École doctorale
Physique, chimie-physique
ED 182
Université de Strasbourg

UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE DE PHYSIQUE ET CHIMIE PHYSIQUE
Institut Pluridisciplinaire Hubert Curien (IPHC), UMR 7178

THÈSE

présentée par:

Tristan Fillinger

soutenue le : **30 Septembre 2022**

pour obtenir le grade de: **Docteur de l'Université de Strasbourg**
Discipline/Spécialité: Physique des particules

Mesure de l'asymétrie CP dépendante du temps pour les désintégrations de mésons B en K^* gamma avec l'expérience Belle II

THÈSE dirigée par:

Pr. Jérôme Baudot	Institut Pluridisciplinaire Hubert Curien, France

RAPPORTEURS:

Pr. Adrian Bevan	Queen Mary University of London, United Kingdom
Dr. Marie-Hélène Schune	Laboratoire de physique des 2 infinis - Irène Joliot-Curie, France

EXAMINATEURS:

Pr. Boris Hippolyte	Institut Pluridisciplinaire Hubert Curien, France
Dr. Stefano Lacaprara	INFN Sezione di Padova, Italy

École doctorale
Physique, chimie-physique
ED 182
Université de Strasbourg

Tristan FILLINGER

Measurement of the time dependent CP Asymmetry of $B \to K^*\gamma$ decays in the Belle II experiment

Résumé

Le modèle standard de physique des particules, qui décrit les particules fondamentales et leurs interactions, est la théorie la plus réussie de l'histoire de la science. Pendant des décennies, il a été testé par de nombreuses collaborations scientifiques avec succès. Cependant, la théorie demeure incomplète. Afin de trouver des preuves de physique au-delà du modèle standard, les expériences sont améliorées pour recueillir de plus en plus de données avec une précision de plus en plus élevée. Cette thèse présente la proposition d'un nouveau détecteur de vertex pour l'expérience Belle II, ainsi que son implémentation logicielle. Avec Belle, cette thèse présente une nouvelle mesure dépendante du temps de la violation CP avec $B^0 \to K^0_S\pi^+\pi^-\gamma$, qui peut améliorer les contraintes de certains paramètres du modèle standard. Avec Belle II, la redécouverte de $B^0 \to K^0_S\pi^+\pi^-\gamma$ et la mesure du rapport d'embranchement de $B^0 \to K^0_S\pi^0\gamma$ avec environ 190 fb^{-1} sont décrites. Le résultat de $\mathcal{B}\left(B^0 \to K^0_S\pi^0\gamma\right) = (7.28 \pm 1.75(\text{stat}) \pm 1.03(\text{syst})) \times 10^{-6}$ est compatible avec les prédictions.

Abstract

The Standard Model of particle physics, which describes the fundamental particles and their interactions, is the most successful theory in the history of science. For decades it has been tested by numerous scientific collaborations with great success. However, the theory remains incomplete. In order to find evidence of physics beyond the Standard Model, the experiments are upgraded to gather more and more data with higher accuracy. This thesis presents the proposition of a new vertex detector for the Belle II experiment, along with its software implementation. With Belle, this thesis presents a new time-dependent measurement of the CP violation with $B^0 \to K^0_S\pi^+\pi^-\gamma$, which can improve the constrains on some Standard Model parameters. With Belle II, the rediscovery of $B^0 \to K^0_S\pi^+\pi^-\gamma$ and the branching fraction measurement of $B^0 \to K^0_S\pi^0\gamma$ with around 190 fb^{-1} is described. The result of $\mathcal{B}\left(B^0 \to K^0_S\pi^0\gamma\right) = (7.28 \pm 1.75(\text{stat}) \pm 1.03(\text{syst})) \times 10^{-6}$ is compatible with the predictions.

Remerciements

Il est crucial durant sa thèse d'avoir un bon entourage, et je pense avoir été chanceux d'en avoir eu un d'aussi grande qualité. Malgré la pandémie, ces trois dernières années étaient incroyablement riches, tant sur le plan académique que le plan amical. Je vais essayer dans ces quelques lignes de rendre hommage à toutes ces personnes qui m'ont grandement aidé dans cette aventure.

Mes premiers remerciements vont bien évidemment à Jérôme Baudot. Merci de m'avoir accompagné depuis le stage de licence et de m'avoir fait prendre goût au domaine passionnant qu'est la physique subatomique. Apprendre et travailler avec lui est une formidable expérience, de par sa passion pour le domaine, sa pédagogie, son soutien et sa bonne humeur. Sous sa direction, parfois intense mais enrichissante, j'ai pu participer et contribuer à de nombreux projets, et je lui en suis infiniment reconnaissant. Merci de m'avoir fait confiance.

Un grand merci à Giulio Dujany qui m'a énormément aidé à mener à bout mes différentes analyses. Nos nombreux échanges, souvent jusqu'à tard le soir, m'ont beaucoup apporté.

Merci Isabelle Ripp-Baudot pour son aide, ses anecdotes, ses conseils et pour son excellente gestion du groupe. J'ai vécu des choses que je n'aurai pu imaginer, et pour cela merci.

Un grand merci à tout le groupe Belle II de Strasbourg et ses anciens membres, pour leur sympathie et l'ambiance de travail chaleureuse. Merci Reem et ses tips sur le doctorat ; merci Sviat pour sa grande aide sur l'analyse de Belle et le superbe tuto "How-to do the Belle TDCPV analysis and keep yourself mentally healthy at the same time".

Je tiens à remercier la collaboration Belle II, les reviewers de mes analyses et les quelques personnes avec qui j'ai pu échanger.

Merci à Marie-Hélène Schune, Adrian Bevan, Boris Hippolyte et Stefano Lacaprara d'avoir accepté d'être dans mon jury de thèse, et pour tous leurs commentaires constructifs.

Un immense merci à mes parents, leur soutien a été proportionnel à leur non-compréhension du sujet, mais ils m'ont fait confiance jusqu'au bout et je leur en suis grandement reconnaissant.

Enfin, je remercie tous mes amis et groupes d'amis ; b 78[, Conversation éorme, Remington nous les modales avec civil, mercredi, pizzeria peperonni maccaroni al capone, aaa, cenao, 2 docteurs & 1 ingénieur, les étoiles montantes de... Je me sens très chanceux de vous avoir à mes côtés.

Contents

List of Figures

List of Tables

Introduction

The Standard Model of particle physics, which describes the fundamental particles and their interactions, is the most successful theory in the history of science. For decades it has been tested by various scientific collaborations and experiments and it predicts experimental results with a high level of precision, *e.g.* the fine structure constant at 81 parts per trillion accuracy. However, the theory is believed to be incomplete as it does not, for example, explain the presence of dark matter or dark energy in the Univers. Particle physicists now search for hints of physics beyond the Standard Model by looking at experimental deviations from the theoretical predictions.

The radiative $b \to s\gamma$ decays, which proceeds through quantum loops, are prone to reveal new effects from physics beyond the Standard Model. This process have been studied by experiments at previous B-factories (Belle and Babar) and is the topic of various measurements by the LHCb experiment. In this work, we focus on the photon polarization. Due to the V-A structure of the Standard Model, the helicity of the photon is mostly left-handed (for b-quark, and right-handed for anti-b). Measuring an increased number of right-handed photons would invalidate the Standard Model predictions. This has been done in Belle and BaBar by looking at the CP symmetry violation, like the analysis presented in Chapter 5 or Chapter 6, with no conclusive results as the measurements were not sensitive enough [1-4].

The particle physics experiments entered a precision era. Large collaborations continue to improve their complex accelerator and detectors to get large amount of data with high accuracy. The Belle experiment and the KEKB collider were upgraded to be more precise and gather fifty times more statistics. The first data were taken in 2019, and upgrades of the Belle II experiment are already planned, as discussed Chapter 4.

The thesis is organized as follows.

The first chapter summarizes the theoretical framework used in the next chapters. The Standard Model is briefly described, with a focus on the CP asymmetry, the CKM matrix and the photon polarization in radiative $b \to s\gamma$ decays.

The second chapter describes the different sources of CP violation in the B meson system. It contains the equations of the CP violation parameters used in the time dependent analysis presented in the next chapters. It also explains the usage of a new method to better constraint the modelization of new physics.

The third chapter is devoted to the description of the Belle II experiment. The SuperKEKB accelerator is described, along with each sub-detector of the Belle II detector and its trigger system.

The fourth chapter presents the plan and the work done for the upgrade of the vertex detector of Belle II. The software of Belle II and its tracking algorithm are explained. Then, the software implementation and the performances of new vertex detector geometries are shown.

The fifth chapter presents the time-dependent CP asymmetry measurement of the $B^0 \to K^0_S\pi^+\pi^-\gamma$ decay using 1 ab^{-1} of Monte-Carlo data in Belle. The analysis uses, for the first time, the new Dalitz-splitting method to further constrain the Standard Model. The chapter goes through the whole reconstruction, selection and the fitting strategy to extract the CP violating parameters.

Finally, the sixth chapter presents the rediscovery of two radiative decays with 190 fb^{-1} of Belle II data, $B^0 \to K^0_S\pi^+\pi^-\gamma$ and $B^0 \to K^0_S\pi^0\gamma$, and the branching fraction measurement of $B^0 \to (K^{*0} \to K^0_S\pi^0)\gamma$. This represents the first step to the full time-dependent CP analysis. All the details of the analysis are illustrated, from the data sample used, the reconstruction and pre-selection strategy, the selection optimization, the determination of the systematic uncertainties to the fitting procedure to extract the yields.

Chapter 1

The Standard Model and CP violation

In this chapter, the history of the discovery and the inclusion of CP violation in the Standard Model is presented. The CKM matrix is then introduced and described in Section 1.3 in order to understand how CP violation manifests itself. Finally, in Section 1.4, we focus on the $b \to s\gamma$ mathematical description since the photon polarization is of great interest to discover physics beyond the one predicted by the Standard Model.

1.1 Historical introduction

The Standard Model (SM) of particle physics is the theory describing the fundamental interactions and propagation of all elementary particles. There are 3 generations spin 1/2 fermions:

$$\begin{pmatrix} u \\ d \end{pmatrix} \begin{pmatrix} c \\ s \end{pmatrix} \begin{pmatrix} t \\ b \end{pmatrix}, \tag{1.1}$$

$$\begin{pmatrix} e \\ \nu_e \end{pmatrix} \begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix} \begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix}, \tag{1.2}$$

and their anti-particles. The quarks are designated in Eq. 1.1, each representing a different flavor, and the ordinary matter is only composed of combination of elements in the first column. The interactions manifest themselves *via* the propagation of a spin 1 boson: the γ for the electromagnetic force, interacting with photons and leptons; the $W^\pm$ and Z^0 for the weak interaction for the leptons or quarks; the gluons for the strong force, interacting with quarks. The gravitational interaction is neglected as its relative strength is 10^{25} times lower than the weak interaction and the graviton, the hypothetical boson carrying the interaction [5], is not yet discovered. Finally, the Higgs boson, which is responsible for the masses of the fundamental particles [6], was discovered in 2012 [7, 8].

Symmetries are important properties in physics. They give rise to conservation laws, *e.g.* Noether's theorem [9], structure the formalism of theories and play a major role in our understanding of physics. In this thesis, we will focus on two discrete symmetries:

- **C**: The charge conjugation transforms a particle into its anti-particle, and vice-versa : $q \to \bar{q}$, with q a fundamental particle,
- **P**: The parity transformation defined as an inversion of the spatial coordinates at the origin: $(x, y, z) \to (-x, -y, -z)$.

Parity and charge symmetries are conserved in strong and electromagnetic interaction. It was also supposed to be conserved for the weak interaction, but no experiments existed to assess this. In the mid 1956, Chen Ning Yang and Tsung-Dao Lee questioned the validity of this hypothesis [10]. A year later, Chien Shiung Wu demonstrated, using ^{60}Co weak nuclear decay, that parity is not conserved for the charged weak interaction [11]. Lee and Yang received the Nobel Prize that year for their theoretical predictions. A year later, Goldhaber showed by measuring neutrinos helicity that charge symmetry is also broken for the charged weak interaction [12]. After numerous experiments confirmed the violation, it was also demonstrated that not only the weak interaction breaks C and P, but it was also showed that the violation is maximal, meaning that only left (right) handed[1] particles (anti-particles) interact with the weak force. Physicists then put their hopes on the conservation of CP. It is also conserved for strong and electromagnetic interaction, but the weak interaction seemed to violate P and C in a way that combined CP is conserved. In 1964, James Cronin and Val Fitch demonstrated the opposite: by studying the kaons decay, they discovered a small CP violation [13] in decay of neutral kaons. This work also lead to a Nobel Prize in 1980. This result shook the scientific community and it took nearly ten years to correctly understand and add to the SM this CP violation in the weak sector. In 1973, Makoto Kobayashi and Toshihide Maskawa pointed out that in the case of three quark generations[2] the number of degrees of freedom in the SM naturally gives rise to the violating complex phase in the quark mixing matrix, now known as the Cabibbo-Kobayashi-Maskawa (CKM) matrix [14]. This work also lead to a Nobel Prize in 2008, after confirmation of their predictions by so-called B factory, the Belle and BaBar experiments, constructed to measure the large CP violation predicted in the B meson system. After a year of operation, they already observed time dependent CP violation (TDCPV) in $B^0 \to J/\psi K^0_S$ [15, 16].

Although the SM is a good description of Nature up to energies of nearly 1 TeV. But it is not sufficient to describe the entire subatomic world. For example, it does not describe neutrino masses and their oscillations, the hierarchy described by the free parameters of the CKM matrix, how gravity interacts with elementary particles. Also, Sakharvov demonstrated that CP violation is one of the necessary conditions for the observed dominance of matter over antimatter in the Universe [17], but the CP violating phase in the CKM matrix measured today in not enough, by many orders of magnitude [18]. Therefore, yet unknown additional CP violating phases must exist from the

[1] The helicity (equals to the chirality for massless particle, as photons) is right (left) handed if the direction of its spin is the same (opposite) as the direction of its motion.

[2] Only 3 quarks were discovered at that time. This matrix predicted 3 more quarks, that were discovered afterwards.

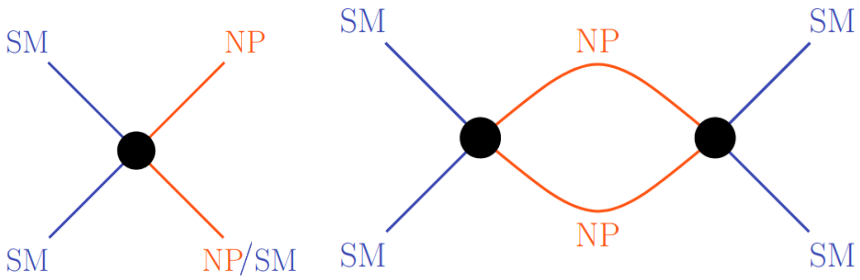

FIGURE 1.1: *Schematics of direct (left) and indirect (right) detection of New Physics.*

cosmological point of view.

Schematically, there are two ways of looking for New Physics (NP) [3], as schematized Fig. 1.1. Either by producing the NP particles (*direct* search), the Large Hadron Collider have in recent years started to probe physics at the TeV energy scale and the only limit is the center of mass beam energy; or by measuring the effect that NP can have if it contributes through virtual particles to the observed process, as high mass particles can replace SM particles in the interaction, following $\Delta E \Delta t \geqslant \hbar/2$. This later method (*indirect search*) is particularly relevant at B-factories and was employed by Belle and BaBar.

This thesis contributes to the development of experimental methods to find hints of NP by measuring the CP violation in B^0 ($\bar{B}^0$) transition with radiated photons, so-called $b \to s\gamma$ radiative penguin. The SM predicts a prevalently right (left) helicity of the photon for b-quark (anti-b-quark) decay. Measuring the photon polarization would be a strong test of the SM, since non-SM processes can introduce diagrams with different polarization associated to each b-quark flavor. Measurement of the photon polarization is possible with specific decays allowing angular distribution reconstruction [19,20]. Information on polarization can also be accessed through other phenomena. We exploit the mixing-induced CP asymmetry $\mathcal{S}$ in $B^0 \to K^0_S \pi^+ \pi^- \gamma$ and $B^0 \to K^0_S \pi^0 \gamma$, which is predicted in the SM to be small.

1.2 Theoretical overview

The SM mathematical description is the Quantum Field Theory (QFT), composed of Quantum Mechanics, Field theory and Special relativity. It is a gauge theory based on the group $SU(3)_c \otimes SU(2)_L \otimes U(1)_Y$. The electromagnetic interaction symmetry group $U(1)_{EM}$ is describing Quantum Electrodynamics (QED) theory and characterizes how electrically charged particles interact by means of exchange of photons. The strong interaction symmetry group $SU(3)_c$, with SU for Special Unitary group and c for color, is describing the Quantum Chromodynamics (QCD) theory. It characterizes the strong interaction between

[3] Physics beyond the Standard Model.

quarks mediated by gluons. Quarks can gather to form composite hadrons such as the proton (uud), neutron (udd) or pions (*e.g.* $u\bar{d}$). The QCD analog of electric charge is a property called color charge, which is conserved when quarks couple. Lastly, the weak interaction symmetry group $SU(2)_L$, with L for left-handed, is best described along QED, forming the Electroweak theory. Its interaction group is $SU(2)_L \otimes U(1)_Y$, with Y for hypercharge. The weak interaction is the only one able to change the flavor of the leptons *via* a W boson.

In the SM, mass of particles come from the Higgs mechanism [6], confirmed by the discovery of the Higgs boson in 2012 by CMS and ATLAS [7]. A complex coupling constant between the mass terms in the Lagrangian and the flavor sector in the weak interactions is responsible for the CP violation in the SM.

The Lagrangian of the SM can be expressed as:

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{Kin}} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{Yukawa}}\,, \tag{1.3}$$

where $\mathcal{L}_{\text{Kin}}$ describes the dynamics of the spinor fields Ψ representing the three fermion generations, $\mathcal{L}_{\text{Higgs}}$ describing the Higgs mechanism.

The quarks acquire mass through their Yukawa couplings to the Higgs field ϕ:

$$\mathcal{L}_{\text{Yukawa}} = -Y_u^{ij}\overline{Q_{L,i}}\epsilon\phi^* u_{R,j} - Y_d^{ij}\overline{Q_{L,i}}\phi d_{R,j} + \text{leptons term} + \text{h.c.}\,, \tag{1.4}$$

where $Q_{L,i}$, $u_{R,j}$ and $d_{R,i}$ are the quarks fields, where i and j label the generation and L (R) labels the left (right) handed fields, $Y_{u,d}$ are 3×3 complex matrices and ϵ is the anti-symmetric tensor.

The Higgs field can be expressed as:

$$\phi = \frac{1}{\sqrt{2}}\begin{pmatrix}\phi^+\\ \phi^0\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ \text{v}\end{pmatrix}, \tag{1.5}$$

with v being the vacuum expectation value. The mass matrices are then:

$$M_u^{ij} = (v/\sqrt{2})Y_u^{ij}, \quad M_d^{ij} = (v/\sqrt{2})Y_d^{ij}, \tag{1.6}$$

where matrices M are 3×3 and generally non-diagonal matrices since there is no prescription or symmetry constraining the content of Y. To obtain the mass eigenstates, we have to diagonalize each of the mass matrices with the help of two unitary matrices:

$$\mathcal{U}_L^\dagger M \mathcal{U}_R = M_{\text{diag}}\,, \tag{1.7}$$

We obtain the mass eigenstates as:

$$\begin{aligned} u^f_{L(R)} &= \mathcal{U}^u_{L(R)} u^m_{L(R)}, \\ d^f_{L(R)} &= \mathcal{U}^d_{L(R)} d^m_{L(R)} \end{aligned} \tag{1.8}$$

where $u^f = \begin{pmatrix} u \\ c \\ t \end{pmatrix}$ and $d^f = \begin{pmatrix} d \\ s \\ b \end{pmatrix}$ represent the up or $down-type$ quarks in the flavor space (f), and the u^m and d^m represent the corresponding mass (m)

eigenstates. The unitary matrices are also chosen to give:

$$\begin{aligned} \mathcal{U}_R^{u\dagger} M_u \mathcal{U}_L^d &= \mathrm{diag}(m_u, m_c, m_t), \\ \mathcal{U}_R^{d\dagger} M_d \mathcal{U}_L^d &= \mathrm{diag}(m_d, m_s, m_b). \end{aligned} \tag{1.9}$$

When we rewrite the Lagrangian in the mass basis, we find that this change of basis does not change the form of the electromagnetic and neutral-current (Z and γ), thanks to the unitary property of $\mathcal{U}$. For the charged-current ($W^{\pm}$) we have:

$$\mathcal{J}_\mu^+ = \bar{u}_L^f \gamma_\mu d_L^f = \bar{u}_L^m \gamma_\mu \mathcal{U}_R^{u\dagger} \mathcal{U}_L^d d_L^m \equiv \bar{u}_L^m \gamma_\mu V_{\mathrm{CKM}} d_L^m, \tag{1.10}$$

where V_{CKM} is the *Cabibbo-Kobayashi-Maskawa* matrix [14].

1.3 The CKM matrix

The CKM matrix describes the coupling of quarks from different flavors and can be represented as:

$$V_{\mathrm{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}, \tag{1.11}$$

where V_{xy} represents the coupling value from the quark x to the quark y.

By construction, V_{CKM} is a 3×3 complex and unitary matrix and has at maximum 18 real parameters. By using the unitary condition we can reduce them by 9, and by using the properties that the phases of the quarks can be redefined, we remove additional 5 parameters. This leads to three mixing angles and one irreducible phase, the later being the main origin of CP violation in the SM.

The *Standard parametrization* of the V_{CKM} matrix was proposed by Chau and Keung [21], using rotation matrices and an overall phase:

$$V_{\mathrm{CKM}} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{13}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{13}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{13}} & c_{23}c_{13} \end{pmatrix} \tag{1.12}$$

where $c_{ij} = \cos\theta_{ij}$, $s_{ij} = \sin\theta_{ij}$ for $i < j = 1, 2, 3$, θ_{ij} are the rotation angles between families and δ_{13} is the irreducible phase.

Thanks to experimental results, we know that this matrix is dominated by its diagonal. Transitions among quarks from the same generation (V_{ud}, V_{cs}, V_{tb}) are far more likely than transitions among quarks from different flavor doublets. A new parametrization, named the *Wolfenstein parametrization* [22] referring its author, takes into account this observed hierarchy by expanding the elements of the CKM matrix in powers of $\lambda \equiv |V_{us}|$:

$$V_{\text{CKM}} = \begin{pmatrix} 1-\frac{1}{2}\lambda^2 & \lambda & \lambda^3 A(\rho - i\eta) \\ -\lambda & 1-\frac{1}{2}\lambda^2 & \lambda^2 A \\ \lambda^3 A(1-\rho-i\eta) & -\lambda^2 A & 1 \end{pmatrix} + \mathcal{O}\left(\lambda^4\right) \tag{1.13}$$

where the real parameters A, ρ and η are at the order of unity, η being non-zero introduces the CP violation, and $\lambda \simeq 0.22$ [23]. By measuring precisely those parameters, we can get more information on the CP violation and on charged-current.

A nice way of visualizing the CKM matrix is to use its equations derived from the unitarity condition. Out of the 9 possible relations, 3 of them are of particular interest to study CP violation:

$$V_{ud}V_{us}^* + V_{cd}V_{cs}^* + V_{td}V_{ts}^* = 0 \quad \left(\propto \lambda\lambda\lambda^5\right), \tag{1.14}$$

$$V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{tb}^* = 0 \quad \left(\propto \lambda^3\lambda^3\lambda^3\right), \tag{1.15}$$

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \quad \left(\propto \lambda^4\lambda^2\lambda^2\right). \tag{1.16}$$

Each term in these equations can be interpreted geometrically as a vector in the complex plane, thus each relation can be visualized as a triangle, also known as the *unitarity triangles*. By looking at the quarks involved in each equation, we can see that Eq. 1.14 and Eq. 1.15 describes the $K^0 - \bar{K}^0$ and $B_s^0 - \bar{B}_s^0$ systems respectively. If we draw the two triangles, we observe that one side is much shorted than the other two, resulting in a collapse of the triangle to a line. Eq. 1.16, related to the $B^0 - \bar{B}^0$ system, is the most opened triangle. This illustrates a large CP asymmetry for B^0 decays, which are studied in B factories.

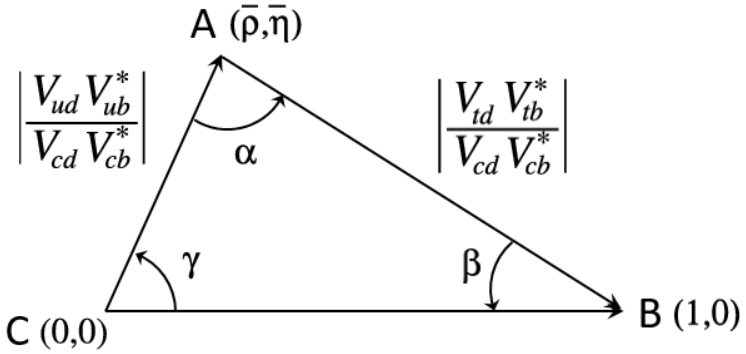

FIGURE 1.2: *Representation in the complex plane of the Eq. 1.16 divided by $V_{cd}V_{cb}^*$. Their sum being equal to zero implies that they have to form a closed triangle.*

Dividing the last equation by $V_{cd}V_{cb}^*$ gives the equation used to represent the triangle shown Fig. 1.2. The vertices are now at $(0,0)$, $(1,0)$ and $(\bar{\rho},\bar{\eta})$, with $\bar{\rho} = \rho\left(1-\lambda^2/2\right)$ and $\bar{\eta} = \eta\left(1-\lambda^2/2\right)$. The side length are expressed with CKM matrix elements:

$$\overline{AB} = \left| \frac{V_{td} V_{tb}^*}{V_{cd} V_{cb}^*} \right|, \tag{1.17}$$

$$\overline{AC} = \left| \frac{V_{ud} V_{ub}^*}{V_{cd} V_{cb}^*} \right|, \tag{1.18}$$

$$\overline{CB} = 1, \tag{1.19}$$

and the three angles:

$$\alpha = \arg \left(\frac{V_{cd} V_{cb}^*}{V_{td} V_{tb}^*} \right), \tag{1.20}$$

$$\beta = \arg \left(\frac{V_{td} V_{tb}^*}{V_{cd} V_{cb}^*} \right), \tag{1.21}$$

$$\gamma = \arg \left(\frac{V_{ud} V_{ub}^*}{V_{cd} V_{cb}^*} \right). \tag{1.22}$$

Measuring and constraining these five parameters of the triangle is of great importance: they constrain the free parameters in the SM, they provide an experimental test of the CP violation mechanism and the flavor structure of the SM, and they can provide a way to discover NP by measuring discrepancies *wrt.* the SM. B factories have, in the last two decades, measured decay rates and CP asymmetries in the B system with great precision to put constraints on the position of $(\bar{\rho}, \bar{\eta})$, shown in Fig. 1.3. So far, all constraints coming from the measurements of very different processes are in very good agreement, meaning that the SM flavor structure is very well described.

1.4 Photon polarization and Wilson coefficient

The NP can appear in the study of flavor-changing neutral current (FCNC) processes such as $b \to s\gamma$ transitions, as illustrated Fig. 1.4. The FCNC are forbidden at tree level and thanks to the GIM mechanism only occur at one loop level *via* a penguin or box diagrams (as in Fig. 2.1), where a new heavy particle can enter the loop and alter the SM predictions. As stated before, the photon emitted in $b \to s\gamma$ decays are left-handed, with a small right-handed portion of the order m_s/m_b. In some NP models [25-27], the right-handed component can be enhanced and reach the same magnitude as the left-handed component, thanks to the presence of NP particles in the electroweak loop process.

In the SM, the quark level $b \to s\gamma$ vertex is given, without QCD corrections, by:

$$\bar{s}\Gamma_\mu^{b\to s\gamma} b = \frac{e}{(4\pi)^2} \frac{g^2}{2M_W^2} V_{ts}^* V_{tb} F_2 \bar{s} i\sigma_{\mu\nu} q^\nu \left(m_b P_R + m_s P_L \right) b, \tag{1.23}$$

where $q = q_b - q_s$ with q_b and q_s the four-momenta of the b and s quarks respectively, F_2 is the loop function defined in [28], V_{ts} and V_{tb} are the leading

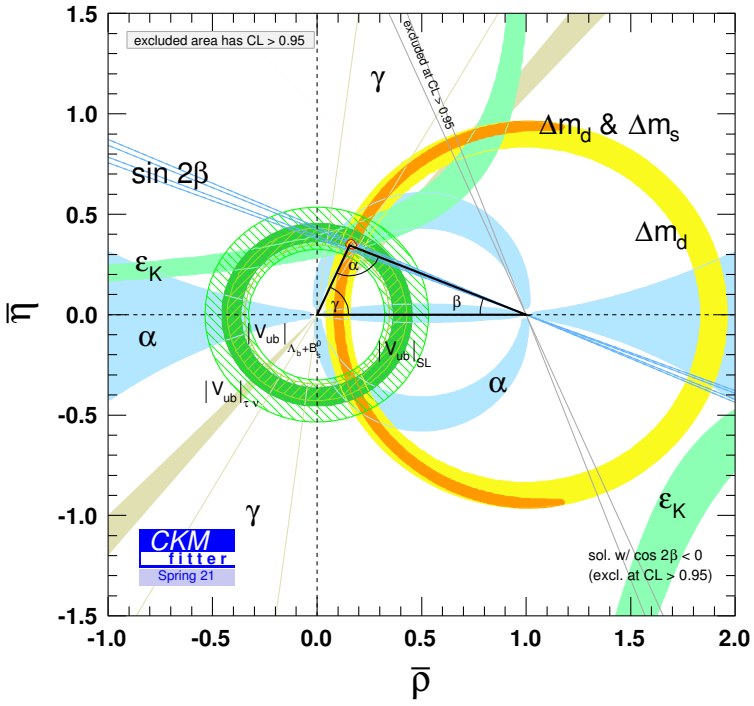

FIGURE 1.3: *Constraints on the CKM $(\bar{\rho}, \bar{\eta})$ coordinates from the global SM CKM-fit by the CKMfitter group [24]. Each band color represents a different measurements to constrain the $(\bar{\rho}, \bar{\eta})$ apex, which is given by their overlap.*

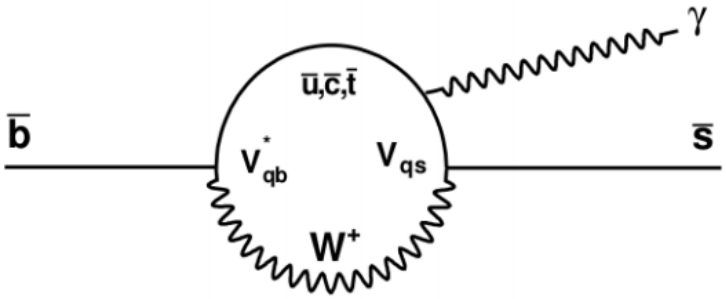

FIGURE 1.4: *Feynman diagram of $\bar{b} \to \bar{s}\gamma$ through an electroweak loop process, expressed with CKM matrix element at the vertex. Due to GIM mecanism, the loop is dominated by the top quark. In this two-body back-to-back decay transition in the $\bar{b}$ rest frame, the helicity of the γ must be right-handed as the helicity is conserved in the SM.*

CKM matrix elements appearing in the loop diagram, $\sigma_{\mu\nu}$ are the commutators of the Dirac γ 4×4 matrices and P_L and P_R are the projection operators defined as:

$$P_L = \frac{1}{2}(1 - \gamma_5), \quad P_R = \frac{1}{2}(1 + \gamma_5). \tag{1.24}$$

In this equation, we observe that the term proportional to m_b describes $b \to s\gamma_L$ transitions, while the one proportional to m_s describes $b \to s\gamma_R$ transitions. This is why, in the SM, the $b \to s\gamma_L$ transitions are predominant, with a small fraction of right-handed photons, as the b-quark is much heavier than the s quark. Similarly, $\bar{b} \to \bar{s}\gamma_R$ transitions are predominant, with a small fraction of left-handed photons.

If we now go from the quark level transition to a cascade decays like $B \to K_{\text{res}}\gamma \to K_S^0\pi^+\pi^-\gamma$ or $K_S^0\pi^0\gamma$, the weak $B \to K_{\text{res}}\gamma$ amplitudes need to include such QCD corrections:

$$\mathrm{H}_{\text{eff}} = -\frac{4G_F}{\sqrt{2}} V_{ts}^* V_{tb} \sum_{i=1}^{8} C_i(\mu)\mathcal{O}_i(\mu) + C_i'(\mu)\mathcal{O}_i'(\mu), \tag{1.25}$$

where the C_i and $\mathcal{O}_i$ are the short-distance Wilson coefficients and local long distance operators related to $b \to s\gamma_L$ transitions respectively, the primed operators being related to $b \to s\gamma_R$ transitions. G_F is the Fermi constant, which is related to the coupling constant of the weak interaction (g) and the mass of the W boson (M_W) such as $G_F \propto g^2/M_W^2$, and μ is the renormalization scale. The Wilson coefficients C_i can be regarded as coupling constants associated to the effective point-like vertices which are represented by the local operators $\mathcal{O}_i$.

The interesting coefficients and operators for the description in $b \to s\gamma$ for left and right handed photon are $\mathcal{O}_7$ and $\mathcal{O}_7'$ respectively. In Eq. 1.25, $\mathcal{O}_7$ is the dominant operator at the leader order, and we have

$$\frac{C_7'}{C_7} = \frac{m_s}{m_b} \sim 0.02. \tag{1.26}$$

In the experiment results, C_7 (C_7') is usually referred as $C_7'^{\text{eff}}$ and corresponds to the effective coefficient which absorbs all the subleading contribution from $\mathcal{O}_{i\neq 7}$. Eq. 1.26 still holds with this approximation.

Constraining C_7 and C_7' to see if they deviate from the SM is a good way to probe NP, as they are experimental observables sensitive to the structure of the effective Hamiltonian for $b \to s\gamma$.

Several methods and channels have been studied in the past to put constraints on the $\mathrm{Im}(C_7'/C_7)$ and $\mathrm{Re}(C_7'/C_7)$ plane. The covered area is shown Fig. 1.5. Belle and BaBar performed inclusive $B \to X_s\gamma$ decay [30-33], where NP can enhance its branching ratio, and mixing-induced CP asymmetry of $B^0 \to K_S^0\pi^0\gamma$ [1,3,29], whereas LHCb performed also a mixing-induced CP asymmetry measurement on $B_s^0 \to \phi\gamma$ [34] and an angular analysis of $B^0 \to K^{*0}e^-e^-$ at low q^2 [35], where its angular distribution is sensitive to photon polarization. This last channel is the most recent and constraining measurement to date. A good agreement is observed with respect to the SM, but the CP asymmetry measurement in the $B^0\overline{B}^0$ system, described in detail in the next chapter, is a promising way to constrain the Wilson coefficients. All the measurements are currently statistically limited. LHCb will double its data in the next three years, and the next statistical jump will be with the run4 along with its upgrade program around 2030. For Belle II, it will accumulate around ten times the Belle and BaBar data in the next three years.

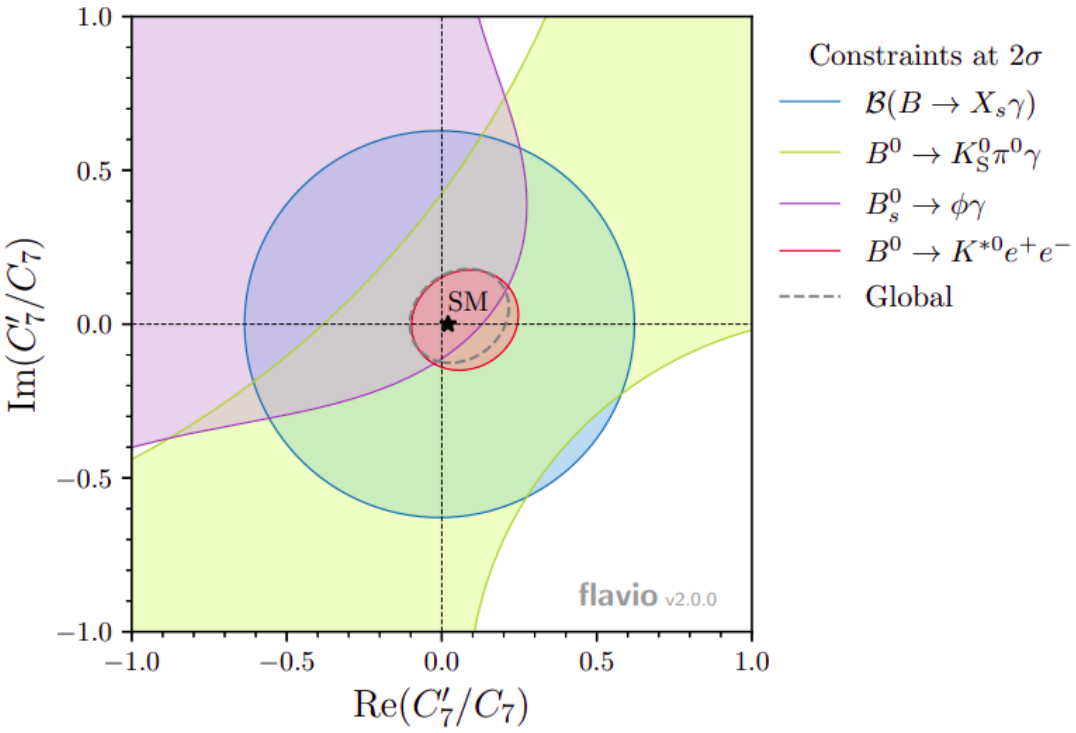

FIGURE 1.5: *Summarized constraints on Wilson coefficients on* $\mathrm{Re}(C_7'/C_7)$ *and* $\mathrm{Im}(C_7'/C_7)$ *with different methods and channels, using Belle [1, 29, 30], BaBar [3, 31-33] and LHCb data [34, 35], the latter being the most constraining measurement to date. From [35].*

Chapter 2

CP violation in the B meson system

This chapter is focused on the neutral B meson system and demonstrates how to extract CP violation observables from different processes, using features of the flavor and weak interaction structure of the SM. The theory is described in detail in [36] and [37]. In Section 2.3, we will also go into details on a new method, tested for the first time in the analysis described Chapter 5, to constrain SM parameters to probe NP.

2.1 $B^0\overline{B}^0$ mixing

The B^0 ($\bar{B}^0$) meson, with a flavor eigenstate $\bar{b}d$ ($\bar{d}b$), is a hadron decaying due to the weak interaction, with a lifetime of $\tau_B = (1.519 \pm 0.007)$ ps [23]. A neutral meson can be described in two ways: the flavor eigenstates, describing the quarks content and useful to understand the weak decay, and the mass eigenstates, solutions of the Hamiltonian, describes the mass, lifetime and the propagation in space and time. As we saw in the previous chapter, CP and mass eigenstates can differ since the CP symmetry is broken in the SM.

To get the mass eigenstates, let's consider an arbitrary combination of the neutral B-meson flavor eigenstates:

$$a(t)\left|B^0\right\rangle + b(t)\left|\bar{B}^0\right\rangle \tag{2.1}$$

which is governed by the time-dependent Schrödinger equation:

$$i\frac{d}{dt}\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = \mathrm{H}\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} \equiv (M - i\Gamma)\begin{pmatrix} a(t) \\ b(t) \end{pmatrix}, \tag{2.2}$$

where M and Γ are two 2×2 hermitian matrices known as mass and decay matrices respectively. The assumption of CPT invariance constraints the Hamiltonian and imposes: $M_{11} = M_{22} = M$, $\Gamma_{11} = \Gamma_{22} = \Gamma$, $M_{12} = M_{21}$ and $\Gamma_{12} = \Gamma_{21}$. The off-diagonal terms are important in the discussion of CP violation: if it is conserved, then $\mathrm{Im}(M_{12}) = 0 = \mathrm{Im}(\Gamma_{12})$.

Diagonalizing the Hamiltonian leads to the expression of the mass eigenstates of the neutral B meson system, which consist in a superposition of the B^0 and $\bar{B}^0$ flavor eigenstates, such as:

$$|B_L\rangle = p\left|B^0\right\rangle + q\left|\bar{B}^0\right\rangle, \tag{2.3}$$

$$|B_H\rangle = p\left|B^0\right\rangle - q\left|\bar{B}^0\right\rangle, \tag{2.4}$$

where the index $L(H)$ refers to the lighter (heavier) mass eigenstate and the complex coefficients p and q are constrained by a normalization condition

$$|p|^2 + |q|^2 = 1. \tag{2.5}$$

In the expression of the mass eigenstates, we see that the flavor eigenstates get mixed as they propagate in time and space. This phenomenon has been observed in Kaons [38], B_d and B_s mesons [39-41] and D mesons [42-44]. The mass matrix M in Eq. 2.2 describes the oscillation between the flavor eigenstates: the diagonal terms are dominated by the flavor eigenstates masses while the off-diagonal terms represent the transitions *via* virtual intermediate states. For the decay matrix Γ, its diagonal is also linked to the flavor states and describes the decay into different final states, while its off-diagonal terms describe transitions *via* real intermediate states. The $B^0\overline{B}^0$ mixing, where the B meson change its flavor with a weak transition, can be visualized by Feynman diagrams shown in Figure 2.1, referred as box diagrams.

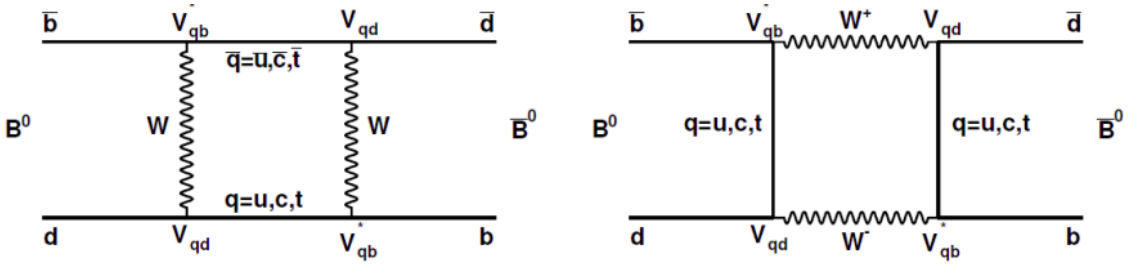

FIGURE 2.1: *Feynman diagrams representing the $B^0\overline{B}^0$ mixing phenomenon.*

The three up-type quarks involved in the loop lead to a CKM factor that is proportional to λ^6 and this loop should be suppressed. However, this transition is enhanced by the GIM mechanism [45], and the contribution of each quarks in the intermediate state is suppressed by the ratio of its mass and the weak boson [36]. The top quark is then leading this transition and the corresponding CKM factors gives a phase of $\left(V_{td}V_{tb}^*\right)^2 \sim e^{-i2\beta}$, with β called the *mixing angle*. When the B^0 oscillates into a $\bar{B}^0$, it picks up an extra -2β phase *wrt.* to a B^0 that does not oscillate. This is measurable if both flavors decay into a CP final state, common to B^0 and $\bar{B}^0$, and if the flavor at production is known.

The eigenstates of Eq. 2.2 gives the following relation on M and Γ matrix elements:

$$\frac{q}{p} = \pm\sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12}^* - \frac{i}{2}\Gamma_{12}}}, \tag{2.6}$$

$$(\Delta m_d)^2 - \frac{1}{4}(\Delta\Gamma)^2 = 4\left(|M_{12}|^2 - \frac{1}{4}|\Gamma_{12}|^2\right), \tag{2.7}$$

$$\Delta m_d \Delta\Gamma = 4\,\mathrm{Re}\left(M_{12}\Gamma_{12}^*\right), \tag{2.8}$$

with

$$\Delta m_d \equiv M_H - M_L > 0, \quad \Delta\Gamma \equiv \Gamma_H - \Gamma_L. \tag{2.9}$$

In Eq. 2.6, choosing the negative sign rather than the positive is equivalent to interchanging the labels 1 and 2 of the mass eigenstates. Thus, the ratio q/p is finally given by:

$$\frac{q}{p} = -\frac{\Delta m_d - \frac{i}{2}\Delta\Gamma}{2\left(M_{12} - \frac{i}{2}\Gamma_{12}\right)}. \tag{2.10}$$

If we look at these equations with the B^0 parameters, we can make the assumption that $\Delta\Gamma_{B^0} \ll \Gamma_{B^0}$, as $\Delta\Gamma_{B_d}$ value is produced by decay amplitudes of CP channels, with corresponding branching fractions at the level of 10^{-3}. And since many channels contribute with different signs, their sum is not expected to exceed the individual level.

Also, we have the measurement of the B^0 oscillation parameter x_d [46]:

$$x_d \equiv \frac{\Delta m_d}{\Gamma_{B_d}} = 0.775 \pm 0.006, \tag{2.11}$$

which implies $\Delta\Gamma_{B_d} \ll \Delta m_d$.

By putting the two model-independent assumptions in Eq. 2.7 and Eq. 2.9, we get:

$$|\Gamma_{12}| \ll |M_{12}|, \tag{2.12}$$

$$|q/p| \sim 1. \tag{2.13}$$

We can express the time evolution of a $B^0(t)$ ($\bar{B}^0(t)$) by using the eigenvalues of the Hamiltonian and the p and q parameters:

$$\left|B^0(t)\right\rangle = g_+(t)\left|B^0\right\rangle + \frac{q}{p}g_-(t)\left|\bar{B}^0\right\rangle, \tag{2.14}$$

$$\left|\bar{B}^0(t)\right\rangle = g_+(t)\left|\bar{B}^0\right\rangle + \frac{p}{q}g_-(t)\left|B^0\right\rangle, \tag{2.15}$$

with

$$g_+(t) = e^{-\frac{i}{2}(M_H+M_L)t} e^{-\frac{\Gamma}{2}t} \cos\left(\frac{\Delta m_d t}{2}\right) \tag{2.16}$$

$$g_-(t) = e^{-\frac{i}{2}(M_H+M_L)t} e^{-\frac{\Gamma}{2}t} i \sin\left(\frac{\Delta m_d t}{2}\right). \tag{2.17}$$

The probability of *e.g.* a B^0 to change or not its flavor is then:

$$P_{B^0\to\bar{B}^0}(t) = \left|\frac{q}{p}\right|^2 |g_-(t)|^2 \propto \left|\frac{q}{p}\right|^2 e^{-\Gamma t}\left(1-\cos\left(\Delta m_d t\right)\right), \tag{2.18}$$

$$P_{B^0\to B^0}(t) = |g_+(t)|^2 \propto e^{-\Gamma t}\left(1+\cos\left(\Delta m_d t\right)\right) \tag{2.19}$$

Since the B meson lifetime and the frequency of oscillation $\Delta m_d = (0.507 \pm 0.005)$ ps^{-1} [23] are close, the amplitude for the mixing process is rather large. This leads to large time-dependent CP asymmetries in the $B^0\bar{B}^0$ mixing phenomenon, described in the next section.

2.2 Source of CP violation

We can find three manifestations of CP violation for the B^0 system, pictured Fig. 2.2, each with an associated observable.

2.2.1 CP violation in decay

The first CP violation can occur when the amplitude for a decay and its CP conjugate process have different magnitudes. If we note $A_f = \langle f|\mathrm{H}|B^0\rangle$ $\left(\bar{A}_{\bar{f}} = \langle \bar{f}|\mathrm{H}|\bar{B}^0\rangle\right)$ the amplitude of the $B^0(\bar{B}^0)$ decay to a final state $f(\bar{f})$, the CP observables associated are:

$$\frac{\Gamma\left(B^0\to f\right)}{\Gamma\left(\bar{B}^0\to \bar{f}\right)} \neq 1 \Leftrightarrow \frac{A_f}{\bar{A}_{\bar{f}}} \neq 1 \Leftrightarrow CP \text{ violation.} \tag{2.20}$$

or the CP asymmetry definition:

$$\mathcal{A}_{CP} = \frac{\Gamma(\bar{B}\to\bar{f}) - \Gamma(B\to f)}{\Gamma(\bar{B}\to\bar{f}) + \Gamma(B\to f)} \neq 0 \Leftrightarrow CP \text{ violation.} \tag{2.21}$$

This type of CP violation is referred to as *direct CP violation* and can be found to not only $B^0\bar{B}^0$ system but to any b-hadron.

2.2.2 CP violation in mixing

The second type of CP violation occurs when the neutral B meson oscillates into its CP conjugate. As we saw in Sec. 2.1, the CP and mass eigenstate are not identical and an asymmetry arises between their transition probabilities, *i.e.* $B^0 \to$

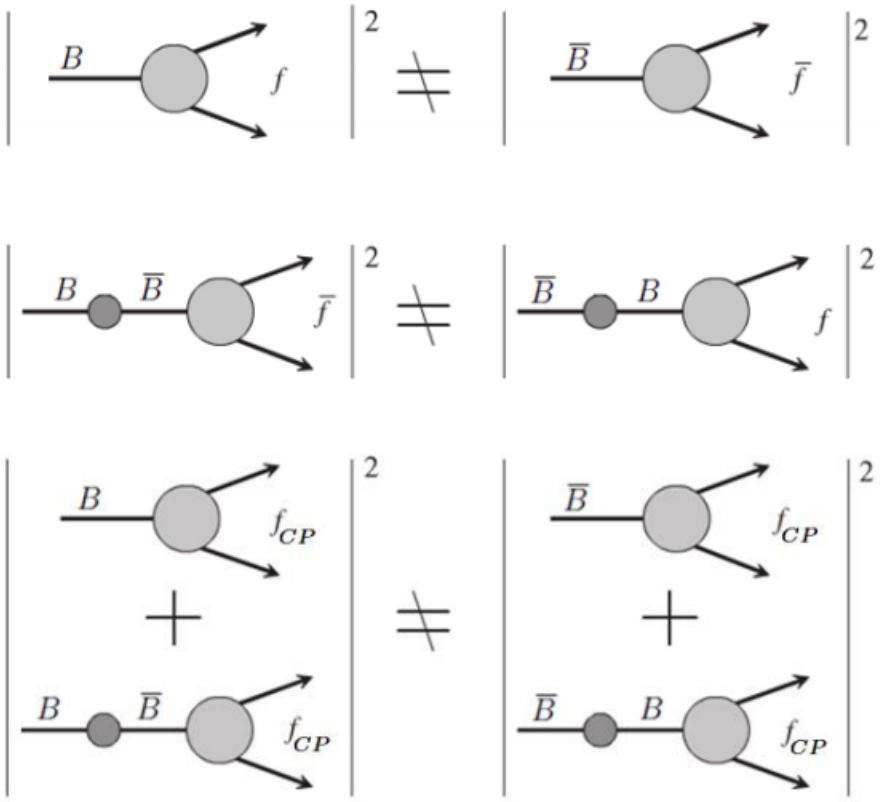

FIGURE 2.2: *Pictural representation of the different CP violation manifestation in the $B^0\bar{B}^0$ system: direct, indirect and in the interference (from top to bottom); with f representing any final state and f_{CP} a CP final state accessible to both B^0 and $\bar{B}^0$. Adapted from [36].*

$\bar{B}^0$ and $\bar{B}^0 \to B^0$. The description of such transition is carried by the off-diagonal matrix elements of M and Γ in Eq. 2.2. If CP is conserved, the relative phase between M_{12} and Γ_{12} vanishes. Thus, Eq. 2.6 gives the CP observable associated to this process:

$$\left|\frac{q}{p}\right| \neq 1 \Leftrightarrow CP \text{ violation} \tag{2.22}$$

This type of CP violation is referred to as *indirect CP violation*. The first CP violation observed in the Kaon system described in Chapter 1 was an indirect CP violation. For the B^0 system, this CP violation process is expected to be $\mathcal{O}\left(10^{-4}\right)$ [47] and not accessible by the current B factory sensitivities, so the approximation $|q/p| \approx 1$ remains.

2.2.3 CP violation in the interference between decays with and without mixing

Finally, the last possibility of CP violation is through the interference between the mixing amplitude $A\left(B^0 \to \bar{B}^0 \to f_{CP}\right)$ and the decay amplitude $A\left(B^0 \to f_{CP}\right)$

and can be resumed as

$$\frac{\Gamma\left(B^0 \to f_{CP}\right)}{\Gamma\left(\overline{B}^0 \to f_{CP}\right)} \neq 1, \tag{2.23}$$

with f_{CP} a CP final state, accessible to both B^0 and $\overline{B}^0$.

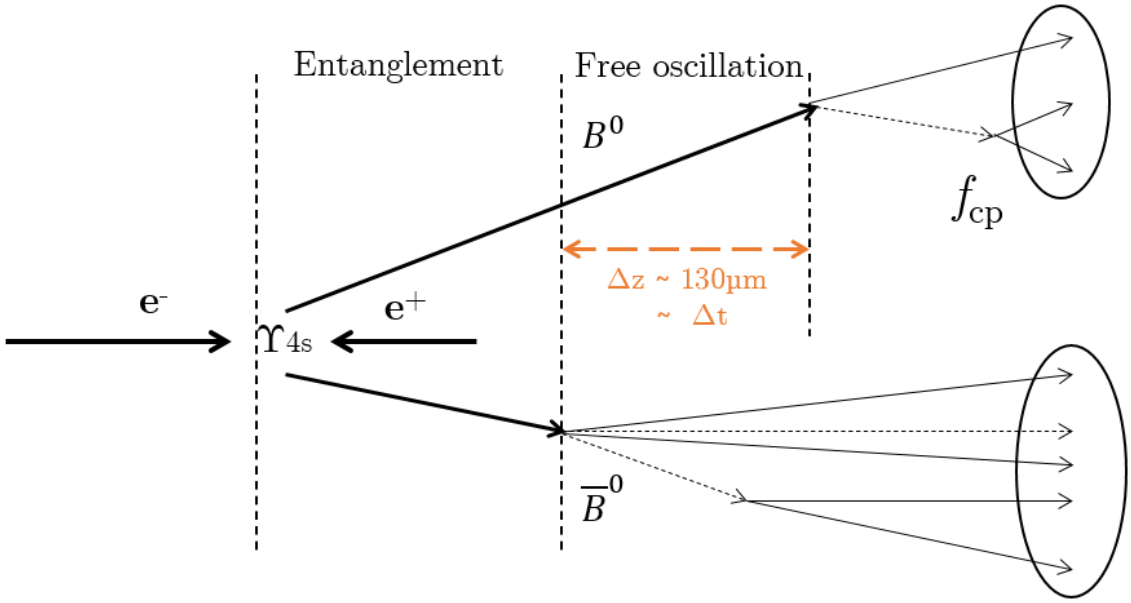

FIGURE 2.3: *Pictural representation of a e^+e^- collision producing a $B^0\overline{B}^0$ pair through a $\Upsilon(4S)$ resonnance. First, the B^0 and $\overline{B}^0$ are entangled as they come from the same resonnance, and they oscillate at the same time as the other. Then, when one of the B decays (here the $\overline{B}^0$), the other can oscillate freely until it decays into a CP final state.*

An observable similar to the one used for direct CP violation could be used here, but it is not trivial to know the flavor of the B meson in this case as its flavor changes over time. To get this information, the B factories benefit from the fact that the B meson pair is produced from a single resonance $\Upsilon(4S) \to B\overline{B}$, as shown in Fig. 2.3. They are then *entangled* until one of them decays, implying that at an instant t, the probability of one of the mesons to be a particle equals the probability for the other meson to be an antiparticle. This way, if we know the flavor of the B meson at an instant t, the other one must be at the opposite flavor at the same instant t. This process stops only when one of the B decays, the other is then oscillating freely, following the mixing principle, until its decay. If a B, noted B_{tag}, is decaying into a final state that can unambiguously determine its flavor, the flavor of the other B, noted B_{rec}, can be determined with a certain probability by taking into account a precise measurement of the time elapsed (Δt) between its decay and the moment that its flavor is known, corresponding to the moment when the other B decays. This method is known as *flavor tagging* [48] (see Sec. 5.2.6). Thus, the CP asymmetry Eq. 2.21 can be defined with a time-dependence:

$$\mathcal{A}_{CP}(\Delta t) = \frac{\Gamma\left(B_{\text{tag}=B^0}(\Delta t) \to f_{CP}\right) - \Gamma\left(B_{\text{tag}=\bar{B}^0}(\Delta t) \to f_{CP}\right)}{\Gamma\left(B_{\text{tag}=B^0}(\Delta t) \to f_{CP}\right) + \Gamma\left(B_{\text{tag}=\bar{B}^0}(\Delta t) \to f_{CP}\right)}. \tag{2.24}$$

If we neglect experimental biases, such as the probability to wrongly determine the flavor of the B_{tag}, called *mistagging*, the amplitudes for the processes of a B^0 ($\bar{B}^0$) going to a CP final state f_{CP} are:

$$A_{f_{CP}}(\Delta t) = A_{f_{CP}}\left(g_+(\Delta t) + \lambda_{f_{CP}} g_-(\Delta t)\right), \tag{2.25}$$

$$\bar{A}_{f_{CP}}(\Delta t) = \frac{p}{q} A_{f_{CP}}\left(g_-(\Delta t) + \lambda_{f_{CP}} g_+(\Delta t)\right), \tag{2.26}$$

with $g_+(\Delta t)$ and $g_-(\Delta t)$ defined Eq. 2.16 and Eq. 2.17 respectively and the CP observable associated to this process is defined as:

$$\lambda_{f_{CP}} = \frac{q}{p}\frac{\bar{A}_{f_{CP}}}{A_{f_{CP}}} \neq 1 \Leftrightarrow CP \text{ violation} \tag{2.27}$$

By construction of $\lambda_{f_{CP}}$, we can see that the previous inequality holds if CP violation occurs in the decay or the mixing, $\left|\bar{A}_{\bar{f}}/A_f\right| \neq 1$ and $|q/p| \neq 1$ respectively. In addition, even if $\left|\bar{A}_{\bar{f}}/A_f\right| = |q/p| = 1, CP$ violation could still occur through:

$$\left|\lambda_{f_{CP}}\right| = 1, \quad \text{Im}\left(\lambda_{f_{CP}}\right) \neq 0. \tag{2.28}$$

The time-dependent decay rate expression used to measure CP asymmetry in B factories is expressed as:

$$f_{q_{\text{tag}}}(\Delta t) = \frac{e^{-|\Delta t|/\tau}}{4\tau}\left[1 + q_{\text{tag}}\left(\mathcal{S}\sin\left(\Delta m_d \Delta t\right) + \mathcal{A}\cos\left(\Delta m_d \Delta t\right)\right)\right], \tag{2.29}$$

where $q_{\text{tag}} = +1(-1)$ when the B_{tag} is identified as a B^0 $\left(\bar{B}^0\right)$, and the expression of the CP asymmetry in mixing and decay can be found as:

$$\mathcal{A}_{CP}(\Delta t) = \mathcal{S}\sin\left(\Delta m_d \Delta t\right) + \mathcal{A}\cos\left(\Delta m_d \Delta t\right) \tag{2.30}$$

where

$$\mathcal{S} = \frac{2\,\text{Im}\left(\lambda_{f_{CP}}\right)}{1 + \left|\lambda_{f_{CP}}\right|^2}, \quad \mathcal{A} = \frac{\left|\lambda_{f_{CP}}\right|^2 - 1}{\left|\lambda_{f_{CP}}\right|^2 + 1}. \tag{2.31}$$

The coefficients $\mathcal{S}$ and $\mathcal{A}$ are different from zero when CP violation occurs in the interference or in direct CPV respectively.

Eq. 2.29 may be re-written in terms of $\langle w \rangle$ and $\Delta w/2$ factors, described in the next chapter Sec. 5.2.6, accounting for the probability for a B meson to be experimentally mistagged:

$$f_{q_{\text{tag}}}(\Delta t) = \frac{e^{-|\Delta t|/\tau}}{4\tau}\left[1 + q_{\text{tag}}\frac{\Delta w}{2} + q_{\text{tag}}\langle w \rangle \left(\mathcal{S}\sin\left(\Delta m_d \Delta t\right) + \mathcal{A}\cos\left(\Delta m_d \Delta t\right)\right)\right], \tag{2.32}$$

where $q_{\text{tag}} = +1(-1)$ when the B_{tag} is identified as a B^0 $\left(\bar{B}^0\right)$.

$\mathcal{A}_{CP}(\Delta t)$ becomes:

$$\mathcal{A}_{CP}(\Delta t) = \langle w \rangle \left(\mathcal{S}\sin\left(\Delta m_d \Delta t\right) + \mathcal{A}\cos\left(\Delta m_d \Delta t\right)\right). \tag{2.33}$$

Equations 2.21 and 2.30 or 2.33 are strictly equivalent. Thus, by measuring experimentally $\mathcal{A}_{CP}(\Delta t)$ with a precise Δt resolution and tagging efficiency, $\mathcal{A}$ and $\mathcal{S}$ can be deduced.

If there is only one SM contribution to the amplitudes A_f and $\bar{A}_{\bar{f}}$, the expectations are that $\mathcal{S} = -\sin(2\beta)$ and $\mathcal{A} = 0$. Deviations from that imply the existence of unaccounted amplitudes that, depending on the characteristics of the mode, could be an indication of physics beyond the SM.

For $B^0 \to K^0_S\pi^0\gamma$ decay process, where $B^0 \to K^0_S\pi^0\gamma$ goes through a $K^*(892)^0$ resonance, the SM predicts a CP asymmetry of $\mathcal{S} \simeq -0.028$ [49, 50]. For $B^0 \to K^0_S\rho^0\gamma$ decay process where ρ^0 goes to $\pi^+\pi^-$, the SM predicts $\mathcal{S} \simeq 0.03$ [51]. Both BaBar and Belle have performed measurements of the direct and mixing-induced CP asymmetry parameter $\mathcal{A}$ and $\mathcal{S}$. The results are shown Table 2.1. All the values are statistically limited and are in good agreement with the SM prediction, this is why we can benefit from new B factories like Belle II to get more precise measurements and see if there are deviations from the SM.

TABLE 2.1: *Measurements of $\mathcal{A}$ and $\mathcal{S}$ for $B^0 \to K^0_S\pi^0\gamma$ and $B^0 \to K^0_S\rho^0\gamma$ in Belle and BaBar.*

	Belle	BaBar
$B^0 \to K^0_S\pi^0\gamma$	$\mathcal{S} = -0.10 \pm 0.31 \pm 0.07$ [1]	$\mathcal{S} = -0.78 \pm 0.59 \pm 0.09$ [3]
	$\mathcal{A} = -0.20 \pm 0.20 \pm 0.06$ [1]	$\mathcal{A} = -0.36 \pm 0.33 \pm 0.04$ [3]
$B^0 \to K^0_S\rho^0\gamma$	$\mathcal{S} = -0.11 \pm 0.33^{+0.05}_{-0.09}$ [2]	$\mathcal{S} = 0.14 \pm 0.25 \pm 0.03$ [4]
	$\mathcal{A} = -0.05 \pm 0.18 \pm 0.06$ [2]	$\mathcal{A} = -0.39 \pm 0.2^{+0.03}_{-0.02}$ [4]

2.3 New method of constraining C_7 and C_7'

A new method published in 2019 in [52] proposes to study $B^0 \to K^0_S\pi^+\pi^-\gamma$ to get more information on the $\text{Im}(C_7'/C_7)$ and $\text{Re}(C_7'/C_7)$ plane described Sec. 1.4.

The CP asymmetry measurement is directly related to the photon polarization:

$$\mathcal{S} = -\frac{2\,\mathrm{Im}\left(\frac{q}{p}C_7 C_7'\right)}{|C_7|^2 + |C_7'|^2}\mathcal{D}, \tag{2.34}$$

with C_7 and C_7' the Wilson coefficients, q and p defined in Eq. 2.6.

To take into account the pollution of the non CP eigenstates in the final measurement of $\mathcal{S}$, the dilution factor $\mathcal{D}$ is defined as:

$$\mathcal{D} \equiv \frac{\mathcal{S}_{\text{all decays}}}{\mathcal{S}_{CP\text{ decay}}} \tag{2.35}$$

$$\equiv \frac{\mathcal{S}_{K_S^0\pi^+\pi^-\gamma}}{\mathcal{S}_{K_S^0\rho^0\gamma}} \quad \text{for } B^0 \to K_S^0\pi^+\pi^-\gamma \tag{2.36}$$

since $B^0 \to K_S^0\pi^+\pi^-\gamma$ can occur *via* different Kaonic resonances:

$$B^0 \to K_{\text{res}}\gamma \to (K_S^0\rho^0)\gamma \to K_S^0(\pi^+\pi^-)\gamma \tag{2.37}$$

$$\to (K^{*+}\pi^-)\gamma \to (K_S^0\pi^+)\pi^-\gamma \tag{2.38}$$

$$\to ((K\pi)_0^+\pi^-)\gamma \to (K_S^0\pi^+)\pi^-\gamma \tag{2.39}$$

where only the decay 2.37 is a CP eigenstate. To measure this factor, a detailed amplitude analysis is needed, where the different resonances are fitted in the Dalitz plot[1], on a large data sample, *e.g.* $B^+ \to K_{\text{res}}^+\gamma \to K^+\pi^+\pi^-\gamma$ assuming isospin symmetry. The current constraints on C_7 and C_7' using $B^0 \to K_S^0\pi^0\gamma$ for 711 fb^{-1} correspond to the yellow area shown Fig. 1.5, and the same area shape is predicted using $B^0 \to K_S^0\pi^+\pi^-\gamma$.

The new idea developed in [52] is to take advantage of the Dalitz structure of the $B^0 \to K_S^0\pi^+\pi^-\gamma$ decay to further constrain the Wilson coefficients. We can extract two new hadronic parameters, a and b, which depend on the different amplitudes of the different Kaonic contributions during the decay. They can also be measured with an amplitude analysis.

By splitting the Dalitz plane in two along the bisector $m_{12} - m_{23}$ as shown in Fig. 2.4, we can measure $\mathcal{S}$ in both regions to get two more parameters, noted $\mathcal{S}^I$ for the one above the line and $\mathcal{S}^{\bar{I}}$ for the one below.

The relation linking the Wilson coefficients and the four new parameter expresses as:

[1] In a three-body decay $P \to p_1p_2p_3$, the 2 dimension Dalitz plot is the name given to the m^2_{p1p2} vs m^2_{p2p3} invariant mass squared plot.

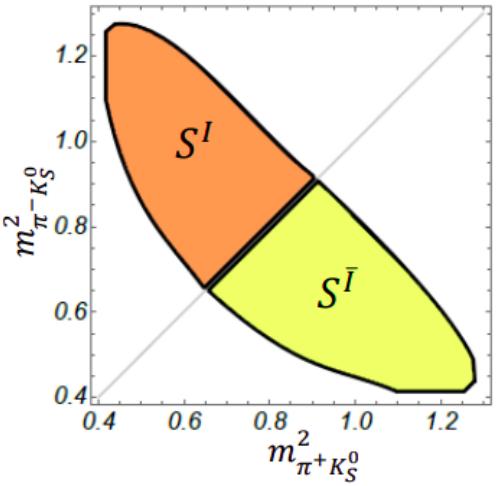

FIGURE 2.4: *Dalitz plot for $B^0 \to K^0_S\pi^+\pi^-\gamma$ split in two by the bisector $m_{12} - m_{23}$. $\mathcal{S}$ is measured in the two new regions to get $\mathcal{S}^I$ for the one above the line and $\mathcal{S}^{\bar{I}}$ for the one below.*

$$\frac{\mathrm{Re}\left(\frac{C_7'}{C_7^*}\right)}{1+\left|\frac{C_7'}{C_7^*}\right|^2} = \frac{1}{8}\left(\frac{\mathcal{S}^-}{\mathrm{b}^I}\cos 2\beta - \frac{\mathcal{S}^+}{\mathrm{a}^I}\sin 2\beta\right) \tag{2.40}$$

$$\frac{\mathrm{Im}\left(\frac{C_7'}{C_7^*}\right)}{1+\left|\frac{C_7'}{C_7^*}\right|^2} = \frac{1}{8}\left(\frac{\mathcal{S}^-}{\mathrm{b}^I}\sin 2\beta + \frac{\mathcal{S}^+}{\mathrm{a}^I}\cos 2\beta\right) \tag{2.41}$$

where β is the mixing angle and

$$\mathcal{S}^+ \equiv \mathcal{S}^I + \mathcal{S}^{\bar{I}}, \tag{2.42}$$

$$\mathcal{S}^- \equiv \mathcal{S}^I - \mathcal{S}^{\bar{I}}. \tag{2.43}$$

The new constraints on the $\mathrm{Re}(C_7'/C_7)$ and $\mathrm{Im}(C_7'/C_7)$ plane obtain with this new set of parameters are shown in Fig. 2.5. The area is now constrained on both real and imaginary part of C_7'/C_7, compared to the area covered in Fig. 1.5. This type of measurement can now be competitive with respect to the $B \to Kee$ done at LHCb. The first implementation for the measurement of $\mathcal{S}^+$ and $\mathcal{S}^-$ using the full Belle data was one of the main work in this thesis and is presented in the next chapter.

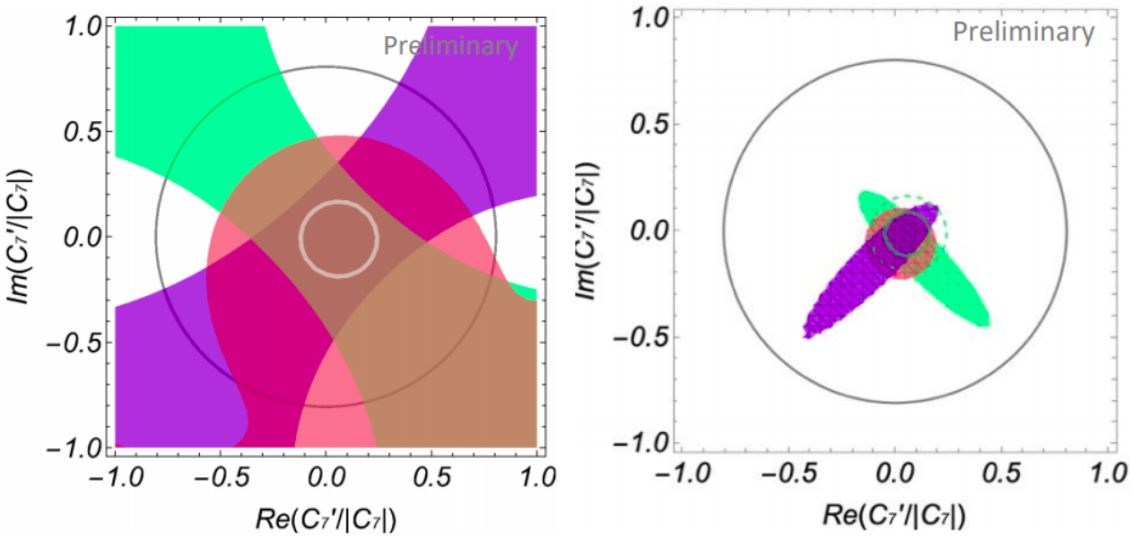

FIGURE 2.5: *Prospect for constraints on* $\mathrm{Re}(C_7'/C_7)$ *and* $\mathrm{Im}(C_7'/C_7)$*, at the three standard deviation level, obtained with three different arbitrary set of values for* $[\mathcal{S}^+, \mathcal{S}^-, \mathrm{a}, \mathrm{b}]$ *(in pink, violet and green), the gray circular contour is the constraint obtained from the branching fraction measurement of the inclusive* $B \to X_s\gamma$ *[46] and the white (left) or green (right) circles are the angular analysis measurement of* $B \to Kee$ *at low* q^2 *[35] (same as in Fig. 1.5). All the uncertainties are further scaled according to the increase of integrated luminosity: the left and right plots correspond to the current LHCb data / projection for* 22 fb^{-1} *for the* $B \to Kee$ *measurement and a luminosity of* 1.5 ab^{-1}/10 ab^{-1} *for Belle II. Updated from [52] thanks to one of the author.*

Chapter 3

The Belle II experiment

In order to measure with a high level of precision observables of the Standard Model for discovering New Physics, new B-factories benefit from improving technologies to reach higher and higher luminosities. NP particles can then be created in the final state, so-called *on-shell*, therefore a higher energy means reaching higher mass particles. This direct-search of NP, as described in Chapter 1, is the method used at LHC, and the future HL-LHC with increased accelerator radius and beam-energies. The other method is to measure the deviation of a prediction in the SM caused by NP. This way, the more data the more precise the measurement can be. Such indirect-search was and is the main technique used in Belle and Belle II to probe NP.

The Belle II experiment, located in Tsukuba in Japan, benefits from the upgraded Belle II detector and SuperKEKB accelerator, which has higher collision rates with a lower boost. Section 3.1 describes SuperKEKB, its nano-beam scheme and the beam background it creates. Each sub-detector of the Belle II detector is presented Section 3.2, with a description of the upgrade from Belle. Finally, the trigger system of Belle II is described Section 3.3.

3.1 SuperKEKB accelerator

3.1.1 Luminosity and Interaction point

The SuperKEKB is a 3 km circular energy-asymmetric electron-positron collider. It is the successor of KEKB, which collected data for about a decade, from 1999 to 2010. In order to achieve asymmetric energies, the accelerator is composed of two rings: the High Energy Ring (HER) for electrons, and the Low Energy Ring (LER) for positrons. A schematic view of the SuperKEKB colliders is shown in Figure 3.1.

Both beams interact in a single Interaction Region (IR). The event rate for a process with a cross section σ can be written:

$$\frac{dN}{dt} = \mathcal{L}.\sigma, \tag{3.1}$$

where the instantaneous luminosity $\mathcal{L}$ for two beams with a Gaussian profile of horizontal and vertical size σ_x and σ_y, is given by:

$$\mathcal{L} = \frac{N_{e^+} N_{e^-} f_c}{4\pi\sigma_x\sigma_y} \cdot R_{L'} \tag{3.2}$$

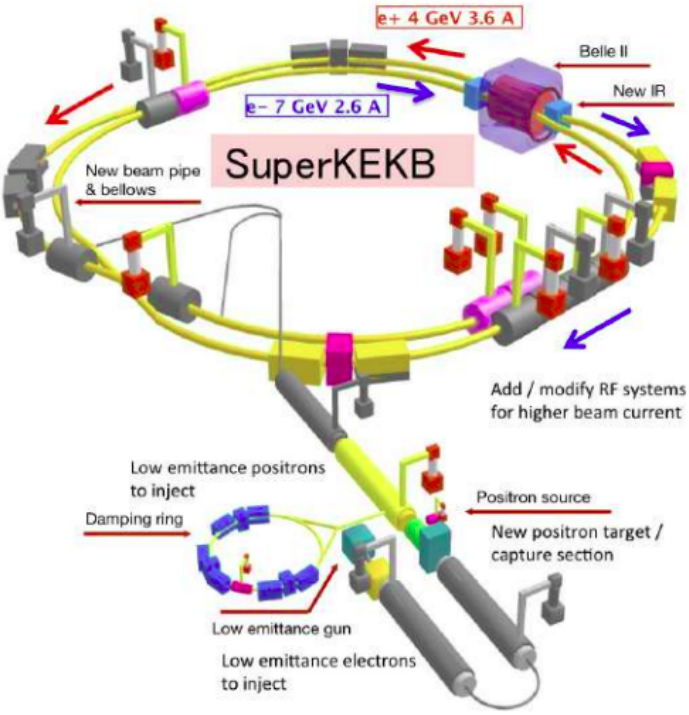

FIGURE 3.1: *Schematic representation of the SuperKEKB accelerator.*

where N_{e^+} and N_{e^-} are the number of particles in a positron and in an electron bunch, f_c is the bunch crossing frequency, and R_L is a reduction factor accounting for a geometrical effects associated with the finite crossing angle and the bunch length. To reach very high luminosity, SuperKEKB increases the beam currents $I_{e^+/e^-} \propto N_{e^-/e^+} f_c$ by a factor of two compared to KEKB and reduces the beam sizes:

$$\sigma_{x/y}(s) = \sqrt{\epsilon_{x/y}\beta_{x/y}(s)}, \tag{3.3}$$

where ϵ (in nm) is the emittance being a constant of motion along the orbit. $\beta_{x/y}$ (in mm) is the so-called "β function" that describes the displacement of a particle beam due to oscillations in the transverse plane xy with respect to its nominal trajectory in a given point (s) along the orbit (see Table. 3.1). To achieve a low emittance, a new damping ring was constructed for the positrons in order to reduce the emittance with radiation damping.

The beam energies have been chosen to operate at a center-of mass energy of 10.58 GeV, corresponding to the invariant mass of the $\Upsilon(4S)$ resonance, which is a bound state (bottomonium) of the beauty quark b and its antiparticle. It decays around 96% (at 90% CL) of the time into a $B\bar{B}$ charged or neutral pair, roughly in the same proportion. Since the mass of the B meson is 5.279 GeV/c^2 [23], they are produced almost at rest in the center of mass frame and because of their short lifetime ($\tau = 1.519 \times 10^{-12}$ s), they travel a very reduced distance in the laboratory. Therefore, the accelerator uses different energies for the electron and

positron beams so that the $\Upsilon(4S)$ acquires a boost in the laboratory frame and the decay products are pushed forward in the direction of the electron beam. The decay length is calculated as $\Delta z = \beta\gamma c\tau$, where $\beta = v/c$ is the ratio between the velocity of the particle v and the speed of light c, and γ is the Lorentz factor. The average distance Δl (Δz in boost direction) was about 190 μm at KEKB, and is now reduced to about 130 μm at SuperKEKB, due to different beam energies (see Tab. 3.1). With this lower boost, the vertex detector of Belle now in Belle II had to be upgraded.

3.1.2 Nano-beam scheme

The main modification in the nano-beam scheme approach regards the minimization of the longitudinal size of the beams overlap d (a schematic view of the beam intersection is reported in Fig. 3.2), which can be shown to be the lower bound for β_y^*. The formula for d is the following:

$$d \simeq \frac{\sigma_x^*}{\phi} \tag{3.4}$$

The numerator σ_x^* represents the horizontal beam size at the IR and ϕ is the half-crossing angle ($\sim$ 41.5 mrad). If the assumption of flat beams is verified ($\sigma_x^* = 9$ μm and $\sigma_y^* = 60$ μm for SuperKEKB), the luminosity formula at the IR can be written as:

$$\mathcal{L} = \frac{\gamma_\pm}{2er_e}\left(1 + \frac{\sigma_y^*}{\sigma_x^*}\right)\left(\frac{I_\pm \xi_{y\pm}}{\beta_{y\pm}^*}\right)\frac{R_L}{R_{\xi_y}} \tag{3.5}$$

where γ is the Lorentz factor, e the elementary electric charge and r_e the electron classical radius. R_L, R_{ξ_y} are the luminosity and vertical beam-beam reduction SuperKEKB accelerator factors, whose ratio can be approximately considered as unity. Since the ratio σ_y^*/σ_x^* is of the order of 10^{-3}, the first term into brackets can be considered to be equal to unity. Indices $(+,-)$ refer to electron and positron beam, respectively. From this formula, it can be noticed that there are three fundamental parameters for manipulating the luminosity and they are the beam current (I), the vertical beam-beam parameter (ξ_y) and the beta function $\left(\beta_y^*\right)$ at the IP. A comparison of the major upgraded parameters between KEKB and SuperKEKB is given in Table 3.1.

The $\Upsilon(4S)$ resonance is not the dominant particle produced when the e^+e^- beams collide. Table 3.2 resumes the production cross-sections for the main processes. The interaction is dominated by Bhabha scattering $e^+e^- \to e^+e^-$. The cross-sections of the other $q\bar{q}$ processes ($u\bar{u}$, $d\bar{d}$, $s\bar{s}$ and $c\bar{c}$) have the same order of magnitude than the $\Upsilon(4S)$ one, which leads to a high background contamination, called *continuum background*. To only select the interesting physics events, a trigger system has been designed (see Sec. 3.3).

TABLE 3.1: *Parameters of SuperKEKB and KEKB.*

		KEKB achieved	SuperKEKB target
Energy (LER/HER) [GeV]	E	3.5/8.0	4.0/7.0
Instantaneous luminosity $[10^{34}\ \mathrm{cm}^{-2}\ \mathrm{s}^{-1}]$	$\mathcal{L}$	2.11	80
Beam current (LER/HER) [A]	I	1.64/1.19	3.60/2.62
Vertical Beta Function at IP (LER/HER) [mm]	β_y^*	5.9/5.9	0.27/0.41
Beam crossing angle [mrad]	ϕ_c	22 ($\approx 1.3°$)	83 ($\approx 4.8°$)
Horizontal beam size at IP [μm]	σ_x^*	90	9
Vertical beam size at IP [nm]	σ_y^*	1900	60
Vertical beam-beam parameter (LER/HER)	ξ_y	0.129/0.090	0.090/0.088
Boost	$\beta\gamma$	0.425	0.284

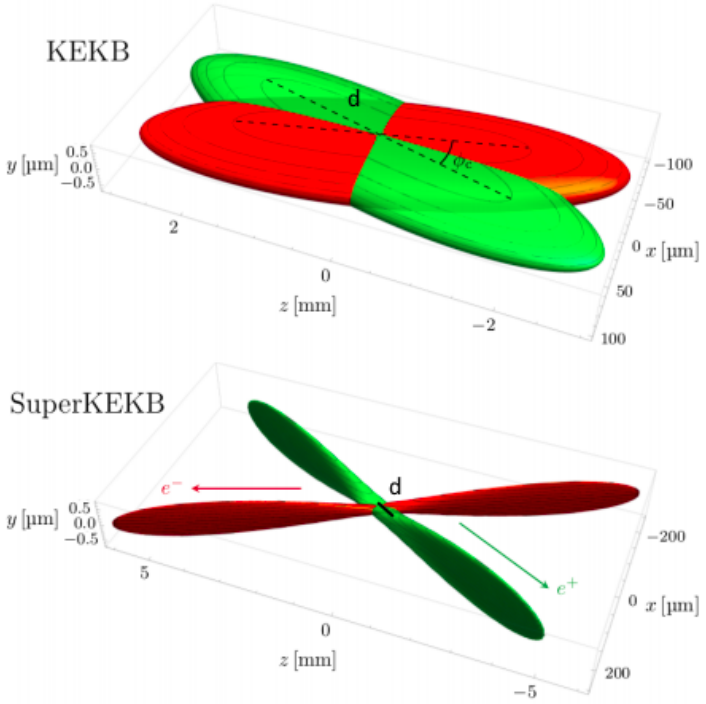

FIGURE 3.2: *Schematics of the beam size at the interaction point for KEKB (up) and SuperKEKB (down).*

3.1.3 Beam-induced background

There are five different types of background coming from the beams. Figure 3.3 schematizes each type:

- (a) Synchrotron radiation: When charged particles (e^+ or $e^\pm$ here) are moving in an electric field with a curved trajectory, they are accelerated by this field and then emit an electromagnetic radiation (X-rays). This is especially the case in the interaction region with its strong focusing magnets. The radiation level is proportional to the current of the beam.

TABLE 3.2: Production cross-sections for the main processes in e^+e^- collisions at a center-of-mass energy $\sqrt{s} = m(\Upsilon(4S)) = 10.58$ GeV

$e^+e^- \rightarrow$	Cross-section (nb)
$\Upsilon(4S)$	1.05
$u\bar{u}(\gamma)$	1.61
$d\bar{d}(\gamma)$	0.40
$s\bar{s}(\gamma)$	0.38
$c\bar{c}(\gamma)$	1.30
$\tau^+\tau^-$	0.94
$\mu^+\mu^-$	1.16
e^+e^-	~ 40

- (b) Beam-gas scattering: Residual gas in the beam-pipe can interact with the beam electrons, which will lose some of their momentum. Due to the magnetic field, those particles will then be deviated, hit the walls of the beam pipe or surrounding magnet which leads to production of secondary showers of particles that can leave traces in the detector. The scattering is proportional to the square of the current of the beam and the pressure in the beam-pipe.

- (c) Touschek scattering: The beam is so squeezed that Coulomb scattering can happen inside it. The electron's momentum and energy will change, which leads to the same effects as the Beam-gas scattering. This type of background dominates over the two above, and is expected to be 20× higher than in KEKB at nominal luminosity, due to the nano-beam scheme. The rate of Touschek scattering scales with the beam energy as E^{-3}, increasing the energy of LER in Belle II was a solution to reduce it.

- (d) Radiative-Bhabha process and (e) Electron-Positron pair production: Most of the time, as shown Tab. 3.2, the collision does not create an $\Upsilon(4S)$ but the two beams interact in two distinct way. An electromagnetic interaction happens between the electrons or positrons, deviating them and creating a photon or creating a low momentum e^+e^- pair. For both, the production rate is proportional to the luminosity, therefore it is expected to be 40× the background level at KEKB at nominal luminosity.

The first three types of beam-induced background are only coming from one beam and are characterized by very random processes, which can produce parasitic particles far from the center of the detector. When those particles reach the detector, they pass through a lot of materials, thus are hard to simulate. During the commissioning of SuperKEKB, two phases with specific instrumentation (BEAST) were used to control the beam background (see Sec. 3.1.4). They were acceptable to install the full detector. The Touschek scattering limits the beam lifetime (to about 10 minutes) and hence requires a continuous (top-up regime) machine. Collimators are placed along the

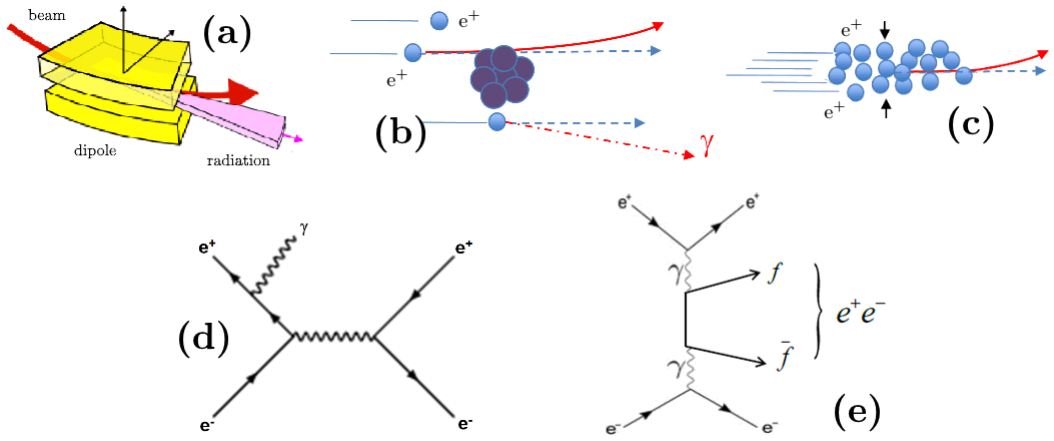

FIGURE 3.3: *Schematics of the five different type of beam-induced background. Above, the three types from only one beam: (a) Synchrotron radiation, (b) Beam-gas scattering and (c) Touschek scattering; under, the Feynman diagram of the two types from the two-beam interaction: (d) Radiative-Bhabha process and (e) Electron-Positron pair production.*

beam-pipe and a shield is in place to mitigate the beam-induced background. The collaboration constantly monitors the beam-induced background rates and perform simulations to reproduce the measured level. This strategy intends to understand the background and allow further mitigation measures as the luminosity increases. Indeed, the beam currents and sizes are modified to obtain higher luminosity only if the background in Belle II is not too large. Also these simulations allow to extrapolate the expected level of background at the nominal luminosity.

The rate of the last two beam-induced types of background, concerning the two-beam interaction, can be estimated by QED, as they are well defined by quantum effect and Feynman diagram. Those two types of background, which rates are proportional to the instantaneous luminosity, have been measured [53] but they will only be predominant at higher luminosity.

3.1.4 SuperKEKB running phases

The SuperKEKB had three phases during its commissioning. The Phase 1, started in 2016, consisted of testing the new nano-beam scheme and measuring the first backgrounds level, to make sure that the Belle II detector will be safe to install. No focusing magnets were in place, so no collisions were possible. Thanks to the BEAST II detectors system [54], the impact of the machine parameters on the beam-induced background and its level were measured.

The Phase 2, started in 2018, had the Belle II detector installed except the vertex detector (VXD) (described Sec. 3.2.2). Instead, a special detector which included BEAST II and one ladder of each subdetector of the VXD were installed close to the interaction point. The goals of this setup was to see if the VXD will not be damaged during the physics runs, and to further measure the beam-induced

backgrounds rates. The focusing magnets, along with the positron damping ring, were in place. The first collisions were detected on April 26th 2018. Until July 2018, Belle II was running and collected around 0.5 fb^{-1} of data, with SuperKEKB reaching $5 \times 10^{33}\,\mathrm{cm}^{-1}s^{-1}$ instantaneous luminosity. The data have been used in order to prove that the operation of the detector is understood.

The Phase 3, started in March 2019, is the beginning of the physics run with the full Belle II detector. Only the first layer of the VXD is not installed. The current luminosity projections are shown Figure 3.4. Since it started, around 400 fb^{-1} have been collected, with a maximum of $2.40 \times 10^{34}\,\mathrm{cm}^{-1}s^{-1}$ instantaneous luminosity. The luminosity increases with the capacity to limit the beam-induced background. A first shutdown starting in the summer of 2022 is planned to mainly install the whole VXD. The physics runs will resume in Fall 2023, until a long shutdown around 2026-2027 to improve the machine and upgrade the subdetectors. The SuperKEKB is designed to reach 50 ab^{-1} to do complete the Belle II physics program [55].

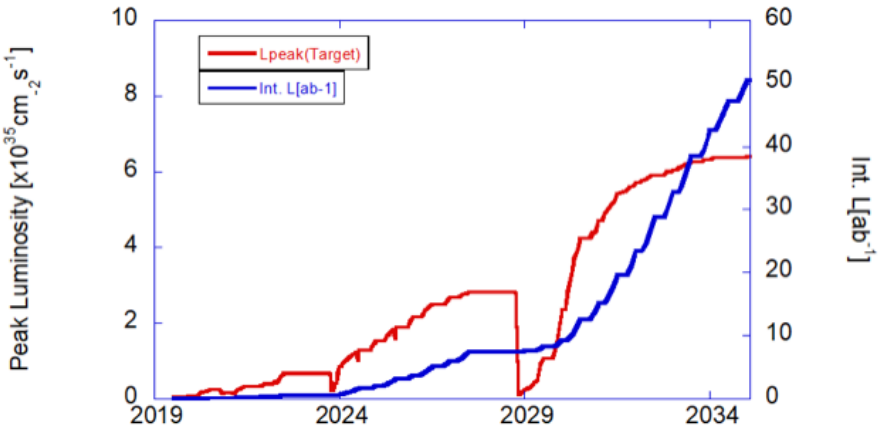

FIGURE 3.4: *Projection of the Belle II peak instantaneous (in red) and integrated (in blue) luminosity, as of June 2022.*

3.2 Belle II detector

The Belle II detector, successor of Belle, is located around one of the two interaction points of the SuperKEKB collider. A general overview of all the sub-detectors is provided in the next sections, along with the description of the improvement made from Belle.

3.2.1 Belle upgrade to Belle II

A typical hadronic event in Belle II is composed in average of 11 charged tracks, mostly low momentum pions as shown Fig. 3.5 and Tab. 3.3. The boost in Belle II is lower than in Belle (see Tab. 3.1). Moreover, with the new nano-beam scheme, beam-induced background levels rise significantly. The idea behind the Belle upgrade to Belle II is to improve or maintain the performance compared to Belle with this new experimental environment.

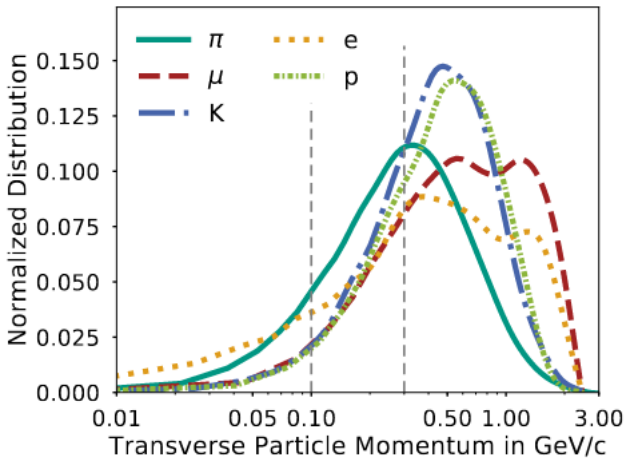

FIGURE 3.5: *Transverse momentum distribution of the most commonly produced particles in an e^+e^- collision in Belle II [56].*

TABLE 3.3: *Average fractions of the charged particles present in a $\Upsilon(4S)$ decay [56].*

Particle	Fraction
$\pi^{\pm}$	72.8%
$K^{\pm}$	14.9%
$e^{\pm}$	5.8%
$\mu^{\pm}$	4.7%
$p^{\pm}$	1.8%

A schematic view of the Belle II detector (top half) in comparison to the previous Belle detector (bottom half) is presented in Fig. 3.6. The modification on each sub-detector are the following:

- the radius of the closest detector to the beam-pipe, composed of silicon strip sensors, is moved closer to achieve the same proper time resolution, as the SuperKEKB boost is reduced *wrt.* KEKB. Also, the inner part of this detector is replaced by two new layers of silicon pixel detector based on DEPFET technology (PXD). This allows to have lighter material budget, thus less Coulomb scattering and better vertex resolution with the pixel technology;
- the silicon strip detector (SVD) is moved after the new pixel one, and layers are added to reach a larger radius than in Belle;
- to cope with higher beam-induced background, the silicon strips readout is changed to APV25 chips, which have a shorter shaping time;
- the central drift chamber (CDC) extends to a larger radius and has a larger number of small-cell layers;

- a new particle identification system (PID) in the endcaps and the barrel exploit the Cherenkov image detection method. They are more compact, highly performing and have a very fast readout electronics;
- the electronics of the electromagnetic calorimeter (ECL) is replaced with a fast wave-form sampling circuit;
- in the first two layers of the barrel part and in the endcaps of the muon and KL detector, Resistive Plate Chambers are replaced by scintillators readout with wavelength shifting fibers and Silicon Photomultiplier;
- the data acquisition system was changed to handle the higher event rate;
- the final focusing magnet (QCS) was the same for the incoming and outgoing beams. With SuperKEKB, two separate quadrupole magnets are in place. It reduces the number of over-bent electrons and positrons, which then hit the wall of the magnets and create electromagnetic showers.

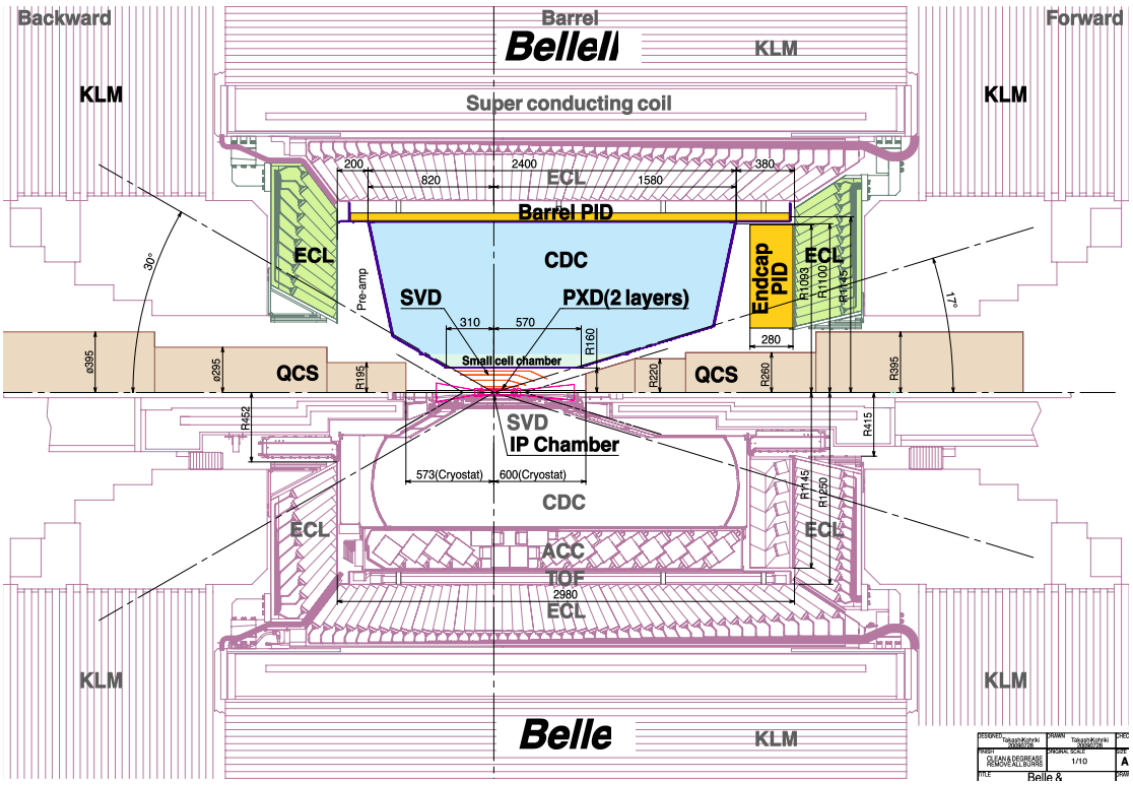

FIGURE 3.6: *Schematic view of the Belle II detector (top half) in comparison to the previous Belle detector (bottom half). From [57].*

In conclusion, the Belle II detector was designed to have the same performance under higher background levels thanks to better performing electronics. It even achieves better performances thanks to the new pixel detector and larger SVD radius to get better K_S^0 reconstruction, and the new PID system to ensure better pion/kaon separation.

3.2.2 Tracking Detectors

The vertex detector (VXD) is composed of a silicon pixel detector (PXD) and a silicon strip vertex detector (SVD). With the central drift chamber (CDC), they compose the tracking system in Belle II and reconstruct charged particle tracks and get the momentum of the particles thanks to the 1.5 T magnetic field.

Pixel Detector

The PXD is located closest to the 10 mm radius beam-pipe. The fine segmentation of the pixel technology allows it to be efficient in this high occupancy environment.

The two layers are at 14 mm and 22 mm (see Fig. 3.7) and are composed of 8 and 12 ladders respectively, with an acceptance going from 17° (in the forward region) to 155° (backward region). The asymmetry is due to the Lorentz boost in the laboratory frame, which favors events boosted in the forward direction. A ladder, with a detection area of 15 mm $\times$ 100 to 130 mm, is composed of two depleted field effect transistor pixels (DEPFET) [58]. The main goal of this subdetector is to have a good interaction point resolution in a high occupancy environment, hence the choice of this monolithic, light pixel technology. This technology consists of a p-channel MOSFET structure integrated onto a fully depleted silicon substrate. When a charged particle passes through the depleted area, it creates electron-hole pairs. Then the electrons drift toward the measuring electronics. A module is composed of 205 $\times$ 756 pixels, making around eight million pixels in total. Thanks to the internal amplification of the DEPFET, the sensitive volume of each pixel can be very thin (around 50 μm), which reduce greatly the multiple scattering effect. To further reduce this effect, the light materials of each PXD layer contribute to about 0.2% of a radiation length (X_0) and the readout electronics are located outside of the acceptance region.

The vertex resolution requirement of 20 μm leads to a pixel size of 50 $\times$ 50 μm^2 and 50 $\times$ 75 μm^2 for the inner and outer layers respectively. All the pixels can be read in a total time frame of 20 μs as the high occupancy environment requires a fast readout time.

Due to problems in the assembling of the PXD ladders in 2019, only the 8 ladders from the first layer and 2 ladders from the second layers are currently installed. This does not affect the track finding efficiency at low machine background, but will be useful when the instantaneous luminosity will increase. The full PXD installation is planned during the 2022 shutdown.

Silicon Vertex Detector

The SVD is composed of four double-sided silicon strip layers at 38, 80, 115 and 140 mm radius (see Fig. 3.7) with the same acceptance as the PXD. The position resolution of the strips, around 20 μm, is enough at these radii. In the forward region, the layer 4, 5 and 6 are slanted in the boost direction in order to reduce the material budget and the number of sensors needed. Its main roles are to link the track finding from the CDC to the VXD and to track low momentum charged

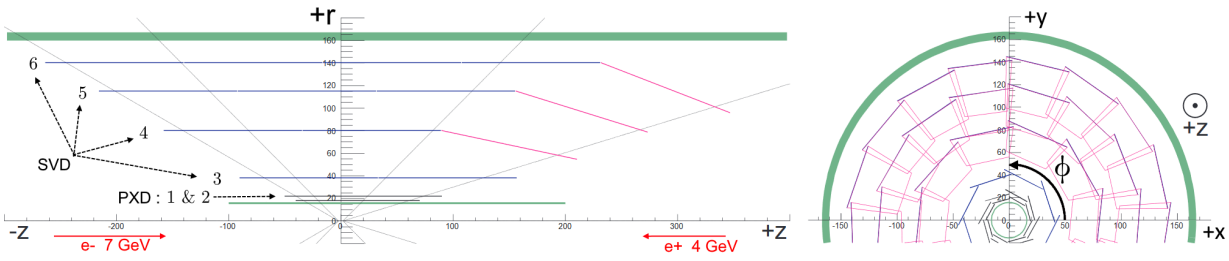

FIGURE 3.7: *Schematic of the VXD: PXD (layer 1 and 2) and SVD (layer 3, 4, 5 and 6), in the* $r = \sqrt{x^2 + y^2}$ *and* $\phi = \arctan(y/x)$ *coordinates. The z axis is oriented to match the boost direction. Adapted from [59].*

particles (from 50 MeV/c). It can also be used as a standalone tracking device, as particles with transverse impulsion below 600 MeV/c do not reach the CDC.

With its 172 sensors, the SVD is composed of around 220000 reading strips of 320 μm width. In order to limit the material budget, the read-out circuits are placed directly on top of the sensors (so-called origami scheme) in order to not exceed 0.6% X_0. The cooling system, a dual phase CO_2 system, was also designed to minimize the material budget. APV25 chips [60] are used (same one as in the CMS experiment at LHC), each having 128 channels and a resolution time of the order of 3 ns. The pitch of each sensors are shown Figure 3.8.

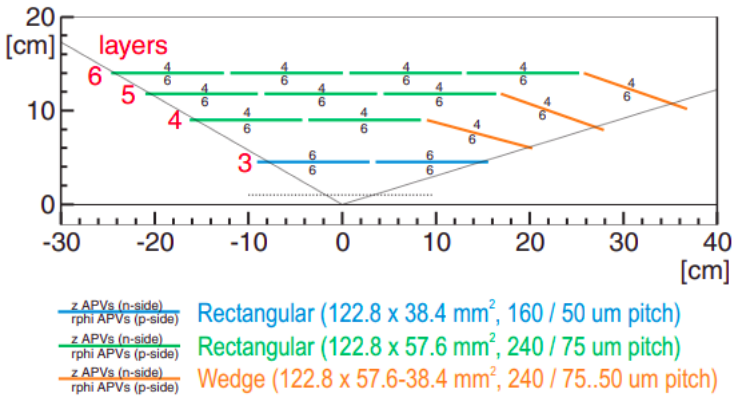

FIGURE 3.8: *Schematic configuration showing the three different sensor geometries and the number of APV25 read-out chips for each detector [59].*

The VXD determines the position in space where the charged particles interact inside the detector. The local information they leave by ionization in the layers are called *hits*. These traces are regroup to form *clusters* in order to reconstruct the track of the particle and its trajectory. By fitting all the tracks in the event, the decay vertices can be reconstructed. Having a good resolution is especially important in CP violation measurement. In fact, the B mesons are flying for a short period of time before they decay, of the order of $d_z = 150\,\mu$m. In a TDCPV measurement, we want to measure with great precision the time

Δt between the decay of the two B, and Δt is directly linked to d_z through their boost. The longitudinal resolution of the vertex detector is shown Figure 3.9. The incoming angle of the particle decreases the resolution by a factor $1/\sin^2(\theta)$, and the thickness crossed, which enhance multiple scattering effects, brings a factor $1/\sqrt{\sin(\theta)}$ It reaches around 10 μm for particles high momentum tracks.

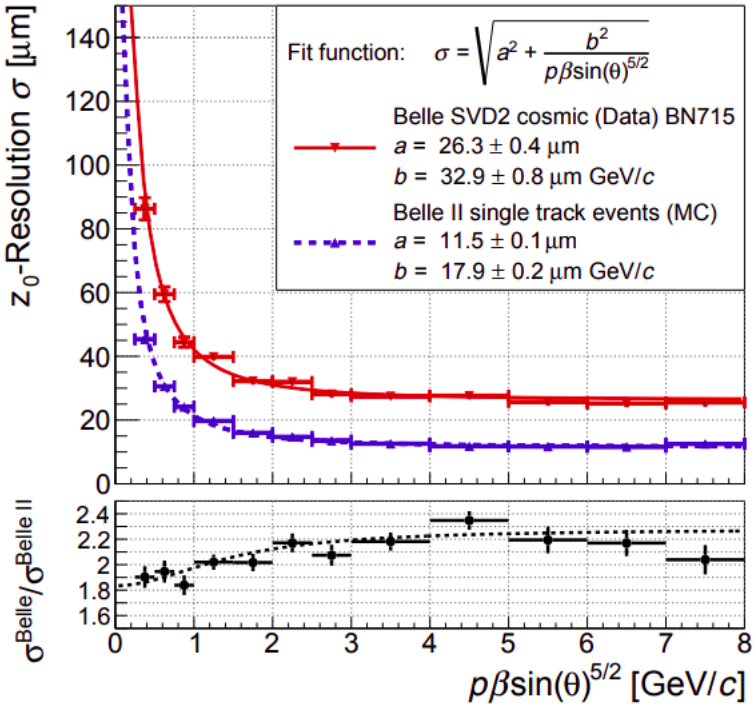

FIGURE 3.9: *Resolution of the longitudinal impact parameters (d_z for tracks) for Belle (in red) and Belle II (in blue) [55].*

Central Drift Chamber

The CDC is the main tracking device of the Belle II detector. The track finding algorithm begins by fitting the hits in the CDC and then extrapolates to the VXD. It is also one of the particle identification detector thanks to the measurements of the characteristic energy loss (dE/dx) of charged particle due to ionization in the gas volume. Finally, it provides reliable trigger signals for charged particle events. A large gas volume is best adapted to get an excellent momentum resolution, as it gives many points to extrapolate the charged particle tracks and it has a low material budget.

The CDC is filled with a 50% helium - 50% ethan gas mixture and is composed of 14336 sense wires of 30 μm diameter arranged in 56 layers. The length of the squared cells varies from 10 mm for the innermost super-layer to 18.2 mm for the outermost ones. There are two kinds of layers: the axial layer which is composed of wires aligned with the solenoidal magnetic field and the stereo layer in which wires are skewed with respect to the axial layer. As a result, it is possible to

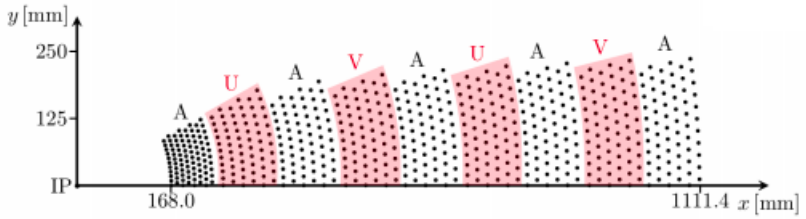

FIGURE 3.10: *Layer configuration of the CDC at Belle II. The letter A stays for axial superlayers and U and V for stereo superlayers with positive stereo angle and negative stereo angle respectively.*

reconstruct a 3D helix track and deduce its momentum. The configuration of the CDC is shown Fig. 3.10. Its spacial resolution is about 100 μm.

3.2.3 Particle Identification detectors

Two detectors are dedicated to particle identification: the Time-Of-Propagation counter (TOP) located in the barrel and the Aerogel Ring-Imaging Cherenkov counter (ARICH) in the forward endcap. Both sub-detectors exploit the Cherenkov effect to identify charged particles.

Time-Of-Propagation counter

The TOP is used for the particle identification in the barrel region by measuring the time of propagation of the Cherenkov photons internally reflected in the quartz bar.

It is composed of 16 modules surrounding the CDC. Each module is made of 2.5 m long quartz bar, a prism, a focusing mirror, and a Multi-Channel Plate Photo Multipliers (MCP-PMT). Figure 3.11 shows a schematic view of the TOP along with the path of the internal reflecting photons.

When they propagate and reflect in the quartz, the photons preserve their initial angular information. These photons have different path lengths inside the quartz bar, and thus different times of propagation until they arrive at the MCP-PMT, depending on their Cherenkov angle. By knowing the particle momentum and the location of the impact point on the quartz bar using the CDC, a likelihood for different mass assignment hypothesis can be calculated.

Aerogel Ring-Imaging Cherenkov counter

The particle identification in the forward endcap region is provided by the proximity a focusing aerogel ring imaging Cherenkov counter. When a charged particle passes through the aerogel radiator, Cherenkov photons are produced and developed into rings until they reach the hybrid avalanche photon detector (HAPD). A schematic view is shown Figure 3.12.

Increasing the thickness of the aerogel radiator provides a larger number of produced and detected photons, but the resolution of Cherenkov ring becomes worse due to the uncertainty of the emission point. To prevent this degradation,

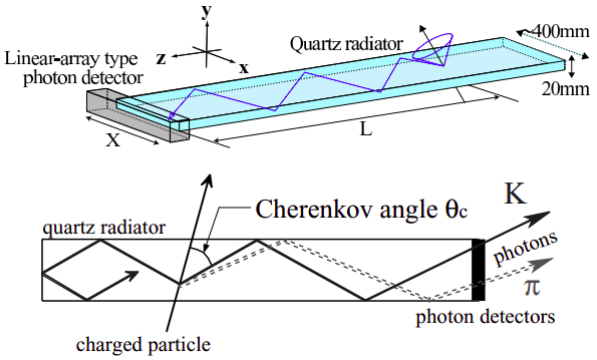

FIGURE 3.11: *Conceptual overview of TOP counter (top) and schematic side-view of TOP counter and internal reflecting Cherenkov photons (bottom). [59]*

two 2 cm layers of aerogel with different refractive indexes are used. The produced rings are overlapped on the detection surface, achieving a resolution on the Cherenkov angle and on the tracks of around 16 mrad and 4.5 mrad respectively.

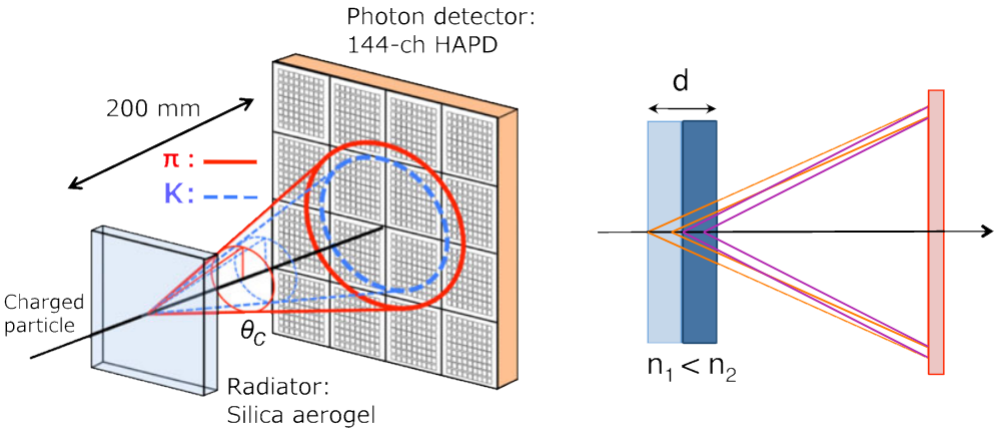

FIGURE 3.12: *On the left, the principle of the particle identification of the ARICH. On the right, the principle of operation of the proximity focusing with non-homogeneous aerogel radiator. [61]*

Figure shows the latest performances of the charged kaon and pion identification. In the plots presented, the binary likelihood ratio is defined as:

$$\mathcal{R}_{K/\pi} = \frac{\mathcal{L}_K}{\mathcal{L}_K + \mathcal{L}_\pi}. \tag{3.6}$$

The Particle Identification (PID) performance of the charged kaon and pion separation are calculated using $D^{*+} \rightarrow D^0 [K^- \pi^+] \pi^+$ decays and its charge conjugated mode. Slow pions can be used to tag the flavor of the D^0, which is finally used to identify the kaons and pions. With this information K/π PID efficiency and mis-ID rate can be studied with data. The kaon identification efficiency ϵ_K (ϵ_π) is defined as:

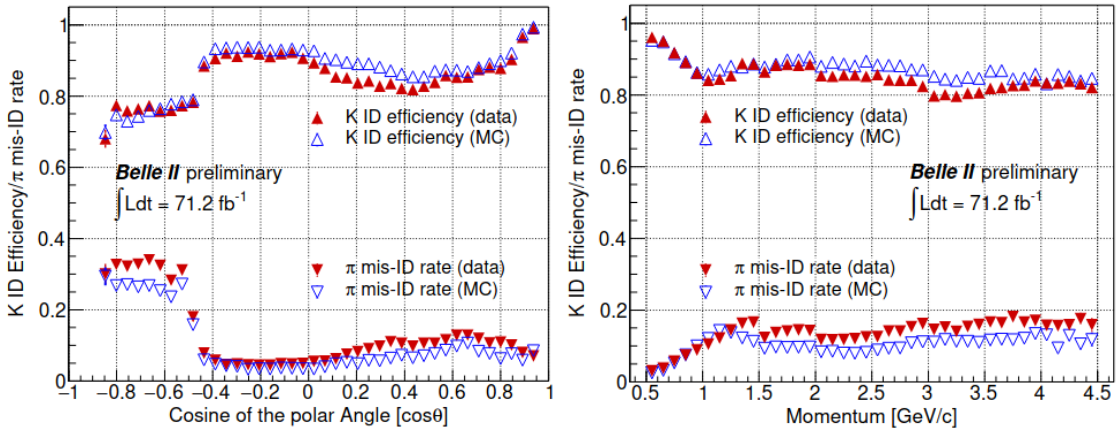

FIGURE 3.13: *Kaon efficiency and pion mis-ID rate for the PID criterion* $\mathcal{R}_{K/\pi} > 0.5$ *using the decay in the bins of polar angle (left) and laboratory frame momentum (right) of the tracks. The discontinuity around* $\cos(\theta) = -0.5$ *corresponds to the ARICH and TOP separation.*

$$\epsilon_K\,(\epsilon_\pi) = \frac{\text{number of kaon (pion) tracks identified as kaon (pion)}}{\text{number of kaon (pion) tracks}}, \tag{3.7}$$

while the pion mis-identification rate (miss-ID rate) is defined as:

$$\pi \text{ mis-ID rate } = \frac{\text{number of pion tracks identified as kaon}}{\text{number of pion tracks}}. \tag{3.8}$$

A good agreement is observed between Monte-Carlo and data, as shown Figure 3.13.

3.2.4 Electromagnetic Calorimeter

The ECL is used to measure the energy and angle of photons, while also generating a photon trigger. Since a third of the B meson decay products consist of neutral pions π^0 decaying into two photons, a high energy resolution and over a wide energy range, from 20 MeV to 4 GeV, is necessary. It can also measure the energy and the radiative energy deposition caused by bremsstrahlung for electrons and positrons to gain additional information for particles which were already reconstructed by the tracking system. Together with the KLM, it ensures K^0_L detection. Finally, the calorimeter provides luminosity measurements.

The ECL consists of a barrel section of 3 m length and 1.25 m inner radius, composed of 8736 CsI(Tl) crystals, each having a shape of a truncated pyramid with a length of 30 cm and a 6×6 cm^2 cross section, equivalent to 16.1 radiation length. The light generated by the initial interaction of photons with crystals are then detected by photo-diodes. The ECL geometry covers about 90% of the solid angle in the center-of-mass-system.

In the Belle experiment, the energy resolution observed with the same calorimeter was $\sigma_E / E = 4\%$ at 100 MeV, 1.6% at 8 GeV, and the angular resolution was 13 mrad (3 mrad) at low (high) energies; π^0 mass resolution was 4.5 MeV/c^2. In absence of backgrounds, a very similar performance would also be expected in Belle II. However, due to the relatively long decay time of scintillation in CSI(Tl) (1 μs), the overlapping of pulses are considerably increased at the SuperKEKB luminosity. The ECL electronics have been significantly upgraded compared to Belle. The scintillator photo-sensors are equipped with wave-form-sampling read-out electronics to mitigate the pile-up noise. From the 2019 spring run, the ECL performances were confirmed by $e^+e^- \to \mu^+\mu^-\gamma$, π^0 or $\eta \to \gamma\gamma$ reconstruction [62]. The mass resolution obtained are 5.4 MeV/c^2 and 12 MeV/c^2 for π^0 and η mesons respectively.

3.2.5 Neutral kaon and muon detector

The K_L^0 and muon detector (KLM) consists of alternating 4.7 cm thick iron plates and active material detectors located outside the superconducting solenoid. In the barrel region there are 15 layers, while in the forward endcap 14 layers are placed and in the backward endcap there are 12 layers.

The iron plates serve as the magnetic flux return and provides 3.9 interaction lengths (λ) in total, in addition to the 0.8 λ of the ECL, in which K_L^0 can shower hadronically. Most of barrel layers is equipped with glass electrode Resistive Plate Chambers (RPC). Because RPCs have the long dead time during the recovery of the electric field after a discharge, the inner two barrel layers and the endcaps employ layers of scintillator strips with silicon photomultiplier (SiPM). The angular coverage of the KLM is comparable to that of the ECL. Polyethylene sheets shield the KLM from multiple backgrounds including neutrons produced from beam background processes.

The muon identification also mostly relied on the KLM, since the mass of muon and that of pion are very close. Muons often traverse the ECL and the KLM with minimal energy loss, since they are most often low ionizing particle. When hadrons deposit energy into the ECL or collide with the iron plates, they interact strongly such that a characteristic shower of secondary hadrons is produced. The identification of muon and K_L^0 is performed by extrapolating tracks reconstructed by CDC and comparing with the actual measured hit in the KLM or if no secondary shower is present in the KLM. For high-purity samples of muons (*e.g.* $e^+e^- \to \mu^+\mu^-(\gamma)$) it has been estimated that the barrel section of the KLM provides a muon reconstruction efficiency of 89% for particle momenta above 1 GeV/c.

3.3 The trigger system at Belle II

As observed from Table 3.2, not all events in the collision are interesting to register. Most of the time, elastic Bhabha scattering processes happens, which are useful for calibration of detectors and luminosity measurement but not for the physics studies performed on the Belle II data. The background rates are

expected to be 20-30 times larger than at KEKB, resulting to a total expected event rate of about 15 kHz, while the limitation of the data acquisition system (DAQ) is 30 kHz. In addition, the detector readout of the sub-detectors needs to be synchronized to make sure that each device records the data of the same collision event. To preserve high efficiency on the processes of interest, an effective trigger system is essential. The trigger system is composed of the Level 1 trigger (L1), implemented in hardware, and the software implementation of the High Level Trigger (HLT).

3.3.1 L1 Trigger

The L1 trigger system gathers four sub-trigger systems which provide information from various subdetectors (see Fig. 3.14). The global decision logic (GDL) receives all sub-trigger information and makes the final trigger.

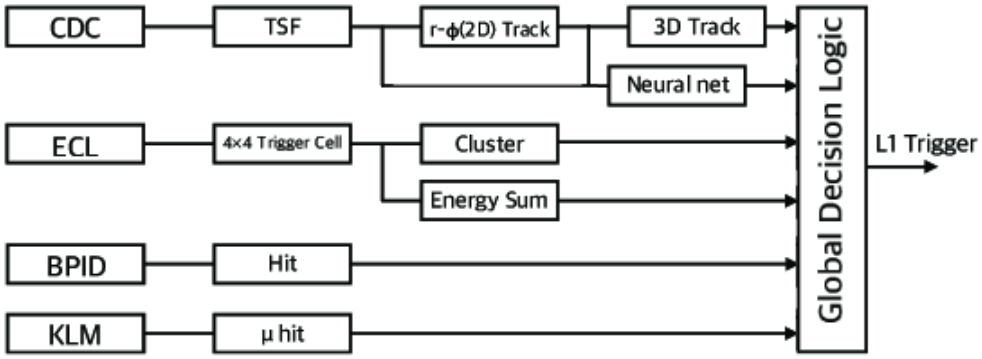

FIGURE 3.14: *Schematic view of L1 trigger system. CDC finds a track using small segment information and reconstructs 2D track. In addition, 3D track information which is constructed from 2D track and beam along information. Neural network based track finder is going to be implemented. ECL provides both each cluster information and total energy deposit in ECL to GRL. TOP and KLM information helps to identify particle type at the trigger level. Adapted from [63].*

The CDC trigger finds and characterizes the charged tracks detected in the drift chamber and gets their information such as momentum, position, and charge by using 2D and 3D tracking algorithms. Thanks to the 3D tracking, the vertex position along the beam direction is known and the beam-induced backgrounds in which tracks do not originate from the interaction point can be eliminated. In Belle, only the 2D tracking algorithm was used. The maximum 5 μs latency, limited by the buffer size in the SVD read-out ASICs, of the L1 trigger system requires a very fast z-vertex reconstruction, to be performed in no more than 1 μs, which is achieved thanks to neural networks techniques.

The calorimeter trigger provides energy deposit and cluster information to be sensitive to events with high electromagnetic energy deposition or multi-hadronic physics events with low energy clusters and/or minimum ionizing

particles. It can also detect Bhabha scattering and $e^+e^- \to \gamma\gamma$, which have large energy deposit and a back-to-back event topology.

The Barrel PID (BPID) trigger gives precise timing and hit topology information. With the KLM trigger, which gives μ track information, they are sensitive to radiative two-track processes and non radiative dimuon events such as $e^+e^- \to \mu^+\mu^-$.

The maximum instantaneous rate of the L1 trigger is 30 kHz. The goal is to keep the trigger rate as low as possible for physics targets, mainly hadronic events, while ensuring a high efficiency (larger than 99%). The L1 output is then forwarded to the High Level Trigger for further background rejection based on fully reconstructed objects.

3.3.2 High Level Trigger

In order to further reduce the information we need to store online or in tapes, the HLT reduces the event rates to the maximum design output event rate of 10kHz. It uses full detector information, except for the PXD, to calculate the event topology and the physical properties of the particles in order to reconstruct offline the events passing the L1 trigger. The events are categorized into each physics process ($B\bar{B}$, $e^+e^- \to \tau^+\tau^- \ldots$). Combining all the track information gathered, the PXD is then used to define so-called *Region-Of-Interest* (ROI). On each sensor plane where the particle is expected to go through, a rectangular-shaped region is defined around the extrapolated positions. Only the pixels belonging to the ROI are then readout and used to complete the event reconstruction.

Currently, all events are kept in the HLT since the total event rates are acceptable. With the increase in luminosity, the software for the HLT is being developed to tackle the rising event rate [53].

Chapter 4

Upgrade of the Belle II vertex detector

The current VXD was conceived to have good performances and resolutions until the nominal luminosity of 8×10^{35} cm^{-2}s^{-1} with certain beam optic parameters and without safety factors. Reaching such instantaneous luminosity requires new (yet unknown) beam parameters, hence it is not guaranteed that the current VXD is well suited. The beam-induced background is up to five times higher than the rate predicted by the MC simulations [54], but this single-beam induced background will be dominated by the beam-beam induced background at high luminosity. However, due to the extreme environment of the VXD, it is usual to change the pixel detector multiple times during the B-factory lifetime. The collaboration has to plan a possibility of an upgrade in the coming years to be able to replace the vertex detector. The upgrade plan to takes advantages of monolithic active pixel sensor (MAPS) to build a new VTX operating at low occupancy, which brings safety margin.

My work consisted of implementing new fully pixelated geometries in the Belle II software and optimizing them to reproduce the measured performances of existing MAPS. An overview of the requirements for the upgrade is given Section 4.1. The Belle II software framework and the tracking algorithm are described Section 4.2. Finally, the work on implementation of the upgraded vertex detector in the Belle II framework is presented Section 4.3.

4.1 Belle II upgraded tracking detector requirements

The Belle II physics program aims to accumulate a total integrated luminosity of 50 ab^{-1}, as shown in Fig. 3.4, which requires a peak instantaneous luminosity of SuperKEKB of 6×10^{35} cm^{-2}s^{-1}, approximately thirty times higher than the current one. To achieve that, a long shutdown is scheduled around 2026 to replace the focusing magnets (QCS), which provides an opportunity to upgrade the tracking detector of Belle II.

The estimate of the beam-induced background rate at peak luminosity has large uncertainty since various key factors are still unknown. The machine optics needed to reach peak luminosity is not yet confirmed. Also the interaction region geometry may change significantly. Finally the background generated during continuous injection has not yet been robustly assess, since it is necessary to know perfectly how the mismatch between the injection optics and the optics of the

main ring affects the dynamics of the beam, hence limiting severely the precision of any extrapolation to higher luminosity. Current simulations with the existing geometry predict that at peak luminosity, occupancies for the various layers of the present VXD get close to or exceed the limit beyond which tracking performance starts to degrade (see Fig. 4.1).

With uncertainty on the level of background noise which will accompany the rise in brightness beyond the target instantaneous luminosity, the collaboration wishes to have a solution which presents a more important security factor than that of the current VXD detector. It is therefore necessary to start developing upgrade solutions over the years to come. The construction decision will have to take place in 2024 at the latest in order to allow installation in 2026-27.

The goal is to get better performance and background handling compared to the current VXD. This can be achieved by designing a new VXD with significantly lower occupancy (which drives pattern recognition performance). The occupancy of the layers can be expressed as:

$$\text{occupancy} \propto \frac{t_{\text{integration}}}{\text{granularity}}. \tag{4.1}$$

By increasing the granularity of the detector and by reducing the integration time, the occupancy can be greatly reduced. A fast and fully pixelated detector, for instance based on the CMOS pixel sensor technology [64], to replace the strip and pixel detectors is then an appealing solution.

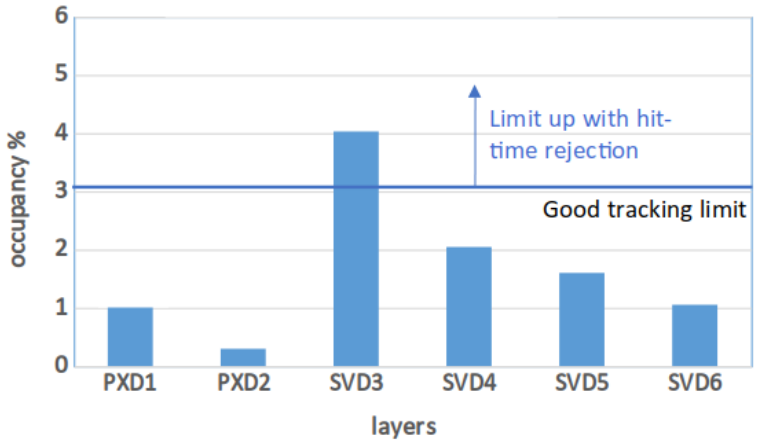

FIGURE 4.1: Occupancy extrapolation at peak luminosity for the current VXD, as of February 2021. The horizontal line indicates the occupancy limit beyond which tracking performance is expected to degrade. However, instrumental and algorithmic effort with the current instrument may push this limit a few percent higher.

As described in the Section 3.2.2, the VXD is currently composed of two pixelated layers using the DEPFET technology and four silicon strip layers. The upgrade consists of replacing the six layers with five to seven pixelated ones, to get the same technology on every layer. The new detector should at least match the acceptance of the current one. Its sensitive volume has then to extend from 1.4 cm to 14 cm in radii and from 12 cm to 72 cm in length.

Reducing the occupancy while providing similar or better tracking performance requires a pixel pitch in the range from 30 to 40 μm and an integration time under 100 ns. The material budget should target values of 0.1% of a radiation length (X_0) for the inner layers and 0.3% of X_0 for the outer layers, which implies a small sensor thickness, under 50 μm, and a low power dissipation, under 200 mW/cm^2, to limit the material required for cooling.

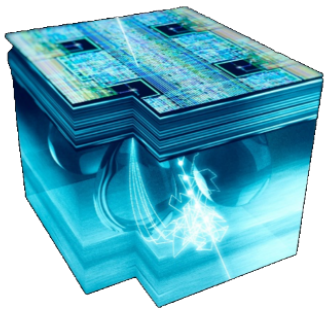

FIGURE 4.2: *Representation of the CMOS-MAPS sensors. From Magnus Mager at CERN.*

These requirements match the capabilities of CMOS-MAPS [64]. As depicted in Figure 4.2, MAPS combines the sensitive and signal treatment elements, allowing powerful detection performance with a low material budget. In particular, the TJ-MonoPix1 [65] and TJ-MonoPix2 [66] prototypes are good candidates as they feature a pixel size of 40×36 μm^2 and 33×33 μm^2, respectively, and an integration time of 25 ns which will be able to cope with the increased background at the maximal instantaneous luminosity. TJ-MonoPix1 already exists and was characterized at the DESY test beam facility. TJ-MonoPix2 is being fabricated and its pixel matrix is a good candidate for the sensor needed for the upgrade. It is then necessary to implement the response of these sensors in a so-called digitizer module in order to assess the overall performance of the upgraded VXD.

To predict the performance of the TJ-MonoPix2, we developed a test-beam simulation using the Belle II software. By modifying and simplifying the existing digitizer used for the DEPFET technology to fit the properties of a MonoPix chip, we tuned the digitizer with the data of a test-beam experiment done at DESY with TJ-MonoPix1 chips, a predecessor of the MonoPix-2.

4.2 Detailed simulation development

In order to estimate the performances of new geometries in Belle II, they have to be developed and validated in the Belle II analyses software. To do this, the geometries have to be correctly modeled and connected to the tracking scheme to be used in Belle II simulations. The major part of my work was to implement two geometries, called VTX, develop all the necessary infrastructure in the framework and link them to the existing tracking algorithms.

4.2.1 The Belle II software

The Belle II Analysis Software (basf2) [67] is the modular analysis framework in the Belle II collaboration and is used to do Monte-Carlo (MC) simulation, any physics analysis or data processing. It is based on C++ and Python and used ROOT [68] for analysis or graphics, EvtGen [69] and PYTHIA [70] to simulate physics events and Geant4 [71] to modelize the Belle II geometry and the interaction of particles with matter. basf2 is divided into packages, each containing the code relevant for a subdetector or data processing code (*e.g.* tracking package containing the whole tracking scheme).

A typical way to process data is building in a Python script a chain of smaller processing blocks, called modules, performing very specific tasks. A module can be responsible for loading the geometry with Geant4, generate particles with EvtGen, apply the tracking algorithms to the data or even save the analysis information in a root file.

Monte-Carlo method is used to do simulations. The particles are generated at the center of collision following a defined distribution, they decays into modes that can be modified by the user and their decay products then interacts with the detector. MC simulations lets us access all the data of generated particles (quark composition, impulsion, energy ...), which we can then compare with the reconstructed data from the detector. All the information is sorted in so-called *datastore*, in particular the hits from the various subdetectors before the reconstruction and the information of the reconstructed tracks from the tracking algorithms. Finally, the information from the datastore can be saved in root files as *NTuples*.

4.2.2 Track Finding in Belle II

The whole tracking scheme is explained in detail in [56]. Here, only a brief description is given in order to understand how the new VTX detector was implemented and connected to the existing tracking system.

The tracking scheme

In order to reconstruct the tracks, the track finding algorithm needs information from the CDC, the SVD and PXD.

For the CDC, by combining the time between the event's trigger signal and the arrival of the drift electrons at the sense wire, the relative distance between the sense wire and a passing particle can be computed. The amplitude information from the front-end read-out is also used to separate signal and background hits.

In the SVD, clusters are formed by combining adjacent strips with a signal-over noise ratio above three. Then, 3D space points are deduced by combining all clusters on one side of a sensor with clusters on the other side. Along with the space points, cluster charge and time information is used in the tracking algorithms.

For the PXD, neighboring pixels in the ROIs (see Sec. 3.2.2) with a charge above a threshold are combined into clusters and their positions and charges are used in the tracking algorithms.

With all the inputs from the tracking detectors, the track finding algorithm will try to return all fitted tracks. Figure 4.3 shows all necessary steps to reconstruct the tracks.

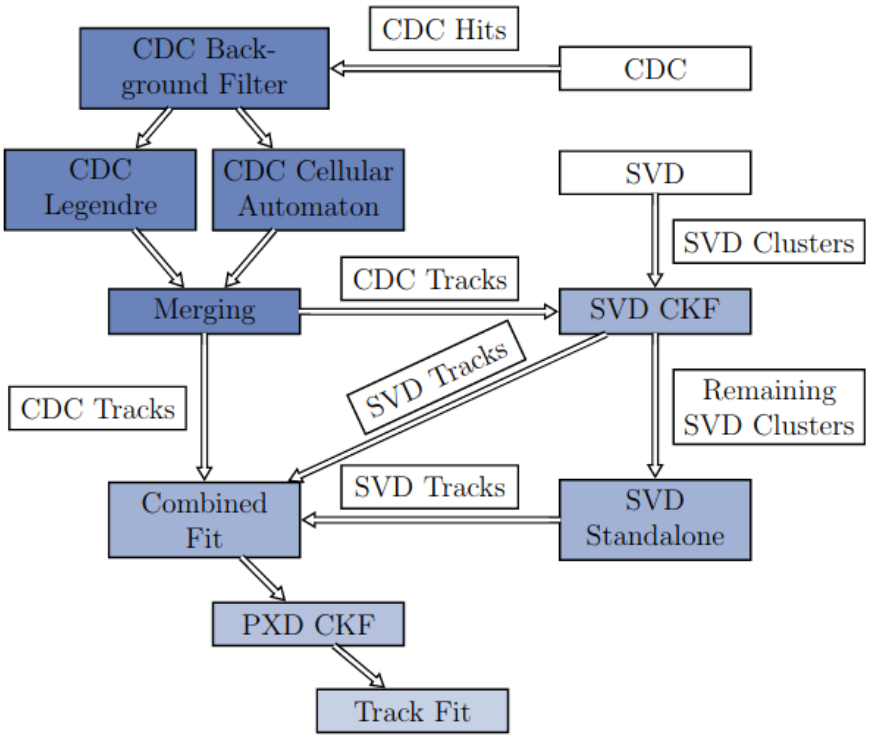

FIGURE 4.3: *Overview of the steps performed for track reconstruction at Belle II. From [56].*

The tracking begins with the CDC, where all the measured signals are filtered and reconstructed globally by a track finding algorithm based on the Legendre transformation [72] and locally by using a cellular automaton, which complement the global fit by finding the short tracks and tracks displaced from the interaction point. The results of the two method are merged and fitted together using a deterministic annealing filter (DAF) [73], based on boosted decision trees (FastBDT) [74], to finally get the tracks from the CDC. Then, the tracks are extended to the SVD by using its clusters with a combinatorial Kalman filter (CKF) [75].

Particles with a transverse momentum of less than 100 MeV/c do not deposit enough hits in the CDC to be fitted, so a dedicated standalone algorithm was developed to reconstruct those tracks within the SVD. It uses a cellular automaton and an advanced filter concept called Sector Map. For my service task[1], I optimized the Sector Map to reduce its memory usage while having the same performances. The work is detailed Appendix A. The tracks coming from the CDC, the SVD CKF and the SVD standalone algorithms are then merged together with an other DAF.

[1] In order to be author, a 6-months work on a general tool useful to the whole collaboration.

The PXD hits are not used for the track finding, but are finally added to the merged tracks with a second CKF. The final fit, on all the found tracks, is done by a third DAF. The results of the pattern recognition is stored to be used in physics analyses.

Performance evaluation

Three types of efficiency will be interesting to discuss the performances of the new geometries.

In a Monte-Carlo simulation, the generated parameters for each particle along with the Geant4 simulation of the interaction between them and the detector can be accessed. Using these information, an ideal track finder, called MC track finder, was implemented. Its performance cannot be surpassed by definition as it is limited only by the acceptance, efficiency and resolution of the detector.

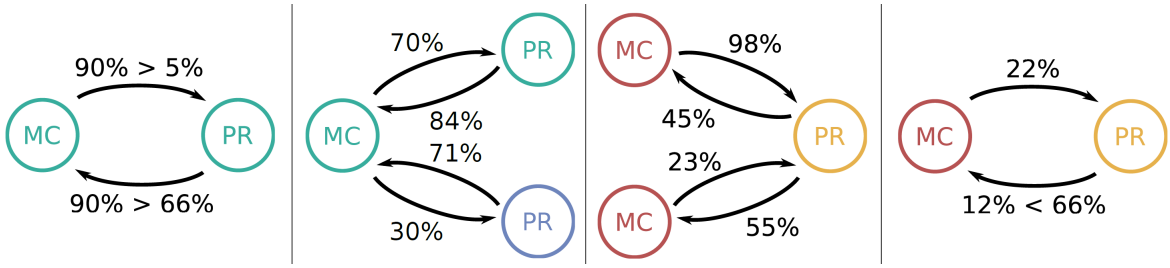

FIGURE 4.4: *Different classification of MC (left side) and PR (right side) tracks based on the efficiency and purity calculated between them as indicated by the arrows (top arrow is efficiency, bottom is purity) [76].*

The performance of the Track Finder described in the last section should behave as closely as possible to the MC one. The number of shared hits are used to compare the tracks from the pattern recognition (PR-track) with the MC track finder (MC-track). Figure 4.4 summarizes the four cases:

- A PR-track share more than 66%[2] of its hits with a MC-track with a minimum efficiency of 5%: the two tracks are then linked together and the PR-track is called *matched*. The tracking algorithm correctly reconstructed the track.
- Two PR-tracks share more than 66% of their hits with a MC-track: The one with the highest efficiency of linked to the MC-track and the other one is called *clone*. The two reconstructed tracks correspond to the same MC-track, but one took some hits from other tracks and/or background.
- A PR-track share hits with a lot of MC-track without going over 66% purity: it is marked as *fake*. This reconstructed track should not exist.

With this classification, three figure of merits car be defined:

[2] 66% purity was chosen as a minimum to get a good resolution on the impulsion of the particle.

- Finding efficiency : $\frac{\text{Number of matched tracks}}{\text{Number of generated tracks}}$
- Clone rate : $\frac{\text{Number of clone tracks}}{\text{Number of matched tracks}}$
- Fake rate : $\frac{\text{Number of fake tracks}}{\text{Number of reconstructed tracks}}$

Ideally, the reconstruction efficiency should equal to 1 and the clone and fake rate equal to 0. In reality, the low momentum particles don't leave enough hits in the tracking detector and are affected by multiple scattering effect. They do not go through a lot of tracking layers, thus are difficult to reconstruct. Also, they can curl in the detector due to the magnetic field, which increase the clone rate. The fake tracks are mainly due to background hits.

In the next section, along with momentum resolution, these figure of merits will be used to evaluate the Track Finder for a specific geometry compared to the current Belle II setup.

4.3 VTX implementation in basf2

I was responsible for the implementation in basf2 of a new pixelated detector, replacing the current VXD, in order to estimate how much Belle II can benefit from this technology. The description of the new VTX tracking detectors had to be developed then connected to Geant4 and the existing tracking algorithm. Finally, it had to be tuned to mimic TJ-MonoPix1 and TJ-MonoPix2 performances.

4.3.1 New five and seven layer geometries

Two new geometries were fully implemented: one 5-layer and one 7-layer pixelated detector (see Fig. 4.5). A 5-layer with disks in the forward region, to track lower momentum particles, was also developed but later dropped as the mechanical part is too difficult to realize and install in Belle II in the short span of time before the long shutdown.

The properties of each detector are shown in Tables 4.1 and 4.2 for the five and seven layer respectively. The sensor properties are shown in Table 4.3. The sensor type A and A′ are respectively placed in the inner and outer ladders of the detectors. Some simplifications were made in order to reduce the computing power needed in the simulation: the length of the sensors is equal to the length of one ladder (12 cm), in the final design the sensors will have a length of about 2 cm; the thickness of the sensors is a bit larger to simulate some support materials that are not modelized in basf2; for the outer layer, one sensor corresponds to two sensor placed next to each other. In the final design, the five and seven layer detectors are expected to have around 2000 and 3000 CMOS sensors respectively.

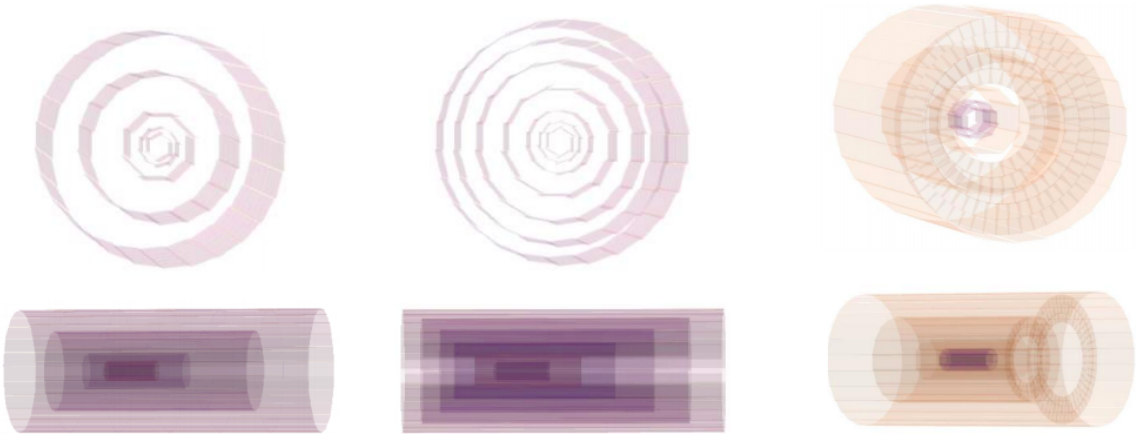

FIGURE 4.5: *Front and side view of the 5-layer, 7-layer and 5-layer with disks geometries (from left to right) implemented in basf2. Only the first two are connected to the tracking.*

TABLE 4.1: *Five layer VTX detector properties. See Tab. 4.3 for the sensor characteristics.*

Layers	1	2	3	4	5
Radius (cm)	1.4	2.2	3.9	8.9	14.0
# ladders	6	1 0	8	1 8	2 6
Sensor type	A	A	A′	A′	A′
# Sensor rows (z direction)	1	1	2	4	6

TABLE 4.2: *Seven layer VTX detector properties. See Tab. 4.3 for the sensor characteristics.*

Layers	1	2	3	4	5	6	7
Radius (cm)	1.4	2.2	3.5	6.0	9.0	11.5	13.5
# ladders	6	1 0	1 4	1 2	1 8	2 2	2 6
Sensor type	A	A	A	A′	A′	A′	A′
# Sensor rows (z direction)	1	1	2	3	4	5	6

The VTX package is following the PXD package structure, designed and described in [77]. Figure 4.6 summarizes each developed steps in the VTX package in order to be used by basf2.

First, the geometry is configured in files containing definition parameters under the general purpose `xml` format. A lot of parameters car be tuned. One file is responsible for the number of layers, ladders and the angle between the ladders; an other for the number of sensors per ladder, the radius and shift (*wrt.* the z axis origin) of each layer, the material of each component and all the parameters for the pixels (number of, size, integration time . . .); and a final file to characterize the mounting support. With all this information, Geant4 is now able to simulate matter-matter interaction in the new geometries.

Specific containers for all the VTX related objects had to be developed to then be used in the simulation and reconstruction algorithm. These two works the

TABLE 4.3: *Sensor properties for the two new geometries.*

	Sensor A	Sensor A′
Width (cm)	2.0	2.0×2
Thickness (mm)	0.4	1.0
Length (cm)	12	
Radiation length X_0	0.1%	0.3%

	Sensitive area	
Width (cm)	1.8	1.75×2
Height (μm)	40	100
Length (cm)	12	
# Pixels U	512	
# Pixels V	3584	
Charge Threshold (ENC)	150	
Integration time (ns)	100	

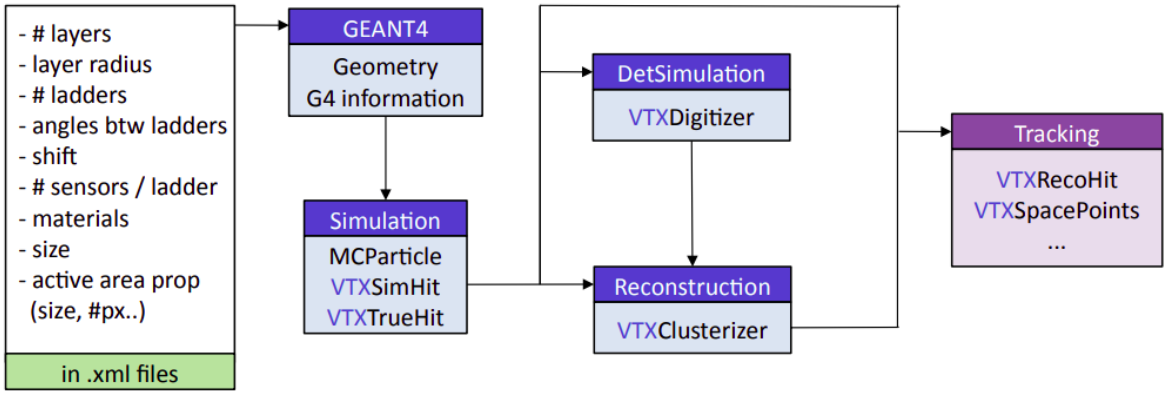

FIGURE 4.6: *Overview of all the necessary steps for the VTX package to be used in the Belle II software.*

same way as in the PXD: the particles going through the layer of the VTX create hits (`VTXSimHits` and `VTXTrueHits`), which are then group together into clusters and digitized into signal (detailed description in the next section). All these information will then be used by the tracking.

The tracking algorithms have also been modify to be used with VTX. Figure 4.7 shows all the steps with the VTX detector implemented, similarly as Fig. 4.3 for the standard geometry. The CDC standalone tracking was not changed. To extend the CDC tracks to the VTX ones, a new CKF had to be trained. The remaining hits from the VTX go to a standalone track finding, similar to the SVD one. The 3D space points are now easier to get as neighboring hits in the pixels are combined. Finally, to merge the VTX and CDC hits, another CKF had to be trained.

Two setup can be used to compare the performance of the VTX: the full tracking scheme, from CDC to VTX, and VTX standalone track finding, where no

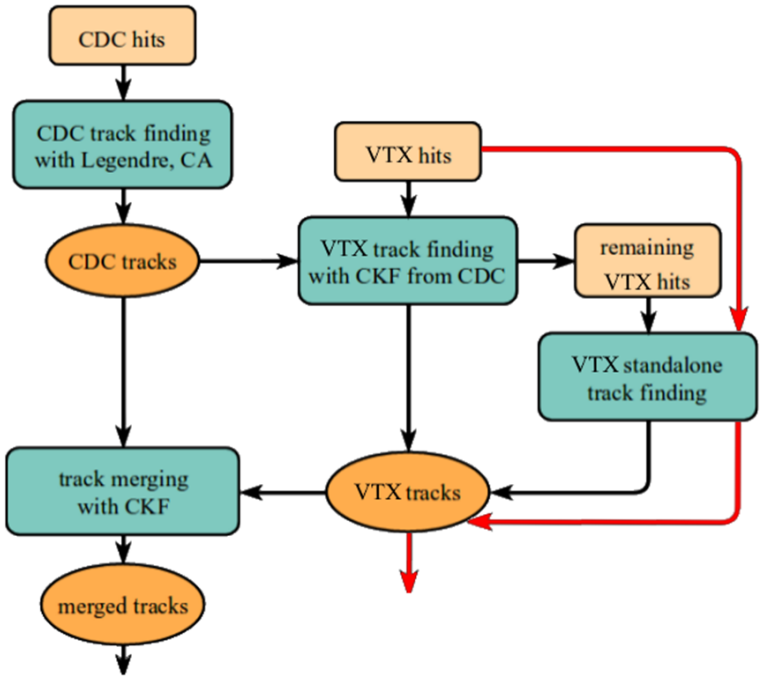

FIGURE 4.7: *Overview of the steps performed for track reconstruction with the new pixelated detector. Black arrows show the full tracking path, while red arrows show the VTX standalone path.*

information from the CDC are used (black and red arrows in Fig. 4.7 respectively).

4.3.2 Tuning of the digitizer

When a charged particle traverses the sensor, secondary charges are generated by ionization along the particle trajectory within the active thickness of the sensor. Then, their movement driven by the sensor biasing generate voltage pulses over a number of pixels, which the electronic front-end microcircuits concert into a digital quantity, representing the pixel signals. The so-called digitizer module simulates the whole process according to a number of parameters. I was responsible to tune those parameters to match a test-beam experiment made at DESY with TJ-MonoPix1 chips.

Description of the digitizer

Figure 4.8 shows a schematics on how the digitizer works. First, it checks if the particle hit in inside the integration time window $T_{int} = 100$ ns. Then, it emulates the collection of the drifting charges over the pixels in the vicinity of the particle hit. It splits the path of the particle in the VTX active thickness e into segments and drift the charges from the center of each segments. Assuming the active volume is fully depleted, charges are randomly spread according to a Gaussian, with a width that depends linearly with the charge depth creation. Then, the digital value of each pixel is decided taking into account an equivalent

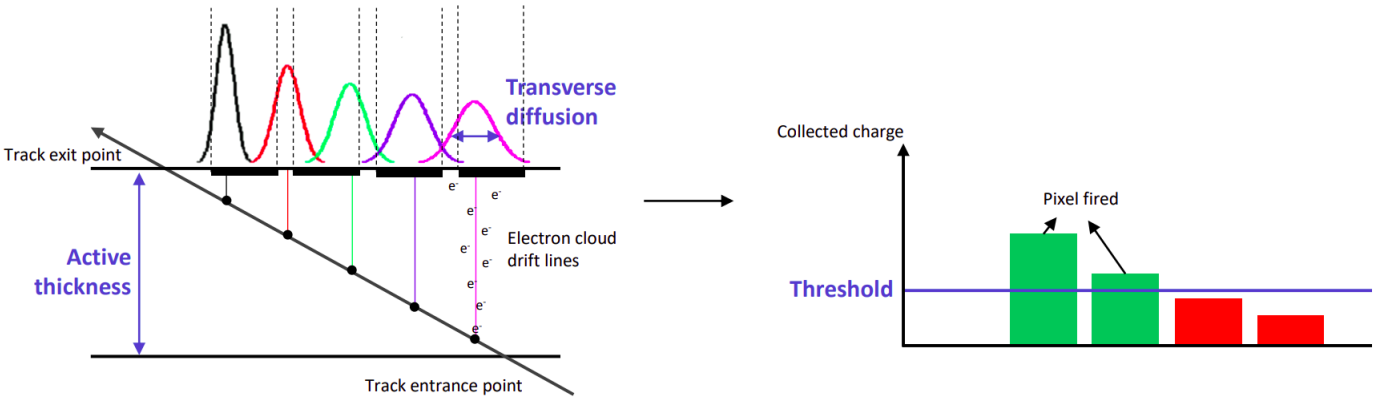

FIGURE 4.8: *Schematization of the VTX digitizer operation. The blue parameters have been tuned to match TJ-MonoPix1 and 2 performances.*

noise charge and a charge-to-digital conversion factor. Finally, if the number of electrons is above a detection threshold, the pixel is fired.

The digitized signal follows [78] and is given by:

$$S = \frac{Q(l) - t}{c} = \frac{e(100.6 + 35.35\ln(e))}{3.6c} - \frac{t}{c}, \tag{4.2}$$

with S the width of the cluster charge, $e = 30\ \mu$m and the threshold $t = 500\, e^-$. After a Landau fit of the cluster charge from the test-beam measurement, the conversion factor is extracted. We get $c = 120\ e^-/\text{ADU}$ (analogue to digital converter).

For a normal incidence (see Fig. 4.9 for the test-beam setup), the transverse diffusion coefficient D is obtained from:

$$\sigma_{\text{diffusion}} = \sqrt{\frac{De}{2}}, \tag{4.3}$$

with $\sigma_{\text{diffusion}}$ size of the cluster, obtained with the cluster size plot Figure 4.10. We finally get $D = 0.6\ \mu$m.

Along with the charge threshold, those parameters have then been tweaked to match the observed distributions as TJ-MonoPix1 test-beam data.

Results

In the end, the tuned digitizer reproduces fairly well the test-beam data. It leads to a good agreement with the cluster size and cluster charge distribution for every test-beam angle (see Fig. 4.11). The small discrepancies are not an issue to estimate the performance of TJ-MonoPix1. A more precise tuning will be done when TJ-MonoPix2 will be implemeted.

Also the spatial resolution is reproduced within 20%, as shown Figure 4.12.

These results show that the digitizer, in particular the tuned value of the transverse diffusion, is validated for the TJ-MonoPix1, can be used to predict

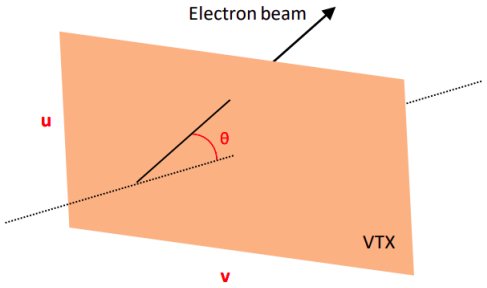

FIGURE 4.9: *Schematics of the test-beam setup. The u and v coordinates corresponds to the horizontal and longitudinal axis of the sensor respectively.*

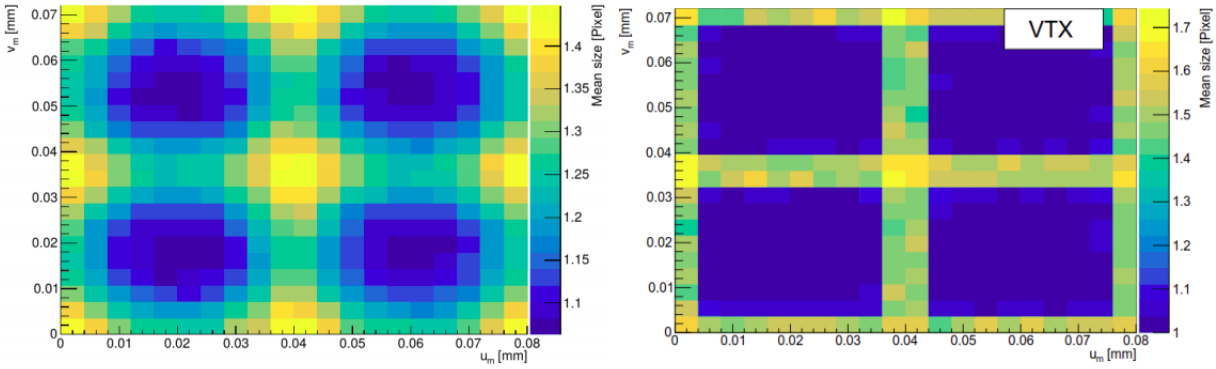

FIGURE 4.10: *Cluster size map in u and v coordinates for the test-beam experiment (left) and for the VTX simulation in basf2 after tuning (right).*

the TJ-MonoPix2 behavior in the full simulation of the upgraded geometries presented in the next section.

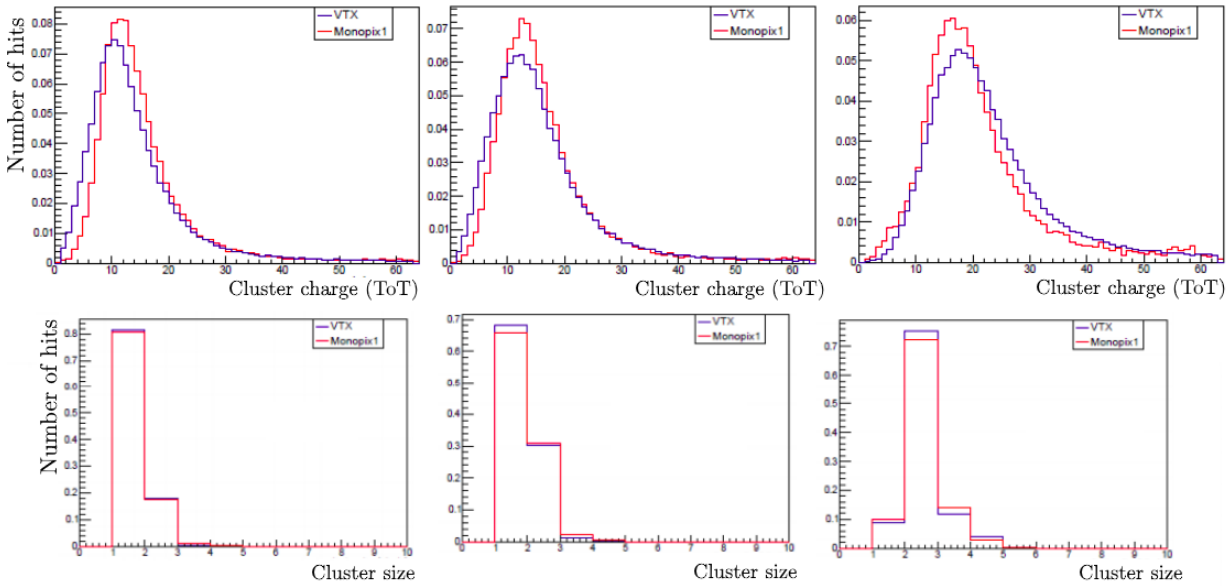

FIGURE 4.11: *Cluster charge (up) and size (down) distribution for the test-beam (in red) and the tuned digitizer in the Belle II software (in blue) with, from left to right: a normal incidence beam to the detector, an angle of 30° and 60° (see Fig. 4.9).*

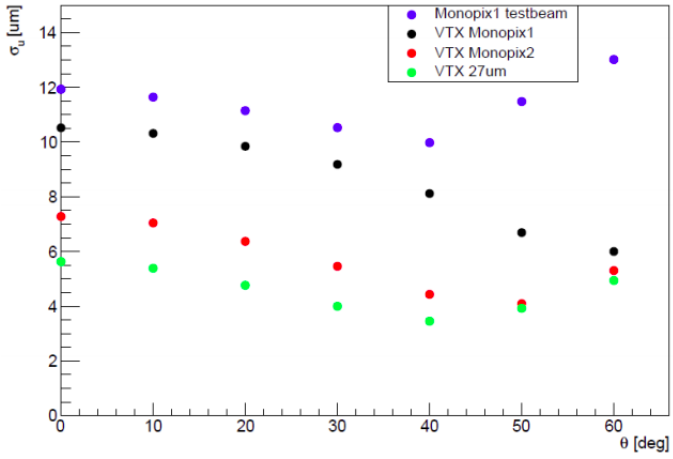

FIGURE 4.12: *Pointing resolution on the u side as a function of the beam angle (see Fig. 4.9) for the test-beam data (in blue), the tuned digitizer to reproduce the TJ-MonoPix1 (in black) and the predicted TJ-MonoPix2 (in red) and a theoretical sensor with a 27 μm pixel size (in green).*

4.4 Performance results

To see how the upgraded geometries perform, several resolution parameters have been measured and compared with the current Belle II setup.

When using a muon particle gun with nominal background predictions from 2019 overlaid, the CMOS five and seven layers have better transverse momentum resolution compared to the current Belle II VXD, in particular at low p_t, as shown in Fig. 4.13. Also, the absolute full-track finding efficiency (CDC + VTX) is

improved. When averaged on a flat momentum distribution is 2-3% better, as shown in Fig. 4.14. However, an increase in the fake rate is observed. This can be reduced with a better optimization of the tracking algorithm with the VTX detector, *e.g.* with a better tuning of the sector map (see Appendix A).

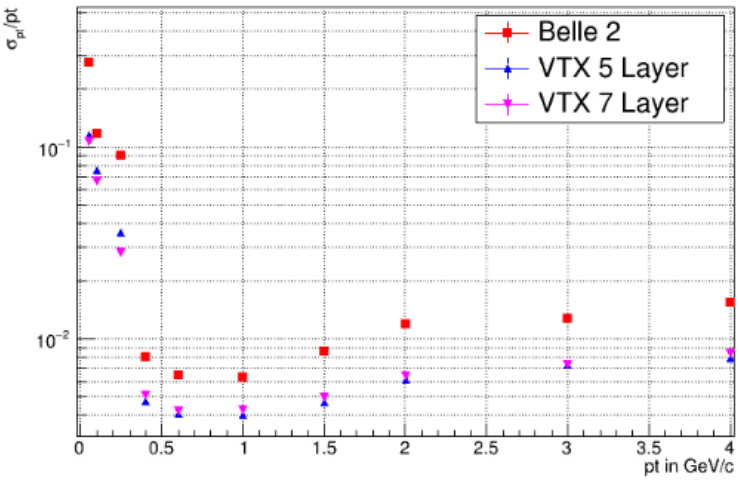

FIGURE 4.13: *Transverse momentum resolution as a function of p_t, with in red the current Belle II VXD, in blue the five layer and in pink the seven layer geometries.*

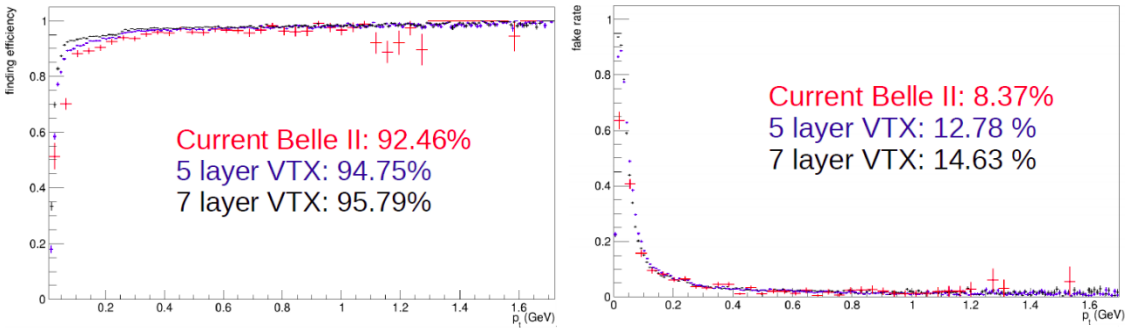

FIGURE 4.14: *Finding efficiency (left) and fake rate (right) for the current Belle II tracking system (in red), the VTX five (in blue) and seven (in black) layer detector as a function of transverse momentum. The percentages correspond to the mean over the whole region.*

For the VTX-only figure of merits, the average layer one occupancy is 0.0016%, which is three orders of magnitude lower than the current VXD, due to its pixelated design and its fast integration time. Also, as shown Table 4.4, the current SVD efficiency drops by 5% on average when the nominal background is multiplied by five as a safety factor. This is not the case for the new geometries. They only drop in absolute tracking efficiency by 0.5% to 1%, which shows that this technology is very robust to the increase of the background.

TABLE 4.4: *Standalone average tracking efficiencies for the current SVD and the two new geometries (5 and 7-layer) for a nominal and a five times rescaled nominal beam-induced background.*

	Background × 1	Background × 5
Current SVD	0.961	0.907
5-layer	0.984	0.979
7-layer	0.987	0.978

4.5 Conclusion

To upgrade the SuperKEKB collider for its peak luminosity, a long shutdown around 2026 is foreseen, which is a good opportunity to also upgrade the Belle II vertex detector. Current estimates indeed show that its occupancies are close to a limit where its performances can degrade. Two new geometries and a fully-pixelated technology were implemented in the Belle II software. The digitizer was tuned to reproduce the TJ-MonoPix1, thus predicting the MonoPix-2 performance. Also, according to the first Monte Carlo simulations, the new designs are more robust against beam-induced background, have a lower occupancy and have better tracking performances. The results were presented at the XXVII Cracow EPIPHANY Conference and lead to a proceeding [79]. Currently, the impact of this upgrade on the physics channels are studied, to further assess the benefits for Belle II from a fully-pixelated detector [80, 81].

Chapter 5

Dalitz-separated analysis of the Time-dependent CP asymmetries in $B^0 \to K^0_S \pi^+ \pi^- \gamma$ decays in Belle

In this chapter, I report the development of the time-dependent asymmetry in the decay $B^0 \to K^0_S \pi^+ \pi^- \gamma$ measurement, done previously in Belle [2], with a new method developed by S. Akar *et al.* [52] to extract more information using Dalitz plots and get new constraints on the imaginary and real parts of $\mathcal{C}'_7$.

This measurement is based on a MC sample that is about $1\times$ the Belle data set size, that contains around 1×10^6 $B\bar{B}$ pairs and it was collected near the $\Upsilon(4S)$ resonance, with the Belle detector at the KEKB asymmetric energy e^+e^- collider.

Section 5.1 describes the MC samples used in the analysis. The reconstruction and the selection strategy is presented Section 5.2. The fit strategy to get the CP violating parameters is explained Section 5.3. Finally, a conclusion and the perspectives of the new measurement and method in Belle is given Section 5.4.

5.1 Data samples used

All the samples used in this analysis are generated and analyzed using the Belle Analysis Software (basf) [82]. basf works similarly to basf2, described in Section 4.2.1, but with older version of its components. EvtGen and Pythia are the generators used to simulate the events and the physical processes, and Geant3 provides the response of the Belle detector when the generated particles go though its layers. The output is written in the format of root files, which can be analyzed with the ROOT software.

To study the signal, 500000 events composed with only $B^0 \to [K^0_1 \to (K^0_S \to \pi^+\pi^-)\ (\rho_0 \to \pi^+\pi^-)]\ \gamma$ were generated, corresponding to a luminosity of about 250 ab^{-1}. This sample, called the signal-only sample, was used to determine the efficiency of the selection and the signal and cross-feed shapes (see Sec. 5.3.2).

Charged (B^+B^-), mixed ($B^0\bar{B}^0$), charm and *uds* background were generated using the Monte-Carlo samples made available by the Belle collaboration and they amount to six times the Belle data set, so around 6 ab^{-1}. The first two backgrounds are merged together to get the $B\bar{B}$ background sample and the last two to get the continuum background sample. The two samples are used in the fit to get the shape of all these background contributions.

The final sample used contains all the $B\bar{B}$ rare decays ($BR < 10^{-4}$) from PDG2004 revision and HFAG Winter 2005. In particular, our signal is a possible decay in this sample. It was also provided by the Belle collaboration and amount for about 50 ab^{-1}. This rare sample was used to fit the rare contribution background.

To do the final fit and measure $\mathcal{A}$ and $\mathcal{S}$ (see Sec. 2.2), the continuum, $B\bar{B}$ and rare samples were resized to 1 ab^{-1} and merged together in a unique sample.

5.2 Reconstruction and selection

Since the idea is to redo the measurement in [2], most of the event selection cuts were not re-optimized in order to compare our results. Our signal is composed of four charged tracks and one high energy photon. The charged are first reconstructed, then combined with vertex fits to get the K^0_S and combined with the photon to get the B^0 candidate.

5.2.1 Charged track reconstruction

From all the charged particles candidates, we select all the reconstructed pions tracks and we ensure that the tracks originate from near the interaction point. We require $|dr| < 1\,\text{cm}$ and $|dz| < 3\,\text{cm}$, where dr (dz) is the distance of the point of the closest approach (in the z axis) in respect to the interaction point.

With the information of the CDC, Cherenkov counter and time-of-flight detector, we can get a likelihood on the nature of the charged tracks. The PID likelihood is defined as:

$$\mathcal{P}(K:\pi) = \frac{\mathcal{L}_K}{\mathcal{L}_K + \mathcal{L}_\pi} \tag{5.1}$$

where $\mathcal{P}(K:\pi)$ is the likelihood that a track is a kaon compared against a pion hypothesis, and $\mathcal{L}_p$ is the likelihood that the particle is of type p. To ensure good pions candidates and to disambiguate them from kaons and electrons, we require $\mathcal{P}(K^\pm:\pi^\pm) < 0.25$ and $\mathcal{P}(e^\pm:\pi^\pm) < 0.25$, which keeps 72% of the signal with a background rejection of 54%.

5.2.2 Photon reconstruction

To select only high energy photons, we get all the reconstructed photons from the events and require $1.5\,\text{GeV} < E^{\text{CMS}}_\gamma < 3.5\,\text{GeV}$. They also have to be in the ECL barrel region ($-0.65 < \cos(\theta) < 0.86$). Since the electromagnetic shower has a narrower shape than the hadronic shower, we require that E9/E25$>$ 0.95, where E9 and E25 are energies summed over the ECL crystal arrays with 3×3 and 5×5 surrounding respectively, including the center of the shower.

One of the important background of radiative decays comes from the decay of the light meson π^0 and η going to $\gamma\gamma$. 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. A

tool has been developed [83] to distinguish and remove these events. Using discriminating variables, like the diphoton mass or the energy, the π^0/η veto tool returns the probability $\mathcal{P}$ that the photon pair originates from a π^0 or η decay. We require $\mathcal{P}_{\pi^0} < 0.2$ and $\mathcal{P}_{\eta} < 0.25$, which keeps 87% of the signal while removing 61% of the background candidates.

5.2.3 K_S^0 reconstruction

K_S^0 decays into two charged pions 69.2% of the time. The two charged tracks are used to reconstruct so called V0-particles. The second dominant decay, $\pi^0\ \pi^0$, in not used. These candidates can come from K_S^0, $\Lambda \to \pi p$ decays, other two tracks decays, combinatorial background or fake tracks. A mass cut within $\pm 20\,\mathrm{MeV}/c^2$ of the K_S^0 mass is applied on all the V0 reconstructed particles.

Then, to have better reconstruction efficiency and purity and to further discriminate Λ decays and K_S^0, a neural network trained on event-shape variables and PID is applied [84]. The network outputs two variables: a probability of the candidate being a V0-particle like and a probability of the candidate not being Lambda like. We require that the probabilities are above 0.25 and above 0.3 for the first and second probability respectively; which yield a purity around 93% and an efficiency around 85%.

5.2.4 $K_S^0\ \pi^+\pi^-$ combination

The $K_S^0\ \pi^+\pi^-$ system can come from a lot of resonances (see Fig. 5.1), as $K^{*0}(892)$ resonance and higher K resonances, like $K_1^0(1270)$ and $K^{*0}(1410)$, which are often denoted as X_{sd}.

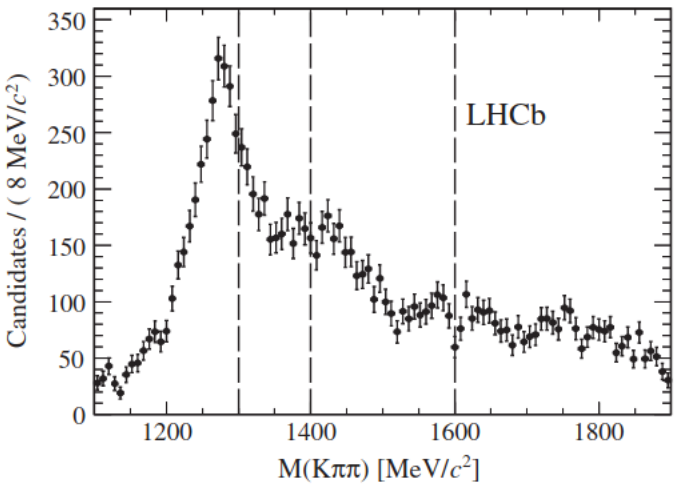

FIGURE 5.1: *Background-subtracted $K^+\ \pi^+\pi^-$ mass distribution of the $B^+ \to K^+\ \pi^+\pi^-\ \gamma$ signal. The same distribution is expected for $B^0 \to K_S^0\pi^+\pi^-\gamma$, assuming isospin symmetry. The dashed lines correspond to different intervals of interest in [20].*

The two reconstructed pions and the K_S^0 are combined together, and the resulting particle is called X_{sd}. To further reduce miss-reconstructed candidates,

we apply a loose cut on its mass, $0.9 < \mathrm{M}_{X_{sd}} < 1.8\,\mathrm{GeV}/c^2$ and a hard cut on its CMS energy, $1 < \mathrm{E}^{CMS}_{X_{sd}} < 3.5\,\mathrm{GeV}$.

5.2.5 B^0 reconstruction

To get the B^0 and $\bar{B}^0$ candidates, we use the same method as in the last subsection: we combine the X_{sd} and the γ candidates.

Since $\Upsilon(4S) \to B^0\overline{B}^0$ is a 2-body decay, we know that the energy of one B is the $\Upsilon(4S)$ frame, E_B^{CMS}, must be half of the total energy of the e^+e^- system. For each event, we know with high precision the energy of the collision. We can then define a new variable m_{bc}, called the beam constrained mass, which have a better resolution than the invariant mass of the B:

$$m_{\mathrm{bc}} = \sqrt{\left(E^{CMS}_{\mathrm{beam}}\right)^2 - |\vec{p}^{\,CMS}_B|^2} \tag{5.2}$$

In addition, we can also define the difference between the reconstructed energy and the beam energy:

$$\Delta E = E_B^{CMS} - E^{CMS}_{\mathrm{beam}} \tag{5.3}$$

We require that the B^0 candidates have $m_{\mathrm{bc}} > 5.2$ GeV/c^2 and $|\Delta E| < 0.2$ GeV. In addition, the two final state pions has to come from a ρ^0 meson, which peaks at 0.78 GeV/c^2 as shown Fig. 5.2. We select the B candidate which fulfill $0.6 < \mathrm{m}_{\pi^+\pi^-} < 0.9\,\mathrm{GeV}/c^2$.

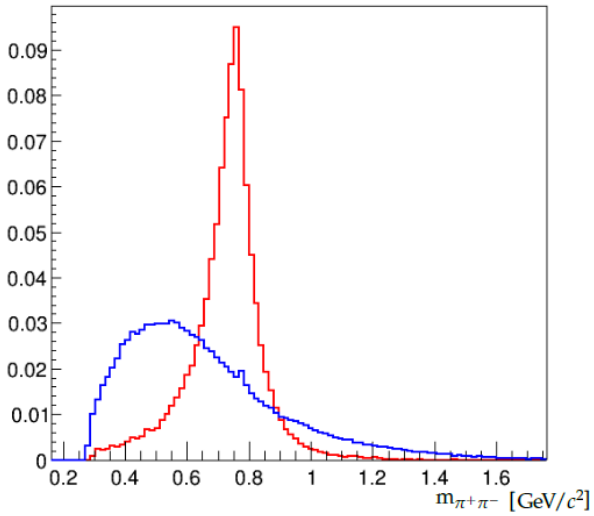

FIGURE 5.2: *Distribution of $m_{\pi^+\pi^-}$ for the $B^0 \to K^0_S \pi^+ \pi^- \gamma$ signal only sample, with the $\rho^0(770)$ resonnance (in red) and all the others (in blue).*

5.2.6 Flavor tagging

In order to measure $\mathcal{A}$ and $\mathcal{S}$, we have to know the flavor of the reconstructed B mesons (see Eq. 2.32). Because of its oscillation before the decay of the B, we

cannot directly measure the flavor. However, we can use information on the second B, called B_{tag}, and take into account the B^0 $\bar{B}^0$ mixing to get the flavor of the signal B meson, called B_{rec} or B_{sig}. This process is called B tagging or flavor tagging in the context of Belle and Belle II.

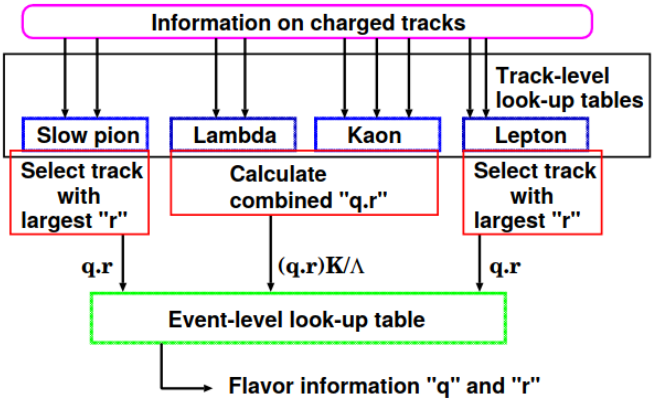

FIGURE 5.3: *Schematic diagram of the workflow of the flavor tagging procedure. q holds the information of the flavor of the B, while and r represents the quality of this information. From [48].*

The procedure used in Belle is described in detail in [48] and a schematic diagram of the workflow of the algorithm is shown Figure 5.3. The algorithm looks at the decay products of the event and try to deduce the flavor information, like the charge of the lepton or kaons in semi-leptonic decays or B^0 to charged kaon decays. Two parameters, q and r, are returned after the algorithm has combined the result of all the tracks. The first one can take only two values: $+1$ if B_{tag} is preferably B^0 and -1 if it is preferably $\bar{B}^0$. The second parameter can take a value between 0 and 1 and summarizes the quality of the information on q. If $r = 0$, we have no information on the flavor; if $r = 1$, the flavor was unambiguously assigned.

In addition, we have to take into account the wrongly tagged candidates w, and Δw the difference in fraction of wrongly tagged B^0 and $\bar{B}^0$. Their values, expected to be decay mode independent, have been estimated using self-tagged decays like $B^0 \to D^{*-} l^+ \nu$, $D^{*-}\pi^+$ and $D^{*-}\rho^+$. The values of w and Δw depend on the quality of the information we have on the flavor of the B. We divide the reconstructed B candidate into seven bins according to the value of r and obtain the values of w and Δw.

The miss-tag fractions range from 2% at high r-bins to 40% at low r-bins. The differences in the wrong tag fractions between B^0 and $\bar{B}^0$ are around 1%. In total, the effective tagging efficiency obtained in Belle is around 29%.

5.2.7 Vertex fitting

In order to get the time interval between the decay of the reconstructed B_{rec} meson and the tagged B_{tag} meson, we have to get a precise measurement of the distance between their decay vertices. The two quantities are related by:

$$\Delta t \simeq \frac{z_{rec} - z_{tag}}{\gamma \beta c} \tag{5.4}$$

where z_{rec} and z_{tag} are the z coordinate of the B_{rec} and B_{tag} decay vertex, and $\beta\gamma = 0.425$ is the boost of the B meson pair in Belle. Having a precise vertex resolution is important, since the asymmetry depends on the decay time.

The vertexing algorithm of Belle tries to get the best estimation of the common origin point of the tracks by minimizing the candidate track χ^2 and using a Kalman filter [85]. To get more constrain for the kinematical fit, we also exploit the beam interaction point (IP). We require the fit of the vertex to be consistent the IP profile, which size is typically around $\sigma_x \simeq 100\,\mu$m, $\sigma_y \simeq 5\,\mu$m, and $\sigma_z \simeq 3mm$. To take into account the B decay length and to avoid introducing a bias in the z position, the constrain is not modeled by a three dimensional Gaussian, but rather a tube shape profile called IP tube [86], as depicted Fig. 5.4. It has the same size in the xy plane, and it is smeared by an additional 21 μm.

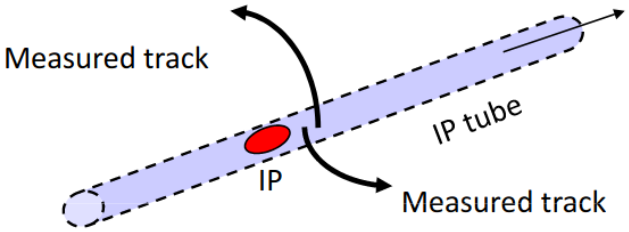

FIGURE 5.4: *Schematics of the IP tube constrain [87].*

However, the IP constrain biases the standard quality indicator of the vertex fit χ^2/ndf, as the tube makes it correlated with the z-position. We define a new vertex quality indicator:

$$h = \frac{\sum_{k=1}^{n} \chi^2_k}{2n - 2}, \tag{5.5}$$

with n the number of tracks and χ^2_k the χ^2 of the k-th track. This unbiased indicator corresponds to the reduced χ^2 without the IP constrain taken into account. If the vertexing has failed, the B candidate is rejected.

For the signal side, we determine the vertex using the two final state pions. We require them to have a minimal number of hits in the vertex detector $N^{SVD}_{r\phi} >= 1$ and $N^{SVD}_{z} >= 2$ in the $r\phi$ plane and z axis. We also require that the candidates have a vertex quality $h < 50$.

For the tag side, every remaining tracks not used for the signal side are taken for the vertexing. We require $h < 50$ for multiple tracks candidates.

5.2.8 Continuum suppression

To separate the signal from continuum background events ($e^+e^- \to q\bar{q}$ ($u\bar{u}$, $c\bar{c}$, $d\bar{d}$, $s\bar{s}$)), we use a fast boosted decision tree (fastBDT), described in [74].

Given a set of variables (like B properties, kinematics, vertex information ...) and a discriminating one to know if the candidate is signal or background, the fastBDT tries to optimize the cuts on all the variables and combine them in one output variable, called CSMVA. Doing a cut on this variable removes 80% of the background, while only removing 15% of the signal candidates.

For the MC data, all the B candidates and their children are linked to their corresponding MC information, coming from the generator. We can then know if the B candidate is our $B^0 \to K^0_S \pi^+ \pi^- \gamma$ signal decay or not.

As input, we give two samples: one with around 110000 signal only events and one with around 500000 continuum background events. We train the BDT on eleven event-shape variables, described below. Two third of the events in the two samples are used to train the BDT, and the remaining third is used as a testing sample to see if the BDT was not over-trained and to compare the performance with the training sample.

Thrust axis

The thrust of an ensemble of particles is defined as the direction in which the sum of the projections of the momenta of the particles is maximized. As shown Fig. 5.5, the $q\bar{q}$ events are created back-to-back in the center-of-mass frame, while the $B\bar{B}$ events are mostly isotropic.

We input in the BDT the transverse energy, the $\cos(\theta)$ of the B candidate in the CMS frame and the cosine of the angle between the candidate and the thrust axis of the event. For $B\bar{B}$ events, their distribution are flat, while for $q\bar{q}$ they peaks mostly at large values.

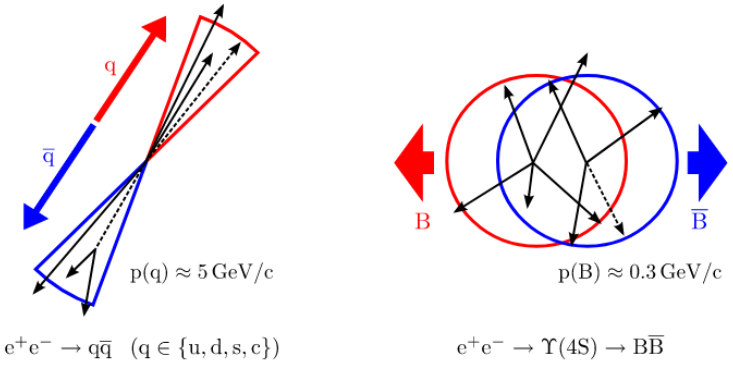

FIGURE 5.5: *Illustration of the shapes of continuum (left) and $B\bar{B}$ (right) events. Light quark pairs produced in the e^+e^--annihilation are produced back-to-back and give rise to a jet-like structure of continuum events, while $B\bar{B}$ events have spherical shape [88].*

Modified Fox-Wolfram moments

To describe the event topology of e^+e^- annihilation, we can use the k-th Fox-Wolfram moment H_k is defined as:

$$H_k = \sum_{i,j} \frac{|\mathbf{p}_i| \, |\mathbf{p}_j| \, P_k(\cos\theta_{ij})}{E^2_{vis}}, \tag{5.6}$$

where P_k denotes the k-th order Legendre polynomial, θ_{ij} is the opening angle between the momentum $\mathbf{p}$ of the i-th and the j-th particle and E_{vis} is the sum of the measured energy in an event. We also define the k-th normalized Fox-Wolfram moment R_k as:

$$R_k = \frac{H_k}{H_0}. \tag{5.7}$$

This provides a good continuum to $B\bar{B}$ event separation. For example, the continuum events with 2 jet-like structures will have R_k values close to zero or one, and spherical $B\bar{B}$ event topologies will have different values.

The Belle collaboration refined the definition, to develop so-called Modified or Super-Fox-Wolfram moments [89]. The idea is, in the summation over the momenta of the i-th and the j-th particle, to not sum over all final state particles. Instead, we define H^{ij}_{xk}, where the superscripts i and j can be: s indicates that one of the indices in the summation takes the final state particles of the reconstructed B candidate and o indicated that one of the indices run over the remaining final-state particles in the event. In addition, the summation only runs on the x category: charged ($x = 0$), neutral ($x = 1$) or missing ($x = 2$) group of particles. k corresponds to the k-th order Legendre polynomial. For H^{oo}_k, the x category is not taken into account. There are in total 16 KSFW moments in basf: $H^{so}_{\{0,1,2\}\times\{0,2,4\}}$ (9), $H^{so}_{\{0\}\times\{1,3\}}$ (2) and $H^{oo}_{\{0-4\}}$ (5).

The most important discriminating variables are shown Figure 5.6. The output of the fastBDT is shown Fig. 5.7. A good discrimination between signal and background and no over-training is observed. We remove every events which have a *CSMVA* value below zero.

5.2.9 Best candidate selection

After the selection, some events still have multiple B candidates. We select the candidate which have the highest vertex confidence level. This cut removes candidates in about 6% of the events.

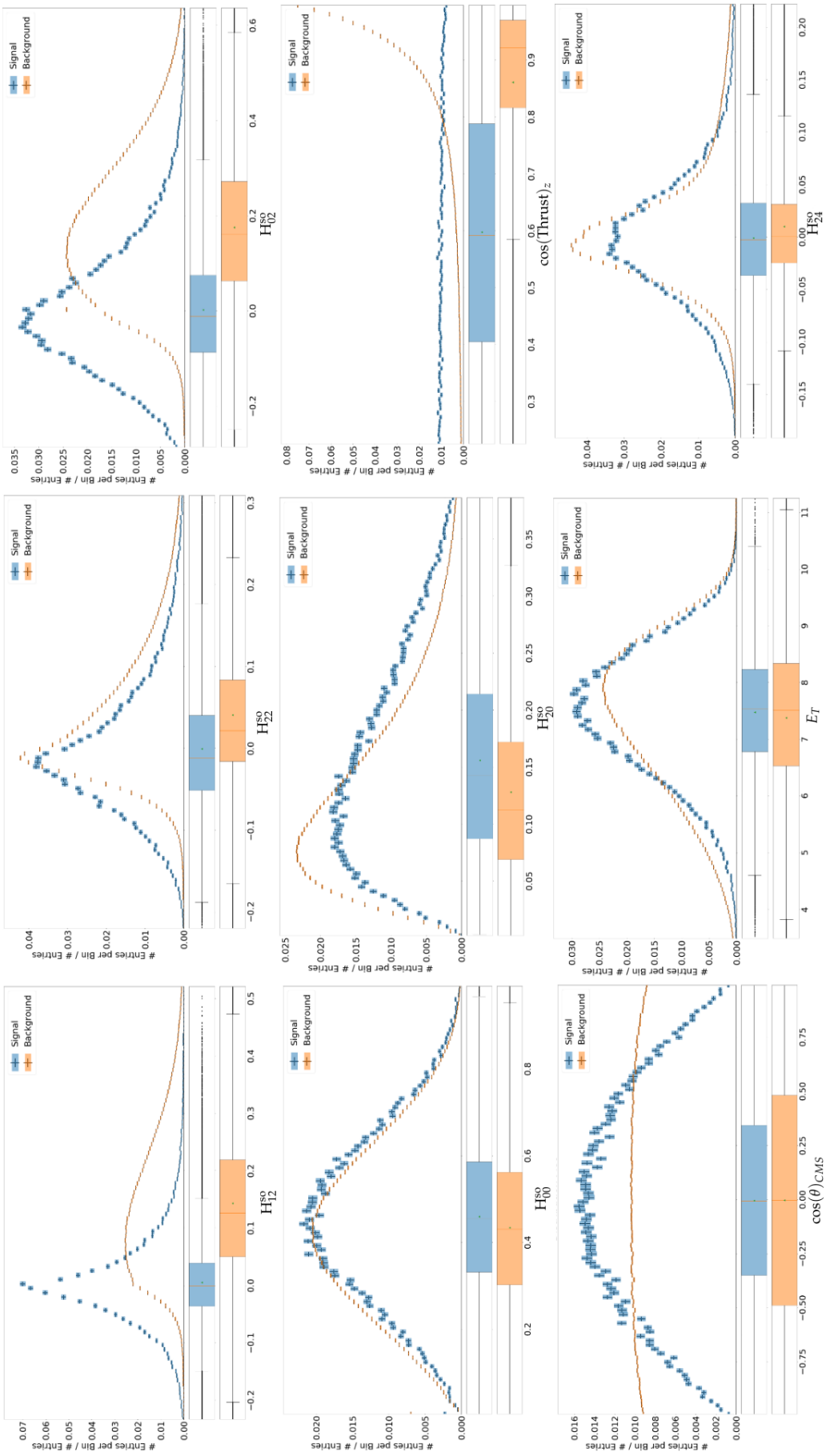

FIGURE 5.6: *Distributions of signal (in blue) and background (in orange) of the nine most discriminating variables used as an input for the fastBDT training.*

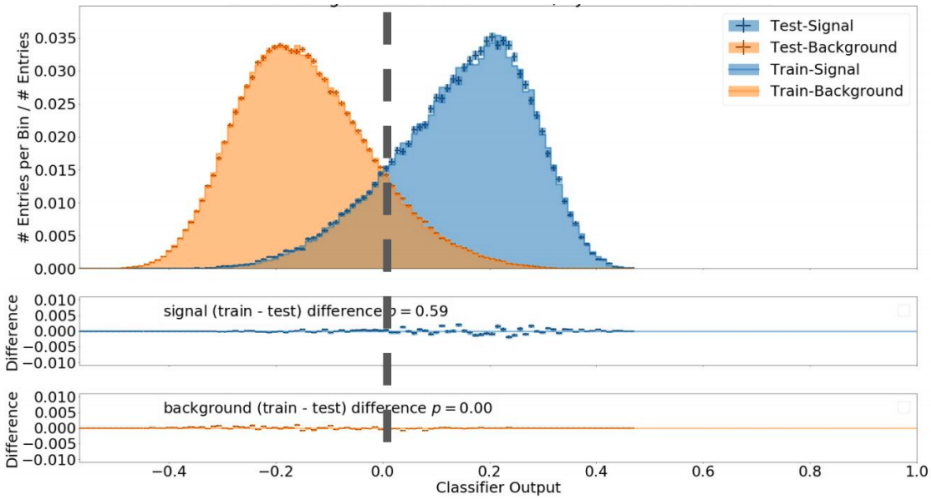

FIGURE 5.7: *Output of the CSMVA training for signal (in blue) and background (in orange). The black dashed line correspond to the selection cut. The overtraining is checked using the difference plot between the test and train sample. The p-value comes from a Kolmogorov-Smirnov test.*

5.3 Fit strategy

To extract $\mathcal{A}$ and $\mathcal{S}$, we perform a three dimensional extended unbinned maximum likelihood fit to $m_{\rm bc}$, ΔE and Δt. The fit is done by RooFit [90] and the likelihood is computed using the MINUIT software [91]. We implemented 3D fits on these three variables on the signal and each individual background MC samples to get values of the probability density functions parameters. Finally, we combined them all into a $\sim$1 ab^{-1} MC sample to do the 3D fit.

5.3.1 Probability Density Functions

A probability density function (PDF) describes the relative likelihood for a continuous random variable to take a given value. A PDF is, by construction, non-negative and normalized to unity over the whole range. The three-dimensional PDFs are build by the product of three unidimensional PDF for each variable, $m_{\rm bc}$, ΔE and Δt. For the final fit, we split the Dalitz plot to fit two $\mathcal{A}$ and $\mathcal{S}$ in the two regions. We will quickly define the different base PDFs used to fit the different MC samples.

Gaussian Distribution

The standard Gaussian PDF takes two input parameters and is defined as:

$$\mathrm{G}(x;\mu,\sigma) = \frac{1}{N}\cdot e^{-\frac{(x-\mu)^2}{2\sigma}} \tag{5.8}$$

where μ and σ are the mean and the width, respectively, of the function. N is a normalization factor. It is used for some $m_{\rm bc}$ and ΔE distributions, and in the Δt resolution function.

Exponential Distribution

The standard exponential PDF takes one input parameter and is defined as:

$$\mathrm{E}(x;\lambda) = \begin{cases} \frac{1}{N}e^{-\lambda x} & x \geq 0, \\ 0 & x < 0. \end{cases} \tag{5.9}$$

where λ describes the slope of the function. N is a normalization factor. It is used in some ΔE distributions, and in the Δt resolution function.

Chebyshev Distribution

Chebyshev Polynomials are an orthogonal set of polynomials which can be used to approximate any function in a given interval using a sufficient high order. The n-th order Chebyshev takes n input parameters. We use Chebyshev Polynomials

of the first kind which are defined as:

$$\text{Cheb}_n\left(x; c_1, c_2, \ldots, c_n\right) = \frac{1}{N} \cdot \left(T_0(x) + \sum_{i=1}^{n} c_i \cdot T_i(x) \right) \tag{5.10}$$

where N is a normalization factor and the $T_i(x)$ are defined using the recurrence relation:

$$T_0(x) = 1 \quad T_1(x) = x \quad T_{n+1}(x) = 2x \cdot T_n(x) - T_{n-1}(x). \tag{5.11}$$

It is mainly used for m_{bc} and ΔE.

Argus Distribution

The Argus PDF is an empirical function introduced by the ARGUS collaboration [92] to model the invariant mass distribution for the continuum component to the background. It takes two input parameters and is defined as:

$$\text{Argus}(x; a, c) = \frac{1}{N} \cdot x \cdot \sqrt{1 - \frac{x^2}{c^2}} e^{a \cdot \left(1 - \frac{x^2}{c^2}\right)} \tag{5.12}$$

with the curvature parameter $a < 0$ and the cut-off energy $c > 0$. N is a normalization factor.

In our case, the cut-off parameter is the beam energy, so we only have to determine the curvature parameter:

$$\text{Argus}(x; a) = \text{Argus}\left(x; a, E_{\text{beam}}\right) \tag{5.13}$$

Crystal Ball Distribution

The Crystal Ball PDF [93], named after the Crystal Ball Collaboration, consists of a Gaussian core portion and a power-law low-end tail, below a certain threshold. It is mainly used to fit m_{bc}. It takes four parameters as input and is defined as:

$$f(x; \alpha, n, \mu, \sigma) = N \cdot \begin{cases} \exp(-\frac{(x-\mu)^2}{2\sigma^2}), & \text{for } \frac{x-\mu}{\sigma} > -\alpha \\ A \cdot (B - \frac{x-\mu}{\sigma})^{-n}, & \text{for } \frac{x-\mu}{\sigma} \leqslant -\alpha \end{cases} \tag{5.14}$$

where

$$A = \left(\frac{n}{|\alpha|}\right)^n \cdot \exp\left(-\frac{|\alpha|^2}{2}\right), \tag{5.15}$$

$$B = \frac{n}{|\alpha|} - |\alpha|, \tag{5.16}$$

N is a normalization factor, μ and σ drive the mean and the width, respectively, of the Gaussian part, and α and n drive the power-law end tail and their values are greatly correlated to each other.

Δt distribution

For Δt, a specific software tool called *Tatami* developed in Belle is used to fit the signal and all the background contributions [86, 94-96].

The typical background PDF is composed of one physics-like-part P_{bkg} convoluted to one resolution-like-part R_{bkg}. It is defined as:

$$F_{\text{bkg}}(\Delta t) = \int_{-\infty}^{\infty} P_{\text{bkg}}(\Delta t')\, R_{\text{bkg}}(\Delta t - \Delta t')\, d\Delta t' \tag{5.17}$$

where the physics and resolution parts are defined as:

$$P_{\text{bkg}}(\Delta t) = f_\delta \delta(\Delta t - \mu_d) + (1 - f_\delta) \exp\left(-\frac{|\Delta t - \mu_l|}{\tau_{\text{bkg}}}\right) \tag{5.18}$$

$$R_{\text{bkg}}(\Delta t) = (1 - f_{tail})\, G(\Delta t; S_{main}\sigma_{vtx}) + f_{tail} G(\Delta t; S_{tail}\sigma_{vtx}) \tag{5.19}$$

and

$$G(x;\sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp^{-\frac{x^2}{\sigma^2}} \tag{5.20}$$

$$\sigma_{vtx} = \sqrt{\sigma_{rec}^2 + \sigma_{asc}^2} \tag{5.21}$$

The input parameters are as follow: S_{main} and S_{tail} are the scale factor for the main or tail part of the resolution function, μ_d and μ_l the mean shift of the prompt component, τ_{bkg} the mean lifetime of the exponential component and f_δ and f_{tail} the fractions of the delta and Gaussian components.

For the signal, the function used to fit Δt can be decomposed into four factors that affect the resolution:

$$F_{\text{sig}}(\Delta t) = P_{\text{phys}} \otimes R_{\text{sig}}(\Delta t) \otimes R_{\text{tag}}(\Delta t) \otimes R_{\text{np}}(\Delta t) \otimes R_{\text{kin}}(\Delta t), \tag{5.22}$$

where P_{phys} is the time-dependent decay rate in Eq. 2.32, reminded here:

$$P_{\text{phys}}(\Delta t) = \frac{1}{4\tau_{B^0}} \exp^{-\frac{|\Delta t|}{\tau_{B^0}}} \left[1 - q\Delta w + q(1 - 2w) \times (\mathcal{S}\sin(\Delta m \Delta t) + \mathcal{A}\cos(\Delta m \Delta t))\right], \tag{5.23}$$

which takes as input a lot of event-by-event information like the flavor q of the B, the miss-tag fraction w or information on CP and tag side. R_{vtx} and R_{tag} describe the vertex resolutions of the B_{sig} and B_{tag} candidate. R_{np} represents misplacing of the reconstructed vertex due to non-primary tracks. Finally, $R_{\text{kin}}(\Delta t)$ describes the kinematic approximation coming from the analytic formula relating the relation between Δt and Δz.

5.3.2 Results

The results of the signal and all the background fits are described below. In all the Δt plot, we only plot the events in the signal region in $m_{\rm bc}$ and ΔE, *i.e.* events in $5.27 < m_{\rm bc} < 5.29\,{\rm GeV}/c^2$ and $-0.1 < \Delta E < 0.05\,{\rm GeV}$.

Signal and cross-feed fit

Using the MC signal only sample after the selection, the cuts and the single candidate selection, 13380 signal events remain. The 3D fit is performed with only B candidate labeled as signal on the MC sample. After this cut, 11417 signal events remain, which gives a efficiency of $2.29 \pm 0.02\%$.

A Crystal Ball shape is used for $m_{\rm bc}$ and ΔE while the tatami function $F_{\rm sig}$ is used for Δt. All parameters are left floating in the fit.

The fit is shown on the Fig. 5.8. We observe a slight disagreement in Δt. Since this fit uses more than 100 times more data than what we will expect, some non-linearity behavior of the tatami function appears. No discrepancies are observed when the luminosity corresponds to two times the expected signal (see Appendix B.1.2). As we do not fix Δt parameters from this fit, this doesn't impact the final measurement.

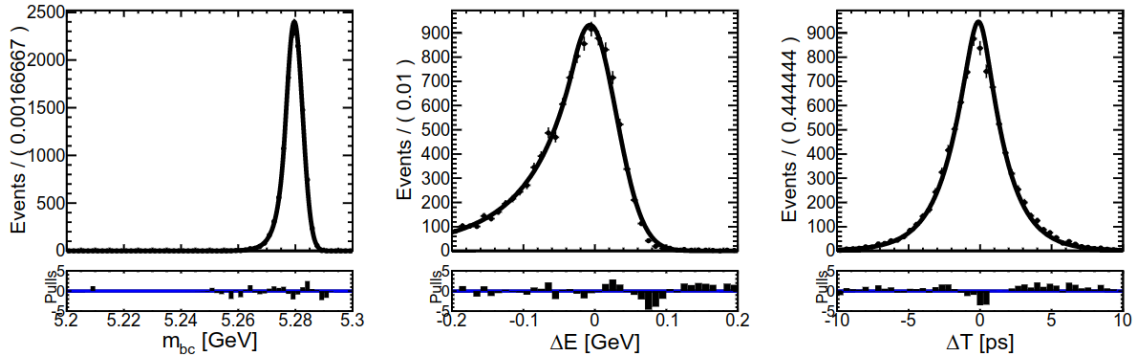

FIGURE 5.8: *3D fit of $m_{\rm bc}$, ΔE and Δt for the signal component on a private MC signal sample with 11417 signal events after the selection. Table with every parameters of the functions used can be found in Table B.1.1.*

The self cross-feed background, or *cross-feed*, described wrongly identified K or π tracks in the three track combination.

The cross-feed is fitted using the same data set, but on all the B candidate labeled as not our signal. 1963 events are remaining, thus 14.7% of the signal is self cross-feed. The $m_{\rm bc}$ shape is composed of a sum of a Gaussian and an Argus as:

$$F_{m_{\rm bc}} = f \times {\rm G} + (1-f) \times {\rm Argus}, \tag{5.24}$$

where f is a coefficient between zero and one. ΔE is fitted with the sum of an exponential and a Gaussian, and Δt uses the tatami $F_{\rm bkg}$ function. All the parameters are left floating. The results are shown in Fig. 5.9.

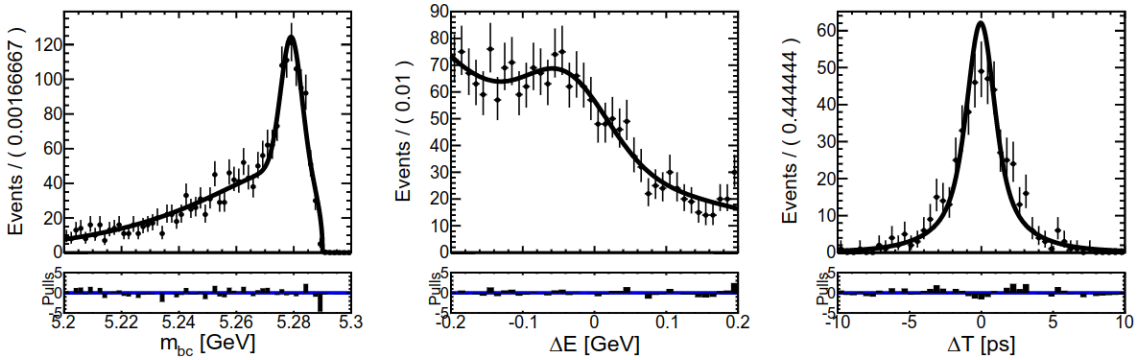

FIGURE 5.9: *3D fit of m_{bc}, ΔE and Δt for the cross-feed component on a private MC signal sample with 1963 signal events after the selection. Table with every parameters of the functions used can be found in Tab. B.2.*

Continuum background fit

To fit the $q\bar{q}$ events, we use the continuum background data set, corresponding to a luminosity of six times the final Belle data set. After the selection, the cuts and the single candidate selection, 40287 continuum events are remaining. For the 3D fit, an Argus shape is used for m_{bc}, an exponential is used for ΔE and the tatami function for Δt. All the parameters are left floating in the fit. The results are shown on the Fig. 5.10.

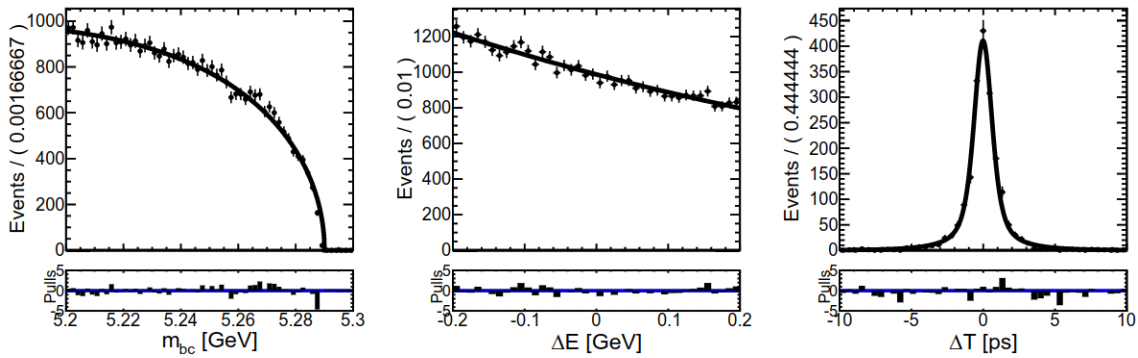

FIGURE 5.10: *3D fit of m_{bc}, ΔE and Δt for the continuum component (charm and uds contributions) on a MC generated sample, with 40287 events after the selection. Table with every parameters of the functions used can be found in Tab. B.3.*

$B\bar{B}$ background fit

The fit the $B^0\bar{B}^0$ or B^+B^- background, we use the $B\bar{B}$ background data set corresponding to a luminosity of six times the final Belle data set. After the selection, the cuts and the single candidate selection, 1075 $B\bar{B}$ background events are remaining. For the 3D fit, the sum of a Gaussian and an Argus shape is used for m_{bc}, an exponential is used for ΔE and the tatami function for Δt. All the parameters are left floating in the fit. The results is shown on the Fig. 5.11.

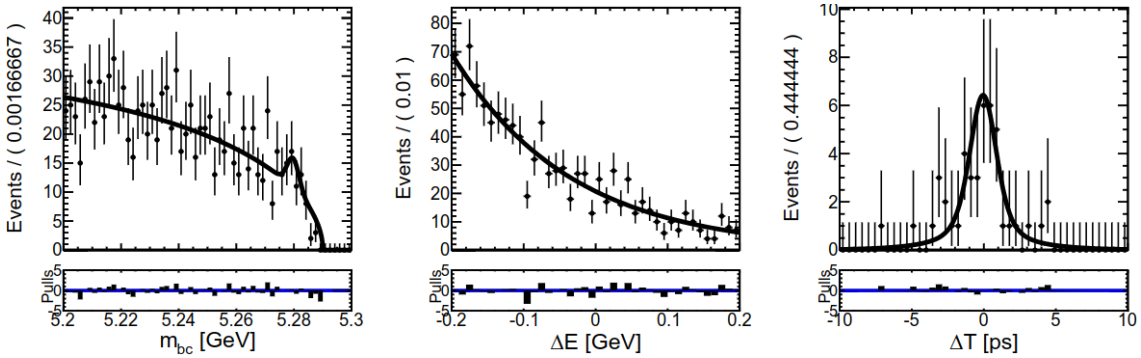

FIGURE 5.11: *3D fit of m_{bc}, ΔE and Δt for the $B\bar{B}$ component (mixed and charged contributions) on a MC generated sample, with 1075 events after the selection. Table with every parameters of the functions used can be found in Tab. B.4.*

Rare decays fit

To fit the rare decays contribution, we use the rare data set corresponding to fifty times the final Belle data set. After the selection, the cuts and the single candidate selection, 75364 events are remaining. This includes the rare background contribution and our signal. To only fit the rare and remove the signal and the cross-feed background coming with it, a new variable *isSgnEvt* was developed. If the event contains our signal or not, the variable is equal to one or zero respectively. It checks if one of the most important kaonic decay is present in the event by looking at MC information. This includes most of the $B^0 \to K_{\text{res}} \to K^0_S \pi^+ \pi^- \gamma$ decays, where $K_1(1270)^0$ or $K^*_2(1430)^\pm$ go to $\pi^\pm$ and $K^*(892)^\pm$, $K^*_0(1430)^\pm$; or go to K^0_S and $\rho(770)^0$, $\omega(782)$ or $f_0(1370)$. By cutting out events with *isSgnEvt* equals to one, we make sure that there are only rare events in the selection. 11753 events remain after this cut.

For the 3D fit, the sum of a Crystal Ball and an Argus shape is used for m_{bc}, an exponential is used for ΔE and the tatami function is used for Δt. All the parameters are left floating in the fit. The results are shown on the Fig. 5.12.

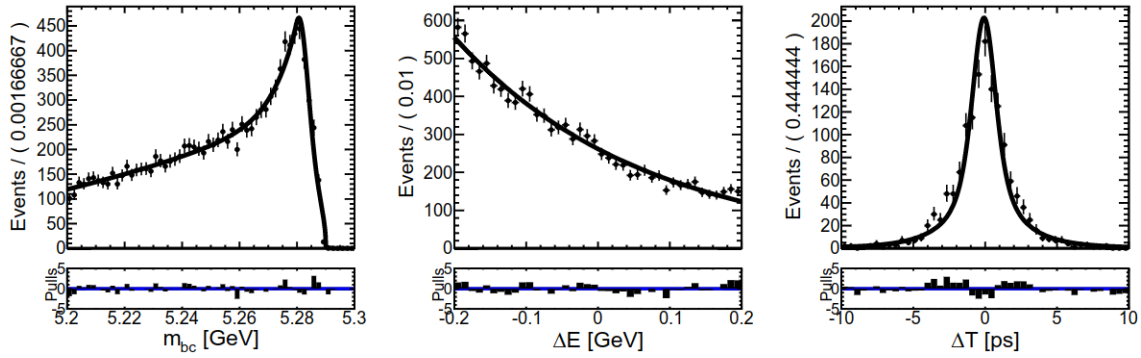

FIGURE 5.12: *3D fit of m_{bc}, ΔE and Δt for the rare components on a MC generated sample, with 11753 events after the selection. Table with every parameters of the functions used can be found in Tab. B.5.*

5.3.3 Final A and S fit

A and S are fitted using the resized one time Belle data set, composed by rare, $B\bar{B}$ and continuum contribution.

The $B\bar{B}$ and continuum fractions are automatically determined using their number of events. For the rare sample, the number of events with the B candidate labeled as our signal correspond to signal, and 14.7% of those were measured as cross-feed. The remaining events are considered as rare.

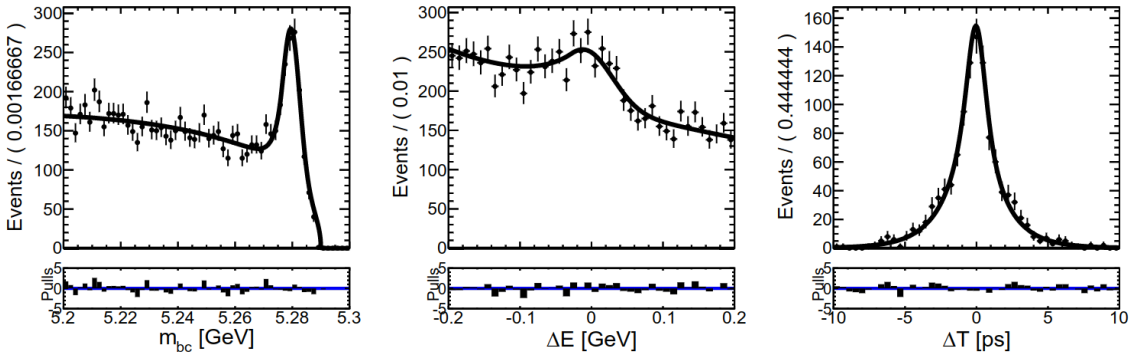

FIGURE 5.13: *3D fit of m_{bc}, ΔE and Δt after the selection.*

In the final sample, there are 6714 continuum background, 179 $B\bar{B}$ background and 1507 rare contribution, which contains 646 rare events, 734 signal and 127 cross-feed events. The PDF of $m_{\rm bc}$, ΔE and Δt are the convolution of all the PDF used for the individual contributions. The fractions are fixed, and every parameters of every shapes are fixed to the values that have been fitted in the previous plots. Only $\mathcal{A}$ and $\mathcal{S}$ are free. The result are shown in Fig. 5.13. The CP-violation parameters measured values are:

$$\mathcal{A} = 0.07 \pm 0.10 \tag{5.25}$$
$$\mathcal{S} = -0.09 \pm 0.12. \tag{5.26}$$

The statistical uncertainty comes from the fitting procedure. The result is compatible with zero, as expected from the inputs into the simulation. This result predicts a statistical uncertainty which is a good improvement (around a factor two) compared to the last Belle measurement [2].

The fit with the Dalitz separation (see Fig. 2.4) allows to extract the two new parameters $\mathcal{S}^+$ and $\mathcal{S}^-$, defined Eq. 2.43:

$$\mathcal{S}^I = -0.07 \pm 0.17 \tag{5.27}$$
$$\mathcal{S}^{\bar{I}} = -0.1 \pm 0.17, \tag{5.28}$$

which gives:

$$\mathcal{S}^+ = -0.17 \pm 0.24 \tag{5.29}$$

$$\mathcal{S}^- = 0.03 \pm 0.24. \tag{5.30}$$

These results are also compatible with the zero CP violation present in the simulated data used for the fit. Of course, we observe a larger statistical uncertainty due to the separation of the sample in two equal parts.

5.4 Conclusion and perspectives

I developed on simulated data a measurement of CP violation parameters for the $B^0 \to K^0_S \pi^+ \pi^- \gamma$ on MC simulations in Belle, resulting in an excepted improved statistical uncertainties compared to the last measurement. I also implemented on the same simulated sample the first measurement of the two new parameters $\mathcal{S}^+$ and $\mathcal{S}^-$ that give better constrain on C_7 and C'_7, detailed in [52]. All the work is resumed in an internal Belle Note [97].

Before measuring the parameters in data, the study of all the systematic uncertainty has to be done. In addition, the interest of the new method based on a separation of the Dalitz plot requires more data to reach interesting constraints. As Belle II collected more and more data, I switched my focus to the Belle II rediscoveries. The systematic study was resumed in 2022 by a postdoc in the Strasbourg group.

To finally use this measurement in the constrain of C_7 and C'_7, we have to combine our efforts with a detailed amplitude analysis of $B^+ \to K^+_{\text{res}} \gamma \to K^+ \pi^+ \pi^- \gamma$, which is under study in a different Belle group.

Chapter 6

Rediscovery and branching fraction measurements for $B^0 \to K^0_S \pi^+ \pi^- \gamma$ and $B^0 \to K^0_S \pi^0 \gamma$

This chapter presents the work I did for the rediscovery of $B^0 \to K^0_S \pi^+ \pi^- \gamma$ and $B^0 \to K^{*0} \gamma \to K^0_S \pi^0 \gamma$ channels based on a full Monte-Carlo (MC) simulation and data of the Belle II experiment. Measurements of the respective branching fractions are targeted with the data samples from 2019 to 2021 at the $\Upsilon(4S)$ resonance, which corresponds to an integrated luminosity of 189.26 fb^{-1}. I optimize the particle selections over a simulated data sample equivalent to 900 fb^{-1}. I use one dimension ΔE fits to determine the signal yields for the $K^0_S \pi^+ \pi^- \gamma$ and $K^0_S \pi^0 \gamma$ final states. Efficiencies are extracted from dedicated simulation samples featuring 2M signal-only decays.

The next section describes the different data set used. Then, the skimming and reconstruction procedure is explained Section 6.2, followed by the selection strategy and its results. After that, the remaining background contribution is shown in Section 6.5 . The final fit with yield extractions on MC samples are shown in Section 6.6. The systematic uncertainties are presented in Section 6.7, before the unblinding strategy and the final branching fraction measurement in Section 6.9.

6.1 Data sample used

The main simulation sample[1] corresponds to 900 fb^{-1} integrated luminosity combining the generic events $e^+ e^- \to B^0\overline{B}^0$, B^+B^-, $u\overline{u}$, $d\overline{d}$, $c\overline{c}$ and $s\overline{s}$. The $B^0\overline{B}^0$ sample, which contains our two decay channels, is referred in the note as mixed-MC sample. The $B^0\overline{B}^0$, without our signal, and B^+B^- combined is referred as $B\overline{B}$ background MC sample, and the $u\overline{u}$, $d\overline{d}$, $c\overline{c}$ and $s\overline{s}$ combined sample is referred as the $q\overline{q}$ MC sample. We divide the 900 fb^{-1} sample into a 700 fb^{-1}, used to optimize the selection, to determine the fit functions and background rejection; and a 200 fb^{-1} sample, which is about the same size as data, used to estimate the expected yields.

[1] In Belle II, a new Monte-Carlo production is planned every year. For this analysis, we used the *MC*14 production from the end of 2020 / beginning of 2021.

To calculate efficiencies, we used two large samples of signal only events, each with 2 million $B \to K_{res}\gamma$ decays. For the $B^0 \to K^0_S\pi^+\pi^-\gamma$ event set, K_{res} decay to different kaonic resonances (mainly $K_1(1270)^0$ with $BF < 5.8 \times 10^{-5}$ [23], $K_1(1400)^0$ with $BF < 1.2 \times 10^{-5}$, $K_2(1430)^0$ with $BF = (1.24 \pm 0.24) \times 10^{-5}$ and $K^*(1680)^0$ with $BF < 2.0 \times 10^{-3}$) and X_{sd}. For the $B^0 \to K^0_S\pi^0\gamma$ event set, only $B \to (K^{*0}(892) \to K^0_S\pi_0)\gamma$ are simulated.

The signal and $e^+\,e^- \to \Upsilon(4S) \to B\bar{B}$ samples are generated using the EvtGen package [69], while the continuum background events are generated with the KKMC [98] generator interfaced with Pythia [70]; the detector response is then simulated by the Geant4 package [71].

The data used in this analysis were collected between March 2019 and July 2021. They correspond to a total integrated luminosity of 190 fb^{-1} obtained at the center-of-mass energies at or near the $\Upsilon(4S)$ resonance (10.58 GeV). The sample corresponds to $N_{B\overline{B}} = (197 \pm 6) \times 10^6$ $B^0\overline{B}^0$ and B^+B^- events determined following Ref. [99].

Both data and simulated data sets are analyzed with Belle II analysis software framework, basf2 [67] (see Sec. 4.2.1).

6.2 Reconstruction and pre-selection strategy

The first step of the analysis is to reconstruct both $B^0 \to K^0_S\pi^+\pi^-\gamma$ and $B^0 \to K^0_S\pi^0\gamma$ in the MC samples. Two similar reconstruction scripts have been developed. To skim a large part of the background, we apply some lose cuts on track parameters, mass energies or PID of the reconstructed tracks. This step is called pre-selection, or skimming.

6.2.1 Reconstruction

Both $b \to s\gamma$ channels discussed share two final state particles: K^0_S and high energy photon γ. The unified reconstruction scheme of these particles will help to evaluate the systematic uncertainties for both analyses. The particle are reconstructed as follow.

Charged track reconstruction

To skim some of the events, we require that the event must have a least 3 tracks which satisfy $|d_r| < 0.5$ cm, $|d_z| < 2.0$ cm, where dr (dz) is the distances of the point of the closest approach (in the z axis) in respect to the interaction point, and at least 20 hits in the CDC. We also require a visible CMS energy above 4 GeV. In data, a small shift in momentum was observed. We apply a correction factor of 0.99971 on all the tracks.

All pions tracks that would belong to our signal must be in the CDC acceptance and the track fit has to have a positive χ^2. We also require $|d_r| < 0.5\,\text{cm}$, $|d_z| < 3\,\text{cm}$. Finally, to discriminate fake pion recontruction, we apply a global PID cut, defined as:

$$\mathcal{P}(\pi) = \frac{\mathcal{L}_\pi}{\mathcal{L}_e + \mathcal{L}_\mu + \mathcal{L}_\pi + \mathcal{L}_K + \mathcal{L}_p + \mathcal{L}_d} > 0.01, \tag{6.1}$$

where, in the denominator, are summed the likelihoods of the reconstructed track to be an electron, muon, pion, kaon, proton and deutron.

Finally, a list composed of pions that passes $p_t > 0.1$ GeV/c, $|d_0| < 0.5$ cm, $|z_0| < 2,0$ cm and at least 20 hits in the CDC is used to create the event-shaped variables, used in the continuum suppression.

Photon reconstruction

Events with at least one ECL cluster in the tracking acceptance and $E_{\text{cluster}} > 0.2$ GeV are kept.

To select only high energy photons, we get all the reconstructed photons from the events and require $1.4\,\text{GeV} < E_\gamma < 4\,\text{GeV}$ in the laboratory frame. They also have to be in the ECL and CDC tracking acceptance. On data, a small bias compared to MC in the photon energy was measured by the collaboration. We correct it by applying a small scaling factor in bins of E_γ.

As in Belle, one of the background of radiative decays comes from the decay of the light meson π^0 and η going to $\gamma\gamma$. To reject such events, a π^0-η veto algorithm was developed in basf2 and is applied on the high-energy photon candidates. The output is a probability $\mathcal{P}_{\pi^0}$, called π^0 likeness, and is equal to one if the photon comes from a π^0 and zero if not. The value of this cut is defined in the cut optimization Section 6.4.2.

For the event-shape variables, we select photons that are in the tracking acceptance and have an energy above 0.1 GeV. We also require that the total deposited energy in the ECL are below 9 GeV.

K^0_S reconstruction

Two reconstructed charged pions are used to fit their vertex using TreeFitter [100], which is a Kalman based filter for a whole particle decay chain. Only candidates with an invariant mass in the range $0.450 < M < 0.550$ GeV/c^2, and for which the vertex fit did not fail, are kept.

π^0 reconstruction

A vertex fit using KFit [101], single vertex fit inherited from Belle, is performed to reconstruct $\pi^0 \to \gamma\gamma$. Only candidates with an invariant mass in the range $0.120 < M < 0.145$ GeV/c^2, and for which the vertex fit did not fail, are kept. The photons used in the fit must be in the CDC acceptance with at least one ECL cluster, which energy must be at least 30 MeV in barrel, 80 MeV in forward endcap and 60 MeV backward endcap.

B^0 reconstruction

The B^0 and $\bar{B}^0$ candidates are reconstructed using two charged pions, a K^0_S and a photon for $B^0 \to K^0_S\pi^+\pi^-\gamma$; and a π^0, K^0_S and a photon for $B^0 \to K^0_S\pi^0\gamma$.

Their vertex is fitted using TreeFitter. Also, the reconstructed B must satisfy $5.2 < m_{\rm bc} < 5.29$ GeV/c^2, $|\Delta E| < 0.5$ GeV and the event must have at least 5 reconstructed tracks. This last cut is not necessary for the moment, since we do not use the tag-side, but it will be in the TDCPV analysis.

6.2.2 Systematic cuts

In order to reduce the continuum background and the miss-reconstructed candidates, we apply a cut on pion PID and π^0 veto before the optimised selection. These so-called systematic cuts feature the following values:

- $\mathcal{P}(\pi) > 0.1$ for $K^0_S\pi^+\pi^-\gamma$.
- $\mathcal{P}_{\pi^0} < 0.7$ for the photon for both channels.

For the pion PID cut, since looser cut $\mathcal{P}(\pi) > 0.01$, is applied in the skimming for the pions in the final state and there are no systematic associated, we took the lowest pion PID cut ($\mathcal{P}(\pi) > 0.1$) with associated systematic (see Sec. 6.7). For the pions coming from the K^0_S, no cut is applied on the skim and in the pre-selection.

For the π^0 likeness cut, its value was optimized in the cut optimization procedure before (see Sec. 6.3.2), but its cut value as a function of the figure of merit plot is mostly flat (see Appendix C.2). The value of 0.7 is compatible with the best cut value found in the optimization process for both channels and has an associated systematic.

The performance of all this procedure is shown in Table 6.2 and Table 6.4 for $B^0 \to K^0_S\pi^+\pi^-\gamma$ and $B^0 \to K^0_S\pi^0\gamma$. The efficiency of the pre-selection and the reconstruction is obtained using the number of truth-matched reconstructed candidates after this procedure divided by the number of generated signal in the 2 M signal only sample. The final efficiency of the pre-selection and the reconstruction is (27.867±0.031)% and (28.865±0.032)% for $B^0 \to K^0_S\pi^+\pi^-\gamma$ and $B^0 \to K^0_S\pi^0\gamma$ respectively.

6.3 Selection strategy

The work flow of the selection strategy consist of the following steps. We first reduce the continuum using a boosted decision tree; then we optimized cuts on specific variables to maximize the significance; finally we select a unique candidate in the event.

6.3.1 Continuum Suppression

The dominated background comes from $e^+e^- \to q\bar{q}$ ($u\bar{u}$, $c\bar{c}$, $d\bar{d}$, $s\bar{s}$). In order to separate $q\bar{q}$ background and $B^0\bar{B}^0$ signal events, we use a binary boosted decision-tree classifier (fastBDT) [74], which combines information in a nonlinear way from several variables and output one CSMVA variable that discriminates signal and background. The tree is trained with 30 event-shape variables on 60000 signal and 60000 $q\bar{q}$ background, and tested on a 30000 signal and 30000 $q\bar{q}$ events.

Thrust axis

As explained in Section 5.2.8, we use the cosine of the angle between thrust axis of the signal B and thrust axis of the rest of the event, and the cosine of angle between thrust axis of the signal B and z-axis. In addition, we input the magnitudes of the rest of the event and the signal B thrust axis.

Fox-Wolfram moments

As also explained in Section 5.2.8, we input 16 Kakuno-Super-Fox-Wolfram moments and the ratio between the second and zeroth Fox-Wolfram moments.

CLEO cones

The 9 left variables that are input in the fastBDT were developed by the CLEO Collaboration to suppress continuum background [102]. They are based on the sum of the absolute values of the momenta of all particles within angular sectors around the thrust axis in intervals of 10°, resulting in 9 concentric cones.

Figure 6.1 shows the distributions of the nine most important variables used in the training of the BDT for $B^0 \to K^0_S\pi^+\pi^-\gamma$ continuum suppression. Similar distributions are observed for $B^0 \to K^0_S\pi^0\gamma$, as shown Fig. 6.2.

6.3.2 Cut optimization

In order to reduce significantly the background, while keeping a maximum number of signal, a selection on a few variables is done for each channel. Using the significance $\frac{S}{\sqrt{S+B}}$ as Figure Of Merit (FOM), the selection criteria is optimized on the 700 fb^{-1} MC sample in the signal range, *i.e.* $5.26 < m_{\rm bc} < 5.29$ GeV/c^2 and $-0.2 < \Delta E < 0.1$ GeV. The procedure is as follow:

1. Take a variable and measure the FOM as a function of the variable cut values and the cut which gives the highest FOM is chosen.
2. The found cut is applied to the data set, then a new variables is optimized the same way.
3. The procedure is repeated until no variable is left.
4. After all the cuts are found and applied, the order of the variables is changed. For each order, the FOM of the whole selection is calculated and the final selection result is the order which gives the highest FOM. All possible orders, 6 for $B^0 \to K^0_S\pi^+\pi^-\gamma$ and 24 for $B^0 \to K^0_S\pi^0\gamma$, are tested.

An example of a cut on K^0_S decay length significance is shown on Figure 6.3 and all the plots of the selection can be found in Appendix C.1.

This procedure is applied on

- the CSMVA output variable to reduce the continuum background,

- the K^0_S decay length significance $d_{K^0_S}/\sigma(d_{K^0_S})$, with $d_{K^0_S}$ the decay length of the K^0_S and σ its resolution, to have a good K^0_S selection and to further reduce the continuum,
- the invariant mass $m_{\rm bc}$ to reduce any background type,
- the $\mathrm{m}_{K^0_S\pi^0}$ invariant mass for $K^0_S\pi^0\gamma$, to remove other kaonic resonances and only keep $K^{*0}(892) \to \pi^0 K^0_S$.

For $B^0 \to K^0_S\pi^+\pi^-\gamma$, the $\mathrm{m}_{K^0_S\pi^+\pi^-}$ invariant mass is not optimized. It is instead fixed to the value used in the last Belle analysis [2] to efficiently select all the $K_{\rm res}$ contributions. The cut $\mathrm{m}_{K^0_S\pi^+\pi^-} < 1.8\,\mathrm{GeV}/c^2$ is thus chosen and applied before the optimization.

The result of the optimization is described in Sec. 6.4.2.

6.3.3 Single candidate selection

After reconstruction and selection is applied, all candidates with a failed B^0 vertex fit are removed. For events containing multiple candidates, the one featuring the lowest χ^2 from the B^0 vertex fit is chosen and the other discarded. If two candidates have the same χ^2, a random one is selected. The plots of the number of candidates per event before and after applying the candidate selection are shown in Sec. 6.4.3.

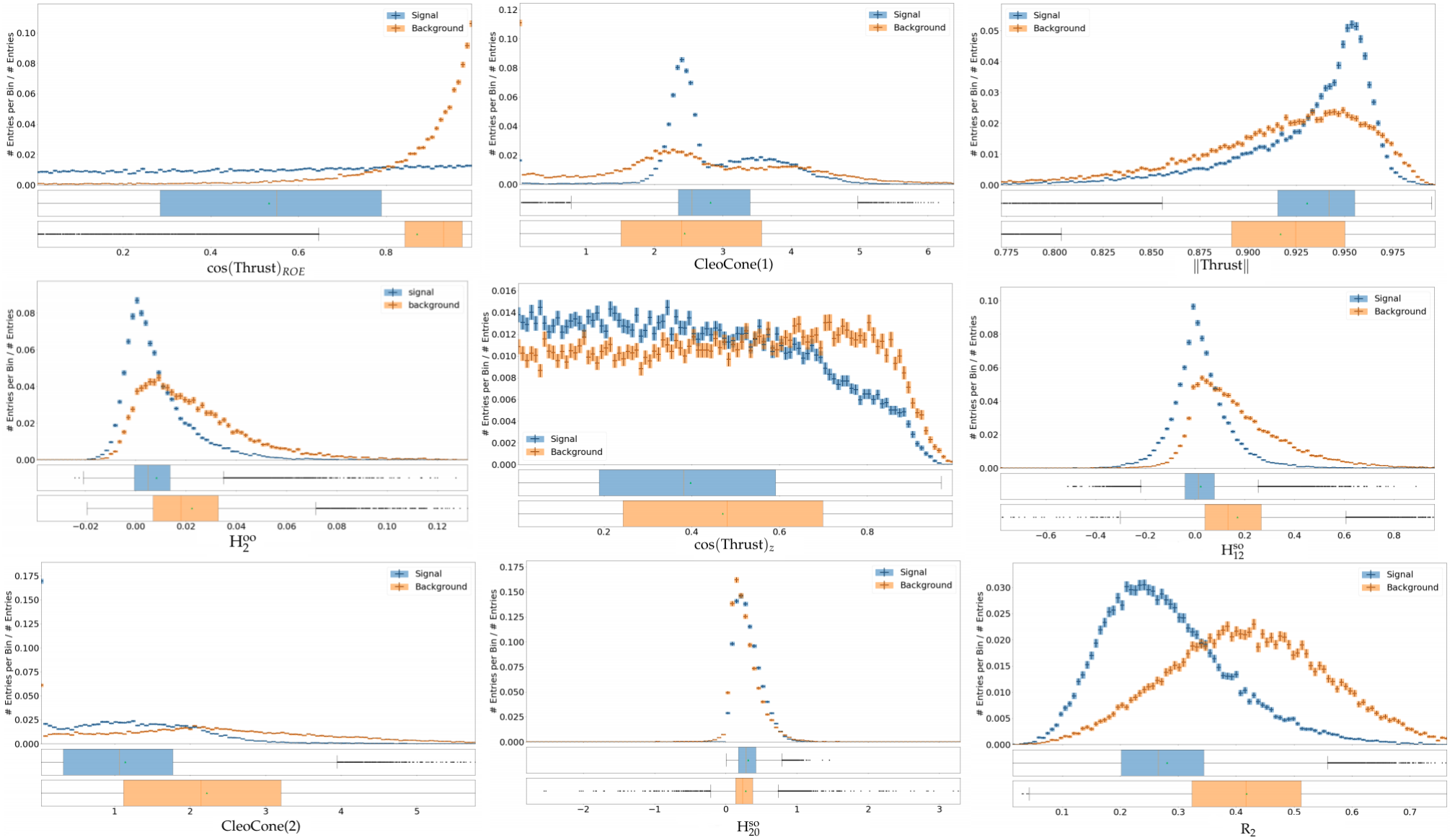

FIGURE 6.1: *Distributions of $B^0 \to K_S^0\pi^+\pi^-\gamma$ signal (in blue) and background (in orange) of the nine most important variables used in the CSMVA training.*

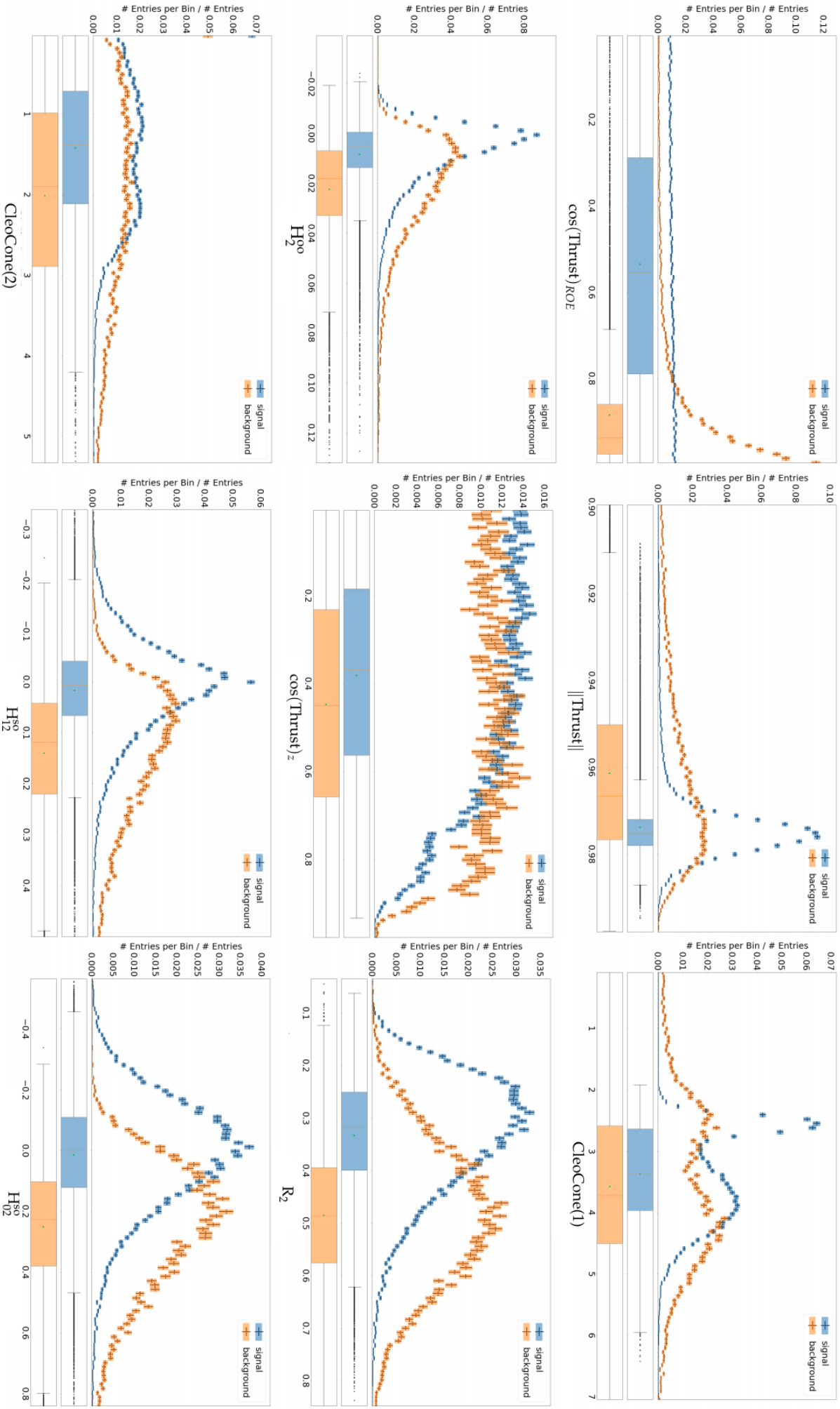

FIGURE 6.2: *Distributions of $B^0 \to K^0_S\pi^0\gamma$ signal (in blue) and background (in orange) of the nine most important variables used in the CSMVA training.*

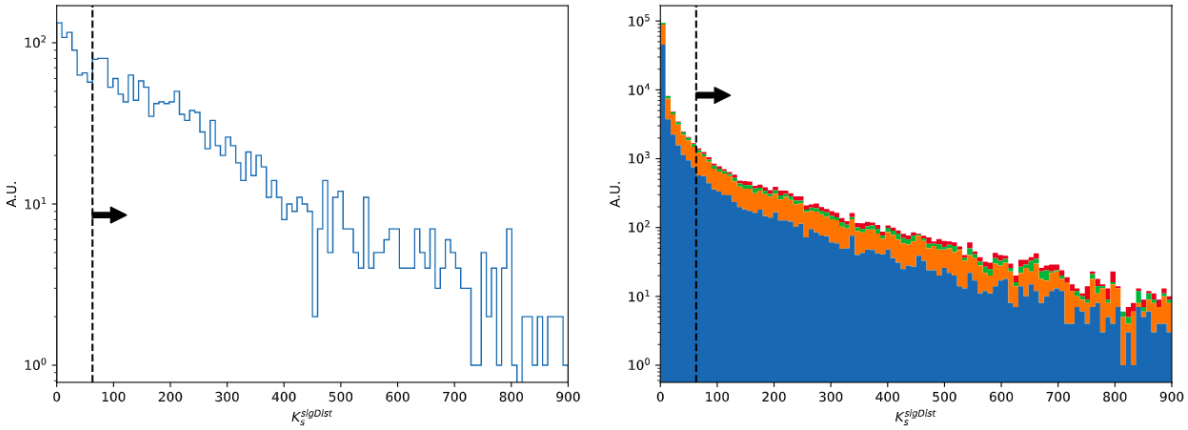

FIGURE 6.3: *K_S^0 decay length significance distribution with all contribution (right) and the signal only plot (left), with the cut corresponding to the highest FOM (see Figure 6.4) for $B^0 \to K_S^0\pi^+\pi^-\gamma$. Data on the right of the dashed line is kept. Blue is $q\bar{q}$, orange $c\bar{c}$ and green $B\bar{B}$.*

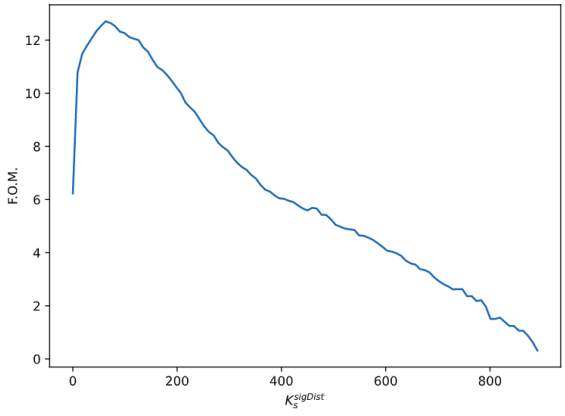

FIGURE 6.4: *Plot of the FOM as a function of the K_S^0 decay length significance for $B^0 \to K_S^0\pi^+\pi^-\gamma$. The cut on this variable corresponds to the maximal of this function.*

6.4 Selection results

Here we present the results of each step of the selection procedure.

6.4.1 Continuum suppression

The optimization result of the CSMVA with the fastBDT is shown in Figure 6.5 and Figure 6.6 for $K^0_S\pi^+\pi^-\gamma$ and $K^0_S\pi^0\gamma$. The output variable gives a good discrepancy between the background and the signal for both channels. Also, the testing sample, when comparing to the training one show that there is no over-training of the boosted decision tree.

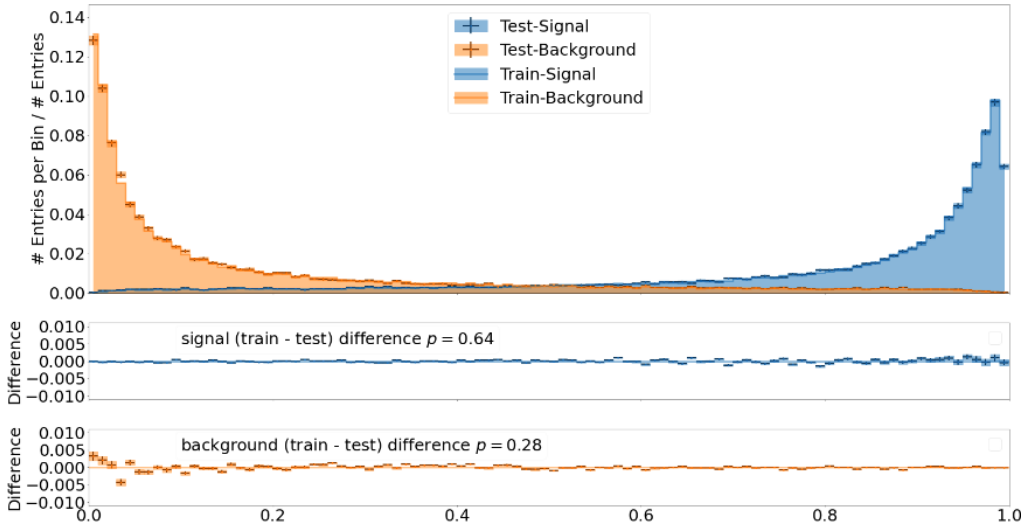

FIGURE 6.5: *Output of the CSMVA training for signal (in blue) and background (in orange) for $B^0 \to K^0_S\pi^+\pi^-\gamma$. The over training is checked using the difference plot between the test and train sample. The p-value comes from a Kolmogorov-Smirnov test.*

6.4.2 Cut optimization

The results tables of the optimization of the selection is shown in Tables 6.1 and 6.2 for $B^0 \to K^0_S\pi^+\pi^-\gamma$ and Tables 6.3 and 6.4 for $B^0 \to K^0_S\pi^0\gamma$.

For $B^0 \to K^0_S\pi^+\pi^-\gamma$, the efficiency of the four selection cuts combined is 34%. With the reconstruction and pre-selection, the final efficiency of the selection is 9.5%, with a FOM of 22.3. The distribution for each background and signal contribution before and after the selection is presented in Figure 6.7.

For $B^0 \to K^0_S\pi^0\gamma$, the efficiency of the four selection cuts combined is 34%. With the reconstruction and pre-selection, the final efficiency of the selection is 9.9%, with a FOM of 11.7. The distribution for each background and signal contribution before and after the selection is presented in Figure 6.8. In particular, the $m_{K^0_S\pi^0}$ cut value remove the $B^0 \to X_{sd}\gamma$ contribution from the background (see Figure 6.10 in Section 6.5), so the final ΔE fit will not have this peaking background.

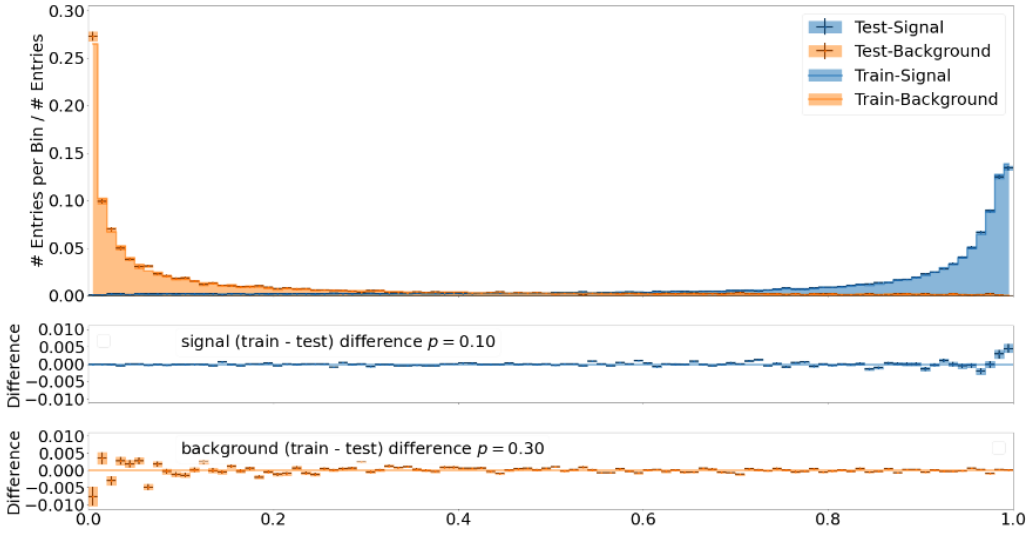

FIGURE 6.6: *Output of the CSMVA training for signal (in blue) and background (in orange) for $B^0 \to K^0_S \pi^0 \gamma$. The over training is checked using the difference plot between the test and train sample. The p-value comes from a Kolmogorov-Smirnov test.*

TABLE 6.1: Value of the cut and the evolution of the FOM for each variable after the optimization process for $B^0 \to K^0_S \pi^+ \pi^- \gamma$ in the signal region for 700 fb^{-1}.

Cut	cut value	FOM
$m_{K_S\pi\pi}$	$m_{K_S\pi\pi} < 1.8$ GeV/c^2	4.43 ± 0.09
m_{bc}	$m_{\text{bc}} > 5.275$ GeV/c^2	6.40 ± 0.13
$K_S^{sigDist}$	$K_S^{sigDist} > 63$	13.09 ± 0.30
CSMVA	CSMVA > 0.88	22.3 ± 0.5

TABLE 6.2: Signal efficiency, calculated on the 2M signal only sample in the whole region, and background retention, calculated on MC in the signal region, for each variable after the optimization process for $B^0 \to K^0_S \pi^+ \pi^- \gamma$ on 700 fb^{-1}. The efficiency of the single candidate selection is calculated in the whole region.

Cut	sgn eff [%]	bkg retention [%]
Reconstruction + pre-selection	27.867 ± 0.031	–
$m_{K_S\pi\pi}$	80.64 ± 0.05	19.990 ± 0.031
m_{bc}	86.90 ± 0.05	38.85 ± 0.08
$K_S^{sigDist}$	72.53 ± 0.07	11.74 ± 0.09
CSMVA	67.20 ± 0.09	9.91 ± 0.24
Combined selection	34.16 ± 0.06	0.452 ± 0.011
Single candidate selection	87.08 ± 0.08	82.5 ± 0.6
Total efficiency	8.288 ± 0.019	–

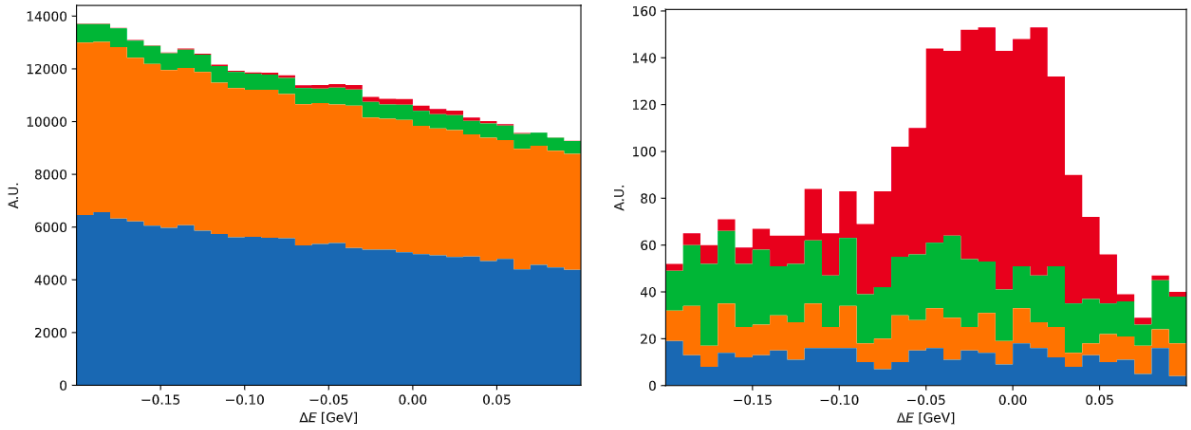

FIGURE 6.7: *ΔE distribution in the signal region before (left) and after (right) the selection optimization for $B^0 \to K^0_S\pi^+\pi^-\gamma$ on the 700 fb^{-1}. Blue is $q\bar{q}$, orange $c\bar{c}$, green $B\bar{B}$ and red is signal.*

TABLE 6.3: Value of the cut and the evolution of the FOM for each variable after the optimization process on MC for $B^0 \to K^0_S\pi^0\gamma$ in the signal region on 700 fb^{-1}.

Cut	cut value	FOM
$K_S^{sigDist}$	$K_S^{sigDist} > 36$	2.11 ± 0.07
m_{bc}	$m_{\text{bc}} > 5.275$ GeV/c^2	2.99 ± 0.11
CSMVA	CSMVA > 0.96	10.5 ± 0.5
$m_{K_S\pi^0}$	$m_{K_S\pi^0} < 1.1$ GeV/c^2	11.7 ± 0.5

TABLE 6.4: Signal efficiency, calculated on the 2M signal only sample in the whole region, and background retention, calculated on MC in the signal region, for each variable after the optimization process for $B^0 \to K^0_S\pi^0\gamma$ on 700 fb^{-1}. The efficiency of the single candidate selection is calculated in the whole region.

Cut	sgn eff [%]	bkg retention [%]
Reconstruction + pre-selection	28.865 ± 0.032	–
$K_S^{sigDist}$	88.33 ± 0.04	14.624 ± 0.035
m_{bc}	86.32 ± 0.05	39.51 ± 0.13
CSMVA	45.34 ± 0.08	1.14 ± 0.04
$m_{K_S\pi^0}$	99.190 ± 0.020	68.3 ± 1.8
Combined selection	34.29 ± 0.06	0.0450 ± 0.0021
Single candidate selection	87.74 ± 0.07	57.0 ± 1.2
Total efficiency	8.685 ± 0.020	–

6.4.3 Single candidate selection

The number of candidates per event after the selection is shown on Figure 6.9 for both $B^0 \to K^0_S\pi^+\pi^-\gamma$ and $B^0 \to K^0_S\pi^0\gamma$. Efficiencies on the signal are reported in Table 6.2 and Table 6.4, they reach 87% for the first channel and 88% for the second. Before the selection, a large amount of multiple candidates is present

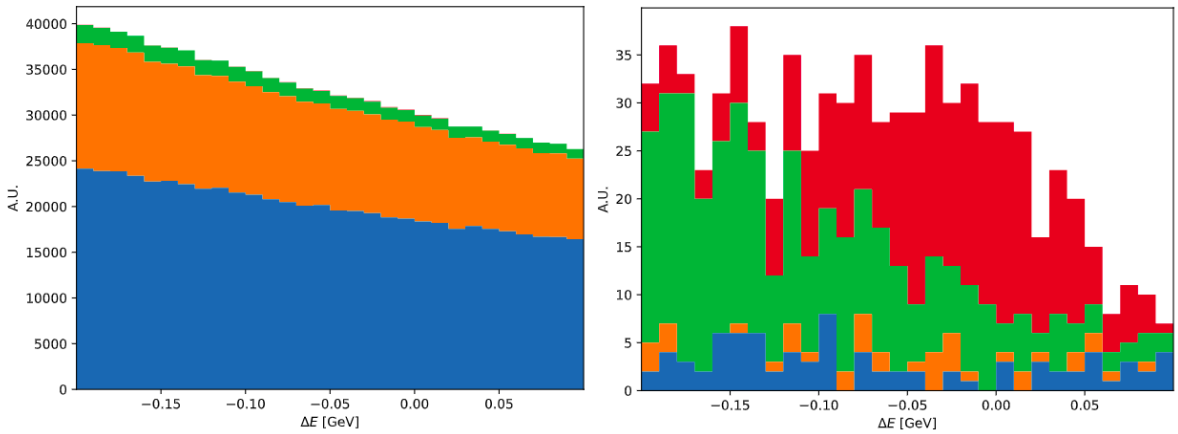

FIGURE 6.8: *ΔE distribution in the signal region before (left) and after (right) the selection optimization for $B^0 \to K^0_S\pi^0\gamma$ on the 700 fb^{-1}. Blue is $q\bar{q}$, orange $c\bar{c}$, green $B\bar{B}$ and red is signal.*

on both channels. This might comes from the fact that we take all pions to reconstruct $K^0_S\pi^+\pi^-\gamma$, and for $K^0_S\pi^0\gamma$ the photons from the π^0 might be miss-interpreted as the one from the final state.

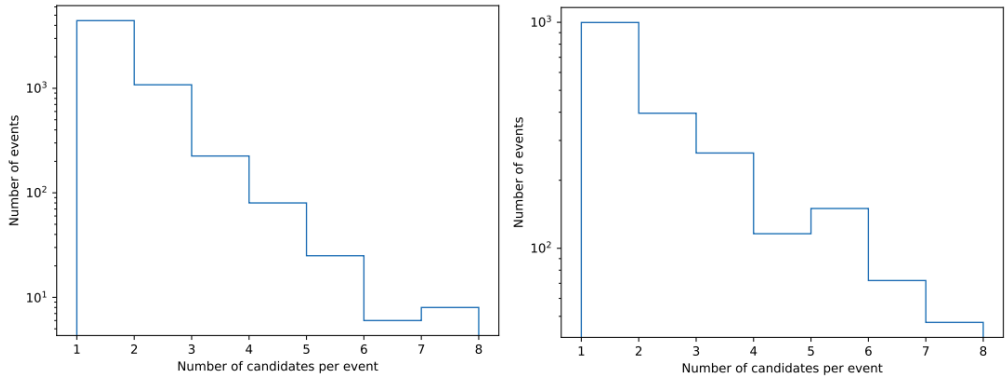

FIGURE 6.9: *Distribution of the number of candidates per events after applying the selection for $B^0 \to K^0_S\pi^+\pi^-\gamma$ (left) and $B^0 \to K^0_S\pi^0\gamma$ (right).*

6.5 Background contributions

As explained in the cut optimization Section 6.4.2, the $m_{K^0_S\pi^0} < 1.1$ GeV/c^2 cut for $B^0 \to K^0_S\pi^0\gamma$ decays correctly removes the major peaking background contribution, where K^0_S π^0 come from a X_{sd} resonance instead of a $K^{*0}(892)$. Distributions for both contributions are shown in Figure 6.10.

The ΔE final distributions of all backgrounds and signal contribution before the fit for both channels are shown on Figure 6.11. There is no significant peaking background in the signal region in both channels.

For $K^0_S\pi^+\pi^-\gamma$, the peak beyond $\Delta E > 0.1$ GeV from the B^+B^- contribution is coming from $B^\pm \to K^{*\pm}\gamma$ and cannot be removed by adding additional cuts.

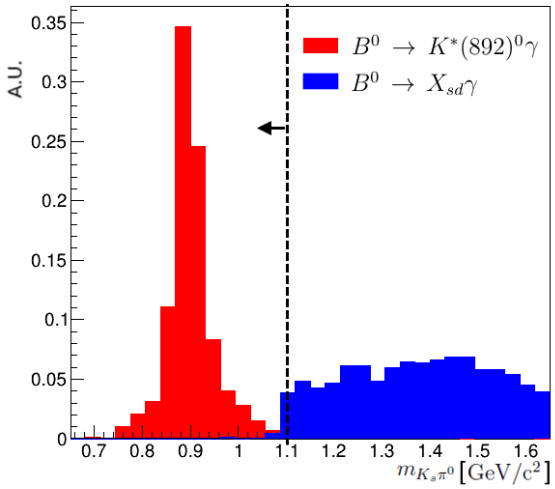

FIGURE 6.10: *Distribution of $m_{K^0_S\pi^0}$ for $B^0 \to K^0_S\pi^0\gamma$ before the selection, showing the contribution of $B^0 \to K^{*0}(892)\gamma$ in blue and $B^0 \to X_{sd}\gamma$ in red.*

For $K^0_S\pi^0\gamma$, the peaking background in $-0.3 < \Delta E < -0.2$ GeV range comes from quark fragmentation channel $B^\pm \to X^+_{su}\gamma$. Attempts to fit this contribution has been made (see Appendix C.4) but it was not kept as no statistical accuracy improvement of the fit was observed, there is not enough stats to handle it at 200 fb^{-1}.

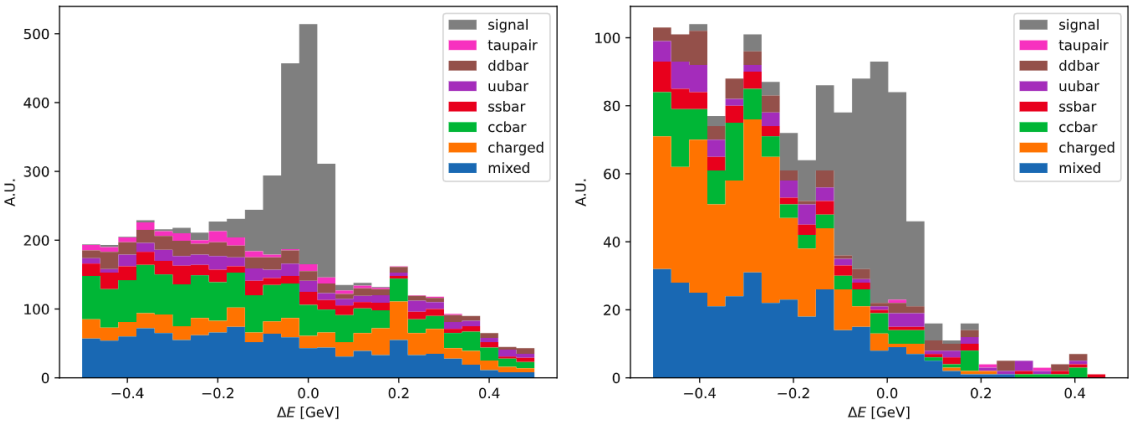

FIGURE 6.11: *Distributions for $B^0 \to K^0_S\pi^+\pi^-\gamma$ (left) and $B^0 \to K^0_S\pi^0\gamma$ (right) showing the contribution of each background types and the signal after the selection and the single candidate selection.*

6.6 Yield extraction

Yields are extracted from fits to ΔE distributions. First, templates for each contribution types are established on dedicated simulated samples. Then these templates are assembled into a PDF aiming to reproduce the complete ΔE distribution. In addition to the yield, some line shape parameters for the signal contribution are free in the final fit.

6.6.1 ΔE distribution templates

A Johnson function [103], which acts like a double-sided Crystal Ball (see Sec. 5.3.1), describes the signal contribution with four parameters, one for the central value and three for the shape.

The resulting Johnson PDF is:

$$\text{PDF}\left(x=\Delta E;\mu,\sigma,\gamma,\lambda\right)=\frac{\lambda}{\sigma\sqrt{2\pi}}\frac{1}{\sqrt{1+\left(\frac{x-\mu}{\sigma}\right)^2}}\exp\left[-\frac{1}{2}\left(\gamma+\lambda\sinh^{-1}\left(\frac{x-\mu}{\sigma}\right)\right)^2\right] \tag{6.2}$$

The fit results to candidates recognized as signal on the 700 fb^{-1} mixed-MC sample are shown in the left plots of Figures 6.12 and 6.13. Table 6.5 lists the fitted parameters.

For $K^0_S\pi^0\gamma$, all background contributions are fitted with a single Chebyshev polynomial of order two. For $B^0\to K^0_S\pi^+\pi^-\gamma$ the background is fitted using two Gaussian functions, one for the continuum and one to fit the charged peaking background in $\Delta E > 0.1$. The fit results to candidates recognized as not-signal for an equivalent of 700 fb^{-1} from the generic-MC sample are shown in the right plots of Figures 6.12 and 6.13. Table 6.5 lists the fitted parameters.

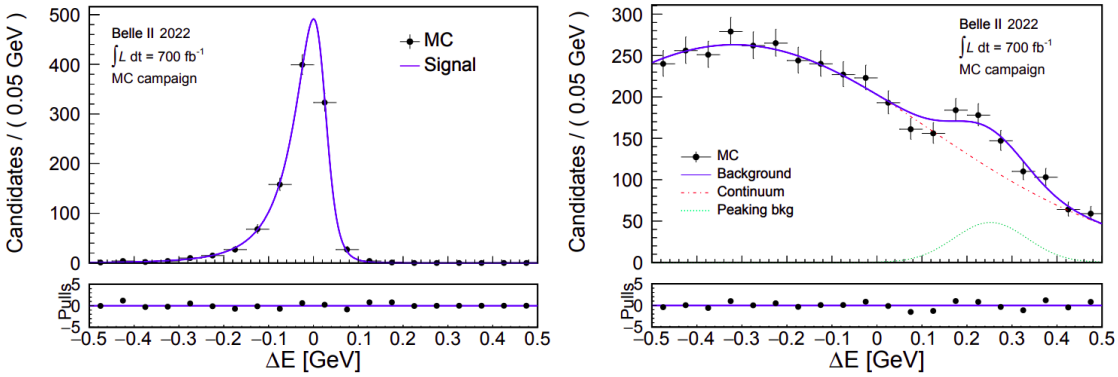

FIGURE 6.12: *Fit of the signal (left) and the background (right) on the simulated mixed-MC sample corresponding to 700 fb^{-1} for $B^0\to K^0_S\pi^+\pi^-\gamma$.*

TABLE 6.5: *Values of the parameters for the signal and background templates as fitted over the mixed-MC sample (700 fb^{-1}).*

$B^0\to K^0_S\pi^+\pi^-\gamma$				$B^0\to K^0_S\pi^0\gamma$			
Signal		Background		Signal		Background	
γ	1.10 ± 0.12	μ_{bkg}	-0.318 ± 0.030	γ	1.18 ± 0.30	a_0	-1.323 ± 0.019
λ	1.20 ± 0.08	σ_{bkg}	0.439 ± 0.033	λ	1.47 ± 0.24	a_1	0.398 ± 0.031
μ_{sgn}	0.024 ± 0.004	$\mu_{K*\gamma}$	0.252 ± 0.021	μ_{sgn}	0.036 ± 0.017		
σ_{sgn}	0.0401 ± 0.0035	$\sigma_{K*\gamma}$	0.081 ± 0.021	σ_{sgn}	0.080 ± 0.016		
		$f_{Bkg-K*\gamma}$	0.051 ± 0.019				

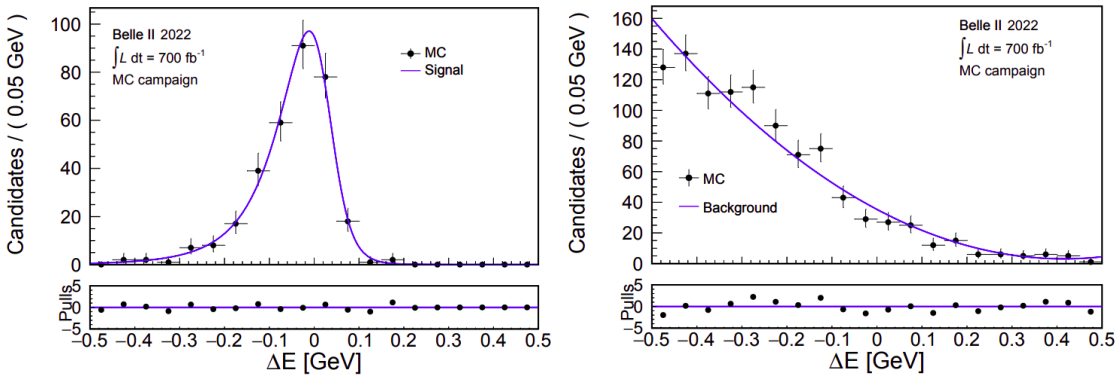

FIGURE 6.13: *Fit of the signal (left) and the background (right) on the simulated mixed-MC sample corresponding to 700 fb^{-1} for $B^0 \to K^0_S\pi^0\gamma$.*

6.6.2 ΔE final fit

The PDF for the final fit is built from the sum of the signal and background templates, from the previous Section 6.6.1, weighted by respectively the signal and background yields.

For both channel, the mean μ_{sgn} and the width σ_{sgn} are left free for the signal template, while λ and γ are fixed to the values from Table 6.5.

For the background templates, the parameters of the continuum contribution are fixed to the values fitted on the mixed-MC sample and listed in Table 6.5 for both channels. For the $K^*\gamma$ peaking background in $B^0 \to K^0_S\pi^+\pi^-\gamma$, the ratio between the continuum component and the $K^*\gamma$ one is let free. For the Gaussian function, the mean and the sigma are defined as:

$$\mu'_{K^*\gamma} = \mu_{sgn} + d\mu \tag{6.3}$$
$$\sigma'_{K^*\gamma} = \sigma_{sgn} \times \text{scale}, \tag{6.4}$$

with

$$d\mu = \mu_{K^*\gamma} - \mu_{sgn} \tag{6.5}$$
$$\text{scale} = \sigma_{K^*\gamma}/\sigma_{sgn}, \tag{6.6}$$

with the parameters corresponding to the values from Table 6.5. This way, by only letting the signal mean, if there is a shift in ΔE in data, the same shift will be applied on this peaking background. Same goes for a different resolution of the peak in data compared to MC, the scaling factor for the sigma will correct the shape.

The final fit is exercised over a part of the generic-MC sample equivalent to 200 fb^{-1}, corresponding to approximately the size of the data currently available. This sample of 200 fb^{-1} is disjoint from the sample of 700 fb^{-1} used to fix the

templates. The fit results are displayed in Figures 6.14 and 6.15 respectively for the $B^0 \to K^0_S \pi^+ \pi^- \gamma$ and $B^0 \to K^0_S \pi^0 \gamma$ channels.

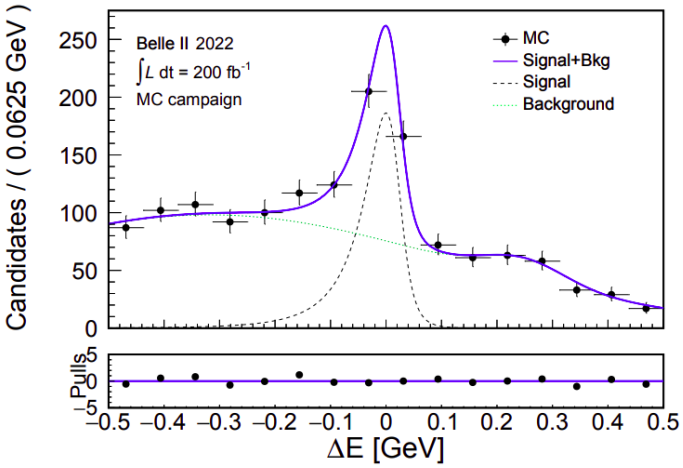

FIGURE 6.14: *Fit of signal + background on 200 fb^{-1} for $B^0 \to K^0_S \pi^+ \pi^- \gamma$. The yields, mean and sigma of the signal function and the continuum-$K^*\gamma$ ratio are let free.*

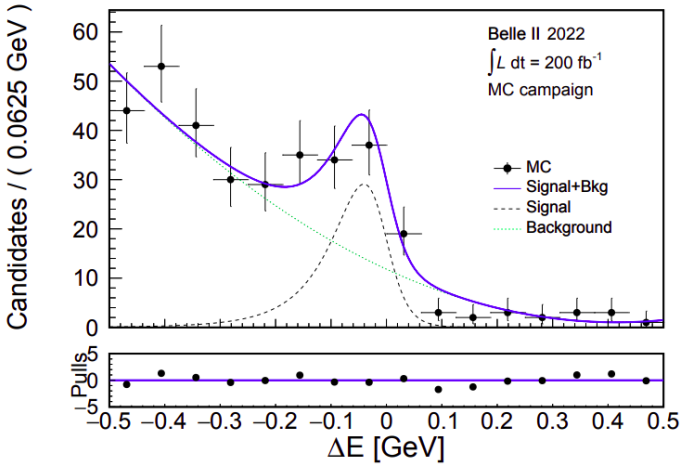

FIGURE 6.15: *Fit of signal + background on 200 fb^{-1} for $B^0 \to K^0_S \pi^0 \gamma$. The yields, mean and sigma of the signal function are left free.*

Table 6.6 and 6.7 summarize the fitted values of the different parameters and the raw yields estimated through the fit procedure compared to the Monte Carlo truth. The error estimation is provided by HESSE [104]. No statistical incompatibility is observed.

6.6.3 Fit robustness

ToyMC procedure

To estimate the fit robustness, we perform a Toy Monte-Carlo study of the fit on both channels. 500 data sets, all corresponding to integrated luminosity of

TABLE 6.6: *Fitted values of μ and σ on a generic-MC sample equivalent to 200 fb^{-1}.*

$B^0 \to K^0_S\pi^+\pi^-\gamma$		$B^0 \to K^0_S\pi^0\gamma$	
μ_{sgn}	0.022 ± 0.004	μ_{sgn}	-0.000 ± 0.011
σ_{sgn}	0.037 ± 0.005	σ_{sgn}	0.068 ± 0.016
$f_{Bkg-K*\gamma}$	0.048 ± 0.016		

TABLE 6.7: *Results of the yield extraction from the overall fit on generic-MC sample equivalent to 200 fb^{-1}.*

	$B^0 \to K^0_S\pi^+\pi^-\gamma$		$B^0 \to K^0_S\pi^0\gamma$	
	Signal	Background	Signal	Background
MC truth yields	316	1117	75	264
fit yields	292 ± 31	1140 ± 40	67 ± 15	272 ± 21

200 fb^{-1}, are created by bootstrapping the initial 700 fb^{-1} MC sample. The total number of events for each data set are determined from a Poisson distribution which average is set to the true number of events expected for a 200 fb^{-1} data set. Then the final fit procedure described in Sec. 6.6.2 is performed for each data set. The final values of the free parameters obtained in each fit are stored.

Finally, the differences and the pulls for the signal yield are calculated following Eq. 6.7 and Eq. 6.8. The true signal and background yields considered ($n_{\text{sgn true}}$ and $n_{\text{bkg true}}$) are the expected ones and are common to all toy data sets to correctly take into account the Poissonian fluctuation in the estimation of the statistical uncertainties. The same equations are used for the background yield.

$$n_{\text{sgn diffs}} = n_{\text{sgn fit}} - n_{\text{sgn true}} \tag{6.7}$$

$$n_{\text{sgn pulls}} = \frac{n_{\text{sgn diffs}}}{n_{\text{sgn fit-error}}} \tag{6.8}$$

Additionally, the evolution of the fit parameters are studied with varying true signal yields $n_{\text{sgn true}}$ while keeping the true background $n_{\text{bkg true}}$ constant at the level expected for 200 fb^{-1}.

Results

Pull and the difference distributions are shown for the $B^0 \to K^0_S\pi^+\pi^-\gamma$ channel on Fig. 6.16 and Fig. 6.17, and for the $B^0 \to K^0_S\pi^0\gamma$ channel on Fig. 6.18 and Fig. 6.19. We observe a slight over-estimation of the signal yield in both cases.

The dependence with $n_{\text{sgn true}}$ is shown on Fig. 6.20 and Fig. 6.21. We decide to not correct for this small effect but rather treat the average difference as a source of systematic uncertainty. For $B^0 \to K^0_S\pi^+\pi^-\gamma$, the relative uncertainty reaches 2.7% and for $B^0 \to K^0_S\pi^0\gamma$, it amounts to 11.5%.

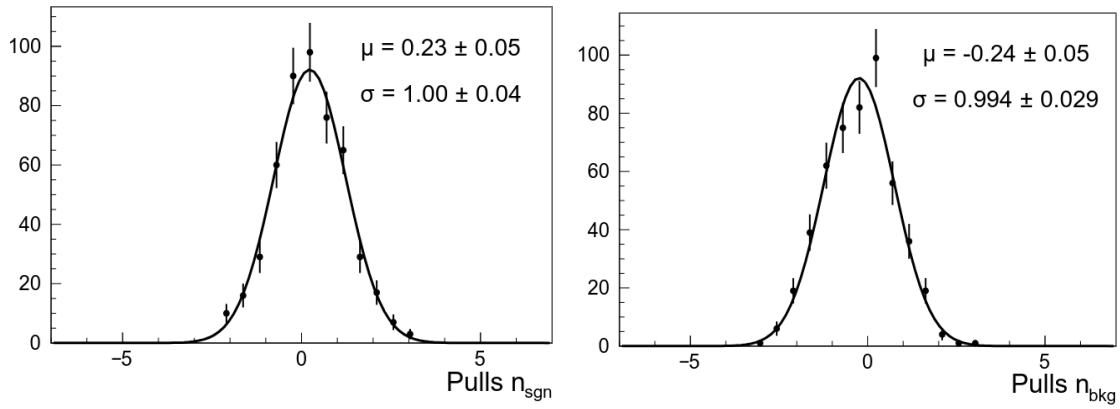

FIGURE 6.16: *Pulls (see Eq. 6.8) for the signal (left) and background (right) yields for* $B^0 \to K^0_S \pi^+ \pi^- \gamma$.

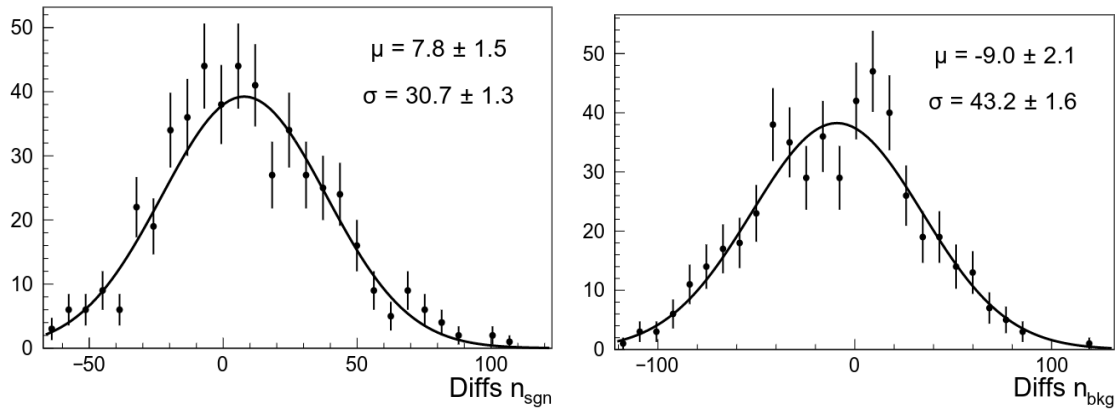

FIGURE 6.17: *Differences (see Eq. 6.7) for the signal (left) and background (right) yields for* $B^0 \to K^0_S \pi^+ \pi^- \gamma$.

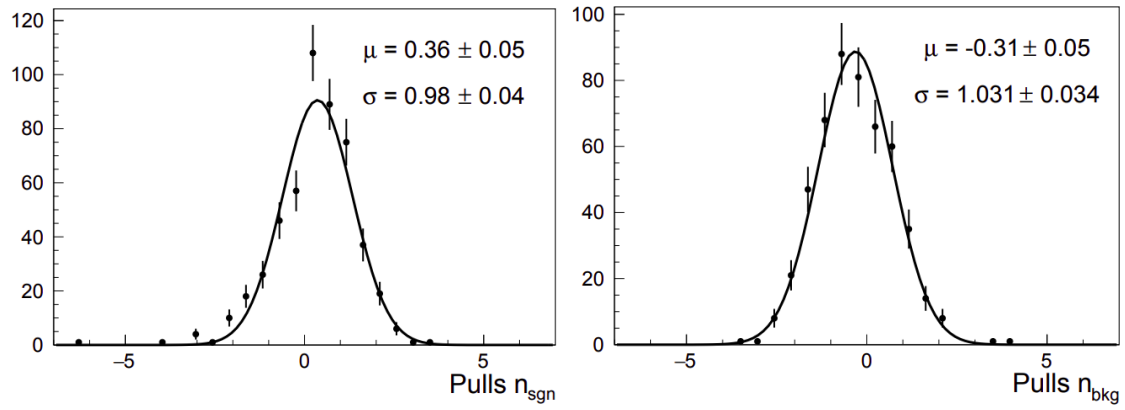

FIGURE 6.18: *Pulls (see Eq. 6.8) for the signal (left) and background (right) yields for* $B^0 \to K^0_S \pi^0 \gamma$.

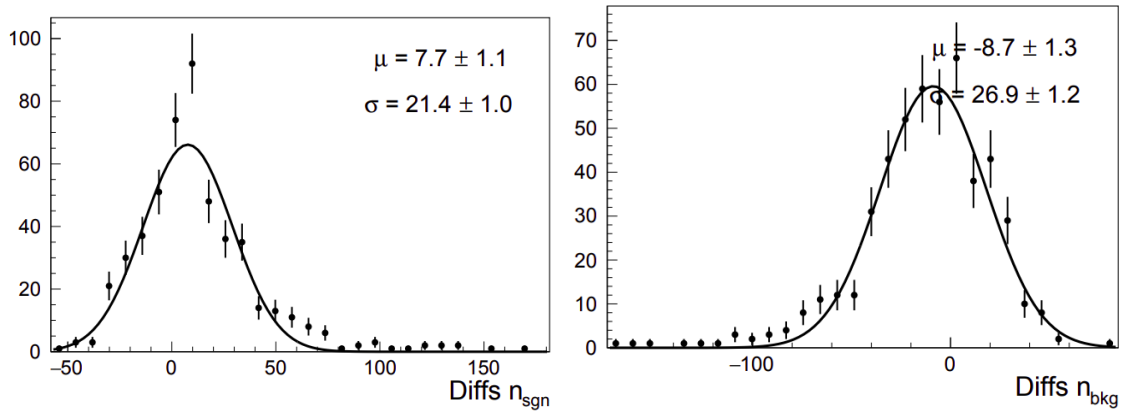

FIGURE 6.19: *Differences (see Eq. 6.7) for the signal (left) and background (right) yields for* $B^0 \to K^0_S\pi^0\gamma$.

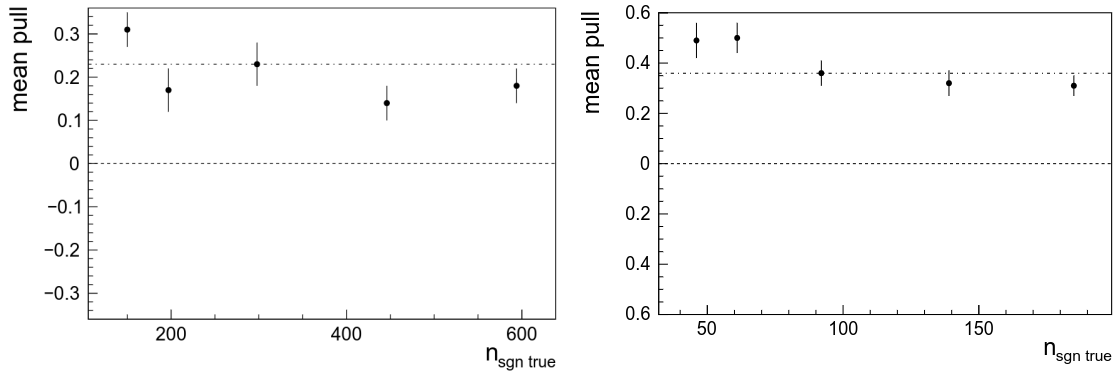

FIGURE 6.20: *Mean of the pull as a function of the* $n_{sgn\ true}$ *distribution for* $B^0 \to K^0_S\pi^+\pi^-\gamma$ *(left) and* $B^0 \to K^0_S\pi^0\gamma$ *(right). The middle point correspond to 200* fb^{-1} *of signal. For all points,* $n_{bkg\ true}$ *correspond to 200* fb^{-1}*. The dot-dashed line corresponds to the correction used for the systematic.*

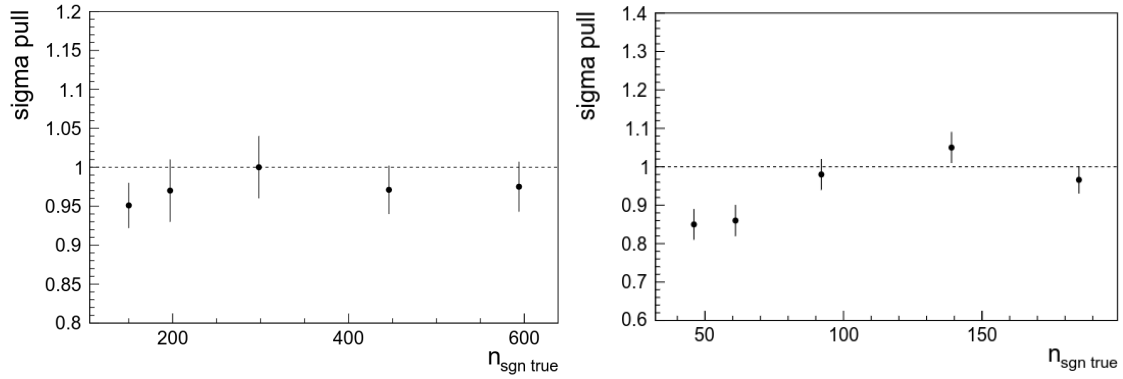

FIGURE 6.21: *Sigma of the pull as a function of the* $n_{sgn\ true}$ *distribution for* $B^0 \to K^0_S\pi^+\pi^-\gamma$ *(left) and* $B^0 \to K^0_S\pi^0\gamma$ *(right). The middle point correspond to 200* fb^{-1} *of signal. For all points,* $n_{bkg\ true}$ *correspond to 200* fb^{-1}*.*

6.7 Systematic uncertainties

The systematic uncertainties arising from the generation, reconstruction, selection, fit and yield extraction procedures are listed and quantified below. All relative systematic uncertainty values are summarized in Table 6.9.

MC sample size

A systematic uncertainty of 0.2% is assigned due to the finite sample size of the MC signal sample.

MC generation

The $B^0 \to K^0_S\pi^+\pi^-\gamma$ signal is composed of different components that have different efficiencies, it is thus important that the relative contributions of the different component is well simulated. To account for a miss modeling of the resonances, which can lead to a different efficiency in the $m_{K^0_S\pi^+\pi^-}$ cut, we assign a systematic by varying each branching fractions inside its uncertainties between one K_{res} and all the other, which gives:

$$\sum_{K_{\text{res}}} \left(\sigma \times \left| \frac{\tilde{\Delta}\epsilon}{1+\tilde{\Delta}\epsilon} \right| \right), \tag{6.9}$$

where

$$\tilde{\Delta}\epsilon = f_{K_{\text{res}}} \times \frac{(\epsilon_{K_{\text{res}}} - \epsilon_{\text{all other res}})}{\epsilon_{\text{all other res}}}, \tag{6.10}$$

$$\sigma = \frac{\sigma_{\mathcal{B}}}{\mathcal{B}(K_{\text{res}})}, \tag{6.11}$$

with $\mathcal{B}(K_{\text{res}})$ and $\sigma_{\mathcal{B}}$ the branching fraction of the K_{res} and its uncertainty, $\epsilon_{\text{all other res}}$ and $\epsilon_{K_{\text{res}}}$ are the efficiencies of the $m_{K^0_S\pi^+\pi^-}$ cut on all the other resonances and on the K_{res} population respectively, and $f_{K_{\text{res}}}$ is the fraction of the K_{res} population. The detailed of the calculation can be found Appendix C.5. Table 6.8 details the different components considered together with the branching fraction assumed in the simulation, the most up-to-date branching fraction derived assuming isospin symmetry and the individual channel efficiency. A correction factor of 1.13 for the efficiency is derived to correct for the wrong branching fraction used in the simulation. The associated systematic relative uncertainty of 4.2% is obtained by summing the difference in efficiency obtained by varying each branching fractions inside its uncertainties. The choice to make a simple sum instead of a sum in quadrature is to consider the most conservative case with 100% correlation between the variations due to the various components.

For $B^0 \to K^0_S\pi^0\gamma$, the K^* component is dominant in our selection. We checked the efficiencies of the $M(K^0_S\pi^0)$ restriction after the pre-selection and reconstruction and after the whole selection except for this criterion. We assign a relative systematic uncertainty of 2.0% to the efficiency.

TABLE 6.8: *Branching fractions of the different components of the $B^0 \to K^0_S\pi^+\pi^-\gamma$ signal from the PDG '21 [23] assuming isospin symmetry versus the ones used in the signal MC and total efficiencies excluding the single candidate selection.*

Component	BF PDG '21 [23] ($\times 10^{-5}$)	BF in basf2 ($\times 10^{-5}$)	Efficiency [%]
$K_1(1270)^0$	$4.4^{+0.7}_{-0.6}$	4.3	12.4 ± 0.04
$K_1(1400)^0$	10^{+5}_{-4}	0.65	11.43 ± 0.09
$K_2(1430)^0$	1.28 ± 0.17	1.237	10.80 ± 0.12
$K^*(1680)^0$	$6.7^{+1.7}_{-1.4}$	0.17	6.04 ± 0.19
X_{sd}	21.5 ± 1.2	27.4	7.94 ± 0.02

Pion PID

Studies have been done on $D^{*+} \to D^0[K^-\pi^+]\pi^+$ decays to measure the pion identification performances in data. The results are, for different values of pion PID, 2D tables in 7 momentum bins and 8 polar angle bins, and for each bin the ratio of the efficiencies in data and MC is calculated. For each final state pions with a PID cut, we realize the 2D tables with the same axis and bins configuration, and the systematic uncertainty is then calculated with:

$$\frac{\epsilon_{DATA}}{\epsilon_{MC}} = \epsilon \prod_i \Gamma_i, \tag{6.12}$$

where

$$\epsilon = \frac{N_{\text{after PID cut}}}{N_{\text{before PID cut}}}, \tag{6.13}$$

$$\Gamma_i = \frac{\sum_j (n_j c_j)}{\sum_j n_j}, \tag{6.14}$$

with $N_{\text{tot before PID cut}}$ and $N_{\text{after PID cut}}$ is the total number of events before and after all the PID respectively, n_j the number of event in the i-th bin of our table, and c_j the ratio of the efficiencies in data and MC in the i-th bin.

For the global PID $\mathcal{P}(\pi) > 0.1$ for the two final state pions in $B^0 \to K^0_S\pi^+\pi^-\gamma$, we get $\frac{\epsilon_{DATA}}{\epsilon_{MC}} = 0.9158 \pm 0.0020$. The correction is applied on the selection efficiency and the corresponding relative systematic uncertainty is $0.0020/0.9158 = 0.2\%$.

Tracking

The track reconstruction efficiency in data was measured using τ-pair events, in which one tau lepton decays leptonically while the other decays hadronically into three charged pions. The result is that for each track in the final state, a relative systematic uncertainty of 0.69% is applied. It gives a global value of 1.38% for $B^0 \to K^0_S\pi^+\pi^-\gamma$. The systematic uncertainty associated to the reconstruction of K^0_S is computed separately.

π^0 reconstruction

For the π^0 selection, the difference in reconstruction efficiency between data and simulation is estimated by comparing the $\eta \to \gamma\gamma$ and $\eta \to \pi^0\pi^0\pi^0$ yields. For $B^0 \to K^0_S\pi^0\gamma$, it gives a systematic uncertainty of $\frac{\epsilon_{DATA}}{\epsilon_{MC}} = 0.945 \pm 0.004 \pm 0.041$, which is compatible with 1. Hence, we do not correct the efficiency and take as a relative uncertainty $|1 - 0.945| = 5.5\%$.

K^0_S reconstruction

The uncertainty on the K^0_S reconstruction comes from the study on displaced vertexes on $B^+ \to \phi[K^{*+} \to K^0_S\pi^+]$ and $B^0 \to \pi^+[\overline{D}^0 \to K^0_S\pi^-]$. The result is that we assign a relative uncertainty computed as average 3D distance (in cm) $\times$ 0.413%. The average K^0_S decay distance after selections are 8.71 cm and 8.37 cm for $B^0 \to K^0_S\pi^+\pi^-\gamma$ and $B^0 \to K^0_S\pi^0\gamma$, which give the corresponding relative uncertainties of 3.60% and 3.46% respectively and are taken both as systematic and corrections factors.

π^0 veto

The same idea as pion PID is performed to get the π^0 veto systematic uncertainty. For a given $\mathcal{P}^0_\pi$ cut, the ration of efficiency of that veto between real and simulated is evaluated on $B^0 \to D^-\pi^+$ and $B^+ \to \overline{D}^0\pi^+$ samples, in bins of the energy of the photon. By doing the same histogram for our sample, we multiply each bin by the data/MC ratio to get the Γ from Eq. 6.12. The ϵ is the ratio between the number of events after and before the π^0 veto cut.

We get $\frac{\epsilon_{DATA}}{\epsilon_{MC}} = 1.036 \pm 0.018$ for $B^0 \to K^0_S\pi^+\pi^-\gamma$ and $\frac{\epsilon_{DATA}}{\epsilon_{MC}} = 1.034 \pm 0.020$ for $B^0 \to K^0_S\pi^0\gamma$. The correction is applied on the selection efficiency and the corresponding relative systematic uncertainty is 1.7% and 1.9% for $B^0 \to K^0_S\pi^+\pi^-\gamma$ and $B^0 \to K^0_S\pi^0\gamma$ respectively.

Photon selection

According to the measurement of the data to MC ratio of photon reconstruction efficiency in the calorimeter, using radiative muon pair events, we get $\frac{\epsilon_{DATA}}{\epsilon_{MC}} = 1.0026^{+0.0015+0.0007}_{-0.0041-0.0007}$. The efficiency is not corrected and the relative uncertainty is taken as 0.3%.

Continuum suppression

We looked at the off-resonance sample, recorded 60 MeV below $\Upsilon(4S)$ resonance, and the continuum events in the MC sample before and after the CSMVA cut to extract a systematic on the efficiency. We get a background rejection fraction of $\frac{r_{DATA}}{r_{MC}} = 0.969 \pm 0.009$ for $B^0 \to K^0_S\pi^+\pi^-\gamma$ and $\frac{r_{DATA}}{r_{MC}} = 1.028 \pm 0.029$ for $B^0 \to K^0_S\pi^0\gamma$. Since this study concerns the background, we decided to assign

a conservative 3% relative systematic uncertainty on the signal efficiency for the BDT selection.

Fit bias

As discussed in Sec. 6.6.3, we observed a small overestimation of the signal yield by the final fit procedure. It is taken into account with relative uncertainties of 2.7% for the $B^0 \to K^0_S\pi^+\pi^-\gamma$ channel and 11.5% for the $B^0 \to K^0_S\pi^0\gamma$ channel.

Number of $B\bar{B}$ pairs

For the data sample we have, $N_{B\overline{B}} = (197.2 \pm 5.7(syst.)) \times 10^6$, the systematic uncertainty was estimated at 2.9%.

f^{00} systematic

A systematic uncertainty of 1.2% is assigned due to the uncertainty in the branching fraction of $\Upsilon(4S)$ decaying to $B\bar{B}$ from Ref. [23].

TABLE 6.9: *List of all the systematic relative uncertainties affected the branching ratio estimation for $B^0 \to K^0_S\pi^+\pi^-\gamma$ (left) and $B^0 \to K^0_S\pi^0\gamma$ (right). The total efficiency systematic corresponds to the quadratic sum of all the systematic uncertainties for the selection efficiency.*

Efficiency systematics	$B^0 \to K^0_S\pi^+\pi^-\gamma$	$B^0 \to K^0_S\pi^0\gamma$
MC sample size (stat error)	0.2%	
MC generation	4.2%	2.0%
pion PID	0.2%	-
Tracking	1.38%	-
π^0 reconstruction	-	5.5%
K^0_S reconstruction	3.60%	3.46%
π^0 veto	1.7%	1.9%
γ selection	0.3%	
Continuum suppression	3.0%	
Total efficiency	6.67%	7.68%

Yields systematic	$B^0 \to K^0_S\pi^+\pi^-\gamma$	$B^0 \to K^0_S\pi^0\gamma$
Fit bias	2.7%	11.5%

Number of $B^0\overline{B}^0$ pairs syst	2.9%

f^{00} systematic	1.2%

	$B^0 \to K^0_S\pi^+\pi^-\gamma$	$B^0 \to K^0_S\pi^0\gamma$
Total	7.86%	14.2%

6.8 Measurements of Branching Fractions

The branching fraction is computed according to:

$$\mathcal{B} = \frac{N}{\epsilon 2 f^{00} N_{B\overline{B}}}, \tag{6.15}$$

where N is the yield of a given decay channel and extracted from the procedure described in Section 6.6.2, ϵ is the overall efficiency to reconstruct the decay, $f^{00} = (48.6 \pm 0.6)\%$ [23] the ratio of $\Upsilon(4S)$ going to $B^0\overline{B}^0$ and $N_{B\overline{B}} = (197.2 \pm 5.7(syst.)) \times 10^6$ is the estimated number of $B\overline{B}$ pairs produced in the data sample listed in Section 6.1.

Overall efficiencies are taken from Tables 6.2 and 6.4 with the corrections from Sec. 6.7. They amount to:

$$\epsilon(B^0 \to K^0_S\pi^+\pi^-\gamma) = (8.566 \pm 0.571)\% \tag{6.16}$$
$$\epsilon(B^0 \to K^0_S\pi^0\gamma) = (8.669 \pm 0.666)\% \tag{6.17}$$

6.9 Data unblinding

In Belle II, all official measurements are done following an *unblinding* process. Two steps are required to make sure that the analysis done on MC is under control. First we compare various distributions, mainly variables used in the selection and the fit, between data and MC in a kinematical region that does not contain our signal, to see if everything agrees and to assign systematic uncertainties if not. This region is the named data sideband in the following. Then, we perform the same comparison including the signal region, this is called *unblinding*. If everything is under control, we can then proceed to the fitting of ΔE and the BR measurement.

6.9.1 Data-MC sideband check

To proceed to a Data-MC comparison before unblinding, the data were processed with the strategy described in Section 6.2. In addition, a cut on $m_{\text{bc}} < 5.27$ GeV/c^2 is applied to look only at the sideband regions and keep blinding the signal region.

The Data-MC sideband comparisons for variables used in the selection are shown in Fig. 6.22, Fig. 6.23 and Fig. 6.24 for $B^0 \to K^0_S\pi^+\pi^-\gamma$ and Fig. 6.25, Fig. 6.26 and Fig. 6.27 for $B^0 \to K^0_S\pi^0\gamma$.

Fair agreements are observed for the K^0_S decay distance significance. A discrepancy is observed for the invariant mass $m_{K^0_S\pi^+\pi^-}$, but it disappears when we compare the distributions after applying all other selection cuts, as shown in Fig. 6.28. We do not thus apply any correction. The same discrepancy is observed for $m_{K^0_S\pi^0}$, but is not visible in the off-resonance[2] plot (Fig. C.30) and in the plot

[2] Around 18 fb^{-1} of data recorded 60 MeV below $\Upsilon(4S)$ resonance.

after applying all the selection (Fig. 6.28). The CSMVA discrepancy are covered by the systematic defined Sec. 6.7. For distributions related to the π^0 veto and the photon energy E_γ, discrepancies are expected and systematic uncertainties provided by the different collaboration studies described Sec. 6.7 are used to take them into account.

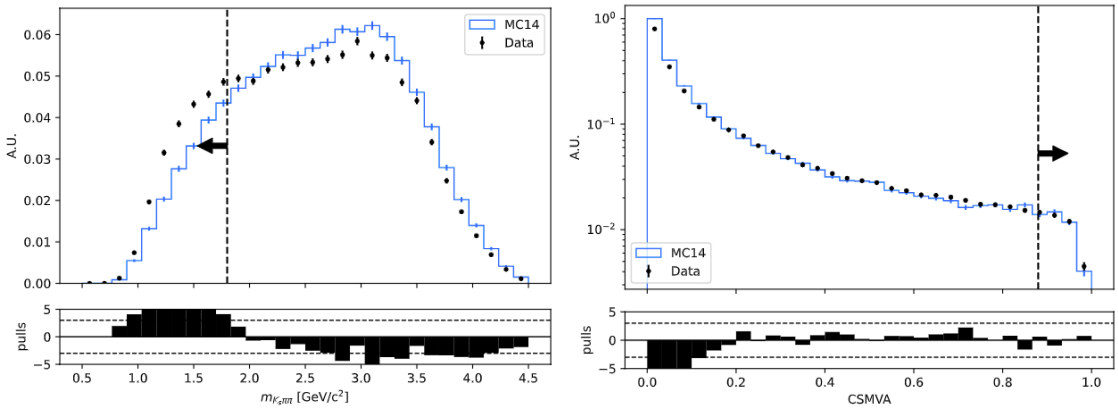

FIGURE 6.22: *Normed $m_{K^0_S\pi^+\pi^-}$ (left) and CSMVA output (right) distributions for MC and data sidebands for $B^0 \to K^0_S\pi^+\pi^-\gamma$. The bar indicates the value of the cut and the arrow is oriented to the kept part.*

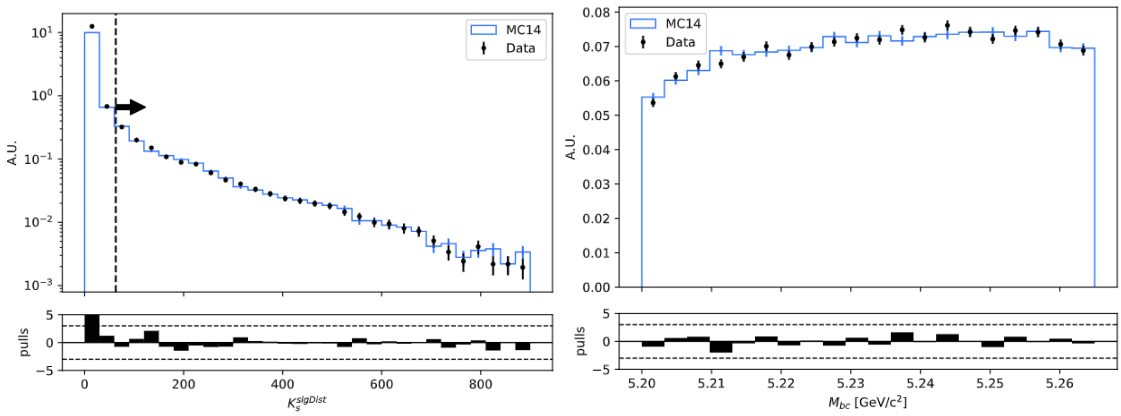

FIGURE 6.23: *Normed K^0_S significance of distance (left) and m_{bc} (right) distributions for MC and data sidebands for $B^0 \to K^0_S\pi^+\pi^-\gamma$. The bar indicates the value of the cut and the arrow is oriented to the kept part.*

The ΔE distribution for both channel is shown Fig. 6.29. No discrepancies are observed.

Finally, the last check consists of applying all the cut from the selection except on m_{bc} (but still keep $m_{\mathrm{bc}} < 5.27\,\mathrm{GeV}/c^2$) and look at the ΔE sideband distribution, shown in Fig. 6.30. The two distributions are compatible for both channels.

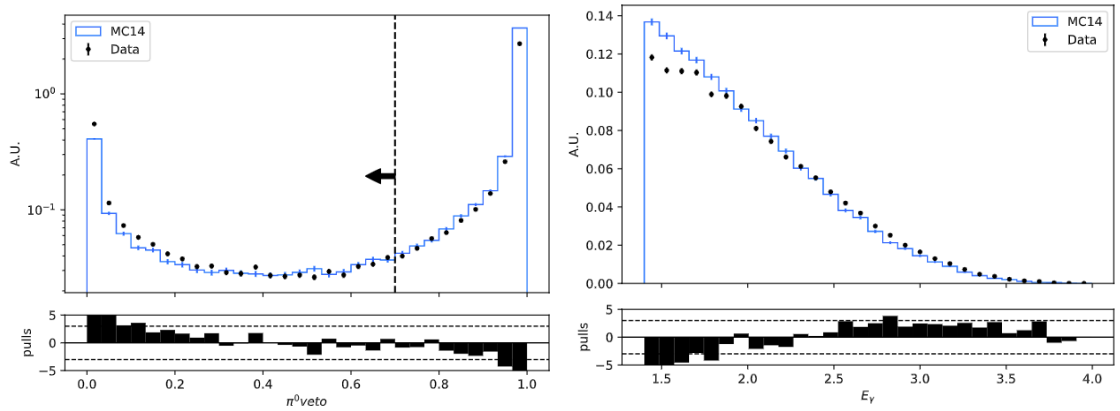

FIGURE 6.24: *Normed π^0 veto variable (left) and E_γ (right) distributions for MC and data sidebands for $B^0 \to K^0_S\pi^+\pi^-\gamma$. The bar indicates the value of the cut and the arrow is oriented to the kept part.*

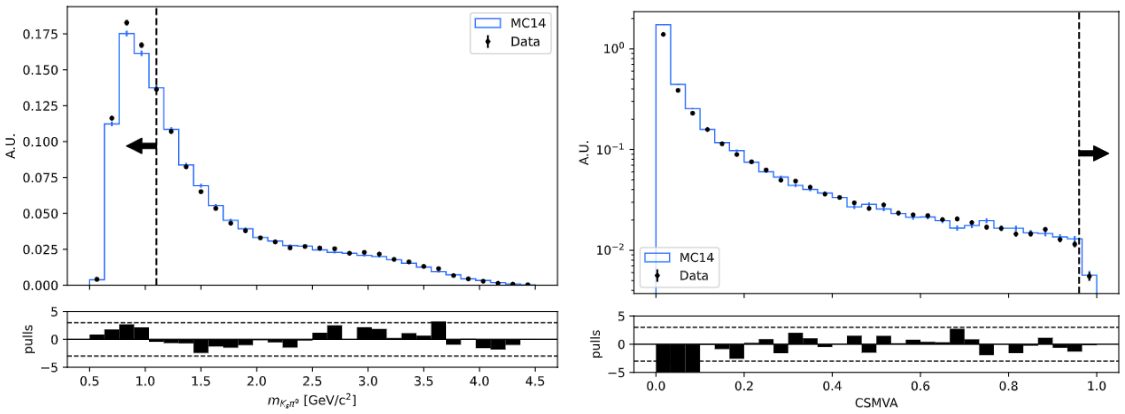

FIGURE 6.25: *Normed $m_{K^0_S\pi}$ (left) and CSMVA output (right) distributions for MC and data sidebands for $B^0 \to K^0_S\pi^0\gamma$. The bar indicates the value of the cut and the arrow is oriented to the kept part.*

Off-resonance data and sideband comparison with MC is shown in Appendix C.6. There are no differences in the plots compared to the study in this section.

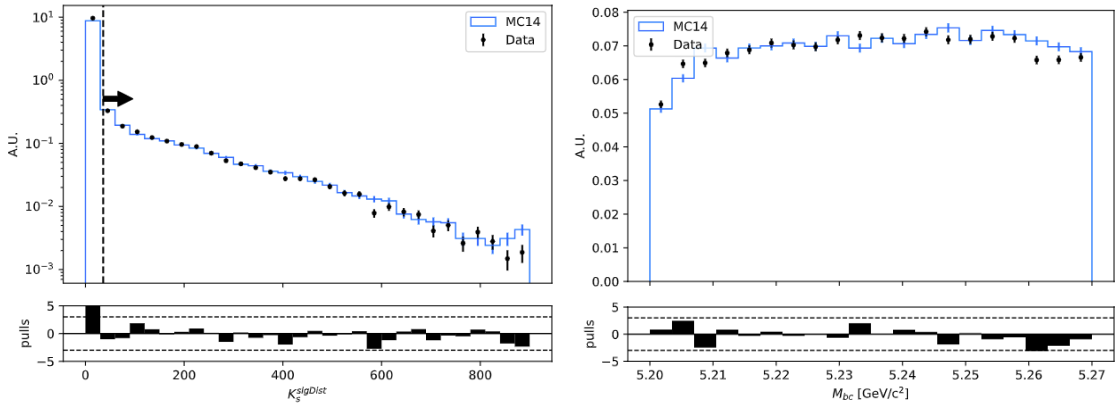

FIGURE 6.26: *Normed K_S^0 significance of distance (left) and m_{bc} (right) distributions for MC and data sidebands for $B^0 \to K_S^0\pi^0\gamma$. The bar indicates the value of the cut and the arrow is oriented to the kept part.*

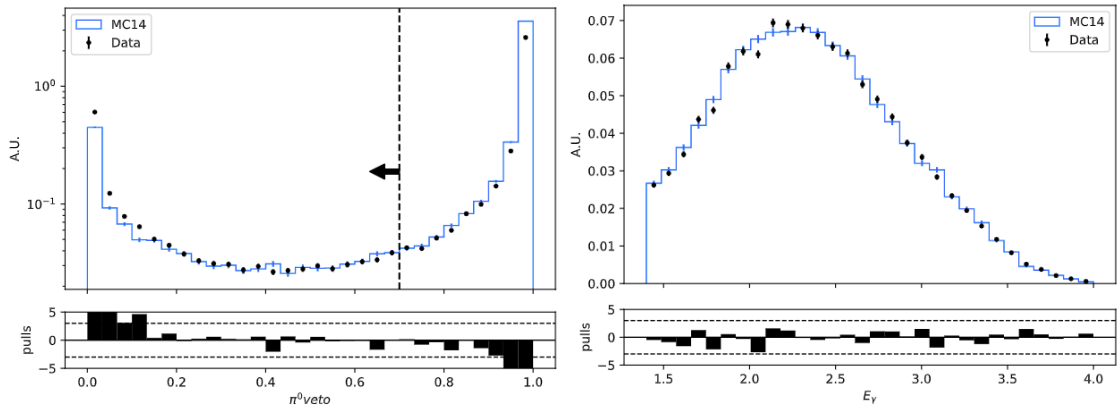

FIGURE 6.27: *Normed π^0 veto variable (left) and E_γ (right) distributions for MC and data sidebands for $B^0 \to K_S^0\pi^0\gamma$. The bar indicates the value of the cut and the arrow is oriented to the kept part.*

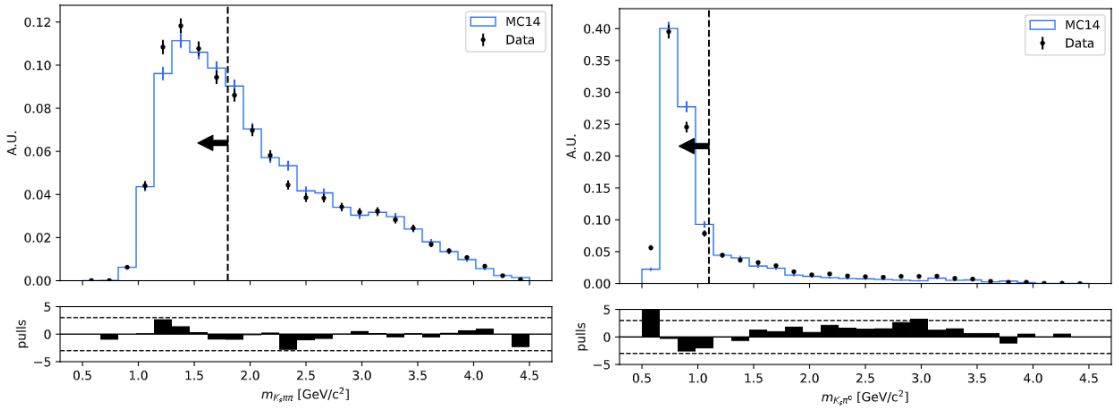

FIGURE 6.28: *Normed $m_{K_S^0\pi^+\pi^-}$ and $m_{K_S^0\pi^0}$ distributions for MC and data sidebands for $B^0 \to K_S^0\pi^+\pi^-\gamma$ (left) and $B^0 \to K_S^0\pi^0\gamma$ (right) after applying all the cut from the selection except on this variable.*

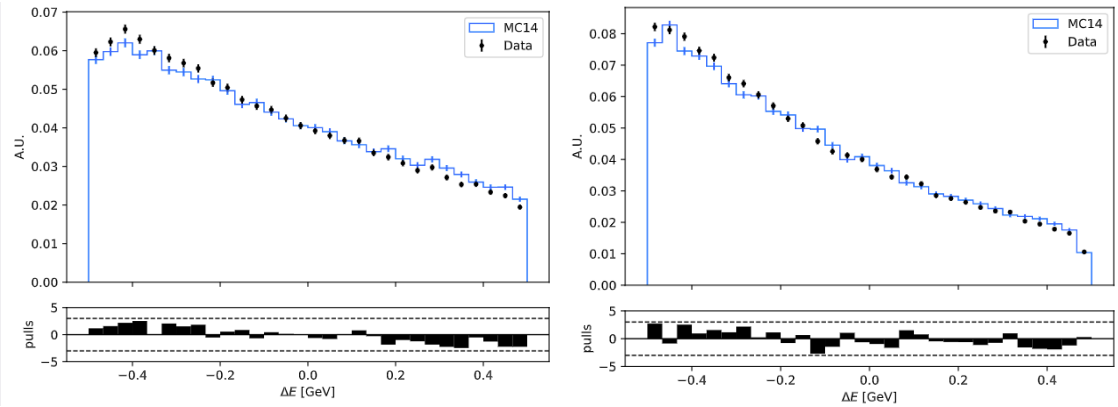

FIGURE 6.29: *Normed ΔE distributions for MC and data sidebands for $B^0 \to K^0_S \pi^+ \pi^- \gamma$ (left) and $B^0 \to K^0_S \pi^0 \gamma$ (right).*

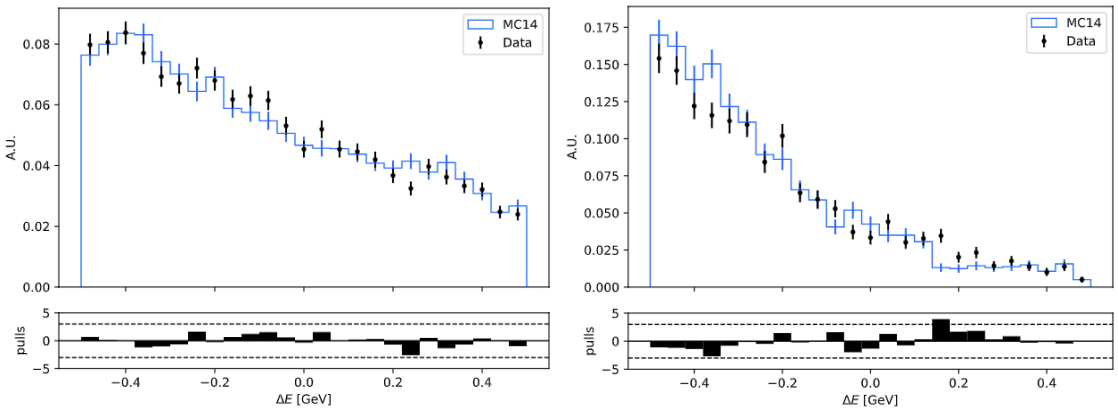

FIGURE 6.30: *Normed ΔE distributions for MC and data sidebands for $B^0 \to K^0_S \pi^+ \pi^- \gamma$ (left) and $B^0 \to K^0_S \pi^0 \gamma$ (right) after applying all the cut of the selection except m_{bc}.*

6.9.2 Data-MC check

The first check in the unblinding procedure is to redo the same distribution comparison as in Sec. 6.9.1 without the m_{bc} cut.

The Data-MC comparisons for variables used in the selection are shown in Fig. 6.31, Fig. 6.32 and Fig. 6.34 for $B^0 \to K^0_S\pi^+\pi^-\gamma$ and Fig. 6.35, Fig. 6.36 and Fig. 6.37 for $B^0 \to K^0_S\pi^0\gamma$.

The discrepancy at the tail of the m_{bc} variable is related to the fact that the beam energy in our MC is fixed to 10.58 GeV. Some MC samples are linked to some data samples, where various parameters, *e.g.* the beam energy, are fixed to their true values. if we use this type of MC sample and do the same comparison, it shows a fair agreement, see Fig. 6.33. Since the $B^0\overline{B}^0$ population does not depend on the beam energy, no correction is applied. For the other variables, we see no unexpected behavior compared to the data-MC sidebands study.

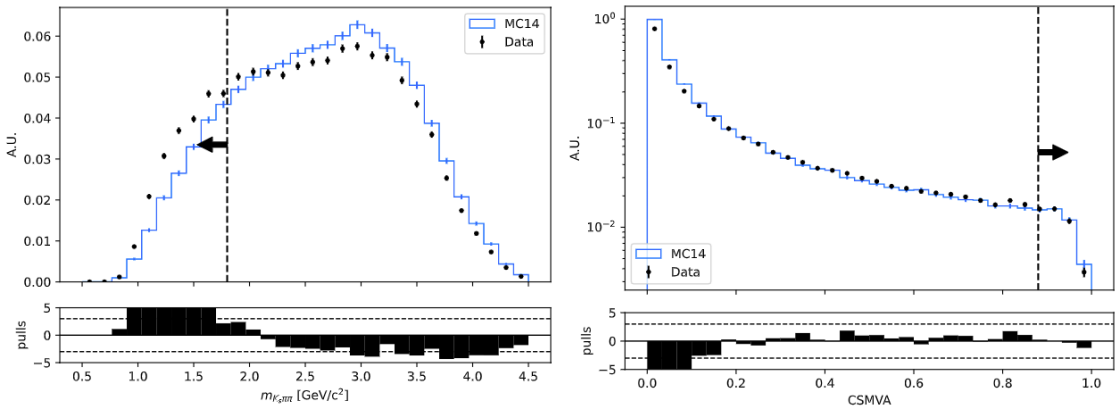

FIGURE 6.31: *Normed $m_{K^0_S\pi^+\pi^-}$ (left) and CSMVA output (right) distributions for MC and data for $B^0 \to K^0_S\pi^+\pi^-\gamma$. The bar indicates the value of the cut and the arrow is oriented to the kept part.*

The ΔE distribution for both channel is shown Fig. 6.38. No discrepancy are observed.

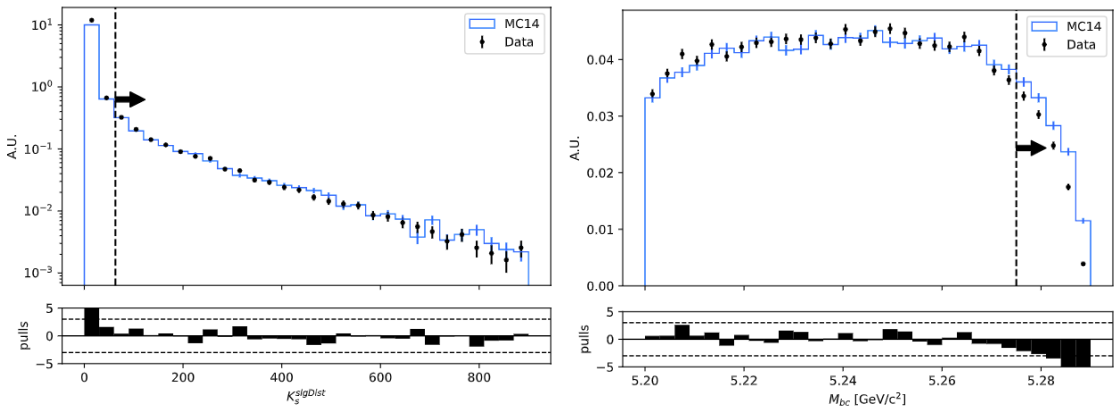

FIGURE 6.32: *Normed K^0_S significance of distance (left) and m_{bc} (right) distributions for MC and data for $B^0 \to K^0_S\pi^+\pi^-\gamma$. The bar indicates the value of the cut and the arrow is oriented to the kept part.*

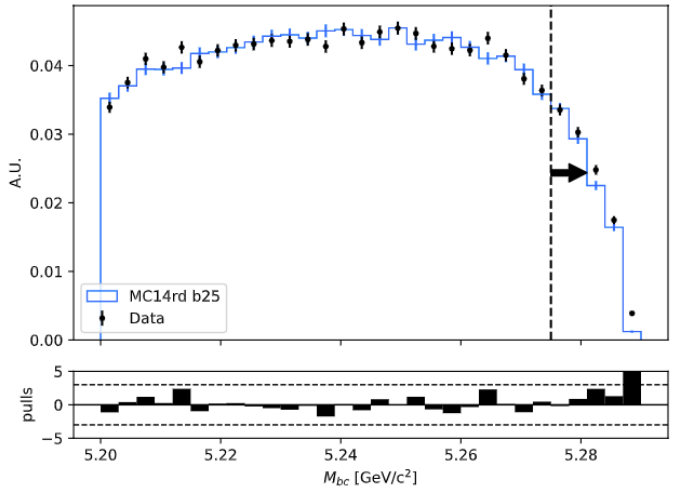

FIGURE 6.33: *Normed m_{bc} distribution for MC run-dependent and data sidebands for $B^0 \to K^0_S\pi^+\pi^-\gamma$. The bar indicates the value of the cut and the arrow is oriented to the kept part.*

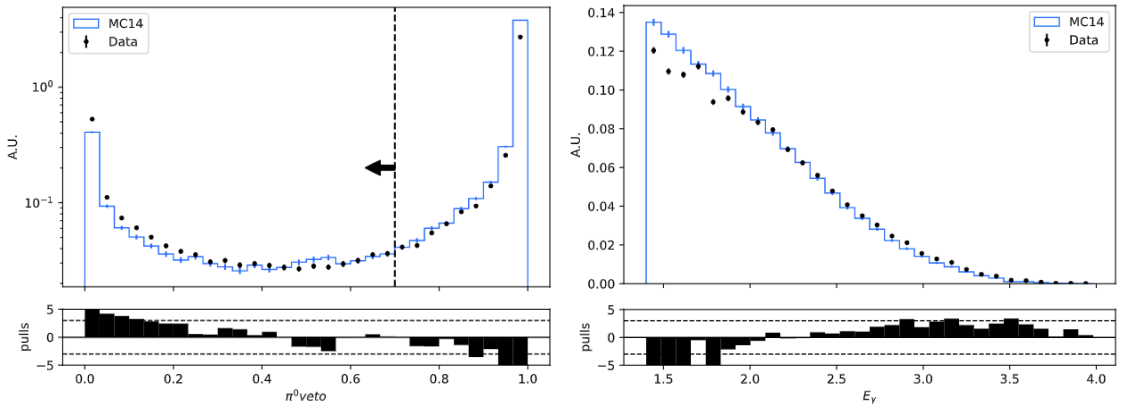

FIGURE 6.34: *Normed π^0 veto variable (left) and E_γ (right) distributions for MC and data for $B^0 \to K^0_S\pi^+\pi^-\gamma$. The bar indicates the value of the cut and the arrow is oriented to the kept part.*

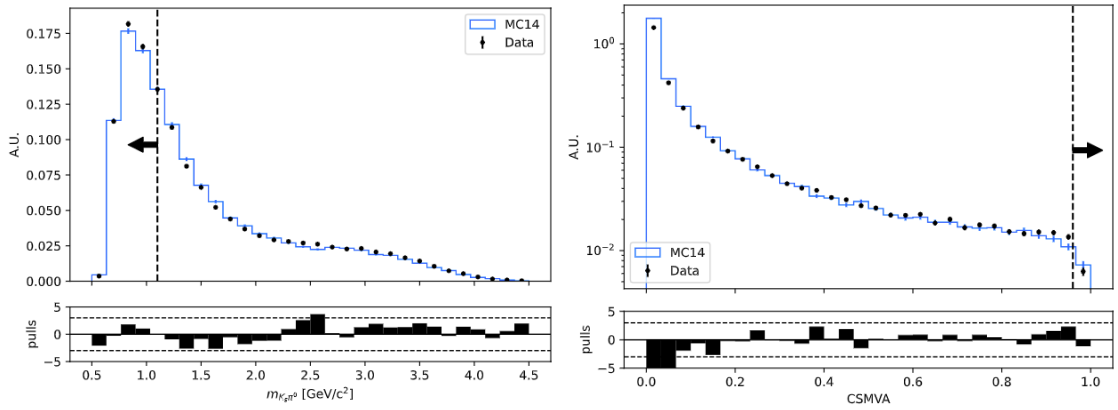

FIGURE 6.35: *Normed $m_{K^0_S\pi}$ (left) and CSMVA output (right) distributions for MC and data for $B^0 \to K^0_S\pi^0\gamma$. The bar indicates the value of the cut and the arrow is oriented to the kept part.*

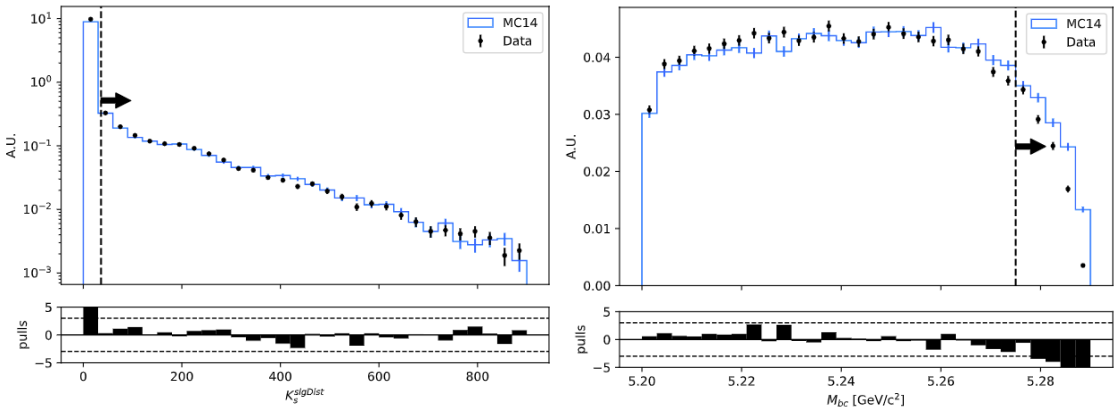

FIGURE 6.36: *Normed K^0_S significance of distance (left) and m_{bc} (right) distributions for MC and data for $B^0 \to K^0_S\pi^0\gamma$. The bar indicates the value of the cut and the arrow is oriented to the kept part.*

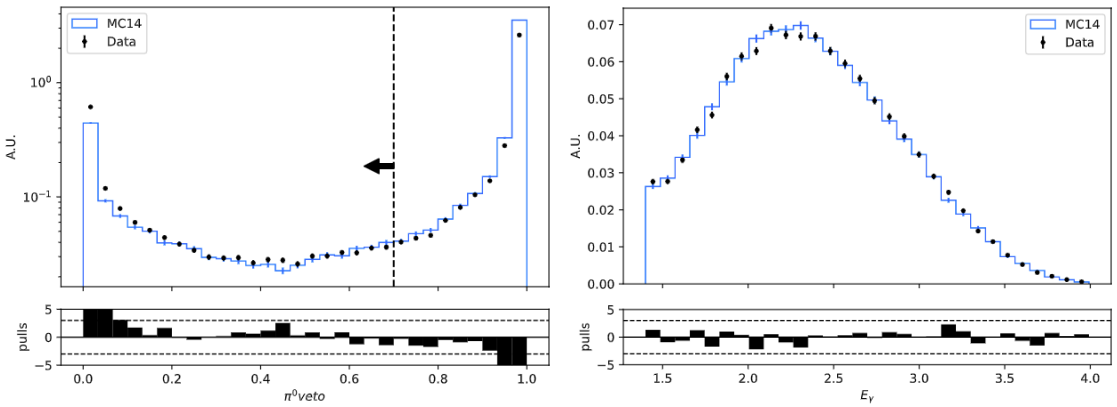

FIGURE 6.37: *Normed π^0 veto variable (left) and E_γ (right) distributions for MC and data for $B^0 \to K^0_S\pi^0\gamma$. The bar indicates the value of the cut and the arrow is oriented to the kept part.*

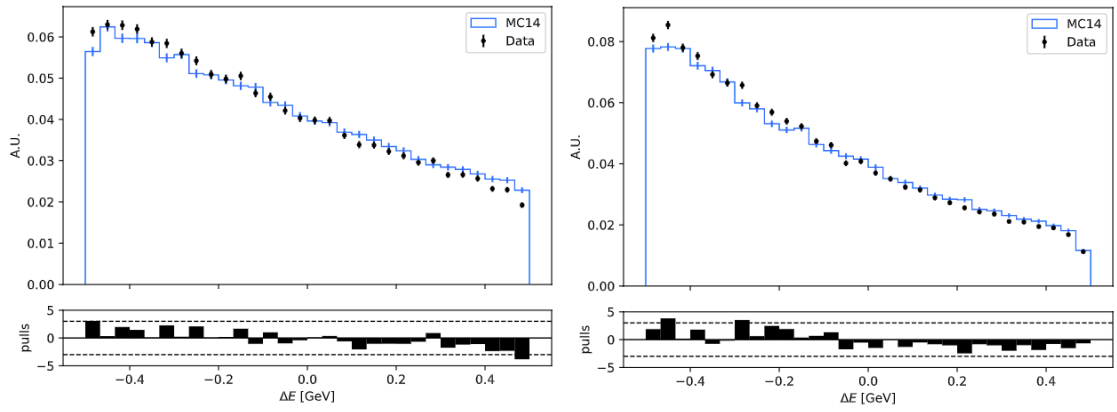

FIGURE 6.38: *Normed ΔE distributions for MC and data for $B^0 \to K^0_S\pi^+\pi^-\gamma$ (left) and $B^0 \to K^0_S\pi^0\gamma$ (right).*

Finally, the last check consists of applying all the cut from the selection and look at the ΔE distribution, shown in Fig. 6.39.

For $B^0 \to K^0_S\pi^+\pi^-\gamma$, we observe a peak in the signal region, but the amplitude is significantly lower than expected. We tried to understand this discrepancy by looking at other kinematic or vertexing variables. No discrepancy were observed. We also tried to redo the same selection on a control channel $B^+ \to K^+\pi^+\pi^-\gamma$, but the same discrepancy in ΔE is observed after the unblinding. Further studies beyond the time line of this thesis should be done to fully understand the source of the missing signal for $B^0 \to K^0_S\pi^+\pi^-\gamma$.

A good agreement is observed for $B^0 \to K^0_S\pi^0\gamma$. Once the unblind process is done, we can proceed to the ΔE fit for this channel.

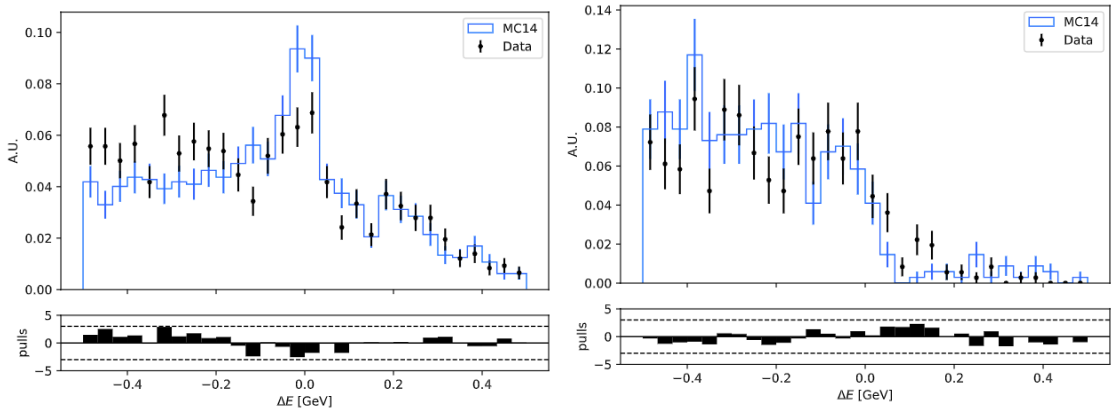

FIGURE 6.39: *Normed ΔE distributions for MC and data for $B^0 \to K^0_S\pi^+\pi^-\gamma$ (left) and $B^0 \to K^0_S\pi^0\gamma$ (right) after applying all the cut of the selection except m_{bc}.*

6.9.3 Fit and BR measurement

After applying the selection and the best candidate selection to data, we perform the fit described in Sec. 6.6.2 for $B^0 \to K^0_S\pi^0\gamma$. The result is shown Fig. 6.40. Table 6.10 summarizes the fitted values of the different parameters and the raw yields estimated.

TABLE 6.10: *Fitted values for $B^0 \to K^0_S\pi^0\gamma$ on data.*

$B^0 \to K^0_S\pi^0\gamma$	
μ_{sgn}	0.034 ± 0.018
σ_{sgn}	0.136 ± 0.040

Signal yields	121 ± 29
Background yields	158 ± 29

Using Eq. 6.15, we find a branching fraction of:

$$\mathcal{B}\left(B^0 \to K^0_S\pi^0\gamma\right) = (7.28 \pm 1.75(\text{stat}) \pm 1.03(\text{syst})) \times 10^{-6}, \tag{6.18}$$

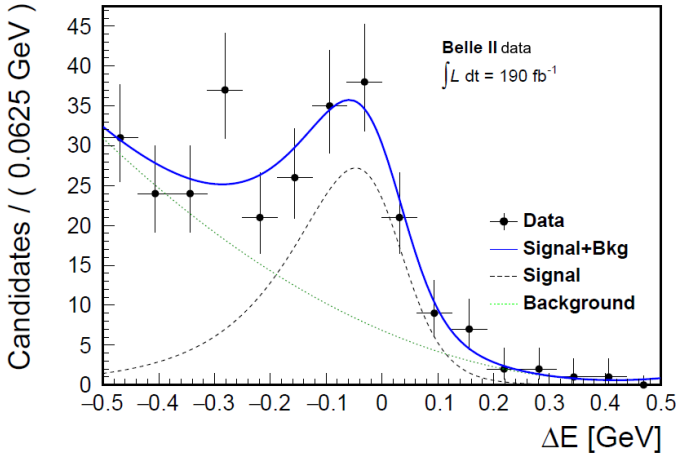

FIGURE 6.40: *Fit of the data sample for $B^0 \to K^0_S\pi^0\gamma$. The yields, mean and sigma of the signal function are left free.*

which is compatible with the known value of $(7.0 \pm 0.4) \times 10^{-6}$ [23].

6.10 Summary

I developed two rediscovery analysis in Belle II, one on $B^0 \to K^0_S \pi^+ \pi^- \gamma$ decay and one on $B^0 \to K^0_S \pi^0 \gamma$ decay. As preparation for a future time-dependent CP asymmetry measurement, both analysis showed reconstruction efficiencies inline with previous studies conducted in Belle [1,2]. An inconsistency between MC and data has been revealed in the first channel and requires further investigations. The second one lead to a branching fraction measurement with around 190 fb^{-1} of data, taken by Belle II between March 2019 and July 2021. The result is compatible with the world average, and was presented at the Moriond 2022 conferences [105] and published on arXiv in June 2022 [106].

Conclusion

In this thesis, I developed the first implementation in Belle of the time dependent CP violation measurement of $B^0 \to K^0_S \pi^+ \pi^- \gamma$ using a new method of splitting the Dalitz plane to further constrain the Standard Model. Using around 1 ab^{-1} of Monte-Carlo simulation of $\Upsilon(4S) \to e^+ e^-$, I was able to extract the following statistical sensitivity to CP violation parameters:

$$\mathcal{A} = 0.07 \pm 0.10 \tag{6.19}$$
$$\mathcal{S} = -0.09 \pm 0.12 \tag{6.20}$$

and the new parameters from [52]:

$$\mathcal{S}^+ = -0.17 \pm 0.24 \tag{6.21}$$
$$\mathcal{S}^- = 0.03 \pm 0.24. \tag{6.22}$$

where the central value is the outcome of the measurement on the simulated sample with no CP-violation included and where the statistical uncertainty comes from the fitting procedure. The result is compatible with zero, as the SM predicts. The uncertainty obtained is around a factor two improvement compared to the last Belle measurement of this channel [1]. To actually measure these parameters on Belle data, a thorough study on the systematic is needed. The work on this analysis is currently continued within the Strasbourg group, and we expect a measurement on the full Belle data by the end of 2022. To finally constrain C_7 and C'_7, we need the measurement of the dilution factor and the hadronic parameters. A good candidate to provide these inputs is the charged mode $B^+ \to K^+ \pi^+ \pi^- \gamma$, as it is easier to reconstruct and has better yields. Some studies [37, 107-109] on this decay have been done on Belle II and LHCb, and a combination of the effort to constrain the SM is currently being developed.

I also developed the rediscovery of $B^0 \to K^0_S \pi^+ \pi^- \gamma$ and $B^0 \to K^0_S \pi^0 \gamma$ decays. These analysis are the first steps toward the TDCPV measurement with Belle II data. The discrepancies between data and MC for the first channel are still under investigations, with no conclusive results yet. The second channel lead to a branching fraction measurement [106]. We found:

$$\mathcal{B}\left(B^0 \to K^0_S \pi^0 \gamma\right) = (7.28 \pm 1.75(\text{stat}) \pm 1.03(\text{syst})) \times 10^{-6}, \tag{6.23}$$

which is compatible with the world average [23]. With the updated projection

of the luminosity (see Fig. 6.41), we expect to have at least the same data set as Belle for around 2024. We can then have a TDCPV analysis of both channels to compare our sensitivity with the world average. Then, by the year 2026 (5 ab^{-1}) and 2034 (50 ab^{-1}), we expect to reduce the expected uncertainties of CP violation parameters by three and five respectively [55]. Regarding the constraint on the Wilson coefficient parametrizing new physics, C_7 and C_7', this set of measurements from Belle II will have to be combined with studies conducted on LHCb on $B \to Ke^-e^-$ channels.

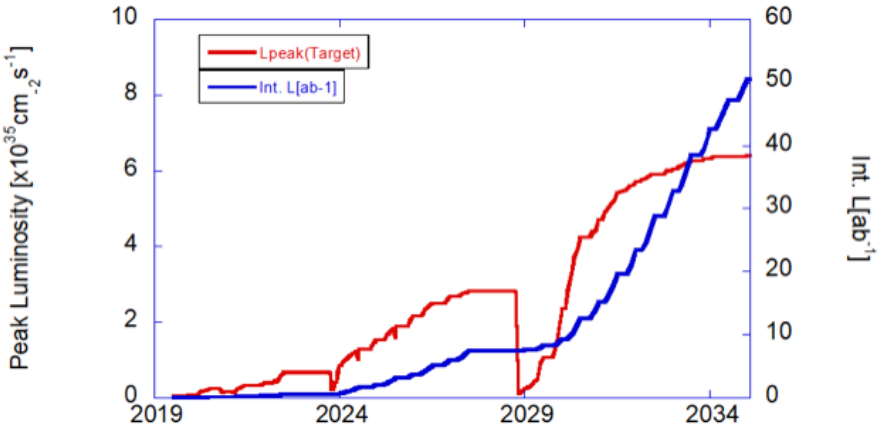

FIGURE 6.41: *Projection of the Belle II peak instantaneous (in red) and integrated (in blue) luminosity, as of June 2022.*

A long shutdown is planned around 2027, as seen in the luminosity projection Fig. 6.41, allowing for an upgrade of parts of the Belle II subdetectors.

For the upgrade of the current vertex detector of Belle II, I implemented a full package in the Belle II analysis software to simulate the new geometries [79]. Two configurations, one with five and one with seven fully pixelated layers, have been developed and connected to the tracking package in order to be used in physics simulations. The performance of the new geometries have been compared with the current performance of the Belle II trajectometer. They have a better track reconstruction efficiency and better resolution for tracks with low transverse momentum. The hit occupancy for the closest layer of the beam is reduced by three order of magnitude, due to better granularity in space and especially time. Consequently, this technology is quite robust regarding the increase in background noise.

Since its development, a six fully pixelated layers geometry has been implemented in the package, and more studies have been done in order to have a better simulation of the material support. In addition, a new technology based on PIXOR striplet detector [53] is being developed in the upgrade package. In the end, the performance comparison with the full Belle II software of the various proposed technologies and their various geometries will bring key inputs for the choice of the final upgrade setup.

Résumé en Français

Mon travail de thèse s'est déroulé dans le cadre d'une grande expérience de physique des particules : la collaboration Belle II. En m'appuyant sur les outils d'analyse des données mises à disposition par la collaboration, j'ai développé un programme de recherche qui constitue le thème principal de ma thèse. Par ailleurs, j'ai contribué aux développements de la collaboration avec une seconde étude sur la possibilité d'une jouvence du trajectomètre de l'expérience. Ces deux thématiques sont résumées successivement ci-dessous.

Analyse dépendante du temps de la violation de la symétrie CP avec Belle et Belle II

L'objectif principal de cette thèse est de découvrir des processus physiques qui ne sont pas décrits par le Modèle Standard (MS) de la physique des particules, qualifiés de *Nouvelle Physique*. Le MS, dans le canal de désintégration $B \to K_{res}\gamma$ que j'étudie, prédit une asymétrie CP[3] nulle. Mesurer une valeur non nulle peut alors mettre en défaut le MS et permettre la découverte de nouvelle physique.

Belle était un détecteur situé placé autour des collisions e^+e^- réalisées par l'accélérateur de particules KEKB. L'énergie de collision des deux faisceaux a été choisie afin de produire une résonance $\Upsilon(4S)$, qui donne dans environ 96% des cas une paire de mésons $B\bar{B}$. Belle a déjà réalisé avec une statistique proche d'un millions de paires $B\bar{B}$ (correspondant à une luminosité intégrée d'environ 1 ab^{-1}) une mesure de l'asymétrie CP sur le canal $B^0 \to K^0_S\pi^+\pi^-\gamma$, non conclusive car son niveau de précision n'était pas assez bon pour obtenir un écart significatif avec le MS [1, 2].

Belle II, l'expérience succédant Belle, visible Figure 42, a démarré sa prise de donnée fin 2018 et il faudra encore accumuler des données jusqu'en 2023 pour commencer une mesure statistiquement pertinente de l'asymétrie. De plus, il existe depuis peu une nouvelle méthode d'analyse des plans de Dalitz proposée par S. Akar *et al.* [52] qui permet d'extraire davantage d'informations afin de mieux contraindre les modèles physiques.

De ce fait, j'ai décidé d'utiliser le plus grand échantillon de données disponible, celui de Belle, afin d'implémenter pour la première fois cette analyse. J'ai redéveloppé les outils de la mesure, vérifié le temps de vie des mésons B et j'ai pu obtenir sur des données Monte-Carlo une première validation de la nouvelle méthode pour une mesure de l'asymétrie du canal de désintégration $B^0 \to K^0_S\pi^+\pi^-\gamma$. La Figure 43 illustre la dernière étape de cette analyse.

[3] La symétrie CP est une symétrie interne des systèmes élémentaires. Sa légère violation n'est autorisée que dans quelques processus du Modèle Standard, et constitue une marque distinctive de la structure du modèle.

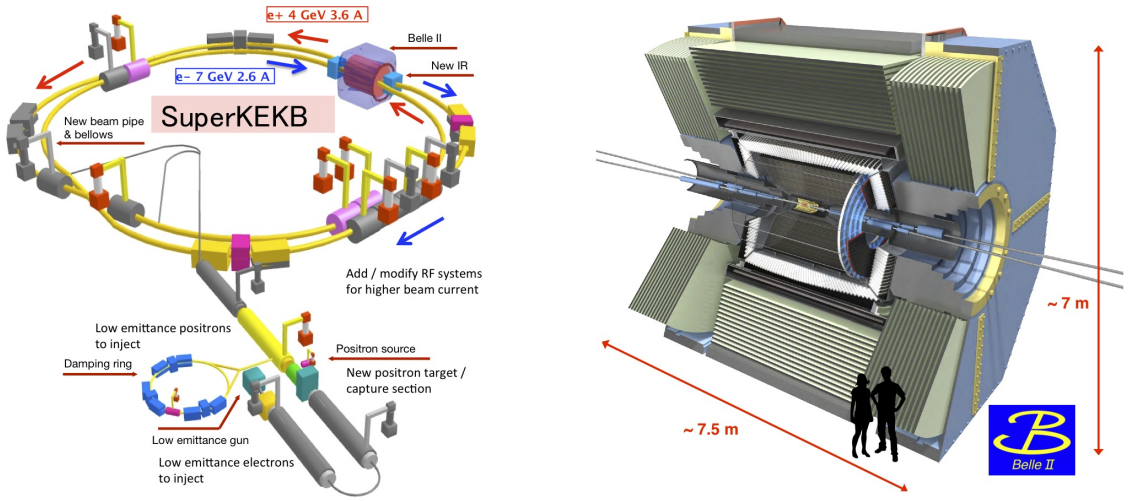

FIGURE 42: *Schéma représentant l'accélérateur SuperKEKB (à gauche) avec, au niveau de la collision, le détecteur Belle II (représenté à droite).*

Les distributions des grandeurs cinématiques m_{bc} et ΔE (respectivement la masse invariante et l'énergie du méson B reconstruit, contrainte par l'énergie du faisceau) font ressortir la présence du signal recherché, qui est responsable des bosses bien visibles sur la figure. La grandeur Δt représente la mesure de la différence de temps de désintégration entre les deux mésons B produits dans chaque événement. Cette grandeur est un élément indispensable pour l'extraction du paramètre $\mathcal{S}$ qui est obtenu suivant l'équation de la désintégration des mésons B :

$$f_{q_{\mathrm{tag}}}(\Delta t) = \frac{e^{-|\Delta t|/\tau}}{4\tau}\left[1 + q_{\mathrm{tag}}\frac{\Delta w}{2} + q_{\mathrm{tag}}\langle w\rangle\left(\mathcal{S}\sin\left(\Delta m_d \Delta t\right) + \mathcal{A}\cos\left(\Delta m_d \Delta t\right)\right)\right], \tag{24}$$

avec $\mathcal{A}$ et $\mathcal{S}$ les coefficients quantifiant la violation de CP, τ le temps de vie d'un méson B, Δm_d la différence de masse entre les états de propagation des mésons B, q_{tag} indique la saveur des B et w de combien on se trompe sur l'identification de cette saveur.

J'ai estimé la sensibilité de la mesure sur des données Monte-Carlo correspondant à environ 1 ab^{-1} :

$$\mathcal{A} = 0.07 \pm 0.10 \tag{25}$$
$$\mathcal{S} = -0.09 \pm 0.12. \tag{26}$$

La validation est très compétitive par rapport à l'ancienne mesure de Belle [2]. J'ai également estimé la sensibilité, pour la première fois, de la mesure sur le nouveau paramètre $\mathcal{S}^+$, nécessaire dans la nouvelle méthode pour contraindre le MS [52]:

$$\mathcal{S}^+ = -0.17 \pm 0.24 \quad (27)$$

$$\mathcal{S}^- = 0.03 \pm 0.24. \quad (28)$$

J'ai rédigé une note interne [97] et j'ai présenté ces résultats au *Belle Analysis Meeting* début 2021.

Afin de mesurer $\mathcal{A}$ et $\mathcal{S}$ sur les données de Belle, il faut encore réaliser une étude des incertitudes systématiques présentes pour cette mesure, et vérifier si la sélection concorde entre les données et le MC. Pour finalement contraindre le MS avec la nouvelle méthode, il faut combiner cette mesure a une analyse d'amplitude du Dalitz du canal similaire $B^+ \to K^+ \pi^+\pi^- \gamma$.

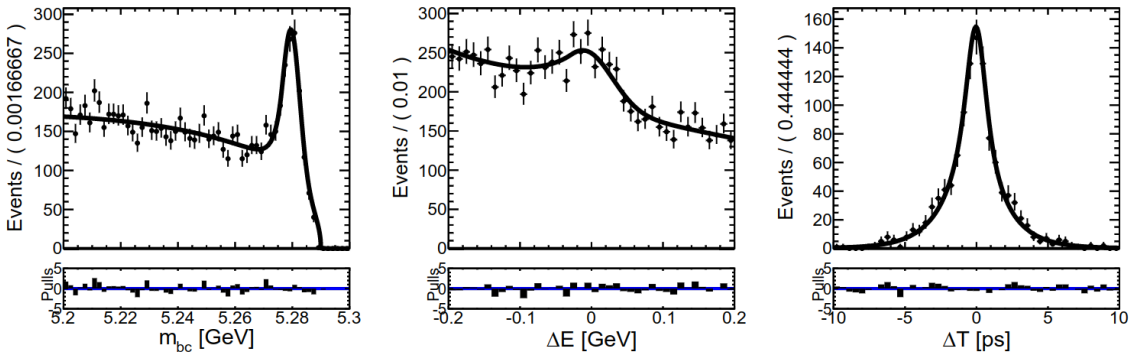

FIGURE 43: *Ajustement des fonctions sur le MC de Belle pour trois distributions utilisées pour obtenir les paramètres de la violation CP.*

Une fois cette analyse bien entamée, j'ai pu démarrer l'analyse dans Belle II. Comme Belle II n'a pas encore accumulé assez de données pour pouvoir réaliser une mesure de l'asymétrie, je me suis concentré dans un premier temps sur le développement de la redécouverte des deux canaux radiatifs $B^0 \to K^0_S\pi^+\pi^-\gamma$ et $B^0 \to K^0_S\pi^0\gamma$ ainsi que la mesure de leur rapport d'embranchement (BR). Les outils que j'ai développés pour cette analyse seront également utilisés pour la mesure de l'asymétrie. Comme ces deux canaux sont rares, ils présentent un rapport d'embranchement très faible $BR < 10^{-6}$. Pour que notre pic de signal soit visible parmi tous les autres canaux et le bruit de fond machine, j'ai du développer un algorithme de tri capable d'exploiter un grand nombre de variables discriminantes. Sur des données Monte-Carlo (MC) correspondant à 700 fb^{-1}, j'ai entraîné un arbre de décision boosté pour pouvoir discerner le bruit de fond du signal et j'ai optimisé les valeurs des coupures appliquées sur des variables d'importance. La figure de mérite[4] obtenue à la fin de la sélection est de l'ordre de 22 et 12, pour une efficacité de rétention du signal de 9,5% et 9,9%, pour $B^0 \to K^0_S\pi^+\pi^-\gamma$ et $B^0 \to K^0_S\pi^0\gamma$ respectivement.

Après la sélection, une "bosse" devient visible sur la distribution de la grandeur cinématique ΔE. Pour estimer la quantité de signal restante, j'ai du

[4] $N_S/\sqrt{N_S+N_B}$, avec N_S et N_B le nombre de candidats identifiés comme signal et bruit de fond respectivement.

ensuite ajuster des fonctions sur ΔE pour ce pic de signal et pour le bruit de fond restant sur un ensemble de donnée Monte-Carlo correspondant à la luminosité que Belle II a accumulée pour le début de ma 3e année, soit 190 fb^{-1}, afin d'obtenir une estimation du nombre de signal attendu pour les deux canaux. Figure 44 et Figure 45 montrent le résultat de l'ajustement. Une erreur statistique liée à l'incertitude statistique des paramètres de l'ajustement est associée à ce nombre.

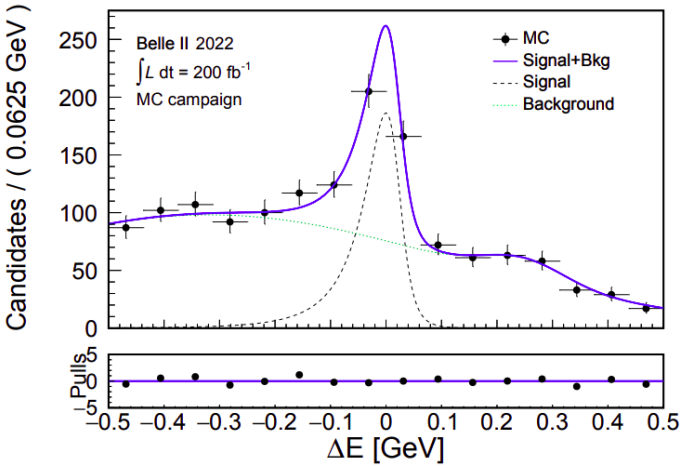

FIGURE 44: *Ajustement des fonctions de bruit de fond et du signal sur des données MC pour* $B^0 \to K^0_S \pi^+ \pi^- \gamma$.

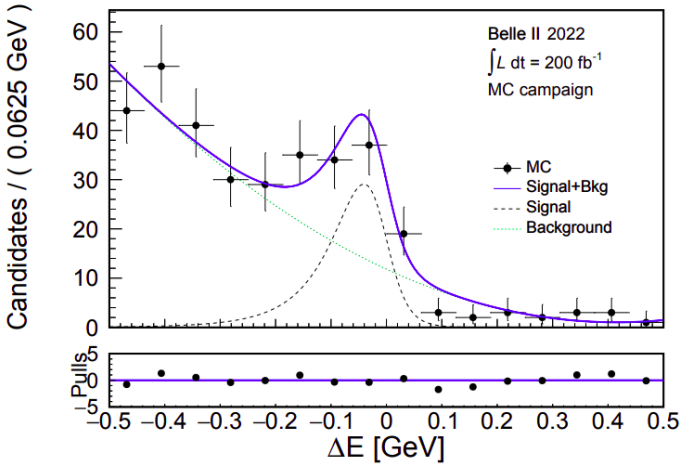

FIGURE 45: *Ajustement des fonctions de bruit de fond et du signal sur des données MC pour* $B^0 \to K^0_S \pi^0 \gamma$.

J'ai ensuite réalisé une étude des erreurs systématiques pour chaque paramètre du calcul du rapport d'embranchement. Ceux-ci proviennent de la reconstruction des traces, de la façon d'optimiser la sélection ou encore de l'ajustement des fonctions si les distributions MC sont légèrement différentes.

Dans Belle II, toute mesure officielle suit une procédure dite *à l'aveugle*. On ne regarde les données que lorsque l'optimisation sur le MC est réalisée. S'en suit une comparaison du MC avec une partie des données, dans une zone cinématique qui ne correspond pas à celle du signal, dans le but de vérifier si

le MC reproduit bien les données et appliquer des corrections systématiques si ce n'est pas le cas. Ensuite, on révèle la zone de signal et on compare à nouveau les distributions avant et après la sélection pour enfin réaliser la mesure du BR. J'ai donc comparé les formes des distributions entre les données et le MC pour vérifier leur concordance pour les deux canaux et j'ai finalement pu ajuster le pic de signal pour $B^0 \to K^0_S \pi^0 \gamma$ obtenu dans les données (voir Fig. 46) pour mesurer un rapport d'embranchement de :

$$\mathcal{B}\left(B^0 \to K^0_S \pi^0 \gamma\right) = (7.28 \pm 1.75(\text{stat}) \pm 1.03(\text{syst})) \times 10^{-6}, \tag{29}$$

qui est compatible avec la valeur du MS $(7.0 \pm 0.4) \times 10^{-6}$ [23].

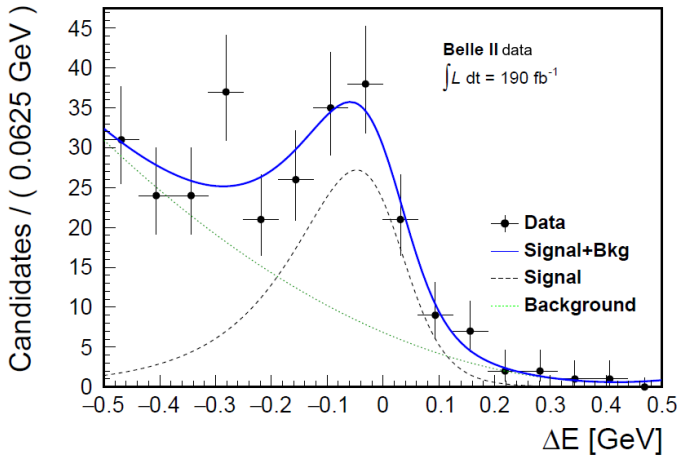

FIGURE 46: *Ajustement des fonctions de bruit de fond et du signal sur les données pour* $B^0 \to K^0_S \pi^0 \gamma$.

J'ai rédigé une note interne à Belle II [110], qui a été examinée par un groupe de travail spécialement désigné. Après avoir répondu et implémenté différents commentaires et suggestions, la note a ensuite été passée en revue par la collaboration afin de valider mon analyse et de pouvoir la présenter à des conférences. Mon résultat a été présenté aux rencontres de Moriond [105] et la note a été publiée sur arXiv [106].

Amélioration du trajectomètre de Belle II

Un deuxième sujet de ma thèse consiste à participer à une amélioration à long terme (2026) de l'expérience Belle II par le renouvellement du trajectomètre interne via des couches entièrement pixelisées. Détecteur le plus proche du faisceau, son environnement radiatif extrême fait qu'il est l'un des premiers à devoir être changés pour répondre aux défis qu'apporte une augmentation de la luminosité. Le premier objectif était d'explorer différentes configurations (géométrie du détecteur, performances des capteurs) afin de bien caractériser le nouveau détecteur. À l'aide d'un logiciel de calculs numériques simplifiés pour les performances de trajectométrie, j'ai pu conclure qu'une très petite taille

de pixel n'augmente pas énormément les performances à cause de la diffusion multiple des particules à travers la matière du détecteur, mais qu'une réduction de la taille ou de la longueur de radiation du tube à vide permet un gain d'un facteur deux sur les résolutions des paramètres des traces.

Le second objectif consistait à contribuer à la mise en place d'une simulation détaillée de ces géométries. J'ai été responsable de l'implémentation, dans l'infrastructure logicielle complète de Belle II (basf2) [67], de trois configurations à cinq et sept couches de capteurs pixelisés, avec et sans disques à l'avant, visibles Figure 47. Seules les deux premières géométries ont été complètement implémentées dans basf2.

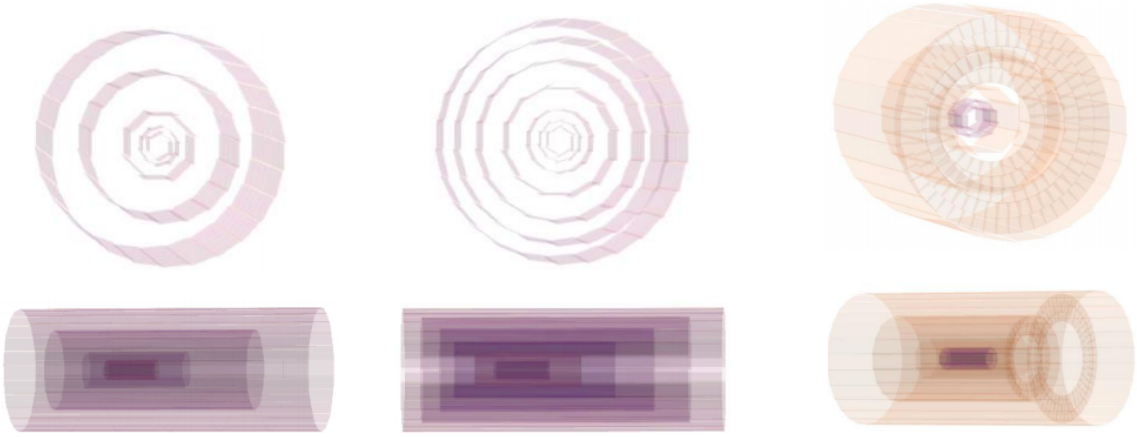

FIGURE 47: *Vue de face (haut) et de profil (bas) de la configuration à cinq, sept couches de capteurs pixelisés et cinq couches avec disques à l'avant (de gauche à droite).*

J'ai aussi été responsable de l'ajustement du simulateur de la réponse détaillée des capteurs (*digitizer*). Avec les données recueillies lors d'un test en faisceau des capteurs adaptés pour notre trajectomètre, j'ai pu obtenir dans basf2 des distributions de charges et des résolutions similaires, montrant ainsi que notre modèle de *digitizer* est adapté pour simuler ce type de capteur pixelisé. Les résultats de l'ajustement sont montrés Figure 48.

J'ai également participé au travail nécessaire pour connecter les géométries et l'algorithme de reconnaissance des traces afin de pouvoir rendre opérationnel les géométries. Une fois que cette dernière étape a été réalisée, j'ai pu comparer les performances de nos nouvelles configurations avec les performances actuelles du trajectomètre de Belle II. Les nouvelles géométries ont une meilleure efficacité de reconstruction des traces, ou encore une meilleure résolution pour des traces à faible impulsion transverse, comme montré Figure 49.

L'occupation en hit pour la couche la plus proche du faisceau est diminuée de trois ordre de grandeur, due à une meilleure pixelisation et un plus petit temps d'intégration. De plus, comme montré Tableau 11, l'efficacité du SVD actuel décroit de 6% en moyenne lorsque le bruit de fond nominal est multiplié par cinq, pris comme facteur de sécurité. Ce n'est pas le cas pour les nouvelles géométries. Elles ne perdent que 0.5% à 1% en moyenne, ce qui montre que cette technologie est assez robuste quant à l'augmentation du bruit de fond.

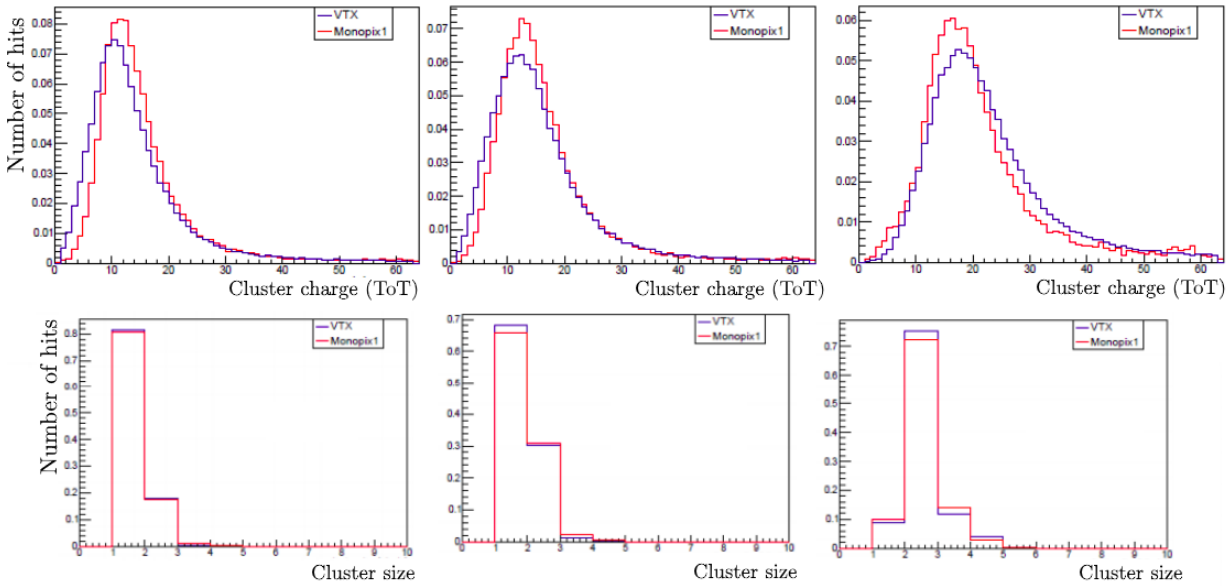

FIGURE 48: *Distribution de la charge (haut) et de la taille (bas) des clusters pour l'étude en faisceau-test (en rouge) et le résultat de l'ajustement du digitizer dans basf2 (en bleu), pour différents angle du faisceau (de gauche à droite): incidence normale, 30° et 60° (détails Fig. 4.9).*

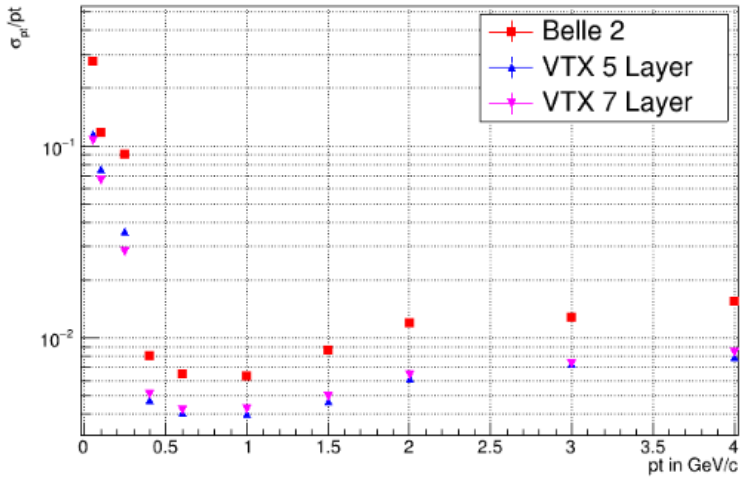

FIGURE 49: *Résolution sur l'impulsion transverse en fonction de p_t, pour en rouge le trajectomètre de Belle II, en bleu et en rose la configuration à cinq et sept couches respectivement.*

TABLE 11: *Efficacité de la reconstruction des traces pour le SVD et les deux nouvelles géométries (cinq et sept couches) pour un bruit de fond nominal et multiplié par cinq.*

	Bruit de fond × 1	Bruit de fond × 5
SVD actuel	0.961	0.907
5 couches	0.984	0.979
7 couches	0.987	0.978

Ces travaux ont été présentés dans le groupe *Upgrade* de la collaboration et ont permis de valider l'utilisation de mon outil pour les études d'optimisation à venir.

J'ai également travaillé sur l'amélioration de l'algorithme de reconstruction des trajectoires pour le préparer à l'utilisation des capteurs pixelisés qui requièrent une subdivision plus détaillée. Cette méthode utilise un algorithme de reconnaissance de motifs qui, au lieu de les chercher à travers les pixels d'un même plan image, essaye de mettre en relation les pixels des différentes couches des capteurs. J'ai démontré qu'on peut utiliser une division plus fine de ces plans pour les mettre plus rapidement en relation sans pénaliser les performances de l'algorithme. Ce travail, qui compte comme un *service task* nécessaire pour être auteur des publications de la collaboration, est résumé dans une note interne et je suis également auteur d'un des articles du groupe de travail sur la reconstruction des traces [56].

Le travail concernant l'amélioration du trajectomètre ainsi que le *service task* ont pu être présenté à la *XXVII Cracow EPIPHANY Conference*, qui a donné lieu à une publication [79].

Conclusion

La mesure de la violation de la symétrie CP était et demeure l'un des programmes principaux et prometteurs de Belle et Belle II visant à découvrir de la nouvelle physique. Durant ma thèse, j'ai développé dans Belle un algorithme suivant une nouvelle méthode de mesure proposée dans [52] pour contraindre le modèle standard en utilisant le canal $B^0 \to K^0_S \pi^+ \pi^- \gamma$, et j'ai pu valider la méthode pour la première fois sur des données Monte-Carlo.

Par ailleurs, j'ai implémenté dans l'infrastructure logicielle de Belle II une simulation détaillée et adaptable pour des détecteurs candidats à l'amélioration du détecteur de vertex actuel de Belle II. Finalement, comme les données que Belle II a accumulé en 2022 ne sont pas suffisantes pour démarrer les analyses dépendantes du temps de la violation de la symétrie CP (TDCPV), j'ai développé l'analyse de la redécouverte pour les deux canaux $B^0 \to K^0_S \pi^+ \pi^- \gamma$ et $B^0 \to K^0_S \pi^0 \gamma$, qui a mené vers une mesure du rapport d'embranchement pour le second. Cette mesure est une première étape nécessaire pour mettre en place des mesures de type TDCPV.

Appendix A

Sector Map pruning

A.1 Introduction

The Sector Map is a key element of the Belle II tracking strategy. It's an object that contains information about how space points in different regions of the detector can be related by different tracks. All the sensors are subdivided into smaller sections, called sectors. This way, we can define relations between sectors that usually are crossed by the same tracks, as shown in Fig. A.1, where the sectors 6 and 15 are related due to the track traversing both sectors. The sector maps holds a mapping of those relations, which reduces a lot the combinatorial required for the reconstruction algorithm. To reject background hits, the sector map also stores so called "filters" that combine different space points information, mostly geometrical quantities. The sector map are described in detail in the Track Finding at Belle II paper [56].

My work consisted of increasing the segmentation of the sensors to see the impact on the tracking performance, while also pruning the least used relations to decrease the size of the sector map.

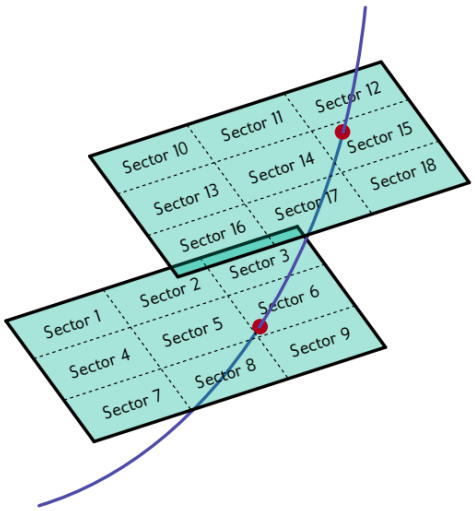

FIGURE A.1: Standard 3x3 division for the sensors and the relation between the sectors 6 and 15 which are traversed by the same track [56].

A.2 Training of the sector map

A.2.1 Standard procedure

Three steps are required to get the sector map:

- 10 M Monte-Carlo events composed by generic $\Upsilon(4S)$ and high-momentum muon events are simulated. Bhabha events are also added to further improve the tracking in the forward and backward region of the detector.
- The events are then processed and the cellular automaton stores the relations between pairs of sectors which have been traversed by at least one track as well as the minima and maxima or the quantiles of the distributions of each filter variable.
- Finally, the sector map is trained used all the relation from the cellular automaton. Each time a relation is used in the training, a proper variable called `n_found` is increased. This will be used for the pruning of the map. At the end, a root file is created with contains all the information needed to be used with the tracking algorithm.

The standard division of each sector is 3x3 in u and v direction, but this can be changed in the training. To modify this, one can change the values in the $]0, 1]$ interval of the `u` and `v` sensor divider. For the results presented in the next section, the segmentation was done as follows:

- 3x3 : {.3, .7, 1.} in u and v
- 4x4 : {.25, .5, .75, 1.} in u and v
- 5x5 : {.2, .4, .6, .8, 1.} in u and v
- 6x6 : {.165, .33, .5, .670, .835, 1.} in u and v
- 6x4 : {.165, .33, .5, .670, .835, 1.} in u and {.25, .5, .75, 1.} in v

A.2.2 Pruning

It's in the last step of the training of the sector map that the pruning is done. Each relation and filter are stored in a `SubGraph`, and the whole set of `SubGraphs` is stored in one `SectorGraph`. The `SectorGraph` is build for two hit-combination scenarios (2 or 3).

After the training, all the `n_found` for each relation have been increased to their final value. A histogram and its normalized cumulative (see Fig. A.2) of all the `n_found` is created. We can see that around 50 % of the `SubGraph` have a `n_found` < 10, meaning 50 % of all the relations have been used less than 10 $\times$ in the training of the map, and some can be found up to 3000000 times (not on the plot).

To prune the least used graphs, a small function get the maximal $\texttt{n\_found}_{max}$ below the threshold entered by the user and every relation with a `n_found` below the $\texttt{n\_found}_{max}$ is removed from the `SectorGraph`.

Launching the usual script to train a map, the user can now specify a relative threshold (in %) used to prune the sector maps.

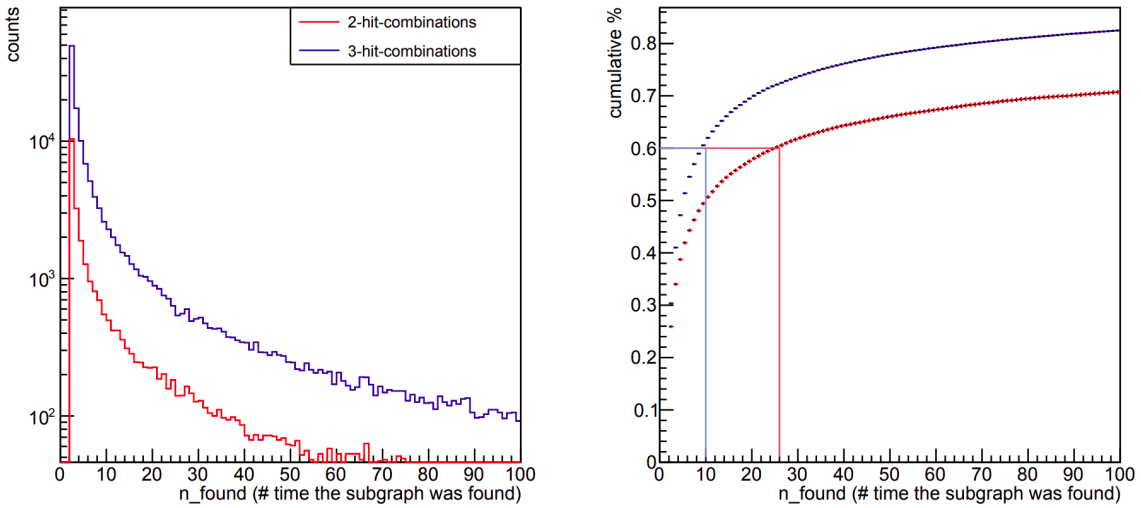

FIGURE A.2: Number of times that subgraphs were found `n_found` times (left plot) and its normalized cumulative function (right plot) for 2-hit and 3-hit combinations for a 3x3 division with 10M events trained. For example, if a 60% threshold is applied, subgraphs with `n_found` $<$ 11 for the 3-hit and `n_found` $<$ 27 for the 2-hit combination will be removed (respectively straight blue and red line in the right plot).

A.2.3 Validation

To validate a sector map, the standard SVD standalone validation script is used. 1000 $\Upsilon(4S)$ events, with and without the nominal bean-induced background, are generated and reconstructed using the map trained with a certain segmentation and threshold in order to get tracking efficiencies, clone and fake rates (see Section 4.2.2 for the definitions).

A.3 Results

Here are presented the validation results (Finding and hit efficiencies, clone and fake rates) for five different divisions (3x3, 4x4, 5x5, 6x6, 6x4) and seven different threshold (no threshold, 40%, 50%, 60%, 70%, 80%, 90%).

A.3.1 Without background

For the finding and hit efficiencies in Fig. A.3, with a maximum difference between every subdivision of 0.4% and 0.1% respectively, all configurations have more or less the same efficiency. For the clone and fake rate in Fig. A.4, the worst

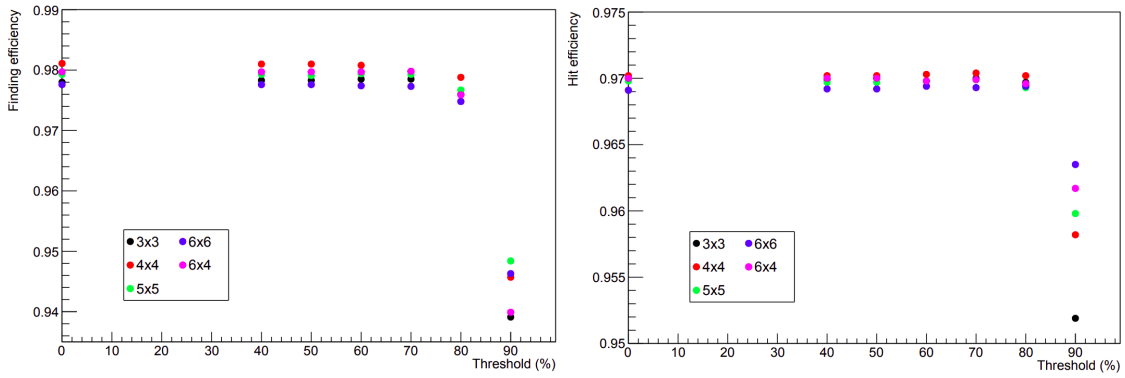

FIGURE A.3: Finding (left) and hit (right) efficiencies for different subdivisions and different thresholds, no background.

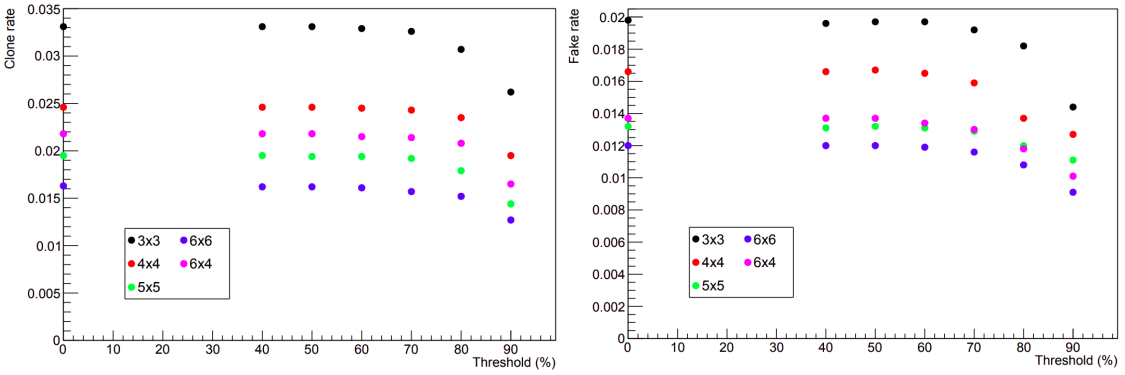

FIGURE A.4: Clone (left) and fake (right) rate for different subdivisions and different thresholds, no background.

setup is the standard 3x3 and the best is 6x6 with a $\sim$ 45% decrease. We also observe that for every figure of merit, everything is constant until 70% threshold.

The size of the map (Fig. A.5) decreases linearly with the threshold. With a 70% pruning, the loss in efficiency is negligible while the size of the sector map is divided by three.

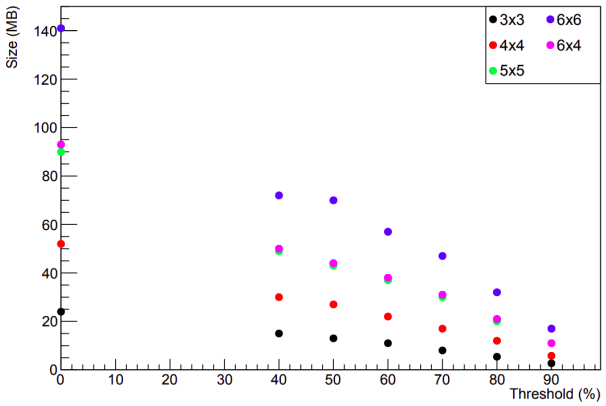

FIGURE A.5: Size of the root file containing the sector map for different configurations and subdivisions.

Overall, large division does not harm efficiencies but decreases fake and clone rate. Pruning the map can help maintain its size low. For a giver map size, the largest division below threshold can thus be selected.

A.3.2 With nominal background

With nominal background, everything behave the same way as without background. The finding efficiency is down 3.5% in average and the hit efficiency 2%, and every configurations are still really close to each other (see Fig. A.6). The clone rate is at the same level as without background, but the fake rate is five times higher for every configuration (see Fig. A.7). We still have this $\sim$ 45% decrease between the 3x3 and 6x6 segmentation. There is also little to no drop in efficiencies, clone or fake rates before 70% pruning.

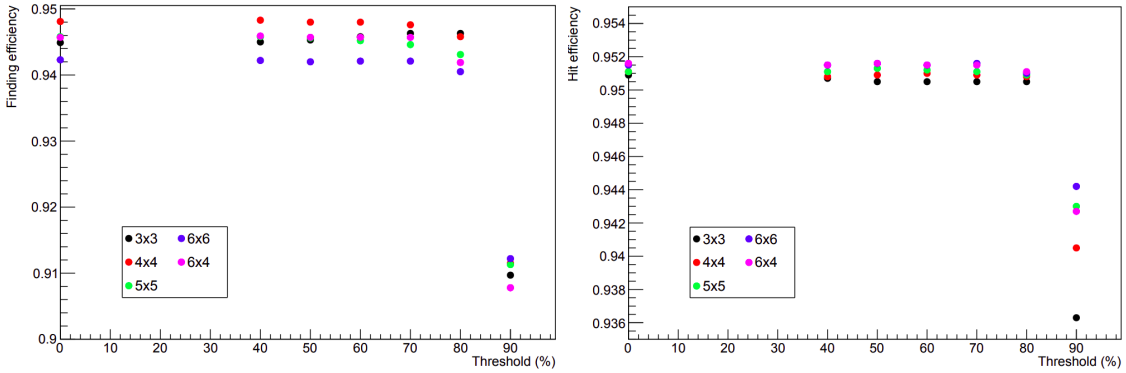

FIGURE A.6: Finding (left) and hit (right) efficiencies for different subdivisions and different thresholds with nominal background.

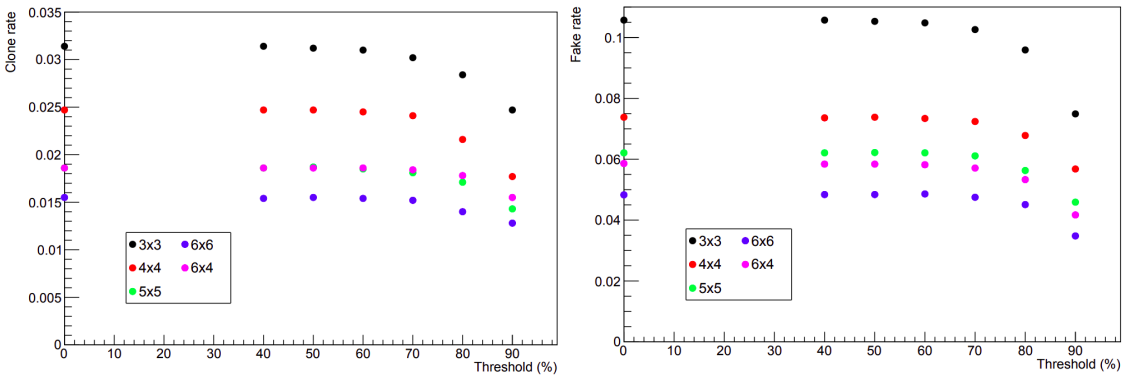

FIGURE A.7: Clone (left) and fake (right) for different subdivisions and different thresholds with nominal background.

Overall, the pruning does not affect differently the efficiency or the clone and fake rate with or without background.

The sector map used in basf2 before this study was 15MB in size (3x3 trained with 2 M events). Looking at Fig. A.5, the only sector maps trained with 10 M

events that are around this limit are the 3x3 after 50% pruning, 4x4 after 70% pruning, the 5x5 and 6x6 after 90% pruning. However, the efficiencies begin to drop after 70% pruning, so the only viable candidates are 3x3 and 4x4.

A.4 Conclusion

The sector map with 4x4 segmentation and 70% pruning seems to be the best compromise between its size and its performances. It has since replaced the 3x3 sector map used by the tracking package before this study.

Appendix B

Belle analysis, complementary fit tables

In this appendix are detailed all the values for all the PDF parameters fixed in the fit.

B.1 Signal

B.1.1 Function tables

Crystal Ball in m_{bc}	Fit value $\pm$ statistical uncertainty
μ	5.27959 ± 0.00004
σ	0.002881 ± 0.000032
α	1.16 ± 0.05
n	19 ± 16
Crystal Ball in ΔE	
μ	-0.0067 ± 0.0008
σ	0.0366 ± 0.0006
α	0.600 ± 0.018
n	6.2 ± 2.6

For Δt, $\mathcal{A}$ and $\mathcal{S}$ are left free in the final fit, and the event-by-event parameters input in the function are scaled by internal model independent coefficients, determined by other analysis before. More information in [96].

B.1.2 Low statistic fit

Figure B.1 shows the behavior of the Signal fit functions with two times the Belle dataset. No discrepancies are observed.

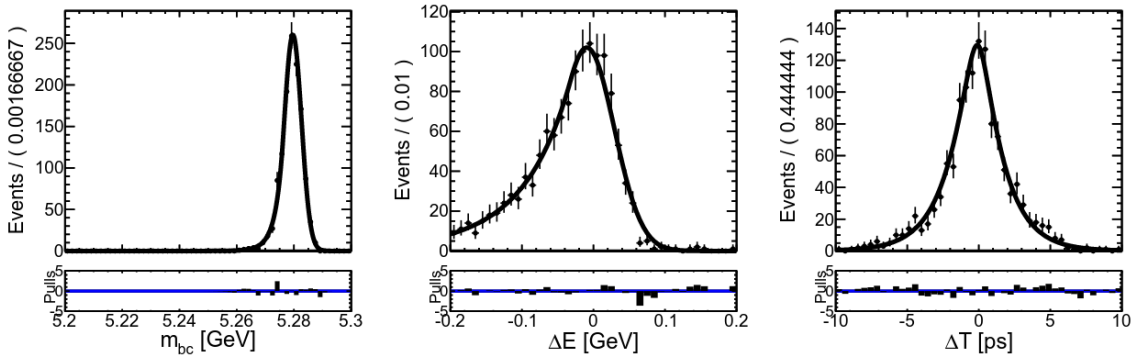

FIGURE B.1: *3D fit of m_{bc}, ΔE and Δt for the signal component on a private MC signal sample corresponding to two times the Belle dataset.*

B.2 Cross-feed background

$f\times$ Gaussian + $(1-f)\times$ Argus in m_{bc}	Fit value $\pm$ statistical uncertainty
μ	5.27910 ± 0.00030
σ	0.00360 ± 0.00022
a	-96.9 ± 3.4
f	0.217 ± 0.015
$f\times$ Exponential + $(1-f)\times$ Gaussian in ΔE	
α	-3.72 ± 0.24
μ	-0.038 ± 0.007
σ	0.058 ± 0.005
f	0.793 ± 0.025
Tatami function in Δt	
τ	1.4 ± 0.5
S_{main}	2.84 ± 0.06
S_{tail}	3.11 ± 0.10
μ_d	-0.040 ± 0.035
μ_l	-0.690 ± 0.024
f_{delt}	0.15 ± 0.18
f_{tail}	0.319 ± 0.018

B.3 Continuum background

Argus in m_{bc}	Fit value $\pm$ statistical uncertainty
a	-9.8 ± 0.6
Exponential in ΔE	
α	-1.07 ± 0.04
Tatami function in Δt	
S_{main}	1.825 ± 0.009
S_{tail}	3.77 ± 0.04
τ	1.4 ± 0.5
μ_d	-0.011 ± 0.004
μ_l	-0.70 ± 0.10
f_{delt}	0.11 ± 0.20
f_{tail}	0.1778 ± 0.0029

B.4 BB background

$f\times$ Gaussian + $(1-f)\times$ Argus in m_{bc}	Fit value $\pm$ statistical uncertainty
μ	5.2797 ± 0.0013
σ	0.0020 ± 0.0007
a	-5 ± 4
f	0.015 ± 0.008
Exponential in ΔE	
α	-6.04 ± 0.31
Tatami function in Δt	
τ	1.4 ± 0.5
S_{main}	2.79 ± 0.09
S_{tail}	3.43 ± 0.18
μ_d	-0.04 ± 0.04
μ_l	-0.70 ± 0.09
f_{delt}	0.15 ± 0.18
f_{tail}	0.218 ± 0.020

B.5 Rare background

$f\times$ Crystal Ball + $(1-f)\times$ Argus in m_{bc}	Fit value $\pm$ statistical uncertainty
μ	5.28079 ± 0.00016
σ	0.00309 ± 0.00009
α	0.322 ± 0.015
n	9.8 ± 3.5
a	-45.9 ± 1.3
f	0.243 ± 0.007
Exponential in ΔE	
α	-3.76 ± 0.09
Tatami function in Δt	
τ	1.4 ± 0.5
S_{main}	2.432 ± 0.023
S_{tail}	3.32 ± 0.04
μ_d	-0.077 ± 0.013
μ_l	-0.6 ± 0.6
f_{delt}	0.15 ± 0.18
f_{tail}	0.329 ± 0.007

B.6 CP violation parameters

TABLE B.1: CP-violation parameters measured with the Dalitz separation in an equivalent of 1 ab^{-1} of Belle MC.

$\mathcal{S}^+$	-0.17 ± 0.24
$\mathcal{S}^-$	0.03 ± 0.24
$\mathcal{A}^+$	0.13 ± 0.20
$\mathcal{A}^-$	-0.05 ± 0.20

Appendix C

Belle II analysis, additional materials

C.1 Plots of the optimization of the selection

All those distributions are from the 700 fb^{-1} data set.

C.1.1 $B^0 \to K^0_S \pi^+ \pi^- \gamma$

Distributions before the selection

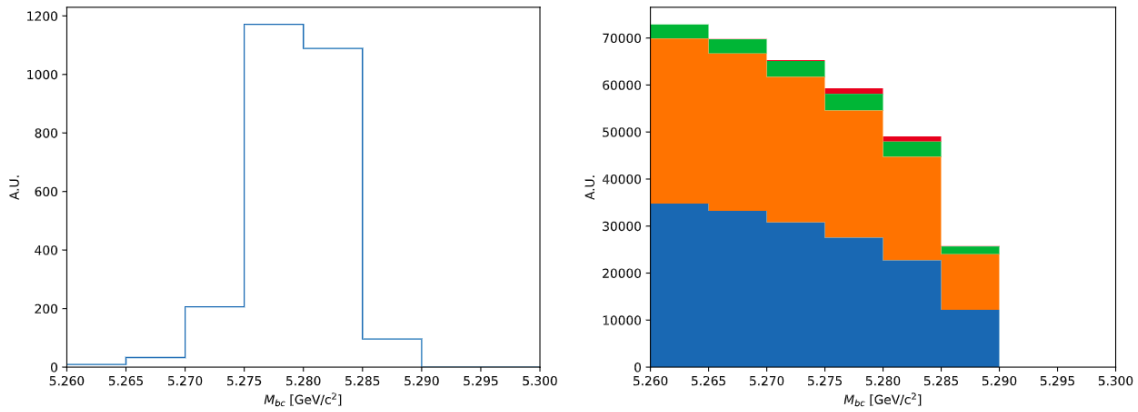

FIGURE C.1: *m_{bc} distribution on signal (left) and background+signal (right). Blue is $q\bar{q}$, orange $c\bar{c}$, green $B\bar{B}$ and red is signal.*

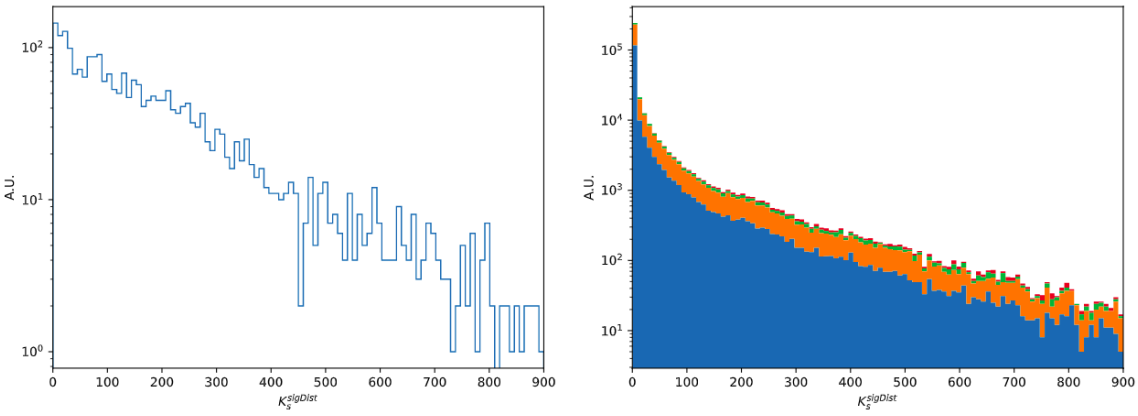

FIGURE C.2: *K_S^0 significance of distance distribution on signal (left) and background+signal (right). Blue is $q\bar{q}$, orange $c\bar{c}$, green $B\bar{B}$ and red is signal.*

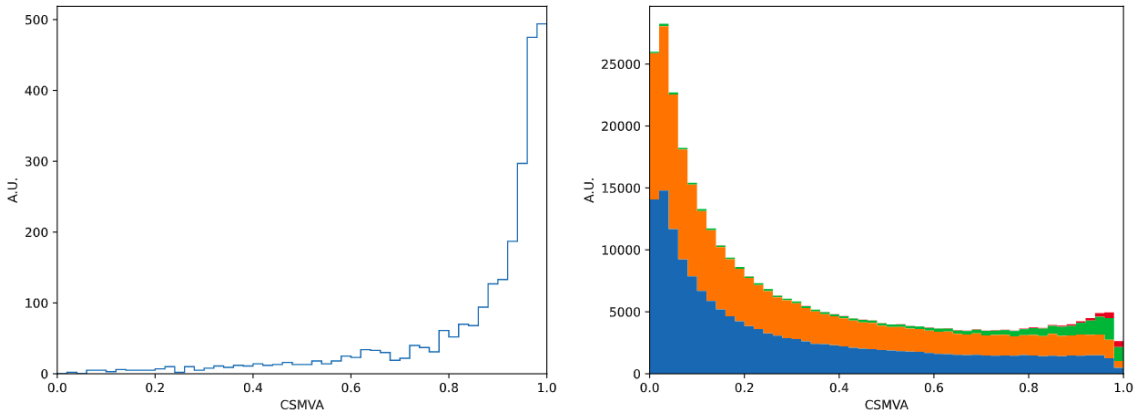

FIGURE C.3: *CSMVA output distribution on signal (left) and background+signal (right). Blue is $q\bar{q}$, orange $c\bar{c}$, green $B\bar{B}$ and red is signal.*

Distributions during the selection

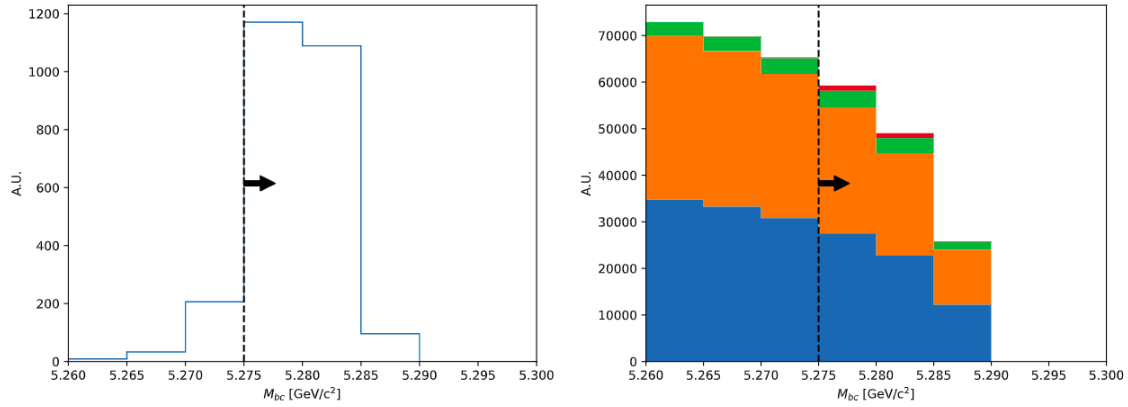

FIGURE C.4: m_{bc} distribution on signal (left) and background+signal (right). Blue is $q\bar{q}$, orange $c\bar{c}$, green $B\bar{B}$ and red is signal. The bar indicates the value of the cut and the arrow is oriented to the kept part.

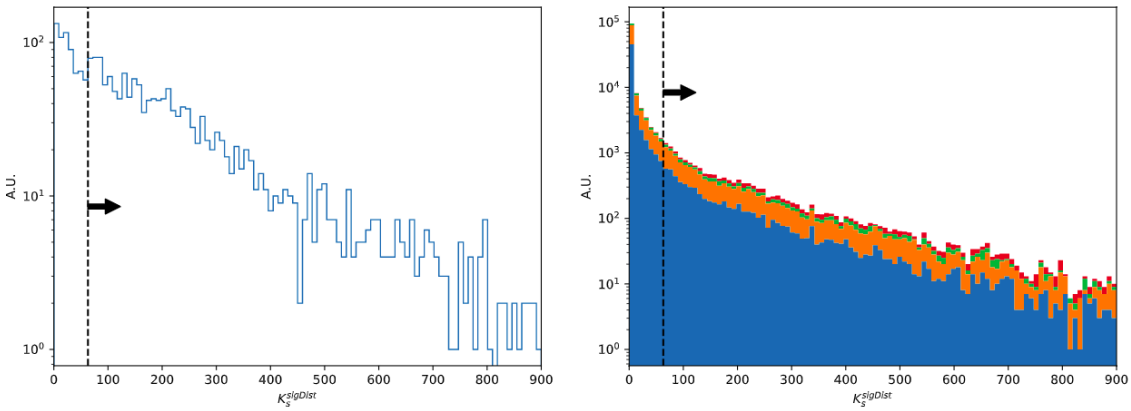

FIGURE C.5: K_S^0 significance of distance distribution on signal (left) and background+signal (right). Blue is $q\bar{q}$, orange $c\bar{c}$, green $B\bar{B}$ and red is signal. The bar indicates the value of the cut and the arrow is oriented to the kept part.

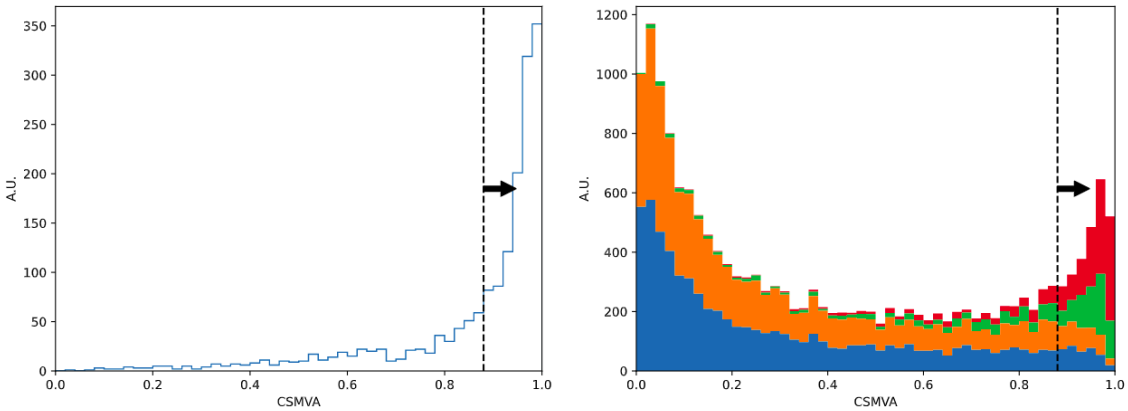

FIGURE C.6: CSMVA output distribution on signal (left) and background+signal (right). Blue is $q\bar{q}$, orange $c\bar{c}$, green $B\bar{B}$ and red is signal. The bar indicates the value of the cut and the arrow is oriented to the kept part.

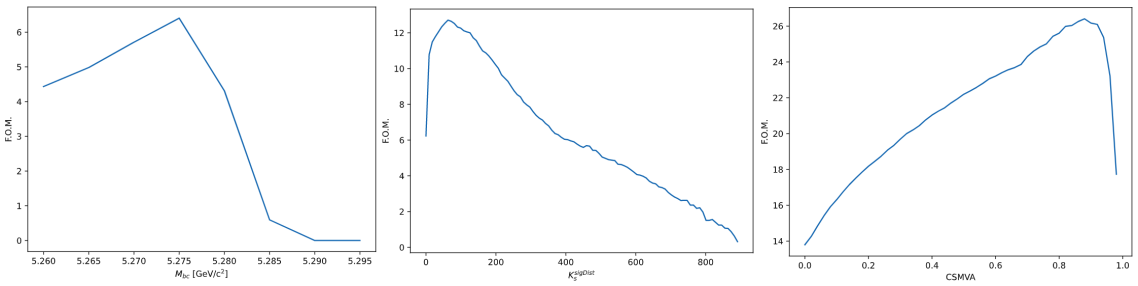

FIGURE C.7: *Evolution of the FOM as a function of the m_{bc}, K_S^0 significance of distance or CSMVA cut (from left to right).*

C.1.2 $B^0 \to K^0_S \pi^0 \gamma$

Distributions before the selection

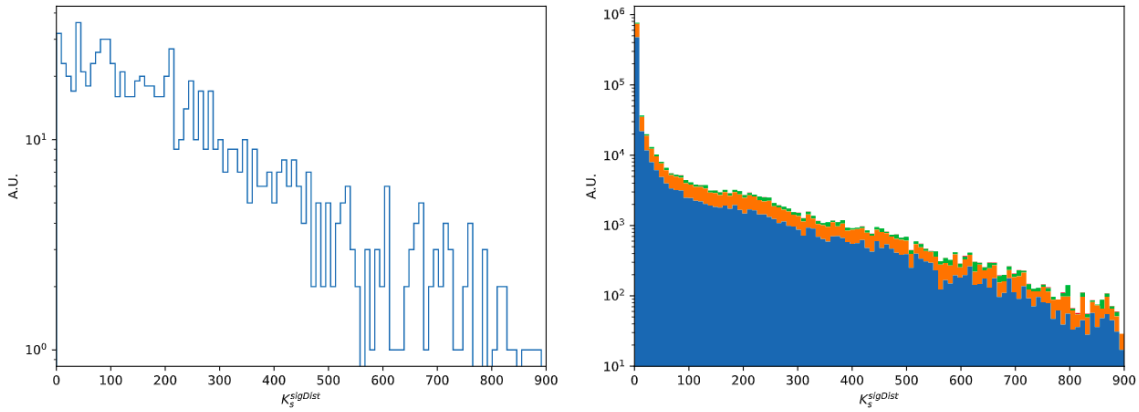

FIGURE C.8: *K^0_S significance of distance distribution on signal (left) and background+signal (right). Blue is $q\bar{q}$, orange $c\bar{c}$, green $B\bar{B}$ and red is signal.*

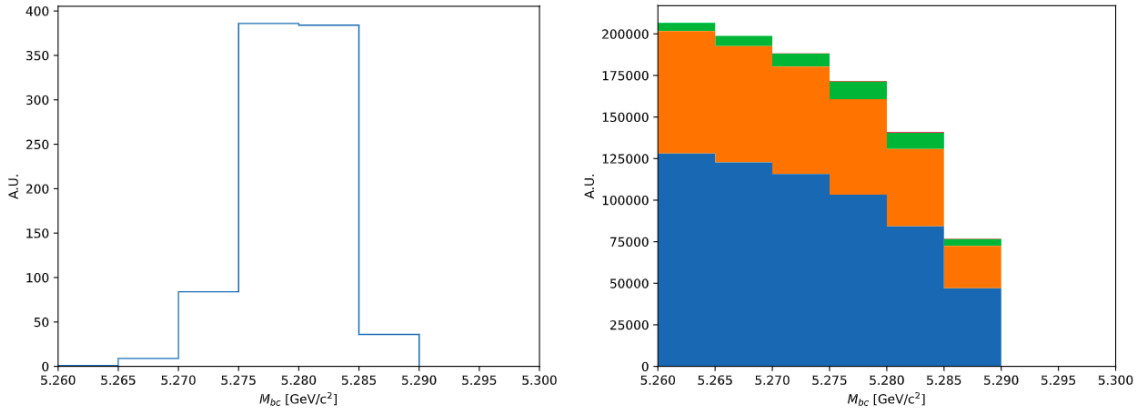

FIGURE C.9: *m_{bc} distribution on signal (left) and background+signal (right). Blue is $q\bar{q}$, orange $c\bar{c}$, green $B\bar{B}$ and red is signal.*

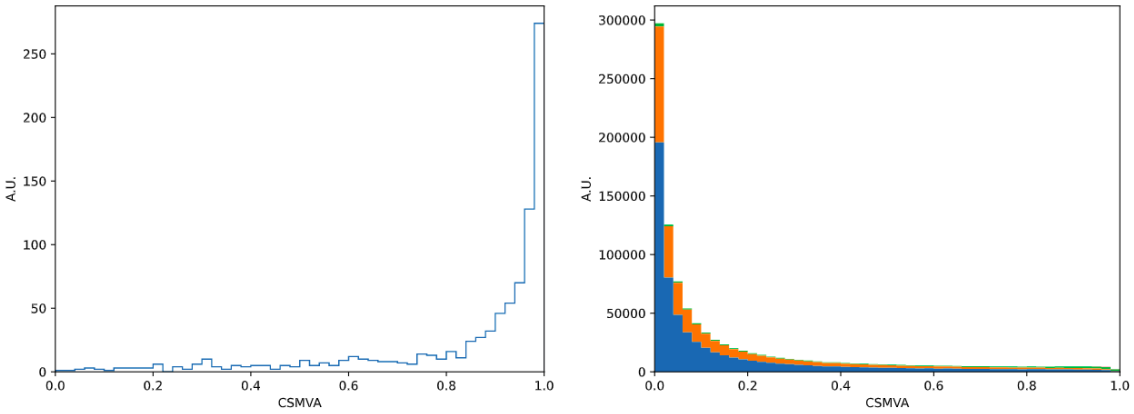

FIGURE C.10: *CSMVA output distribution on signal (left) and background+signal (right). Blue is $q\bar{q}$, orange $c\bar{c}$, green $B\bar{B}$ and red is signal.*

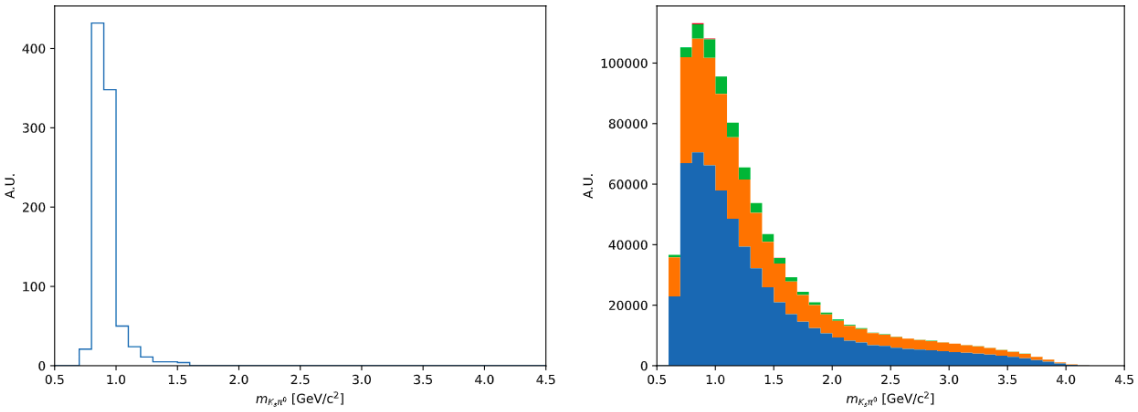

FIGURE C.11: *$m_{K^0_S\pi^0}$ distribution on signal (left) and background+signal (right). Blue is $q\bar{q}$, orange $c\bar{c}$, green $B\bar{B}$ and red is signal.*

Distributions during the selection

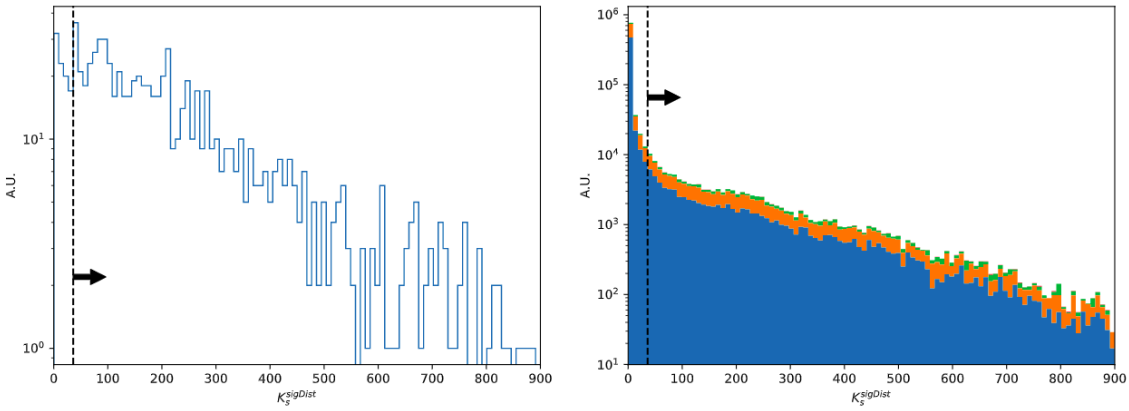

FIGURE C.12: K_S^0 *significance of distance distribution on signal (left) and background+signal (right). Blue is* $q\bar{q}$*, orange* $c\bar{c}$*, green* $B\bar{B}$ *and red is signal. The bar indicates the value of the cut and the arrow is oriented to the kept part.*

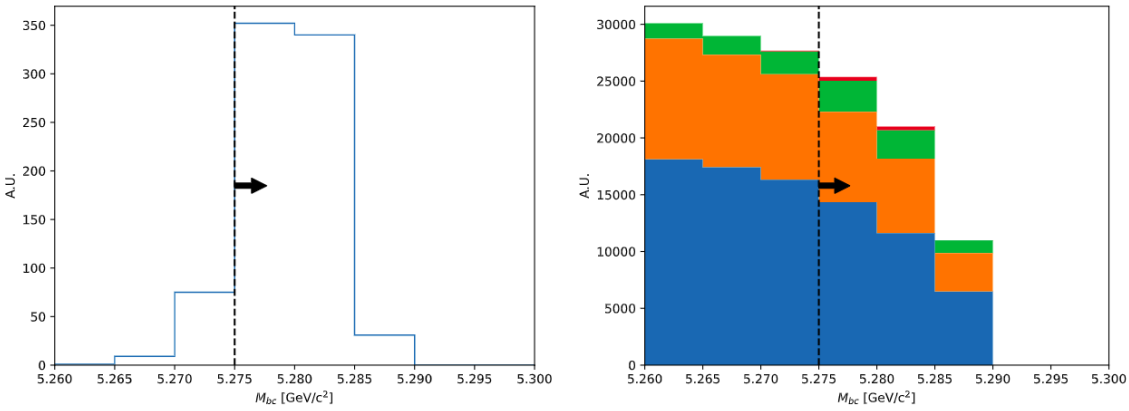

FIGURE C.13: m_{bc} *distribution on signal (left) and background+signal (right). Blue is* $q\bar{q}$*, orange* $c\bar{c}$*, green* $B\bar{B}$ *and red is signal. The bar indicates the value of the cut and the arrow is oriented to the kept part.*

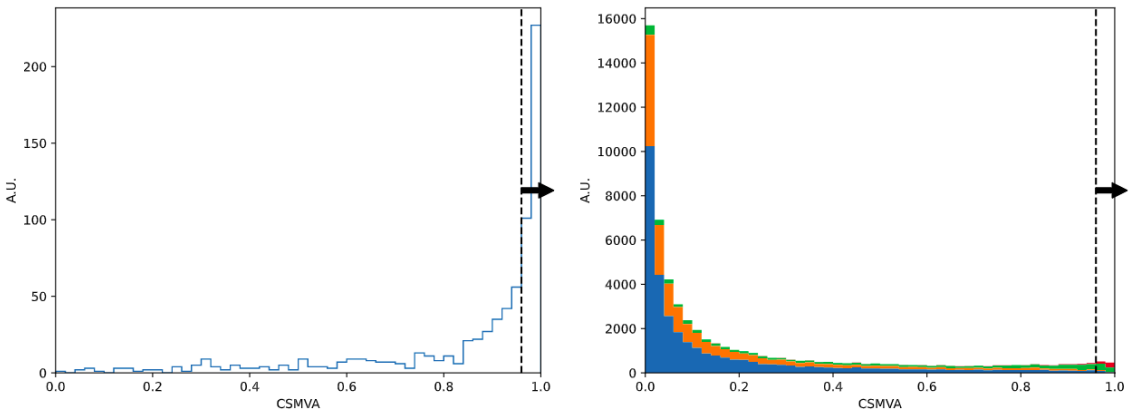

FIGURE C.14: *CSMVA output distribution on signal (left) and background+signal (right). Blue is* $q\bar{q}$*, orange* $c\bar{c}$*, green* $B\bar{B}$ *and red is signal. The bar indicates the value of the cut and the arrow is oriented to the kept part.*

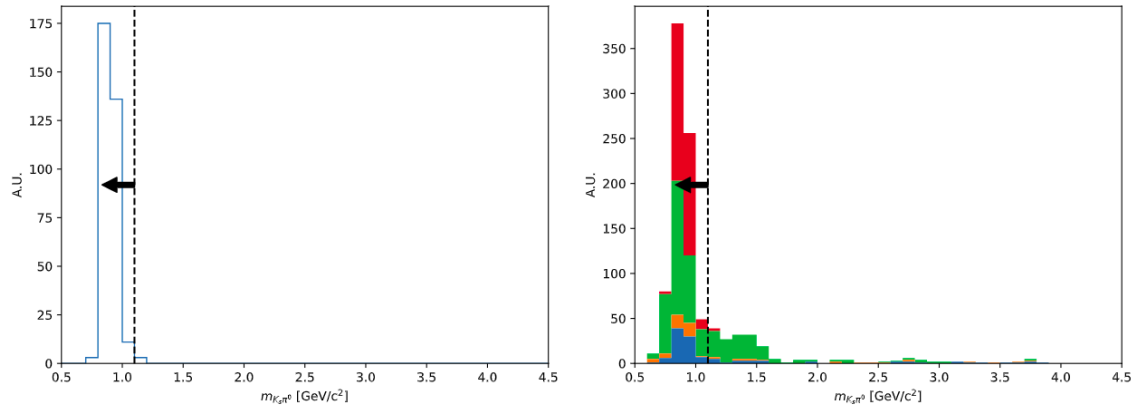

FIGURE C.15: *$m_{K^0_S\pi^0}$ distribution on signal (left) and background+signal (right). Blue is $q\bar{q}$, orange $c\bar{c}$, green $B\bar{B}$ and red is signal. The bar indicates the value of the cut and the arrow is oriented to the kept part.*

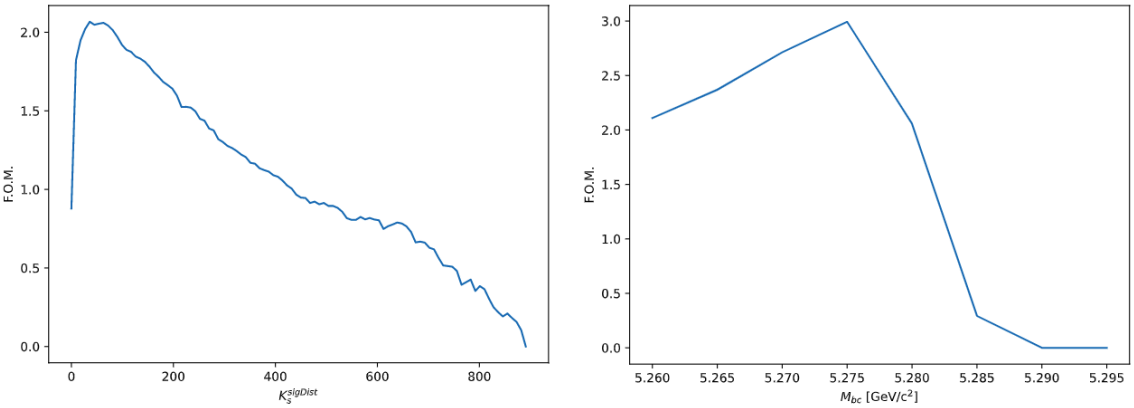

FIGURE C.16: *Evolution of the FOM as a function of the K^0_S significance of distance (left) or m_{bc} cut (right).*

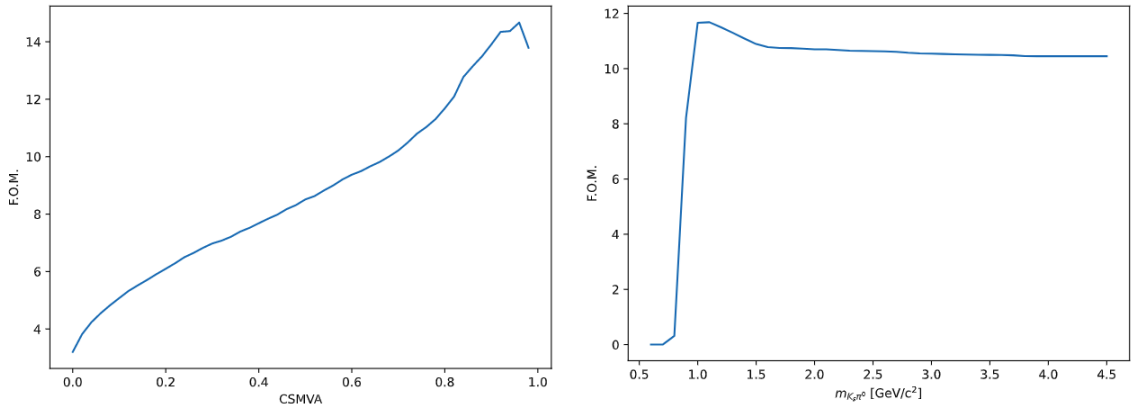

FIGURE C.17: *Evolution of the FOM as a function of the CSMVA (left) or $m_{K^0_S\pi^0}$ cut (right).*

C.2 π^0 veto FOM

π^0 veto cut was a variable optimized in the selection, but its figure of merit as a function of the cut value was mostly flat (Fig. C.18).

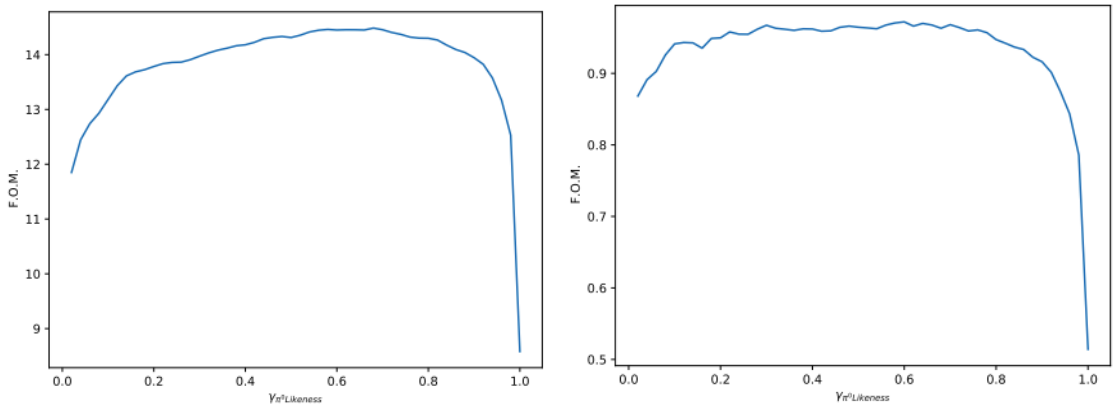

FIGURE C.18: *Evolution of the FOM as a function of the π^0 veto cut value for $B^0 \to K^0_S\pi^+\pi^-\gamma$ (left) and $B^0 \to K^0_S\pi^0\gamma$ (right).*

C.3 Pulls, additional plots

The pulls of the other variables that are left free in the fit Sec. 6.6.2 are shown Fig. C.19 and Fig. C.20 for $B^0 \to K^0_S\pi^+\pi^-\gamma$ and Fig. C.21 for $B^0 \to K^0_S\pi^0\gamma$.

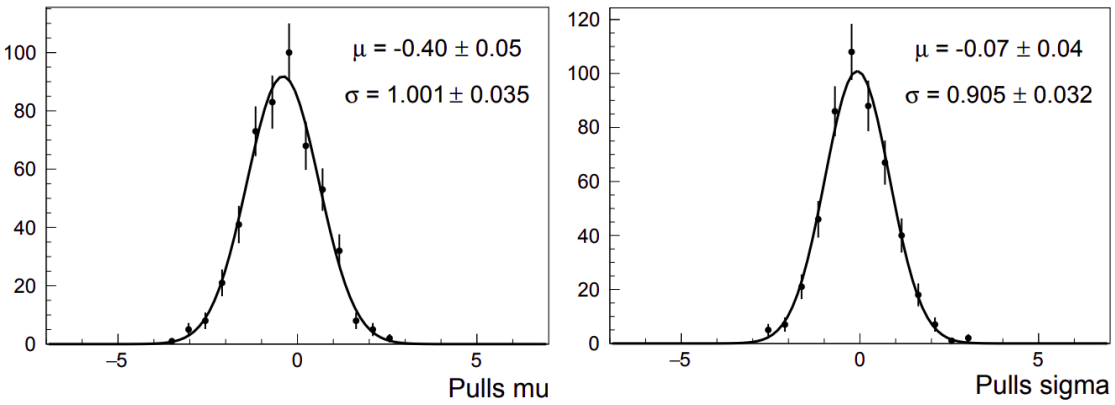

FIGURE C.19: *Pulls (see Eq. 6.8) for μ_{sgn} (left) and σ_{sgn} (right) for $B^0 \to K^0_S\pi^+\pi^-\gamma$.*

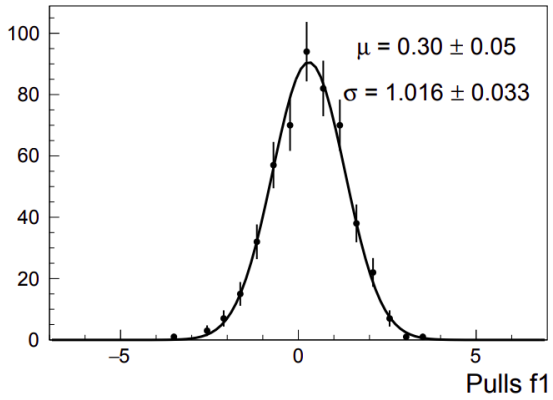

FIGURE C.20: *Pulls (see Eq. 6.8) for* $f_{Bkg-K*\gamma}$ *for* $B^0 \to K^0_S\pi^+\pi^-\gamma$.

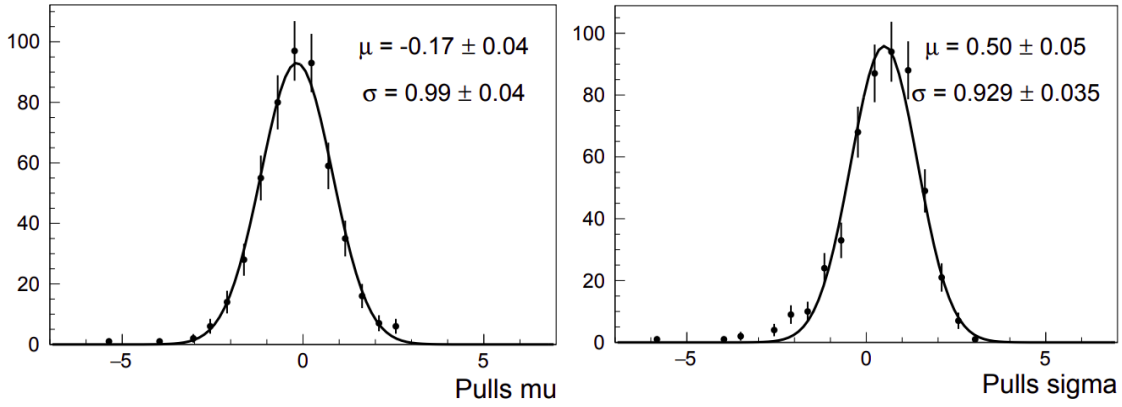

FIGURE C.21: *Pulls (see Eq. 6.8) for* μ_{sgn} *(left) and* σ_{sgn} *(right) for* $B^0 \to K^0_S\pi^0\gamma$.

C.4 $K^0_S\pi^0\gamma$ fit: complementary tests

To try to reduce the fit systematic uncertainty for $B^0 \to K^0_S\pi^0\gamma$, some tests have been made to take into account the peaking background in $-0.3 < \Delta E < -0.2$ GeV range.

C.4.1 B^+B^- peaking background fit

The first test consists of adding to the Chebyshev a Gaussian function. The background only fit on the 700 fb^{-1} MC data set is shown Fig. C.22 on the left and the signal and background fit on the 200 fb^{-1} is shown on the right. For the signal, the same Johnson function is used as in Sec. 6.6.2. For the new Gaussian, the mean and sigma are defined with the same technique as in Eq. 6.6. If the fraction between the Gaussian and the Chebyshev is let free, the fit does not converge (the Gaussian disappears, as there is not enough stats in the 200 fb^{-1} to get a correct fit). To handle this, the fraction was fixed to the value from the background only fit.

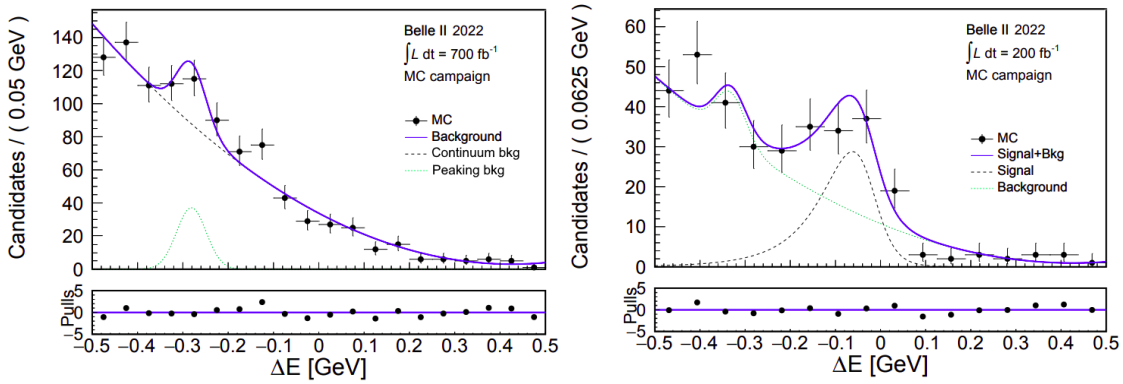

FIGURE C.22: *Fit of the background (left) and the final fit (right) on the simulated MC sample, corresponding to 700 fb^{-1} and 200 fb^{-1} respectively, for $B^0 \to K^0_S\pi^0\gamma$.*

The result of the toy study with this new background contribution is shown Fig. C.23 and Fig. C.24. The shift in the signal pull is worse. The Chebyshev is enough to fit the background at this statistic.

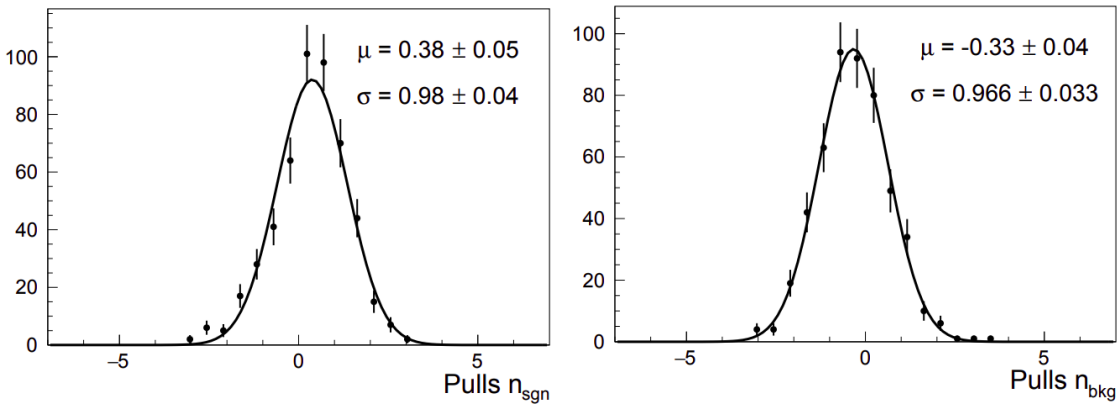

FIGURE C.23: *Pulls (see Eq. 6.8) for the signal (left) and background (right) yields for $B^0 \to K^0_S\pi^0\gamma$ for the fit Fig. C.22.*

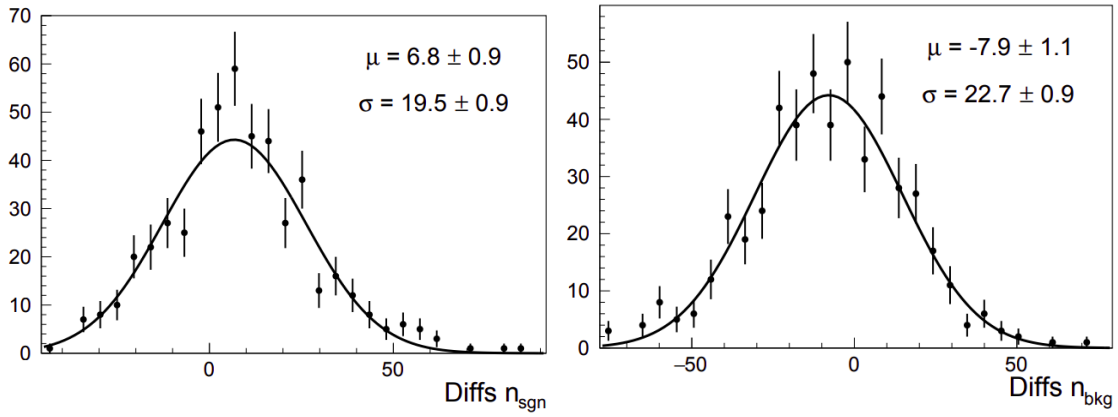

FIGURE C.24: *Differences (see Eq. 6.7) for the signal (left) and background (right) yields for $B^0 \to K^0_S \pi^0 \gamma$ for the fit Fig. C.22.*

C.4.2 Background parameters free

In order to reduce the toy systematic, we tried to let the background parameters free and rerun the toys study. The differences and pulls of the signal yield is shown Fig. C.25. We observe a smaller overestimation of the signal, reducing the systematic from 11.5% to 6%, but increasing the statistical uncertainty on the final ΔE data fit from 24% to 31.5% as seen Fig. C.26. Also, the Chebyshev function tends to go negative above $\Delta E > 0.35$ GeV. We decided to keep the background parameters fixed, as we have a good agreement on ΔE in the data-MC sideband comparison (see Fig. 6.39).

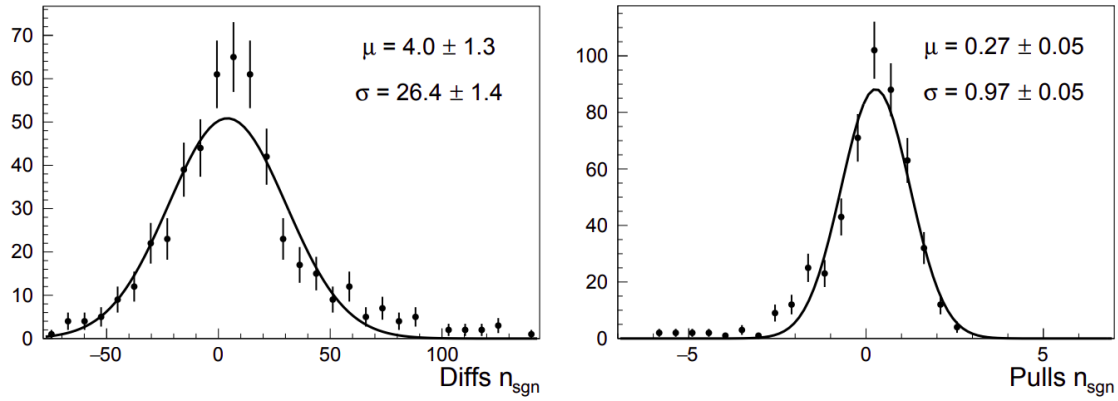

FIGURE C.25: *Differences (see Eq. 6.7) and pulls (see Eq. 6.8) for the signal yield for $B^0 \to K^0_S \pi^0 \gamma$, where the background parameters are left free.*

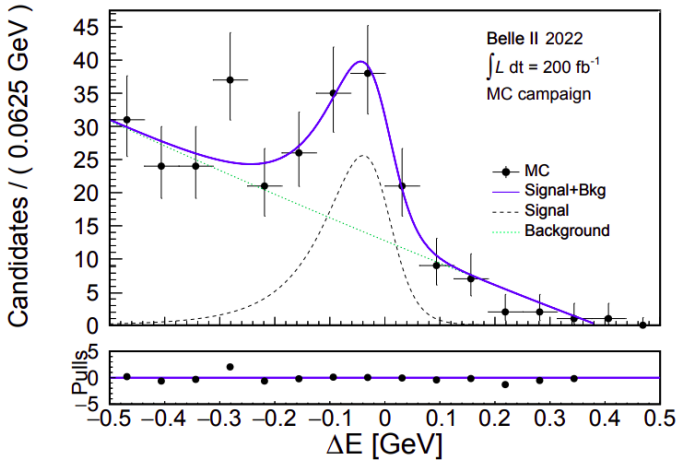

FIGURE C.26: *Fit of the data sample for $B^0 \to K^0_S\pi^0\gamma$. The yields, mean and sigma of the signal function and the two Chebyshev parameters for the background function are let free.*

C.5 Continuum suppression systematic equation

We define the indices as:

$$1 = \text{All other kaonic resonances except } K_{\text{res}}, \tag{C.1}$$
$$2 = K_{\text{res}} \text{ population.} \tag{C.2}$$

The efficiency on one cut in $m_{K^0_S\pi^+\pi^-}$ is expressed as:

$$\epsilon = f_2\epsilon_2 + (1 - f_2)\epsilon_1, \tag{C.3}$$

where f_2 is the fraction of K_{res} events, and ϵ_1 and ϵ_2 are the efficiency of the $m_{K^0_S\pi^+\pi^-}$ cut on 1 and 2.

If we vary the fraction of the number of K_{res} event by a σ amount:

$$\tilde{f_2} = f_2 + \sigma, \tag{C.4}$$

we get a new expression for the efficiency:

$$\tilde{\epsilon} = \tilde{f_2}\epsilon_2 + (1 - f_2)\epsilon_1 - \sigma\epsilon_1 f_2 \tag{C.5}$$

We can then calculate the relative uncertainty on this variables:

$$\frac{\tilde{\epsilon}-\epsilon}{\epsilon} = \frac{\sigma f_2(\epsilon_2-\epsilon_1)}{f_2\epsilon_2+(1-f_2)\epsilon_1} \tag{C.6}$$

$$= \frac{\sigma f_2 \frac{(\epsilon_2-\epsilon_1)}{\epsilon_1}}{1+f_2\frac{(\epsilon_2-\epsilon_1)}{\epsilon_1}} \tag{C.7}$$

$$= \sigma \times \frac{\tilde{\Delta}\epsilon}{1+\tilde{\Delta}\epsilon} \tag{C.8}$$

where

$$\tilde{\Delta}\epsilon = f_2\frac{(\epsilon_2-\epsilon_1)}{\epsilon_1}. \tag{C.9}$$

C.6 Off-resonance data-MC comparison

In this section are shown the off-resonnance, around 18 fb^{-1} of data recorded 60 MeV below $\Upsilon(4S)$ resonance, data-MC comparison, used as a complementary check with the sidebands data-MC comparison.

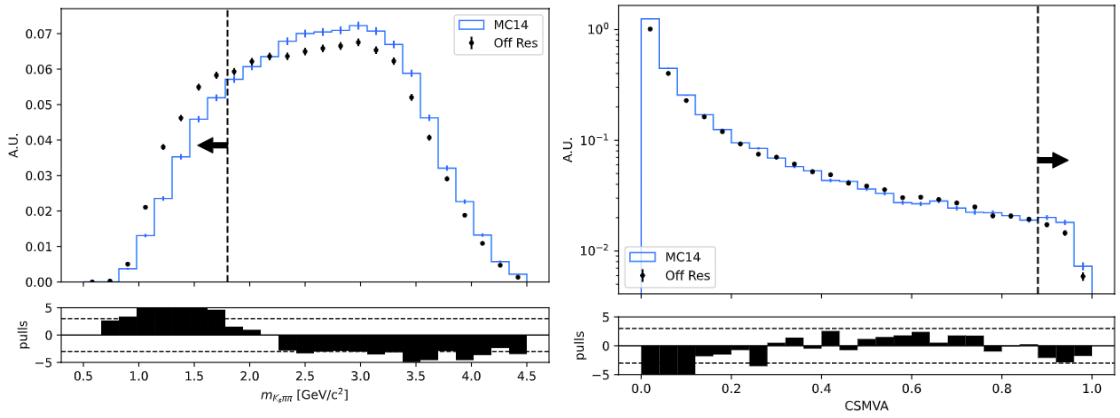

FIGURE C.27: *Normed $m_{K^0_S\pi^+\pi^-}$ (left) and CSMVA output (right) distributions for MC and off-resonance data for $B^0 \to K^0_S\pi^+\pi^-\gamma$. The bar indicates the value of the cut and the arrow is oriented to the kept part.*

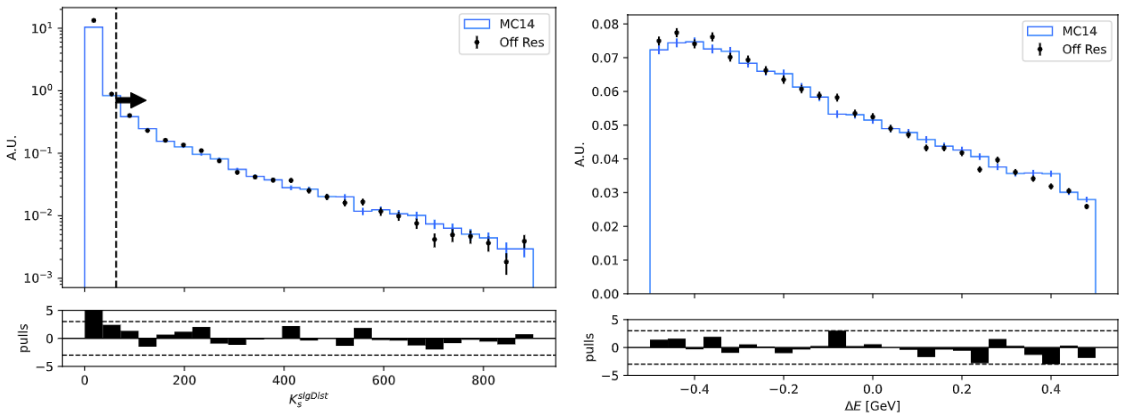

FIGURE C.28: *Normed K_S^0 significance of distance (left) and m_{bc} (right) distributions for MC and off-resonance data for $B^0 \to K_S^0\pi^+\pi^-\gamma$. The bar indicates the value of the cut and the arrow is oriented to the kept part.*

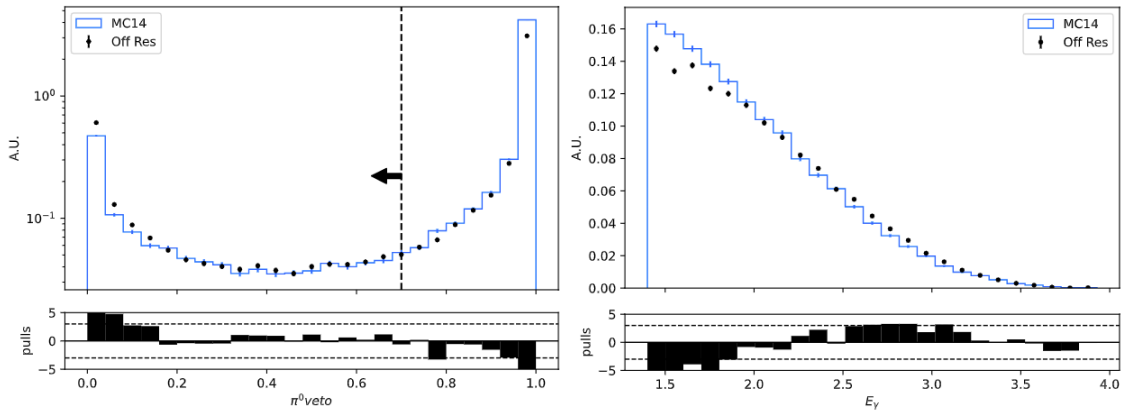

FIGURE C.29: *Normed π^0 veto variable (left) and E_γ (right) distributions for MC and off-resonance data for $B^0 \to K_S^0\pi^+\pi^-\gamma$. The bar indicates the value of the cut and the arrow is oriented to the kept part.*

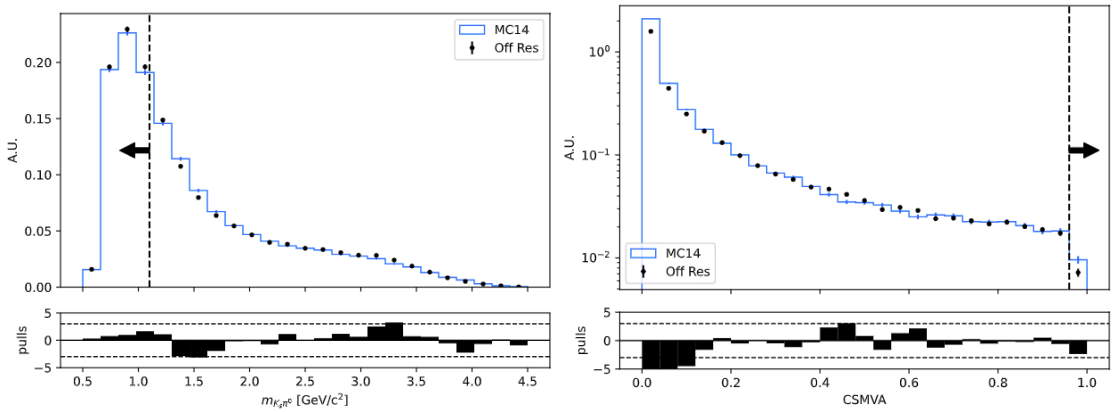

FIGURE C.30: *Normed $m_{K_S^0\pi}$ (left) and CSMVA output (right) distributions for MC and off-resonance data for $B^0 \to K_S^0\pi^0\gamma$. The bar indicates the value of the cut and the arrow is oriented to the kept part.*

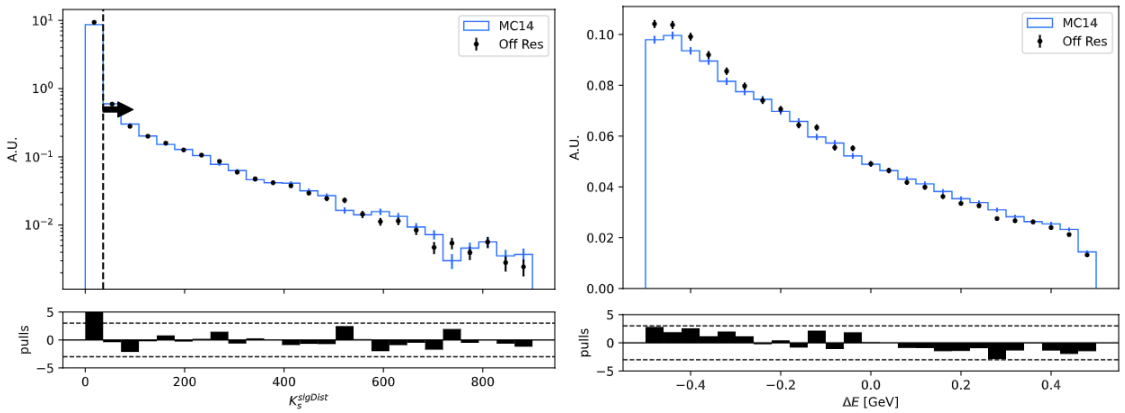

FIGURE C.31: *Normed K_S^0 significance of distance (left) and m_{bc} (right) distributions for MC and off-resonance data for $B^0 \to K_S^0\pi^0\gamma$. The bar indicates the value of the cut and the arrow is oriented to the kept part.*

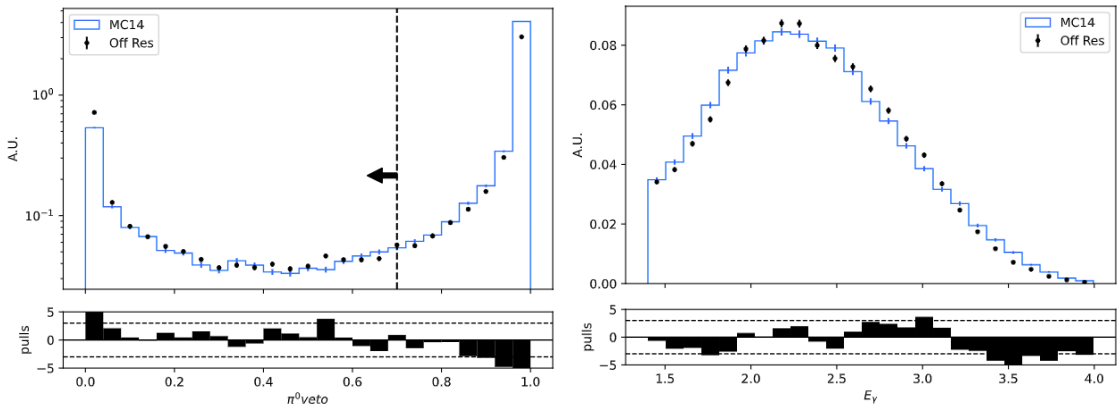

FIGURE C.32: *Normed π^0 veto variable (left) and E_γ (right) distributions for MC and off-resonance data for $B^0 \to K^0_S \pi^0 \gamma$. The bar indicates the value of the cut and the arrow is oriented to the kept part.*

Bibliography

[1] Y. Ushiroda et al., Belle collaboration, *Time-Dependent CP Asymmetries in $B^0 \to K^0_S \pi^0 \gamma$ transitions*, Phys. Rev. D **74** (2006) 111104, arXiv:hep-ex/0608017.

[2] J. Li et al., Belle collaboration, *Time-dependent CP Asymmetries in $B^0 \to K^0_S rho^0 \gamma$ Decays*, Phys. Rev. Lett. **101** (2008) 251601, arXiv:0806.1980 [hep-ex].

[3] B. Aubert et al., BaBar collaboration, *Measurement of Time-Dependent CP Asymmetry in $B^0 \to K^0_S \pi^0 \gamma$ Decays*, Phys. Rev. D **78** (2008) 071102, arXiv:0807.3103 [hep-ex].

[4] P. del Amo Sanchez et al., BaBar collaboration, *Time-dependent analysis of $B^0 \to K^0_S \pi^- \pi^+ \gamma$ decays and studies of the $K^+ \pi^- \pi^+$ system in $B^+ \to K^+ \pi^- \pi^+ \gamma$ decays*, Phys. Rev. D **93** (2016) no. 5, 052013, arXiv:1512.03579 [hep-ex].

[5] D. Blokhintsev and F. Gal'perin, *Neutrino Hypothesis and Conservation of Energy*, Pod Znamenem Marxisma **6** (1934) 147.

[6] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508–509. https://link.aps.org/doi/10.1103/PhysRevLett.13.508.

[7] 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*, Physics Letters B **716** (2012) no. 1, 1–29. https://doi.org/10.1016%2Fj.physletb.2012.08.020.

[8] S. Chatrchyan et al., CMS collaboration, *Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC*, Phys. Lett. B **716** (2012) 30–61, arXiv:1207.7235 [hep-ex].

[9] E. Noether, *Invariante Variationsprobleme*, Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse **1918** (1918) 235–257. http://eudml.org/doc/59024.

[10] T. D. Lee and C. N. Yang, *Question of Parity Conservation in Weak Interactions*, Phys. Rev. **104** (1956) 254–258. https://link.aps.org/doi/10.1103/PhysRev.104.254.

[11] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson, *Experimental Test of Parity Conservation in Beta Decay*, Phys. Rev. **105** (1957) 1413–1415. https://link.aps.org/doi/10.1103/PhysRev.105.1413.

[12] M. Goldhaber, L. Grodzins, and A. W. Sunyar, *Helicity of Neutrinos*, Phys. Rev. **109** (1958) 1015–1017. https://link.aps.org/doi/10.1103/PhysRev.109.1015.

[13] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Evidence for the* 2π *Decay of the* K_2^0 *Meson*, Phys. Rev. Lett. **13** (1964) 138–140. https://link.aps.org/doi/10.1103/PhysRevLett.13.138.

[14] M. Kobayashi and T. Maskawa, *CP-Violation in the Renormalizable Theory of Weak Interaction*, Progress of Theoretical Physics **49** (1973) no. 2, 652–657, https://academic.oup.com/ptp/article-pdf/49/2/652/5257692/49-2-652.pdf. https://doi.org/10.1143/PTP.49.652.

[15] K. Abe et al., Belle Collaboration, *Observation of Large CP Violation in the Neutral B Meson System*, Phys. Rev. Lett. **87** (2001) 091802. https://link.aps.org/doi/10.1103/PhysRevLett.87.091802.

[16] B. Aubert and other, BaBar Collaboration, *Observation of CP Violation in the* B^0 *Meson System*, Phys. Rev. Lett. **87** (2001) 091801. https://link.aps.org/doi/10.1103/PhysRevLett.87.091801.

[17] A. D. Sakharov, *Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe*, Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32–35.

[18] M. Gavela, P. Hernandez, J. Orloff, O. Péne, and C. Quimbay, *Standard model CP-violation and baryon asymmetry (II). Finite temperature*, Nuclear Physics B **430** (1994) no. 2, 382–426. https://doi.org/10.1016%2F0550-3213%2894%2900410-2.

[19] M. Gronau, Y. Grossman, D. Pirjol, and A. Ryd, *Measuring the Photon Polarization in* $B \to K\pi\pi\gamma$, Phys. Rev. Lett. **88** (2002) 051802. https://link.aps.org/doi/10.1103/PhysRevLett.88.051802.

[20] R. Aaij et al., LHCb Collaboration, *Observation of Photon Polarization in the* $b \to s\gamma$ *Transition*, Phys. Rev. Lett. **112** (2014) 161801. https://link.aps.org/doi/10.1103/PhysRevLett.112.161801.

[21] L.-L. Chau and W.-Y. Keung, *Comments on the Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. **53** (1984) no. 19, 1802.

[22] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. **51** (1983) 1945–1947. https://link.aps.org/doi/10.1103/PhysRevLett.51.1945.

[23] P. Zyla et al., Particle Data Group, *Review of Particle Physics*, PTEP **2020** (2020) no. 8, 083C01. and 2021 update.

[24] J. Charles et al., *Current status of the Standard Model CKM fit and constraints on* $\Delta F = 2$ *New Physics*, Phys. Rev. D **91** (2015) no. 7, 073007, arXiv:1501.05013 [hep-ph].

[25] K. Fujikawa and A. Yamada, *Test of the chiral structure of the top-bottom charged current by the process $b \to s\gamma$*, Phys. Rev. D **49** (1994) 5890–5893. https://link.aps.org/doi/10.1103/PhysRevD.49.5890.

[26] K. S. Babu, K. Fujikawa, and A. Yamada, *Constraints on left-right symmetric models from the process $b \to s$ gamma*, Phys. Lett. B **333** (1994) 196–201, arXiv:hep-ph/9312315.

[27] P. L. Cho and M. Misiak, *$b \to s$ gamma decay in SU(2)-L x SU(2)-r x U(1) extensions of the Standard Model*, Phys. Rev. D **49** (1994) 5894–5903, arXiv:hep-ph/9310332.

[28] T. Inami and C. S. Lim, *Effects of Superheavy Quarks and Leptons in Low-Energy Weak Processes $KL \to \mu\mu$, $K+ \to \pi + \nu\nu$ and $K0 \leftrightarrow K0$*, Progress of Theoretical Physics **65** (1981) no. 1, 297–314, https://academic.oup.com/ptp/article-pdf/65/1/297/5252099/65-1-297.pdf. https://doi.org/10.1143/PTP.65.297.

[29] T. Saito et al., Belle collaboration, *Measurement of the $\bar{B} \to X_s\gamma$ Branching Fraction with a Sum of Exclusive Decays*, Phys. Rev. D **91** (2015) no. 5, 052004, arXiv:1411.7198 [hep-ex].

[30] Belle collaboration, A. Abdesselam et al., *Measurement of the inclusive $B \to X_{s+d}\gamma$ branching fraction, photon energy spectrum and HQE parameters*, in *38th International Conference on High Energy Physics*. 8, 2016. arXiv:1608.02344 [hep-ex].

[31] B. Aubert et al., BaBar collaboration, *Measurement of the $B \to X_s\gamma$ branching fraction and photon energy spectrum using the recoil method*, Phys. Rev. D **77** (2008) 051103, arXiv:0711.4889 [hep-ex].

[32] J. P. Lees et al., BaBar collaboration, *Precision Measurement of the $B \to X_s\gamma$ Photon Energy Spectrum, Branching Fraction, and Direct CP Asymmetry $A_{CP}(B \to X_{s+d}\gamma)$*, Phys. Rev. Lett. **109** (2012) 191801, arXiv:1207.2690 [hep-ex].

[33] J. P. Lees et al., BaBar collaboration, *Exclusive Measurements of $b \to s\gamma$ Transition Rate and Photon Energy Spectrum*, Phys. Rev. D **86** (2012) 052012, arXiv:1207.2520 [hep-ex].

[34] R. Aaij et al., LHCb collaboration, *Measurement of CP-violating and mixing-induced observables in $B_s^0 \to \phi\gamma$ decays*, Phys. Rev. Lett. **123** (2019) no. 8, 081802, arXiv:1905.06284 [hep-ex].

[35] R. Aaij et al., LHCb collaboration, *Strong constraints on the $b \to s\gamma$ photon polarisation from $B^0 \to K^{*0}e^+e^-$ decays*, JHEP **12** (2020) 081, arXiv:2010.06011 [hep-ex].

[36] I. I. Bigi and A. I. Sanda, *CP Violation*, vol. 2. Cambridge University Press, 2009.

[37] S. Akar, *Etudes des désintégrations $B \to K\pi^+\pi^-\ \gamma$ avec l'expérience BABAR : hélicité du photon et structure résonante du système $K\pi^+\pi^-$*. PhD thesis, 2013. http://www.theses.fr/2013PA066399/document. Thèse de doctorat dirigée par Ben-Haim, Eli Particules. Notaux. Cosmologie Paris 6 2013.

[38] K. Lande, E. T. Booth, J. Impeduglia, L. M. Lederman, and W. Chinowsky, *Observation of Long-Lived Neutral V Particles*, Phys. Rev. **103** (1956) 1901–1904. https://link.aps.org/doi/10.1103/PhysRev.103.1901.

[39] H. Albrecht et al., *Observation of B0-B0 mixing*, Physics Letters B **192** (1987) no. 1, 245–252. https://www.sciencedirect.com/science/article/pii/0370269387911774.

[40] A. Abulencia et al., CDF collaboration, *Measurement of the $B_s^0 - \bar{B}_s^0$ Oscillation Frequency*, Phys. Rev. Lett. **97** (2006) 062003, arXiv:hep-ex/0606027.

[41] R. Aaij et al., LHCb collaboration, *Observation of B_s^0-$\bar{B}_s^0$ mixing and measurement of mixing frequencies using semileptonic B decays*, Eur. Phys. J. C **73** (2013) no. 12, 2655, arXiv:1308.1302 [hep-ex].

[42] B. Aubert et al., BaBar collaboration, *Evidence for $D^0 - \overline{D}^0$ Mixing*, Phys. Rev. Lett. **98** (2007) 211802, arXiv:hep-ex/0703020.

[43] M. Staric et al., Belle collaboration, *Evidence for D^0 - $\bar{D}^0$ Mixing*, Phys. Rev. Lett. **98** (2007) 211803, arXiv:hep-ex/0703036.

[44] R. Aaij et al., LHCb collaboration, *Measurement of $D^0\bar{D}^0$ Mixing Parameters and Search for CP Violation Using $D^0 \to K^+\pi^-$ Decays*, Phys. Rev. Lett. **111** (2013) no. 25, 251801, arXiv:1309.6534 [hep-ex].

[45] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry*, Phys. Rev. D **2** (1970) 1285–1292. https://link.aps.org/doi/10.1103/PhysRevD.2.1285.

[46] Y. Amhis et al., HFLAV, *Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2016*, Eur. Phys. J. C **77** (2017) no. 12, 895, arXiv:1612.07233 [hep-ex].

[47] J. Charles et al., *Predictions of selected flavour observables within the Standard Model*, Phys. Rev. D **84** (2011) 033005, arXiv:1106.4041 [hep-ph].

[48] H. Kakuno et al., Belle collaboration, *Neutral B flavor tagging for the measurement of mixing induced CP violation at Belle*, Nucl. Instrum. Meth. A **533** (2004) 516–531, arXiv:hep-ex/0403022.

[49] M. Matsumori and A. I. Sanda, *Mixing-induced CP asymmetry in $B \to K * \gamma$ decays with perturbative QCD approach*, Physical Review D **73** (2006) no. 11, . https://doi.org/10.1103%2Fphysrevd.73.114022.

[50] P. Ball and R. Zwicky, *Time-dependent CP asymmetry in $B \to K * \gamma$ as a (quasi)null test of the Standard Model*, Physics Letters B **642** (2006) no. 5-6, 478–486. https://doi.org/10.1016%2Fj.physletb.2006.10.013.

[51] D. Atwood, T. Gershon, M. Hazumi, and A. Soni, *Mixing-induced CP violation in $B \to P(1)\, P(2)$ gamma in search of clean new physics signals*, Phys. Rev. D **71** (2005) 076003, arXiv:hep-ph/0410036.

[52] S. Akar, E. Ben-Haim, J. Hebinger, E. Kou, and F.-S. Yu, *The time-dependent CP asymmetry in $B^0 \to K_{\rm res}\gamma \to \pi^+\pi^- K^0_S\gamma$ decays*, JHEP **09** (2019) 034, arXiv:1802.09433 [hep-ph].

[53] Belle II collaboration, F. Forti, *Snowmass Whitepaper: The Belle II Detector Upgrade Program*, in *2022 Snowmass Summer Study*. 3, 2022. arXiv:2203.11349 [hep-ex].

[54] P. M. Lewis et al., *First Measurements of Beam Backgrounds at SuperKEKB*, Nucl. Instrum. Meth. A **914** (2019) 69–144, arXiv:1802.01366 [physics.ins-det].

[55] W. Altmannshofer et al., Belle II collaboration, *The Belle II Physics Book*, PTEP **2019** (2019) no. 12, 123C01, arXiv:1808.10567 [hep-ex]. [Erratum: PTEP 2020, 029201 (2020)].

[56] V. Bertacchi et al., *Track finding at Belle II*, Computer Physics Communications **259** (2021) 107610. https://www.sciencedirect.com/science/article/pii/S0010465520302861.

[57] Belle-ECL, *Electromagnetic calorimeter for Belle II*, . https://doi.org/10.1088/1742-6596/587/1/012045.

[58] F. Luetticke et al., *The ultralight DEPFET pixel detector of the Belle II experiment*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **845** (2017) 118–121. https://www.sciencedirect.com/science/article/pii/S0168900216306799. Proceedings of the Vienna Conference on Instrumentation 2016.

[59] T. Abe et al., Belle II collaboration, *Belle II Technical Design Report*, arXiv:1011.0352 [physics.ins-det].

[60] M. French et al., *Design and results from the APV25, a deep sub-micron CMOS front-end chip for the CMS tracker*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **466** (2001) no. 2, 359–365. https://www.sciencedirect.com/science/article/pii/S0168900201005897. 4th Int. Symp. on Development and Application of Semiconductor Tracking Detectors.

[61] S. Iwata et al., *Particle identification performance of the prototype Aerogel RICH counter for the Belle II experiment*, Progress of Theoretical and Experimental Physics **2016** (2016) 033H01.

[62] K. Miyabayashi, *Belle II electromagnetic calorimeter and its performance during early SuperKEKB operation*, Journal of Instrumentation **15** (2020) no. 10, C10016–C10016. https://doi.org/10.1088/1748-0221/15/10/c10016.

[63] S. Kim, I. Lee, Y. Unno, and B. Cheon, *Status of the electromagnetic calorimeter trigger system at Belle II.*, Journal of Physics: Conference Series **928** (2017) 012022.

[64] W. Snoeys, *Monolithic pixel detectors for high energy physics*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **731** (2013) 125–130. https://www.sciencedirect.com/science/article/pii/S0168900213006840. PIXEL 2012.

[65] L. Flores Sanz de Acedo et al., *Design of large scale sensors in 180 nm CMOS process modified for radiation tolerance*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **980** (2020) 164403. https://www.sciencedirect.com/science/article/pii/S0168900220308007.

[66] N. Wermes, *Monopix - radiation hard monolithic CMOS pixel detectors*, CERN detector seminar (2020) . https://indico.cern.ch/event/884089/.

[67] T. Kuhr et al., Belle II collaboration, *The Belle II Core Software*, Comput. Softw. Big. Sci. **3** (2018) , arXiv:1809.04299 [physics.comp-ph].

[68] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, Nucl. Instrum. Meth. A **389** (1997) 81–86.

[69] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. A **462** (2001) 152–155.

[70] T. Sjostrand et al., *A brief introduction to PYTHIA 8.1*, Comp. Phys. Comm. **178** (2008) 852–867.

[71] S. Agostinelli et al., GEANT4, *GEANT4–a simulation toolkit*, Nucl. Instrum. Meth. A **506** (2003) 250–303.

[72] T. Alexopoulos, M. Bachtis, E. Gazis, and G. Tsipolitis, *Implementation of the Legendre Transform for track segment reconstruction in drift tube chambers*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **592** (2008) no. 3, 456–462. https://www.sciencedirect.com/science/article/pii/S0168900208005780.

[73] T. Bilka et al., *Implementation of GENFIT2 as an experiment independent track-fitting framework*, arXiv:1902.04405 [physics.data-an].

[74] T. Keck, *FastBDT: A speed-optimized and cache-friendly implementation of stochastic gradient-boosted decision trees for multivariate classification*, arXiv e-prints (2016) arXiv:1609.06119, arXiv:1609.06119 [cs.LG].

[75] R. Frühwirth, *Application of Kalman filtering to track and vertex fitting*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **262** (1987) no. 2, 444–450. https://www.sciencedirect.com/science/article/pii/0168900287908874.

[76] N. Braun, *Combinatorial Kalman Filter and High Level Trigger Reconstruction for the Belle II Experiment*. Springer Theses, 01, 2019.

[77] M. Ritter, *Measurement of the branching fraction and time dependent CP asymmetry in B0 → D*-D*+K0s decays at the belle experiment*. Ludwig-Maximilians-Universität München, 2014. http://nbn-resolving.de/urn:nbn:de:bvb:19-172823.

[78] H. Bichsel, *Straggling in thin silicon detectors*, Rev. Mod. Phys. **60** (1988) 663–699. https://link.aps.org/doi/10.1103/RevModPhys.60.663.

[79] T. Fillinger, *Detector Simulation for a Potential Upgrade of the Vertex Detector of the Belle II Experiment*, Acta Phys. Pol. **B52** (2021) 909.

[80] L. Massaccesi, *Simulation of an all-layer monolithic pixel vertex detector for the Belle II upgrade*, PM2021 - 15th Pisa Meeting on Advanced Detectors (2022) . https://agenda.infn.it/event/22092/contributions/166670/attachments/90806/122531/poster_Massaccesi_15thPisaMeeting_V4.pdf.

[81] P. Stavroulakis, *Optimisation of an upgrade to the Belle II experiment's inner tracker using simulated $B^0 \to K^0_S \pi^+ \pi^- \gamma$ decays*, Belle II master thesis (internal document) (2022) . https://docs.belle2.org/record/3089?ln=en.

[82] R. Itoh, Belle collaboration, *BELLE AnalysiS Framework*, 9th International Conference on Computing in High-Energy Physics (1997) . http://www.ifh.de/CHEP97/paper/244.ps.

[83] P. Koppenburg, Belle collaboration, *An improved π^0 and η veto*, Belle Note 665 (2003) .

[84] H. Nakano et al., Belle collaboration, *Ks selection with NeuroBayes and nisKsFinder class*, Belle Note 1253 (2012) .

[85] R. Mankel, *Pattern recognition and event reconstruction in particle physics experiments*, Reports on Progress in Physics **67** (2004) no. 4, 553–622. https://doi.org/10.1088/0034-4885/67/4/r03.

[86] H. Tajima et al., *Proper time resolution function for measurement of time evolution of B mesons at the KEK B factory*, Nucl. Instrum. Meth. A **533** (2004) 370–386, arXiv:hep-ex/0301026.

[87] T. Higuchi, *Vertexing*, . https://kds.kek.jp/event/3808/contributions/126511/attachments/97844/116455/BAS2009.pdf.

[88] M. Röhrken, *Time-Dependent CP Violation Measurements in Neutral B Meson to Double-Charm Decays at the Japanese Belle Experiment*, .

[89] S. H. Lee et al., Belle Collaboration, *Evidence for $B^0 \to \pi^0\pi^0$*, Phys. Rev. Lett. **91** (2003) 261801. https://link.aps.org/doi/10.1103/PhysRevLett.91.261801.

[90] W. Verkerke and D. Kirkby, *The RooFit toolkit for data modeling*, . https://arxiv.org/abs/physics/0306116.

[91] F. James and M. Roos, *Minuit - a system for function minimization and analysis of the parameter errors and correlations*, Computer Physics Communications **10** (1975) no. 6, 343–367. https://www.sciencedirect.com/science/article/pii/0010465575900399.

[92] H. Albrecht et al., *Search for hadronic $b \to u$ decays*, Physics Letters B **241** (1990) no. 2, 278–282.

[93] T. Skwarnicki, *A study of the radiative CASCADE transitions between the Upsilon-Prime and Upsilon resonances*. PhD thesis, Cracow, INP, 1986. https://inspirehep.net/literature/230779.

[94] B. Kronenbitter and M. Prim, *The Belle "Fit Library Collection" - About libRooComplexPDF, libRooTatami and others*, . https://kds.kek.jp/event/9641/contributions/154048/attachments/122377/145198/BAS_2012_05_11.pdf.

[95] K. Sumisawa, *Time-dependent CP Violation*, . https://kds.kek.jp/event/9641/contributions/154059/attachments/122374/145192/BAS2012_2_sumisawa.pdf.

[96] T. Higuchi, Belle collaboration, *Resolution Function Update for ICHEP06*, Belle Note 924b, internal document (2009) .

[97] T. Fillinger et al., Belle collaboration, *Dalitz-separated analysis of the Time-dependent CP asymmetries in $B^0 \to K^0_S\pi^+\pi^-\gamma$ transitions*, Belle Note 1524, internal document (2020) .

[98] B. F. L. Ward, S. Jadach, and Z. Was, *Precision calculation for e+ e- $\to$ 2f: The KK MC project*, Nucl. Phys. B Proc. Suppl. **116** (2003) 73–77, arXiv:hep-ph/0211132.

[99] C. Cecchi et al., Belle II collaboration, internal, *B counting measurement in "Moriond 2022" Belle II dataset*, BELLE2-NOTE-PH-2022-007 (2022) . https://docs.belle2.org/record/2846/.

[100] J.-F. Krohn et al., Belle II collaboration, *Global decay chain vertex fitting at Belle II*, Nucl. Instrum. Meth. **A976** (2020) 164269, arXiv:1901.11198 [hep-ex].

[101] J. Tanaka, Belle collaboration, *Kinematic Fitting*, Belle Note 194, internal document (2000) .

[102] D. M. Asner et al., *Search for exclusive charmless hadronic B decays*, Phys. Rev. **D53** (1996) 1039.

[103] N. L. Johnson, *Systems of Frequency Curves Generated by Methods of Translation*, Biometrika **36** (1949) 149–176.

[104] F. James, *MINUIT: Function Minimization and Error Analysis Reference Manual*, CERN-D-506 (1994) .

[105] T. Humair, *CP Violation in charmless decays at Belle2*, Rencontres de Moriond (2022) . https://moriond.in2p3.fr/2022/EW/slides/5/1/7_THumair-v1.pdf.

[106] F. Abudinén et al., Belle II collaboration, *Measurement of the branching fraction of the $B^0 \to K^0_S \pi^0 \gamma$ decay using 190 fb^{-1} of Belle II data*, arXiv:2206.08280 [hep-ex].

[107] G. Veneziano, *Towards the measurement of photon polarisation in the decay $B^+ \to K^+ \pi^+ \pi^- \gamma$*. PhD thesis, 2016.

[108] V. Bellée, *Amplitude analysis for the measurement of the photon polarisation in $B \to K\pi^+\pi^- \gamma$ decays*. PhD thesis, 2020.

[109] B. Knysh, *Analysis of the $B^+ \to K^+ \pi^+ \pi^- \gamma$ decay with early Belle II data*. PhD thesis, 2021. https://tel.archives-ouvertes.fr/tel-03538092/document. High Energy Physics - Experiment [hep-ex]. Université Paris-Saclay, 2021. English. ⟨NNT : 2021UPASP101⟩. ⟨tel-03538092⟩.

[110] T. Fillinger et al., *Rediscovery and branching fraction measurements for $B^0 \to K^0_S \pi^+ \pi^- \gamma$ and $B^0 \to K^0_S \pi^0 \gamma$*, Belle II Note 2021-05, internal document (2022) .