

# UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Fisica e Astronomia “Galileo Galilei”

Master Degree in Physics

Final Dissertation

Study of charmless  $B \rightarrow \eta' K$  decays  
with the Belle II experiment

Thesis supervisor

Dr. Stefano Lacaprara

Candidate

Valeria Fioroni

Academic Year 2020/2021



## Abstract

The study of penguin-dominated  $B^0 \rightarrow \eta' K_S^0$  decays provides a good test of the Standard Model, since the  $CP$  violation measurement in this mode is sensitive to the presence of new physics effects in the loop. In this work, a precise description of the signal extraction for both charged and neutral  $B \rightarrow \eta' K$  decays is presented, focusing on the suppression of the main background component: the continuum. A FastBDT classifier is trained to separate continuum and signal candidates, using variables related to the event topology. Furthermore, a study on multivariate Monte Carlo reweighting has been performed, in order to reduce the impact of the Monte Carlo mismodelling of the continuum suppression variables on the systematic uncertainty. From the measured signal yield, a branching fraction measurement has been performed, obtaining results consistent with world averages. Finally, a study on the signal  $\Delta t$  resolution function is presented.

Lo studio del decadimento  $B^0 \rightarrow \eta' K_S^0$ , dominato da diagrammi a pinguino, permette di testare le previsioni del Modello Standard, poiché la misura della violazione di  $CP$  in questa transizione è sensibile ad effetti di nuova fisica nel loop. In questo lavoro è presentata l'estrazione del segnale per i decadimenti  $B \rightarrow \eta' K$  carichi e neutri, concentrandosi sulla soppressione della principale componente di fondo: il fondo continuo. Per separare i candidati di segnale da quelli di continuo è stato effettuato il training di un classificatore FastBDT, usando variabili legate alla topologia dell'evento. Inoltre è stato effettuato uno studio per la riponderazione del Monte Carlo, attraverso tecniche di analisi multivariata, per ridurre l'impatto dell'erronea modellazione Monte Carlo delle variabili di soppressione del continuo sull'incertezza sistematica. Dalla misura del numero di candidati di segnale nel dataset è stata effettuata una misura della branching fraction, ottenendo risultati consistenti con le medie mondiali. Infine è stato presentato uno studio della funzione di risoluzione in  $\Delta t$  per il segnale.



# Contents

<b>1</b>	<b>Motivation and theoretical framework</b>	<b>3</b>
1.1	Introduction . . . . .	3
1.2	The Standard Model . . . . .	3
1.2.1	The CKM matrix . . . . .	5
1.2.2	Measurement of the Unitarity triangle angles . . . . .	6
1.2.3	Flavour oscillation and mixing . . . . .	7
1.3	Discrete symmetries $\mathcal{P}$ , $C$ and $\mathcal{T}$ . . . . .	9
1.4	$C\mathcal{P}$ violation . . . . .	9
1.4.1	$C\mathcal{P}$ violation in the $B^0 \rightarrow \eta' K_S^0$ decay . . . . .	9
<b>2</b>	<b>SuperKEKB and Belle II</b>	<b>13</b>
2.1	The SuperKEKB collider . . . . .	13
2.2	Physics program of Belle II at SuperKEKB . . . . .	14
2.3	The Belle II detector . . . . .	14
2.3.1	Vertex detector (VXD) . . . . .	15
2.3.2	Central Drift Chamber (CDC) . . . . .	15
2.3.3	Particle Identification (PID) . . . . .	16
2.3.4	Electromagnetic Calorimeter (ECL) . . . . .	16
2.3.5	$K_L$ and Muon detector (KLM) . . . . .	17
2.3.6	Trigger . . . . .	17
<b>3</b>	<b>Introduction to the analysis of the <math>B \rightarrow \eta' K</math> decay</b>	<b>19</b>
3.1	Signal and decay modes . . . . .	19
3.2	Background sources . . . . .	19
3.3	Analysis strategy . . . . .	21
3.4	Data and Monte Carlo samples . . . . .	21
3.5	Reconstruction . . . . .	22
3.6	Signal selection . . . . .	23
<b>4</b>	<b>Continuum Suppression</b>	<b>29</b>
4.1	Variables . . . . .	29
4.1.1	Thrust . . . . .	29
4.1.2	Fox-Wolfram moments . . . . .	29
4.1.3	CLEO cones . . . . .	31
4.2	FastBDT . . . . .	31
4.3	Training, validation and test . . . . .	33
4.4	Validation with the off-resonance data . . . . .	35
4.5	Multivariate Monte Carlo reweighting . . . . .	38
<b>5</b>	<b>Branching Fractions</b>	<b>41</b>
5.1	Fit procedure for signal extraction . . . . .	41
5.1.1	Fitting tools . . . . .	44

5.2	Fit validation and results . . . . .	45
5.2.1	Extraction of the PDFs from Monte Carlo . . . . .	45
5.2.2	Test of the fit procedure . . . . .	47
5.2.3	Fit results . . . . .	49
5.3	Branching fractions measurement . . . . .	50
5.3.1	Systematic uncertainties . . . . .	55
<b>6</b>	<b>Time-dependent CP Violation</b>	<b>57</b>
6.1	Description of the measurement . . . . .	57
6.2	Previous results . . . . .	58
6.2.1	BaBar . . . . .	58
6.2.2	Belle . . . . .	58
6.3	Vertices reconstruction . . . . .	59
6.4	Flavour tagger . . . . .	60
6.5	Study of the signal resolution function . . . . .	60
6.5.1	Fit of the resolution function . . . . .	62
6.5.2	Lifetime fit on Monte Carlo . . . . .	64
<b>7</b>	<b>Conclusion</b>	<b>67</b>
	<b>Appendices</b>	<b>73</b>
<b>A</b>	<b>Fit for signal extraction with zfit</b>	<b>75</b>
A.0.1	Extraction of the PDFs from Monte Carlo . . . . .	75
A.0.2	Test of the fit procedure . . . . .	75
<b>B</b>	<b>RooFit fit test with toy Monte Carlo</b>	<b>85</b>





# Chapter 1

## Motivation and theoretical framework

### 1.1 Introduction

The Standard Model provides a precise description of elementary particles and their interactions, but there are many fundamental questions still unsolved. Among them there is the matter-antimatter asymmetry observed in the universe.  $CP$  violation, discussed in section 1.4, is a necessary condition for the evolution to a matter-dominated universe [1], however, the  $CP$  violation predicted by the Standard Model is several orders of magnitude too small to explain the observed asymmetry. Furthermore, the flavour mixing Cabibbo-Kobayashi-Maskawa matrix (section 1.2.1) is roughly diagonal and the origin of this hierarchy is unknown, since its elements are free parameters of the theory. This may suggest the presence of some new flavour symmetry at higher energy scales. Many new physics theories, including new particles and processes, have been proposed to explain the effects not described by the Standard Model.

Experiments at the Large Hadron Collider (LHC) [2], [3], allow to directly search for new physics effects at the TeV scale. The Belle II experiment, introduced in chapter 2, uses a different but complementary approach. It aims to search for new physics at the intensity frontier, measuring many processes with unprecedented precision and allowing to appreciate any eventual discrepancy with the Standard Model predictions, due to new particles and processes occurring at mass scales higher than the ones reached by the LHC.

Precision measurements of  $CP$  violation in the  $B$  sector are performed in the decay channel  $B^0 \rightarrow J/\psi K^0$ , which has a relatively large branching fraction since it is mediated by a tree diagram. The comparison with a penguin-mediated decay, such as  $B^0 \rightarrow \eta' K^0$ , is potentially sensitive to new physics occurring in the loop, as described in section 1.4.1. The presence of many neutral particles in the final states of these decays makes their reconstruction at hadron colliders extremely challenging. Furthermore, the current world average for  $CP$  violation in this mode is dominated by the statistical uncertainty. For these reasons, the measurement of this  $CP$  violating process at Belle II, given the clean environment at  $e^+e^-$  colliders and the large amount of data that this experiment aims to collect, will be crucial to assess the presence of new physics in the  $B^0 \rightarrow \eta' K^0$  process. The study of such decays is the goal of this work. Specifically, the signal extraction for the  $B \rightarrow \eta' K$  decays will be discussed, providing a measurement of the branching fractions for these processes. Finally, in view of the future measurement of  $CP$  violation in these transitions, a study on the signal time resolution function  $\mathcal{R}(\Delta t)$  will be presented.

### 1.2 The Standard Model

The Standard Model is the theory describing the interaction of particles through the strong, weak and electromagnetic fundamental forces.

The particle content of the theory is summarized in Fig. 1.1. There are 12 fermions, 6 leptons and 6 quarks, categorized in 3 generations. The force carriers are the spin-1 gauge bosons, namely the photon, gluons and the  $W^\pm$  and  $Z^0$  bosons. Finally there is the Higgs boson, the particle associated to the Higgs field.

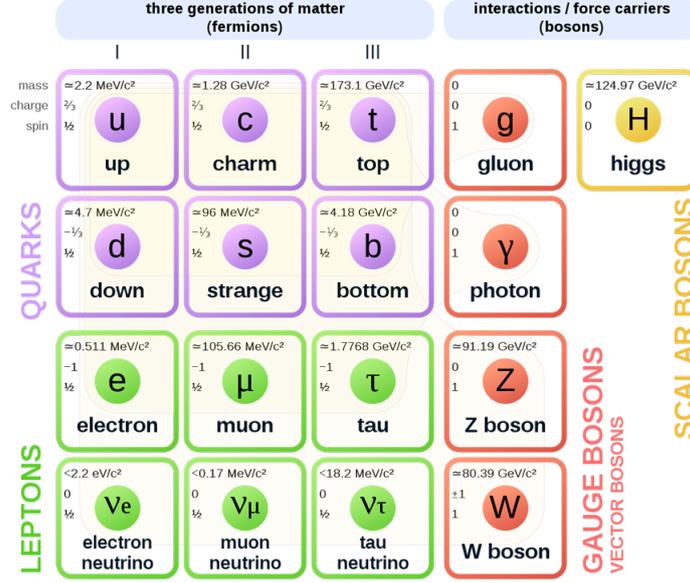


Figure 1.1: Elementary particle content of the Standard Model

The Standard Model lagrangian can be written as the sum of two components:

$$\mathcal{L}_{SM} = \mathcal{L}_{QCD} + \mathcal{L}_{EW} \quad (1.1)$$

where  $\mathcal{L}_{QCD}$  is the QCD lagrangian, describing strong interactions and invariant under local  $SU(3)_C$  transformations, and  $\mathcal{L}_{EW}$  is the electro-weak lagrangian, which includes electromagnetic and weak interactions (both neutral and charged current), with symmetry group  $SU(2)_L \otimes U(1)_Y$ . An important ingredient of the theory is the possibility of spontaneous symmetry breaking of the gauge group into a smaller group: this spontaneous symmetry breaking plays a key role in the Higgs mechanism. Gauge bosons and fermions interact with the Higgs field, acquiring their mass. The choice of the ground state breaks the Standard Model symmetry group

$$G_{SM} = SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$$

into

$$G_{SM}^{SSB} = SU(3)_C \otimes U(1)_{em}$$

resulting in three massive bosons,  $W^\pm$  and  $Z$ , 8 massless gluons and a massless photon.

The interaction of fermions with the Higgs field is described by the Yukawa term of the Standard Model lagrangian, that, considering only the quark terms, reads:

$$\mathcal{L}_Y = -\bar{Q}_L^i Y_{ij}^d \Phi d_R^j - \bar{Q}_L^i Y_{ij}^u \tilde{\Phi} u_R^j + \text{h.c.}$$

where  $\bar{Q}_L = \begin{pmatrix} u' \\ d' \end{pmatrix}_L$  is the left-handed  $SU(2)_L$  doublet,  $u'_R, d'_R$  are right-handed  $SU(2)_L$  singlets,  $\Phi$  is the Higgs field,  $\tilde{\Phi} = i\sigma_2 \Phi$ , the  $Y^{u,d}$  are the complex  $3 \times 3$  Yukawa matrices and  $i$  and  $j$  are flavour indices. After symmetry breaking, it is possible to diagonalize the  $Y$  matrices by means of a biunitary transformation  $\hat{Y} = LYR^\dagger$ . Diagonalizing the  $Y^{u,d}$  and rotating the fields  $u'_L = L_u u_L$ ,  $d'_L = L_d d_L$ ,  $u'_R = R_u u_R$  and  $d'_R = R_d d_R$ , we pass to the fermion mass basis, with fermion masses  $m_f^{ij} = \frac{v}{\sqrt{2}} \hat{Y}_f^{ij}$

(where  $v = 246$  GeV is the vacuum expectation value of the Higgs field). This change of basis does not affect the other fermionic sectors, with the exception of the weak charged current interaction sector: here the interaction is among up-type and down-type fermions and the unitarity property of the transformation cannot be exploited. The charged current interaction lagrangian in the mass basis reads

$$\mathcal{L}_{CC}^{mass} = -\frac{g}{\sqrt{2}} [\bar{u}_L^i (L_u^{i\dagger} L_d^j) W^+ d_R^j + \text{h.c.}] = -\frac{g}{\sqrt{2}} [\bar{u}_L^i V^{ij} W^+ d_R^j + \text{h.c.}]$$

where the flavour mixing Cabibbo-Kobayashi-Maskawa matrix  $V \equiv L_u^\dagger L_d$  has been introduced.

### 1.2.1 The CKM matrix

The Cabibbo-Kobayashi-Maskawa (CKM) matrix is a unitary matrix with flavour mixing properties that gives the relation between the weak eigenstates  $d', s', b'$  and the mass eigenstates  $d, s, b$ :

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \quad (1.2)$$

and appears in the couplings of the  $W$  boson to quarks in the mass basis. For a theory with three generations of quarks, the  $V$  can be parametrized by three mixing angles and one irreducible phase, that is responsible for  $CP$  violation in the Standard Model 1.4. The *standard parametrization* of the CKM matrix is obtained as the product of three rotation matrices, one of which contains the phase  $\delta$ :

$$V = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (1.3)$$

$$= \begin{pmatrix} c_{13}c_{12} & c_{13}s_{12} & s_{13}e^{i\delta} \\ -c_{23}s_{12} - s_{23}c_{12}s_{13}e^{i\delta} & c_{23}c_{12} - s_{23}s_{13}s_{12}e^{i\delta} & c_{13}s_{23} \\ s_{23}s_{12} - c_{23}c_{12}s_{13}e^{i\delta} & -s_{23}c_{12} - c_{23}s_{12}s_{13}e^{i\delta} & c_{13}c_{23} \end{pmatrix} \quad (1.4)$$

where  $\theta_{ij}$  are the mixing angles between the  $i$ th and  $j$ th generation and  $c_{ij} = \cos\theta_{ij}$ ,  $s_{ij} = \sin\theta_{ij}$ . The rotation angle between the first two generations,  $\theta_{12}$ , is known as the *Cabibbo angle* [4]. In the case of two generations, the  $2 \times 2$  CKM complex matrix has 4 parameters: one mixing angle and three phases. The  $U(1)$  chiral symmetry allows to remove the three phases, leaving only one parameter, the Cabibbo angle. Thus, the 2-generations CKM matrix can be taken real and no  $CP$  violation can take place.

From experimental observations  $|V_{ub}|^2 \ll |V_{cb}|^2 \ll |V_{us}|^2$ , hence an expansion was introduced. By defining

$$s_{12} \equiv \lambda, \quad s_{23} \equiv A\lambda^2, \quad s_{13}e^{i\delta} \equiv A\lambda^3(\rho - i\eta)$$

it is possible to write the *Wolfenstein parameterization* [5] of the CKM matrix:

$$V = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4)$$

From the unitarity condition  $VV^\dagger = V^\dagger V = I$  it is possible to obtain nine independent equations, one of which is particularly relevant for the study of  $B$  physics:

$$V_{ud}^* V_{ub} + V_{cd}^* V_{cb} + V_{td}^* V_{tb} = 0$$

Since the terms are of the same order in  $\lambda$  (i.e.  $\lambda^3$ ), this equation defines a non-degenerate triangle in the  $(\bar{\rho}, \bar{\eta})$  plane, the so called *unitarity triangle* shown in fig. 1.2, where the relation between  $\rho$ ,  $\eta$  and  $\bar{\rho}$ ,  $\bar{\eta}$  is

$$\rho + i\eta = \frac{\sqrt{1 - A^2\lambda^4}(\bar{\rho} + i\bar{\eta})}{\sqrt{1 - \lambda^2}[1 - A^2\lambda^4(\bar{\rho} + i\bar{\eta})]}$$

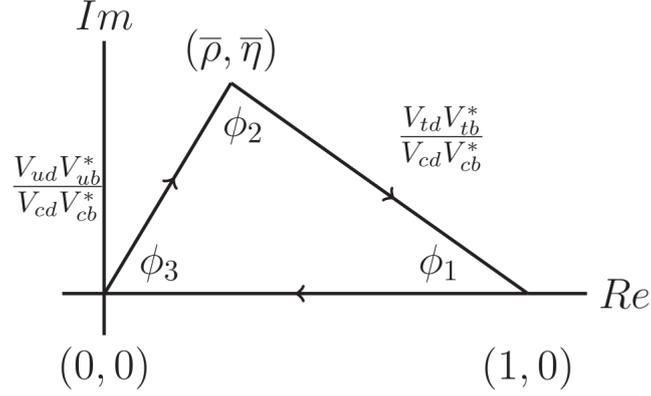


Figure 1.2: Unitarity triangle

The three angles are defined as

$$\phi_1 = \arg \left[ -\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*} \right], \quad \phi_2 = \arg \left[ -\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*} \right], \quad \phi_3 = \arg \left[ -\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} \right]$$

and are also known as  $\phi_1 = \beta$ ,  $\phi_2 = \alpha$  and  $\phi_3 = \gamma$ .

### 1.2.2 Measurement of the Unitarity triangle angles

The measurement of the unitarity triangle is an important test of the Standard Model and of the CKM mechanism. The  $V$  matrix is unitary by construction but if, for example, there were a fourth generation, the three-generations subsector of the CKM matrix would not be unitary. The study of  $B$  physics plays a major role in the determination of the angles  $\phi_1$ ,  $\phi_2$  and  $\phi_3$  and of the sides of the unitarity triangle. This is one of the main goals of experiments at the  $B$  factories. Through a set of independent measurements from several experiments it was possible to prove that the unitarity triangle is in fact a closed triangle, in agreement with the Standard Model prediction, within the experimental uncertainties.

The angle  $\phi_1$  has been precisely measured from  $CP$  violation in  $b \rightarrow c\bar{c}s$  decays to  $CP$  eigenstates, for example  $B^0 \rightarrow J/\psi K_S^0$ . Also charmless penguin-mediated  $b \rightarrow s$  transitions allow to measure  $\phi_1$  and their relevance is described in section 1.4.1.

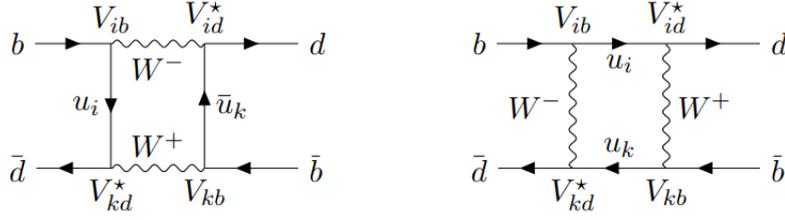
The measure of  $\phi_2$  from  $b \rightarrow d$  decays is more complicated since there is a sizeable contribution from penguin decays carrying a different phase with respect to the tree amplitudes  $b \rightarrow u\bar{u}d$ . The decay modes used to determine  $\phi_2$  are  $B \rightarrow \pi\pi$ ,  $\rho\rho$  and  $\pi\rho$ .

The angle  $\phi_3$  can be measured from tree level processes, since it does not depend on CKM elements involving the top quark, thus it is unlikely to be affected from physics beyond the Standard Model. It can be measured from the interference between  $b \rightarrow c$  and  $b \rightarrow u$  transitions, for example in  $B^- \rightarrow D^0 K^- / B^- \rightarrow \bar{D}^0 K^-$  and  $\bar{B}^0 \rightarrow D^\pm \pi^\mp / \bar{B}^0 \rightarrow B^0 \rightarrow D^\pm \pi^\mp$  decays. A limit to these analyses is the large difference in magnitude of the two interfering amplitudes, since the sensitivity depends on their ratio. For the measurement of  $\phi_3$  Belle II will be competitive but the LHCb [6] experiment is expected to dominate with an integrated luminosity of  $300 \text{ fb}^{-1}$  [7], foreseen to be collected by 2030, given the high cross section for  $B$  production at LHC.

In fig. 1.3, the current results for the unitarity triangle parameters are shown.

The Belle II experiment, described in chapter 2, aims to collect a  $B\bar{B}$  sample corresponding to  $50 \text{ ab}^{-1}$  by 2031, allowing to precisely measure the unitarity triangle. In fig. 1.4 the expected precision on the unitarity triangle parameters with  $50 \text{ ab}^{-1}$  is shown.



Figure 1.5: Box diagrams for  $B^0$  mesons oscillation

As a consequence,  $M$  and  $\Gamma$  diagonal elements are real and  $M_{21} = M_{12}^*$ ,  $\Gamma_{21} = \Gamma_{12}^*$ . Furthermore,  $CPT$  invariance (section 1.3) requires  $M_{11} = M_{22}$  and  $\Gamma_{11} = \Gamma_{22}$ . Thus, the  $B^0 - \bar{B}^0$  mixing can be parametrized by five real parameters:  $M_{11}$ ,  $\Gamma_{11}$ ,  $|M_{12}|$ ,  $|\Gamma_{12}|$  and the phase difference between  $M_{12}$  and  $\Gamma_{12}$ ,  $\phi = \arg(-\frac{M_{12}}{\Gamma_{12}})$ . At time  $t$ , a generic state of the meson-antimeson system can be decomposed in the flavour eigenstate:

$$\psi(t) = a(t) |B^0\rangle + b(t) |\bar{B}^0\rangle$$

and its time evolution is described by the Schrödinger equation

$$i \frac{\partial \psi}{\partial t} = H \psi$$

Diagonalizing the hamiltonian, the mass eigenstates  $|B_L\rangle$  and  $|B_H\rangle$  can be written as a combination of the flavour eigenstates  $|B^0\rangle$  and  $|\bar{B}^0\rangle$  as follows:

$$|B_L\rangle = p |B^0\rangle + q |\bar{B}^0\rangle$$

$$|B_H\rangle = p |B^0\rangle - q |\bar{B}^0\rangle$$

where  $p$  and  $q$  are complex numbers, with  $\frac{q}{p} = \sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}}$ , and  $|p|^2 + |q|^2 = 1$ . The eigenvalues corresponding to the heavy and light states are

$$\lambda_{H,L} = m_{H,L} - \frac{i}{2}\Gamma_{H,L}$$

The time evolution of the mass eigenstates is

$$|B_{H,L}(t)\rangle = e^{i\lambda_{H,L}t} |B_{H,L}(0)\rangle$$

and the time evolution of a  $B$  produced in a  $B^0$  or  $\bar{B}^0$  flavour eigenstate at time  $t = 0$  is

$$|B^0(t)\rangle = g_+(t) |B^0\rangle + \frac{q}{p} g_-(t) |\bar{B}^0\rangle$$

$$|\bar{B}^0(t)\rangle = g_-(t) |B^0\rangle + \frac{p}{q} g_+(t) |\bar{B}^0\rangle$$

where

$$g_{\pm}(t) = \frac{1}{2} (e^{-im_H t - \frac{1}{2}\Gamma_H t} \pm e^{-im_L t - \frac{1}{2}\Gamma_L t})$$

An initial  $B^0$  (or  $\bar{B}^0$ ) evolves in time oscillating between the two flavours with frequency  $\Delta m = m_H - m_L$ . For the  $B$  meson system the number of oscillations in a lifetime is of the order of unity, allowing to experimentally observe this phenomenon.

One of the consequences of  $B$  mesons mixing is  $CP$  violation. A precise description of this phenomenon can be found in the next section.

### 1.3 Discrete symmetries $\mathcal{P}$ , $C$ and $\mathcal{T}$

Symmetries are a very important concept in physics since the invariance of physics equations under a certain transformation is connected to conservation laws by the Noether theorem.

The process studied in this work allows to evaluate the violation of two discrete symmetries: the parity  $\mathcal{P}$  and the charge conjugation  $C$ . A parity transformation  $\mathcal{P}$  consists in a simultaneous change in the sign of the spatial coordinates  $\vec{x} \rightarrow -\vec{x}$ . The charge conjugation operation  $C$  turns a particle into its antiparticle by changing the sign of all its charges. Another important discrete symmetry is the time reversal  $\mathcal{T} : t \rightarrow -t$ .

### 1.4 $C\mathcal{P}$ violation

$\mathcal{P}$  and  $C$  are good symmetries of the strong and electromagnetic interactions, while the weak interactions violate  $\mathcal{P}$ ,  $C$  and  $C\mathcal{P}$ . The parity violation was firstly observed in 1957 in the  $\beta$  decay of  $^{60}\text{Co}$  [10].  $C\mathcal{P}$  violation was then observed in the neutral kaon system, with the detection of the  $C\mathcal{P}$  violating  $K_L \rightarrow \pi^+\pi^-$  decay [11]. The Standard Model of elementary particles accounts for  $C\mathcal{P}$  violation in the weak sector through the presence of a complex phase in the three-generations flavour mixing CKM matrix, introduced in 1973 by Kobayashi and Maskawa [12], expanding the 2-generation  $C\mathcal{P}$  conserving Cabibbo rotation matrix, discussed in section 1.2.1. In 2001, the observation of  $C\mathcal{P}$  violation in the  $B$  meson system by Belle and Babar experiments, provided a clear proof of the Kobayashi and Maskawa scheme for  $C\mathcal{P}$  violation in the Standard Model. The precise measurement of  $C\mathcal{P}$  violating processes is extremely important and might be sensitive to contribution from new physics.

In the Standard Model there are three types of  $C\mathcal{P}$  violation:

- **Violation in the wave function**, also called **violation in mixing**

It happens when the wave functions of the free Hamiltonian are not  $C\mathcal{P}$  eigenstates. It is a small effect that has been observed in the neutral kaon system.

- **Violation in decays** (or **direct  $C\mathcal{P}$  violation**)

Let  $M$  be a meson decaying into the final state  $f$ . If  $C\mathcal{P}$  is conserved the amplitude  $A(M \rightarrow f)$  should be equal to the amplitude  $A(\bar{M} \rightarrow \bar{f})$ , where  $\bar{M}$  is the antimeson and  $\bar{f}$  the conjugate state of  $f$ . This type of  $C\mathcal{P}$  violation can be observed in both neutral and charged mesons, for example in  $B^0 \rightarrow K^+\pi^-$  decays, where direct  $C\mathcal{P}$  violation has been observed with a significance of  $5\sigma$  [8].

- **Violation in the interference between decays with and without oscillation**

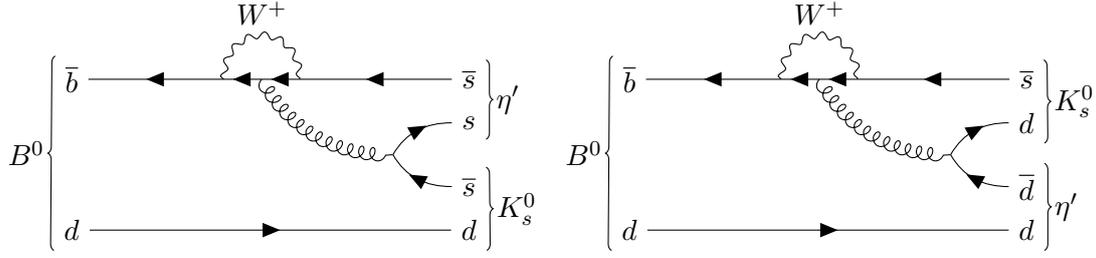
This type of  $C\mathcal{P}$  violation is due to the interference of the decay without mixing  $M \rightarrow f$  and the decay with mixing  $M \rightarrow \bar{M} \rightarrow f$ . This effect can be observed in the decay to final states common to both  $M$  and  $\bar{M}$ . It was observed for the first time in 2001 by Belle and BaBar experiments [13], [14] in the decay of the  $B^0/\bar{B}^0$  to the  $C\mathcal{P}$  eigenstate  $J/\psi K^0$ . This type of violation can also be measured in loop-mediated  $b \rightarrow s$  transitions, such as the  $B \rightarrow \eta' K$  decay, as discussed in the following section.

#### 1.4.1 $C\mathcal{P}$ violation in the $B^0 \rightarrow \eta' K_S^0$ decay

The  $B^0 \rightarrow \eta' K_S^0$  decay proceeds through a  $b \rightarrow s$  penguin loop transition (fig. 1.6), since the tree level  $b \rightarrow u$  process is both color and CKM suppressed.

$C\mathcal{P}$  violation in this mode is of the third type presented in section 1.4. In particular, the time-dependent  $C\mathcal{P}$  asymmetry reads

$$A_{CP}(t) \equiv \frac{\Gamma(B^0 \rightarrow \eta' K_S^0)(t) - \Gamma(\bar{B}^0 \rightarrow \eta' K_S^0)(t)}{\Gamma(B^0 \rightarrow \eta' K_S^0)(t) + \Gamma(\bar{B}^0 \rightarrow \eta' K_S^0)(t)} = \mathcal{S}_{\eta' K_S^0} \sin(\Delta m t) + \mathcal{A}_{\eta' K_S^0} \cos(\Delta m t)$$

Figure 1.6: Feynman diagrams contributing to  $B^0 \rightarrow \eta' K_S^0$  process

where  $\Delta m = m_H - m_L$  is the mixing frequency and the  $\mathcal{A}_{\eta' K_S^0}$  and  $\mathcal{S}_{\eta' K_S^0}$  parameters are connected to  $CP$  violation. The Standard Model prediction for  $\mathcal{A}_{\eta' K_S^0}$  is 0, since this parameter is related to direct  $CP$  violation, while  $\mathcal{S}_{\eta' K_S^0} \sim \sin 2\phi_1$ . Hence, the measurement of  $A_{CP}$  in  $B^0 \rightarrow \eta' K_S^0$  decays allows to measure the unitarity triangle angle  $\phi_1$ .

The value of  $\sin 2\phi_1$  has been precisely measured in the ‘‘golden channel’’  $B^0 \rightarrow J/\psi K_S^0$ , a tree level decay with a suppressed contribution from penguin processes. The result obtained by Belle and BaBar experiments for  $\mathcal{S}_{J/\psi K_S^0}$  are given in table 1.1.

	$\mathcal{S}_{J/\psi K_S^0}$
Belle	$0.670 \pm 0.029 \pm 0.13$
BaBar	$0.687 \pm 0.028 \pm 0.012$

Table 1.1: Belle and BaBar results for the  $CP$  violating parameter  $\mathcal{S}_{J/\psi K_S^0}$  [8]

The Standard Model predictions for  $\mathcal{A}_{J/\psi K_S^0}$  and  $\mathcal{S}_{J/\psi K_S^0}$  are almost the same as in the  $B^0 \rightarrow \eta' K_S^0$  process, because of the suppression of the tree contribution to  $B^0 \rightarrow \eta' K_S^0$  and of the loop contribution to  $B^0 \rightarrow J/\psi K_S^0$ . Hence, any difference between  $\mathcal{S}_{J/\psi K_S^0}$  and  $\mathcal{S}_{\eta' K_S^0}$  would be a sign of physics beyond the Standard Model. In presence of new physics, in fact, new heavy particles in the penguin loop could contribute to the decay amplitude, potentially carrying new  $CP$  violating phases. This effect would be detectable in the  $B^0 \rightarrow \eta' K_S^0$  loop-dominated process, while  $\mathcal{S}_{J/\psi K_S^0}$ , from the tree-dominated  $B^0 \rightarrow J/\psi K_S^0$  decay, would not be significantly modified.

The current world averages for  $\mathcal{A}_{\eta' K^0}$  and  $\mathcal{S}_{\eta' K^0}$ <sup>1</sup> [8] are:

$$\mathcal{A}_{\eta' K^0} = -0.06 \pm 0.04$$

$$\mathcal{S}_{\eta' K^0} = 0.63 \pm 0.06$$

This results, based on the the measurements performed at Babar and Belle experiments [15], [16], described in section 6.2 and given in table 1.2, are still dominated by the statistical uncertainty. The Belle II detector, described in section 2.3, aims to collect a larger data sample, up to  $50 \text{ ab}^{-1}$ , allowing to perform a precise  $CP$  violation measurement in this kind of transitions.

	$\mathcal{A}_{\eta' K^0}$	$\mathcal{S}_{\eta' K^0}$
Belle	$0.03 \pm 0.05 \pm 0.04$	$0.68 \pm 0.07 \pm 0.03$
BaBar	$-0.08 \pm 0.06 \pm 0.02$	$0.57 \pm 0.08 \pm 0.02$

Table 1.2: Belle and BaBar results for the  $CP$  violating parameters in  $B^0 \rightarrow \eta' K^0$  decays

The expected sensitivity on the  $CP$  violation parameter  $\mathcal{S}$  at  $50 \text{ ab}^{-1}$  for both  $B^0 \rightarrow \eta' K^0$  and  $B^0 \rightarrow J/\psi K_S^0$  (and for the combined  $b \rightarrow c\bar{c}s$  transitions measurement) can be found in [9] and is given in tables 1.3 and 1.4.

<sup>1</sup>The world averages are given for the  $B \rightarrow \eta' K^0$  process, which combines the results obtained for  $B^0 \rightarrow \eta' K^0$  and  $B^0 \rightarrow \eta' K_L^0$

	$B^0 \rightarrow J/\psi K_S^0$	all $b \rightarrow c\bar{c}s$
statistical	0.0035	0.0027
systematic (reducible)	0.0012	0.0026
systematic (irreducible)	0.0044	0.0036

Table 1.3: Expected sensitivity on the  $CP$  violation parameters  $\mathcal{S}_{J/\psi K_S^0}$  and  $\mathcal{S}_{c\bar{c}s}$  at  $50 \text{ ab}^{-1}$

	$B^0 \rightarrow \eta' K^0$	
	$5 \text{ ab}^{-1}$	$50 \text{ ab}^{-1}$
statistical	0.027	0.0085
systematic	0.020	0.013

Table 1.4: Expected sensitivity on  $\mathcal{S}_{\eta' K^0}$  at  $5$  and  $50 \text{ ab}^{-1}$



## Chapter 2

# SuperKEKB and Belle II

### 2.1 The SuperKEKB collider

SuperKEKB is an asymmetric-energy electron-positron collider, obtained by upgrading the KEKB B-factory [17]. It aims to reach a peak luminosity of  $60 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ , 30 times higher than its predecessor. In fig. 2.1, the luminosity profile of SuperKEKB is shown. Currently, a world record peak luminosity of  $3 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$  has been achieved and the Belle II detector, described in section 2.3, has collected  $\sim 230 \text{ fb}^{-1}$ . The final luminosity goal of Belle II is  $50 \text{ ab}^{-1}$ , while the integrated luminosity collected by the Belle experiment at KEKB was  $1 \text{ ab}^{-1}$ . The higher luminosity is achieved increasing

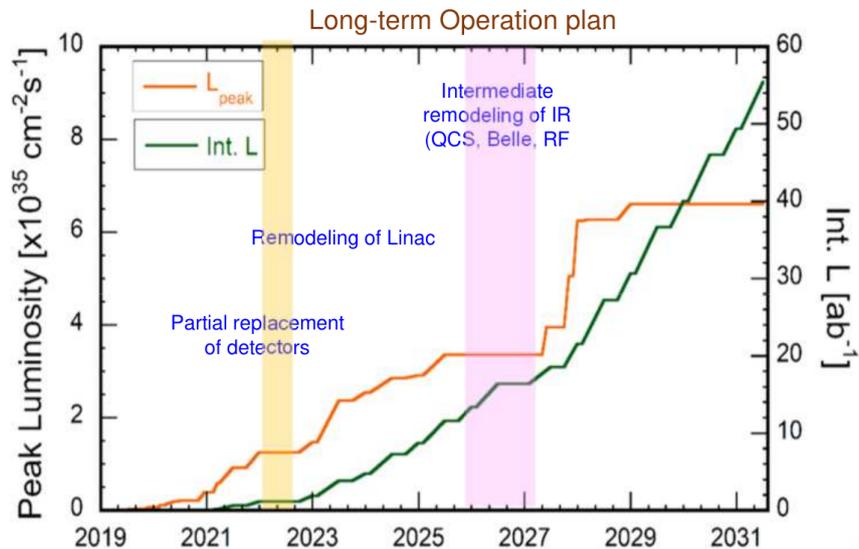


Figure 2.1: Luminosity profile of SuperKEKB [18]

the beam currents and reducing the beam dimension at the Interaction Point (IP), with the use of the nano-beam scheme [19]. The SuperKEKB configuration is shown in Fig. 2.2. The two separate rings are dedicated to the 7-GeV electron beam (HER, high-energy ring) and the 4-GeV positron ring (LER, low-energy ring). The resulting center of mass energy is  $\sqrt{s} = 10.58 \text{ GeV}$ , corresponding to the  $\Upsilon(4S)$  resonance. The asymmetric beam energies provide a Lorentz boost to the center-of-mass system with  $\beta\gamma = 0.28$ . As a consequence, the  $B$  or  $D$  mesons produced in the collision, travel an appreciable distance before decaying, allowing to precisely measure lifetimes, mixing parameters and CP violation.

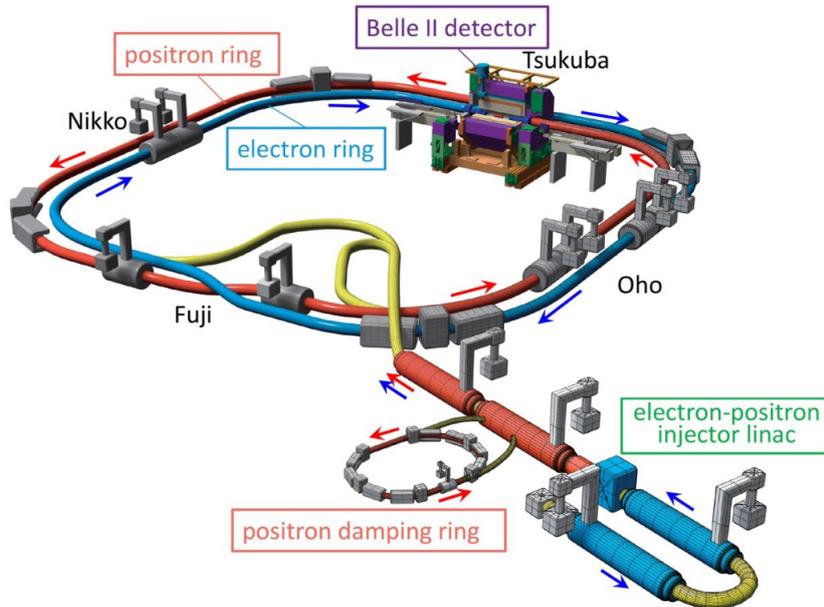


Figure 2.2: Schematic view of SuperKEKB

## 2.2 Physics program of Belle II at SuperKEKB

The Belle II detector, at SuperKEKB, aims to search for new physics at the intensity frontier, measuring many processes with unprecedented precision and allowing to appreciate discrepancies with the Standard Model predictions due to new particles and processes occurring at mass scales higher than the ones reached by the Large Hadron Collider (LHC).

One of the questions addressed by the Belle II experiment is the presence of new  $C\mathcal{P}$  violating phases in the Standard Model, for example with the measurement of time-dependent  $C\mathcal{P}$  violation in penguin transitions, such as the  $b \rightarrow s$  transitions described in section 1.4.1. Furthermore, the measurement of flavour transitions to  $\tau$  leptons, such as  $B \rightarrow \tau\nu$  and  $B \rightarrow D^{(*)}\tau\nu$  will allow to search for charged Higgs bosons, that are predicted to exist by many extensions of the Standard Model. Left-right symmetry and the presence of new flavour changing neutral currents from physics beyond the Standard Model are investigated measuring processes such as  $b \rightarrow s\nu\bar{\nu}$  transitions and precisely measuring  $b \rightarrow d$ ,  $b \rightarrow s$  and  $c \rightarrow u$  transitions. The decay of  $\tau$  leptons are studied to search for sources of lepton flavour violation beyond the Standard Model. Moreover, besides the measurements in the flavour sector, Belle II is sensitive to dark matter searches looking at missing energy decays and it will contribute to the study of quarkonia. These analyses require a  $4\pi$  spectrometer with good detection capabilities for all neutral and charged particles and excellent PID capabilities. In the next section a detailed description of the Belle II detector is given.

The main competitor of the Belle II experiment is LHCb [6] at LHC, where the  $B$  production cross section is large and more statistics can be collected. Nonetheless, the clean environment at the  $e^+e^-$  SuperKEKB collider gives Belle II an edge in the study of decays involving neutral particles in the final state or missing energy (e.g.  $B \rightarrow K\nu\bar{\nu}$ ), since the full kinematic information of the event can be retrieved knowing the initial state and reconstructing the final state.

## 2.3 The Belle II detector

The Belle II detector is designed for the SuperKEKB collider and it is optimized to operate at a higher event rate and reduced center-of-mass boost with respect to its predecessor Belle. The detector has a

cylindrical structure around the beam pipe and it is almost entirely located inside a superconducting solenoid, providing a 1.5 T magnetic field. The structure of the detector is shown in Fig. 2.3. In

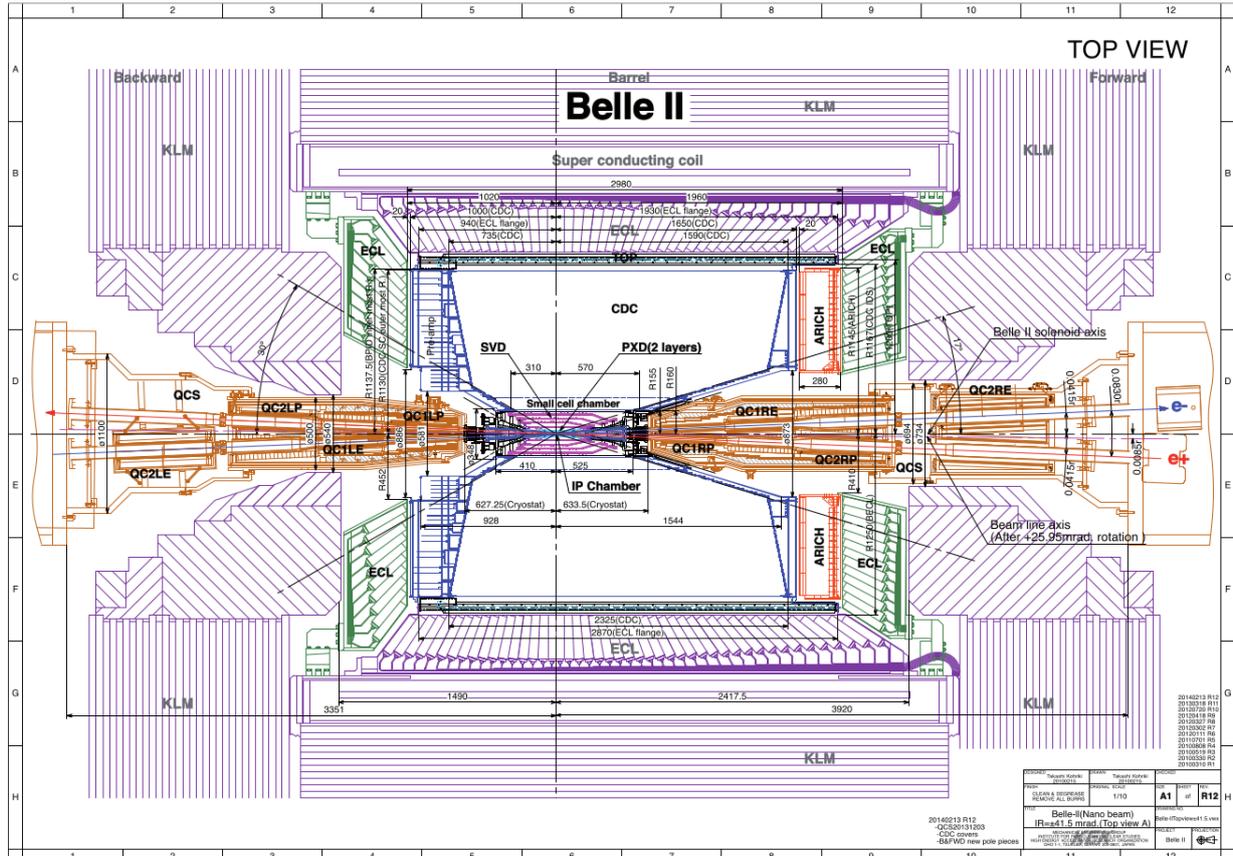


Figure 2.3: Top view of the Belle II detector

the following sections, the Belle II subdetectors are presented. A detailed description of the Belle II detector can be found in [20] and [9].

### 2.3.1 Vertex detector (VXD)

The innermost component of the Belle II detector is the vertex detector, and it comprises of two devices: a silicon pixel detector (PXD) and a silicon vertex detector (SVD). The main purpose of this subdetector is to measure the  $B$  decays vertices for the measurement of  $CP$  asymmetry. Currently, the PXD detector consists in two layers at  $r = 14$  mm and  $r = 22$  mm from the IP of DEPFET (DEPLETED Field Effect Transistor) pixels. The pixelated sensors allow a fine granularity, needed to sustain the high hit rate close to the IP and the higher luminosity with respect to Belle. A third PXD layer will be installed during the 2022 shutdown. The SVD is composed of 4 layers placed at  $r = 39$  mm,  $r = 80$  mm,  $r = 104$  mm and  $r = 135$  mm from the IP, equipped with double-sided silicon strip sensors.

### 2.3.2 Central Drift Chamber (CDC)

The CDC is a wire drift chamber filled with a mixture of helium and ethane (50% He, 50% C<sub>2</sub>H<sub>6</sub>). This subdetector is dedicated to the reconstruction of charged tracks and the precise measurement of their momenta, to particle identification, using measurements of energy loss within its gas volume, and it provides trigger signals for charged particles. In order to operate at high event rates, the chamber has smaller drift cells than the ones used in Belle. The CDC contains 14336 sense wires arranged in 56 layers, either in *axial* orientation (aligned with the solenoidal magnetic field) or *stereo* (skewed with

respect to the axial wires). Combining information from the axial and stereo layers it is possible to reconstruct a full three-dimensional helix track.

### 2.3.3 Particle Identification (PID)

Additional information on particle identification is provided by two dedicated subdetectors, based on the Cherenkov effect: the Time-Of-Propagation counter (TOP) and the Aerogel Ring-Imaging Cherenkov detector (ARICH).

#### Time of Propagation counter (TOP)

This subdetector provides PID information in the barrel region of the Belle II detector. It is composed of 16 modules arranged around the CDC. Each module consists in a quartz bar with a focusing mirror in the forward end and an array of MCP-PMTs (microchannel-plate photomultipliers), capable of excellent time resolution ( $\sim 70$  ps) and with fine granularity to measure also position, placed at the end of a trapezoidal quartz expansion box. For a charged particle crossing the quartz bar at a given position and angle, the propagation of the Cherenkov light inside the quartz is a function of the Cherenkov emission angle  $\theta_C$  (Fig. 2.4). The position of arrival of the photons on the array of PMTs, as well as the time of arrival, thus depends of the ID of the incoming particle. For each crossing particle the number of emitted Cherenkov photons which are detected by the MCP-PMT, taking into account their quantum efficiency, is  $\mathcal{O}(10)$ . Comparing the distributions of time and position of the arrival of photons with the expected PDFs corresponding to the particle hypotheses  $e, \mu, \pi, K, p, d$ , it is possible to obtain identification probabilities from the likelihood ratios. The detector is mostly dedicated to  $\pi$   $K$  separation, but can provide some information also for other particles.

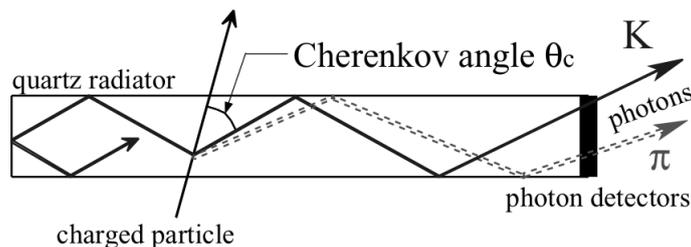


Figure 2.4: Schematic side-view of the TOP counter

#### Aerogel Ring-Imaging Cherenkov detector (ARICH)

The Aerogel Ring-Imaging Cherenkov detector (ARICH) covers the forward region of the detector. This subdetector is dedicated to the separation of pions and kaons and to the discrimination between low momentum pions, electrons and muons. A schematic view of the ARICH detector can be found in Fig. 2.5. It is composed of an aerogel radiator, where the Cherenkov photons are emitted and an expansion volume, to allow the Cherenkov photons to form rings on the photon detector surface.

### 2.3.4 Electromagnetic Calorimeter (ECL)

The electromagnetic calorimeter is used to detect photons and electrons and to measure their energy. Belle II ECL consists in a highly segmented array of thallium-doped caesium iodide CsI(Tl) crystals, covering about 90% of the solid angle in the centre-of-mass system. The CsI(Tl) crystals, preamplifiers, and support structures have been reused from Belle detector, while the readout electronics and the reconstruction software have been upgraded. The energy resolution of the calorimeter can be approximated as:

$$\frac{\sigma_E}{E} = \sqrt{\left(\frac{0.066\%}{E}\right)^2 + \left(\frac{0.81\%}{\sqrt[4]{E}}\right)^2 + (1.34)^2}$$

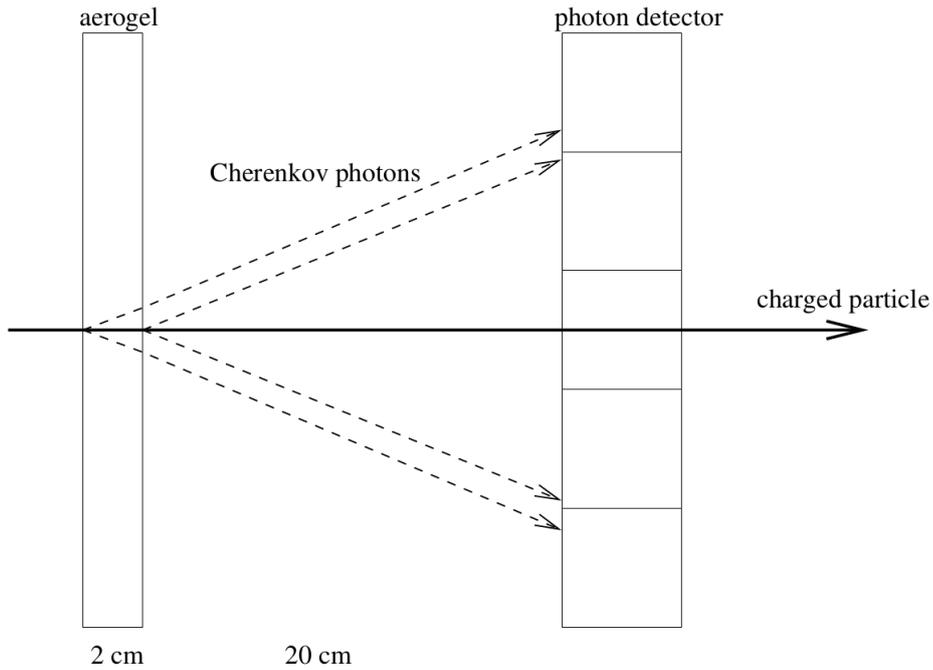


Figure 2.5: Schematic view of ARICH detector

where  $E$  is in GeV.

### 2.3.5 $K_L$ and Muon detector (KLM)

The KLM detector is located outside of the superconducting solenoid and it is composed of 4.7 cm iron plates alternated with active detector elements. It is used to measure long living particles such as muons and  $K_L$ s. The iron plates, besides providing 3.9 interaction lengths of material, serve as the magnetic flux return for the solenoid.

### 2.3.6 Trigger

The Belle II trigger system is composed of two levels: a hardware-based low-level trigger (L1) and a software-based high-level trigger (HLT). The L1 trigger mainly uses the information coming from the CDC and the ECL. It has a latency of  $5 \mu\text{s}$  and a maximum output rate of 30kHz. The efficiency for  $B\bar{B}$  events is close to 100%.

The HLT goal is to reduce online event rates to 10 kHz for offline storage. It uses all the subdetectors (except the PXD) to perform an offline reconstruction of the event, applying physics requirements to reduce the event rate.



# Chapter 3

## Introduction to the analysis of the $B \rightarrow \eta' K$ decay

In this chapter, the datasets used for this analysis are described, introducing the signal and the main background sources. A description of the analysis strategy is presented with information on the reconstruction and the signal selection.

### 3.1 Signal and decay modes

At SuperKEKB a  $B\bar{B}$  meson pair is produced via  $e^+e^- \rightarrow \Upsilon(4S) \rightarrow B\bar{B}$  process, with  $B = B^0, B^\pm$ . The signal for this study consists in events where one of the two  $B$  mesons decays into the  $C\mathcal{P}$  eigenstate  $\eta' K$ .

The  $K^0$  can be seen as a mixture of  $K_S$  and  $K_L$  states, that can be approximated to  $C\mathcal{P}$  eigenstates with opposite eigenvalues:  $\xi_{K_S} = +1$  and  $\xi_{K_L} = -1$ . In this analysis only the final state with a  $K_S$  decaying into two charged pions is considered, since the  $K_L$  requires a different reconstruction technique and the  $K_S \rightarrow \pi^0\pi^0$  final state has low reconstruction efficiency and large combinatorial background. These final states will be included in future analyses.

The  $\eta'$  meson is reconstructed using the following (most probable) decay modes:

- $\eta' \rightarrow \eta\pi^+\pi^-$
- $\eta' \rightarrow \rho\gamma$

with  $\eta \rightarrow \gamma\gamma$  and  $\rho \rightarrow \pi^+\pi^-$ . The final state  $\eta' \rightarrow \eta(\pi^+\pi^-\pi^0)\pi^+\pi^-$  has been excluded on account of its low reconstruction efficiency. This final state has about half the branching fraction and a preliminary study [21] showed that the reconstruction efficiency is significantly lower than the sibling channel  $\eta' \rightarrow \eta(\gamma\gamma)\pi^+\pi^-$  ( $\frac{\epsilon_{3\pi}}{\epsilon_{2\gamma}} = 0.14 \pm 0.01$ ), mostly because of the lower momenta of the three pions, especially the  $\pi^0$ , from its decay. So, the signal yield is expected to be significantly lower. With larger integrated luminosity, and optimized reconstruction and selection also this channel can be included in the future.

The inclusive Branching Fraction for  $B \rightarrow \eta' K$  and the Branching fractions for each  $\eta'$  decay mode can be found in Table 3.1.

### 3.2 Background sources

The production of a  $\Upsilon(4S)$ , decaying into a  $B\bar{B}$  meson pair, is not the most probable result of an  $e^+e^-$  collision at  $\sqrt{s} = 10.58$  GeV. Cross sections for the most important physics processes, occurring at the default center of mass energy, are given in Table 3.2. The  $e^+e^- \rightarrow e^+e^-$  process (*Bhabha scattering*) has the largest cross section but these background events are efficiently discarded at trigger level. The

Mode	Decay channel	Branching fraction
$B^+ \rightarrow \eta' K^+$	inclusive	$7.06 \times 10^{-5}$
	$\eta' \rightarrow \eta(\gamma\gamma)\pi^+\pi^-$	$1.19 \times 10^{-5}$
	$\eta' \rightarrow \rho(\pi^+\pi^-)\gamma$	$2.04 \times 10^{-5}$
	total	$3.23 \times 10^{-5}$
$B^0 \rightarrow \eta' K$	inclusive	$6.6 \times 10^{-5}$
	$\eta' \rightarrow \eta(\gamma\gamma)\pi^+\pi^-$	$5.54 \times 10^{-6}$
	$\eta' \rightarrow \rho(\pi^+\pi^-)\gamma$	$9.54 \times 10^{-6}$
	total	$1.51 \times 10^{-5}$

Table 3.1: Branching fractions for the final states relevant to this analysis [8]

Physics process	Cross section [nb]
$\Upsilon(4S)$	$1.110 \pm 0.008$
$u\bar{u}(\gamma)$	0.61
$d\bar{d}(\gamma)$	0.40
$s\bar{s}(\gamma)$	0.38
$c\bar{c}(\gamma)$	1.30
$e^+e^-(\gamma)$	$300 \pm 3$ (MC stat.)
$\gamma\gamma(\gamma)$	$4.99 \pm 0.05$ (MC stat.)
$\tau^+\tau^-(\gamma)$	0.919
$\mu^+\mu^-(\gamma)$	1.148

Table 3.2: Total production cross section for several physics processes from  $e^+e^-$  collisions at  $\sqrt{s} = 10.58$  GeV [9]

main background source for this analysis is the *continuum background*, i.e. non-resonant  $e^+e^- \rightarrow q\bar{q}$  hadronic events ( $q = u, d, c, s$ ) and  $e^+e^- \rightarrow \tau^+\tau^-$ . These events have several tracks and neutrals in the final state and it is possible to find a combination of tracks that mimics the signal without being from a real  $B$  decay. In particular, the  $c\bar{c}$  events have a relative large probability to produce an  $\eta'$  in the final state, which can combine to a spurious or misreconstructed kaon to mimic the signal. Since the cross section for these processes is large, the amount of continuum candidates misidentified as signal candidates is significant and a precise continuum suppression strategy needs to be applied. Continuum and  $B\bar{B}$  events can be separated exploiting the information on event topology: continuum candidates are strongly collimated, due to the large momentum available for the decay to light quarks, while, in contrast,  $B\bar{B}$  decay products are uniformly distributed, as shown in Fig. 3.1. A detailed description of the continuum suppression can be found in Chapter 4.

Another contribution to the background comes from true  $B\bar{B}$  decays, both charged and neutral, decaying to final states different from the signal one. This background is referred to as *peaking background*. To discuss peaking background discrimination two variables are needed: the beam constrained mass  $M_{bc} = \sqrt{E_{beam}^2 - P_B^2}$  and the energy difference  $\Delta E = E_B - E_{beam}$ , with  $E_{beam} = \frac{\sqrt{s}}{2}$  and  $(E, P)_B$  energy/momentum of the  $B$  candidate in the center of mass system. This two variables are built to exploit the kinematical constraints of the  $\Upsilon(4S)$  decay to  $B\bar{B}$  pairs: if the  $B$  meson is correctly reconstructed the energy of its decay products is half the center of mass energy  $\sqrt{s}$ , or, equivalently, equal to the beam energy in the  $\Upsilon(4S)$  rest frame. The  $\Delta E$  variable strongly depends on the mass hypothesis for the final state particles and is very effective in discriminating signal candidates from peaking candidates. If, for example, a kaon is mistaken for a pion, its reconstructed energy will be smaller than its true energy, shifting the event towards the negative tail of the  $\Delta E$  distribution. In contrast,  $M_{bc}$  does not depend on the mass hypothesis for the particles, leading to a distribution peaked at the nominal mass of the  $B$  for both signal  $B$  decays and background  $B$  decays (hence the name *peaking* background).

Both  $M_{bc}$  and  $\Delta E$  are very effective also for continuum background discrimination and they will be used to extract the signal candidates, as described in section 3.3. A detailed description of  $M_{bc}$  and  $\Delta E$  variables can be found in [22].

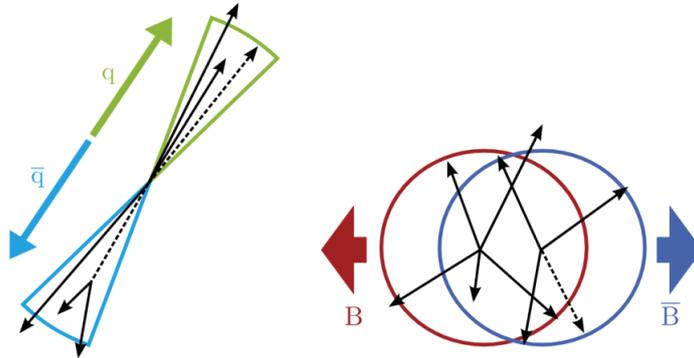


Figure 3.1: Schematic representation of event shape for continuum (left) and  $B\bar{B}$  (right) events

### 3.3 Analysis strategy

The first step of the analysis is the reconstruction of the  $B$  candidates for each decay mode applying a preliminar selection, as discussed in section 3.5. After the reconstruction, a signal selection on a set of kinematical variables is performed, to reduce the amount of background. A multivariate discriminator is trained on variables related to the event topology to separate signal and continuum background. The output variable of the discriminator ( $CS_{var}$ ) is validated on the off-resonance dataset (data collected 60 MeV below the resonance, see section 3.4) and a multivariate Monte Carlo reweighting procedure is applied to reduce the discriminator input variables mismodelling in Monte Carlo. No cut is applied to  $CS_{var}$ : the variable is included in the Unbinned Maximum Likelihood fit for the signal extraction. The signal yield is obtained fitting the observables  $M_{bc}$ ,  $\Delta E$  and  $CS_{var}$ . The fit procedure is tested and validated using toy Monte Carlo samples. From the signal yield a measurement of the branching fraction for each decay mode has been obtained. Finally, in view of the future measurement of the time dependent  $CP$  violation in the  $B \rightarrow \eta'K$  decay, a preliminar study of the resolution function  $\mathcal{R}(\Delta t)$  is presented.

### 3.4 Data and Monte Carlo samples

In this section the data and Monte Carlo samples used for this analysis are presented. Only events passing the HLT\_HADRON selection criteria have been considered. The HLT\_HADRON requires at least 3 good tracks coming from the interaction point ( $p_t > 0.2 \frac{GeV}{c}$ ,  $|d_0| < 2.0$  cm,  $|z_0| < 4.0$  cm) and uses a Bhabha event skim as veto.

#### Data

The dataset used in this analysis was collected at Belle II in 2019 and 2020 and consists of two samples:

- $\mathcal{L} = 62.8 \pm 0.5 \text{ fb}^{-1}$  collected at the  $\Upsilon(4S)$  resonance [23], corresponding to  $N(B\bar{B}) = 68.21 \times 10^6 \pm 0.09\%(\text{stat}) \pm 1.3\%(\text{syst})$   $B\bar{B}$  pairs [24], where the uncertainties include the uncertainty on cross section, integrated luminosity and possible shifts of the collision energy;
- $\mathcal{L} = 9.2 \pm 0.2 \text{ fb}^{-1}$  off-resonance dataset, i.e. data collected approximately 60 MeV below the  $\Upsilon(4S)$  resonance. At this energy no  $B$  mesons can be produced, giving a pure sample of continuum candidates, used for continuum suppression studies.

## Monte Carlo

A Monte Carlo simulated dataset is used to validate the analysis strategy and perform background rejection studies. The full simulation uses `KKMC` [25] for the generation of quark-antiquark pairs from the  $e^+e^-$  collisions and `PYTHIA8` [26] for the hadronization. The decays of the produced hadrons are simulated using `EVTGEN` [27] while `GEANT4` [28] is used for the detector response. The simulation includes also beam background events. The peaking background candidates are reconstructed from a sample of generic  $B\bar{B}$  Monte Carlo, where all the possible final states are included, removing events where one  $B$  meson decays to the signal final state. The signal Monte Carlo, instead, is obtained simulating  $e^+e^- \rightarrow \mathcal{T}(4S) \rightarrow B\bar{B}$  events and forcing one of the two  $B$  mesons to decay into the  $\eta'K$  final state. The equivalent integrated luminosity of the continuum background Monte Carlo sample is  $\mathcal{L}^{eq} = 0.5 \text{ ab}^{-1}$ , while for the peaking background  $\mathcal{L}^{eq} = 2 \text{ ab}^{-1}$  are used. For each signal decay channel,  $1 \times 10^6$  simulated signal events are reconstructed.

## 3.5 Reconstruction

This study originates from the successful rediscovery of  $\eta$  and  $\eta'$  mesons in early Belle II data [29], [21].

In order to describe the reconstruction technique, the following variables are needed:

- `clusterE1E9`: ratio between the energy deposit in the central crystal (E1) and the sum of the energies in a  $3 \times 3$  matrix of crystals around the central one (E9). Its value is higher for the photons and lower for the hadrons.
- `clusterE9E21`: ratio of the sum of the energies in the inner  $3 \times 3$  crystals matrix, E9, and  $5 \times 5$  crystals around the central crystal, excluding the corners (E21).
- `clusterTiming`: is the difference between the photon timing and the event time. Photons produced at the interaction point should have `clusterTiming` close to zero.
- `clusterErrorTiming`: ECL cluster's timing uncertainty that contains 99% of true photons
- `clusterNHits`: is the sum of the weights  $w_i$  of all the crystals in an ECL cluster. In case of non-overlapping clusters it is equal to the number of crystals in the clusters, otherwise, in case of energy splitting among different clusters, it can be a non-integer value.

### Reconstruction of the $\eta' \rightarrow \eta(\gamma\gamma)\pi^+\pi^-$ channel

The photons are required to be within the CDC geometrical acceptance, with a loose cut on `clusterErrorTiming` and with `clusterE1E9`  $> 0.4$  or  $E > 0.075 \text{ GeV}$ . Furthermore each photon is selected to have  $60 \text{ MeV} < E_\gamma < 6 \text{ GeV}$ , `clusterE9E21`  $> 0.9$ , `clusterNHits`  $> 1.5$  and `clusterTiming`  $< 600 \text{ ns}$ . The  $\eta$  meson candidates are reconstructed from the two photons requiring their invariant mass to be  $0.4 \frac{\text{GeV}}{c^2} < M_{\gamma\gamma} < 0.7 \frac{\text{GeV}}{c^2}$ . The  $\eta$  candidate is combined with two oppositely charged pions in the CDC acceptance region to form an  $\eta'$  candidates. The  $\eta'$  invariant mass is required to be  $0.9 \frac{\text{GeV}}{c^2} < M_{\eta'} < 1.1 \frac{\text{GeV}}{c^2}$ .

### Reconstruction of the $\eta' \rightarrow \rho(\pi^+\pi^-)\gamma$ channel

The  $\rho$  meson is reconstructed from a  $\pi^+\pi^-$  pair with invariant mass  $0.47 \frac{\text{GeV}}{c^2} < M_\rho < 1.07 \frac{\text{GeV}}{c^2}$ . A photon, with the same requirements described in section 3.5, is combined to the  $\rho$  to form an  $\eta'$  candidate with invariant mass  $0.9 \frac{\text{GeV}}{c^2} < M_{\eta'} < 1.1 \frac{\text{GeV}}{c^2}$ .

### Reconstruction of the $B$ meson

The  $B$  meson candidates are reconstructed combining the  $\eta'$  candidate with a  $K^\pm$  or  $K_S^0$ . The  $K^\pm$  are required to have  $\text{PID} > 0.1$  for the kaon hypothesis, also using a PID likelihood which uses all detectors' information. The  $K_S^0$  is reconstructed from a  $\pi^+\pi^-$  pair with invariant mass  $450 \frac{\text{MeV}}{c^2} < M_{\pi^+\pi^-} < 550$

$\frac{\text{MeV}}{c^2}$ . The decay chain is fitted with the `treeFit` algorithm, updating the momenta of the daughter particles. A description of the vertex fit can be found in section 6.3.

For each event, more than one  $B$  candidate can be reconstructed for the specified decay. In table 3.3 the average number of  $B$  candidates per event is given.

Decay channel	$N_{\text{cands}}$
$B^0 \rightarrow \eta'(\eta_{\gamma\gamma}\pi^+\pi^-)K_S^0$	2.23
$B^\pm \rightarrow \eta'(\eta_{\gamma\gamma}\pi^+\pi^-)K^\pm$	1.77
$B^0 \rightarrow \eta'(\rho_{\pi^+\pi^-}\gamma)K_S^0$	6.19
$B^\pm \rightarrow \eta'(\rho_{\pi^+\pi^-}\gamma)K^\pm$	5.65

Table 3.3: Average number of  $B$  candidates per event ( $N_{\text{cands}}$ ) for each decay channel

To choose the best candidate, the information on the quality of the  $B$ -vertex fit is used, retaining the candidate with best  $\chi^2$  probability. In about 95% of the cases, the candidate is correctly reconstructed, according to MC truth. The misreconstructed signal candidates are called *self cross-feed* (SxF) candidates. In Fig. 3.2 the results of the study on SxF origin performed on  $B^0 \rightarrow \eta'(\eta_{\gamma\gamma}\pi^+\pi^-)K_S^0$  Monte Carlo candidates are shown. The incorrect reconstruction of the  $B^0$  is almost entirely due to the  $\eta'$  reconstruction. The misreconstruction of the  $\eta'$  is mainly due to the  $\eta$  reconstruction (wrong photon selection) but also pions are frequently misidentified.

In Table 3.4 the Signal reconstruction efficiency  $\varepsilon_{\text{reco}}$  (%) and the SxF fraction, computed as  $\frac{N_{\text{SxF}}}{N_{\text{sig}}+N_{\text{SxF}}}$ , for each final state is given.

Decay	$\varepsilon_{\text{reco}}$ (%)	SxF fraction (%)
$B^0 \rightarrow \eta'(\eta_{\gamma\gamma}\pi^+\pi^-)K_S^0$	$39.7 \pm 0.1$	$18.4 \pm 0.1$
$B^\pm \rightarrow \eta'(\eta_{\gamma\gamma}\pi^+\pi^-)K^\pm$	$40.4 \pm 0.1$	$15.8 \pm 0.1$
$B^0 \rightarrow \eta'(\rho_{\pi^+\pi^-}\gamma)K_S^0$	$30.1 \pm 0.1$	$30.1 \pm 0.2$
$B^\pm \rightarrow \eta'(\rho_{\pi^+\pi^-}\gamma)K^\pm$	$30.7 \pm 0.1$	$27.1 \pm 0.2$

Table 3.4: Signal reconstruction efficiency and SxF fraction for each decay channel after reconstruction and before the signal selection described in 3.6

## 3.6 Signal selection

As discussed in section 3.2, it is possible to distinguish B events from continuum events from the event shape. However they have also different kinematics distribution which can be exploited. A selection, aimed at reducing the fraction of backgrounds and SxF preserving most of the signal, is applied to a set of kinematic variables such as photon energies and  $\eta'$ ,  $\eta$  and  $\rho$  invariant masses. The selections have been defined optimizing the figure of merit  $\frac{S}{\sqrt{S+B}}$ , where  $S$  and  $B$  are the number of signal and background candidates in the signal region (section 5.1), respectively. The signal selections applied to the kinematic variables are:

- $\eta' \rightarrow \eta(\gamma\gamma)\pi^+\pi^-$ 
  - $E_\gamma > 150 \text{ MeV}$
  - $0.5 \frac{\text{GeV}}{c^2} < M_\eta < 0.57 \frac{\text{GeV}}{c^2}$
  - $0.92 \frac{\text{GeV}}{c^2} < M_{\eta'} < 1.0 \frac{\text{GeV}}{c^2}$
- $\eta' \rightarrow \rho(\pi^+\pi^-)\gamma$ 
  - $E_\gamma > 150 \text{ MeV}$
  - $\cos\theta_\gamma > -0.64$

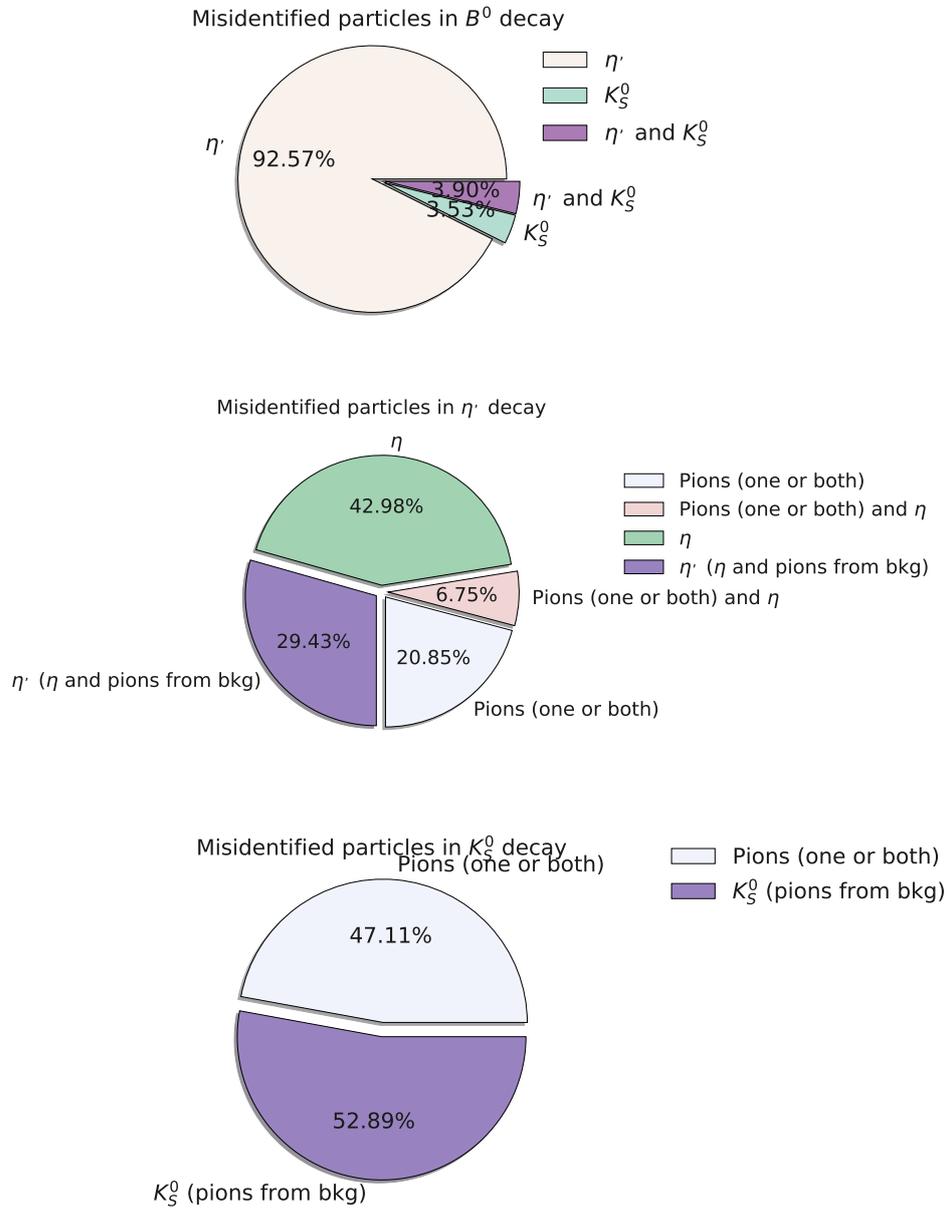


Figure 3.2: Results of SxF origin study for  $B^0 \rightarrow \eta'(\eta_{\gamma\gamma}\pi^+\pi^-)K_S^0$

- $0.51 \frac{\text{GeV}}{c^2} < M_\rho < 1.0 \frac{\text{GeV}}{c^2}$
- $0.92 \frac{\text{GeV}}{c^2} < M_{\eta'} < 1.0 \frac{\text{GeV}}{c^2}$
- $K^\pm$ 
  - $\cos\theta_{K^\pm} > -0.5$
- $K_S^0$ 
  - $\cos\theta_{p,v} > 0.99$ , where  $\theta_{p,v}$  is the angle between the momentum and the vertex vector
  - $0.49 \frac{\text{GeV}}{c^2} < M_{K_S^0} < 0.51 \frac{\text{GeV}}{c^2}$
- $B$ 
  - $M_{bc} > 5.2 \frac{\text{GeV}}{c^2}$
  - $|\Delta E| < 0.25 \text{ GeV}$

In fig. 3.3, 3.4, 3.5, 3.6, 3.7 and 3.8 the distribution of the signal selection variables, for the signal, continuum, peaking and SxF component are shown.

In table 3.5 the signal efficiency  $\varepsilon_{sel}$  and SxF fraction, after reconstruction and signal selection, is given.

Decay	$\varepsilon_{sel}(\%)$	SxF fraction(%)
$B^0 \rightarrow \eta'(\eta_{\gamma\gamma}\pi^+\pi^-)K_S^0$	$31.0 \pm 0.1$	$8.7 \pm 0.1$
$B^\pm \rightarrow \eta'(\eta_{\gamma\gamma}\pi^+\pi^-)K^\pm$	$31.8 \pm 0.1$	$7.2 \pm 0.1$
$B^0 \rightarrow \eta'(\rho_{\pi^+\pi^-}\gamma)K_S^0$	$23.6 \pm 0.1$	$8.3 \pm 0.1$
$B^\pm \rightarrow \eta'(\rho_{\pi^+\pi^-}\gamma)K^\pm$	$24.2 \pm 0.1$	$7.3 \pm 0.1$

Table 3.5: Signal efficiency and SxF fraction for each decay channel after reconstruction and signal selection

The fraction of the reconstructed continuum candidates that survive the signal selection is of the order of a few % for the neutral decays and of the order of 10% for the charged ones. However, because of the large amount of continuum background that affects the signal searched for in this analysis, a multivariate method will be used to assess this component, as discussed in the next chapter. The signal selection is also effective in reducing the peaking background, retaining only  $\sim 15$ -30% of the reconstructed peaking candidates.

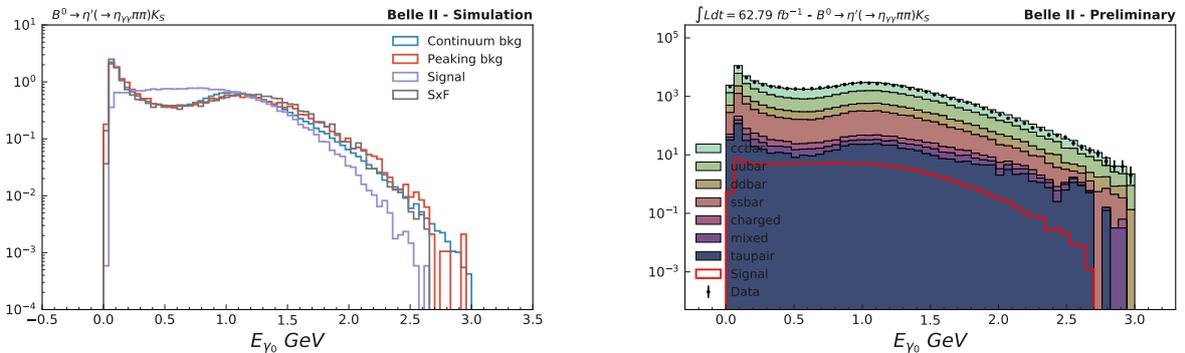


Figure 3.3: Monte Carlo distributions of  $E_\gamma$ , normalized to unit area, (left) and data-Monte Carlo comparison (right) for  $B^0 \rightarrow \eta'(\eta_{\gamma\gamma}\pi^+\pi^-)K_S^0$  decays.

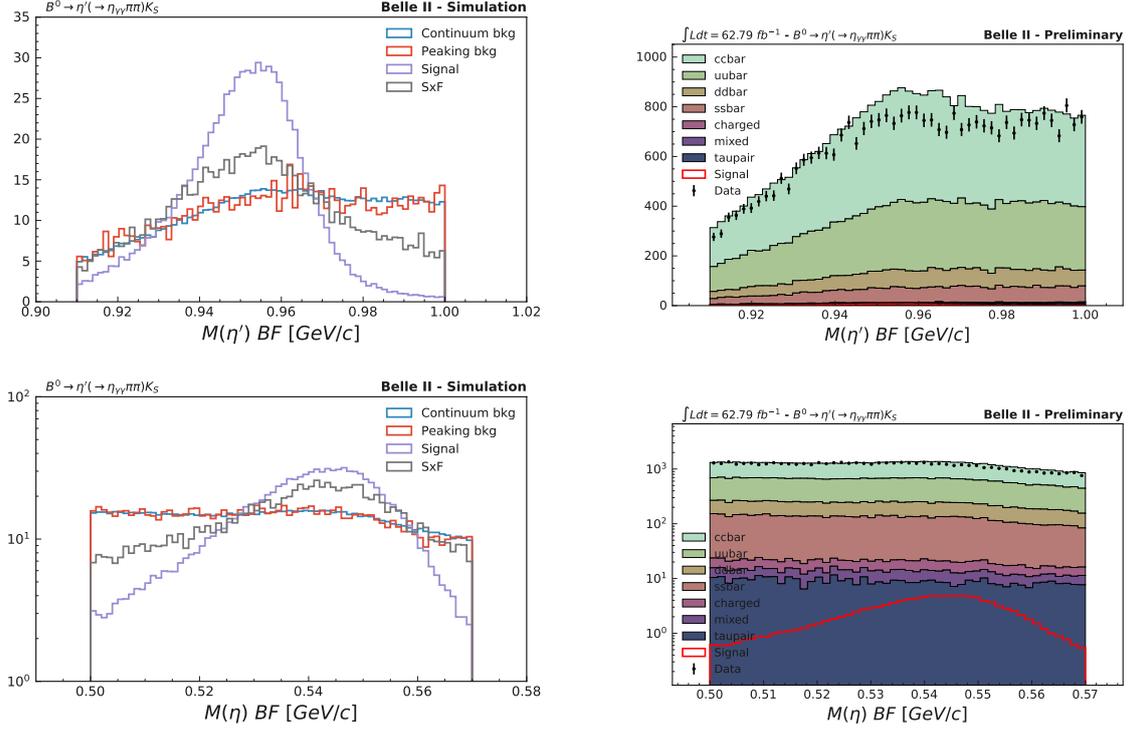


Figure 3.4: Monte Carlo distributions of  $\eta'$  (top) and  $\eta$  (bottom) invariant masses before fit constraint, normalized to unit area, (left) and data-Monte Carlo comparison (right) for  $B^0 \rightarrow \eta'(\eta_{\gamma\gamma}\pi^+\pi^-)K_S^0$  decays.

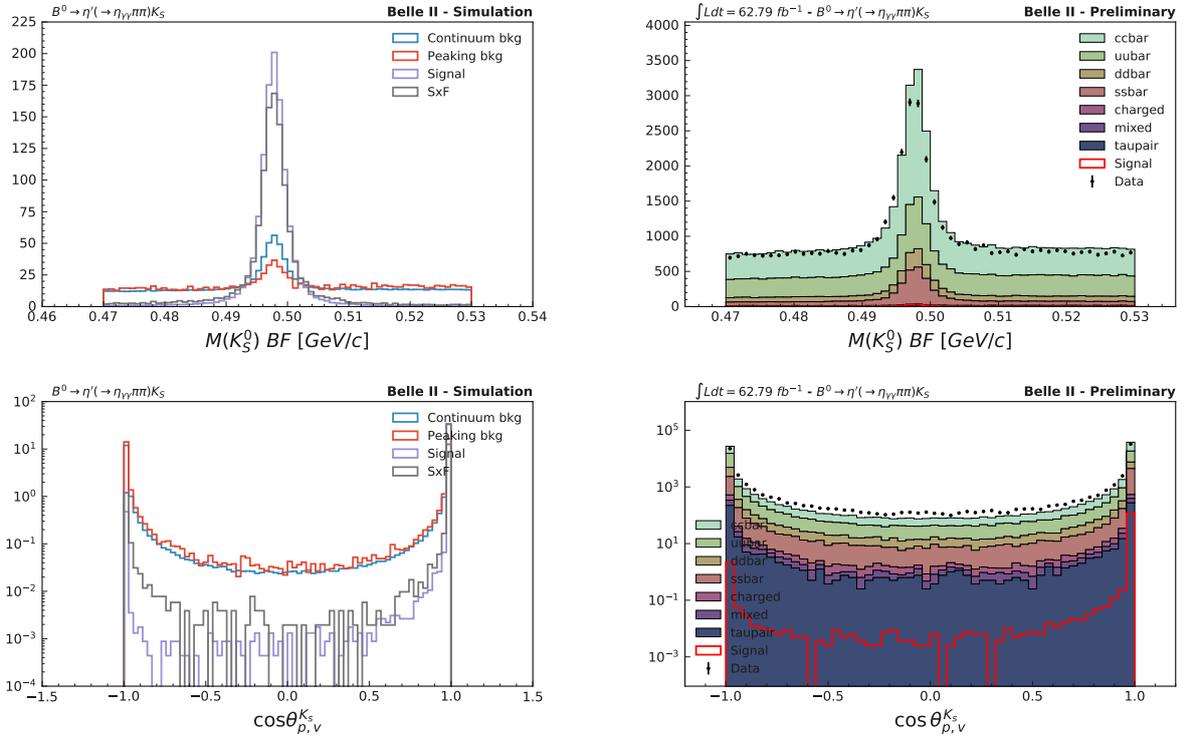


Figure 3.5: Monte Carlo distributions of  $K_S^0$  invariant mass before fit constraint (top) and cosine of the angle between  $K_S^0$  momentum and vertex vector (bottom), normalized to unit area, (left) and data-Monte Carlo comparison (right) for  $B^0 \rightarrow \eta'(\eta_{\gamma\gamma}\pi^+\pi^-)K_S^0$  decays.

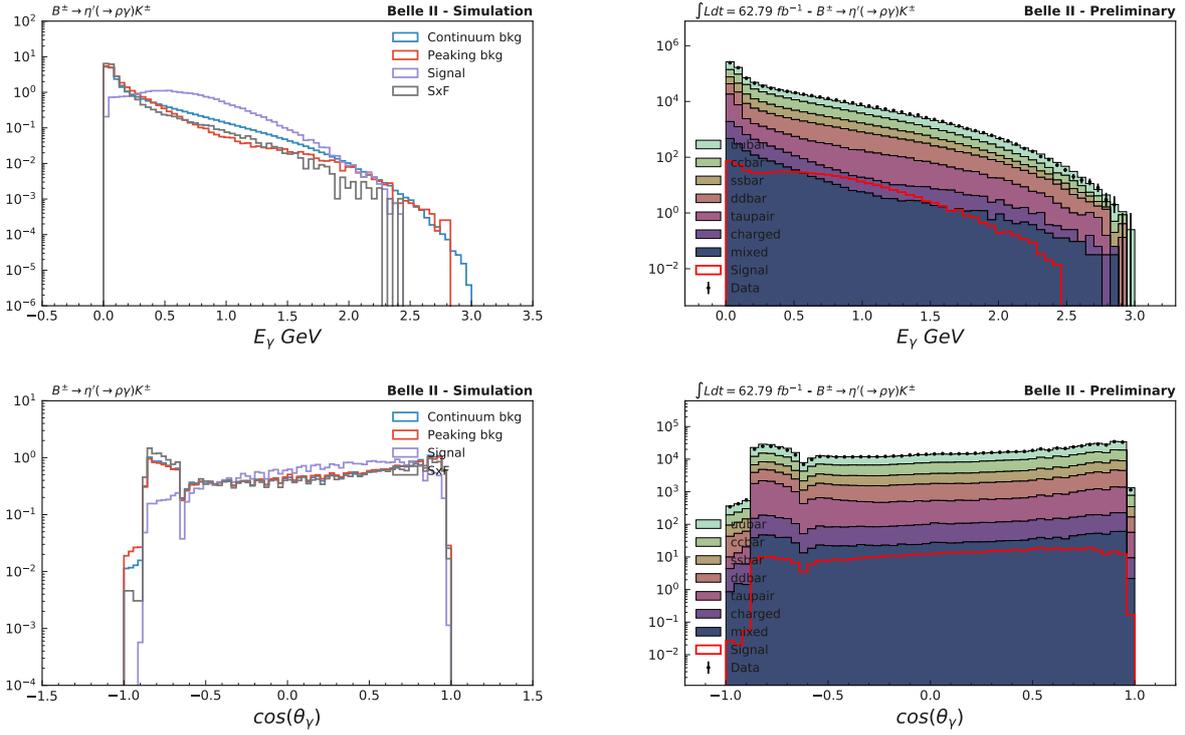


Figure 3.6: Monte Carlo distributions of  $E_\gamma$  (top) and  $\cos(\theta_\gamma)$  (bottom), normalized to unit area, (left) and data-Monte Carlo comparison (right), for  $B^\pm \rightarrow \eta'(\rho\gamma)K^\pm$  decays.

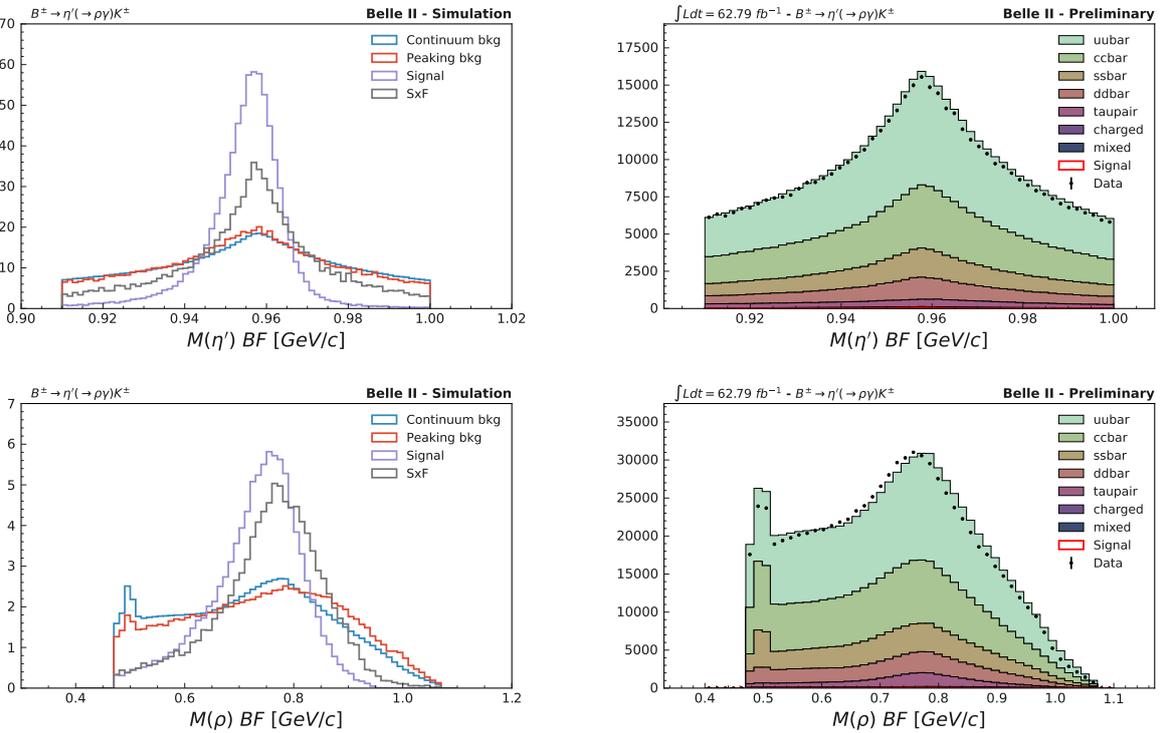


Figure 3.7: Monte Carlo distributions of  $\eta'$  invariant mass before fit constraint (top) and  $\rho$  (bottom) invariant mass, normalized to unit area (left) and data-Monte Carlo comparison (right), for  $B^\pm \rightarrow \eta'(\rho\gamma)K^\pm$  decays. The  $M_\rho$  distribution shows a peak corresponding to the  $K_S^0$ .

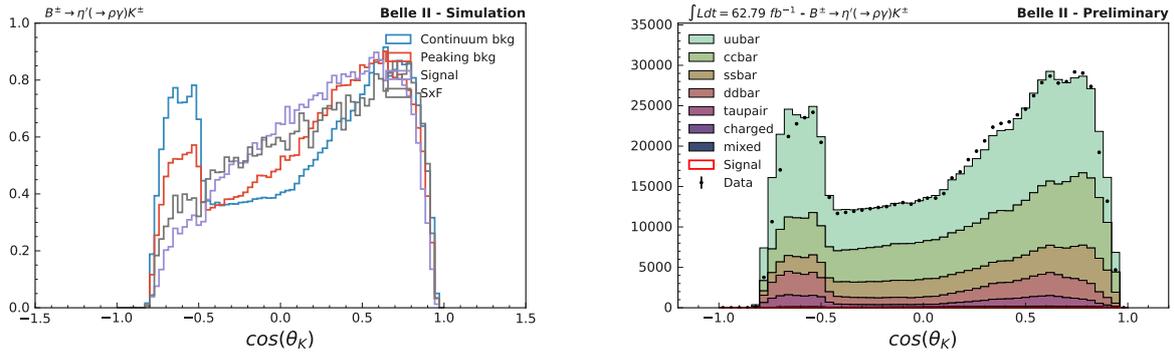


Figure 3.8: Monte Carlo distributions of  $\cos(\theta_K)$ , normalized to unit area (left) and data-Monte Carlo comparison (right), for  $B^\pm \rightarrow \eta'(\rho\gamma)K^\pm$  decays.

# Chapter 4

## Continuum Suppression

In the case of charmless  $b \rightarrow s$  decays, the background dominates over the signal by orders of magnitude. In order to measure this kind of transition it is necessary to be able to separate signal from continuum candidates. Continuum background assessment is performed using multivariate analysis (MVA) methods, exploiting information on event topology that allows to separate jet-like continuum candidates from spherical  $B\bar{B}$  candidates.

In the following sections a precise description of the algorithms, variables and strategies used for continuum suppression can be found. More details on background rejection at  $B$ -Factories can be found in Ref. [22].

### 4.1 Variables

A set of 27 features sensible to event shape and with low correlation with  $M_{bc}$ ,  $\Delta E$  and  $\Delta t$  is used to train the algorithm.

#### 4.1.1 Thrust

The thrust axis for a set of  $N$  particles of momenta  $\mathbf{p}_i$ , with  $i = 1, \dots, N$ , is defined as the vector along which their total projection is maximal. In a  $B\bar{B}$  event two thrust axes can be defined, one for particles from the signal  $B$  candidate and one for all the other particles in the event (rest-of-event or ROE). Three variables related to the thrust axes of the event are used for continuum suppression:

- $|\cos(\theta_B - \theta_O)|$ , i.e the cosine of the angle between the thrust axis of the momenta of the  $B$  candidate decay particles and the thrust axis of ROE. Since in  $B\bar{B}$  events the  $B$  mesons decay products are isotropically distributed in the  $\mathcal{T}(4S)$  system, their thrust axes are randomly distributed, and thus  $|\cos(\theta_B - \theta_O)|$  is uniformly distributed between 0 and 1. In contrast,  $q\bar{q}$  events have thrust axes pointing back-to-back, giving a  $|\cos(\theta_B - \theta_O)|$  distribution strongly peaked around 1, as shown in Fig. 4.1.
- $|\cos(\theta_B z)|$ , the cosine between the thrust axis of the signal  $B$  and the z-axis. For the  $B$  decay particles this variable is uniformly distributed, while the continuum candidates thrust axes follow the angular distribution of the jets, resulting in the distribution shown in Fig. 4.1.
- $\text{Thrust}_{Om}$  is the magnitude of the ROE thrust axis.

#### 4.1.2 Fox-Wolfram moments

For a collection of  $N$  particles with momenta  $\mathbf{p}_i$ , the  $k$ -th order Fox Wolfram moment is defined as

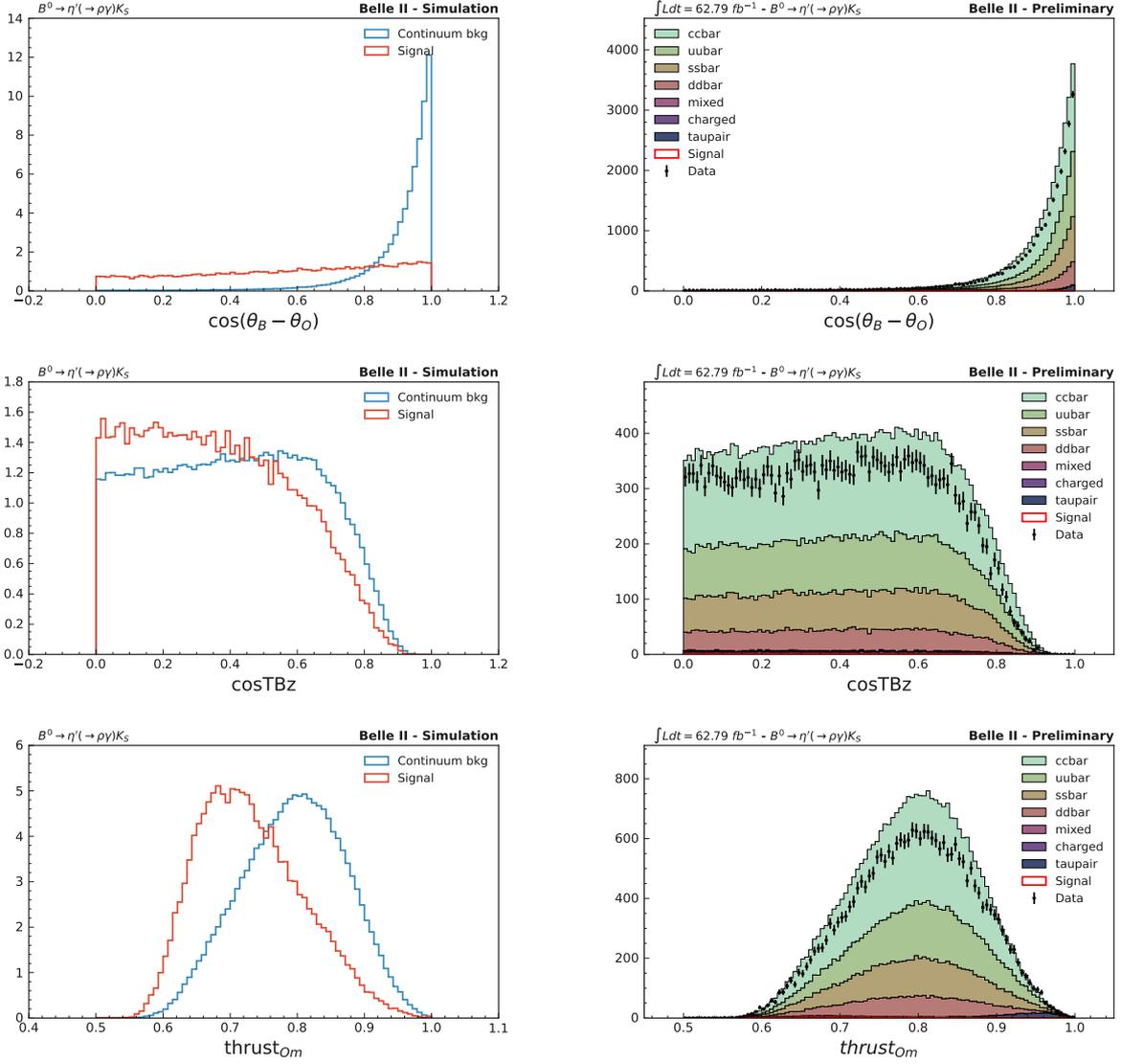


Figure 4.1: Distribution of the thrust axis variables for signal and continuum Monte Carlo candidates, normalized to unit area (left) and data-Monte Carlo comparison (right).

$$H_k = \sum_{i,j}^N |\mathbf{p}_i| |\mathbf{p}_j| P_k(\cos\theta_{ij}) \quad (4.1)$$

where  $P_k$  is the  $k$ -th Legendre polynomial and  $\theta_{ij}$  is the angle between the two momenta  $\mathbf{p}_i$  and  $\mathbf{p}_j$ . From the Fox-Wolfram moments it is possible to define the normalized ratio

$$R_k = \frac{H_k}{H_0} \quad (4.2)$$

Fifteen variables constructed from the Fox-Wolfram moments are included in this analysis:

- $R_2$ , the normalized second Fox-Wolfram moment.
- Kakuno-Super-Fox-Wolfram moments (KSFW)  $H_{c,k}^g$  [30]. These complex variables are constructed performing the sum in Eq. 4.1 on different sets of particles, as specified by the  $g$  parameter:  $g = so$  indicates that one of the indices in the sum runs on the signal side particles and the other index runs on the ROE particles, if  $g = oo$  both indices run on ROE particles set.

The  $c$  parameter is used only in the case  $g = so$  and describes whether the considered particles are charged ( $c = 0$ ), neutral ( $c = 1$ ) or missing ( $c = 2$ ). An example of the distribution of the KSFW moment  $H_{02}^{so}$  can be found in Fig. 4.2.

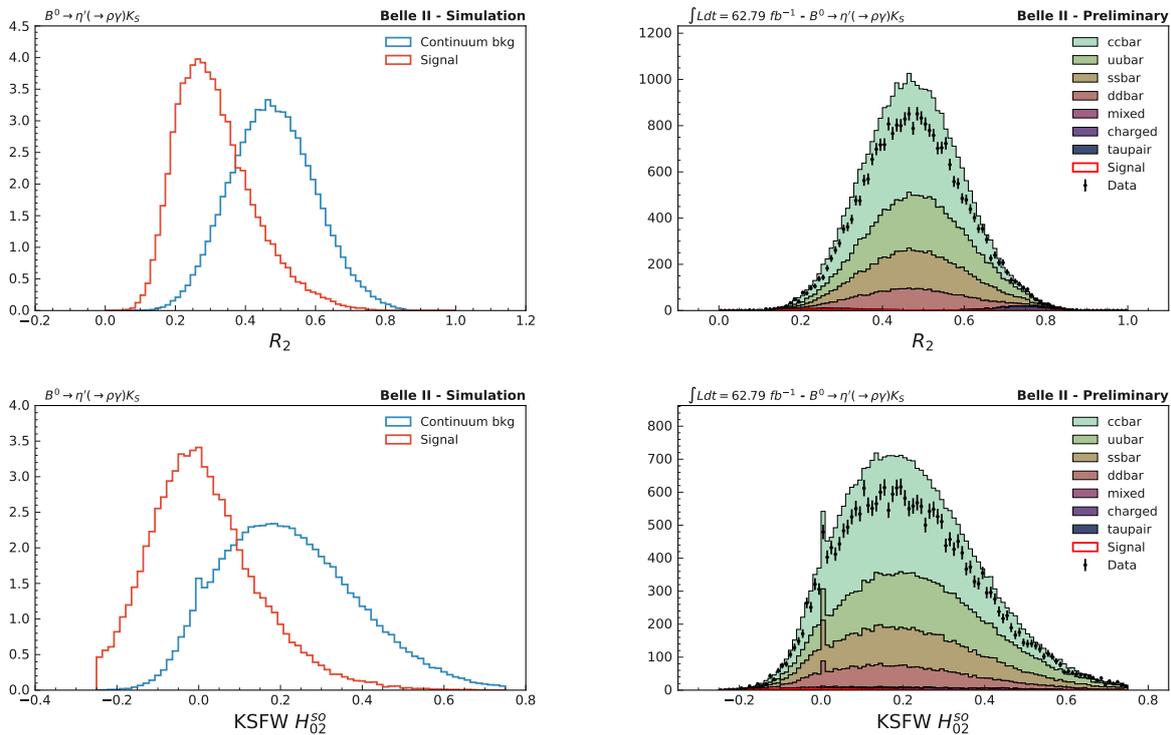


Figure 4.2: Distribution of the variables  $R_2$  and KSFW moment  $H_{02}^{so}$  for signal and continuum Monte Carlo candidates, normalized to unit area (left) and data-Monte Carlo comparison (right).

### 4.1.3 CLEO cones

The CLEO cones have been introduced by the CLEO Collaboration [31]. These variables measure the scalar momentum flow into concentric cones in angular intervals of  $10^\circ$  around the thrust axis of the  $B$  candidate. The distributions of the first and second CLEO cones for  $B^0 \rightarrow \eta'_{\rho\gamma} K_S$  can be found in Fig 4.3.

## 4.2 FastBDT

The signal/background classification is performed using the FastBDT algorithm [32]. A *Decision Tree* (DT) is a classifier that uses a series of cuts on the input features to hierarchically partition the data, as schematized in Fig 4.4.

From the root node a series of binary splits, based on conditions on the input variables, are applied. Each split uses the variable that at that node provides the best separation between signal and background when being cut on. This evaluation is performed computing for each feature at each node a cumulative probability histogram (CPH) for both signal and background. These histograms allow to determine the feature and cut-position that maximize the separation gain between signal and background. The same variable may thus be used at several nodes, while others might not be used at all. The final goal of this approach is to construct a partitioning of the feature space that is informative about the label of events in each partition. More complex trees (e.g. trees with many layers and many nodes) are able to perform a finer partitioning, maximizing the classification performances on the training dataset. However this approach has low out-of-sample performances and leads to over-fitting. For this reason, usually, a set of hyper-parameters such as the maximum number of consecutive cuts

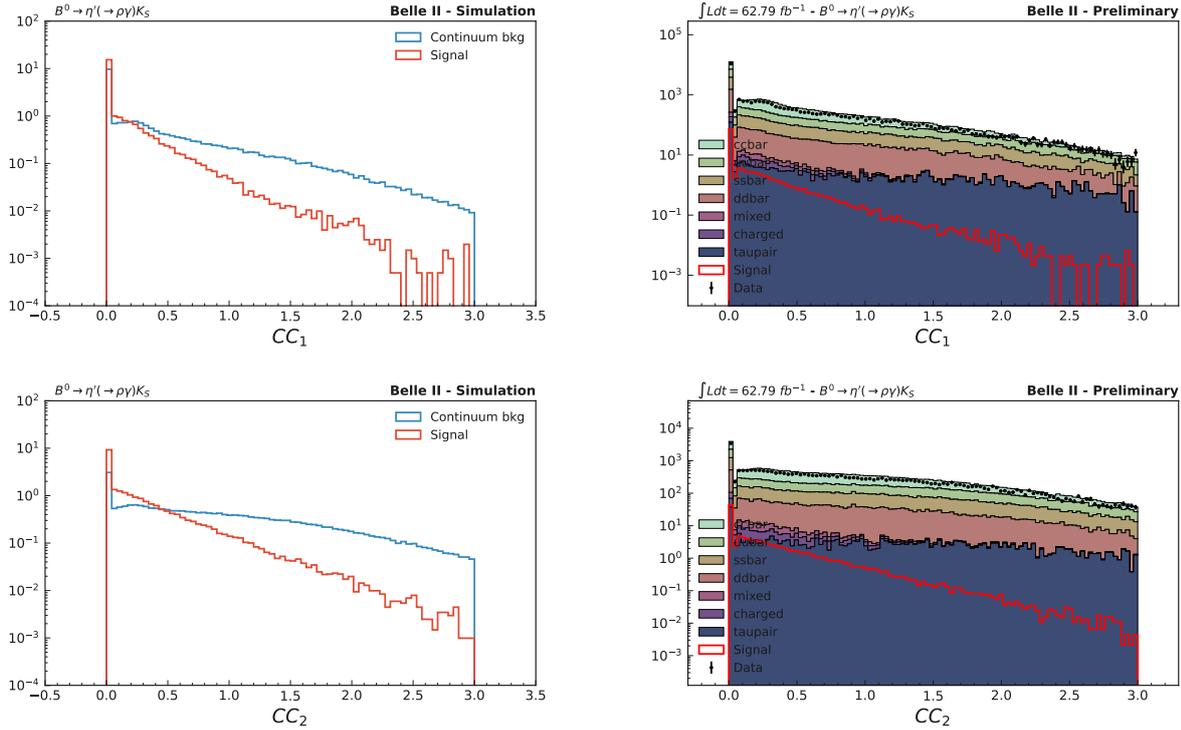


Figure 4.3: Distribution of the first two CLEO cones variables (CC) for signal and continuum Monte Carlo candidates, normalized to unit area (left) and data-Monte Carlo comparison (right).

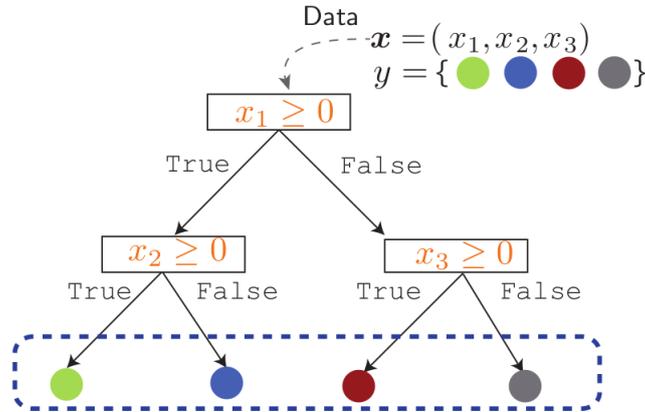


Figure 4.4: Schematic view of a Decision Tree [33].

(depth of the tree) is used to control the complexity and reduce the risk of over-fitting. Furthermore DTs have very high variance, they are sensible to many details in the training data and may be affected by finite-sample effects. For this reasons individual DTs are weak classifiers and they are usually combined in an ensemble method, such as *Boosted Decision Trees* (BDT).

The boosting procedure allows to construct a more robust classification model by sequentially constructing shallow DTs during the training phase, minimizing a negative binomial log-likelihood loss-function. Additional hyperparameters, such as the number of trees and the learning rate, also referred to as *shrinkage*, are needed to control the complexity of the algorithm.

**FastBDT** provides an optimized implementation of the BDT algorithm that allows to have a shorter execution time. More information on Boosted Decision Trees can be found in Ref. [33].

### 4.3 Training, validation and test

A sample of roughly  $1.4 \times 10^6$  labeled Monte Carlo events from all the final states considered in this study is used to train the classifier and assess its performances. Information on the Monte Carlo sample composition can be found in Fig. 4.5.

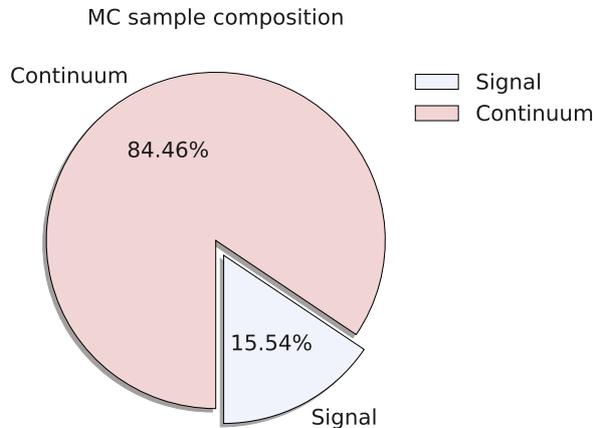


Figure 4.5: Fraction of signal and continuum candidates in the Monte Carlo sample used for continuum suppression

The MC sample is split into 50% training, 30% validation and 20% test sample. The training sample is used to train the algorithm and define its internal parameters. The classifier hyperparameters are the default ones given in Table 4.1: no optimization has been performed.

Hyperparameter	Default value
number of trees	100
depth of the tree	3
shrinkage	0.1
subsample	0.5

Table 4.1: Default hyperparameters configuration for the **FastBDT** algorithm

In Fig. 4.6 the **FastBDT** output variable ( $CS_{var}$ ) distribution for signal and continuum training candidates is shown. The classifier provides a good separation between the two categories.

Using the internal feature ranking provided by **FastBDT** it is possible to evaluate the discriminating power of each input variable. The Intern Feature Importance is computed using the sum of the information gains provided by each feature in all decision trees (Fig. 4.7).

The variable  $|\cos(\theta_B - \theta_O)|$  is the most discriminating, while most of the variables have very low Feature Importance and could be removed without any loss in performances.

After the training the model is applied to the validation sample to evaluate out-of-sample performances and check for overfitting. Finally the independent test sample is used to assess the performances of the final model. The performances of the **FastBDT** classifier are shown in Fig. 4.8 and Fig. 4.9.

The *Receiver Operating Characteristic curves* (or ROC curves) for training, validation and test samples are in good agreement and the *Area Under the Curve* (or AUC) is  $\sim 0.93$  for all the samples. The distribution of  $CS_{var}$  for training and test samples are consistent for all the decay channels considered in this analysis.

No selection is applied to the continuum suppression variable:  $CS_{var}$  is included in the Unbinned

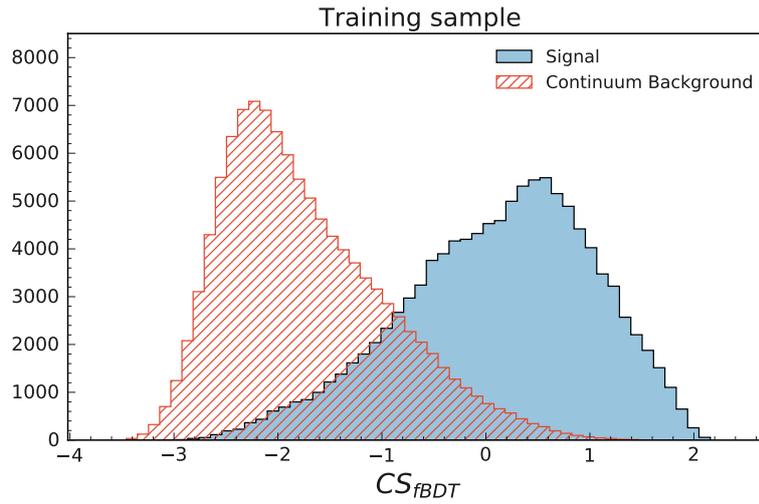


Figure 4.6: Distribution of the FastBDT output variable for signal and continuum training candidates. The histograms are normalized to have the same area.

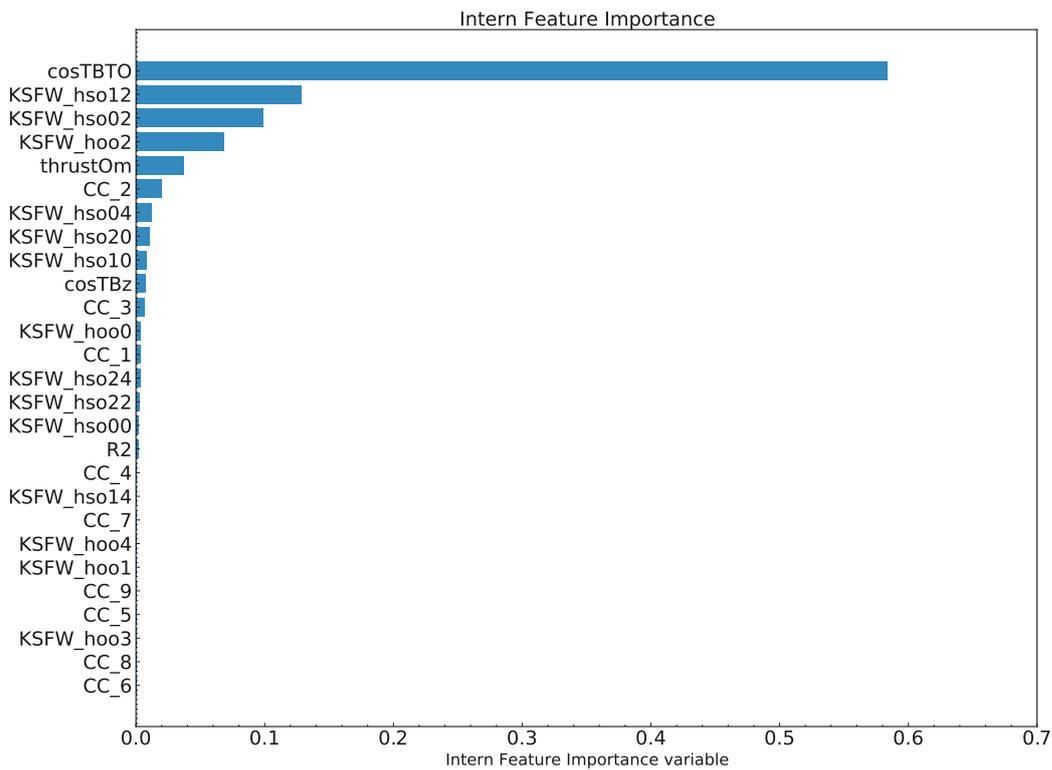


Figure 4.7: Intern feature importance of the training variables.

Maximum Likelihood fit for the signal extraction. In this way it is possible to avoid any further efficiency loss for signal.

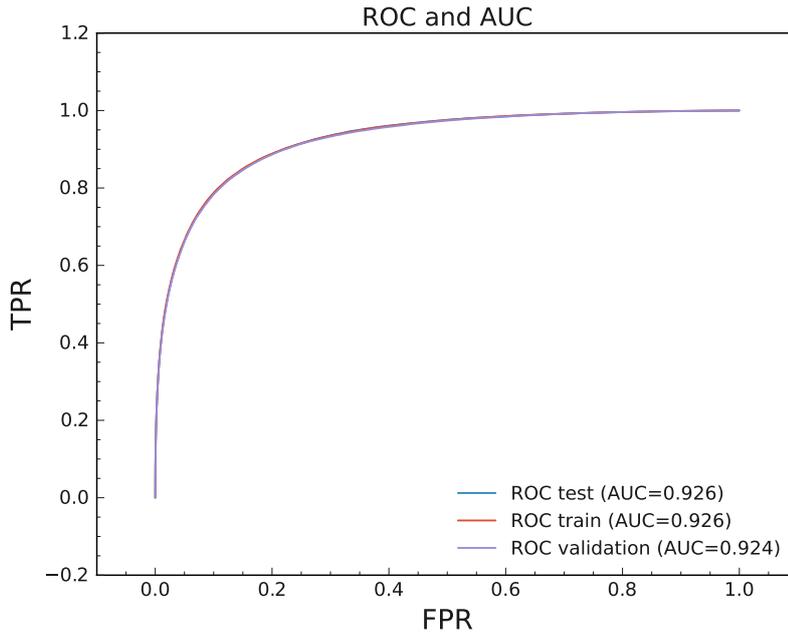


Figure 4.8: ROC curve and AUC for training, validation and test samples.

#### 4.4 Validation with the off-resonance data

The off-resonance dataset is used to validate the continuum suppression variable. In Fig. 4.10, for each final state, a comparison between the MC and off-resonance distribution of  $CS_{var}$  is given.

Overall the agreement between data and MC is good, however some deviation and trend can be seen in all channels. This is due to the Monte Carlo mismodelling of the continuum suppression training feature presented in Section 4.1. This issue is covered by a dedicated systematic uncertainty. Furthermore a Monte Carlo multivariate reweighting procedure has been studied to reduce the impact of Monte Carlo mismodelling on the final results.

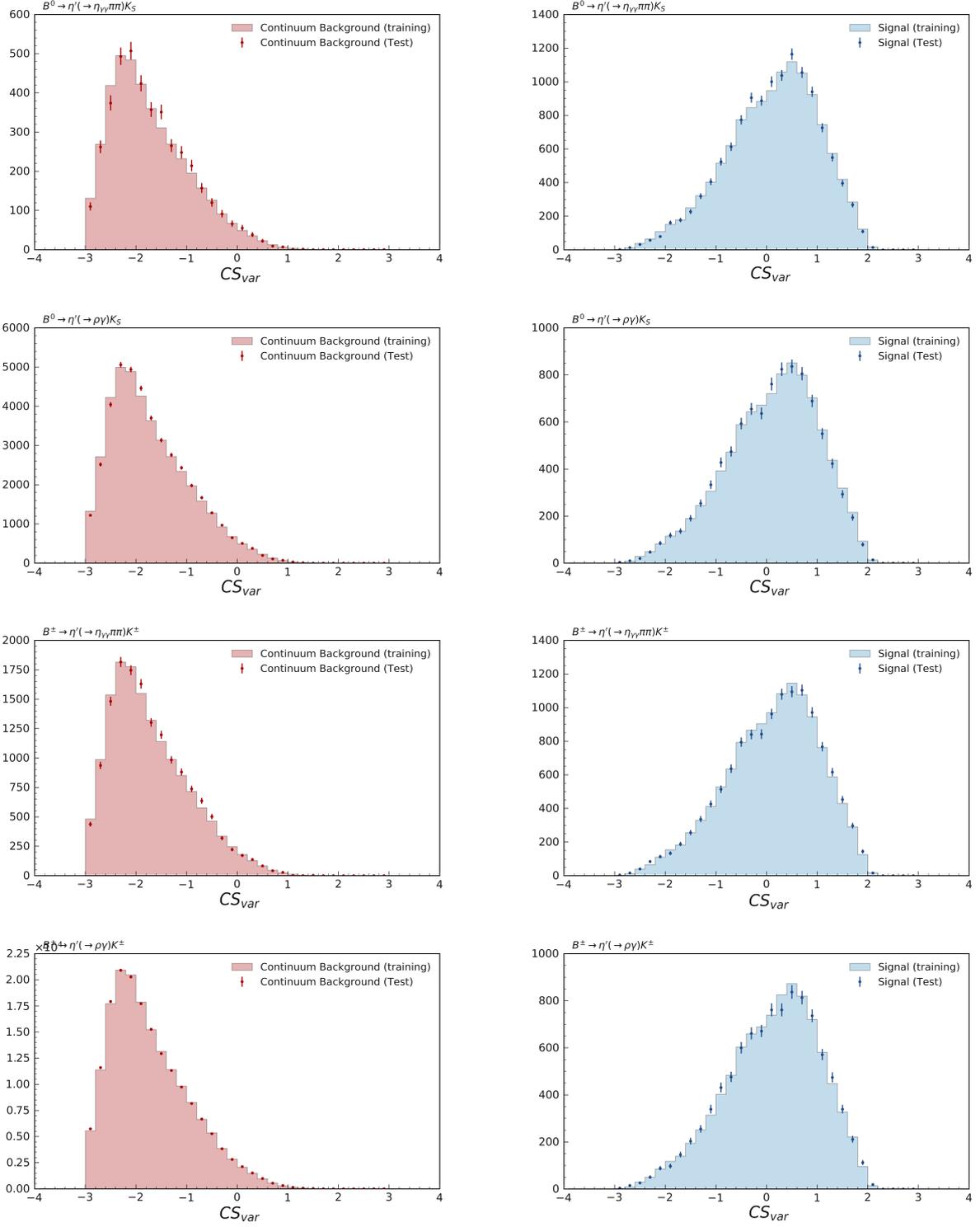


Figure 4.9:  $CS_{var}$  variable distribution for training sample and for each final state test sample (from top to bottom:  $B^0 \rightarrow \eta'(\eta_{\gamma\gamma}\pi^+\pi^-)K_S^0$ ,  $B^0 \rightarrow \eta'(\rho_{\pi^+\pi^-}\gamma)K_S^0$ ,  $B^\pm \rightarrow \eta'(\eta_{\gamma\gamma}\pi^+\pi^-)K^\pm$  and  $B^\pm \rightarrow \eta'(\rho_{\pi^+\pi^-}\gamma)K^\pm$ ) for the continuum background (left) and Signal (right). The distributions are normalized to the same area.

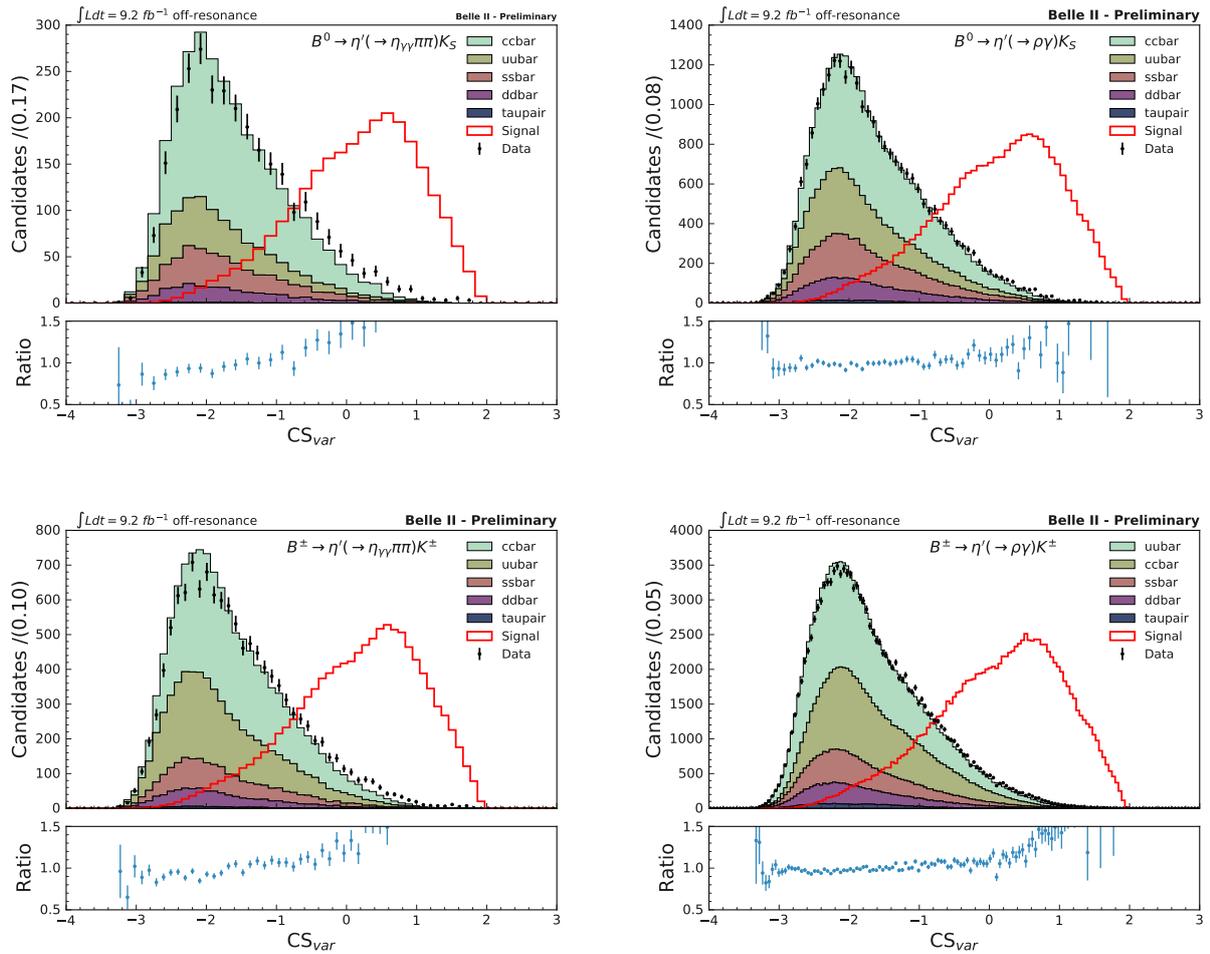


Figure 4.10:  $CS_{var}$  distribution for Monte Carlo and off-resonance data.

## 4.5 Multivariate Monte Carlo reweighting

The general idea of Multivariate Monte Carlo reweighting is discussed in Ref. [34]. The goal of this procedure is to determine weights in order to emphasize those events in the MC sample which have high probability to be found in data. It is possible to determine these weights performing a training using real data as target and Monte Carlo as background. If, ideally, there were no discrepancies between Monte Carlo and data distributions, the algorithm would learn nothing, everything it learns is due to mismodelling in MC. This method allows to derive individual per event weights that can outweigh differences in all the input variables simultaneously.

For this study the off-resonance data sample has been used. As discussed in section 3.4 this data, collected  $\sim 60$  MeV below the  $\Upsilon(4S)$  resonance, are a sample of pure continuum candidates. Training a **FastBDT** classifier using the off-resonance data as target and the continuum Monte Carlo as background, the algorithm should be able to learn the difference between the two samples, due to Monte Carlo mismodelling. The training of the algorithm is performed using the entire off-resonance dataset and 10% of the continuum Monte Carlo sample, splitting this sample into 90% training and 10% test sample. The input features are those used for the built of the continuum suppression discriminator, introduced in section 4.1. In figure 4.11 the Monte Carlo and off-resonance data distribution of the output variable of the algorithm is shown for the training and test samples. The small shift between the two distributions is related to the Monte Carlo mismodelling.

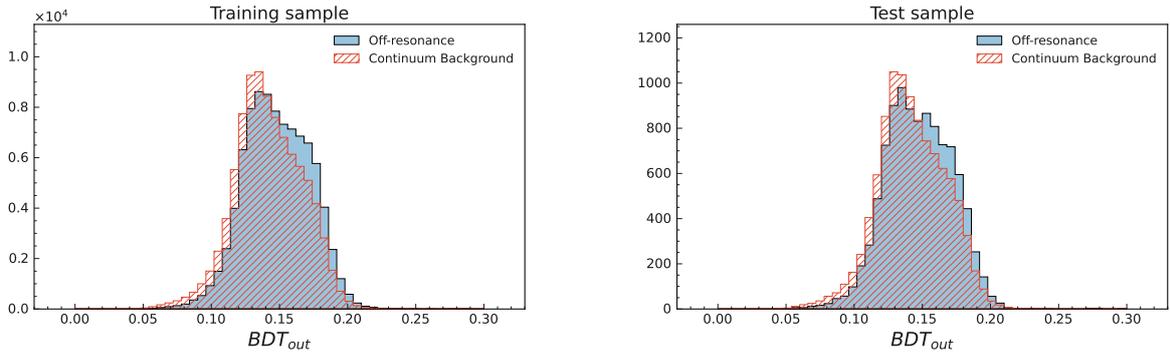


Figure 4.11: Output variable distribution of the reweighting algorithm for off-resonance data and continuum Monte Carlo for the training (left) and test (right) samples in  $B^\pm \rightarrow \eta'(\rho\gamma)K^\pm$  decays.

The algorithm is then applied to MC events and the discriminator output variable ( $BDT_{out}$ ) is transformed into an event weight:

$$p = BDT_{out} = \frac{data}{data + MC} \quad (4.3)$$

$$w = \frac{data}{MC} = \frac{BDT_{out}}{1 - BDT_{out}} \quad (4.4)$$

A separate training has been performed for each decay channel.

In Figure 4.12 the histogram reweighting results for some continuum suppression variables are shown for  $B^\pm \rightarrow \eta'(\rho_{\pi^+\pi^-}\gamma)K^\pm$  decay channel.

In order to test the effectiveness of the reweighting procedure on an independent data sample, since the model was trained targeting the off-resonance dataset, the continuum suppression variable distribution is shown in figure 4.13 for data in the control region (or side band, defined in section 5.1) and the reweighted continuum Monte Carlo.

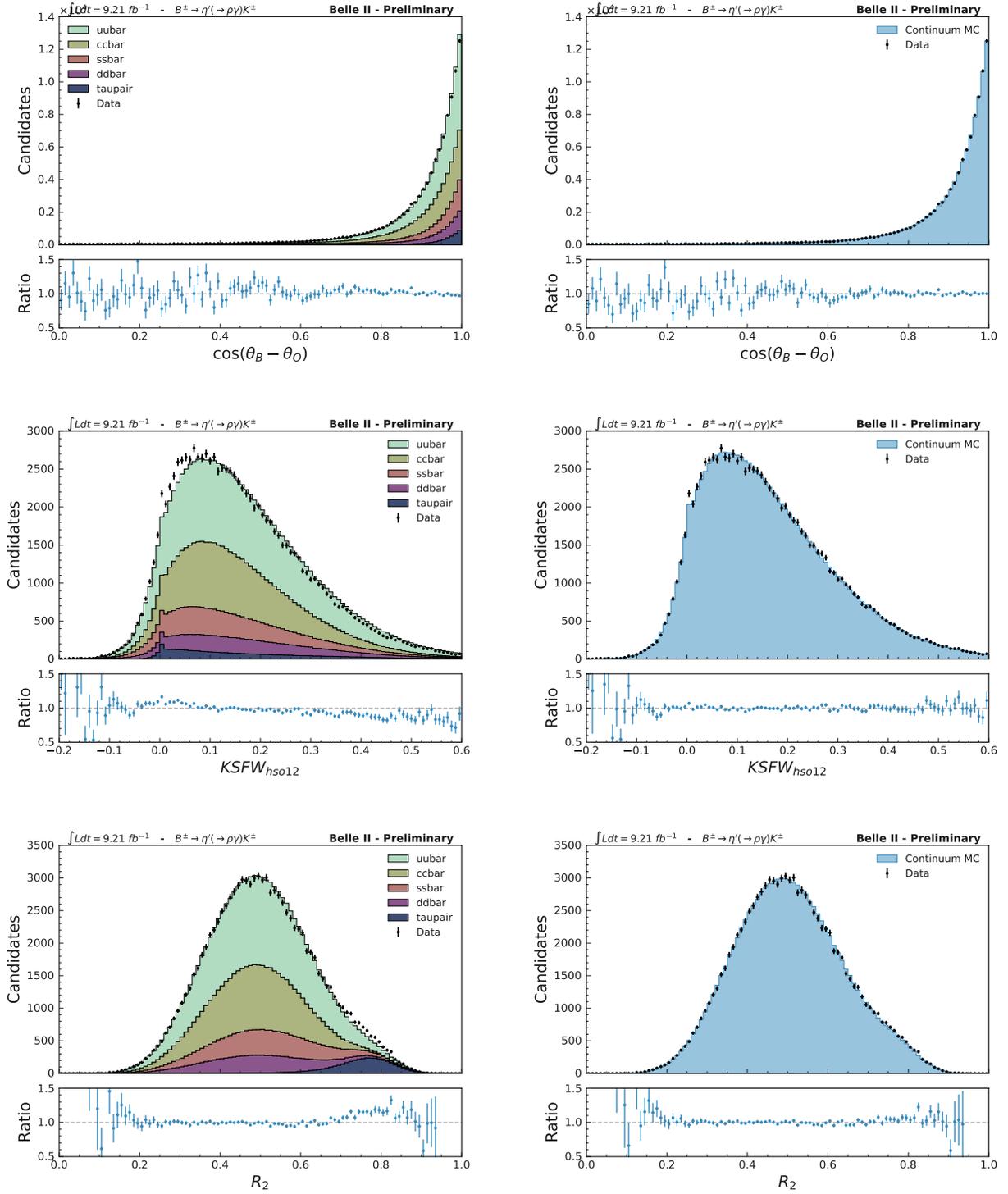


Figure 4.12: Monte Carlo and off-resonance data distribution of (from top to bottom)  $|\cos(\theta_B - \theta_O)|$ ,  $H_{12}^{so}$  and  $R_2$ . The distributions are shown before the reweighting (left) and after the reweighting (right).

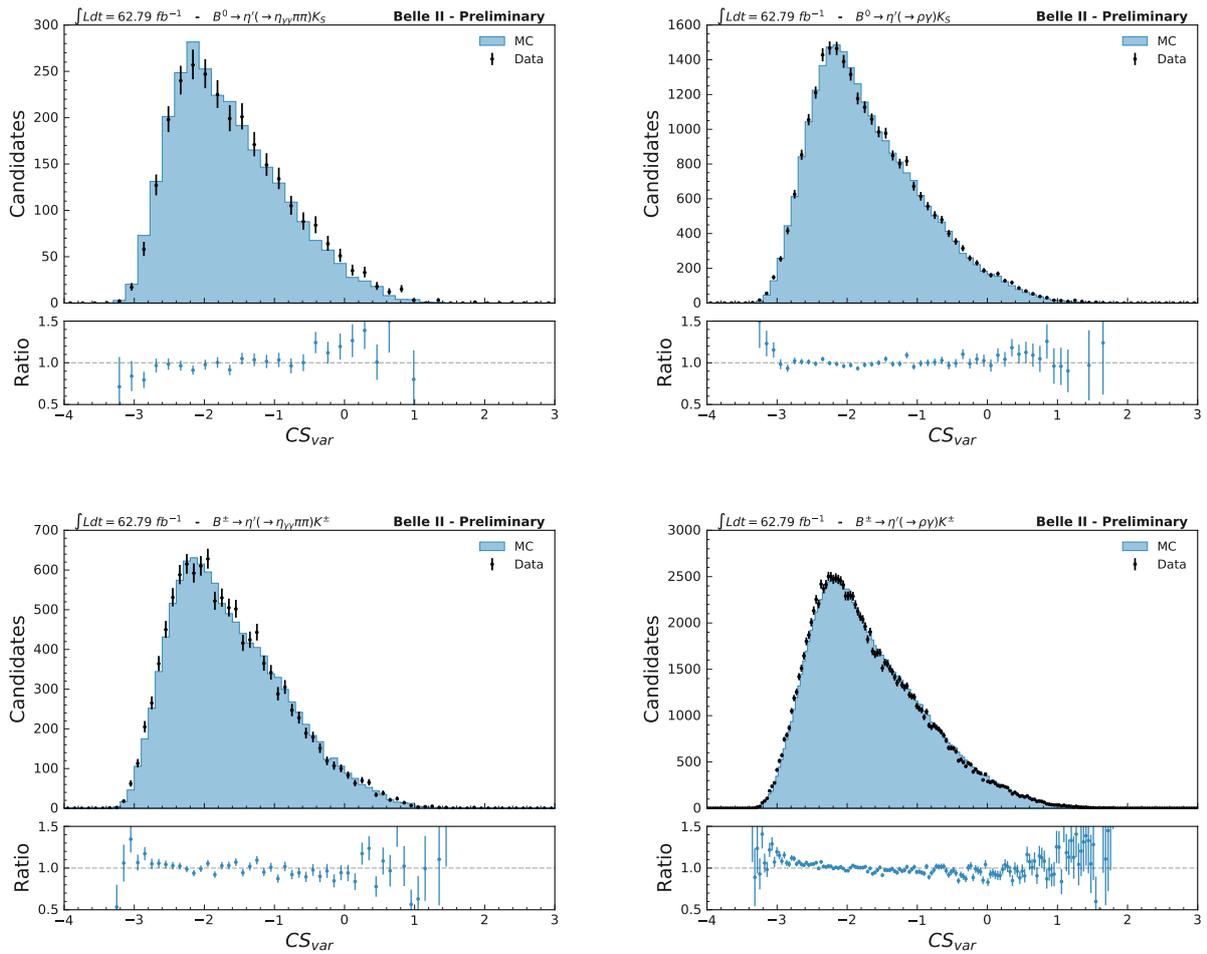


Figure 4.13:  $CS_{var}$  distribution for reweighted Monte Carlo and data in the control region. Since in the control region there are also peaking candidates, the mean value of the continuum weights has been used as a scale factor for the peaking background.

# Chapter 5

## Branching Fractions

### 5.1 Fit procedure for signal extraction

In order to measure the branching fractions for the studied processes, the number of signal candidates has been extracted through an extended unbinned maximum likelihood fit of the observables  $M_{bc}$ ,  $\Delta E$  and  $CS_{var}$  for the components signal, continuum background, peaking background and SxF, described in section 3.2. In Fig. 5.1 the Monte Carlo distributions of the fit variables for each component are shown. The likelihood  $\mathcal{L}_i$  of the  $i^{th}$  event, with set of observables  $\vec{x}_i$  is defined as

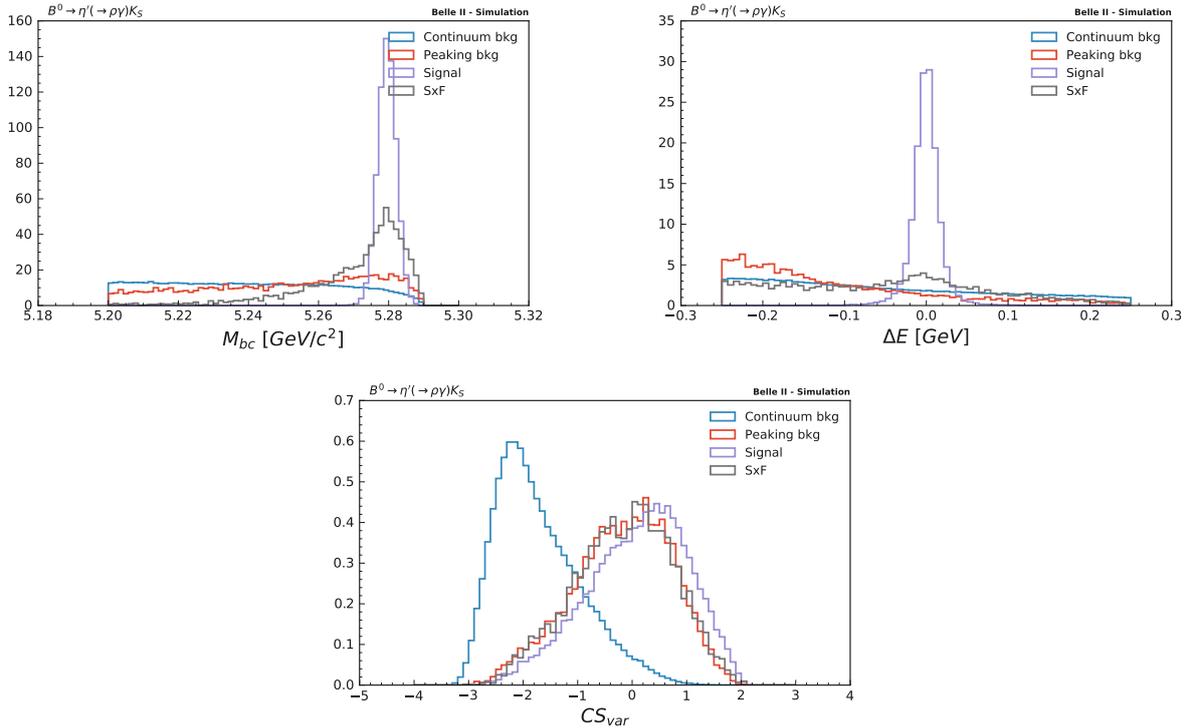


Figure 5.1: Fit variables distributions for signal, SxF, continuum background and peaking background components

$$\mathcal{L}_i = \sum_j^m n_j \mathcal{P}_j(\vec{x}_j) \quad (5.1)$$

where  $\mathcal{P}_j$  is the probability computed for the input observables  $\vec{x}_i$  and  $n_j$  is the number of events in the dataset. The low correlations between the input observables are neglected (fig. 5.2) and the

probabilities  $\mathcal{P}_j$  are assumed to be the product of the 1-dimensional probability density functions of the input variables.

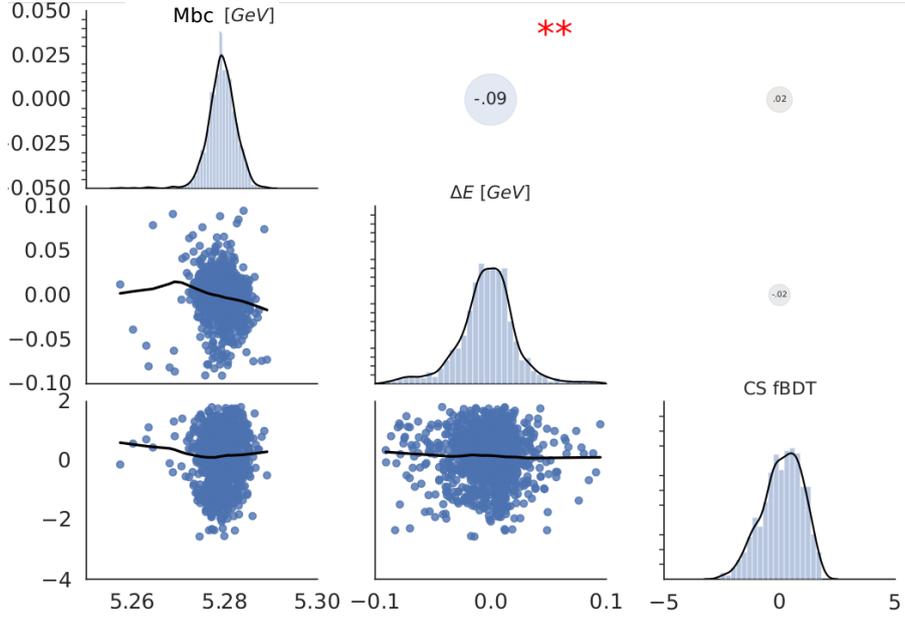


Figure 5.2: Correlation among the three fitting variables  $CS_{var}$ ,  $M_{bc}$ , and  $\Delta E$  for  $B^0 \rightarrow \eta'(\eta_{\gamma\gamma}\pi^+\pi^-)K_S^0$  signal Monte Carlo. The number in the upper diagonal represents the correlation, the lines are KDE representations of the distribution and the lower diagonal has the scatter plot.

For a dataset of  $N$  events, where  $N$  fluctuates according to the Poisson statistics, the likelihood is

$$\mathcal{L}(N, \vec{x}_i) = \frac{e^{-\sum n_j}}{N!} \prod_i^N \mathcal{L}_i \quad (5.2)$$

Two separate regions have been defined: the signal region, including candidates with  $M_{bc} > 5.27$  GeV/ $c^2$  and  $-0.07 < \Delta E < 0.05$  GeV and the side band, defined as the complementary region up to  $M_{bc} > 5.2$  GeV/ $c^2$  and  $|\Delta E| < 0.25$  GeV. The agreement of the fit variables distributions between Monte Carlo and data has been checked keeping the signal region hidden. In fig. 5.3, 5.4, 5.5 and 5.6 the side band distribution of the fit variables for data and Monte Carlo are shown.

The Monte Carlo distributions are in good agreement with the data (apart from a difference in the normalization in some of the channels), hence many parameters of the probability density function (PDF) shapes of the observables, for each component, are obtained from a fit to Monte Carlo samples. Since the expected number of SxF candidates is low and their  $M_{bc}$ ,  $\Delta E$  and  $CS_{var}$  distributions are very similar to the signal ones, signal and SxF are fitted together, fixing their relative ratio from Monte Carlo. The contribution to the systematic uncertainty of this choice is discussed in section 5.3.1.

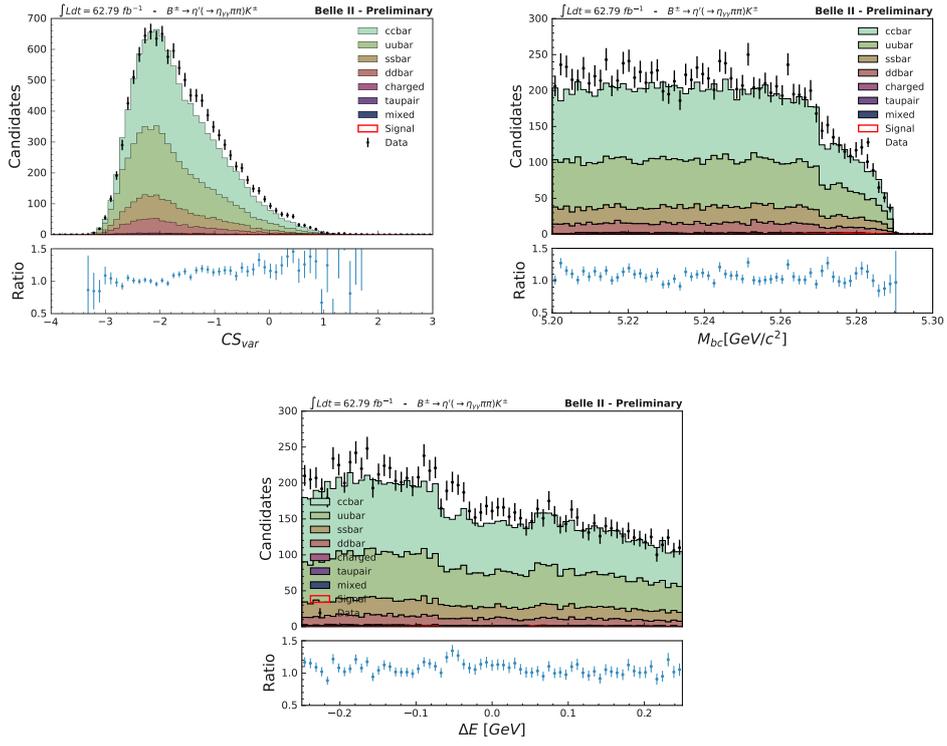


Figure 5.3: Distribution of the fit variables  $CS_{var}$ ,  $M_{bc}$ , and  $\Delta E$  in the side band for data and Monte Carlo, normalized to data integrated luminosity, for the decay channel  $B^\pm \rightarrow \eta'(\eta_{\gamma\gamma}\pi^+\pi^-)K^\pm$

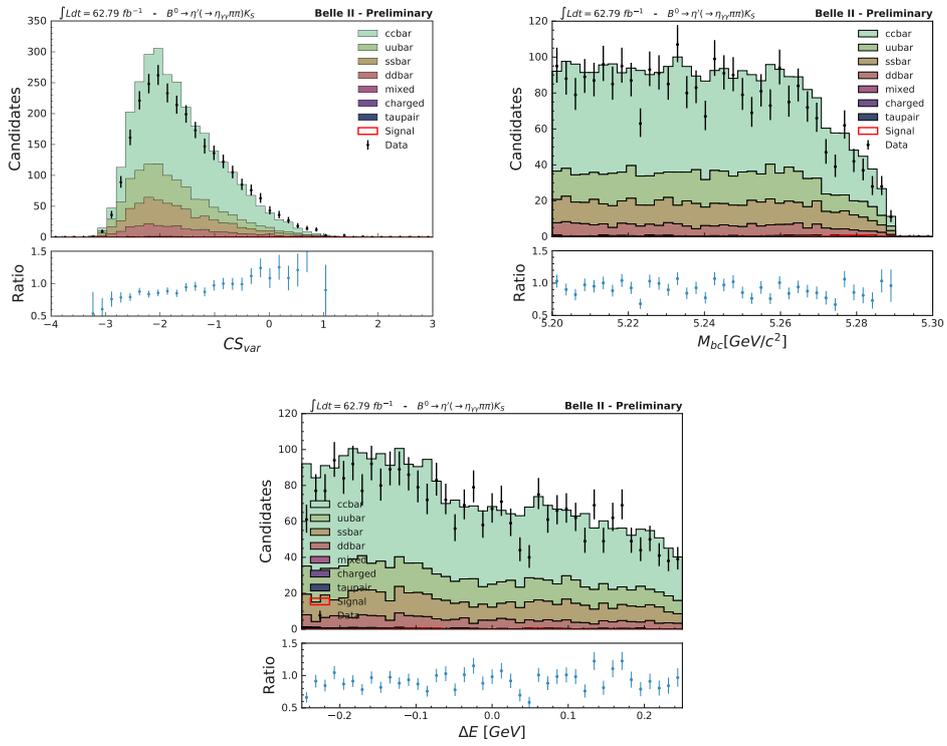


Figure 5.4: Distribution of the fit variables  $CS_{var}$ ,  $M_{bc}$ , and  $\Delta E$  in the side band for data and Monte Carlo, normalized to data integrated luminosity, for the decay channel  $B^0 \rightarrow \eta'(\eta_{\gamma\gamma}\pi^+\pi^-)K_S^0$

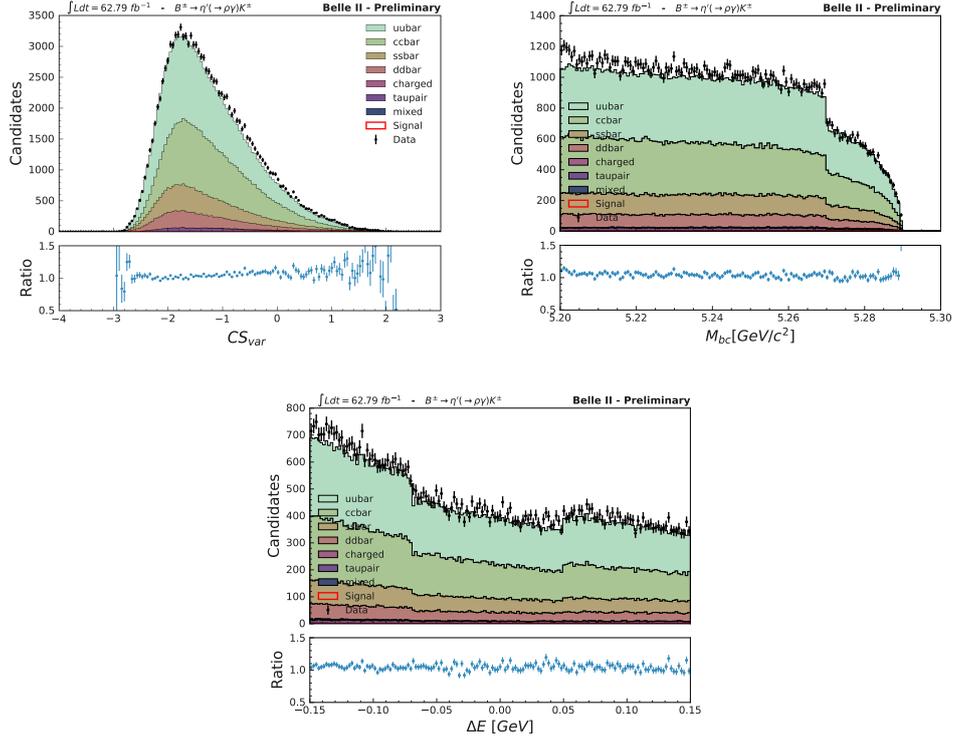


Figure 5.5: Distribution of the fit variables  $CS_{var}$ ,  $M_{bc}$ , and  $\Delta E$  in the side band for data and Monte Carlo, normalized to data integrated luminosity, for the decay channel  $B^\pm \rightarrow \eta'(\rho_{\pi^+\pi^-}\gamma)K^\pm$

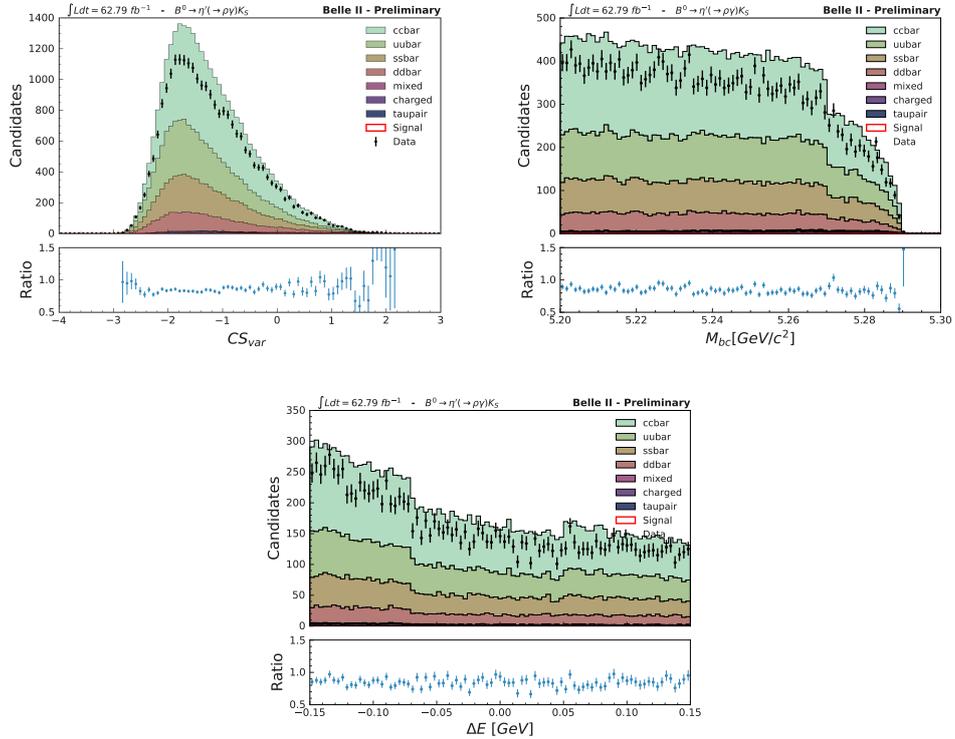


Figure 5.6: Distribution of the fit variables  $CS_{var}$ ,  $M_{bc}$ , and  $\Delta E$  in the side band for data and Monte Carlo, normalized to data integrated luminosity, for the decay channel  $B^0 \rightarrow \eta'(\rho_{\pi^+\pi^-}\gamma)K_S^0$

### 5.1.1 Fitting tools

The fit has been performed using the RooFit [35] toolkit, using RooRarFit [36] interface. Since the time-dependent  $CP$  violation study uses a different software, zfit [37], based on Tensorflow, an

attempt at building the UML fit using also this toolkit is presented in appendix A.

In both cases the minimization is based on `Minuit` [38] (or its python equivalent `iminuit`), using the `MIGRAD` minimizer. The error computation is performed with the `MINOS` algorithm. This algorithm gives approximate confidence intervals performing a scan of the likelihood along one parameter, while maximizing the likelihood with respect to all the other parameters. This approach allows to construct, for each parameter  $\theta_i$ , the negative log-likelihood  $-\ln\mathcal{L}(\theta_i, \hat{\theta}_k(\theta_i))$  for  $k \neq i$ , where  $\hat{\theta}_k$  is the estimate for  $\theta_k$ . The interval is defined from the values at which  $-\ln\mathcal{L}$  increases of 0.5 with respect to the minimum.

## 5.2 Fit validation and results

### 5.2.1 Extraction of the PDFs from Monte Carlo

The first step of the fit is the extraction of the PDFs parameters from Monte Carlo. In table 5.1 and table 5.2 the PDF shapes used for each variable and each category can be found, where `Gauss(N)` is a sum of N gaussian functions, `CB` is a Crystal Ball function [39], `Argus` is an Argus function [40], `pol(N)` is a Chebyshev polynomial of degree N and `BG` indicates a Bifurcated Gaussian.

	Signal	SxF	$q\bar{q}$	$B\bar{B}$
$M_{bc}$	<code>Gauss(2)</code>	<code>CB</code>	<code>Argus</code>	<code>Argus+Gauss(1)</code>
$\Delta E$	<code>Gauss(2)</code>	<code>CB+pol(2)</code>	<code>pol(3)/pol(2)</code>	<code>pol(2)</code>
$CS_{var}$	<code>BG</code>	<code>BG</code>	<code>BG</code>	<code>BG</code>

Table 5.1: PDF shapes for  $B \rightarrow \eta'(\eta\pi^+\pi^-)K$ . When two functions are indicated, the former is for the charged final state and the latter for the neutral one

	Signal	SxF	$q\bar{q}$	$B\bar{B}$
$M_{bc}$	<code>Gauss(2)</code>	<code>CB</code>	<code>Argus</code>	<code>Argus+CB/Argus+Gauss(1)</code>
$\Delta E$	<code>Gauss(2)</code>	<code>Gauss(1)+pol(2)/Gauss(1)+pol(1)</code>	<code>pol(2)</code>	<code>pol(2)</code>
$CS_{var}$	<code>BG</code>	<code>BG</code>	<code>BG</code>	<code>BG</code>

Table 5.2: PDF shapes for  $B \rightarrow \eta'(\rho\gamma)K$ . When two functions are indicated, the former is for the charged final state and the latter for the neutral one

In Fig. 5.7, 5.8, 5.9 and 5.10 the results of the PDFs fits to Monte Carlo samples are shown for all the final states considered in this analysis.

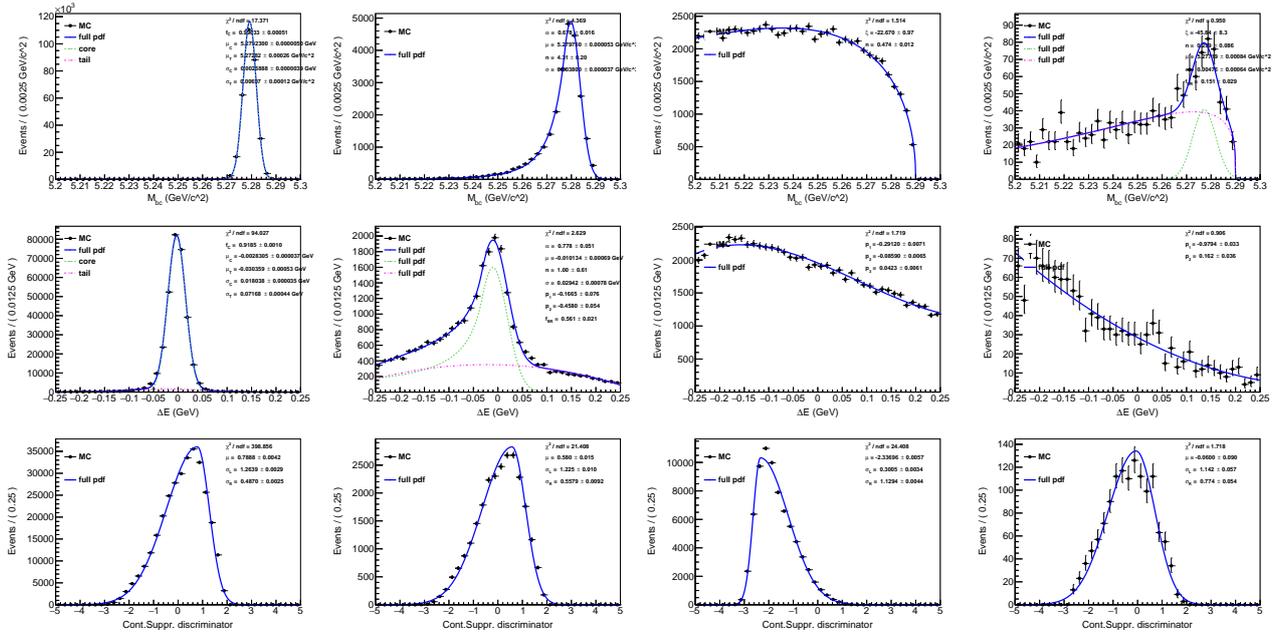


Figure 5.7: Distribution of the fit variables  $M_{bc}$  and  $\Delta E$ , and  $CS_{var}$  (top to bottom) for signal, SxF, continuum background, and peaking background (left to right) and fitted PDF as described in table 5.1 for channel  $B^\pm \rightarrow \eta'(\eta\pi^+\pi^-)K^\pm$

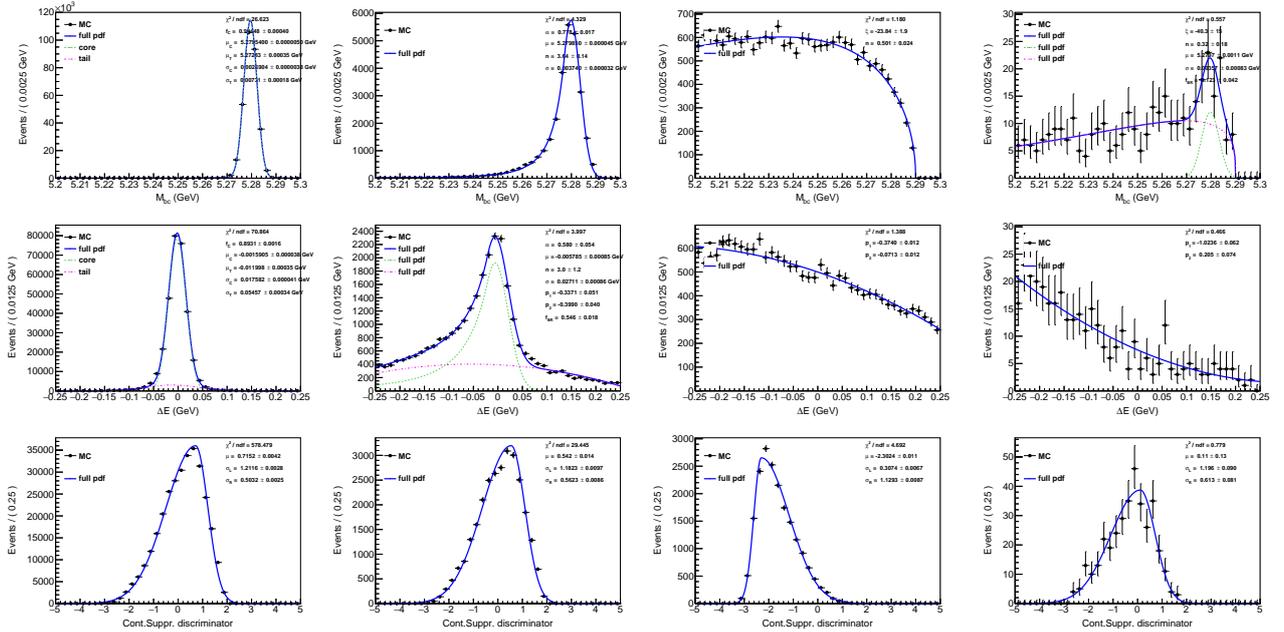


Figure 5.8: Distribution of the fit variables  $M_{bc}$  and  $\Delta E$ , and  $CS_{var}$  (top to bottom) for signal, SxF, continuum background, and peaking background (left to right) and fitted PDF as described in table 5.1 for channel  $B^0 \rightarrow \eta'(\eta\pi^+\pi^-)K_S^0$

Most of the parameters of the PDFs are fixed from this fit to Monte Carlo. The only parameters left floating in the final fit are the core parameters of the  $M_{bc}$  and  $\Delta E$  distributions for the signal component, the slope of the *Argus* function modelling the  $M_{bc}$  distribution for continuum background and the coefficients of the polynomial modelling the continuum  $\Delta E$  distribution.

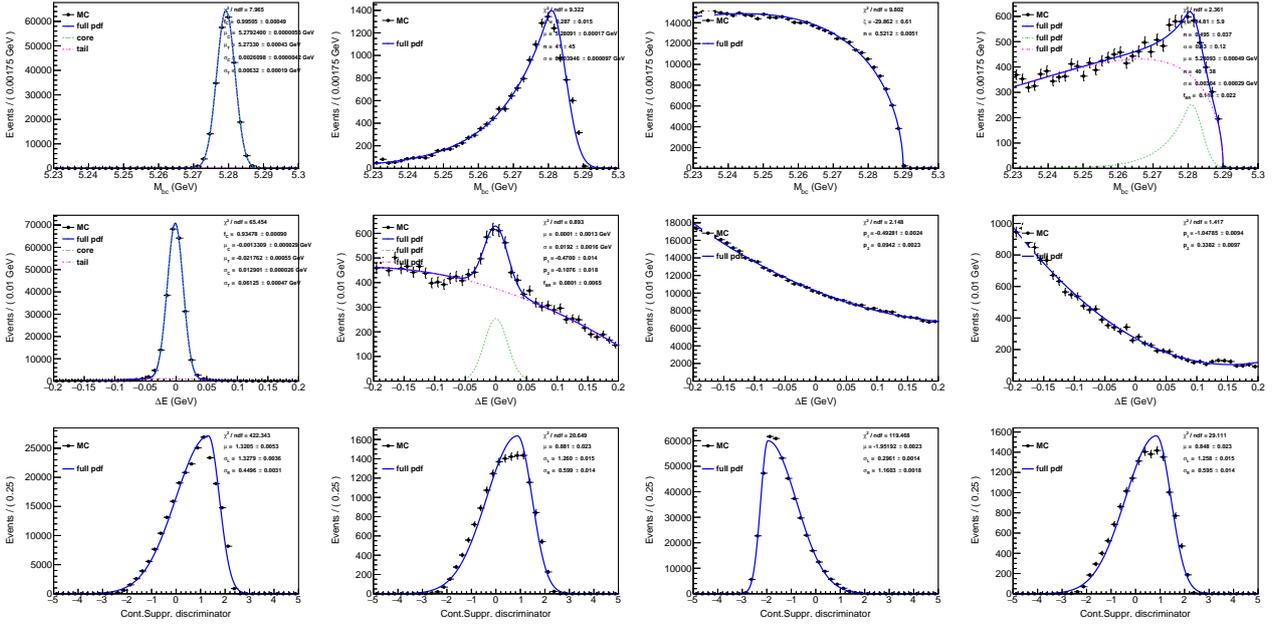


Figure 5.9: Distribution of the fit variables  $M_{bc}$  and  $\Delta E$ , and  $CS_{var}$  (top to bottom) for signal, SxF, continuum background, and peaking background (left to right) and fitted PDF as described in table 5.2 for channel  $B^\pm \rightarrow \eta'(\rho\gamma)K^\pm$

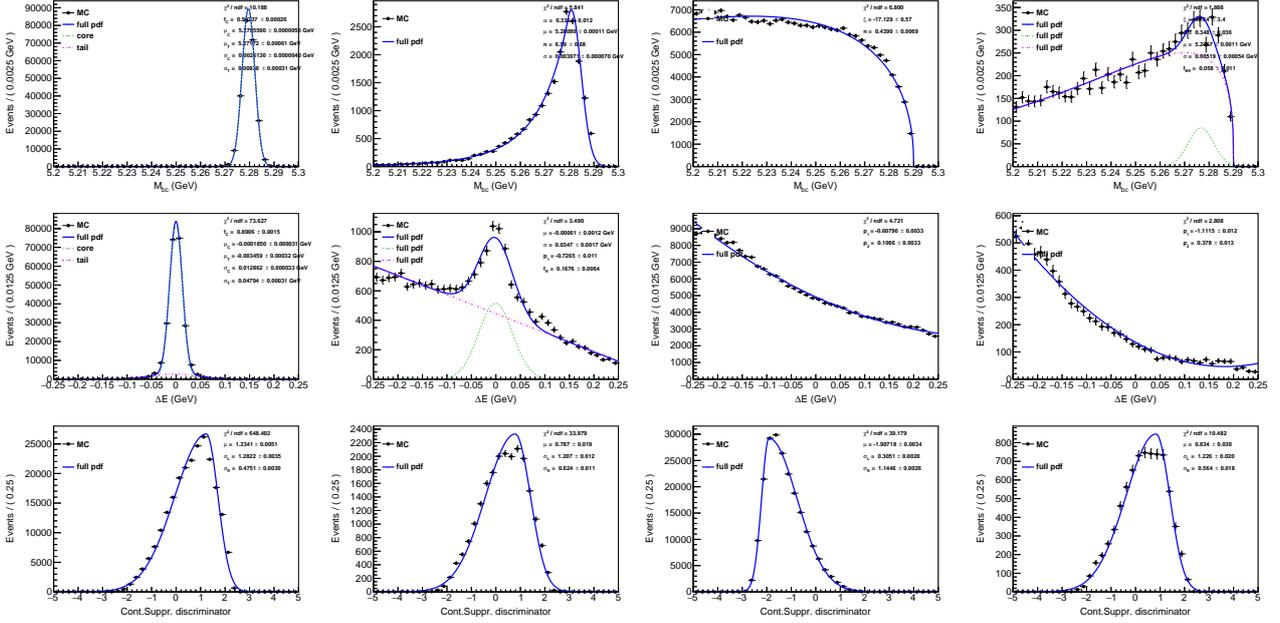


Figure 5.10: Distribution of the fit variables  $M_{bc}$  and  $\Delta E$ , and  $CS_{var}$  (top to bottom) for signal, SxF, continuum background, and peaking background (left to right) and fitted PDF as described in table 5.2 for channel  $B^0 \rightarrow \eta'(\rho\gamma)K^0$

### 5.2.2 Test of the fit procedure

The fit procedure has been tested using toy Monte Carlo samples. The samples for the continuum background have been generated according to its PDFs, while the signal and peaking candidates have been sampled from the large  $B\bar{B}$  Monte Carlo set.

The number of continuum and peaking background candidates has been fixed normalizing the Monte Carlo equivalent integrated luminosity to the luminosity of the data sample; the expected number of candidates for signal and background categories can be found in table 5.3. The linearity of the fit has

Decay channel	$N_{\text{sig}}$	$N_{\text{cont}}$	$N_{\text{peak}}$
$B^0 \rightarrow \eta'(\eta_{\gamma\gamma}\pi^+\pi^-)K_S^0$	$85.82 \pm 0.01$	$3100 \pm 20$	$13.0 \pm 0.6$
$B^\pm \rightarrow \eta'(\eta_{\gamma\gamma}\pi^+\pi^-)K^\pm$	$274.64 \pm 0.06$	$11040 \pm 30$	$42 \pm 1$
$B^0 \rightarrow \eta'(\rho_{\pi^+\pi^-\gamma})K_S^0$	$111.92 \pm 0.02$	$30310 \pm 60$	$242 \pm 2$
$B^\pm \rightarrow \eta'(\rho_{\pi^+\pi^-\gamma})K^\pm$	$353.1 \pm 0.1$	$121800 \pm 100$	$865 \pm 5$

Table 5.3: Expected number of candidates for each component and each decay channel with  $\mathcal{L} = 62.8 \text{ fb}^{-1}$ . The uncertainties are statistical only.

been tested varying the number of injected signal candidates in the toy Monte Carlo sample. For each value of the number of injected signal candidates, 1000 toy Monte Carlo samples have been generated and the average yield has been computed. The results of the linearity test are shown in figure 5.11. No significant bias has been observed, except for the  $B^\pm \rightarrow \eta'(\rho\gamma)K^\pm$  decay mode, where the signal yield is underestimated of approximately  $1\sigma$ .

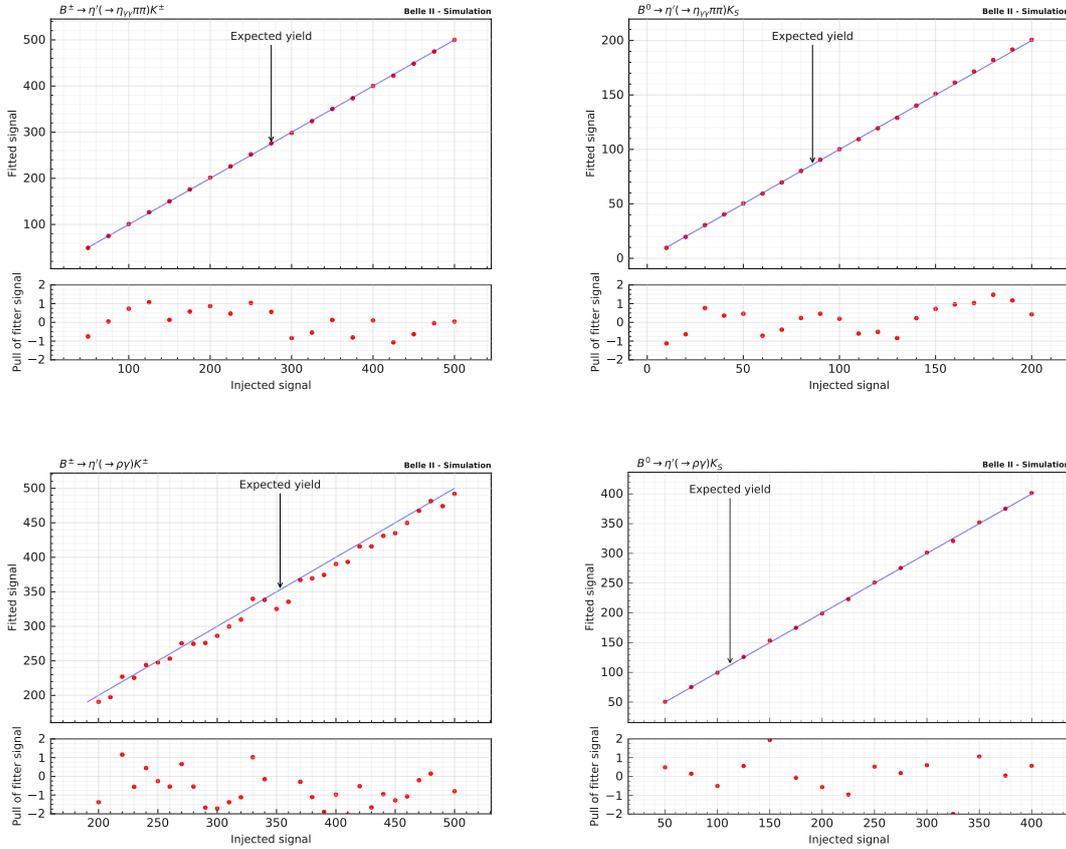
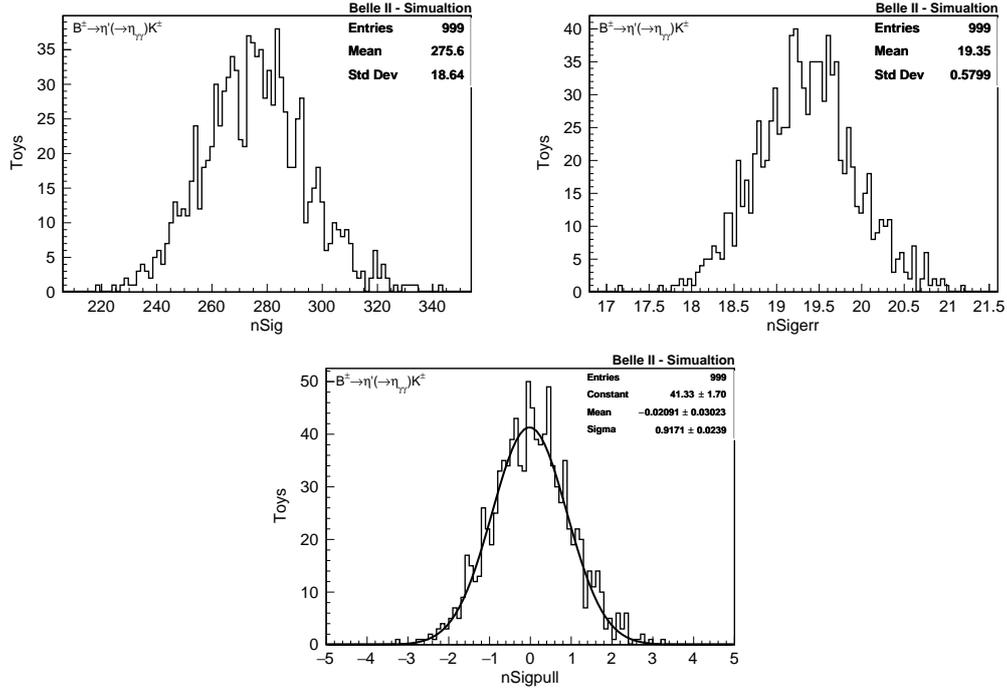
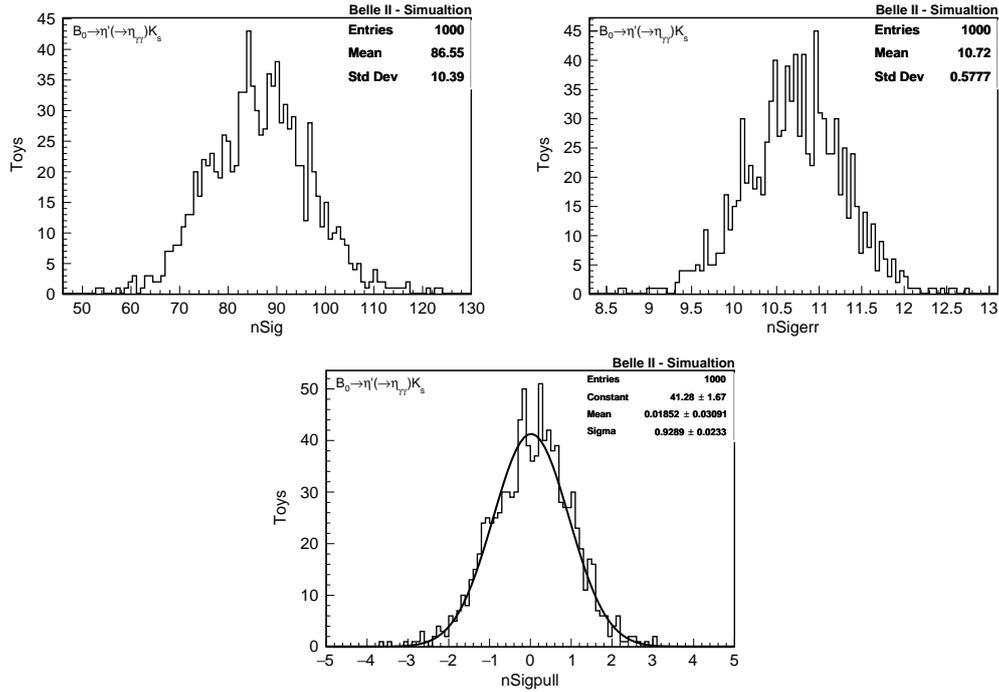


Figure 5.11: Linearity test: average fit results for the signal yield as a function of the injected signal and relative pulls. The results are shown for the  $B \rightarrow \eta'(\eta_{\gamma\gamma}\pi^+\pi^-)K$  decay (top) and for the  $B \rightarrow \eta'(\rho\gamma)K$  (bottom), both for charged (left) and neutral (right) modes.

Another test has been performed fitting 1000 toy Monte Carlo samples fixing the yield for each category to the expected ones, predicted from Monte Carlo, given in table 5.3. The results of this test for the signal yield are shown in fig. 5.12, 5.13, 5.14 and 5.15, while the distributions for the backgrounds yield can be found in appendix B.

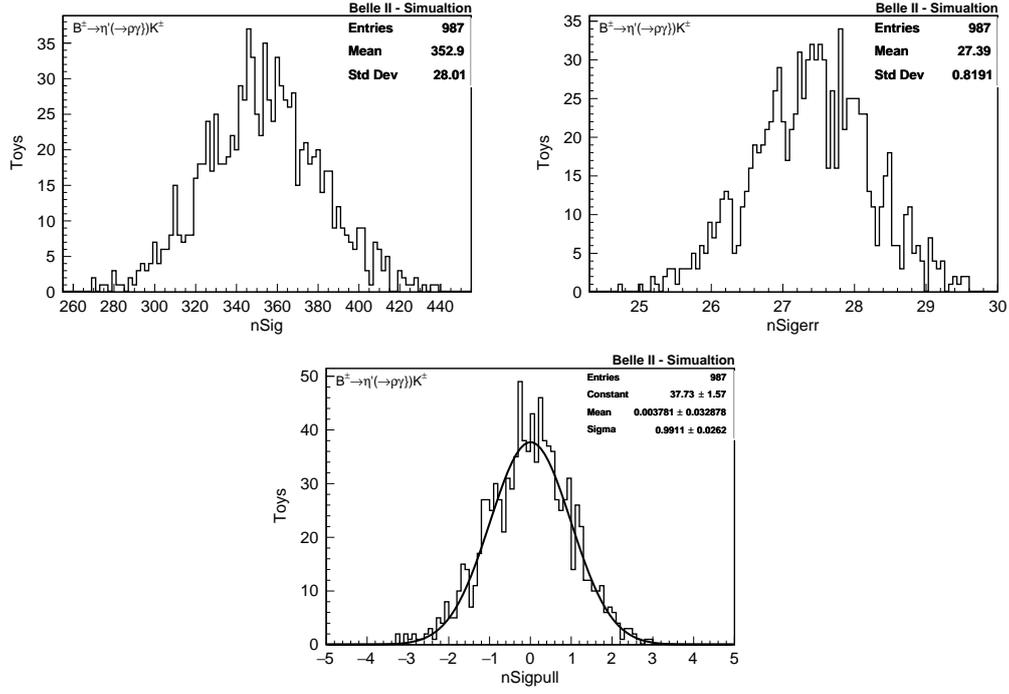
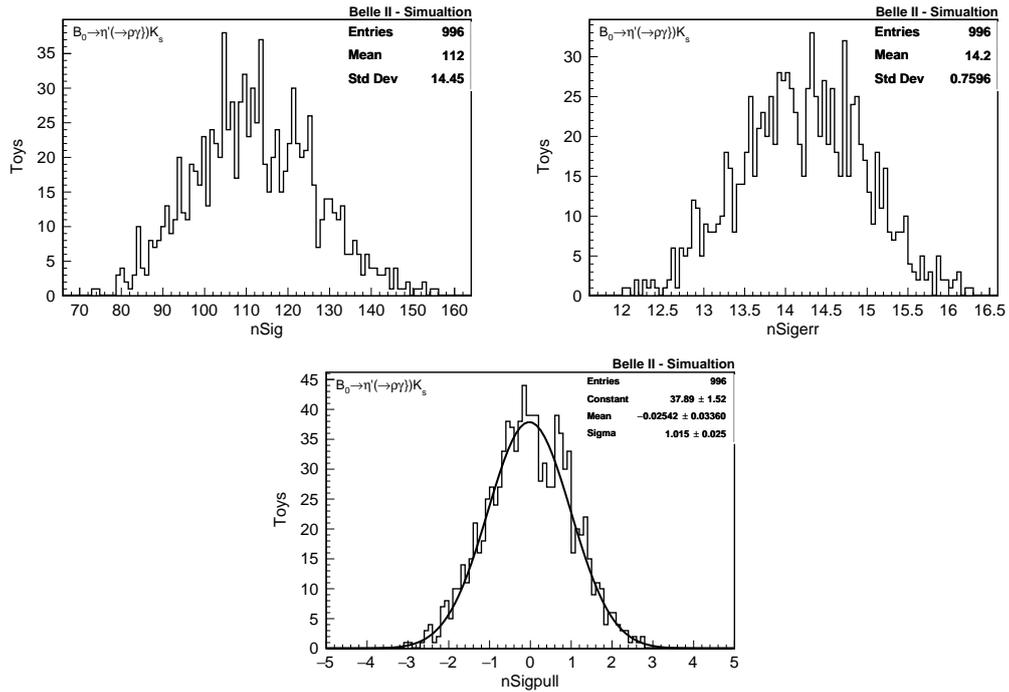
The mean of the yields agrees with the injected one, and the pulls are normally distributed, with mean compatible with zero and  $\sigma \sim 1$ . For the decays with  $\eta' \rightarrow \eta_{\gamma\gamma}\pi^+\pi^-$  an underestimation of the uncertainties on signal yield of about 7.0-8.0% has been observed.

Figure 5.12: Distribution of signal yield, error, and pulls for 1000 toys for  $B^\pm \rightarrow \eta'(\eta_{\gamma\gamma}\pi^+\pi^-)K^\pm$ Figure 5.13: Distribution of signal yield, error, and pulls for 1000 toys for  $B^0 \rightarrow \eta'(\eta_{\gamma\gamma}\pi^+\pi^-)K_S^0$ 

### 5.2.3 Fit results

The fit procedure, described and validated in the previous sections, has been applied to the data samples, measuring the signal yield for each decay channel. In fig. 5.16, 5.17, 5.18 and 5.19 the fit results are shown, for each fit variable, applying a cut on the signal-over-background likelihood ratio of 0.7. The 2-dimensional scatter plot of  $M_{bc}$  and  $\Delta E$  is shown for candidates passing a selection on the continuum suppression variable, defined optimizing the usual figure of merit  $\frac{S}{\sqrt{S+B}}$ .

In table 5.4 the results for the signal yield  $N_{sig}$  and the relative statistical error are given, for each

Figure 5.14: Distribution of signal yield, error, and pulls for 1000 toys for  $B^\pm \rightarrow \eta'(\rho\gamma)K^\pm$ Figure 5.15: Distribution of signal yield, error, and pulls for 1000 toys for  $B^0 \rightarrow \eta'(\rho\gamma)K_S^0$ 

final state.

### 5.3 Branching fractions measurement

The measurement of the branching fractions for the processes studied in this work has been obtained from the number of signal candidates extracted through the fit with RooFit, described in section 5.2.

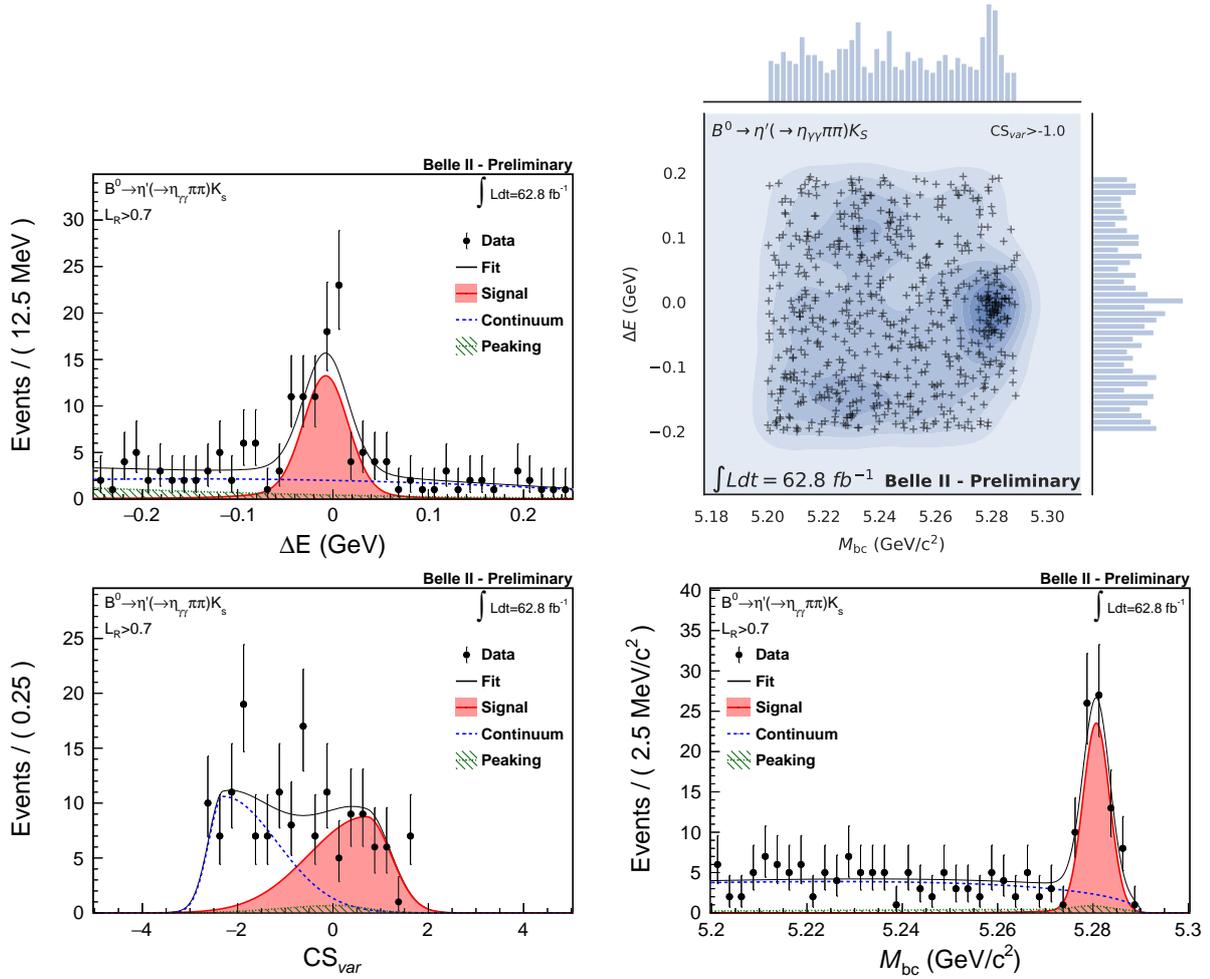


Figure 5.16:  $B^0 \rightarrow \eta'(\eta_{\gamma\gamma} \pi^+ \pi^-) K_S^0$  decay: distribution of the fit variables for candidates with signal-over-background likelihood ratio greater than 0.7 and 2-dimensional plot of  $M_{bc}$  and  $\Delta E$  for candidates with  $CS_{var} > -1.0$

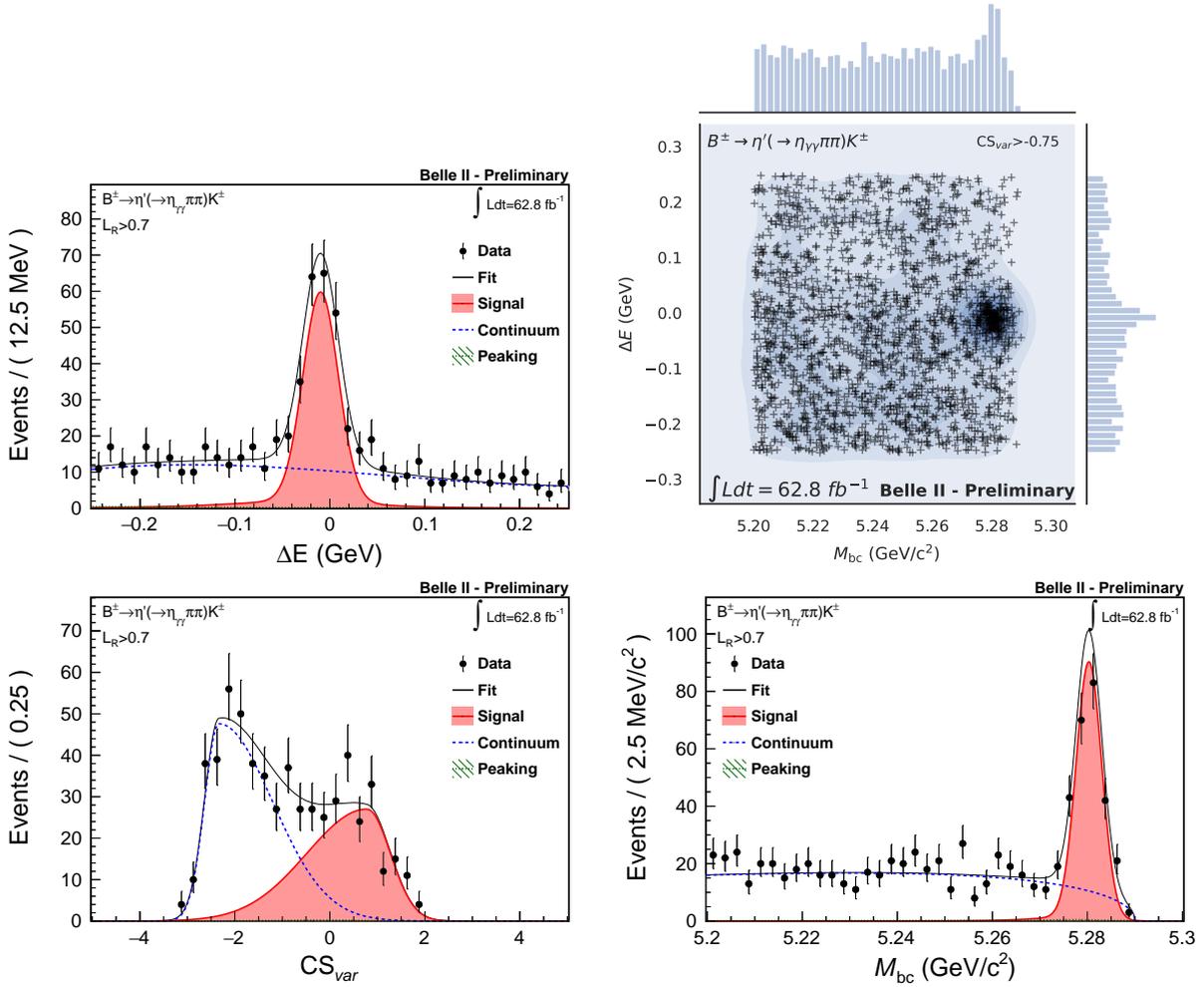


Figure 5.17:  $B^\pm \rightarrow \eta'(\eta_{\gamma\gamma} \pi^+ \pi^-) K^\pm$  decay: distribution of the fit variables for candidates with signal-over-background likelihood ratio greater than 0.7 and 2-dimensional plot of  $M_{bc}$  and  $\Delta E$  for candidates with  $CS_{var} > -0.75$

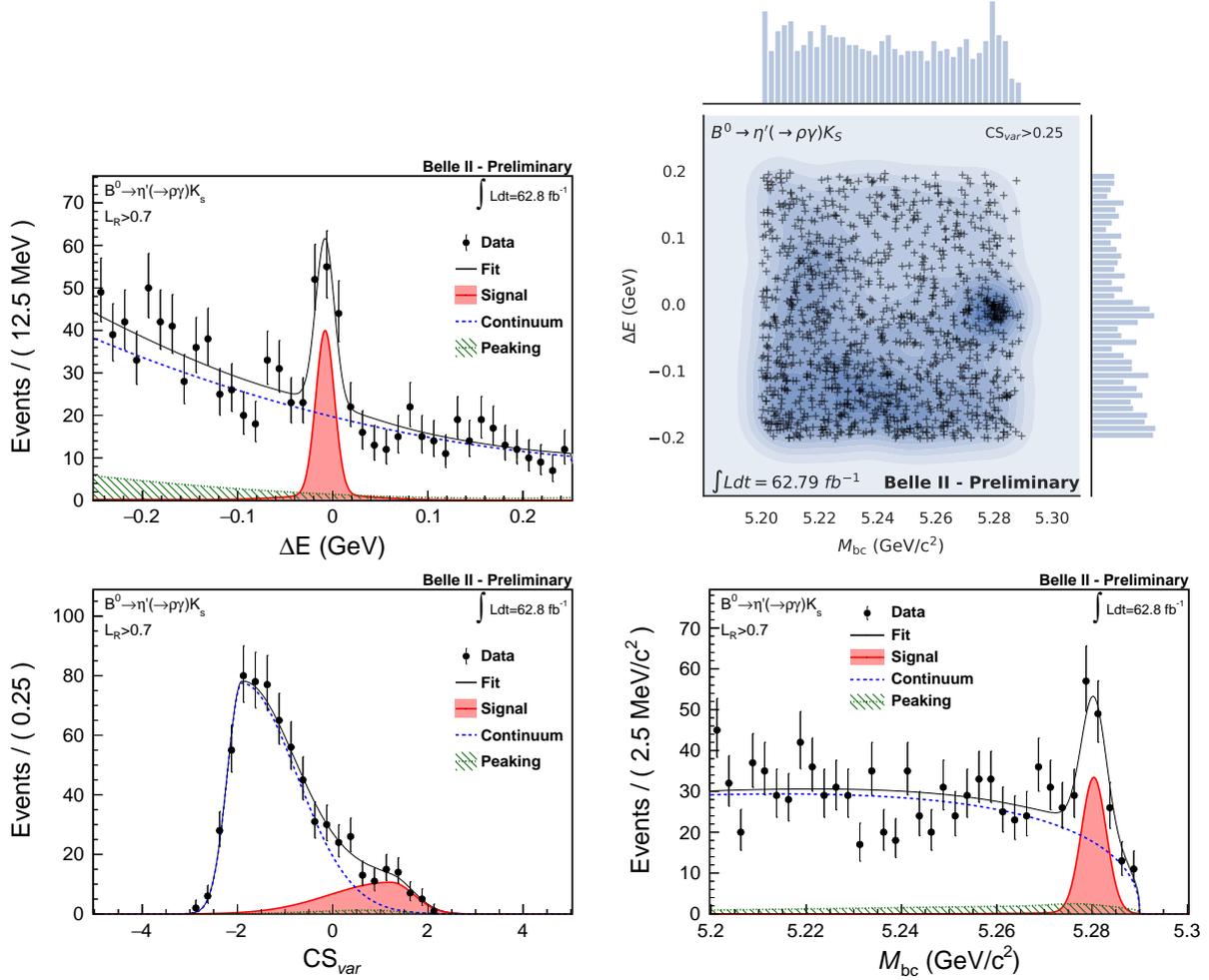


Figure 5.18:  $B^0 \rightarrow \eta'(\rho\gamma)K_S^0$  decay: distribution of the fit variables for candidates with signal-over-background likelihood ratio greater than 0.7 and 2-dimensional plot of  $M_{bc}$  and  $\Delta E$  for candidates with  $CS_{var} > 0.25$

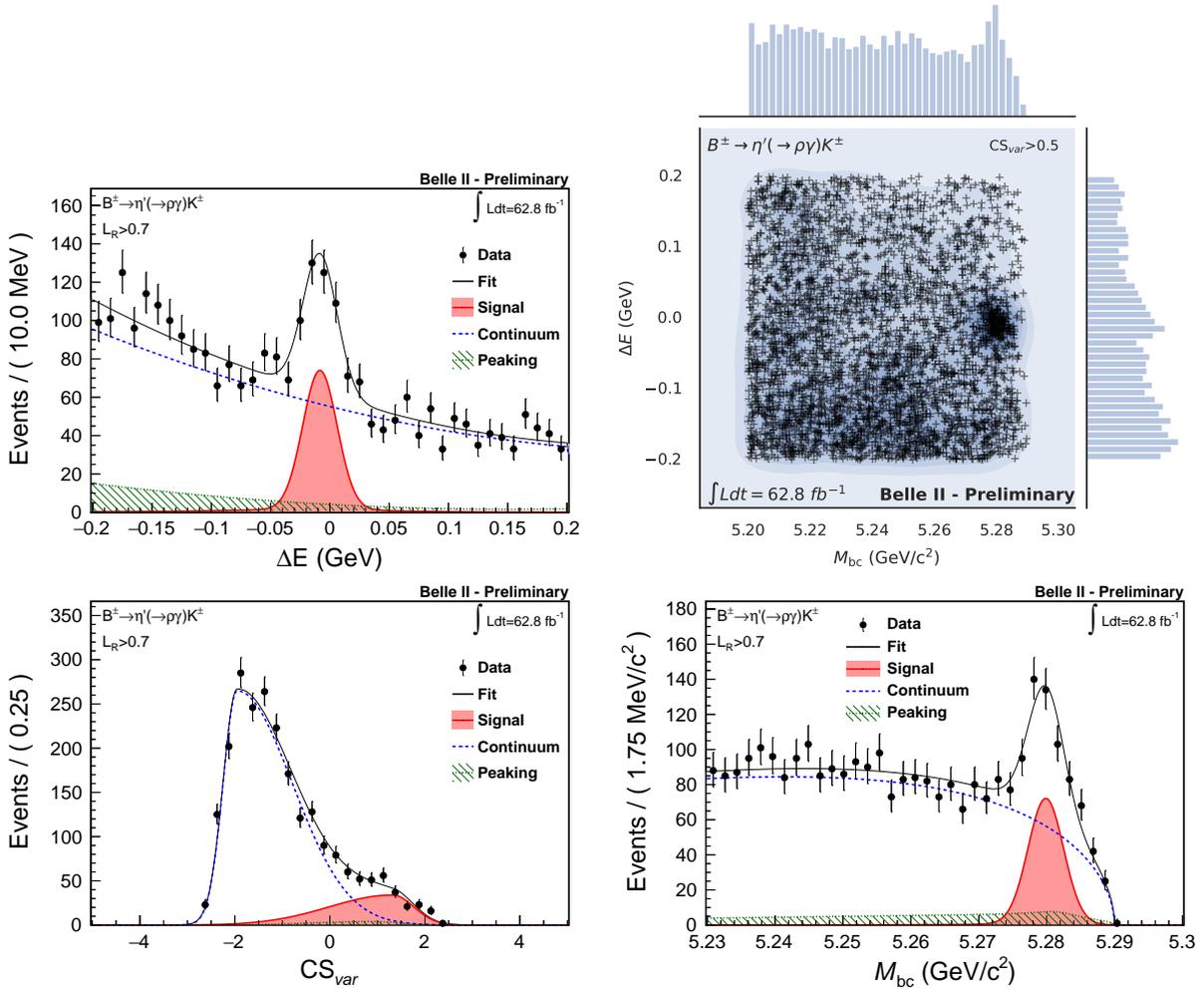


Figure 5.19:  $B^\pm \rightarrow \eta'(\rho\gamma)K^\pm$  decay: distribution of the fit variables for candidates with signal-over-background likelihood ratio greater than 0.7 and 2-dimensional plot of  $M_{bc}$  and  $\Delta E$  for candidates with  $CS_{var} > 0.5$

Decay channel	$N_{sig}$
$B^0 \rightarrow \eta'(\eta_{\gamma\gamma}\pi^+\pi^-)K_S^0$	$80.0^{+11.2}_{-10.4}$
$B^\pm \rightarrow \eta'(\eta_{\gamma\gamma}\pi^+\pi^-)K^\pm$	$263^{+18}_{-19}$
$B^0 \rightarrow \eta'(\rho_{\pi^+\pi^-}\gamma)K_S^0$	$99.7^{+14.2}_{-12.7}$
$B^\pm \rightarrow \eta'(\rho_{\pi^+\pi^-}\gamma)K^\pm$	$335^{+26}_{-25}$

Table 5.4: Fit result for the signal yield. The uncertainties are statistical only.

The branching ratios are computed from the measured signal yield  $N_{sig}$  as

$$\mathcal{B}(B \rightarrow X) = \frac{N_{sig}}{2 \cdot N(B\bar{B}) \cdot f_{00/+} \cdot \varepsilon\mathcal{B}},$$

where  $N(B\bar{B})$  is the number of  $B\bar{B}$  pairs in the dataset,  $f_{00/+}$  is the fraction of  $B^0\bar{B}^0$  and  $B^+B^-$ , respectively, and  $\varepsilon\mathcal{B}$  is the product of the signal reconstruction and selection efficiency, given in table 3.5, and the branching ratio of the considered final state (table 3.1).

The results for the branching fractions measurement [41], obtained for charged and neutral channels, are the following:

$$\begin{aligned} \mathcal{B}(B^\pm \rightarrow \eta'K^\pm) &= (63.4^{+3.4}_{-3.3} \text{ (stat)} \pm 3.2 \text{ (syst)}) \times 10^{-6} \\ \mathcal{B}(B^0 \rightarrow \eta'K^0) &= (59.9^{+5.8}_{-5.5} \text{ (stat)} \pm 2.9 \text{ (syst)}) \times 10^{-6} \end{aligned}$$

where the contributions to the systematic uncertainty are described in section 5.3.1. The results are consistent with the world averages  $\mathcal{B}(B^\pm \rightarrow \eta'K^\pm) = 70.6 \pm 2.5$  and  $\mathcal{B}(B^0 \rightarrow \eta'K^0) = 66 \pm 4$  [8], obtained from the measurements described in [42], [43], and [44].

### 5.3.1 Systematic uncertainties

The signal efficiency is extracted from Monte Carlo, leading to a systematic uncertainty associated to this quantity. The main contributions to the systematic uncertainty on signal efficiency are the following:

- **photons selection:** the uncertainty on photon selection has been obtained from the photon detection efficiency data-to-simulation ratio presented in [45]. This study was performed on a  $e^+e^- \rightarrow \mu^+\mu^-\gamma_{ISR}$ , where  $\gamma_{ISR}$  is an initial-state high energy photon. The ratios are given as a function of the photon energy and its angular direction. The contribution of the selection on the photon timing has been neglected.
- **tracking:** for each charged track in the final state a systematic uncertainty of 0.69% is added linearly, as described in [46]
- **$K_S^0$  reconstruction efficiency:** an uncertainty of 0.31% per cm of 3D flight distance is used, plus a 15% for the  $K_S^0$ s with  $x$ - $y$  distance between 8-10.4 cm, corresponding to layer 2 and 3 of the Silicon Vertex Detector, due to mismodelling of the material in this region. In this analysis, the latter condition applies to 10% of the signal candidates.
- **PID:** the difference in efficiency for data and Monte Carlo has been computed using a sample of  $D^* \rightarrow D(K\pi)\pi_{soft}$  [47] as a function of the momentum and angular distribution of the  $K$  or  $\pi$  particle.

The systematic uncertainty on  $N(B\bar{B})$ , introduced in section 3.4, is 1.4% and it accounts for uncertainties on the cross section, integrated luminosity, and spread of collision energy.

Given the limited statistic available, many parameters of the extended unbinned maximum likelihood fit for the signal extraction have been fixed from Monte Carlo. This approach leads to a systematic uncertainty due to possible mismodelling in the simulated samples. The main contribution to this systematic come from the following effects:

- **$CS_{var}$  modelling:** in section 4.4 the residual mismodelling of the Monte Carlo of the continuum suppression variable has been discussed. The reweighting procedure described in section 4.5 seems to be effective in reducing the discrepancies between simulation and data distributions, but it has not been included in the fit yet. Currently, to account for this modelling problem, a dedicated systematic uncertainty has been introduced repeating the fit while modelling the  $CS_{var}$  PDF distribution on the off-resonance dataset, and taking the difference in yield as systematic uncertainty.
- **SxF fraction:** as described in section 5.1, the SxF fraction has been fixed from Monte Carlo. The systematic uncertainty associated to this procedure has been computed varying the fraction of SxF by  $\pm 50\%$  and evaluating the difference in yield.

The systematic uncertainties are summarized in table 5.5.

Channel	$B^\pm \rightarrow \eta' K^\pm$	$B^0 \rightarrow \eta' K_S^0$	$B^\pm \rightarrow \eta' K^\pm$	$B^0 \rightarrow \eta' K_S^0$
Source	$\eta' \rightarrow \eta(\gamma\gamma)\pi^+\pi^-$		$\eta' \rightarrow \rho\gamma$	
Tracking efficiency	2.1	2.8	2.1	2.8
Photon efficiency	0.5	0.5	0.5	0.5
$K_S^0$ efficiency	-	4.5	-	4.5
$\pi^\pm$ PID	-	-	2.4	2.4
$K^\pm$ PID	2.5	-	2.5	-
$CS_{var}$ modelling	5.0	1.0	5.5	2.3
SxF fraction	2.6	1.8	5.9	3.2
$N(B\bar{B})$			1.4	
Total	6.6	5.9	9.1	7.2

Table 5.5: Contributions to the systematic uncertainty

A large contribution to the systematic uncertainties comes from the effect of the Monte Carlo mismodelling on the signal yield extraction. The source of uncertainty related to the SxF fraction fixing may be removed in future analyses using a larger data sample, that would allow to extract the SxF candidates directly from data. The systematic associated to the modelling of  $CS_{var}$  could be reduced including the Monte Carlo reweighting in the final fit or, with a larger data sample, the PDF parameters could be extracted directly from the off-resonance data sample or from the side band of the dataset, thus solving the mismodelling problem. Belle II has already tripled the statistic of the dataset, with respect to the integrated luminosity used for this analysis, hence new, more precise, results could be obtained using the full dataset and including also the decay channels that were not considered in this study, such as the final states with a  $K_L$ , the  $K_S^0 \rightarrow \pi^0\pi^0$  and the  $\eta' \rightarrow \eta(\pi^0\pi^+\pi^-)\pi^+\pi^-$  decays.

## Chapter 6

# Time-dependent CP Violation

### 6.1 Description of the measurement

The  $\Upsilon(4S)$  produced in the  $e^+e^-$  collisions at Belle II, decays into a  $B\bar{B}$  pair. In this study, the signal events are the ones where a  $B$  meson decays, at time  $t = t_{CP}$ , into the  $\eta'K_S^0$  final state. The other  $B$  meson decays generically at time  $t = t_{tag}$  and, from its final state, the flavour of the  $B_{tag}$  can be determined, as discussed in 6.4. Since the two  $B$  mesons are produced in a coherent quantum state, the determination of the flavour of the  $B_{tag}$  projects the  $B_{CP}$ , at time  $t = t_{tag}$ , to the opposite flavour. Measuring the distance along the beam axis ( $z$ ) between the two decay vertices  $\Delta z = z_{CP} - z_{tag}$ , it is possible to estimate the time difference  $\Delta t$  between the two decays that, neglecting the motion of the  $B$  in the center of mass frame, is approximately  $\Delta t \sim \frac{\Delta z}{\beta\gamma c}$ . The principle of this measurement is schematized in fig. 6.1.

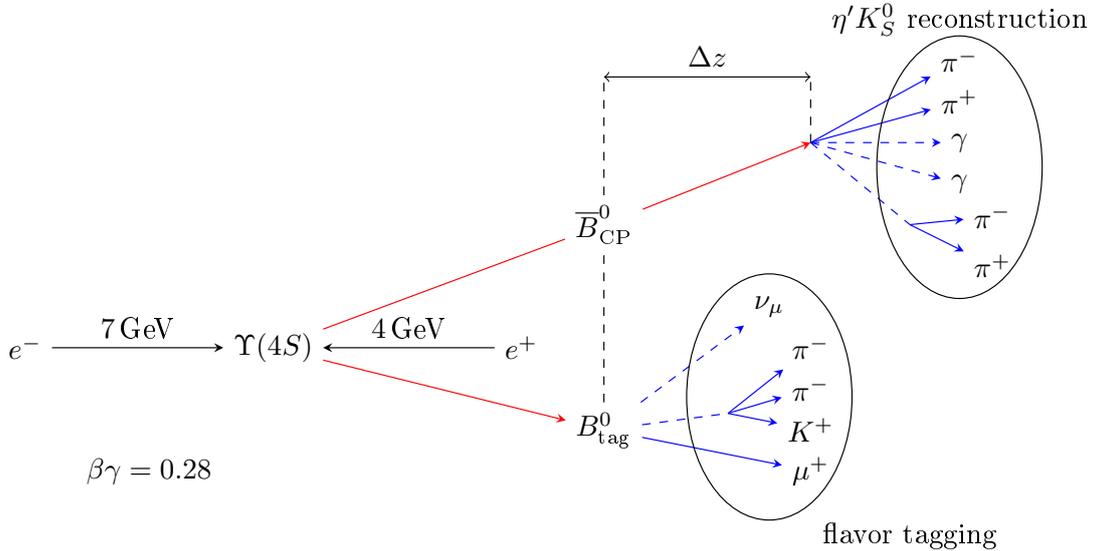


Figure 6.1: Scheme for time-dependent  $CP$  violation studies

The  $CP$  violating parameters are measured from the time-dependent  $CP$  asymmetry:

$$A_{CP} = \frac{\Gamma_{\bar{B}^0 \rightarrow \eta' K_S^0}(\Delta t) - \Gamma_{B^0 \rightarrow \eta' K_S^0}(\Delta t)}{\Gamma_{\bar{B}^0 \rightarrow \eta' K_S^0}(\Delta t) + \Gamma_{B^0 \rightarrow \eta' K_S^0}(\Delta t)} \sim \mathcal{S}_{\eta' K_S^0} \sin(\Delta m \Delta t) + \mathcal{A}_{\eta' K_S^0} \cos(\Delta m \Delta t)$$

## 6.2 Previous results

### 6.2.1 BaBar

The latest BaBar measurement [15] has been performed on a sample of  $(467 \pm 5) \times 10^6$   $B\bar{B}$  pairs, corresponding to the entire dataset collected at the  $\Upsilon(4S)$  resonance. The  $\Delta t$  distribution for  $B$  and  $\bar{B}$  tagged events and the corresponding asymmetry are shown in fig. 6.2 for both  $B^0 \rightarrow \eta' K_S^0$  and  $B^0 \rightarrow \eta' K_L^0$  decays.

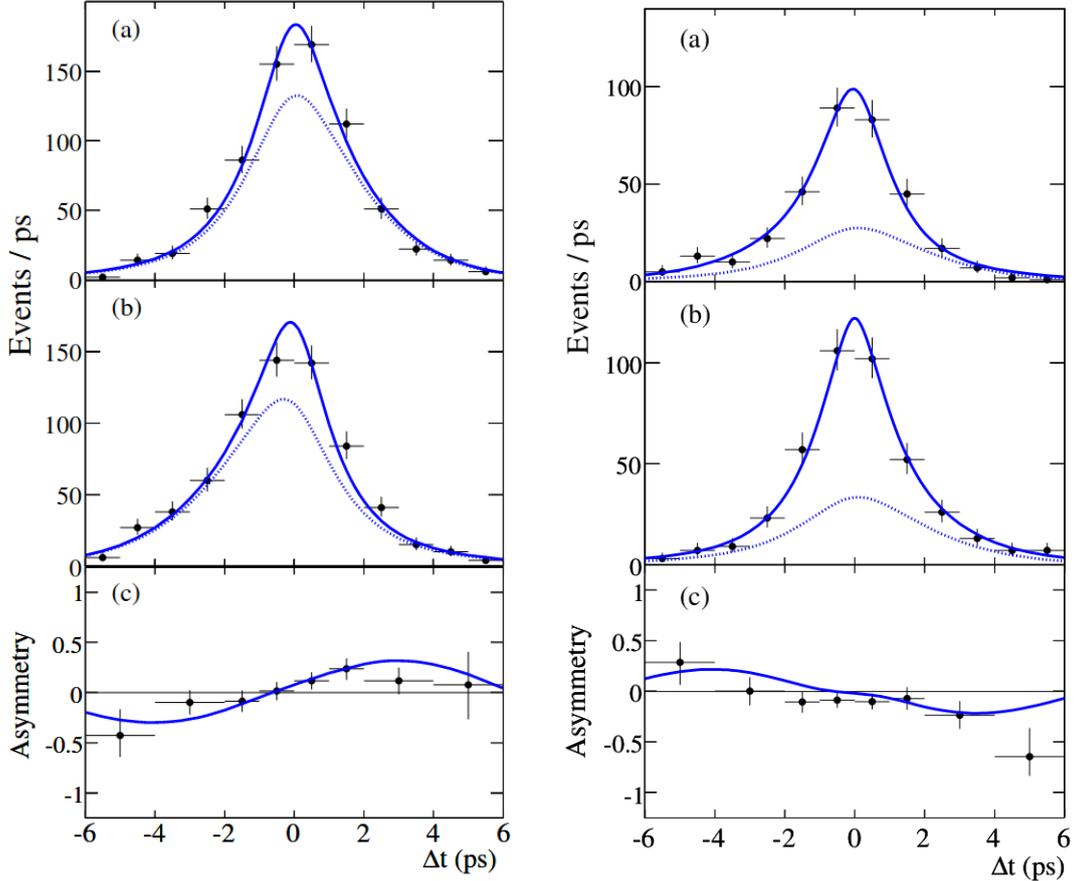


Figure 6.2: Data and model for the  $\Delta t$  distribution for  $B^0$  (a) and  $\bar{B}^0$  (b) tags and asymmetry (c) for both  $B^0 \rightarrow \eta' K_S^0$  (left) and  $B^0 \rightarrow \eta' K_L^0$  (right) decays. These results have been presented by the BaBar Collaboration [15].

The resulting measurements of the  $CP$  violation parameters are the following:

$$\mathcal{A}_{\eta' K^0} = -0.08 \pm 0.06 \pm 0.02$$

$$\mathcal{S}_{\eta' K^0} = 0.57 \pm 0.08 \pm 0.02$$

### 6.2.2 Belle

In 2014 the Belle Collaboration presented a measurement of the  $CP$  violation parameters in  $B^0 \rightarrow \eta' K^0$  decays based on its full data sample, containing  $772 \times 10^6$   $B\bar{B}$  pairs collected at the  $\Upsilon(4S)$  resonance using the Belle detector and the KEKB collider [16]. In fig. 6.3 the  $\Delta t$  distribution and the asymmetry for both  $B^0 \rightarrow \eta' K_S^0$  and  $B^0 \rightarrow \eta' K_L^0$  decays are shown.

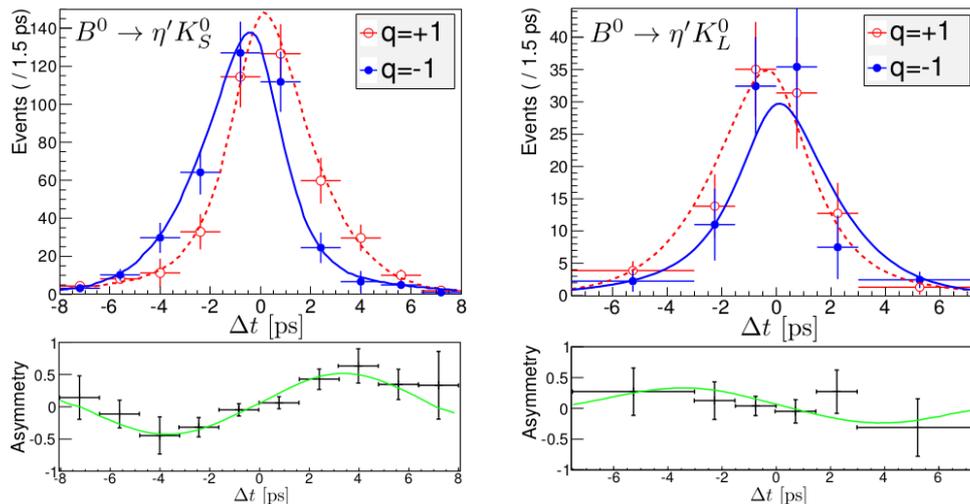


Figure 6.3: Background-subtracted  $\Delta t$  distribution (top) and asymmetry (bottom) for both  $B^0 \rightarrow \eta' K_S^0$  (left) and  $B^0 \rightarrow \eta' K_L^0$  (right) decays. The red and blue curves describe the  $\Delta t$  distribution for the two flavours  $q$ . These results have been presented by the Belle Collaboration [16].

The results obtained for the  $C\mathcal{P}$  violation parameters are the following:

$$\mathcal{A}_{\eta' K^0} = 0.03 \pm 0.05 \pm 0.04$$

$$\mathcal{S}_{\eta' K^0} = 0.68 \pm 0.07 \pm 0.03$$

As discussed in section 1.4.1, this measurement is dominated by the statistical uncertainty and a 10 times larger sample (corresponding to  $\sim 10 \text{ ab}^{-1}$ ) is expected to be necessary to perform a systematic dominated measurement at Belle II.

Belle II has currently collected an integrated luminosity of  $\sim 230 \text{ fb}^{-1}$ . Even though the dataset is not yet large enough to give a significant result, a first time-dependent  $C\mathcal{P}$  violation study could be performed to assess Belle II capability to perform this kind of measurement on  $b \rightarrow s$  transitions. Comparing this preliminary result with the one obtained by the Belle Collaboration with a similar amount of data (for example with  $140 \text{ fb}^{-1}$  [48]) it is also possible to estimate Belle II performances with respect to its predecessor.

For these reasons, since the first crucial step for this measurement is the determination of the  $\Delta t$  resolution function, a study on the signal resolution is presented in the following.

## 6.3 Vertices reconstruction

In order to measure  $\Delta t$ , both  $B$  mesons vertices need to be reconstructed. The reconstruction of the signal side is described in section 3.5 and the vertex fitting uses the `Tree Fitter` algorithm [49], very effective in modes that contain neutral particles. The `Tree Fitter` algorithm performs a global fit of the entire decay tree, taking into account all intermediate particles, extracting the particle's four-momenta, vertex positions, flight lengths and their covariance matrices. The vertex fitting consists in a least square minimization and the best least squares estimator is found by means of the Kalman filter iterative approach, described in [50]. The fit is performed using an Interaction Point constraint for the  $\Upsilon(4S)$  and constraining the masses of the  $\eta$ ,  $\eta'$  and  $K_S^0$  particles to their PDG values. The momenta and vertex position of the daughter particles are updated with the fitted values.

The tag side vertex fit uses the `TagV` module, with the `RAVE` fit algorithm [51]. `TagV` performs the fit over the tracks in the rest-of-event to determine the tag decay vertex, iteratively downweighting the contribution from outlier tracks. The fit is performed applying the `BTube` constraint. Given a fully

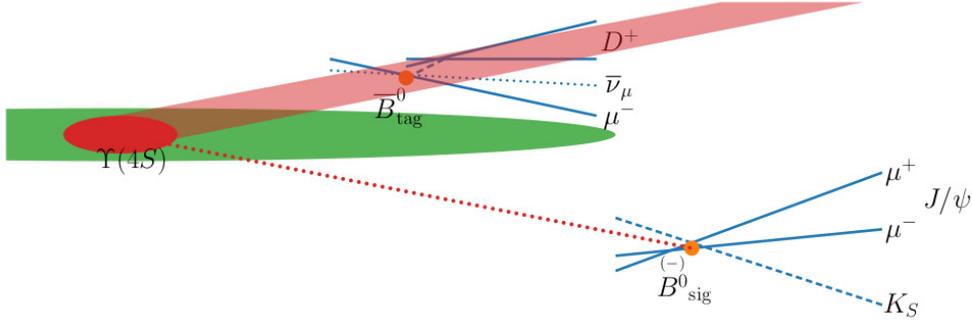


Figure 6.4: Schematic representation of vertices fit strategy.

reconstructed signal candidate, its direction can be propagated to the beam spot, applying a vertex fit to determine the vertex from which both  $B$  mesons originated. From this vertex information and from four-momentum conservation it is possible to determine the flight direction of the  $B_{tag}$  and construct a “tube” constraint for the  $B_{tag}$ , as shown in fig. 6.4.

## 6.4 Flavour tagger

The flavour tagging procedure exploits the fact that many decay modes of the neutral  $B$  mesons provide flavour signatures through flavour-specific final states. For example, in the case of semileptonic decays such as  $B \rightarrow D^* l \bar{\nu}_l$ , the sign of the charge of the lepton  $l$  is related to the flavour of the decaying  $B$ .

Belle II flavour taggers are multivariate algorithms that, from information on tag side kinematic, track-hit, and PID, provide as output the product  $q \cdot r$ , where  $q$  is the tagged flavour ( $-1$  for  $\bar{B}^0$  and  $+1$  for  $B^0$ ) and  $r$  is a dilution factor that ranges from 0, when no flavour tagger information is available, to 1, for an unambiguous flavour assignment. The  $r$  parameter can also be expressed as  $r = 1 - 2w$ , where  $w$  is the fraction of wrongly tagged events. To account for possible differences in the mistag fraction for  $\bar{B}^0$  and  $B^0$  mesons, the  $\Delta w = w_{B^0} - w_{\bar{B}^0}$  has been introduced, while  $w = \frac{w_{B^0} + w_{\bar{B}^0}}{2}$ , where the subscript corresponds to the true flavour. The total effective tagging efficiency is

$$\varepsilon_{eff} = (30.0 \pm 1.2(\text{stat}) \pm 0.4(\text{syst}))\%$$

for a category-based algorithm and

$$\varepsilon_{eff} = (28.8 \pm 1.2(\text{stat}) \pm 0.4(\text{syst}))\%$$

for a deep-learning-based algorithm. More information on the flavour tagging algorithms can be found in [52].

## 6.5 Study of the signal resolution function

The parametrization of the  $\Delta t$  resolution function, for both signal and background components, is an important ingredient of the time-dependent  $CP$  violation measurement. For each component  $i$  (signal and backgrounds) the  $\mathcal{P}_i(\Delta t)$  PDF for the  $\Delta t$  distribution can be defined as the convolution of a PDF that describes the physics of the process ( $\mathcal{P}_i^{phys}(\Delta t)$ ) and the  $\Delta t$  resolution function  $\mathcal{R}_i(\Delta t)$ :

$$\mathcal{P}_i(\Delta t) = \mathcal{P}_i^{phys}(\Delta t) * \mathcal{R}_i(\Delta t)$$

The signal resolution function can be estimated from Monte Carlo, fitting the  $\delta\Delta t = \Delta t - \Delta t_{MC}$  distribution, where  $\Delta t_{MC}$  is the Monte Carlo truth for the decay time difference. The signal physics PDF,  $\mathcal{P}_{sig}^{phys}$ , contains the dependence on the  $CP$  violating parameters:

$$\mathcal{P}_{sig}^{phys}(\Delta t) = \frac{1}{4\tau} e^{-|\Delta t|/\tau} [1 - q\Delta w + q(1 - 2w) \cdot (\mathcal{S}_{\eta'K_S^0} \sin(\Delta m\Delta t) + \mathcal{A}_{\eta'K_S^0} \cos(\Delta m\Delta t))]$$

where  $\tau$  is the lifetime of the  $B$  meson and  $\Delta m$  is the  $B$ -mixing frequency.

The resolution function is different for signal, for continuum, and for peaking background. The continuum resolution can be modelled fitting the  $\Delta t$  distribution, convolving the resolution with a physics PDF defined as the sum of a prompt and a lifetime component:

$$\mathcal{P}_{bkg}^{phys} = f_{\delta} \delta(\Delta t - \mu_{\delta}) + (1 - f_{\delta}) \exp\left(-\frac{|\Delta t - \mu_{\tau}|}{\tau_{bkg}}\right)$$

where  $f_{\delta}$  is the fraction of prompt component,  $\mu_{\delta}$  and  $\mu_{\tau}$  are mean values and  $\tau_{bkg}$  is an effective lifetime. This PDF accounts for resolution effect on prompt component, namely for those particles with short lifetime, as well as those with long lifetime, typically charm mesons.

A similar PDF can model also the peaking background, where the prompt component is expected to be negligible ( $f_{\delta} = 0$ ), since the peaking comes from true  $B\bar{B}$  events.

Since in both cases an inclusive reconstruction is not performed, the lifetime is an effective one.

For this studies, we concentrated on the signal resolution function  $\mathcal{R}_{sig}(\Delta t)$ . The signal resolution is studied using a sample of 200000 simulated signal candidates for the  $B^0 \rightarrow \eta'(\eta_{\gamma\gamma}\pi^+\pi^-)K_S^0$  decay, corresponding to an equivalent integrated luminosity of  $\mathcal{L}_{eq} \sim 50 \text{ ab}^{-1}$ .

In order to exclude pathological events, the following selections have been applied on variables related to the vertices:

- error on  $\Delta t$   $err(\Delta t) < 1.5 \text{ ps}$
- p-value of the tag vertex fit  $> 0.01$

In fig 6.5 the distribution of the two selection variables is shown, showing also the cut region for the p-value distribution. This selection has a  $\sim 95\%$  efficiency on the signal candidates.

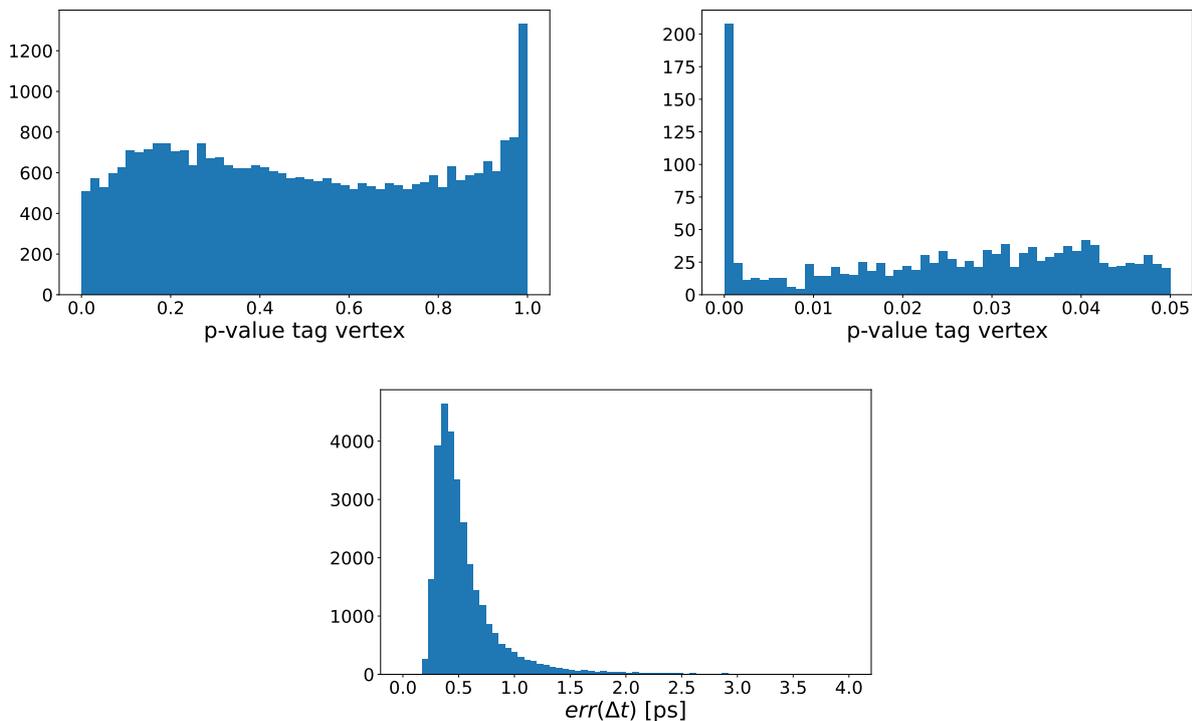


Figure 6.5: Distribution of the p-value of the tag vertex fit (top) and of  $err(\Delta t)$  (bottom). The p-value distribution is shown for its full range (left) and zooming in on the selection region (right).

As discussed above, the signal resolution function is obtained from a fit to the  $\delta\Delta t$  distribution, showed in figure 6.6. In fig. 6.7 the 2-dimensional plot of  $\delta\Delta t$  and  $\Delta t_{MC}$  is shown. The two variables are not

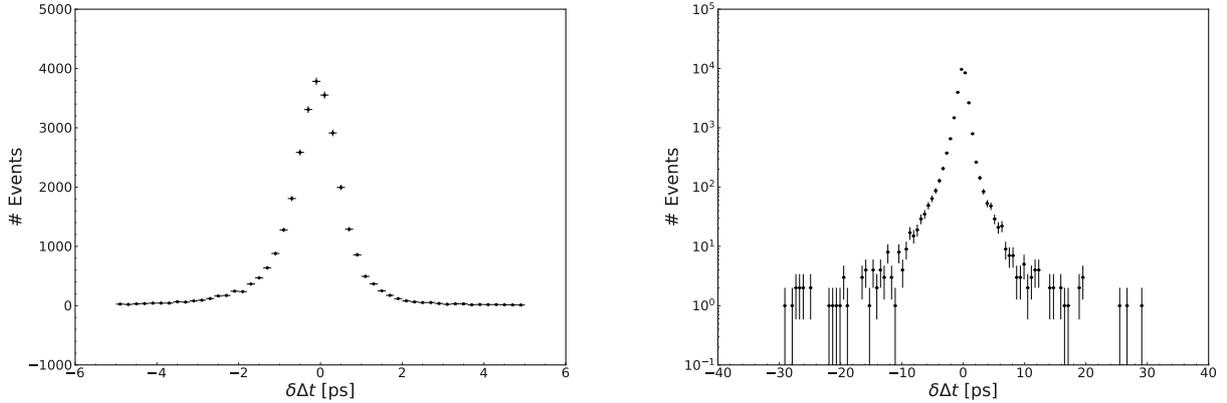


Figure 6.6: Distribution of  $\delta\Delta t = \Delta t - \Delta t_{MC}$

correlated, as expected, with correlation coefficient (Pearson)  $r = 0.00058$ .

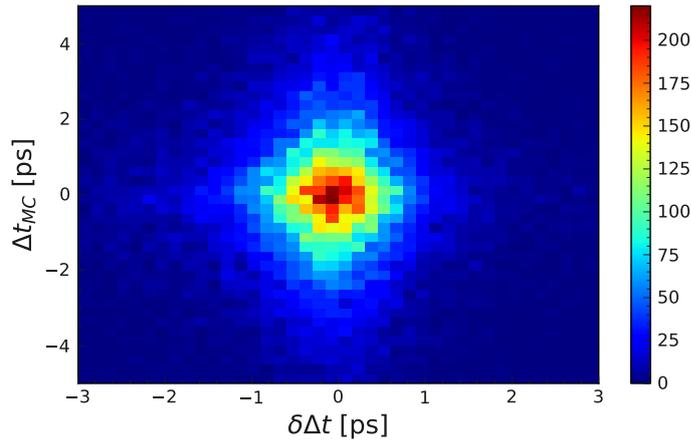


Figure 6.7: Two-dimensional plot of  $\delta\Delta t$  and  $\Delta t_{MC}$  variables.

In a first approximation, the  $\delta\Delta t$  distribution can be described as the sum of three gaussian functions: one for the core of the distribution and two to describe the long tails. Both Belle and BaBar experiments described the resolution function using the error on the measured  $\Delta t$  as an input. In fig. 6.8 the distribution of the residuals  $\delta\Delta t$  for different  $err(\Delta t)$  bins is shown and, in fig. 6.9, the distribution of the pulls  $\frac{\delta\Delta t}{err(\Delta t)}$  is given.

The resolution function is different in bins of  $err(\Delta t)$  but the pulls distribution appears to be constant in bins of  $err(\Delta t)$ .

For this reason, the shape of the resolution function is extracted from a 2-dimensional fit of  $\delta\Delta t = \Delta t - \Delta t_{MC}$  and the error on  $\Delta t$ ,  $err(\Delta t)$ , as described in the next section.

### 6.5.1 Fit of the resolution function

The study on the signal resolution function has been performed using the `zfit` toolkit.

The  $err(\Delta t)$  distribution has been fitted with a double inverse gaussian and the result of the fit has been stored as an histogram. In fig. 6.10 the  $err(\Delta t)$  PDF fit result is shown.

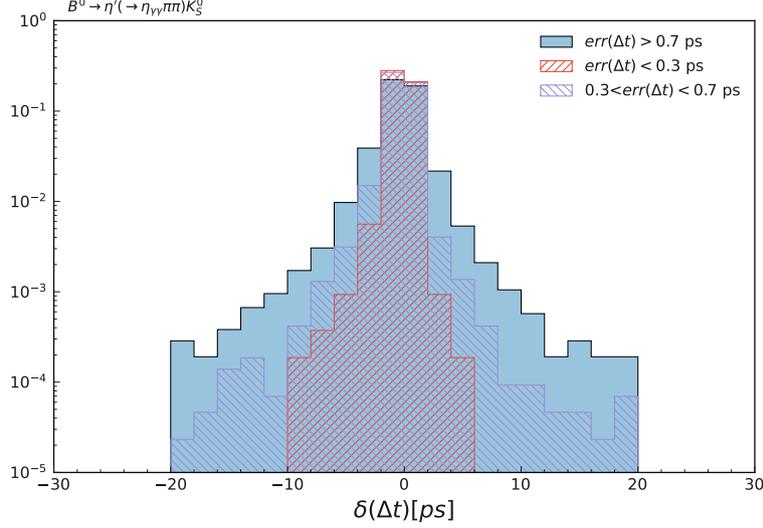


Figure 6.8: Distribution of  $\delta\Delta t = \Delta t - \Delta t_{MC}$  for three  $err(\Delta t)$  bins:  $err(\Delta t) > 0.7$  ps,  $err(\Delta t) < 0.3$  ps and  $0.3 < err(\Delta t) < 0.7$  ps. The distributions are normalized to unit area.

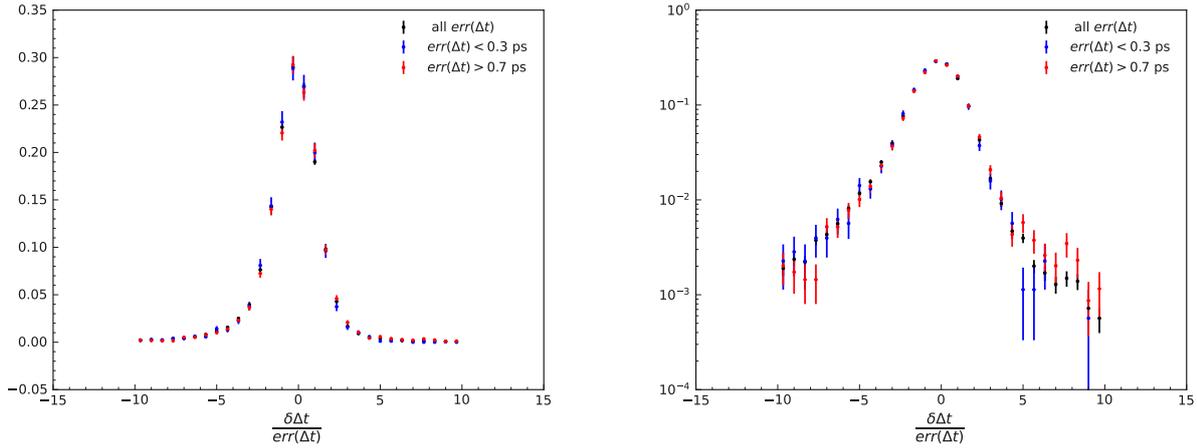


Figure 6.9: Distribution of  $\frac{\Delta t - \Delta t_{MC}}{err(\Delta t)}$  for two different bins of  $err(\Delta t)$ :  $err(\Delta t) < 0.3$  ps (blue) and  $err(\Delta t) > 0.7$  ps (red). The distributions are normalized to unit area

The PDF of the  $\delta\Delta t$  distribution has been defined as the sum of three gaussian functions, one gaussian to describe the core and two gaussians for the tails of the distribution. The mean and  $\sigma$  of the gaussians are not constant and they are defined as follows:

$$\mu = \mu_1 \cdot err(\Delta t)$$

$$\sigma = \sigma_1 \cdot err(\Delta t)$$

where  $err(\Delta t)$  indicates the fitted  $err(\Delta t)$  histogram. In this way we account for the  $err(\Delta t)$  dependence of the width of the resolution distribution, and also allow for a similar dependence of the mean, even if this is not observed. Hence for the resolution PDF we have

$$\begin{aligned} \mathcal{R}(\Delta t) = & (1 - f_{tail,1} - f_{tail,2}) \cdot \text{Gauss}(\delta\Delta t; \mu_{main} \cdot err(\Delta t), \sigma_{main} \cdot err(\Delta t)) \\ & + f_{tail,1} \cdot \text{Gauss}(\delta\Delta t; \mu_{tail,1} \cdot err(\Delta t), \sigma_{tail,1} \cdot err(\Delta t)) \\ & + f_{tail,2} \cdot \text{Gauss}(\delta\Delta t; \mu_{tail,2} \cdot err(\Delta t), \sigma_{tail,2} \cdot err(\Delta t)) \end{aligned} \quad (6.1)$$

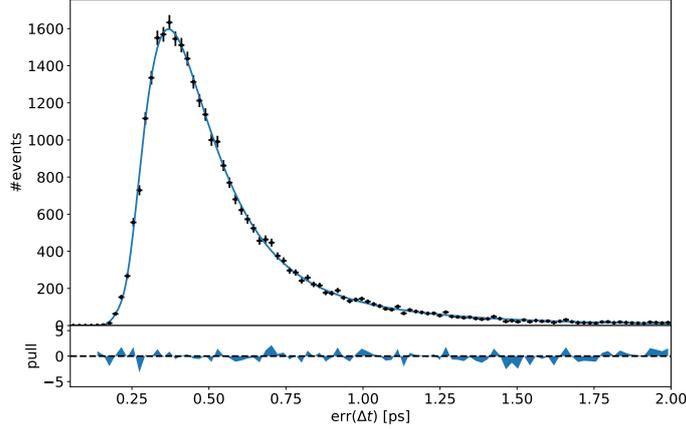


Figure 6.10: Distribution of  $err(\Delta t)$  and fitted PDF

The signal Monte Carlo sample has been divided in two independent subsamples and one of them is used to fix the resolution function parameters. In fig. 6.11 the result of the resolution function fit is shown.

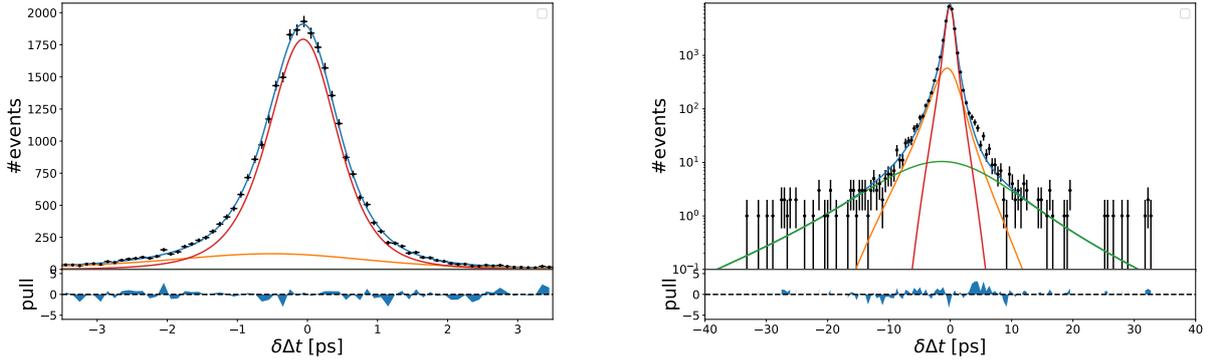


Figure 6.11: Resolution function fit. The blue distribution is the total fitted resolution function. The core Gaussian distribution (red) and the tail distribution (orange and green) are also shown.

The resolution function describes well both the peak and the tails of the distribution, with some possible problems in the high tail as shown in fig. 6.11.

To test the resolution function, it is possible to perform a lifetime fit on  $\Delta t$  on the signal Monte Carlo subsample not used to define  $\mathcal{R}(\Delta t)$ . The results of this test are discussed in the next section.

### 6.5.2 Lifetime fit on Monte Carlo

A  $B$  lifetime measurement performed on  $B \rightarrow \eta' K$  decays cannot produce world leading results, since the branching fractions are sensibly smaller than other hadronic or semileptonic decays, such as  $B \rightarrow DX$  or  $B \rightarrow Dl\nu_l X$ . However, since the lifetime measurement strongly depends on the correct parametrization of the  $\mathcal{R}(\Delta t)$  function, it provides a good test of the resolution function.

The lifetime fit test has been performed using a signal Monte Carlo sample corresponding to three different equivalent integrated luminosities:

- $\mathcal{L}_{eq} \sim 200 \text{ fb}^{-1}$ , corresponding to the integrated luminosity currently collected at Belle II
- $\mathcal{L}_{eq} \sim 1 \text{ ab}^{-1}$ , expected to be collected before the first shutdown in 2023

- $\mathcal{L}_{eq} \sim 10 \text{ ab}^{-1}$ , corresponding to the luminosity needed to have a systematic dominated measurement of the  $CP$  violation parameters

The lifetime PDF has been defined as the convolution of an exponential  $\mathcal{P}(\Delta t) = e^{-|\Delta t|/\tau}$  with the resolution function.

### Lifetime fit with $\mathcal{L}_{eq} \sim 200 \text{ fb}^{-1}$

Given the low statistics of the fitting sample ( $\sim 200$  candidates for  $B^0 \rightarrow \eta'(\eta_{\gamma\gamma}\pi^+\pi^-)K_S^0$ ), the fit has been repeated using 20 samples of signal Monte Carlo. The average and standard deviation of the fit results have been computed:

$$\tau = 1.55 \pm 0.10 \text{ ps}$$

The result is consistent with the Monte Carlo truth  $\tau_{true} = 1.525 \text{ ps}$ .

This result is affected by a large statistical uncertainty. The same is expected for the measurement of the  $CP$  violation parameters  $\mathcal{A}$  and  $\mathcal{S}$ .

### Lifetime fit with $\mathcal{L}_{eq} \sim 1 \text{ ab}^{-1}$

The lifetime fit result on the Monte Carlo sample with  $\mathcal{L}_{eq} \sim 1 \text{ ab}^{-1}$  is shown in fig. 6.12. The resulting

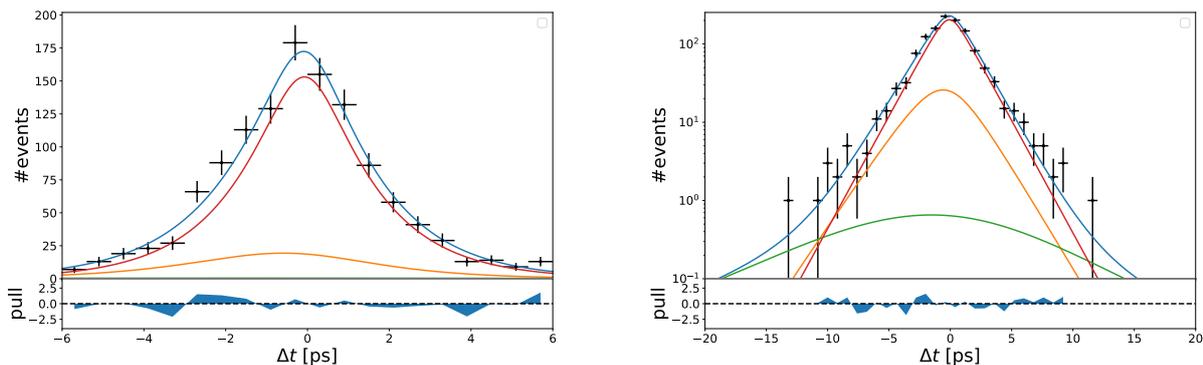


Figure 6.12: Lifetime fit results on signal Monte Carlo sample with  $\mathcal{L}_{eq} \sim 1 \text{ ab}^{-1}$

lifetime is

$$\tau = 1.511 \pm 0.052 \text{ ps}$$

and it is consistent with the Monte Carlo truth  $\tau_{true} = 1.525 \text{ ps}$ .

### Lifetime fit with $\mathcal{L}_{eq} \sim 10 \text{ ab}^{-1}$

The lifetime fit result on the Monte Carlo sample with  $\mathcal{L}_{eq} \sim 10 \text{ ab}^{-1}$  is shown in fig. 6.13.

The resulting lifetime is

$$\tau = 1.531 \pm 0.017 \text{ ps}$$

consistent with the Monte Carlo truth  $\tau_{true} = 1.525 \text{ ps}$ .

Using this large sample, the statistical error is small and the systematic contribution becomes relevant. In particular, computing  $\Delta t$  as  $\Delta t \sim \frac{\Delta z}{\beta\gamma c}$  the motion of the  $B$  in the center of mass frame is being neglected and this could induce a bias on the measurements performed fitting the  $\Delta t$  distribution. A study on  $B \rightarrow D^{(*)-}\pi^+$  has shown that the neglect of this *kinematic approximation* can induce a bias on the measured lifetime of the order of 10 fs. In view of the future analysis with larger samples, it will be necessary to correct for this kinematic approximation and also to evaluate the contribution to the systematic uncertainty of other effects, such as the alignment of the vertex detector.

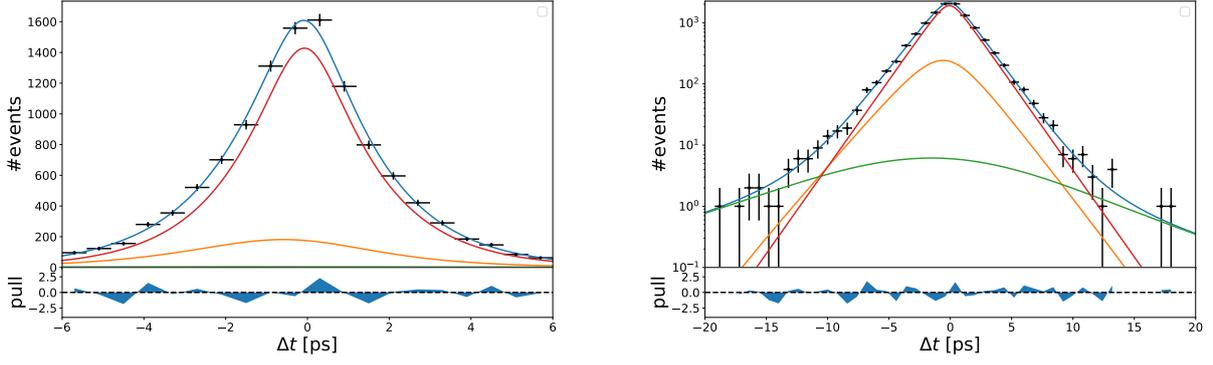


Figure 6.13: Lifetime fit results on signal Monte Carlo sample with  $\mathcal{L}_{eq} \sim 10 \text{ ab}^{-1}$

The signal resolution function described in this chapter seems to provide a good description of the Monte Carlo signal  $\Delta t$  resolution, and it is a good starting point for the study of  $\mathcal{R}(\Delta t)$ . Other studies are needed to describe the  $\Delta t$  distribution also for the background components. After completing the building of the resolution function for all the components, it could be tested performing a lifetime measurement on the data sample. As discussed before, this measurement would be useful only as a test of the resolution function, since the lifetime measurement is precisely performed on final states with higher branching fractions.

# Chapter 7

## Conclusion

The measurement of  $C\mathcal{P}$  violation in loop-mediated  $B^0 \rightarrow \eta' K_S^0$  decays is a good test of the Standard Model, since it is sensitive to the presence of new physics effects in the loop.

In this work, a study of this decay in Belle II 2019-2020 data ( $62.8 \text{ fb}^{-1}$ ) has been presented, focusing on the signal extraction and the suppression of the continuum  $q\bar{q}$  background. A multivariate signal-continuum classification has been performed, training the model with variables related to the topology of the event. The output of the classifier,  $CS_{var}$ , provides a good separation between signal and continuum candidates, but its Monte Carlo distribution shows a small disagreement with the data distribution. To account for this discrepancy a dedicated systematic uncertainty has been introduced, and a study on a Monte Carlo multivariate reweighting procedure has been performed. The Monte Carlo reweighting showed to be effective in reducing the  $CS_{var}$  mismodelling in Monte Carlo and could be used in future analyses to reduce the systematic uncertainty due to this effect.

From the signal yield, extracted through an extended unbinned maximum likelihood fit, a measurement of the branching fractions has been obtained for charged and neutral channels:

$$\mathcal{B}(B^\pm \rightarrow \eta' K^\pm) = (63.4_{-3.3}^{+3.4}(\text{stat}) \pm 3.2(\text{syst})) \times 10^{-6}$$

$$\mathcal{B}(B^0 \rightarrow \eta' K^0) = (59.9_{-5.5}^{+5.8}(\text{stat}) \pm 2.9(\text{syst})) \times 10^{-6}$$

where the main sources of systematic uncertainty are related to the  $K_S^0$  reconstruction efficiency (4.5%), continuum suppression modelling (up to  $\sim 5\%$  for charged channels) and SxF fraction (from 1.8% to 5.9%). The measured branching fractions are consistent with the world averages. The current Belle II dataset, corresponding to an integrated luminosity of  $230 \text{ fb}^{-1}$ , could allow to perform a new branching fractions measurement in the near future. With a larger dataset, both the statistical and systematic uncertainties are expected to be reduced, since many quantities could be extracted directly from data, without relying on Monte Carlo.

Finally, a preliminar study on the  $\Delta t$  resolution function has been presented, defining the PDF shape of the signal resolution function and testing it through a lifetime fit to signal Monte Carlo candidates. The resolution function describes well the peak of the  $\delta\Delta t$  distribution, while the tails show some disagreement with the fitted model. The lifetime test was performed for different integrated luminosities and the resulting lifetime was consistent with the Monte Carlo truth value  $\tau_{true} = 1.525 \text{ ps}$ .

This result provides a good starting point for the modelling of the resolution function, in view of the future time-dependent  $C\mathcal{P}$  violation measurement in the  $B^0 \rightarrow \eta' K_S^0$  decay.



# Bibliography

- [1] A. D. Sakharov. Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe. *Pisma Zh. Eksp. Teor. Fiz.*, 5:32–35, 1967.
- [2] The ATLAS Collaboration, G Aad, E Abat, J Abdallah, A A Abdelalim, A Abdesselam, et al. The ATLAS experiment at the CERN large hadron collider. *Journal of Instrumentation*, 3(08):S08003–S08003, aug 2008.
- [3] The CMS Collaboration, S Chatrchyan, G Hmayakyan, V Khachatryan, A M Sirunyan, W Adam, et al. The CMS experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08004–S08004, aug 2008.
- [4] Nicola Cabibbo. Unitary symmetry and leptonic decays. *Phys. Rev. Lett.*, 10:531–533, Jun 1963.
- [5] Lincoln Wolfenstein. Parametrization of the Kobayashi-Maskawa Matrix. *Phys. Rev. Lett.*, 51:1945–1947, Nov 1983.
- [6] The LHCb Collaboration, A Augusto Alves, L M Andrade Filho, et al. The LHCb detector at the LHC. *Journal of Instrumentation*, 3(08):S08005–S08005, aug 2008.
- [7] Ckm fitter, [http://ckmfitter.in2p3.fr/www/html/ckm\\_results.html](http://ckmfitter.in2p3.fr/www/html/ckm_results.html).
- [8] P.A. Zyla et al. Review of Particle Physics. *PTEP*, 2020(8):083C01, 2020.
- [9] W. Altmannshofer et al. The Belle II Physics Book. *PTEP*, 2019(12):123C01, 2019.
- [10] 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.*, 105:1413–1415, Feb 1957.
- [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, Jul 1964.
- [12] Makoto Kobayashi and Toshihide Maskawa. CP-Violation in the Renormalizable Theory of Weak Interaction. *Progress of Theoretical Physics*, 49(2):652–657, 02 1973.
- [13] K. Abe, K. Abe, R. Abe, I. Adachi, Byoung Sup Ahn, H. Aihara, M. Akatsu, G. Alimonti, K. Asai, M. Asai, et al. Observation of Large CP Violation in the Neutral  $B$  Meson System. *Physical Review Letters*, 87(9), Aug 2001.
- [14] B. Aubert, D. Boutigny, J.-M. Gaillard, A. Hicheur, Y. Karyotakis, J. P. Lees, P. Robbe, V. Tisserand, A. Palano, G. P. Chen, et al. Observation of CP Violation in the  $B^0$  Meson System. *Physical Review Letters*, 87(9), Aug 2001.
- [15] Bernard Aubert et al. Measurement of time dependent CP asymmetry parameters in  $B^0$  meson decays to  $\omega K_S^0$ ,  $\eta' K^0$ , and  $\pi^0 K_S^0$ . *Phys. Rev. D*, 79:052003, 2009.
- [16] L. Šantelj et al. Measurement of Time-Dependent CP Violation in  $B^0 \rightarrow \eta' K^0$  Decays. *JHEP*, 10:165, 2014.
- [17] KEKB B-Factory Design Report. *KEK Report 95-7*, 1995.

- [18] M. Tobiayama Y. Suetsugu and K. Furukawa. Long-term SuperKEKB operation and upgrade plan. *39th B2GM*. <https://indico.belle2.org/event/4490/contributions/21664/>.
- [19] Kazunori Akai, Kazuro Furukawa, and Haruyo Koiso. SuperKEKB Collider. *Nucl. Instrum. Meth.*, A907:188–199, 2018.
- [20] T. Abe. Belle II Technical Design Report. 2010. KEK-REPORT-2010-1.
- [21] Stefano Lacaprara. Rediscovery of  $\eta$  and  $\eta'$  mesons in belle ii data. *BELLE2-NOTE-PH-2018-038*, Nov 2018.
- [22] A. J. Bevan et al. The Physics of the B Factories. *Eur. Phys. J.*, C74:3026, 2014.
- [23] F. Abudinén, I. Adachi, P. Ahlburg, H. Aihara, N. Akopov, A. Aloisio, F. Ameli, L. Andricsek, N. Anh Ky, D. M. Asner, et al. Measurement of the integrated luminosity of the phase 2 data of the belle ii experiment. *Chinese Physics C*, 44(2):021001, Jan 2020.
- [24] The Belle II Collaboration. Approved B counting analysis results with ICHEP2020 dataset. *BELLE2-NOTE-PL-2020-006*, Jul 2020.
- [25] S. Jadach, B.F.L. Ward, and Z. Was. The precision monte carlo event generator for two-fermion final states in collisions. *Computer Physics Communications*, 130(3):260–325, Aug 2000.
- [26] Torbjörn Sjöstrand, Stefan Ask, Jesper R. Christiansen, Richard Corke, Nishita Desai, Philip Ilten, Stephen Mrenna, Stefan Prestel, Christine O. Rasmussen, and Peter Z. Skands. An introduction to pythia 8.2. *Computer Physics Communications*, 191:159–177, Jun 2015.
- [27] David J. Lange. The evtgen particle decay simulation package. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 462(1):152–155, 2001. BEAUTY2000, Proceedings of the 7th Int. Conf. on B-Physics at Hadron Machines.
- [28] S. Agostinelli et al. Geant4-a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250–303, 2003.
- [29] Stefano Lacaprara. Rediscovery of  $\eta$  and  $\eta'$  mesons in early phase 3 belle ii data. *BELLE2-NOTE-PL-2020-003*, Apr 2020.
- [30] S. H. Lee, K. Suzuki, K. Abe, K. Abe, T. Abe, I. Adachi, et al. Evidence for  $B^0 \rightarrow \pi^0\pi^0$ . *Phys. Rev. Lett.*, 91:261801, Dec 2003.
- [31] D. M. Asner, M. Athanas, D. W. Bliss, W. S. Brower, G. Masek, et al. Search for exclusive charmless hadronic b decays. *Phys. Rev. D*, 53:1039–1050, Feb 1996.
- [32] Thomas Keck. FastBDT: A Speed-Optimized Multivariate Classification Algorithm for the Belle II Experiment. *Comput. Softw. Big Sci.*, 1(1):2, 2017.
- [33] Pankaj Mehta, Marin Bukov, Ching-Hao Wang, Alexandre G.R. Day, Clint Richardson, Charles K. Fisher, and David J. Schwab. A high-bias, low-variance introduction to machine learning for physicists. *Physics Reports*, 810:1–124, May 2019.
- [34] D Martschei, M Feindt, S Honc, and J Wagner-Kuhr. Advanced event reweighting using multivariate analysis. *Journal of Physics: Conference Series*, 368:012028, jun 2012.
- [35] Wouter Verkerke and David P. Kirkby. The RooFit toolkit for data modeling. *eConf*, C0303241:MOLT007, 2003.
- [36] Lei Zhang. The RooRarFit Document. Technical report. <http://rarfit.sourceforge.net/>.
- [37] <https://github.com/zfit/zfit/>.

- [38] F. James. MINUIT Function Minimization and Error Analysis: Reference Manual Version 94.1. Technical report, 1994. CERN-D-506, CERN-D506.
- [39] John Erthal Gaiser. Charmonium Spectroscopy From Radiative Decays of the  $J/\psi$  and  $\psi'$ . Master thesis, Aug. 1982. SLAC-0255, UMI-83-14449-MC, SLAC-R-0255, SLAC-R-255.
- [40] H. Albrecht et al. Search for hadronic  $b \rightarrow u$  decays. *Phys. Lett. B*, 241(2):278 – 282, 1990.
- [41] The Belle II Collaboration. Measurement of the branching fractions of  $B \rightarrow \eta'K$  decays using 2019/2020 Belle II data. arXiv: 2104.06224.
- [42] K Abe, K Abe, R Abe, et al. Measurement of the branching fraction for  $B \rightarrow \eta'K$  and search for  $B \rightarrow \eta'\pi^+$ . *Physics Letters B*, 517(3):309–318, 2001.
- [43] B. Aubert, Y. Karyotakis, J. P. Lees, V. Poireau, E. Prencipe, X. Prudent, V. Tisserand, J. Garra Tico, E. Grauges, M. Martinelli, et al.  $B$  meson decays to charmless meson pairs containing  $\eta$  or  $\eta'$  mesons. *Physical Review D*, 80(11), Dec 2009.
- [44] J. Schümann, C. H. Wang, K. Abe, I. Adachi, H. Aihara, D. Anipko, K. Arinstein, Y. Asano, T. Aushev, A. M. Bakich, et al. Evidence for  $B \rightarrow \eta'\pi$  and Improved Measurements for  $B \rightarrow \eta'K$ . *Physical Review Letters*, 97(6), Aug 2006.
- [45] Henrikas Svidras, Markus Röhrken, Simon Wehle, and Kerstin Tackmann. Measurement of the data to mc ratio of photon reconstruction efficiency of the belle ii calorimeter using radiative muon pair events. *BELLE2-NOTE-TE-2020-026*, Sep 2021.
- [46] Alessandro GAZ. Rediscovery of  $B^+ \rightarrow \phi K^{(*)+}$  and  $B^0 \rightarrow \phi K^{(*)0}$  using the Summer 2020 dataset. *BELLE2-NOTE-PH-2020-019 (Internal)*, May 2020.
- [47] The Belle II Collaboration. Kaon and pion identification performances in phase iii data. *BELLE2-NOTE-PL-2020-024*, Jul 2020.
- [48] Katsushige Abe, Toshi Abe, I Adachi, Hiroaki Aihara, K Akai, M Akatsu, Yoshihiro Asano, T Aso, Tagir Aushev, Seema Bahinipati, AM Bakich, Y Ban, Sampriti Banerjee, I Bedny, I Bizjak, A Bondar, Andrzej Bozek, M Bracko, J Brodzicka, and Belle Collaboration. Measurement of Time-Dependent CP-Violating Asymmetries in  $B^0 \rightarrow \phi K_S^0$ ,  $K^+K^-K_S^0$ , and  $\eta'K_S^0$  Decays. *Physical Review Letters*, 91, 12 2003.
- [49] J.-F. Krohn, F. Tenchini, P. Urquijo, F. Abudinén, S. Cunliffe, T. Ferber, M. Gelb, J. Gemmler, P. Goldenzweig, T. Keck, and et al. Global decay chain vertex fitting at belle ii. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 976:164269, Oct 2020.
- [50] Wouter D. Hulsbergen. Decay chain fitting with a kalman filter. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 552(3):566–575, Nov 2005.
- [51] Wolfgang Waltenberger. Rave—a detector-independent toolkit to reconstruct vertices. *Nuclear Science, IEEE Transactions on*, 58:434 – 444, 05 2011.
- [52] F. Abudinén, N. Akopov, et al.  $B$ -flavor tagging at Belle II. 2021. arXiv: 2110.00790.



# Appendices



# Appendix A

## Fit for signal extraction with `zfit`

In this section an attempt at extracting the number of signal candidates using the `zfit` analysis toolkit is presented. Since `zfit` will be used to perform the time-dependent  $CP$  violation analysis, it is useful to built the UML fit for the signal extraction using this toolkit.

The fit strategy is described in chapter 5.

### A.0.1 Extraction of the PDFs from Monte Carlo

The PDFs parameters for the fit variables are extracted through a fit to Monte Carlo samples. In table A.1 and table A.2 the PDF shapes used for each variable and each category can be found, where `Gauss(N)` is a sum of N gaussian functions, `CB` is a Crystal Ball function [39], `Argus` is an Argus function [40], `pol(N)` is a Chebyshev polynomial of degree N and `BG` indicates a Bifurcated Gaussian.

	Signal	SxF	$q\bar{q}$	$B\bar{B}$
$M_{bc}$	CB	CB	Argus	Argus+Gauss(1)
$\Delta E$	Gauss(2)	BG+pol(2)	pol(3)/pol(2)	pol(2)
$CS_{var}$	BG	BG	BG	BG

Table A.1: PDF shapes for  $B \rightarrow \eta'(\eta\pi^+\pi^-)K$ . When two functions are indicated, the former is for the charged final state and the latter for the neutral one

	Signal	SxF	$q\bar{q}$	$B\bar{B}$
$M_{bc}$	CB	CB	Argus	Argus+Gauss(1)
$\Delta E$	Gauss(2)	Gauss(1)+pol(2)/Gauss(1)+pol(1)	pol(2)	pol(2)
$CS_{var}$	BG	BG	BG	BG

Table A.2: PDF shapes for  $B \rightarrow \eta'(\rho\gamma)K$ . When two functions are indicated, the former is for the charged final state and the latter for the neutral one

### A.0.2 Test of the fit procedure

The fit procedure has been tested using toy Monte Carlo samples. The samples for the continuum background have been generated according to its PDFs, while the signal and peaking candidates have been sampled from the large  $B\bar{B}$  Monte Carlo set.

The number of continuum and peaking background candidates has been fixed to the expected number of candidates in the data sample, given in table 5.3.

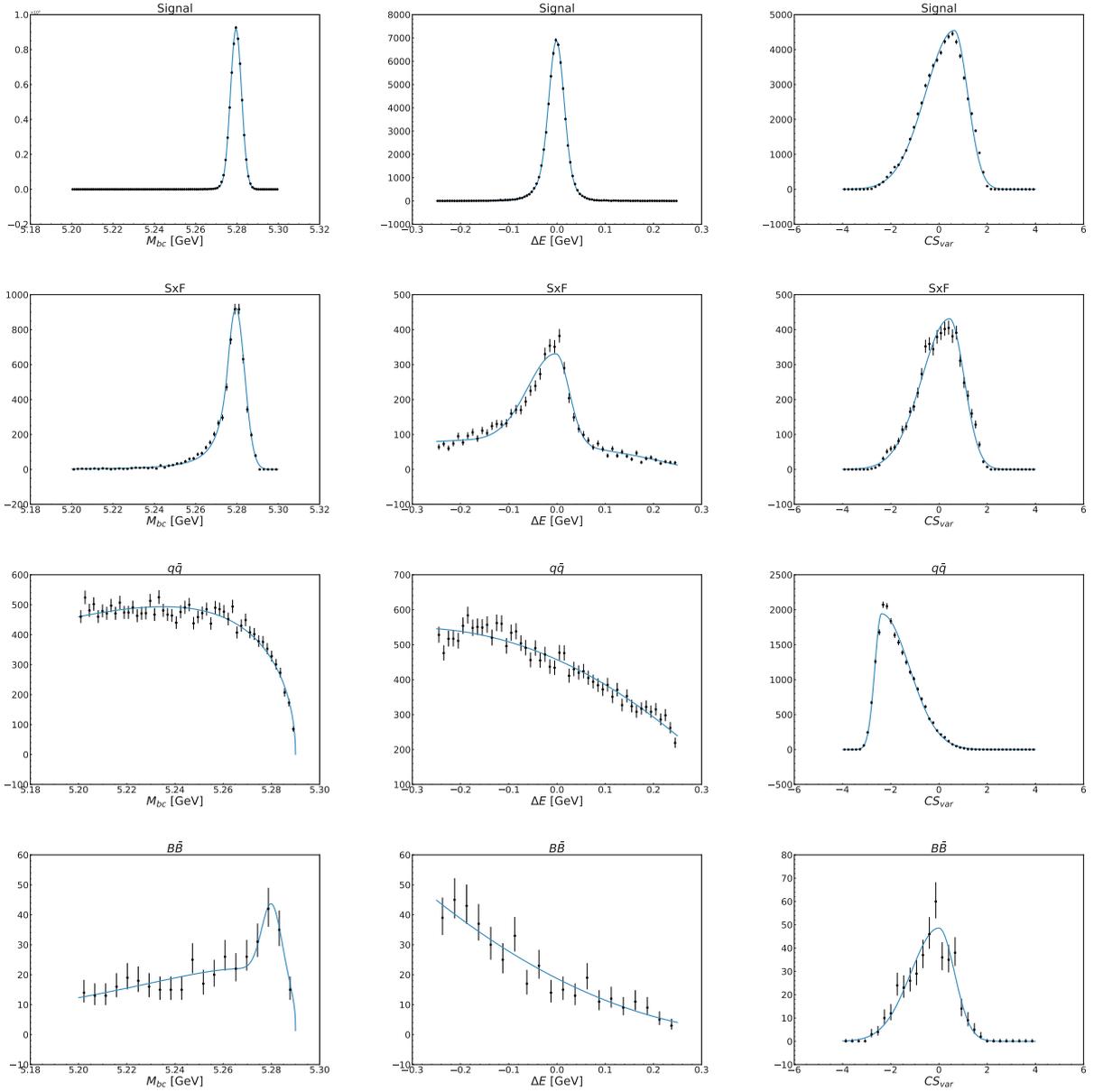


Figure A.1: Monte Carlo distribution of the fit variables  $M_{bc}$  and  $\Delta E$ , and  $CS_{var}$  (left to right) for signal, SxF, continuum background, and peaking background (top to bottom) and fitted PDF for channel  $B^0 \rightarrow \eta'(\eta\pi^+\pi^-)K_S^0$

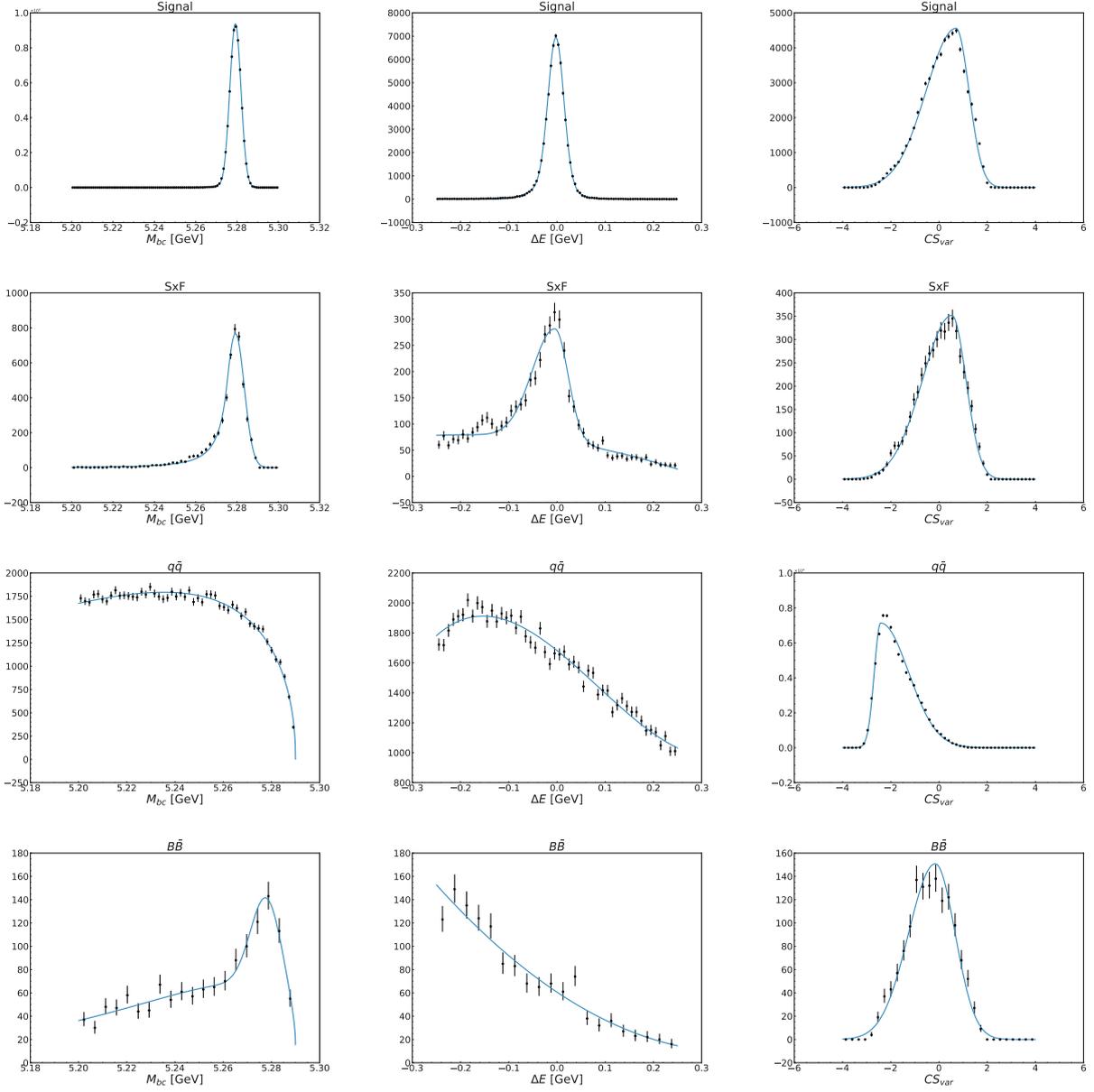


Figure A.2: Monte Carlo distribution of the fit variables  $M_{bc}$  and  $\Delta E$ , and  $CS_{var}$  (left to right) for signal, SxF, continuum background, and peaking background (top to bottom) and fitted PDF for channel  $B^\pm \rightarrow \eta'(\eta_{\gamma\gamma}\pi^+\pi^-)K^\pm$

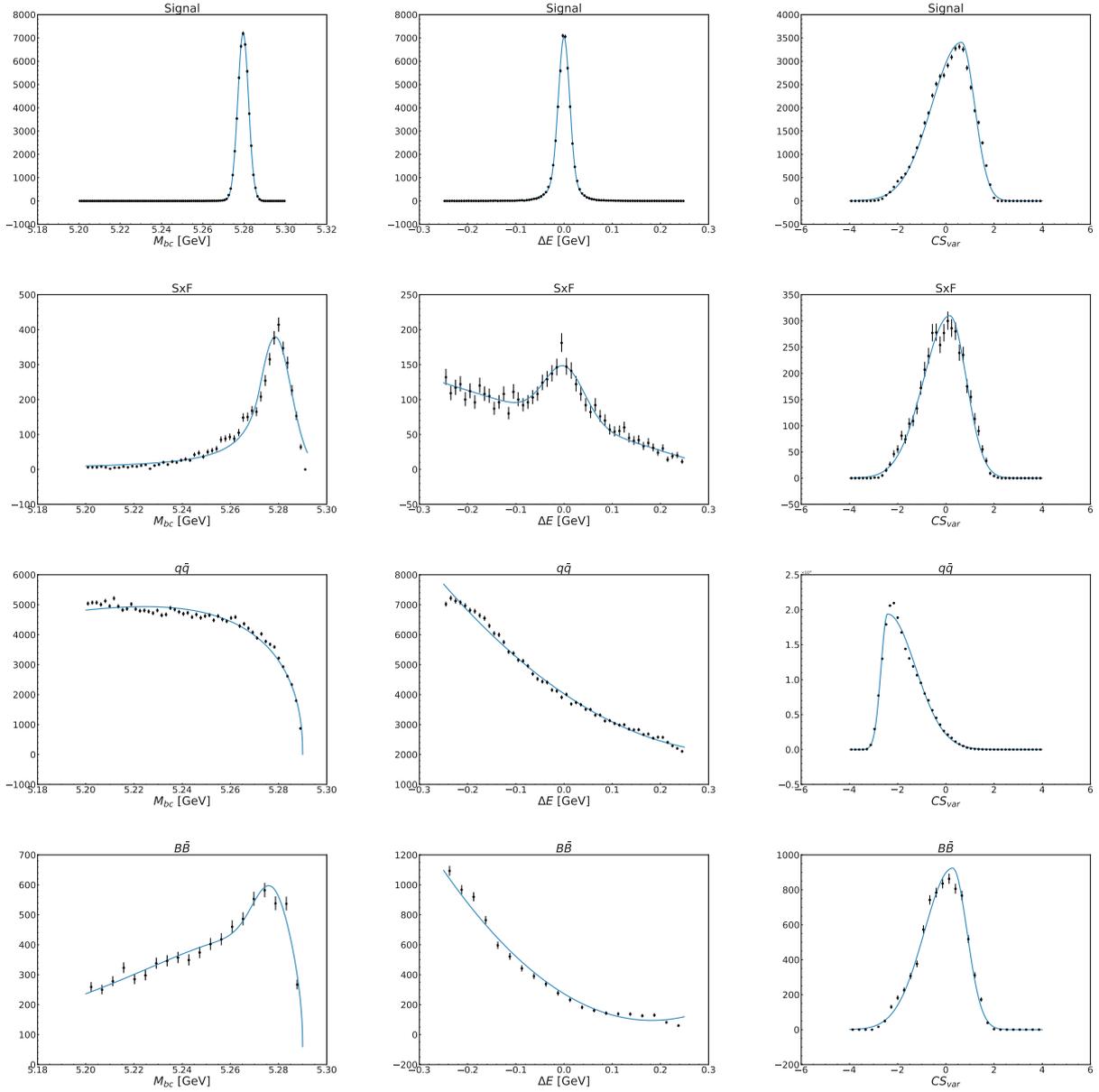


Figure A.3: Monte Carlo distribution of the fit variables  $M_{bc}$  and  $\Delta E$ , and  $CS_{var}$  (left to right) for signal, SxF, continuum background, and peaking background (top to bottom) and fitted PDF for channel  $B^0 \rightarrow \eta'(\rho\gamma)K_S^0$

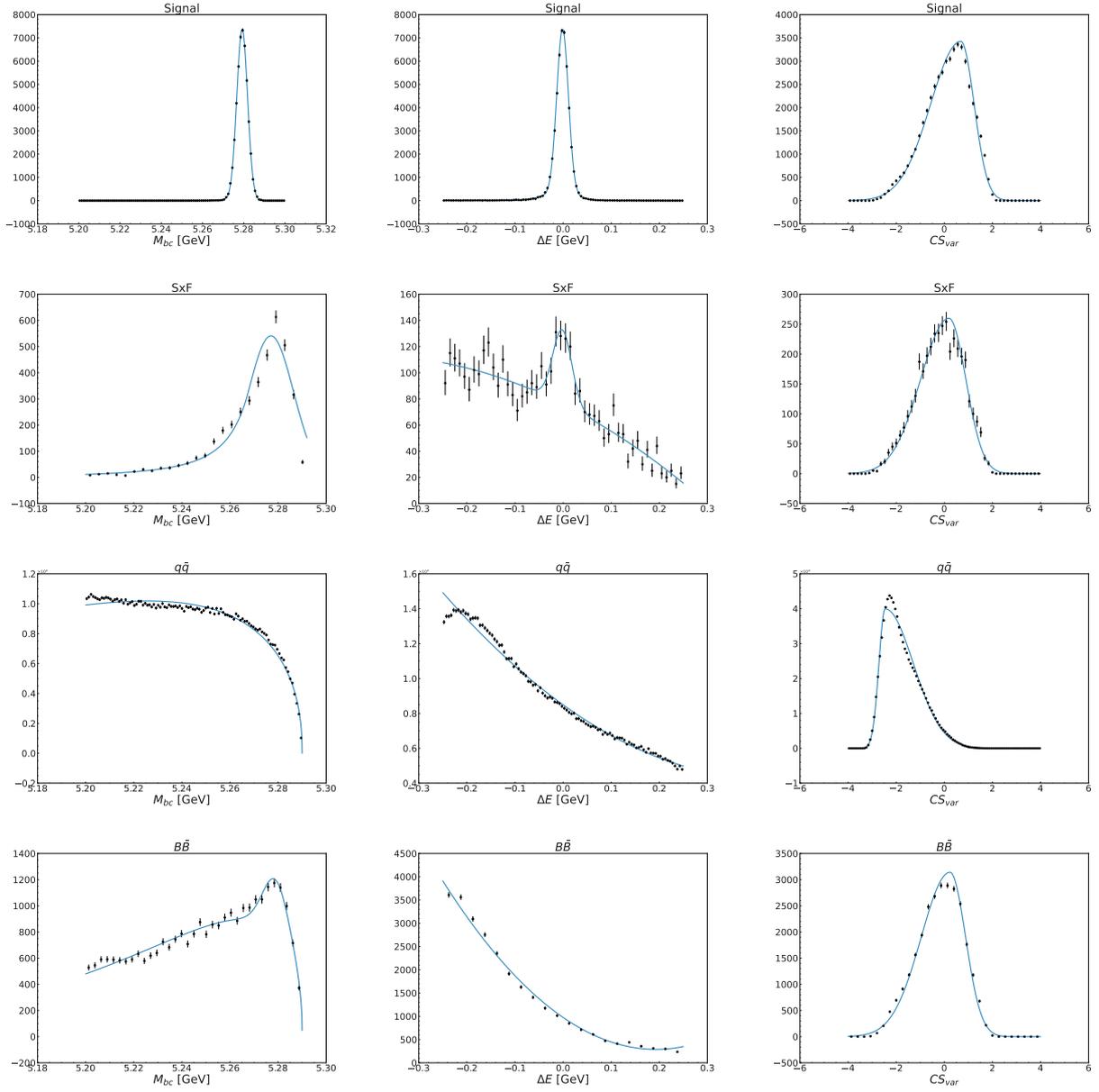


Figure A.4: Monte Carlo distribution of the fit variables  $M_{bc}$  and  $\Delta E$ , and  $CS_{var}$  (left to right) for signal, SxF, continuum background, and peaking background (top to bottom) and fitted PDF for channel  $B^\pm \rightarrow \eta'(\rho\gamma)K^\pm$

The linearity of the fit has been tested varying the number of injected signal candidates in the toy Monte Carlo sample. For each value of the number of injected signal candidates, 100 toy Monte Carlo samples have been generated and the average yield and standard deviation of the results have been computed. The results of the linearity test are shown in figure A.5. No significant bias has been observed, but the standard deviation of the results is large, especially for the channels with  $\eta' \rightarrow \rho\gamma$ , affected by a larger amount of background.

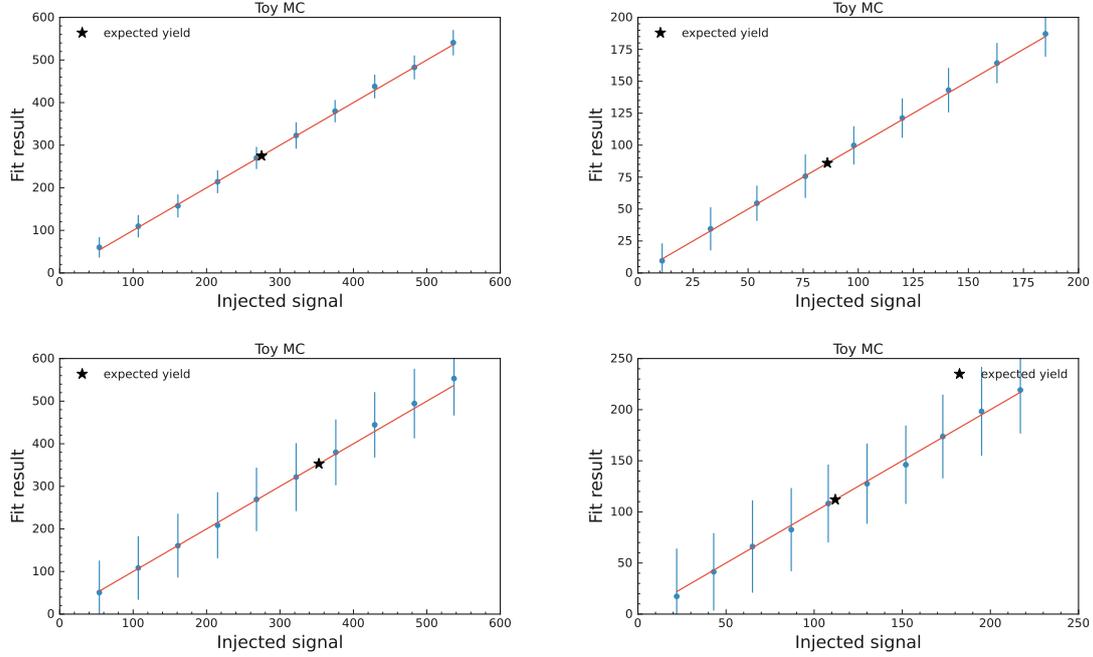


Figure A.5: Linearity test: average fit results for the signal yield as a function of the injected signal and relative pulls. The results are shown for the  $B \rightarrow \eta'(\eta_{\gamma\gamma}\pi^+\pi^-)K$  decay (top) and for the  $B \rightarrow \eta'(\rho\gamma)K$  (bottom), both for charged (left) and neutral (right) modes.

Another test has been performed fitting 600 toy Monte Carlo samples fixing the yield for each category to the expected ones, predicted from Monte Carlo, given in table 5.3. The results of this test for the signal yield are shown in fig. A.6, A.7, A.8 and A.9, while the distributions for the backgrounds yield can be found in fig. A.10, A.11 A.12 and A.13.

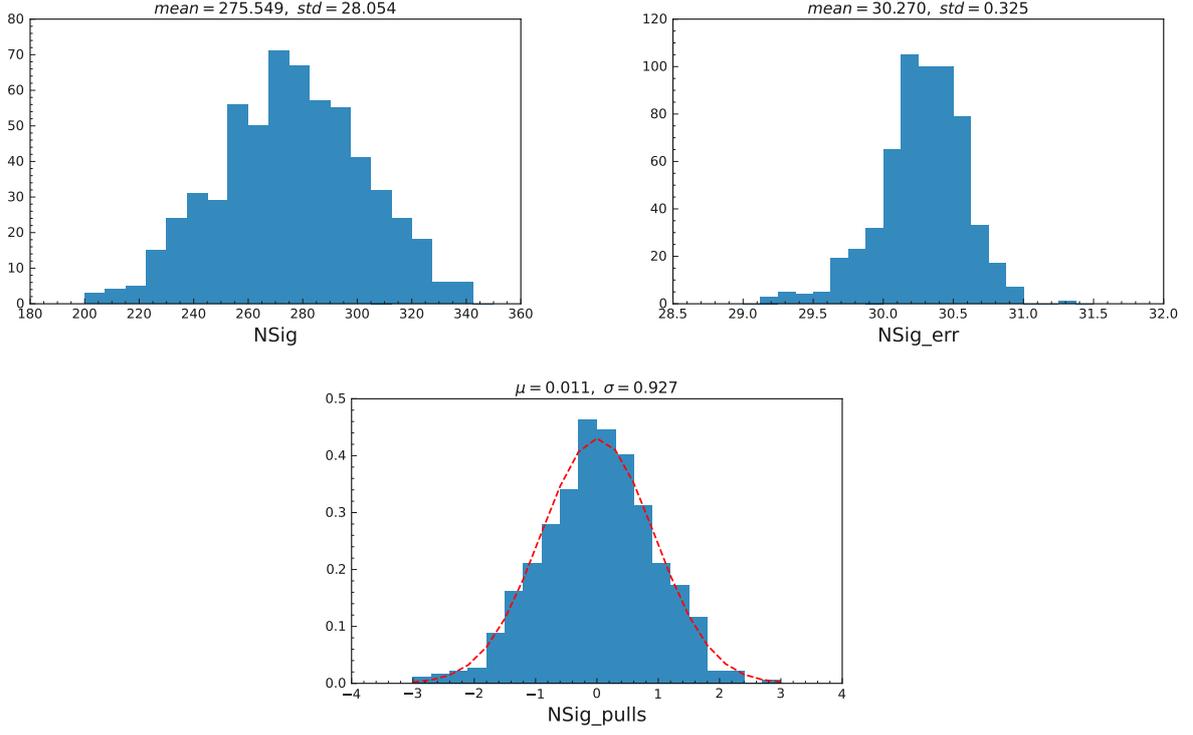


Figure A.6: Distribution of signal yield, error, and pulls for 600 toys for  $B^\pm \rightarrow \eta'(\eta_{\gamma\gamma}\pi^+\pi^-)K^\pm$ . The number of injected signal candidates is  $N_{sig} = 275$ .

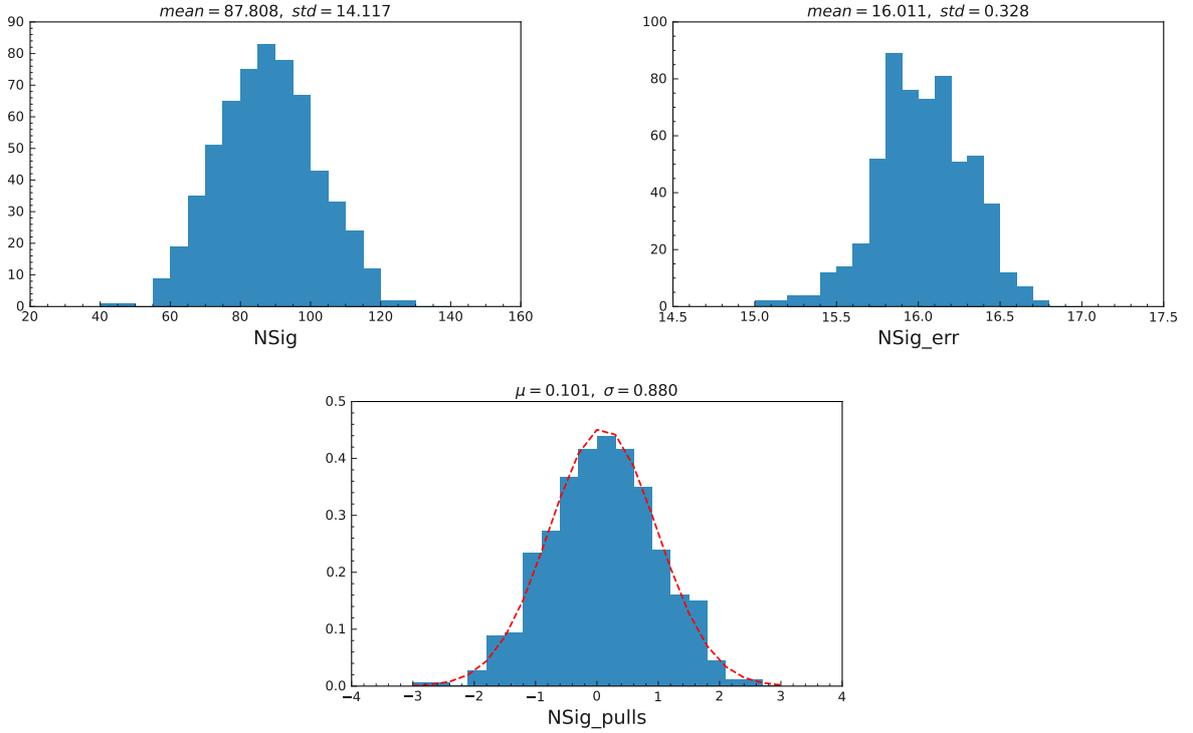


Figure A.7: Distribution of signal yield, error, and pulls for 600 toys for  $B^0 \rightarrow \eta'(\eta_{\gamma\gamma}\pi^+\pi^-)K_S^0$ . The number of injected signal candidates is  $N_{sig} = 86$ .

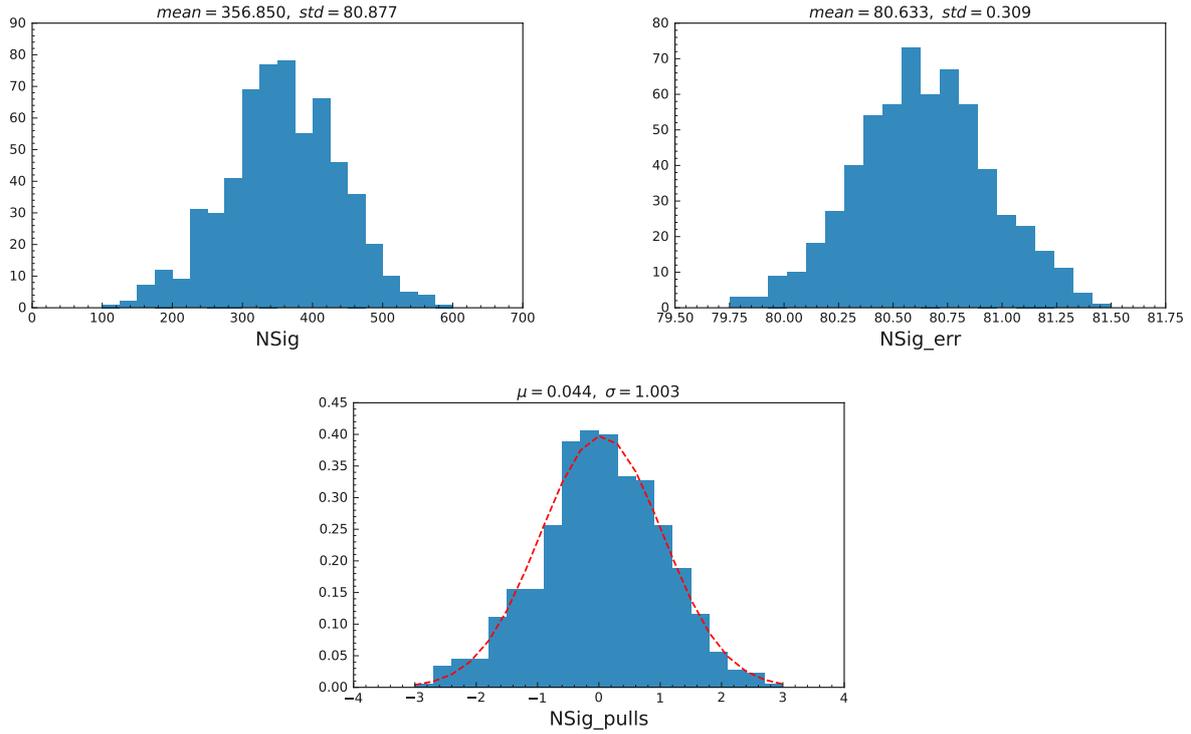


Figure A.8: Distribution of signal yield, error, and pulls for 600 toys for  $B^\pm \rightarrow \eta'(\rho\gamma)K^\pm$ . The number of injected signal candidates is  $N_{sig} = 353$ .

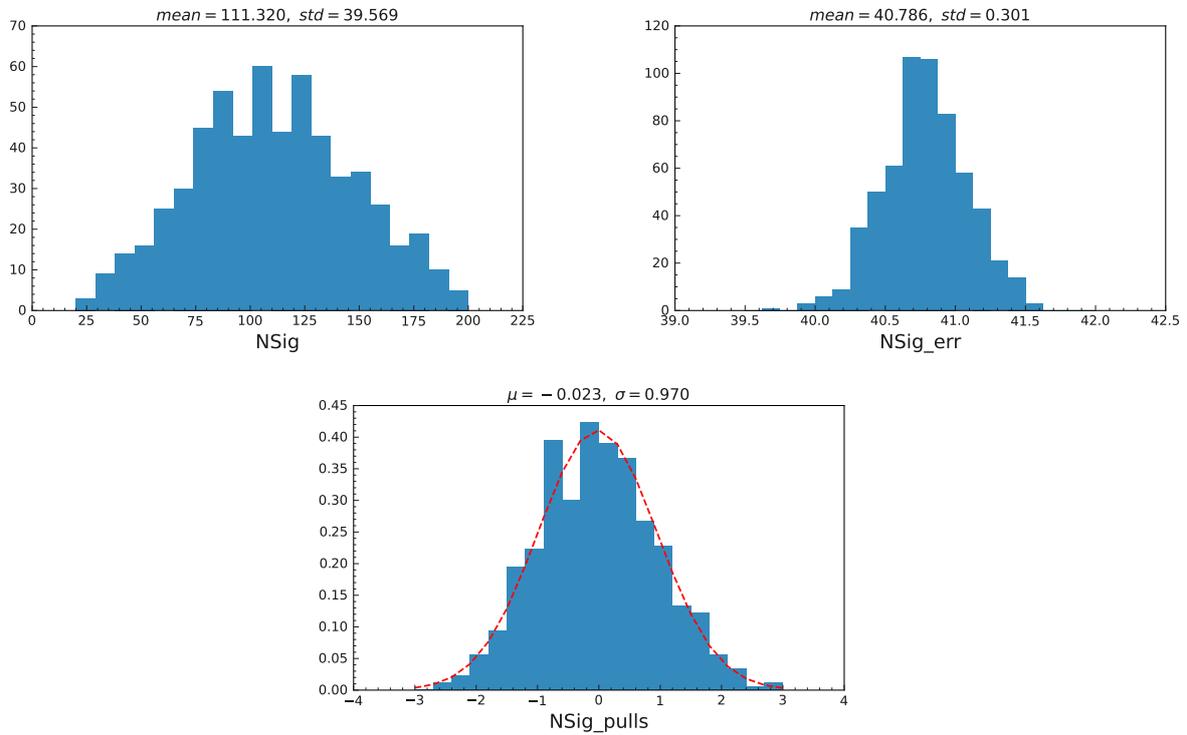


Figure A.9: Distribution of signal yield, error, and pulls for 600 toys for  $B^\pm \rightarrow \eta'(\rho\gamma)K^\pm$ . The number of injected signal candidates is  $N_{sig} = 112$ .

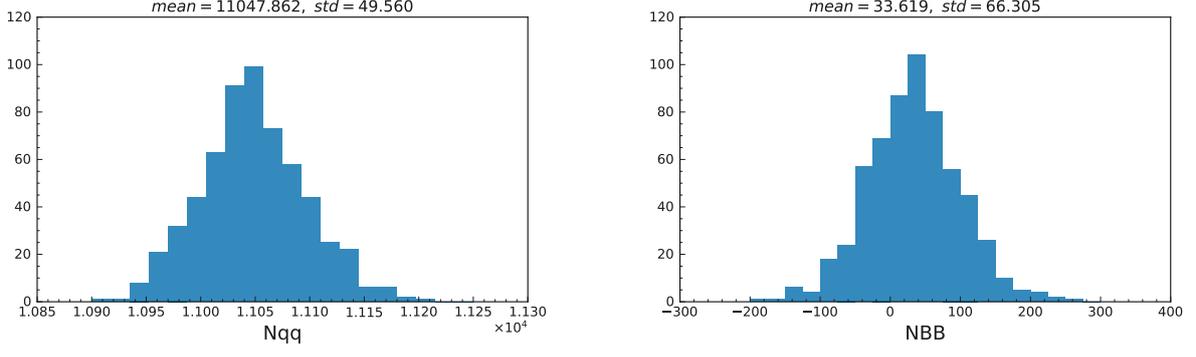


Figure A.10: Distribution of continuum background (right) and peaking background (left) yield for 600 toys for the  $B^\pm \rightarrow \eta'(\eta_{\gamma\gamma}\pi^+\pi^-)K^\pm$  decay channel

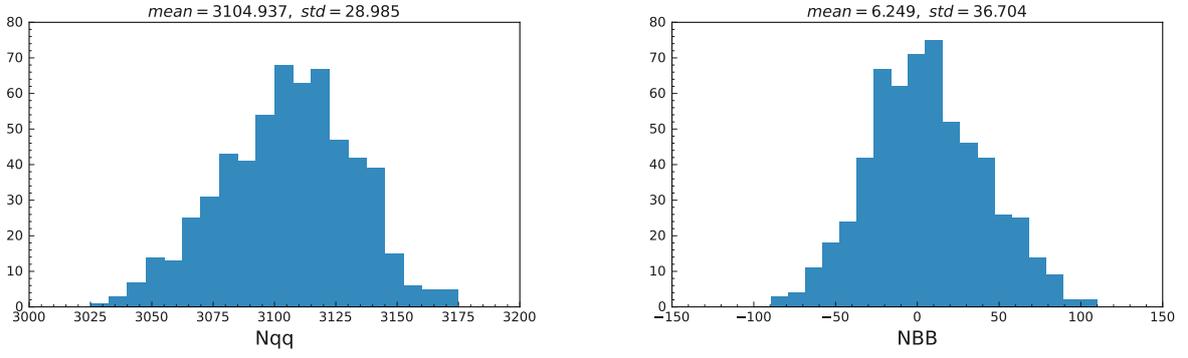


Figure A.11: Distribution of continuum background (right) and peaking background (left) yield for 600 toys for the  $B^0 \rightarrow \eta'(\eta_{\gamma\gamma}\pi^+\pi^-)K_S^0$  decay channel

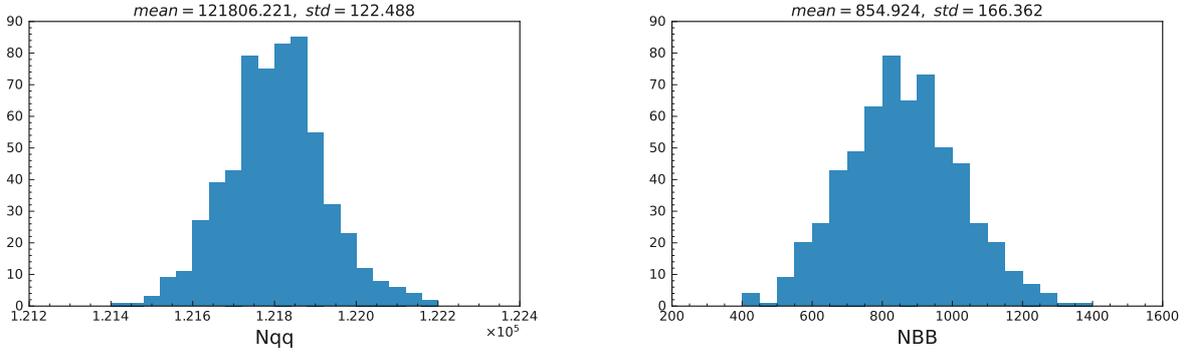


Figure A.12: Distribution of continuum background (right) and peaking background (left) yield for 600 toys for the  $B^\pm \rightarrow \eta'(\rho\gamma)K^\pm$  decay channel

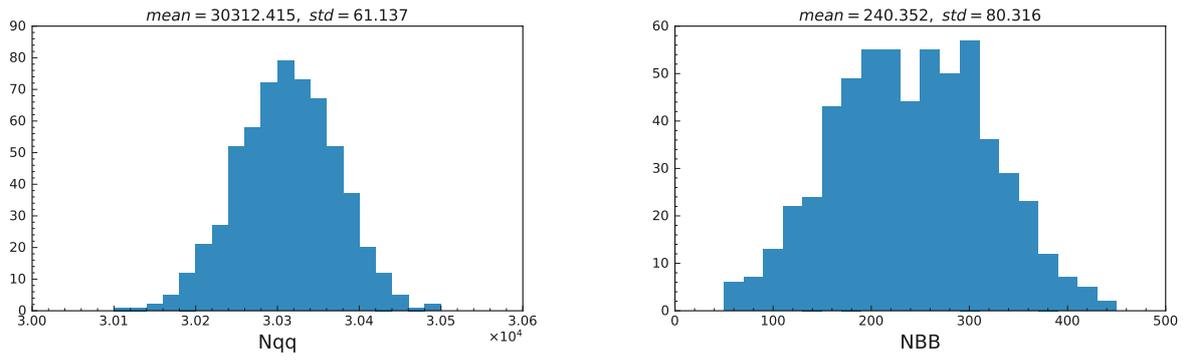


Figure A.13: Distribution of continuum background (right) and peaking background (left) yield for 600 toys for the  $B^0 \rightarrow \eta'(\rho\gamma)K_S^0$  decay channel

# Appendix B

## RooFit fit test with toy Monte Carlo

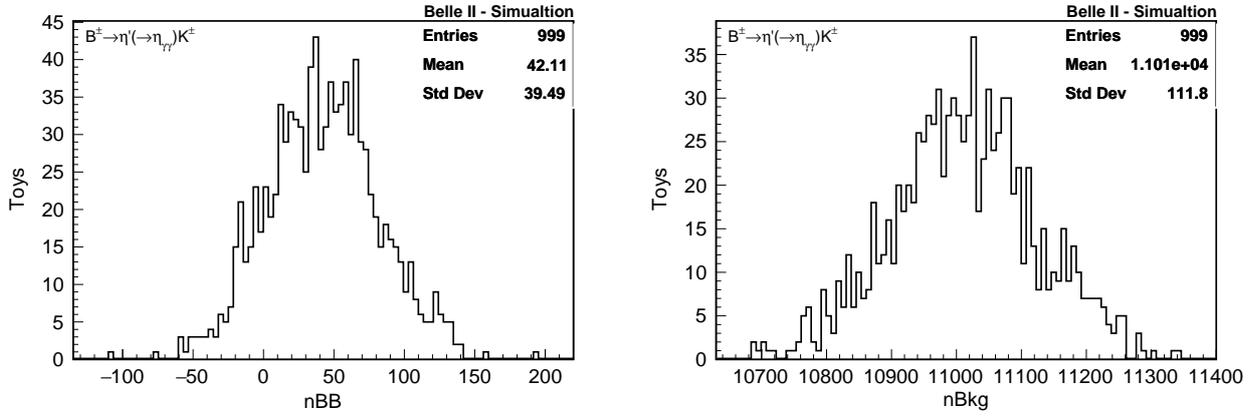


Figure B.1: Distribution of continuum background (right) and peaking background (left) yield for 1000 toys for the  $B^\pm \rightarrow \eta'(\eta_{\gamma\gamma}\pi^+\pi^-)K^\pm$  decay channel

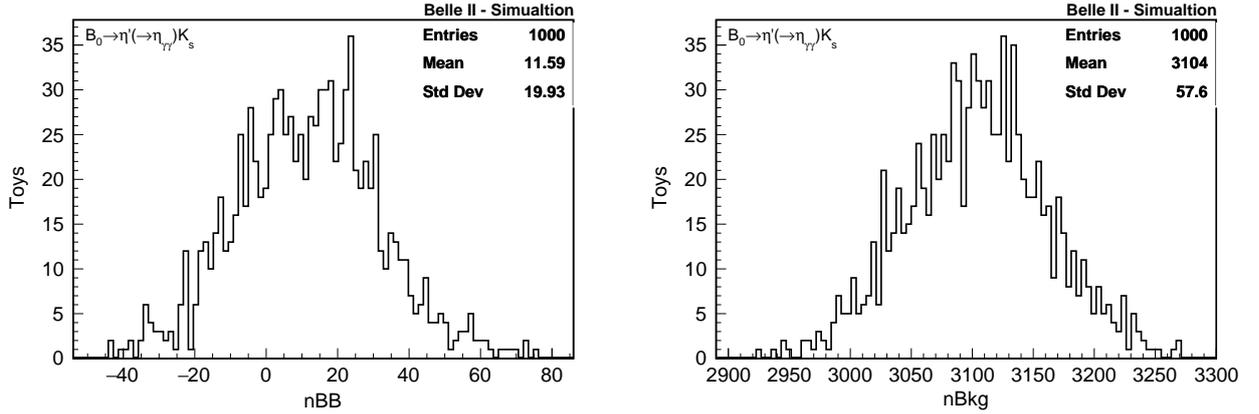


Figure B.2: Distribution of continuum background (right) and peaking background (left) yield for 1000 toys for the  $B^0 \rightarrow \eta'(\eta_{\gamma\gamma}\pi^+\pi^-)K_S^0$  decay channel

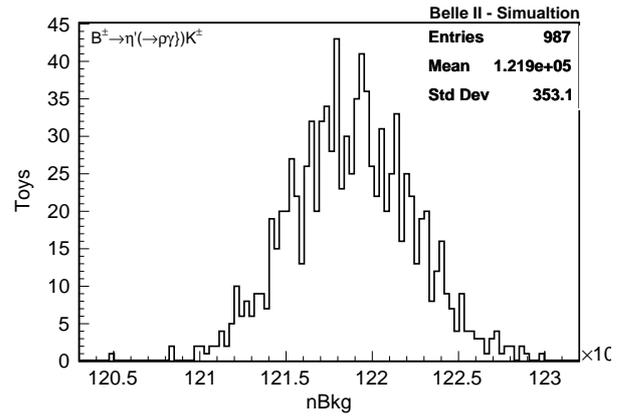
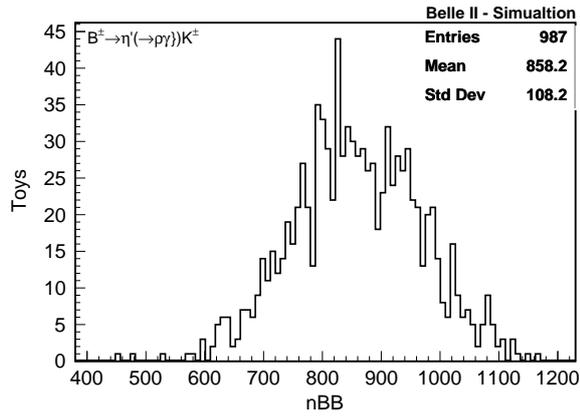


Figure B.3: Distribution of continuum background (right) and peaking background (left) yield for 1000 toys for the  $B^{\pm} \rightarrow \eta'(\rho\gamma)K^{\pm}$  decay channel

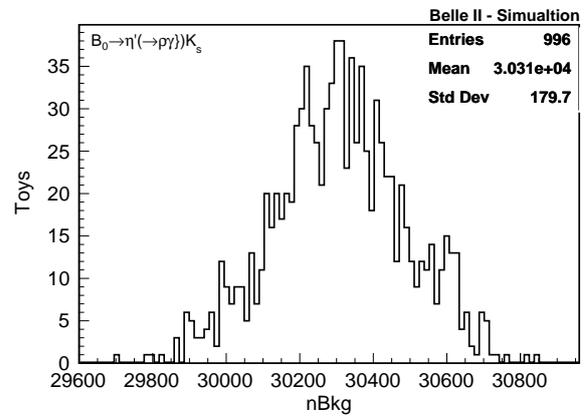
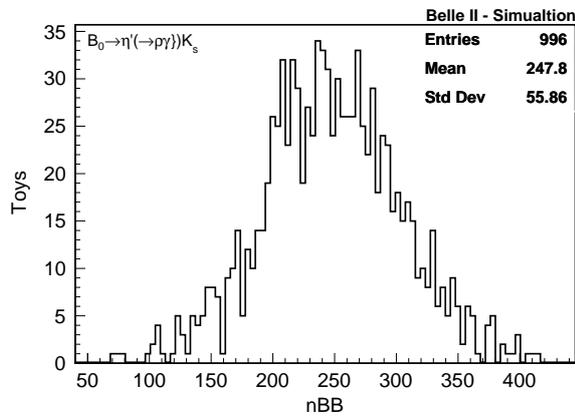


Figure B.4: Distribution of continuum background (right) and peaking background (left) yield for 1000 toys for the  $B^0 \rightarrow \eta'(\rho\gamma)K^0$  decay channel