Analysis of the $B^+ \to K^+\pi^+\pi\,\gamma$ decay with early Belle II data

Borys Knysh

université
PARIS-SACLAY

Analysis of the $B^+ \to K^+\pi^+\pi^-\gamma$ decay with early Belle II data

Analyse de la désintégration $B^+ \to K^+\pi^+\pi^-\gamma$ avec les premières données Belle II

Thèse de doctorat de l'université Paris-Saclay

École doctorale n° 576, Particules, hadrons, énergie et noyau : instrumentation, imagerie, cosmos et simulation
Spécialité de doctorat : Physique des particules
Graduate School : Physique. Référent : Faculté des sciences d'Orsay

Thèse préparée dans l'unité de recherche **IJCLab** (Université Paris-Saclay, CNRS), sous la direction de **François LE DIBERDER**, Professeur des universités, et le co-encadrement d'**Emi KOU**, Directrice de recherche.

Thèse soutenue à Paris-Saclay, le 7 octobre 2021, par

Borys KNYSH

Composition du jury

Marie-Hélène SCHUNE Directrice de recherche, IJCLab	Présidente
Eli BEN-HAIM Professeur, Sorbonne Université	Rapporteur & Examinateur
Arantza OYANGUREN CAMPOS Professeure, IFIC de Valencia	Rapportrice & Examinatrice
Elisa MANONI Directrice de recherche, INFN de Perugia	Examinatrice
François LE DIBERDER Professeur, Université Paris-Diderot	Directeur de thèse

THESE DE DOCTORAT
NNT : 2021UPASP101

université PARIS-SACLAY

ÉCOLE DOCTORALE
Particules, hadrons, énergie et noyau : instrumentation, imagerie, cosmos et simulation (PHENIICS)

Titre: Analyse de la désintégration $B^+ \to K^+\pi^+\pi^-\gamma$ avec les premières données Belle II
Mots clés: Modèle standard et au-delà, Désintégration faible radiative des mésons B, Analyse des données.

Résumé: Le document présente l'analyse de la désintégration du méson B^+ dans l'état final $K^+\pi^+\pi^-\gamma$ en utilisant les données initiales de l'expérience Belle II. Le détecteur Belle II est situé auprès du collisionneur SuperKEKB dans le laboratoire KEK, au Japon.

Le but de cette thèse est d'ouvrir la voie à la mesure de la polarisation du photon γ présent dans l'état final. Le modèle standard de la physique des particules prédit que le photon est pratiquement complètement polarisé gauche. Une mesure de la polarisation en désaccord avec cette prédiction signalerait directement l'intervention d'une physique au-delà du Modèle Standard.

L'analyse repose sur le logiciel de simulation Monte-Carlo "GamPola" qui a été développé par l'auteur pour incorporer l'état de l'art de la compréhension théorique des désintégrations des mésons B neutres et chargés dans les divers états finals $K\pi\pi\gamma$.

Les données utilisées correspondent à une luminosité de $62.7\ \mathrm{fb}^{-1}$ qui a été accumulée par l'expérience Belle II jusqu'au printemps 2021. Cette luminosité étant environ 50 fois plus faible que celle nécessaire pour aborder la mesure de la polarisation du photon, d'une part, et d'autre part, le détecteur Belle II n'étant pas encore complètement maitrisé à ce stade préliminaire, cette analyse initiale est focalisée sur la mise au point d'une procédure de sélection robuste des événements $K^+\pi^+\pi^-\gamma$. Cette sélection pourra être utilisée plus tard comme point de départ pour conduire à la mesure de la polarisation du photon.

Une efficacité de sélection de $25\ \%$ est obtenue tout en maintenant les événements de bruit de fond à un niveau suffisamment bas pour permettre une mesure du rapport d'embranchement avec une précision statistique de $10\ \%$. Une approche plus raffinée, basée sur une segmentation de l'espace de phase qui réduit fortement la dépendance dans le modèle théorique utilisé, conduit à estimer qu'une précision statistique de $3\ \%$ sera atteinte pour $1\ \mathrm{ab}^{-1}$. Comme étape préliminaire en vue de la mesure de la polarisation, il est montré que l'asymétrie A_{ud} sera possible avec une précision du pourcent pour $1\ \mathrm{ab}^{-1}$. L'analyse des données est encore en mode aveugle à ce stade, mais les données de contrôles ont pu être dévoilées.

Title: Analysis of the $B^+ \to K^+\pi^+\pi^-\gamma$ decay with early Belle II data.
Keywords: Standard Model and beyond, Weak radiative decays of B mesons, Data analysis.

Abstract: The document presents the analysis of the decay of the B^+ meson into the $K^+\pi^+\pi^-\gamma$ final state using the early data of the Belle II experiment, that is located at the SuperKEKB collider in the KEK laboratory, Japan.

The aim of the present work is to pave the way leading to the measurement of the polarization of the photon γ involved in the decay. The Standard Model of particle physics predicts the photon polarization to be overwhelmingly left-handed. A measurement of the polarization away from this prediction would provide a direct indication of Physics beyond the Standard Model.

The analysis relies on a detailed Monte-Carlo generator named "GamPola" that was developed by the author for this purpose to include state of the art theoretical understanding of the charged and neutral B decays into the various possible $K\pi\pi\gamma$ final states.

The data sample used in the analysis corresponds to a luminosity of $62.7\ \mathrm{fb}^{-1}$ that was accumulated by the Belle II experiment until spring 2021. This luminosity being about 50 times smaller than what is needed to perform the measurement of the photon polarization, and the Belle II detector being not yet fully understood in this early stage, the aim of the initial analysis is focused on defining a robust procedure to select the $K^+\pi^+\pi^-\gamma$ events. This selection procedure can be used later as a starting point to perform the photon polarization measurement.

An overall selection efficiency of $25\ \%$ is achieved while maintaining the contribution of background events to a low enough level allowing a $10\ \%$ statistical precision on the branching ratio measurement. A refined model-independent approach based on phase-space binning leads to a $3\ \%$ precision for $1\ \mathrm{ab}^{-1}$. As a preliminary step towards the photon polarization, the measurement of the up-down asymmetry A_{ud} is assessed to reach the percent level, for $1\ \mathrm{ab}^{-1}$. While the analysis is still blind at the present stage, the unblinding of the control samples has been performed.

Acknowledgements

In the beginning, I would like to thank Laboratoire de Physique des 2 Infinis Irène Joliot Curie, Université Paris-Saclay, and PHENIICS doctoral school for the hospitality and the possibility for accomplishing this thesis. I express my gratitude to the Belle II Collaboration for taking me as part of such a large Physics Society.

I am grateful to my supervisor Francois Le Diberder without whom this work will not be possible. Thank you for the proposed topic and such a great opportunity to dive into the scientific world, for your patience and huge efforts made while supporting me, for the knowledge you transferred, and for believing in me.

I would like to sincerely thank Emi Kou, who was guiding, supporting, and helping me during the whole period of three years. In particular, your expertise greatly influenced building the theoretical part of this thesis, and your help in that domain is invaluable.

Many thanks I am delighted to express to Karim Trabelsi for directing and motivating me in the performing of this work. The general approach I learned from you of looking at the subject from a new angle and from the point of view of the "third dimension" helped me a lot to finish this work and will become credo in my whole professional career.

I want especially to thank Elisa Manoni and Saurabh Sandilya, the Conveners of EWP group, for your advice and recommendations concerning the analysis and the feedbacks which inspired the development of many things included in this thesis.

The topic covered in this work is not only conceptually complicated, yet technically. I am willing to express my appreciation to Sviatoslav Bilokin and Shun Watanuki-san for getting me familiar with the software and computational parts of the analysis, for your prompt answers to my questions.

Towards finalizing the document a huge contribution was obtained from the Jury Members. I thank you for your comments and the questions you asked: they all were very valuable and greatly increased the maturity of the manuscript.

I thank Leonid Burmistrov, Dmytro Hohov, Oleksandr Hryhorenko for giving help hand in hard times and your company, that I enjoined a lot.

I am grateful to my parents, who bore, rose, and supported me at all times. The infinite patience, support, and understanding of my beloved wife played a tremendous role in reaching this stage, and am grateful to you for this. It is such a milestone for both of us.

Apologize to those who have not been mentioned among these lines. Thank you to all of you!

for Lili

Contents

Introduction

Particle Physics has rapidly arisen as a direction of Physics in XX-th century [1]. Theoretical [2–11] and experimental [12–19] results have been combined and the Standard Model has appeared. The theory was extending and capturing description of electromagnetic, strong and weak fundamental interations and the mechanism of mass generation of elementary particles [8, 9].

It has been tested with a high precision [20] and treated as a baseline to compare with. If performed measurement deviates from the Standard Model prediction with the predefined confidence level, it can be evidence of the New Physics.

Still, there are certain known topics that the Standard Model doesn't address. For instance, a theoretical description of gravitational forces is not included in it, though attempts to create high-level theory, which would include all four interactions are being made [21–24]. Also, physicists understand that nearly $95\ \%$ of the Universe is not made of ordinary matter. Instead, much of the Universe consists of dark matter and dark energy that does not fit into the Standard Model. Also, the fact that the Universe is made predominantly of matter and not anti-matter is not explained.

High energy physics experiments are performed to solve the above problems and search for New Physics. Dark matter searches can be classified as indirect, direct, and collider experiments. Indirect approaches search for signs of dark matter annihilation [25–27]. Within direct experiments, the background level of 1 event per year per tonne of detecting material is reached by conducting measurements deeply under a layer of ground or water [28–31]. Such a low level of background is required to increase the sensitivity for the new phenomena. The evidence of the presence of dark matter at colliders can be missing transversal energy of the reconstructed particles [32].

At Large Hadron Collider (LHC), the energy frontier experiments directly search for the presence of massive particles produced in proton-proton collisions at a center of mass energy of up to $14\ \mathrm{TeV}$. For Future Circular Collider [33] energies will be pushed even further (up to $100\ \mathrm{TeV}$), allowing a much broader search for New Physics. Purely intensity frontier experiments take advantage of the fact, that sensitivity to the direct production of a specific new particle depends on the cross section and on the size of the data sample. Within this approach, the signatures of new particles or processes can be observed through measurements of suppressed flavour physics reactions or from deviations from SM predictions. For instance, B-factories [34] have proven their feasibility and discovered CP-violation [35, 36] in the B-meson decays. The Belle II Experiment [37] with 50 times larger integrated luminosity and substantial upgrade of detector system compared to Belle, not only will improve the measurements conducted by predecessor but, in addition, allow for a broader search of New Physics.

Rare $b \to s\gamma$ flavour-changing neutral-current transitions are expected to be sensitive to NP effects that may arise from the exchange of heavy fermions in the electroweak penguin loop. The emission of right-handed photons is suppressed by a factor $\frac{m_s}{m_b}$. This suppression can be relieved in some New Physics models resulting in a factor of $\frac{m_{NP}}{m_b}$ instead of $\frac{m_s}{m_b}$. These effects might result in the enhancement of the photon right-handed component, whereas the SM predicts it to be mostly left-handed.

The polarisation of the photon, accessible through the study of radiative $b \to s\gamma$ decays, is one of the predictions of the SM that have been searched for during the last decades [39–48]. The most recent result from late 2020 produced by LHCb through angular analysis has put very strong constraints on the photon

polarization parameter with 5 % of precision [48]. The results given in the current work don't aim for comparing with LHCb, instead, they demonstrate another approach that can be used in the future of Belle II given a significant amount of data.

The process $b \to s\gamma$ can be hadronized in several ways. For instance, the possibility is $B \to K_{res}\gamma \to K\pi\pi\gamma$, where an interesting hadronic structure can be displayed in the K_{res} spectrum: there are contributions from several kaonic resonances decaying to $K\pi\pi$. The decays of these resonances themselves exhibit a resonant structure, with $K^*\pi$, $K\rho$ combinations.

This thesis presents a study of $B \to K\pi\pi\gamma$ channel using two approaches. The first one is the model-dependent approach, which is implemented in "GamPola" software, incorporating generator and fitter of $B \to K\pi\pi\gamma$ events. Second one is analysis of $B^+ \to K^+\pi^+\pi^-\gamma$ decay, with $K^+\pi^+\pi^-$ mass being restricted to < 1.8 GeV/c^2 using 1 ab^{-1} of Monte-Carlo data, 63 fb^{-1} ($\Upsilon(4S)$ center-of-mass energy) and 9.2 fb^{-1} (off-resonance) data collected with the Belle II detector. The analysis presented in this thesis is applied to the control samples — samples of real data, not containing signal contribution.

In Chapter 1 more details about Standard Model are given: continuous and discrete symmetries, gauge invariance of SM Lagrangian, mass generation of gauge bosons using Higgs mechanism, weak and strong interactions and potential sources of New Physics contributions. The second part of the chapter lists state-of-art (SOA) approaches of measuring the photon polarization along with experimental status.

Chapter 2 addresses a question of how to measure photon polarization in $B \to K\pi\pi\gamma$ decay. The approach is formalized in the software package "GamPola" (Gamma Polarization). It includes the modelling of the decay [49,50], and in particular, incorporates several resonances representing the K_{res} above. Based on the model and its adjustable parameters, "GamPola"-generator produces $B \to K\pi\pi\gamma$ events. Given a set of $B \to K\pi\pi\gamma$ generated events, it is described how "GamPola"-fitter performs the inverse procedure of extracting the model parameters using likelihood fit. An approach for normalizing probability density function efficiently within the likelihood function is covered. Being simple conceptually, but sophisticated technically it allows also to reduce the computational time complexity during the fit. Based on generator and fitter the sensitivity study is discussed. It includes the Baseline Model definition based on the plots of LHCb data. The impact of the invariant mass of $K\pi\pi$ system on photon polarization measurement is addressed. Photon polarization sensitivity for different decay modes and generated data samples is presented. Being multi-dimensional, "GamPola"-fitter is probed on the subject of stability using a set of events generated using the Baseline Model. The estimation of required integrated luminosity for competitive precision of photon polarization measurement is performed.

In the Chapter 3 the SuperKEKB collider [117] is described. The upgrade with respect to KEKB and corresponding improvements are highlighted. A brief overview of the Belle II detector and its components are given. The role of each sub-detector system (in measuring signal coming from the interaction of particles with matter [51]) is clarified.

Chapter 4 covers $B^+ \to K^+\pi^+\pi^-\gamma$ decay from experimental point of view. The description of the analysis procedure is given with accompanying results using generic Monte-Carlo data of 1 ab^{-1}. Being orientated for the photon polarization measurement, the analysis addresses a kinematically correctly reconstructed target set of events defined in this chapter.

Hence, further reasoning is done aiming to obtain such a set of events. As the initial step, the event selection is described in terms of efficiencies of corresponding cuts on physical measurables. Various backgrounds are studied, the main sources are identified and suppressed. Though, a more comprehensive

study can be done to further purify obtained dataset. This part of the work explains the fitting procedure of obtained sample in order to measure yields of the signal and other event species. The procedure of accessing background-subtracted distributions of variables on the B-meson and $K\pi\pi$ system levels is discussed and applied. Measurement of branching ratios of $B^+ \to K^+\pi^+\pi^-\gamma$ decay on MC-samples corresponding to $62.7\ \mathrm{fb}^{-1}$ and $1\ \mathrm{ab}^{-1}$ is done using two different approaches and results are compared to the input PDG value. Discussed an essential step towards photon polarization measurement, where the preferred direction of the photon in the $K\pi\pi$ rest frame is measured. Unblinding on the real data of $62.7\ \mathrm{fb}^{-1}$ ($\Upsilon(4S)$) and $9.2\ \mathrm{fb}^{-1}$ (off-resonance) is performed for several controls samples and results are compared with Monte-Carlo.

1 - Theoretical overview

1.1 . The Standard Model of particle physics

The Standard Model unifies electromagnetic, weak, and strong interactions. Within this theory particles are divided into two families: fermions (building blocks of the matter) and bosons (responsible for the interactions between fermions). Fundamentally, the fermions are six quarks and six leptons grouped into three generations. There are twelve bosons responsible for fundamental interactions: eight gluons standing for strong interaction, a photon exchanged in electromagnetic transitions, and three vector bosons $W^{\pm}$, Z^0 responsible for weak interaction. In addition, the discovery of spinless massive boson at 127 GeV by ATLAS [52] and CMS [53] experimentally confirmed Brout-Englert-Higgs mechanism [8, 9] as a part of SM. Graphically all these particles can be summarized in Fig. 1.1.

Being experimentally validated and tested, the SM represents the scientists' best understanding of the particles and the fields. However, it doesn't describe gravitational forces, addresses the nature of dark matter and dark energy. Moreover, SM can't explain the existing asymmetry between matter and anti-matter observed in the Universe. In the minimalistic conventional SM, the neutrinos are massless while several experiments [54, 55] have demonstrated the neutrinos' oscillation, which appears to be possible if the neutrinos have masses.

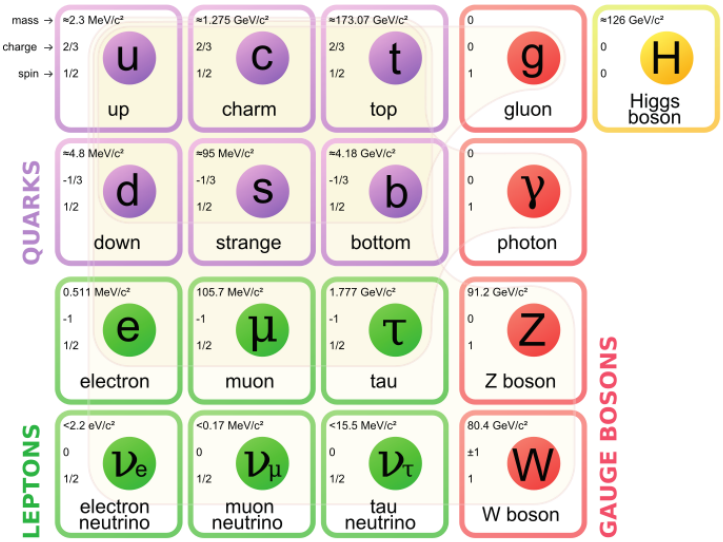


Figure 1.1. Graphical representation of the Standard Model

1.1.1 . Symmetries in the Standard Model

The results of combining quantum mechanics and Einstein's theory of special relativity led Paul Dirac to the Lorentz invariant Lagrangian for fermions [56][1]. In quantum mechanics, the probability of a certain

[1]The theory predicted the existence of positron and described fermions as spinors.

process is described by $|\psi(x)|^2$, where $\psi(x)$ — wave function of this process. Thus, multiplying wave function on a complex number $\psi(x) \to e^{i\phi(x)}\psi(x)$ will not change the corresponding probability and should not change the Lagrangian of the system. Such invariance of physics laws, described by Lagrangian is called gauge invariance. One can notice, that function $e^{i\phi(x)}$ represents the element of the group $\mathrm{U}(1)$ and in order to satisfy the gauge invariance, Dirac theory of free fermion field was extended to include massless boson field.

Moreover, shortly after the experimental discovery of quarks, in 1964 Oscar W. Greenberg introduced the term "color" [57] for quarks in order to explain coexistence inside hadrons in otherwise identical quantum states without violating the Pauli exclusion principle [58]. Since color was not experimentally observed, all stable strongly interacting particles were considered in a color singlet state. Three colors are combined into eight linearly independent combinations, which can be represented by Gell-Mann matrices [2]. These matrices are the generators of SU(3) group and correspond to physical fields of gluons. Thereby, in the following the gauge invariance of the SM is discussed in terms of group theory.

Fundamental interactions within SM are represented by local $SU(3)_C \times SU(2)_L \times U(1)_Y$ gauge symmetry group, where $SU(3)_C$ — the symmetry of strong interactions, $SU(2)_L \times U(1)_Y$ — electroweak interaction group.

Groups in the Standard Model

A group is a collection of elements unified by common properties, where operator $*$ defines an operation on this group. The statement can be formalized by the following rules:

- if a and b are any elements of the group, $a * b$ is also the element of this group;
- $a * (b * c) = (a * b) * c$ — associative property;
- the identity element should exist, satisfying $\hat{I} * a = a * \hat{I} = a$;
- each element of the group has an inverse element: $a * a^{-1} = \hat{I}$.

The dimension of a group equals the number of its independent parameters. The generators of a group are the elements from which all the group elements can be obtained. If the symmetries are continuous, then the group is a Lie group.

Examples of relevant groups for particle physics are:

- the unitary group $U(n)$, containing all $n \times n$ unitary matrices, with the matrix multiplication operation. In particular, for $n = 1$ it is a group of complex numbers with modulus equals 1. The dimension of this group is 1.
- the special unitary group $SU(n)$ contains all $n \times n$ unitary matrices, with determinant equals 1, the group operation is matrix multiplication. Such groups have dimensions $n^2 - 1$ with non-commuting symmetries.

  For instance, if $n = 2$ the generators representing $SU(2)$ can be chosen as on Eq. 1.1.

$$\tau^1 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \quad \tau^2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \tau^3 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \tag{1.1}$$

For $n = 3$, the generators are of the form $T_a = \frac{\lambda_a}{2}$

$$\begin{aligned}
\lambda_1 &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\
\lambda_4 &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} \quad \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \\
\lambda_7 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \quad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}
\end{aligned} \tag{1.2}$$

It can be shown, that the identity generator of $U(1)$ symmetry, three generators of $SU(2)$ and eight generators of $SU(3)$ are the mathematical representation of exchange bosons carrying electromagnetic, weak, and strong interactions respectively.

C, P, T transformations

Three additional symmetries are accounted in the SM preserving the invariance of the physics laws:

- Parity (P) transformation — spacial coordinates are flipped $\overrightarrow{r} \to -\overrightarrow{r}$;
- Charge conjugation (C) — charge of the particle is flipped $e \to -e$;
- Time reversal (T) — time is reversed $t \to -t$.

Electromagnetic and strong interactions conserve C, P and T separately, meanwhile weak interactions violate each of them separately [59, 60]. The combined CP symmetry is violated in the weak interactions as well. The symmetry which is always conserved in the SM is combined CPT [61].

Gauge invariance

Within a principle of gauge invariance the wave function of generic field is defined up to the local tranformations:

$$\psi(x)' \to e^{i\xi^a(x)T^a}\psi(x) \tag{1.3}$$

where $\xi^a(x)$ — set of local transformations, T^a — the generators of Lie group and $\psi(x)$ — wave function that describes a field.

Free Dirac field is described by the following Lagrangian:

$$\mathcal{L}_\psi = \bar{\psi}(i\partial\!\!\!/ - m)\psi \tag{1.4}$$

Substituting Eq. 1.3 into Eq. 1.4 one obtains Lagrangian which is not invariant, i.e $\mathcal{L}_{\psi'} \neq \mathcal{L}_\psi$. It becomes gauge invariant after introducing covariant derivative D_μ that transforms as

$$D_\mu\psi \to (D_\mu\psi)' = e^{ie\xi(x)}(D_\mu\psi) \tag{1.5}$$

where e — charge.

The covariant derivative is represented as

$$\begin{aligned} D_\mu\psi &= (\partial + ieA_\mu(x))\psi \quad (1.6)\\ D_\mu\psi \to (D_\mu\psi)' &= e^{ie\xi(x)}(\partial_\mu - i\partial_\mu\xi(x) + ieA'_\mu(x))\psi \quad (1.7)\end{aligned}$$

where A_μ is transformed as

$$A_\mu \to A'_\mu = A_\mu + \frac{1}{e}\partial_\mu\xi(x) \quad (1.8)$$

Obtained gauge invariant lagrangian is symmetric to local phase transformations (see Eq. 1.9).

$$\begin{aligned} \mathcal{L}_\psi &= \bar{\psi}(i\not{D} - m)\psi \\ &= \bar{\psi}(i\not{\partial} - m)\psi - e\bar{\psi}\gamma^\mu\psi A_\mu \\ &= \bar{\psi}(i\not{\partial} - m)\psi - j^\mu A_\mu \quad (1.9)\end{aligned}$$

The field A_μ can be interpreted as an exchange of boson between fermions. Apart from Lagrangian of free field, an additional term $j^\mu A_\mu$ appeared, which is responsible for interaction between fermion and boson fields.

The compelete Lagrangian including gauge invariant free boson contribution is given as following

$$\mathcal{L}_\psi = \bar{\psi}(i\not{\partial} - m)\psi - j^\mu A_\mu - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \quad (1.10)$$

The mass term of the type $\frac{1}{2}m^2 A_\mu A^\mu$ is dropped since it doesn't satisfy the principle of gauge invariance. In the following, the proper mechanism of giving masses to particles will be described.

Having started from the lagrangian of free Dirac field, and by adding the principle of gauge invariance to it, the Lagrangian describing fermions and bosons was introduced. Considering A_μ as electromagnetic vector potential, the Eq. 1.10 describes electrodynamics with $F_{\mu\nu}$ being electromagnetic tensor, e — electric charge, j_μ — conserved electromagnetic current.

Consider now free Dirac lagrangian, where fields are spinors $\psi = \begin{pmatrix}\psi_1\\ ...\\ \psi_N\end{pmatrix}$ and local gauge transformation is defined by $U(x) = e^{i\overrightarrow{T}\overrightarrow{\xi}(x)}\psi$. Here $\overrightarrow{\xi}(x)$ are arbitrary local functions, $\overrightarrow{T}$ — $N \times N$ dimensional generators of the group satisfying the relation

$$[T^i, T^j] = if_{ijk}T^k \quad (1.11)$$

where f_{ijk} — structure constants of the group.

The new gauge invariant Lagrangian can be expressed by

$$\mathcal{L}_\psi = \psi(i\not{D} - m)\psi + \frac{1}{4g^2}Tr(G_{\mu\nu}G^{\mu\nu}) \quad (1.12)$$

where g — coupling constant, $G_{\mu\nu} = \partial_\nu\Gamma_\mu - \partial_\mu\Gamma_\nu - ig[\Gamma_\nu, \Gamma_\mu]$, $D_\mu(x) = \partial_\mu + i\Gamma_\mu(x)$ and gauge transformation of $\Gamma_\mu(x)$ can be written as following.

$$\Gamma_\mu(x) \to \Gamma_\mu(x)' = U(x)\Gamma_\mu(x)U^\dagger(x) + i(\partial_\mu U(x))U^\dagger(x) \quad (1.13)$$

In the case of $U(1)$, coupling g is a charge, $\Gamma_\mu = eA_\mu$, and $G_{\mu\nu} = F_{\mu\nu}$ giving the case of electrodynamic theory.

In the case of $SU(2)$ and $SU(3)$ the lagrangian from Eq. 1.12 describes a system of fermions interacting via 3 and 8 exchange bosons respectively. From the experimental results follows that these bosons can be massive, and there must be an additional mechanism giving masses to them.

Spontaneous symmetry breaking: mass generation of gauge bosons

Considered previously the principle of gauge invariance is a mechanism of introducing to the theory bosons' and fermions' gauge fields corresponding to the interaction between particles. Applying this principle causes that quants of gauge fields are massless, since $\sim A_\mu A^\mu$ terms violate the gauge invariance.

From experimental data it is known that photons are massless, however, for instance, $W^\pm$ and Z^0 are massive. The theory being gauge invariant and containing massive terms can be built having in the system the boson field, for which spontaneous symmetry breaking occurs. The mechanism of generating masses for gauge fields is called Higgs mechanism [62–64].

Lagrangian of Higgs field is defined as following:

$$\mathcal{L}_H = (D_\mu H)^\dagger D_\mu H - \lambda\Big(H^\dagger H - \frac{\nu^2}{2}\Big)^2 \tag{1.14}$$

where $\dagger$ — unitary conjugation, Higgs field is an isotopic doublet[2] of $SU_W(2)$

$$H = \begin{pmatrix} 0 \\ \frac{\nu+h}{\sqrt{2}} \end{pmatrix} \tag{1.15}$$

and $\lambda > 0$ — parameter defining the shape of potential of H-field, ν — vacuum average of the field.

Kinetic term of the Lagrangian

$$\mathcal{L}^H_{kin} = (D^\mu H)^\dagger D_\mu H \tag{1.16}$$

contains an interaction of gauge bosons with Higgs field.

$$D_\mu H = \Big(\partial_\mu - i\frac{g}{2}\overrightarrow{\tau}\cdot\overrightarrow{V}_\mu - i\frac{g'}{2}B_\mu\Big)H \tag{1.17}$$

where V^i_μ $(i = 1, 2, 3)$ with coupling g being three fields of $SU_W(2)$; B_μ with coupling $\frac{g'}{2}$ — one field of $U_Y(1)$. Taking into account Eqs. 1.1, 1.15, Eq. 1.17 can be rewritten in the form of Eq. 1.18.

$$D_\mu H = \begin{pmatrix} -i\frac{g}{2}(V^1_\mu - iV^2_\mu)\frac{\nu+h}{\sqrt{2}} \\ \frac{\partial_\mu h}{2} + i\frac{g}{2}V^3_\mu\frac{\nu+h}{\sqrt{2}} - i\frac{g'}{2}B_\mu\frac{\nu+h}{\sqrt{2}} \end{pmatrix} \tag{1.18}$$

Substituting Eq. 1.18 into Eq. 1.16 gives

$$(D^\mu H)^\dagger D_\mu H = \frac{\partial^\mu h\partial_\mu h}{2} + \frac{g^2}{8}(\overrightarrow{V}^\mu\cdot\overrightarrow{V}_\mu)(\nu+h)^2 + \frac{g'^2}{8}B^\mu B_\mu(\nu+h)^2 - \frac{gg'}{4}B^\mu V^3_\mu(\nu+h)^2 \tag{1.19}$$

[2]The Higgs field given here is in a unitary calibration, which otherwise introduces much more additional computations, but doesn't change final result.

Massive terms of gauge fields can be combined into the following lagrangian

$$\mathcal{L}_m^{gauge} = \frac{g^2\nu^2}{8}(V^{1,\mu}V_\mu^1 + V^{2,\mu}V_\mu^2) + \left(B^\mu\, V^{3,\mu}\right)\begin{pmatrix} \frac{g^2\nu^2}{8} & \frac{-gg'\nu^2}{8} \\ \frac{-gg'\nu^2}{8} & \frac{g'^2\nu^2}{8} \end{pmatrix}\begin{pmatrix} B^\mu \\ V^{3,\mu} \end{pmatrix} \tag{1.20}$$

Mass part of gauge fields contains non-diagonal terms. In order to get fields in massive basis, the following transformation is performed

$$\begin{pmatrix} B_\mu \\ V_{3,\mu} \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} \tag{1.21}$$

Re-expressing V_μ^1, V_μ^2 gives

$$V_\mu^1 = \frac{W_\mu^- + W_\mu^+}{\sqrt{2}} \qquad V_\mu^2 = \frac{W_\mu^- - W_\mu^+}{i\sqrt{2}} \tag{1.22}$$

Substituting Eqs. 1.21-1.22 into Eq. 1.20 non-diagonal terms can be removed by choosing angle θ to be

$$\tan(2\theta) = \frac{2gg'}{g^2 + g'^2} \tag{1.23}$$

It gives the following expression for Eq. 1.20

$$\mathcal{L}_m^{gauge} = m_W^2 W^{-,\mu}W_\mu^+ + \frac{m_A^2}{2}A^\mu A_\mu + \frac{m_Z^2}{2}Z^\mu Z_\mu \tag{1.24}$$

where $m_W = \frac{g\nu}{2}$ $\quad m_Z = \frac{\sqrt{g^2+g'^2}\nu}{2}$ $\quad m_A = 0.$

As a result of spontaneous symmetry breaking Higgs's field obtains a non-zero vacuum average, which results in the masses generation for fields $\overrightarrow{V}_\mu$, B_μ. Computations are shown that V_μ^3 and B_μ aren't fields on a mass basis. After transformation to mass basis massless electromagnetic and massive $W^\pm$, Z bosons are obtained.

The gauge field of a group $SU_C(3)$ does not contribute to the Eq. 1.17, since the Higgs field is a singlet with respect to the $SU_C(3)$. Correspondingly, the gauge field of strong interactions is massless.

1.1.2 . Weak interactions

One of the base of the Standard Model is an experimental fact of the leptons' (e μ τ), non-charged neutrinos (ν_e ν_μ ν_τ) and quarks' (u d c s t b) existense. Leptons with left chirality $L_n = \begin{pmatrix} \nu_n \\ e_n \end{pmatrix}$ create isotopic dublets of $SU_W(2)$, charged right $E_n = (e_n)_R$ leptons create singlets of $SU_W(2)$, neutrinos field with right chirality don't exist. Left quarks $Q_n = \begin{pmatrix} u_n \\ d_n \end{pmatrix}_L$ create isotopic dublets of $SU_W(2)$, right quarks $U_n = (u_n)_R$, $D_n = (d_n)_R$ — singlets of $SU_W(2)$.

Left and right wave functions are defined using projection operators P_L P_R.

$$P_{L,R} = \frac{1}{2}(1 \pm \gamma^5) \tag{1.25}$$

Gauge invariant lagrangian of fermionic fields interacting with Higgs field in unitary gauge can be written as follows.

$$\mathcal{L}_m = - \frac{\nu}{\sqrt{2}}(Y^l_{mn}\bar{e}_{L_m}e_{R_n} + Y^d_{mn}\bar{d}_{L_m}d_{R_n} + Y^u_{mn}\bar{u}_{L_m}u_{R_n} + h.c) \tag{1.26}$$

$$- \frac{h}{\sqrt{2}}(Y^l_{mn}\bar{e}_{L_m}e_{R_n} + Y^d_{mn}\bar{d}_{L_m}d_{R_n} + Y^u_{mn}\bar{u}_{L_m}u_{R_n} + h.c) \tag{1.27}$$

One notes that neutrino field is absent in the above lagrangian, hence in SM neutrino is massless and doesn't interact with Higgs field h. In order to remove mixing in massive terms of fermions, one should change the basis using unitarity transformation ($U^\dagger U = \hat{1}$)

$$e_{R_m} = U^{e_R}_{mn}\tilde{e}_{R_n} \qquad d_{R_m} = U^{d_R}_{mn}\tilde{d}_{R_n} \qquad u_{R_m} = U^{u_R}_{mn}\tilde{u}_{R_n} \tag{1.28}$$

$$e_{L_m} = U^{e_L}_{mn}\tilde{e}_{L_n} \qquad d_{L_m} = U^{d_L}_{mn}\tilde{d}_{L_n} \qquad u_{L_m} = U^{u_L}_{mn}\tilde{u}_{L_n} \tag{1.29}$$

In this new basis mixing matrices U can be chosen to diagonalize $\tilde{Y}^l$, $\tilde{Y}^d$, $\tilde{Y}^u$ so that they are real

$$\tilde{Y}^l = (U^{e_L})^\dagger Y^l U^{e_R}, \qquad \tilde{Y}^d = (U^{d_L})^\dagger Y^d U^{d_R}, \qquad \tilde{Y}^u = (U^{u_L})^\dagger Y^u U^{u_R} \tag{1.30}$$

Then masses of fermions are determined by diagonal matrices

$$m_{e_n} = \frac{\nu}{\sqrt{2}}\tilde{Y}^l_{nn}; \qquad m_{d_n} = \frac{\nu}{\sqrt{2}}\tilde{Y}^d_{nn}; \qquad m_{u_n} = \frac{\nu}{\sqrt{2}}\tilde{Y}^u_{nn} \tag{1.31}$$

In the quark sector transition to a massive basis causes quarks mixing of different generations in the vertices, where interaction with $W^\pm$ takes place.

$$\mathcal{L}_{q-W} = -\frac{g}{2\sqrt{2}}\left[\bar{\tilde{u}}_m\gamma^\mu(1-\gamma^5)V_{mn}\tilde{d}_n W^+_\mu + \bar{\tilde{d}}_m\gamma^\mu(1-\gamma^5)V^+_{mn}\tilde{u}_n W^-_\mu\right] \tag{1.32}$$

where $V = (U^{u_L})^{-1}U^{d_L}$ — unitary 3×3 matrix of mixing Cabbibo-Kobayashi-Masakawa (CKM).

There are several ways to parametrize CKM matrix. Standard parametrization is set up using several sequential Euler's rotations $V = R_{23}W_{13}R_{12}$, where R_{ij} — matrix of real rotations in the ij plane ($i,j = 1,2,3$ corresponds to x,y,z), W_{13} — complex rotations in the xz plane:

$$\begin{aligned} V &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix}\begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix}\begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} c_{12}c_{13} & c_{13}s_{12} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} \end{aligned} \tag{1.33}$$

where δ — phase of Kobayashi-Masakawa; $c_{ij} = \cos\theta_{ij}$, $s_{ij} = \sin\theta_{ij}$, $\theta_{12}, \theta_{13}, \theta_{23}$ — three mixing angles.

In the Wolfenstein parametrization, CKM matrix looks as follows

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + O(\lambda^4) \tag{1.34}$$

where $\lambda = s_{12}$, $A = \frac{s_{23}}{\lambda^2}$ and ρ, η may be extracted through the measurements of V_{ub} and V_{td}. This parametrization is useful to understand the order of suppression of flavour changing process.

Returning to the interaction of leptons with W-bosons, the lagrangian is the following

$$\mathcal{L}_{l-W} = -\frac{g}{2\sqrt{2}}\left[\bar{\nu}_n U^{e_L}_{mn}\gamma^\mu(1-\gamma^5)\tilde{e}_m W^+_\mu + \bar{\tilde{e}}_m\gamma^\mu(1-\gamma^5)(U^{e_L})^+_{mn}\nu_n W^-_\mu\right] \tag{1.35}$$

Reparametrizing $\tilde{\nu} = (U^{e_L})^+\nu$, Eq. 1.35 becomes

$$\mathcal{L}_{l-W} = -\frac{g}{2\sqrt{2}}\left[\bar{\tilde{\nu}}_n\gamma^\mu(1-\gamma^5)\tilde{e}_n W^+_\mu + \bar{\tilde{e}}_n\gamma^\mu(1-\gamma^5)\tilde{\nu}_n W^-_\mu\right] \tag{1.36}$$

The functions $\tilde{\nu}_n$ are the flavour states (e, ν, τ neutrinos). In the description of interaction of W-bosons with leptons mixing matrix analogous to CKM is avoided, only because in SM there are no mass terms for neutrinos. If they existed reparametrization $\tilde{\nu} = (U^{e_L})^+\nu$ wouldn't be possible.

1.1.3 . Strong interactions

Lagrangian of quantum chromodynamics (QCD) [2, 65] is as following

$$\mathcal{L}_{QCD} = -\frac{1}{4}\sum_{a=1}^{8} G^a_{\mu\nu}G^{a,\mu\nu} + \sum_{quarks}\bar{q}\Big(i\gamma^\mu\partial_\mu - m_q + g_s\sum_{a=1}^{8}\frac{\lambda^a}{2}G^a_\mu\Big)q \tag{1.37}$$

which includes 6 quark fields (u, d, c, s, t, b) and 8 masseless gluons (quants of gauge fields G^a_μ). In Eq. 1.37

$$\sum_{a=1}^{8}\frac{\lambda^a}{2}G^a_\mu = \frac{1}{2}\begin{pmatrix} G^3_\mu + \frac{1}{\sqrt{3}}G^8_\mu & G^1_\mu - iG^2_\mu & G^4_\mu - iG^5_\mu \\ G^1_\mu + iG^2_\mu & -G^3_\mu + \frac{1}{\sqrt{3}}G^8_\mu & G^6_\mu - iG^7_\mu \\ G^4_\mu + iG^5_\mu & G^6_\mu + iG^7_\mu & -\frac{2}{\sqrt{3}}G^8_\mu \end{pmatrix} = \hat{\Gamma} \tag{1.38}$$

is a matrix of 3×3 with λ^a being $SU(3)$ generators defined in Eq. 1.2, and q — quark function (row of three color components of a quark). As can be seen interaction of quarks of different colors occurs through combinations of gluons' fields. Interaction of quarks in the QCD can be written as

$$\mathcal{L}^{QCD}_{inter} = g_s\sum_{a,b=r,g,b}\bar{q}_a\Gamma_{ab}q_b \tag{1.39}$$

where the elements $g_s\Gamma_{ab}$ define intensity of interactions of quarks of different colors.

Quarks and gluons can be observed as jets of particles, however, their detection as isolated objects is forbidden. For the study of QCD processes the experiments need to have access to very large energies, corresponding to very small distances between quarks in a hadron, so that the quarks can be studied as quasi-free objects.

1.1.4 . Parametrising new Physics within the Standard Model

The Standard Model has been considered as the most fundamental theory in the sense that it is valid up to higher energies, than other theories [66, 67]. Yet, it is not a theory of everything and the existence of new particles and interactions beyond its formalism is a motivated hypothesis.

Also the CKM matrix has a hierarchical structure, where diagonal terms are close to unity and mixing angles are $\theta_{13} \ll \theta_{23} \ll \theta_{12}$. The experimental observations of neutrino oscillations signify that there is also a rich flavour structure in the leptonic sector. All of these masses and mixings are free parameters in SM, which should be explained by a higher scale theory.

Additional CP-violation is needed to produce the observed baryon asymmetry in the universe. Within the SM, from Eq. 1.34 CKM matrix elements V_{ub} and V_{td} contain complex phases and provide the only source of CP-violation. However, models of baryogenesis suggest that it is quantitatively insufficient, therefore another reason to consider the new physics models arises.

Flavour changing neutral current processes, such as $B-\bar{B}$ missing and $b \to s\gamma$ transition, provide strong constraints on new physics beyond the SM. If there is no suppression mechanism for flavour changing neutral current (FCNC) processes, the new physics contribution can become consistent with the experimental data. Hence, the measurement of FCNC processes provides a test of the flavour structure in the NP models.

Inspired by lectures of A. Falkowski [68], this part explains parametrizing of New Physics using SM building blocks. At the time of writing the thesis, there is no confident direct or indirect evidence of new particles beyond SM. Hence, one can assume, that searched new particles are much heavier than the weak scale. Integrating out the beyond-SM particles from the theory, their effects can be described using effective field theory. Such theory is analogous to SM: it has the same degrees of freedom, local symmetry group, mechanism of mass generation and is constructed from fermion, gauge, and Higgs fields. This difference with the SM, that it has arbitrary large mass dimensions D. The interactions can be described by the Lagrangian as in Eq. 1.40.

$$\mathcal{L}_{SMEFT} = \mathcal{L}_{SM} + \sum_{n=1}^{\infty} \frac{1}{\Lambda^n} \sum C_i^{(n+4)} \mathcal{O}^{D=n+4} \tag{1.40}$$

where Λ — mass scale of a new particle, which has been integrated out, C_i — Wilson coefficients, reflecting interaction couplings, $\mathcal{O}_i$ — operators constructed from SM fields.

All odd dimension operators $\mathcal{O}_i$ violate the lepton number. Leading contribution comes from $D = 5$ and gives rise to Majorana-type neutrino [69] after spontaneous symmetry breaking. On the other side, not all even-dimensional operators violate the B-L number. Restricting the search for New Physics to lepton conservating processes, EFT lagrangian will start from $D = 6$ operators, since all particles violating this number have been integrated out.

The SMEFT approach though does not allow to estimate the Λ scale by its own, but only the ratio $\frac{C_i^{(n+4)}}{\Lambda^n}$ as in the same way effective Fermi theory of weak interactions did not allow to infer the mass of the W boson. The searches for carries of weak interaction were performed starting from few GeV. But once $SU(2)_L \times U(1)_Y$ was proposed, weak coupling g was derived from electromagnetic coupling e, and subsequently, masses of the W $(80\ \mathrm{GeV/c^2})$ and Z $(90\ \mathrm{GeV/c^2})$ bosons were obtained.

Having covered this historical background of SM advancement, it is essential to assume that the current theory is an edge case (up to TeV scale) of a higher dimensional model of everything, and searches for the particles and processes predicted by such theories are actively being conducted [70–72].

1.2 . The photon polarization in the radiative B decays

Consider the process of $b \to s\gamma$ in the Standard Model. There is no term in the Lagrangian directly associated with the spinor functions of two quarks of the same charge and different flavours. Thus, such an FCNC process is forbidden on the tree level. However, there is a term (see Eq. 1.32) representing interaction with W-boson, which violates the flavour of quark and changes its charge. Based on this term, in the SM $b \to s\gamma$ proceeds mostly through a loop of W-q pair [73] as given on the Fig. 1.2, where the photon can be emitted by internal quark. The result of computation $b \to s\gamma$ vertex is given as in Eq. 1.41 [49, 73].

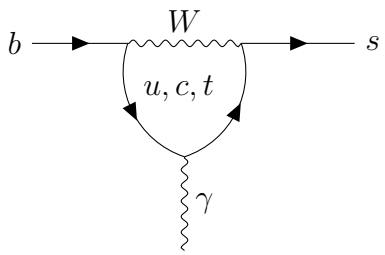


Figure 1.2. In the SM internal quark line is mostly t-quark, since $V_{ts}^* V_{tb}$ and $V_{cs}^* V_{cb}$ nearly of the same order, suppressing $V_{us}^* V_{ub}$. Moreover loop function F_2 is small for c-quark.

$$\bar{s}\Gamma_\mu^{b\to s\gamma} b = \frac{e}{(4\pi)^2}\frac{g^2}{2M_W^2} V_{ts}^* V_{tb} F_2 i\bar{s}\sigma_{\mu\nu}q^\nu(m_b P_R + m_s P_L)b \tag{1.41}$$

where $q^\nu = p_b^\nu - p_s^\nu$, F_2 — loop function whose expression can be found in [74], $V_{ts}.V_{tb}$ — CKM matrix elements, defined in Eq. 1.34, $\sigma_{\mu\nu} = \frac{i}{2}(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu)$, m_b, m_s — masses of b and s quarks respectively.

Helicity amplitude can be obtained by multiplying Eq. 1.41 on photon polarization vector $\varepsilon_{R,L}^\mu$. Considering only operator part of the above expression, one obtains:

$$\mathcal{M} \sim \bar{s}\sigma_{\mu\nu}q^\nu(m_b P_R + m_s P_L)b\varepsilon_{R,L}^{\mu *} \tag{1.42}$$

Explicit calculations of helicity amplitude can be made in the b-quark rest frame, where the photon 4-momentum is defined as $q^\mu = (|q|, 0, 0, |q|)$ and polarization vector lies in the transverse plane to photon direction $\varepsilon_{R,L}^\mu = \mp\frac{1}{\sqrt{2}}(0, 1, \pm i, 0)$. From Eq. 1.42 the following relations can be obtained (see Eqs. 1.43-1.44).

$$\bar{s}_L\sigma_{\mu\nu}q^\nu b_R\varepsilon_R^{\mu *} = 0, \quad \bar{s}_L\sigma_{\mu\nu}q^\nu b_R\varepsilon_L^{\mu *} \neq 0 \tag{1.43}$$

$$\bar{s}_R\sigma_{\mu\nu}q^\nu b_L\varepsilon_L^{\mu *} = 0, \quad \bar{s}_R\sigma_{\mu\nu}q^\nu b_L\varepsilon_R^{\mu *} \neq 0 \tag{1.44}$$

One may conclude that the polarization of s-quark equals polarization of the photon for $b \to s\gamma$ process. Using Eqs. 1.42- 1.44 one obtains Eqs. 1.45-1.46.

$$\mathcal{M}(b_R \to s_L\gamma_L) \sim m_b \tag{1.45}$$

$$\mathcal{M}(b_L \to s_R\gamma_R) \sim m_s \tag{1.46}$$

Another conclusion is that, within the SM photon is predominantly left-handed and the right-handed photon is suppressed by a factor of $\frac{m_s}{m_b} \approx 0.02$.

Including QCD corrections allows to obtain the effective Hamiltonian on the weak scale (see Eq. 1.47) [75].

$$\begin{aligned}\mathcal{H}_{eff} \quad \sim \quad & \Big[\sum_{i=1}^{6}(C_i(\mu)\mathcal{O}_i(\mu) + C'_i(\mu)\mathcal{O}'_i(\mu)) + \underbrace{C_{7\gamma}(\mu)\mathcal{O}_{7\gamma}(\mu) + C'_{7\gamma}(\mu)\mathcal{O}'_{7\gamma}(\mu)}_{b \to s\gamma} \\ + \quad & C_{8g}(\mu)\mathcal{O}_{8g}(\mu) + C'_{8g}(\mu)\mathcal{O}'_{8g}(\mu)\Big] \end{aligned} \tag{1.47}$$

where μ — renormalization scale, C_i are the short-distance Wilson coefficients, $\mathcal{O}_i$ are the local long-distance operators. The electromagnetic operators $\mathcal{O}_7$ and $\mathcal{O}'_7$ describe left- and right handed photon respectively, and defined in Eqs. 1.48-1.49.

$$\mathcal{O}_7 = \frac{e^2}{16\pi^2} m_b \bar{s}_{\alpha L} \sigma_{\mu\nu} b_{\alpha R} F^{\mu\nu} \tag{1.48}$$

$$\mathcal{O}'_7 = \frac{e^2}{16\pi^2} m_b \bar{s}_{\alpha R} \sigma_{\mu\nu} b_{\alpha L} F^{\mu\nu} \tag{1.49}$$

where α — color index, $F^{\mu\nu}$ — electromagnetic field tensor.

Matching Eqs. 1.41, 1.47 and applying Fourier transform $-\sigma_{\mu\nu}F^{\mu\nu} \to 2i\sigma_{\mu\nu}q^\nu\varepsilon^{\mu *}$ gives relation as in Eq. 1.50.

$$\frac{C'_7}{C_7} = \frac{m_s}{m_b} \tag{1.50}$$

Having obtained C_7 and C'_7, the photon polarisation parameter can be defined as in Eq. 1.51.

$$\lambda_\gamma = \frac{1 - \left|\frac{C'_7}{C_7}\right|^2}{1 + \left|\frac{C'_7}{C_7}\right|^2} \tag{1.51}$$

In the experimental results one usually refers to C_7^{eff} and $C_7^{\prime eff}$. Wilson coefficients C_7 and C'_7 absorb short-distance QCD corrections and in the SM one obtains Eq. 1.52.

$$\frac{C_7^{\prime eff}}{C_7^{eff}} \approx \frac{m_b}{m_s} \approx 0.02 \tag{1.52}$$

Though, explicit calculations of QCD corrections effect performed by [49, 73] show a substantial difference between C_7 and C_7^{eff} concluding that the photon polarization is a nontrivial experimental observable sensitive to the structure of effective Hamiltonian.

1.2.1 . Motivation for measuring the photon polarization

Flavour changing neutral current transitions as $b \to s\gamma$ are expected to be very sensitive to NP effects. These processes are allowed at loop level in the SM suppressed due to the GIM mechanism [87–89]. In this case, New Physics may be observed in the exchange of heavy particles in the electroweak penguin loop.

Several theories beyond the SM [87–89] suggest that there is still the possibility of non-SM photon polarization. In some scenarios, the photon may acquire a significant right-handed component, because of a heavy-fermion being exchanged in the electroweak penguin loop. In addition, in grand unification models, right-handed neutrinos (and the consequent right-handed quark coupling) are expected to enhance the right-handed photon component [90].

The contribution of a right-handed photon may be enhanced in the Left-Right Symmetric Model (LRSM) [76], caused by $W_L - W_R$ mixing and chirality flip along the internal t-quark line in the loop. Another example is the unconstrained minimal supersymmetric model [77], where a strong enhancement of order $\frac{m_{\tilde{g}}}{m_b}$ appears due to chirality flip along the gluino line and left-right squark mixing. In this case, the photon polarization parameter λ_γ may have any value between [-1; 1].

Various theories beyond SM predict possible deviation from the left-handed polarization component. It confirms the fact of the high sensitivity of λ_γ and $b \to s\gamma$ to the New Physics.

Experimental status of $b \to s\gamma$ decays

The current status of the $b \to s\gamma$ decays can be summarized by the Fig. 1.3. The approaches applied within each of these works significantly differ but allow to constrain electroweak Wilson coefficients.

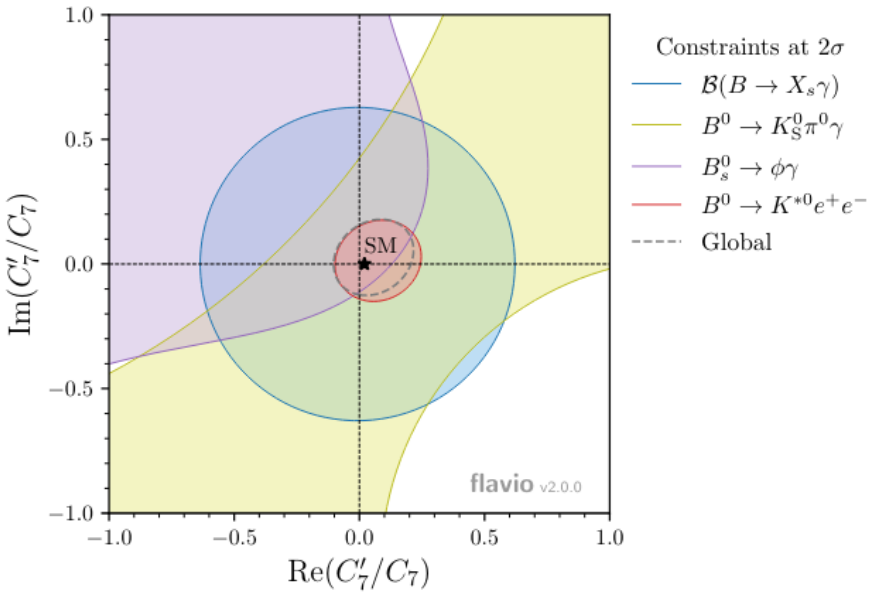


Figure 1.3. Summarized constraints on Wilson coefficients C_7 and C'_7. Despite of past measurements, performed by Belle [78, 79], BaBar [80–83] and LHCb [84] recent results of LHCb [48] demonstrate strong constraints on the photon polarization. The figure has been taken from the latter work.

Branching ratio measurements of inclusive $B \to X_s\gamma$ decay performed by Belle [78] and BaBar [80–82]

allowed to experimentally restrict C_7 and C_7' (see Eq. 1.53).

$$\mathrm{BR(B \to X_s\gamma)} \propto |\mathrm{C_7}|^2|\mathcal{M}_\mathrm{L}|^2 + |\mathrm{C_7'}|^2|\mathcal{M}_\mathrm{R}|^2 \tag{1.53}$$

where left- and right-handed components sum up because the photon is in different helicity states.

Another approach was implemented by Belle and BaBar for $B^0 \to K^0_S\pi^0\gamma$ [79, 83] and LHCb for $B^0_s \to \phi\gamma$ [84]. Constraints on C_7 and C_7' are infered from time dependent CP-asymmerty measurements.

The generic expression of decay rate can be written as in Eq. 1.54.

$$\Gamma(B(\bar{B}) \to f_{CP}\gamma) \propto e^{-\Gamma t}\Big[\cosh\big(\frac{\Delta\Gamma}{2}t\big) - H\sinh\big(\frac{\Delta\Gamma}{2}t\big) \pm C\cos(\Delta m t) \mp S\sin(\Delta m t)\Big] \tag{1.54}$$

where $\Delta\Gamma$ and Δm are the decay width and mass differences between the B CP eigenstates; H, C, S — the parameters related to the polarisation of the photon, direct CP asymmetry and the CP asymmetry associated with $B_{(s)} - \bar{B}_{(s)}$ mixing.

Consider B meson decay to a hadronic CP eigenstate f_{CP} as $B(t) \to f_{CP}\gamma$ with $\xi = \pm 1$. If B state identified (tagged) as a B_0 (and not a $\bar{B}^0$) at time t=0, one obtains:

$$\mathcal{M}(\bar{B} \to f_{CP}\gamma_L) = A\cos\psi e^{i\phi_L}, \quad \mathcal{M}(\bar{B} \to f_{CP}\gamma_R) = A\cos\psi e^{i\phi_R} \tag{1.55}$$

$$\mathcal{M}(B \to f_{CP}\gamma_R) = \xi A\cos\psi e^{-i\phi_R}, \quad \mathcal{M}(B \to f_{CP}\gamma_L) = \xi A\cos\psi e^{-i\phi_L} \tag{1.56}$$

where ϕ_L, ϕ_R — CP-odd weak phases, ξ — CP-eigenvalues, A — amplitude controlling the overall rate , ψ — the relative amount of left-polarized photons and right-polarized photons in $\bar{B}$ decays.

Then at time t, the time-dependent CP asymmetry is given by Eq. 1.57, where effects of small width difference between B-meson states and direct CP-violation are neglected in Eq. 1.54.

$$\begin{aligned}\mathcal{A}_{CP}(t) &= \frac{\Gamma(B^0(t) \to f_{CP}\gamma) - \Gamma(\bar{B}^0(t) \to f_{CP}\gamma)}{\Gamma(B^0(t) \to f_{CP}\gamma) + \Gamma(\bar{B}^0(t) \to f_{CP}\gamma)} \\ &= \xi\sin(2\psi)\sin(\phi_M - \phi_L - \phi_R)\sin(\Delta m t) \\ &= S_{CP}\sin(\Delta m t)\end{aligned} \tag{1.57}$$

where $\Gamma(B(t) \to f_{CP}\gamma)$ — decay rate, ϕ_M — phase of $B^0 - \bar{B}^0$ mixing, Δm — mass difference between neutral B-meson states.

In particular, one expects the following $\mathcal{A}(t)$ for the above mentioned decays (see Eqs. 1.58-1.59).

$$B^0 \to K^0_S\pi^0\gamma \quad : \qquad \mathcal{A}(t) = \frac{2m_s}{m_b}\sin(2\beta)\sin(\Delta m t) \tag{1.58}$$

$$B^0_s \to \phi\gamma \quad : \qquad \mathcal{A}(t) \approx 0 \tag{1.59}$$

where β — angle taken from the relations for CKM matrix.

Obtained results are summarized in Table. 1.1.

Experiment	Decay channel	$B-\bar{B}$ mixing	Direct CP	Asymmetry due to $\Delta\Gamma$
Belle	$B^0 \to K^0_S\pi^0\gamma$	$-0.10\pm0.31\pm0.07$	$-0.20\pm0.20\pm0.06$	
BaBar	$B^0 \to K^0_S\pi^0\gamma$	$-0.78\pm0.59\pm0.09$	$-0.36\pm0.33\pm0.04$	
LHCb	$B^0_s \to \phi\gamma$	$0.43 \pm 0.30 \pm 0.11$	$0.11 \pm 0.29 \pm 0.11$	$-0.67^{+0.37}_{-0.41} \pm 0.17$

Table 1.1. Time-dependent CP-asymmetry measurements. Direct CP and $B-\bar{B}$ mixing have been measured by Belle, BaBar and LHCb. On the other side, only LHCb has measured asymmetry related to the photon polarization.

The contributions from S and C vanish when considering the inclusive $\Gamma_{B_s}+\Gamma_{\bar{B}_s}$ assuming the $\bar{B}_s/B_s$ production asymmetry vanish. Therefore no flavor tagging is required for the extraction of H [44].

Another method to constraint the photon polarization is to perform an angular analysis with respect to the photon direction [39]. For instance, $B^0 \to K^{*0}(\to K^+\pi^-)\gamma$ can not be used in the scope of this approach, since in the B-meson rest frame $K-\pi$ plane is symmetric with respect to the photon direction. On the other hand, $B \to K_{res}(\to K\pi\pi)\gamma$ decay has 4 particles in the final state and photon helicity can be measured. Another promising decay process, which has been successfully applied for angular analysis is $B^0 \to K^{*0}\gamma(\to e^+e^-)$, where the photon is virtual.

Recently LHCb [48] has measured photon polarization with 5 % sensitivity using angular analysis of $B^0 \to K^{*0}e^+e^-$, where the dilepton pair e^+e^- originates from a virtual photon. In general, the analysis with muons in the final state produces a higher yield at LHCb, due to the distinctive signal that muons provide in the selection, better mass, and energy resolutions, and higher reconstruction efficiency of dimuon decays [47]. However, dielectron decays at low squared invariant masses of virtual photons provide higher sensitivity for the photon polarization measurements: due to muon mass, virtual photon contribution is suppressed compared to the one in dielectron decays. In addition, neglecting electron mass allows to significantly simplify the formalism.

Obtained results are summarized on the Fig. 1.3 in the form of red "disk". One can observe, that the photon polarization in $b \to s\gamma$ is constrained through C_7 and C_7' coefficients with significantly better precision than the combination of previous measurements.

Several works [45, 46, 93] demonstrate the promising results of studying photon polarization through $B \to K\pi\pi\gamma$ decay. In the first part of the work [93] the probability density function (p.d.f) of $K\pi\pi$ invariant mass is modeled as a squared linear combination of relativistic Breit-Wigner functions with complex couplings. It allows introducing interference between various kaonic resonances. In the second part of the work time-dependent analysis of $B^0 \to K^0_S\pi^+\pi^-\gamma$ is performed with the purpose of probing the photon helicity.

Similar to the first part of the previous case, in [45] amplitude fit is performed for $B^+ \to K^+\pi^+\pi^-\gamma$ modeling, including Dalitz dimensions (squared invariant masses of $K\pi$ and $\pi\pi$). It allowed enhancing the power of the fit. Photon polarization was accessed through up-down A_{ud} asymmetry measurements. Such quantity is proportional to λ_γ, and is interpreted as the asymmetry between the measured signal rates with photons emitted above and below the $K\pi\pi$ decay plane in the K_{res} reference frame. Obtained results with 5σ level demonstrate that A_{ud} is not zero, thus the photon is polarized.

In the work [46] the analysis has been extended to include 2 additional angular variables, describing the photon orientation with respect to $K\pi\pi$ plane and rotations within this plane. Within the work, the photon polarization parameter was directly introduced into the formalism. Obtained statistical uncertainty $\sigma_{\lambda_\gamma} = 0.018$ for 14 000 signal events demonstrates that the method itself offers a sensitivity to this parameter on the generator level (without taking into account detector effect and background).

2 - $B \to K\pi\pi\gamma$ decay modelling

2.1 . Theory

The main purpose of particle decays modeling is to verify an understanding of the Physics of a system under study. Model is a function of several kinematical variables and parameters. The latter is adjusted to reproduce as close as possible a description of the real data collected by a detector.

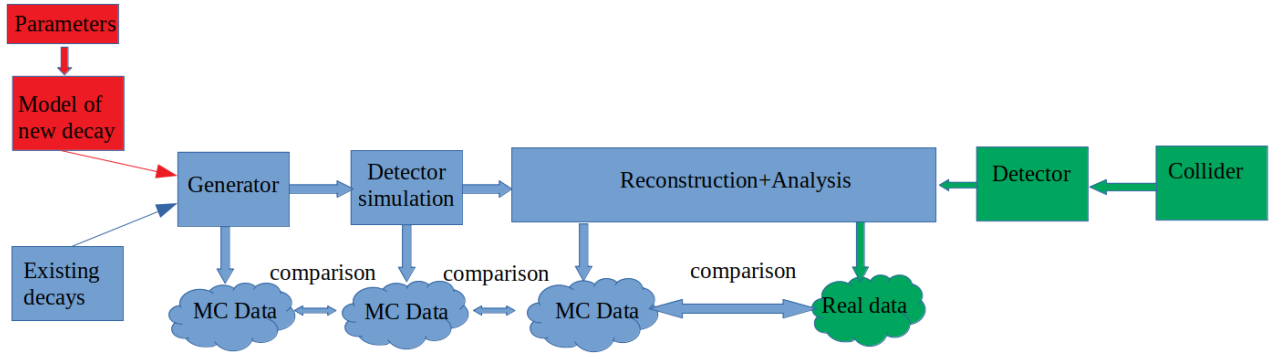


Figure 2.1. Pipeline of data processing. Model is an essential source of artificial Monte-Carlo data produced for comparison with real data.

The typical scenario of adding new decay is given in Fig. 2.1. The sampled dataset represents a snapshot of the incorporated physics properties of the predefined model and is obtained as an output of a generator engine. Further, this data is passed through the pipeline including simulation of detector effect, reconstruction software, and particle physics analysis. After each stage, the generated data is being modified by each component. The final stage is the comparison of the Monte-Carlo and real dataset recorded by the detector. Obtained differences may signify a non-realistic new model or chosen parameters, assuming perfect alignment of simulation with detector setup and ability of existing decays to properly model real data (for example, all backgrounds are properly understood).

Sanity check of the modelling pipeline is made to verify the ability of parameters extraction using generated data (see Fig. 2.2).

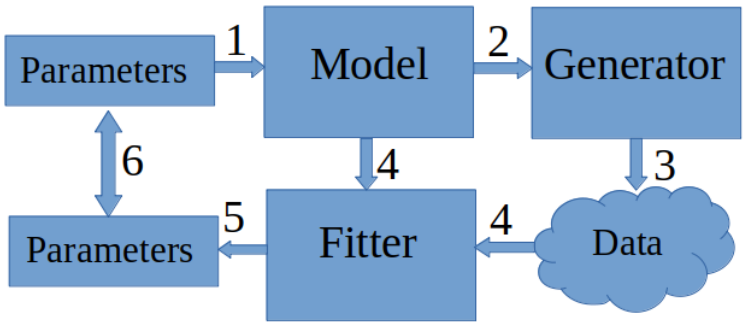


Figure 2.2. Self-consistency of the modelling pipeline: 1 — parameters are inputed to the model function; 2 — model is an input to the generator engine; 3 — generator produces artificial MC data; 4 — obtained data and model function are used as the inputs to the fitter; 5 — extraction model parameters from the fit; 6 — comparison of input and output parameters.

In the present work $B \to K\pi\pi\gamma$ modeling is performed for two cases: in the frame of a standalone package called "GamPola" [112], incorporating generator and fitter, and as a part of the external generator engine called "EvtGen" [95]. The integration within the latter is crucial, since it is a commonly used generator engine in the B-Physics. Moreover, it serves as a link for further data processing according to Fig. 2.1.

2.1.1 . Motivation for $B \to K\pi\pi\gamma$ decay modelling

The decay $B \to K_{res}\gamma$ proceeds through several kaonic resonances K_{res} followed by the decay of K_{res} to the $K\pi\pi$ final state. The existing approach used within the "EvtGen" [95] framework doesn't allow to describe the $K\pi\pi$ system taking into account the interference of kaonic resonances. It is rather assumed that the probability of the decay is proportional to a sum of squared amplitudes $\mathcal{A}^i_{K_{res}}$ corresponding to the intermediate resonances K_{res} as in Eq. 2.1.

$$P_{EvtGen} \sim \sum_i |\mathcal{A}^i_{K_{res}}|^2 \quad \text{No Interference} \tag{2.1}$$

$$P_{Int} \sim |\sum_i \mathcal{A}^i_{K_{res}}|^2 \quad \text{Interference} \tag{2.2}$$

The differential decay rate must be described as a sum of the contributions of kaonic resonances at amplitude level as in Eq. 2.2. Hence, an amplitude-based approach is needed to properly describe this system. This section contains an overview of the theoretical formulas used for describing the differential decay rate of $B \to K\pi\pi\gamma$ channel.

2.1.2 . Kinematics of $B \to K\pi\pi\gamma$ decay

In the given frame, the 4-momenta of the final state particles are constrained under kinematical bounds. In this subsection, more details are given about the description of these constraints.

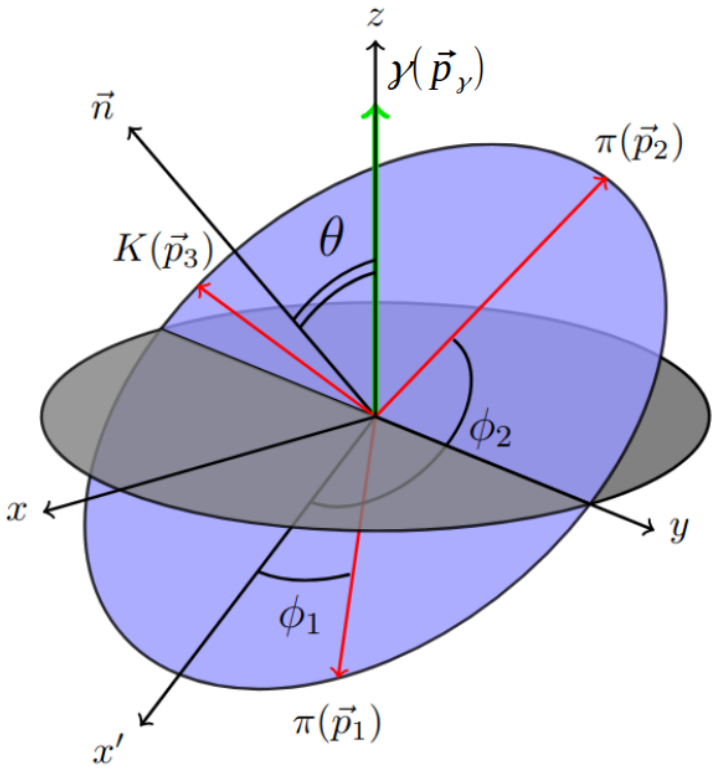


Figure 2.3. Kinematics of the $B \to K\pi\pi\gamma$ decay: three directions of K, π, π (red arrows) represent $K\pi\pi$ plane and the photon direction (green arrow along z-axis) has an angle θ with a normal to the plane.

The kinematics of $B \to K\pi\pi\gamma$ decay can be described with the following variables defined in the $K\pi\pi$ rest frame (see Fig. 2.3):

- $s_{K\pi\pi} = M^2_{K\pi\pi}$ — squared invariant mass of $K\pi\pi$;
- $s_{K\pi} = M^2_{K\pi}$ — squared invariant mass of $K\pi$;
- $s_{\pi\pi} = M^2_{\pi\pi}$ — squared invariant mass of $\pi\pi$;
- θ — angle between normal vector to the $K\pi\pi$ plane and the photon direction;
- $\phi = \frac{\phi_1+\phi_2}{2}$ — rotation angle in the $K\pi\pi$ plane between x'-axis and bissectrice of angle between two pions, where x'-axis is defined (see Fig. 2.3);

Three variables $s_{K\pi\pi}, s_{K\pi}, s_{\pi\pi}$ define the Dalitz plot. The angular components cover 4π angular space.

Kinematics in the $K\pi\pi$ rest frame (frame used in theoretical computation)

The final state particle's momenta are defined in the $K\pi\pi$ rest frame: $\pi^+(\overrightarrow{p}_{\pi^+})$, $\pi^-(\overrightarrow{p}_{\pi^-})$, $K^+(\overrightarrow{p}_{K^+})$, $\gamma(\overrightarrow{p}_\gamma)$. The notations are presented for the charged mode ($B^+ \to K^+\pi^+\pi^-\gamma$ decay), however expressions derived in this subsection are not changed considering neutral mode ($B^0 \to K^+\pi^0\pi^-\gamma$).

The z-axis in this case is defined as photon's direction by Eq. 2.3.

$$\overrightarrow{e}_z = \frac{\overrightarrow{p}_\gamma}{|\overrightarrow{p}_\gamma|} \tag{2.3}$$

The normal is defined as orthogonal vector to the $K\pi\pi$ plane by Eq. 2.4.

$$\overrightarrow{n} = \frac{[\overrightarrow{p}_{\pi^+} \times \overrightarrow{p}_{\pi^-}]}{|[\overrightarrow{p}_{\pi^+} \times \overrightarrow{p}_{\pi^-}]|} \tag{2.4}$$

The y-axis is chosen to be orthogonal to the z and $\overrightarrow{n}$ directions and defined by Eq. 2.5.

$$\overrightarrow{e}_y = \frac{[\overrightarrow{e}_z \times \overrightarrow{n}]}{|[\overrightarrow{e}_z \times \overrightarrow{n}]|} \tag{2.5}$$

Finally, x-axis is chosen to be orthogonal to y and z-axes: $\overrightarrow{e}_x = [\overrightarrow{e}_y \times \overrightarrow{e}_z]$.

Hence, the polar angle associated with the photon direction with respect to the $K\pi\pi$ plane is obtained from Eq. 2.6.

$$\cos\theta = \overrightarrow{e}_z \cdot \overrightarrow{n} \tag{2.6}$$

The allowed range of this angle is $[0;\pi]$.

Rotating the defined frame $\overrightarrow{e}_x$, $\overrightarrow{e}_y$, $\overrightarrow{e}_z$ around y-axis, such that $\overrightarrow{e}_z \to \overrightarrow{n}$, allows to define new axis $\overrightarrow{e}'_x$ placed in the $K\pi\pi$ plane (see Eq. 2.7).

$$\overrightarrow{e}'_x = \overrightarrow{e}_y \times \overrightarrow{n} \tag{2.7}$$

In this way a second orthogonal frame $\overrightarrow{e}'_x$, $\overrightarrow{e}_y$, $\overrightarrow{n}$ is defined.

In this rotated frame ($\overrightarrow{e}'_x$, $\overrightarrow{e}_y$, $\overrightarrow{n}$) the pions' three momenta are calculated with Eq. 2.8 and can be expressed in the original frame ($\overrightarrow{e}_x$, $\overrightarrow{e}_y$, $\overrightarrow{e}_z$) by Eq. 2.9.

$$\overrightarrow{p}_{\pi^+,\pi^-} = |\overrightarrow{p}_{\pi^+,\pi^-}|(\cos\phi_{\pi^+,\pi^-}\overrightarrow{e}'_x + \sin\phi_{\pi^+,\pi^-}\overrightarrow{e}_y) \tag{2.8}$$

$$\overrightarrow{p}_{\pi^+,\pi^-} = |\overrightarrow{p}_{\pi^+,\pi^-}|(\cos\theta\cos\phi_{\pi^+,\pi^-}; \sin\phi_{\pi^+,\pi^-}; -\sin\theta\cos\phi_{\pi^+,\pi^-}) \tag{2.9}$$

The difference of the ϕ_{π^+,π^-} angles (see Eq. 2.10) doesn't depend on angular variables but is a function of the Dalitz variables from Eqs. 2.11-2.14.

$$\delta = \phi_{\pi^-} - \phi_{\pi^+} \tag{2.10}$$

$$s_{K^+\pi^+\pi^-} = (p_{\pi^+} + p_{\pi^-} + p_{K^+})^2 \tag{2.11}$$

$$s_{K^+\pi^+} = (p_{\pi^+} + p_{K^+})^2 \tag{2.12}$$

$$s_{K^+\pi^-} = (p_{\pi^-} + p_{K^+})^2 \tag{2.13}$$

$$s_{\pi^+\pi^-} = (p_{\pi^+} + p_{\pi^-})^2 \tag{2.14}$$

In the $K\pi\pi$ rest frame $\overrightarrow{p}_{\pi^+} + \overrightarrow{p}_{\pi^-} + \overrightarrow{p}_{K^+} = 0$ and $|\overrightarrow{p}_{\pi^+,\pi^-,K^+}|$ can be expressed in terms of the Dalitz varibles. Since only two of them are independent, $s_{K^+\pi^-}$, $s_{K^+\pi^+}$ can be chosen. The relative angle between the three momenta of the two pions is given by Eq. 2.15.

$$\begin{aligned}\cos\delta &= \frac{\overrightarrow{p}_{\pi^+} \cdot \overrightarrow{p}_{\pi^-}}{|\overrightarrow{p}_{\pi^+}||\overrightarrow{p}_{\pi^-}|} = \frac{|\overrightarrow{p}_{K^+}|^2 - |\overrightarrow{p}_{\pi^+}|^2 - |\overrightarrow{p}_{\pi^-}|^2}{2|\overrightarrow{p}_{\pi^+}||\overrightarrow{p}_{\pi^-}|} \\ &= \frac{m^2_{\pi^+} + m^2_{\pi^-} + 2\cdot E_{\pi^+}\cdot E_{\pi^-} - s_{\pi^+\pi^-}}{2|\overrightarrow{p}_{\pi^+}||\overrightarrow{p}_{\pi^-}|}\end{aligned} \tag{2.15}$$

where E_{π^+,π^-}, m_{π^+,π^-} — energies and masses of π^+, π^-.

The angle defining the rotation in the $K\pi\pi$ plane is defined in Eq. 2.16.

$$\phi = \frac{\phi_{\pi^+} + \phi_{\pi^-}}{2} \tag{2.16}$$

where it is decided, that $\phi \in [0; 2\pi]$.

Hence, ϕ_{π^+,π^-} can be re-expressed as in Eq. 2.17.

$$\phi_{\pi^+,\pi^-} = \phi \mp \frac{\delta}{2} \tag{2.17}$$

These 5 variables ($s_{K^+\pi^+\pi^-}$, $s_{K^+\pi^+}$, $s_{K^+\pi^-}$, $\cos\theta$, ϕ) allow to fully describe $B \to K\pi\pi\gamma$ decay in the $K\pi\pi$ rest frame and can be used express $p_{K^+,\pi^-,K^+,\gamma}$ 4-momenta.

Transformations between the $K\pi\pi$ and the B-meson rest frame

The theoretical computation is done in terms of two angles $\theta \in [0; \pi]$, $\phi \in [0; 2\pi]$ and three Dalitz variables $s_{K^+\pi^+\pi^-}$, $s_{K^+\pi^+}$, $s_{K^+\pi^-}$. One needs to generate the events in terms of four momenta of the final state particle (FSP) as data analysis in the particle physics (see Fig. 2.1) is done in this conventional representation of the kinematics (cut on high energy photon, quality cut on the calorimeter's acceptance, etc). Once defined, these 4-momenta can be boosted to any other frame (laboratory, center of mass of the collision, etc), assuming boosting vector is given. In particular, boosting to the B-meson rest frame is a necessary step, since within this frame the final state particles 4-momenta are to the "EvtGen" generator. Moreover, since a photon direction is defined along z-axis in the $K\pi\pi$ rest frame, additional simultaneous randomization of absolute directions of $K\pi\pi\gamma$ is required using three Euler angles.

Below the following steps are given for the purpose of $((s_{K^+\pi^+\pi^-}, s_{K^+\pi^+}, s_{K^+\pi^-}, \phi, \cos\theta) \to (p^*_{\pi^+}, p^*_{\pi^-}, p^*_{K^+}, p^*_\gamma))$ conversion, where $p^*_{...}$ stands for the 4-momentum in the B-meson rest frame.

- Define momenta and energies of FSP in the $K\pi\pi$ reference frame using Eqs. 2.18-2.19;

$$|\overrightarrow{p}_i| = \frac{\left[(s_{K^+\pi^+\pi^-} - (\sqrt{s_{jk}} + m_i)^2)(s_{K^+\pi^+\pi^-} - (\sqrt{s_{jk}} - m_i)^2)\right]^{1/2}}{2s_{K^+\pi^+\pi^-}} \tag{2.18}$$

$$|\overrightarrow{p}_\gamma| = \frac{m_B^2 - s_{K^+\pi^+\pi^-}}{2\sqrt{s_{K^+\pi^+\pi^-}}} \tag{2.19}$$

  where $i, j, k = (\pi^+, \pi^-, K^+), (\pi^-, \pi^+, K^+), (K^+, \pi^+, \pi^-)$. The energies can be computed by $E_i = \sqrt{m_i^2 + |\overrightarrow{p}_i|^2}$:

- Three momenta of two pions can be calculated using Eq. 2.9. Kaon's three momentum is calculated from the 4 momenta conservation in the $K\pi\pi$ rest frame: $\overrightarrow{p}_{K^+} = -\overrightarrow{p}_{\pi^+} - \overrightarrow{p}_{\pi^-}$. As mentioned above, a photon is emitted along the z-axis, which together with Eq. 2.19 fully defines its 4-momentum. At this stage, the 4-momenta of $K^+, \pi^+, \pi^-, \gamma$ are defined.

- Define the boost along z-axis to the B-meson rest frame using the Lorentz transformations from Eqs. 2.20-2.22.

$$\overrightarrow{\beta} = \left(0;0;\frac{m_B^2 - s_{K^+\pi^+\pi^-}}{m_B^2 + s_{K^+\pi^+\pi^-}}\right) \quad (2.20)$$

$$E_i^* = \gamma(E_i + (\overrightarrow{\beta} \cdot \overrightarrow{p}_i)) \quad (2.21)$$

$$\overrightarrow{p}_i'^* = \overrightarrow{p}_i + (\gamma - 1)\overrightarrow{\beta}\frac{(\overrightarrow{\beta} \cdot \overrightarrow{p}_i)}{\beta^2} + \gamma\overrightarrow{\beta}E_i \quad (2.22)$$

where $i \in [K^+, \pi^+, \pi^-]$; m_B — mass of the B-meson; $\gamma = \frac{1}{\sqrt{1-\beta^2}}$.

- Randomize the directions of all FSP particles according to the Euler's transformation $\overrightarrow{p}_i^* = \hat{R}\overrightarrow{p}_i'^*$ with three random angles defined in Eq. 2.23-2.25.

$$\varphi^{rand} = 2\pi \cdot uniform(0,1) \quad (2.23)$$

$$\theta^{rand} = \arccos(uniform(-1,1)) \quad (2.24)$$

$$\chi^{rand} = 2\pi \cdot uniform(0,1) \quad (2.25)$$

where $\hat{R}$ — rotation matrix, and $uniform(x_{min}, x_{max})$ — function sampling random variable according to the uniform distribution from the range $[x_{min}, x_{max}]$.

On the other side, conversion of B rest frame to $K\pi\pi$ rest frame $(p^*_{\pi^+}, p^*_{\pi^-}, p^*_{K^+}, p^*_\gamma) \to (s_{K^+\pi^+\pi^-}, s_{K^+\pi^+}, s_{K^+\pi^-}, \phi$ allows to get the kinemtics of $B \to K\pi\pi\gamma$ decay in the representation convenient for performing an amplitude analysis. For instance, "GamPola"-fitter (to be introduced later) calculates model parameter values using 5-dimensional kinematics in the $K\pi\pi$ rest frame.

The steps for such conversion are given below:

- Calculate the Dalitz variables using Eqs. 2.11-2.13;
- Define the boost vector along the photon's direction to the $K\pi\pi$ rest frame:

$$\overrightarrow{\beta}^* = \frac{\overrightarrow{p}^*_\gamma}{E^*_{\pi^+} + E^*_{\pi^-} + E^*_{K^+}} \quad (2.26)$$

and perform Lorentz transformation according to Eqs. 2.21-2.22

- Calculate $\cos\theta$ using Eq. 2.6;
- Using Eq. 2.27 calculate ϕ_{π^+,π^-};

$$\phi_{\pi^+,\pi^-} = \arctan\left[\frac{\overrightarrow{e}_y \cdot \overrightarrow{p}_{\pi_{\pi^+,\pi^-}}}{\overrightarrow{e}'_x \cdot \overrightarrow{p}_{\pi_{\pi^+,\pi^-}}}\right] \quad (2.27)$$

- Calculate ϕ using Eq. 2.16;

- Angle $\phi \in [-\pi;\pi]$, while it should be in $[0;2\pi]$. Thus, correction for the range of the calculated ϕ is required (see Eq. 2.28).

$$\phi = \begin{cases} \phi + 2\pi & \delta = \phi_{\pi^-} - \phi_{\pi^+}, \phi < 0 \\ \phi & \delta = \phi_{\pi^-} - \phi_{\pi^+}, \phi \geq 0 \\ \phi + \pi & \delta \neq \phi_{\pi^-} - \phi_{\pi^+} \end{cases}, \tag{2.28}$$

  This correction is the case in C++ when $\arctan(x) \in [-\pi;\pi]$ (see Eq. 2.27) and $\arccos(x) \in [0;\pi]$ (follows from solving Eq. 2.15 with respect to δ).

2.1.3 . Total amplitude of $B \to K\pi\pi\gamma$ decay

Defined kinematical variables

$$(s_{K^+\pi^+\pi^-}, s_{K^+\pi^+}, s_{K^+\pi^-}, \phi, \cos\theta) = (s_{K\pi_1\pi_2}, s_{K\pi_1}, s_{K\pi_2}, \phi, \cos\theta)$$

represent a set of physical observables up to transformations described previously. The process $B \to K_{res}\gamma \to K\pi\pi\gamma$ is defined in terms of modelled decay rate as a function of these variables and model parameters. Total amplitude of $B \to K_{L,R}\gamma_{L,R} \to K\pi\pi\gamma_{L,R}$ is defined by Eq. 2.29.

$$\mathcal{M}(B \to K^{res}_{L,R}\gamma_{L,R}) \to K\pi\pi\gamma_{L,R} = M^{L,R}_{K_{res}} M(K_{res} \to K\pi\pi) BW_{K_{res}} \tag{2.29}$$

where $M^{L,R}_{K_{res}}$ — matrix element of $B \to K_{res}\gamma$ process, $BW_{K_{res}}$ — relativistic Breit-Wigner function of $s_{K\pi\pi}$ describing a resonance.

$\bar{B} \to \bar{K}_{res}\gamma$ decay amplitude

In general form an amplitude of $\bar{B} \to \bar{K}_{res}\gamma$ decay is defined by Eq. 2.30.

$$M(\bar{B} \to \bar{K}_{res}\gamma) = -4\frac{G_F}{\sqrt{2}} V^*_{ts} V_{tb} \Big(C_L \langle \bar{K}^{res}\gamma | \mathcal{O}_7 | \bar{B} \rangle + C_R \langle \bar{K}^{res}\gamma | \mathcal{O}'_7 | \bar{B} \rangle \Big) \tag{2.30}$$

where C_L, C_R — redefined C_7 and C'_7 correspondingly [73]; G_F — fermi coupling;

Substituting $\mathcal{O}_7$ and $\mathcal{O}'_7$ defined by Eqs. 1.48-1.49 followed by Fourier transform, matrix elements as in Eq. 2.31 can be obtained.

$$\langle \bar{K}^{res} | \bar{s}\sigma_{\mu\nu}(1 \pm \gamma^5) q^\nu b | \bar{B} \rangle \tag{2.31}$$

They are parametrized using 4-momenta and polarizations of B-meson and K_{res} so that they satisfy Ward identity for on-shell photon. In Eq. 2.31 operators $\sigma^{\mu\nu}$ and $\sigma^{\mu\nu}\gamma^5$ correspond to different parity, therefore final matrix element should contain the terms with different parity. For instance, if K_{res} is represented by 1^+, 1^- or 2^+ resonance, Egs. 2.32-2.34 are obtained.

$$M^{L,R}_{1^+} = M(B \to 1^+_{L,R}\gamma_{R,L}) = iV_{tb}V^*_{ts}\frac{\sqrt{2}em_b G_F}{2\pi^2} C_{L,R}(m_B^2 - s_{K\pi\pi}) T_1^{1^+} \tag{2.32}$$

$$M^{L,R}_{1^-} = M(B \to 1^-_{L,R}\gamma_{R,L}) = \pm iV_{tb}V^*_{ts}\frac{\sqrt{2}em_b G_F}{2\pi^2} C_{L,R}(m_B^2 - s_{K\pi\pi}) T_1^{1^-} \tag{2.33}$$

$$M^{L,R}_{2^+} = M(B \to 2^+_{L,R}\gamma_{R,L}) = \pm iV_{tb}V^*_{ts}\frac{\sqrt{2}em_b G_F}{2\pi^2} C_{L,R}\frac{(m_B^2 - s_{K\pi\pi})^{3/2}}{m_B} T_1^{2^+} \tag{2.34}$$

where $T_1^{1^+}, T_1^{1^-}, T_1^{2^+}$ — coefficients taken from parametrization of $B \to K_{res}\gamma$ matrix element. Although above expressions contain dependence of kinematics ($s_{K\pi\pi}$) since $s_{K\pi\pi} \ll m_B^2$, $s_{K\pi\pi}$ is replaces by nominal masses of decaying resonances in these expressions.

$K_{res} \to K\pi\pi$ decay amplitude

Consider the decay $K_{res} \to VP_k \to P_iP_jP_k$, where V — intermediate vector resonance with spin 1, P_i, P_j, P_k — final state pseudo scalar particles with spin 0 corresponding to K, π, π. The amplitude of such decay can be expressed by Eq. 2.35.

$$M^V_{(P_iP_j)P_k} = M(K_{res} \to VP_k)M(V \to P_iP_j)BW_V(s_{ij}) \tag{2.35}$$

where $BW_V(s_{ij})$ — Breit-Wigner function of subresonance.

Summation over all possible subresonances V gives the full expression for the amplitude of $K_{res} \to K\pi\pi$ decay (see Eq. 2.36).

$$M(K_{res} \to K\pi\pi) \quad = \quad \sum_V c_{ijk} M^V_{(P_iP_j)P_k} \tag{2.36}$$

where c_{ijk} — product of Clebsh-Gordon coefficients [96] defined by Eq. 2.37.

$$c_{ijk} = \langle J^{K_{res}}, M^{K_{res}} | J^V, M^V; J^{P_k}, M^{P_k} \rangle \langle J^V, M^V | J^{P_i}, M^{P_i}; J^{P_j}, M^{P_j} \rangle \tag{2.37}$$

where J — angular momentum, M — projection of angular momentum on z-axis.

For instance, in the case $B^+ \to K^+\pi^+\pi^-\gamma$ intermediate subresonances V are represented by $K^{*0} \to K^+\pi^-$ or $\rho^0 \to \pi^+\pi^-$ giving non-zero contributions from $c_{\pi^-K^+\pi^+}$ and $c_{\pi^+\pi^-K^+}$.

Another example is $B^0 \to K^+\pi^0\pi^-\gamma$ decay where non-zero contributions come from $c_{\pi^0\pi^-K^+}$, $c_{\pi^0K^+\pi^-}$ and $c_{\pi^-K^+\pi^0}$ corresponding to ρ^-, K^{*+} and K^{*0} intermediate subresonances correspondingly. Determination of c_{ijk} coefficients is done using tables from [97].

An amplitude of $V \to P_iP_j$ decay can be calculated using Eq. 2.38.

$$M(V \to P_iP_j) = g_{VP_iP_j}\varepsilon_V^\mu (p_i - p_j)_\mu \tag{2.38}$$

where $g_{VP_iP_j}$ — coupling constant calculated using Eq. 2.39.

$$g^2_{VP_iP_j} = \frac{48\pi M_V^5 \Gamma(V \to P_iP_j)}{\left[(M_V^2 - (m_i + m_j)^2)(M_V^2 - (m_i - m_j)^2)\right]^{3/2}} \tag{2.39}$$

where $\Gamma(V \to P_iP_j)$, M_V — nominal width and mass of V-resonance and sign of the coupling is defined from quark-pair-creation model [49]. For instance, in case of ρ and K^* corresponding couplings equal $g_{\rho\pi\pi} = -5.98$ and $g_{K^*K\pi} = 5.68$ respectively.

The $K_{res} \to VP_k$ subdecay is expressed differently depending on spin and parity of K_{res}. Experimentally observed resonances populating $K\pi\pi$ invariant mass spectrum include K_1^{1270} [98], K_1^{1400} [99] and resonances having strong evidences for existing: K_1^{*1410}, K_1^{*1680}, K_2^{1430} [100]. The properties of these resonances are summarized in the Table 2.1.

Resonance K_{res}	Spin	Parity	nominal mass, MeV	nominal width, MeV
K_1^{1270}	1	+	1253 ± 7	101 ± 12
K_1^{1400}	1	+	1403 ± 7	174 ± 13
K^{*1410}	1	−	1414 ± 15	232 ± 21
K^{*1680}	1	−	1718 ± 18	322 ± 10
K_2^{1430}	2	+	1427.3 ± 1.5	100 ± 2.1

Table 2.1. Properties of considered resonances

Information about intermediate resonances V can be found in the Table 2.2. In the present work the following decay chains of K_{res} are considered:

- $K_{res} \to K^{*892}\pi \to (K\pi)\pi$;
- $K_{res} \to \rho^{770}K \to (\pi\pi)K$;

where the decay channels $K_1(1270) \to K^{*0}(800)\pi$ [49, 73] and $K_1(1270) \to K^{*0}(1430)\pi$ [103] are not taken into account. Although existence of $K^{*0}(800)$ [101] has been confirmed [102] and $K_1(1270) \to K^{*0}(800)\pi$ is kinematically allowed, according to the Particle Data Group (PDG) [147] $K_1(1270)$ does not contain such contribution. The decay $K_1(1270) \to K^{*0}(1430)\pi$ is neglected, since its relative branching ratio is 2 % according to the measurements of Belle [104] and BaBar [105].

Subresonance V	Spin	Parity	nominal mass, MeV	nominal width, MeV
K^{*892}	1	−	891.66 ± 0.26	50.8 ± 0.9
ρ^{770}	1	−	775.26 ± 0.25	147.8 ± 0.9

Table 2.2. Properties of intermediate subresonances

In the following, the intermediated mesons are described by the relativistic Breit-Wigner functions with a fixed width for kaonic resonances :

$$BW_{K_{res}}(s_{K\pi\pi}) = \frac{1}{s_{K\pi\pi} - M^2_{K_{res}} + i\Gamma_{K_{res}}M_{K_{res}}} \tag{2.40}$$

where $M_{K_{res}}$ and $\Gamma_{K_{res}}$ stand for nominal mass and width of K_{res} resonance respectively, as well as for K^{*892} resonance

$$BW_{K^{*892}}(s_{K\pi}) = \frac{1}{s_{K\pi} - M^2_{K^{*892}} + i\Gamma_{K^{*892}}M_{K^{*892}}} \tag{2.41}$$

where M_{K^*} and Γ_{K^*} stand for nominal mass and width of K^* respectively. On the other hand, the parametrization of ρ^{770} has energy-dependent width according to [109].

$$BW_{\rho^{770}} = \frac{1}{s_{\pi\pi} + M^2_{\rho^{770}} + i\Gamma_{\rho^{770}}(s_{\pi_1\pi_2})M_{\rho^{770}}} \tag{2.42}$$

$$\Gamma_{\rho^{770}}(s_{\pi_1\pi_2}) = \Gamma_{\rho^{770}}\Big[\frac{s_{\pi_1\pi_2} - (m_{\pi_1} + m_{\pi_2})^2}{M_{\rho^{770}} - (m_{\pi_1} + m_{\pi_2})^2}\Big]^{3/2} \tag{2.43}$$

where M_ρ and Γ_ρ stand for nominal mass and width of ρ respectively.

$K_1^{1270/1400} \to K\pi\pi$ decay amplitude

According to the Table. 2.1, K_1^{1270} and K_1^{1400} have the same spin and parity. These resonances both decay to $K^*(892)$ and $\rho(770)$. Hence, decays of K_1^{1270} and K_1^{1400} are formalised in terms of the same functional expression representing the decay $1^+ \to 1^-0^- \to 0^-0^-0^-$. The matrix element of $1^+ \to 1^-0^-$ decay is defined as in Eq. 2.44.

$$M(1^+ \to 1^-0^-) = \varepsilon^\mu_{K_1} T_{\mu\nu} \varepsilon^{\nu*}_V \tag{2.44}$$

where $\varepsilon^\mu_{K_1}$, $\varepsilon^{\nu*}_V$ — polarization vectors of K_1 and V respectively; $T_{\mu\nu}$ — hadronic tensor.

Formula for $T_{\mu\nu}$ is obtained based on the properties of $1^+ \to 1^-0^-$. In particular, possible values of angular momentum between initial 1^+ and final 1^-0^- states can be $L \in 0, 1, 2$. However, from the parity conservation: $+1 = (-1)^L \cdot (-1) \cdot (-1)$ follows, that $L = 0, 2$. One can conclude that $T_{\mu\nu}$ should contain parity-even terms corresponding to mixture of $L = 0$ (S-wave) and $L = 2$ (D-wave) (see Eq. 2.45).

$$T_{\mu\nu} = f^V g_{\mu\nu} + h^V p^V_\mu p^{K_1}_\nu \tag{2.45}$$

where $g_{\mu\nu}$ — Minkowski tensor, f^V, h^V — hadronic form-factors. Therefore, substituting Eq. 2.45 $\to$ Eq. 2.44 and Eqs. 2.38, 2.44 $\to$ Eq. 2.35 $\to$ Eq. 2.36 expression for $M(K_1 \to K\pi\pi)$ is obtained as in Eq. 2.46.

$$M(K_1 \to K\pi\pi) = \sum_V c_{ijk} M^V_{(P_iP_j)P_k} = (\overrightarrow{\varepsilon}_{K_1} \cdot \overrightarrow{J}) \tag{2.46}$$

$$\overrightarrow{J} = C_1 \overrightarrow{p}_{\pi_1} - C_2 \overrightarrow{p}_{\pi_2} \tag{2.47}$$

with $\overrightarrow{\varepsilon}_{K_1} = \pm\frac{1}{\sqrt{2}}(1, \mp, 0)$ and $C_{1,2}$ defined by Eqs. 2.48. In Eq. 2.47 indexes $(1, 2)$ correspond to π_1, π_2 respectively and appear due to momenta conservation in the $K\pi\pi$ rest frame, where $\overrightarrow{p}_K = -\overrightarrow{p}_{\pi_1} - \overrightarrow{p}_{\pi_2}$.

$$\begin{aligned} C_{1,2} = {} & \pm f^{K_1} \sum_{V=\rho,K^*} c_{ijk} g_{VP_iP_j} BW_V(s_{ij}) \Big[\pm f^V + h^V \sqrt{s_{K\pi\pi}}(E_i - E_j) \\ - {} & \frac{m_i^2 - m_j^2}{s_{ij}} [f^V + h^V \sqrt{s_{K\pi\pi}}(E_i + E_j)] \Big] \end{aligned} \tag{2.48}$$

where f^{K_1} — contribution of each K_1 resonance within total amplitude: $f^{K_1^{1270}} = 1$ is fixed, while $f^{K_1^{1400}}$ is the model parameter, which can be adjusted; $E_{i,j}$ — energies of i, j particles respectively. Depending on

the decay of charged or neutral B-meson to ,for instance, $K^+\pi^+\pi^-\gamma$ or $K^+\pi^0\pi^-\gamma$ final state, $V=\rho^0, K^{*0}$ or $V=\rho^-, K^{*+}, K^{*0}$ respectively, and notation $V=\rho, K^*$ in the Eq. 2.48 covers such cases.

Final expressions for the form factors f^V, h^V are calculated in [106] using partial waves analysis [107] and listed in Eqs. 2.49-2.50.

$$f^V = -A_S^V - \frac{1}{\sqrt{2}}A_D^V \tag{2.49}$$

$$h^V = \frac{E_V}{\sqrt{s_{K\pi\pi}}|\overrightarrow{p}_V|^2}\Big[(1-\frac{\sqrt{s_V}}{E_V})A_S^V + (1+\frac{2\sqrt{s_V}}{E_V})\frac{1}{\sqrt{2}}A_D^V\Big] \tag{2.50}$$

where $E_V, \overrightarrow{p}_V$ — energy and 3-momentum of V-resonance; $A_{S,D}^V$ — partial amplitudes of S and D waves for $K_1 \to VP_k$ decay.

The observed K_1^{1270} and K_1^{1400} are considered to be mixtures of non mass eigenstates K_{1A} and K_{1B} and their mixing angle θ_{K_1} is defined according to Eq. 2.51 [108].

$$\begin{pmatrix}|K_1^{1270}\rangle\\ |K_1^{1270}\rangle\end{pmatrix} = \begin{pmatrix}\sin\theta_{K_1} & \cos\theta_{K_1}\\ \cos\theta_{K_1} & -\sin\theta_{K_1}\end{pmatrix}\cdot\begin{pmatrix}|K_{1A}\rangle\\ |K_{1B}\rangle\end{pmatrix} \tag{2.51}$$

Applying quark pair creation model [49], partial wave amplitudes $S^{K_1K^*\pi/K_1K\rho}$ and $D^{K_1K^*\pi/K_1K\rho}$ are calculated for K_{1A} and K_{1B} decays. In this case $A_{S,D}^V$ can be obtained from Eqs. 2.52-2.55.

$$A_S(K_1^{1270}\to K^*\pi/K\rho) = S^{K_1K^*\pi/K_1K\rho}(\sqrt{2}\sin\theta_{K_1} \mp \cos\theta_{K_1}) \tag{2.52}$$

$$A_D(K_1^{1270}\to K^*\pi/K\rho) = D^{K_1K^*\pi/K_1K\rho}(-\sin\theta_{K_1} \mp \sqrt{2}\cos\theta_{K_1}) \tag{2.53}$$

$$A_S(K_1^{1400}\to K^*\pi/K\rho) = S^{K_1K^*\pi/K_1K\rho}(\sqrt{2}\cos\theta_{K_1} \pm \sin\theta_{K_1}) \tag{2.54}$$

$$A_D(K_1^{1400}\to K^*\pi/K\rho) = D^{K_1K^*\pi/K_1K\rho}(-\cos\theta_{K_1} \pm \sqrt{2}\sin\theta_{K_1}) \tag{2.55}$$

In addition to θ_{K_1} and $f^{K_1^{1400}}$ model parameters, three hadronic phases $\varphi_S^{K^*}, \varphi_S^\rho, \varphi_D^\rho$ parametrizing f^V, h^V [112] are introduced in Eqs. 2.56-2.59. These phases responsible for the directions of S and D waves in the complex space.

$$f^{K^*} = -A_S^{K^*} - \frac{1}{\sqrt{2}}A_D^{K^*}e^{i\varphi_D^{K^*}} \tag{2.56}$$

$$f^\rho = (-A_S^\rho - \frac{1}{\sqrt{2}}A_D^\rho e^{i\varphi_D^\rho})e^{i\varphi_S^\rho} \tag{2.57}$$

$$h^{K^*} = \frac{E_{K^*}}{\sqrt{s_{K\pi\pi}}|\overrightarrow{p}_{K^*}|^2}\Big[(1-\frac{\sqrt{s_{K^*}}}{E_{K^*}})A_S^{K^*} + (1+\frac{2\sqrt{s_{K^*}}}{E_{K^*}})\frac{1}{\sqrt{2}}A_D^{K^*}e^{i\varphi_D^{K^*}}\Big] \tag{2.58}$$

$$h^\rho = \frac{E_\rho}{\sqrt{s_{K\pi\pi}}|\overrightarrow{p}_\rho|^2}\Big[(1-\frac{\sqrt{s_\rho}}{E_\rho})A_S^\rho + (1+\frac{2\sqrt{s_\rho}}{E_\rho})\frac{1}{\sqrt{2}}A_D^\rho e^{i\varphi_D^\rho}\Big]e^{i\varphi_S^\rho} \tag{2.59}$$

$K^{*1410/1680} \to K\pi\pi$ decay amplitude

Decays of $K^{*1410/1680} \to VP_k$ can be jointly described since K^{*1410} and K^{*1680} have the same spin and parity (1^-) and decay to the 1^-0^- state. The corresponding amplitude is given as in Eq. 2.60.

$$M(1^-\to 1^-0^-) = \varepsilon_{K_1^*}^\mu T_{\mu\nu}\varepsilon_V^{\nu *} \tag{2.60}$$

From angular momentum conservation between 1^- and 1^-0^- follows that $L = 0, 1, 2$. From parity conservation: $-1 = (-1)^L \cdot (-1) \cdot (-1)$ one obtains $L = 1, 3, 5, ...$ Thus allowed value of angular mometum is $L = 1$, which corresponds to P-wave and $T_{\mu\nu}$ is defined as in Eq. 2.61 [107].

$$T_{\mu\nu} = i f^V_{K_1^*} \epsilon_{\mu\nu\alpha\beta} p^\alpha_{K_1^*} p^\beta_V \tag{2.61}$$

where $\epsilon_{\mu\nu\alpha\beta}$ — fully anti-symmetric tensor; $f^V_{K_1^*}$ — model parameters of $K_1^* \to V P_k$ decay amplitude and it is assumed that $f^{K^{*0}}_{K_1^*} = f^{K^{*+}}_{K_1^*}$ and $f^{\rho^0}_{K_1^*} = f^{\rho^+}_{K_1^*}$. Substituting Eq. 2.61 $\to$ Eq. 2.60 and Eqs. 2.60,2.38 $\to$ Eq. 2.35 $\to$ Eq. 2.36 one obtains Eq. 2.62 where decay amplitude includes $f^{K^{*892}}_{K^{*1410}}, f^{K^{*892}}_{K^{*1680}}, f^{\rho}_{K^{*1410}}, f^{\rho}_{K^{*1680}}$ model parameters.

$$M(K_1^* \to K\pi\pi) = (\overrightarrow{\varepsilon}_{K_1^*} \cdot \overrightarrow{L}) = \overrightarrow{\varepsilon}_{K_1^*} \cdot \sum_{V=\rho,K^*} 2ic_{ijk} BW_V(s_{ij}) f^V_{K_1^*} g_{VP_iP_j} \sqrt{s_V} [\overrightarrow{p}_i \times \overrightarrow{p}_j] \tag{2.62}$$

where $\overrightarrow{\varepsilon}_{K_1^*} = \pm\frac{1}{\sqrt{2}}(1, \mp i, 0)$.

$K_2^{1430} \to K\pi\pi$ decay amplitude

According to the Table. 2.1, K_2^{1430} is in 2^+ spin-parity state and decays to 1^-0^-. Allowed values of angular momentum are $L = 1, 2, 3$ and from parity conservation: $+1 = (-1)^L \cdot (-1) \cdot (-1) \to L = 0, 2, 4, ...$ Only $L = 2$ satisfies both conservation laws and corresponds to D-wave, so the amplitude is defined by Eq. 2.63.

$$M(2^+ \to 1^-0^-) = \varepsilon^{\mu\nu} T_{\mu\nu\rho} \varepsilon_V^{\rho *} \tag{2.63}$$

where $\varepsilon^{\mu\nu}$, $T_{\mu\nu\rho}$ — polarization and hadronic tensors defined by Eqs. 2.64-2.65.

$$\varepsilon^{\mu\nu} = \frac{1}{\sqrt{2}}(\varepsilon^\mu_\pm \varepsilon^\nu_0 + \varepsilon^\nu_\pm \varepsilon^\mu_0) \tag{2.64}$$

$$T_{\mu\nu\rho} = i f^V_{K_2} \epsilon_{\mu\nu\alpha\beta} P^\alpha q^\beta P_\rho \tag{2.65}$$

where $\varepsilon^\mu_0 = (0, 0, 0, 1)$ — zero component of K_2 polarization; $\varepsilon^\mu_\pm = \pm\frac{1}{\sqrt{2}}(0, 1, \mp i, 0)$ — transverse polarization of K_2; $f^V_{K_2}$ — adjustable model parameters with $f^{K^{*0}}_{K_2} = f^{K^{*+}}_{K_2}$ and $f^{\rho^0}_{K_2} = f^{\rho^+}_{K_2}$; p_{K_2},p_V 4-momenta are re-defined as $P = p_{K_2} + p_V$, $q = p_{K_2} - p_V$.

Applying the chain of substitutions: Eqs. 2.65,2.64 $\to$ Eq. 2.63 and (Eq. 2.63,2.38) $\to$ Eq. 2.35 $\to$ Eq. 2.36 allows to obtain Eq. 2.66 with model parameters $f^{K^*}_{K_2}$ and $f^{\rho}_{K_2}$.

$$\begin{aligned} M(K_2 \to K\pi\pi) &= (\overrightarrow{\varepsilon}_\pm \cdot \overrightarrow{K}) = \overrightarrow{\varepsilon}_\pm \cdot \Big(\sum_{V=\rho,K^*} 2\sqrt{2} i g_{VP_iP_j} c_{ijk} f^V_{K_2} \sqrt{s_{K\pi\pi}} BW_V(s_{ij}) \cdot \\ &\cdot \quad \big((\overrightarrow{\varepsilon}_0 \cdot \overrightarrow{p}_V)[\overrightarrow{p}_i \times \overrightarrow{p}_i] + (\overrightarrow{\varepsilon}_0 \cdot [\overrightarrow{p}_i \times \overrightarrow{p}_i])\overrightarrow{p}_V \big) \Big) \end{aligned} \tag{2.66}$$

2.1.4 . Decay rate

Helicity currents $\overrightarrow{J}, \overrightarrow{L}, \overrightarrow{K}$ defined by Eqs. 2.47, 2.62, 2.66 of $K_{res} \to K\pi\pi$ do not contain information about photon helicity, however they can be redefined as in Eqs. 2.67-2.71 taking into account $M^{L,R}_{K_{res}}$ (see Eqs. 2.32-2.34).

$$\overrightarrow{J}^{1270} \to \overrightarrow{J}^{1270}_{L,R} = \overrightarrow{J}^{1270}(\theta_{K_1}, \varphi_S^{K^*}, \varphi_S^{\rho}, \varphi_D^{\rho}) \quad (2.67)$$

$$\overrightarrow{J}^{1400} \to \overrightarrow{J}^{1400}_{L,R} = \frac{m_B^2 - M^2_{K_1^{1400}}}{m_B^2 - M^2_{K_1^{1270}}} \overrightarrow{J}^{1400}(\theta_{K_1}, \varphi_S^{K^*}, \varphi_S^{\rho}, \varphi_D^{\rho}, f^{K_1^{1400}}) \quad (2.68)$$

$$\overrightarrow{L}^{1410} \to \overrightarrow{L}^{1410}_{L,R} = \pm\left[\frac{m_B^2 - M^2_{K_1^{1410}}}{m_B^2 - M^2_{K_1^{1270}}}\right] \overrightarrow{L}^{1410}(f^{K^*}_{K^{*1410}}, f^{\rho}_{K^{*1410}}) \quad (2.69)$$

$$\overrightarrow{L}^{1680} \to \overrightarrow{L}^{1680}_{L,R} = \pm\left[\frac{m_B^2 - M^2_{K_1^{1680}}}{m_B^2 - M^2_{K_1^{1270}}}\right] \overrightarrow{L}^{1680}(f^{K^*}_{K^{*1680}}, f^{\rho}_{K^{*1680}}) \quad (2.70)$$

$$\overrightarrow{K}^{1430} \to \overrightarrow{K}^{1430}_{L,R} = \pm\left[\frac{(m_B^2 - M^2_{K_2^{1430}})^{\frac{3}{2}}}{m_B(m_B^2 - M^2_{K_1^{1270}})}\right] \overrightarrow{K}^{1430}(f^{K^*}_{K_2}, f^{\rho}_{K_2}) \quad (2.71)$$

Since radiative $b \to s\gamma$ transition occurs either through $b \to s_L\gamma_L$ or $b \to s_R\gamma_R$, the final states are different and differential decay rate is a sum of two decay rates (see Eq. 2.72).

$$\frac{d\Gamma}{d\cos\theta d\phi ds_{K\pi_1} ds_{\pi_1\pi_2} ds} = \Big[\frac{d\Gamma_L}{d\cos\theta d\phi ds_{K\pi_1} ds_{\pi_1\pi_2} ds}\Big] + \Big[\frac{d\Gamma_R}{d\cos\theta d\phi ds_{K\pi_1} ds_{\pi_1\pi_2} ds}\Big] \quad (2.72)$$

The L/R differential decay rates in this case are defined in Eq. 2.73.

$$\begin{aligned}\frac{d\Gamma_{L/R}}{d\cos\theta d\varphi ds_{K\pi_1} ds_{\pi_1\pi_2} ds} &\propto \frac{1 \pm \lambda}{2}(m_B^2 - m^2_{K_1^{1270}})^2\Big(\overrightarrow{\varepsilon}_{L,R} \cdot (\overrightarrow{J}^{1270}_{L,R} + \\ &+ \overrightarrow{J}^{1400}_{L,R} + \overrightarrow{L}^{1410}_{L,R} + \overrightarrow{L}^{1680}_{L,R} + \overrightarrow{K}^{1430}_{L,R})\Big)^2 \end{aligned} \quad (2.73)$$

where λ is the absolute value of a photon polarization ($\lambda = 1$ for SM); $\overrightarrow{\varepsilon}_{L,R} = \overrightarrow{\varepsilon}_{K_1} = \overrightarrow{\varepsilon}_{K_1^*} = \overrightarrow{\varepsilon}_{\pm} = \frac{1}{\sqrt{2}}(\pm 1; -i; 0)$ — K_{res} polarization vector.

The expressions defined by Eq. 2.72-2.73 form the master formula used for the $B \to K\pi\pi\gamma$ modelling with several free parameters summarized in the Table. 2.3.

Label	Origin	Type
λ	Decay rate	Real
θ_{K_1}	$K_1^{1270/1400}$	Real
$\varphi^D_{K^*}$		Real
φ^S_ρ		Real
φ^D_ρ		Real
$f^{K_1^{1400}}$	K_1^{1400}	Complex
$f^{K^*}_{K^{*1410}}$	K_1^{*1410}	Complex
$f^{\rho}_{K^{*1410}}$		Complex
$f^{K^*}_{K^{*1680}}$	K_1^{*1680}	Complex
$f^{\rho}_{K^{*1680}}$		Complex
$f^{K^*}_{K_2}$	K_2^{1430}	Complex
$f^{\rho}_{K_2}$		Complex

Table 2.3. Summary of the model parameters describing $B \to K\pi\pi\gamma$ decay

2.2 . "GamPola" software

Having defined the model of $B \to K\pi\pi\gamma$ decay, it is applied to produce Monte-Carlo data within the generator engine. In addition, on the basis of the model, the likelihood function is built and the fit of the external dataset is performed.

In the present work, the software incorporating both generator and fitter capabilities is called "GamPola" (Gamma Polarization) and designed for generating and analysing events in the process $B \to K_{res}\gamma$, where K_{res}: $K_1^{1270}, K_1^{1400}, K^{*1410}, K^{*1680}, K_2^{1430}$. The software was developed according to the general principles of the Object-Oriented Design [113].

The overall functionality of "GamPola" can be split on several zones of responsibility:

- Within a block of classes inheriting from *GInterfaceForMathFunctions* differential decay rate as in Eq. 2.73 is calculated. The following sub-classes are inherited:
  - *GKinematics* — class implementing all kinematics described in the Subsection 2.1.2 and the Algorithm 1 for generating events within kinematical bounds;
  - *GBreitWigner* — implements relativistic Breit-Wigner functions of Eqs. 2.40-2.43;
  - *GCouplingConstants* — implements Eq. 2.39 taking into account a correct sign of the couplings;
  - *GQuarkPairCreationModel* — includes theoretical expressions defined in Eqs. 2.52-2.55 and formulas for quark pair creation model [49];
  - Within the class *GFormFactors* Eqs. 2.56-2.59 are defined, representing form-factors of $K_1 \to VP_k$ decay amplitude;
  - *GResVector* contains expressions for $C_{1,2}$, $\overrightarrow{J}_{L,R}$ and $M(K_1 \to K\pi\pi)$ defined in Eqs. 2.46-2.48, where all helicity currents are redefined according to Eqs. 2.32-2.34;

  - In the similar way, *GResPseudoVector* implemets Eqs. 2.62,2.69,2.70 defining the amplitude of 1^- resonances;
  - *GResTensor* class implements the amplitude of $2^+ \to 1^-0^- \to 0^-0^-0^-$ decay through Eqs. 2.66,2.71;
  - Differential decay rate as in Eqs. 2.72,2.73 is defined within *GDecayRate* class. While calculating the amplitudes of K_{res} within above classes, the side functionality is present as well: lengthy expressions of functions of kinematical variables are groupped near model parameters and cached into multidimensional array. It is done in order to reuse them during the fit procedure;

- *GEventsGenerator* implements methods for producing samples of events in terms of distributions of $(s_{K\pi_1\pi_2}, s_{K\pi_1}, s_{K\pi_2}, \phi, \cos\theta)$ kinematical variables. It includes the Algorithm 2 and serves as a core of generator engine;

- *GGenerator* is a wrapper around *GEventsGenerator* and implements an interface to the generator core: setting model parameter values, launching event generation, writing output distributions to the file;

- *GEventsAnalyzer* — core of the fitter (to be discussed below) and implements methods for reading and writing normalization part of the decay rate, initializes starting point of the model parameters, performs likelihood fit based on *Minuit* [111];

- *GFitter* class is an interface class to *GEventsAnalyzer* and allows to set data for the fit, forward model parameters initialization to the fitter, generate normalization part of the decay rate;

- Class *GSymbolicMathFunctions* synchronously shadows decay rate calculation from *GInterfaceForMathFunctions* block of classes. It redefines the lengthy functions of kinematical variables as the symbolic variables (which are cached during decay rate calculation) using 3rd party library [110];

- Having obtained symbolic expression of the decay rate from *GSymbolicMathFunctions* and defined approach of caching of functions of kinemtical variables within *GInterfaceForMathFunctions*, *GSymbolicExpressionsAnalyzer* performs a split of the symbolic decay rate in two sets: expressions $g_i(\overrightarrow{\alpha})$ purely depending on model parameters and functions $f_i(\overrightarrow{\Omega})$ depending only on kinematical variables as in Eq. 2.74.

$$\frac{d\Gamma}{d\cos\theta d\phi ds_{K\pi_1} ds_{\pi_1\pi_2} ds_{K\pi\pi}} = \sum_i^{N_\Sigma} f_i(\overrightarrow{\Omega}) g_i(\overrightarrow{\alpha}) \tag{2.74}$$

  where N_Σ — number of terms within expanded expression of the decay rate, $\overrightarrow{\Omega} = (s_{K\pi\pi}, s_{K\pi_1}, s_{\pi_1\pi_2}, \phi, \cos\theta)$, $\overrightarrow{\alpha}$ — set of model parameters listed in the Table. 2.3;

- Utility class *GExCompiler* creates two *.cxx* files filled with functions depending purely on kinematics and model parameters respectively. It responsible for creating dynamic library of *.so* files and loading functions from these libraries to the program scope;

- *GBF2Ana* performs χ^2 fit of the histograms of external data and adjusts $\overrightarrow{\alpha}$ accordingly.

- In order to access the goodness of fit, *GGOFKolmogorovTest* implements two-sided Kolmogorov-Smirnov test for unbinned data;

2.2.1 . GamPola event generator

This section contains the description of the software used for the generation of $B \to K\pi\pi\gamma$ events. The parameters $\overrightarrow{\alpha}$ from Eq. 2.74 remain unchanged during events generation process, while $\overrightarrow{\Omega}$ is sampled from allowed kinematical region. The generator originally outputs a set of $s_{K\pi\pi}, s_{K\pi_1}, s_{\pi_1\pi_2}, \theta, \phi$ variables, however this representation can be transformed to the 4-momenta using *GKinematics*.

Generation of kinematically allowed events

Without introducing any physical interactions, kinematics of $K\pi\pi\gamma$ final state is constrained. The Algorithm 1 allows to obtain $\overrightarrow{\Omega}$ satisfying these constraints and outputs 1 event.

Algorithm 1: Kinematically allowed event

```
Result: (cos θ, φ, s_{Kπ1}, s_{π1π2}, s_{Kππ})
while 1 do
    s_{Kππ} = uniform(s_min, s_max);
    s_{π1π2} = uniform((m_{π1} + m_{π2})^2, (√s_max − m_K)^2);
    s_{Kπ1} = uniform((m_K + m_{π1})^2, (√s_max − m_{π2})^2);
    s^min_{π1π2} and s^max_{π1π2} calculations using Eqs. 2.78-2.79;
    if E*_{π2} > m_{π2} and s_{π1π2} ≥ s^min_{π1π2} and s_{π1π2} ≤ s^max_{π1π2} then
        break the loop;
    end
end
cos θ = uniform(−1, 1);
φ = uniform(0, 2π);
```

where in this work $s_{min} = 1\ \mathrm{GeV}^2/\mathrm{c}^4$ and $s_{max} = 4\ \mathrm{GeV}^2/\mathrm{c}^4$. Within if-block, first condition corresponds to the fact that $s_{K\pi_1} < (\sqrt{s_{K\pi\pi}} - m_{\pi_2})^2$ is not necessarily hold, because of s_{max} constant instead of $s_{K\pi\pi}$ variable. The second and third conditions define the boundaries of the Dalitz plot. Having obtained a set of the Dalitz variables $(s_{K\pi_1}, s_{\pi_1\pi_2}, s_{K\pi\pi})$, $s_{K\pi_2}$ can be obtained from Eq. 2.75.

$$s_{K\pi_1} + s_{K\pi_2} + s_{\pi_1\pi_2} = s_{K\pi\pi} + m_K^2 + m_{\pi_1}^2 + m_{\pi_2}^2 \tag{2.75}$$

The formulas for $s^{min}_{\pi_1\pi_2}$ and $s^{max}_{\pi_1\pi_2}$ are defined as in Eqs. 2.78-2.79. Those bounds are the functions of $s_{K\pi_1}$ and $s_{K\pi\pi}$ defining the limits of the Dalitz plot [114] on the $s_{K\pi_1}$—$s_{\pi_1\pi_2}$ plane. In the abscence of any physical interaction one expects the Dalitz plot to be flat, however since dependence on $s_{K\pi\pi}$ is present, then for instance: $s_{K\pi\pi} \downarrow\to E^*_{\pi_2} \downarrow\to (E^*_{\pi_1} + E^*_{\pi_2})^2 \downarrow\to s^{min}_{\pi_1\pi_2} \downarrow, s^{max}_{\pi_1\pi_2} \downarrow$ leads to decreasing of the region surrounded by Dalitz plot. The decreasing of a region is limited and $s^{min}_{\pi_1\pi_2}, s^{max}_{\pi_1\pi_2} \nrightarrow 0$. The overall effect of changing $s_{K\pi\pi}$ is shown on the Fig. 2.8, where three contours corresponding to three different constant values of $s_{K\pi\pi}$ are given. And since the Dalitz plot with high $s_{K\pi\pi}$ overlaps with the one having low $s_{K\pi\pi}$ value, one expects increasing the density of events in the overlapping region of $s_{K\pi_1}$—$s_{\pi_1\pi_2}$ plane. The

same reasoning holds for $s_{K\pi_2}$—$s_{\pi_1\pi_2}$ and $s_{K\pi_1}$—$s_{K\pi_2}$ planes (see Fig. 2.7).

$$E^*_{\pi_1} = \frac{s_{K\pi_1} - m^2_K + m^2_{\pi_1}}{2\sqrt{s_{K\pi_1}}} \tag{2.76}$$

$$E^*_{\pi_2} = \frac{s_{K\pi\pi} - s_{K\pi_1} - m^2_{\pi_2}}{2\sqrt{s_{K\pi_1}}} \tag{2.77}$$

$$s^{min}_{\pi_1\pi_2} = (E^*_{\pi_1} + E^*_{\pi_2})^2 - (\sqrt{E^{2*}_{\pi_1} - m^2_{\pi_1}} + \sqrt{E^{2*}_{\pi_2} - m^2_{\pi_2}})^2 \tag{2.78}$$

$$s^{max}_{\pi_1\pi_2} = (E^*_{\pi_1} + E^*_{\pi_2})^2 - (\sqrt{E^{2*}_{\pi_1} - m^2_{\pi_1}} - \sqrt{E^{2*}_{\pi_2} - m^2_{\pi_2}})^2 \tag{2.79}$$

The distributions of 5D kinematics in the absence of physics interactions are depicted in Figs. 2.4-2.7.

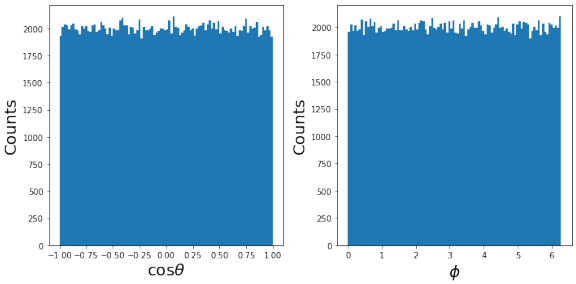


Figure 2.4. Angular distributions are flat as they are generated once condition of while-loop termination is met in the Algorithm 1.

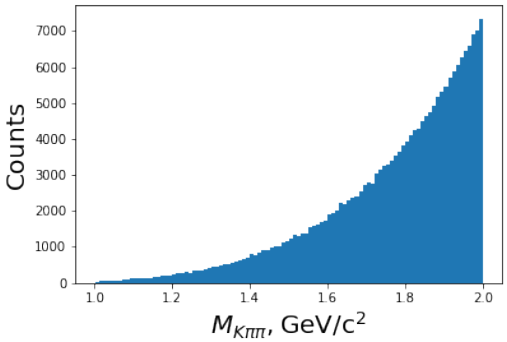


Figure 2.5. $M_{K\pi\pi} = \sqrt{s_{K\pi\pi}}$ distribution is not flat as "if"-condition in the Algorithm 1 depends on $s_{K\pi\pi}$.

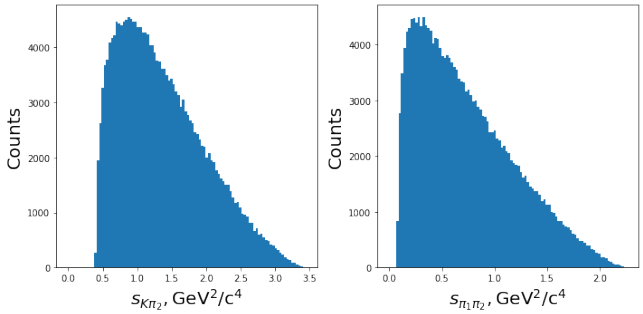


Figure 2.6. Distributions of $s_{K\pi_2}$ and $s_{\pi_1\pi_2}$ variables in the abscence of physics interaction have highest values near $s_{K\pi_2} = M^2_{K\pi_2}$ and $s_{\pi_1\pi_2} = M^2_{\pi_1\pi_2}$ correspondingly.

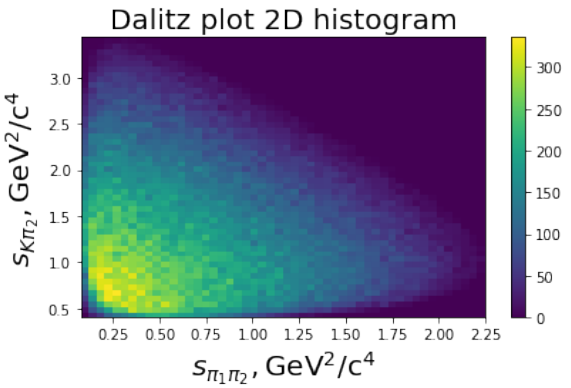


Figure 2.7. The dalitz plot is not flat since $s_{K\pi\pi}$ is not a constant. Dense region is created in the lower region of the plane due to continous overlapping of the flat Dalitz plots corresponding to different $s_{K\pi\pi}$ values.

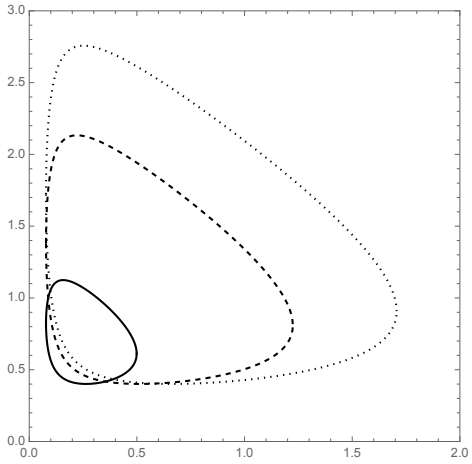


Figure 2.8. Contours of Dalitz plots given for different $s_{K\pi\pi}$ values, where the smallest $s_{K\pi\pi}$ corresponds to the smallest contour.

2.2.2 . Event generation

Event in the kinematically allowed region is further submitted to the physics modelling part, where decay rate is evaluated as a function of this event and a set of predefined model parameter values $\overrightarrow{\alpha}$. The set of $\overrightarrow{\alpha}$ is obtained to visually describe the real-data kinematical distributions $\overrightarrow{\Omega}$ and the procedure is given in the following. Sampling from decay rate function is a two-step process. First step is defined by Eq. 2.80 and made only once for given set of $\overrightarrow{\alpha}$. The second step is an event generation itself and summarized by Algorithm 2, where $\mathrm{DR}_{max}(\overrightarrow{\alpha})$ is assumed to be computed from Eq. 2.80.

$$\mathrm{DR}_{max}(\overrightarrow{\alpha}) = \max_{\overrightarrow{\Omega}} \left[\frac{d\Gamma}{d\overrightarrow{\Omega}}(\overrightarrow{\Omega}, \overrightarrow{\alpha}) \right] \tag{2.80}$$

Algorithm 2: Event sampling from the decay rate

```
Result: Ω⃗
while 1 do
    Ω⃗* = (s_Kππ, s_Kπ1, s_π1π2, cos θ, φ) ← Algorithm 1
    F(Ω⃗*, α⃗) = dΓ/dΩ⃗;
    u = uniform(0, DR_max(α⃗));
    if F(Ω⃗*, α⃗) > u then
        Ω⃗ = Ω⃗*
        break the loop
    end
end
```

Applying Algorithm 2 N times one obtains a set of N events simulating the physics of $B \to K\pi\pi\gamma$ decay.

Having introduced physics interactions through decay rate with predefined model parameters allows to get the following distributions of the kinematical variables for $B^+ \to K^+\pi_1^+\pi_2^-\gamma$ (charged) and $B^0 \to K^+\pi_1^0\pi_2^-\gamma$ (neutral) modes (see Figs. 2.9-2.18). In particular, for charged mode on the Figs. 2.9,2.11,2.13 one observes contributions from $K^{*0} \to K^+\pi^-$ and $\rho \to \pi^+\pi^-$ decays. In the case of neutral mode on the Figs. 2.10,2.14 additional peak associated with $K^{*+} \to K^+\pi^0$ decay is observed. The angular distributions of the neutral mode (Fig. 2.16) are modified comparing to the charged one (Fig. 2.15), meanwhile invariant mass distribution of $K\pi\pi$ remains almost unchanged for the two cases.

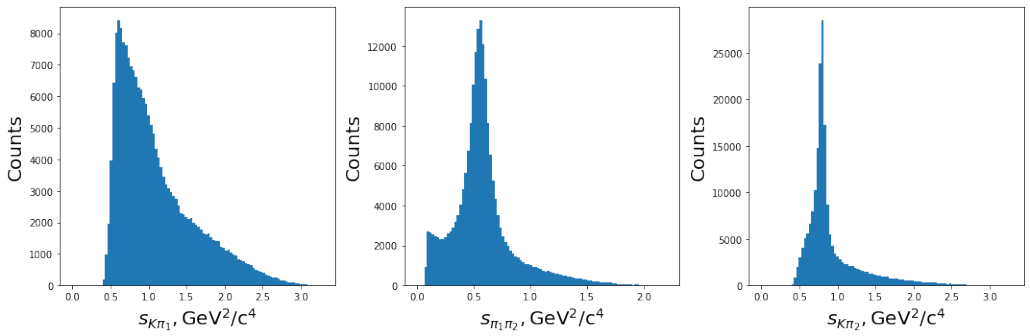


Figure 2.9. $B^+ \to K^+\pi_1^+\pi_2^-\gamma$: bulk of events is observed in the low $s_{K\pi_1}$ (left) region, Dalitz variables $s_{\pi_1\pi_2}$ (middle) and $s_{K\pi_2}$ (right) are peaked around nominal values of $M^2_{\rho^0}$ and $M^2_{K^{0*}}$ respectively.

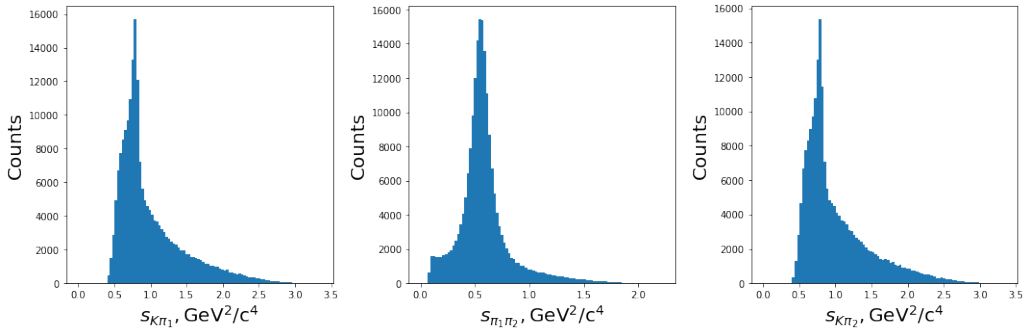


Figure 2.10. $B^0 \to K^+\pi_1^0\pi_2^-\gamma$: the peak corresponding to K^{*+} appears on the left histogram. On the right distribution $s_{K\pi_2}$ an excess of events appears further away from $s_{K\pi_2} = M^2_{K^{*0}}$ region, comparing to the charged case.

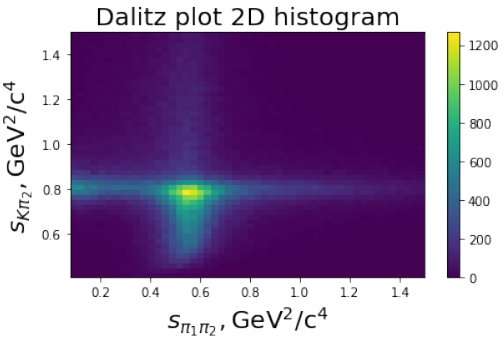


Figure 2.11. $B^+ \to K^+\pi_1^+\pi_2^-\gamma$: the Dalitz plot in these coordinates is plotted to compare with Fig. 2.7.

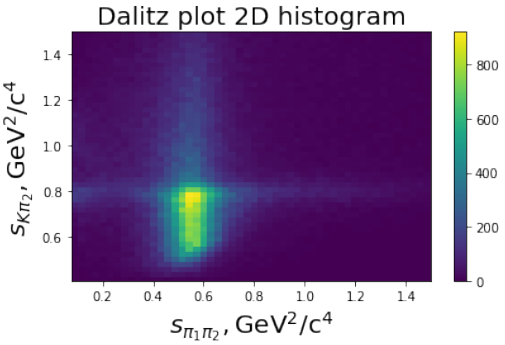


Figure 2.12. $B^0 \to K^+\pi_1^0\pi_2^-\gamma$: comparing to the charged case, excess of events is present and associated with additional $K^{*+} \to K^+\pi^0$ decay.

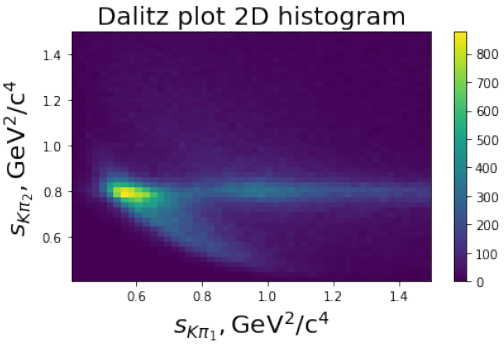


Figure 2.13. $B^+ \to K^+\pi_1^+\pi_2^-\gamma$: The dalitz plot contains resonances structure corresponding to sub-decays into K^{0*} and ρ^0

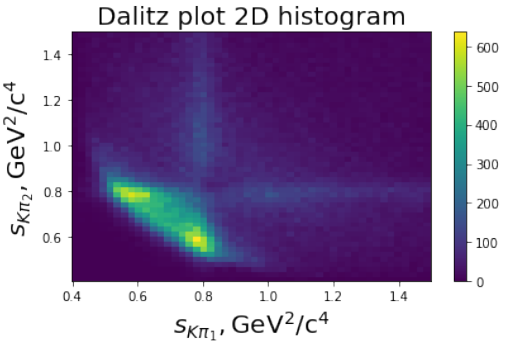


Figure 2.14. $B^0 \to K^+\pi_1^0\pi_2^-\gamma$: The dalitz plot also contains resonance corresponding to sub-decay into K^{*+}.

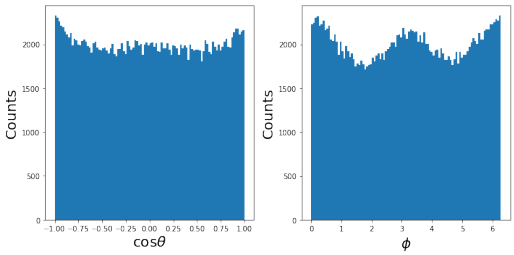


Figure 2.15. $B^+ \to K^+\pi_1^+\pi_2^-\gamma$: angular distributions

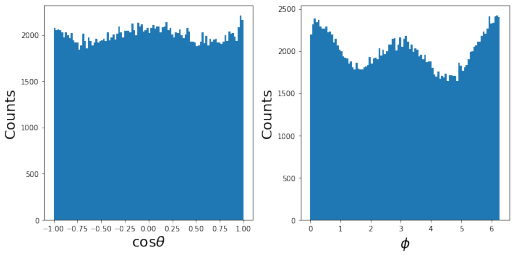


Figure 2.16. $B^0 \to K^+\pi_1^0\pi_2^-\gamma$: angular distributions

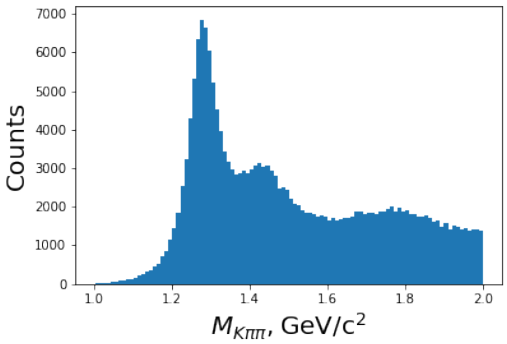


Figure 2.17. $B^+ \to K^+\pi_1^+\pi_2^-\gamma$ channel: $M_{K\pi\pi} = \sqrt{s_{K\pi\pi}}$ distribution now includes several K_{res}.

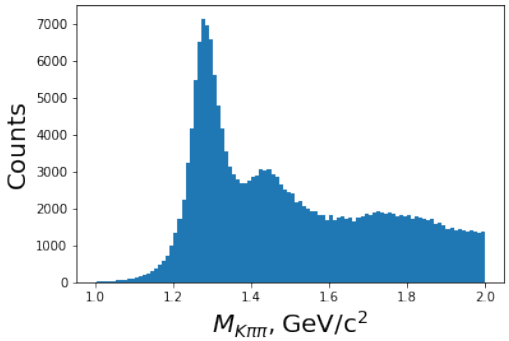


Figure 2.18. $B^0 \to K^+\pi_1^0\pi_2^-\gamma$ channel: distribution is very similar to the charged case.

2.3 . GamPola fitter

The fitting procedure is an optimization of the objective function of multiple variables and parameters with respect to the variables. In the context of the probability density function modelling (normalized decay rate as in Eq. 2.81), one refers to the optimization of the likelihood (see Eq. 2.82) as a function of $\vec{\alpha}$ model parameters and sample of events represented by kinematical variables $\vec{\Omega}_i$. The product of values of the same probability density functions signifies that events within a sample are independent and identically distributed. In the frame of maximum likelihood approach [128] one estimates a set of model parameter values $\vec{\alpha}^*$ maximizing joint probability of a given sample of events $\hat{\Omega}$ to occur (see Eq. 2.83).

$$\text{p.d.f}(\vec{\Omega}, \vec{\alpha}) = \frac{\frac{d\Gamma}{d\vec{\Omega}}(\vec{\Omega}, \vec{\alpha})}{\int_{\vec{\Omega}} \frac{d\Gamma}{d\vec{\Omega}}(\vec{\Omega}, \vec{\alpha}) d\vec{\Omega}} \tag{2.81}$$

$$L(\hat{\Omega}|\vec{\alpha}) = \prod_i^{N_{\text{evt}}} \text{p.d.f}(\vec{\Omega}_{\text{i}}, \vec{\alpha}) \tag{2.82}$$

$$\vec{\alpha}^* = \underset{\vec{\alpha}}{\text{argmax}}\, L(\hat{\Omega}|\vec{\alpha}) \tag{2.83}$$

where L — likelihood function; $\hat{\Omega}$ — matrix of $5 \times N_{\text{evt}}$ size, describing kinematics of data sample (number 5 stands for number of kinematical variables $\vec{\Omega}$); N_{evt} — number of events within a dataset; i — event index; $\vec{\alpha}^*$ — estimated set of the model parameters from likelihood function. On practice, integration over $\vec{\Omega}$ in Eq. 2.81 is replaced by summation over large number N_Σ of kinematically allowed values $\vec{\Omega}_i$ (see Eq. 2.84).

$$\text{Norm} = \int \frac{d\Gamma}{d\vec{\Omega}}(\vec{\Omega}, \vec{\alpha}) d\vec{\Omega} \to \sum_i^{N_\Sigma} \frac{d\Gamma}{d\vec{\Omega}}(\vec{\Omega}, \vec{\alpha}) \Delta\vec{\Omega}(\vec{\Omega}_i) = \sum_i^{N_\Sigma} \frac{d\Gamma}{d\vec{\Omega}}(\vec{\Omega}, \vec{\alpha}) \frac{\Omega}{N_\Sigma} \tag{2.84}$$

where $N_\sum$ — large number of kinematically allowed events; $\Delta\overrightarrow{\Omega}_i$ — infinitisemately small volume of kinematical region, which size depends on $\overrightarrow{\Omega}$; Ω — full volume of kinematically allowed region. It is assumed that having large enough $N_\sum$ leads to $\Delta\overrightarrow{\Omega} = \Delta\overrightarrow{\Omega}_1 = \Delta\overrightarrow{\Omega}_2 = ... = \Delta\overrightarrow{\Omega}_{N_\sum}$ and breaks dependence on $\overrightarrow{\Omega}$.

Computation of Eq. 2.82 usually leads to the underflow of a floating point precision. In addition, taking derivates $\partial L(\hat{\Omega}|\overrightarrow{\alpha})/\partial\alpha_j$ or $\partial^2 L(\hat{\Omega}|\overrightarrow{\alpha})/\partial\alpha_j\partial\alpha_k$ during the fit is complicated since it requires iterative or recursive evaluation. Thus one replaces $L(\hat{\Omega}|\overrightarrow{\alpha}) \to \log L(\hat{\Omega}|\overrightarrow{\alpha})$, since logarithm of positively defined function preserves all its extreme points. In most cases the sotware used for the fit provides an interface for minimization of the given function and not maximization, hence one minimizes $-\log L(\hat{\Omega}|\overrightarrow{\alpha})$. Substituting Eq. 2.84 $\to$ Eq. 2.81 $\to$ Eq. 2.82 and applying $-\log$ yields Eq. 2.85.

$$\mathcal{L} = -\log L(\hat{\Omega}|\overrightarrow{\alpha}) = -\sum^{N_{\text{evt}}} \log\Big[\frac{d\Gamma}{d\overrightarrow{\Omega}}(\overrightarrow{\Omega}_i, \overrightarrow{\alpha})\Big] + N_{\text{evt}} \log\Big[\underbrace{\sum_i^{N_\sum} \frac{d\Gamma}{d\overrightarrow{\Omega}}(\overrightarrow{\Omega}_i, \overrightarrow{\alpha})}_{\text{heavy computations}}\Big] \tag{2.85}$$

where term $N_{\text{evt}} \log\frac{\Omega}{N_\sum}$ has been dropped as it doesn't depend on the model parameters $\overrightarrow{\alpha}$.

2.3.1 . Decay rate factorization

Fit is an iterative process and on each step n it evaluates $\mathcal{L}$ for current set of $\overrightarrow{\alpha}_n$ estimated from prrevious steps. The most computationally expensive part of Eq. 2.85 is a calculation of normalization of the decay rate. In order to cover full kinematical region $N_\sum$ should be large enough and condition $\Delta\overrightarrow{\Omega} = \Delta\overrightarrow{\Omega}_1 = \Delta\overrightarrow{\Omega}_2 = ... = \Delta\overrightarrow{\Omega}_{N_\sum}$ should be satisfied. Suppose time complexity of $\frac{d\Gamma}{d\overrightarrow{\Omega}}$ evaluation is $T_{\text{p.d.f}}$ arbitrary units, then normalization part calculation takes $N_\sum \cdot T_{\text{p.d.f}}$. In order to correctly estimate normalization part one needs to perform on average N_Ω evaluations of the decay rate along each dimension of $\overrightarrow{\Omega}$, which in total yields time complexity of $N_\Omega^5 \cdot T_{\text{p.d.f}}$ arbitrary units, where 5 stands for the size of $\overrightarrow{\Omega}$: $(s_{K\pi_1\pi_2}, s_{K\pi_1}, s_{K\pi_2}, \phi, \cos\theta)$ and $N_\sum = N_\Omega^5$. Assuming that decay rate can be decomposed as in Eq. 2.74, one expectecs the time complexity of normalization part to be $N_{\text{terms}} \cdot T_{\text{term}}$ according to Eq. 2.86.

$$\sum_i^{N_\sum} \frac{d\Gamma}{d\overrightarrow{\Omega}}(\overrightarrow{\Omega}_i, \overrightarrow{\alpha}) = \sum_i^{N_\sum}\sum_j^{N_{\text{terms}}} f_j(\overrightarrow{\Omega}_i) g_j(\overrightarrow{\alpha}) = \sum_j^{N_{\text{terms}}} \tilde{f}_j \cdot g_j(\overrightarrow{\alpha}) \tag{2.86}$$

$$\tilde{f}_j = \sum_i^{N_\sum} f_j(\overrightarrow{\Omega}_i) \tag{2.87}$$

where N_{terms} — number of additive terms within decay rate expression; $\tilde{f}_j$ — normalization integrals of functions depending on kinematics; T_{term} — average time complexity of $g_j(\overrightarrow{\alpha})$ evaluation.

Considering $(N_\Omega = 10)$, $(T_{\text{p.d.f}} \approx T_{\text{term}})$[1] and $(N_{\text{terms}} \approx 1000)$[2] one obtains a ratio of the time complexities as $\frac{N_\Omega^5 \cdot T_{\text{p.d.f}}}{N_{\text{terms}} \cdot T_{\text{term}}} \approx 100$, which leads to reducing of calculation of normalization part in 100

[1] On practice within present work $T_{\text{term}} < T_{\text{p.d.f}}$ since each $g_j(\overrightarrow{\alpha})$ is a product of single variables and at most sin and cos functions of single variables, whereas decay rate according to Eq. 2.73 is overwhelmed with non-linear functions.

[2] The number inferred experimentally within current work using 3rd part library for symbolic computations.

times in the case of decay rate factorization assuming that $\tilde{f}_i$ have been computed before performing the fit.

The formulas for $g_i(\overrightarrow{\alpha})$ and $f_i(\overrightarrow{\Omega})$ are obtained from expanded expression of the decay rate. For instance, within the works [45, 46, 93] the decay rates are defined as in Eqs. 2.88-2.89.

$$\frac{d\Gamma}{d\overrightarrow{\Omega}'} = \sum_{n,m}^{N_{\text{res}}} f_n f_m^* \mathcal{A}_n(\overrightarrow{\Omega}')\mathcal{A}_m^*(\overrightarrow{\Omega}') \tag{2.88}$$

$$\frac{d\Gamma}{d\overrightarrow{\Omega}'} = \frac{1+\lambda}{2}\sum_{n,m}^{N_{\text{res}}} f_n f_m^* \mathcal{A}_n^L(\overrightarrow{\Omega}')\mathcal{A}_m^{L*}(\overrightarrow{\Omega}') + \frac{1-\lambda}{2}\sum_{n,m}^{N_{\text{res}}} f_n f_m^* \mathcal{A}_n^R(\overrightarrow{\Omega}')\mathcal{A}_m^{R*}(\overrightarrow{\Omega}') \tag{2.89}$$

where $\overrightarrow{\Omega}'$ — vector describing the kinematics, and in general $\overrightarrow{\Omega} \neq \overrightarrow{\Omega}'$; N_{res} — number of resonances within the model; $\mathcal{A}_n$ — decay amplitude through n-th resonance; f_n — complex model parameter associated with n-th resonance;

One can analytically infer expressions for $g_i(\overrightarrow{\alpha})$ and $f_i(\overrightarrow{\Omega}')$ as in Eqs. 2.90-2.92 for [45, 93].

$$i = m \cdot N_{\text{res}} + n \tag{2.90}$$

$$g_i(\overrightarrow{\alpha}) = f_n f_m^* \tag{2.91}$$

$$f_i(\overrightarrow{\Omega}') = \mathcal{A}_n(\overrightarrow{\Omega}')\mathcal{A}_m^*(\overrightarrow{\Omega}') \tag{2.92}$$

where first equation represents a mapping from 2D grid ($n, m \in [1; N_{\text{res}}]$) of integer points to an index of i-th term within expanded decay rate expression;

In the case of [46] presence of λ parameter complicates the expressions taking into account right and left helicity states of the final photon (see Eq. 2.89). One notes that grouping within Eq. 2.95 is done analytically by collecting terms near $f_n f_m^*$ and $f_n f_m^* \lambda$.

$$i = h \cdot (m \cdot N_{\text{res}} + n) + (m \cdot N_{\text{res}} + n) \tag{2.93}$$

$$g_i(\overrightarrow{\alpha}) = f_n f_m^* \lambda^h \tag{2.94}$$

$$f_i(\overrightarrow{\Omega}') = \frac{1}{2}\Big[\mathcal{A}_n^L(\overrightarrow{\Omega}')\mathcal{A}_m^{L*}(\overrightarrow{\Omega}') + (-1)^h \mathcal{A}_n^R(\overrightarrow{\Omega}')\mathcal{A}_m^{R*}(\overrightarrow{\Omega}')\Big] \tag{2.95}$$

where first equation represents a mapping from 3D grid ($h \in [0; 1]$, $n, m \in [1; N_{\text{res}}]$) of integer points to an index of i-th term within expanded decay rate expression;

In the present work the derivation of g_i and f_i functions is not straight-forward because of the way how $\theta_{K_1}, \varphi_S^{K^*}, \varphi_S^{\rho}, \varphi_D^{\rho}$ enter the model of $B \to K\pi\pi\gamma$ decay (see Eq. 2.96).

$$\frac{d\Gamma_{L,R}}{d\overrightarrow{\Omega}} = \Big| \underbrace{A_0^{L,R}\sin\theta_{K_1} + A_1^{L,R}\cos\theta_{K_1}}_{K_1^{1270}} + \underbrace{f^{K_1^{1400}} \cdot (A_2^{L,R}\sin\theta_{K_1} + A_3^{L,R}\cos\theta_{K_1})}_{K_1^{1400}}$$
$$+ \underbrace{\sum_j \sum_V f_j^V B_j^{L,R}(\overrightarrow{\Omega})}_{K^{*1410/1680}, K_2^{1430}} \Big|^2 \tag{2.96}$$

$$A_i = A_i^{L,R}(\overrightarrow{\Omega}, \varphi_S^{K^*}, \varphi_S^{\rho}, \varphi_D^{\rho}) = a_{i0}^{L,R}(\overrightarrow{\Omega}) + a_{i1}^{L,R}(\overrightarrow{\Omega}) \cdot e^{i\varphi_S^{K^*}}$$
$$+ a_{i2}^{L,R}(\overrightarrow{\Omega}) \cdot e^{i\varphi_S^{\rho}} + a_{i3}^{L,R}(\overrightarrow{\Omega}) \cdot e^{i(\varphi_S^{\rho} + \varphi_D^{\rho})} \tag{2.97}$$

where $a_{i0}^{L,R}, a_{i1}^{L,R}, a_{i2}^{L,R}, a_{i3}^{L,R}, B_j^{L,R}$ — complex functions depending only on kinematics; $i \in [0;3]$, $j \in (K_1^{*1410}, K_1^{*1680}, K_2^{1430})$, $V \in (K^{*892}, \rho)$;

While computing $\overrightarrow{J}, \overrightarrow{L}, \overrightarrow{K}$ within corresponding sub-modules of "GamPola" a side functionality is established: coefficients depending only on $\overrightarrow{\Omega}$ or its part are cached into multidimensional array $\mathcal{C}$. Forwarding computation of helicity amplitudes from S, D partial waves to $\overrightarrow{J}$, updates the values of the cached array correspondingly. As an example consider the general case of Eqs. 2.52-2.55 represented by Eqs. 2.98-2.101.

$$A_{\mathcal{W}} = \mathcal{W}(\overrightarrow{\Omega}, \overrightarrow{\alpha}) \cdot (\mathcal{U}(\overrightarrow{\Omega}) \sin\theta_{K_1} + \mathcal{V}(\overrightarrow{\Omega}) \cos\theta_{K_1}) \quad (2.98)$$

$$\mathcal{C}_{\mathcal{W}\sin} = \mathcal{C}_{\mathcal{W}} \cdot \mathcal{U}(\overrightarrow{\Omega}) \quad (2.99)$$

$$\mathcal{C}_{\mathcal{W}\cos} = \mathcal{C}_{\mathcal{W}} \cdot \mathcal{V}(\overrightarrow{\Omega}) \quad (2.100)$$

$$\mathcal{C}' = [\mathcal{C}_{\mathcal{W}\sin}; \mathcal{C}_{\mathcal{W}\cos}] \quad (2.101)$$

where $\mathcal{W}$ — intermediate function depending on kinematics and model parameters; $\mathcal{U}, \mathcal{V}$ — coefficients depending only on kinematics (can be numerical constants as well); $\mathcal{C}_{\mathcal{W}}$ — all elements of the cache previously modified while calculating $\mathcal{W}(\overrightarrow{\Omega}, \overrightarrow{\alpha})$; $\mathcal{C}'$ — all elements of the cache after calculation of $A_{\mathcal{W}}$;

Having built left and right helicity amplitudes as inner expression in Eq. 2.73 before applying square operation, the procedure of cache creation is defined according to simple rules similar to the ones on Eqs. 2.98-2.101. At this stage $\mathcal{C}_{ijklmn}$ is a multidimensional array with the axes as following: $i \in (0_{\sin\theta_{K_1}}, 1_{\cos\theta_{K_1}})$, $j \in (0_{C_1}, 1_{C_2})$, $k \in (0_L, 1_R)$, $l \in (0_{K^*}, 1_\rho, 2_\kappa)$, $m \in (0_{K_1^{1270}}, 1_{K_1^{1400}}, 2_{K_1^{*1410}}, 3_{K_1^{*1680}}, 4_{K_2^{1430}})$, $n \in (0_S, 1_D)$. It is sparse array most dimensions of which introduced for caching kinematical coefficients associated with the amplitudes of 1^+ resonances ($K_1^{1270/1400}$). Along dimension l there is a kinematical term associated with an internediate resonance κ: in the present work this resonance do not contribute to the model as is explained previously and its coupling constant manually set to 0, but in order to keep structure of the cache not modified, κ formally is part of the decay rate expression and part of the cache correspondingly. The structure of the cache is an open subject for the further improvements. Multidimensional array $\mathcal{C}_{ijklmn}$ is flattened into $\mathcal{C}_I^{1D}$, where mappiing $(i,j,k,l,m,n) \to I$ is applied and I — index of kinematical coeffcient within helicity amplitude.

The coefficients $\mathcal{C}_I^{1D}$ are related to $f_i(\overrightarrow{\Omega})$, but functional view of $g_i(\overrightarrow{\alpha})$ is still unknown. In order to address the latter, symbolic computations are used within "GiNAC" library [110]. Applying its functionality allows to define symbolic cache ${}_s\mathcal{C}_I^{1D}$ having exactly the same structure as numerical $\mathcal{C}_I^{1D}$ one and set of symbolic model parameters $\overrightarrow{\alpha}^s$. Then symbolic decay rate is build out of ${}_s\mathcal{C}_I^{1D}$ and $\overrightarrow{\alpha}^s$ as a computational graph. The nodes are either elements of ${}_s\mathcal{C}_I^{1D}$ or simple functions of elements of $\overrightarrow{\alpha}^s$. The vetrices are represented by addition and multiplication operations. "GiNAC" is able to infer expanded symbolic expression from computational graph as in Eq. 2.102.

$$DR = \sum_i^{N_{\text{terms}}} f_i({}_s\mathcal{C}_I^{1D}) \cdot g_i(\overrightarrow{\alpha}^s) \quad (2.102)$$

where DR — symbolic decay rate;

In the symbolic representation $f_i({}_s\mathcal{C}_I^{1D})$ and $g_i(\overrightarrow{\alpha}^s)$ can be separated. Substituting numerical values for the arguments $\mathcal{C}_I^{1D}$ (or $\overrightarrow{\alpha}$) into Eq. 2.102 allows to collect symbolic expressions near $g_i(\overrightarrow{\alpha}^s)$ (or $f_i({}_s\mathcal{C}_I^{1D})$).

2.3.2 . Normalization and fitting

As it is explained in the Subsection 2.3.1, $\tilde{f}_i$ should be known in advance in order to perform the fit efficiently. Thus, two steps are involved in the procedure:

- Caching — stage, where normalization integrals $\tilde{f}_i$ are calculated and correspondings of these integrals and couplings $g_i(\overrightarrow{\alpha})$ are created for each term among N_{terms} possible;
- Fitting — using cached normalization integrals, perform $\mathcal{L}$ minimization (gradient descent fully relies on the ROOT [129] minimizer *Minuit* [111]).

For the demonstration purpose, the scenario as in Fig. 2.2 can be implemented by generating a sample of 50000 $B^+ \to K^+\pi^+\pi^-\gamma$ events with "GamPola"-generator followed by the fit using "GamPola"-fitter. It allows performing a sanity check of the tool. As a result, the following distributions of generated and fit projections of kinematical variables are obtained (see Figs. 2.19-2.23). The Table. 2.4 summarizes the comparison between truth ($\overrightarrow{\alpha}$) and estimated ($\overrightarrow{\alpha}^*$) values. Fit distributions are obtained by re-generating a sample of 50000 events using $\overrightarrow{\alpha}^*$. The distributions are consistent and parameter values are equal within the statistical uncertainty of the fit. Truth model parameters $\overrightarrow{\alpha}$ used within sanity check are obtained from the binned χ^2 fit of the LHCb histograms of real data as will be explained further. Photon polarization parameter value is assigned to an artificially low value, with the purpose to simulate a new physics effect, although according to the results of [48] such value looks unrealistic.

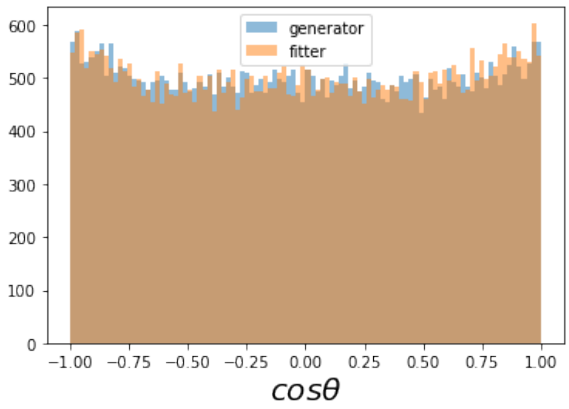


Figure 2.19. $\cos\theta$ distributions comparison

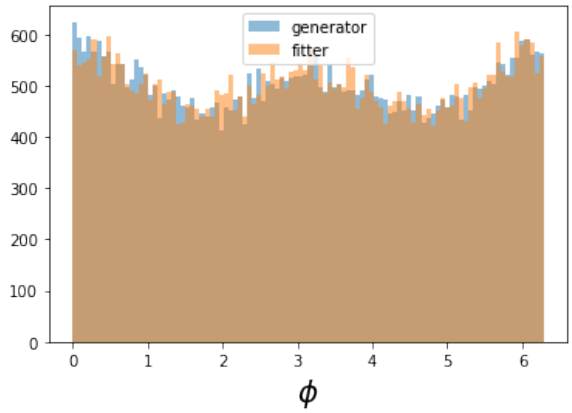


Figure 2.20. ϕ distributions comparison

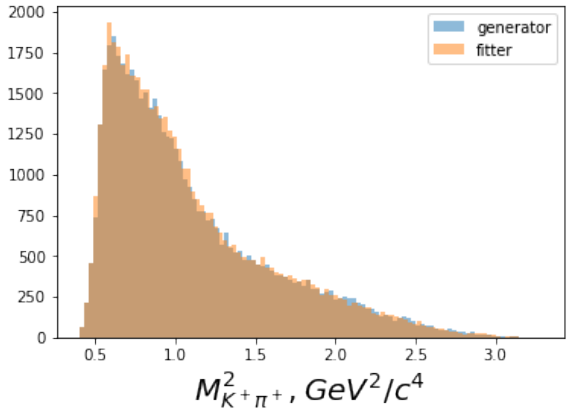


Figure 2.21. $s_{K\pi_1}$ distributions comparison

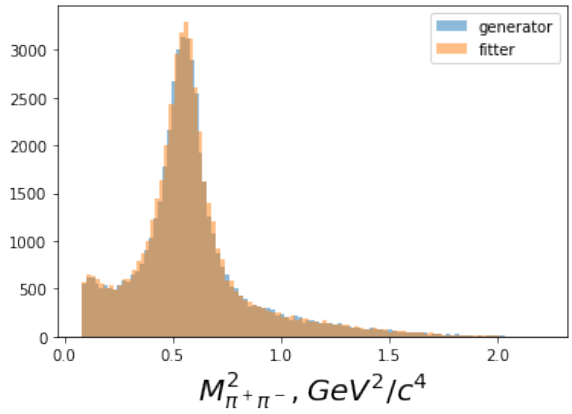


Figure 2.22. $s_{\pi_1\pi_2}$ distributions comparison

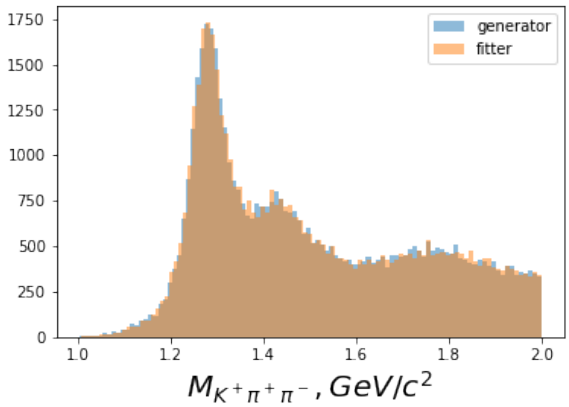


Figure 2.23. $M_{K\pi_1\pi_2}$ distributions comparison

Label	α	α^*	σ_{α^*}
θ_{K_1}	1.024	1.020	0.005
$\varphi^D_{K^*}$	-0.13	-0.11	0.02
φ^S_ρ	0.0	-0.01	0.02
φ^D_ρ	0.1	0.07	0.06
λ	0.65	0.63	0.01
$f^{K_1^{1400}}$	$-0.31 + i \cdot -0.38$	$-0.31 + i \cdot -0.38$	$0.005 + i \cdot 0.006$
$f^{K^*}_{K^{*1410}}$	$0.19 + i \cdot 0.28$	$0.177 + i \cdot 0.282$	$0.014 + i \cdot 0.014$
$f^{\rho}_{K^{*1410}}$	$-1.22 + i \cdot 1.87$	$-1.25 + i \cdot 1.83$	$0.04 + i \cdot 0.03$
$f^{K^*}_{K^{*1680}}$	$0.13 + i \cdot 0.55$	$0.14 + i \cdot 0.55$	$0.02 + i \cdot 0.1$
$f^{\rho}_{K^{*1680}}$	$-0.39 + i \cdot 0.02$	$-0.37 + i \cdot 0.04$	$0.02 + i \cdot 0.03$
$f^{K^*}_{K_2^{1430}}$	$0.24 + i \cdot 0.031$	$0.243 + i \cdot 0.038$	$0.006 + i \cdot 0.007$
$f^{\rho}_{K_2^{1430}}$	$-0.38 + i \cdot -0.66$	$-0.37 + i \cdot -0.65$	$0.02 + i \cdot 0.01$
$f^{K_1^{1270}}$	1	NA	NA
f^{1410}	$5 + i \cdot 0$	NA	NA
f^{1680}	$5 + i \cdot 0$	NA	NA
f^{1430}	$5 + i \cdot 0$	NA	NA

Table 2.4. Comparison of generator and fit results for the baseline model.

where NA — not applicable; constants in the bottom of the table are multiplied on the $\overrightarrow{J}^{1270}$, $\overrightarrow{L}^{1410/1680}$, $\overrightarrow{K}^{1430}$ to additionally weight helicity amplitudes.

2.4 . Sensitivity study

Having introduced a model of $B^+ \to K^+\pi^+\pi^-\gamma$ decays, generator, and fitter tools, the photon polarization sensitivity study is performed within this section. The model parameter values are determined and the Baseline model is defined. The sensitivity studies are done using the Baseline model.

2.4.1 . The Baseline model

The Baseline model is defined for the Monte-Carlo studies and relies on the binned χ^2 fit of the input histograms of $B^+ \to K^+\pi^+\pi^-\gamma$ decay. The model parameters $\overrightarrow{\alpha}$ are chosen to roughly reproduce the LHCb histograms of $M_{K\pi\pi} = \sqrt{s_{K^+\pi^+\pi^-}}$ (100 bins), $M_{\pi\pi} = \sqrt{s_{\pi^+\pi^-}}$ (44 bins), $M_{K\pi} = \sqrt{s_{K^+\pi^-}}$ (37 bins), $\cos\theta_\gamma = \cos\theta$ (20 bins) extracted from [94], where $\cos\theta_\gamma$ is taken for the following bins of $M_{K\pi\pi}$: $[1.1; 1.3]$ GeV/c^2, $[1.3; 1.4]$ GeV/c^2, $[1.4; 1.6]$ GeV/c^2. For this purpose $N^{evt} = 10^6$ of events are produced using "GamPola"-generator with arbitrary initial model parameter values $\overrightarrow{\alpha}_0$. For each event i, $\overrightarrow{f}(\overrightarrow{\Omega}_i)/DR(\overrightarrow{\Omega}_i, \overrightarrow{\alpha}_0)$ is evaluated and submitted to the bin j according to generated value of Ω_{ik}, where k — index of kinematical variable within $\overrightarrow{\Omega}$ $(s_{K\pi\pi}, s_{K\pi}, s_{\pi\pi}, \cos\theta_{K_1}, \phi)$. Histogramization

of 1-dimensional data as in Eqs. 2.103-2.104 compresses number of data points used by the fit.

$$\overrightarrow{r}_{jk} = \sum_{i}^{N_{jk}^{evt}} \frac{\overrightarrow{f}(\overrightarrow{\Omega}_i)}{DR(\overrightarrow{\Omega}_i, \overrightarrow{\alpha}_0)} \tag{2.103}$$

$$N^{evt} = \sum_{j}^{N_k^{bins}} N_{jk}^{evt} \tag{2.104}$$

where N_{jk}^{evt} — number of events in the bin j of the k-th histogram; $\overrightarrow{f}(\overrightarrow{\Omega}_i)$ — vector of values of functions defined by Eq. 2.74 for i-th event; $DR = \frac{d\Gamma}{d\overrightarrow{\Omega}}$ — decay rate defined by Eq. 2.72; N_k^{bins} — number of bins in the k-th histogram; N^{evt} — total number of generated events.

Minimization function is defined as in Eq. 2.105 and represented by the sum of χ^2 values of several histograms. As can be seen from Eqs. 2.105-2.108, $\overrightarrow{\alpha}$ is extracted based on weights of the bins of LHCb histograms.

$$\chi^2 = \sum_{k}^{N_{\text{hist}}} N_{k,\text{LHCb}}^2 \sum_{j}^{N_k^{bins}} \frac{\left(w_{jk}^{\text{LHCb}} - w_{jk}^{\text{Fit}}\right)^2}{\sigma_{jk}^2} \tag{2.105}$$

$$w_{jk}^{\text{LHCb}} = \frac{H_{jk}^{\text{LHCb}}}{\sum_{j'}^{N_k^{bins}} H_{jk}^{\text{LHCb}}} \tag{2.106}$$

$$w_{jk}^{Fit} = \frac{(\overrightarrow{r}_{jk} \cdot \overrightarrow{g}(\overrightarrow{\alpha}))}{\sum_{j'}^{N_k^{bins}} (\overrightarrow{r}_{j'k} \cdot \overrightarrow{g}(\overrightarrow{\alpha}))} \tag{2.107}$$

$$\sigma_{jk}^2 = \sqrt{\left[\frac{H_{jk}^{\text{LHCb}}}{\sum_{j'}^{N_k^{bins}} (\overrightarrow{r}_{j'k} \cdot \overrightarrow{g}(\overrightarrow{\alpha}))}\right]^2 (\overrightarrow{r}_{jk} \cdot \overrightarrow{g}(\overrightarrow{\alpha})) + \sigma_{\text{LHCb}}^2} \tag{2.108}$$

where k — index of histogram; j — index of bin within histogram; N_{hist} — number of histograms; $N_{k,\text{LHCb}}$ — number of entries in the k-th LHCb histogram; H_{jk}^{LHCb} — height of j-th bin of k-th LHCb histogram; $w_{jk}^{\text{LHCb}}, w_{jk}^{Fit}$ — weights of LHCb k-th histogram and the fit for j-th bin; $g(\overrightarrow{\alpha})$ — vector of functions defined by Eq. 2.74 and depending on the model parameter values; σ_{LHCb} — statistical uncertainties of the heights of the bins for LHCb histograms;

The results of the fit are given on the Figs. 2.24-2.30. While for the first two bins of $M_{K\pi\pi} \in [1.1; 1.4]$ GeV/c^2 (see Figs. 2.24-2.25) the argeement is reasonable, it starts to degrade when going towards higher resonances, where $M_{K\pi\pi} \in [1.4; 1.6]$ GeV/c^2 (Fig. 2.26) and gives a huge discrepancy for $M_{K\pi\pi} \in [1.6; 1.9]$ GeV/c^2 (see Fig. 2.27).

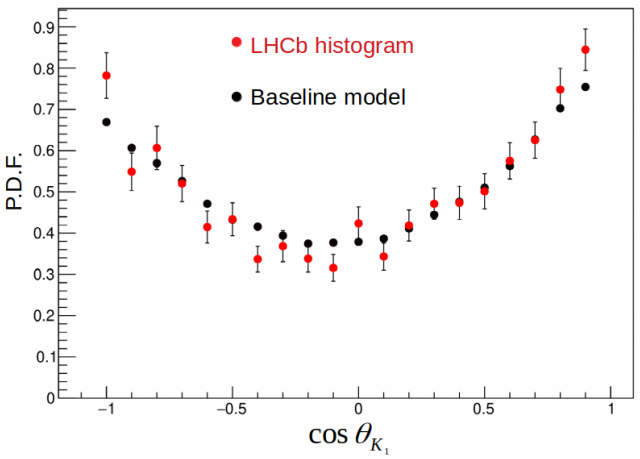


Figure 2.24. $\cos\theta$ distributions comparison for the LHCb data and the Baseline model for $M_{K\pi\pi} \in [1.1; 1.3]$ GeV/c^2

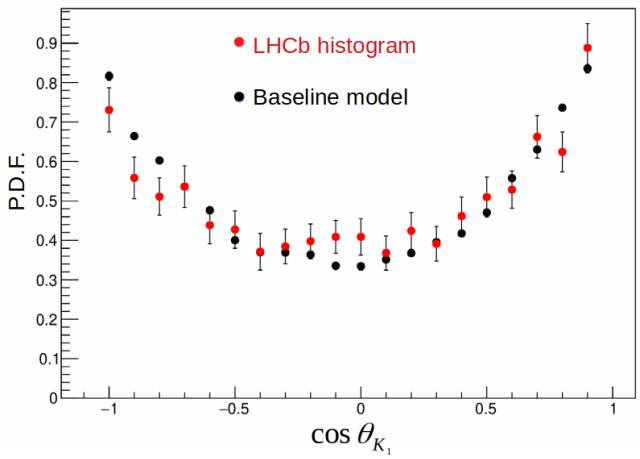


Figure 2.25. $\cos\theta$ distributions comparison for the LHCb data and the Baseline model for $M_{K\pi\pi} \in [1.3; 1.4]$ GeV/c^2

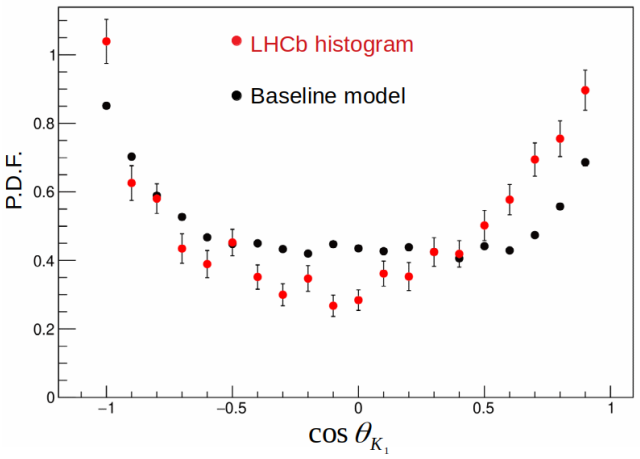


Figure 2.26. $\cos\theta$ distributions comparison for the LHCb data and the Baseline model for $M_{K\pi\pi} \in [1.4; 1.6]$ GeV/c^2

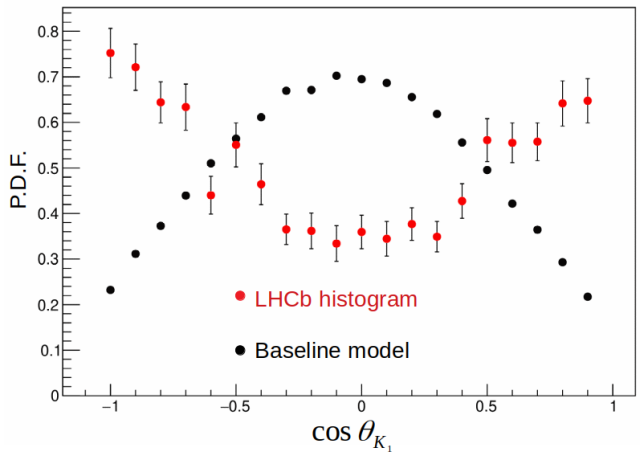


Figure 2.27. $\cos\theta$ distributions comparison for the LHCb data and the Baseline model for $M_{K\pi\pi} \in [1.6; 1.9]$ GeV/c^2

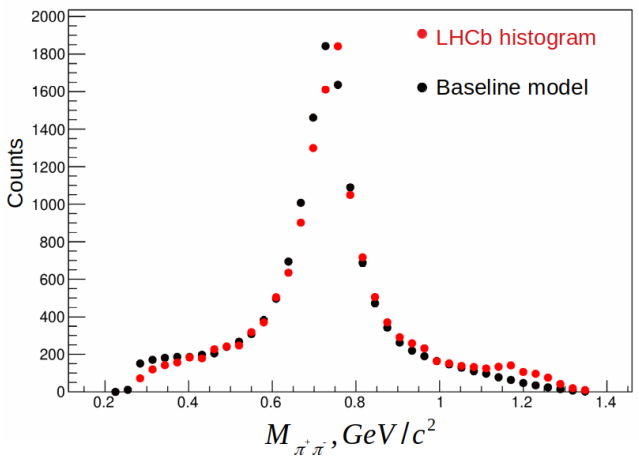


Figure 2.28. $M_{\pi\pi}$ distributions comparison for the LHCb data and the Baseline model

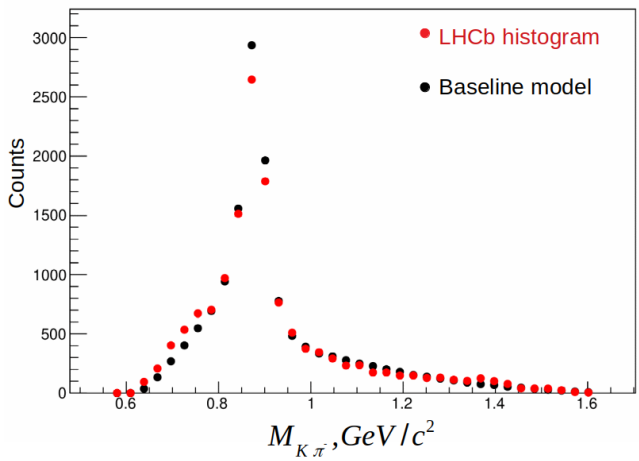


Figure 2.29. $M_{\pi\pi}$ distributions comparison for the LHCb data and the Baseline model

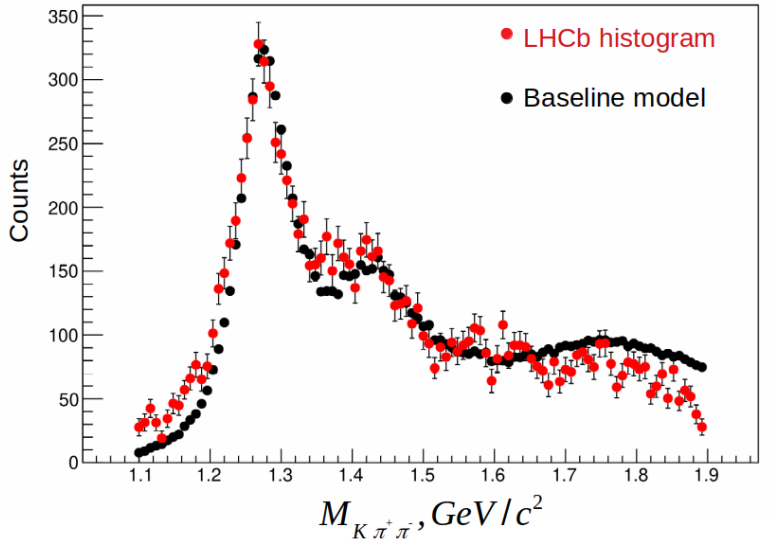


Figure 2.30. $M_{K\pi\pi}$ distributions comparison for the LHCb data and the Baseline model

Within the binned χ^2 fit, $\cos\theta_\gamma$ distribution for $M_{K\pi\pi} \in [1.6; 1.9]$ GeV/c^2 is not considered, because "GamPola" does not describe all main resonances beyond $M_{K\pi\pi} > 1.6$ GeV/c^2. In particular, decay of K_2^{1770} resonance is not included in the model. Hence, interference in the higher range of $M_{K\pi\pi}$ spectrum is underestimated. However, the full region of $M_{K\pi\pi} \in [1.1; 1.9$ GeV$]$ is considered for the fit.

The Baseline model parameters are defined in Table 2.5. Obtained values do not represent realistic model parameter values for $B \to K\pi\pi\gamma$ decay, but values that are able to simulate a sample of events appearing visually similar to the real data for several kinematical variables. In addition, $\lambda_\gamma = 0.7 < 1$ value is taken to simulate the New Physics effect.

Label	α
θ_{K_1}	1.042
$\varphi^D_{K^*}$	0
φ^S_ρ	0
φ^D_ρ	0
λ	0.7
$f^{K_1^{1400}}$	$-0.34 - i \cdot 0.41$
$f^{K^*}_{K^{*1410}}$	$0.2 + i \cdot 0.26$
$f^{\rho}_{K^{*1410}}$	$-1.2 + i \cdot 1.95$
$f^{K^*}_{K^{*1680}}$	$0.01 + i \cdot 0.64$
$f^{\rho}_{K^{*1680}}$	$-0.33 + i \cdot 0.08$
$f^{K^*}_{K_2^{1430}}$	$0.22 + i \cdot 0.03$
$f^{\rho}_{K_2^{1430}}$	$-0.42 - i \cdot 0.72$
$f^{K_1^{1270}}$	1
f^{1410}	$5 + i \cdot 0$
f^{1680}	$5 + i \cdot 0$
f^{1430}	$5 + i \cdot 0$

Table 2.5. The Baseline model parameters

2.4.2 . Effect of cut on the $M_{K\pi\pi}$ spectrum

As it is explained in the Subsection 2.4.1, $\cos\theta_\gamma$ histogram for $M_{K\pi\pi} \in [1.6; 1.9]$ GeV/c^2 is not used for definition of the Baseline model due to contribution of kaonic resonances in this region which are not included in "GamPola". The purpose of this study is to analyze the photon polarization sensitivity while excluding this region and, in general, sequentially putting upper cut on $M_{K\pi\pi}$ spectrum.

The generator produces 5000 events according to the Baseline model. Applying upper cut on invariant mass of $K\pi\pi$ from 1.3 GeV/c^2 to 2 GeV/c^2 with step 0.04 GeV/c^2 yields reduced samples, which are further fitted using likelihood approach within "GamPola". Statistical uncertainties σ_{λ_γ} of photon polarization parameter λ_γ are obtained from the fit.

The plots on the Figs. 2.31-2.32 include two effects at once: increasing the statistical uncertainty due to reducing of a number of events after the cuts and decreasing complexity of the data by cutting out kaonic resonances in the higher range of $M_{K\pi\pi}$. Obtained results demonstrate, that σ_{λ_γ} almost doesn't decrease for $M_{K\pi\pi} > 1.6$ GeV/c^2 and this region can be safely excluded. Huge uncertainties for $M_{K\pi\pi} < 1.4$ GeV/c^2 correspond to the case where model greatly overfits the data.

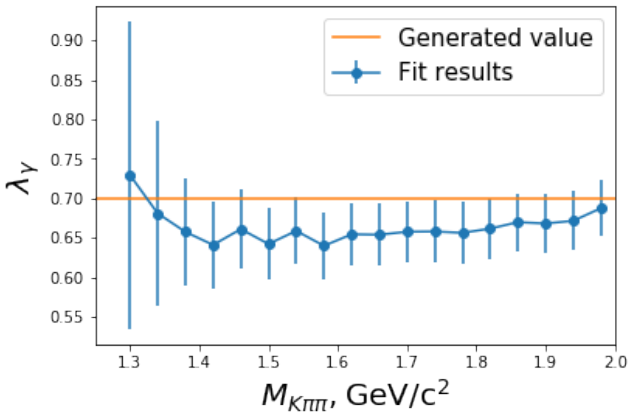


Figure 2.31. Photon polarization dependency on $M_{K\pi\pi}$ cut value

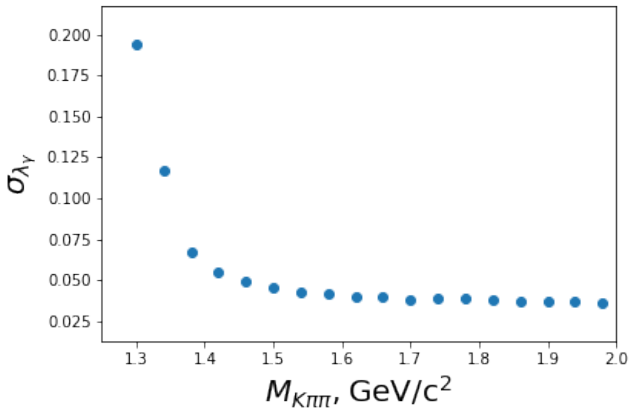


Figure 2.32. Photon polarization uncertainty as a function of $M_{K\pi\pi}$ cut.

2.4.3 . Photon polarization sensitivity study

In this subsection comparison the sensitivities of the photon polarization parameter with respect to the decay channel ($B^+ \to K^+\pi^+\pi^-\gamma$ and $B^0 \to K^+\pi^0\pi^-\gamma$) and the Baseline model choice (the one defined above or another alternative without imaginary parts) is described.

Events generation with two types of models is performed:

- Model1: all model parameters are complex (the Baseline model);
- Model2: all model parameters are real;

Each generated set contains 5000 events, corresponding to the 680 fb^{-1} of data being collected at Belle II after all efficiency selections. Having applied the cut $M_{K\pi\pi} < 1.6$ GeV/c^2, several fits have been made using the Baseline model. Statistical uncertainties σ_{λ_γ} of λ_γ are placed to the Table 2.6.

	$\sigma_{\lambda_\gamma}(B^+ \to K^+\pi^+\pi^-\gamma)$	$\sigma_{\lambda_\gamma}(B^0 \to K^+\pi^0\pi^-\gamma)$
Model1 (baseline model)	0.0399	0.0405
Model2 (all real)	0.045	0.030

Table 2.6. Comparison of λ_γ for different generation models and different decay channels. Fit is done using baseline model.

While for the Model1 σ_λ is $\sigma_\lambda(B^+ \to K^+\pi^+\pi^-\gamma) \simeq \sigma_\lambda(B^0 \to K^+\pi^0\pi^-\gamma)$, for Model2 $\sigma_\lambda(B^+ \to K^+\pi^+\pi^-\gamma) > \sigma_\lambda(B^0 \to K^+\pi^0\pi^-\gamma)$. It is expected that due to the additional K^{*0}_{892} subresonance, in the neutral channel the uncertainty (sensitivity) is lower (higher), than in the case of charged channel. This is the case for the Model2, but not for the Baseline model. Though the sensitivity is highly depends on the model parameters values as can be seen from Table. 2.6.

2.4.4 . Stability of the fit: mirror and fake solutions

The presence of mixing angle θ_{K_1} and hadronic phases affects the stability of the fit, since they enter into decay rate as the arguments of $\sin$ and $\cos$ functions, causing additional entangling between model parameters compared to Eqs. 2.88-2.89.

In order to probe the stability using one sample of events[3], a lot of fits need to be done with random seeds. The measure of estimation of the best fit results in this case is $\Delta\chi^2_i = \mathcal{L}_i - \min(\mathcal{L}_i)$, where $\mathcal{L}_i$ — the likelihood of the i^{th} fit defined by Eq. 2.82, $\min$ — minimum fitted $\mathcal{L}_i$ outcome among all minimization trials.

While repeating the fit of 5000 events 250 times with random seed, 3 types of solutions are observed:

- Truth: $\Delta\chi^2 < \Delta\chi^2_{thrs}$ and $\overrightarrow{\alpha}_{gen} \approx \overrightarrow{\alpha}_{fit}$; ($\Delta\chi^2_{thrs} = 16$ is considered)
- Fake: $\Delta\chi^2 \gg \Delta\chi^2_{thrs}$ and $\overrightarrow{\alpha}_{gen} \neq \overrightarrow{\alpha}_{fit}$;
- Mirror: $\Delta\chi^2 < \Delta\chi^2_{thrs}$ and $\overrightarrow{\alpha}_{gen} \neq \overrightarrow{\alpha}_{fit}$ (majority of α_i are different);

Fake solutions correspond to high values of $\Delta\chi^2$ (see Fig. 2.33) and are suppressed by applying a tight cut on this variable. Truth and mirror solutions both correspond to the low $\Delta\chi^2$ values. The latter ones as shown on the Fig. 2.34 appear from invariance of the decay rate of Eq. 2.96 with respect to the model parameters transformation. One observes that hadronic phases are transformed as $\varphi_S^{K^*} \to \varphi_S^{K^*} + 2\pi, \varphi_S^{\rho} \to \varphi_S^{\rho} + 2\pi, \varphi_D^{\rho} \to \varphi_D^{\rho} + 2\pi$. Excluding hadronic phases, $\frac{d\Gamma_{L,R}}{d\overrightarrow{\Omega}}$ contains terms: $f^{1400}\sin\theta_{K_1}\cos\theta_{K_1}$, $\tilde{g}_i \sin\theta_{K_1}$, $\tilde{g}_i \cos\theta_{K_1}$, $\tilde{g}_i f^{1400}\sin\theta_{K_1}$, λ... from which follow simultaneous transformations as in Eqs. 2.109-2.114. The results obtained analytically are numerically confirmed on the Figs. 2.35-2.37. In particular, the important conclusion follows from Eq. 2.114 and Fig. 2.37, where photon polarization parameter is not affected by the mirror solutions.

$$\theta_{K_1} \to \theta_{K_1} + \pi \tag{2.109}$$
$$f^{K^*/\rho}_{K_2^{1430}} \to -f^{K^*/\rho}_{K_2^{1430}} \tag{2.110}$$
$$f^{K^*/\rho}_{K^{*1410}} \to -f^{K^*/\rho}_{K^{*1410}} \tag{2.111}$$
$$f^{K^*/\rho}_{K^{*1680}} \to -f^{K^*/\rho}_{K^{*1680}} \tag{2.112}$$
$$f^{K_1^{1400}} \to f^{K_1^{1400}} \tag{2.113}$$
$$\lambda \to \lambda \tag{2.114}$$

[3]One expects that only one sample of events will be available in the final stage of analysis.

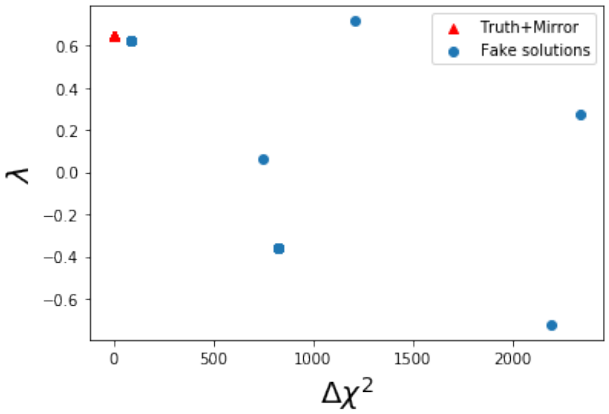


Figure 2.33. Several clusters of solutions are present on the $\lambda - \Delta\chi^2$ plane.

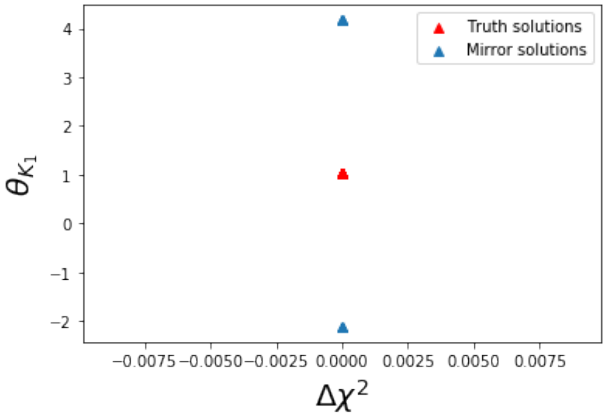


Figure 2.34. For the same value of $\Delta\chi^2$ three different values of θ_{K_1} are observed.

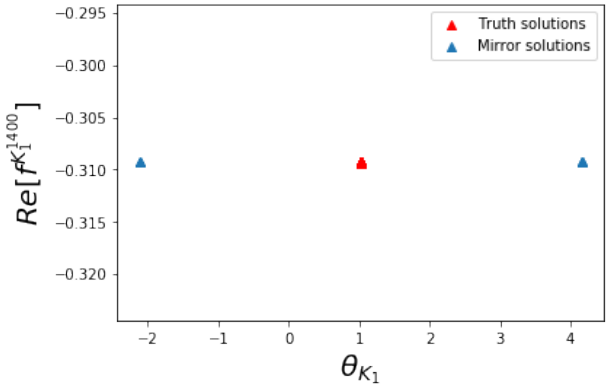


Figure 2.35. Illustration of simultaneous transformation according to Eqs. 2.109, 2.113: mixing angle change doesn't affect coupling of K_1^{1400}.

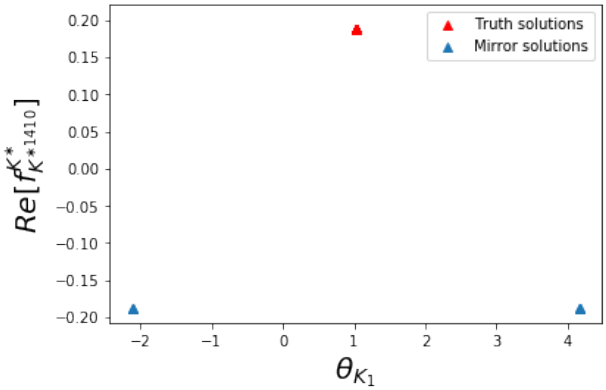


Figure 2.36. Illustration of simultaneous transformation according to Eqs. 2.109, 2.111: both mixing angle and coupling K^{*1410} are discretely changed.

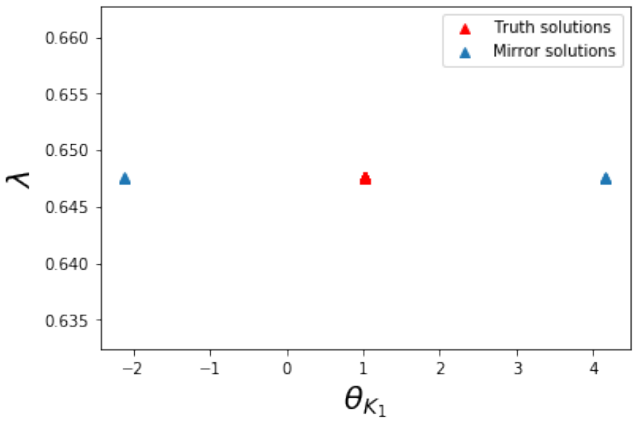


Figure 2.37. Simultaneous transformations according to Eqs. 2.109, 2.114: when tight cut on χ^2 is applied, only mirror solutions are left and λ_γ is not affected by them.

Numerically, the fit stability can be estimated by the relative fraction of each outcome among 250 trials and is summarized below:

- Truth solutions: 45 %;
- Mirror solutions: 5 %;
- Fake solutions: 50 %;

One concludes that the probability to obtain truth or mirror solution is 50 %.

2.4.5 . Competitivity of the proposed approach for the photon polarization measurement

The present work within Belle II Experiment aims measuring photon polarization parameter defined by Eq. 1.51 in the similar way as in [46]. On the Fig. 1.3 it is equivant to putting a circle $[\mathrm{Re}(\mathrm{C}_7'/\mathrm{C}_7)]^2 + [\mathrm{Im}(\mathrm{C}_7'/\mathrm{C}_7)]^2 = \frac{1-\lambda}{1+\lambda}$. The radius of this circle $\sqrt{\frac{1-\lambda}{1+\lambda}}$ and its half-width can be expressed on the Figs. 2.38-2.39 as a functions of λ and integrated luminosity within Belle II. While smaller the radius, higher polarization, smaller ΔR implies higher confidence in the obtained λ_γ value. The Fig. 2.39 is obtained using calibration: $\sigma[\lambda(\mathcal{L}_0 = 680\ \mathrm{fb}^{-1})] = \sigma[\lambda(\varepsilon \cdot \mathrm{N}^{\mathrm{evt}}_{680\ \mathrm{fb}^{-1}})] = 0.04$ and assuming that all selection efficiencies ε of the analysis were applied. One can observe that $\Delta R_{\lambda_\gamma}$ is smallest for $\lambda_\gamma < 0.9$ and it can be understood from Eq. 2.116, where $\lambda_\gamma = 1$ corresponds to the pole of $\Delta R_{\lambda_\gamma}$ function.

$$R = \sqrt{\frac{1-\lambda_\gamma}{1+\lambda_\gamma}} \tag{2.115}$$

$$\Delta R = \frac{1}{R} \cdot \frac{1}{(1+\lambda_\gamma)^2} \cdot \frac{\sigma[\lambda^0_\gamma]}{\sqrt{\frac{\mathcal{L}}{\mathcal{L}_0}}} = \frac{\sigma[\lambda^0_\gamma]\sqrt{\mathcal{L}_0}}{\sqrt{\mathcal{L}}} \cdot \frac{1}{\sqrt{1-\lambda_\gamma}(1+\lambda_\gamma)^{\frac{3}{2}}} \tag{2.116}$$

The area surrounded by red disk on the Fig. 1.3 can be approximated by a circle with radius $\approx 0.15-0.2$. The smaller square of this area, higher precision of measurements of the photon polarization. At 2σ level

within Belle II, the area constraining possible λ_γ values on the $C_7 - C_7'$ plane is defined by $S_{Belle\ II} = 8\pi R_{\lambda_\gamma} \Delta R_{\lambda_\gamma}$ and represented by a double circle with inner and outer radii equal to $R_{\lambda_\gamma} - \Delta R_{\lambda_\gamma}$ and $R_{\lambda_\gamma} + \Delta R_{\lambda_\gamma}$ correspondingly. Hence, in order to compete with the LHCb results [48], inequality as in Eq. 2.117 should be satisfied. From the Fig. 2.40 follows that having $> 3\ \mathrm{ab}^{-1}$ the sensitivity (area) to the λ_γ in the $B \to K\pi\pi\gamma$ will be similar to the one in the $B^0 \to K^{*0} e^+ e^-$.

$$8\pi R_{\lambda_\gamma} \Delta R_{\lambda_\gamma} \leq \pi \cdot 0.15^2 \approx 0.1 \tag{2.117}$$

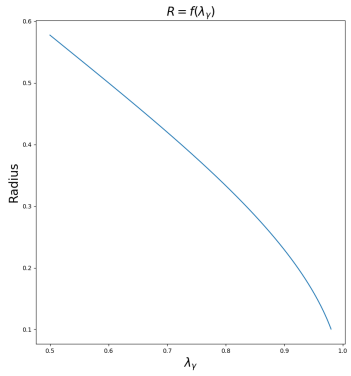


Figure 2.38. Radius of the circle on the $C_7 - C_7'$ plane as a function of the photon polarization parameter

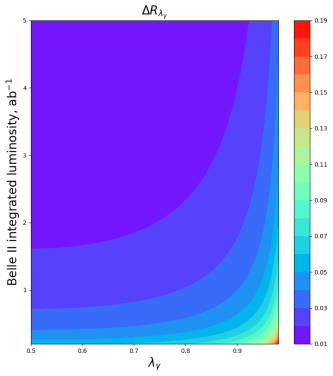


Figure 2.39. Statistical uncertainty of the circle radius as a function of λ_γ and integrated Belle II luminosity.

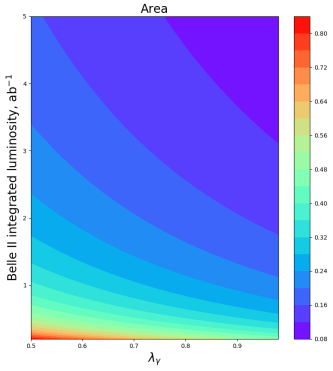


Figure 2.40. Square of potential Belle II region in the $C_7 - C_7'$ plane as a function of λ_γ and integrated Belle II luminosity.

3 - The Belle II Experiment

Belle II Experiment is a next-generation particle physics experiment aiming to search for New Physics in the flavour sector and improve existing measurements within Standard Model. It is designed for colliding electrons and positrons at the $\Upsilon(4S)$ center of mass, which is located just above a threshold of $B\bar{B}$ pair creation. SuperKEKB collider was designed with asymmetric beam energies to boost to a center of the mass system and allow for time-dependent charge-parity symmetry violation measurements. The boost is slightly less than in the case of KEKB, but it is preferable for the analyses with neutrino in the final state, where detector hermeticity is required. On the other side, with such changes, better resolution of vertex reconstruction is needed. SuperKEKB has a design luminosity of $8 \times 10^{35}\mathrm{cm}^{-2}\mathrm{s}^{-1}$, which is nearly 40 times larger than the recorded peak of its predecessor KEKB and 80 times KEKB's design luminosity. The amount of data collected at the time of writing the thesis corresponds to $213\ \mathrm{fb}^{-1}$ integrated luminosity.

The sensitivity of Belle II to NP depends on the values of the corresponding couplings of the NP. The mass of the new physics particle can be up to $O(100\ \mathrm{TeV})$ if the corresponding couplings are not as suppressed as in the SM. Belle II and SuperKEKB explore the intensity frontier of NP searches by moving beyond a simple observation of an NP effect to its detailed studies through constraining measurements in several flavour physics processes. This section contains a brief description of the Belle II Experiment, extensively covered in [117].

3.1 . Overview of SuperKEKB

The target luminosity of SuperKEKB is 40 times greater than the recorded peak of KEKB. It requires a large upgrade of the accelerator complex [137]. The key factor of luminosity increasing is the implementation of the "nano-beam" scheme proposed in [138]. The main modifications are 2 times higher currents and 20 times smaller beam size ($1\ \mu\mathrm{m} \to 50\ \mathrm{nm}$) at the collision point compared to the KEKB values. The two beams also collide at a larger angle of $83\ \mathrm{mrad}$ (compared to $22\ \mathrm{mrad}$ in KEKB). Collision energies have been changed: $8\ \mathrm{GeV} \to 7\ \mathrm{GeV}$ for electron and $3.5\ \mathrm{GeV} \to 4\ \mathrm{GeV}$ for positron beam. A lower beam energy asymmetry was decided to reduce the beam losses due to the Touschek effect. In addition, it leads to slight improvements in solid angle acceptance for the decays involving missing particles. On the other side, the spatial separation between B-mesons is reduced, which concerns the analysis programs targeting time-dependent CP violation measurements.

The modifications to the accelerator complex cover a new electron-injection gun, a new the target for positron production, and a new additional damping ring for the positron beam [117].

The range of beam energies, which SuperKEKB is able to produce covers the $\Upsilon(1S) - \Upsilon(6S)$ resonance states. The maximum achievable center of mass energy of $11.24\ \mathrm{GeV}$ at SuperKEKB is restricted by the power of the injector linac.

It is worth mentioning that in June 2021 SuperKEKB achieved a new luminosity world record, $3.1 \times 10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$.

3.2 . Belle II Detector

The facility used for capturing products of e^+e^- collisions is the Belle II detector (see Fig. 3.1). It has the same spatial size as its predecessor Belle detector with length 7.5 m and height 7 m. The components of the new detector are either completely new or significantly upgraded [139].

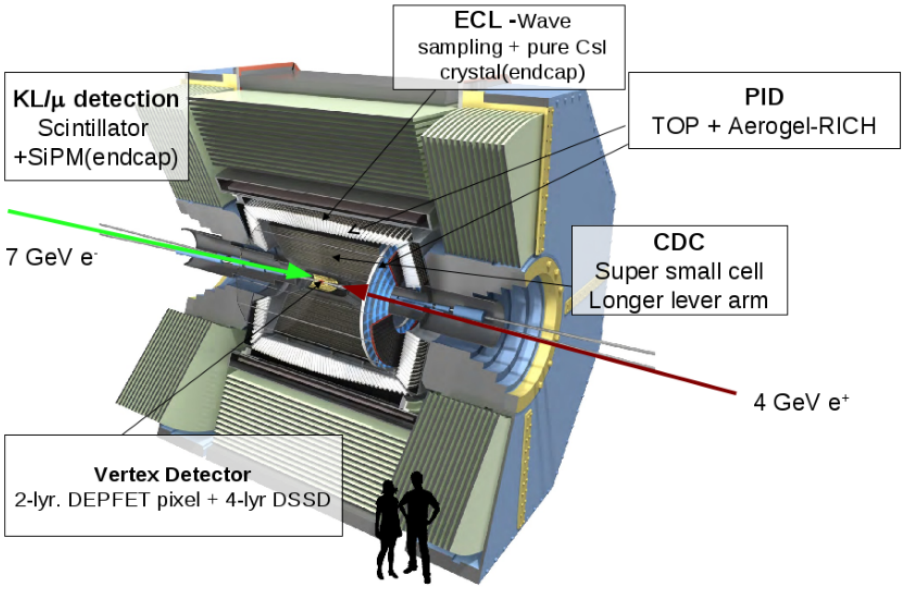


Figure 3.1. Belle II detector.

The advantage of the Belle II detector, that it is able to operate at ≈ 40 times higher event rates, due to increasing luminosity. On the other side, the backgrounds rates are enhanced in 10-20 times [139]. The descent performance of the spectrometer is achieved after strongly suppressing the effects of an increase in radiation damage, occupancy, fake hits, pile-up noise, and neutron-induced hits in the muon detection system. In addition, the trigger scheme, data acquisition system, and computing have been modified due to higher event rates. In particular, the trigger and data acquisition systems have been adapted to support broader low-multiplicity physics analysis programs. Hadron identification and hermeticity have been improved as well.

The main high-level properties of the Belle II detector can be summarized as follows:

- Vertex resolution of ≈ 50 μm;
- Very high reconstruction efficiencies for charged particles with momenta lower than a few hundred $\mathrm{MeV/c}$;
- Very good momentum resolution over the whole kinematic range of the experiment, i.e. up to ≈ 8 $\mathrm{GeV/c}$;
- Precise measurements of photon energy and direction from a few tens of MeV to ≈ 8 GeV, and efficient detection from 30 MeV onwards;
- Highly efficient particle identification system to separate pions, kaons, protons, electrons, and muons over the full kinematic range of the experiment;

- Cover the almost full solid angle;
- Fast and efficient trigger system, as well as a data acquisition system capable of storing a large amounts of data.

3.2.1 . Vertex detector (VXD)

The new vertex detector consists of two devices, the silicon Pixel Detector (PXD) and Silicon Vertex Detector (SVD), with six layers (Fig. 3.2) around Be beam pipe of 10 mm radius. The first two layers at r $= 14$ mm and $r = 22$ mm use pixelated sensors of the DEPFET type [118]. The remaining four layers at radii of 38 mm, 80 mm, 115 mm, and 140 mm contain double-sided silicon strip sensors. In general, compared to the Belle vertex detector, the beam pipe, and the first two detector layers are closer to the interaction point, while the outermost layer is at a considerably larger radius. Altogether, it leads to considerable improvement with respect to Belle in the vertex resolution and reconstruction efficiency for decays with hits in the vertex detector.

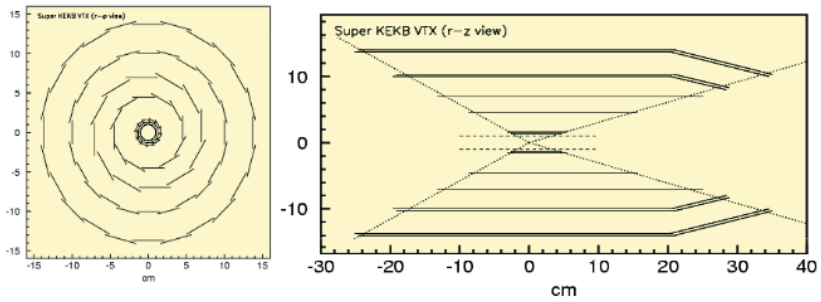


Figure 3.2. A schematic view of the Belle II vertex detector with a Be beam pipe: two pixelated layers and four layers with silicon strip sensors.

3.2.2 . Central Drift Chamber (CDC)

One of the key elements of the Belle II detector is the central tracking device. It is represented by a large volume drift chamber with small drift cells. Compared to Belle, it has a larger radius of 880 mm $\rightarrow 1130$ mm due to the upgrade to a much thinner particle identification (PID) device in the barrel region. In order to operate at higher event rates with increased backgrounds, the chamber has smaller drift cells than the one used in Belle. The CDC contains nearly 14k sense wires grouped in 56 layers, either in "axial" or "stereo" orientation. By combining information from axial and stereo layers allows reconstructing a full 3D helix track. The chamber is filled with $He - C_2H_6$ gas mixture of $50 : 50$ proportion with an average drift velocity of 3.3 cm$/\mu$s and a maximum drift time of about 350 ns for 17 mm cell size.

3.2.3 . Particle identification system (TOP and ARICH)

The time-of-propagation (TOP) counter is used [140,141] as particles identification system in the barrel region. It is a type of Cherenkov detector where the 2D information of a Cherenkov ring image is given by the time of arrival and impact position of Cherenkov photons at the photo-detector [141] at one end of a 2.6 m long quartz bar (see Fig. 3.3). Each of 16 detector modules consists of a quartz bar (45 cm wide and 2 cm thick) with a 10 cm long expansion volume at the sensor end of the bar. The expansion

wedge introduces additional pinhole imaging, relaxes the precision of timing requirements, and reduces the hit occupancy at the photo-detector. At the exit window of the wedge, two rows of sixteen fast multi-anode photon detectors are mounted. The required time resolution of photo-sensors with single-photon is ≈ 100 ps. It is achieved with a 16-channel MCP PMT [141, 142] specially developed for this purpose.

A custom-made waveform sampling read-out electronics is used [143] to achieve such time resolution. In the frame of particle identification method with TOP, the particle production time has to be known with a precision of 50 ps.

Another type of identification system is represented by ARICH, proximity focusing Cherenkov ring imaging detector with aerogel. It is used in the forward end-cap region to identify charged particles. It is able to identify low-momentum particles and has good separation power for pions and kaons from 0.4 GeV/c up to ≈ 4 GeV/c.

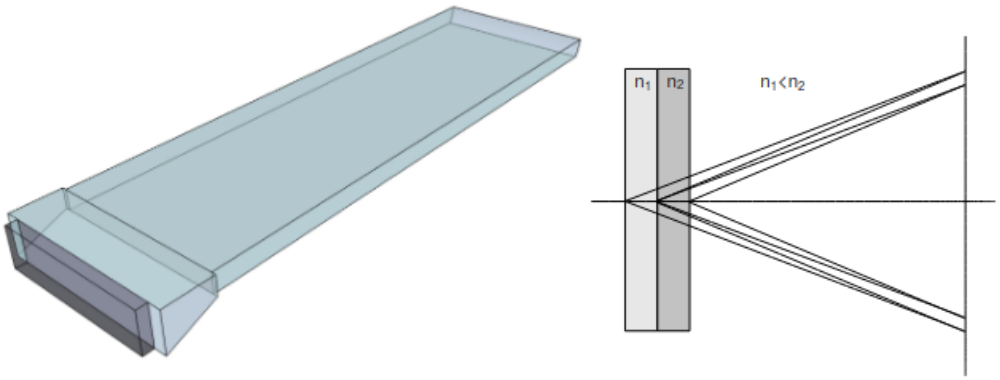


Figure 3.3. Belle-II PID systems: one of the modules of the TOP counter (left), principle of operation of the proximity focusing RICH (right)

Within ARICH, the number of detected Cherenkov photons is increased by a new method (see right part of Fig. 3.3): two 2 cm thick layers of aerogel with different refractive indices are used to increase the yield without degrading the Cherenkov angle resolution. The hybrid avalanche photon detector (HAPD) was developed in a joined effort with Hamamatsu [144, 145]. It is used as a high granularity sensor sensitive to a single photon. In this sensor, photo-electrons are accelerated over a potential difference of 8 kV and are detected in avalanche photodiodes (APD). Sensor production was optimized following radiation tolerance tests [145] with neutrons and gamma rays. It results in good separation of pions and kaons from 0.4 GeV/c up to ≈ 4 GeV/c.

3.2.4 . Electromagnetic Calorimeter (ECL)

The electromagnetic calorimeter (ECL) is used for the photons detection and separation of electrons and hadrons. The system is a highly-segmented array of $CsI(Tl)$ crystals assembled together. There are around 8k crystals split among three detector regions (barrel, forward and backward end-caps) covering nearly 90 % of the solid angle in the center of mass of e^+e^-. The crystals themselves, along with preamplifiers and support structures have been reused from Belle. Meanwhile, the readout electronics and reconstruction software have been upgraded.

In the abscence of backgroounds, the energy resolution for ECL is $\sigma E/E = 4\ \%$ at $100\ \mathrm{MeV}$ and $1.6\ \%$ at $8\ \mathrm{GeV}$. The angular resolution is $13\ \mathrm{mrad}$ ($3\ \mathrm{mrad}$) at low (high) energies. In the presence of a significant amount of background, scintillations in $CsI(Tl)$ crystals have high decay time and increase the overlapping of pulses from neighboring background events. The resulting noise is mitigated by using wave-form-sampling read-out electronics.

3.2.5 . K_L-Moun Detector (KLM)

The detector is made of a sequence of $4.7\ \mathrm{cm}$ thick iron plates and active detector elements located outside the superconducting solenoid. The iron plates serve as the magnetic flux return for the solenoid and have a 3.9 interaction length of material, compared to ECL, where this quantity is 0.8 for K_L^0. Originally, the Belle KLM was based on glass-electrode resistive plate chambers (RPC) and demonstrated good performance during the data-taking period. In the Belle II, the large background occurs from neutrons produced in the electromagnetic showers from background reactions in some detector areas (both endcaps and the inner-most layers in the barrel region). The issue of reducing the detection efficiency and fake muon identification probability was coming from the long-dead time of the resistive plate chambers (RPC). Consequently, they have been replaced by layers of scintillator strips with wavelength shifting fibers, read out by silicon photomultiplier (SiPMs, Geiger mode operated APDs) as light sensors [119].

3.2.6 . Trigger System

On the one hand, the trigger system of Belle II operates in the presence of 10-20 times higher background rates produced by SuperKEKB and on the other side satisfies the limitations of the data acquisition system (DAQ).

Several examples of triggers for new physics are single-photon triggers for dark sector searches, and the two- and three-photon triggers for axion-like particle searches. The dominant beam backgrounds are coming from the Touschek effect, Beam-gas scattering, synchrotron radiation, radiative Bhabha process, two-photon process, and beam-beam effects. The rates of these background processes, for instance, depend on beam size, beam current, luminosity, accelerator status, vacuum conditions. Most of these processes have less than two charged particle tracks in CDC and one or two clusters in the ECL. Such topologies are similar to low-multiplicity production modes and are therefore a large problem for such studies.

For events that are reconstructed by offline algorithms the efficiency of trigger for most B-decays is close to $100\ \%$. In addition to B physics, Belle II aims to perform studies on τ physics, dark sector, two-photon physics, and precision measurements of low-multiplicity processes. The trigger system of Belle II has two levels: hardware-based low-level trigger (L1) and software-based high-level trigger (HLT). Key design features of each level are as follows:

- The 3D tracking algorithms based on 3D-fitting, and Multi-Layer Perceptron (MLP) have been developed to compute the vertex position in the direction of the beam-line (z-axis). This is used to suppress beam background that is not coming from the interaction point. Using the 3D track information allows to match the CDC track with associated ECL clusters, and thereby improve PID at the trigger level;

- The backgrounds from radiative Bhabha scattering are reduced with improved online reconstruction techniques. For instance, applying Bhabha vetoes removes a substantial rate of interesting low-multiplicity events. In order to better suppress Bhabha events a 3D logic has been applied using 3D

ECL clustering information;

- The trigger information from each detector trigger is combined with an FPGA-based Global Reconstruction Logic (GRL) to match between tracks found in the CDC and clusters found by the ECL trigger. GRL is critical for controlling rates at high luminosity;
- Trigger menu has been extended and satisfies a variety of physics analysis targets. B decays and continuum are triggered with high efficiency by requiring that there are at least three tracks in CDC. Low multiplicity processes are mimicked by radiative Bhabha or beam background events, and are therefore difficult to efficiently trigger;
- Dark matter searches are characterized by the presence of only one energetic photon in the final state. For such events, Bhabha and $e^+e^- \to \gamma\gamma$ are the dominant background in the endcaps and at high luminosity;

3.3 . Reconstruction software

The data processing within Belle II is done by software framework *basf2* [131]. The software code operates on the principle of lazy computations: the user specifies several blocks, which are executed within a defined path, and only after explicitly specifying the processing command the script starts the procedure. The blocks are communicating through a common object store, which keeps track of their relationships. The datasets are processed in several phases. For instance, the raw data is obtained from the detector in the form of track hits and calorimeter clusters. This information is further used to reconstruct higher-level objects such as charged tracks. At this point, the information about low-level objects is discarded leading to a reduction of event size. The high-level information is further used to determine particle level information: 4-momentum, event shape variables, etc.

3.3.1 . Tracks reconstruction

Track reconstruction procedure can be performed in two steps. Initially, VXD and CDC hits are identified. They occur due to the ionization of a given particle and are distinguished among ionization of other particles and different types of background by the algorithm. In the second step, the trajectory is obtained from a fit of hits positions. Almost all tracks originate from the beam region, however, charged particles originating from decays of Λ, K_s^0, converted photons are produced outside the beam pipe. The task of the tracker algorithm is to identify all tracks and combine them.

Reconstruction of charged tracks

Charged particle reconstruction is done in two stages: track finding and track fitting. The first one refers to the procedure of collecting together appropriate hits belonging to a single track. The second one determines the trajectory of the track by fitting the tracking candidate.

The VXD track finder is based on cellular automaton [120]. In the frame of this algorithm pairs of hits in neighbor, layers are grouped together. Obtained cells are validated using a look-up table, called the sector map, which is created by simulating a large number of tracks in the VXD. In the next stage cells

that share a hit are submitted to the set of geometrical cuts. A Hopfield network is further applied to obtain a set of non-overlapping track candidates.

The sector map is not guaranteed to be stable for different values of momentum of the particle. Thus, several runs of the track finder algorithm are needed for different momentum regions.

Within CDC local and global track finders are used. Similar to CDC, local track finder searches for the tracks, track segments, and relation between neighboring hits using CA. The global track finder uses all hits at the same time. Its efficiency is the highest for the tracks with high transverse momentum coming from the origin. The algorithm applying both track finders is robust with respect to particles' energy losses and also the tracks not coming from the interaction point.

As a final step of track finder, VXD and CDC track candidates are combined together using the distance between track candidates in two sub-detectors.

Taking into account that in a vacuum in a constant magnetic field track moves along a helix parametrized by several quantities [117], the trajectories of the tracks in the Belle II are fitted. However, in reality, these trajectories are not perfect helices, since charged particles interact with the detector environment depositing their energy and scattering. In addition, the magnetic field of the superconducting solenoid is not constant in space. All these effects are properly taken into account during fitting and extrapolating the tracks.

In particular, the interaction of the charged particle with matter is treated by assuming a hypothesis about a certain mass of this article. There are five different hypotheses: electron, muon, pion, kaon, and proton.

The algorithm used for track fitting is called deterministic annealing filter (DAF) [121]. In order to deal with false hits assignments within this algorithm, hits with high residuals from the smoothest track are discarded.

Reconstruction of V^0

Long-lived neutral particles V^0 decay to two charged particles and are reconstructed by pairing all charged tracks with opposite charges to the most initial hit of one of the tracks. The combination is accepted if such extrapolation succeeds. If vertex fit of survived pair fails the combination is rejected by RAVE reconstruction software [122].

3.3.2 . Calorimeter reconstruction

The electromagnetic calorimeter reconstructs the coordinates and energy of the particles, performing its main task. For instance, summing up energies of all reconstructed showers allows limiting missing energy range in decays involving neutrinos. The second task of the device is particles identification based on variables representing the shower shape and tracks matched with clusters of energy depositions.

The calorimeter has three regions: barrel, forward and backward encaps. In the barrel, the size of a cluster is limited to a square of 5×5 crystals. The center $\overrightarrow{r}$ of each cluster is calculated as following:

$$\overrightarrow{r} = \frac{\sum_i E_i \overrightarrow{r}_i}{\sum_i E_i} \tag{3.1}$$

where sum goes over crystals within cluster, E_i — energy deposited in i-th crystal; $\overrightarrow{r}_i$ — spacial position of i-th crystal. The reconstruction of the position according to this equation is biased towards the highest energy crystal in the shower. The cluster energy is obtained as a sum of energies deposited in each crystal.

The cluster time is defined as the time between the moment of collision and the moment of reaching the highest energetic crystal. Clusters and tracks are matched together using GEANT functionality: cluster, which contains a crystal hit consistent with an extrapolated track, is matched to the track.

3.3.3 . Identification of charged particles

Particle identification allows to suppress background, separate hadronic final states, and use flavour-tagging algorithms. Within Belle II, upgraded TOP and ARICH detector systems provide the information for charged particles identification over the full kinematical range. It is combined with dE/dx measurements performed in SVD and CDC.

Information obtained from each detector sub-system is treated independently and likelihood for each particle hypothesis is provided. The likelihoods may be used together for creating a combined likelihood ratio. Binary PID value for two-particle hypotheses is defined using sigmoid function as follows:

$$L(\alpha:\beta) = \frac{1}{1+e^{\sum_{det}(\log \mathcal{L}_{\beta}^{det} - \log \mathcal{L}_{\alpha}^{det})}} = \frac{\prod_{det} \mathcal{L}_{\alpha}^{det}}{\prod_{det} \mathcal{L}_{\alpha}^{det} + \prod_{det} \mathcal{L}_{\beta}^{det}} \tag{3.2}$$

where det — set of detector subsystems used for particle identification; α, β — six long-lived charged particle hypotheses: electron, muon, pion, kaon, proton and deuteron; $\mathcal{L}_{\alpha}^{det}$ — likelihood of hypothesis α for subdetector det. The value $L(\alpha:\beta) > 0.5$ signifies that track of given particle is more likely to belong to particle of type α, rather than β.

The ionization energy loss (dE/dx) measurements are obtained from VXD and CDC. Such measurements should depend on particle velocity ($\beta\gamma = p/m$). However, the proper calibration of these subsystems is required to eliminate systematics, which could violate such dependence [117].

Charged hadrons (π, K, p, d) identification depends initially on information from the CDC, TOP, and ARICH. These detectors are also responsible for the identification of electrons and muons.

The main method of muon identification within KLM relies on the differences in longitudinal penetration depth and transverse scattering of the extrapolated track. There are two steps involved: extrapolation of the track taking into account muon hypothesis and likelihood calculation for each of six possible particle hypotheses: μ, K, π, p, d, e. The log-likelihood differences further can be used to accept or reject the muon hypothesis for a given track.

Electron identification applies likelihoods from ECL, CDC, SVD, TOP, and ARICH. The value E/p is the main variable for separating electrons from other particle hypotheses (predominantly μ and π). At low values of the electron's momentum, the radius of the curvature in the magnetic field is small. It doesn't allow to reach ECL and causes a drop in identification efficiency for the electrons.

3.3.4 . Identification of neutral particles

Within ECL the photon is identified using the shower shape of the clusters, which are not matched to a reconstructed track. The identification procedure applies the property that electromagnetic shower caused by a photon is cylindrically symmetric around its direction. In addition, the energy deposition decreases exponentially from the incident axis. The dominant background for the photon identification comes from hadron interactions, which produce asymmetric showers and several ECL clusters not associated with any charged track.

Consequently, π^0 reconstruction in the process $\pi^0 \to \gamma\gamma$ combines two-photon candidates. Depending on the range of π^0 energy, the topology of the decay is different. In the case of low energy (< 1 GeV),

the angle between two photons in the laboratory frame is sufficient to observe two non-overlapping ECL clusters. In the middle range ($1 < E_{\pi^0} < 2.5$ GeV) two ECL clusters overlap but can be reconstructed as separate photon candidates. Resolution of π^0 energy is further improved by performing mass constrained fit of the two photons to the nominal mass of π^0. For highly energetic π^0 (> 2.5 GeV), two photons are indistinguishable and are reconstructed as one photon.

The identification of K_L^0 mesons is done using ECL and KLM detectors. Based on boosted decision trees (BDT) [123], ECL and KLM clusters are independently classified on the subject of originating from K_L^0. The main background contributions come from neutrons and photons. The latter ones are predominantly originating from beam interactions with detector and beam-pipe followed by neutral particle products of such processes. BDT input is composed of several features using the KLM detector. The strongest features are as follows:

- distance to nearest track should be large for neutral clusters;
- fake clusters from the beam background should appear much earlier than the ones from primary collisions;
- hadronic clusters more likely have a larger radius than electromagnetic clusters;

In the ECL the most significant features have the following properties:

- neutral clusters should have rarely the tracks nearby;
- the shape of the distribution of the ratio of energies in the inner 3×3 and outer $5 \times 5 - 4$ clusters depends on the shower's origin: hadronic or electromagnetic;
- K_L^0 usually deposits < 50 MeV energy in the cluster;

4 - Study of the $B^+ \to K^+\pi^+\pi^-\gamma$ signal

4.1 . Introduction

The "radiative penguin" transition $b \to s\gamma$ proceeds through flavor-changing neutral current, which is described by loop diagram in Standard Model (SM) and therefore results in a left-handed photon emission (but for small QCD corrections). Therefore it is a promising process for searching for New Physics that may occur within the loop.
If the hadronization of the s quark and the spectator quark entails at least three particles, it becomes possible to infer the polarization state of the photon, and to detect possible New Physics contributions since the latter can strongly affect the photon polarization.
In this analysis, we are considering the decay $B \to K_{\rm res}\gamma$, where $K_{\rm res}$ denotes a $K\pi\pi$ final state assumed to result from a set of resonances, each able to decay into $K\pi\pi$, through $K^*\pi$ or $K\rho$ intermediate decays. The p.d.f. distribution of events depends on the photon polarization, but also crucially on the resonance structure of $K_{\rm res}$. This structure must be derived simultaneously from the data analysis, and it has been shown [49] that the photon polarization can be measured in such $B \to K\pi\pi\gamma$ decays, provided one restricts the analysis to the range $M_{K\pi\pi} < 1.6\ {\rm GeV}/c^2$ to limit the number of resonances contributing to the decay. To begin with, the selection will apply a looser cut $M_{K\pi\pi} < 1.8\ {\rm GeV}/c^2$, to be tightened later.

The present work aims to provide the first step of this long-term analysis program, and to provide in passing an estimate of the branching ratio of the $B^\pm \to K^\pm\pi^\pm\pi^\mp\gamma$ decay. The analysis strategy developed in the following is not meant primarily to allow for the branching ratio measurement, it is meant to allow for the second step of the analysis: namely, using ${}_s\mathcal{P}lot$ to analyze the distribution of events and infer the photon polarization together with the resonance structure.
Because of that, the analysis stress is not on optimizing the selection efficiency, but on optimizing the selection potential to isolate a clean set of events, where backgrounds ($q\bar{q}$, $B\bar{B}$, self-cross feed) are strongly suppressed and truth-matched events is a dominant contribution. Although the set of events is considered clean, it still contains contributions from the above-mentioned background types. Having performed fit for the yields, a background subtraction using ${}_s\mathcal{P}lot$ allows recovering the truth-matched distributions. This clean set of events is referred to as Truth Match (TM) in the following, since, for the Monte-Carlo data, these are events where the full $K^+\pi^+\pi^-\gamma$ final state is matched to the generated one.
The Standard Model expectations according to the PDG [147] for the branching ratio of $B^+ \to K^+\pi^+\pi^-\gamma$ is $2.58 \pm 0.15 \cdot 10^{-5}$.

4.2 . Samples and event selection

The Belle-II data consists of large amount of events coming from $e^+e^- \to \Upsilon(4S) \to B\bar{B}$ and from background processes : $q\bar{q}$ pairs (so-called continuum events), lepton-pairs, two-photons and beam-gas interactions. In this study only $e^+e^- \to B\bar{B}$, $q\bar{q}$, and $\tau^+\tau^-$ events are considered, since the other processes do not contribute.
At this stage the analysis is blind and uses only Monte-Carlo simulated events, corresponding to $1\ {\rm ab}^{-1}$, namely the two samples:

1) A generic Monte-Carlo sample that includes: $q\bar{q}$, $\tau\bar{\tau}$, B^+B^-, and $B^0\bar{B}^0$ events, except that $B^\pm \to K^\pm\pi^\pm\pi^\mp\gamma$ events have been removed;

2) A signal Monte-Carlo sample obtained using EvtGen but with the "GamPola" model (described in the Chapter 2) in place of the default model.

The generator used by default in EvtGen to produce $B^+ \to K^+\pi^+\pi^-\gamma$ events is using a simplified model for $K_{\rm res}$ that doesn't take into account interfering of multiple resonances (see Eq. 2.1). To include such effect (see Eq. 2.2) in the signal MC generation, "GamPola" is used.
For $\mathcal{L} = 1\ \mathrm{ab}^{-1}$, the number of signal events expected to be produced is:

$$\begin{aligned} N &= \mathcal{L} \cdot \sigma_{\Upsilon(4s)} \cdot \mathrm{Br}_{B^+B^-} \cdot \mathrm{Br}_{K^+\pi^+\pi^-\gamma} \\ &= 10^9 \times 1.110 \times 0.514 \times 2.58\ 10^{-5} = 1.47 \cdot 10^4 \end{aligned} \tag{4.1}$$

where the branching ratio values of $\mathrm{Br}_{K^+\pi^+\pi^-\gamma}$ and $\mathrm{Br}_{B^+B^-}$ [147] are the PDG averages that correspond to $K^+\pi^+\pi^-\gamma$ and $B\bar{B}$ final states with an invariant mass $M_{K^+\pi^+\pi^-} < 1.8\ \mathrm{GeV/c^2}$.
Considering both B^+ and B^- decays, it is expected about

$$N = 3 \cdot 10^4 \quad \text{events} \tag{4.2}$$

signal events in the $M_{K^+\pi^+\pi^-} < 1.8\ \mathrm{GeV}/c^2$ region. This analysis is targeting real data (Moriond dataset) consisting of $9.2\ \mathrm{fb}^{-1}$ of off-resonance and $62.7\ \mathrm{fb}^{-1}$ of $\Upsilon(4S)$ data.

Using a rather large range of $K^+\pi^+\pi^-$ invariant mass, random combinations of one kaon and two pions lead to multiple B-meson candidates both in signal and in background events. The photon polarization measurement imposes that a selection is made among the candidates to select the "best" one, once the number of candidates per event ($N_{\rm can}$) is small enough (of order unity). The "best candidate selection" (BCS) is designed to identify the candidate most likely to be TM.

Two powerful and theoretically weakly correlated variables used in the B-Physics to discriminate the signal and background events are beam constraint mass (M_{bc}) and the energy difference between reconstructed B-meson and e^+e^- beam (ΔE). They are defined using Eqs. 4.5-4.6 and follow from the constraints in the $\Upsilon(4S)$ rest frame. In the case when B-meson candidate is perfectly reconstructed, in the absence of beam energy spread, reconstructed energy of B should be equal to the energy of the beam and reconstructed invariant mass equals the nominal mass of the B.

$$E^*_{rec} = E^*_{beam} \tag{4.3}$$
$$M_{rec} = M_B^{\rm nominal} \tag{4.4}$$
$$\Delta E = E^*_B - E^*_{beam} \tag{4.5}$$
$$M_{bc} = \sqrt{E^{*2}_{beam} - \overrightarrow{p}^{\,2}_B} \tag{4.6}$$

where E^*_{beam} — energy of the e^+e^- beam; $M_B^{\rm nominal}$ — nominal mass of B-meson; E^*_B — reconstructed energy of the B; $\overrightarrow{p}_B$ — 3-momentum of B-meson.

The two variables are weakly correlated since, for instance, if pion was misidentified with kaon, ΔE moves away from 0, but M_{bc} value doesn't change.

As discussed below, the ΔE variable proves to be a powerful element of BCS, with the result that ΔE is not available anymore for the yield measurement, which is using only the $M_{\rm bc}$ variable.

4.2.1 . Truth-Matched candidates

At all stages of the event selection described below, a given event is associated with N_{can} candidates. Because of combinatorics, in the early stage of the analysis where few cuts have been applied the number of candidates, N_{can} is very large. This is illustrated in Fig.(4.1) that is obtained after reconstruction of the B-candidates, just before the first selection cut on M_{bc} of Subsection (4.2.5) (see also Table (4.1)).

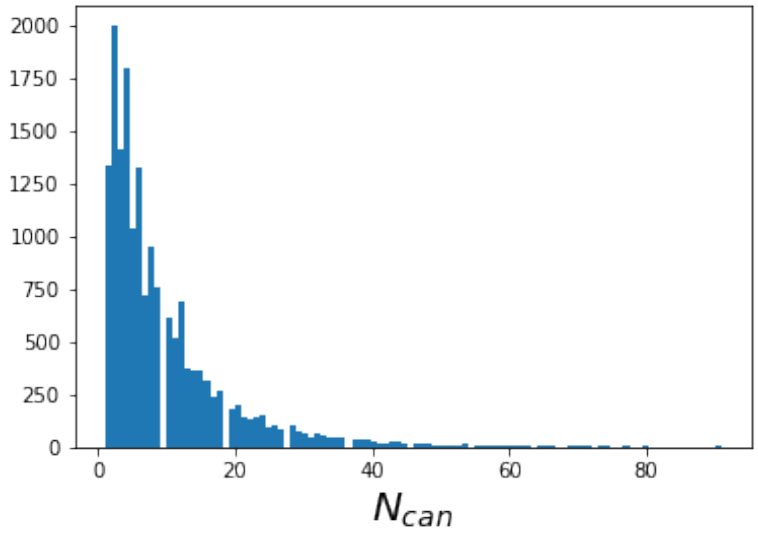


Figure 4.1. The distribution of N_{can}, the number of candidates per event for signal $B^+ \to K^+\pi^+\pi^-\gamma$ events.

Because the aim of the analysis is to search for signs of new physics in the process through a multi-dimensional analysis (cf. section(4.1)) the proper kinematical reconstruction of the final state is instrumental. For instance, a signal event where none of the N_{can} candidates is made of the correct particles is not useful for the analysis, although it would be useful for an analysis aiming at measuring the branching fraction of the decay.
Therefore, $B^+ \to K^+\pi^+\pi^-\gamma$ event can be:

1) TM (truth-matched) event, which, as mentioned in the introductory part, is an event including one candidate for which all four particles are associated with the generated ones. All other candidates within events are called false-matched candidates.

2) FM (false-matched) is an event having all candidates with at least one of the four particles not associated with the generated one. Every candidate within such an event is a false-matched candidate.

The event selection described below ultimately targets the TM events since the selection of events with all correctly reconstructed particles in the final state is the final goal of the analysis. At the skimming and preselection stages, most cuts are applied to ensure a good quality of reconstruction, not to eliminate FM candidates. The cuts are applied in sequence, each figure corresponding to a given variable is obtained after having applied all the previous cuts.

4.2.2 . Skimming cuts

Skimming cuts are represented by general selection prior to the cuts specifically applied within the particular analysis of the B-decay. They allow significantly reduce CPU consumption while reconstructing a particular decay mode. In the case of $B^+ \to K^+\pi^+\pi^-\gamma$ at least three charged tracks ($nTracks \geq 3$), and highly energetic photon are required in the final state. In addition, the cut on the ratio of second and zeroth order of Fox-Wolfram variables [115] is applied. Fox-Wolfram moments defined by Eq. 4.7 are based on information coming from the rest of the event and help to discriminate between $B\bar{B}$ and continuum events.

$$H_l = \sum_{i,j} \frac{|\overrightarrow{p}_i| \cdot |\overrightarrow{p}_j|}{E^2} P_l(\cos\theta_{ij}) \tag{4.7}$$

$$\cos\theta_{ij} = \cos\theta_i \cos\theta_j + \sin\theta_i \sin\theta_j \cos(\phi_i - \phi_j) \tag{4.8}$$

$$\tag{4.9}$$

where $\overrightarrow{p}_i, \overrightarrow{p}_j$ — indices of all final state particles obtained from rest of event; θ_i, ϕ_i — angles of i-th particle in spherical coordinates; P_l — Legendre polynomial of order l. In the case of jet-like event foxWolframR2 values are closer to 1.

Evidence of the photon presence in the final state is enhanced by applying the cut on the $clusterE9E21$ variable. According to [116] it is defined within electromagnetic calorimeter as a ratio of energies in the inner 3x3 crystals (E_9) and 5x5 crystals around the central crystal without corners (E_{21}). The ratio is expected to be high for the photons and small for the hadrons.

Skimming cuts can be summarized as follows:

- $nTracks \geq 3$ ($p_T > 0.1$ GeV/c and $E_\gamma > 0.1$ GeV) — event level cut constructed using tracks with transverse momentum p_T greater than 0.1 GeV/c and clusters with $E_\gamma > 0.1$ GeV in the laboratory frame;
- foxWolframR2 $= \dfrac{H_2}{H_0} < 0.5$ ($p_T > 0.1$ GeV/c and $E_\gamma > 0.1$ GeV) — event level cut on the event shape to reduce continuum background;
- $1.4 < E^\gamma_{\mathrm{CMS}} < 3.4$ GeV, where E^γ_{CMS} — energy of the photon in the center of mass of B-meson;
- $clusterE9E21 = \dfrac{E_9}{E_{21}} > 0.9$;

4.2.3 . Photon selection

The signals coming from the photons are detected by ECL as clusters of activated crystals. The photon energy along with cluster information is used for the selection:

- $clusterE9E21 > 0.95$ — tighten skimming cut suppressing signals from hadron showers and keeping signals of photon detection;

- $1.8 < E^{\gamma}_{\rm CMS} < 3.4$ GeV retains photons with high energy in the CMS, since the photon's energy from $K^+\pi^+\pi^-\gamma$ decays is

$$E^{\gamma}_{\rm CMS} \simeq \frac{m_B^2 - {M_{K\pi\pi}}^2}{2m_B} \simeq 2.5 \text{ GeV} \tag{4.10}$$

- $clusterNHits = \sum_{i=0}^{n} w_i \geq 8$ — at least 8 effective crystals should detect a signal from photon, where w_i — weight term of each crystal: integer for non-overlapping crystals and can be floating in case of energy splitting among nearby clusters; n — number of the crystals forming signal cluster;

- $clusterSecondMoment = \frac{\sum_{i=0}^{n} w_i E_i r_i^2}{\sum_{i=0}^{n} w_i E_i} \leq 1.5$ — variable associated with spread of cluster within ECL, where E_i are the single crystal energies, and r_i is the distance of the i-th cluster to the shower center in the plane ortogonal to the shower axis.

The distributions of the above four quantities are shown in the Figs.(4.2)-(4.3) for TM candidates, after applying the skimming cuts.

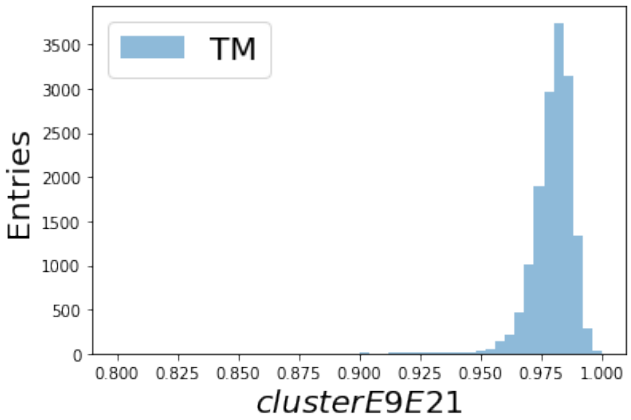


Figure 4.2. The distribution of $clusterE9E21$ for TM candidates.

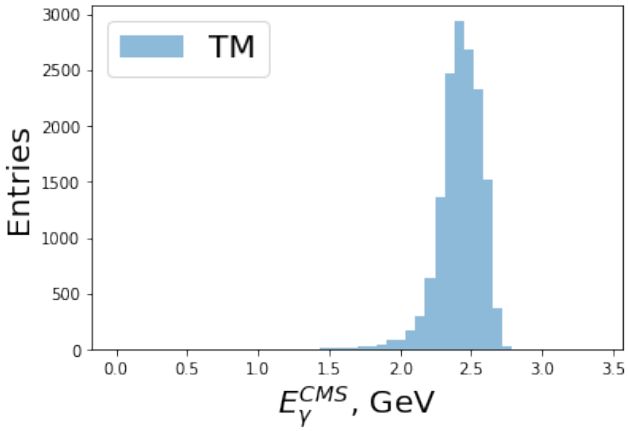


Figure 4.3. $E^{\gamma}_{\rm CMS}$ distribution: the upper bound 3.4 GeV of the associated cut is further tighten within selection cuts which follow.

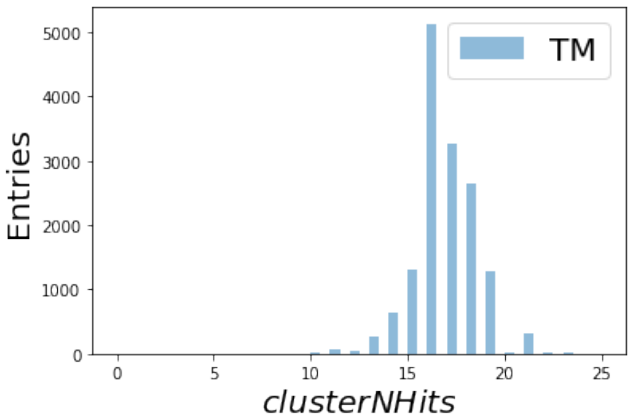


Figure 4.4. The distribution of $clusterNHits$ for TM candidates.

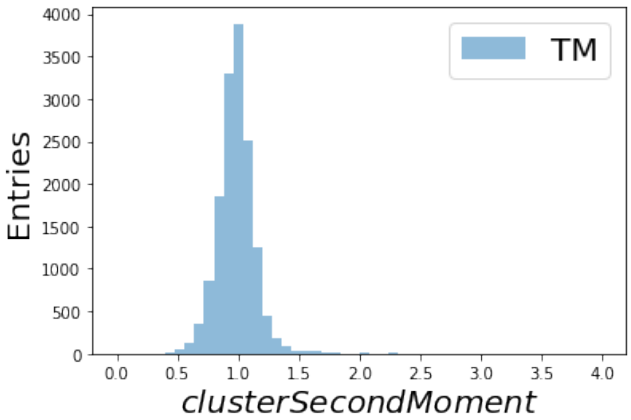


Figure 4.5. The distribution of $clusterSecondMoment$ for TM candidates.

The width of the $E^{\gamma}_{\rm CMS}$ distribution is driven by the resolution on the photon energy: at truth level the full range is approximatively $E^{\gamma}_{\rm CMS} \in [2.10, 2.70]$ GeV (see Fig.(4.22)). More generally, for the key quantities $M_{\rm bc}$, ΔE, and $m_{K^+\pi^+\pi^-\gamma}$ belonging to TM candidates, the photon reconstruction, and especially its energy reconstruction is the leading factor driving the difference between truth-level and reconstruction level distributions. This is illustrated later in section (4.2.5).

Additional quality cuts related to photon selection and applied in the late preselection stage are $0 < P(\gamma|\pi_0)$ and $0 < P(\gamma|\eta)$. They represent cuts on probabilities of detecting the final state photon from either $\pi_0 \to \gamma\gamma$ or $\eta \to \gamma\gamma$. The probabilities should essentially be positively defined.

4.2.4 . K and π selection

A track propagating in a vacuum in a constant magnetic field moves along a helix trajectory [117]. The point of closest approach (POCA) for a given track is chosen in the cylindrical system of coordinates (r, φ, z) to minimize the distance from an interaction point to the track. The distance contains unsigned transversal (dr) and signed longitudinal (dz) components with respect to the beam axis z.

Then charged tracks are selected applying the default quality cuts on the POCA:

- $dr < 0.5$ cm
- $|dz| < 2$ cm

Charged particles assignment to K or π is defined through either:

- $\dfrac{\mathcal{L}_\pi}{\mathcal{L}_K + \mathcal{L}_\pi} > 0.6$
- $\dfrac{\mathcal{L}_K}{\mathcal{L}_K + \mathcal{L}_\pi} > 0.6$

where $\mathcal{L}_{K,\pi}$ — combined PID likelihoods of K, π from all subdetector systems: SVD, TOP, CDC, ARICH, ECL, and KLM. The distributions of the above quantities are shown in Figs.(4.6)-(4.11) for signal events.

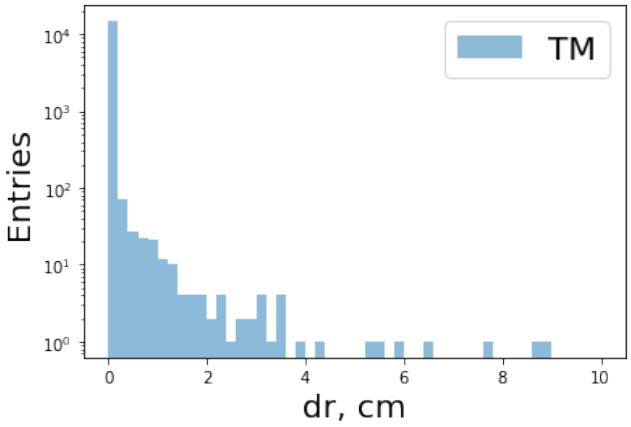


Figure 4.6. The distribution of dr_π for TM candidates.

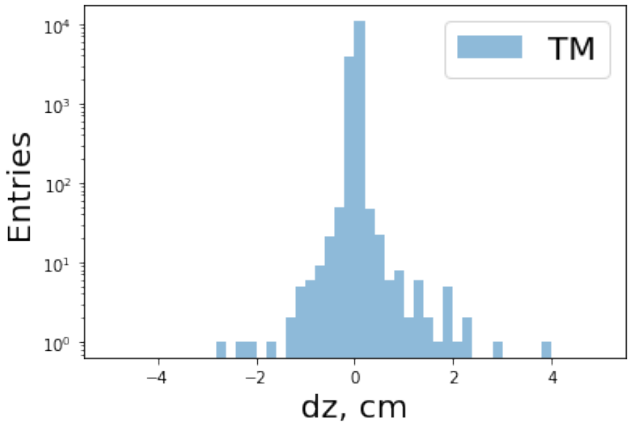


Figure 4.7. The distribution of dz_π for TM candidates.

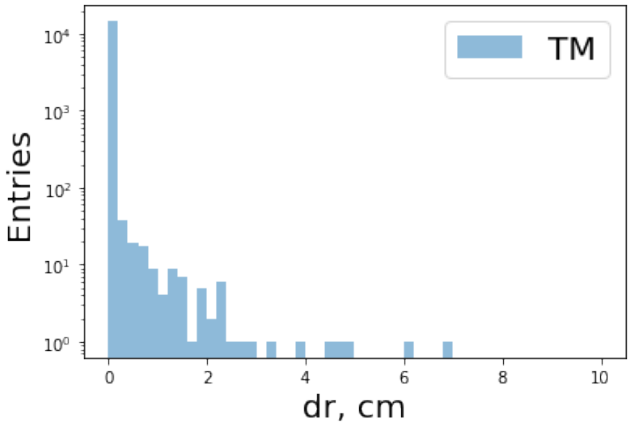


Figure 4.8. The distribution of dr_K for TM candidates.

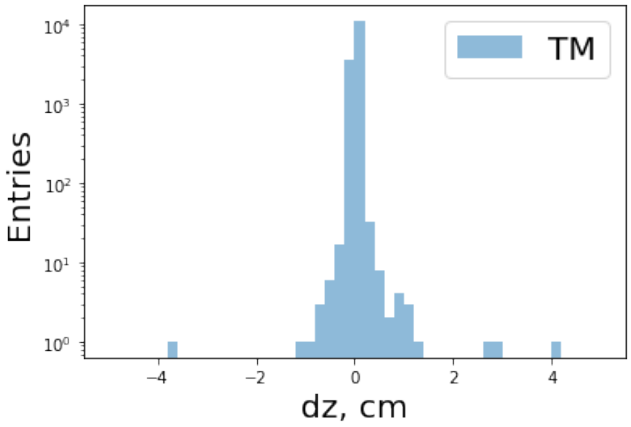


Figure 4.9. The distribution of dz_K for TM candidates.

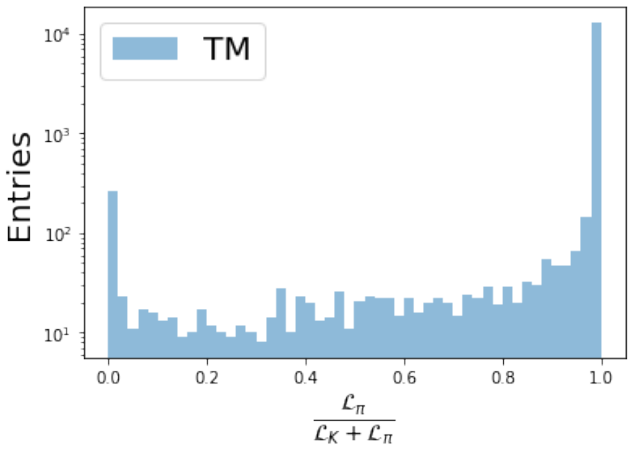


Figure 4.10. The distribution of PID_π for TM candidates.

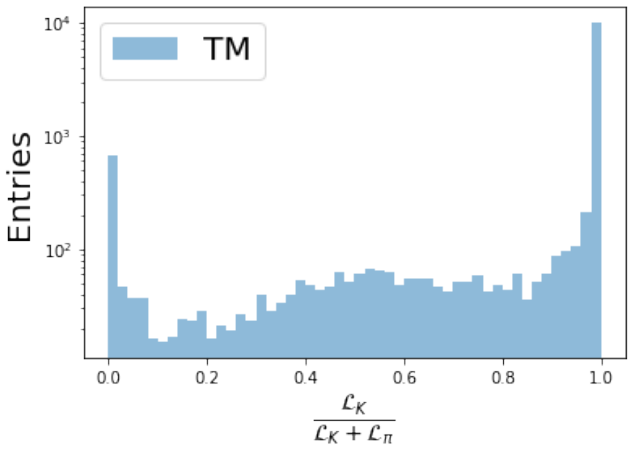


Figure 4.11. The distribution of PID_K for TM candidates.

4.2.5 . Selection cuts

To reduce background (and in particular to reduce N_{can}), events are further submitted to the following cuts:

(1) $5.2 < M_{bc} < 5.29\ \mathrm{GeV}/c^2$;

(2) $0 < M_{K\pi\pi} < 1.8\ \mathrm{GeV}/c^2$;

(3) $-0.2 < \Delta E < 0.1\ \mathrm{GeV}$;

These cuts are still quite loose for what concerns TM candidates, as illustrated in Figs.(4.12)-(4.14). The M_{bc} distribution in the Fig. 4.12 is depicted after applying cut (1) and $M_{K\pi\pi}$, ΔE are given before applying (2) and (3) respectively. The cuts (1) and (3) can be tightened on the stage of the photon polarization measurements. In addition, such loose selection leaves a window for the backgrounds which will be estimated better given a sufficient range of discriminating variables. Although the branching ratio is not the final concern of this analysis, its measurement is a required intermediate step. Thus precise measuring of yields for signal and background events is an important step.

The efficiency of the cut (2) is 100 % for $B^+ \to K^+\pi^+\pi^-\gamma$ events since according to introductory part input PDG value of branching ratio and consequently $N = 3 \cdot 10^4$ events are given in this range. On the Fig. 4.13 $M_{K\pi\pi}$ distribution for TM corresponds to "GamPola" output before applying cut (2) and extends up to $2\ \mathrm{GeV/c^2}$. Thus, even though applying this cut removes signal events (TM and FM) in the range $M_{K\pi\pi} > 1.8\ \mathrm{GeV/c^2}$, they are not taken into account for the efficiency calculation for signal events. However, when cut (2) is applied for $q\bar{q}$, $B\bar{B}$ and $\tau^+\tau^-$ backgrounds, its efficiency is taken into account.

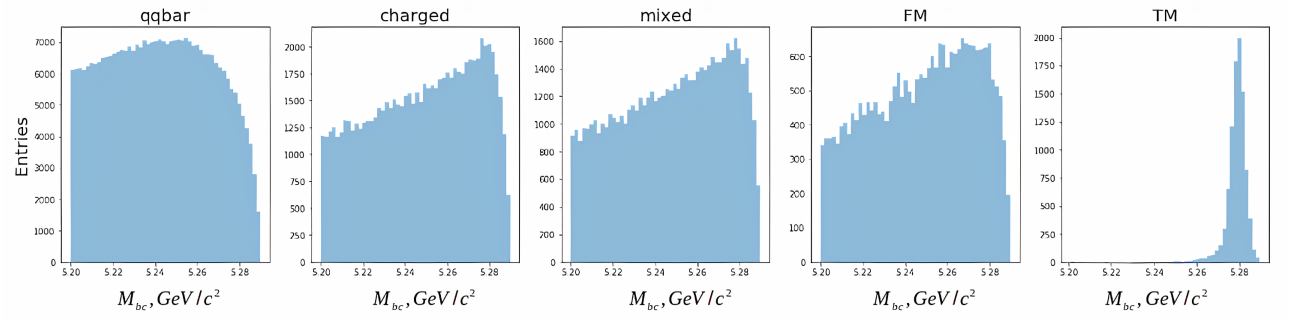


Figure 4.12. The distribution of M_{bc} for TM and main background categories.

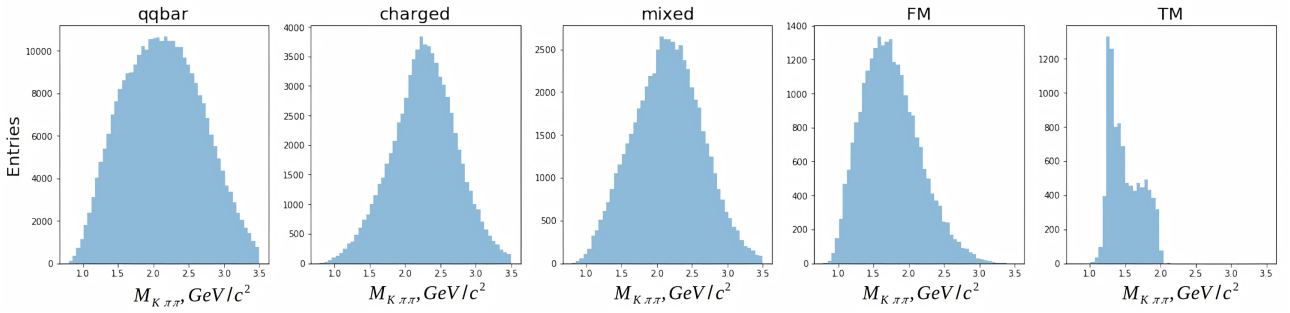


Figure 4.13. The distribution of $M_{K\pi\pi}$ for TM and main background categories.

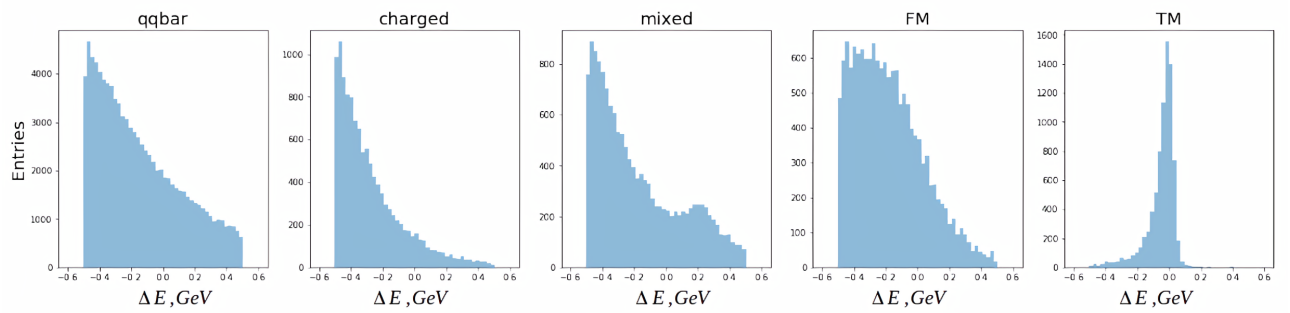


Figure 4.14. The distribution of ΔE for TM and main background categories.

The vertex reconstruction for the three charged tracks must satisfy the quality cut:

(4) $\chi_{\text{prob}} > 0.0001$ (see Fig. 4.15), where χ_{prob} — p-value of B-meson vertex fit.

The cut was not decided on firm ground, it was chosen to keep approximately 95% of the TM events, that is to say, accepting a reasonable cost inefficiency, but expecting in return a background rejection for data not too degraded compared to the background rejection Monte-Carlo. A cut optimization based on a figure of merit is not possible at this stage of the selection, but the subject will be addressed again briefly after selection, in Subsection (4.3.3).

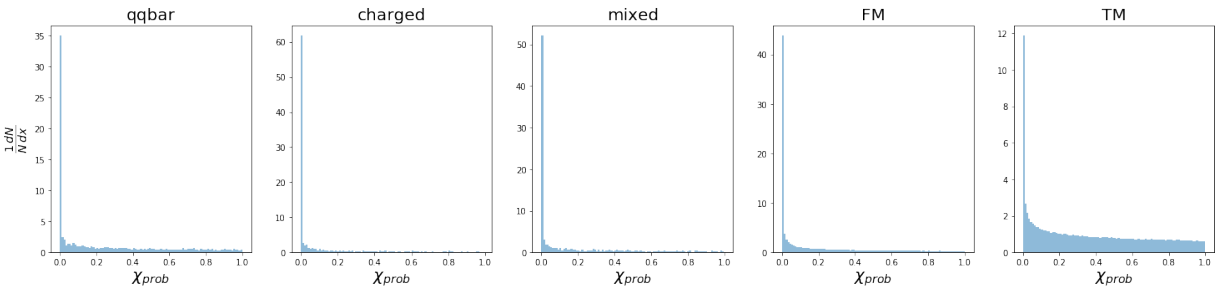


Figure 4.15. The distribution of $\chi_{\rm prob}$ for TM candidates.

To reduce background events where the photon comes from a π^0 decay, one uses the probability cut:

(5) $P(\gamma|\pi^0) < 0.8$,

where the probability for hard photon $P(\gamma|\pi^0)$ is internally calculated by reconstruction software using:

- invariant mass of π^0;
- soft photon energy in the laboratory frame;
- soft photon ECL cluster's polar angle;
- soft photon output of MVA using Zernike moments [124] of the cluster;
- soft photon distance from ECL Cluster to the nearest point on nearest Helix at the ECL cylindrical radius;

In addition, in the default weight files a value of 1.4 GeV is set as the lower limit for the hard photon energy in the CMS frame.

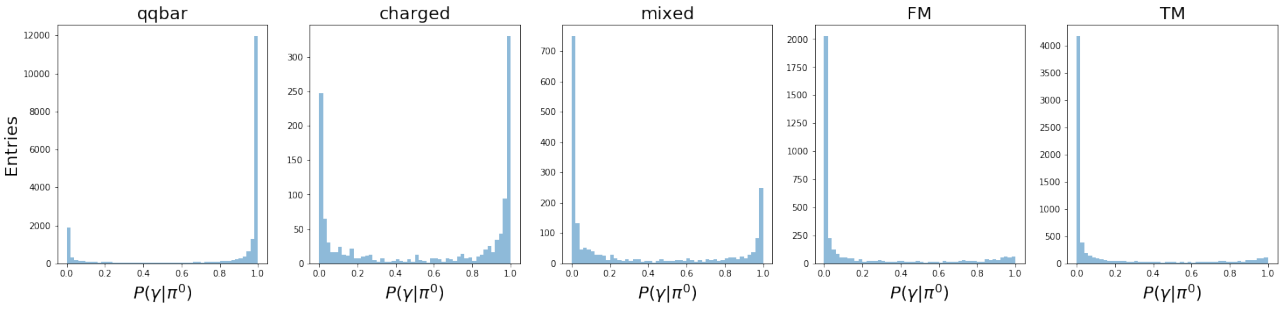


Figure 4.16. The distribution of $P(\gamma|\pi^0)$ for background and TM candidates.

The cut values were chosen from Fig.(4.16). The cut on $P(\gamma|\pi^0) < 0.8$ is chosen to get rid of the backgrounds while keeping the bulk of TM events. It appears to provide a powerful means to suppress generic background from $q\bar{q}$ and $B\bar{B}$. A cut optimization based on a figure of merit is not possible at this stage of the selection where the background level is very large, while it will be reduced by an order

of magnitude later. Furthermore, at this stage of Belle II life, it might be premature to attach too much credit to optimization using Monte-Carlo based figure of merit. This subject will be addressed again briefly after the final selection, in section (4.3.3).

(6) $2.1 < E_{\gamma}^{CMS} < 2.75$ GeV From the Fig. 4.2 it can be seen, that cut on E_{γ}^{CMS} can be tighten and according to Fig. 4.17 it removes noticeable amount of background events.

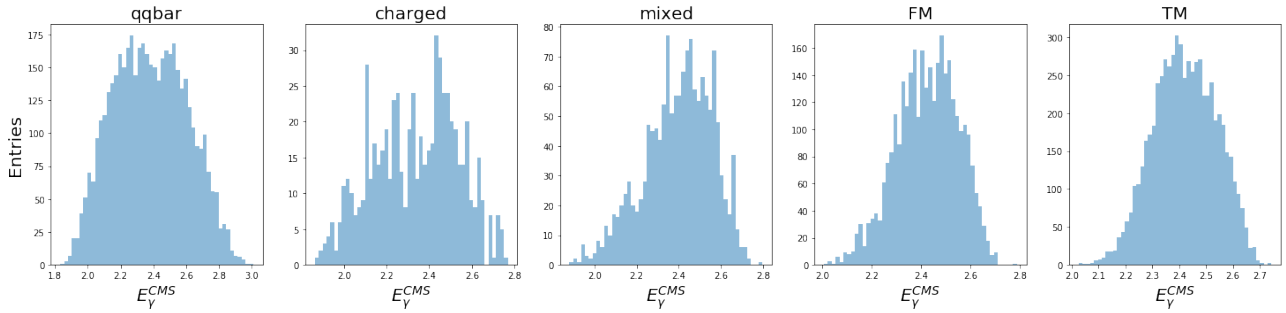


Figure 4.17. The distribution of E_{γ}^{CMS} for background and TM candidates.

Preselection and selection cuts' efficiency

The tables below summarize the efficiency of the cuts for signal and background events. In Table (4.1) the two values are the fraction of signal events and the fraction of events with a TM candidate that are kept by a given cut, while all previous cuts have been satisfied. In Table (4.2) the values correspond to the selection efficiency cuts for various backgrounds.

Cuts	All events	TM events
Reconstruction	97.25%	66.33%
Skimming cuts		
$nTracks \geq 3$	100.0%	100.0%
$foxWolframR2 < 0.5$	97.07%	97.06%
$clusterE9E21 > 0.9$	99.97%	99.53%
$1.4 < E_{\gamma}^{CM} < 3.4$ GeV	86.01%	99.49%
Summary skimming	83.47%	96.12%
Preselection cuts		
$clusterE9E21 > 0.95$	98.4%	98.75%
$1.8 < E_{\gamma CM} < 3.4$ GeV	98.31%	98.73%
$clusterNHits \geq 8$	100.0%	100.0%
$clusterSecondMoment \leq 1.5$	98.82%	99.16%
$dr_{\pi_1} < 0.5$	100.0%	99.63%
$\lvert dz_{\pi_1} \rvert < 2$	100.0%	99.84%
$dr_{\pi_2} < 0.5$	100.0%	99.39%
$\lvert dz_{\pi_2} \rvert < 2$	99.97%	99.79%
$dr_K < 0.5$	100.0%	99.71%
$\lvert dz_K \rvert < 2$	99.94%	100.0%
$\frac{\mathcal{L}_{\pi_1}}{\mathcal{L}_K + \mathcal{L}_{\pi_1}} > 0.6$	100.0%	93.32%
$\frac{\mathcal{L}_{\pi_2}}{\mathcal{L}_K + \mathcal{L}_{\pi_2}} > 0.6$	98.35%	95.01%
$\frac{\mathcal{L}_K}{\mathcal{L}_K + \mathcal{L}_{\pi}} > 0.6$	94.0%	87.81%
$0 < P(\gamma \lvert \pi^0)$	99.87%	99.84%
$0 < P(\gamma \lvert \eta \rvert)$	99.87%	99.95%
Summary preselection	72.91%	73.25%
	Selection cuts	
$5.20 < M_{bc} < 5.29$ GeV/c^2	86.64%	100.0%
$M_{K\pi\pi} < 1.8$ GeV/c^2	100%	100%
$-0.2 < \Delta E < 0.1$ GeV	81.72%	91.72%
$0.0001 < chiProb$	92.65%	94.41%
$0 < P(\gamma \lvert \pi^0) < 0.8$	90.0%	89.83%
$2.1 < E_{\gamma}^{CMS} < 2.75$ GeV	99.58%	99.84%
Summary selection	71%	78.33%
Overall	42.0%	36.58%

Table 4.1. Cuts' efficiencies for signal (GamPola) events

The values in the second and third columns are computed by dividing the number of events before the given cut by the number of events after this cut[1]. The same applies for Table (4.2). The summaries of skimming, preselection and selection cuts are obtained by multiplying the efficiencies of all cuts within the corresponding groups of cuts. The overall efficiency is obtained by multiplying the total efficiencies of the reconstruction and the three groups of cuts.

Cut/Events kept %	B^+B^-	$B^0\bar{B}^0$	$u\bar{u}$	$c\bar{c}$	$d\bar{d}$	$s\bar{s}$	$\tau^+\tau^-$
$M_{K\pi\pi} < 1.8$ GeV/c^2	26.38%	31.07%	37.01%	38.66%	37.14%	38.31%	43.1%
$-0.2 < \Delta E < 0.1$ GeV	23.32%	28.57%	36.08%	34.94%	35.84%	35.68%	33.2%
$0.0001 < \chi_{\rm prob}$	54.81%	60.85%	77.83%	70.1%	77.48%	76.31%	65.5%
$0 < P(\gamma\|\pi^0) < 0.8$	51.43%	73.48%	21.87%	27.1%	21.05%	23.06%	17.86%
$2.1 < E_\gamma^{CMS} < 2.75$ GeV	85.25%	94.09%	82.45%	82.38%	83.83%	79.77%	80.0%
Summary	1.5%	3.7%	1.9%	2.1%	1.8%	1.9%	1.3%

Table 4.2. Cuts' efficiencies for backgrounds: although obtained selection efficiencies (without taking into account preselection and skimming cuts) are much smaller than for the signal events, associated number of background events is much larger than signal events.

Photon reconstruction performance studies

As mentioned in section (4.2.3), the reconstruction of the photon determines the width of the $E^{\gamma}_{\rm CMS}$ distribution. The same is true for the distribution of $M_{\rm bc}$, ΔE and $m_{K^+\pi^+\pi^-\gamma}$, as illustrated on Figs.(4.18)-(4.21) : essentially all the difference between Monte-Carlo truth and reconstruction level for the three variables comes from the energy reconstruction (cf. Fig.(4.19)) where notation $E_\gamma^{truth} \to E_\gamma^{reco}$ means replacing truth photon energy by reconstructed one. In this case truth 4-momentum of the photon is modified as following.

$$(E_\gamma^{truth}; E_\gamma^{truth} \cdot \overrightarrow{e}_\gamma^{truth}) \to (E_\gamma^{reco}; E_\gamma^{reco} \cdot \overrightarrow{e}_\gamma^{truth}) \tag{4.11}$$

On the other side, replacing truth direction of the photon by reconstructed one ($\overrightarrow{e}_\gamma^{truth} \to \overrightarrow{e}_\gamma^{reco}$) modifies 4-momentum:

$$(E_\gamma^{truth}; E_\gamma^{truth} \cdot \overrightarrow{e}_\gamma^{truth}) \to (E_\gamma^{truth}; E_\gamma^{truth} \cdot \overrightarrow{e}_\gamma^{reco}) \tag{4.12}$$

According to the Fig. 4.20 such replacement almost doesn't affect M_{bc}, ΔE and m_B distributions.

The third option is to replace truth 4-momentum of the photon by reconstructed one as in Fig. 4.21. Combined effect of two previous cases ($E_\gamma^{truth} \to E_\gamma^{reco}$ and $\overrightarrow{e}_\gamma^{truth} \to \overrightarrow{e}_\gamma^{reco}$) is undistinguishable from $E_\gamma^{truth} \to E_\gamma^{reco}$ case.

Most of the width of the $M_{\rm bc}$ distribution is due to the beam energy spread, not to detector effects. But among the latter, the leading effect is the reconstruction of the photon energy. In contrast the shape of the ΔE and $m_{K^+\pi^+\pi^-\gamma}$ is almost completely due to the reconstruction of the photon energy.

[1]The cut on $M_{K\pi\pi} < 1.8$ GeV/c^2 has 100% efficiency for signal because the signal Monte-Carlo events are produced in this range.

The invariant mass is considered along with the usual variables $M_{\rm bc}$ and ΔE, because at truth level the distribution of $m_{K^+\pi^+\pi^-\gamma}$ is a Dirac δ function, not affected by the beam energy spread.

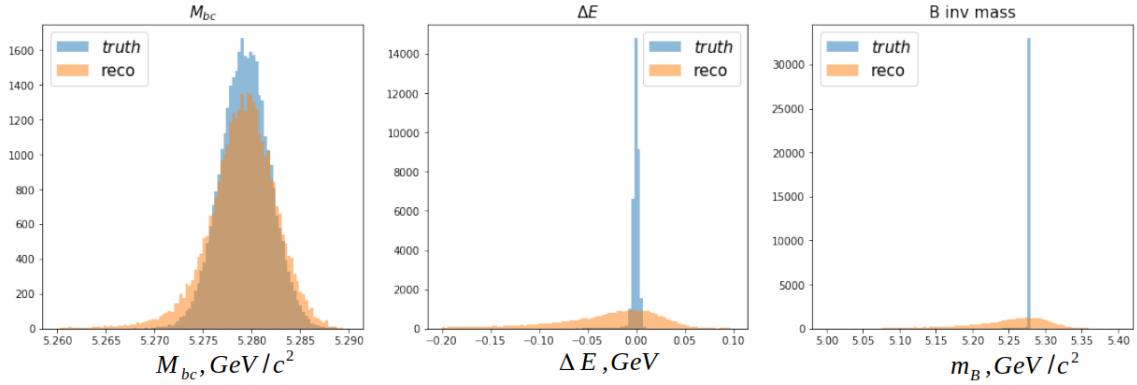


Figure 4.18. Distributions of $M_{\rm bc}$, ΔE, and $m_{K^+\pi^+\pi^-\gamma}$ for TM candidates at reconstruction level and at truth-level.

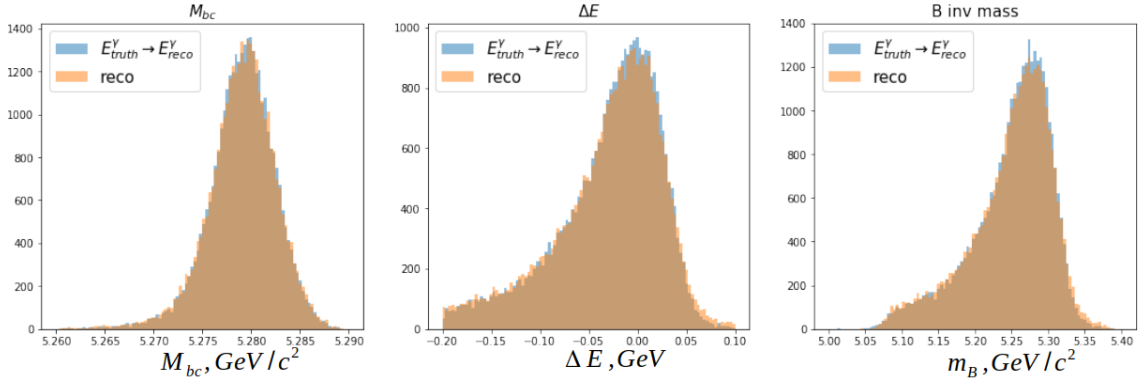


Figure 4.19. Distributions of $M_{\rm bc}$, ΔE, and $m_{K^+\pi^+\pi^-\gamma}$ for TM candidates at reconstruction level and at almost truth-level: for the latter, the energy of the photon has been replaced by the reconstructed energy.

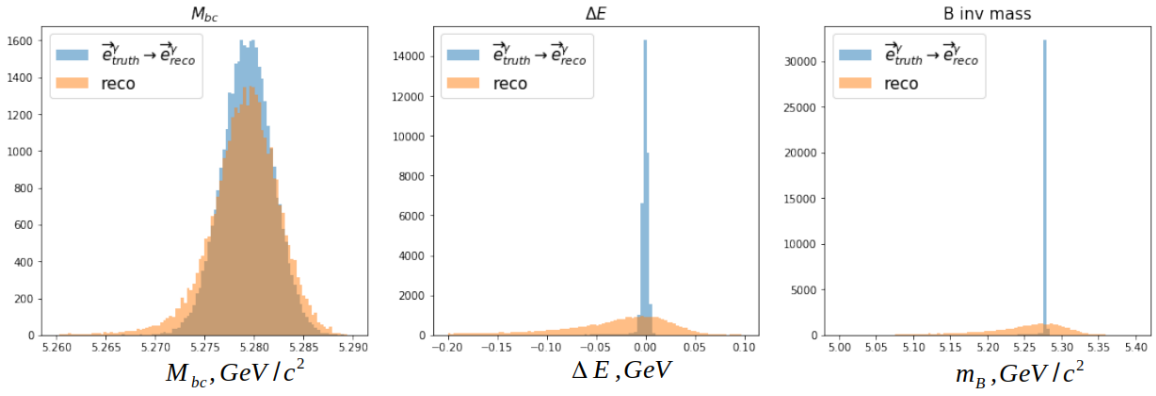


Figure 4.20. Distributions of $M_{\rm bc}$, ΔE, and $m_{K^+\pi^+\pi^-\gamma}$ for TM candidates at reconstruction level and at almost truth-level: for the latter, the direction of the photon has been replaced by the reconstructed direction.

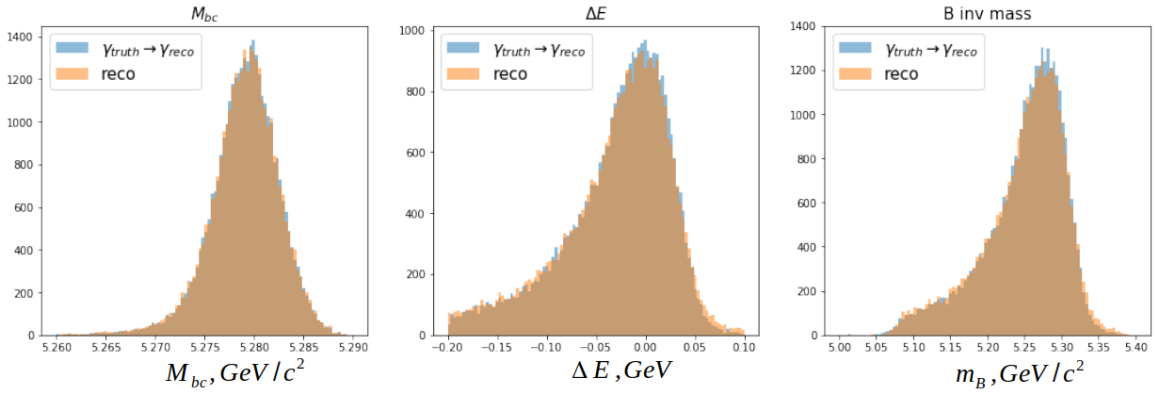


Figure 4.21. Distributions of $M_{\rm bc}$, ΔE, and $m_{K^+\pi^+\pi^-\gamma}$ for TM candidates at reconstruction level and at almost truth-level: for the latter, the energy-momentum of the photon has been replaced by the reconstructed energy-momentum.

Figs.(4.22)-(4.24) illustrate the detector effects by presenting, for TM events, the distribution of the photon energy at truth-level and at reconstruction-level, and their scatter-plot.

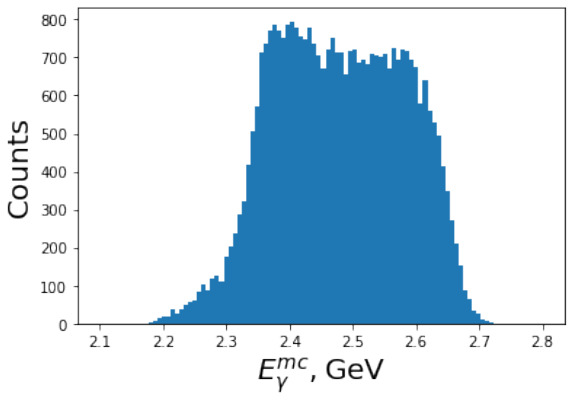


Figure 4.22. E_γ^{mc} distribution at selection level.

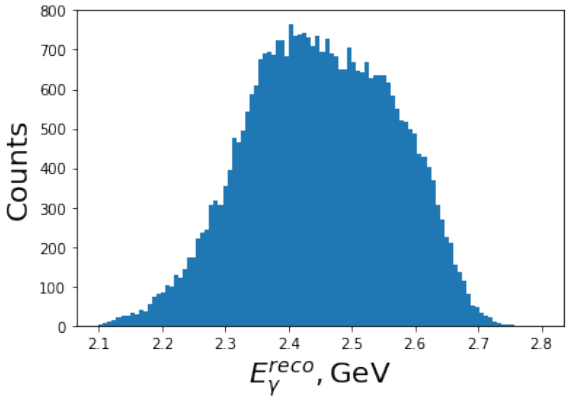


Figure 4.23. E_γ^{reco} distribution at selection level.

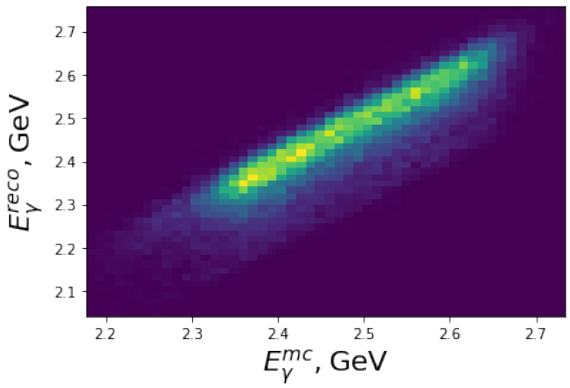


Figure 4.24. A systematic bias towards lower reconstructed energies is observed.

In order to check the behavior in the different parts of the calorimeter Fig.(4.25) present the resolution $\frac{E_\gamma^{mc}-E_\gamma^{reco}}{E_\gamma^{mc}}$, after selection cuts. The clear asymmetric tails of the distributions confirm that the reconstructed energy underestimates the true energy: part of the energy leak is not corrected. The shapes of the distributions for TM candidates of (forward and backward) endcaps and barrel are roughly similar.

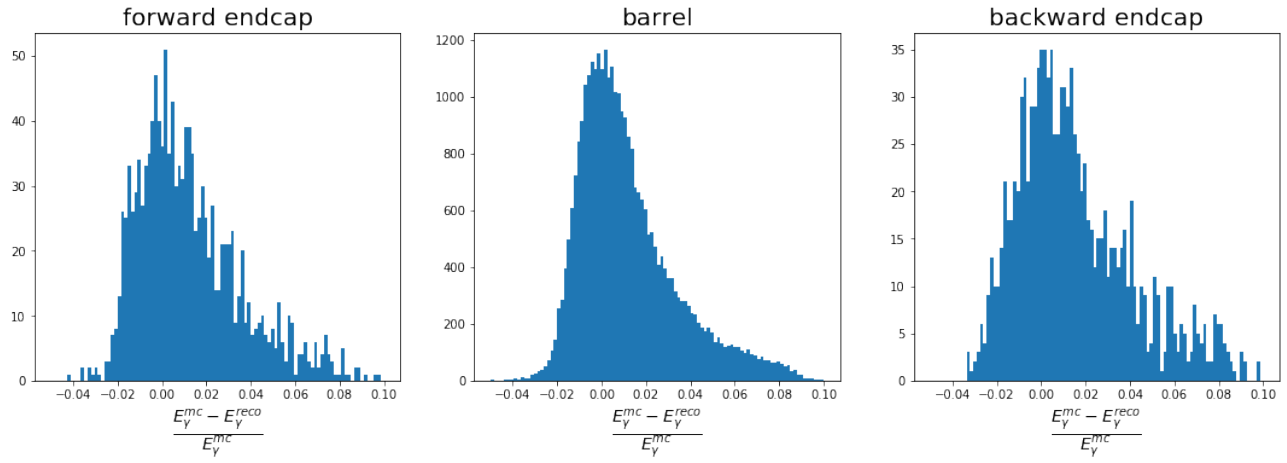


Figure 4.25. Resolution of the photon energy for different parts of electromagnetic calorimeter.

One concludes that percent level difference between truth and reconstruction energies of the final state photon can cause simultaneous changes in M_{bc} and ΔE distributions, which leads to correlation between these two quantities. Thus, joint usage both variables within extended likelihood fit for yields measurement seems to be unfeasible and only one variable can be considered.

4.2.6 . Expected signal and background contributions at selection level

Applying all the above cuts to the sample of background and signal events of $1\ \mathrm{ab}^{-1}$ one obtains 12600 signals events and 331614 background events. The composition of the sample is given in the table below. In effect, as shown below, $\tau\bar{\tau}$ pairs could have been neglected.

event type	$K\pi\pi\gamma$	TM	~~TM~~	$q\bar{q}$	B^+B^-	$B^0\bar{B}^0$	$\tau\bar{\tau}$
N_{event}	12600	11110	1490	316046	4489	10922	157
%	3.7%	3.2%	0.4%	91.8%	1.3%	3.2%	0.05%
$\langle N_{\mathrm{can}} \rangle$	1.35	1.35	1.30	1.24	1.29	1.31	1.06

Table 4.3. Composition of events at selection level: $K\pi\pi\gamma$ are signal events split as the sum of TM (events with a TM candidate) and ~~TM~~ (events with no TM candidate). For each type of events are quoted the number of events (N_{event}), the fraction they represent over the whole sample, and the average number of candidates they contain ($\langle N_{\mathrm{can}} \rangle$).

The table (4.4) below corresponds to the signal region $M_{\mathrm{bc}} \in [5.27, 5.29]\ \mathrm{GeV/c^2}$ with 11583 signals events and 68948 background events. The dominant background is due to continuum $q\bar{q}$ pairs.

event type	$K\pi\pi\gamma$	TM	~~TM~~	$q\bar{q}$	B^+B^-	$B^0\bar{B}^0$	$\tau\bar{\tau}$
N_{event}	11583	10928	655	63134	1452	4322	40
%	14.4%	13.6%	0.8%	78.4%	1.8%	5.4%	0.05%
$\langle N_{\text{can}} \rangle$	1.14	1.14	1.17	1.11	1.15	1.17	1.08

Table 4.4. Composition of events in the signal region $M_{\text{bc}} \in [5.27, 5.29]$ GeV/c^2 (cf. table (4.3)).

At this point of the selection, one may expect to already be able to observe a peak hinting for signal as illustrated on Fig.(4.26).

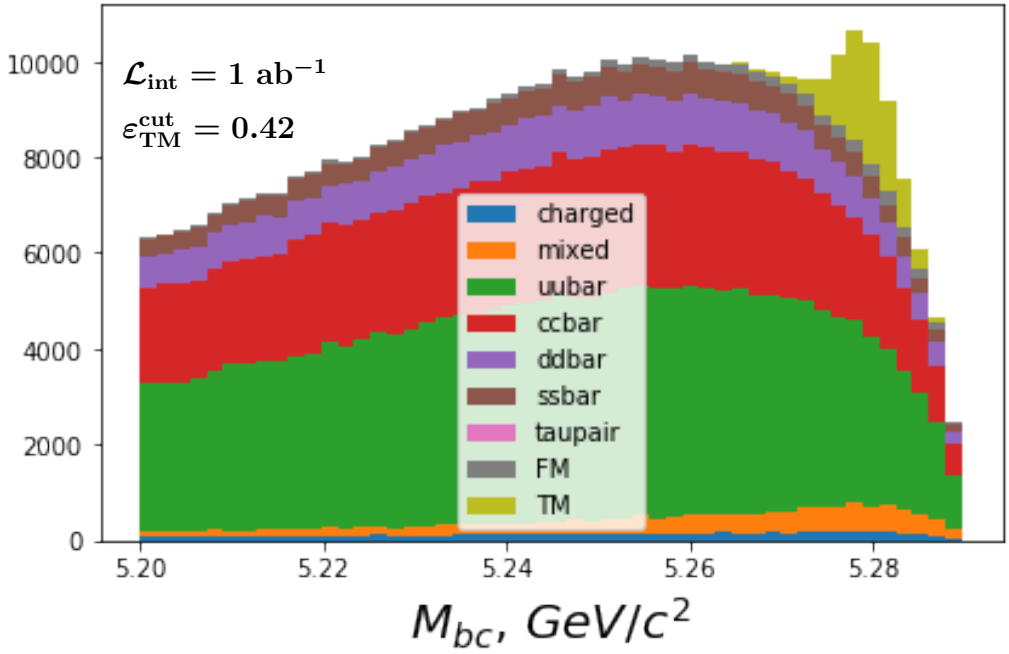


Figure 4.26. The distribution of M_{bc} for the candidates of the sample, at selection level. The aliases charged and mixed correpond to B^+B^- and $B^0\bar{B}^0$ events respectively.

4.3 . Background reduction

As seen above (cf. Table(4.4)), the dominant background comes from continuum $q\bar{q}$ pairs production. A dedicated artificial Neural Net (NN) is used to reduce its level, and in passing the level of other sources of background, including FM candidates. The model is depicted on Fig. 4.27.

Due to a large amount of training data, it is not feasible to use all dataset with size N at once for the training of neural nets. Such models are trained iteratively: on each iteration, a sample of training examples of size m (mini-batch of data with $m \ll N$) is used to update the weight for each layer using mini-batch gradient descent [133].

Each dense layer l, taking as input vector $\overrightarrow{x}^{l-1}$ of length N^{l-1}, has N^l neurons and is parametrized by matrix $\hat{W}^l$ of $N^{l-1} \times N^l$ adjustable weights. Such layer also contains non-linear function $F_l(\cdot)$ applied element-wise and producing output vector $\overrightarrow{x}^l$ (see Eq. 4.13).

$$x_i^l = \sum_{j=1}^{N^{l-1}} F^l(w_{ij}^l \cdot x_j^{l-1}) \tag{4.13}$$

$$F_l(x) = \begin{cases} x & x \geq 0 \\ 0 & x < 0 \end{cases}, \quad \forall l \in 1; L-1 \tag{4.14}$$

$$F_L(x) = \frac{1}{1+e^{-x}} \tag{4.15}$$

where w_{ij}^l — weights of matrix $\hat{W}^l$; Eq. 4.14 models non-linearities of intermediate layers and represents ReLU activation function [126]; Eq. 4.15 models output of neural net as a probability that input is a signal and represents sigmoid activation function [127]; L — index of last layer on neural net.

In between dense layers, there are batch normalization layers. They are applied to avoid internal covariate shift [125]. Such problem appears if, for instance, neuron previously receiving input distributed normally as $\mathcal{N}(0,1)$ after several iterations of weights $\hat{W}^l$ update starts receiving inputs distributed rather differently $\mathcal{N}(1,1)$. It is an essential process in the training of neural networks. In this case, a lot of information a given neuron learned so far, is not needed anymore making the training not efficient. Batch normalization solves this by introducing the mapping as in Eq. 4.16.

$$\hat{x}_i = \gamma \frac{x_i - \mu_{\mathcal{B}}}{\sqrt{\sigma_{\mathcal{B}} + \varepsilon}} + \beta \tag{4.16}$$

$$\tag{4.17}$$

where $\mu_{\mathcal{B}}$ — mean of x_i among mini batch of training examples for given neuron; $\sigma_{\mathcal{B}}$ — standard deviation of x_i within a mini batch for given neuron; γ, β — parameters adjusted during the training of neural net for given neuron;

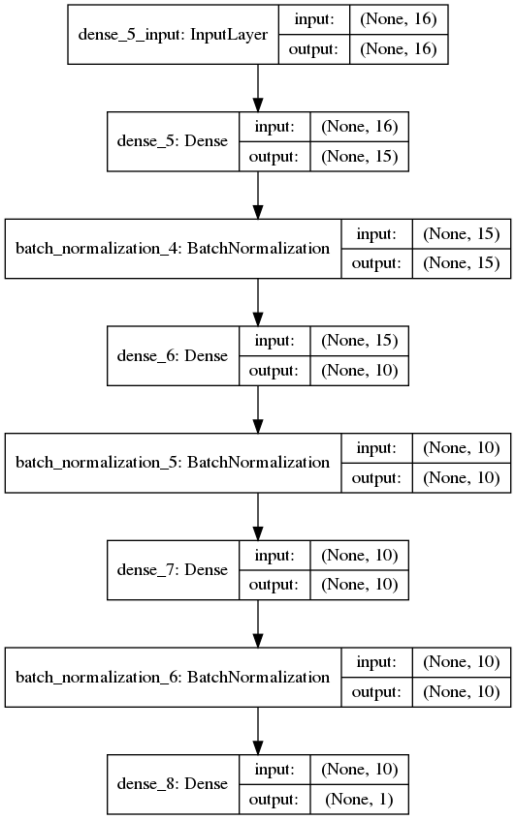


Figure 4.27. The architecture of artificial neural net used for continuum suppression. Each layer has input and output and two types of layers are considered: dense and batch normalization. Output dense layer classify an input as signal or background.

Most of the NN inputs are event shape variables referring implicitly to the candidates : thus, they are different for each candidate within the event. Taking into account that this analysis is untagged (only the signal B meson is reconstructed), the Rest Of Event (ROE) is built around a signal B and it is a collection of all particles coming from the accompanying B meson. Thus, depending on the considered candidate for the event, ROE is changed accordingly representing a different set of particles for each signal B-meson candidate.

- Thrust of set of particles is defined as an axis $\overrightarrow{t}$ for which sum of projections of these particles on the axis is maximized:

$$\hat{T} = \max\Big(\frac{\sum_i \overrightarrow{t} \cdot \overrightarrow{p}_i}{\sum_i |\overrightarrow{p}_i|}\Big) \tag{4.18}$$

  where $\overrightarrow{p}_i$ — 3-momentum of final state particle within a set. Using Eq. 4.18, the variables derived from thrust can be defined:

  - T_{ROE}, the thrust of the "rest of the event" (ROE) obtained by summing over ROE particles for given B-meson candidate in Eq. 4.18;
  - $\cos T_B z = |\overrightarrow{t}_B \cdot \overrightarrow{e}_z|$ — the cosine of angle between the $\overrightarrow{t}_B$ direction and the z-axis;

– $\cos T_B T_{ROE} = |\overrightarrow{t}_B \cdot \overrightarrow{t}_{ROE}|$ — the cosine between the $\overrightarrow{t}_B$ and $\overrightarrow{t}_{ROE}$ directions;

For instance, for $q\bar{q}$ event with two jets as in Fig. 4.28 $K\pi\pi\gamma$ final state is selected randomly, but the thrust axis $\overrightarrow{t}_B$ is correlated with $\overrightarrow{t}_{ROE}$. For $B\bar{B}$ events thrust of signal and accompanying B-meson are uncorrelated. The effect can be illustrated on Fig. 4.3 where distribution of $\cos T_B T_{ROE}$ is depicted. For $q\bar{q}$ events the distributions is peaked towards zero signifying about correlations, meanwhile for signal $B\bar{B}$ events it is almost flat.

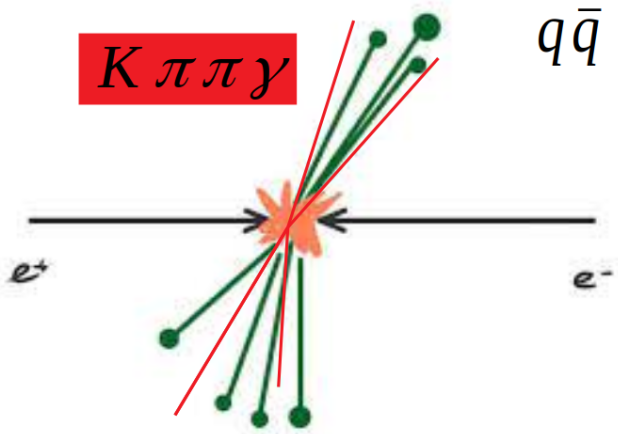


Figure 4.28. $q\bar{q}$ event is represented by two jets: even though $K\pi\pi\gamma$ final state is selected randomly, their thrust is correlated with thrust of ROE.

- Kakuno Super Fox-Wolfram moments [130] — continuum suppression variables developped by Belle Collaboration. They provide higher grained information about event than Fox-Wolfram moments used in Section 4.2. The moments are calculated using: reconstracted signal and ROE separately; charged, neutral and missing particles; even and odd degrees of Legendre polynomial. The moments H^{so}_{xl} are represented by the following equations:

$$H^{so}_{xl} = \begin{cases} \sum_{i,jx} |\overrightarrow{p}_{jx}| P_l(\cos\theta_{i,jx}) & l = 0,2,4 \\ \sum_{i,jx} Q_i Q_{jx} |\overrightarrow{p}_{jx}| P_l(\cos\theta_{i,jx}) & l = 1,3 \end{cases} \tag{4.19}$$

where i — index running over B daughters; jx runs over the ROE in the category x; x — charged ($x = 0$), neutral ($x = 1$) or missing ($x = 2$) group of particles; $\overrightarrow{p}_{jx}$ — momentum of particle jx; $P_l(\cdot)$ — l-th order Legendre polynomial; $\theta_{i,jx}$ — angle between i-th and jx-th particle.

Considering only ROE, corresponding moments H^{oo}_l are computed as follows:

$$H^{oo}_l = \begin{cases} \sum_{j,k} |\overrightarrow{p}_j||\overrightarrow{p}_k| P_l(\cos\theta_{j,k}) & l = 0,2,4 \\ \sum_{j,k} Q_j Q_k |\overrightarrow{p}_j||\overrightarrow{p}_k| P_l(\cos\theta_{j,k}) & l = 1,3 \end{cases} \tag{4.20}$$

where j, k run over ROE, Q_j, Q_k — charges of j-th and k-th particle;

The H^{so}_{xl} and H^{oo}_l moments are normalized to H^{max}_0 and $(H^0_{max})^2$, respectively, where $H^0_{max} = 2(E^*_{beam} - \Delta E)$ to avoid dependence on ΔE. There are 16 KSFW moments: $H^{so}_{\{0,1,2\}\times\{0,2,4\}}$ (9), $H^{so}_{\{0\}\times\{1,3\}}$ (2), $H^{oo}_{\{0-4\}}$ (5) calculated internally within *basf2* software [131].

- Sphericity of event is defined from eigenvalues of the symmetric matrix created from outer vector products of 3-momenta of final state particles:

$$M_{xyz} = \sum_i \overrightarrow{p}_i \otimes \overrightarrow{p}_i \tag{4.21}$$

$$\lambda_1, \lambda_2, \lambda_3 \leftarrow M_{xyz}\lambda = \mathbb{I}\lambda \tag{4.22}$$

$$S = \frac{3}{2}(\lambda_2 + \lambda_3) \tag{4.23}$$

  where $\overrightarrow{p}_i$ — 3-momenta of final state particle; $\mathbb{I}$ — 3×3 identity matrix; λ — eigen value of matrix M_{xyz}. The allowed range of S values is $0 \leq S < 1$ and typically is higher for $B\bar{B}$ than $q\bar{q}$ events.

- Since B-meson has relatively higher life-time relatively to strongly decaying lighter mesons, it has a longer flight distance. The quantity $\Delta z = z_{B_{sig}} - z_{B_{ROE}}$ is a distance along beam direction between vertices of signal and other B decays and can be used to separate continuum and $B\bar{B}$ events.

- $R_2 = \frac{H_2}{H_0}$ variable is built out of Fox-Wolfram moments defined in Eq. 4.7. It indicates about jettiness of an event shape;

- Often, on the late stage of analysis ΔE and M_{bc} are jointly used as discriminating variables in extended likelihood fit for yields extraction. In the present work ΔE is considered for neural net input in order to improve its discriminating power against backgrounds and in particular against FM candidates. Being present in the NN, ΔE will not be used in the yield measurements through a likelihood fit. In addition, from (4.2.5) M_{bc} and ΔE are correlated due to the misreconstructed photon energy, which complicates modelling of the joint distribution [93].

- B flavour tagging variable is defined as $q \cdot r$, where q is the flavor of the other B, while r is an expected flavor dilution factor ranging from zero for no flavor to unity for non-ambiguity flavor assignment. For continuum events, $|q \cdot r|$ peaks at 0, and for $B\bar{B}$ events is expected to peak at 1.

The discriminating power of the above variables between $q\bar{q}$ candidates and TM candidates is illustrated in Figs.(4.29)-(4.35). In the figure captions are quoted the separation power of the corresponding signal ($S(x)$) and background ($B(x)$) p.d.f. 's.

$$\langle S^2 \rangle = \frac{1}{2} \int \frac{(S(x) - B(x))^2}{S(x) + B(x)} \, \mathrm{d}x \tag{4.24}$$

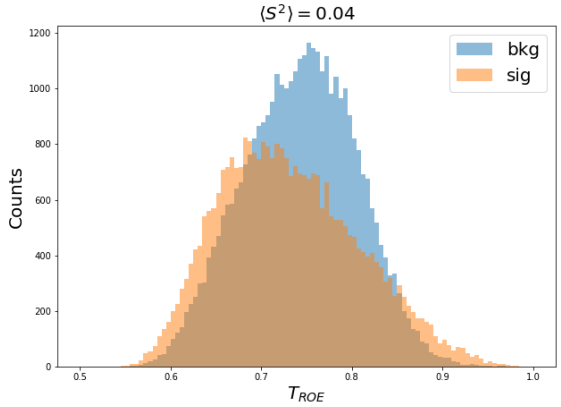


Figure 4.29. Thrust variable for the ROE for signal and $q\bar{q}$ events

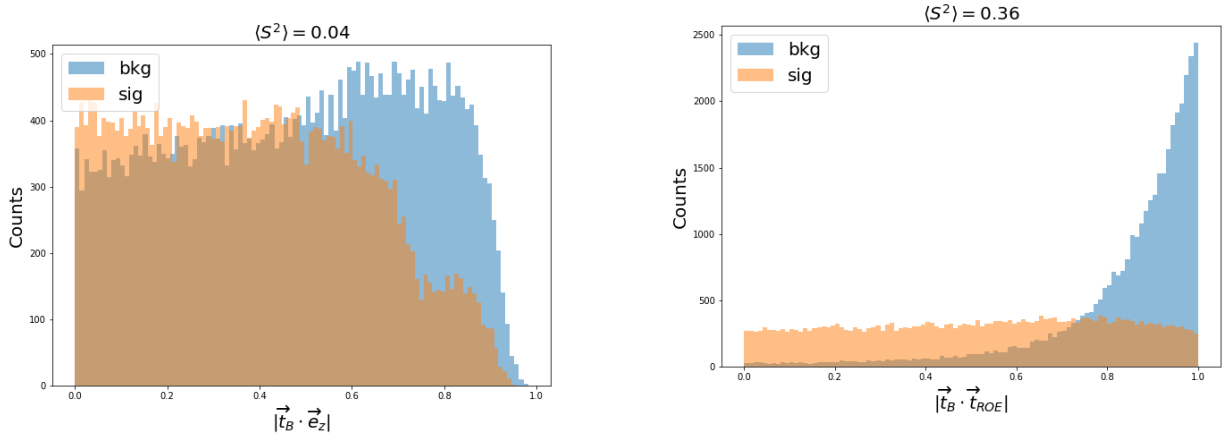


Figure 4.30. Cosines of angle between the thrust axis of the candidates and the z-axis (left) and between the thrust axis of the candidates and the ROE thrust axis (right) for signal and $q\bar{q}$ events

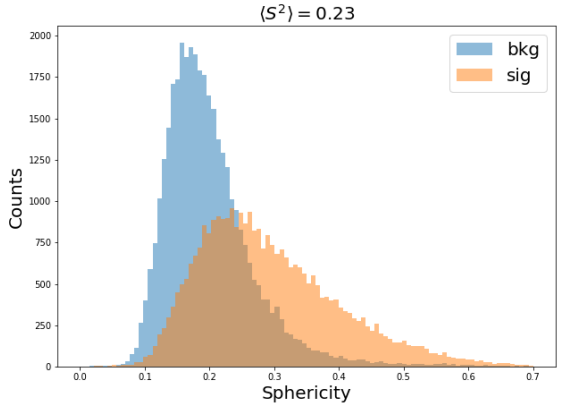


Figure 4.31. Sphericity distribution for signal and $q\bar{q}$ events.

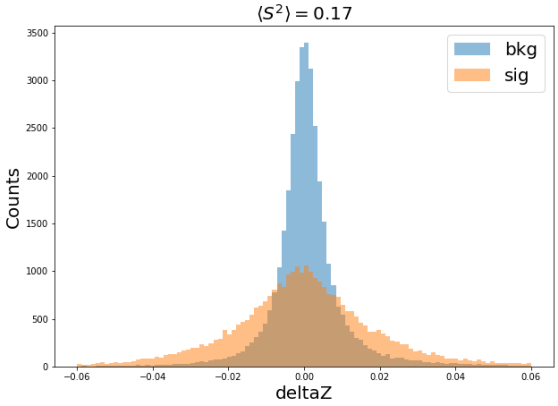


Figure 4.32. ΔZ for signal and $q\bar{q}$ events

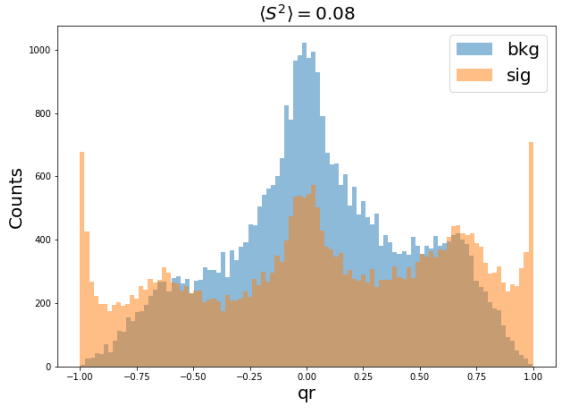


Figure 4.33. qr tagging information based on FastBDT [] for signal and $q\bar{q}$ events

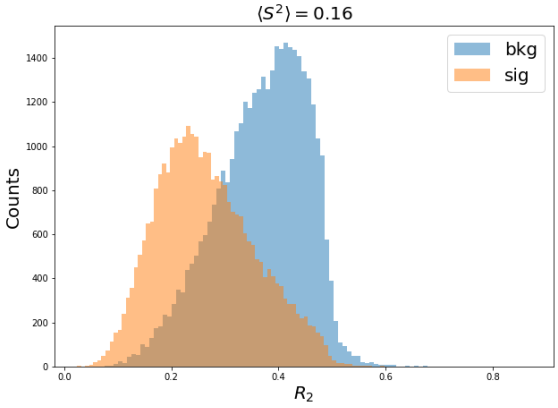


Figure 4.34. R_2 for signal and $q\bar{q}$ events

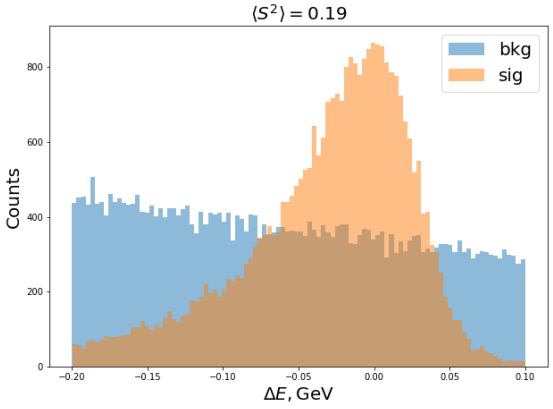


Figure 4.35. ΔE for signal and $q\bar{q}$ events.

On practice though not all quantities introduced above are used as a neural net input. Only the most important, sufficiently separated and less correlated variables are considered: $\cos T_B T_{ROE}$, ΔE, R_2, ΔZ, H_{12}^{so}, Sphericity, $\cos T_B z$, H_{02}^{so}, H_{10}^{so}, $q \cdot r$, T_{ROE}, H_0^{oo}, H_{04}^{so}, H_{14}^{so}, H_{22}^{so}, H_{20}^{so}. Corresponding procedure is given in Appendix5.

The training of the Neural Net[2] defined in Fig. 4.27 is based on two samples, with an equal number of events:

- A sample of 309943 $q\bar{q}$ events with $\langle N_{\rm can} \rangle = 1.22$,
- A sub-sample of 309943 TM candidates (the FM candidates being ignored[3]).

The above two samples were both split into training sample (64%), validation sample (16%) and testing sample (20%) following recommendation of [146]. The training is done using Adam adaptive gradient descent algorithm [132] with learning rate 0.001 and binary cross-entropy loss defined as following:

$$\text{Log loss} = -\frac{1}{m}\sum_{i=1}^{m}(y_i \log p_i + (1-y_i)\log(1-p_i)) \tag{4.25}$$

where y_i — ground-truth value of given set of input variables within a batch (1 for TM candidate and 0 for $q\bar{q}$ event); p_i — neural net output for given input; m — number of training examples within a batch. Training is stopped once Log loss doesn't decrease after 5 passes through the data set.

The histograms of the NN output for the training part and for the testing part are shown in Fig.(4.36): they appear to be in very good agreement with no hint for over-or under-training, which is confirmed in Fig.(4.37).

[2]The NN is developed within the Keras framework.

[3]The option of merging the FM candidates with the $\Sigma_{q\bar{q}}$ candidates was considered as an overkill complication.

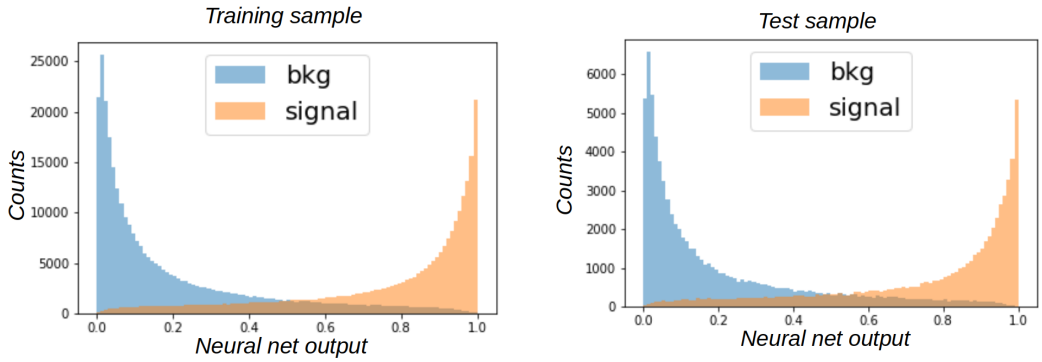


Figure 4.36. Distributions of NN, the output of the trained network for the training sample (left) and the testing sample (right).

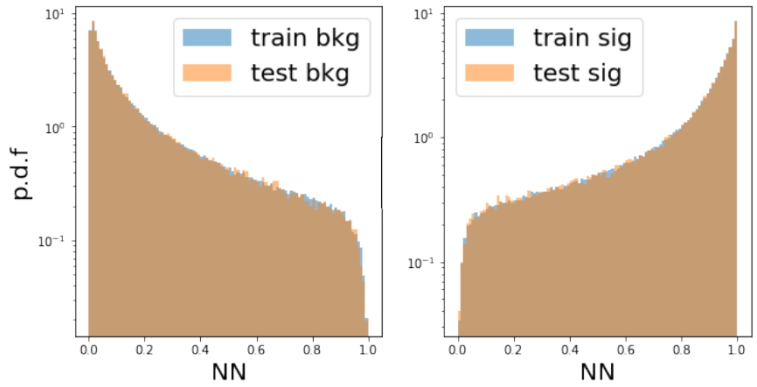


Figure 4.37. Comparison of distributions of NN outputs for the training and testing samples, for the background (left) and TM events (right). Here, the distributions are normalized to unity.

4.3.1 . Best candidate selection using Neural Net

The candidate within event being any combination of $K^+\pi^+\pi^-\gamma$ final state particles which pass given selection criteria (see Table 4.1) is further compared to other candidates. The criteria aiming to select the candidate closest to the produced generated event is called best candidate selection (BCS). After passing detector and reconstruction software steps, generated Monte-Carlo event is translated to TM candidate if generator level information is matched with reconstructed one. Thus BCS is done to get the candidate closest to the TM candidate.

The NN is trained using TM candidates as a signal sample. It provides a means to reduce $q\bar{q}$ background and also to perform BCS by defining the "best candidate" as the one with the largest NN output. The higher output corresponds to a higher probability that given input values of NN correspond to TM candidate. For signal events, if the best candidate is truth matched, the event is termed TM, while if the BCS is FM, the event is termed SCF, for self crossfeed.

For instance, consider three signal events: $E_1 : (TM, FM)$, $E_2 : (TM, FM)$, $E_3 : (FM, FM)$. Suppose that each event has two candidates. First two events are TM-events and have both truth-matched

and false-matched candidate, while third one is FM-event and has only false-matched candidates. Consider that after BCS the events are as following: $E_1 : (TM)$, $E_2 : (FM)$, $E_3 : (FM)$. Then E_1 remains TM-event, while E_2, E_3 are SCF, even though before BCS E_2 was TM-event.
Several options can be studied to handle the issue of multiple candidates:

- (All): no BCS is applied and several entries per event are possible. The fit becomes very involved and difficult to validate with control samples.
- ($N_{\text{can}} = 1$) : Keep only events with a unique candidate (Signal events fraction kept 74%). This is the simplest approach, which makes sense since the fraction of events with a unique candidate is large (see Fig. 4.38).

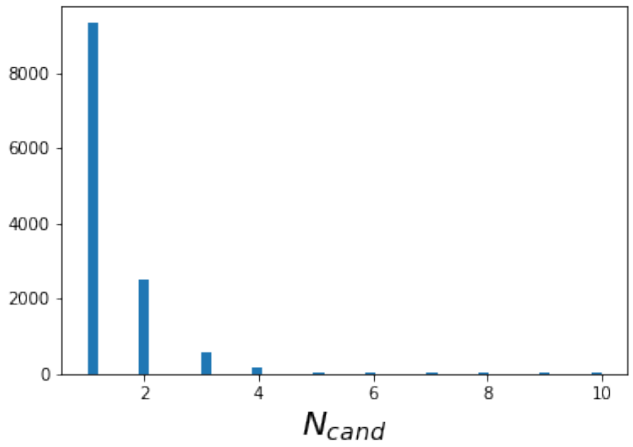


Figure 4.38. Distribution of number of candidates per signal event before BCS.

- (pValue) : Define the best candidate as the one with the highest probability of $K^+\pi^+\pi^-\gamma$ vertex fit.
- (random): Select randomly a single candidate within the event.
- (NN) : Define the best candidate as the one with the largest NN output.

The table 4.5 summarizes the features of the above options for simulated signal events. The sample consists of the 12600 signal events remaining at selection level (cf. Table(4.3)) comprising 16966 candidates, 11117 of which are TM and 5849 are FM (i.e. 66% of the candidates are TM). One observes that the NN-based BCS yields the largest selection efficiency for TM (93%). Its purity with respect to SCF (82%) is close to the largest purity (88%) provided by the $N_{\text{can}} = 1$-based BCS, but the latter entails a significantly lower TM efficiency.
As indicated in the lower part of the table below, this conclusion holds as well in the signal region, wherein addition the NN-based BCS purity is even closer to the $N_{\text{can}} = 1$-based BCS.
Accordingly, it is decided to choose the NN-based BCS definition.

	All	$N_{\rm can}=1$	pValue	random	NN
Signal events kept	100%	74%	100%	100%	100%
% of TM kept	100%	73%	90%	85%	93%
TM/(TM+SCF)	66%	88%	79%	75%	82%
TM	11117	8164	10002	9481	10365
SCF	5849	1161	2598	3119	2235
% of TM kept	100%	73%	90%	85%	93%
TM/(TM+SCF)	83%	95%	90%	89%	92%
TM	10935	8008	9828	9317	10198
SCF	2249	454	1042	1200	905

Table 4.5. Effect of different BCS criteria. The first column refers to all candidates, without any BCS applied. The number of TM candidates (i.e. 11117) is larger than the number of TM events quoted in Table(4.3) (i.e. 11110). This is because the matching is imperfect and a handful of TM events have more than one TM candidate. In the following columns the results for the four BCS criteria, detailed in the text, are reported. In the upper (resp. lower) part of the table, events are in the range $5.20 < M_{\rm bc} < 5.29$ GeV/c^2 (resp. $5.27 < M_{\rm bc} < 5.29$ GeV/c^2).

The $M_{\rm bc}$ distribution of events after BCS is shown in Fig.(4.39). It is to be compared to the distribution of candidates shown in Fig.(4.26). The improvement is marginal because only BCS has been performed, the sample is still overwhelmed by $q\bar{q}$ events.

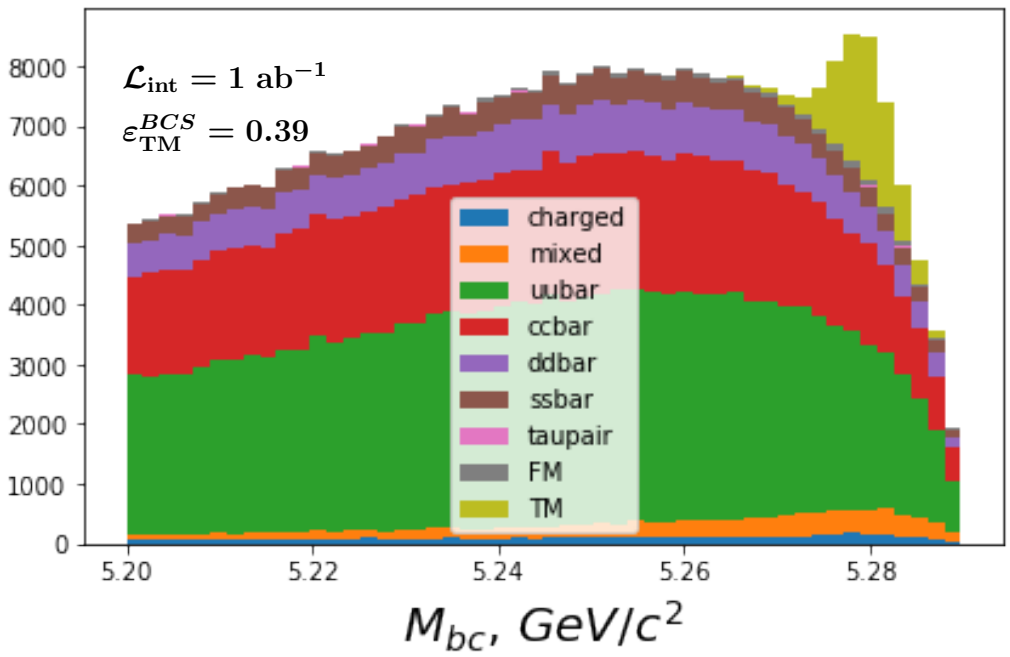


Figure 4.39. $M_{\rm bc}$ distribution after NN-based BCS.

4.3.2 . Choice of the Neural Net cut

After the BCS selection has been performed, a cut on the NN output is applied to drastically reduce the contamination from the $q\bar{q}$ background. The cut is not applied on the original NN outputs, but rather on their transformed version, where cut value can be interpreted as a fraction of discarded TM events.

μ (flatness)-transform

A change of variable is performed on the NN output to define a secondary variable μ_{NN} such that its distribution is uniform between 0 and 1. Although it does not improve the rejection power, this change of variable presents three useful features:

1) the ordering being the same for NN and μ_{NN}, BCS can be defined on μ_{NN} as well,
2) by construction the p.d.f. for the signal is known and easy to describe: it is just flat.
3) a cut on $\mu_{\mathrm{NN}} > \mu_{\mathrm{NN_{cut}}}$ implies a well defined selection efficiency $\epsilon = 1 - \mu_{\mathrm{NN_{cut}}}$.

The μ_{NN}-flatness transform with respect to the sample of events consists of the following steps:

- Considering only TM events, sort its neural net outputs is ascending order ($\hat{NN}_{sorted}$);
- Store the sorted array of neural net outputs $\hat{NN}_{sorted}$ along with its length L;
- Take the neural net output for an arbitrary event (TM, $q\bar{q}$, $B\bar{B}$, FM) and search in the sorted array for an index i of the closest element to the left.
- Transformed neural net output is defined as $\mu_{NN} = \begin{cases} \dfrac{i}{L}, & NN \geq \min(\hat{\mathrm{NN}}_{\mathrm{sorted}}) \\ NN, & NN < \min(\hat{\mathrm{NN}}_{\mathrm{sorted}}) \end{cases}$;

This is illustrated on Fig.(4.40) where the distributions of the original NN output and μ_{NN} are shown for the training samples. For instance, the cut $\mu_{\mathrm{NN}} > 0.8$ means that 80% of TM signal is discarded. The distributions of the individual components are indicated on Fig.(4.41).

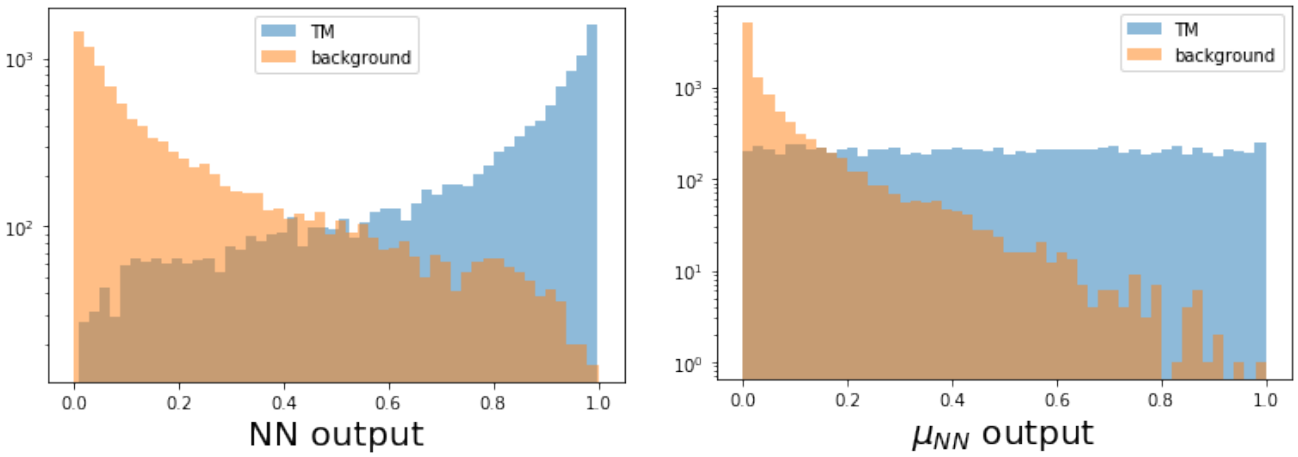


Figure 4.40. Distributions of the Neural Network output (left) and of its μ-transform μ_{NN} (right), both in log-scale.

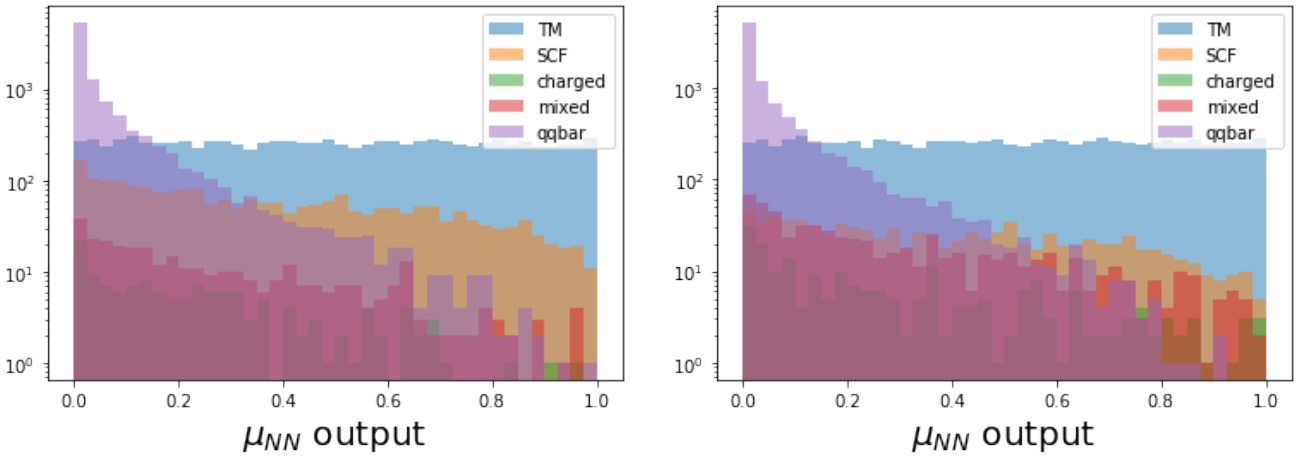


Figure 4.41. Distribution of μ_{NN} for TM, SCF, and the various backgrounds for $M_{\rm bc} \in [5.20, 5.29]$ GeV/c^2 (left) and in the signal region for $M_{\rm bc} \in [5.27, 5.29]$ GeV/c^2.

Optimal cut on μ_{NN}

In order to choose the optimal cut on $\mu_{\rm NN}$ the Monte-Carlo sample of 1 $\rm ab^{-1}$ was used.

The signal significance is used to express the cleanliness of the signal in the presence of statistical fluctuations of observed signal and background:

$$\mathrm{FOM}(\mu_{\rm NN}) = \frac{\int_{\mu_{\rm NN}}^{1} \frac{dN_{\rm TM}}{dx}(x|\mathcal{S})dx}{\sqrt{\int_{\mu_{\rm NN}}^{1} \left(\frac{dN_{\rm TM}}{dx}(x|\mathcal{S}) + \frac{dN_{\rm bkg}}{dx}(x|\mathcal{S})\right)dx}} = \frac{N_{\rm TM}(\mu_{\rm NN})}{\sqrt{N_{\rm TM}(\mu_{\rm NN}) + N_{\rm bkg}(\mu_{\rm NN})}} \tag{4.26}$$

where $\frac{dN_{\rm TM}}{dx}(x|\mathcal{S})$, $\frac{dN_{\rm bkg}}{dx}(x|\mathcal{S})$ — distributions of $\mu_{\rm NN}$ variable for signal and background in the signal region ($M_{\rm bc} \in \mathcal{S} = [5.27; 5.29]$ GeV/$\rm c^2$); considered background events are $q\bar{q}$, $B\bar{B}$ and SCF.

The optimal value of $\mu_{\rm NN}$ is obtained by maximizing signal significance as following:

$$\mu_{\rm NN}^{\rm cut} = \underset{\mu}{\mathrm{argmax}}\, \mathrm{FOM}(\mu_{\rm NN}) \tag{4.27}$$

This figure of merit is shown in Fig.(4.42) as function of $\mu_{\rm NN}^{\rm cut}$. The maximum is reached for $\mu_{\rm NN}^{\rm cut} \simeq 0.28$, which is, therefore, the value chosen to apply background suppression. Selection efficiencies of μ_{NN} cut are defined for signal and background as follows:

$$\varepsilon_{\mu_{\rm NN}}^{\rm evt} = \frac{\int_{\mu_{\rm NN}}^{1} \frac{dN_{\rm evt}}{dx}dx}{\int_{0}^{1} \frac{dN_{\rm evt}}{dx}dx} \tag{4.28}$$

where $\rm evt \in [TM, bkg = SCF + q\bar{q} + B\bar{B}]$; $\frac{dN_{\rm evt}}{dx}$ — μ_{NN} distributions for event species evt.

By construction, $\mu_{\rm NN}^{\rm cut}$ amounts to a selection efficiency for TM events $\varepsilon_{\mu_{\rm NN}}^{\rm TM} = 0.72$. Two receiver operating characteristic (ROC) curves are presented in Figs.(4.43)-(4.44). One observes on Fig.(4.44) that the background selection efficiency $\varepsilon_{\mu_{\rm NN}}^{\rm bkg}$ drops more than exponentially when reducing $\varepsilon_{\mu_{\rm NN}}^{\rm TM}$.

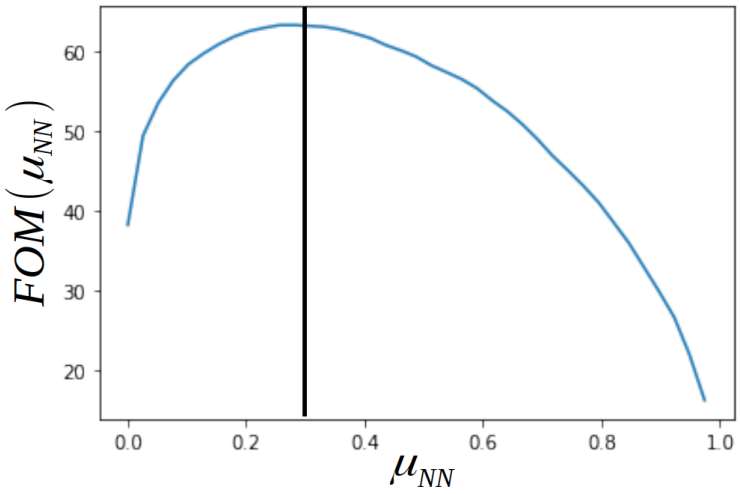


Figure 4.42. The figure of merit (Eq.(4.26)) as a function of the cut on $\mu_{\rm NN}$. The final selection cut is set as $\mu_{\rm NN} = 0.28$ (vertical line), which keeps 72% of the TM events.

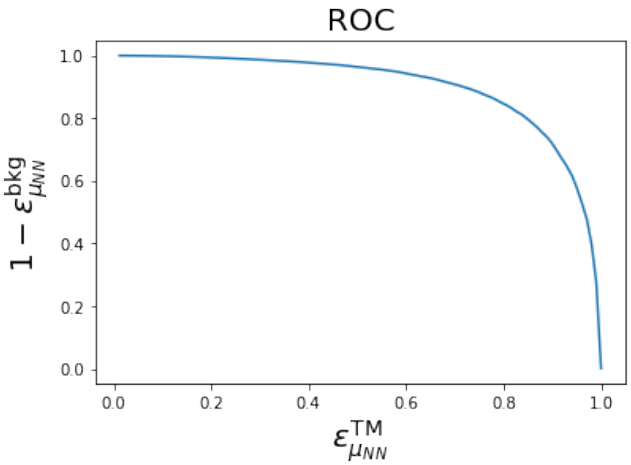


Figure 4.43. The fraction of rejected background events, $1-\varepsilon_{\mu_{\rm NN}}^{\rm bkg}$ (linear scale) as a function of the selection efficiency $\varepsilon_{\mu_{\rm NN}}^{\rm TM}$.

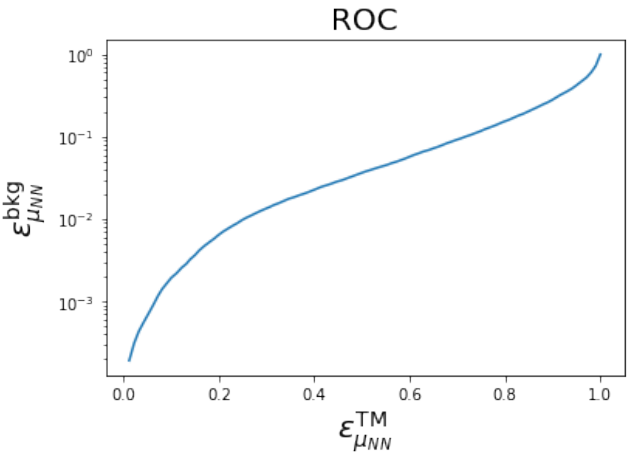


Figure 4.44. The selection efficiency background events, $\varepsilon_{\mu_{\rm NN}}^{\rm bkg}$ (log scale) as a function of the selection efficiency $\varepsilon_{\mu_{\rm NN}}^{\rm TM}$.

Because of the correlation between $\mu_{\rm NN}$ and $M_{\rm bc}$, one should not expect the selection efficiency to be constant as a function of $M_{\rm bc}$, neither for TM events, nor for background events. Efficiency as a function

of M_{bc} distributions before and after neural net cut for $\mathrm{evt} \in [\mathrm{TM}, \mathrm{SCF}, \mathrm{q\bar{q}}, \mathrm{B\bar{B}}]$ is defined as:

$$\frac{d\varepsilon^{\mathrm{evt}}_{\mu_{\mathrm{NN}}}}{dM_{bc}} = \frac{\frac{dN_{\mathrm{evt}}}{dM_{bc}}(M_{bc}|\mu > \mu^{\mathrm{cut}}_{\mathrm{NN}})}{\frac{dN_{\mathrm{evt}}}{dM_{bc}}(M_{bc}|\mu > 0)} \tag{4.29}$$

and illustrated in Figs(4.45)-(4.48). While for $q\bar{q}$ events efficiency is decreasing it is increasing for $B\bar{B}$ and SCF, which is undesirable.

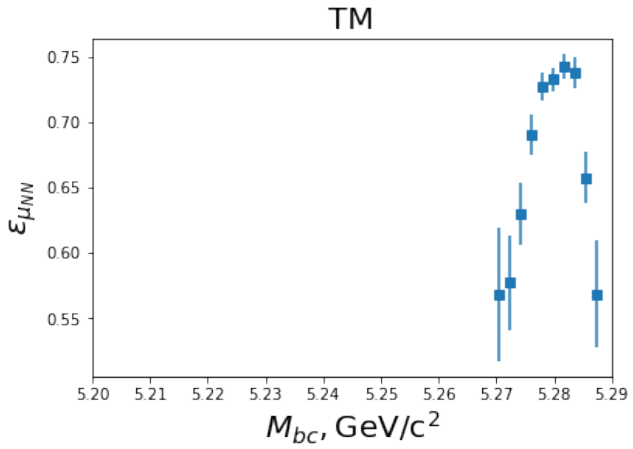


Figure 4.45. $\frac{d\varepsilon^{\mathrm{TM}}_{\mu_{\mathrm{NN}}}}{dM_{bc}}$ as a function of M_{bc}: the efficiency is significantly higher in the signal region $M_{\mathrm{bc}} \in [5.27, 5.29]$ GeV/c^2

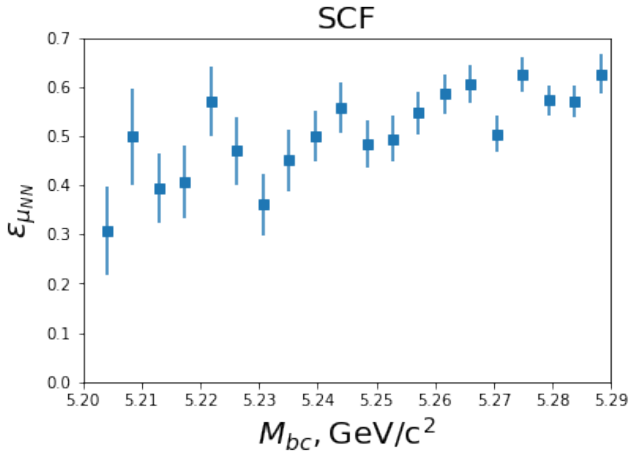


Figure 4.46. $\frac{d\varepsilon^{\mathrm{SCF}}_{\mu_{\mathrm{NN}}}}{dM_{bc}}$ as a function of M_{bc}: the efficiency is linearly increasing.

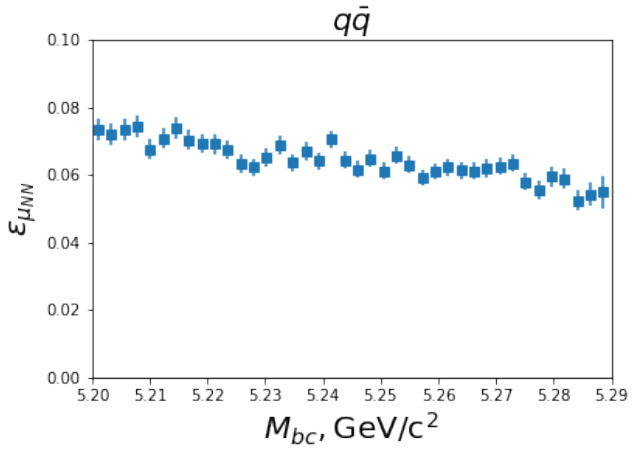


Figure 4.47. $\frac{d\varepsilon^{q\bar{q}}_{\mu_{\mathrm{NN}}}}{dM_{bc}}$ as a function of M_{bc}: the efficiency is linearly decreasing.

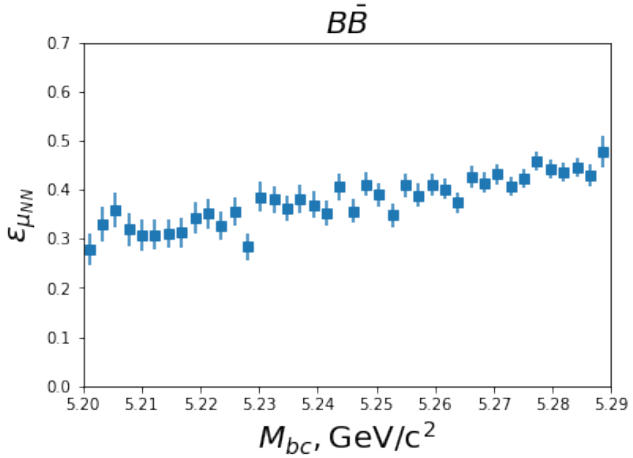


Figure 4.48. $\frac{d\varepsilon^{B\bar{B}}_{\mu_{\mathrm{NN}}}}{dM_{bc}}$ as a function of M_{bc}: the efficiency is linearly increasing.

From tables (4.1) and (4.5) the overall selection efficiency for TM events is

$$\varepsilon_{\rm TM}^{\rm tot} = 0.245 \simeq 0.42 \times 0.93 \times 0.72 \tag{4.30}$$

$$\sigma[\varepsilon_{\rm TM}^{\rm tot}] = \sqrt{\frac{\varepsilon_{\rm TM}^{\rm tot} \cdot (1 - \varepsilon_{\rm TM}^{\rm tot})}{N}} \simeq \sqrt{\frac{0.245 \times 0.755}{3 \cdot 10^4}} = 0.003 \tag{4.31}$$

where the first number corresponds to the overall selection efficiency, the second to the BCS, and the third to the NN cut. This total efficiency corresponds to 7350 ± 86 TM events.

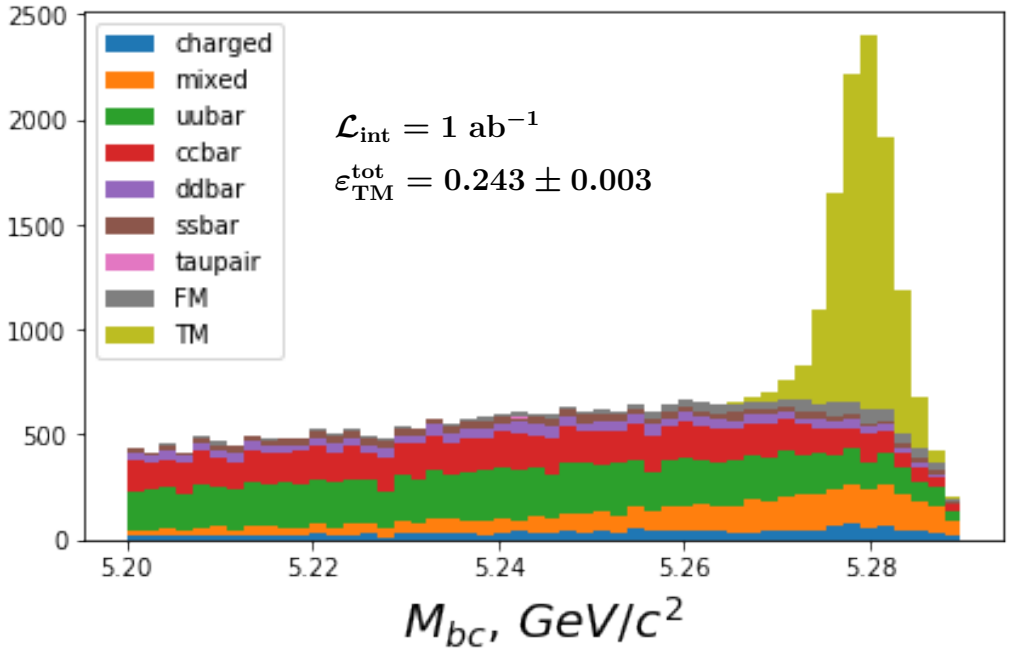


Figure 4.49. M_{bc} distribution corresponding to the optimal cut value $\mu_{\rm NN} = 0.28$. The TM component appears clearly above the background.

The distribution depicted in Fig.(4.49) should be compared to the one in Fig.(4.39), which was obtained after BCS, but before applying the cut at $\mu_{\rm NN}^{\rm cut} = 0.28$. One observes that the background level has been strongly suppressed by about an order of magnitude, and that its distribution is more uniform than before the $\mu_{\rm NN}$ cuts, reflecting the increase of $\varepsilon_{\mu_{\rm NN}}^{\rm bkg}$ with $M_{\rm bc}$.
To illustrate the sharp suppression of background with $\mu_{\rm NN}$, the sample of events is split into the two bins $\mu_{\rm NN} \in [0.28, 0.64]$ and $\mu_{\rm NN} \in [0.64, 1.0]$ which are almost equally populated with TM events. The corresponding distributions of $M_{\rm bc}$ are shown in Figs.(4.50)-(4.51). As expected, one observes that the background level in Fig.(4.51) is strongly reduced compared to the background level in Fig.(4.50).

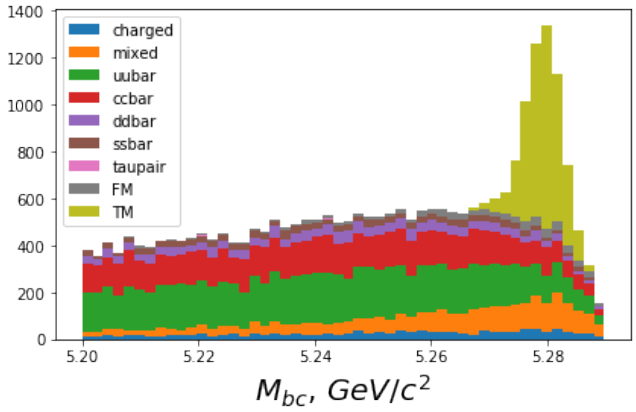


Figure 4.50. The M_{bc} distribution corresponding to $\mu_{\mathrm{NN}} \in [0.28, 0.64]$.

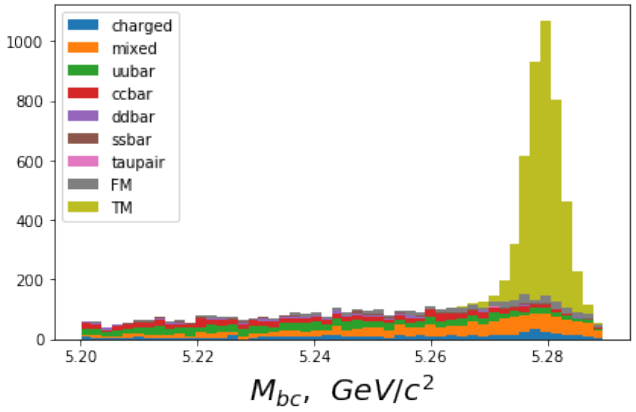


Figure 4.51. The M_{bc} distribution corresponding to $\mu_{\mathrm{NN}} \in [0.64, 1.0]$.

4.3.3 . Check on optimization of some selection cuts

As mentioned in section (4.2.5) the cuts applied at the selection level are not optimized, since there is no reliable figure of merit available at that stage. In particular, one may wonder if the cuts on $P(\gamma|\pi^0) < 0.8$ and $\chi_{\mathrm{prob}} > 0.0001$ are too tight, or too loose.
The Figs.(4.52)-(4.53) both show that the maximum of the figure of merit is not yet attained when one reaches the cut values: it follows that the cuts should not be tightened, but loosened.
However, as already stated in section (4.2.5), at this stage of the Belle II life, optimization of cuts should be taken with a grain of salt (μ_{NN} included) and it was not deemed necessary to pursue further on this path.

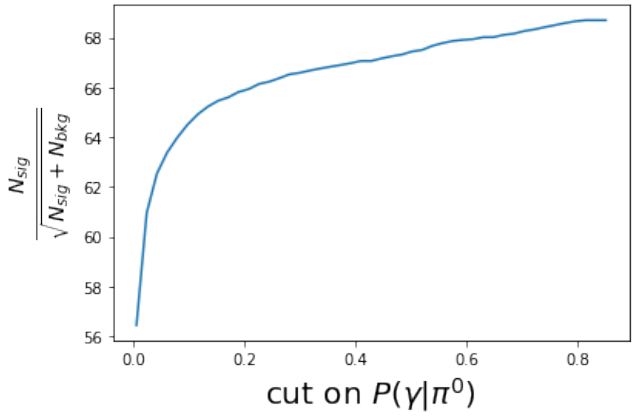


Figure 4.52. The figure of merit Eq.(4.26) as a function of the cut on $P(\gamma|\pi^0)$.

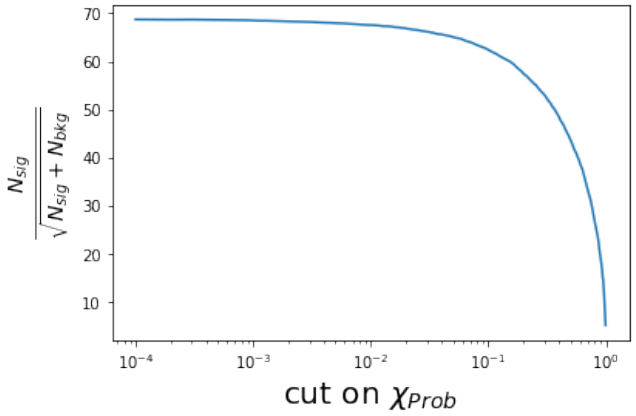


Figure 4.53. The figure of merit Eq.(4.26) as a function of the cut on χ_{prob}.

4.3.4 . Expected signal and background contributions

Applying the selection cuts, BCS and the $\mu_{\mathrm{NN}} > 0.28$ cut to the sample of background events and signal events one obtains a sample of events which composition is given in table (4.6), for $M_{\mathrm{bc}} \in$

$[5.20, 5.29]$ GeV/c^2. This sample of events is referred to as the Physics Sample in the following[4].

event type	$K\pi\pi\gamma$	TM	SCF	$q\bar{q}$	B^+B^-	$B^0\bar{B}^0$	$\tau\bar{\tau}$
N_{event}	8602	7386	1216	20381	1607	4459	16
%	24.5	21.1	3.5	58.1	4.6	12.7	0.05

Table 4.6. Composition of events with $M_{\text{bc}} \in [5.20, 5.29]$ GeV/c^2. At this stage, all events have a unique candidate. Signal events ($K\pi\pi\gamma$) are split as the sum of TM (the BCS is TM) and SCF (self crossfeed, events where the BCS is not TM). For each type of events, the number of events (N_{event}) and the fraction they represent over the whole sample of 8602 signal events plus 26463 background events are quoted.

The table (4.7) below corresponds to the signal region $M_{\text{bc}} \in [5.27, 5.29]$ GeV/c^2.

event type	$K\pi\pi\gamma$	TM	SCF	$q\bar{q}$	B^+B^-	$B^0\bar{B}^0$	$\tau\bar{\tau}$
N_{event}	7846	7309	537	3206	515	1764	3
%	58.8	54.8	4.0	24.0	3.9	13.2	0.02

Table 4.7. Composition of events in the signal region $M_{\text{bc}} \in [5.27, 5.29]$ GeV/c^2 (cf. table (4.6)).

4.3.5 . Peaking backgrounds

Peaking backgrounds are a major issue in this analysis since the final fit is performed on the M_{bc} distribution, where they tend to mimic the signal and can induce significant biases on the yield determination. Peaking backgrounds arise because of tracks lost or swapped from events with kinematics similar to the kinematics of signal events. While the reduction of $q\bar{q}$ background was given priority above, it is also essential to ensure that the peaking background in the signal region is under control. The SCF background is further discussed in section (4.4.2), where a dedicated control sample is presented.

Two sources of peaking backgrounds are discussed below.

Self crossfeed

The FM candidates selected as BCS in signal events (i.e. SCF) is one source of peaking background. For such FM candidates, typically a low momentum charged pion is not coming from the $B^+ \to K^+\pi^+\pi^-\gamma$ decay, but from the ROE (details are given in section (4.4.2)). As a result, the bulk of the kinematics of the FM candidate is able to roughly mimic a TM candidate, and in particular, the reconstructed M_{bc} will tend to populate the signal region $M_{\text{bc}} \in [5.27, 5.29]$ GeV/c^2. Fig.(4.54) shows the distribution of M_{bc}

[4]The detailed composition of the Physics Sample in the two bins $\mu_{\text{NN}} \in [0.28, 0.64]$ and $\mu_{\text{NN}} \in [0.64, 1.0]$ can be found in Appendix (5).

for SCF events, overlaid with TM events. About half of the SCF events are in the signal region.
As a matter of principle, the shape of the $M_{\rm bc}$ distribution of SCF events can be taken from Monte-Carlo, as well as its relative normalization with respect to TM events. However, as discussed in section (4.5) this is not needed.

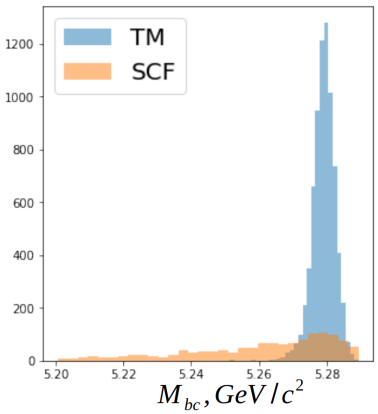


Figure 4.54. $M_{\rm bc}$ distributions for SCF and TM events.

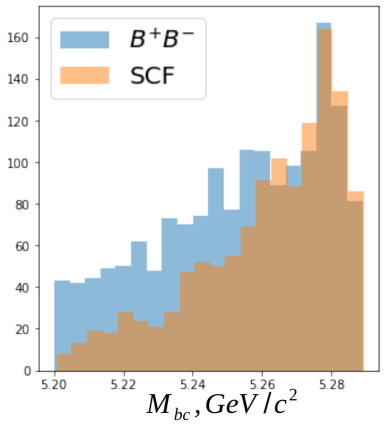


Figure 4.55. $M_{\rm bc}$ distributions for SCF and B^+B^- events.

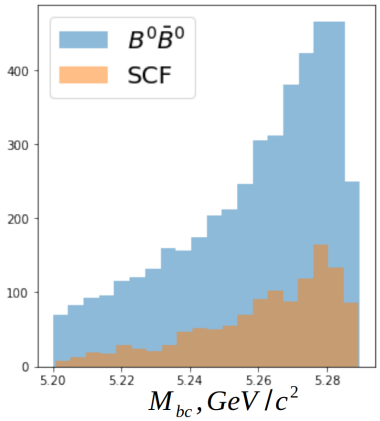


Figure 4.56. $M_{\rm bc}$ distributions for SCF and $B^0\bar{B}^0$ events.

$B\bar{B}$ background

Figs. 4.55-4.56 show the distribution of $M_{\rm bc}$ for B^+B^- and $B^0\bar{B}^0$ events overlaid with self-crossdfeed events. Both types of background have a peaking background component. One observes that while the size of SCF and B^+B^- distributions are comparable, the $B^0\bar{B}^0$ background is the leading background. Furthermore, its $M_{\rm bc}$ distribution appears quite similar to the SCF distribution. Peaking background events, akin to SCF, can appear from $B\bar{B}$ background (predominantly from $B^0\bar{B}^0$) due to final states close to $K^+\pi^+\pi^-\gamma$, when one (or more) soft particle is either added, replaced, or removed as for example:

1) $B^0 \to K^{*0}\gamma \to K^+\pi^-\gamma$, with a soft π^+ added from the ROE,

2) $B^0 \to K^0_{\rm res}\gamma \to K^+\pi^0\pi^-\gamma$, with a soft π^+ from the ROE replacing a soft π^0.

If such clearly identified decays were accounting for most of the $B\bar{B}$ peaking background, it could be attempted to veto them. However, such a veto would induce a phase-space dependent selection efficiency, which one should avoid in the photon polarization analysis.

Detailed sampling of the $B\bar{B}$ background sources corresponding to $\approx 300\ {\rm ab}^{-1}$ is provided by Fig.(4.57) where the sources are ranked according to their importance.

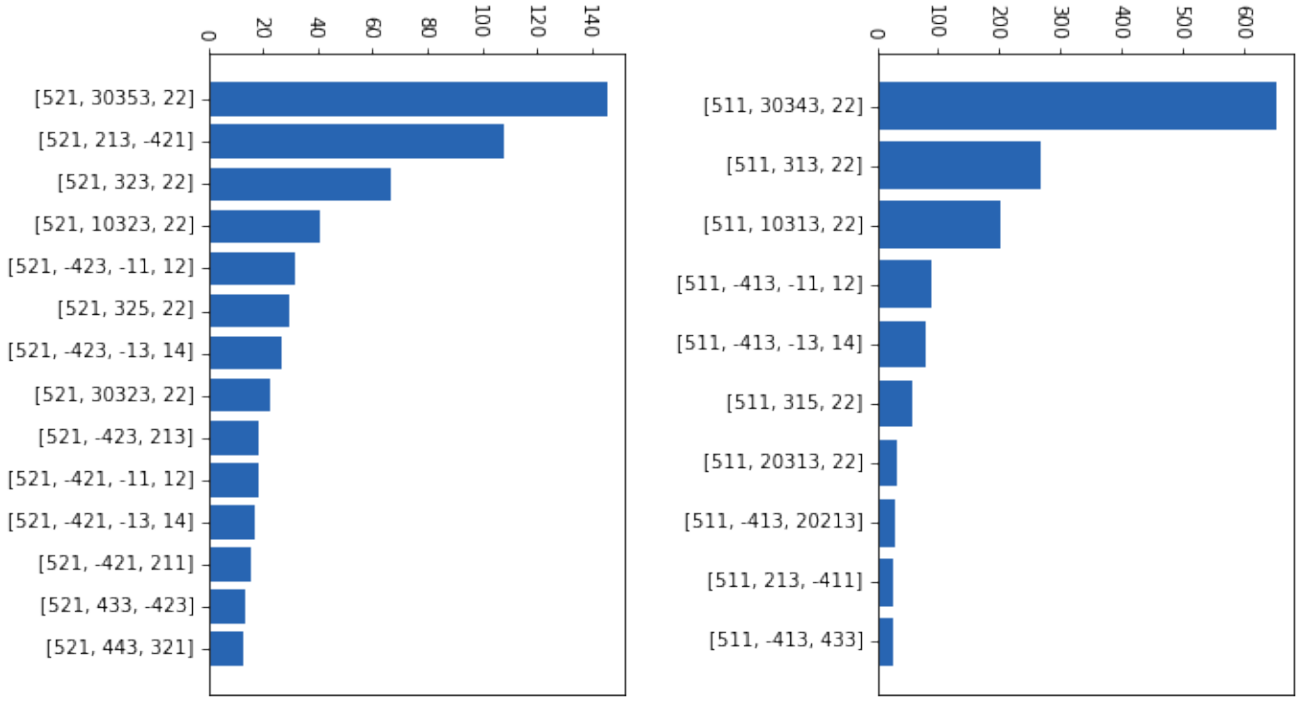


Figure 4.57. Sorted frequencies of decay for $B^{\pm}$ (left) and B^0 ($\bar{B}^0$) (right). The numbers on the vertical axis are PDG codes of corresponding partciles within Monte-Carlo simulation. The mapping from PDG code to the particle can be found in [147]. First number within each list corresponds to the B-meson, the rest — daughters of B-meson decay.

The five leading sources are

1) $B^0 \to X_s\gamma$,

2) $B^0 \to K^{0*}\gamma$,

3) $B^0 \to K_1^0\gamma$,

4) $B^+ \to X_s^+\gamma$,

5) $B^+ \to \bar{D}^0\rho^+$,

For instance, in the case of $B^0 \to K^{0*}\gamma$, followed by $K^* \to K^+\pi^-$ decay, soft π^+ can be captured to create $K^+\pi^+\pi^-\gamma$ final state.

Except for the last one, leading background sources are all peaking in the signal region of $M_{bc} \in [5.27; 5.29]$ GeV/c^2 and will be described by a dedicated PDF in the $M_{\rm bc}$ fit for the yield extraction.

Their $M_{\rm bc}$, ΔE, and $M_{K\pi\pi}$ distributions are shown in Figs(4.58)-(4.62).

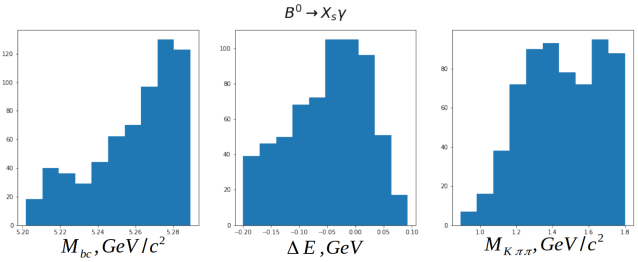


Figure 4.58. $B^0 \to X_s\gamma$ decays : $M_{\rm bc}$, ΔE and $M_{K\pi\pi}$ distributions.

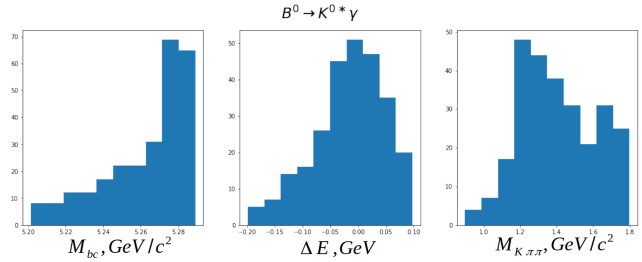


Figure 4.59. $B^0 \to K^{0*}\gamma$ decays : $M_{\rm bc}$, ΔE and $M_{K\pi\pi}$ distributions.

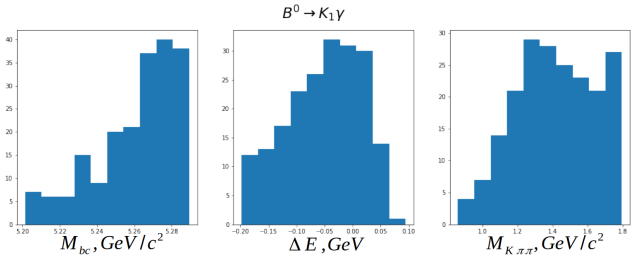


Figure 4.60. $B^0 \to K_1^0\gamma$ decays : $M_{\rm bc}$, ΔE and $M_{K\pi\pi}$ distributions.

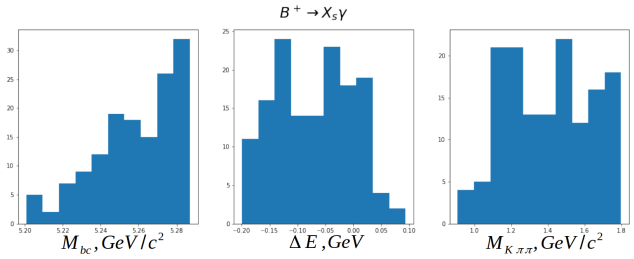


Figure 4.61. $B^+ \to X_s^+\gamma$ decays : $M_{\rm bc}$, ΔE and $M_{K\pi\pi}$ distributions.

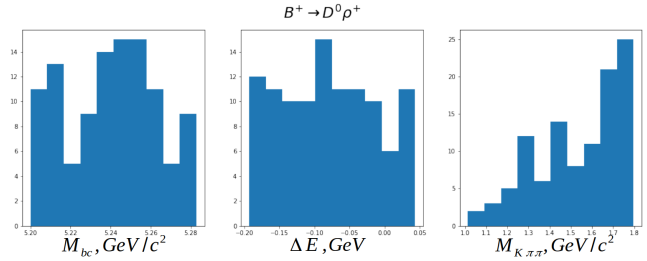


Figure 4.62. $B^+ \to D^0\rho^+$ decays : $M_{\rm bc}$, ΔE and $M_{K\pi\pi}$ distributions.

4.4 . Control samples

Typical analysis path within Belle II can be visualized as in Fig. 4.63 and contains three main stages:

- Monte-Carlo study establishes the analysis procedure using artificial data coming from the software simulating e^+e^- collisions. For instance, one the chooses the following steps: Monte-Carlo dataset, types, and values of selection cuts, continuum suppression variables and corresponding model, best candidate selection if applicable, probability density functions for signal and background description, control samples for further unblinding;

- Control Samples unblinding procedure validates the decisions taken on the previous step. It is done using the sets of real data collected by the Belle II detector, which should not include signal events, but only background;
- Full-unblinding ("opening the box") allows to access the real data captured by the Belle II detector in the signal region hidden in the previous step. The measurements of various physics observables targeting specific analysis are performed in this last stage.

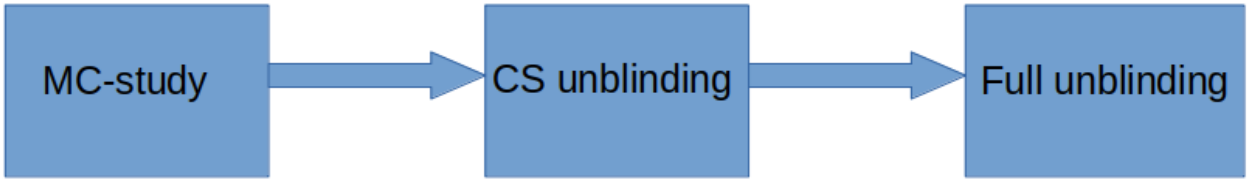


Figure 4.63. Typical analysis path.

The primary purpose of the control samples is to validate the analysis procedure by comparing data and Monte-Carlo at various stages of the analysis. In the present work, the function of the control sample is also to set additional constraints on the final fit, as explained in section (4.5.8). The control samples are discussed below.

4.4.1 . Non-peaking backgrounds

For continuum events ($q\bar{q}$), off-resonance data was collected from runs where the center-of-mass energy was $60\ \mathrm{MeV}$ below the $\Upsilon(4S)$ resonance, provides a control sample of about 10% of the statistics of the on-resonance data sample. While insufficient for a detailed analysis, this control sample allows performing sanity checks, as for example with the distributions of M_{bc} and μ_{NN}.
More importantly, the consistency of the Monte-Carlo description of the non-peaking background can be checked by the sidebands outside the signal range $M_{\mathrm{bc}} \in [5.2, 5.27]\ \mathrm{GeV/c^2}$.

4.4.2 . Peaking backgrounds

As mentioned in section (4.3.5), because peaking backgrounds are located in the signal region, they can directly affect the yield determinations. In contrast with continuum background and non-peaking $B\bar{B}$ events that are controlled by side-bands (mostly $M_{\mathrm{bc}} \in [5.2, 5.27]\ \mathrm{GeV/c^2}$) the level of peaking background cannot be assessed from kinematical side-bands, since by definition peaking background events are kinematically signal-like. However, a side-band is provided by the control samples presented in this section.

The control samples are defined making use of the origin of the peaking background that arises because of track lost or swapped from events with kinematics similar to the kinematics of signal events. The track lost or swapped can lead to candidates which track charges are consistent with signal events, or inconsistent with signal events. In the latter case, the events are not selected as background, but they are kinematically signal-like, and thus provide control samples. This is summarized in table (4.8) where seven classes of events are listed:

- Class 0 : TM events,
- Class 1a : SCF where the π^+ is coming from the ROE,
- Class 1b : SCF where the π^- is coming from the ROE,
- Class 1c : SCF where the K^+ is coming from the ROE,
- Class 1d : SCF for $K^-\pi^-\pi^+\gamma$ events, where the K^+ is replaced by a K^- of the ROE,
- Class CS1 : $K^+\pi^-\pi^-\gamma$ events where the π^+ is replaced by a π^- of the ROE,
- Class CS2 : $K^+\pi^+\pi^+\gamma$ events where the π^- is replaced by a π^+ of the ROE.

				Name	Class	N_1	N_2
K^+	π^+	π^-	γ	TM	0	-	-
K^+	π^+	π^-	γ	SCF	1a	477	849
K^+	π^+	π^-	γ	SCF	1b	474	843
K^+	π^+	π^-	γ	SCF	1c	19	32
K^-	π^+	π^-	γ	SCF	1d	61	108
K^+	π^-	π^-	γ	$K^+\pi^-\pi^-$	CS1	-	976
K^+	π^+	π^+	γ	$K^+\pi^+\pi^+$	CS2	-	880

Table 4.8. Truth level particles are indicated in black, and swapped particles are indicated in red. A swapped particle comes from the ROE, with or without its correct PID. The numbers in the column N_1 are the number of SCF events (from signal Monte-Carlo) where only the particle in red has been swapped. The sum of the quoted numbers (1031) is lower than the number of SCF events (1216) because for 185 SCF events more than one particle is not truth-matched. The numbers in the last column N_2 are about twice larger: they are obtained as for N_1 (i.e. from signal Monte-Carlo) but taking away the TM events when applying BCS (see text).

The two control samples discussed in this section are events of the last two classes which are selected following the same procedure as for the signal sample events, but for the charges of the pion candidates that have to be equal, instead of the opposite. These types of events can only arise from combinatorics, not from physics processes:

CS1: this type of events could come from a $B^- \to K^+\pi^-\pi^-\gamma$ decay, but this decay is extremely suppressed in the Standard Model because, at quark level, it involves the intricate transition $b \to d(d\bar{u})(u\bar{s})$,

CS2: this type of events would violate charge conservation if it was coming from a B decay.

As stated above, events of the CS1 type (resp. CS2 type) can arise from signal events where a π^+ (resp. a π^-) has been swapped with a π^- (resp. a π^+) from the ROE : in that case, they are akin to SCF events of class 1a (resp. 1b). They can also arise from $B\bar{B}$ peaking background like $B^0 \to K^0_{\rm res}\gamma \to K^+\pi^0\pi^-\gamma$,

with a soft charged π from the ROE replacing a soft π^0: in that case, they are akin to the $B\bar{B}$ peaking background. Finally, they may also arise from $q\bar{q}$ and non-peaking $B\bar{B}$ backgrounds: in that case, they are akin to the non-speaking background.

A priori, the $M_{\rm bc}$ distributions of these two control samples cannot be assumed to be identical to the $M_{\rm bc}$ distributions of the peaking and non-peaking backgrounds. However,

1) the underlying physics and detector effects involved are similar. Therefore if the data distributions of CS1 and CS2 are satisfactorily described by Monte-Carlo, one can expect that the corresponding distributions in the signal sample are also satisfactorily described by Monte-Carlo.

2) the component of peaking background in these control samples can be assumed to be proportional to the peaking background in the Physics Sample. Therefore a combined analysis of the Physics Sample and the Control Samples can help to pin down the peaking background contribution (cf. section(4.5)).

For $1\ {\rm ab}^{-1}$ of Monte-Carlo data the composition of these two control samples are indicated in Tables (4.9)-(4.10). The distributions of $M_{\rm bc}$ for these two types of events are shown in Figs(4.64)-(4.65).

event type	$K^+\pi^+\pi^-\gamma$	$q\bar{q}$	B^+B^-	$B^0\bar{B}^0$
$N_{\rm event}$	976	9690	806	2457
%	7.0	69.6	5.8	17.6

Table 4.9. Composition of events in the control sample CS1 ($K^+\pi^-\pi^-\gamma$) after the full selection procedure.

event type	$K^+\pi^+\pi^-\gamma$	$q\bar{q}$	B^+B^-	$B^0\bar{B}^0$
$N_{\rm event}$	880	4516	417	763
%	13.3	68.6	6.3	11.6

Table 4.10. Composition of events in the control sample CS2 ($K^+\pi^+\pi^+\gamma$) after the full selection procedure.

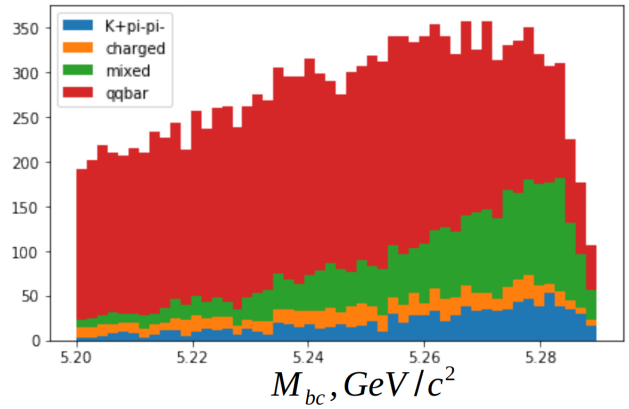


Figure 4.64. M_{bc} distributions for $B \to K^+\pi^-\pi^-\gamma$. The contribution from signal events is indicated as $K^+\pi^-\pi^-$ in the figure.

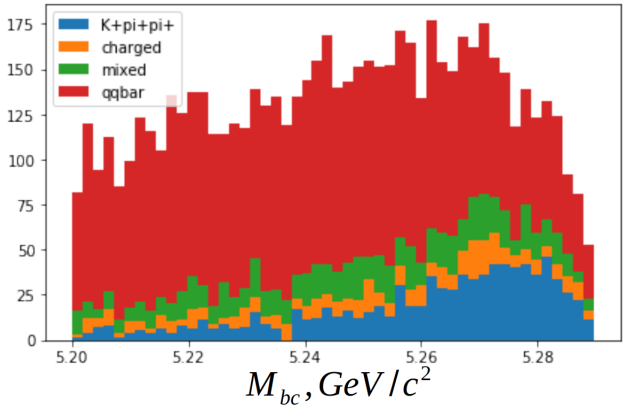


Figure 4.65. M_{bc} distributions for $B \to K^+\pi^+\pi^+\gamma$. The contribution from signal events is indicated as $K^+\pi^+\pi^+$ in the figure.

The contribution from $K^+\pi^+\pi^-\gamma$ events to the control sample (976 events for CS1 and 880 for CS2) is similar to the number of SCF events in the signal sample (1216 events cf. table (4.6)). This is quantitatively well understood because one expects to have the same number of events in classes 1a and CS1 and in classes 1b and CS2. This is what is observed if one takes care to remove TM events when counting events in classes 1a and 1b as shown in column N2 of table (4.8). Otherwise, mostly TM candidates are retained when BCS is applied, not the SCF events, as shown in column N1 of the table (4.8) where the number of events is about half the one of column N2.
Ignoring the TM candidates, the 7386 TM events (cf. table (4.6)) become 831 SCF events that added to the 1216 SCF events become 2047 SCF events. This is to be compared with the 1856=976+880 events of classes CS1 and CS2. The remaining difference 2047-1856=191 is explained by the fact the SCF also receives contributions from $K^{\pm}$ swaps. For instance, among the 2047 SCF events, 849 (resp. 843) come from a π^+ (resp. π^-) swap, to be compared to the 976 CS1 (resp. 880 CS2) events (cf. table (4.8)).

The contribution from background events, by contrast, is lower than the number of the background of events in the signal sample (18649 to be compared with 26463 cf. table (4.6)) because

- combinatorics favors the reconstruction of decay with total charge -1 (as for CS1) than a decay with total charge +3 (as for CS2),
- the $B\bar{B}$ peaking background is almost absent in CS2.

In summary, the composition of CS1 is close to the background composition of the Physics Sample, and its size is about half the size of the Physics Sample (if one replaces SCF events with "SCF" events). Both CS1 and CS2 will be used as a control sample but in addition, CS1 will also be considered in the fit, as a complement to the Physics Sample (cf. section (4.5.8)).

4.5 . Yields measurement

This section is organized as follows. After an overall presentation of the extended maximum likelihood fit followed by ${}_s\mathcal{P}lot$ notations, including the p.d.f. 's used in the fit in section (4.5.3), a welcome technical

simplification is explained in section (4.5.4), the TM-oriented baseline fit is defined in section (4.5.5), the foreseen validation of the fit with ${}_s\mathcal{P}lot$ distributions is discussed in section (4.5.6), and finally in section (4.5.7) a toy Monte-Carlo study is performed to assess the behavior of the fit with a reduced Physics Sample of $63\ \mathrm{fb}^{-1}$. Also the possible extension to the baseline fit including the control sample CS1 is presented in section (4.5.8).

4.5.1 . Extended maximum likelihood fit and ${}_s\mathcal{P}lot$

Having applied all selections, the resulting data sample, apart from signal, contains contributions from different backgrounds (for instance $q\bar{q}, B\bar{B}$, SCF). The total distribution of a random variable x, representing a physics observable within the data sample is modeled using probability density functions of different species of events present in the data sample. The functional view $\mathcal{P}_j(x, \overrightarrow{a}_j)$ of the p.d.f for given event type is usually selected based on Monte-Carlo data. The values of some model parameters $\overrightarrow{a}_j$ can be also fixed from Monte-Carlo or control samples. For instance, distribution of a random variable of event type j is fitted using maximum likelihood method defined by Eqs. 2.82-2.83 and parameters $\overrightarrow{a}_j$ are obtained along with their covariance matrix. At this stage one may decide which parameters out of $\overrightarrow{a}_j$ should be fixed for each $\mathcal{P}_j(x, \overrightarrow{a}_j)$.

Because of the stochastic nature of underlying Particle Physics processes, it is in principle unknown in advance how many signal or background events were produced in the given sample of data corresponding to predefined integrated luminosity. For the Monte-Carlo data, using generator-level information, yields of different species of events can be counted but not in the case of real data.

The procedure allowing simultaneously calculate expected yields, $\overrightarrow{a}_j$ for each species j along with their covariance matrix is referred to as extended maximum likelihood fit. In the frame of this approach likelihood function is defined as follows:

$$L(\{N_j; \overrightarrow{a}_j\}) = \underbrace{\frac{e^{-\sum_j^M N_j}(\sum_j^M N_j)^N}{N!}}_{\text{probability of outcome } N} \prod_i^N \underbrace{\frac{\sum_j^M N_j \mathcal{P}_j(x_i, \overrightarrow{a}_j)}{\sum_j^M N_j}}_{\text{total p.d.f of all processes}} \tag{4.32}$$

where it is used that number of events for each species and total number of events are distributed according to the Poisson law [134]:

$$n_j \sim \frac{e^{-N_j} N_j^{n_j}}{n_j!} \tag{4.33}$$

$$N = \sum_j^M n_j \sim \frac{e^{-\sum_j^M N_j}(\sum_j^M N_j)^N}{N!} \tag{4.34}$$

where N — total number of events in a given datasample; n_j — number of events of j-th species in a given datasample; N_j — expected yield for j-th event species; M — number of species of events; x — discriminating variable; $\mathcal{P}_j(x, \overrightarrow{a}_j)$ — p.d.f of variable x for j-th event species.

The Eq. 4.32 can be simplified:

$$L(\{N_j; \overrightarrow{a}_j\}) = e^{-\sum_j^M N_j} \prod_i^N \sum_j^M N_j \mathcal{P}_j(x_i, \overrightarrow{a}_j) \tag{4.35}$$

The minimization is done with respect to $-\log L$ and optimal parameter values $\{N_j; \overrightarrow{a}_j\}$ are computed. Then covariance matrix V mentioned above is calculated as following:

$$V_{jk}^{-1} = \frac{\partial^2(-\log L)}{\partial b_i \partial b_j} \tag{4.36}$$

where $b_i \in \{N_j; \overrightarrow{a}_j\}$.

Having performed extended maximum likelihood fit using x variable from the data sample, it is possible to recover the distribution of variable y from the same dataset, assuming it is not statistically correlated with x. The algorithm is referred to as ${}_s\mathcal{P}lot$ and fully described in [135]. In the frame of ${}_s\mathcal{P}lot$ extended likelihood fit is repeated one more time with all fixed $\overrightarrow{a}_j$ values obtained from the previous fit and a new covariance matrix is computed with respect to the yields N_j. Then for event e, j-th weight is defined as follows:

$${}_s\mathcal{P}_j(x_e) = \frac{\sum_{k=1}^{M} V_{jk}\mathcal{P}_k(x_e, \overrightarrow{a}_k)}{\sum_{k=1}^{M} N_k \mathcal{P}_k(x_i, \overrightarrow{a}_k)} \tag{4.37}$$

The histogram of y distribution for event species j can be obtained by incrementing ${}_s\mathcal{P}_j(x_e)$ to the height H_{nj} of bin n, such that y_e falls within its lower and upper bounds:

$$H_{nj} = \sum_{e \subset \delta y} {}_s\mathcal{P}_j(y_e) \tag{4.38}$$

In this work, the fit and ${}_s\mathcal{P}lot$ are performed in the frame of RooFit package [136].

4.5.2 . Choice of discriminating variable for the fit

Only the variable $M_{\rm bc}$ is used to perform the measurement of the yields. In principle, even though the $\mu_{\rm NN}$ variable is correlated to $M_{\rm bc}$, its correlation is mild enough that it would not be difficult to implement a more powerful 2D-fit using both $M_{\rm bc}$ and $\mu_{\rm NN}$. However, at this stage of the life of the Belle II experiment, it was deemed premature to assume being able to rely on a Monte-Carlo description of the $\mu_{\rm NN}$ distribution to be used in the fit. Rather, the $\mu_{\rm NN}$ distribution for signal and background events is planned to be obtained from the fit using ${}_s\mathcal{P}lot$.

Another option that has been explored is an analysis using two bins in the NN variable: $\mu_{\rm NN} \in [0.28, 0.64]$ and $\mu_{\rm NN} \in [0.64, 1.0]$ as shown in Figs.(4.50)-(4.51). While the two bins contain about the same numbers of TM events, the background in the second bin is strongly reduced. This allows a significant improvement of the fit, provided the two bins are distinguished in a combined fit. But, here again, it was concluded that the baseline analysis at this stage should be kept as simple as possible.

4.5.3 . Notations

Events within the Physics Sample are initially split into the four different groups previously defined: TM, SCF, $q\bar{q}$ and $B\bar{B}$ (cf. Table (4.6)).
The probability density functions (p.d.f.) that are used to describe the individual distributions of events are functions of $M_{\rm bc}$ of the following types:

$$F_{\rm TM} = \text{CB} \tag{4.39}$$
$$F_{\rm SCF} = \text{Argus} \tag{4.40}$$
$$F_{q\bar{q}} = \text{Argus} \tag{4.41}$$
$$F_{B\bar{B}} = \eta\ \tilde{\text{G}} + (1-\eta)\ \text{Argus} \tag{4.42}$$

where in the last expression, the fraction η of $\tilde{\mathrm{G}}$ accounts for the peaking background. The notations are[5]

1) CB: Crystal Ball p.d.f. (four parameters) see Eq.(4),

2) Argus: Argus p.d.f. (two parameters) see Eq.(1),

3) $\tilde{\mathrm{G}}$: modified Gaussian (μ, σ, κ parameters), defined as

$$\tilde{\mathrm{G}}(M_{\mathrm{bc}}) \;=\; \mathrm{cst}\;\exp\left(-\frac{(\mathrm{M_{bc}}-\mu)^2}{2\sigma^2}\right)\;\mathrm{Erf}\left((\frac{\mathrm{M_{bc}^{max}}-\mathrm{M_{bc}}}{\sigma})^{\kappa}\right) \tag{4.43}$$

where the normalization constant cst must be computed numerically. This modified Gaussian is designed to drop to zero when M_{bc} is close to $5.29\ \mathrm{GeV/c^2}$.

The complete distribution of events is the yield-weighted sum:

$$\mathcal{F}_{\mathrm{tot}} \;=\; N_{\mathrm{TM}}F_{\mathrm{TM}} + N_{\mathrm{SCF}}F_{\mathrm{SCF}} + N_{q\bar{q}}F_{q\bar{q}} + N_{B\bar{B}}F_{B\bar{B}} \tag{4.44}$$

$$\int \mathcal{F}_{\mathrm{tot}}\ \mathrm{d}M_{\mathrm{bc}} \;=\; N_{\mathrm{TM}} + N_{\mathrm{SCF}} + N_{q\bar{q}} + N_{B\bar{B}} \tag{4.45}$$

The distributions of the components are shown in Fig.(4.66), with the $B\bar{B}$ component split into B^+B^- and $B^0\bar{B}^0$.

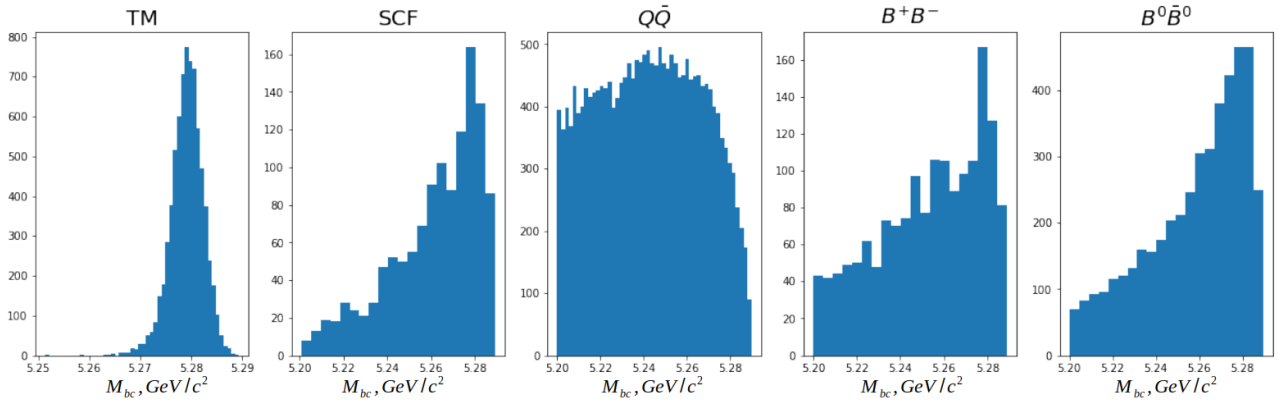


Figure 4.66. The M_{bc} distribution for signal and main background categories.

As already signaled in section (4.3.5), one observes that the distributions of SCF and the peaking component of $B^0\bar{B}^0$ are similar: their respective contributions are difficult to disentangle by the fit. Accordingly, two options can be considered.

- Merging the F_{SCF} and $F_{B\bar{B}}$ p.d.f. 's, one rewrites Eq.(4.44) to define the tFit (for truth-match Fit) distribution as:

$$\mathcal{F}_{\mathrm{tot}}^{\mathrm{tFit}} \;=\; N_{\mathrm{TM}}F_{\mathrm{TM}} + N_{q\bar{q}}F_{q\bar{q}} + N_{\mathrm{bkg}}F_{\mathrm{bkg}} \tag{4.46}$$

where F_{bkg} is the same functional as $F_{B\bar{B}}$ (cf. Eq.(4.42)).

[5]The expressions of the normalized Crystal Ball and Argus p.d.f. 's can be found in section(5).

- Merging the two $N_{\rm TM}$ and $N_{\rm SCF}$ yields into a unique $N_{\rm signal}$ signal yield, one rewrites Eq.(4.44) to define the sFit (for signal Fit) distribution as:

$$\mathcal{F}_{\rm tot}^{\rm sFit} \;=\; N_{\rm signal}(\gamma F_{\rm TM} + (1-\gamma)F_{\rm SCF}) + N_{q\bar{q}}F_{q\bar{q}} + N_{B\bar{B}}F_{B\bar{B}} \tag{4.47}$$

  with $\gamma = 0.86$ (refers to 7386/8602 from Table (4.6)).

Keeping in mind that access to the TM events through ${}_s\mathcal{P}lot$ is compulsory for the photon polarization measurement, this tFit option is preferred.

In both tFit and sFit approaches, a difficulty arises because of the presence of two Argus functions, one for $F_{q\bar{q}}$ and one for $F_{\rm bkg}$ (or $F_{B\bar{B}}$), which makes the fits behave badly. A much welcome simplification occurs, as explained in the next section.

4.5.4 . Argus feature

The aim of this section is to demonstrate that a single Argus p.d.f. can be used to describe $q\bar{q}$ and the non-peaking component of $B\bar{B}$ in tFit (the same holds for sFit).
For a given "threshold" ($M_{\rm bc}^{\rm max} = 5.29$ GeV/c^2), but two χ values, even widely different, the Argus p.d.f. presents the peculiarity:

$$\beta{\rm Argus}_{\chi_1} + (1-\beta){\rm Argus}_{\chi_2} \simeq {\rm Argus}_{\chi_{12}} \tag{4.48}$$

where the parameter χ_{12} is a function of χ_1, χ_2 and β. The approximation is excellent in a wide range of parameter values to the point that it cannot be detected even with a few tens of ab^{-1}. This is illustrated below, by considering two toy Monte-Carlo datasets generated with very different χ-parameters, namely $\chi_1 = -20.8$ and $\chi_2 = -79.8$ correspond respectively to the $q\bar{q}$ and a non-peaking component of the $B\bar{B}$ background distributions. The two datasets are merged into a single dataset with a dominant $q\bar{q}$ contribution ($\beta = 0.75$). Two luminosity values are considered, ~ 50 ab^{-1} for Fig.(4.67) and $\sim .5$ ab^{-1} for Fig.(4.68).

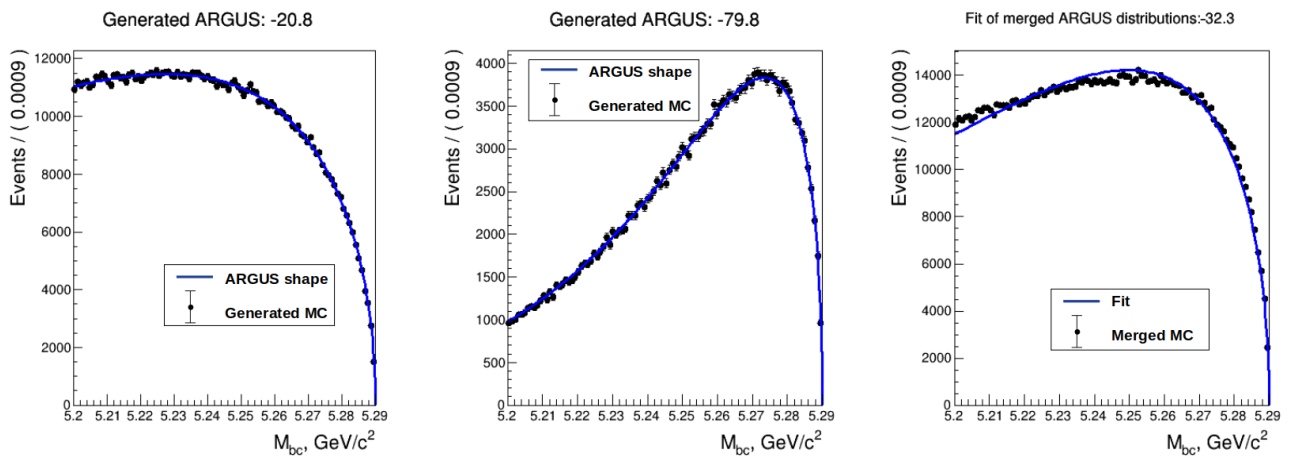


Figure 4.67. Argus distributions of $q\bar{q}$ events, with $\chi_1 = -20.8$ (left), non-peaking $B\bar{B}$ events, with $\chi_2 = -79.8$ (center) and their sum (right). The latter is well described by a single Argus function with $\chi_{12} = -32.3$. The statistics correspond to a luminosity of ~ 50 ab^{-1}.

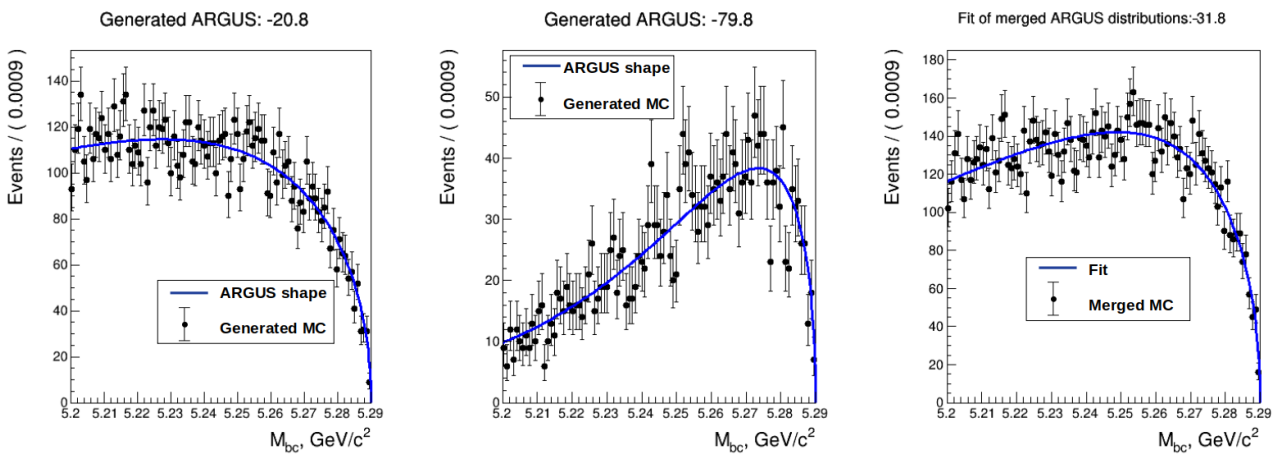


Figure 4.68. Argus distributions of $q\bar{q}$ events, with $\chi_1 = -20.8$ (left), non-peaking $B\bar{B}$ events, with $\chi_2 = -79.8$ (center) and their sum,(right) that is well described by a single Argus function with $\chi_{12} = -31.8$. The statistics correspond to a luminosity of ~ 0.5 ab^{-1}.

Superimposed on the toy Monte-Carlo distributions are fitted using a single Argus function. By construction, the fit is perfect for the two datasets taken separately, since they were generated precisely with Argus functions. The point is that the fit with a single Argus function remains excellent for the merged dataset. Only in Fig.(4.67) which corresponds to a huge statistics does one start to perceive that the fit is not perfect : for ~ 0.5 ab^{-1} (see Fig. 4.68) the agreement is excellent.

This feature of the Argus p.d.f. is important since

- If not recognized, it leads to unstable fit with large correlations between χ_1, χ_2 and β.
- If recognized, it allows to reduce the number of parameters by two units; χ_{12} replacing χ_1, χ_2 and β.

Merging the two Argus components, the two options of Eqs.(4.46)-(4.47) become:

$$\mathcal{F}_{\rm tot}^{\rm tFit} = N_{\rm TM}F_{\rm TM} + N_{\rm Argus}^{\rm tFit}{\rm Argus} + {\rm N}_{\rm Peak}^{\rm tFit}\tilde{\rm G}^{\rm tFit} \quad (4.49)$$

$$\mathcal{F}_{\rm tot}^{\rm sFit} = N_{\rm signal}(\gamma F_{\rm TM} + (1-\gamma)F_{\rm SCF}) + N_{\rm Argus}^{\rm sFit}{\rm Argus} + {\rm N}_{\rm Peak}^{\rm sFit}\tilde{\rm G}^{\rm sFit} \quad (4.50)$$

4.5.5 . Baseline fit: Truth-Match fit (tFit)

Having access to the MC-truth information of signal and backgrounds, a series of likelihood fits can be performed to make sure that considered p.d.fs can properly describe the data in the absence of other event species. In addition, several p.d.fs parameters can be fixed in order to avoid over parametrization on the stage of the full fit with $\mathcal{F}_{\rm tot}^{\rm tFit}$.

Given sample of Monte-Carlo data, it is separated as following: TM, $q\bar{q}$, $B\bar{B}$+SCF. First of all, such separation allows essentially fit the M_{bc} distribution of signal events and fix all parameters of the $F_{\rm TM}$. Full background represented by $q\bar{q}$, $B\bar{B}$ and SCF is further splitted in order to extract parameter values of $\tilde{\rm G}^{\rm tFit}$. Since $q\bar{q}$ background is not peaking anyway, only merged $B\bar{B}$ + SCF sample is considered. The

likelihood fit of $B\bar{B}$+SCF is performed using Eq. 4.42 followed by fixing parameters of $\tilde{\mathrm{G}}^{\mathrm{tFit}}$ to the fitted values. Other parameters of F_{bkg} are not further considered. On this stage Argus p.d.f within F_{bkg} is served as "springboard" for fitting peaking part of peaking background and χ-parameter is re-adjusted on the full fit stage. Thereby, proposed events splitting allows to define the values of model parameters for F_{TM} and $\tilde{\mathrm{G}}^{\mathrm{tFit}}$ p.d.fs within Eq. 4.49. The remaining part is described by a generic Argus function, with a free χ ad-hoc parameter and fixed threshold $M_{bc}^{\mathrm{max}} = 5.29$ GeV/c^2.

The M_{bc} distributions of TM, $q\bar{q}$ and $B\bar{B}$+SCF events are shown in Fig.(4.69) together with their corresponding adjusted p.d.f. 's: F_{TM}, $F_{q\bar{q}}$ and $F_{B\bar{B}} = \eta\ \tilde{\mathrm{G}}^{\mathrm{tFit}} + (1-\eta)$ Argus.

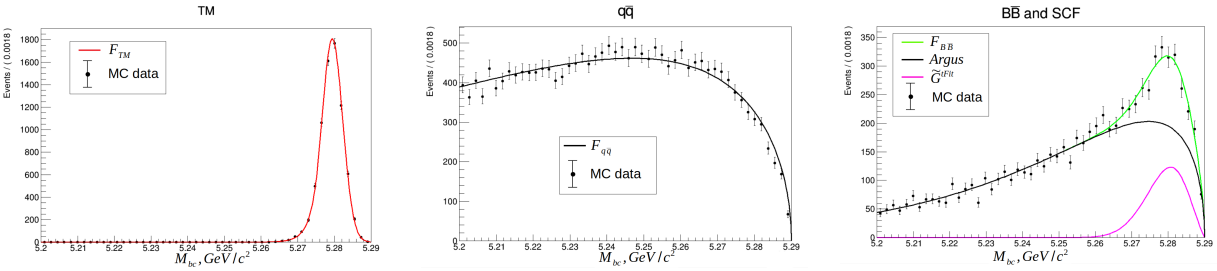


Figure 4.69. The M_{bc} distributions of TM, $q\bar{q}$ and the combination of SCF and $B\bar{B}$ events, together with their corresponding p.d.f. 's $F_{\mathrm{TM}} = \mathrm{CB}$, $F_{q\bar{q}} = \mathrm{Argus}$ and $F_{B\bar{B}} = \eta\ \tilde{\mathrm{G}}^{\mathrm{tFit}} + (1-\eta)$ Argus.

The extended likelihood fit is further performed with four floating parameters: N_{TM}, N_{Argus} and N_{Peak} yields, plus χ (see Table 4.11).

	par name	par type	par value	fixed from
F_{TM}	α	fixed	1.382	MC
	μ	fixed	5.27942	
	σ	fixed	0.002844	
	n	fixed	25.5	
	N_{TM}	floating		
Argus	χ	floating		
	μ_{thrs}	fixed	5.29	MC
	N_{Argus}	floating		
$\tilde{\mathrm{G}}^{\mathrm{tFit}}$	σ	fixed	0.0077	MC
	μ	fixed	5.2829	
	κ	fixed	1.0	
	N_{Peak}	floating		

Table 4.11. List of the parameters needed to define $\mathcal{F}_{\mathrm{tot}}^{\mathrm{tFit}}$. The floating parameters are the yields N_{TM}, N_{Argus} and N_{Peak}, plus the χ parameter of the Argus p.d.f.

The fit leads to a satisfactory description of the Physics Sample, as shown in Fig.(4.70)

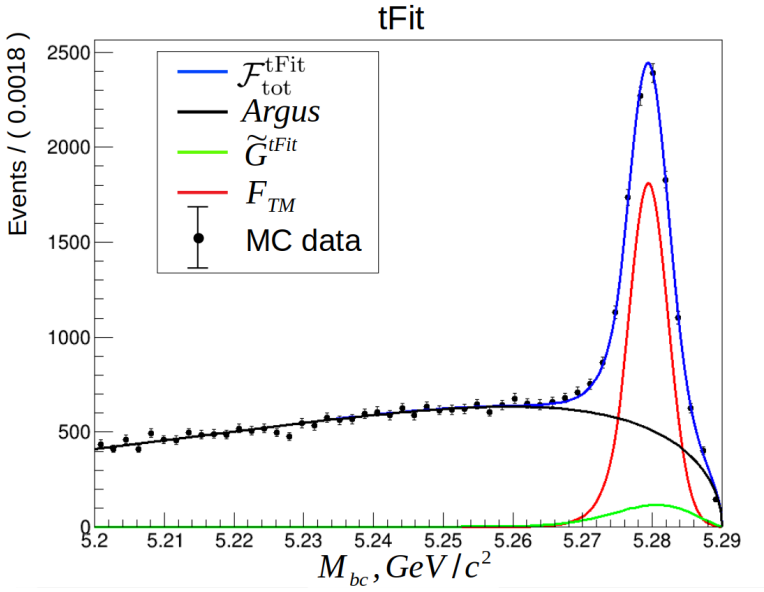


Figure 4.70. The fit of M_{bc} distribution of Physics Sample of 1 ab^{-1}: ≈ 100 % of overlapping between $F_{\rm TM}$ and $\tilde{\rm G}^{\rm tFit}$ causes significant correlation between corresponding expected yields despite on visual agreement between fit and data.

The fit results are given in Tables (4.12)-(4.13). In particular the result for TM events $N_{\rm TM} = 7400 \pm 203$ is fairly consistent with the expected value $N_{\rm TM} = 7386$.

parameter	value	σ
$N_{\rm TM}^{\rm tFit}$	7400	203
$N_{\rm Argus}^{\rm tFit}$	26670	224
$N_{\rm Peak}^{\rm tFit}$	981	293
$\chi^{\rm tFit}$	-43.8	1.0

Table 4.12. Parameters values resulting from the fit.

parameter	$N_{\rm TM}^{\rm tFit}$	$N_{\rm Argus}^{\rm tFit}$	$N_{\rm Peak}^{\rm tFit}$	$\chi^{\rm tFit}$
$N_{\rm TM}^{\rm tFit}$	-	+0.33	-0.81	-0.29
$N_{\rm Argus}^{\rm tFit}$		-	-0.58	-0.51
$N_{\rm Peak}^{\rm tFit}$			-	+0.58
$\chi^{\rm tFit}$				-

Table 4.13. Correlation matrix of the fit. To ease the reading, only the upper part of the symmetric matrix is shown. Very high correlation is observed between $N_{\rm TM}^{\rm tFit}$ and $N_{\rm Peak}^{\rm tFit}$: $\tilde{\rm G}^{\rm tFit}$ and $F_{\rm TM}$ are both peaking in the signal M_{bc} region, which complicates their separation.

4.5.6 . ${}_s\mathcal{P}lot$ validation

The ${}_s\mathcal{P}lot$ method is a subtraction method that allows to reconstruct signal and background distributions of variables that are in principle statistically independent of the variables used in the fit, in our case $M_{\rm bc}$. The statistical independence is a strict requirement if one plans to use the reconstructed distribution - as for measurements related to the photon polarization - but it is not a strict requirement if one merely wants to provide a validation of the analysis: namely to compare ${}_s\mathcal{P}lot$ obtained with data to ${}_s\mathcal{P}lot$ obtained with Monte-Carlo.
Here we are only interested in the validation of the analysis, using the baseline fit. In all figures, the ${}_s\mathcal{P}lot$ for TM events is superimposed on the histogram of TM events. The normalizations are very close: 7400 events for the ${}_s\mathcal{P}lot$ (cf. Table (4.16)) and 7386 events (cf. Table (4.6)) for the histograms. For the sake of illustration, six variables are considered, and only for TM events:

1) $\mu_{\rm NN}$,

2) ΔE,

3) $\cos T_{\rm B} z$,

4) $M_{K\pi\pi}$,

5) $\cos\theta_\gamma$ defined in Eq. 2.6 is the cosine of the photon direction with respect to the normal of the $K^+\pi^+\pi^-$ plane, in the rest frame of the latter,

6) $\cos\theta_B = p_{Bz}/p_B$ as computed in the CMS,

The first three variables are NN variables: they are not fully independent of $M_{\rm bc}$. However, their ${}_s\mathcal{P}lot$'s show a good agreement with the MC expectations except for the ΔE distribution.
The last three variables are expected to be independent of $M_{\rm bc}$ and their ${}_s\mathcal{P}lot$'s show a good agreement with the MC expectations. The ${}_s\mathcal{P}lot$ for $\cos\theta_B$ is particularly interesting since the $(1-\cos^2\theta_B)$ shape is a model-independent prediction for the $e^+e^- \to B^+B^-$ process.

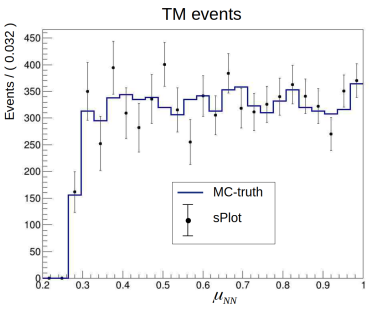


Figure 4.71. The comparison of ${}_s\mathcal{P}lot$ and truth MC distribution of $\mu_{\rm NN}$ for TM events.

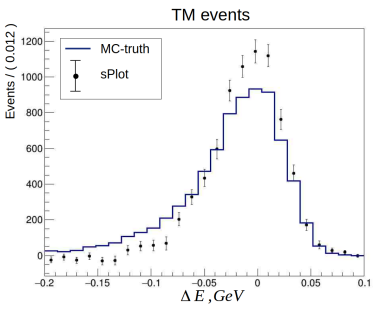


Figure 4.72. The comparison of ${}_s\mathcal{P}lot$ and truth MC distribution of ΔE for TM events.

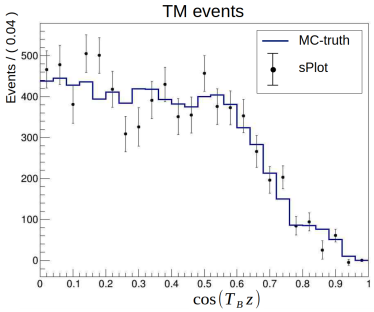


Figure 4.73. The comparison of ${}_s\mathcal{P}lot$ and truth MC distribution of $\cos T_{\rm B} z$ for TM events.

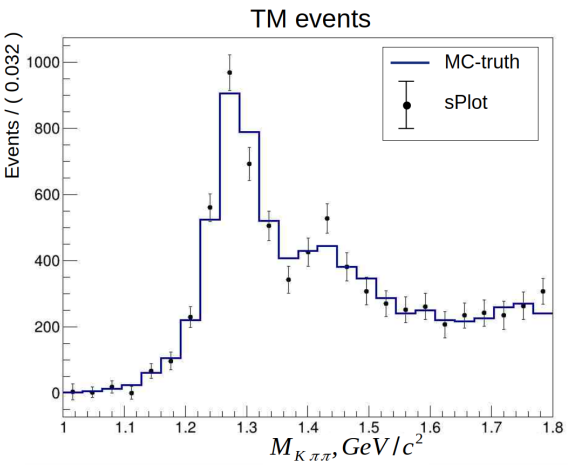


Figure 4.74. The comparison of ${}_s\mathcal{P}lot$ and truth MC distribution of $M_{K\pi\pi}$ for TM events.

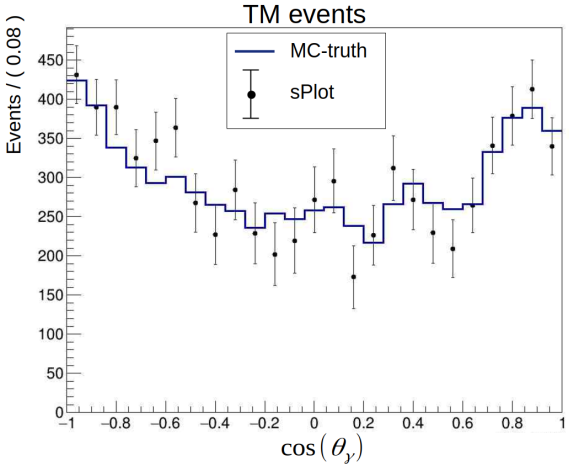


Figure 4.75. The comparison of ${}_s\mathcal{P}lot$ and truth MC distribution of $\cos\theta_\gamma$ for TM events.

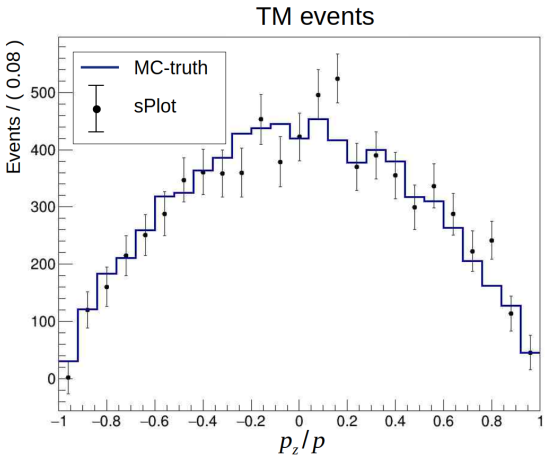


Figure 4.76. The comparison of ${}_s\mathcal{P}lot$ and truth MC distribution of $\cos\theta_B$ for TM events.

4.5.7 . tFit toy Monte-Carlo studies

In this section the toy Monte-Carlo technique will be used for two studies of tFit behavior:

- the cross-check of the tFit results,
- an assessment of the systematical bias.

tFit validation with toy MC

To cross-check the tFit results, a dedicated toy Monte-Carlo study is presented in this section. The parameters of the p.d.f.'s obtained from the Physics Sample are used to produce 10^3 toy Monte-Carlo "experiments", each corresponding to 63 fb^{-1} (i.e. 465 TM events). The Physics Sample of each experiment is then analyzed using tFit and the p.d.f.'s used for the generation of events, with N_{TM}, N_{Argus}, N_{Peak} and χ as floating parameters. For illustration, an example of toy experiment is shown in Fig.(4.77).

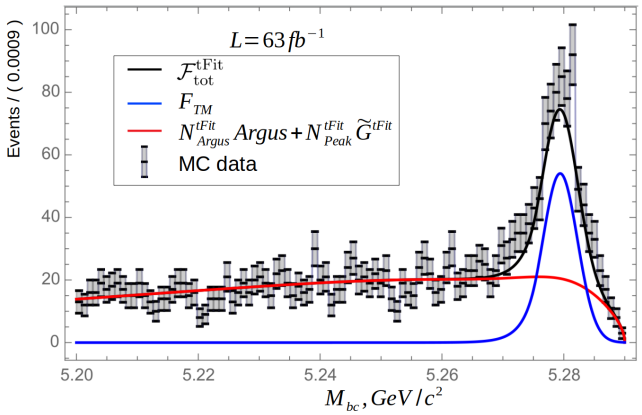


Figure 4.77. An example of a toy experiment corresponding to a luminosity of 63 fb^{-1}. For this experiment one has $N_{\mathrm{TM}} = 442$, $N_{\mathrm{Argus}} = 1730$, and $N_{\mathrm{Peak}} = 61$, while the tFit results in $N_{\mathrm{TM}}^{\mathrm{tFit}} = 442.8 \pm 47.3$, $N_{\mathrm{Argus}}^{\mathrm{tFit}} = 1718.3 \pm 52.6$, $N_{\mathrm{Peak}}^{\mathrm{tFit}} = 71.9 \pm 62.9$, and $\chi^{\mathrm{tFit}} = -40.3 \pm 3.8$. The complete distribution and the TM and background contributions are indicated respectively by the black, blue and red line.

For each experiment, the following quantities derived from the covariance matrix are computed

$$\delta = \frac{\sigma^{\mathrm{tFit}}[N_{\mathrm{TM}}^{\mathrm{tFit}}]}{\sqrt{N_{\mathrm{TM}}^{\mathrm{tFit}}}} \tag{4.51}$$

$$\delta_{\chi} = \frac{\sigma^{\mathrm{tFit}}[N_{\mathrm{TM}}^{\mathrm{tFit}}]}{\sigma^{\mathrm{tFit}}[N_{\mathrm{TM}}^{s\mathcal{P}lot}]} \tag{4.52}$$

$$z = \frac{N_{\mathrm{TM}}^{\mathrm{tFit}} - N_{\mathrm{TM}}}{\sigma^{\mathrm{tFit}}[N_{\mathrm{TM}}^{\mathrm{tFit}}]} \tag{4.53}$$

The lowest possible value of the first quantity (below it is referred to as the dilution) is unity. But the presence of backgrounds, and more generally the presence of nuisance parameters in the fit amplify the dilution. The dilution δ weekly depends on the available luminosity, and its spread is quite limited: therefore, it is convenient to evaluate the accuracy expected for a given luminosity.
The second quantity δ_{χ} is the ratio between the tFit statistical uncertainty on N_{TM} and the statistical uncertainty obtained when the nuisance parameter χ is fixed to the value obtained from tFit (it is used to apply the ${}_s\mathcal{P}lot$ method as in section (4.5.6)). This quantity measures the dilution due to χ, which is needed in section (4.6).

From the experiment distributions one obtains:

$$\langle\delta\rangle \;=\; 2.23 \quad ; \quad \sigma[\delta] = 0.12 \quad ; \quad \langle\delta_\chi\rangle = 1.030 \quad ; \quad \sigma[\delta_\chi] = 0.003 \tag{4.54}$$

The average statistical uncertainties and correlation coefficients are quoted in Table (4.14).

parameter	$N_{\rm TM}$	$N^{\rm tFit}_{\rm Argus}$	$N^{\rm tFit}_{\rm Peak}$	$\chi^{\rm tFit}$
$N_{\rm TM}$	48.2	+0.217	-0.776	+0.237
$N^{\rm tFit}_{\rm Argus}$		52.8	-0.488	+0.460
$N^{\rm tFit}_{\rm Peak}$			63.8	-0.556
$\chi^{\rm tFit}$				4.0

Table 4.14. Average values of the statistical uncertainties and correlation coefficients obtained with 10^3 toy Monte-Carlo experiments. The diagonal terms are the statistical uncertainties for $\mathcal{L} = 63$ fb^{-1}. The off-diagonal terms are the correlation coefficients. To ease the reading, only the upper part of the symmetric matrix is shown.

The distribution of the third quantity (the pull) is ideally expected to be described by a standard Gaussian of mean zero and standard deviation unity, at least for large enough statistics. The mean and the standard deviation of the experiment values are found to be:

$$\langle z\rangle \;=\; +0.01 \pm 0.03 \quad \text{and} \quad \sigma[z] = 0.975 \pm 0.02. \tag{4.55}$$

Hence, for 63 fb^{-1}, the likelihood estimator is not significantly biased.

The histogram of the pull is shown in Fig.(4.78) superimposed with two Gaussians, one standard Gaussian ($\langle z\rangle = 0$; $\sigma[z] = 1$) and one Gaussian with the values of Eq.(4.55).

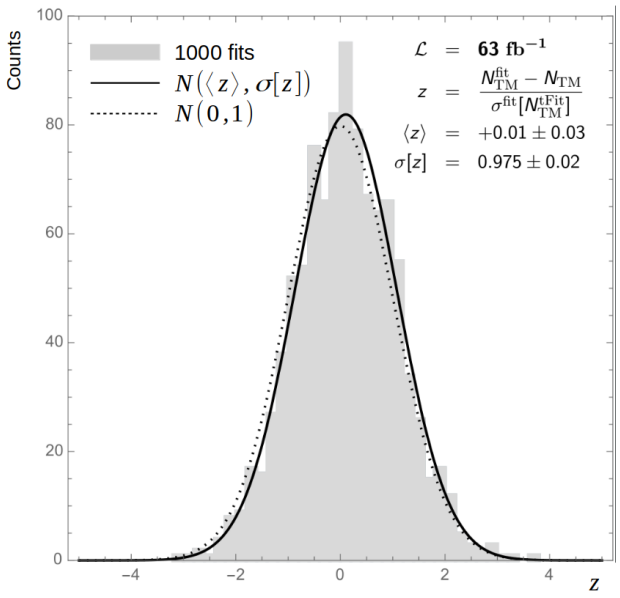


Figure 4.78. The pull (cf. Eq.(4.53)) distribution obtained with 10^3 toy Monte-Carlo experiments. N stands for Gaussian distribution.

Asymptotic behavior

For large enough statistics, the asymptotic covariance matrix can be computed analytically without having recourse to toy Monte-Carlo simulation, but using the same approximation: to describe the analysis using the same toy framework. The covariance matrix is obtained using :

$$V_{ij}^{-1} = \sum_{e=1}^{N} \frac{\partial \mathcal{F}}{\partial x_i} \frac{\partial \mathcal{F}}{\partial x_j} \frac{1}{\mathcal{F}^2} \simeq N \int \frac{\partial \mathcal{F}}{\partial x_i} \frac{\partial \mathcal{F}}{\partial x_j} \frac{1}{\mathcal{F}^2}\ \hat{\mathcal{F}}_0(M_{\mathrm{bc}})\ \mathrm{d}M_{\mathrm{bc}} \tag{4.56}$$

where

- $\mathcal{F} = \mathcal{F}_{\mathrm{tot}}^{\mathrm{tFit}}$ (cf. Eq.(4.46)) is the distribution of events that depends on the fit parameters x_i (N_{TM}, etc.),
- $\mathcal{F}$ and its derivatives are computed for $x_i = x_i^{\mathrm{tFit}}$, the fit outputs,
- $\hat{\mathcal{F}}_0$ is the p.d.f. used in the Monte-Carlo generation of events (it is normalized to unity).
- The fit output x_i^{tFit} are obtained by maximizing the extended likelihood:

$$\begin{aligned} \mathcal{L}(x_i) &= \sum_{e=1}^{N} \ln \mathcal{F}(M_{\mathrm{bc}}(e)) - N^{\mathrm{tFit}} \\ &\simeq N \int \ln \mathcal{F}(M_{\mathrm{bc}})\ \hat{\mathcal{F}}_0(M_{\mathrm{bc}})\ \mathrm{d}M_{\mathrm{bc}} - N^{\mathrm{tFit}} \end{aligned} \tag{4.57}$$

where $N^{\rm tFit} = N^{\rm tFit}_{\rm TM} + N^{\rm tFit}_{\rm Argus} + N^{\rm tFit}_{\rm Peak}$ is the sum of the fitted yields. Asymptotically it leads to $x_i^{\rm tFit} = x_i^0$, if the generation p.d.f. $\hat{\mathcal{F}}_0$ is chosen identical to $\hat{\mathcal{F}}(x_i^0)$, since the likelihood provides asymptotically-unbiased estimators. Therefore, this step is not needed for the current study, but it was followed as a cross-check.

The asymptotic values of the correlation coefficients and the statistical uncertainties (for $1\ {\rm ab}^{-1}$) obtained using the above equations are quoted in Table (4.15). The numerical values of the correlations coefficients are found identical to the ones of Table (4.13) which were obtained by performing tFit on the $1\ {\rm ab}^{-1}$ Monte-Carlo sample of events. This is expected owing to the large statistics entering into the fit for $1\ {\rm ab}^{-1}$.
In effect, the asymptotic values of the dilutions (cf. Eq.(4.51)) are $\delta = 1.72$ and $\delta_\chi = 1.008$ very close to the values of Eq.(4.54), obtained with $63\ {\rm fb}^{-1}$.

parameter	$N^{\rm tFit}_{\rm TM}$	$N^{\rm tFit}_{\rm Argus}$	$N^{\rm tFit}_{\rm Peak}$	$\chi^{\rm tFit}$
$N^{\rm tFit}_{\rm TM}$	190.7	+0.216	-0.776	+0.237
$N^{\rm tFit}_{\rm Argus}$		209.8	-0.487	+0.458
$N^{\rm tFit}_{\rm Peak}$			254.1	-0.557
$\chi^{\rm tFit}$				1.0

Table 4.15. Asymptotic values of the statistical uncertainties and correlation coefficients. The diagonal terms are the statistical uncertainties, expected for $\mathcal{L} = 1\ {\rm ab}^{-1}$. Their values scale as $1/\sqrt{\mathcal{L}}$ for the yields and as $\sqrt{\mathcal{L}}$ for χ, with the result that for $\mathcal{L} = 63\ {\rm fb}^{-1}$ the yield uncertainties are four times smaller ($1/\sqrt{0.063} = 3.98$), while the χ uncertainty is four times larger. The off-diagonal terms are the correlation coefficients; they do not depend on $\mathcal{L}$. To ease the reading, only the upper part of the symmetric matrix is shown. The dilution factor δ_χ can be computed from the correlation coefficient $\rho_{14} = +0.237$ between $N_{\rm TM}$ and χ as $\delta_\chi = 1/\sqrt{1-\rho_{14}^2} = 1.029$.

Potential systematical bias

The toy Monte-Carlo technique can also be used to assess the bias obtained if one analyzes events with a likelihood that is not based on the true distribution of events. The purpose of this section is to probe the systematical bias due to the p.d.f. 's functional chosen to describe the peaking background. The default functional used in tFit (referred to as the "modified Gaussian" of Eq.(4.43)) was chosen because of its simplicity and its ability to describe the whole background $M_{\rm bc}$ distribution once combined with an Argus function. But there is nothing special about this default functional: one could use any function that is peaking in the signal region, vanishing at $M_{\rm bc} = 5.29\ {\rm GeV/c^2}$, and describing the whole background $M_{\rm bc}$ distribution once combined with an Argus function.
We use the second functional defined in Eq.(8), which we refer to as the modified Gaussian of the second type. The parameters of this function are adjusted on the Monte-Carlo as is done for the default modified Gaussian, with the resulting p.d.f. 's shown in Figs. (4.79)-(4.80).

From 10^3 toy Monte-Carlo experiments produced using the modified Gaussian of the second type but analyzed with the default modified Gaussian one obtains $\langle z \rangle = +0.07 \pm 0.03$, which corresponds to a bias of 3 events for $63\ \mathrm{fb}^{-1}$.

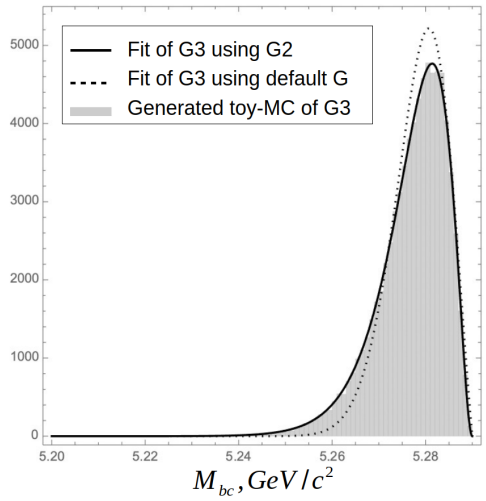


Figure 4.79. The M_{bc} distributions of the two modified Gaussians; default (dotted line) and second type (full line). The grey histogram is a toy Monte-Carlo generation of M_{bc} using the modified Gaussian of the third type.

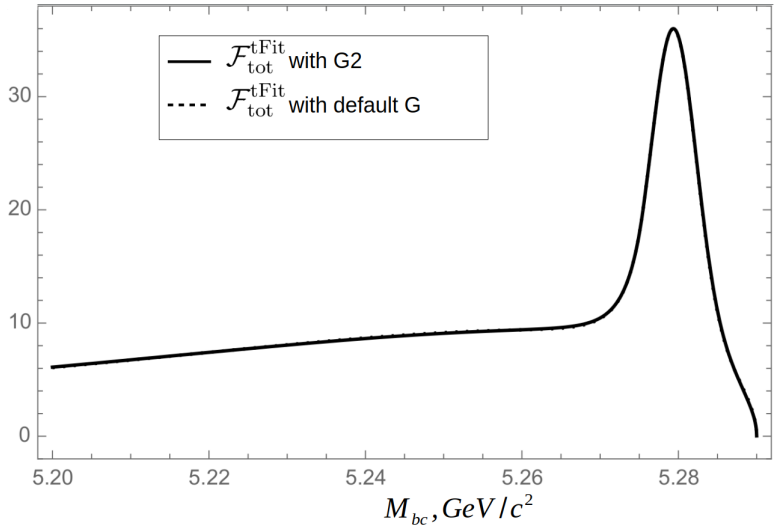


Figure 4.80. The complete M_{bc} distributions $\mathcal{F}$ as adjusted, using the default modified Gaussian (dotted line), and the modified Gaussian of second type (full line). The two PDF's appear undistinguishable.

4.5.8 . tFit with CS1 control sample

The tFit discussed in the Section 4.5.5 provides a correct determination of N_{TM} in presence of a peaking background. However, it appears useful to refine the procedure with a guardrail meant to further protect against the peaking background. To achieve this, assuming that the CS1 and CS2 Monte-Carlo description has been validated by data, the CS1 control sample is proposed to be explicitly included in the baseline fit.

The CS1 events are described by the distribution

$$\mathcal{F}^{\mathrm{CS1}} \quad = \quad N^{\mathrm{CS1}}_{\mathrm{Argus}} \mathrm{Argus}^{\mathrm{CS1}} + \mathrm{N}^{\mathrm{CS1}}_{\mathrm{Peak}} \tilde{\mathrm{G}}^{\mathrm{CS1}} \tag{4.58}$$

The parameters of the modified Gaussian p.d.f are first adjusted on the CS1 peaking background (i.e. ignoring $q\bar{q}$ events). These parameters are then fixed to perform a second fit using all CS1 events in order to obtain the χ parameter describing the overall $\mathrm{Argus}^{\mathrm{CS1}}$ p.d.f. together with $N^{\mathrm{CS1}}_{\mathrm{Peak}}$. Besides the yields, the three parameters needed to specify the CS1 pdf's (χ for the Argus function, σ and μ for the modified Gaussian) are found to be identical within errors with the parameters used in tFit. While this property is not used in the baseline fit, it provides strong support for the approach of combining Physics Sample and CS1 into a single fit.

Having described independently CS1, one defines the ratio :

$$\rho \quad = \quad \frac{N_{\rm Peak}^{\rm CS1}}{N_{\rm Peak}} = 0.46 \pm 0.02 \tag{4.59}$$

where $N_{\rm Peak}$ is obtained as in the previous section, on the Monte-Carlo Physics Sample corresponding to the same luminosity as the above Monte-Carlo CS1 (i.e. $1\ {\rm ab}^{-1}$ presently).
The combined fit uses the above prediction for the ρ ratio to enforce proportionality between the floating numbers of peaking background in the Physics Sample and of peaking background in the CS1 sample: $N_{\rm Peak} = \rho\ N_{\rm Peak}^{\rm CS1}$. To be able to use ${}_s\mathcal{P}lot$, the yield entering the fit should be the sum

$$\hat{N}_{\rm Peak} \quad = \quad N_{\rm Peak} + N_{\rm Peak}^{\rm CS1} = (1+\rho)\ N_{\rm Peak} \tag{4.60}$$

and p.d.f should be defined as following:

$$\begin{aligned} \mathcal{F}_{\rm tot}^{\rm tFit+CS1} \quad &= \quad {\rm y} N_{\rm TM} F_{\rm TM} + {\rm y} N_{\rm Argus}^{\rm tFit} {\rm Argus}^{\rm tFit} + \\ &+ \quad (1-{\rm y}) N_{\rm Argus}^{\rm CS1} {\rm Argus}^{\rm CS1} + \hat{\rm N}_{\rm Peak} \left({\rm y} \frac{\tilde{\rm G}^{\rm tFit}}{1+\rho} + (1-{\rm y}) \frac{\rho \tilde{\rm G}^{\rm CS1}}{1+\rho} \right) \end{aligned} \tag{4.61}$$

where ${\rm y} \in [0;1]$ — discrete variable, which equals 1 for Physics Sample and 0 for CS1 such that $\mathcal{F}_{\rm tot}^{\rm tFit+CS1}({\rm y}=1) = \mathcal{F}_{\rm tot}^{\rm tFit}$ and $\mathcal{F}_{\rm tot}^{\rm tFit+CS1}({\rm y}=0) = \mathcal{F}^{\rm CS1}$.

For data analysis, the combined fit should include the statistical and systematical uncertainty on ρ, but these are omitted at this stage. The tFit can be recovered by using a very large systematical uncertainty, thereby letting $N_{\rm Peak}$ and $N_{\rm Peak}^{\rm CS1}$ free to vary independently.

The baseline fit leads to a satisfactory combined description of the Physics Sample and CS1, as shown in Figs.(4.81)-(4.82)

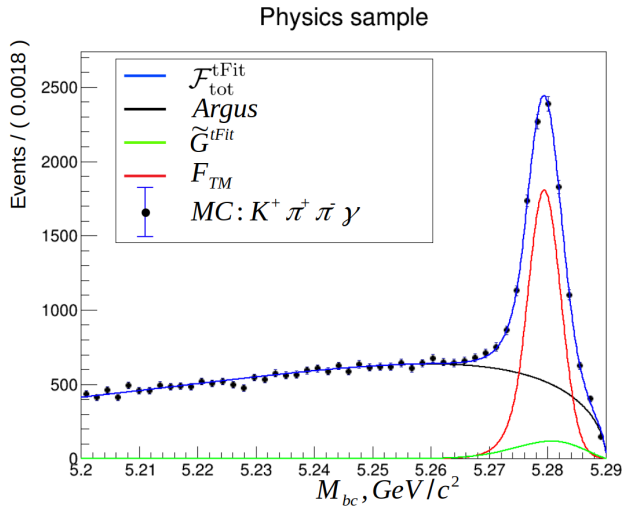


Figure 4.81. Baseline fit results for the Physics Sample.

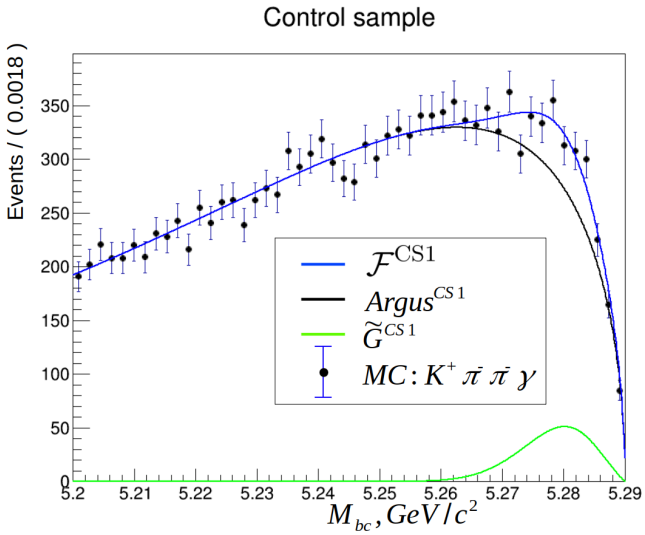


Figure 4.82. Baseline fit results for CS1.

The fit results are given in Tables (4.16)-(4.17). In particular the result for TM events $N_{\rm TM} = 7384 \pm 152$ is fairly consistent with the expected value $N_{\rm TM} = 7386$. Although it is not the motivation for its introduction, this fit leads to a 25% reduction of the statistical uncertainty with respect to tFit.

parameter	value	σ
$N_{\rm TM}$	7384	152
$N_{\rm Argus}^{\rm tFit}$	26758	192
$\chi^{\rm tFit}$	-43.2	0.9
$\hat{N}_{\rm Peak}$	13224	226
$N_{\rm Argus}^{\rm CS1}$	13504	136
$\chi^{\rm CS1}$	-32.83	1.5

Table 4.16. Parameters values resulting from the fit. The first three values refer to the Physics Sample and the last two values refer to CS1. The value $\hat{N}_{\rm Peak}$ is the total number of peaking background present in the Physics Sample and CS1.

parameter	$N_{\rm TM}$	$N_{\rm Argus}^{\rm tFit}$	$\chi^{\rm tFit}$	$\hat{N}_{\rm Peak}$	$N_{\rm Argus}^{\rm CS1}$	$\chi^{\rm CS1}$
$N_{\rm TM}$	-	-0.03	+0.05	-0.62	+0.30	-0.37
$N_{\rm Argus}^{\rm tFit}$		-	-0.33	-0.30	+0.15	-0.18
$\chi^{\rm tFit}$			-	+0.36	-0.18	+0.21
$\hat{N}_{\rm Peak}$				-	-0.49	+0.59
$N_{\rm Argus}^{\rm CS1}$					-	-0.32
$\chi^{\rm CS1}$						-

Table 4.17. Correlation matrix of the fit. To ease the reading, only the upper part of the symmetric matrix is shown. The correlation between $N_{\rm TM}$ and $\hat{N}_{\rm Peak}$ is smaller on $\sim 20\%$ comparing to the case of $N_{\rm TM}$ and $N_{\rm Peak}$ in the Table 4.13.

Although results obtained with this fit look promising, at this stage of the Belle II Experiment amount of data available is not as huge as is used in the tFit+CS1 and it has been decided to move on with the baseline fit.

4.6 . Branching ratio measurement

Having obtained the yield for 63 fb^{-1}: $N_{\rm TM}^{\rm tFit} = 465 \pm 48$, a first estimate of the branching ratio of the decay $B^+ \to K^+\pi^+\pi^-\gamma$ is given by:

$$\mathrm{Br}_{K^+\pi^+\pi^-\gamma} = \frac{N_{\rm TM}^{\rm tFit}}{\varepsilon_{\rm TM}^{\rm tot}} \frac{1}{2 \cdot N_{B^+B^-}} = 2.65 \cdot 10^{-5} \tag{4.62}$$

with

$$N_{B^+B^-} = \mathcal{L} \cdot \sigma_{\Upsilon(4s)} \cdot \mathrm{Br}_{B^+B^-} = 0.0627 \times 10^9 \times 1.110 \times 0.514 = 35.8 \cdot 10^6 \tag{4.63}$$

$$\tag{4.64}$$

where

- the value $\varepsilon_{\rm TM}^{\rm tot} = 0.245(0.003)$ is taken from Eqs.(4.30-4.31);
- the systematical uncertainties remain to be evaluated.

The statistical uncertainty on the branching ratio of Eq.(4.62) is given by:

$$\sigma[\mathrm{Br}_{K^+\pi^+\pi^-\gamma}] = \frac{\sigma[N_{\rm TM}^{\rm tFit}]}{\varepsilon_{\rm TM}^{\rm tot}} \frac{1}{2 \cdot N_{B^+B^-}} \tag{4.65}$$

$$= 0.27 \cdot 10^{-5} \tag{4.66}$$

In the Eq. 4.62 the systematics is expected to be weak since it has been checked for the absence of correlation between μ_{NN} and $M_{K\pi\pi}$. For instance, thrust of signal B exhibits strong correlation with $M_{K\pi\pi}$ on the Fig. 4.83. Consider the case when T_B is included in the Neural Net: it leads to a strong correlation between μ_{NN} and $M_{K\pi\pi}$. Applying cut on μ_{NN} means removing more events in one kinematical region and less in another. The effect introduces dependence of selection efficiency of μ_{NN} cut, and hence total $\varepsilon_{\rm TM}^{\rm tot}$ efficiency on kinematics on the Fig. 4.84. However, once the thrust of signal B variable is removed from the list of neural net's input variables, dependence is strongly suppressed (see Fig. 4.85 for comparison). Three-dimensional binning is explained below.

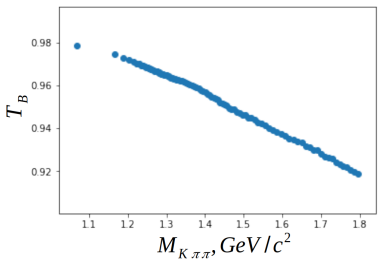


Figure 4.83. Mean value of thrust as a function of $M_{K\pi\pi}$: high correlation is observed.

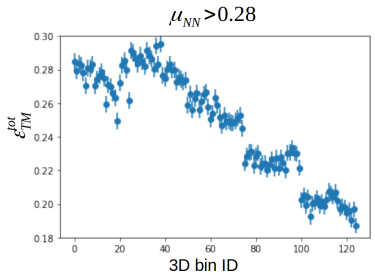


Figure 4.84. Total efficiency as a function of 3D bin, where thrust of signal B is inside the neural net

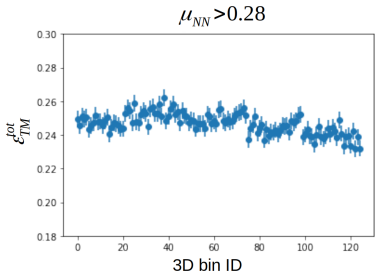


Figure 4.85. Total efficiency as a function of 3D bin, where thrust of signal B is not inside the neural net

The kinematical space used for branching ratio measurement is restricted by the following cut $M_{K^+\pi^+\pi^-} < 1.8$ GeV/c^2 introduced in Section 4.2 and defined in terms of Dalitz variables $M_{K^+\pi^+\pi^-}, s_{K^+\pi^-}, s_{\pi^+\pi^-}$. The Dalitz variables $s_{K^+\pi^-}$ and $s_{\pi^+\pi^-}$ can be further transformed to another set of variables $\theta_{K^+\pi^-}, \theta_{\pi^+\pi^-}$ (see Eqs. 4.67-4.68), which have numerical range $[0;\pi]$ independent on $M_{K\pi\pi}$ range. Their 2D distribution is more flat than 2D distribution of $s_{K^+\pi^-}, s_{\pi^+\pi^-}$ (see Figs. 4.86-4.87).

$$\theta_{K^+\pi^-} = \arccos\left(\frac{m_{K^+}^2 + m_{\pi^-}^2 + 2E_{K^+}E_{\pi^-}}{2|\overrightarrow{p}_{K^+}| \cdot |\overrightarrow{p}_{\pi^-}|}\right) \tag{4.67}$$

$$\theta_{\pi^+\pi^-} = \arccos\left(\frac{m_{\pi^+}^2 + m_{\pi^-}^2 + 2E_{\pi^+}E_{\pi^-}}{2|\overrightarrow{p}_{\pi^+}| \cdot |\overrightarrow{p}_{\pi^-}|}\right) \tag{4.68}$$

where 4-momenta of K^+, π^+, π^- are calculated in the $K\pi\pi$ rest frame.

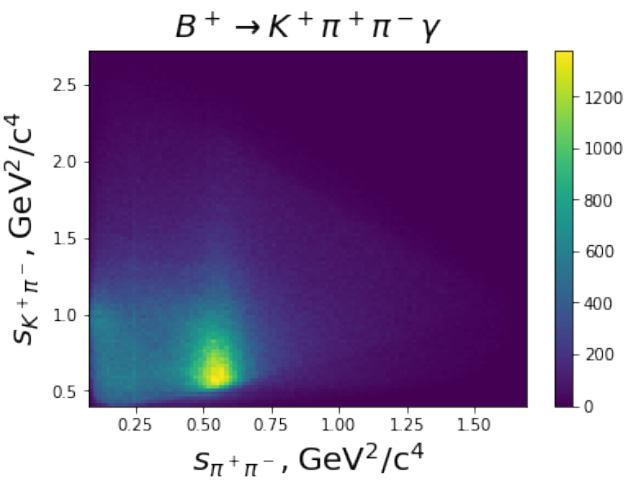


Figure 4.86. 2D histogram of Dalitz variables for $B^+ \to K^+\pi^+\pi^-\gamma$ events (TM+FM) after preselection cuts and $M_{K^+\pi^+\pi^-} < 1.8$ GeV/c^2: large excess of events on small area is observed.

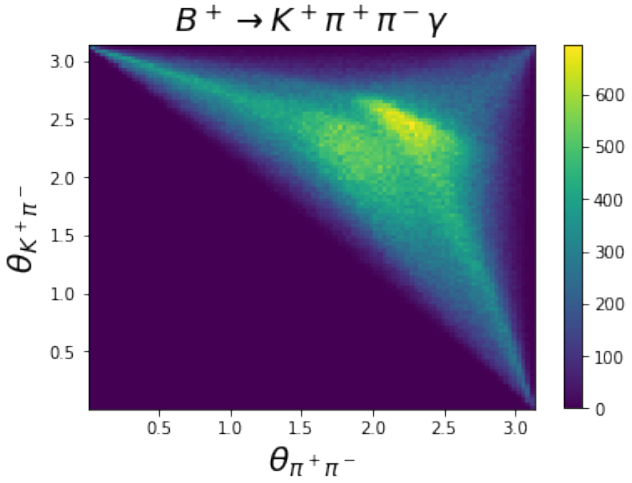


Figure 4.87. 2D histogram of $\theta_{K^+\pi^-}, \theta_{\pi^+\pi^-}$ for $B^+ \to K^+\pi^+\pi^-\gamma$ events (TM+FM) after preselection cuts and $M_{K^+\pi^+\pi^-} < 1.8$ GeV/c^2: the distribution visually looks more flat and less events are within the peak.

The 3-dimensional space can be split in 3D bins such that each bin contains equal amount of TM candidates and is represented by rectangular parallelepiped in $(M_{K^+\pi^+\pi^-}, \theta_{K^+\pi^-}, \theta_{\pi^+\pi^-})$ coordinates. The splitting is done for a large sample (corresponding to 30 ab^{-1} integrated luminosity) of TM candidates obtained after preselection cuts and $M_{K\pi\pi} < 1.8$ GeV/c^2 cut satisfied. Along each dimension of kinematical phase space 5 bins are selected giving in total $5 \times 5 \times 5 = 125$ 3D bins. The examples of 2D binnings in $\theta_{K^+\pi^-}, \theta_{\pi^+\pi^-}$ for several bins (slices) of $M_{K^+\pi^+\pi^-}$ are given on Fig. 4.88.

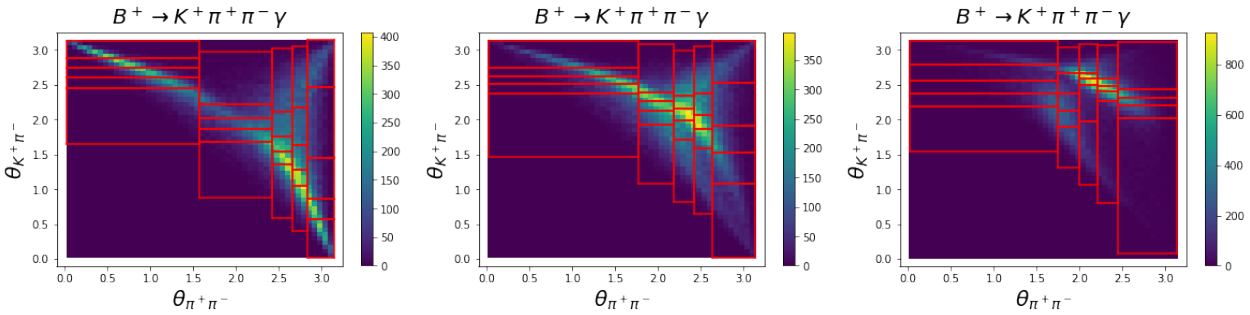


Figure 4.88. Examples of 2D binning for $\theta_{K^+\pi^-}, \theta_{\pi^+\pi^-}$ distribution in $0.97 < M_{K^+\pi^+\pi^-} < 1.28$ GeV/c^2 (left), $1.34 < M_{K^+\pi^+\pi^-} < 1.45$ GeV/c^2 (center) and $1.61 < M_{K^+\pi^+\pi^-} < 1.8$GeV/c^2 (right) slices. The area of the 2D bin is larger for the regions with lower events density. While distribution is changed from slice to slice, 2D bins are changed accordingly.

Denoting $i \in [1; N_{M_{K^+\pi^+\pi^-}} = 5], j \in [1; N_{\theta_{K^+\pi^-}} = 5], k \in [1; N_{\theta_{\pi^+\pi^-}} = 5]$ as indices of 1D bins in $M_{K^+\pi^+\pi^-}, \theta_{K^+\pi^-}, \theta_{\pi^+\pi^-}$ coordinates, 3D bin ID is obtained as following:

$$\text{3D bin ID} = k + (j-1) \cdot N_{\theta_{\pi^+\pi^-}} + (i-1) \cdot N_{\theta_{\pi^+\pi^-}} \cdot N_{\theta_{K^+\pi^-}} \tag{4.69}$$

The above estimate of branching ratio ignores the fact that the selection efficiency varies across the kinematical phase space, with the result that the data distribution in phase-space should be accounted for. This is achieved by replacing Eq.(4.62) by :

$$\mathrm{Br}_{K^+\pi^+\pi^-\gamma} = \sum_{\mathrm{3DbinID}} \frac{\sum_{e:\mathrm{3DbinID}} {}_s\mathcal{P}_{\mathrm{TM}}(e)}{\varepsilon^{\mathrm{tot}}_{\mathrm{TM}}(\mathrm{3DbinID})} \frac{1}{2 \cdot N_{B^+B^-}} \tag{4.70}$$

where

- the first sum runs over an appropriate binning of the phase-space,
- the second sum runs over the events e in a given phase-space bin,
- ${}_s\mathcal{P}_{\mathrm{TM}}(e)$ is the ${}_s\mathcal{P}lot$ weight for TM, corresponding to event e,
- $\varepsilon^{\mathrm{tot}}_{\mathrm{TM}}(\mathrm{3DbinID})$ is the selection efficiency in a given bin.

The statistical uncertainty on the branching ratio of Eq.(4.70) is given by:

$$\sigma[\mathrm{Br}_{K^+\pi^+\pi^-\gamma}] = \delta_\chi \sqrt{\sum_{\mathrm{3DbinID}} \sum_{e:\mathrm{3DbinID}} \left(\frac{{}_s\mathcal{P}^2_{\mathrm{TM}}(e)}{\varepsilon^{\mathrm{tot}}_{\mathrm{TM}}(\mathrm{3DbinID})} \right)^2} \frac{1}{2 \cdot N_{B^+B^-}} \tag{4.71}$$

$$\delta_\chi = \frac{\sigma[N^{\mathrm{tFit}}_{\mathrm{TM}}]}{\sigma^{\mathrm{fit}}[N^{{}_s\mathcal{P}lot}_{\mathrm{TM}}]} = \frac{\sigma[N^{\mathrm{tFit}}_{\mathrm{TM}}]}{\sqrt{\sum_e {}_s\mathcal{P}^2_{\mathrm{TM}}(e)}} = 1.03 \tag{4.72}$$

where the δ_χ value is taken from Eq.(4.54). If the bin dependence of $\epsilon^{\mathrm{tot}}_{\mathrm{TM}}(\mathrm{3DbinID})$ is ignored, Eq.(4.71) is identical to Eq.(4.65).

Selection efficiency is calculated per 3D bin as following:

$$\varepsilon^{\mathrm{tot}}_{\mathrm{TM}}(\mathrm{3DbinID}) = \varepsilon^{\mathrm{presel}}_{\mathrm{TM}} \varepsilon^{\mathrm{presel}:\mu_{\mathrm{NN}}}_{\mathrm{TM}}(\mathrm{3DbinID}) \tag{4.73}$$

$$\varepsilon^{\mathrm{presel}:\mu_{\mathrm{NN}}}_{\mathrm{TM}}(\mathrm{3DbinID}) = \frac{N_{\mu^{\mathrm{TM}}_{NN}}(\mathrm{3DbinID})}{N^{\mathrm{TM}}_{presel}(\mathrm{3DbinID})} \tag{4.74}$$

where $\varepsilon^{\mathrm{presel}}_{\mathrm{TM}}$ — product of reconstruction, skimming and preselection efficiencies for TM events listed in Table 4.1; $\varepsilon^{\mathrm{presel}:\mu_{\mathrm{NN}}}_{\mathrm{TM}}$ — efficiency of TM computed between preselection cuts and μ_{NN} cut; $N_{\mu^{\mathrm{TM}}_{NN}}$ — number of TM events obtained after applying the full selection procedure to the sample of $30\ \mathrm{ab}^{-1}$ of $B^+ \to K^+\pi^+\pi^-\gamma$ events; $N^{\mathrm{TM}}_{presel}(\mathrm{3DbinID})$ — number of TM events obtained after applying all cuts up-to the selection cuts level and corresponding to $30\ \mathrm{ab}^{-1}$.

Since $\varepsilon_{\rm TM}^{\rm presel}$ and $\frac{N_{\mu_{NN}^{\rm TM}}(\rm 3DbinID)}{N_{presel}^{\rm TM}(\rm 3DbinID)}$ are calculated independently for different integrated luminosity levels ($1\ \rm ab^{-1}$ and $30\ \rm ab^{-1}$ correspondingly) variance of $\varepsilon_{\rm TM}^{\rm tot}(\rm 3DbinID)$ is computed as following:

$$Var[\varepsilon_{\rm TM}^{\rm tot}(\rm 3DbinID)] = (\varepsilon_{\rm TM}^{\rm presel:\mu_{NN}}(\rm 3DbinID))^2\frac{\varepsilon_{\rm TM}^{\rm presel}(1-\varepsilon_{\rm TM}^{\rm presel})}{N} + \quad (4.75)$$

$$+ (\varepsilon_{\rm TM}^{\rm presel})^2\frac{\varepsilon_{\rm TM}^{\rm presel:\mu_{NN}}(\rm 3DbinID)(1-\varepsilon_{\rm TM}^{\rm presel:\mu_{NN}}(\rm 3DbinID))}{N_{presel}^{\rm TM}(\rm 3DbinID)} \quad (4.76)$$

The dominant contribution is coming from the second term corresponding to the statistical uncertainty of $\varepsilon_{\rm TM}^{\rm presel:\mu_{NN}}(\rm 3DbinID)$ in the global 3D bin. On the Fig. 4.89 statistical uncertainties are very small comparing to the mean value of $\varepsilon_{\rm TM}^{\rm tot}$.

Meanwhile ${}_s\mathcal{P}_{\rm TM}(e)$ and ${}_s\mathcal{P}^2_{\rm TM}(e)$ have been already computed when perfoming ${}_s\mathcal{P}lot$ on $1\ \rm ab^{-1}$ of MC sample in Section 4.5.

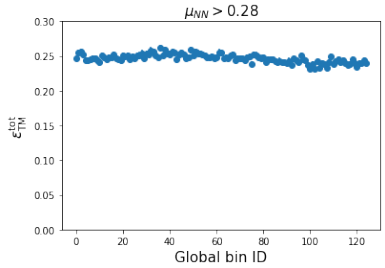


Figure 4.89. Efficiency as a function of 3D bin number.

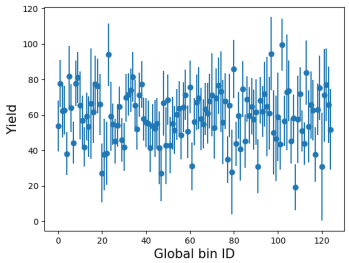


Figure 4.90. Yield as a function of bin number.

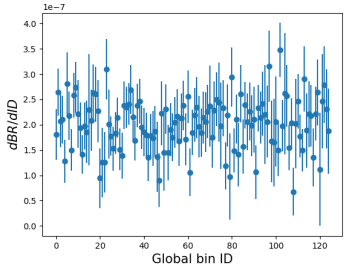


Figure 4.91. Differential branching ratio as a function of phase space.

From Fig. 4.89 one may conclude that weak dependency on the phase-space is present. From Fig. 4.90 one can see that binning has been chosen with a target of uniform splitting of the phase-space.

Differential branching ratio as a function of global bin ID (phase space) is given in Fig. 4.91 and doesn't depend on phase space with binning applied. The dominant source of statistical uncertainly is coming from yields measurements for each bin, for this reason, statistical uncertainty of efficiency has been neglected in the Eq. 4.71.

Hence, using Eqs. 4.70-4.71 the branching ratio can be computed:

$$\rm Br_{K^+\pi^+\pi^-\gamma} = 2.52 \pm 0.07 \cdot 10^{-5} \quad (4.77)$$

Obtained values of branching ratios using both $63\ fb^{-1}$ and $1\ ab^{-1}$ from Eqs. 4.62,4.77 are in a good agreement with input PDG value $2.58 \cdot 10^{-5}$ given in the Section. 4.1. Computed uncertainties from Eqs. 4.65,4.71 are in very nice agreement as well: $\frac{0.27\cdot 10^{-5}}{\sqrt{\frac{1000\ fb^{-1}}{63\ fb^{-1}}}} \simeq 0.07 \cdot 10^{-5}$.

As an illustration on the Figs. 4.92-4.95 efficiencies as a functions of phase space variables are given for $M_{K^+\pi^+\pi^-}$, $M_{K^+\pi^-}$, $M_{\pi^+\pi^-}$, $\cos\theta_\gamma$. One can see that efficiencies are flat.

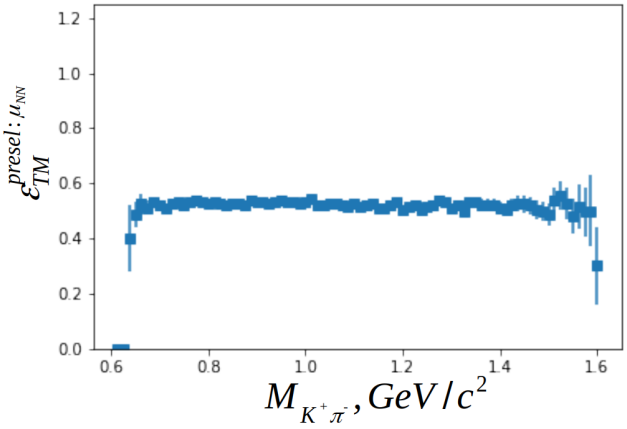


Figure 4.92. Efficiency as a function of $M_{K\pi}$

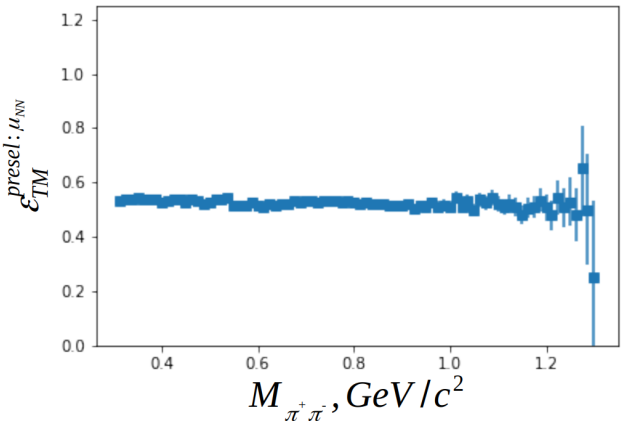


Figure 4.93. Efficiency as a function of $M_{\pi\pi}$.

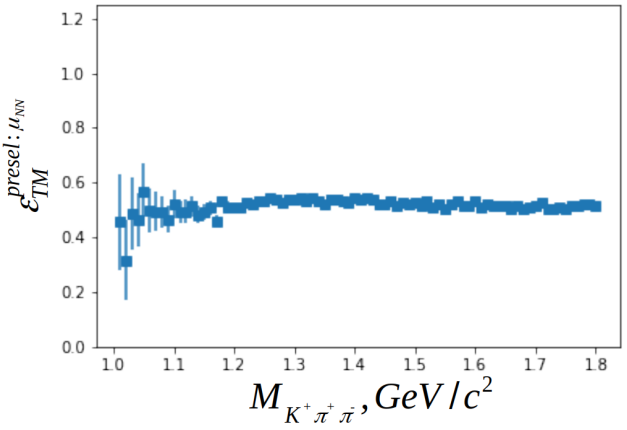


Figure 4.94. Efficiency as a function of $M_{K\pi\pi}$

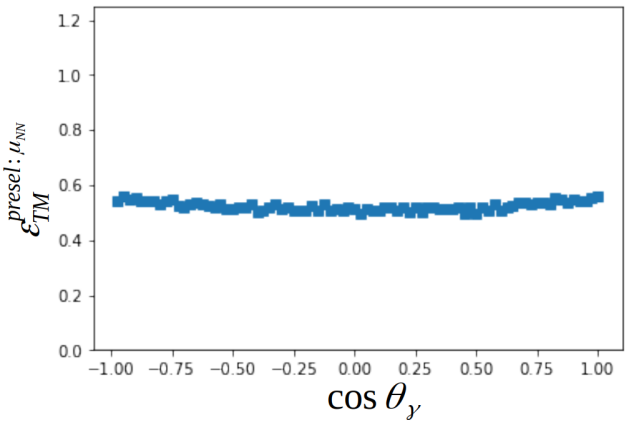


Figure 4.95. Efficiency as a function of $\cos\theta_\gamma$.

4.6.1 . Correction factor calculation

While performing cuts efficiency comparison between two data samples it is possible that the same cut applied on the same selection level for both these samples lead to different selection efficiencies. Such a situation may appear when comparing data and Monte-Carlo. Although it is not the treatment of the source of disagreement but allows to account for different cuts' efficiencies. Within Eqs. 4.78-4.81 the

formal definitions are given for basic measures and their statistical uncertainties.

$$\varepsilon = \frac{N_{after}}{N_{before}} \tag{4.78}$$

$$\sigma[\varepsilon] = \sqrt{\frac{\varepsilon \cdot (1-\varepsilon)}{N_{before}}} \tag{4.79}$$

$$f = \frac{N_{data}}{N_{MC}} \tag{4.80}$$

$$\sigma[f] = \sqrt{\frac{f \cdot (1+f)}{N_{MC}}} \tag{4.81}$$

$$\tag{4.82}$$

where N_{before} — number of events before a cut, N_{after} — number of events after cut, ε — selection efficiency of cut, N_{data} (N_{MC}) — number of events in the data (MC data) sample after a cut, f — disagreement coefficient:

- $f \simeq 1$ — cuts' efficiencies are in good agreement;
- $f < 1$ — cut removes more events in data, than in MC;
- $f > 1$ — cut removes more events in MC, than in data;

In the present work, f can be found by comparing off-resonance data and MC. Performing the calculations described as in Eqs. 4.83-4.86 one can find the disagreement coefficients h for any MC sample, given the selection efficiency of the cut on the truth level.

$$g = \frac{1-\varepsilon_{off} \cdot f}{1-\varepsilon_{off}} \tag{4.83}$$

$$\sigma[g] = \frac{1}{1-\varepsilon_{off}} \sqrt{\frac{(1-f)^2}{(1-\varepsilon_{off})^2}\sigma^2[\varepsilon_{off}] + \varepsilon_{off}^2 \sigma^2[f]} \tag{4.84}$$

$$h = \frac{1-(1-\varepsilon)\cdot g}{\varepsilon} \tag{4.85}$$

$$\sigma[h] = \frac{1}{\varepsilon} \sqrt{\frac{(1-g)^2}{\varepsilon^2}\sigma^2[\varepsilon] + (1-\varepsilon)^2 \sigma^2[g]} \tag{4.86}$$

where ε_{off} — selection efficiency of cut for MC off-resonance sample, g — disgreement of cuts inefficiencies, h — disgreement of cuts efficiencies (if $\varepsilon = \varepsilon_{off} \to h = f$).

After introducing the factor $h_{\pi^0}^{TM}$ total efficiency can be corrected as $\varepsilon_{\rm TM}^{\rm data} = \varepsilon_{\rm TM}^{\rm tot} \cdot h_{\pi^0}^{TM}$. Even though such correction for selection efficiency is not applied in this work, it can be useful for future studies on the topic.

4.7 . Towards photon polarization measurements

Ultimately, the photon polarization parameter can be measured by applying the fitting procedure described in Section 2.3. In this section as a starting point, up-down asymmetry is measured for probing the sensitivity to the photon polarization. According to [40], up-down asymmetry is defined as follows:

$$A_{ud} = \frac{\int_0^1 \frac{d\Gamma}{d\cos\theta} d\cos\theta - \int_{-1}^0 \frac{d\Gamma}{d\cos\theta} d\cos\theta}{\int_0^1 \frac{d\Gamma}{d\cos\theta} d\cos\theta + \int_{-1}^0 \frac{d\Gamma}{d\cos\theta} d\cos\theta} \tag{4.87}$$

$$\frac{d\Gamma}{d\cos\theta} = \frac{|\mathcal{A}_L|^2 + |\mathcal{A}_R|^2}{2} + \lambda_\gamma \frac{|\mathcal{A}_L|^2 - |\mathcal{A}_R|^2}{2} \tag{4.88}$$

$$\tag{4.89}$$

where $\mathcal{A}_{L,R}$ — left and right decay amplitudes; $\frac{d\Gamma}{d\cos\theta}$ — differential decay rate integrated over all kinematical variables except $\cos\theta$.

For instance, in the case of 1^+ resonance the differential decay rate is defined as following:

$$\frac{d\Gamma}{d\cos\theta} = A(1 + \cos^2\theta) + \lambda_\gamma B \cos\theta \tag{4.90}$$

where A, B — constants obtained after integration $\frac{d\Gamma}{d\vec{\Omega}}$ over all kinematical variables (Dalitz plot, $M_{K\pi\pi}$, ϕ) except $\cos\theta$. Substituting Eq. 4.90 into Eq. 4.87 one obtains:

$$A_{ud} = \lambda_\gamma \frac{3B}{8A} \tag{4.91}$$

which demonstrates the relation between up-down asymmetry and photon polarization parameter. In the present work according to Eq. 2.72, $\frac{d\Gamma}{d\cos\theta} = A(\theta) + \lambda_\gamma B(\theta)$. If $A(\theta)$ and $B(\theta)$ are even and odd functions of $\cos\theta$ respectively, A_{ud} is proportional to λ_γ. Although analytical check of the functional view has not been performed, A_{ud} study is related to the study of the photon polarization parameter.

4.7.1 . Up-down asymmetry measurements

The up-down assymetry A_{ud} for $B^+ \to K^+\pi^+\pi^-\gamma$ decay is measured (see Eq. 4.92) with respect to the $K\pi\pi$ decay plane (see Fig. 4.97) by counting the number of photons going "up" and "down".

$$A_{ud} = \frac{N_+ - N_-}{N_+ + N_-} \tag{4.92}$$

$$\sigma[A_{ud}] = \frac{1}{N_+ + N_-} \sqrt{(1 - A_{ud})^2 \cdot \sigma^2[N_+] + (1 + A_{ud})^2 \cdot \sigma^2[N_-]} \tag{4.93}$$

$$\tag{4.94}$$

where N_+ and N_- — numbers of photons going "up" and "down" respectively; $\theta_\gamma = \theta$ — photon's angle (see Eq. 2.6) with respect to the normal to the $K\pi\pi$ decay plane; $\sigma[N_\pm]$ — statistical uncertainties of $N_\pm$.

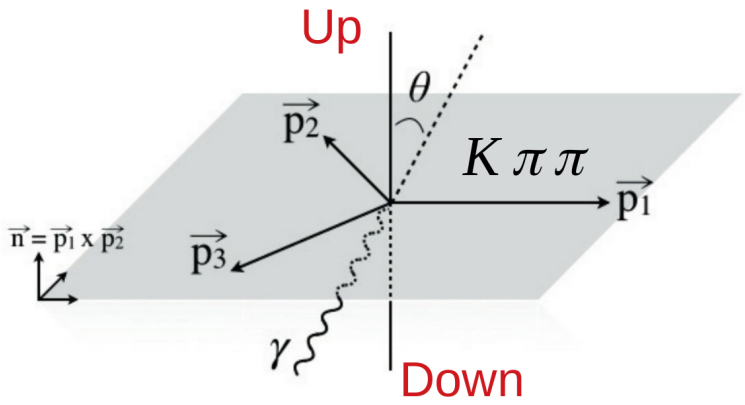


Figure 4.96. Photon direction with respect to the $K\pi\pi$ decay plane

The quantity A_{ud} is measured for every bin of $M_{K\pi\pi}$ spectrum, where the binning is the same as in the Section. 2.4.1, except for the last bin where $M_{K\pi\pi} \in [1.6; 1.9]$ GeV/c^2 has been ignored. The later fact is concluded from the studies performed in the Section. 2.4.2 and large inconsistency between the Baseline model's and LHCb distributions of $\cos\theta_\gamma$ for this bin (see Fig. 2.27).

Having access to the MC truth information, A_{ud}^{MC} and $\sigma[A_{ud}^{MC}]$ are computed for each $M_{K\pi\pi}$ bin for TM events directly using Eqs. 4.92-4.93 with $\sigma[N_\pm] = \sqrt{N_\pm}$.

On the other side after performing the extended likelihood fit of 1 ab^{-1} of MC data, $N_+^{s\mathcal{P}}$ and $\sigma[N_+^{s\mathcal{P}}]$ are obtained from ${}_s\mathcal{P}_{\rm TM}(e)$ for each bin of $M_{K\pi\pi}$:

$$N_\pm^{s\mathcal{P}} = \sum_{e:\delta M_{K\pi\pi}\cap\delta\cos\theta_\gamma^\pm} {}_s\mathcal{P}_{\rm TM}(e) \tag{4.95}$$

$$\sigma[N_+^{s\mathcal{P}}] = \sqrt{\sum_{e:\delta M_{K\pi\pi}\cap\delta\cos\theta_\gamma^\pm} {}_s\mathcal{P}^2_{\rm TM}(e)} \tag{4.96}$$

where notation $e : \delta M_{K\pi\pi} \cap \delta\cos\theta_\gamma^\pm$ means that event should fall into some $M_{K\pi\pi}$ bin $\delta M_{K\pi\pi}$ and also satisfy either $\cos\theta_\gamma < 0$ or $\cos\theta_\gamma \geq 0$. Table. 4.18 summarizes the comparison between MC-truth and ${}_s\mathcal{P}lot$ quantities obtained for each bin of $M_{K\pi\pi}$.

Bin	N_+^{MC}	$N_+^{s\mathcal{P}}$	N_-^{MC}	$N_-^{s\mathcal{P}}$	A_{ud}^{MC}	$A_{ud}^{s\mathcal{P}}$	$\sigma[A_{ud}^{MC}]$	$\sigma[A_{ud}^{s\mathcal{P}}]$	χ^2
$M_{K\pi\pi} \in [1.1; 1.3]$ GeV/c^2	1070	1079	1110	1144	-0.018	-0.029	0.021	0.032	0.21
$M_{K\pi\pi} \in [1.3; 1.4]$ GeV/c^2	804	752	785	745	0.012	0.004	0.025	0.041	0.06
$M_{K\pi\pi} \in [1.4; 1.6]$ GeV/c^2	1003	992	1109	1086	-0.05	-0.045	0.022	0.034	0.04

Table 4.18. Comparative summary of the up-down assymetry measurements between MC-truth and ${}_s\mathcal{P}lot$ values for 1 ab^{-1} of MC data (at this stage of the analysis corresponds to 7386 TM events). The fitness between MC-truth and ${}_s\mathcal{P}lot$ values of A_{ud} is defined as $\chi^2 = \frac{(A_{ud}^{s\mathcal{P}} - A_{ud}^{MC})^2}{[\sigma(A_{ud}^{s\mathcal{P}})]^2 - [\sigma(A_{ud}^{MC}]^2)}$

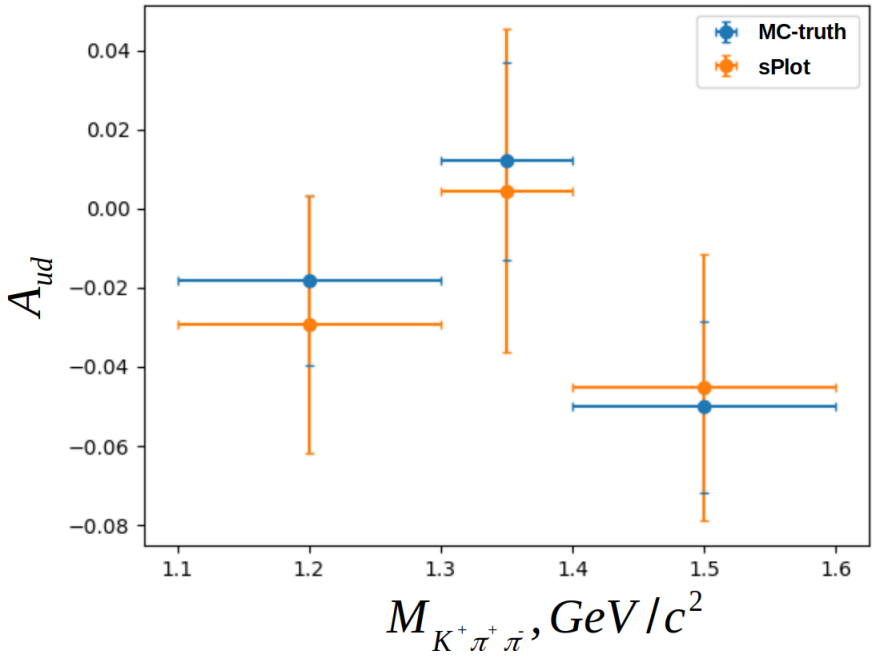


Figure 4.97. A_{ud} as a function of $M_{K\pi\pi}$.

From the Table. 4.18 and Fig. 4.97 one can see that the MC-truth and reconstructed values of up-down asymmetry are consistent.

The more granular results rather than N_+ and N_- can be obtained for $\cos\theta_\gamma$ as on Figs. 4.98-4.100. Monte-Carlo truth and reconstructed by ${}_s\mathcal{P}lot$ distributions of $\cos\theta_\gamma$ are in a very good agreement.

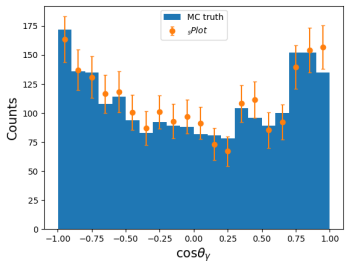


Figure 4.98. $\cos\theta_\gamma$ distribution for $M_{K\pi\pi} \in [1.1; 1.3]$ GeV/c^2

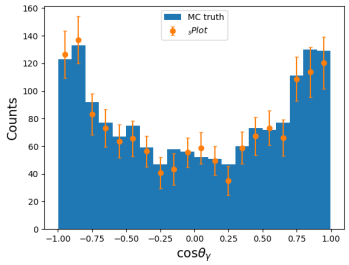


Figure 4.99. $\cos\theta_\gamma$ distribution for $M_{K\pi\pi} \in [1.3; 1.4]$ GeV/c^2

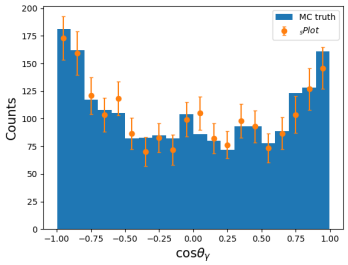


Figure 4.100. $\cos\theta_\gamma$ distribution for $M_{K\pi\pi} \in [1.4; 1.6]$ GeV/c^2

LHCb [45] has been measured the up-down assymetry through curve fitting using Legendre's polynomial. Obtained results along with accompanied uncertainties (statistics and systematics) are given in the Fig. 4.101. Rescaling down $\sigma[A_{ud}^{s\mathcal{P}}]$ by a factor of $\sqrt{\frac{13876}{7386}} = 1.37$ gives a typical value of $\sigma[A_{ud}^{s\mathcal{P}}] = 0.025$ which is still higher than obtained in the LHCb analysis.

	$[1.1, 1.3]$	$[1.3, 1.4]$	$[1.4, 1.6]$	$[1.6, 1.9]$
c_1	6.3 ± 1.7	5.4 ± 2.0	4.3 ± 1.9	-4.6 ± 1.8
c_2	31.6 ± 2.2	27.0 ± 2.6	43.1 ± 2.3	28.0 ± 2.3
c_3	-2.1 ± 2.6	2.0 ± 3.1	-5.2 ± 2.8	-0.6 ± 2.7
c_4	3.0 ± 3.0	6.8 ± 3.6	8.1 ± 3.1	-6.2 ± 3.2
$\mathcal{A}_{\rm ud}$	6.9 ± 1.7	4.9 ± 2.0	5.6 ± 1.8	-4.5 ± 1.9

Figure 4.101. Up-down asymmetry measurements using LHCb data with 13876 ± 153 signal events.

Comparing A_{ud} (Table 4.18) and $\mathcal{A}_{ud}$ (Table 4.101) one observes large divergence for each bin of $M_{K\pi\pi}$. Despite satisfactory description of LHCb histograms by "GamPola" fit while defining the Baseline Model in the Section 2.4.1, it turns out A_{ud} was not fitted properly.

4.8 . Control samples unblinding

In this section, the real data used for unblinding is discussed. Several datasets incorporating both peaking and non-peaking backgrounds are built. Tables and plots for MC/data comparison are presented.

4.8.1 . Dataset

The dataset used for the unblinding contains $62.7\ fb^{-1}$ of $\Upsilon(4S)$ data and $9.2\ fb^{-1}$ of off-resonance data. On-resonance data is used to build $B^+ \to K^+\pi^-\pi^-\gamma$ (CS1) and $B^+ \to K^+\pi^+\pi^+\gamma$ (CS2) datasets in the full region of $M_{bc} \in [5.2; 5.29]$ GeV and also for reconstruction $B^+ \to K^+\pi^+\pi^-\gamma$ in the sideband region ($M_{bc} \in [5.2; 5.27]$ GeV). In the off-resonance case Monte-Carlo contains only $q\bar{q}$ events and for all other cases complete generic sample of events.

4.8.2 . Off-resonance data

Since energy of the collisions is smaller on 60 MeV in the case of off-resonace data taking, it leads to the effect where M_{bc}-distribution does not extend up to 5.29 GeV, but stops nearly 5.26 GeV as can be seen from Fig. 4.102. In this case M_{bc} and others energy dependent variables are multiplied by a factor of $\frac{M_{\Upsilon(4S)}}{M_{\Upsilon(4S)}-60\ \rm MeV} = 1.0057$ in order to account for boost changing (see Fig. 4.103) and properly compare with continuum background of Monte-Carlo.

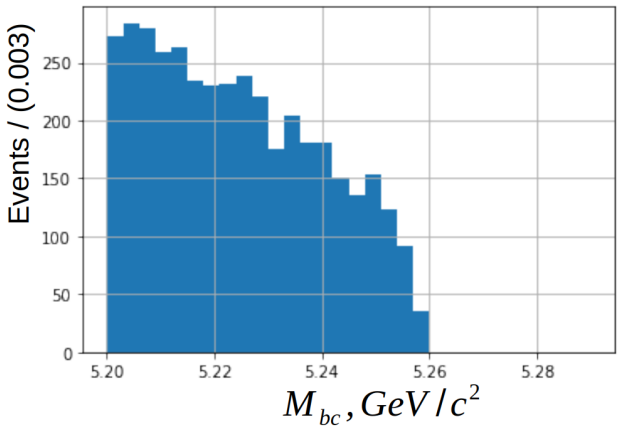


Figure 4.102. M_{bc} distribution without rescaling factor

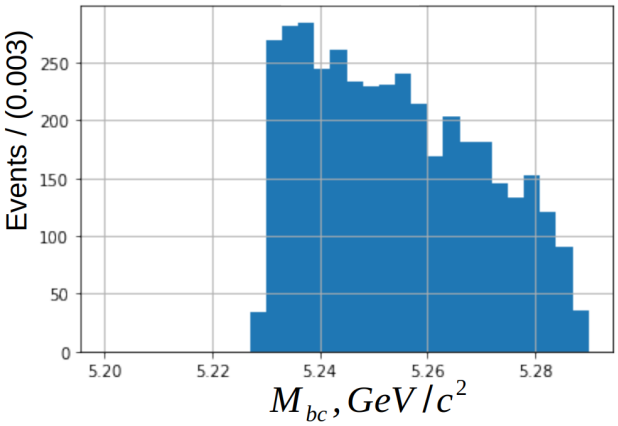


Figure 4.103. M_{bc} distribution with scaling factor.

Having applied the scale factor mentioned above, on the Figs. 4.104-4.109 the comparison of data and Monte-Carlo is given in two ways: shapes comparison — $\frac{1}{N}\frac{dN}{dx}$ (P.D.Fs) are compared, statistics comparison — $\frac{dN}{dx}$ (Counts) of two histograms are compared.

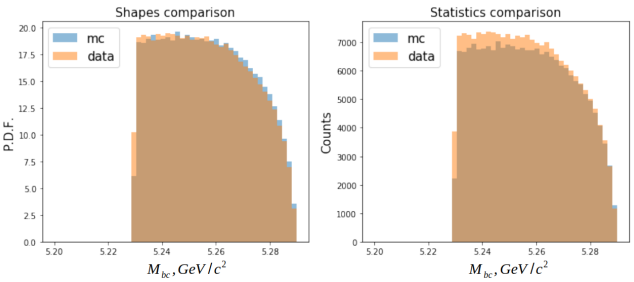


Figure 4.104. M_{bc} distributions after preselection cuts

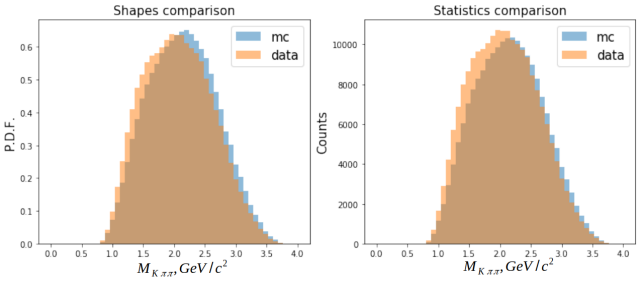


Figure 4.105. $M_{K\pi\pi}$ distributions after preselection cuts

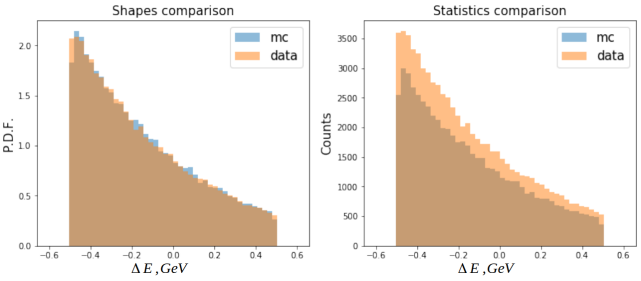


Figure 4.106. ΔE distributions after $M_{K\pi\pi} < 1.8$ GeV

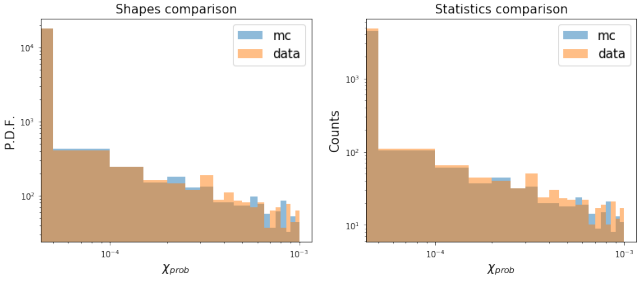


Figure 4.107. Part of χ_{prob} distributions after $-0.2 < \Delta E < 0.1$ GeV.

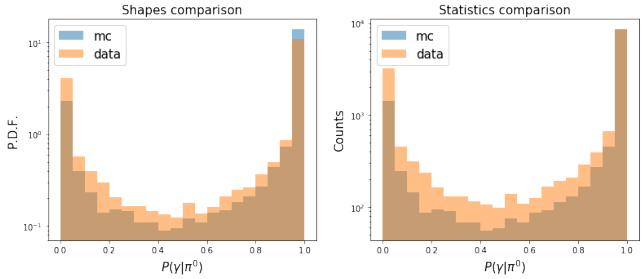


Figure 4.108. $P(\gamma|\pi^0)$ distributions after $\chi_{prob} > 0.0001$

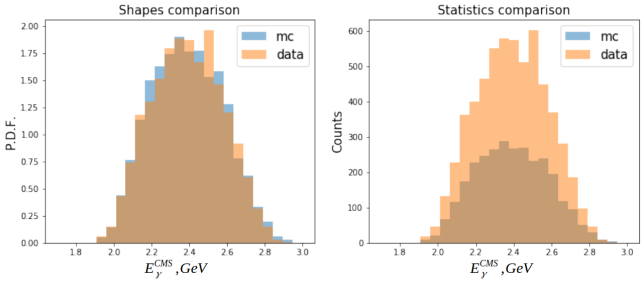


Figure 4.109. E_γ^{CMS} distributions after $P(\gamma|\pi^0) < 0.8$

Cut/Events kept %	Monte-Carlo	Off-resonance data
$0 < m_{K\pi\pi} < 1.8$ GeV/c^2	34.3%	39.8%
$-0.2 < \Delta E < 0.1$ GeV	34.4%	34.5%
$0.0001 < \chi_{\text{prob}}$	75.6%	78%
$0 < P(\gamma\|\pi^0) < 0.8$	23.2%	37.6%
$2.1 < E_\gamma^{CMS} < 2.75$ GeV	90.5%	90.8%
Summary	1.75%	3.4%

Table 4.19. Selection cuts efficiencies for off-resonance data

From Table 4.19 one can see that cut on $P(\gamma|\pi^0)$ is a major source of divergence between data and Monte-Carlo. Fig. 4.110 illustrates the efficiency of the cut as a function of E_γ^{CMS}: although the efficiencies are correlated and have the same trend, the bias between them is significant.

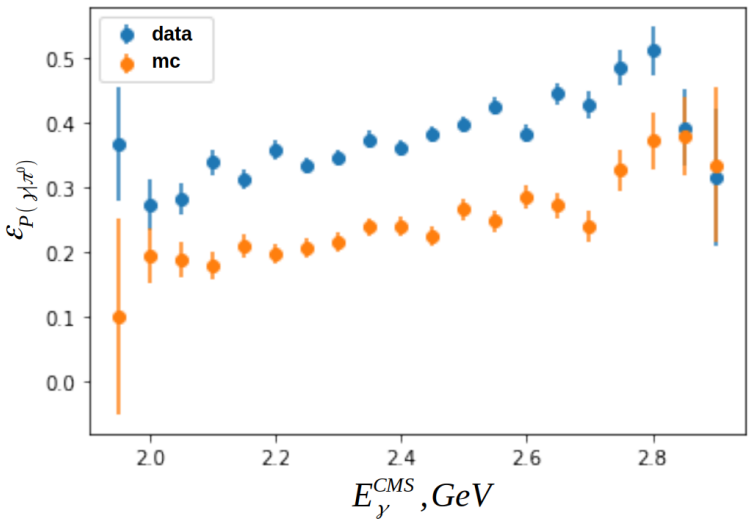


Figure 4.110. Eficiency of $P(\gamma|\pi^0) < 0.8$ cut as a function of E_{γ}^{CMS}

Post selection results

As a further stage of comparison of real data and Monte-Carlo neural net μ_{NN} distributions are depicted in two representations:

- Luminosity re-scaling — distribution of random variable (corresponding to huge MC statistics) is rescaled by luminosity factor when plotting. This type of rescaling was applied for comparing the statistics of the distributions;
- Statistics re-scaling — the same procedure as in luminosity re-scaling except that once scaling factor is applied both histograms have the same amount of entries. This type of rescaling was applied for comparing shapes of the distributions;

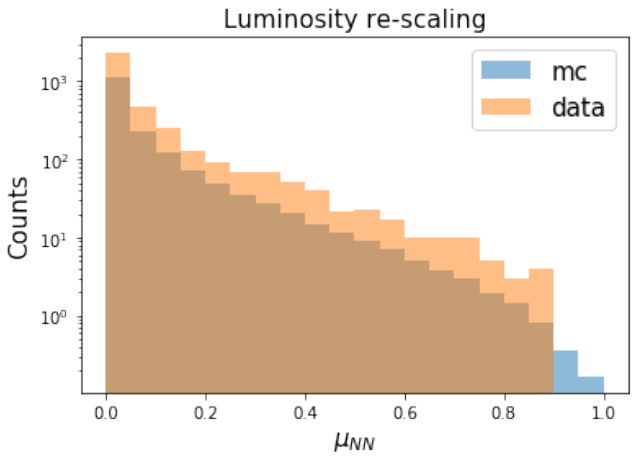


Figure 4.111. μ_{NN} distribution after BCS with luminosity re-scaling

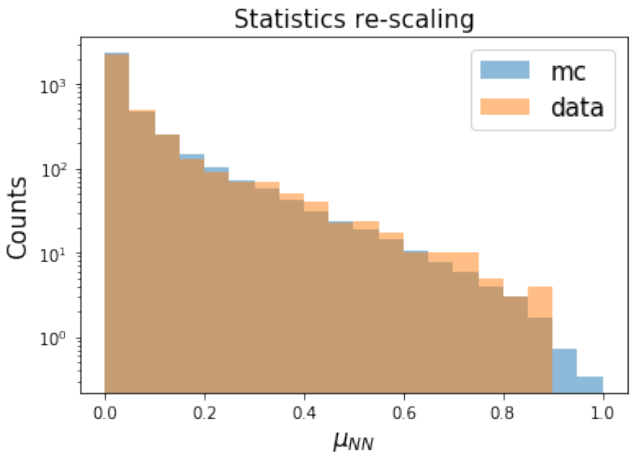


Figure 4.112. μ_{NN} distribution after BCS, scaling = 2.04

The distributions of M_{bc} after $\mu_{NN} > 0.28$ cut is plotted on the Figs. 4.113-4.114).

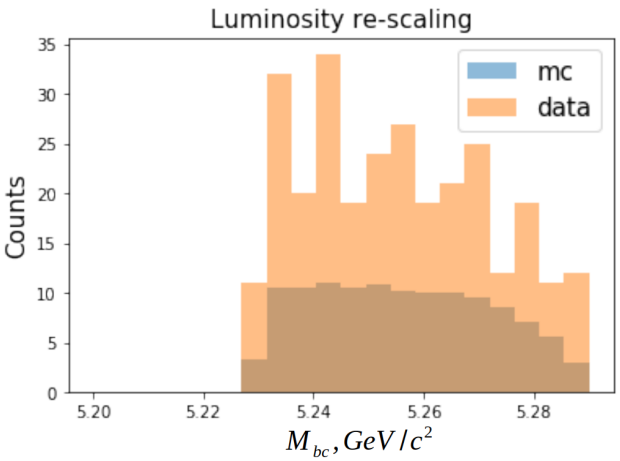


Figure 4.113. M_{bc} distribution after μ_{NN} cut with luminosity re-scaling

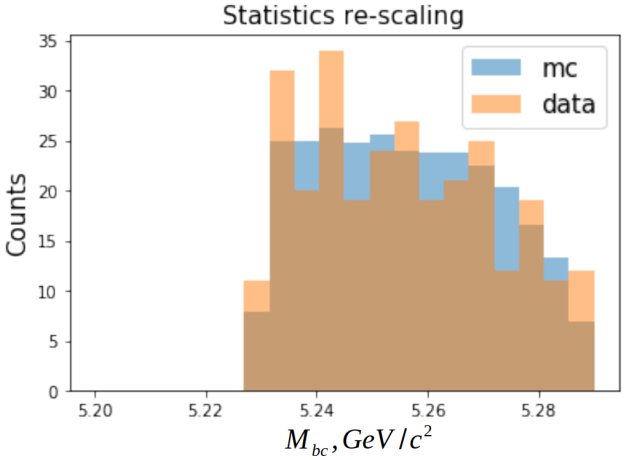


Figure 4.114. M_{bc} distribution after μ_{NN} cut with scaling = 2.38

Although the shapes of distributions are in good agreement, statistics of the Monte-Carlo data sample is a factor of two less, than of the off-resonance data. The same is concluded from the selection efficiencies in Table 4.19.

4.8.3 . CS1

A control sample of type I stands for $K^+\pi^-\pi^-\gamma$ final state and is discussed in detail in Section. 4.4. In this part, the comparison between data and MC is discussed.

As in the previous section, starting from selection cuts intermediate comparisons (see Figs. 4.126-4.120) are given.

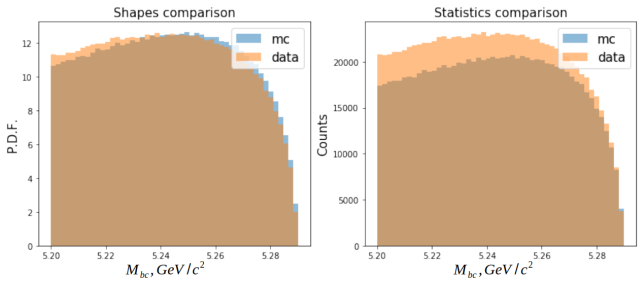


Figure 4.115. M_{bc} distributions after preselection cuts

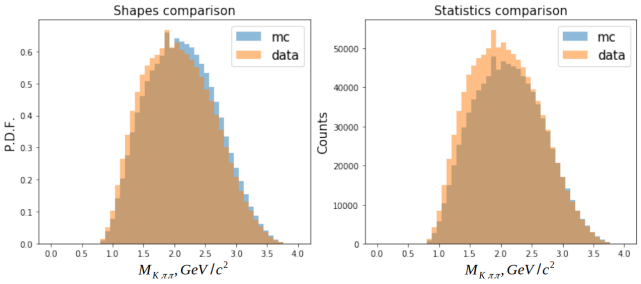


Figure 4.116. $M_{K\pi\pi}$ distributions after preselection cuts

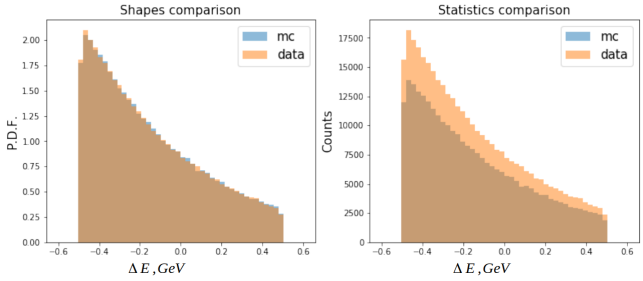


Figure 4.117. ΔE distributions after $M_{K\pi\pi} < 1.8$ GeV

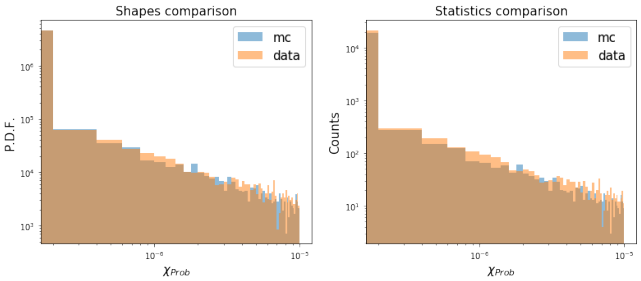


Figure 4.118. Part of χ_{prob} distributions after $-0.2 < \Delta E < 0.1$ GeV.

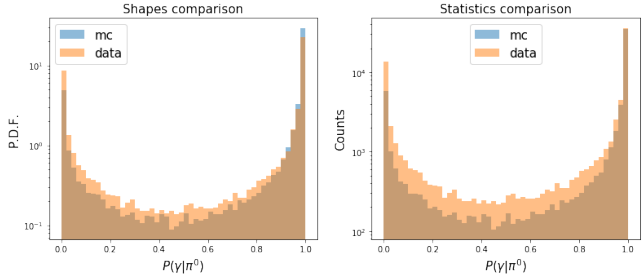


Figure 4.119. $P(\gamma|\pi^0)$ distributions after $\chi_{prob} > 0.0001$

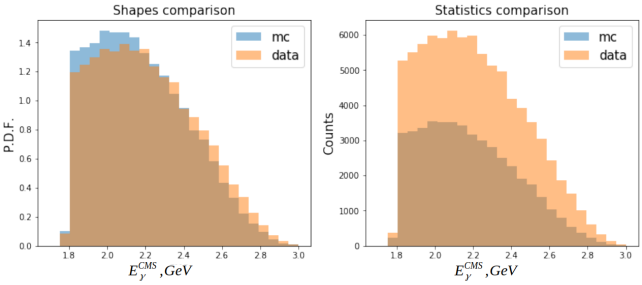


Figure 4.120. E_γ^{CMS} distributions after $P(\gamma|\pi^0) < 0.8$

Cut/Events kept %	CS1 Monte-Carlo	CS1 data
$0 < m_{K\pi\pi} < 1.8$ GeV/c^2	35.8%	40.3%
$-0.2 < \Delta E < 0.1$ GeV	32.9%	33.4%
$0.0001 < \chi_{\text{prob}}$	74.5%	77.7%
$0 < P(\gamma\|\pi^0) < 0.8$	24.6%	38.3%
$2.1 < E_\gamma^{CMS} < 2.75$ GeV	55%	58%
Summary	1.2%	2.3%

Table 4.20. Selection cuts efficiencies for CS1

At the preselection level $M_{K\pi\pi}$ distribution contains the peaks (see Fig. 4.127) nearly 1.87 GeV and 2 GeV corresponding to the decay channels of D^- and D^{*-} mesons. The selection efficiencies for the $2.1 < E_\gamma^{CMS} < 2.75$ GeV cut are in very good areement for data and MC, however significantly different from the ones listed in the Table. 4.19. This can be also seen when comparing the Figs. 4.109, 4.120.

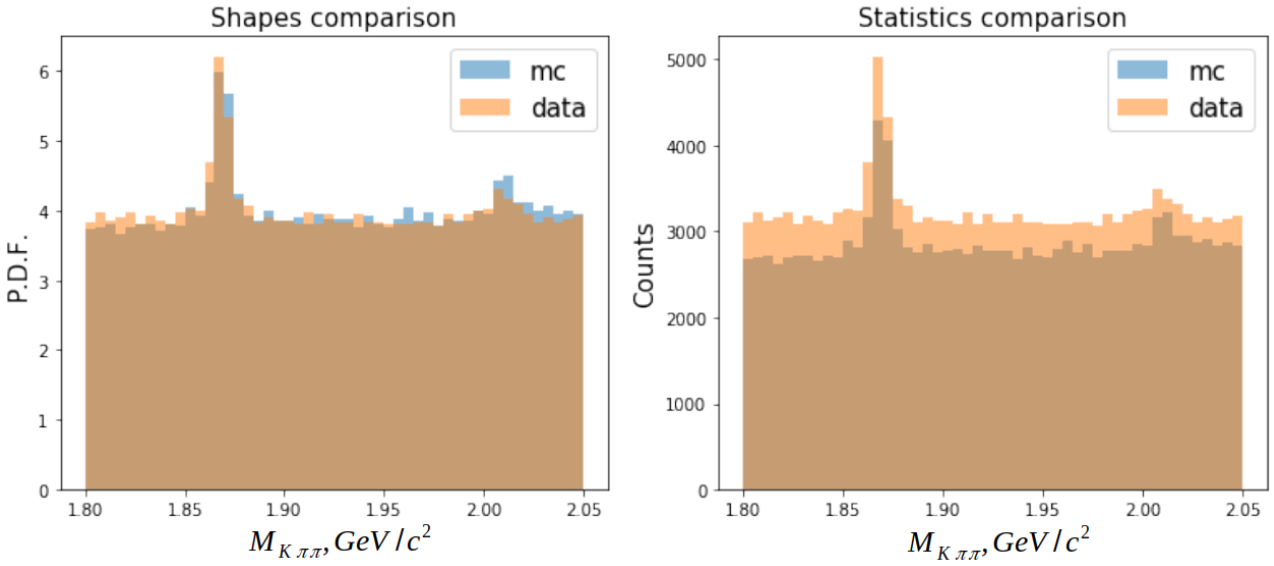


Figure 4.121. $M_{K\pi\pi}$ distribution after preselection (zoom nearly peaks)

In order to understand whether cut on $P(\gamma|\pi^0)$ has a global scaling effect (different types of events' species are affected in the same way), extended likelihood fit of the data and Monte-Carlo in the range $M_{K\pi\pi} \in [1.85; 1.9]$ GeV has been performed at the preselection level. The fit is performed using gaussian P.D.F. for signal and first-order polynomial for background (see Figs. 4.122-4.123). The shape parameters of these P.D.Fs have been fixed. After applying selection cuts up to cut on $0 < P(\gamma|\pi^0) < 0.8$ extended likelihood fit was repeated with fixed P.D.Fs parameters and corresponding yields were measured (see Figs. 4.124-4.125). Table 4.21 summarizes these results. Each number was obtained by dividing the

yields of events' species before specified selection and after. Obtained results demonstrate that within a sample (MC or data) the efficiencies are close for peaking and flat parts of $M_{K\pi\pi}$, which signifies that the divergence between data and MC has a global effect.

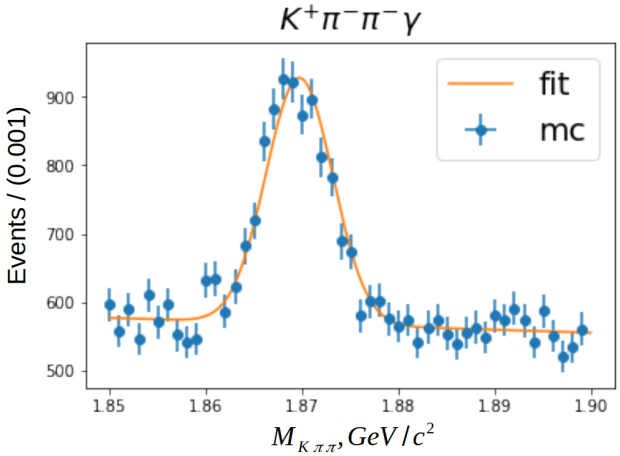


Figure 4.122. Fit of $M_{K\pi\pi}$ distribution at the preselection level for MC.

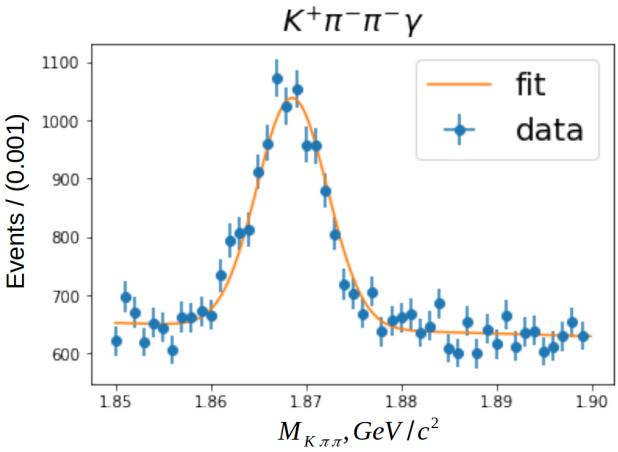


Figure 4.123. Fit of $M_{K\pi\pi}$ distribution at the preselection level for data.

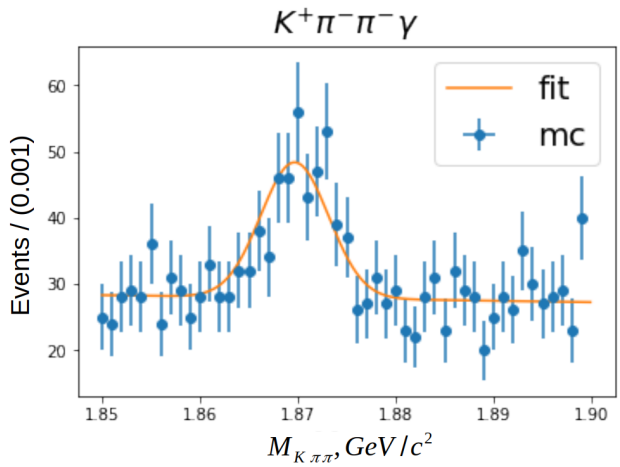


Figure 4.124. Fit of $M_{K\pi\pi}$ distribution after π^0-veto cut for MC

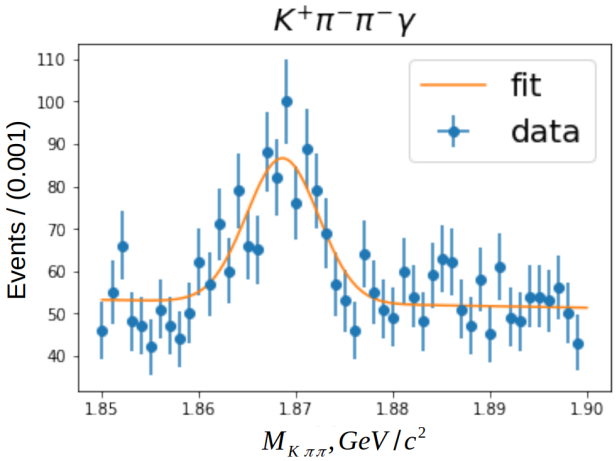


Figure 4.125. Fit of $M_{K\pi\pi}$ distribution after π^0-veto cut for data

Efficiencies %	Peak	Flat
Monte-Carlo	5.7 (0.4)%	4.9 (0.1)%
Data	8.7 (0.5)%	8.1 (0.2)%

Table 4.21. Efficiencies of peaking and flat components after $P(\gamma|\pi^0)$ cut

Post selection plots

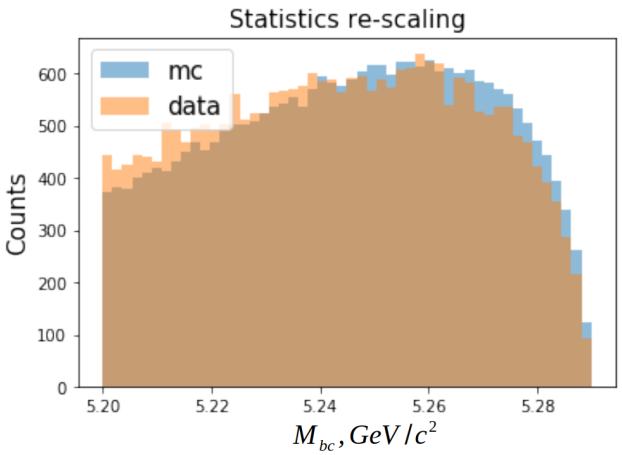


Figure 4.126. M_{bc} distributions after selection cuts, scaling = 2.05

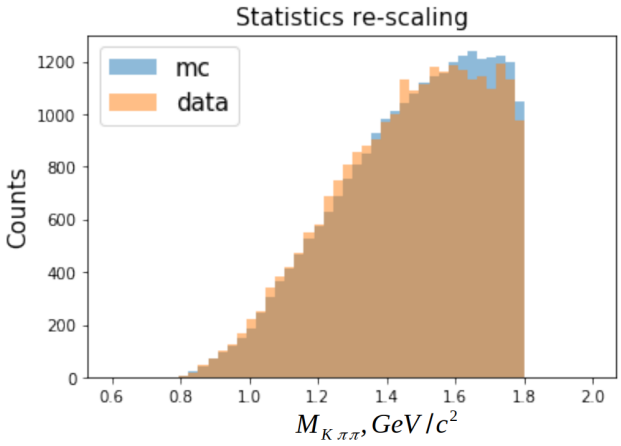


Figure 4.127. $M_{K\pi\pi}$ distributions after selection cuts, scaling = 2.05

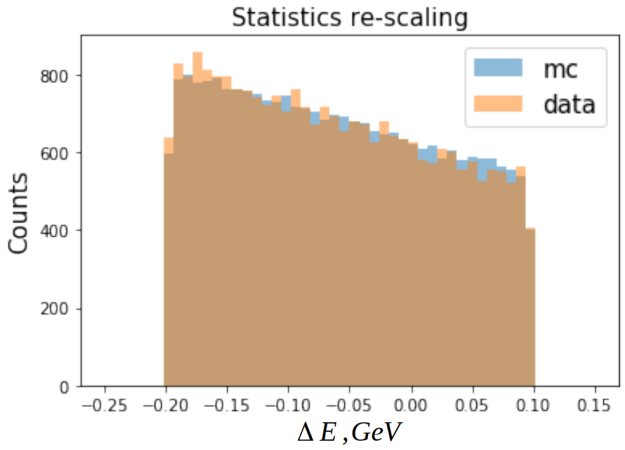


Figure 4.128. ΔE distributions after selection cuts, scaling = 2.05

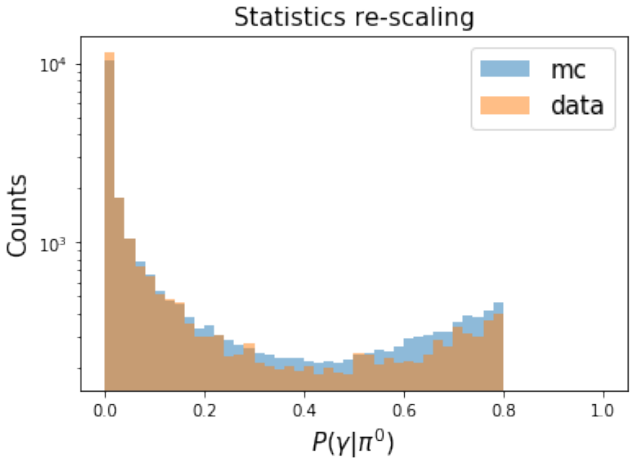


Figure 4.129. Part of χ_{prob} distributions after selection cuts, scaling = 2.05

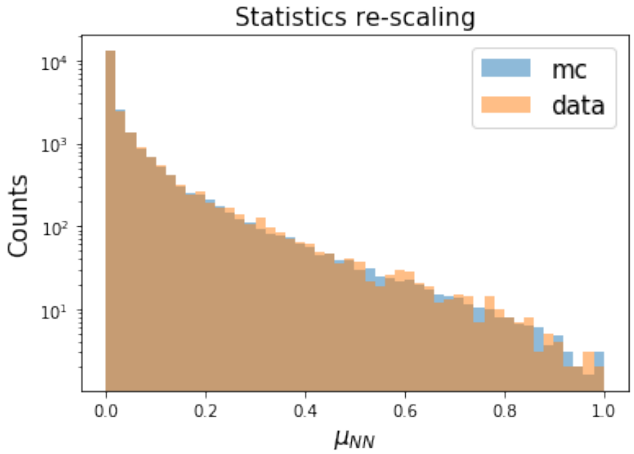


Figure 4.130. μ_{NN} distribution after BCS, scaling = 1.97

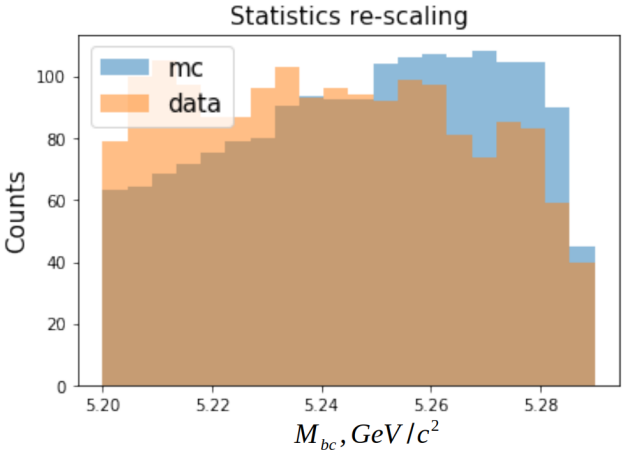


Figure 4.131. M_{bc} distributions after $\mu_{NN} > 0.28$, scaling = 1.95

4.8.4 . CS2

Control sample of type II stands for $K^+\pi^+\pi^+\gamma$ final state and discussed in details in the Section. 4.4. In this section the comparison between data and MC is given: Table. 4.22 summarizes selection cuts efficiencies, Figs. 4.132-4.135 illustrate the comparison after selection cuts, Fig. 4.136 — μ_{NN} distribution after BCS and Fig. 4.137 — M_{bc} disribution after $\mu_{NN} > 0.28$ cut.

Cut/Events kept %	Monte-Carlo	CS2 data
$0 < m_{K\pi\pi} < 1.8$ GeV/c^2	33.8%	36.7%
$-0.2 < \Delta E < 0.1$ GeV	30.9%	31.5%
$0.0001 < \chi_{\text{prob}}$	69.3%	73.1%
$0 < P(\gamma\|\pi^0) < 0.8$	24.5%	38.7%
$2.1 < E_\gamma^{CMS} < 2.75$ GeV	84.45%	84.3%
Summary	1.5%	2.8%

Table 4.22. Cut efficiencies for CS2

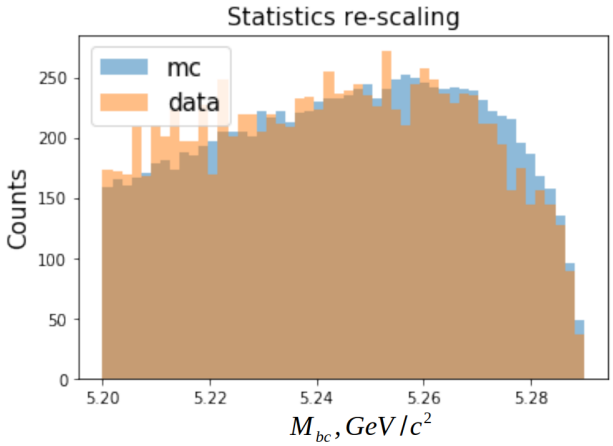


Figure 4.132. M_{bc} distributions after selection cuts, scaling factor = 2.21

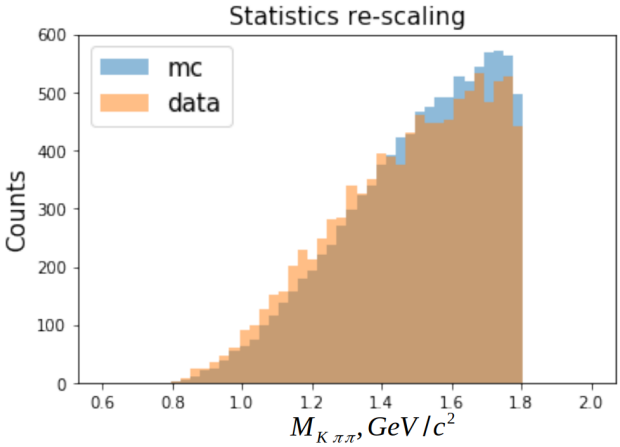


Figure 4.133. $M_{K\pi\pi}$ distributions after selection cuts, scaling factor = 2.21

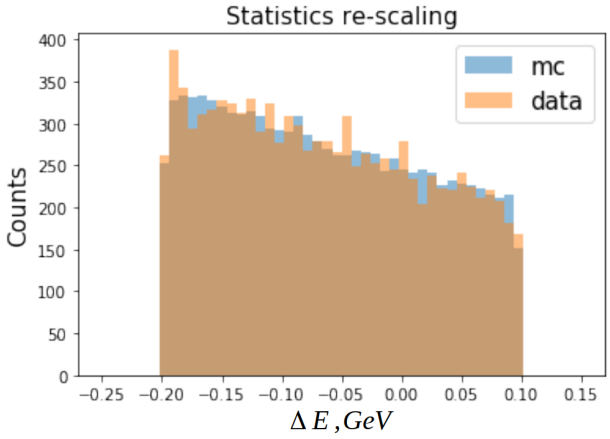


Figure 4.134. ΔE distributions after selection cuts, scaling factor = 2.21

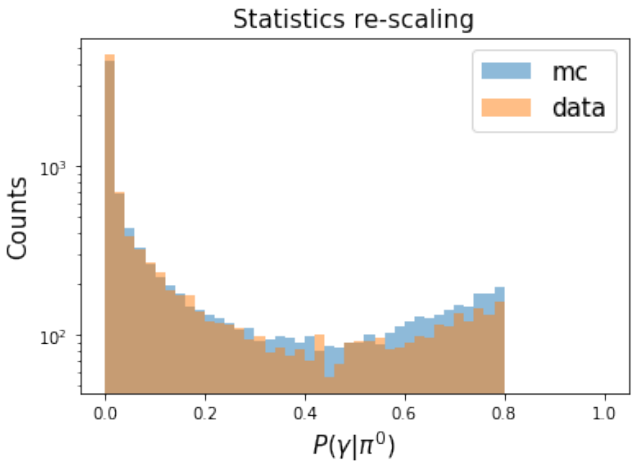


Figure 4.135. Part of χ_{prob} distributions after selection cuts, scaling factor = 2.21

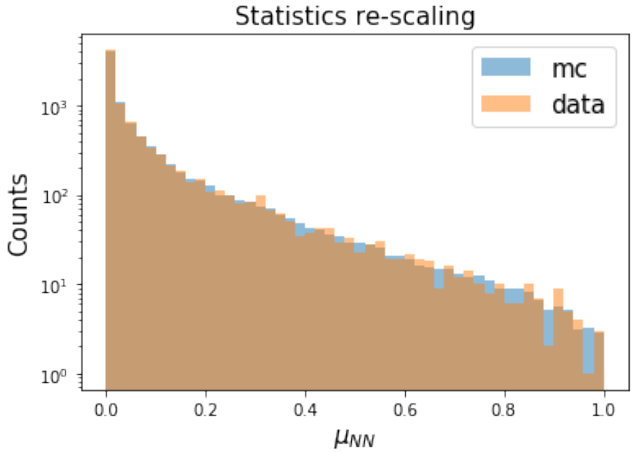


Figure 4.136. μ_{NN} distribution after BCS, scaling factor = 2.11

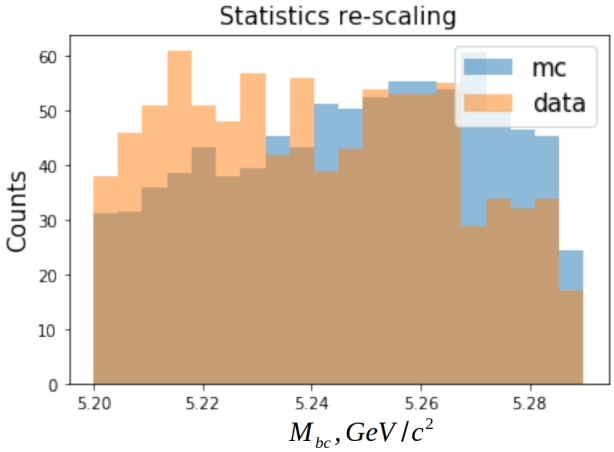


Figure 4.137. M_{bc} distributions after $\mu_{NN} > 0.28$, scaling factor = 2.13

Scaling factor $\simeq 2$ is applied for the statistics rescaling and is accumulated during the whole selection process, but the main source of inconsistency remains $P(\gamma|\pi^0) < 0.8$ cut.

4.8.5 . M_{bc}-sideband

In this part of the work, the comparison between data and MC is done for sideband data.

On the Table. 4.23 selection efficiencies comparison is given, Figs. 4.138-4.141 illustrate the comparison after selection cuts, Fig. 4.142 — after BCS and Fig. 4.143 shows M_{bc} distribution after μ_{NN} cut.

Cut/Events kept %	Monte-Carlo	Sideband data
$0 < m_{K\pi\pi} < 1.8$ GeV/c^2	36.5%	40.4%
$-0.2 < \Delta E < 0.1$ GeV	34.4%	34.9%
$0.0001 < \chi_{\text{prob}}$	74.8%	77.6%
$0 < P(\gamma\|\pi^0) < 0.8$	23.8%	38.7%
$2.1 < E_\gamma^{CMS} < 2.75$ GeV	80.1%	80.2%
Summary	1.8%	3.5%

Table 4.23. Cut efficiencies for sideband before re-weightening

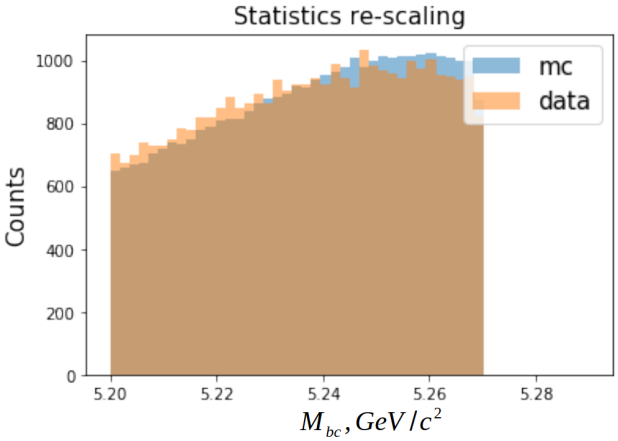


Figure 4.138. M_{bc} distributions after selection cuts, scaling = 1.91

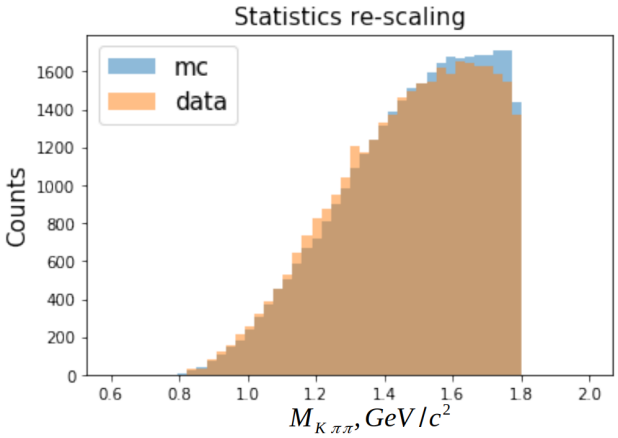


Figure 4.139. $M_{K\pi\pi}$ distributions after selection cuts, scaling = 1.91

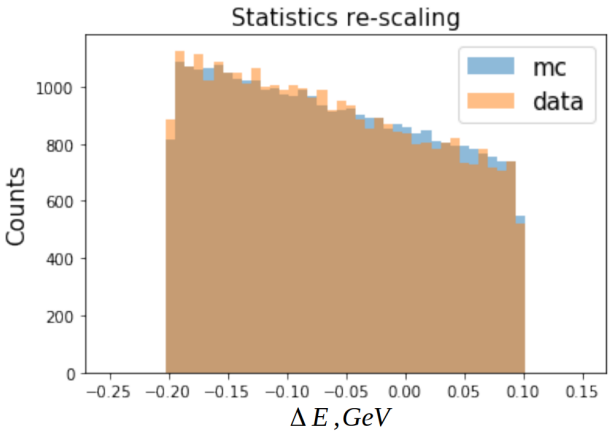


Figure 4.140. ΔE distributions after selection cuts, scaling = 1.91

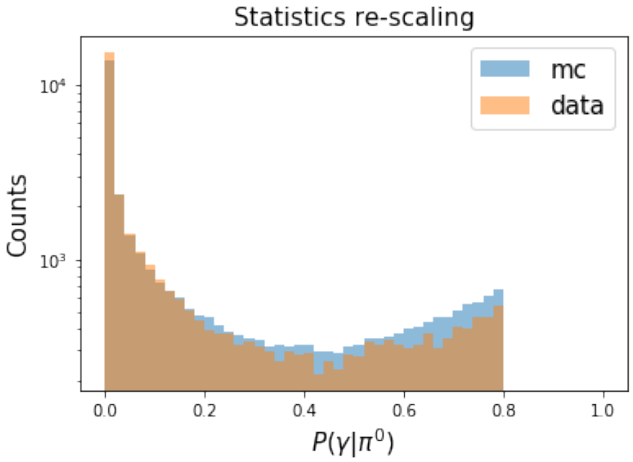


Figure 4.141. Part of χ_{prob} distributions after selection cuts, scaling = 1.91

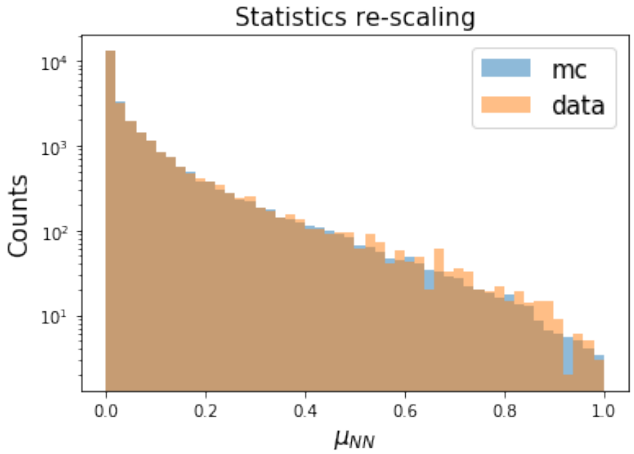


Figure 4.142. μ_{NN} distribution after BCS, scaling = 1.81

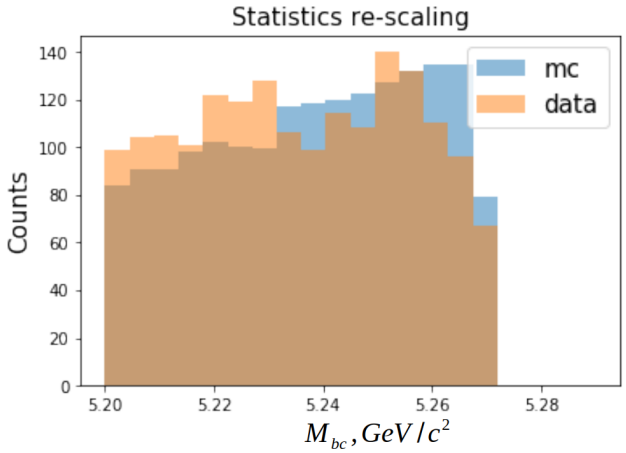


Figure 4.143. M_{bc} distribution after cut on $\mu_{NN} > 0.28$, scaling = 1.95

Unlike off-resonance data, representing only $q\bar{q}$ events, sideband data contains also $B\bar{B}$ events in the $M_{bc} \in [5.2; 5.27]$ GeV region, the later holds for CS1 and CS2 as well with $M_{bc} \in [5.2; 5.29]$ GeV/c^2. Apart from persisting π^0-veto cut's efficiencies disagreement (see Tables. 4.19-4.23) on the Figs. 4.131,4.137,4.143 one can observe excess of data events for lower values of M_{bc}.

By splitting MC sample on two parts: $B\bar{B}$ only and $q\bar{q}$ only the M_{bc} distributions' comparisons are made (see Figs. 4.144-4.145). From these figures one can conclude that within MC $B\bar{B}$ events are the source of divergence of the shapes.

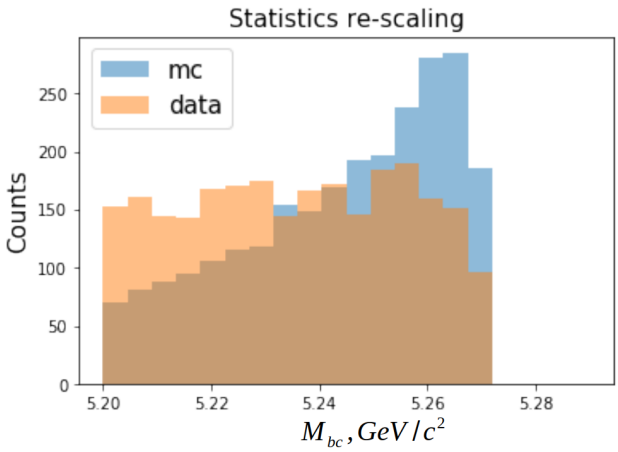


Figure 4.144. M_{bc} distribution after cut on $\mu_{NN} > 0.28$ (only $B\bar{B}$ MC events), scaling = 9.23

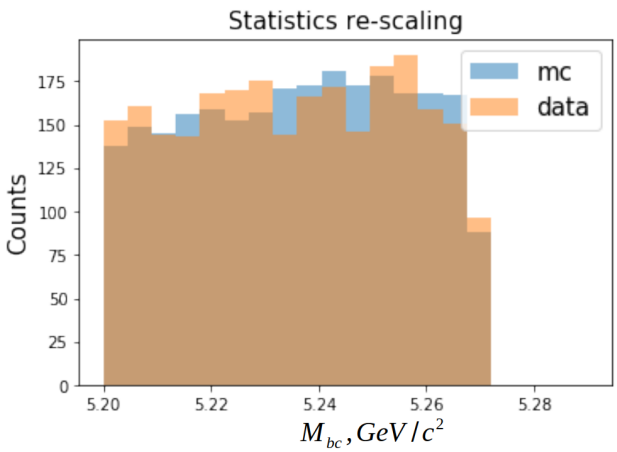


Figure 4.145. M_{bc} distribution after cut on $\mu_{NN} > 0.28$ (only $q\bar{q}$ MC events), scaling = 2.47

5 - Conclusions

In the frame of the present work $B^+ \to K_{res}\gamma \to K^+\pi^+\pi^-\gamma$ decay study was addressed. The theoretical model has been implemented within the developed software package "GamPola". The model describes both neutral and charged $B \to K\pi\pi\gamma$ modes through decays of several kaonic resonances K_{res}: $K_1^{1270}, K_1^{1400}, K_1^{*1410}, K_1^{*1680}, K_2^{1430}$ populating spectrum of invariant mass of $K\pi\pi$. The need for such a model is motivated by the fact that within "EvtGen" interference of kaonic resonances is not taken into account.

The kinematics of decay in the $K\pi\pi$ rest frame is described in terms of 5 independent variables: $(s_{K\pi\pi}, s_{K\pi}, s_{\pi\pi}, \theta, \phi)$. The expressions for conversion between B and $K\pi\pi$ rest frames are provided: while transforming 5D kinematics in the $K\pi\pi$ rest frame to a set of 4-momenta of $K\pi\pi\gamma$ in the B rest frame, 3 additional degrees of freedom are introduced, representing random Euler's angles of the absolute direction of the system.

The amplitude of B-meson decay is factorized as a sum of products of the amplitudes of $B \to K_{res}\gamma$ and $K_{res} \to K\pi\pi$ described separately. The decay of K_{res} proceeds either through subsequent $K_{res} \to K^{*892}\pi$ or $K_{res} \to \rho^{770}K$ interfering decays. In case of 1^+ resonance, K_{res} is represented by K_1^{1270} and K_1^{1400} mixing using angle θ_{K_1}. Form factors of $K_1 \to K^{*892}\pi$ and $K_1 \to \rho^{770}K$ are parametrized with hadronic phases $\varphi_S^{K^*}, \varphi_S^{\rho}, \varphi_D^{K^*}$. Complex coupling is introduced to the model, standing for the fraction of K_1^{1400} in the total amplitude. In the case of $K_1^{*1410/1430}, K_2^{1430}$ the parametrization is done by complex model parameters representing by the couplings of $K_{res} \to VP$ sub-decays. The photon polarization parameter λ_γ is introduced to the model on the stage of decay rate definition in the $\frac{1\pm\lambda_\gamma}{2}$ weights of left and right decay rates.

The "GamPola" software relies on the defined model and implements generator and fitter of $B \to K\pi\pi\gamma$ events. The generator produces events in two stages: firstly is produces a kinematically allowed event that would occur in the absence of any physics interaction, then it uses the event's kinematics for generating the final event with physics included. Illustration of the generator output is done, and distributions of 5D kinematics are compared for charged and neutral decay modes. The "GamPola"-fitter is based on maximum likelihood fit, which fully relies on the ROOT framework. The normalization part of the decay rate is calculated based on the factorization of decay rate in two parts: functions of model parameters and functions of kinematical variables. The factorization is done using the "GiNAC" library for symbolic computations and allows to greatly reduce the fitting time. Generator and fit level distributions of 5D kinematics are in good agreement as well as model parameter values.

Having implemented generator and fitter, the Baseline model of $B^+ \to K^+\pi^+\pi^-\gamma$ decay is defined using LHCb histograms of $M_{K^+\pi^+\pi^-}, M_{K^+\pi^-}, M_{\pi^+\pi^-}$ and $\cos\theta_\gamma$. The latter are the histograms in three bins of $M_{K^+\pi^+\pi^-}$: $1.1 < M_{K^+\pi^+\pi^-} < 1.3$ GeV/c^2, $1.3 < M_{K^+\pi^+\pi^-} < 1.4$ GeV/c^2 and $1.4 < M_{K^+\pi^+\pi^-} < 1.6$ GeV/c^2. The last bin $1.6 < M_{K^+\pi^+\pi^-} < 1.9$ GeV/c^2 of $\cos\theta_\gamma$ is not included since "GamPola" unable to fit the histograms in this region. The Baseline model is further used for the sensitivity study. In particular it is shown that cut on $M_{K^+\pi^+\pi^-} < 1.6$ GeV almost doesn't affect the photon polarization sensitivity and can be safely applied on the stage of the photon polarization measurement. In the frame of another test, the fit using the Baseline model of four samples is performed: for both neutral

and charged decay modes two models with complex and real parameters are used. Obtained results show mild dependence of the photon polarization sensitivity on the model choice and decay mode with typical value of $\sigma_{\lambda_\gamma}(B^+ \to K^+\pi^+\pi^-\gamma) \approx 0.04$ for 5000 events. The fit stability is probed by repeatedly fitting the same sample of events with a random seed. Three types of solutions are identified: truth — corresponding to the generator level parameter values, fake — corresponding to the high values of $\Delta\chi^2 > 16$ and mirror — corresponding to low $\Delta\chi^2 \leq 16$. The fake solutions are discarded by applying tight cut on the fit outcome $\Delta\chi^2 \leq 16$ and produced in 50 % of fit trials. Mirror solutions are not dangerous and do not affect the photon polarization parameter, which is the main concern of $B \to K\pi\pi\gamma$ decay modeling. Analyzing $C_7 - C_7'$ plane, it is concluded that measurement of λ_γ is represented on this plane by a double-circle with inner and outer radii. Required integrated luminosity for the competitive photon polarization measurements is estimated and equals nearly $3\ \mathrm{ab}^{-1}$. In such a case, the area of the LHCb disk is expected to be less or equal to the area of the double circle of Belle II. However, calculation of systematics is needed and will affect such an estimate.

In the present work the analysis of $B^+ \to K^+\pi^+\pi^-\gamma$ decay with a $K^+\pi^+\pi^-$ mass in the $[1, 1.8]\ \mathrm{GeV/c^2}$ range is done for the Monte-Carlo dataset of size $1\ \mathrm{ab}^{-1}$ and real-world dataset consisting of $9.2\ \mathrm{fb}^{-1}$ of off-resonance and $62.7\ \mathrm{fb}^{-1}$ of $\Upsilon(4S)$ data collected by Belle II at the end of summer 2020.

For the large sample ($1\ \mathrm{ab}^{-1}$) of generic MC-data and input PDG value of branching ratio $2.58 \pm 0.15 \cdot 10^{-5}$, $B^+ \to K^+\pi^+\pi^-\gamma$ events generated by "EvtGen" are replaced by "GamPola" model incorporated in the "EvtGen". The replacement procedure is done after preselection cuts are applied. An obtained new sample of generic MC data is further used within the analysis. It is concluded that for the final goal of this analysis program the main concern for the MC-study should be truth-matched (TM) events.

Several groups of cuts are considered, including skimming, preselection, and selection giving in total 42 % of selection efficiency of $B^+ \to K^+\pi^+\pi^-\gamma$ events. Applying the cuts already allows observing a peak hinting for signal. The possible source of correlation between M_{bc} and ΔE is identified representing photon energy, which is reconstructed with % level difference with respect to its truth value. Changing the photon energy from truth to reconstructed value explains the difference of both M_{bc} and ΔE distributions between truth and reconstructed ones.

Continuum background is suppressed using artificial Neural Net taking as input mostly candidate level variables. The model is trained using an equal amount of signal TM candidates and $q\bar{q}$ events. Best candidate selection is done using NN output which allows keeping 93 % of TM events. Original Neural Net output is further transformed using μ-transfrom, such that μ_{NN}^{TM} distribution is flat. In this new representation applying optimal cut $\mu_{NN} > 0.28$ allows to keep 7386 TM events corresponding to $1\ \mathrm{ab}^{-1}$ of MC-data. The main sources of peaking background in the signal region of M_{bc} are identified. The highest contribution is coming from $B^0\bar{B}^0$ events. While $B^+\bar{B}^-$ and SCF are peaking as well, their contributions are much smaller.

Several control samples are selected for analysis validation. Among standard ones, such as M_{bc} side-band ($M_{bc} \in [5.2; 5.27]\ \mathrm{GeV/c^2}$) and off-resonance data, there are two specific for $B^+ \to K^+\pi^+\pi^-\gamma$ analysis: with $K^+\pi^+\pi^+\gamma$ and $K^+\pi^-\pi^-\gamma$ in the final state. The latter one exhibits additional property: it is used in the joint extended likelihood fit along with the Physics Sample.

Extended likelihood fit is performed using Physics Sample and three components of the total p.d.f. Introducing separate p.d.f for TM events allows to directly measure the expected yield of these events and more importantly recover TM distributions of various observables using ${}_s\mathcal{P}lot$. The fitted yield is

7400 ± 203 events which is in very good agreement with the MC-truth value, but the correlation between this quantity and yield of peaking background is very high. Although tFit+CS1 ($K^+\pi^-\pi^-\gamma$ in the final state) reduces statistical uncertainty of TM yield on $25\ \%$ compared to the tFit, the latter is considered as the baseline fit on this stage of the Belle II. Toy Monte-Carlo study is done for $63 fb^{-1}$ demonstrating that the likelihood estimator is not significantly biased with $\langle z \rangle = +0.01 \pm 0.03$ and $\sigma[z] = 0.975 \pm 0.02$.

For the sample of $63\ \mathrm{fb}^{-1}$, branching ratio equals $\mathrm{Br}_{K^+\pi^+\pi^-\gamma} = 2.65 \pm 0.27 \cdot 10^{-5}$ and is calculated without taking into account possible dependence on kinematical phase space. Such dependence is taken into account by splitting sample of $1\ \mathrm{ab}^{-1}$ into 125 bins in Dalitz variables and calculating yield and efficiency for each bin. The obtained value of branching ratio within the second approach is $\mathrm{Br}_{K^+\pi^+\pi^-\gamma} = 2.52 \pm 0.07 \cdot 10^{-5}$ and the uncertainties of branching ratios obtained by the two approaches are in good agreement.

As intermediate step, up-down asymmetry is calculated in the three bins of $M_{K\pi\pi}$ using ${}_s\mathcal{P}lot$ technique and MC-truth counting. Obtained values of A_{ud} are equal within statistical uncertainties. However, comparison with LHCb results gives large inconsistency. It can be explained by the fact, that the Baseline Model obtained from the fit of LHCb histograms reproduces these histograms only visually.

Control samples unblinding is performed using off-resonance, sideband, CS1, and CS2. The unblinding shows good agreement between the shapes of the distributions of Neural Net outputs but demonstrates a factor of two global differences between data and Monte-Carlo mainly due to the π^0-veto ($P(\gamma|\pi^0)$) cut. In addition, the peaking background contribution is enhanced for the MC, compared to the real data.

A further step for this analysis would be to understand the difference between data and Monte-Carlo appearing in the control samples unblinding and proceed with signal unblinding and branching ratio estimation using available real-world datasets. The study of systematic uncertainties for selection applied will add more meaningful information to the branching ratio measurements and A_{ud} estimation.

As the next step after up-down asymmetry measurements, the fit of the large amount of Monte-Carlo events using "GamPola" can be performed to extract photon polarization along with other hadron parameters on the reconstruction level. Afterward, having a sufficient amount of real data of order $\sim 1\ \mathrm{ab}^{-1}$ (≈ 7500 events) the statistical uncertainty for the photon polarization λ_γ is estimated to be 0.03. Though, in order to compete with recent LHCb results, where polarization has been measured with $5\ \%$ of precision using $B^0 \to K^* e^+ e^-$ decay, an even more, larger sample $> 3\ \mathrm{ab}^{-1}$ is required.

Above the mentioned level of integrated luminosity should be reachable in a few years, in 2024 according to Fig. 5.1.

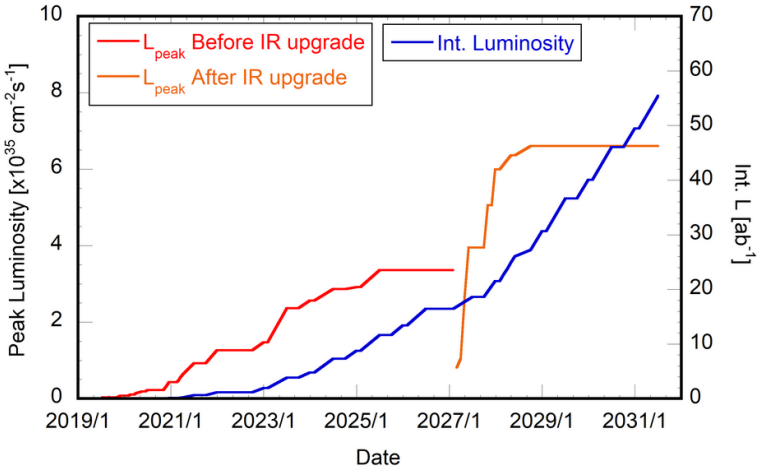


Figure 5.1. Expectations of Belle II luminosity recording.

In this thorny way towards the determination of the photon polarization, the results are given in this work represent an attempt for the establishment of the analysis procedure for the study of the $B^+ \to K^+\pi^-\pi^+\gamma$ decay. Within a future study, given the fantastic opportunity to further increase the Belle II statistics in this decay channel, it will be possible to tighten the determination of photon polarisation in $B \to K\pi\pi\gamma$ decays, providing a further test to the SM predictions.

Appendices

Replacing procedure for EvtGen signal events

This section explains extraction of the original EvtGen $B \to K\pi\pi\gamma$ events from the generic Monte-Carlo sample introduced in Sec. 4.2.

Within Belle II, the Monte-Carlo is produced in a centralized way. Generic Monte-Carlo is versioned and stored as shared read-only for the analyst resource. The size of such "raw" data is very large and is processed through steering *basf2*-script written by an analyst.

Since generic Monte-Carlo also includes signal B-decays, replacement original "EvtGen" $B^+ \to K^+\pi^+\pi^-\gamma$ events by "GamPola" events is not feasible. However, such replacement is possible after having processed "raw" (mdst) files through *basf2*-script. At this stage amount of obtained data is manageable even on a PC with 16 GB of RAM.

The procedure concerns $B^+ \to K^+\pi^+\pi^\gamma$ events obtained after selection cuts. The number of such events for given luminosity is defined by Eq. 4.1. And total efficiency is estimated in Sec. 4.2 for the sample of "GamPola" events. Thereby multiplying these two quantities gives a number of events which is put back to the sample after removing original "EvtGen" signal events.

Physics Sample split in $\mu_{\rm NN}$ bins

The Physics Sample obtained with the cut $\mu_{\rm NN} > 0.28$ can be split into the two bins $\mu_{\rm NN} \in [0.28, 0.64]$ and $\mu_{\rm NN} \in [0.64, 1.0]$ that are equally populated in TM events. The composition of these two bins is detailed here.

event type	$K\pi\pi\gamma$	TM	SCF	$q\bar{q}$	B^+B^-	$B^0\bar{B}^0$	$\tau\bar{\tau}$
$N_{\rm event}$	4426	3667	759	18248	1159	3129	13
%	16.4	13.6	2.8	68.0	4.2	11.6	0.04

Table 1. Composition of events with $M_{\rm bc} \in [5.20, 5.29]$ GeV/c^2 with cut $0.28 < \mu_{NN} < 0.64$.

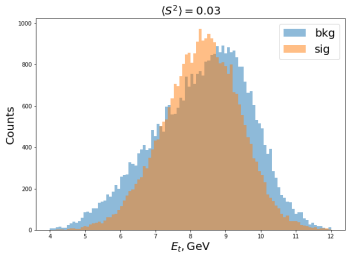
event type	$K\pi\pi\gamma$	TM	SCF	$q\bar{q}$	B^+B^-	$B^0\bar{B}^0$	$\tau\bar{\tau}$
$N_{\rm event}$	3953	3624	329	2891	338	1215	2
%	47.1	43.1	3.9	34.4	4.0	14.5	0.02

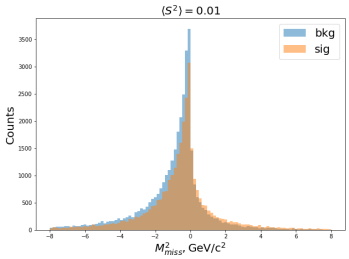
Table 2. Composition of events with $M_{\rm bc} \in [5.20, 5.29]$ GeV/c^2 with cut $0.64 < \mu_{NN} < 1$.

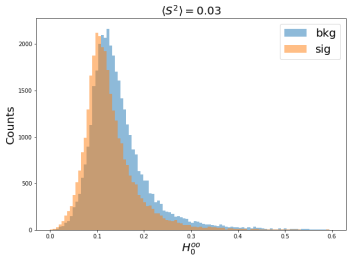
Kakuno Super Fox-Wolfram (KSFW) moments

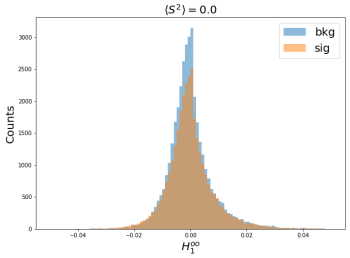
The Fox-Wolfram moments are used to describe the event shapes and defined in Eq. 4.19. In total there are 16 moments. In addition, there are scaled sum of transverse energy (E_T) and M^2_{miss} — missing squared invariant mass. The distributions of KSFW moments for signal and background components are

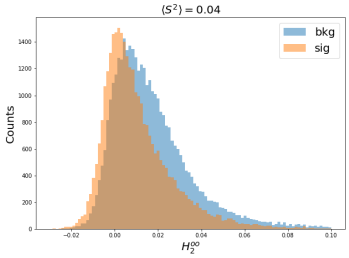
given below without any annotations. The decisions regarding whether to keep or discard a particular variable are taken in the following section.

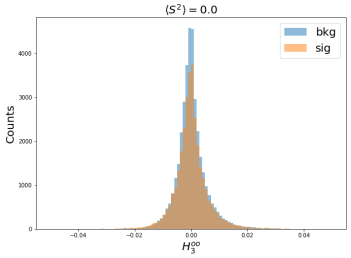


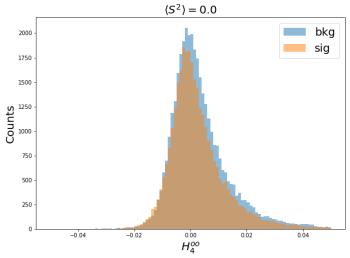


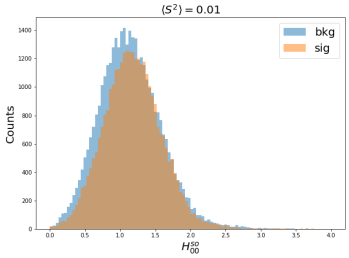


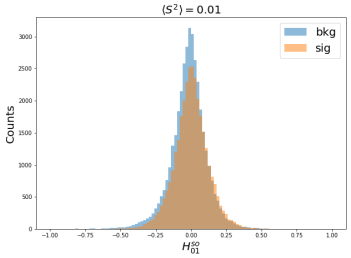


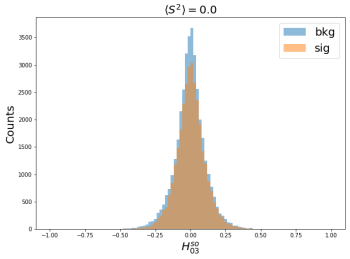


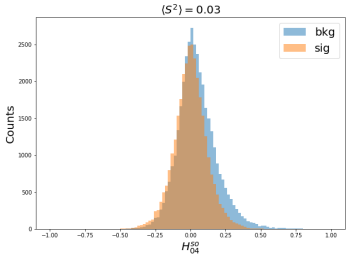


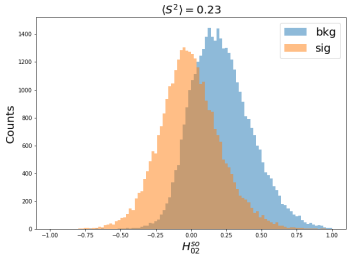


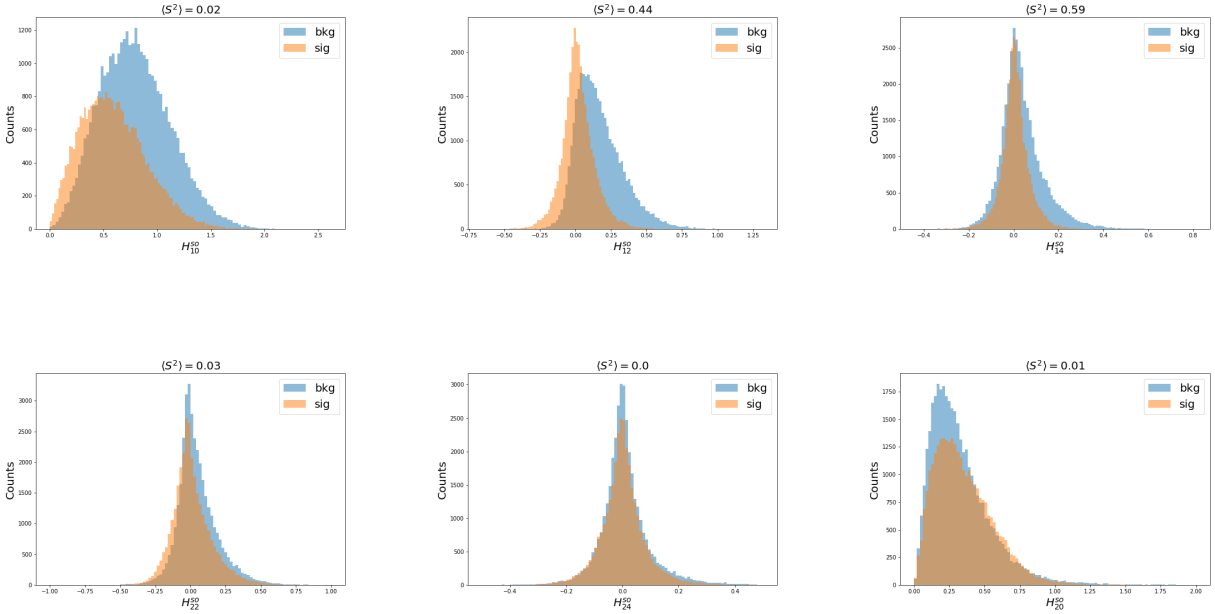












Features selection

Before any features' selection, correlation matrices (Figs. 2-3) for TM signal and continuum background show a strong correlation between some of the variables that can be used as NN inputs. The highest correlations are observed for certain Kakuno-Super-Fox-Wolfram moments and R_2 variables.

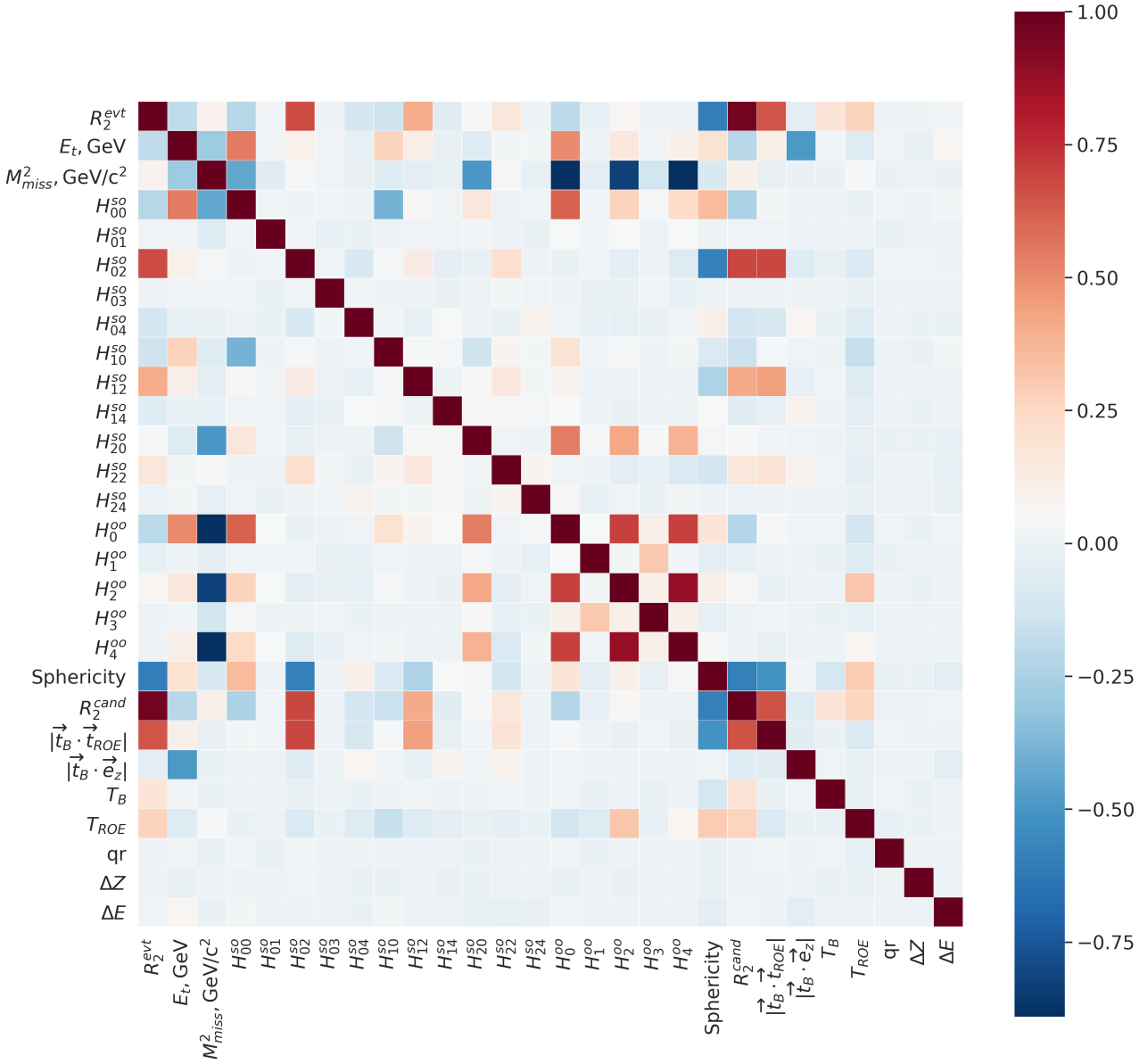


Figure 2. Correlation matrix of features for TM candidates of $B^+ \to K^+\pi^+\pi^-\gamma$

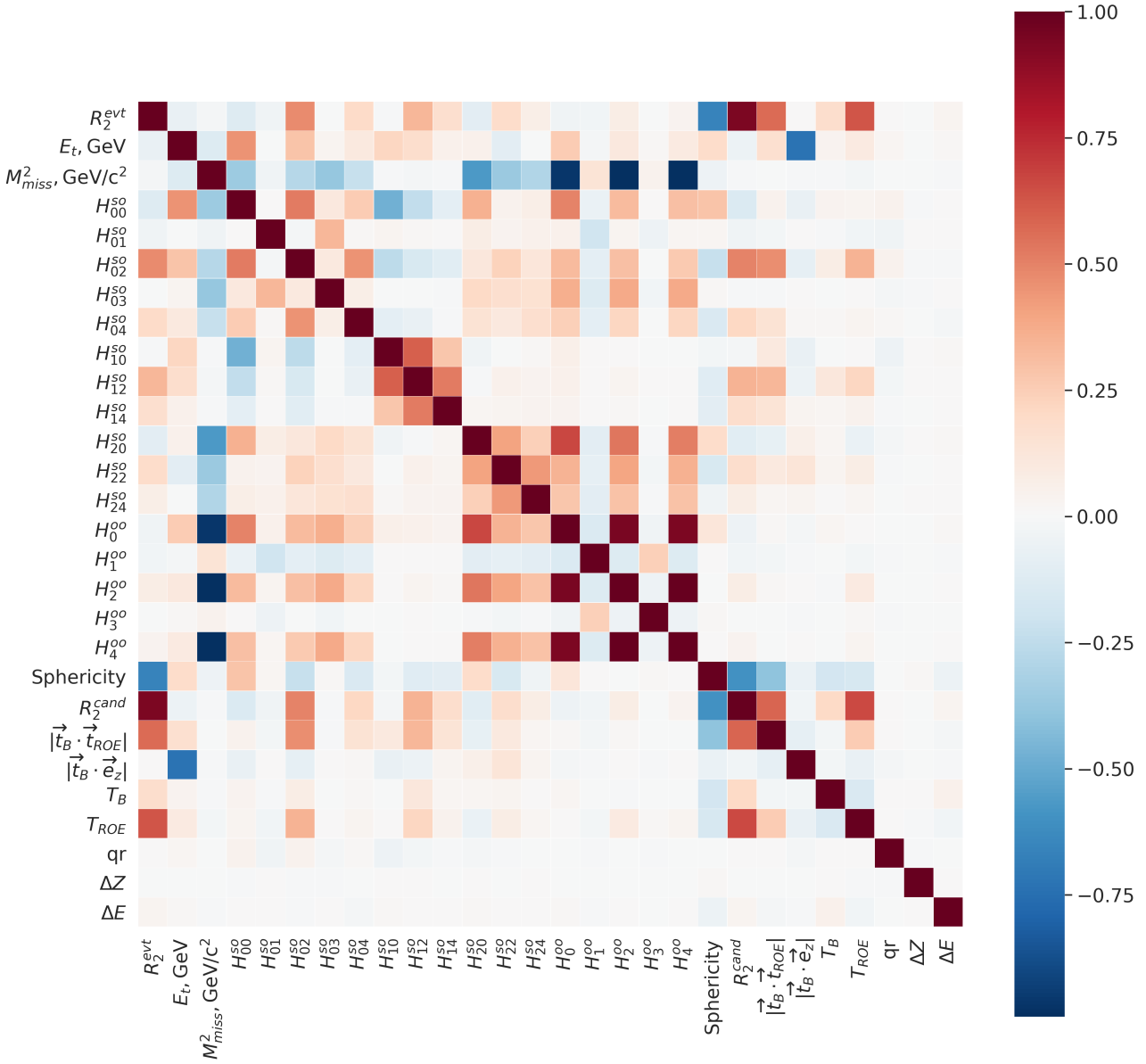


Figure 3. Correlation matrix of features for continuum background

The importance of each variable within a dataset containing 31409 signal TM events and 36882 continuum candidates is obtained using a random forest binary classifier [148]. Random forest is usually built out of decision trees. Each decision tree is represented by a directed acyclic graph and is grown recursively based on splitting criteria: the discriminating variable which maximizes the information gain after each split is selected first, thus such variable is more important than the others. The procedure of recursive splitting is repeated for the remaining variables. The growing is stopped once stopping criteria is met:

- Number of cases in the node is less than some pre-specified limit;

- Purity of the node is more than some pre-specified limit;
- Depth of the node is more than some pre-specified limit;
- Predictor values for all records are identical - in which no rule could be generated to split them;

In the case of highly correlated KSFW moments, most of these variables have low importances (see Fig. 4). In case of R_2 variables, R_2^{evt} calculated on the event level is dropped since it exhibits large correlation with R_2^{cand} and has less importance. Thrust value (T_B) of signal B-meson is excluded from the list of resulting Neural Net variables as it is strongly correlated with $M_{K^+\pi^+\pi^-}$ (see Fig. 4.83).

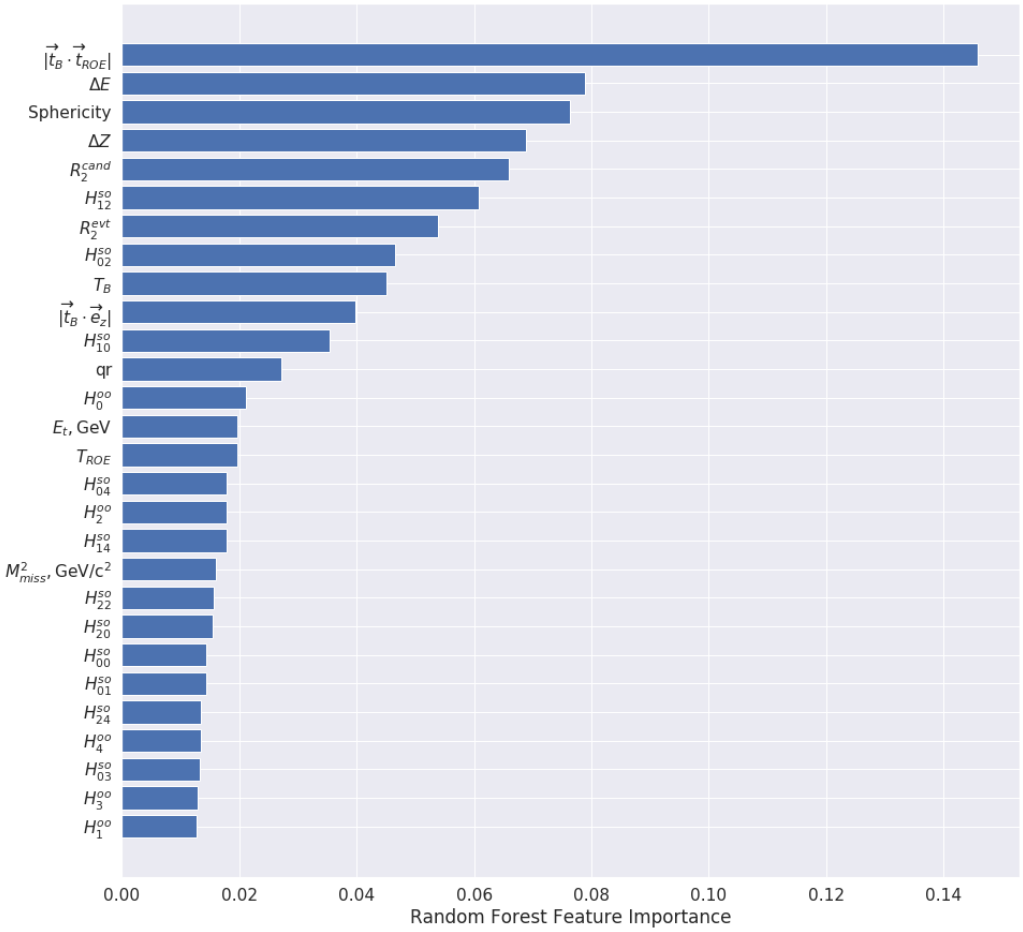


Figure 4. Features importance

Taking into account results obtained from Sec. 5 with separation defined by Eq. 4.24, it has been decided, that the similarity of distributions between signal and background should be $\langle S^2 \rangle > 0.01$.

After dropping weak and strongly correlated variables, a set of variables for the Neural Net input is obtained and sorted according to their strengths (see Fig. 7).

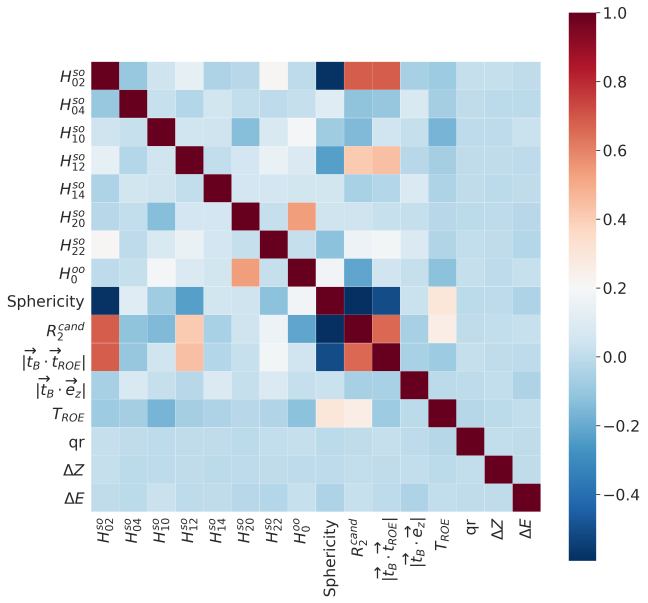


Figure 5. Correlation matrix of discriminating variables for TM candidates of $B^+ \to K^+\pi^+\pi^-\gamma$ events

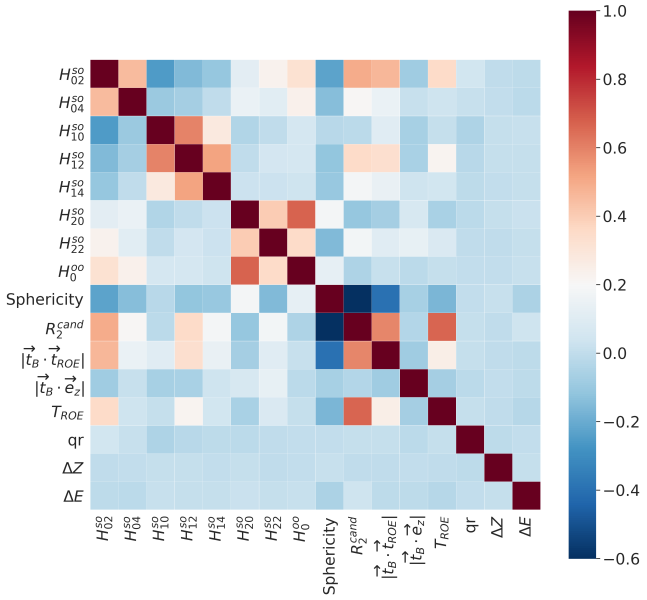


Figure 6. Correlation matrix of discriminating variables for continuum background

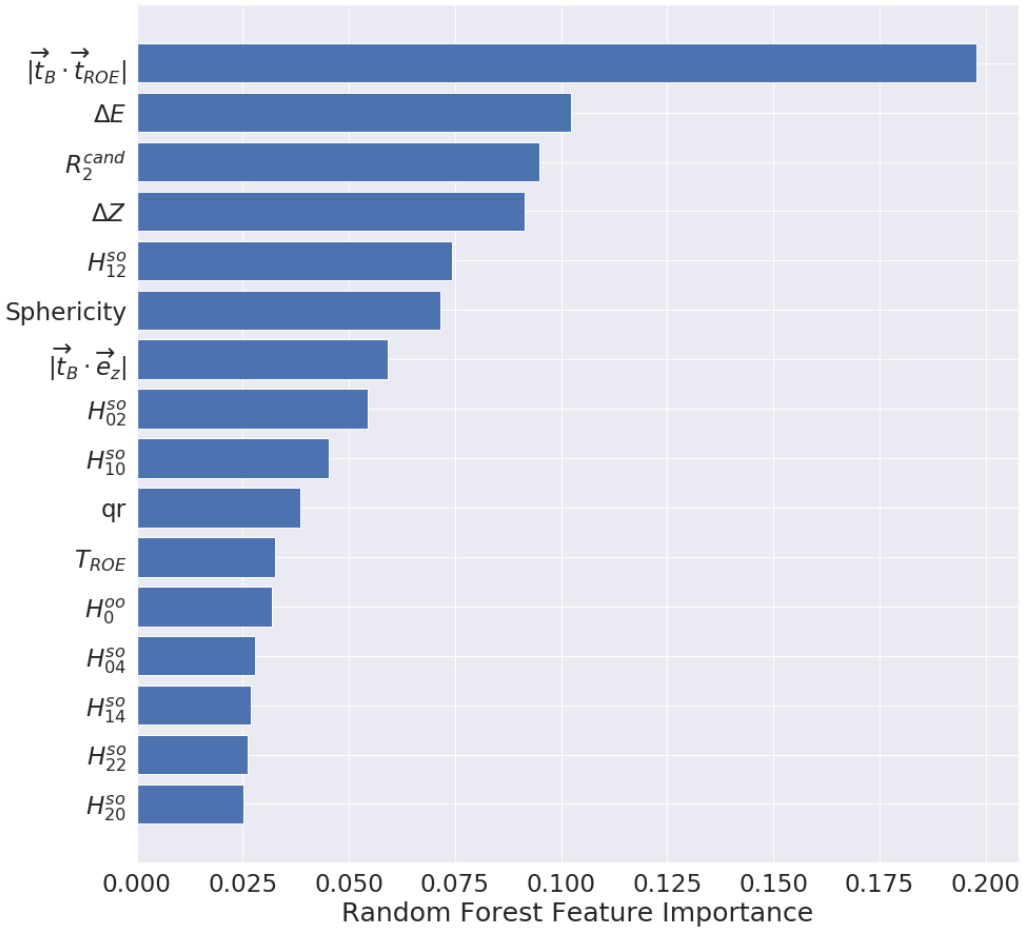


Figure 7. Sorted according to their importances, list of Neural Net variables.

P.D.F description

The probability density functions of the variable $M_{\rm bc}$ that are used in the analysis are normalized in the range $M_{\rm bc} \in [M_{\rm bc}^{\rm min}, M_{\rm bc}^{\rm max}]$ with $M_{\rm bc}^{\rm min} = 5.2$ GeV/c^2, and $M_{\rm bc}^{\rm max} = 5.29$ GeV/c^2. Their expressions are listed below:

1) The Argus p.d.f. can be written as:

$$\mathrm{Argus(M_{bc})} = \mathrm{c_{Argus}}\ \frac{\mathrm{M_{bc}}}{\mathrm{M_{bc}^{max}}}\ \sqrt{1-\left(\frac{\mathrm{M_{bc}}}{\mathrm{M_{bc}^{max}}}\right)^2}\ \exp\left[\chi\ (1-\left(\frac{\mathrm{M_{bc}}}{\mathrm{M_{bc}^{max}}}\right)^2)\right] \tag{1}$$

where χ is negative (in our case) and the normalization constant is

$$c_{\rm Argus} = \frac{(-2\chi)^{\frac{3}{2}}}{M_{\rm bc}^{\rm max}\ \Psi(u)} \tag{2}$$

$$\Psi(u) = \sqrt{\frac{\pi}{2}}\mathrm{Erf}(\sqrt{\frac{u}{2}}) - \sqrt{u}\ e^{-\frac{u}{2}} \quad ; \quad u = -2\chi\ (1-(\frac{M_{\rm bc}^{\rm min}}{M_{\rm bc}^{\rm max}})^2) \tag{3}$$

2) The Crystall Ball p.d.f. can be expressed as follows (for $\alpha > 0$ and $M_{\rm bc}^{\rm max} > \mu > M_{\rm bc}^{\rm min}$):

$$\mathrm{CB}(M_{\rm bc}) \;=\; \mathrm{cst}_{\rm CB} \begin{cases} \exp(-\frac{1}{2}t^2) & \text{if } t \geq -\alpha \\ \exp(-\frac{1}{2}\alpha^2)\ (1 - \frac{\alpha}{n}(\alpha + t))^{-n} & \text{if } t \leq -\alpha \end{cases}$$

where

$$t \;\equiv\; \frac{M_{bc} - \mu}{\sigma} \tag{4}$$

$$\mathrm{cst}_{\rm CB}^{-1} \;=\; \exp(-\frac{1}{2}\alpha^2)\frac{n}{n-1}\ \frac{\sigma}{\alpha}(1 - \left(\frac{n\sigma}{\alpha(\mu - M_{\rm bc}^{\rm min}) + \sigma(n - \alpha^2)}\right)^{n-1}) \tag{5}$$

$$+\sqrt{\frac{\pi}{2}}\sigma\ (\mathrm{Erf}(\frac{\alpha}{\sqrt{2}}) + \mathrm{Erf}(\frac{M_{\rm bc}^{\rm max} - \mu}{\sqrt{2}\sigma}) \tag{6}$$

The p.d.f. is illustrated in Fig.(8) for parameters values typical for the analysis.

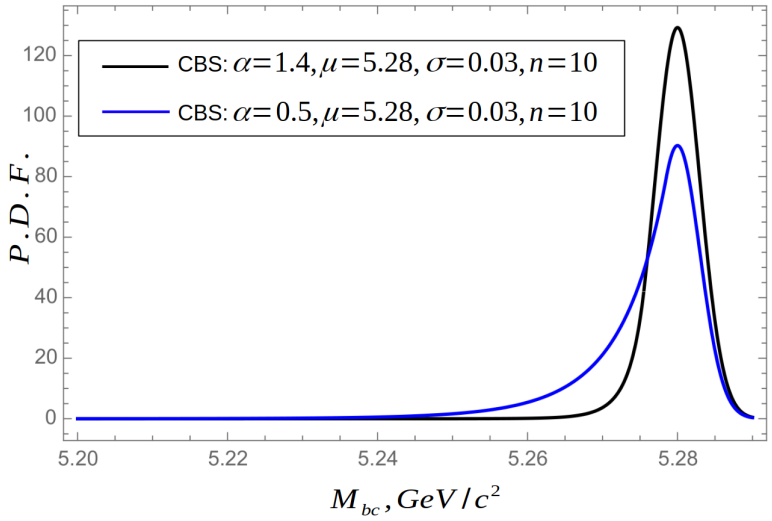


Figure 8. The CB p.d.f. , for typical values (black line) $\alpha = 1.4$, $\mu = 5.28$ GeV, $\sigma = 0.003$ and $n = 10$, and for the same values, except $\alpha = 0.5$.

3) Three p.d.f. 's are considered to describe the peaking background:

- The expression of the default p.d.f. (so-called modified Gaussian) was previously given in Eq.(4.43). It is repeated below:

$$\tilde{\mathrm{G}}(M_{\rm bc}) \;=\; \mathrm{cst}\ \exp\left(-\frac{(\mathrm{M_{bc}} - \mu)^2}{2\sigma^2}\right)\ \mathrm{Erf}\left((\frac{\mathrm{M_{bc}^{max}} - \mathrm{M_{bc}}}{\sigma})^{\kappa}\right) \tag{7}$$

where the normalization constant $\tilde{\mathrm{cst}}$ must be computed numerically.

- The Modified Gaussian of the second type (the generalized Argus p.d.f.) defined by:

$$\tilde{\mathrm{G}}_{(2)}(M_{\mathrm{bc}}) = \mathrm{cst}_{(2)}\ \frac{M_{\mathrm{bc}}}{M_{\mathrm{bc}}^{\max}}\ \left(1-\left(\frac{M_{\mathrm{bc}}}{M_{\mathrm{bc}}^{\max}}\right)^2\right)^p \exp\left[\chi\ (1-\left(\frac{M_{\mathrm{bc}}}{M_{\mathrm{bc}}^{\max}}\right)^2)\right] \tag{8}$$

$$\mathrm{cst}_{(2)} = \frac{1}{M_{\mathrm{bc}}^{\max}}\ \frac{2(-\chi)^{p+1}}{\Gamma(p+1)-\Gamma(p+1,-\chi(1-(\frac{M_{\mathrm{bc}}^{\min}}{M_{\mathrm{bc}}^{\max}})^2))} \tag{9}$$

  that is identical to the $\mathrm{Argus}(\mathrm{M_{bc}})$ p.d.f. for $p = 1/2$.

- The Modified Gaussian of the third type (hand-engineered p.d.f.) and defined by:

$$\tilde{\mathrm{G}}_{(3)}(M_{\mathrm{bc}}) = \mathrm{cst}\ \exp\left(-\frac{(M_{\mathrm{bc}}-\mu)^2}{2\sigma^2(1-\alpha(M_{\mathrm{bc}}-\mu))}\right) \tag{10}$$

$$\alpha = \begin{cases} 0 & \text{if } M_{\mathrm{bc}} < \mu \\ \frac{0.975}{M_{\mathrm{bc}}^{\max}-\mu} & \text{if } M_{\mathrm{bc}} \geq \mu \end{cases} \tag{11}$$

In Fig.(9) are represented $\tilde{\mathrm{G}}_{(3)}$ for $\mu = 5.2825$ and $\sigma = 0.01$; $\tilde{\mathrm{G}}$ for $\mu = 5.284$, $\sigma = 0.01$ and $\kappa = 0.5$; and $\tilde{\mathrm{G}}_{(2)}$ for $\chi = -500$ and $p = 1.5$.

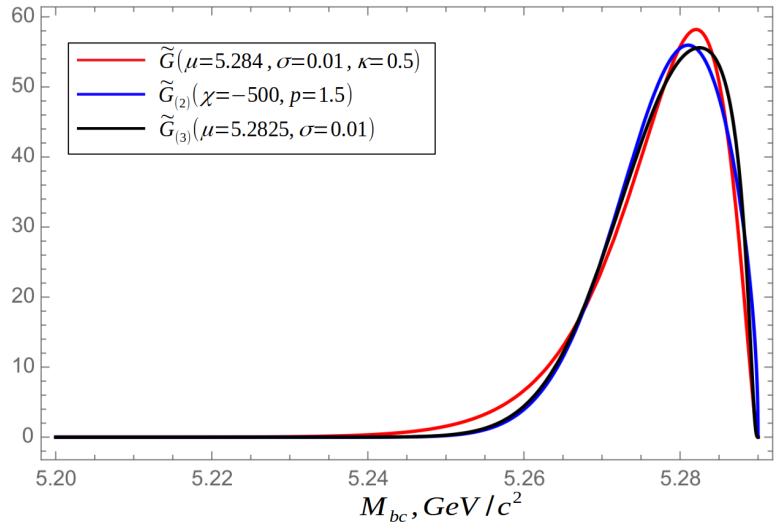


Figure 9. The parameters of $\tilde{\mathrm{G}}_{(3)}$ and $\tilde{\mathrm{G}}_{(2)}$ have been chosen to reproduce closely the $\tilde{\mathrm{G}}$ p.d.f.

Bibliography

[1] F. Close, *Particle Physics: A Very Short Introduction.*, ISBN 978-0-19-280434-1, Oxford University Press (2004)

[2] M. Gell-Mann, *The Eightfold Way: A Theory of Strong Interaction Symmetry*, doi:10.2172/4008239, Synchrotron Laboratory Report CTSL-20. California Institute of Technology (1961)

[3] G. Zweig, *An SU(3) Model for Strong Interaction Symmetry and its Breaking: II* CERN-TH-412, (1964)

[4] S. L. Glashow; J. Iliopoulos; L. Maiani . *Weak Interactions with Lepton–Hadron Symmetry*. Physical Review D. 2 (7): 1285–1292. doi:10.1103/PhysRevD.2.1285 (1970)

[5] Yang, C. N.; Mills, R. (1954), *Conservation of Isotopic Spin and Isotopic Gauge Invariance* Physical Review. 96 (1): 191–195. doi:10.1103/PhysRev.96.191

[6] S.L. Glashow (1961). *Partial-symmetries of weak interactions* Nuclear Physics. 22 (4): 579–588. doi:10.1016/0029-5582(61)90469-2

[7] S. Weinberg (1967). *A Model of Leptons* Physical Review Letters. 19 (21): 1264–1266. doi:10.1103/PhysRevLett.19.1264

[8] F. Englert; R. Brout (1964) *Broken Symmetry and the Mass of Gauge Vector Mesons* Physical Review Letters. 13 (9): 321–323. doi:10.1103/PhysRevLett.13.321

[9] P.W. Higgs (1964). *Broken Symmetries and the Masses of Gauge Bosons*. Physical Review Letters. 13 (16): 508–509. doi:10.1103/PhysRevLett.13.508

[10] Fermi, Enrico (1934). *Versuch einer Theorie der β-Strahlen. I.* Zeitschrift für Physik A. 88 (3–4): 161–177. doi:10.1007/BF01351864

[11] Kobayashi, M.; Maskawa, T. (1973) *CP-Violation in the Renormalizable Theory of Weak Interaction* Progress of Theoretical Physics. 49 (2): 652–657. doi:10.1143/PTP.49.652

[12] E. D. Bloom; et al. (1969), *High-Energy Inelastic e–p Scattering at 6^{deg} and 10^{deg}*. Physical Review Letters. 23 (16): 930–934. doi:10.1103/PhysRevLett.23.930

[13] J.J. Thomson *The Discharge of Electricity through Gases (1898)* Archibald Constable and Co.; 1st British Edition (January 1, 1898)

[14] Perrin, J. (1908). *Agitation moléculaire et mouvement brownien.* Comptes Rendus de l'Académie des Sciences, 146, 967–970.

[15] E. Rutherford *The scattering of alpha and beta particles by matter and the structure of the atom*, Philosophical Magazine, volume 21 (1911), pages 669-688.

[16] Davisson, C. J.; Germer, L. H. (1928). *Reflection of Electrons by a Crystal of Nickel*. Proceedings of the National Academy of Sciences of the United States of America. 14 (4): 317–322. doi:10.1073/pnas.14.4.317

[17] Chadwick, James (1932). *Existence of a Neutron.* Proceedings of the Royal Society A. 136 (830): 692–708. doi:10.1098/rspa.1932.0112

[18] Anderson, C. D. (1933). *The Positive Electron.* Physical Review. 43 (6): 491–494. doi:10.1103/PhysRev.43.491

[19] Casimir, H. B. G.; Polder, D. (15 February 1948). *The Influence of Retardation on the London-van der Waals Forces.* Physical Review. 73 (4): 360–372. Bibcode:1948PhRv...73..360C. doi:10.1103/PhysRev.73.360. ISSN 0031-899X

[20] *Testing the Standard Model.* 1990 Theoretical Advanced Study Institute in Elementary Particle Physics, Boulder, Colorado, 3 – 27 June 1990, https://doi.org/10.1142/1175 — June 1991

[21] Gliozzi, F.; Scherk, J.; Olive, D. I. (1977). *Supersymmetry, Supergravity Theories and the Dual Spinor Model.* Nucl. Phys. B. 122 (2): 253. doi:10.1016/0550-3213(77)90206-1

[22] S. Dimopoulos; H. Georgi (1981). *Softly Broken Supersymmetry and SU(5).* Nuclear Physics B. 193: 150–162. doi:10.1016/0550-3213(81)90522-8. hdl:2027.42/24165

[23] Witten, Edward (1995). *String theory dynamics in various dimensions*. Nuclear Physics B. 443 (1): 85–126. arXiv:hep-th/9503124. Bibcode:1995NuPhB.443...85W. doi:10.1016/0550-3213(95)00158-O. S2CID 16790997

[24] Randall, Lisa; Sundrum, Raman (1999). *Large Mass Hierarchy from a Small Extra Dimension*. Physical Review Letters. 83 (17): 3370–3373. arXiv:hep-ph/9905221. Bibcode:1999PhRvL..83.3370R. doi:10.1103/PhysRevLett.83.3370

[25] S. Fukuda; et al. (April 2003), *The Super-Kamiokande detector*. Nuclear Instruments and Methods in Physics Research A, 501 (2–3): 418–462, Bibcode:2003NIMPA.501..418F, doi:10.1016/S0168-9002(03)00425-X

[26] Hyper-Kamiokande collaboration *Hyper-Kamiokande Design Report*, arXiv:1805.04163 [physics.ins-det]

[27] IceCube Collaboration *The Design and Performance of IceCube DeepCore*, arXiv:1109.6096 [astro-ph.IM]

[28] Dama/Libra Collaboration *The DAMA/LIBRA apparatus*, arXiv:0804.2738 [astro-ph]

[29] DUNE collaboration *Deep Underground Neutrino Experiment (DUNE), Far Detector Technical Design Report, Volume II: DUNE Physics*, arXiv:2002.03005 [hep-ex]

[30] COSINE-100 collaboration *Design and Initial Performance of the COSINE-100 Experiment*, DOI:10.1140/epjc/s10052-018-5590-x

[31] NEWAGE collaboration *Direction-sensitive dark matter search with a low-background gaseous detector NEWAGE-0.3b* arXiv:2101.09921 [hep-ex]

[32] Antonio Boveia, Caterina Doglioni *Dark Matter Searches at Colliders*, arXiv:1810.12238 [hep-ex]

[33] Future Circular Collider collaboration *Future Circular Collider: Conceptual Design Report*. FCC Study Office. CERN. 2018. Retrieved 15 January 2019.

[34] Bevan, A. J.; Golob, B.; Mannel, Th; Prell, S.; Yabsley, B. D.; Aihara, H.; Anulli, F.; Arnaud, N.; Aushev, T. (2014-11-01). *The Physics of the B Factories*. The European Physical Journal C. 74 (11): 3026. arXiv:1406.6311. Bibcode:2014EPJC...74.3026B. doi:10.1140/epjc/s10052-014-3026-9. ISSN 1434-6044. S2CID 9063079

[35] Abe K., et al. Belle Collaboration (2001) *Observation of large CP violation in the neutral B meson system*. Phys. Rev. Lett. 87, 091802

[36] Aubert B., et al. Babar Collaboration (2001) *Observation of CP violation in the B^0 meson system*. Phys. Rev. Lett. 87, 091801.

[37] Belle II Collaboration *The Belle II Physics Book*. arXiv:1808.10567 [hep-ex]

[38] The Belle Collaboration *KEKB B-Factory Design Report* KEK Report 95-7, August 1995, NATIONAL LABORATORY FOR HIGH ENERGY PHYSICS

[39] Michael Gronau, Dan Pirjol *Photon polarization in radiative B decays*. arXiv:hep-ph/0205065

[40] M. Gronau, Y. Grossman, D. Pirjol, and A. Ryd, *Measuring the photon helicity in radiative B decays*, Phys. Rev. Lett. 88 (2002) 051802, arXiv:hep-ph/0107254.

[41] M. Gronau and D. Pirjol, *Photon polarization in radiative B decays*, Phys. Rev. D66 (2002) 054008, arXiv:hep-ph/0205065.

[42] Emi Kou, Alain Le Yaouanc, Andrey Tayduganov *Determining the photon polarization of the $b \to s\gamma$ using the $B \to K_1(1270)\gamma \to (K\pi\pi)\gamma$ decay*. arXiv:1011.6593 [hep-ph]

[43] The LHCb Collaboration Observation of Photon Polarization in the $b \to s\gamma$ Transition. DOI: 10.1103/PhysRevLett.112.161801

[44] Mostafa Hoballah *Measurement of the photon polarization using $B_s^0 \to \phi\gamma$ at LHCb*. Other [cond-mat.other]. Université Blaise Pascal - Clermont-Ferrand II, 2015. English. NNT: 2015CLF22555. tel-01176223.

[45] Giovanni Veneziano *Towards the measurement of photon polarisation in the decay $B^+ \to K^+\pi^-\pi^+\gamma$*. DOI:10.5075/epfl-thesis-6896

[46] Violaine Bellée *Amplitude analysis for the measurement of the photon polarisation in $B \to K\pi\pi\gamma$ decays*. DOI 10.5075/epfl-thesis-10183

[47] The LHCb Collaboration *Angular analysis of the* $B^0 \to K^{*0}e^+e^-$ *decay in the low-*q^2 *region* arxiv.org/pdf/1501.03038

[48] The LHCb Collaboration *Strong constraints on the* $b \to s\gamma$ *photon polarisation from* $B^0 \to K^{*0}e^+e^-$ *decays*. arXiv:2010.06011 [hep-ex]

[49] Andrey Tayduganov. *Electroweak radiative B-decays as a test of the Standard Model and beyond*. Université Paris Sud - Paris XI, 2011. English. NNT: 2011PA112195.tel-00648217

[50] Andrii Kotenko $B \to K\pi\pi\gamma$ *angular analysis*. 5-6 Octobre 2017, Lausanne, Switzerland

[51] Oleg Bezshyyko *Interaction of ionizating radiation with matter* Course of lectures, physics faculty, Taras Shevchenko National University of Kyiv

[52] The ATLAS Collaboration *Advanced Series on Directions in High Energy Physics: Volume 30* doi.org/10.1142/11030

[53] The CMS Collaboration *CMS Physics : Technical Design Report Volume 1: Detector Performance and Software* CERN-LHCC-2006-001, CMS-TDR-8-1

[54] SNO collaboration, B. Aharmim et al., *Electron energy spectra, fluxes, and day-night asymmetries of* 8^B *solar neutrinos from measurements with NaCl dissolved in the heavy-water detector at the Sudbury Neutrino Observatory*, Phys. Rev. C72 (2005) 055502.

[55] Super-Kamiokande collaboration, Y. Ashie et al., *Evidence for an oscillatory signature in atmospheric neutrino oscillation* Phys. Rev. Lett. 93 (2004) 101801, arXiv:hep-ex/0404034.

[56] Paul Dirac *The quantum theory of the electron*. doi.org/10.1098/rspa.1928.0023

[57] Greenberg, O. W. (1964-11-16). *Spin and Unitary-Spin Independence in a Paraquark Model of Baryons and Mesons*. Physical Review Letters. 13 (20): 598–602. doi:10.1103/PhysRevLett.13.598. ISSN 0031-9007

[58] Pauli, W. (1925). *Über den Zusammenhang des Abschlusses der Elektronengruppen im Atom mit der Komplexstruktur der Spektren*. Zeitschrift für Physik. 31 (1): 765–783. Bibcode:1925ZPhy...31..765P. doi:10.1007/BF02980631. S2CID 122941900

[59] C. S. Wu et al., *Experimental test of parity conservation in* β *decay* Phys. Rev. 105 (1957) 1413.

[60] R. L. Garwin, L. M. Lederman, and M. Weinrich, *Observations of the failure of conservation of parity and charge conjugation in meson decays: the magnetic moment of the free muon*, Phys. Rev. 105 (1957) 1415.

[61] G. Lüders, *Proof of the TCP theorem*, Ann. Phys. 2 (1957) 1.

[62] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, Phys. Rev. Lett. 13 (1964) 321.

[63] P. W. Higgs, *Spontaneous symmetry breakdown without massless bosons*, Phys. Rev. 145 (1966) 1156.

[64] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys. Lett. 12 (1964) 132.

[65] G. Zweig, *An SU(3) model for strong interaction symmetry and its breaking*, CERN-TH-401, 1964.

[66] B. Holstein *A brief introduction to chiral perturbation theory*

[67] R. P. Feynman (1950). *Mathematical Formulation of the Quantum Theory of Electromagnetic Interaction*. Physical Review. 80 (3): 440–57. Bibcode:1950PhRv...80..440F. doi:10.1103/PhysRev.80.440

[68] A. Falkowski *Lecture on Higgs Effective Field Theory* https://indico.ijclab.in2p3.fr/event/2565/

[69] Majorana, Ettore (1937). *Teoria simmetrica dell'elettrone e del positrone* Il Nuovo Cimento (in Italian). 14 (4): 171–84.

[70] Ross, G. (1984). *Grand Unified Theories*. Westview Press. ISBN 978-0-8053-6968-7.

[71] Wang, Juven (May 2021). *Ultra Unification*. Physical Review D. 103 (10): 105024. arXiv:2012.15860. Bibcode:2020arXiv201215860W. doi:10.1103/PhysRevD.103.105024. ISSN 2470-0029. S2CID 229923679

[72] J.L.Chkareuli, *SU(N) SUSY GUTS WITH STRING REMNANTS: MINIMAL SU(5) AND BEYOND*, Invited Talk given at 29th International Conference on High-Energy Physics (ICHEP 98), Vancouver, 23–29 July 1998. In *Vancouver 1998, High energy physics, vol. 2 1669–73

[73] Jeremy Hebinger *B Mesons Phenomenology to the Search of a Signal Beyond the Standard Model*. tel-01915429, 2017SACLS260

[74] Takeo Inami, C. S. Lim *Effects of Superheavy Quarks and Leptons in Low-Energy Weak Processes* $K_L \to \mu\bar{\mu}$, $K^+ \to \pi^+\nu\bar{\nu}$ *and* $K^0 \leftrightarrow \bar{K}^0$ doi.org/10.1143/PTP.65.297

[75] A. J. Buras, *Weak Hamiltonian, CP violation and rare decays*, arXiv:hep-ph/9806471v1.

[76] P.S. Bhupal Dev, Rabindra N. Mohapatra, Werner Rodejohann, Xun-Jie Xu, *Vacuum structure of the left-right symmetric model* arXiv:1811.06869 [hep-ph]

[77] Lisa Everett et al *Alternative approach to* $b \to s\gamma$ *in the uMSSM* JHEP01(2002)022

[78] The Belle Collaboration *Measurement of the* $B \to X_s\gamma$ *Branching Fraction with a Sum of Exclusive Decays* arXiv:1411.7198 [hep-ex]

[79] The Belle Collaboration *Time-Dependent CP Asymmetries in* $B^0 \to K^0_S\pi^0\gamma$ *Transitions* arXiv:hep-ex/0608017

[80] The BaBar Collaboration *Exclusive Measurements of* $b \to s\gamma$ *Transition Rate and Photon Energy Spectrum* arXiv:1207.2520 [hep-ex]

[81] The BaBar Collaboration *Precision Measurement of the $B \to X_s\gamma$ Photon Energy Spectrum, Branching Fraction, and Direct CP Asymmetry $A_{CP}(B \to X_s^+ d\gamma)$* arXiv:1207.2690 [hep-ex]

[82] The BaBar Collaboration *Measurement of the $B \to X_s\gamma$ Branching Fraction and Photon Energy Spectrum using the Recoil Method* arXiv:0711.4889 [hep-ex]

[83] The BaBar Collaboration *Measurement of Time-Dependent CP Asymmetry in $B^0 \to K_S^0\pi^0\gamma$ Decays* arXiv:0807.3103 [hep-ex]

[84] The LHCb Collaboration, *Measurement of CP-violating and mixing-induced observables in $B_s^0 \to \phi\gamma$ decays* arXiv:1905.06284 [hep-ex]

[85] S. Weinberg, *The quantum theory of fields. Vol. 2: Modern applications*, Cambridge University Press, 1996.

[86] J. C. Pati and A. Salam, *Lepton number as the fourth color*, Phys. Rev. D10 (1974) 275, erratum ibid. D11 (1975) 703.

[87] R. N. Mohapatra and J. C. Pati, *Left-right gauge symmetry and an isoconjugate model of CP violation*, Phys. Rev. D11 (1975) 566.

[88] G. Senjanovic and R. N. Mohapatra, *Exact left-right symmetry and spontaneous violation of parity*, Phys. Rev. D12 (1975) 1502.

[89] C. Lim and T. Inami, *Lepton flavor nonconservation and the mass generation mechanism for neutrinos*, Prog. Theor. Phys. 67 (1982) 1569.

[90] D. Becirevic, E. Kou, A. Le Yaouanc, A.Tayduganov, *Future prospects for the determination of the Wilson coefficient $C'_{7\gamma}$*, JHEP 08 (2012) 090.

[91] F. Muheim, Y. Xie, and R. Zwicky, *Exploiting the width difference in $B_s^0 \to \phi\gamma$*, Phys. Lett. B664 (2008) 174 .

[92] Belle collaboration, J. Li and et al., *Measurement of the time-dependent CP asymmetries in $B^0 \to K_S^0\rho^0\gamma$ decays*, Phys. Rev. Lett. 101 (2008) 251601.

[93] S. Akar, *Study of $B^+ \to K^+\pi^-\pi^+\gamma$ decays with the BaBar experiment: the photon helicity and the resonant structure of the $K^+\pi^-\pi^+$ system*, PhD thesis, Université Pierre et Marie Curie - Paris VI, 2013.

[94] The LHCb collaboration *Phys. Rev. Lett. 112, 161801 (2014)*

[95] A.Ryd et al. *EvtGen A Monte Carlo Generator for B-Physics* BAD 522 V6 EvtGen V00-11-06

[96] Clebsch, Alfred and Paul Gordan (1866). *Theorie der Abelschen Functionen.* Leipzig: Teubner.

[97] Baird, C.E.; L. C. Biedenharn (October 1964). *On the Representations of the Semisimple Lie Groups. III. The Explicit Conjugation Operation for SUn.* J. Math. Phys. 5 (12): 1723–1730. Bibcode:1964JMP.....5.1723B. doi:10.1063/1.1704095

[98] The Belle Collaboration Observation of $B+ \to K_1^{1270+}\gamma$. arXiv:hep-ex/0408138

[99] The BaBar Collaboration *Measurement of branching fractions of B decays to $K_1^{1270}\pi$ and $K_1^{1400}\pi$ and determination of the CKM angle α from $B_0 \to a_1^{1260\pm}\pi\mp$.* arXiv:0909.2171 [hep-ex]

[100] L. Burakovsky, T. Goldman *On D-wave meson spectroscopy and the $K^{*1410} - K^{*1680}$ problem.* doi.org/10.1016/S0375-9474(97)81461-3

[101] S. Descotes-Genon and B. Moussallam, *The $K^{0*}(800)$ scalar resonance from Roy-Steiner representations of πK scattering.* Eur. Phys. J. C 48 (2006) 553, arXiv:hep-ph/0607133v2

[102] J.R. Pelaez, A. Rodas, J. Ruiz de Elvira *Strange resonance poles from $K\pi$ scattering below 1.8 GeV* Eur. Phys. J. C 77, 91 (2017). doi.org/10.1140/epjc/s10052-017-4668-1

[103] The ACCMOR Collaboration, *Diffractive Production of Strange Mesons at 63 GeV* DOI:10.1016/0550-3213(81)90114-0

[104] The Belle Collaboration *Study of the $K^+\pi^+\pi^-$ Final State in $B^+ \to J/\psi K^+\pi^+\pi^-$ and $B^+ \to \psi' K^+\pi^+\pi^-$.* arXiv:1009.5256v2 [hep-ex]

[105] The 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* arXiv:1512.03579v1 [hep-ex]

[106] Emi Kou, Alain Le Yaouanc, Andrey Tayduganov *Determining the photon polarization of the $b \to s\gamma$ using the $B \to K_1(1270)\gamma \to (K\pi\pi)\gamma$ decay* arXiv:1011.6593 [hep-ph]

[107] S. U. Chung, *Spin formalisms.* Physics Department, Brookhaven National Laboratory, Upton, NY 11973

[108] M. Suzuki, *Strange axial-vector mesons*, Phys. Rev. D47 (1993) 1252–1255.

[109] G. J. Gounaris and J. J. Sakurai, *Finite-width corrections to the vector-meson-dominance prediction for $\rho \to e^-e^+$* Phys. Rev. Lett. 21, 244 (1968).

[110] *GiNAC* https://www.ginac.de/

[111] James, Fred and Winkler, Matthias *MINUIT User's Guide* CERN, Geneva, 6-th June, 2004

[112] Borys Knysh *"GamPola" software package (manual included)* https://github.com/basista94/GamPola

[113] E. Gamma, R. Helm, R. Johnson, and J. Vlissides. *Design Patterns: Elements of Reusable Object-Oriented Software* Addison-Wesley Professional, 1 edition, (1994)

[114] R. H. Dalitz (1953). *On the analysis of τ-meson data and the nature of the τ-meson.* Philosophical Magazine. 44 (357): 1068–1080. doi:10.1080/14786441008520365

[115] Geoffrey C. Fox and Stephen Wolfram *Observables for the Analysis of Event Shapes in e^+e^- Annihilation and Other Processes.* Phys. Rev. Lett. 41, 1581

[116] basf2 documentation: clusterE9E21

[117] The Belle II Collaboration *The Belle II Physics Book*

[118] P. Fischer et al., Nucl. Instrum. Meth., A582, 843–848 (2007).

[119] V. Balagura et al., Nucl. Instrum. Meth., A564, 590–596 (2006), arXiv:physics/0504194.

[120] R. Fruhwirth et al., Nucl. Instrum. Meth., A732, 95–98 (2013).

[121] R. Fruhwirth and A. Strandlie, Comput. Phys. Commun., 120(2-3), 197–214 (1999).

[122] W. Waltenberger et al., IEEE Transactions on Nuclear Science, 58, 434 – 444 (2011).

[123] T. Keck, CoRR, arXiv:1609.06119 (2016).

[124] Zernike, von F. *Beugungstheorie des schneidenver-fahrens und seiner verbesserten form, der phasenkontrastmethode* 10.1016/S0031-8914(34)80259-5

[125] Sergey Ioffe, Christian Szegedy *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift* arXiv:1502.03167 [cs.LG]

[126] Nair, Vinod and Hinton, Geoffrey E. *Rectified linear units improve restricted Boltzmann machines.* In ICML, pp. 807–814, 2010

[127] SridharNarayan *The generalized sigmoid activation function: Competitive supervised learning* doi.org/10.1016/S0020-0255(96)00200-9

[128] R. A. Fisher *On the mathematical foundations of theoretical statistics* doi/pdf/10.1098/rsta.1922.0009

[129] Rene Brun and Fons Rademakers, *ROOT - An Object Oriented Data Analysis Framework*, Proceedings AIHENP'96 Workshop, Lausanne, Sep. 1996, Nucl. Inst. & Meth. in Phys. Res. A 389 (1997) 81-86.

[130] K. Abe et al. (Belle Collaboration), Phys. Lett. B 511, 151 (2001).

[131] Kuhr, T., Pulvermacher, C., Ritter, M. et al. *The Belle II Core Software. Comput Softw Big Sci 3, 1 (2019)* doi.org/10.1007/s41781-018-0017-9

[132] Diederik P. Kingma, Jimmy Ba *Adam: A Method for Stochastic Optimization* arXiv:1412.6980 [cs.LG]

[133] Xin Qian, Diego Klabjan *The Impact of the Mini-batch Size on the Variance of Gradients in Stochastic Gradient Descent* arXiv:2004.13146 [math.OC]

[134] Dr. Krishna Jagannathan *Lecture 15: Sums of Random Variables* EE5110 : Probability Foundations for Electrical Engineers

[135] Muriel Pivk, Francois R. Le Diberder *sPlot: a statistical tool to unfold data distributions* arXiv:physics/0402083 [physics.data-an]

[136] Wouter Verkerke, David Kirkby *The RooFit toolkit for data modeling* arXiv:physics/0306116 [physics.data-an]

[137] Y. Ohnishi et al., PTEP, 2013, 03A011 (2013).

[138] M. Bona et al., SuperB (2007), arXiv:0709.0451.

[139] T. Abe, Belle II (2010), arXiv:1011.0352.

[140] M. Akatsu et al., Nucl. Instrum. Meth., A440, 124–135 (2000), arXiv:physics/9904009.

[141] M. Staric et al., Nucl. Instrum. Meth., A595, 252–255 (2008).

[142] M. Akatsu et al., Nucl. Instrum. Meth., A528, 763–775 (2004).

[143] L. L. Ruckman and G. S. Varner, Nucl. Instrum. Meth., A602, 438–445 (2009), arXiv:0805.2225.

[144] S. Nishida et al., Nucl. Instrum. Meth., A595, 150–153 (2008).

[145] S. Nishida et al., Nucl. Instrum. Meth., A766, 28–31 (2014).

[146] Ben Allison *Is there a rule-of-thumb for how to divide a dataset into training and validation sets?* https://stackoverflow.com/questions/13610074/is-there-a-rule-of-thumb-for-how-to-divide-a-dataset-into-training-and-validatio

[147] P.A. Zyla et al. *(Particle Data Group), Prog. Theor. Exp. Phys.2020, 083C01 (2020)*

[148] Breiman, L. *Random Forests.* Machine Learning 45, 5–32 (2001). https://doi.org/10.1023/A:1010933404324

Résumé étendu en français de la thèse

Cette thèse présente une étude du canal $B \to K\pi\pi\gamma$ en utilisant deux méthodes. La première est l'approche dépendante du modèle, qui est mise en œuvre dans le logiciel "GamPola", incorporant le générateur et l'ajusteur d'événements $B \to K\pi\pi\gamma$. La deuxième est l'analyse de la désintégration $B^+ \to K^+\pi^+\pi^-\gamma$, la masse de $K^+\pi^+\pi^-$ étant limitée à $< 1.8\ \mathrm{GeV/c^2}$ en utilisant $1\ \mathrm{ab}^{-1}$ de données Monte-Carlo, $63\ \mathrm{fb}^{-1}$ (énergie au centre de masse de $\Upsilon(4S)$) et $9.2\ \mathrm{fb}^{-1}$ (hors résonance) collectées avec le détecteur Belle II. L'analyse présentée dans cette thèse est appliquée aux échantillons de contrôle — échantillons de données réelles, ne contenant pas de contribution de signal.

Dans le chapitre "Theoretical overview", plus de détails sur le modèle standard sont donnés. On décrit un résumé de la théorie des groupes appliquée au SM (les groupes $U(1), SU(2), SU(3)$ sont discutés), les symétries discrètes et le théorème CPT sont couverts. L'impact de l'invariance de jauge du Lagrangien du SM est montré, mais la génération de la masse des bosons de jauge par le mécanisme de Higgs, les interactions faibles et fortes et les sources potentielles de contributions de la Nouvelle Physique sont présentées.

La deuxième partie du chapitre énumère les approches de l'état de l'art pour mesurer la polarisation des photons ainsi que le statut expérimental. Les mesures du rapport de branchement de la désintégration inclusive $B \to X_s\gamma$ effectuées par Belle et BaBar ont permis de restreindre expérimentalement C_7 et C_7'. Une autre approche a été mise en œuvre par Belle et BaBar pour $B^0 \to K^0_S\pi^0\gamma$ et LHCb pour $B^0_s \to \phi\gamma$. Les contraintes sur C_7 et C_7' sont déduites des mesures de l'asymétrie CP en fonction du temps.

Une autre méthode pour contraindre la polarisation du photon consiste à effectuer une analyse angulaire par rapport à la direction du photon. Par exemple, $B^0 \to K^{*0}(\to K^+\pi^-)\gamma$ ne peut pas être utilisé dans le cadre de cette approche, puisque dans le cadre de repos du méson B le plan $K-\pi$ est symétrique par rapport à la direction du photon. En revanche, la désintégration $B \to K_{res}(\to K\pi\pi)\gamma$ comporte 4 particules dans l'état final et l'hélicité du photon peut être mesurée. Un autre processus de désintégration prometteur, qui a été appliqué avec succès à l'analyse angulaire, est le processus $B^0 \to K^{*0}\gamma(\to e^+e^-)$, où le photon est virtuel.

Le chapitre "$B \to K\pi\pi\gamma$ decay modelling" aborde la question de la mesure de la polarisation des photons dans la désintégration $B \to K\pi\pi\gamma$. La désintégration $B \to K_{res}\gamma$ passe par plusieurs résonances kaoniques K_{res} suivies par la désintégration de K_{res} vers l'état final $K\pi\pi$. L'approche existante utilisée dans le cadre d'"EvtGen" ne permet pas de décrire le système $K\pi\pi$ en tenant compte de l'interférence des résonances kaoniques. On suppose plutôt que la probabilité de la désintégration est proportionnelle à une somme de carrés d'amplitudes $\mathcal{A}^i_{K_{res}}$ correspondant aux résonances intermédiaires K_{res}. Le taux de désintégration différentiel doit être décrit comme une somme des contributions des résonances kaoniques au niveau de l'amplitude. Par conséquent, une approche basée sur l'amplitude est nécessaire pour décrire correctement ce système.

L'approche est formalisée dans le progiciel "GamPola" (Gamma Polarization). Il inclut la modélisation de la désintégration, et en particulier, incorpore plusieurs résonances représentant les K_{res} ci-dessus. Sur la base du modèle et de ses paramètres ajustables, le générateur "GamPola" produit des événements $B \to K\pi\pi\gamma$. Le modèle décrit à la fois les modes neutres et chargés de la flèche $B \to K\pi\pi\gamma$ à travers les désintégrations de plusieurs résonances kaoniques K_{res} : $K_1^{1270}, K_1^{1400}, K_1^{*1410}, K_1^{*1680}, K_2^{1430}$ peuplant le

spectre de masse invariante de $K\pi\pi$.

La cinématique de la désintégration dans le cadre de repos de $K\pi\pi$ est décrite en termes de 5 variables indépendantes : $(s_{K\pi\pi}, s_{K\pi}, s_{\pi\pi}, \theta, \phi)$. Les expressions pour la conversion entre les cadres de repos B et $K\pi\pi$ sont fournies : tout en transformant la cinématique 5D dans le cadre de repos $K\pi\pi$ en un ensemble de 4-momenta de $K\pi\pi\gamma$ dans le cadre de repos B, 3 degrés de liberté supplémentaires sont introduits, représentant les angles d'Euler aléatoires de la direction absolue du système.

Le calcul théorique est effectué en termes de deux angles $\theta \in [0;\pi]$, $\phi \in [0;2\pi]$ et trois variables de Dalitz $s_{K^+\pi^+\pi^-}$, $s_{K^+\pi^+}$, $s_{K^+\pi^-}$. Il faut générer les événements en termes de quatre moments de la particule de l'état final car l'analyse des données en physique des particules se fait dans cette représentation conventionnelle de la cinématique (coupure sur le photon de haute énergie, coupure de qualité sur l'acceptation du calorimètre, etc). Une fois définis, ces 4-momenta peuvent être amplifiés dans n'importe quel autre cadre (laboratoire, centre de masse de la collision, etc), en supposant que le vecteur d'amplification soit donné. En particulier, le renforcement vers le cadre de repos du méson B est une étape nécessaire, puisque dans ce cadre les 4-momenta des particules de l'état final sont vers le générateur "EvtGen". De plus, puisque la direction d'un photon est définie le long de l'axe z dans le cadre de repos $K\pi\pi$, une randomisation supplémentaire simultanée des directions absolues de $K\pi\pi\gamma$ est nécessaire en utilisant trois angles d'Euler.

L'amplitude de la désintégration du méson B est factorisée comme une somme de produits des amplitudes de $B \to K_{res}\gamma$ et de $K_{res} \to K\pi\pi$ décrites séparément. La désintégration de K_{res} se produit soit par des désintégrations interférentes ultérieures de $K_{res} \to K^{*892}\pi$ ou de $K_{res} \to \rho^{770}K$. En cas de résonance 1^+, K_{res} est représenté par un mélange de K_1^{1270} et K_1^{1400} utilisant l'angle θ_{K_1}. Les facteurs de forme de $K_1 \to K^{*892}\pi$ et $K_1 \to \rho^{770}K$ sont paramétrés avec les phases hadroniques $\varphi_S^{K^*}, \varphi_S^{\rho}, \varphi_D^{K^*}$. Un couplage complexe est introduit dans le modèle, correspondant à la fraction de K_1^{1400} dans l'amplitude totale. Dans le cas de $K_1^{*1410/1430}, K_2^{1430}$ la paramétrisation est faite par des paramètres complexes du modèle représentant par les couplages des sous-décisions de $K_{res} \to VP$. Le paramètre de polarisation des photons λ_γ est introduit dans le modèle au stade de la définition du taux de désintégration dans les poids $\frac{1\pm\lambda_\gamma}{2}$ des taux de désintégration gauche et droit.

Le logiciel "GamPola" s'appuie sur le modèle défini et met en œuvre un générateur et un ajusteur d'événements $B \to K\pi\pi\gamma$. Le générateur produit des événements en deux étapes : il produit d'abord un événement cinématiquement autorisé qui se produirait en l'absence de toute interaction physique, puis il utilise la cinématique de l'événement pour générer l'événement final avec la physique incluse. La sortie du générateur est illustrée, et les distributions de la cinématique 5D sont comparées pour les modes de désintégration chargés et neutres.

Étant donné un ensemble d'événements générés par $B \to K\pi\pi\gamma$ on décrit comment l'ajusteur "GamPola" effectue la procédure inverse d'extraction des paramètres du modèle en utilisant l'ajustement de vraisemblance. Une approche pour normaliser efficacement la fonction de densité de probabilité dans la fonction de vraisemblance est réalisée en utilisant la bibliothèque "GiNAC" pour les calculs symboliques et permet de réduire considérablement le temps d'ajustement. Les distributions du générateur et du niveau d'ajustement de la cinématique 5D sont en bon accord ainsi que les valeurs des paramètres du modèle.

Après avoir mis en œuvre le générateur et le dispositif d'ajustement, le modèle de référence de la désintégration $B^+ \to K^+\pi^+\pi^-\gamma$ est défini à l'aide des histogrammes LHCb de $M_{K^+\pi^+\pi^-}, M_{K^+\pi^-}, M_{\pi^+\pi^-}$ et $\cos\theta_\gamma$. Ces derniers sont les histogrammes dans trois bins de $M_{K^+\pi^+\pi^-}$: $1.1 < M_{K^+\pi^+\pi^-} <$

$1.3\ \mathrm{GeV/c^2}$, $1.3 < M_{K^+\pi^+\pi^-} < 1.4\ \mathrm{GeV/c^2}$ et $1.4 < M_{K^+\pi^+\pi^-} < 1.6\ \mathrm{GeV/c^2}$. La dernière case $1.6 < M_{K^+\pi^+\pi^-} < 1,9\ \mathrm{GeV/c^2}$ de $K^+\pi^+\pi^-$ n'est pas inclus car "GamPola" est incapable d'ajuster les histogrammes dans cette région. Le modèle de base est ensuite utilisé pour l'étude de sensibilité.

L'impact de la masse invariante du système $K\pi\pi$ sur la mesure de la polarisation des photons est abordé. En particulier, il est montré que la coupure sur $M_{K^+\pi^+\pi^-} < 1.6\ \mathrm{GeV}$ n'affecte presque pas la sensibilité à la polarisation des photons et peut être appliquée sans risque au stade de la mesure de la polarisation des photons.

La sensibilité à la polarisation des photons pour différents modes de désintégration et les échantillons de données générés sont présentés. Dans le cadre de ce test, l'ajustement à l'aide du modèle Baseline de quatre échantillons est effectué : pour les modes de désintégration neutre et chargé, deux modèles avec des paramètres complexes et réels sont utilisés. Les résultats obtenus montrent une faible dépendance de la sensibilité à la polarisation des photons par rapport au choix du modèle et au mode de désintégration, avec une valeur typique de $\sigma_{\lambda_\gamma}(B^+ \to K^+\pi^+\pi^-\gamma) \approx 0.04$ pour 5000 événements.

Étant multidimensionnel, l'ajustement de "GamPola" est sondé sur le sujet de la stabilité en utilisant un ensemble d'événements générés à l'aide du modèle de base. La stabilité de l'ajustement est vérifiée en ajustant de manière répétée le même échantillon d'événements avec une graine aléatoire. Trois types de solutions sont identifiés : vrai — correspondant aux valeurs des paramètres du niveau du générateur, faux — correspondant aux valeurs élevées de $\Delta\chi^2 > 16$ et miroir — correspondant aux faibles $\Delta\chi^2 \leq 16$. Les fausses solutions sont éliminées en appliquant une coupe stricte sur le résultat de l'ajustement $\Delta\chi^2 \leq 16$ et produites dans $50\ \%$ des essais d'ajustement. Les solutions miroirs ne sont pas dangereuses et n'affectent pas le paramètre de polarisation des photons, qui est la principale préoccupation de la modélisation de la désintégration de $B \to K\pi\pi\gamma$.

L'estimation de la luminosité intégrée requise pour une précision compétitive de la mesure de la polarisation des photons est effectuée. En analysant le plan $C_7 - C_7'$, on conclut que la mesure de λ_γ est représentée sur ce plan par un double cercle avec des rayons intérieur et extérieur. La luminosité intégrée requise pour les mesures de polarisation des photons en compétition est estimée et est égale à près de $3\ \mathrm{ab^{-1}}$. Dans ce cas, la surface du disque de LHCb devrait être inférieure ou égale à la surface du double cercle de Belle II. Cependant, le calcul des systématiques est nécessaire et affectera une telle estimation.

Le chapitre "The Belle II Experiment" décrit le collisionneur SuperKEKB. La mise à niveau par rapport à KEKB et les améliorations correspondantes sont soulignées. Un bref aperçu du détecteur Belle II et de ses composants est donné. Le rôle de chaque système de sous-détecteur (dans la mesure du signal provenant de l'interaction des particules avec la matière) est clarifié.

Le chapitre "Study of the $B^+ \to K^+\pi^+\pi^-\gamma$ signal" couvre la désintégration $B^+ \to K^+\pi^+\pi^-\gamma$ d'un point de vue expérimental. L'analyse de la désintégration de $B^+ \to K^+\pi^+\pi^-\gamma$ avec une masse de $K^+\pi^+\pi^-$ dans la gamme $[1, 1.8]\ \mathrm{GeV/c^2}$ est effectuée pour l'ensemble de données de Monte-Carlo de taille $1\ \mathrm{ab^{-1}}$ et l'ensemble de données du monde réel composé de ensemble de données du monde réel composé de $9.2\ \mathrm{fb^{-1}}$ de données hors résonance et de $62.7\ \mathrm{fb^{-1}}$ de données $\Upsilon(4S)$ collectées par Belle II à la fin de l'été 2020. Pour le grand échantillon ($1\ \mathrm{ab^{-1}}$) de données MC génériques et la valeur PDG d'entrée du rapport de branchement $2.58 \pm 0.15 \cdot 10^{-5}$, les événements $B^+ \to K^+\pi^+\pi^-\gamma$ générés par "EvtGen" sont remplacés par le modèle "GamPola" incorporé dans le "EvtGen". La procédure de remplacement est effectuée après l'application des coupes de présélection. Un nouvel échantillon obtenu de données MC génériques est ensuite utilisé dans l'analyse.

Orientée pour la mesure de la polarisation des photons, l'analyse porte sur un ensemble d'événements cibles correctement reconstruits du point de vue cinématique et définis dans ce chapitre. Il est conclu que pour l'objectif final de ce programme d'analyse, la principale préoccupation de l'étude MC devrait être les événements appariés à la vérité (TM). Par conséquent, un raisonnement supplémentaire est effectué afin d'obtenir un tel ensemble d'événements.

Deux variables puissantes et théoriquement faiblement corrélées utilisées dans la physique des B pour discriminer les événements de signal et de fond sont la masse de contrainte du faisceau (M_{bc}) et la différence d'énergie entre le méson B reconstruit et le faisceau e^+e^- (ΔE). Dans le cas où le candidat méson B est parfaitement reconstruit, en l'absence de dispersion de l'énergie du faisceau, l'énergie reconstruite de B devrait être égale à l'énergie du faisceau et la masse invariante reconstruite est égale à la masse nominale de B.

Dans un premier temps, la sélection des événements est décrite en termes d'efficacité des coupes correspondantes sur les mesurables physiques. Plusieurs groupes de coupes sont considérés, y compris l'écrémage, la présélection et la sélection, ce qui donne au total $42\ \%$ d'efficacité de sélection des événements $B^+ \to K^+\pi^+\pi^-\gamma$.

Les coupes d'écrémage sont représentées par une sélection générale avant les coupes appliquées spécifiquement dans le cadre de l'analyse particulière de la désintégration de B. Elles permettent de réduire de manière significative la consommation du CPU lors de la reconstruction d'un mode de désintégration particulier. Dans le cas de la décroissance $B^+ \to K^+\pi^+\pi^-\gamma$, au moins trois pistes chargées ($nTracks \geq 3$), et un photon hautement énergétique sont nécessaires dans l'état final.

Les signaux provenant des photons sont détectés par l'ECL sous forme d'amas de cristaux activés. L'énergie des photons ainsi que les informations sur les clusters sont utilisées pour la sélection. Une piste se propageant dans le vide dans un champ magnétique constant se déplace le long d'une trajectoire en hélice. Le point d'approche le plus proche pour une piste donnée est choisi dans le système cylindrique de coordonnées (r, φ, z) pour minimiser la distance entre un point d'interaction et la piste. La distance contient des composantes transversales non signées (dr) et longitudinales signées (dz) par rapport à l'axe du faisceau z. Pour réduire le bruit de fond, les événements sont soumis à des coupes de sélection.

Bien que le rapport d'embranchement ne soit pas la préoccupation finale de cette analyse, sa mesure est une étape intermédiaire nécessaire. La mesure précise des rendements des événements de signal et de fond est donc une étape importante.

L'application des coupes permet déjà d'observer un pic indiquant un signal. La source possible de corrélation entre M_{bc} et ΔE est identifiée comme représentant l'énergie du photon, qui est reconstruite avec une différence de niveau de $\ \%$ par rapport à sa valeur réelle. Le passage de l'énergie des photons de la valeur réelle à la valeur reconstruite explique la différence des distributions de M_{bc} et de ΔE entre la valeur réelle et la valeur reconstruite.

Différents arrière-plans sont étudiés, les principales sources sont identifiées et supprimées. Cependant, une étude plus complète peut être réalisée pour purifier davantage l'ensemble des données obtenues. Le fond continu est supprimé à l'aide d'un réseau neuronal artificiel. La plupart des entrées du réseau neuronal sont des variables de forme d'événement se référant implicitement aux candidats : elles sont donc différentes pour chaque candidat au sein de l'événement. En tenant compte du fait que cette analyse est "untagged" (seul le signal B est reconstruit), le reste de l'événement (ROE) est construit autour d'un signal B et est une collection de toutes les particules provenant du méson B qui l'accompagne. Ainsi, en fonction du candidat

considéré pour l'événement, le ROE est modifié en conséquence, représentant un ensemble différent de particules pour chaque candidat méson B du signal. Le modèle est entraîné en utilisant une quantité égale de candidats mésons signal et d'événements $q\bar{q}$.

Le candidat au sein de l'événement, qui est une combinaison quelconque de particules d'état final $K^+\pi^+\pi^-\gamma$ et qui satisfait à des critères de sélection donnés, est ensuite comparé aux autres candidats. Le critère visant à sélectionner le candidat le plus proche de l'événement produit est appelé sélection du meilleur candidat (BCS). Après avoir passé les étapes du logiciel de détection et de reconstruction, l'événement de Monte-Carlo généré est traduit en candidat TM si les informations du niveau du générateur correspondent à celles de la reconstruction. Ainsi, la sélection du meilleur candidat est effectuée pour obtenir le candidat le plus proche du candidat TM.

Le NN est entraîné en utilisant les candidats TM comme échantillon de signal. Il fournit un moyen de réduire le bruit de fond $q\bar{q}$ et également d'effectuer un BCS en définissant le "meilleur candidat" comme celui ayant la plus grande sortie NN. La sortie la plus élevée correspond à une plus grande probabilité que des valeurs d'entrée données de NN correspondent au candidat TM. Pour les événements de signal, si le meilleur candidat correspond à la vérité, l'événement est appelé TM, tandis que si le BCS est FM, l'événement est appelé SCF, pour self cross-feed. La sélection du meilleur candidat est effectuée à l'aide de la sortie du NN, ce qui permet de conserver $93\ \%$ des événements TM.

Une fois la sélection BCS effectuée, une coupure est appliquée à la sortie du NN afin de réduire de manière drastique la contamination provenant du fond $q\bar{q}$. La coupure n'est pas appliquée sur les sorties originales du réseau neuronal, mais plutôt sur leur version transformée, où la valeur de la coupure peut être interprétée comme une fraction d'événements TM rejetés. La sortie du réseau neuronal d'origine est ensuite transformée à l'aide de μ-transfrom, de sorte que la distribution de μ_{NN}^{TM} soit plate. Dans cette nouvelle représentation, l'application de la coupure optimale $\mu_{NN} > 0.28$ permet de conserver 7386 événements TM correspondant à $1\ \mathrm{ab}^{-1}$ de données MC. En raison de la corrélation entre μ_{NN} et M_{bc}, il ne faut pas s'attendre à ce que l'efficacité de la sélection soit constante en fonction de M_{bc}, ni pour les événements TM, ni pour les événements de fond. Alors que pour les événements $q\bar{q}$ l'efficacité est décroissante, elle est croissante pour $B\bar{B}$ et SCF, ce qui n'est pas souhaitable.

Les principales sources de bruit de fond dans la région du signal de M_{bc} sont identifiées. Ces sources sont dues à la perte ou à l'échange de pistes provenant d'événements dont la cinématique est similaire à celle des événements du signal. Bien que la réduction du bruit de fond $q\bar{q}$ ait été considérée comme prioritaire ci-dessus, il est également essentiel de s'assurer que le fond de pic dans la région du signal est sous contrôle. La contribution la plus importante provient des événements $B^0\bar{B}^0$. Bien que les événements $B^+\bar{B}^-$ et SCF présentent également des pics, leurs contributions sont beaucoup plus faibles.

Cette partie du travail explique également la procédure d'ajustement de l'échantillon obtenu afin de mesurer les rendements du signal et des autres espèces d'événements. La procédure d'accès aux distributions des variables soustraites du bruit de fond au niveau des mésons B et des systèmes $K\pi\pi$ est discutée et appliquée. L'ajustement de la vraisemblance étendue est effectué en utilisant l'échantillon physique et trois composantes de la p.d.f. totale. L'introduction de p.d.f. séparées pour les événements TM permet de mesurer directement le rendement attendu de ces événements et, plus important encore, de récupérer les distributions TM de diverses observables en utilisant ${}_s\mathcal{P}lot$. Le rendement ajusté est de 7400 ± 203 événements, ce qui est en très bon accord avec la valeur MC-vérité, mais la corrélation entre cette quantité et le rendement du fond de pic est très élevée. Bien que tFit+CS1 ($K^+\pi^-\pi^-\gamma$ dans l'état final) réduise

l'incertitude statistique du rendement TM de $25\ \%$ par rapport au tFit, ce dernier est considéré comme l'ajustement de référence à ce stade de Belle II.

Une étude de Monte-Carlo est effectuée pour $63\ \mathrm{fb}^{-1}$ démontrant que l'estimateur de vraisemblance n'est pas significativement biaisé avec $\langle z \rangle = +0.01 \pm 0.03$ et $\sigma[z] = 0.975 \pm 0.02$.

La mesure des rapports d'embranchement de la désintégration de $B^+ \to K^+\pi^+\pi^-\gamma$ sur des échantillons MC correspondant à $62.7\ \mathrm{fb}^{-1}$ et $1\ \mathrm{ab}^{-1}$ est effectuée en utilisant deux approches différentes et les résultats sont comparés à la valeur PDG d'entrée. Pour l'échantillon de $62.7\ \mathrm{fb}^{-1}$, le rapport d'embranchement est égal à $\mathrm{Br}_{K^+\pi^+\pi^-\gamma} = 2.65 \pm 0.27 \cdot 10^{-5}$ et est calculé sans tenir compte de la dépendance possible de l'espace des phases cinématique. Une telle dépendance est prise en compte en divisant l'échantillon de $1\ \mathrm{ab}^{-1}$ en 125 bins dans les variables de Dalitz et en calculant le rendement et l'efficacité pour chaque bin. La valeur obtenue du rapport d'embranchement dans le cadre de la deuxième approche est $\mathrm{Br}_{K^+\pi^+\pi^-\gamma} = 2.52 \pm 0.07 \cdot 10^{-5}$ et les incertitudes des rapports d'embranchement obtenus par les deux approches sont en bon accord.

Discussion d'une étape essentielle vers la mesure de la polarisation des photons, où l'on mesure la direction préférée du photon dans le cadre de repos $K\pi\pi$. En tant qu'étape intermédiaire, l'asymétrie haut-bas est calculée dans les trois bins de $M_{K\pi\pi}$ en utilisant la technique ${}_s\mathcal{P}lot$ et le comptage MC-vérité. Les valeurs obtenues de A_{ud} sont égales dans les limites des incertitudes statistiques. Cependant, la comparaison avec les résultats de LHCb donne une grande incohérence. Cela peut s'expliquer par le fait que le modèle de ligne de base obtenu à partir de l'ajustement des histogrammes de LHCb ne reproduit ces histogrammes que visuellement.

La levée de doute sur les données réelles de $62.7\ \mathrm{fb}^{-1}$ ($\Upsilon(4S)$) et $9.2\ \mathrm{fb}^{-1}$ (hors résonance) est effectuée pour plusieurs échantillons de contrôle et les résultats sont comparés avec Monte-Carlo. Le démêlage des échantillons de contrôle est effectué en utilisant la résonance externe, la bande latérale, CS1 et CS2. Le débouclage montre une bonne concordance entre les formes des distributions des sorties du réseau neuronal, mais présente des différences globales d'un facteur deux entre les données et Monte-Carlo, principalement dues à la coupure π^0-veto ($P(\gamma|\pi^0)$). En outre, la contribution du bruit de fond est renforcée pour le MC, par rapport aux données réelles.

Une étape supplémentaire de cette analyse consisterait à comprendre la différence entre les données et les résultats de Monte-Carlo dans le démêlage des échantillons de contrôle et à procéder au démêlage du signal et à l'estimation du rapport de branchement en utilisant les ensembles de données réelles disponibles. L'étude des incertitudes systématiques pour la sélection appliquée ajoutera des informations plus significatives aux mesures du rapport de branchement et à l'estimation de A_{ud}.

Après les mesures de l'asymétrie haut-bas, l'ajustement de la grande quantité d'événements Monte-Carlo à l'aide de "GamPola" peut être effectué pour extraire la polarisation des photons ainsi que d'autres paramètres hadroniques au niveau de la reconstruction. Ensuite, avec une quantité suffisante de données réelles de l'ordre de $\sim 1\ \mathrm{ab}^{-1}$ (≈ 7500 événements), l'incertitude statistique pour la polarisation des photons λ_γ est estimée à 0.03. Cependant, afin de rivaliser avec les résultats récents du LHCb, où la polarisation a été mesurée avec une précision de $5\ \%$ en utilisant la désintégration $B^0 \to K^* e^+ e^-$, un échantillon encore plus grand $> 3\ \mathrm{ab}^{-1}$ est nécessaire. Ce niveau de luminosité intégrée devrait pouvoir être atteint en 2024.

Dans ce chemin épineux vers la détermination de la polarisation des photons, les résultats sont donnés dans ce travail représentent une tentative pour l'établissement de la procédure d'analyse pour l'étude de la

désintégration $B^+ \to K^+\pi^-\pi^+\gamma$. Dans le cadre d'une étude future, étant donné la fantastique opportunité d'augmenter encore les statistiques de Belle II dans ce canal de désintégration, il sera possible de renforcer la détermination de la polarisation des photons dans les désintégrations $B \to K\pi\pi\gamma$, fournissant un test supplémentaire aux prédictions du SM.