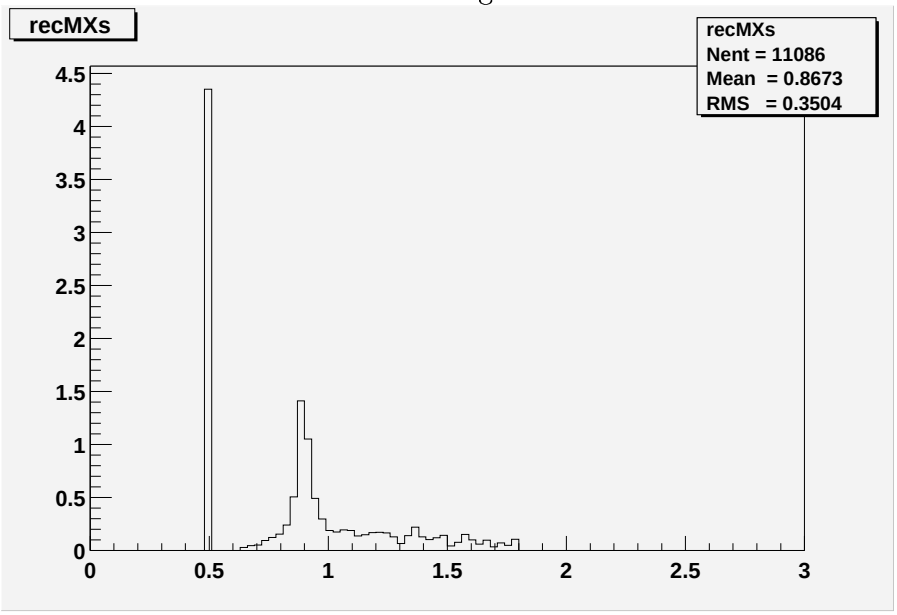

Figure 46: Distributions of hadronic mass for correctly reconstructed $B \to X_s\, \mu^+\mu^-$ signal after the likelihood ratio cuts are applied. The selection also excludes final states containing 3 pions. The Belle model is used for the signal MC.

CHAPTER 6

Cross checks and validation

In this analysis we use MC generated signal and background events and detector simulation to tune the background suppression cuts and estimate the final signal efficiency. In order to test the validity of such approach we use the following control samples to compare directly the real detector response and cut effectiveness to the ones obtained from simulation:

1. Charmonium sample.
2. $X_s e^{\pm}\mu^{\mp}$ sample.
3. Off-resonance sample.
4. On-resonance sideband sample.

6.1 Charmonium control sample

The charmonium decays $B \to J/\psi X_s$ and $B \to \psi(2S)X_s$ have signatures identical to that of our signal and provide an excellent control sample to compare signal reconstruction efficiencies, albeit in restricted dilepton mass regions. The control sample is obtained at the stage when the charmonium veto is applied, after the

Table 6: Definition of the topology number of the hadronic system X_s.

X_s Topology	Id #
K	1
$K\pi^0$	2
$K\pi$	3
$K\pi\pi^0$	4
$K\pi\pi$	5
K_S^0	20
$K_S^0\pi^0$	21
$K_S^0\pi$	22
$K_S^0\pi\pi^0$	23
$K_S^0\pi\pi$	24

postselection and before the likelihood ratio cuts. The events in the sample are required to satisfy one of the following dilepton mass cuts:

$$\begin{aligned} 2.70 < m(e^+e^-) < 3.25 \ \ \text{GeV}, \\ 2.85 < m(\mu^+\mu^-) < 3.15 \ \ \text{GeV}, \\ 3.50 < m(e^+e^-) < 3.76 \ \ \text{GeV}, \\ 3.60 < m(\mu^+\mu^-) < 3.74 \ \ \text{GeV}, \end{aligned} \tag{6.1}$$

where the e^+e^- invariant mass is computed after Bremsstrahlung recovery. In order to obtain a sample with higher charmonium purity the above cuts are tighter than the ones used for the veto. First we compare data and MC for this *raw* sample, which includes B candidates with $m_{\rm ES}$ values between 5.20 and 5.29 GeV implying that the hadronic system is not correctly reconstructed for a significant fraction of the sample, see Figs. 47 to 49. Figure 48 shows the X_s system topology following the definition in Table 6.

Second, we compare data and MC for fully reconstructed $B \to J/\psi X_s$ and $B \to \psi(2S)X_s$ decays by only considering events from the signal region in $m_{\rm ES}$:

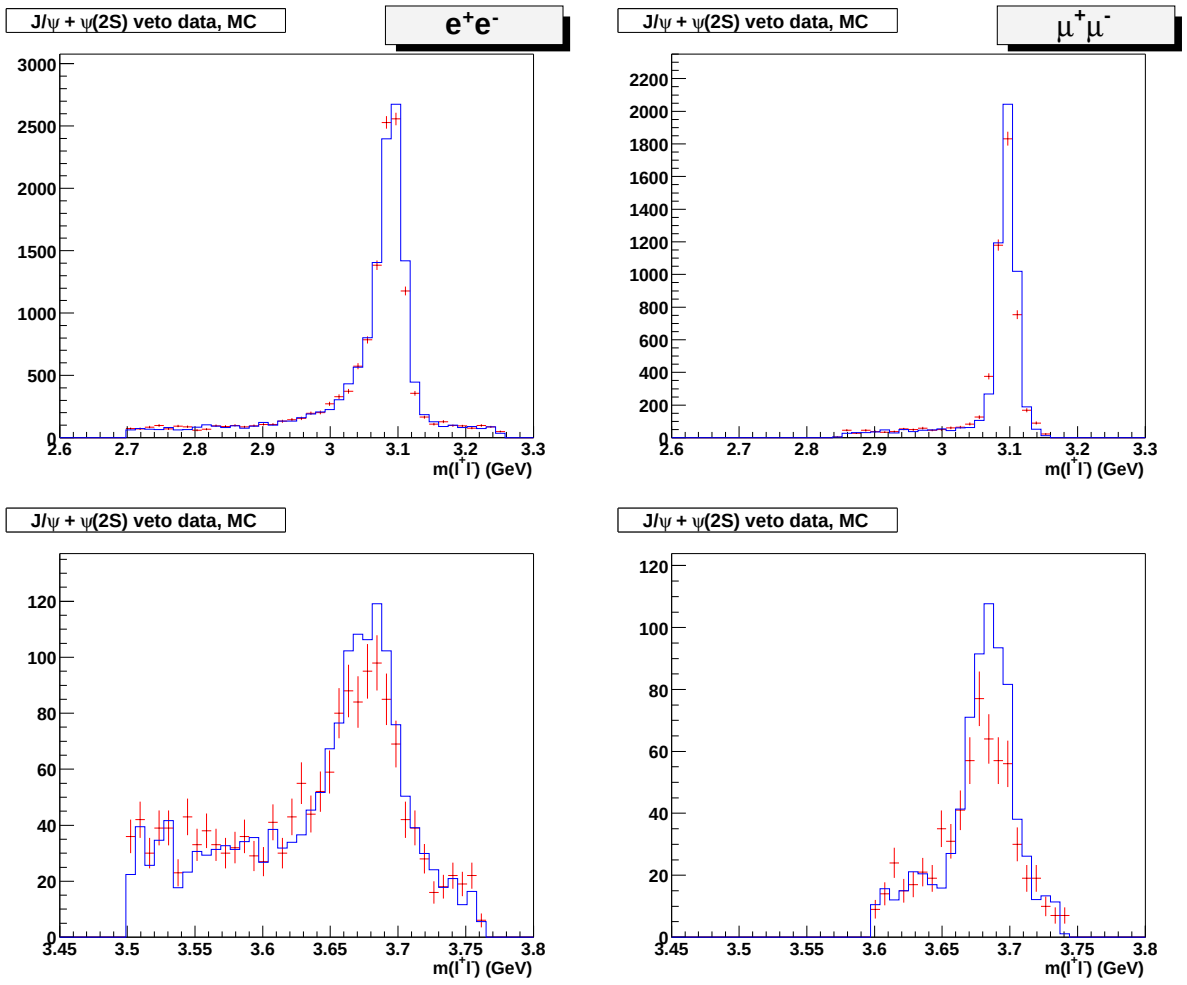

Figure 47: Distributions of dilepton invariant mass for $B \to X_s\, \ell^+\ell^-$ candidates in the charmonium veto sample for electrons (left) and muons (right). The top and bottom rows are for the J/ψ and $\psi(2S)$ mass regions, respectively. The points represent the data and the histograms the MC, scaled to correspond to an integrated luminosity of 81.9 fb^{-1}.

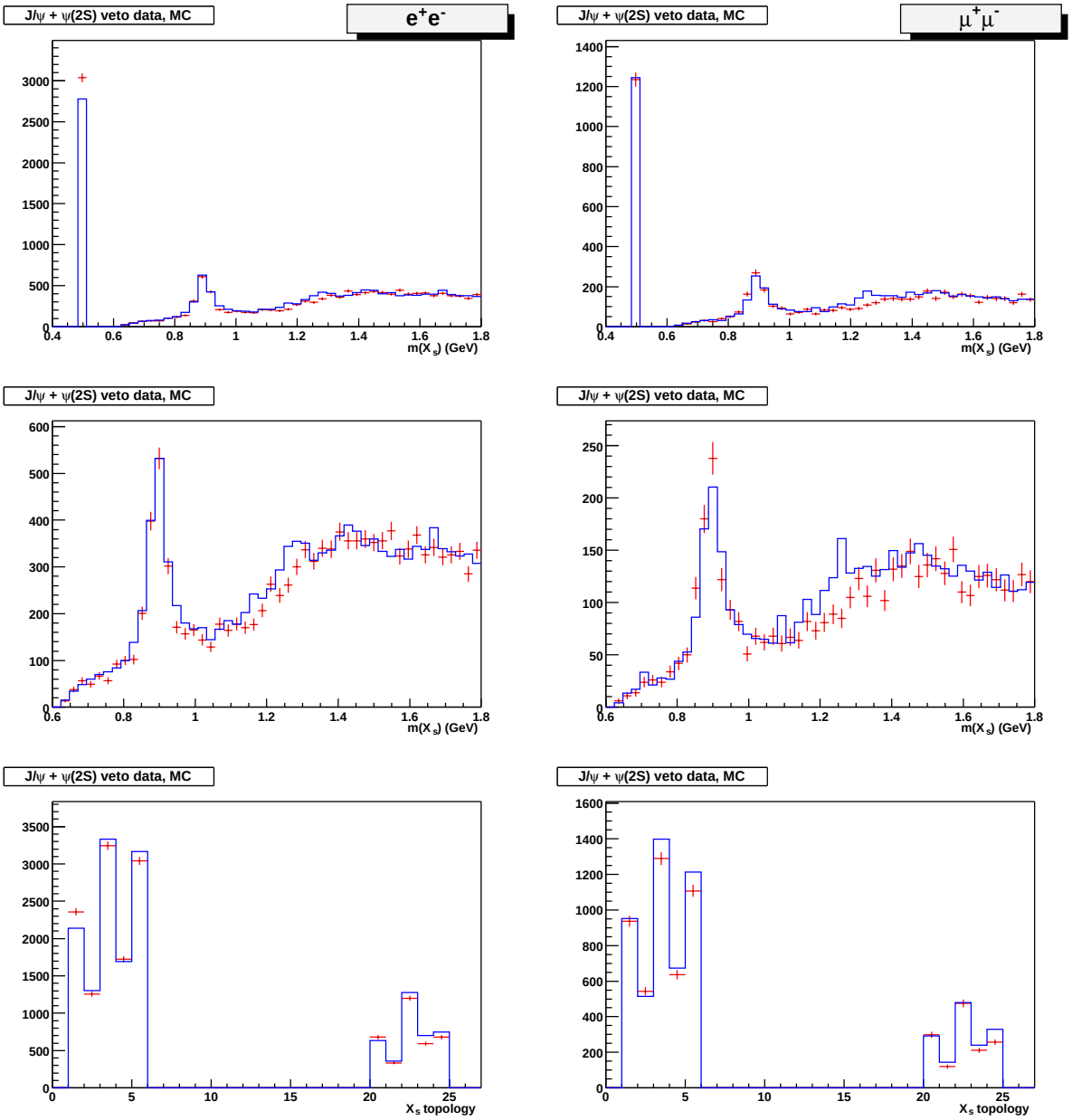

Figure 48: Distributions of hadronic invariant mass and final state topology (see Table 6) for $B \to X_s\, \ell^+\ell^-$ candidates in the charmonium veto sample for electrons (left) and muons (right). The points represent the data and the histograms the MC, scaled to correspond to an integrated luminosity of 81.9 fb^{-1}.

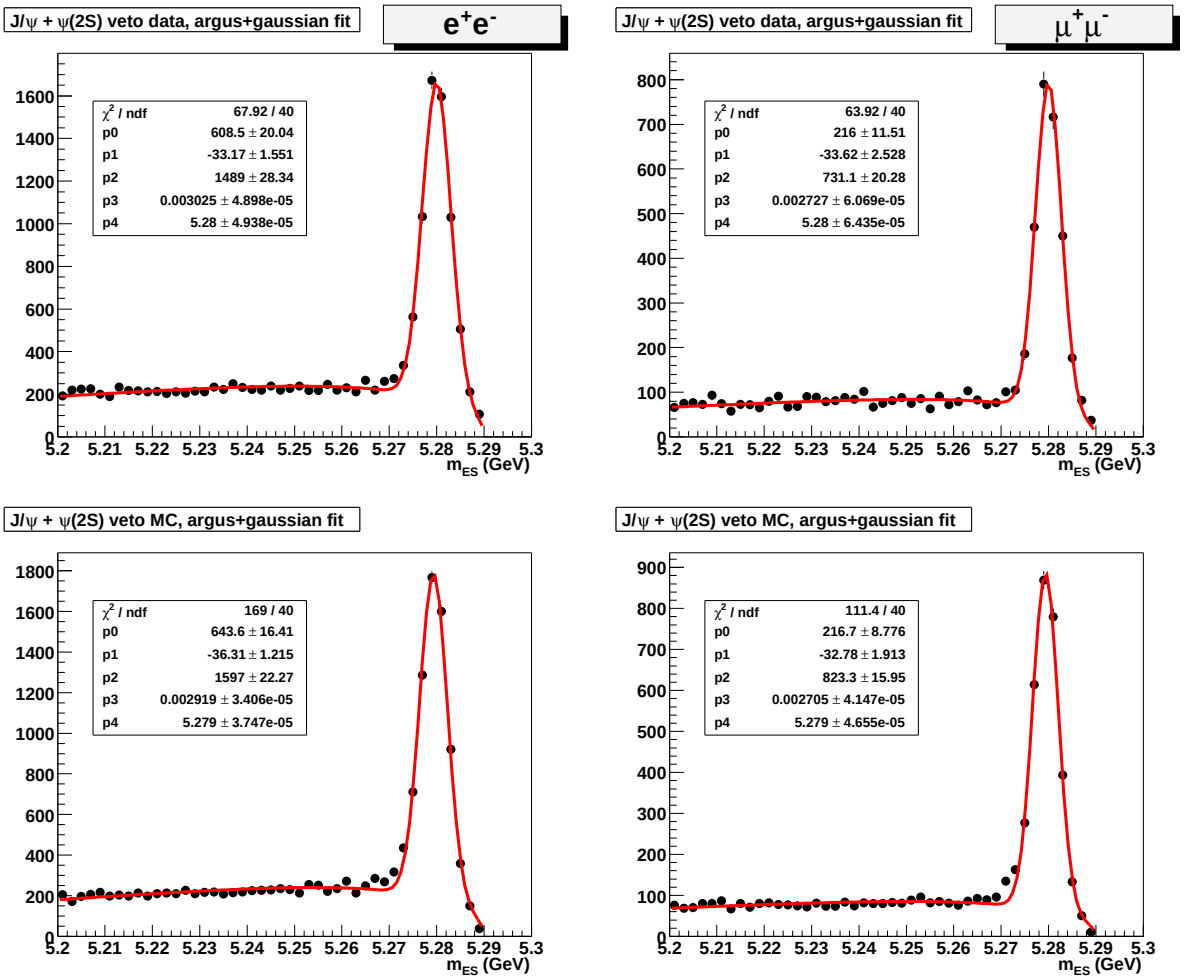

Figure 49: Distributions of m_{ES} for $B \to X_s\, \ell^+\ell^-$ candidates in the charmonium veto sample for electrons (left) and muons (right). The top row corresponds to data and the bottom shows the MC, scaled to correspond to an integrated luminosity of 81.9 fb^{-1}.

$5.27 < m_{\rm ES} < 5.29$ GeV. We estimate and subtract the combinatorial background in the signal region by using a scale factor obtained from a χ^2 fit to the $m_{\rm ES}$ distributions, see Fig. 49. The maximum likelihood fit uses a Gaussian to describe the signal component and an Argus function for the combinatorial background. The end point of the Argus function is fixed to 5.29 GeV, all other fit parameters are allowed to float. Distributions obtained after background subtraction approximate the distributions to be expected from correctly reconstructed charmonium B decays. Figures 6.1 and 6.2, in particular, show comparisons for the nine variables entering the background-suppression likelihood ratio. Other distributions are shown in Figs. 50–51. A slight excess is observed in the data for the Ke^+e^- final state and a slight deficit in the $K_1(1270)^0\ell^+\ell^-$ mode, see below.

Third, we apply the nominal values of the likelihood ratio cut. The resulting distributions are shown in Figs. 52–56. Of particular interest is the ratio between the absolute yields in the data and the MC. This ratio is shown in Tables 7 and 8 for each reconstructed $B \to J/\psi\, X$ topology in the electron and muon channels. Tables 9 and 10 show the same results for $B \to \psi(2S)\, X$ candidates. This comparison is important since it tests whether the simulation reliably estimates the efficiency to reconstruct B decays into final states with a signature identical to that of the $B \to X_s\, \ell^+\ell^-$ signal. However, the current knowledge of inclusive $B \to J/\psi\, X$ and $B \to \psi(2S)\, X$ decays is limited: $\mathcal{B}(B \to J/\psi\, X) = (1.15 \pm 0.06)\%$ and $\mathcal{B}(B \to \psi(2S)\, X) = (0.35 \pm 0.05)\%$ [2], i.e. the total branching fractions are known with a relative uncertainty of 5.2% and 14%, respectively. Furthermore, only a few exclusive decay channels have been measured with a level of accuracy close to this. The branching fractions for decays into $J/\psi K$ and $J/\psi K^*$ are known to 5-6% and 7-9%, respectively (the uncertainties vary between B^0 and B^+ decays). Higher mass final states are rather unknown. This means that the comparison be-

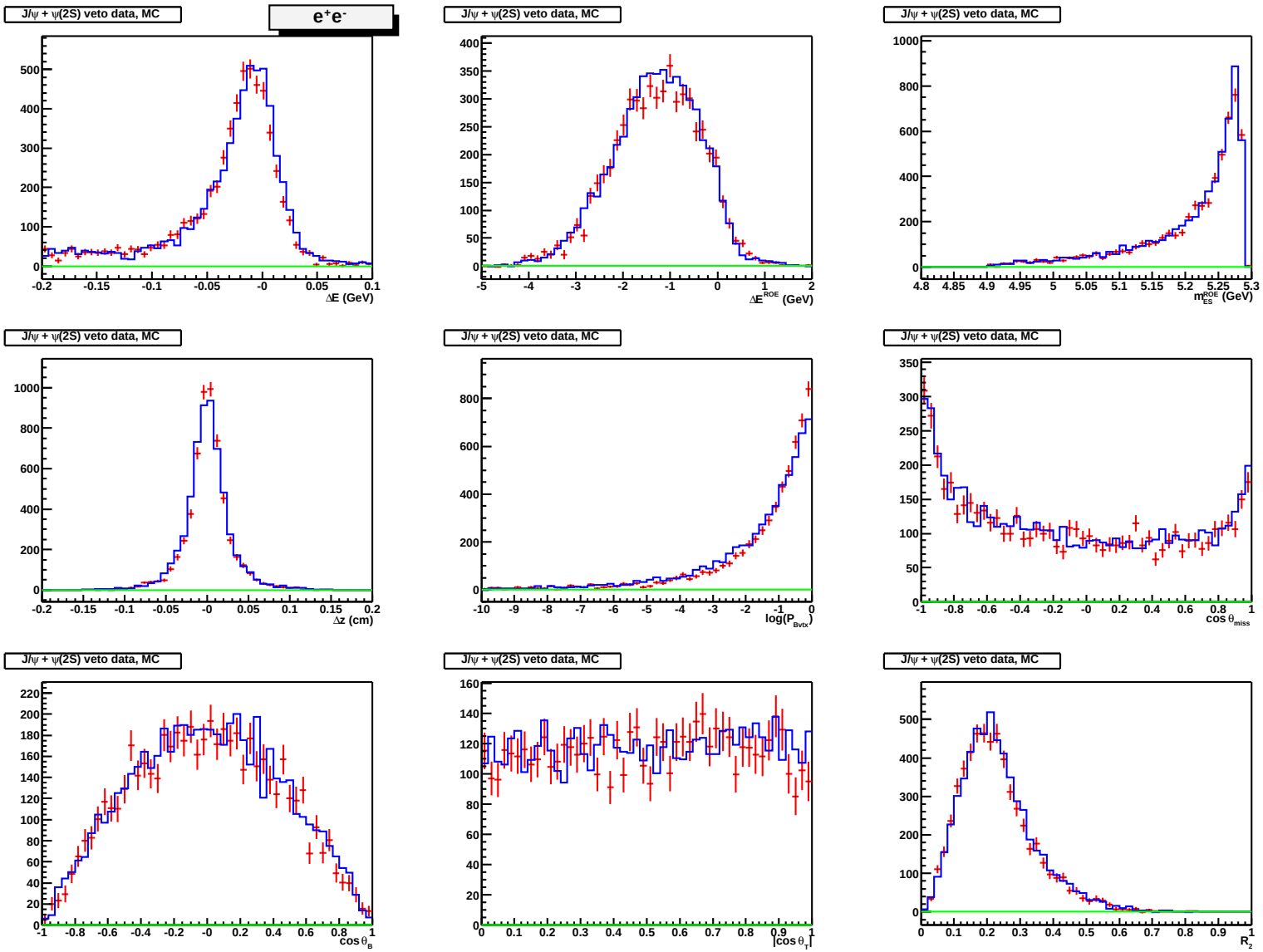

Figure 6.1: Distributions of the background-suppression variables entering the likelihood ratio for the background-subtracted charmonium events in the electron channel. The points represent the data and the histograms the MC, scaled to correspond to an integrated luminosity of 81.9 fb^{-1}.

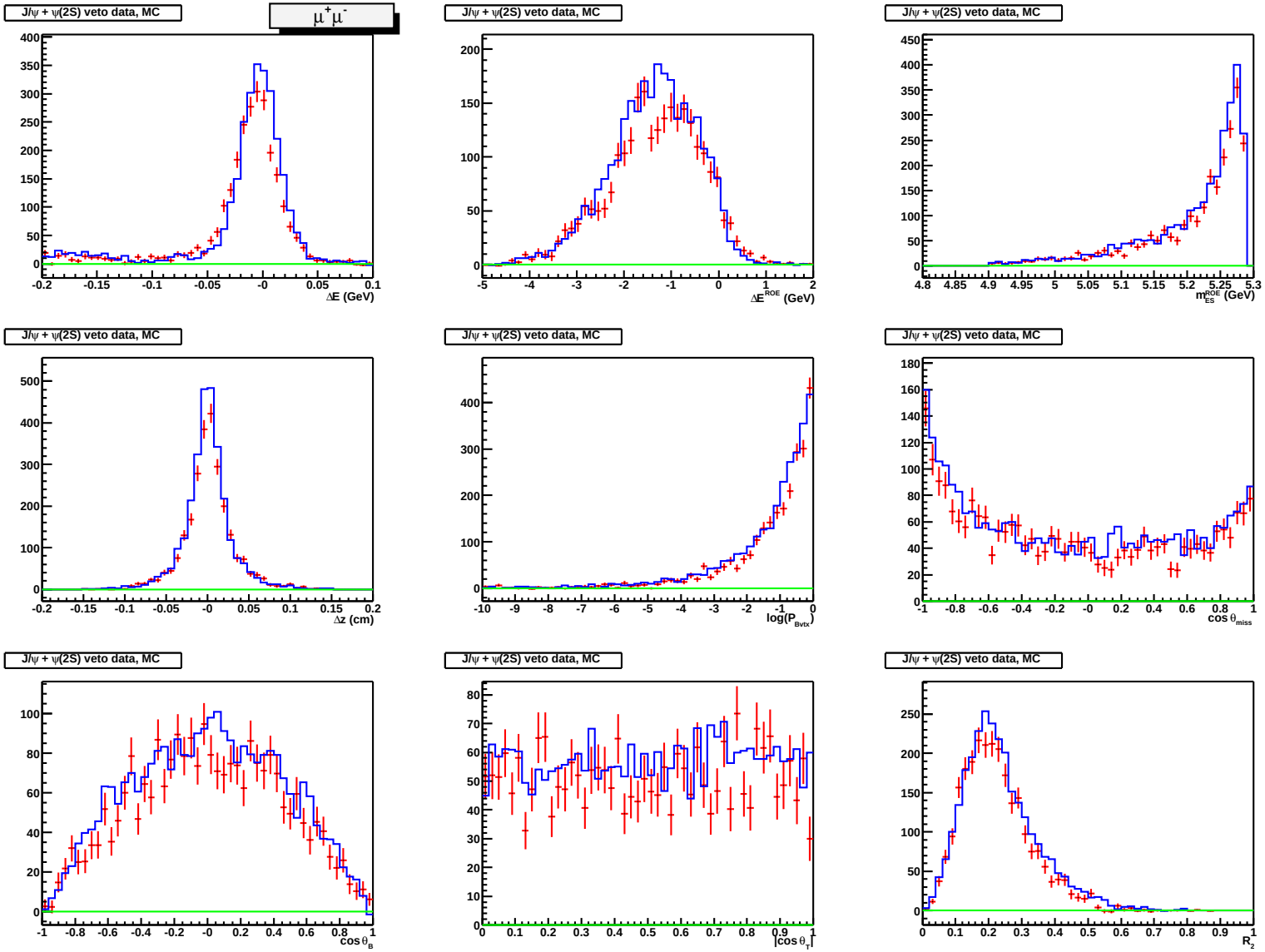

Figure 6.2: Distributions of the background-suppression variables entering the likelihood ratio for the background-subtracted charmonium events in the muon channel. The points represent the data and the histograms the MC, scaled to correspond to an integrated luminosity of 81.9 fb^{-1}.

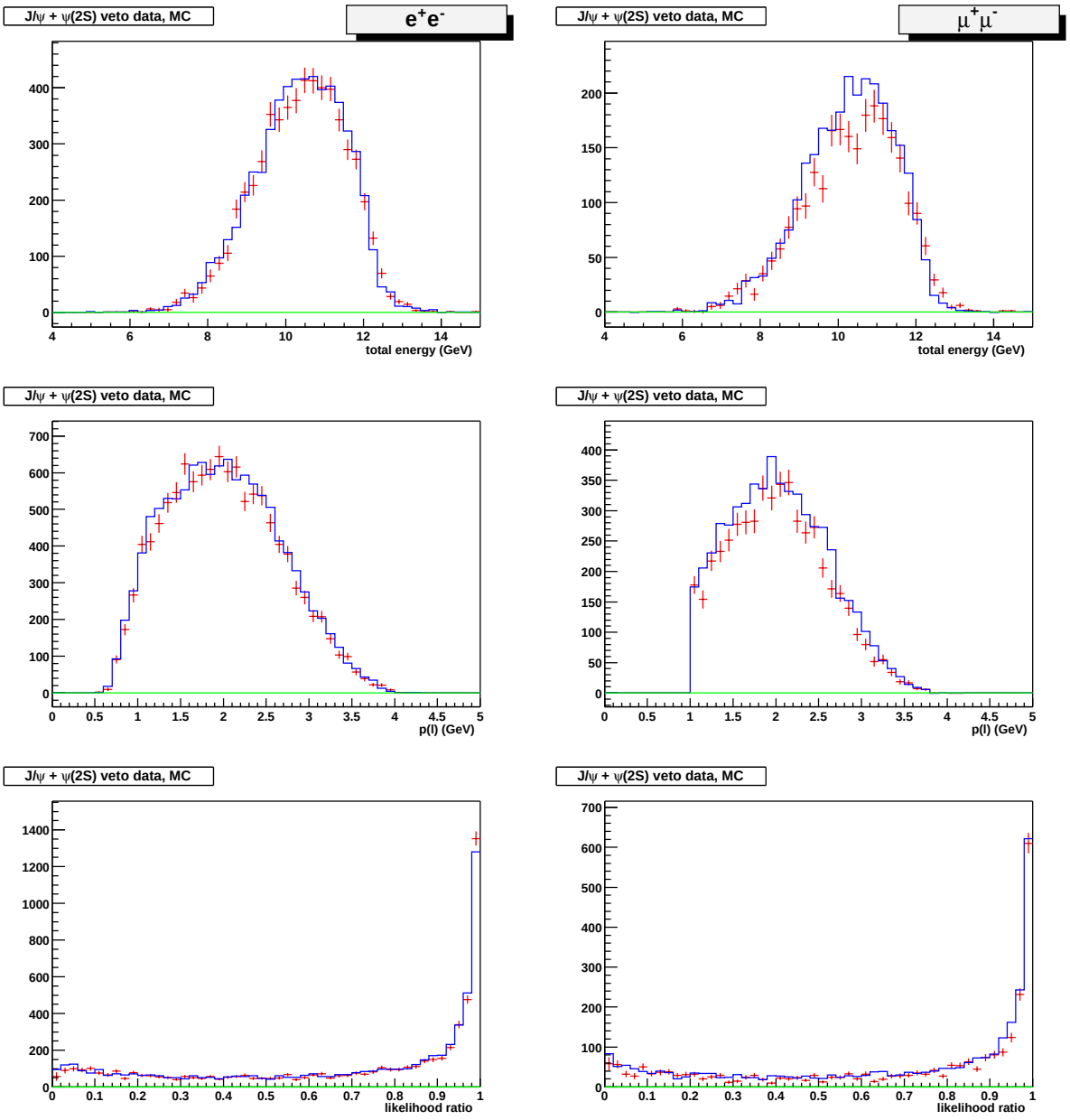

Figure 50: Distributions of total event energy, lepton momentum (both in the lab frame), and likelihood ratio for $B \to X_s \ell^+ \ell^-$ candidates in the charmonium veto sample for electrons (left) and muons (right) after background subtraction. The points represent the data and the histograms the MC, scaled to correspond to an integrated luminosity of 81.9 fb^{-1}.

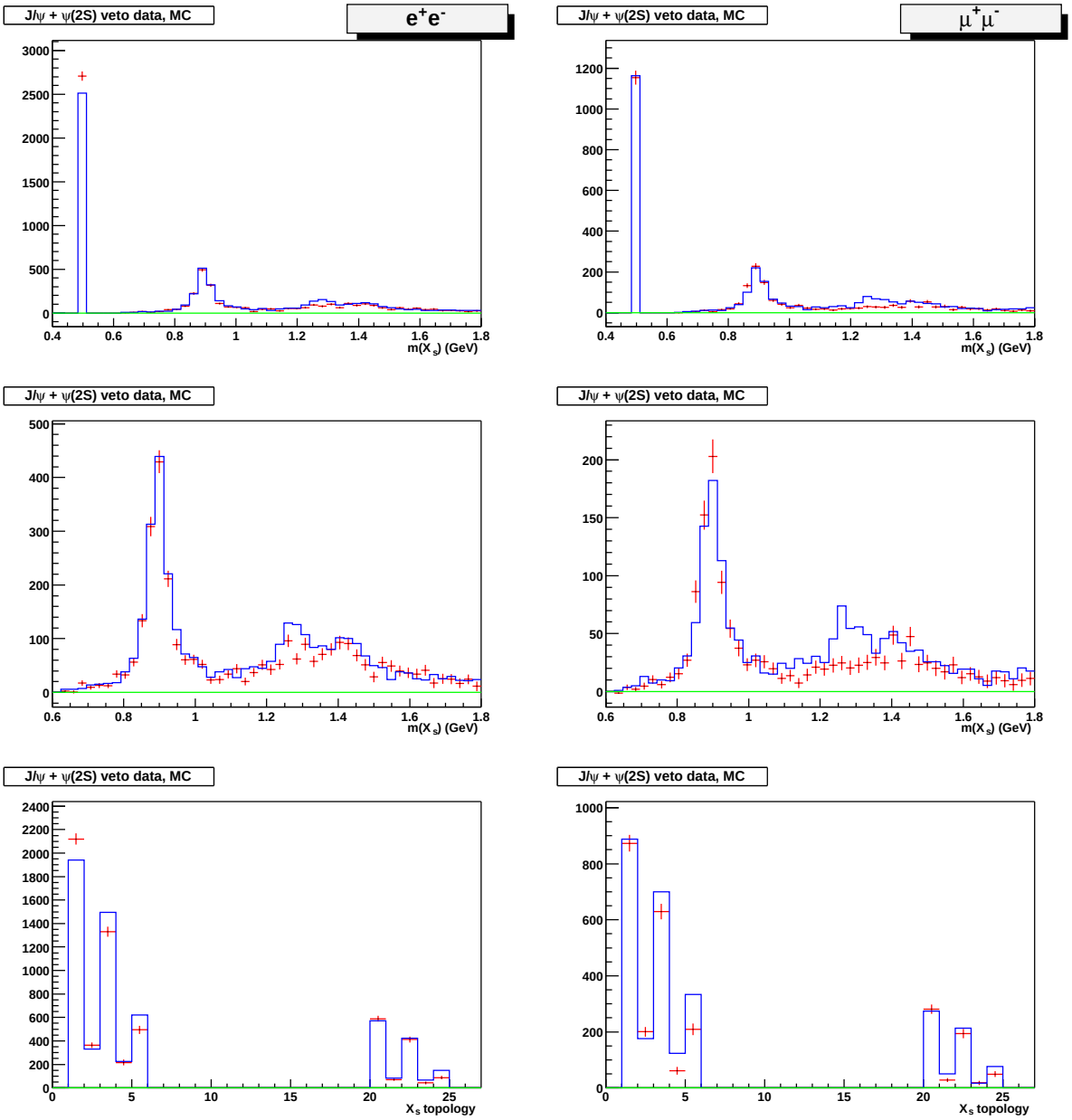

Figure 51: Distributions of hadronic invariant mass and final state topology (see Table 6) for $B \to X_s\, \ell^+\ell^-$ candidates in the charmonium veto sample for electrons (left) and muons (right) after background subtraction. The points represent the data and the histograms the MC, scaled to correspond to an integrated luminosity of 81.9 fb^{-1}.

Table 7: Ratio of background-subtracted $B \to J/\psi\, X$ yields for data and MC in the different topologies for the electron channels.

X_s Topology	Data/MC	Data/MC(corrected)	MC correction
K	1.118 ± 0.036	0.991 ± 0.061	1.128 ± 0.059
$K\pi^0$	1.16 ± 0.13	1.08 ± 0.16	1.07 ± 0.10
$K\pi$	1.013 ± 0.048	1.00 ± 0.09	1.01 ± 0.07
$K\pi\pi^0$	1.07 ± 0.24		
$K\pi\pi$	0.67 ± 0.08		
K_S^0	1.107 ± 0.069	1.14 ± 0.10	0.971 ± 0.058
$K_S^0\pi^0$	0.79 ± 0.20	0.78 ± 0.21	1.01 ± 0.07
$K_S^0\pi$	1.01 ± 0.09	0.94 ± 0.12	1.07 ± 0.10
$K_S^0\pi\pi^0$	0.64 ± 0.35		
$K_S^0\pi\pi$	0.49 ± 0.16		
All above $J/\psi X$ modes	1.035 ± 0.024		
$J/\psi X$ (0 or 1π)	1.078 ± 0.025	1.008 ± 0.041	

tween data and MC is most useful in the lower multiplicity $J/\psi\, K$ and $J/\psi\, K^*$ final states. Tables 7 and 8 indicate good agreement between data and MC. In the third column of these tables, the MC yields have been corrected to account for differences between branching fractions used in the MC generator and those listed in the PDG [2]. (Here, we assumed that the $K\pi$ and $K\pi^0$ modes are dominated by the K^* resonance.) Tables 11 and 12 compare the branching fractions used in the MC generator with those from the PDG [2]. The exclusive decay modes listed in the table are supplemented with inclusive quark processes hadronized via `Jetset`, the result of this process is tuned such as to reproduce the fully inclusive branching fraction, as shown in the last row of Tables 11 and 12. It should be noted that the corrected ratio data/MC for combined 0 and 1 pion modes given in Tables 7–10 is computed assuming the total uncertainties in the individual modes are uncorrelated. The uncertainty in that combined mode ratio is thus a lower estimate.

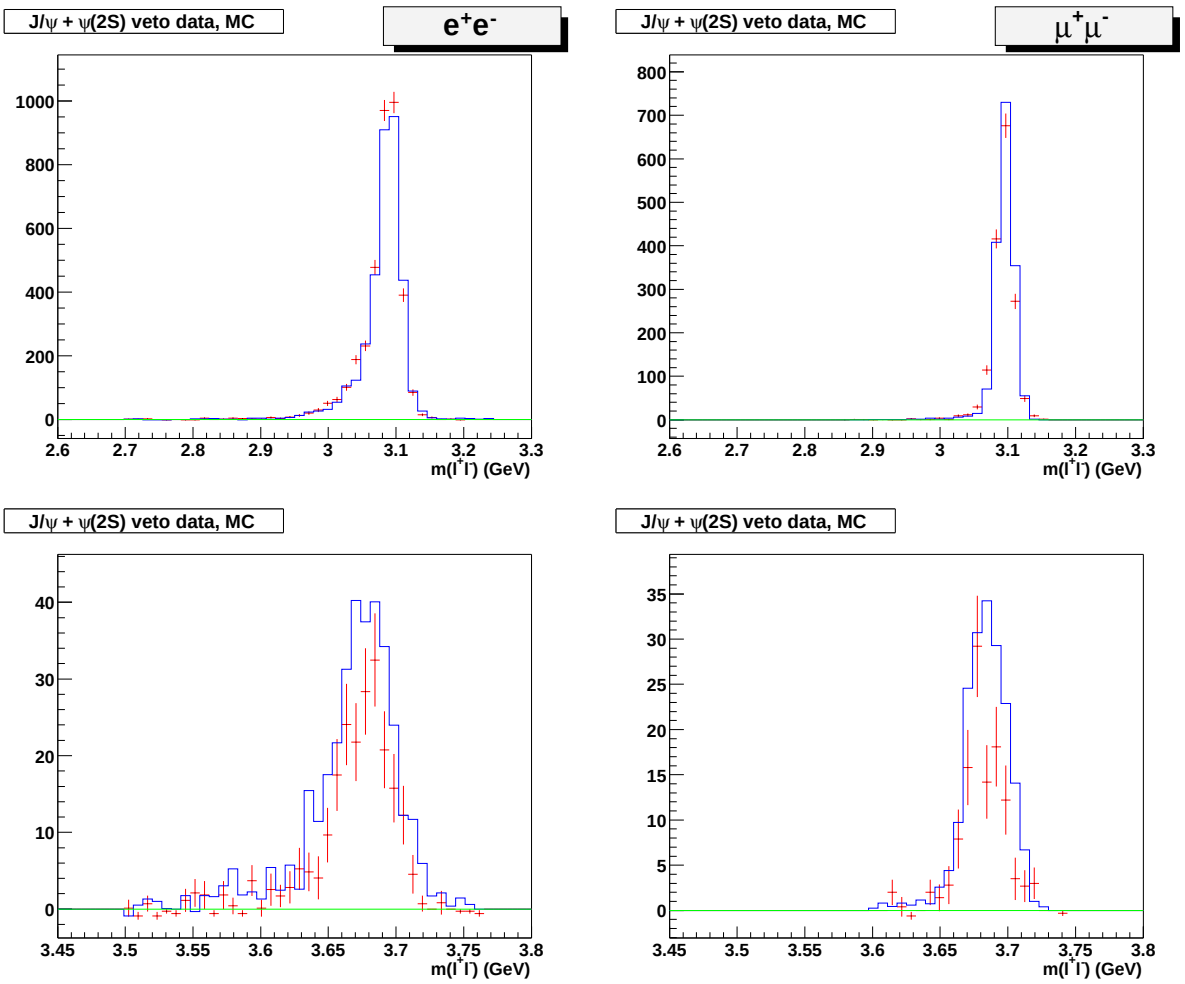

Figure 52: Distributions of dilepton invariant mass for $B \to X_s\, \ell^+\ell^-$ candidates in the charmonium veto sample for electrons (left) and muons (right) after likelihood ratio cuts and background subtraction. The top and bottom rows are for the J/ψ and $\psi(2S)$ mass regions, respectively. The points represent the data and the histograms the MC, scaled to correspond to an integrated luminosity of 81.9 fb^{-1}.

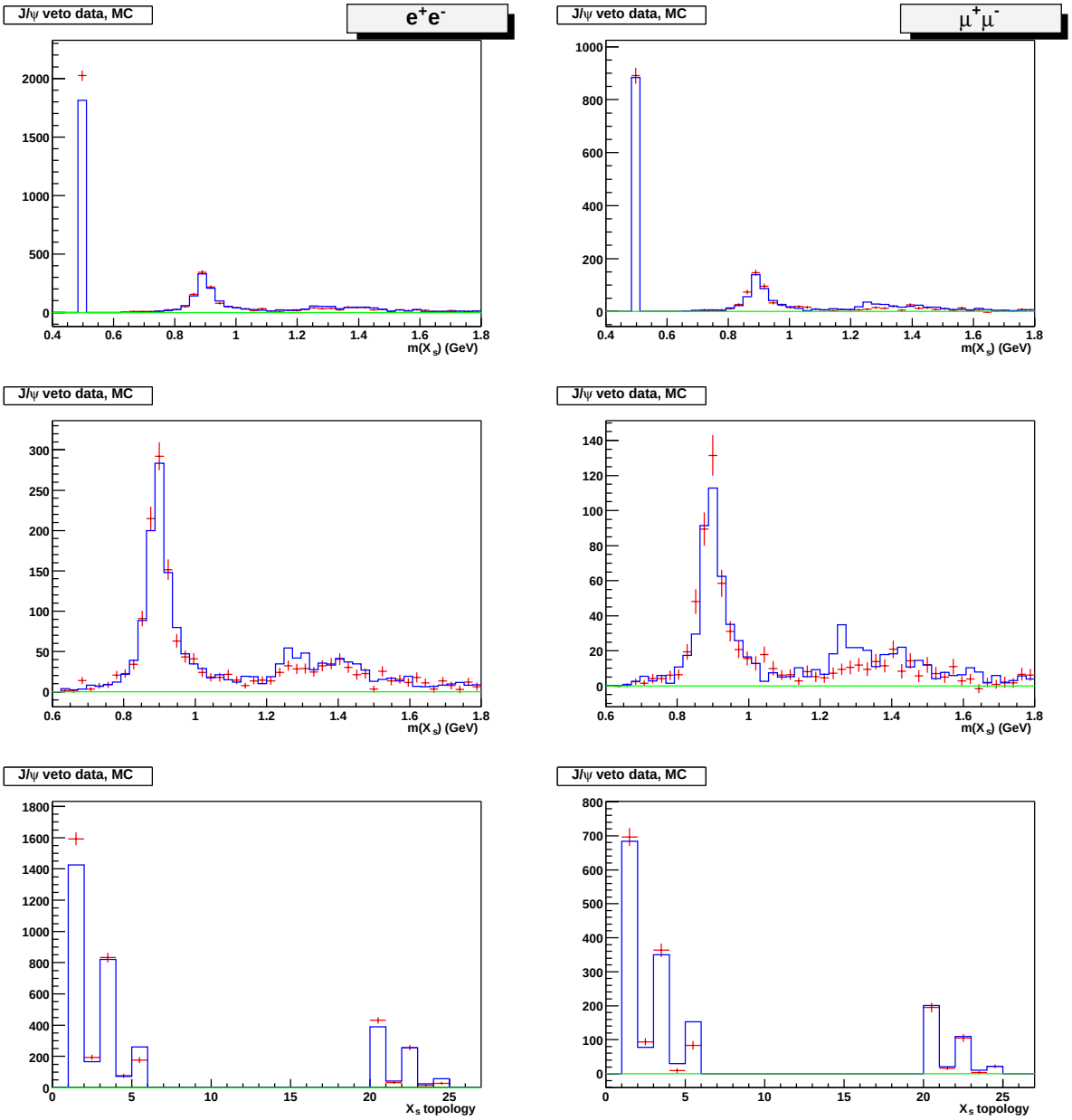

Figure 53: Distributions of hadronic invariant mass and final state topology (see Table 6) for $B \to X_s\, \ell^+\ell^-$ candidates in the $B \to J/\psi\, X$ veto sample for electrons (left) and muons (right) after likelihood ratio cuts and background subtraction. The points represent the data and the histograms the MC, scaled to correspond to an integrated luminosity of 81.9 fb^{-1}.

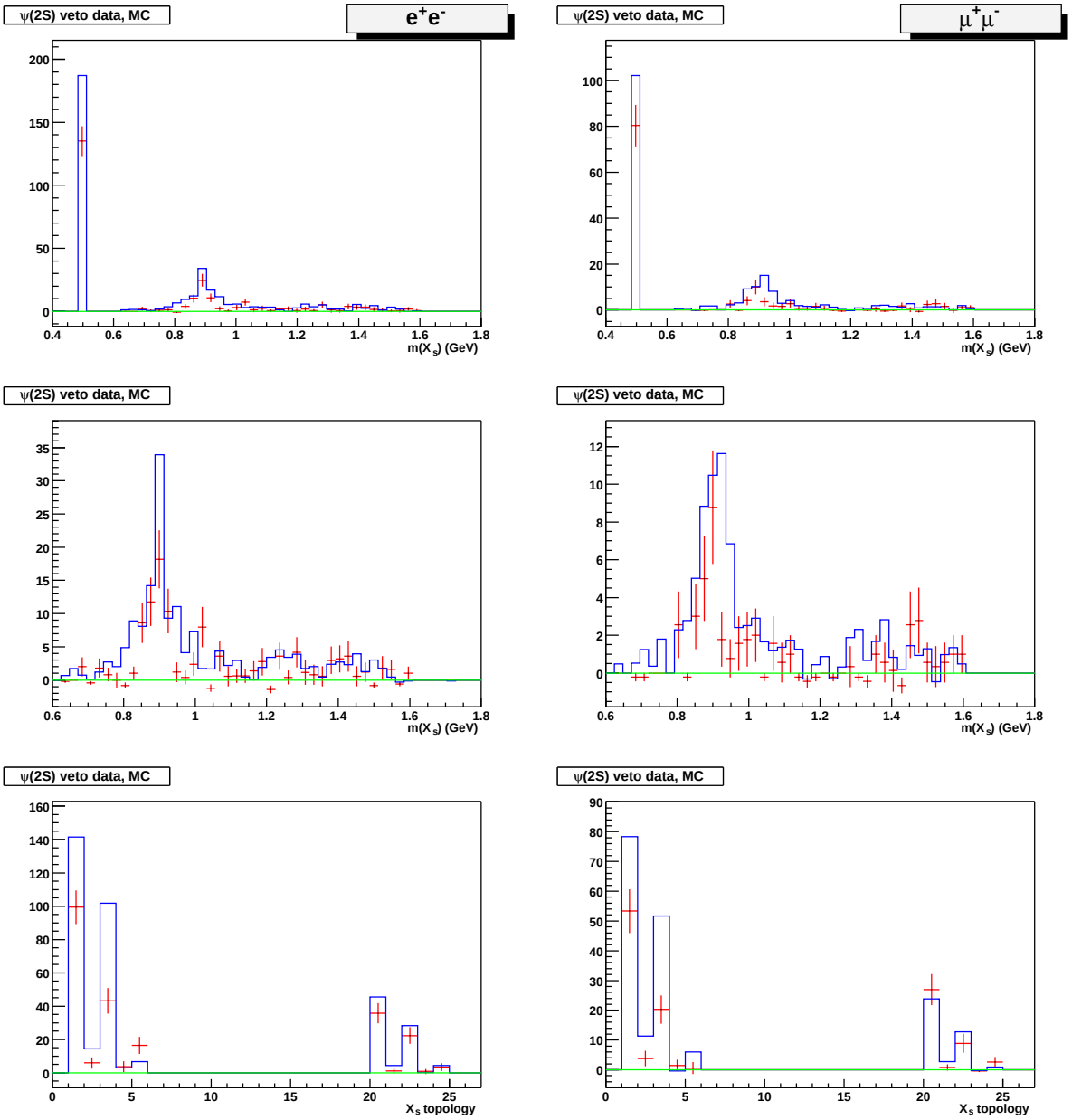

Figure 54: Distributions of hadronic invariant mass and final state topology (see Table 6) for $B \to X_s \ell^+ \ell^-$ candidates in the $B \to \psi(2S) X$ veto sample for electrons (left) and muons (right) after likelihood ratio cuts and background subtraction. The points represent the data and the histograms the MC, scaled to correspond to an integrated luminosity of 81.9 fb^{-1}.

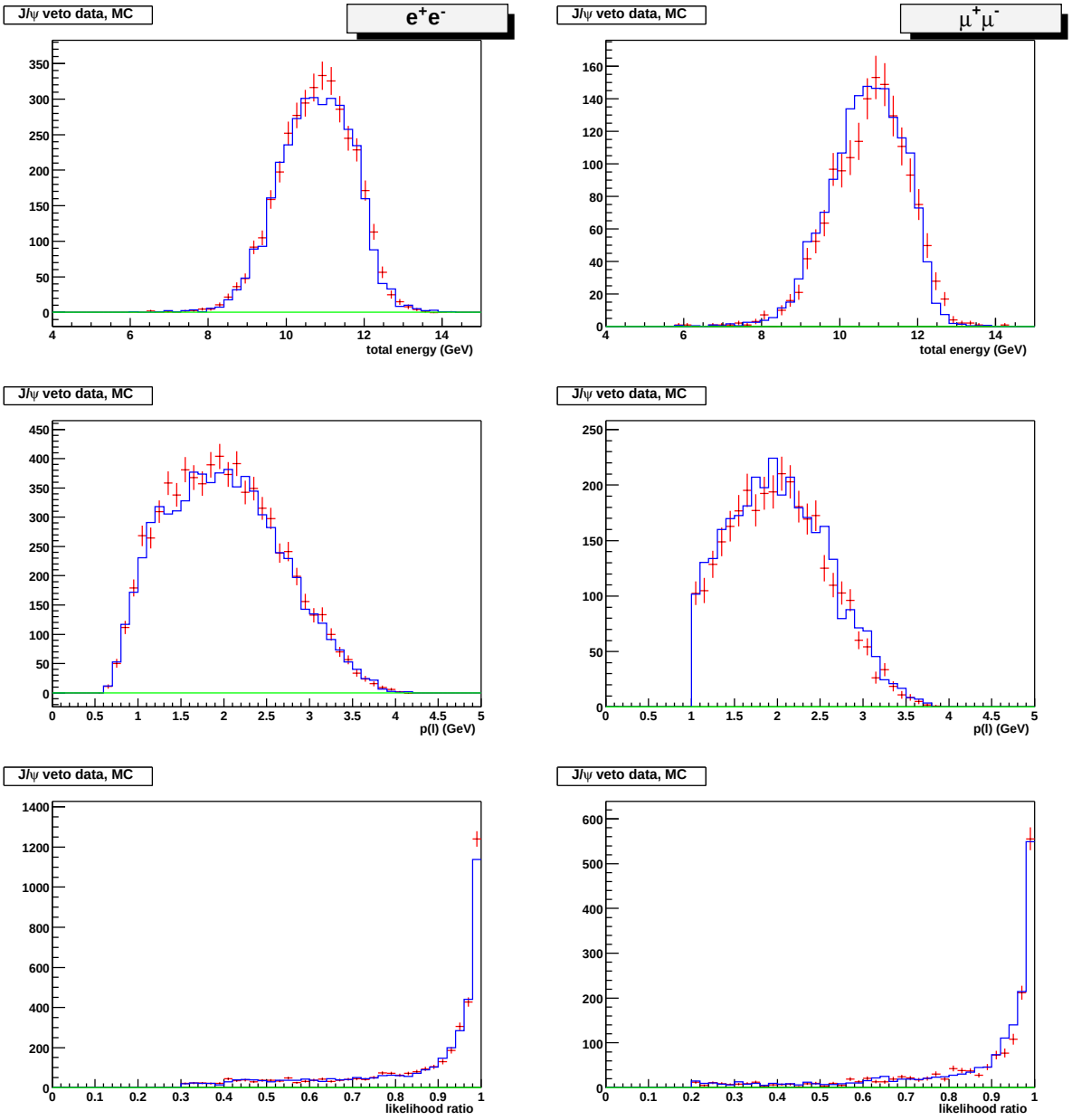

Figure 55: Distributions of total event energy, lepton momentum (both in the lab frame), and likelihood ratio for $B \to X_s\, \ell^+\ell^-$ candidates in the $B \to J/\psi\, X$ veto sample for electrons (left) and muons (right) after likelihood ratio cuts and background subtraction. The points represent the data and the histograms the MC, scaled to correspond to an integrated luminosity of 81.9 fb^{-1}.

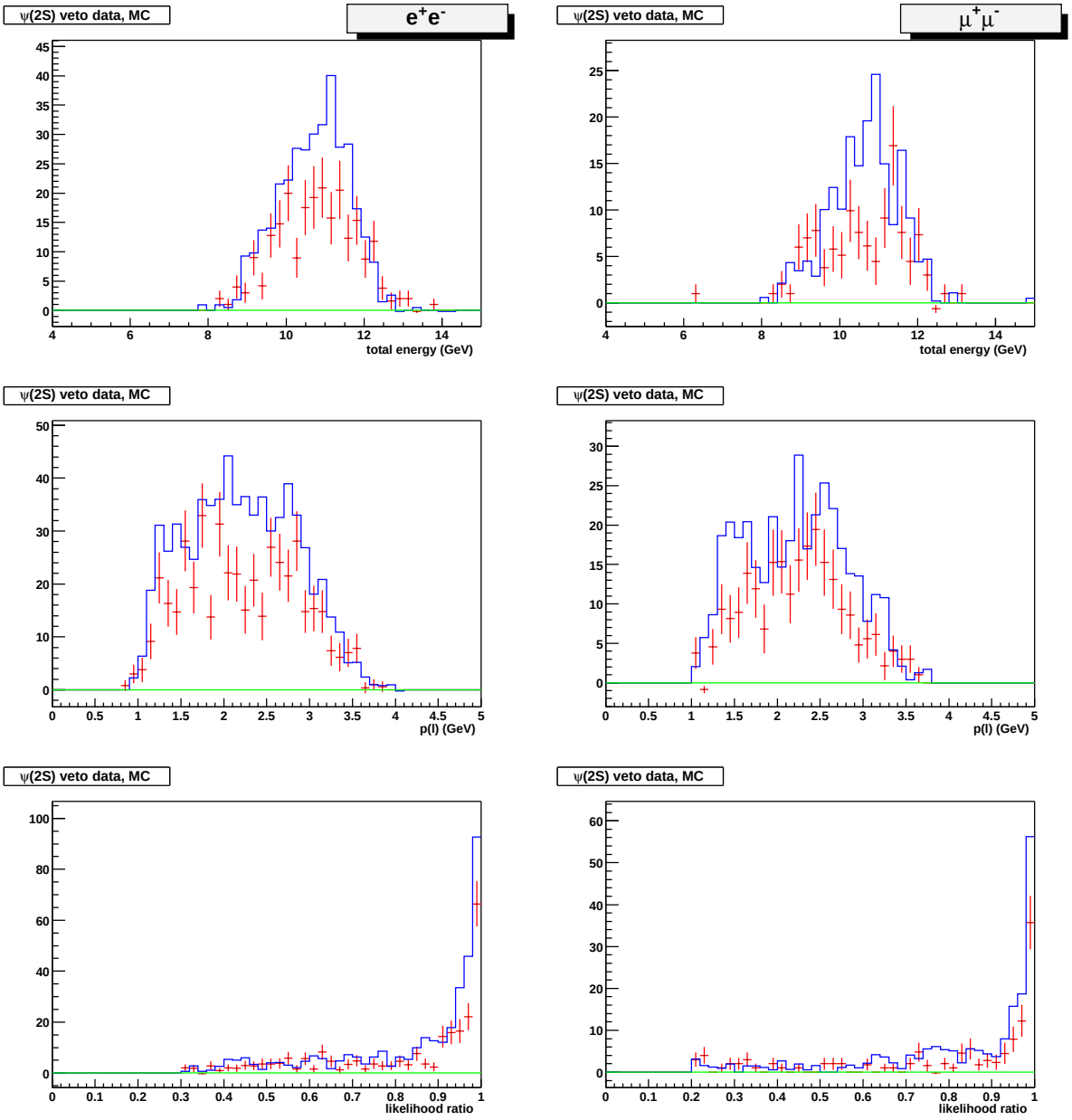

Figure 56: Distributions of total event energy, lepton momentum (both in the lab frame), and likelihood ratio for $B \to X_s \, \ell^+ \ell^-$ candidates in the $B \to \psi(2S) \, X$ veto sample for electrons (left) and muons (right) after likelihood ratio cuts and background subtraction. The points represent the data and the histograms the MC, scaled to correspond to an integrated luminosity of 81.9 fb^{-1}.

Table 8: Ratio of background-subtracted $B \to J/\psi\, X$ yields for data and MC in the different topologies for the muon channels.

X_s Topology	Data/MC	Data/MC(corrected)	MC correction
K	1.018 ± 0.048	0.910 ± 0.064	1.118 ± 0.059
$K\pi^0$	1.21 ± 0.18	1.14 ± 0.20	1.07 ± 0.10
$K\pi$	1.040 ± 0.072	1.04 ± 0.10	1.00 ± 0.07
$K\pi\pi^0$	0.31 ± 0.24		
$K\pi\pi$	0.54 ± 0.09		
K_S^0	0.97 ± 0.09	1.01 ± 0.11	0.963 ± 0.058
$K_S^0\pi^0$	0.78 ± 0.26	0.78 ± 0.27	1.00 ± 0.07
$K_S^0\pi$	0.96 ± 0.13	0.90 ± 0.15	1.07 ± 0.10
$K_S^0\pi\pi^0$	0.32 ± 0.42		
$K_S^0\pi\pi$	0.99 ± 0.33		
All above $J/\psi X$ modes	0.957 ± 0.032		
$J/\psi X$ (0 or 1π)	1.019 ± 0.034	0.961 ± 0.045	

Table 9: Ratio of background-subtracted $B \to \psi(2S)\, X$ yields for data and MC in the different topologies for the electron channels.

X_s Topology	Data/MC	Data/MC(corrected)	MC correction
K	0.70 ± 0.08	0.90 ± 0.15	0.78 ± 0.08
$K\pi^0$	0.41 ± 0.27	0.55 ± 0.38	0.75 ± 0.18
$K\pi$	0.42 ± 0.08	0.65 ± 0.17	0.65 ± 0.11
$K\pi\pi^0$	1.3 ± 1.4		
$K\pi\pi$	2.4 ± 1.5		
K_S^0	0.79 ± 0.16	1.17 ± 0.32	0.67 ± 0.12
$K_S^0\pi^0$	0.27 ± 0.35	0.41 ± 0.55	0.65 ± 0.11
$K_S^0\pi$	0.79 ± 0.22	1.06 ± 0.39	0.75 ± 0.18
$K_S^0\pi\pi^0$	insuff.stats.		
$K_S^0\pi\pi$	0.79 ± 0.65		
All above $\psi(2S)X$ modes	0.663 ± 0.056		
$\psi(2S)X$ (0 or 1π)	0.619 ± 0.053	0.86 ± 0.10	

Table 10: Ratio of background-subtracted $B \to \psi(2S)\, X$ yields for data and MC in the different topologies for the muon channels.

X_s Topology	Data/MC	Data/MC(corrected)	MC correction
K	0.68 ± 0.11	0.91 ± 0.21	0.75 ± 0.12
$K\pi^0$	0.33 ± 0.24	0.47 ± 0.36	0.72 ± 0.19
$K\pi$	0.39 ± 0.10	0.63 ± 0.21	0.62 ± 0.13
$K\pi\pi^0$	insuff.stats.		
$K\pi\pi$	0.10 ± 0.34		
K_S^0	1.13 ± 0.28	1.76 ± 0.58	0.64 ± 0.14
$K_S^0\pi^0$	0.28 ± 0.40	0.46 ± 0.65	0.62 ± 0.13
$K_S^0\pi$	0.70 ± 0.30	0.97 ± 0.49	0.72 ± 0.19
$K_S^0\pi\pi^0$	insuff.stats.		
$K_S^0\pi\pi$	insuff.stats.		
All above $\psi(2S)X$ modes	0.633 ± 0.071		
$\psi(2S)X$ (0 or 1π)	0.631 ± 0.071	0.91 ± 0.14	

6.2 $X_s e^{\pm}\mu^{\mp}$ sample

Since $X_s\, e^{\pm}\mu^{\mp}$ combinations cannot result from a B decay, they do not peak in m_{ES} and therefore, are a good control sample of combinatorial background events. The sample is collected in parallel with the standard analysis described above, the only difference being the requirement for a different flavor of the dilepton candidates (1 e and 1 μ). For the charmonium veto and the calculation of the likelihood ratio, we treat the $X_s\, e^{\pm}\mu^{\mp}$ candidate as if it were an $X_s\, \mu^+\mu^-$ candidate. Final states with three pions are rejected and nominal likelihood ratio cuts are applied.

Distributions of $m_{\rm ES}$ for data and MC, fitted with an Argus function, are shown in Fig. 57. The fitted parameters obtained from the two distributions are consistent within their errors. Other data and MC comparisons are shown in Fig. 58. The agreement is good and the ratio between the number of events in the data and MC is 0.986 ± 0.050, which shows that the combinatorial background is simulated accurately and no additional scaling is necessary. It should be noted that the

Table 11: Branching fractions for B decays to J/ψ in SP4 MC (EvtGen) and data (PDG) [2]. The last entry gives the fully inclusive branching fraction at the $\Upsilon(4S)$ where the MC uncertainty corresponds to the statistical uncertainty obtained in a large sample of generated events.

Decay mode	EvtGen MC	PDG
$B^0 \to J/\psi K^0$	9.0×10^{-4}	$(8.7 \pm 0.5) \times 10^{-4}$
$B^0 \to J/\psi K^{*0}$	1.3×10^{-3}	$(1.31 \pm 0.09) \times 10^{-3}$
$B^0 \to J/\psi K^+ \pi^-$	2.0×10^{-4}	
$B^0 \to J/\psi K^0 \pi^0$	1.0×10^{-4}	
$B^0 \to J/\psi \phi K^0$	9.0×10^{-5}	$(8.8^{+3.7}_{-3.3}) \times 10^{-5}$
$B^0 \to J/\psi K_1(1270)^0$	1.5×10^{-3}	$(1.3 \pm 0.5) \times 10^{-3}$
$B^0 \to J/\psi K_1(1400)^0$	1.0×10^{-4}	
$B^0 \to J/\psi K_2^*(1430)^0$	5.0×10^{-4}	
$B^0 \to J/\psi K^0 \pi^+ \pi^-$		$(1.0 \pm 0.4) \times 10^{-4}$
$B^0 \to J/\psi K^{*+} \pi^-$		$(0.8 \pm 0.4) \times 10^{-3}$
$B^0 \to J/\psi K^{*0} \pi^+ \pi^-$		$(6.6 \pm 2.2) \times 10^{-4}$
$B^0 \to J/\psi \pi^0$	2.0×10^{-5}	$(2.1 \pm 0.5) \times 10^{-4}$
$B^0 \to J/\psi \rho^0$	3.0×10^{-5}	$< 2.5 \times 10^{-4}$
$B^0 \to J/\psi \omega$	3.0×10^{-5}	$< 2.7 \times 10^{-4}$
$B^+ \to J/\psi K^+$	9.0×10^{-4}	$(1.01 \pm 0.05) \times 10^{-3}$
$B^+ \to J/\psi K^{*+}$	1.3×10^{-3}	$(1.39 \pm 0.13) \times 10^{-3}$
$B^+ \to J/\psi K^+ \pi^0$	1.0×10^{-4}	
$B^+ \to J/\psi K^0 \pi^+$	2.0×10^{-4}	
$B^+ \to J/\psi \phi K^+$	9.0×10^{-5}	$(8.8^{+3.7}_{-3.3}) \times 10^{-5}$
$B^+ \to J/\psi K_1(1270)^+$	1.5×10^{-3}	$(1.8 \pm 0.5) \times 10^{-3}$
$B^+ \to J/\psi K_1(1400)^+$	1.0×10^{-4}	$< 5 \times 10^{-4}$
$B^+ \to J/\psi K_2^*(1430)^+$	5.0×10^{-4}	
$B^+ \to J/\psi K^+ \pi^+ \pi^-$		$(1.4 \pm 0.6) \times 10^{-3}$
$B^+ \to J/\psi \pi^+$	4.0×10^{-5}	$(4.2 \pm 0.7) \times 10^{-5}$
$B^+ \to J/\psi \rho^+$	6.0×10^{-5}	$< 7.7 \times 10^{-4}$
$B \to J/\psi X$	$(1.10 \pm 0.02) \times 10^{-2}$	$(1.15 \pm 0.06) \times 10^{-2}$

Table 12: Branching fractions for B decays to $\psi(2S)$ in SP4 MC (EvtGen) and data (PDG) [2]. The last entry gives the fully inclusive branching fraction at the $\Upsilon(4S)$ where the MC uncertainty corresponds to the statistical uncertainty obtained in a large sample of generated events.

Decay mode	EvtGen MC	PDG
$B^0 \to \psi(2S)K^0$	6.2×10^{-4}	$(5.7 \pm 1.0) \times 10^{-4}$
$B^0 \to \psi(2S)K^{*0}$	9.0×10^{-4}	$(8.0 \pm 1.3) \times 10^{-4}$
$B^0 \to \psi(2S)K^+\pi^-$	4.0×10^{-4}	
$B^0 \to \psi(2S)K^0\pi^0$	2.0×10^{-4}	
$B^0 \to \psi(2S)K^0\pi^+\pi^-$	2.0×10^{-4}	
$B^0 \to \psi(2S)K^0\pi^0\pi^0$	1.0×10^{-4}	
$B^0 \to \psi(2S)K^+\pi^-\pi^0$	1.0×10^{-4}	
$B^0 \to \psi(2S)K_1(1270)^0$	4.0×10^{-4}	
$B^+ \to \psi(2S)K^+$	6.2×10^{-4}	$(6.6 \pm 0.6) \times 10^{-4}$
$B^+ \to \psi(2S)K^{*+}$	9.0×10^{-4}	$(9.2 \pm 2.2) \times 10^{-4}$
$B^+ \to \psi(2S)K^+\pi^0$	2.0×10^{-4}	
$B^+ \to \psi(2S)K^0\pi^+$	4.0×10^{-4}	
$B^+ \to \psi(2S)K^+\pi^+\pi^-$	2.0×10^{-4}	$(1.9 \pm 1.2) \times 10^{-3}$
$B^+ \to \psi(2S)K^+\pi^0\pi^0$	1.0×10^{-4}	
$B^+ \to \psi(2S)K^0\pi^+\pi^0$	1.0×10^{-4}	
$B^+ \to \psi(2S)K_1(1270)^+$	4.0×10^{-4}	
$B \to \psi(2S)X$	$(3.2 \pm 0.1) \times 10^{-3}$	$(3.5 \pm 0.5) \times 10^{-3}$

Table 13: Branching fractions for charmonium decays in SP4 MC (EvtGen) and data (PDG) [2].

Decay mode	EvtGen MC	PDG
$J/\psi \to e^+e^-$	5.90×10^{-2}	$(5.93 \pm 0.10) \times 10^{-2}$
$J/\psi \to \mu^+\mu^-$	5.90×10^{-2}	$(5.88 \pm 0.10) \times 10^{-2}$
$\psi(2S) \to e^+e^-$	1.00×10^{-2}	$(0.73 \pm 0.04) \times 10^{-2}$
$\psi(2S) \to \mu^+\mu^-$	1.00×10^{-2}	$(0.70 \pm 0.09) \times 10^{-2}$

fit for the signal yield does not rely on this sample to describe the shape of the combinatorial background.

6.3 Off-resonance sample

In order to cross-check the combinatorial background contribution due to continuum events, as obtained from MC, we use off-resonance data, taken 40 MeV below the $\Upsilon(4S)$ peak. When comparing the data with MC, the reconstructed m_{ES} value in the MC is decreased by 20 MeV and all analysis cuts are applied. Data and MC are compared in Figs. 59 and 60. Agreement is good but statistics are poor. The ratio between the number of events in the data and MC is 1.78 ± 0.86 in the electron channel and 0.57 ± 0.42 in the muon channel.

6.4 On-resonance sideband sample

On-resonance sideband data, with $m_{ES} < 5.27$ GeV provides another check of the combinatorial background. Comparison between data and MC shows good agreement, see Figs. 61–6.4. The ratio between the number of events in the data and MC is 0.96 ± 0.08 in the electron channel and 1.17 ± 0.13 in the muon channel.

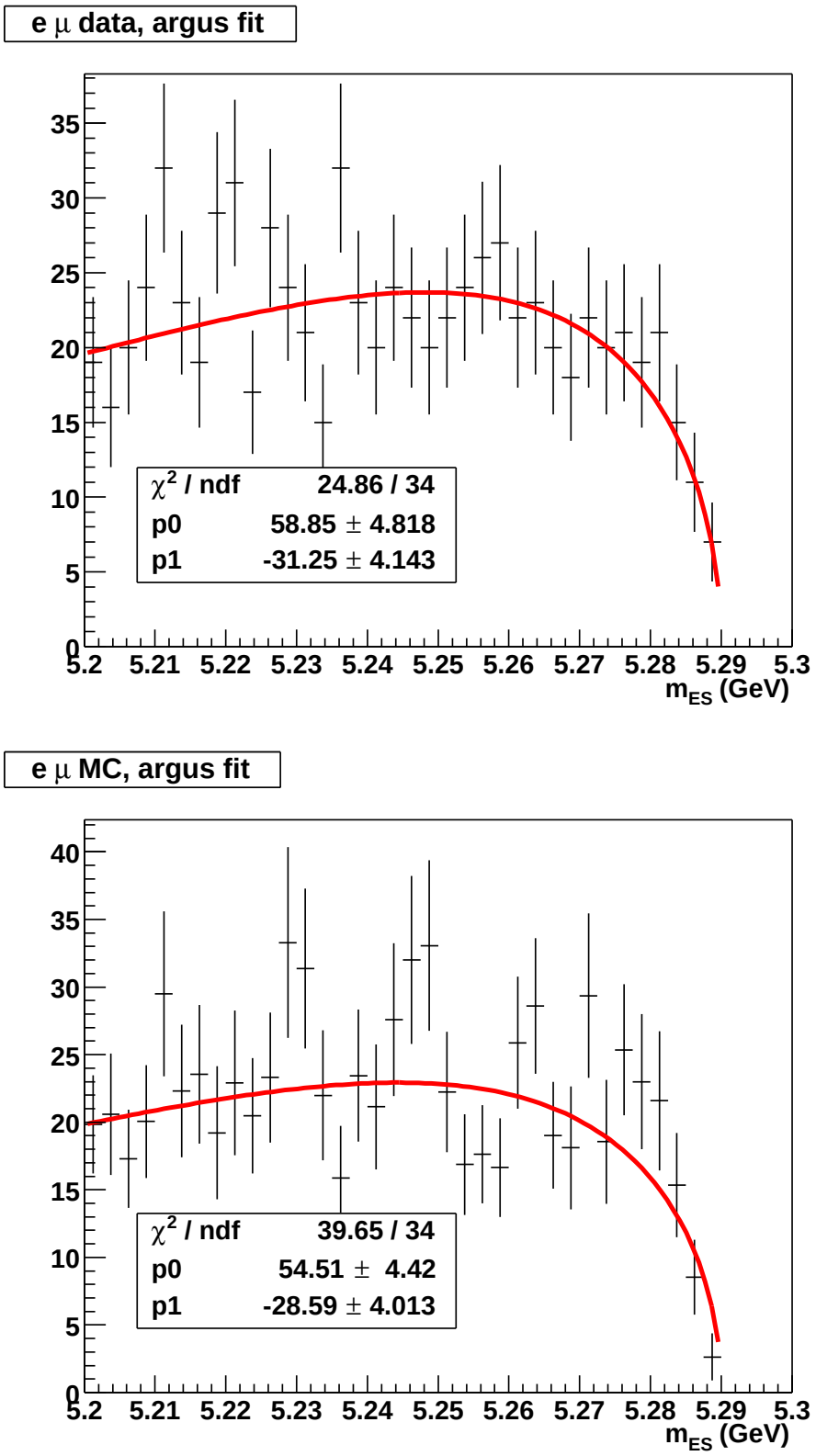

Figure 57: Distributions of m_{ES} for B decay candidates in the $B \to X_s\, e^{\pm}\mu^{\mp}$ sample for data (top) and MC (bottom). The MC distribution has been scaled to correspond to an integrated luminosity of 81.9 fb^{-1}.

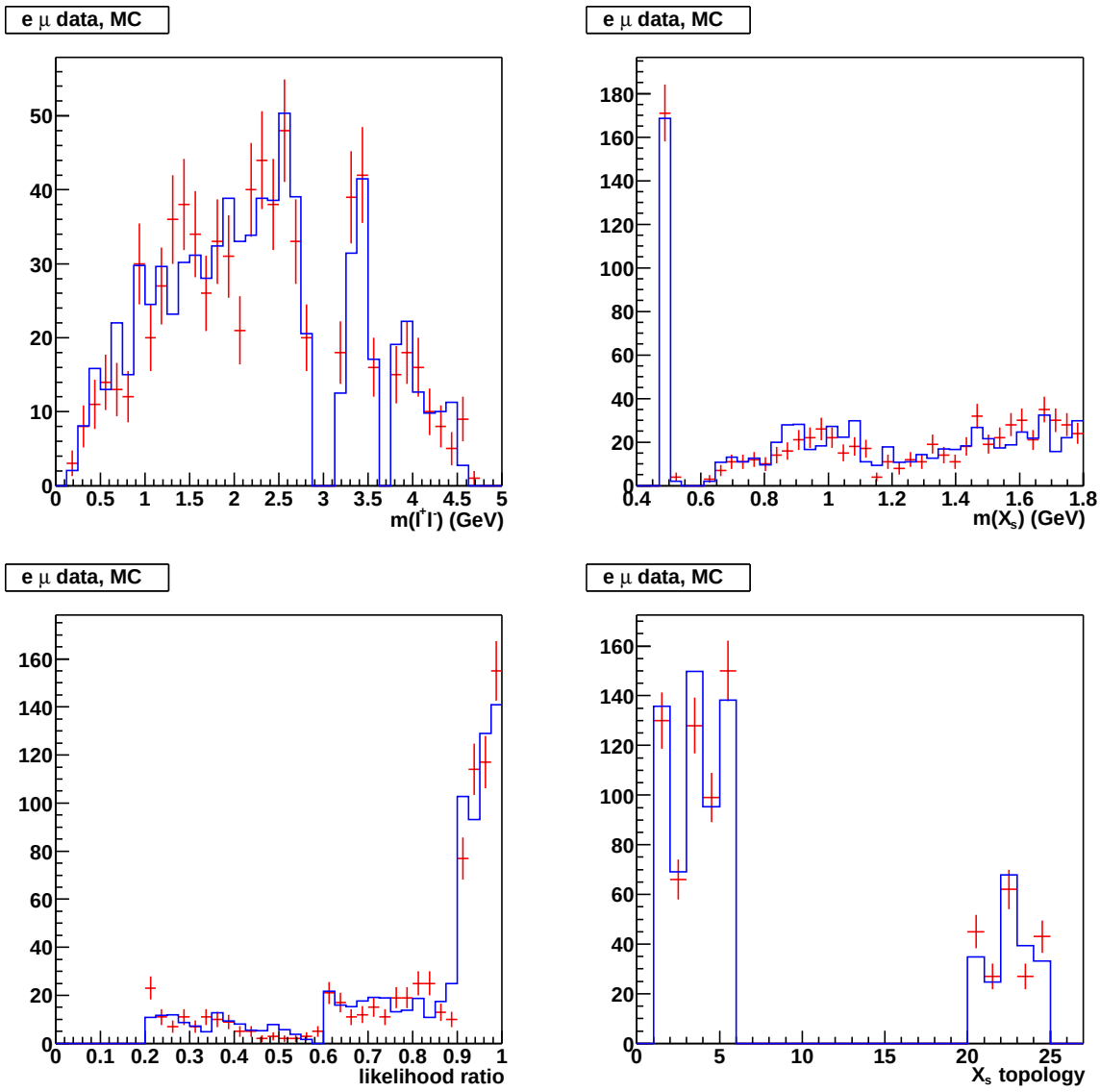

Figure 58: Distributions of dilepton mass, hadronic mass, likelihood ratio and final state topology for B decay candidates in the $B \to X_s\, e^{\pm}\mu^{\mp}$ sample. The points represent the data and the histograms the MC, scaled to correspond to an integrated luminosity of 81.9 fb^{-1}.

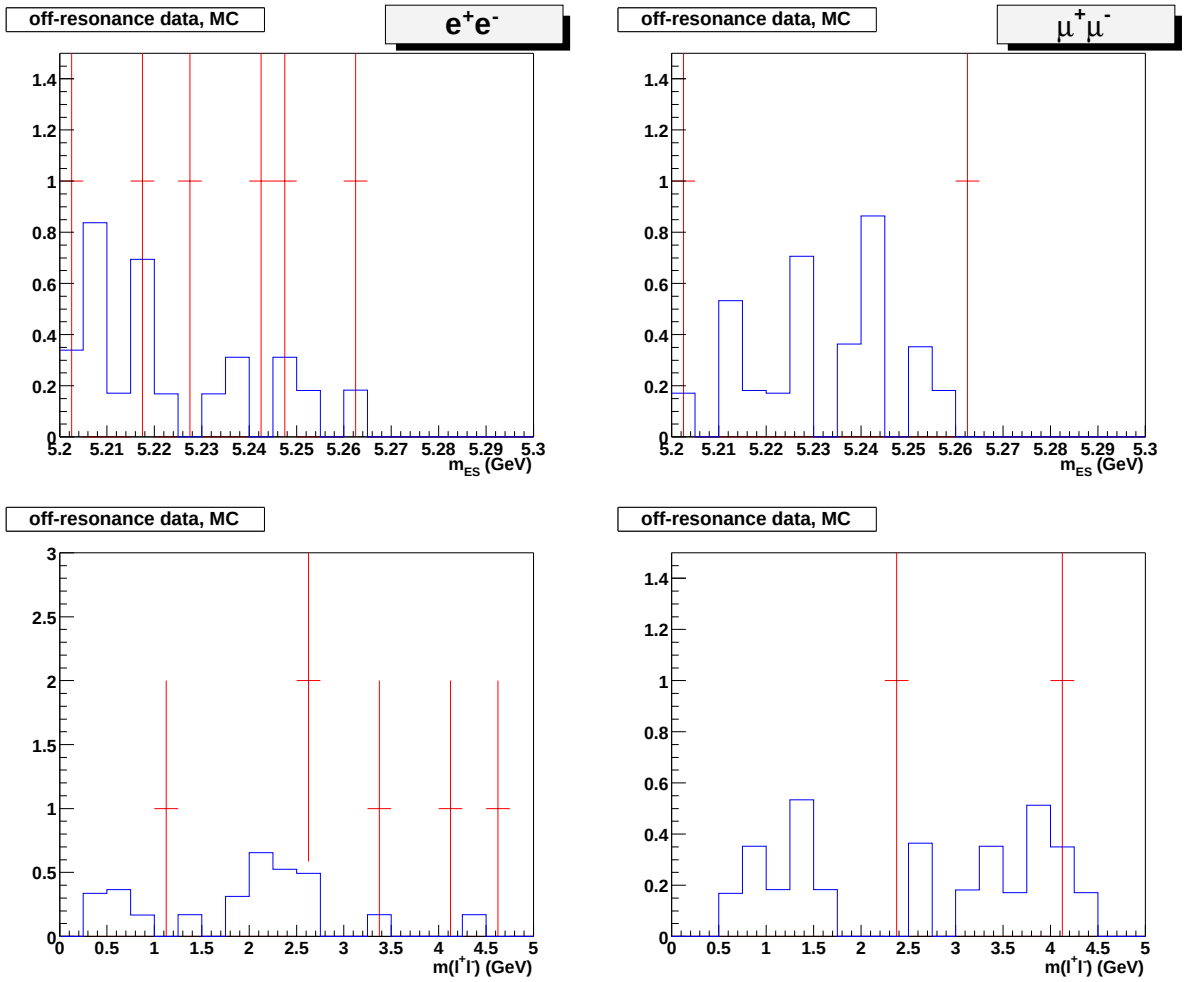

Figure 59: Distributions of m_{ES} and dilepton mass for $B \to X_s\, \ell^+\ell^-$ candidates in the off-resonance data sample for electrons (left) and muons (right). The points represent the data and the histograms the MC, scaled to correspond to an integrated luminosity of 9.6 fb^{-1}.

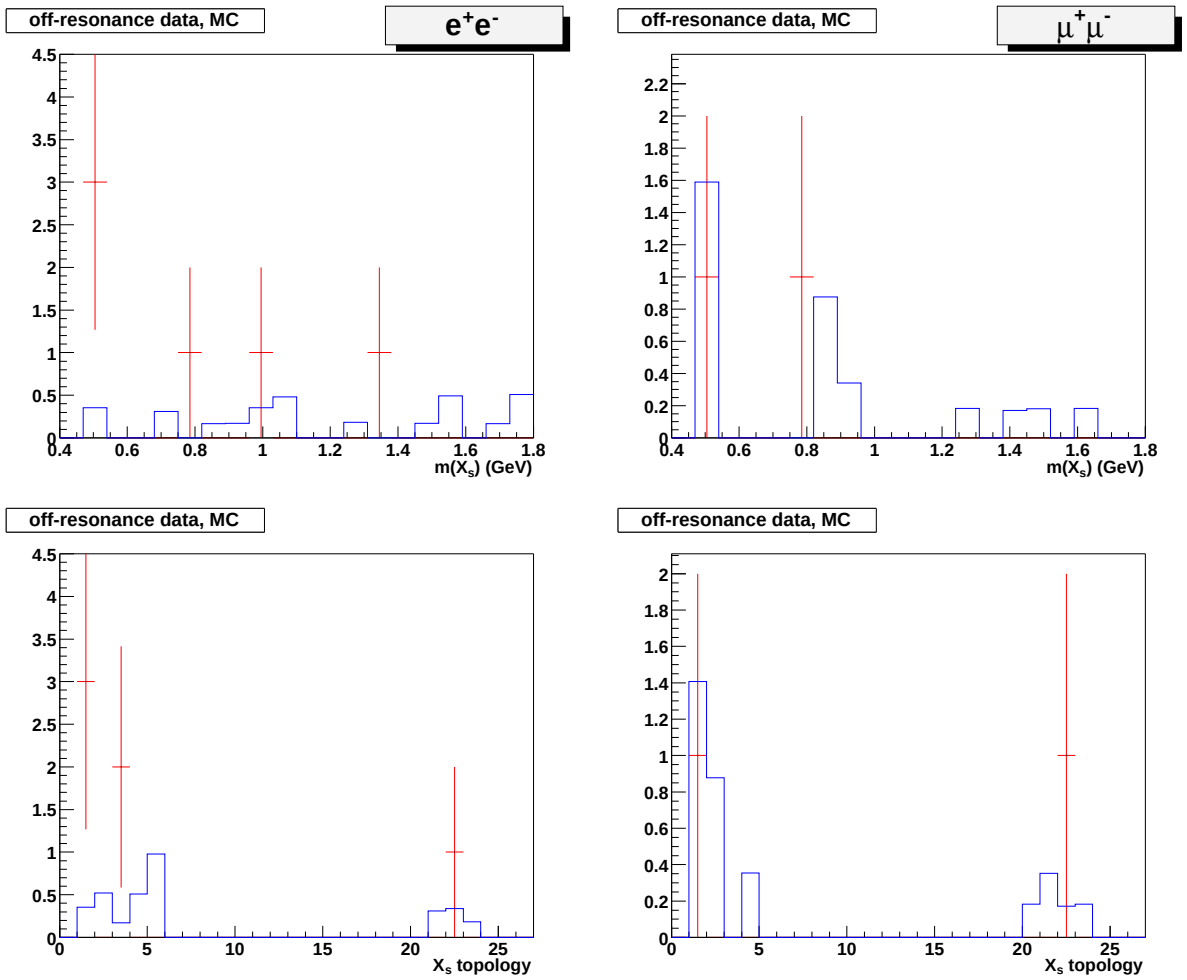

Figure 60: Distributions of hadronic mass and final state topology for $B \to X_s\, \ell^+\ell^-$ candidates in the off-resonance data sample for electrons (left) and muons (right). The points represent the data and the histograms the MC, scaled to correspond to an integrated luminosity of 9.6 fb^{-1}.

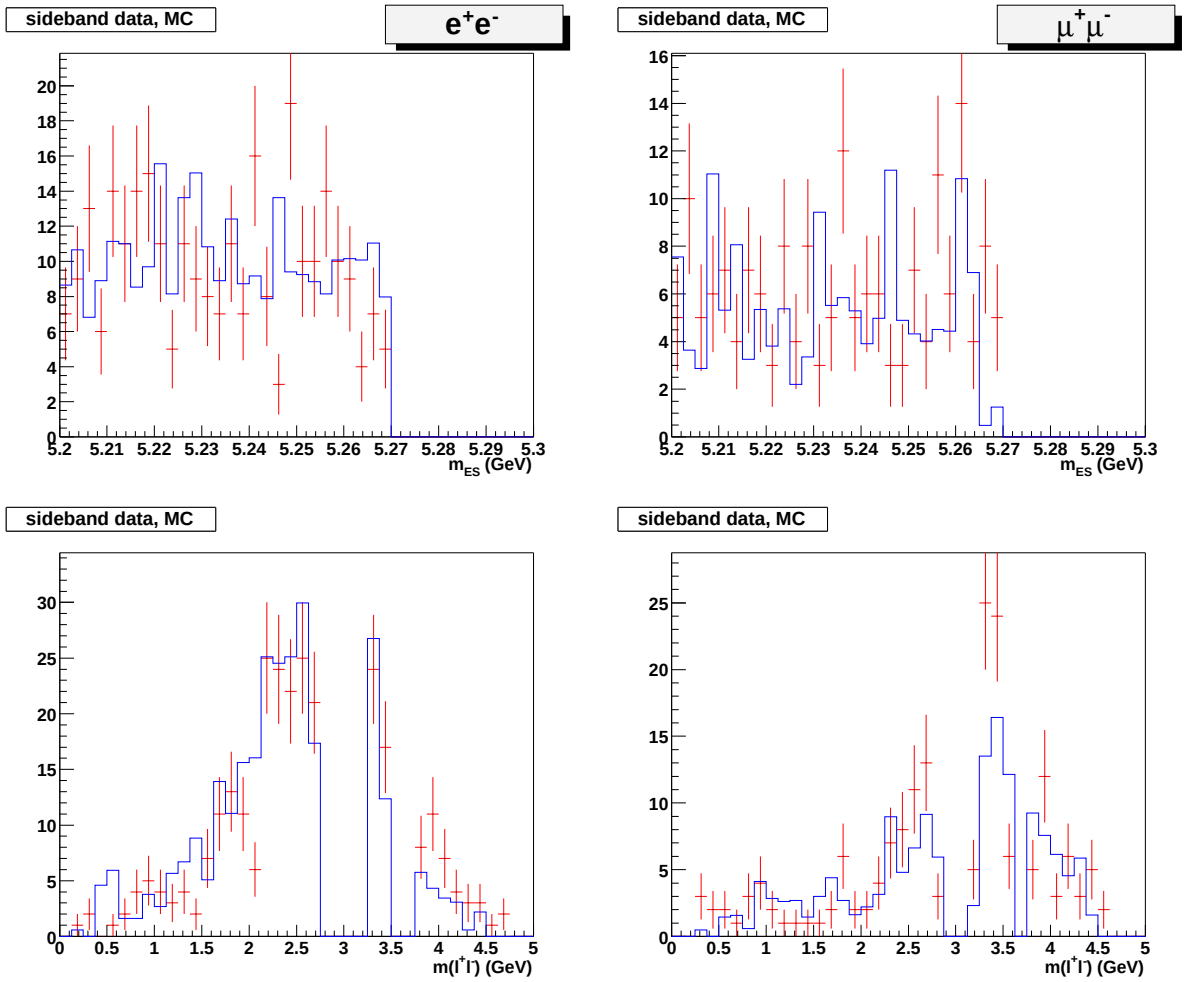

Figure 61: Distributions of m_{ES} and dilepton mass for $B \to X_s \ell^+ \ell^-$ candidates in the sideband data sample for electrons (left) and muons (right). The points represent the data and the histograms the MC, scaled to correspond to an integrated luminosity of 81.9 fb^{-1}.

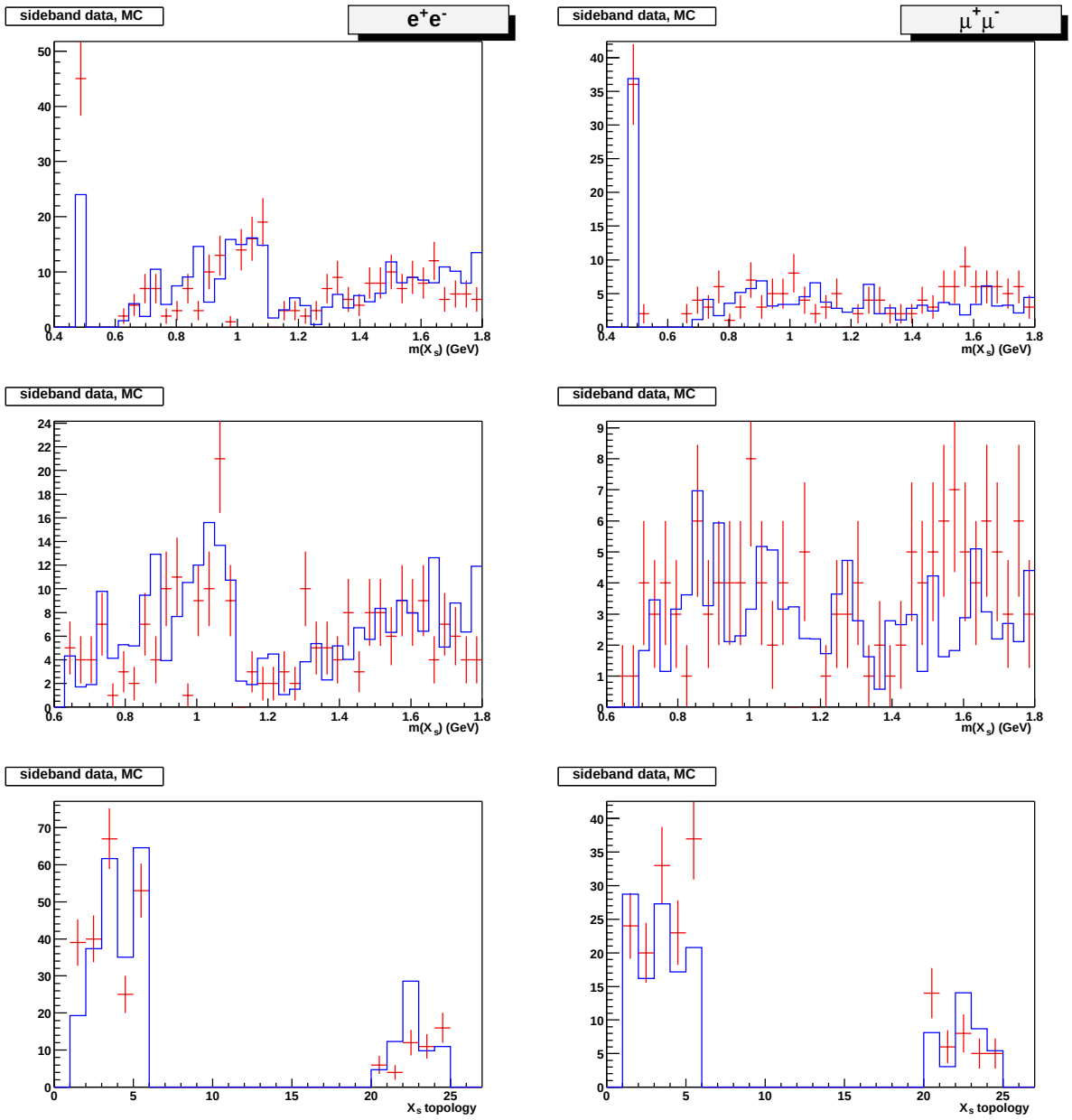

Figure 62: Distributions of hadronic mass and final state topology for $B \to X_s\, \ell^+\ell^-$ candidates in the sideband data sample for electrons (left) and muons (right). The points represent the data and the histograms the MC, scaled to correspond to an integrated luminosity of 81.9 fb^{-1}.

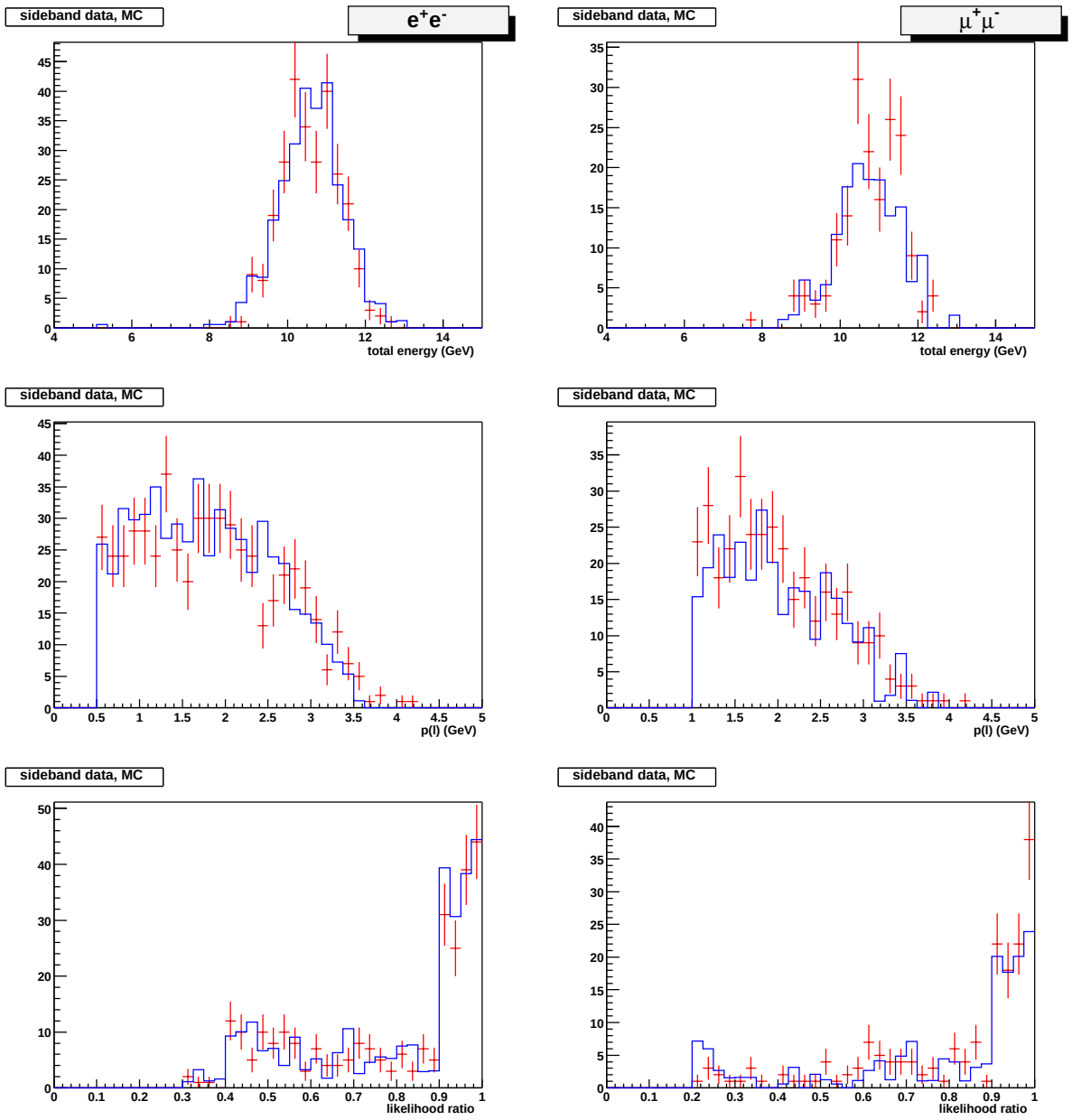

Figure 63: Distributions of total event energy, lepton momentum (both in the lab frame) and likelihood ratio for $B \to X_s \ell^+\ell^-$ candidates in the sideband data sample for electrons (left) and muons (right). The points represent the data and the histograms the MC, scaled to correspond to an integrated luminosity of 81.9 fb^{-1}.

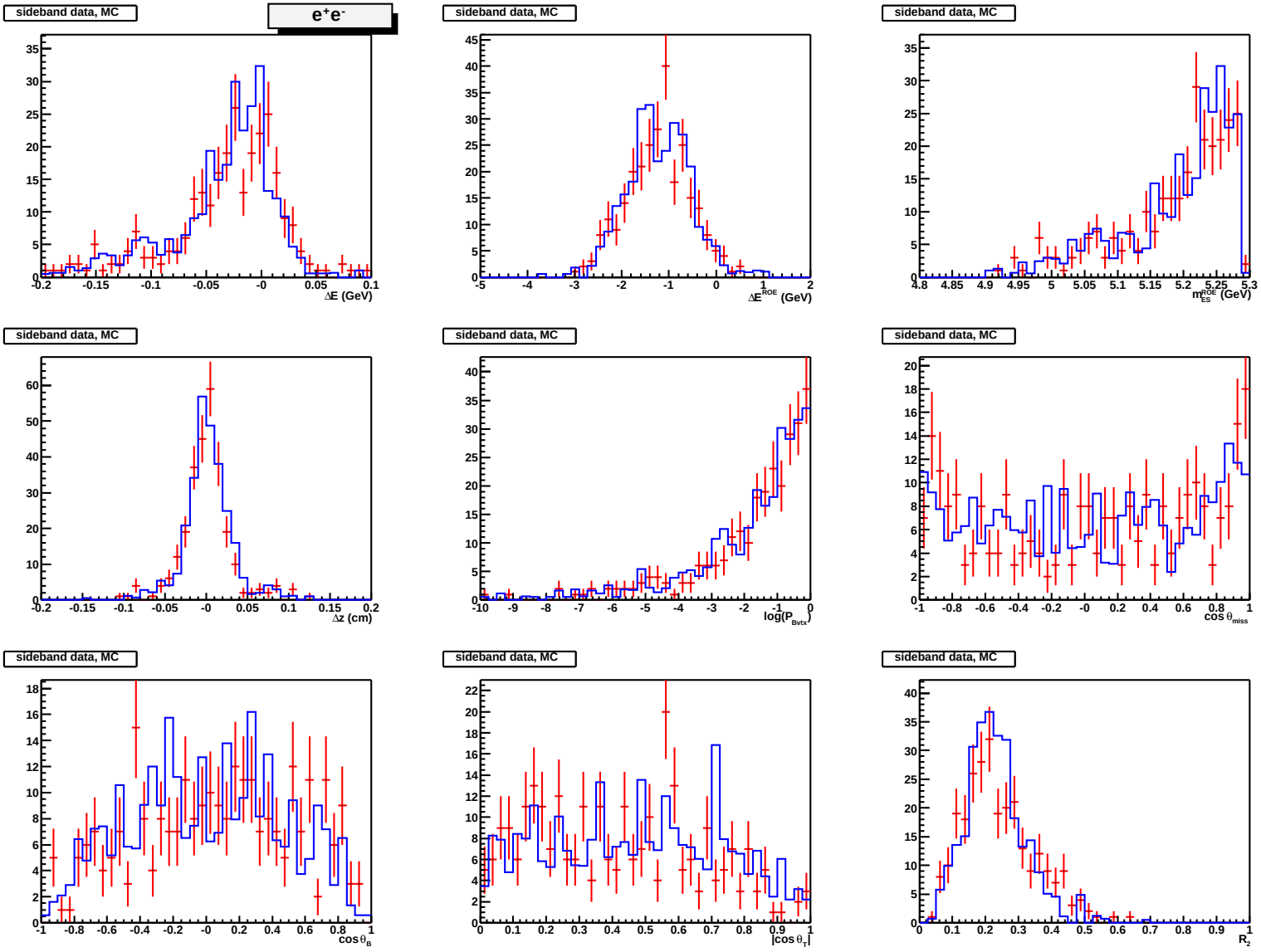

Figure 6.3: Distributions of the background-suppression variables entering the likelihood ratio for $B \to X_s\, e^+e^-$ candidates in the sideband data sample. The points represent the data and the histograms the MC, scaled to correspond to an integrated luminosity of 81.9 fb^{-1}.

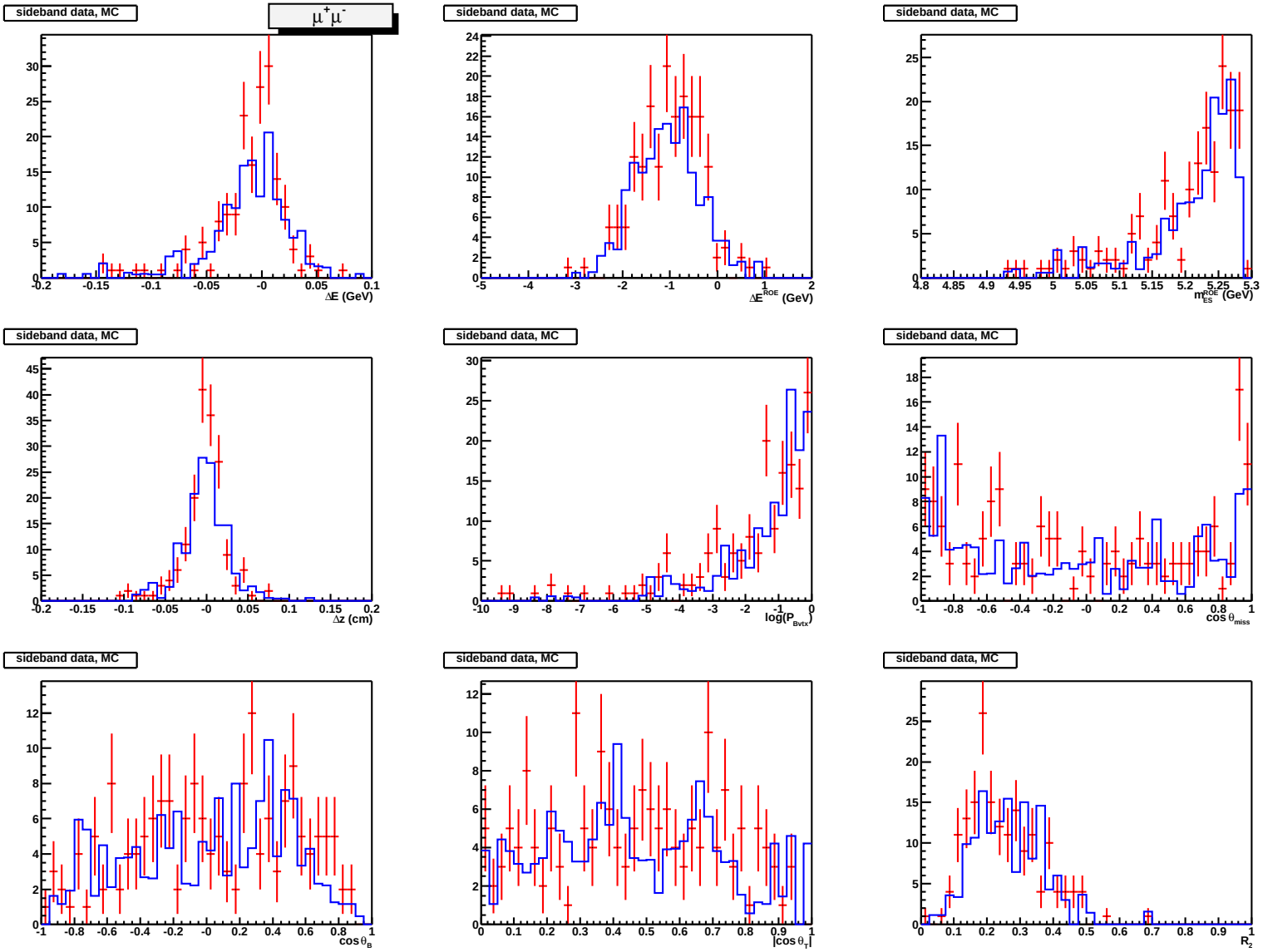

Figure 6.4: Distributions of the background-suppression variables entering the likelihood ratio for $B \to X_s\,\mu^+\mu^-$ candidates in the sideband data sample. The points represent the data and the histograms the MC, scaled to correspond to an integrated luminosity of 81.9 fb^{-1}.

CHAPTER 7

Peaking backgrounds

The three different peaking background types and the methods we use to suppress them were described in Sec. 5.3. In this chapter we estimate the remaining peaking background components after all analysis cuts have been applied.

7.1 Charmonium peaking backgrounds

Peaking backgrounds from $B \to J/\psi\, X$ and $B \to \psi(2S)\, X$ decays are efficiently suppressed by the dilepton mass cuts. However, a very small contribution remains. To estimate this contribution, we used MC generated $B \to J/\psi X$ and $B \to \psi(2S)X$ samples. Table 14 shows the raw numbers of fully reconstructed $B \to J/\psi\, X$ and $B \to \psi(2S)\, X$ decays, in the signal region $m_{\rm ES} > 5.27$ GeV, and the corresponding numbers scaled to an integrated luminosity of 81.9 fb^{-1}, before applying a cut on the likelihood ratio. We found a number of fully reconstructed $B \to J/\psi X$ decays with $J/\psi \to \mu^+\mu^-$ in which one charged pion from the hadronic system was misidentified as a muon and one of the muons from the J/ψ decay was not identified and therefore considered to be a pion. We refer to this situation as a case of $\mu \leftrightarrow \pi$ swap, and include these events in the total charmonium peaking background. These types of decays are likely to evade the charmonium veto cuts

Table 14: Number of charmonium peaking background events from Monte Carlo simulation, before applying a cut on the likelihood ratio. The entries are scaled to an integrated luminosity of 81.9 fb^{-1}. The numbers given in parentheses are the raw event counts.

Mode	Generic $B\overline{B}$ MC	$B \to J/\psi\, X$ MC	$B \to \psi(2S)\, X$ MC	Sum
$J/\psi \to e^- e^-$	0.57 (1)	0.00 (0)	—	0.57 (1)
$J/\psi \to \mu^+ \mu^-$	0.00 (0)	0.00 (0)	—	0.00 (0)
$\psi(2S) \to e^- e^-$	0.40 (1)	—	0.21 (1)	0.61 (2)
$\psi(2S) \to \mu^+ \mu^-$	0.40 (1)	—	0.21 (1)	0.61 (2)
$J/\psi \to \mu^+ \mu^-$	2.28 (4)	1.20 (2)	—	3.48 (6)
$(\mu \leftrightarrow \pi)$ swap				
$\psi(2S) \to \mu^+ \mu^-$	0.00 (0)	—	0.00 (0)	0.00 (0)
$(\mu \leftrightarrow \pi)$ swap				

since the veto is applied on the wrong pair of particles (the reconstructed dilepton consists of one of the J/ψ daughters and a pion from the hadronic system). In a future version of this analysis, one might consider vetoing against such an occurence, but no such requirement is imposed in the present analysis.

After scaling the raw number of peaking background decays to correspond to an integrated luminosity of 81.9 fb^{-1}, we find 0.59 and 2.05 events in the electron and muon channels, respectively. (The numbers in the right-hand column of Table 14 need to be divided by two.) Finally, we scale these numbers by a factor of 0.68 to account for the average efficiency for charmonium decays to pass the likelihood ratio cuts, as measured in the charmonium veto data. The final estimate for the number of charmonium peaking background events is thus 0.40 ± 0.23 in the electron channel and 1.39 ± 0.49 in the muon channel.

For the measurement of partial branching fractions in several bins of dilepton and hadronic mass, we need to estimate the fraction of background in each bin. For the distribution in terms of dilepton mass, we use the dilepton mass distribution of the small number of peaking background events found in the simulation. For

Table 15: Number of charmonium peaking background events from Monte Carlo simulation for electron and muon modes combined in various bins of dilepton and hadronic mass, scaled to an integrated luminosity of 81.9 fb^{-1}.

m_{ll} (GeV)	# bkg decays	m_{Xs} (GeV)	# bkg decays
0.2 – 1.0	0.00 ± 0.00	0.4 – 0.6	0.69 ± 0.21
1.0 – 2.0	0.19 ± 0.19	0.6 – 0.8	0.05 ± 0.02
2.0 – $m_{J/\psi}$	1.18 ± 0.48	0.8 – 1.0	0.37 ± 0.11
$m_{J/\psi}$ – $m_{\psi'}$	0.41 ± 0.21	1.0 – 1.8	0.67 ± 0.20
$m_{\psi'}$ – 5.0	0.00 ± 0.00		
1.0 – 2.45	0.98 ± 0.44		
3.8 – 5.0	0.00 ± 0.00		

the distribution in terms of hadronic mass, we use the distribution found in the charmonium-veto data sample. The estimated numbers of background events are given in Table 15.

7.2 Hadronic peaking backgrounds

Another peaking background due to particle misidentification arises from hadronic B meson decays where two hadrons of opposite charge have been misidentified as leptons. Since the $\pi \to e$ misidentification rate is negligible, this background is only considered in the muon channel. For example, the decay $B^+ \to \bar{D}^0\pi^+$ with $\bar{D}^0 \to K^+\pi^-$ is a background to $B^+ \to K^+\mu^+\mu^-$ when both pions are misidentified as muons.

These backgrounds are estimated using a sample of hadronic B decays in BaBar data. The sample, later on referred to as $X_s hh$, is collected by imposing requirements identical to those for signal candidates, except that no lepton identification

is applied. This results in a large sample of predominantly hadronic B candidates. Each event is given a weight

$w = P(\pi^+ \to \mu^+)P(\pi^- \to \mu^-)$, where $P(\pi^{+(-)} \to \mu^{+(-)})$ is the probability for positive (negative) pions to be misidentified by the positive (negative) muon signal selector.

The probabilities above are obtained from standard PID tables, which contain separate rates for each charge and each run condition block, as well as for varying degrees of binning in momentum and polar angle.

After applying the nominal likelihood ratio cuts, a weighted distribution in m_{ES} is obtained, and a one-dimensional binned χ^2 fit for signal plus combinatorial background gives the number of hadronic peaking background events.

The fits for each mode are shown in Figures 64, 65, and 66, and the signal yield for each fit is shown in Table 16. The hadronic peaking background for the muon mode is estimated to be 2.39 ± 0.75. The uncertainty is conservatively estimated by increasing all misidentification rates provided in the PID tables by $\pm 1\sigma$, i.e. assuming the rates in the different bins of angle and momentum are fully correlated.

For the measurement of partial branching fractions in several bins of dilepton and hadronic mass, we need to estimate the fraction of background in each bin. The above procedure is repeated in each of the bins and the results for the $X_s h^+ h^-$ topology are summarized in Table 17.

Mode	$X_s h^+ h^-$
$X_s \mu^+ \mu^-$	2.39 ± 0.75
$K^\pm \mu^+ \mu^-$	1.04 ± 0.38
$K^\pm \pi^0 \mu^+ \mu^-$	0.24 ± 0.06
$K^\pm \pi^\pm \mu^+ \mu^-$	0.43 ± 0.13
$K^\pm \pi^\pm \pi^0 \mu^+ \mu^-$	0.07 ± 0.01
$K^\pm \pi^\pm \pi^\pm \mu^+ \mu^-$	0.15 ± 0.04
$K_s^0 \mu^+ \mu^-$	0.19 ± 0.06
$K_s^0 \pi^0 \mu^+ \mu^-$	0.02 ± 0.01
$K_s^0 \pi^\pm \mu^+ \mu^-$	0.13 ± 0.04
$K_s^0 \pi^\pm \pi^0 \mu^+ \mu^-$	0.04 ± 0.01
$K_s^0 \pi^\pm \pi^\pm \mu^+ \mu^-$	0.02 ± 0.01

Table 16: Hadronic peaking backgrounds for $X_s \mu^+ \mu^-$ and its individual event topologies with the $X_s h^+ h^-$ sample.

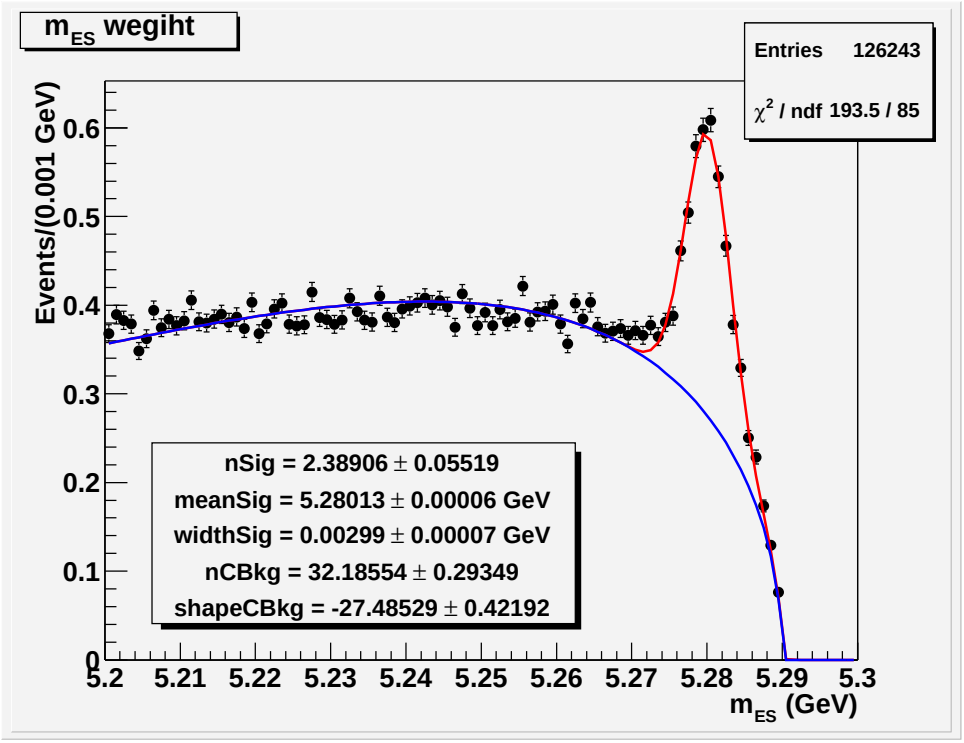

Figure 64: Weighted m_{ES} distribution of $X_s h^+ h^-$ events. The red line is the total fit to a Gaussian signal plus an Argus background; the blue line is the Argus background portion of the fit.

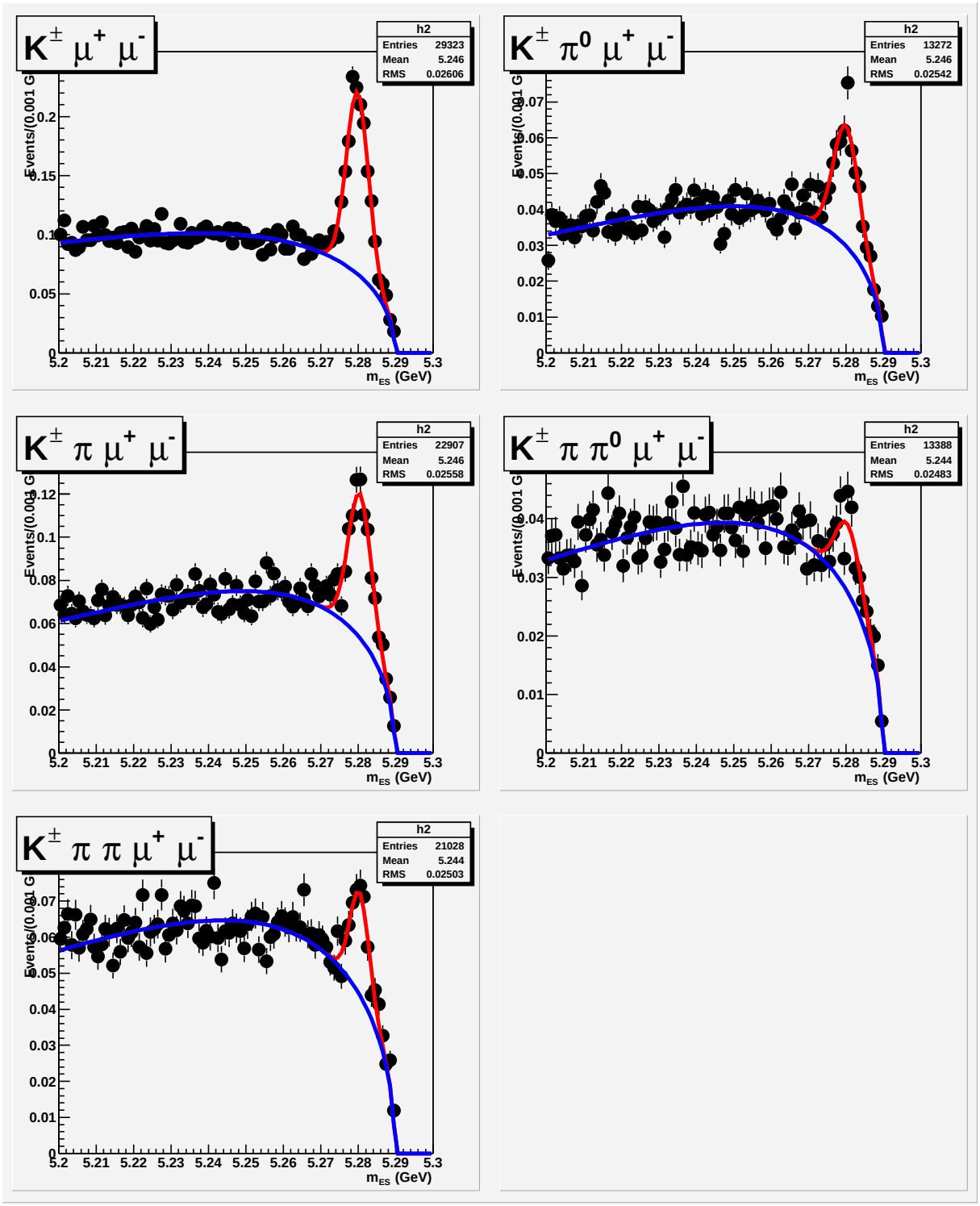

Figure 65: Weighted m_{ES} distributions for $K^{\pm}n\pi h^{+}h^{-}$ final state topologies. The red line is the total fit to a Gaussian signal plus an Argus background; the blue line is the Argus background portion of the fit.

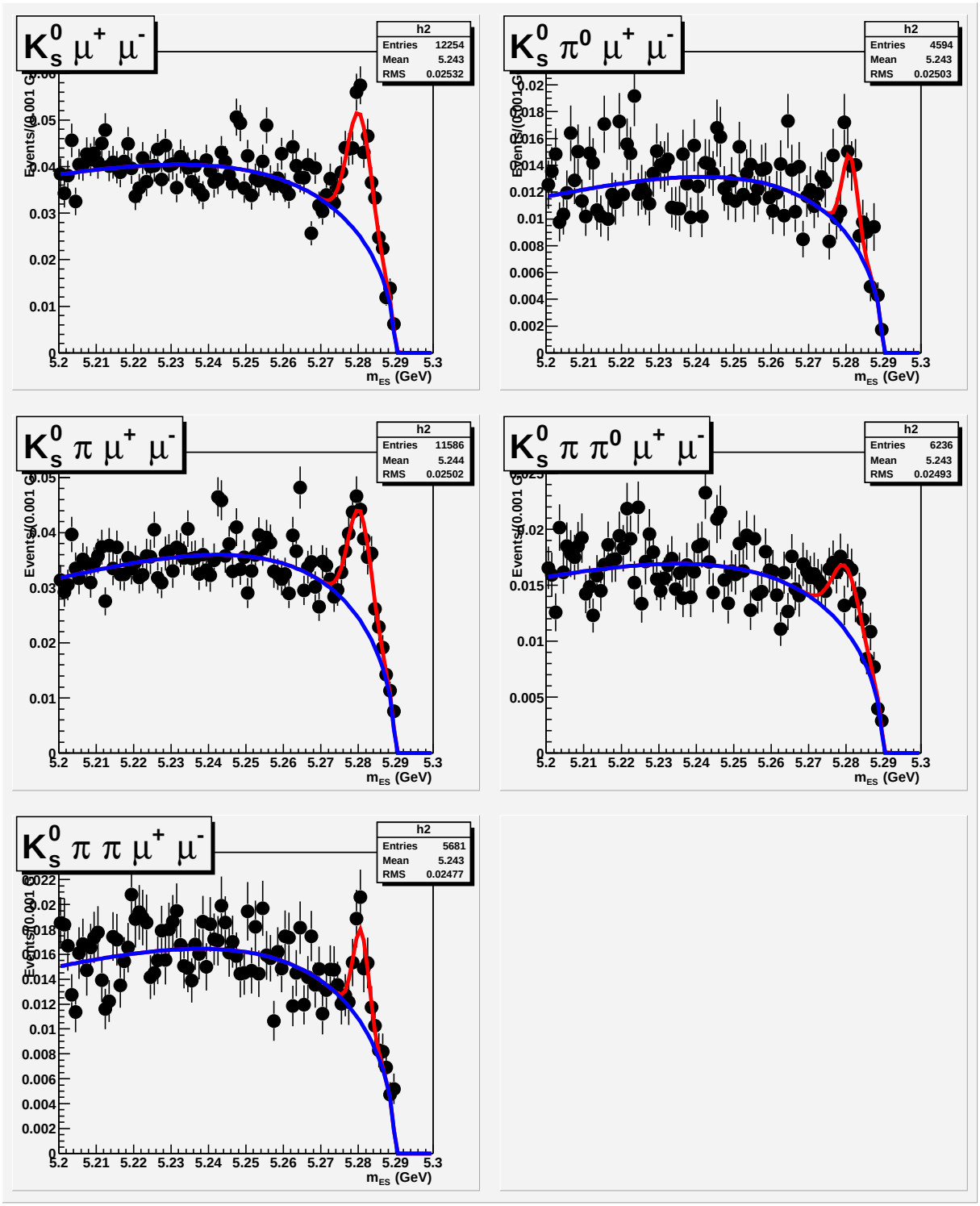

Figure 66: Weighted m_{ES} distribution for $K^0_s n\pi h^+ h^-$ final state topologies. The red line is the total fit to a Gaussian signal plus an Argus background; the blue line is the Argus background portion of the fit.

Table 17: Number of $X_s h^+ h^-$ hadronic peaking background events for muon modes in various bins of dilepton and hadronic mass.

m_{ll} (GeV)	# bkg decays	m_{Xs} (GeV)	# bkg decays
0.2 – 1.0	0.06 ± 0.03	0.4 – 0.6	1.24 ± 0.45
1.0 – 2.0	0.11 ± 0.03	0.6 – 0.8	0.22 ± 0.06
2.0 – $m_{J/\psi}$	0.29 ± 0.08	0.8 – 1.0	0.37 ± 0.11
$m_{J/\psi}$ – $m_{\psi'}$	0.58 ± 0.16	1.0 – 1.8	0.55 ± 0.15
$m_{\psi'}$ – 5.0	1.32 ± 0.47		
1.0 – 2.45	0.22 ± 0.07		
3.8 – 5.0	1.32 ± 0.47		

CHAPTER 8

Maximum likelihood fit

We extract the signal yield with an extended unbinned maximum likelihood fit to the m_{ES} distribution in the region $m_{ES} > 5.2$ GeV. The likelihood function consists of 4 components: signal, charmonium peaking background, hadronic peaking background and continuum,

$$\mathcal{L} = \frac{e^{-(N_{sig}+N_{c\bar{c}s}+N_{had}+N_{comb})}}{N!} \prod_{k=1}^{N} [(N_{sig}+N_{c\bar{c}s})\mathcal{P}^{sig}(\mathrm{m_{ES}^k}) + \mathrm{N_{had}}\mathcal{P}^{\mathrm{had}}(\mathrm{m_{ES}^k}) + \mathrm{N_{comb}}\mathcal{P}^{\mathrm{comb}}(\mathrm{m_{ES}^k})]. \tag{8.1}$$

The product runs over the total number – N – of reconstructed candidate events. We perform separate fits for electron and muon modes, as well as a third fit where both samples are combined. The parameters N_{sig}, $N_{c\bar{c}s}$, N_{had} and N_{comb} represent the signal yield, the charmonium peaking background, the hadronic peaking background and the combinatorial background, respectively, with corresponding PDFs given by $\mathcal{P}^{sig}$, $\mathcal{P}^{had}$ and $\mathcal{P}^{comb}$. The fitted parameters are N_{sig}, N_{comb} and the shape parameter of the PDF describing the combinatorial background $\mathcal{P}^{comb}$. The remaining parameters are fixed during the fit with values determined as follows:

- Signal and charmonium peaking background shapes:

  Both the signal and the charmonium peaking background shapes are described by one PDF, $\mathcal{P}^{sig}$. For the functional form we use a Gaussian for the muon as well as the electron modes, since the Bremsstrahlung recovery leads to a negligible radiative tail in the m_{ES} distribution of the latter. The shape parameters of the Gaussian are determined from Gaussian + Argus fits to the charmonium veto data sample as shown in Fig. 67. The fits yield signal peak positions at $m_{sig} = 5.280036 \pm 0.000051$ GeV, 5.280045 ± 0.000070 GeV and 5.280040 ± 0.000041 GeV in the electron, muon and electron+muon modes, respectively. The corresponding m_{ES} resolutions are $\sigma_{sig} = 2.804 \pm 0.046$ MeV for electrons, $\sigma_{sig} = 2.613 \pm 0.061$ MeV for muons and $\sigma_{sig} = 2.745 \pm 0.037$ MeV for electrons and muons combined. The m_{ES} distributions for truth-matched signal modes are illustrated in Fig. 68, and their Gaussian fit results are in good agreement with the shape parameters extracted from the fits to the charmonium veto data sample.

- Charmonium peaking background amount

  The expected number of fully reconstructed $B \to J/\psi\, X$ and $B \to \psi(2S)\, X$ decays – $N_{c\bar{c}s}$ – passing through the charmonium veto is estimated to be 0.40 and 1.39 events in the electron and muon channels, respectively, as determined from the MC, see Sec. 7.1.

- Hadronic peaking background

  The amount of this background, $N_{had} = 2.39$, is derived directly from data by performing the analysis without the lepton identification requirements, as described in Sec. 7.2. Parameters for the PDF $\mathcal{P}^{had}$ are also taken from that fit: the Gaussian mean is $m_{had} = 5.28013 \pm 0.00006$ GeV and the Gaussian

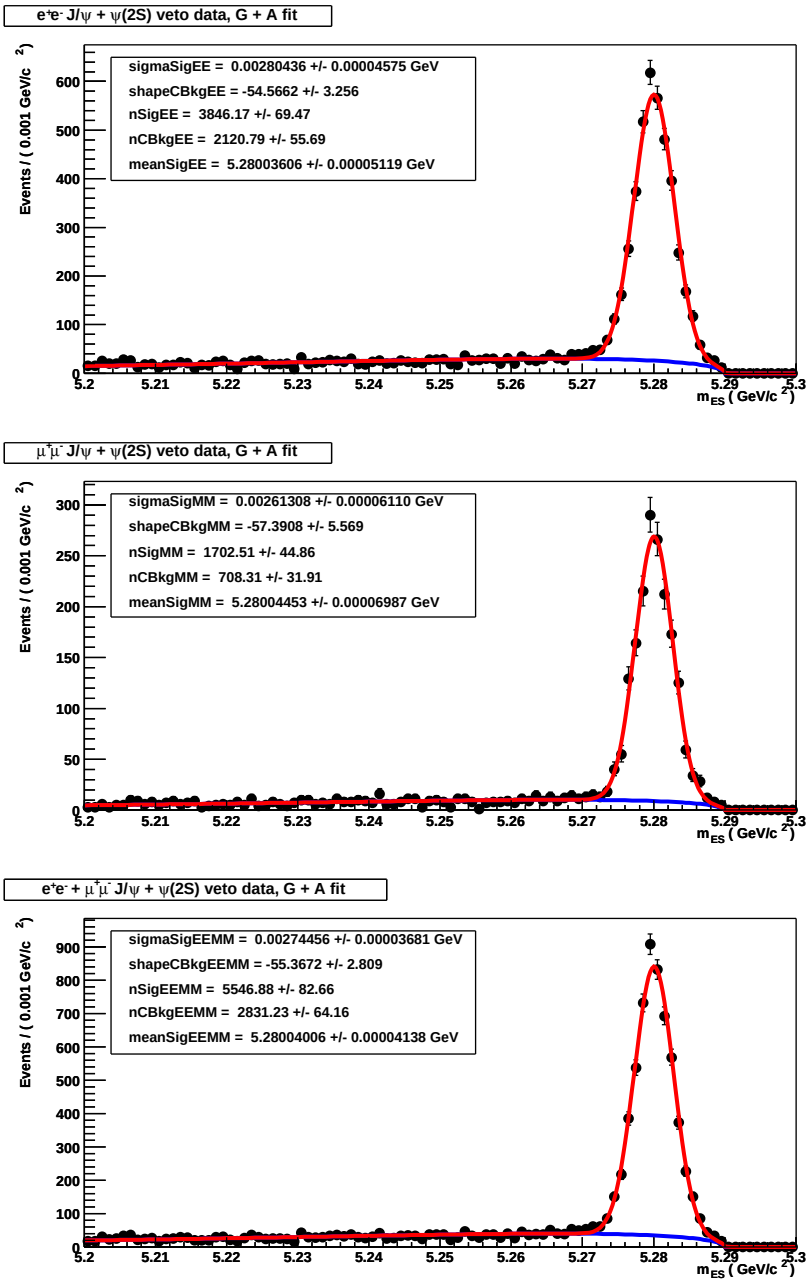

Figure 67: Distributions of m_{ES} for $B \to X_s\, \ell^+\ell^-$ candidates in the charmonium veto data sample for electrons (top), muons (middle), and both combined (bottom) after all cuts (but the inverted charmonium veto) are applied.

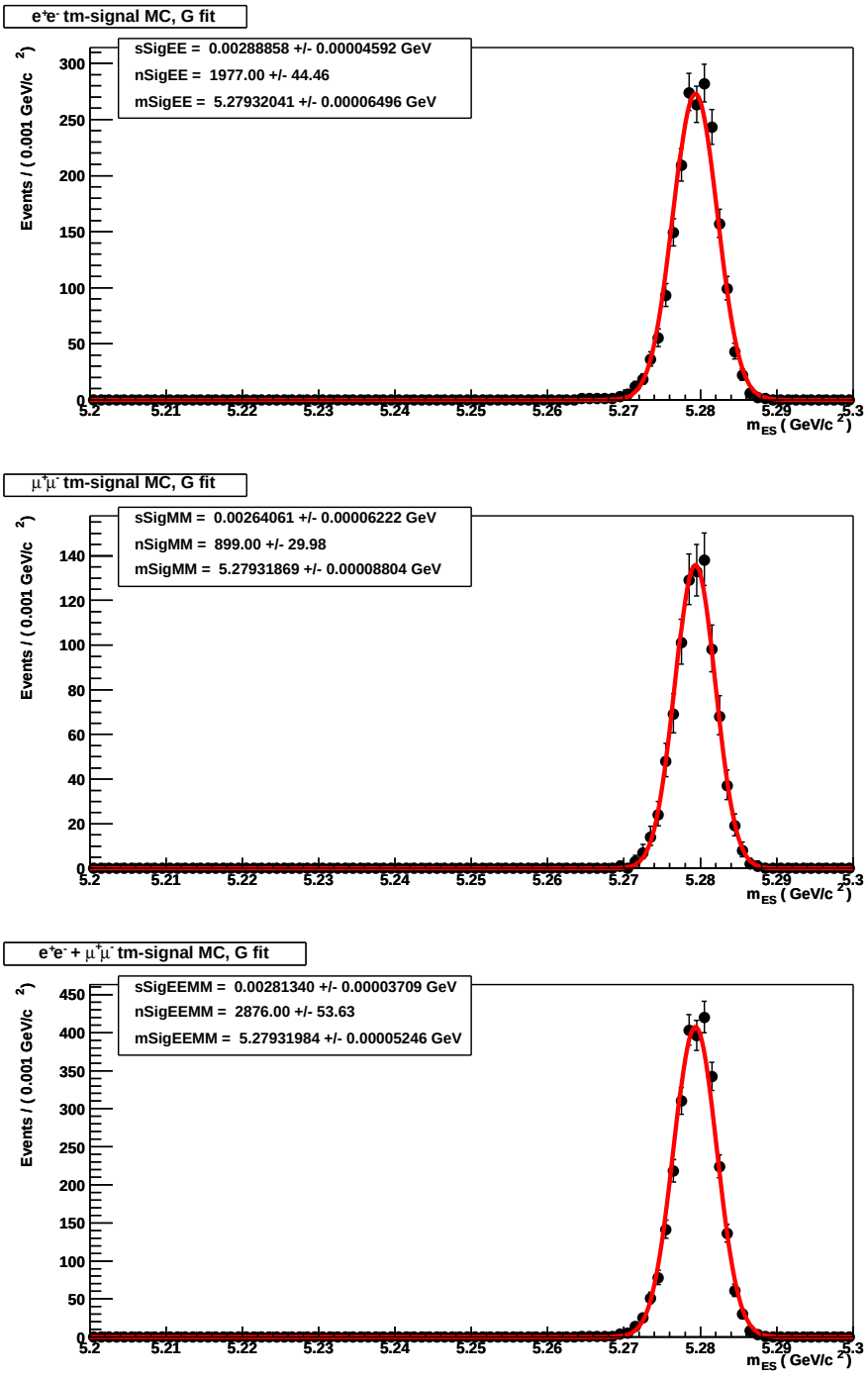

Figure 68: Distributions of m_{ES} for truth-matched $B \to X_s\, \ell^+\ell^-$ candidates for electrons (top) and muons (bottom).

width is $\sigma_{had} = 2.99 \pm 0.07$ MeV, see Fig. 64.

- Combinatorial background cutoff

  The combinatorial background PDF, $\mathcal{P}_k^{comb}$, is given by an $Argus$ shape which describes the combinatorial contribution from continuum events and $B\bar{B}$ events. The $Argus$ cutoff is determined by the beam energy in the $\Upsilon(4S)$ rest frame, $E_{beam} = 5.290$ GeV.

The signal yield obtained with the outlined procedure contains a small contribution from $B \to X_s\, \ell^+ \ell^-$ cross-feed events which is estimated from Monte Carlo, as will be shown in Section 9.1. The overall signal efficiency needed to compute branching fractions includes this small contribution.

CHAPTER 9

Results

In this chapter we present the results from the measurements of the total branching fraction, the branching fractions in dilepton and hadronic mass bins and a measurement of the CP asymmetry.

9.1 Total branching fractions

The final $m_{\rm ES}$ histograms fitted as described in Chapter 8 are shown in Fig. 69. We measure a branching fraction of $\mathcal{B}(B \to X_s\,\ell^+\ell^-) = (5.64 \pm 1.46) \times 10^{-6}$ with statistical significance of $4.3\,\sigma$, see Table 18. The significance is computed according to Signif $= \sqrt{2(\log \mathcal{L}_{max} - \log \mathcal{L}^0_{max})}$, where $\mathcal{L}_{max}$ is the maximum likelihood for the nominal fit (with the signal yield as in Table 18) and $\mathcal{L}^0_{max}$ is a maximum likelihood obtained with a different fit for which the signal yield is fixed at $N_{sig} = 0$.

The branching fraction $\mathcal{B}$ for the signal is calculated with

$$\mathcal{B} = \frac{N_{sig}}{2N_{B\overline{B}}\,\epsilon}, \tag{9.1}$$

where $N_{B\overline{B}} = (88.9 \pm 1.0) \times 10^6$ is the number of $B\overline{B}$ pairs produced in 81.9 fb^{-1} and ϵ is a signal efficiency calculated from the total truth-matched signal

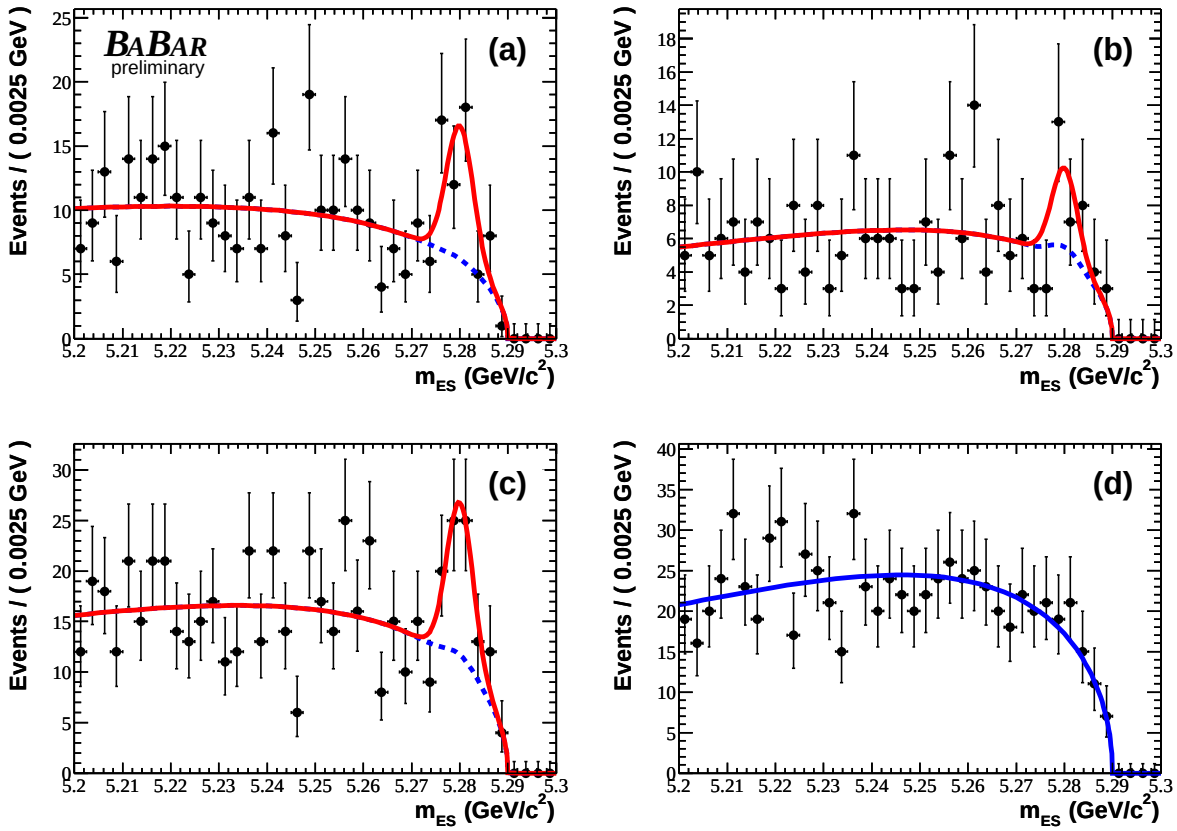

Figure 69: Distributions of m_{ES} for data in the (a) e^+e^-, (b) $\mu^+\mu^-$, (c) e^+e^- $+\mu^+\mu^-$, and (d) $e^\pm\mu^\mp$ channels. The red line is the result of the fit, with the total background component shown in the dashed (or solid for $e^\pm\mu^\mp$) blue line.

Table 18: Results of the fit to data: signal yield, peaking background (fixed in the fit), combinatorial background, signal significance, signal efficiency and branching fraction.

Sample	N_{sig}	$N_{c\bar{c}s}$	N_{peak}	N_{comb}	Signif.	ϵ	$\mathcal{B}(\times 10^{-6})$
$X_s\, e^+e^-$	29.22 ± 8.28	0.40	0.0	319.4 ± 18.9	4.0	2.74%	6.00 ± 1.70
$X_s\, \mu^+\mu^-$	11.25 ± 6.19	1.39	2.39	207.0 ± 15.2	2.0	1.26%	5.04 ± 2.77
$X_s\, \ell^+\ell^-$	40.06 ± 10.35	1.79	2.39	526.8 ± 24.3	4.3	2.00%	5.64 ± 1.46

reconstruction efficiency, multiplied by a factor to account for the small cross-feed component picked up by the signal Gaussian in the fit process.

The truth-matched signal reconstruction efficiency is shown in the last column of Tables 28 and 29, and is calculated as the ratio between the number of fully reconstructed signal decays in the ten considered modes and the total number of generated signal events in all topologies. The additional cross-feed factor is determined by comparing the fit results on the data obtained with the likelihood function from Chapter 8, see Fig. 69, to those obtained with a modified likelihood function that includes an extra PDF for the cross-feed contribution, the green line in Fig. 70. The shape of the cross-feed PDF is determined from MC, see Fig. 71 and is fixed during the maximum likelihood fit, the relative cross-feed yield, with respect to the signal yield, is derived from simulation and fixed to 0.47 and 0.37 for electron and muon modes, respectively. The additional signal yield obtained in the nominal fit due to cross-feed is determined to be 1.67, 0.56, and 2.10 events in the electron, muon, and electron+muon channels, respectively. This corresponds to a relative increase of 6.1%, 5.3%, and 5.5% in the signal efficiency with respect to the efficiency obtained with truth-matched decays only. The signal efficiency calculated for truth-matched signal decays is therefore increased by a factor of 1.055, as determined from the larger statistics electron+muon channel. The overall signal efficiencies are thus estimated to be 2.74%, 1.26%, and 2.00% in the electron, muon, and electron+muon channels, respectively. The efficiency for the electron+muon channel is computed as a simple average of the electron and muon channel efficiencies, since the branching fractions are expected to be equal after the requirement $m_{ll} > 0.2$ GeV. Branching fractions are given in Table 18. For the combined electron and muon channels, the efficiency and branching fraction are to be understood as averages over the electron and muon channels.

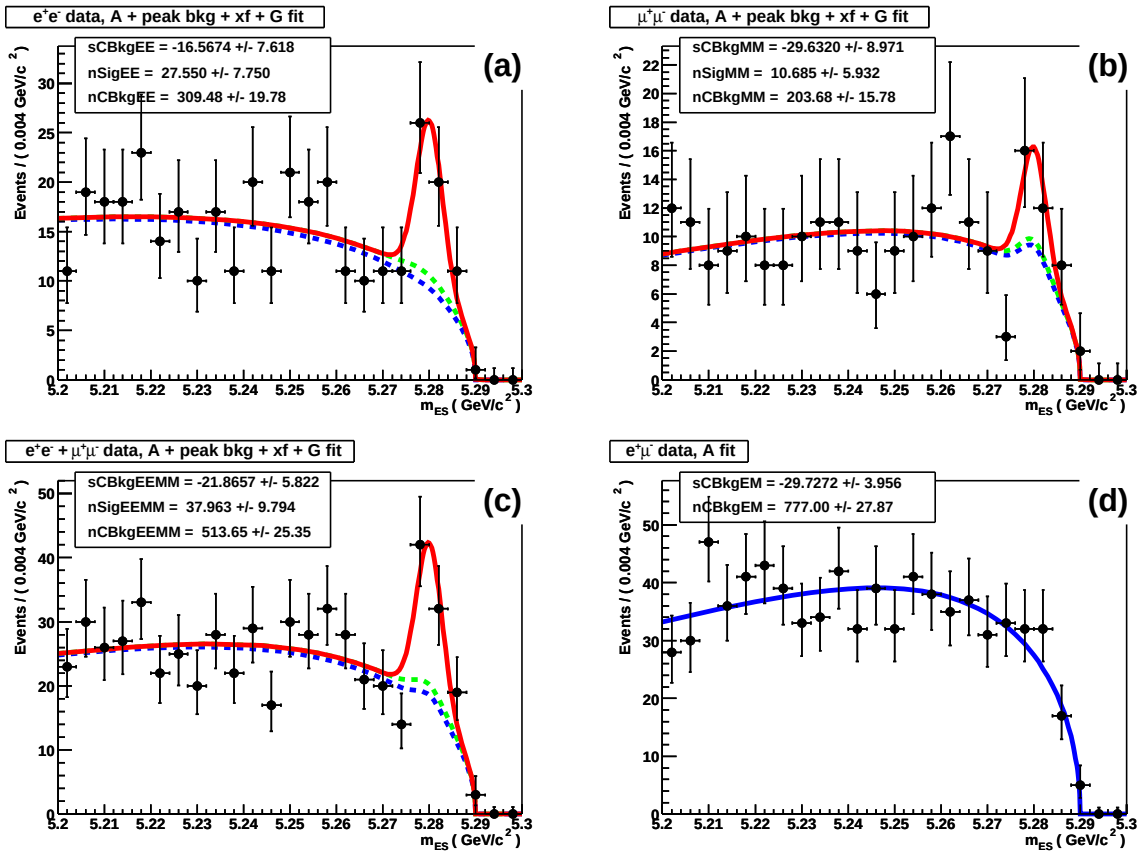

Figure 70: Distributions of m_{ES} for data in the (a) e^+e^-, (b) $\mu^+\mu^-$, (c) e^+e^- $+\mu^+\mu^-$, and (d) $e^\pm\mu^\mp$ channels. The red line is the result of the fit, with the total background component shown in the dashed blue line and the sum of total background + signal cross-feed contribution in the dashed green line; a solid blue line is used for the total background in the $e^\pm\mu^\mp$ channel.

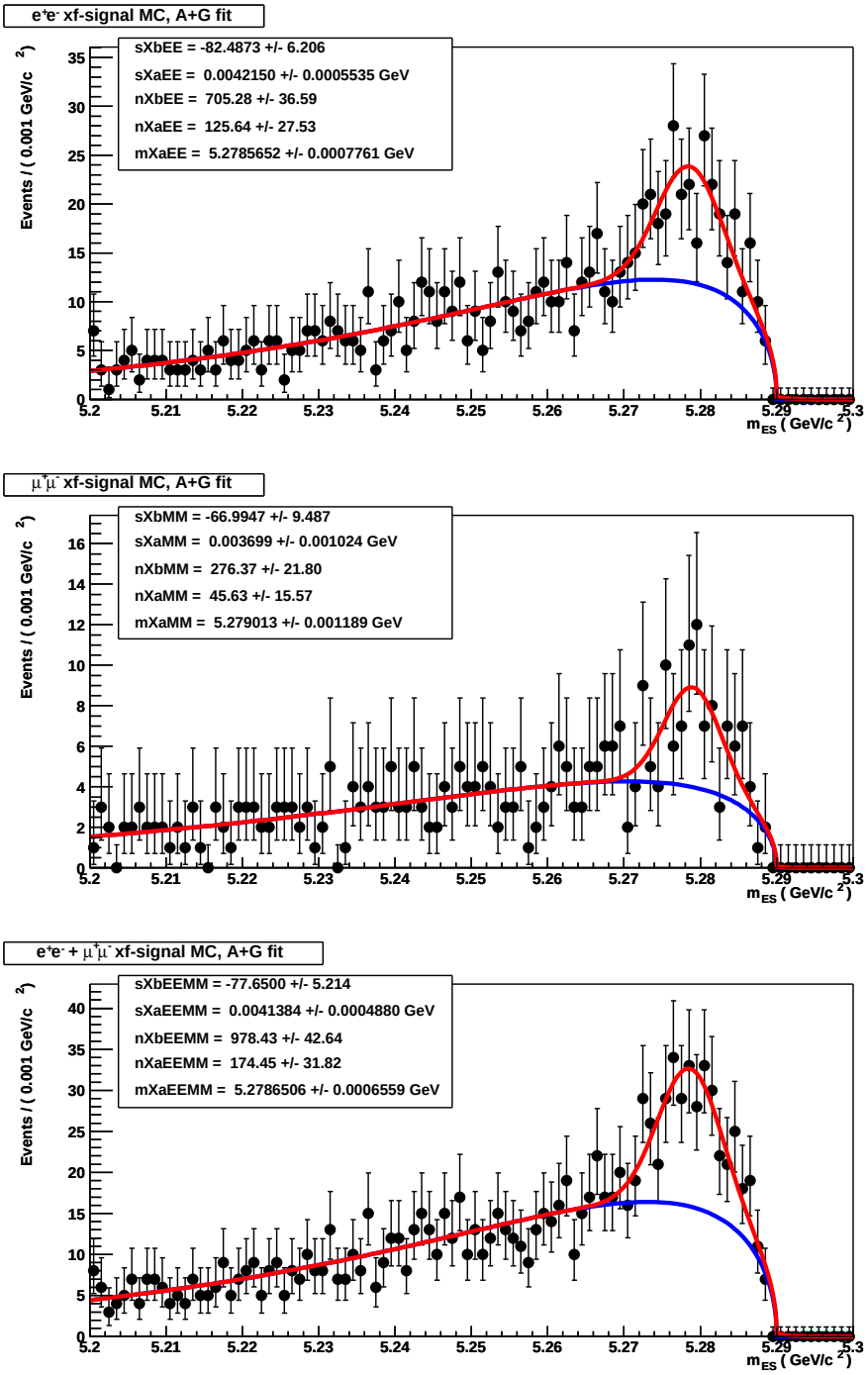

Figure 71: Distributions of m_{ES} for $B \to X_s \ell^+ \ell^-$ candidates in the signal MC that fail truth-matching, i.e. cross-feed events for electrons (top) and muons (bottom). A $Gaussian + Argus$ is used to parametrize the cross-feed PDF for both dilepton modes.

Table 19: Signal yields for electron and muon modes combined in various bins of dilepton and hadronic mass.

m_{ll} (GeV/c^2)	N_{sig}	m_{Xs} (GeV/c^2)	N_{sig}
0.2 – 1.0	0.4 ± 2.0	0.4 – 0.6	16.3 ± 5.3
1.0 – 2.0	12.2 ± 4.8	0.6 – 0.8	3.5 ± 2.6
2.0 – $m_{J/\psi}$	12.6 ± 6.0	0.8 – 1.0	6.9 ± 4.3
$m_{J/\psi}$ – $m_{\psi'}$	5.8 ± 5.0	1.0 – 1.8	12.3 ± 7.3
$m_{\psi'}$ – 5.0	8.6 ± 4.3		
1.0 – 2.45	15.9 ± 6.2		
3.8 – 5.0	8.6 ± 4.3		

9.2 Partial branching fractions in dilepton and hadronic mass bins

Signal yields are also extracted for the electron and muon channels combined in a series of dilepton and hadronic mass ranges. The individual m_{ES} distributions and respective fits are shown in Figs. 72 and 73. The nominal fit procedure is performed with the nominal signal Gaussian shape parameters and numbers of peaking background decays as described in Secs. 7.1 and 7.2. The signal yields are given in Table 19.

To compute the partial branching fraction in each of the bins, we need to estimate the additional efficiency due to cross-feed $B \to X_s\, \ell^+\ell^-$ events. Overall, the number of signal events attributed to cross-feed events is estimated to be 5.5% of the number of fully reconstructed events (i.e. truth-matched decays). To distribute these events among m_{Xs} and m_{ll} bins, we use the simulated m_{Xs} and m_{ll} distributions for cross-feed $B \to X_s\, \ell^+\ell^-$ events. This approach assumes that the m_{ES} shape of the cross-feed events does not change as a function of m_{Xs} and m_{ll}. Table 20 summarizes the efficiency correction factors used to account for the cross-feed

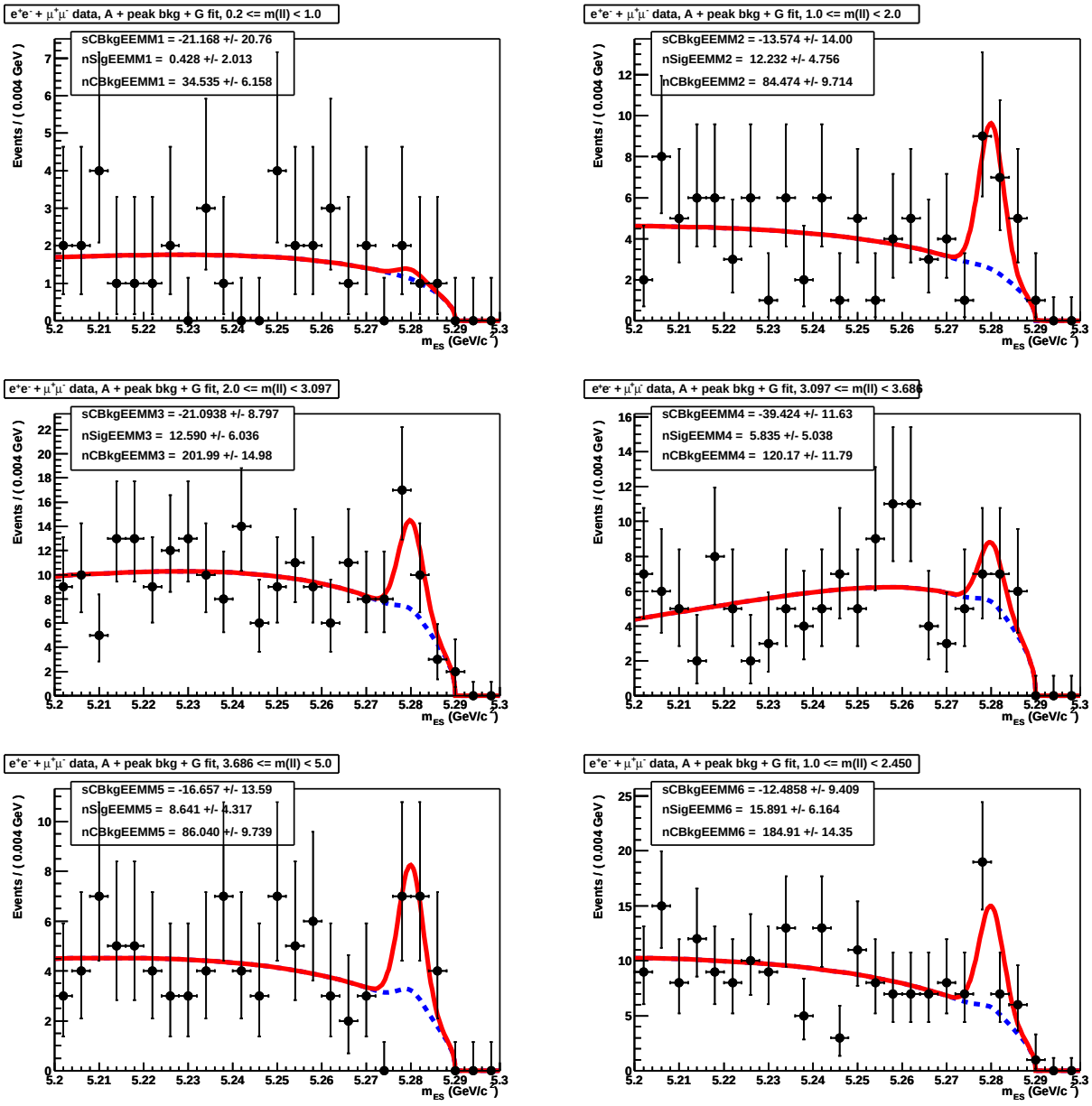

Figure 72: Distributions of m_{ES} for $B \to X_s \ell^+\ell^-$ candidates (electron+muon channels combined) in different ranges of dilepton mass (see range in the box above each plot). The red line is the result of the fit, with the total background component shown in the dashed blue line.

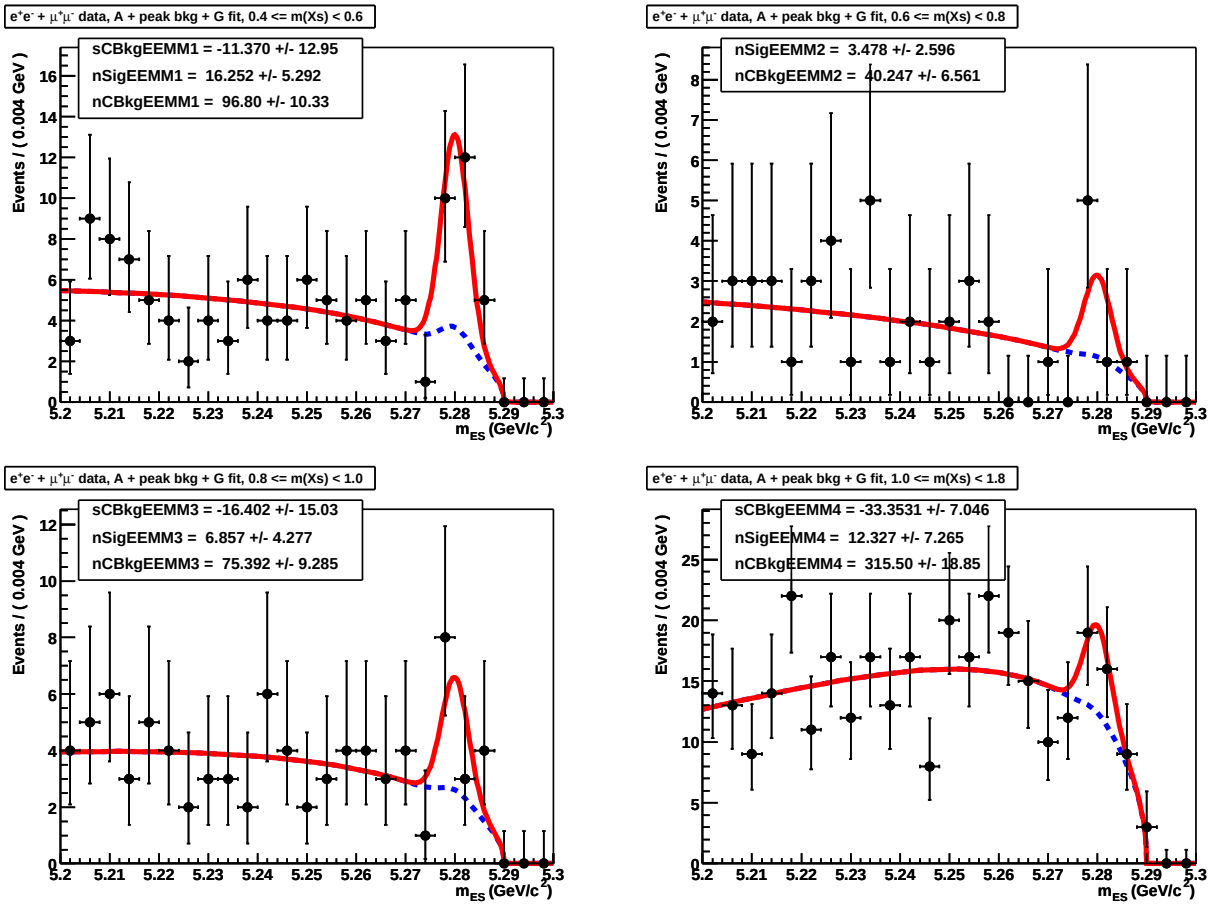

Figure 73: Distributions of m_{ES} for $B \to X_s \ell^+\ell^-$ candidates (electron+muon channels combined) in different ranges of hadronic mass (see range in the box above each plot). The red line is the result of the fit, with the total background component shown in the dashed blue line.

Table 20: Efficiency (in percent) for truth-matched $B \to X_s\, \ell^+\ell^-$ decays, enhancement factor to account for the cross-feed contribution to the signal yield, and final efficiency (in percent).

Sample	Raw efficiency	Factor	Efficiency
e^+e^-	2.60	1.06	2.74
$\mu^+\mu^-$	1.19	1.06	1.26
$\ell^+\ell^-$	1.89	1.06	2.00
$m(X_s)(\,\mathrm{GeV}/c^2)$			
$0.4 - 0.6$	8.49	1.02	8.66
$0.6 - 0.8$	2.85	1.06	3.03
$0.8 - 1.0$	2.76	1.09	2.99
$1.0 - 1.8$	0.91	1.09	0.99
$m(\ell^+\ell^-)(\,\mathrm{GeV}/c^2)$			
$0.2 - 1.0$	1.07	1.45	1.55
$1.0 - 2.0$	2.05	1.03	2.10
$2.0 - m_{J/\psi}$	1.94	1.04	2.02
$m_{J/\psi} - m_{\psi'}$	1.57	1.07	1.68
$m_{\psi'} - 5.0$	3.49	1.10	3.82
$1.0 - 2.45$	2.33	1.04	2.42
$3.795 - 5.0$	4.45	1.10	4.88

contribution to the signal yield in the different m_{Xs} and m_{ll} bins.

9.3 CP asymmetry

Using the fit described above, we can also extract the number of signal events in separate samples of B and $\overline{B}$ decays, and thus measure the direct CP asymmetry in the $B \to X_s\, \ell^+\ell^-$ decay process. To do so, we use only self-tagging decay modes, i.e., the 3 modes with hadronic systems consisting of $X_s = K_s$, $K_s\pi^0$, and $K_s\pi^+\pi^-$ are removed. The m_{ES} distributions for $\overline{B}$ and B decays are shown in Fig. 74. As for the nominal fit, we extract the signal Gaussian shape parameters from the corresponding 7 modes in the charmonium veto sample. Furthermore, we

assume identical shapes for the B and $\overline{B}$ samples, as well as peaking backgrounds, i.e., we divide the amount of peaking background equally between the B and $\overline{B}$ subsamples. Using our control samples, we found no evidence for these asumptions to be incorrect. Fit results are $N_{sig}^{\overline{B}} = 14.7 \pm 6.5$ and $N_{sig}^{B} = 22.9 \pm 7.4$, which corresponds to a charge asymmetry of $A_{CP} \equiv \frac{N_{sig}^{\overline{B}} - N_{sig}^{B}}{N_{sig}^{\overline{B}} + N_{sig}^{B}} = -0.22 \pm 0.26$.

This result is consistent with the expectation of no asymmetry in the Standard Model. Ali and Hiller [25] predict $A_{CP} = (0.19^{+0.17}_{-0.19})\%$.

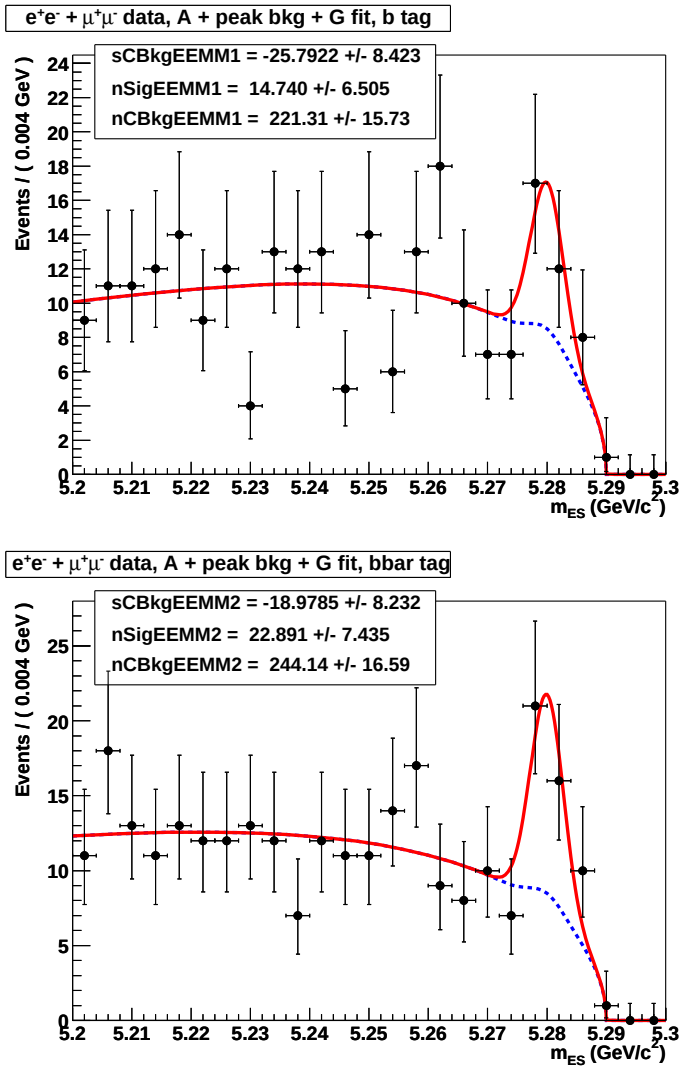

Figure 74: Distributions of m_{ES} for combined electron and muon data in self-tagged $\overline{B}$ (top) and B (bottom) decays. The red line is the result of the fit, with the total background component shown in the dashed blue line.

CHAPTER 10

Systematic uncertainties

The systematic uncertainties are of two different types, those that affect the extraction of the number of signal decays and those that affect the calculation of the branching fraction.

10.1 Signal yield systematics

Uncertainties affecting the extraction of the signal yield are as follows:

- Signal Gaussian shape parameters

  Signal shape parameters are taken from a fit to the charmonium veto sample, see Fig. 67. The mean and width of the signal Gaussian are varied by $\pm 1\sigma$, where σ is the statistical uncertainty from the fit.

- Signal shape

  Fits can be performed with different signal function shapes. Both Crystal Ball and Novosibirsk functions provide adequate description of the signal and allow for an asymmetric m_{ES} distribution. The difference (from the nominal fit) in signal yields is taken as a systematic uncertainty.

- Hadronic peaking background

  The amount of hadronic peaking background ($B \to X_s\, h^+ h^-$) is appreciable only in the muon channels and is estimated directly from data (see Table 16). The associated uncertainty is dominated by the uncertainty in the $\pi \to \mu$ misidentification probability. Gaussian shape parameters (mean and width) are also varied according to the fit result to the sample obtained without lepton identification.

- Charmonium peaking background

  The other source of peaking background originates from $B \to J/\psi X$ and $B \to \psi(2S)\, X$ decays that evade the charmonium veto. The amount of background is estimated from Monte Carlo simulation. This amount is varied according to the error expected for the number of selected events in the simulation.

The total uncertainties in the signal yield are ± 1.26 and ± 0.90 events in the electron and muon channels, respectively. See Table 21 for a summary of the branching fraction systematic errors due to the signal yield uncertainties.

10.2 Branching fraction systematics

The uncertainties affecting the estimate of the signal selection efficiency (or the number of B mesons in the sample) are as follows:

- Tracking efficiency

  The systematic uncertainties of the charged track reconstruction efficiencies are determined by the *Tracking Efficiency Task Force* at *BABAR*. For the charged track reconstruction efficiency we apply a flat systematic uncertainty of $\pm 1.3\%$ per track for lepton and kaon tracks, and $\pm 0.8\%$ per track for the

pion tracks. These values were chosen based on the PID selectors we use, and the average track multiplicity in our MC samples.

- K_S^0 efficiency

  For K_S^0 the uncertainty is 1.5% per K_S^0 estimated using the official recipe [26].

- π^0 efficiency

  For neutral pions, the uncertainty is $\pm 5.0\%$ per π^0. This increases to 5.6% per π^0 after taking care of an additional uncertainty added in quadrature corresponding to whether one accounts for uncorrelated errors or not.

- Charged particle ID efficiency

  The efficiency systematics for charged particle identification (e, μ and K) can be estimated in two ways. First, we take uncertainties from the PID tables and properly average them using signal MC. This method yields uncertainties of 0.56% per electron, 3.85% per muon and 1.06% per kaon. Another approach has been taken for the study of exclusive $B \to K\ell^+\ell^-$ and $B \to K^*\ell^+\ell^-$ decays [27], where the efficiencies and associated uncertainties are computed from the charmonium veto sample. This second approach yields uncertainties of 0.65% per electron, 1.6% per muon and 1.0% per kaon. The latter approach is used for this analysis. In the case of the electron selection, we apply an additional uncertainty of 0.2% per electron (added linearly) to account for the difference in PID efficiency observed between electron candidates in the ChargedTracks selector (as used in this analysis) and the GoodTracksLoose selector (as used by the exclusive $B \to K^{(*)}\ell^+\ell^-$ analysis) lists. The uncertainty is thus 0.85% per electron.

- Likelihood ratio efficiency

The efficiency of the final step of background rejection is checked with the charmonium veto sample. After background subtraction, the efficiency of the likelihood ratio cuts is found to be 68.0% (68.1%) in the electron (muon) data and 65.5% (65.9%) in the electron (muon) MC. Given the available statistics the level of disagreement between data and MC is about 2 σ. We take the discrepancy of 3.9% (3.3%) in the electron (muon) channel to the systematic uncertainty in the likelihood ratio cut efficiency.

- Cross-feed efficiency

  The fraction of cross-feed that ends up in the signal Gaussian is determined by comparing the fit results on the data obtained with either the nominal likelihood function or with a modified version that incorporates an extra PDF that represents the shape of the cross-feed contribution, as determined from MC. The additional signal yield obtained in the nominal fit is taken to correspond to an increase of 5.5% in the signal efficiency with respect to the efficiency obtained with truth-matched decays only. The number of cross-feed events included in the signal Gaussian are consistent with those obtained in toy MC studies. These studies indicate that the distribution in the number of cross-feed events has an R.M.S. close to the average number of cross-feed events. Therefore, we assign an uncertainty of $\pm 100\%$ in the additional efficiency due to the signal cross-feed contribution.

- Fermi motion model

  The parameters of the Fermi motion model are set to the values that best describe inclusive $b \to s\gamma$ decays measurements at CLEO [28]: $p_F = 410\,\mathrm{MeV}/c$ and $m_q = 0$. Uncertainties in the parameter p_F can be extracted using measurements of HQET parameters from measurements of hadronic moments in

semileptonic decays: $\lambda_1 = -0.24 \pm 0.11\,\mathrm{GeV}^2$ [29] and from inclusive $b \to s\gamma$ decays: $\overline{\Lambda} = 0.35 \pm 0.13\,\mathrm{GeV}$ [28]. These can be related to the parameters of the Fermi motion model, see Ref. [30], to yield $p_F = \sqrt{\frac{2}{3}(-\lambda_1)} = 0.40^{+0.08}_{-0.11}$ GeV and $p_F = (\sqrt{\pi}/2)\overline{\Lambda} = 0.31 \pm 0.11$ GeV. As a result, the range of Fermi motion parameter p_F is thus 200 to 480 MeV.

- Fraction of $B \to K\ell^+\ell^-$ and $B \to K^*\ell^+\ell^-$ decays

  The signal model requires knowledge of the ratio between the branching fractions for exclusive $B \to K\ell^+\ell^-$ and $B \to K^*\ell^+\ell^-$ decays and inclusive $b \to s\ell^+\ell^-$ decays. It relies on the calculations and uncertainties from Ref. [1], as shown in Table 2. We assume the branching fractions for $B \to K\ell^+\ell^-$ and $B \to K^*\ell^+\ell^-$ to be fully correlated (Gudrun Hiller indicates that these are more correlated than they are uncorrelated). We therefore vary both branching fractions together by $\pm 1\sigma$.

- Transition point between K^* to higher mass hadronic states

  The signal model assumes that the hadronic system near the K^* resonance is saturated by it below $M_{X_s} = 1.1$ GeV (and above the kaon mass). The transition point is varied by ± 0.1 GeV.

- Hadronization

  The inclusive Monte Carlo generator relies on Jetset to fragment and hadronize the system consisting of the final state s quark and the spectator quark from the B meson. Since the signal efficiencies depend strongly on the particle content of the final state, uncertainties in the fraction of charged and neutral kaon and pions translate into a significant uncertainty in the signal efficiency (for $m_{X_s} > 1.1$ GeV). We use measurements of various particle production observables (see below) performed in the semi-inclusive $B \to X_s\,\gamma$ BaBar

analysis [31] to determine this uncertainty.

First, the ratio between modes containing a K_S^0 and those containing a charged K is varied according to $K_S^0/K = 0.50 \pm 0.05$, to account for possible isospin violation. The resulting relative uncertainty is $\pm 0.5\%$ in both electron and muon channels. Second, the ratio between modes containing one π^0 and those containing none is varied according to $\pi^0/\pi = 1.0 \pm 0.5$. Third, the ratio between two-body and three-body hadronic systems is varied according to $2-\text{body}/3-\text{body} = 0.5 \pm 0.3$. The overall hadronization uncertainty in the electron and muon channels is $\pm 4.0\%$ and $\pm 3.3\%$, respectively.

The last two variations are motivated by the fact that the $b \to s\gamma$ data indicates the need for the fragmentation model to increase the ratio between modes including a π^0 and those without π^0 by a factor of 1.5, and decrease the ratio between 2- and 3-body hadronic states by a factor of 0.4.

- Missing modes

  The 10 modes selected in this analysis only capture about 50% of the full set of final states. Approximately half of the missing modes is due to final states with a K_L^0 meson and can be determined from the K_S^0 modes. However, we need to account for the uncertainty in modes with too many pions or kaons (2 extra kaons may be produced via $s\bar{s}$ popping), as well as for modes with photons that do not originate from π^0 decays but rather from η, η', ω, etc.

  The fraction of modes with ≥ 2 π^0 is varied around the generator MC value of 0.20 (for $m_{X_s} > 1.1$ GeV/c^2) by $\pm 50\%$.

  The fraction of modes with either 0 π^0 and ≥ 3 charged pions or 1 π^0 and ≥ 2 charged pions is varied around the generator MC value of 0.26 (for $m_{X_s} > 1.1$ GeV/c^2) by $\pm 50\%$.

The fraction of modes with more than 1 kaon is varied around the generator MC value of 0.034 (for $m_{X_s} > 1.1$ GeV/c^2) by $\pm 50\%$.

The fraction of modes with at least one photon that does not originate from the decay of a π^0 is varied around the generator MC value of 0.011 (for $m_{X_s} > 1.1$ GeV/c^2) by $\pm 50\%$.

The overall systematic error due to missing mode uncertainties is 3.2% and 2.5% in the electron and muon modes, respectively.

- Monte Carlo statistics

  The finite signal MC statistics translates into a relative uncertainty of 1.6%, 1.8%, and 1.3% in the electron, muon, and electron+muon channels.

- Number of $B\overline{B}$ events

  This is taken to be $(88.9 \pm 1.0) \times 10^6$.

The signal model uncertainties amount to $^{+19.6}_{-16.3}\%$ in the electron modes, $^{+20.8}_{-18.3}\%$ in the muon modes, and $^{+19.9}_{-16.9}\%$ in the combined $e+\mu$ modes. The other uncertainties in the efficiency amount to $\pm 9.9\%$, $\pm 9.5\%$, and $\pm 9.7\%$, respectively. The total uncertainties in the signal efficiency is $^{+21.9}_{-19.1}\%$ (e), $^{+22.9}_{-20.7}\%$ (μ), and $^{+22.2}_{-19.5}\%$ (e+μ). Relative uncertainties in the signal yields are 4.3% (e), 8.0% (μ), and 4.4% (e+μ). Tables 21, 22 and 10.1 summarize the results of the systematics studies.

Table 21: List of systematic uncertainties (in %) affecting the branching fraction calculation: uncertainties in the signal yield, the signal efficiency, and the number of B mesons.

Source	$X_s\, e^+e^-$	$X_s\, \mu^+\mu^-$	$X_s\, \ell^+\ell^-$
Signal shape	± 4.23	± 2.77	± 3.74
Peaking bkg	± 0.78	± 7.52	± 2.25
Signal yield total	± 4.31	± 8.02	± 4.36
Tracking efficiency	± 4.10	± 4.05	± 4.08
Lepton ID efficiency	± 1.70	± 3.20	± 2.17
Kaon ID efficiency	± 0.76	± 0.77	± 0.76
$\pi^\pm$ ID efficiency	± 0.94	± 0.90	± 0.93
K^0_S efficiency	± 0.85	± 0.77	± 0.82
π^0 efficiency	± 0.67	± 0.57	± 0.64
LR cut efficiency	± 3.86	± 3.32	± 3.69
Cross-feed efficiency	± 5.53	± 5.53	± 5.53
Detector model subtotal	± 8.23	± 8.40	± 8.25
Fermi motion model	$^{-1.04}_{+10.90}$	$^{-1.09}_{+9.92}$	$^{-1.06}_{+10.59}$
$\mathcal{B}(B \to K^*\ell^+\ell^-)$	± 15.58	± 17.86	± 16.30
K^*–X_s transition	± 4.68	± 4.08	± 4.49
Signal model subtotal	$^{+19.58}_{-16.30}$	$^{+20.83}_{-18.35}$	$^{+19.95}_{-16.94}$
Hadronization	± 4.05	± 3.28	± 3.80
Missing modes	± 3.24	± 2.52	± 3.01
MC statistics	± 1.63	± 1.84	± 1.26
Efficiency total	$^{+21.93}_{-19.05}$	$^{+22.91}_{-20.68}$	$^{+22.16}_{-19.50}$
$B\overline{B}$ counting	± 1.13	± 1.13	± 1.13

Table 22: List of systematic uncertainties (in %) affecting the branching fraction calculation in the different hadronic mass ranges (see Table 19): uncertainties in the signal yield, the signal efficiency, and the number of B mesons.

Source	$m(X_s)$ bin 1	$m(X_s)$ bin 2	$m(X_s)$ bin 3	$m(X_s)$ bin 4
Signal shape	± 4.79	± 0.72	± 5.09	± 3.31
Peaking bkg	± 2.93	± 1.77	± 2.21	± 1.98
Signal yield total	± 5.61	± 1.91	± 5.55	± 3.86
Tracking efficiency	± 3.62	± 4.42	± 4.26	± 4.55
Lepton ID efficiency	± 2.22	± 2.29	± 2.15	± 2.10
Kaon ID efficiency	± 0.77	± 0.83	± 0.74	± 0.76
$\pi^{\pm}$ ID efficiency	± 0.88	± 0.62	± 1.00	± 0.91
K^0_S efficiency	± 0.00	± 1.24	± 1.17	± 1.61
π^0 efficiency	± 0.00	± 0.57	± 0.87	± 1.35
LR cut efficiency	± 3.69	± 3.69	± 3.69	± 3.69
Cross-feed efficiency	± 1.93	± 6.32	± 8.54	± 9.26
Detector model subtotal	± 6.06	± 9.02	± 10.63	± 11.42
Fermi motion model	$^{0.00}_{0.00}$	$^{0.00}_{0.00}$	$^{0.00}_{0.00}$	$^{-0.44}_{+7.60}$
$\mathcal{B}(B \to K^* \ell^+ \ell^-)$	± 0.00	± 0.00	± 0.00	± 2.34
K^*–X_s transition	± 0.00	± 0.00	± 0.01	± 10.61
Signal model subtotal	$^{+0.00}_{-0.00}$	$^{+0.00}_{-0.00}$	$^{+0.01}_{-0.01}$	$^{+13.26}_{-10.87}$
Hadronization	± 0.00	± 0.00	± 0.00	± 17.08
Missing modes	± 0.00	± 0.00	± 0.00	± 12.62
MC statistics	± 0.47	± 8.44	± 1.35	± 5.37
Efficiency total	$^{+6.08}_{-6.08}$	$^{+12.35}_{-12.35}$	$^{+10.71}_{-10.71}$	$^{+28.03}_{-26.99}$
$B\overline{B}$ counting	± 1.12	± 1.12	± 1.12	± 1.12

Table 10.1: List of systematic uncertainties (in %) affecting the branching fraction calculation in different dilepton mass ranges (see Table 19): uncertainties in the signal yield, the signal efficiency, and the number of B mesons.

Source / $m(\ell^+\ell^-)$ bin	bin 1	bin 2	bin 3	bin 4	bin 5	bin 6	bin 7
Signal shape	±6.14	±4.75	±3.99	±2.49	±3.25	±5.11	±3.25
Peaking bkg	±6.64	±1.57	±3.98	±3.96	±5.21	±2.80	±5.21
Signal yield total	±9.04	±5.00	±5.64	±4.68	±6.14	±5.83	±6.14
Tracking efficiency	±4.28	±4.06	±4.10	±4.10	±3.96	±4.08	±3.95
Lepton ID efficiency	±2.02	±2.00	±2.12	±2.47	±2.29	±2.03	±2.29
Kaon ID efficiency	±0.75	±0.72	±0.76	±0.79	±0.77	±0.75	±0.77
$\pi^\pm$ ID efficiency	±0.95	±1.07	±0.91	±0.80	±0.87	±0.98	±0.87
K^0_S efficiency	±1.17	±0.86	±0.84	±0.78	±0.57	±0.85	±0.57
π^0 efficiency	±0.86	±0.81	±0.69	±0.54	±0.30	±0.76	±0.30
LR cut efficiency	±3.69	±3.69	±3.69	±3.69	±3.69	±3.69	±3.69
Cros-feed efficiency ±44.87	±2.75	±4.32	±6.53	±9.66	±3.98	±9.66	
Detector model subtotal	±45.31	±6.69	±7.49	±9.01	±11.38	±7.28	±11.38
Fermi motion model	$^{-6.31}_{+41.90}$	$^{-3.54}_{+2.61}$	$^{0.83}_{+5.86}$	$^{+3.40}_{-11.25}$	$^{+1.65}_{-9.61}$	$^{-2.43}_{+3.14}$	$^{+1.11}_{-7.26}$
$\mathcal{B}(B \to K^*\ell^+\ell^-)$	±18.52	±18.84	±15.99	±12.69	±5.59	±18.64	±3.38
K^*–X_s transition	±11.37	±8.02	±3.60	±1.76	±3.61	±6.64	±3.45
Signal model subtotal	$^{+47.20}_{-2.63}$	$^{+6.00}_{-2.78}$	$^{+17.43}_{-6.39}$	$^{-3.26}_{-17.05}$	$^{+6.86}_{-11.69}$	$^{+0.44}_{-9.93}$	$^{+4.96}_{-8.72}$
Hadronization	±5.84	±5.03	±4.95	±2.41	±0.34	±4.94	±0.44
Missing modes	±8.17	±5.91	±2.94	±1.72	±1.74	±5.09	±1.22
MC statistics	±4.63	±2.47	±2.22	±4.12	±1.46	±1.62	±1.14
Efficiency total	$^{+6.35}_{-51.84}$	$^{+37.51}_{-2.30}$	$^{-9.95}_{-9.05}$	$^{-6.81}_{-9.94}$	$^{-3.49}_{-6.48}$	$^{+2.13}_{-2.43}$	$^{-2.53}_{-14.44}$
$B\overline{B}$ counting	±1.12	±1.12	±1.12	±1.12	±1.12	±1.12	±1.12

10.3 CP asymmetry systematics

All systematic errors cancel by definition in the measurement of the *CP* asymmetry, except for systematic effects which are different for b and $\overline{b}$ final states. The effect of these differences is to fake a *CP* asymmetry. There are many known effects which can give rise to the fake *CP* asymmetry. We determine the systematic uncertainty using the charmonium veto sample. As a cross check, we also study the impact of individual detector modeling uncertainties.
In the following section, we will first describe the charmonium veto sample method and then the itemized method.

Charmonium veto sample method

We use the charmonium veto sample to estimate the *CP* asymmetry systematic uncertainty. We consider the difference in the peaking component of B and $\overline{B}$ charmonium veto sample as the uncertainty, assuming the *CP* asymmetry in B decays to charmonium to be zero. The advantage of this method is that since the uncertainty is directly estimated from data, all effects that may fake a *CP* asymmetry are included, except mis-ID biases (however this effect is very small as described later). This is of course only strictly valid in the dilepton mass regions of the charmonium veto cuts but we assume this restricted range to be representative of the full range included in the analysis. The result of separate fits to the B and $\overline{B}$ data is shown in Fig. 75. We find the bias to be $A_{CP}(J/\psi\ \psi(2S)) = -0.005 \pm 0.016$, where the error is statistical only.

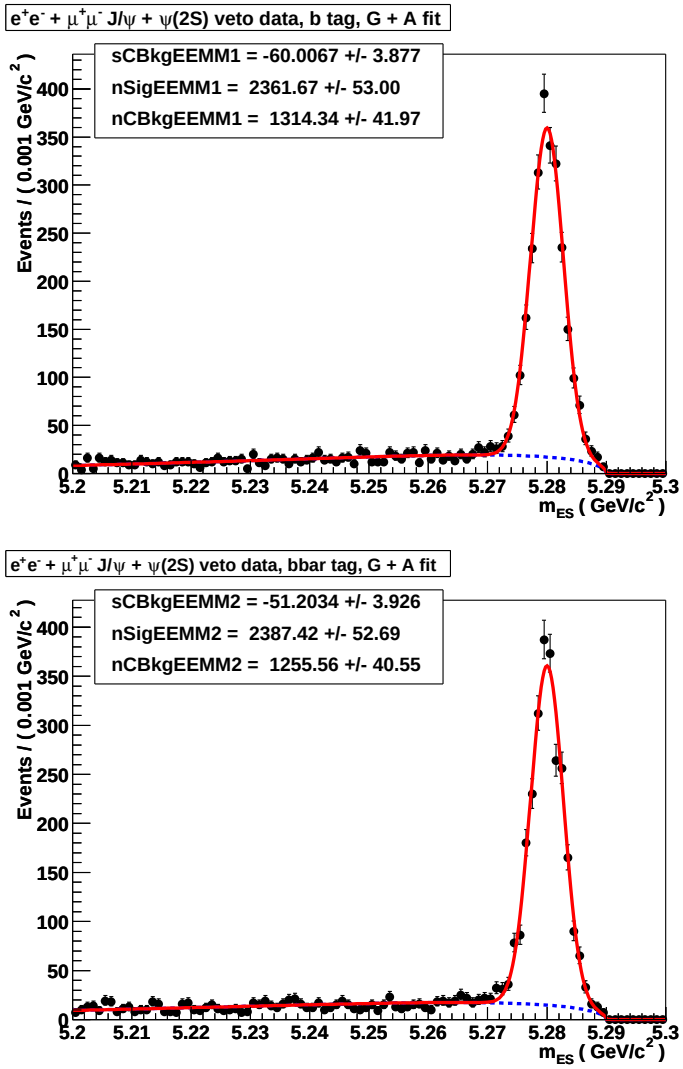

Figure 75: Distributions of m_{ES} for combined electron and muon data in self-tagged $\overline{B}$ (top) and B (bottom) decays from the charmonium veto sample. The red line is the result of the fit, with the total background component shown in the dashed blue line.

Table 23: Apparent CP asymmetry in reconstructed signal events due to tracking, kaon ID and pion ID efficiencies biases. The corresponding statistical error and the error due to tracking, kaon ID and pion ID efficiencies errors are shown for the computed CP asymmetry.

	A_{CP}
Trk eff.	$0.0094 \pm 0.0076 \pm 0.0017$
Kaon ID	$0.0027 \pm 0.0076 \pm 0.0002$
Pion ID	$0.0002 \pm 0.0076 \pm 0.0016$

Itemized method

We will first describe the fake CP asymmetry due to detector biases which affects the signal events, then we will describe the mis–tagged events and the corresponding asymmetry and finally we will investigate the CP asymmetry in cross–feed events.

- **CP asymmetry in signal events due to efficiency biases** The detector efficiency is not perfectly symmetric for positive and negative charged tracks. Tracking efficiency differences can arise from the opposite curvature of the tracks in the solenoidal field. In the Drift Chamber this couples to the drift electron direction, which is distorted by the Lorentz angle effect. This can affect both the track reconstruction and the dE/dx particle identification information. The other major source of asymmetry is the different strong interaction rates of positive and negative particles. This leads to tracking efficiency asymmetries and particle identification asymmetries.

  We use the efficiency tables for positive and negative charged tracks, kaons and electrons, provided by the Tracking Efficiency Task Force, to analyze the apparent CP asymmetry due to the particle reconstruction bias. The results are shown in Table 23. There is no evidence of bias within the errors.

- **CP asymmetry in mis–tagged signal MC events due to mis–ID biases**

Some signal events are not properly tagged in the correct flavor due to detector effects. This can generate a fake *CP* asymmetry. We study only those events that are still classified as "true" signal events, i.e. reconstructed in the same or opposite flavor. Events reconstructed with opposite flavor are what we call "mistagged" events.

In the following, first we study the mis-tag rate in our events, then we investigate the corresponding flavor asymmetry.

We estimate the mis–tag rate using reconstructed signal MC event. The mis-tag rate is defined as the percentage of times the wrong b–flavor is reconstructed with respect to all the events which pass the selection. We find the mis-tagged rate $\alpha_{tag} = 0.0002 \pm 0.0076$, where the error is statistical only. The mis-tagged rate is very small.

In order to estimate the flavor asymmetry in the misidentified events, we compute the asymmetry of the wrongly reconstructed b flavor events versus the wrongly reconstructed $\overline{b}$ events, normalized by the total events which pass the selection. The result is $A_{CP}(\text{Mis-tagged rate}) = -0.0002 \pm 0.0076$, where the error is statistical only. The asymmetry is consistent with zero. We also check the fake *CP* asymmetry in mis-tagged events due to $K^{\pm} \leftrightarrow \pi^{\pm}$ and $K^{\pm} \leftrightarrow \mu^{\pm}$ mis-IDs by using PID look up tables for positive and negative charged particles. The results are summarized in Table 24. The asymmetries are zero.

- ***CP* asymmetry in cross–feed MC events due to detector effects** There can also be a *CP* asymmetry in the cross–feed events due to detector biases. In this section we consider only the "peaking" component of the cross–feed.

  There are no recipes available for estimating the rates for replacing or adding

Table 24: Fake A_{CP} in mis-tagged events due to $K^{\pm} \leftrightarrow \pi^{\pm}$ and $K^{\pm} \leftrightarrow \mu^{\pm}$ mis-IDs. The corresponding statistical error and the error due to $K^{\pm} \leftrightarrow \pi^{\pm}$ and $K^{\pm} \leftrightarrow \mu^{\pm}$ mis-ID errors are shown for the computed *CP* asymmetry.

	A_{CP}
$K^{\pm} \leftrightarrow \pi^{\pm}$ mis-id	$-0.0000 \pm 0.0076 \pm 0.0000$
$K^{\pm} \leftrightarrow \mu^{\pm}$ mis-id	$-0.0000 \pm 0.0076 \pm 0.0000$

Systematic	
Trk eff.	1.12
Kaon ID	0.28
Pion ID	0.18
Mis-tag rate	0.02
$K^{\pm} \leftrightarrow \pi^{\pm}$ mis-id	0.00
$K^{\pm} \leftrightarrow \mu^{\pm}$ mis-id	0.00
MC stat	0.76
Total	1.49

Table 25: List of fractional systematic uncertainties (in %) affecting the A_{CP} calculation.

particles, so we use the peaking component of the cross–feed events from the reconstructed signal MC samples. Result of the difference in the peaking component of b and $\overline{b}$ cross-feed events is considered as a bias of the asymmetry. We obtained the bias from the cross–feed events to be $A_{CP}(\text{cross–feed}) = 0.14 \pm 0.18.$, where the error is statistical only. As the statistical error is very large (i.e. we do not expect many events from the peaking background components) we do not observe any asymmetry.

We summarize the list of systematic uncertainties for *CP* asymmetry with the itemized method in Table 25.

CHAPTER 11

Summary

We have observed the inclusive $b \to s\ell^+\ell^-$ process in the $B \to X_s\, e^+e^-$ and $B \to X_s\, \mu^+\mu^-$ channels with a sum of exclusive modes technique. The inclusive branching fraction is determined to be

$$\mathcal{B}(B \to X_s\, \ell^+\ell^-) = (5.6 \pm 1.5(\text{stat}) \pm 0.6(\text{exp syst}) \pm 1.1(\text{model syst})) \times 10^{-6}, \tag{11.1}$$

for $m_{ll} > 0.2$ GeV. Tables 26 and 27 summarize the other results of the analysis. These agree with the theoretical prediction of $(4.2 \pm 0.7) \times 10^{-6}$ [32] for $m_{\ell\ell} > 0.2$ GeV, $(1.63 \pm 0.20) \times 10^{-6}$ [33] for $1 < m_{\ell\ell} < 2.45$ GeV, and $(0.40 \pm 0.08) \times 10^{-6}$ [33] for $m_{\ell\ell} > 3.795$ GeV. We determine the direct CP asymmetry to be $A_{CP} = -0.22 \pm 0.26(stat) \pm 0.02(syst)$, in agreement with the vanishingly small asymmetry expected in the SM [25].

Figure 76 shows the results in terms of number of signal events (errors are statistical only) as a function of hadronic mass and dilepton mass. The figure shows that higher mass hadronic states contribute a significant portion of the total signal. Similarly, Fig. 77 shows the differential branching fraction as a function of hadronic and dilepton mass. In this case, systematic uncertainties in the branching fractions are included in quadrature in the outermost vertical error bars.

Table 26: Summary of results: signal yield, signal significance, signal efficiency and branching fraction. In the case of the signal yield, the first error is statistical and the second error is systematic. In the case of the signal efficiency, the first error corresponds to the experimental systematic uncertainty arising from detector modeling, hadronization, $B\overline{B}$ counting, and Monte Carlo statistics, whereas the second error corresponds to the uncertainties in the signal model. In the case of the branching fraction, the first, second, and third errors correspond to statistical, experimental systematic, and signal model systematic uncertainties, respectively.

Sample	N_{sig}	Signif.	ϵ (%)	$\mathcal{B}$ ($\times 10^{-6}$)
$X_s\, e^+e^-$	$29.22 \pm 8.28 \pm 1.26$	4.0	$2.74 \pm 0.27 \pm 0.49$	$6.00 \pm 1.70 \pm 0.65 \pm 1.09$
$X_s\, \mu^+\mu^-$	$11.25 \pm 6.19 \pm 0.90$	2.0	$1.26 \pm 0.12 \pm 0.25$	$5.04 \pm 2.77 \pm 0.63 \pm 1.01$
$X_s\, \ell^+\ell^-$	$40.06 \pm 10.35 \pm 1.75$	4.3	$2.00 \pm 0.19 \pm 0.37$	$5.64 \pm 1.46 \pm 0.60 \pm 1.06$

Table 27: Summary of branching fractions, where the first error is statistical, the second is systematic excluding signal model uncertainties, and the third is from signal model systematics.

Sample	$\mathcal{B}$ ($\times 10^{-6}$)	Total σ_{syst}
e^+e^-	$5.9987 \pm 1.6991 \pm^{0.6729}_{0.6275} \pm^{1.1475}_{1.0362}$	$\pm^{1.3302}_{1.2114}$
$\mu^+\mu^-$	$5.0373 \pm 2.7742 \pm^{0.6459}_{0.6164} \pm^{1.1172}_{0.9099}$	$\pm^{1.2905}_{1.0990}$
$\ell^+\ell^-$	$5.6396 \pm 1.4571 \pm^{0.6222}_{0.5813} \pm^{1.1312}_{0.9874}$	$\pm^{1.2910}_{1.1458}$
$m(X_s)(\,\mathrm{GeV}/c^2)$		
$0.4 - 0.6$	$0.5279 \pm 0.1719 \pm^{0.0449}_{0.0434} \pm^{0.0000}_{0.0000}$	$\pm^{0.0449}_{0.0434}$
$0.6 - 0.8$	$0.3223 \pm 0.2407 \pm^{0.0432}_{0.0380} \pm^{0.0000}_{0.0000}$	$\pm^{0.0432}_{0.0380}$
$0.8 - 1.0$	$0.6442 \pm 0.4018 \pm^{0.0826}_{0.0743} \pm^{0.0001}_{0.0001}$	$\pm^{0.0826}_{0.0743}$
$1.0 - 1.8$	$3.4960 \pm 2.0605 \pm^{1.0092}_{0.7769} \pm^{0.4234}_{0.4240}$	$\pm^{1.0944}_{0.8851}$
$m(\ell^+\ell^-)(\,\mathrm{GeV}/c^2)$		
$0.2 - 1.0$	$0.0776 \pm 0.3650 \pm^{0.0644}_{0.0267} \pm^{0.0209}_{0.0271}$	$\pm^{0.0677}_{0.0381}$
$1.0 - 2.0$	$1.6343 \pm 0.6354 \pm^{0.1988}_{0.1849} \pm^{0.4097}_{0.4704}$	$\pm^{0.4554}_{0.5054}$
$2.0 - m_{J/\psi}$	$1.7508 \pm 0.8394 \pm^{0.2036}_{0.1921} \pm^{0.3396}_{0.2675}$	$\pm^{0.3960}_{0.3293}$
$m_{J/\psi} - m_{\psi'}$	$0.9786 \pm 0.8450 \pm^{0.1162}_{0.1076} \pm^{0.1895}_{0.1161}$	$\pm^{0.2223}_{0.1582}$
$m_{\psi'} - 5.0$	$0.6356 \pm 0.3175 \pm^{0.0894}_{0.0793} \pm^{0.0809}_{0.0416}$	$\pm^{0.1206}_{0.0895}$
$1.0 - 2.45$	$1.8466 \pm 0.7163 \pm^{0.2268}_{0.2127} \pm^{0.4452}_{0.4666}$	$\pm^{0.4996}_{0.5128}$
$3.795 - 5.0$	$0.4976 \pm 0.2486 \pm^{0.0695}_{0.0616} \pm^{0.0462}_{0.0239}$	$\pm^{0.0835}_{0.0661}$

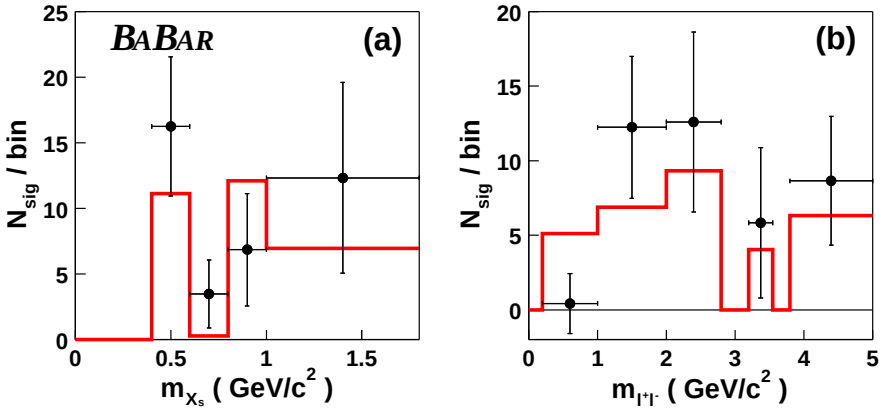

Figure 76: Number of fitted signal events in bins of m_{Xs} (left) and m_{ll} for electron and muon channels combined for data (points) and MC signal (histogram). The error bars represent the statistical uncertainty only.

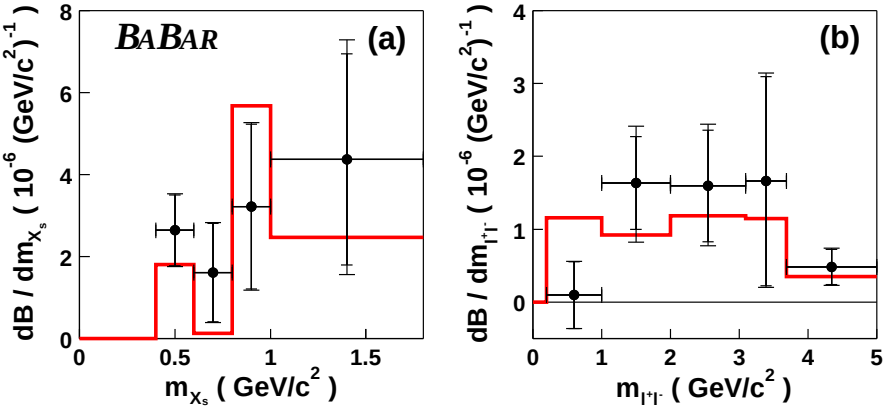

Figure 77: Differential branching fraction as a function of hadronic mass (left) and dilepton mass (right), averaged over electron and muon channels for data (points) and signal Monte Carlo (histogram).

APPENDIX A

Signal efficiency tables

The tables in this section show the MC signal efficiency for different stages of the analysis. The numbers correspond to number of events in 81.9 fb^{-1}. The entries in [%] show the efficiency of the corresponding step. The last column lists the overall efficiencies.

Table 28: Signal MC efficiency for electron modes.

Topol	gen	skim + prVtx	in lists	preSel	lept Sel	had Sel	postSel	jpsi veto	LR cut	tot Eff
other	298.102	279.981 93.921%	56.539 20.194%	1.101 1.947%	0.900 81.793%	0.000 0.000%	0.000 0.000%	0.000 0.000%	0.000 0.000%	0.000%
K	3.405	32.104 96.107%	17.795 55.430%	14.977 84.164%	14.590 97.415%	13.661 93.633%	11.250 82.348%	7.256 64.502%	5.588 77.009%	16.728%
K pi0	53.557	51.498 96.155%	8.238 15.997%	6.866 8.344%	6.315 91.972%	4.637 73.431%	3.778 81.476%	2.482 65.687%	1.361 54.825%	2.540%
K pi	95.148	90.858 95.491%	35.978 39.598%	29.011 80.635%	27.333 94.215%	20.876 76.377%	16.619 79.609%	10.781 64.873%	6.410 59.458%	6.737%
K pi pi0	42.139	40.451 95.994%	4.639 11.468%	3.707 79.913%	3.270 8.203%	1.997 61.063%	1.379 69.073%	1.069 77.521%	0.419 39.189%	0.994%
K pi pi	37.254	35.066 94.127%	10.132 28.895%	7.717 76.157%	6.979 90.438%	4.592 65.804%	3.282 71.460%	2.509 76.454%	1.109 44.195%	2.976%
K pi pi pi0	27.568	25.785 93.532%	1.391 5.393%	1.020 73.331%	0.866 84.882%	0.421 48.614%	0.209 49.653%	0.177 84.770%	0.000 0.000%	0.000%
K pi pi pi	7.853	7.198 91.653%	1.294 17.983%	0.854 65.996%	0.699 81.841%	0.414 59.260%	0.187 45.098%	0.144 77.009%	0.000 0.000%	0.000%
Ks+Kl	30.880	29.873 96.739%	4.822 16.143%	4.206 87.209%	4.087 97.170%	3.816 93.389%	3.157 82.728%	2.017 63.875%	1.519 75.304%	4.918%
Ks pi0	24.119	23.172 96.075%	2.110 9.105%	1.873 8.763%	1.794 95.798%	1.207 67.294%	0.959 79.427%	0.625 65.163%	0.353 56.481%	1.463%
Ks pi	52.629	50.269 95.516%	10.699 21.283%	8.867 82.883%	8.288 93.471%	6.297 75.972%	5.127 81.426%	3.353 65.386%	2.074 61.861%	3.941%
Ks pi pi0	23.240	22.187 95.469%	1.425 6.422%	1.168 81.976%	1.027 87.943%	0.580 56.428%	0.419 72.210%	0.316 75.423%	0.111 35.040%	0.476%
Ks pi pi	17.396	16.310 93.758%	2.709 16.611%	2.212 81.633%	2.023 91.483%	1.357 67.064%	0.955 70.390%	0.770 80.617%	0.303 39.334%	1.741%
Ks pi pi pi0	11.655	10.934 93.817%	0.255 2.337%	0.220 86.259%	0.198 90.043%	0.100 50.572%	0.044 43.739%	0.044 100.000%	0.000 0.000%	0.000%
Ks pi pi pi	4.452	4.114 92.412%	0.462 11.242%	0.309 66.775%	0.264 85.329%	0.170 64.335%	0.078 46.128%	0.060 77.208%	0.000 0.000%	0.000%
Total	759.397	719.800 94.786%	158.490 22.019%	84.107 53.068%	78.633 93.491%	60.125 76.463%	47.443 78.906%	31.602 66.612%	19.246 60.899%	2.534%

Table 29: Signal MC efficiency for muon modes.

Topol	gen	skim + prVtx	in lists	preSel	lept Sel	had Sel	postSel	jpsi veto	LR cut	tot Eff
other	295.036	278.068 94.249%	18.516 6.659%	0.385 2.079%	0.282 73.360%	0.000 0.000%	0.000 0.000%	0.000 0.000%	0.000 0.000%	0.000%
K	33.422	32.125 96.119%	6.884 21.430%	6.446 93.640%	5.886 91.298%	5.673 96.386%	4.625 81.535%	3.441 74.393%	2.968 86.268%	8.881%
K pi0	54.157	52.071 96.148%	2.946 5.658%	2.746 93.211%	2.286 83.245%	1.854 81.110%	1.584 85.428%	1.125 71.001%	0.592 52.625%	1.093%
K pi	95.483	91.194 95.508%	12.732 13.961%	11.340 89.067%	9.780 86.241%	8.391 85.801%	6.765 80.627%	4.857 71.788%	2.825 58.164%	2.959%
K pi pi0	40.036	38.495 96.152%	1.385 3.599%	1.224 88.365%	0.883 72.104%	0.666 75.443%	0.488 73.202%	0.367 75.380%	0.146 39.851%	0.366%
K pi pi	38.799	36.632 94.415%	2.967 8.100%	2.462 82.964%	1.867 75.847%	1.434 76.796%	1.118 77.957%	0.869 77.716%	0.388 44.627%	0.999%
K pi pi pi0	27.086	25.371 93.666%	0.470 1.853%	0.366 77.809%	0.253 69.080%	0.168 66.630%	0.106 63.003%	0.091 85.364%	0.000 0.000%	0.000%
K pi pi pi	7.690	7.049 91.664%	0.423 6.000%	0.367 86.667%	0.298 81.373%	0.189 63.237%	0.104 5.172%	0.084 80.248%	0.000 0.000%	0.000%
Ks+Kl	30.866	29.891 96.842%	1.831 6.126%	1.754 95.789%	1.576 89.840%	1.525 96.774%	1.261 82.676%	0.936 74.255%	0.796 85.021%	2.579%
Ks pi0	24.368	23.462 96.283%	0.664 2.830%	0.639 96.201%	0.532 83.268%	0.418 78.494%	0.325 77.936%	0.222 68.120%	0.120 54.114%	0.492%
Ks pi	52.470	50.230 95.731%	3.813 7.591%	3.511 92.066%	2.956 84.201%	2.439 82.502%	2.005 82.205%	1.448 72.232%	0.880 60.756%	1.677%
Ks pi pi0	22.780	21.993 96.548%	0.426 1.935%	0.396 93.030%	0.286 72.318%	0.189 66.003%	0.141 74.343%	0.104 74.354%	0.045 42.679%	0.196%
Ks pi pi	16.708	15.877 95.030%	0.679 4.277%	0.607 89.354%	0.465 76.606%	0.324 69.681%	0.249 76.794%	0.197 79.330%	0.086 43.661%	0.516%
Ks pi pi pi0	12.245	11.494 93.867%	0.080 0.700%	0.080 100.000%	0.055 68.644%	0.044 79.386%	0.028 64.850%	0.028 100.000%	0.000 0.000%	0.000%
Ks pi pi pi	4.569	4.224 92.448%	0.074 1.754%	0.074 100.000%	0.052 70.613%	0.026 49.137%	0.020 78.937%	0.020 100.000%	0.000 0.000%	0.000%
Total	75.714	718.177 95.033%	53.892 7.504%	32.397 60.114%	27.457 84.752%	23.339 85.001%	18.819 80.636%	13.789 73.271%	8.846 64.150%	1.171%

A.1 Background rejection tables

In this section we list the background suppression rates for different stages of the analysis, as estimated in the MC. The table entries are in the $m_{\rm ES}$ range $5.20 < m_{\rm ES} < 5.29$ GeV. The numbers correspond to number of events in 81.9 fb^{-1}. The entries in [%] show the efficiency of the corresponding step. The last column lists the overall efficiencies.

Table 30: Background MC rejection for electron modes. The table entries are in the $m_{\rm ES}$ range $5.20 < m_{\rm ES} < 5.29$ GeV.

Topol	best Combo	preSel	postSel	jpsi veto	LR cut	tot Eff
other	0.00	0.00 0.0%	0.00 0.0%	0.00 0.0%	0.00 0.0%	0.0%
K	3391.32	3391.32 100.0%	2463.31 72.6%	247.13 10.0%	22.86 9.3%	0.7%
K pi0	3929.69	3929.69 100.0%	1996.85 50.8%	622.57 31.2%	42.32 6.8%	1.1%
K pi	9057.81	9057.81 100.0%	4941.15 54.6%	1341.56 27.2%	70.97 5.3%	0.8%
K pi pi0	7796.38	7796.38 100.0%	2852.11 36.6%	1099.09 38.5%	35.45 3.2%	0.5%
K pi pi	13523.41	13523.41 100.0%	5350.83 39.6%	1853.75 34.6%	69.80 3.8%	0.5%
K pi pi pi0	6678.81	6678.81 100.0%	1223.63 18.3%	566.38 46.3%	0.00 0.0%	0.0%
K pi pi pi	7171.13	7171.13 100.0%	1525.27 21.3%	682.33 44.7%	0.00 0.0%	0.0%
Ks	1037.72	1037.72 100.0%	730.58 70.4%	76.91 10.5%	4.95 6.4%	0.5%
Ks pi0	1105.39	1105.39 100.0%	514.17 46.5%	167.16 32.5%	12.01 7.2%	1.1%
Ks pi	3786.50	3786.50 100.0%	1937.12 51.2%	577.60 29.8%	31.79 5.5%	0.8%
Ks pi pi0	3209.66	3209.66 100.0%	1071.68 33.4%	401.65 37.5%	11.06 2.8%	0.3%
Ks pi pi	3370.94	3370.94 100.0%	1244.41 36.9%	442.52 35.6%	12.98 2.9%	0.4%
Ks pi pi pi0	1783.55	1783.55 100.0%	312.21 17.5%	155.51 49.8%	0.00 0.0%	0.0%
Ks pi pi pi	2710.58	2710.58 100.0%	500.38 18.5%	235.09 47.0%	0.00 0.0%	0.0%
Total	68552.88	68552.88 100.0%	26663.70 38.9%	8469.24 31.8%	314.19 3.7%	0.5%

Table 31: Background MC rejection for muon modes. The table entries are in the $m_{\rm ES}$ range $5.20 < m_{\rm ES} < 5.29$ GeV.

Topol	best Combo	preSel	postSel	jpsi veto	LR cut	tot Eff
other	0.00	0.00 0.0%	0.00 0.0%	0.00 0.0%	0.00 0.0%	0.0%
K	1430.27	1430.27 100.0%	1185.85 82.9%	193.07 16.3%	30.69 15.9%	2.1%
K pi0	1314.68	1314.68 100.0%	880.04 66.9%	333.57 37.9%	18.35 5.5%	1.4%
K pi	3289.86	3289.86 100.0%	2336.29 71.0%	775.51 33.2%	31.80 4.1%	1.0%
K pi pi0	2428.60	2428.60 100.0%	1251.32 51.5%	496.38 39.7%	18.32 3.7%	0.8%
K pi pi	4352.43	4352.43 100.0%	2321.76 53.3%	919.33 39.6%	25.15 2.7%	0.6%
K pi pi pi0	1845.01	1845.01 100.0%	502.26 27.2%	231.44 46.1%	0.00 0.0%	0.0%
K pi pi pi	1992.88	1992.88 100.0%	620.60 31.1%	261.02 42.1%	0.00 0.0%	0.0%
Ks	417.80	417.80 100.0%	350.80 84.0%	39.92 11.4%	8.92 22.4%	2.1%
Ks pi0	396.67	396.67 100.0%	249.09 62.8%	99.27 39.9%	3.63 3.7%	0.9%
Ks pi	1331.26	1331.26 100.0%	837.06 62.9%	299.54 35.8%	14.44 4.8%	1.1%
Ks pi pi0	939.24	939.24 100.0%	424.45 45.2%	177.16 41.7%	6.69 3.8%	0.7%
Ks pi pi	1075.50	1075.50 100.0%	562.19 52.3%	198.85 35.4%	4.56 2.3%	0.4%
Ks pi pi pi0	452.19	452.19 100.0%	127.28 28.1%	60.04 47.2%	0.00 0.0%	0.0%
Ks pi pi pi	747.28	747.28 100.0%	197.39 26.4%	85.17 43.1%	0.00 0.0%	0.0%
Total	22013.65	22013.65 100.0%	11846.38 53.8%	4170.26 35.2%	162.54 3.9%	0.7%

APPENDIX B

Background suppression variables

Figures 78–94 show the background suppression variables used in the likelihood ratio, for $B \to X_s e^+ e^-$ and $B \to X_s \mu^+ \mu^-$ candidates in the four different categories of events: signal, cross-feed, $B\overline{B}$ and continuum MC.

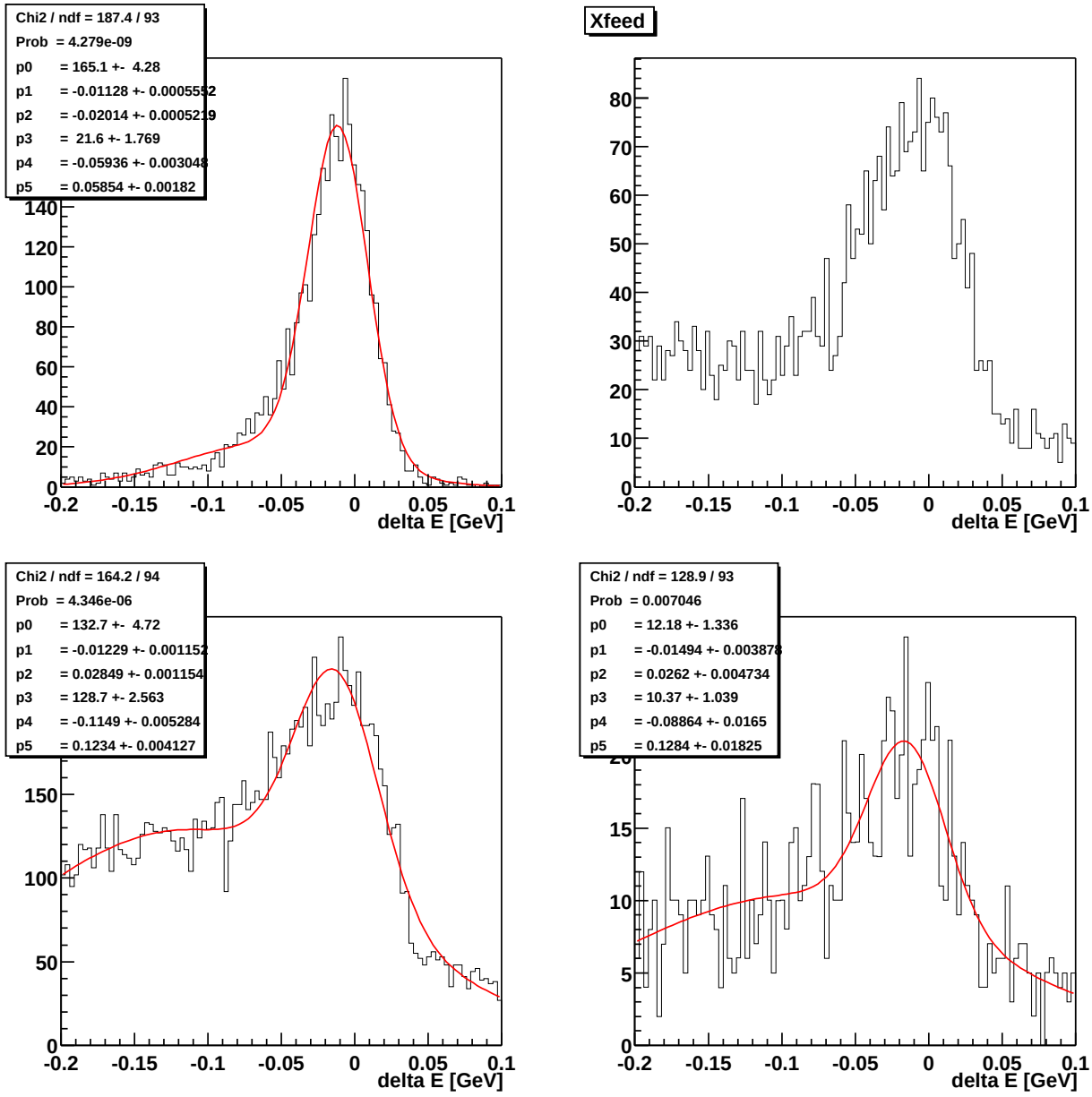

Figure 78: Distributions of ΔE for $B \to X_s\, e^+ e^-$ candidates in truth-matched signal (top left), cross-feed (top right), $B\overline{B}$ (bottom left) and continuum (bottom right) events. The curves show the result of the fits to these distributions using the functions listed in Table 5.

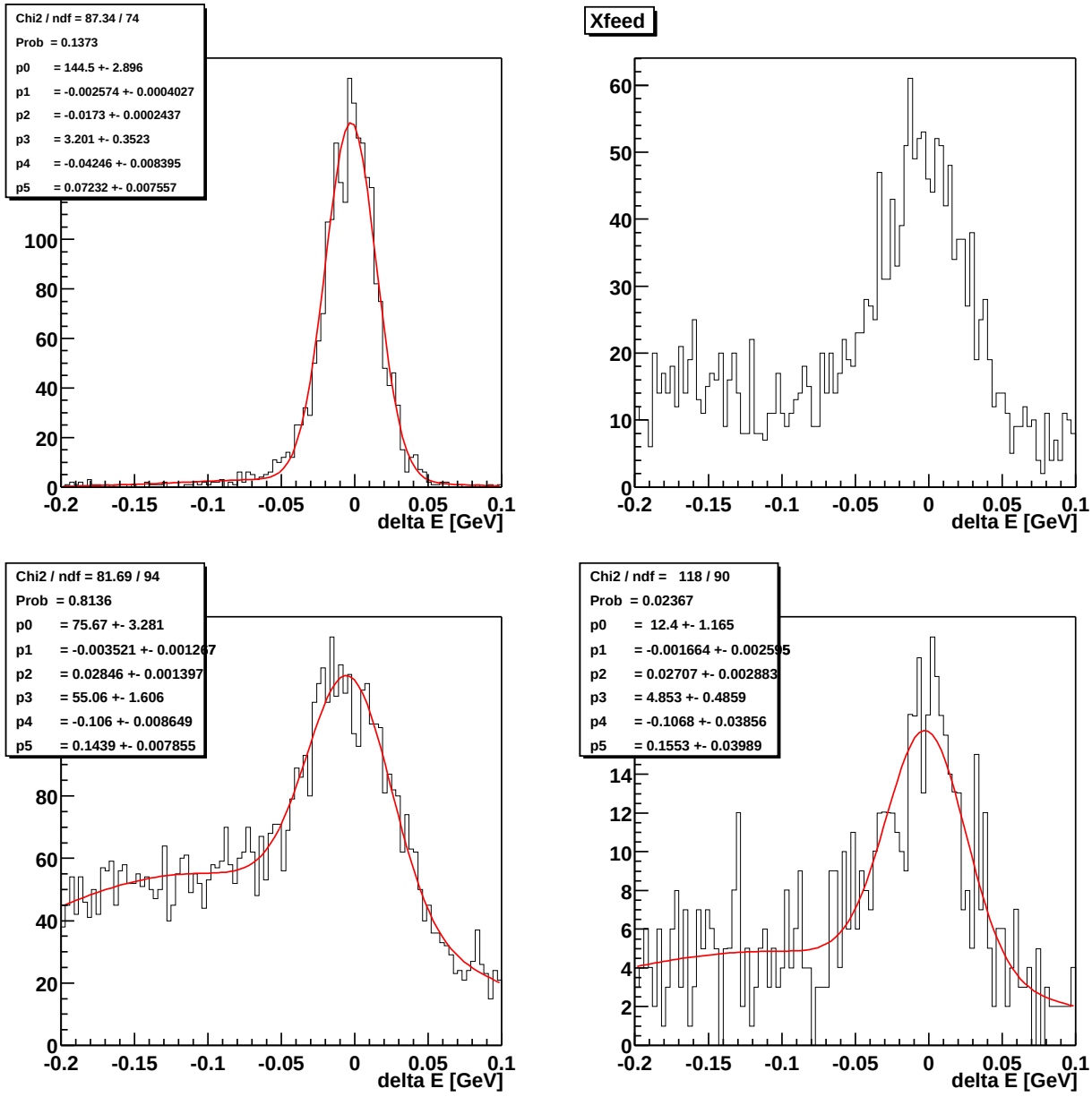

Figure 79: Distributions of ΔE for $B \to X_s \mu^+ \mu^-$ candidates in truth-matched signal (top left), cross-feed (top right), $B\overline{B}$ (bottom left) and continuum (bottom right) events. The curves show the result of the fits to these distributions using the functions listed in Table 5.

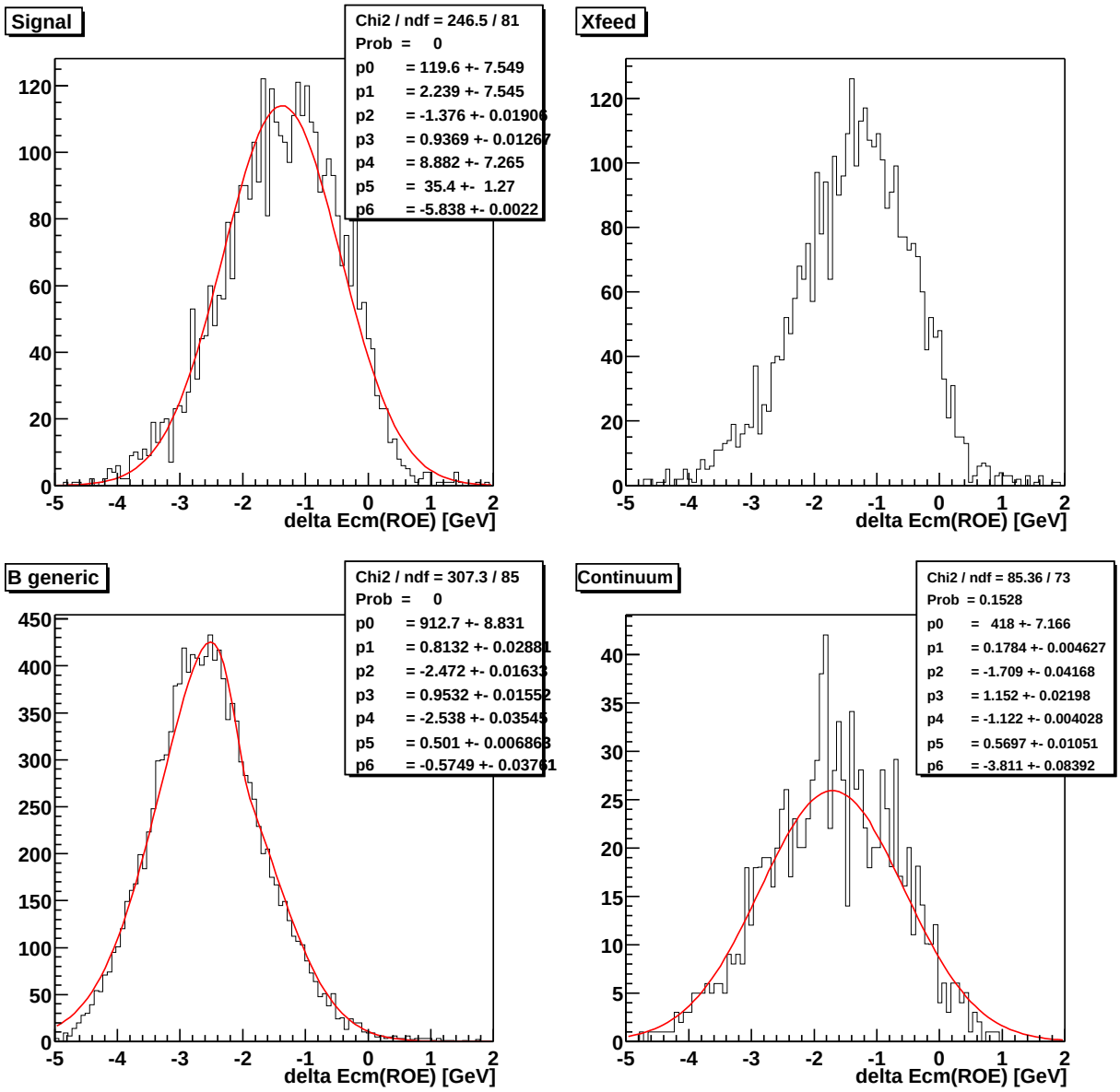

Figure 80: Distributions of ΔE^{ROE} for $B \to X_s\, e^+e^-$ candidates in truth-matched signal (top left), cross-feed (top right), $B\overline{B}$ (bottom left) and continuum (bottom right) events. The curves show the result of the fits to these distributions using the functions listed in Table 5.

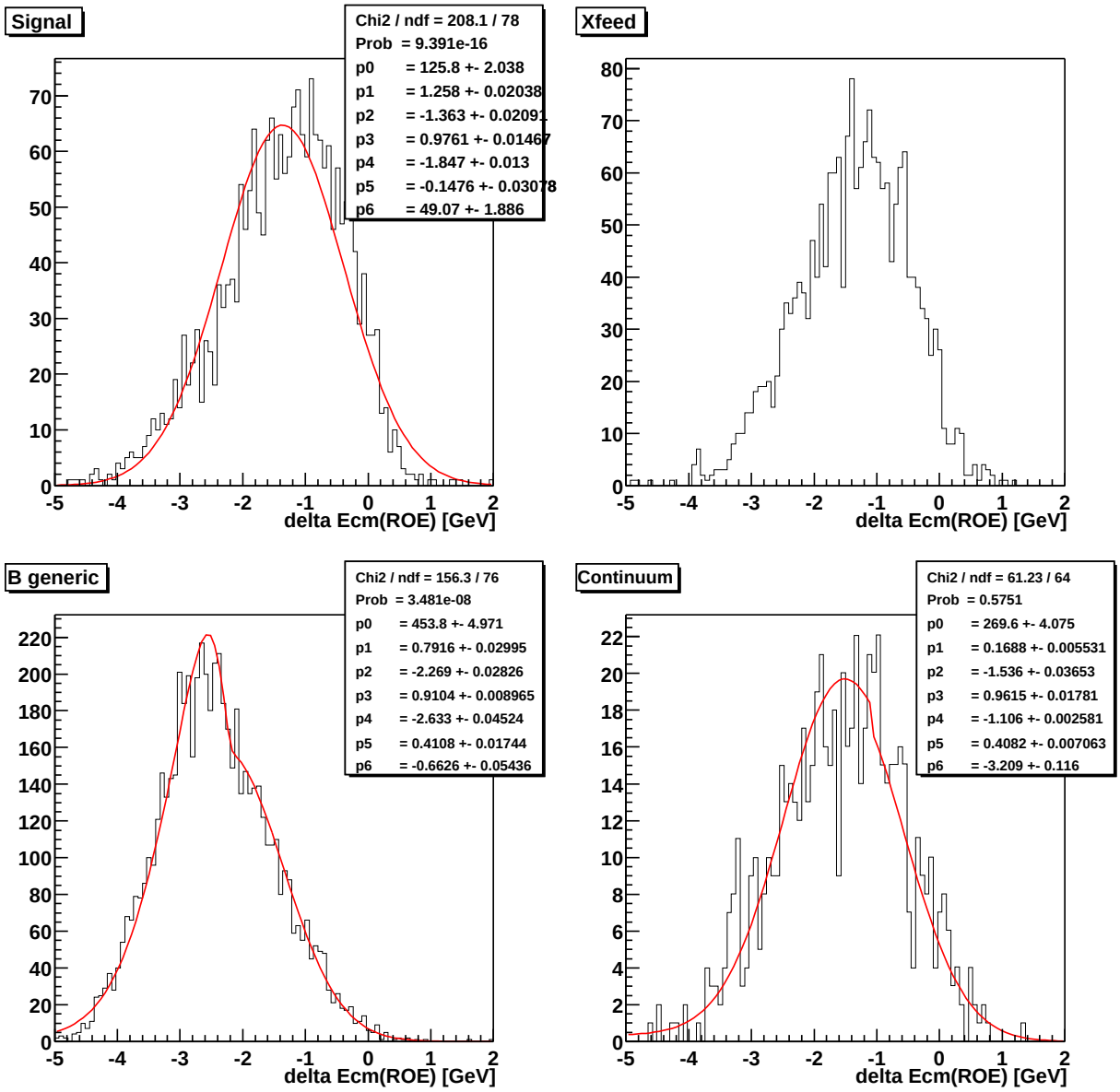

Figure 81: Distributions of ΔE^{ROE} for $B \to X_s\, \mu^+\mu^-$ candidates in truth-matched signal (top left), cross-feed (top right), $B\overline{B}$ (bottom left) and continuum (bottom right) events. The curves show the result of the fits to these distributions using the functions listed in Table 5.

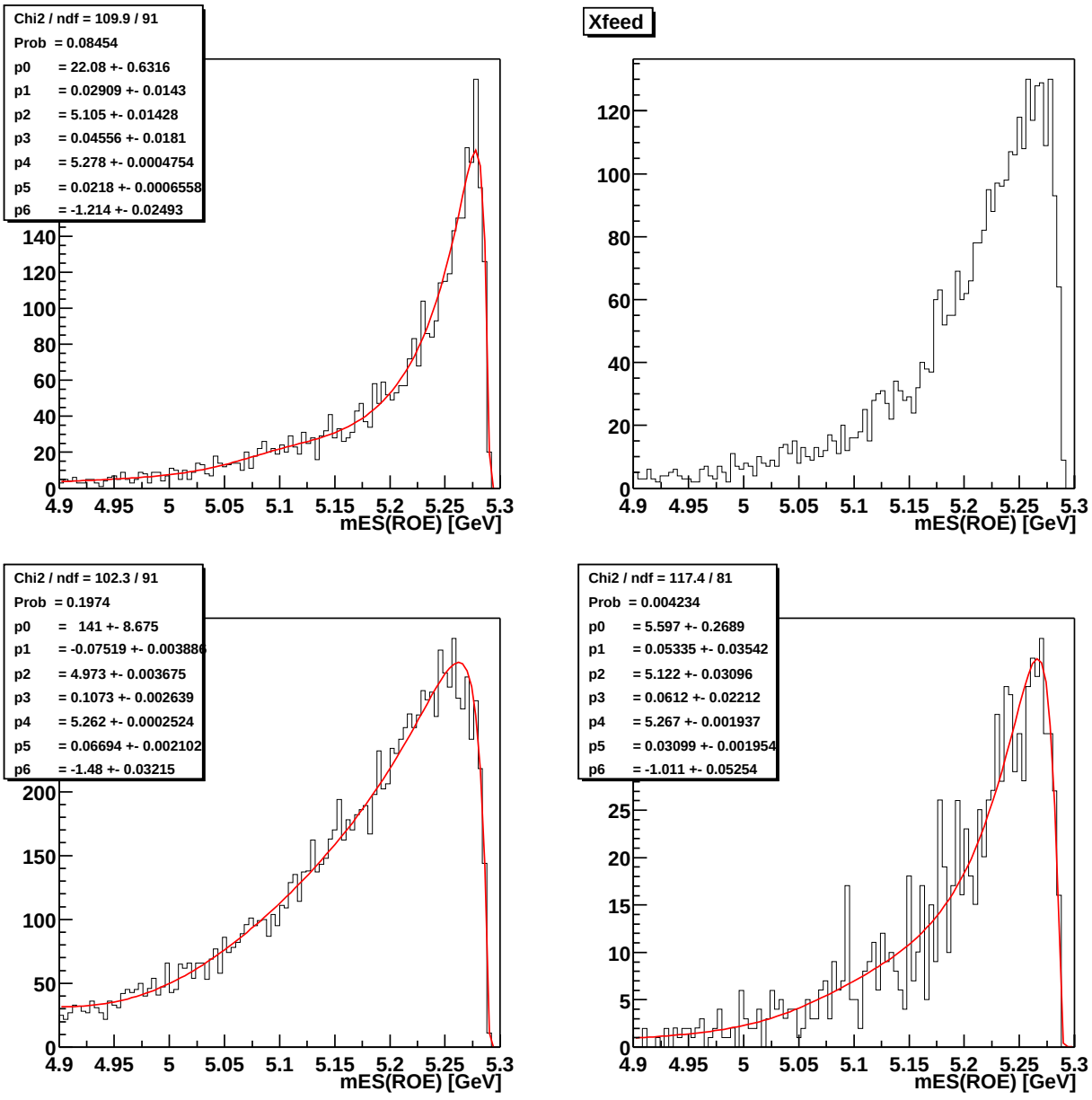

Figure 82: Distributions of m_{ES}^{ROE} for $B \to X_s\, e^+e^-$ candidates in truth-matched signal (top left), cross-feed (top right), $B\overline{B}$ (bottom left) and continuum (bottom right) events. The curves show the result of the fits to these distributions using the functions listed in Table 5.

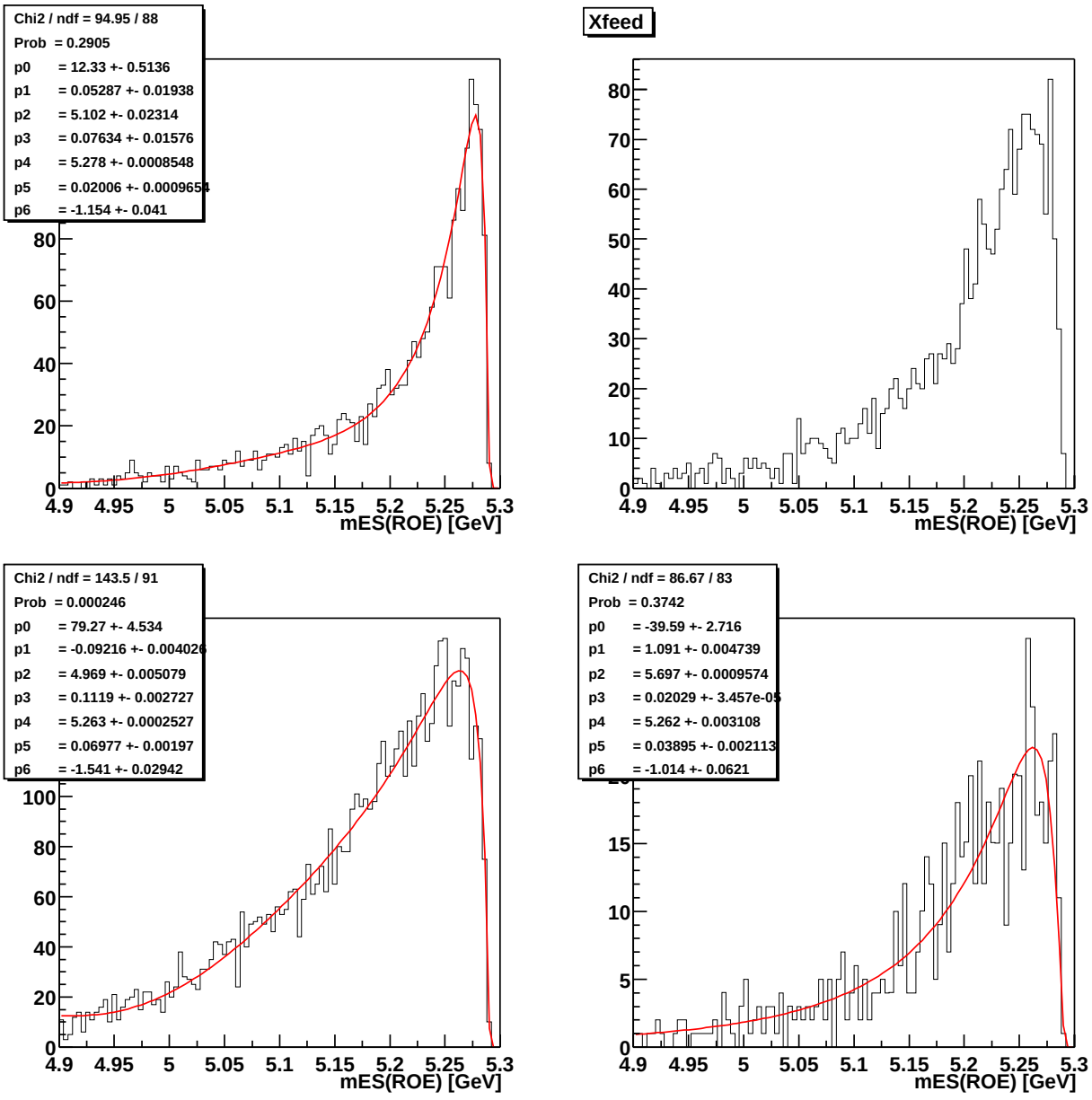

Figure 83: Distributions of m_{ES}^{ROE} for $B \rightarrow X_s\ \mu^+\mu^-$ candidates in truth-matched signal (top left), cross-feed (top right), $B\overline{B}$ (bottom left) and continuum (bottom right) events. The curves show the result of the fits to these distributions using the functions listed in Table 5.

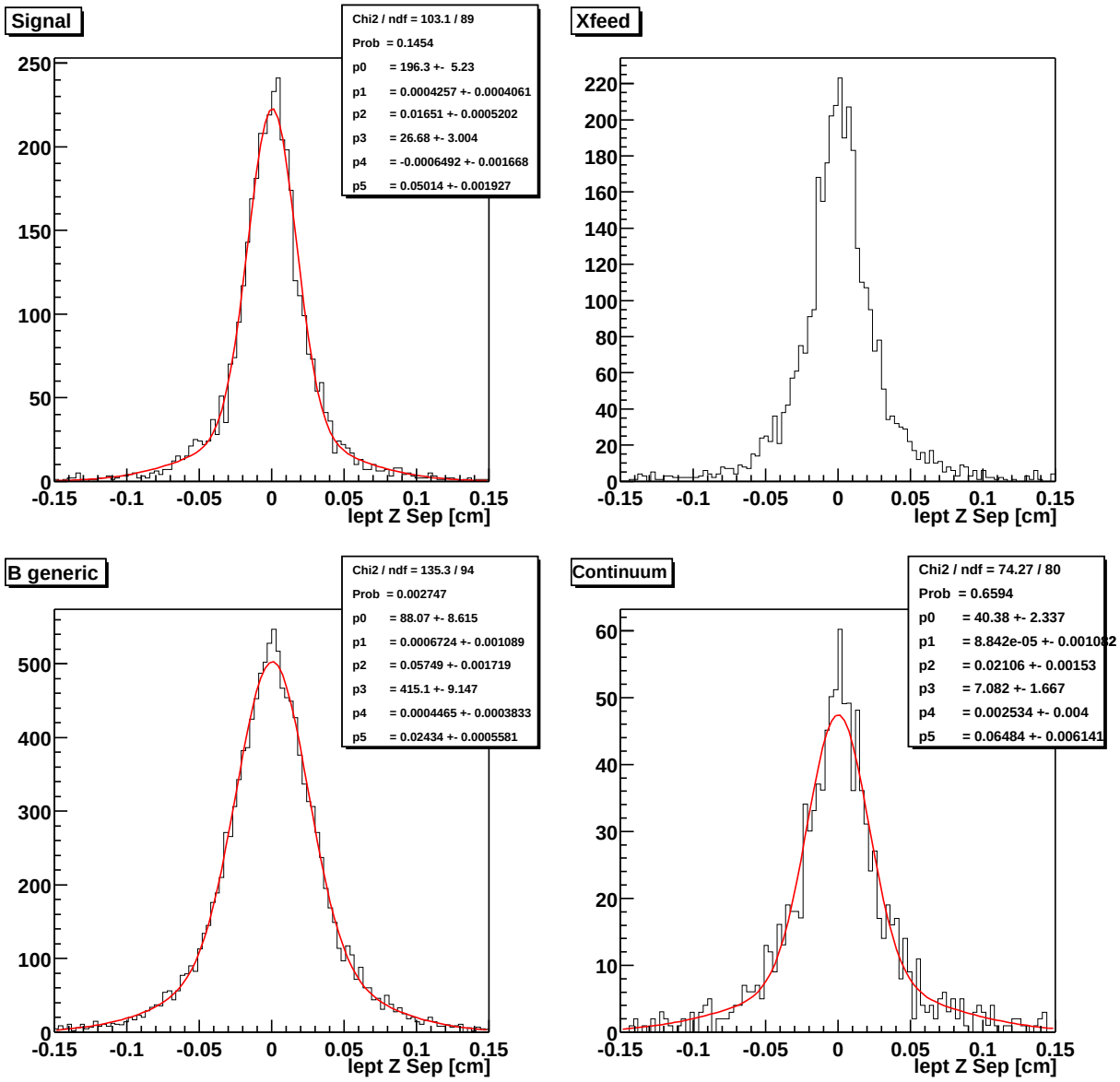

Figure 84: Distributions of Δz between the leptons for $B \to X_s\, e^+e^-$ candidates in truth-matched signal (top left), cross-feed (top right), $B\overline{B}$ (bottom left) and continuum (bottom right) events. The curves show the result of the fits to these distributions using the functions listed in Table 5.

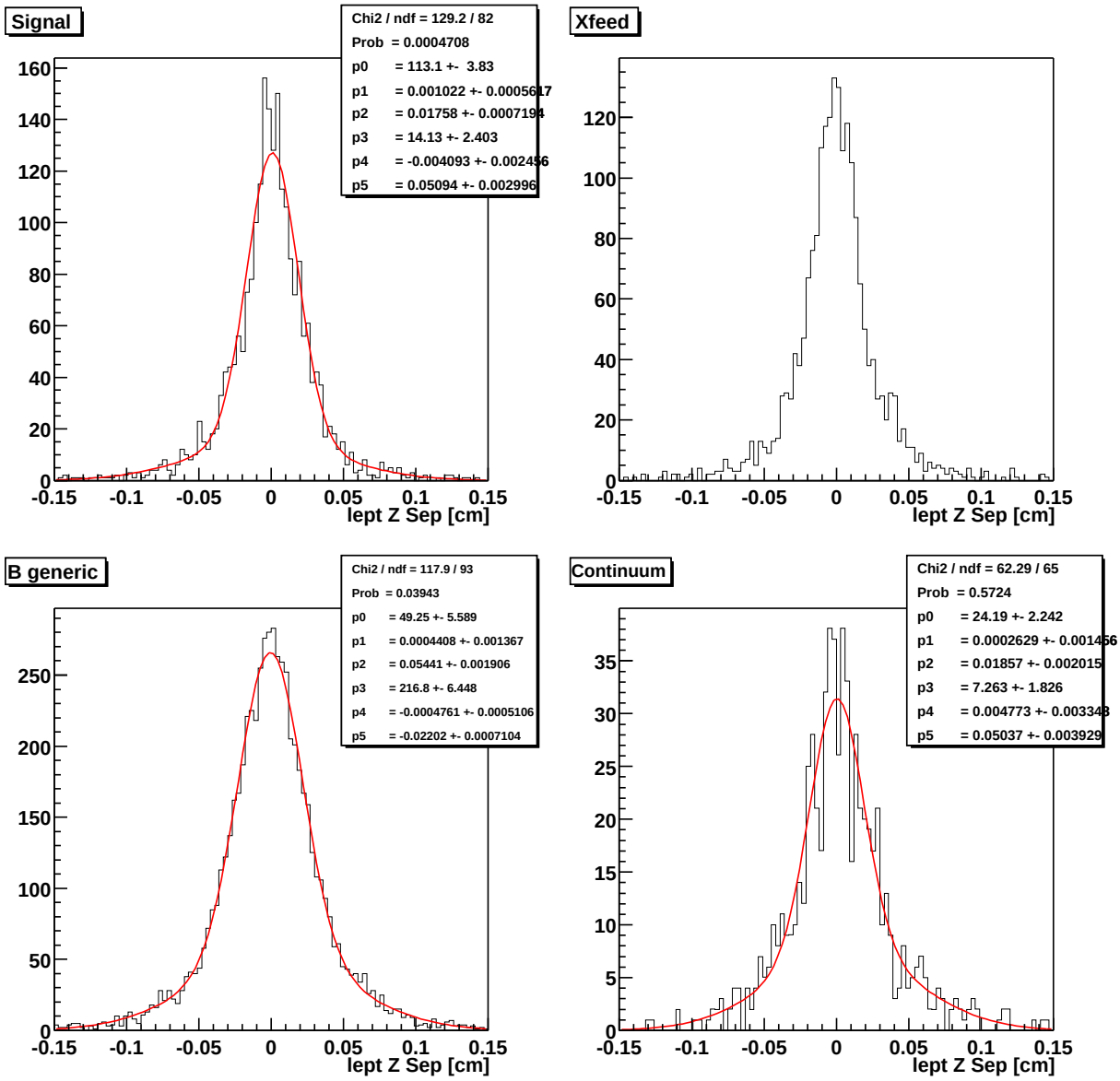

Figure 85: Distributions of Δz between the leptons for $B \to X_s\,\mu^+\mu^-$ candidates in truth-matched signal (top left), cross-feed (top right), $B\overline{B}$ (bottom left) and continuum (bottom right) events. The curves show the result of the fits to these distributions using the functions listed in Table 5.

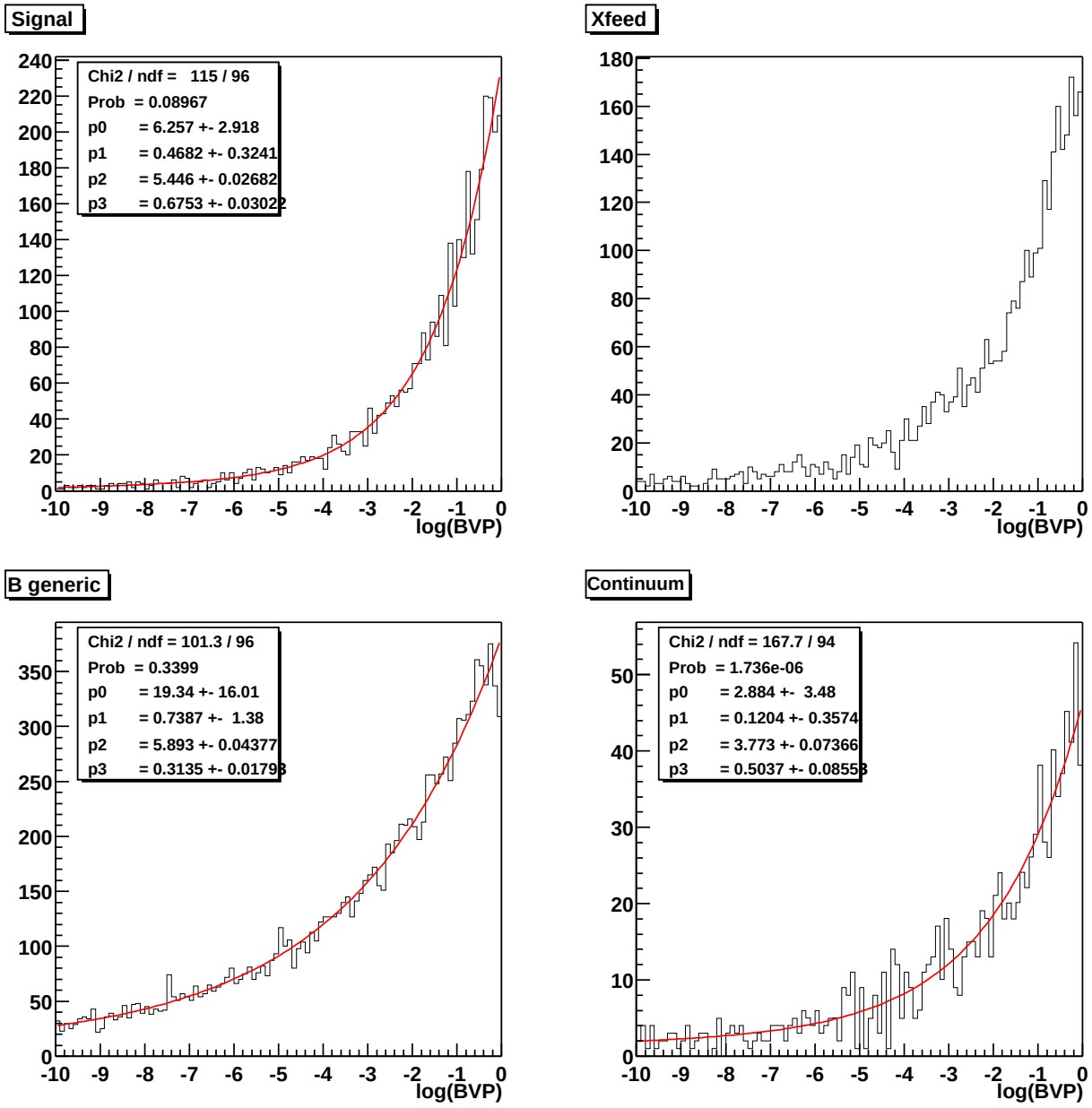

Figure 86: Distributions of $\log(P_{Bvtx})$ for $B \to X_s\, e^+e^-$ candidates in truth-matched signal (top left), cross-feed (top right), $B\overline{B}$ (bottom left) and continuum (bottom right) events. The curves show the result of the fits to these distributions using the functions listed in Table 5.

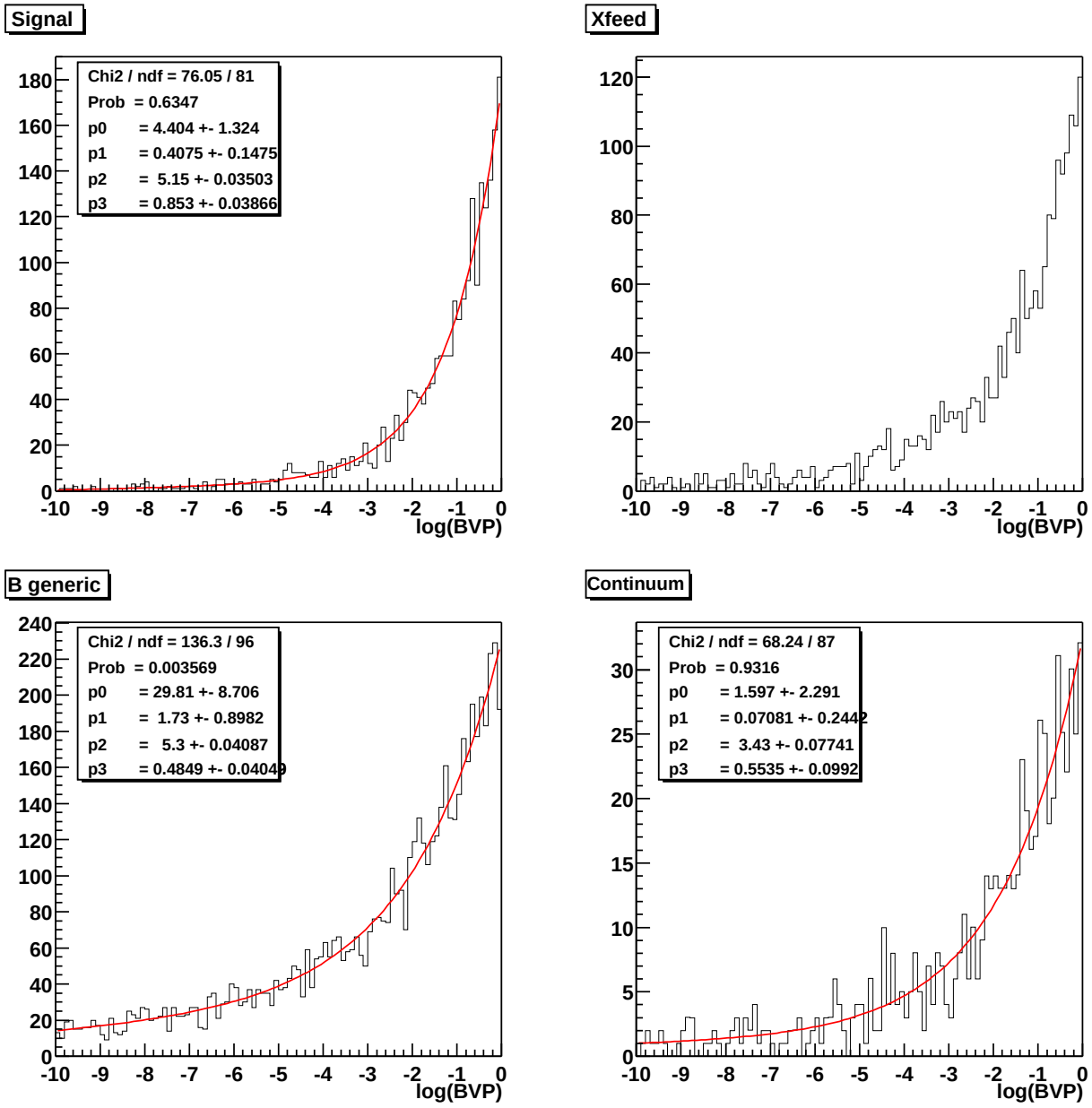

Figure 87: Distributions of $\log(P_{Bvtx})$ for $B \to X_s\, \mu^+\mu^-$ candidates in truth-matched signal (top left), cross-feed (top right), $B\overline{B}$ (bottom left) and continuum (bottom right) events. The curves show the result of the fits to these distributions using the functions listed in Table 5.

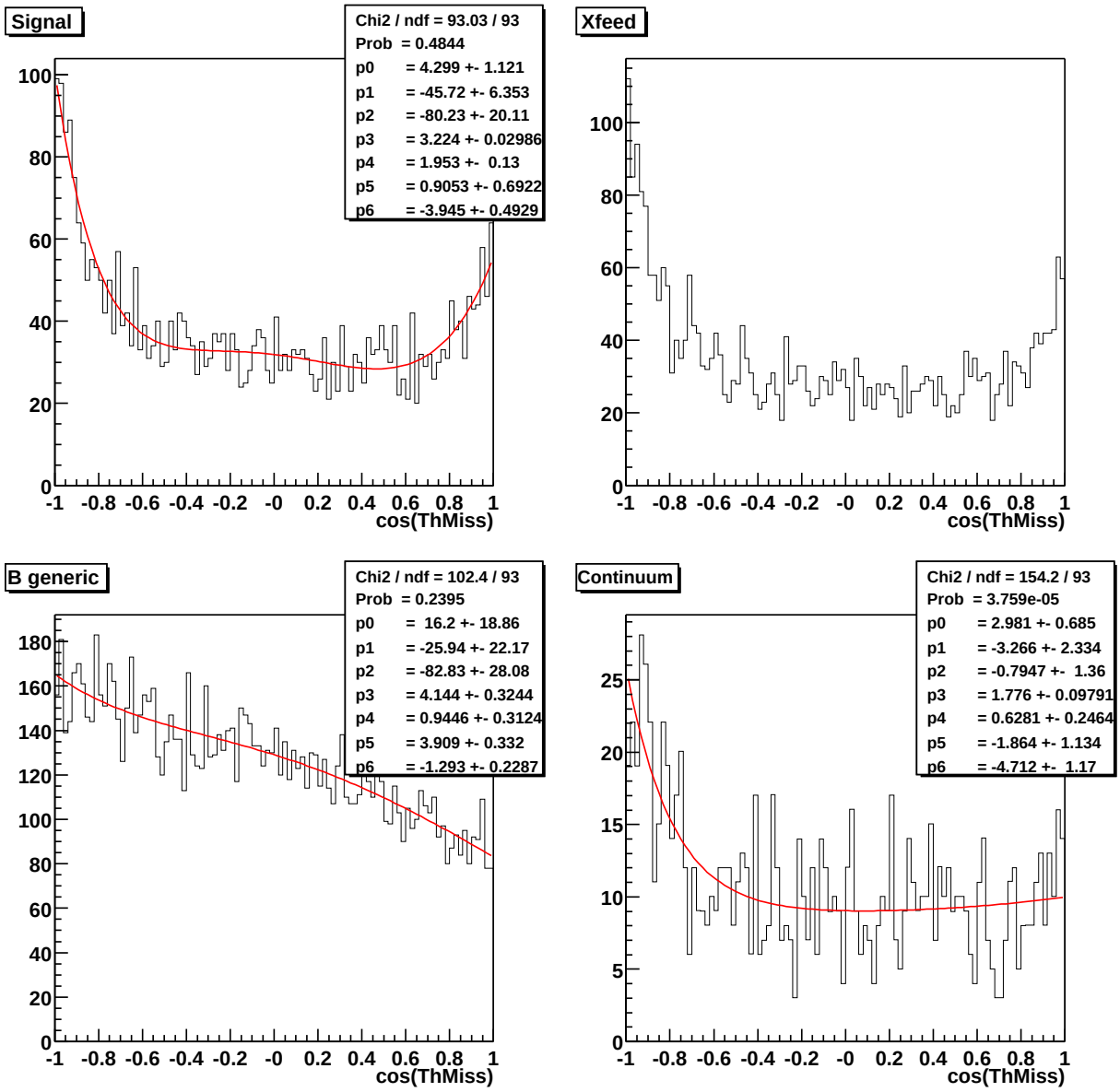

Figure 88: Distributions of $\cos\theta_{miss}$ for $B \to X_s\, e^+e^-$ candidates in truth-matched signal (top left), cross-feed (top right), $B\overline{B}$ (bottom left) and continuum (bottom right) events. The curves show the result of the fits to these distributions using the functions listed in Table 5.

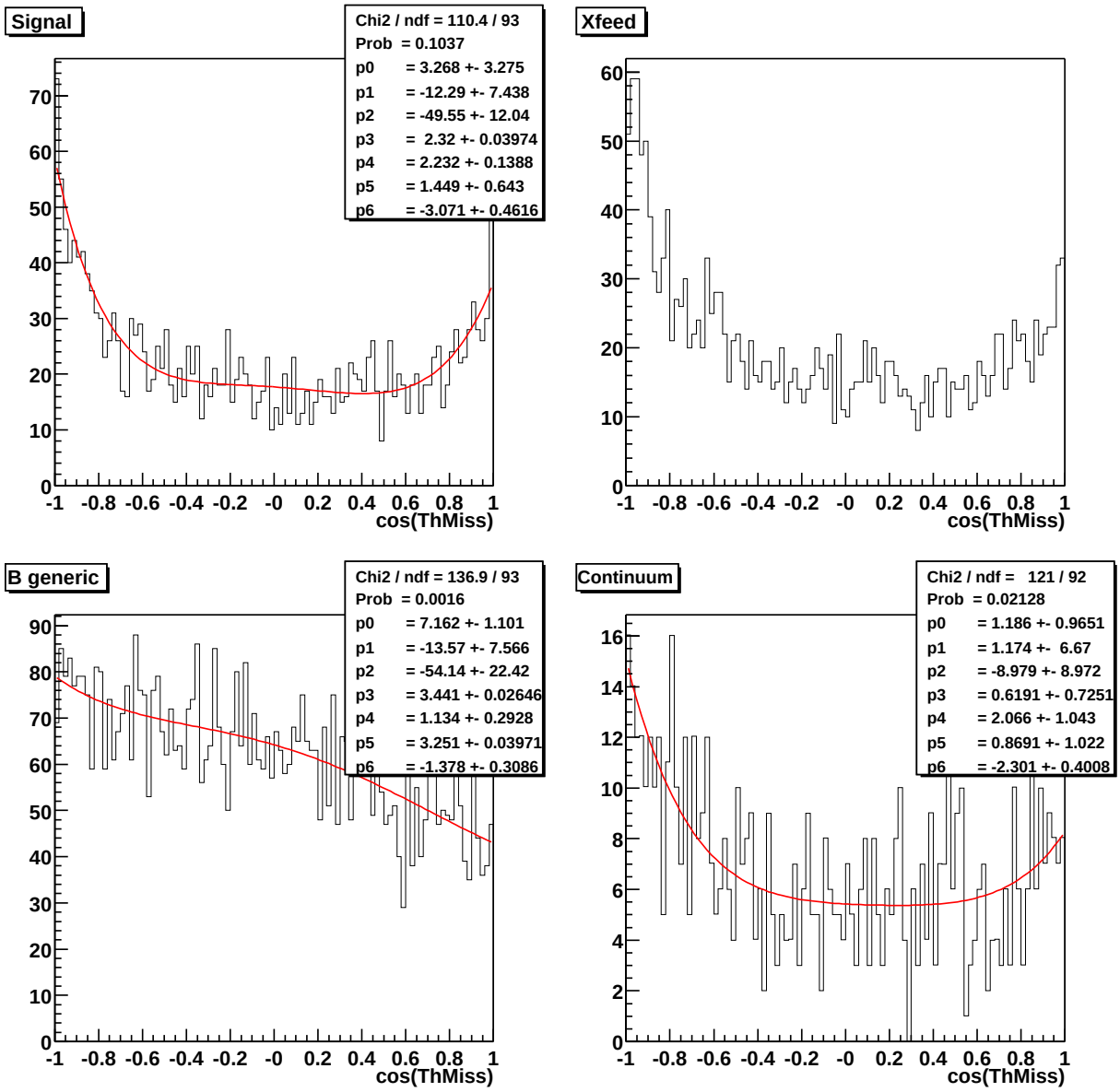

Figure 89: Distributions of $\cos\theta_{miss}$ for $B \to X_s\ \mu^+\mu^-$ candidates in truth-matched signal (top left), cross-feed (top right), $B\overline{B}$ (bottom left) and continuum (bottom right) events. The curves show the result of the fits to these distributions using the functions listed in Table 5.

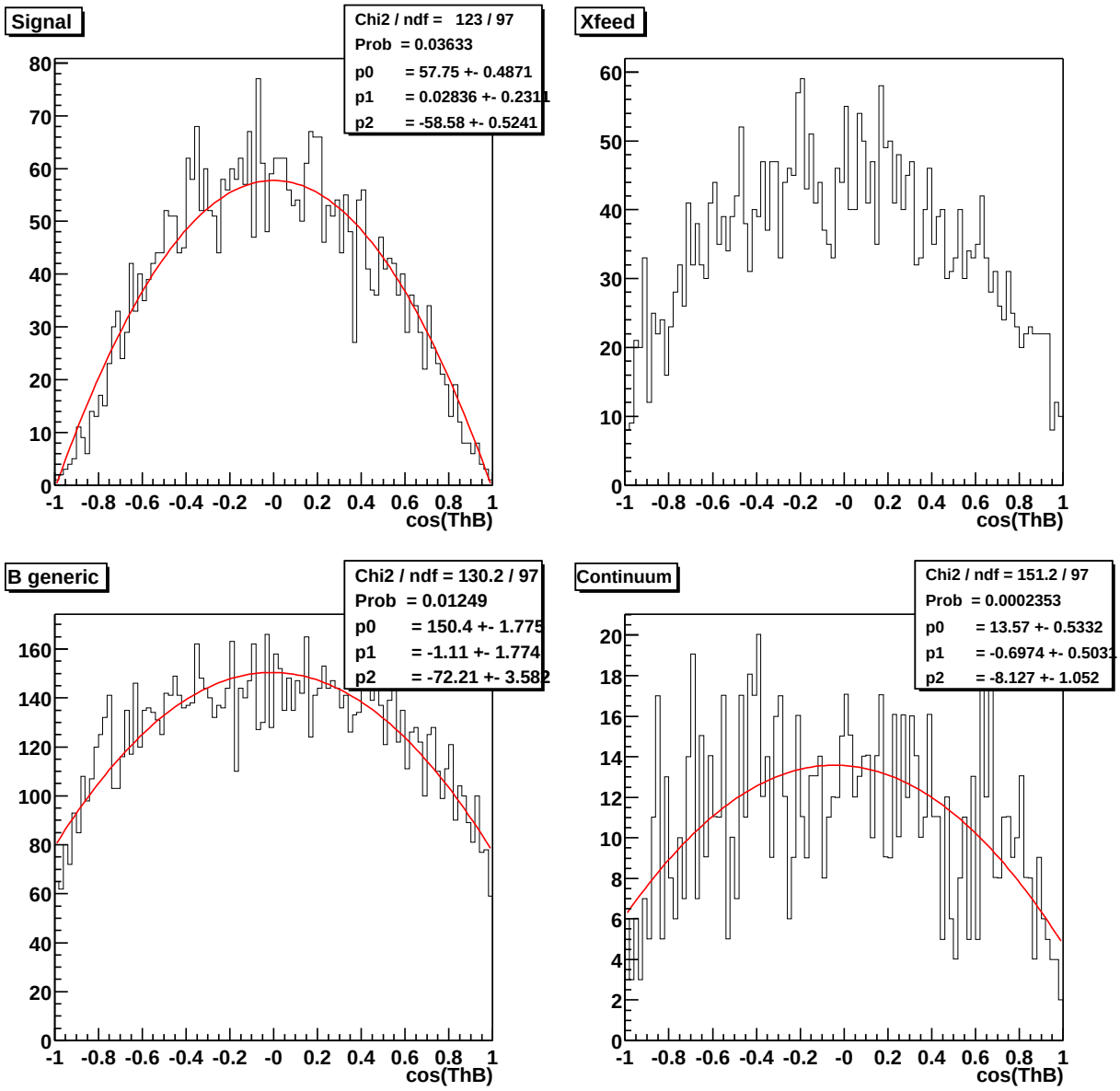

Figure 90: Distributions of $\cos\theta_B$ for $B \to X_s\, e^+e^-$ candidates in truth-matched signal (top left), cross-feed (top right), $B\overline{B}$ (bottom left) and continuum (bottom right) events. The curves show the result of the fits to these distributions using the functions listed in Table 5.

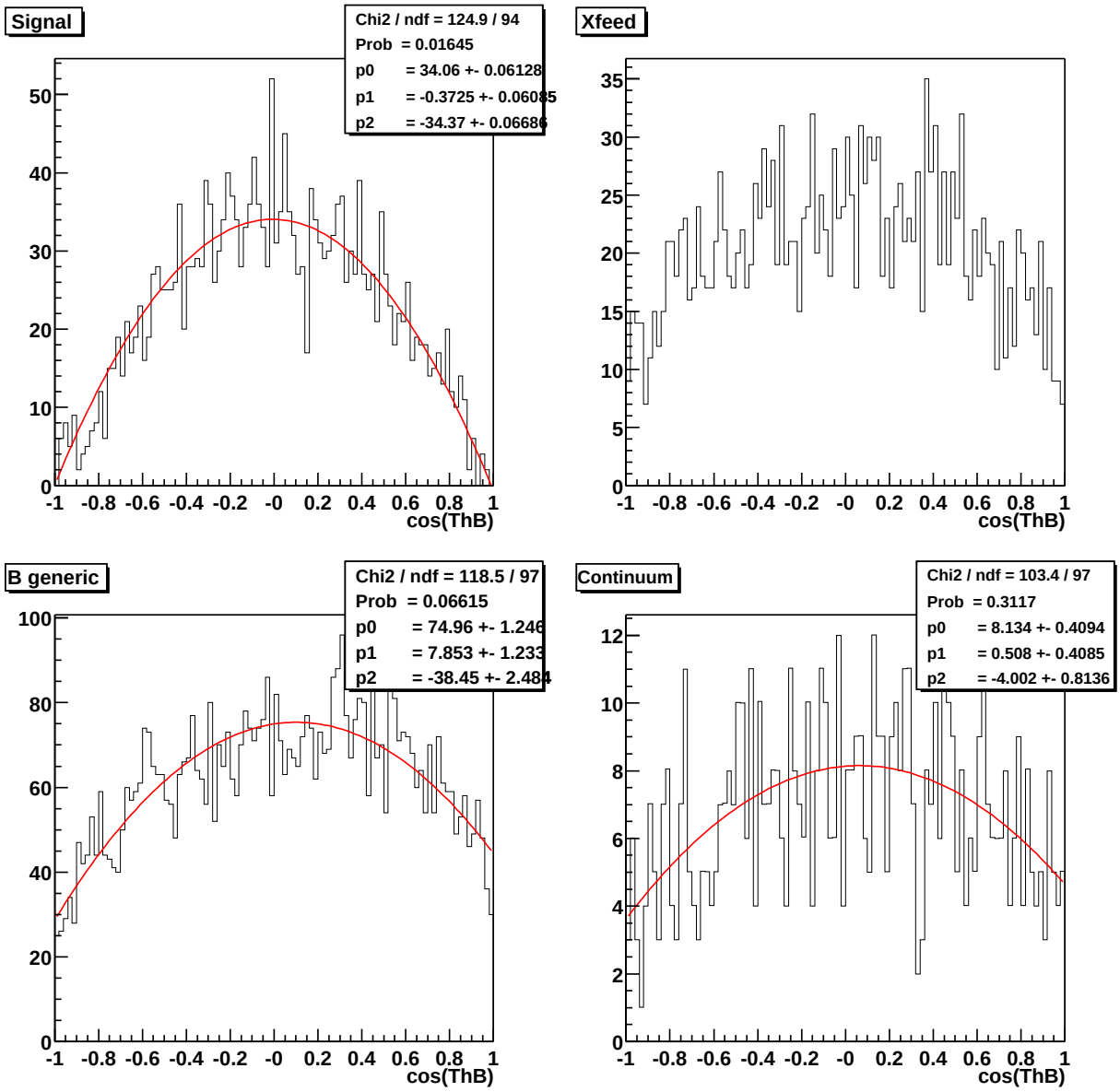

Figure 91: Distributions of $\cos\theta_B$ for $B \to X_s\,\mu^+\mu^-$ candidates in truth-matched signal (top left), cross-feed (top right), $B\overline{B}$ (bottom left) and continuum (bottom right) events. The curves show the result of the fits to these distributions using the functions listed in Table 5.

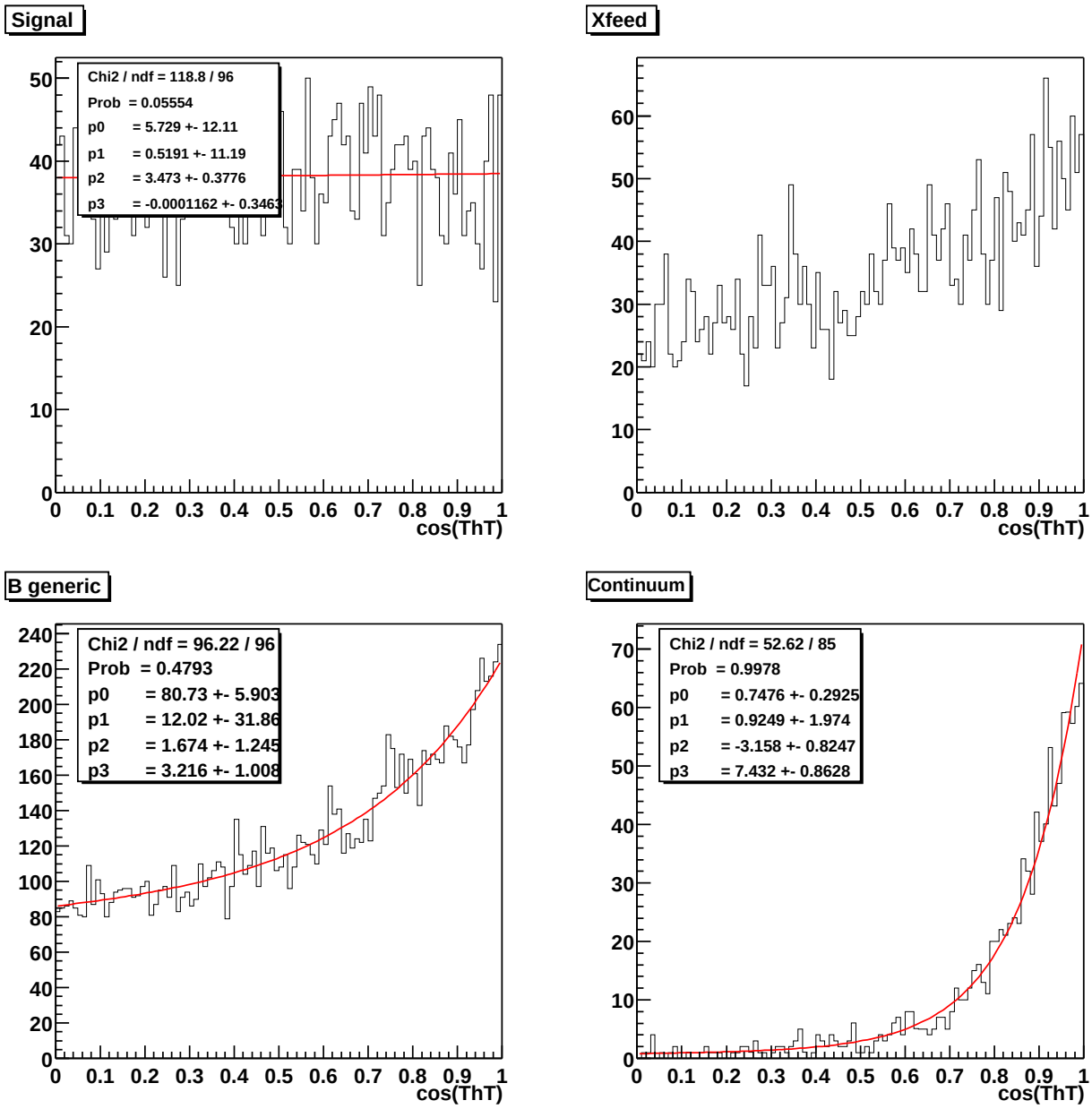

Figure 92: Distributions of $|\cos\theta_T|$ for $B \to X_s\, e^+e^-$ candidates in truth-matched signal (top left), cross-feed (top right), $B\overline{B}$ (bottom left) and continuum (bottom right) events. The curves show the result of the fits to these distributions using the functions listed in Table 5.

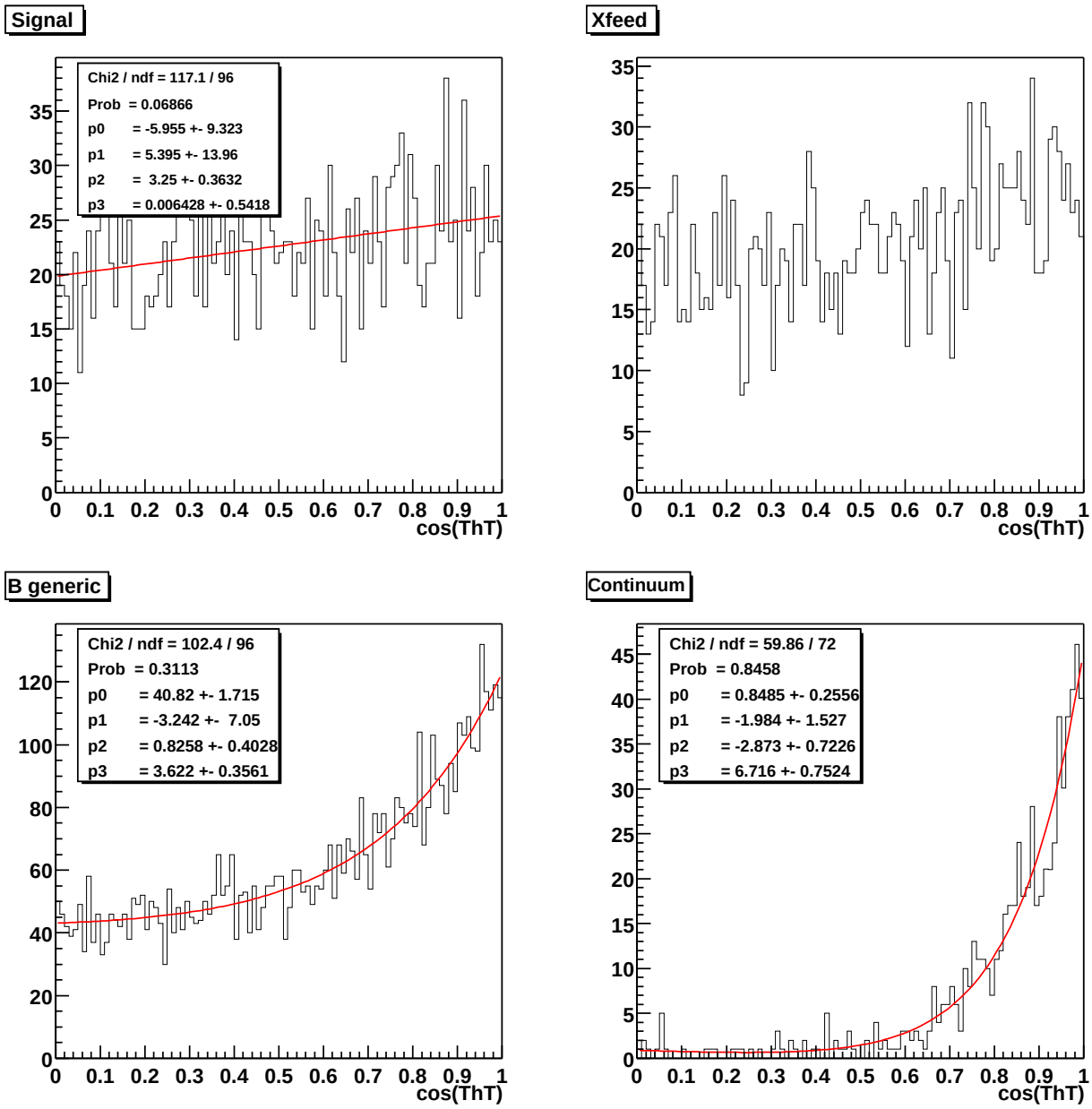

Figure 93: Distributions of $|\cos\theta_T|$ for $B \to X_s\,\mu^+\mu^-$ candidates in truth-matched signal (top left), cross-feed (top right), $B\overline{B}$ (bottom left) and continuum (bottom right) events. The curves show the result of the fits to these distributions using the functions listed in Table 5.

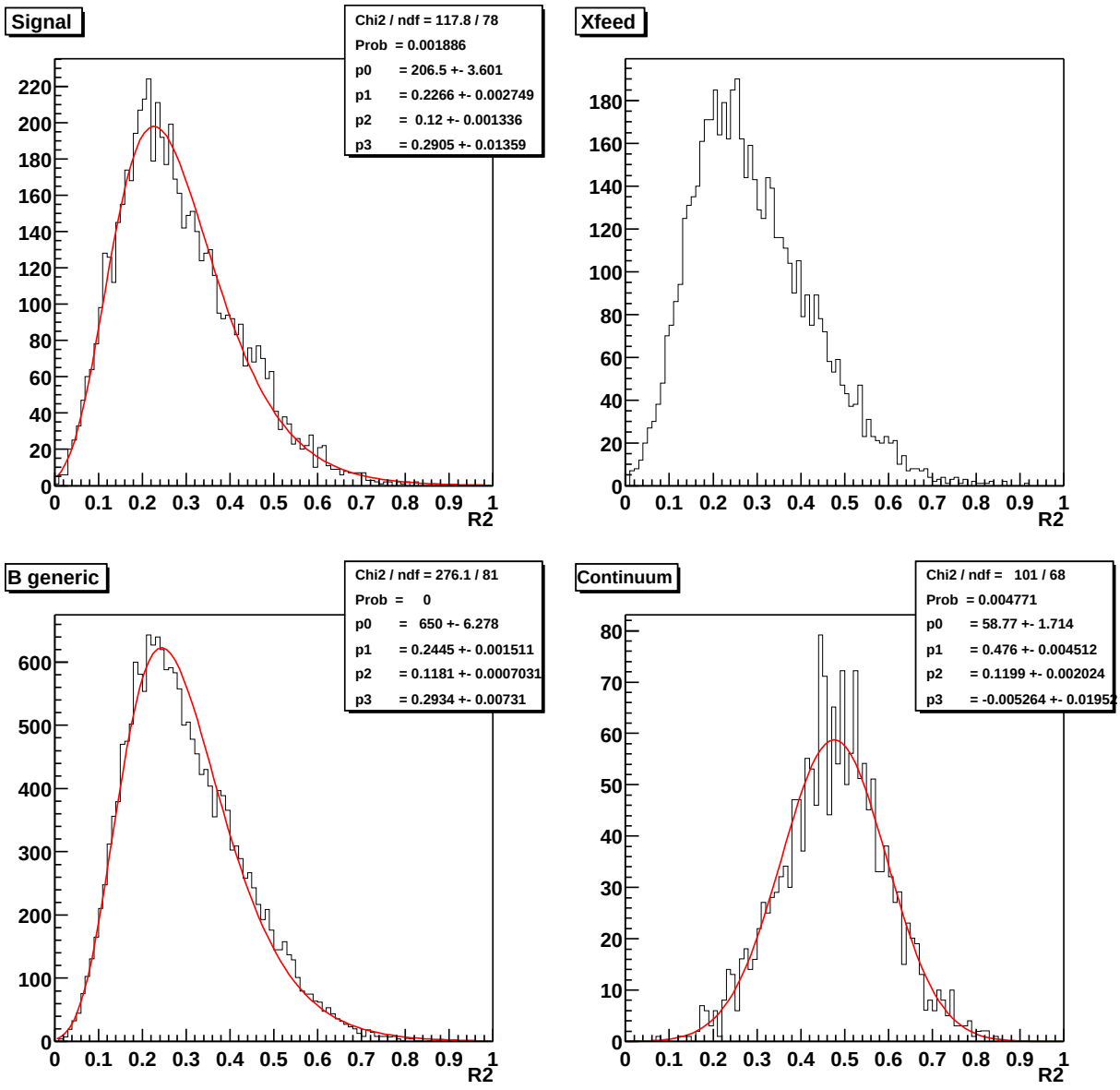

Figure 94: Distributions of R_2 for $B \to X_s\, \ell^+\ell^-$ candidates in truth-matched signal (top left), cross-feed (top right), $B\overline{B}$ (bottom left) and continuum (bottom right) events. The curves show the result of the fits to these distributions using the functions listed in Table 5.

APPENDIX C

Gamma conversion veto

The e^+e^- pairs originating from photon conversions in the detector material must be suppressed in the analysis. The γ conversion code makes use of the decay topology of γ conversions by finding e^+e^- pairs that have tangential tracks and small $m_{e^+e^-}$.

The module `GammaToEE_Default` uses m_e-assigned `ChargedTracks` to generate a list of γ conversion candidates, `gammaConversionDefault`, with cuts on the distance between the two $POCA$ P_1 and P_2 of the two tracks in the (x, y) plane, $|\delta_{xy}| < 0.5\ cm$, the distance along the z axis between P_1 and P_2, $|\delta_z| < 1.0\ cm$, and $m_{e^+e^-} < 30\ MeV$. Taking into account the location of the beam pipe, we also require that the radial distance in the (x, y) plane between the decay vertex of the γ conversion and the primary IP be $R_{xy} > 2.0\ cm$.

The γ conversion veto is applied after the best B candidate selection. In case at least one lepton in an e^+e^- pair of the B candidate is part of a γ conversion candidate's e^+e^- pair, the event is discarded. This veto removes about $x.x$ % of reconstructed, truth matched signal $B \to X_s e^+e^-$ candidates while decreasing the contamination in those decays by γ conversion candidates to a negligible level.

BIBLIOGRAPHY

[1] A. Ali, E. Lunghi, C. Greub, and G. Hiller. Improved model-independent analysis of semileptonic and radiative rare b decays. *Phys. Rev.*, D66:034002, 2002.

[2] K. Hagiwara et al. Review of particle physics. *Phys. Rev.*, D66:010001, 2002.

[3] Ahmed Ali, Patricia Ball, L. T. Handoko, and G. Hiller. A comparative study of the decays b $\rightarrow$ (k,k*) l+ l- in standard model and supersymmetric theories. *Phys. Rev.*, D61:074024, 2000.

[4] S. Weinberg. A model of leptons. *Phys. Rev. Lett.*, 19:1264–1266, 1967.

[5] A. Salam. *Elementary Particle Theory.* Almquist and Wiksell, Stockholm, 1968.

[6] Elliot Leader and Enrico Predazzi. *An Introduction to Gauge Theories and Modern Particle Physics*, volume 1. Cambridge University Press, 1996.

[7] N. Cabibbo. Unitary symmetry and leptonic decays. *Phys. Rev. Lett.*, 10:531–532, 1963.

[8] M. Kobayashi and T. Maskawa. Cp violation in the renormalizable theory of weak interaction. *Prog. Theor. Phys.*, 49:652–657, 1973.

[9] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson. Experimental test of parity conservation in beta decay. *Phys. Rev.*, 105:1413–1414, 1957.

[10] M. Goldhaber, L. Grodzins, and A. W. Sunyar. Helicity of neutrinos. *Phys. Rev.*, 109:1015–1017, 1958.

[11] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay. Evidence for the 2 pi decay of the k(2)0 meson. *Phys. Rev. Lett.*, 13:138–140, 1964.

[12] B. Aubert et al. Observation of cp violation in the b0 meson system. *Phys. Rev. Lett.*, 87:091801, 2001.

[13] K. Abe et al. Observation of large cp violation in the neutral b meson system. *Phys. Rev. Lett.*, 87:091802, 2001.

[14] A. D. Sakharov. Violation of cp invariance, c asymmetry, and baryon asymmetry of the universe. *Pisma Zh. Eksp. Teor. Fiz.*, 5:32–35, 1967.

[15] Andrzej J. Buras. Weak hamiltonian, cp violation and rare decays. 1998.

[16] Karen Lingel, Tomasz Skwarnicki, and James G. Smith. Penguin decays of b mesons. *Ann. Rev. Nucl. Part. Sci.*, 48:253, 1998.

[17] Tobias Hurth. Present status of inclusive rare b decays. *Rev. Mod. Phys.*, 75:1159–1199, 2003.

[18] Christoph Bobeth, Mikolaj Misiak, and Jorg Urban. Photonic penguins at two loops and m(t)-dependence of br(b $\to$ x(s) l+ l-). *Nucl. Phys.*, B574:291–330, 2000.

[19] Gudrun Hiller. b $\to$ s l+ l- decays in and beyond the standard model. *AIP Conf. Proc.*, 549:614–618, 2002.

[20] B. Aubert et al. The *BABAR* detector. *Nucl. Instrum. Meth.*, A479:1–116, 2002.

[21] U. Langenegger. *"Cut-based electron identification"*, bad #90 (2001).

[22] S. Chen et al. Branching fraction and photon energy spectrum for b –¿ s gamma. *Phys. Rev. Lett.*, 87:251807, 2001.

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

[24] H. Albrecht et al. Reconstruction of b mesons. *Phys. Lett.*, B185:218, 1987.

[25] A. Ali and G. Hiller. Eur. Phys. Jour. C **8**. page 619, 1999.

[26] A. Telnov and A. Gritsan. *"$K^0_S \to \pi^+\pi^-$ Reconstruction Efficiency in BABAR: Release-10 Data vs. SP4 Monte Carlo"*, bad #677 (2003).

[27] J. Berryhill, F. Porter, J. Richman, A. Ryd, and A. Samuel. *"Observation of the rare decays $B \to K\ell^+\ell^-$ and $B \to K^*\ell^+\ell^-$ "*, bad #559 (2003).

[28] S. Chen *et al.* Phys. Rev. Lett. **87**, 251807. 2001.

[29] D. Cronin-Hennessy *et al.* Phys. Rev. Lett. **87**, 251808. 2001.

[30] A. Ali and G. Hiller. Phys. Rev. D **60**, 034017. 1999.

[31] F. Di Lodovico, A. Eichenbaum, E. Grauges, S. Playfer, T. Pulliam, and J. Tinslay. *"Study of $b \to s\gamma$ using the sum of exclusive modes"*, bad #220 (2002).

[32] A. Ali. `hep-ph/0210183`, cern-th/2002-284, october 2002.

[33] A. Ghinculov, T. Hurth, G. Isidori, and Y.P. Yao. `hep-ph/0312128`, slac–pub–10265 (2003).