Search for $B^0 \to \rho^0 \gamma$ decay at the LHCb experiment

Giulio Gazzoni

ED SF numéro d'ordre : 934

UNIVERSITÉ CLERMONT AUVERGNE

U.F.R. Sciences et Technologies

ÉCOLE DOCTORALE DES SCIENCES FONDAMENTALES

THÈSE

présentée pour obtenir le grade de

Docteur d'Université

Spécialité : PHYSIQUE DES PARTICULES

par

GAZZONI Giulio

Search for $B^0 \to \rho^0\gamma$ decay at the LHCb experiment

Recherche de la désintégration $B^0 \to \rho^0\gamma$ dans l'expérience LHCb

(sous la direction de Régis LEFÈVRE et Pascal PERRET)

Soutenue publiquement le 15 décembre 2017, devant la commission d'examen :

Président :	Dominique PALLIN	
Examinateurs :	Angelo CARBONE	Rapporteur
	Maria Aranzazu DE OYANGUREN CAMPOS	Rapporteur
	Emi KOU	
	Mossadek TALBY	
Directeur de thèse :	Pascal PERRET	

Résumé

L'expérience LHCb installée auprès du LHC, grâce à la grande section efficace de production de paires $b\bar{b}$, offre la possibilité de conduire des mesures de précision dans le secteur des hadrons beaux. Dans cette thèse une étude de la désintégration radiative $B^0 \to \rho^0\gamma$ est présentée. Cette désintégration fait intervenir une transition $b \to d\gamma$, un processus d'ordre supérieur qui correspond à un changement de saveur par courants neutres. L'objectif de ce travail est de mesurer le rapport des rapports d'embranchements $\mathcal{B}(B^0 \to \rho^0\gamma)/\mathcal{B}(B^0 \to K^{*0}\gamma)$. Ce dernier est lié au rapport des éléments de la matrice de Cabibbo-Kobayashi-Maskawa $|V_{td}/V_{ts}|$ et par conséquent permet de tester le Modèle Standard de la Physique des Particules. La mesure utilise la statistique complète du Run 1 collectée par LHCb. Elle est conduite à l'aveugle pour éviter d'introduire de possibles biais d'analyse. Une sélection dédiée à maximiser la significance attendue de $B^0 \to \rho^0\gamma$ ainsi qu'une modélisation simultanée des spectres de masse invariante $B^0 \to \rho^0\gamma$ et $B^0 \to K^{*0}\gamma$ ont été développées. Les échantillons de données 2011 et 2012 sont modélisés simultanément afin d'extraire directement le rapport des rapports d'embranchements. L'étude des principales incertitudes systématiques est présentée. La mesure préliminaire est compatible avec la combinaison des mesures de BaBar et Belle et significativement plus précise.

Mots Clés:

Expérience LHCb - Matrice CKM - Modèle Standard - Physique des Saveurs Lourdes - Courants Neutres Changeant la Saveur (FCNC) - Désintégrations radiatives des mésons beaux - $B^0 \to \rho^0\gamma$ **-** $B^0 \to K^{*0}\gamma$.

Abstract

The LHCb experiment at the LHC, by exploiting the high production cross section of $b\bar{b}$ pairs, offers the possibility to perform precision measurements in the sector of b-hadrons. In this thesis a study of the $B^0 \to \rho^0\gamma$ radiative decay is presented. Such decay occurs via a $b \to d\gamma$ transition, which is a non tree-level process corresponding to a Flavour Changing Neutral Current transition. The goal of this work is to measure the ratio of branching fractions $\mathcal{B}(B^0 \to \rho^0\gamma)/\mathcal{B}(B^0 \to K^{*0}\gamma)$. This ratio of branching fractions is related to the ratio of the Cabibbo-Kobayashi-Maskawa matrix elements $|V_{td}/V_{ts}|$ and therefore allows to test the Standard Model of Particle Physics. The measurement is performed using the full Run 1 statistics recorded at LHCb and putting a blind on the signal region in order to avoid possible analysis biases. A dedicated selection to improve the $B^0 \to \rho^0\gamma$ expected significance and a simultaneous fit to the $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ mass spectra have been developed. The 2011 and 2012 data samples are fit simultaneously in order to directly extract the ratio of branching fractions. The study of the main systematic uncertainties is presented. The preliminary measurement is compatible with the combination of BaBar and Belle measurements and significantly more precise.

Keywords:

LHCb experiment - CKM matrix - Standard Model - Heavy Flavour Physics - Flavour Changing Neutral Currents (FCNC) - Radiative B decays - $B^0 \to \rho^0\gamma$ - $B^0 \to K^{*0}\gamma$.

Contents

Introduction

The Standard Model of Particle Physics is a very successful theory that provides predictions in very good agreement with the experimental observations. However still numerous open questions remain to justify the search for New Physics (NP). The most straightforward way to search for NP consists of producing and observing directly new particles in high energy colliders like the Large Hadron Collider (LHC). Nevertheless, the maximum center-of-mass energy available is the limiting factor: the higher the energy, the heavier the particle that one can produce, thus probing higher NP scales. An alternative way to search for NP is measuring the effects of virtual heavy particles exchanged at loop level in weak decays of SM particles, where the production energy threshold is not an issue: the higher the precision of the measurements, the higher the scale of NP that can be probed.

The LHCb experiment installed at the LHC conducts an extensive campaign of precision measurement in the field of Flavour Physics, by exploiting the high production cross section of $c\bar{c}$ and $b\bar{b}$ quark pairs. Its approach is complementary to the one of the main purpose LHC experiments, ATLAS and CMS.

Flavour Changing Neutral Currents (FCNC), in particular decays of B mesons with a photon in the final state, are a promising place to probe for NP. These transitions are non tree-level processes that correspond at quark level to $b \to q\gamma$ transitions and appear in the SM through electroweak penguin diagrams in which the main contribution is represented by the propagation of a top quark in the loop. $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ radiative decays, which respectively correspond to a $b \to d\gamma$ and $b \to s\gamma$ transition, allow to measure the magnitudes of the Cabibbo-Kobayashi-Maskawa (CKM) matrix elements V_{td} and V_{ts}. These matrix elements also contribute to the oscillations of $B^0 - \overline{B}^0$ and $B_s^0 - \overline{B}_s^0$ systems.

In this thesis the Run 1 (2011 and 2012) data collected at LHCb, corresponding to an integrated luminosity of 3 fb^{-1}, are used to measure the ratio of branching fractions $\mathcal{B}(B^0 \to \rho^0\gamma)/\mathcal{B}(B^0 \to K^{*0}\gamma)$, which is proportional to the square of the ratio $|V_{td}/V_{ts}|$. The result coming from such measurement can be compared with the ones arising from the measurement of the oscillation frequencies of $B^0 - \overline{B}^0$ and $B_s^0 - \overline{B}_s^0$ systems, thus allowing to test the standard model consistency.

This thesis manuscript is organized as follows. In the first chapter a theoretical overview of the Standard Model and the CKM matrix is given. Additionally two sections are dedicated to the description of $B - \overline{B}$ oscillations and $b \to q\gamma$ transitions, putting emphasis on the observables of interest.

The second chapter presents a description of the general characteristics of the LHC accelerator and of the LHCb detector. Concerning LHCb the focus is put on the various subdetectors and in particular on the ones playing a major role in the photons reconstruction.

The third chapter describes the basic concepts of neutral objects reconstruction, namely photons (γ) and neutral pions (π^0). A description of the procedure implemented to provide the Run 1 calibration samples necessary to evaluate the γ/π^0 separation efficiency with a data-driven approach is also given. In particular we discuss how these samples were produced detailing both the event selection and the background subtraction procedure.

The fourth chapter gives a brief overview of the experimental status of the ratio of branching fractions $\mathcal{B}(B^0 \to \rho^0\gamma)/\mathcal{B}(B^0 \to K^{*0}\gamma)$ measurement and describes the overall strategy used to perform the measurement. A key point is that, in order to avoid possible biases, the analysis is conducted blindly. This is a standard procedure in LHCb and aims at preventing both analysts and reviewers to introduce systematic biases on the measurements by knowing the measured value.

The fifth chapter describes the full event selection developed in order to achieve the best possible rejection of the various backgrounds. This is a crucial step in the analysis given the $B^0 \to \rho^0\gamma$ decay is a particularly rare mode.

The sixth chapter details the study of the various background sources and the first validation of the mass fit model exploiting the $B^0 \to K^{*0}\gamma$ control channel.

The seventh chapter details the baseline mass fit model used to extract the measurement. In particular, a simultaneous fit to the $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ mass spectra is implemented and 2011 and 2012 data samples are fitted at the same time. This is aimed to improve the statistical precision and to have a better control on the final result. The validation of the simultaneous fit model and the expected sensitivity are also discussed.

Finally, the last chapter details the preliminary results on the measurement of the ratio of branching fractions $\mathcal{B}(B^0 \to \rho^0\gamma)/\mathcal{B}(B^0 \to K^{*0}\gamma)$. Such results were obtained after having received the authorization of the LHCb collaboration to perform the private unblinding of the simultaneous mass fit. Several crosschecks needed to confirm the consistency of the measurement are also performed and a preliminary study of the main systematic uncertainties is presented.

Chapter 1

Theoretical framework

1.1 The Standard Model

The Standard Model (SM) of particle physics is the theory used to describe interactions among the fundamental constituents of matter, namely quarks, leptons and their antiparticles [1]. The SM provides predictions which are in very good agreement with experimental data and it has been successfully confirmed through four decades. This theory allows to describe in a unified way the strong, weak and electromagnetic interactions within the extended gauge symmetry group

$$G_{SM} = SU(3)_C \times [SU(2)_L \times U(1)_Y] \ , \tag{1.1}$$

composed by:

- the $SU(3)_C$ group which describes the strong interaction mediated by 8 massless gluons carrying color charge;
- the group $SU(2)_L \times U(1)_Y$ which describes the electroweak interaction associated to the three massive gauge bosons $W^{\pm}$, Z^0 and the photon γ.

The subscript Y stands for the weak hypercharge defined through the Gell-Mann-Nishijima relation $Q = I_3 + Y/2$: Q is the electric charge and I_3 is the third component of the weak isospin. The SM is a gauge theory invariant under *local transformations* of its symmetry group and renormalizable in such a way that couplings and masses correspond to experimental data.

The SM contains three fermion generations, each consisting of five representations of G_{SM} [2]

$$Q^I_{Li}(3,2)_{+1/6}, \quad u^I_{Ri}(3,1)_{+2/3}, \quad d^I_{Ri}(3,1)_{-1/3}, \quad L^I_{Li}(1,2)_{-1/2}, \quad l^I_{Ri}(1,1)_{-1} \tag{1.2}$$

and a single scalar representation composed by two complex scalar fields

$$\phi(1,2)_{+1/2} = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \ . \tag{1.3}$$

This notation means for instance that left-handed quarks $Q_L^I = (u_L^I, d_L^I)^T$ (the superscript T stands for transposed) are triplets of $SU(3)_C$, doublets of $SU(2)_L$ and carry hypercharge $Y = +1/6$. The super-index I denotes interaction eigenstates. The sub-index $i = 1, 2, 3$ is the flavour (or generation) index. The u_R^I and d_R^I singlets are generic right-handed up-type and down-type quarks, $L_L^I = (\nu_L^I, e_L^I)^T$ are three $SU(2)_L$ leptonic doublets and l_R^I are three right-handed charged leptonic singlets. The summary table of the fundamental particles of the SM is shown in Fig. 1.1. As can be noticed, the table also shows the Higgs

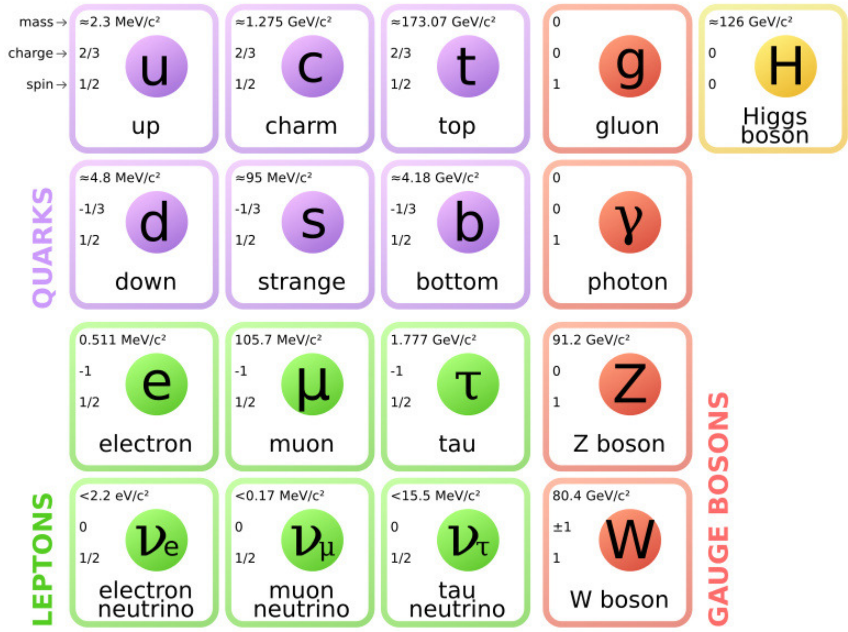


Figure 1.1: Standard Model fundamental fermions and bosons and their properties.

boson that was observed for the first time in 2012 at the LHC by ATLAS [3] and CMS [4] experiments. Afterwards, both ATLAS [5] and CMS [6] confirmed the scalar nature of the new particle. The value of the Higgs boson mass reported on the Particle Data Group (PDG) is $m_H = 125.09 \pm 0.24$ GeV [7].

The most general renormalizable lagrangian that describes the SM interactions can be divided into four different parts:

$$\mathcal{L}_{SM} = \mathcal{L}_{Gauge} + \mathcal{L}_{Kinetic} + \mathcal{L}_{Higgs} + \mathcal{L}_{Yukawa} \ . \tag{1.4}$$

The first term is related to the propagation of the gauge fields and is given by

$$\mathcal{L}_{Gauge} = -\frac{1}{4}G^a_{\mu\nu}(G^a)^{\mu\nu} - \frac{1}{4}W^d_{\mu\nu}(W^d)^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}, \tag{1.5}$$

with

$$G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu + g_s f_{abc} G^b_\mu G^c_\nu \ , \tag{1.6}$$
$$W^d_{\mu\nu} = \partial_\mu W^d_\nu - \partial_\nu W^d_\mu + g\epsilon_{def} W^e_\mu W^f_\nu \ , \tag{1.7}$$
$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \ , \tag{1.8}$$

where:

- $G^a_{\mu\nu}$ is the Yang-Mills tensor which involves eight $(a = 1, 2, \ldots, 8)$ gluon fields G^a_μ, the strong coupling constant g_s and the $SU(3)_C$ structure constants f_{abc}.
- $W^d_{\mu\nu}$ is the weak field tensor which involves three $(d = 1, 2, 3)$ gauge fields W^d_μ, the weak coupling constant g and the $SU(2)_L$ structure constants ϵ_{def}.
- $B_{\mu\nu}$ is the electromagnetic tensor which involves the $U(1)_Y$ gauge field B_μ.

Permutation terms $f_{abc}G^b_\mu G^c_\nu$ and $\epsilon_{def}W^e_\mu W^f_\nu$ in $G^a_{\mu\nu}$ and $W^d_{\mu\nu}$ reflect a crucial property of strong and weak interactions, namely they are non-abelian theories in which gauge fields can directly couple and self-interact. This is not true for the electromagnetic interaction which is an abelian theory and forbids the photon self-interaction.

The second term of $\mathcal{L}_{SM}$ describes the kinetic energy of fermions and their interaction with the gauge fields. It is given by

$$\mathcal{L}_{Kinetic} = \overline{\psi}\gamma^\mu i D_\mu \psi \ , \tag{1.9}$$

where γ^μ are the Dirac matrices, $\overline{\psi} = \psi^\dagger\gamma^0$ and the spinor fields ψ are the three fermion generations of Eq. 1.2. D_μ is the total covariant derivative defined as

$$D_\mu = \underbrace{\partial_\mu + \frac{ig}{2}W^d_\mu\sigma_d + \frac{ig'}{2}B_\mu Y}_{\nabla_\mu} + \frac{ig_s}{2}G^a_\mu\lambda_a \ . \tag{1.10}$$

where σ_d, Y and λ_a are respectively the generators of $SU(2)_L$, $U(1)_Y$ and $SU(3)_C$ symmetry groups, g' is the hypercharge coupling constant and ∇_μ stands for the covariant derivative only in the electroweak sector. σ_d are the 2×2 Pauli matrices and λ_a are the 3×3 Gell-Mann matrices.

The third part of Eq. 1.4, the Higgs lagrangian, describes the mechanism of spontaneous electroweak symmetry breaking (EWSB) through which the gauge bosons mediators of the weak interaction acquire masses. This lagrangian is written as

$$\mathcal{L}_{Higgs} = (\nabla_\mu \phi)^\dagger (\nabla^\mu \phi) + \underbrace{\mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2}_{V(\phi^\dagger \phi)} , \tag{1.11}$$

where the first term expresses the kinetic energy of the Higgs field and its gauge interactions. The two other terms, the mass term and the self-interaction term, represent the Higgs potential $V(\phi^\dagger\phi)$: μ and λ are free parameters named respectively "mass" and "quartic coupling". The essential aspect of the Higgs mechanism is the introduction of the complex scalar doublet ϕ of Eq. 1.3, which modifies the vacuum state making it not symmetrical. This field is present everywhere in the space-time and weakly self-interacting. By means of this field the masses of all particles are dynamically generated through their interaction with ϕ. The Higgs boson is nothing but the excitation of this field. For $\mu^2 < 0$ and $\lambda > 0$ the potential $V(\phi^\dagger\phi)$ has the shape of a "mexican hat" (see Fig. 1.2) and the vacuum state $\phi = 0$ becomes a local maximum which disturbs the symmetry of the system, making the configuration unstable. This way, by setting $\phi^+ = 0$, $\phi^0 = v$ and $Y = 1$, the Higgs field acquires a non-trivial vacuum expectation value $\langle\phi\rangle_0$ given by

$$\langle\phi\rangle_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \qquad \text{with} \qquad v = \sqrt{\frac{-\mu^2}{\lambda}} \simeq 246 \text{ GeV} , \tag{1.12}$$

that causes the EWSB and the SM gauge symmetry breaking $G_{SM} \longrightarrow SU(3)_C \times U(1)_{EM}$.

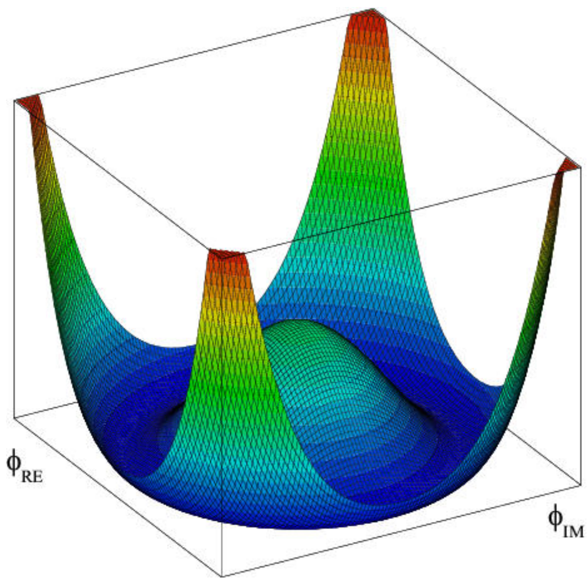


Figure 1.2: Shape of the Higgs potential for $\mu^2 < 0$ and $\lambda > 0$ as a function of $\phi_{RE} = \mathfrak{Re}\,(\phi)$ and $\phi_{IM} = \mathfrak{Im}\,(\phi)$.

By means of this dynamics the $W^\pm$ and Z^0 bosons acquire mass together with the Higgs boson H, that has mass $m_H = \sqrt{-2\mu^2} = \sqrt{2\lambda}v$. Tab. 1.1 shows the masses and mass-eigenstates of the weak bosons together with the ones of the photon ($A = \gamma$) [1].

Boson	Mass-eigenstate	Mass
$W^\pm_\mu$	$\frac{1}{\sqrt{2}}(W^1_\mu \mp iW^2_\mu)$	$m_W = \frac{gv}{2}$
Z_μ	$\cos\theta_w W^3_\mu - \sin\theta_w B_\mu$	$m_Z = \frac{m_W}{\cos\theta_w}$
A_μ	$\sin\theta_w W^3_\mu + \cos\theta_w B_\mu$	$m_A = 0$

Table 1.1: Mass-eigenstates and masses of the fundamental vector bosons; θ_w is the Weinberg mixing angle.

The last term of $\mathcal{L}_{SM}$ is the Yukawa lagrangian which describes the interactions between fermions and the Higgs field. It is given by

$$\mathcal{L}_{Yukawa} = -Y^d_{ij}\overline{Q^I_{Li}}\phi d^I_{Rj} - Y^u_{ij}\overline{Q^I_{Li}}\tilde{\phi} u^I_{Rj} - Y^l_{ij}\overline{L^I_{Li}}\phi l^I_{Rj} + h.c. \;, \tag{1.13}$$

where $\tilde{\phi} = i\sigma_2\phi^\dagger$, $Y^{u,d,l}$ are 3×3 complex matrices and i, j are generation labels. Neglecting leptons, when EWSB occurs the lagrangian of Eq. 1.13 yields mass terms for quarks. The physical states can be obtained diagonalizing the Y^f matrix using four unitary matrices $V^f_{L,R}$ such that

$$M^f_{diag} = \frac{v}{\sqrt{2}} V^f_L Y^f V^{f\dagger}_R \;, \tag{1.14}$$

where $f = u, d$ and $v/\sqrt{2} = \langle\phi\rangle_0$. As a consequence of this diagonalization, the charged current interactions are then given by

$$\mathcal{L}_{W^\pm} = -\frac{g}{\sqrt{2}}\overline{u}_{Li}\gamma^\mu \left(V^u_L V^{d\dagger}_L\right)_{ij} d_{Lj} W^+_\mu + h.c. \;, \tag{1.15}$$

and weak-eigenstates and mass-eigenstates of quarks become mixed. The product $\left(V^u_L V^{d\dagger}_L\right) = V_{CKM}$, that contains the couplings of an up-type antiquark and a down-type quark to the charged W bosons, is called Cabibbo-Kobayashi-Maskawa (CKM) matrix [8], [9] and allows to write the interaction-eigenstates as

$$\begin{pmatrix} d^I \\ s^I \\ b^I \end{pmatrix} = V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} . \tag{1.16}$$

By convention, the interaction-eigenstates and the mass-eigenstates are chosen to be equal for the up-type quarks, whereas the down-type quarks are chosen to be rotated. This important matrix will be extensively discussed in the following.

1.2 Discrete symmetries: the $\mathcal{C}, \mathcal{P}$ and $\mathcal{T}$ transformations

Symmetries are a very important topic in quantum field theory since they help restricting classes of models, providing stability and simplifying calculations as well as results. Symmetries, which are broadly divided into spacetime and internal ones, are some transformations of fields and coordinates that map solutions of the equations of motion to other, allowing to generate a whole class of solutions from a single one [10].

A spacetime symmetry is a transformation that acts directly on spacetime coordinates, like a Poincaré translation

$$x'^{\mu} = \Lambda^{\mu}{}_{\nu}(x^{\nu} + a^{\nu}),$$

where a^{ν} is an arbitrary constant four vector and $\Lambda^{\mu}{}_{\nu}$ is the Lorentz matrix. An example of internal symmetry is a global phase transformation which maps a field into another without making reference to its spacetime dependence:

$$\psi'(x) = e^{iq\alpha}\psi(x), \qquad \overline{\psi'}(x) = \overline{\psi}(x)e^{-iq\alpha}, \qquad 0 \leq \alpha < 2\pi,$$

where q is the particle electrical charge. Unlike spacetime symmetries, internal symmetry transformations involve peculiar field degrees of freedom such as the electric charge, the weak charge and the colour charge. From the Noether's theorem, if these transformations of coordinates and fields make null the variation of the action $S = \int \mathcal{L}\ \mathrm{d}^4x$, they are called *continuous symmetries* and correspond to definite constants of motion like energy or electrical charge.

In addition to these important symmetries, there is another relevant class of transformations that belongs to the category of internal ones, namely the *discrete symmetries* such as $\mathcal{C}$, $\mathcal{P}$ and $\mathcal{T}$ transformations. These operations are named charge conjugation, parity transformation and time reversal: their transformation rules will be shown only for the spinor field, neglecting bilinears like mass terms $\overline{\psi}(x)\psi(x)$ and current terms $\overline{\psi}(x)\gamma^{\mu}\psi(x)$.

The charge conjugation $\mathcal{C}$ is the operation under which a particle is transformed into its antiparticle of equal mass, momentum and spin, but opposite quantum numbers like electric charge. Its action is given by $\psi^{\mathcal{C}}(x) = \mathcal{C}\psi(x)\mathcal{C}^{\dagger}$ with $\mathcal{C} = \mathcal{C}^{\dagger} = \mathcal{C}^{-1}$ and $\mathcal{C}^2 = \mathbb{I}$ and the transformation rules for the fields are [11]

$$\begin{aligned} \psi^{\mathcal{C}}(x) &= \mathcal{C}\psi(x)\mathcal{C} = \gamma^2\left(\psi^{\dagger}(x)\right)^T = \left(\psi^{\dagger}(x)\gamma^2\right)^T = \left(\overline{\psi}(x)\gamma^0\gamma^2\right)^T, \\ \overline{\psi^{\mathcal{C}}}(x) &= \mathcal{C}\overline{\psi}(x)\mathcal{C} = \psi^{\mathcal{C}\dagger}(x)\gamma^0 = \left(-\gamma^2\psi(x)\right)^T\gamma^0 = \left(-\gamma^0\gamma^2\psi(x)\right)^T. \end{aligned} \tag{1.17}$$

The parity transformation $\mathcal{P}$ or spatial inversion is the operation that reflects the space coordinates $\mathbf{x}$ into $-\mathbf{x}$ and is equivalent to a mirror reflection followed by a rotation. Its action is given by $\psi^{\mathcal{P}}(x) = \mathcal{P}\psi(t,\mathbf{x})\mathcal{P}^{\dagger}$ with $\mathcal{P} = \mathcal{P}^{\dagger} = \mathcal{P}^{-1}$ and $\mathcal{P}^2 = \mathbb{I}$ and the transformation rules are

$$\begin{aligned} \psi^{\mathcal{P}}(x) &= \mathcal{P}\psi(t,\mathbf{x})\mathcal{P} = \gamma^0\psi(t,-\mathbf{x}), \\ \overline{\psi^{\mathcal{P}}}(x) &= \mathcal{P}\overline{\psi}(t,\mathbf{x})\mathcal{P} = \overline{\psi}(t,-\mathbf{x})\gamma^0. \end{aligned} \tag{1.18}$$

Since the spin direction is left unaltered a very important consequence of this operation is that the helicity of the particle, or the projection of the spin onto the direction of momentum, is reversed: this means that under the parity transformation a left-handed particle becomes right-handed.

Finally, we have the time reversal transformation $\mathcal{T}$ that reflects the time coordinate t into $-t$ leaving $\mathbf{x}$ unchanged: this means that while spatial relations must be the same, all momenta and angular momenta must be reversed. For this reason the $\mathcal{T}$ operation represents the reversal of motion. Its action is given by $\psi^{\mathcal{T}}(x) = \mathcal{T}\psi(t, \mathbf{x})\mathcal{T}^\dagger$ and defining the matrix Θ

$$\Theta = -\gamma^1\gamma^3 = -i \begin{pmatrix} \sigma_2 & 0 \\ 0 & \sigma_2 \end{pmatrix}, \tag{1.19}$$

that fulfills the relations $\Theta\Theta^* = -\mathbb{I}$ and $\Theta\Theta^\dagger = \mathbb{I}$, we find the transformation rules for the fields

$$\begin{aligned} \psi^{\mathcal{T}}(x) &= \mathcal{T}\psi(t, \mathbf{x})\mathcal{T}^\dagger = \Theta\psi(-t, \mathbf{x}), \\ \overline{\psi^{\mathcal{T}}}(x) &= \mathcal{T}\overline{\psi}(t, \mathbf{x})\mathcal{T}^\dagger = \overline{\psi}(-t, \mathbf{x})\Theta^\dagger \quad . \end{aligned} \tag{1.20}$$

The experimental evidence actually shows that the electromagnetic and strong interactions as well as classical gravity respect $\mathcal{C}$ and $\mathcal{P}$ symmetries and, therefore, their combination $\mathcal{CP}$. On the other hand weak interactions violate both $\mathcal{C}$ and $\mathcal{P}$ in the strongest possible way. For example the $W^\pm$ bosons couple only to left-handed particles and right-handed antiparticles, but neither to right-handed particles nor to left-handed antiparticles. Parity violation, that was prompted first by C. N. Yang and T. D. Lee in 1956 [12], was confirmed the next year in both nuclear [13] and pion beta decay [14], [15]. After this discovery, the combined $\mathcal{CP}$ symmetry was proposed as a symmetry of Nature. Indeed, if we consider the pion decay $\pi^+ \to \mu^+\nu_\mu$ involving a left-handed neutrino and its permitted $\mathcal{CP}$ counterpart $\pi^- \to \mu^-\overline{\nu}_\mu$ involving a right-handed antineutrino, we find that the combined $\mathcal{CP}$ transformation is an exact symmetry (see Fig. 1.3).

This aspect is crucial because if $\mathcal{CP}$ was an exact symmetry, the laws of Nature would be the same for matter and antimatter. However, $\mathcal{CP}$ is violated in certain rare processes and thus only the combined discrete $\mathcal{CPT}$ symmetry transformation is an exact symmetry of Nature. Historically the first experimental evidence of $\mathcal{CP}$ violation was pointed out in 1964 by J. W. Cronin and V. L. Ficht [16] who analyzed the K^0 meson decay.

1.3 The CKM matrix

In the SM, the $\mathcal{CP}$ symmetry is broken by the quarks' Yukawa couplings with the Higgs field which are described by the lagrangian of Eq. 1.13 limited to the quark sector. After the EWSB and the mass diagonalization of Eq. 1.14, the weak-eigenstates and mass-eigentstates of quarks become mixed and the charged current interactions are given by

$$\mathcal{L}_{W^\pm} = -\frac{g}{\sqrt{2}}\overline{u}_{Li}\gamma^\mu \left(V_{CKM}\right)_{ij} d_{Lj} W_\mu^+ + h.c.,$$

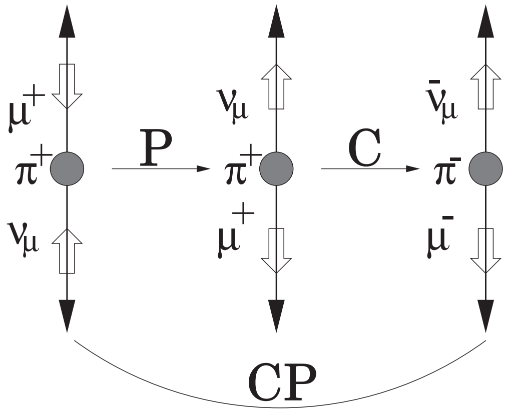


Figure 1.3: The physical π^+ decay is transformed via the parity operation to an unphysical decay, but the further charge conjugation operation transforms it to a physically allowed situation for π^- decay. The solid arrows denote momentum vectors, the open arrows the spin.

where V_{CKM} is the Cabibbo-Kobayashi-Maskawa matrix, previously defined in Eq. 1.16. The fundamental property of the CKM matrix is its unitarity: $V_{CKM}^\dagger V_{CKM} = V_{CKM} V_{CKM}^\dagger = \mathbb{I}$. This condition determines the number of free parameters of the matrix. A generic $N \times N$ unitary matrix contains N^2 independent parameters, $2N-1$ of which can be eliminated redefining the phase of the N up-type and N down-type quarks as

$$u_{Li} \to e^{i\varphi_i^u} u_{Li} \ , \qquad d_{Lj} \to e^{i\varphi_j^d} d_{Lj} \ , \tag{1.21}$$

in such a way to reduce free parameters to $(N-1)^2$. The remaining parameters can be further splitted into mixing angles and complex phases as follows

$$\underbrace{\frac{1}{2}N(N-1)}_{\text{mixing angles}} + \underbrace{\frac{1}{2}(N-1)(N-2)}_{\text{complex phases}} = (N-1)^2 \ . \tag{1.22}$$

As can be noticed, the case $N = 2$ leads to only one free parameter that is the Cabibbo mixing angle θ_C contained in the Cabibbo matrix [8]

$$V_C = \begin{pmatrix} \cos\theta_C & \sin\theta_C \\ -\sin\theta_C & \cos\theta_C \end{pmatrix} \ . \tag{1.23}$$

The $N = 3$ case leads to the generalization of the V_C matrix, the V_{CKM} matrix, which contains four physical parameters, namely three mixing angles and one complex phase. This phase causes the $\mathcal{CP}$ violation in the SM.

Between the many possible conventions, a standard choice to parametrize V_{CKM} is [7]

$$V_{CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} \quad (1.24)$$

where $s_{ij} = \sin\theta_{ij}$, $c_{ij} = \cos\theta_{ij}$ and δ is the $\mathcal{CP}$ violating phase. All the angles θ_{ij} can be chosen to lie in the first quadrant, so that $s_{ij} \geq 0$ and $c_{ij} \geq 0$. They must vanish if there is no mixing between two quark generations i, j: in particular, in the limit $\theta_{13} = \theta_{23} = 0$ the CKM matrix would reduce to V_C. The presence of a complex phase in the mixing matrix is a necessary but not sufficient condition for having $\mathcal{CP}$ violation. As pointed out in [17], another key condition is

$$(m_t^2 - m_c^2)(m_t^2 - m_u^2)(m_c^2 - m_u^2)(m_b^2 - m_s^2)(m_b^2 - m_d^2)(m_s^2 - m_d^2) \times J_{CP} \neq 0 \quad , \quad (1.25)$$

where

$$J_{CP} = |\mathfrak{Im}(V_{ij}V_{kl}V_{il}^*V_{kj}^*)| \qquad (i \neq k, j \neq l) \quad (1.26)$$

is the "Jarlskog invariant". The mass terms in Eq. 1.25 reflect the fact that the CKM phase could be eliminated through a transformation of the quark fields like Eq. 1.21 if any two up-type or down-type quarks were degenerate in mass. Consequently, the origin of $\mathcal{CP}$ violation is deeply connected to other important questions of particle physics like the "flavour problem", the hierarchy of quark masses, and the number of fermion generations. The Jarlskog invariant J_{CP} is a phase-convention-independent measure of the size of $\mathcal{CP}$ violation, that according to the standard parametrization of Eq. 1.24 can be written as

$$J_{CP} = s_{12}s_{13}s_{23}c_{12}c_{23}c_{13}^2 \sin\delta \ . \quad (1.27)$$

According to the tiny $\mathcal{CP}$ violation effects expected within the SM, it corresponds to the very small value [18]

$$J_{CP} = (3.099^{+0.052}_{-0.063}) \times 10^{-5} \ . \quad (1.28)$$

1.3.1 Magnitude of the matrix elements

Now we give an overview of the main measurements leading to the experimental values of the CKM matrix elements. The size of $|V_{ij}|$ elements can be directly determined from the following tree-level processes [7]:

$|V_{ud}|$: Nuclear β decays or neutron decay $n \to pe^-\overline{\nu}_e$;

$|V_{us}|$: Semileptonic kaon decays $K \to \pi l \nu_l$ and hadronic tau decays $\tau^- \to K^-\nu_\tau$. It is also accessible through $|V_{us}/V_{ud}|$ measuring the ratios $K^+ \to \mu^+\nu_\mu(\gamma)/\pi^+ \to \mu^+\nu_\mu(\gamma)$ (where (γ) indicates that radiative decays are included) and $\tau^- \to K^-\nu_\tau/\tau^- \to \pi^-\nu_\tau$;

$|V_{cd}|$: Semileptonic D meson decays $D \to \pi l \nu_l$ and leptonic decay $D^+ \to \mu^+\nu_\mu$. It can also be accessed through measurement of charm-production fractions in neutrino interactions;

$|V_{cs}|$: Semileptonic D decays and semileptonic D_s decays;

$|V_{cb}|$: Exclusive and inclusive semileptonic decays of B mesons to charm;

$|V_{ub}|$: Semileptonic $B \to X_u l\nu$ decays;

$|V_{tb}|$: Branching fractions $\mathcal{B}(t \to Wb)/\mathcal{B}(t \to Wq)$ with $q = b, s, d$ from top decays and single top-quark-production cross section.

The $|V_{td}|$ and $|V_{ts}|$ elements cannot be measured using tree-level processes and need to be determined from the measurement of the oscillation frequencies of the $B^0 - \overline{B}^0$ and $B_s^0 - \overline{B}_s^0$ systems. They can also be accessed through $|V_{td}/V_{ts}|$ as extracted from decays occurring via $b \to q\gamma$ radiative transitions, where $q = b, d$. In Fig. 1.4 are shown the Feynman diagrams for some important processes that allow to extract the experimental values of the first five CKM matrix elements of the list above.

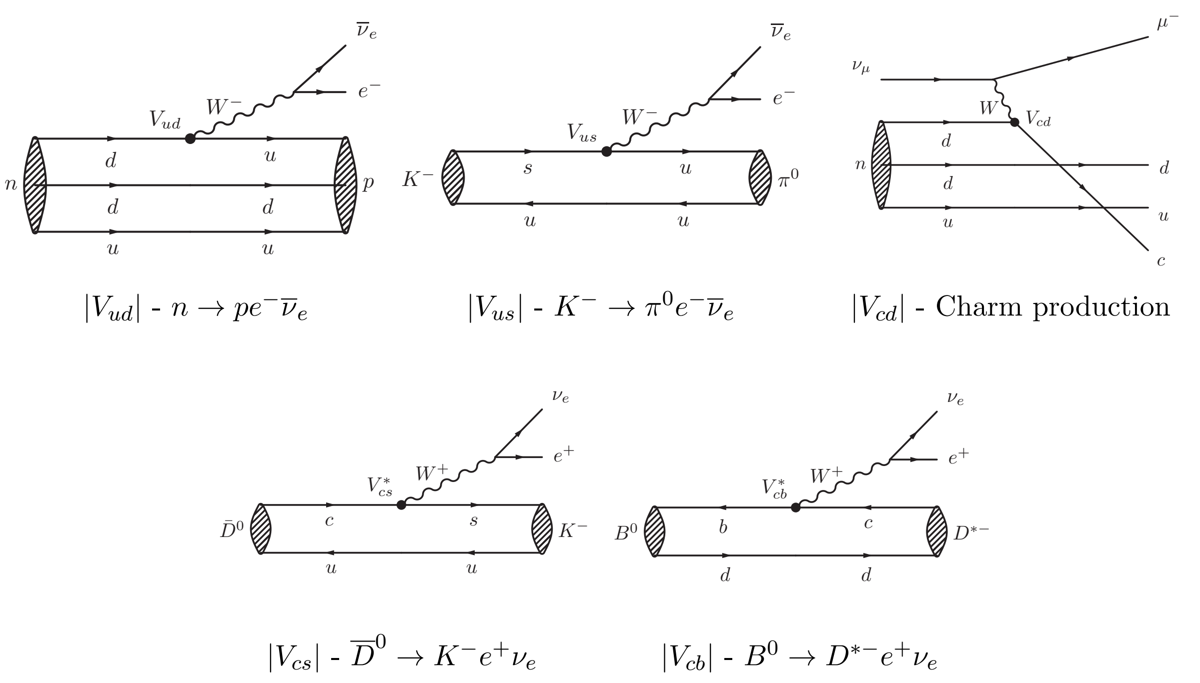


Figure 1.4: Feynman diagrams of some specific processes important for determining five of the $|V_{ij}|$ matrix elements.

Since the probability of a particular quark transition is proportional to the square modulus $|V_{ij}|^2$ of the relative matrix element, the knowledge of these elements is very important to determine the amplitudes of weak decays. The CKM matrix elements can be most precisely determined by a global fit that combines all the available measurements and imposes the SM constraints. Tab. 1.2 summarizes the current knowledge of the magnitudes of the CKM matrix elements using the latest CKMfitter results [18]. The uncertainties

CKM matrix element	Fitted value
$\|V_{ud}\|$	$0.974334^{+0.000064}_{-0.000068}$
$\|V_{us}\|$	$0.22508^{+0.00030}_{-0.00028}$
$\|V_{cd}\|$	$0.22494^{+0.00029}_{-0.00028}$
$\|V_{cs}\|$	$0.973471^{+0.000067}_{-0.000067}$
$\|V_{cb}\|$	$0.04181^{+0.00028}_{-0.00060}$
$\|V_{ub}\|$	$0.003715^{+0.000060}_{-0.000060}$
$\|V_{td}\|$	$0.008575^{+0.000076}_{-0.000098}$
$\|V_{ts}\|$	$0.04108^{+0.00030}_{-0.00057}$
$\|V_{tb}\|$	$0.999119^{+0.000024}_{-0.000012}$

Table 1.2: Magnitudes of the CKM matrix elements [18].

reported in this table correspond to a 68% confidence level (CL). Transitions within the same quark generation, corresponding to the diagonal elements of the CKM matrix, are strongly favoured while transitions between different generations, corresponding to the off-diagonal elements, are disfavoured (see Fig. 1.5). Transitions between the first and second generations are suppressed by a factor $\mathcal{O}(10^{-1})$, those between the second and third generations are suppressed by a factor $\mathcal{O}(10^{-2})$ and those between the first and third generations are strongly suppressed by a factor $\mathcal{O}(10^{-3})$.

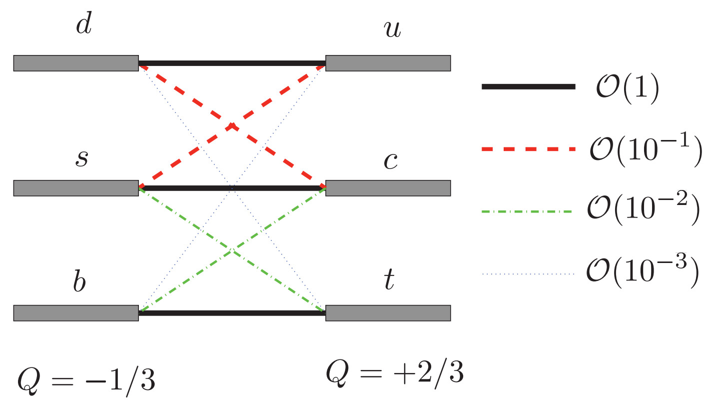


Figure 1.5: Graphical representation of the transition hierarchy between the different quark generations.

1.3.2 Wolfenstein parametrization and unitarity triangles

Given the experimental knowledge of the $|V_{ij}|$ matrix elements it can be stated that exists a clear hierarchy in the mixing of quarks, namely

$$1 \gg s_{12} \gg s_{23} \gg s_{13} \ . \tag{1.29}$$

From this evidence, by introducing the following identities known as "Wolfenstein parametrization" [19],

$$s_{12} = \lambda = \frac{|V_{us}|}{\sqrt{|V_{ud}|^2 + |V_{us}|^2}} \ , \tag{1.30}$$

$$s_{23} = A\lambda^2 = \lambda \left| \frac{V_{cb}}{V_{us}} \right| \ , \tag{1.31}$$

$$s_{13}e^{i\delta} = V_{ub} = A\lambda^3(\rho - i\eta) \ , \tag{1.32}$$

we can rewrite the CKM matrix of Eq. 1.24 in power series of the parameter λ obtaining a different parametrization of V_{CKM}. This way, expanding up to the fifth order, we have [20]

$$V_{CKM} = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 - \frac{1}{8}\lambda^4 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda + \frac{1}{2}A^2\lambda^5[1 - 2(\rho + i\eta)] & 1 - \frac{1}{2}\lambda^2 - \frac{1}{8}\lambda^4(1 + 4A^2) & A\lambda^2 \\ A\lambda^3[1 - (1 - \frac{1}{2}\lambda^2)(\rho + i\eta)] & -A\lambda^2 + \frac{1}{2}A\lambda^4[1 - 2(\rho + i\eta)] & 1 - \frac{1}{2}A^2\lambda^4 \end{pmatrix} + \mathcal{O}(\lambda^6). \tag{1.33}$$

In this parametrization the Jarlskog invariant defined in Eq. 1.27 becomes

$$J_{CP} = \lambda^6 A^2 \eta, \tag{1.34}$$

where η is the $\mathcal{CP}$ violating parameter.

A very important aspect of the CKM matrix is that the unitarity property, $(V^\dagger_{CKM}V_{CKM})_{ij} = (V_{CKM}V^\dagger_{CKM})_{ij} = \delta_{ij}$, corresponds to a set of nine complex equations that connects the V_{ij} elements: three of these involve the diagonal terms and are equal to 1, while the other six equations, those ones involving the off-diagonal terms, vanish. The equations for the off-diagonal elements ($i \neq j$) can be represented as triangles of area

$J_{CP}/2$ and are respectively:

$$\underbrace{V_{ud}V_{us}^*}_{\mathcal{O}(\lambda)} + \underbrace{V_{cd}V_{cs}^*}_{\mathcal{O}(\lambda)} + \underbrace{V_{td}V_{ts}^*}_{\mathcal{O}(\lambda^5)} = 0, \tag{1.35}$$

$$\underbrace{V_{ud}V_{ub}^*}_{\mathcal{O}(\lambda^4)} + \underbrace{V_{cs}V_{cb}^*}_{\mathcal{O}(\lambda^2)} + \underbrace{V_{ts}V_{tb}^*}_{\mathcal{O}(\lambda^2)} = 0, \tag{1.36}$$

$$\underbrace{V_{ud}V_{ub}^*}_{\mathcal{O}(\lambda^3)} + \underbrace{V_{cd}V_{cb}^*}_{\mathcal{O}(\lambda^3)} + \underbrace{V_{td}V_{tb}^*}_{\mathcal{O}(\lambda^3)} = 0, \tag{1.37}$$

$$\underbrace{V_{ud}^*V_{cd}}_{\mathcal{O}(\lambda)} + \underbrace{V_{us}^*V_{cs}}_{\mathcal{O}(\lambda)} + \underbrace{V_{ub}^*V_{cb}}_{\mathcal{O}(\lambda^5)} = 0, \tag{1.38}$$

$$\underbrace{V_{cd}^*V_{td}}_{\mathcal{O}(\lambda^4)} + \underbrace{V_{cs}^*V_{ts}}_{\mathcal{O}(\lambda^2)} + \underbrace{V_{cb}^*V_{tb}}_{\mathcal{O}(\lambda^2)} = 0, \tag{1.39}$$

$$\underbrace{V_{ud}^*V_{td}}_{\mathcal{O}(\lambda^3)} + \underbrace{V_{us}^*V_{ts}}_{\mathcal{O}(\lambda^3)} + \underbrace{V_{ub}^*V_{tb}}_{\mathcal{O}(\lambda^3)} = 0, \tag{1.40}$$

where each product $|V_{ij}V_{kl}^*|$ represent the lenght of the corresponding triangle side. These triangles are the so-called "unitarity triangles". Two triangles, namely those of Eq. 1.37 and 1.40, have all their sides length of the order of $\mathcal{O}(\lambda^3)$. The other four equations contain terms with different powers of λ and hence give rise to "squashed" triangles.

The triangle of Eq. 1.37 is of particular importance and is commonly known as the Unitarity Triangle (UT). It can be represented in the complex plane $(\overline{\rho}, \overline{\eta})$, where the $\overline{\rho}$ and $\overline{\eta}$ parameters contain corrections in λ and are defined as

$$\overline{\rho} = \rho\left(1 - \frac{\lambda^2}{2}\right) \quad \overline{\eta} = \eta\left(1 - \frac{\lambda^2}{2}\right) . \tag{1.41}$$

The UT angles are given by

$$\alpha \equiv \phi_2 = \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right), \quad \beta \equiv \phi_1 = \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right), \quad \gamma \equiv \phi_3 = \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right), \tag{1.42}$$

where two conventions exist in the literature, and their sides are defined as

$$R_u = \left|\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right| = \sqrt{\overline{\rho}^2 + \overline{\eta}^2}, \quad R_t = \left|\frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*}\right| = \sqrt{(1-\overline{\rho})^2 + \overline{\eta}^2}. \tag{1.43}$$

The apex of the triangle is then given by

$$\overline{\rho} + i\overline{\eta} \equiv -\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} \equiv 1 + \frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*} = \frac{\sqrt{1-\lambda^2}(\rho + i\eta)}{\sqrt{1-A^2\lambda^4} + \sqrt{1-\lambda^2}A^2\lambda^4(\rho+i\eta)} . \tag{1.44}$$

The graphical representation of the unitarity triangle is shown in Fig. 1.6. In Tab. 1.3 are summarized the values of the A, λ, $\overline{\rho}$ and $\overline{\eta}$ parameters, where the uncertainties correspond

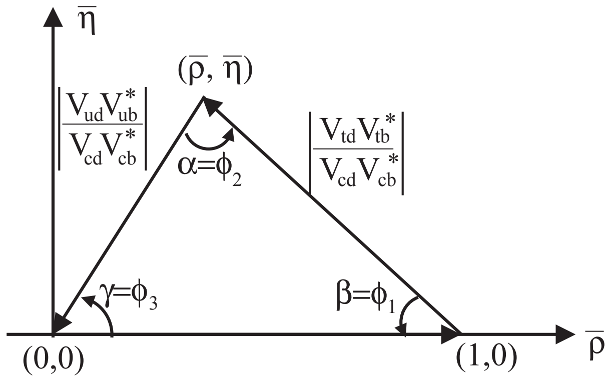


Figure 1.6: Graphical representation in the complex plane $(\overline{\rho}, \overline{\eta})$ of the unitarty triangle of Eq. 1.37 as normalized to $V_{cd}V_{cb}^*$: $1 + V_{ud}V_{ub}^*/V_{cd}V_{cb}^* + V_{td}V_{tb}^*/V_{cd}V_{cb}^* = 0$.

to a 68% CL. In Fig. 1.7 are superimposed all the latest CKM constraints determined under the SM hypothesis in the $(\overline{\rho}, \overline{\eta})$ plane.

As a last remark, let us notice that the two matrix elements $|V_{ts}|$ and $|V_{cb}|$ have the same order of magnitude. Looking at Tab. 1.2 we can see that their values are indeed extremely close. As a consequence, the R_t side of Eq. 1.43 can be rewritten as

$$R_t \approx \left|\frac{V_{td}V_{tb}^*}{V_{cd}V_{ts}^*}\right| \quad . \tag{1.45}$$

This means that, by measuring the ratio $|V_{td}/V_{ts}|$, we can constrain the right side of the unitarity triangle.

Parameter	Fitted value
A	$0.8250^{+0.0071}_{-0.0111}$
λ	$0.22509^{+0.00029}_{-0.00028}$
$\overline{\rho}$	$0.1598^{+0.0076}_{-0.0072}$
$\overline{\eta}$	$0.3499^{+0.0063}_{-0.0061}$

Table 1.3: Current values of the Wolfenstein parameters.

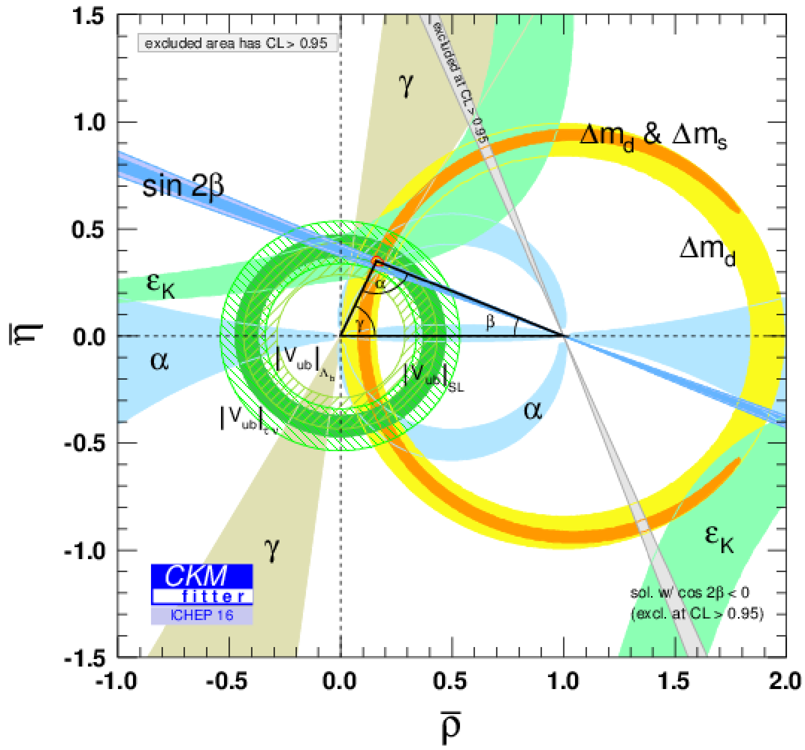


Figure 1.7: Constraints on the unitarity triangle at a 95% confidence level.

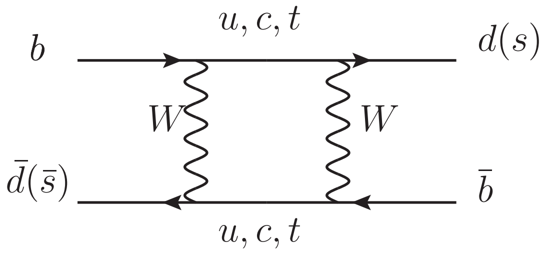
1.4 Neutral B meson oscillations

The neutral B^0 (B_s^0) meson is a pseudoscalar particle with bottomness $B = +1$ composed by the $b\bar{d}$ ($b\bar{s}$) quark-antiquark pair, that can oscillate into its antiparticle $\overline{B}^0$ ($\overline{B}_s^0$) composed by the $\bar{b}d$ ($\bar{b}s$) pair and with $B = -1$. These $B^0 \longleftrightarrow \overline{B}^0$ ($B_s^0 \longleftrightarrow \overline{B}_s^0$) oscillations, also referred to as mixing, are due to second order virtual transitions $\Delta B = 2$ associated to the box diagrams shown in Fig. 1.8. In this section we discuss only the mixing of the B^0 and B_s^0 mesons, but the very same formalism can be applied to K^0 and D^0 mesons.

Because of the oscillations, at any time t the B_q^0 generic meson (where $q = d, s$ and $B_d^0 = B^0$) can be seen as a superposition of states described by the following wave function

$$|B_q^0(t)\rangle = a(t)|B_q^0\rangle + b(t)|\overline{B}_q^0\rangle + \sum_f c_f(t)|f\rangle \ , \tag{1.46}$$

where $|B_q^0\rangle$ and $|\overline{B}_q^0\rangle$ represent the particle and antiparticle states of the B_q^0 meson, $|f\rangle$ are the possible finale states in which the meson is allowed to decay and $c_f(t)$ are the coefficients of each final state. In the case of a time range much larger than the typical



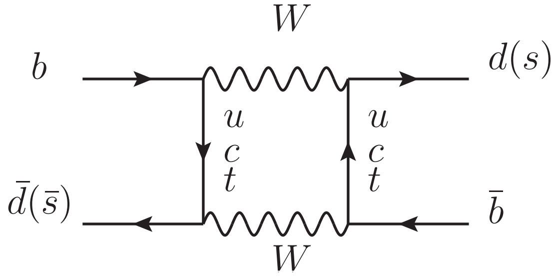


Figure 1.8: Box diagrams describing the $B^0 - \overline{B}^0$ and $B_s^0 - \overline{B}_s^0$ mixing.

strong interaction scale, we can exploit the so called Weisskopf-Wigner approximation that allows to use the simplified formalism [21], [22]

$$|B^0(t)\rangle = a(t)|B^0\rangle + b(t)|\overline{B}^0\rangle \ , \tag{1.47}$$

with $|a(t)|^2 + |b(t)|^2 = 1$. The subscript q is here suppressed to simplify the discussion. This wave function obeys to the Schrödinger equation $i(\partial/\partial t)|B^0(t)\rangle = \mathcal{H}_{eff}|B^0(t)\rangle$ where the effective Hamiltonian is given by

$$\mathcal{H}_{eff} = \mathbf{M} - \frac{i}{2}\mathbf{\Gamma} = \begin{pmatrix} M & M_{12} \\ M_{12}^* & M \end{pmatrix} - \frac{i}{2}\begin{pmatrix} \Gamma & \Gamma_{12} \\ \Gamma_{12}^* & \Gamma \end{pmatrix}. \tag{1.48}$$

The diagonal elements of $\mathcal{H}_{eff}$ are associated to flavour-conserving transitions $B^0 \to B^0$ and $\overline{B}^0 \to \overline{B}^0$, while the off-diagonal ones are associated to flavour-changing transitions $B^0 \longleftrightarrow \overline{B}^0$. The diagonal elements of $\mathcal{H}_{eff}$ are equal because of the assumption of $\mathcal{CPT}$ invariance [23]. This Hamiltonian is not Hermitian otherwise mesons would not oscillate and not decay, but the $\mathbf{M}$ and $\mathbf{\Gamma}$ matrices are. The $\mathbf{M}$ matrix represents transitions via dispersive intermediate state ("off-shell" or short-range transitions), and $\mathbf{\Gamma}$ represents transitions via absorptive intermediate states ("on-shell" transition). The solution of the eigenvalue equation for $\mathcal{H}_{eff}$ gives two eigenstates

$$|B_H^0\rangle = p|B^0\rangle - q|\overline{B}^0\rangle, \qquad |B_L^0\rangle = p|B^0\rangle + q|\overline{B}^0\rangle, \tag{1.49}$$

which corresponds to the eigenvalues

$$\begin{aligned} \lambda_H &= M - \frac{i}{2}\Gamma + \frac{q}{p}\left(M_{12} - \frac{i}{2}\Gamma_{12}\right) = M_H - \frac{i}{2}\Gamma_H, \\ \lambda_L &= M - \frac{i}{2}\Gamma - \frac{q}{p}\left(M_{12} - \frac{i}{2}\Gamma_{12}\right) = M_L - \frac{i}{2}\Gamma_L, \end{aligned} \tag{1.50}$$

where p and q are complex coefficients satisfying $|p|^2 + |q|^2 = 1$ and whose ratio is given by

$$\frac{q}{p} = \sqrt{\frac{M_{12}^* - (i/2)\Gamma_{12}^*}{M_{12} - (i/2)\Gamma_{12}}} \ . \tag{1.51}$$

The real parts of the eigenvalues $\lambda_{H,L}$ represent the masses, $M_{H,L}$, and their imaginary parts represent the widths, $\Gamma_{H,L}$, of the two eigenstates $|B^0_{H,L}\rangle$: by convention the subscripts H and L label respectively the heavy and the light eigentstates. The mass and width differences between the two eigenstates are

$$\Delta m = M_H - M_L, \qquad \Delta\Gamma = \Gamma_L - \Gamma_H, \tag{1.52}$$

where Δm is positive by definition and the sign of $\Delta\Gamma$, which is unknown a priori, has to be experimentally determined and is expected to be positive within the Standard Model. Now, defining the functions

$$g_+(t) = \left(\frac{e^{-i\lambda_H t} + e^{-i\lambda_L t}}{2}\right), \qquad g_-(t) = \left(\frac{e^{-i\lambda_H t} - e^{-i\lambda_L t}}{2}\right), \tag{1.53}$$

and inserting Eq. 1.49 we can write the expressions for the time evolution of pure B^0 and $\overline{B}^0$ particle states as follows

$$|B^0(t)\rangle = g_+(t)|B^0\rangle + \frac{q}{p}g_-(t)|\overline{B}^0\rangle, \qquad |\overline{B}^0(t)\rangle = g_+(t)|\overline{B}^0\rangle + \frac{p}{q}g_-(t)|B^0\rangle, \tag{1.54}$$

with the square modulus of the $g_\pm(t)$ functions given by

$$\begin{aligned} |g_\pm(t)|^2 &= \frac{1}{4}\left(e^{-\Gamma_H t} + e^{-\Gamma_L t} \pm 2e^{-\Gamma t}\cos\Delta m t\right) \\ &= \frac{e^{-\Gamma t}}{2}\left(\cosh\frac{\Delta\Gamma t}{2} \pm \cos\Delta m t\right) \ . \end{aligned} \tag{1.55}$$

This quantity represents the time-dependent probability to conserve the initial flavour (+) or oscillate into the opposite one (−). We now recall the subscript q and write the Γ, $\Delta\Gamma$ and Δm expression for the generic B^0_q meson

$$\Gamma_q = \frac{\Gamma_{qH} + \Gamma_{qL}}{2}, \qquad \Delta\Gamma_q = \frac{\Gamma_{qL} - \Gamma_{qH}}{2}, \qquad \Delta m_q = M_{qH} - M_{qL} \ , \tag{1.56}$$

where Γ_q fulfills the natural role of decay constant, $\Gamma_q = 1/\tau_q$.

The $B^0 - \overline{B}^0$ and $B^0_s - \overline{B}^0_s$ systems have very different oscillation frequencies and these are determined by the Δm_q terms, which contain the short-distance contributions associated to the Feynman diagrams of Fig. 1.8. The main contribution to the box is that one coming from the top quark. The SM transition amplitude for the B^0_q mixing is given by [24]

$$\langle B^0_q|\mathcal{H}^{\Delta B=2}_{eff}|\overline{B}^0_q\rangle = \langle B^0_q| \ \frac{G_F^2 M_W^2}{16\pi^2}(\mathcal{F}^0_d\mathcal{Q}^d_1 + \mathcal{F}^0_s\mathcal{Q}^s_1) + h.c. \ |\overline{B}^0_q\rangle, \tag{1.57}$$

where $\mathcal{H}_{eff}^{\Delta B=2}$ is the simplest effective Hamiltonian and the four-fermion operators $\mathcal{Q}_1^q$ are defined as

$$\mathcal{Q}_1^q = [\bar{b}\gamma_\mu(1-\gamma_5)q][\bar{b}\gamma^\mu(1-\gamma_5)q]. \tag{1.58}$$

G_F is the Fermi coupling constant, M_W is the mass of the $W^\pm$ boson and $\gamma_5 = i\gamma^0\gamma^1\gamma^2\gamma^3$. The short distance function $\mathcal{F}_q^0$ in Eq. 1.57 is

$$\mathcal{F}_q^0 = \lambda_{tq}^2 S_0(x_t) \tag{1.59}$$

with

$$\lambda_{tq} = V_{tq}^* V_{tb} \tag{1.60}$$

and where $S_0(x_t)$ is an Inami-Lim function with $x_t = m_t^2/M_W^2$. It describes the basic electroweak box contributions without QCD. This leads to the SM prediction of the Δm_q mass difference

$$\Delta m_q = \frac{G_F^2 M_W^2}{6\pi^2} m_{B_q} \eta_{2B} S_0(x_t) f_{B_q}^2 \hat{B}_{B_q} |V_{tq}V_{tb}^*|^2 \ , \tag{1.61}$$

where m_{B_q} is the mass of the oscillating meson, η_{2B} is a perturbative QCD correction factor calculated to the next to leading order (NLO), f_{B_q} is the decay constant of the meson and $\hat{B}_{B_q}$ is the "bag-factor". The product $f_{B_q}\sqrt{\hat{B}_{B_q}}$ is related to non-perturbative corrections, then is the most important source of uncertainty in the Δm_q prediction. On the other hand the ratio

$$\xi = \frac{f_{B_d}\sqrt{\hat{B}_{B_d}}}{f_{B_s}\sqrt{\hat{B}_{B_s}}} \tag{1.62}$$

is calculated with a better precision in Lattice QCD (LQCD) than the individual $f_{B_q}\sqrt{\hat{B}_{B_q}}$ because statistical and systematic uncertainties cancel in part. With this ratio, $|V_{td}/V_{ts}|$ can be determined from the ratio of oscillation frequencies $\Delta m_d/\Delta m_s$ and used to constrain the side R_t. The present constraint as reported on the PDG is $|V_{td}/V_{ts}| = 0.215 \pm 0.001 \pm 0.011$ [7].

A graphical representation of the oscillation probabilities of the $B_q^0 - \overline{B}_q^0$ mesons is shown in Fig. 1.9. It can be noticed that the oscillation frequency of the $B_s^0 - \overline{B}_s^0$ system is much higher than the $B^0 - \overline{B}^0$ one. This depends on the magnitude of the CKM matrix elements that contribute to the transition probability: in fact while $\Delta m_d \propto |V_{td}V_{tb}^*|^2 m_t^2 \sim \lambda^6 m_t^2$, $\Delta m_s \propto |V_{ts}V_{tb}^*|^2 m_t^2 \sim \lambda^4 m_t^2$.

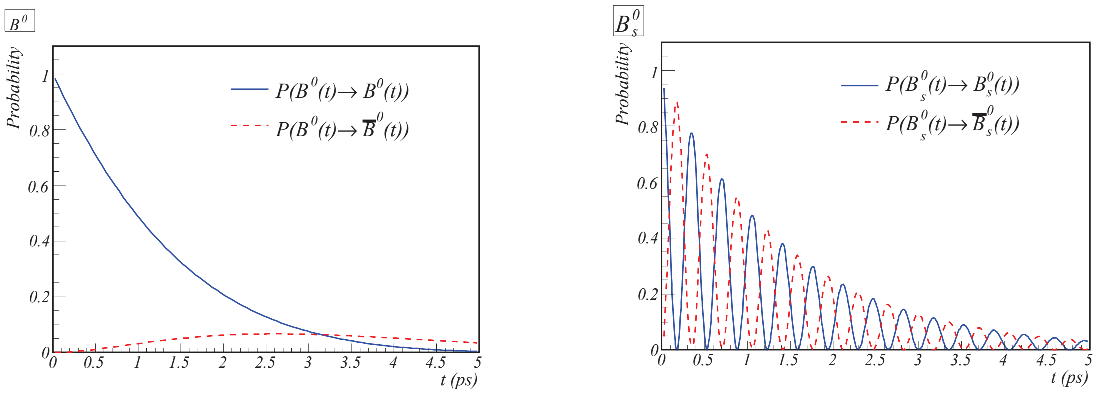


Figure 1.9: Time evolution of the $B^0 - \overline{B}^0$ (left) and $B^0_s - \overline{B}^0_s$ (right) systems.

1.5 Radiative B decays

Decays of B mesons with a photon in the final state are non tree-level processes that correspond at quark level to $b \to q\gamma$ transitions occurring via Flavour Changing Neutral Currents (see Fig. 1.10). These transitions appear in the SM through electroweak penguin diagrams in which the main contribution is represented by the propagation of a top quark in the loop. In this section we will consider only the $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ radiative decays, which respectively correspond to a $b \to d\gamma$ and $b \to s\gamma$ transition. Due to the presence of the V_{td} matrix element in one of the vertices of the penguin diagram on Fig. 1.10, the $B^0 \to \rho^0\gamma$ decay results to be highly suppressed. The $B^0 \to K^{*0}\gamma$ mode, even being a rare decay, instead is strongly favoured with respect to $B^0 \to \rho^0\gamma$ since it involves the V_{ts} matrix element: $|V_{ts}| \sim 5\ |V_{td}|$.

In order to compute the observables of interest, let us start defining the effective SM

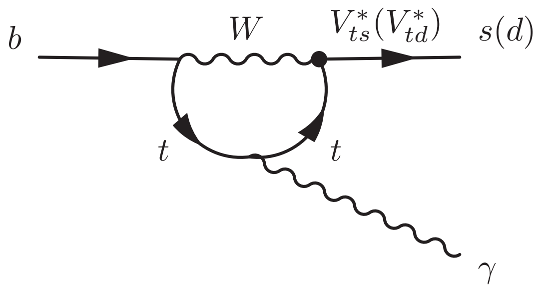


Figure 1.10: Feynman penguin diagram for $b \to q\gamma$ transitions in the SM ($q = d, s$).

Hamiltonian for the $b \to d\gamma$ transition, which is given by [25]

$$\mathcal{H}_{eff}^{b\to d} = \frac{G_F}{\sqrt{2}}\Big\{ V_{ub}V_{ud}^* \left[C_1^u(\mu)\mathcal{Q}_1^u(\mu) + C_2^u(\mu)\mathcal{Q}_2^u(\mu)\right] \tag{1.63}$$

$$+ V_{cb}V_{cd}^* \left[C_1^c(\mu)\mathcal{Q}_1^c(\mu) + C_2^c(\mu)\mathcal{Q}_2^c(\mu)\right] \tag{1.64}$$

$$- V_{tb}V_{td}^* \left[C_7^{eff}(\mu)\mathcal{Q}_7(\mu) + C_8^{eff}(\mu)\mathcal{Q}_8(\mu)\right] + \ \ldots \Big\}, \tag{1.65}$$

where μ is the scale $\mu = O(m_b)$ with m_b the mass of the b quark. The operators $\mathcal{Q}_1^q$ and $\mathcal{Q}_2^q$ $(q = u, c)$ are the four-fermion operators

$$\mathcal{Q}_1^q = [\bar{d}_\alpha\gamma_\mu(1-\gamma_5)q_\beta][\bar{q}_\beta\gamma^\mu(1-\gamma_5)b_\alpha], \qquad \mathcal{Q}_2^q = [\bar{d}_\alpha\gamma_\mu(1-\gamma_5)q_\alpha][\bar{q}_\beta\gamma^\mu(1-\gamma_5)b_\beta], \tag{1.66}$$

and $\mathcal{Q}_7$ and $\mathcal{Q}_8$ are the electromagnetic and chromomagnetic penguin operators respectively defined as

$$\mathcal{Q}_7 = \frac{em_b}{8\pi^2}[\bar{d}_a\sigma^{\mu\nu}(1+\gamma_5)b_a]F_{\mu\nu}, \qquad \mathcal{Q}_8 = \frac{g_s m_b}{8\pi^2}[\bar{d}_\alpha\sigma^{\mu\nu}(1+\gamma_5)T^a_{\alpha\beta}b_\beta]G^a_{\mu\nu}\ . \tag{1.67}$$

Here, e and g_s are the electric and colour charges, $F_{\mu\nu}$ and $G^a_{\mu\nu}$ are the electromagnetic and gluonic field tensors of Eq. 1.6 and 1.7, $T^a_{\alpha\beta}$ are the $SU(3)_C$ generators of Eq. 1.10 in a slightly different notation and $\sigma^{\mu\nu} = \frac{i}{4}[\gamma^\mu, \gamma^\nu]$. The quark indices α, β and the colour index a are written explicitly. The coefficients $C_1^q(\mu)$ and $C_2^q(\mu)$ in Eq. 1.65 are the Wilson coefficients corresponding to the operators $\mathcal{Q}_1^q$ and $\mathcal{Q}_2^q$, while the coefficients $C_7^{eff}(\mu)$ and $C_8^{eff}(\mu)$ include also the effects of the QCD penguin four-fermion operators which are assumed to be present in the effective Hamiltonian. For the $b \to s\gamma$ transition the effective hamiltonian $\mathcal{H}_{eff}^{b\to s}$ can be obtained by replacing the quark field d_α by s_α and by replacing the CKM factors $V_{qb}V_{qd}^*$ by $V_{qb}V_{qs}^*$ $(q = u, c, t)$ in $\mathcal{H}_{eff}^{b\to d}$.

This leads to the following ratio of branching fractions [26]

$$\frac{\mathcal{B}(B^0 \to \rho^0\gamma)}{\mathcal{B}(B^0 \to K^{*0}\gamma)} = \frac{1}{2}\left|\frac{V_{td}}{V_{ts}}\right|^2 \left(\frac{1 - m_\rho^2/m_B^2}{1 - m_{K^*}^2/m_B^2}\right)^3 \zeta_\rho[1 + \Delta R_\rho] \tag{1.68}$$

that allows to constrain the $|V_{td}/V_{ts}|$ ratio. ζ_ρ is the ratio of the $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ form factors and ΔR_ρ accounts for differences in the decay dynamics including weak annihilation contributions.

Given the high suppression of the $B^0 \to \rho^0\gamma$ decay, this measurement is not competitive with the output of $B^0 - \overline{B}^0$ and $B_s^0 - \overline{B}_s^0$ oscillation measurements in terms of precision. However $b \to d\gamma$ transitions are quite sensitive to many extensions of the SM [27]. Therefore, if these two different measurements gave results in disagreement this would be a hint of the existence of New Physics. One in fact can imagine new heavy particles other than the top quark, or the W boson, propagating in the loop of Fig. 1.10. In this sense, the measurement of the ratio of branching fractions in Eq. 1.68 represents an interesting test of the standard model consistency. The discussion of the experimental status is postponed to Chap. 4.

Chapter 2

The LHCb detector at the LHC

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) at CERN [28], in Geneva, is the biggest and most powerful particle accelerator ever built. The LHC is a two-ring collider with a circumference of 27 km placed inside the tunnel which originally contained the LEP (Large Electron Positron Collider), at an average depth of 100 m (see Fig. 2.1). It is designed to collide protons up to a center-of-mass energy of 14 TeV, with an instantaneous luminosity of 10^{34} cm^{-2}s^{-1}, and heavy ions (Pb-Pb) at a center-of-mass energy of 2.8 TeV per nucleon, with a peak luminosity of 10^{27} cm^{-2}s^{-1}.

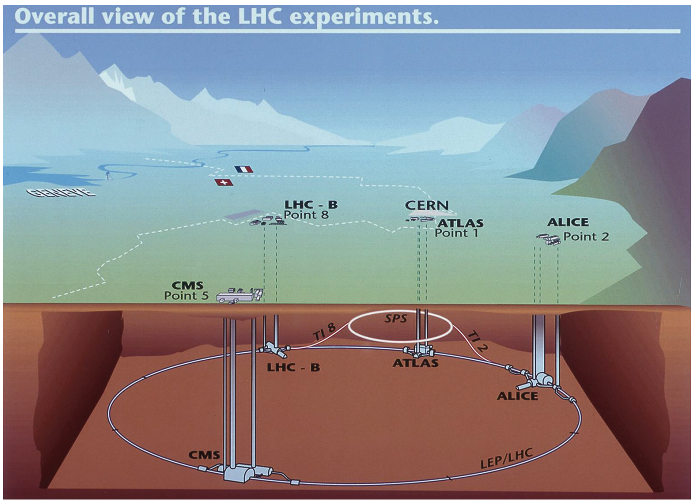


Figure 2.1: Schematic view of the LHC collider. The figure also shows the four main experiments (ALICE, ATLAS, CMS and LHCb).

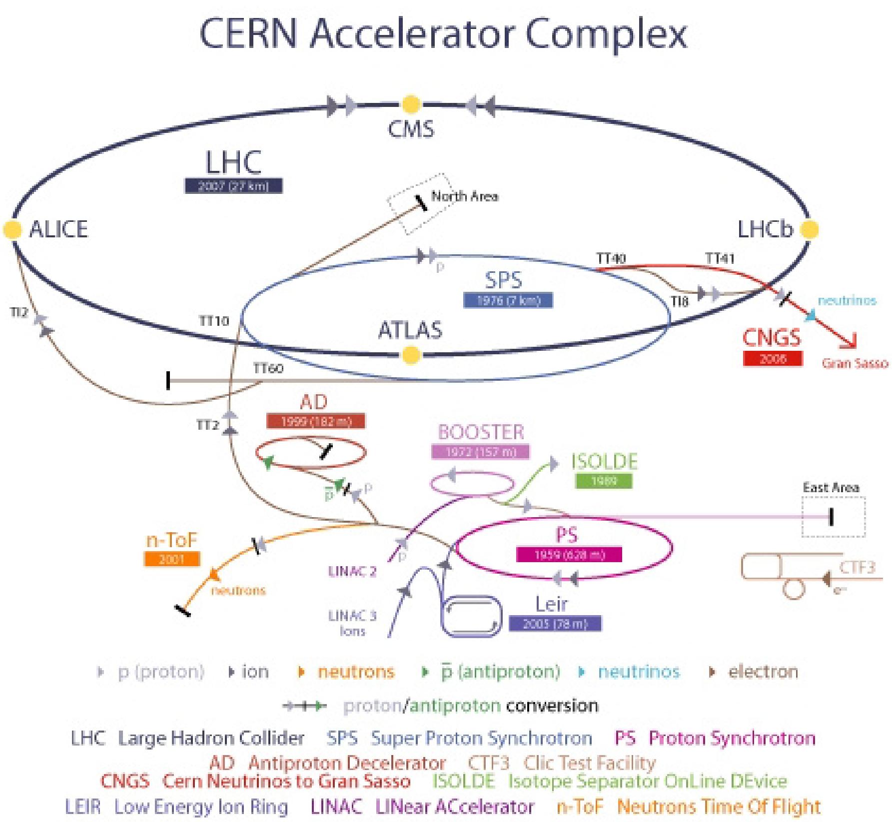


Figure 2.2: Sketch representing the various pre-accelerating machines. The four main detectors (yellow points) are asymetrically positioned along the beam line.

The protons used in the collisions, which are obtained from ionized hydrogen atoms, require to be pre-accelerated before being injected into the main rings. In Fig. 2.2 is schematically shown the CERN accelerator complex. Firstly, protons are accelerated passing through the LINAC2, from which they come out with an energy of 50 MeV. Then they pass through the PSB (Proton Synchrotron Booster) and the PS (Proton Synchtotron), reaching an energy of 1.4 GeV and 26 GeV respectively. Finally, before the injection in the LHC, the SPS (Super Proton Synchrotron) increases the energy of protons up to 450 GeV. Once in the collider, the protons are kept in their orbits by means of super-conducting magnets providing a magnetic field of 8.34 T. At the nominal operation regime, the LHC rings store 2808 proton bunches per ring, each of them containing 1.1×10^{11} protons, colliding with a frequency of 40 MHz. The two general purpose detectors ATLAS and CMS are dedicated to the direct search of new particles and to study the properties of the Higgs boson. ALICE is a specialized detector dedicated to the study of the quark-gluon plasma, while LHCb is dedicated to the study of heavy flavour physics and $\mathcal{CP}$ violation.

2.2 The LHCb detector

The LHCb experiment is designed to exploit the great production cross sections of $b\bar{b}$ pairs in pp collisions at the LHC energies. This has been measured at $\sqrt{s} = 7$ TeV to be [29]

$$\sigma(pp \to b\bar{b}X)_{4\pi} = (284 \pm 20 \pm 49)\ \mu\text{b}$$

where X stands for the other collision products. Thanks to this and to the excellent detector performances, LHCb can perform high precision studies of flavour physics.

The integrated luminosities corresponding to the various years of data taking are shown in Fig. 2.3. A full data set of about 3 fb^{-1} was collected between 2010 and 2012, the so called Run 1 period. The center-of-mass energy was $\sqrt{s} = 7$ TeV in 2010 and 2011 and $\sqrt{s} = 8$ TeV in 2012. The analysis presented in this thesis exploits the 2011 and 2012 data samples. After a shutdown period from the end of 2012 to summer 2015, the Long Shutdown 1 (LS1), the LHC restarted and since then delivers pp collisions at a center-of-mass energy of $\sqrt{s} = 13$ TeV: this phase is referred to as Run 2. The instantaneous luminosity at the LHCb experiment is $4 \times 10^{32}\text{cm}^{-2}\text{s}^{-1}$, one order of magnitude lower than the one achieved for the general purpose experiments. This is due to the fact that LHCb could not efficiently operate in the same pile-up conditions as ATLAS and CMS.

The LHCb detector [30] is a single-arm spectrometer with a forward geometry that covers a region of angular acceptance approximately between 10 mrad and 300 mrad in the horizontal plane (xz) and between 10 mrad and 250 mrad in the vertical plane (yz). The detector has been designed with such a geometry because at high energies $b\bar{b}$ pairs are mostly produced with a strong boost along the beam line and, as a consequence, B hadrons are predominantly produced with a small angle in the same forward or backward cone. The difference between horizontal and vertical angular acceptances is justified by the

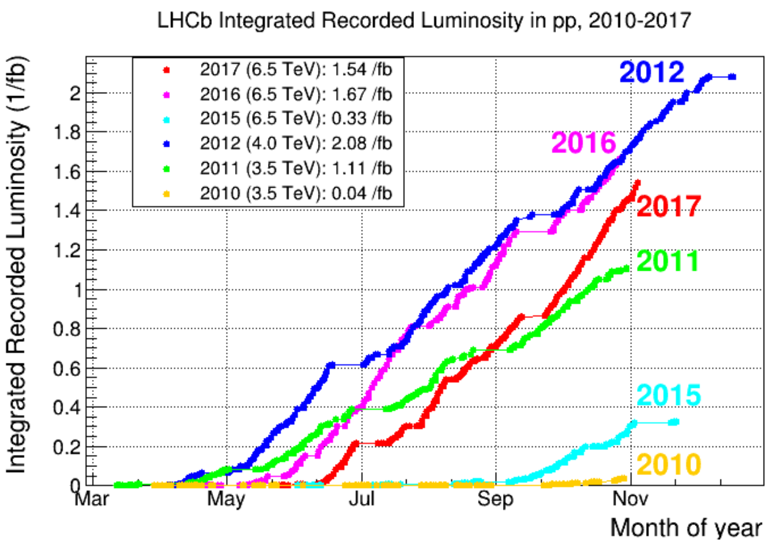


Figure 2.3: Integrated luminosities recorded by the LHCb experiment during Run 1 (2010-2012) and Run 2 (since 2015) data taking periods.

fact that the horizontal plane is the bending plane in which charged particles are deflected by the dipole magnetic field of LHCb. Such a geometrical acceptance corresponds to a pseudorapidity between about 2 and 5. The pseudorapidity is defined as

$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right] \tag{2.1}$$

where θ is the angle between the beam direction and the particle direction.

The LHCb physics program needs the detector to satisfy the following requirements:

- A great precision in the reconstruction of pp interaction vertices and B hadrons decay vertices. This is crucial to perform studies like the measurement of neutral B meson oscillations, for which it is necessary to have an adequate proper-time resolution.

- An excellent particle identification (PID) system in order to discriminate between charged pions, charged kaons and protons with momentum in a range between few GeV/c up to 100 GeV/c. Furthermore the analysis of final states containing leptons requires an optimal PID of muons and electrons.

- The invariant mass resolution must be as small as possible in order to discriminate the signals from the combinatorial background and in order to distinguish between B^0 and B_s^0 decays. For these reasons, the momentum of charged tracks must be measured with a relative precision of $\sim 10^{-3}$.

- Since the production cross sections of $b\bar{b}$ pairs is considerable, the trigger system must have a very high background rejection efficiency in order to reduce the acquired data samples to a manageable size. In order to reach this purpose, the LHCb trigger is organized in multiple levels, each of them processing the output of the previous level.

- The large amount of data collected by the experiment requires efficient and reliable computing resources, both needed for the data processing and for their storage.

An overview of the entire LHCb detector is shown in Fig. 2.4, where from left to right the following elements are visible:

VELO : the VErtex LOcator is placed around the interaction region and provides the reconstruction of primary and secondary vertices;

RICH1 : the first Ring Imaging CHerenkov detector is placed just after the VELO and provides information for the identification of charged particles;

TT : the Trigger Tracker is placed after the first RICH and is the former tracking system;

Magnet : the dipole magnet of LHCb provides the magnetic field used to bend particles in order to evaluate their charge and momentum;

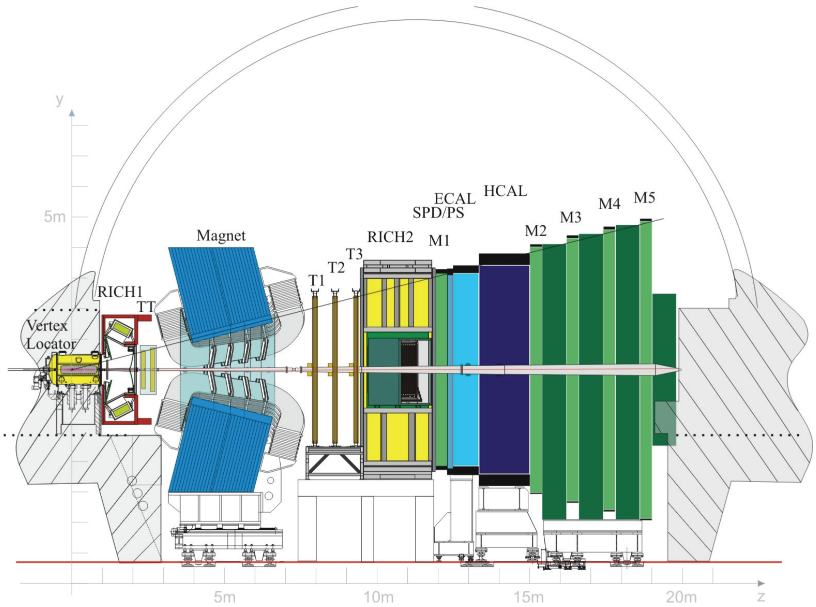


Figure 2.4: Overview of the entire LHCb detector.

Tracking Stations : the three tracking stations (T1 to T3) are placed behind the magnet and accomplish the second stage of tracks reconstruction;

RICH2 : the second Ring Imaging CHerenkov detector is designed to provide efficient particle identification in a different momentum range with respect to RICH1;

Electromagnetic Calorimeter : the Electromagnetic CALorimeter (ECAL) system is placed after the RICH2 and is necessary for an efficient trigger and for the identification of electrons and photons. It is preceded by the Scintillating Pad Detector (SPD) and the PreShower (PS) auxiliary sub-detectors;

Hadronic Calorimeter : the Hadronic CALorimeter (HCAL) is placed behind the ECAL and is exploited for the hadronic trigger;

Muon Stations : the Muon Stations are placed just before the SPD/PS (M1) and at the end of the detector (M2 to M5), where only muons can arrive without being stopped by the calorimeter system. They are used both for an efficient trigger on decays with muons in the final state and for muon identification.

The VELO, the Trigger Tracker and the three tracking stations together with the magnetic dipole form the LHCb tracking system, while the two RICH detectors, the Electromagnetic and Hadronic calorimeters and the muon stations form the LHCb particle identification system.

2.3 The LHCb tracking system

The purpose of the tracking system is to identify the interaction and decay vertices, reconstruct the particles trajectories and measure their momenta. The first task is accomplished by the VELO, which also provides the track reconstruction together with the Trigger Tracker and the three tracking stations.

2.3.1 The Vertex Locator

The VErtex LOcator provides precise measurements of the tracks coordinates close to the interaction region. These coordinates are used to identify the displaced secondary vertices which are the distinctive feature of B hadron decays. In fact, since B hadrons have on average lifetimes of the order of $\sim 10^{-12}$ s, they cover a mean distance of about 1 cm inside the detector giving rise to secondary vertices well spaced from the primary pp interaction vertex. For this reason, in order to select signals and reject most of the combinatorial background, it is necessary to have a vertex detector with a micrometric precision.

The VELO [31] consists of a series of 21 circular silicon modules, each providing a measure of the $R = \sqrt{x^2 + y^2}$ and ϕ coordinates, arranged perpendicularly along the beam line direction as shown in Fig. 2.5. Each module is divided in two halves which can be moved far from or close to the beam, depending on need. In fact the VELO aperture varies from an open position, required during the beam stabilization phase, to a closed position which is maintained during data-taking. For this reason, VELO modules are mounted on a moveable device inside a vessel that maintains the vacuum, and each half can be moved between 3 cm and 8 mm away from the beam. The module halves are composed of two planes of 300 μm thick silicon microstrip sensors that provide a measure of radial (R sensors) and polar (ϕ sensors) coordinates of the hits generated by ionizing particles that cross the VELO. The sketch of R and ϕ sensors is shown in Fig. 2.6. The z coordinate is measured knowing which modules provided a signal for a given particle hit.

The R sensors are divided in four parts per halves of about 45° each. The microstrips are modeled in a semi-circular shape and their width varies from 38 μm (close to the beam) to 102 μm (far from the beam): the smaller width close to the interaction region is related to the higher number of particles expected in that zone.

The ϕ sensors are divided in two regions, inner and outer. The pitch size of the inner region increases linearly as a function of the radius, ranging from 38 μm to 78 μm. The outer region, that starts at a radius of 17.25 mm, has instead a pitch ranging from 39 μm to 97 μm. Inner and outer regions have different skew to the radial direction in order to improve pattern recognition: they are tilted of 20° and 10° respectively. Furthermore, in order to achieve a better track reconstruction, the longitudinally adjacent ϕ sensors have opposite skew to each other.

In Fig. 2.7 a 3-dimensional overview of the entire VELO apparatus is shown. Detection stations and dedicated read-out electronics are mounted on a moveable device and placed inside the vacuum vessel (10^{-4} mbar).

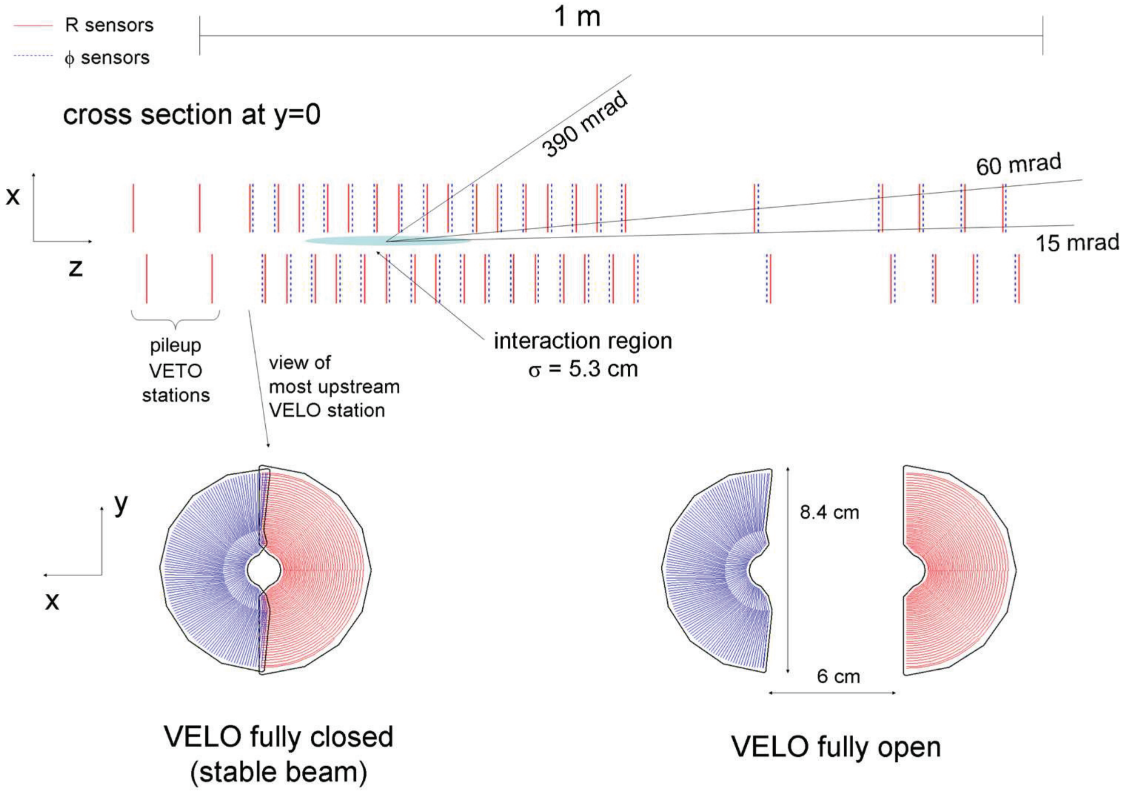


Figure 2.5: Top view of the VELO silicon modules, with the detector in the fully closed position (top). Front view of the modules in both the closed (bottom left) and open positions (bottom right). As can be noticed, in the fully closed configuration the two halves of each silicon module partially overlap.

The performances of the VELO detector have been extensively studied using the data collected in 2011 [32]. In particular a primary vertex resolution of 13 μm in the transverse plane (x, y) and 71 μm along the beam axis is achieved for vertices with 25 tracks or more. An impact parameter resolution of less than 35 μm is achieved for particles with transverse momentum greater than 1 GeV/c.

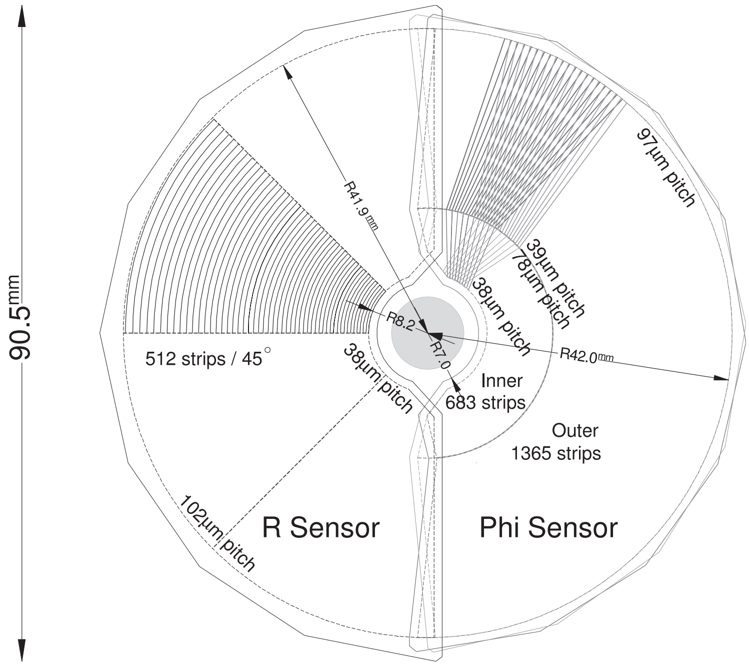


Figure 2.6: Scheme of the geometry of R (left) and ϕ (right) silicon sensors of the VELO detector. In order to show their different orientation, the microstrips of the ϕ sensors for two adjacent modules are drawn.

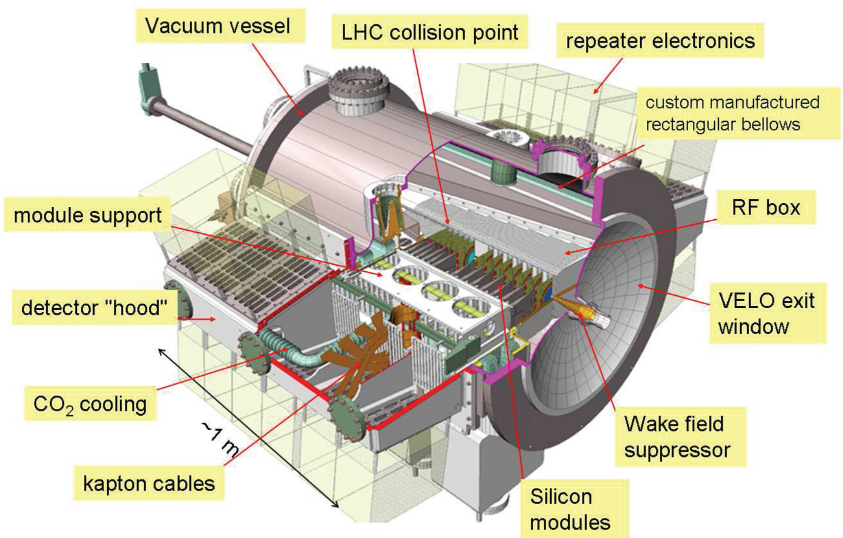


Figure 2.7: 3-dimensional overview of the entire VELO apparatus.

2.3.2 The Trigger Tracker

The Trigger Tracker [33] is placed between the first RICH detector and the dipole magnet in a region where a residual magnetic field is present (0.15 Tm). The TT purpose is to provide reference segments which are used to combine the tracks reconstructed in the tracking stations after the magnet and those reconstructed in the VELO: this allows to improve the resolution on momentum and trajectory of the reconstructed tracks. The system consists of four stations, divided in two groups called respectively TTa and TTb, spaced by approximately 30 cm and placed at a distance of about 2.4 m from the interaction region. Each of the four stations cover a rectangular region of about 130 cm in height and about 150 cm wide. The scheme of the TT sub-detector is shown in Fig. 2.8. Each TT station is made up of silicon microstrip sensors with a pitch of about 200 μm and is arranged into up to 38 cm long readout strips. In the first and fourth station the strips are parallel to the vertical plane, while in the second and third station they are tilted by $+5°$ (u-layer) and $-5°$ (v-layer) respectively. This arrangement allows to obtain a better precision in the reconstruction. The TT has a typical single-hit resolution of about 50 μm.

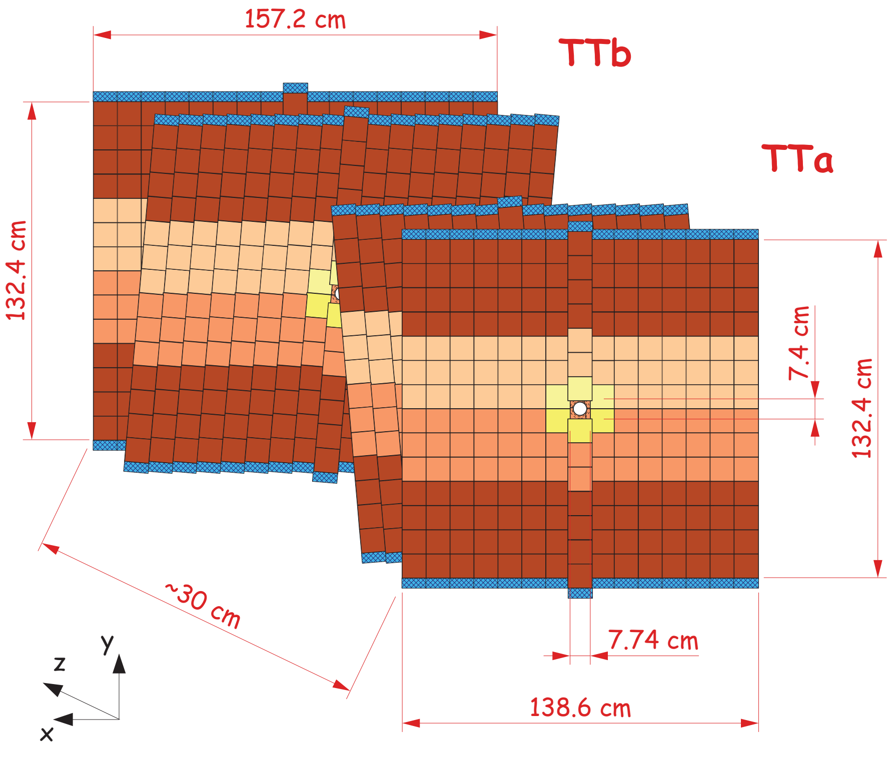


Figure 2.8: Scheme of the four TT stations.

2.3.3 The tracking stations T1-T2-T3

The three tracking stations T1, T2 and T3 are placed just after the dipole magnet and they are divided in two main parts. The inner part of the tracking stations, the Inner Tracker (IT), uses silicon microstrip sensors while the outer part, the Outer Tracker (OT), exploits drift straw tubes. As shown in Fig. 2.9 the IT part of each station is placed in front of the OT part.

The Inner Tracker [34] covers the region around the beam pipe and consists of four detection planes arranged as shown in Fig. 2.9 and 2.10. Similarly to the TT, in the first and fourth planes the silicon sensors are parallel to the vertical plane (x-planes), while in the second and third plane sensors are tilted respectively by $+5°$ (u-plane) and $-5°$ (v-plane). The features of microstrip sensors are analogous to those used for the Trigger Tracker since they have a pitch of about 200 μm and they are up to 22 cm long. The total size of the IT sub-detector is about 1.2 m in the bending plane and about 40 cm in the

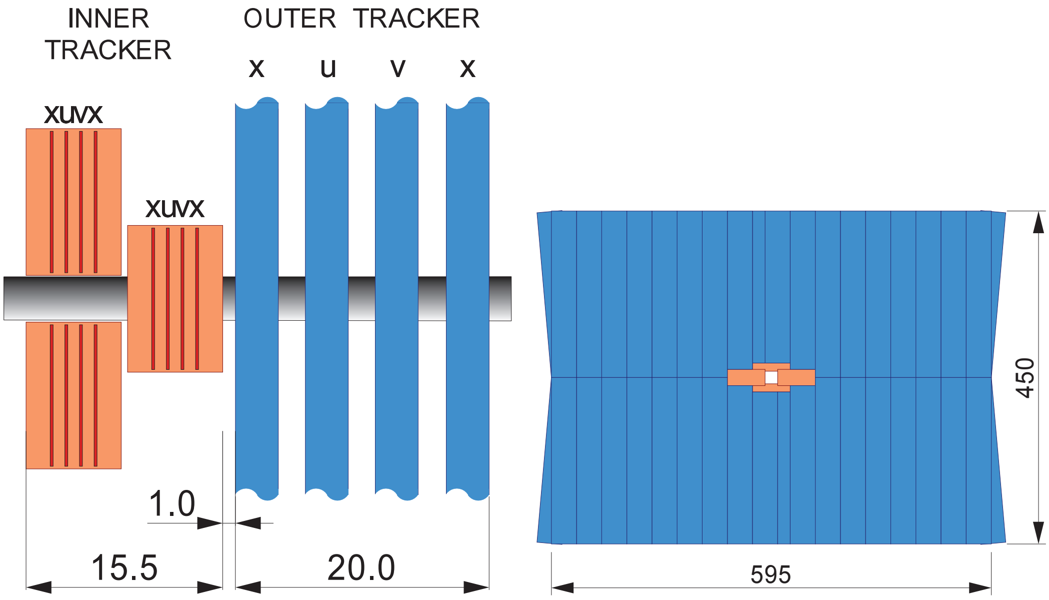


Figure 2.9: Layout of a T station from a side view (left) and from a front view (right). The dimensions are in centimeters. In the left part it can be seen that the IT sub-detector is placed in front of the OT sub-detector and the x- u- v-planes mentioned in the text are shown. In the right part it can be seen that the IT sub-detector (in orange) is placed around the beam pipe, while the OT sub-detector covers the outer region of the station.

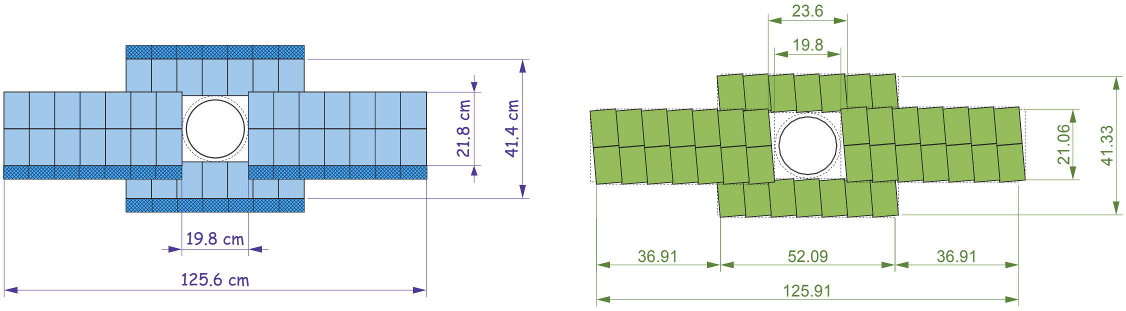


Figure 2.10: Frontal view of the x-plane (left) and u-plane (right) of the IT sub-detector.

vertical plane. The hit resolution in the IT is about the same as in the TT.

The Outer Tracker [35] is realized using gas-filled straw tubes detectors and consists of four planes of tubes arranged in the same way as the TT and IT sensors. In fact the first and fourth planes have tubes parallel to the vertical plane (x-planes), while the second and third planes have tubes tilted by $+5°$ (u-plane) and $-5°$ (v-plane). Furthermore each plane has two rows of tubes arranged in a honeycomb structure (see Fig. 2.11) in order to optimize the sensitive area. The straw tubes have a radius of 5 mm and are filled with a mixture of $Ar/CO_2/O_2$ (70/28.5/1.5 %) which guarantees a fast drift-time, below 50 ns. The OT resolution on the coordinate of the hits is typically around 200 μm [36].

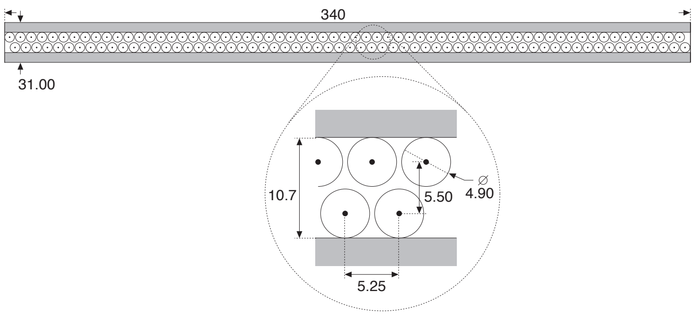


Figure 2.11: Cross section of a straw tubes plane in the OT. The zoomed part shows the honeycomb structure of the two rows of tubes.

2.3.4 The LHCb magnet

The magnetic field of LHCb is provided by a warm dipole magnet (*i.e.* non superconducting) placed between the TT and the first tracking station T1 [37]. As already mentioned the magnetic field is needed to identify the particles charge and to measure their momentum. The magnet (see Fig. 2.12) is formed by two coils inclined by a small angle with respect to the beam direction, in order to become wider with the increase of the z coordinate. The magnetic field is directed along the y coordinate perpendicular to the xz bending plane. In Fig. 2.13 the y component of the magnetic field is reported as a function of the z coordinate. The maximum intensity of the magnetic field is about 1 T, while the magnetic field integral is approximately 4 Tm. During the data-taking, the polarity of the magnetic field has been flipped several times in order to allow the evaluation of any left-right asymmetry induced by the detector. In fact, since the positive and negative charged particles are bent to different directions by the magnetic field, any variation in the detection efficiency between the left and the right part of the detector could affect $\mathcal{CP}$ asymmetry measurements.

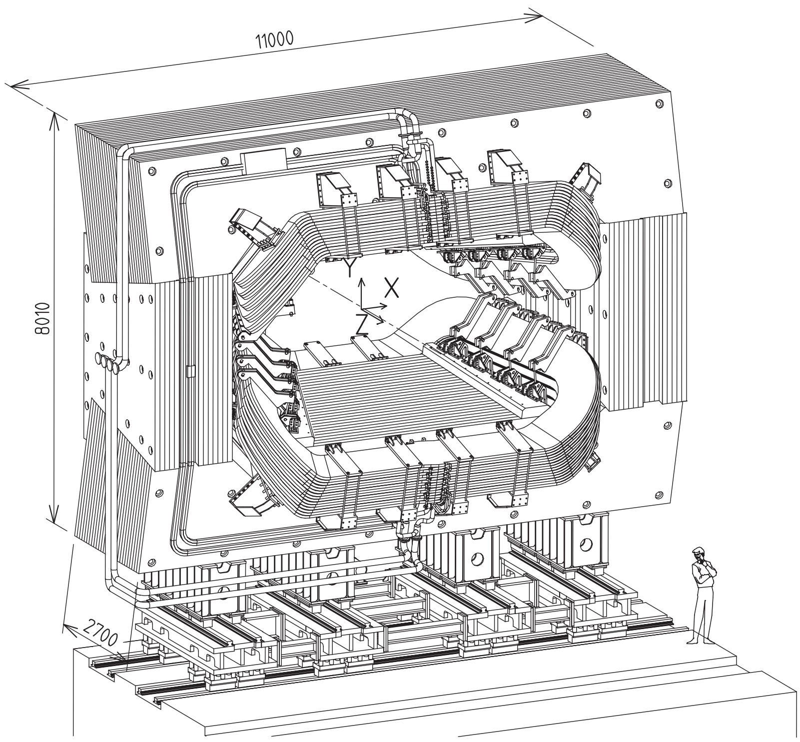


Figure 2.12: Sketch of the dipole magnet of LHCb.

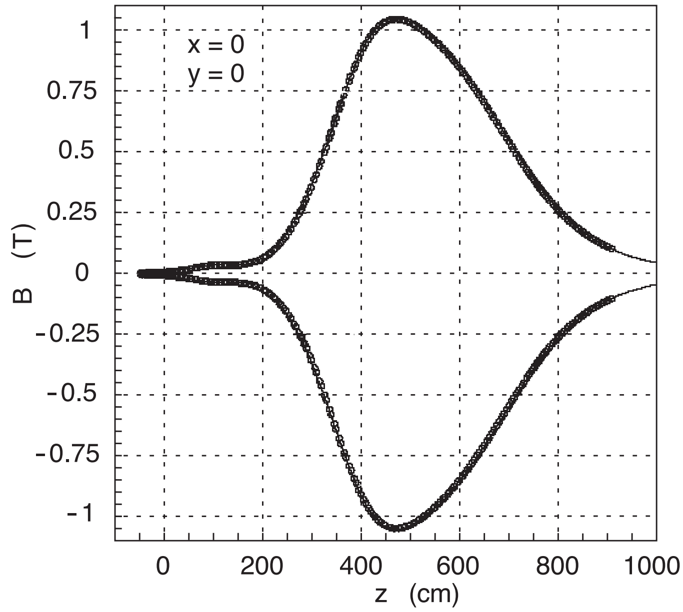


Figure 2.13: y component of the magnetic field as a function of the z coordinate.

2.4 The LHCb particle identification system

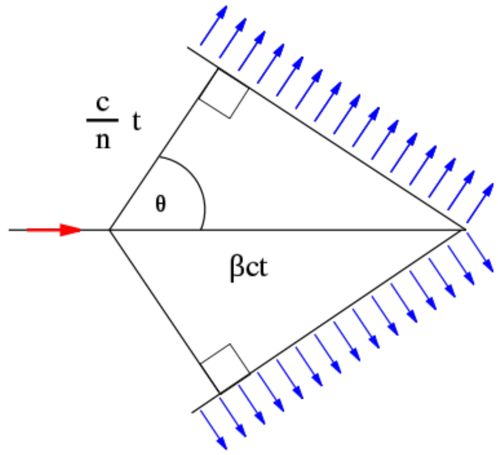
The purpose of the particle identification system is to provide an efficient and reliable identification of charged leptons and hadrons as well as photons and neutral pions. This task is accomplished by the two Ring Imaging Cherenkov detectors (RICH1 and RICH2), the electromagnetic (ECAL) and hadronic (HCAL) calorimeters and by the muon stations.

2.4.1 The RICH detectors

The Cherenkow effect is the emission of electromagnetic radiation that occurs when a charged particle passes through a dielectric medium at a speed greater than the phase velocity of light in that medium. This effect is exploited to discriminate charged pions, kaons and protons in a momentum range between few GeV/c up to about 100 GeV/c. A schematic representation of this effect is shown on the left side of Fig. 2.14. Cherenkov light detectors exploit the following relation between the particle momentum and the emission angle of Cherenkov photons

$$\cos(\theta_C) = \frac{1}{\beta n}, \tag{2.2}$$

where θ_C is the emission angle of Cherenkov photons with respect to the particle direction of flight, $\beta = v/c$ is the particle velocity normalized with respect to the speed of light in the vacuum and n is the refraction index of the radiator medium. Measuring this angle together with the momentum, it is possible to determine the mass of the particle.



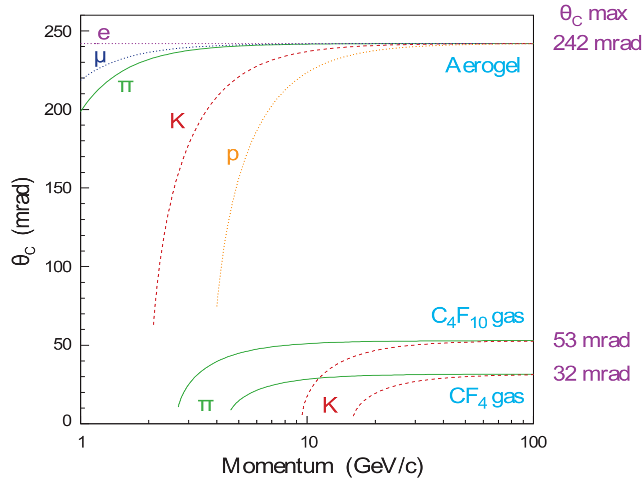


Figure 2.14: Left: geometric representation of the Cherenkov emission. Right: Cherenkov angle θ_C as a function of the particle momentum. The curves correspond to the different radiators used in RICH1 and RICH2 and to the various particle types.

The RICH detectors exploit different radiators because for particles approaching the speed of light the Cherenkov angle saturate at the value $\theta_C = \arccos(1/n)$. This behaviour is shown on the right side of Fig. 2.14.

The RICH1 [33], which is placed immediately after the VELO and has a geometrical acceptance from 25 mrad to 330 mrad, is optimized to identify low momentum tracks, in a range between 1 GeV/c and about 60 GeV/c. The RICH1 exploits two different types of radiators: the first is a 5 cm thick Aerogel layer with $n = 1.03$ optimal for low momentum particles (up to 10 GeV/c), while the second, gaseous C_4F_{10} with $n = 1.0015$ that fills a gap of about 85 cm, is dedicated to particles with higher momenta (up to 60 GeV/c).

The RICH2 is instead optimized for the identification of particles with higher momenta, from 15 GeV/c up to 100 Gev/c, and hence is complementary to the RICH1. It is placed after the last tracking station and has a geometrical acceptance of about 100 mrad in the vertical plane and about 120 mrad in the horizontal plane. The radiator used for the RICH2 is gaseous CF_4 that has a refraction index $n = 1.00046$.

The schematic picture of the two RICH detectors of LHCb is reported in Fig. 2.15. Both RICH detectors exploit an optical system made of spherical and plane mirrors in order to convey the emitted Cherenkov light on a lattice of photo-detectors (Hybrid Photon Detector, HPD). The HPD planes are placed out of the detector acceptance and are carefully shielded from the residual magnetic field.

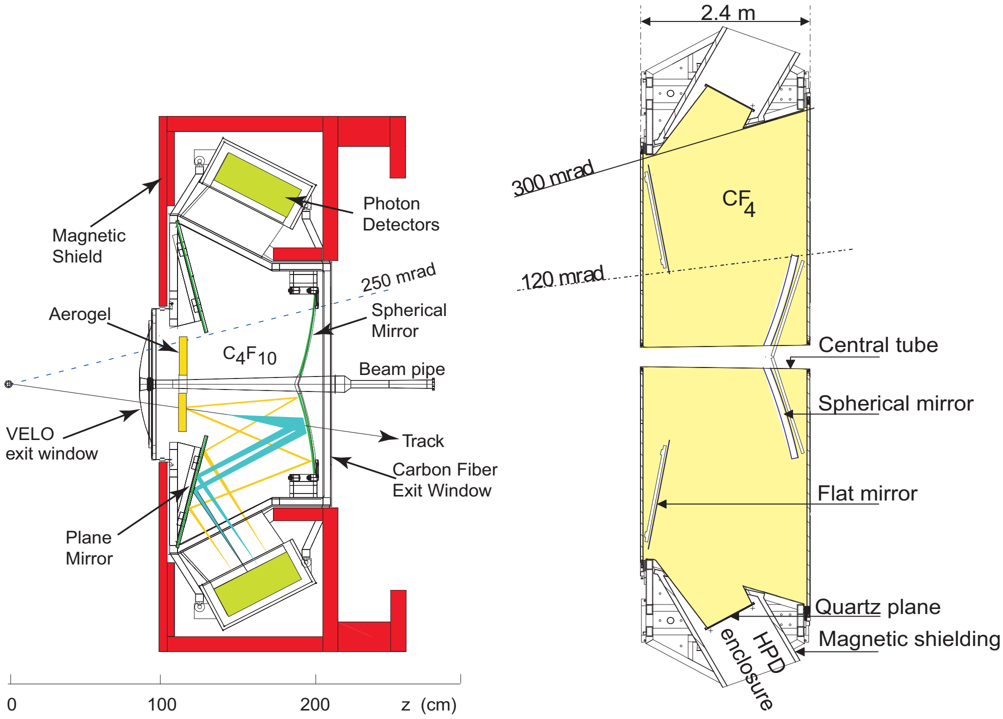


Figure 2.15: Schematic view of the RICH1 (left) and RICH2 (right) detectors. The different radiators and the optical systems are also shown.

The LHCb RICH detectors have excellent particle identification performances and provide a very clear discrimination of charged pions, kaons and protons. Fig. 2.16 shows the Cherenkov angle as a function of particle momentum using information from the C_4F_{10} radiator for isolated tracks selected in data (a track is here defined as isolated when its Cherenkov ring does not overlap with any other ring from the same radiator) [38]. As can be seen from the figure, events are distributed into distinct bands according to their mass. Although the RICH detectors are primarily used for hadron identification, it is worth noting that a distinct muon band can also be observed.

2.4.2 The calorimeter system

The calorimeter system [39] is used to identify electrons, photons and neutral pions and to measure their energy. Furthermore, it provides crucial information for the Level-0 trigger (L0-trigger), evaluating the transverse energy E_T of hadrons, electrons and photons. The calorimeter system is divided into four sub-detectors:

- Scintillator Pad Detector (SPD);
- PreShower (PS);

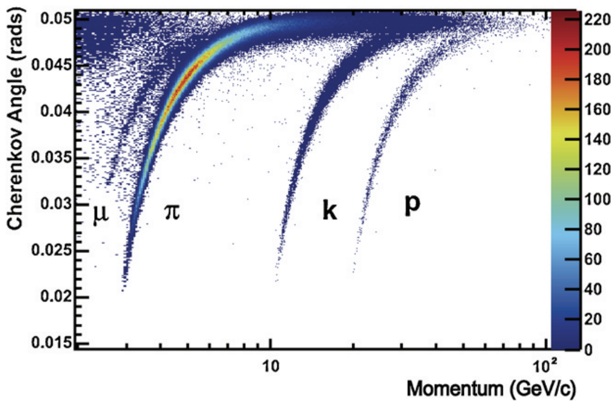


Figure 2.16: Reconstructed Cherenkov angle as a function of track momentum in the C_4F_{10} radiator [38].

- Electromagnetic CALorimeter (ECAL);
- Hadronic CALorimeter (HCAL).

In Fig. 2.17 is schematically represented the interaction of electrons, hadrons and photons with the various sub-detectors. Each sub-detector is divided into regions where different cell sizes are used. Indeed, aiming to reach a compromise between occupancy and a reasonable number of read-out channels, the size of the sensor elements increases going far from the beam-pipe and the high occupancy region. SPD, PS and ECAL are divided into three regions (inner, middle and outer) while the HCAL is composed of only two regions (inner and outer). A schematic overview of these subdivisions is shown in Fig. 2.18.

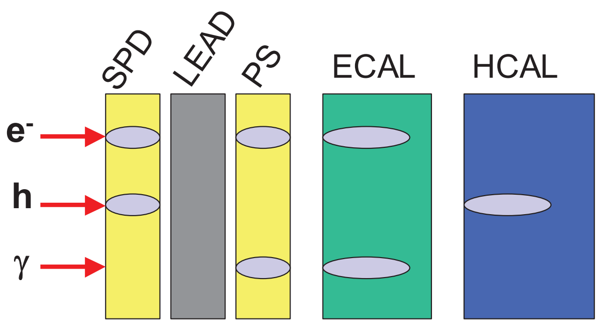


Figure 2.17: Energy deposited in the different parts of the calorimeter by an electron (e), a hadron (h) and a photon (γ).

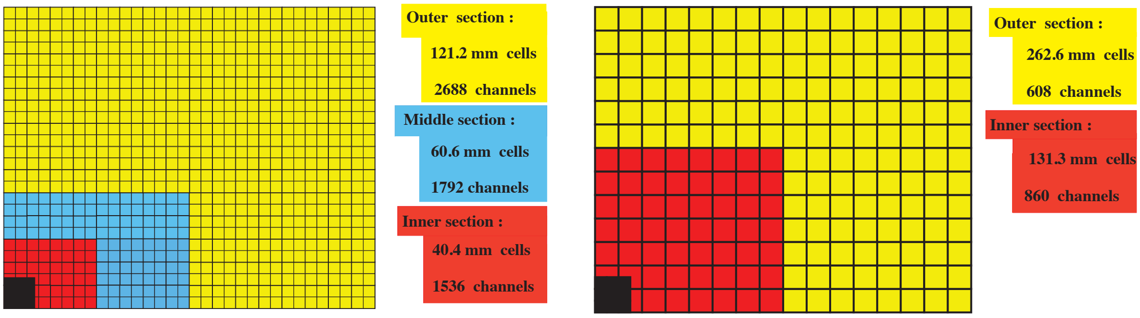


Figure 2.18: Left: frontal view of the SPD/PS and ECAL detectors where the three regions mentioned in the text are shown. Right: frontal view of the HCAL where there are only two regions.

The SPD and the PS are auxiliary sub-detectors of the electromagnetic calorimeter and they are placed in front of it. Both the SPD and the PS are composed by a scintillator plane about 15 mm thick. They are separated by a 2.5 radiation lengths [1] lead converter layer. The light emitted by the scintillator material is collected by means of wavelenght-shifting optical fibers (WLS) connected to clear fibers which convey the light to multi-anode photomultipliers placed outside the detector. The SPD determines if the particle hitting the calorimeter system is charged or neutral, since the former produces light when passing through a scintillator material while the latter does not. The PS determines the electromagnetic character of the particle, namely whether it is an electron/photon or not.

The ECAL is a sampling calorimeter using the Shashlik technology and separated into different independent modules. This particular type of calorimeters exploit WLS optical fibers which cross longitudinally the entire module and carry the scintillation light to the read-out photomultipliers. The ECAL modules are composed by 66 lead converter layers of thickness of about 2 mm separated by plastic scintillator layers which are about 4 mm thick. ECAL modules have a total size of about 25 radiation lengths and 1.1 nuclear interaction lengths [2]. Each module has a section of 12×12 cm^2. In the inner region, a module corresponds to 9 read-out channels of 4×4 cm^2. In the middle region, a module

[1]The radiation lenght is defined as follows

$$X_0 = \frac{A \cdot 716.4}{Z(Z+1)\ln(287/\sqrt{Z})}\ g/\mathrm{cm}^2 \tag{2.3}$$

where A is the mass number and Z is the atomic number of the material considered. This quantity corresponds to the distance over which the energy of an electron is reduced by a factor $1/e$ only due radiation loss [40].

[2]The nuclear interaction length $\lambda_I \propto A^{1/3}$, similarly to the radiation length X_0, is the mean path length required to reduce the energy of a relativistic charged particle passing through matter by a factor $1/e$.

has 4 read-out channels of 6×6 cm^2. In the outer region, a module corresponds to a single read-out channel. A schematic view of an ECAL module is shown in Fig. 2.19.

The HCAL provides the measurement of the energies of hadronic showers, which is the fundamental information for the Level-0 hadronic trigger. The HCAL modules consist of 4 and 6 mm thick iron plates interspaced with scintillating tiles arranged parallel to the beam pipe. HCAL modules have a total size of approximately 5.6 nuclear interaction lengths. Each module has a section of 13×13 cm^2 in the inner region and 26×26 cm^2 in the outer region. An HCAL module corresponds to a single read-out channel.

The performances of the calorimeter system have been evaluated with various tests performed before the start of the data taking. The energy resolutions of the calorimeter system are given by [41]

$$\begin{aligned} \text{ECAL}: \quad \frac{\sigma(E)}{E} &= \frac{(8.5-9.5)\%}{\sqrt{E}} \oplus 0.8\%, \\ \text{HCAL}: \quad \frac{\sigma(E)}{E} &= \frac{(69 \pm 5)\%}{\sqrt{E}} \oplus (9 \pm 2)\%. \end{aligned} \tag{2.4}$$

The ECAL calibration is performed by reconstructing resonances decaying to two photons like $\pi^0 \to \gamma\gamma$ and $\eta \to \gamma\gamma$. The calibration of the HCAL is achieved by measuring the ratio E/p between the energy E, as measured in the calorimeter, for a hadron with momentum p, as measured by the tracking system.

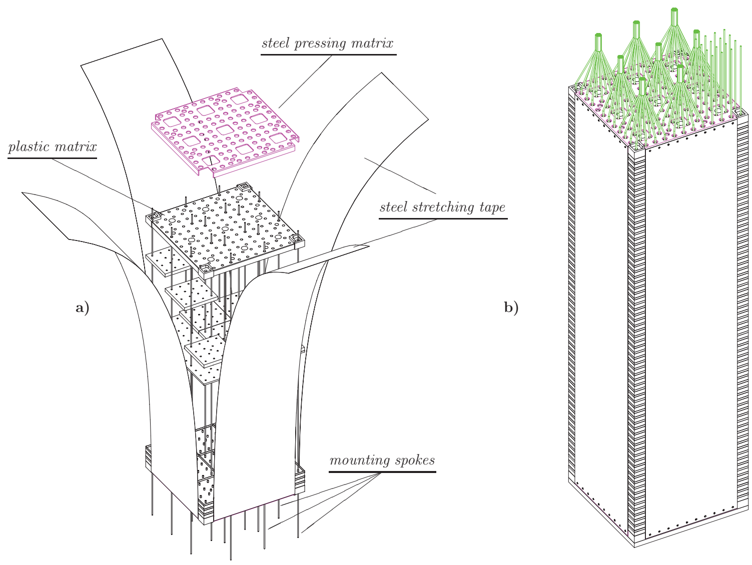


Figure 2.19: Left: picture of an ECAL module during the assembly phase, the lead/scintillator layers are also shown. Right: representation of an assembled ECAL module of the inner region, the green lines represent the optical fibers conveying the light to photo-multipliers.

2.4.3 The muon system

The final part of the LHCb detector is the muon system which provides the identification of muons. Muons are present as final decay products in different fundamental LHCb measurements such as $B^0 \to K^{*0}\mu^+\mu^-$ [42], $B_s^0 \to J/\Psi(\mu^+\mu^-)\phi$ [43] or $B_s \to \mu^+\mu^-$ [44], [45].

The muon system [46] (see Fig. 2.20) is made of five stations (M1 to M5) covering an angular acceptance of ± 300 mrad in the horizontal plane and ± 200 mrad in the vertical plane. This corresponds to a geometrical efficiency of approximately 46% for the detection of muons arising from B hadrons. The first muon station, M1, is placed before the calorimeters in order to avoid possible multiple scattering effects that could modify the particle trajectory. The remaining stations, M2 to M5, are placed after the calorimeter system, at the end of the LHCb detector, and are separated by iron planes 80 cm thick. Each muon station is divided into four regions (R1-R4) as shown in Fig. 2.21. The R1 region is the closest to the beam-pipe and has the most dense segmentation while the R4 region is the farther. The segmentation defined per region is such that the charged particle occupancy is expected to be approximately the same in each region. All the muon chambers are composed by Multi-Wire Proportional Chambers, except for the inner region

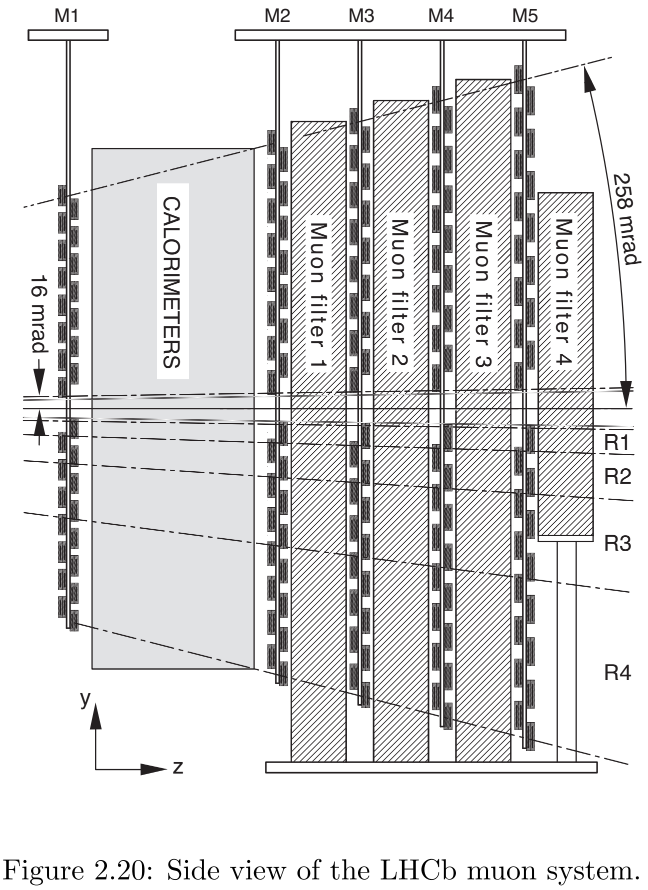


Figure 2.20: Side view of the LHCb muon system.

of the M1 station, which exploits three gas electron multiplier foils sandwiched between anode and cathode planes (GEM detectors). In total, the muon system consist of 1368 MWPC and 12 GEM detectors.

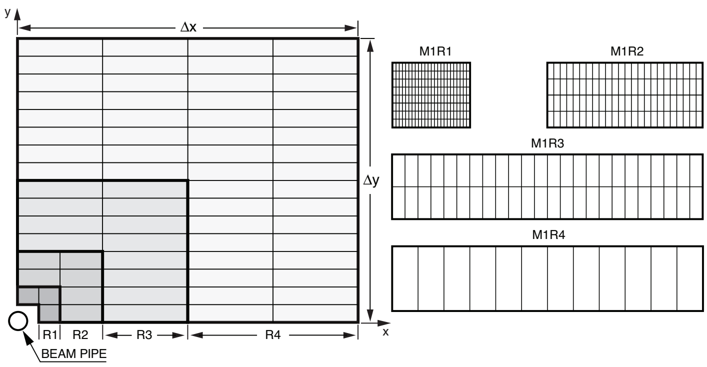


Figure 2.21: Left: front view of a quadrant of a muon station, each rectangle represents a chamber. Right: segmentation of the four types of chambers corresponding to the four region of M1. In M2, M3 (M4, M5), the number of columns per cell is double (half) with respect to M1, while the number of rows is the same.

2.4.4 Particle identification methods

The particle identification (PID) in LHCb is achieved by combining the information coming from the various sub-detectors. The RICH detectors, the calorimeters and the muon stations are used for the identification of charged particles (e, μ, π, K and p), while photons (γ) and neutral pions (π^0) are identified using the calorimeter system.

For each particle the available PID information is elaborated in two variables of different nature, but with the same purpose: the log-likelihood difference (DLL) and the ProbNN variable which has been introduced later in the collaboration.

The first variable, the DLL, is defined as the difference between a given PID hypothesis (x) and the pion hypothesis as

$$DLL_{x\pi} = \ln \mathcal{L}_x - \ln \mathcal{L}_\pi = \ln \left(\frac{\mathcal{L}_x}{\mathcal{L}_\pi} \right) \ , \tag{2.5}$$

where each likelihood function $\mathcal{L}_i$ ($i = x$ or π) combines the information coming from the various PID sub-detectors. The DLL_{xy} related to any particle hypotheses x and y can then be defined accordingly. The higher the variable DLL_{xy} is, the higher the probability of the candidate to be x.

The second kind of variable, the ProbNN, is built by running multivariate analysis tools (in particular Neural Networks) based on the detector PID information. Differently from the likelihood functions, the multivariate analyses take into account the correlations between the information coming from the different detectors. The ProbNN variables produced as output are defined between 0 and 1, as a probability would be, and can be used to separate between different charged tracks: in particular they are referred to as ProbNNx according to the particle hypothesis which is tested.

In the analysis presented in this thesis, ProbNN variables are used to distinguish between charged hadrons (kaon, pion and protons). For what concerns the identification of neutral particles, which is very important in analyses involving radiative decays, this is achieved using dedicated PID variables described in detail in Chap. 3.

2.5 The LHCb trigger

As already said, the production cross section of $b\bar{b}$ pairs is rather large but still two orders of magnitude smaller than the inelastic pp cross section $\sigma_{\rm inel} = (66.9 \pm 2.9 \pm 4.4)$ mb [47]. Furthermore the capabilities to store data are naturally limited by cost and technological reasons. This means that the LHCb trigger has to efficiently select the interesting events while rejecting most of the background events. The necessary performances have been achieved by separating the trigger into different levels, each processing the output of the previous level as shown in Fig. 2.22. The LHCb trigger system is divided into three stages:

Level-0 (L0): this is the first level and consists of a hardware trigger based on custom electronics. It is designed to accomplish a first filtering of the events and to reduce significantly the input rate of 40 MHz to an output rate of only 1 MHz;

High Level Trigger 1 (HLT1): the second trigger level is software based. The HLT1 purpose it to filter heavy hadron events in an inclusive way and to reduce the rate to 50 kHz;

High Level Trigger 2 (HLT2): this is the last trigger level and it is also software based. The HLT2 trigger further reduces the output rate to 5 kHz applying exclusive and inclusive selections of beauty and charm decays. The output is sent to mass storage.

The L0 trigger operates synchronously with the 40 MHz LHC clock and a fixed latency of 4 μs. It reduces the rate to about 1 MHz at which the entire detector may be read out. The L0 information is coming from the pile-up sensors of the VELO, the calorimeters and the muon system. It is sent to the *Level 0 Decision Unit* (L0DU) where the L0 selection algorithms are run.

The L0 calorimeter trigger system uses information from the four components of the calorimeter system, SPD, PS, ECAL and HCAL. It computes the transverse energy deposited in 2×2 clusters. Then the clusters with the largest E_T are selected and identified either as hadron, photon or electron candidates. The hadron candidate is defined from the highest E_T HCAL cluster. The photon candidate is the highest E_T ECAL cluster

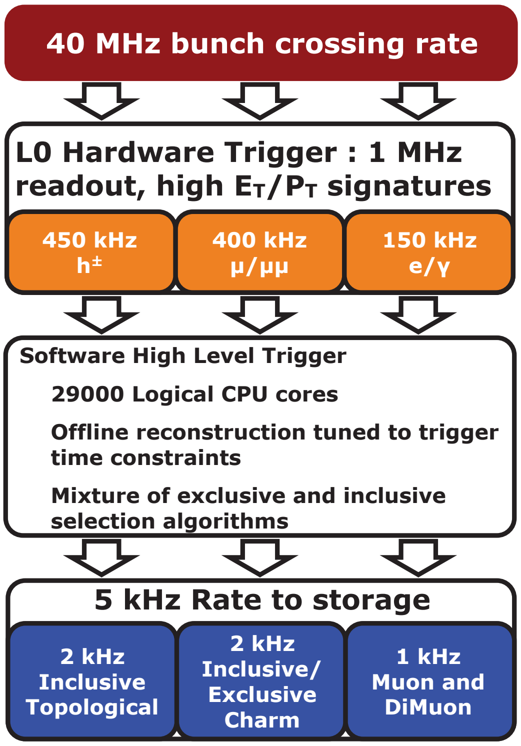


Figure 2.22: Flow-diagram of the different trigger stages in Run 1. Software High Level Trigger indicates HLT1 and HLT2 stages.

with at least one PS cell with energy higher than a 5 MIP (Minimum Ionizing Particle) threshold and no hit in the corresponding SPD cells. The electron candidate is defined as the photon candidate except that at least one hit is required in the SPD cells in front of the PS cells above the 5 MIP threshold. The total number of hits in the SPD is also computed. It is used to reject high multiplicity events which would proportionally take too much time to be processed in the High Level Trigger.

The L0 muon trigger system performs a stand-alone reconstruction of muon tracks with a p_T resolution of $\sim 25\%$. Tracks are searched combining the pad data from the five muon stations to form towers pointing towards the interaction region. The muon stations are divided in quadrants and there is no exchange of information between the quadrants. In each quadrant, the two muon candidates with highest p_T are selected.

The High Level Trigger (HLT) is a fully software trigger. It consists of a C++ application executed on an Event Filter Farm (EFF) composed of 29000 CPU cores running up to about 26000 copies of the application. The HLT application has access to all the data of a given event and runs the selection algorithms, called “trigger lines”, optimized to cover a certain class of events of interest. In Run 1 the HLT processing time per event was of the order of 30 ms. The HLT reduces the rate from the level 0 output to ~ 5 kHz. For timing reasons, the HLT is divided into two stages, HLT1 and HLT2.

The HLT1 reduces the level 0 input rate by a factor of about 20. It performs a partial event reconstruction limited to high transverse momentum tracks ($p_T > 1$ GeV/c or so) using information from the VELO and the tracking stations. The selection requires a single high p_T track displaced for all primary vertices (PV). Along with some track quality requirements, the track should have an IP larger than 125 μm with respect to any PV, $p_T > 1.8$ GeV/c and $p > 12.5$ GeV/c. For events triggered by the Level 0 photon and electron lines, the p_T requirement is relaxed to 0.8 GeV/c.

At the HLT2 stage a simplified full event reconstruction is performed. HLT1 rate allows an HLT2 tracking close to the offline one. Only tracks with $p > 5$ GeV/c and $p_T > 0.5$ GeV/c are reconstructed. The HLT2 runs exclusive and inclusive selections. Special inclusive lines have been developed to trigger on partially reconstructed b hadron decays. These so-called topological lines [48] are based on displaced vertices with 2, 3 or 4 associated tracks. The topological lines were first implemented as cut based selections. To improve the performances, additional lines using a multivariate approach were then added [49].

2.6 Luminosity measurements in LHCb

The number of selected events of a given process per unit of time, denoted as dn/dt, is given by

$$\frac{\mathrm{d}n}{\mathrm{d}t} = \sigma\mathcal{L}\epsilon \tag{2.6}$$

where σ is the process cross section, $\mathcal{L}$ is the instantaneous luminosity and ϵ is the total efficiency accounting for the detector acceptance as well as the reconstruction and selection efficiencies. The determination of the luminosity is essential for the measurement of any cross section. The average instantaneous luminosity of two colliding bunches can be expressed as

$$\mathcal{L} = \frac{fN_1N_2}{4\pi\sigma_x\sigma_y} \tag{2.7}$$

where f is the revolution frequency (11245 Hz at the LHC), N_1 and N_2 are the number of protons in the two bunches, σ_x and σ_y are the transverse sizes of the bunch at the interaction point along the x and y axis respectively.

At LHCb two methods to determine the absolute luminosity have been implemented [50]: the Van der Meer scan and the beam-imaging gas method. In the first method the beams are moved in transverse directions in order to investigate the beam transverse profiles counting the interaction rate as a function of the beam offsets. In the second one the high acceptance of the VELO around the interaction point is used to reconstruct beam-gas vertices produced by the collision of protons in the beam with molecules in the remaining gas of the beam pipe. The positions of the beam-gas interactions are used to determine beam angles and profiles [51]. Combining the two methods, the absolute luminosity can be determined with a relative precision of 3.5%, this allows to calculate a reference cross section of visible interactions. Then, dedicated luminosity counters are defined in order to follow the evolution of the instantaneous luminosity during the data taking.

Chapter 3

Neutral PID and calibration samples

The reconstruction of neutral particles is achieved by combining the information coming from the Electromagnetic (ECAL) and Hadronic (HCAL) CALorimeters together with the information coming from the Scintillating Pad Detector (SPD) and the PreShower (PS) auxiliary sub-detectors. The identification of neutral objects and their separation is a key element when performing physics analyses with decays involving photons or neutral pions in the final state: in particular, a good γ/π^0 separation is essential for the study of radiative decays such as $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$. This is something very challenging since LHCb has not been specifically designed to accomplish it. Neutral pions decay into two photons with a branching of $(98.823 \pm 0.034)\%$ [7] and for high p_{T} values, typically $p_{\mathrm{T}} > 2.5$ GeV/c, the two resulting γ start to merge in a single ECAL cell faking a true radiative photon. These neutral pions are called *merged* π^0 (see Fig.3.1). Decays that involve merged π^0 are an important source of background for radiative channels. Conversely, when the two photons of a π^0 decay are detected in two separate ECAL cells, we have a *resolved* π^0.

In this chapter, after an introduction on the basic concepts of neutral objects reconstruction, we describe the two main tools used in the collaboration to achieve a suitable photon identification and γ/π^0 separation. In order to perform a complete physics analysis, it is necessary to evaluate the γ/π^0 separation efficiency with a data-driven approach by means of dedicated calibration samples. In the second part of the chapter, we discuss how these samples were produced detailing both the event selection and the fit procedure.

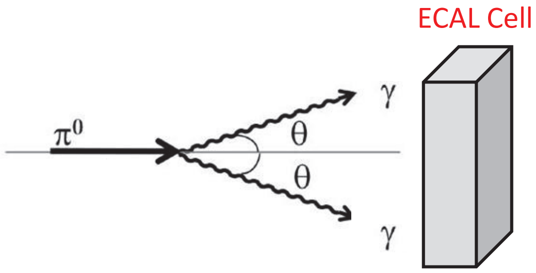


Figure 3.1: Sketch illustrating a merged π^0.

3.1 Photon reconstruction

The energy from the interaction of photons and electrons with the detector is mainly absorbed in the ECAL. The reconstruction process of electromagnetic showers begins with the identification of the ECAL cell that has an excess in energy deposition compared to all its direct neighbors. Such cell is also called *seed*. The number of neighbor cells is 8 per 'regular' cells which are far from boundaries of ECAL detector, and it varies from 3 to 9 for cells close to the boundaries. These cells are selected only if the transverse energy is larger than 50 MeV [52]. The reference cell will originate the cluster according to the clusterisation procedure adopted in the Cellular Automaton algorithm [53]. As direct consequence of these formal definitions, the seed cells of two reconstructed clusters are always separated at least by one cell. If a calorimeter cell is shared between two overlapping clusters, the energy of the cell is shared between the clusters depending on the cluster energies and its distance to the clusters' barycenters.

In order to determinate the 4-momentum of the corresponding particle, the total energy of the electromagnetic shower and the transversal barycenter are calculated according to [54]

$$\epsilon = \sum_i \epsilon_i \quad x_b = \frac{1}{\epsilon}\sum_i x_i\epsilon_i \quad y_b = \frac{1}{\epsilon}\sum_i y_i\epsilon_i \ , \tag{3.1}$$

where ϵ_i stands for the energy deposit in each cell of the cluster, x_i and y_i are the coordinates of the cell center and the sum runs over all cells forming the cluster.

The selection of neutral clusters ("photon candidates") is performed using anti-coincidence techniques with reconstructed tracks. In order to discriminate photons from charged particles, the main criterion is the isolation of the cluster with respect to the tracks extrapolated to the ECAL reference plane. This requirement is implemented defining the bi-dimensional χ^2

$$\chi^2_{2D}(\vec{\mathbf{p}}) = (\vec{\mathbf{p}}_{tr} - \vec{\mathbf{p}})^T C_{tr}^{-1}(\vec{\mathbf{p}}_{tr} - \vec{\mathbf{p}}) + (\vec{\mathbf{p}}_{cl} - \vec{\mathbf{p}})^T S_{cl}^{-1}(\vec{\mathbf{p}}_{cl} - \vec{\mathbf{p}}) \ , \tag{3.2}$$

where $\vec{\mathbf{p}}_{tr}$ is the extrapolated track 2D-point to the calorimeter plane, C_{tr} is the covariance matrix associated to $\vec{\mathbf{p}}_{tr}$ parameters, $\vec{\mathbf{p}}_{cl}$ is the cluster barycenter position and S_{cl} is the 2×2 cluster second momenta matrix. For each track the χ^2_{2D} is minimized with respect to the 2D-point $\vec{\mathbf{p}}$ in the calorimeter plane. The smallest value over all the tracks is used to distinguish neutral particles clusters (large χ^2_{2D}) from charged particles ones (low χ^2_{2D}): in particular clusters with $\chi^2_{2D} > 4$ are selected as photon candidates. This criterion is highly efficient on clusters related to photons and allows to significantly suppress the clusters due to charged particles such as electrons.

We can distinguish two types of photons, namely converted and unconverted. Due to the material placed between the interaction point and the calorimeter system, 30% of the photons coming from the interaction region convert into a e^+e^- pair before the calorimeter front face. There are two kinds of converted photons, namely photons converted before the magnet (about 23% of the converted photons [55]) or after the magnet. In the first case the tracks associated to the electrons pair can be reconstructed, then the photon information is coming from two separated electron tracks. In the second case the electrons mostly end up in a single ECAL cluster and the charged nature of the conversion can be identified using the signal left in the SPD. A cluster with no matching tracks but with a deposit in the SPD is the signature of a photon converted after the magnet. Dedicated calibrations are determined for those converted photons.

The geometry of the calorimeter system has been designed in such a way that the ECAL cells coincide with the PS and SPD cells. The photon energy is evaluated by summing the ECAL cluster energy with the energy deposit in the PS cells in front. The main energy loss arises from the fact that the cluster is reconstructed as a 3×3 matrix of ECAL cells centered around the cluster seed and the transversal energy deposited outside this matrix is not taken into account. Additionally, the dead material between the modules of the calorimeter cells introduces intrinsic losses that need to be compensated. These leakages are corrected according to

$$E_c = \alpha E_{ECAL} + \beta E_{PS} \ , \tag{3.3}$$

where E_c is the corrected photon energy, E_{ECAL} stands for the energy deposit in the 3×3 ECAL cluster and E_{PS} is the measured energy in the PS cells. The α and β coefficients are extracted from data and depend on the ECAL region.

The transversal barycenter of the shower, (x_c, y_c), is evaluated from the energy-weighted barycenter of the cluster (see Eq. 3.1) corrected from the non linear transversal profile of the shower shape (hereafter named as S-correction). The transversal shower shape is described with a single exponential in the first order approximation $E(r) \sim E_0 e^{-r/\mathbf{b}}$, where r is the relative position inside the cluster and $\mathbf{b}$ is a constant tuned on data. The $\mathbf{b}$ parameter is found to be about 10%, 13% and 15% of the cell size for the outer, middle and inner ECAL regions respectively. In this approach, the S-corrected barycenter (x_c, y_c) is given by the single parameter S-function [56]

$$(x_c, y_c) = \mathcal{S}_0[(x_b, y_b), \mathbf{b}] = \mathbf{b} \ \mathrm{asinh} \left[\frac{(x_b, y_b)}{(\Delta/2)} \cosh \frac{(\Delta/2)}{\mathbf{b}} \right] \ , \tag{3.4}$$

where (x_b, y_b) are the coordinates of the energy-weighted barycenter and $\Delta/2$ is the half cell size. The transversal barycenter position needs to be further corrected from left/right asymmetries in the X-direction (bottom/up in the Y-direction) which are due to the incidence angle of the photon that induces a non spherical profile of the energy deposit. A correction to $\mathcal{S}_0$ that accounts for the assumption of first order exponential profile of the shower shape is also considered. The total S-correction is then given by

$$(x_c, y_c) = \mathcal{S}_0[(x_b, y_b), \mathbf{b}] + \mathcal{S}_1[(x_b, y_b)] + \mathcal{S}_2[(x_b, y_b)] \ , \tag{3.5}$$

where $\mathcal{S}_1$ and $\mathcal{S}_2$ correct from residual S-shape and geometrical asymmetries respectively. These corrections allow to improve the resolution on the impact position of the photon [54].

The longitudinal z coordinate needed to identify the three-dimensional barycenter of the cluster (x_c, y_c, z_c) needs to be corrected as well. The final z_c position, which has to account for the penetration depth of the photon, is defined as

$$z_c = z_{ECAL} + \alpha_c \ln(E_c) + \beta_c E_{PS} \ , \tag{3.6}$$

where z_{ECAL} is the position of the front face of the ECAL along the beam axis with respect to the nominal collision point. The coordinate z_c depends on the logarithm of the energy and, in order to take into account the position of the beginning of the shower, a dependance with E_{PS} is included. The photon 4-momentum is finally evaluated from (E_c, x_c, y_c, z_c) assuming the photon is coming from the nominal interaction point, which is $(0, 0, 0)$ by definition.

The performance of high energy photon reconstruction is illustrated by the reconstructed $B^0 \to K^{*0}\gamma$ mass distribution displayed in Fig. 3.2. The resolution on the reconstructed B^0 mass using calorimetric photons is found to be 93 MeV/c^2 and is dominated by the ECAL energy resolution.

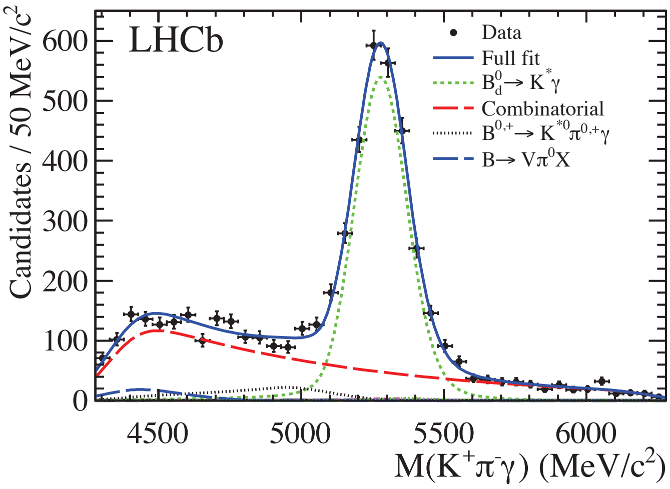


Figure 3.2: Mass distribution of reconstructed $B^0 \to K^{*0}$ $(K^+\pi^-)\gamma$ candidates obtained using 2011 data. The blue curve corresponds to the mass shape fit. The $K^*\gamma$ signal (green dotted line) and the various background contaminations are shown [57].

3.2 π^0 reconstruction

The π^0 signature in the ECAL depends on its kinematics, the higher is the momentum of the π^0 the closer the two photons are at the entry of the calorimeter. These two photons can then produce two separated clusters or share a single cluster in which their individual signals are not clearly distinguishable. The π^0 are classified as resolved π^0 in the former case and as merged π^0 in the latter one. The transverse momentum spectrum of merged π^0 starts around 2 GeV/c.

3.2.1 Resolved π^0

Resolved π^0 are reconstructed pairing the two photons in the final state and requiring their reconstructed invariant mass $m_{\gamma\gamma}$ to be in the range [105; 165] MeV/c^2, with the π^0 mass. Only photons with transverse momentum $p_T > 200$ MeV/c and with a track matching χ^2_{2D} greater than 4 are taken into account. Using those criteria and according to the simulation, the global reconstruction and identification efficiency of resolved π^0, with respect to events where both photons are in the ECAL acceptance, both with transverse momentum greater than 200 MeV/c, is about 50% [54]. The inefficiency is mostly due to photons showering in the material upstream the calorimeter. Part of this inefficiency can be recovered by considering converted photons reconstructed as a pair of electrons. Fig. 3.3 displays the invariant mass distribution for $\pi^0 \to \gamma\gamma$ candidates obtained using Run 1 LHCb data [58]. The plot shows the $\gamma\gamma$ mass distribution before (red) and after (blue) applying the fine ECAL calibration procedure. The details of this procedure are not reported here and can be found in Ref. [59]. Looking at the blue curve in Fig. 3.3 we

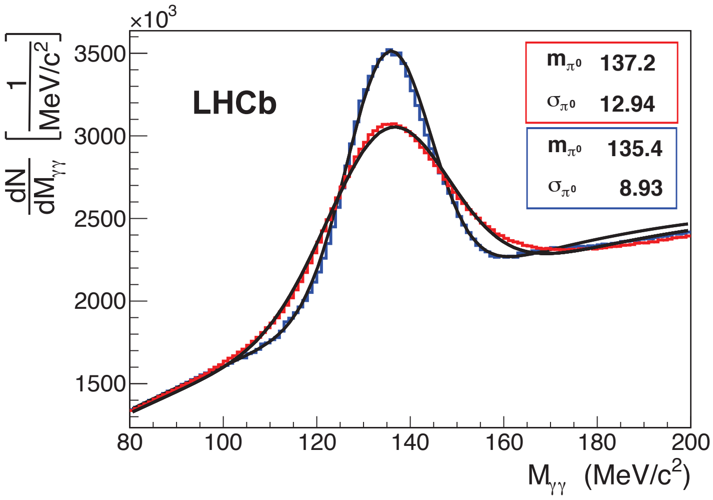


Figure 3.3: Invariant mass distribution for $\pi^0 \to \gamma\gamma$ candidates before (red curve) and after (blue curve) applying the fine ECAL calibration procedure. Values in the red (blue) box are the mean and sigma of the signal peak distribution in MeV/c^2 before (after) applying the calibration.

can notice that the final resolution on $m_{\gamma\gamma}$ is about 9 MeV/c^2.

3.2.2 Merged π^0

Each electromagnetic cluster is split in two subclusters defined from the two most energetic cells in the cluster. An algorithm calculates the barycenter of each sub-cluster using the expected transverse shower shape of individual photons. The positions of the two barycenters depend on the energy sharing between the two sub-clusters, which itself depends on the positions of the two barycenters. According to that, the calculation is done using an iterative procedure [54]. After the preparation of the two photon sub-clusters, the following criteria are applied to identify the cluster as arising from a π^0.

- The cluster is identified as coming from a neutral particle requiring a track matching $\chi^2_{2D} > 1$.
- The cluster energy has to be compatible with the merged π^0 hypothesis. To ensure that, a cut is applied on the minimal distance allowed by the kinematics between the impacts of the two photons on the ECAL front face

  $$d_{\gamma\gamma} = 2 \times z_{ECAL} \times m_{\pi^0}/E_{\pi^0} < 1.8\Delta \ , \tag{3.7}$$

  where Δ is the transverse size of the calorimeter cell, m_{π^0} is fixed to 135 MeV/c^2, E_{π^0} is the π^0 energy and z_{ECAL} is the position of the ECAL front face with respect to the nominal interaction point.

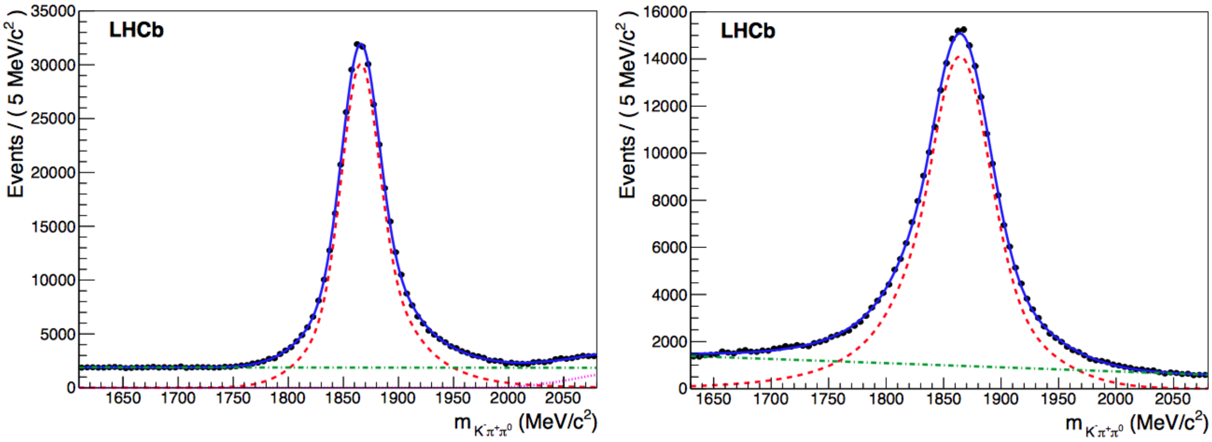


Figure 3.4: Mass distributions of reconstructed $D^0 \to K^-\pi^+\pi^0$ candidates with resolved π^0 (left) and merged π^0 (right). Both are obtained using the 2011 data sample. The overall mass fit [60] is represented by the blue curve, with the signal (red dashed line) and background (green dash-dotted line and purple dotted lines) contributions also shown.

- The reconstructed merged π^0 mass is required to be in the range 75 to 195 MeV/c^2. The wider mass range compared to resolved π^0 is due to the wider $m_{\gamma\gamma}$ resolution, which is typically 12-15 MeV/c^2.

The performance of neutral pion reconstruction is illustrated in Fig. 3.4, which shows the invariant mass distribution of $D^0 \to K^-\pi^+\pi^0$ candidates for resolved (left plot) and merged (right plot) π^0 candidates [60]. In this example, the estimated resolution on the invariant mass of the reconstructed D^0 is about 20 MeV/c^2 for the resolved π^0 candidates and 30 MeV/c^2 for merged ones.

3.3 Photon identification

The default algorithm currently used in the collaboration in order to achieve the photon identification (ID) is a Neural Network (specifically, a Multi-Layer Perceptron) discriminant. Two independent estimators, IsNotH and IsNotE, are built to establish the photon hypothesis: the former is aimed to distinguish the γ from non-electromagnetic deposits associated to hadrons while the latter allows to separate between photons and electrons. These classifiers are described in Ref. [61] and represent a significant improvement compared to the previous photon identification procedure used in LHCb.

Here we discuss only the IsNotH discriminant (also called γ_{CL}). It exploits the following variables:

- The track matching χ^2_{2D}.

- The energy deposits in the PS: E_{PS}.

- The ratio of the energy of the seed cell and that of the 3×3 cluster in the ECAL: E_{seed}/E_{ECAL}.
- The ratio of the energy in the PS cell in front of the seed cell and the total energy measured in the PS: E_1/E_{PS}.
- The ratio of the energy in the HCAL, in the projective area matching the cluster, and the cluster energy in the ECAL: E_{HCAL}/E_{ECAL}.
- The 2×2 matrix of PS cells with the highest energy deposit in front of the reconstructed cluster: PSE4Max.
- The shower shape variables

$$S_{XX} = \frac{\sum_{i=1}^{N} \epsilon_i (x_i - x_c)^2}{\sum_{i=1}^{N} \epsilon_i}, \quad S_{YY} = \frac{\sum_{i=1}^{N} \epsilon_i (y_i - y_c)^2}{\sum_{i=1}^{N} \epsilon_i}, \tag{3.8}$$

$$S_{XY} = S_{YX} = \frac{\sum_{i=1}^{N} \epsilon_i (x_i - x_c)(y_i - y_c)}{\sum_{i=1}^{N} \epsilon_i}, \tag{3.9}$$

  where $(x_c,\, y_c)$ is the transversal barycenter of the cluster defined in Eq. 3.4.
- The second order momenta $r2$ related to the spread of the shower and defined using the shape variables of Eq. 3.8

$$r2 = \langle r^2 \rangle = S_{XX} + S_{YY} = \frac{\sum_{i=1}^{N} \epsilon_i [(x_i - x_c)^2 + (y_i - y_c)^2]}{\sum_{i=1}^{N} \epsilon_i} \tag{3.10}$$

- The number of PS cells in front of the cluster with non-zero energy deposit: PSMulti.
- The multiplicity of hits in the 3×3 SPD cells matrix in front of the reconstructed cluster: SPDMulti.

The combination of these variables in the Neural Network classifier gives as output the IsNotH distributions displayed on the left side of Fig. 3.5 for $B^0 \to K^{*0}\gamma$ data (red dots) and MC (blue triangles). The efficiencies as a function of the IsNotH (NN in the plot) applied cut for $B^0 \to K^{*0}\gamma$ data (dashed red line) and MC (solid blue line) are shown on the right side of Fig. 3.5. In both plots data points are background subtracted. The IsNotH distributions are obtained using $B^0 \to K^{*0}\gamma$ events selected according to the selection described in Sec. 3.5.1 while the efficiency plot is taken from [61]. As can be noticed, for low values of the IsNotH cut data and MC exhibit a good agreement. This will be one of the criteria when optimizing the IsNotH cut in the $B^0 \to \rho^0\gamma$ analysis (see Sec. 5.3).

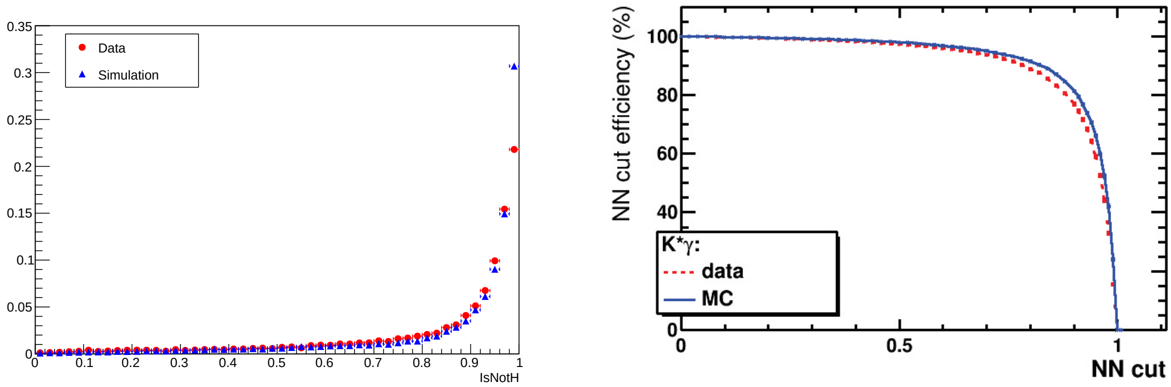


Figure 3.5: IsNotH distributions (left) and corresponding efficiencies [61] (right) for $B^0 \to K^{*0}\gamma$ data (red dots/dashed line) and MC (blue triangles/solid line). Data points are background subtracted.

3.4 γ/π^0 separation

The γ/π^0 separation, needed to reject backgrounds coming from high p_{T} neutral pions, is achieved using a dedicated multivariate-based tool developed in LHCb [62]. Such tool exploits a specific set of variables describing the shape of the cluster similarly for the PreShower (PS) and the ECAL. In addition multiplicity variables in the PS are also used. All the variables are combined in a Neural Network called IsPhoton, whose output allows to obtain a good discrimination between γ and π^0.

The discriminating variables used for the ECAL, some of which are defined according to the shower shape variables of Eq. 3.8 and 3.9, are:

- The second order momenta r2 of Eq. 3.10.
- The quantity

$$r2r4 = 1 - \frac{\langle r^2 \rangle^2}{\langle r^4 \rangle} \ , \tag{3.11}$$

  $\langle r^4 \rangle$ being defined similarly to $\langle r^2 \rangle$ (see Eq. 3.10), that informs about the importance of the tails.

- The κ variable related to the ratio of the eigenvalues of the matrix S, which is $(1+\kappa)/(1-\kappa)$. This way κ relates to the major and minor semiaxes of an ellipse

$$\kappa = \sqrt{1 - 4\frac{S_{XX}S_{YY} - S_{XY}^2}{(S_{XX}+S_{YY})^2}} = \sqrt{1 - 4\frac{\det S}{\mathrm{Tr}^2 S}} \ . \tag{3.12}$$

  This variable describes how squeezed is the cluster [63].

- The $asym$ variable which provides information about the orientation of the ellipse related to the correlation between X and Y coordinates

$$asym = \frac{S_{XY}}{S_{XX}S_{YY}} \ . \tag{3.13}$$

- The ratio E_{seed}/E_{ECAL} between the energy of the seed cell and the energy of the cluster in the ECAL.
- The ratio $(E_{seed} + E_{2nd})/E_{ECAL}$, where E_{2nd} is the energy of the ECAL cell with the second highest energy deposit in the cluster.

Analogous variables are defined for the PS:

- $r2PS$ and $asymPS$ using the 3×3 PS cells matrix in front of the ECAL cluster.
- The energy ratios E_{max}/E_{PS} and E_{2nd}/E_{PS}, where E_{max} (E_{2nd}) is the energy measured in the PS cell with the (second) highest energy deposit and E_{PS} is the total energy deposited in the PS cluster.

In addition, a set of variables related to the multiplicity of hits in the 3×3 PS cells matrix is used. These variables, which have different requirements on the minimum energy deposited in the cells, are $multi$, $multi15$, $multi30$ and $multi45$, where the number refers to the energy threshold in MeV (the threshold for $multi$ is 0).

The combination of these variables in the Neural Network classifier gives as output the IsPhoton distributions displayed on the left side of Fig. 3.6 for $B^0 \to K^{*0}\gamma$ (red dots) and $D^0 \to K^-\pi^+\pi^0$ (blue triangles) background subtracted data. We can notice that the IsPhoton distributions for high p_{T} photons and merged π^0 look very different, allowing to achieve a good γ/π^0 separation. On the right side of Fig. 3.6 is shown the photon identification efficiency with respect to the π^0 rejection efficiency for simulation and data. We can notice that a photon identification efficiency of 93% can be obtained while rejecting 50% of the merged π^0 reconstructed as photons. The IsPhoton distributions are obtained using $B^0 \to K^{*0}\gamma$ events selected according to the selection described in Sec. 3.5.1 and $D^0 \to K^-\pi^+\pi^0$ events selected according to the selection described in Sec. 3.5.2. The plot on the right side of Fig. 3.6 is taken from Ref. [58].

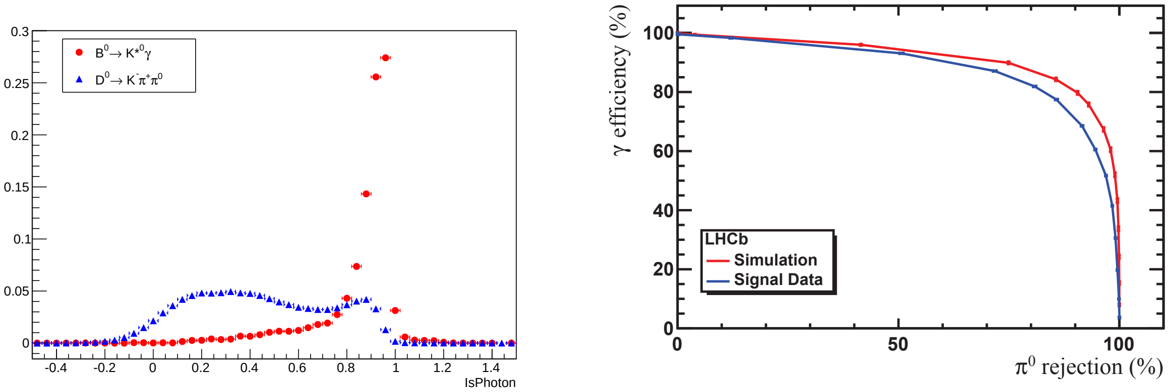


Figure 3.6: IsPhoton distributions (left) for $B^0 \to K^{*0}\gamma$ (red dots) and $D^0 \to K^-\pi^+\pi^0$ (blue triangles) background subtracted data. Photon identification efficiency as a function of π^0 rejection efficiency [58] (right) for the γ/π^0 separation tool for simulation (red curve) and data (blue curve).

3.5 Calibration samples

The IsPhoton variable is not correctly reproduced by the simulation then, when performing a physics analysis, we cannot rely on MC to extract the efficiency of a certain cut. In Fig. 3.7 is shown the comparison of IsPhoton distributions in data and MC for $B^0 \to K^{*0}\gamma$ (left) and $D^0 \to K^-\pi^+\pi^0$ (right) modes. We can notice that the plots exhibit a poor data/MC agreement. For this reason we use a data-driven approach using dedicated calibration samples that can be accessed with the `GammaPi0SeparationCalib` tool [64] available in the v5r0 release of the Urania package. The tool allows to extract the real IsPhoton efficiency as a function of the IsNotH cut, the pseudorapidity η and the p_{T} of the photon or the π^0 candidates.

The calibration samples have been produced exploiting the $B^0 \to K^{*0}(K^+\pi^-)\gamma$ decays for photons and $D^{*+} \to D^0(K^-\pi^+\pi^0)\pi^+$ decays for neutral pions. These samples contain the signal weights (sWeights) together with the variables relevant for the efficiency computation, namely IsPhoton, IsNotH, η and p_{T}. The signal weights, which are calculated after performing a mass fit and using the ${}_s\mathcal{P}lot$ technique [65], are needed to get background subtracted data samples that can be compared to the simulation. The calibration samples have been provided as part of the work carried out in this thesis and are the reference samples used in the collaboration to extract the γ/π^0 separation efficiency for Run 1 data. The events are reconstructed using the `Reco14` version of the reconstruction software for both 2011 and 2012 data taking campaigns. The data are stripped using `Stripping21r1` for 2011 and `Stripping21` for 2012. In the following we describe the event selection and the fit procedure used to produce the photon and neutral pions calibration samples.

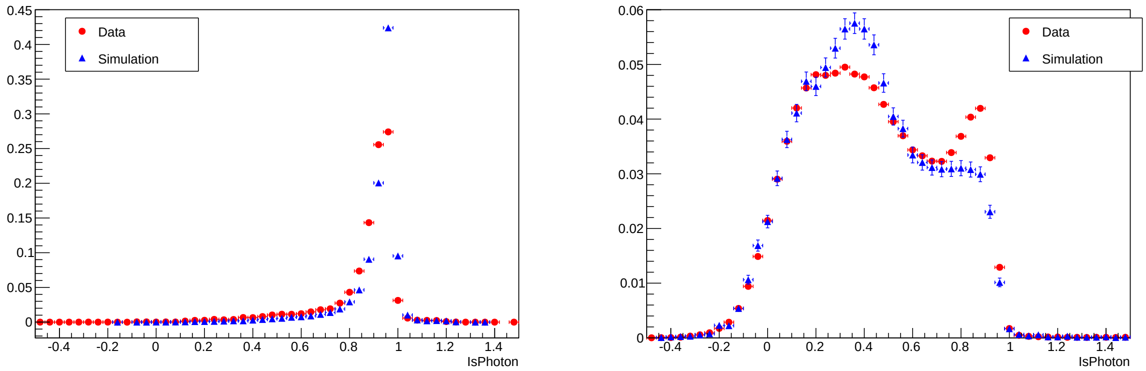


Figure 3.7: Data/MC comparison of the IsPhoton distributions for $B^0 \to K^{*0}\gamma$ (left) and $D^0 \to K^-\pi^+\pi^0$ (right) channels. Red dots stand for background subtracted data while blue triangles are for MC events.

3.5.1 High p_{T} photons

As already mentioned, the reference channel used to produce the calibration samples for calorimetric photons is the $B^0 \to K^{*0}\gamma$ mode. The first event selection is performed exploiting the inclusive stripping line `StrippingB2XGamma2pi_Line` for both 2011 and 2012 data (see Sec. 5.1). This algorithm allows to select a generic $B \to V(h^+h'^-)\gamma$ candidate with an intermediate vector resonance V, decaying into two charged hadrons, and a high transverse momentum photon in the final state.

The full trigger sequence used to select the $B^0 \to K^{*0}\gamma$ events is displayed in Tab. 3.1, where all the lines are “Triggered On Signal” (TOS). The L0 trigger conditions together with the HLT1 requirements are explained in Sec. 5.2.

The HLT2 trigger exploits the exclusive `Hlt2Bd2KstGamma` line that contains the requirements listed in Tab. 3.2. The definitions of the topological variables are given in Tab. 5.12. In order to fire this HLT2 line, events are required to pass the `L0Photon` or `L0Electron` lines and to have triggered any HLT1 line. The two tracks used to build the K^{*0} candidate are required to have both $p_T > 300$ MeV/c and $p > 3000$ MeV/c. They also have to be of good quality, namely having a track χ^2 per number of degrees of freedom lower than 5, and must point away from the pp interaction vertices by requiring

Trigger level	Standard path	High photon E_{T} path
L0	`L0Photon` || `L0Electron`	`L0PhotonHi` || `L0ElectronHi`
HLT1	`Hlt1TrackAllL0`	`Hlt1TrackPhoton`
HLT2	`Hlt2Bd2KstGamma`	

Table 3.1: Summary of the trigger conditions used to select $B^0 \to K^{*0}\gamma$ events.

		`Hlt2Bd2KstGamma`
L0 lines		`L0Photon` or `L0Electron`
HLT1		HLT1 Physics
Track $p_{\rm T}$	MeV/c	> 300
Track p	MeV/c	> 3000
Track χ^2/ndf		< 5
Track $\chi^2_{\rm IP}$		> 20
$m(K^{*0})$	MeV/c^2	$[m^{PDG}_{K^{*0}} - 100; m^{PDG}_{K^{*0}} + 100]$
$\chi^2_{\rm vtx}(K^{*0})$		< 16
$E_{\rm T}(\gamma)$	MeV	> 2600
$p_{\rm T}(B)$	MeV/c	> 3000
$\chi^2_{\rm IP}(B)$		< 12
$\theta_{DIRA}(B)$	mrad	< 45
$m(B)$	MeV/c^2	$[m^{PDG}_{B^0} - 1000; m^{PDG}_{B^0} + 1000]$

Table 3.2: Selection cuts implemented in the $B^0 \to K^{*0}\gamma$ HLT2 exclusive line.

the minimum Impact Parameter (IP) χ^2 with respect to any Primary Vertex (PV) to be greater than 20. The K^{*0} candidate is selected in a mass window of ± 100 MeV/c^2 around the nominal mass ($m^{PDG}_{K^{*0}} = 895.55 \pm 0.20$ MeV/c^2 [7]) and is required to have a decay vertex with χ^2 lower than 16. The γ candidate is required to have $E_{\rm T} > 2600$ MeV/c. The B candidate must have a transverse momentum greater than 3000 MeV/c and an IP χ^2 lower than 12. In addition it is required to have a direction angle $\theta_{DIRA}(B)$ lower than 45 mrad and it is selected in a mass window of ± 1000 MeV/c^2 around the nominal mass ($m^{PDG}_{B^0} = 5279.63 \pm 0.15$ MeV/c^2 [7]).

The events passing the trigger conditions of Tab. 3.1 and the inclusive stripping line are further filtered applying the event selection of Tab. 5.5, except for the IsPhoton variable for which no cut is applied here.

The $B^0 \to K^{*0}\gamma$ selected events are used to perform the mass fit needed to calculate the sWeights. The fit model consists of a double-tail Crystal Ball Probability Density Function (PDF) to describe the signal shape, an exponential function for the combinatorial background and two ARGUS functions convoluted with a Gaussian PDF to describe the partially reconstructed backgrounds (see Chap. 6 for more details about the PDFs definition). For these last ones we considered two generic contributions associated respectively to one and two missing π in the final state. The shapes of the signal and of the first ARGUS (one missing π) are fixed on MC, while the parameters of the other PDFs are left free to float in the fit. A further clarification concerning the B mass window is needed. As already mentioned, the mass window for the B candidate required by the HLT2 trigger line is ± 1000 MeV/c^2 around the B^0 nominal mass. However, this mass window cannot be fitted as it is due to the effects of calorimeter miscalibration and ageing: in particular a

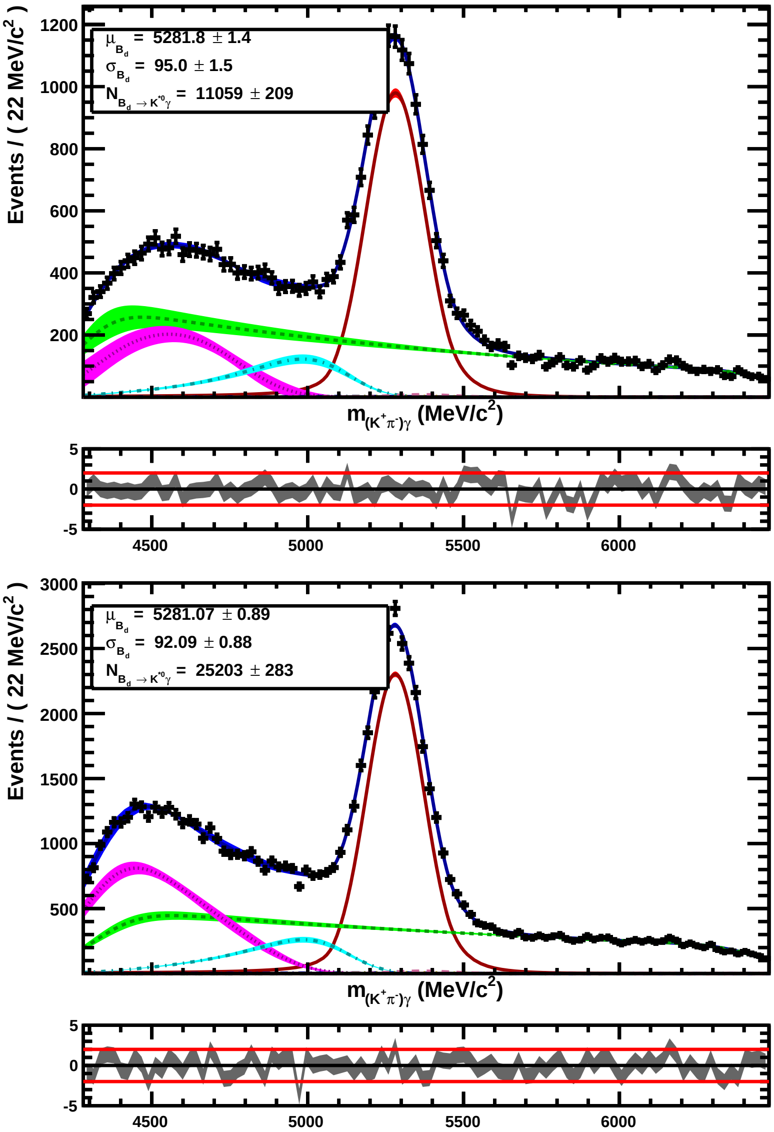


Figure 3.8: $B^0 \to K^{*0}\gamma$ mass fits for 2011 (top) and 2012 (bottom) data. The black points represent the data and the solid blue line is the total fit function. The signal is described with a double-tail Crystal Ball PDF (solid red line), the combinatorial background with an exponential function (dashed green line), and the partially reconstructed backgrounds for one (dashed cyan line) and two (dashed magenta line) missing π with ARGUS functions convoluted with a Gaussian PDF.

significant ageing of the ECAL response was experienced along the 2011 data taking period. Such effect was corrected by applying a set of time-dependent cell-by-cell corrections in the offline data reconstruction processing. These corrections were not applied at the trigger level and therefore data were collected with an incomplete ECAL calibration. As a result, the ± 1000 MeV/c^2 mass window requirement imposed by the trigger causes a bias in acceptance near the limits of this window. The inefficiency at the edges of the mass window is modeled by including a three-parameter threshold function in the fit model [57]

$$T(m_B) = \left[1 - \text{erf}\left(\frac{m_B - t_L}{\sqrt{2}\sigma_d}\right)\right] \times \left[1 - \text{erf}\left(\frac{t_U - m_B}{\sqrt{2}\sigma_d}\right)\right] , \tag{3.14}$$

where erf is the Gauss error function, the parameter t_L (t_U) represents the actual lower (upper) mass threshold and σ_d is the resolution.

The 2011 and 2012 mass fits to the $B^0 \to K^{*0}\gamma$ spectrum used to compute the sWeights are shown in Fig. 3.8. They correspond to 36×10^3 signal events over the full Run 1 statistics. The distribution in the $(p_{\rm T}, \eta)$ plane of the background subtracted high $p_{\rm T}$ photons from the calibration samples is shown on Fig. 3.9.

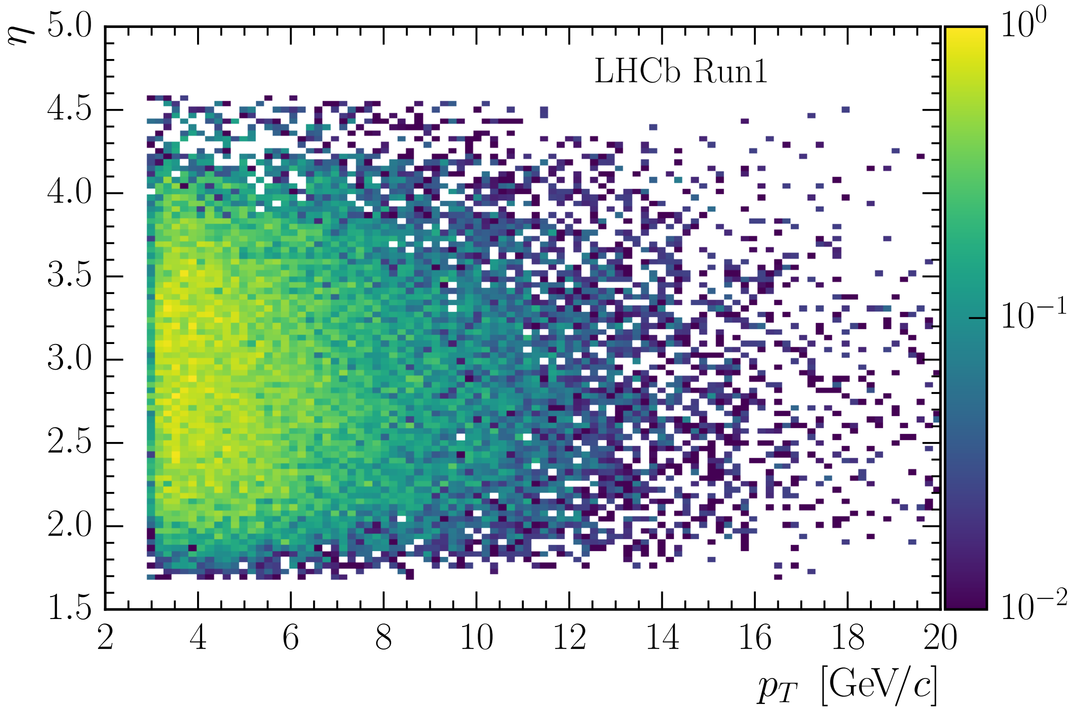


Figure 3.9: Distribution of Run1 high $p_{\rm T}$ photons from the calibration samples in the $(p_{\rm T}, \eta)$ plane. The plot is normalized.

3.5.2 Merged π^0

The reference channel used to produce the calibration samples for merged π^0 is the $D^0 \to K^-\pi^+\pi^0$ mode. The high branching fraction of the $D^0 \to K^-\pi^+\pi^0$ process, namely $(14.2 \pm 0.5)\%$ [7], allows to get a high statistics sample with low background level. The first event selection is performed exploiting the dedicated stripping line `StrippingD02KPiPi0_M` developed in Ref. [60] for both 2011 and 2012 data. The requirements contained in this line are displayed in Tab. 3.3.

The tracks must be of good quality requiring a track fit probability greater than 10^{-6}. They also shall not come from a primary vertex asking their minimum IP χ^2 with respect to any PV to be greater than 16 and are required to have a transverse momentum greater than 300 MeV/c. Finally, tracks must have a ghost probability lower than 0.3 (this variable is defined in Sec. 5.1). The identification of the kaon and pion candidates is made by applying cuts on the particle identification (PID) information provided by the RICH system: $DLL_{K\pi} > 0$ for kaons and $DLL_{\pi K} < 5$ for pions (DLL variables are defined in Sec. 2.4.4). The neutral pions must have a large transverse momentum, namely $p_{\mathrm{T}} > 2000$ MeV/c. The K^- and π^+ tracks are requested to form a good secondary vertex checking the D^0 vertex fit probability. The D^0 candidate is required to point to a PV: its minimal IP χ^2 has to be lower than 9. The Flight Distance (FD) χ^2 must be greater than 64 and the cosine of the direction angle has to be close to 1, which corresponds to require that the primary and secondary vertices are detached and aligned with the reconstructed D^0 momentum. Finally a relatively large mass window for the reconstructed D^0 is used.

In order to improve the signal over background ratio, the $D^0 \to K^-\pi^+\pi^0$ candidates are requested to come from the $D^{*+} \to D^0\pi^+$ decay chain. This is a standard way to clean out a D^0 sample and has for example been used in the study of time-integrated

Variable		Cut
Tracks fit probability		$> 10^{-6}$
Tracks χ^2_{IP}		> 16
Tracks p_{T}	MeV/c	> 300
Tracks GhostProb		< 0.3
$K^{\mp}$ PID		$DLL_{K\pi} > 0$
$\pi^{\pm}$ PID		$DLL_{\pi K} < 5$
$p_T(\pi^0)$	MeV/c	> 2000
D^0 vertex fit probability		$> 10^{-3}$
$\chi^2_{IP}(D^0)$		< 9
$\chi^2_{FD}(D^0)$		> 64
$\cos\theta_{DIRA}(D^0)$		> 0.9999
$m(K^-\pi^+\pi^0)$	MeV/c^2	$[1600; 2100]$

Table 3.3: Cuts implemented in the `StrippingD02KPiPi0_M` stripping line.

CP asymmetries in $D^0 \to K^+K^-$ and $D^0 \to \pi^+\pi^-$ decays [66]. The requirements used to match the D^0 candidates with the D^{*+} and the π^+ (also called soft pion) are displayed in Tab. 3.4.

The soft pion track from the D^{*+} is required to be of good quality and to come from a PV, cutting on the track fit probability and on the minimum IP χ^2 with respect to any PV. It must also have a ghost probability lower than 0.3. The D^{*+} candidate is requested to have a good end vertex and to decay close to the PV, checking the D^{*+} vertex fit probability together with its minimum IP χ^2 and FD χ^2 with respect to the associated PV. Transverse momentum requirements are applied to both the D^0 and the D^{*+} candidates to reduce the combinatorial background. Finally events are combined in a window of ± 10 MeV/c^2 around the $m(D^{*+}) - m(D^0)$ nominal value, namely (145.421 ± 0.010) MeV/c^2 [7].

The $D^0 \to K^-\pi^+\pi^0$ events selected this way are then further filtered with few additional requirements. The window around the expected value of the mass difference $m(D^{*+}) - m(D^0)$ is restricted to ± 2 MeV/c^2, allowing to reject most of the background. The three tracks, the charged kaon and pion from the D^0 decay as well as the soft charged pion from the D^{*+}, are required to not have associated hits in the muon chambers (`isMuon` $= 0$). Finally, we keep only events with a single D^{*+} candidate.

The $D^0 \to K^-\pi^+\pi^0$ selected events are used to perform the mass fit needed to calculate the sWeights. The fit model consists of a double-tail Crystal Ball PDF to describe the signal shape, a first order polynomial for the combinatorial background and an ARGUS function convoluted with a Gaussian PDF to describe the partially reconstructed background associated to the $D^+ \to K^-2\pi^+\pi^0$ decay. In the beginning the idea was to fix the signal tail parameters on MC as done for the $B^0 \to K^{*0}\gamma$ channel. However, the simulation was found to not properly describe the D^0 mass shape and then we opted for a different strategy. We decided to perform a simultaneous fit on the 2011 and 2012 samples building the signal PDF to have tail parameters shared between 2011 and 2012 but independent μ and σ. This way we get a better control on the tails. All the parameters related to the other PDFs of the fit are left free to float separately for 2011 and 2012.

Variable		Cut
Soft π^+ fit probability		$> 10^{-6}$
Soft π^+ χ^2_{IP}		< 16
Soft π^+ GhostProb		< 0.3
D^* vertex fit probability		$> 10^{-2}$
$\chi^2_{IP}(D^{*+})$		< 16
$\chi^2_{FD}(D^{*+})$		< 16
D^{*+} and D^0 p_{T}	MeV/c	> 4000
$\lvert m(D^{*+}) - m(D^0) - 145.421 \rvert$	MeV/c^2	< 10

Table 3.4: Cuts implemented in the selection of $D^{*+} \to D^0(K^-\pi^+\pi^0)\pi^+$ decays.

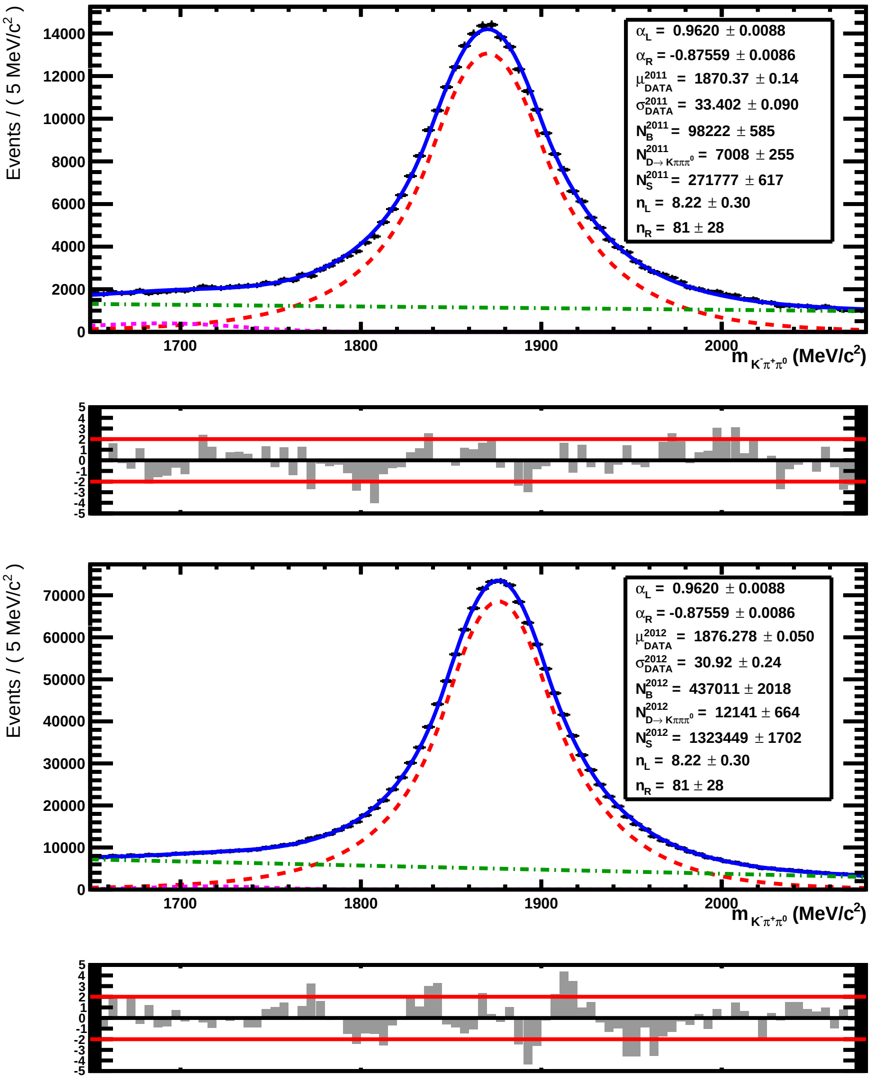


Figure 3.10: $D^0 \to K^-\pi^+\pi^0$ mass fit projections for 2011 (top) and 2012 (bottom) data. The black points represent the data and the solid blue line is the total fit function. The signal is described with a double-tail Crystal Ball PDF sharing the tails parameters between 2011 and 2012 (red dashed line), the combinatorial background with a first order polynomial (dashed green line), and the partially reconstructed $D^+ \to K^-2\pi^+\pi^0$ background (dashed magenta line) with an ARGUS function convoluted with a Gaussian PDF.

The 2011 and 2012 projections of the mass fit to the $D^0 \to K^-\pi^+\pi^0$ spectrum used to compute the sWeights are shown in Fig. 3.10, where the main fit parameters are displayed on the plots. They correspond approximately to 1.6×10^6 signal events over the full Run 1 statistics. The distribution in the (p_{T}, η) plane of the background subtracted merged π^0 from the calibration samples is shown on Fig. 3.11.

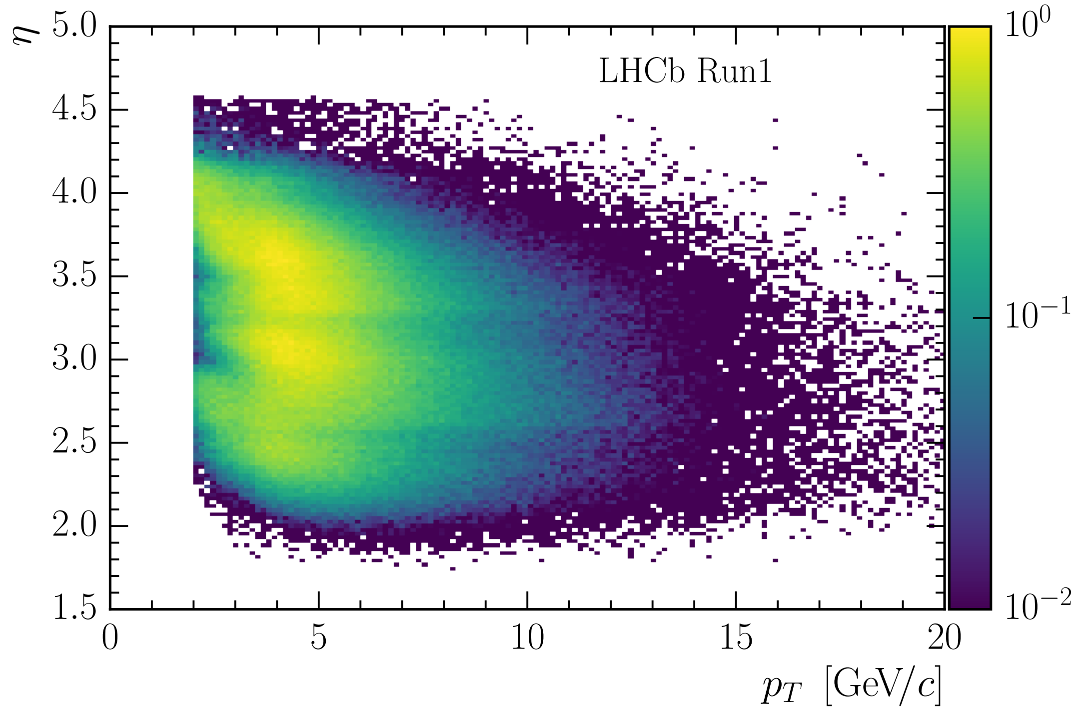


Figure 3.11: Distribution of Run1 merged π^0 from the calibration samples in the (p_T, η) plane. The plot is normalized.

γ/π^0 separation performances

We can now estimate the IsPhoton performances exploiting the Stripping 21 calibration samples. In Fig. 3.12 are shown the γ and π^0 selection efficiencies as a function of the IsPhoton cut for $B^0 \to K^{*0}\gamma$ (red dots) and $B^0 \to K^+\pi^-\pi^0$ (blue triangles) simulated events. These MC samples, that are also used in the $B^0 \to \rho^0\gamma$ analysis, have been exploited to extract the IsPhoton efficiency computed with the `GammaPi0SeparationCalib` tool [64] and using the 2012 binning configuration in Tab. 5.11. MC events are selected according to the trigger conditions of Tab. 5.4 and to the pre-selection cuts of Tab. 5.7. Looking at Fig. 3.12, we can notice that for IsPhoton > 0.6, which is the standard cut applied so far in all radiative analyses, photons are selected with an efficiency of about 94% while π^0 are selected approximately with a 44% efficiency. This corresponds to a 56% rejection of the merged π^0 background.

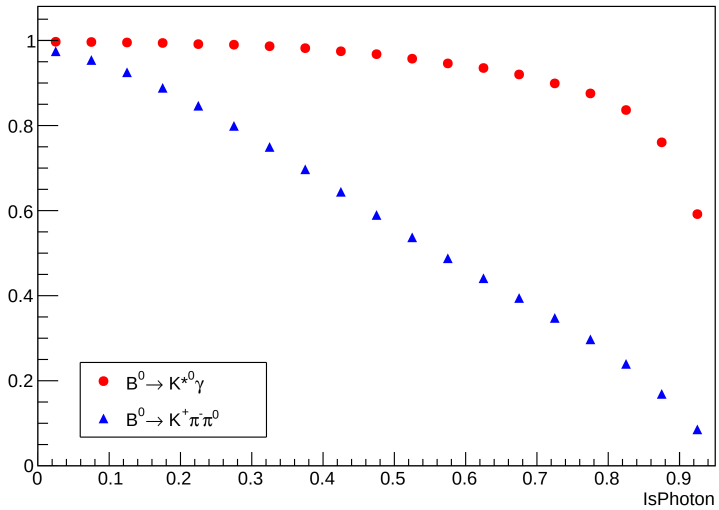


Figure 3.12: Photon and π^0 selection efficiencies as a function of the IsPhoton cut for $B^0 \to K^{*0}\gamma$ (red dots) and $B^0 \to K^+\pi^-\pi^0$ (blue triangles) MC samples. The experimental points correspond to the average IsPhoton efficiencies calculated with the γ/π^0 separation tool.

3.5.3 Resolved π^0

Calibration samples for resolved π^0 have also been produced using the $D^0 \to K^-\pi^+\pi^0$ mode. These samples are not useful to extract the γ/π^0 separation efficiency, but could be exploited as soft photons calibration samples. The first event selection is performed exploiting the dedicated stripping line `StrippingD02KPiPi0_R` developed in [60] for both 2011 and 2012 data. The cuts contained in this line are those displayed in Tab. 3.3, except for the p_T of the tracks that here is required to be greater than 600 MeV/c and the p_T of the π^0 that has to be greater than 1000 MeV/c. In addition a minimal cut, IsNotH > 0.2, is imposed to the two photons used to build the resolved π^0. The stripped events are then matched to the $D^{*+} \to D^0\pi^+$ decay and further filtered the very same way as for merged π^0 events.

Selected events are used to perform the mass fit and calculate the sWeights. The fit model consists of a double-tail Crystal Ball PDF to describe the signal shape, a second order polynomial for the combinatorial background and a Gaussian PDF to describe the $D^0 \to K^-\pi^+$ background in the right side of the mass spectrum. As done for merged π^0 the tail parameters of the signal PDF are shared between the 2011 and 2012 samples, while all the other fit components are left free to float independently for 2011 and 2012.

The 2011 and 2012 projections of the mass fit to the $D^0 \to K^-\pi^+\pi^0$ spectrum used to compute the sWeights are shown in Fig. 3.13, where the main fit parameters are displayed on the plots. They correspond approximately to 2.6×10^6 signal events over the full Run 1 statistics. The distribution in the (p_T, η) plane of the background subtracted resolved π^0 from the calibration samples is shown on Fig. 3.14.

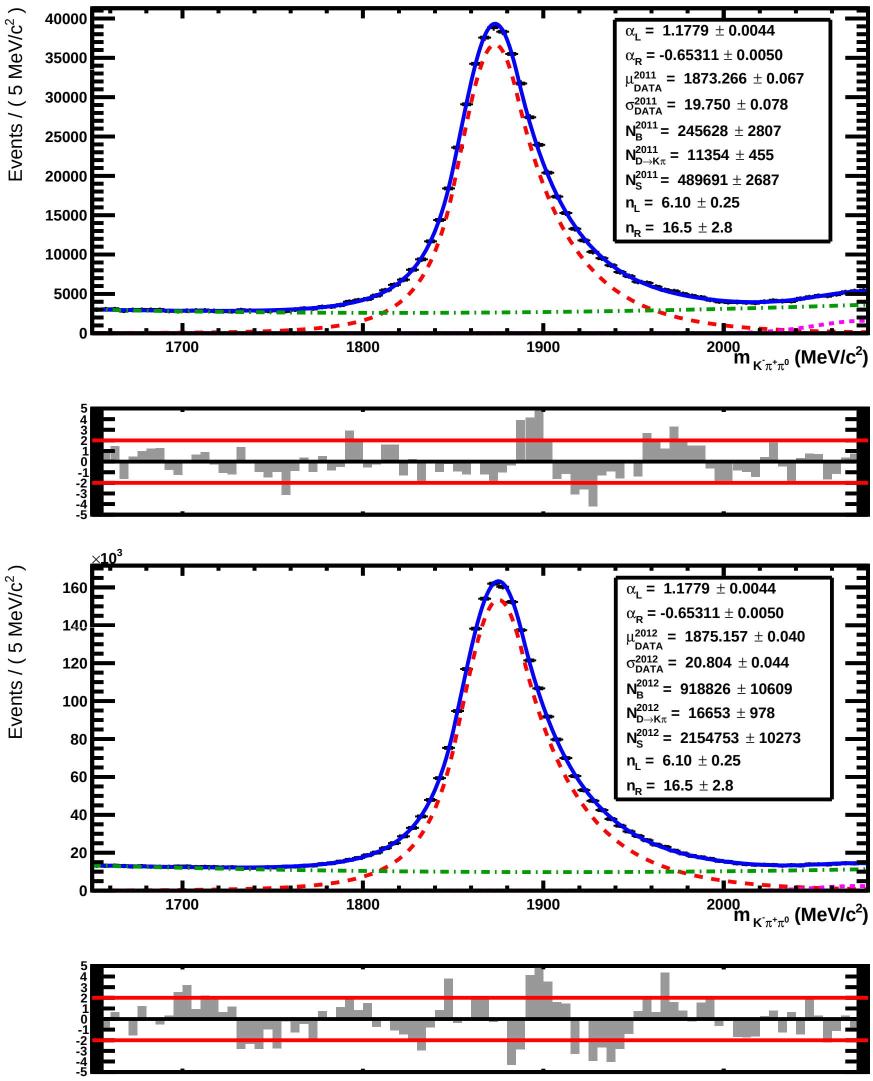


Figure 3.13: $D^0 \to K^-\pi^+\pi^0$ mass fit projections for 2011 (top) and 2012 (bottom) data. The black points represent the data and the solid blue line is the total fit function. The signal is described with a double-tail Crystal Ball PDF sharing the tails parameters between 2011 and 2012 (red dashed line), the combinatorial background with a second order polynomial (dashed green line), and the $D^0 \to K^-\pi^+$ background (dashed magenta line) with a Gaussian PDF.

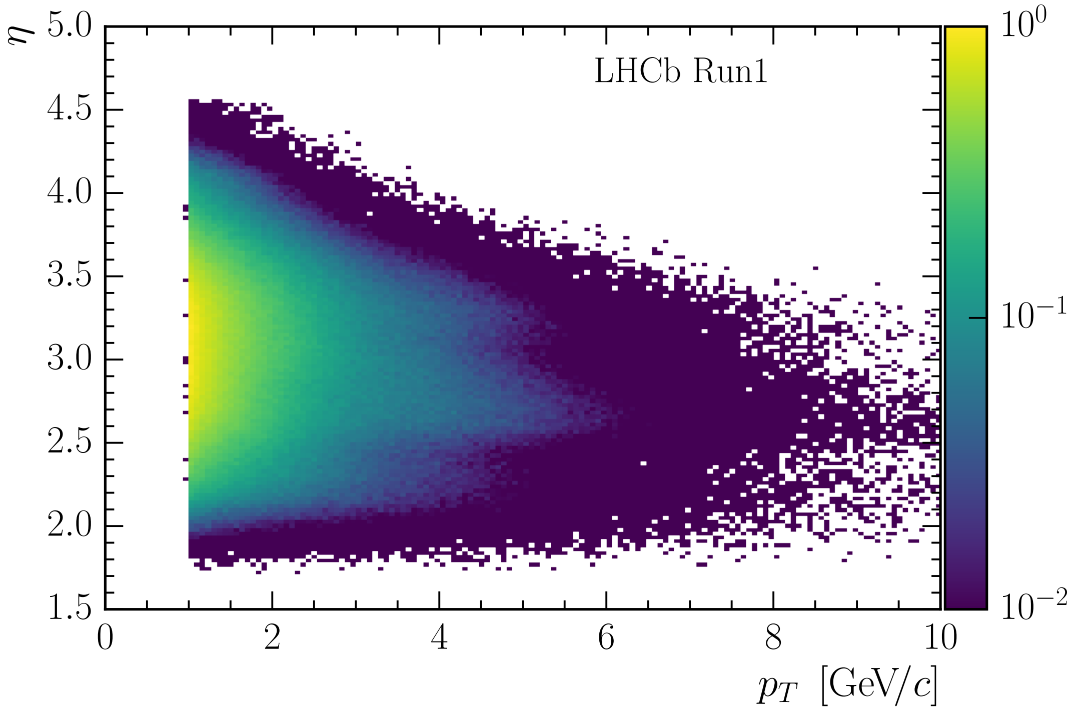


Figure 3.14: Distribution of Run1 resolved π^0 from the calibration samples in the (p_{T}, η) plane. The plot is normalized.

Chapter 4

Introduction to the measurement

The radiative $B^0 \to \rho^0\gamma$ decay is a rare decay which corresponds at quark level to a $b \to d\gamma$ transition occurring via Flavour Changing Neutral Currents. This transition appears in the SM through electroweak penguin diagrams where the main contribution to the loop is represented by the propagation of a top quark (Fig. 4.1). $b \to d\gamma$ transitions lead to very rare decays in the SM and those are quite sensitive to many extensions of the SM [27].

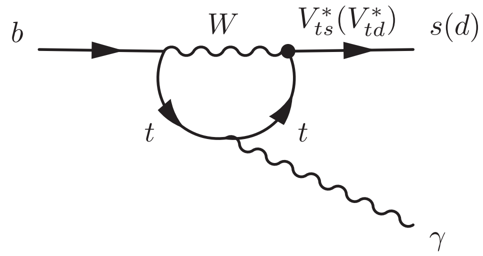


Figure 4.1: Feynman diagram for $b \to q\gamma$ transitions in the SM ($q = d, s$).

The $B^0 \to \rho^0\gamma$ decay has already been observed at BaBar [26] and Belle [67] experiments, but not yet at LHCb. The measured branching fraction reported on the PDG is $\mathcal{B}\ (B^0 \to \rho^0\gamma) = (8.6 \pm 1.5) \times 10^{-7}$ [7]. Studying this decay provides a way to extract the modulus of the V_{td} element of the CKM matrix, thus allowing to constrain the unitarity triangle. The objective of this study is to measure the ratio of branching fractions

$$\frac{\mathcal{B}(B^0 \to \rho^0\gamma)}{\mathcal{B}(B^0 \to K^{*0}\gamma)} \propto \left|\frac{V_{td}}{V_{ts}}\right|^2 \tag{4.1}$$

where $B^0 \to K^{*0}\gamma$, which is the reference channel for B radiative decays, is used as control mode in the analysis. This measurement allows to extract the ratio $|V_{td}/V_{ts}|$, which is related to the length of the right side of the unitarity triangle, as a very interesting test of the SM: in particular we can compare the result with the ones arising from the measurement of the oscillation frequencies of the $B^0 - \overline{B}^0$ [68] and $B_s^0 - \overline{B}_s^0$ [69] systems.

4.1 Analysis strategy

Experimentally we look for the following final state

$$B^0 \to \rho^0\gamma \to (\pi^+\pi^-)\gamma\ .$$

This is something very challenging at LHCb basically for two reasons: first of all we are dealing with a particularly rare decay which is intrinsically difficult to observe and, secondly, we need to efficiently distinguish between high energy π^0 and γ. Around 88% of the photons detected by the calorimeter are originated from a π^0 decay. When the two photons of the $\pi^0 \to \gamma\gamma$ decay are reconstructed as a single cluster in the ECAL and they cannot be resolved due to the detector granularity, this is known as a *merged* π^0 [62]. These pions, that typically have a transverse momentum $p_T(\pi^0) > 2500$ MeV/c, are then particularly dangerous since the π^0 can be misidentified as a radiative photon, which is the signature of our decay. Decays with merged π^0 are the dominant backgrounds of radiative channels.

In order to perform the analysis we have to set up a thorough selection to reduce and control the potentially very large background contaminations. These backgrounds can be divided in three categories:

- Peaking backgrounds
- Partially reconstructed backgrounds
- Combinatorial background

The peaking backgrounds, or cross-feeds, are other decays peaking under the signal peak and can be of two types: either $h^+h^{'-}\gamma$ decays for which at least one track is misidentified as a pion such as $B^0 \to K^{*0}(K^+\pi^-)\gamma$ or $\Lambda_b^0 \to \Lambda^{*0}(pK^-)\gamma$, or $h^+h^{'-}\pi^0$ decays for which a high p_T π^0 is misidentified as a photon. They are the most dangerous backgrounds since they cannot be distinguished from signal using the reconstructed B mass. It is then crucial to carefully describe their shapes in the mass fit and to precisely evaluate their actual contaminations to the invariant mass spectrum. The cross-feeds contamination can be reduced applying tight charged and neutral particle identification (PID) requirements.

Partially reconstructed backgrounds are decays with at least one additional particle (such as π^0, $\pi^\pm$, K or γ) in the final state with respect to the signal. The contributions coming from these decays, which are dominant in the left side of the reconstructed B mass spectrum, can be reduced using isolation variables.

The combinatorial background is associated to random combinations of reconstructed particles that pass the signal selection. These events have significantly different topology and kinematics compared to the signal and can be reduced using multivariate analysis techniques.

In order to avoid introducing any bias, the analysis is conducted blindly and exploiting $B^0 \to K^{*0}\gamma$ as control channel to fix the working point and tune all the selection requirements. In particular the goal is to sequentially fix all the neutral and charged PID requirements before developing and optimizing the multivariate classifier (MVA). This way we aim to select the most relevant sample of combinatorial background to fight against in the MVA. The $B^0 \to K^{*0}\gamma$ control channel has a branching ratio about 50 times larger than the $B^0 \to \rho^0\gamma$ mode. It is also the dominant source of peaking background.

4.2 Data and Monte Carlo samples

This analysis is performed using the full Run I data collected by the LHCb experiment at a center-of-mass energy of $\sqrt{s} = 7$ TeV for the 2011 campaign and $\sqrt{s} = 8$ TeV for the 2012 campaign. These samples correspond respectively to integrated luminosities of $\int \mathcal{L}\mathrm{d}t = 1\,\mathrm{fb}^{-1}$ and $\int \mathcal{L}\mathrm{d}t = 2\,\mathrm{fb}^{-1}$. The events are reconstructed using the `Reco14` version of the reconstruction software for both years campaign. The data are stripped using `Stripping21r1` for 2011 and `Stripping21` for 2012 and exploiting the inclusive stripping line `StrippingB2XGamma2pi_Line` for both years (see Sec. 5.1). Monte Carlo (MC) generated samples are produced using Gauss with different simulation (Sim) versions and configurations. They are needed to study the behaviour of the signal and background events (peaking backgrounds and partially reconstructed backgrounds) and to extract the shapes to be included in the mass fit model. The number of MC events produced for each decay mode and year is summarized in Table 4.1. The Sim versions are also reported. All the MC samples used in the analysis were reconstructed under both the $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ hypotheses. This means that when our signal is the $B^0 \to \rho^0\gamma$ mode all the MC samples we mention were reconstructed as $\rho\gamma$ and when our signal is $B^0 \to K^{*0}\gamma$ all the MC samples were reconstructed as $K^*\gamma$.

Data and MC ntuples were produced using DaVinci v36r5 and by selecting the intermediate vector meson candidate in a wide mass range of ± 500 MeV/c^2 around the ρ^0 PDG mass for the $B^0 \to \rho^0\gamma$ hypothesis and ± 300 MeV/c^2 around the K^{*0} PDG mass for the $B^0 \to K^{*0}\gamma$ hypothesis.

Decay mode	Event type	Year	Sim version	Generated events
$B^0 \to \rho^0\gamma$	11102222	2011 (2012)	Sim08i (Sim08a)	1583901 (2918486)
$B^0 \to K^{*0}\gamma$	11102202	2011 (2012)	Sim08f (Sim08d)	9283600 (6053226)
$B_s^0 \to \phi\gamma$	13102202	2011 (2012)	Sim08f (Sim08d)	4045112 (6075698)
$\Lambda_b^0 \to \Lambda^*(1520)^0\gamma$	15102203	2011 (2012)	Sim08h (Sim08c)	2044218 (2734988)
$B^0 \to \pi^+\pi^-\pi^0$	11102400	2012	Sim08a	2017990
$B^0 \to K^+\pi^-\pi^0$	11202401	2012	Sim08a	1521986
$B^0 \to K^{*0}(K_S^0\pi^0)\rho^0$	11100001	2012	Sim08g	4053754
$B^0 \to K^{*0}\eta(\gamma\gamma)$	11102441	2011 (2012)	Sim08e (Sim08e)	3049210 (5608620)
$B^0 \to \omega\gamma$	11102600	2012	Sim06b	2012979
$B^0 \to K^{*0}\eta(\pi^+\pi^-\pi^0)$	11104401	2011 (2012)	Sim08e (Sim08e)	3110923 (6087634)
$B^0 \to D^-(K^+\pi^-\pi^-)\rho^+$	11164406	2012	Sim08d	1020999
$B^0 \to K_1(1270)^0\gamma$	11202602	2012	Sim08c	1520242
$B^+ \to K^*(1410)^+\gamma$	12103221	2012	Sim06b	2027491
$B^+ \to K^+\pi^-\pi^+\gamma$	12103240	2012	Sim08h	968560
$B^+ \to \rho^+\rho^0$	12103401	2011 (2012)	Sim08e (Sim08e)	2030917 (3227630)
$B^+ \to K^{*+}\pi^+\pi^-$	12103421	2011 (2012)	Sim08e (Sim08e)	1044087 (2011561)
$B^+ \to a_1(1260)^0(\pi^+\pi^-\pi^0)K^+$	12103490	2011 (2012)	Sim08i (Sim08i)	491322 (1092340)
$B^+ \to D^0(K^-\pi^+\pi^0)\rho^+$	12163472	2012	Sim08e	2118714
$B^+ \to D^0(K^-\pi^+)\rho^+$	12163473	2011 (2012)	Sim08h (Sim08h)	2065779 (4056653)
$B^+ \to D^0(K_S^0\pi^+\pi^-)\rho^+$	12165511	2012	Sim08a	1518492
$B^+ \to K^{*0}\pi^+\gamma$	12203203	2012	Sim08c	1523248
$B^+ \to K_2^*(1430)^+\gamma$	12203212	2012	Sim08a	1030247
$B^+ \to K_1(1270)^+\gamma$	12203224	2012	Sim08a	3080237
$B^+ \to K_1(1400)^+\gamma$	12203225	2012	Sim08a	1035245
$B^+ \to a_1(1270)^+(\pi^+\pi^-\pi^+)\gamma$	12203226	2012	Sim08c	3076489
$B^+ \to K^{*+}(K_S^0\pi^+)\gamma$	12203303	2012	Sim08e	7937039
$B^+ \to K^+\pi^-\pi^+\pi^0$	12203401	2012	Sim08c	1516496
$B^+ \to K_1(1270)^+\eta(\gamma\gamma)$	12203410	2012	Sim08c	1521743
$\Lambda_b^0 \to \Lambda^*(1670)^0\gamma$	15102228	2011 (2012)	Sim08h (Sim08h)	2026725 (2085195)
$\Lambda_b^0 \to \Lambda^*(1820)^0\gamma$	15102230	2011 (2012)	Sim08h (Sim08g)	2008328 (3056353)
$\Lambda_b^0 \to \Lambda^*(1830)^0\gamma$	15102240	2011 (2012)	Sim08h (Sim08g)	2036065 (3045037)

Table 4.1: Simulated samples used in the analysis.

Chapter 5

Event reconstruction and selection optimization

5.1 Stripping

After being collected and stored, raw data are reconstructed and a loose preselection is applied in order to minimize the computing time and the required storage space. In this preselection step, called stripping, the candidates to be used in the analyses are created filtering the raw events with a combination of different inclusive and exclusive lines that contain a set of dedicated cuts fixed according to the physics targets. This way the candidates can be accessed and handled in a faster and more efficient way.

As already mentioned, the stripping line used in this analysis is the `StrippingB2XGamma2pi_Line` inclusive line. It allows to select a generic $B \to V(h^+h'^-)\gamma$ candidate with an intermediate vector resonance V, decaying into two charged hadrons, and a high transverse momentum photon in the final state. When producing the ntuples, events can be reprocessed adding further requirements, like for instance the V mass window and the type of charged tracks, in order to build up the decay of interest.

The B candidate is selected in a large invariant mass range [3280; 9000] MeV/c^2. It is also required to have a vertex χ^2 per degree of freedom lower than 9, an impact parameter (IP) χ^2 lower than 9 and a positive direction angle θ_{DIRA}. The precise definition of these variables is given in Sec. 5.6. The sum of the transverse momenta of the daughter particles (h^+, h'^- and γ) is required to be greater than 5000 MeV/c.

The two charged tracks must have both $p_T > 300$ MeV/c and $p > 1000$ MeV/c. The sum of their transverse momenta has to be greater than 1500 MeV/c. The minimum of the IP χ^2 with respect to any primary vertex (PV) shall be greater than 16 and the χ^2 per degree of freedom of the track fit shall be below 3. Tracks must also have a ghost probability (probability for this track to be reconstructed from hits in the tracking system that belong to two separated tracks) lower than 0.4.

The photon is required to have a transverse momentum greater than 2000 MeV/c and a γ_{CL} (see Sec. 5.3) greater than 0. The stripping cuts are reported in Tab. 5.1.

Variable		Cut
$m(B)$	MeV/c^2	$[3280; 9000]$
Daughters $\sum p_T(B)$	MeV/c	> 5000
$\theta_{DIRA}(B)$		> 0
$\chi^2_{\rm vtx}/{\rm ndf}(B)$		< 9
$\chi^2_{\rm IP}(B)$		< 9
Track p_T	MeV/c	> 300
Track p	MeV/c	> 1000
Tracks $\sum p_T$	MeV/c	> 1500
Track $\chi^2/{\rm ndf}$		< 3
Track $\chi^2_{\rm IP}$		> 16
Tracks GhostProb		< 0.4
$p_T(\gamma)$	MeV/c	> 2000
γ_{CL}		> 0

Table 5.1: Selection requirements applied in the `StrippingB2XGamma2pi_Line` stripping line to select $B \to V\gamma$ events.

5.2 Radiative trigger

For radiative decays, the trigger relies in part on the photon which is required to have large E_T in order to reduce the large combinatorics coming from pp collisions.

At the L0 trigger level, this is achieved by requiring a high E_T electromagnetic cluster, with thresholds between 2.5 and 2.96 GeV, and an SPD multiplicity below 600 [70]. As around $\sim 40\%$ of the photons reconstructed as a single cluster in the LHCb calorimeter are converted into electrons before the SPD detector, both the `L0Photon` and `L0Electron` lines are included in the trigger selection. Additionally, a version of these L0 lines with a higher threshold of $E_{\rm T} > 4.2$ GeV, `L0PhotonHi` and `L0ElectronHi`, is used as input for the track photon HLT1 line.

The `Hlt1TrackAllL0` line (see Tab. 5.2) imposes standard track requirements. In radiative decays, the efficiency is improved by an additional $\sim 20\%$ by also using the `Hlt1TrackPhoton` line which imposes looser track requirements for events passing a higher photon E_T threshold in the L0 trigger. Events are required to have triggered any L0 line for the `Hlt1TrackAllL0` line while the `Hlt1TrackPhoton` line is limited to events that have passed the `L0PhotonHi` or `L0ElectronHi` lines. For the sake of brevity, the meaning of every cut entering Tab. 5.2 is not explained. However let us mention that offline, in order to align the various trigger thresholds, we require that at least one of the two tracks from the vector meson candidate have a p_T greater than 1800 MeV/c and that the photon candidate have an E_T greater than 3000 MeV.

		Hlt1TrackAllL0	Hlt1TrackPhoton
L0 channel		L0 physics	L0PhotonHi or L0ElectronHi
VELO track hits		> 9	
VELO missed hits		< 3	< 4
VELO track IP	µm	$--$	(> 100)
Track p	GeV/c	> 10 (3)	> 6 (3)
Track p_{T}	GeV/c	> 1.7 (1.6)	> 1.2
Track χ^2/ndf		< 2	
Track χ^2_{IP}		> 16	

Table 5.2: Selection requirements applied in 2011 (2012) on the HLT1 lines relevant to radiative decays (more detailed values can be found in Ref. [70]). Missed VELO track hits refer to the difference between the number of hits assigned to the VELO track and the number of hits expected given the track direction and the first measured point on the track.

The HLT2 selection exploits the inclusive radiative topological lines `Hlt2RadiativeTopoTrack(TOS)` and `Hlt2RadiativeTopoPhoton(L0)` [71]. These lines are designed to efficiently trigger on any B decay with at least 2 tracks and one high p_T photon. The B candidates are selected using the criteria of Tab. 5.3. In this table the first and third sections give the cuts on the input tracks and photons respectively. The second corresponds to the combination of the two tracks. At least one of the two tracks is required to have a track fit χ^2 per degree of freedom lower than 3. The 2 tracks should have a Distance Of Closest Approach (DOCA) lower than 0.15 mm and form a good quality vertex. The tracks four-momenta, under the pion hypothesis, are added and the combination is required to have an invariant mass below 2000 MeV/c^2 and a transverse momentum above 1500 MeV/c. The cuts on the B candidate are reported in the fourth section of Tab. 5.3. The B decay vertex should be separated from the associated PV requiring a minimal flight distance (see Sec. 5.6) χ^2 of 64. The last section of the table is a Global Event Cut (GEC). Since these inclusive lines were introduced later during the 2011 data taking, the actual integrated luminosity available for this year is 0.7 fb^{-1}.

The full trigger sequence used in the analysis is displayed in Tab. 5.4. All the lines are required to be “Triggered On Signal” (TOS): the two tracks and the photon from the reconstructed B decay are sufficient to fulfill the trigger line requirement.

Variable		Cut
Track p_T	MeV/c	> 700
Track p	MeV/c	> 5000
Track χ^2_{IP}		> 10
Track χ^2/ndf		< 5
Min track χ^2/ndf		< 3
2-track DOCA	mm	< 0.15
2-track χ^2_{vtx}/ndf		< 10
2-track m	MeV/c^2	< 2000
2-track p_T	MeV/c	> 1500
$E_T(\gamma)$	MeV	> 2500
$p_T(B)$	MeV/c	> 1000
$m(B)$	MeV/c^2	[4000; 7000]
$\chi^2_{FD}(B)$		> 64
Daughters $\sum p_T(B)$	MeV/c	> 5000
# of forward tracks with $p_T > 500$ MeV/c		< 120

Table 5.3: Selection criteria for the radiative topological lines.

Trigger level		Standard path	High photon E_{T} path
L0		`L0Photon` ‖ `L0Electron`	`L0PhotonHi` ‖ `L0ElectronHi`
HLT1		`Hlt1TrackAllL0`	`Hlt1TrackPhoton`
HLT2	(2011)	`Hlt2RadiativeTopoTrackTOS` ‖	`Hlt2RadiativeTopoPhotonL0`
	(2012)	`Hlt2RadiativeTopoTrack` ‖	`Hlt2RadiativeTopoPhoton`

Table 5.4: Summary of the trigger conditions used throughout the analysis.

5.3 Cut based selection and γ_{CL} optimization

The first step in the selection optimization process is tuning the requirement on the photon identification, namely fixing the cut on the γ_{CL} variable. This variable, also called IsNotH, is aimed to distinguish the γ cluster from non-electromagnetic deposits associated to hadrons. IsNotH is a Neural Network variable [61] that combines cluster information from the PreShower and ECAL detectors in a set of shape parameters together with the multiplicity in the SPD and the ratio $E_{\text{HCAL}}/E_{\text{ECAL}}$. The last quantity is the ratio of the energy in the hadronic calorimeter, in the projected area matching the cluster, and the cluster energy in the electromagnetic calorimeter.

Since a dedicated calibration tool to extract the efficiency for a given γ_{CL} cut is not available, we adopt a straight cut approach exploiting the $B^0 \to K^{*0}\gamma$ control channel. We perform a series of fits to the B mass spectrum in data varying the IsNotH requirement and evaluate the corresponding efficiencies. To select the events to be used in the optimization we apply the set of cuts listed in Tab. 5.5. This selection is close to those used to perform the LHCb measurements of the ratio of branching ratios $\mathcal{B}(B^0 \to K^{*0}\gamma)/\mathcal{B}(B_s^0 \to \phi\gamma)$ [57] and of the photon polarization in $B_s^0 \to \phi\gamma$ decay [72].

Variable		Cut
Track p	MeV/c	$[3000; 100000]$
Track p_T	MeV/c	> 500
Track η		$[1.5; 5]$
Track χ^2_{IP}		> 55
$K^{\pm}$PID		ProbNNk > 0.2
$\pi^{\mp}$PID		ProbNNpi > 0.2 & ProbNNk < 0.2
$m(K^{*0})$	MeV/c^2	$[m_{K^{*0}}^{PDG} - 100; m_{K^{*0}}^{PDG} + 100]$
$\chi^2_{\text{vtx}}(K^{*0})$		< 9
$\lvert\cos\theta_H\rvert$		< 0.8
γ_{CL}		> 0
IsPhoton		> 0.6
$m(B)$	MeV/c^2	$[4000; 7000]$
$p_T(B)$	MeV/c	> 3000
$\chi^2_{\text{IP}}(B)$		< 9
$\chi^2_{FD}(B)$		> 100
Smallest $\Delta\chi^2_{\text{vtx}}(B)$		> 2

Table 5.5: Cut based selection applied to tune the γ_{CL} cut. The trigger requirements of Tab. 5.4 as well as the stripping selection of Tab. 5.1 are applied.

The two tracks used to build the K^{*0} meson are required to both have $p_T > 500$ MeV/c, p in the range $[3000; 100000]$ MeV/c and pseudorapidity within $[1.5; 5]$. They must also point away from the pp interaction vertices by requiring the IP χ^2 to be greater than 55. The identification of the kaon and pion tracks is made by applying cuts on the particle identification (PID) information provided by the RICH system: ProbNNk > 0.2 for the kaon, and ProbNNpi > 0.2 together with ProbNNk < 0.2 for the pion (ProbNN variables are defined in Sec 5.4). The ranges used for the tracks momenta and pseudorapidities correspond to the phase space covered by the PID calibration samples (see Tab. 5.6).

The K^{*0} meson is selected in a mass window of ± 100 MeV/c^2 around the PDG mass ($m_{K^{*0}}^{PDG} = 895.55 \pm 0.20$ MeV/c^2 [7]) and is required to have a vertex with χ^2 lower than 9. The distribution of the helicity angle θ_H, defined as the angle between the momentum of the positively charged daughter of the vector meson V and the momentum of the B candidate in the rest frame of the vector meson, follows a $\sin^2\theta_H$ function for $B \to V\gamma$ modes and a $\cos^2\theta_H$ function for $B \to V\pi^0$ backgrounds. A requirement of $|\cos\theta_H| < 0.8$ is therefore made to reduce backgrounds from $B \to V\pi^0$ decays in which the neutral pion is misidentified as a photon [57].

The B candidate is selected in the mass range $[4000; 7000]$ MeV/c^2 and required to have a transverse momentum greater than 3000 MeV/c, an IP χ^2 lower than 9 and a the flight distance χ^2 greater than 100. In addition a requirement on the isolation of the reconstructed B vertex is applied by imposing smallest $\Delta\chi^2_{\text{vtx}}(B) > 2$. The precise definition of this variable is given in Sec. 5.6.

Two cuts are applied on the photon identification: $\gamma_{CL} > 0$ as already imposed at the stripping level and IsPhoton > 0.6. The IsPhoton variable is dedicated to the discrimination between clusters from photons and clusters from neutral pions. It is discussed in details in Sec. 5.5. Here an additional comment is needed since in principle γ_{CL} and IsPhoton are correlated variables. This correlation, however, starts to appear only when applying very tight cuts (greater than 0.8) on both variables at the same time: this corresponds to a situation in which one selects almost unambiguously a real photon. Besides, not applying any cut on the IsPhoton variable would require an extra care to take into account backgrounds from $B \to V\pi^0$ decays. For this reason we require IsPhoton > 0.6, which is the standard cut applied so far in all radiative analyses. The event selection shown in Tab. 5.5 will be reused throughout the analysis to perform various crosschecks.

In order to optimize the γ_{CL} cut we exploit the standard Figure of Merit (FoM)

$$\text{FoM}\ (\gamma_{CL}) = \frac{S}{\sqrt{S+B}} \tag{5.1}$$

where S and B are respectively the number of signal and background events extrapolated in the $\pm 3\sigma$ signal region defined as [5000; 5580] MeV/c^2. The various mass fits are performed varying the γ_{CL} cut by steps of 0.05 in the range [0; 0.7] fixing the μ and the σ of the signal PDF to the values obtained for $\gamma_{CL} > 0$. The fit model consists of a double-tail Crystal Ball Probability Density Function (PDF) to describe the signal shape, an exponential function for the combinatorial background and two ARGUS functions convoluted with

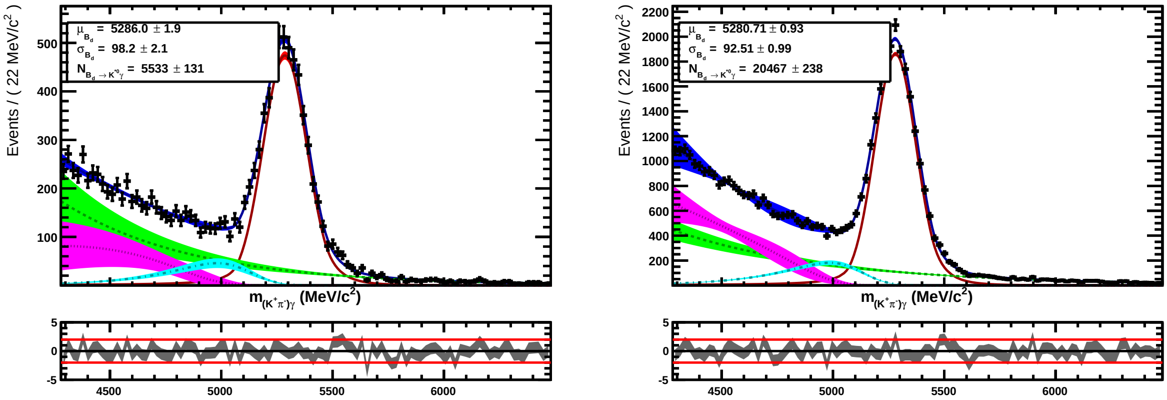


Figure 5.1: $B^0 \to K^{*0}\gamma$ mass fits for 2011 (left) and 2012 (right) data as selected according to the cut-based approach and with $\gamma_{CL} > 0$.

a Gaussian PDF to describe the partially reconstructed backgrounds (see Chap. 6 for more details about the PDFs definition). For these last ones we considered two generic contributions associated respectively to one and two missing π in the final state. No cross-feeds contributions are included in the fit and the shapes of the signal and of the first ARGUS PDFs are fixed on MC. The parameters of the other PDFs are left free to float. This is clearly a simplified fit model but a complete description of all the background sources is not needed for our purpose. The 2011 and 2012 mass fits of the $B^0 \to K^{*0}\gamma$ mass spectrum for $\gamma_{CL} > 0$ are shown in Fig. 5.1. The efficiencies are calculated as

$$\epsilon^{DATA} = \frac{S(\gamma_{CL} > ...)}{S(\gamma_{CL} > 0)} \tag{5.2}$$

where $S(\gamma_{CL} > ...)$ is the fitted signal yield for a given γ_{CL} cut.

The 2011 and 2012 FoMs and efficiency plots are shown in Fig. 5.2. The efficiencies obtained on MC are also reported. For low values of the γ_{CL} cut, data and MC efficiencies are in very good agreement, and the FoMs vary only slightly. Aiming to benefit from the suitable data/MC agreement, we take as reference for the analysis $\gamma_{CL} > 0.35$ which corresponds approximately to a 96% selection efficiency for both 2011 and 2012 data as well as MC. Beyond the loss in term of FoM becomes more significant and the spread between data and MC efficiencies increases.

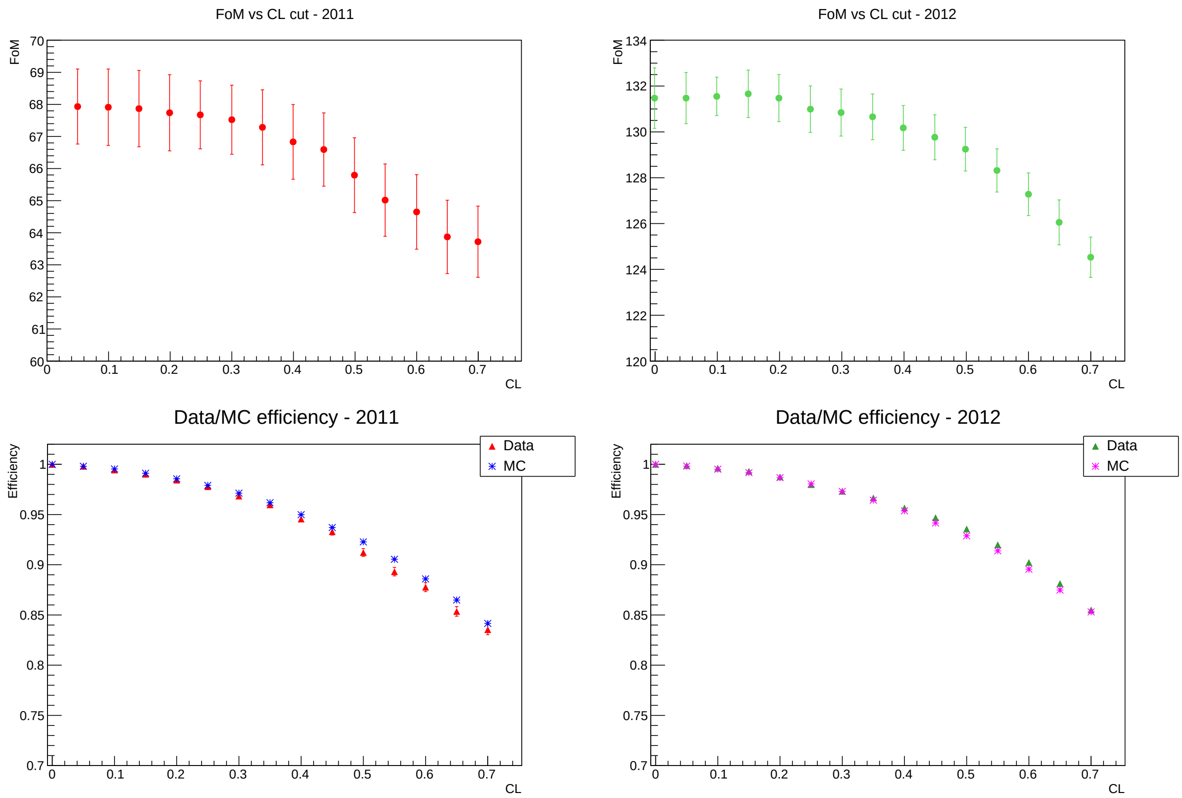


Figure 5.2: $B^0 \to K^{*0}\gamma$ figures of merit (top) as a function of the γ_{CL} cut and corresponding efficiency plots (bottom) for 2011 (left) and 2012 (right).

5.4 Charged PID optimization

In order to separate between different tracks and to get the best possible background rejection, we use the dedicated ProbNN PID variables. These variables are the result of the training of Neural Networks putting together the information coming from the various PID sub-detectors (essentially the RICHes for the $p/K/\pi$ separation). For a given track, values associated to the various PID hypotheses are given (ProbNNk for K, ProbNNpi for π and ProbNNp for p). ProbNN variables can be combined in different ways according to the PID approach used in the analysis.

Aiming to reduce the cross-feeds contamination and having mutually exclusive events in each spectrum we opted for a bi-dimensional cut in the (ProbNNpi-ProbNNk, ProbNNp) plane (Fig. 5.3), where the ProbNNp variable is added in the picture in order to prevent contaminations coming from baryonic modes such as $\Lambda_b^0 \to \Lambda^{*0}\gamma$. The underlying idea is to define optimal and independent regions where to select the charged pions and kaons. The goal of this approach is to improve the performances of the simultaneous fit to the $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ mass spectra. At the beginning of the optimization process, the actual contamination coming from baryonic modes was unknown. For this reason, in case of high contamination, we considered the possibility to define a region dedicated to protons in order to include $\Lambda_b^0 \to \Lambda^{*0}\gamma$ as a third channel in the simultaneous fit. However, as it will be shown in the following, the contributions of these modes to the $\rho\gamma$ and $K^*\gamma$ spectra are small and a region dedicated to protons is not needed. The complete set of cuts to select pions and kaons in the (ProbNNpi-ProbNNk, ProbNNp) plane is given by

$$\pi : \text{ProbNNpi-ProbNNk} > X \quad \text{and} \quad \text{ProbNNp} < Y \tag{5.3}$$

$$K : \text{ProbNNpi-ProbNNk} < X' \quad \text{and} \quad \text{ProbNNp} < Y \tag{5.4}$$

where X, X' and Y are the values that we want to optimize.

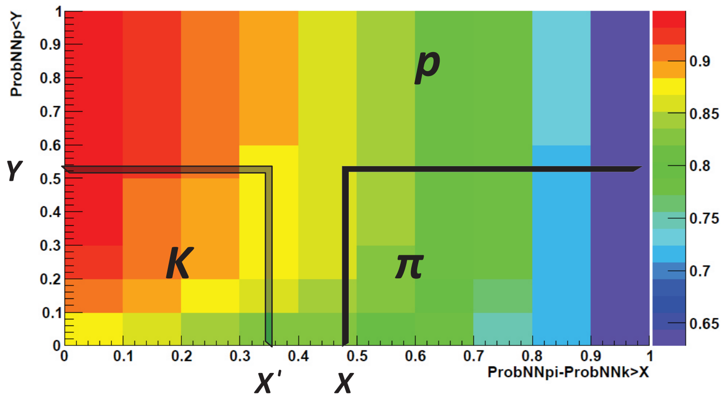


Figure 5.3: Illustrative plot of the PID strategy in the (ProbNNpi-ProbNNk, ProbNNp) plane. Colors are related to the overall PID efficiency of a given combination of cuts, in this case for $B^0 \to \rho^0\gamma$ events, the cuts being applied to both pions.

However the simulation does not reproduce such variables with sufficient accuracy and we cannot rely on MC to extract the efficiency for a given cut. Therefore we use a data-driven approach using dedicated high statistics calibration samples that can be accessed thanks to the `PIDCalib` tool [73] available in the LHCb software. These samples, that contain millions of events of high-purity pion, kaon and proton candidates, are produced exploiting the channels displayed in Tab. 5.6. In this analysis we used the PIDCalib version integrated in the v5r0 release of the Urania package. For a given set of cut we use PIDCalib to generate efficiency maps for the individual tracks (π, K and p) and then we combine them to build up the candidates for which we want to know the real PID efficiency. In this last step we extract the efficiencies by applying the generated maps to the MC simulated events containing the full event information (multiplicity for instance) and the kinematics of the individual tracks. At this stage the correlations between tracks in the MC events are taken into account. These event-by-event efficiencies are averaged to calculate the probability of a certain decay mode to pass a given set of PID cuts.

Species	Soft p_T	Hard p_T
$\pi^{\pm}$	$K^0_S \to \pi^+\pi^-$	$D^{*+} \to [K^-\pi^+]_{D^0}\pi^+$
$K^{\pm}$	$D^+_s \to [K^+K^-]_{\phi}\pi^+$	$D^{*+} \to [K^-\pi^+]_{D^0}\pi^+$
$p/\overline{p}$	$\Lambda^0 \to p^+\pi^-$	$\Lambda^0 \to p^+\pi^-$, $\Lambda^+_c \to p^+K^-\pi^+$, $\Lambda^0_b \to [p^+K^-\pi^+]_{\Lambda^+_c}\mu^-\overline{\nu}_{\mu}$

Table 5.6: Summary of the modes collected in the pion, kaon and proton PIDCalib calibration samples.

5.4.1 nTracks reweighting and choice of the binning

To produce the PIDCalib efficiency maps we define a three-dimensional binning in the variables {nTracks, p, η} (event multiplicity, momentum and pseudorapidity of the tracks) for a total number of bins of $\{5 \times 8 \times 8\}$. Given the nTracks distribution is not properly reproduced by MC, we perform a reweighting to correct the simulation exploiting the $B^0 \to K^{*0}\gamma$ control channel. To do this we use the cut-based selection of Tab. 5.5 together with the cut on the γ_{CL} and perform a mass fit to calculate signal weights (sWeights) using the ${}_s\mathcal{P}lot$ technique [65]. This way we get a background subtracted $B^0 \to K^{*0}\gamma$ data sample that can be compared to the simulation. The nTracks weights are calculated as the normalized ratio of the nTracks distributions in sweighted data and MC and are used to reweight all the MC samples considered in the analysis. Separated weights are calculated for 2011 and 2012. The nTracks distributions in data and MC before and after the reweighting are displayed in Fig. 5.4. As a last step we apply some pre-selection requirements to filter the MC samples we use. These cuts, which are summarized in Table 5.7, are applied to signal and background MC according to the $B^0 \to \rho^0\gamma$ or $B^0 \to K^{*0}\gamma$ reconstruction hypotheses. In the following, this pre-selection will also be applied to data.

Variable		$B^0 \to \rho^0\gamma$	$B^0 \to K^{*0}\gamma$
Track p	MeV/c	[3000; 100000]	
Track η		[1.5; 5]	
nTracks		[0; 450]	
γ_{CL}		> 0.35	
m_V ($V = \rho^0,\ K^{*0}$)	MeV/c^2	[400; 1200]	$[m^{PDG}_{K^{*0}} - 100; m^{PDG}_{K^{*0}} + 100]$

Table 5.7: Pre-selection cuts.

The cuts in p, η and nTracks correspond to the upper and lower boundaries of each dimension of the PIDCalib binning. The cuts in pseudorapidity and momentum are standard fiducial cuts aimed to ensure a reliable $\pi/K/p$ separation. The event multiplicity is cut at 450 since beyond this threshold no candidates are observed in $K^*\gamma$ sweighted data sample. The ρ^0 mass window is particularly large as the $\pi^+\pi^-$ reconstructed mass spectrum will be used to crosscheck the understanding of the backgrounds. In addition, we require MC events to be truth-matched by matching the mother particle (B) and

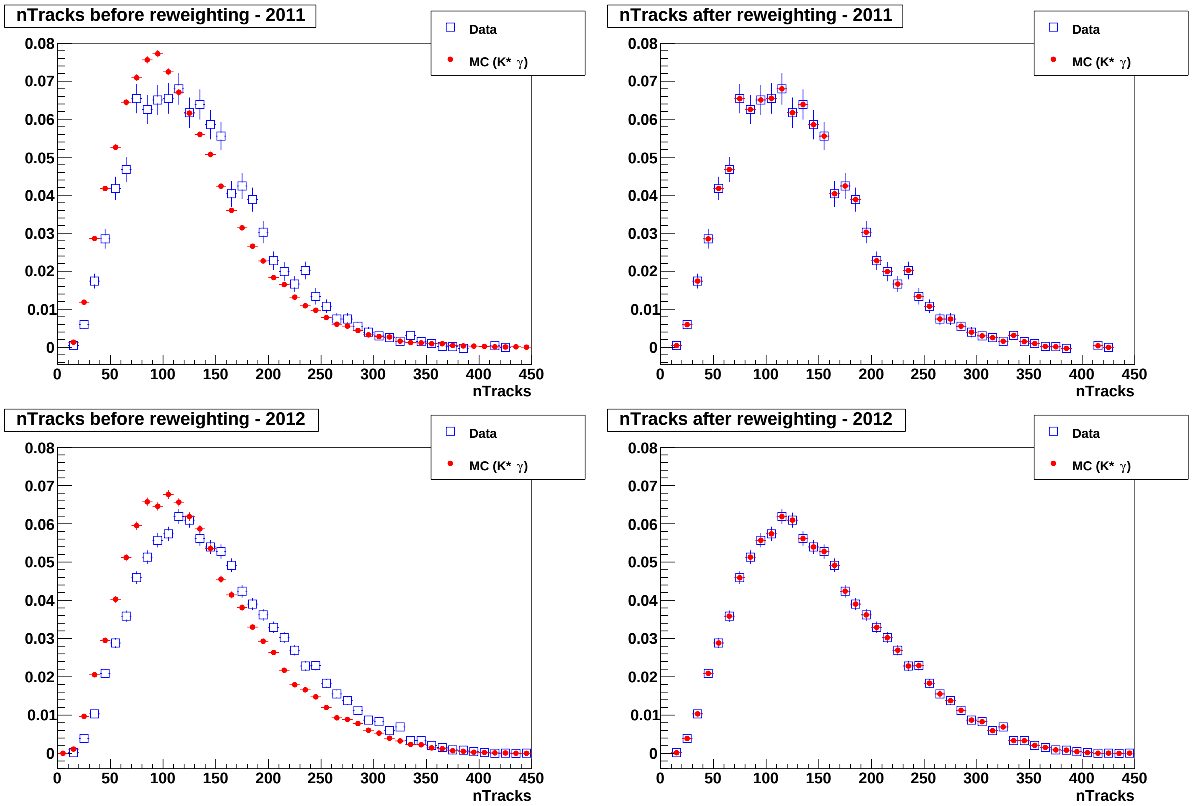


Figure 5.4: $B^0 \to K^{*0}\gamma$ nTracks distribution in 2011 (top) and 2012 (bottom) data and MC before (left) and after (right) performing the reweighting. Empty blue squares stand for sweighted data and red dots for MC.

the charged tracks (π, K or p) on the MC TRUEID. This allows to select the relevant signal and cross-feed candidates to extract the PID efficiency. The TRUEID value of the B candidate depends on the MC sample: for instance B is required to be a true B^0 if we consider the $B^0 \to K^{*0}\gamma$ mode or to be a true Λ_b^0 if we take $\Lambda_b^0 \to \Lambda^{*0}\gamma$. We use the $B^0 \to \rho^0\gamma$ MC samples to set the binning configuration by fixing the boundaries in order to have homogeneously populated bins for signal pions. The same binning configuration as the one extracted for pions is used to produce the K and p efficiency maps. The binning schemes for 2011 and 2012 are reported in Table 5.8.

Year	nTracks boundaries	p boundaries (in MeV/c)	η boundaries
2011	{0; 79; 112; 144; 186; 450}	{3000; 8244; 11595; 15265; 19600; 25180; 33385; 47720; 100000}	{1.5; 2.4055; 2.6390; 2.8415; 3.0340; 3.2330; 3.4685; 3.7750; 5.0}
2012	{0; 94; 130; 169; 223; 450}	{3000; 8300; 11970; 15880; 20530; 26530; 35170; 50300; 100000}	{1.5; 2.3875; 2.6150; 2.8090; 3.0010; 3.2030; 3.4260; 3.7370; 5.0}

Table 5.8: Boundaries of the binning scheme used to produce the PID (mis)identification efficiency maps in the (ProbNNpi-ProbNNk, ProbNNp) plane.

5.4.2 Optimal cut to select pions

The optimization of the cut to select pions is carried out by maximizing the figure of merit

$$\text{FoM (PID)} = \frac{S}{\sqrt{S + \sum_i B_i^{Peak}}} \tag{5.5}$$

where the S term stands for the $B^0 \to \rho^0(\pi^+\pi^-)\gamma$ expected yields written as follows

$$S = \mathcal{L} \times 2\sigma(b\bar{b}) \times f_d \times \mathcal{B}(B^0 \to \rho^0\gamma) \times \epsilon^{MC} \times \epsilon^{Sel} \times \epsilon^{PID}. \tag{5.6}$$

$\mathcal{L}$ is the integrated luminosity, $\sigma(b\bar{b})$ is the production cross-section of $b\bar{b}$ pairs, f_d is the $b \to B^0$ hadronization fraction, $\mathcal{B}$ $(B^0 \to \rho^0\gamma)$ is the branching fraction, ϵ^{MC} is the overall generation efficiency of the related MC sample, ϵ^{Sel} is the MC selection efficiency (reconstruction, trigger and preselection requirements) and ϵ^{PID} is the PID efficiency for the corresponding PID cuts. The values of the $\sigma(b\bar{b})$ cross-section at $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV are respectively $297 \pm 5 \pm 32\,\mu$b [74] and $298 \pm 2 \pm 36\,\mu$b [75], while the value of the hadronization fraction f_d is $(33.7 \pm 2)\%$ [76]. The B_i^{Peak} terms in Eq. 5.5 stand for the main peaking backgrounds that we consider in the optimization procedure, namely the $B^0 \to K^{*0}(K^+\pi^-)\gamma$, $B_s^0 \to \phi(K^+K^-)\gamma$ and $\Lambda_b^0 \to \Lambda^*(pK^-)\gamma$ modes. In order to estimate the expected yields for these channels we have to replace the hadronization fraction, the branching fraction and the efficiencies in Eq. 5.6 according to the mode of interest. While for $\rho\gamma$ we have two charged pions in the final state, for $K^*\gamma$, $\phi\gamma$ and $\Lambda^*\gamma$ we have respectively a $K\pi$, KK and pK final state. The corresponding PID efficiencies are then calculated using pions, kaons and protons efficiency maps according to each final state. Concerning the $K^*\gamma$ and $\phi\gamma$ modes the branching fractions $\mathcal{B}(B^0 \to K^{*0}\gamma) = (4.33 \pm 0.15) \times 10^{-5}$ [7] and $\mathcal{B}(B_s^0 \to \phi\gamma) = (3.52 \pm 0.34) \times 10^{-5}$ (see next paragraph) have to be multiplied by the branching fractions of the vector mesons to the observed final states. These branching fractions are respectively $\mathcal{B}(K^{*0} \to K^+\pi^-) = (66.507 \pm 0.014) \times 10^{-2}$ and $\mathcal{B}$ $(\phi \to K^+K^-) = (48.9 \pm 0.5) \times 10^{-2}$ [7].

For the $\phi\gamma$ and $\Lambda^*\gamma$ modes the expression $f \times \mathcal{B}(....)$ in Eq. 5.6 is written as

$$f_s \times \mathcal{B}(B_s^0 \to \phi\gamma) = f_d \times \mathcal{B}(B^0 \to K^{*0}\gamma) \times r^{\mathcal{B}}_{\phi\gamma}\ , \tag{5.7}$$
$$f_d \times \mathcal{B}(\Lambda_b^0 \to \Lambda^{*0}\gamma) = f_d \times \mathcal{B}(B^0 \to K^{*0}\gamma) \times r^{\mathcal{B}}_{\Lambda^*\gamma}\ , \tag{5.8}$$

with

$$r^{\mathcal{B}}_{\phi\gamma} = \frac{f_s}{f_d} \cdot \frac{\mathcal{B}(B^0_s \to \phi\gamma)}{\mathcal{B}(B^0 \to K^{*0}\gamma)} \ , \tag{5.9}$$

$$r^{\mathcal{B}}_{\Lambda^*\gamma} = \frac{f_{\Lambda^0_b}}{f_d} \cdot \frac{\mathcal{B}(\Lambda^0_b \to \Lambda^*\gamma)}{\mathcal{B}(B^0 \to K^{*0}\gamma)} \ . \tag{5.10}$$

The ratios $r^{\mathcal{B}}_{\phi\gamma}$ and $r^{\mathcal{B}}_{\Lambda^*\gamma}$ come from two LHCb measurements: $r^{\mathcal{B}}_{\phi\gamma} = 0.217 \pm 0.013$ and $r^{\mathcal{B}}_{\Lambda^*\gamma} = 0.297 \pm 0.015$ as extracted from [57] and [77] respectively. To estimate the $\Lambda^*\gamma$ expected yields we use the $\Lambda^0_b \to \Lambda^*(1520)^0\gamma$ generated MC sample. This is a conservative choice because, while we use the full $\Lambda^0_b \to \Lambda^{*0}\gamma$ branching fraction, we perform the calculation for the $\Lambda^*(1520)^0$ which is more likely to be in the ρ^0 mass window. This means that by construction we somehow overestimate the Λ^{*0} contamination.

Every efficiency term ϵ^{PID} is related to a PIDCalib efficiency map generated in the (ProbNNpi-ProbNNk, ProbNNp) plane. The numerical value that we consider to extract the expected yields is the average value of the PID efficiency of the hh' candidate ($\pi\pi$, $K\pi$, KK or pK) corrected by the reweighting in nTracks: this value is obtained from the total PIDCalib efficiency histogram of the decay mode after having merged events with MagUp and MagDown polarities. The effect of the correction arising from the reweighting in nTracks corresponds to an average relative loss of $\sim 1\%$ in the final PID efficiencies.

Before evaluating the figure of merits we test if the expected yields estimated on MC correspond to what is observed on data. To do this we use the $B^0 \to K^{*0}\gamma$ control channel and estimate the signal yields fitting the reconstructed B^0 mass. We apply the cut-based selection of Tab. 5.5 together with the pre-selection of Tab. 5.7, adding for MC the truth-matching requirement. The MC gives a significant yields overestimation for 2011, for which we introduce a correction factor $c_{2011} = 0.724$, while it properly reproduces the observed yields for 2012, leading to $c_{2012} = 1.011$. All the yields needed to extrapolate the figure of merit of Eq. 5.5 are then corrected applying these coefficients. In order to understand the origin of this difference in 2011 we inspected the two main TCKs (*Trigger Configuration Key*), namely `0x760037` and `0x790038`, and crosschecked the results using the exclusive `Hlt2Bd2KstGammaDecision` HLT2 line. We found that inclusive and exclusive HLT2 configurations give compatible results and that MC introduces a yields overestimation of about 20% in `0x760037` and 35% in `0x760038`. For 2012, the various TCKs give compatible results. A further investigation was not carried out but the difference observed in 2011 can be due to several factors such as data taking conditions or online calibrations. However, this does not have a real impact on the rest of the analysis. Indeed, the most important thing is that MC has to correctly reproduce the kinematical variables (p and η) needed to extract the PID efficiencies and the variables used to build up the multivariate classifier (see Sec. 5.6). In both cases MC is found to be reliable.

To optimize the PID cuts, we first performed a scan in the (ProbNNpi-ProbNNk, ProbNNp) plane by steps of 0.1 in X and Y (see Eq. 5.3), identified the region where the FoM is maximal, and then made a further scan by steps of 0.02 in X. Given the $\rho\gamma$ FoM varies only marginally with Y the proton rejection cut will be tuned on the $B^0 \to K^{*0}\gamma$ mode (see Sec. 5.4.3). In the PIDCalib software different tunings (M12TuneVX with X=2,

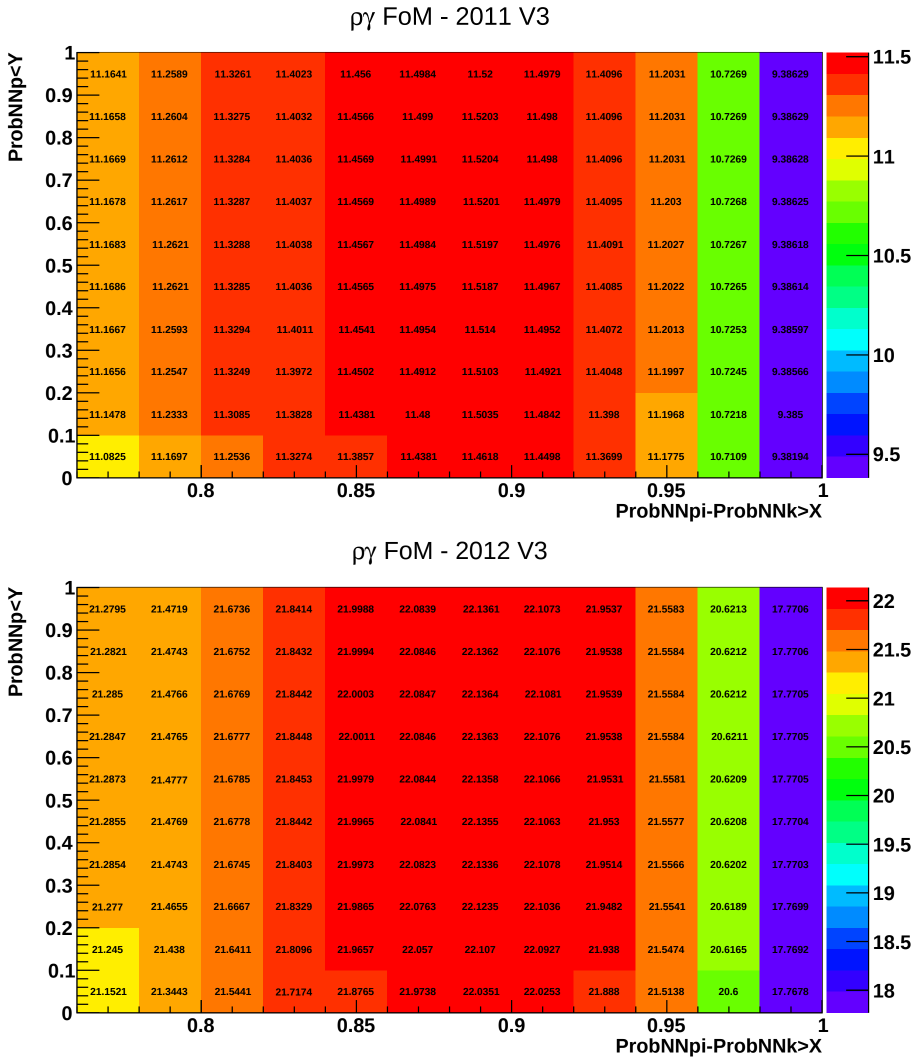


Figure 5.5: $B^0 \to \rho^0\gamma$ figure of merit maps for 2011 (top) and 2012 (bottom) after performing the zoom.

3 and 4) of the ProbNN variables are available for Run 1. Aiming to have the best possible performances we tested all the tunings and found that V3 is the one giving here the best FoM. It is then used as default in the analysis. All the FoM that we evaluated, together with the reference 2D efficiency maps for signal and cross-feeds (V3 only), are displayed in Appendix A. The final figure of merits for 2011 and 2012 are shown in Fig. 5.5.

We choose as optimal region to select pions the FoMs columns corresponding to ProbNNpi-ProbNNk>0.86. This region corresponds for 2012 (2011) to 615 (165) $B^0 \to \rho^0\gamma$ events selected with a 68% PID efficiency and to an intrinsic $B^0 \to K^{*0}\gamma$ contamination of 155 (40) events selected with a 0.5% PID efficiency. As already mentioned the cut in ProbNNp does not significantly affect the value of the figure of merit because the ProbNNpi-ProbNNk cut is already very effective in rejecting the background coming from the $\Lambda_b^0 \to \Lambda^{*0}\gamma$ mode.

5.4.3 Optimal cut to select kaons and reject protons

In order to improve the precision on the ratio of branching fractions measurement, the remaining PID cuts are optimized minimizing the relative uncertainty $\Delta N_{K^*\gamma}/N_{K^*\gamma}$ on the $K^*\gamma$ signal yields which can be extracted as

$$N_{K^*\gamma} = N_T - N^{\pi K}_{K^*\gamma} - N_{\rho\gamma} - N_{\phi\gamma} - N^{Kp}_{\Lambda^*\gamma} - N^{pK}_{\Lambda^*\gamma} \ . \tag{5.11}$$

N_T is the overall yield of selected events, $N^{\pi K}_{K^*\gamma}$ is the $B^0 \to K^{*0}\gamma$ signal reflection contribution involving two misidentications ($\pi \to K$ and $K \to \pi$), $N^{Kp}_{\Lambda^*\gamma}$ and $N^{pK}_{\Lambda^*\gamma}$ are the $\Lambda^*\gamma$ selected yields respectively in case of one ($p \to \pi$) and two misidentications ($p \to K$ and $K \to \pi$). The other terms are self-explaining.

This way, we take into account already at this stage the sources of uncertainty related to branching fractions (and hadronization fractions) aiming to improve the expected precision of the measurement also considering those systematic uncertainties. The uncertainty on the $K^*\gamma$ signal yields can be written as $\Delta N_{K^*\gamma} = \sqrt{\sum_i \Delta N_i^2}$ where the ΔN_i^2 terms are

$$(\Delta N_T)^2 = N_T, \tag{5.12}$$

$$(\Delta N^{PID}_{K^*\gamma})^2 = (N_{K^*\gamma})^2 \times \left(\frac{\Delta\epsilon^{PID}_{K\pi}}{\epsilon^{PID}_{K\pi}}\right)^2 , \tag{5.13}$$

$$(\Delta N^{\pi K}_{K^*\gamma})^2 = (N^{\pi K}_{K^*\gamma})^2 \times \left(\frac{\Delta\epsilon^{PID}_{\pi K}}{\epsilon^{PID}_{\pi K}}\right)^2 , \tag{5.14}$$

$$(\Delta N_{\rho\gamma})^2 = N^2_{\rho\gamma} \times \left(\frac{\Delta\epsilon^{PID}_{\pi\pi}}{\epsilon^{PID}_{\pi\pi}}\right)^2 , \tag{5.15}$$

$$(\Delta N_{\phi\gamma})^2 = N^2_{\phi\gamma} \times \left[\left(\frac{\Delta\epsilon^{PID}_{KK}}{\epsilon^{PID}_{KK}}\right)^2 + \left(\frac{\Delta r^{\mathcal{B}}_{\phi\gamma}}{r^{\mathcal{B}}_{\phi\gamma}}\right)^2\right] , \tag{5.16}$$

$$(\Delta N^{Kp}_{\Lambda^*\gamma})^2 = (N^{Kp}_{\Lambda^*\gamma})^2 \times \left[\left(\frac{\Delta\epsilon^{PID}_{Kp}}{\epsilon^{PID}_{Kp}}\right)^2 + \left(\frac{\Delta r^{\mathcal{B}}_{\Lambda^*\gamma}}{r^{\mathcal{B}}_{\Lambda^*\gamma}}\right)^2 + \left(\frac{\Delta f_{Kp}}{f_{Kp}}\right)^2\right] , \tag{5.17}$$

$$(\Delta N^{pK}_{\Lambda^*\gamma})^2 = (N^{pK}_{\Lambda^*\gamma})^2 \times \left[\left(\frac{\Delta\epsilon^{PID}_{pK}}{\epsilon^{PID}_{pK}}\right)^2 + \left(\frac{\Delta r^{\mathcal{B}}_{\Lambda^*\gamma}}{r^{\mathcal{B}}_{\Lambda^*\gamma}}\right)^2 + \left(\frac{\Delta f_{pK}}{f_{pK}}\right)^2\right] . \tag{5.18}$$

The first term represents the statistical uncertainty on the overall yields and the second one corresponds to the uncertainty on the signal PID selection efficiency (in fact this term does not affect the extracted $N_{K^*\gamma}$ yield itself but it anyhow translates in an uncertainty on the ratio of branching fractions measurement). The other terms are related to the individual background channels to be subtracted for which we propagate the relative uncertainty on PID efficiency and on other sources of systematic uncertainties, as explained in the following. Here a digression on how the $\epsilon^{PID}_{hh'}$ terms and the associated relative uncertainties are computed is needed. Differently from the previous section, $B^0 \to K^{*0}\gamma$ is

	Associated yields	K candidate	π candidate
$B^0 \to K^{*0}\gamma$	$N_{K^*\gamma}$ $N^{\pi K}_{K^*\gamma}$	$K^\pm$ $\pi^\pm$	$\pi^\mp$ $K^\mp$
$B^0 \to \rho^0\gamma$	$N_{\rho\gamma}$	$\pi^\pm$	$\pi^\mp$
$B^0_s \to \phi\gamma$	$N_{\phi\gamma}$	$K^\pm$	$K^\mp$
$\Lambda^0_b \to \Lambda^{*0}\gamma$	$N^{Kp}_{\Lambda^*\gamma}$ $N^{pK}_{\Lambda^*\gamma}$	K^+/K^- $p/\overline{p}$	$\overline{p}/p$ K^-/K^+
Applied cut		ProbNNpi-ProbNNk $< X'$ & ProbNNp $< Y$	ProbNNpi-ProbNNk > 0.86 & ProbNNp $< Y$

Table 5.9: Splitting of the candidates, associated N_i yields and cuts applied in K and π hypotheses.

here considered as the signal mode, then we have to deal with an asymmetric final state where different type of tracks can be (mis)identified not only as a π like in $B^0 \to \rho^0\gamma$. For the $B^0 \to K^{*0}\gamma$ and $\Lambda^0_b \to \Lambda^{*0}\gamma$ channels, we have to distinguish events according to the possible misidentifications. Table 5.9 shows how candidates are split in order to take into account the various reconstruction hypotheses. The related Ni yields contributing to the absolute uncertainties terms in Eq. 5.13-5.18 are also shown. Let us notice that pions are now selected according to the cut optimized in the previous section. The scan of the (ProbNNpi-ProbNNk $< X'$, ProbNNp $< Y$) plane is then made by steps of 0.1 in X' and Y. When applying the generated maps to the MC simulated events, PIDCalib does not only compute the final PID efficiency even-by-event, but also calculates the associated statistical uncertainty. However, while this uncertainty is reliable for what concerns the signal and the main cross-feed modes, the respective histograms for cross-feeds with very low selection efficiency result to be plenty of overflow bins. Therefore, in order to have a coherent approach everywhere and to not define an arbitrary uncertainty range for each cross-feed, we decided to take as absolute uncertainty $\Delta\epsilon^{PID}_{hh'}$ the error on the average value of the PID efficiency of the hh' candidate corrected by the reweighting in nTracks.

The terms $r^{\mathcal{B}}_{\phi\gamma}$ and $r^{\mathcal{B}}_{\Lambda^*\gamma}$in Eq. 5.16-5.18 correspond to the ratios introduced in Eq. 5.9 and 5.10. They are respectively associated to a relative uncertainty of 5.9% and 5.1%.

No amplitude analysis of the $\Lambda^0_b \to \Lambda^{*0}\gamma$ channels exists, so there is no estimation of the contributions of the various Λ^{*0} resonances or of their interferences. To be conservative, the $\Lambda^0_b \to \Lambda^{*0}\gamma$ contribution to the $B^0 \to K^{*0}\gamma$ spectrum is estimated from the various $\Lambda^0_b \to \Lambda^{*0}\gamma$ MC samples and the full spread of those estimations is accounted for in the last terms of Eq. 5.17 and 5.18. We exploit the $\Lambda^0_b \to \Lambda^*(1670)^0\gamma$, $\Lambda^0_b \to \Lambda^*(1820)^0\gamma$ and $\Lambda^0_b \to \Lambda^*(1830)^0\gamma$ MC simulated samples listed on the bottom of Table 4.1 together with the baseline $\Lambda^0_b \to \Lambda^*(1520)^0\gamma$ one. We extract the expected yields for each MC sample, without propagating the PID efficiency, and define

$$f_{hh'} = \frac{f_{hh'}^{MAX}[\Lambda^*(m_1)^0\gamma] + f_{hh'}^{MIN}[\Lambda^*(m_2)^0\gamma]}{2}\ , \tag{5.19}$$

$$\Delta f_{hh'} = \frac{f_{hh'}^{MAX}[\Lambda^*(m_1)^0\gamma] - f_{hh'}^{MIN}[\Lambda^*(m_2)^0\gamma]}{2}\ , \tag{5.20}$$

where m_1 and m_2 represent the Λ^{*0} resonance corresponding respectively to the highest and lowest selection efficiency. The various yields obtained for the different modes and the relative uncertainties entering Eq. 5.17-5.18 are reported in Table 5.10. In those equations, the $\Lambda^*\gamma$ yields are then calculated as $N_{pK\gamma} = f_{Kp}\epsilon_{Kp}^{PID}$ and $N_{pK\gamma} = f_{pK}\epsilon_{pK}^{PID}$.

	Channel	Yields 2011	Yields 2012
Kp	$\Lambda_b^0 \to \Lambda^*(1520)^0\gamma$	3421	6013
	$\Lambda_b^0 \to \Lambda^*(1670)^0\gamma$	1402	2463
	$\Lambda_b^0 \to \Lambda^*(1820)^0\gamma$	798	1430
	$\Lambda_b^0 \to \Lambda^*(1830)^0\gamma$	750	1448
$\Delta f_{Kp}/f_{Kp}$		64%	61.6%
pK	$\Lambda_b^0 \to \Lambda^*(1520)^0\gamma$	4461	8165
	$\Lambda_b^0 \to \Lambda^*(1670)^0\gamma$	2352	4534
	$\Lambda_b^0 \to \Lambda^*(1820)^0\gamma$	581	1100
	$\Lambda_b^0 \to \Lambda^*(1830)^0\gamma$	559	1131
$\Delta f_{pK}/f_{pK}$		77.7%	76.3%

Table 5.10: $\Lambda_b^0 \to \Lambda^{*0}\gamma$ yields estimated without propagating the PID efficiency and associated relative uncertainties for the Kp and pK hypotheses.

To estimate $\Delta N_{K^*\gamma}/N_{K^*\gamma}$, all the MC yields are corrected by the very same $c_{2011} = 0.724$ and $c_{2012} = 1.011$ scaling factors obtained in the previous section. The relative uncertainty maps obtained are shown in Fig. 5.6. We select $X' = 0.5$ and $Y = 0.3$, which allows to minimize $\Delta N_{K^*\gamma}/N_{K^*\gamma}$ while using a tight cut to reject protons as contaminations from Λ_b^0 decays are the less well known. It corresponds to a relative uncertainty of $\sim 1.4\%$ for 2011 and $\sim 0.8\%$ for 2012. It can be noticed that the maps look different between 2011 and 2012 and that the minimal point is located in a different region of the (ProbNNpi-ProbNNk, ProbNNp) plane. This can be explained by the differences between 2011 and 2012 on the signal yields (see Fig. 5.1) and on the ProbNN shapes. In any case the relative increase in $\Delta N_{K^*\gamma}/N_{K^*\gamma}$ between the minimal point and the point that we take as reference is about 2.5% for both 2011 and 2012. The final set of charged PID cuts used in the analysis to select pions and kaons is

$$\pi : \text{ProbNNpi-ProbNNk} > 0.86 \quad \text{and} \quad \text{ProbNNp} < 0.3 \tag{5.21}$$

$$K : \text{ProbNNpi-ProbNNk} < 0.5 \quad \text{and} \quad \text{ProbNNp} < 0.3\ . \tag{5.22}$$

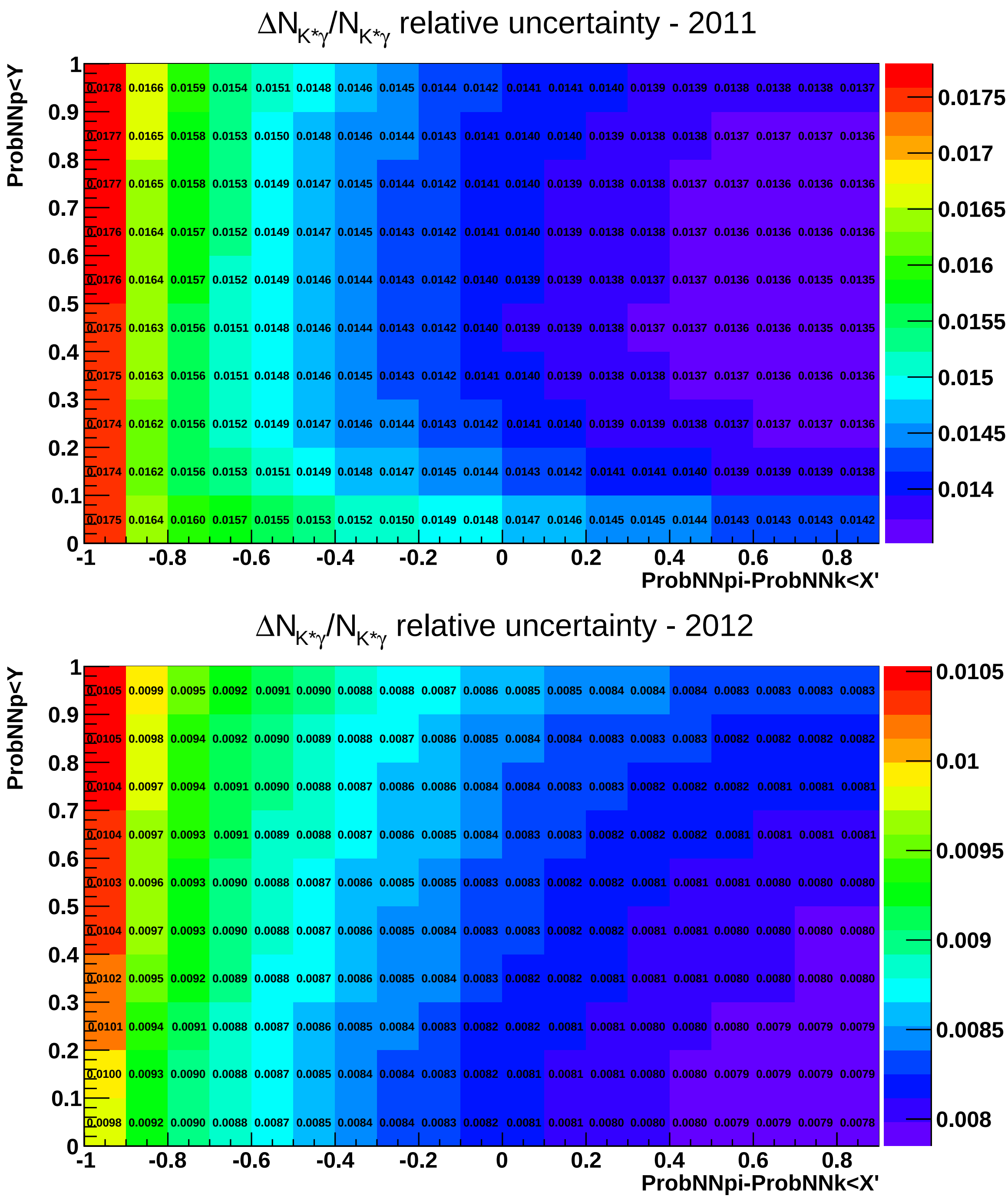


Figure 5.6: $\Delta N_{K^*\gamma}/N_{K^*\gamma}$ relative uncertainty maps for 2011 (top) and 2012 (bottom).

5.5 γ/π^0 separation optimization

The last PID cut that remains to be optimized is the one related to the γ/π^0 separation. In order to reject backgrounds coming from high p_T neutral pions, we use a dedicated Neural Network called IsPhoton [62] that combines a specific set of variables describing the shape of the cluster similarly for the PreShower (PS) and the ECAL together with multiplicity variables in the PS. This IsPhoton variable gives a good discrimination between γ and π^0. However it is not correctly reproduced by the MC simulation and then, to extract the real efficiency for a given cut, calibration samples have to be used as for the charged PID. To do so we use the `GammaPi0SeparationCalib` tool [64] available in Urania v5r0. The calibration samples that can be accessed with the tool exploit the channels $B^0 \to K^{*0}(K^+\pi^-)\gamma$ for photons and $D^{*+} \to D^0(K^-\pi^+\pi^0)\pi^+$ for neutral pions.

The binning that we use to compute the IsPhoton cut efficiency is bi-dimensional in the variables $\{p_T, \eta\}$ for a total number of bins of $\{12 \times 4\}$. Due to the low statistics of the γ samples (36×10^3 signal events over the full Run 1 statistics) and the lack of coverage at low/high pseudorapidities, we choose a higher number of bins in transverse momentum. The binning scheme used to extract the γ/π^0 separation cut efficiency for each year is summarized in Table 5.11. Similarly to what we did for charged PID, the binning is optimized for photons and then we use the same configuration for π^0.

The two peaking background modes with π^0 that we considered in the optimization are $B^0 \to \pi^+\pi^-\pi^0$ and $B^0 \to K^+\pi^-\pi^0$ which have the following branching fractions $\mathcal{B}(B^0 \to \pi^+\pi^-\pi^0) = (2.50 \pm 0.24) \times 10^{-5}$ and $\mathcal{B}(B^0 \to K^+\pi^-\pi^0) = (3.78 \pm 0.32) \times 10^{-5}$. Since the non-resonant $\pi^+\pi^-\pi^0$ mode is still unobserved we took as reference the sum of the $B^0 \to \rho^0\pi^0$ and $B^0 \to \rho^\mp\pi^\pm$ branching fractions. The optimization was carried out only for 2012 since no 2011 MC was available for π^0 modes. Nevertheless, the IsPhoton distributions from γ and π^0 calibration samples are very similar for 2011 and 2012 and the working points are expected to be similar.

The $B^0 \to \rho^0\gamma$ figure of merit is calculated as done so far with Eq. 5.5, where now $\sum_i B_i^{Peak}$ contains also the terms related to the two π^0 modes. The PID efficiency term accounts here not only for charged PID but also for neutral PID. The very same procedure is applied also when calculating the relative uncertainty on the $B^0 \to K^{*0}\gamma$ yield. Now we have three more yields terms contributing to Eq. 5.11, namely $N^{K\pi}_{K\pi\pi^0}$, $N^{\pi K}_{K\pi\pi^0}$ and $N_{\pi\pi\pi^0}$,

Year	p_T boundaries (in MeV/c)	η boundaries
2011	{3000; 3250; 3500; 3800; 4150; 4550; 4950; 5350; 5900; 6550; 7550; 9400; 25000}	{1.5; 2.5; 2.95; 3.4; 5.0}
2012	{3000; 3500; 3780; 4100; 4450; 4800; 5200; 5650; 6200; 6900; 8000; 9800; 25000}	{1.5; 2.5; 2.95; 3.4; 5.0}

Table 5.11: Binning scheme used to extract the γ/π^0 (mis)identification efficiency.

that are associated to the following absolute uncertainties

$$(\Delta N^{K\pi}_{K\pi\pi^0})^2 = (N^{K\pi}_{K\pi\pi^0})^2 \times \left(\frac{\Delta\epsilon^{PID}_{K\pi\pi^0}}{\epsilon^{PID}_{K\pi\pi^0}}\right)^2, \tag{5.23}$$

$$(\Delta N^{\pi K}_{K\pi\pi^0})^2 = (N^{\pi K}_{K\pi\pi^0})^2 \times \left(\frac{\Delta\epsilon^{PID}_{\pi K\pi^0}}{\epsilon^{PID}_{\pi K\pi^0}}\right)^2, \tag{5.24}$$

$$(\Delta N_{\pi\pi\pi^0})^2 = N^2_{\pi\pi\pi^0} \times \left(\frac{\Delta\epsilon^{PID}_{\pi\pi\pi^0}}{\epsilon^{PID}_{\pi\pi\pi^0}}\right)^2. \tag{5.25}$$

We perform a scan in the IsPhoton variable by steps of 0.05 and get the $B^0 \to \rho^0\gamma$ figure of merit as well as the relative uncertainty $\Delta N_{K^*\gamma}/N_{K^*\gamma}$ maps displayed in Fig. 5.7. The cut that maximizes the $\rho\gamma$ FoM, also minimizing the $K^*\gamma$ relative uncertainty, is IsPhoton>0.65. It corresponds for 2012 to an average selection efficiency of $\sim 92\%$ for photon modes and of $\sim 41\%$ for π^0 modes and leads to an overall relative uncertainty on the $K^*\gamma$ yields of $\sim 0.9\%$. For 2012, it corresponds to 566 $B^0 \to \rho^0\gamma$ events selected with an overall PID efficiency of 62% and to an intrinsic $B^0 \to K^{*0}\gamma$ contamination of 143 events selected with an overall PID efficiency of 0.5%. For what concerns π^0 modes, the most significant peaking background for $B^0 \to \rho^0\gamma$ is the $B^0 \to \pi^+\pi^-\pi^0$ channel that corresponds for 2012 to a contamination of 159 events selected with an overall PID efficiency of 26%.

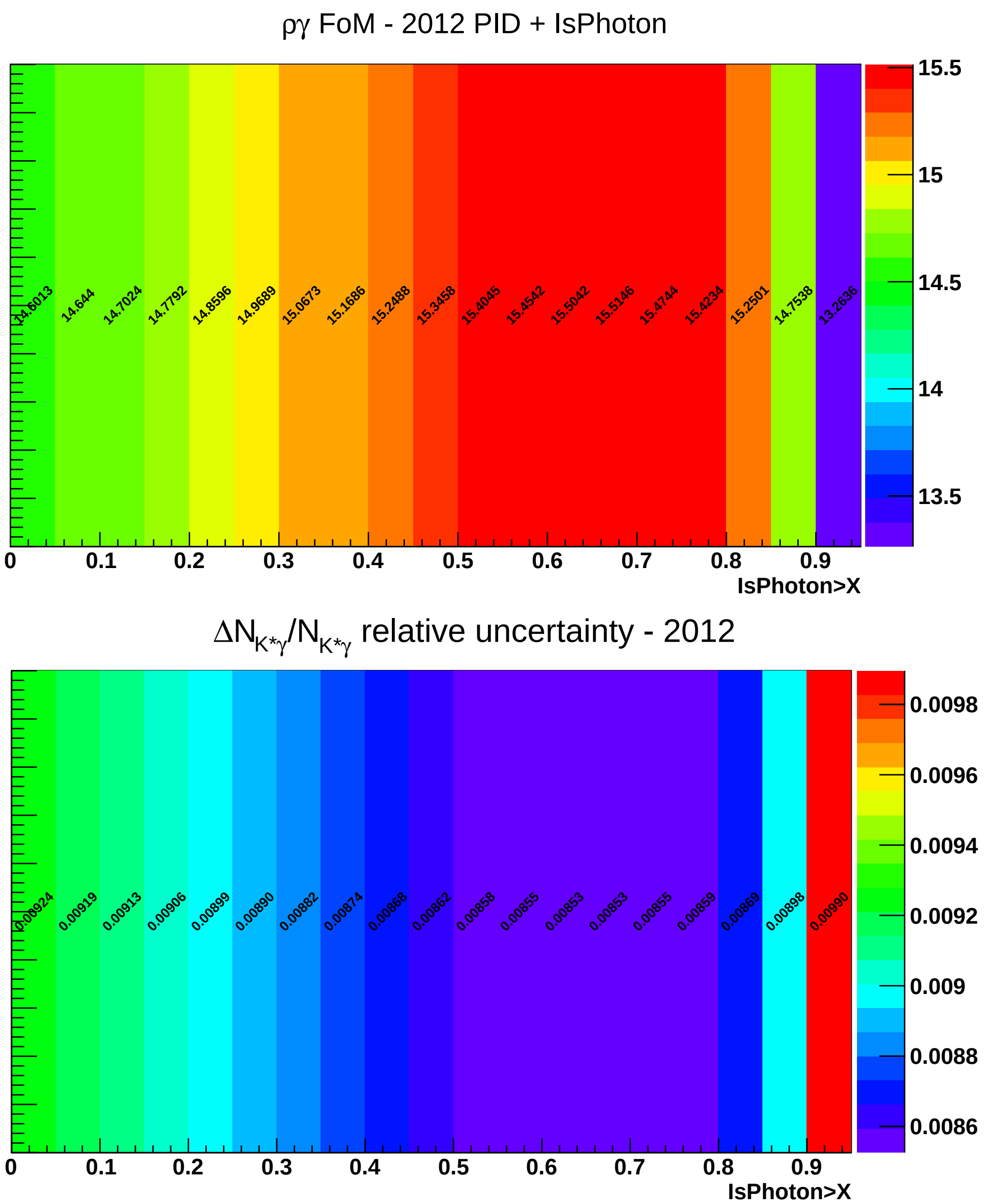


Figure 5.7: Final 2012 $B^0 \to \rho^0\gamma$ figure of merit (top) and relative uncertainty $\Delta N_{K^*\gamma}/N_{K^*\gamma}$ map (bottom) corresponding to the full PID selection.

5.6 MVA selection optimization

In order to achieve the best possible rejection of the combinatorial background we build up a multivariate based classifier exploiting the TMVA package [78] (*Toolkit for Multivariate Data Analysis*). The signal events used in the MVA training are taken from $B^0 \to \rho^0\gamma$ MC simulated events, while combinatorial background events are taken from the right-hand sideband (RHSB) of real data in a reconstructed B mass range within [5580, 6600] MeV/c^2. We decided to discard events beyond 6600 MeV/c^2 since at higher masses the shape of the RHSB is affected by differences between online and offline calibrations of the calorimeters. Aiming to choose the optimal set of variables to be used in the final discriminator we select only those which exhibit a good data/MC agreement and which have weak linear correlations or different correlations for background and signal events. These two aspects are essential because poorly modeled or highly correlated variables would negatively affect the discriminating power of the classifier. We also limit as much as possible the usage of kinematical variables to not restrict the phase space of the decay. In order to check the data/MC agreement for the different variables we exploit the same background subtracted $B^0 \to K^{*0}\gamma$ data sample used to perform the nTracks reweighting and we compare it with $B^0 \to K^{*0}\gamma$ and $B^0 \to \rho^0\gamma$ MC samples. The data/MC comparison plots for 2011 and 2012 are shown in Appendix B.

We consider a set of nine discriminating variables for the MVA training: two kinematical, six topological and an isolation one. Those are defined in Tab. 5.12. The kinematical variables are the p_T and η of the B candidate; the topological variables are the θ_{DIRA}, χ^2_{IP}, FD and $\chi^2_{\mathrm{vtx}}/\mathrm{ndf}$ of the B candidate complemented by the χ^2_{IP} of the two charged tracks; the isolation variable is the smallest $\Delta\chi^2_{\mathrm{vtx}}$ of the B candidate. Apart from the p_T and η of the B, we consider in fact the $\log_{10}$ of the variables to avoid very long tails that could compromise the training.

Before moving to the final MVA training we performed an extensive series of studies in order to find the best possible configuration:

- We tested different algorithms and found that those providing the best performances are the two Boosted Decision Trees, BDT (adaptive boosting) and BDTG (gradient boosting), and the MLP neural network (Multi-Layer Perceptron). We always used the default settings except for BDT for which we had to reduce the number of trees to avoid overtraining.

- We verified that our set of variable is minimal considering that when removing a variable the performances worsen significantly.

- In order to maximally exploit the RHSB statistics we tested configurations with $k = 2, 5$ and 10 folds and found that a higher folding does not improve the expected significance. Then we took two folds as default. The k-folding technique, already used in [79], consists on dividing the RHSB in a certain number of samples (folds) in order to benefit from a higher statistics in the training phase (for instance for $k = 4$ we use three folds for training and one for testing). Such technique can be useful when dealing with low statistics.

Variable	Description
$p_T(B)$	The transverse momentum of the reconstructed B candidate.
$\eta(B)$	The pseudo-rapidity of the reconstructed B candidate.
$\theta_{DIRA}(B)$	The direction angle defined as the angle between the reconstructed momentum of the B candidate and the line connecting its primary and secondary vertices.
$\chi^2_{\mathrm{IP}}(B)$	χ^2 increase of the primary vertex when adding the tracks of the reconstructed B candidate.
$FD(B)$	Flight distance of the reconstructed B candidate.
$\chi^2_{\mathrm{vtx}}/\mathrm{ndf}(B)$	χ^2 of the vertex of the reconstructed B candidate divided by the associated number of degrees of freedom.
$\chi^2_{\mathrm{IP}}(\pi^{\pm})$	χ^2 increase of the primary vertex when adding one of the two charged tracks of the reconstructed B candidate.
Smallest $\Delta\chi^2_{\mathrm{vtx}}(B)$	χ^2 increase of the vertex of the reconstructed B candidate when adding an extra most compatible track.

Table 5.12: List of variables used as inputs in the MVA discriminant.

- We observed that it is better to train on samples with a preselection cut such that (Smallest $\Delta\chi^2_{\mathrm{vtx}} > 1$ || Smallest $\Delta\chi^2_{\mathrm{vtx}} == -1$). We decided to keep the -1 events since they correspond to B candidates which are not compatible with any additional track: these events are then associated to the most isolated B vertex.

The efficiencies on signal MC and background as a function of the smallest $\Delta\chi^2_{\mathrm{vtx}}$ preselection cut are displayed in Fig. 5.8. Putting smallest $\Delta\chi^2_{\mathrm{vtx}} > 1$ for 2012 (2011) corresponds to select background data events with a 72% (69%) efficiency and MC signal events with a 96% efficiency for both years. Including the candidates with smallest $\Delta\chi^2_{\mathrm{vtx}} == -1$ we recover about 1% of the MC signal events. As a crosscheck we fitted the $B^0 \to K^{*0}\gamma$ control channel and observed a consistent increase of the signal yields out of the mass fit when keeping those events.

Events in the RHSB to be used as background sample in the training are required to pass the trigger conditions of Tab. 5.4, the $B^0 \to \rho^0\gamma$ preselection cuts displayed in Tab. 5.7, the cut on the smallest $\Delta\chi^2_{\mathrm{vtx}}$ and the full PID selection. MC signal events are additionally required to have a B mass within the range [4000; 6600] MeV/c^2, to be truth-matched (see Sec. 5.4.1) and to have a value of the background category variable such that `BKGCAT==0` or `BKGCAT==50`. The `BKGCAT==0` events are signal events where the reconstructed decay matches perfectly with the generated decay, while the `BKGCAT==50` ones stand for the so-called low mass background in which a Bremsstrahlung photon is emitted and not reconstructed. In this background category, all candidate's final state daughters are well identified and the difference between the true mother mass and the

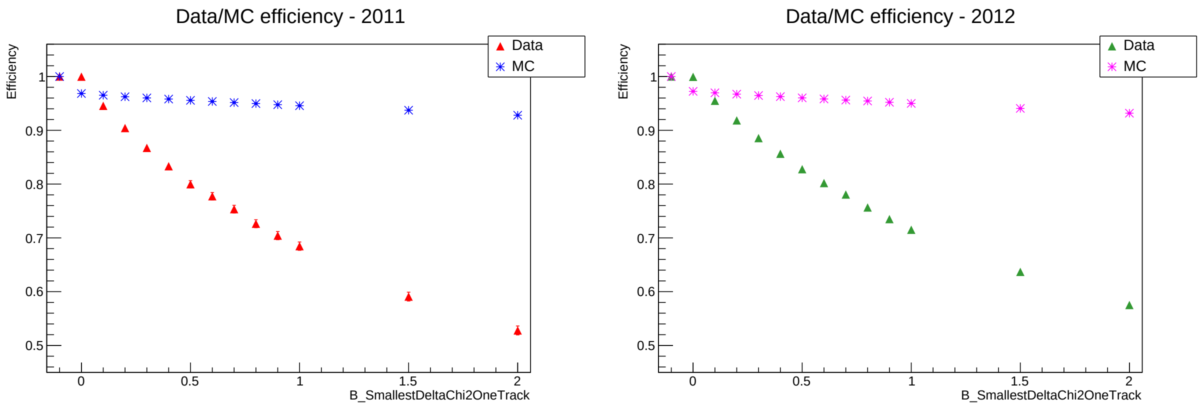


Figure 5.8: Efficiencies on signal MC and background as a function of the B smallest $\Delta\chi^2_{\rm vtx}$ preselection cut for 2011 (left) and 2012 (right). Data events are taken from the $B^0 \to \rho^0\gamma$ RHSB, MC ones are from the $B^0 \to \rho^0\gamma$ MC samples.

candidate invariant mass is below 100 MeV/c^2. As a final clarification, since we cannot directly cut on the PID variables for simulated events, MC events are used with a global per-event weight which is the overall PID efficiency term $\epsilon^{PID}_{\pi\pi\gamma}$ already containing the correction in nTracks.

In order to split the background and signal samples in two subsamples for the training and testing of the multivariate discriminant we use a pseudo-random number generated for every candidate as a function of (`eventNumber` + `runNumber`). We generate a reproducible random number r ranging from 0 and 1 on an event-by-event basis thanks to the `SetSeed` method. The r distributions were checked and found to be flat in both data and MC. Events having $r \leqslant 0.5$ are used to train MVA$_1$ and the remaining ones are used to train MVA$_2$. The events are then crossed-over in the testing of the MVA response for overtraining and when calculating the efficiencies for a given MVA cut. When applying the MVA response to data, RHSB events used to train MVA$_1$ are selected using MVA$_2$, and vice versa. The data events in the rest of the B mass spectrum, that were not used in the training of the two MVAs, are analogously cut either using MVA$_1$ or MVA$_2$ with the same random selection algorithm as done for the RHSB events. This way the full Run 1 statistics can be used in the analysis without introducing any bias.

As a result of the MVA trainings for different years and folds, the BDTG and MLP algorithms were found to give the best performances and, since those are very similar, we carried out the optimization for both before making the final choice. However, as better detailed afterwards, we can already mention here that the chosen baseline is the MLP neural network. The histograms of the variable distributions superimposing RHSB background events and MC signal events as well as the correlation maps of the variables used to train MVA$_1$ in 2012 are shown respectively in Fig. 5.9 and 5.10. In Fig. 5.11 are shown the 2012 MLP$_1$ response with the related overtraining test and the background-rejection efficiency versus signal selection efficiency curves (ROC curves). Looking at Fig. 5.10 we can notice that the majority of the variables exhibit weak linear correlations. However, for those variables that have high linear correlations, we can see that these significantly differ when

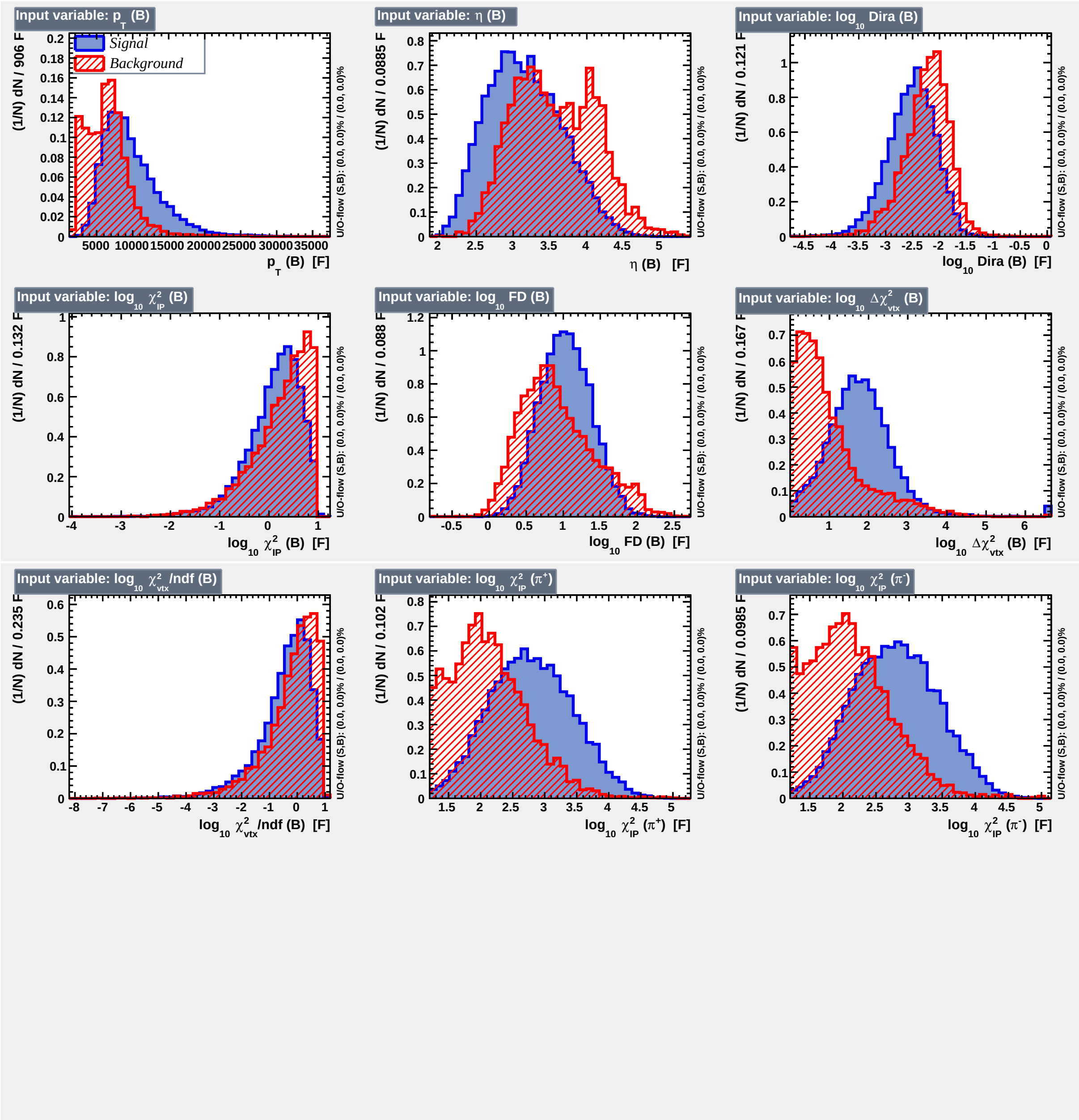


Figure 5.9: Distribution of variables used in the training of MVA_1 for 2012, superimposing RHSB background events (in red) and MC signal events (in blue). The distribution of the other MVAs for 2011 and 2012 are displayed in Appendix B.

comparing signal and background events. Looking at Fig. 5.11 we can see that there is no evident sign of overtraining, as also indicated by the Kolmogorov-Smirnov test [80].

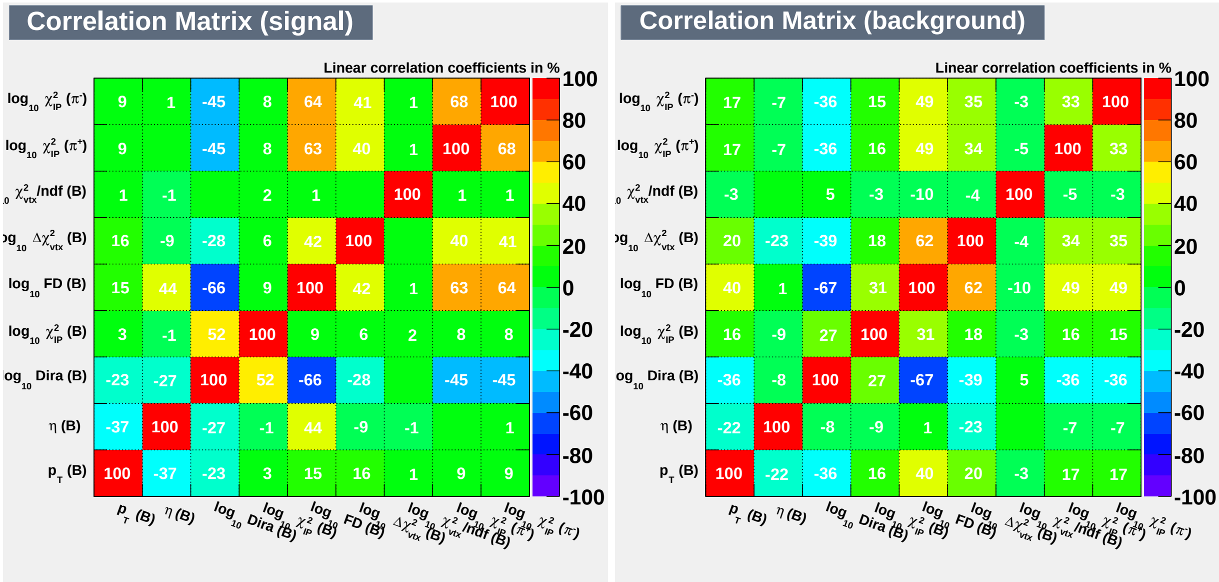


Figure 5.10: Linear correlation of variables used in the training of MVA$_1$ for 2012 for (left) signal events and (right) background events. See Appendix B for the other MVAs.

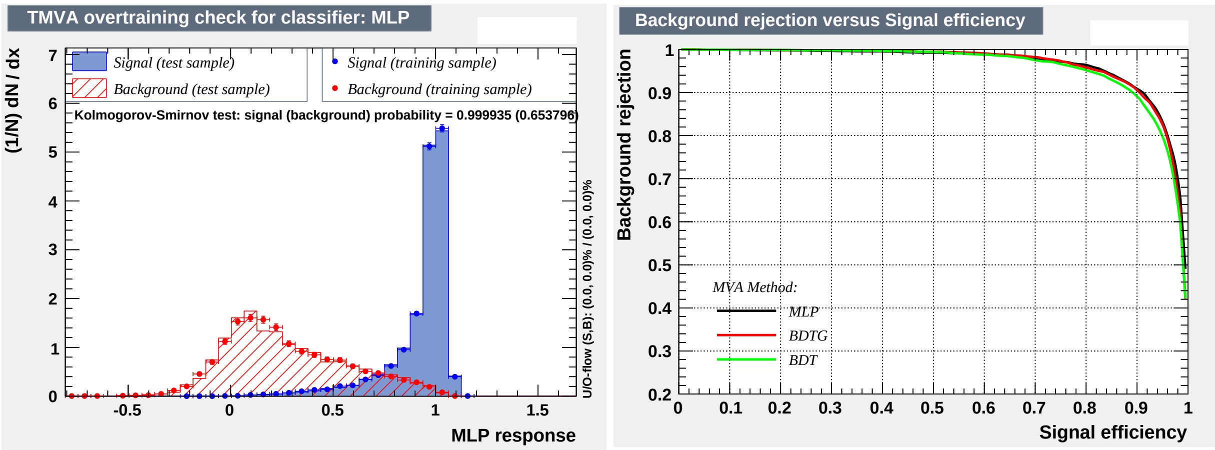


Figure 5.11: 2012 MLP$_1$ response with overtraining test (left) and 2012 MVA$_1$ ROC-curves (right) for the best performing algorithms. See Appendix B for the remaining plots.

The rankings of the different variables are given in Tab. 5.13. The variable ranking of the MLP neural network uses the sum of the weights-squared of the connections between the variable's neuron in the input layer and the first hidden layer [78]. The importance of the i variable is given by

$$I_i = \overline{x}_i^2 \sum_{j=1}^{n_{\mathrm{h}}} \left[w_{ij}^{(1)} \right]^2 , \quad i = 1, \ldots, n_{\mathrm{var}} \ , \tag{5.26}$$

where $\overline{x}_i$ is the mean of the input variable i, n_{var} and n_{h} are the number of neurons in

Variable	Importance			
	2011		2012	
	MLP_1	MLP_2	MLP_1	MLP_2
$p_T(B)$	8.247	11.21	12.54	14.03
$\eta(B)$	1.656	7.176	3.068	3.73
$\log_{10}\theta_{DIRA}(B)$	0.556	1.237	0.721	0.66
$\log_{10}\chi^2_{\text{IP}}(B)$	10.46	2.725	9.642	7.11
$\log_{10} FD(B)$	2.696	2.753	3.602	2.97
$\log_{10}\chi^2_{\text{vtx}}/\text{ndf}(B)$	17.95	8.428	9.844	16.32
$\log_{10}\chi^2_{\text{IP}}(\pi^+)$	2.617	1.474	2.153	2.75
$\log_{10}\chi^2_{\text{IP}}(\pi^-)$	1.258	0.992	1.775	1.79
$\log_{10}$Smallest $\Delta\chi^2_{\text{vtx}}(B)$	5.248	3.898	3.908	6.84

Table 5.13: Importance of the variables used in the MLP training.

the input and inner layer respectively and the apex $^{(1)}$ indicates that connections (w_{ij}) between the input layer and the first hidden layer are considered. However, such variable importances should be taken with care as they depend on the mean of the variable and only on a fraction of the weights. For this reason they are given here only as illustration of the individual power of the discriminating variables and they were not considered as selection criterion when choosing the optimal set of variables for the final MVA training.

The ROC integrals and the signal efficiencies for background efficiencies of 1%, 10% and 30% are reported in Tab. 5.14. The performances of the different MLPs are very similar.

In order to optimize the MVA cut and take into account the correct normalizations for both signal and background events we maximize the figure of merit

$$\text{FoM (MVA)} = \frac{S}{\sqrt{S + \sum_i B_i^{Peak} + cB^{Comb}}} \quad \text{with} \quad c = \frac{N_{bkg}^{3\sigma}}{N_{bkg}^{RHSB}}. \tag{5.27}$$

All the yield terms S, B^{Peak} and cB^{Comb} are here calculated in the $\pm 3\sigma$ signal region, *i.e.* [5000; 5580] MeV/c^2. B^{Comb} is the number of combinatorial background events in the RHSB passing the MVA cut and c is a scaling factor used to estimate the number of expected background events cB^{Comb} under the signal peak. The factor c, which is calculated without applying any MVA cut, is given by the ratio $N_{bkg}^{3\sigma}/N_{bkg}^{RHSB}$ where N_{bkg}^{RHSB} is the total number of events in the RHSB and $N_{bkg}^{3\sigma}$ is the number of events extrapolated under the signal region. Aiming to not bias the optimization with a single assumption on the combinatorial background shape we considered two extrapolations: we fit the RHSB either with an exponential function or with a first order polynomial. The fits to the RHSB and the parameters needed to calculate the c factors are shown in Appendix B. The stability of the RHSB shape was checked for different values of the MVA cut and the fitted shape parameters were found to be consistent with the baseline ones (no MVA cut). The S and

MLP classifier	Year	ROC integral	Signal efficiencies (for a given background efficiency)		
			$\epsilon_{Comb} = 1\%$	$\epsilon_{Comb} = 10\%$	$\epsilon_{Comb} = 30\%$
MLP_1	2011	0.968	0.505	0.913	0.982
MLP_2	2011	0.966	0.526	0.899	0.983
MLP_1	2012	0.966	0.582	0.911	0.979
MLP_2	2012	0.964	0.565	0.903	0.977

Table 5.14: ROC integrals and signal selection efficiencies for background selection efficiencies of 1%, 10% and 30%.

B^{Peak} terms in Eq. 5.27 are rescaled as usual taking into account the overall PID efficiency of the MC events that pass the MVA cut.

The optimization procedure leading to the final figure of merits is carried out in two main steps: in the first phase we calculate the FoMs in the various configurations without applying any correction, while in the second phase we use the $B^0 \to K^{*0}\gamma$ control channel to test if the expected yields estimated on MC correspond to what we actually observe in data. We perform this crosscheck analogously to what we did for the PID optimization, but in this case we apply the selection tuned for the analysis (including the MVA cut fixed in the first step) instead of the cut-based selection. By fitting the $K^*\gamma$ mass spectrum we extract the definitive correction factors to be applied to all the MC yields: these factors, which are compatible with the ones extracted previously, are $C_{2011} = 0.705$ and $C_{2012} = 1.04$. We checked the compatibility of the optimization of the MVA cuts from the FoMs obtained with and without applying the yields corrections and found that they are consistent. The MVA cut has been optimized for both BDTG and MLP algorithms: the MLP gives the best figure of merit overall and is then taken as reference for the analysis. After having performed various additional crosschecks to test the FoM stability we found that the optimal cuts are compatible:

- in the different folds;
- with the linear and exponential extrapolations of the combinatorial background;
- with and without considering the $hh\gamma$ or $hh\pi^0$ background sources;
- when comparing 2011 and 2012.

Even though slight differences between 2011 and 2012 can be spotted, for the sake of simplicity we consider a unique cut for both years which is MLP > 0.96. The final figure of merits obtained considering only combinatorial background events are displayed in Fig. 5.12. The other FoMs are shown in Appendix B.

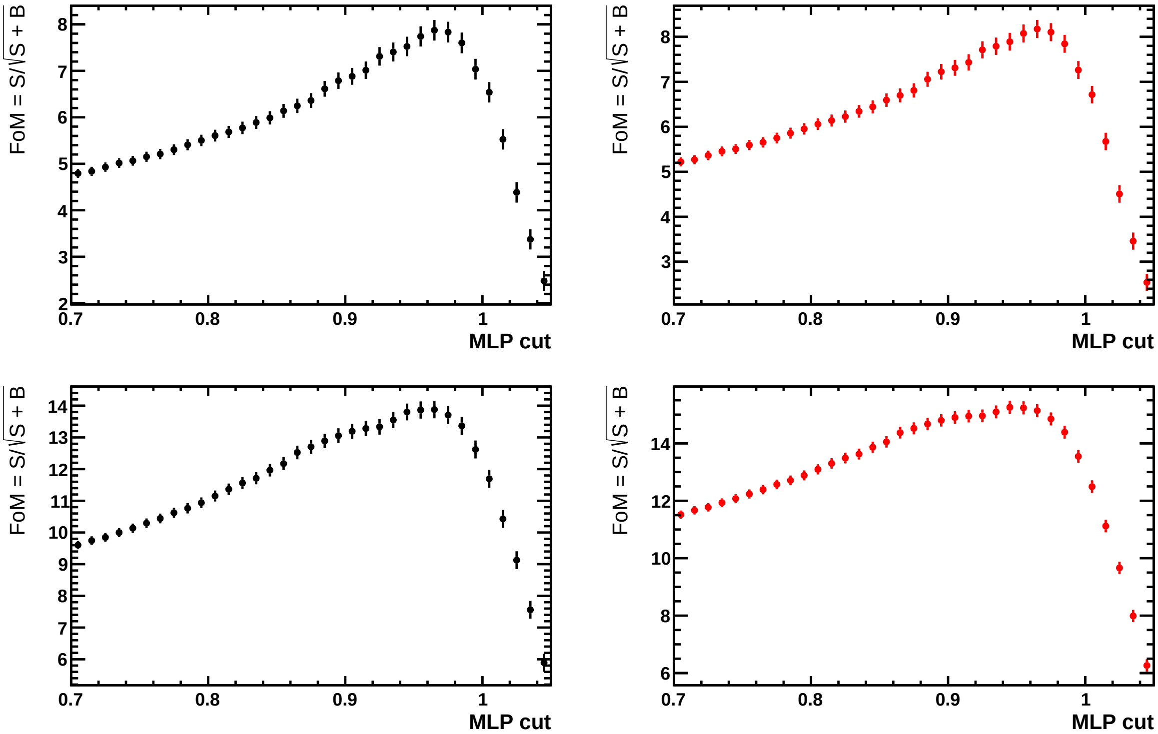


Figure 5.12: $B^0 \to \rho^0\gamma$ figure of merits obtained considering only combinatorial background events for exponential (left) and linear (right) extrapolation hypotheses in 2011 (top) and 2012 (bottom).

The expected yields for signal, peaking backgrounds and combinatorial background considering the optimal MLP cut are displayed in Tab. 5.15 and 5.16, respectively for $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ signal hypotheses. These numbers, that include the MC yields corrections previously mentioned ($C_{2011} = 0.705$ and $C_{2012} = 1.04$), are limited to the $\pm 3\sigma$ signal region and do not include the associated uncertainties whose propagation is carried out in the next section where the relative contaminations are derived. The yields associated to a double misidentification of the tracks, $N_{\phi\gamma}$ for the $\rho^0\gamma$ signal hypothesis, $N^{\pi K}_{K^*\gamma}$, $N^{pK}_{\Lambda^*\gamma}$ and $N^{\pi K}_{K\pi\pi^0}$ for the $K^{*0}\gamma$ signal hypothesis, are found to be negligible. The comprehensive study of the various background sources and the extraction of their shapes is discussed in Sec. 6.2. The number of expected signal events for the $B^0 \to K^{*0}\gamma$ control channel in the full mass range [4000; 6600] MeV/c^2 is 3481 ± 251 and 12002 ± 838 respectively for 2011 and 2012. The errors on the yields include the uncertainty related to the MC statistics and all the uncertainties associated to the various terms entering Eq. 5.6.

Mode	Associated yields	Year	
		2011	2012
$B^0 \to \rho^0\gamma$	$N_{\rho\gamma}$	95	335
$B^0 \to K^{*0}\gamma$	$N_{K^*\gamma}$	19	74
$B^0_s \to \phi\gamma$	$N_{\phi\gamma}$	–	–
$\Lambda^0_b \to \Lambda^{*0}\gamma$	$N_{\Lambda^*\gamma}$	0.3	2.2
$B^0 \to \pi^+\pi^-\pi^0$	$N_{\pi\pi\pi^0}$		76
$B^0 \to K^+\pi^-\pi^0$	$N_{K\pi\pi^0}$		1.3
Comb. Bkg.	$c_{exp}B^{Comb}$	61	247
	$c_{lin}B^{Comb}$	49	154

Table 5.15: Expected yields when applying the optimized cut MLP > 0.96 for the $B^0 \to \rho^0\gamma$ channel. These yields are the ones entering the figure of merits.

Mode	Associated yields	Year	
		2011	2012
$B^0 \to K^{*0}\gamma$	$N_{K^*\gamma}$	3304	11411
	$N^{\pi K}_{K^*\gamma}$	0.8	2.9
$B^0 \to \rho^0\gamma$	$N_{\rho\gamma}$	6.1	23
$B^0_s \to \phi\gamma$	$N_{\phi\gamma}$	1.2	4.9
$\Lambda^0_b \to \Lambda^{*0}\gamma$	$N^{Kp}_{\Lambda^*\gamma}$	19	83
	$N^{pK}_{\Lambda^*\gamma}$	1.1	8.1
$B^0 \to \pi^+\pi^-\pi^0$	$N_{\pi\pi\pi^0}$		5.6
$B^0 \to K^+\pi^-\pi^0$	$N^{K\pi}_{K\pi\pi^0}$		158
	$N^{\pi K}_{K\pi\pi^0}$		0.1

Table 5.16: Expected yields when applying the optimized cut MLP > 0.96 for the $B^0 \to K^{*0}\gamma$ channel.

5.7 Derivation of peaking backgrounds contaminations

Since all the selection requirements are fixed we can now extract the relative contaminations to the $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ spectra for the various peaking backgrounds. The contamination of the $H_b \to X$ decay is computed as a fraction of the signal yields and defined as

$$\mathcal{C}^{signal}_{H_b\to X} = \frac{N^{\rm sel}(H_b\to X)}{N^{\rm sel}(signal)} = \frac{\mathcal{E}^{signal}_{H_b\to X}}{\mathcal{E}^{signal}} \times \frac{f_{H_b}}{f_d} \times \frac{\mathcal{B}(H_b\to X)}{\mathcal{B}(signal)} \times \frac{\mathcal{B}^{Eff}_{H_b\to X}}{\mathcal{B}^{Eff}_{signal}}, \tag{5.28}$$

where *signal* corresponds to either the $B^0\to \rho^0\gamma$ or $B^0\to K^{*0}\gamma$ mode, f_{H_b}/f_d is the ratio of hadronization fractions, and $\mathcal{E}^{signal}$ and $\mathcal{E}^{signal}_{H_b\to X}$ are the global efficiencies for the signal and background of interest. The notation $\mathcal{B}^{Eff}_{H_b\to X}$ is clarified at the level of Eq. 5.32. The global efficiency $\mathcal{E}$ can be written as

$$\mathcal{E} = \epsilon^{MC} \times \epsilon^{Sel} \times \epsilon^{PID}(\times f^{Signal}_{\Lambda^*\gamma}) \ , \tag{5.29}$$

following the notation introduced in Eq. 5.6: here ϵ^{Sel} includes also the MVA cut and ϵ^{PID} stands for the overall PID efficiency. The $f^{signal}_{\Lambda^*\gamma}$ term between parenthesis in Eq. 5.29 is only relevant for the $\Lambda^0_b\to \Lambda^{*0}\gamma$ mode. Is is used to account for the different evaluations of the product $\epsilon^{MC}\epsilon^{Sel}$ obtained from the various $\Lambda^0_b\to \Lambda^{*0}\gamma$ MC samples: $\Lambda^0_b\to \Lambda^*(1520)^0\gamma$, $\Lambda^0_b\to \Lambda^*(1670)^0\gamma$, $\Lambda^0_b\to \Lambda^*(1820)^0\gamma$ and $\Lambda^0_b\to \Lambda^*(1830)^0\gamma$. This quantity and its associated uncertainty are calculated as

$$f^{signal}_{\Lambda^*\gamma} = \frac{(\epsilon^{MC}\epsilon^{Sel})^{MAX}[\Lambda^{*0}\gamma] + (\epsilon^{MC}\epsilon^{Sel})^{MIN}[\Lambda^{*0}\gamma]}{2\times(\epsilon^{MC}\epsilon^{Sel})[\Lambda^*(1520)^0\gamma]} \ , \tag{5.30}$$

$$\Delta f^{signal}_{\Lambda^*\gamma} = \frac{(\epsilon^{MC}\epsilon^{Sel})^{MAX}[\Lambda^{*0}\gamma] - (\epsilon^{MC}\epsilon^{Sel})^{MIN}[\Lambda^{*0}\gamma]}{2\times(\epsilon^{MC}\epsilon^{Sel})[\Lambda^*(1520)^0\gamma]} \ . \tag{5.31}$$

The values of the $(\epsilon^{MC}\epsilon^{Sel})$ products for the different $\Lambda^0_b\to \Lambda^{*0}\gamma$ MC samples together with the values of $f^{signal}_{\Lambda^*\gamma}$ and the relative uncertainty $\Delta f^{signal}_{\Lambda^*\gamma}/f^{signal}_{\Lambda^*\gamma}$ for both $B^0\to \rho^0\gamma$ and $B^0\to K^{*0}\gamma$ hypotheses are displayed in Tab. 5.17.

The individual efficiencies are given in Tab. 5.18 and 5.19 for the $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ hypotheses respectively. The global efficiencies are reported in Tab. 5.20 together with the $r^{Eff}_{H_b\to X}$ ratios defined as

$$r^{Eff}_{H_b\to X} = \frac{f_{H_b}}{f_d} \times \frac{\mathcal{B}(H_b\to X)}{\mathcal{B}(B^0\to K^{*0}\gamma)} \times \frac{\mathcal{B}^{Eff}_{H_b\to X}}{\mathcal{B}(K^{*0}\to K^+\pi^-)} \ , \tag{5.32}$$

where $\mathcal{B}^{Eff}_{H_b\to X}$ is $\mathcal{B}(K^{*0} \to K^+\pi^-) = (66.507 \pm 0.014) \times 10^{-2}$ for $B^0 \to K^{*0}\gamma$, $\mathcal{B}(\phi\to K^+K^-) = (48.9\pm 0.5)\times 10^{-2}$ for $B^0_s\to \phi\gamma$ and 1 for the other modes. The branching fraction $\mathcal{B}(\Lambda^{*0} \to pK^-)$ is in fact already included in the measurement of the $\Lambda^0_b\to \Lambda^{*0}\gamma$ branching fraction we are using. For the $B^0\to K^{*0}\gamma$ spectra, one has:

$$\mathcal{C}^{K^*\gamma}_{H_b\to X} = \frac{\mathcal{E}^{K^*\gamma}_{H_b\to X}}{\mathcal{E}^{K^*\gamma}} \times r^{Eff}_{H_b\to X} \ . \tag{5.33}$$

Hypothesis		Channel	$\epsilon^{MC}\epsilon^{Sel}$ (%) 2011	2012
$B^0 \to \rho^0\gamma$		$\Lambda_b^0 \to \Lambda^*(1520)^0\gamma$	0.135 ± 0.001	0.120 ± 0.001
		$\Lambda_b^0 \to \Lambda^*(1670)^0\gamma$	0.192 ± 0.002	0.157 ± 0.001
		$\Lambda_b^0 \to \Lambda^*(1820)^0\gamma$	0.162 ± 0.002	0.130 ± 0.001
		$\Lambda_b^0 \to \Lambda^*(1830)^0\gamma$	0.153 ± 0.002	0.124 ± 0.001
$f_{\Lambda^*\gamma}^{\rho\gamma}$	(%)		121.25 ± 1.07	115.37 ± 1.04
$\Delta f_{\Lambda^*\gamma}^{\rho\gamma}/f_{\Lambda^*\gamma}^{\rho\gamma}$	(%)		17.52 ± 0.89	13.32 ± 0.91
$B^0 \to K^{*0}\gamma$		$\Lambda_b^0 \to \Lambda^*(1520)^0\gamma$	0.079 ± 0.001	0.063 ± 0.001
		$\Lambda_b^0 \to \Lambda^*(1670)^0\gamma$	0.037 ± 0.001	0.031 ± 0.001
		$\Lambda_b^0 \to \Lambda^*(1820)^0\gamma$	0.020 ± 0.001	0.016 ± 0.001
		$\Lambda_b^0 \to \Lambda^*(1830)^0\gamma$	0.019 ± 0.001	0.016 ± 0.001
$f_{\Lambda^*\gamma}^{K^*\gamma}$	(%)		62.15 ± 1.73	62.46 ± 1.58
$\Delta f_{\Lambda^*\gamma}^{K^*\gamma}/f_{\Lambda^*\gamma}^{K^*\gamma}$	(%)		60.89 ± 3.26	60.11 ± 2.95

Table 5.17: Summary of the $\epsilon^{MC}\epsilon^{Sel}$ efficiency products for the different $\Lambda^*\gamma$ resonances and values of $f_{\Lambda^*\gamma}^{signal}$ and $\Delta f_{\Lambda^*\gamma}^{signal}/f_{\Lambda^*\gamma}^{signal}$ for both the $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ signal hypotheses.

We now derive how the peaking background contaminations are constrained in the $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ simultaneous fit. We introduce $\mathcal{Y}$ the ratio of the fitted signal yields for $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$:

$$\mathcal{Y} = \frac{N^{\rho\gamma}}{N^{K^*\gamma}} = \frac{\mathcal{E}^{\rho\gamma}}{\mathcal{E}^{K^*\gamma}} \times r_{\rho\gamma}^{Eff} \ . \tag{5.34}$$

$r_{\rho\gamma}^{Eff}$, the ratio of branching fractions we aim to measure, can then be expressed as

$$r_{\rho\gamma}^{Eff} = \frac{\mathcal{E}^{K^*\gamma}}{\mathcal{E}^{\rho\gamma}} \times \mathcal{Y} \ . \tag{5.35}$$

The contamination of $B^0 \to \rho^0\gamma$ in the $B^0 \to K^{*0}\gamma$ spectrum is

$$\mathcal{C}_{\rho\gamma}^{K^*\gamma} = \frac{\mathcal{E}_{\rho\gamma}^{K^*\gamma}}{\mathcal{E}^{K^*\gamma}} \times r_{\rho\gamma}^{Eff} = \frac{\mathcal{E}_{\rho\gamma}^{K^*\gamma}}{\mathcal{E}^{K^*\gamma}} \frac{\mathcal{E}^{K^*\gamma}}{\mathcal{E}^{\rho\gamma}} \times \mathcal{Y} \ , \tag{5.36}$$

that simplifies as

$$\mathcal{C}_{\rho\gamma}^{K^*\gamma} = \frac{\mathcal{E}_{\rho\gamma}^{K^*\gamma}}{\mathcal{E}^{\rho\gamma}} \times \mathcal{Y} = \mathcal{D}_{\rho\gamma}^{K^*\gamma} \times \mathcal{Y} \ . \tag{5.37}$$

We introduce here the notation $\mathcal{D}_{H_b \to X}^{signal}$, which corresponds to the terms used to constrain the peaking background contaminations in the simultaneous fit. Let us remind that the efficiencies only with a superscript are related to the signal reconstructed as signal, while

Channel	Year	ϵ^{MC} (%)	ϵ^{Sel} (%)	ϵ^{PID} (%)
$B^0 \to \rho^0\gamma$	2011	20.322 ± 0.030	0.931 ± 0.008	62.270 ± 0.186
	2012	20.593 ± 0.063	0.764 ± 0.005	62.780 ± 0.163
$B^0 \to K^{*0}\gamma$	2011	21.274 ± 0.037	0.867 ± 0.003	0.398 ± 0.001
	2012	21.479 ± 0.055	0.737 ± 0.003	0.428 ± 0.002
$\Lambda_b^0 \to \Lambda^{*0}\gamma$	2011	23.073 ± 0.050	0.585 ± 0.005	0.023 ± 0.001
	2012	23.266 ± 0.080	0.516 ± 0.004	0.044 ± 0.006
$B^0 \to \pi^+\pi^-\pi^0$	2012	33.015 ± 0.078	0.009 ± 0.001	28.270 ± 1.349
$B^0 \to K^+\pi^-\pi^0$	2012	17.279 ± 0.064	0.032 ± 0.001	0.214 ± 0.017
$\Lambda_b^0 \to \Lambda^*(1670)^0\gamma$	2011	25.142 ± 0.058	0.765 ± 0.006	
	2012	25.355 ± 0.048	0.619 ± 0.005	
$\Lambda_b^0 \to \Lambda^*(1820)^0\gamma$	2011	39.952 ± 0.075	0.406 ± 0.005	
	2012	39.980 ± 0.100	0.325 ± 0.003	
$\Lambda_b^0 \to \Lambda^*(1830)^0\gamma$	2011	39.958 ± 0.075	0.383 ± 0.004	
	2012	39.885 ± 0.110	0.312 ± 0.003	

Table 5.18: Summary table of all the individual efficiencies in the $B^0 \to \rho^0\gamma$ signal hypothesis. Let us remind that the PID efficiencies for the $\Lambda^*(1670)^0\gamma$, $\Lambda^*(1820)^0\gamma$ and $\Lambda^*(1830)^0\gamma$ modes are not needed since they are used only to calculate the $f^{\rho\gamma}_{\Lambda^*\gamma}$ term.

the ones with both a subscript and a superscript correspond to peaking backgrounds (subscript) reconstructed under a given signal hypothesis (superscript).

The other $\mathcal{D}^{K^*\gamma}_{H_b\to X}$ terms are simply the contamination fractions:

$$\mathcal{C}^{K^*\gamma}_{\phi\gamma} = \frac{\mathcal{E}^{K^*\gamma}_{\phi\gamma}}{\mathcal{E}^{K^*\gamma}} \times r^{Eff}_{\phi\gamma} = \mathcal{D}^{K^*\gamma}_{\phi\gamma} \ , \tag{5.38}$$

$$\mathcal{C}^{K^*\gamma}_{\Lambda^*\gamma} = \frac{\mathcal{E}^{K^*\gamma}_{\Lambda^*\gamma}}{\mathcal{E}^{K^*\gamma}} \times r^{Eff}_{\Lambda^*\gamma} = \mathcal{D}^{K^*\gamma}_{\Lambda^*\gamma} \ , \tag{5.39}$$

$$\mathcal{C}^{K^*\gamma}_{K\pi\pi^0} = \frac{\mathcal{E}^{K^*\gamma}_{K\pi\pi^0}}{\mathcal{E}^{K^*\gamma}} \times r^{Eff}_{K\pi\pi^0} = \mathcal{D}^{K^*\gamma}_{K\pi\pi^0} \ , \tag{5.40}$$

$$\mathcal{C}^{K^*\gamma}_{\pi\pi\pi^0} = \frac{\mathcal{E}^{K^*\gamma}_{\pi\pi\pi^0}}{\mathcal{E}^{K^*\gamma}} \times r^{Eff}_{\pi\pi\pi^0} = \mathcal{D}^{K^*\gamma}_{\pi\pi\pi^0} \ . \tag{5.41}$$

Channel	Year	ϵ^{MC} (%)	ϵ^{Sel} (%)	ϵ^{PID} (%)
$B^0 \to K^{*0}\gamma$	2011	21.274 ± 0.037	0.805 ± 0.003	71.210 ± 0.085
	2012	21.479 ± 0.055	0.662 ± 0.003	70.810 ± 0.109
$B^0 \to \rho^0\gamma$	2011	20.322 ± 0.030	0.560 ± 0.006	6.911 ± 0.093
	2012	20.593 ± 0.063	0.466 ± 0.004	7.345 ± 0.086
$B_s^0 \to \phi\gamma$	2011	23.203 ± 0.041	0.322 ± 0.003	0.389 ± 0.003
	2012	23.519 ± 0.059	0.293 ± 0.002	0.400 ± 0.003
$\Lambda_b^0 \to \Lambda^{*0}\gamma$	2011	23.073 ± 0.050	0.342 ± 0.004	2.182 ± 0.065
	2012	23.266 ± 0.080	0.269 ± 0.003	2.919 ± 0.088
$B^0 \to K^+\pi^-\pi^0$	2012	17.279 ± 0.064	0.026 ± 0.001	28.590 ± 0.811
$B^0 \to \pi^+\pi^-\pi^0$	2012	33.015 ± 0.078	0.006 ± 0.001	3.561 ± 0.477
$\Lambda_b^0 \to \Lambda^*(1670)^0\gamma$	2011	25.142 ± 0.058	0.148 ± 0.003	
	2012	25.355 ± 0.048	0.122 ± 0.002	
$\Lambda_b^0 \to \Lambda^*(1820)^0\gamma$	2011	39.952 ± 0.075	0.050 ± 0.002	
	2012	39.980 ± 0.100	0.039 ± 0.001	
$\Lambda_b^0 \to \Lambda^*(1830)^0\gamma$	2011	39.958 ± 0.075	0.048 ± 0.002	
	2012	39.885 ± 0.110	0.041 ± 0.001	

Table 5.19: Summary table of all the individual efficiencies in the $B^0 \to K^{*0}\gamma$ signal hypothesis. Let us remind that the PID efficiencies for the $\Lambda^*(1670)^0\gamma$, $\Lambda^*(1820)^0\gamma$ and $\Lambda^*(1830)^0\gamma$ modes are not needed since they are used only to calculate the $f^{K^*\gamma}_{\Lambda^*\gamma}$ term.

For the $B^0 \to \rho^0\gamma$ mass spectrum, the contamination fractions can be written as

$$\begin{aligned}\mathcal{C}^{\rho\gamma}_{H_b\to X} = & \frac{\mathcal{E}^{\rho\gamma}_{H_b\to X}}{\mathcal{E}^{\rho\gamma}} \times \frac{f_{H_b}}{f_d} \times \frac{\mathcal{B}(H_b \to X)}{\mathcal{B}(B^0 \to K^{*0}\gamma)} \times \frac{\mathcal{B}^{Eff}_{H_b\to X}}{\mathcal{B}(K^{*0} \to K^+\pi^-)} \\ & \times \frac{\mathcal{B}(B^0 \to K^{*0}\gamma)}{\mathcal{B}(B^0 \to \rho^0\gamma)} \times \mathcal{B}(K^{*0} \to K^+\pi^-) = \frac{\mathcal{E}^{\rho\gamma}_{H_b\to X}}{\mathcal{E}^{\rho\gamma}} \times \frac{r^{Eff}_{H_b\to X}}{r^{Eff}_{\rho\gamma}}\ .\end{aligned} \tag{5.42}$$

Using Eq. 5.35, one has

$$\mathcal{C}^{\rho\gamma}_{H_b\to X} = \frac{\mathcal{E}^{\rho\gamma}_{H_b\to X}}{\mathcal{E}^{\rho\gamma}} \frac{\mathcal{E}^{\rho\gamma}}{\mathcal{E}^{K^*\gamma}} \times \frac{r^{Eff}_{H_b\to X}}{\mathcal{Y}}\ . \tag{5.43}$$

This simplifies as

$$\mathcal{C}^{\rho\gamma}_{H_b\to X} = \frac{\mathcal{E}^{\rho\gamma}_{H_b\to X}}{\mathcal{E}^{K^*\gamma}} \times r^{Eff}_{H_b\to X} \times \mathcal{Y}^{-1}. \tag{5.44}$$

Channel	Year	$\mathcal{E}^{\rho\gamma}_{H_b\to X}$ ($\times10^{-4}$)	$\mathcal{E}^{K^*\gamma}_{H_b\to X}$ ($\times10^{-4}$)	$r^{Eff}_{H_b\to X}$
$B^0\to\rho^0\gamma$	2011	11.7813 ± 0.1086	0.7865 ± 0.0136	0.0299 ± 0.0053
	2012	9.8772 ± 0.0758	0.7049 ± 0.0105	
$B^0\to K^{*0}\gamma$	2011	0.0734 ± 0.0003	12.1951 ± 0.0522	1
	2012	0.0678 ± 0.0005	10.0685 ± 0.0547	
$B^0_s\to\phi\gamma$	2011	$-$	0.0291 ± 0.0004	0.1596 ± 0.0095
	2012	$-$	0.0276 ± 0.0003	
$\Lambda^0_b\to\Lambda^{*0}\gamma$	2011	0.0038 ± 0.0007	0.1070 ± 0.0652	0.4468 ± 0.0227
	2012	0.0061 ± 0.0012	0.1141 ± 0.0687	
$B^0\to\pi^+\pi^-\pi^0$	2012	0.0840 ± 0.0102	0.0071 ± 0.0015	0.8681 ± 0.0886
$B^0\to K^+\pi^-\pi^0$	2012	0.0012 ± 0.0001	0.1284 ± 0.0062	1.3125 ± 0.1201

Table 5.20: Global efficiencies and values of the $r^{Eff}_{H_b\to X}$ ratios. For the $\Lambda^0_b\to\Lambda^{*0}\gamma$ mode, the $\mathcal{E}^{signal}_{\Lambda^*\gamma}$ global efficiencies include the $f^{signal}_{\Lambda^*\gamma}$ term and its uncertainty. $r^{Eff}_{\rho\gamma}$ corresponds to the ratio of branching ratios we aim to measure: the PDG value is reported here.

The various terms $\mathcal{D}^{\rho\gamma}_{H_b\to X}$ are then given by:

$$\mathcal{C}^{\rho\gamma}_{K^*\gamma} = \frac{\mathcal{E}^{\rho\gamma}_{K^*\gamma}}{\mathcal{E}^{K^*\gamma}} \times \mathcal{Y}^{-1} = \mathcal{D}^{\rho\gamma}_{K^*\gamma} \times \mathcal{Y}^{-1} \ , \tag{5.45}$$

$$\mathcal{C}^{\rho\gamma}_{\Lambda^*\gamma} = \frac{\mathcal{E}^{\rho\gamma}_{\Lambda^*\gamma}}{\mathcal{E}^{K^*\gamma}} \times r^{Eff}_{\Lambda^*\gamma} \times \mathcal{Y}^{-1} = \mathcal{D}^{\rho\gamma}_{\Lambda^*\gamma} \times \mathcal{Y}^{-1} \ , \tag{5.46}$$

$$\mathcal{C}^{\rho\gamma}_{\pi\pi\pi^0} = \frac{\mathcal{E}^{\rho\gamma}_{\pi\pi\pi^0}}{\mathcal{E}^{K^*\gamma}} \times r^{Eff}_{\pi\pi\pi^0} \times \mathcal{Y}^{-1} = \mathcal{D}^{\rho\gamma}_{\pi\pi\pi^0} \times \mathcal{Y}^{-1} \ , \tag{5.47}$$

$$\mathcal{C}^{\rho\gamma}_{K\pi\pi^0} = \frac{\mathcal{E}^{\rho\gamma}_{K\pi\pi^0}}{\mathcal{E}^{K^*\gamma}} \times r^{Eff}_{K\pi\pi^0} \times \mathcal{Y}^{-1} = \mathcal{D}^{\rho\gamma}_{K\pi\pi^0} \times \mathcal{Y}^{-1} \ , \tag{5.48}$$

All the values of the $\mathcal{D}^{signal}_{H_b\to X}$ terms are reported in Tab. 5.21. The expressions of the contaminations as well as their values assuming for $r^{Eff}_{\rho\gamma}$ the value of the PDG, are also shown. The expected $B^0\to\rho^0\gamma$ yields corresponding to this hypothesis are 101 and 352 for 2011 and 2012 respectively. All the uncertainties related to the efficiencies and the $r^{Eff}_{H_b\to X}$ ratios have been propagated when calculating the $\mathcal{D}^{signal}_{H_b\to X}$ associated errors. For the $B^0\to\rho^0\gamma$ spectrum, an additional contribution related to $\Lambda^0_b\to N^{*0}(p\pi^-)\gamma$ decays has to be considered. This mode has never been observed and we will assume $\mathcal{C}^{K^*\gamma}_{N^*\gamma} = \mathcal{C}^{K^*\gamma}_{\Lambda^*\gamma}$. One may notice that, even if $\Lambda^0_b\to\Lambda^*\gamma$ is one of the dominant contribution in the $B^0\to K^{*0}\gamma$ spectrum, $B^0\to K^{*0}\gamma$ and $B^0\to\pi^+\pi^-\pi^0$ are expected to be by far the dominant contaminations in the $B^0\to\rho^0\gamma$ spectrum.

Signal	$\mathcal{C}^{signal}_{H_b \to X}$	Year	$\mathcal{D}^{signal}_{H_b \to X}$ (%)	Contamination assuming $r^{Eff}_{\rho\gamma}$ as in the PDG (%)
$B^0 \to \rho^0\gamma$	$\mathcal{D}^{\rho\gamma}_{K^*\gamma} \times \mathcal{Y}^{-1}$	2011	0.602 ± 0.004	20.744
		2012	0.673 ± 0.006	22.960
	$\mathcal{D}^{\rho\gamma}_{\Lambda^*\gamma} \times \mathcal{Y}^{-1}$	2011	0.014 ± 0.003	0.480
		2012	0.027 ± 0.006	0.923
	$\mathcal{D}^{\rho\gamma}_{\pi\pi\pi^0} \times \mathcal{Y}^{-1}$	2012	0.724 ± 0.115	24.963
	$\mathcal{D}^{\rho\gamma}_{K\pi\pi^0} \times \mathcal{Y}^{-1}$	2012	0.016 ± 0.002	0.533
$B^0 \to K^{*0}\gamma$	$\mathcal{D}^{K^*\gamma}_{\rho\gamma} \times \mathcal{Y}$	2011	6.676 ± 0.131	0.194
		2012	7.137 ± 0.120	0.209
	$\mathcal{D}^{K^*\gamma}_{\phi\gamma}$	2011	0.038 ± 0.002	
		2012	0.044 ± 0.003	
	$\mathcal{D}^{K^*\gamma}_{\Lambda^*\gamma}$	2011	0.392 ± 0.245	
		2012	0.506 ± 0.313	
	$\mathcal{D}^{K^*\gamma}_{K\pi\pi^0}$	2012	1.674 ± 0.173	
	$\mathcal{D}^{K^*\gamma}_{\pi\pi\pi^0}$	2012	0.061 ± 0.014	

Table 5.21: Summary table gathering all the values of the $\mathcal{D}^{signal}_{H_b \to X}$ terms to be used to constrain the peaking background contaminations in the simultaneous fit. The expressions of the contaminations as well as their values, assuming $r^{Eff}_{\rho\gamma}$ as in the PDG, are also reported. $\mathcal{Y}$ it the ratio of the fitted signal yields for $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$: $\mathcal{Y} = N^{\rho\gamma}/N^{K^*\gamma}$.

Chapter 6

Fit of the $B^0 \to K^{*0}\gamma$ mass spectrum

After applying the full selection various background contributions remain and populate the $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ mass spectra. In order to perform a reliable measurement all these backgrounds have to be properly described and included in the final B mass fit model. In this chapter we describe how the signal and background shapes are extracted in both $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ reconstruction hypotheses and how the total PDF used to fit the data is built up. The fit stability is also checked exploiting the $B^0 \to K^{*0}\gamma$ control channel. The implementation of the final simultaneous fit model and its validation are discussed in Chap. 7.

The shapes to be included in the mass fit model are extracted through a fit to the invariant mass distribution of MC events passing the full selection except for the PID cuts, since the ProbNN and IsPhoton variables are not well described in the simulation. However, the PID requirements affect the shape of the mass distribution because every candidate has a certain probability to pass the cuts: this probability corresponds to the overall event-by-event PID efficiency. In order to take into account this effect we reweight the simulated events propagating the PID efficiencies ϵ^{PID} for every MC sample we use. The estimation of the correct PID efficiency for signal and cross-feed modes is obtained requiring the matching on the MC TRUEID of the charged tracks (π, K or p) but, differently from what we did so far, here the matching on the B candidate is not applied. The reason behind this choice is that non-matched B events mainly populate the sides of the mass peak, therefore by including them we aim to obtain the most realistic description of the tails. Let us remind that when considering the $B^0 \to K^{*0}\gamma$ signal mode the πK and pK cross-feed reflections are neglected, then the corresponding shapes are not extracted. For what concerns the partially reconstructed backgrounds the PID weights are calculated requiring the $\pi\pi$ matching in the $B^0 \to \rho^0\gamma$ reconstruction hypothesis and the $K\pi$ matching in the $B^0 \to K^{*0}\gamma$ one.

All the signal and background shapes are extracted through an unbinned extended maximum likelihood fit to the reconstructed B mass spectrum in the channel of interest. However, performing a likelihood fit to weighted events without any correction would lead to a wrong evaluation of the uncertainties on the fitted parameters. Those would indeed scale not with the actual MC statistics but with the sum of the weights. As suggested in [81], we decided to perform the fits scaling the log-likelihood functions by a factor α defined as

$$\alpha = \frac{\sum_i^N W_i}{\sum_i^N W_i^2}, \tag{6.1}$$

where N is the number of events in the sample and W_i is the overall PID weight of the i^{th} event. This log-likelihood scaling procedure only affects the evaluated uncertainties not the fitted central values. Given the scaling we use, the number of yields extracted in the fit, which should vary according to the average PID efficiency, is not shown with the other fit parameters since it does not actually reflect the real corrected yields.

6.1 Signal shapes

The signal invariant mass distribution is extracted fitting the $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ simulated events in the full mass range [4000; 6600] MeV/c^2 and using an asymmetric double tail Crystal-Ball PDF [82] defined as

$$\mathrm{DCB}(m;\mu,\sigma,\alpha_L,n_L,\alpha_R,n_R) = \begin{cases} A_L\left(B_L - \dfrac{m-\mu}{\sigma}\right)^{-n_L} & \text{for } \dfrac{m-\mu}{\sigma} \leq -\alpha_L \\ \exp\left\{-\dfrac{(m-\mu)^2}{2\sigma^2}\right\} & \text{for } -\alpha_L < \dfrac{m-\mu}{\sigma} < \alpha_R \\ A_R\left(B_R + \dfrac{m-\mu}{\sigma}\right)^{-n_R} & \text{for } \dfrac{m-\mu}{\sigma} \geq \alpha_R \end{cases} \tag{6.2}$$

where $\alpha_{L(R)} > 0$ and $A_{L(R)}$ and $B_{L(R)}$ are

$$\begin{aligned} A_i &= \left(\frac{n_i}{|\alpha_i|}\right)^{n_i} \exp\left\{-\frac{|\alpha_i|^2}{2}\right\} \\ B_i &= \frac{n_i}{|\alpha_i|} - |\alpha_i| \quad . \end{aligned} \tag{6.3}$$

The difference between left and right resolutions is considered by defining two independent sigmas, σ_L and σ_R, according to

$$\sigma = \begin{cases} \sigma_L & \text{if } m < \mu \\ \sigma_R & \text{else} \end{cases} \quad . \tag{6.4}$$

If we do not take into account such difference, the function of Eq. 6.2 reduces to a symmetric double-tail Crystal Ball with a single resolution σ. We chose an asymmetric PDF as reference to extract the signal shapes because this function allows to obtain a better fit on MC. The left tail describes the low mass region and accounts for the Bremsstrahlung and possible losses in the photon energy due to the fiducial volume of the calorimeter. The tail at high masses models the imperfections of the tracking. However, in the case of radiative decays large pile-up deposits in the ECAL cluster forming the photon candidate are also likely to contribute. The invariant mass resolution of the B candidate depends on the ECAL resolution as it is dominated by the γ contribution. The value of the μ parameter depends on the ECAL average calibration.

In Table 6.1 we summarize the values of the fit parameters as extracted from the $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ MC samples. The fits to the $\rho\gamma$ and $K^*\gamma$ mass spectra for 2011 and 2012 MC are shown respectively in Fig. 6.1 and 6.2. In the fit to the data, the tail parameters $n_{L(R)}$ and $\alpha_{L(R)}$ will be fixed to the values extracted in the simulation while the μ and $\sigma_{L(R)}$ will be left free to float.

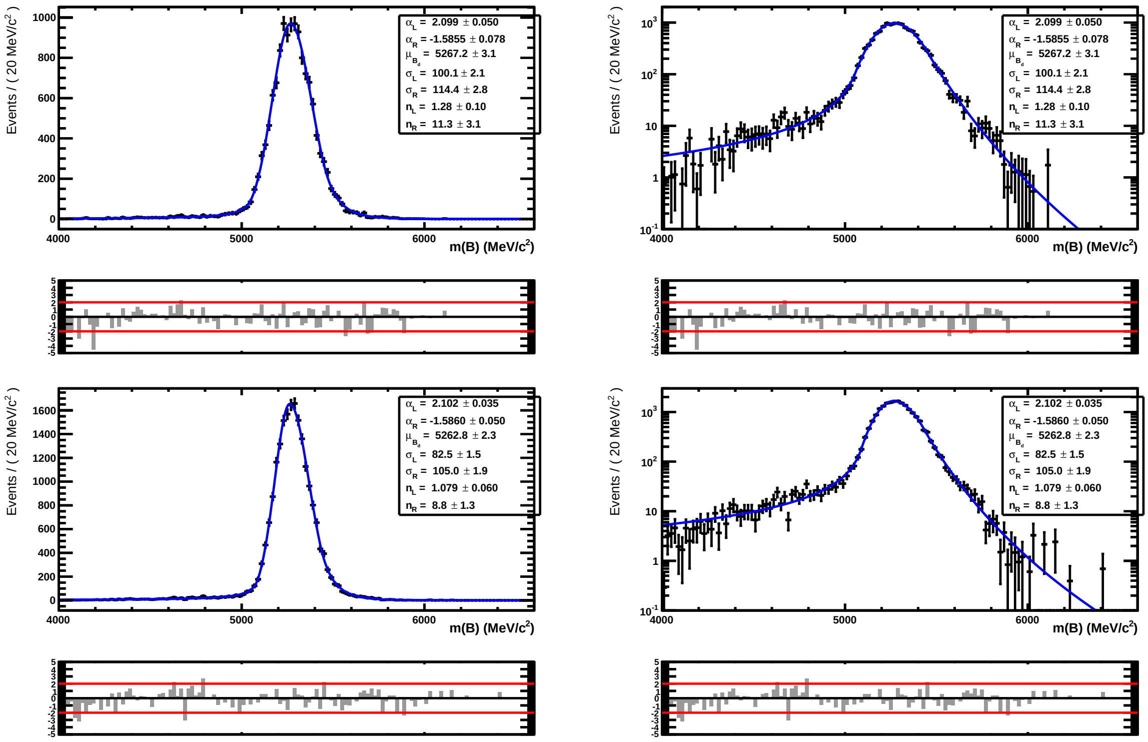


Figure 6.1: $B^0 \to \rho^0\gamma$ signal shapes for 2011 (top) and 2012 (bottom) MC samples in linear (left) and semi-logarithmic (right) scale.

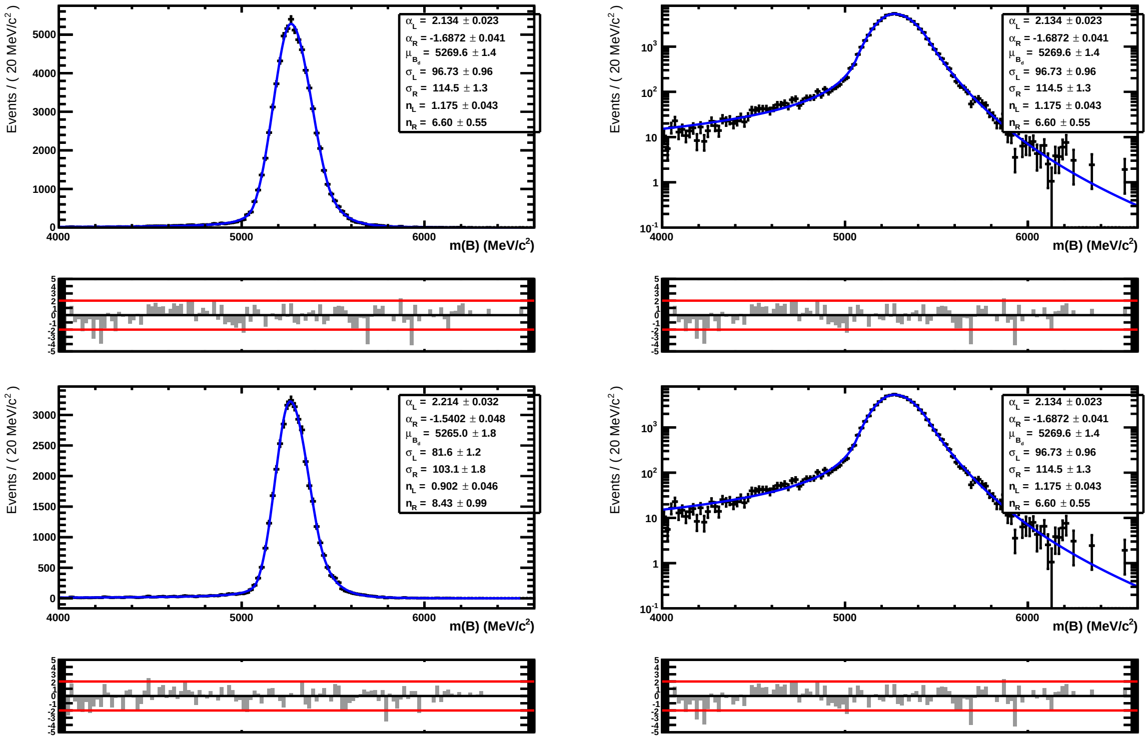


Figure 6.2: $B^0 \to K^{*0}\gamma$ signal shapes for 2011 (top) and 2012 (bottom) MC samples in linear (left) and semi-logarithmic (right) scale.

Fit parameters		Channel			
		$B^0 \to \rho^0\gamma$		$B^0 \to K^{*0}\gamma$	
		2011	2012	2011	2012
μ	MeV/c^2	5267.15 ± 3.14	5262.82 ± 2.25	5269.60 ± 1.42	5265.04 ± 1.82
σ_L	MeV/c^2	100.10 ± 2.14	82.52 ± 1.51	96.73 ± 0.96	81.58 ± 1.21
σ_R	MeV/c^2	114.40 ± 2.84	104.98 ± 1.93	114.52 ± 1.33	103.10 ± 1.77
α_L		2.099 ± 0.050	2.102 ± 0.035	2.134 ± 0.023	2.214 ± 0.032
α_R		-1.586 ± 0.078	-1.586 ± 0.050	-1.687 ± 0.041	-1.540 ± 0.048
n_L		1.280 ± 0.105	1.079 ± 0.060	1.175 ± 0.043	0.902 ± 0.046
n_R		11.29 ± 3.13	8.76 ± 1.35	6.60 ± 0.55	8.43 ± 0.99

Table 6.1: Summary table of the fitted parameters for $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ signal modes.

6.2 Background shapes

In this section we extract the shapes of the various peaking and partially reconstructed backgrounds to the $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ spectra.

All the $hh'\gamma$ and $hh'\pi^0$ peaking backgrounds are fitted in the full mass range [4000; 6600] MeV/c^2 and using a symmetric double tail Crystal-Ball PDF. The various cross-feed contaminations have already been calculated in Sec. 5.7 and will be used to constrain the respective yields in the final mass fit.

The partially reconstructed backgrounds are fitted in a limited mass range [4000; 5600] MeV/c^2 and using an ARGUS function [83] convoluted with a Gaussian PDF needed to model the detector resolution. The generalized ARGUS function has three parameters (m_t, c, p) and is given by

$$\mathrm{ARGUS}(m; m_t, c, p) = \frac{2^{-p}c^{2(p+1)}}{\Gamma(p+1) - \Gamma(p+1, c^2/2)} \cdot \frac{m}{m_t}\left(1 - \frac{m^2}{m_t^2}\right)^p \exp\left[-\frac{1}{2}c^2\left(1 - \frac{m^2}{m_t^2}\right)\right], \tag{6.5}$$

where $\Gamma(n)$ and $\Gamma(n, x)$ are respectively the usual Gamma function and incomplete Gamma function. The parameter m_t describes the threshold: if $m > m_t$ the function evaluates zero. The parameter p controls the curvature of the function and the parameter c controls the falling of the slope. The total function used to fit the partially reconstructed backgrounds is then defined as

$$\mathrm{Partial}(m; c, p, \Delta M, \sigma) = \mathrm{ARGUS}(m; c, \mu + \Delta M, p) \otimes \mathrm{Gauss}(0, \sigma), \tag{6.6}$$

where μ is fixed to the value of μ extracted in the fit to MC signal events, ΔM is the mass shift. The resolution of the Gaussian, σ, is fixed to the value of σ_R extracted in the fit to MC signal events. The threshold of the ARGUS in Eq. 6.5 is then $m_t = \mu + \Delta M$, where ΔM can be 0, $-m_{\pi^0}$, $-m_{\pi^+}$ or $-m_K$ according to the missing particles ($\Delta M = 0$ in case of one or more unreconstructed photons). As it will be shown in Sec. 6.2.2, many different modes with similar shapes compete in the same mass region of the $\rho\gamma$ and $K^*\gamma$ spectra. Additionally, the branching fractions of most of these decays are unknown. For this reason, apart from one single exception in the $B^0 \to K^{*0}\gamma$ reconstruction hypothesis, the contaminations of the various partially reconstructed backgrounds are not calculated. An intermediate shape between all the ones corresponding to the same ΔM is then taken as reference to extract the ARGUS shape parameters. Its relative contamination will be left free to float in the fit to the data.

6.2.1 Peaking backgrounds

The $hh'\gamma$ modes contributing to the $B^0 \to \rho^0\gamma$ mass spectrum are $B^0 \to K^{*0}\gamma$ and $\Lambda_b^0 \to \Lambda^*\gamma$, while the $hh'\pi^0$ ones are $B^0 \to \pi^+\pi^-\pi^0$ and $B^0 \to K^+\pi^-\pi^0$. The $hh'\gamma$ mass fits for 2011 and 2012 MC samples are shown in Fig. 6.3 and the $hh'\pi^0$ fits for 2012 MC samples are shown in Fig. 6.4. The summary of the fitted parameters is displayed in Tab. 6.2.

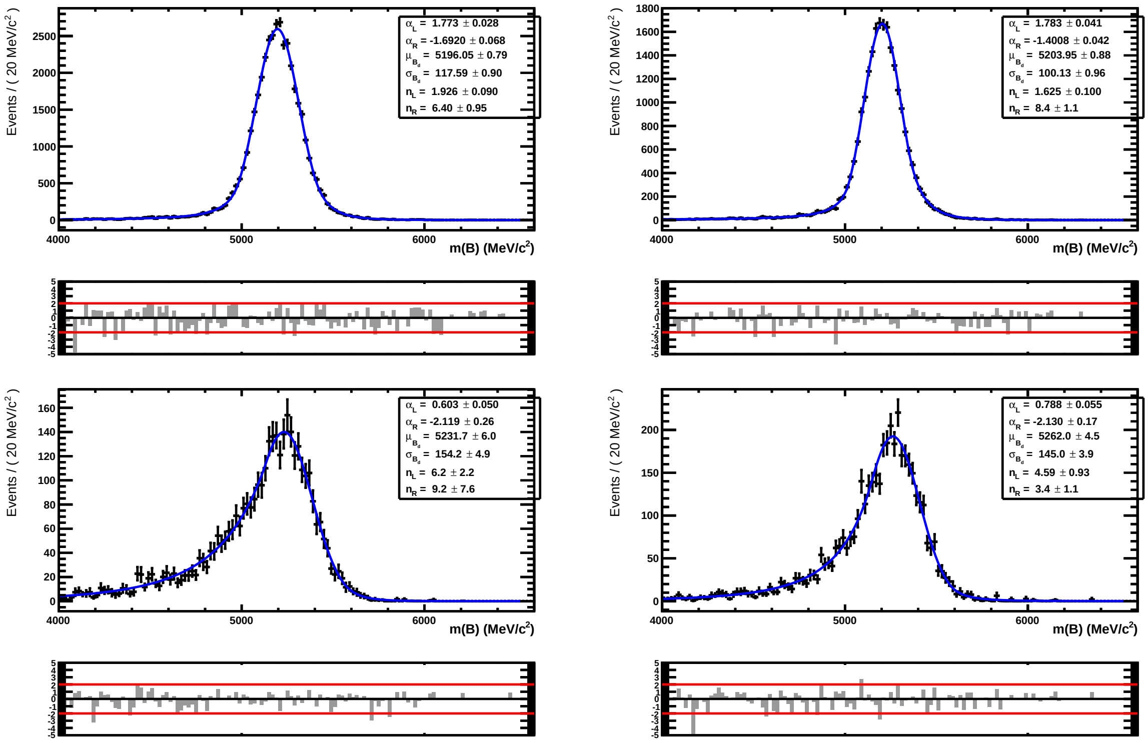


Figure 6.3: Fits to the invariant mass distributions of the $B^0 \to K^{*0}\gamma$ (top) and $\Lambda_b^0 \to \Lambda^*\gamma$ (bottom) modes reconstructed as $B^0 \to \rho^0\gamma$ for 2011 (left) and 2012 (right) MC samples.

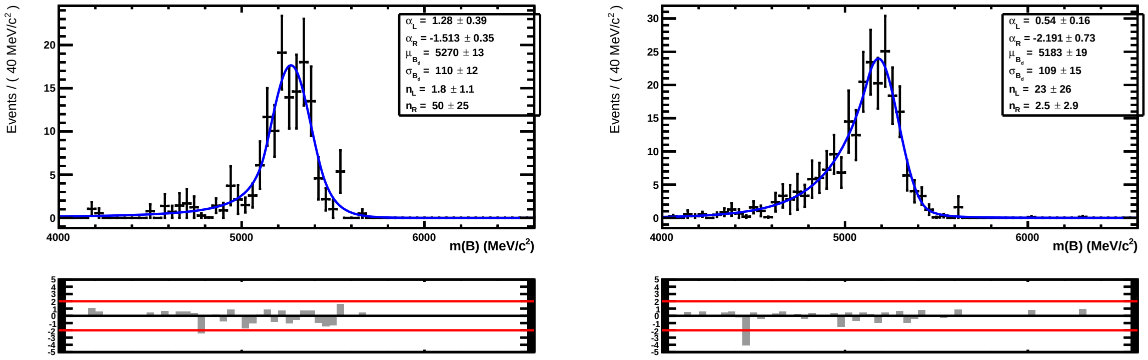


Figure 6.4: Fits to the invariant mass distributions of the $B^0 \to \pi^+\pi^-\pi^0$ (left) and $B^0 \to K^+\pi^-\pi^0$ (right) modes reconstructed as $B^0 \to \rho^0\gamma$ for 2012 MC samples.

Cross-feeds	Fit parameters						
	Year	μ (in MeV/c^2)	σ (in MeV/c^2)	α_L	n_L	α_R	n_R
$B^0 \to K^{*0}\gamma$	2011	5196.05 ± 0.79	117.59 ± 0.90	1.773 ± 0.028	1.926 ± 0.090	-1.692 ± 0.068	6.40 ± 0.95
	2012	5203.95 ± 0.88	100.12 ± 0.96	1.783 ± 0.041	1.625 ± 0.099	-1.401 ± 0.042	8.41 ± 1.12
$\Lambda_b^0 \to \Lambda^*\gamma$	2011	5231.70 ± 5.96	154.16 ± 4.86	0.603 ± 0.050	6.19 ± 2.23	-2.119 ± 0.260	9.19 ± 7.61
	2012	5262.00 ± 4.48	144.97 ± 3.94	0.788 ± 0.055	4.59 ± 1.11	-2.130 ± 0.169	3.36 ± 1.11
$B^0 \to \pi^+\pi^-\pi^0$	2012	5269.63 ± 13.06	109.51 ± 12.36	1.280 ± 0.393	1.81 ± 1.13	-1.513 ± 0.350	50.00 ± 25.19
$B^0 \to K^+\pi^-\pi^0$	2012	5183.01 ± 19.08	109.31 ± 15.14	0.539 ± 0.159	23.36 ± 25.81	-2.191 ± 0.729	2.51 ± 2.93

Table 6.2: Summary table of the fitted parameters for the $hh'\gamma$ and $hh'\pi^0$ peaking backgrounds in the $B^0 \to \rho^0\gamma$ channel.

The significant $hh'\gamma$ modes for the $B^0 \to K^{*0}\gamma$ mass spectrum are $B^0 \to \rho^0\gamma$, $B_s^0 \to \phi\gamma$ and $\Lambda_b^0 \to \Lambda^*\gamma$, while the $hh'\pi^0$ ones are $B^0 \to K^+\pi^-\pi^0$ and $B^0 \to \pi^+\pi^-\pi^0$. The $hh'\gamma$ mass fits for 2011 and 2012 MC samples are shown in Fig. 6.5 and the $hh'\pi^0$ fits for 2012 MC samples are shown in Fig. 6.6. The summary of the fitted parameters is displayed in Tab. 6.3.

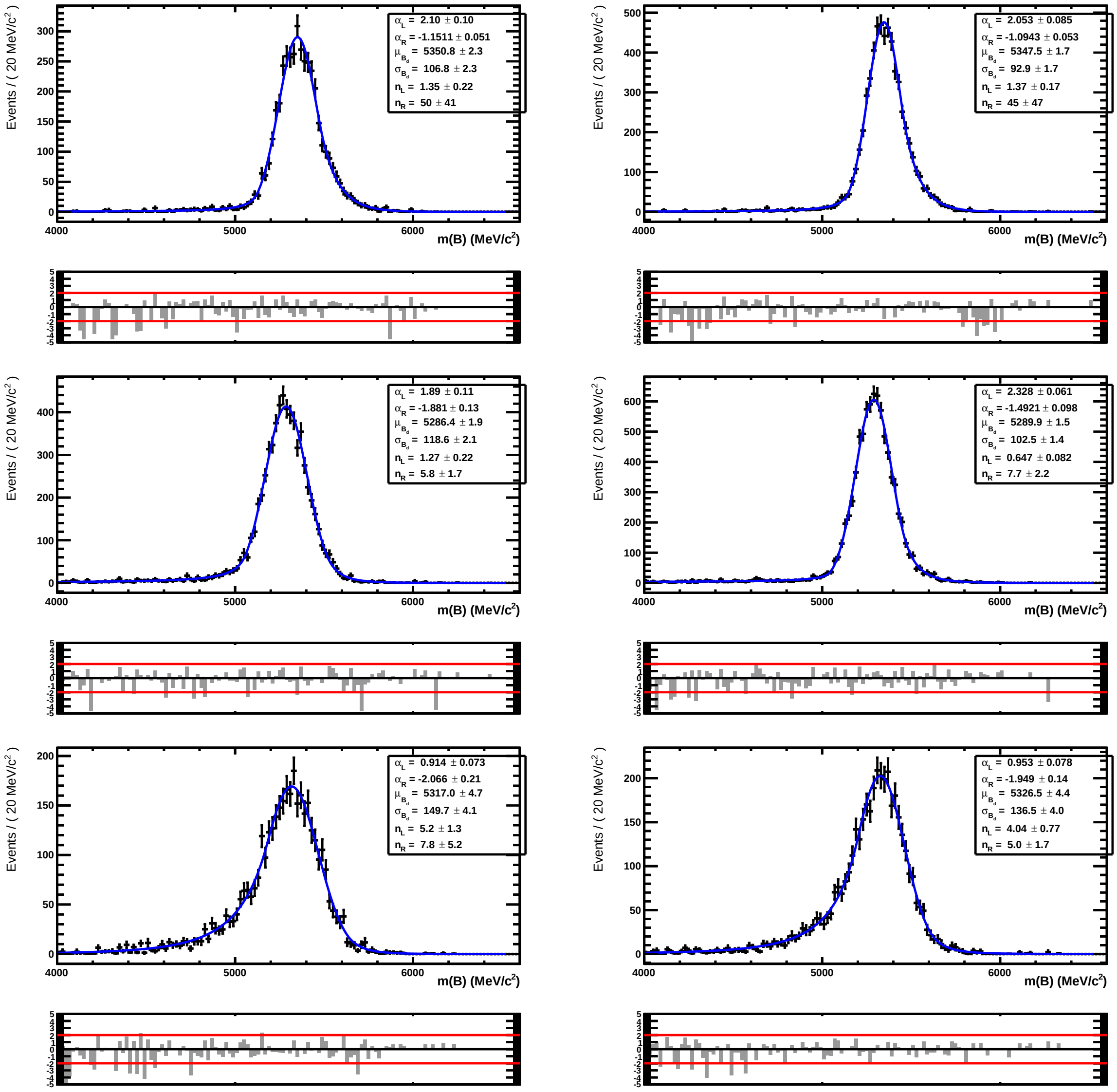


Figure 6.5: Fits to the invariant mass distributions of the $B^0 \to \rho^0\gamma$ (top), $B_s^0 \to \phi\gamma$ (middle) and $\Lambda_b^0 \to \Lambda^*\gamma$ (bottom) modes reconstructed as $B^0 \to K^{*0}\gamma$ for 2011 (left) and 2012 (right) MC samples.

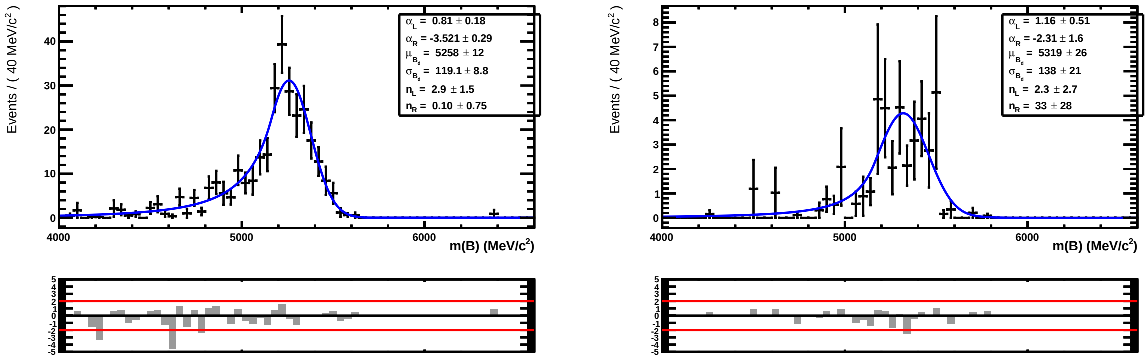


Figure 6.6: Fits to the invariant mass distributions of the $B^0 \to K^+\pi^-\pi^0$ (left) and $B^0 \to \pi^+\pi^-\pi^0$ (right) modes reconstructed as $B^0 \to K^{*0}\gamma$ for 2012 MC samples.

Cross-feeds	Fit parameters						
	Year	μ (in MeV/c^2)	σ (in MeV/c^2)	α_L	n_L	α_R	n_R
$B^0 \to \rho^0\gamma$	2011	5350.76 ± 2.32	106.83 ± 2.28	2.104 ± 0.102	1.352 ± 0.220	-1.151 ± 0.051	50.00 ± 41.06
	2012	5347.46 ± 1.69	92.93 ± 1.71	2.053 ± 0.085	1.375 ± 0.171	-1.094 ± 0.053	44.88 ± 46.98
$B_s^0 \to \phi\gamma$	2011	5286.37 ± 1.88	118.56 ± 2.12	1.890 ± 0.112	1.267 ± 0.216	-1.881 ± 0.133	5.83 ± 1.71
	2012	5289.87 ± 1.53	102.47 ± 1.41	2.328 ± 0.061	0.647 ± 0.082	-1.492 ± 0.098	7.72 ± 2.16
$\Lambda_b^0 \to \Lambda^*\gamma$	2011	5317.03 ± 4.67	149.73 ± 4.08	0.914 ± 0.073	5.169 ± 1.268	-2.066 ± 0.210	7.81 ± 5.15
	2012	5326.46 ± 4.43	136.55 ± 4.01	0.953 ± 0.078	4.037 ± 0.078	-1.950 ± 0.142	5.02 ± 1.69
$B^0 \to K^+\pi^-\pi^0$	2012	5258.44 ± 11.79	119.08 ± 8.81	0.815 ± 0.183	2.91 ± 1.48	-3.521 ± 0.294	0.10 ± 0.75
$B^0 \to \pi^+\pi^-\pi^0$	2012	5319.26 ± 25.88	138.25 ± 20.95	1.161 ± 0.513	2.35 ± 2.71	-2.312 ± 1.626	33.10 ± 27.59

Table 6.3: Summary of the fitted parameters for the $hh'\gamma$ and $hh'\pi^0$ peaking backgrounds to the $B^0 \to K^{*0}\gamma$ channel.

6.2.2 Partially reconstructed backgrounds

Backgrounds to $B^0 \to \rho^0\gamma$

Partially reconstructed backgrounds to the $B^0 \to \rho^0\gamma$ channel could be any B decays leading to $\pi^+\pi^-\gamma X$ or $\pi^+\pi^-\pi^0 X$ final states where X could be one or more not reconstructed charged or neutral particle. Based on the additional particle not reconstructed in the final state, we can group the various backgrounds into the following general categories

- $B \to \pi^+\pi^-\pi\gamma$ decays which are reconstructed with a missing pion ($\pi^\pm$ or π^0);
- $B \to K\pi^+\pi^-\gamma$ decays which are reconstructed with a missing kaon;
- $B \to \pi^+\pi^-\pi^0 X$ decays in which the π^0 is reconstructed as a photon.

The decays contributing to the first category are $B^0 \to \omega(\pi^+\pi^-\pi^0)\gamma$, $B^+ \to a_1(1270)^+(\pi^+\pi^-\pi^+)\gamma$ and $B^+ \to K^{*+}(K^0_S\pi^+)\gamma$. The comparison of the respective mass distributions is displayed on the left side of Fig. 6.7. The first two channels have very similar kinematics and shapes, while the last one exhibits a different mass distribution due to the presence of the intermediate K^0_S which decays into two pions. If needed, it is foreseen to apply dedicated cuts to veto those $K^0_S \to \pi^+\pi^-$ decays. Since the $a_1(1270)^+\gamma$ mode is still unobserved, we take as reference channel for this category the $B^0 \to \omega\gamma$ mode. When fitting this mode we fix ΔM to -134.977 MeV/c^2, which corresponds to the π^0 mass.

In the $B \to K\pi^+\pi^-\gamma$ category there are many competing modes. The MC samples we have are $B^+ \to K^*(1410)^+\gamma$, $B^+ \to K_1(1270)^+\gamma$, $B^+ \to K_1(1400)^+\gamma$, $B^+ \to K_2^*(1430)^+\gamma$, $B^+ \to K^+\pi^-\pi^+\gamma$ and $B^+ \to K^{*0}\pi^+\gamma$. The comparison of the respective mass distributions is displayed on the right side of Fig. 6.7. As can be seen, even with slight differences, all the channels give compatible shapes as the kinematics are similar. The MC sample that we take as reference to extract this shape is the non-resonant $B^+ \to K^+\pi^-\pi^+\gamma$ mode: its mass distribution is intermediate compared to the others and the available statistics is

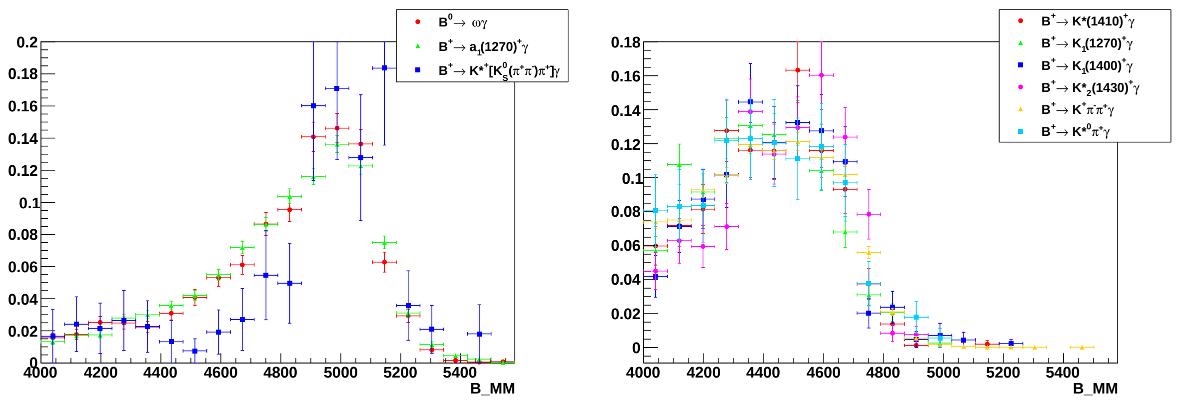


Figure 6.7: Invariant mass distributions of $B \to \pi^+\pi^-\pi\gamma$ (left) and $B \to K\pi^+\pi^-\gamma$ (right) partially reconstructed decays.

the highest one. When performing the fit on this mode we use $\Delta M = -493.677$ MeV/c^2, which corresponds to the charged kaon mass.

The decays that belong to the $B \to \pi^+\pi^-\pi^0 X$ category have considerably different kinematics and then different mass distributions. The MC samples we have are $B^+ \to \rho^+\rho^0$, $B^+ \to K^+\pi^-\pi^+\pi^0$, $B^0 \to K^{*0}(K^0_S\pi^0)\rho^0$ and $B^+ \to D^0(K^0_S\pi^+\pi^-)\rho^+$. The comparison of the respective mass distributions is displayed on the left side of Fig. 6.8. Concerning the two first decays, these can be discarded because the $B^+ \to \rho^+\rho^0$ mode has a shape which is very similar to the $B^0 \to \omega\gamma$ one and the $B^+ \to K^+\pi^-\pi^+\pi^0$ channel has a mass distribution which is compatible with the one of the $B^+ \to K^+\pi^-\pi^+\gamma$ background. The decay modes $B^+ \to D^0\rho^+$ and $B^0 \to K^{*0}\rho^0$ that populate the low mass region have similar mass distributions but, given the higher statistics, we take as reference the second one. Actually the $B^0 \to K^{*0}\rho^0$ mode competes with another channel for which we have not defined a dedicated category, namely $B^+ \to K_1(1270)^+\eta(\gamma\gamma)$ (see right side of Fig. 6.8). However, since this mode is still unobserved and $B^0 \to K^{*0}(K^0_S\pi^0)\rho^0$ guarantees a better statistics, we discard it. When fitting this mode we fix the value of the ARGUS mass shift to $\Delta M = -497.611$ MeV/c^2, which corresponds to the K^0 mass.

To summarize, the three generic reference shapes used to describe the partially reconstructed backgrounds to the $B^0 \to \rho^0\gamma$ signal and the associated channels that we fitted are shown in Fig. 6.9. The mass fits to the invariant mass distributions of the reference modes are shown in Fig. 6.10, while the summary of all the parameters of the fits are listed in Tab. 6.4.

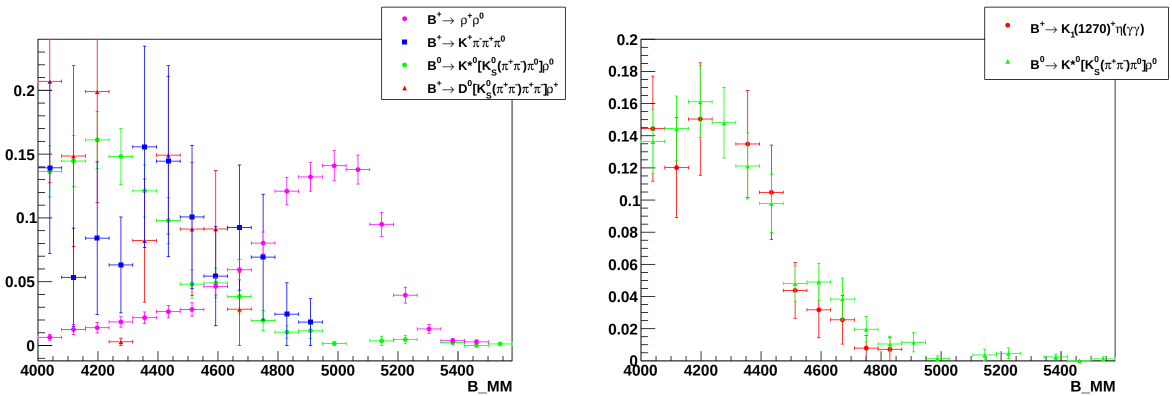


Figure 6.8: Invariant mass distributions of $B \to \pi^+\pi^-\pi^0 X$ partially reconstructed decays (left) and comparison of $B^0 \to K^{*0}\rho^0$ and $B^+ \to K_1(1270)^+\eta$ modes (right).

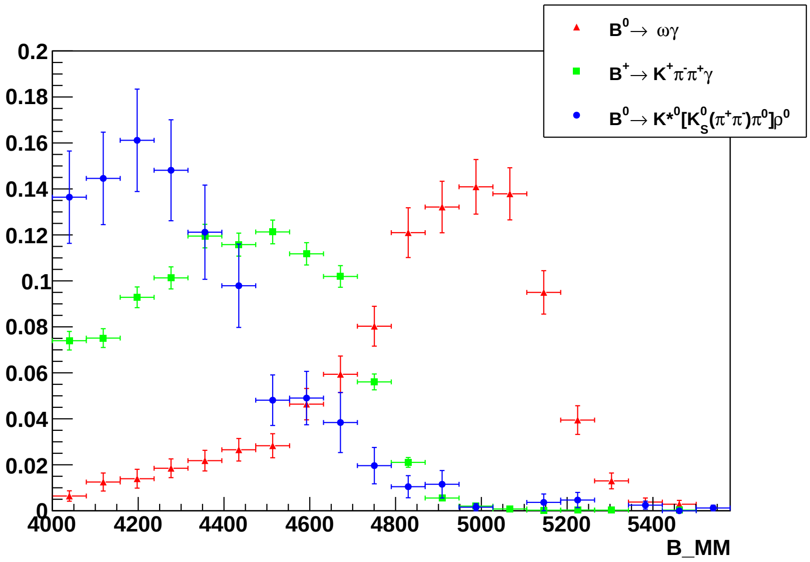


Figure 6.9: Reference shapes and respective channels used to describe the partially reconstructed background contributions to the $B^0 \to \rho^0\gamma$ mass spectrum.

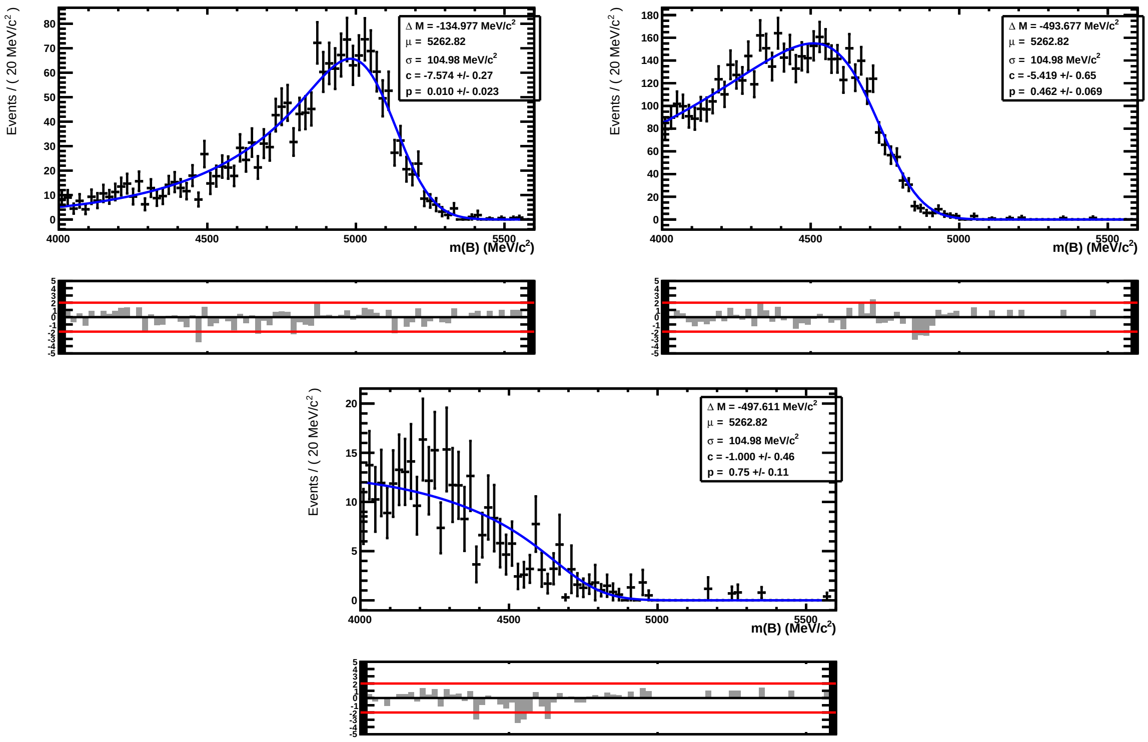


Figure 6.10: Fits to the invariant mass distributions of the $B^0 \to \omega\gamma$ (top left), $B^+ \to K^+\pi^-\pi^+\gamma$ (top right) and $B^0 \to K^{*0}\rho^0$ (bottom) decays reconstructed as $B^0 \to \rho^0\gamma$.

Channel	Year	ΔM (in MeV/c^2)	Fit parameters μ (in MeV/c^2)	σ (in MeV/c^2)	c	p
$B^0 \to \omega\gamma$	2012	-134.977	5262.82	104.98	-7.574 ± 0.271	0.010 ± 0.023
$B^+ \to K^+\pi^-\pi^+\gamma$	2012	-493.677	5262.82	104.98	-5.419 ± 0.650	0.462 ± 0.069
$B^0 \to K^{*0}(K^0_S\pi^0)\rho^0$	2012	-497.611	5262.82	104.98	-1.000 ± 0.460	0.749 ± 0.110

Table 6.4: Summary of the fit parameters for the partially reconstructed modes in the $B^0 \to \rho^0\gamma$ channel. The μ and σ parameters come from the fit of the signal 2012 MC sample.

Backgrounds to $B^0 \to K^{*0}\gamma$

Partially reconstructed backgrounds to the $B^0 \to K^{*0}\gamma$ channel could be any B decays leading to $K^+\pi^-\gamma X$ or $K^+\pi^-\pi^0 X$ final states where X could be one or more not reconstructed charged or neutral particle. B decays with a misidentified kaon are highly suppressed thanks to the tight PID selection. Based on the additional particle not reconstructed, we can group the various backgrounds into the following general categories

- $B \to K^+\pi^-\pi\gamma$ decays which are reconstructed with a missing pion;
- $B \to K^+\pi^-\pi^0 X$ in which the π^0 is reconstructed as a photon;
- Decays with η in the final state.

There are many competing modes that contribute to the first category. The MC samples we have are $B^+ \to K^*(1410)^+\gamma$, $B^+ \to K_1(1270)^+\gamma$, $B^+ \to K_1(1400)^+\gamma$, $B^+ \to K_2^*(1430)^+\gamma$, $B^+ \to K^+\pi^-\pi^+\gamma$ and $B^+ \to K^{*0}\pi^+\gamma$. The comparison of the corresponding mass distributions is displayed on the top left side of Fig. 6.11. Apart from the non-resonant $B^+ \to K^+\pi^-\pi^+\gamma$ channel, even though some differences due to various intermediate resonances can be spotted, the shapes of the other modes are very similar. The MC sample that we take as reference for this shape is the $B^+ \to K_1(1270)^+\gamma$ mode: its mass distribution is intermediate compared to the other ones and the available statistics is better. When performing the fit we fix ΔM to -139.570 MeV/c^2, which is the value of the $\pi^\pm$ mass.

Only two MC samples enter the $B \to K^+\pi^-\pi^0 X$ category, namely $B^0 \to K_1(1270)^0\gamma$ and $B^+ \to K^+\pi^-\pi^+\pi^0$. The associated shapes (see top right side of Fig. 6.11) are similar to the ones in the $B \to K^+\pi^-\pi\gamma$ category, therefore no additional reference shape is considered for this category.

Decays involving a η meson in the final state have considerably different kinematics. The MC samples we have are $B^0 \to K^{*0}\eta(\gamma\gamma)$, $B^+ \to K_1(1270)^+\eta(\gamma\gamma)$ and $B^0 \to K^{*0}\eta(\pi^+\pi^-\pi^0)$. The comparison of their mass distributions is shown on the bottom of Fig. 6.11. As can be seen the various distributions are significantly different, then for all these modes we extract a dedicated shape. The $B^+ \to K_1(1270)^+\eta$ is still unobserved but accounts for any partially reconstructed decay involving one missing pion and one missing photon from the $\eta \to \gamma\gamma$

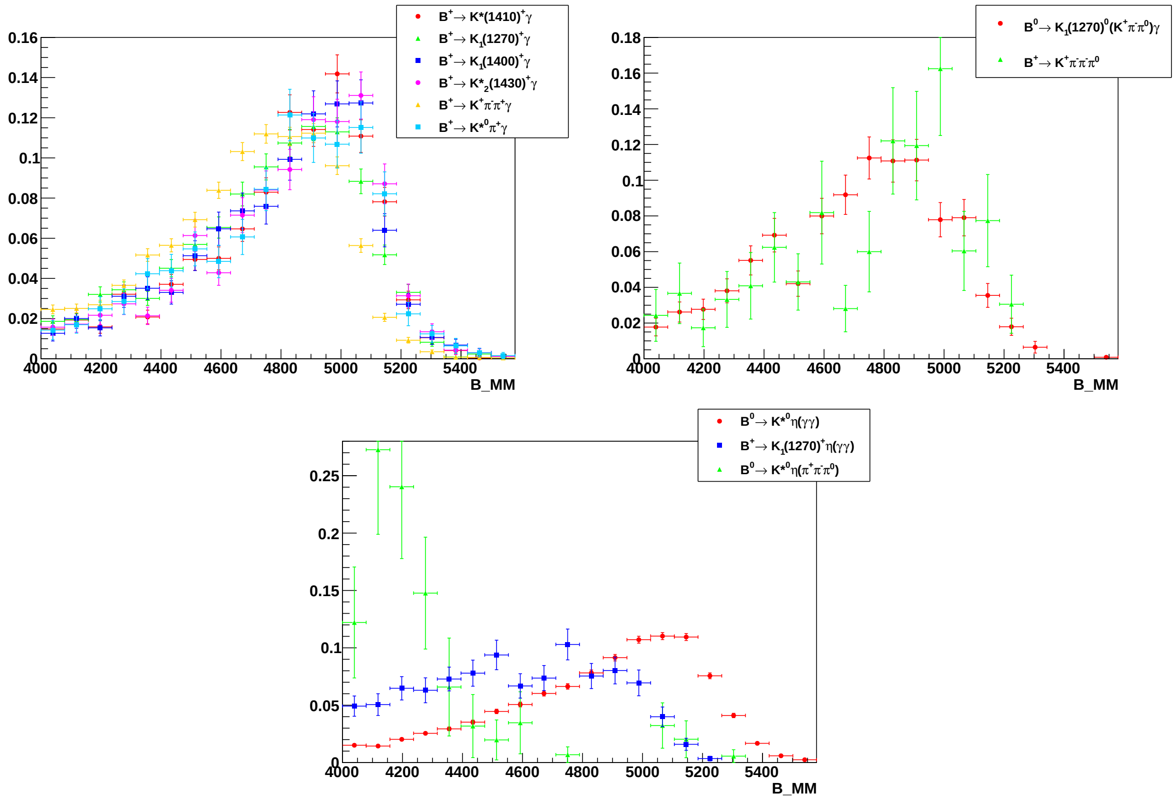


Figure 6.11: Invariant mass distributions of $B \to K^+\pi^-\pi\gamma$ (top left) and $B \to K^+\pi^-\pi^0 X$ (top right) partially reconstructed backgrounds. The shapes associated to decays with a η meson in the final state are shown on the bottom.

decay. For what concerns the $B^0 \to K^{*0}\eta(\gamma\gamma)$ channel, which has a specif shape, we can fix its relative contamination to the $B^0 \to K^{*0}\gamma$ mass spectrum using the effective branching fraction given by $\mathcal{B}(B^0 \to K^{*0}\eta) = (4.17 \pm 0.26) \times 10^{-6}$. The contamination is computed according to what has been done in Sec. 5.7 for peaking backgrounds. The resulting contaminations for 2011 and 2012, together with all the parameters used to extract them, are displayed in Tab 6.5. Let us notice that the MC generation efficiency used for both years is the one of 2012 since the corresponding 2011 generator statistics tables are not available. This does not have a significant impact on the final value of the contamination since 2011 and 2012 generation efficiencies are very close. We can also notice that the average $K^*\eta$ PID efficiency is about 66%, while the $K^*\gamma$ one is about 71% (see Tab. 5.19). This difference comes from the fact that the kinematical distributions of the charged tracks for a partially reconstructed background are different from those of the signal. In principle it would be possible to fix the contamination also for the $B^0 \to K^{*0}\eta(\pi^+\pi^-\pi^0)$ mode because the effective branching fraction is known. However, this background populates the low invariant mass region where very few signal events are expected and its shape can suitably describe other generic partially reconstructed backgrounds that could contribute to the very left part of the B mass spectrum. Then we do not compute this contamination.

Channel	Year	ϵ^{MC}	ϵ^{Sel}	ϵ^{PID}	$r^{Eff}_{K^*\eta}$	$\mathcal{C}^{K^*\gamma}_{K^*\eta}$
		(%)	(%)	(%)		(%)
$B^0 \to K^{*0}\eta(\gamma\gamma)$	2011	16.640 ± 0.037	0.282 ± 0.003	66.480 ± 0.228	0.145 ± 0.010	3.699 ± 0.268
	2012		0.232 ± 0.003	65.700 ± 0.192		3.646 ± 0.265

Table 6.5: Summary table of the individual efficiencies and $r^{Eff}_{K^*\eta}$ used to derive the $B^0 \to K^{*0}\eta(\gamma\gamma)$ relative contaminations $\mathcal{C}^{K^*\gamma}_{K^*\eta}$ in the $B^0 \to K^{*0}\gamma$ signal hypothesis.

When performing the fits to the MC samples we fix ΔM to 0 for $B^0 \to K^{*0}\eta(\gamma\gamma)$ since we have a missing photon, -139.570 MeV/c^2 for $B^+ \to K_1(1270)^+\eta(\gamma\gamma)$ since we have a missing $\pi^\pm$ and a missing photon and -277.140 MeV/c^2 for $B^0 \to K^{*0}\eta(\pi^+\pi^-\pi^0)$ which stands for the two missing $\pi^\pm$ mass.

To summarize, the four generic reference shapes used to describe the partially reconstructed backgrounds to the $B^0 \to K^{*0}\gamma$ signal and the associated channels that we fitted are shown in Fig. 6.12. The fits to the invariant mass distributions of the reference modes are shown in Fig. 6.13, while the summary of all the parameters of the fits are listed in Tab. 6.6.

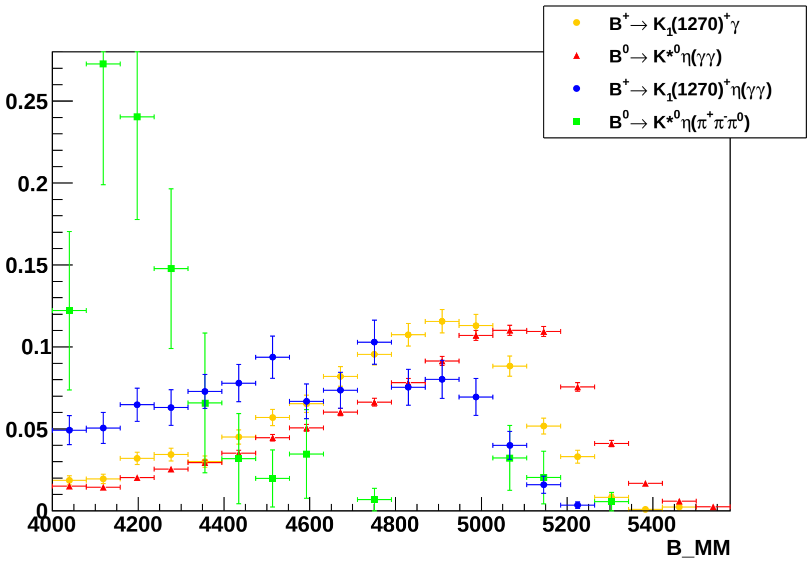


Figure 6.12: Reference shapes and respective channels used to describe the partially reconstructed background contributions to the $B^0 \to K^{*0}\gamma$ mass spectrum.

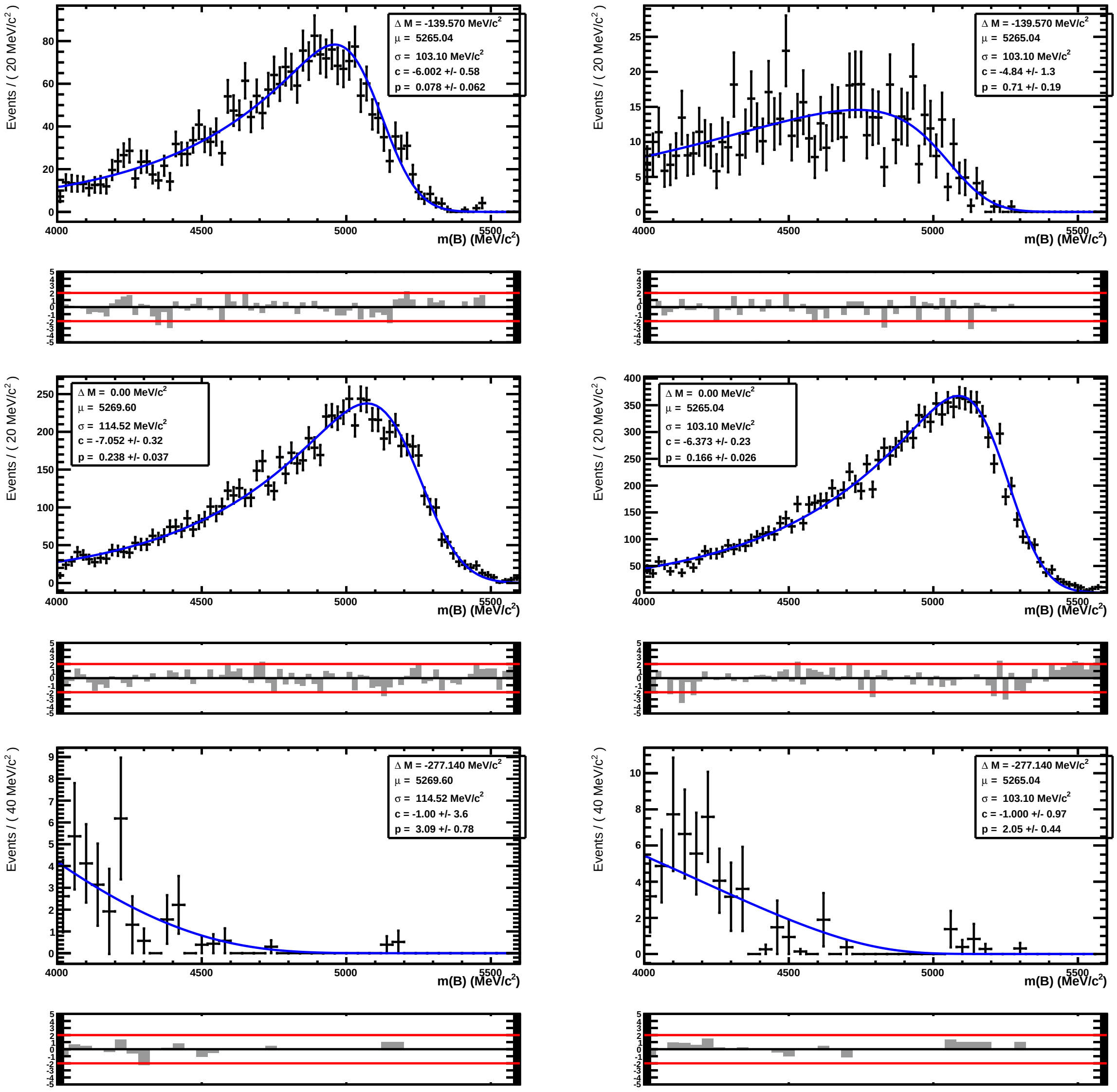


Figure 6.13: On the top we show the 2012 fits to the invariant mass distributions of the $B^+ \to K_1(1270)^+\gamma$ (left) and $B^+ \to K_1(1270)^+\eta(\gamma\gamma)$ (right) modes. The 2011 (left) and 2012 (right) fits to the $B^0 \to K^{*0}\eta(\gamma\gamma)$ (middle) and $B^0 \to K^{*0}\eta(\pi^+\pi^-\pi^0)$ (bottom) channels are also displayed. All the MC samples are reconstructed as $B^0 \to K^{*0}\gamma$.

Channel	Year	Fit parameters				
		ΔM (in MeV/c^2)	μ (in MeV/c^2)	σ (in MeV/c^2)	c	p
$B^+ \to K_1(1270)^+\gamma$	2012	-139.570	5265.04	103.10	-6.002 ± 0.584	0.078 ± 0.062
$B^0 \to K^{*0}\eta(\gamma\gamma)$	2011	0	5269.60	114.52	-7.052 ± 0.315	0.238 ± 0.037
	2012	0	5265.04	103.10	-6.373 ± 0.232	0.166 ± 0.026
$B^+ \to K_1(1270)^+\eta(\gamma\gamma)$	2012	-139.570	5265.04	103.10	-4.843 ± 1.272	0.708 ± 0.189
$B^0 \to K^{*0}\eta(\pi^+\pi^-\pi^0)$	2011	-279.140	5269.60	114.52	-1.000 ± 3.563	3.087 ± 0.776
	2012	-279.140	5265.04	103.10	-1.000 ± 0.974	2.049 ± 0.439

Table 6.6: Summary of the fit parameters for the partially reconstructed modes in the $B^0 \to K^{*0}\gamma$ channel. The μ and σ parameters come from the fits on the signal 2011 and 2012 MC sample.

6.3 Fit to the $B^0 \to K^{*0}\gamma$ mass distribution in data

Once the shapes and expected background contaminations are determined, we perform a mass fit to the $B^0 \to K^{*0}\gamma$ spectrum as a preliminary check to test the fit model before building up the final simultaneous fit model. The total function used to fit the data is built as an extended PDF given by

$$\mathcal{F}(m) = N_S \cdot S(m) + \sum_k N_k \cdot P_k(m) + \sum_i N_i \cdot B_i(m) \tag{6.7}$$

where $S(m)$ is the signal PDF, $P_k(m)$ is the peaking background PDF where k stands for the different species and $B_i(m)$ is the partially reconstructed or combinatorial background PDF where i stands for the different background contributions. N_S is the total number of signal event to be extracted in the fit and N_k is the number of peaking background events constrained to their relative contamination with respect to the signal as $N_k = \mathcal{C}_k^{K^*\gamma} N_S$. The relative contamination coming from the $\rho\gamma$ channel is fixed in the fit assuming for $r_{\rho\gamma}^{Eff}$ the value of the PDG (see Tab. 5.21). N_i is the number of events from either the combinatorial background or one of the partially reconstructed backgrounds. Let us remind that for the $B^0 \to K^*\eta(\gamma\gamma)$ partially reconstructed background N_i is fixed to $\mathcal{C}_{K^*\eta}^{K^*\gamma} N_S$ as for peaking backgrounds. All the other N_i are left free.

We perform an unbinned extended maximum likelihood fit according to the following prescriptions:

- The signal is modeled with an asymmetric double-tail Crystal Ball PDF (see Sec. 6.1) with the μ, σ_L and σ_R parameters free to float.
- The peaking backgrounds are modeled with a double-tail Crystal Ball PDF (see

Sec. 6.2.1) with the μ and σ of every channel constrained according to:

$$\mu_{Bkg}^{DATA} = \mu_{Bkg}^{MC} + (\mu_{Sig}^{DATA} - \mu_{Sig}^{MC}), \quad \sigma_{Bkg}^{DATA} = \sigma_{Bkg}^{MC} \times \frac{\sigma_{Sig}^{DATA}}{\sigma_{Sig}^{MC}} \tag{6.8}$$

$$\sigma_{Sig} = \frac{1}{2}\sqrt{\sigma_L^2 + \sigma_R^2} \tag{6.9}$$

where Bkg refers to the peaking background species and σ_{Sig} is the combination of σ_L and σ_R defined for both data and MC.

- The combinatorial background shape is modeled with a first order polynomial

$$\mathrm{Comb}(m; p_0) = 1 + p_0 \cdot m \ . \tag{6.10}$$

- The partially reconstructed backgrounds are modeled with an ARGUS function convoluted with a Gaussian PDF (see Sec. 6.2.2), the μ and σ parameters are set to be the μ and σ_R of the signal PDF.

The fits to the $B^0 \to K^{*0}\gamma$ full mass spectrum for 2011 ($0.7\,\mathrm{fb}^{-1}$) and 2012 ($2\,\mathrm{fb}^{-1}$) are shown respectively in Fig. 6.14 and 6.15, while the values of the fit parameters, including the fixed contaminations, are shown in Tab. 6.7.

Channel	Parameter		Year 2011	Year 2012
$B^0 \to K^{*0}\gamma$	N_S		3446 ± 63	11854 ± 119
	μ	MeV/c^2	5270.9 ± 4.8	5281.8 ± 2.4
	σ_L	MeV/c^2	90.5 ± 3.8	90.9 ± 2.2
	σ_R	MeV/c^2	108.8 ± 3.7	83.5 ± 1.6
$B^0 \to \rho^0\gamma$	$\mathcal{C}$	%	0.194	0.209
$B_s^0 \to \phi\gamma$	$\mathcal{C}$	%	0.038	0.044
$\Lambda_b^0 \to \Lambda^*\gamma$	$\mathcal{C}$	%	0.392	0.506
$B^0 \to K^+\pi^-\pi^0$	$\mathcal{C}$	%	1.674	1.674
$B^0 \to \pi^+\pi^-\pi^0$	$\mathcal{C}$	%	0.061	0.061
Combinatorial	N_{Comb}		59 ± 60	411 ± 80
	p_0	$(\mathrm{MeV}/c^2)^{-1}$	-0.33 ± 0.79	-0.96 ± 0.06
$B^0 \to K^{*0}\eta(\gamma\gamma)$	$\mathcal{C}$	%	3.699	3.646
$B^+ \to K_1(1270)^+\gamma$	N_{Part}		0 ± 33	0 ± 50
$B^+ \to K_1(1270)^+\eta(\gamma\gamma)$	N_{Part}		2727 ± 90	7180 ± 276
$B^0 \to K^{*0}\eta(\pi^+\pi^-\pi^0)$	N_{Part}		1115 ± 62	5218 ± 145

Table 6.7: Summary of the parameters entering the fit model and values extracted fitting the $B^0 \to K^{*0}\gamma$ mass spectrum separately for 2011 and 2012 data.

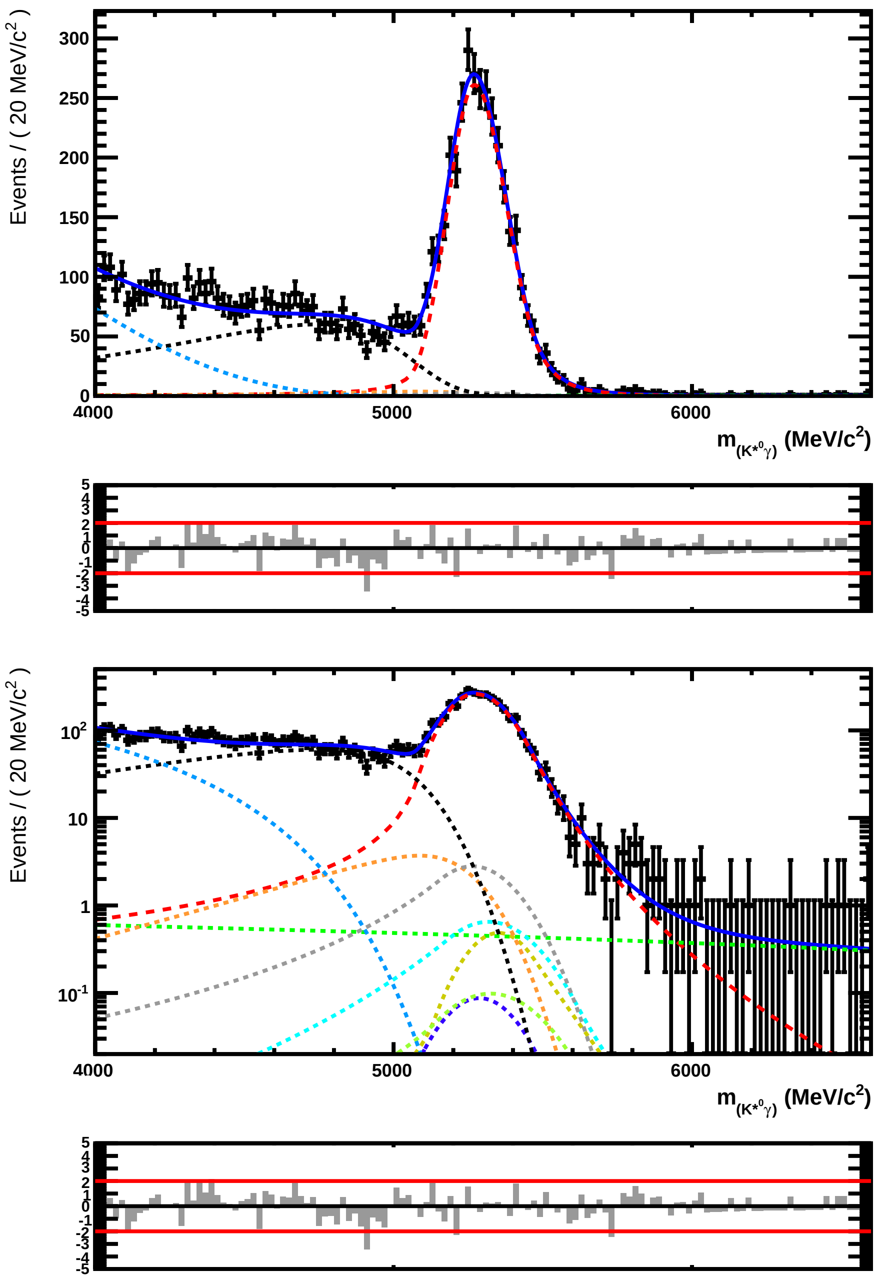


Figure 6.14: 2011 $B^0 \to K^{*0}\gamma$ full mass fit in linear (top) and semi-logarithmic scale (bottom). The black points represent the data and the solid blue line is the total fit function. The signal is fitted with an asymmetric double-tail Crystal Ball PDF (dashed red line) and the combinatorial background is described with a first order polynomial (dashed green line). The peaking background modes are $B^0 \to \rho^0\gamma$ (dashed dark yellow line), $B_s^0 \to \phi\gamma$ (dashed dark blue line), $\Lambda_b^0 \to \Lambda^*\gamma$ (dashed cyan line), $B^0 \to K^+\pi^-\pi^0$ (dashed gray line) and $B^0 \to \pi^+\pi^-\pi^0$ (dashed light green line). The partially reconstructed modes are $B^0 \to K^{*0}\eta(\gamma\gamma)$ (dashed orange line), $B^+ \to K_1(1270)^+\eta(\gamma\gamma)$ (dashed black line) and $B^0 \to K^{*0}\eta(\pi^+\pi^-\pi^0)$ (dashed light blue line).

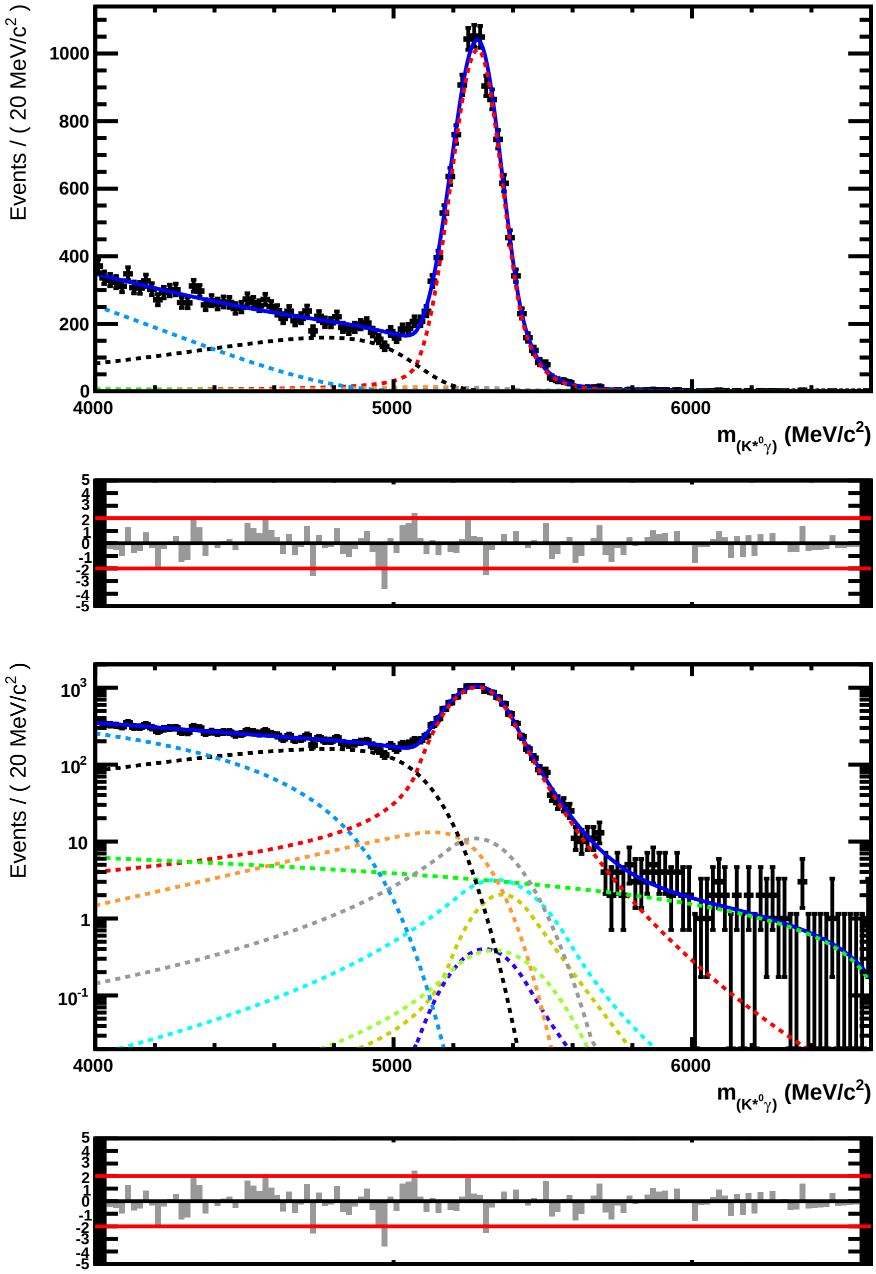


Figure 6.15: 2012 $B^0 \to K^{*0}\gamma$ full mass fit in linear (top) and semi-logarithmic scale (bottom). The black points represent the data and the solid blue line is the total fit function. The signal is fitted with an asymmetric double-tail Crystal Ball PDF (dashed red line) and the combinatorial background is described with a first order polynomial (dashed green line). The peaking background modes are $B^0 \to \rho^0\gamma$ (dashed dark yellow line), $B_s^0 \to \phi\gamma$ (dashed dark blue line), $\Lambda_b^0 \to \Lambda^*\gamma$ (dashed cyan line), $B^0 \to K^+\pi^-\pi^0$ (dashed gray line) and $B^0 \to \pi^+\pi^-\pi^0$ (dashed light green line). The partially reconstructed modes are $B^0 \to K^{*0}\eta(\gamma\gamma)$ (dashed orange line), $B^+ \to K_1(1270)^+\eta(\gamma\gamma)$ (dashed black line) and $B^0 \to K^{*0}\eta(\pi^+\pi^-\pi^0)$ (dashed light blue line).

As we can see from the pulls under the mass plots, the agreement between the data and the total fit function is very good. However, in both 2011 and 2012 the fitted partially reconstructed background yields are not coherent with the number of reference shapes identified on MC. In particular, the fits give 0 events for the shape associated to the $B^+ \to K_1(1270)^+\gamma$ mode. Even though we tried to keep the number of partially reconstructed background shapes to a minimum when fitting the data, too many contributions are competing in the low mass region. Selecting which shape to use or not would arbitrarily affect the fitted signal yields leading to systematic uncertainties related to the fit model. Those would add to the ones associated to the choices of specific modes to extract the various reference shapes. One can see on Fig. 6.14 and 6.15 that the contribution associated to the $B^0 \to K^{*0}\eta(\pi^+\pi^-\pi^0)$ shape is negligible under the signal peak: evaluating this contribution is not really mandatory for the extraction of the signal yields. A way to simplify the fit is then to ignore this contribution fitting the data in a limited mass range where it can be neglected.

For this reason we decided to restrict the fit range to the [4800; 6600] MeV/c^2 mass region and to define a generic $K^+\pi^-\gamma X$ contribution that accounts for any partially reconstructed background with at least one missing meson. This background is modeled setting ΔM to -134.977 MeV/c^2, corresponding to the π^0 mass, and fixing $p = 0.5$ in the generalized ARGUS of Eq. 6.5, which then becomes a regular ARGUS with c as free parameter. The resolution of the convoluted Gaussian is always set to be the σ_R of the signal PDF.

The fits to the $B^0 \to K^{*0}\gamma$ reduced mass spectrum for 2011 ($0.7\,\text{fb}^{-1}$) and 2012 ($2\,\text{fb}^{-1}$) are shown respectively in Fig. 6.16 and 6.17, while the values of the fit parameters, including the fixed contaminations, are shown in Tab. 6.8.

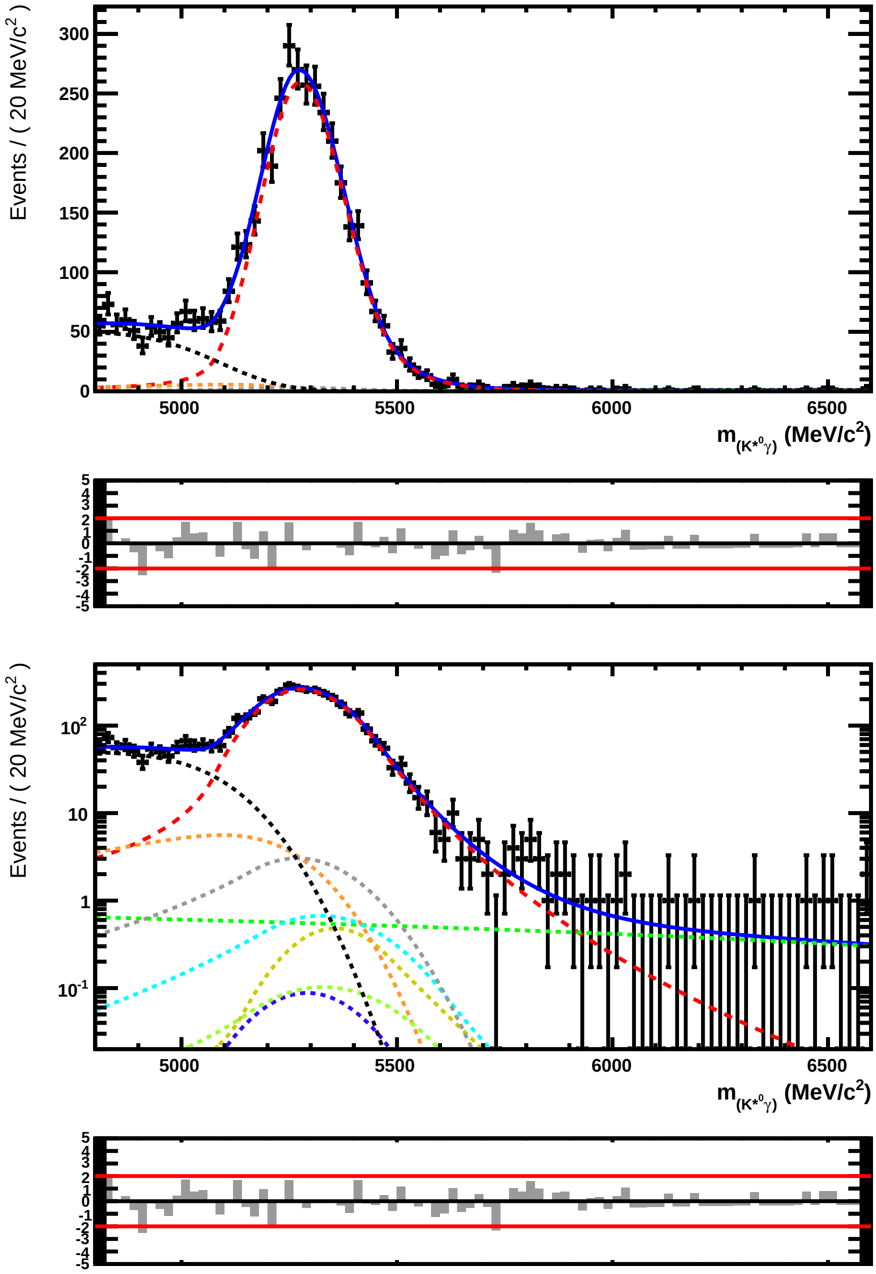


Figure 6.16: 2011 $B^0 \to K^{*0}\gamma$ mass fit in linear (top) and semi-logarithmic scale (bottom); the fit is performed in the restricted range [4800; 6600] MeV/c^2. The black points represent the data and the solid blue line is the total fit function. The signal is fitted with an asymmetric double-tail Crystal Ball PDF (dashed red line) and the combinatorial background is described with a first order polynomial (dashed green line). The peaking background modes are $B^0 \to \rho^0\gamma$ (dashed dark yellow line), $B_s^0 \to \phi\gamma$ (dashed dark blue line), $\Lambda_b^0 \to \Lambda^*\gamma$ (dashed cyan line), $B^0 \to K^+\pi^-\pi^0$ (dashed gray line) and $B^0 \to \pi^+\pi^-\pi^0$ (dashed light green line). The partially reconstructed modes are $B^0 \to K^{*0}\eta(\gamma\gamma)$ (dashed orange line) and $B^0 \to K^+\pi^-\gamma X$ (dashed black line).

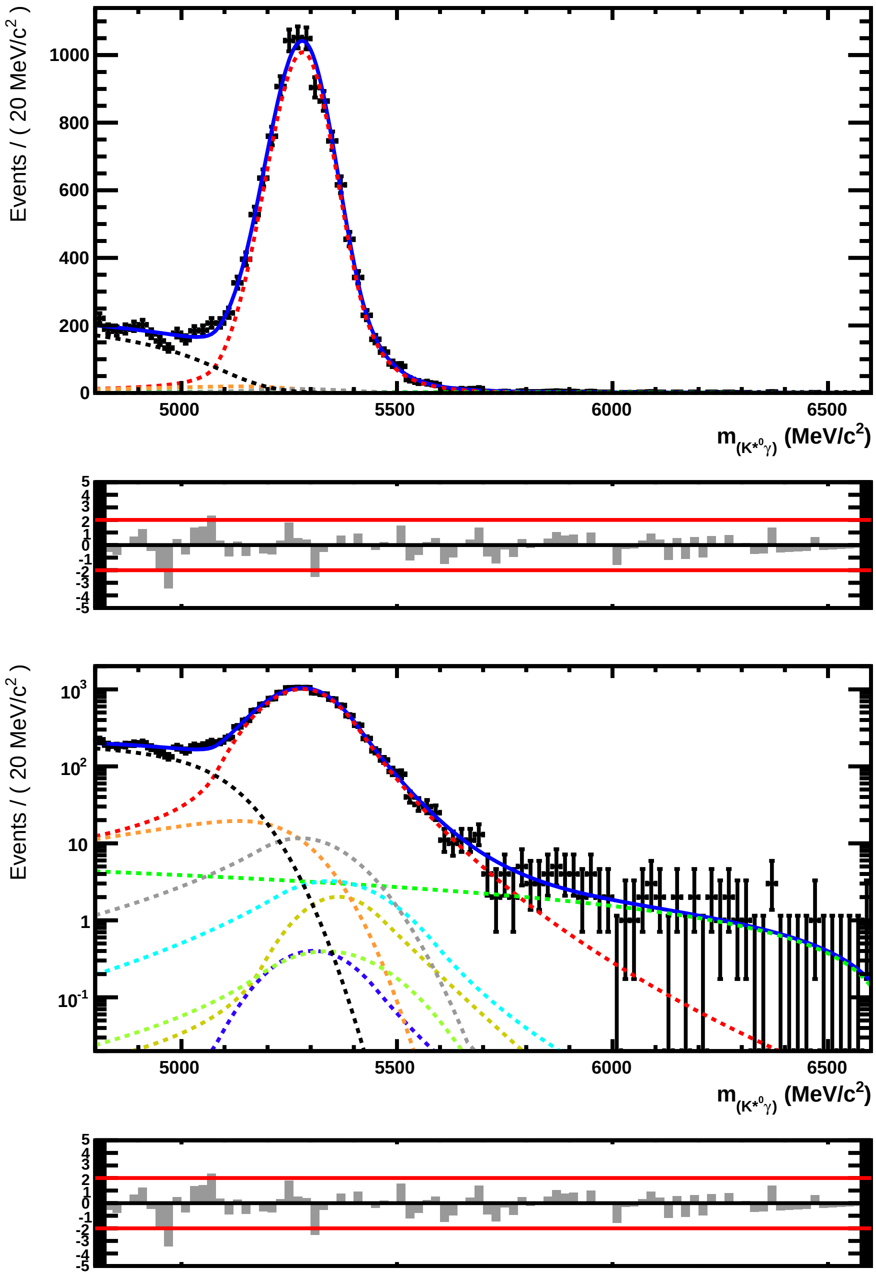


Figure 6.17: 2012 $B^0 \to K^{*0}\gamma$ mass fit in linear (top) and semi-logarithmic scale (bottom); the fit is performed in the restricted range [4800; 6600] MeV/c^2. The black points represent the data and the solid blue line is the total fit function. The signal is fitted with an asymmetric double-tail Crystal Ball PDF (dashed red line) and the combinatorial background is described with a first order polynomial (dashed green line). The peaking background modes are $B^0 \to \rho^0\gamma$ (dashed dark yellow line), $B_s^0 \to \phi\gamma$ (dashed dark blue line), $\Lambda_b^0 \to \Lambda^*\gamma$ (dashed cyan line), $B^0 \to K^+\pi^-\pi^0$ (dashed gray line) and $B^0 \to \pi^+\pi^-\pi^0$ (dashed light green line). The partially reconstructed modes are $B^0 \to K^{*0}\eta(\gamma\gamma)$ (dashed orange line) and $B^0 \to K^+\pi^-\gamma X$ (dashed black line).

Channel	Parameter		Year	
			2011	2012
$B^0 \to K^{*0}\gamma$	N_S		3374 ± 68	11489 ± 118
	μ	MeV/c^2	5275.3 ± 5.3	5281.9 ± 2.5
	σ_L	MeV/c^2	93.4 ± 4.9	90.0 ± 2.4
	σ_R	MeV/c^2	106.0 ± 3.9	83.4 ± 1.7
$B^0 \to \rho^0\gamma$	$\mathcal{C}$	%	0.194	0.209
$B_s^0 \to \phi\gamma$	$\mathcal{C}$	%	0.038	0.044
$\Lambda_b^0 \to \Lambda^*\gamma$	$\mathcal{C}$	%	0.392	0.506
$B^0 \to K^+\pi^-\pi^0$	$\mathcal{C}$	%	1.674	1.674
$B^0 \to \pi^+\pi^-\pi^0$	$\mathcal{C}$	%	0.061	0.061
Combinatorial	N_{Comb}		43 ± 27	201 ± 37
	p_0	$(\text{MeV}/c^2)^{-1}$	-0.36 ± 0.58	-0.94 ± 0.08
$B^0 \to K^{*0}\eta(\gamma\gamma)$	$\mathcal{C}$	%	3.699	3.646
$B^0 \to K^+\pi^-\gamma X$	N_{Part}		701 ± 48	2160 ± 76
	c		-2.98 ± 2.56	-0.54 ± 1.77

Table 6.8: Summary of the parameters entering the fit model and values extracted fitting the $B^0 \to K^{*0}\gamma$ mass spectrum in the limited mass range separately for 2011 and 2012 data.

As we can see looking at the pulls under the mass plots the fits are still very good. The values of the different parameters are coherent with the ones of the previous fits in the full mass range. The signal yields decrease by about 3% as expected from MC given the reduction of the mass range. The simultaneous fit will be based upon this simpler fit in a limited mass range.

Chapter 7

Simultaneous fit to the $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ mass spectra

In this chapter the implementation of the simultaneous fit to the $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ invariant mass spectra is discussed. The robustness of the fit model is tested using pseudo-experiments and preliminary results are presented. The mass fits are performed in the restricted range [4800; 6600] MeV/c^2 and the measurement is conducted putting a blind on the $B^0 \to \rho^0\gamma$ signal yields. Once received the permission from the LHCb collaboration, it will be possible to unblind the fit results and to extract the ratio of branching fractions $\mathcal{B}(B^0 \to \rho^0\gamma)/\mathcal{B}(B^0 \to K^{*0}\gamma)$ that we aim to measure.

Before fitting the data, we checked the $\pi^+\pi^-$ and $K^+\pi^-$ mass distributions for events passing the full $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ selections respectively. As the $\pm 3\sigma$ region is blind, the $\pi^+\pi^-$ mass distributions (see Fig. 7.1) are shown only in the left handed sideband (LHSB), corresponding to the range [4800; 5000] MeV/c^2, and in the right handed sideband (RHSB), corresponding to the range [5580; 6600] MeV/c^2. The $K^+\pi^-$ mass (see Fig. 7.2) is instead shown in the three mass intervals. No specific background is visible. In particular, looking at the $\pi^+\pi^-$ plots, we can notice that in both LHSB and RHSB (here the statistics is particularly low) no peak associated to the K^0_S is visible, indicating that there is no need to apply dedicated cuts to veto $K^0_S \to \pi^+\pi^-$ decays.

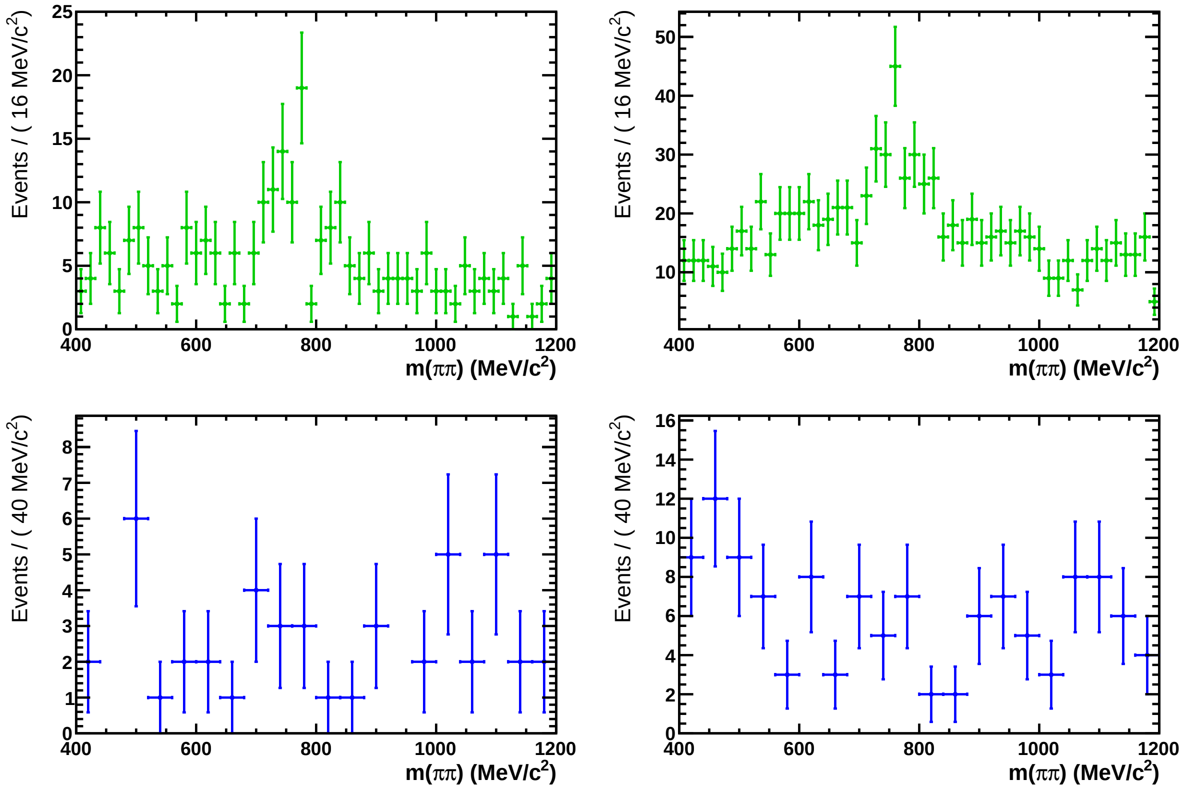


Figure 7.1: 2011 (left) and 2012 (right) reconstructed $\pi^+\pi^-$ invariant mass distributions in the LHSB (top) and RHSB (bottom) for events passing the full $\rho^0\gamma$ selection requirements.

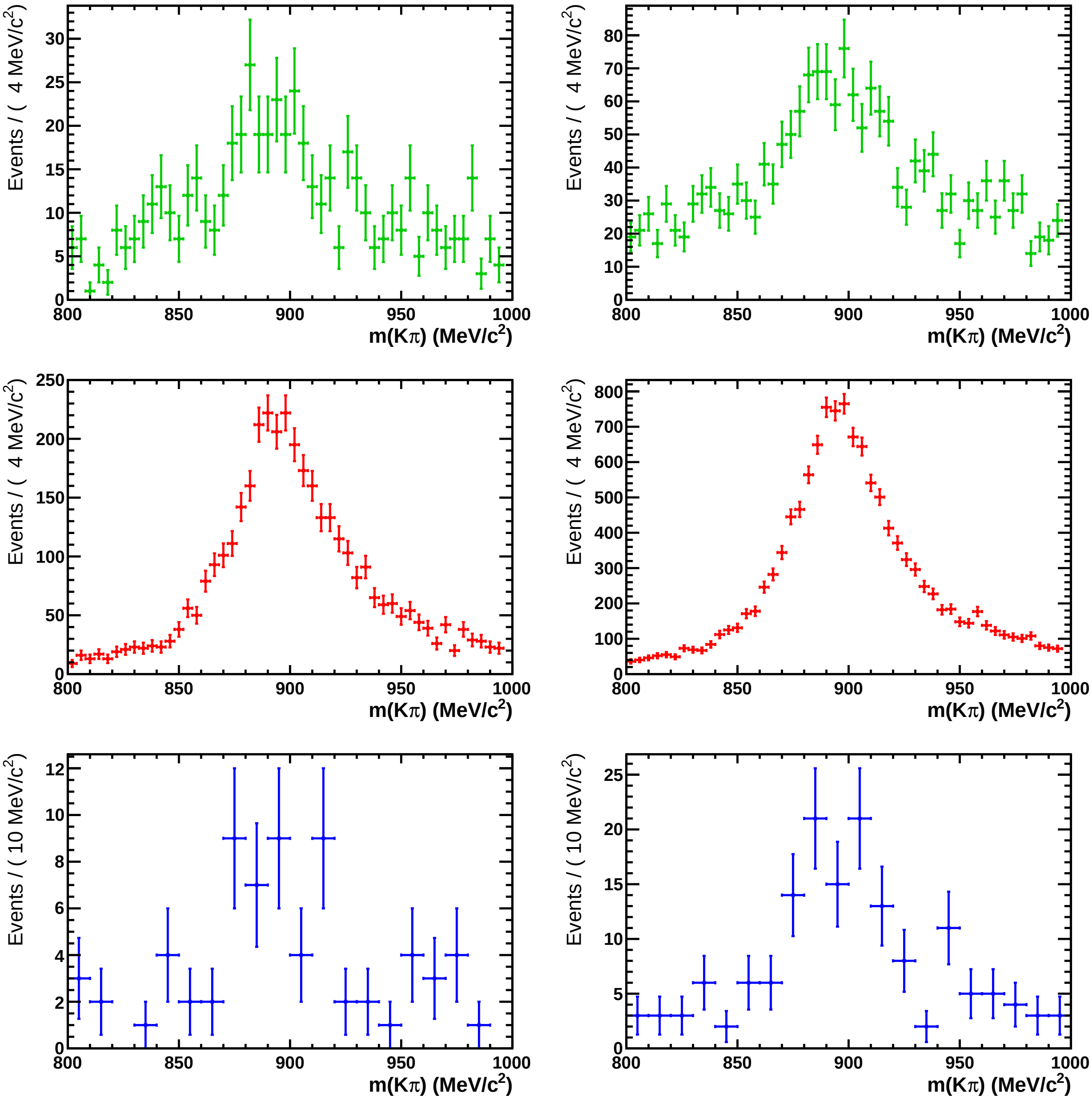


Figure 7.2: 2011 (left) and 2012 (right) reconstructed $K^+\pi^-$ invariant mass distributions in the LHSB (top), $\pm 3\sigma$ region (middle) and RHSB (bottom) for events passing the full $K^{*0}\gamma$ selection requirements.

7.1 Simultaneous fit strategy

The total function used to fit the $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ mass spectra is built as an extended PDF given by

$$\mathcal{F}(m) = \mathcal{F}^{\rho\gamma}(m) + \mathcal{F}^{K^*\gamma}(m) \tag{7.1}$$

where the individual $\mathcal{F}^{\rho\gamma}(m)$ and $\mathcal{F}^{K^*\gamma}(m)$ functions are defined as in Eq. 6.7. Here the number of peaking background events in each spectrum is written according to Tab. 5.21, without any assumption on $r^{Eff}_{\rho\gamma}$, introducing the ratio of fitted signal yields $\mathcal{Y} = N_S^{\rho\gamma}/N_S^{K^*\gamma}$, which is shared between the $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ spectra.

The simultaneous fit is performed according to the same prescriptions discussed in Sec. 6.3 except for the following details:

- The μ, σ_L and σ_R parameters in the $B^0 \to \rho^0\gamma$ spectrum are constrained to those floating in the $B^0 \to K^{*0}\gamma$ spectrum as

$$\mu^{DATA}_{\rho\gamma} = \mu^{DATA}_{K^*\gamma} + (\mu^{MC}_{\rho\gamma} - \mu^{MC}_{K^*\gamma}), \tag{7.2}$$

$$\sigma^{DATA}_{\rho\gamma L} = \sigma^{MC}_{\rho\gamma L} \times \frac{\sigma^{DATA}_{K^*\gamma L}}{\sigma^{MC}_{K^*\gamma L}}, \quad \sigma^{DATA}_{\rho\gamma R} = \sigma^{MC}_{\rho\gamma R} \times \frac{\sigma^{DATA}_{K^*\gamma R}}{\sigma^{MC}_{K^*\gamma R}}. \tag{7.3}$$

- The fitted $B^0 \to \rho^0\gamma$ signal yield, $N_S^{\rho\gamma}$, is blinded using the `RooUnblindUniform` blinding tool of `RooFit`.

- The peaking backgrounds in each spectrum and the related μ and σ for every channel are constrained as in Eq. 6.8 and 6.9 according to the signal hypothesis.

Concerning the other background sources, both the combinatorial and the partially reconstructed background shapes are left free to float independently in the two mass spectra. As the generic $B^0 \to K^+\pi^-\gamma X$ background, the $B^0 \to \pi^+\pi^-\gamma X$ contribution to the $B^0 \to \rho^0\gamma$ spectrum is modeled setting ΔM to -134.977 MeV/c^2 and fixing $p = 0.5$ in the generalized ARGUS of Eq. 6.5.

The projections of the simultaneous mass fits to the $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ mass spectra for 2011 ($0.7\,\text{fb}^{-1}$) and 2012 ($2\,\text{fb}^{-1}$) are shown respectively in Fig. 7.3 - 7.4 and 7.5 - 7.6, while the values of the fit parameters, including the fixed contaminations, are shown in Tab. 7.1 and 7.2. The $B^0 \to \rho^0\gamma$ signal yields are blind and the μ, σ_L and σ_R are not displayed as they are constrained in the fit.

As we can see from the residuals under the mass plots, the agreement between the data and the individual fit functions is very good in both mass spectra. The fitted parameters in the $B^0 \to K^{*0}\gamma$ spectrum are compatible with the ones shown in Tab. 6.8 for both 2011 and 2012.

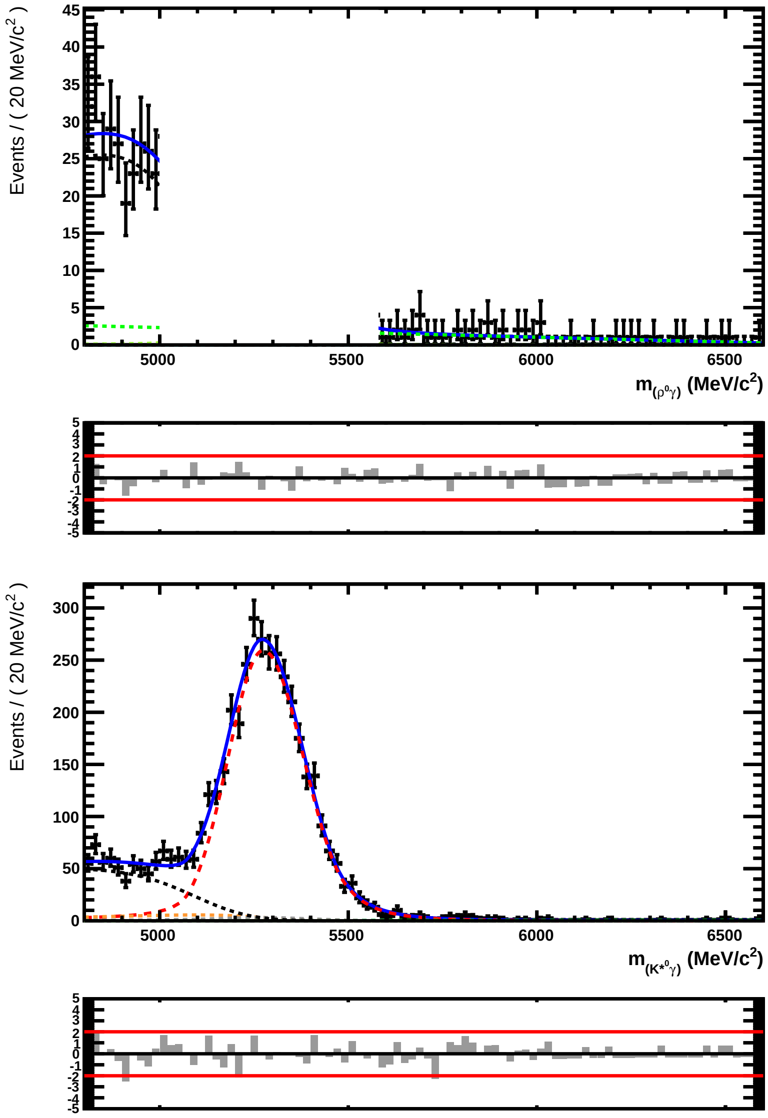


Figure 7.3: $B^0 \to \rho^0\gamma$ (top) and $B^0 \to K^{*0}\gamma$ (bottom) mass fit projections for 2011 data in linear scale. The black points represent the data in the respective mass spectrum and the solid blue line is the total fit function. The signal is fitted with an asymmetric double-tail Crystal Ball PDF and the combinatorial background is described with a first order polynomial (dashed green line). In the $B^0 \to \rho^0\gamma$ spectrum the $\pm 3\sigma$ region is blind and the signal PDF is not shown. The peaking background modes are respectively $B^0 \to K^{*0}\gamma$ (dashed dark yellow line), $\Lambda_b^0 \to \Lambda^*\gamma$ (dashed cyan line), $B^0 \to \pi^+\pi^-\pi^0$ (dashed light green line) and $B^0 \to K^+\pi^-\pi^0$ (dashed gray line). The partially reconstructed contribution is associated to $B^0 \to \pi^+\pi^-\gamma X$ decays (dashed black line). In the $B^0 \to K^{*0}\gamma$ spectrum the signal PDF (dashed red line) is visible while the $B^0 \to \rho^0\gamma$ cross-feed is not plotted. The peaking background modes, apart from $B_s^0 \to \phi\gamma$ (dashed dark blue line), are the same as in the $B^0 \to \rho^0\gamma$ spectrum and are plotted with the same colors. The partially reconstructed modes are $B^0 \to K^{*0}\eta(\gamma\gamma)$ (dashed orange line) and $B^0 \to K^+\pi^-\gamma X$ (dashed black line). Several background contributions in both mass spectra are not visible here (see plots in semi-logarithmic scale).

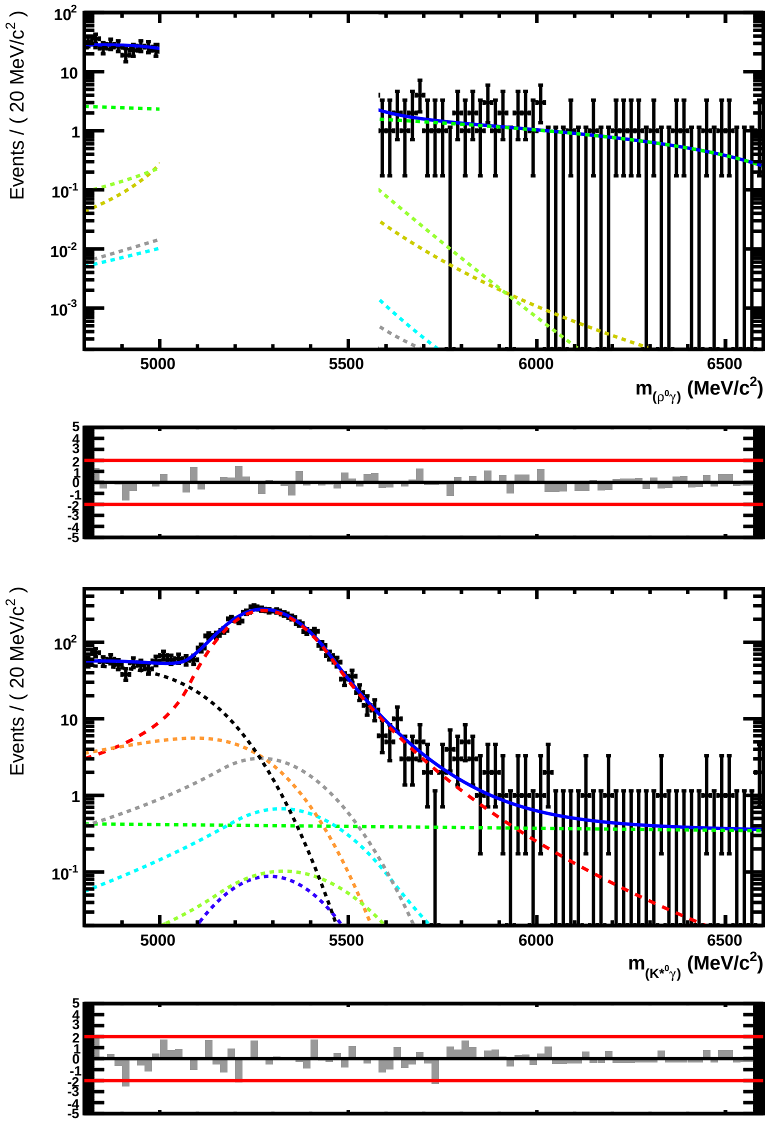


Figure 7.4: $B^0 \to \rho^0\gamma$ (top) and $B^0 \to K^{*0}\gamma$ (bottom) mass fit projections for 2011 data in semi-logarithmic scale. The black points represent the data in the respective mass spectrum and the solid blue line is the total fit function. The signal is fitted with an asymmetric double-tail Crystal Ball PDF and the combinatorial background is described with a first order polynomial (dashed green line). In the $B^0 \to \rho^0\gamma$ spectrum the $\pm 3\sigma$ region is blind and the signal PDF is not shown. The peaking background modes are respectively $B^0 \to K^{*0}\gamma$ (dashed dark yellow line), $\Lambda_b^0 \to \Lambda^*\gamma$ (dashed cyan line), $B^0 \to \pi^+\pi^-\pi^0$ (dashed light green line) and $B^0 \to K^+\pi^-\pi^0$ (dashed gray line). The partially reconstructed contribution is associated to $B^0 \to \pi^+\pi^-\gamma X$ decays (dashed black line). In the $B^0 \to K^{*0}\gamma$ spectrum the signal PDF (dashed red line) is visible while the $B^0 \to \rho^0\gamma$ cross-feed is not plotted. The peaking background modes, apart from $B_s^0 \to \phi\gamma$ (dashed dark blue line), are the same as in the $B^0 \to \rho^0\gamma$ spectrum and are plotted with the same colors. The partially reconstructed modes are $B^0 \to K^{*0}\eta(\gamma\gamma)$ (dashed orange line) and $B^0 \to K^+\pi^-\gamma X$ (dashed black line).

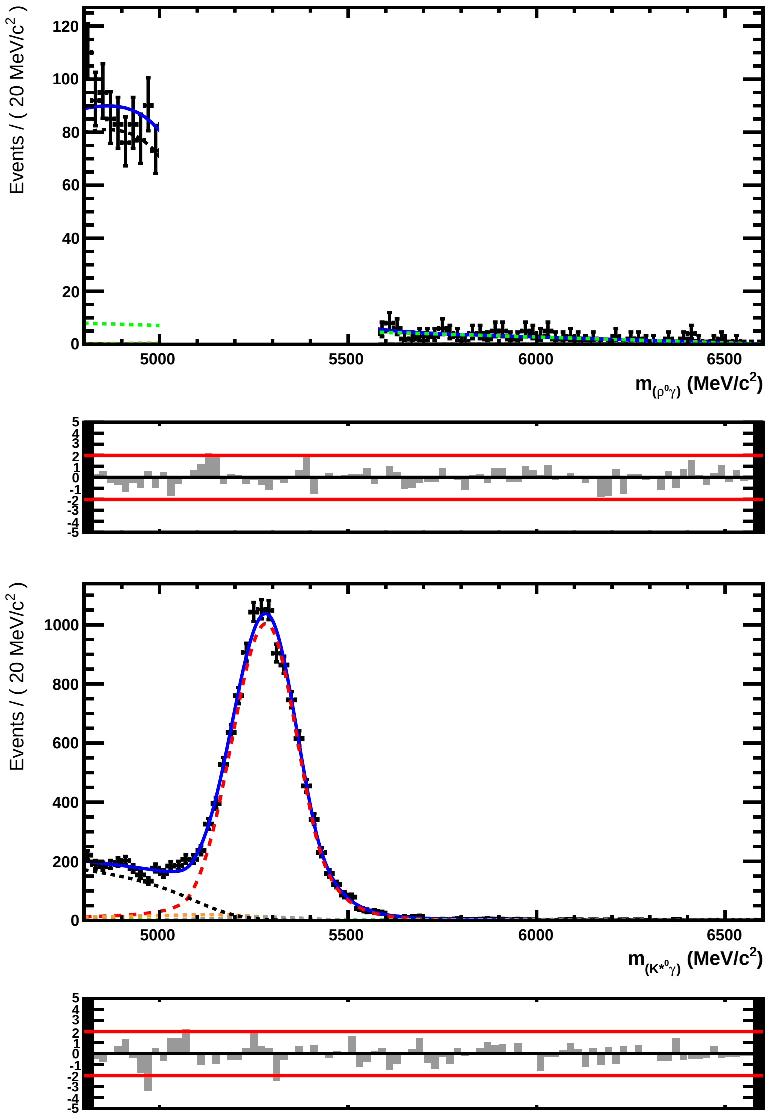


Figure 7.5: $B^0 \to \rho^0\gamma$ (top) and $B^0 \to K^{*0}\gamma$ (bottom) mass fit projections for 2012 data in linear scale. The black points represent the data in the respective mass spectrum and the solid blue line is the total fit function. The signal is fitted with an asymmetric double-tail Crystal Ball PDF and the combinatorial background is described with a first order polynomial (dashed green line). In the $B^0 \to \rho^0\gamma$ spectrum the $\pm 3\sigma$ region is blind and the signal PDF is not shown. The peaking background modes are respectively $B^0 \to K^{*0}\gamma$ (dashed dark yellow line), $\Lambda_b^0 \to \Lambda^*\gamma$ (dashed cyan line), $B^0 \to \pi^+\pi^-\pi^0$ (dashed light green line) and $B^0 \to K^+\pi^-\pi^0$ (dashed gray line). The partially reconstructed contribution is associated to $B^0 \to \pi^+\pi^-\gamma X$ decays (dashed black line). In the $B^0 \to K^{*0}\gamma$ spectrum the signal PDF (dashed red line) is visible while the $B^0 \to \rho^0\gamma$ cross-feed is not plotted. The peaking background modes, apart from $B_s^0 \to \phi\gamma$ (dashed dark blue line), are the same as in the $B^0 \to \rho^0\gamma$ spectrum and are plotted with the same colors. The partially reconstructed modes are $B^0 \to K^{*0}\eta(\gamma\gamma)$ (dashed orange line) and $B^0 \to K^+\pi^-\gamma X$ (dashed black line). Several background contributions in both mass spectra are not visible here (see plots in semi-logarithmic scale).

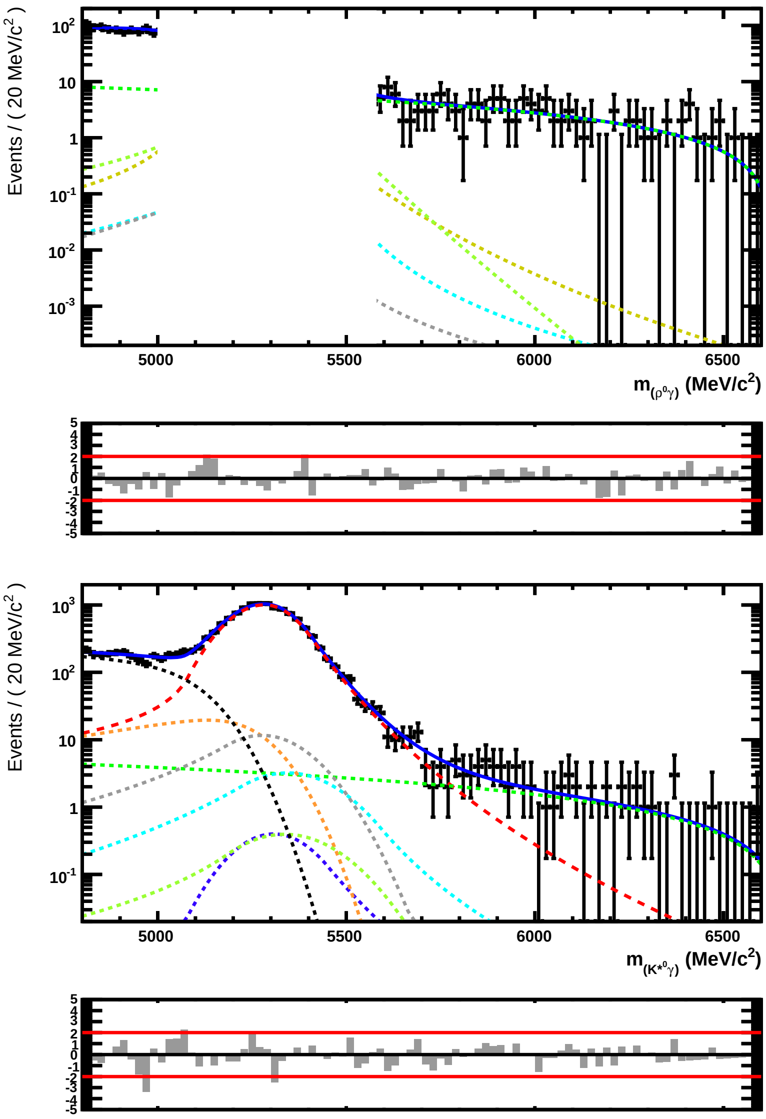


Figure 7.6: $B^0 \to \rho^0\gamma$ (top) and $B^0 \to K^{*0}\gamma$ (bottom) mass fit projections for 2012 data in semi-logarithmic scale. The black points represent the data in the respective mass spectrum and the solid blue line is the total fit function. The signal is fitted with an asymmetric double-tail Crystal Ball PDF and the combinatorial background is described with a first order polynomial (dashed green line). In the $B^0 \to \rho^0\gamma$ spectrum the $\pm 3\sigma$ region is blind and the signal PDF is not shown. The peaking background modes are respectively $B^0 \to K^{*0}\gamma$ (dashed dark yellow line), $\Lambda_b^0 \to \Lambda^*\gamma$ (dashed cyan line), $B^0 \to \pi^+\pi^-\pi^0$ (dashed light green line) and $B^0 \to K^+\pi^-\pi^0$ (dashed gray line). The partially reconstructed contribution is associated to $B^0 \to \pi^+\pi^-\gamma X$ decays (dashed black line). In the $B^0 \to K^{*0}\gamma$ spectrum the signal PDF (dashed red line) is visible while the $B^0 \to \rho^0\gamma$ cross-feed is not plotted. The peaking background modes, apart from $B_s^0 \to \phi\gamma$ (dashed dark blue line), are the same as in the $B^0 \to \rho^0\gamma$ spectrum and are plotted with the same colors. The partially reconstructed modes are $B^0 \to K^{*0}\eta(\gamma\gamma)$ (dashed orange line) and $B^0 \to K^+\pi^-\gamma X$ (dashed black line).

Channel	Parameter		Signal mode	
			$B^0 \to \rho^0\gamma$	$B^0 \to K^{*0}\gamma$
	N_S		XXX.X±XXX.X	3378 ± 69
	μ	MeV/c^2		5274.4 ± 7.1
	σ_L	MeV/c^2		93.0 ± 6.0
	σ_R	MeV/c^2		106.5 ± 5.1
$B^0 \to K^{*0}\gamma$	$\mathcal{C}$	%	$0.602 \times \mathcal{Y}^{-1}$	1
$B^0 \to \rho^0\gamma$	$\mathcal{C}$	%	1	$6.676 \times \mathcal{Y}$
$\Lambda_b^0 \to \Lambda^*\gamma$	$\mathcal{C}$	%	$0.014 \times \mathcal{Y}^{-1}$	0.392
$B_s^0 \to \phi\gamma$	$\mathcal{C}$	%	–	0.038
$B^0 \to \pi^+\pi^-\pi^0$	$\mathcal{C}$	%	$0.724 \times \mathcal{Y}^{-1}$	0.061
$B^0 \to K^+\pi^-\pi^0$	$\mathcal{C}$	%	$0.016 \times \mathcal{Y}^{-1}$	1.674
Combinatorial	N_{Comb}		127 ± 22	35 ± 37
	p_0	(MeV/c^2)$^{-1}$	-0.82 ± 0.13	-0.10 ± 1.53
$B^0 \to \pi^+\pi^-\gamma X$	N_{Part}		391 ± 25	–
	c		-5.06 ± 2.41	–
$B^0 \to K^{*0}\eta(\gamma\gamma)$	$\mathcal{C}$	%	–	3.699
$B^0 \to K^+\pi^-\gamma X$	N_{Part}		–	704 ± 52
	c		–	-3.08 ± 2.62

Table 7.1: Summary of the parameters entering the fit model and values extracted fitting simultaneously the $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ mass spectra with 2011 data.

Channel	Parameter		Signal mode	
			$B^0 \to \rho^0\gamma$	$B^0 \to K^{*0}\gamma$
	N_S		XXX.X±XXX.X	11497 ± 118
	μ	MeV/c^2		5282.8 ± 2.2
	σ_L	MeV/c^2		91.2 ± 2.2
	σ_R	MeV/c^2		83.1 ± 1.5
$B^0 \to K^{*0}\gamma$	$\mathcal{C}$	%	$0.673 \times \mathcal{Y}^{-1}$	1
$B^0 \to \rho^0\gamma$	$\mathcal{C}$	%	1	$7.137 \times \mathcal{Y}$
$\Lambda_b^0 \to \Lambda^*\gamma$	$\mathcal{C}$	%	$0.027 \times \mathcal{Y}^{-1}$	0.506
$B_s^0 \to \phi\gamma$	$\mathcal{C}$	%	–	0.044
$B^0 \to \pi^+\pi^-\pi^0$	$\mathcal{C}$	%	$0.724 \times \mathcal{Y}^{-1}$	0.061
$B^0 \to K^+\pi^-\pi^0$	$\mathcal{C}$	%	$0.016 \times \mathcal{Y}^{-1}$	1.674
Combinatorial	N_{Comb}		366 ± 36	201 ± 36
	p_0	$(\text{MeV}/c^2)^{-1}$	-0.97 ± 0.01	-0.94 ± 0.08
$B^0 \to \pi^+\pi^-\gamma X$	N_{Part}		1241 ± 43	–
	c		-4.66 ± 1.17	–
$B^0 \to K^{*0}\eta(\gamma\gamma)$	$\mathcal{C}$	%	–	3.646
$B^0 \to K^+\pi^-\gamma X$	N_{Part}		–	2146 ± 73
	c		–	-0.30 ± 0.79

Table 7.2: Summary of the parameters entering the fit model and values extracted fitting simultaneously the $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ mass spectra with 2012 data.

7.2 Validation of the fit

In order to validate the fitting procedure toy Monte Carlo studies are conducted to check for possible fit biases. A thousand of pseudo-experiments (toy samples) has been generated for the reconstructed $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ mass spectra using the parameters extracted from the simultaneous fit of the data. The $B^0 \to \rho^0\gamma$ signal yield, $N^{\rho\gamma}$, which is blind in the fit of the data, is set assuming for $r^{Eff}_{\rho\gamma}$ the value of the PDG. Each sample is then fitted with the same model used to fit the data. The pull distribution $\mathcal{P}(x)$ of a given parameter x is defined as

$$\mathcal{P}(x) = \frac{x_{fit} - x_{gen}}{\sigma_x} \tag{7.4}$$

where x_{fit} is the value of the parameter x extracted fitting the toy sample, x_{gen} the value used to generate the sample and σ_x the uncertainty returned by the fit. If the fit is well behaved, the distribution of $\mathcal{P}(x)$ should be Gaussian with mean zero and width one.

As examples, the 2011 and 2012 fits to the $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ invariant mass distributions for a single generated toy sample are shown in Fig. 7.7 and 7.8. The pull distributions of the $B^0 \to \rho^0\gamma$ signal yields for 2011 and 2012 are displayed respectively in Fig. 7.9 and 7.10, together with the distributions of the significance $N^{\rho\gamma}/\Delta N^{\rho\gamma}$ and relative uncertainty $\Delta N^{\rho\gamma}/N^{\rho\gamma}$. Looking at the results of the fits to the pull distributions, we can see that no significant bias is observed as the central values of the Gaussians are consistent with zero and the resolutions are compatible with one.

Looking at the $N^{\rho\gamma}/\Delta N^{\rho\gamma}$ and $\Delta N^{\rho\gamma}/N^{\rho\gamma}$ plots we can notice that pseudo-experiments giving a significance greater than 5σ are about 71% of the generated ones in 2011 and 100% in 2012, while the expected relative uncertainty is on average about 19% in 2011 and 9% in 2012.

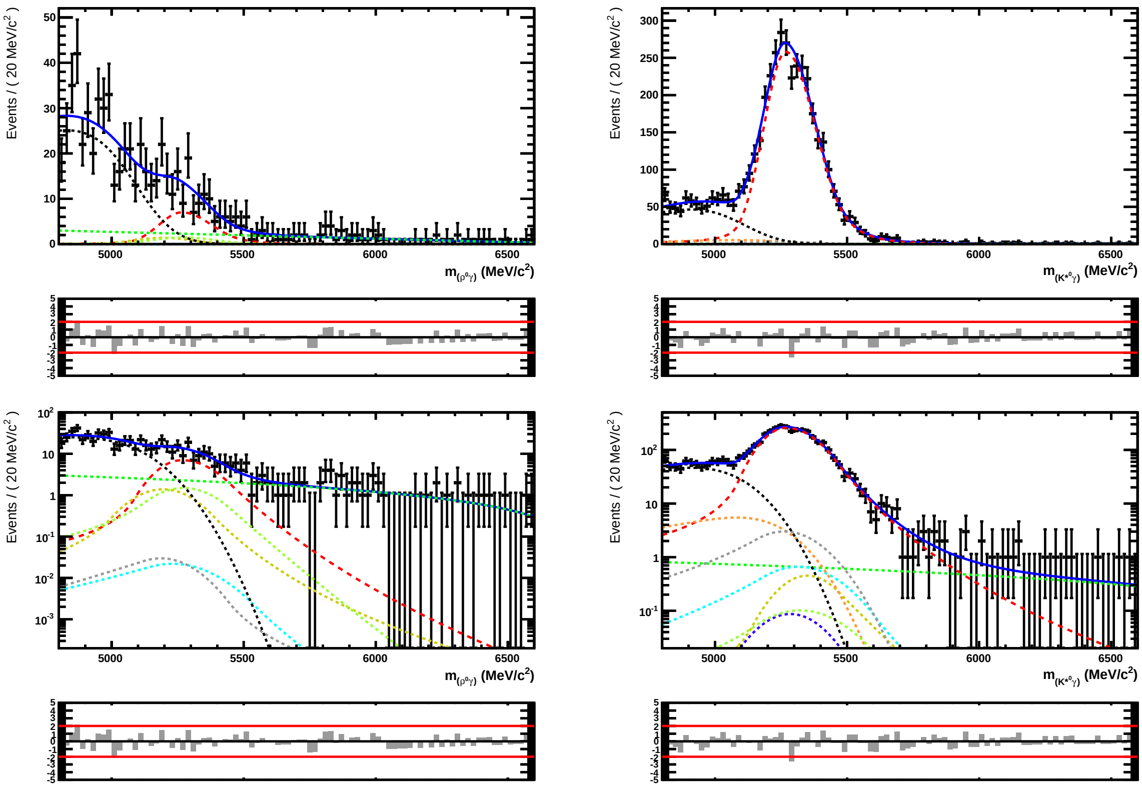


Figure 7.7: 2011 $B^0 \to \rho^0\gamma$ (left) and $B^0 \to K^{*0}\gamma$ (right) mass fit projections for one generated toy sample in linear (top) and semi-logarithmic (bottom) scale. Here the signal PDF (dashed red line) is shown in both spectra. Let us notice that in the $B^0 \to \rho^0\gamma$ spectrum the dashed dark yellow line stands for the $B^0 \to K^{*0}\gamma$ cross-feed, while in the $B^0 \to K^{*0}\gamma$ spectrum it stands for the $B^0 \to \rho^0\gamma$ cross-feed.

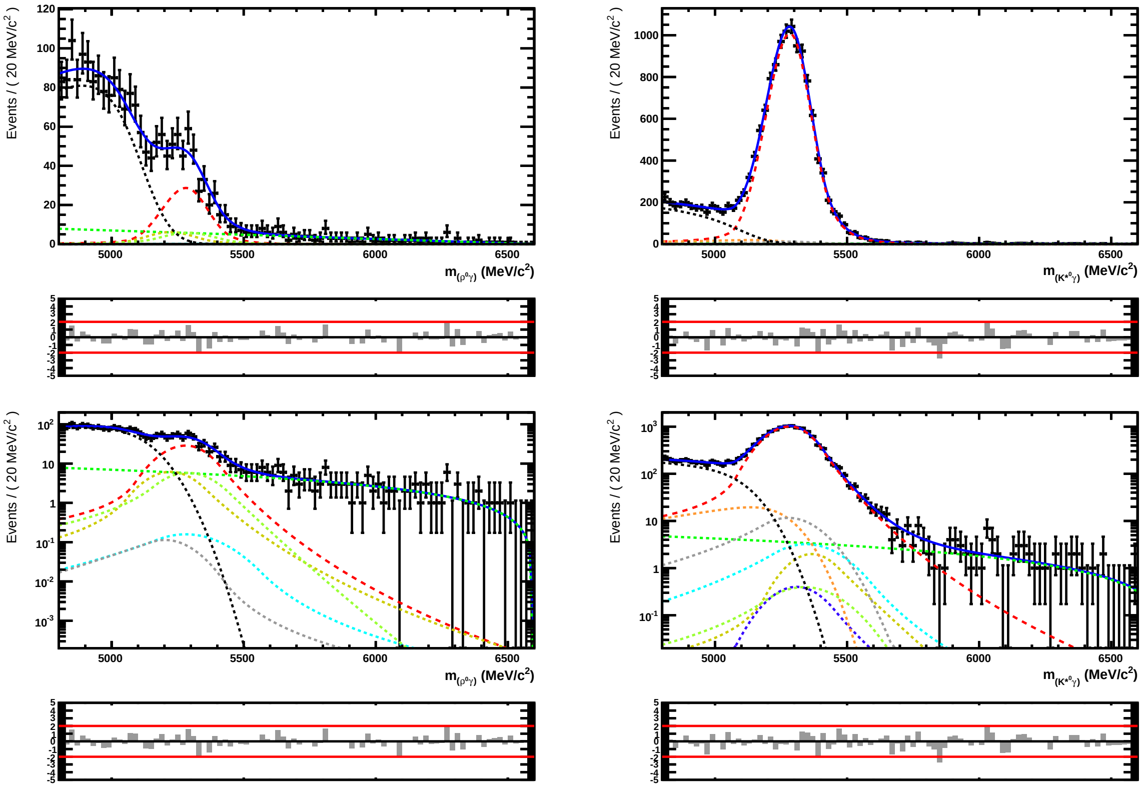


Figure 7.8: 2012 $B^0 \to \rho^0\gamma$ (left) and $B^0 \to K^{*0}\gamma$ (right) mass fit projections for one generated toy sample in linear (top) and semi-logarithmic (bottom) scale. Here the signal PDF (dashed red line) is shown in both spectra. Let us notice that in the $B^0 \to \rho^0\gamma$ spectrum the dashed dark yellow line stands for the $B^0 \to K^{*0}\gamma$ cross-feed, while in the $B^0 \to K^{*0}\gamma$ spectrum it stands for the $B^0 \to \rho^0\gamma$ cross-feed.

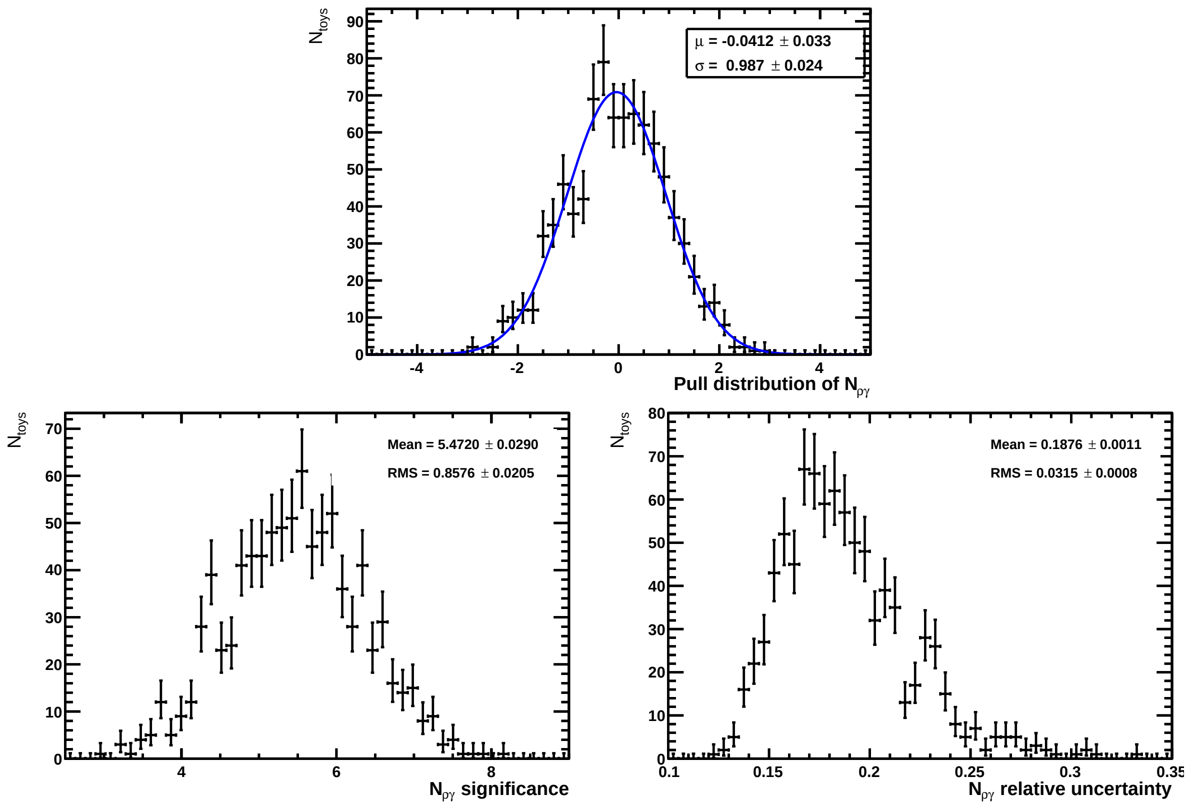


Figure 7.9: Pull (top), significance (bottom left) and relative uncertainty (bottom right) distributions for 2011 $B^0 \to \rho^0\gamma$ yields. The blue curve is the result of a gaussian fit to the pull distribution.

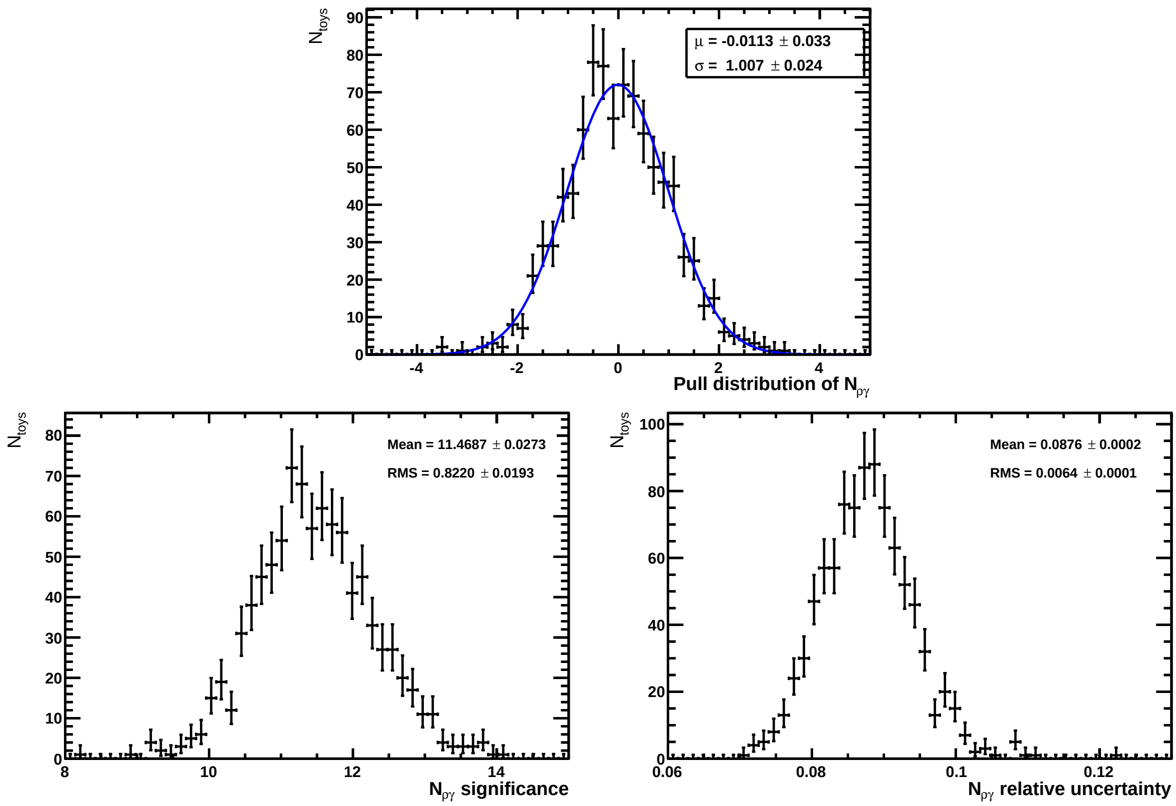


Figure 7.10: Pull (top), significance (bottom left) and relative uncertainty (bottom right) distributions for 2012 $B^0 \to \rho^0\gamma$ yields. The blue curve is the result of a gaussian fit to the pull distribution.

7.3 Towards a measurement of the ratio of branching fractions

For the measurement of the ratio of branching ratios we plan to fit simultaneously the 2011 and 2012 datasets. The fit model is the same except that, instead of fitting the $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ yields, we fit the $B^0 \to K^{*0}\gamma$ yield ($N^{K^*\gamma}$) and the ratio of branching ratios ($r^{Eff}_{\rho\gamma}$) which is the only fitted parameter common between the 2011 and 2012 fit models. The 2011 and 2012 yield ratios ($\mathcal{Y}$) are then given by Eq. 5.35. The 2011 and 2012 $B^0 \to \rho^0\gamma$ signal yields ($N^{\rho\gamma}$) are simply the product of the yield ratio and of the $B^0 \to K^{*0}\gamma$ yield ($\mathcal{Y} \times N^{K^*\gamma}$). This way, the final result of the measurement can be obtained from

$$\frac{\mathcal{B}(B^0 \to \rho^0\gamma)}{\mathcal{B}(B^0 \to K^{*0}\gamma)} = r^{Eff}_{\rho\gamma} \times \mathcal{B}(K^{*0} \to K^+\pi^-) \; . \tag{7.5}$$

To validate this 2011+2012 simultaneous fit, another set of 1200 pseudo-experiments is generated, still assuming for $r^{Eff}_{\rho\gamma}$ the value of the PDG. As an example, the 2011 and 2012 simultaneous fit to the $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ invariant mass distributions for a single generated toy sample are shown in Fig. 7.11. The distribution of the $r^{Eff}_{\rho\gamma}$ values as fitted on the toy samples and the pull distribution are displayed on Fig. 7.12. The mean of the fitted ratio $r^{Eff}_{\rho\gamma}$ distribution is in very good agreement with the input value (0.0299). The root mean square corresponds to a relative uncertainty of 7.8%. Looking at the fit of the pull distribution we notice that no significant bias is observed since the central value of the Gaussian is consistent with zero and the resolution is compatible with one. It has to be mentioned that the pull distribution is limited to samples in which the fit gives a full accurate covariance matrix (only about half of the pseudo-experiments) as we did not use the Minos method to have a better evaluation of the uncertainties to speed up the validation procedure. In particular, the range in which the input parameters are allowed to float in the toy generation was not optimized. When performing the fit on real data, the error matrix results to be full accurate.

The projections of the 2011+2012 simultaneous mass fit to the $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ mass spectra for 2011 and 2012 data samples are shown respectively in Fig. 7.13 and 7.14. The values of the parameters allowed to float in the fit are reported in Tab. 7.3. The ratio $r^{Eff}_{\rho\gamma}$ is blind and the μ, σ_L and σ_R for the $B^0 \to \rho^0\gamma$ channel are not displayed as they are constrained to the $B^0 \to K^{*0}\gamma$ parameters.

As can be seen from the residuals under the mass plots, the agreement between the data and the individual fit functions is very good in both mass spectra and years. All the fitted parameters for both $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ modes are compatible with the ones shown in Tab. 7.1 and 7.2. This means that the yield ratios are in agreement between 2011 and 2012.

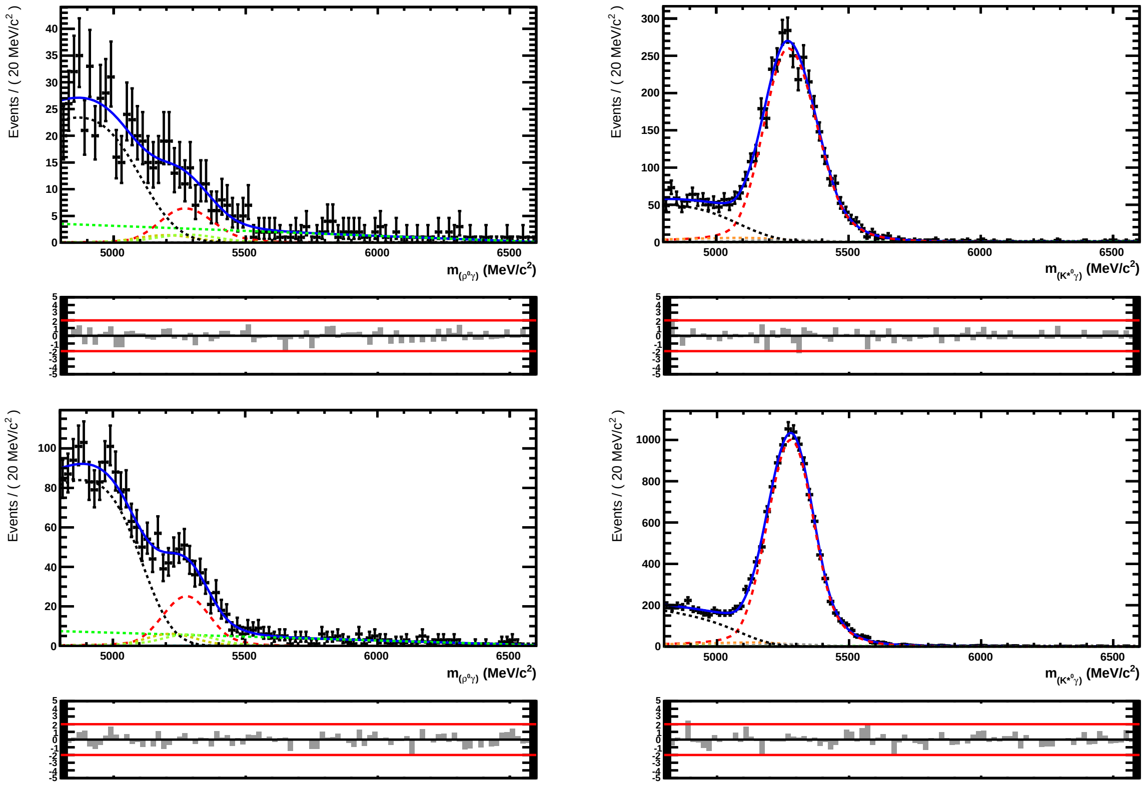


Figure 7.11: $B^0 \to \rho^0\gamma$ (left) and $B^0 \to K^{*0}\gamma$ (right) mass fit projections for one generated (2011+2012) toy sample for 2011 (top) and 2012 (bottom).

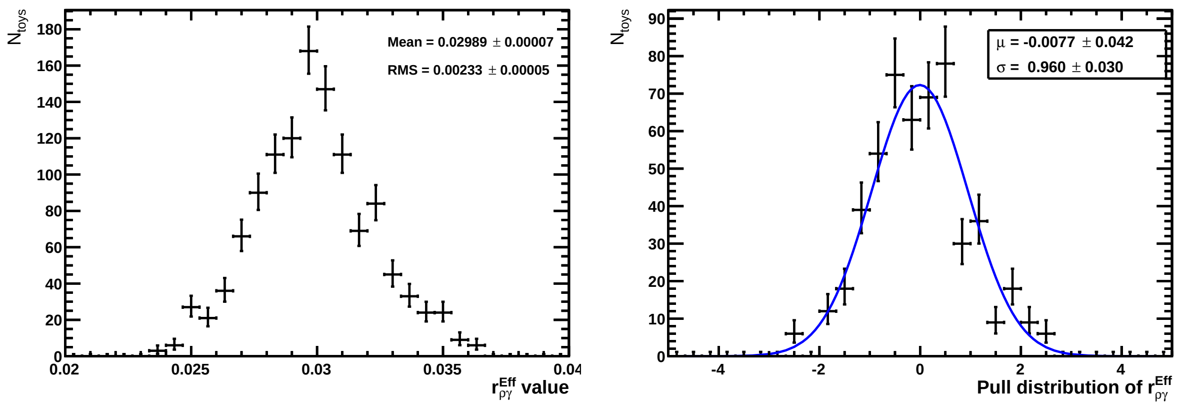


Figure 7.12: Fitted values (left) and pull (right) distributions for the ratio of branching fractions $r^{Eff}_{\rho\gamma}$. The blue curve is the result of a gaussian fit to the pull distribution.

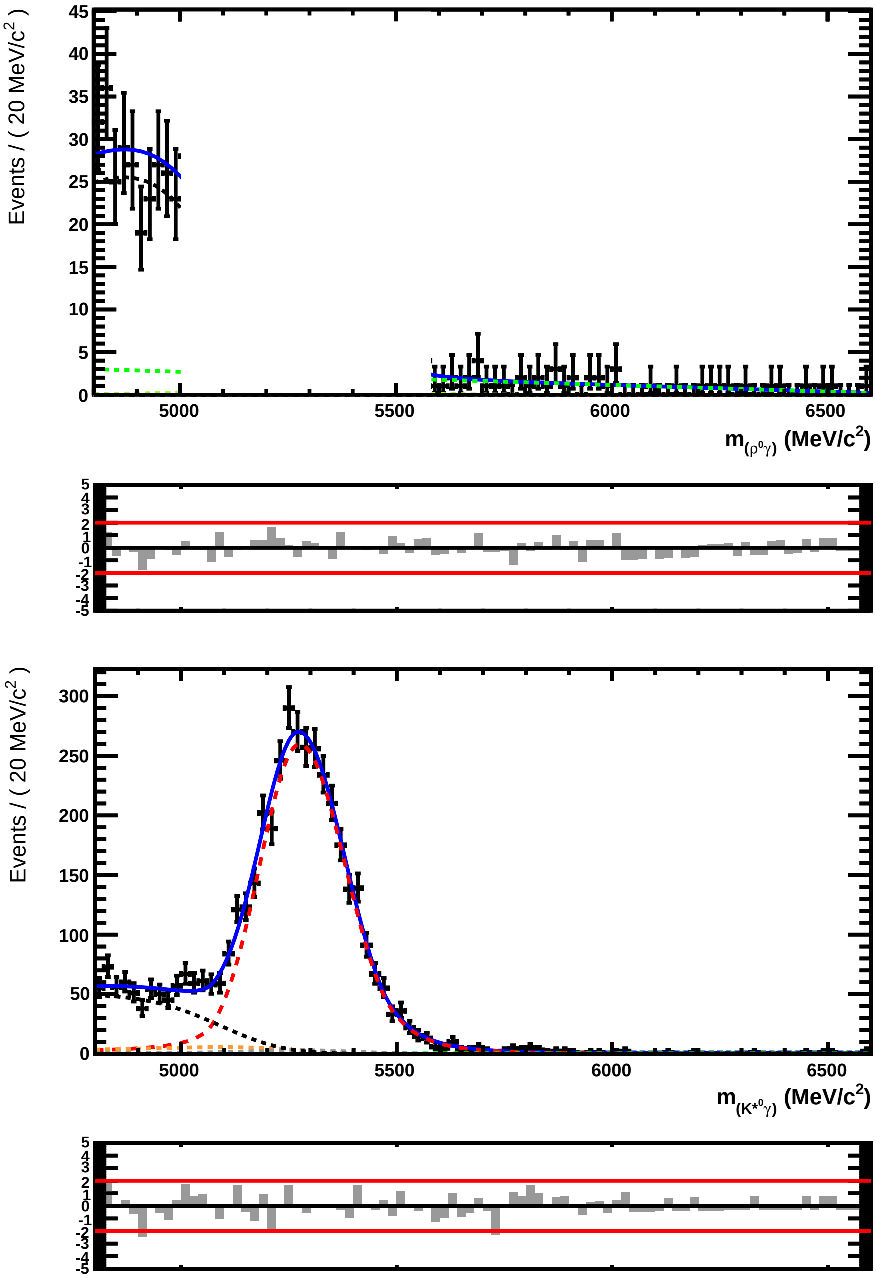


Figure 7.13: $B^0 \to \rho^0\gamma$ (top) and $B^0 \to K^{*0}\gamma$ (bottom) 2011+2012 simultaneous mass fit projections for 2011 data in linear scale. The black points represent the data in the respective mass spectrum and the solid blue line is the total fit function. In the $B^0 \to \rho^0\gamma$ spectrum the $\pm 3\sigma$ region is blind and the signal PDF is not shown. The description of the various background contributions visible in the plots is already detailed in previous figures.

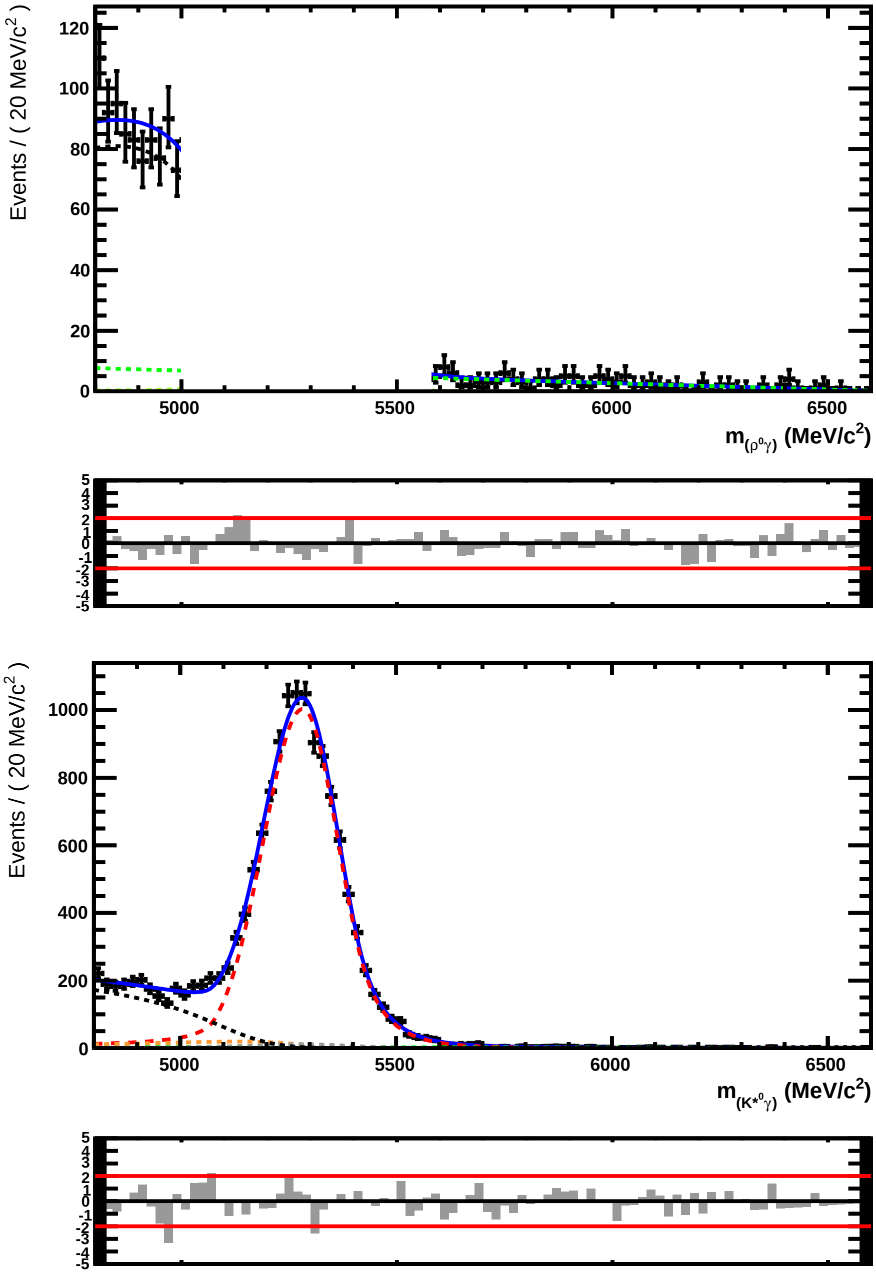


Figure 7.14: $B^0 \to \rho^0\gamma$ (top) and $B^0 \to K^{*0}\gamma$ (bottom) 2011+2012 simultaneous mass fit projections for 2012 data in linear scale. The black points represent the data in the respective mass spectrum and the solid blue line is the total fit function. In the $B^0 \to \rho^0\gamma$ spectrum the $\pm 3\sigma$ region is blind and the signal PDF is not shown. The description of the various background contributions visible in the plots is already detailed in previous figures.

Channel	Parameter		Year	Signal mode	
				$B^0 \to \rho^0\gamma$	$B^0 \to K^{*0}\gamma$
Signals	$r_{\rho\gamma}^{Eff}$			XXX.X±XXX.X	
	N_S		2011		3392 ± 67
			2012		11484 ± 119
	μ	MeV/c^2	2011		5274.9 ± 4.9
			2012		5283.6 ± 2.3
	σ_L	MeV/c^2	2011		93.2 ± 4.6
			2012		92.0 ± 2.4
	σ_R	MeV/c^2	2011		106.3 ± 3.7
			2012		82.6 ± 1.5
Combinatorial	N_{Comb}		2011	142 ± 21	38 ± 29
			2012	355 ± 35	206 ± 36
	p_0	(MeV/c^2)$^{-1}$	2011	-0.87 ± 0.11	-0.24 ± 0.85
			2012	-0.95 ± 0.07	-0.95 ± 0.07
$B^0 \to \pi^+\pi^-\gamma X$	N_{Part}		2011	394 ± 25	–
			2012	1241 ± 43	–
	c		2011	-5.78 ± 2.37	–
			2012	-4.38 ± 1.16	–
$B^0 \to K^+\pi^-\gamma X$	N_{Part}		2011	–	701 ± 48
			2012	–	2135 ± 76
	c		2011	–	-2.98 ± 2.53
			2012	–	-0.05 ± 1.30

Table 7.3: Summary of the free parameters entering the fit model and values extracted fitting the $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ mass spectra simultaneously on 2011 and 2012 data.

Chapter 8

Preliminary results

In this chapter we present the results obtained following the private unblinding of the simultaneous mass fit to the $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ mass spectra. At a first stage, we disclose the $\pm 3\sigma$ signal region, perform the mass fit and calculate signal weights to get background subtracted data samples. This way we can check the reconstruted $\pi^+\pi^-$ and $K^+\pi^-$ mass spectra in order to verify the understanding of the backgrounds. The large ρ^0 mass window was found to be not suitable. In a second iteration, we then tighten the ρ^0 mass window, re-evaluate all the contaminations fixed in the fit and perform the final mass fit to extract the results. As further crosschecks, we verify again the reconstruted $\pi^+\pi^-$ and $K^+\pi^-$ mass spectra and inspect the helicity distributions to check if they are compatible with $B \to V\gamma$ decays. In the second part of the chapter we present the preliminary evaluation of systematic uncertainties and briefly discuss the results.

8.1 Unblinded mass fit

Before unblinding the $\pm 3\sigma$ $\rho\gamma$ signal region, we re-evaluated the $B^0 \to K^{*0}\eta(\gamma\gamma)$ contamination to the $B^0 \to K^{*0}\gamma$ spectrum (see Tab. 8.1). This is the only mode significantly affected by the restriction of the mass fit range $[4000 \to 4800;\ 6600]$ MeV/c^2.

We also decided to add the $B^0 \to \eta(\gamma\gamma)\rho^0$ channel as additional background contribution to the $B^0 \to \rho^0\gamma$ spectrum. This mode is still unobserved but potentially contributes. The respective contamination is fixed in the fit and is estimated using the branching ratio prediction from calculations in the framework of flavour symmetry [84]. The theoretical prediction is $\mathcal{B}(B^0 \to \eta\rho^0) = (0.54 \pm 0.32) \times 10^{-6}$, whereas the experimental limit from BaBar [85] is $\mathcal{B}(B^0 \to \eta\rho^0) < 1.5 \times 10^{-6}$ at a 90% CL. Given no $B^0 \to \eta\rho^0$ MC sample is available, we computed the contamination assuming for $\eta\rho$ in the $B^0 \to \rho^0\gamma$ spectrum the same ratio of efficiencies as $K^*\eta$ in the $B^0 \to K^{*0}\gamma$ spectrum. According to what has been done in Sec. 5.7 and considering $\mathcal{E}^{\rho\gamma}_{\eta\rho}/\mathcal{E}^{\rho\gamma} = \mathcal{E}^{K^*\gamma}_{K^*\eta}/\mathcal{E}^{K^*\gamma}$, we find

$$\mathcal{C}^{\rho\gamma}_{\eta\rho} = \frac{\mathcal{E}^{K^*\gamma}_{K^*\eta}\mathcal{E}^{\rho\gamma}}{(\mathcal{E}^{K^*\gamma})^2} \times r^{Eff}_{\eta\rho} \times \mathcal{Y}^{-1} = \mathcal{D}^{\rho\gamma}_{\eta\rho} \times \mathcal{Y}^{-1}\ , \tag{8.1}$$

where $r^{Eff}_{\eta\rho} = (7.39 \pm 2.60) \times 10^{-3}$. The resulting contamination is

$$\mathcal{C}^{\rho\gamma}_{\eta\rho} = (0.182 \pm 0.108) \times \mathcal{Y}^{-1}\ \%$$

for both 2011 and 2012. This contamination is computed taking as global $K^*\eta$ efficiency, $\mathcal{E}^{K^*\gamma}_{K^*\eta}$, the product of the individual efficiencies of Tab. 8.1, while the other global efficiencies are taken from Tab. 5.20. The associated shape to be included in the fit is assumed to be the same as the one extracted for $K^*\eta$ reconstructed as $B^0 \to K^{*0}\gamma$. Even though the precise $\eta\rho$ shape cannot be determined, such choice is realistic because the $\eta\rho$ kinematics in the $B^0 \to \rho^0\gamma$ spectrum are expected to be very similar to the ones of $K^*\eta$ in the $B^0 \to K^{*0}\gamma$ spectrum.

The projections of the 2011+2012 unblinded simultaneous mass fit to the $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ mass spectra for 2011 and 2012 data samples are shown respectively in Fig. 8.1 and 8.2 for $B^0 \to \rho^0\gamma$ and in Fig. 8.3 for $B^0 \to K^{*0}\gamma$ (linear scale only). The values of the parameters allowed to float in the fit, including the ratio $r^{Eff}_{\rho\gamma}$ which is now

Channel	Year	ϵ^{MC} (%)	ϵ^{Sel} (%)	ϵ^{PID} (%)	$r^{Eff}_{K^*\eta}$	$\mathcal{C}^{K^*\gamma}_{K^*\eta}$ (%)
$B^0 \to K^{*0}\eta(\gamma\gamma)$	2011	16.640 ± 0.037	0.177 ± 0.002	66.710 ± 0.289	0.145 ± 0.010	2.378 ± 0.173
	2012		0.146 ± 0.002	66.030 ± 0.242		2.351 ± 0.170

Table 8.1: Summary table of the individual efficiencies and $r^{Eff}_{K^*\eta}$ used to derive the $B^0 \to K^{*0}\eta(\gamma\gamma)$ relative contamination $\mathcal{C}^{K^*\gamma}_{K^*\eta}$ in the $B^0 \to K^{*0}\gamma$ signal hypothesis when restricting the B^0 mass range to $[4800;\ 6600]$ MeV/c^2.

unblinded, are reported in Tab. 8.2. The μ, σ_L and σ_R for the $B^0 \to \rho^0\gamma$ channel are not displayed as they are constrained to the $B^0 \to K^{*0}\gamma$ parameters.

As can be seen from the residuals under the mass plots, the agreement between the data and the individual fit functions is very good in both mass spectra and years. As expected, all the fitted parameters for both $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ modes are compatible with the ones shown in Tab. 7.3.

Channel	Parameter		Year	Signal mode	
				$B^0 \to \rho^0\gamma$	$B^0 \to K^{*0}\gamma$
	$r_{\rho\gamma}^{Eff}$			0.0417 ± 0.0027	
Signals	N_S		2011		3405 ± 67
			2012		11550 ± 119
	μ	MeV/c^2	2011		5275.5 ± 3.4
			2012		5282.6 ± 2.2
	σ_L	MeV/c^2	2011		94.0 ± 3.9
			2012		91.8 ± 1.5
	σ_R	MeV/c^2	2011		105.7 ± 3.0
			2012		83.1 ± 0.9
Combinatorial	N_{Comb}		2011	142 ± 21	42 ± 26
			2012	355 ± 35	201 ± 35
	p_0	$(\text{MeV}/c^2)^{-1}$	2011	-0.87 ± 0.11	-0.35 ± 0.62
			2012	-0.96 ± 0.07	-0.94 ± 0.08
$B^0 \to \pi^+\pi^-\gamma X$	N_{Part}		2011	389 ± 25	–
			2012	1241 ± 42	–
	c		2011	-5.70 ± 2.37	–
			2012	-4.31 ± 1.17	–
$B^0 \to K^+\pi^-\gamma X$	N_{Part}		2011	–	728 ± 47
			2012	–	2229 ± 75
	c		2011	–	-3.24 ± 2.41
			2012	–	-0.59 ± 1.20

Table 8.2: Summary of the free parameters entering the fit model and values extracted fitting the $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ mass spectra simultaneously on 2011 and 2012 data. The measured value of the ratio $r_{\rho\gamma}^{Eff}$ is unblinded.

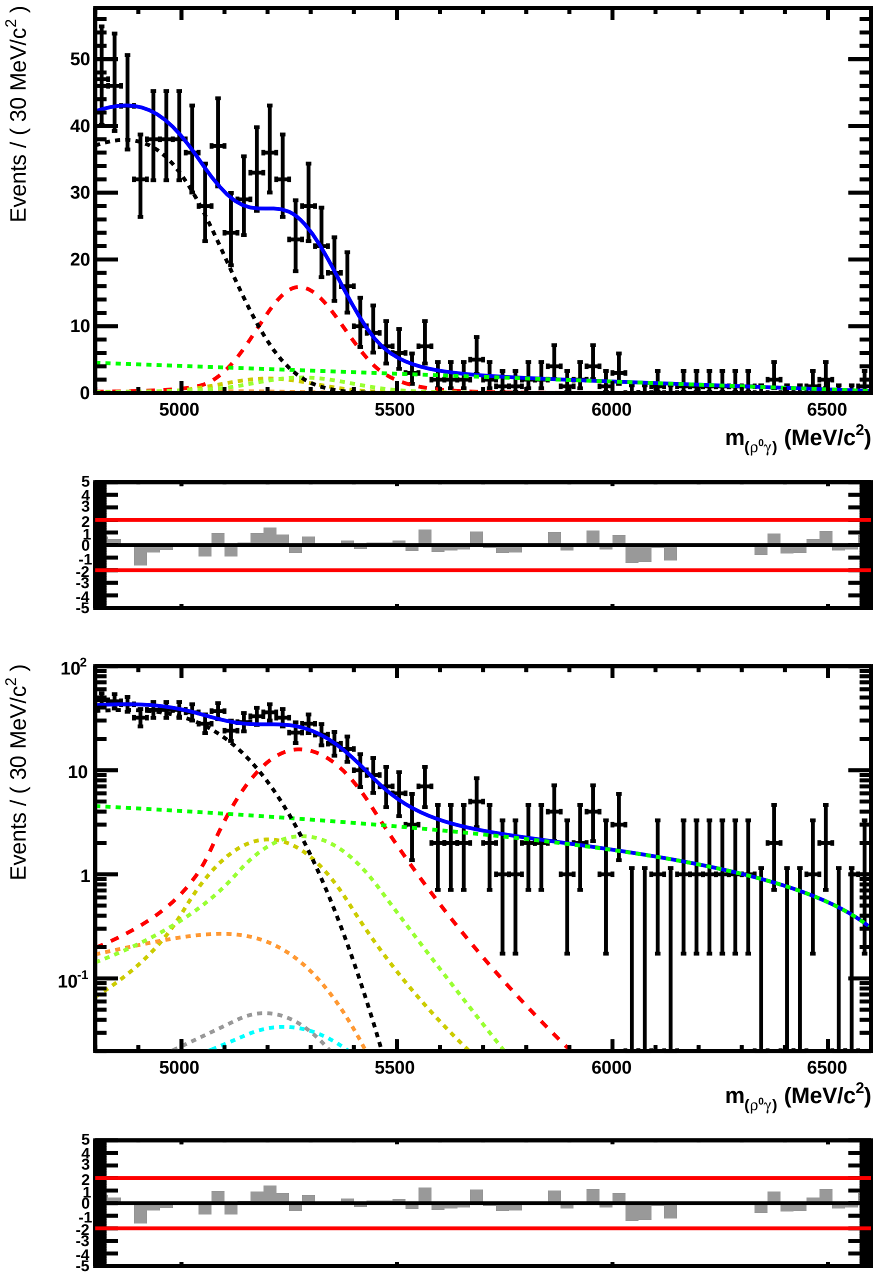


Figure 8.1: $B^0 \to \rho^0\gamma$ 2011+2012 simultaneous mass fit projections for 2011 data in linear (top) and semi-logarithmic scale (bottom). The black points represent the data and the solid blue line is the total fit function. The signal is fitted with an asymmetric double-tail Crystal Ball PDF and the combinatorial background is described with a first order polynomial (dashed green line). The peaking background modes are $B^0 \to K^{*0}\gamma$ (dashed dark yellow line), $\Lambda_b^0 \to \Lambda^*\gamma$ (dashed cyan line), $B^0 \to \pi^+\pi^-\pi^0$ (dashed light green line) and $B^0 \to K^+\pi^-\pi^0$ (dashed gray line). The partially reconstructed modes are $B^0 \to \eta(\gamma\gamma)\rho^0$ (dashed orange line) and $B^0 \to \pi^+\pi^-\gamma X$ (dashed black line).

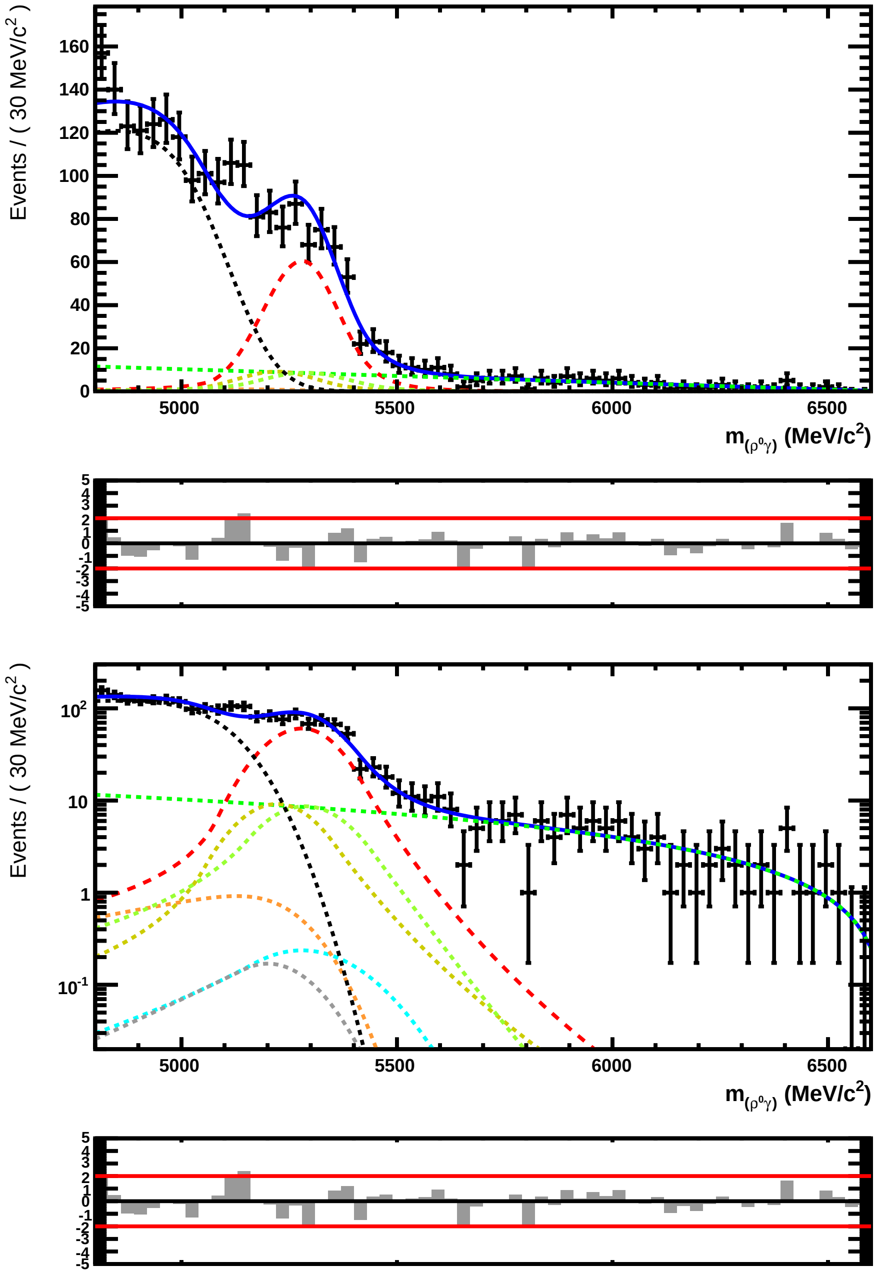


Figure 8.2: $B^0 \to \rho^0\gamma$ 2011+2012 simultaneous mass fit projections for 2012 data in linear (top) and semi-logarithmic scale (bottom). The black points represent the data and the solid blue line is the total fit function. The signal is fitted with an asymmetric double-tail Crystal Ball PDF and the combinatorial background is described with a first order polynomial (dashed green line). The peaking background modes are $B^0 \to K^{*0}\gamma$ (dashed dark yellow line), $\Lambda_b^0 \to \Lambda^*\gamma$ (dashed cyan line), $B^0 \to \pi^+\pi^-\pi^0$ (dashed light green line) and $B^0 \to K^+\pi^-\pi^0$ (dashed gray line). The partially reconstructed modes are $B^0 \to \eta(\gamma\gamma)\rho^0$ (dashed orange line) and $B^0 \to \pi^+\pi^-\gamma X$ (dashed black line).

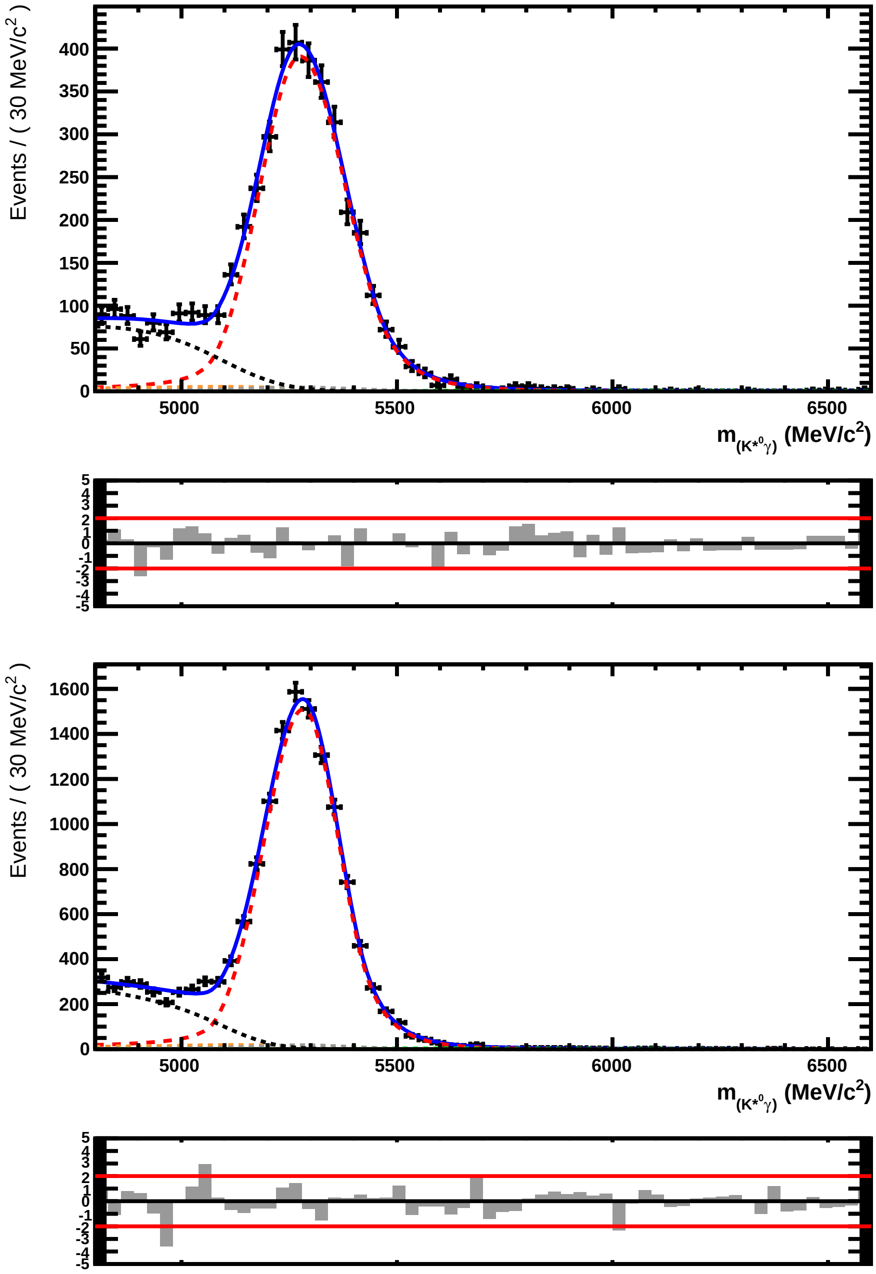


Figure 8.3: $B^0 \to K^{*0}\gamma$ 2011+2012 simultaneous mass fit projections for 2011 (top) and 2012 (bottom) data in linear scale. The black points represent the data and the solid blue line is the total fit function. The signal is fitted with an asymmetric double-tail Crystal Ball PDF and the combinatorial background is described with a first order polynomial (dashed green line). The peaking background modes are $B^0 \to \rho^0\gamma$ (dashed dark yellow line), $B^0_s \to \phi\gamma$ (dashed dark blue line), $\Lambda^0_b \to \Lambda^*\gamma$ (dashed cyan line), $B^0 \to K^+\pi^-\pi^0$ (dashed gray line) and $B^0 \to \pi^+\pi^-\pi^0$ (dashed light green line). The partially reconstructed modes are $B^0 \to K^{*0}\eta(\gamma\gamma)$ (dashed orange line) and $B^0 \to K^+\pi^-\gamma X$ (dashed black line).

The measured value of the ratio of branching fractions is $r_{\rho\gamma}^{Eff} = (4.17 \pm 0.27) \times 10^{-2}$, to be compared with $r_{\rho\gamma}^{Eff}(\mathrm{PDG}) = (2.99 \pm 0.53) \times 10^{-2}$. We can notice that the measured value is 2σ away from the PDG one.

In order to verify the consistency of the measurement, after fitting the data we calculate signal weights and check the reconstructed $\pi^+\pi^-$ and $K^+\pi^-$ mass spectra for events passing the $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ selections respectively.

The reconstructed $K^+\pi^-$ mass distributions in sweighted data and MC are shown in Fig. 8.4. We can see that for $B^0 \to K^{*0}\gamma$ signal events the mass distributions exhibit a very good data/MC agreement.

The reconstructed $\pi^+\pi^-$ mass distributions for sweighted data allows to crosscheck the understanding of the backgrounds for $B^0 \to \rho^0\gamma$ events. Indeed, given the initial choice of a particularly large ρ^0 mass window, $[400; 1200]$ MeV/c^2, different backgrounds may contribute to the $\pi^+\pi^-$ mass resulting in a poor data/MC agreement. For this comparison, the relative fraction of $B^0 \to K^{*0}\gamma$ events is not negligible and has to be included in the description. The shapes associated to the ρ^0 and K^{*0} mesons are extracted through a fit to the $\pi^+\pi^-$ mass distribution of $\rho\gamma$ and $K^*\gamma$ MC events reconstructed as $B^0 \to \rho^0\gamma$ using an asymmetric double tail Crystal-Ball PDF. The $\pi^+\pi^-$ invariant mass distributions are shown in Fig. 8.5. At high $\pi^+\pi^-$ mass, the data exhibit an excess of events with respect to what would be expected just for $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$. This excess is well modeled considering two additional contributions associated to the $f_0(980)$ and $f_2(1270)$ resonances. These are described by relativistic P-wave Breit-Wigner distributions [86]

$$BW(m) = \frac{m_0 \cdot m \cdot \Gamma(m)}{(m_0^2 - m^2)^2 + (m_0 \cdot \Gamma(m))^2}, \tag{8.2}$$

$$\Gamma(m) = \Gamma_0 \cdot \left(\frac{q}{q_0}\right)^{(2l+1)} \cdot \frac{m_0}{m}, \tag{8.3}$$

where m_0 is the resonance mass, $\Gamma_0 \equiv \Gamma(m_0)$ its natural width, l the transferred angular

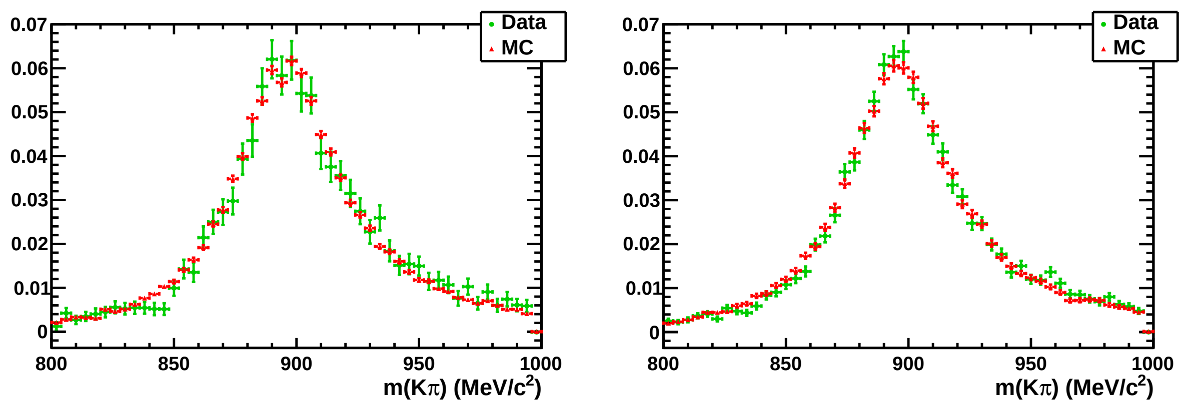


Figure 8.4: 2011 (left) and 2012 (right) reconstructed $K^+\pi^-$ invariant mass distributions for $B^0 \to K^{*0}\gamma$ signal events. The green dots stand for sweighted data and red triangles for signal MC.

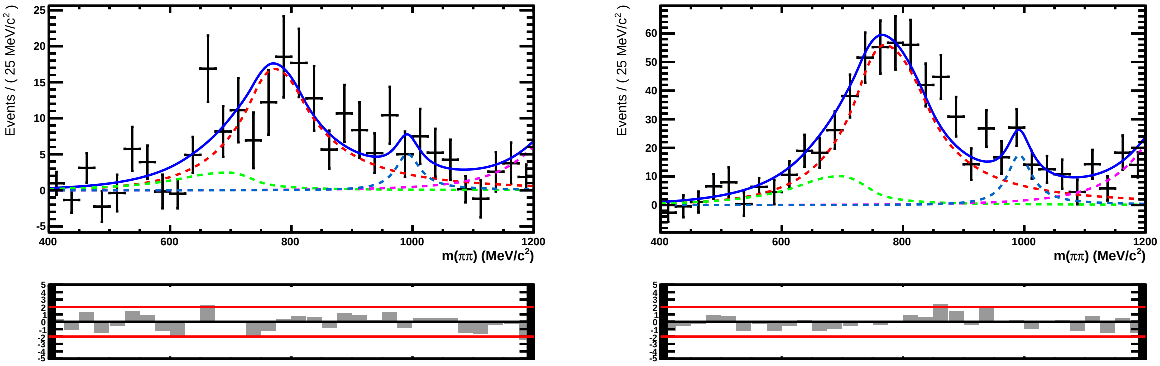


Figure 8.5: 2011 (left) and 2012 (right) fit of the reconstructed $\pi^+\pi^-$ invariant mass distributions for $B^0 \to \rho^0\gamma$ signal events. The black points represent the data and the solid blue line is the total fit function. The ρ^0 (dashed red line) and K^{*0} (dashed green line) shapes are both described with an asymmetric double-tail Crystal Ball PDF, while the $f_0(980)$ (dashed light blue line) and $f_2(1270)$ (dashed magenta line) resonances are described by two Breit-Wigner PDFs with free independent fractions.

momentum (0 for the $f_0(980)$, 2 for the $f_2(1270)$), $q(m)$ the momentum of the decay products in the rest frame of the mother particle and $q_0 \equiv q(m_0)$. Since we are dealing with a $\pi^+\pi^-$ final states, the expression of $q(m)$ is given by

$$q(m) = \frac{\sqrt{m^2 - 4m_\pi^2}}{2}. \tag{8.4}$$

The masses and the natural widths of the $f_0(980)$ and $f_2(1270)$ resonances are fixed to their respective PDG values. The two PDFs are included in the description of the reconstructed $\pi^+\pi^-$ mass using independent fractions allowed to float in the fit to the sweighted samples. The $B^0 \to K^{*0}\gamma$ fraction is fixed according to the various contaminations used in the simultaneous fit to the data as $\mathcal{C}^{\rho\gamma}_{K^*\gamma}/(1 + \sum_i \mathcal{C}^{\rho\gamma}_i)$, where the various contaminations $\mathcal{C}^{\rho\gamma}_i$ contributing to the sum are reported in Tab. 5.21 and Eq. 8.1. As one may notice the dominant contributions are $B^0 \to K^{*0}\gamma$ and $B^0 \to \pi^+\pi^-\pi^0$. However, the $\pi^+\pi^-\pi^0$ contamination is mainly associated to the $B^0 \to \rho^0\pi^0$ mode giving rise to the same $\pi^+\pi^-$ mass spectrum as for the $B^0 \to \rho^0\gamma$ signal mode. Therefore no dedicated shape is needed for $B^0 \to \pi^+\pi^-\pi^0$.

The fit to the $\pi^+\pi^-$ sweighted data distributions is good but is only considered as indicative of the fact that higher resonances may contribute given the large $\pi^+\pi^-$ mass window initially used. To reject them, we decided to set a higher cut at 925 MeV/c^2 given that, according to the fit, they seem to be negligible below this value. We also set a lower cut at 625 MeV/c^2 as the $B^0 \to K^{*0}\gamma$ contamination dominates at lower masses. Therefore the new baseline ρ^0 mass window is [625; 925] MeV/c^2.

We now re-evaluate all the efficiencies and contaminations to the $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ spectra taking into account the restricted ρ^0 mass window and the reduced B^0 mass range [4800; 6600] MeV/c^2. The values of the $(\epsilon^{MC}\epsilon^{Sel})$ products for the different $\Lambda_b^0 \to \Lambda^{*0}\gamma$ MC samples together with the values of $f_{\Lambda^*\gamma}^{signal}$ and the relative uncertainty $\Delta f_{\Lambda^*\gamma}^{signal}/f_{\Lambda^*\gamma}^{signal}$ for both $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ hypotheses are displayed in Tab. 8.3. The individual efficiencies are given in Tab. 8.4 and 8.5 for the $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ hypotheses respectively. The global efficiencies are reported in Tab. 8.6 together with the $r_{H_b \to X}^{Eff}$ ratios. The new contaminations together with the values of the $\mathcal{D}_{H_b \to X}^{signal}$ terms are reported in Tab. 8.7.

Hypothesis		Channel	$\epsilon^{MC}\epsilon^{Sel}$ (%) 2011	$\epsilon^{MC}\epsilon^{Sel}$ (%) 2012
$B^0 \to \rho^0\gamma$		$\Lambda_b^0 \to \Lambda^*(1520)^0\gamma$	0.015 ± 0.001	0.013 ± 0.001
		$\Lambda_b^0 \to \Lambda^*(1670)^0\gamma$	0.124 ± 0.001	0.101 ± 0.001
		$\Lambda_b^0 \to \Lambda^*(1820)^0\gamma$	0.034 ± 0.001	0.028 ± 0.001
		$\Lambda_b^0 \to \Lambda^*(1830)^0\gamma$	0.032 ± 0.001	0.028 ± 0.001
$f_{\Lambda^*\gamma}^{\rho\gamma}$	(%)		467.16 ± 2.93	433.20 ± 2.82
$\Delta f_{\Lambda^*\gamma}^{\rho\gamma}/f_{\Lambda^*\gamma}^{\rho\gamma}$	(%)		78.59 ± 0.80	76.92 ± 0.82
$B^0 \to K^{*0}\gamma$		$\Lambda_b^0 \to \Lambda^*(1520)^0\gamma$	0.074 ± 0.001	0.059 ± 0.001
		$\Lambda_b^0 \to \Lambda^*(1670)^0\gamma$	0.034 ± 0.001	0.028 ± 0.001
		$\Lambda_b^0 \to \Lambda^*(1820)^0\gamma$	0.017 ± 0.001	0.013 ± 0.001
		$\Lambda_b^0 \to \Lambda^*(1830)^0\gamma$	0.016 ± 0.001	0.013 ± 0.001
$f_{\Lambda^*\gamma}^{K^*\gamma}$	(%)		60.86 ± 1.74	61.22 ± 1.71
$\Delta f_{\Lambda^*\gamma}^{K^*\gamma}/f_{\Lambda^*\gamma}^{K^*\gamma}$	(%)		64.30 ± 3.40	63.35 ± 3.31

Table 8.3: Summary of the new $\epsilon^{MC}\epsilon^{Sel}$ efficiency products for the different $\Lambda^*\gamma$ resonances and values of $f_{\Lambda^*\gamma}^{signal}$ and $\Delta f_{\Lambda^*\gamma}^{signal}/f_{\Lambda^*\gamma}^{signal}$ for both the $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ signal hypotheses.

Channel	Year	ϵ^{MC} (%)	ϵ^{Sel} (%)	ϵ^{PID} (%)
$B^0 \to \rho^0\gamma$	2011	20.322 ± 0.030	0.744 ± 0.007	62.760 ± 0.205
	2012	20.593 ± 0.063	0.614 ± 0.005	62.350 ± 0.174
$B^0 \to K^{*0}\gamma$	2011	21.274 ± 0.037	0.531 ± 0.002	0.445 ± 0.002
	2012	21.479 ± 0.055	0.456 ± 0.003	0.449 ± 0.003
$\Lambda_b^0 \to \Lambda^{*0}\gamma$	2011	23.073 ± 0.050	0.064 ± 0.002	0.023 ± 0.003
	2012	23.266 ± 0.080	0.057 ± 0.001	0.041 ± 0.002
$B^0 \to \pi^+\pi^-\pi^0$	2012	33.015 ± 0.078	0.006 ± 0.001	29.890 ± 1.696
$B^0 \to K^+\pi^-\pi^0$	2012	17.279 ± 0.064	0.011 ± 0.001	0.235 ± 0.024
$\Lambda_b^0 \to \Lambda^*(1670)^0\gamma$	2011	25.142 ± 0.058	0.492 ± 0.005	
	2012	25.355 ± 0.048	0.398 ± 0.005	
$\Lambda_b^0 \to \Lambda^*(1820)^0\gamma$	2011	39.952 ± 0.075	0.086 ± 0.002	
	2012	39.980 ± 0.100	0.069 ± 0.003	
$\Lambda_b^0 \to \Lambda^*(1830)^0\gamma$	2011	39.958 ± 0.075	0.081 ± 0.002	
	2012	39.885 ± 0.110	0.070 ± 0.003	

Table 8.4: Summary table of all the new individual efficiencies in the $B^0 \to \rho^0\gamma$ signal hypothesis. Let us remind that the PID efficiencies for the $\Lambda^*(1670)^0\gamma$, $\Lambda^*(1820)^0\gamma$ and $\Lambda^*(1830)^0\gamma$ modes are not needed since they are used only to calculate the $f^{\rho\gamma}_{\Lambda^*\gamma}$ term.

Channel	Year	ϵ^{MC} (%)	ϵ^{Sel} (%)	ϵ^{PID} (%)
$B^0 \to K^{*0}\gamma$	2011	21.274 ± 0.037	0.791 ± 0.003	71.130 ± 0.076
	2012	21.479 ± 0.055	0.648 ± 0.003	70.800 ± 0.108
$B^0 \to \rho^0\gamma$	2011	20.322 ± 0.030	0.552 ± 0.006	7.125 ± 0.100
	2012	20.593 ± 0.063	0.459 ± 0.004	7.603 ± 0.092
$B_s^0 \to \phi\gamma$	2011	23.203 ± 0.041	0.312 ± 0.003	0.355 ± 0.003
	2012	23.519 ± 0.059	0.284 ± 0.002	0.381 ± 0.003
$\Lambda_b^0 \to \Lambda^{*0}\gamma$	2011	23.073 ± 0.050	0.320 ± 0.004	2.354 ± 0.072
	2012	23.266 ± 0.080	0.254 ± 0.003	2.970 ± 0.091
$B^0 \to K^+\pi^-\pi^0$	2012	17.279 ± 0.064	0.023 ± 0.001	29.100 ± 0.869
$B^0 \to \pi^+\pi^-\pi^0$	2012	33.015 ± 0.078	0.005 ± 0.001	3.505 ± 0.456
$\Lambda_b^0 \to \Lambda^*(1670)^0\gamma$	2011	25.142 ± 0.058	0.135 ± 0.003	
	2012	25.355 ± 0.048	0.109 ± 0.002	
$\Lambda_b^0 \to \Lambda^*(1820)^0\gamma$	2011	39.952 ± 0.075	0.042 ± 0.002	
	2012	39.980 ± 0.100	0.033 ± 0.001	
$\Lambda_b^0 \to \Lambda^*(1830)^0\gamma$	2011	39.958 ± 0.075	0.040 ± 0.001	
	2012	39.885 ± 0.110	0.034 ± 0.001	

Table 8.5: Summary table of all the new individual efficiencies in the $B^0 \to K^{*0}\gamma$ signal hypothesis. Let us remind that the PID efficiencies for the $\Lambda^*(1670)^0\gamma$, $\Lambda^*(1820)^0\gamma$ and $\Lambda^*(1830)^0\gamma$ modes are not needed since they are used only to calculate the $f_{\Lambda^*\gamma}^{K^*\gamma}$ term.

Channel	Year	$\mathcal{E}^{\rho\gamma}_{H_b\to X}$ $(\times 10^{-4})$	$\mathcal{E}^{K^*\gamma}_{H_b\to X}$ $(\times 10^{-4})$	$r^{Eff}_{H_b\to X}$
$B^0 \to \rho^0\gamma$	2011	9.4897 ± 0.0938	0.7994 ± 0.0142	0.0299 ± 0.0053
	2012	7.8785 ± 0.0673	0.7188 ± 0.0109	
$B^0 \to K^{*0}\gamma$	2011	0.0502 ± 0.0004	11.9687 ± 0.0505	1
	2012	0.0440 ± 0.0004	9.8578 ± 0.0578	
$B^0_s \to \phi\gamma$	2011	–	0.0257 ± 0.0003	0.1596 ± 0.0095
	2012	–	0.0254 ± 0.0003	
$\Lambda^0_b \to \Lambda^{*0}\gamma$	2011	0.0016 ± 0.0013	0.1059 ± 0.0682	0.4468 ± 0.0227
	2012	0.0023 ± 0.0021	0.1076 ± 0.0682	
$B^0 \to \pi^+\pi^-\pi^0$	2012	0.0616 ± 0.0065	0.0061 ± 0.0010	0.8681 ± 0.0886
$B^0 \to K^+\pi^-\pi^0$	2012	0.0005 ± 0.0001	0.1176 ± 0.0072	1.3125 ± 0.1201
$B^0 \to \eta(\gamma\gamma)\rho^0$	2011	1.5584 ± 0.0282	–	0.0074 ± 0.0026
	2012	1.2790 ± 0.0243	–	
$B^0 \to K^{*0}\eta(\gamma\gamma)$	2011	–	1.9655 ± 0.0287	0.145 ± 0.010
	2012	–	1.6003 ± 0.0192	

Table 8.6: New global efficiencies and values of the $r^{Eff}_{H_b\to X}$ ratios. For the $\Lambda^0_b \to \Lambda^{*0}\gamma$ mode, the $\mathcal{E}^{signal}_{\Lambda^*\gamma}$ global efficiencies include the $f^{signal}_{\Lambda^*\gamma}$ term and its uncertainty.

Signal	$\mathcal{C}^{signal}_{H_b\to X}$	Year	$\mathcal{D}^{signal}_{H_b\to X}$ (%)
$B^0\to\rho^0\gamma$	$\mathcal{D}^{\rho\gamma}_{K^*\gamma}\times\mathcal{Y}^{-1}$	2011	0.420 ± 0.004
		2012	0.446 ± 0.005
	$\mathcal{D}^{\rho\gamma}_{\Lambda^*\gamma}\times\mathcal{Y}^{-1}$	2011	0.006 ± 0.005
		2012	0.011 ± 0.009
	$\mathcal{D}^{\rho\gamma}_{\pi\pi\pi^0}\times\mathcal{Y}^{-1}$	2012	0.543 ± 0.080
	$\mathcal{D}^{\rho\gamma}_{K\pi\pi^0}\times\mathcal{Y}^{-1}$	2012	0.006 ± 0.001
	$\mathcal{D}^{\rho\gamma}_{\eta\rho}\times\mathcal{Y}^{-1}$	2011 2012	0.096 ± 0.0057
$B^0\to K^{*0}\gamma$	$\mathcal{D}^{K^*\gamma}_{\rho\gamma}\times\mathcal{Y}$	2011	8.424 ± 0.171
		2012	9.123 ± 0.159
	$\mathcal{D}^{K^*\gamma}_{\phi\gamma}$	2011	0.034 ± 0.002
		2012	0.041 ± 0.003
	$\mathcal{D}^{K^*\gamma}_{\Lambda^*\gamma}$	2011	0.395 ± 0.255
		2012	0.488 ± 0.317
	$\mathcal{D}^{K^*\gamma}_{K\pi\pi^0}$	2012	1.566 ± 0.172
	$\mathcal{D}^{K^*\gamma}_{\pi\pi\pi^0}$	2012	0.054 ± 0.010
	$\mathcal{D}^{K^*\gamma}_{K^*\eta}$	2011	2.378 ± 0.173
		2012	2.351 ± 0.170

Table 8.7: Summary table gathering all the new values of the $\mathcal{D}^{signal}_{H_b\to X}$ terms to be used to constrain the background contaminations in the simultaneous fit. The expressions of the contaminations are also reported. $\mathcal{Y}$ it the ratio of the fitted signal yields for $B^0\to\rho^0\gamma$ and $B^0\to K^{*0}\gamma$: $\mathcal{Y}=N^{\rho\gamma}/N^{K^*\gamma}$.

The projections of the new 2011+2012 unblinded simultaneous mass fit to the $B^0\to\rho^0\gamma$ and $B^0\to K^{*0}\gamma$ mass spectra for 2011 and 2012 data samples are shown respectively in Fig. 8.6 and 8.7 for $B^0\to\rho^0\gamma$ and in Appendix C for $B^0\to K^{*0}\gamma$. The values of the parameters allowed to float in the fit are reported in Tab. 8.8. The μ, σ_L and σ_R for the $B^0\to\rho^0\gamma$ channel are not displayed as they are constrained to the $B^0\to K^{*0}\gamma$ parameters.

As can be seen from the residuals under the mass plots, the agreement between the data and the individual fit functions is very good in both mass spectra and years. The fitted parameters for both $B^0\to\rho^0\gamma$ and $B^0\to K^{*0}\gamma$ modes are compatible with the ones shown in Tab. 8.2. However, thanks to the restricted ρ^0 mass window, we can notice that in the $B^0\to\rho^0\gamma$ spectrum the combinatorial background is reduced by a factor three

Channel	Parameter		Year	Signal mode $B^0 \to \rho^0\gamma$	$B^0 \to K^{*0}\gamma$
Signals	$r^{Eff}_{\rho\gamma}$			0.0352±0.0025	
	N_S		2011		3405 ± 68
			2012		11561 ± 120
	μ	MeV/c^2	2011		5275.8 ± 5.1
			2012		5282.6 ± 2.3
	σ_L	MeV/c^2	2011		94.2 ± 4.8
			2012		91.6 ± 2.3
	σ_R	MeV/c^2	2011		105.4 ± 3.7
			2012		83.2 ± 1.6
Combinatorial	N_{Comb}		2011	52 ± 14	43 ± 27
			2012	109 ± 20	201 ± 36
	p_0	(MeV/c^2)$^{-1}$	2011	-0.99 ± 0.16	-0.37 ± 0.61
			2012	-1.05 ± 0.08	-0.94 ± 0.08
$B^0 \to \pi^+\pi^-\gamma X$	N_{Part}		2011	197 ± 18	–
			2012	622 ± 30	–
	c		2011	-4.82 ± 3.32	–
			2012	-4.38 ± 1.65	–
$B^0 \to K^+\pi^-\gamma X$	N_{Part}		2011	–	727 ± 48
			2012	–	2233 ± 77
	c		2011	–	-3.22 ± 2.47
			2012	–	-0.69 ± 1.26

Table 8.8: Summary of the free parameters entering the fit model and values extracted fitting the $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ mass spectra simultaneously on 2011 and 2012 data. Events in the $B^0 \to \rho^0\gamma$ mass spectrum are selected in the restricted ρ^0 mass window [625; 925] MeV/c^2.

and the partially reconstructed background is reduced by a factor two. Correlatively the signal peaks are more evident. The measured value of the ratio of branching fractions, $r^{Eff}_{\rho\gamma} = (3.52 \pm 0.25) \times 10^{-2}$, is compatible with the combination of the BaBar and Belle measurements, $r^{Eff}_{\rho\gamma}(\text{PDG}) = (2.99 \pm 0.53) \times 10^{-2}$. It corresponds to about 95 and 325 $B^0 \to \rho^0\gamma$ signal events for 2011 and 2012 respectively.

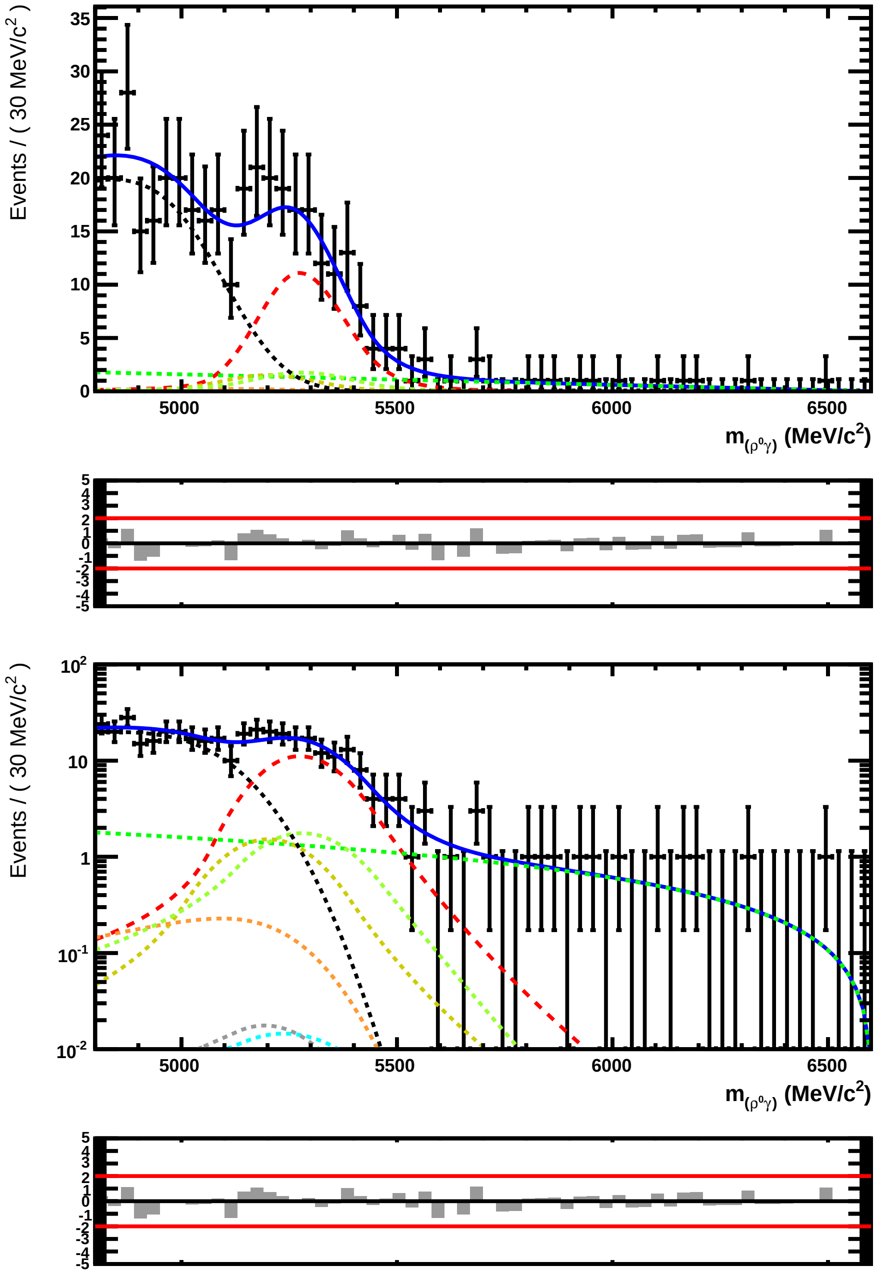


Figure 8.6: $B^0 \to \rho^0\gamma$ 2011+2012 simultaneous mass fit projections for 2011 data in linear (top) and semi-logarithmic scale (bottom). The black points represent the data and the solid blue line is the total fit function. The signal is fitted with an asymmetric double-tail Crystal Ball PDF and the combinatorial background is described with a first order polynomial (dashed green line). The peaking background modes are $B^0 \to K^{*0}\gamma$ (dashed dark yellow line), $\Lambda_b^0 \to \Lambda^*\gamma$ (dashed cyan line), $B^0 \to \pi^+\pi^-\pi^0$ (dashed light green line) and $B^0 \to K^+\pi^-\pi^0$ (dashed gray line). The partially reconstructed modes are $B^0 \to \eta(\gamma\gamma)\rho^0$ (dashed orange line) and $B^0 \to \pi^+\pi^-\gamma X$ (dashed black line).

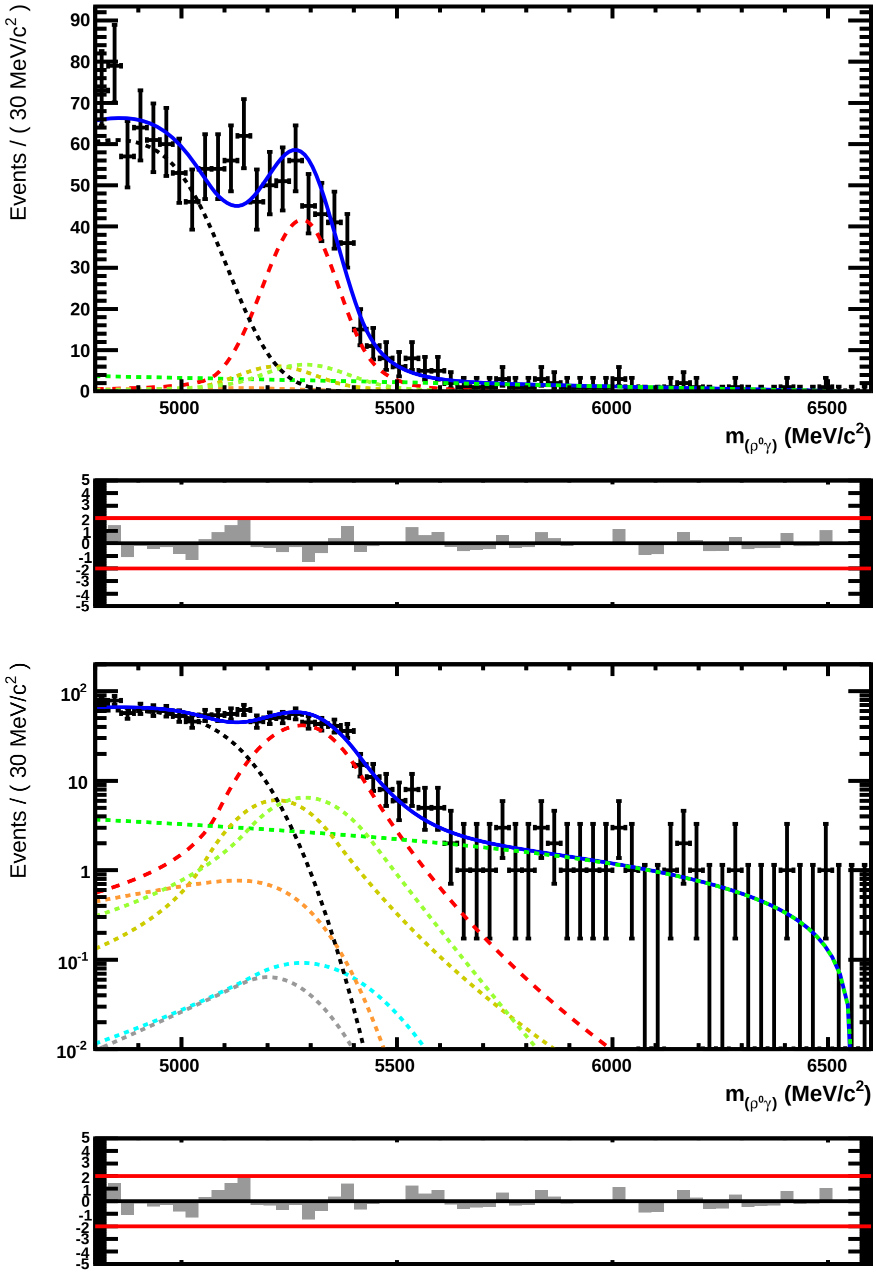


Figure 8.7: $B^0 \to \rho^0\gamma$ 2011+2012 simultaneous mass fit projections for 2012 data in linear (top) and semi-logarithmic scale (bottom). The black points represent the data and the solid blue line is the total fit function. The signal is fitted with an asymmetric double-tail Crystal Ball PDF and the combinatorial background is described with a first order polynomial (dashed green line). The peaking background modes are $B^0 \to K^{*0}\gamma$ (dashed dark yellow line), $\Lambda_b^0 \to \Lambda^*\gamma$ (dashed cyan line), $B^0 \to \pi^+\pi^-\pi^0$ (dashed light green line) and $B^0 \to K^+\pi^-\pi^0$ (dashed gray line). The partially reconstructed modes are $B^0 \to \eta(\gamma\gamma)\rho^0$ (dashed orange line) and $B^0 \to \pi^+\pi^-\gamma X$ (dashed black line).

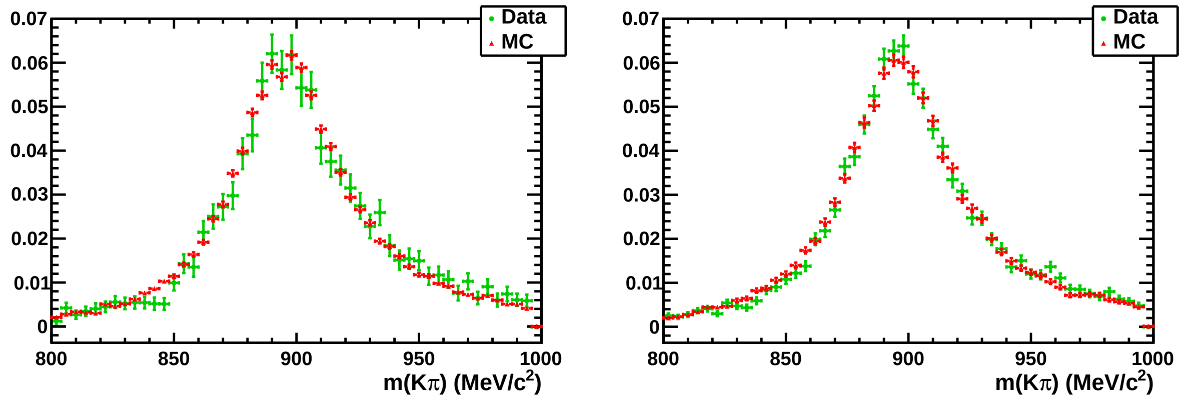


Figure 8.8: 2011 (left) and 2012 (right) reconstructed $K^+\pi^-$ invariant mass distributions for $B^0 \to K^{*0}\gamma$ signal events. The green dots stand for sweighted data and red triangles for signal MC.

We calculate the new signal weights and check again the reconstructed $\pi^+\pi^-$ and $K^+\pi^-$ mass spectra. The reconstructed $K^+\pi^-$ mass distributions in sweighted data and MC are shown in Fig. 8.8 and as shown previously the mass distributions are in a very good agreement. The reconstructed $\pi^+\pi^-$ mass distributions are displayed in Fig. 8.9 where now the overall modelisation only consists of the sum of the ρ^0 and K^{*0} shapes. The $B^0 \to K^{*0}\gamma$ fraction is still fixed according to the contaminations used in the simultaneous fit to the data (see Tab. 8.7): there is no free parameter in the modelisation. The description of the $\pi^+\pi^-$ mass spectrum is suitable, meaning that the understanding of the backgrounds seems correct.

As a final test, we check the helicity distributions of the ρ^0 and K^{*0} vector mesons in order to verify if they follow a $\sin^2\theta_H$ function as it should for a $B \to V\gamma$ decay. The distributions of the cosine of the helicity angles for $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ are

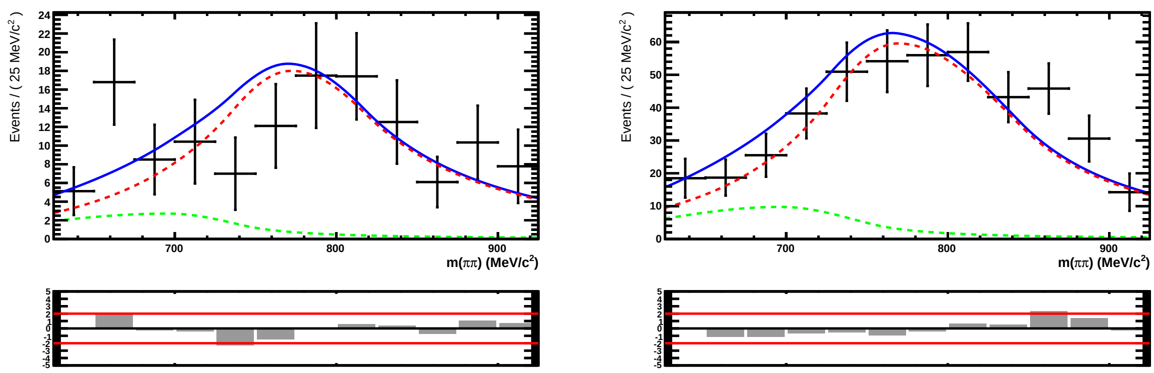


Figure 8.9: 2011 (left) and 2012 (right) reconstructed $\pi^+\pi^-$ invariant mass distributions for $B^0 \to \rho^0\gamma$ signal events. The black points represent the data and the solid blue line is the overall modelisation. The ρ^0 (dashed red line) and K^{*0} (dashed green line) shapes are both described with an asymmetric double-tail Crystal Ball PDF.

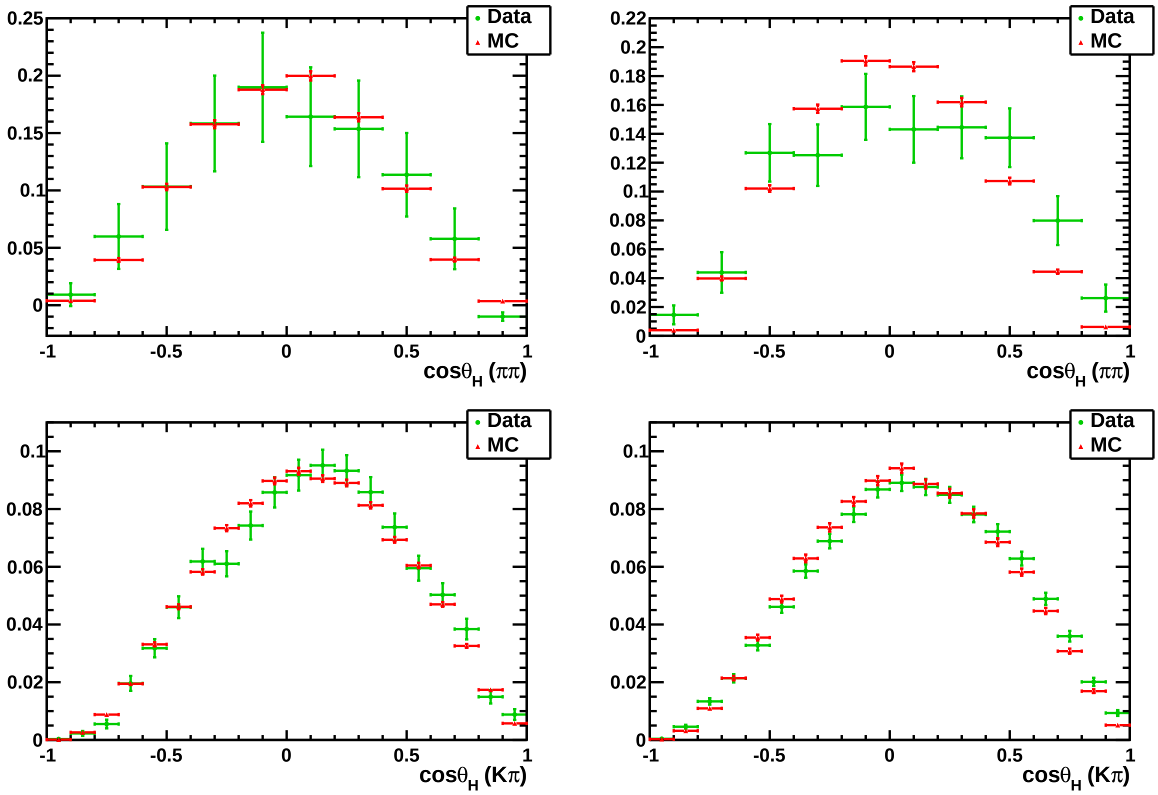


Figure 8.10: 2011 (left) and 2012 (right) helicity distributions for $B^0 \to \rho^0\gamma$ (top) and $B^0 \to K^{*0}\gamma$ (bottom) signal events. The green dots stand for sweighted data and red triangles for signal MC.

shown in Fig. 8.10 for sweighted data and MC. The distributions exhibit a good data/MC agreement.

8.2 Systematic uncertainties and results

The definitive study of systematic uncertainties requires the validation of the various steps of the analysis and the definitive approval of the fit model to be carried out during the official LHCb review process. However a preliminary estimation based on the sources which are expected to be the most relevant can be performed. In most cases the quantitative estimate of the systematic uncertainties relies on toy experiments, but at this level we perform series of independent fits to the real data in order to be as conservative as possible. After each fit we note the variation of the $r_{\rho\gamma}^{Eff}$ measured value and take it as the uncertainty associated to the source we are considering. The systematic uncertainties that we evaluated are related to the following.

- **Background contaminations constrained in the fit**. Each contribution is calculated increasing and reducing the respective contamination by the associated relative uncertainty. The most relevant contribution is expected to be the one related to the $B^0 \to \pi^+\pi^-\pi^0$ decay. Indeed, in order to be conservative, we assumed for this contamination a variation of 25% that comes from the relative uncertainty on the branching ratio $\mathcal{B}(B^0 \to \rho^0\pi^0) = (2.0 \pm 0.5) \times 10^{-5}$. For what concerns the $B^0 \to \eta\rho^0$ mode, given it is unobserved, we considered two extreme cases: either no contamination at all or a contamination that would correspond to the experimental limit from BaBar instead of the theoretical prediction used nominally, the ratio between the two being almost a factor 3. The variations associated to the other contaminations are the relative uncertainties on the $\mathcal{D}^{signal}_{H_b \to X}$ terms of Tab. 8.7 calculated using the full precision and not the approximated values displayed in the table.

- **Signal, combinatorial and partially reconstructed background shapes**. The contributions associated to the signal and combinatorial background shapes are evaluated switching respectively to a symmetric double-tail Crystal Ball and to an exponential function. Since the only constraint on the partially reconstructed background shape is the ARGUS threshold, to evaluate this contributions we perform a fit using the $\pi^\pm$ mass as threshold instead of the π^0 mass.

Another source of systematic uncertainty, even if certainly less significant compared to the others, is the one related to the choice of the charged (see Tab. 5.8) and neutral (see Tab. 5.11) PID binning. Its contribution can be computed changing the binning scheme and re-evaluating the PID efficiencies. However this uncertainty is not calculated here.

The various systematic uncertainties together with the associated variations are shown in Tab. 8.9. The most significant contributions associated to the contaminations come from the $B^0 \to \pi^+\pi^-\pi^0$ and $B^0 \to \eta\rho^0$ modes. They respectively correspond to a 4% and 1% relative uncertainty on the ratio of branching fractions. For what concerns the systematic uncertainties related to the fit model the most relevant ones are those associated to the signal and background shapes. They respectively correspond to a 3% and 4% relative uncertainty on the ratio of branching fractions. It has to be mentioned that to evaluate the contribution of the combinatorial background we fixed the slope of the exponential in the $B^0 \to \rho^0\gamma$ spectrum to the value extracted fitting the left handed sideband in the range [4000; 4800] MeV/c^2. This choice is related to the fact that the competition between combinatorial and partially reconstructed background shapes (both free in the fit) is particularly relevant in the restricted B^0 mass range given there is no sufficient leverage for the exponential slope to be properly determined. For this reason, during the review process, it will be probably needed to extend the mass fit range in the $B^0 \to \rho^0\gamma$ spectrum to better control the combinatorial background shape.

The measured value of the effective ratio of branching fractions, including the systematic uncertainty, is

$$r^{Eff}_{\rho\gamma} = (3.52 \pm 0.25(\text{stat.}) \pm 0.24(\text{syst.})) \times 10^{-2}. \tag{8.5}$$

Combining the statistical and systematic uncertainties we get $r^{Eff}_{\rho\gamma} = (3.52 \pm 0.35) \times 10^{-2}$.

Type	Source	Variation	Uncertainty
$\mathcal{C}^{\rho\gamma}_{H_b\to X}$	$B^0\to K^{*0}\gamma$	$\pm$ 1.23%	+0.0051 / − 0.0061
	$\Lambda^0_b\to\Lambda^{*0}\gamma$	$\pm$ 90.10%	+0.0069 / − 0.0085
	$B^0\to\pi^+\pi^-\pi^0$	$\pm$ 25.00%	+0.1513 / − 0.1518
	$B^0\to K^+\pi^-\pi^0$	$\pm$ 15.70%	+0.0006 / − 0.0006
	$B^0\to\eta(\gamma\gamma)\rho^0$	$\times$(0 or 2.78)	+0.0336 / − 0.0609
$\mathcal{C}^{K^*\gamma}_{H_b\to X}$	$B^0\to\rho^0\gamma$	$\pm$ 1.82%	+0.0002 / − 0.0002
	$B^0_s\to\phi\gamma$	$\pm$ 6.10%	+0.0001 / − 0.0001
	$\Lambda^0_b\to\Lambda^{*0}\gamma$	$\pm$ 64.60%	+0.0125 / − 0.0104
	$B^0\to K^+\pi^-\pi^0$	$\pm$ 11.00%	+0.0071 / − 0.0068
	$B^0\to\pi^+\pi^-\pi^0$	$\pm$ 25.00%	+0.0003 / − 0.0005
	$B^0\to K^{*0}\eta(\gamma\gamma)$	$\pm$ 7.34%	+0.0016 / − 0.0026
Shapes	Signal	Symmetric DCB	+0.1020 / (−0.1020)
	Comb. Bkg.	Exponential	(+0.1416) / − 0.1416
	Part. Rec. Bkg.	ARGUS threshold	+0.0240 / (−0.0240)

Table 8.9: Summary table gathering all the systematic uncertainties together with the associated variations. The uncertainties related to the shapes are evaluated with a single fit. The values between parentheses do not come from the fits and mean that, to be conservative, we assumed a symmetric uncertainty when combining them with the other contributions.

This is compatible with the combination of BaBar and Belle measurements, $r^{Eff}_{\rho\gamma} = (2.99 \pm 0.53) \times 10^{-2}$, the absolute precision being 34% better.

Using Eq. 7.5 the ratio of branching ratios is found to be

$$\frac{\mathcal{B}(B^0\to\rho^0\gamma)}{\mathcal{B}(B^0\to K^{*0}\gamma)} = (2.34 \pm 0.23) \times 10^{-2}. \tag{8.6}$$

The relative precision is about 10%. This corresponds to an uncertainty of $\sim 5\%$ on the ratio $|V_{td}/V_{ts}|$ of CKM matrix elements.

Assuming for $B^0 \to K^{*0}\gamma$ the branching ratio reported on the PDG, $\mathcal{B}(B^0\to K^{*0}\gamma) = (4.33 \pm 0.15) \times 10^{-5}$, we get

$$\mathcal{B}(B^0\to\rho^0\gamma) = (1.01 \pm 0.11) \times 10^{-6} \tag{8.7}$$

to be compared to the PDG value $\mathcal{B}(B^0\to\rho^0\gamma) = (0.86 \pm 0.15) \times 10^{-6}$.

Conclusions

This thesis presents the analysis procedure developed to perform the measurement of the ratio of branching fractions $\mathcal{B}(B^0 \to \rho^0\gamma)/\mathcal{B}(B^0 \to K^{*0}\gamma)$. The full Run 1 statistics collected at the LHCb experiment is used, namely $\int \mathcal{L}\mathrm{d}t = 1\,\mathrm{fb}^{-1}$ at a center-of-mass energy of $\sqrt{s} = 7$ TeV for 2011 and $\int \mathcal{L}\mathrm{d}t = 2\,\mathrm{fb}^{-1}$ at $\sqrt{s} = 8$ TeV for 2012. The measurement is performed using the $B^0 \to K^{*0}\gamma$ decay as control mode.

This analysis is very challenging since the $B^0 \to \rho^0\gamma$ decay is a particularly rare mode and because we need to efficiently distinguish between high energy photons and neutral pions. The γ/π^0 separation is a crucial point since LHCb has not been specifically designed to accomplish it. As a part of the work carried out during this thesis, dedicated calibration samples that allow to evaluate the γ/π^0 separation efficiency on a data-driven basis have been produced. These samples are the official Run 1 calibration samples used in LHCb analyses involving high energy γ and π^0.

Aiming to observe the $B^0 \to \rho^0\gamma$ decay with the best possible significance, a thorough event selection that allow to efficiently reject the various sources of background has been implemented. In particular, special care has been taken in the conception of the strategy and in the optimization phase in order to reject background events. Peaking backgrounds are notably the most dangerous source of background as, lying in the signal region, they could lead to a significant over(under)-estimation of the fitted signal yields. It is then very important to describe their shapes in the mass fit and to precisely evaluate their actual contaminations to the invariant mass spectrum.

The ratio of branching fractions analysis is performed blindly, namely without knowing the central value of the measurement and without displaying the $\pm 3\sigma$ signal region. A simultaneous fit to the $B^0 \to \rho^0\gamma$ and $B^0 \to K^{*0}\gamma$ mass spectra has been developed in order to directly extract the ratio of branching fractions. The fit is performed simultaneously on 2011 and 2012 data samples.

The results obtained following the private unblinding of the mass fit on real data are presented and the preliminary study of systematic uncertainties is also discussed. The preliminary measurement is compatible with the value of the ratio of branching fractions obtained combining BaBar and Belle measurements and significantly more precise. It corresponds to about 400 $B^0 \to \rho^0\gamma$ signal events. Combining the statistical and systematic uncertainties we obtain a relative precision of about 10%.

Appendix A

PID optimization extra plots

In this appendix we show the 2D PID efficiency maps for signal and cross-feeds used to set the optimal PID cut to select pions. Only the MC12TuneV3 maps are shown. In addition also the figure of merits obtained using the different MC12TuneVX tunings (X=2, 3 and 4) available for Run 1 are displayed. As can be noticed the V3 tuning provides the best performances.

2D PID efficiency maps

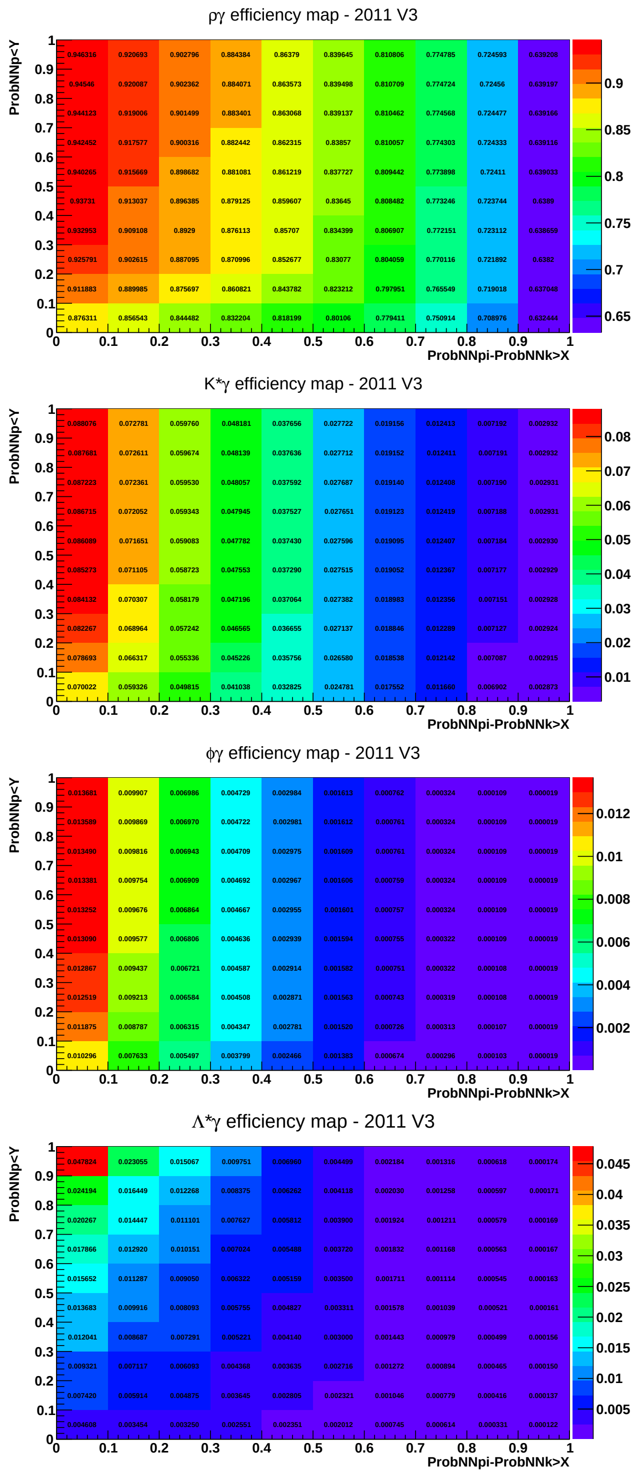


Figure A.1: From top to bottom $\pi\pi$ ($\rho\gamma$), $K\pi$ ($K^{*}\gamma$), KK ($\phi\gamma$) and pK ($\Lambda^{*}\gamma$) 2011 efficiency maps before performing the zoom. The numbers stand for the average value of the selection efficiency for each PID cut.

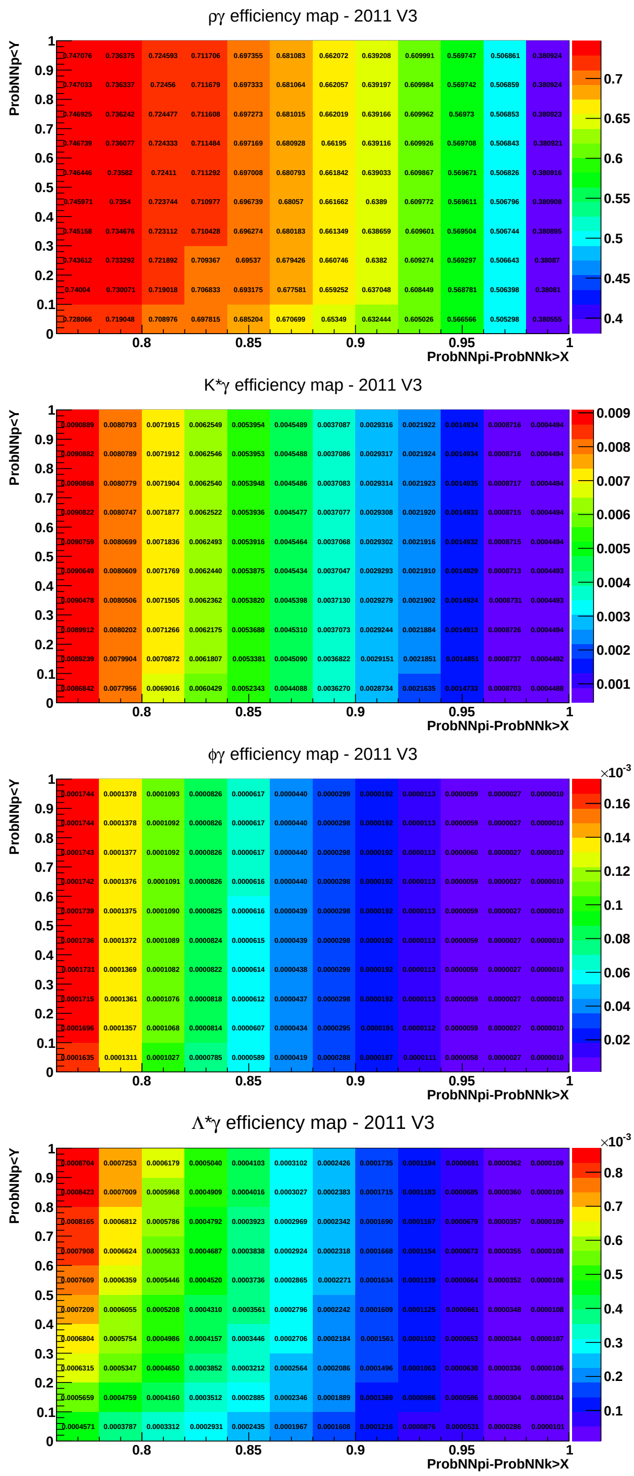


Figure A.2: From top to bottom $\pi\pi$ ($\rho\gamma$), $K\pi$ ($K^*\gamma$), KK ($\phi\gamma$) and pK ($\Lambda^*\gamma$) 2011 efficiency maps after performing the zoom. The numbers stand for the average value of the selection efficiency for each PID cut.

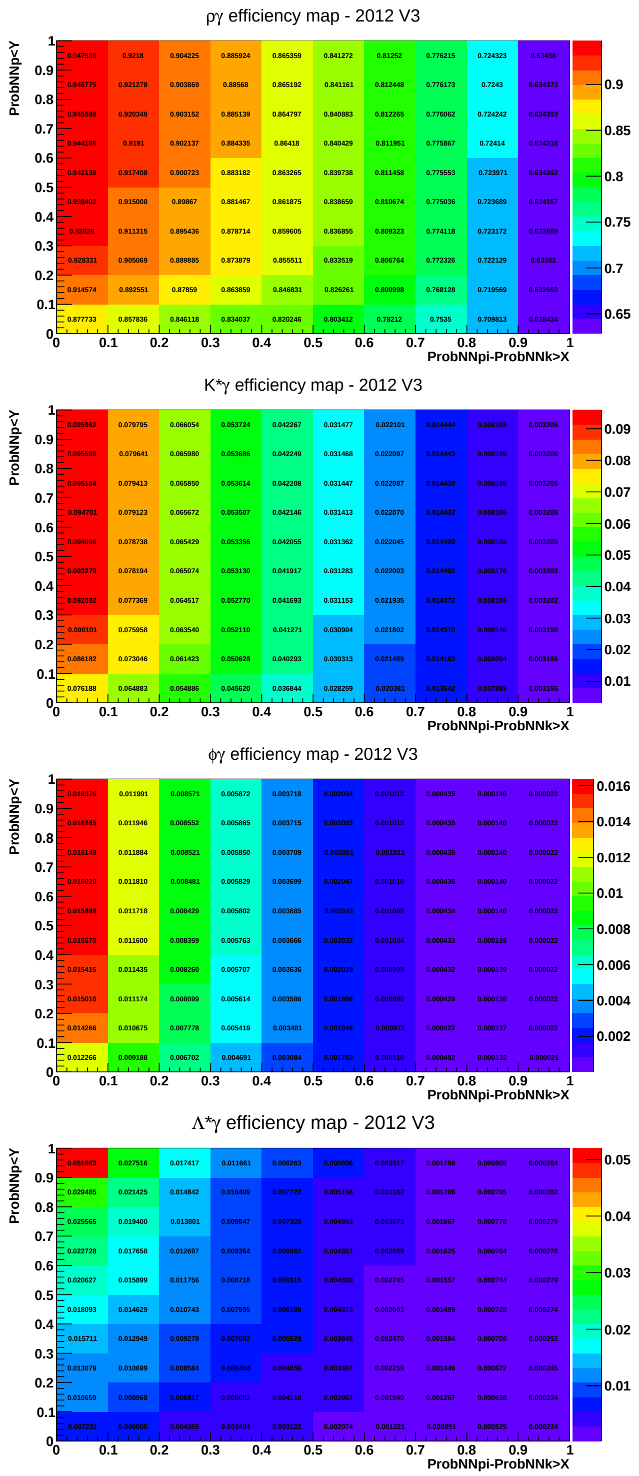


Figure A.3: From top to bottom $\pi\pi$ ($\rho\gamma$), $K\pi$ ($K^*\gamma$), KK ($\phi\gamma$) and pK ($\Lambda^*\gamma$) 2012 efficiency maps before performing the zoom. The numbers stand for the average value of the selection efficiency for each PID cut.

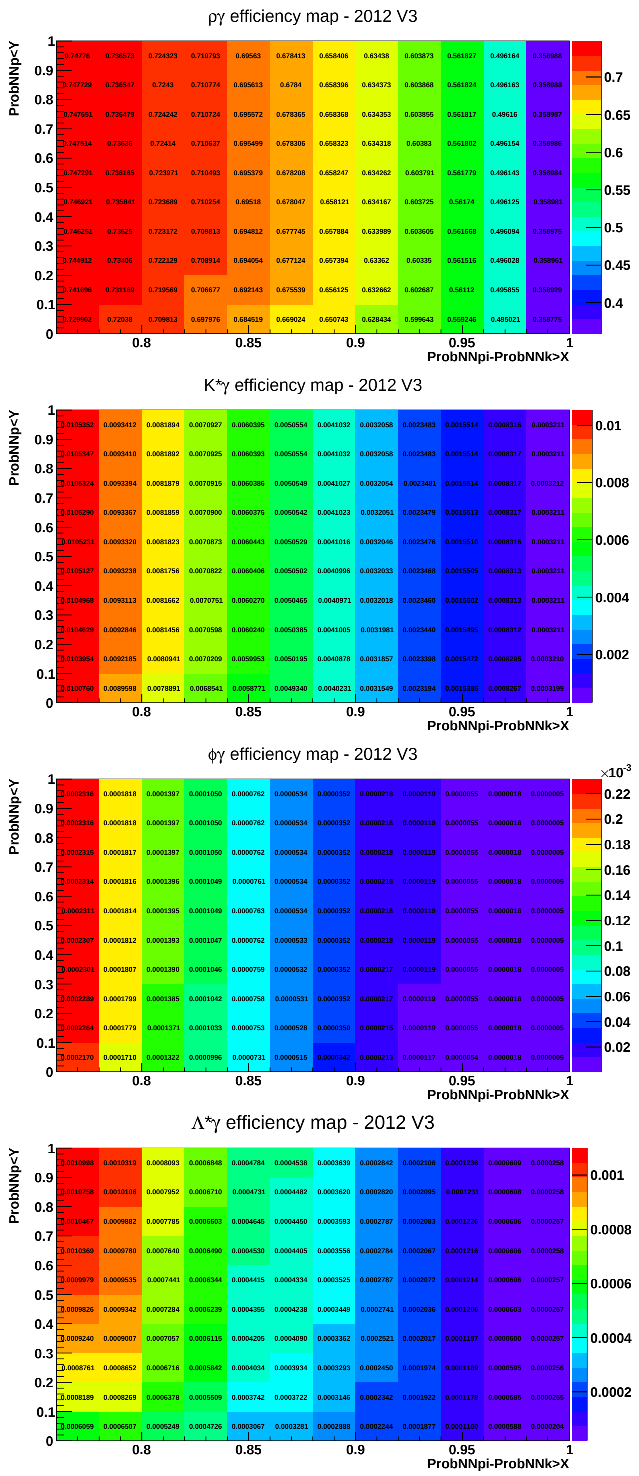


Figure A.4: From top to bottom $\pi\pi$ ($\rho\gamma$), $K\pi$ ($K^{*}\gamma$), KK ($\phi\gamma$) and pK ($\Lambda^{*}\gamma$) 2012 efficiency maps after performing the zoom. The numbers stand for the average value of the selection efficiency for each PID cut.

Figure of merits for the various ProbNN tunings

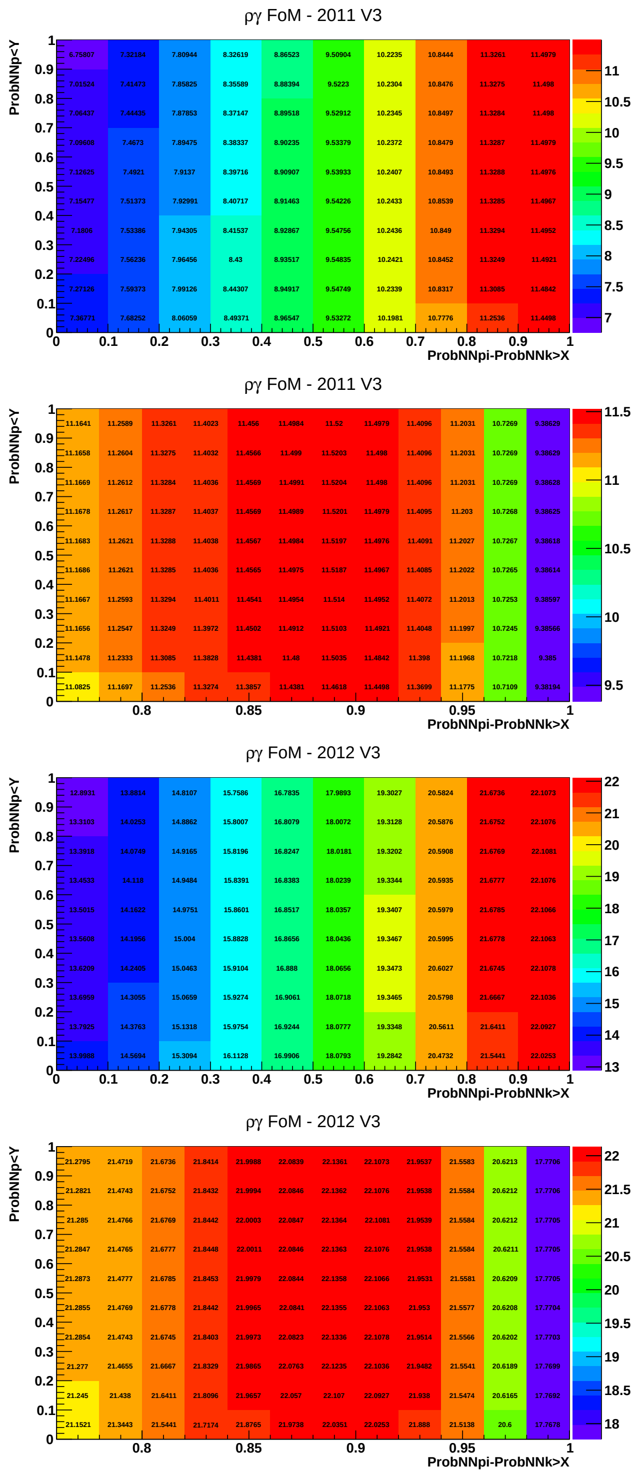


Figure A.5: From top to bottom MC12TuneV3 $B^0 \to \rho^0\gamma$ figure of merit maps for 2011 (first two) and 2012 (last two) Monte Carlo before (first and third) and after (second and fourth) performing the zoom.

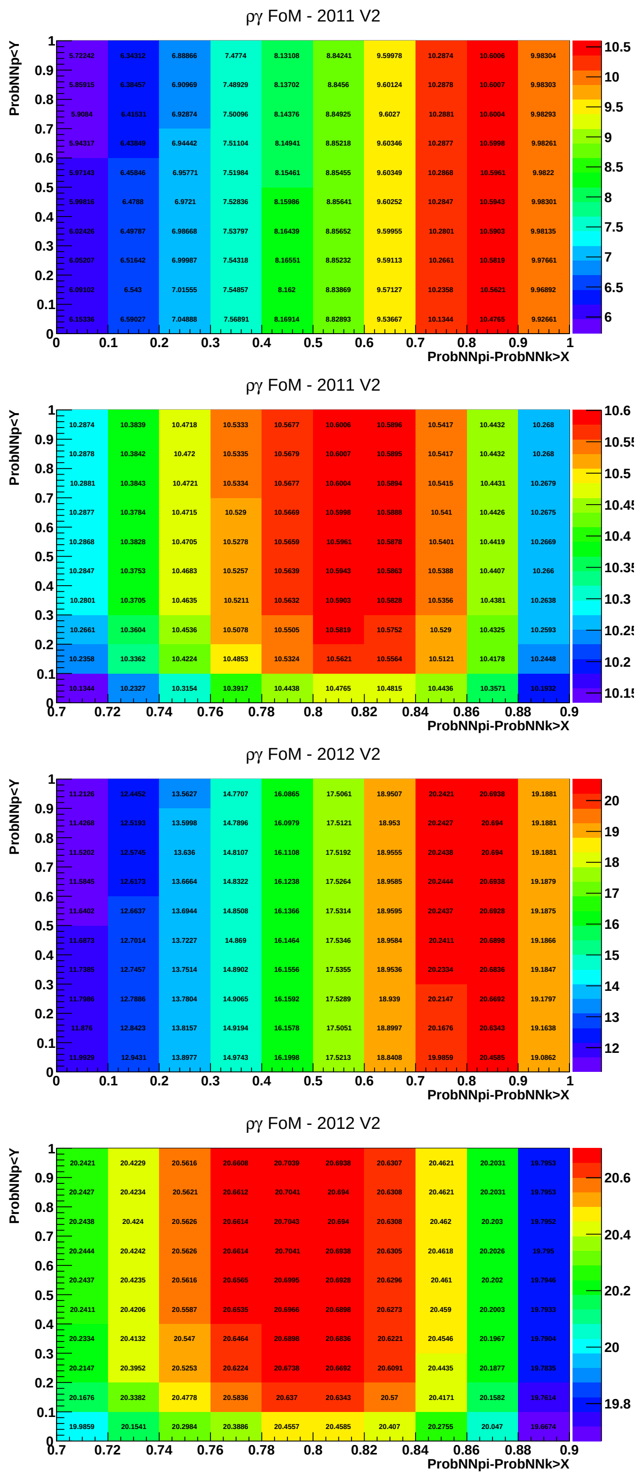


Figure A.6: From top to bottom MC12TuneV2 $B^0 \to \rho^0\gamma$ figure of merit maps for 2011 (first two) and 2012 (last two) Monte Carlo before (first and third) and after (second and fourth) performing the zoom.

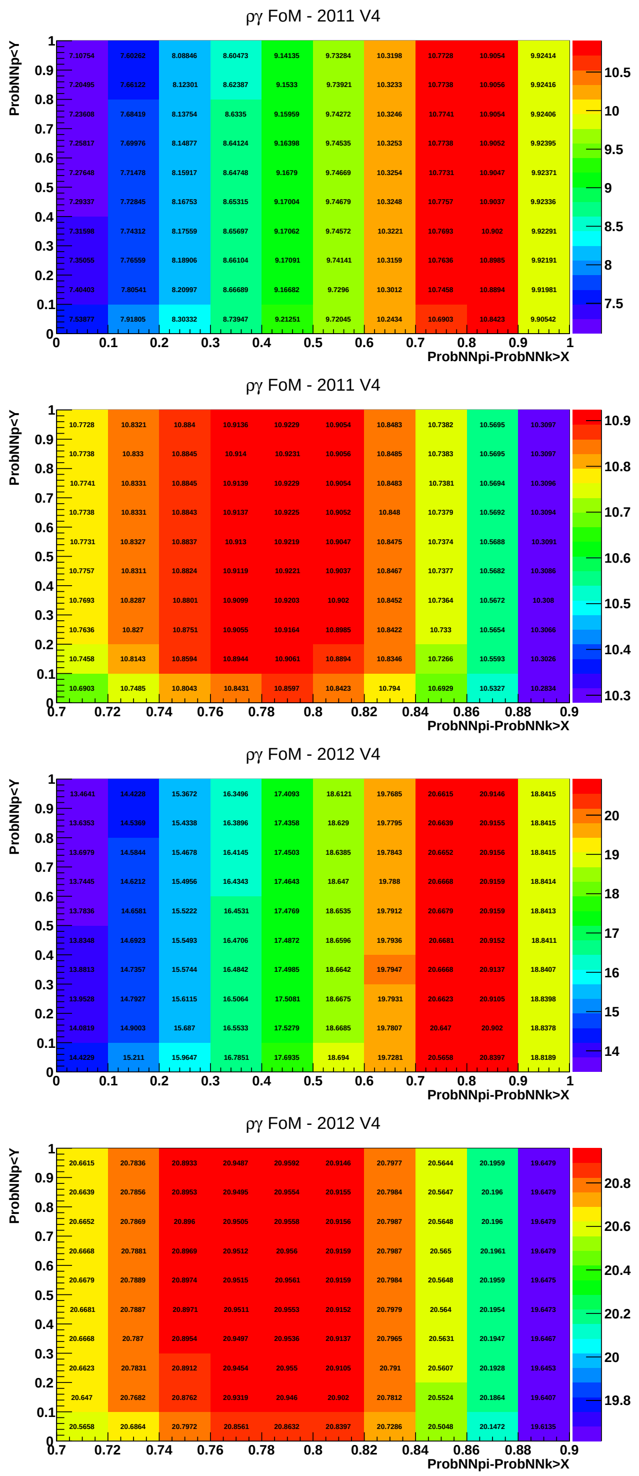


Figure A.7: From top to bottom MC12TuneV4 $B^0 \to \rho^0\gamma$ figure of merit maps for 2011 (first two) and 2012 (last two) Monte Carlo before (first and third) and after (second and fourth) performing the zoom.

Appendix B

MVA optimization extra plots

In this appendix we show the additional plots produced in the optimization of the multivariate classifier. The plots contained in the appendix are in order: data/MC comparison plots for the final list of variables used in the training; the distributions of the variables and their correlations for signal and background events after performing the final MLP training; the MLP overtraining response plots and the ROC curves for the main algorithms that we tested; the fits to the RHSB used to extract the slopes of the combinatorial background; the additional figure of merits calculated to verify the MLP stability.

Data/MC matching

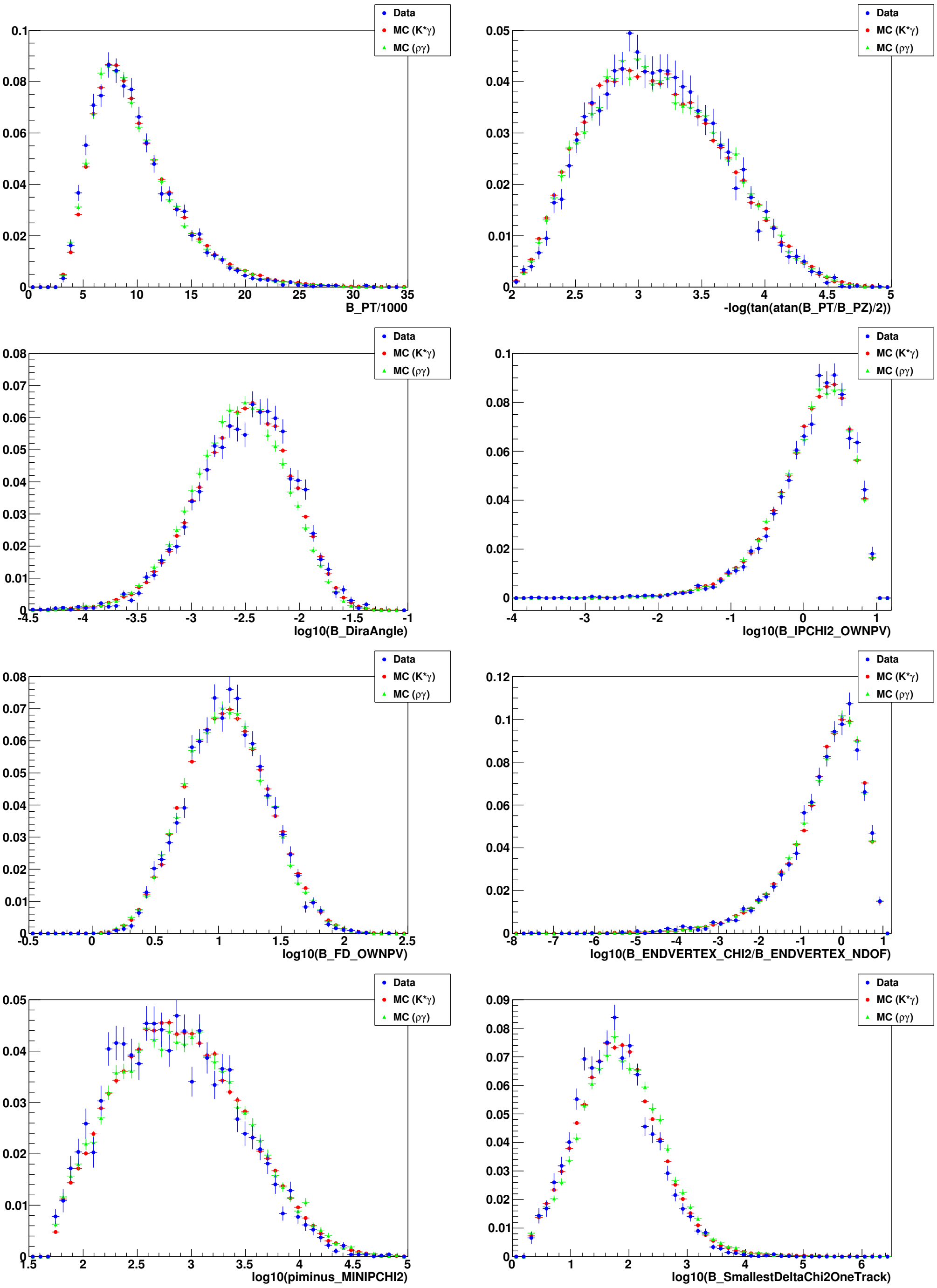


Figure B.1: 2011 data/MC agreement plots for the variables used in the final MVA training. $K^*\gamma$ sWeighted data (blue dots) are compared to simulated $K^*\gamma$ (red dots) and $\rho\gamma$ (green dots) MC events. The χ^2_{IP} of the positive charged track is not displayed to avoid comparing different objects (K and π). The χ^2_{IP} (K^+) data/MC agreement was separately checked on the $K^*\gamma$ channel and resulted to be fine.

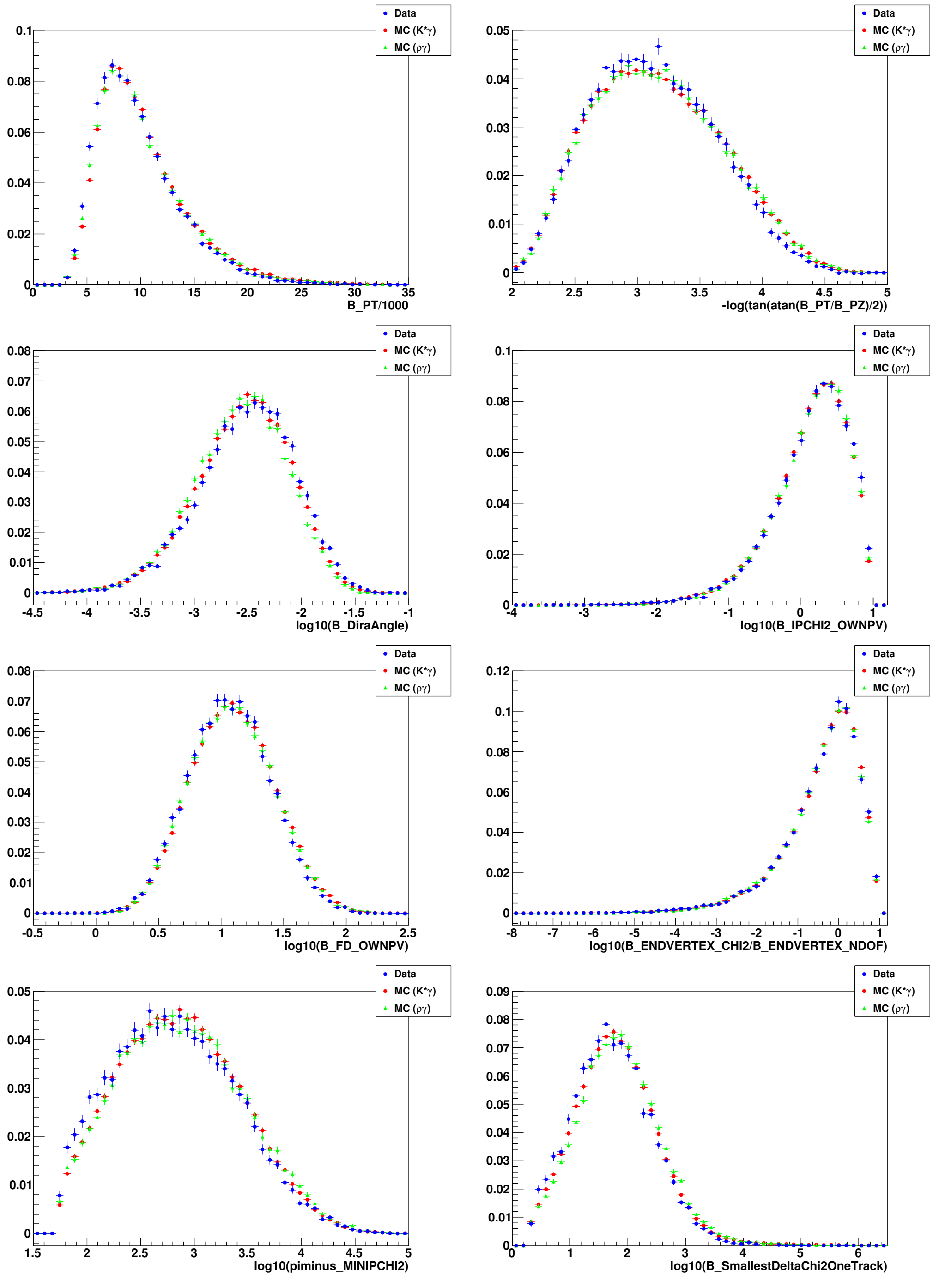


Figure B.2: 2012 data/MC agreement plots for the variables used in the final MVA training. $K^*\gamma$ sWeighted data (blue dots) are compared to simulated $K^*\gamma$ (red dots) and $\rho\gamma$ (green dots) MC events. The $\chi^2_{\rm IP}$ of the positive charged track is not displayed to avoid comparing different objects (K and π). The $\chi^2_{\rm IP}$ (K^+) data/MC agreement was separately checked on the $K^*\gamma$ channel and resulted to be fine.

Distribution of variables used in the MLP training

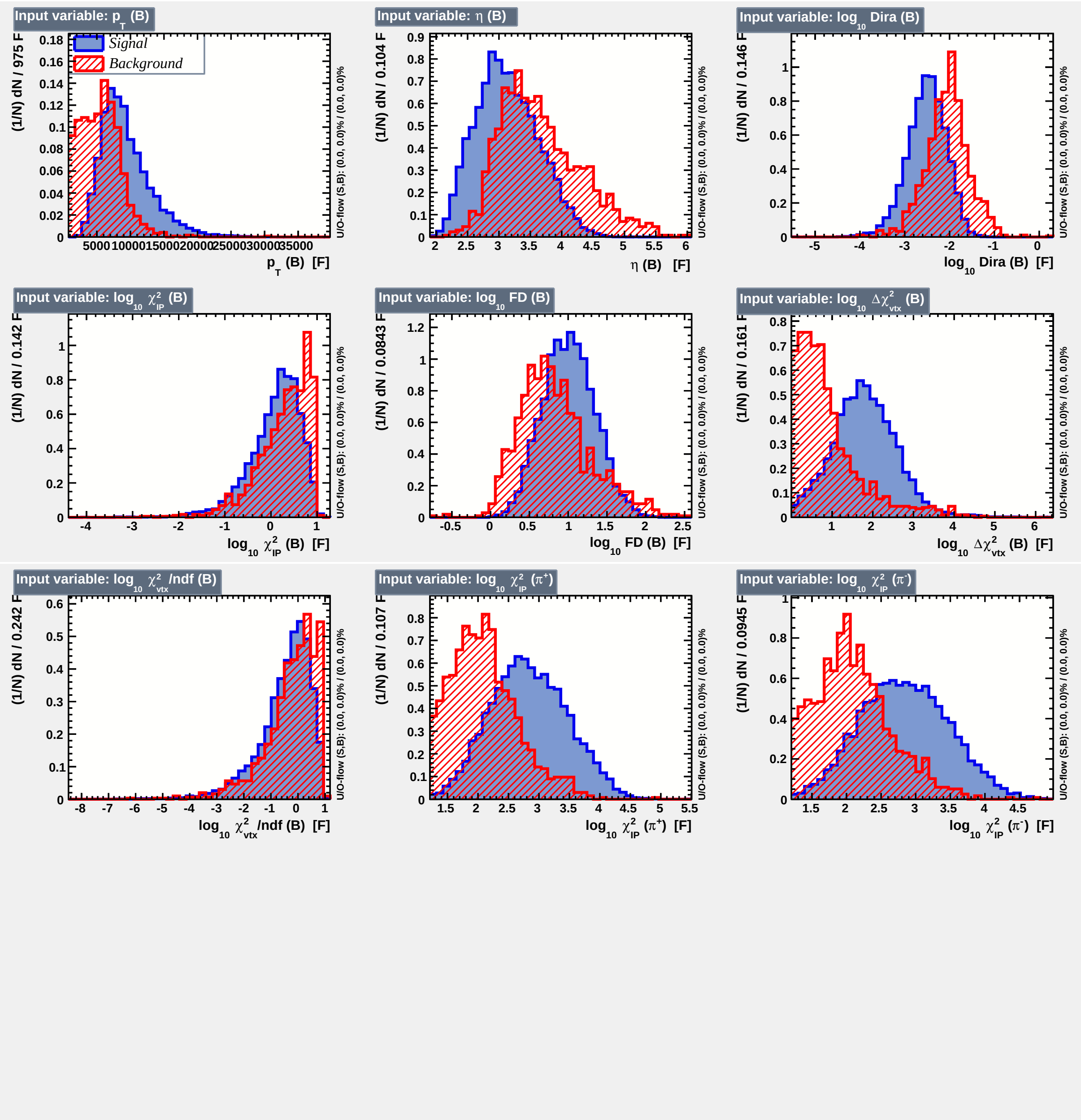


Figure B.3: Distribution of variables used in the training of MVA_1 for 2011, superimposing RHSB background events (in red) and MC-generated signal events (in blue).

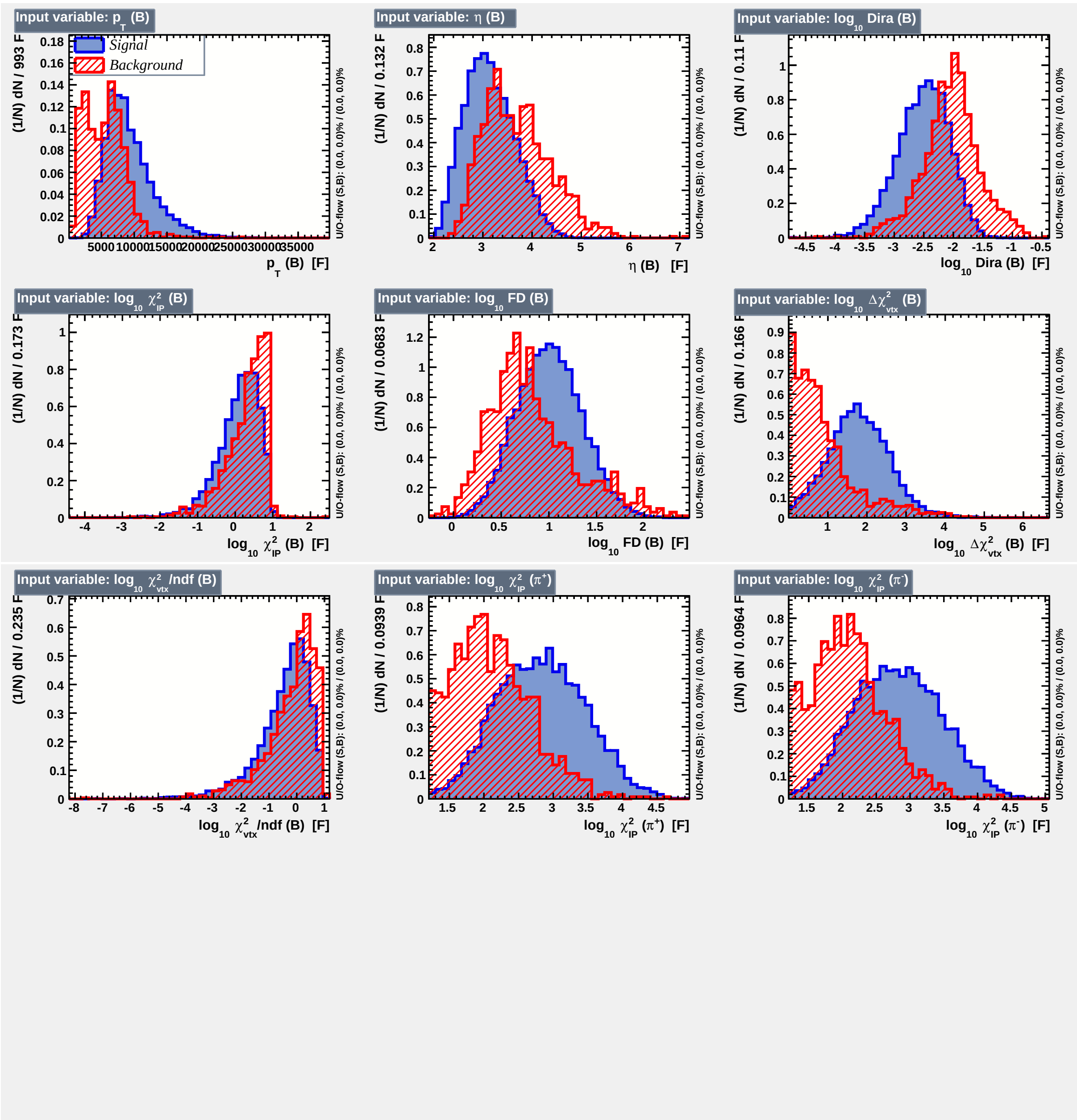


Figure B.4: Distribution of variables used in the training of MVA_2 for 2011, superimposing RHSB background events (in red) and MC-generated signal events (in blue).

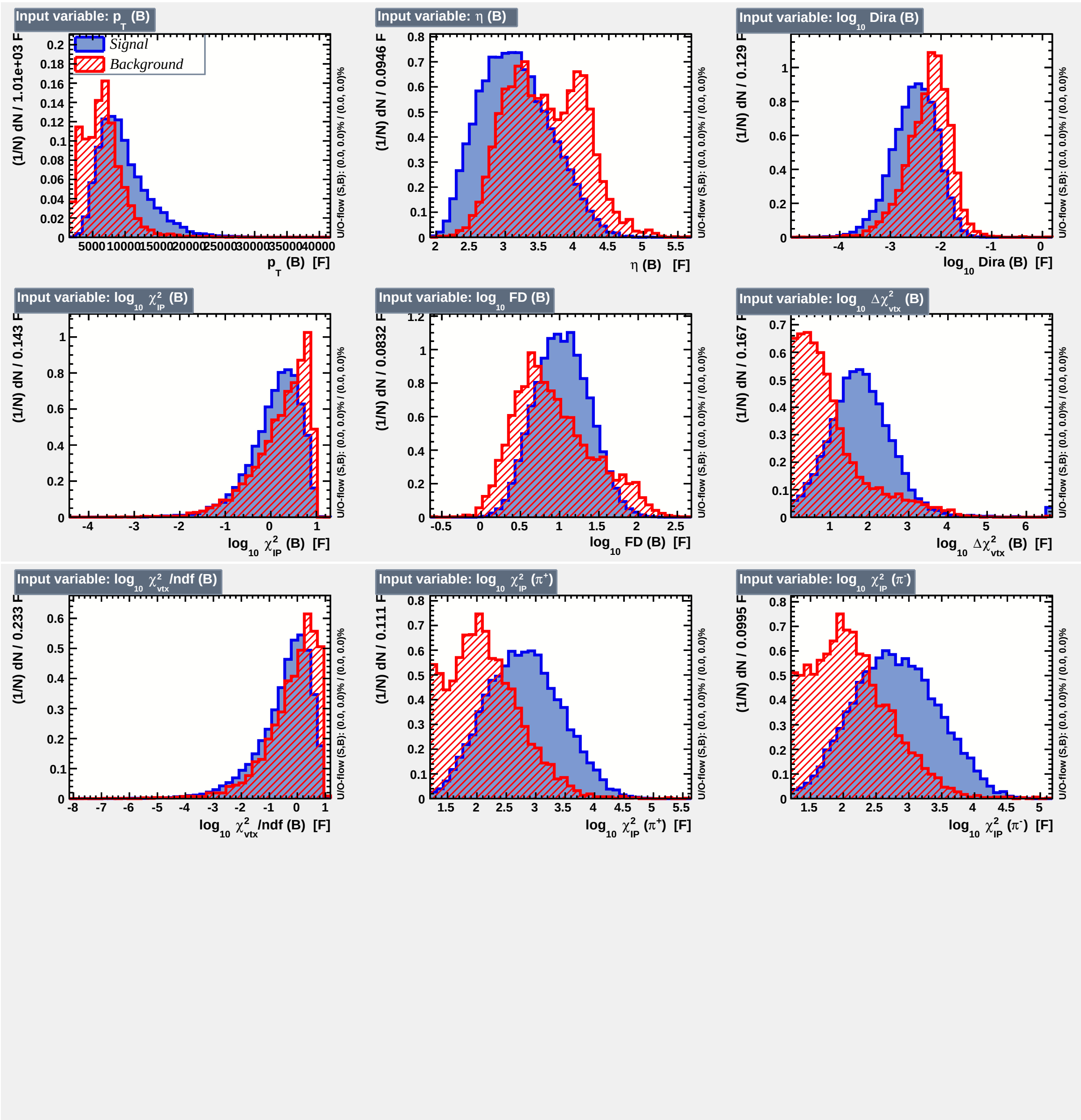


Figure B.5: Distribution of variables used in the training of MVA$_2$ for 2012, superimposing RHSB background events (in red) and MC-generated signal events (in blue).

Correlations of variables used in the MLP training

Figure B.6: Linear correlation of variables used in the training of MVA_1 for 2011 (top), MVA_2 for 2011 (middle) and MVA_2 for 2012 (bottom). The plots on the left column are for the signals and on the right column are for the background events.

Overtraining plots and ROC curves

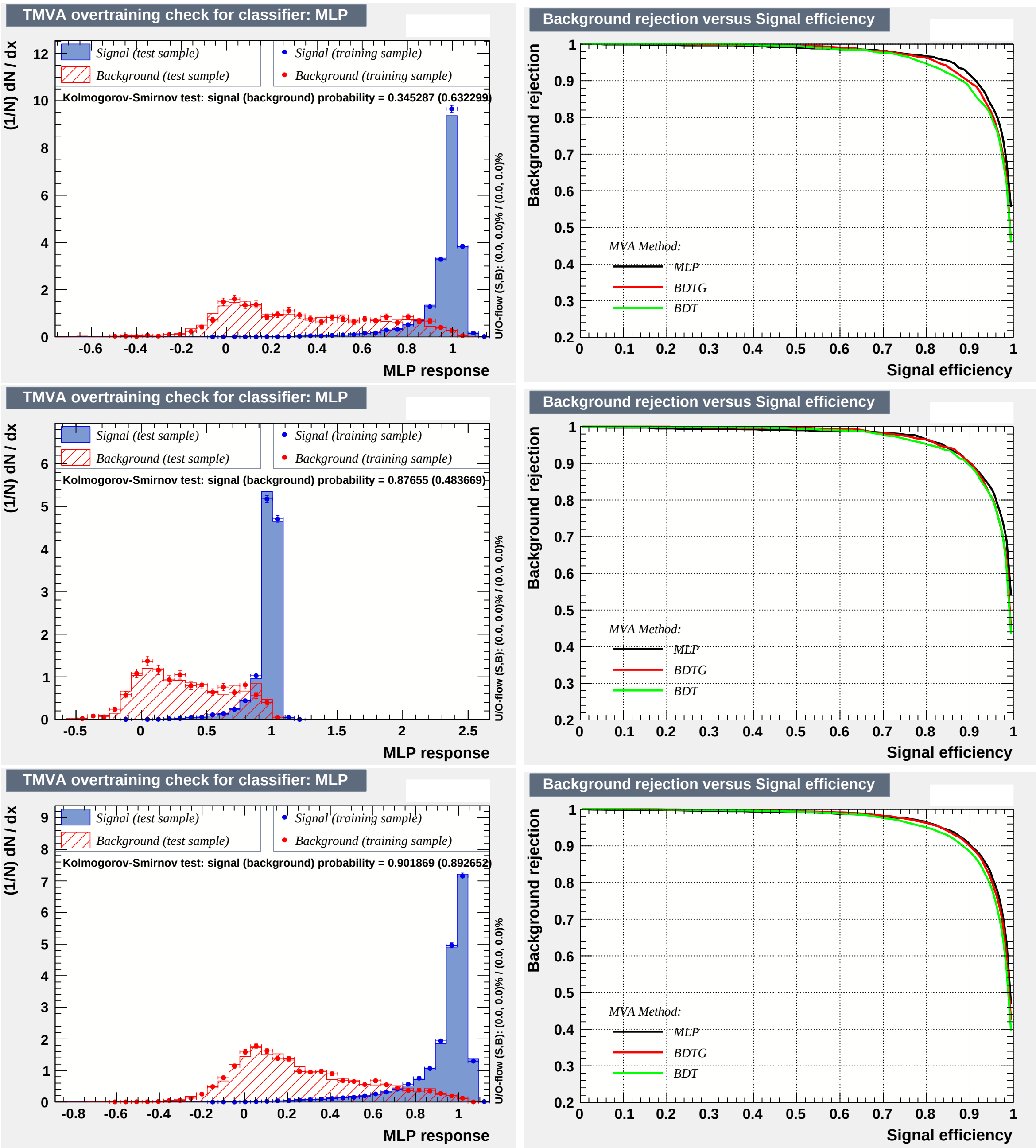


Figure B.7: $MLP_{1/2}$ response with overtraining test (left) and $MVA_{1/2}$ ROC-curves (right) for the best performing algorithms for 2011 (top/middle). Plots related to MVA_2 for 2012 are displayed on the bottom.

Fits to the RHSB without MLP cuts

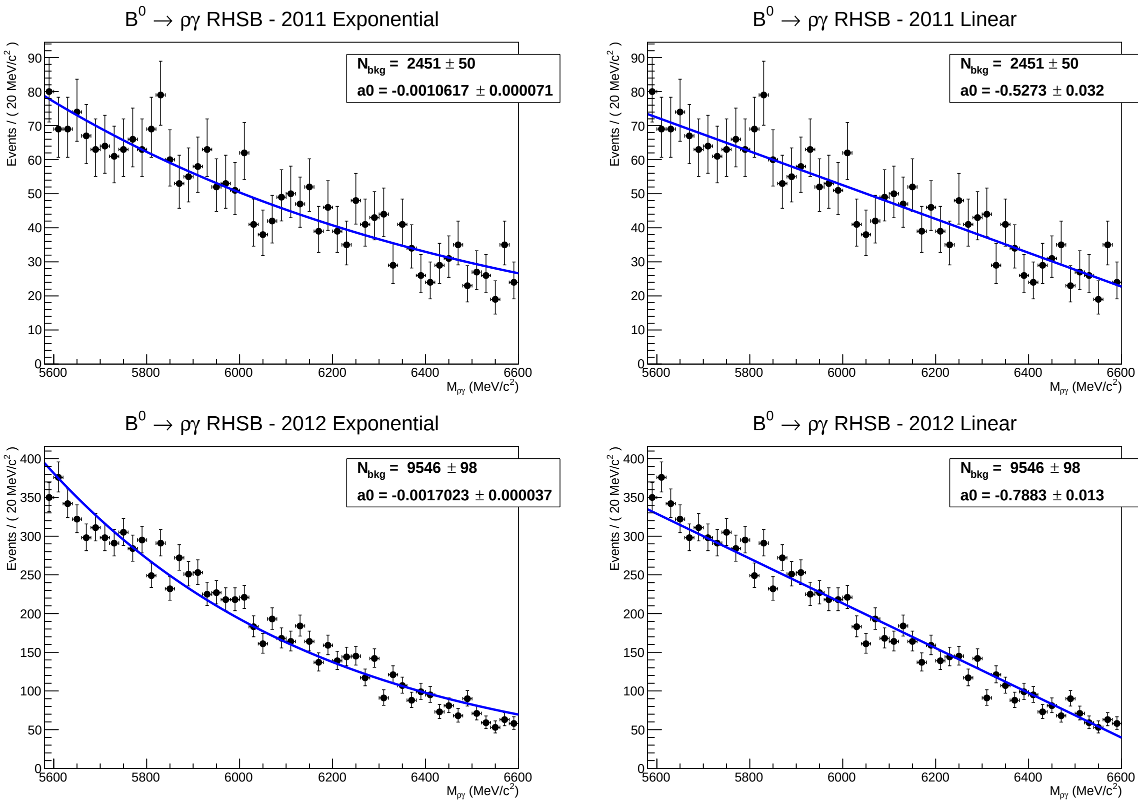


Figure B.8: Exponential (left) and linear (right) fits to the $B^0 \to \rho^0\gamma$ RHSB for 2011 (top) and 2012 (bottom). The fitted parameters were used to extract the scaling factors needed to estimate the expected combinatorial background yields in the $\pm 3\sigma$ region.

Additional figure of merits for MLP cut stability checks

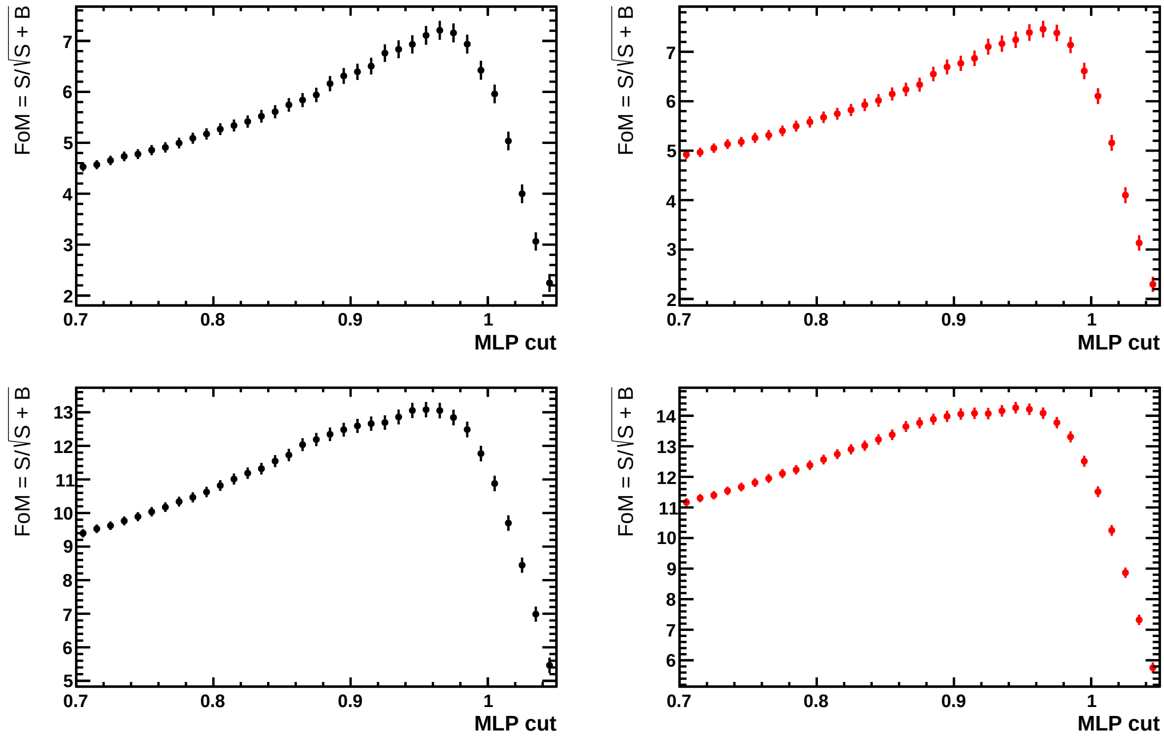


Figure B.9: $B^0 \to \rho^0\gamma$ figure of merits obtained considering combinatorial background events and $hh\gamma$ cross-feeds for exponential (left) and linear (right) extrapolation hypotheses in 2011 (top) and 2012 (bottom).

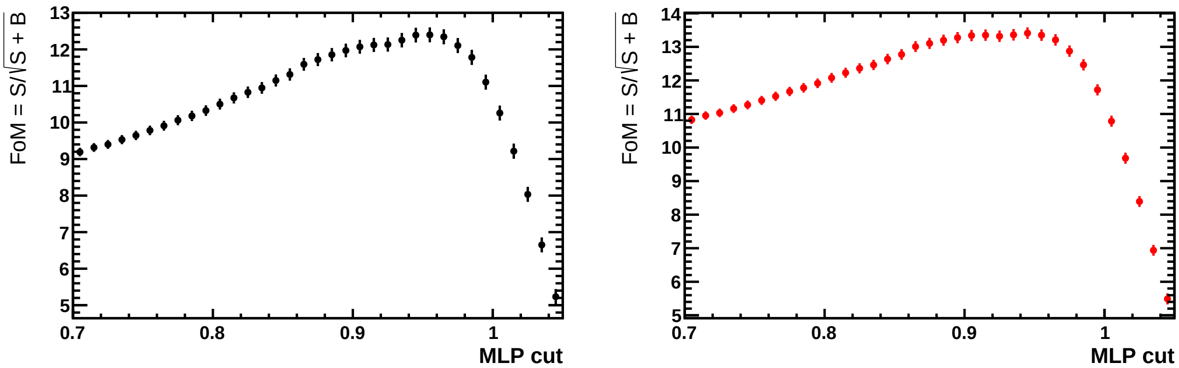


Figure B.10: 2012 $B^0 \to \rho^0\gamma$ figure of merits obtained considering combinatorial background events, $hh\gamma$ and $hh\pi^0$ cross-feeds for exponential (left) and linear (right) extrapolation hypotheses.

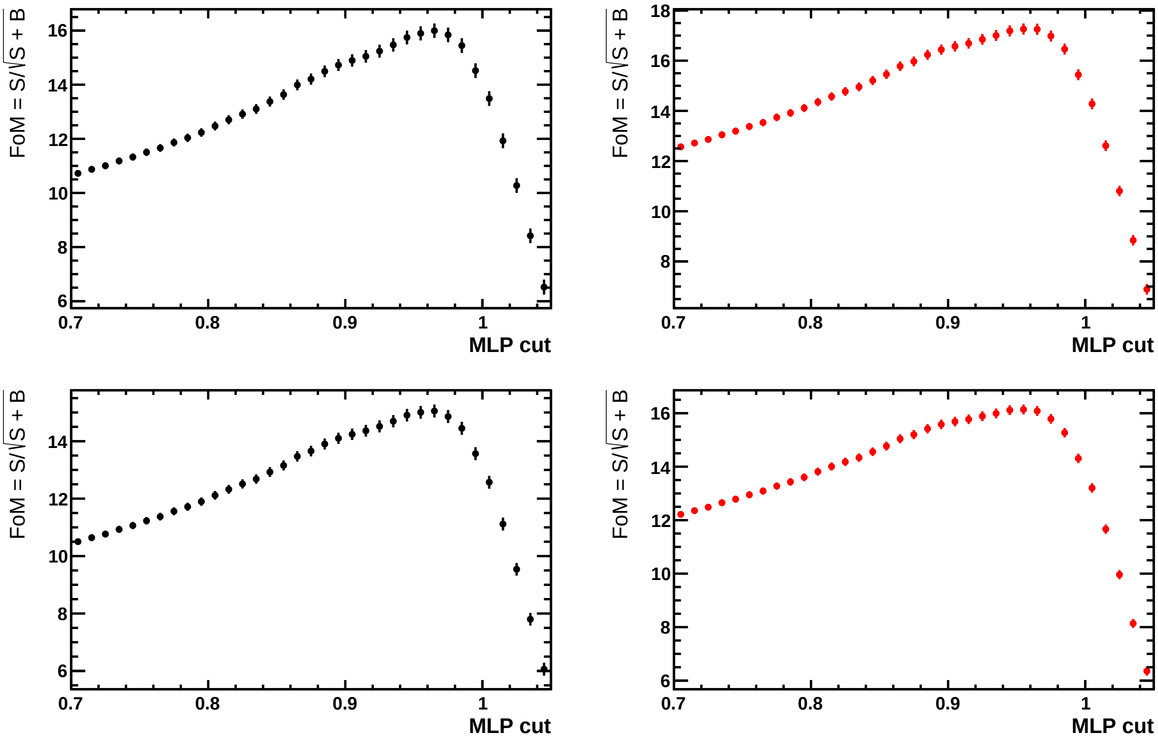


Figure B.11: $B^0 \to \rho^0\gamma$ total figure of merits (2011 + 2012) obtained considering combinatorial background events only (top) and adding also $hh\gamma$ cross-feeds (bottom) for exponential (left) and linear (right) extrapolation hypotheses.

Appendix C

Post-unblinding extra plots

In this appendix we show the $B^0 \to K^{*0}\gamma$ simultaneous mass fit projections associated to the final fit used to perform the measurement. Events in the $B^0 \to \rho^0\gamma$ mass spectrum are selected in the restricted ρ^0 mass window [625; 925] MeV/c^2.

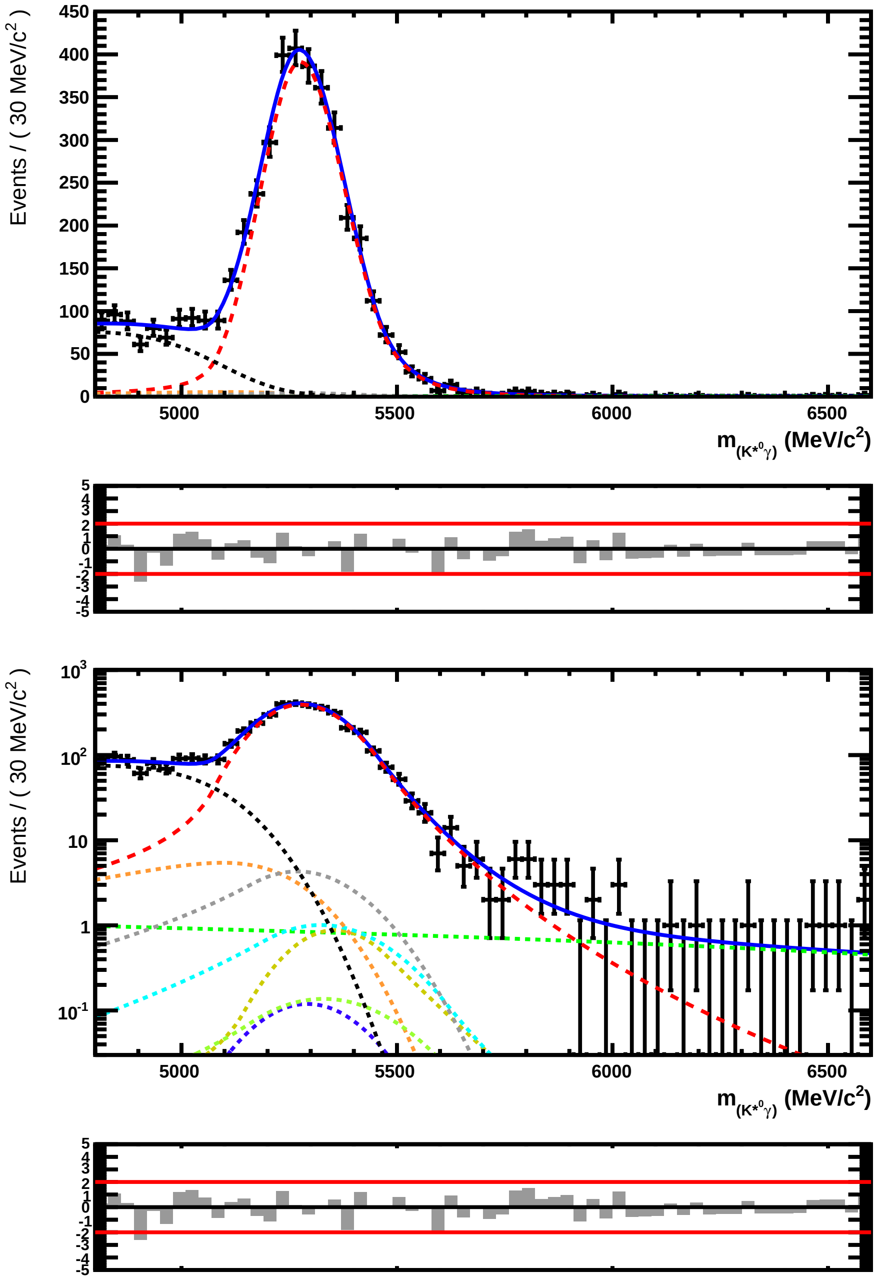


Figure C.1: $B^0 \to K^{*0}\gamma$ 2011+2012 simultaneous mass fit projections for 2011 data in linear (top) and semi-logarithmic scale (bottom). The black points represent the data and the solid blue line is the total fit function. The signal is fitted with an asymmetric double-tail Crystal Ball PDF and the combinatorial background is described with a first order polynomial (dashed green line). The peaking background modes are $B^0 \to \rho^0\gamma$ (dashed dark yellow line), $B_s^0 \to \phi\gamma$ (dashed dark blue line), $\Lambda_b^0 \to \Lambda^*\gamma$ (dashed cyan line), $B^0 \to K^+\pi^-\pi^0$ (dashed gray line) and $B^0 \to \pi^+\pi^-\pi^0$ (dashed light green line). The partially reconstructed modes are $B^0 \to K^{*0}\eta(\gamma\gamma)$ (dashed orange line) and $B^0 \to K^+\pi^-\gamma X$ (dashed black line).

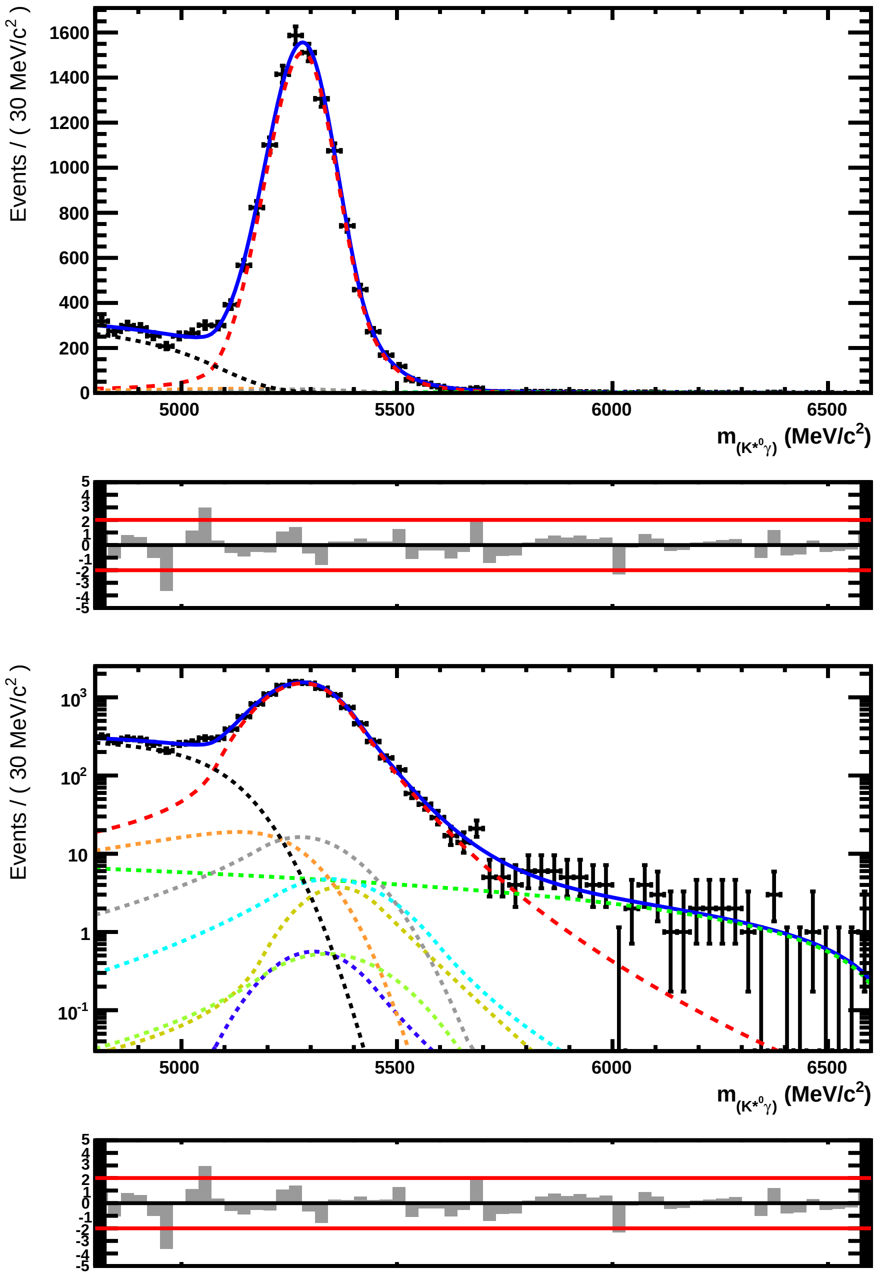


Figure C.2: $B^0 \to K^{*0}\gamma$ 2011+2012 simultaneous mass fit projections for 2012 data in linear (top) and semi-logarithmic scale (bottom). The black points represent the data and the solid blue line is the total fit function. The signal is fitted with an asymmetric double-tail Crystal Ball PDF and the combinatorial background is described with a first order polynomial (dashed green line). The peaking background modes are $B^0 \to \rho^0\gamma$ (dashed dark yellow line), $B_s^0 \to \phi\gamma$ (dashed dark blue line), $\Lambda_b^0 \to \Lambda^*\gamma$ (dashed cyan line), $B^0 \to K^+\pi^-\pi^0$ (dashed gray line) and $B^0 \to \pi^+\pi^-\pi^0$ (dashed light green line). The partially reconstructed modes are $B^0 \to K^{*0}\eta(\gamma\gamma)$ (dashed orange line) and $B^0 \to K^+\pi^-\gamma X$ (dashed black line).

Bibliography

[1] S.Braibant, G.Giacomelli, M.Spurio, *Particles and Fundamental Interactions*, Springer (2012).

[2] Y. Nir, *CP violation in meson decays*, in *High-energy physics. Proceedings, 3rd Latin American CERN-CLAF School, Malargue, Argentina, February 27-March 12, 2005*, pp. 79–145, 2006. `arXiv:hep-ph/0510413`.

[3] ATLAS, G. Aad *et al.*, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716** (2012) 1, `arXiv:1207.7214`.

[4] CMS, S. Chatrchyan *et al.*, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. **B716** (2012) 30, `arXiv:1207.7235`.

[5] ATLAS, G. Aad *et al.*, *Evidence for the spin-0 nature of the Higgs boson using ATLAS data*, Phys. Lett. **B726** (2013) 120, `arXiv:1307.1432`.

[6] CMS, S. Chatrchyan *et al.*, *Study of the Mass and Spin-Parity of the Higgs Boson Candidate Via Its Decays to Z Boson Pairs*, Phys. Rev. Lett. **110** (2013), no. 8 081803, `arXiv:1212.6639`.

[7] Particle Data Group, C. Patrignani *et al.*, *Review of Particle Physics*, Chin. Phys. **C40** (2016), no. 10 100001.

[8] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10** (1963) 531, [,648(1963)].

[9] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652.

[10] N. Beisert, *Quantum Field Theory I*, https://people.phys.ethz.ch/.

[11] R. Soldati, *Introduction to Quantum Field Theory (A Primer for a Basic Education)*, http://www.robertosoldati.com.

[12] T. D. Lee and C. N. Yang, *Question of parity conservation in weak interactions*, Phys. Rev. **104** (1956) 254.

[13] C. S. Wu *et al.*, *Experimental test of parity conservation in beta decay*, Phys. Rev. **105** (1957) 1413.

[14] R. L. Garwin, L. M. Lederman, and M. Weinrich, *Observations of the failure of conservation of parity and charge conjugation in meson decays: the magnetic moment of the free muon*, Phys. Rev. **105** (1957) 1415.

[15] J. I. Friedman and V. L. Telegdi, *Nuclear emulsion evidence for parity nonconservation in the decay chain* $\pi^+ \to \mu^+ \to e^+$, Phys. Rev. **106** (1957) 1290.

[16] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Evidence for the* 2π *decay of the* k_2^0 *meson*, Phys. Rev. Lett. **13** (1964) 138.

[17] C. Jarlskog, *Commutator of the quark mass matrices in the standard electroweak model and a measure of maximal* CP *nonconservation*, Phys. Rev. Lett. **55** (1985) 1039.

[18] The CKMfitter Group, http://ckmfitter.in2p3.fr.

[19] L. Wolfenstein, *Parametrization of the kobayashi-maskawa matrix*, Phys. Rev. Lett. **51** (1983) 1945.

[20] A. J. Buras, M. E. Lautenbacher, and G. Ostermaier, *Waiting for the top quark mass,* $K^+ \to \pi^+ \nu\overline{\nu}$*,* $B_s^0 - \overline{B}_s^0$ *mixing and CP asymmetries in B decays*, Phys. Rev. **D50** (1994) 3433, `arXiv:hep-ph/9403384`.

[21] V. Weisskopf and E. P. Wigner, *Calculation of the natural brightness of spectral lines on the basis of Dirac's theory (In German)*, Z. Phys., **63**, 54 (1930).

[22] V. Weisskopf and E. P. Wigner, *Over the natural line width in the radiation of the harmonius oscillator (In German)*, Z. Phys., **65**, 18 (1930).

[23] I. I. Bigi and A. I. Sanda, *CP violation*, Cambridge University Press, (2000).

[24] S. Aoki *et al.*, *Review of lattice results concerning low-energy particle physics*, Eur. Phys. J. **C74** (2014) 2890, `arXiv:1310.8555`.

[25] A. Ali, E. Lunghi, and A. Ya. Parkhomenko, *Implication of the* $B \to (\rho, \omega)\gamma$ *branching ratios for the CKM phenomenology*, Phys. Lett. **B595** (2004) 323, `arXiv:hep-ph/0405075`.

[26] BABAR Collaboration, B. Aubert *et al.*, *Measurements of Branching Fractions for* $B^+ \to \rho^+\gamma$*,* $B^0 \to \rho^0\gamma$*, and* $B^0 \to \omega\gamma$, Phys. Rev. **D78** (2008) 112001, `arXiv:0808.1379`.

[27] P. Ball, G. W. Jones, and R. Zwicky, $B \to V\gamma$ *beyond QCD factorisation*, Phys. Rev. **D75** (2007) 054004, `arXiv:hep-ph/0612081`.

[28] L. Evans and P. Bryant, *LHC Machine*, vol. 3, 2008. doi: 10.1088/1748-0221/3/08/S08001.

[29] LHCb, R. Aaij *et al.*, *Measurement of $\sigma(pp \to b\bar{b}X)$ at $\sqrt{s} = 7$ TeV in the forward region*, Phys. Lett. **B694** (2010) 209, `arXiv:1009.2731`.

[30] LHCb, A. A. Alves, Jr. *et al.*, *The LHCb Detector at the LHC*, JINST **3** (2008) S08005.

[31] LHCb Collaboration, *LHCb VELO (VErtex LOcator): Technical Design Report*, CERN-LHCC-2001-011, CERN, Geneva, 2001.

[32] R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, JINST **9** (2014) 09007, `arXiv:1405.7808`.

[33] LHCb, *LHCb technical design report: Reoptimized detector design and performance*, CERN-LHCC-2003-030, 2003.

[34] LHCb, *LHCb: Inner tracker technical design report*, CERN-LHCC-2002-029, 2002.

[35] LHCb, *LHCb: Outer tracker technical design report*, CERN-LHCC-2001-024, 2001.

[36] LHCb Outer Tracker Group, R. Arink *et al.*, *Performance of the LHCb Outer Tracker*, JINST **9** (2014), no. 01 P01002, `arXiv:1311.3893`.

[37] LHCb, *LHCb magnet: Technical design report*, CERN-LHCC-2000-007, 2000.

[38] LHCb RICH Group, M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. **C73** (2013) 2431, `arXiv:1211.6759`.

[39] LHCb, *LHCb calorimeters: Technical design report*, CERN-LHCC-2000-036, 2000.

[40] William R. Leo, *Techniques for Nuclear and Particle Physics Experiments - 2nd Revised Edition*, Springer (1994).

[41] LHCb, P. Perret, *First Years of Running for the LHCb Calorimeter System*, PoS **TIPP2014** (2014) 030, `arXiv:1407.4289`.

[42] LHCb, R. Aaij *et al.*, *Angular analysis of the $B^0 \to K^{*0}\mu^+\mu^-$ decay using 3 fb^{-1} of integrated luminosity*, JHEP **02** (2016) 104, `arXiv:1512.0444`.

[43] LHCb, R. Aaij *et al.*, *Measurement of the CP-violating phase ϕ_s in the decay $B_s^0 \to J/\psi\phi$*, Phys. Rev. Lett. **108** (2012) 101803, `arXiv:1112.3183`.

[44] LHCb, R. Aaij *et al.*, *First Evidence for the Decay $B_s^0 \to \mu^+\mu^-$*, Phys. Rev. Lett. **110** (2013), no. 2 021801, `arXiv:1211.2674`.

[45] LHCb, R. Aaij *et al.*, *Measurement of the $B_s^0 \to \mu^+\mu^-$ branching fraction and search for $B^0 \to \mu^+\mu^-$ decays at the LHCb experiment*, Phys. Rev. Lett. **111** (2013) 101805, `arXiv:1307.5024`.

[46] LHCb, *LHCb: Addendum to the muon system technical design report*, CERN-LHCC-2003-002, 2003.

[47] LHCb, R. Aaij *et al.*, *Measurement of the inelastic pp cross-section at a centre-of-mass energy of $\sqrt{s} = 7$ TeV*, JHEP **02** (2015) 129, `arXiv:1412.2500`.

[48] M. Williams *et al.*, *The HLT2 Topological Lines*, LHCb-PUB-2011-002, CERN-LHCb-PUB-2011-002.

[49] H.-J. Yang, B. P. Roe, and J. Zhu, *Studies of boosted decision trees for MiniBooNE particle identification*, Nucl. Instrum. Meth. **A555** (2005) 370, `arXiv:physics/0508045`.

[50] LHCb, R. Aaij *et al.*, *Absolute luminosity measurements with the LHCb detector at the LHC*, JINST **7** (2012) P01010, `arXiv:1110.2866`.

[51] LHCb, R. Aaij *et al.*, *Precision luminosity measurements at LHCb*, JINST **9** (2014), no. 12 P12005, `arXiv:1410.0149`.

[52] O. Callot, *Calorimeter data format and 2-D zero suppression scheme*, Tech. Rep. LHCb-99-037, CERN, Geneva, Oct, 1999.

[53] V. Breton, N. Brun, and P. Perret, *A clustering algorithm for the LHCb electromagnetic calorimeter using a cellular automaton*, Tech. Rep. LHCb-2001-123, CERN, Geneva, Sep, 2001.

[54] O. Deschamps *et al.*, *Photon and neutral pion reconstruction*, Tech. Rep. LHCb-2003-091, CERN, Geneva, Sep, 2003.

[55] L. Beaucourt and E. Tournefier, *Study of $B^0 \to K^{*0}\gamma$, $B_s^0 \to \phi\gamma$ and $B_s^0 \to K^{*0}\gamma$ decays using converted photons with the LHCb detector.*, Oct, 2016. Presented 08 Dec 2016.

[56] R. Hollebeek, *Topics in Calorimetry for High Energy Physics*, SLAC Summer School Lectures (1991).

[57] LHCb, R. Aaij *et al.*, *Measurement of the ratio of branching fractions $B(B^0 \to K^{*0}\gamma)/B(B_s^0 \to \phi\gamma)$ and the direct CP asymmetry in $B^0 \to K^{*0}\gamma$*, Nucl. Phys. **B867** (2013) 1, `arXiv:1209.0313`.

[58] LHCb, R. Aaij *et al.*, *LHCb Detector Performance*, Int. J. Mod. Phys. **A30** (2015), no. 07 1530022, `arXiv:1412.6352`.

[59] I. Belyaev *et al.*, *Kali: The framework for fine calibration of the lhcb electromagnetic calorimeter*, Journal of Physics: Conference Series **331** (2011), no. 3 032050.

[60] D. A. Roa Romero and P. Perret, *Study of charmless $B^0_{d,s} \to h^+h'^-\pi^0$ decays in LHCb*, May, 2013. Presented 24 May 2013.

[61] M. Hoballah, P. Perret, and O. Deschamps, *Measurement of the photon polarization using $B^0_s \to \phi\gamma$ at LHCb.*, Apr, 2015. Presented 03 Mar 2015.

[62] M. Calvo Gomez *et al.*, *A tool for γ/π^0 separation at high energies*, Tech. Rep. LHCb-PUB-2015-016. CERN-LHCb-PUB-2015-016, CERN, Geneva, Aug, 2015.

[63] K. G. Sobczak and S. Monteil, *Study of charmless three-body decays of neutral B mesons with the LHCb spectrometer*, 2011. Presented 15 Dec 2011.

[64] M. Calvo Gomez, *Calibration tools for the γ/π^0 separation tool*, `https://twiki.cern.ch/twiki/bin/view/LHCbPhysics/GammaPi0SeparationCalib`.

[65] M. Pivk and F. R. Le Diberder, *SPlot: A Statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, `arXiv:physics/0402083`.

[66] LHCb, R. Aaij *et al.*, *Measurement of the difference of time-integrated CP asymmetries in $D^0 \to K^-K^+$ and $D^0 \to \pi^-\pi^+$ decays*, Phys. Rev. Lett. **116** (2016), no. 19 191601, `arXiv:1602.0316`.

[67] Belle Collaboration, N. Taniguchi *et al.*, *Measurement of Branching Fractions, Isospin, and CP-Violating Asymmetries for Exclusive $b \to d\gamma$ Modes*, Phys. Rev. Lett. **101** (2008) 111801, `arXiv:0804.4770`.

[68] LHCb, R. Aaij *et al.*, *A precise measurement of the B^0 meson oscillation frequency*, Eur. Phys. J. **C76** (2016), no. 7 412, `arXiv:1604.0347`.

[69] LHCb, R. Aaij *et al.*, *Observation of B^0_s-$\bar{B}^0_s$ mixing and measurement of mixing frequencies using semileptonic B decays*, Eur. Phys. J. **C73** (2013), no. 12 2655, `arXiv:1308.1302`.

[70] A. Puig, *The LHCb trigger in 2011 and 2012*, LHCb-PUB-2014-046. CERN-LHCb-PUB-2014-046.

[71] A. Puig, *The HLT2 Radiative Topological Lines*, Tech. Rep. LHCb-PUB-2012-002. CERN-LHCb-PUB-2012-002, CERN, Geneva, Feb, 2012.

[72] LHCb, R. Aaij *et al.*, *First experimental study of photon polarization in radiative B^0_s decays*, Phys. Rev. Lett. **118** (2017), no. 2 021801, `arXiv:1609.2032`.

[73] Sneha Malde, *PIDCalib Packages*, `https://twiki.cern.ch/twiki/bin/view/LHCb/PIDCalibPackage`.

[74] Y. Gao, J. He, and B. Liu, *Measurement of the B^+, B^0 and B^0_s mesons production cross-sections in pp collisions at $\sqrt{s} = 7$ TeV*, LHCb-ANA-2012-046.

[75] LHCb, R. Aaij *et al.*, *Production of J/psi and Upsilon mesons in pp collisions at sqrt(s) = 8 TeV*, JHEP **06** (2013) 064, `arXiv:1304.6977`.

[76] LHCb, R. Aaij *et al.*, *Measurement of B meson production cross-sections in proton-proton collisions at $\sqrt{s}$ = 7 TeV*, JHEP **08** (2013) 117, `arXiv:1306.3663`.

[77] V. J. Rives Molina, A. Puig, and M. Calvo, *Study of b-hadron decays into two hadrons and a photon at LHCb and first observation of b-baryon radiative decays*, Jul, 2016. Presented 17 Oct 2016.

[78] A. Hoecker *et al.*, *TMVA - Toolkit for Multivariate Data Analysis*, `arXiv:physics/0703039`.

[79] S. Coquereau *et al.*, *The $B^0 \to K^{*0}\mu^+\mu^-$ selection for 3fb^{-1} of LHCb data*, Tech. Rep. LHCb-INT-2013-058. CERN-LHCb-INT-2013-058, CERN, Geneva, Nov, 2013.

[80] G. W. Corder, D. I. Foreman, *Nonparametric Statistics for Non-Statisticians: A Step-by-Step Approach*, John Wiley and Sons, Inc. (2009).

[81] Y. Xie, *Some physics and technical issues in analysis of the decay $B_s \to J/\psi K^+K^-$*, LHCb-INT-2012-017.

[82] J. E. Gaiser, *Charmonium Spectroscopy From Radiative Decays of the J/ψ and ψ'*, PhD thesis, SLAC, 1982.

[83] ARGUS, H. Albrecht *et al.*, *Search for Hadronic $b \to u$ Decays*, Phys. Lett. **B241** (1990) 278.

[84] H.-Y. Cheng, C.-W. Chiang, and A.-L. Kuo, *Updating BPP, VP decays in the framework of flavor symmetry*, Phys. Rev. **D91** (2015), no. 1 014011, `arXiv:1409.5026`.

[85] BaBar, B. Aubert *et al.*, *Search for Neutral B-Meson Decays to a0 pi, a0 K, eta rho0, and eta f0*, Phys. Rev. **D75** (2007) 111102, `arXiv:hep-ex/0703038`.

[86] HERA-B, I. Abt *et al.*, *K^{*0} and ϕ meson production in proton-nucleus interactions at $\sqrt{s} = 41.6$ GeV*, Eur. Phys. J. **C50** (2007) 315, `arXiv:hep-ex/0606049`.