Universität Bonn

Physikalisches Institut

Study of radiative $B \to X_{s+d}\gamma$ decays at Belle

Luis Pesántez

An analysis of the decay $B \to X_{s+d}\gamma$ using the complete $\Upsilon(4S)$ data sample of the Belle experiment is presented. This inclusive analysis, reconstructs only the high energy photon from the signal B meson and a lepton from the partner B, with the purposes of flavor tagging and background suppression. The branching fraction is measured for several photon energy (E_γ^*) thresholds and extrapolated to the 1.6 GeV point. For the $E_\gamma^* \geq 1.8$ GeV the measured value is $\mathcal{B}_{s\gamma} = (3.12 \pm 0.10(\text{stat}) \pm 0.19(\text{syst}) \pm 0.08(\text{model}) \pm 0.04(\text{extrap})) \times 10^{-4}$, consistent with Standard Model predictions. It allows to set a lower bound on the mass of a charged Higgs boson on the Two-Higgs-Doublet-Model of type II of $m_{H^+} \gtrsim 580$ GeV. The CP asymmetry ($\mathcal{A}^{\text{CP}}_{(s+d)\gamma}$) is measured for several photon energy thresholds, testing its independence from this choice. The result is consistent with Standard Model expectations, for the $E_\gamma^* \geq 2.1$ GeV the asymmetry is $\mathcal{A}^{\text{CP}}_{(s+d)\gamma} = (2.2 \pm 3.9(\text{stat}) \pm 0.9(\text{syst}))\%$. An analysis of the photon energy spectrum allows to extract the HQE parameters $m_b = (4.627^{+0.030}_{-0.031})$ GeV and $\mu_\pi^2 = (0.301^{+0.070}_{-0.066})$ GeV^2. The values for the first and second spectral moments are also measured for several photon energy thresholds, for $E_\gamma^* \geq 1.8$ GeV the measured values are $\langle E_\gamma^* \rangle = (2.320 \pm 0.016 \pm 0.047 \pm 0.001)$ GeV and $\Delta E_\gamma^* = (4.258 \pm 0.525 \pm 1.498 \pm 0.050) \times 10^{-2}$ GeV^2.

Physikalisches Institut der
Universität Bonn
Nussallee 12
D-53115 Bonn

BONN-IR-2017-03
February 2017
ISSN-0172-8741

Study of radiative $B \to X_{s+d}\gamma$ decays at Belle

Dissertation
zur
Erlangung des Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität
Bonn

von
Luis Pesántez
aus
Quito, Ecuador

Bonn, 12.2016

Dieser Forschungsbericht wurde als Dissertation von der Mathematisch-Naturwissenschaftlichen Fakultät der Universität Bonn angenommen und ist auf dem Hochschulschriftenserver der ULB Bonn `http://hss.ulb.uni-bonn.de/diss_online` elektronisch publiziert.

1. Gutachter:	Prof. Dr. Jochen Dingfelder
2. Gutachter:	Dr. Phillip Urquijo

Tag der Promotion:	17.02.2017
Erscheinungsjahr:	2017

Contents

CHAPTER 1

Introduction

Flavor-changing-neutral-currents have been of interest as far back as 1975 [1, 2]. This kind of decay can only proceed through loop diagrams mediated by heavy particles in the Standard Model. They are interesting decays to study CP violation and provide the possibility to search for new heavy particles appearing in the loops. Decays containing loops thus address two of the questions that occupy particle physicists: why is the amount of CP violation that we observe in nature much smaller than what one would need to create a universe dominated by matter, and how would physics beyond the Standard Model manifest itself.

In the loop-mediated $B \to X_s\gamma$ and $B \to X_d\gamma$ decays, where X_s and X_d represent all possible hadronic final states containing as s or d quark, a highly energetic photon is radiated from the loop, granting a clear signature to perform experimental measurements. In 1993 the decay $B \to K^*(892)\gamma$, one of the $B \to X_s\gamma$ final states, was observed for the first time by the CLEO collaboration [3], opening the doors to over two decades of analyses of radiative decays of the B meson at several experimental facilities. The cleanest theoretical predictions are, however, given for inclusive decays. Several experimental approaches have been devised to study radiative decays inclusively as opposed to reconstructing individual final states. In this thesis an inclusive study of the admixture of $B \to X_s\gamma$ and $B \to X_d\gamma$ decays, $B \to X_{s+d}\gamma$, will be presented.

Theoretical predictions of the $B \to X_s\gamma$ branching fraction have reached a 7 % precision, whereas the inclusive measurements have uncertainties around 11 %. The branching fraction has established itself over many years as the most powerful tool to set limits on the mass of a hypothetical charged Higgs bosons predicted by extensions of the Standard Model with and extended Higgs sector, and provides even better constrains than direct searches at LEP and the LHC. This demonstrates not only the strength and importance of flavor physics measurements, but also the complementarity of precision studies of B mesons to high-energy searches at hadron colliders.

A very interesting observable is the CP asymmetry in $B \to X_{s+d}\gamma$ decays, where the theoretical prediction is zero, as well as very small uncertainties. This means that any deviation from zero, can be interpreted as a new physics effect. Before the present analysis, the experimental precision on the CP asymmetry was around 6 %, a precision still far beyond what would be needed to observe the small deviations from zero some new physics model predict.

Additionally to these observables, with which on can test the validity of the Standard Model by performing accurate measurements, the shape of the $B \to X_s\gamma$ photon energy spectrum can be studied. The spectrum is not expected to offer a path to look for new physics, but is necessary to understand the kinematics of the b quark bound within the B meson. This is important because the selection procedure in an inclusive analysis can have different efficiencies for different photon energy regions, so accurate branching fraction measurements require reliable descriptions of the spectrum. The most relevant properties of the b quark are its mass (m_b) and mean kinetic energy inside the B meson (μ_π^2). Usually these quantities have been obtained from fits using the moments of the $B \to X_s\gamma$ photon energy spectrum (and from other B decays). The moments will be measured in this analysis too. Besides the moment measurements, a different approach is presented, in which the complete information of the shape of the spectrum is used to extract m_b and μ_π^2.

The initial question that can be asked when approaching the analysis is, of course, what is the precision required to detect new phenomena? The answer to this question is, however, not simple and presents us with different questions: what is the best precision that can be achieved with current data and where can improvements be made? Taking these two questions as a starting point, a detailed look into previous measurements has permitted the design of an analysis that builds upon current knowledge, improves some experimental techniques and presents new experimental ideas. The present analysis aims to improve the precision on several quantities and is expected to serve as a foundation for studies of radiative decays in future B physics experiments, which will certainly further improve the experimental precision and advance our knowledge of flavor physics.

This thesis starts by reviewing the theoretical framework of the Standard Model and radiative decays in Chapter 2. Subsequently the Belle detector is described in Chapter 4. The used data samples and the selection procedure are explained in Chapter 5, while the studies necessary for the correction of the various simulated samples are described in Chapter 6. The measurement of the branching fraction and evaluation of its uncertainties are included in Chapter 7. The measurement of the CP asymmetry can be found in Chapter 8. The results are discussed in Chapter 9 and the impact of the measurement on new physics scenarios is explored. In addition and an outlook on future measurements is given. This thesis closes with a summary and concluding remarks in Chapter 10.

CHAPTER 2

Theoretical framework

THE Standard Model (SM) of particle physics is a well established model that describes the properties of particles that exist in nature and the interactions between them. The particles that compose matter are fermions, particles of spin $\frac{1}{2}$, the mediators of the interactions between them are bosons, particles of spin 1. Although this model has been extremely successful and has withstood all experimental test so far, it has some shortcomings: it does not include gravitation, one of the four fundamental forces, it does not contain a large enough matter-antimatter asymmetry to explain the present composition of the universe and it does not provide explanations for dark matter and dark energy.

Several extensions to the SM have been proposed throughout the years to complement the current theory. The testing of these new theories can be performed in experimental facilities that provide large energies to directly produce new particles, or can be done by carrying out very precise measurements of known processes and looking for discrepancies from the SM expectations.

2.1 The Standard Model

Noether's theorem states that any operation that leaves a system invariant corresponds to a conservation law [4], e.g. time invariance of a system corresponds to the conservation of its energy. Two important operations that can be performed on a system are charge conjugation and parity inversion:

- `Charge conjugation` (C) inverts the sign of all additive quantum numbers of a system such as electric charge and lepton or baryon number, converting a particle into its antiparticle. It does not affect the mass, spin or momentum:

$$C\,|\psi\rangle \mapsto |\bar{\psi}\rangle \tag{2.1}$$

- `Parity inversion` (P) flips the sign of the spatial coordinates of a vector, transforming a left-handed system into a right-handed one:

$$P\psi(\mathbf{x}) \mapsto \psi(-\mathbf{x}) \tag{2.2}$$

Although it was long believed that parity was a fundamental symmetry of nature, in 1956, two experiments showed that it is maximally violated in weak decays [5, 6]. It was later thought that the combination of both operations, CP, would be conserved, but the Cronin and Fitch experiment found CP violation in the neutral kaon system [7]. These ideas and findings, together with several other theoretical and experimental outcomes, lay the foundations for the formulation of a general model that describes particles in nature.

The Standard Model of particle physics is a relativistic quantum field theory. It is motivated by both experimental findings and theoretical developments and utilizes the idea of symmetry conservation as its backbone. The local gauge invariance of the model is given by the symmetry group $\mathrm{SU}_c(3) \times \mathrm{SU}_L(2) \times \mathrm{U}_Y(1)$, it describes the transformations under the color, weak isospin and hypercharge symmetries. The group $\mathrm{SU}_L(2) \times \mathrm{U}_Y(1)$ unifies electromagnetism and the weak interaction, which is responsible for radioactive decays. The weak force carrier to couple only to left-handed particles due to parity violation. The gauge bosons that mediate electromagnetic and weak interactions are the charged $W^\pm$ bosons with a mass around 80 GeV, the neutral Z^0 boson with a mass around 91 GeV and the massless photon (γ). These bosons are superpositions of the gauge bosons of the symmetry group: $W_{1,2,3}$ and B. The strong interaction is described by the SU(3) group and its force carriers are the massless gluons [8].

The last piece of the model is the Higgs mechanism. The Higgs field ϕ is an $\mathrm{SU}_L(2)$ doublet and acquires a vacuum expectation value $\langle\phi_0\rangle$ after spontaneous symmetry breaking. The coupling of the Higgs field to the gauge bosons and the fermions gives rise to their mass terms in the SM Lagrangian, leaving the neutrinos and the photon massless. The scalar Higgs boson is the quantum excitation of the Higgs field, and has a mass around 125 GeV [9, 10].

Quarks and leptons are fermions and are organized in three "generations" composed by an up-type quark, a down-type quark, a charged lepton and a neutrino. All generations are almost identical replicas of each other, and the fermions differ above all in their masses:

$$\begin{pmatrix} u \\ d \\ e \\ \nu_e \end{pmatrix}, \quad \begin{pmatrix} c \\ s \\ \mu \\ \nu_\mu \end{pmatrix}, \quad \begin{pmatrix} t \\ b \\ \tau \\ \nu_\tau \end{pmatrix} \tag{2.3}$$

Only particles from the first generation do not decay, they make up ordinary matter. Neutrinos of all three generations do not decay. These fermion fields can be decomposed into left and right-handed components $\psi = \psi_L + \psi_R$. The main reason to do so is the left-handed nature of the weak interaction, meaning it only couples to the left-handed component of the fields. The operator that projects a field into its left or right-handed components is:

$$P_{L,R} = \frac{1}{2}\left(1 \mp \gamma^5\right), \tag{2.4}$$

where the matrix $\gamma^5 := i\gamma^0\gamma^1\gamma^2\gamma^3$. The gamma or Dirac matrices $\{\gamma^0, \gamma^1, \gamma^2, \gamma^3\}$ are 4×4 matrices which obey the anticommutation rules $\{\gamma^\mu, \gamma^\nu\} = -2\eta^{\mu\nu} I_4$ and generate a representation of the

Clifford algebra that acts on the four-dimensional Minkowsky space [11].

The charged weak currents (j^μ and J^μ), mediated by the $W^\pm$ bosons, describe the transition between charged and neutral leptons, and between up and down-type quarks. This kind of decay is called $\Delta F = 1$ because one flavor transition occurs (e.g. $c \to s$ or $e \to \nu_e$):

$$j^\mu = \bar{\ell}\gamma^\mu P_L \nu_\ell, \tag{2.5}$$

$$J^\mu = V_{q'q}\bar{q}'\gamma^\mu P_L q. \tag{2.6}$$

The Dirac fields ℓ and ν_ℓ represent the charged lepton and neutrino and the fields q' and q the quarks. Here, $V_{q'q}$ is a Cabibbo-Kobayashi-Maskawa (CKM) matrix element, which describes the strength of the transitions between up and down-type quarks. The CKM matrix is a 3 × 3 unitary matrix which describes the rotation between the "mass eigenstates", that are physically observed, and the "weak eigenstates", that take part in weak interactions:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix}_L = \begin{pmatrix} \mathrm{V}_{ud} & \mathrm{V}_{us} & \mathrm{V}_{ub} \\ \mathrm{V}_{cd} & \mathrm{V}_{cs} & \mathrm{V}_{cb} \\ \mathrm{V}_{td} & \mathrm{V}_{ts} & \mathrm{V}_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}_L \tag{2.7}$$

The matrix has nine complex entries, i.e. 18 real numbers are needed to fully describe it. Unitarity constraints, $\sum\limits_k V_{ik}V^*_{kj} = \delta_{ij}$, reduce the number of parameters to 9. The six quark fields can absorb a complex phase each, but a global phase remains, so from all these considerations, the CKM matrix can be fully described by four parameters: three rotation angles and a complex phase. Direct CP violation can be present in the Standard Model only under the presence of a complex phase, which represents the last of the four parameters of the CKM matrix.

The CKM matrix shows a strong hierarchy for the magnitude of its elements: diagonal entries have values close to unity, while off-diagonal terms are smaller, as can be seen in Eq. (2.8) where the magnitudes $|V_{ij}|$ are shown. This means that weak transitions between different generations are suppressed. It is convenient to find a parametrization that reflects these properties. The Wolfenstein parametrization as shown in Eq. (2.9) serves this purpose. Taking into account that all four parameters used in the parametrization, A, ρ, η, and λ are of comparable magnitude ($\mathcal{O}(1)$), the matrix written in this way reveals not only the magnitude of its elements but also in which elements the CP violating phase plays a significant role.

$$V_{\mathrm{CKM}} \approx \begin{pmatrix} 0.974 & 0.225 & 0.004 \\ 0.225 & 0.973 & 0.041 \\ 0.009 & 0.040 & 0.999 \end{pmatrix} \tag{2.8}$$

$$V_{\mathrm{Wolf}} = \begin{pmatrix} 1-\frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta + i\eta\frac{1}{2}\lambda^2) \\ -\lambda & 1-\frac{1}{2}\lambda^2 - i\eta A^2\lambda^4 & A\lambda^4(1+i\lambda^2\eta) \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^5) \tag{2.9}$$

One of the unitarity relations for the off-diagonal elements represents a particularly interesting triangle in the complex (ρ-η) plane. Its sides are all of comparable length, and both its angles and sides can be accessed by many different measurements. It is simply called "Unitary Triangle",

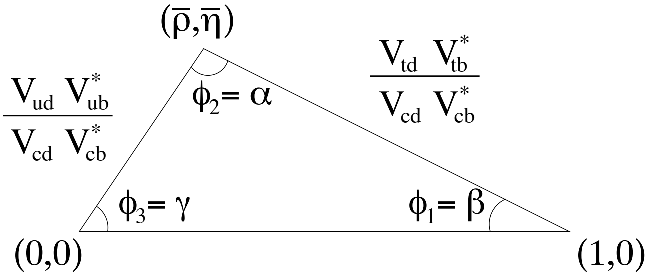


Figure 2.1: The Unitary Triangle [from 12, Figure 16.5.1]. The naming conventions for the angles correspond to those of the Belle (ϕ_1, ϕ_2, ϕ_3) and *BABAR* (α, β, γ) collaborations.

shown in Fig. 2.1 and given by the unitary condition:

$$\mathrm{V}_{ud}\,\mathrm{V}_{ub}^* + \mathrm{V}_{cd}\,\mathrm{V}_{cb}^* + \mathrm{V}_{td}\,\mathrm{V}_{tb}^* = 0. \tag{2.10}$$

The naming conventions for the angles of the Unitary Triangle correspond to those of the Belle (ϕ_1, ϕ_2, ϕ_3) and *BABAR* (α, β, γ) collaborations. The lengths of the sides and the angles can be constrained from experimental measurements such as semileptonic B decays ($|\,\mathrm{V}_{ub}\,|$, $|\,\mathrm{V}_{cb}\,|$), CP and isospin asymmetry measurements, B and B_s mixing (Δm_d, Δm_s) and others. The current status of the global determination, meaning a fit that uses as input lattice predictions as well as measurements of B, D and K meson decays, is presented in Fig. 2.2. The agreement of all measurements at the apex of the triangle is remarkable and is a strong proof of the validity of the CKM mechanism and the SM. A detailed description of the used measurements and the current status of this determination by the CKMFitter collaboration can be found in [13].

As previously mentioned, CP violation is a fundamental component of the SM. Experimentally it can be studied specifying a particular decay channel. The transition amplitude of a system ψ into a final state f, and the transition amplitude of its CP conjugate $\bar\psi$ into a final state $\bar f$, are given by:

$$\mathcal{A}_f = \langle f|\mathcal{H}|\psi\rangle, \qquad \bar{\mathcal{A}}_{\bar f} = \langle \bar f|\mathcal{H}|\bar\psi\rangle. \tag{2.11}$$

Here $\mathcal{H}$ denotes the Hamiltonian that governs the weak decay. If the magnitudes of these two amplitudes differ from each other, direct CP violation is present in the decay [14]. A distinction is made with respect to indirect CP violation, which occurs when the states ψ and $\bar\psi$ can mix and transform into one another.

A useful observable to quantify CP violation is the CP asymmetry ($\mathcal{A}^{\mathrm{CP}}$) defined as:

$$\mathcal{A}^{\mathrm{CP}} = \frac{\Gamma(\psi \to f) - \Gamma(\bar\psi \to \bar f)}{\Gamma(\psi \to f) + \Gamma(\bar\psi \to \bar f)}, \tag{2.12}$$

where the decay widths, Γ_i, are proportional to the squared amplitudes: $\Gamma_i \propto |\mathcal{A}_i|^2$.

The decay amplitudes $\mathcal{A}_f$ have an overall phase composed by a CP-odd (ϕ) and a CP-even (δ) phase. The first arises from the weak coupling constants (CKM matrix elements) and the second from the strong interaction between the quarks inside the hadron ("strong phase"). The phases

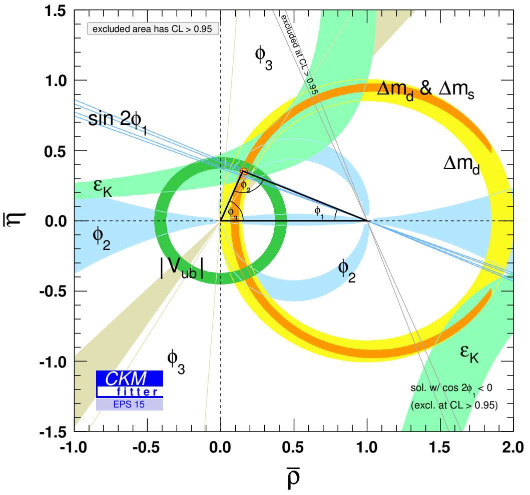


Figure 2.2: Constraints on the Unitary Triangle from a fit to several experimental measurements performed by the CKMfitter collaboration [13].

can be factored out of the amplitudes, resulting in $\mathcal{A}_f = |\mathcal{A}_f|e^{i(\phi+\delta)}$ and $\bar{\mathcal{A}}_{\bar{f}} = |\bar{\mathcal{A}}_{\bar{f}}|e^{i(-\phi+\delta)}$. The quadratic dependence of the decay width on the decay amplitudes would effectively mean the CP violating phases would always vanish for a single decay mode. For this reason it is only under the presence of both weak and strong phases and the interference of more than one decay diagram that CP violation can be realized.

2.2 The radiative $B \to X_s\gamma$ and $B \to X_d\gamma$ decays

The decays $B \to X_s\gamma$ and $B \to X_d\gamma$ are considered to be very sensitive to contributions from physics beyond the SM. They involve flavor-changing-neutral-currents (FCNC) that can only proceed via loops in the SM, as shown in Fig. 2.3. Heavy particles from new theories could appear as corrections to the SM calculation. FCNC are also called $\Delta F = 2$ decays, meaning there are two changes of flavor, the decaying quark and its decay product are both either up- or down-type ($b \to u, c, t \to s$). The X_s and X_d mean the decays are considered inclusively, i.e. all possible final states containing an s or d quark are included.

The calculation of the rate requires a careful accounting of strong (long-distance) and weak (short-distance) contributions, the latter can be calculated perturbatively. The decays of B mesons can be described by effective point-like vertices in a technique called Operator Product Expansion (OPE) [15]. It relies on integrating out contributions from heavy particles such as the top quark and electroweak bosons. The long-distance contributions are contained in the operator matrix elements O_i, which are usually of non-perturbative nature. The short-distance effects are perturbative in nature and are contained inside the corresponding Wilson coefficients C_i. In this approach, the effective Hamiltonian governing the decays can be written as [16]:

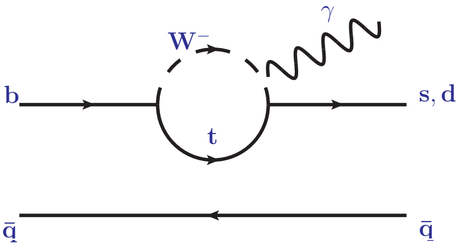


Figure 2.3: The $B \to X_s\gamma$ and $B \to X_d\gamma$ Feynman diagram.

$$\mathcal{H}_{\text{eff}} = -\frac{4G_F}{\sqrt{2}}\left[\lambda_q^t \sum_{i=1}^{8} C_i(\mu)O_i(\mu) + \lambda_q^u \sum_{i=1}^{2} C_i(\mu)\Big((O_i(\mu) - O_i^u(\mu)\Big)\right], \tag{2.13}$$

where $q = s$ or d, the CKM factors $\lambda_q^t = \mathrm{V}_{tb}\,\mathrm{V}_{tq}^*$ and $\lambda_q^u = \mathrm{V}_{ub}\,\mathrm{V}_{uq}^*$ and G_F is the Fermi constant. The renormalization scale μ separates long-distance contributions, at scales lower than μ, and short-distance contributions, at scales higher than μ. The scale is typically chosen to be of the order of the b quark's mass, m_b.

The most important contributions in radiative B decays come from the electroweak "penguin"[†] operators ($O_{7,8}$) and current-current operators ($O_{1,2}$). The QCD-penguin operators O_{3-6} have very small Wilson coefficients and can be neglected [18]. The significant operators are given by Eq. (2.14).

$$\begin{aligned}
O_1 &= (\bar{s}_L\gamma_\mu T^a c_L)(\bar{c}_L\gamma^\mu T^a b_L),\\
O_2 &= (\bar{s}_L\gamma_\mu c_L)(\bar{c}_L\gamma^\mu b_L),\\
O_7 &= \frac{e}{16\pi^2}m_b(\bar{s}_L\sigma^{\mu\nu}b_R)F_{\mu\nu},\\
O_8 &= \frac{g}{16\pi^2}m_b(\bar{s}_L\sigma^{\mu\nu}T^a b_R)G^a_{\mu\nu},
\end{aligned} \tag{2.14}$$

where T^a are the generators of the SU(3) group, $F_{\mu\nu}$ and $G^a_{\mu\nu}$ are the field strength tensors of the electromagnetic and strong fields, and $\sigma^{\mu\nu} = \frac{i}{2}[\gamma^\mu, \gamma^\nu]$. The operators $O^u_{1,2}$ are generated by replacing the charm-quark spinor c_L by an up-quark spinor u_L.

Given that the b quark is heavy, one can make use of quark-hadron duality and approximate the inclusive decay rate of the B meson by the perturbatively calculable decay rate of the b quark to partons. For this purpose, corrections for the quarks and gluons dressing the heavy b inside the meson must be calculated. These corrections are of order Λ_{QCD}/m_b or higher [19, 20]:

$$\Gamma(B \to X_s\gamma)^{E_\gamma \geq E_0} = \Gamma(b \to s\gamma)^{E_\gamma \geq E_0} + \delta\Gamma. \tag{2.15}$$

It is worthwhile noting that the decay rate is calculated for a certain photon energy threshold (E_0). This is motivated both theoretically since one needs to avoid charmonium bound states, and

---

[†] An amusing account on how the name "penguin" came to be can be read in [17].

experimentally since large $B\overline{B}$ backgrounds arise at low photon energies. The usual threshold chosen for theoretical determinations is $E_0 = 1.6\,\text{GeV}$.

The calculation of the perturbative terms that enter $\Gamma(b \to s\gamma)$ begins with the evaluation of the leading-oder (LO) diagram, which is the one loop diagram of Fig. 2.4(a). Dressing the diagram with one or two virtual gluons are examples of the diagrams found at next-to-leading order (NLO) and next-to-next-to-leading-order (NNLO) in QCD, as in Fig. 2.4(b) and (c) [21, 22]. The higher order corrections should also include gluon and light-quark bremsstrahlung. These corrections to the one-loop diagram enhance the decay rate by a factor of ~ 2, therefore it is important to determine them [21]. The determination of all contributions at NNLO is an effort that has taken several years and has been recently completed [20].

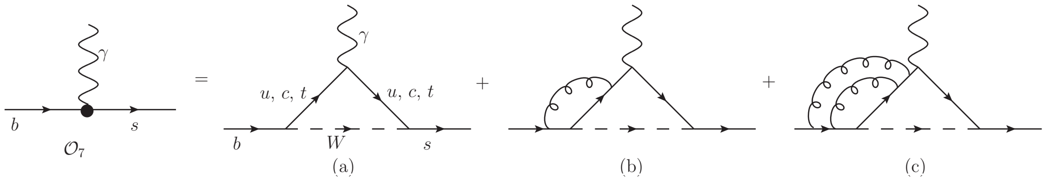


Figure 2.4: Examples of Feynman diagrams contributing to $b \to s\gamma$ at various orders.

The decay rate of the b quark to partons is proportional to $|\langle s\gamma|\mathcal{H}|b\rangle|^2$. When making use of the OPE, it will take the form:

$$\Gamma(b \to s\gamma) \propto \sum_{i,j} C_i C_j \widetilde{\Lambda}_{ij}, \tag{2.16}$$

where the functions $\widetilde{\Lambda}_{ij}$ represent the interference between the different operators $O_{i,j}$ and $C_{i,j}$ are Wilson coefficients. Particularly important is the interference between the electromagnetic dipole operator O_7 and other operators. This interference involves subprocesses, in which the photon couples to light partons via conversion, instead of connecting directly to the effective weak-interaction vertex; this effect is called the "resolved photon" contribution [23]. So far, the calculation of the Wilson coefficients up to NNLO has been completed with the calculation of the four-loop contribution [24]. Furthermore the contributions from $\widetilde{\Lambda}_{77}$, $\widetilde{\Lambda}_{88}$, $\widetilde{\Lambda}_{(1,2)7}$ and $\widetilde{\Lambda}_{(1,2)8}$ have been calculated up to NNLO, taking into account three- and four-body contributions and considering the effects of the charm and bottom masses in loops [20, and references therein]. If one assumes that the $B \to X_s\gamma$ decay is mediated by the O_7 operator alone and neglects perturbative QCD corrections, the decay rate can be written as:

$$\Gamma(B \to X_s\gamma)^{(O_7,O_7)} = \frac{G_F^2 m_b^5 \alpha_{\text{EM}}}{32\pi^4} |\,\text{V}_{ts}\,\text{V}_{tb}^*\,|^2 |C_7|^2(m_b) \times \left(1 - \frac{1}{m_b^2}\left[\frac{1}{2}\mu_\pi^2 + \frac{3}{2}\mu_G^2\right]\right), \tag{2.17}$$

where α_{EM} is the fine-structure constant and μ_π^2 and μ_G^2 are the Heavy Quark Expansion (HQE) parameters for the kinetic energy and the chromomagnetic energy [12, section 17.9].

The non-perturbative corrections to the decay rate, $\delta\Gamma$ in Eq. (2.15), receive large contributions from the resolved photon terms. The calculation of the resolved photon terms, which effectively is hadronic substructure of the photon, introduces uncertainties in the description of the photon spectrum. The functions that describe them cannot be extracted from experiment nor be calculated with current techniques, leading to an irreducible uncertainty of ±(4 - 5)%.

The $B \to X_s\gamma$ decay rate is usually normalized to a theoretically cleaner decay to cancel some terms and reduce uncertainties. Normalization to the charmless semileptonic $B \to X_u e\nu$ decays cancels non-perturbative corrections of order $1/m_b^2$, however the best choice is to normalize to the much better experimentally constrained semileptonic $B \to X_c\ell\nu$ decays. The latter choice cancels the m_b^5 dependence on Eq. (2.17) [12, section 17.9]. Following the notation of [25] the $B \to X_s\gamma$ branching fraction for an energy cut $E_\gamma \geq E_0$, and using $B \to X_c\ell\nu$ to normalize it, can be written as:

$$\mathcal{B}_{s\gamma}^{E_\gamma \geq E_0} = \mathcal{B}_{B\to X_c\ell\nu}\left|\frac{\mathrm{V}_{ts}\,\mathrm{V}_{tb}^*}{\mathrm{V}_{cb}}\right|^2 \frac{6\alpha_{\mathrm{EM}}}{\pi C}\left[P(E_0) + N(E_0)\right], \tag{2.18}$$

where the factor C accounts for the phase space difference from the choice of $B \to X_c e\nu$ over $B \to X_u e\nu$ as a normalization channel and is given by:

$$C = \left|\frac{\mathrm{V}_{ub}}{\mathrm{V}_{cb}}\right|^2 \frac{\Gamma(B \to X_c\ell\nu)}{\Gamma(B \to X_u e\nu)}. \tag{2.19}$$

The functions $P(E_0)$ and $N(E_0)$ account for the perturbative contributions in the expansion and the non-perturbative corrections, respectively. As discussed above, $P(E_0)$ can now be calculated up to NNLO, while $N(E_0)$ is still poorly known. The function $P(E_0)$ can be extracted from the ratio:

$$\frac{\Gamma(b \to s\gamma)}{|\,\mathrm{V}_{cb}\,/\,\mathrm{V}_{ub}\,|^2\,\Gamma(b \to ue\nu)} = \left|\frac{\mathrm{V}_{ts}\,\mathrm{V}_{tb}^*}{\mathrm{V}_{cb}}\right|^2 \frac{6\alpha_{\mathrm{EM}}}{\pi C} P(E_0). \tag{2.20}$$

At leading order, the perturbative function receives contributions from the Wilson coefficients $C_{(7,8)}$ and is equivalent to the effective coefficient C_7^{eff} of the operator O_7 [26]:

$$C_7^{\mathrm{eff},(0)}[\mu_0] = \eta^{\frac{16}{23}} C_{7,\mathrm{SM}}^{(0)}[\mu_0] + \frac{8}{3}\left(\eta^{\frac{14}{23}} - \eta^{\frac{16}{23}}\right) C_{8,\mathrm{SM}}^{(0)}[\mu_0] + \sum_{i=1}^{8} h_i \eta^{a_i}. \tag{2.21}$$

The values of h_i and a_i are given in Table 2.1 and $\eta = \alpha_s(\mu_0)/\alpha_s(\mu_b)$, where α_s is the running strong coupling constant evaluated at the matching scale μ_0 and the scale μ_b, where the operators are evaluated. At the renormalization scale μ_0, the Wilson coefficients take values:

$$\begin{aligned} C_{7,\mathrm{SM}}^{(0)}[\mu_0] &= -\frac{x}{2}\left(2F_1(x) + 3F_2(x)\right), \\ C_{8,\mathrm{SM}}^{(0)}[\mu_0] &= -\frac{3x}{2}F_1(x), \end{aligned} \tag{2.22}$$

with the ratio between the masses of the top quark and the $W^\pm$ boson $x = m_t[\mu_0]/m_W[\mu_0]$, a common choice for the renormalization scale is $\mu_0 = m_W$. The functions $F_{1,2}$ are given by:

i	1	2	3	4	5	6	7	8
h_i	$\frac{626126}{272277}$	$-\frac{56281}{51730}$	$-\frac{3}{7}$	$-\frac{1}{14}$	−0.6494	−0.0380	−0.0185	−0.0057
a_i	$\frac{14}{23}$	$\frac{16}{23}$	$\frac{6}{23}$	$-\frac{12}{23}$	0.4086	−0.4223	−0.8994	0.1456

Table 2.1: Coefficients for the calculation of the leading order contributions to the $b \to s\gamma$ decay rate [from 13, Table 3].

$$\begin{aligned} F_1(x) &= \frac{x^3 - 6x^2 + 3x + 2 + 6x\ln(x)}{12(x-1)^4}, \\ F_2(x) &= \frac{2x^3 + 3x^2 - 6x + 1 - 6x^2\ln(x)}{12(x-1)^4}. \end{aligned} \tag{2.23}$$

The complete form of the perturbative functions takes into account the top-quark contribution (K_t), contributions from charm loops (K_c), electroweak corrections (ε_{EM}) and the contributions from all operators from the OPE ($B(E_0)$):

$$P(E_0) = \left| K_c + \left(1 + \frac{\alpha_s(\mu_0)}{\pi} r(\mu_0) K_t + \varepsilon_{\text{EM}}\right) \right| + B(E_0). \tag{2.24}$$

Here, the factor $r(\mu_0) = m_b^{\overline{MS}}(\mu_0)/m_b^{1S}$ takes care of the different renormalization scales chosen for some kinematic factors. The function $B(E_0)$ is given by [25]:

$$B(E_0) = \frac{\alpha_s(\mu_0)}{\pi} \sum_{i,j=1,2,7,8} C_i^{(0)\text{eff}}[\mu_b] C_j^{(0)\text{eff}}[\mu_b] \widetilde{\Lambda}_{ij} \left(1 - \frac{2E_0}{m_b}\right). \tag{2.25}$$

The calculations of the $B \to X_s\gamma$ and $B \to X_d\gamma$ decay rates are identical, provided one takes care of the necessary mass and CKM element replacements. Looking back at Eq. (2.13), while the contributions λ_s^u are rather small in $B \to X_s\gamma$, λ_d^u are of the same order as λ_d^t for $B \to X_d\gamma$ [12]. The consequence of the absence of a CKM suppression in some terms is that, unlike the $B \to X_s\gamma$ case, long-distance effects play an important role in $B \to X_d\gamma$ decays and can give rise to large CP and isospin asymmetries.

Currently, the total theory uncertainty in $B \to X_s\gamma$ is around ±7 %, which can be broken down as follows: non-perturbative ±5 %, higher-order ±3 %, interpolation ±3 % and parametric ±2 % [20]. The first component is the non-perturbative part related to resolved photon contributions. The second uncertainty is related to higher-order perturbative effects in the OPE. The interpolation uncertainty is due to the m_c dependence of the NNLO corrections, which in some places is assumed to be zero. Finally, the last uncertainty comes from the precision with which input parameters such as CKM matrix elements, quark masses and others have been measured. The current theoretical predictions for the $B \to X_s\gamma$ and $B \to X_d\gamma$ branching fractions, given in Eq. (2.26) [20], show a very good agreement with the current experimental world average of Eq. (2.27) [27]. For the branching fraction $\mathcal{B}_{s\gamma}$ the agreement is within a 1σ level. The branching fraction $\mathcal{B}_{d\gamma}$ is consistent within $\sim 2.5\sigma$. The experimental value still has a large uncertainty and consists so far of a single measurement by the *BABAR* collaboration [28].

$$\begin{aligned} \mathcal{B}_{s\gamma}^{\mathrm{SM}} &= (3.36 \pm 0.23) \times 10^{-4}, \\ \mathcal{B}_{d\gamma}^{\mathrm{SM}} &= (1.73^{+0.12}_{-0.22}) \times 10^{-5}. \end{aligned} \tag{2.26}$$

$$\begin{aligned} \mathcal{B}_{d\gamma}^{\mathrm{exp}} &= (3.43 \pm 0.21(\mathrm{stat}) \pm 0.07(\mathrm{syst})) \times 10^{-4}, \\ \mathcal{B}_{d\gamma}^{\mathrm{exp}} &= (0.92 \pm 0.30) \times 10^{-5}. \end{aligned} \tag{2.27}$$

2.2.1 $B \to X_s\gamma$ photon energy spectrum

As pointed out in Eq. (2.15), a requirement on the minimum photon energy is necessary both for theoretical calculations and measurements of $\mathcal{B}_{s\gamma}$. Experimental thresholds that aim to reduce background typically range between 1.7 to 2.1 GeV. A theoretical description of the spectrum in the B rest frame, namely $d\Gamma/dE_\gamma^B$, is necessary to handle these cuts. In the free quark model, the photon from the $b \to s\gamma$ decay has a fixed energy $E_\gamma^B = m_b/2$, since in the two-body decay the photon recoils against a single quark. Corrections such as the Fermi motion of the b quark inside the B meson and gluon bremsstrahlung lead to a broadening of the photon spectrum [29, 30].

The total $B \to X_s\gamma$ decay rate can be calculated using the OPE, but once an energy cut is required several terms of the expansion have to be resumed into a non-perturbative "shape function". This function describes the light-cone momentum distribution k of the b quark inside the B meson. The shape function is, at leading order, a universal property of the B meson, and can be related to the function that describes the endpoint of the lepton spectrum in $B \to X_u e\nu$ decays at leading order in $1/m_b$. In terms of the shape function $f(x)$, the spectrum can be written as [12, section 17.9]:

$$\frac{d\Gamma}{dx} = \frac{G_F^2 m_b^6 \alpha_{\mathrm{EM}}}{32\pi^4} |\mathrm{V}_{ts}\,\mathrm{V}_{tb}^*|^2 |C_7|^2 f(1-x), \tag{2.28}$$

where $x = 2E_\gamma^B/m_b$. The $B \to X_s\gamma$ shape function leads to unavoidable uncertainties from hadronic contributions and the assumed modeling of it. These effects become more important if a higher photon energy threshold is chosen. It is therefore important that measurements have an as low as possible photon energy cut. In any case, the necessary extrapolation of the measurement to the 1.6 GeV point used in theoretical calculations also induces model-dependent uncertainties.

Several models to describe the $B \to X_s\gamma$ spectrum exist, they make different assumptions about the functional form of the shape function, and use different renormalization schemes. The schemes differ in the renormalization scale that separates soft scales (included in the matrix elements of the local operators) and hard scales (included in the Wilson coefficients), are implemented. The choice of scheme naturally affects the interpretation and evolution of quark masses. Certainly, the assumed form of the shape function can introduce model dependencies. In this study three models have been chosen to describe the photon energy spectrum: Kagan-Neubert, BLNP and BBU [31–33]. The choice of the shape function in these models is usually constrained by the expectation values of the hadronic moments defined as:

$$A_n = \int dx\, x^n\, f(x). \tag{2.29}$$

Additionally to the different models available, a model-independent determination of the shape function has been proposed [34]. In this approach the shape function $f(x)$ is expanded in a complete set of orthonormal basis functions, f_n, with coefficients c_n as:

$$f(x) = \frac{1}{\lambda}\left[\sum_{n=0}^{\infty} c_n f_n(x)\right]^2, \tag{2.30}$$

where $\lambda \sim \Lambda_{\mathrm{QCD}}$ is a dimensional parameter of the basis. The authors propose to use a basis of functions based on the Legendre polynomials. The coefficients c_n can be determined directly from data. Due to the limited experimental information, the expansion has to be truncated after a few terms, leading to a model-dependence. The authors argue that specifying a shape function that depends on few parameters underestimates the true model uncertainties, so this more generic approach, that determines most information directly from data, would not suffer from this problem. This promising method has not been used for the results of this work.

The Kagan-Neubert model (KN) [31] describes the $B \to X_s\gamma$ spectrum at NLO. Although more recent models are available, this model has historically been used in most analyses and is arguably the easiest to implement. The moments of the shape function satisfy the relations $A_0 = 1, A_1 = 0, A_2 = \frac{1}{3}\mu_\pi^2$, and the ansatz for its form is:

$$f(k) = N(1-x)^a e^{(1+a)x}, \tag{2.31}$$

where N is a normalization factor that can be fixed from the moment A_0, $A_2 = \bar{\Lambda}^2/(1+a)$, $\bar{\Lambda} = m_B - m_b$ and $x = k/\bar{\Lambda}$ with k being the momentum of the b quark. From these relations one can see that the shape function can be described by only two parameters m_b, and μ_π^2. This model choses the renormalization scale μ_b to be m_b. The KN shape function is shown in Fig. 2.5, the parameter a is chosen to be $a = 1.29$. The value of a is obtained from $m_b = 4.8\,\mathrm{GeV}$ and $\mu_\pi^2 = 0.3\,\mathrm{GeV}^2$, which are the reference values in the original paper [31].

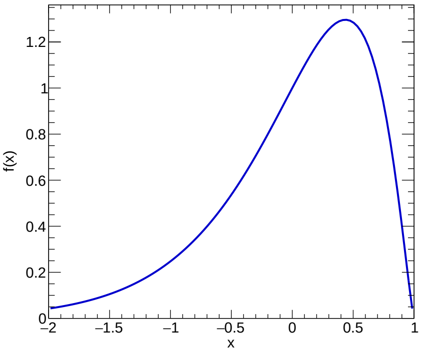


Figure 2.5: Shape function in the Kagan-Neubert model for $a = 1.29$.

The BLNP model [32, 35] is also a NLO calculation of $B \to X_s\gamma$ and uses a renormalization scheme typically known as the “shape-function scheme”, thus both names will be used indistinctly for this model. The resummation of terms in OPE is modified in this model, such that the light-cone distribution function only resums the most singular terms of the OPE in the shape function region [36]. The remaining terms are treated as sub-leading, this gives rise to four additional functions called sub-leading shape functions. These functions provide the dominant hadronic power corrections and two of them are sensitive to the heavy quark spin. There is little information about the sub-leading shape functions, beyond the values of their first few moments. The authors propose three functional forms for the leading shape function, which are more complex in form than that of the KN model: exponential form, Gaussian form and cosh form [from 35, Equation 46]. The first one will be the default used in this study. The functions depend on two parameters that can also be interpreted as the HQE parameters m_b and μ_π^2. In Fig. 2.6, the leading shape function forms are plotted as a function of $\hat{\omega}/\Lambda$ and $\hat{\omega}$. Here $\hat{\omega}$ can be related to the photon energy in the B rest frame $\hat{\omega} = m_B - 2E_\gamma^B$ and $\Lambda = \langle\hat{\omega}\rangle$.

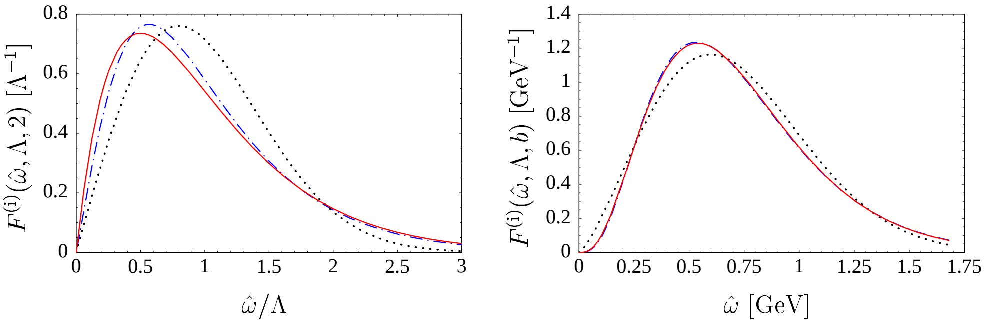


Figure 2.6: Leading shape function forms in the BLNP model. The exponential form is shown as the solid red line, the Gaussian as the dotted black line and the cosh form as the blue dash-dotted line. The left plots are given as a function of $\hat{\omega}/\Lambda$, the right plot as a function of $\hat{\omega}$ and shows the three shape forms with parameters tuned to $m_b = 4.61\,\mathrm{GeV}$ and $\mu_\pi^2 = 0.2\,\mathrm{GeV}^2$ [from 35, Figure 2].

The BBU model uses two ansätze for the shape function [33, 37], one exponential in k_+ ($F_1(k_+)$) and one exponential in k_+^2 ($F_2(k_+)$), where k_+ is the light-cone momentum carried by the b quark. The shape function ansatz $F_1(k_+)$ is the default for this study. The BBU model uses the "kinetic scheme" for the renormalization, the renormalization scale is chosen to be 1 GeV. Three parameters describe this model: m_b, μ_π^2 and μ_G^2. The comparison between the two shape function ansätze to the spectrum measured by the CLEO collaboration [38] is shown in Fig. 2.7.

Given the broad choices for models and shape functions, several of them have been considered in this study to minimize the model dependence of the result. As mentioned previously, the spectral moments are used to constrain the choice of a shape function. The definitions from Eq. (2.29) require a parametric expression and cannot be used for measured spectra which are binned. The zeroth, first and second spectral moments, representing the total number of events (which is an input to the inclusive branching fraction), the mean energy ($\langle E_\gamma^*\rangle$), and the variance of the

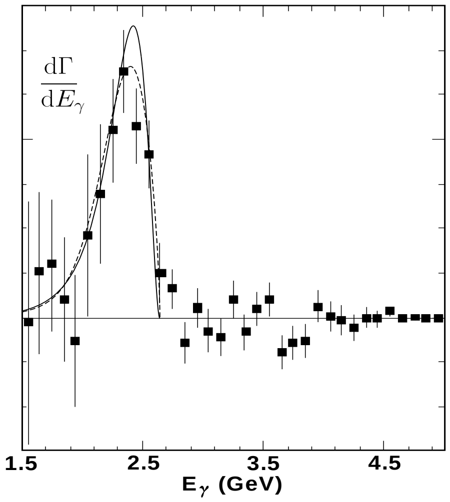


Figure 2.7: The two ansätze for the shape function in the BBU model compared to the CLEO $B \to X_s\gamma$ spectrum [38]. The solid line represents the choice $F_1(k_+)$ and the dashed line $F_2(k_+)$ [from 37, Figure 4]. The normalization of the data spectrum is arbitrary.

spectrum (ΔE^*_γ) are studied. For binned spectra they are given by:

$$N_{\text{total}} = \sum_i N_i, \tag{2.32}$$

$$\langle E^*_\gamma \rangle = \frac{\sum_i N_i (E^*_\gamma)_i}{\sum_i N_i}, \tag{2.33}$$

$$\Delta E^*_\gamma = \langle \left(E^*_\gamma\right)^2 \rangle - \langle E^*_\gamma \rangle^2 = \frac{\sum_i N_i (E^*_\gamma)^2_i}{\sum_i N_i} - \left(\frac{\sum_i N_i \left(E^*_\gamma\right)_i}{\sum_i N_i} \right)^2 . \tag{2.34}$$

2.2.2 CP violation in $B \to X_s\gamma$ and $B \to X_d\gamma$

CP violation effects in $B \to X_s\gamma$ decays are not expected to be very large due to three effects: an $\alpha_s(m_b)$ factor arising from strong phases, CKM suppression ($|\,V_{ub}\,/\,V_{tb}\,|$), and GIM suppression (m_c^2/m_b^2) [39]. If only short-distance effects are taken into account, the expected CP asymmetry $\mathcal{A}^{\text{CP}}_{s\gamma}$ is about 0.5 %. It is, however, the resolved photon contributions that ultimately dominate in the $B \to X_s\gamma$ and $B \to X_d\gamma$ asymmetries [40]. Taking all known effects into account one can write the asymmetry as:

$$\mathcal{A}^{\text{CP}}_{s\gamma} \approx \pi \left| \frac{C_1}{C_7} \right| \text{Im}\epsilon_s \left(\frac{\Lambda^u_{17} - \Lambda^c_{17}}{m_b} + \frac{40\alpha_s}{9\pi} \frac{\Lambda_c}{m_b} \right). \tag{2.35}$$

The parameters Λ_{17}^{q} represent, as discussed in Eq. (2.16), the resolved photon contributions related to the interaction between the operators O_1 and O_7 from OPE, the scale $\Lambda_c \approx 0.38\,\mathrm{GeV}$ and $\epsilon_s = \frac{\mathrm{V}_{ub}\mathrm{V}_{us}^*}{\mathrm{V}_{tb}\mathrm{V}_{ts}^*}$. The equivalent CKM factor for $B \to X_d\gamma$ would be $\epsilon_d = \frac{\mathrm{V}_{ub}\mathrm{V}_{ud}^*}{\mathrm{V}_{tb}\mathrm{V}_{td}^*}$. With the latest determination of Wolfenstein parameters [13], their central values would be $\epsilon_s \approx 0.018$ and $\epsilon_d \approx 0.285$. This absence of a CKM suppression in the $\mathcal{A}_{d\gamma}^{\mathrm{CP}}$ asymmetry explains why large asymmetries are allowed. It has been shown that the asymmetry is not very sensitive to the choice of photon energy threshold [39], the SM predicts the following ranges for the asymmetries at the $E_0 = 1.9\,\mathrm{GeV}$ threshold:

$$\begin{aligned} -0.6\,\% < \mathcal{A}_{s\gamma}^{\mathrm{CP}} < 2.8\,\%, \\ -62\,\% < \mathcal{A}_{d\gamma}^{\mathrm{CP}} < 14\,\%. \end{aligned} \tag{2.36}$$

The 2014 experimental average for the direct $\mathcal{A}_{s\gamma}^{\mathrm{CP}}$ is [27]:

$$\mathcal{A}_{s\gamma}^{\mathrm{CP}} = (1.5 \pm 2.0)\%. \tag{2.37}$$

An interesting observable is the sum of the asymmetries in X_s and X_d. In the SM there is an almost perfect cancellation of both contributions, up to small U-symmetry breaking corrections[‡], that arises from CKM unitarity:

$$\begin{aligned} \mathcal{A}_{(s+d)\gamma}^{\mathrm{CP}} &\propto \Delta\Gamma(B \to X_s\gamma) + \Delta\Gamma(B \to X_d\gamma) \\ &\propto \mathrm{Im}\left[\mathrm{V}_{ub}\,\mathrm{V}_{us}^* \cdot \mathrm{V}_{cb}^*\,\mathrm{V}_{cs} + \mathrm{V}_{ub}\,\mathrm{V}_{ud}^* \cdot \mathrm{V}_{cb}^*\,\mathrm{V}_{cd}\right] = 0. \end{aligned} \tag{2.38}$$

As a consequence of this CKM cancellation, theoretical uncertainties also cancel in the sum, making $\mathcal{A}_{(s+d)\gamma}^{\mathrm{CP}}$, which is thus predicted to be zero with negligible uncertainty in the SM, a very powerful observable to search for new physics.

2.3 New physics in $B \to X_s\gamma$ and $B \to X_d\gamma$

2.3.1 New physics in the branching fraction

New physics (NP) could modify the observables in loop-mediated FCNC transitions if heavy new particles enter in the loops. This would be translated to an enhancement of either the branching fractions or CP asymmetries predicted by the SM. A popular and simple NP model is the Two-Higgs-Doublet-Model (THDM). It introduces and additional Higgs SU(2) doublet and is the simplest extension to the Higgs sector of the SM [42]. THDMs are rather simple extensions to the SM, but their Higgs phenomenology resembles that of more complex supersymmetric models. In Supersymmetry (SUSY), scalars and their complex conjugates belong to multiplets of different chirality that cannot mix. For this reason a single Higgs doublet is not sufficient. The Minimal Supersymmetric Standard Model (MSSM) contains two Higgs doublets, so THDMs can explore the MSSM-like Higgs structure.

[‡] The U-symmetry condition used in the calculations is that $m_d = m_s$ [41].

An additional Higgs doublet gives rise to additional degrees of freedom that, after spontaneous symmetry breaking, become five massive Higgs fields: two charged scalars ($H^{\pm}$), two neutral scalars (H^0, h) and a pseudoscalar (A). THDMs can in general show very different phenomenologies, the most general scalar potential contains 14 parameters and can have CP-conserving, CP-violating and charge-violating minima. The THDM of type II (THDM-II) is built such that tree-level FCNC are not allowed. CP-violating interactions arise only from the Yukawa couplings. Up-type quarks are chosen to couple to one of the Higgs doublets and down-type quarks and leptons to the other. The new charged Higgs boson would enter in the leading-order FCNC loop diagram as shown in Fig. 2.8.

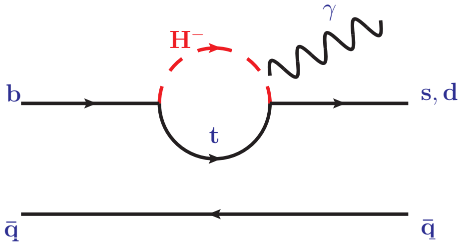


Figure 2.8: The leading order THDM contribution to $B \to X_{s+d}\gamma$.

After spontaneous symmetry breaking, the two Higgs doublets acquire vacuum expectation values υ_1 and υ_2. The observable $\tan\beta$ is defined as the ratio $\tan\beta = \upsilon_2/\upsilon_1$. Since the additional degrees of freedom of the THDM are all heavy, they modify the Wilson coefficients C_i, and the NP contribution can be factored out as [43]:

$$C_i = C_i^{\mathrm{SM}} + C_i^{\mathrm{THDM}}. \tag{2.39}$$

The additional NP contributions depend on two parameters: $\tan\beta$ and the mass of the charged Higgs m_{H^+}. The observable of choice is $\mathcal{B}_{s\gamma}$, which can be strongly modified by m_{H^+}, but is fairly insensitive to $\tan\beta$ larger than 2, as can be seen in Fig. 2.9.

A dedicated study of the THDM-II phenomenology in the light of $B \to X_s\gamma$ physics [44] has

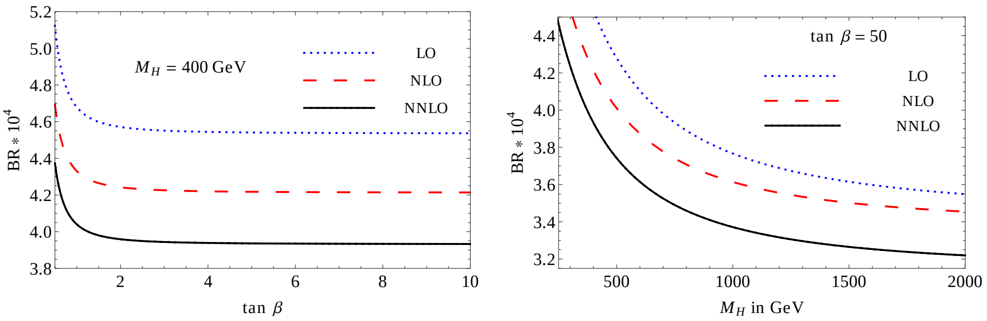


Figure 2.9: $\mathcal{B}_{s\gamma}$ in the THDM-II as a function of $\tan\beta$ (left) and m_{H^+} (right) [from 43, Figure 5].

made use of a convenient parametrization to express the branching fraction as given in Eq. (2.40):

$$P(E_0) + N(E_0) = \left(C^{\mathrm{eff},(0)}_{7,\mathrm{SM}} + B\Delta C^{\mathrm{eff},(0)}_{7,H^+}\right)^2 + A. \tag{2.40}$$

Here the leading-order coefficient $C^{\mathrm{eff},(0)}_{7,\mathrm{SM}}$ is the same as in Eq. (2.21), the coefficients A and B are chosen such that this parametric form matches the full calculation given that only the leading-order terms are used, and $\Delta C^{\mathrm{eff},(0)}_{7,H^+}$ is introduced by the new diagrams with the heavy H^+. The modification to the Wilson coefficients given in Eq. (2.41) introduces some new terms and the ratio $x_{tH} = (m_t[\mu_0]/m_{H^+})^2$, which are summarized in Eqs. (2.42) and (2.43):

$$\Delta C^{\mathrm{eff},(0)}_{7,H^+} = \eta^{\frac{16}{23}} C^{(0)}_{7,H^+}[\mu_0] + \frac{8}{3}\left(\eta^{\frac{14}{23}} - \eta^{\frac{16}{23}}\right) C^{(0)}_{8,H^+}[\mu_0], \tag{2.41}$$

where

$$\begin{aligned} C^{(0)}_{7,H^+}[\mu_0] &= -\frac{x_{tH}}{2}\left(\frac{1}{\tan^2\beta}\left(\frac{2}{3}F_1(x_{tH}) + F_2(x_{tH})\right) + \frac{2}{3}F_3(x_{tH}) + F_4(x_{tH})\right), \\ C^{(0)}_{8,H^+}[\mu_0] &= -\frac{x_{tH}}{2}\left(\frac{F_1(x_{tH})}{\tan^2\beta} + F_3(x_{tH})\right), \end{aligned} \tag{2.42}$$

and

$$\begin{aligned} F_3(x) &= \frac{x^2 - 4x + 3 + 2\ln(x)}{2(x-1)^3}, \\ F_4(x) &= \frac{x^2 - 1 - 2x\ln(x)}{2(x-1)^3}. \end{aligned} \tag{2.43}$$

In the perturbative regime $\tan\beta$ takes values between 0 and ~60. The branching fraction is virtually insensitive to most values of $\tan\beta$, except for those below ~ 2. The contributions that enter the terms in Eq. (2.42) are plotted in Fig. 2.10 for m_{H^+} up to roughly 1 TeV. The functions that enter $C^{(0)}_{8,H^+}[\mu_0]$ drop off very quickly, while the functions that enter $C^{(0)}_{7,H^+}[\mu_0]$ stabilize at around $m_{H^+} \approx 4m_t[\mu_0]$. The flat behavior of $C^{(0)}_{7,H^+}[\mu_0]$ hints that when constraining m_{H^+} from experimental measurements, the limit could have large shifts even for small variations on the central value of the measured $\mathcal{B}_{s\gamma}$. From the current experimental average [27] and the latest theoretical predictions, a 95 % confidence limit of $m_{H^+} \geq 480\,\mathrm{GeV}$ is found [20].

2.3.2 New physics in *CP* and isospin asymmetry

The direct *CP* asymmetries for the decays $B \to X_{s+d}\gamma$, $B \to X_s\gamma$ and $B \to X_d\gamma$, in a general NP model such as a flavor-blind MSSM, follow a strict proportionality to one another [45]. New *CP* violating effects enter the C_7 and C_8 coefficients, which are real in the SM, but receive an imaginary part in NP scenarios. Although in very generic models $\mathcal{A}^{\mathrm{CP}}_{s\gamma}$ could be as large ±10 %, constraints from electron and neutron electric dipole moment reduce the possible *CP* asymmetry to the range −2 % to 2 %. For generic supersymmetric models the electric dipole moment constraints are less severe since additional phases have to be taken into account. The

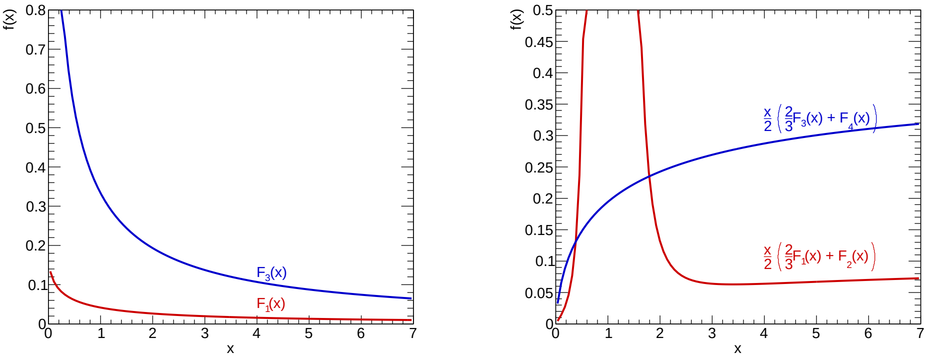


Figure 2.10: Functions that describe the Wilson coefficients in the SM (left) and in the THDM-II (both images).

results of a numerical analysis of the CP asymmetries is shown in Fig. 2.11.

The isospin asymmetry in the SM defined in Eq. (2.44) is also an important observable since the resolved photon contributions introduce flavor-dependent terms. This observable could be as large as 10 % in some NP scenarios if C_7 or C_8 receive new CP violating phases [40]. The challenge with this observable is the difficulty on achieving a precise determination of the operator $\widetilde{\Lambda}_{78}$.

$$\Delta\mathcal{A}^{\mathrm{CP}}_{s\gamma} \equiv \mathcal{A}^{\mathrm{CP}}_{B^+\to X_s^+\gamma} - \mathcal{A}^{\mathrm{CP}}_{B^0\to X_s^0\gamma} \approx 4\pi^2\alpha_s \frac{\widetilde{\Lambda}_{78}}{m_b}\mathrm{Im}\frac{C_8}{C_7}. \tag{2.44}$$

The C_7 and C_8 coefficients have recently been investigated in more exotic scenarios such as the the minimal Randall-Sundrum (RS) model [46, 47]. In this model the penguin loop receives corrections from the scalar parts of the $W^{\pm}$-boson penguin diagrams, the scalar component of the 5D gauge-boson field and from the charged Higgs boson. Given that many unconstrained parameters arise, it is hard to draw conclusions about the probed parameter space that can be rejected based on the prediction and measurements, $\mathcal{B}_{s\gamma}$ is sensitive to the mass of the first Kaluza-Klein gluon resonance. A bound of 3.8 TeV is found for $y_* = 3$, where y_* is an upper bound for the 5D-Yukawa couplings of the model. Possible corrections to $\mathcal{A}^{\mathrm{CP}}_{s\gamma}$ are too small to constraint the RS parameter space.

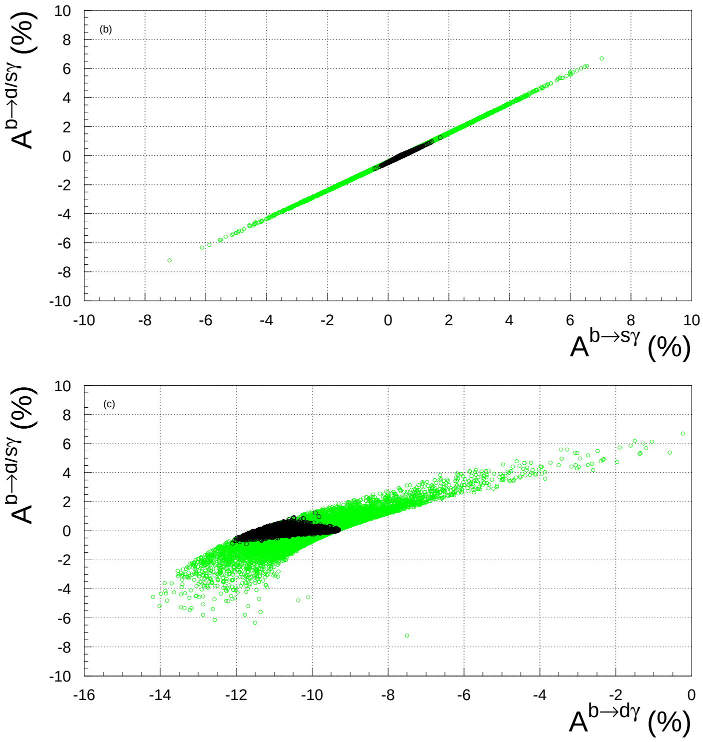


Figure 2.11: Direct CP asymmetry in the minimal flavor violating scenario [from 45, Figure 2]. The constraint from electron and neutron electric dipole moments is relaxed for the green points and imposed on the black ones. The figures display the clear proportionality between CP asymmetries in $B \to X_{s+d}\gamma$ vs. $B \to X_s\gamma$ and in $B \to X_{s+d}\gamma$ vs. $B \to X_d\gamma$.

CHAPTER 3

Overview of experimental approaches and previous measurements

The X_s and X_d final states are a short way of writing “all final states with an s or d quark from a $b \to s$ or $b \to d$ transition”. They are composed of resonances (like $K^*(892)$ in X_s or ω in X_d), and non-resonant states (such as combinations of a kaon and one or more pions for X_s, or several pions for X_d). The resonant and non-resonant states often share phase space regions, so it is not easy to disentangle them. The X_s final state has a total strangeness $S = \pm 1$, strangeness is a quantum number defined as 1 for the s quark, -1 for the $\bar{s}$ quark and 0 for all others. The branching fractions of single resonant and non-resonant states have been measured and are summarized in Tables 3.1 and 3.2. The most important X_s exclusive final state is the $K^*(892)$ resonance, which makes up about 12 % of the total X_s phase space. This resonance is very narrow and is easy to identify, its branching fraction has been, consequently, precisely measured. Measurements of the $B \to X_d\gamma$, $B \to X_s\gamma$ and $B \to X_{s+d}\gamma$ decays can be exclusive, meaning that specific final states are reconstructed, or inclusive, meaning that no particular X_s or X_d final state is specified and the selection aims to include all possible decays.

The decays have been studied at the B-factories, where B mesons are created in pairs from the reaction $e^-e^+ \to \Upsilon(4S) \to B\overline{B}$, the $\Upsilon(4S)$ is a $b\bar{b}$ bound state with a mass $m_{\Upsilon(4S)} = 10.58\,\mathrm{GeV}$. The partner B can be used to improve the suppression of background. The full or partial reconstruction of the partner B meson is called “tagging” and can be used to reduce background. The partner B meson in the reaction is hence called “tag B”, whereas the B meson of interest is called “signal B”. An important background source for the measurement of interest is continuum processes $e^-e^+ \to q\bar{q}$, which produce light quarks (u, d, s, c) in e^-e^+ annihilation. Hadron colliders cannot study $B \to X_s\gamma$ decays inclusively as can be done with e^-e^+ machines with a well defined initial energy and momentum.

The inclusive and sum of exclusive final states methods have been exploited by Belle and *BABAR* and have reached a very high precision, currently comparable to the precision of theoretical determinations. In the inclusive approach only the photon is reconstructed, comprising thus all possible X_{s+d} final states. In the sum of exclusive final states, several final states are reconstructed, in an attempt to make the measurement as inclusive as possible. The third method, recoil-tag method, uses a full reconstruction of the partner B and has only been studied by *BABAR*. It

Mode	$\mathcal{B} \times 10^6$
$B \to X_s\gamma$ †	349 ± 19
$B \to X_s\gamma$ §	345 ± 30
$B^+ \to K^{*+}(1680)\gamma$	$66.7\ ^{+17.1}_{-13.8}$
$B^+ \to K^0\pi^+\pi^0\gamma$	45.6 ± 5.2
$B^+ \to K_1^+(1270)\gamma$	$43.8\ ^{+7.1}_{-6.3}$
$B^+ \to K^{*+}(892)\gamma$	42.1 ± 1.8
$B^+ \to K^{*+}(1410)\gamma$	$27.1\ ^{+8.0}_{-6.1}$
$B^+ \to K^+\pi^-\pi^+\gamma$	24.6 ± 1.3
$B^+ \to K^{*0}\pi^+\gamma$	$23.3\ ^{+1.2}_{-1.1}$
$B^0 \to K^{*0}\gamma$	43.3 ± 1.5
$B^0 \to K^+\pi^-\pi^0\gamma$	40.7 ± 3.8
$B^0 \to K^0\pi^+\pi^-\gamma$	19.5 ± 2.2
$B^0 \to K_2^*(1430)^0\gamma$	12.4 ± 2.4

Table 3.1: Summary of measured branching fractions of exclusive $B \to X_s\gamma$ states [48]. Only modes with measured branching fractions larger than 10^{-5} are included. The † and § $\mathcal{B}_{s\gamma}$ averages differ, since the first one takes the lowest threshold from each measurement, while the second one averages all measurements at the 1.9 GeV threshold.

Mode	$\mathcal{B} \times 10^6$
$B \to X_d\gamma$	9.2 ± 3.0
$B^+ \to \rho^+\gamma$	$0.98^{+0.25}_{-0.24}$
$B^0 \to \rho^0\gamma$	$0.86^{+0.15}_{-0.14}$
$B^0 \to \omega\gamma$	$0.44^{+0.18}_{-0.16}$

Table 3.2: Summary of measured branching fractions of exclusive $B \to X_d\gamma$ states [48].

is far from reaching the precision of the other methods with current data sets due to the low selection efficiency inherent to the full reconstruction. The recoil-tag method is expected to yield the most precise measurements in the future, once the new generation B-factory SuperKEKB starts delivering collisions and the Belle II experiment, starts taking data. Figure 3.1 shows the possibilities for the selection of particles from the signal and tag B mesons, which are described in detail below.

3.1 Inclusive method

The inclusive selection is essentially based on the selection of a high-energy photon from the signal B. A high-energy lepton from the tag side can also be required, this serves as a way to reduce continuum background and specify the flavor of the B mesons in order to measure $\mathcal{A}^{\mathrm{CP}}$. A lepton is an obvious choice since it could not be part of the X_d or X_s hadronic components and can only arise from the tag side or background. Given that neither B meson is fully reconstructed, the analysis has to be performed in the $\Upsilon(4S)$ rest frame (or center-of-mass (CM) frame). The inclusive analysis also cannot distinguish between $B \to X_s\gamma$ and $B \to X_d\gamma$ contributions, thus

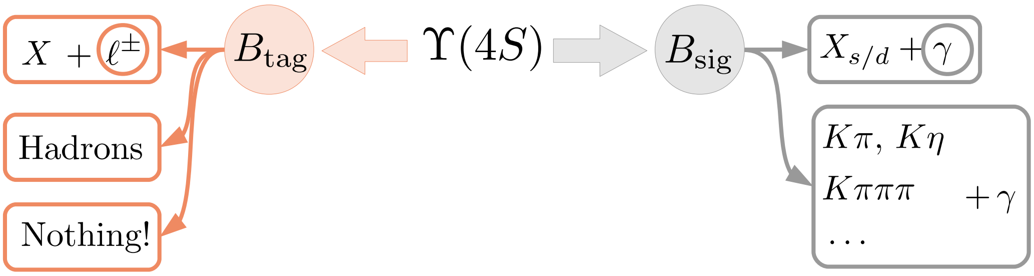


Figure 3.1: Possible selection on the signal and tag B mesons in the $B \to X_s\gamma$ and $B \to X_{s+d}\gamma$ analyses. The signal B decay can either be identified by reconstructing solely the high energy photon or reconstructing several exclusive decay modes. The tag side can be used either fully reconstructed in hadronic modes, only by selecting a charged lepton, or not reconstructed at all.

one can only expect to measure $B \to X_{s+d}\gamma$ yields, and correct for the $B \to X_d\gamma$ admixture if desired. The $B \to X_d\gamma$ amounts to roughly $|\mathrm{V}_{td}\,/\,\mathrm{V}_{ts}|^2 \approx 4.5\,\%$.

This approach has been used by the CLEO, Belle and *BABAR* collaborations. The CLEO analysis additionally uses a pseudo-reconstruction of the X_s states, and the presence of a lepton in the event as background suppression parameters. The pseudo-reconstruction basically consist of combining a kaon and one to four pions to form the X_s hadronic part and check if its combination with the photon is compatible with the mass and energy of a B meson. An score for the reconstruction is calculated based on the nominal B mass and the beam energy, only this score is used for continuum suppression and is not used in later stages of the analysis. The CLEO analysis measures $\mathcal{B}_{s\gamma}$ and the spectral moments [38]. An additional inclusive analysis from CLEO also measures $\mathcal{A}^{\rm CP}_{s\gamma}$ [49]. The latest Belle study performs in parallel an analysis with and without a tag-lepton. It measures $\mathcal{B}_{s\gamma}$ and the spectral moments using a data sample $\sim 15\,\%$ smaller than the present analysis [50]. The most recent *BABAR* analysis includes a lepton requirement and, additionally to the measurement of $\mathcal{B}_{s\gamma}$ and spectral moments, measures $\mathcal{A}^{\rm CP}_{(s+d)\gamma}$ [51].

The large photon background is a major challenge in this method, particularly in the low energy region (photon energies ~ 2.0 GeV and below). The background from B decays drops off very quickly beyond ~ 2.0 GeV due to kinematic constraints from B decays, however photons from continuum processes can extend even beyond ~ 4.0 GeV as shown in Fig. 3.2. The photon must be geometrically isolated from other calorimeter clusters and tracks, also when combined with other photons it should not be compatible with the decay of the light mesons $\pi^0/\eta \to \gamma\gamma$.

Background from continuum events is hard to model and is thus estimated using data samples taken below the $B\overline{B}$ production threshold. Since these samples are not very interesting from a physics point of view beyond background modeling, only samples roughly 10 times smaller than the $\Upsilon(4S)$ sample for Belle and *BABAR* were recorded. The drawback is that this background is also the largest one for the analysis so it limits the precision of the measurement. For this reason dedicated suppression techniques must be applied. There are many event topology variables that can be defined and used to very effectively reduce continuum background with the use of machine learning techniques. Besides continuum, the main background sources are $\pi^0/\eta \to \gamma\gamma$ decays,

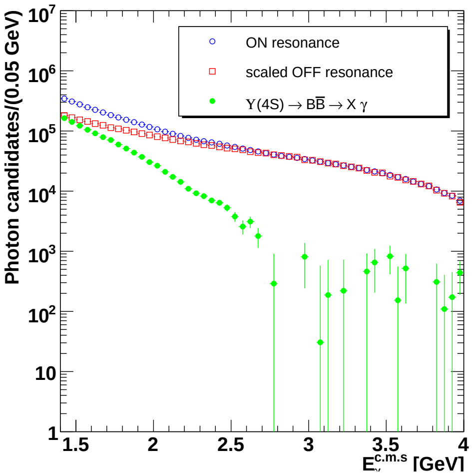


Figure 3.2: Photon energy spectrum before background subtraction [from 12, Figure 17.9.2]. The plot shows the relation between $B\overline{B}$ and $q\overline{q}$ events in the Belle analysis without a lepton tag.

true photons from ω, η' and radiative semileptonic decays. Alongside calorimeter clusters from electrons, anti-neutrons and K_L can be misidentified as photon candidates.

CLEO measures the branching fraction for a single threshold $E_\gamma^B \geq 2.0\,\text{GeV}$:

$$\mathcal{B}_{s\gamma}(E_\gamma^B \geq 2.0) = \left(3.21 \pm 0.43(\text{stat}) \pm 0.27(\text{syst})^{+0.18}_{-0.10}(\text{theo})\right) \times 10^{-4}. \tag{3.1}$$

The Belle analysis measures $\mathcal{B}_{s\gamma}$ for thresholds $E_\gamma^B = 1.7\,\text{GeV}$ to $2.1\,\text{GeV}$ [50]:

$$\mathcal{B}_{s\gamma}(E_\gamma^B > 1.7) = (3.41 \pm 0.15(\text{stat}) \pm 0.40(\text{syst})) \times 10^{-4}. \tag{3.2}$$

BaBar measures the branching fraction for thresholds $E_\gamma^B = 1.8\,\text{GeV}$ to $2.1\,\text{GeV}$ [51]:

$$\mathcal{B}_{s\gamma}(E_\gamma^B > 1.8) = (3.21 \pm 0.15(\text{stat}) \pm 0.29(\text{syst}) \pm 0.08(\text{extrap})) \times 10^{-4}, \tag{3.3}$$

where the last uncertainty corresponds to the model uncertainty from the extrapolation of the result from the 1.8 GeV threshold to 1.6 GeV, which is the threshold used for theoretical calculations.

For the asymmetry measurement, the dominant systematic uncertainties from background subtraction vanish in the ratio. The only relevant systematic effects come from possible asymmetries in the detection and reconstruction of particles, and asymmetries in the background. A systematic uncertainty arises from the probability of a wrong determination of the B flavor. CLEO measures $\mathcal{A}^{\text{CP}}_{(s+d)\gamma}$ for the $E_\gamma^B \geq 2.2\,\text{GeV}$ threshold, where a multiplicative uncertainty is presented separately:

$$\mathcal{A}^{\text{CP}}_{(s+d)\gamma} = (-7.9 \pm 10.8(\text{stat}) \pm 2.2(\text{syst}))\,(1.00 \pm 0.03)\%. \tag{3.4}$$

BABAR measures $\mathcal{A}^{\mathrm{CP}}_{(s+d)\gamma}$ for the $E^B_\gamma \geq 2.1\,\mathrm{GeV}$ threshold:

$$\mathcal{A}^{\mathrm{CP}}_{(s+d)\gamma} = (5.7 \pm 6.0(\mathrm{stat}) \pm 1.8(\mathrm{syst}))\,\%. \tag{3.5}$$

The inclusive experimental technique with a lepton tag has been chosen to perform the study presented here. It builds upon the previous Belle measurement of [50, 52]. For this reason it is important to take a more detailed look at the achieved precision. As pointed out previously, the old measurement performs in parallel analyses with and without a lepton tag, so it is only the uncertainties of the later that are considered in this discussion. The total uncertainties on the branching fraction for the 1.7 GeV threshold are 6.1 %(stat) and 14.4 %(syst). The most important systematic uncertainty is related to continuum background, which amounts to 10.3 %. This uncertainty is related to the correction factors applied to this sample, and are proportional to the total $q\overline{q}$ yield inside the signal region, so it is clear that a better reduction of this background component could greatly improve the precision of the measurement. Continuum can be better reduced making use of the topological differences with respect to $B\overline{B}$ events and exploring machine learning techniques.

The prospect is different for other uncertainties related to the irreducible $B\overline{B}$ background. There is no topological handle for background photons from $B\overline{B}$ decays, which usually arise from the decays of π^0, η, ω, η' mesons and others. Additional background arises from beam and bremsstrahlung photons, and from misidentified hadron or electron clusters in the calorimeter. Some of these background sources can be vetoed by e.g. trying to reconstruct the mother meson or studying the cluster properties, but an irreducible component is always present. The uncertainty related to π^0/η background is 1.9 % and several other $B\overline{B}$ components 7.2 %. The trade-off between better experimental precision by choosing a higher photon threshold against the increase in model uncertainty because of this, is a point where theorists and experimentalists do not completely agree. For this reason it is important to obtain results for several photon energy thresholds.

3.2 Sum of exclusive final states

In this approach as many exclusive X_s final states as possible are reconstructed. The latest Belle and *BABAR* analyses reconstruct 38 exclusive modes, amounting to roughly 70 % of the total branching fraction. Isospin relations are considered when estimating this 70 %, since one usually reconstructs K_S final states and assumes that final states containing a K_L are of equal importance. The decay modes are reconstructed as combinations of one or 3 kaons, up to four pions and at most one η meson. The main challenges of this method are the cross-feed between the modes and the proper determination of selection efficiencies for each reconstructed mode, which rely on a good understanding of the hadronization of the s quark. This analysis approach does not include an X_d contribution, and allows for the direct measurement of $\mathcal{A}^{\mathrm{CP}}$ and isospin asymmetry for the $B \to X_s\gamma$ transition.

The main discriminating variable to separate signal and background is the beam-constrained mass, m_{bc}[†] . Additionally one can use the difference ΔE between the reconstructed and expected

---

[†] *BABAR* uses the energy-substituted mass m_{ES}, which is equivalent to the m_{bc} used by Belle.

B meson energy to constrain the kinematics of the reconstructed B. They are defined as:

$$\begin{aligned} \Delta E &\equiv E_B - E_{\text{beam}}/2 \\ m_{\text{bc}} &\equiv \sqrt{(E_{\text{beam}}/2)^2 - (p_B)^2}. \end{aligned} \tag{3.6}$$

Here E_B and p_B are the reconstructed B meson energy and momentum, and E_{beam} is the sum of energies of the electron and positron. For correctly reconstructed signal events, ΔE peaks around 0 and m_{bc} around the nominal B mass of 5.28 GeV. An example distribution for m_{bc} from [53] is presented in Fig. 3.3. Important systematic uncertainties are related to the proper determination of the shapes of signal and background distributions. The analyses that use this approach also makes use of machine learning techniques for background suppression.

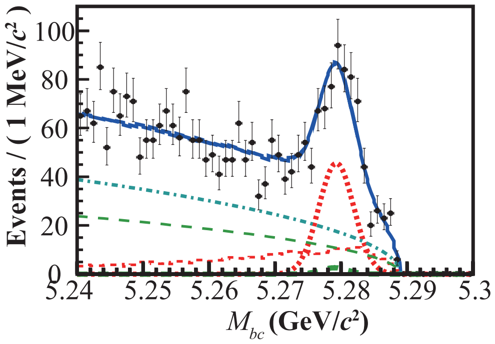


Figure 3.3: Beam-constrained mass in the Belle analysis with sum of exclusive final states [53]. The figure corresponds to the hadronic mass bin 1.1 GeV $\leq m_{X_s} \leq$ 1.1 GeV. The plot shows the data (black points), and the total MC (blue solid line). The MC components correspond to signal (red, thick short-dashed line), cross-feed (red, thin short-dashed line), peaking $B\overline{B}$ background (green thick long dashed line), non-peaking $B\overline{B}$ background (green, thin long-dashed line) and continuum background (blue, dot-dashed line)

The result of the Belle branching fraction measurement is given for the $E_\gamma^B \geq 1.9$ GeV threshold [53]:

$$\mathcal{B}_{s\gamma}(E_\gamma^B > 1.9) = (3.51 \pm 0.17(\text{stat}) \pm 0.33(\text{syst})) \times 10^{-4}. \tag{3.7}$$

BABAR measures the branching fraction for the same threshold [54], as well as CP and isospin asymmetries for 16 of the 38 reconstructed exclusive states [55]:

$$\begin{aligned} \mathcal{B}_{s\gamma}(E_\gamma^B > 1.9) &= (3.29 \pm 0.19(\text{stat}) \pm 0.48(\text{syst})) \times 10^{-4}, \\ \mathcal{A}_{s\gamma}^{\text{CP}} &= (1.7 \pm 1.9(\text{stat}) \pm 1.0(\text{syst}))\,\%, \\ \Delta\mathcal{A}_{s\gamma}^{\text{CP}} &= (5.0 \pm 3.9(\text{stat}) \pm 1.5(\text{syst}))\,\%. \end{aligned} \tag{3.8}$$

BABAR studied $B \to X_d\gamma$ decays using the sum of exclusive final states [28]. The X_d final state is reconstructed in seven final states in which two to four pions or a pion and an η are combined with a γ. This result is statiscally limited, which is not surprising since the branching fraction for this decay is expected to be ~ 25 times smaller than that of $B \to X_s\gamma$. The analysis uses photons with energies in the range $1.15 \leq E_\gamma^* \leq 3.50$GeV, but the signal region is constrained based on the mass of the reconstructed hadronic system m_{X_d}. Belle has not studied the $B \to X_d\gamma$ decay inclusively, but only its resonant components: $\rho^{\pm,0}\gamma$ and $\omega\gamma$ [56]. Similarly to the *BABAR* measurement, a broad energy range is allowed for the selected photons. The measured branching fraction values are given in Table 3.2. For comparison purposes the $B \to K^*\gamma$ branching fraction [27], which is part of $B \to X_s\gamma$, is also included in the table. Belle has also presented measurements of $\mathcal{A}^{\mathrm{CP}}$ and the isospin asymmetry (Δ) in the $B \to \rho\gamma$ decay channel:

$$\begin{aligned} \mathcal{A}^{\mathrm{CP}}_{\rho\gamma} &= -0.11 \pm 0.32 \pm 0.09, \\ \Delta(\rho\gamma) &= -0.48\ {}^{+0.21}_{-0.19}\ {}^{+0.08}_{-0.09}. \end{aligned} \tag{3.9}$$

3.3 Recoil-tag method

B mesons decay into purely hadronic final states about 5 % of the time. Sophisticated methods have been devised to reconstruct a B meson into hundreds of hadronic final states, such that the kinematics and properties of one B meson are fully determined. The remaining particles correspond to the decay of the other B in the event [12, 57]. Full hadronic reconstruction suffers from a very small reconstruction efficiency, typically 0.2 % to 0.4 %. Analyses using this method typically have a smaller $q\overline{q}$ background component, but the study of rare decays using a method with low efficiency results in poor statistical precision. Only *BABAR* has studied $B \to X_s\gamma$ decays with this approach, the analysis reaches a 0.3 % signal selection efficiency [58]. Given that one has information of the B flavor, additionally to the branching fraction, isospin and CP asymmetries can readily be measured. Both asymmetries are measured for $E_\gamma^B \geq 2.2$ GeV, while the branching fraction is measured for $E_\gamma^B \geq 1.9$ GeV:

$$\begin{aligned} &\mathcal{B}_{s\gamma} = (3.66 \pm 0.85(\mathrm{stat}) \pm 0.60(\mathrm{syst})) \times 10^{-4}, \\ &\mathcal{A}^{\mathrm{CP}}_{(s+d)\gamma} = (10 \pm 18(\mathrm{stat}) \pm 5(\mathrm{syst}))\,\%, \\ &\Delta(B \to X_{s+d}\gamma) = (6 \pm 15(\mathrm{stat}) \pm 7(\mathrm{syst}))\,\%. \end{aligned} \tag{3.10}$$

CHAPTER 4

The Belle experiment

B-factories are dedicated colliders that produce large quantities of $B\overline{B}$ pairs with the aim of carrying out high-precision measurements of the Standard Model using B decays. They are electron-positron (e^-e^+) colliders with energy tuned to the mass of the $\Upsilon(4S)$ resonance. The $\Upsilon(4S)$ resonance is a $b\bar{b}$ bound state and decays over 96 % of the time into a $B\overline{B}$ pair. Three e^-e^+ B-factories have been built until now [59]: CESR with the CLEO experiment at Cornell, the PEP-II collider *BABAR* experiment at Stanford and the KEKB collider with the Belle experiment in Tsukuba. Belle completed its operation in 2010 after recording $\sim 1\,\mathrm{ab}^{-1}$ of data, $710\,\mathrm{fb}^{-1}$ of it on the $\Upsilon(4S)$ resonance [60]. In the last years of operation, the KEKB collider achieved a world-record instantaneous luminosity of $2.11 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. Currently the only running B-physics experiment is the LHCb experiment at CERN, which studies B mesons produced in proton-proton collisions. The successor to the Belle experiment, the Belle-II experiment, is currently under construction and is expected to start taking data in 2018.

4.1 The KEK-B accelerator

The KEKB accelerator is an asymmetric-energy e^-e^+ collider [61]. The beam energy asymmetry permits to produce B mesons with a large boost. This is important to distinguish the decay vertices, which is necessary for time dependent measurements, e.g. time-dependent CP violation. The accelerator runs nominally at a center-of-mass (CM) energy $\sqrt{s} = 10.58\,\mathrm{GeV}$, with an 8 GeV e^- high-energy ring (HER) and a 3.5 GeV e^+ low-energy ring (LER), both 3 km in length. At this CM energy corresponding to the mass of the $\Upsilon(4S)$, $710\,\mathrm{fb}^{-1}$ of "on-resonance" data were recorded. Additionally $\sim 90\,\mathrm{fb}^{-1}$ of "off-resonance" data were recorded at $\sqrt{s} = 10.52\,\mathrm{GeV}$ to study non-$B\overline{B}$ backgrounds. This second data sample is also called the "off-resonance" sample where the CM energy is below the B meson production, consisting mainly of events where light quarks are produced ($q = u, d, s, c$). The beams are designed to cross at an angle of 22 mrad at the Interaction Point (IP), where the Belle detector is located. As of June of 2009, the highest instantaneous luminosity of an e^-e^+ machine was achieved: $2.11 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, which is twice as large as the design luminosity [62]. A schematic of KEK-B is shown in Fig. 4.1.

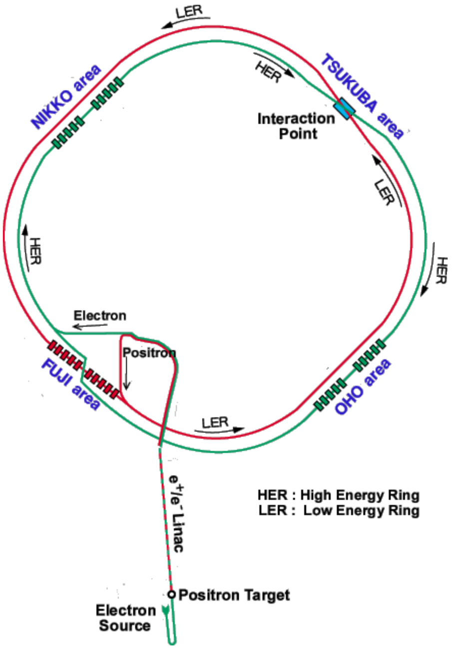


Figure 4.1: The KEK-B accelerator [63].

4.2 The Belle Detector

The Belle detector, alongside with the *BABAR* experiment, led to outstanding observations and discoveries, such as the first observation of CP violation in the B system and the discovery of the neutral and charged tetraquark states $X(3872)$ and $Z(4430)$. The detector is an arrangement of concentric sub-detector layers surrounding the interaction point of KEK-B covering a total of 92 % of the solid angle in the laboratory frame. The sub-detectors are designed for a precise reconstruction of decay vertexes, and charged and neutral particles. A solenoid magnet provides the necessary magnetic field for the distinction of positive and negatively charged particles and the measurement of their momentum. Figure 4.2 shows the configuration and components of the detector. The direction of the HER defines the direction of the z axis in the cylindrical coordinate system (ρ, θ, z) used at Belle, with origin at the IP. The polar angle θ is measured with respect to z. A detailed description of the detector can be found elsewhere [64], here a short overview is presented.

4.2.1 Beam Pipe

The beam pipe surrounding the IP is the first piece of material all particles traverse before reaching the detector. To minimize the impact of the beam-pipe on the trajectories of the particles, it must be as thin as possible and made of a material with low atomic number. It is also important that its diameter is as small as possible since vertexing performance is reduced with increasing distance of the vertex detector from the IP.

The beam pipe was chosen to be constructed of beryllium, consisting of a dual layer cylinder

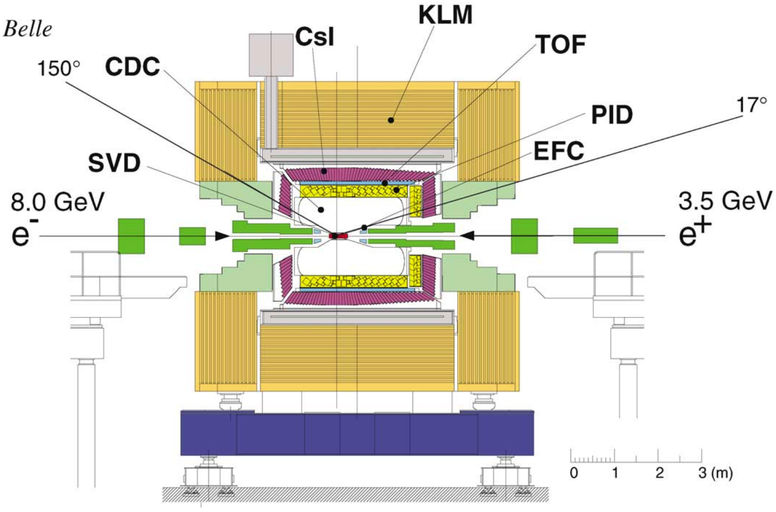


Figure 4.2: The Belle detector and its components [64]. The acronyms and subdetector systems are described in the text.

where the gap between the walls is used as a helium gas channel for cooling. The cylinder walls are 0.5 mm thick, corresponding to 0.3 % of a radiation length. A 20 μm gold foil covers the outer surface to reduce low-energy X-ray background which otherwise might generate background in the vertex detectors, its thickness corresponds to 0.6 % of a radiation length [64]. The first beam pipe had an outer radius of 23.5 mm but was replaced after the first 20% of the data was collected when the inner detector was re-designed and placed closer to the IP. The second beam pipe had an outer radius of 16.25 mm and a gold foil of 10 μm thickness.

4.2.2 Silicon Vertex Detector

Determining the decay vertices of both B mesons is very important, in particular for measuring time dependent quantities. The Silicon Vertex Detector (SVD) is situated directly outside the beam pipe. There were two generations of silicon vertex detectors in Belle which are referred to as SVD1 and SVD2. The first generation, SVD1, covered 86 % of the full solid angle [65], and was made of three layers of double-sided silicon-strip detectors (DSSD). It was used to record the first 140 fb^{-1} of on-resonance data. The upgraded generation installed in 2003 was a four-layer detector [66] with a higher radiation tolerance and better spatial resolution. The SVD2 recorded the last 570 fb^{-1} of $\Upsilon(4S)$ data. There were 246 double-sided silicon sensors in a ladder configuration as shown in Fig. 4.3. The four layers were located at 20, 43.5, 70 and 88 mm from the interaction point, covering the polar angle region $17° \leq \theta \leq 150°$.

A DSSD is a pn-junction with a bias voltage of 75 V applied to the n-side and a grounded p-side. p-implants are interleaved with n strips to separate them. A charged particle passing through the n-bulk silicon liberates electrons from the valence band into the conduction band, creating electron-hole pairs. The electrons and holes drift to their corresponding biased side of the DSSD, making a 2-dimensional hit signal. The p strips were aligned along the beam axis and

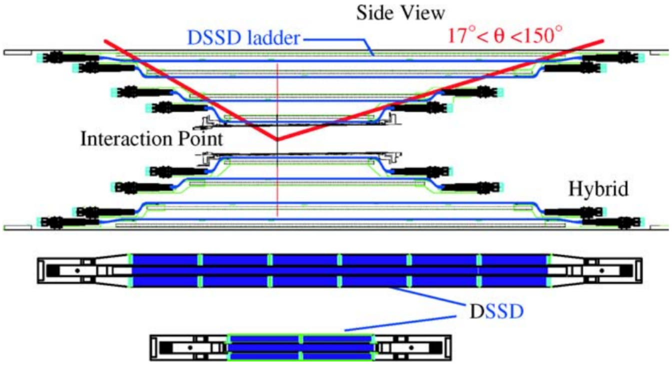


Figure 4.3: Silicon vertex detector, side view of SVD2 [66].

therefore measured the azimuthal angle ϕ. The n strips were aligned perpendicularly to the beam axis and measured z. The readout of the DSSDs was based on CMOS integrated circuits placed outside of the tracking volume.

The performance of SVD1 and SVD2 is summarized in Table 4.1 and Fig. 4.4. Parameters such as signal-to-noise ratio, impact parameter resolution and efficiency are shown. The resolution for the impact parameters dz and $d\rho$ was obtained from cosmic rays measured during the collision data taking.

Parameter	SVD1	SVD2
Signal-to-noise ratio S/N	> 17	> 16
Mean occupancy in layer 1 (%)	~6	~10
Matching efficiency (%)	98	98
Impact parameter resolution for dz ($p = 1\,\mathrm{GeV\,c^{-1}}$, $\theta = 90°$)	61 µm	42 µm
Impact parameter resolution for $d\rho$ ($p = 1\,\mathrm{GeV\,c^{-1}}$, $\theta = 90°$)	57 µm	42 µm

Table 4.1: Performance parameters of SVD1 and SVD2 [67].

4.2.3 Central Drift Chamber

The Central Drift Chamber (CDC) was a cylindrical array of gold-plated tungsten and aluminum wires inside a 50 % : 50 % helium-ethane gas mixture covering the polar angular region $17° \leq \theta \leq 150°$ [68]. When a charged particle traverses a gaseous medium, it loses energy due to interactions with the gas molecules via elastic scattering, excitation and ionization. Together with the SVD, the CDC provided important information for track fitting, particle identification, momentum measurement and vertex reconstruction.

The CDC extended radially from 88 mm to 874 mm, consisting of 50 cylindrical wire layers with a total of 24944 field wires and 8400 readout channels used as sense wires where 2.4 kV bias voltage was applied. The longest wires have a length of 2 400 mm. The structure of the CDC

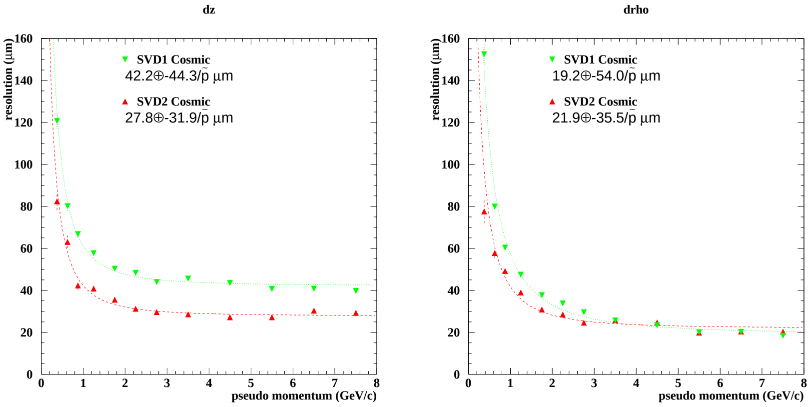


Figure 4.4: Impact parameter resolutions for dz on the left and in $d\rho$ on the right, for the SVD1 and SVD2 [67].

is shown in Fig. 4.5. Approximately half of the wires were in z direction to provide measurement of the transverse momentum p_T with respect to the z-axis, while the other half was slanted by a small angle of 50 mrad. These so-called stereo wires allowed the polar angle of the track to be measured. The longitudinal momentum could be determined from the measured transversal momentum and the polar angle. Schematics of the CDC and the wire configuration can be seen in Figs. 4.5 and 4.6.

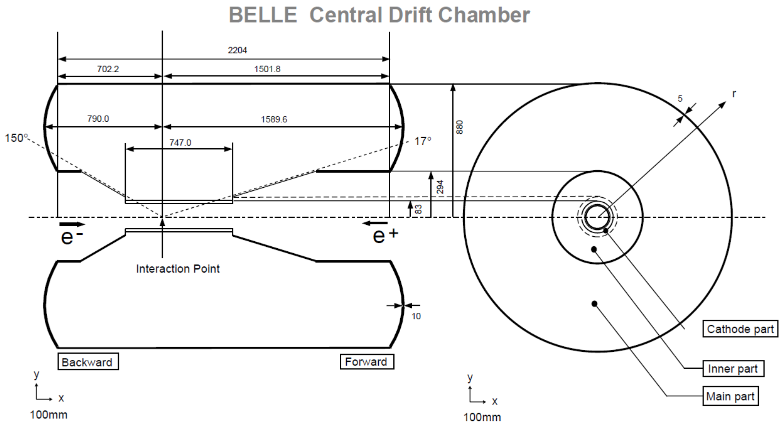


Figure 4.5: Central drift chamber structure in side (left) and frontal (right) projections [64].

The achieved spatial resolution for tracks in the radial direction was $\sigma_{r\phi} = 130\,\mu\text{m}$ and in the z direction $\sigma_z = 200\text{-}1400\ \mu\text{m}$. The distribution of energy loss (dE/dx) as a function of track momentum can be seen in Fig. 4.7. The dE/dx resolution is 6 %. The distributions are clearly different for different particle species, making it an important handle for particle identification. The p_T resolution as a function of p_T was measured using cosmic rays:

$$\sigma(p_T) = (0.201 \pm 0.003)\% p_T \oplus (0.290 \pm 0.006)/\beta\%. \tag{4.1}$$

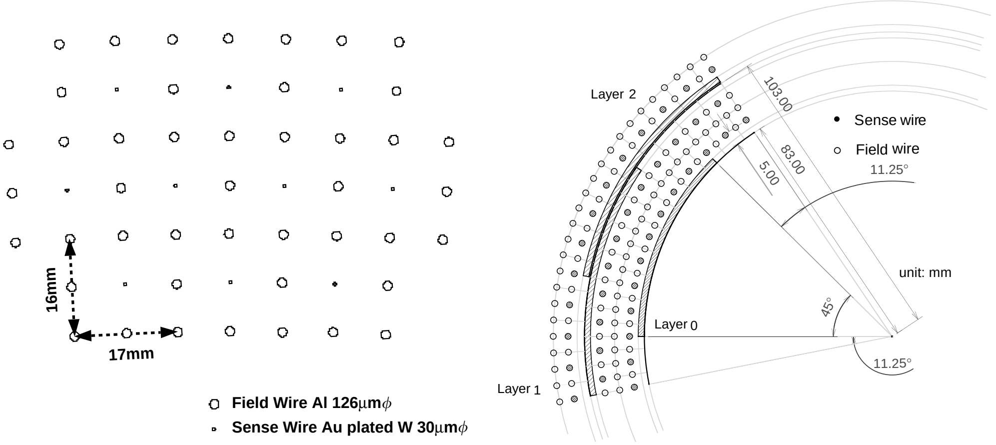


Figure 4.6: Wire configuration in the CDC [64].

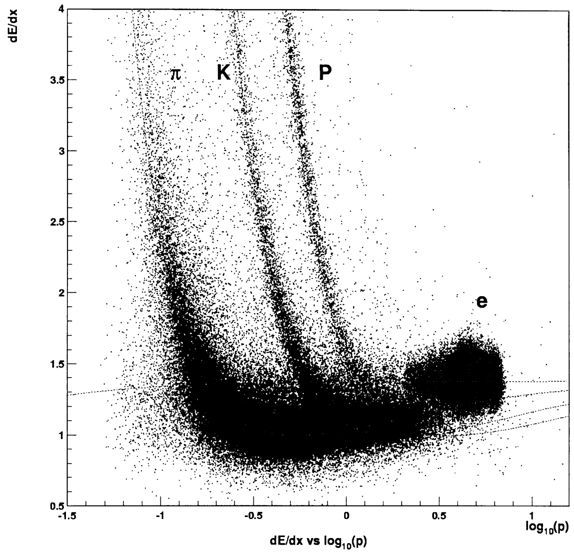


Figure 4.7: dE/dx as a function of track momentum measured in collision data [64].

4.2.4 Aerogel Cerenkov Counter

The silica aerogel Cerenkov counter (ACC) system complemented the capability of other sub-detector systems by offering additional handles to distinguish $\pi^{\pm}$ and $K^{\pm}$, which is crucial to the study of many B decays. When charged particles traverse a transparent material with a velocity greater than the speed of light in the medium, they emit coherent Cerenkov light. The condition for emission relates the speed of the particle $\beta = \frac{v}{c}$ to the refractive index of the material n by:

$$n \geq \frac{1}{\beta} = \frac{1}{\sqrt{1+(m/p)^2}}. \tag{4.2}$$

Given that pions and kaons have different masses, for a certain refractive index of the material and particles of the same momentum, pions emit light while kaons do not.

Belle used a total of $2\,\mathrm{m}^3$ of silica aerogel with a small refractive index. The ACC modules were arranged along the barrel and end-cap regions of the detector with 960 and 228 counters, respectively [69]. The modules point towards the IP and photo-multipliers are attached to them in order to detect the Cerenkov light. Each ACC row in the barrel of the detector region had on average 60 counters which produced between 10 to 20 photoelectrons. Each ACC endcap layer produced on average around 25 to 30 photoelectrons. The K/π separation was possible in the momentum range $1.2 \leq p \leq 3.5\,\mathrm{GeV}$. Additionally to K/π separation, e/K and p/π separation was possible with a significance above 3σ. The schematics of this sub-detector are shown in Fig. 4.8.

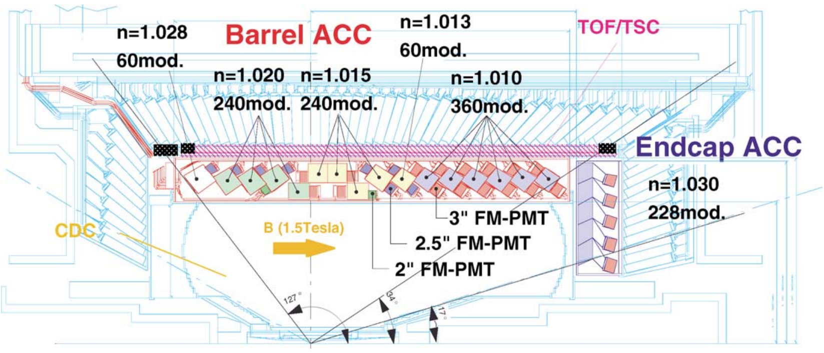


Figure 4.8: The arrangement of ACC modules [64].

4.2.5 Time-of-Flight Counter

The time-of-flight (TOF) detector consisted of plastic scintillators and was used for particle identification and to supply timing information for the trigger system. It measured the arrival time of particles, which was used to calculate their velocity. Its timing resolution was 100 ps for particles with momenta below 1.2 GeV. It additionally provided excellent information for π/p and π/K separation at low momenta. The system was composed of 128 TOF counters and 64 Trigger Scintillation Counters (TSC) which make up a total of 64 modules mounted in the inner part of the calorimeter at a radius of 1.2 m [64]. The trigger information is important to veto off-time events that arise from the high luminosity of the machine and it has been used in this analysis. For a given path length (L), the mass of the detected hadron can be calculated from the time needed to reach the TOF (T):

$$m^2 = \left(\frac{1}{\beta^2} - 1\right) p^2 = \left(\left(\frac{cT}{L}\right)^2 - 1\right) p^2 \tag{4.3}$$

Figure 4.9 shows the distribution of calculated mass for detected tracks according to Eq. (4.3). The peaks correspond to pions, kaons and protons.

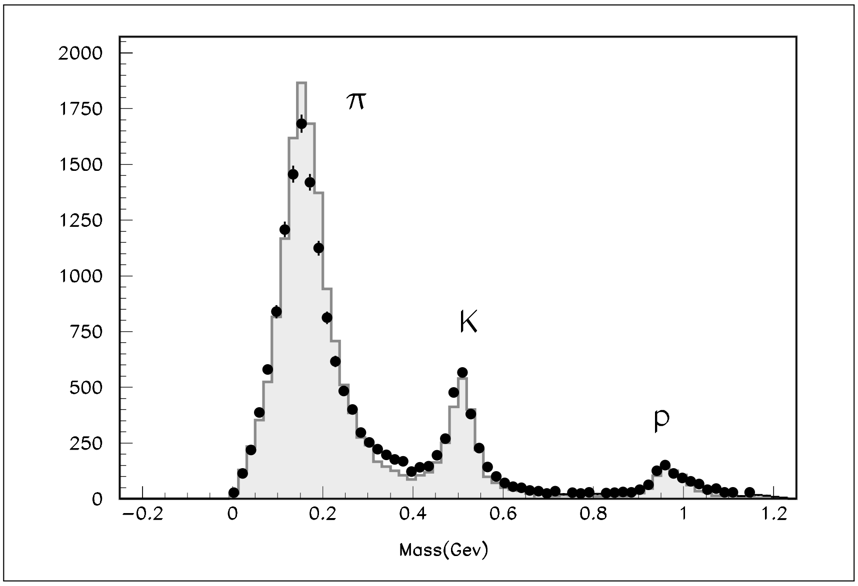


Figure 4.9: Calculated mass for detected tracks using TOF timing information as given by Eq. (4.3) for tracks with $p < 1.2\,\mathrm{GeV\,c^{-1}}$ [64].

4.2.6 Electromagnetic Calorimeter

The electromagnetic calorimeter (ECL) is the most important subdetector for this analysis. Its task was to measure electromagnetic showers from photons and electrons and to provide $e^{\pm}$ and $\pi^{\pm}$ separation. Photons of up to 4 GeV in the laboratory frame can be expected to be produced in decays of interest such as $B \rightarrow X_s\gamma$. At the same time photons originating from $\pi^0 \rightarrow \gamma\gamma$ decays have to be clearly resolved to be able to reconstruct the initial hadron. This translates to the need a of high-resolution detector and fine segmentation in the ECL crystals. The ECL contained 8736 CsI crystals slightly tilted with respect to the direction to the interaction point to avoid photons escaping through the gaps between the crystals. It covered 91 % of the total solid angle. The ECL is subdivided into three regions: forward ($12.4° \leq \theta \leq 31.4°$), barrel ($32.2° \leq \theta \leq 128.7°$) and backward ($130.7° \leq \theta \leq 155.1°$) [70] as sketched in Fig. 4.10.

Photons and electrons entering the ECL crystals develop an electromagnetic showers and deposit their energy. The shower energy is absorbed by the molecules in the crystal and re-emitted in the form of light in a process called scintillation, the scintillation light is in turn detected by photodiodes. The shower energy is estimated from the detected scintillation light. The energy

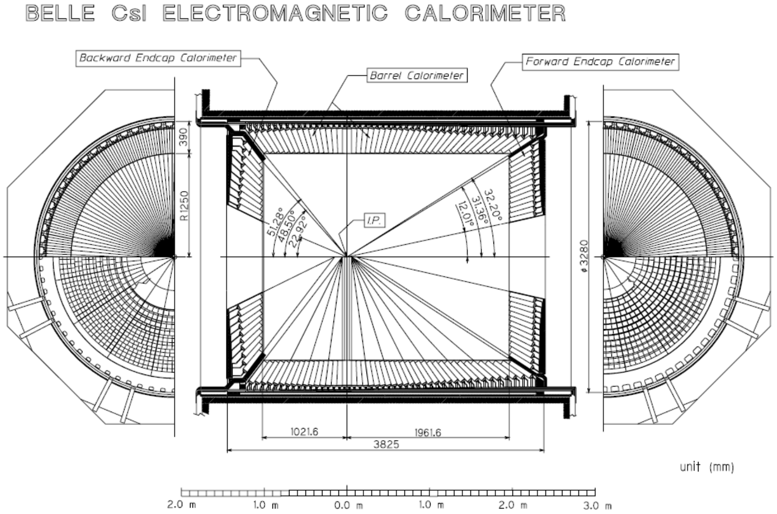


Figure 4.10: Electromagnetic calorimeter [64].

resolution can be affected by electronic readout and thermal noise. The large number of photons produced in the decays and those originating from interaction of electrons and positrons with the detector material can lead with time to a loss of transparency of the crystals, which in turn worsens the energy resolution.

The photon energy resolution of the ECL crystals was measured in two different ways [64]. The first one used photon beams at the ROKK-1M facility in the Budker Institute of Nuclear Physics, produced by the scattering of laser light with a beam of electrons. These Compton photons were detected with a 6×6 array of CsI crystals. The measurement of the spectrum of the scattered electrons allows for the determination of the initial photon energy. The energy can also be calculated as the sum of the energy deposition in each crystal. The comparison from both determinations permits the determination of the resolution. The resolution was found to be of the order $\sigma_E/E \sim 10^{-3}$ in the photon energy range 20 MeV to 5.4 GeV.

The measurement was also performed directly using information extracted from the crystal array, without using the information of the scattered electron, thus providing an independent cross-check. The resolution was determined for the 3×3 and 5×5 cells around the seed cell, performing a fit to the measured energy with a normal-logarithmic function. The resolution for high energy photons (as produced in $b \to s\gamma$ decays), is of the order of 2 % as can be seen in Fig. 4.11. The very precise determination of the photon energy, particularly for higher energies, is of great importance in the study of radiative B decays. It ensures a precise determination of the photon spectrum, allowing for a good comparison between experimental and theoretical results.

4.2.7 Solenoid and K_L^0/μ Detector

A superconducting solenoid surrounded the tracking system and calorimeter providing a 1.5 T magnetic field pointing in the z direction. A strong magnetic field bends charged tracks and allows in this way to measure their momentum and distinguish their charges. The coil consisted of a

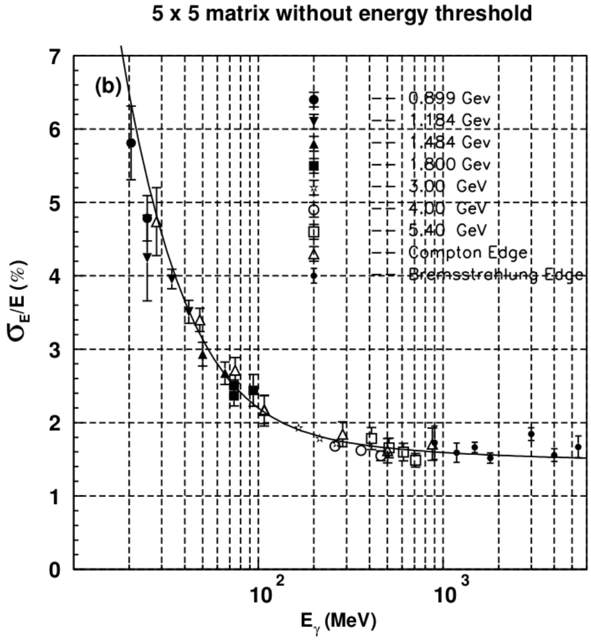


Figure 4.11: Photon energy resolution for the sum of energies in an array of 25 CsI crystals [64].

single layer niobium-titanium-copper alloy embedded in a high-purity aluminum stabilizer and wound around the inner surface of an aluminum support cylinder measuring 3.4 m in diameter and 4.4 m in length. Cooling was provided by circulating liquid helium through a tube on the inner surface of the aluminum cylinder.

Around the solenoid the multilayer iron yoke provided absorbing material for the K_L^0/μ detector (KLM) and mechanical support for the different detector parts (see Fig. 4.12). The KLM was an arrangement of alternating charged particle detectors (glass-electrode resistive plate counters (RPC)) and iron layers. Neutral kaons interact more strongly with the material in comparison to muons, which are therefore expected to travel further. The detection of charged particles is performed RPCs, which enclose a gas gap and detect the ionization in it produced by charged particles. Neutral K_L mesons are identified when a cluster is found in the KLM without an associated track in the CDC [64]. The spatial resolution for K_L was $\Delta\theta = \Delta\phi = 30$ mrad.

4.2.8 Triggering and Data Acquisition

A large fraction of the events occurring during e^-e^+ collisions are not from interesting physics processes, but originating in beam-gas or beam-beam interactions, cosmic rays, synchrotron radiation and others. Alongside $B\overline{B}$ production, events such as Bhabha scattering, μ-pair production and two-photon processes are useful for luminosity measurements and calibration. The cross-sections for the most common processes are summarized in Table 4.2. A multilevel triggering system was used to record the data, it had as efficiency of 99 % for $B\overline{B}$ events and consisted of a level 1 hardware trigger, a real-time level 3 software trigger and an off-line level 4 software trigger.

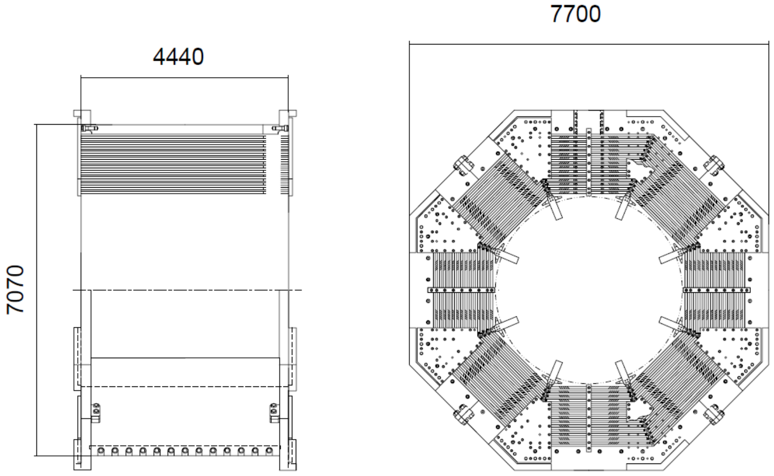


Figure 4.12: Barrel part of the iron yoke, in side (left) and frontal (right) projection [64].

Trigger signals were collected by a Global Decision Logic (GDL) and a level 1 decision was taken within 2.2 µs of the beam crossing.

Process	Cross section (nb)
$B\overline{B}$	1.1
$q\bar{q}$ $(q = u, d, s, c)$	3.3
$\mu^+\mu^-\ +\ \tau^+\tau^-$	1.6
Bhabha (and radiative)	37.8
Two-photon	11.1

Table 4.2: Cross section for several processes running at the $\Upsilon(4S)$ center of mass energy [71].

During normal operation ($\mathcal{L} = 10^{34}\,\text{cm}^{-2}\,\text{s}^{-1}$) the total event rate was around 200 Hz but the system was designed to sustain a maximum trigger rate of 500 Hz. CDC and TOF were used to yield trigger signals for charged particles. The ECL provided triggers based on total energy deposit and number of activated crystal. The KLM gave additional information on muons. The ECL triggers were also used for tagging two-photon events as well as Bhabha events.

After a hardware trigger signal, the Data Acquisition system (DAQ) collected data from all sub-detectors and sent them to the data storage system. The signal from all sub-detectors, except for SVD, were digitized by a unified readout system based on charge-to-time (Q-to-T) conversion and transfered to the on-line event building farm via VME crate. The data from the SVD were processed by a PC-based readout system and sent to the event building farm directly via the network. The online computer farm formats the data into an offline event format and proceeds to level 3 trigger to reduce background. A fast tracking program is used for the level 3 trigger.

CHAPTER 5

Event selection

In this section the used data and Monte Carlo (MC) samples are described. Furthermore, the selection of $B \to X_{s+d}\gamma$ events and the steps taken to reduce the different background components are described. A section is dedicated to the suppression of continuum background, which is the largest background source on this analysis. The variables used in the selection are described and motivated.

5.1 Data and Monte Carlo samples

The Belle data set consists of 710 fb^{-1} of integrated luminosity taken at the $\Upsilon(4S)$ resonance ($\sqrt{s} = 10.58\,\mathrm{GeV}$), 121 fb^{-1} of data at the $\Upsilon(5S)$ resonance, 90 fb^{-1} of integrated luminosity of continuum data taken below the open $B\overline{B}$ production threshold ($\sqrt{s} = 10.52\,\mathrm{GeV}$) and smaller data sets collected at the $\Upsilon(1S)$, $\Upsilon(2S)$ and $\Upsilon(3S)$ resonances. The on-resonance data sample contains $(771.6 \pm 10.6) \times 10^6$ $B\overline{B}$ pairs, determined with a precision of 1.4 %. Monte Carlo (MC) simulation samples are produced independently for continuum events, B^+B^- pairs, $B^0\overline{B}^0$ pairs, $B \to X_s\gamma$, rare B decays (such as other $b \to s$ transitions) and $b \to u\ell\nu$. A summary of all data and MC samples used in this study is given below in Table 5.1. The MC samples are generated in sets corresponding to the total integrated luminosity of the on-resonance data, with a full set denoted as a “stream”.

The decays chains of B mesons are simulated using the EvtGen package [72], while the detector response is simulated using GEANT 3 [64]. Belle data and MC are split into subsets called “experiments”, with the two main data taking periods marked by the upgrade of the SVD. The first one consists of 11 experiments and corresponds to 141 fb^{-1} of on-resonance data, and the second consists of 19 experiments and 569 fb^{-1} of on-resonance data. The experiment numbers do not follow a strict sequence (e.g. 1, 2, 3, etc.) since beam and detector calibration periods are not counted in. The experiment numbers for data used here begin with 7 and end at 73, on-resonance data was recorded up to experiment 65, beyond that only off-resonance data for $\Upsilon(4S)$ analyses was recorded. During most experiments, both on-resonance and continuum data were taken, however after experiment number 65, only off-resonance data were recorded. Data from different experiments are handled differently due to slight differences in detector conditions.

Sample	Size	$\mathcal{L}_{\text{int}}$
On-resonance data		$710\,\text{fb}^{-1}$
Of-resonance data		$90\,\text{fb}^{-1}$
Continuum: u,d,s	1 stream	$1\times 693\,\text{fb}^{-1}$
Continuum: c	1 stream	$1\times 693\,\text{fb}^{-1}$
Generic B^+B^-	10 streams	$10\times 693\,\text{fb}^{-1}$
Generic $B^0\overline{B^0}$	10 streams	$10\times 693\,\text{fb}^{-1}$
Rare $B\overline{B}$	50 streams	$50\times 710\,\text{fb}^{-1}$
$B\to X_u\ell\nu$	20 streams	$20\times 710\,\text{fb}^{-1}$
$B\to X_s\gamma$	$\sim$ 30 streams	$\sim 30\times 710\,\text{fb}^{-1}$

Table 5.1: Summary of Monte-Carlo and data samples used in the analysis. A stream corresponds to the expected number of such events in the full $\Upsilon(4S)$ data sample.

For this analysis the most relevant difference is the availability of a time stamp to veto off-time events in the second data-taking period.

5.1.1 Corrections to the off-resonance sample

Continuum background is the largest background component in this analysis. MC samples for continuum events perform rather poorly so it is necessary to rely on off-resonance data to subtract them. Off-resonance data can describe the continuum contribution during on-resonance data-taking provided it is properly corrected. However, using off-resonance data poses two problems in this analysis. The first and most important one is that the off-resonance sample is about 8 times smaller than the on-resonance sample. Second of all, off-resonance events were produced at a lower center-of-mass energy, 60 MeV below the $\Upsilon(4S)$. This translates in a lower average energy of the particles in the event as well as a lower average track and cluster multiplicity. The latter effects are small and can be easily corrected using studies in MC samples. These corrections are important to be sure that the off-resonance events populate the correct on-resonance energy and momentum phase-space. An appropriate correction means also a correct calculation of the kinematic and event shape variables used in the continuum suppression. In [52] the correction factors for an average energy (F_E) and an average multiplicity (F_N) have been calculated. Since the production of MC samples is identical, the factors previously determined were taken for this analysis; they take the following values:

$$F_E = 1.0036 \pm 0.0036 \tag{5.1}$$

$$F_N = 1.0009 \pm 0.0009. \tag{5.2}$$

These values are determined using MC samples for $e^-e^+ \to q\bar{q}$ at $\sqrt{s} = 10.52\,\text{GeV}$ and $\sqrt{s} = 10.58\,\text{GeV}$. Given that [52] does not provide any error estimation on these correction factors, and given the fact that they are very small assigning a 100 % uncertainty to the corrections is feasible and a conservative guess for the uncertainty. As will be seen later, the uncertainty has an almost negligible impact on the measurement of the $B \to X_s\gamma$ branching fraction.

One additional correction factor takes into account the different integrated luminosities and cross-sections for the off-resonance data and is given by:

$$\alpha'_{\text{off}} = \frac{\int \mathcal{L}^{\text{on}}}{\int \mathcal{L}^{\text{off}}} \frac{s^{\text{off}}}{s^{\text{on}}} = 7.7657.$$

The official Belle luminosity website [73] advices to assign an uncertainty to the luminosities ratio between 0.3 % and 0.5 %. In [52] an alternative uncertainty on the ratio was explored studying $e^-e^+ \to e^-e^+$ and $e^-e^+ \to \mu^+\mu^-$ reactions which resulted in a 0.32 % uncertainty. Given that these determinations are consistent with each other, an uncertainty on α'_{off} of 0.5 % is chosen.

The preselection and selection efficiencies, described in Section 5.2 and Section 7.2, also differ for different CM energies. Different data preselection are available at Belle and are common to several analyses looking for similar signatures. One set of such preselection requirements are referred to as a "skim". Most skims build upon the HadronB skim [71], which consists of requirements to filter low-multiplicity events which are not likely to be $B\overline{B}$ or $q\overline{q}$ events. HadronB is the only skim used in this analysis. The ratios of selection and skim efficiencies for on- and off-resonance continuum MC are taken into account as an additional correction factor:

$$\alpha_{\text{eff}} = \frac{\varepsilon^{\text{on}}_{\text{skim}}}{\varepsilon^{\text{off}}_{\text{skim}}} \frac{\varepsilon^{\text{on}}_{\text{sel}}}{\varepsilon^{\text{off}}_{\text{sel}}}.$$

Here the efficiencies ε correspond to the selection efficiency of the on- or off-resonance continuum MC, and of the used skim or selection procedure. The ratio $\frac{\varepsilon^{\text{on}}_{skim}}{\varepsilon^{\text{off}}_{skim}}$ for the HadronB skim was very precisely measured in [52]. Since the measurement is purely done with MC and neither the generator nor the skim have changed, the same value of 0.9986 ± 0.0001 is taken. For the determination of $\frac{\varepsilon^{\text{on}}_{\text{sel}}}{\varepsilon^{\text{off}}_{\text{sel}}}$, the selection procedure, as described in the following sections, is applied to on- and off-resonance continuum MC and the ratio of efficiencies is calculated. It yields $\frac{\varepsilon^{\text{on}}_{\text{sel}}}{\varepsilon^{\text{off}}_{\text{sel}}} = 0.9680 \pm 0.0248$.

Taking all these factors into account, the correction factor for the off-resonance data normalization as given in Eq. (5.3) is obtained. The uncertainties on α'_{off} and α_{eff} are propagated to the uncertainty on α_{off}. This uncertainty will play an important role for the studies performed on this thesis. It receives contributions proportional to the size of the data samples, meaning that it will be reduced if larger data sets become available.

$$\alpha_{\text{off}} = \alpha'_{\text{off}} \times \alpha_{\text{eff}} = 7.509 \pm 0.196. \tag{5.3}$$

5.1.2 Scaling of generic Monte Carlo samples

The generic $B\overline{B}$ MC luminosity does not exactly match the luminosity of the on-resonance data, therefore it must be scaled. The scaling factors differ for each experiment, and are given in Table 5.2.

Experiment	MC $\mathcal{L}_{\text{int}}$	Scaling factor	Experiment	MC $\mathcal{L}_{\text{int}}$	Scaling factor
7	5.8	1.022	35	16.9	0.989
9	4.1	1.078	37	60.9	1.000
11	7.7	1.056	39	42.4	0.972
13	10.7	1.002	41	57.3	1.025
15	12.7	0.987	43	54.6	1.029
17	9.2	1.212	45	12.8	1.012
19	24.6	1.014	47	36.7	1.013
21	4.3	1.014	49	26.7	1.013
23	6.3	0.992	51	38.8	1.012
25	25.5	1.058	55	71.3	1.012
27	25.4	1.003	61	34.2	0.998
31	17.0	1.041	63	32.2	1.022
33	17.3	1.007	65	37.2	1.016

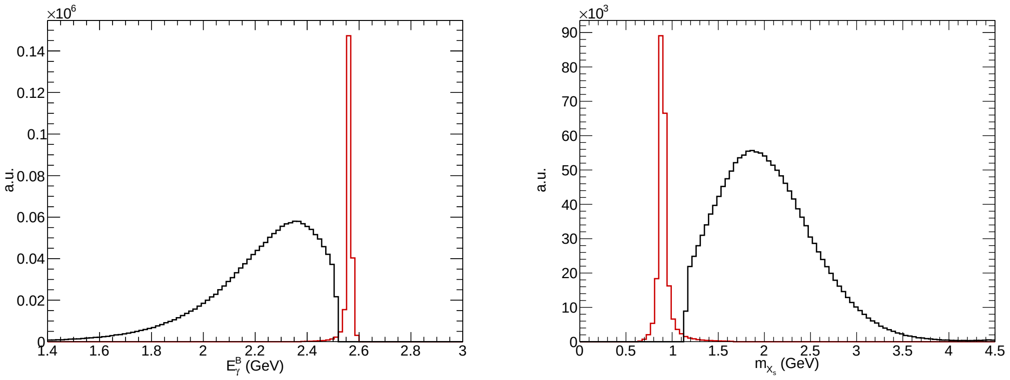
Table 5.2: Recorded and generated on-resonance luminosity with the corresponding experiment numbers. For the numbering of experiments refer to the text on Section 5.1.

5.1.3 $B \to X_s\gamma$ Monte Carlo sample

The signal MC is generated as the sum of an exclusive and an inclusive component. The decay $B \to X_s\gamma$ contains several resonant two-body final states, the most important and most precisely measured is the decay $B \to K^*(892)\gamma$. It has been measured to be $\sim$ 12 % of the full $B \to X_s\gamma$ decay rate, the K^* has a mass of 892 MeV. The K^* mass and photon energy spectra correspond to narrow peaks. The peak broadening is caused by the Fermi motion of the b quark inside the B meson and the interaction of the b quark with quarks and gluons inside the meson. The inclusive $B \to X_s\gamma$ component is generated for hadronic masses above 1.1 GeV. The X_s is produced as a spin 1 particle, whose hadronization is simulated by JETSET [74]. The inclusive and exclusive components are mixed in proportions consistent with the current world average of 88:12 [27]. The hadronic mass and photon energy spectra can be seen in Fig. 5.1.

Initially 3×10^7 inclusive events are generated, according to a relativistic Breit-Wigner distribution for m_{X_s}. They are reweighted to follow the theoretical description for the spectrum. The chosen theory model is the Kagan-Neubert (KN) model since it is the easiest to implement [31]. It has two input parameters m_b and μ_π^2, the first is the mass of the b quark and the later can be understood as the average kinetic energy of the b quark inside the hadron (see Section 2.2.1). The used input values are $m_b = 4.574\,\text{GeV}$ and $\mu_\pi^2 = 0.459\,\text{GeV}^2$ [75]. The events are not strictly reweighted, but only a subset of them is chosen such that the theoretical Kagan-Neubert distribution is reproduced, the remaining events are dismissed.

In the reweighting step, around 53.1 % of the events are retained. After this, $B \to K^*\gamma$ events are generated in the experimentally determined proportion and mixed with the inclusive component. At this stage, the $B \to X_s\gamma$ signal MC is composed of 15.9×10^6 $B \to X_s\gamma$ events in total, which corresponds to roughly 30 times the expected number in data. The expected number of events to be found in data using the 2014 HFAG average [27] for the $B \to X_s\gamma$ branching



(a) (b)

Figure 5.1: (a) Photon energy spectrum in the B rest frame and (b) hadronic Xs mass in $B \to X_s\gamma$ decays showing the $K^*(892)$ component in red and the inclusive X_s component in black.

fraction is:

$$\begin{aligned} N(b \to s\gamma) &= N(B\overline{B}) \times 2 \times \mathcal{B}_{s\gamma} \\ &= 772 \times 10^6 \times 2 \times 3.43 \times 10^{-4} \\ &= 5.29 \times 10^5. \end{aligned} \tag{5.4}$$

5.2 Preselection

Several requirements are imposed for the recorded events in order to ensure that they are consistent with the decay of a $B\overline{B}$ pair. As mentioned in Table 4.2, several processes have cross-sections larger than that of $B\overline{B}$ events, so one must try to preselect hadronic events, as opposed to e.g. Bhabha, $\tau^+\tau^-$, dimuon and two-photon events. This preselection is fully described in [71] and is included inside the HadronA and HadronB skims of the Belle software, and is common to all B analyses. $B\overline{B}$ events have a larger track and cluster multiplicity than other processes, and the energy of the tracks and clusters is thus relatively low compared to processes with very few particles. The main requirements include: at least three tracks, visible energy from good tracks and clusters above 20 % of the total CM energy, total energy in the ECL less than 80 % and more than 18 % of the total CM energy. The HadronA and HadronB requirements are 99.1 % efficient for $B\overline{B}$ events. After HadronB, the light quark background is still roughly 3 times larger than $B\overline{B}$ events, while contributions from other processes such as $\tau^+\tau^-$ and Bhabha scattering are negligible. The cross-sections and efficiencies for various processes after this preselection are taken from [71], and shown in Table 5.3.

Process	$B\overline{B}$	$q\overline{q}$	$\tau^+\tau^-$	QED	$\gamma\gamma$	Beam gas
σ(nb)	1.09	2.62	0.05	0.001	0.04	0.11
$\varepsilon_{\mathrm{skim}}$	0.991	0.795	0.049	0.00002	0.004	$0.09\varepsilon_A$

Table 5.3: Cross sections and efficiencies for various processes after the HadronB skim.

5.3 Selection of objects

In this inclusive analysis the event selection is based on finding a high energy photon from the signal B and a high momentum lepton from the other B, which is used for continuum suppression and flavor tagging. Additionally some selection requirements are imposed for the remaining particles in the event.

5.3.1 Selection of tracks and neutral clusters

Low-momentum charged particles can spiral inside the CDC due to the strong magnetic field. This could cause the track finding algorithm to find two or more tracks for a single particle. Duplicated tracks can be recognized if they have a small momentum difference and are spatially close to each other. Tracks with a momentum difference smaller than 0.1 GeV are rejected if the angle between them is $\vartheta_{12} \leq 15°(\geq 165°)$ for same (opposite) charges.

It is also important to reject tracks that pass far away from the IP as they likely do not belong to the $B\overline{B}$ decay, or are tracks made up of a random combination of hits in the tracking system. This is achieved by imposing requirements in the impact parameters dr and dz, which measure the distance of closest approach to the IP in the radial and z directions. These requirements differ for different transverse momentum regions and are summarized in Table 5.4.

p_T	dr	dz
≥ 0.50 GeV	≤ 10 cm	≤ 20 cm
≤ 0.50 GeV	≤ 15 cm	≤ 50 cm
≤ 0.25 GeV	≤ 20 cm	≤ 100 cm

Table 5.4: Track impact parameter requirements.

A large number of photons are produced due to interactions of charged particles with the detector material, wires and the supporting structures and from beam background. Photons from beam background are most likely found in very forward or backward directions, being produced in the direction of the beams. For ECL clusters a polar-angle-dependent energy threshold is applied, the minimum energies are 100 MeV, 50 MeV and 150 MeV for the forward, barrel and backward regions of ECL, respectively, as defined in Section 4.2.6.

Several event shape variables are constructed using all of the particles in the events. An additional requirement on the energy of the particles ensures that only particles which are likely to come from a B decay are used. Since in average half of the initial beam energy is carried by each B meson, a particle coming from a two-body decay of a B meson has a maximum energy that is roughly $1/4$ of the total energy of the beams. If one allows for an additional 0.5 GeV that can arise from resolution effects, the maximum possible energy for a decay product of a B in the CM frame is 3.2 GeV.

5.3.2 Signal photon

The photon energy distribution for $B \to X_s\gamma$ decays is expected to peak around 2.3 GeV in the B rest frame and extends to about 2.6 GeV. Theoretical calculations are done for a photon with energy greater than 1.6 GeV. This analysis is performed in the CM frame, the additional Doppler broadening of the spectrum due to the choice of frame would cause the endpoint of the CM spectrum to be around 2.8 GeV and would slightly shift the peak towards lower energies. For these reasons high energy photon candidates with CM energies $1.4\,\text{GeV} \leq E^*_\gamma \leq 4.0\,\text{GeV}$ are required. The signal region where the $\mathcal{A}^{\text{CP}}_{(s+d)\gamma}$ and $\mathcal{B}_{s\gamma}$ observables are measured is defined as $1.7\,\text{GeV} \leq E^*_\gamma \leq 2.8\,\text{GeV}$. Photon candidates in the sidebands above and below the signal region are used to calibrate continuum and $B\overline{B}$ background.

Photons from beam background processes are most probably located in the ECL forward and backward end-caps. The signal photon is thus required to be found exclusively in the ECL barrel: $32.2° \leq \theta^{\text{lab}}_\gamma \leq 128.7°$. Finally, a requirement on the transverse shower shape in the ECL (E_9/E_{25}) is applied. Photons are reconstructed from isolated calorimeter clusters in the calorimeter, the clusters are built from 5×5 crystals found around a seed crystal, where the highest energy deposition is found. The variable E_9/E_{25} is the ratio of the energy deposited in the central 3×3 crystals to that in the central 5×5 crystals and is required to be at least 90 %. With this selection it is possible to suppress hadronic clusters from those of photons or electrons.

5.3.3 Tag lepton

The lepton ($e^\pm$ or $\mu^\pm$) selected on the tag side must be consistent with coming from a semileptonic B decay ($B \to X\ell\nu$). In semileptonic decays, the lepton charge is directly correlated to its mother B flavor. Semileptonic decays have a kinematic lepton momentum endpoint at around 2.3 GeV, corresponding to roughly half of the mass of the B meson. Leptons that do not directly come from a semileptonic decay likely have lower momenta: lepton tracks produced in the decay of a B daughter (“secondary leptons”), or hadrons (K/π) misidentified as a lepton (“lepton fakes”), leptons from semileptonic decays are called “prompt leptons”. Similarly, a large number of leptons from continuum processes have low momenta. Leptons from continuum process could also have a very large momenta, since the light quarks are produced with high velocities, unlike B mesons produced almost at rest in the CM frame. The selected CM lepton momentum range for this study is $1.10\,\text{GeV} \leq p^*_\ell \leq 2.25\,\text{GeV}$, it keeps a large fraction of prompt leptons ($\sim 80\,\%$), while greatly reducing continuum, secondaries and fakes as can be observed in Fig. 5.2. Further requirements are applied to the lepton track: for the impact parameters $dr \leq 2\,\text{cm}$ and $dz \leq 4\,\text{cm}$, and at least one hit on the SVD are required.

The identification of electrons at Belle is based on the shape of electromagnetic showers of electron clusters which is different to the shape of hadronic clusters, and properties measured in the CDC and ACC (dE/dx and momentum) [76]. A likelihood ratio is defined using five discriminants: matching between the position of the ECL cluster and the extrapolated track, ratio between the energy measured in the ECL and momentum measured in CDC, transverse shower

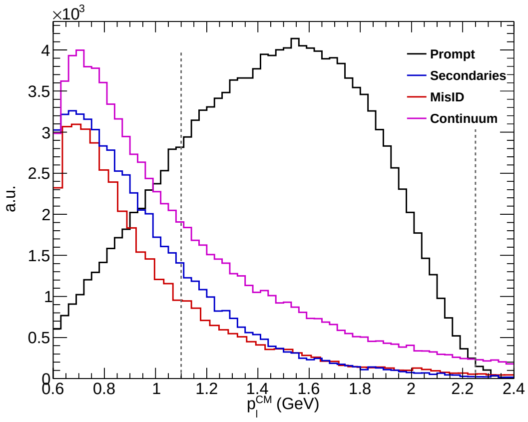


Figure 5.2: Components of the lepton momentum spectrum in the CM frame. An arbitrary normalization is used to compare the shape of each component.

shape in the ECL (E_9/E_{25}), dE/dx in CDC, light yield in ACC. The likelihood ratio (eID) peaks at 1 for electrons, 0 for hadrons and it is exactly 0.5 when there is an identification ambiguity. The requirement in this analysis is $e\text{ID} > 0.90$.

Muon candidates are required to reach the KLM system, therefore they need a momentum of at least 0.6 GeV. The likelihood ratio for muon identification uses the distance difference between the KLM hit and the track extrapolation, and the goodness of fit of all hits in the tracking system associated with this track [77]. With this information a normalized likelihood is calculated with the probabilities of the track being a muon, kaon or pion:

$$\mathcal{L}_\mu = \frac{p_\mu}{p_\mu + p_K + p_\pi} \tag{5.5}$$

The likelihood ratio is also called μID, the value demanded for this analysis is $\mu\text{ID} > 0.90$.

Pions and kaons can be misidentified as a lepton. This effect is small for e but it can be important for μ, particularly for K. The tag lepton requirements are summarized in Table 5.5.

Variable	Cut
Lepton momentum	$1.10\,\text{GeV} \leq p_\ell^* \leq 2.25\,\text{GeV}$
Impact parameter dr	$dr \leq 2\,\text{cm}$
Impact parameter dz	$dz \leq 4\,\text{cm}$
SVD hits	$N_{\text{SVD}} \geq 1$
Electron probability	$e\text{ID} > 0.90$
Electron polar angle	$18° \leq \theta_e^{\text{lab}} \leq 150°$
Muon probability	$\mu\text{ID} > 0.90$
Muon polar angle	$25° \leq \theta_\mu^{\text{lab}} \leq 145°$

Table 5.5: Selection requirements for the tag lepton.

5.3.4 Veto of π^0 and η background

The main sources of $B\overline{B}$ background photons are the decays of the light mesons π^0, $\eta \to \gamma\gamma$ decays. This background is at least two orders of magnitude larger than the signal after the preselection. The decay of a light meson that can produce photons energetic enough to pass the selection criteria must be very asymmetric, meaning that the second photon has a much lower energy, between 30 MeV and 400 MeV as has been found in [78]. The energy of the second photon is often below the threshold chosen for used clusters, therefore it is often not possible to find it and reconstruct the mother meson.

In order to veto some of these events, the prompt photon is combined with all other photons in the event and the combination is checked to see if the two photons come from the same light meson. The vetoing tool used in radiative analyses at Belle uses the reconstructed di-photon mass and the energy of the soft photon as discriminating variables, and separately considers three ECL regions [78]. The veto tool returns the probability that the photon pair originated from the same π^0 or η, i.e. a number between 0 and 1. It clearly distinguishes a true π^0 or η from random combinations, as can be seen in Fig. 5.3. Photon pair combinations that correspond to decays other than π^0 or η decays are accumulated close to zero. The performance of the veto has been tested in data, using a partial reconstruction of D^* mesons in the decay chain $D^* \to D^0(K^-\pi^+\pi^0)\pi^+$ .The veto efficiency was found to be compatible between data and MC [78].

The criteria applied to this analysis are prob(π^0) ≤ 0.10 and prob(η) ≤ 0.20. The π^0 veto is found to be 90 % efficient for $b \to s\gamma$ signal while retaining only 27 % of photons from pions. The η veto is 97 % efficient for signal and retains 58 % of the η background. These values are based on the MC prediction.

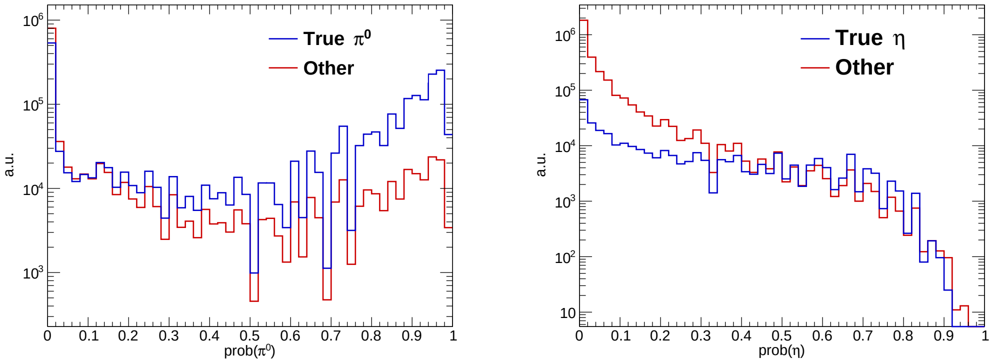


Figure 5.3: Veto for π^0 (left) and η (right), in the generic $B\overline{B}$ MC sample.

5.3.5 Veto for off-time events

Due to the high interaction rate at the Belle experiment, it is likely that the ECL crystals show a signal that belongs to a previous collision. Such occurrences are called “off-time events”. The off-time background consists mainly of beam background, Bhabha or $e^-e^+ \to \gamma\gamma$ events. A time stamp makes it possible to deal with these overlapping QED interactions from previous collisions.

It is constructed with the information of the ECL trigger cells (TC) and is $\sim 100\,\%$ efficient for clusters corresponding to the *current* event [79, 80].

The TC information was initially only available for data taken after the tracking system upgrade (experiment 31 and onwards). A recovery of this information from raw data was later attempted for the SVD1 data sample but part of it could not be reprocessed, resulting in a significant loss of data [81]. Due to this problem it was chosen not to use this variable for the early reprocessed data but to apply the veto only for the SVD2 sample corresponding to 458 fb^{-1}, which is 65 % of the total integrated luminosity. Events with a time stamp between 9 000 ns and 11 000 ns are selected. This cut is applied for on- and off-resonance data. The distribution of this variable is shown in Fig. 5.4. Off-time events are not included in the MC simulation.

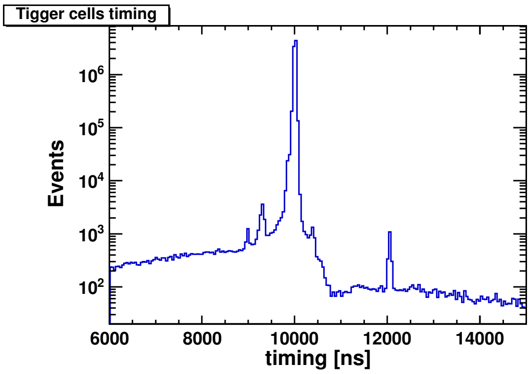


Figure 5.4: Timing information from the calorimeter trigger cells. The time window where the collisions are expected to occur is 9000 ≤ t ≤ 11000.

5.4 Suppression of continuum background

The composition of the photon energy spectrum at this stage of the selection can be seen in Fig. 5.5, with the shape of the $B \to X_s\gamma$ spectrum scaled up by a factor of 50. It is clear that the continuum background poses a major challenge, but topological differences between continuum and $B\overline{B}$ events can be exploited. This can be understood since the heavy $B\overline{B}$ pairs are produced almost at rest in the CM frame, thus their decay products are emitted isotropically. On the contrary, light quark pairs have a large initial momentum since less energy is needed to generate the light quark masses, resulting in more directional decays (“jet-like” event topology). An appropriate choice of discriminating variables makes it possible to suppress continuum. Variables that give the best discrimination while not introducing systematic biases to the photon energy spectrum must be found. Machine learning techniques are used to make optimal use of their separation power.

The chosen variables can be categorized as: cluster isolation, event shape variables and kinematic variables. Those with high separation power, low correlation to the photon energy and good modeling in the MC are chosen.

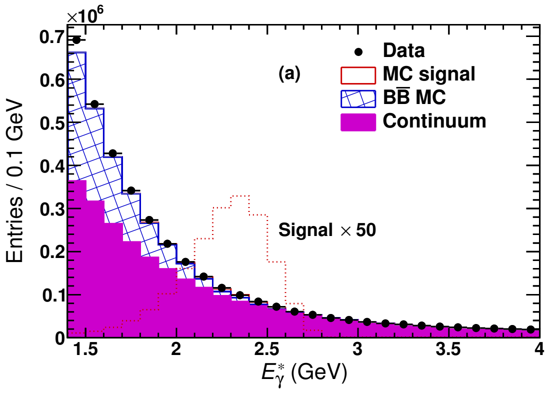


Figure 5.5: Photon energy spectrum before continuum suppression. The $B \to X_s\gamma$ contribution is scaled for comparison.

5.4.1 Kinematic variables

The tag lepton requirement from a semileptonic decay means at least one neutrino is present the event. Neutrinos cannot be detected and result in missing energy and momentum in the reconstructed event. The missing energy is quantified using the missing mass squared M^2_{miss}. The well defined e^-e^+ initial-state four-momentum (P_{o}), makes it easy to infer the missing energy and momentum in the event by subtracting the four-momenta of all visible particles (P_{vis}):

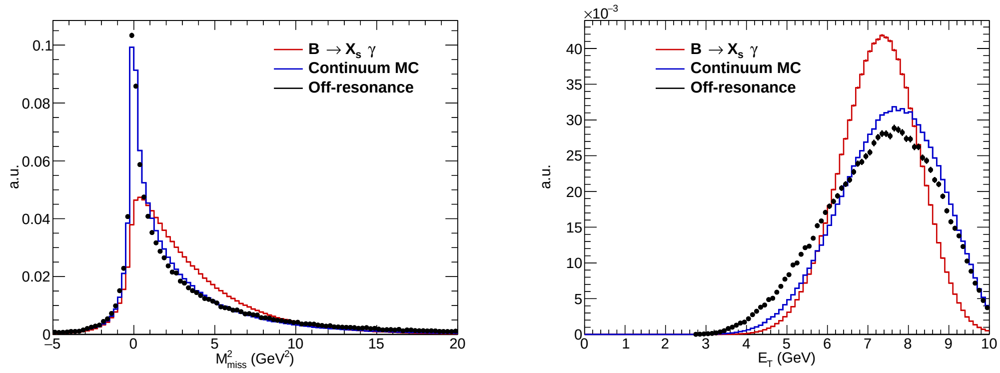
$$\begin{aligned}
P_{\text{o}} &= (E_{\text{o}}, \vec{P}_{\text{o}}) = (E_{\text{LER}} + E_{\text{HER}}, E_{\text{HER}} \sin\psi, 0, E_{\text{HER}} \cos\psi - E_{\text{LER}}) \\
P_{\text{vis}} &= \sum_{i}^{\text{all particles}} P_i \\
P_{\text{miss}} &= P_{\text{o}} - P_{\text{vis}} \\
M^2_{\text{miss}} &= P^2_{\text{miss}}.
\end{aligned} \tag{5.6}$$

Here E_{LER} and E_{HER} are the energies of the low- and high-energy beam particles (the positrons and electrons, with an energies 3.5 GeV and 8 GeV). The angle ψ represents the crossing angle between the two beams at the interaction point and has a value of 22 mrad at Belle.

The highly directional continuum events deposit on average a larger fraction of their energy in the direction of the beams and not as much in the direction transversal to it. This effect is quantified using the transverse energy (E_{T}). The momentum P_{vis} is decomposed into components that are parallel and transverse to the beam pipe as:

$$\begin{aligned}
P_{\text{vis}} &= \left(E_{\text{vis}}, p_x, p_y, p_z\right) \\
E_{\text{T}} &= \sqrt{p_x^2 + p_y^2}.
\end{aligned} \tag{5.7}$$

The distribution of these two variables is shown in Fig. 5.6.



(a)

(b)

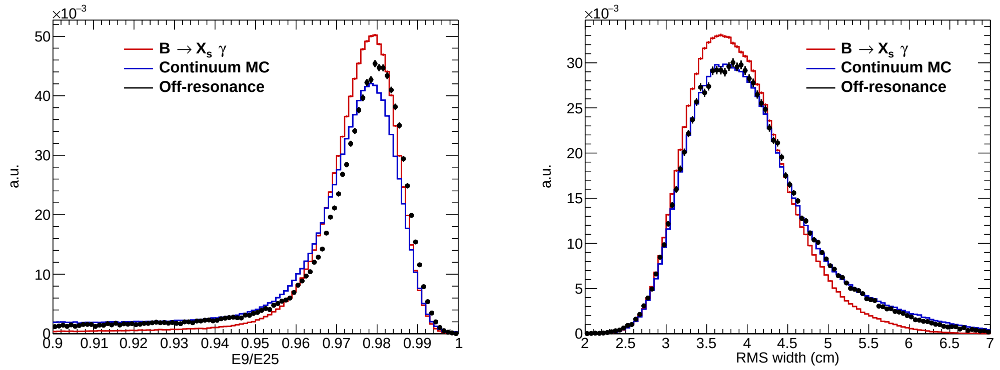
Figure 5.6: Kinematic variables before continuum suppression: (a) missing mass squared and (b) transverse energy.

5.4.2 Calorimeter variables

Photons typically deposit all of their energy in the ECL crystals. Isolated and well reconstructed photons are sought after, so clusters whose energy is regularly distributed around the seed cell are selected. The two variables that are used for this purpose are shown in Fig. 5.7 and are defined as:

- E_9/E_{25}: the ratio of energy deposited in the 9 crystals around the seed cell over the energy deposited in the 25 crystals around it.
- Cluster width: is the average RMS width of the shower, measured in cm.

The selection $E_9/E_{25} > 0.9$ is applied in the preselection stage and it is not used further in the continuum suppression steps. The cluster width is used in the continuum suppression described in the following sections.



(a)

(b)

Figure 5.7: Calorimeter variables before continuum suppression: (a) E_9/E_{25} and (b) width.

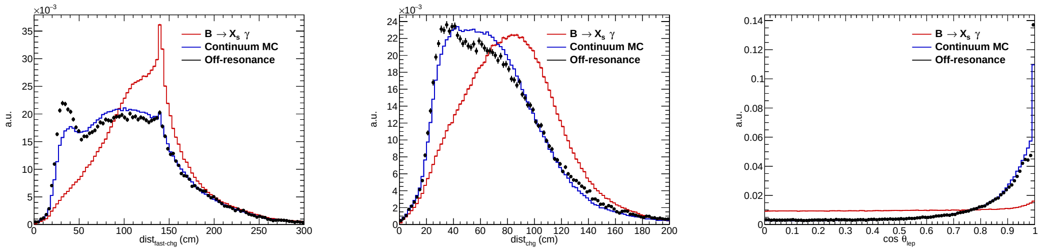
5.4.3 Topological variables

In addition to the variables described so far, several other variables can be studied in order to characterize the spatial distribution of the decay products in the event. They are in general called topological variables. They use all or most particles in the event in order to construct various sets of variables. The following variables are used:

Isolation variables

Most of the photon background consists of photons originating from light meson decays, namely $\pi^0/\eta \to \gamma\gamma$, as well as bremsstrahlung photons. Only light mesons with a large momentum can produce photons with an energy above the threshold for this analysis, thus it is expected that their decay products are produced in the direction of their mother momentum, and close to each other. To veto this kind of events, a number of variables that ensure the prompt photon cluster to be isolated in the ECL are studied. Examples are shown in Fig. 5.8:

- Distance of the cluster in the ECL to the closest track, $dist_{\text{chg}}$.
- Distance of the cluster in the ECL to the closest track with $p^* > 1.0\,\text{GeV}$ (such tracks are here called “fast tracks”), $dist_{\text{fast-chg}}$.
- Cosine of the angle in the CM between the cluster and the tag lepton, $\cos\theta_{\text{lep}}$.



`(a)` `(b)` `(c)`

Figure 5.8: Photon cluster isolation variables before continuum suppression: `(a)` distance to fast track, `(b)` distance to closest track and `(c)` cosine of the angle between the cluster and the tag lepton.

Thrust and thrust axis

The thrust axis, $\vec{T}$, is defined as the unit vector that maximizes the sum of momentum projection of all particles of the event onto its direction. A strong correlation between the direction of $\vec{T}$ and the direction of the decay products is expected for continuum events, while in $B\overline{B}$ events there is no preferred direction. The angle between $\vec{T}$ and the z-axis, $\cos\theta_{thr}$, is considered. The vector $\vec{T}$ is constructed with the CM momenta of all the good neutral clusters and all well reconstructed tracks. An additional variable used is the “thrust”, which uses the unit vector $\vec{T}$ and the three-momentum of all particles $\vec{p}_i$ and is defined as:

$$t = \sum_i \frac{\vec{p}_i \cdot \vec{T}}{|\vec{p}_i|}. \tag{5.8}$$

For perfectly "back-to-back" events the value of thrust is 1, for isotropic events it is 0.5. The distributions can be seen in Fig. 5.9.

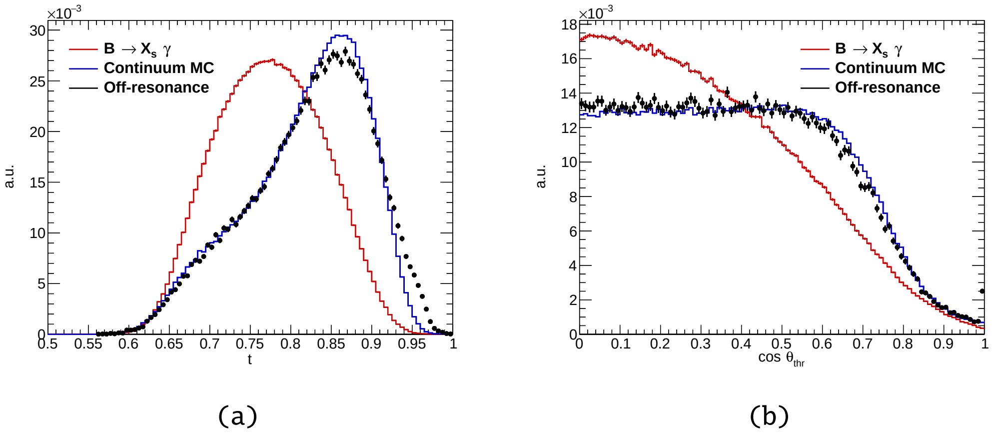


Figure 5.9: Thrust variables before continuum suppression: `(a)` thrust and `(b)` cosine of the thrust angle.

Fox-Wolfram moments

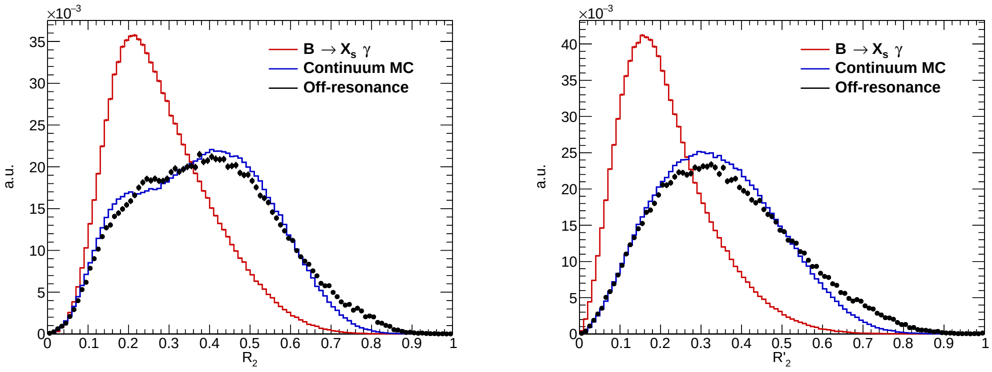
The Fox-Wolfram moments (FWM) were first proposed in 1978 as a background suppression tool at e^-e^+ colliders [82]. They are rotationally invariant variables, constructed with Legendre polynomials, and characterize the distribution of the decay products in the event. They are defined as:

$$H_l = \sum_{i,j} \frac{|\vec{p}_i| \cdot |\vec{p}_j|}{E^2_{\text{vis}}} P_l(\cos\theta_{ij}), \tag{5.9}$$

where the sum includes the momenta of all neutral and charged particles, E_{vis} is the event's visible energy and $P_l(\cos\theta_{ij})$ are the Legendre polynomials with the angle θ_{ij} being measured between the i^{th} and j^{th} particles. It is common practice to normalize them by the 0^{th} component, $R_l = H_l/H_0$, in order to remove the explicit dependence on the energy. From now on these normalized moments will be simply referred to as the Fox-Wolfram moments.

Three different sets of FWM are studied: the first uses all particles in the event and is symbolized by R_l; in the second the prompt photon is excluded and the FWM are called "partial FWM", R'_l; the last one uses all particles in the event as well as the missing momentum $\vec{p}_{\text{miss}}$ and the FWM are denoted by R_l^{miss}. Examples are shown in Fig. 5.10. Not all of these variables are later used in the analysis, since they have large correlations.

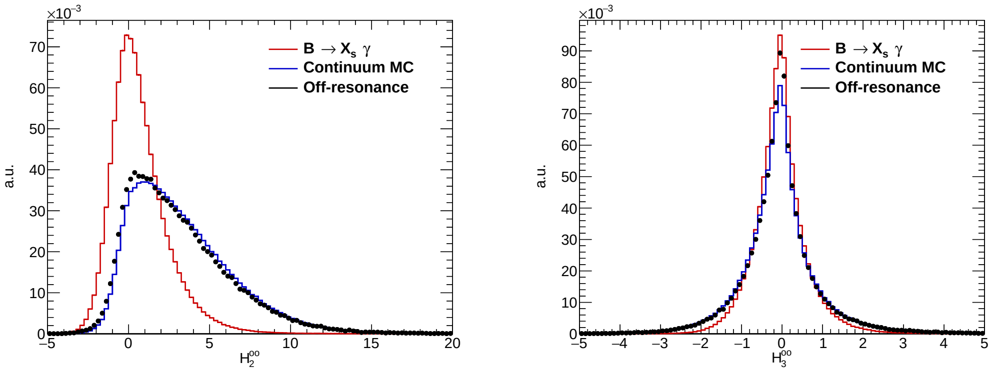
The *Super Fox-Wolfram moments* (SFW) and *Kakuno Super Fox-Wolfram moments* (KSFW) variables are standard continuum suppression tools developed by Belle [83–85]. They are Fisher



(a) (b)

Figure 5.10: Fox Wolfram Moment R_2 before continuum suppression: (a) using all particles in the event and (b) excluding the candidate photon.

discriminants that use modified versions of the FWM. In the case of SFW, particles are split into the products of the "signal" B meson and "others". In the KSFW case, particles are further separated into charged and neutral. These Fisher discriminants are usually combined with other variables to form a likelihood ratio to discriminate signal from background events. Since neither B meson is reconstructed here, the particles are split into signal (s), defined as the signal photon and tag lepton, and all the other particles (o), such that the same categorization as the default SFW and KSFW can be used. New sets of FWM which follow the same definition as Eq. (5.9) are constructed, but are labeled ss, oo and so, depending on which sets particles were used. An example of two of the modified FWM is shown in Fig. 5.11.



(a) (b)

Figure 5.11: Examples of modified Fox Wolfram Moments in the oo category before continuum suppression: (a) H_2^{oo} and (b) H_3^{oo}.

5.4.4 Boosted Decision tree

To reach an optimal separation of signal from continuum background, a multivariate analysis with several input variables is used. In this way, phase space regions where the signal is significantly enhanced relative to the the background can be selected. The best possible way to do this efficiently is using machine learning techniques, such as Boosted Decision Trees. A decision tree is a series of “yes/no” decisions taken on single variables, that permit the categorization of an event in one of two categories, signal-like or background-like, as shown in Fig. 5.12. Signal and background MC samples are used to iteratively determine the structure of the tree that yields the maximum separation. A Boosted Decision Tree (BDT) represents a series of decision trees that form a “forest”. In practical applications, thousands of trees can be created and tens of variables can be used as input parameters. Boosting is the process by which events that were misidentified after an iteration are given a higher weight to enhance the classification performance in the following steps.

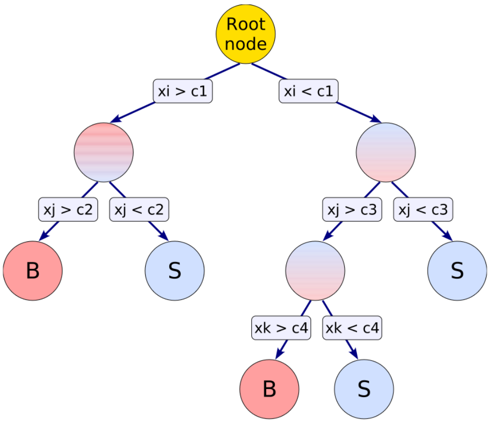


Figure 5.12: Diagram of a decision tree: cuts on single variables lead to a classification of an event as signal-like (S) or background-like (B) [86].

Several BDT parameters can be tuned to improve the performance: the depth indicates the maximum number of levels before the decision is taken; the number of leaf nodes indicates the number of variables used on each level of decision making; shrinkage is related to the “learning speed”, it allows one to control the weight of the single trees, and makes the method more stable by forcing more trees to be grown.

Highly correlated variables do not improve the separation power of the BDT and make the structure unstable with respect to statistical fluctuations [86]. For this reason several of our variables had to be dropped since no new information was provided by them. The R_l^{miss} set of FWM was dismissed due to its large correlation to R_l. Different BDTs were trained using each H_l^{oo} and H_l^{so} modified FWM set, and both together. Only a slight improvement was found when using all these variables, therefore the set H_l^{oo}, consisting of five variables, is the only one chosen.

The kinematic, thrust and calorimeter variables were found to be safe: they are fairly uncorrelated to the others variables and were kept in the training. In total, 19 variables were used in the BDT training. They are summarized in Fig. 5.13, the correlations among them in signal and background are shown in Fig. 5.14, where additionally the correlation to the photon CM energy is given.

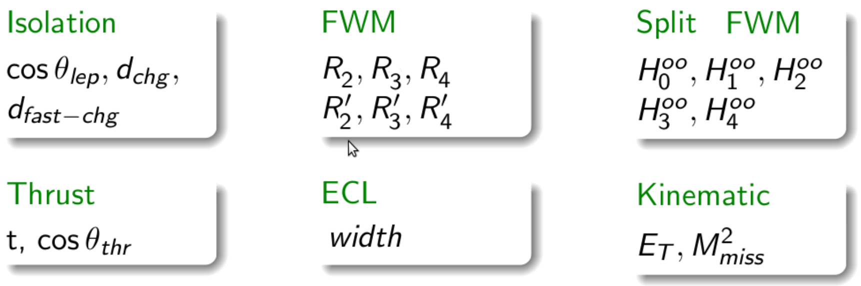


Figure 5.13: Variables used in the continuum suppression BDT.

The BDT is trained using 1000 trees with a maximum depth of four layers and a maximum of five leaf nodes. Additionally, the learning speed was reduced by setting the shrinkage parameter to 0.1. The chosen boosting algorithm was “Gradient Boost” [86, section 7.2]. For training and testing initially 5×10^6 $B \to X_s\gamma$ events and one stream of continuum MC are used, the latter containing $\sim 1.9 \times 10^7$ events. After preselection the signal sample contains $\sim 1.9 \times 10^5$ events and the continuum sample $\sim 9 \times 10^5$. Half of the events are used for training and half for testing. The result of the training is shown in Fig. 5.15 for off-resonance data and the different MC samples. Continuum events peak clearly around -1 while the signal has a roughly uniform distribution. The signal events that peak at low values of the BDT score are due to the reconstructed lepton being a secondary or fake lepton candidate.

A figure-of-merit (FOM) is used to find the optimal selection criterion in the BDT score that maximizes the statistical significance on the $B \to X_s\gamma$ yield. The overall FOM takes into account the statistical uncertainties in the determination of the number of signal, $B\overline{B}$ and continuum events, and the proper scaling of each of the components:

$$\mathrm{FOM} = \frac{N_{B\to X_s\gamma}}{\sqrt{(\delta N_{B\to X_s\gamma})^2 + (\delta N_{\mathrm{off}})^2 + (\delta N_{B\overline{B}})^2 + (\delta N_{\mathrm{rare}})^2 + (\delta N_{\mathrm{ulnu}})^2}}. \tag{5.10}$$

Here $N_{B\to X_s\gamma}$ represents the number of $B \to X_s\gamma$ events, N_{off} the number of off-resonance events, $N_{B\overline{B}}$ the number of $B\overline{B}$ events from generic MC, N_{rare} the number of $B\overline{B}$ events from the rare MC sample, and N_{ulnu} the number of $B\overline{B}$ events from the $B \to X_u\ell\nu$ sample. The uncertainties δN_i take into account the scaling factors corresponding to each sub-sample. $N_{B\to X_s\gamma}$ is calculated by subtracting continuum and $B\overline{B}$ background from the on-resonance sample:

$$N_{B\to X_s\gamma} = N_{\mathrm{data}} - \alpha_{\mathrm{off}}N_{\mathrm{off}} - \alpha_{B\overline{B}}N_{B\overline{B}} - \alpha_{\mathrm{rare}}N_{\mathrm{rare}} - \alpha_{\mathrm{ulnu}}N_{\mathrm{ulnu}}. \tag{5.11}$$

The scaling factors for the MC samples are $\alpha_{B\overline{B}} = 1/10$, $\alpha_{\mathrm{rare}} = 1/50$ and $\alpha_{\mathrm{ulnu}} = 1/20$.

The FOM is shown in Fig. 5.16. The optimization was done for the $1.7\,\mathrm{GeV} \leq E^*_\gamma \leq 2.8\,\mathrm{GeV}$ region. The optimal cut to be used in the analysis is found to be BDT$\geq$ 0.25.It gives an expected statistical precision of 3.2 % on the $B \to X_s\gamma$ yield.

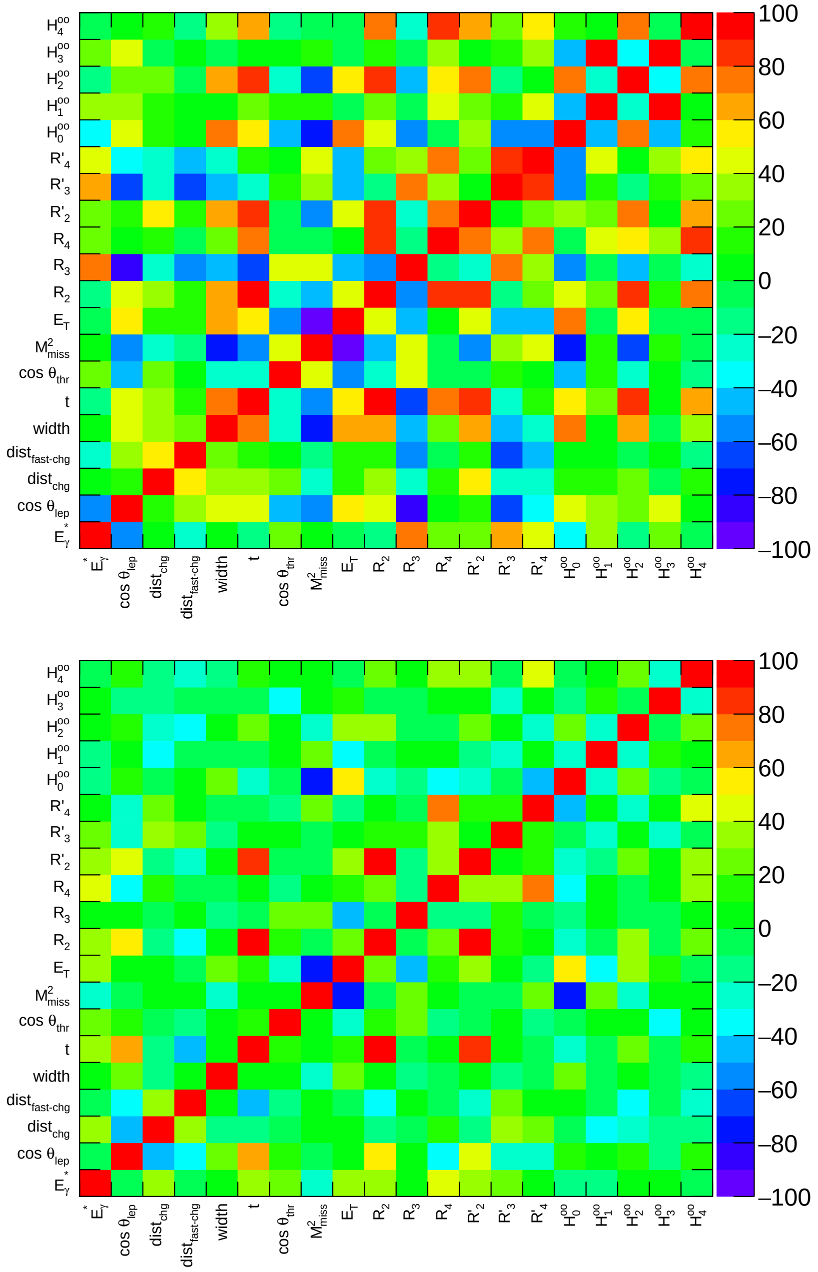


Figure 5.14: Correlations among the variables that are input to the BDT and correlations to the photon CM energy. The top plot is for continuum MC and the bottom plot for signal MC.

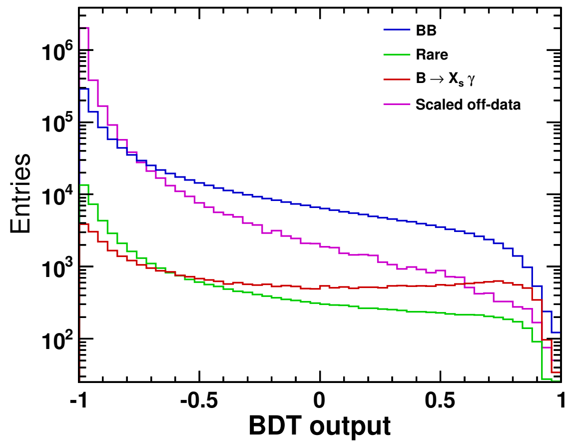


Figure 5.15: BDT output for the different MC samples and off-resonance data.

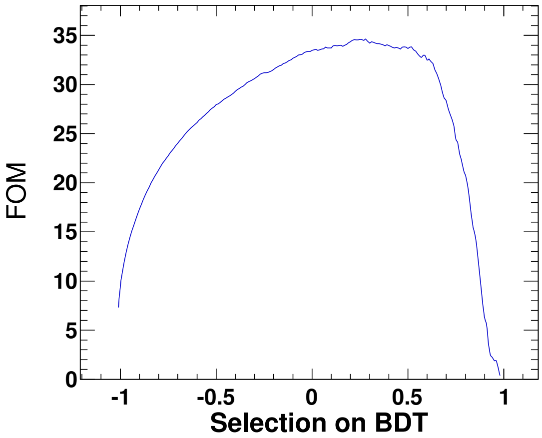


Figure 5.16: Figure of merit for different BDT output cuts. The selection that maximizes the FOM is BDT ≥ 0.25.

5.4.5 Photon spectrum after selection

After applying the complete selection, including BDT selection, a photon spectrum containing 43 008 on-resonance and 702 off-resonance events in the signal region between 1.7 to 2.8 GeV is obtained. In the low-energy sideband between 1.4 GeV and 1.7 GeV, 26 470 on-resonance and 159 off-resonance events are found, while in the high-energy sideband 2.8 GeV and 4.0 GeV 3 018 on-resonance and 371 off-resonance events are obtained. Based on MC simulation, about ~ 20 % of the events in the signal region are expected to come from $B \to X_{s+d}\gamma$ decays. In Fig. 5.17 the composition of the spectrum showing the continuum and $B\overline{B}$ background sources as well as the signal component is shown.

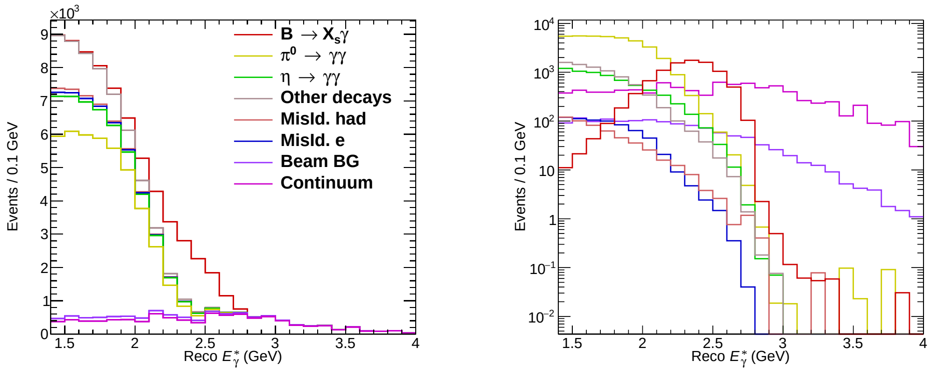


Figure 5.17: Expected composition of the photon spectrum after selection. The $B\overline{B}$ and $B \to X_s\gamma$ components are the MC prediction, while continuum is off-resonance data. The right-hand-side plot is shows the same components using a logarithmic scale for the y-axis and without stacking them, such that their shape and size can be directly compared.

CHAPTER 6

Correction of background components

Monte Carlo methods simulate, to the best of the current knowledge, physical processes that occur in high-energy collisions. There are many effects that play a role in a precise description of the data, so large efforts go in making simulated samples as accurate as possible. Typically several corrections are needed in order to improve the description of the data. In this chapter, the corrections to the background normalization, shape and selection efficiencies are described.

6.1 Correction of background sources

The following corrections are common to the spectra used in the measurement of the inclusive branching fraction and the CP asymmetry. They are obtained using data sidebands and dedicated control samples.

6.1.1 Correction of inclusive semileptonic branching fractions

The inclusive charmed semileptonic $B \to X_c \ell \nu$ branching fractions used in the generation of the MC do not match the latest measured values, since their production precedes them. For this reason the values must be corrected to match the state-of-the-art values as shown in Table 6.1. Since the measurements use a minimum lepton momentum $p_\ell^* > 0.6\,\text{GeV}$, the extrapolation factor to the full phase space of 1.0495 is used.

Component	Belle MC	Belle [87] ($p_\ell > 0.6\,\text{GeV}$)	Belle extrapolated	Correction
$B^0 \to X_c \ell \nu$	10.45 %	(9.80 ± 0.36) %	(10.29 ± 0.38) %	0.9847
$B^+ \to X_c \ell \nu$	11.35 %	(10.34 ± 0.36) %	(10.85 ± 0.36) %	0.9559

Table 6.1: Corrections to the inclusive semileptonic branching fractions in the Belle MC.

6.1.2 Correction of the lepton spectrum from semileptonic decays

Semileptonic decays in the Belle MC were generated in 2009, so the form factors involved in these decays are not up to date. A dedicated Belle study [88] provides tools to reweight the default MC to follow a theoretical prediction [89] that uses state-of-the-art form factor values. The tool described in the note allows to reweight $B \to D\ell\nu$, $B \to D^{(*)}\ell\nu$ and $B \to D^{**}\ell\nu$ events with weights given in bins of p_ℓ^* and q^2, where q^2 is the momentum of the virtual $W^\pm$ boson exchanged in the decay. The momentum spectra for electrons and muons before and after the reweighting are shown in Figs. 6.1 and 6.2, The components of the lepton spectrum have been described in Section 5.3.3. A π^0 control sample is used for the validation of several corrections in this analysis, including this one. It follows the same selection as the signal sample, except for the π^0 veto, which is reversed to require that the photon candidate comes from a π^0 decay. This sample is used here to show the improved agreement between data and MC after correcting the form factors on the generic $B\overline{B}$ MC.

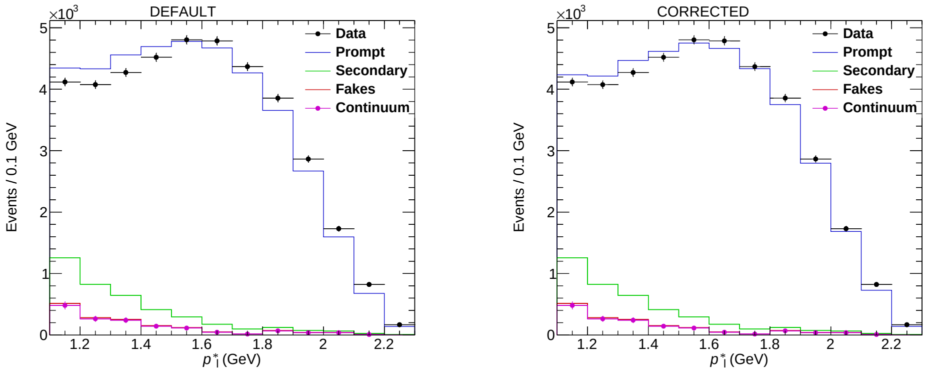


Figure 6.1: Electron spectrum in the π^0 control mode. The left-hand-side plot shows the default Belle MC and the right-hand-side the corrected one.

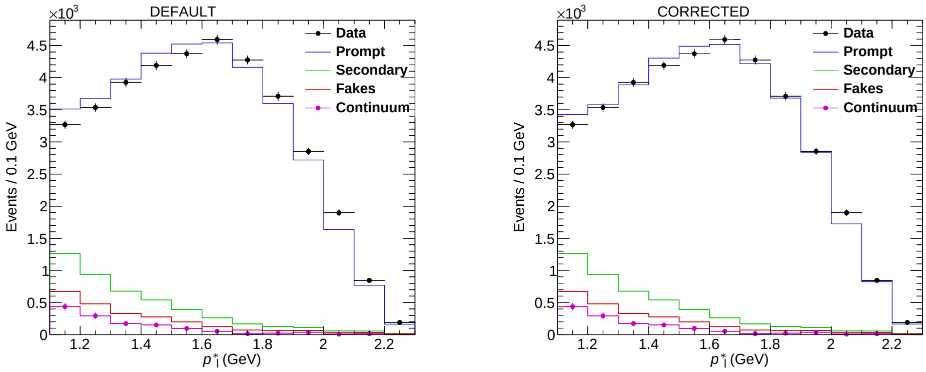


Figure 6.2: Muon spectrum in the π^0 control mode. The left-hand-side plot shows the default Belle MC and the right-hand-side the corrected one.

The correction improves the agreement between data and MC greatly. Several dedicated analyses have shown the difficulty of properly describing the inclusive semileptonic spectrum, particularly due to contributions from poorly known excited D states (D^{**}). It is not the goal of this analysis to study this problem in detail, and for this study the MC describes the data sufficiently well. The normalization of the photon background components is studied in different control samples, therefore the effect of the reweighting procedure is important in the determination of the selection efficiency and the fractions of the different components.

A second correction deals with the fraction of lepton fakes. This contribution is corrected by studying $D^{*+} \to D^0(K^-\pi^+)\pi^+$ decays. The same particle identification selection criteria as used for the tag lepton (eID > 0.9 or muID > 0.9) are applied to the $\pi^\pm$ and $K^\pm$ candidates in the decay. In this way the fraction of $\pi^\pm$ or $K^\pm$ that are misidentified as a lepton are determined, and the MC is corrected accordingly to match the data. This correction factor and its uncertainty depend on the charge, flavor, momentum and polar angle of the misidentified hadrons. It has already been included in Figs. 6.1 and 6.2. Similarly to the previous correction, it is taken into account as a systematic effect that could affect the selection efficiency.

6.1.3 Background from π^0 and η decays

The the dominant $B\overline{B}$ background sources after the selection are photons from π^0 and η decays, making up about 49 % and 8 % of the total yield, respectively.

There are cases in which the one of the photons in a $\pi^0(\eta) \to \gamma\gamma$ decay has a high energy and is identified as our candidate, but the second photon has very low energy and cannot be found due to threshold requirements. This makes the reconstruction of the π^0 or η impossible, hence vetoing such event is impossible. The background from $\pi^0(\eta) \to \gamma\gamma$ decays is normalized using inclusive $B \to X\pi^0(\eta)$ samples. The samples are obtained by removing the veto requirement from the selection chain and combining the prompt photon with any other photon in the event. For all combinations the diphoton mass ($m_{\gamma\gamma}$) is calculated and a fit performed to the $m_{\gamma\gamma}$ peak around the π^0 and η masses to estimate the number of these mesons in data and MC.

The second photon in the event is required to have an energy larger than 50 MeV. All combinations with a mass smaller than 0.7 GeV are considered. The nominal masses of the mesons of interest are $m_{\pi^0} = 0.135$ GeV and $m_\eta = 0.548$ GeV, as can be seen in Fig. 6.3. For each prompt photon there are on average five combinations passing the requirements. All ten streams of generic $B\overline{B}$ MC are used to avoid limitations in the statistical precision of the correction factors due to insufficient MC events.

The diphoton combinations are categorized according to their “mother” using true MC information as described below and shown in Figs. 6.4 and 6.5.

(a) True π^0 (η): both photons come from the same meson.

(b) Cascade combinations: photons come from different π^0 (η) candidates in the B decay chain, or only one of the photons comes from a π^0 (η).

(c) Random combinations: neither one of the photons comes from a π^0 (η).

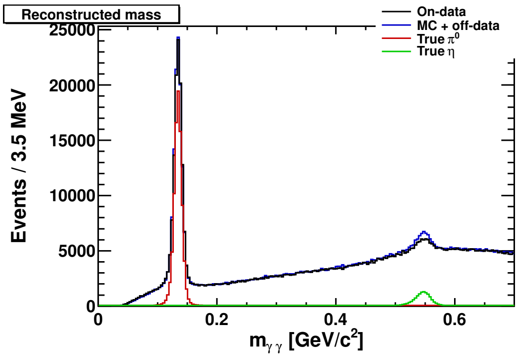
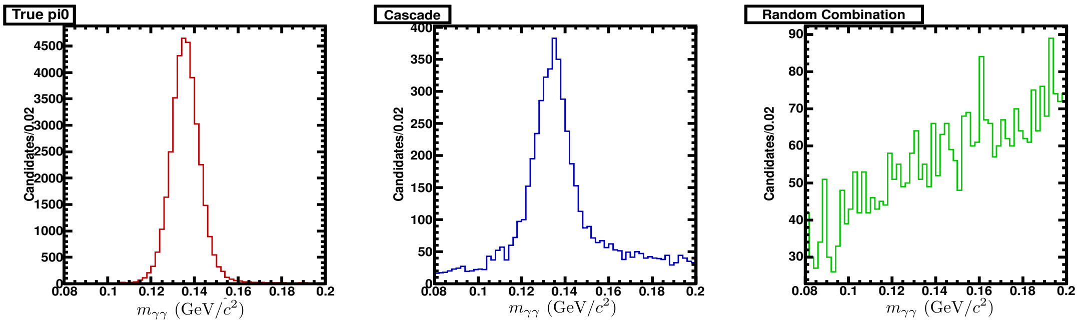


Figure 6.3: Reconstructed diphoton mass, showing the peaks corresponding to the π^0 and η.

The shape of the distribution for the different categories is determined using MC distribution for 11 p_ℓ^* bins and in a mass window around the nominal masses: $0.08\,\mathrm{GeV/c^2} \leq m_{\gamma\gamma} \leq 0.20\,\mathrm{GeV/c^2}$ for π^0 and $0.45\,\mathrm{GeV/c^2} \leq m_{\gamma\gamma} \leq 0.65\,\mathrm{GeV/c^2}$ for η.



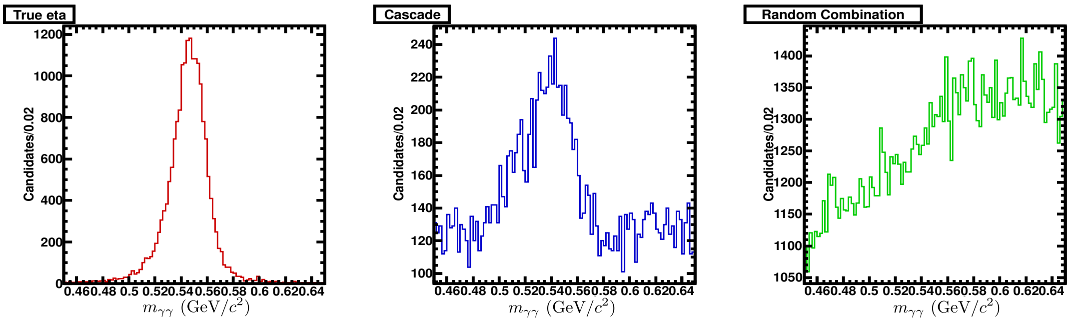
(a) (b) (c)

Figure 6.4: Reconstructed $\pi^0 \rightarrow \gamma\gamma$ diphoton mass categories for (a) true π^0, (b) cascades and (c) random combinations.

The random combinations are fitted using a second-order Chebyshev polynomial. The true mass peak is modeled with a Cruijff function [90], which is a bifurcated Gaussian function with asymmetric tails:

$$f(m; m_0, \sigma_L, \alpha_L, \sigma_R, \alpha_R) = \begin{cases} \exp\left(\frac{(m-m_0)^2}{2\sigma_L^2+\alpha_L(m-m_0)^2}\right) & \text{if } m > m_0 \\ \exp\left(\frac{(m-m_0)^2}{2\sigma_R^2+\alpha_R(m-m_0)^2}\right) & \text{if } m < m_0 \end{cases} \tag{6.1}$$

This function is well suited to perform this study: the diphoton shape is sensitive to the slight asymmetry of the mass distributions due to the photon energy resolution, and to the cascade component which, as seen in Figs. 6.4 and 6.5, broadens the tails of the diphoton peak. The fit is initially performed in MC, allowing us to extract the function parameters $m_0, \sigma_L, \alpha_L, \sigma_R$ and α_R for each meson momentum bin.



(a)　(b)　(c)

Figure 6.5: Reconstructed $\eta \to \gamma\gamma$ diphoton mass categories for (a) true η, (b) cascades and (c) random combinations.

Examples of the fits can be observed in Figs. 6.6 and 6.7. Due to the large sample statistics and limited calorimeter resolution, the tails of the distributions cannot be perfectly described by the fit function. This translates on non-vanishing residuals (the difference between data and the fit function). It was attempted to fit the diphoton mass distributions with more complex functions (e.g. with one or two additional Gauss components) in order to better describe the tails of the distributions. These fit setups were not successful since the increase in the number of parameters made the fit unstable and the fit parameters poorly constrained.

Inclusive $B \to X\pi^0(\eta)$ decays are only found on the $B\overline{B}$ part of the data sample, so the $q\overline{q} \to X\pi^0(\eta)$ contributions must be subtracted using off-resonance data. The correction factor for each meson momentum bin is the ratio of continuum subtracted on-resonance data to MC:

$$c = \frac{N_{\text{on}} - \alpha_{\text{off}} \cdot N_{\text{off}}}{N_{\text{generic MC}} + N_{\text{rare MC}}}. \tag{6.2}$$

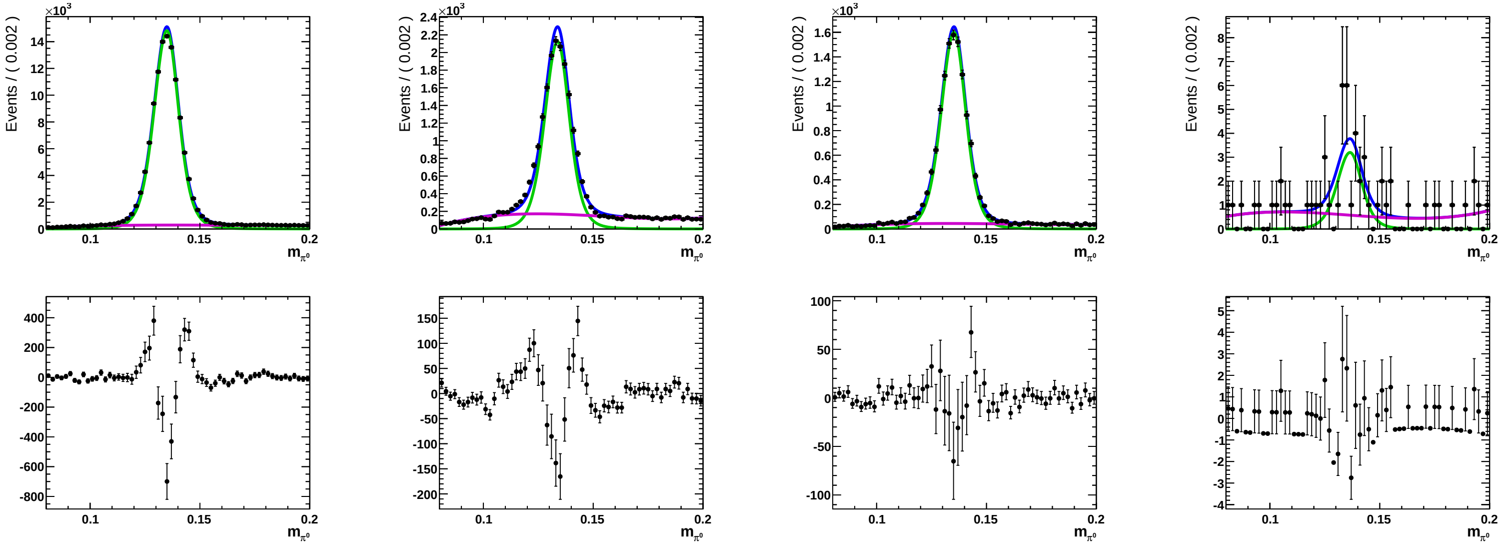
All correction factors for the inclusive $B \to X\pi^0(\eta)$ branching fractions are shown in Fig. 6.8. The yields extracted from the fits and corresponding correction factors are summarized in Tables 6.2 and 6.3.

6.2 Additional corrections used for the measurement of the branching fraction

The following corrections are applied to obtain the photon spectrum used in the measurement of the inclusive $B \to X_s\gamma$ branching fraction.

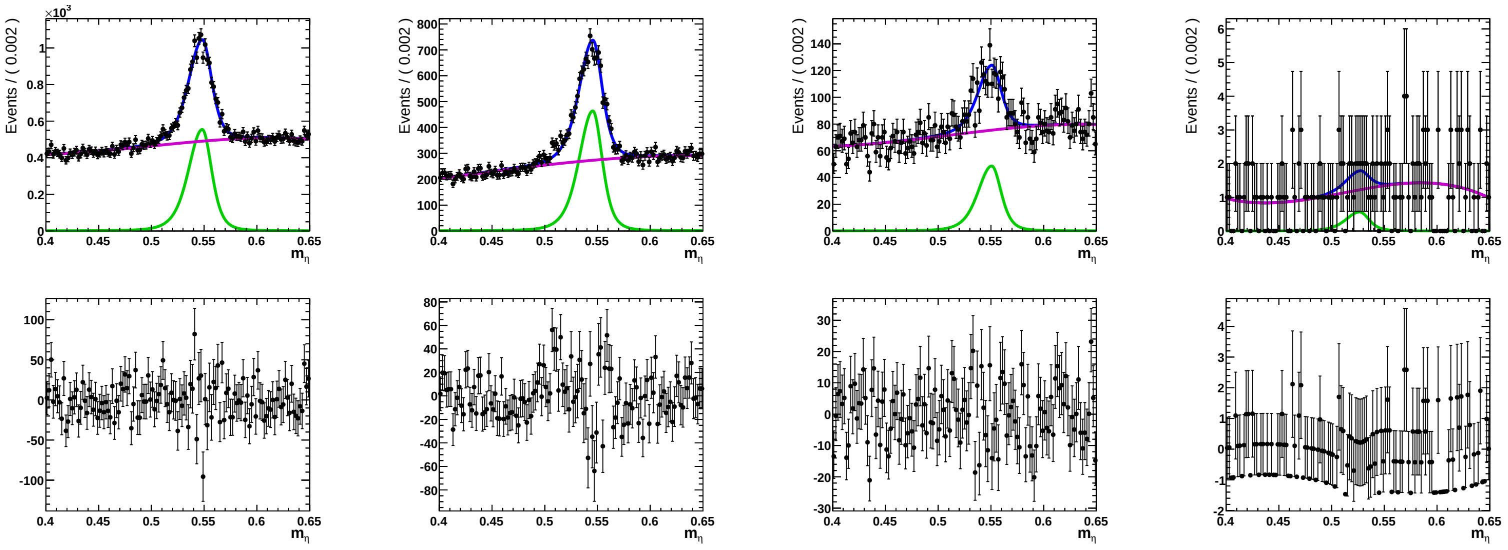
6.2.1 Hadronic background

A small part of the photon background comes from hadronic interactions in the ECL that produce a cluster. This background amounts to $\sim 0.2\%$ in the signal region $1.7\,\text{GeV} < E^*_\gamma < 2.8\,\text{GeV}$, and $\sim 1\%$ for $E^*_\gamma < 1.7\,\text{GeV}$. The hadron-cluster energy spectra in the lab and CM frames are shown



(a) (b) (c) (d)

Figure 6.6: Fits to π^0 mass spectrum in MC and data samples in the region $2.1\,\mathrm{GeV/c} \le |\vec{p}^{\,*}_{\pi}| \le 2.2\,\mathrm{GeV/c}$: (a) $B\overline{B}$ MC, (b) rare MC (c) on-resonance data and (d) off-resonance data. The top plots show the data or MC points and the fit components and the bottom plots the fit residuals.



(a) (b) (c) (d)

Figure 6.7: Fits to η mass spectrum in MC and data samples in the region $2.1\,\mathrm{GeV/c} \leq |\vec{p}^{\,*}_{\pi}| \leq 2.2\,\mathrm{GeV/c}$: (a) $B\overline{B}$ MC, (b) rare MC (c) on-resonance data and (d) off-resonance data. The top plots show the data or MC points and the fit components and the bottom plots the fit residuals.

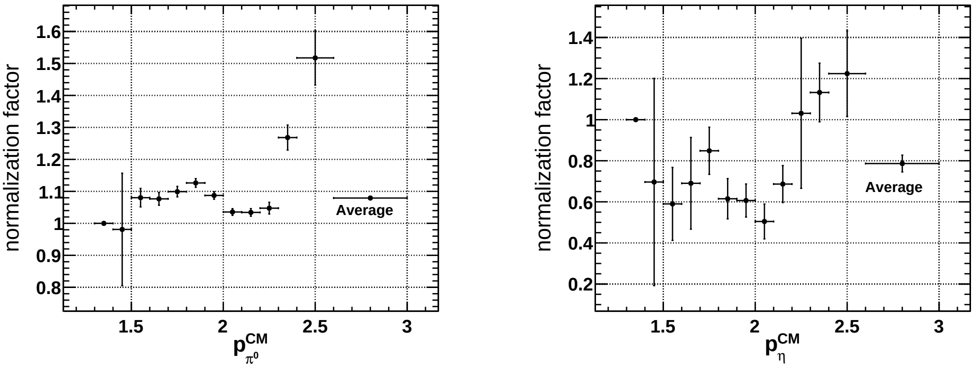


Figure 6.8: Correction factors for (a)$\mathcal{B}[B \to X\pi^0]$ and (b)$\mathcal{B}[B \to X\eta]$. The factors are given in bins of meson momentum and the last bin shows the average over all bins.

Momentum bin [GeV/c]	On data	Off data	Generic MC	Rare MC	Correction
1.4 – 1.5	584 ± 41	0 ± 20	614 ± 13	8 ± 1	0.938 ± 0.076
1.5 – 1.6	3935 ± 91	117 ± 43	3699 ± 27	54 ± 2	1.018 ± 0.028
1.6 – 1.7	6965 ± 112	221 ± 54	6595 ± 34	94 ± 2	1.008 ± 0.019
1.7 – 1.8	9793 ± 125	260 ± 57	8982 ± 37	131 ± 2	1.046 ± 0.016
1.8 – 1.9	12676 ± 136	183 ± 52	11439 ± 41	163 ± 2	1.077 ± 0.013
1.9 – 2.0	14836 ± 142	191 ± 59	13812 ± 43	206 ± 3	1.045 ± 0.011
2.0 – 2.1	15023 ± 138	173 ± 54	14573 ± 42	247 ± 3	1.002 ± 0.010
2.1 – 2.2	11578 ± 118	186 ± 47	11077 ± 36	304 ± 3	1.001 ± 0.012
2.2 – 2.3	6386 ± 88	314 ± 56	5616 ± 25	340 ± 3	1.019 ± 0.018
2.3 – 2.4	2560 ± 49	104 ± 48	1554 ± 13	385 ± 3	1.267 ± 0.037
2.4 – 2.6	1766 ± 50	181 ± 73	384 ± 7	670 ± 4	1.505 ± 0.085

Table 6.2: Yields from the π^0 fits in the different samples, and corresponding correction factors for 11 meson momentum bins.

Momentum bin [GeV/c]	On data	Off data	Generic MC	Rare MC	Correction
1.4 – 1.5	558 ± 206	80 ± 92	644 ± 56	19 ± 4	0.721 ± 0.346
1.5 – 1.6	785 ± 199	122 ± 77	1133 ± 55	24 ± 4	0.573 ± 0.186
1.6 – 1.7	1266 ± 166	317 ± 101	1309 ± 48	42 ± 4	0.703 ± 0.146
1.7 – 1.8	1314 ± 142	67 ± 63	1400 ± 43	54 ± 3	0.857 ± 0.110
1.8 – 1.9	1031 ± 130	95 ± 74	1415 ± 36	74 ± 4	0.629 ± 0.102
1.9 – 2.0	940 ± 106	59 ± 55	1421 ± 32	86 ± 3	0.584 ± 0.080
2.0 – 2.1	932 ± 78	266 ± 68	1218 ± 26	108 ± 3	0.502 ± 0.079
2.1 – 2.2	759 ± 63	65 ± 57	851 ± 21	144 ± 3	0.698 ± 0.087
2.2 – 2.3	692 ± 55	24 ± 55	484 ± 14	173 ± 3	1.017 ± 0.120
2.3 – 2.4	525 ± 47	31 ± 35	218 ± 8	193 ± 3	1.201 ± 0.144
2.4 – 2.6	537 ± 48	69 ± 57	59 ± 6	318 ± 3	1.239 ± 0.199

Table 6.3: Yields from the η fits in the different samples, and corresponding correction factors for 11 meson momentum bins.

in Fig. 6.9. From all clusters, ~ 65% correspond to anti-neutrons and ~ 30% to K_L. There is no easy way to obtain a clean sample of anti-neutrons, therefore clusters from proton and anti-proton interactions in the ECL are studied to try to draw conclusions about anti-neutron clusters. This is achieved by studying a pure sample of $\Lambda \to p^+\pi^-$ ($\bar{\Lambda} \to \bar{p}\pi^+$) decays.

The Λ baryons are relatively "long-lived", so their decay vertex is displaced with respect to the IP. Displaced vertexes of two tracks can correspond to Λ baryons or K_S mesons, therefore special tools have been developed at Belle to find such decays and distinguish them from each other. The K_S/Λ finding tool uses information about the position of the vertex, direction of the tracks, number of SVD and CDC hits and PID score of the tracks. The information is combined in two neural networks that allow neutral particles to be selected from two-track combinations ("V-like" decays), and to discriminate between K_S and Λ. The output of these neural networks are called `nb_vlike`, for V-like decays and `nb_nolam` for the decay being consistent with a K_S (or conversely, *not* been consistent with a Λ). Scores of `nb_vlike` close to 1 mean the two-track combination likely arises from the same neutral particle. A `nb_nolam` score close to -1 means the neutral particle is likely a Λ (a score close to 1 means it is likely a K_S).

The applied selection criteria are select `nb_vlike` < 0.95 and `nb_nolam` < -0.95. The selection achieved with both neural networks is very clean and the obtained Λ samples are background free. In Fig. 6.10, the mass distributions for the $p^+\pi^-$ and $\bar{p}\pi^+$ combinations are shown for the cases where either the proton or pion track has an associated ECL cluster.

For the cases where either the proton or anti-proton track has an associated ECL cluster, the E_9/E_{25} cluster shape is compared between data and MC in Fig. 6.11. The first relevant observation is that the overall number of Λ candidates is underestimated by ~ 30% in the MC, as can be seen in the first column of plots of Fig. 6.11. This number is obtained from scaling the MC to match the yield in data. The second thing to note is that the shape of the E_9/E_{25} distribution is very poorly modeled for anti-protons, but not for protons. This translates to a selection efficiency ($N_{E_9/E_{25}>0.9}/N_{\text{total}}$) in MC that is half of the selection efficiency in data, for energies above 1 GeV.

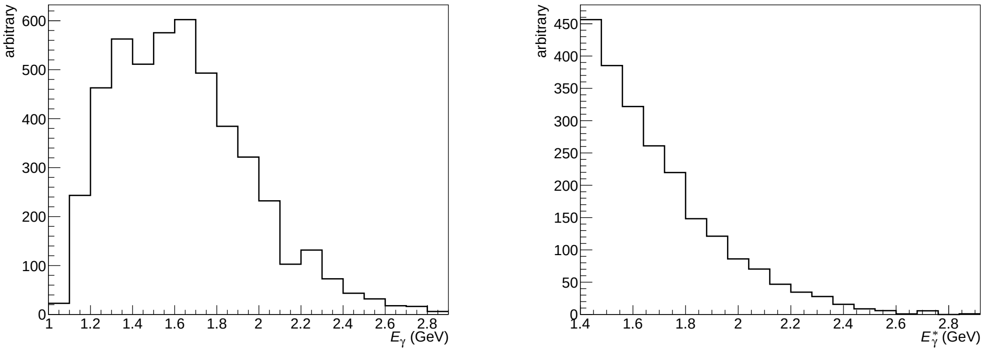


Figure 6.9: Hadron background cluster energy in the lab frame (left) and CM frame (right). The background accumulates at lower energies in the CM frame, and is consistent with the background shape estimation in the $B \to X_{s+d}\gamma$ selection.

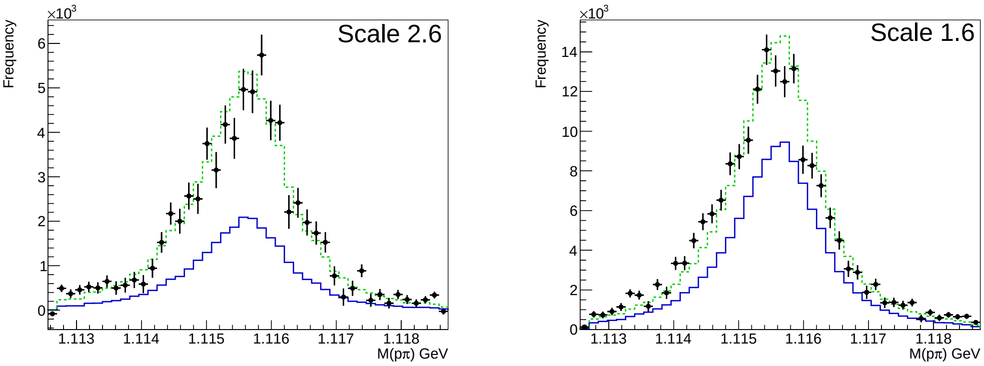


Figure 6.10: Proton-pion invariant mass distributions for the cases where the proton track has an associated ECL cluster. The particle with an associated cluster in the left plot is an anti-proton and in the right plot a proton. The blue distribution corresponds to the MC, the green is the MC scaled by the factor displayed on the plots to match the number of events in data.

Conversely for protons, the ratio of selection efficiencies is consistent with 1. Both these effects, number of Λ baryons and E_9/E_{25} efficiency, are taken into account for correcting the hadronic background.

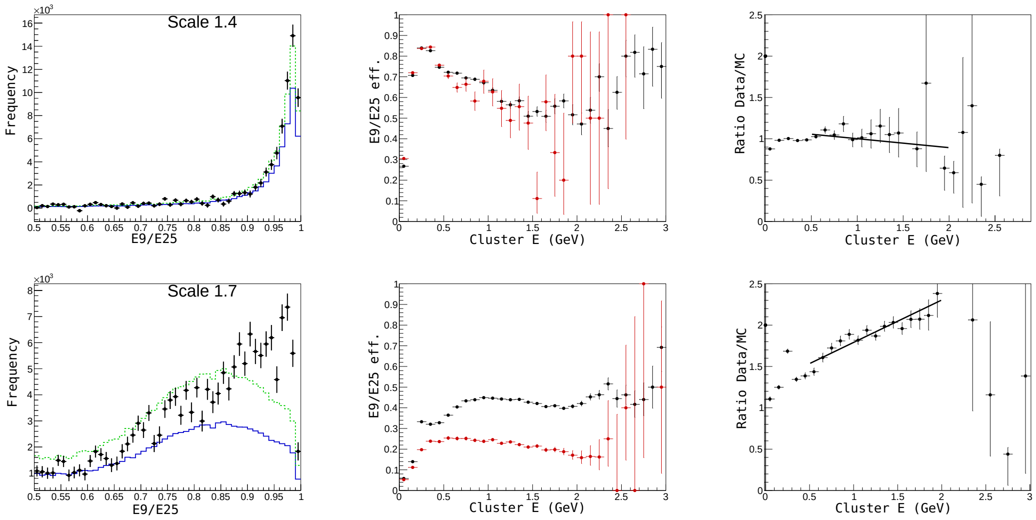


Figure 6.11: Distributions of E_9/E_{25} and selection efficiency for protons (top row) and anti-protons (bottom row) in data and MC, for all cluster energies. The leftmost plot shows the continuum-subtracted data as black points, $B\overline{B}$ MC in blue, and $B\overline{B}$ MC normalized to data in green. The central plots show the selection efficiency for $E_9/E_{25} > 0.9$, the right plots show the ratio of efficiencies of data and MC.

When trying to assess the selection efficiency for anti-neutrons, some assumptions about the validity of the anti-proton study must be made. The annihilation of the anti-nucleons in the calorimeter proceeds through strong processes, therefore it is assumed that the cut efficiency in data is similar for both types of anti-nucleons. The E_9/E_{25} distributions for anti-nucleons in data and MC with a lab energy greater than 1 GeV are shown in Fig. 6.12, for data only anti-protons can be shown. As mentioned previously, the selection efficiency in MC is half of the selection efficiency in data, so the MC must receive a correction of 100 %. A 50 % uncertainty on the correction is assigned to account for possible differences on the E_9/E_{25} shape and efficiency for anti-neutrons.

6.2.2 Background from other decays

The remaining background is composed of photons from several different sources: decays of ω, η' and J/ψ mesons, final state radiation in $B \to X\ell\gamma\nu$ decays, and others. It is simply called "others". None of these single components is dominant, nor is there a strong difference in their shapes. It is thus difficult to find control samples that can serve to study these single components. Given that this in the only remaining component that has not been corrected, the sideband ($1.4\,\text{GeV} < E^*_\gamma < 1.55\,\text{GeV}$) is used to scale it to match the data. In this region one finds 4.5 %

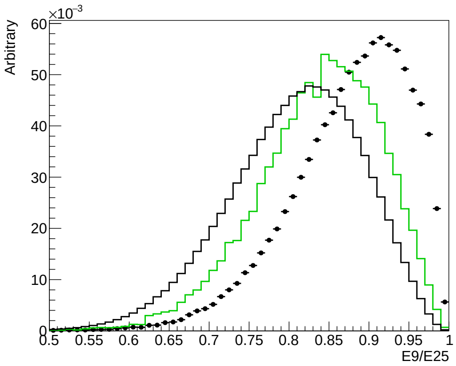


Figure 6.12: Distributions of E_9/E_{25} for MC anti-protons (green), MC anti-neutrons (black) and anti-protons in data (black points). Clusters with lab energies larger than 1 GeV are used.

of continuum and 78 % of $B\overline{B}$ events, without including the “others” component. The scaling factor is 1.30 ± 0.15, where the uncertainty arises from the available on- and off-resonance data statistics.

6.3 Cut efficiency corrections for π^0 and η background

In order to study the efficiency of the π^0 and η vetoes in data, appropriate control regions must be identified. The vetoes are applied in the selected control regions, and the veto efficiency for data and MC are determined. The ratio of selection efficiencies is then fitted using a polynomial of first order ($f(x) = p_0 + p_1 x$). The fit result is used to correct the MC simulation.

To study the π^0 veto the region $-0.25 < \mathrm{BDT} < 0.25$ is chosen. The results to the polynomial fit are $p_0 = 1.145 \pm 0.024$ and $p_1 = -0.061 \pm 0.014$, showing an energy dependence of the correction. The η veto is studied in a control sample obtained by reversing the π^0 veto, prob(π^0) > 0.75. This sample is almost entirely composed of photons from π^0 decays from $B\overline{B}$ events, and a small contribution from continuum. For the η veto the fit results are $p_0 = 1.000 \pm 5.3 \times 10^{-6}$ and $p_1 = (1.160 \pm 0.146) \times 10^{-5}$.

The results for this study are presented in Fig. 6.13. The small uncertainty on the fit parameters of the η veto correction reflects that there is no limitation from data or MC statistics in the sample, and that the fit describes the ratio of efficiencies very well. Consequently, this correction will be only a very small systematic uncertainty on the measurement of $\mathcal{B}_{s\gamma}$.

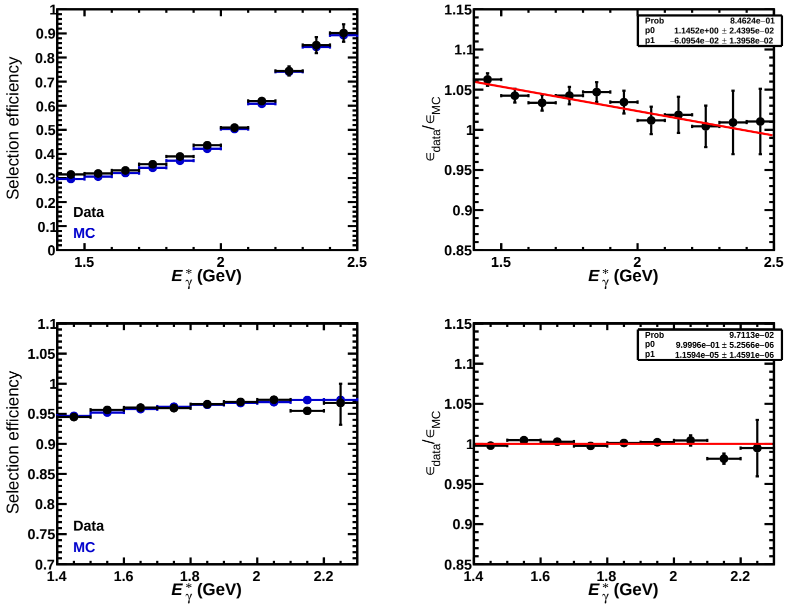


Figure 6.13: Efficiency and correction of the π^0 and η vetoes.

CHAPTER 7

Measurement of the $B \to X_{s+d}\gamma$ branching fraction and spectral moments

THERE are several interesting physical observables that can be studied to deliver information on the $B \to X_{s+d}\gamma$ decay such as the inclusive branching fraction. The shape of the photon energy spectrum can provide information that is not only specific to the $B \to X_{s+d}\gamma$ decay, namely the kinematics and properties of the quarks inside the B meson. The inclusive $B \to X_{s+d}\gamma$ branching fraction is calculated as:

$$\mathcal{B}^{E_\gamma \geq E_0}_{(s+d)\gamma} = \frac{1}{2N_{B\overline{B}}} \frac{1}{\varepsilon_{\mathrm{rec}}} \cdot \frac{\alpha^{E_\gamma \geq E_0}}{\varepsilon_{\mathrm{sel}}^{E^*_\gamma \geq E_0}} N^{E^*_\gamma \geq E_0}. \tag{7.1}$$

The superscript $E_\gamma \geq E_0$ ($E^*_\gamma \geq E_0$) indicates that a quantity is given for a photon energy threshold E_0 in the B (CM) frame. The integral of the photon energy spectrum (or total number of signal events) above the threshold E_0 up to 2.8 GeV is given by $N^{E^*_\gamma \geq E_0}$; the probability that a signal photon lies within the acceptance of the detector, and is found by the reconstruction algorithms is given by $\varepsilon_{\mathrm{rec}}$; the probability that a signal event in this photon energy region passes all selection criteria is given by the selection efficiency $\varepsilon_{\mathrm{sel}}^{E^*_\gamma \geq E_0}$; $N_{B\overline{B}}$ is the total number of $B\overline{B}$ events recorded by Belle, and is $(771.581 \pm 10.566) \times 10^6$; finally $\alpha^{E_\gamma \geq E_0}$ is a factor that transforms the branching fraction measured in the CM frame into the B frame. The full inclusive branching fraction is, by definition, frame-independent, however the integral of the CM-frame spectrum differs from the integral of B-frame spectrum when the same threshold is required in both frames. This is due to the Doppler broadening of the spectrum in the CM frame. All these factors are discussed in detail in the following, and the corresponding systematic uncertainties are presented.

In addition to the inclusive branching fraction, the partial branching fractions as a function of photon energy are calculated as:

$$(\Delta\mathcal{B}_{(s+d)\gamma})_i = \frac{1}{2N_{B\overline{B}}} \frac{1}{\varepsilon_{\mathrm{rec}}} \cdot \frac{1}{\varepsilon_{\mathrm{sel},\,i}} N_i. \tag{7.2}$$

In this equation, the subscript i represents the CM photon energy bin, $\varepsilon_{\mathrm{sel},\,i}$ is the selection efficiency on the bin i. To obtain the partial branching fractions, detector resolution effects are unfolded making use of Singular Value Decomposition Unfolding (SVDU) [91]† . Also, with the unfolded spectrum it is possible to calculate the mean and variance of the $B \to X_{s+d}\gamma$ photon energy spectrum. They correspond to the first and second moment of the spectrum and can be calculated as described in Eq. (2.34).

Finally, the HQE parameters m_b and μ_π^2 in the shape-function scheme are determined by folding the BLNP theoretical prediction and performing a fit to the background-subtracted photon spectrum. This is a novel approach since the HQE parameters are usually extracted from a combined global fit to the spectral moments of semileptonic decays and the $B \to X_s\gamma$ photon energy. With this approach one aims to obtain a more precise determination of the HQE parameters.

The results presented in this chapter are compared to state-of-the-art theoretical predictions, as well as the most recent measurements, with the aim of exploring possible new physics contributions in this decay.

7.1 Background subtraction and systematic uncertainties

7.1.1 background-subtracted spectrum

Having corrected the background components, the $B \to X_{s+d}\gamma$ spectrum is obtained by subtracting the background on a bin-by-bin basis as shown in Eq. (5.11). The background-subtracted spectrum is presented in Fig. 7.1, together with the statistical and systematic uncertainties and the prediction for the HQE parameters in the shape-function scheme that best fits the spectrum. The systematic uncertainties and fit are presented in the following sections. The statistical uncertainty on the spectrum comes from the numbers of events in the on-resonance and off-resonance samples, the latter carries the scaling factor discussed previously. The MC statistics are discussed as part of the systematic uncertainties and therefore not included here. In Table 7.1 the on the on- and off-resonance samples yields, the total $B\overline{B}$ background yield and the background-subtracted signal yield are presented together with their statistical uncertainties.

7.1.2 Systematic uncertainties from background subtraction

The corrections to the MC and off-resonance sample described in Chapter 6 give rise to several systematic uncertainties on the $B \to X_{s+d}\gamma$ spectrum and correlations among uncertainties on CM photon energy bins. The corrections to the major backgrounds, particularly from the π^0 component, give rise to very large systematic uncertainties, particularly at low photon energies where the $B\overline{B}$ background is dominant. Anti-correlation also exists in some high photon CM energy bins. This is because the correction to the off-resonance energies can cause migration of events between neighboring bins.

In addition to the corrections previously described, there are a few other sources of uncertainties:

---

† The Singular Value Decomposition algorithm is usually abbreviated SVD, this clashes however with the previous abbreviation used for the Silicon Vertex Detector, thus the choice SVDU.

Bin (E^*_γ)	N_{on}	N_{off}	N_{MC}	$N_{B\to X_{s+d}\gamma}$
1.4 – 1.5	9074.0	50.0	8588.9	109.3 ± 109.1
1.5 – 1.6	8809.0	57.1	8360.1	20.5 ± 109.7
1.6 – 1.7	8587.0	52.0	8035.6	160.6 ± 107.4
1.7 – 1.8	7988.0	52.0	7577.6	19.6 ± 104.5
1.8 – 1.9	7271.0	57.1	6768.7	73.9 ± 102.4
1.9 – 2.0	6415.0	58.1	5684.9	294.2 ± 98.4
2.0 – 2.1	5318.0	50.0	4234.7	707.5 ± 90.2
2.1 – 2.2	3851.0	81.1	2577.0	665.3 ± 91.8
2.2 – 2.3	3150.0	65.1	1326.7	1334.8 ± 82.6
2.3 – 2.4	2811.0	56.1	627.6	1762.5 ± 77.3
2.4 – 2.5	2459.0	45.0	330.1	1790.7 ± 70.7
2.5 – 2.6	1908.0	83.1	171.9	1112.3 ± 81.2
2.6 – 2.7	1196.0	75.1	89.9	542.4 ± 73.7
2.7 – 2.8	641.0	79.1	55.7	−8.5 ± 71.4
2.8 – 2.9	484.0	64.1	34.1	−31.1 ± 64.0
2.9 – 3.0	407.0	70.1	26.2	−145.3 ± 66.0
3.0 – 3.1	408.0	53.0	19.7	−10.0 ± 58.3
3.1 – 3.2	368.0	35.0	14.0	90.9 ± 48.4
3.2 – 3.3	328.0	31.0	12.5	82.5 ± 45.6
3.3 – 3.4	282.0	33.0	9.1	24.9 ± 46.3
3.4 – 3.5	208.0	17.0	5.3	75.0 ± 34.2
3.5 – 3.6	155.0	28.0	4.2	−59.7 ± 41.7
3.6 – 3.7	137.0	12.0	3.9	42.9 ± 28.5
3.7 – 3.8	109.0	11.0	1.9	24.5 ± 27.0
3.8 – 3.9	73.0	13.0	1.5	−26.2 ± 28.4
3.9 – 4.0	59.0	4.0	1.1	27.8 ± 16.9

Table 7.1: Number of events in the different samples and in the background subtracted sample, with statistical uncertainties

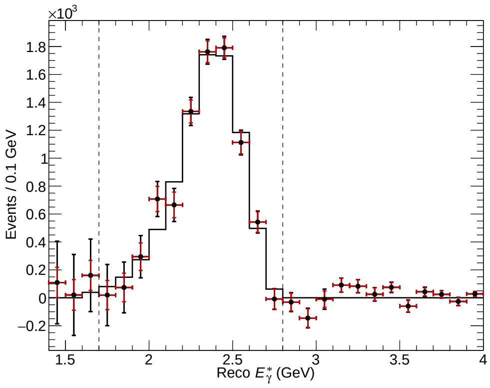


Figure 7.1: background-subtracted $B \to X_{s+d}\gamma$ photon energy spectrum. The inner error bars are statistical, the outer bars contain also the systematic uncertainties from background subtraction. The solid line shows the best fit in the shape-function scheme.

- MC statistics: the statistical uncertainty coming from 10 streams of generic $B\overline{B}$ MC, 50 streams of rare MC and 20 streams of $B \to X_u \ell \nu$ is taken into account as a systematic uncertainty.
- Beam background yield: beam background is taken from a random data trigger and mixed into the MC. A conservative ±20% on its yield is assigned as systematic uncertainty.
- Background from electron clusters without matched track: ±20% on its yield is assigned as systematic uncertainty.

The covariance and correlation matrices between the background-subtracted photon CM energy bins related to all these effects are shown in Appendix A. The correlation matrix for systematic uncertainties is given in Table A.7.

7.1.3 Photon detection efficiency and photon energy resolution

The photon detection efficiency and energy resolution have been studied in [92] using radiative $e^-e^+ \to \mu^+\mu^-\gamma$ events. This study aims to calibrate high-energy photons, with lab energies greater than 1.50 GeV. Four-momentum conservation implies that the four-momentum of the radiated photon is given by:

$$P_\gamma = P_{e^+} + P_{e^-} - P_{\mu^+} - P_{\mu^-}. \tag{7.3}$$

This yields a precise measurement of the photon kinematics, given that the initial state e^-e^+ is precisely known and the momenta of the muons can be very well measured. In an ideal scenario, the difference between measured photon energy and photon momentum inferred from track measurements would be negligible (i.e. resolution would be perfect). The limited resolution of the calorimeter causes, however, a broadening of the measured photon energy spectrum. Table 7.2 shows the mean and standard deviation of the resolution for data and MC as a function of the photon energy in the lab frame. The difference between the means in data and MC is taken as the systematic uncertainty in the measurement.

Energy (in lab frame)	Mean (MeV)		Standard deviation (MeV)	
	Data	MC	Data	MC
1.00–1.25	-2.3	-3.3	35.9	31.1
1.25–1.50	-1.7	-3.9	39.1	33.8
1.50–1.75	1.5	-3.3	40.2	35.9
1.75–2.00	1.5	-3.9	42.9	37
2.00–2.25	4.2	-3.9	43.4	38.6
2.25–2.50	4.2	-3.3	46.1	38.1
2.50–2.75	4.2	-3.3	44.5	39.1
2.75–3.00	10.6	-5.5	46.1	42.4
3.00–3.25	13.9	-1.7	46.7	41.3
3.25–3.50	13.9	-4.4	48.3	44
3.50–3.75	14.4	-2.3	49.4	45.1
3.75–4.00	11.2	-7.6	51	47.7
4.00–4.25	7.4	-7.6	52	49.9
4.25–4.50	3.1	-6	54.7	50.4
4.50–4.75	3.7	-6	55.3	53.7
4.75–5.00	1.5	-4.4	56.9	56.3
5.00–6.20	-4.4	-6	62.3	60.1

Table 7.2: Mean and standard deviation of the photon energy resolution for data and MC.

Similarly, the same study presents photon selection efficiency in bins of laboratory energy and polar angle. All values can be found in Tables 7 and 8 of [92], the tables are not reproduced here. The average efficiency for photons in the final spectrum can be calculated in both data and MC:

$$\varepsilon_{\mathrm{MC}} = 0.920 \pm 0.002$$
$$\varepsilon_{\mathrm{Data}} = 0.925 \pm 0.005,$$

Once again, the difference between data and MC, which is around 0.5%, is taken as a systematic uncertainty. The covariance and correlation matrices for both these effects are shown in Appendix A.

7.2 Reconstruction and selection efficiencies

The reconstruction and selection efficiencies quantify how likely it is for a produced signal photon to be found in the detector and be accepted as a candidate after applying all the selection criteria.

These quantities are estimated using MC samples and are defined as:

$$\varepsilon_{\mathrm{rec}} = \frac{N_{\mathrm{rec}}}{N_{\mathrm{gen}}} \tag{7.4}$$

$$\varepsilon_{\mathrm{sel}}^{E_\gamma^* \geq E_0} = \frac{\sum\limits_i N_{\mathrm{pass},\,i}}{\sum\limits_i N_{\mathrm{rec},\,i}} \tag{7.5}$$

$$\varepsilon_{\mathrm{sel},\,i} = \frac{N_{\mathrm{pass},\,i}}{N_{\mathrm{rec},\,i}}, \tag{7.6}$$

where N_{gen} represents the total number of generated signal events, N_{rec} is the number of generated signal events that have been found and reconstructed in as ECL clusters. $N_{\mathrm{rec},\,i}$ is the number of signal events found and reconstructed in the i^{th} CM photon energy bin and $N_{\mathrm{pass},\,i}$ is the number of signal events passing all selection criteria found in the same bin. The sum is performed for all bins between the threshold energy E_0 and 2.8 GeV. As pointed out previously, $\varepsilon_{\mathrm{sel}}^{E_\gamma^* \geq E_0}$ is used in the measurement of the inclusive branching fraction, where the spectrum is integrated from a certain energy threshold.

The efficiencies $\varepsilon_{\mathrm{sel}}^{E_\gamma^* \geq E_0}$ and $\varepsilon_{\mathrm{sel},\,i}$ are slightly model-dependent due to the inclusion of event shape variables in the continuum suppression BDT. Making the E_γ^* spectrum slightly harder or softer, would affect both the event multiplicity and the spatial distribution of the particles from the X_s decay. A second observation, is that both selection efficiencies show a strong energy dependence. This effect is not unexpected, it is known that in an inclusive analysis where continuum and $B\overline{B}$ background are overwhelmingly large at low energy, a loss of efficiency is unavoidable. Both these findings are not true for $\varepsilon_{\mathrm{rec}}$, which is energy- and model-independent, and is basically determined by the detector geometry and the reconstruction algorithms.

In Fig. 7.2, $\varepsilon_{\mathrm{rec}}$ for signal photons is presented in bins of true E_γ^B and $\varepsilon_{\mathrm{sel},\,i}$ in bins of reconstructed E_γ^*. The efficiency $\varepsilon_{\mathrm{rec}}$ is perfectly described by a constant with value $\varepsilon_{\mathrm{rec}} = 0.7121 \pm 0.0003$, where the uncertainty comes from MC statistics, the uncertainty reflects only the large size of the generated signal MC samples and will only play a minor role as a systematic uncertainty. The efficiency $\varepsilon_{\mathrm{sel},\,i}$ is presented for the pair of m_b and μ_π^2 that best fit the spectrum, for which results are presented in the next section. The uncertainties on $\varepsilon_{\mathrm{sel},\,i}$ also come from MC statistics. In Table B.1 the values for $\varepsilon_{\mathrm{sel},\,i}$ and $\varepsilon_{\mathrm{sel}}^{E_\gamma^* \geq E_0}$ are summarized.

7.2.1 BDT modeling

The BDT modeling for $B\overline{B}$ events is studied in the π^0 control sample. Given that the off-resonance data models the continuum background, one still has to assess if the $B\overline{B}$ MC properly describes the BDT distribution in data. This is studied in two ways: looking at the difference between the selection efficiency in data and MC in a control sample, and calculating the χ^2 between the data and MC BDT distributions in the signal region.

For the first test, the efficiency of the BDT cut in bins of E_γ^* is calculated. Off-resonance data is subtracted from the on-resonance sample to be sure only $B\overline{B}$ events are studied. Fig. 7.3 shows the BDT distributions for the control sample and the signal region. The ratio between data

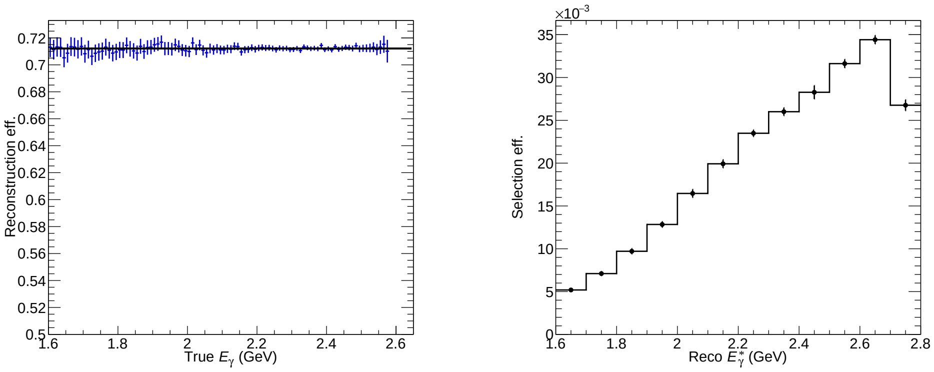


Figure 7.2: Reconstruction (left) and selection efficiencies $\varepsilon_{\mathrm{sel},\,i}$ (right) for the best fit to data in the shape function scheme. The error bars show the uncertainties from MC statistics.

and MC distributions has a uniform distribution, but a more precise assessment of the matching between data and MC is also presented here.

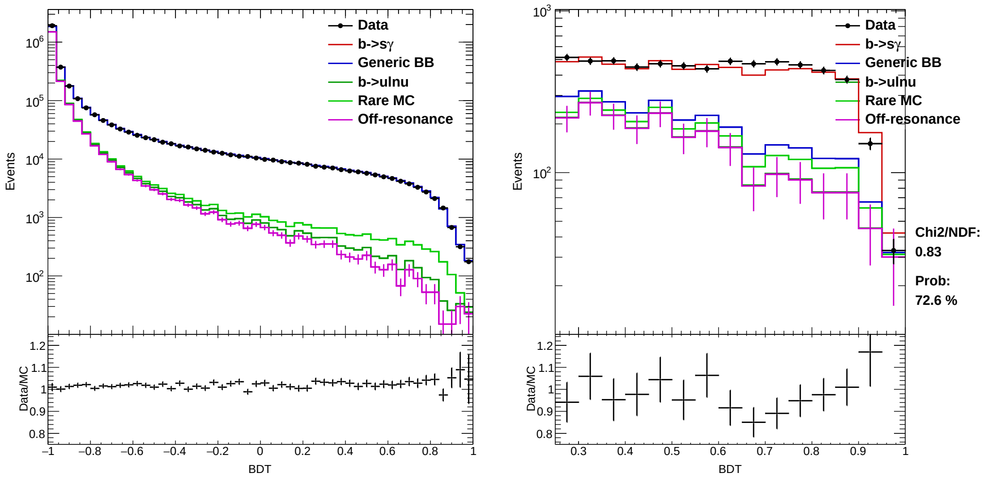


Figure 7.3: BDT distribution in the π^0 control sample (left) and in the signal region (right). The ratio below shows the ratio between data and MC, and the error bars reflect the statistical uncertainties.

The efficiency of the cut BDT $>$ 0.25 for data and MC in the control sample, as well as the difference between data and MC are presented in Fig. 7.4. The plot extends only up to 2.5 GeV because the $B\overline{B}$ contributions vanish beyond that point. The difference plot is fitted using polynomials of zeroth, first and second order to estimate the systematic uncertainty related to the BDT selection. Although all three functions fit the data well, the polynomial of degree one is finally used since it yields information on the energy dependence of the difference and the fit parameters are more precisely determined than the degree two polynomial. The fit results for the three fit functions are shown below, and the corresponding correlation and covariance matrices for the polynomial of first order presented in Appendix A.

- Polynomial 0th order: $p_0 = 0.018 \pm 0.008$
- Polynomial 1st order: $p_0 = -0.151 \pm 0.113$, $p_1 = 0.094 \pm 0.062$
- Polynomial 2nd order: $p_0 = -0.361 \pm 1.371$, $p_1 = 0.324 \pm 1.493$, $p_2 = -0.062 \pm 0.404$

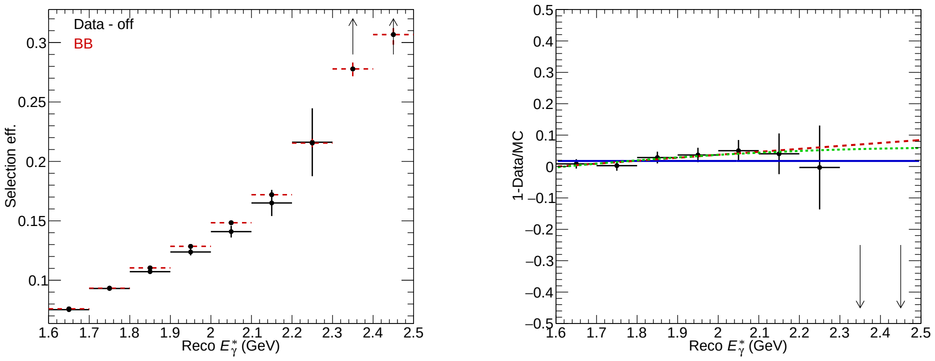


Figure 7.4: Efficiency of the cut BDT > 0.25 and difference $1 - \varepsilon_{\mathrm{Data}}/\varepsilon_{\mathrm{MC}}$ as a function of E^*_γ. The bins outside the range of the plot are indicated by arrows, the statistical uncertainty on these bins is very large and the difference on them is consistent with zero. The difference plot is fitted using polynomials of order zero (blue), one (red) and two (green).

As an additional check of the BDT modeling by the MC, the BDT distribution was studied in bins of CM photon energy in the region 1.8 to 2.8 GeV. The plots are show in Appendix D. The χ^2 is calculated for the BDT distribution before and after applying the BDT selection cut. The results are shown in Table 7.3. The χ^2/NDF is calculated between the on-resonance distribution and the sum of off-resonance and all MC samples. This is done to quantify the agreement between data and MC. For all E^*_γregions it is close to 1, confirming that the MC describes the BDT properly.

	−1. <BDT< 1. NDF = 99			0.25. <BDT< 1. NDF=37		
Bin (E^*_γ)	χ^2	χ^2/NDF	Probability (%)	χ^2	χ^2/NDF	Probability (%)
1.8 – 2.0	134.1	1.4	1.1	23.0	0.6	96.5
2.0 – 2.2	147.1	1.5	0.1	31.6	0.9	72.0
2.2 – 2.4	114.0	1.2	14.4	44.5	1.2	18.5
2.4 – 2.6	106.1	1.1	29.5	41.7	1.1	27.4
2.6 – 2.8	91.5	0.9	69.1	71.8	1.9	0.1

Table 7.3: χ^2 and χ^2/NDF for the BDT distribution in bins of E^*_γ. The values are around 1, confirming that the description of this variable by the MC is correct.

7.2.2 Systematic effects from PYTHIA settings

In the inclusive $B \to X_s\gamma$ decay, the hadronization of the inclusive X_s part is simulated using PYTHIA [93]. Modifying the settings of this MC tool changes the composition of the sample and thus affects the particle multiplicity of the X_s decays. The particle multiplicity affects the event shape variables and consequently the BDT distribution. A dedicated study performed by Belle in the semi-inclusive $B \to X_s\gamma$ analysis [53] served as a starting point for this test. In this analysis the PYTHIA parameters described below are tuned such that the data is best described by the MC, and several samples with various settings are produced to study the effects on reconstruction efficiencies and the fraction of X_s modes that are not reconstructed by the analysis. Uncertainties arising from the PYTHIA settings are is dominant in the sum-of-exclusives analysis, but are expected to be secondary for this analysis since it does explicitly reconstruct hadronic final states and missing modes must not be estimated.

Five parameters are tested and modified and are described below. The 18 MC samples that are used are described in detail in Table 7.4 and were generated by the authors of [53]. The selection efficiencies $\varepsilon_{\mathrm{sel},\,i}$ and $\varepsilon_{\mathrm{sel}}^{E_\gamma^* \geq E_0}$ are determined on each sample, and the difference with respect to the default value given in Table B.1, is taken as the systematic uncertainty. The systematic effect is of order 1 % on $\varepsilon_{\mathrm{sel},\,i}$ and can be seen in Fig. 7.5. The corresponding correlation and covariance matrices are presented in Appendix A. The five PYTHIA parameters are:

- PARJ(2) is the suppression of s quark pair production compared to u or d quark pair production,
- PARJ(11) is the probability that a light meson (containing u and d quarks only) has spin 1,
- PARJ(12) is the probability that a strange meson has spin 1,
- PARJ(15) is the probability that a spin 1 meson is produced with an orbital angular momentum 1, for a total spin of 1,
- PARJ(25) is the extra suppression factor for η production in the hadronization.

7.3 Determination of heavy quark expansion parameters

The HQE parameters m_b and μ_π^2 are extracted for the shape-function scheme by performing a fit to the background-subtracted $B \to X_{s+d}\gamma$ spectrum. There are two assumptions made in this study, the first is that any resonant structure in the $B \to X_s\gamma$ spectrum is smeared out sufficiently by detector resolution and the choice of CM frame, such that an inclusive description of the spectrum is appropriate. The second assumption is that the $B \to X_d\gamma$ and $B \to X_s\gamma$ spectra follow the same shape, making the description of the data spectrum using only the inclusive $B \to X_s\gamma$ theory appropriate. This is a reasonable assumption, given that both are two-body decays, and $B \to X_d\gamma$ is strongly suppressed with respect to $B \to X_s\gamma$.

The comparison of the theoretical spectrum with the experimental one is made possible by applying a “folding” procedure, in which all detector resolution and Doppler broadening effects,

SAMPLE	PARJ(2)	PARJ(11)	PARJ(12)	PARJ(15)	PARJ(25)
Default	0.3	0.5	0.6	0.05	1
Tuned [53]	0.1	0.95	0.7	0.25	0.03
Sample 1		0.95		0.3	0.5
Sample 2	0.1	0.95		0.3	0.03
Sample 3		0.95		0.3	
Sample 4	0.1	0.95		0.3	0.5
Sample 5	0.1	0.95		0.3	0.03
Sample 6	0.1	0.95		0.3	
Sample 7	0.1	0.95		0.2	0.03
Sample 8	0.1	0.95	0.7	0.4	0.03
Sample 9	0.1	0.95			
Sample 10	0.1	0.95		0.25	0.03
Sample 11		0.95			
Sample 12	0.1	0.95		0.25	0.5
Sample 13			0.7		
Sample 14	0.1	0.95		0.2	0.03
Sample 15				0.1	
Sample 16	0.1	0.95	0.7	0.3	0.03

Table 7.4: PYTHIA setup for the different signal MC samples used. The “tuned” values have been determined in the a different Belle analyisis [53]. Empty cells mean that the default value is used.

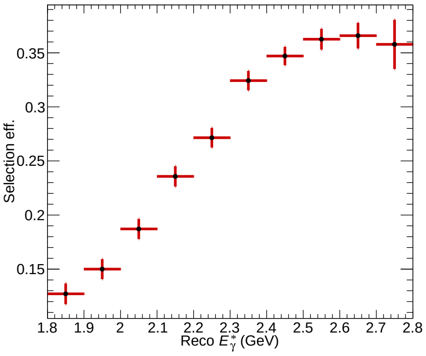


Figure 7.5: Selection efficiency $\varepsilon_{\mathrm{sel},i}$, the error bars represent the systematic uncertainty from the PYTHIA settings.

as well as selection and reconstruction efficiencies, are simulated in the theoretical spectrum. Spectra using the shape-function theoretical prediction (BLNP model) are generated and MC information used to transform the generated photon spectrum from the B rest frame into the CM rest frame. As a second step, resolution effects caused by the detector are introduced. Finally each E^*_γ bin is corrected with $\varepsilon_{\mathrm{rec}}$ and $\varepsilon_{\mathrm{sel},\,i}$. A χ^2 fit between the theoretical spectra and the background-subtracted spectrum is performed using the full experimental covariance matrix. In the fit m_b and μ_π^2 are free parameters and the pair of them that yields the minimum χ^2 is found. The fit is performed in the E^*_γ range 1.8 to 2.8 GeV.

The folding matrices used for transforming to the CM frame and to account for detector resolution are shown in Fig. 7.6, the spectra obtained after each folding step are compared in Fig. 7.7.

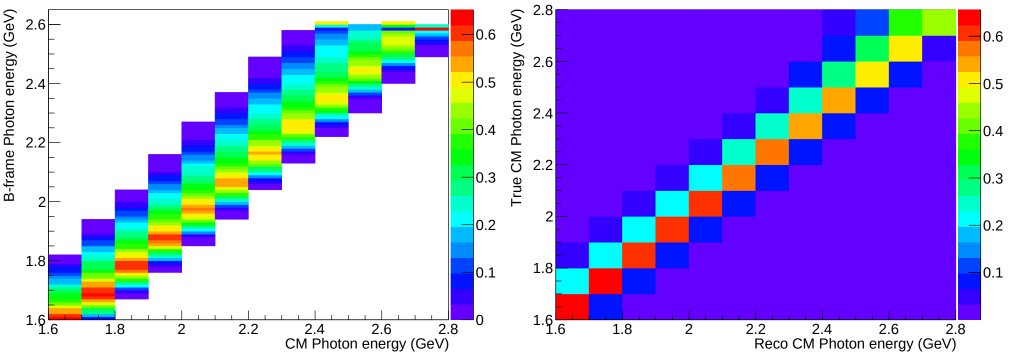


Figure 7.6: Folding matrices: matrix to convert the B frame spectrum into the CM spectrum (left) and matrix to fold in the detector resolution effects (right).

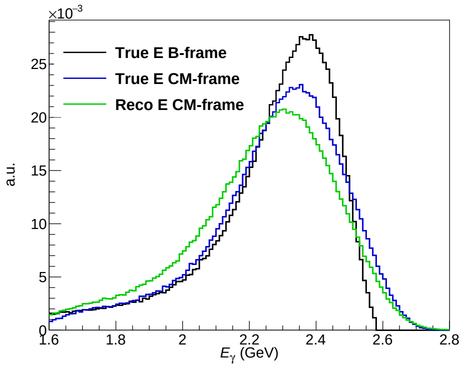


Figure 7.7: Spectra in the B frame, CM frame and CM frame with detector resolution effects. The latter two spectra are derived from the B frame spectrum (in black) using the folding matrices.

The fit parameters are varied in steps of size $O(10^{-5})$. It is technically impossible to calculate the spectrum from the analytical expressions on each step, therefore one has to rely on an interpolation. Spectra are generated in the ranges of $4.400 \leq m_b \leq 4.795\,\mathrm{GeV}$ in steps of

0.05 GeV and $0.005 \leq \mu_\pi^2 \leq 0.690\,\text{GeV}^2$ in steps of $0.05\,\text{GeV}^2$. A fine grid of theoretical spectra is necessary to obtain a smooth interpolation, which in turn means the χ^2 distribution is smooth, the global minimum can be properly found and the fit uncertainties properly estimated. The two-dimensional interpolation depicted in Fig. 7.8, consists of finding the three closest generated spectra to the point that is been tested, and for each bin of the spectrum, find the plane that is spanned by the three points. The interpolated point is extracted from the plane that was found.

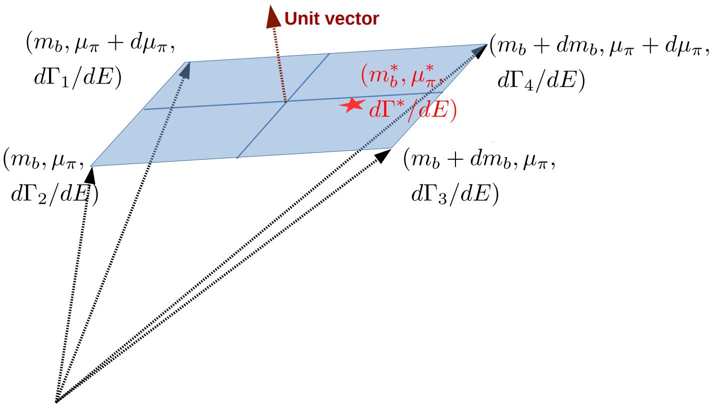


Figure 7.8: Interpolation of generated spectra. The plane spanned by the three closest generated points to the point of interest (in red) is found. The value of the interpolation is found on this plane. This step is performed for each bin of the spectrum.

In order to assess the fit errors from the default fitter, one-dimensional χ^2 scans and two-dimensional contour plot are shown in Figs. 7.9 and 7.10. The default fit result with symmetric parabolic errors is $m_b = (4.627 \pm 0.029)\,\text{GeV}$ and $\mu_\pi^2 = (0.301 \pm 0.063)\,\text{GeV}^2$. However, the assumption that the errors are symmetric does not hold for the one-dimensional scans. They cannot be fitted using a polynomial of second order, and are better described by a polynomial of third order.

The fit results, both for asymmetric and symmetrized uncertainties, are summarized in Table 7.5, where also the latest HFAG fit values [27] are quoted. The HFAG result shows the value obtained from a fit to the moments of semileptonic decay spectra in the kinetic scheme. This result is translated into the shape-function scheme, which is the source of the second uncertainty quoted in the table. One could think of a few reasons why the obtained result is more precise: the spectrum has been extracted with improved statistical and systematic uncertainties, due to a large data sample, as well as improved background suppression techniques, which is reflected by the fit uncertainties. This is the first time the folding approach is used to determine the HQE parameters. If the measured spectrum were to be unfolded to correct for detector resolution, and corrected for the selection efficiency, additional uncertainties and model dependencies would arise. Usually, the HQE parameters have not been directly obtained from a fit to the spectrum, but rather to the measured spectral moments.

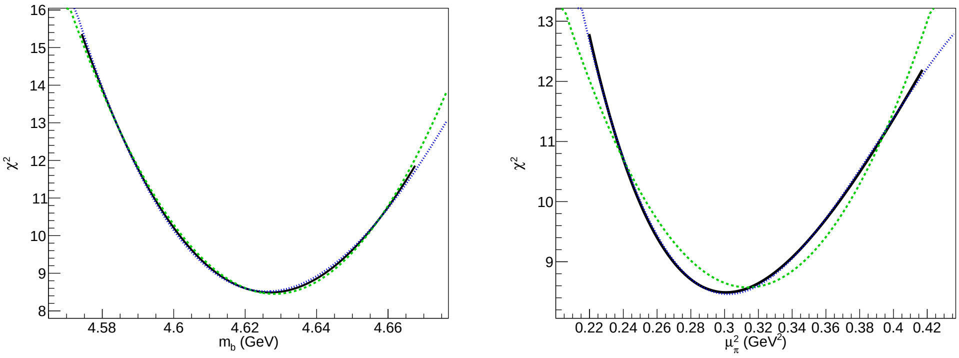


Figure 7.9: One dimensional χ^2 scan for m_b and μ_π^2. The χ^2 scans are fitted with polynomials of order 2 (in green) and 3 (in blue) to assess if the errors are symmetric, the order 2 fit cannot describe the curves.

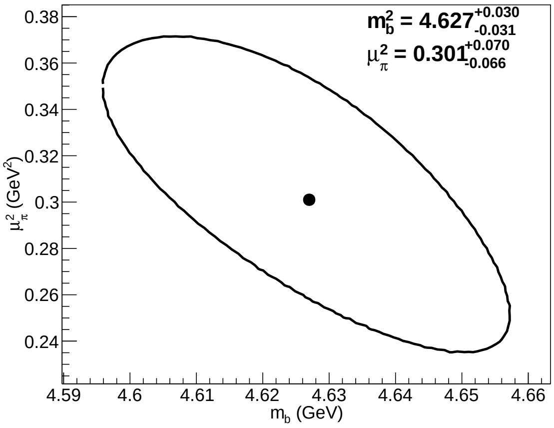


Figure 7.10: HQE fit result and 1 σ fit error contours.

Parameter	This result	HFAG
m_b (GeV)	$4.627 \pm 0.029 \left(\begin{smallmatrix}+0.030\\-0.031\end{smallmatrix}\right)$	$4.569 \pm 0.023 \pm 0.018$
μ_π^2 (GeV^2)	$0.301 \pm 0.063 \left(\begin{smallmatrix}+0.070\\-0.066\end{smallmatrix}\right)$	$0.145 \pm 0.089^{+0.020}_{-0.040}$
Correlation	-0.701	—
Fit χ^2	11.09	—

Table 7.5: Extracted values for the HQE parameters m_b and μ_π^2, the symmetric uncertainties correspond to the parabolic error estimation, while the asymmetric uncertainties in parentheses can be read out form the contour plot. The last column shows the latest HFAG fit results. The results are translated from a fit in the kinetic scheme, the first uncertainty comes from the fit and the second from the scheme translation.

7.3.1 Fit validation with toys

In order to validate the fitting procedure, the fit was repeated on a set of 2000 toy spectra. Each toy was generated from the photon energy distribution obtained with the default parameters $m_b = 4.60\,\mathrm{GeV}$ and $\mu_\pi^2 = 0.30\,\mathrm{GeV}^2$ by including Gaussian fluctuations for each bin content with a size derived from the experimental covariance matrix. For each toy spectrum, the pull of each fit parameter x is calculated. The pull σ_x is defined as the difference between the fit value and the default parameter value divided by the fit error σ_x $p_x = \frac{x_{\text{default}}-x_{\text{fit}}}{\sigma_x}$. The pull distributions are expected to be described by a Gaussian of mean zero and width 1. The results of this study are presented in Fig. 7.11.

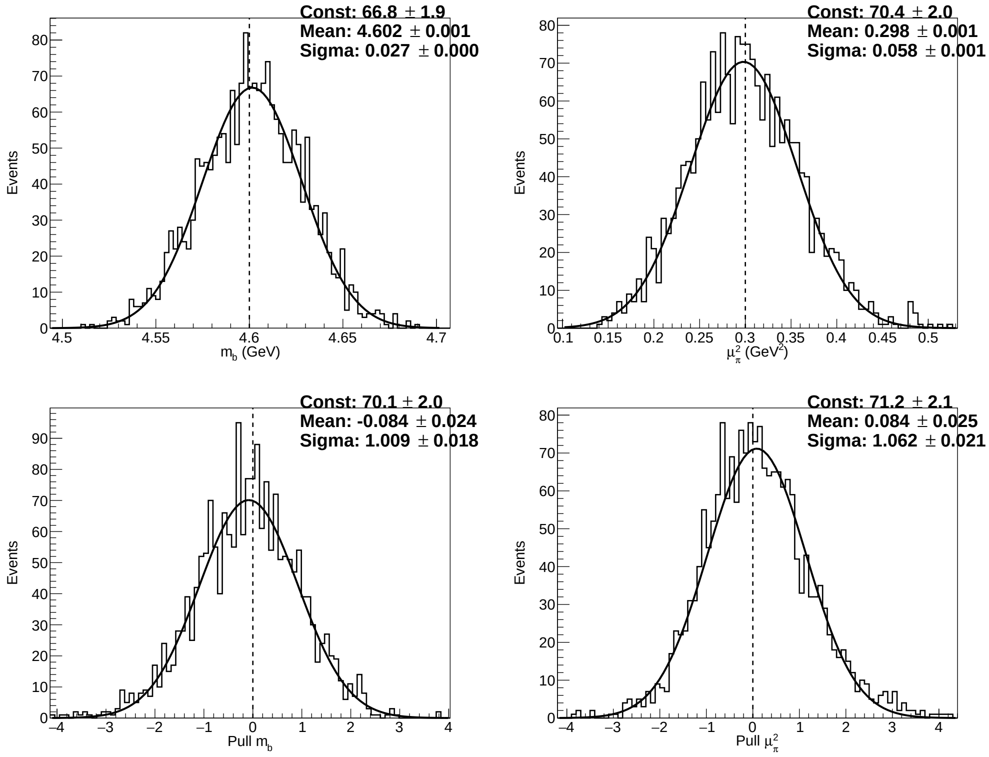


Figure 7.11: Distributions of the m_b and μ_π^2 parameters for 2000 toy spectra and pulls on them. The upper two plots show the distributions of m_b and μ_π^2, whose means agree with the default values of $m_b = 4.60\,\mathrm{GeV}$ and $\mu_\pi^2 = 0.30\,\mathrm{GeV}^2$. The bottom plots are the pulls. The bias on the mean of the pulls is not a concern since the default values of the parameters are reproduced within $O(2 \times 10^{-3})$, it is a result of the small uncertainties on the fit.

Both pull distributions show small deviations from the mean of zero and width of one. The fitted values of m_b and μ_π^2 for the toy spectra agree with the default values of $m_b = 4.60\,\mathrm{GeV}$ and $\mu_\pi^2 = 0.30\,\mathrm{GeV}^2$ within $O(2 \times 10^{-3})$. This agreement is much smaller than the fit uncertainties, but is source to the slight deviations on the pull mean. The reason for the deviations on the pull this is that the pull is ill-defined for asymmetric uncertainties. As observed in Table 7.5, the symmetrized uncertainties obtained from the fit are smaller than the average of the lower and upper values of the obtained asymmetric uncertainties from the scan. As a test, the fit error was changed by ±5%, and the results are compatible with a width of one as seen in Fig. 7.12.

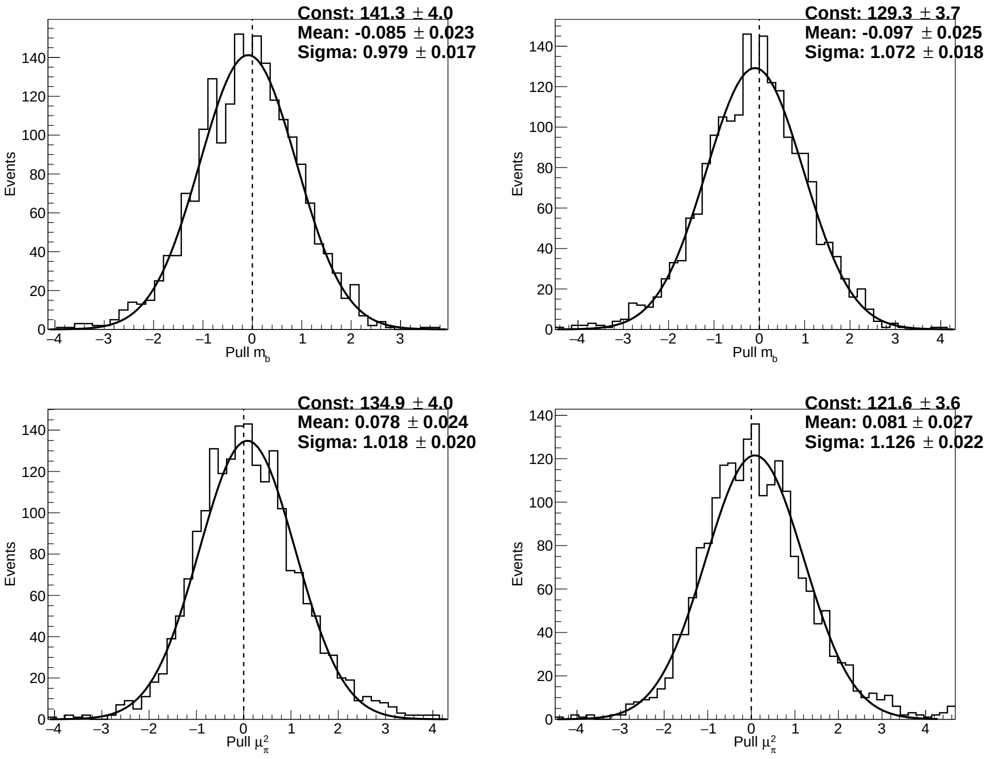


Figure 7.12: Pull distributions for the m_b (top row) and μ_π^2 (bottom row) parameters changing the fit error by +5% (left) and −5% (right) to account for asymmetric uncertainties.

7.3.2 $B \to X_d\gamma$ contribution

The decay $B \to X_d\gamma$ has not been as thoroughly studied as $B \to X_s\gamma$, because its branching fraction is much smaller. Belle studied the resonant final states $B^{\pm,0} \to \rho^{\pm,0}\gamma$ and $B^0 \to \omega\gamma$ [56] and *BABAR* performed a sum of exclusives analysis [28]. The measured branching fractions are shown in Table 3.2. The $B \to X_d\gamma$ resonant contributions are compared to the measured spectrum in Fig. 7.13. In the figure, the resonant $K^*\gamma$ contribution is presented for comparison, and it can be seen that it is much larger than the other contributions.

To evaluate the impact of the $B \to X_d\gamma$ contribution, the $B \to X_d\gamma$ resonances are subtracted from the measured spectrum, and the fit performed again. The changes in the parameters are negligible, and within fit uncertainties: $m_b = (4.628 \pm 0.029)\,\mathrm{GeV}$, $\mu_\pi^2 = (0.282 \pm 0.062)\,\mathrm{GeV}^2$, with a correlation $\rho = -0.724$. No uncertainty related to this contribution was assigned in the fit result, mainly because of the lack of knowledge about the total composition of the $B \to X_d\gamma$ contribution, hence assigning an uncertainty based only on the $B \to X_d\gamma$ resonances would be arbitrary.

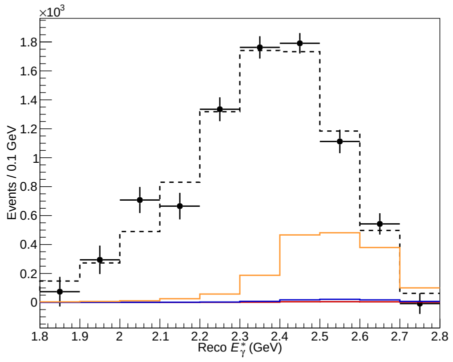


Figure 7.13: Background-subtracted spectrum showing the inclusive $B \to X_s\gamma$ and $B \to X_s\gamma$ and $B \to X_d\gamma$ resonances. The orange component corresponds to $K^*\gamma$, the red component to $\omega\gamma$ and the blue component to $\rho^{\pm,0}\gamma$.

7.3.3 HQE parameters from the Kagan-Neubert and kinetic scheme models

Not to rely on a single theoretical description of the $B \to X_s\gamma$ model for the determination of several quantities relevant for the following measurements, the Kagan-Neubert (KN) model and the kinetic scheme (BBU) models are also considered to find the values of m_b and μ_π^2 that best describe the spectrum. When using different schemes, the HQE parameters are not expected to yield the same values, since the treatment of non-perturbative QCD effects is different and different shape function assumptions are used, but the aim is to find the best description of the data for each model.

Unlike the involved procedure used for the shape-function model, only a χ^2 scan for these two models is performed. Sets of spectra for several values of m_b and μ_π^2 were generated, the χ^2 for each one was calculated, and the minimum value found. There is no need for an extremely precise determination of these parameters since we use a dense enough lattice of templates to perform the scan, such that variations from one point to its neighbor are very small. In addition to m_b and μ_π^2, BBU uses the chromomagnetic operator μ_G^2, which takes a value of $0.35\,\mathrm{GeV}^2$ [94]. The results are displayed in Fig. 7.14. For the kinetic scheme one obtains $m_b = 4.62\,\mathrm{GeV}$ and $\mu_\pi^2 = 0.50\,\mathrm{GeV}^2$ and for KN $m_b = 4.69\,\mathrm{GeV}$ and $\mu_\pi^2 = 0.61\,\mathrm{GeV}^2$.

The determination of the best parameter values for different models is not only important to exercise on a new way to measure them. The three models here determined are crucial to properly determine the selection efficiency and corrections and extrapolation factors used in the measurement of the inclusive branching fraction. Additionally, uncertainties arising from model dependence are calculated based on them.

7.4 Inclusive $B \to X_s\gamma$ branching fraction

Since the selection efficiency shows a dependency on the photon energy, a more precise result on the $B \to X_s\gamma$ branching fraction can be achieved by counting the total number of events for a broad energy range and calculating the selection efficiency for the same range instead of using

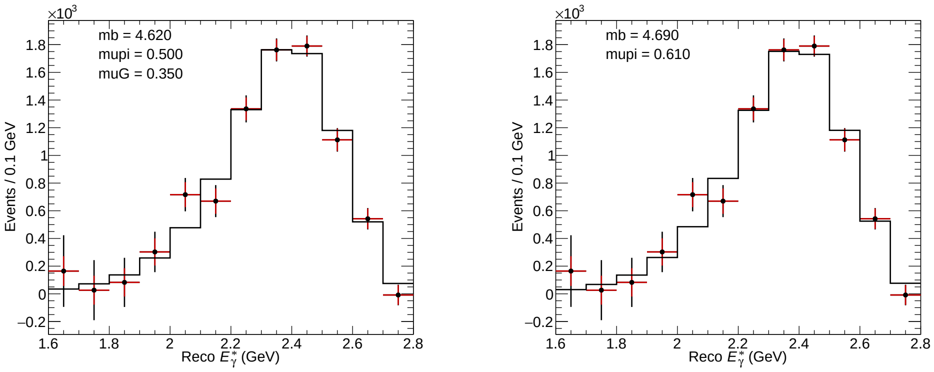


Figure 7.14: Spectra obtained from best values of HQE parameters for kinetic scheme and KN models that describe the data spectrum. The plot corresponding to the kinetic scheme is shown on the left and the plot corresponding to the KN model on the right.

the bin-by-bin efficiency. The branching fraction for a given energy range can be calculated as shown in Eq. (7.1). In this inclusive analysis, the $B \to X_d\gamma$ admixture has to be corrected in order to measure $\mathcal{B}_{s\gamma}$, and an extrapolation factor to the 1.6 GeV threshold is necessary. The branching fraction is then:

$$\mathcal{B}_{s\gamma} = \alpha^{E_0}_{\text{extrap}} \cdot \frac{1}{\alpha_{d\gamma}} \cdot \mathcal{B}^{E_\gamma \geq E_0}_{(s+d)\gamma}, \tag{7.7}$$

The correction factor is proportional to the ratio of CKM matrix elements that dominate the penguin loop and dictate the ratio of both branching fractions $\alpha_{d\gamma} = 1 + |\mathrm{V}_{td}\,/\,\mathrm{V}_{ts}|^2$. This extrapolation factor is discussed in the following.

7.4.1 Systematic uncertainty in the branching fraction from $N_{B\bar{B}}$

The number of $B\bar{B}$ events was determined by Belle to be $N_{B\bar{B}} = (771.851 \pm 10.566) \times 10^6$. The precision on this number of 1.4% enters as a systematic uncertainty in the measured branching fraction.

7.4.2 Systematic uncertainty in the branching fraction from the $B \to X_d\gamma$ admixture

The $B \to X_d\gamma$ component is suppressed by a factor $|\mathrm{V}_{td}\,/\,\mathrm{V}_{ts}|^2$ with respect to $B \to X_s\gamma$. As indicated is Eq. (7.7), the $\mathcal{B}_{(s+d)\gamma}$ is corrected by a factor $1 + |\mathrm{V}_{td}\,/\,\mathrm{V}_{ts}|^2$ to measure $\mathcal{B}_{s\gamma}$. The ratio of $|\mathrm{V}_{td}|$ and $|\mathrm{V}_{ts}|$ has been measured to be $|\mathrm{V}_{td}\,/\,\mathrm{V}_{ts}| = 0.216 \pm 0.011$. The correction is thus $1 + |\mathrm{V}_{td}\,/\,\mathrm{V}_{ts}|^2 = 1.047 \pm 0.007$ The uncertainty of 0.7 % contributes to the systematic uncertainties.

7.4.3 Selection efficiency

The selection efficiencies $\varepsilon_{\mathrm{sel},i}$ and $\varepsilon_{\mathrm{sel}}^{E_\gamma^* \geq E_0}$ as defined in Eq. (7.6) are determined using the best fits to the shape-function scheme, kinetic scheme and Kagan-Neubert models. The central value of the efficiency is the average of the efficiencies for each model, and the model uncertainty is taken as the largest difference between the 1σ fit error and the central value. In this way one can determine the selection efficiency in a bin-by-bin basis and also for the integrated spectrum for different thresholds. Both efficiencies are presented in Fig. 7.15 and summarized in Table B.1 together with the uncertainty due to model dependence.

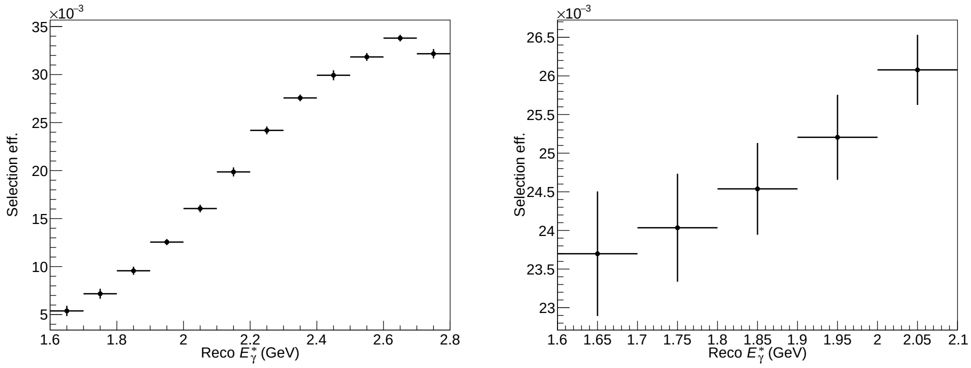


Figure 7.15: Selection efficiency for each bin ($\varepsilon_{\mathrm{sel},i}$) on the left and for a the integrated spectrum from a given E_γ^* threshold $\varepsilon_{\mathrm{sel}}^{E_\gamma^* \geq E_0}$ on the right. The error bars represent model uncertainties.

7.4.4 Conversion factors to the B rest frame

The factors $\alpha^{E_\gamma \geq E_0}$ and their uncertainties are, similarly to the selection efficiencies, determined from the ensemble of three theoretical models. As pointed out before on Eq. (7.1), they transform the branching fraction measured in the CM frame into the B rest frame, for the same energy threshold. The values with their respective model uncertainties are presented in Fig. 7.16 and summarized in Table B.1.

7.4.5 Extrapolation factors

Theoretical predictions for $\mathcal{B}_{s\gamma}$ are given for the 1.6 GeV threshold. Extrapolation factors to transform the measured branching fractions as shown in Eq. (7.7) are calculated from the chosen threshold down to 1.6 GeV ($\alpha_{\mathrm{extrap}}^{E_0}$). The central values and model uncertainties are determined from three models, in the same way as the selection efficiencies. The values with their respective model uncertainties are presented in Fig. 7.17 and summarized in Table B.1.

7.4.6 Results

The observables $\mathcal{B}_{(s+d)\gamma}^{E_\gamma \geq E_0}$, $\mathcal{B}_{s\gamma}^{E_\gamma \geq E_0}$ and $\mathcal{B}_{s\gamma}$ have been calculated. The results with their total, statistical, systematic, model and extrapolation uncertainties are presented in Tables 7.6 to 7.8.

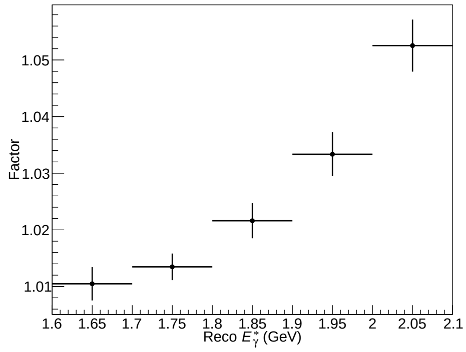


Figure 7.16: Conversion factors to the B rest frame. The error bars represent the model uncertainty.

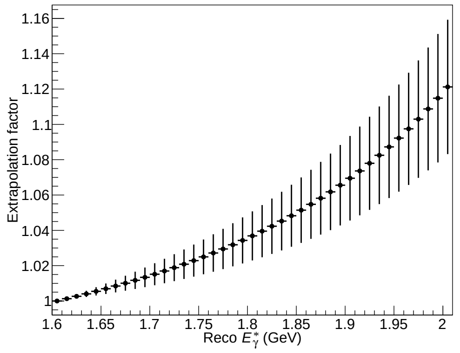


Figure 7.17: Extrapolation factors to the 1.6 GeV threshold. The error bars represent the model uncertainty.

Further tables showing the full breakdown of uncertainties are presented in Appendix A.2. Additionally, the correlations for the measurement of $\mathcal{B}_{(s+d)\gamma}$ for different photon energy thresholds are presented in Table 7.9.

The uncertainties are larger for lower thresholds due to the larger $B\overline{B}$ background, but for higher thresholds the model and extrapolation uncertainties start playing an important role. For this reason, the results are presented for a broad range of thresholds, between 1.6 GeV and 2.1 GeV.

E_γ^B	$\mathcal{B}_{(s+d)\gamma}$	Total	Stat	Syst	Model
1.60	3.28	0.37	0.12	0.33	0.12
1.70	3.18	0.29	0.11	0.25	0.10
1.80	3.14	0.23	0.10	0.19	0.08
1.90	3.06	0.19	0.09	0.15	0.07
2.00	2.90	0.15	0.08	0.11	0.05
2.10	2.61	0.12	0.08	0.09	0.04

Table 7.6: Branching fraction $\mathcal{B}_{(s+d)\gamma}$ for different thresholds in units of 10^{-4}.

E_γ^B	$\mathcal{B}_{s\gamma}$	Total	Stat	Syst	Model
1.60	3.13	0.35	0.11	0.32	0.11
1.70	3.04	0.28	0.11	0.24	0.09
1.80	3.00	0.22	0.10	0.18	0.07
1.90	2.92	0.18	0.09	0.14	0.07
2.00	2.77	0.14	0.08	0.11	0.05
2.10	2.49	0.12	0.07	0.08	0.04

Table 7.7: Branching fraction $\mathcal{B}_{s\gamma}$ for different thresholds in units of 10^{-4}.

E_γ^B	$\mathcal{B}_{s\gamma}$ at 1.6 GeV	Total	Stat	Syst	Model	Extrap
1.60	3.13	0.35	0.11	0.32	0.11	0.00
1.70	3.09	0.28	0.11	0.24	0.09	0.02
1.80	3.11	0.23	0.10	0.19	0.08	0.04
1.90	3.13	0.20	0.09	0.15	0.07	0.07
2.00	3.11	0.19	0.09	0.12	0.06	0.11
2.10	3.01	0.20	0.09	0.10	0.05	0.15

Table 7.8: Branching fraction $\mathcal{B}_{s\gamma}$ extrapolated to 1.60 GeV for different thresholds in units of 10^{-4}.

	1.6	1.7	1.8	1.9	2.0	2.1
1.6	1.00	0.93	0.85	0.76	0.65	0.53
1.7		1.00	0.92	0.83	0.72	0.59
1.8			1.00	0.90	0.79	0.66
1.9				1.00	0.88	0.74
2.0					1.00	0.84
2.1						1.00

Table 7.9: Correlation between $B \to X_{s+d}\gamma$ branching fraction for different thresholds.

7.5 Unfolding of the spectrum

The measured E_γ^* spectrum deviates from the true spectrum due to measurement effects in the calorimeter. These effects can be clearly seen in the folding matrix from Fig. 7.6, where up to 40 % of events belonging to a certain energy bin migrate to other bins. For this reason it is necessary to correct for detector resolution effects to be able to study the true underlying shape of the spectrum. The procedure by which one can revert the calorimeter distortion or any other effect that causes migrations is called “unfolding”. There are many available algorithms, which approach the problem with different mathematical tools, a survey of them can be found in [95]. Unfolding is not compulsory to study the spectrum, as it has been shown in Section 7.3 distortions from the detector can also be included in the theory, but an unfolded spectrum allows for comparison between different experiments.

True ($\boldsymbol{\mu}$) and reconstructed ($\boldsymbol{\nu}$) distributions are related to each other via a response matrix in the following way:

$$\boldsymbol{\mu}_i = R^{-1}_{ij}\boldsymbol{\nu}_j, \tag{7.8}$$

where the response matrix R (and its inverse R^{-1}) gives the probability that an event with true energy belonging to the bin i is reconstructed with an energy belonging to the bin j. In addition to the response matrix, unfolding methods make use of MC information for the true underlying distributions, which are the true and reconstructed distributions that are expected to accurately describe the data. The response matrix simply provides information about the detector reconstruction effects, the true underlying distributions are used to regularize the solution, as described below.

The simplest method to implement, is to invert the response matrix. This method is, however, sensitive to statistical fluctuations. Mathematically, large off-diagonal elements would make the inversion of the matrix difficult, and would give rise to artificially large variances and correlations in the result. Using this algorithm means spreading a statistical fluctuation over all bins. For this reason the idea of "regularization" is introduced. Regularizing the unfolding means requiring certain characteristics of the unfolded solution, such that statistical fluctuations are not interpreted as features of the data. Much like the matrix inversion method, regularized unfolding algorithms try to solve a system of equations of the form:

$$\hat{\boldsymbol{\mu}} = R^{-1}\boldsymbol{\nu}, \tag{7.9}$$

but with additional regularization conditions. The system of equations can be written as:

$$R^{-1}\boldsymbol{\nu} - \hat{\boldsymbol{\mu}} + k \cdot C\hat{\boldsymbol{\mu}} = \text{min}. \tag{7.10}$$

The regularized system of equations makes use of a Lagrange multiplier k, that alongside the function C, determines the *a priori* condition of the solution. In this case the hat on $\hat{\boldsymbol{\mu}}$ shows that this is the estimator of the solution. The Lagrange multiplier k is called the "regularization parameter". A typical choice for the function C is the "curvature" of the spectrum, meaning a smooth solution is required with small bin-to-bin variations.

For this analysis an unfolding algorithm that makes use of a Singular Value Decomposition of the response matrix was chosen. The advantage of this method over others, is that it requires a simpler formulation and implementation. Singular value decomposition means that a matrix is factorized in the form $A = USV^T$, where U is an $m \times m$ orthogonal matrix, V is an $n \times n$ orthogonal matrix, and S is an $m \times n$ diagonal matrix with non-negative diagonal elements s_i called singular values of the matrix A. The columns of U and V are called the left and right singular vectors. From the practical point of view, the method is advantageous since it has been implemented in ROOT[‡] as the TSVDUnfold class [97]. In SVDU the regularization condition C

[‡] ROOT is the data analysis framework that is widely used in experimental particle physics [96].

is defined to be the curvature matrix C[§] .

$$C = \begin{pmatrix} -1 & 1 & 0 & & \cdots & \\ 1 & -2 & 1 & & \cdots & \\ 0 & 1 & -2 & & \cdots & \\ & & & \ddots & & \\ \vdots & \vdots & \vdots & & -2 & 1 \\ & & & & 1 & -1 \end{pmatrix} \tag{7.11}$$

SVDU performs two transformations of the system of equations. The first is a renormalization of the unknowns such that the response matrix R no longer represents probabilities but actually the number of events. This matrix can be readily obtained from MC, is physically motivated, and has no effect on the solution, since for a well determined system of linear equations, a rescaling will leave it unchanged. The second transformation involves a rescaling of the equations. Given that there are uncertainties in the measurement of the spectrum $\boldsymbol{\mu}$, it makes sense to rescale the equations by these errors, or better yet, to rescale them using the covariance matrix in order to have equations with the same weight. Thinking of this as a least square minimization problem, and only considering the first part of Eq. (7.10), the condition without regularization reads:

$$(R\boldsymbol{\nu} - \boldsymbol{\mu})^T B^{-1} (R\boldsymbol{\nu} - \boldsymbol{\mu}) = \min, \tag{7.12}$$

where the covariance matrix, B, is symmetric and positive-definite so it can also be factorized using SVDU. Ultimately the least-square minimization problem that has to be solved is:

$$(\tilde{R}\tilde{\boldsymbol{\nu}} - \tilde{\boldsymbol{\mu}})^T (\tilde{R}\tilde{\boldsymbol{\nu}} - \tilde{\boldsymbol{\mu}}) + k(C\tilde{\boldsymbol{\nu}})^T C\tilde{\boldsymbol{\nu}} = \min, \tag{7.13}$$

where the tilde over the matrices and vectors means they have been renormalized. There is no standard prescription for the choice of k: a small regularization means statistical fluctuations could be important, a highly regularized spectrum relies heavily on the input MC information.

In SVDU the singular values are ranked by size and decrease towards a gaussian-distributed random number around one, k is an integer and tells us which one is the last significant singular value. For an un-regularized solution, divisions by very small and statistically insignificant numbers appear when solving the equations. In the regularized solution, k also appears in the denominator, preventing that the divisions give large and randomly amplified numbers. A more detailed explanation of the regularized solution can be found in [91]. Also some guidelines about the choice of k are found in Chapter 7 of the same reference. In this analysis, k will be chosen such that the bias in the parameters of interest is small. The regularization parameter will be denoted as “kReg” in the following.

[§] The curvature of a function is given by its second derivative, in this case the matrix represents the second derivative of a discrete function: $\ddot{f} = \frac{f(x_{i+1}) - 2f(x_i) + f(x_{i-1})}{(\Delta x)^2}$

7.5.1 Default unfolding

For the analysis of the unfolded spectrum there are three observables of interest: the total number of events (which is directly related to the branching fraction), the mean energy of the spectrum ($\langle E^*_\gamma \rangle$), and the variance of the spectrum (ΔE^*_γ). The observables are used together with an ensemble of theoretical models to understand any bias that the choice of model in the unfolding procedure could introduce. The chosen MC samples for this step cover a broad range of possibilities for the true underlying spectrum. All samples are described in Table 7.10. If a bias is observed in the unfolding, it will be taken into account as a systematic uncertainty in the measurement.

Name	Theory model	Description
BLNP0 (default)	Shape-function	Fit values as in Table 7.5
BLNP1	Shape-function	Fit values: m_b +100 MeV and μ^2_π-100 MeV^2
BLNP2	Shape-function	Fit values: m_b-100 MeV and μ^2_π +100 MeV^2
KN0	Kagan-Neubert	Scan values as in Section 7.3.3
KN1	Kagan-Neubert	Scan values: m_b +100 MeV and μ^2_π-100 MeV^2
KN2	Kagan-Neubert	Scan values: m_b-100 MeV and μ^2_π +100 MeV^2
BBU0	Kinetic	Scan values as in Section 7.3.3
BBU1	Kinetic	Scan values: m_b +100 MeV and μ^2_π-100 MeV^2
BBU2	Kinetic	Scan values: m_b-100 MeV and μ^2_π +100 MeV^2

Table 7.10: MC samples used for the study of the unfolding procedure.

The stability of the unfolded result is investigated using toy experiments. They are derived from the default MC sample (BLNP0) and the complete experimental covariance matrix. The BLPN0 sample is used as default throughout. The toys are then unfolded using different model assumptions for the true underlying distribution and the integral and spectral moments are calculated. The mean and variance of the measured observables are compared to the expected values, and a possible bias induced by the unfolding procedure is estimated. The BLNP0 response matrix is shown in Fig. 7.18, the numerical values for the same matrix are presented in Table C.1.

Unfolding with BLNP0 theory

For this study 10000 toy experiments derived from BLNP0 and the total experimental covariance matrix are generated. Their distribution is shown in Fig. 7.19. Initially a "null test" is performed: the generated toys and true information are both chosen from the default sample. As expected there is no bias in the unfolded spectra and the mean of each bin perfectly reproduces the input information, as seen in Fig. 7.19. One must remember that the complete information about the errors is hidden in the off-diagonal terms of the covariance matrix, and the error bars in e.g. Fig. 7.19 only show part of them. It can be seen that the error bars increase in size with larger kReg. This shows that the correlation of bin uncertainties of the unfolded spectrum changes with

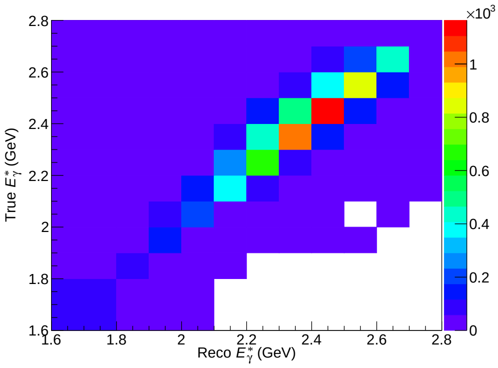


Figure 7.18: Response matrix for the BLNP0 model. The entries of the response matrix are given as number of events and not probabilities.

a different choice of regularization parameter. It is, however, also possible that the unfolding slightly changes the magnitude of the total error.

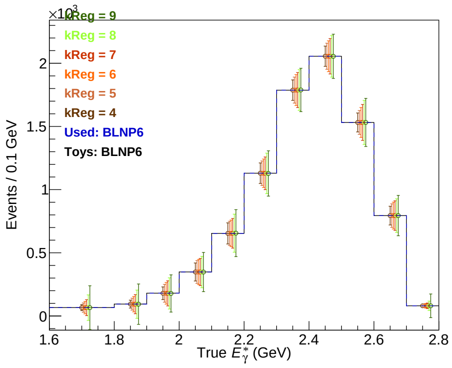


Figure 7.19: Unfolded spectra for different choices of the regularization parameter. The black distribution shows the true E_γ^* spectrum, corresponding to the spectrum used for the generation of toys. The blue distribution corresponds to the distribution used as MC input in the unfolding as true underlying spectrum.

The three observables of interest are calculated after the unfolding and compared with the expectation from the sample the toys are derived from. They are plotted in Fig. 7.20. In the figure corresponding to the integral, the uncertainty of the original distribution (red box for statistical uncertainty, blue box for the total uncertainty) can be compared with the variance of the toys for different regularization parameters (red and black error bars). For the integral and moments the complete information from the off-diagonal elements is also included.

A clear feature is discerned: for all values of kReg the total uncertainty is smaller than the uncertainty of the original distribution. This is particularly clear for the part coming from the systematic uncertainties. Additionally, it can be seen that for strong regularization (low values of kReg) the variance of the toys is smaller. A slight decrease of the uncertainty could be expected due to the additional input of MC information to require smoothness in the solution. What is observed, however, is that the uncertainties are underestimated in the order of $O(25\,\%)$. The observable studied here is the number of events, which is conserved by the unfolding, so the uncertainty on it is expected to be of same size after the unfolding. The strong underestimation of the uncertainty has to be better understood and might indicate some shortcoming of the unfolding algorithm.

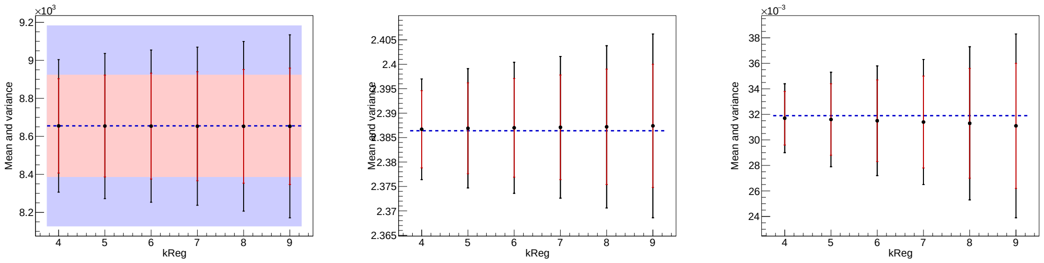


Figure 7.20: Integral (left), mean (center) and variance (right) of the unfolded spectrum for different values of the regularization parameter kReg. The plots show the mean and variance for each observable of the unfolded toys using BLNP0 MC information. The red bars and box correspond to the statistical uncertainty, whereas the blue box and black bars show the total uncertainty. The boxes are the uncertainties for the original distribution, while the bars correspond to the variance of the unfolded toy distributions.

7.5.2 Unfolding tests and Asimov data

It is suspected that the observed performance of the unfolding algorithm is related to the structure of the experimental covariance matrix, with large correlations and uncertainties at low photon energies. Statistical and systematic covariances are considered separetely to generate toys and Asimov data sets [98][¶] . They have different correlation structures and behaved differently in the tests shown in the previous section. The statistical and systematic correlation matrices are shown in Section 7.5.2, the statistical correlation matrix is simply the identity matrix.

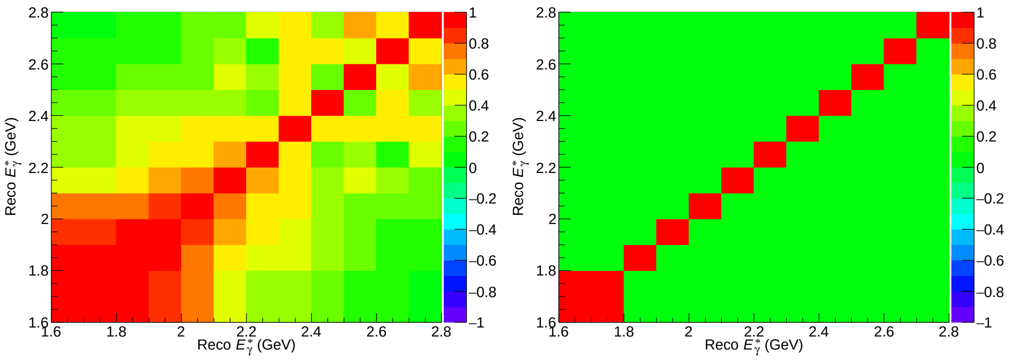


Figure 7.21: Systematic (left) and statistical (right) correlation matrices.

[¶] Asimov data sets are representative sets that can replace an ensemble of simulated experiments.

The same procedure described in the previous section is used to generate toy distributions, unfold them using different values for kReg, and calculate the integral of the unfolded spectrum. The average uncertainty on the integral for all unfolded toy distributions and for different values of kReg is plotted as the error bars in Fig. 7.22, the central value is fixed at zero. It can be observed that while for the toys generated from the statistical covariance matrix the errors before and after unfolding are very similar, for toys generated from the systematic covariance matrix, the error after unfolding is always strongly underestimated.

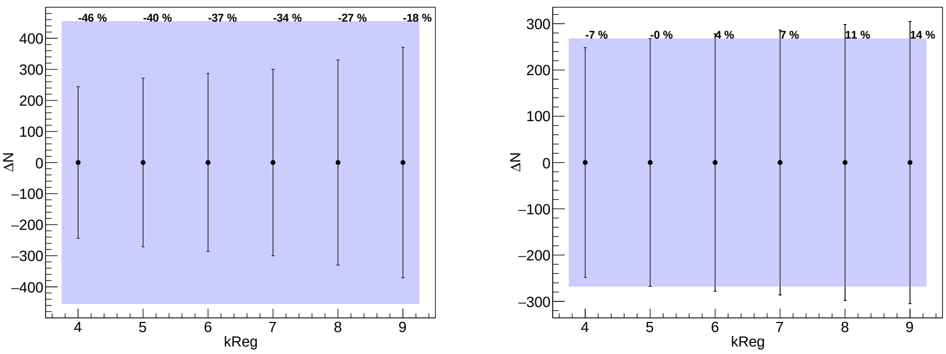


Figure 7.22: Uncertainty on the integral of unfolded toy distributions. The toy distributions are generated from the systematic (left) and statistic (right) covariance matrices. The uncertainty on the integral of the spectrum before unfolding (blue box) and after it (error bars) are compared for several values of kReg. On top of each error bar the difference between both uncertainty values is written.

For the same sources of uncertainty, three sets of Asimov data were generated, corresponding to the central value and the $\pm 1\sigma$ variations. Asimov data are advantageous because they by construction smooth and any features of the unfolding can be directly related to properties of the distributions, disregarding bin-to-bin fluctuations. The tested Asimov data sets are constructed from the $\pm 1\sigma$ fluctuations of the covariance matrices. The results of the unfolding and the uncertainties for each bin before and after the unfolding are presented in Fig. 7.23. The difference plots shown below the spectra in Fig. 7.23 are the difference between the central distribution and the $\pm 1\sigma$ Asimov data sets. From these plots one can understand the regularization and the small uncertainties: SVDU penalizes entries with large uncertainties [91]. Given that the systematic uncertainties throughout the photon energy spectrum are not equal throughout all bins but large at low energy, the regularization cuts away this part of the spectrum, leading to an underestimation of the uncertainties.

7.5.3 Modified unfolding algorithm

As the default SVDU is not well suited for this particular problem, it was attempted to modify the algorithm. As previously pointed out, the system of equations to be solved in SVDU is reweighted by the covariance matrix penalizing the low end of the spectrum and leading to the underestimation of uncertainties. This step was removed from the calculations by suggestion of

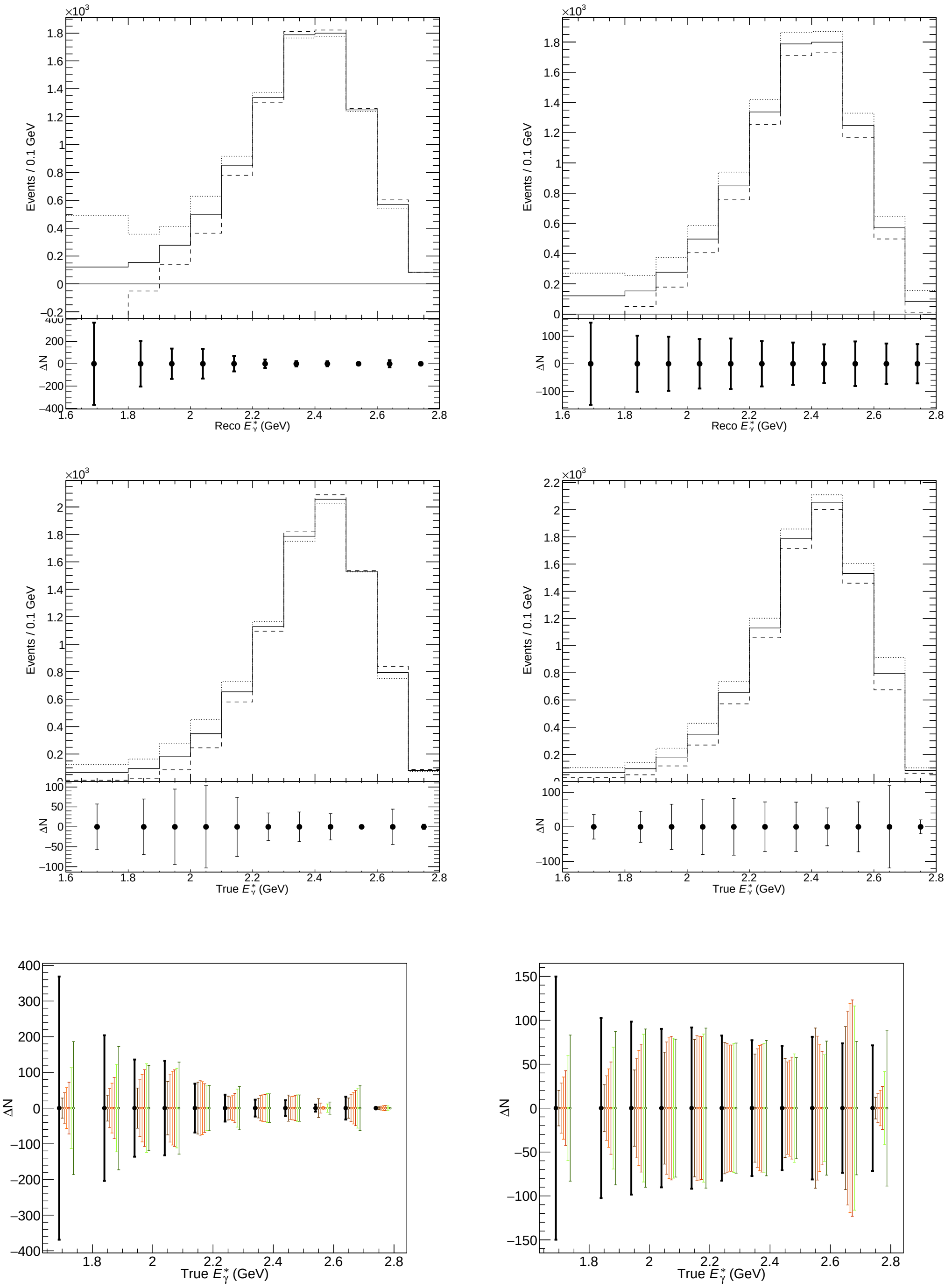


Figure 7.23: Asimov data derived from systematic (left) and statistic (right) covariance matrices. The plots on the top show the default MC spectrum (solid line) as well as the $+1\sigma$ (dotted line) and -1σ (dashed line) Asimov data. Beneath these spectra the difference (ΔN) between the default MC spectrum and the Asimov data on each bin is shown. The plots in the central row show the unfolded Asimov data with kReg=6, and the bottom one compares the differences ΔN for each bin before unfolding (black bars) and after for several kReg values. The colors of the bars are the same as for Fig. 7.19.

the authors [99](∞). A similar test is performed, generating toys from the BLNP0 sample with the total experimental covariance matrix, unfolding them using the modified SVDU for various values of kReg, and looking at the uncertainty on the integral of the unfolded toy distributions. The plots corresponding to this new test are shown in Fig. 7.24. In this case both the statistical and systematic uncertainties after the unfolding are of the same size as the uncertainties before unfolding.

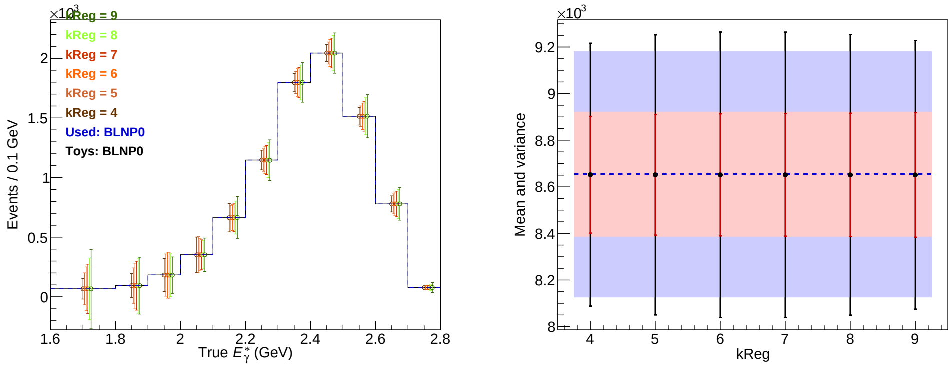


Figure 7.24: Unfolding of BLNP0 toys with BLNP0 true information using the modified SVDU before correcting for selection efficiency. The colors for the right-hand-side plot are explained in Fig. 7.20.

After checking that the unfolding algorithm no longer underestimates the uncertainties, the possible bias introduced by the unfolding can be determined. For this purpose, the procedure is carried out as it would be done for the final measurement: the spectrum is corrected for the selection efficiency (dividing each bin by $\varepsilon_{\mathrm{sel},\,i}$ Fig. 7.15). The default sample BLNP0 is used to generate toys, and all nine samples described in Table 7.10 are used to unfold them and then determine the unfolding uncertainty. In Figs. 7.25 to 7.27, the sample BLNP0 is plotted as a solid line, the sample used as true underlying shape for the regularization is shown as a dashed line, and the mean and variance of the unfolded toys are shown as points with error bars. All used MC samples recover the actual underlying shape from which the toys are produced. The unfolding systematic uncertainty for each bin, the integral, and the spectral moments is taken as the largest deviations between all unfolded spectra for kReg=6 with respect to the default. As seen in Fig. 7.28, the uncertainty related to the unfolding is small. The correlation matrix before and after the unfolding is presented in Fig. 7.29. As can is expected, the uncertainties of the unfolded solution have slightly different correlations compared to the uncertainties before unfolding.

7.5.4 Unfolding of the data distribution

The background-subtracted spectrum of Fig. 7.1 is corrected bin-by-bin for the selection efficiency $\varepsilon_{\mathrm{sel},\,i}$ using Fig. 7.15. The efficiency-corrected spectrum is unfolded with the modified SVDU algorithm which uses BLNP0 true information. The chosen regularization parameter is kReg=6.

(∞)The specific part that is skipped in the algorithm is Equation 34 from [91].

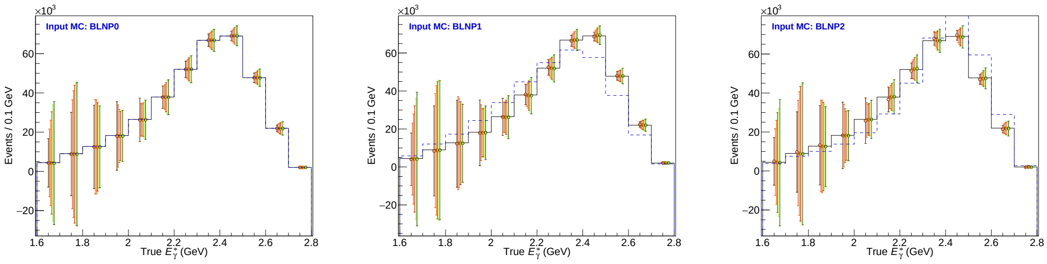


Figure 7.25: Unfolding of BLNP0 toy distributions using BLNP0 (left), BLNP1 (center) and BLNP2 (right) true information. The unfolding is done after correcting for selection efficiencies . The color scheme is the one used in Fig. 7.24.

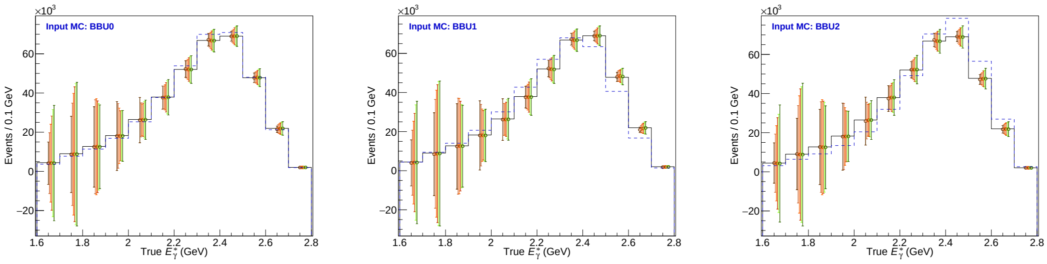


Figure 7.26: Unfolding of BLNP0 toy distributions using BBU0 (left), BBU1 (center) and BBU2 (right) true information. The unfolding is done after correcting for selection efficiencies, the color scheme is the one used in Fig. 7.24.

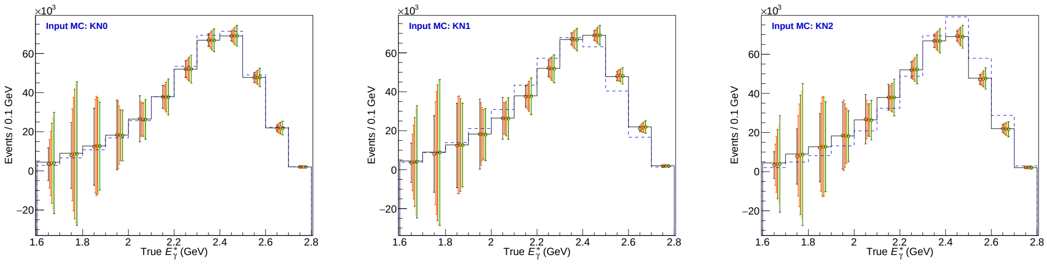


Figure 7.27: Unfolding of BLNP0 toy distributions using KN0 (left), KN1 (center) and KN2 (right) true information. The unfolding is done after correcting for selection efficiencies. The color scheme is the one used in Fig. 7.24.

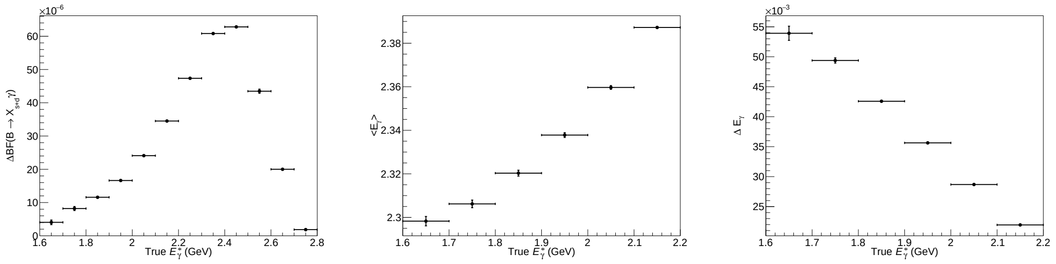


Figure 7.28: Unfolding uncertainty in the unfolded spectrum (left), first moment (center) and second moment (right).

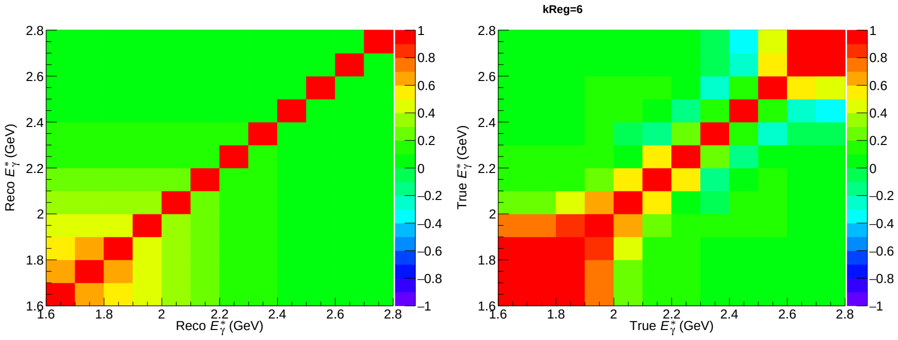


Figure 7.29: Correlation matrix before (left) and after the unfolding with regularization parameter kReg=6 (right).

The result of the unfolding, with statistical and systematic uncertainties is presented in Fig. 7.30. The values are presented in Table A.8.

7.6 Partial branching fraction and spectral moments

Making use of the unfolded spectrum Fig. 7.30, one can calculate the $B \to X_{s+d}\gamma$ partial branching fractions ($\Delta\mathcal{B}_{(s+d)\gamma}$) using Eq. (7.1). Similarly, one can calculate the first and second spectral moments for several photon energy thresholds and the correlations among these measurements. The results of these measurements are shown in Fig. 7.31. The systematic uncertainties that enter these measurements are the same as those already determined for the measurement of the inclusive branching fraction. The full set of uncertainties is presented in Appendices A.3 and A.4, and a summary of them can be found in Tables 7.11 to 7.13. The correlations between the spectral moments for different photon energy thresholds are presented in Table 7.14.

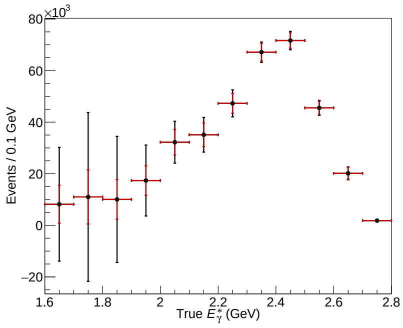


Figure 7.30: Unfolded and efficiency-corrected data spectrum. The red error bars show statistical uncertainties, the black error bars additionally include systematic uncertainties from background subtraction.

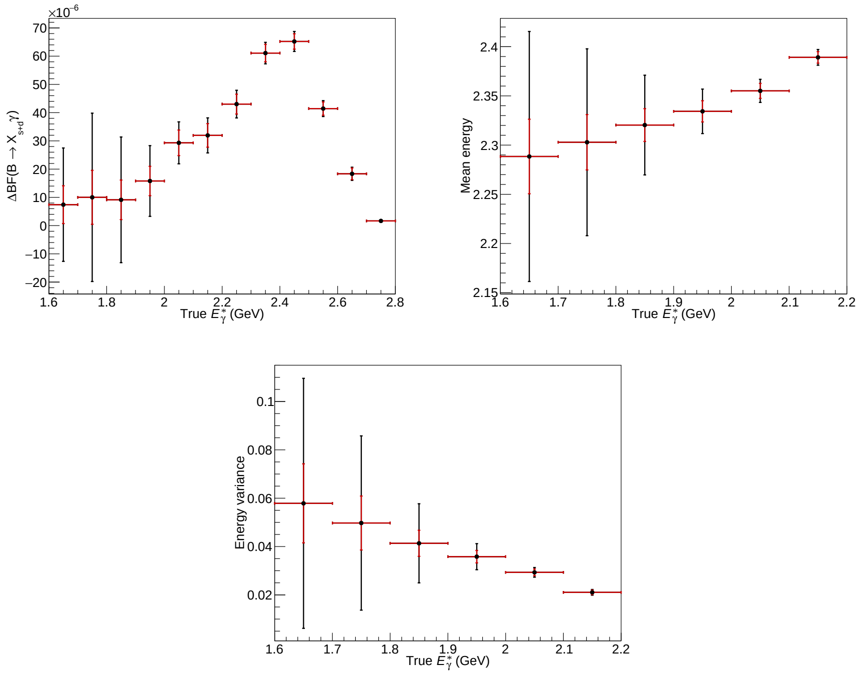


Figure 7.31: Partial $B \to X_{s+d}\gamma$ branching fractions (top left), first moment (top right) and second moment (bottom), derived from the unfolded spectrum in the CM frame.

E_γ^* bin (GeV)	$\Delta\mathcal{B}_{(s+d)\gamma}$	Total	Stat	Syst	Model
1.6 – 1.7	4.1	20.1	6.7	18.9	0.8
1.7 – 1.8	8.2	29.8	9.5	28.2	0.8
1.8 – 1.9	11.6	22.2	7.0	21.1	0.5
1.9 – 2.0	16.6	12.5	5.2	11.4	0.4
2.0 – 2.1	24.1	7.4	4.5	5.9	0.6
2.1 – 2.2	34.5	6.2	4.2	4.5	0.9
2.2 – 2.3	47.4	4.9	3.5	3.3	0.9
2.3 – 2.4	60.8	3.8	3.1	2.0	0.9
2.4 – 2.5	62.8	3.5	2.7	1.9	1.1
2.5 – 2.6	43.5	2.8	2.3	1.3	0.8
2.6 – 2.7	20.0	2.3	2.0	1.0	0.5
2.7 – 2.8	1.9	0.3	0.3	0.1	0.1

Table 7.11: Partial branching fractions of the $B \to X_{s+d}\gamma$ spectrum in units of 10^{-6} and uncertainties.

E_γ^* thr.	$\langle E_\gamma^* \rangle$	Total	Stat	Syst	Model
1.6	2.298	0.129	0.038	0.123	0.003
1.7	2.306	0.094	0.028	0.090	0.002
1.8	2.320	0.050	0.016	0.047	0.001
1.9	2.338	0.023	0.011	0.020	0.001
2.0	2.360	0.012	0.008	0.009	0.001
2.1	2.387	0.008	0.006	0.005	0.001

Table 7.12: Mean energy of the $B \to X_{s+d}\gamma$ spectrum and uncertainties for different thresholds, values in GeV.

E_γ^* thr.	ΔE_γ^*	Total	Stat	Syst	Model
1.6	5.392	5.506	1.728	5.225	0.166
1.7	4.938	3.625	1.121	3.446	0.096
1.8	4.258	1.589	0.525	1.498	0.050
1.9	3.563	0.548	0.260	0.481	0.036
2.0	2.869	0.207	0.141	0.149	0.029
2.1	2.195	0.111	0.087	0.064	0.024

Table 7.13: Variance of the $B \to X_{s+d}\gamma$ spectrum in units of $10^{-2}\mathrm{GeV}^2$ and uncertainties for different thresholds.

		$\langle E^*_\gamma\rangle$						ΔE^*_γ					
		1.6	1.7	1.8	1.9	2.0	2.1	1.6	1.7	1.8	1.9	2.0	2.1
$\langle E^*_\gamma\rangle$	1.6	1.00	0.95	0.85	0.65	0.30	0.15	-0.76	-0.98	-0.92	-0.70	-0.29	-0.09
	1.7		1.00	0.96	0.77	0.38	0.18	-0.55	-0.91	-0.98	-0.82	-0.37	-0.12
	1.8			1.00	0.90	0.53	0.25	-0.41	-0.79	-0.96	-0.92	-0.51	-0.17
	1.9				1.00	0.81	0.45	-0.29	-0.59	-0.79	-0.92	-0.73	-0.29
	2.0					1.00	0.80	-0.12	-0.26	-0.37	-0.56	-0.73	-0.46
	2.1						1.00	-0.05	-0.10	-0.10	-0.11	-0.26	-0.37
ΔE^*_γ	1.6							1.00	0.77	0.52	0.34	0.12	0.03
	1.7								1.00	0.90	0.67	0.28	0.09
	1.8									1.00	0.88	0.44	0.13
	1.9										1.00	0.74	0.29
	2.0											1.00	0.71
	2.1												1.00

Table 7.14: Correlations between first and second $B \rightarrow X_{s+d}\gamma$ spectral moments for different energy thresholds. The moments are given for the unfolded spectrum in the CM frame.

CHAPTER 8

CP asymmetry

MEASURING the *CP* asymmetry $\mathcal{A}^{\text{CP}}$ requires the estimation of additional systematic effects specific to this observable and that do not play a role for the measurement of the branching fraction. Possible asymmetries arising from the reconstruction or selection procedure, or asymmetries that are present in the background could bias the result. Additionally, since the flavor-tagging is based on the lepton charge, if the charge does not correctly correlate to the signal B flavor a dilution of the asymmetry occurs.

The *CP* asymmetry is defined in Eq. (2.12), but can be rewritten it in a more appropriate way, utilizing the number of signal decays from a $B \to X_{s+d}\gamma$ or a $\overline{B} \to \overline{X}_{s+d}\gamma$ as:

$$\mathcal{A}^{\text{CP}}_{(s+d)\gamma} = \frac{N_{\overline{B}\to\overline{X}_{s+d}\gamma} - N_{B\to X_{s+d}\gamma}}{N_{\overline{B}\to\overline{X}_{s+d}\gamma} + N_{B\to X_{s+d}\gamma}} \overset{?}{=} \frac{N^+ - N^-}{N^+ + N^-}. \tag{8.1}$$

Effectively, in this analysis one has only access to the number of events tagged with a positive or negative lepton, so the asymmetry is calculated with these yields. The symbol $\overset{?}{=}$ indicates that although the charge of the chosen lepton, expected to come from a semileptonic decay, is directly correlated to the flavor of its mother B as sketched in Fig. 8.1, this is not always the case. A wrong-tag probability (ω) arises in different cases. The first occurs when the neutral B mesons have oscillated and both B mesons in the event decay having the same flavor. The second case is when a lepton that is not from a semileptonic decay and has a wrong charge-flavor correlation has been chosen. The true value of $\mathcal{A}^{\text{CP}}$ is related to the measured asymmetry ($\mathcal{A}^{\text{CP}}_{\text{meas}}$), the wrong tag probability and possible asymmetries in the background ($\mathcal{A}_{\text{bkg}}$) or from detection ($\mathcal{A}_{\text{det}}$) by the relation:

$$\mathcal{A}^{\text{CP}}_{(s+d)\gamma} = \frac{1}{1-2\omega}\left(\mathcal{A}^{\text{CP}}_{\text{meas}} + \frac{B}{S}\mathcal{A}_{\text{bkg}} + \mathcal{A}_{\text{det}}\right). \tag{8.2}$$

The asymmetry in the $B\overline{B}$ background affects the measurement proportionally to the ratio between $B\overline{B}$ and $B \to X_{s+d}\gamma$ yields in the signal region B/S.

Given that $\mathcal{A}^{\text{CP}}_{(s+d)\gamma}$ is proportional to the difference between positive- (ell^+) and negative-tagged (ell^-) events, systematic uncertainties related to background subtraction cancel. In this chapter the systematic uncertainties are estimated and the asymmetry for different photon energy thresholds are measured. The independence of $\mathcal{A}^{\text{CP}}_{(s+d)\gamma}$ with respect to the choice of energy

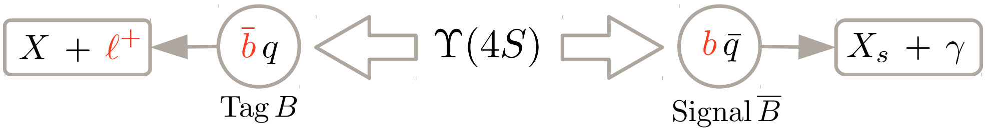


Figure 8.1: Correlation between the flavor of the tag and signal B mesons to the lepton charge (note the particles in red) with a semileptonic B tag.

threshold [39] is tested.

8.1 Detector induced charge asymmetries

A possible charge asymmetry in the identification of e or μ is studied in samples of $J/\psi \to e^+e^- \,/\, \mu^+\mu^-$ decays. To determine the efficiencies a "tag and probe" approach is used. First, one of the leptons in the decay is clearly identified by requiring a high probability that it is compatible with the muon or electron hypothesis: $\mu\text{ID} > 0.99$ and $e\text{ID} > 0.80$. These cuts are the standard used by Belle to estimate PID systematic errors [100]. After this, the same selection criteria as the for $B \to X_s\gamma$ analysis are applied to the second candidate ($\mu\text{ID} > 0.90$ and $e\text{ID} > 0.90$) and the numbers of events that pass and fail the cut ($N^{\pm}_{\text{pass}}$, $N^{\pm}_{\text{fail}}$) are determined.

This study is performed in 11 lepton momentum and 8 polar angle bins to obtain a complete understanding of the kinematic phase space. The yields $N^{\pm}_{\text{pass}}$ and $N^{\pm}_{\text{fail}}$ are determined by performing a fit to the invariant mass of the e^+e^- and $\mu^+\mu^-$ pairs. The mass distribution is modelled using a Crystal-Ball function [101] while the background is modeled using a polynomial. The efficiency is determined as:

$$\varepsilon^{\pm} = \frac{N^{\pm}_{\text{pass}}}{N^{\pm}_{\text{pass}} + N^{\pm}_{\text{fail}}}. \tag{8.3}$$

The fits are performed simultaneously for positive and negative leptons in each polar-angle and lepton-momentum bin and the ratio of efficiencies, $\varepsilon^-/\varepsilon^+$, is extracted directly from the fit. Example fits can be seen in Fig. 8.2 for muons and in Fig. 8.3 for electrons.

The efficiency ratios are plotted in Fig. 8.4 for muons and Fig. 8.5 for electrons. The average ratio of efficiencies is 1.0006 ± 0.0007 for muons and 0.9948 ± 0.0029 for electrons. The ratio is consistent with no asymmetry within 0.1 % for muons. In the electron case the ratio is 1.8σ below unity. This effect is not very significant and any asymmetry effectively vanishes when averaged over both lepton flavors.

The quantity of interest is the asymmetry of efficiencies which is a correction to the measured value of $\mathcal{A}^{\text{CP}}$ and is calculated as:

$$\mathcal{A}_{\text{LID}} = \frac{\varepsilon^+ - \varepsilon^-}{\varepsilon^+ + \varepsilon^-} = \frac{1 - \frac{\varepsilon^-}{\varepsilon^+}}{1 + \frac{\varepsilon^-}{\varepsilon^+}}. \tag{8.4}$$

The asymmetries for both lepton flavors and the average over electrons and muons was measured very precisely and is consistent with no asymmetry. The results are summarized in Table 8.1. The average asymmetry over both lepton flavors, $\mathcal{A}_{\text{LID}}$, is calculated taking into account the number

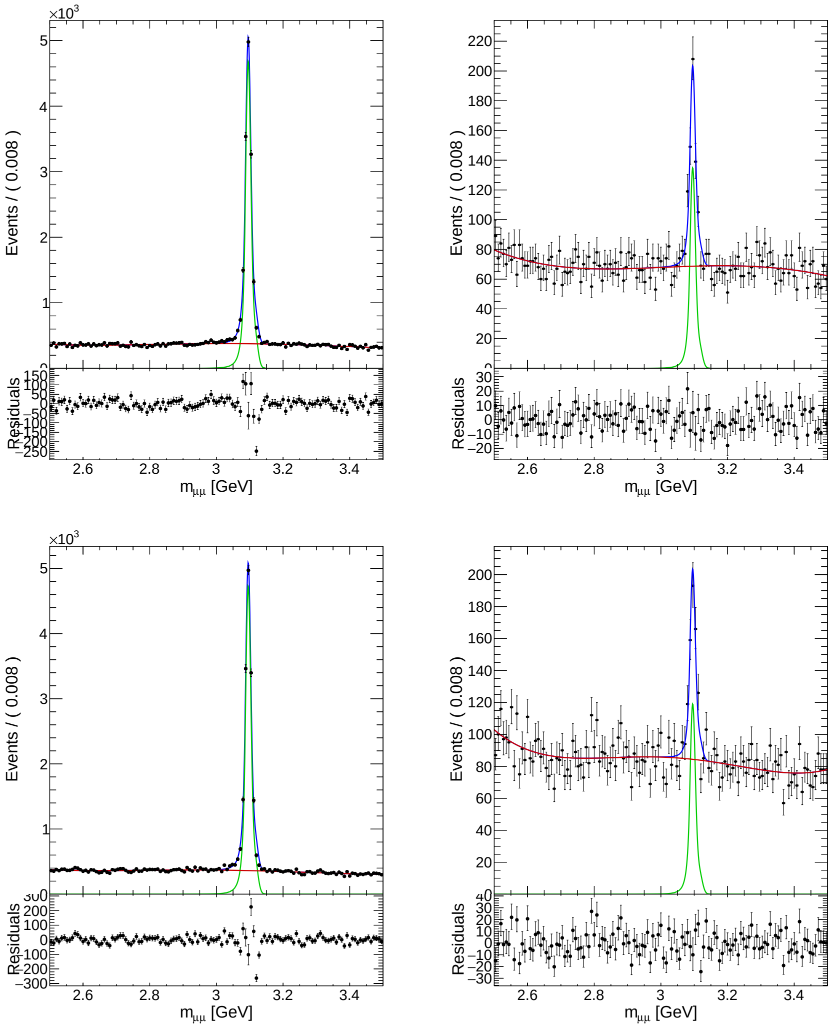


Figure 8.2: Example of a fit to the J/ψ mass peak for $\mu^{+}\mu^{-}$ pairs. The red line corresponds to the polynomial background, the green line to the signal Crystal-Ball function and the blue line the the sum of both. The fit residuals are shown below each plot. The top plots correspond to negative charges (μ^{+} tagged) and the lower to positive charges (μ^{-} tagged). The plots on the left is for events passing the electron ID criterion and the plots on right shows events failing the criterion.

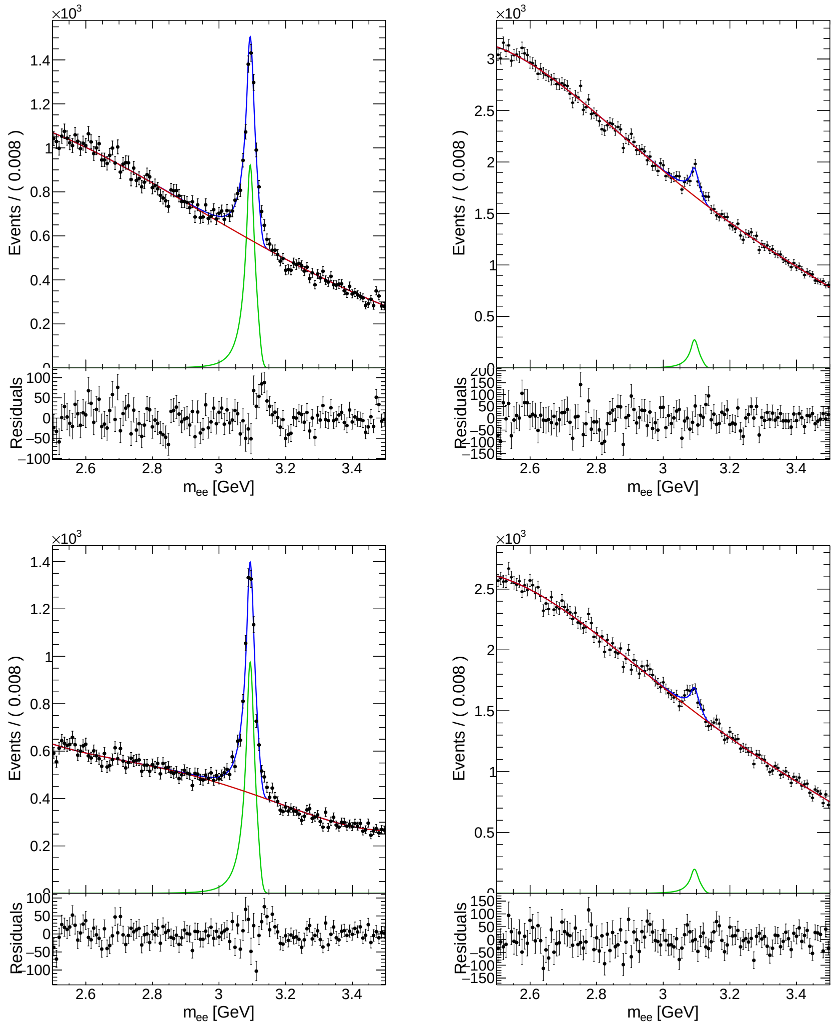


Figure 8.3: Example of a fit to the J/ψ mass peak for e^-e^+ pairs. The red line corresponds to the polynomial background, the green line to the signal Crystal-Ball function and the blue line to the sum of both. The fit residuals are shown below each plot. The top plots correspond to negative charges (e^+ tagged) and the lower to positive charges (e^- tagged). The plots on the left is for events passing the electron ID criterion and the plots on right shows events failing the criterion.

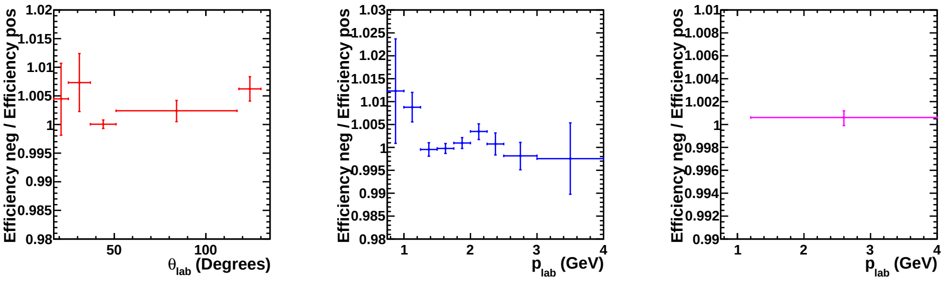


Figure 8.4: Ratio of efficiencies for $\varepsilon^{-}/\varepsilon+$ μ ID criterion. The plot on the left shows the ratio integrated over all momentum bins, the plot in the center shows the ratio integrated over polar angle bins and the plot on the right the ratio integrated over all bins. The ratio over all polar angle and momentum regions is $\varepsilon^{-}/\varepsilon+ = 1.0006 \pm 0.0007$.

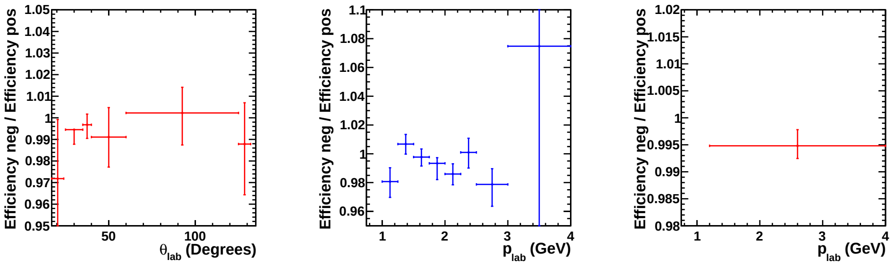


Figure 8.5: Ratio of efficiencies $\varepsilon^{-}/\varepsilon+$ for the eID cut. The plot on the left shows the ratio integrated over all momentum bins, the plot in the center shows the ratio integrated over all polar angle bins and the plot in the right the ratio integrated over all bins. The ratio over all polar angle and momentum regions is $\varepsilon^{-}/\varepsilon+ = 0.9948 \pm 0.0029$.

of e and μ found after the selection.

In a Belle study [102], the tracking charge asymmetry ($\mathcal{A}_{\text{track}}$) was measured in data and MC, using partially reconstructed D^* decays in experiments 07 to 71. The decay chain used is $D^* \to \pi D$, $D \to \pi\pi K^0_S$, $K^0_S \to \pi^+\pi^-$. The efficiency of finding a charged pion ($\varepsilon_{\pi^\pm}$) is determined by taking the ratio between the number of events in which a track is found in the detector and those in which a track can be unambiguously identified by making use of kinematic constrains. Making use of the finding efficiencies one defines the tracking charge asymmetry as:

$$\mathcal{A}_{\text{track}} = \frac{\varepsilon_{\pi^+} - \varepsilon_{\pi^-}}{\varepsilon_{\pi^+} + \varepsilon_{\pi^-}}. \tag{8.5}$$

The results are consistent with no asymmetry and are summarized in Table 8.1.

The total particle detection asymmetry, $\mathcal{A}_{\text{det}}$, is the sum of both effects described above: PID and tracking. It is calculated as the sum of both contributions and is also shown in Table 8.1.

Asymmetry	Measured value
$\mathcal{A}_{\mu\text{ID}}$	$(-0.03 \pm 0.03)\%$
$\mathcal{A}_{e\text{ID}}$	$(0.26 \pm 0.14)\%$
$\mathcal{A}_{\text{LID}}$	$(0.11 \pm 0.07)\%$
$\mathcal{A}_{\text{track}}$	$(-0.01 \pm 0.21)\%$
$\mathcal{A}_{\text{det}}$	$(\mathbf{0.10} \pm \mathbf{0.22})\%$

Table 8.1: Charge-asymmetries for lepton identification and tracking. The lepton identification asymmetries are given for μ, e and the average between both ($\mathcal{A}_{\text{LID}}$). $\mathcal{A}_{\text{track}}$ is the tracking asymmetry. $\mathcal{A}_{\text{det}}$ is the total detection charge asymmetry due to tracking and lepton detection.

8.2 CP asymmetry in the background

Rare processes that contribute to the $B \to X_{s+d}\gamma$ background could be CP violating. Subtracting them without properly modeling CP violation in the MC results on an incorrect background subtraction. The different background components were checked in order to see if they have significant CP violating effects that are not modeled in the MC. In Fig. 8.6 the most important components in the rare background sample and their measured $\mathcal{A}^{\text{CP}}$ are shown. The contributions are shown as a percent of the rare MC sample after selection. The most important rare decay background in this analysis the inclusive $B \to X_s\eta$ decays. The CP asymmetry in this decay has been measured by Belle to be $\mathcal{A}^{\text{CP}} = (-13 \pm 5)\%$ [103], accordingly the MC is corrected to reproduce this CP asymmetry. All other components with non-zero CP asymmetries are negligible.

Any CP asymmetry effects in the generic $B\overline{B}$ background should be washed out since they are a mixture of many contributions of which none is dominant. In any case the selection procedure could be slightly asymmetric for positive and negative tagged background events from fluctuations in the data. To study this effect the asymmetry $\mathcal{A}_{\text{bkg}}$ is measured in data and MC in order to determine a possible bias $\mathcal{A}^{\text{CP}}$. The asymmetry $\mathcal{A}_{\text{bkg}}$ is measured in the low energy sideband $E^*_\gamma \leq 1.7\,\text{GeV}$, continuum is subtracted from the on-resonance data, and the assumption that no signal events are found in the sample is made. The asymmetry $\mathcal{A}_{\text{bkg}}$ is measured in the same way as $\mathcal{A}^{\text{CP}}$: $\mathcal{A}_{\text{bkg}} = \frac{N^+ - N^-}{N^+ + N^-}$. The data is split in four statistically independent sub-samples in order to see if a trend shows. The sub-samples were chosen as:

- Set I: corresponding to exp07 to exp27, a total of $140.7\,\text{fb}^{-1}$ on resonance and $15.6\,\text{fb}^{-1}$ off-resonance data.
- Set II: corresponding to exp31 to exp43, a total of $268.9\,\text{fb}^{-1}$ on resonance and $31.2\,\text{fb}^{-1}$ off-resonance data.
- Set III: corresponding to exp45 to exp55, a total of $188.5\,\text{fb}^{-1}$ on resonance and $20.7\,\text{fb}^{-1}$ off-resonance data.
- Set IV: corresponding to exp61 to exp73, a total of $104.7\,\text{fb}^{-1}$ on resonance and $21.9\,\text{fb}^{-1}$ off-resonance data.

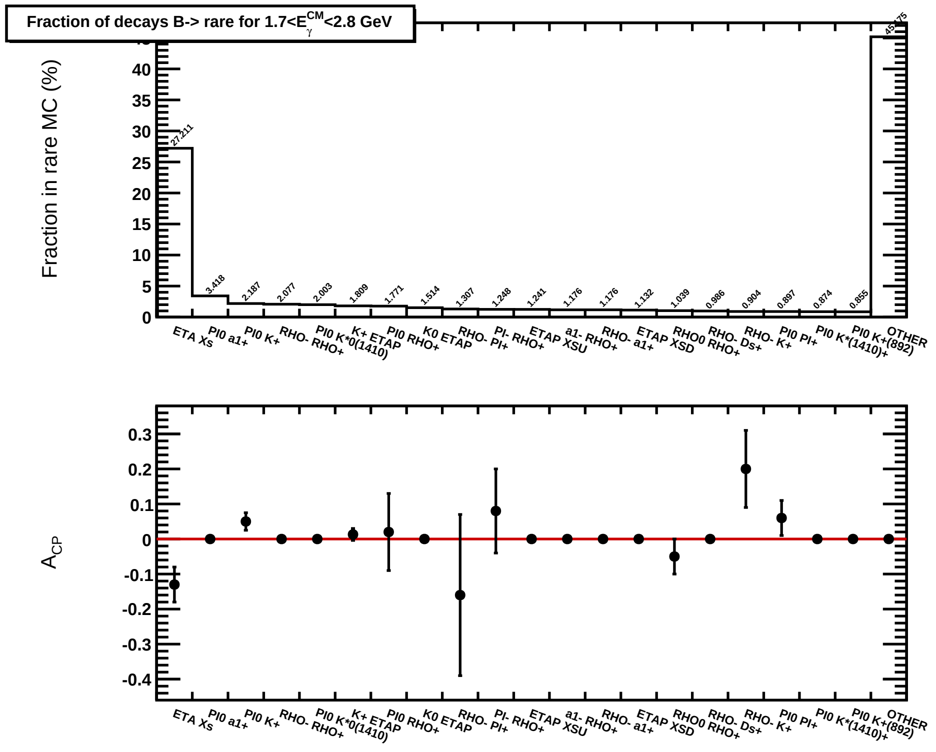


Figure 8.6: Rare background components (top plot) and CP asymmetries (bottom plot).

The results of the measurement are summarized in Fig 8.7, for data and MC. For the MC $B\overline{B}$ sample the asymmetries in the regions 1.7 to 2.8 GeV and 1.4 to 2.8 GeV are included for reference. Each bin corresponds to a data subset and the last bin to the complete data sample. For data, $\mathcal{A}_{\rm bkg}$ is measured to be consistent with zero $\mathcal{A}_{\rm bkg}^{\rm data} = (-0.14 \pm 0.78)\%$. The measured asymmetry in MC for the low-energy sideband is $\mathcal{A}_{\rm bkg}^{\rm MC} = (-0.29 \pm 0.21)\%$. As pointed out, this uncertainty enters the corrections weighted by the ratio of $B\overline{B}$ to $B \to X_{s+d}\gamma$ events in the region of measurement, which is presented in Fig. 8.8 for several photon energy thresholds.

8.3 Wrong tag probabilities

As indicated in Eq. (8.2), the measured value of $\mathcal{A}^{\rm CP}$ differs from the true value $\mathcal{A}^{\rm CP}_{\rm true}$. Apart from corrections due to systematic asymmetries, the measured asymmetry has to be corrected for the dilution factor that arises due to “wrong tags”. The main contribution to wrong tags comes from neutral B oscillation. In this case even if the tag lepton corresponds to a true lepton from a semileptonic decay, the signal B is of the same flavor as the tag B. Leptons from such cases come directly from the decay of a B (prompt leptons). The next contribution in importance is from leptons that are not direct decay products of the B, but come from subsequent decays of B daughter particles. They are referred to as “secondary leptons”. Finally a small factor arises from

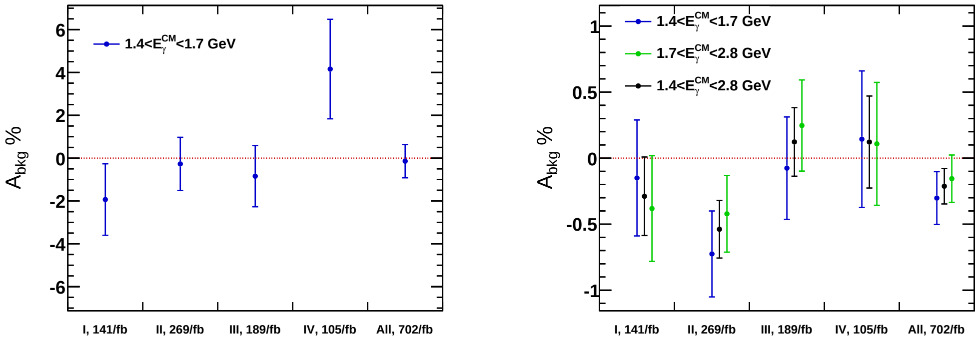


Figure 8.7: Raw asymmetry in the control sample in data (left plot) and MC (right plot).

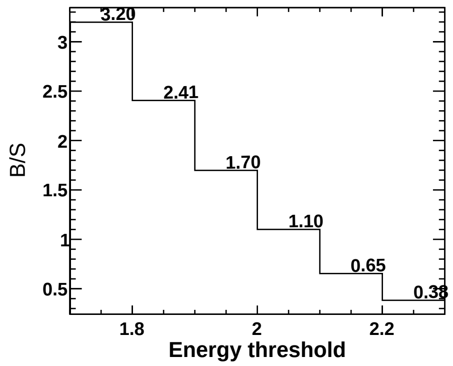


Figure 8.8: Ratio B/S of $B\overline{B}$ to signal for different energy thresholds.

hadrons reconstructed as leptons, referred to as “misidentified hadrons”.

The wrong-tag probability corresponding to neutral B oscillation is given by the fraction of prompt leptons among all lepton candidates (91.1%), the probability that a neutral B oscillates, $\chi_d = 0.1875 \pm 0.0020$, and the fraction of neutral B mesons from the decay of the $\Upsilon(4S)$ $f_{00} = 0.487 \pm 0.006$ [27]:

$$\omega_{\text{osc.}} = 0.911 \cdot \chi_d \cdot f_{00} = 0.0832 \pm 0.0015. \tag{8.6}$$

Secondary leptons with a wrong charge correlation to the B flavor arise mainly from “lower vertex” charm production, meaning that the lepton is a decay product of a D meson with different charge than the B. According to the MC prediction, wrong-tag secondaries occur in 4.4% of muon tagged events and in 4.3% of electron tagged events. The systematic uncertainty on this factor uses the fact that 87 % of all wrong-tag secondaries arise from lower vertex charm, with smaller contributions coming from J/ψ and τ decays. The branching fraction of $B \to \overline{D}X$ decays is measured wit a 5% precision [104], while semileptonic $D \to X\ell\nu$ decays have been measured to a 3% precision [105]. A 30% uncertainty to the remainder of wrong-tag secondaries is assigned.

The average of electrons and muons yields a probability $\omega_{\text{sec.}} = 0.0431 \pm 0.0036$.

The fraction of misidentified hadrons in the MC is corrected to match the data by studying $D^{*+} \to D^0(K^-\pi^+)\pi^+$ decays, as described in Section 6.1.2. With the corrected misidentified hadron yield, the MC prediction is used to determine the fraction of them that produce a wrong tag. A 50% error is assumed as uncertainty, obtaining $\omega_{\text{misID}} = 0.0069 \pm 0.0034$. The different wrong-tag probabilities are summarized in Table 8.2, the total wrong-tag probability is $\omega_{\text{total}} = 0.1332 \pm 0.0052$.

Factor	Value
ω_{misID}	0.0069 ± 0.0034
ω_{2nd}	0.0431 ± 0.0036
ω_{osc}	0.0832 ± 0.0015
ω_{total}	0.1332 ± 0.0052

Table 8.2: Wrong-tag probabilities.

8.4 Results

Similar to the measurement of $\mathcal{B}_{(s+d)\gamma}$, the photon energy spectrum is obtained by subtracting the different photon backgrounds from the selected sample. The selection and corrections to the MC and off-resonance data are identical, except for the corrections corresponding to hadronic photon background and the category “others”, presented in Sections 6.2.1 and 6.2.2 which were not applied for this analysis. The spectra for positive and negative tags and for the sum of both are presented in Fig. 8.9. The uncertainties in the plots are the combination of statistical and systematics from background subtraction. These systematic uncertainties, however, do not play a role in the measurement of $\mathcal{A}^{\text{CP}}$. In the right plot of Fig. 8.9 statistical and total uncertainties are plotted. No striking differences can be observed between the positive and negative tagged spectra, the slight excess of events observed in the lower tail of the spectrum is due to the different corrections applied to the MC. The yield on the low-energy sideband is consistent with zero within 2.4σ. The uncertainties of the low-energy bins have large correlations.

The asymmetry can be calculated taking the integral of the spectra between a certain energy threshold and up to 2.8 GeV, to obtain the number of positive and negative tagged events (N^+ and N^-) and make use of the second part of Eq. (8.1). This “measured” asymmetry ($\mathcal{A}^{\text{CP}}_{\text{meas}}$) is then corrected using all the effects described above and the “true” asymmetry ($\mathcal{A}^{\text{CP}}_{(s+d)\gamma}$) can be obtained. The “measured” and “true” results are shown in Fig. 8.10. In this figure statistical and total uncertainties are presented. In order to account for possible differences in the asymmetry due to the excess of events at low energy, the $B\overline{B}$ MC is normalized to match the data and the asymmetry, the difference between this value and the value from the default selection is taken as an additional systematic uncertainty and is called “$B\overline{B}$ norm”. The results are summarized in Table 8.3 and a breakdown of uncertainties entering the measurement is presented in Table 8.4. These results have been published in [106].

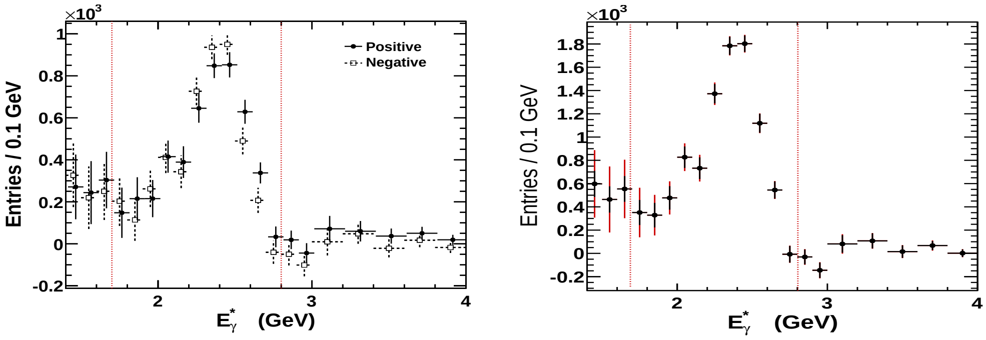


Figure 8.9: Extracted photon energy spectrum for positive and negative tagged events (left) and the sum of both spectra (right). The chosen signal region is delimited by the dotted lines. Statistical and systematic uncertainties are included in the left-hand-side plot, in the right-hand-side plot statistical uncertainties are plotted in red and total uncertainties in black.

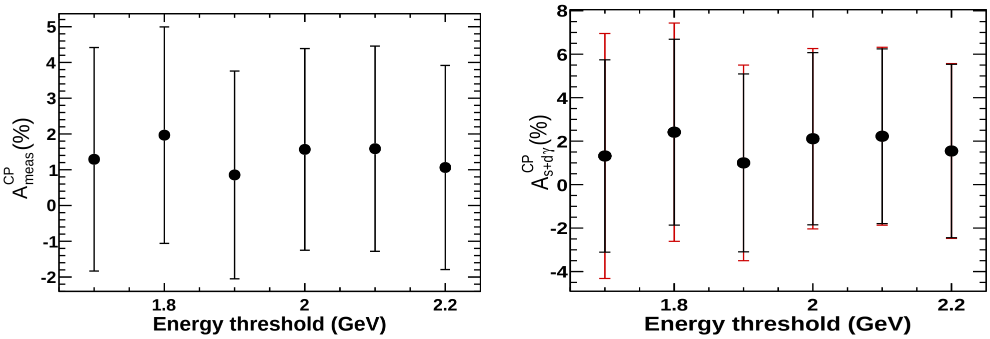


Figure 8.10: “Measured” and “true” values for the $B \to X_{s+d}\gamma$ CP asymmetry for different energy thresholds. For the result $\mathcal{A}^{\text{CP}}_{\text{meas}}$ only statistical uncertainties are plotted. For the “true” asymmetry ($\mathcal{A}^{\text{CP}}_{(s+d)\gamma}$) statistical and total uncertainties are shown in black and red respectively.

E^*_γ thresh.	$\mathcal{A}^{\text{CP}}_{\text{meas}}$	$\mathcal{A}_{\text{bkg}}$	$\mathcal{A}_{\text{det}}$	MC stat.	$B\overline{B}$ norm.	$\mathcal{A}^{\text{CP}}_{(s+d)\gamma}$
1.7 GeV	1.3 ± 3.1	-0.4 ± 2.5	0.1 ± 0.2	± 0.8	± 0.5	$2.2 \pm 4.3 \pm 3.5$
1.8 GeV	2.0 ± 3.0	-0.3 ± 1.9	0.1 ± 0.2	± 0.7	± 0.1	$3.0 \pm 4.1 \pm 2.7$
1.9 GeV	0.9 ± 2.9	-0.2 ± 1.3	0.1 ± 0.2	± 0.6	± 0.3	$1.4 \pm 4.0 \pm 1.9$
2.0 GeV	1.6 ± 2.8	-0.2 ± 0.9	0.1 ± 0.2	± 0.5	± 0.0	$2.2 \pm 3.8 \pm 1.3$
2.1 GeV	1.6 ± 2.9	-0.1 ± 0.5	0.1 ± 0.2	± 0.4	± 0.1	$2.2 \pm 3.9 \pm 0.9$
2.2 GeV	1.1 ± 2.9	-0.1 ± 0.3	0.1 ± 0.2	± 0.3	± 0.2	$1.4 \pm 3.9 \pm 0.6$

Table 8.3: Corrections in the calculation of $\mathcal{A}^{\text{CP}}_{(s+d)\gamma}$. The asymmetry $\mathcal{A}_{\text{bkg}}$ has been scaled by the factor B/S that can be read off in Fig. 8.8. All asymmetries and uncertainties are given in percent.

E_γ^* thresh.	Statistical	Total systematic	$\mathcal{A}_{\mathrm{det}}$	$\mathcal{A}_{\mathrm{bkg}}$	MC stat.	$B\overline{B}$ norm.	Wrong tag
1.70 GeV	4.26	3.52	0.30	3.40	0.76	0.42	0.02
1.80 GeV	4.13	2.72	0.30	2.56	0.68	0.53	0.05
1.90 GeV	3.96	1.92	0.30	1.81	0.58	0.10	0.02
2.00 GeV	3.84	1.32	0.30	1.17	0.48	0.19	0.04
2.10 GeV	3.91	0.86	0.30	0.70	0.39	0.12	0.04
2.20 GeV	3.89	0.59	0.30	0.41	0.30	0.04	0.03

Table 8.4: Summary of uncertainties in the measurement of $\mathcal{A}^{\mathrm{CP}}_{(s+d)\gamma}$. All uncertainties are given in percent.

CHAPTER 9

Discussion of results and constraints on New Physics

THE present results, using the complete Belle data sample, constitute the most precise and final Belle measurements of $B \to X_{s+d}\gamma$ and $B \to X_s\gamma$ decays. A large improvement for the branching fraction measurement has been reached, not only from the exploration of new tools and techniques for background suppression and estimation, but also from making use of all the knowledge gathered by the B-factories in previous measurements. In this chapter the results are discussed in more detail and the potential for future measurements of these decays at future B-factories is presented.

9.1 $B \to X_s\gamma$ branching fraction measurement

The branching fraction $\mathcal{B}_{s\gamma}$ measured in this theses is the most precise measurement if this decay mode to date. It can be compared with previous inclusive measurements from Belle and *BABAR* in Fig. 9.1 and with measurements using other techniques in Fig. 9.2. For the 1.7 GeV threshold, the total statistical uncertainty is 3.46 %, estimated from the number of on- and off-resonance events. Given that the off-resonance sample needs to be scaled up significantly to match the on-resonance statistics, it is clear that this uncertainty could be reduced with a comparatively larger continuum sample. If the off-resonance sample was of the same size as the expected continuum background in the on-resonance sample, the statistical uncertainty would drop to 2.65 %. It is, however, impossible to dedicate large part of the data taking to record background events, since no interesting physics will likely be found there. Even in the future Belle-II experiment one should count on having an off-resonance sample roughly 10 times smaller than the on-resonance one. However with the much larger samples at Belle-II, statistical uncertainties will become much smaller.

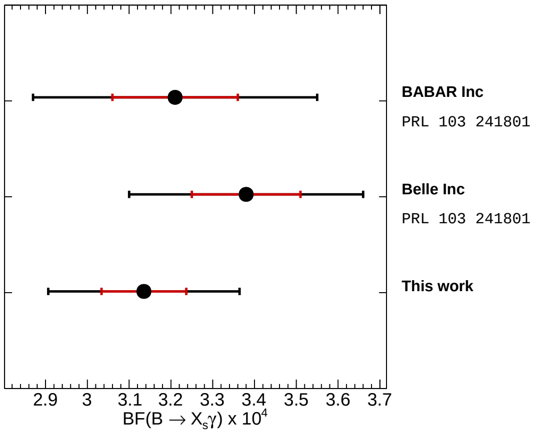


Figure 9.1: Previous results for $\mathcal{B}_{s\gamma}$ from inclusive measurements compared to the present measurement. The results are not extrapolated to lower E_{γ}^{B} of 1.6 GeV and are compared at the same threshold of 1.8 GeV. The inner error bars represent the statistical uncertainties and the external ones additionally include systematic and model uncertainties.

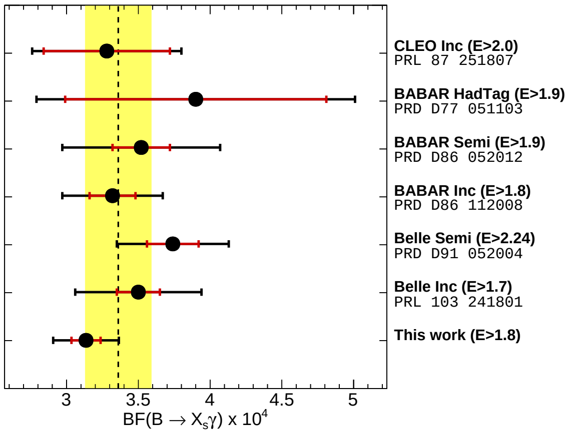


Figure 9.2: Previous results for $\mathcal{B}_{s\gamma}$ from all previous measurements compared to the present measurement. For each measurement the threshold used is indicated in parentheses, the measurements are extrapolated to 1.6 GeV. The theoretical prediction of [16] is shown by the yellow band. The inner error bars represent the statistical uncertainties and the external ones additionally include systematic, model and extrapolation uncertainties.

9.2 $B \to X_{s+d}\gamma$ spectrum and HQE parameters

The measured values for the first and second spectral moments of the E_γ^* spectrum can be compared to measurements by Belle [50] and *BABAR* [51]. The previous Belle analysis initially measured the moments in the CM frame after unfolding detector effects, and used conversion factors to quote values in the B frame. The values in the CM frame can also be found in the internal documentation [52] and are better suited for this comparison. One has to additionally consider that the previous Belle analysis was performed simultaneously for samples with and without a tag lepton, and the final result was presented as the average of both. For this reason the comparison of results is done for the averaged result and the result for the tagged sample. All available measurements for thresholds between 1.7 to 2.1 GeV are presented in Fig. 9.3.

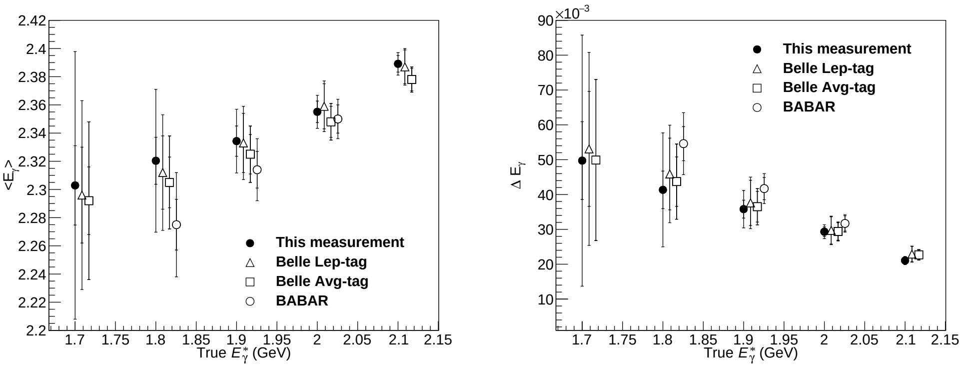


Figure 9.3: Spectral moments in the CM frame for different photon energy thresholds. This measurement (filled circles) is compared with previous determinations by Belle with a lepton tag (triangles) and the average of tagged and untagged analyses (squares) [50] and *BABAR* (open circles) [52]. The left and right plots show the first and second moments respectively. The *BABAR* measurement is only available for the 1.8, 1.9 and 2.0 GeV thresholds. The inner bars represent the statistical uncertainties and the external ones additionally include systematic and model uncertainties.

It is important to point out that although the measured spectrum in this analysis is more precise than that of [50], the measured moments are not more precise for the 1.7 and 1.8 GeV thresholds. This is because the moments are calculated from the efficiency-corrected spectrum, each bin i and bin error are calculated by dividing the background subtracted bin content by the selection efficiency on the bin $\varepsilon_{\text{sel},i}$. The energy dependence of the selection efficiency obtained in this analysis is more pronounced than that of the previous analysis, so the first few bins of the efficiency-corrected spectrum are less precise than in the previous analysis. This problem could be overcome by using a less aggressive continuum suppression, but this would in a less precise branching fraction and CP asymmetry. As was pointed out in Chapter 2, the photon energy spectrum is not expected to be sensitive to new physics, so measuring the branching fraction and CP asymmetry precisely is necessary.

The determination of the HQE parameters from a fit to the experimental background-subtracted spectrum represents an approach different to what has commonly used so far. Folding the theoretical prediction to compare it to the measured spectrum delivers a very precise measurement, yielding $m_b = \left(4.627^{+0.030}_{-0.031}\right)$ GeV and $\mu_\pi^2 = \left(0.301^{+0.070}_{-0.066}\right)$ GeV^2.

9.3 CP asymmetry measurement

The $\mathcal{A}^{\mathrm{CP}}_{(s+d)\gamma}$ measurement presented in this thesis is the most precise to date. All measurements of $\mathcal{A}^{\mathrm{CP}}$ in $B \to X_s\gamma$ and $B \to X_{s+d}\gamma$ decays are consistent with the SM prediction, i.e. no asymmetry. They are all statistically limited and the systematic uncertainties are irreducible for the measurements with the current detectors since they directly dependent on features of the detector and reconstruction algorithms. A summary of all $\mathcal{A}^{\mathrm{CP}}_{(s+d)\gamma}$ measurements is presented in Fig. 9.4. All measurements of $\mathcal{A}^{\mathrm{CP}}_{s\gamma}$ and $\mathcal{A}^{\mathrm{CP}}_{(s+d)\gamma}$ are presented in Fig. 9.5. The present measurement is quoted for the 2.1 GeV threshold to be consistent with the other inclusive determinations. As mentioned in Section 2.3.2, it is important to consider both observables, $\mathcal{A}^{\mathrm{CP}}_{(s+d)\gamma}$ and $\mathcal{A}^{\mathrm{CP}}_{s\gamma}$, since they are directly proportional in generic new physics models and in the SM. If one considers off-resonance samples of the same size as the continuum background in on-resonance data, the statistical uncertainty for the 2.1 GeV threshold would drop from 4.0 % to 3.1 %. As the sample sizes are the limiting factor, large improvements can be expected with Belle-II data.

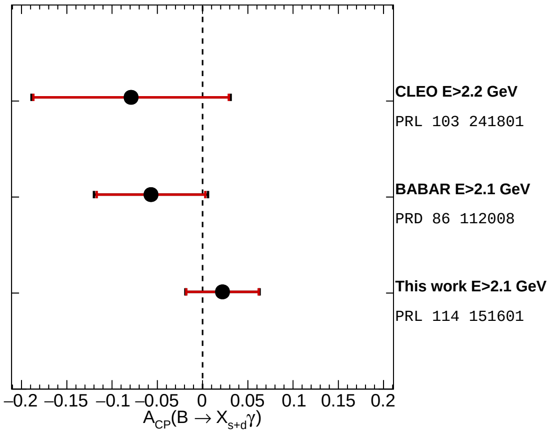


Figure 9.4: Previous $\mathcal{A}^{\mathrm{CP}}_{(s+d)\gamma}$ results from inclusive measurements compared with the present measurement from this thesis. The threshold used for each measurement is indicated in the figure. The SM prediction of zero, with vanishing uncertainties, is indicated by the dashed line.

9.4 Constraints on the Two-Higgs-Doublet Model of type II

The interpretation of the $\mathcal{B}_{s\gamma}$ results in the framework of new physics models has been discussed in Section 2.3.1. The simplified parametric form of Eq. (2.40), written once again here,

$$P(E_0) + N(E_0) = \left(C^{\mathrm{eff},(0)}_{7,\mathrm{SM}} + B\Delta C^{\mathrm{eff},(0)}_{7,H^+}\right)^2 + A,$$

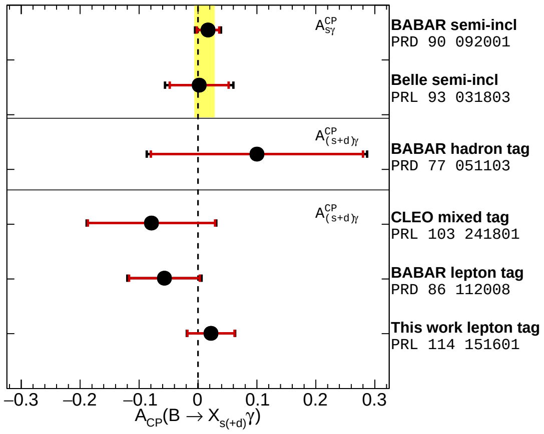


Figure 9.5: Previous $\mathcal{A}^{\mathrm{CP}}_{(s+d)\gamma}$ and $\mathcal{A}^{\mathrm{CP}}_{s\gamma}$ measurements compared to the present measurement. Additionally to the lepton tag, the CLEO result uses a pseudo-reconstruction technique to increase statistics. The present result is shown for the 2.1 GeV threshold. The theoretical prediction of [40] for $\mathcal{A}^{\mathrm{CP}}_{s\gamma}$ is given by the yellow band.

where the functions $P(E_0)$ and $N(E_0)$ represent the perturbative and non perturbative terms in the calculation of $\mathcal{B}_{s\gamma}$ requires the determination of the coefficients A and B that reproduce the full Two-Higgs-Doublet Model of type II (THDM-II) calculation. For this purpose the SusyBSG package [107] is used. The package provides calculations of $\mathcal{B}_{s\gamma}$ for the SM and for several new physics models, such as THDM-II and the MSSM. SusyBSG uses most NNLO QCD corrections, and all leading electroweak corrections in the calculation of $\mathcal{B}_{s\gamma}$, additionally, a complete NLO THDM-II calculation is available. A fit is performed to determine the coefficients A and B of the parametric model, such that the SusyBSG calculation of the THDM-II is reproduced.

The coefficients A and B depend on the input parameters in the $B \to X_s\gamma$ calculation: top quark mass and, W and Z weak boson masses and Λ_{QCD}, their values are the same to those used by [20] for consistency[†] . The coefficients can be expressed by their leading-order Taylor expansion around the parameter central values, in order to be able to update the parameters in the future:

$$X = X_0 * \left(1 + \partial_{m_c}\Delta m_c + \partial_{m_t}\Delta m_t + \partial_{m_Z}\Delta m_Z + \partial_{m_W}\Delta m_W + \partial_{\Lambda_{\mathrm{QCD}}}\Delta\Lambda_{\mathrm{QCD}}\right), \tag{9.1}$$

where X represents A or B, and Δp the difference between the central value of parameter p and the updated value.

This fit allows us to extract the $P(E_0) + N(E_0)$ part of Eq. (2.40) from the SusyBSG calculation, and also all factors ∂_p for the parameters p of Eq. (9.1). The comparison between the parametric model and the SusyBSG calculation can be seen in Fig. 9.6, the corresponding set of factors is summarized in Table 9.1.

[†] For reference, the used values are $m_c = 1.131$ GeV, $m_t = 173.21$ GeV, $m_W = 80.385$ GeV, $m_Z = 91.1876$ GeV and $\alpha_s = 0.1185$. In SusyBSG α_s is calculated from Λ_{QCD} for a renormalization scale μ as $\alpha_s(\mu, \Lambda_{\mathrm{QCD}}) = \frac{12\pi}{23 \cdot L(\mu)} \cdot \left(1 - \frac{348}{529} \cdot \frac{\log(L(\mu))}{L(\mu)}\right)$, where $L(\mu) = 2 \cdot \log(\mu/\Lambda_{\mathrm{QCD}})$

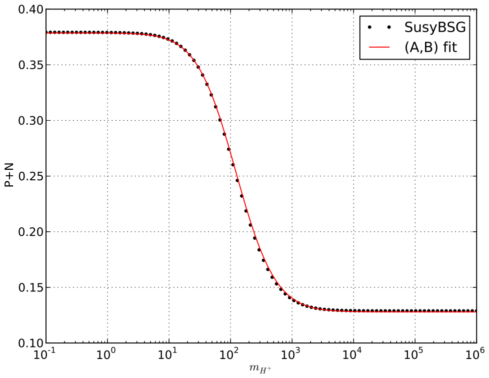


Figure 9.6: $P(E_0) + N(E_0)$ SusyBSG calculation and fitted parametric model as defined in Eq. (2.40), the values are given as a function of m_{H^+}.

X	X_0	∂_{m_c}	∂_{m_t}	∂_{m_Z}	∂_{m_W}	$\partial_{\Lambda_{\mathrm{QCD}}}$
A	-1.388×10^{-2}	6.606	9.514×10^{-4}	-4.665×10^{-3}	-2.648×10^{-2}	13.030
B	0.831	−0.198	-4.717×10^{-4}	1.508×10^{-5}	-1.543×10^{-3}	0.352

Table 9.1: Coefficients that enter the calculation of the $B \to X_s\gamma$ THDM-II parametric model from Eqs. (2.40) and (9.1).

In addition to quark and weak boson masses, CKM matrix elements enter the $\mathcal{B}_{s\gamma}$ calculation. The CKM matrix elements are constrained in the SM using flavor measurements, as presented in Fig. 2.2. Any new physics model that is considered would not affect a single decay, but would have a affect several measurements. For this reason that while the CKM matrix can be very precisely constrained from several flavor-physics measurements. When trying to draw conclusions about the THDM-II, most of them can no longer be utilized since. The charged Higgs contribution would appear at leading order in loop- and box-mediated processes such as $B_{s,d}$ mixing. The angles of the Unitary Triangle cannot be used since they come from CP violation measurements, using interference between decay and mixing. A reduced set of measurements, where the THDM-II has negligible impact, is used to constraint the $(\rho - -\eta)$ apex of the Unitary Triangle [44]:

- the determination of the angle γ from $\alpha + \beta$ Where the contributions with box diagrams cancel. THDM-II contributions enter as $\left(m_b \cdot m_u/m^2_{H^+},\ m_b \cdot m_d/m^2_{H^+}\right)$.
- the determination of $|\,\mathrm{V}_{cb}\,|$ from $B \to X_c\ell\nu$ decays. THDM-II contributions appear at tree level but are negligible since they enter as $\left(m_b \cdot m_e/m^2_{H^+},\ m_c \cdot m_e/m^2_{H^+}\right)$.
- the determination of $|\,\mathrm{V}_{ub}\,|$ from $B \to X_u\ell\nu$ decays. THDM-II contributions appear at tree level but are negligible since enter as $\left(m_b \cdot m_u/m^2_{H^+},\ m_b \cdot m_d/m^2_{H^+}\right)$.
- the determination of $|\,\mathrm{V}_{ud}\,|$ from super-allowed β decays of nuclei where no heavy mass is involved.

The values of these observables are summarized in Table 9.2. When they are used as the only inputs to constraint the Unitary Triangle, the obtained constraint is weaker than that of the global fit and shows a fourfold multiplicity as seen in Fig. 9.7. The fit is performed making with the statistical framework developed by the CKMfitter collaboration [108], which uses a frequentist approach (Rfit [109]).

Parameter	Value	Source
V_{ud}	0.97425 ± 0.00022	From [110]
V_{cb}	$(41.00 \pm 0.33 \pm 0.74) \times 10^{-3}$	CKMfitter average
V_{ub}	$(3.70 \pm 0.12 \pm 0.26) \times 10^{-3}$	CKMfitter average
$\cos 2\beta$	≥ 0	CKMfitter constraint
$\sin 2\beta$	0.682 ± 0.019	CKMfitter average
α	$87.7^{+3.5}_{-3.2} \cup 0.1^{+2.5}_{-\infty}$	CKMfitter scan,
	$\cup\, 179.0^{+\infty}_{-4.1}$	from direct measurements

Table 9.2: Input measurements to the Unitarity Triangle fit for the THDM-II analysis.

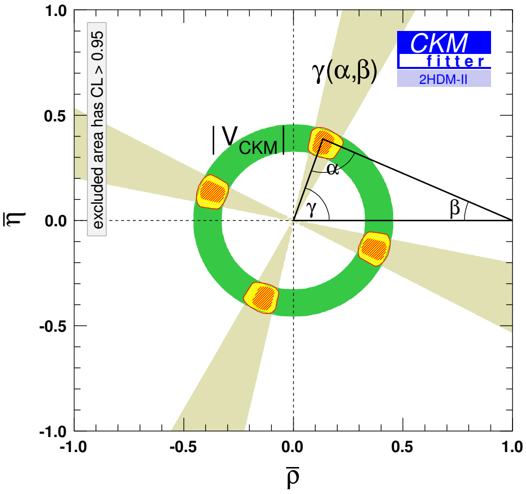


Figure 9.7: Constraints on the Unitarity Triangle for the THDM-II analysis.

For the constraints of the THDM-II, the extrapolated $\mathcal{B}_{s\gamma}$ result for the 1.8 GeV threshold is used. This result is compared with the theoretical prediction of the parametric model, and the Unitarity triangle is constrained as described above. The $B \to X_s\gamma$ branching fraction is normalized to the $B \to X_c\ell\nu$ branching fraction whose world average is $\mathcal{B}_{c\ell\nu} = (10.65 \pm 0.16) \times 10^{-2}$ [104]. From Eq. (2.18), the phase space factor C takes the value $C = 0.568 \pm 0.012$ [111]. Using the CKMfitter framework to perform the fit, one obtains a lower bound on the mass of the THDM-II charged Higgs boson of $m_{H^+} \gtrsim 580\,\mathrm{GeV}$ with a 95 % confidence level. The fit result is presented in Fig. 9.8 the colored region has not been excluded by the measurement. The constraint derived from this indirect search is much more stringent than those by direct searches at ATLAS [112, 113] or previously at LEP [114]. Both results are also included in Fig. 9.8.

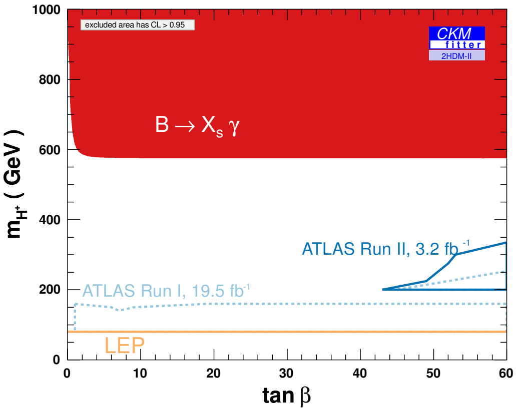


Figure 9.8: Exclusion region in m_{H^+} vs. $\tan\beta$ at 95 % confidence level. This study uses the measured $\mathcal{B}_{s\gamma} = 3.12 \pm 0.10(\text{stat}) \pm 0.19(\text{syst}) \pm 0.08(\text{model}) \pm 0.04(\text{extrap})$. The colored region is not excluded by this measurement, the constraint from $B \to X_s\gamma$ is much more stringent than those from direct searches at ATLAS and LEP. The LEP limit is $m_{H^+} > 80\,\text{GeV}$

9.5 Projections for measurements at Belle II

The Belle II experiment is the successor of Belle and is currently in its final preparation stages. Data taking is planned to begin in 2018. The aim is to record ~50 ab^{-1} of data within five years, i.e. roughly 70 times more than Belle, assuming all the recorded data is taken at the $\Upsilon(4S)$. The doors to many new rare physics processes will be open. Radiative decays will remain one of the most interesting decay modes to look for new physics. The first and most obvious improvement with Belle-II is that the statistical precision for measurements of $\mathcal{A}^{\text{CP}}_{(s+d)\gamma}$ and $\mathcal{B}_{(s+d)\gamma}$ and the photon energy spectrum shape will be greatly improved. Second, some of the systematic uncertainties of the branching fraction scale with luminosity and will thus be reduced. Third, full reconstruction techniques as described in Section 3.3 could be applied, making it possible to have a precise measurement of all the observables with an approach with much less background. The measurement of isospin asymmetries and the difference of CP asymmetries between charged and neutral B decays ($\Delta\mathcal{A}^{\text{CP}}$) will also possible.

A few assumptions are made in the following discussion: the relative sizes of on- and off-resonance samples are assumed to be the same as for Belle, as well as the number of available MC streams; several systematic uncertainties could be improved with the Belle-II detector, e.g. those related to tracking and particle identification algorithms, which means that the uncertainties related to the lepton tag could improve significantly. Assessing these improvements is however, beyond the scope of this study and these uncertainties are assumed to remain unchanged.

Systematic uncertainties in $\mathcal{B}_{(s+d)\gamma}$ that could improve with larger data and MC sample sizes include uncertainties related to the MC statistics and the statistical uncertainty of the selection efficiency. Other uncertainties such as those coming from the normalizations of π/η background, the π/η veto efficiencies and the estimation of the "others" background component scaling, could also improve. These effects include irreducible uncertainties from e.g. MC templates for the fits, but could also marginally improve with more data. All other effects quoted in Appendix A.2 are assumed to be irreducible.

All uncertainties related to the continuum background could improve if better continuum-suppression techniques are devised. It is expected that the studies performed here are a good starting point for future analysis. In addition, a continuum background MC that reliably describes the data would be very beneficial for the analysis since it can be produced in large quantities. This was unfortunately not the case for Belle.

In Fig. 9.9 the $\mathcal{B}_{s\gamma}$ projections for different data sample sizes at Belle-II are presented, for the measurement at the 1.8 GeV threshold. The statistical uncertainty could drop down to $\frac{1}{8}$ of the current value with a full data set. Systematic uncertainties will plateau rapidly and are not expected to drop below $\sim 5.5\,\%$. At this stage the experimental uncertainties become of the same magnitude as the irreducible theoretical uncertainties, so a significant advancement on both the theory side and the experimental techniques would be needed to study this decay more precisely. The experimental side would in any case profit from averaging several measurements with different techniques.

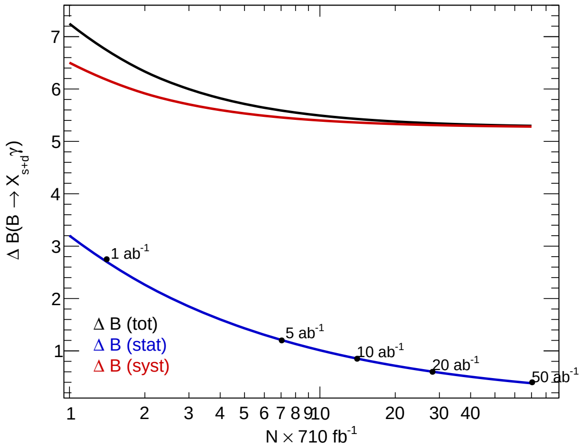


Figure 9.9: Projection of the uncertainties on $\mathcal{B}_{(s+d)\gamma}$ for the 1.8 GeV threshold for various data sample sizes. The x-axis (in logarithmic scale) shows multiples of the recorded Belle integrated luminosity. The uncertainties on the measurement on the y-axis are given in percent.

Systematic uncertainties in $\mathcal{A}^{\mathrm{CP}}_{(s+d)\gamma}$ are irreducible while statistical uncertainties scale with the data sample size. It is expected that statistical uncertainties will be reduced to a third of the current size with only $\sim 5\,\mathrm{ab}^{-1}$ and become of the same size as the systematic uncertainties with $\sim 20\,\mathrm{ab}^{-1}$, as shown in Fig. 9.10. Even at the end of Belle-II data taking, the total uncertainty is not expected to drop below $\sim 1\,\%$, being at this point systematically limited. This precision might be enough to discard very generic models, but not the more constrained MSSM scenarios. It is thus clear that one cannot rely on $\mathcal{A}^{\mathrm{CP}}_{(s+d)\gamma}$ alone to study new CP-violating effects, but certainly also $\mathcal{A}^{\mathrm{CP}}_{s\gamma}$ and $\mathcal{A}^{\mathrm{CP}}_{d\gamma}$. As pointed out previously, the three asymmetries $\mathcal{A}^{\mathrm{CP}}_{s\gamma}$, $\mathcal{A}^{\mathrm{CP}}_{d\gamma}$ and $\mathcal{A}^{\mathrm{CP}}_{(s+d)\gamma}$ are closely related and are proportional to one another in new physics models. For this reason it is important to study all of them, using different measurements, in order to provide constraints on new physics models from different sides.

The flavor specific $\mathcal{A}^{\mathrm{CP}}_{s\gamma}$ is already more precisely determined than $\mathcal{A}^{\mathrm{CP}}_{(s+d)\gamma}$ and it can be expected to be further improved with larger statistics. One should not forget that $\mathcal{A}^{\mathrm{CP}}_{(s+d)\gamma}$ could also be measured with the recoil-tag method, so the combination of all these new measurements could potentially push the uncertainties below $1\,\%$. The Belle-II scenario for new physics in $\mathcal{A}^{\mathrm{CP}}_{(s+d)\gamma}$ and $\mathcal{A}^{\mathrm{CP}}_{s\gamma}$ is illustrated in Fig. 9.11, where the current experimental precision is shown together the SM and MSSM predictions.

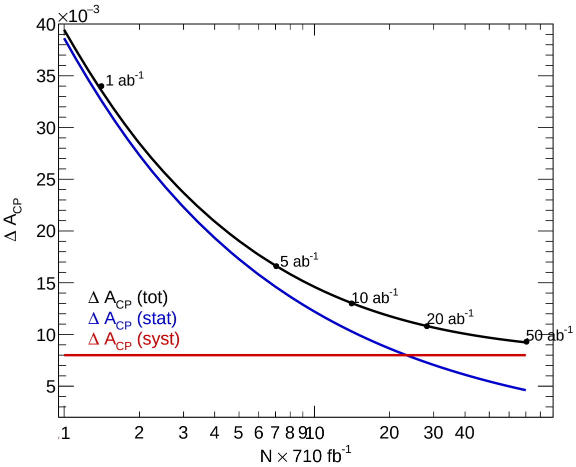


Figure 9.10: Projection of the uncertainties on $\mathcal{A}^{\mathrm{CP}}_{(s+d)\gamma}$ for the 2.1 GeV threshold for various data sample sizes. The x-axis (in logarithmic scale) shows multiples of the recorded Belle integrated luminosity. The uncertainties on the measurement on the y-axis is given in percent.

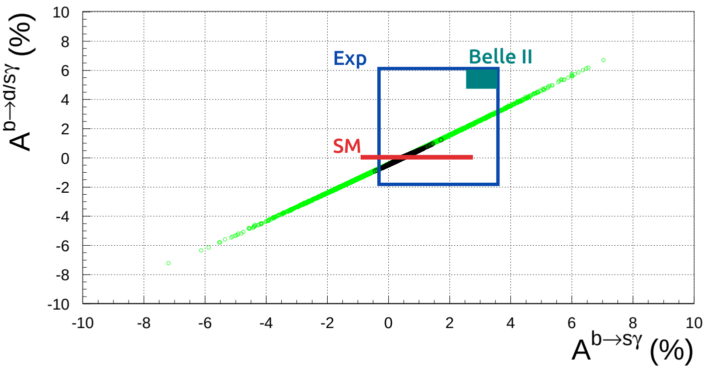


Figure 9.11: *CP* asymmetries $\mathscr{A}_{s\gamma}^{\mathrm{CP}}$ (x-axis) and $\mathscr{A}_{(s+d)\gamma}^{\mathrm{CP}}$ (y-axis) in MFV models compared with the SM expectation (red), current measurement (blue box) and Belle II projected precision (cyan filled box), [from 45, (Figure 2, modified)]. The central value and uncertainties of the presently most precise measurements are used as for the current measurements. For the projected result, projected uncertainties are used, while the central value is arbitrary.

CHAPTER 10

Conclusions and outlook

The decays $B \to X_{s+d}\gamma$ and $B \to X_s\gamma$ have been studied in this thesis, using 710 fb^{-1} of data collected at the $\Upsilon(4S)$ by the Belle experiment. The decays $B \to X_{s+d}\gamma$ and $B \to X_s\gamma$ are flavor-changing-neutral-currents that proceed through loop diagrams in the Standard Model (SM). They are sensitive to particles from New Physics models. The $B \to X_s\gamma$ branching fraction, $\mathcal{B}_{s\gamma}$, has been calculated to a precision of 7 % and strongly contrains certain new physics models. The CP asymmetry in $B \to X_{s+d}\gamma$, $\mathcal{A}^{\mathrm{CP}}_{(s+d)\gamma}$, is predicted in the SM to be zero, with negligible uncertainties. Measurements of $\mathcal{A}^{\mathrm{CP}}_{(s+d)\gamma}$ are complementary to those of the CP asymmetry in $B \to X_s\gamma$. Interesting information about the kinematics of the b quark inside the B-meson can be extracted from the $B \to X_s\gamma$ photon energy spectrum. The spectrum is not expected to be affected by new physics, but the information it delivers can be used to improve our knowledge of SM parameters, relevant to other B decays.

In this thesis the inclusive approach is used to study $B \to X_{s+d}\gamma$, meaning that only the photon from the decay is reconstructed. At the Belle experiment B mesons produced in pairs from the reaction $e^-e^+ \to \Upsilon(4S) \to B\overline{B}$. The partner B-meson is used in this analysis to suppress continuum background ($e^-e^+ \to q\overline{q}$). For this purpose, a lepton from the partner B is reconstructed. Suppressing continuum background is one of the major challenges of this analysis, since it is much larger than the expected signal and only small samples are available to model it. Multivariate analysis techniques were used to achieve an effective continuum suppression. It was necessary to study several variables that characterize the spatial distribution of particles from the decay and that yield good separation between the interesting $B\overline{B}$ events and $q\overline{q}$. The continuum suppression was very efficient and made it possible to greatly reduce uncertainties related to continuum background, compared to the previous Belle analysis.

The background-subtracted photon energy spectrum is used to measure the branching fractions $\mathcal{B}_{(s+d)\gamma}$ and $\mathcal{B}_{s\gamma}$, the CP asymmetry, the first and second spectral moments and to extract the HQE parameters m_b and μ_π^2 from the shape-function model. The branching fractions and CP asymmetry have been measured for different photon energy thresholds ranging from 1.6 GeV to 2.2 GeV. Theoretical predictions of $\mathcal{B}_{s\gamma}$ are given for a photon energy threshold of 1.6 GeV, experimentally low thresholds are challenging since the background is large for low photon energies. Experimental results rely thus on extrapolation factors to express $\mathcal{B}_{s\gamma}$ at the 1.6 GeV threshold. The extrapolation factors rely on a theoretical description of the photon energy spectrum and are thus model dependent. For this reason it is important to present results for several thresholds, such that new theoretical descriptions could in the future be used as well. The measured $\mathcal{B}_{s\gamma}$ is consistent with SM predictions and in agreement with previous measurements.

The SM predicts that the value of $\mathcal{A}^{\mathrm{CP}}_{(s+d)\gamma}$ does no dependent on the photon energy threshold. $\mathcal{A}^{\mathrm{CP}}_{(s+d)\gamma}$ results are presented for several energy thresholds to test this prediction. The measured $\mathcal{A}^{\mathrm{CP}}_{(s+d)\gamma}$ is consistent with the null CP asymmetry predicted in the SM and in agreement with previous measurements. The measured asymmetry is also independent of the chosen energy threshold.

The extraction of the m_b and μ_π^2 parameters uses an approach different to what has been usually done. A folding procedure is used in order to transform theoretical spectra from the shape function model into spectra that include distortion from detector resolution, Doppler broadening and the selection efficiencies. These new spectra can be directly compared to the background-subtracted data and the pair of m_b and μ_π^2 parameter values that best describe it can be determined from a fit.

A representative summary of the results is presented in the table below. The measured branching fractions, CP asymmetry and spectral moments are the most precise measurements to date. The determination of m_b and μ_π^2 follows a novel approach and yields a more precise result than the one obtained from the traditional approach. These are the legacy Belle results using the inclusive method, since they are obtained using the full Belle data sample. With the measured value of $\mathcal{B}_{s\gamma}$, it is possible to obtain a lower bound on m_{H^+}, of $m_{H^+} \gtrsim 580\,\mathrm{GeV}$, which is currently the most constraining limit.

Possible improvements that could be implemented in analyses using an improved detector and larger data samples, as expected for the Belle II experiment, have been discussed. An estimation of experimental uncertainties for larger data samples has been presented. Some methods used in this thesis have been developed over several years, some have been developed in this analysis. The methodology described in this thesis lays the ground for future analyses at the Belle II experiment.

Observable	Value	Comment
$\mathcal{B}_{s\gamma}$	$(3.12 \pm 0.10 \pm 0.19 \pm 0.08 \pm 0.04) \times 10^{-4}$	Extrapolated from 1.8 GeV
$\langle E^*_\gamma \rangle$	$(2.320 \pm 0.016 \pm 0.047 \pm 0.001)$ GeV	For $E^*_\gamma \geq 1.8$ GeV
ΔE^*_γ	$(4.258 \pm 0.525 \pm 1.498 \pm 0.050) \times 10^{-2}$ GeV^2	For $E^*_\gamma \geq 1.8$ GeV
$\mathcal{A}^{\mathrm{CP}}_{(s+d)\gamma}$	$(2.2 \pm 3.9 \pm 0.9)\%$	For $E^*_\gamma \geq 2.1$ GeV
m_b	$\left(4.627^{+0.030}_{-0.031}\right)$ GeV	For shape function model
μ_π^2	$\left(0.301^{+0.070}_{-0.066}\right)$ GeV^2	For shape function model
m_{H^+}	$\gtrsim 580$ GeV	Limit with 95 % C.L.

Table 10.1: Summary of measurements presented in this study. The uncertainties in the $\mathcal{B}_{s\gamma}$ are statistical, systematic, due to model dependence and extrapolation. The uncertainties in the measurement of the photon energy spectrum moments are statistical, systematic and due to model dependence. The uncertainties in $\mathcal{A}^{\mathrm{CP}}_{(s+d)\gamma}$ are statistical and systematic. For the m_b and μ_π^2 measurements, the fit error contains the total experimental uncertainty (statistical and systematic). The lower bound on m_{H^+} of the THDM-II is derived from the measurement of $\mathcal{B}_{s\gamma}$.

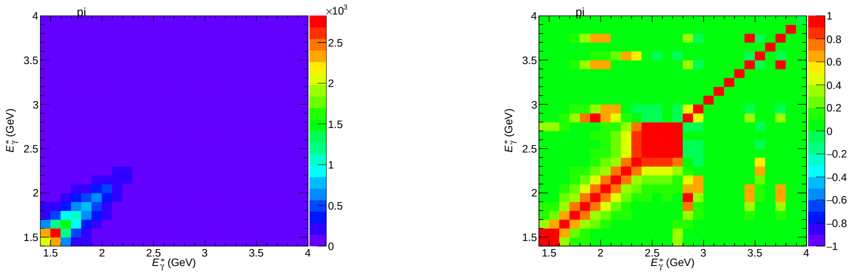
APPENDIX A

Systematic uncertainties

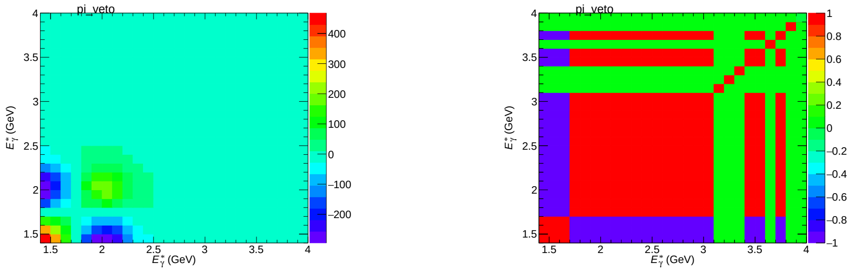
A.1 Background subtraction

In Table A.1 the systematic covariance and correlation matrices for different background sources are presented. These matrices correspond to the uncertainties on the E^*_γspectrum.

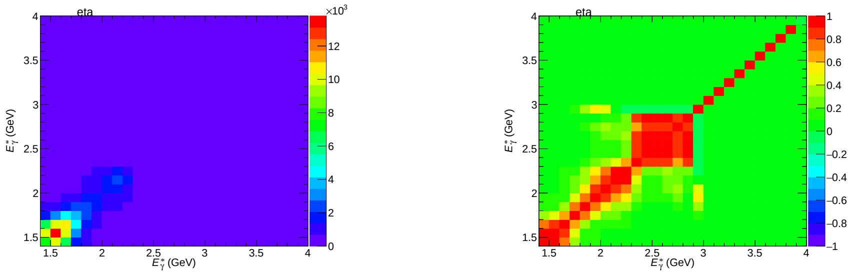
(a)



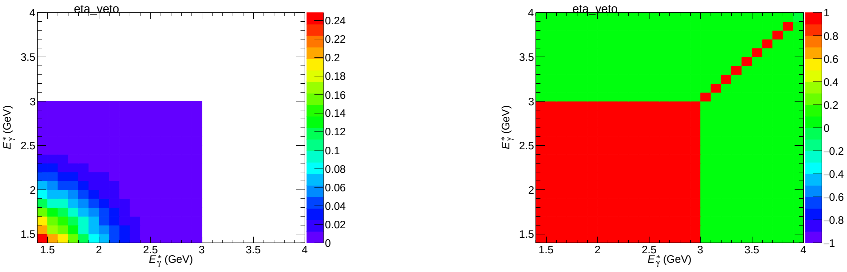
(b)



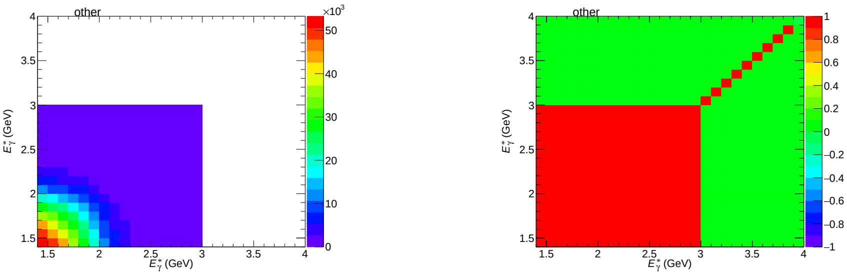
(c)



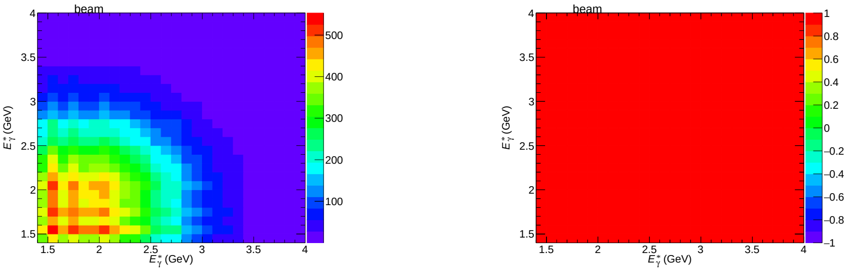
(d)


eta_veto
E^*_γ (GeV)
1.5
2
2.5
3
3.5
4
0.24
0.22
0.2
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
eta_veto
1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−1

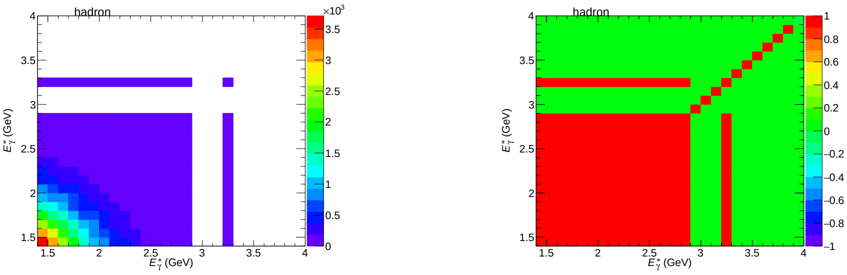
(e)


other
E^*_γ (GeV)
1.5
2
2.5
3
3.5
4
$\times10^3$
50
40
30
20
10
0
other
1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−1

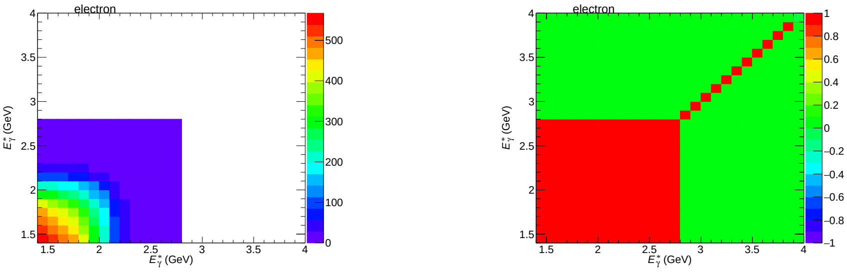
(f)


beam
E^*_γ (GeV)
1.5
2
2.5
3
3.5
4
500
400
300
200
100
beam
1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−1

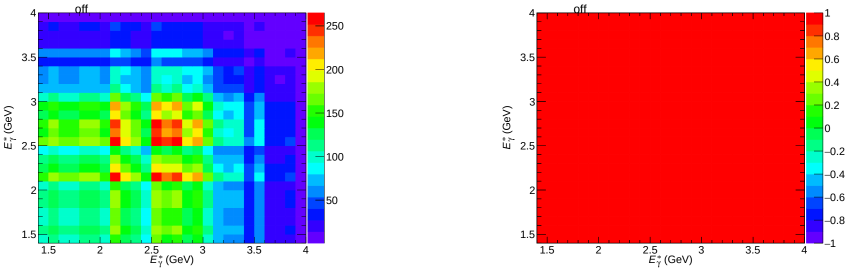
(g)


hadron
E^*_γ (GeV)
1.5
2
2.5
3
3.5
4
$\times10^3$
3.5
3
2.5
2
1.5
1
0.5
0
hadron
1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−1

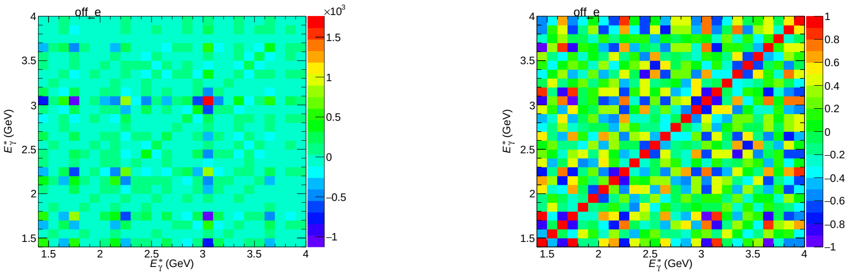
(h)


electron
E^*_γ (GeV)
1.5
2
2.5
3
3.5
4
500
400
300
200
100
0
electron
1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−1

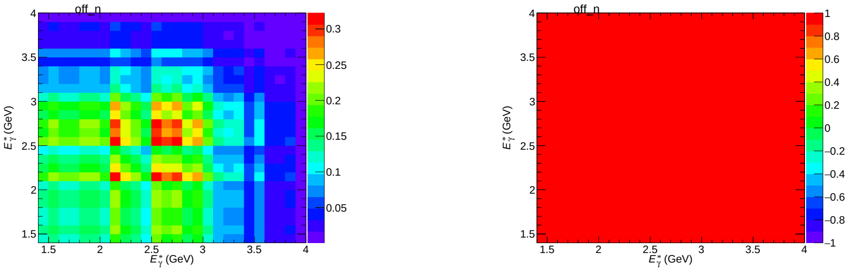
(i)


off
E^*_γ (GeV)
E^*_γ (GeV)
1.5
2
2.5
3
3.5
4
250
200
150
100
50
off
1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−1

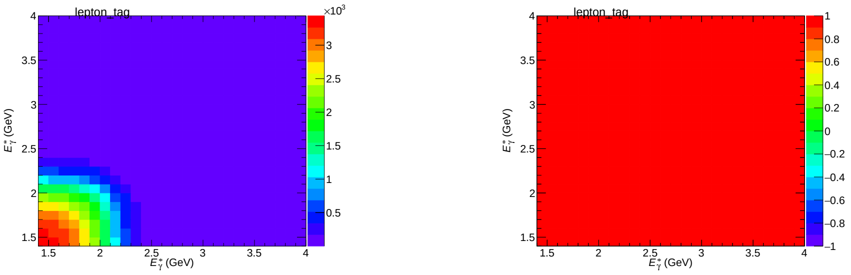
(j)


off_e
$\times10^3$
1.5
1
0.5
0
−0.5
−1
E^*_γ (GeV)
E^*_γ (GeV)
off_e

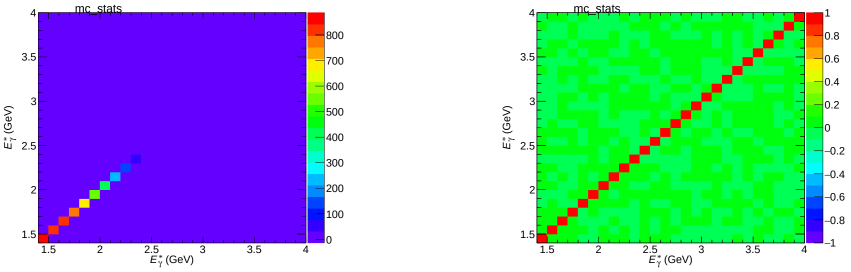
(k)


off_n
0.3
0.25
0.2
0.15
0.1
0.05
E^*_γ (GeV)
E^*_γ (GeV)
off_n

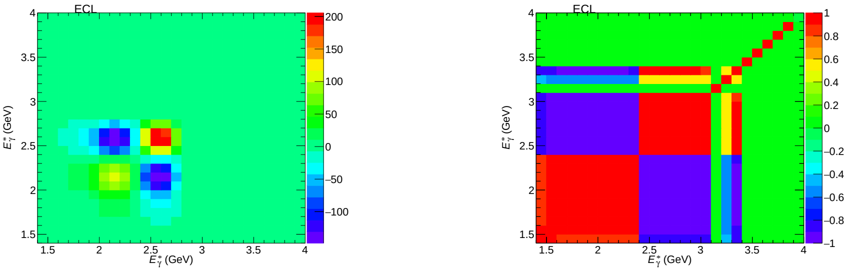
(l)


lepton_tag
$\times10^3$
3
2.5
2
1.5
1
0.5
E^*_γ (GeV)
E^*_γ (GeV)
lepton_tag

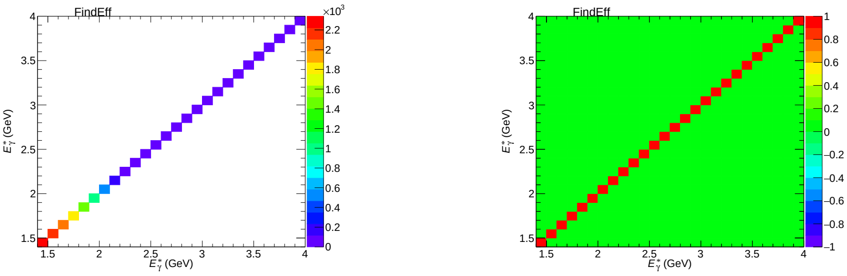
(m)


mc_stats
800
700
600
500
400
300
200
100
0
E^*_γ (GeV)
E^*_γ (GeV)
mc_stats

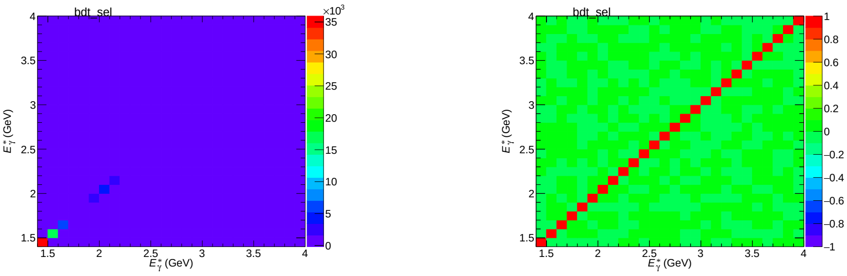
`(n)`



`(o)`



`(p)`



`(q)`

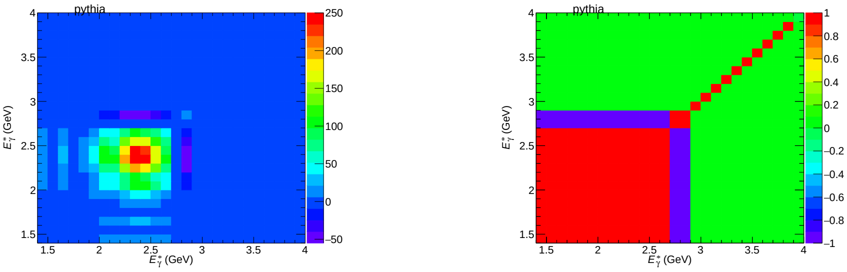


Table A.1: Systematic covariance (left) and correlation (right) matrices for different contributions: `(a)` calibration of $\pi^0 \to \gamma\gamma$ background, `(b)` efficiency of π^0 veto selection, `(c)` calibration of $\eta \to \gamma\gamma$ background, `(d)` efficiency of η veto selection, `(e)` yield of other photon background, `(f)` yield of beam background, `(g)` correction of hadronic background, `(h)` yield of electron background, `(i)` scaling factor for off-resonance data, `(j)` correction for energy of off-resonance particles F_E, `(k)` correction for off-resonance particle multiplicity F_N, `(l)` corrections to the tag lepton spectrum, `(m)` MC statistics, `(n)` photon energy resolution, `(o)` photon finding efficiency, `(p)` BDT efficiency of $B\overline{B}$ background, `(q)` PYTHIA hadronization.

A.2 Uncertainties in the measurement of branching fractions

In the following tables the uncertainties in the measurements of the branching fractions $\mathcal{B}_{(s+d)\gamma}$, $\mathcal{B}_{s\gamma}$ and $\mathcal{B}_{s\gamma}$ extrapolated to 1.6 GeV are shown. The uncertainties are presented in percentage in Tables A.2 to A.4, and each contribution is presented as an absolute uncertainty in Table A.5.

E_γ^B	$\mathcal{B}_{(s+d)\gamma}$	Total	Stat	Syst	Model
1.60	3.28	11.3	3.6	10.1	3.5
1.70	3.18	9.1	3.5	7.9	3.0
1.80	3.14	7.3	3.2	6.0	2.5
1.90	3.06	6.1	3.0	4.7	2.3
2.00	2.90	5.1	2.9	3.9	1.8
2.10	2.61	4.6	2.9	3.3	1.5

Table A.2: Branching fraction $\mathcal{B}_{(s+d)\gamma}$ for different thresholds in units of 10^{-4}. Uncertainties in percentage.

E_γ^B	$\mathcal{B}_{s\gamma}$	Total	Stat	Syst	Model
1.60	3.13	11.3	3.6	10.1	3.5
1.70	3.04	9.1	3.5	7.9	3.0
1.80	3.00	7.3	3.2	6.1	2.5
1.90	2.92	6.1	3.0	4.8	2.3
2.00	2.77	5.2	2.9	3.9	1.8
2.10	2.49	4.7	2.9	3.3	1.5

Table A.3: Branching fraction $\mathcal{B}_{s\gamma}$ for different thresholds in units of 10^{-4}. Uncertainties in percentage.

E_γ^B	$\mathcal{B}_{s\gamma}$ at 1.6 GeV	Total	Stat	Syst	Model	Extrap
1.60	3.13	11.3	3.6	10.1	3.5	0.0
1.70	3.09	9.2	3.5	7.9	3.0	0.6
1.80	3.11	7.4	3.2	6.0	2.5	1.3
1.90	3.13	6.5	3.0	4.7	2.3	2.2
2.00	3.11	6.2	2.9	3.9	1.8	3.4
2.10	3.01	6.8	2.9	3.3	1.5	4.9

Table A.4: Branching fraction $\mathcal{B}_{s\gamma}$ extrapolated to 1.60 GeV for different thresholds in units of 10^{-4}. Uncertainties in percentage.

E_γ^B	$\mathcal{B}_{(s+d)\gamma}$	Total	Stat	Factor	EffMod	EffSta	EffRec	NBB	ECL	FindEff	BDT	Beam	Cont	Electr	Eta	Etavet	Hadron	Leptag	MCstat	Other	Pi	Pivet	Pythia
1.60	3.28	0.37	0.12	0.01	0.12	0.01	0.00	0.04	0.00	0.03	0.05	0.08	0.06	0.03	0.10	0.00	0.06	0.09	0.02	0.24	0.05	0.02	0.03
1.70	3.18	0.29	0.11	0.01	0.10	0.01	0.00	0.04	0.00	0.03	0.04	0.07	0.05	0.02	0.08	0.00	0.04	0.07	0.02	0.17	0.04	0.02	0.03
1.80	3.14	0.23	0.10	0.01	0.08	0.01	0.00	0.04	0.00	0.02	0.04	0.06	0.05	0.02	0.07	0.00	0.03	0.06	0.02	0.12	0.03	0.02	0.03
1.90	3.06	0.19	0.09	0.01	0.07	0.01	0.00	0.04	0.00	0.02	0.04	0.05	0.04	0.01	0.06	0.00	0.02	0.04	0.01	0.07	0.03	0.02	0.03
2.00	2.90	0.15	0.08	0.01	0.05	0.01	0.00	0.04	0.00	0.01	0.04	0.05	0.04	0.01	0.05	0.00	0.01	0.03	0.01	0.04	0.02	0.01	0.03
2.10	2.61	0.12	0.08	0.01	0.04	0.01	0.00	0.04	0.01	0.01	0.03	0.04	0.04	0.00	0.04	0.00	0.01	0.01	0.01	0.03	0.02	0.01	0.03

Table A.5: Branching fraction $\mathcal{B}_{(s+d)\gamma}$ for different thresholds in units of 10^{-4}.

E_γ^B	$\mathcal{B}_{(s+d)\gamma}$	Total	Stat	Factor	EffMod	EffSta	EffRec	NBB	ECL	FindEff	BDT	Beam	Cont	Electr	Eta	Etavet	Hadron	Leptag	MCstat	Other	Pi	Pivet	Pythia
1.60	3.28	11.3	3.6	0.3	3.5	0.4	0.0	1.4	0.0	1.0	1.6	2.3	1.7	1.0	3.1	0.0	1.7	2.9	0.7	7.3	1.4	0.5	0.9
1.70	3.18	9.1	3.5	0.2	3.0	0.4	0.0	1.4	0.0	0.8	1.4	2.2	1.7	0.8	2.5	0.0	1.3	2.4	0.6	5.4	1.2	0.6	0.9
1.80	3.14	7.3	3.2	0.3	2.5	0.4	0.0	1.4	0.1	0.7	1.4	1.9	1.5	0.6	2.2	0.0	0.9	1.8	0.6	3.7	1.0	0.7	0.9
1.90	3.06	6.1	3.0	0.4	2.2	0.4	0.0	1.4	0.1	0.5	1.4	1.8	1.5	0.4	1.9	0.0	0.7	1.3	0.5	2.4	0.9	0.6	0.9
2.00	2.90	5.1	2.9	0.4	1.8	0.4	0.0	1.4	0.2	0.4	1.3	1.6	1.5	0.2	1.7	0.0	0.5	0.9	0.4	1.5	0.7	0.5	1.0
2.10	2.61	4.6	2.9	0.5	1.4	0.4	0.0	1.4	0.3	0.3	1.2	1.6	1.5	0.1	1.5	0.0	0.4	0.6	0.3	1.0	0.6	0.4	1.1

Table A.6: Branching fraction $\mathcal{B}_{(s+d)\gamma}$ for different thresholds in units of 10^{-4}. Uncertainties in percentage.

A.3 Correlations matrices of the photon energy spectrum and the unfolded photon energy spectrum

The tables here present the correlation between systematic uncertainties for the background-subtracted E_{γ}^{*}spectrum and for the total uncertainties of the unfolded E_{γ}^{*}spectrum.

	1.7	1.8	1.9	2.0	2.1	2.2	2.3	2.4	2.5	2.6	2.7
1.7	100	85	71	54	40	26	31	24	14	15	8
1.8		100	76	57	42	34	31	23	15	12	12
1.9			100	58	45	38	32	22	19	11	14
2.0				100	48	38	35	24	15	16	14
2.1					100	40	36	23	22	20	15
2.2						100	41	21	24	9	29
2.3							100	48	44	42	37
2.4								100	38	59	35
2.5									100	56	59
2.6										100	54

Table A.7: Correlation of systematic uncertainties of the background-subtracted E_{γ}^{*} spectrum. The correlations are given in percent.

	1.7	1.8	1.9	2.0	2.1	2.2	2.3	2.4	2.5	2.6	2.7
1.7	100	97	79	29	11	18	7	7	8	3	1
1.8		100	89	40	12	16	9	8	9	4	3
1.9			100	69	20	10	11	10	11	6	5
2.0				100	59	8	-1	12	11	5	3
2.1				100	51	-17	1	12	5	4	
2.2						100	24	-18	2	8	8
2.3							100	13	-21	-5	-1
2.4								100	17	-26	-31
2.5									100	58	42
2.6										100	98
2.7											100

Table A.8: Correlation of uncertainties of the unfolded E_{γ}^{*} spectrum. The correlations are given in percent.

A.4 Uncertainties in the measurement of the partial branching fractions and spectral moments

In the following tables the uncertainties in the measurements of the $B \to X_{s+d}\gamma$ partial branching fractions and spectral moments$\mathcal{B}_{(s+d)\gamma}$, $\mathcal{B}_{s\gamma}$ and $\mathcal{B}_{s\gamma}$ extrapolated to 1.6 GeV are shown. The uncertainties are presented in Tables A.9, A.11 and A.13, and as percentage in Tables A.10, A.12 and A.14.

E_γ^* bin (GeV)	$\Delta\mathcal{B}_{(s+d)\gamma}$	Total	Stat	Det	ECL	EffSel	FindEff	NBB	Unf	BDT	beam	Cont	Elect	Eta	Etavet	Hadron	LepTag	MCstat	Other	Pi	Pivet	Pythia
1.6 – 1.7	4.1	20.1	6.7	0.0	0.1	0.3	2.8	0.1	0.7	4.1	1.8	1.0	1.8	7.0	0.0	3.3	4.7	1.8	15.2	2.9	0.4	0.1
1.7 – 1.8	8.2	29.8	9.5	0.0	0.1	0.5	3.9	0.1	0.6	5.2	2.9	1.8	2.8	9.9	0.0	4.9	7.3	2.5	23.0	4.3	0.4	0.1
1.8 – 1.9	11.6	22.2	7.0	0.0	0.2	0.4	2.8	0.2	0.2	2.6	2.5	1.5	2.2	6.6	0.0	3.6	6.0	1.8	17.4	3.3	0.2	0.1
1.9 – 2.0	16.6	12.5	5.2	0.0	0.3	0.4	1.8	0.2	0.2	2.3	1.8	1.1	1.3	3.6	0.0	1.8	3.7	1.3	8.7	2.1	0.8	0.1
2.0 – 2.1	24.1	7.4	4.5	0.0	0.4	0.5	1.4	0.3	0.3	3.0	1.2	0.8	0.5	2.7	0.0	0.6	1.7	1.1	2.2	1.5	1.0	0.3
2.1 – 2.2	34.5	6.2	4.2	0.0	0.5	0.8	0.8	0.5	0.3	2.6	0.9	1.1	0.1	2.4	0.0	0.2	0.7	0.8	0.4	1.0	0.7	0.3
2.2 – 2.3	47.4	4.9	3.5	0.0	0.4	0.9	0.4	0.6	0.2	1.7	0.7	0.9	0.1	1.8	0.0	0.3	0.3	0.5	0.6	0.6	0.3	0.4
2.3 – 2.4	60.8	3.8	3.1	0.0	0.2	0.9	0.2	0.8	0.3	1.0	0.6	0.6	0.0	0.6	0.0	0.2	0.2	0.3	0.4	0.4	0.1	0.5
2.4 – 2.5	62.8	3.5	2.7	0.0	0.2	1.1	0.1	0.9	0.2	0.6	0.4	0.8	0.0	0.4	0.0	0.1	0.1	0.2	0.2	0.3	0.0	0.5
2.5 – 2.6	43.5	2.8	2.3	0.0	0.5	0.5	0.0	0.6	0.6	0.3	0.4	0.7	0.0	0.1	0.0	0.0	0.0	0.1	0.1	0.1	0.0	0.4
2.6 – 2.7	20.0	2.3	2.0	0.0	0.5	0.3	0.0	0.3	0.4	0.2	0.4	0.7	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.2
2.7 – 2.8	1.9	0.3	0.3	0.0	0.1	0.0	0.0	0.0	0.1	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Table A.9: Partial branching fractions of the $B \to X_{s+d}\gamma$ spectrum in units of 10^{-6} and uncertainties.

E_γ^* bin (GeV)	$\Delta\mathcal{B}_{(s+d)\gamma}$	Total	Stat	Det	ECL	EffSel	FindEff	NBB	Unf	BDT	beam	Cont	Elect	Eta	Etavet	Hadron	LepTag	MCstat	Other	Pi	Pivet	Pythia
1.6 – −1.7	4.1	495.4	165.0	0.1	1.9	7.2	68.9	1.4	17.2	101.9	44.3	25.9	44.3	173.0	0.8	81.0	115.9	43.9	374.0	72.2	9.6	1.9
1.7 – −1.8	8.2	363.3	116.4	0.1	1.8	5.6	48.1	1.4	7.3	63.0	35.1	21.4	34.0	120.1	0.6	59.8	89.4	30.8	280.0	53.0	4.4	1.2
1.8 – −1.9	11.6	192.0	60.5	0.1	1.8	3.5	23.9	1.4	1.9	22.8	21.6	13.3	19.4	56.9	0.3	31.3	52.0	15.8	149.8	28.4	1.8	0.8
1.9 – −2.0	16.6	75.3	31.3	0.1	1.7	2.3	10.9	1.4	1.3	13.7	10.7	6.3	7.7	21.4	0.1	10.9	22.2	7.8	52.0	12.6	4.8	0.9
2.0 – −2.1	24.1	30.8	18.8	0.1	1.8	2.2	5.6	1.4	1.2	12.3	4.9	3.1	2.0	11.0	0.0	2.3	6.9	4.4	9.3	6.1	4.1	1.1
2.1 – −2.2	34.5	18.0	12.1	0.1	1.6	2.4	2.4	1.4	1.0	7.6	2.6	3.1	0.4	7.0	0.0	0.6	2.0	2.3	1.1	2.8	1.9	0.9
2.2 – −2.3	47.4	10.3	7.4	0.1	0.9	2.0	0.9	1.4	0.4	3.6	1.5	1.9	0.1	3.7	0.0	0.5	0.7	1.1	1.2	1.3	0.5	0.8
2.3 – −2.4	60.8	6.2	5.1	0.1	0.3	1.4	0.3	1.4	0.5	1.6	0.9	1.0	0.1	0.9	0.0	0.3	0.3	0.5	0.7	0.6	0.1	0.9
2.4 – −2.5	62.8	5.6	4.3	0.1	0.3	1.8	0.1	1.4	0.3	0.9	0.7	1.2	0.0	0.7	0.0	0.1	0.1	0.3	0.3	0.4	0.1	0.9
2.5 – −2.6	43.5	6.5	5.4	0.1	1.2	1.2	0.1	1.4	1.4	0.6	1.0	1.5	0.0	0.3	0.0	0.1	0.1	0.2	0.2	0.2	0.1	0.9
2.6 – −2.7	20.0	11.6	10.2	0.1	2.4	1.3	0.1	1.4	2.1	0.8	1.8	3.3	0.0	0.1	0.0	0.1	0.1	0.4	0.2	0.0	0.0	1.1

Table A.10: Partial branching fractions of the $B \to X_{s+d}\gamma$ spectrum in units of 10^{-6} and uncertainties in percent.

E_γ^* thr.	$\langle E_\gamma^* \rangle$	Total	Stat	Det	ECL	EffSel	FindEff	NBB	Unf	BDT	beam	Cont	Elect	Eta	Etavet	Hadron	LepTag	MCstat	Other	Pi	Pivet	Pythia
1.6	2.298	0.129	0.038	0.000	0.003	0.002	0.016	0.000	0.002	0.020	0.014	0.007	0.013	0.041	0.000	0.022	0.034	0.010	0.101	0.019	0.001	0.000
1.7	2.306	0.094	0.028	0.000	0.003	0.002	0.011	0.000	0.001	0.013	0.011	0.006	0.010	0.029	0.000	0.016	0.026	0.007	0.074	0.014	0.002	0.000
1.8	2.320	0.050	0.016	0.000	0.002	0.001	0.006	0.000	0.000	0.007	0.006	0.003	0.005	0.014	0.000	0.008	0.015	0.004	0.037	0.008	0.002	0.000
1.9	2.338	0.023	0.011	0.000	0.002	0.001	0.003	0.000	0.000	0.006	0.003	0.001	0.002	0.008	0.000	0.003	0.007	0.002	0.013	0.004	0.002	0.000
2.0	2.360	0.012	0.008	0.000	0.002	0.001	0.002	0.000	0.000	0.005	0.001	0.001	0.001	0.005	0.000	0.001	0.002	0.001	0.003	0.002	0.002	0.000
2.1	2.387	0.008	0.006	0.000	0.002	0.001	0.001	0.000	0.001	0.003	0.000	0.001	0.000	0.003	0.000	0.000	0.001	0.001	0.001	0.001	0.001	0.000

Table A.11: Mean energy of the $B \to X_{s+d}\gamma$ spectrum and uncertainties for different thresholds, values in GeV.

E_γ^* thr.	$\langle E_\gamma^* \rangle$	Total	Stat	Det	ECL	EffSel	FindEff	NBB	Unf	BDT	Beam	Cont	Elect	Eta	Etavet	Hadron	LepTag	MCstat	Other	Pi	Pivet	Pythia
1.6	2.298	5.6	1.7	0.0	0.1	0.1	0.7	0.0	0.1	0.9	0.6	0.3	0.6	1.8	0.0	0.9	1.5	0.4	4.4	0.8	0.0	0.0
1.7	2.306	4.1	1.2	0.0	0.1	0.1	0.5	0.0	0.0	0.6	0.5	0.2	0.4	1.2	0.0	0.7	1.1	0.3	3.2	0.6	0.1	0.0
1.8	2.320	2.1	0.7	0.0	0.1	0.1	0.2	0.0	0.0	0.3	0.3	0.1	0.2	0.6	0.0	0.3	0.6	0.2	1.6	0.3	0.1	0.0
1.9	2.338	1.0	0.5	0.0	0.1	0.1	0.1	0.0	0.0	0.3	0.1	0.1	0.1	0.3	0.0	0.1	0.3	0.1	0.6	0.2	0.1	0.0
2.0	2.360	0.5	0.3	0.0	0.1	0.1	0.1	0.0	0.0	0.2	0.1	0.0	0.0	0.2	0.0	0.0	0.1	0.1	0.1	0.1	0.1	0.0
2.1	2.387	0.3	0.2	0.0	0.1	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Table A.12: Mean energy in GeV of the $B \to X_{s+d}\gamma$ spectrum and uncertainties for different thresholds. Uncertaintites in percent.

E_γ^* thr.	ΔE_γ^*	Total	Stat	Det	ECL	EffSel	FindEff	NBB	Unf	BDT	Beam	Cont	Elect	Eta	Etavet	Hadron	LepTag	MCstat	Other	Pi	Pivet	Pythia
1.6	5.392	5.506	1.728	0.002	0.008	0.089	0.709	0.000	0.140	0.937	0.535	0.323	0.524	1.793	0.009	0.907	1.380	0.456	4.265	0.800	0.053	0.007
1.7	4.938	3.625	1.121	0.002	0.005	0.065	0.451	0.000	0.070	0.532	0.376	0.236	0.357	1.124	0.006	0.596	0.946	0.293	2.829	0.529	0.005	0.003
1.8	4.258	1.589	0.525	0.001	0.002	0.041	0.195	0.000	0.029	0.192	0.190	0.127	0.165	0.448	0.002	0.253	0.447	0.132	1.220	0.240	0.037	0.001
1.9	3.563	0.548	0.260	0.001	0.002	0.028	0.082	0.000	0.023	0.125	0.078	0.060	0.054	0.154	0.001	0.071	0.156	0.060	0.342	0.094	0.041	0.002
2.0	2.869	0.207	0.141	0.001	0.004	0.023	0.034	0.000	0.019	0.081	0.029	0.035	0.011	0.065	0.000	0.011	0.038	0.028	0.045	0.036	0.024	0.002
2.1	2.195	0.111	0.087	0.001	0.006	0.019	0.011	0.000	0.015	0.038	0.012	0.026	0.001	0.029	0.000	0.001	0.007	0.011	0.001	0.012	0.007	0.001

Table A.13: Variance of the $B \to X_{s+d}\gamma$ spectrum in units of $10^{-2}\mathrm{GeV}^2$ and uncertainties for different thresholds.

E_γ^* thr.	ΔE_γ^*	Total	Stat	Det	ECL	EffSel	FindEff	NBB	Unf	BDT	Beam	Cont	Elect	Eta	Etavet	Hadron	LepTag	MCstat	Other	Pi	Pivet	Pythia
1.6	5.392	102.1	32.0	0.0	0.1	1.7	13.2	0.0	2.6	17.4	9.9	6.0	9.7	33.2	0.2	16.8	25.6	8.4	79.1	14.8	1.0	0.1
1.7	4.938	73.4	22.7	0.0	0.1	1.3	9.1	0.0	1.4	10.8	7.6	4.8	7.2	22.8	0.1	12.1	19.2	5.9	57.3	10.7	0.1	0.1
1.8	4.258	37.3	12.3	0.0	0.0	1.0	4.6	0.0	0.7	4.5	4.5	3.0	3.9	10.5	0.1	5.9	10.5	3.1	28.6	5.6	0.9	0.0
1.9	3.563	15.4	7.3	0.0	0.1	0.8	2.3	0.0	0.6	3.5	2.2	1.7	1.5	4.3	0.0	2.0	4.4	1.7	9.6	2.6	1.2	0.1
2.0	2.869	7.2	4.9	0.0	0.1	0.8	1.2	0.0	0.6	2.8	1.0	1.2	0.4	2.3	0.0	0.4	1.3	1.0	1.6	1.3	0.8	0.1
2.1	2.195	5.1	4.0	0.0	0.3	0.8	0.5	0.0	0.7	1.7	0.5	1.2	0.1	1.3	0.0	0.1	0.3	0.5	0.0	0.6	0.3	0.0

Table A.14: Variance of the $B \to X_{s+d}\gamma$ spectrum and uncertainties in units of $10^{-2}\mathrm{GeV}^2$ for different thresholds. Uncertaintites in percent.

APPENDIX B

Selection efficiencies and correction factors

The table below shows the selection efficiency, conversion factor from the CM frame to the B frame and the extrapolation factor. The uncertainties are related to MC statistics and due model dependence.

E_γ	$\varepsilon_{\mathrm{sel},\,i}$	$\varepsilon_{\mathrm{sel}}^{ethr}$	$\alpha^{E_\gamma \geq E_0}$	Extrapolation
1.6	0.539 ± 0.054 ± 0.018	2.370 ± 0.081 ± 0.005	1.047 ± 0.294	0.000 ± 0.000
1.7	0.718 ± 0.052 ± 0.016	2.403 ± 0.070 ± 0.005	1.347 ± 0.235	1.515 ± 0.624
1.8	0.957 ± 0.042 ± 0.016	2.454 ± 0.059 ± 0.005	2.161 ± 0.310	3.686 ± 1.390
1.9	1.256 ± 0.032 ± 0.015	2.521 ± 0.055 ± 0.005	3.336 ± 0.388	6.950 ± 2.394
2.0	1.605 ± 0.039 ± 0.013	2.608 ± 0.045 ± 0.005	5.255 ± 0.460	12.121 ± 3.804
2.1	1.987 ± 0.047 ± 0.012	2.718 ± 0.038 ± 0.006	8.107 ± 0.579	20.776 ± 5.932
2.2	2.419 ± 0.041 ± 0.011	2.850 ± 0.039 ± 0.007	12.288 ± 1.077	36.240 ± 9.326
2.3	2.757 ± 0.034 ± 0.011	—	—	—
2.4	2.993 ± 0.052 ± 0.014	—	—	—
2.5	3.184 ± 0.041 ± 0.021	—	—	—
2.6	3.380 ± 0.034 ± 0.044	—	—	—
2.7	3.217 ± 0.049 ± 0.144	—	—	—

Table B.1: Selection efficiency and correction factors for the calculation of the $B \to X_s\gamma$ branching fraction. The first column shows the corresponding bin (or threshold). The second column shows the selection efficiency $\varepsilon_{\mathrm{sel},\,i}$, in percent, in a bin-by-bin basis, whereas the third one shows it for the integrated spectrum from the given threshold. The fourth column shows the factor to convert the CM measurement in to the B frame for the given threshold, the difference from unit is presented as percent. The last column showns the extrapolation factor from the given threshold to 1.6 GeV, the difference from unit is presented as percent. The central values are accompanied by the model dependence uncertainty, for the selection efficiencies the second uncertainty comes from MC statistics.

APPENDIX C

Response matrix

The response matrix used in the unfolding procedure is presented below.

	1.6	1.7	1.8	1.9	2.0	2.1	2.2	2.3	2.4	2.5	2.6	2.7
1.6	48.76	24.77	6.20	3.37	2.93	3.16	3.29	2.66	1.54	0.62	0.12	0.00
1.7	4.40	49.74	22.35	6.63	4.20	3.87	3.67	2.86	1.55	0.56	0.12	0.00
1.8	0.11	4.43	46.41	23.94	8.54	5.78	4.80	3.45	1.79	0.64	0.10	0.00
1.9	0.03	0.10	4.12	44.66	26.82	10.80	6.56	4.16	1.98	0.65	0.11	0.00
2.0	0.01	0.02	0.08	4.04	45.81	30.09	11.49	5.40	2.25	0.70	0.11	0.00
2.1	0.00	0.01	0.02	0.07	4.61	50.50	31.33	9.54	2.99	0.80	0.12	0.00
2.2	0.00	0.00	0.01	0.02	0.09	6.10	57.21	29.04	6.11	1.26	0.16	0.00
2.3	0.00	0.00	0.00	0.01	0.04	0.15	8.92	63.63	23.57	3.35	0.33	0.01
2.4	0.00	0.00	0.00	0.01	0.02	0.08	0.31	13.36	67.14	17.77	1.27	0.02
2.5	0.00	0.00	0.00	0.01	0.03	0.08	0.27	0.75	19.51	69.26	9.97	0.12
2.6	0.00	0.00	0.01	0.03	0.07	0.20	0.52	1.12	2.07	32.59	61.69	1.69
2.7	0.01	0.03	0.06	0.23	0.47	1.15	3.00	5.95	8.58	9.82	50.66	20.01

Table C.1: Detector response matrix. It gives the probability of an event with energy X (given in columns), to be reconstructed with energy Y (given in rows). All energies given in GeV.

APPENDIX D

BDT in the signal region

The BDT distribution in the signal region is shown below for different E^*_γbins.

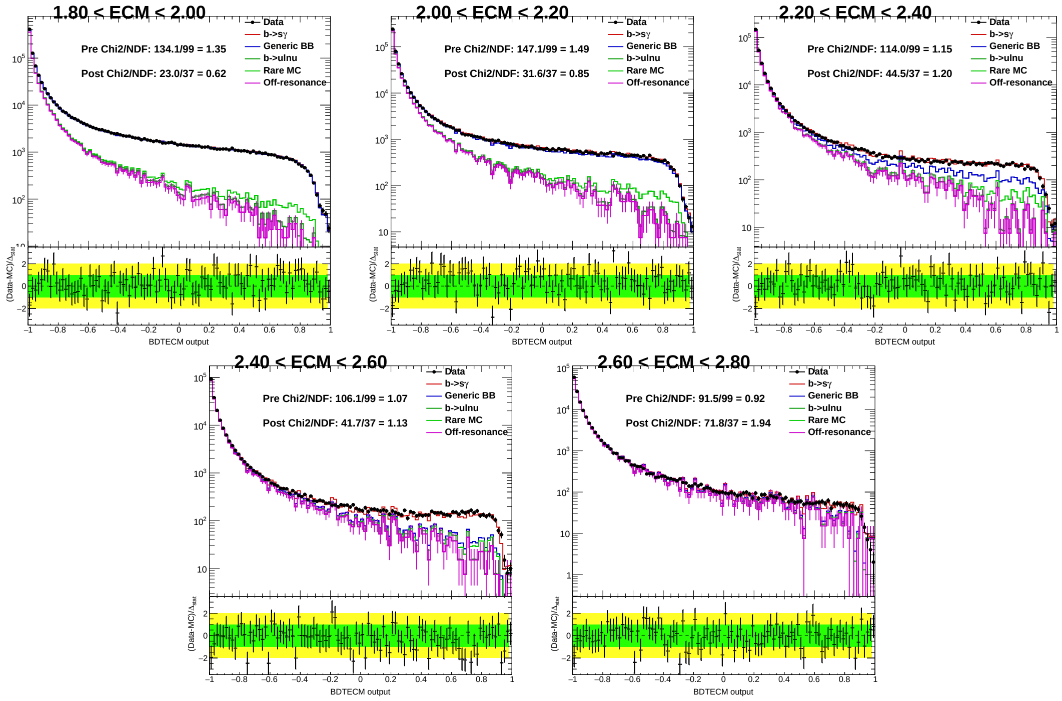


Figure D.1: BDT distribution in the signal region, in bins of E^*_γ. The ratio plot below takes into account the statistical uncertainties of on and off-resonance data.

APPENDIX E

Systematic uncertainty in the unfolding

Comparison between the integral and spectral moments of the unfolded toys and the BLNP0 sample where they are derived from. These plots are used to determine the systematic uncertainty related to the unfolding procedure.

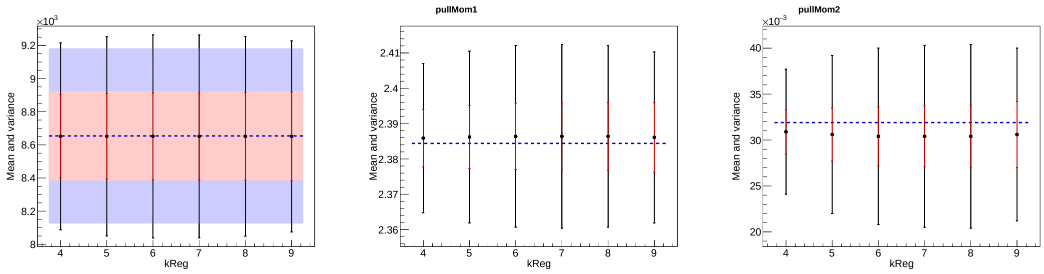


Figure E.1: Toy mean and variance for the spectrum's integral, mean energy and energy variance, for different values of the regularization parameter kReg using the BLNP0 MC sample. The colors are explained in Fig. 7.20.

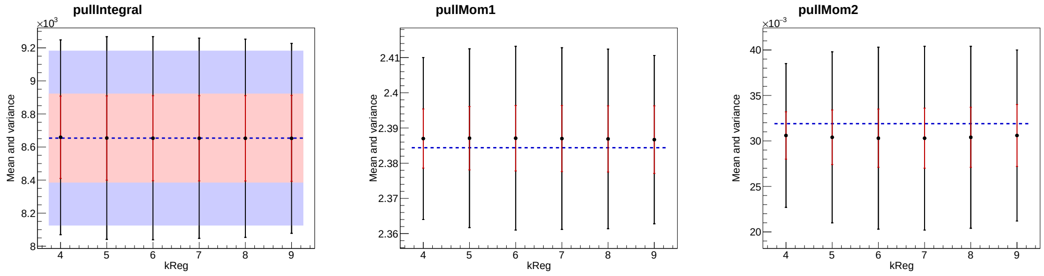


Figure E.2: Toy mean and variance for the spectrum's integral, mean energy and energy variance, for different values of the regularization parameter kReg using the BLNP1 MC sample. The colors are explained in Fig. 7.20.

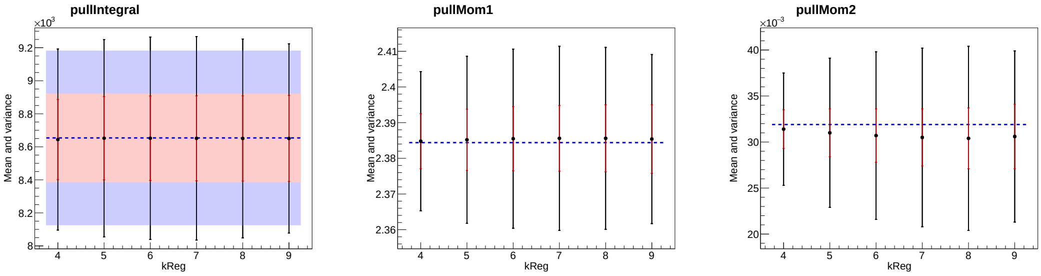


Figure E.3: Toy mean and variance for the spectrum's integral, mean energy and energy variance, for different values of the regularization parameter kReg using the BLNP2 MC sample. The colors are explained in Fig. 7.20.

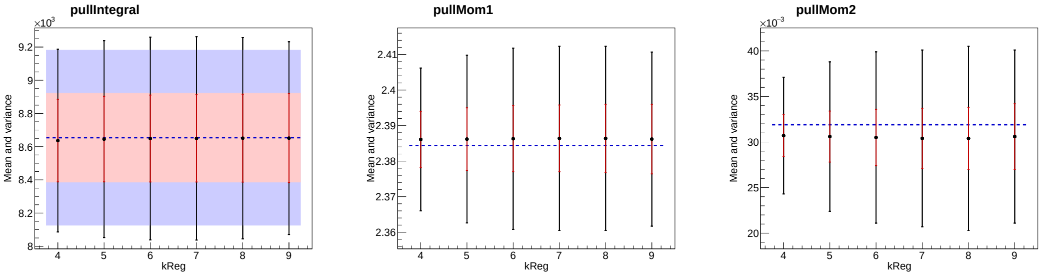


Figure E.4: Toy mean and variance for the spectrum's integral, mean energy and energy variance, for different values of the regularization parameter kReg using the BBU0 MC sample. The colors are explained in Fig. 7.20.

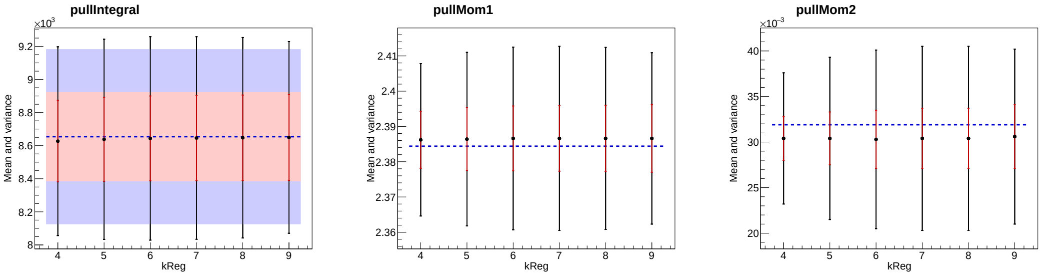


Figure E.5: Toy mean and variance for the spectrum's integral, mean energy and energy variance, for different values of the regularization parameter kReg using the BBU1 MC sample. The colors are explained in Fig. 7.20.

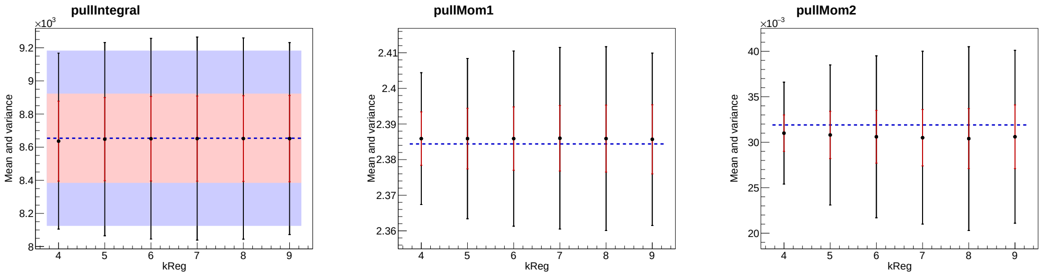


Figure E.6: Toy mean and variance for the spectrum's integral, mean energy and energy variance, for different values of the regularization parameter kReg using the BBU2 MC sample. The colors are explained in Fig. 7.20.

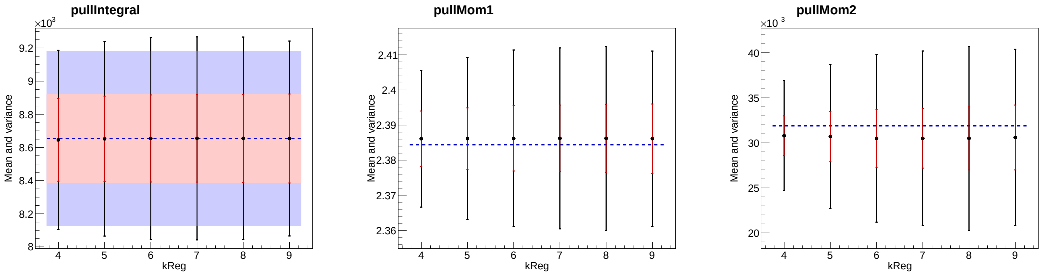


Figure E.7: Toy mean and variance for the spectrum's integral, mean energy and energy variance, for different values of the regularization parameter kReg using the KN0 MC sample. The colors are explained in Fig. 7.20.

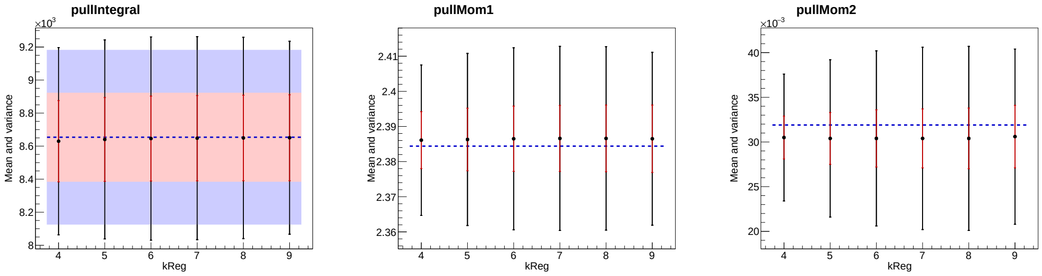


Figure E.8: Toy mean and variance for the spectrum's integral, mean energy and energy variance, for different values of the regularization parameter kReg using the KN1 MC sample. The colors are explained in Fig. 7.20.

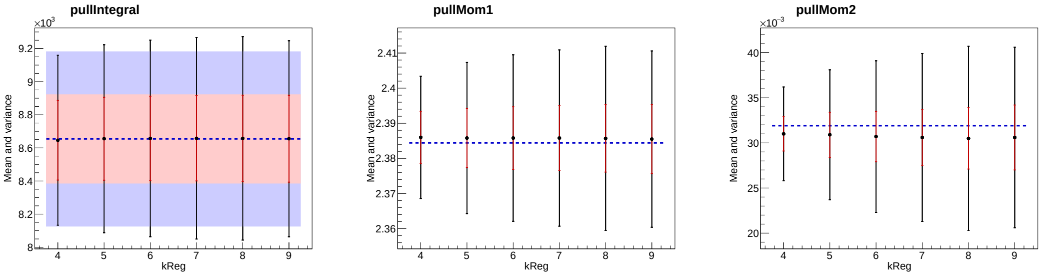


Figure E.9: Toy mean and variance for the spectrum's integral, mean energy and energy variance, for different values of the regularization parameter kReg using the KN2 MC sample. The colors are explained in Fig. 7.20.

Bibliography

[1] A. I. Vainshtein, V. I. Zakharov, and M. A. Shifman, *A Possible mechanism for the Delta T = 1/2 rule in nonleptonic decays of strange particles*, JETP Lett. **22** (1975) p. 55, [Pisma Zh. Eksp. Teor. Fiz.22,123(1975)] (cit. on p. 1).

[2] J. R. Ellis et al., *The Phenomenology of the Next Left-Handed Quarks*, Nucl. Phys. **B131** (1977) p. 285, [Erratum: Nucl. Phys.B132,541(1978)] (cit. on p. 1).

[3] R. Ammar et al., *Evidence for penguins: First observation of* $B \to K^*(892)\gamma$, Phys. Rev. Lett. **71** (1993) p. 674 (cit. on p. 1).

[4] E. Noether, *Invariant Variation Problems*, Gott. Nachr. **1918** (1918) p. 235, [Transp. Theory Statist. Phys.1,186(1971)], arXiv: `physics/0503066 [physics]` (cit. on p. 3).

[5] T. D. Lee and C. N. Yang, *Question of Parity Conservation in Weak Interactions*, Phys. Rev. **104** (1 1956) p. 254, URL: `http://link.aps.org/doi/10.1103/PhysRev.104.254` (cit. on p. 4).

[6] C. S. Wu et al., *Experimental Test of Parity Conservation in Beta Decay*, Phys. Rev. **105** (4 1957) p. 1413, URL: `http://link.aps.org/doi/10.1103/PhysRev.105.1413` (cit. on p. 4).

[7] J. Christenson et al., *Evidence for the* 2π *Decay of the* K_2^0 *Meson*, Phys.Rev.Lett. **13** (1964) p. 138 (cit. on p. 4).

[8] D. Griffiths, *Introduction to elementary particles*, 2008 (cit. on p. 4).

[9] 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) p. 1, arXiv: `1207.7214 [hep-ex]` (cit. on p. 4).

[10] 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) p. 30, arXiv: `1207.7235 [hep-ex]` (cit. on p. 4).

[11] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*, 1995, ISBN: 9780201503975, 0201503972, URL: `http://www.slac.stanford.edu/spires/find/books/www?cl=QC174.45%3AP4` (cit. on p. 5).

[12] A. J. Bevan et al., *The Physics of the B Factories*, Eur. Phys. J. **C74** (2014) p. 3026, arXiv: `1406.6311 [hep-ex]` (cit. on pp. 6, 9–12, 24, 27).

[13] J. Charles et al.,
Current status of the Standard Model CKM fit and constraints on $\Delta F = 2$ New Physics,
Phys. Rev. **D91**.7 (2015) p. 073007, arXiv: `1501.05013 [hep-ph]`
(cit. on pp. 6, 7, 11, 16).

[14] P. Kooijman and N. Tuning, *Lectures Notes on CP Violation*, 2012,
URL: `http://www.nikhef.nl/~h71/Lectures/2012/cp-080212.pdf`
(cit. on p. 6).

[15] K. G. Wilson and W. Zimmermann, *Operator product expansions and composite field operators in the general framework of quantum field theory*,
Commun. Math. Phys. **24** (1972) p. 87 (cit. on p. 7).

[16] M. Misiak et al.,
Updated NNLO QCD predictions for the weak radiative B-meson decays,
Phys. Rev. Lett. **114**.22 (2015) p. 221801, arXiv: `1503.01789 [hep-ph]`
(cit. on pp. 7, 122).

[17] M. A. Shifman, “Foreword to ITEP lectures in particle physics”,
ITEP Lectures in Particle Physics and Field Theory. Edited by M. Shifman. Singapore, World Scientific, 1999. Vol. 1, pp. v-xi, 1995, arXiv: `hep-ph/9510397 [hep-ph]`
(cit. on p. 8).

[18] G. Buchalla, A. J. Buras, and M. E. Lautenbacher,
Weak decays beyond leading logarithms, Rev. Mod. Phys. **68** (1996) p. 1125,
arXiv: `hep-ph/9512380 [hep-ph]` (cit. on p. 8).

[19] A. J. Buras, *Operator product expansion, renormalization group and weak decays*,
Lect. Notes Phys. **558** (2000) p. 65, [,65(1999)], arXiv: `hep-ph/9901409 [hep-ph]`
(cit. on p. 8).

[20] M. Czakon et al., *The $(Q_7, Q_{1,2})$ contribution to $\overline{B} \to X_s\gamma$ at $\mathcal{O}\left(\alpha_s^2\right)$*,
JHEP **04** (2015) p. 168, arXiv: `1503.01791 [hep-ph]` (cit. on pp. 8, 9, 11, 18, 125).

[21] M. Misiak, *Weak radiative B decays*, J. Phys. **G27** (2001) p. 1051 (cit. on p. 9).

[22] M. Misiak et al., *Estimate of $\mathcal{B}(\bar{B} \to X_s\gamma)$ at $O(\alpha_s^2)$*,
Phys. Rev. Lett. **98** (2007) p. 022002, arXiv: `hep-ph/0609232 [hep-ph]` (cit. on p. 9).

[23] M. Benzke et al.,
Factorization at Subleading Power and Irreducible Uncertainties in $\bar{B} \to X_s\gamma$ Decay,
JHEP **08** (2010) p. 099, arXiv: `1003.5012 [hep-ph]` (cit. on p. 9).

[24] M. Czakon, U. Haisch, and M. Misiak,
Four-Loop Anomalous Dimensions for Radiative Flavour-Changing Decays,
JHEP **03** (2007) p. 008, arXiv: `hep-ph/0612329 [hep-ph]` (cit. on p. 9).

[25] P. Gambino and M. Misiak, *Quark mass effects in $B \to X_s\gamma$*,
Nucl. Phys. **B611** (2001) p. 338, arXiv: `hep-ph/0104034 [hep-ph]`
(cit. on pp. 10, 11).

[26] A. J. Buras et al.,
Theoretical uncertainties and phenomenological aspects of B —> X(s) gamma decay,
Nucl. Phys. **B424** (1994) p. 374, arXiv: `hep-ph/9311345 [hep-ph]` (cit. on p. 10).

[27] Y. Amhis et al.,
Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2014, (2014),
arXiv: `1412.7515 [hep-ex]` (cit. on pp. 11, 16, 18, 27, 44, 86, 116).

[28] P. del Amo Sanchez et al., *Study of $B \to X\gamma$ Decays and Determination of $|V_{td}/V_{ts}|$*,
Phys. Rev. **D82** (2010) p. 051101, arXiv: `1005.4087 [hep-ex]` (cit. on pp. 12, 26, 89).

[29] M. Neubert, *Analysis of the photon spectrum in inclusive $B \to X_s\gamma$ decays*,
Phys. Rev. **D49** (1994) p. 4623, arXiv: `hep-ph/9312311 [hep-ph]` (cit. on p. 12).

[30] A. Ali and C. Greub, *Inclusive photon energy spectrum in rare B decays*,
Z. Phys. **C49** (1991) p. 431 (cit. on p. 12).

[31] A. L. Kagan and M. Neubert, *QCD anatomy of $B \to X_s\gamma$ decays*,
Eur.Phys.J. **C7** (1999) p. 5, arXiv: `hep-ph/9805303 [hep-ph]` (cit. on pp. 12, 13, 44).

[32] M. Neubert, *Advanced predictions for moments of the $B \to X_s\gamma$ photon spectrum*,
Phys. Rev. **D72** (2005) p. 074025, arXiv: `hep-ph/0506245 [hep-ph]`
(cit. on pp. 12, 13).

[33] D. Benson, I. I. Bigi, and N. Uraltsev,
On the photon energy moments and their 'bias' corrections in $B \to X_s\gamma$,
Nucl. Phys. **B710** (2005) p. 371, arXiv: `hep-ph/0410080 [hep-ph]`
(cit. on pp. 12, 14).

[34] Z. Ligeti, I. W. Stewart, and F. J. Tackmann,
Treating the b quark distribution function with reliable uncertainties,
Phys. Rev. **D78** (2008) p. 114014, arXiv: `0807.1926 [hep-ph]` (cit. on p. 13).

[35] B. O. Lange, M. Neubert, and G. Paz,
Theory of charmless inclusive B decays and the extraction of V(ub),
Phys. Rev. **D72** (2005) p. 073006, arXiv: `hep-ph/0504071 [hep-ph]`
(cit. on pp. 13, 14).

[36] C. W. Bauer, M. E. Luke, and T. Mannel,
Light cone distribution functions for B decays at subleading order in 1 / m(b),
Phys. Rev. **D68** (2003) p. 094001, arXiv: `hep-ph/0102089 [hep-ph]` (cit. on p. 14).

[37] I. Bigi and N. Uraltsev,
On the expected photon spectrum in B –> X + gamma and its uses,
Int. J. Mod. Phys. **A17** (2002) p. 4709, arXiv: `hep-ph/0202175 [hep-ph]`
(cit. on pp. 14, 15).

[38] S. Chen et al., *Branching fraction and photon energy spectrum for $b \to s\gamma$*,
Phys. Rev. Lett. **87** (2001) p. 251807, arXiv: `hep-ex/0108032 [hep-ex]`
(cit. on pp. 14, 15, 23).

[39] A. L. Kagan and M. Neubert,
Direct CP violation in $B \to X_s\gamma$ decays as a signature of new physics,
Phys.Rev. **D58** (1998) p. 094012, arXiv: `hep-ph/9803368 [hep-ph]`
(cit. on pp. 15, 16, 110).

[40] M. Benzke et al., *Long-Distance Dominance of the CP Asymmetry in $B \to X_s\gamma$ Decays*,
Phys.Rev.Lett. **106** (2011) p. 141801, arXiv: `1012.3167 [hep-ph]`
(cit. on pp. 15, 19, 125).

[41] T. Hurth and T. Mannel,
CP asymmetries in $b \to (s/d)$ transitions as a test of CKM CP violation,
Phys. Lett. **B511** (2001) p. 196, arXiv: `hep-ph/0103331 [hep-ph]` (cit. on p. 16).

[42] G. C. Branco et al., *Theory and phenomenology of two-Higgs-doublet models*,
Phys. Rept. **516** (2012) p. 1, arXiv: `1106.0034 [hep-ph]` (cit. on p. 16).

[43] T. Hermann, M. Misiak, and M. Steinhauser,
$\bar{B} \to X_s\gamma$ in the Two Higgs Doublet Model up to Next-to-Next-to-Leading Order in QCD,
JHEP **11** (2012) p. 036, arXiv: `1208.2788 [hep-ph]` (cit. on p. 17).

[44] O. Deschamps et al., *The Two Higgs Doublet of Type II facing flavour physics data*,
Phys. Rev. **D82** (2010) p. 073012, arXiv: `0907.5135 [hep-ph]` (cit. on pp. 17, 126).

[45] T. Hurth, E. Lunghi, and W. Porod,
Untagged $B \to X_(s+d)\gamma$ CP asymmetry as a probe for new physics,
Nucl.Phys. **B704** (2005) p. 56, arXiv: `hep-ph/0312260 [hep-ph]`
(cit. on pp. 18, 20, 131).

[46] R. Malm, M. Neubert, and C. Schmell,
Impact of Warped Extra Dimensions on the Dipole Coefficients in $\boldsymbol{b \to s\gamma}$ Transitions,
(2015), arXiv: `1509.02539 [hep-ph]` (cit. on p. 19).

[47] P. Moch and J. Rohrwild, *$\bar{B} \to X_s\gamma$ with a warped bulk Higgs*, (2015),
arXiv: `1509.04643 [hep-ph]` (cit. on p. 19).

[48] H. F. A. G. (HFAG),
Electroweak radiative B decays, preliminary results from ICHEP 2016, 2016, URL: `http://www.slac.stanford.edu/xorg/hfag/rare/ICHEP2016/radll/index.html`
(cit. on p. 22).

[49] T. Coan et al., *Bounds on the CP asymmetry in $b \to s\gamma$ decays*,
Phys.Rev.Lett. **86** (2001) p. 5661, arXiv: `hep-ex/0010075 [hep-ex]` (cit. on p. 23).

[50] A. Limosani et al., *Measurement of Inclusive Radiative B-meson Decays with a Photon Energy Threshold of 1.7-GeV*, Phys.Rev.Lett. **103** (2009) p. 241801,
arXiv: `0907.1384 [hep-ex]` (cit. on pp. 23–25, 123).

[51] *Measurement of BR($B \to X_s\gamma$), the $B \to X_s\gamma$ photon energy spectrum, and the direct CP asymmetry $\mathcal{A}^{\mathrm{CP}}_{(s+d)\gamma}$ decays*, Phys. Rev. **D86** (2012) p. 112008,
arXiv: `1207.5772 [hep-ex]` (cit. on pp. 23, 24, 123).

[52] A. Limosani,
Measurement of inclusive radiative B-meson decays with and without a lepton tag,
Belle Note 1035 (2008) (cit. on pp. 25, 42, 43, 123).

[53] T. Saito et al.,
Measurement of the $\bar{B} \rightarrow X_s\gamma$ Branching Fraction with a Sum of Exclusive Decays,
Phys. Rev. **D91**.5 (2015) p. 052004, arXiv: `1411.7198 [hep-ex]`
(cit. on pp. 26, 83, 84).

[54] J. P. Lees et al.,
Exclusive Measurements of $b \rightarrow s\gamma$ Transition Rate and Photon Energy Spectrum,
Phys. Rev. **D86** (2012) p. 052012, arXiv: `1207.2520 [hep-ex]` (cit. on p. 26).

[55] J. P. Lees et al., *Measurements of direct CP asymmetries in $B \rightarrow X_s\gamma$ decays using sum of exclusive decays*, Phys. Rev. **D90**.9 (2014) p. 092001, arXiv: `1406.0534 [hep-ex]` (cit. on p. 26).

[56] N. Taniguchi et al., *Measurement of branching fractions, isospin and CP-violating asymmetries for exclusive $b \rightarrow d\gamma$ modes*,
Phys. Rev. Lett. **101** (2008) p. 111801, [Erratum: Phys. Rev. Lett.101,129904(2008)],
arXiv: `0804.4770 [hep-ex]` (cit. on pp. 27, 89).

[57] M. Feindt et al., *A Hierarchical NeuroBayes-based Algorithm for Full Reconstruction of B Mesons at B Factories*, Nucl. Instrum. Meth. **A654** (2011) p. 432,
arXiv: `1102.3876 [hep-ex]` (cit. on p. 27).

[58] B. Aubert et al., *Measurement of the $B \rightarrow X_s\gamma$ branching fraction and photon energy spectrum using the recoil method*, Phys. Rev. **D77** (2008) p. 051103,
arXiv: `0711.4889 [hep-ex]` (cit. on p. 27).

[59] Y. Watanabe, *B factory detectors*, Nucl.Instrum.Meth. **A453** (2000) p. 9 (cit. on p. 29).

[60] *Integrated luminosity of B-factories*,
URL: `http://belle.kek.jp/bdocs/lumi_belle.png` (cit. on p. 29).

[61] S. Kurokawa and E. Kikutani, *Overview of the KEKB accelerators*,
Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **499**.1 (2003) p. 1, ISSN: 0168-9002,
URL:
`http://www.sciencedirect.com/science/article/pii/S0168900202017710`
(cit. on p. 29).

[62] *KEKB/Belle integrated luminosity reached 1000*, 2009,
URL: `http://legacy.kek.jp/intra-e/press/2009/KEKBluminosity3.html`
(cit. on p. 29).

[63] *KEKB ring*, URL: `http://belle.kek.jp/belle/slides/KEKB/Bellering4.gif`
(cit. on p. 30).

[64] A. Abashian et al., *The Belle Detector*, Nucl.Instrum.Meth. **A479** (2002) p. 117
(cit. on pp. 30, 31, 33–39, 41).

[65] H. Aihara, *The Belle Silicon Vertex Detector*, Nucl.Instrum.Meth. **A466** (2001) p. 268, eprint: `arXiv:hep-ex/0011028` (cit. on p. 31).

[66] R. Stamen et al., *Status of the Belle Silicon Vertex Detector*, Nucl.Instrum.Meth. **A541** (2005) p. 61 (cit. on pp. 31, 32).

[67] K. Abe et al., *Letter of Intent for KEK Super B Factory. Part II: Detector*, 2004, URL: `http://superb.kek.jp/documents/loi/img/LoI_detector.pdf` (cit. on pp. 32, 33).

[68] Shoji and Uno, *The BELLE central drift chamber*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **379**.3 (1996) p. 421, Proceedings of the Sixth International Conference on Instrumentation for Experiments at e+ e- Colliders, ISSN: 0168-9002, URL: `http://www.sciencedirect.com/science/article/pii/0168900296005554` (cit. on p. 32).

[69] T. Sumiyoshi et al., *Silica aerogel Cherenkov counter for the KEK B-factory experiment*, Nucl. Instrum. Methods Phys. Res., A **433**.1-2 (1999) p. 385 (cit. on p. 35).

[70] Kenkichi and Miyabayashi, *Belle electromagnetic calorimeter*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **494**.1-3 (2002) p. 298, Proceedings of the 8th International Conference on Instrumentation for Colliding Beam Physics, ISSN: 0168-9002, URL: `http://www.sciencedirect.com/science/article/pii/S0168900202014833` (cit. on p. 36).

[71] B. Casey, *HadronB*, Belle Note 390 (2001) (cit. on pp. 39, 43, 45).

[72] D. Lange, *The EvtGen particle decay simulation package*, Nucl.Instrum.Meth. **A462** (2001) p. 152 (cit. on p. 41).

[73] V. Zhilich, *Integrated Luminosity for Hadron B(J) skim for reprocessed data (CaseB)*, 2010, URL: `http://belle.kek.jp/group/ecl/private/lum/lum6new.html` (cit. on p. 43).

[74] T. Sjostrand, *PYTHIA 5.7 and JETSET 7.4: Physics and manual*, (1995), arXiv: `hep-ph/9508391 [hep-ph]` (cit. on p. 44).

[75] Y. Amhis et al., *Averages of B-Hadron, C-Hadron, and tau-lepton properties as of early 2012*, (2012), arXiv: `1207.1158 [hep-ex]` (cit. on p. 44).

[76] K. Hanagaki et al., *Electron identification in Belle*, Nucl.Instrum.Meth. **A485** (2002) p. 490, arXiv: `hep-ex/0108044 [hep-ex]` (cit. on p. 47).

[77] A. Abashian et al., *Muon identification in the Belle experiment at KEKB*, Nucl.Instrum.Meth. **A491** (2002) p. 69 (cit. on p. 48).

[78] P. Koppenburg, *An improved π^0 and η veto*, Belle Note 665 (2003) (cit. on p. 49).

[79] I. Nakamura, *Implementation of tracking information of ECL clusters*, Belle Note 774 (2004) (cit. on p. 50).

[80] S. Villa, *Search for the decay $B^0 \to \gamma\gamma$*, Belle Note 831 (2006) (cit. on p. 50).

[81] *SVD1 data containing ECL timing*, 2012, URL: `http://belle.kek.jp/secured/wiki/doku.php?id=software:svd1%5C_data%5C_with%5C_ecl%5C_timing` (cit. on p. 50).

[82] G. C. Fox and S. Wolfram, *Observables for the Analysis of Event Shapes in e+ e- Annihilation and Other Processes*, Phys.Rev.Lett. **41** (1978) p. 1581 (cit. on p. 54).

[83] K. Abe et al., *A Measurement of the branching fraction for the inclusive $B \to X_s\gamma$ decays with BELLE*, Phys.Lett. **B511** (2001) p. 151, arXiv: `hep-ex/0103042 [hep-ex]` (cit. on p. 54).

[84] K. Abe et al., *Measurement of branching fractions for $B \to \pi\pi$, $K\pi$ and KK decays*, Phys.Rev.Lett. **87** (2001) p. 101801, arXiv: `hep-ex/0104030 [hep-ex]` (cit. on p. 54).

[85] K. Abe et al., *Evidence for $B^0 \to \pi^0\pi^0$*, Phys.Rev.Lett. **91** (2003) p. 261801, arXiv: `hep-ex/0308040 [hep-ex]` (cit. on p. 54).

[86] A. Hocker et al., *TMVA - Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) p. 040, arXiv: `physics/0703039 [PHYSICS]` (cit. on pp. 56, 57).

[87] P. Urquijo et al., *Moments of the electron energy spectrum and partial branching fraction of $B \to X_c e\nu$ decays at Belle*, Phys. Rev. **D75** (2007) p. 032001, arXiv: `hep-ex/0610012 [hep-ex]` (cit. on p. 61).

[88] J. Oswald Ch. Hasenbusch and U. P., *MC corrections for semileptonic decays*, Belle Note 1335 (2014) (cit. on p. 62).

[89] A. K. Leibovich et al., *Semileptonic B decays to excited charmed mesons*, Phys. Rev. **D57** (1998) p. 308, arXiv: `hep-ph/9705467 [hep-ph]` (cit. on p. 62).

[90] R. package, *Cruijff function*, 2006, URL: `https://github.com/ekfriis/FinalStateAnalysis/blob/master/StatTools/src/RooCruijff.cc` (cit. on p. 64).

[91] A. Hocker and V. Kartvelishvili, *SVD approach to data unfolding*, Nucl. Instrum. Meth. **A372** (1996) p. 469, arXiv: `hep-ph/9509307 [hep-ph]` (cit. on pp. 76, 96, 100, 102).

[92] G. Majumder, *Linearity test of ECL detector with Radiative Mupair Event*, Belle Note 718 (2005) (cit. on pp. 78, 79).

[93] T. Sjostrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **05** (2006) p. 026, arXiv: `hep-ph/0603175 [hep-ph]` (cit. on p. 83).

[94] P. Gambino and N. Uraltsev, *Moments of semileptonic B decay distributions in the 1/m(b) expansion*, Eur. Phys. J. **C34** (2004) p. 181, arXiv: `hep-ph/0401063 [hep-ph]` (cit. on p. 90).

[95] G. Cowan, *A survey of unfolding methods for particle physics*, Conf. Proc. **C0203181** (2002) p. 248, [,248(2002)] (cit. on p. 94).

[96] R. Brun, F. Rademakers, and S. Panacek, *ROOT, an object oriented data analysis framework*, Conf. Proc. **C000917** (2000) p. 11 (cit. on p. 95).

[97] K. Tackmann and A. Hoecker, "SVD-based unfolding: implementation and experience", *Proceedings, PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding, CERN,Geneva, Switzerland 17-20 January 2011*, CERN, Geneva: CERN, 2011 p. 276, arXiv: `1112.2226 [physics.data-an]`, URL: `http://inspirehep.net/record/1081446/files/arXiv:1112.2226.pdf` (cit. on p. 95).

[98] G. Cowan et al., *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. **C71** (2011) p. 1554, [Erratum: Eur. Phys. J.C73,2501(2013)], arXiv: `1007.1727 [physics.data-an]` (cit. on p. 99).

[99] K. Tackmann and A. Hoecker, private communication, 2016 (cit. on p. 102).

[100] L. Hinz, *Lepton ID efficiency correction and systematic error*, Belle Note 954 (2006) (cit. on p. 110).

[101] J. E. Geiser, *Charmonium spectroscopy from radiative decaysof the J/Ψ and Ψ'*, PhD thesis: Stanford University, 1982 (cit. on p. 110).

[102] B. Bhuyan, *High P_T tracking efficiency using partially reconstructed D^* decays*, Belle Note 1165 (2010) (cit. on p. 113).

[103] K. Nishimura et al., *First Measurement of Inclusive $B \to X_s\eta$ Decays*, Phys.Rev.Lett. **105** (2010) p. 191803, arXiv: `0910.4751 [hep-ex]` (cit. on p. 114).

[104] K. A. Olive et al., *Review of Particle Physics*, Chin. Phys. **C38** (2014) p. 090001 (cit. on pp. 116, 127).

[105] D. Asner et al., *Measurement of absolute branching fractions of inclusive semileptonic decays of charm and charmed-strange mesons*, Phys.Rev. **D81** (2010) p. 052007, arXiv: `0912.4232 [hep-ex]` (cit. on p. 116).

[106] L. Pesántez et al., *Measurement of the direct CP asymmetry in $\bar{B} \to X_{s+d}\gamma$ decays with a lepton tag*, Phys. Rev. Lett. **114**.15 (2015) p. 151601, arXiv: `1501.01702 [hep-ex]` (cit. on p. 117).

[107] G. Degrassi, P. Gambino, and P. Slavich, *SusyBSG: A Fortran code for BR[B —> X(s) gamma] in the MSSM with Minimal Flavor Violation*, Comput. Phys. Commun. **179** (2008) p. 759, arXiv: `0712.3265 [hep-ph]` (cit. on p. 125).

[108] J. Charles et al., *CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories*, Eur. Phys. J. **C41** (2005) p. 1, arXiv: `hep-ph/0406184 [hep-ph]` (cit. on p. 127).

[109] A. Hocker et al., *A New approach to a global fit of the CKM matrix*, Eur. Phys. J. **C21** (2001) p. 225, arXiv: `hep-ph/0104062 [hep-ph]` (cit. on p. 127).

[110] J. C. Hardy and I. S. Towner, *Superallowed 0+ —> 0+ nuclear beta decays: A New survey with precision tests of the conserved vector current hypothesis and the standard model*, Phys. Rev. **C79** (2009) p. 055502, arXiv: `0812.1202 [nucl-ex]` (cit. on p. 127).

[111] A. Alberti et al., *Precision Determination of the Cabibbo-Kobayashi-Maskawa Element* V_{cb}, Phys. Rev. Lett. **114**.6 (2015) p. 061802, arXiv: `1411.6560 [hep-ph]` (cit. on p. 127).

[112] G. Aad et al., *Search for charged Higgs bosons decaying via* $H^{\pm} \rightarrow \tau^{\pm}\nu$ *in fully hadronic final states using pp collision data at* $\sqrt{s} = 8$ *TeV with the ATLAS detector*, JHEP **03** (2015) p. 088, arXiv: `1412.6663 [hep-ex]` (cit. on p. 128).

[113] M. Aaboud et al., *Search for charged Higgs bosons produced in association with a top quark and decaying via* $H^{\pm} \rightarrow \tau\nu$ *using pp collision data recorded at* $\sqrt{s} = 13$ *TeV by the ATLAS detector*, Phys. Lett. **B759** (2016) p. 555, arXiv: `1603.09203 [hep-ex]` (cit. on p. 128).

[114] G. Abbiendi et al., *Search for Charged Higgs bosons: Combined Results Using LEP Data*, Eur. Phys. J. **C73** (2013) p. 2463, arXiv: `1301.6065 [hep-ex]` (cit. on p. 128).

List of Figures

List of Tables

Acknowledgement

Cada paso de este largo camino ha sido acompañado por mi mamá, mi papá y toda mi familia. A ellos les agradezco por todo el apoyo y el amor sobre un océno de distancia.

I have to specially thank Jochen and Phill for sharing their knowledge and experiences, and for providing guidance in moments of scientific confusion. Thank you for the valuable tips and the help to improve my analysis and my writings. And certainly a big thank you for all the spared time writings letters and contacting stubborn administrative staff.

A particular acknowledgement to all my Bonn colleagues. The camaraderie that one experiences in this group is certainly unique, and has made the rough path of completing this thesis more enjoyable. Christian, Benedict, Will, Federico, Jan, Bruno and everyone else with whom I have shared an office, a fun lunch or a problem-plagued movie night, thank you for the good times.

To my Melbourne colleagues, thank you for being so welcoming and for providing a fantastic working environment. It was you who also added up to the charm of that city and made my not-so-short stay unforgettable. Hopefully I will be visiting soon again.

Finally, a big thank you to all the Belle colleagues. Your guidance, support and criticism have helped me greatly and made this work possible.