UNIVERSITÀ DEGLI STUDI DI MILANO

Department of Physics

PHD SCHOOL IN PHYSICS, ASTROPHYSICS AND APPLIED PHYSICS

Cycle XXXIII

Search for pentaquark candidates in $B^0_{(s)} \to J/\psi\, p\bar{p}$ decays at LHCb

Disciplinary Scientific Sector FIS/01 and FIS/04

Supervisor: Professor Nicola NERI

Director of the School: Professor Matteo PARIS

PhD Thesis of:

Elisabetta SPADARO NORELLA

A. Y. 2019-20

Commission of the final examination:

External Referees:
Dr. Mikhail MIKHASHENKO - CERN
Prof. Adam SZCZEPANIAK - Indiana University

External Members:
Prof. Marta CALVI - Università di Milano Bicocca
Prof. Antonio Davide POLOSA - Università di Roma "La Sapienza"
Prof. Sebastian NEUBERT - Universität Bonn

Final examination:

March 25th, 2021

Università degli Studi di Milano, Dipartimento di Fisica, Milano, Italy

Cover illustration:

Pentaquark candidates at LHCb
(copyright: Song Chen/Tsinghua Spotlights)

MIUR subjects:

FIS/01 - FIS/04

PACS:

13.25.Hw

Contents

Abstract

This thesis presents the analysis of the $B^0_{(s)} \to J/\psi p\bar{p}$ decays, which have been performed on the data collected by the LHCb experiment at CERN, during Run 1 and Run 2. These decay modes are particularly attractive for the searches of exotic states, such as pentaquarks and glueballs, which could occur in the final states $J/\psi p$ ($J/\psi \bar{p}$) and $p\bar{p}$, respectively. Pentaquarks are states composed of five quarks, observed for the first time by LHCb in the decay of $\Lambda_b \to J/\psi p K^-$ in 2015. Glueballs are states made up only of gluons, which are the mediators of the strong force. Although they are predicted by the Standard Model of particle physics, they have not been observed yet. The field of exotic spectroscopy has gained increasing interest over the last few years, notably after the first observation of pentaquark states made by LHCb. Since Quantum Chromodynamics (QCD) does not work in the perturbative regime at distance scales that are characteristic of observable hadrons, a number of non-perturbative QCD-inspired models have been developed to describe the substructures of these exotic states. Depending on the model, multiquark states could either be tight bound states or molecules made up of hadrons. Therefore, it is crucial to find new experimental evidence that could help discriminate among different theoretical models.

In this thesis, the observation of $B^0_{(s)} \to J/\psi p\bar{p}$ decays [1] is presented using data collected by LHCb between 2011 and 2016 and corresponding to a luminosity of 5.2 fb^{-1}. The branching ratio of these modes is measured to be:

$$\mathcal{B}(B^0 \to J/\psi p\bar{p}) = (4.51 \pm 0.40\ (\text{stat}) \pm 0.44\ (\text{syst})) \times 10^{-7},$$
$$\mathcal{B}(B^0_s \to J/\psi p\bar{p}) = (3.58 \pm 0.19\ (\text{stat}) \pm 0.33\ (\text{syst})) \times 10^{-6},$$

where the first uncertainty is statistical and the second one is systematic. A statistical significance greater than 10σ is obtained for these modes. The limited phase space, which characterizes this decay, allowed us to perform the most precise $B^0_{(s)}$ mass measurements at the publication date. The $B^0_{(s)}$ masses are measured to be $5366.85 \pm 0.19\ (\text{stat}) \pm 0.13\ (\text{sys})$ MeV and $5279.74 \pm 0.30\ (\text{stat}) \pm 0.10\ (\text{sys})$ MeV for the B^0_s and B^0, respectively.

To investigate the presence of exotic resonances, we perform a full amplitude analysis of the $B^0_s \to J/\psi p\bar{p}$ decay using the helicity formalism and study the properties of intermediate resonances, such as masses, widths and quantum numbers, from the description of the angular distribution of the decays.

The study is based on about 900 signal candidates, selected from the data collected by LHCb in the period of time between 2011 to 2018 and corresponding to 9 fb^{-1} of luminosity. We find evidence of a new state decaying to $J/\psi p$ and $J/\psi \bar{p}$ with a statistical

significance of 3.1σ, including systematic uncertainties. The excess can be interpreted as a pentaquark state candidate with quark content $c\bar{c}uud$. Its mass, width and fit fraction are measured to be:

$$\begin{aligned} \mathrm{M}_{P_c} &= 4337^{+7}_{-4}(\mathrm{stat})^{+3}_{-3}(\mathrm{syst})\,\mathrm{MeV}, \\ \Gamma_{P_c} &= 29^{+26}_{-12}(\mathrm{stat})^{+14}_{-13}(\mathrm{syst})\,\mathrm{MeV}, \\ f(P_c) &= 14.4^{+8.0}_{-4.0}(\mathrm{stat}) \pm 8.2(\mathrm{syst})\%, \end{aligned}$$

where the first uncertainty is statistical and the second one is systematic. The preferred spin hypothesis for $J/\psi p$ corresponds to quantum numbers equal to $J^P = 1/2^+$, even though other J^P assignments cannot be excluded. Concerning the theoretical interpretation, the P_c state occurs around 10 MeV below the threshold of an hidden-charm molecule composed by $\chi_{c0}(1P)p$, which has preferred spin-parity value equal to $1/2^+$. This fact could favour the molecular interpretation for this state. However, given the limited statistics up to date, it is not possible to rule out other hypotheses. In order to confirm the nature of such excess we need to wait for more data to be collected by the experiment. Finally, we find no evidence for glueball states decaying to $p\overline{p}$, which were predicted by Ref. [2] to have mass around 2.23 GeV.

During the period of the PhD, I have also been involved in hardware activities related to the upgrade of the LHCb detector to be installed in 2021. My main contribution was to the test of the front-end electronics of the Upstream Tracker (UT) silicon strip detector, in particular to the SALT ASIC chip designed to operate at the pp collision frequency of 40 MHz. We have found a suboptimal performance in terms of signal-to-noise ratio, which led to redesigning the chip to meet the requirements.

Thesis overview

The thesis is organized as follows. Chapter 1 presents an introduction to the Standard Model (SM) of particle physics, focusing on the state of the art of exotic heavy-hadron spectroscopy and B-meson decays.

Chapter 2 reports an introduction to the Large Hadron Collider (LHC) accelerator at CERN and the LHCb experiment, which collected the data that have been analyzed in this thesis.

Chapter 3 discusses a theoretical overview of helicity formalism developed to describe the relativistic processes of spin particles. We used this formalism to build the amplitude model of the $B_s^0 \to J/\psi p\bar{p}$ decays. Critical aspects in the definition of the amplitudes have been found during the work and are here carefully detailed.

Chapter 4 and Chapter 5 describe the main topics of this thesis, illustrating the $B_{(s)}^0 \to J/\psi p\bar{p}$ branching ratio measurements and the amplitude analysis of the B_s^0 decays.

Finally, Chapter 6 details the front-end electronics tests for the Upstream Tracker detector.

CHAPTER 1

Flavour physics in the Standard Model

The Standard Model of particle physics (SM) is the most reliable theory that describes strong, weak and electromagnetic interactions to date. It is a Lorentz-invariant quantum field theory based on the gauge group, $SU(3) \otimes SU(2)_L \otimes U(1)_Y$ [3, 4, 5]. The $SU(3)$ group describes the theory of strong interactions (QCD), acting on a three-dimensional space of color charges, whose generators are eight fields of gluons. Quarks are the fundamental representations of the $SU(3)$ group and combine, together with gluons, in color-singlet states to form hadrons, either in $q\bar{q}$ pairs to form mesons or in qqq triplets to form baryons.

The invariance of $SU(2)_L \otimes U(1)_Y$ gauge transformation describes the electroweak interactions. The $SU(2)_L$ group, defined on the two-dimensional space of weak isospin, transforms only the left-handed chirality states. They are defined in doublets of weak isospin. The leptons are paired with their associated neutrinos and the up-type quarks with the down-type ones:

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix}, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}; \quad \begin{pmatrix} u \\ d' \end{pmatrix}, \begin{pmatrix} c \\ s' \end{pmatrix}, \begin{pmatrix} t \\ b' \end{pmatrix}, \tag{1.1}$$

where the weak down-type eigenstates, $d'_i = (d', s', b')$, are a combination of the mass eigenstates, $d_i = (d, s, b)$, and are defined as $d'_i = \sum_j V_{ij} d_j$, where V is the Cabibbo-Kobayashi-Maskawa (CKM) mixing matrix, given below in Eq. 1.6. Finally, the $U(1)$ group acts on the one-dimensional space of the hypercharge Y. The generators of the electroweak interaction are 4 vector fields, associated to the mediators of the interactions: the photon γ for the electromagnetic and $W^\pm$ and Z^0 for the weak interaction.

The SM Lagrangian describing the interactions with the bosonic fields and their dynamics is the following:

$$\mathcal{L} = \sum_{j=1,\dots,3} \sum_{\psi=Q^j,E^j} \bar{\psi} i \not{D} \psi - \frac{1}{4} \sum_{a=1,\dots,8} G^a_{\mu\nu} G^{a\mu\nu} - \frac{1}{4} \sum_{a=1,\dots,3} W^a_{\mu\nu} W^{a\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} \tag{1.2}$$

where j runs over the three quark and lepton families, ψ are the quark and lepton singlets and doublets, $\not{D}$ is the covariant derivative, $G_{\mu\nu}$ is the field strength tensor for the gluon field (with a running over all eight gluons), and $W_{\mu\nu}$ and $B_{\mu\nu}$ are the electroweak fields. The first part of the Lagrangian describes the interaction between the fermionic matter and the bosonic fields. While, the second part corresponds to the field strength of the bosons. It describes the propagation of the gauge bosons, the self interaction between gluons and the interaction among the electroweak bosons themselves.

A complex scalar Higgs doublet [6], ϕ, is added to the model for mass generation through spontaneous symmetry breaking with potential given by,

$$V(\phi) = \mu^2\phi^\dagger\phi + \frac{\lambda^2}{2}\left(\phi^\dagger\phi\right)^2, \quad \phi \equiv \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \tag{1.3}$$

where, for negative μ^2, ϕ assumes the vacuum expectation value, $v/\sqrt{2} =\mid \mu \mid /\lambda$, where $v = 246.22\,\text{GeV}$.

1.1 Quark mixing and CKM matrix

In the quark sector, it is known experimentally that the quark mass eigenstates are not the same as the quark weak eigenstates. The masses and mixing of quarks arise from the following Lagrangian, which describes the Yukawa interactions with the Higgs boson,

$$\mathcal{L}_Y = -Y^d_{ij}\overline{Q^I_{Li}}\phi d^I_{Rj} - Y^u_{ij}\overline{Q^I_{Li}}\epsilon\phi^* u^I_{Rj} + h.c., \tag{1.4}$$

where $Y^{u,d}$ are 3×3 complex matrices, ϕ is the Higgs field, i, j are generation labels, and ϵ is the 2×2 antisymmetric tensor. Q^I_L are left handed quark doublets, and d^I_R and u^I_R are right handed down- and up-type quark singlets, respectively, in the weak-eigenstate basis. The physical states are then obtained by diagonalizing the Y^{ud} matrix using four unitary matrixes, $V^{u,d}_{L,R}$. Indeed, the mixing between the down-type part of the weak doublets (d', s', b') and the mass eigenstates (d, s, b) is expressed by a unitary matrix, called CKM matrix and defined as $V_{CKM} = V^u_L V^{d\dagger}_L$. The quark mixing can then be parametrized by three real parameters and a complex phase, which accounts for CP violation in the SM. The Wolfenstein approximation of CKM is used, up to the third order of the hierarchical parameter λ, [7]:

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{1.5}$$

$$= \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1/2 \end{pmatrix} \tag{1.6}$$

where A, λ, ρ and η are universal constants ($\lambda = 0.226$ and $A \sim 0.8$). The determination of ρ and η is constraint by experimental results and by the combined analysis of the unitary triangle [8]. Decays usually proceed within generation, so the up-type quarks (u, c, t) decays predominantly in their down-type quarks (d, s, b), *i.e.* $c \to W^+s$, although also decays between different generations are possible, $c \to W^+d$. The CKM matrix rules the favored (diagonal matrix elements) and suppressed (out-of-diagonal) weak interaction processes. With the inclusion of the third quark generation (t, b), an irreducible phase, δ, which is defined as $\tan^{-1}\left(\frac{\eta}{\rho}\right)$, enters in the V_{td} and V_{ub} matrix elements, which are suppressed of a factor λ^3, and accommodates for a small amount of CP violation in the SM.

1.2 Flavour oscillation and CP violation

Flavour oscillation occurs when the flavour eigenstates are different from the Hamiltonian eigenstates with well-defined masses and widths. The CP transformation is the

composition of charge conjugation C and parity inversion P [9]. The first symmetry reverts particle with antiparticles, inverting some internal quantum numbers, *i.e.* the electromagnetic charge $Q \to -Q$. The second one reverts the spatial components of vectors, *i.e.* $x \to -x$, where x being a space vector. Hence, under CP, for example, a left-handed electron e^-_L becomes a right-handed positron e^+_R. The aim of this section is not to give an exhaustive overview of CP measurements in the B sector but a small introduction on CP and mixing formalism applied to neutral B meson decays. Therefore, it has the sole purpose of defining CP formalism useful to answer the questions raised during the work presented in this thesis, in particular relative to Sec. 5.1.6.

1.2.1 Amplitude definition

Considering an initial flavour state of B meson (*i.e.*B^0 or B^0_s), denoted as $|B^0_q\rangle$, where $q = d, s$, and its CP conjugate $\bar{B}$, decaying to a multi-particle final state f and its CP conjugate $\bar{f}$, the decay amplitudes are defined as:

$$\begin{aligned} A_f &= \langle f|\mathcal{H}|B^0_q\rangle, \qquad \bar{A}_f = \langle f|\mathcal{H}|\bar{B}^0_q\rangle \\ A_{\bar{f}} &= \langle \bar{f}|\mathcal{H}|B^0_q\rangle, \qquad \bar{A}_{\bar{f}} = \langle \bar{f}|\mathcal{H}|\bar{B}^0_q\rangle \end{aligned} \tag{1.7}$$

where $\mathcal{H}$ is the weak interaction Hamiltonian. Applying CP on the initial and final states, one obtains:

$$\begin{aligned} CP|B^0_q\rangle &= e^{+i\xi_B}|\bar{B}^0_q\rangle, \qquad CP|f\rangle = e^{+i\xi_f}|\bar{f}\rangle \\ CP|\bar{B}^0_q\rangle &= e^{-i\xi_B}|B^0_q\rangle, \qquad CP|\bar{f}\rangle = e^{-i\xi_f}|f\rangle, \end{aligned} \tag{1.8}$$

where ξ_B and ξ_f are phases depending on the flavour content of the states. If CP is conserved in the decay, the amplitudes A_f and $\bar{A}_{\bar{f}}$ are identical, except for an arbitrary unphysical relative phase:

$$\bar{A}_{\bar{f}} = e^{i(\xi_f - \xi_B)} A_f. \tag{1.9}$$

This relationship will be exploited in the definition of the amplitude model for the untagged $B^0_s \to J/\psi\, p\bar{p}$ decays (Sec. 5.1.6), where CP is assumed to be a good symmetry.

1.2.2 Neutral B meson mixing

For neutral B meson, a CP eigenstate can be constructed as a combination of B^0 and $\overline{B}^0$:

$$|B_{\mathrm{L,H}}\rangle = p\left|B^0_q\right\rangle \pm q\left|\bar{B}^0_q\right\rangle \tag{1.10}$$

where $|B_{\mathrm{L,H}}\rangle$ are the light (L) and heavy (H) eigenstates of this system.

Since B^0 and $\overline{B}^0$ can oscillate one into the other, the evolution of the state $|B(t)\rangle$ in time is expressed through the Hamiltonian $M - i\Gamma/2$ [9] as:

$$i\frac{d}{dt}\begin{pmatrix} |B^0_q(t)\rangle \\ |\bar{B}^0_q(t)\rangle \end{pmatrix} = \left(M - i\frac{\Gamma}{2}\right)\begin{pmatrix} |B^0_q(t)\rangle \\ |\bar{B}^0_q(t)\rangle \end{pmatrix}, \tag{1.11}$$

where the off-diagonal elements $M_{12} - i\Gamma_{12}$ represent the off-shell (M_{12}) and on-shell (Γ_{12}) mixing between the two flavours.

The evolution of the flavour state $|B_q^0\rangle$ or $|B_q^0\rangle$ at time t can then be expressed as

$$\left|B_q^0(t)\right\rangle = g_+(t)\left|B_q^0\right\rangle + \frac{q}{p}g_-(t)\left|\bar{B}_q^0\right\rangle, \tag{1.12}$$

$$\left|\bar{B}^0(t)\right\rangle = g_+(t)\left|\bar{B}_q^0\right\rangle + \frac{p}{q}g_-(t)\left|B_q^0\right\rangle \tag{1.13}$$

where the $g_+(t)$ represents the flavour-conserving evolution, while $g_-(t)$ the flavour-changing evolution and their expression is the following:

$$|g_\pm(t)|^2 = \frac{e^{-\Gamma_q t}}{2}\left[\cosh\left(\frac{\Delta\Gamma_q}{2}t\right) \pm \cos\left(\Delta m_q t\right)\right] \tag{1.14}$$

Δm_q and $\Delta\Gamma_q$ are the difference between the real and immaginary part of the CP eigenvalues: $\Delta\Gamma_q = \Gamma_H - \Gamma_L$, $\Delta m_q = m_H - m_L$, obtained by diagonalizing Eq. 1.11.

At this point we can evaluate the decay rate in time $\frac{d\Gamma}{dt}(B_q^0(t) \to f) = |\langle f|B^0(t)\rangle|^2$ and substituting Eq. 1.12, 1.13 and 1.14, we obtain:

$$\begin{aligned} d\Gamma/dt(B_q^0(t) \to f) = e^{-\Gamma t}N_f\Bigg[& \left(|A_f|^2 + \left|(q/p)\bar{A}_f\right|^2\right)\cosh(y\Gamma t) \\ & + \left(|A_f|^2 - \left|(q/p)\bar{A}_f\right|^2\right)\cos(x\Gamma t) \\ & + 2\mathcal{R}e\left((q/p)A_f^*\bar{A}_f\right)\sinh(y\Gamma t) - 2\mathcal{I}m\left((q/p)A_f^*\bar{A}_f\right)\sin(x\Gamma t)\Bigg] \end{aligned} \tag{1.15}$$

where $y = \Delta\Gamma/2\Gamma$, $x = \Delta m/\Gamma$ and N_f is a normalization factor. This expression can be written in terms of a key quantity for CP violation, λ_f:

$$\lambda_f = \frac{q}{p}\frac{\bar{A}_f}{A_f} \tag{1.16}$$

and Eq. 1.15 simplifies to:

$$d\Gamma/dt(B_q^0(t) \to f) = e^{-\Gamma t}N_f\,|A_f|^2\Bigg[\left(1 + |\lambda_f|^2\right)\cosh(y\Gamma t) + \left(1 - |\lambda_f|^2\right)\cos(x\Gamma t) \tag{1.17}$$

$$+ 2\mathcal{R}e\left(\lambda_f\right)\sinh(y\Gamma t) - 2\mathcal{I}m\left(\lambda_f\right)\sin(x\Gamma t)\Bigg] \tag{1.18}$$

Terms proportional to 1 are associated with decays that occur without any net $B_q^0 \to \bar{B}_q^0$ oscillation, while terms proportional to $|\lambda_f|^2$ are associated with decays following a net oscillation. The $\sinh(y\Gamma t)$ and $\sin(x\Gamma t)$ terms are associated with the interference between these two cases.

The same decay rate can be derived for decay of $\bar{B}_q^0$ to the same final state f, obtaining:

$$d\Gamma/dt(\bar{B}_q^0 \to f) = e^{-\Gamma t}N_f(1+a)|\bar{A}_f|^2\Bigg[\left(\frac{1}{|\lambda_f|^2} + 1\right)\cosh(y\Gamma t) - \left(\frac{1}{|\lambda_f|^2} - 1\right)\cos(x\Gamma t) \tag{1.19}$$

$$+ 2\frac{\mathcal{R}e\left(\lambda_f\right)}{|\lambda_f|^2}\sinh(y\Gamma t) + 2\frac{\mathcal{I}m\left(\lambda_f\right)}{|\lambda_f|^2}\sin(x\Gamma t)\Bigg] \tag{1.20}$$

Then, extracting $|\lambda_f|^2$ in Eq. 1.18 and $1+|\lambda_f|^2$ in both eq. 1.18 and 1.20, the expressions for the decay rates become:

$$d\Gamma/dt(B_q^0 \to f) = e^{-\Gamma t} N_f \left|A_f\right|^2 \frac{1+\left|\lambda_f\right|^2}{2} \tag{1.21}$$
$$\left[\cosh\frac{\Delta\Gamma t}{2} + \mathcal{A}_{\rm CP}^{\rm dir}\cos(\Delta m t) + \mathcal{A}_{\Delta\Gamma}\sinh\frac{\Delta\Gamma t}{2} + \mathcal{A}_{\rm CP}^{\rm mix}\sin(\Delta m t)\right]$$

$$d\Gamma/dt(\bar{B}_q^0 \to f) = e^{-\Gamma t} N_f \left|A_f\right|^2 \frac{1+\left|\lambda_f\right|^2}{2}(1+a) \tag{1.22}$$
$$\left[\cosh\frac{\Delta\Gamma t}{2} - \mathcal{A}_{\rm CP}^{\rm dir}\cos(\Delta m t) + \mathcal{A}_{\Delta\Gamma}\sinh\frac{\Delta\Gamma t}{2} - \mathcal{A}_{\rm CP}^{\rm mix}\sin(\Delta m t)\right]$$

where the factors are defined as:

$$\mathcal{A}_{\rm CP}^{\rm dir} = \frac{1-\left|\lambda_f\right|^2}{1+\left|\lambda_f\right|^2} \qquad \mathcal{A}_{\rm CP}^{\rm mix} = -\frac{2\,{\rm Im}\,\lambda_f}{1+\left|\lambda_f\right|^2} \qquad \mathcal{A}_{\Delta\Gamma} = -\frac{2\,{\rm Re}\,\lambda_f}{1+\left|\lambda_f\right|^2} \tag{1.23}$$

which represent the sources of CP violation if f is a CP eigenstate, $CP|f\rangle = \pm|f\rangle$: in direct decay, $\mathcal{A}_{CP}^{dir}$, in the mixing-induced CP violation and in the interference between $B \to f$ and $\overline{B} \to f$, $\mathcal{A}_{CP}^{mix}$. The third term becomes significant only for a sizeable value of $\Delta\Gamma$, $\mathcal{A}_{\Delta\Gamma}$. Then, the factor $a = |\frac{\Gamma_{12}}{M12}|\sin\phi|$ is less than 0.01 and can be neglected.

For the $B_s^0 \to J/\psi p\overline{p}$ decays, which are of interest for this thesis, CP can be assumed to be a good symmetry; no CP violation can occur neither in direct decays nor indirectly, through mixing and interference. Indeed, the former is only present when the CKM element, V_{ub} or V_{td}, contributes the process. While, the latter cannot be seen for construction in the case of untagged decay rate integrated in time, as discussed further below, in Sec. 1.2.3. Therefore, the parameter λ_f is equal to 1, which implies the following relation between the CP-conjugated amplitudes: $|\bar{A}_{\bar{f}}| = |A_f|$, where f is the final state.

1.2.3 Untagged rate

When B_s^0's and $\bar{B_s^0}$'s are produced in equal numbers, the untagged decay rate integrated in time for the decay $B_s^0 \to f$ reads [10]

$$\begin{aligned}
\Gamma_{tot} &= \int dt \frac{1}{2}\left(\frac{d\Gamma}{dt}(B_q^0 \to f) + \frac{d\Gamma}{dt}(\bar{B}_q^0 \to f)\right) \\
&= \int dt \mathcal{N}_f \left[e^{-\Gamma_L t}\left|\langle f|B_L\rangle\right|^2 + e^{-\Gamma_H t}\left|\langle f|B_H\rangle\right|^2\right] + \mathcal{O}(a) \\
&= 2N_f\left|A_f\right|^2(1+\left|\lambda_f\right|^2)\int dt\cdot e^{-\Gamma t}(\cosh(y\Gamma t) + \mathcal{A}_{\Delta\Gamma}\sinh(y\Gamma t)) \\
&= \frac{\mathcal{N}_f}{2}\left|A_f\right|^2\left[1+\left|\lambda_f\right|^2\right]\frac{\Gamma + \mathcal{A}_{\Delta\Gamma}\Delta\Gamma/2}{\Gamma^2 - (\Delta\Gamma/2)^2} + \mathcal{O}(a)
\end{aligned} \tag{1.24}$$

where we have omitted terms of order a. This equation probes that the dependence to the interference between the direct decay $B_s \to f$ and the decay through mixing $B_s \to \bar{B_s} \to f$ (term proportional to $\sinh(y\Gamma t)$) is removed after integrating in time. What remains is only a constant term proportional to the width Γ, the difference between the width of the light and heavy mass eigenstate and λ_f. This argument will be particularly relevant when considering the decay rate of untagged B_s^0 sample, as discussed in Chapter 5, Sec. 5.1.6.

1.3 Quark model and spectroscopy of heavy hadrons

The first model describing subatomic hadrons was made by Gell-Mann [11] and Zweig [12] in 1964 when they stated that hadrons could be described as composites of fractionally charged fermions with baryon number $B = 1/3$, called "quarks". These can combine together in the most simple configuration to form mesons ($B = 0$) and baryons ($B = 1$), which are composite states of $q\bar{q}$ and qqq, respectively. Nevertheless, both authors claimed that more complex structures with integer charges and $B = 0$ or $B = 1$ could exist, such as $qq\bar{q}\bar{q}$ "tetraquark" mesons and $qqq\bar{q}q$ "pentaquark" baryons. However, for long time, no evidence for such candidates was observed.

Quantum Chromodynamics (QCD) (1973) is the generalization of the idea proposed by Gell-Mann to a gauge theory with quarks of fractional electric charge. Despite QED, in which the photons are electrically neutral and do not interact with each other, the gluons of QCD have color charges and, thus, interact with each other. This interaction causes virtual gluon loop corrections in the vacuum polarization diagrams that makes the behaviour of the QCD coupling strenght α_s opposite to that of its QED counterpart [13]. Indeed, α_s decreases at short distances and increases at long distances as illustrated in Fig. 1.1. At short distances, due to a small value of α_s ($\alpha_s = 0.1185 \pm 0.0006$ at $r = 0.002$ fm), the regime of "asymptotic freedom" allows the use of perturbation expansions to make reliable first-principle calculations for high-momentum-transfer processes. In contrast, for distance scales of the characteristic size of hadrons, $r \sim 1$ fm, the coupling constant is of order 1, $\alpha_s \sim O(1)$, which causes divergent perturbation expansions. This increase in the coupling strength for large quark separations is the source of "confinement", which is the reason why isolated colored particles, be the quarks or gluons, are never observed alone. The only strongly interacting particles that can exist in isolation are color-charge-neutral hadrons.

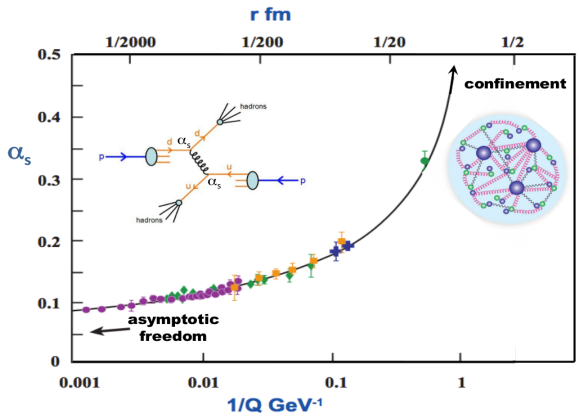


Figure 1.1 The behavior of the QCD coupling strength α_S as a function of the inverse momentum transfer 1/Q or, equivalently, the quark separation distance r. From Ref. [14].

Since QCD works at distance scales that are characteristic of observable hadrons, there are no rigorous, first-principle methods to derive the spectrum and properties of the hadrons from the QCD Lagrangian, except for lattice QCD simulations which are not able to describe all aspects of hadronic states but are developing fast in recent years,

as shortly discussed in Sec. 1.3.3. Instead, a variety of "QCD inspired" phenomenological models have been proposed. All these models have in common the prediction for the existence of hadrons with complex substructures, called "exotic states". They can be included into three categories: multiquark hadrons (tetraquark mesons and pentaquark baryons) formed from tightly bound colored diquarks; hybrid mesons and baryons comprised of color-singlet combinations of quarks and one or more "valence" gluons; and glueball mesons that are comprised only of gluons (with no quarks). The search for non-standard hadrons can be referred to as exotic spectroscopy.

Searches for exotic states were initially conducted in the light-quark sector, *i.e.* u, d, s quarks, but the results were controversial. The primary reason was that the excitation spectrum of light quarks is more complex than in the heavy-quark sector (see sec. 1.3.1), due to the need of a relativistic treatment in the first case. In addition, the spectrum presents too many broad overlapping states, not yet observed, that makes more ambiguous the identification of non-standard hadrons than in the heavy-quark case. Among the candidates of non-$q\bar{q}$ states, we enumerate the scalar mesons $f_0(500)$, $f_0(980)$, $K^{(600)}$ [15] and the hybrid states $\pi_1(1600)$, $\pi_2(1880)$ [16].

In contrast to experiments in the light-quark sector, recent searches for non-standard hadrons containing heavy quark pairs, *i.e.* hadrons that contain a $c\bar{c}$ quark pair, have uncovered a number of intriguing states. In the following an introduction on exotic charm spectroscopy will be presented.

1.3.1 Exotic charm spectroscopy

A phenomenological model that well describes the excitation spectrum of a charm-anticharm state is the semi-relativistic Cornell potential [17]. It is composed by three terms:

$$V(r) = -\frac{4}{3}\frac{\alpha_s(r)}{r} + \sigma r + \delta(1/r^2), \tag{1.25}$$

where the first term corresponds to the short-distance QCD potential, the second one corresponds to the long-distance confinement and the third one includes spin-spin and spin-orbit corrections. This model is found to be precise for the description of high-mass $q\bar{q}$ excited bound states, either $c\bar{c}$ or $b\bar{b}$ spectrum. Indeed, it also predicts states before their observation. However, in the last decades a lot of states, which contain a $c\bar{c}$ quark pair, have been observed that cannot fit in the conventional excitation spectrum. They are composed by a $c\bar{c}$ pair because they decay to charmonia, but have different properties with respect to the predicted ones, such as mass, decay or production rate and quantum numbers. These states are indicated as exotic candidates. Fig. 1.2 shows the $c\bar{c}$ spectrum filled with the exotic states observed before 2017.

1.3.2 Multiquark candidates

The first exotic state in the charmonium sector ever discovered was the neutral $X(3872)$, from the Belle experiment in 2003 [18], decaying to $J/\psi\rho^0$ and $J/\psi\omega$. Its mass, around $3872\,\mathrm{MeV}/c^2$, cannot be associated with any expected $c\bar{c}$ levels and its width does not match predictions [19]. In particular, the width, limited by the experimental resolution of $1.2\,\mathrm{MeV}/c^2$, is extremely narrow for being a state above the open charm threshold. For these reasons it has been interpreted as an exotic state, just below the $D^0\overline{D}^{*0}$ threshold.

Other interesting discoveries include the $Z(4430)^{\pm}$, observed in the decay of $B_s^0 \to \psi(2S)K^+\pi^-$, which is electrically charged. It is the first evidence for a four-quark meson (tetraquark) that decays to $\psi(2S)\pi^-$ [20], which must have minimal quark content of

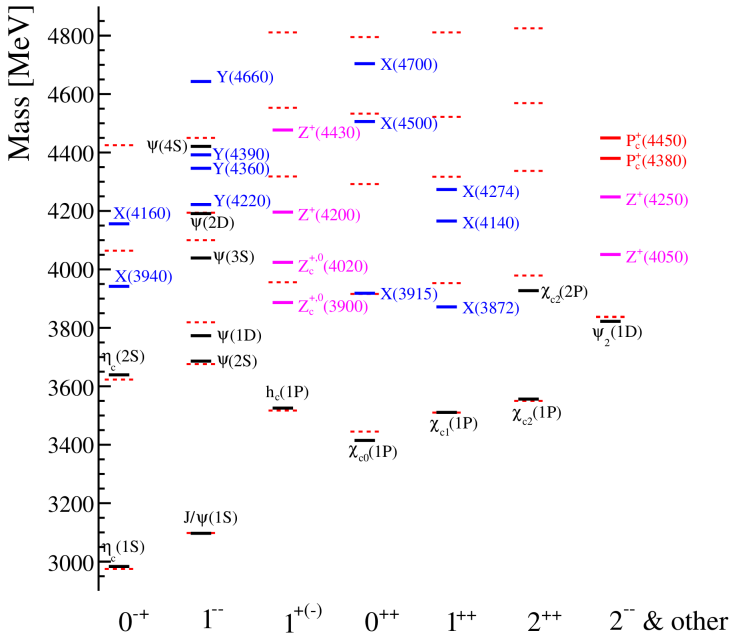


Figure 1.2 The charmonium spectrum (2017): conventional $c\bar{c}$ states in black and predictions in dotted red, neutral (charged) exotics in blue (magenta), pentaquarks in red. Adapted from Ref. [14].

$c\bar{c}d\bar{u}$. Then, pentaquark states, decaying to $J/\psi p$ with a quark content of $uudc\bar{c}$, have been observed by LHCb [21].

In addition to these, about twenty other candidate non-standard hadron states containing $c\bar{c}$ quarks have been found and studied by the BESIII experiment at the BEPCII τ-charm factory in Beijing, the Belle and BaBar experiments at the KEKB and PEPII B factories, the CDF and D0 experiments at the Tevatron, and the LHCb, ATLAS and CMS experiments at the LHC.

1.3.3 Theoretical models

Since there are no rigorous first-principle methods to describe strong interactions at long-scale distances, unless lattice QCD, several phenomenological models have been proposed.

We classify them into two categories according to how quarks bind together inside of non-standard hadrons; they could be either compact multiquark states or hadronic molecules.

A complete description of the models is proposed in Ref. [14], together with an interpretation of the main experimental results in Ref. [22]. Here we want to highlight some experimental signatures that could help distinguish among these models.

In the *compact multiquark* picture, quarks are tightly bound in color-neutral states. Most models predict large widths and isospin multiplets with similar masses. However, they fail to describe narrow decay widths, in contrast to recent experimental discoveries. In addition, no isospin partners have been observed so far.

In the *molecular* picture, the exotic states are loosely bound in mesonic or baryonic molecules. The meson-meson (or meson-baryon) systems are connected with a Yukawa-type meson-exchange force. Since the binding energy is of the order of few MeV, their masses are expected a little below the constituent particles' mass threshold; therefore, no strong decays into its constituents is allowed. The decaying mechanism favorites narrow widths, because the heavy Q and $\overline{Q}$ quarks are typically well separated in space resulting in highly suppressed decay to a hidden-flavour quarkonium state and, thus, leading to long-lived particles. Due to the low binding energy, the few predicted states are preferred to have low orbital momentum (S-wave mode). In addition, spin doublets are expected with masses close one to each other, as for the deuteron binding.

Other models and interpretations for these states are also allowed. Among them, we cite the explanation for which these resonances are not real states but are due to near-threshold kinematic effects. These include threshold "cusps" and anomalous triangle singularities.

Lattice QCD (LQCD) is the most rigorous non-perturbative tool to derive QCD calculations in the low momentum regime and, therefore, extract the properties of hadrons. In LQCD, the space-time is discretised into a lattice, where quark fields placed at lattice sites interact with each other by interconnected gluon links [14]. The resulting equations are then solved numerically thanks to Monte Carlo simulations which are computationally demanding and, therefore, represent one of the major difficulties of these techniques. Thanks to developments in programming techniques [23] and in LQCD algorithms [24], several results have been achieved recently related to hadron masses and excited states [25, 26]. Lattice QCD in the area of nonstandard hadron spectroscopy is still in its infancy [23], but the encouraging results make it possible to become the most powerful theoretical tool for describing the properties of the discovered states in the following years.

1.4 Pentaquark candidates and searches in beauty hadron decays

Pentaquark searches became of great scientific interest in early 2000s, after the observation of an exotic baryon, called Θ^+, compatible with being composed of five quarks $uudd\bar{s}$. This state was firstly observed by the LEPS experiment [27] in photo-production decays ($\gamma n \to nK^+K^-$) as a resonance decaying to $nK^\pm$ final states with mass around 1540 MeV/c^2 and width smaller than 25 MeV. The same state was seen by other nine experiments in similar processes in the following years, including the CLAS experiment at JLAB [28]. However, the measured masses and widths were found to be inconsistent among different experiments, not compatible within the uncertainties. In addition, not all the experiments showed an evidence for it. Indeed, few years after it was disclaimed by the CLAS Collaboration [29] because the excess disappeared with 30 times more statistics than the original dataset.

The first discovery of pentaquark canditates was carried out in July 2015 by the LHCb collaboration in the study of $\Lambda_b \to J/\psi p K^-$ decays [21], analysing data collected during Run 1 and corresponding to 3 fb^{-1}. The pentaquark-like structure with a minimal quark content of $c\bar{c}uud$ is observed decaying into a J/ψ meson and a proton final state. The data contain a narrow peak in the $J/\psi p$ mass distribution that is evident as a distinct horizontal band in the $M^2(J/\psi p)$ *vs.* $M^2(K^-p)$ Dalitz plot shown in Fig. 1.3 and in the $J/\psi p$ invariant masses shown in Fig. 1.5.

In order to clarify the nature of this band, an amplitude six-dimensional analysis was performed. The signal statistics amounts to 26.000 ± 170 signal events with the 5.4%

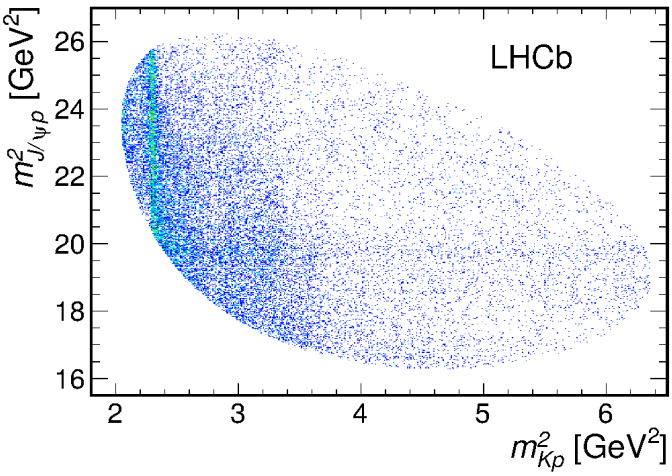


Figure 1.3 Dalitz plot distribution for $\Lambda_b \to J/\psi p K^-$ decays as observed by LHCb [21].

of background. The six-dimensional kinematic space is described by six independent variables, including the $M(Kp)$ invariant mass, the decay helicity angles[1] (θ) of Λ_b, Λ^* and J/ψ, and the angles between the decay planes, ϕ_K and ϕ_μ. The decay angles in the Λ^* chain are defined in Fig. 1.4. Several Λ^* resonances in the K^-p final states are considered in the amplitude model, setting their masses and widths to PDG table [30].

However, these Λ^* contributions taken by themselves fail to describe the data. Two non-standard $P_c^+ \to J/\psi p$ pentaquark contributions were necessary to reasonably describe the narrow structure seen in $M(J/\psi p)$, as illustrated in Fig. 1.5.

The two P_c^+ states were found to have masses of $4380 \pm 8 \pm 29$ MeV/c^2 and $4449.8 \pm 1.7 \pm 2.5$ MeV/c^2, widths of $205 \pm 18 \pm 86$ MeV and $39 \pm 5 \pm 19$ MeV, and quantum numbers

[1]The decay helicity angle is the angle between one of the decay products and the boost direction in the rest frame of the parent particle.

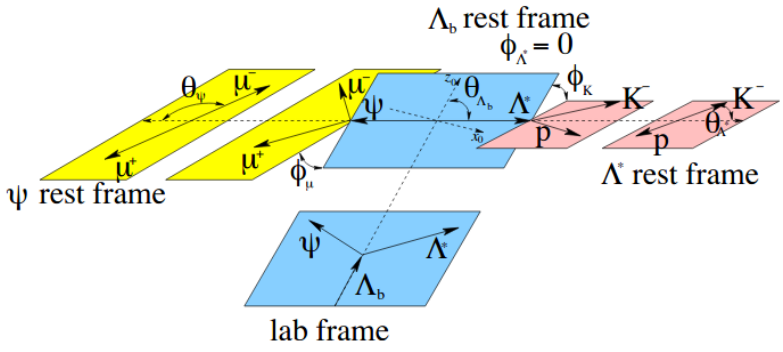


Figure 1.4 Definition of the decay angles in the Λ^* decay chain. From Ref. [21]

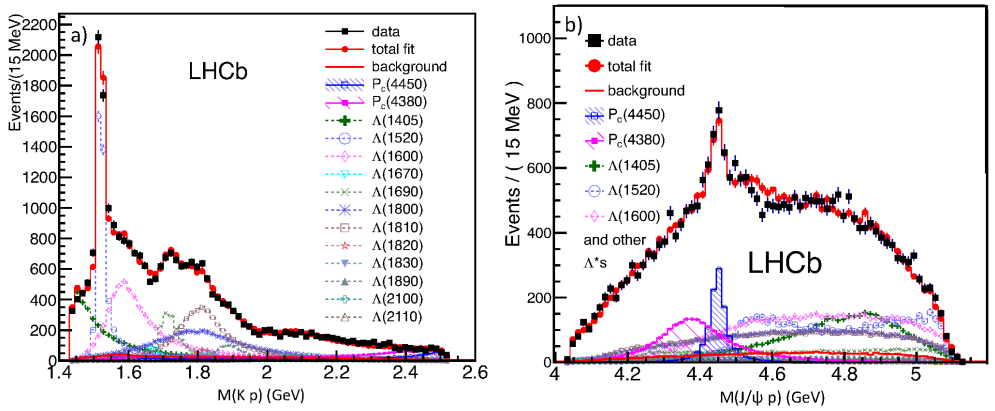


Figure 1.5 Projections of the amplitude fit with $P_c(4380)^+$ and $P_c(4450)^+$ states included (red histogram) onto the a) $M(Kp)$ and b) $M(J/\psi p)$ reconstructed invariant mass distributions (black squares with error bars) for $\Lambda_b \to J/\psi p K^-$ decays by LHCb [21]. Contributions from individual components of the fit are indicated as different color/line-style histograms as labeled.

$J^P = \frac{3}{2}^-$ and $J^P = \frac{5}{2}^+$, respectively. However, other spin hypotheses corresponding to combinations $(3/2^+, 5/2^-)$ and $(5/2^+, 3/2^-)$ cannot be excluded. The statistical significance of the P_c^+ is measured to be 9σ and 12σ for each contribution, respectively.

Since the Λ^* spectroscopy is a complex subject and it is not well-established experimentally, LHCb has investigated the possibility of other resonances in K^-p system exploiting a model-independent approach [31]. The data in the $(m_{K^-p}, \theta_{\Lambda^*})$ distributions are described by an expansion of the θ_{Λ^*} angle in Legendre polynomials:

$$\frac{dN}{d\cos\theta_{\Lambda^*}} = \sum_{l}^{l_{max}} a_l P_l(\cos\theta_{\Lambda^*}), \tag{1.26}$$

where P_l is the Legendre polynomial of order l and a_l are the coefficients, defined as:

$$a_l = \frac{2l+1}{2} \int_{-1}^{+1} P_l(\cos\theta_{\Lambda^*}) \frac{dN}{d\cos\theta_{\Lambda^*}} d\cos\theta_{\Lambda^*} \tag{1.27}$$

and were determined using efficiency-corrected and background-subtracted events by quadratically interpolating the m_{Kp} histogram. The model generated by the K^-p contributions alone is expanded up to rank, l_{max}, which depends on the spin of the intermediate resonance at a certain value of $m(K^-p)$, as summarized in Fig. 1.6. The value of l_{max} in Eq. 1.26 cannot be higher than $2J_{max}$, where J_{max} is the highest spin of any K^-p contributions at a specific value of $m(K^-p)$. In the K^-p model, both resonant Λ^+, Σ^* and non-resonant contributions can be included. As shown in Fig. 1.7, this description is not enough to describe the data, leading to the necessity of additional $J/\psi p$ resonances.

Recently, LHCb released an update of the analysis including the Run 2 data (2015-2018) and corresponding to a total luminosity of $9\,\text{fb}^{-1}$ [32]. Due to an increase in luminosity (three times the Run 1 data), a doubled cross-section in Run 2 and an improved selection, the signal yield increased of a factor nine with respect to the previous analysis. Thanks to the precision gained with significantly higher yield, it was possible to

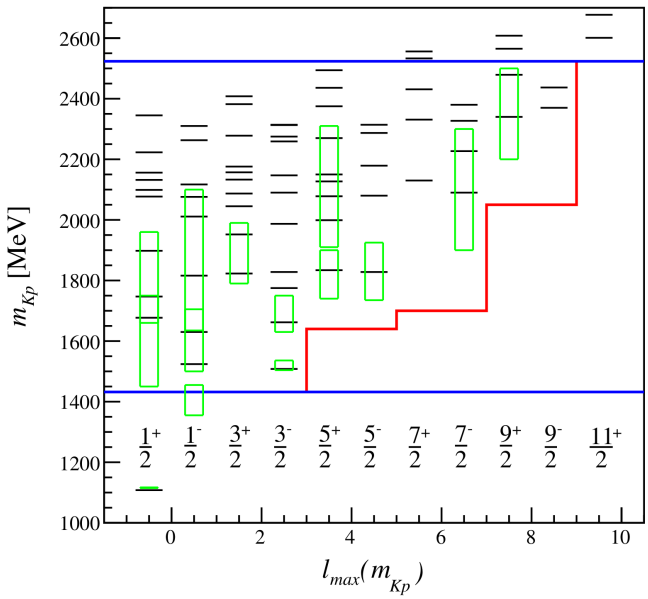


Figure 1.6 Definition of l_{max} as function of the Λ^* resonances available in a certain range of $m(K^-p)$ (predicted resonances are in black and observed ones in green). The blue lines correspond to the mass range, while the red line to the l_{max} values. From Ref. [31].

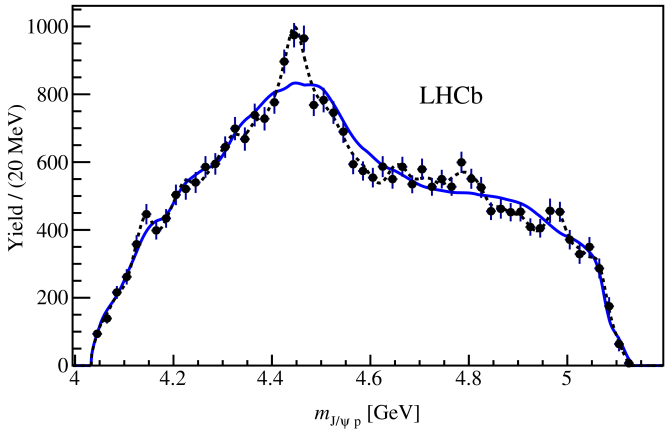


Figure 1.7 Efficiency-corrected and background-subtracted $M(K^-p)$ reconstructed invariant mass distribution, with Legendre expansion (blue). From Ref. [31].

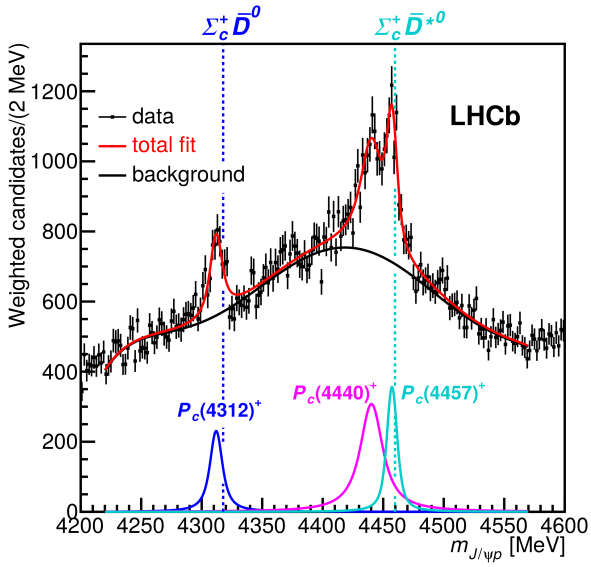


Figure 1.8 One-dimensional fit of $m(J/\psi p)$ invariant mass modelled with three BW functions and a six-dimensional polynomial background. $\Sigma_c^+ \bar{D}^{(*)0}$ mass thresholds are also drawn. From Ref. [32].

resolve the peak at mass $\sim 4450\,\text{MeV}/c^2$ in two separate peaks. A new resonance was also seen, at mass $4312\,\text{MeV}/c^2$. Since all resonances are very narrow, of the order of few MeV, the peaks cannot be due to reflections from K^-p invariant mass. Therefore, before performing a full amplitude analysis, a one-dimensional fit to the $m(J/\psi p)$ is used to determine the masses and widths of the states. The result of the fit is shown in Fig. 1.8, where the model is composed by three Breit-Wigner functions and a six-order polynomial background. The events are also weighted by the inverse of the $\cos\theta_{\Lambda^*}$ distribution, to suppress possible Λ^* resonances which dominates the region of large $\cos\theta_{\Lambda^*}$ values. Looking at Fig. 1.8, the peaks are all few MeV below the mass threshold of $\Sigma_c^+ \bar{D}^{(*)0}$ systems, which could lead to the interpretation of these states as molecules of Σ_c^+ and $\bar{D}^{(*)0}$ particles. However, this puzzle is not resolved yet. Independent confirmation of pentaquark candidates in other decays or production mechanisms is needed.

1.5 B meson decays

B meson decays can occur through various processes, according to the structure of the charged-current interaction. They can be divided into three classes regarding to the final particles: leptonic decays, in which an annihilation of the quarks of the decaying hadron occurs and only leptons appear in the final state; semi-leptonic decays, in which both leptons and hadrons are present in the final state; and non-leptonic decays, in which the final state consists of hadrons only. Representative examples of these three types of decays are shown in Fig. 1.9. On one hand, the first classes of decays are employed for SM tests and measurements of its parameters, as the weak mixing angles of the CKM matrix, and CP violation. On the other hand, hadronic weak decays can provide a unique

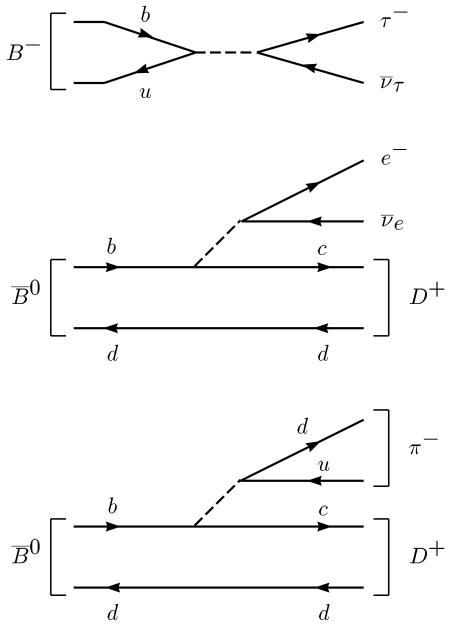


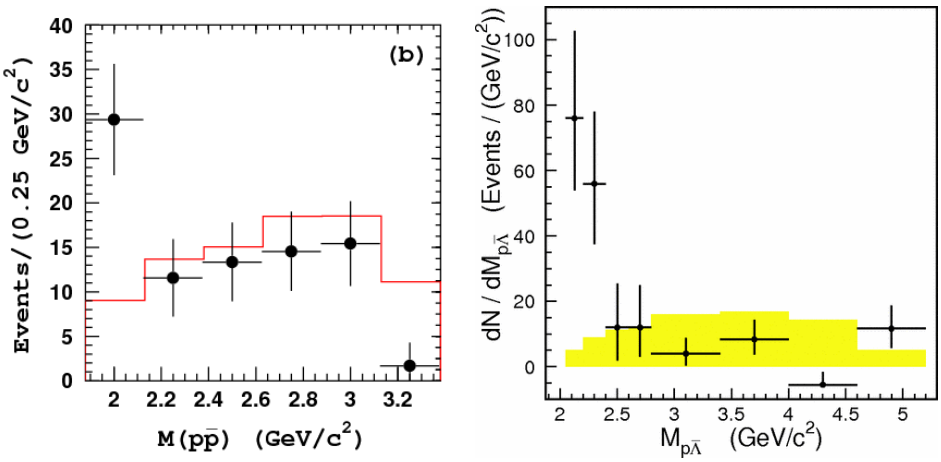
Figure 1.9 Three classes of B decays: (a) leptonic, (b) semi-leptonic and (c) non-leptonic decays.

laboratory for the investigation of heavy baryon spectroscopy, contributing to the understanding strong-interactions. Since the last category is of interest for this thesis, a summary of hadronic B decays is presented in Sec. 1.5.1, followed by an introduction of $B^0_{(s)} \to J/\psi\, p\overline{p}$ decays in Sec. 1.5.2.

1.5.1 Hadronic B decays

Decays of B mesons to final states containing baryons has a long and interesting history, starting with the first observation of the baryonic $B^+ \to \Lambda_c^+ \overline{p}\pi^-$ decay at CLEO [35]. Numerous studies of B^+ and B^0 baryonic decays were performed at the B-factories. The branching fractions for multi-body are typically larger than the two-body final states. It was only with the advent of data from LHCb that the first two-body baryonic mode $B^+ \to p\overline{\Lambda}(1520)$ was observed [36]. More recently, LHCb has also reported the first observations of baryonic decays for the $B_s^0 \to p\Lambda_c^+ K^-$ [37] and $B_c^+ \to J/\psi\, p\overline{p}\pi^+$ [38] mesons.

Baryonic decays of B mesons are interesting for several reasons. First, there seems to be an ubiquitous "threshold enhancement" for dibaryon B decays in several modes. For $p\overline{p}$ pair production, the observed decays $B^0 \to \overline{D}^{(*)0} p\overline{p}$ [33, 39], $B^+ \to K^{(*)+} p\overline{p}$ [40, 41], $B^0 \to K^{(*)0} p\overline{p}$ [42, 43] and $B^+ \to \pi^+ p\overline{p}$ [42, 40] all have an enhancement near the $p\overline{p}$ threshold. A similar threshold effect is seen in $B^0 \to p\bar{\Lambda}\pi^-$ [34]. This indicates that the B meson prefers to decay into a dibaryon system of low invariant mass, with a fast recoiling meson. This can also explain the low branching fractions of two-body dibaryon B decays. It also seems that in the dibaryon rest-frame, the outgoing meson is correlated either the baryon or the anti-baryon [44]. Thus, in $B^+ \to p\overline{p}\{\pi^+, K^+\}$ [40], the proton



(a) (b)

Figure 1.10 The dibaryon threshold effect in: (a) $\overline{B}^0 \to D^0 p\bar{p}$ [33] and (b) $B^0 \to p\bar{\Lambda}\pi^-$ [34]. The histograms indicate expected distributions from phase-space.

helicity angle is forward-peaked for the K^+ and backward-peaked for the π^+. Therefore, a Dalitz analysis of such three-body decays could be very promising.

The same threshold behavior has also been observed in the baryonic J/ψ decays $J/\psi \to \gamma p\bar{p}$ [45] and $J/\psi \to K^- p\bar{\Lambda}$ [46], while no effect has been seen in other decays, as $B^- \to J/\psi\,\Lambda\bar{p}$ [47].

A lot of theoretical ideas have been put forward to explain this effect but none of them is found to be exhaustive. One of the most favorable interpretations comes from the perturbative QCD (pQCD) effect [49]. As illustrated in Fig. 1.11a, one energetic $q\bar{q}$ pair must be emitted back to back by a hard gluon, which hadronizes to produce a baryon and an antibaryon in the two-body decay. Since the hard gluon is highly off mass shell, the two-body decay amplitude is suppressed by order of α_s/q^2, where q is the four-momentum squared transferred by the gluon. This effect is called short-distance picture. While in the 3-body decay processes, as shown in Fig. 1.11b, the $B\bar{B}$ pair, where B being a baryon, is emitted in the opposite direction with respect to the meson. Hence, the gluon which produces the quark and antiquark is close to its mass shell and the short-distance picture does not apply any more. For this reason, $B\bar{B}$ pairs carrying low momentum are more favored. Although this interpretation explains the threshold enhancement in most of the observed decays, it fails to explain the angular correlation seen in the penguin-dominated processes $B^- \to p\bar{p}K^-$ and $B^- \to \Lambda\bar{p}\pi^-$, where the $p(\Lambda)$ is highly correlated with the K^- (π^-) instead of the $\bar{p}$ [40]. Among possible solutions, we enumerate the work presented in Ref. [48], where the inclusion of bound states $N\bar{N}$ may solve this "angular correlation puzzle" and in Ref. [50], where an hint for glueball of mass ~ 2.2 GeV is seen. However, these hypotheses are not yet proven experimentally. New experimental evidences are therefore needed.

1.5.2 The $B_s^0 \to J/\psi\, p\bar{p}$ decays

The present project is focused on the study of neutral $B^0_{(s)}$ mesons decays decaying into a J/ψ meson, a proton p and an antiproton $\bar{p}$.

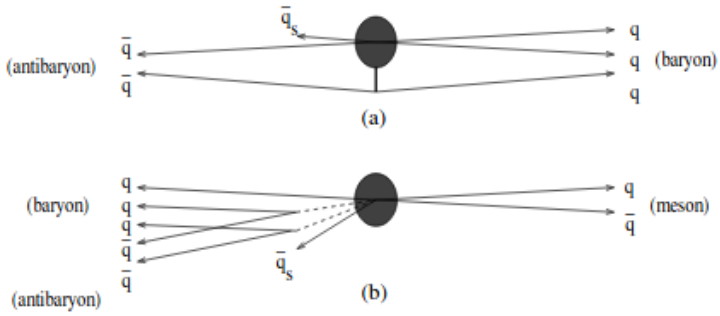


Figure 1.11 Short-distance picture in quarks and antiquarks for (a) two-body baryonic decay and (b) three-body baryonic decay. In the two-body decay (a), the hard gluon is represented by the thick vertical solid line, while in the three-body decay, gluons are represented by dashed lines. $\bar{q_s}$ is the slow spectator anti-quark. From Ref. [48].

These decays belong to the category of hadronic B decays to baryonic final states. They are suppressed compared to the mesonic $B_s^0 \to J/\psi K^+K^-$ or $B^0 \to J/\psi \pi^+\pi^-$ decays due to the much smaller available phase-space. The processes occur via a electroweak interaction where a $\bar{b}$ quark goes into a $\bar{c}$ exchanging a W^+ propagator. The two decays are further reduced due to Cabibbo suppression for the B^0 process and OZI-suppression for the B_s^0 process [2], as described in the Feynman diagrams in Fig. 1.12 and 1.13(a), for the B^0 and B_s^0, respectively.

The OZI suppression was independently proposed by Okubo, Zweig and Iizuka [51, 12, 52] in the 1960s, to explain why certain strong decay processes were suppressed with respect to the expectations. They stated that any strongly occurring process is suppressed if, through only the removal of internal gluon lines, its Feynman diagram can be separated into two disconnected diagrams: one containing all of the initial-state particles and one containing all of the final-state particles. The most notable example where this suppression takes place is the decays of ϕ meson to three pions ($\phi \to \pi^+\pi^-\pi^0$) and the J/ψ decay to hadrons. The former is expected to be favored with respect to the decay into kaons due its larger Q-value, while the rate is measured to be lower. In the latter decay, the suppression explains the increase in lifetime of the J/ψ particle and, as a consequence, its very narrow decay width of just 93.2 ± 2.1 keV. Therefore, the J/ψ electromagnetic decays becomes of the same order than the hadronic ones. The original explanation of the OZI suppression is then supported by the QCD formulation, with the decrease of the QCD coupling constant with increasing energy. In these processes, gluons must carry large transverse momentum to generate pairs of $q\bar{q}$, at least equal to the rest mass of quarks in which they decay into, and since α_s decreases with energy, the processes are suppressed at order $(\alpha_s)^{n_g}$, where n_g is the number of gluons in which a state can annihilate. For the J/ψ decays, it is possible to derive the number of allowed gluons equal to three, as the consequence of the forbidden processes involving one and two gluons. The process through one virtual gluon is forbidden by color and a two-gluon final state is excluded for C conservation. Indeed, the initial state has negative charge conjugation C, while two gluons have positive value of C. While, in the $B_s^0 \to J/\psi p\bar{p}$ processes we could simply exclude the emission of a single gluon but not the one of two gluons, because no C-conjugation relations can be enforced.

Previous attempts have been made to search for $B_{(s)}^0 \to J/\psi p\bar{p}$ decays. Belle reported

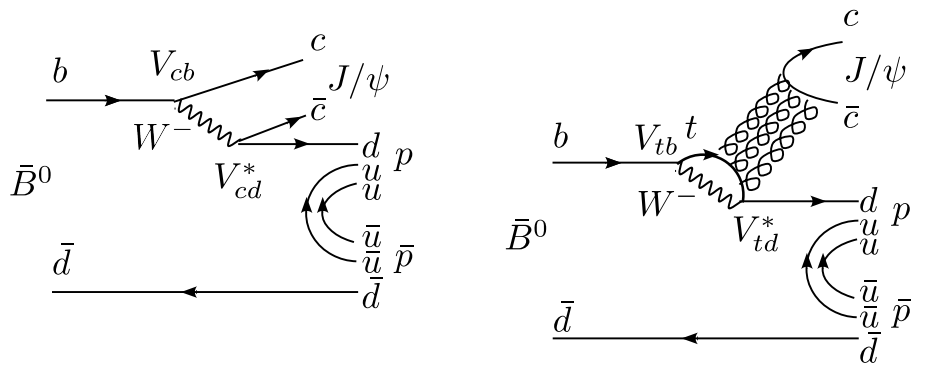


Figure 1.12 Leading diagrams for $B^0 \to J/\psi\, p\bar{p}$ are Cabibbo suppressed: (a) tree and (b) penguin.

a search for $B^0 \to J/\psi\, p\bar{p}$ [47] but no signal has been observed and an upper limit on the branching ratio is imposed: $\mathcal{B}(B^0 \to J/\psi\, p\bar{p}) < 8.3 \cdot 10^{-7}$, at 90% confidence level.

Afterwards, LHCb has published an analysis of $B^0_{(s)} \to J/\psi\, p\bar{p}$ decays with 1 fb^{-1} [53] but also in this case no observation of either modes has been made, with only a $2.8\ \sigma$ evidence seen for the $B^0_s \to J/\psi\, p\bar{p}$ mode. The following upper limits have been placed

$$\begin{aligned} \mathcal{B}(B^0 \to J/\psi\, p\bar{p}) &< 5.2 \times 10^{-7}, \\ \mathcal{B}(B^0_s \to J/\psi\, p\bar{p}) &< 4.8 \times 10^{-6}, \end{aligned} \tag{1.28}$$

at the 90 % confidence level. The mass spectrum is shown in Fig. 1.14 and the measured branching fractions are:

$$\begin{aligned} \mathcal{B}(B^0 \to J/\psi\, p\bar{p}) &= (2.0^{+1.9}_{-1.7}\,(\text{stat}) \pm 0.9\,(\text{syst}) \pm 0.1(\text{norm})) \times 10^{-7}, \\ \mathcal{B}(B^0_s \to J/\psi\, p\bar{p}) &= (3.0^{+1.2}_{-1.1}\,(\text{stat}) \pm 0.6\,(\text{syst}) \pm 0.3(\text{norm})) \times 10^{-6}, \end{aligned} \tag{1.29}$$

where the third uncertainty originates from the control channel ($B^0_s \to J/\psi\, \pi^+\pi^-$) branching fraction measurement.

These decays are very interesting for several reasons. First of all, it could be intriguing to see whether they exhibit the same feature of other hadronic B decays, as the "threshold enhancement" in the di-baryon invariant mass [54], that were mentioned earlier (Sec. 1.5.1). A possible observation could help clarify its origin. In addition, we expect that this effect does not occur for the B^0_s mode because, due to the OZI suppression, the diagram does not correspond to the family of the three-body baryonic decays shown in Fig. 1.11b. As explained above, in the B^0_s process hard gluons generate $q\bar{q}$ pairs which hadronize into a baryon-antibaryon, $B\bar{B}$, while in the other decays slow gluons are responsible of the $B\bar{B}$ production. Therefore, it could be interesting to see if this prediction is satisfied.

Secondly, while for the B^0 decay the limit on the branching ratio is consistent with the expectation from the threshold enhancement that set it to $\sim 10^{-7}$ [54], this seems

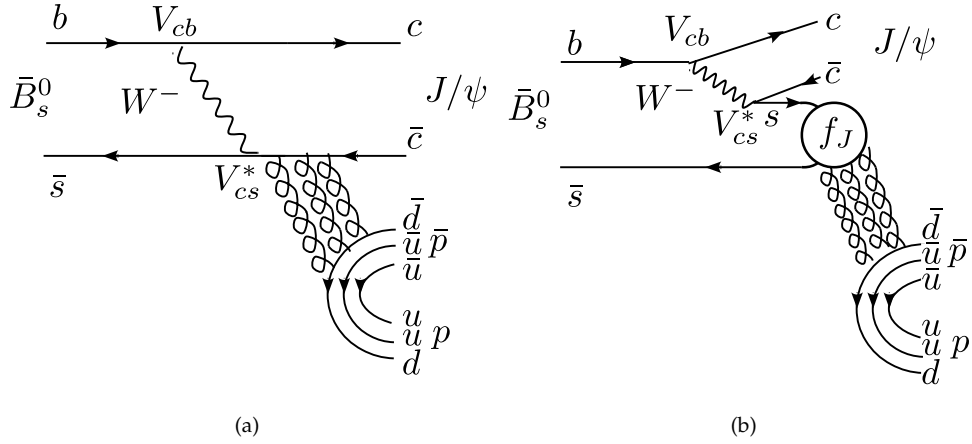


Figure 1.13 Leading diagrams for $B_s^0 \to J/\psi p\bar{p}$ are OZI suppressed: (a) non-resonant process, (b) through production of an intermediate resonance f_J as in Hsiao [2].

not to hold for the B_s^0 decay. Indeed, the LHCb data points to a larger branching fraction for this mode. Assuming that the threshold enhancement cannot occur for the B_s^0 decay, the enhancement in the branching ratio could be explained as due to the formation of an intermediate glueball resonance $f_J(2220)$, predicted by Hsiao [2], that lifts the OZI suppression, as shown in Fig. 1.13b. A glueball is a state composed only by gluons, predicted by the QCD but not confirmed experimentally. In Fig. 1.16, predictions for masses and quantum numbers of these states are extracted from lattice QCD calculations [55]. The suitable glueball candidate for our decay process, which falls in the allowed mass range, is indeed the one at mass 2220 MeV. This is the baryonic equivalent of $B_s^0 \to J/\psi f_0(980)(\to \pi^+\pi^-)$ that has been studied at LHCb [56] earlier. The amplitude distribution with respect to $m_{p\bar{p}}$ is expected to have a narrow and sharp peak corresponding to the value of $m_{p\bar{p}} = 2.23$ GeV, as shown in Fig. 1.15. The glueball is predicted to have quantum numbers of $J^{PC} = 2^{++}$ or 4^{++} and a width less than 30 MeV. A status of the f_J state is reviewed in Ref. [57]. A similar structure has been observed in radiative J/ψ decays from the Mark III [58] and the BES Collaboration [59]. In addition, an hint of it has been seen in charmless B meson decay $B \to Kh^+h^-$, where h being K, π, p [50], even if it is very weak. However, the results were controversial because no evidence for this state was found in other production processes, as $p\bar{p}$ annihilation [60] and 2γ processes [61]. Hence, the $B_s^0 \to J/\psi p\bar{p}$ could be an independent channel for observing this state. If the $f_J(2220)$ contributes to the decay, it can also potentially explain the "threshold enhancement".

Moreover, another interesting prospect for these modes is the possibility of pentaquarks occurring in the $[J/\psi p]$ or $[J/\psi \bar{p}]$ systems. Since the B^0 mode has a quite restricted phase-space, the B_s^0 mode is more favored for observing these exotic structures. LHCb observed [21, 32] at least three resonances decaying to $J/\psi p$ in the $\Lambda_b \to J/\psi pK^-$ decays, as described in Sec. 1.4. The $B_{(s)}^0 \to J/\psi p\bar{p}$ decays are particularly attractive for confirming the ground state pentaquark $P_c^+(4312)$ and the broad $P_c^+(4380)$, which are directly accessible for the B_s^0 mode, as shown in Fig. 1.17, while the $P_c^+(4450)$ can contribute only through its tails.

A full amplitude analysis would be needed to search for pentaquark and glueball

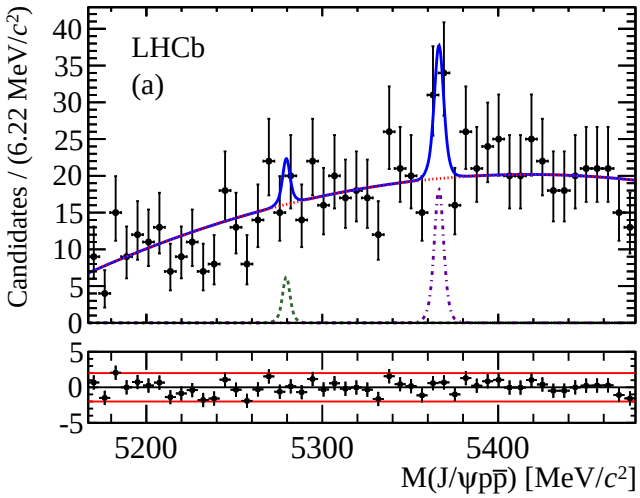


Figure 1.14 The invariant $m(J/\psi p\overline{p})$ mass from the 1 fb^{-1} LHCb analysis [53]. No observation for the $B^0_{(s)} \to J/\psi p\overline{p}$ decays was seen, with a 2.8 σ evidence for the $B^0_s \to J/\psi p\overline{p}$ mode.

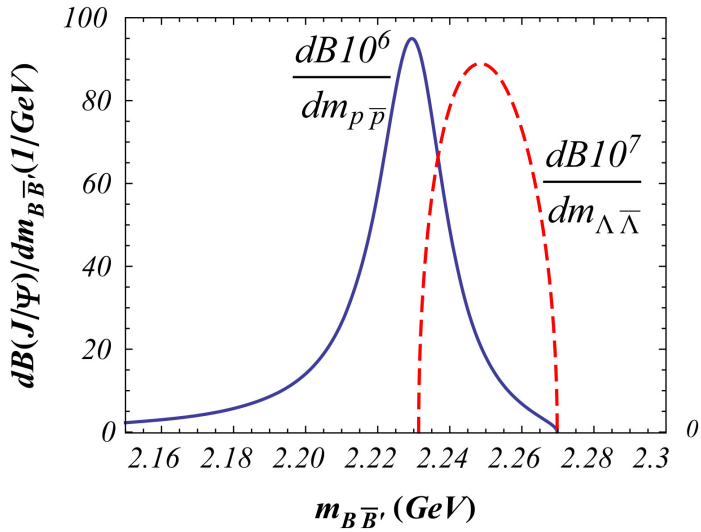


Figure 1.15 The invariant $m_{B\bar{B}}$ for a glueball candidate in blue and for a simple threshold enhancement effect in red. From Ref. [2].

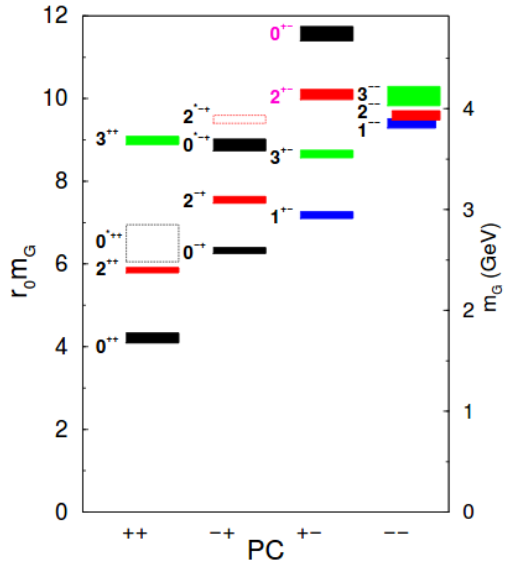


Figure 1.16 The mass spectrum of glueballs in the pure $SU(3)$ gauge theory, taken from Ref. [55]. The masses are given in terms of the hadronic scale r_0 along the left vertical axis and in terms of GeV along the right vertical axis.

existence as will be detailed in the second part of this thesis, Ch. 5.

Finally, the Q-value, or 3-body breakup momentum for the $B_s^0 \to J/\psi\, p\overline{p}$ decay, is only $\sim$ 309 and $\sim$ 393 MeV for B^0 and B_s^0, respectively. This gives a very good resolution for the kinematics and, therefore, it allows to make precise measurements of the $B_{(s)}^0$ masses.

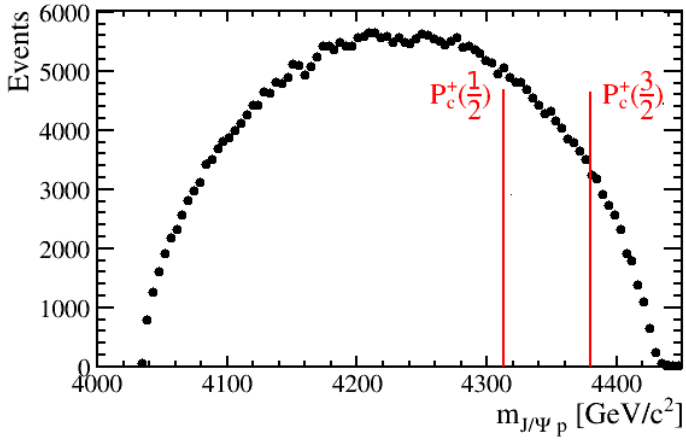


Figure 1.17 The allowed invariant mass of $[J/\psi p]$ for B_s^0 covers the window $[4025, 4450]$ MeV and includes the spin $1/2$ and $3/2$ pentaquark states.

CHAPTER 2

LHCb experiment at LHC

2.1 Large Hadron Collider

The Large Hadron Collider (LHC) particle accelerator [62] is the most powerful accelerator built up to date. It is located in the CERN (Conseil Européen pour la Recherche Nucléaire) laboratory, across the border between Switzerland and France in the Geneva area.

The LHC is a circular collider with a circumference of 26.7 km, built at a depth of around 100 m underground and housed in the former LEP tunnel. The accelerator provided proton-proton collisions at a center of mass energy $\sqrt{s}$ of 7 TeV from 2010 to 2011, of 8 TeV in 2012 and up to 13 TeV in Run 2 (2015-2018). During Run 3, after the ongoing Long Shutdown 2 scheduled to end in 2022, it will reach the designed pp collision energy of 14 TeV.

To achieve the design energy of 7 TeV per beam, the protons are accelerated by Radio Frequency (RF) cavities, which operate at a frequency of 400 MHz and accelerate the protons using a longitudinally oscillating electric field. To achieve a constant acceleration over many turns of the protons, the frequency of the RF must be precisely tuned to the revolution frequency of the LHC. Protons are maintained in circular trajectories thanks to super-conducting dipole magnets, which have a field strength of ~ 8T and are cooled down to 1.9 K due to a superfluid helium-based cooling system. While, quadrupole and octupole magnets are exploited to focus the particle bunches to minimise beam losses.

The LHC is only the last piece of a complicated chain of devices, Fig. 2.1, to produce bunches and accelerate the protons until they finally collide at one of the four collision points. The chain starts with the extraction of protons (50 keV) from hydrogen nuclei. Hydrogen gas is injected into a metal cylinder, which is surrounded by an electric field to separate the gas constituents, electrons and protons. Protons are then guided to a linear accelerator (LINAC), where the energy is increased to 50 MeV and are then injected into a booster synchrotron, reaching an energy of 1.4 GeV, and subsequently into the Proton Synchrotron (PS) with a final energy of 26 GeV. In this machine the protons are also grouped into train of bunches, a structure that is kept until the beams finally collide in the LHC. After leaving the PS, these bunch-trains are injected into the Super Proton Synchrotron (SPS), where they are further accelerated up to 450 GeV. Finally, protons are transferred to the LHC into two opposite directions in separate beam-pipes. After the filling of both counter rotating beams with bunch-trains, the protons are accelerated while the magnetic field, in the bending dipoles, is simultaneously rising until the collision energy is reached. The beams are forced to collide in four interaction points, where the main experiments are located: ATLAS (IP1), ALICE (IP2), CMS (IP5), LHCb (IP8). The collisions occur at a bunch crossing rate of 40 MHz, corresponding to a distance between bunches of 25 ns. The LHC is designed to contain ~ 2808 simultaneous bunches

circulating in the ring, each one filled with up to 1.15×10^{11} protons. The instantaneous luminosity $\mathcal{L}$, which is a key parameter in accelerators, is an estimate of the number of collisions as a function of time and is given by:

$$\mathcal{L} = \frac{f_{BC} N_1 N_2 F}{4\pi\sigma_x\sigma_y} \tag{2.1}$$

where f_{BC} is the bunch crossing frequency, $N_{1,2}$ are the numbers of protons per beam and $\sigma_x\sigma_y$ is the cross-sectional area in the transverse plane, which can be expressed as $\epsilon\beta^*$, where ϵ is the normalised emittance and β^* the betatron parameter. Therefore, the instantaneous luminosity does not depend only on the frequency of interactions and the number of protons inside each beam but varies inversely with the cross-sectional area of the colliding beams. The transverse size of the beam can be reduced by using a lower emittance (ϵ) beam, or by squeezing the beam more with the focusing magnets (reducing β^*). The collisions happen at a crossing-angle between the beams, which is embedded in the geometric factor F. This angle is needed to avoid parasitic collisions due to the short distance between bunches and reduces the luminosity. Another important parameter for LHC is the pileup, which is a measure of the number of inelastic pp interactions that occur per bunch crossing. The relevant parameters are reported in Table 2.1.

So far, LHC reaches an instantaneous peak luminosity of $2 \cdot 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ for proton-proton collisions and $10^{27}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ for heavy-ion collisions (Pb-Pb).

Parameter	Design	Run 1	Run 2
Energy [TeV]	14	7/8	13
Bunch spacing [ns]	25	50	25
Bunch Intensity [10^{11} ppb]	1.15	1.6	1.2
Number of bunches	2800	1400	2500
Emittance [μm]	3.5	2.2	2.2
β^* [cm]	55	80	30→25
Crossing angle [μ rad]	285	-	300→260
Peak Luminosity [$10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$]	1.0	0.8	2.0
Peak pileup	25	45	60

Table 2.1 Summary of the main parameters for the LHC, including the design values and those reached during Run 1 and Run 2.

2.2 LHCb experiment and detector

The LHCb experiment is designed to study heavy flavour physics at the LHC [63]. Its primary goal is to look for new physics by studying CP violation and rare decays of beauty and charm hadrons. The measurement of theoretically clean observables, not compatible with the SM, would represent a clear signature of new physics, also called indirect evidence for new physics to be distinguished from the direct search of new particles not predicted in the SM. However, LHCb also contributed to other studies. In particular, it achieved excellent results in conventional and exotic spectroscopy, with the discovery of new bound states; in searches for possible new light particles, and in studies of heavy-ion processes.

LHCb is built with a design similar to fixed target experiments since pairs of $b\bar{b}$ quark are mostly created with large boost in the forward region, with strong correlations between the two quarks due to the gluon-fusion dominant production process, shown in

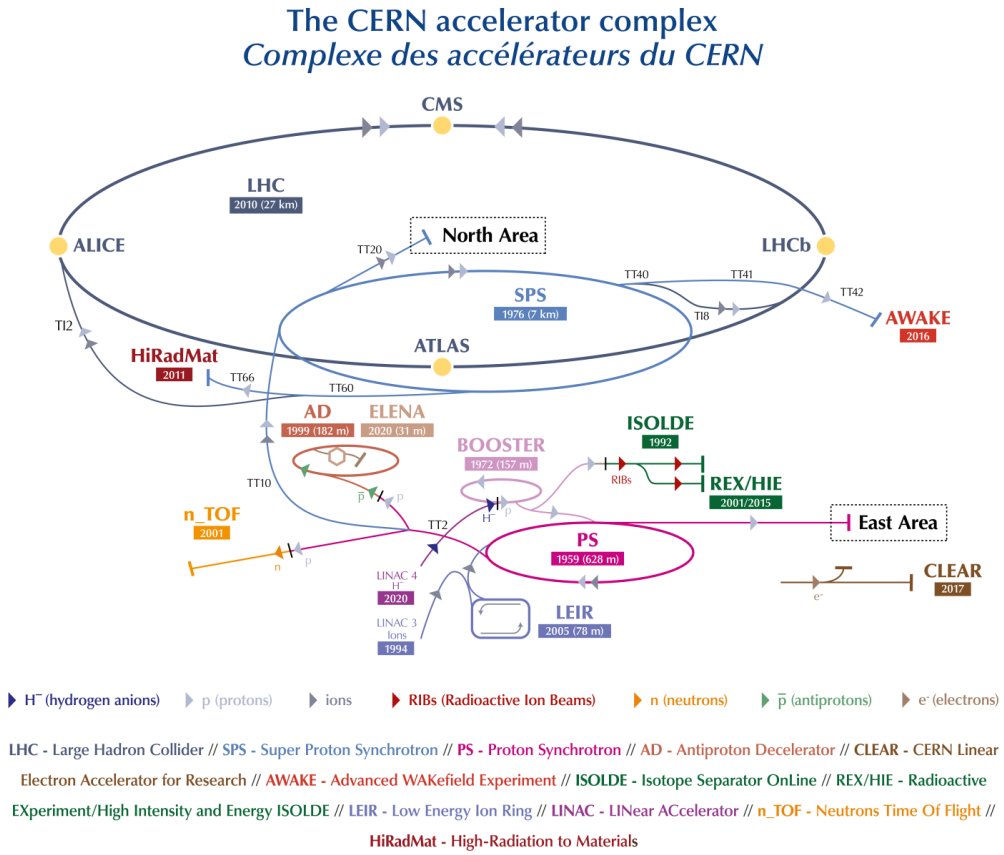


Figure 2.1 The CERN accelerator complex, from `cdsweb.cern.ch`.

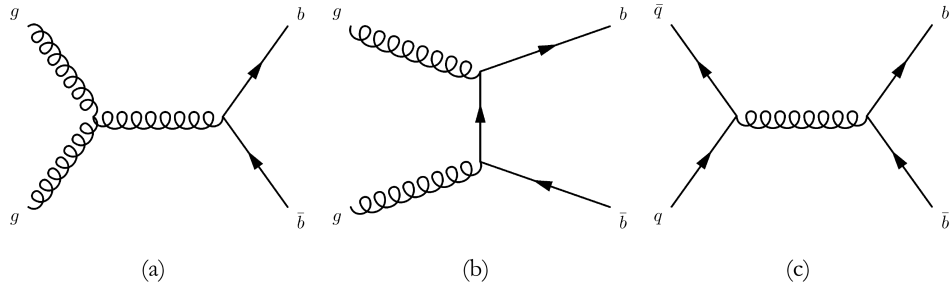


Figure 2.2 Feynmann diagrams at leading order in pp collisions: (a) and (b) shows gluon-gluon fusion processes and (c) shows quark-antiquark annhilation processes.

Table 2.2 Fragmentation fractions for different hadrons containing b quark. For the first row, the B and B^+ fraction are set equal and the number only applies for one type of B meson and not for both together. The numbers are taken from Ref. [66].

Fraction	Value
B^0 or B^+ fraction	0.404 ± 0.006
B_s^0 fraction	0.103 ± 0.005
b-barions fraction	0.088 ± 0.012

Fig. 2.2. For this reason, beauty hadrons are predominantly produced in the forward and backward region with small angles with respect to the beam axis, as shown in Fig. 2.3.

The $b\bar{b}$ cross-section at TeV energy scale increases approximately linearly with energy. In the forward region of the LHCb experiment it is measured to be [65]:

$$\sigma(pp \to H_b X)(7\,\text{TeV}) = (72.0 \pm 0.3 \pm 6.8)\,\mu\text{b} \tag{2.2}$$

$$\sigma(pp \to H_b X)(13\,\text{TeV}) = (144.3 \pm 1 \pm 21)\,\mu\text{b} \tag{2.3}$$

at 7 and 13 TeV, respectively. The cross section is much larger than the one of the B-factories, which exploit the electromagnetic process and is around 1 nb. The relative production of different beauty hadrons, *i.e.* $B^0_{(s)}$ and B^+ mesons and b-baryons, can be extracted from the measurement of the fragmentation fractions, which are associated to the probability that individual quarks fragment into colorless bound states. Values are reported in Table 2.2.

The LHCb detector [63, 64] is a single arm spectrometer with a forward geometry, covering angles from about 10 mrad to approximately 300 mrad in the horizontal plane, which is the magnet bending plane, and 250 mrad in the non-bending plane and covers a pseudorapidity[1] range of $2 < \eta < 5$. The b and $\bar{b}$ quark fraction produced in a pp collision inside the LHCb acceptance is 27%. This geometry allows LHCb to reconstruct a large fraction of the produced particles containing a b and $\bar{b}$ quark while only covering

[1]Pseudorapidity is defined as $\eta = -\ln\tan(\frac{\theta}{2})$, where θ is the angle of the b-hadron with respect to the proton direction.

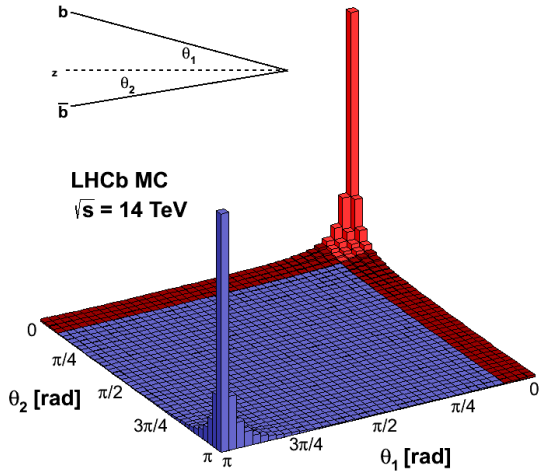


Figure 2.3 Simulated $b\bar{b}$ events at $\sqrt{s}$= 14 TeV. The red region corresponds to the LHCb acceptance. From Ref. [64].

a small solid angle. The LHCb detector layout is shown in Fig. 2.4. It includes a high-precision tracking system composed by the Vertex Locator (VELO) to measure heavy hadron displaced vertices detached from primary pp collisions and TT and T1,T2, T3 stations upstream and downstream of the dipole magnet, to provide particle momentum and charge, see Sec. 2.2.1. Different types of charged hadrons are distinguished using information from two ring-imaging Cherenkov detectors, see Sec. 2.2.2. Photons, electrons and hadrons are identified by a calorimeter system consisting of scintillating-pad and preshower detectors, an electromagnetic calorimeter and a hadronic calorimeter. Muons are identified by a system composed of alternating layers of iron and multiwire proportional chambers, explained in Sec. 2.2.3 [67].

The design luminosity of LHCb in Run 2 is 4×10^{32} cm^{-2} s^{-1}, two orders of magnitude lower than the maximum capacity of the LHC. LHCb has decided against using the full luminosity provided by the LHC for several reasons. The forward region is dominated by a very high flux of particles which creates high occupancies in the detectors and induces radiation damage. Running at full luminosity would have put severe constraints on the choice of detector materials and segmentations to use. Moreover, separation of primary and secondary vertices is crucial for many analyses in LHCb, a task which is more difficult with large pile-up. Additionally, the ability to reconstruct all tracks in the event degrades with increasing number of interactions. The instantaneous luminosity for LHCb was therefore chosen as a compromise of all these parameters. As a measure of the pile-up, we use a quantity called μ_{vis}, which represents the average number of inelastic interactions per bunch crossing, defined as:

$$\mu_{vis} = \frac{\mathcal{L}_{tot} \cdot \sigma_{inel}}{N_{BC}}, \tag{2.4}$$

where σ_{inel} is the inelastic cross-section, $\mathcal{L}_{tot}$ the total luminosity and N_{BC} the number

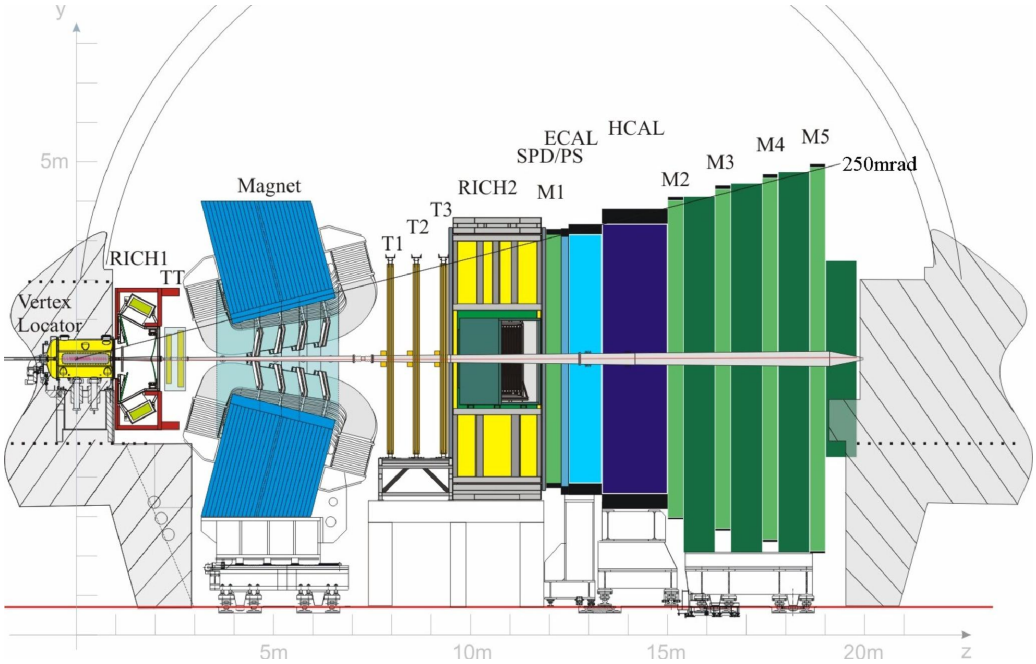


Figure 2.4 Lateral view of the LHCb detector, from `cdsweb.cern.ch`.

of bunches. This quantity is estimated to be 1.79 in Run 1 and ~ 1 in Run 2 at LHCb.

2.2.1 Tracking system

The Vertex Locator (VELO) [68] is the subdetector closest to the pp interaction point. It is designed to precisely measure the position of primary and secondary vertices, which is crucial for most LHCb analyses. B hadrons at LHCb have a mean flight distance from about 10 mm to 1 cm. Due to high track multiplicity in LHC collisions, it is crucial to have a vertex locator with a micrometer precision in order to select signal events and reject most of the background.

The VELO is composed of 21 circular "stations" of silicon modules in a $r-\phi$ geometry installed along the beam line, as shown in Fig. 2.5. This geometry allows to measure the distance from the beam and the azimuthal angle of hits, namely r and ϕ, generated by the ionizing particles that cross the VELO. The third coordinate z is simply measured knowing what modules give a signal for a particular particle hit.

Each VELO station is divided into two halves; detectors modules are installed on the left and on the right side of the beam axis and a hole is left free in the middle for the beams. During beam injection, the two halves are retracted to about 6 cm away from each other; while, when the LHC is in proton-proton collision, the halves are moved in with the sensors at a distance from the beam of only 8 mm.

The sensors are built with 220 µm-thick silicon microstrip $n^+ - in - n$ technology; in the R-sensors the strips have a semi-circular shape with an increasing pitch from 40 µm in the innermost part to 101.6 µm for the outermost strip. The strips in the ϕ-sensor have a radial shape and are split into an inner and an outer region, chosen to equalize the occupancy. With the $r\phi$ geometry the 2D (rz) tracking allows a fast reconstruction in

the Level 1 trigger with sufficient impact parameter resolution to efficiently select events containing b-hadrons.

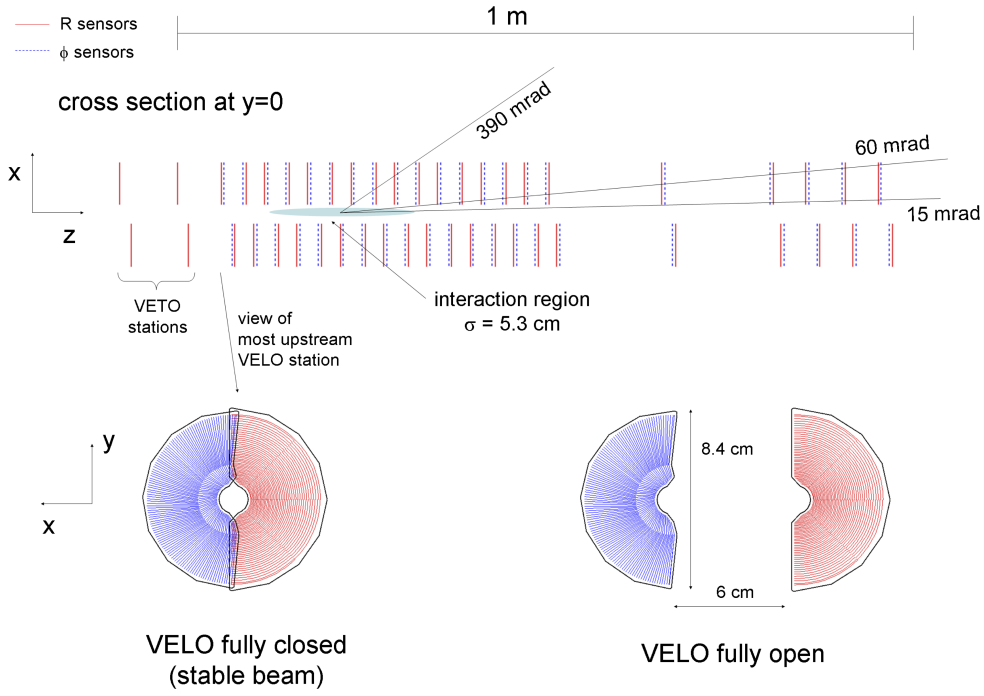


Figure 2.5 Top: top view of the VELO silicon sensors. Bottom: frontal view of the modules in the closed and open positions. From Ref. [63].

The performance of the VELO detector has been analysed using the data collected in 2011-2012 [69]. The resolution on the primary vertex (PV) in the x and y coordinates ranges from 40 μm to 10 μm depending on the number of tracks produced in the pp collision, while the resolution on the z coordinate ranges from 250 μm to 50 μm and the trend is shown in Fig. 2.6. The track impact parameter (IP), defined as the distance between the track and the PV at the track's point of closest approach to the PV, has a resolution of tens of μm, inversely proportional to the track transverse momentum, Fig. 2.7. The IP resolution is governed by multiple scattering of particles due to the detector material; the resolution on the position of hits in the detector from which tracks are reconstructed; and the extrapolation distance of a track from its first hit to the PV.

Besides the VELO detector, the tracking system is composed by the Tracker Turinces (TT), located between RICH1 and the magnet, and three tracking stations T1, T2, T3, located after the magnet. This system provides an high momentum resolution on tracks traversing the whole detector, an efficient track reconstruction and information for the second level trigger.

The TT consists of four planes of silicon microstrip sensors with a pitch of 183 μm, strip length up to 37 cm and thickness of 500 μm, see Fig. 2.8. While the first and the last layer have sensors with strips running vertically, the second and third layer are tilted

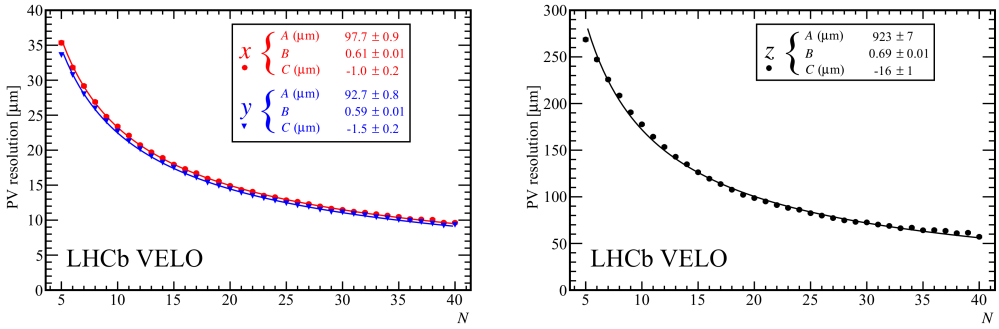


Figure 2.6 PV resolution of events with exactly one PV in 2011 data as a function of track multiplicity. (left) x (red) and y (blue) resolution and (right) z resolution. The fit parameters A, B and C for each coordinate are given. From Ref. [69].

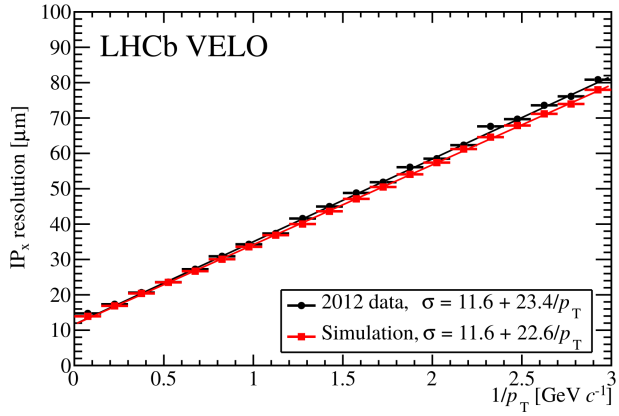


Figure 2.7 IP_x as a function of $1/p_T$, determined with 2012 data and compared with simulation. From Ref. [69].

with an angle of +5° and -5° with respect to the vertical axis, in order to provide tracking information in the vertical direction. The hit resolution is of about 50 µm. The TT is mainly used to improve momentum resolution providing reference segments used to combine the tracks reconstructed in the tracking stations with those reconstructed in the VELO. It also improves the rejection of ghost tracks, which are tracks that do not belong to a real particle.

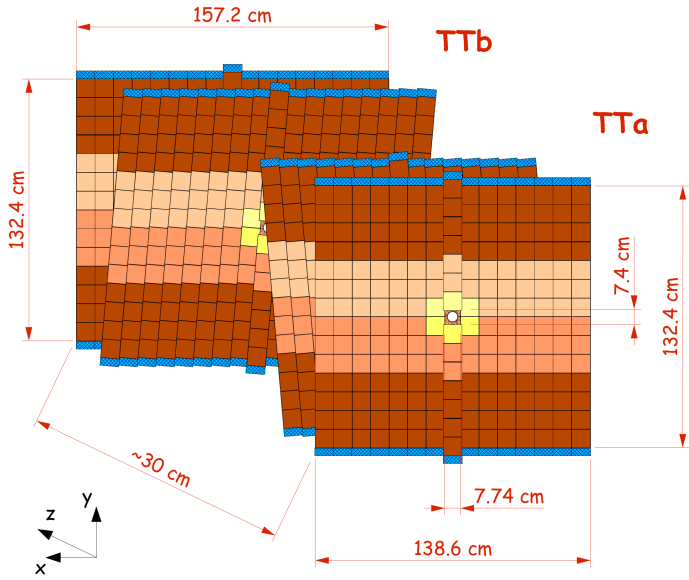


Figure 2.8 TT illustration. The first and the fourth stations have sensors parallel to the vertical axis, while the second and the third stations (called u-plane and v-plane) have sensors titled of +5° and -5°, respectively. The different colours correspond to different readout sectors. The blue region is the position of the readout electronics. From `lhcb-public.web.cern.ch`.

Three tracking stations are placed downstream of the magnet, and are divided in two main parts: the Inner Tracker (IT) [70] and the Outer Tracker (OT) [71]. They are arranged as in Fig. 2.9. The IT forms the inner part closest to the beam pipe. As the particle flux is higher in these regions, silicon sensors were chosen as detecting devices. The tracking stations are composed by four layers of silicon sensors with the same characteristics as the TT. The strip pitch is of 198 µm which leads to a similar resolution as in TT. The total IT size is about 1.2 m in the bending plane and about 40 cm in the vertical plane. The OT consists of an array of straw tubes modules; each plane is composed by two rows of straw tubes, arranged in a honeycomb structure. A single drift tube has a cylindrical shape and its volume is filled with a gas mixture of argon (70%) and carbon dioxide (28.5%) and oxygen (1.5%), which allows a short drift time. The distance of closest approach of the particle to the anode wire is used to measure the coordinate of the particle. This allows a position resolution of 200 µm for a single cell with a hit efficiency of more than 99%. The largest OT station covers a sensitive area of $600\,\text{cm} \times 490\,\text{cm}$.

In order to measure the momentum of charged tracks, a warm (i.e. non supercon-

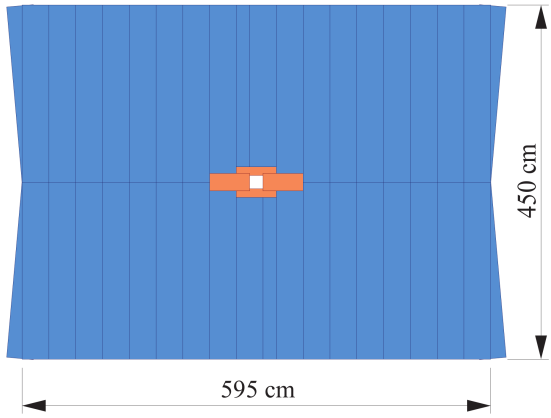
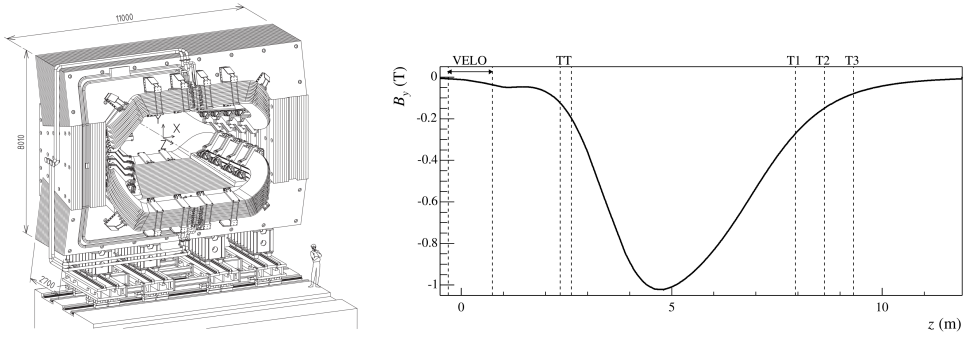


Figure 2.9 Sketch of one of the T stations with the OT in blue and the IT in orange. From Ref. [72]

ducting) dipole magnet [73] with two aluminium coils inside an iron yoke is placed between the TT and T stations. Its shape and field strength are illustrated in Fig. 2.10. The magnet has an integrated field of $\int \vec{B}\, \vec{dl} = 4$ Tm, where the main field component is along the y axis with a maximum intensity of 1 T. Note that there is practically no field at the position of the VELO, while there is only a small field at the position of the TT and the T stations. The magnetic field polarity can be reversed, a procedure which is regularly undertaken during operation of the detector to minimize systematic effects. A very good momentum resolution of charged particles is reached by the LHCb tracking system, with a relative uncertainty of 0.5% at low momentum up to 1.0% for $p = 200$ GeV [64], Fig. 2.11.



(a) (b)

Figure 2.10 Technical drawing of the dipole magnet in LHCb (a) and strength of the y direction of the magnet field as a function of the z coordinate (b). From Ref. [73].

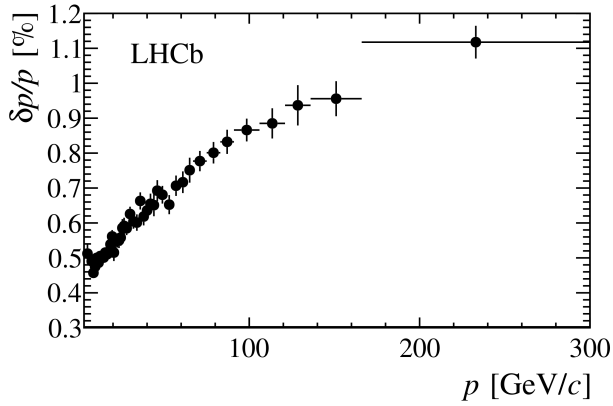


Figure 2.11 Relative momentum resolution versus particle momentum. From Ref. [64].

Track reconstruction

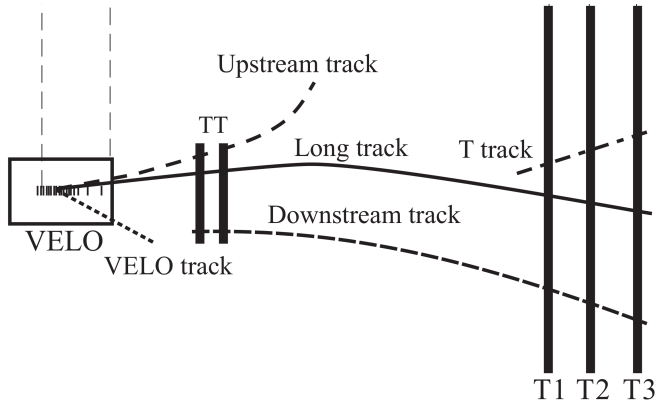


Figure 2.12 Illustration of various track types.

Tracks are reconstructed from hits in the VELO, TT and T stations detectors. Different types of tracks are defined depending on their path in the detector, as shown in Fig. 2.12:

- **Long tracks** generate hits in all tracking sub-detectors and are used for most physics analyses due to the most precise momentum estimate.
- **Upstream tracks** are formed by particles with a very low momentum that are swept out by the magnetic field from the detector acceptance before they can reach T1. It is possible to measure their momentum thanks to the residual magnetic field present between the VELO and the TT, even if the measurement is affected by a 20% relative uncertainty.
- **Downstream tracks** are generated from long-lived neutral particles (*i.e.* K_s^0 or

Λ) that decay between the VELO and the TT, and produce charged particles with associated hits in the TT and in the three tracking stations.

- **VELO tracks** pass only through the VELO and are typically large-angle or backward tracks, which are useful for the primary vertex reconstruction.
- **T tracks** are track segments with hits in the tracking stations only.

The track types used for physics studies are the Long and the Downstream tracks. Track finding and reconstruction algorithms are organized in different steps. The first starts with definition of segments in the various sub-detectors. Inside the VELO, the segments are created matching all hits that lie on a straight line; at least three hits in the R-sensors and three hits in the Φ-sensors are needed to form a segment. In the tracking stations, a segment is created matching the hits contained in a section of the first and third stations, using the information given only by one plane of vertically orientated microstrip sensors. Then, under the hypothesis of a parabolic trajectory, the algorithm calculates the position of the hit in the middle stations and searches for compatible hits. If a signal is found, it is added to the segment and it is used to better determine the parameters of the trajectory. Finally, the compatible hits coming from the tilted layers are also added, in order to have a 3-dimensional segment. The reconstruction process is organized in a hierarchical way: the algorithm tries firstly to reconstruct long tracks and then it picks up unused segments to reconstruct downstream and upstream tracks. After finding segments compatible with straight lines, Long tracks are reconstructed with two algorithms: the first extrapolates VELO segments to the tracking stations, adding to the track the compatible hits in the TT, the second matches VELO and tracking stations segments one to each other, extrapolating VELO segments in the forward direction and tracking stations segments in the backward direction. Downstream tracks are reconstructed starting from the T stations segments and then adding the compatible hits in the TT to those segments. Upstream tracks are obtained extrapolating VELO segments to the TT, adding compatible hits and requiring a non compatibility with any of the tracking station segments.

The tracks are then fitted with a Kalman filter [74] which computes the χ^2 as an estimate of the track quality. Finally a clone killer compares the reconstructed tracks, two by two: if a pair of tracks shares more than a 70% of hits in VELO and in the T-seeding regions, they are considered clones and only the one with the best χ^2 is stored.

2.2.2 Particle identification system

Most of the LHCb data analyses, and in particular those presented in this thesis, require the identification of charged particles. This task is accomplished by some dedicated sub-detectors.

Two Ring Imaging Čerenkov Detectors (RICH) [75] are employed in LHCb for charged hadrons identification and cover different momentum ranges. The structure of the two apparatus is shown in Fig. 2.13. The working principle is based on the Čerenkov effect, namely particles traversing a medium with velocity v faster than the speed of light in this medium c/n emit photons at an angle

$$\cos(\theta_c) = \frac{1}{n\beta}, \tag{2.5}$$

where n is the refraction index of the material and $\beta = v/c$. Knowing the refractive index of the material being traversed, this angle can be related to the speed of the particle.

With the knowledge of the particle's momentum, the invariant mass of the particle can be determined.

RICH1 is located between the VELO and the TT and identifies low momentum particles between $2\,\text{GeV}/c < p < 60\,\text{GeV}/c$. It covers the full LHCb acceptance and during Run 1 it uses two kind of radiators: the first is a 5 cm thick Aerogel layer, with $n = 1.03$, to provide separation for particles with a momentum of few GeV/c, while the second, gaseous C_4F_{10} with $n = 1.0015$, detects higher momentum particles (up to 60 GeV/c). In Run 2, the Aerogel is removed because its capability to identify kaons was degraded due to the high multiplicity environment. The second Čerenkov detector, RICH2, is located between T stations and the first muon station. It has a reduced angular acceptance of 15 mrad to 120 (100) mrad in the bending (non bending) plane and uses CF_4 as a radiator, with $n = 1.0005$ which exploits to identify high-momentum particles ($15\,\text{GeV}/c < p < 100\,\text{GeV}/c$).

The Čerenkov light is reflected via a spherical and a plane mirror into an array of Pixel Hybrid Photon Detectors (HPDs) which provide a photon position measurement, allowing the reconstruction of the "rings" created by the Čerenkov light. The radius of these rings is proportional to the Čerenkov angle, which is shown in Fig. 2.14 (top) for isolated tracks in C_4F_{10}. The ability of identifying protons with respect to pion tracks is shown in Fig. 2.14 (bottom) as the difference of log-likelihood between the two hypotheses.

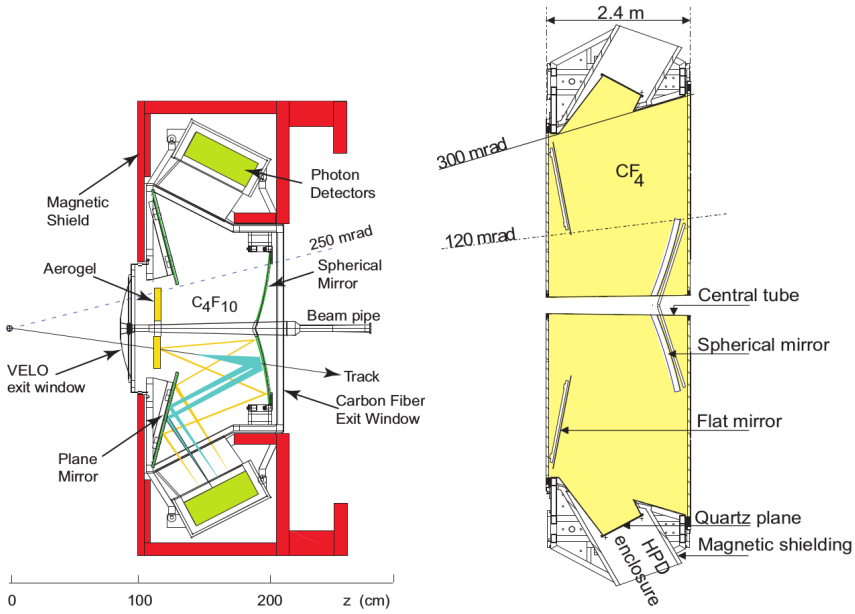


Figure 2.13 Left: schematic layout from a side view of the RICH1 detector. Right: schematic layout from a top view of the RICH2 detector. From Ref. [75].

The calorimetry system [77] has been devised to reconstruct high transverse energy hadrons, electrons and photons (particularly π^0) with good precision, providing infor-

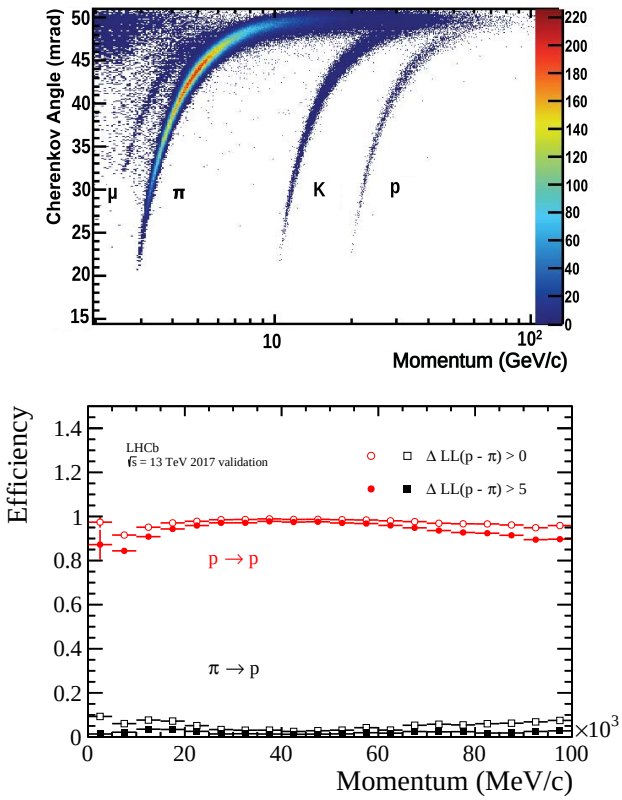


Figure 2.14 Čerenkov angle for C_4F_{10} radiator versus the particle momentum (top) from Ref. [76]; proton identification efficiency and pion misidentification rate as measured using 2017 data (bottom), from `twiki.cern.ch`.

mation about the energy and position to be used in the L0-trigger, for fast decisions.

It consist of three parts: a scintillating pad/preshower detector (SPD/PS), consisting of 15 thick lead absorbers corresponding to 2.5-3 X_0 radiation lenghts, followed by an electromagnetic calorimeter (ECAL) and an hadronic calorimeter (HCAL). The SPD/PS and the ECAL exploit the same substructure by dividing the active region in three parts, with a more coarse granulation when moving away from the beam pipe. The HCAL has a further division into two subparts.

The structure of the ECAL is an alternating planes with absorbing layers made of lead (2 mm thick) and detecting layers made of scintillating material (4 mm thick, polystyrene). In total, there are 66 Pb and scintillating layers that form a stack of 42 cm length, which corresponds to 25 X_0.

The HCAL main task is to measure the energies of hadronic showers, which is fundamental for the Level-0 trigger. The HCAL is built as a sampling device as well, made of layers of scintillators 4 mm thick separated by layers of iron 16 mm thick. The total material budget corresponds to 5.6 nuclear interaction lengths. The readout of the scintillating parts is done via wavelenght shifting fibers guiding the light to photomultipliers. The energy resolution are measured to be $\sigma(E)/E = 8 - 10\%/\sqrt{E}/\,\text{GeV} \oplus 0.9\%$

for ECAL and $\sigma(E)/E = 69\%/\sqrt{E}/\,\text{GeV} \oplus 9\%$ for HCAL.

2.2.3 Muon system

The muon system [78] consists of five stations, the first one (M1) placed upstream of the calorimeters to mitigate muon multiple scattering effects, and four stations (M2 to M5) situated after the HCAL, and separated from each other with layers of 80 cm of iron, which serve as an absorber for all particles not being muons or neutrinos (see Fig. 2.15 (left)). Similar to the calorimeters, the muon system uses different segmentations to cope with the varying particle flux over the full area: all stations are therefore divided into four regions (R1 to R4) with increasing distance from the beam pipe, which linear dimension and logical pad size scale with factors of two with the beam distance, Fig. 2.15 (right). The muon system covers an acceptance between 20 (16) mrad and 306 (258) mrad in the bending (non-bending) plane. It exploits multi-wire proportional chambers (MWPC) for all stations, except for the innermost region of M1, where triple-GEM detectors are used due to the high particle flux and its demand for radiation hardness. MWPCs have four overlapped gaps, each one 5 mm thick and with a distance between wires of about 2 mm. They are built with a gas mixture of Ar-CO_2-CF_4 $(40 : 50 : 10)$ and operate at a voltage of $3.0 - 3.2$ kV. In M1, the MWPCs have only two gaps to reduce the amount of material in front of the calorimeters. The triple-GEM detector consists of three gas electron multiplier foils sandwiched between anode and cathode planes. The muon system is part of the LHCb trigger and is designed to reach a 95% trigger efficiency for events containing one muon track. A muon track needs to have a minimum momentum of 6 GeV/c to be efficiently triggered. Muon efficiency and hadron-muon misidentification rates are illustrated in Fig. 2.16.

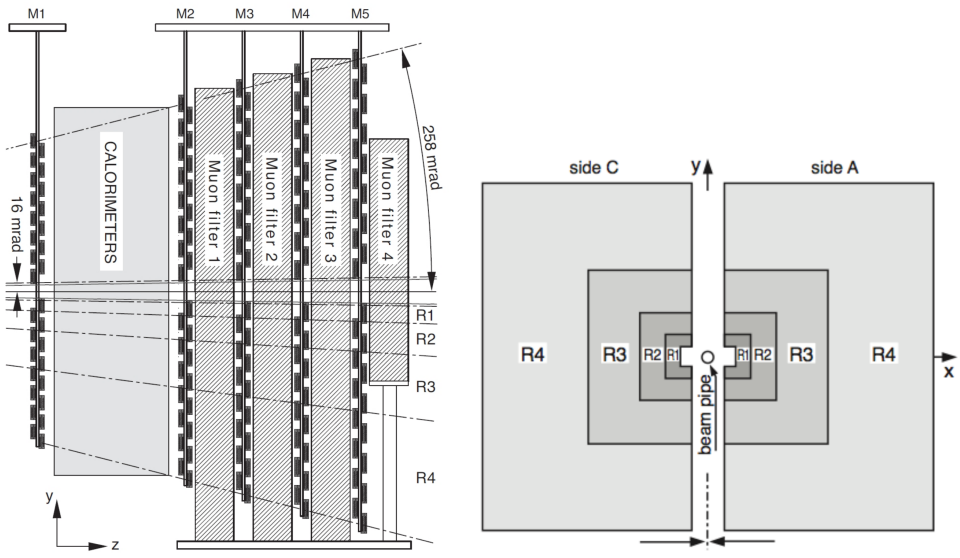


Figure 2.15 Side view of the five chambers of the muon system with the interleaved calorimeters and iron absorbers. From Ref. [78, 79].

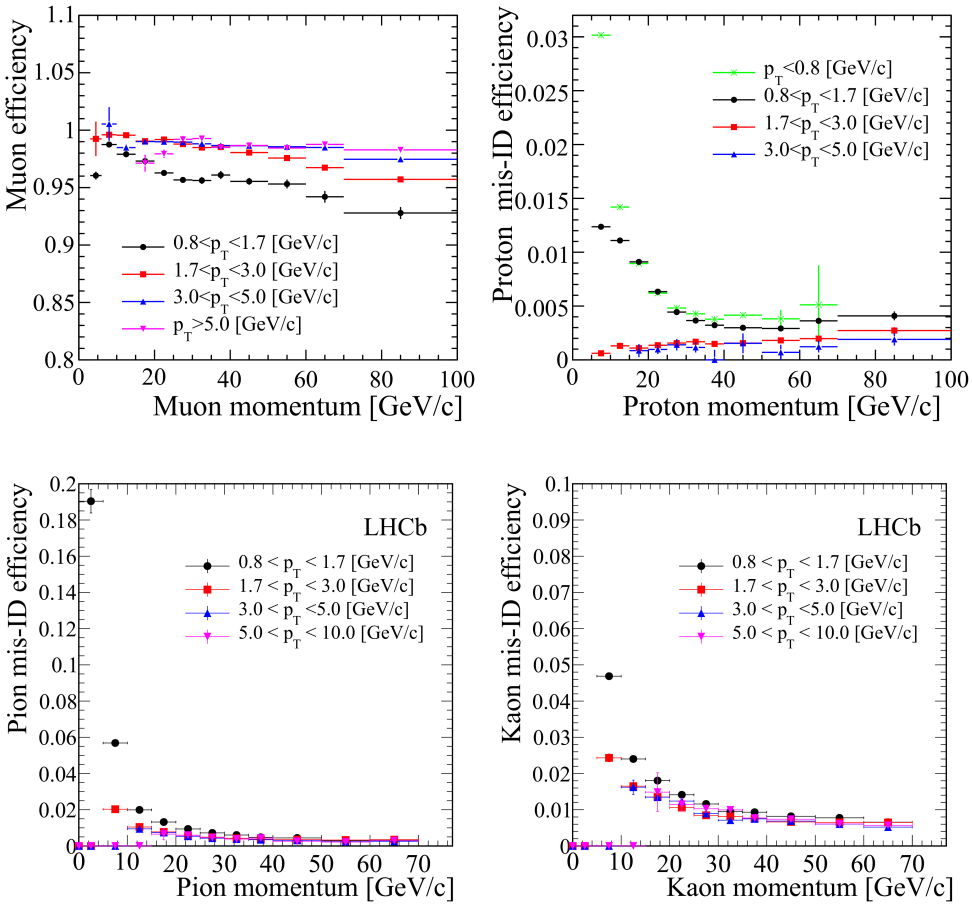


Figure 2.16 Muon efficiency and hadron-muon misidentification probabilities, as a function of momentum for different p_T ranges. From Ref. [64].

2.2.4 Trigger

The pp bunch crossing rate occurs at a frequency of 40 MHz. At this frequency, the raw data storage is too big to be recorded and only few pp interactions produce interesting events. For this reason, a trigger system is built to reduce the event rate down to 3 kH and 12.5 kH in Run 1 [80] and Run 2 [81], respectively. The trigger is composed by two levels: the Level-0 trigger (L0), the first stage of the trigger implemented on hardware level, and the High-Level trigger (HLT), a two stage software trigger executed on event filter farms, to confirm the L0 decisions and to perform the event reconstruction for further data reduction. A sketch of the trigger levels is shown in Fig. 2.17 for Run 1 and Run 2. The LHCb trigger in Run2 has been improved including real-time detector alignment and calibration, using information directly from the trigger system.

The Level-0 trigger (L0) is in common between the two run periods. It runs synchronously with the LHC clock and reduces the rate of pp collisions to about 1 MHz,

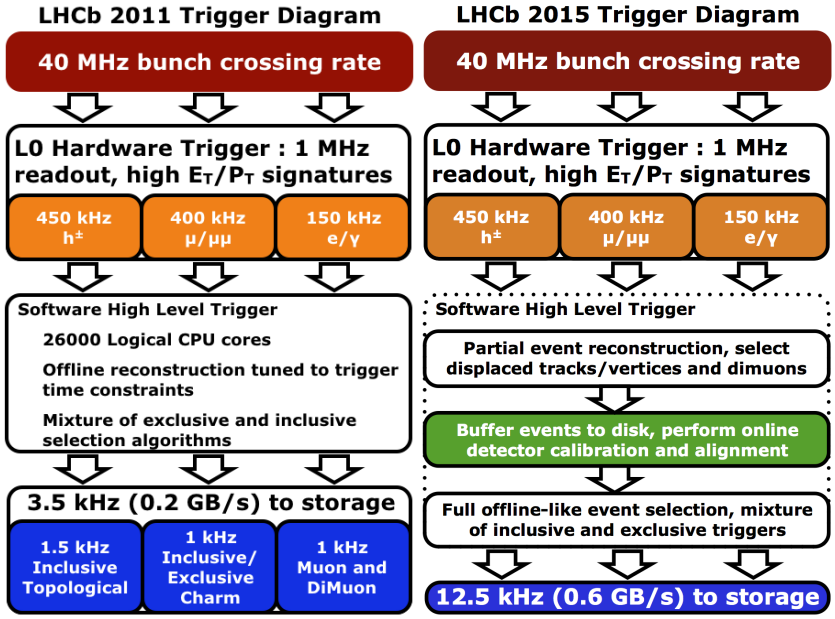


Figure 2.17 LHCb trigger diagram in Run 1 (left) and Run 2 (right). From Ref. [80] and [81].

which is the maximum rate at which the full detector can be read out. It combines the information from the calorimeters and the muon system, selecting only candidates with high p_T or high transverse energy at few GeV level.

The High Level trigger (HLT) is software-based and is divided into two stages, HLT1 and HLT2. HLT1 is faster than HLT2 and performs an inclusive selection of events based on one- or two-track signatures, on the presence of muon tracks displaced from the PVs, or on dimuon combinations. In Run2 the events triggered by the HLT1 are buffered to disk storage, in order to be processed during inter-fill periods, and align the detector before the HLT2 stage. HLT2 runs on events triggered by HLT1 and performs a partial events reconstruction in Run1 and a full reconstruction in Run2. It allows to reconstruct a lot of inclusive and exclusive decays due to a great variety of trigger, ranging from lines selecting prompt and detached muons to "topological lines" selecting n-body decays starting from two charged daughter tracks. The output rate of HLT2 was up to 3 kHz and 12.5 kH in Run1 and Run2, respectively.

2.3 LHCb data sample and analysis software

LHCb has collected data from 2011 to 2018, at the center-of mass energy of 7 and 8 TeV in Run 1 (2011-2012), and of 13 TeV in Run 2 for pp collisions. The total integrated luminosity in the two run periods amounts to 3 fb^{-1} and 5.5 fb^{-1} in Run 1 and Run 2, respectively. In Fig. 2.18, the collected luminosity per year of data-taking is summarized.

The LHCb analysis software is based on the GAUDI [82] framework, which is developed to handle event data processing applications for High Energy Physics experiments. It includes different packages that take care of the full process of the event reconstruc-

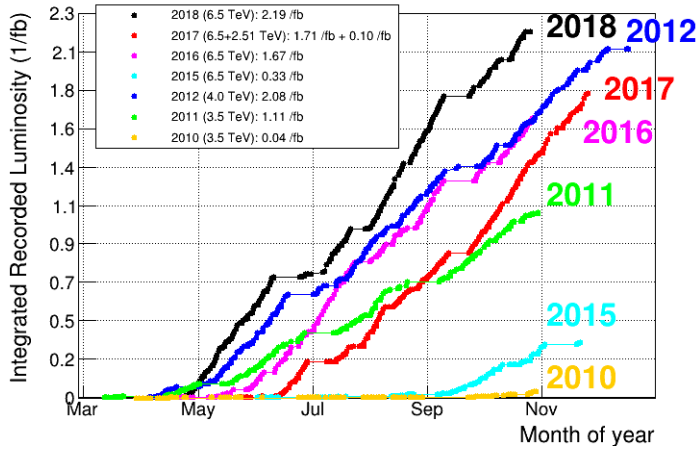


Figure 2.18 LHCb integrated luminosity for each year of data-taking, from `twiki.cern.ch`.

tion, starting from the trigger stage to the vertex fitting, as well as event simulation.

The data processing consists of several steps, each one being controlled by an 'application' that processes the data event-by-event. The first step is the trigger, controlled by the MOORE [83] application. The data reconstruction is then performed by the BRUNEL [84] package, which takes as input the triggered events ready for the offline analysis. It includes all the reconstruction process: the creation of tracks from the hits in the subdetectors, the reconstruction of the particle momentum as well as primary and secondary vertices and the PID identification step. Afterwords, data are filtered further through a set of selections called the *stripping*, controlled by the DAVINCI [85] application. Decays that exhibits common features are also grouped into *streams*, in order to speed up the analysis process.

Simulations with Monte Carlo (MC) methods are also handled by the same framework and are processed in a very similar way to real data. There are two simulation steps which replace the proton-proton collisions and the detector response. The simulation of proton-proton collisions, and the hadronisation and decay of the resulting particles, are controlled by the GAUSS [86] application, which is interfaced with specific packages for handling generation, such as PYTHIA [87], and simulation phases, such as EVTGEN [88] and GEANT4 [89]. The generation of physical processes, as the proton-proton interaction and the hadronization process, is performed by PYTHIA. EVTGEN is used to describe the decays of simulated particles, while GEANT4 is used to simulate the propagation and interaction of particles through and with the detector. The simulated hits made in the virtual detector are converted to signals that mimic the real detector by the BOOLE [90] application. The output of BOOLE is designed to closely match the output of the real detector, and so the simulated data can then be passed through the usual data processing chain described above, beginning with the trigger.

2.4 LHCb upgrade

The LHCb experiment has achieved very important results during Run 1 and Run 2. The most relevant ones include the observation of $B_s^0 \to \mu^-\mu^+$ decays [91], the measurement of the anomalies in the flavour changing neutral currents $B^0 \to K^{0*}\mu^+\mu^-$ [92], the CP violation measurements in B^+ and D^0 decays as well as in B_s^0 decays [93] and the observation of pentaquarks in $\Lambda_b^0 \to J/\psi pK$ decays [32]. However, many crucial measurements are statistically limited [94], with experimental precisions not directly comparable to the uncertanties of the theoretical predictions. Therefore, in order to reach the theoretical uncertainties and search for new physics, an upgrade is needed. The LHCb detector is currently undergoing a major upgrade in order to be able to run at a luminosity of $2 \times 10^{33}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, five times higher than the present one [95], and to integrate a luminosity of $50\,\mathrm{fb}^{-1}$ during Run 3 and Run 4. In order to achieve this goal, major changes in the LHCb design have been implemented. Firstly, the LHCb sub-detectors will be read out every 25 ns, or equivalently at the pp-collision frequency of around 40 MHz. This implies that the hardware trigger $L0$, which now limits the read-out at 1 MHz, will be replaced by a full software trigger in two stages [96]. This trigger will select the interesting events to be stored and will perform real-time detector calibration and alignment. After the second software level the data will be ready for offline analysis. This trigger scheme will allow to increase significantly the efficiency for a broad range of LHCb physics channels. This change requires the replacement of all front-end electronics, in particular of the tracking devices.

Secondly, the higher luminosity will increase the occupancy of the detectors; the average number of (visible) interactions per bunch crossing is estimated to be $\mu \sim 5$, instead of the actual value of ~ 1. Therefore, finer module granularity will be needed to cope with it, together with more efficient radiation hard sensors.

For these reasons, most of the LHCb subdetectors will be replaced. The tracking systems will be composed by: a new vertex locator of $55\,\mu\mathrm{m} \times 55\,\mu\mathrm{m}$ pixel sensors [97], an upstream tracker (UT), consisting of silicon microstrips with finer segmentation, as described in Sec. 6.1, and a scintillating fiber tracker (Scifi) [98], able to achieve spatial resolutions of $\sim 70\ \mu\mathrm{m}$. Particle identification detectors will also be updated [99]. The RICH1 detector will use modified optics and mechanics to reduce occupancy, together with new multianode photomultipliers with finer granularity. The calorimeter system will lower the photomultiplier tubes gain in order to withstand the higher particle fluxes. The SPD and PS detectors, as well as the first muon station, will be removed.

Besides the current upgrade, future detector upgrades will take place during Long Shutdown 3 (Upgrade Ib) and 4 (Upgrade II) [100]. The goal is to fully exploit the High Luminosity LHC (HL-LHC) [101] potential by reaching an integrated luminosity of 300 fb^{-1} during Run 5 and 6. Different projects are under consideration to build detectors with better granularity and radiation hardness, to improve coverage for low p_T tracks and to include timing information in tracking (4D tracking). This will allow to effectively distinguish different collision vertices.

The current and foreseen LHCb upgrades within LHC schedule are summarised in Table 2.3, along with the recorded and planned instantaneous and integrated luminosities in pp collisions.

During Run 3, LHCb will be collecting data at the same time as the Belle II experiment [102] at the SuperKEKB asymmetric e^+e^- collider. Although the two experiments have similar flavour physics programme, they have complementary strength due to the different collision environments. On one hand, LHCb can study all beauty hadrons, in particular B_s^0 and beauty baryons as Λ_b, and it will have much larger yields and lower

backgrounds for decays involving charged particles in the final state. On the other hand, Belle II has better capability for channels involving neutral particles or missing energy and can study all properties of the τ lepton thanks to the $e^+e^- \to \tau^+\tau^-$ production mode. Belle II is expected to collect data corresponding to $50\,\text{ab}^{-1}$ at around the $\Upsilon(4S)$ resonance by 2025. However, there are no current plans for any subsequent upgrade. Therefore, LHCb will probably be the only flavour physics experiment in the HL-LHC era.

	LHC			HL-LHC		
	current LHCb		Upgrade I	Upgrade Ib	Upgrade II	
Runs	Run 1	Run 2	Run 3	Run 4	Run 5	Run 6
(years)	2011-2012	2015-2018	2022-2024	2027-2030	2032-2034	2036-2040
$\int \mathcal{L}dt$	$3\,\text{fb}^{-1}$	$6\ \text{fb}^{-1}$	$25\,\text{fb}^{-1}$	$50\,\text{fb}^{-1}$	$300\,\text{fb}^{-1}$	
$\mathcal{L}$	$2 \times 10^{32}\,\text{cm}^{-2}\,\text{s}^{-1}$		$2 \times 10^{33}\,\text{cm}^{-2}\,\text{s}^{-1}$		$2 \times 10^{34}\,\text{cm}^{-2}\,\text{s}^{-1}$	

Table 2.3 LHCb schedule for current and foreseen upgrades divided by Runs, with corresponding instantaneous and integrated luminosity.

CHAPTER 3

Helicity formalism

The helicity formalism was developed by Jackob and Wick in 1956 [103, 104] for the analysis of decays involving relativistic particles with spin. With this approach, the decay amplitude is parametrized by a functional form depending on a set of free parameters, called couplings, which can be extracted from a fit to data. As a result, the helicity amplitude analysis allows to extract information about intermediate resonances, giving access to fundamental quantities in spectroscopy, as the mass, the width, the spin of the resonant states and the mutual interference among them.

In the following, an introduction on the helicity formalism is reported, along with some preliminary definitions used throughout the document.

The helicity formalism is based on the property that the helicity, defined as the projection of spin along the momentum $\vec{p}$ direction $\lambda = \vec{S} \cdot \hat{p}$, is invariant under Lorentz boosts and rotations along $\hat{p}$, that do not invert the particle's direction of motion. As a consequence, the relativistic basis vectors can be built as simultaneous eigenstates either of total angular momentum and helicity or of the momentum and the helicity. In the following a specific relation will be derived between states of particles with definite momentum, called plane-wave helicity states, and states with definite total angular momentum, called spherical-wave helicity states.

An alternative approach to describe the decay amplitude is based on the tensorial formarlism introduced by Zemach [105], exploiting the LS partial wave model, and then expanded with the usage of explicit covariant expressions [106, 107].

3.1 Rotations

Let us start with the introduction of some basic definitions used in the formulation of the helicity formalism related to rotation of reference frames and of spin eigenstates.

In the following, we will refer to rotations as active transformations between reference frames. The normalised vectors describing the initial reference frame, $(\hat{x}_0, \hat{y}_0, \hat{z}_0)$, are rotated to those describing the final reference system, $(\hat{x}_3, \hat{y}_3, \hat{z}_3)$, by means of a rotation operator $\mathcal{R}(\alpha, \beta, \gamma)$, defined as:

$$\mathcal{R}(\alpha, \beta, \gamma) = \mathcal{R}_{z_2}(\gamma)\mathcal{R}_{y_1}(\beta)\mathcal{R}_{z_0}(\alpha), = e^{-i\gamma J_{z_2}} e^{-i\beta J_{y_1}} e^{-i\gamma J_{z_0}} \tag{3.1}$$

which is composed by a first rotation of an angle α around the $\hat{z}_0$ axis, a second rotation of an angle β around the rotated $\hat{y}_1$ axis and a third rotation of an angle γ around the doubled rotated $\hat{z}_2$ axis. The rotation are then expressed in terms of angular momentum operators J.

In the helicity formalism, for each two body decay $A \to 12$, the initial coordinate system is defined in the rest frame of the decaying particle A, aligning the $\hat{z}$ axis to the

quantization axis of its spin. This initial system will be denoted as $(\hat{x}_0^{\{A\}}, \hat{y}_0^{\{A\}}, \hat{z}_0^{\{A\}})$, where the superscript $\{A\}$ is related to the particle rest frame, while the subscript "0" means the initial coordinates. For the first particle in the decay sequence (*i.e.*B_s^0), the choice of the initial coordinate system is arbitrary. However, once defined, these coordinates must be used consistently between all decay sequences described by the matrix element. A rotation $\mathcal{R}(\alpha, \beta, \gamma)$ defined in Eq. (3.1) is usually exploited in order to align the z axis to the momentum direction of one daughter particle, *i.e.* bringing the initial coordinate system to the final one: $(\hat{x}_3^{\{A\}}, \hat{y}_3^{\{A\}}, \hat{z}_3^{\{A\}})$. In this way, the action of the rotation operator between spin eigenstates of particle A, $|J_A, \lambda_A\rangle$, where J_A is the total angular momentum of A and λ_A its helicity, in the initial coordinate system and $|J_A, \lambda'_A\rangle$ in the rotated $(\hat{x}_3^{\{A\}}, \hat{y}_3^{\{A\}}, \hat{z}_3^{\{A\}})$ coordinate system, is obtained:

$$\mathcal{R}(\alpha, \beta, \gamma)|J_A, \lambda_A\rangle = \sum_{\lambda'_A} D^{J_A}_{\lambda'_A, \lambda_A}(\alpha, \beta, \gamma)|J_A, \lambda'_A\rangle \tag{3.2}$$

in which the Wigner D-matrices $D^{J_A}_{\lambda'_A, \lambda_A}(\alpha, \beta, \gamma)$ are the matrix elements of the rotation operator in the given spin reference frame:

$$D^{J_A}_{\lambda'_A, \lambda_A}(\alpha, \beta, \gamma) = \langle J_A, \lambda'_A|\mathcal{R}(\alpha, \beta, \gamma)|J_A, \lambda_A\rangle \tag{3.3}$$

$$= e^{-i\lambda'_A \alpha} d^{J_A}_{\lambda'_A, \lambda_A}(\beta) e^{-i\lambda_A \gamma}. \tag{3.4}$$

and the Wigner d-matrices elements are known combinations of trigonometric functions of β depending on $J_A, \lambda_A, \lambda'_A$ parameters [108]. To achieve the rotation of the original $\hat{z}_0^{\{A\}}$ axis onto the particle 1 momentum $(p_1^{\{A\}})$, as shown in Fig. 3.1, it is sufficient to rotate by $\alpha = \phi_1^{\{A\}}, \beta = \theta_1^{\{A\}}$, where $\phi_1^{\{A\}}, \theta_1^{\{A\}}$ are the azimuthal and polar angles of the particle 1 momentum vector in the original coordinates

$$\begin{aligned}
\phi_1^{\{A\}} &= \operatorname{atan2}\left(p_{1y}^{\{A\}}, p_{1x}^{\{A\}}\right) \\
&= \operatorname{atan2}\left(\hat{y}_0^{\{A\}} \cdot \vec{p}_1^{\{A\}}, \hat{x}_0^{\{A\}} \cdot \vec{p}_1^{\{A\}}\right) \\
&= \operatorname{atan2}\left(\left(\hat{z}_0^{\{A\}} \times \hat{x}_0^{\{A\}}\right) \cdot \vec{p}_1^{\{A\}}, \hat{x}_0^{\{A\}} \cdot \vec{p}_1^{\{A\}}\right) \qquad (3.5) \\
\cos\theta_1^{\{A\}} &= \hat{z}_0^{\{A\}} \cdot \hat{p}_1^{\{A\}}. \qquad (3.6)
\end{aligned}$$

By convention, the third rotation angle γ is set to 0 throughout this thesis, since the helicity remains invariant under a phase definition.

3.2 Spin states and helicity states

3.2.1 One-particle spin state

Now, the definition of a one-particle spin state is introduced. There are two possible choices in literature for spin states: the canonical and the helicity choices [110]. Let us consider an observer with a given coordinate system $S = (x, y, z)$, and a particle having momentum $\vec{p}$ in the system S.

The canonical choice corresponds to making a direct boost along the particle momentum direction, p, from the particle rest frame to the S frame. Therefore, a canonical state with definite angular momentum is defined starting from the particle rest frame as

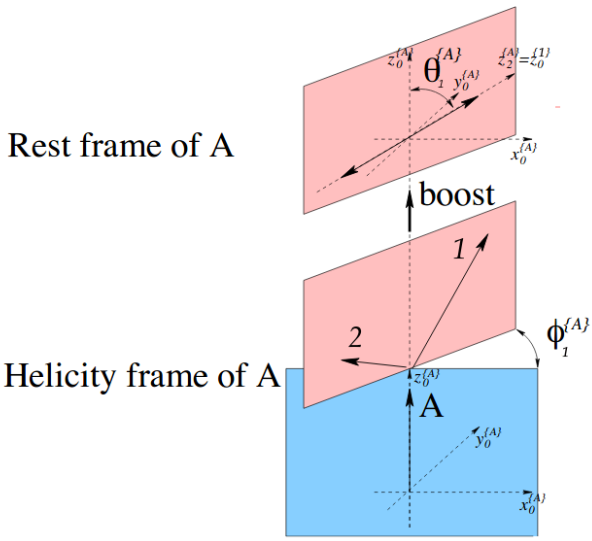


Figure 3.1 Coordinate axes for the spin quantization of particle A (bottom part), equal to the helicity frame of A ($z_0 \parallel p_A$ in the rest frame of its mother particle or in the laboratory frame), together with the polar ($\theta_1^{\{A\}}$) and azimuthal ($\phi_1^{\{A\}}$) angles of the momentum of its daughter particle 1 in the A rest frame (top part). Notice that the directions of these coordinate axes, denoted as $\hat{x}_3^{\{A\}}, \hat{y}_3^{\{A\}}, \hat{z}_3^{\{A\}}$, do not change when boosting from the helicity frame of A to its rest frame. After the Euler rotation $\mathcal{R}(\alpha = \phi_1^{\{A\}}, \beta = \theta_1^{\{A\}}, \gamma = 0)$ (see the text), the rotated z-axis, $\hat{z}_2^{\{A\}}$, is aligned with the particle 1 momentum; thus the rotated coordinates become the helicity frame of 1. If 1 has a sequential decay, then the same boost-rotation process is repeated to define the helicity frame for its daughters. Taken from Ref. [109].

$|0, s, m_z\rangle$, where s is the spin and m_z is the spin projection along its quantization axis, and then boosting along the direction of motion $L(\vec{p})$ to get the state:

$$|\vec{p}, s, m_z\rangle = L(\vec{p})|\vec{p} = 0, s, m_z\rangle. \tag{3.7}$$

While the helicity choice corresponds to aligning the particle spin quantization axis z along the particle momentum. Therefore, an helicity state, which is defined with definite momentum as $|\vec{p}, s, \lambda\rangle$, with $\vec{p}$ denoted by the polar and azimuthal angles θ, ϕ, and with helicityλ, is derived from the particle reference frame as:

$$|\vec{p}, s, \lambda\rangle = \mathcal{R}(\phi, \theta, 0)L(p\hat{z})|\vec{p} = 0, s, \lambda\rangle, \tag{3.8}$$

where the state $|\vec{p} = 0, s, \lambda\rangle$, which is a canonical state since $\lambda = m_z$ in the particle rest frame, is first boosted along the quantization axis $\hat{z}$ and then rotated by $\mathcal{R}(\phi, \theta, 0)$ to obtain the helicity state. This transformation is equivalent to making a rotation so that the quantization axis is along the momentum direction $\vec{p}$ and then boosting the system along p, due to the relation $L(\vec{p}) = R(\phi, \theta, 0)L(p\hat{z})R^{-1}(\phi, \theta, 0)$.

Comparing the definition in Eq. 3.7 and 3.8, the helicity and the canonical states are related by a rotation:

$$|\vec{p}, s, \lambda\rangle = \mathcal{R}(\phi, \theta, 0)|\vec{p}, s, m_z\rangle. \tag{3.9}$$

The advantage of using helicity states is that helicity is invariant under rotations and boosts along $\vec{p}$. It is worth to notice that any rotation of the reference frame around the momentum direction, represented by the third Euler angle ψ, would lead to equivalent helicity states $|p, s, \lambda_i\rangle$. Indeed, the helicity choice does not fix the axes orthogonal to the spin quantisation direction. These axes must be therefore defined by fixing a convention: we decide to choose $\psi = 0$ throughout the thesis.

3.2.2 Two-particle helicity state

In a two-body decay $A \to 12$, a state of the final two particles needs to be defined. The two-particle plane-wave helicity state is the product of the one-particle plane-wave helicity states:

$$|p, \theta_1, \phi_1, \lambda_1, \lambda_2\rangle = |\vec{p_1}, s_1, \lambda_1\rangle \otimes |\vec{p_2}, s_2, \lambda_2\rangle', \tag{3.10}$$

where in the decaying particle center-of-mass, the momenta of the final particles lie back to back $\vec{p_2} = -\vec{p_1}$ and the state of particle-2 is defined with a specific phase convention, called particle-2 convention, as described in Sec. 3.2.3. This state can be related to a canonical state $|0, J, \lambda_1 - \lambda_2\rangle$ of total angular momentum J and spin projections along $\hat{z}$ equal to the difference between the helicities of the final particles $M = \lambda_1 - \lambda_2$, through a sequence of boost and rotation, as done for the 1-particle case. So, we can define the two-particle helicity state $|p, \theta_1, \phi_1, \lambda_1, \lambda_2\rangle$ using the spherical coordinates of particle-1, the modulus $p = |\vec{p_1}| = |\vec{p_2}|$ and the angles θ_1 and ϕ_1, being

$$\begin{aligned} \theta_1 &= \arccos(\hat{z}_0^{\{A\}} \cdot \vec{p_1}), \\ \phi_1 &= \text{atan2}(\hat{y}_0^{\{A\}} \cdot \vec{p_1}, \hat{x}_0^{\{A\}} \cdot \vec{p_1}). \end{aligned} \tag{3.11}$$

where $(x_0^{\{A\}}, y_0^{\{A\}}, z_0^{\{A\}})$ is the A particle rest frame (RF).

The relation between the plane-wave and the spherical-wave is then the following:

$$|p, \theta_1, \phi_1, \lambda_1, \lambda_2\rangle = \sum_{J,M} \sqrt{\frac{2J+1}{4\pi}} D^J_{M,\lambda_1-\lambda_2}(\phi_1, \theta_1, 0)|p, J, M, \lambda_1, \lambda_2\rangle. \tag{3.12}$$

This derivation is possible thanks to the invariance under rotations of the helicity. In fact, as done above for the case of one-particle state, the state $|p, 0, 0, \lambda_1, \lambda_2\rangle$, quantized along the $\hat{z}$ axis, is eigenstate of J_z with eigenvalue $\lambda_1 - \lambda_2$. Then this state is rotated to $|p, \theta_1, \phi_1, \lambda_1, \lambda_2\rangle$ with unchanged helicities.

3.2.3 Particle-2 definition

A problem has been found when Eq. (3.12) is used in the helicity formalism to incorporate different decay chains in the case of final particles with spin. Indeed, the state of particle-2, which is moving in the opposite direction with respect to particle-1 ($\vec{p_2} = -p_z$), is defined up to a phase convention, which needs to be used consistently throughout. The phase convention introduced by Jakob-Wick [103] defines the state vector of particle-2 from Eq. 3.8 for $\theta = \pi, \phi = 0$ by adding a specific phase $(-1)^{J_2-\lambda_2}$,

$$|\vec{p_2} = -p_z, J_2, \lambda_2\rangle' = (-1)^{J_2-\lambda_2} e^{-i\pi J_y} |\vec{p_2} = p_z, J_2, \lambda_2\rangle. \tag{3.13}$$

where the transformation $e^{-i\pi J_y}$ is a rotation of π around the y axis and bring the momentum p_z into $-p_z$. The choice of the phase factor $(-1)^{J_2-\lambda_2}$ is made in order to obtain

a state with opposite helicity in the limit of $p_z \to 0$:

$$|-p_z = 0, J_2, \lambda_2\rangle' = |p_z = 0, J_2, -\lambda_2\rangle, \tag{3.14}$$

which follows from the Wigner-d matrix relation, $d^J_{\lambda',\lambda}(\pi) = (-1)^{J-\lambda}\delta_{\lambda',-\lambda}$.

This convention, referred to as particle-2 convention, is usually exploited to extract the constraint due to parity (Eq. 3.23), as well as the relation between the helicity and the LS couplings (Eq. 3.16), as derived in Sec. 3.2.4.

Although this specific phase convention naturally matches the canonical two-particle states, it treats the two particles in the two body decay differently. This is a problem in a 3-body decay when different decay chains are added, and where the role of particle-2 is played by a different particle in each chain, *i.e.* a proton or a J/ψ. Therefore, an additional phase, equal to $(-1)^{J_2-\lambda_2}$, needs to be added for each two-body amplitude, in order to restore the particle-1 convention, as discussed in Sec. 3.3.

3.2.4 Two-particle states in the LS-coupling basis

The two-particle helicity states, $|p, J, M, \lambda_1, \lambda_2\rangle$, can be written in the simultaneous basis of LS, $|p, J, M, L, S\rangle$, where L and S denote the quantum numbers of the relative orbital angular momentum and the total spin of the two particles, respectively. This basis is useful for certain applications, *i.e.* when the lineshape of a resonance is explicitly dependent on the orbital momentum.

However, there was a lot of debate whether the LS representation holds for relativistic processes. Indeed, in the LS basis, one has to couple the total spin, which is the sum of the individual spins of the two particles and is only defined in the particles' rest frame, with the angular momentum, which is defined in the CM frame. As pointed out in Ref. [111], given the amplitude in a fixed system (*i.e.* the helicity representation) the values of orbital momentum L and total spin S are always defined without loosing relativistic properties. Indeed, all the amplitudes (helicity and LS amplitudes) are relativistic because they are well defined under Lorentz transformations.

Therefore, it is demonstrated that even in relativistic processes it is possible to relate the helicity couplings to the LS couplings, by making an orthogonal transformation between the canonical representation, used in the LS basis, and the helicity representation [104].

The relation between the helicity and the LS basis, derived in Ref. [104], is as follows:

$$\begin{aligned}|JM;\lambda_1,\lambda_2\rangle &= \sum_{L,S}\langle JM;LS|JM;\lambda_1,\lambda_2\rangle|JM;LS\rangle \\ &= \sum_{L,S}\left(\frac{2L+1}{2J+1}\right)^{1/2}\langle L,S,0,\lambda|J\lambda\rangle\langle s_1,s_2,\lambda_1,-\lambda_2|s\lambda\rangle|JM;LS\rangle\end{aligned} \tag{3.15}$$

where $\lambda = \lambda_1 - \lambda_2$ and the expressions inside the brackets are the Clebsch-Gordan coefficients. The transformation matrix between the two representations is therefore:

$$\langle JM;LS|JM;\lambda_1,\lambda_2\rangle = \left(\frac{2L+1}{2J+1}\right)^{1/2}\langle L,S,0,\lambda|J\lambda\rangle\langle s_1,s_2,\lambda_1,-\lambda_2|s\lambda\rangle. \tag{3.16}$$

This relation will be used in the derivation of the $B^0_s \to J/\psi p\overline{p}$ amplitude model to reduce the number of effective couplings, as described in Sec. 5.1.2.

3.3 Two-body helicity amplitudes and decay rate

Finally, the amplitude of a two body decay, $A \to 12$, can be obtained in the A rest frame, as the product through a transition operator $\hat{T}$ of the initial spin state of particle A, $|J_A, \lambda_A\rangle$, and a plane-wave helicity state defined in Eq. 3.12:

$$\mathcal{A}^{A\to 12}_{\lambda_A,\lambda_1,\lambda_2}(\theta_1,\phi_1) = (-1)^{J_2-\lambda_2}\langle p,\theta_1,\phi_1,\lambda_1,\lambda_2|\hat{T}|J_A,\lambda_A\rangle \tag{3.17}$$

where the amplitude is now defined in the particle-1 convention. Inserting a basis of the spherical-wave helicity states $|p, J, M, \lambda_1, \lambda_2\rangle$ the amplitude becomes:

$$\begin{aligned}
\mathcal{A}^{A\to 12}_{\lambda_A,\lambda_1,\lambda_2}(\theta_1,\phi_1) &= (-1)^{J_2-\lambda_2}\langle p,\theta_1,\phi_1,\lambda_1,\lambda_2|\hat{T}|J_A,\lambda_A\rangle \\
&= \sum_{J,M}(-1)^{J_2-\lambda_2}\langle p,\theta_1,\phi_1,\lambda_1,\lambda_2|p,J,M,\lambda_1,\lambda_2\rangle\langle p,J,M,\lambda_1,\lambda_2|\hat{T}|J_A,\lambda_A\rangle \\
&= \sum_{J,M}(-1)^{J_2-\lambda_2}\sqrt{\frac{2J+1}{4\pi}}D^{*J}_{M,\lambda_1-\lambda_2}(\phi_1,\theta_1,0)\delta_{J,J_A}\delta_{M,\lambda_A}\times \\
&\quad \langle p,J,M,\lambda_1,\lambda_2|\hat{T}|J_A,\lambda_A\rangle \\
&= (-1)^{J_2-\lambda_2}\mathcal{H}_{\lambda_1,\lambda_2}\cdot D^{*J_A}_{\lambda_A,\lambda_1-\lambda_2}(\phi_1,\theta_1,0)\cdot\mathcal{R}(m^2_{12}).
\end{aligned} \tag{3.18}$$

The three components in Eq. (3.18) are the following:

- the helicity couplings $\mathcal{H}_{\lambda_1,\lambda_2}$, defined as

$$\mathcal{H}_{\lambda_1,\lambda_2} = \langle p, J_A, \lambda_A, \lambda_1, \lambda_2|\hat{T}|J_A,\lambda_A\rangle, \tag{3.19}$$

  are complex numbers to be determined from the fit to data that depend on the kinematics of the decay. The number of helicity couplings derived from angular momentum conservation, knowing that:

$$|\lambda_1| \le s_1, \qquad |\lambda_2| \le s_2, \qquad |\lambda_1 - \lambda_2| \le J_A, \tag{3.20}$$

  are reduced for parity-conserving decays, following the relation shown in Ref. [103]:

$$\begin{aligned}
\mathcal{H}_{\lambda_1,\lambda_2} &\equiv \langle p, J_A, \lambda_A, \lambda_1, \lambda_2|\hat{T}|J_A,\lambda_A\rangle &(3.21)\\
&= \langle p, J_A, \lambda_A, \lambda_1, \lambda_2|P\hat{T}P|J_A,\lambda_A\rangle &(3.22)\\
&= \eta_A\eta_1\eta_2(-1)^{s_1+s_2-J_A}\mathcal{H}_{-\lambda_1,-\lambda_2} &(3.23)
\end{aligned}$$

  in which η are the intrinsic parities of the particles.

- $(-1)^{J_2-\lambda_2}$ is the phase factor that restores the particle-1 convention and will be incorporated in the definition of helicity as $\tilde{\mathcal{H}}_{\lambda_1,\lambda_2} = (-1)^{J_2-\lambda_2}\mathcal{H}_{\lambda_1,\lambda_2}$.

- Wigner D-matrix, containing the angular dependence and corresponding to the one in Eq. (3.12).

- $\mathcal{R}(m^2_{12})$: a possible lineshape of the resonance, which depends on the invariant mass squared of particle-1 and particle-2, m^2_{12}.

In multi-body decay processes, the decay amplitude can be decomposed in subsequent two-body decays mediated by intermediate particles, until the final state particles are reached. For example, for a three-body decay, a decay of a particle A happens through the sequence: $A \to R(\to 12)3$, where R decays to 12 particles. Following Eq. (3.18), the amplitude of the first two body decay, $A \to R3$, can be written as:

$$\mathcal{A}^{A\to R3}_{\lambda_A,\lambda_R,\lambda_3}(\theta_R,\phi_R) = \tilde{\mathcal{H}}_{\lambda_R,\lambda_3} \cdot D^{*J_A}_{\lambda_A,\lambda_R-\lambda_3}(\phi_R,\theta_R,0) \tag{3.24}$$

where the angles ϕ_R and θ_R are computed in the A reference frame as the direction of the R momentum $p_R^{\{A\}}$. The second decay is expressed as:

$$\mathcal{A}^{R\to 12}_{\lambda_R,\lambda_1,\lambda_2}(\theta_1,\phi_1) = \tilde{\mathcal{H}}_{\lambda_1,\lambda_2} \cdot D^{*J_R}_{\lambda_R,\lambda_1-\lambda_2}(\phi_1,\theta_1,0) \cdot \mathcal{R}(m_{12}^2), \tag{3.25}$$

where the angles ϕ_1, θ_1 are the azimuthal and polar angles of $\vec{p}_1^{\{R\}}$ in the R helicity frame. The lineshape for the resonance R, dependent to the invariant mass squared m_{12}^2, is also included. The total amplitude is then obtained as the sum over the allowed helicity values λ_R for R:

$$\begin{aligned}\mathcal{A}^{A\to R3\to 123}_{\lambda_A,\lambda_1,\lambda_2\lambda_3}(\theta_R,\phi_R,\theta_1,\phi_1) &= \sum_{\lambda_R} \mathcal{A}^{A\to R3}_{\lambda_A,\lambda_R,\lambda_3}(\theta_R,\phi_R) \times \mathcal{A}^{R\to 12}_{\lambda_R,\lambda_1,\lambda_2}(\theta_1,\phi_1) \\ &= \sum_{\lambda_R} \tilde{\mathcal{H}}_{\lambda_R,\lambda_3} \cdot d^{J_A}_{\lambda_A,\lambda_R-\lambda_3}(\theta_R) e^{-i\lambda_A\phi_R} \\ &\times \tilde{\mathcal{H}}_{\lambda_1,\lambda_2} \cdot d^{J_R}_{\lambda_R,\lambda_1-\lambda_2}(\theta_1) e^{-i\lambda_R\phi_1} \cdot \mathcal{R}(m_{12}^2)\end{aligned} \tag{3.26}$$

where the Wigner D-matrices are expressed as polar rotations with small Wigner d-matrices and azimuthal rotations with phase factors.

The amplitude can also occur through resonances that decay to different final states. We will refer to these amplitude as different decay chains.

After the work presented in Ref. [112], it is established that in order to guarantee a correct amplitude definition, the cyclic particle ordering must be ensured. Hence, the resonances that enter the process are defined as: $R = 12$, $S = 23$ and $U = 31$, where the first particle of the decay is the one with respect to which the angles are defined. This definition is mandatory in order to ensure a proper spin matching between different chains, as detailed in Sec. 3.3.1 and for our case study in Sec. 5.1.2.

Let's suppose that the resonance S is now decaying to 23 final states; the amplitude can be written as:

$$\mathcal{A}^{A\to 1S\to 123}_{\lambda_A,\lambda_1',\lambda_2',\lambda_3'}(\theta_S,\phi_S,\theta_2,\phi_2) = \sum_{\lambda_R} \mathcal{A}^{A\to 1S}_{\lambda_A,\lambda_S,\lambda_1'}(\theta_S,\phi_S) \times \mathcal{A}^{S\to 23}_{\lambda_S,\lambda_2',\lambda_3'}(\theta_2,\phi_2) \tag{3.27}$$

which has the same form of Eq. (3.26) but the final particle helicities, labeled λ_i', and the angles, θ_2, ϕ_2, are computed in different reference frames with respect to the amplitude of the R resonance. Indeed, the reference frame of the final particles are reached from a different resonance helicity frame, for instance the R or S helicity frame.

3.3.1 Spin rotation and total amplitude

When intermediate states in different decay chains contribute to the decay, the total amplitude is the incoherent sum of each decay chain amplitude over all allowed final state

helicities. However, as said above, the final state helicities are computed in different frames for each decay sequence. Therefore, a reference helicity frame for each final particle must be chosen, and all the different helicity frames must be rotated to the chosen one. The choice of the reference frame is arbitrary. For instance, it can be the one of the S decay sequence. Hence, the R intermediate state must be rotated to the helicity frames defined for S intermediate state, as:

$$|\lambda_i\rangle_S = \sum_{\lambda_i'} D^J_{\lambda_i',\lambda_i}(\phi_i^R, \theta_i^R, 0)|\lambda_i'\rangle_R \tag{3.28}$$

where the index i refers to all final particle with spin and ϕ_i, θ_i are the azimuthal and polar angles in the particle i rest frame between the boost directions of resonance S and resonance R. Eq. 3.28 can be simplified to a small Wigner-d matrix, $d^J_{\lambda_i',\lambda_i}(\theta_i^R)$, in the case of a three-body decay, where final particles lie on a plane, and when cyclic particle ordering is ensured. Indeed, the cyclic particle ordering guarantees that the spin states are defined with parallel y axes. Therefore, the azimuthal rotation cancels and only polar rotations are performed.

In Fig. 3.2, a sketch of the Wigner rotations is depicted for the case of particle 1. The figure visualizes the transformation of the momenta in a three body decay $A \to 123$ from the A rest frame to the p_1 rest frame. The rows correspond to the three decay sequences through the intermediate resonance $R \to 12$, $S \to 23$ and $U \to 31$, while the columns refer to the aligned rest frame of particle A[1] on the left, and to the p_1 rest frame on the right. In order to reach the p_1 frame from the A one, it is sufficient to boost along the direction of the p_1 momentum. As represented by the sketch on the right, in order to align the particle-1 helicity frame in the R chain to the S one, it is needed to rotate anti-clockwise by the angle θ_1^R around the y axis. The angle θ_1^R is obtained as a scalar product between the boost directions of resonance R and resonance S. Indeed, the former is parallel to particle 3 and the latter to particle A. The same can be applied for the third decay sequence of resonance U, where a clockwise rotation of an angle θ_1^U must be applied. In the following, for brevity, we incorporate the sign of the rotation inside the angle definition, $\theta_1^U = -\theta_1^U$. However, this sign is important to ensure a correct alignment and needs to be accounted for properly, in a similar way as done in Fig. 3.2.

Once the final helicity states are rotated, different decay chains can be summed up incoherently over the final state particles for unpolarized decays:

$$\begin{aligned}\mathcal{A}_{tot}(\Omega) = \sum_{\lambda_1,\lambda_2,\lambda_3} \Bigg(& \mathcal{A}^{A\to 1S\to 123}_{\lambda_A,\lambda_1,\lambda_2\lambda_3}(\Omega) \\ & + \sum_{\lambda_1'^R,\lambda_2'^R,\lambda_3'^R} d^{J_1}_{\lambda_1'^R,\lambda_1}(\theta_1^R) d^{J_2}_{\lambda_2'^R,\lambda_2}(\theta_2^R) d^{J_3}_{\lambda_3'^R,\lambda_3}(\theta_3^R) \mathcal{A}^{A\to R3\to 123}_{\lambda_A,\lambda_1,\lambda_2\lambda_3}(\Omega) \\ & + \sum_{\lambda_1'^U,\lambda_2'^U,\lambda_3'^U} d^{J_1}_{\lambda_1'^U,\lambda_1}(\theta_1^U) d^{J_2}_{\lambda_2'^U,\lambda_2}(\theta_2^U) d^{J_3}_{\lambda_3'^U,\lambda_3}(\theta_3^U) \mathcal{A}^{A\to U2\to 123}_{\lambda_A,\lambda_1,\lambda_2\lambda_3}(\Omega) \Bigg)\end{aligned} \tag{3.29}$$

It is worth noticing that the Euler angles of resonances in different decay chains can be obtained from the set of phase space variables Ω and hence do not constitute a set of independent variables.

Eq. (3.29) can be generalized for the case of subsequent decays of one final particle, *i.e.* particle 3 decays to other two particles, $3 \to 45$. In this case, the Wigner-d matrix of

[1]The aligned configuration corresponds to aligning the spectator particle of the decay to the negative z direction.

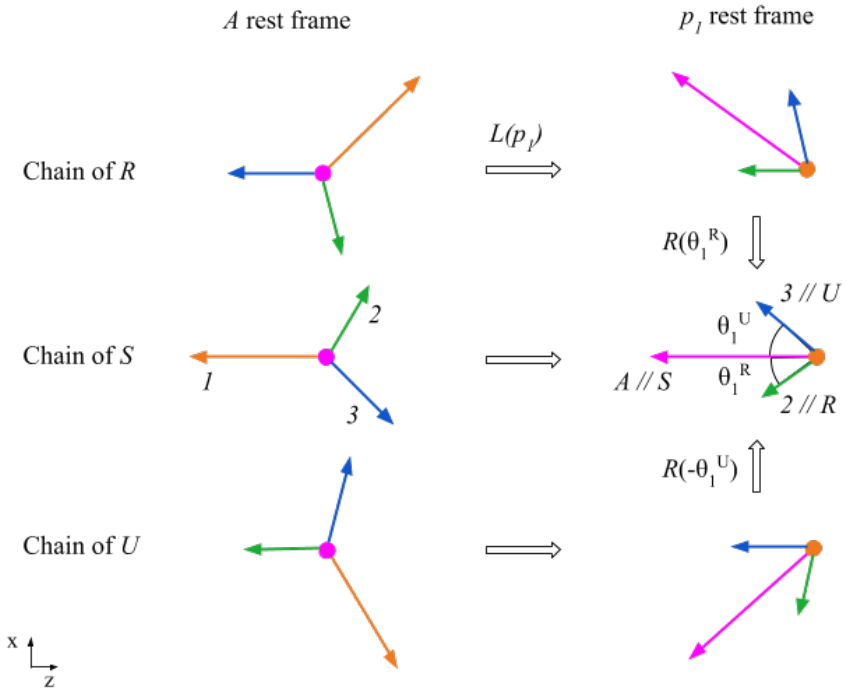


Figure 3.2 Transformation of the momentum vectors of a three body decay $A \rightarrow p_1(\text{orange})p_2(\text{green})p_3(\text{blue})$ in the A rest frame (left) and in the p_1 rest frame (right). The rows show the three decay sequences: $R(12)3, 1S(23), U(31)2$. Image inspired by Ref. [112].

particle 3 will be omitted and replaced by a rotation of the spin state of the final particles it decays into (where also azimuthal angles come into play), as will be detailed for the decay under study during this work.

3.3.2 Decay rate

The differential decay rate of an unpolarized particle A with mass M and 4-momentum $p = (E, \vec{p})$ is computed in its CM as:

$$d\Gamma \propto \frac{|\mathcal{A}_{tot}|^2}{2M} d\Phi, \tag{3.30}$$

where $\mathcal{A}_{tot}$ is the matrix element defined in Eq. (3.29), $2M$ is the normalization factor in the CM of A and $d\Phi$ is the phase space element for a three body decay defined as:

$$d\Phi = |p_1^*||p_3| dm_{12} d\Omega_1^* d\Omega_3, \tag{3.31}$$

where $(|p_1^*|, \Omega_1^*)$ is the momenta and the angle of particle 1 in the rest frame of particle 1 and 2, while $(|p_3|, \Omega_3)$ are the momenta and angles of particle 3 in the rest frame of the decaying particle.

CHAPTER 4

Observation of $B^0_{(s)} \to J/\psi\, p\bar{p}$ decays and precise mass measurements of $B^0_{(s)}$ mesons

In this Chapter, we report the first observation of $B^0_{(s)} \to J/\psi\, p\bar{p}$ decays using the data collected by LHCb between 2011 and 2016 and corresponding to a luminosity of 5.2 fb^{-1}. The results are published in Ref. [1].

The $B^0_{(s)} \to J/\psi\, p\bar{p}$ decays are suppressed due to limited available phase-space, as well as due to OZI or Cabibbo suppression, as shown in the Feynam diagrams in Fig. 1.12 and Fig. 1.13 for B^0 and B^0_s, respectively. These decays are particularly attractive for pentaquark and glueball searches, as described in the introduction in Sec. 1.5.2.

First attempts for observing these modes have been made by Belle [47] and LHCb [53] but no evidence have been found and only upper limits on their branching ratios have been placed. The naive theoretical expectation for the branching fraction, $\mathcal{B}(B^0_s \to J/\psi\, p\bar{p})$, is at the level of 10^{-9} [2], while the previous LHCb measurement estimates it to be of the order of 10^{-7}. An enhancement in the branching ratios could be explained by the presence of exotic resonances, and therefore it is interesting to measure them.

Finally, the limited phase space, which characterizes this decay, allows us to perform the most precise $B^0_{(s)}$ mass measurements.

The Chapter describes in detail all the steps needed for performing a branching ratio measurement, from the selection of the signal candidates (Sec. 4.1) to the efficiency estimation (Sec. 4.4). The corrections applied to Monte Carlo simulations to be representative of real data are also described in Sec. 4.2, followed by the optimization of the selection strategy which exploits multivariate techniques (Sec. 4.3). The invariant mass fit is then performed in Sec. 4.5. The results on the branching ratios and the associated systematic uncertainties are reported in Sec. 4.6 and 4.7, while the mass measurement is given in Sec. 4.8.

4.1 Data sample and event selection

This analysis is based on the 3 fb^{-1} data collected in Run 1 during 2011-2012 at the collision energy of 7 TeV and 8 TeV, respectively, and 2 fb^{-1} data collected in Run 2 during 2015-2016 at the energy of 13 TeV, as depicted in Fig. 2.18. Although the Run 2 integrated luminosity is lower than in Run 1, the higher $\sqrt{s}$ meant that the cross-section is about doubled with respect to Run 1 (Eq. 2.3), so that the signal yield is expected to increase more than twice with the addition of Run 2 data. The run conditions, *e.g.* trigger, background level, are sufficiently different between Runs 1 and 2 so that the datasets are processed independently.

Table 4.1 Monte Carlo samples used in Run 1 and Run 2. See also Tables 4.12, 4.13 and 4.14.

Decay	Event-Type	Statistics	Data Conditions
$B_s^0 \to J/\psi p\overline{p}$	13144061	7.5 M	Run I
$B^0 \to J/\psi p\overline{p}$	11144120	7.5 M	Run I
$B_s^0 \to J/\psi \phi$	13144001	2 M	Run I
$B_s^0 \to J/\psi p\overline{p}$	13144061	7.5 M	Run II
$B^0 \to J/\psi p\overline{p}$	11144120	7.5 M	Run II
$B_s^0 \to J/\psi \phi$	13144011	10 M	Run II

For efficiency calculations, Monte Carlo (MC) samples (see Sec. 2.3) were generated for the $B^0_{(s)} \to J/\psi p\overline{p}$ signal and the $B^0_s \to J/\psi \phi(\to K^+K^-)$ control mode. The sample of $B^0_s \to J/\psi \phi(\to K^+K^-)$ is needed as normalization sample for the branching ratio measurement and for MC corrections. This channel was preferred with respect $B^0_s \to J/\psi \pi^+\pi^-$ used in the previous LHCb analysis [53] for two reasons: $B^0_s \to J/\psi \phi$ is a high statistics and well-studied mode at LHCb, and the kaon mass is closer to the proton than the much lighter pion. The number of events generated after the stripping line and trigger requirements are listed in Table 4.1 for Run 1 and Run 2 (see also Sec. 4.4.1).

4.1.1 Signal characterization and pre-selection

The decay of $B^0_{(s)} \to J/\psi p\overline{p}$ is reconstructed by exploiting specific properties of the decay concerning kinematic and topology. First of all, the B meson is a particle which decays on flight due to its long lifetime of $\sim$ ps and the large boost felt in the LHCb detector. As a consequence, the secondary vertex (of the B decay) is detached from the primary vertex (PV), which is the beam interaction point, by order of 10 mm-1 cm. The separation between vertices is reflected in a large value of the impact parameter (IP) of the final tracks with respect to the PV, defined as the minimum distance between the extrapolated track and the PV. Other important B variables concern the kinematic of the decay, like the total and the tranverse momentum, p and p_T, and the topology: all final tracks may come from the B secondary vertex, which corresponds to a low χ^2 of the PV fit. Secondly, in order to characterize the decay mode some variables regarding the final state have to be accounted for. The final state is composed of four particles, two μ and two protons, which are well-identified in the detector. Their impact parameters and the momenta are important discriminating variables which identify the topology. In addition, the information of the particle identification is essential for the rejection of background as explained in Sec. 4.2.2.

Since these decays have not been observed yet, no dedicated *stripping* line exists, which is a set of selection requirements as defined in Sec. 2.3. Therefore, we combined two muon tracks selected by the `StrippingFullDSTDiMuonJpsi2MuMuDetached` line, with two hadrons tracks which are loosely selected by `StdAllLooseProtons` and `StdAllLooseKaons` for protons and kaons, respectively. In the following, a description of the preliminary selection for building up the $B^0_{(s)} \to J/\psi p\overline{p}$ signal and the $B^0_s \to J/\psi \phi$ control mode candidates is reported.

The signal candidates are built from two charged tracks identified as muons and coming from the J/ψ candidates, selected with filtering cuts described in Table 4.2. At

Table 4.2 Selection requirements for the `StrippingFullDSTDiMuonJpsi2MuMu Detached` line. The selection cuts are associated to the particle listed in the left column. Here some variable definition: Track χ^2/ndf is computed from the track fit done at event reconstruction stage; DLL(h) is the delta log-likelihood between h and pion identities obtained from particle identification systems; DOCA is the distance of closest approach to the vertex between a pair of tracks and the χ^2-distance is the difference between two histograms; the decay length significance is the decay lenght divided by its uncertainty.

Particle	Variable Definition	Selection Requirement
J/ψ	Mass difference wrt nominal J/ψ mass	$\left\|m_{\mu\mu} - m_{J/\psi}\right\| < 100$ MeV
	DOCA χ^2-distance of daughters	< 30
	χ^2 of J/ψ vertex fit	$\chi^2_{J/\psi\ \mathrm{vtx}}/\mathrm{ndof} < 20$
	J/ψ decay length significance	> 3
μ	p_T	$p_\mathrm{T} > 550(500)$ MeV
	track fit χ^2 per degree of freedom	$\chi^2_\mathrm{trk}/\mathrm{ndof} < 5$
	μ PID	$\mathrm{DLL}_{\mu\pi} > 0$

the next stage, the χ^2 of J/ψ vertex fit was tightened to be < 20, and a mass-fit (`MFIT`) constraint was imposed on the J/ψ to improve the μ momentum resolution.

In the second stage, another pair of charged tracks consistent with p and $\overline{p}$, with the selections as in Table 4.3, is combined with the di-muon tracks to form a $B^0_{(s)}$ candidate with the entire decay topology constrained by a kinematic fit `DecayTreeFit(DTF)` [113]. The χ^2 probability from this fit is required to be $> 10^{-6}$ for the selected candidates. The motivation behind this loose cut is to trim out only non-convergent fits at this stage. The actual optimization of the `DTF` cut is performed through the multivariate selection (MVA) by means of a boosted decision tree alghoritm, as shown in Sec. 4.3. The dimuon invariant mass is constrained to be equal to the nominal J/ψ mass [114]. A cut on the proton identification neural-network based PROBNN($p \to p$) variables is also placed to be larger than 0.2. This variable is the output of a neural-network classifier [115] (see also Sec. 4.2.2) and represents the probability for a proton to be identified as proton. The distributions for each PID variable is resampled using calibration samples extracted directly from data, as explain in Sec. 4.2.2.

For the $B^0_s \to J/\psi\phi$ control mode a similar procedure is followed, replacing the $p\overline{p}$ combination with a pair of charged tracks associated to K^+K^- candidates, satisfying $|m(K^+K^-) - 1020| < 5$ MeV, to select the $\phi(1020)$ resonance region [114]. All charged tracks are required to be of good quality, with $p_\mathrm{T} > 300$ MeV ($p_\mathrm{T} > 550$ MeV) for K or μ. Additional vetoes are placed for specific backgrounds from $B^0 \to J/\psi K^+\pi^-$, as listed in Table 4.5.

At this stage, the background is predominantly made up by combinatorial tracks, composed by other hadrons (π, K and p) misidentified as protons. Background studies are detailed in Sec. 4.1.4.

The selection will be optimized by means of a multivariate analysis as described in Sec. 4.3.

Table 4.3 Selection requirements for the proton tracks and the $p\bar{p}$ candidate. The cuts on p and p_{T} are in accordance with the proton PID calibration samples. Ghost probability refers to the probability of the track to be a fake one; `HASRICH` is true when RICH subsystem registers a track in the candidate event.

Particle	Variable Definition	Selection Requirement
$p/\bar{p}$	Track fit χ^2 per degree of freedom	$\chi^2_{\mathrm{trk}}/\mathrm{ndof} < 5$
	Track ghost probability	< 0.6
	`HASRICH`	`True`
	PID(p)	$\mathrm{DLL}_{pK}(p) > -5$, $\mathrm{DLL}_{p\pi}(p) > -5$
	ProbNN($p \to p$)	$\mathrm{ProbNN}_p > 0.2$
$p\bar{p}$	p_{T}	$p_{\mathrm{T}} > 300$ MeV
	p	$p > 6$ GeV
	p of $p\bar{p}$ candidate (Run 2)	$p < 100$ GeV
	χ^2 of $p\bar{p}$ vertex fit	$\chi^2_{p\bar{p}\,\mathrm{vtx}}/\mathrm{ndof} < 25$
B	χ^2 of B vertex fit	$\chi^2_{B\,\mathrm{vtx}}/\mathrm{ndof} < 30$
	χ^2 probability of B vertex fit	`DTF_prob`$> 10^{-6}$

Table 4.4 Selection requirements for the kaon tracks and the K^+K^- candidate.

Particle	Variable Definition	Selection Requirement
K	Maximum track fit χ^2 per degree of freedom	$\chi^2_{\mathrm{trk}}/\mathrm{ndof} < 5$
	Track ghost probability	< 0.6
	`HASRICH`	`True`
	PID	$\mathrm{DLL}_{K\pi}(K) > -5$
	ProbNN($K \to K$)	$\mathrm{ProbNN}_K > 0.2$
K^+K^-	Mass difference wrt nominal ϕ mean mass	$\lvert m_{KK} - 1020\rvert < 5$ MeV
	p_{T}	$p_{\mathrm{T}} > 300\ MeV$
	p	$p > 6\ GeV$
	χ^2 of K^+K^- vertex fit	$\chi^2_{KK\,\mathrm{vtx}}/\mathrm{ndof} < 25$
B	χ^2 of B vertex fit	$\chi^2_{B\,\mathrm{vtx}}/\mathrm{ndof} < 30$

Table 4.5 For the specific background from $B^0 \to J/\psi K^+\pi^-$ for the control mode, after a $K \to \pi$ PID substitution, candidates fulfilling the the listed requirements are vetoed. Both $K^{\pm}$ combinations are checked.

Variable Definition	Background Requirement
Mass difference wrt nominal B^0 mass	$\lvert m_{B^0} - 5280\rvert < 22$ MeV
kaon PID	$\mathrm{DLL}_{K\pi}(K) < 1$
χ^2 probability of B vertex fit	`DTF_prob`$> 10^{-6}$

Table 4.6 Trigger requirements for this analysis – everything is `TOS` on the B head, except for `L0Global`.

Stage	Run I Triggers	Run II Triggers
L0	`L0Muon`, or `L0DiMuon`, or	`L0Muon`, or `L0DiMuon`
HLT 1	`Hlt1TrackAllL0`, or `Hlt1TrackMuon`, or `Hlt1DiMuonHighMass`, or `Hlt1DiMuonLowMass`	`Hlt1DiMuonHighMass`, or `Hlt1DiMuonLowMass`, or `Hlt1SingleMuonHighPT`, or `Hlt1TrackMuon` or `Hlt1TrackMVA`
HLT 2	`Hlt2Topo[2,3,4]BodyBBDT`, or `Hlt2TopoMu[2,3,4]BodyBBDT`, or `Hlt2SingleMuon`, or `Hlt2DiMuonDetached`, or `Hlt2DiMuonDetachedJPsi`	`Hlt2DiMuonJPsi`, or `Hlt2DiMuonDetachedJPsi` or `Hlt2TopoMu[2,3,4]Body`, or `Hlt2Topo[2,3,4]Body`, or `Hlt2SingleMuon`, or `Hlt2DiMuonDetached`, or `Hlt2DiMuonDetachedJPsi`

4.1.2 Trigger selections

Candidates are required to pass the trigger requirements listed in Table 4.6. At each stage the candidates are required to be "triggered on signal" (`TOS`), meaning the trigger decision is due to the candidate. The full trigger sequence comprised "`L0 && HLT1 && HLT2`", as described in Sec. 2.2.4. The trigger is predominantly based on muon lines, the most important are: `L0Muon/L0DiMuon`, `Hlt1TrackMuon` and `hlt2DiMuonDetachedJPsi`. The `L0Muon` and `L0DiMuon` hardware trigger lines require the event to contain one or two muons with sufficiently high p_T; the `Hlt1TrackMuon` trigger is devised for selecting heavy hadrons decaying to at least one muon by exploiting the muon p_T and χ^2 impact parameter with respect to the primary vertex; `Hlt2DiMuonDetachedJPsi` selects events with a J/ψ detached from the primary vertex. In addition, a trigger on the B decays into charged hadrons is considered, with the topological lines `Hlt2Topo[2,3,4]Body` in an inclusive selection based on two- three- and four-prong vertices. These trigger lines are based on a Boosted Decision Tree (BDT) classifier [116].

4.1.3 Signal fit model

In all modes, i.e. in both the signal channels and the normalisation channel, the signal distributions of the $B^0_{(s)}$ candidate invariant mass are modelled by two Crystal Ball (CB) PDFs functions [117], which are made of a Gaussian connected with a power-law tail in such a way to have continuous derivative,

$$CB(t|\sigma,\alpha,n) = N \begin{cases} \exp(-\frac{t^2}{2\sigma^2}) & \text{if} \quad t/\sigma > -\alpha \\ (\frac{n}{|\alpha|})^n \exp(-\frac{\alpha^2}{2}(\frac{n-\alpha^2}{|\alpha|} - \frac{t}{\sigma})) & \text{if} \quad t/\sigma \leq -\alpha, \end{cases} \tag{4.1}$$

where $t = m - \mu$ is the difference between the reconstructed mass m and the mean of the Gaussian μ, σ is the width of the Gaussian and N is the normalisation factor. The parameters α and n determine the behaviour of the power law tail. The sign of α, for example, determines on which side of the Gaussian the tail appears. The freedom in the behaviour of the two tails makes this function the most appropriate shape to describe the

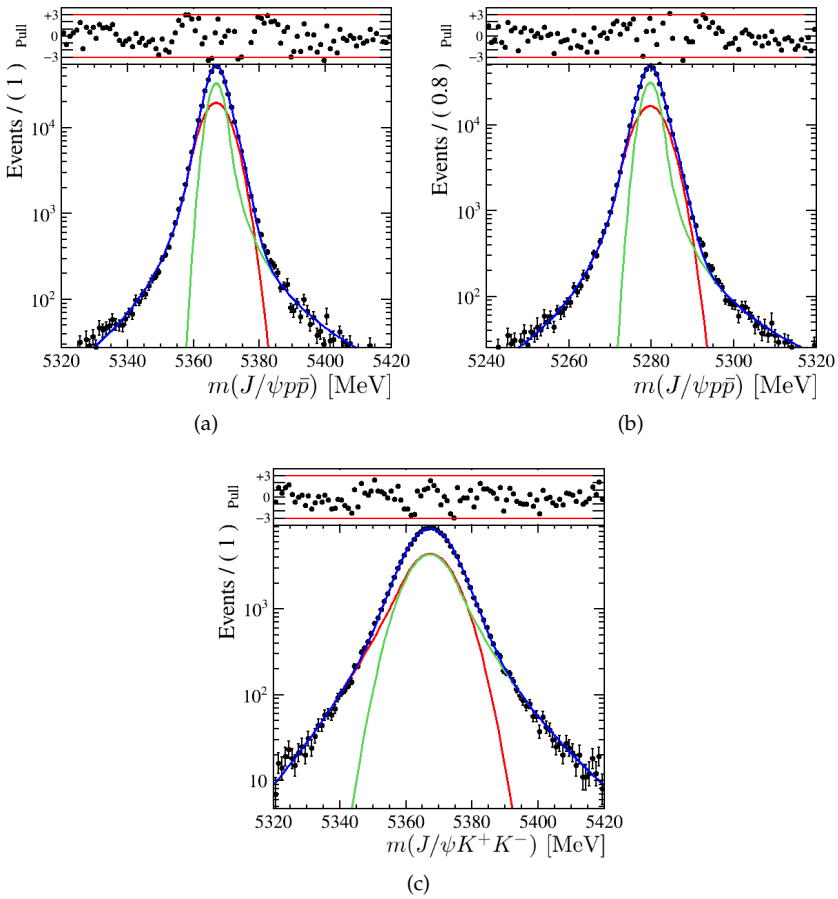


Figure 4.1 Run 1 signal lineshapes from invariant mass fits to truth-matched MC: (a) $B_s^0 \to J/\psi\, p\bar{p}$ (b) $B^0 \to J/\psi\, p\bar{p}$ (c) $B_s^0 \to J/\psi\, \phi$.

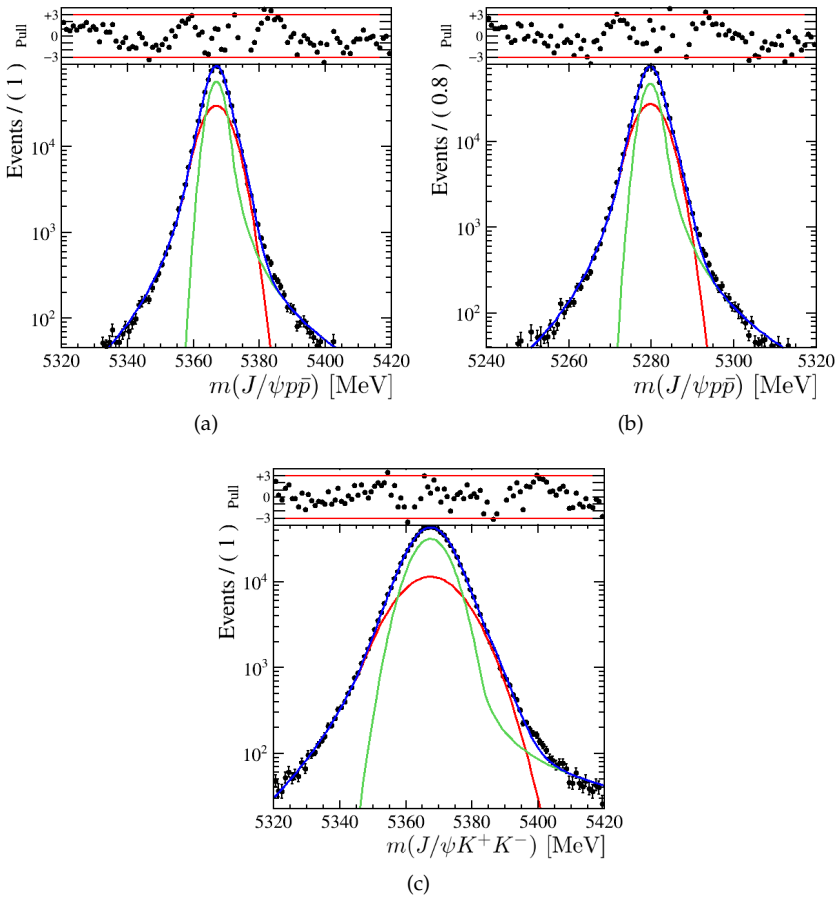


Figure 4.2 Run 2 signal lineshapes from invariant mass fits to truth-matched MC: (a) $B^0_s \to J/\psi\, p\bar{p}$ (b) $B^0 \to J/\psi\, p\bar{p}$ (c) $B^0_s \to J/\psi\, \phi$.

Table 4.7 Fit results for signal lineshapes in $B^0_{(s)} \to J/\psi\, p\overline{p}$ MC, after stripping line, trigger and PID requirements in Run 1 and Run 2.

Floating parameters	Run I	Run II
μ [MeV]	5279.66 $\pm$ 0.01	5279.66 $\pm$ 0.01
$\alpha_R^{B_d}$	2.03 $\pm$ 0.03	1.90 $\pm$ 0.02
$\alpha_L^{B_d}$	-1.84 $\pm$ 0.05	-1.65 $\pm$ 0.06
$n_R^{B_d}$	2.25 $\pm$ 0.08	2.81 $\pm$ 0.07
$n_L^{B_d}$	1.91 $\pm$ 0.07	2.28 $\pm$ 0.09
$\sigma_R^{B_d}$ [MeV]	3.71 $\pm$ 0.05	3.76 $\pm$ 0.05
$\sigma_L^{B_d}$ [MeV]	1.98 $\pm$ 0.03	2.08 $\pm$ 0.03
$frac$	0.49 $\pm$ 0.02	0.52 $\pm$ 0.02
μ [MeV]	5366.89 $\pm$ 0.01	5366.91 $\pm$ 0.02
$\alpha_R^{B_s}$	2.07 $\pm$ 0.02	1.98 $\pm$ 0.02
$\alpha_L^{B_s}$	-1.95 $\pm$ 0.04	-1.91 $\pm$ 0.05
$n_R^{B_s}$	2.17$\pm$ 0.08	2.57 $\pm$ 0.08
$n_L^{B_s}$	1.78 $\pm$ 0.05	1.91 $\pm$ 0.06
$\sigma_R^{B_s}$ [MeV]	4.45 $\pm$ 0.05	4.51 $\pm$ 0.06
$\sigma_L^{B_s}$ [MeV]	2.34 $\pm$ 0.03	2.48 $\pm$ 0.06
$frac$	0.47 $\pm$ 0.01	0.48 $\pm$ 0.01

radiative and misreconstruction effects present in these decays. In our model we sum two CB functions with tails on opposite sides, hence we use the nomenclature $\sigma_L^{B_i}$, $\alpha_L^{B_i}$ and $n_L^{B_i}$ for the parameters of the CB function with the tail on the left and $\sigma_R^{B_i}$, $\alpha_R^{B_i}$ and $n_R^{B_i}$ for that with the tail on the right, where B_i stands for the $B^0_{(s)}$ type. The parameter $frac$ is the fraction of events in one CB with respect to the other. We constrain the means of the two CB functions to be the same.

The fitted MC mass distributions of $B^0_{(s)} \to J/\psi\, p\overline{p}$ and $B^0_s \to J/\psi\,\phi$ are shown in Fig. 4.1 and 4.2 for Run 1 and Run 2, respectively. The two CB are in well-agreement with the shapes in the whole fit range. The results of the fit are reported in Tab. 4.7 and Tab. 4.8 for both Run 1 and Run 2.

The α and n factors responsible for the tail are extracted from the MC distribution and fixed for the fits to data. In the fit, the mean is left floating in the case of a significant excess and the sigma is fixed to the MC value instead. Since small variations are observed in the mean and width parameters between the fits to the $B^0_s \to J/\psi\,\phi$ MC and data, we decide to assign a systematic uncertainty to account for this effect. Therefore, toy MC studies are performed in order to investigate the difference from a fit where the sigma is scaled by the ratio of data/MC values found in the $B^0_s \to J/\psi\,\phi$ control channel, as detailed in Sec. 4.7.2.

Since the $B^0_s \to J/\psi\,\phi$ control sample is already clean after the preselection and trigger cuts, the mass fit can be performed, as reported in Fig. 4.3 for Run 1 and Run 2 data. The $B^0_s \to J/\psi\,\phi$ signal yield amounts to 62461 ± 306 events for Run 1 and 81153 ± 421 events for Run 2. This fit will be used to perform the reweighting of the MC distributions (see Sec. 4.2.1), after making a separation between signal and background components with the `sPlot` statistical technique, described in Ref. [118]. The background-subtracted data

Table 4.8 Fit results for signal lineshapes in $B^0_s \to J/\psi\,\phi$ from MC, after stripping line, trigger and PID requirements in Run I and Run II.

Floating parameters	Run I	Run II
μ [MeV]	5367.20 $\pm$ 0.05	5367.41 $\pm$ 0.04
α_R	2.0 $\pm$ 0.1	1.90 $\pm$ 0.05
α_L	-1.9 $\pm$ 0.2	-2.00 $\pm$ 0.03
n_R	4.3 $\pm$ 0.9	5.7 $\pm$ 0.2
n_L	2.4 $\pm$ 0.4	5.4 $\pm$ 0.1
σ_R [MeV]	8.2 $\pm$ 0.3	5.55 $\pm$ 0.05
σ_L [MeV]	4.9 $\pm$ 0.1	8.73 $\pm$ 0.13
$frac$	0.55 $\pm$ 0.05	0.57 $\pm$ 0.02

will be referred to as sWeighted data throughout.

4.1.4 Background study

The presence of physical backgrounds, coming from other decays that can peak below the $B^0_{(s)}$ invariant mass peaks, is investigated in this section. The main potential sources come from:

- $B^0_{d,s}(\Lambda^0_b) \to J/\psi\, hh'$ decays. These correspond to cross-feed contributions from similar modes, after a misidentification of hadron tracks, *i.e.* $p \leftrightarrow K$ or $p \leftrightarrow \pi$. They are obtained by reconstructing the invariant mass of the proton daughter particles with different mass hypothesis. In particular, the main candidates come from the $B^0_{d,s} \to J/\psi\, K^+\pi^-$, $B^0_s \to J/\psi\, K^+K^-$, $\Lambda^0_b \to J/\psi\, p\pi^-$ and $\Lambda^0_b \to J/\psi\, pK^-$ decays. These specific backgrounds do not peak in the corresponding $B^0_{(s)}$ or Λ^0_b signal region, as shown in Fig. 4.4 using the Run 1 data after the preselection cuts described in Sec. 4.1.1. There is only a slight visible peak in Figs. 4.4(a) around 4700 MeV due to reflection from the signal $B^0_s \to J/\psi\, p\bar{p}$ peak. This hypothesis has been validated on the signal $B^0_s \to J/\psi\, p\bar{p}$ MC by plotting different invariant masses after the corresponding mass-substitutions, as shown in Fig. 4.6, confirming that the peak is a reflection from $B^0_s \to J/\psi\, p\bar{p}$. The background components are also validated after the application of the full selection (comprised of the selection optimization described in Sec. 4.3): no different features are seen, as shown in Fig. 4.5.

- $H_b \to H^{\pm,0}h[h']$ decays. They include the most similar contributions without the J/ψ in the final states. The canonical example for that are the $B_{d,s} \to K^0_S p\bar{p}$ decays, where the two pions from the K^0_S are misidentified as muons. In Fig. 4.4(d), it is clear that this potential background can be considered as combinatorial.

- hadron double mis-ID's from exotics. Along with the narrow $\psi(2S), X(3782) \to J/\psi\,\pi^+\pi^-$, there are broad exotic states that decay to either $X \to J/\psi\,\pi^+\pi^-$ or $X \to J/\psi\, K^+K^-$. If the di-pions or di-kaons are mis-ID-ed as a di-proton, these can still survive for the $B^0_{(s)} \to J/\psi\, p\bar{p}$ mode. However, these exotics arise from partially reconstructed $B^+ \to XK^+$ decays, with the K^+ as missing. Since the B^+ mass is very close to the B^0 mass, these will show up partially reconstructed components below the B^0, but will not peak under the narrow $B^0_{(s)} \to J/\psi\, p\bar{p}$ signal

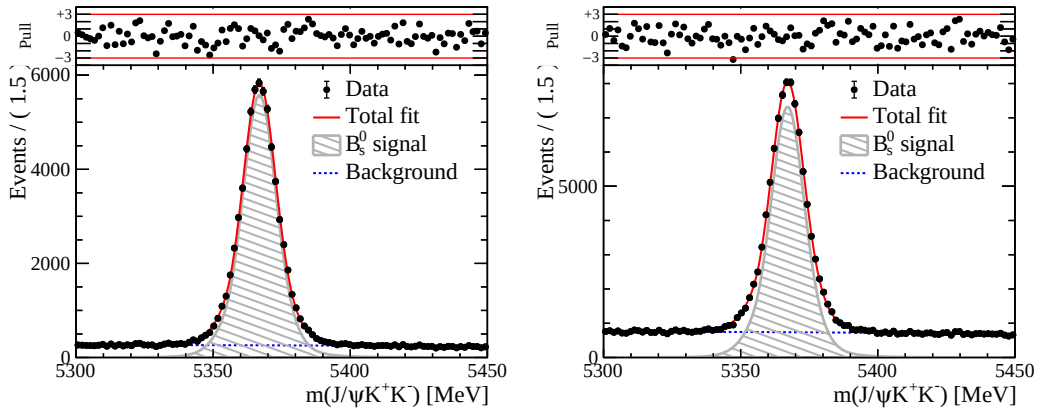


Figure 4.3 Fit to the m($J/\psi\, K^+K^-$) invariant mass for (a) $B_s^0 \to J/\psi\,\phi$ Run I and (b) Run II (only 2016) signal candidates retained after the pre-selection and trigger cuts, prior to any BDT cut. From the fit are extracted the $sWeights$ used to plot the distributions of signal and background variables.

regions. A 2-d scatter plot of $m(J/\psi\, KK)$ vs. $m(KK)$ is shown in Fig. 4.7 for the $B^0_{(s)} \to J/\psi\, p\bar{p}$ data (a) and signal MC (b). The narrow $\phi(1020)$ can be clearly seen in data, while is not present in signal MC. For this reason, we can conclude that it most probably coincides with the peak in $m(J/\psi\, KK) \sim 4700$ MeV, associated with the broad $X(4700) \to J/\psi\,\phi$ resonance, as seen in a previous LHCb analysis [119]. The $m(J/\psi\, p\bar{p})$ spectrum corresponding to a narrow band around the ϕ region is consistent with a flat distribution, as shown in Fig. 4.8. This background component is not removed, and is parameterized as combinatorial background.

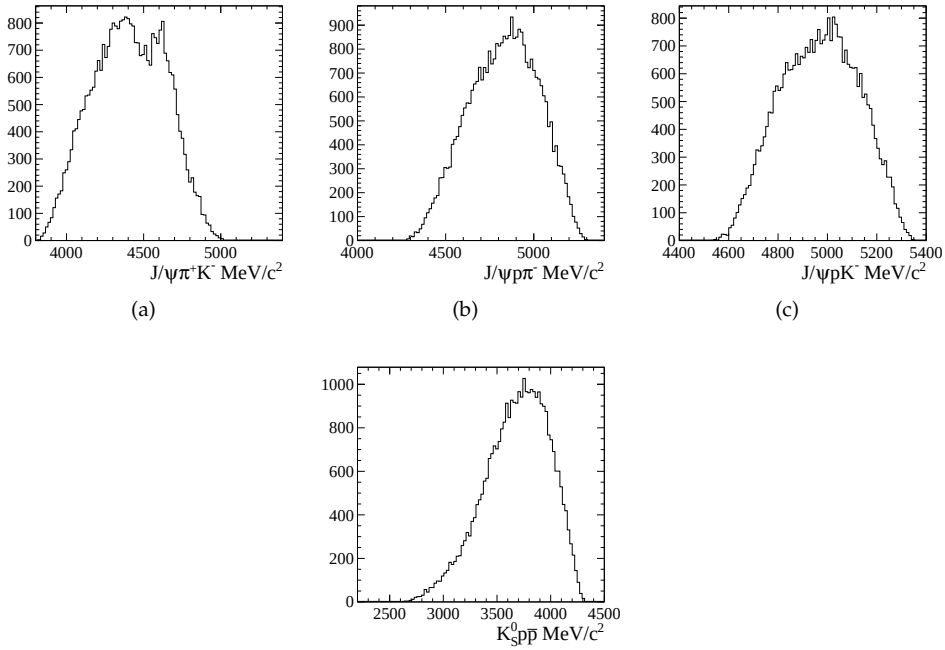


Figure 4.4 Invariant mass distributions reconstructed with different mass hypotheses for the final particles: (a) $B^0_{d,s} \to J/\psi K^+\pi^-$, (b) $\Lambda^0_b \to J/\psi p\pi^-$, (c) $\Lambda^0_b \to J/\psi pK^-$ and (d) $B_{d,s} \to K^0_S p\bar{p}$ decays studied from Run 1 data after the preselection cuts described in Sec. 4.1.1.

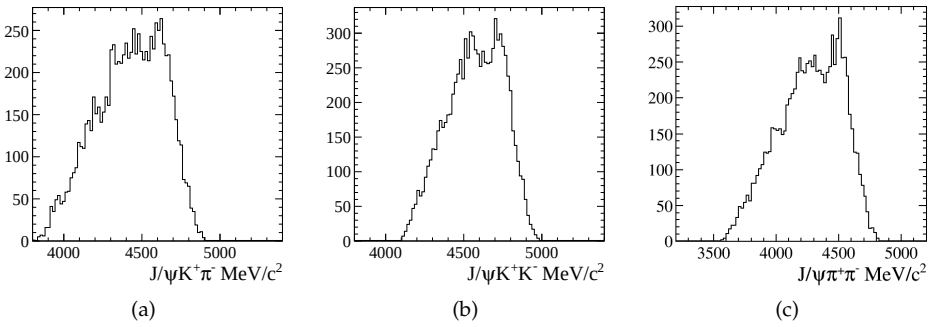
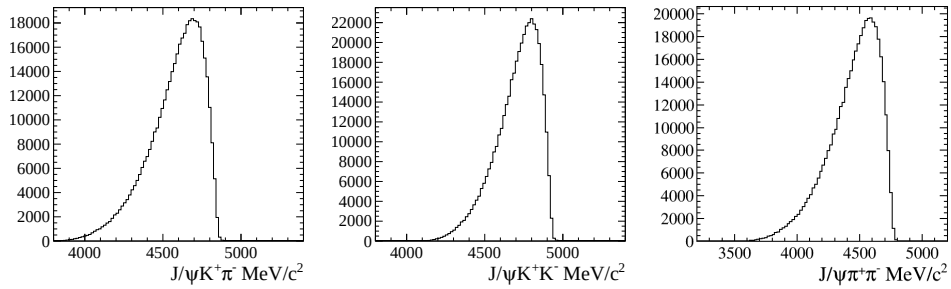
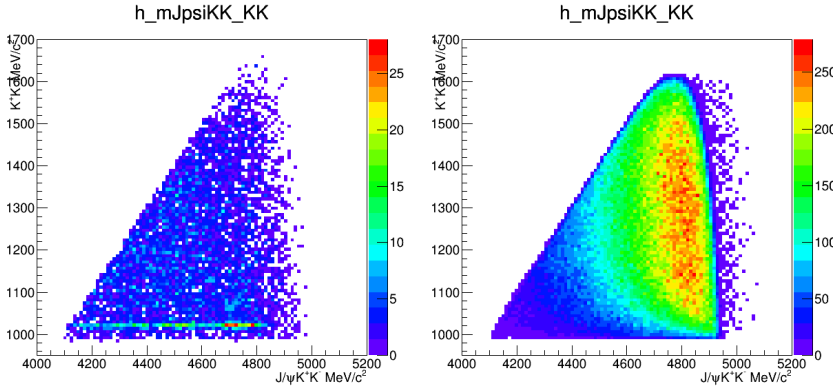


Figure 4.5 Invariant mass distributions reconstructed with different mass hypotheses for the final particles: (a) $B^0_{d,s} \to J/\psi K^+\pi^-$, (b) $B^0_{d,s} \to J/\psi K^+K^-$, (c) $B^0_{d,s} \to J/\psi \pi^+\pi^-$ decays, studied from Run 1 data after the entire selection, comprised the preselection and BDT optimization as in Sec. 4.3.



(a) (b) (c)

Figure 4.6 Invariant mass distributions reconstructed with different mass hypotheses for the final particles: (a) $B^0_{d,s} \to J/\psi K^+\pi^-$, (b) $B^0_{d,s} \to J/\psi K^+K^-$, (c) $B^0_{d,s} \to J/\psi \pi^+\pi^-$decays studied from Run 1 MC samples.



(a) (b)

Figure 4.7 Two-dimensional scatter plots of $m(J/\psi K^-K^+)$ vs $m(K^+K^-)$ invariant masses in data (a) and in Monte Carlo (b), where the kaon mass hypothesis is assigned to the proton and antiproton candidates. The data shows an horizontal band corresponding to the $\phi(1020)$.

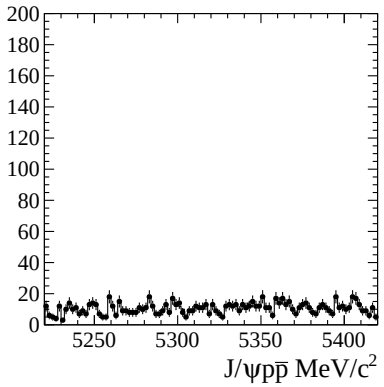


Figure 4.8 The $m(J/\psi p\bar{p})$ invariant mass distribution for signal candidates selected in correspondence of the ϕ region ($|m(K^+K^-)-1020| < 10$), where $m(K^+K^-)$ is evaluated by assigning the kaon mass hypothesis to the proton and antiproton candidates.

4.2 Data-MC comparison

4.2.1 Reweighting the kinematic variables

In general simulated data do not perfectly reproduce real data. Usually the discrepancy arises from limited knowledge of the underlying physics processes and detector mismodeling but can be also related to other issues. In order to be able to use simulated samples, corrections must be applied to make distributions of the observable variables agree better between MC and data.

Reweighting MC distributions is the general procedure to reduce the disagreement between MC and data. The problem is equivalent to estimating the density ratio between the PDFs of data and MC, $f_{data}(x)/f_{MC}(x)$, as a function of the variables participating in reweighting. Typically, the reweighting is performed in a binned fashion. For each bin, the weight is estimated as the ratio of the number of events in data, w^i_{data}, and in MC, w^i_{MC}:

$$w_i = \frac{w^i_{data}}{w^i_{MC}}. \tag{4.2}$$

Each MC event will then be multiplied by this weight according to the bin in which the event is located. However, the binned approach has strong limitations due to the amount of data needed to perform a reliable reweighting, which increases exponentially with the number of variables. This problem is referred to as "curse of dimensionality".

To avoid the binned approach, many alternative methods based on multivariate analysis (MVA)[1] have been proposed. Here, we exploit a method, called gradient boosting reweighting (GB-reweighting), as implemented in the `hep_ml` package[2] [120]. This procedure makes use of boosted decision trees (BDT) to split the space (of the relevant variables being reweighted) into a few large regions in correspondence with the problem. After the choice of the regions, the reweighting follows the same steps as in the binned technique, the distinctions being that the bins are selected differently. Before explaining the main steps of this method, let's recall the basic definition of a decision tree. A decision tree is a machine learning algorithm that splits the space in regions by applying simple selection cuts at each node of the tree. In practise, these regions are selected by checking simple conditions, such as $feature_i > threshold$. Then, it computes a prediction associated to the terminal node, called leaf. If the prediction is a continuous variable, then the tree is called regression tree, while, if it is discrete, classification tree.

In the present context, we use decision trees to divide the space of the selected variables into a few large regions (denoted by leaf of the BDT) in which the difference between data and MC is high.

To build the tree and find the regions (bins) important for reweighting, the following symmetrized χ^2 is maximized:

$$\chi^2 = \sum_{bin} \frac{(w_{bin,MC} - w_{bin,data})^2}{w_{bin,MC} + w_{bin,data}}. \tag{4.3}$$

In so doing, the $thresholds$ of the BDT dividing the different regions are obtained as the ones for which the discrepancy between the number of events in data and MC is

[1] Multivariate Analysis refers to any statistical technique used to analyze data that arises from more than one variable.

[2] https://github.com/arogozhnikov/hep_ml

large. Therefore, no reweighting will be applied to the regions in which the data and MC weights are equal. On the other hand, reweighting will be necessary for regions whose summand in Eq. 4.3 becomes high.

Once the regions have been identified by the BDT, weights equal to the density ratio $w_{leaf,MC}/w_{leaf,data}$ are assigned to MC events, depending on the leaf the events belong to. This procedure will be iterated as many times as the number of trees specifying as options in the boosting.

For the $B^0_s \to J/\psi\,\phi$ control mode, comparisons between $B^0_s \to J/\psi\,\phi$ MC and $B^0_s \to J/\psi\,\phi$ data highlight sizeable differences in some of the variables later employed for the MVA selection. This is a known feature of the MC (see Ref. [121], for example) and stems partly from the poorly understood hadronization of the B^0_s mesons in `Pythia` [122].

Typically, the MC is reweighted in the total and transverse momentum of the $B^0_{(s)}$ ($p(B), p_{\rm T}(B)$) and in track multiplicity `nTracks`, to reproduce the distributions in the background-subtracted data, which are obtained with the $sPlot$ technique from the fits reported in Fig. 4.3. In addition to these three variables $\{p(B), p_{\rm T}(B), \texttt{nTracks}\}$, we also reweight in other variables $\{\texttt{DTF\_prob}, \texttt{Mup\_IP\_OWNPV}, \texttt{Mum\_IP\_OWNPV}\}$, defined as the probability of the fit to the B vertex and the $\mu^{\pm}$ impact parameters, respectively. These variables are then employed in the MVA selection described in Sec. 4.3.1. The reweighting in all the MVA variables helps in preserving correlations among the variables that the MVA is sensitive to, thereby yielding a better agreement between data and MC in the MVA itself. Figures 4.9 and 4.10 show the comparisons in the kinematic variables for Run 1 and Run 2, respectively.

For the $B^0_{(s)} \to J/\psi p\bar{p}$ signal modes, comparisons between (unweighted) signal MC of the control and signal modes indicate potential differences after reconstruction, as shown in Fig. 4.11. This might stem from the fact that the proton in the signal mode is significantly heavier than the kaon in the control mode. For this reason, the reweighting is set up on the assumption that the difference between data and MC is the same for different decay modes. A weight is then computed for each signal candidate, in the space of the kinematic $\{p(B), p_{\rm T}(B), \texttt{nTracks}\}$ variables, as the ratio between simulation and background-subtracted data of the $B^0_s \to J/\psi\,\phi$ control mode. Figures 4.13 and 4.14 show the distributions before and after reweighting for Run 1 and Run 2, respectively.

The assumption is also verified employing a sample of $B \to J/\psi K^*$ decays, which have a similar topology with respect to the other modes. In Fig. 4.12, we compared the $p_{\rm T}(B)$ distribution using different weights: the first weights are obtained reweighting the signal MC to the $B^0_s \to J/\psi\,\phi$ data directly, the second ones are extracted from $B^0_s \to J/\psi\,\phi$ as the ratio between $B^0_s \to J/\psi\,\phi$ MC and data and the third weights are extracted from $B \to J/\psi K^*$ decays in the same manner. The distribution obtained from the first method is completely different from the other two cases, proving that our assumption is solid. The weights extracted from $B^0_s \to J/\psi\,\phi$ and $B \to J/\psi K^*$ present a small difference in the total efficiency of the order of 1%, which will be considered as systematic uncertainty, as described in Sec. 4.7.6.

4.2.2 PID resampling

Particle identification (PID) is an important aspect of the LHCb detector capabilities and it is based on two Čerenkov sub-detectors (RICH1, RICH2) which cover different momentum acceptance regions, as described in Sec. 2.2.2. The PID variables describe how much the particle in a track behaves like a certain particle type. For example `proton_PIDp` is the probability that a proton can be identified as a proton.

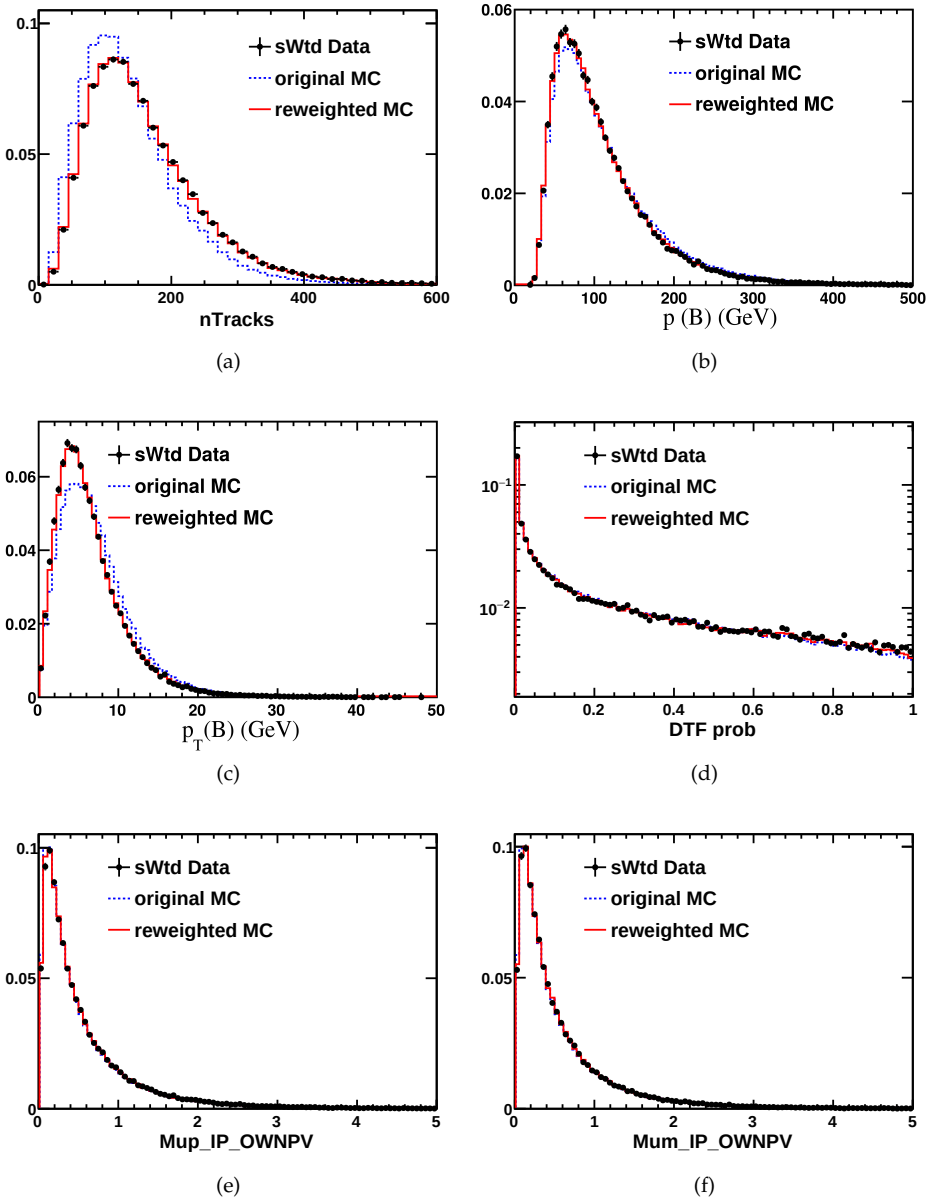


Figure 4.9 Run 1 $B_s^0 \to J/\psi\phi$: Data-MC comparison after reweighting in the kinematic variables. The black markers correspond to background-subtracted data. The blue dashed (red) curves are truth-matched MC before (after) after reweighting as described in Sec. 4.2.1. All histograms are normalized to unity.

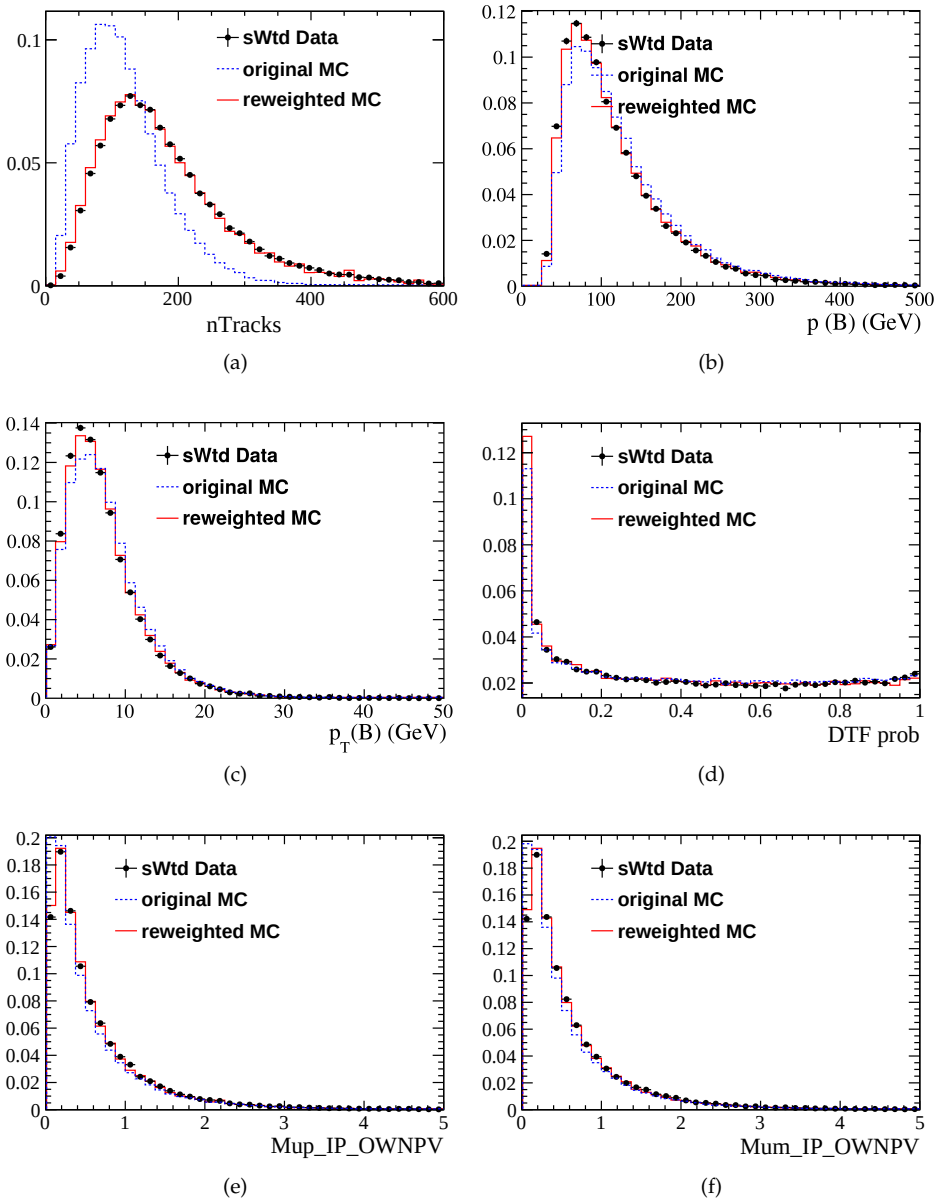
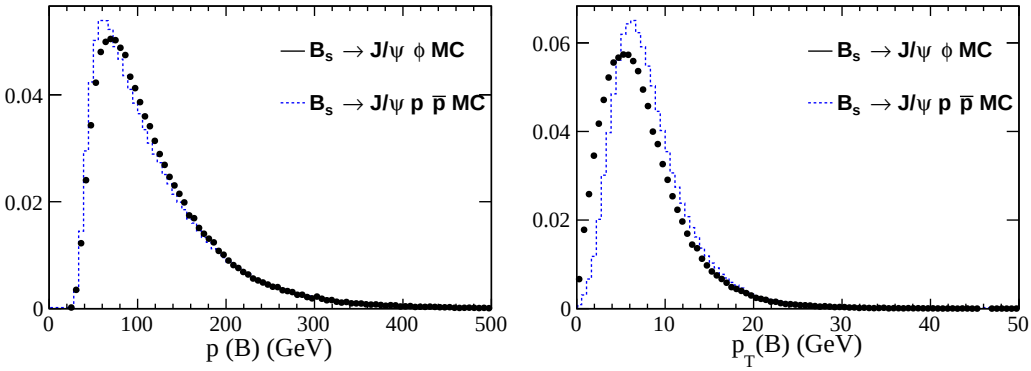


Figure 4.10 Run 2 $B^0_s \to J/\psi\,\phi$: Data-MC comparison after reweighting in the kinematic variables. The black markers correspond to background-subtracted data. The blue dashed (red) curves are truth-matched MC before (after) after reweighting as described in Sec. 4.2.1. All histograms are normalized to unity.



(a) (b)

Figure 4.11 Comparisons between $B_s^0 \to J/\psi\,\phi$ and $B_s^0 \to J/\psi\,p\bar{p}$ unweighted MC in the variables (a) p (B) and (b) $p_{\rm T}$ (B) using Run I signal MC events.

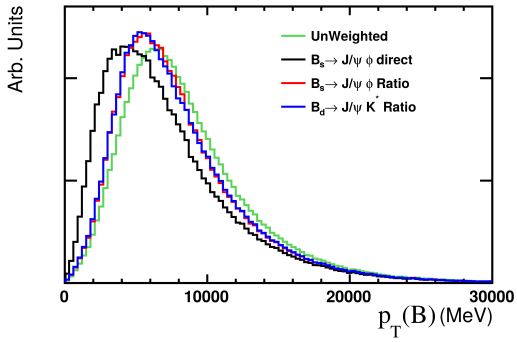


Figure 4.12 Comparisons between $B_s^0 \to J/\psi\,p\bar{p}$ weighted MC in $p_{\rm T}$ (B) for Run I using different weights: unweighted MC (green), weighted to $B_s^0 \to J/\psi\,\phi$ data (black), weighted with weights extracted from $B_s^0 \to J/\psi\,\phi$ (red), weighted with weights extracted from $B \to J/\psi K^*$ (blue).

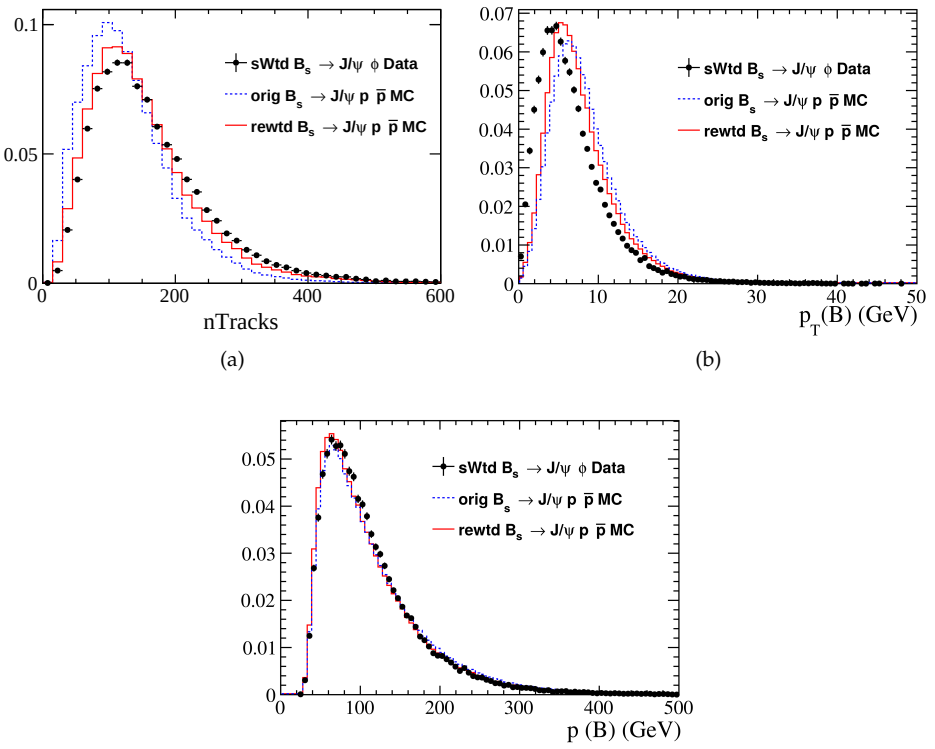


Figure 4.13 Run 1 comparisons between $B^0_s \to J/\psi p\bar{p}$ MC before and after reweighting and $B^0_s \to J/\psi\,\phi$ data.

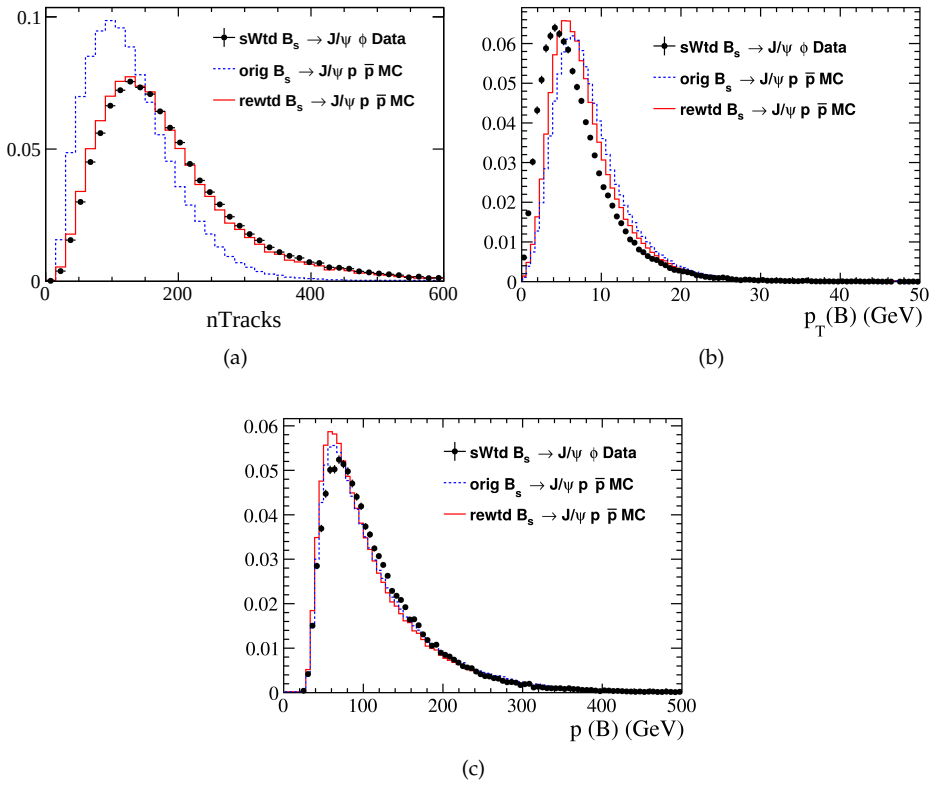


Figure 4.14 Run 2 comparisons between $B_s^0 \to J/\psi p\bar{p}$ MC before and after reweighting and $B_s^0 \to J/\psi\,\phi$ data.

The probability distribution of the PID response is independent of the decay and is approximately a function of

- the true particle type
- the particle momentum: `p`
- the event multiplicity (how many charged tracks are in the event): `nTracks`.
- the pseudo rapidity: η [3].

In this analysis, PID is exploited to obtain a clean sample of signal B^0_s events as a powerful input to the multivariate classifier described in Sec. 4.3.2. In particular, the most important PID variable is called `ProbNN`. It is obtained as the output of a Neural Network based on a recursive method called a Hopfield network [115]. This is machine learning method which uses adaptive training and testing techniques to find patterns in data. It is optimised to have a very high performance in identifying particles.

However, before their usage in the analysis, it is important to ensure good agreement between data and MC, so as to not introduce biases. In order to improve the agreement, the distributions for each PID variable is resampled using the LHCb package, called `PIDGen`, by means of calibration samples extracted directly from data. Unlike the `PIDCalib` package, which exploits a binned approach, the correction is done unbinned, where the calibration PDFs in four dimensions (`ProbNN`, `nTracks`, η and p_{T}) are described by a kernel density estimation procedure using the `Meerkat` library [123]. The samples are obtained from $\Lambda_c^+ \to pK^-\pi^+$ and $D^{*+} \to D^0(\to K^-\pi^+)\pi^+$ decays for proton and kaon, respectively. The proton calibration sample has been validated directly on data, exploiting the Run 2 data sample of $\Lambda_b \to pK^\pm\pi^\mp\pi^-$ decay. The agreement between data and MC after resampling is good, as shown in Fig. 4.17.

While, the kaon calibration samples are validated comparing the $B^0_s \to J/\psi\,\phi$ MC to the background-subtracted data distributions, as shown in Figures 4.15 and 4.16 for Run 1 and Run 2, respectively.

[3] $\eta = 2\ln(\tan\theta/2)$ where θ is the angle with respect to the z axis.

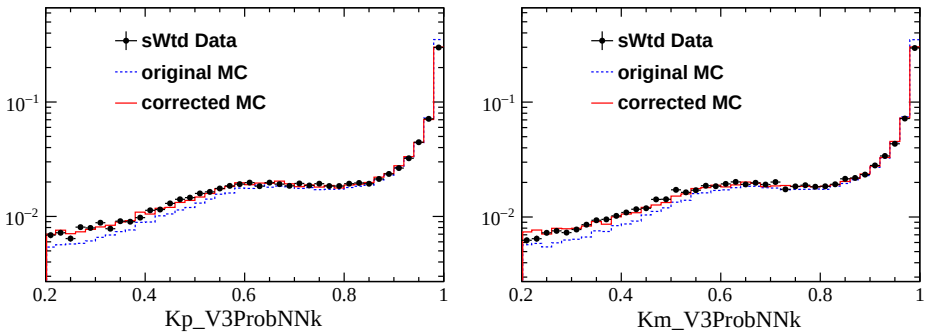


Figure 4.15 Run 1 $B_s^0 \to J/\psi\phi$: Data-MC comparison after resampling in the PID variables. The black markers correspond to background-subtracted data. The blue dashed (red) curves are truth-matched MC before (after) reweighting (Sec. 4.2.1) and PID resampling (Sec. 4.2.2). All histograms are normalized to unity.

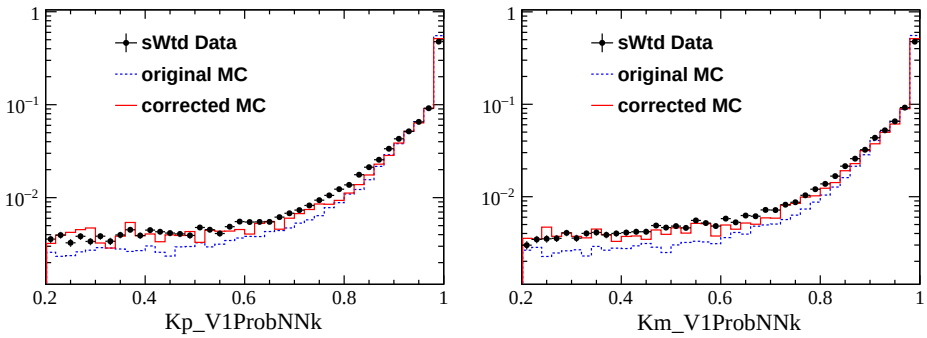


Figure 4.16 Run 2 $B_s^0 \to J/\psi\phi$: Data-MC comparison after resampling in the PID variables. The black markers correspond to background-subtracted data. The blue dashed (red) curves are truth-matched MC before (after) reweighting (Sec. 4.2.1) and PID resampling (Sec. 4.2.2). All histograms are normalized to unity.

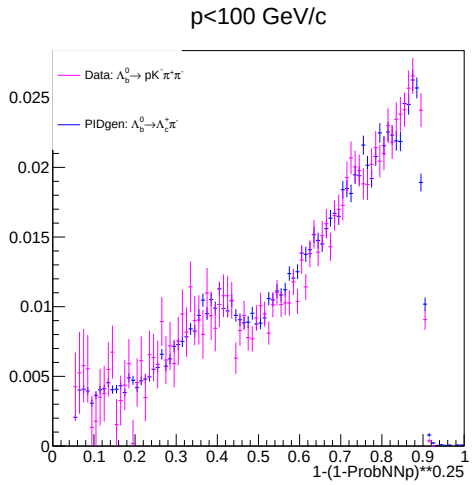


Figure 4.17 Validation of the `PIDGen` calibration sample from $\Lambda_b \to \Lambda_c \pi$ for protons: comparison between data (purple) and MC (blue) from $\Lambda_b \to pK^{\pm}\pi^{\mp}\pi^{-}$ for the transformed variable $(1 - (1 - \texttt{ProbNNp})^{0.25})$.

4.3 Optimization of selection criteria

In order to achieve a good background suppression, the selection is optimized with multivariate techniques based on Boosted Decision Tree (BDT) algorithms [124].

BDT's are one of the two commonly used classifiers in HEP analyses, along with artificial neural networks (NN), to distinguish between signal and background categories. During a training stage, the algorithm learns how to separate the events by exploiting differences between signal and background distributions of a set of appropriately chosen variables. The decision tree is built as a sequence of nodes corresponding to selection criteria. At each node, the split is determined by finding the variable and corresponding cut value that provides the best separation between signal and background. The spilt ends up once the tree reaches the final node, called $leaf$, which categories the events in signal and background according to the majority of events that fall in. The procedure is then iterated on an ensemble of trees, which are combined together in a single classifier through the so-called $boosting$ procedure. An event is then classified on a weighted majority vote of the classifications done by each individual tree.

Since the $B^0_{(s)} \to J/\psi p\overline{p}$ decays are rare modes, after the preselection described in Sec. 4.1.1, no signal events are observed. In order to have a more powerful background rejection, we have optimized the selection using BDT in two stages, as will be described in the following. The first BDT acts as a preselection in order to be able to see a preliminary signal peak. While, the second BDT acts as the main classifier to further reduce the background contribution.

4.3.1 First BDT as pre-selection

In the first BDT stage, the BDT 1 classifier is trained using the copious MC $B^0_s \to J/\psi\phi$ control mode as signal proxy ($\sim$ 100000 events) and $B^0_s \to J/\psi\phi$ data in a mass region above 5450 MeV, called sideband, as background proxy. It employs only kinematic variables whose distributions are similar between the signal and the control mode, in order to keep the selection as similar as possible between the two modes. This helps reducing systematic uncertainties on the efficiencies that enter the branching ratio measurement and cancel in the ratio. The discriminating variables are chosen as the ones that give the best separation power[4] between signal and background, as shown in Table 4.9. These include the p, $p_{\rm T}$, and $\chi^2_{\rm IP}$ of the B^0_s (χ^2 difference adding/removing final tracks to the fit on the associated PV), the χ^2 probability from a kinematic fit [113] to the decay topology, and the IP distances from the PV for the two muons. The BDT performance is also checked looking at the compatibility of the BDT output obtained for training and testing datasets, using the Kolmogorov-Smirnov test [125], which returns the probability that the two distributions are compatible within statistical fluctuations. Failure of this test (probability $< 5\%$) would imply that BDT classifier has become too specific for the samples used for building it. This effect is called $overtraining$. BDT booking options are carefully chosen as a compromise between discrimination power and overtraining probability. Good agreement is found as shown in Fig. 4.18.

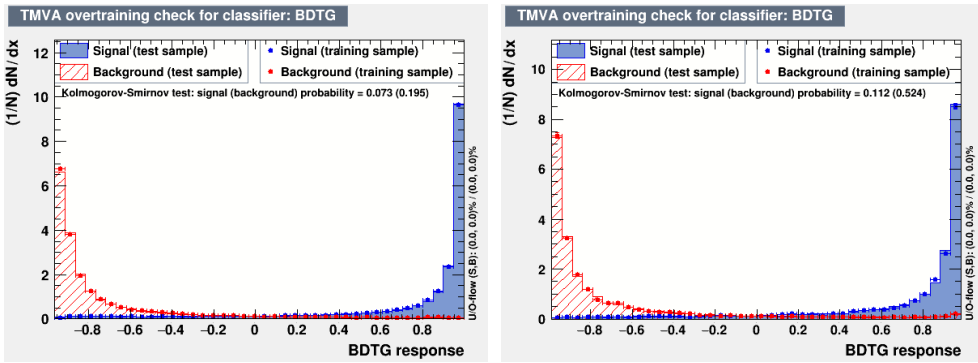
To choose the BDT 1 working point in an optimal but unbiased fashion, the $B^0_s \to J/\psi p\overline{p}$ signal significance, $\mathcal{S}/\sqrt{\mathcal{S}+\mathcal{B}}$, is computed from toy simulations, by generating

[4]Separation power is defined for a training variable x as the average squared difference between the signal $s(x)$ and background $b(x)$ PDFs,

$$< S^2 >= \frac{1}{2}\int \frac{(s(x)-b(x))^2}{s(x)+b(x)}dx. \tag{4.4}$$

Table 4.9 Statistical separation power, $< S^2 >$, for input variables of the BDT I.

Rank	Variable	Run 1 (%)	Run 2 (%)
1	$log(\texttt{IP\_OWNPV}(\mu^-))$	23.96	21.55
2	$log(\texttt{IP\_OWNPV}(\mu^+))$	23.24	21.34
3	$log(\texttt{DTF\_prob})$	19.75	19.61
4	$p_T(B^0_s)$	18.56	19.16
5	$log(\chi^2_{IP}(B^0_s))$	9.05	10.03
6	$p(B^0_s)$	5.43	7.33



(a) (b)

Figure 4.18 Distribution of the BDT response for (a) Run 1 and (b) Run 2 and Kolmogorov-Smirnov test.

and fitting the invariant mass spectrum with the model extracted from MC for signal (see Sec. 4.1.3) and from the sideband region above 5450 MeV for background. The background yield, $\mathcal{B}$, is estimated from a fit to the $B^0_s \to J/\psi p\bar{p}$ data prior to any BDT selection requirement, multiplied by the background efficiency extracted from the sideband. While, the signal yield, $\mathcal{S}$, is obtained from the $B^0_s \to J/\psi p\bar{p}$ branching fraction quoted in Ref. [53], as:

$$S_i = \frac{BR(B^0_s \to J/\psi p\bar{p}) \times \epsilon_{signal,i}}{BR(B^0_s \to J/\psi \phi) \times \epsilon_{CS}} \times N_{CS} \tag{4.5}$$

where $\epsilon_{signal,i}$ and ϵ_{CS} are the total efficiencies for the signal and control sample (CS) estimated from MC and N_{CS} is the number of signal events reported in Sec. 4.1.3. The index i corresponds to different BDT cut values. The S_i yield is then multiplied by the signal efficiency obtained from simulation.

In order to see a reasonable peak, the significance is required to just exceed 5. This choice is made to guarantee a preliminary selection that will be optimized only in the second stage. In Table 4.10, the significance for different BDT cut values is reported; the cuts are therefore chosen to be 0.4 and 0.7 for Run 1 and Run 2, respectively.

The number of signal and background events are extracted from a fit to the invariant mass after BDT 1, as shown in Fig. 4.19, and amount to: $n_{sig} = 395 \pm 54$ and $n_{bkg} = 3246 \pm 17$ in Run 1 and $n_{sig} = 634 \pm 83$ and $n_{bkg} = 7879 \pm 27$ in Run 2, respectively.

Table 4.10 Significance with respect to increasing cut values in Run 1 and Run 2. The 0.4 (0.7) cut is the best one for Run 1 (Run 2) since it reaches a significance larger than 5.

Run 1				
Cut	0.3	0.4	0.5	0.6
Significance	4.88 ± 0.82	5.35 ± 0.85	6.24 ± 0.88	4.81 ± 0.88
Run 2				
Cut	0.5	0.6	0.7	0.8
Significance	4.84 ± 0.69	4.3 ± 0.69	5.22 ± 0.69	5.03 ± 0.69

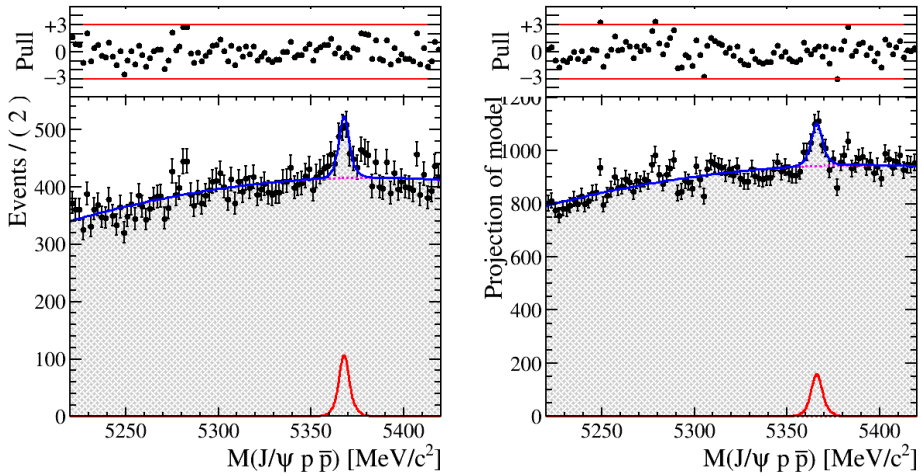


Figure 4.19 Fit results on $B^0_{(s)} \to J/\psi p\bar{p}$ after the pre-selection cuts: (a) Run 1 with BDT 1>0.4 cut, (b) Run 2 BDT 1>0.7 cuts.

4.3.2 Second BDT using hadron PID

In the final selection stage, the hadron PID information (`ProbNNp`) from the RICH detector system is considered in a second classifier, BDT 2, in order to distinguish among hadrons. The BDT 2 also includes the p, $p_{\rm T}$ and $\chi^2_{\rm IP}$ of the protons as training variables, whose statistical separation is summarized in Table 4.11. The $B^0_s \to J/\psi p\bar{p}$ simulation is used as signal sample ($\sim$ 40000 events), while the events in data with $m(J/\psi p\bar{p}) \in [5450, 5500]$ MeV are used as training for the background. The cut on the second BDT is chosen optimizing the significance as figure of merit (FoM), defined as $\mathcal{S}/\sqrt{\mathcal{S}+\mathcal{B}}$, where the signal and background yields are obtained multiplying the initial yields ($\mathcal{S}_0$ and $\mathcal{B}_0$) after BDT 1 by the signal and background efficiencies, taken from MC and sideband data, respectively. The maximum significance corresponds to a cut of 0.6567 for Run 1 and 0.7041 for Run 2, as shown in Fig. 4.20.

After the final selection is applied, most of the events have one signal candidate. A fraction of 0.3% of the events contains multiple candidates and in those cases one candidate per event is selected randomly. A systematic uncertainty is assigned to this

Table 4.11 Statistical separation power, $< S^2 > (\%)$, of input variables of the BDT 2.

Rank	Variable	Run 1 (%)	Run 2(%)
1	$log(\chi^2_{IP}(p))$	20.00	17.59
2	$log(\chi^2_{IP}(\bar{p}))$	19.43	16.50
3	probNNp($\bar{p}$)	13.62	15.94
4	probNNp(p)	13.23	15.37
5	$p_T(\bar{p})$	8.84	12.10
6	$p(p)$	8.44	11.51
7	$p_T(p)$	8.34	6.31
8	$p(\bar{p})$	8.10	4.69

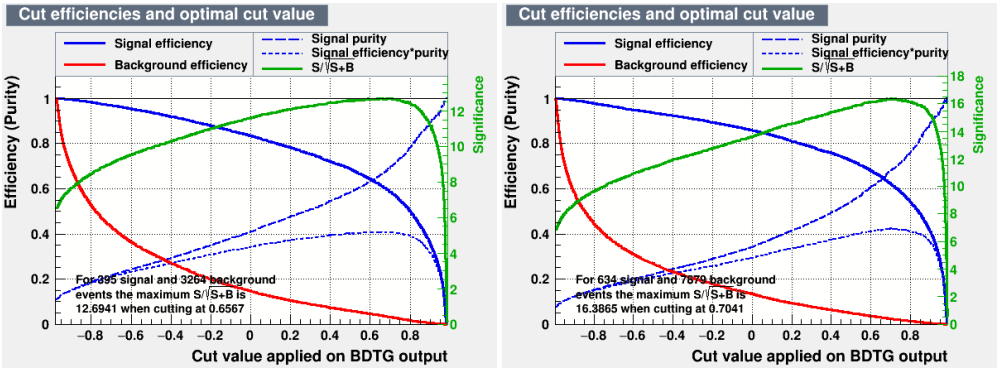


Figure 4.20 Optimisation of the cut on the BDT response for Run 1 (left) and Run 2 (right).

choice as described in Sec. 4.7.7.

4.3.3 Control sample selection

In order to select the $B^0_s \to J/\psi\,\phi$ control sample as similar as possible to the $B^0_{(s)} \to J/\psi p\bar{p}$ signal mode, another BDT is trained on the control mode. For simplicity, a single BDT is used in this case, where the training is performed including both the kinematic variables and PID information (ProbNNk), and using the same samples used in Sec. 4.3.1 for both the signal MC and the data sideband.

The BDT optimization is extracted maximizing the significance as FoM (Fig. 4.21), where the initial signal and background yields are obtained from the $sWeights$ calculation (see Sec. 4.1.3) and are multiplied by the efficiencies. The optimization suggests a cut on the BDT variable at -0.4231 for Run 1 and -0.4464 for Run 2. The performance of the new BDT including the ProbNNk variable is better than BDT 1, as proved by the higher significance of the former, which corresponds to a significance value of 237 with respect to 235 of BDT 1 (see comparisons between Fig. 4.22 and Fig. 4.21).

The use of the hadron PID is also validated on the control mode thanks to this BDT. The distribution of the classifier response is found to be in excellent agreement between the weighted simulation and background-subtracted data of $B^0_s \to J/\psi\,\phi$, as shown in Fig. 4.23. This ensures that the BDT 2 is selecting signal events correctly.

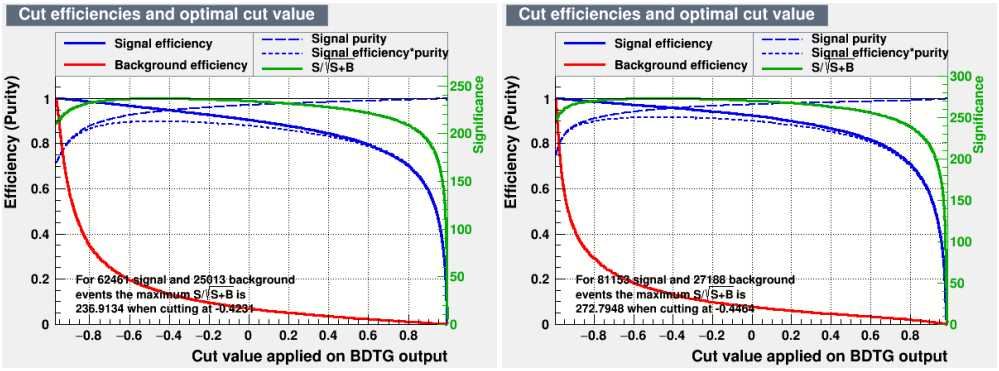


Figure 4.21 Run 1 (left) and Run 2 (right), $B_s^0 \to J/\psi\,\phi$ final BDT: signal efficiency with respect to different BDT cut value for a signal yield of 62461 (81153) in Run 1 (Run 2) and a background yield of 25013 (27188) in Run 1 (Run 2).

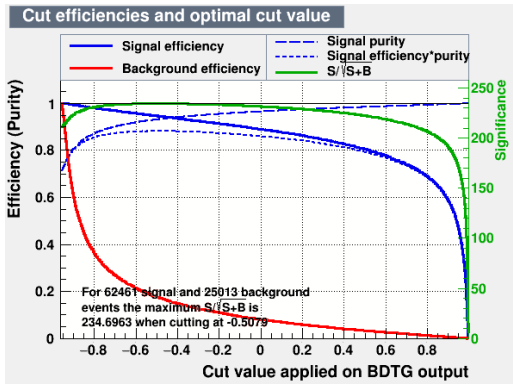


Figure 4.22 BDT 1, $B_s^0 \to J/\psi\,\phi$, Run 1: Signal efficiency and corresponding significance for different cut values optimized with the same initial yields used in the final $B_s^0 \to J/\psi\,\phi$ BDT (BDT 2), see Fig. 4.21. The significance is lower with respect to using BDT 2 as described in section.

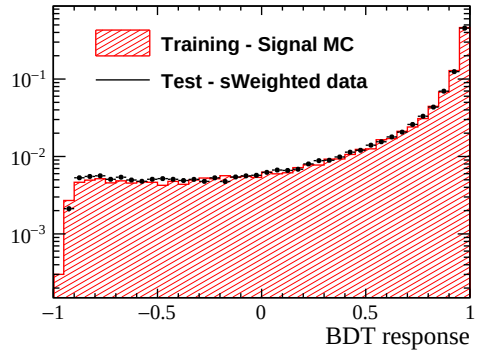


Figure 4.23 Comparison between BDT output on $B_s^0 \to J/\psi\,\phi$ MC for signal and $B_s^0 \to J/\psi\,\phi$ sWeighted data for the final BDT using `ProbNNk` in Run 2.

4.4 Detection Efficiency

In order to perform a branching ratio measurement, efficiency related to trigger, detector acceptance, reconstruction and selection procedure need to be carefully evaluated from MC simulations for the signal and control modes.

In this section, we will present, first of all, the average efficiency at each step of the selection in order to understand the difference between the modes of interest. Then, an event-by-event efficiency is computed to account for variations of the efficiency over the phase space as detailed in Sec. 4.4.5.

The detection efficiency is the product of the efficiencies at each analysis stage. We separate the efficiency into four components:

- ϵ_{geom} the efficiency of signal events to enter the acceptance of the detector;
- $\epsilon_{trigger}$ the efficiency of signal events to pass the trigger requirements;
- ϵ_{reco} the efficiency of signal events to be reconstructed and to pass the stripping and preselections;
- ϵ_{sel} the efficiency of signal events to pass the offline selection, including PID and BDT cuts.

4.4.1 Geometric efficiencies

The generator-level cut efficiency corresponds to the efficiency for each of the daughters of the given decay that satisfies the LHCb acceptance of $10 < \theta < 300$ mrad. In these samples, the B^0_s meson is required to be produced in the LHCb acceptance but the decay products may fall outside of the detector. Tables 4.12, 4.13 indicate the geometric efficiencies for B^0_s and B^0 mesons divided between years and magnet polarities, and the number of events accepted by LHCb, named $N_{\rm bkk}$ (book-keeping) and listed in the right column.

The MC samples used in this analysis were described earlier in Sec. 4.1. Tables 4.12, 4.13 and 4.14 list the actual values of MC sample sizes in the LHCb book-keeping for the signal modes in Run 1 (2011-2012) and for Run 2 (2016).

The total generator level events, $N_{\rm gen}$, is obtained by summing over the different samples

$$N_{\rm gen} = \sum_i \frac{N_{\rm bkk}}{\epsilon_{{\rm geo},i}}. \tag{4.6}$$

Therefore, for Run 1, we obtain:

$$N^{B^0_s \to J/\psi\, p\bar{p}}_{\rm gen} = (44.5 \pm 0.1) \cdot 10^6 \tag{4.7}$$

$$N^{B^0 \to J/\psi\, p\bar{p}}_{\rm gen} = (44.6 \pm 0.1) \cdot 10^6 \tag{4.8}$$

$$N^{B^0_s \to J/\psi\, \phi}_{\rm gen} = (12.03 \pm 0.01) \cdot 10^6 \tag{4.9}$$

$$\tag{4.10}$$

Similarly, for Run 2:

$$N_{\rm gen}^{B_s^0\to J/\psi\, p\bar{p}} = (33.62 \pm 0.08)\cdot 10^6 \tag{4.11}$$

$$N_{\rm gen}^{B^0\to J/\psi\, p\bar{p}} = (33.53 \pm 0.08)\cdot 10^6 \tag{4.12}$$

$$N_{\rm gen}^{B_s^0\to J/\psi\, \phi} = (56.17 \pm 0.09)\cdot 10^6 \tag{4.13}$$

$$\tag{4.14}$$

where the uncertainty takes into account the geometric efficiency uncertainty.

4.4.2 Reconstruction efficiency

The reconstruction efficiency is determined using MC samples as the ratio between the number of reconstructed events which pass the stripping line pre-selections over the initial number of events provided in the LHCb book-keeping, $N_{\rm reco}/N_{\rm bkk}$. The reconstruction and stripping efficiencies are listed in Table 4.15. The value of $\epsilon_{\rm reco}$ includes both the effects that the daughter tracks must be tagged as reconstructible as `Long tracks`, and must also be reconstructed in the detector. In addition, it is required that each event in MC is truth-matched which means that candidate's traces are originated from the B_s^0 vertex. The uncertainty on the efficiency is only statistical and it is evaluated using the binomial formula as

$$\sigma_\epsilon = \sqrt{\frac{\epsilon(1-\epsilon)}{N}}. \tag{4.15}$$

4.4.3 Trigger efficiency

The trigger efficiency is defined as the ratio between the number of signal events which pass a specific trigger line with respect to events without that trigger. Since the trigger requirements are applied offline after the reconstruction, the overall efficiency of the trigger requirements is computed with respect to the reconstruction, as $N_{\rm trig}/N_{\rm reco}$.

4.4.4 Selection efficiency

The selection efficiency is defined as the ratio of the number of events passing cuts on the PID and the BDT selection with respect to the number of triggered events: $N_{\rm sel}/N_{\rm trig}$. In order to produce a cleaner sample of events a cut on $\texttt{ProbNNp}(\texttt{p}) > 0.2$ and $\texttt{ProbNNp}(\bar{\texttt{p}}) > 0.2$ are considered before training the BDT. The final BDT cuts for Run 1 (Run 2) are 0.4

Table 4.12 Geometric efficiencies for $B_s^0 \to J/\psi\, p\bar{p}$ of EventType `13144061`.

i	Sample	$\epsilon_{{\rm geo},i}$	$N_{{\rm bkk},i}$
1	Sim09b 2011MU	0.16679 ± 0.00038	1256906
2	Sim09b 2011MD	0.16679 ± 0.00038	1250139
3	Sim09b 2012MU	0.17023 ± 0.00040	2505832
4	Sim09b 2012MD	0.17042 ± 0.00038	2507230
5	Sim09c 2016MU	0.17860 ± 0.00062	3002196
6	Sim09c 2016MD	0.17860 ± 0.00062	3002793

Table 4.13 Geometric efficiencies for $B^0 \to J/\psi\, p\bar{p}$ of EventType `11144120`.

i	Sample	$\epsilon_{\mathrm{geo},i}$	$N_{\mathrm{bkk},i}$
1	Sim09b 2011MU	0.16716 ± 0.00038	1265959
2	Sim09b 2011MD	0.16681 ± 0.00038	1250455
3	Sim09b 2012MU	0.17030 ± 0.00039	2525446
4	Sim09b 2012MD	0.17057 ± 0.00040	2505129
5	Sim09c 2016MU	0.17920 ± 0.00061	3002323
6	Sim09c 2016MD	0.17893 ± 0.00062	3001678

Table 4.14 Geometric efficiencies for $B^0_s \to J/\psi\, \phi$ of EventType `13144001`.

i	Sample	$\epsilon_{\mathrm{geo},i}$	$N_{\mathrm{bkk},i}$
1	Sim09b 2012MU	0.16721 ± 0.00031	1008750
2	Sim09b 2012MD	0.16766 ± 0.00031	1005503
3	Sim09b 2016MU	0.17743 ± 0.00042	4988476
4	Sim09b 2016MD	0.17745 ± 0.00041	4977902

(0.7) for BDT 1 and 0.6567 (0.7041) for BDT 2, as described in Sec. 4.3.1 and 4.3.2. The second BDT is optimized exploiting the significance as `FoM` (sec. 4.3.2).

Moreover, additional cuts on $m(p\bar{p})$ and $m(B)$ are applied at this stage for $B^0_{(s)} \to J/\psi\, p\bar{p}$ mode. A cut on $m(p\bar{p})$ is considered in order to select events in the kinematic region between $[1877, 2269]$ MeV for B^0_s and $[1877, 2181]$ MeV for B^0 in order to leave out the outliers which are unphysical events and could affect the efficiency calculation. The cut on $m(B)$ is the consequence of the $sWeights$ calculation done in a narrow fit window of $\pm$ 30 MeV around the peak (Sec. 4.5). The number of events removed with these additional cuts is less than 1%, so it is negligible and does not introduce any systematic effect.

All the efficiencies are listed in Table 4.15 for Run 1 and in Table 4.16 for Run 2. The overall average efficiency is around 1% for all decay modes.

4.4.5 Event-by-event efficiency correction

Since the physics models for the $B^0_{(s)} \to J/\psi\, p\bar{p}$ decay processes are unknown, for the acceptance correction, we employ an event-by-event correction to the data using an effi-

Table 4.15 Geometric, reconstruction, trigger and selection integrated phase space efficiencies from Monte Carlo simulations in Run 1. The errors collected in this table are only due to the statistical uncertainties of the Monte Carlo sample. In the last column is also reported the overall efficiency.

Mode	ϵ_{geom}	$\epsilon_{\mathrm{reco/geom}}$	$\epsilon_{\mathrm{trig/reco}}$	$\epsilon_{\mathrm{sel/trig}}$	ϵ_{Tot}
$B^0_s \to J/\psi p\bar{p}$	0.1691 ± 0.0002	0.1537 ± 0.0001	0.7749 ± 0.0004	0.3368 ± 0.0006	0.00678 ± 0.00001
$B^0_d \to J/\psi p\bar{p}$	0.1693 ± 0.0002	0.1523 ± 0.0001	0.7749 ± 0.0004	0.3121 ± 0.0006	0.00623 ± 0.00001
$B^0_s \to J/\psi\, \phi$	0.1674 ± 0.0003	0.1017 ± 0.0002	0.7612 ± 0.0009	0.8749 ± 0.0008	0.01134 ± 0.00003

Table 4.16 Geometric, reconstruction, trigger and selection integrated phase space efficiencies from Monte Carlo simulations in Run 2. The errors collected in this table are only due to the statistical uncertainties of the Monte Carlo sample. In the last column is also reported the overall efficiency.

Mode	ϵ_{geom}	$\epsilon_{\text{reco/geom}}$	$\epsilon_{\text{trig/reco}}$	$\epsilon_{\text{sel/trig}}$	ϵ_{Tot}
$B_s^0 \to J/\psi p\bar{p}$	0.17860 ± 0.00062	0.1684 ± 0.0002	0.8288 ± 0.0004	0.3939 ± 0.0005	0.00986 ± 0.00002
$B_d^0 \to J/\psi p\bar{p}$	0.17920 ± 0.00061	0.1656 ± 0.0002	0.8285 ± 0.0004	0.3712 ± 0.0005	0.00912 ± 0.00002
$B_s^0 \to J/\psi \phi$	0.17740 ± 0.00042	0.1128 ± 0.0003	0.8188 ± 0.0004	0.9564 ± 0.0003	0.01484 ± 0.00002

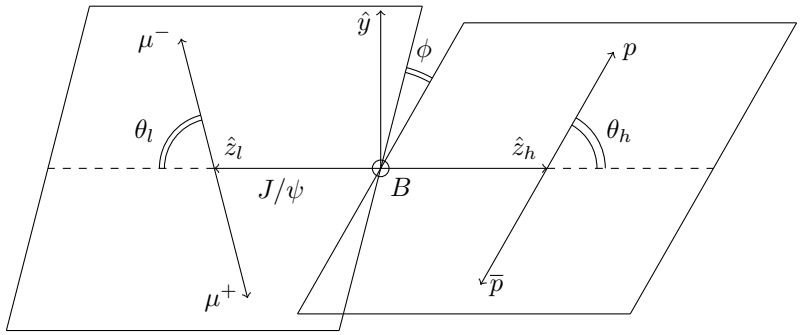


Figure 4.24 The three angular variables $\{\theta_\ell, \theta_h, \chi\}$ for the decay $B \to J/\psi(\to \mu^+\mu^-)h^-h^+$, where $h \in \{p, K\}$. The dihadron and dilepton coordinate systems lie back-to-back with a common vertical $\hat{y}$ axis. The angle between the decay planes is $\phi \in (-\pi, \pi]$, while the two helicity angles, θ_h and θ_ℓ, are defined in the dihadron and dilepton rest frames, respectively.

ciency parameterization based on the decay kinematics. The 4-body phase-space of the topology $B \to J/\psi(\to \mu^+\mu^-)h^+h^-$, where $h \in \{p, K\}$, is fully described by four independent kinematic variables. One of them is the dihadron invariant mass $m_{h^+h^-}$. For a given $m_{h^+h^-}$, the topology can be described by three angles, shown in Fig. 4.24:

- θ_ℓ and θ_h: the helicity angles defined in the dimuon and dihadron rest frames, respectively;
- ϕ: the azimuthal angle between the two decay planes of the dilepton and dihadron systems.

Since the final state is self-conjugate, we choose the h^+ and the μ^- to define the angles, for both B^0_s and $\overline{B^0_s}$. For the signal mode, the overall efficiency, including trigger, detector acceptance and selection procedure, is obtained from simulation as a function of the four kinematic variables, $\vec{\varphi} \equiv \{m'_{p\bar{p}}, \cos\theta_\ell, \cos\theta_h, \phi'\}$. Here, $m'_{p\bar{p}}$ and ϕ' are normalized such that all four variables in $\vec{\varphi}$ lie in the range $(-1, 1]$. The efficiency is parameterized as the product of Legendre polynomials

$$\varepsilon(\vec{\varphi}) = \sum_{i,j,k,l} c_{i,j,k,l}\, P(\cos\theta_\ell, i)P(\cos\theta_h, j)P(\phi', k)P(m'_{p\bar{p}}, l),$$

where $P(x, n)$ is a Legendre polynomial of order n in $x \in (-1, 1]$. Employing the order of the polynomials as $\{3, 7, 7, 5\}$ for $\{m'_{p\bar{p}}, \cos\theta_\ell, \cos\theta_h, \chi'\}$, respectively, was found to give a good parameterization. The coefficients, $c_{h,i,j,k,l}$, are determined from the simulation using a moments technique employing the orthogonality of Legendre polynomials

$$c_{i,j,k,l} = \frac{C}{\sum w_n} \sum_{n=0}^{N_{\rm recon}} w_n \left(\frac{2i+1}{2}\right)\left(\frac{2j+1}{2}\right)\left(\frac{2k+1}{2}\right)\left(\frac{2l+1}{2}\right) \tag{4.16}$$

$$\times P(\cos\theta_\ell, i)P(\cos\theta_h, j)P(\phi', k)P(m', l) \tag{4.17}$$

where w_n is the per-event weight taking into account both the generator level phase-space element, $1/d\Phi$, computed as in Eq. 3.31, and the kinematic event weights described in Sec. 4.2.1. Simulation samples are employed where $B^0_{(s)} \to J/\psi p\bar{p}$ events are generated uniformly in phase-space. In order to render the MC flat also in $m(p\bar{p})$, the inverted phase-space factor, $1/d\Phi$, is considered. The factors of $(2a+1)/2$ arise from the orthogonality of the Legendre polynomials,

$$\int_{-1}^{+1} P(x, a)P(x, a')\mathrm{d}x = \frac{2}{2a+1}\delta_{aa'}\,. \tag{4.18}$$

The sum in Eq. 4.16 is over the reconstructed events in the simulation sample after all selection criteria. The prefactor C ensures appropriate normalization and it is computed such that

$$\sum_{n=0}^{N_{\rm gen}} \varepsilon(\vec{x}_n) = N_{\rm recon}. \tag{4.19}$$

We note that the efficiency here takes into account the geometric, trigger, reconstruction, and selection efficiencies together. Figures 4.25, 4.26 show the comparisons in the kinematic variables between the reconstructed MC and its parametrization for $B^0_s \to J/\psi p\bar{p}$

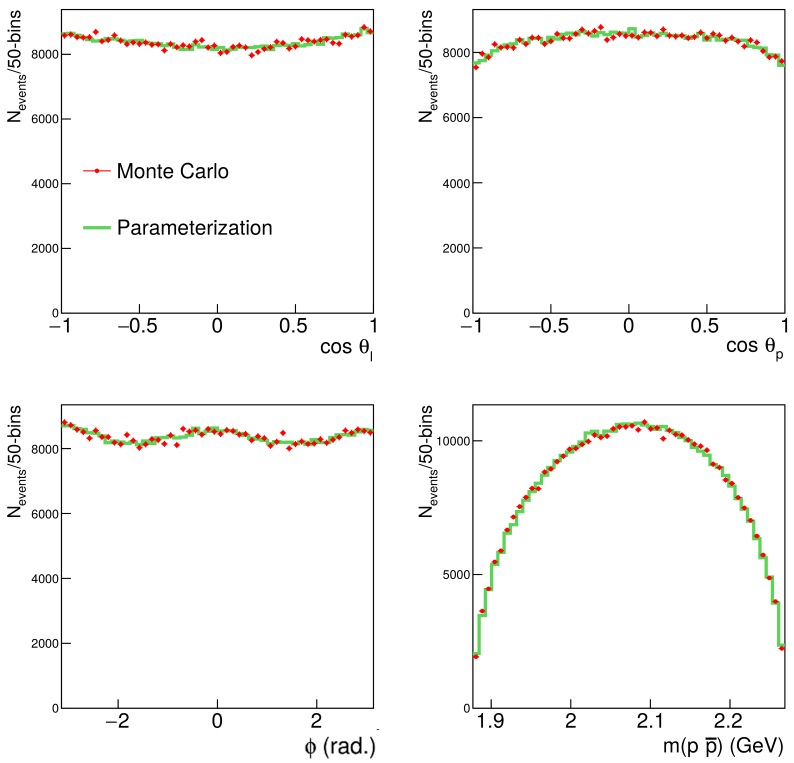
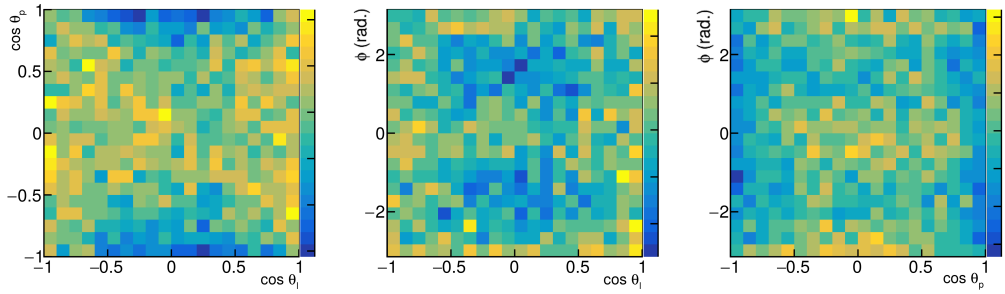
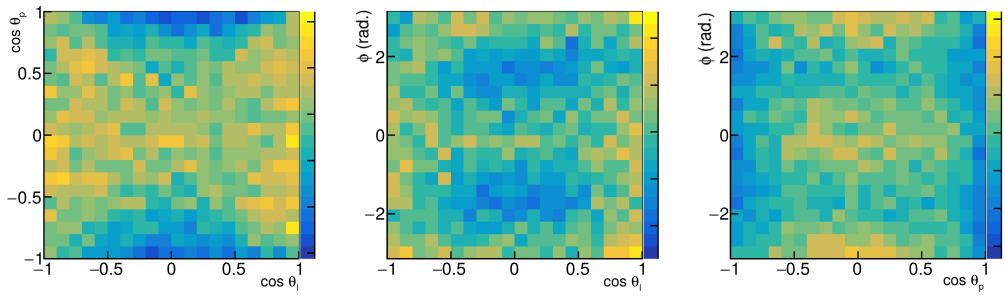


Figure 4.25 Run 1 $B_s^0 \to J/\psi p\bar{p}$: 1-d comparions in the kinematic variables between the reconstructed MC distributions and the efficiency parametrization.



(a)



(b)

Figure 4.26 Run 1 $B^0_s \to J/\psi p\bar{p}$: 2-d scatter plots in the angular variables for the (a) reconstructed MC distributions and the (b) efficiency parametrization. The features in (a) are reproduced in (b), up to statistical fluctuations.

in Run 1. The other comparisons for $B^0_{(s)}$ in Run 1 and Run 2 are reported in Appendix A.1 for completeness. Up to statistical fluctuations, the parametrization follows the simulation data in all the distributions.

For a given data event, the corresponding kinematic variables, $\vec{\varphi}$, are reconstructed and the efficiency, $\varepsilon(\vec{\varphi})$, is computed according to the parameterization. The event is subsequently assigned a weight, $1/\varepsilon(\vec{\varphi})$, to account for the detector efficiency.

4.5 Invariant mass fit

The B^0 and B^0_s signal and background yields are determined via an extended maximum likelihood fit to the $J/\psi p\bar{p}$ invariant mass distribution in the range $[5220,\ 5420]$ MeV. Each signal shape is modeled as the sum of two Crystal Ball [117] functions sharing a common peak position, with tails on either sides of the peak to describe the radiative and misreconstruction effects. The background shape is modeled by a first order polynomial with parameters determined from the fit to data. The signal model parameters are determined from simulation, as reported in Sec. 4.1.3 and only the B^0 and B^0_s central mass values are left as free parameters in the fit to data. The detector invariant mass resolution is in agreement with simulations within a factor of 1.007 ± 0.004 as determined on the control mode. Residual discrepancies are accounted for in the systematic uncertainties. Effects due to potential differences in the signal lineshapes between data and simulation are accounted for in the systematic uncertainties.

The corresponding fit to the signal mode candidates is shown in Fig. 4.27 with the results reported in Table 4.17 for the combination of Run 1 and Run 2 data. In Table 4.18, the results for Run 1 and Run 2 separately are also reported for computing the branching ratio measurement.

The overall statistical significance is computed using Wilks' theorem [126], resulting in 14.6σ and 26.6σ for the B^0 and B^0_s mesons, respectively.

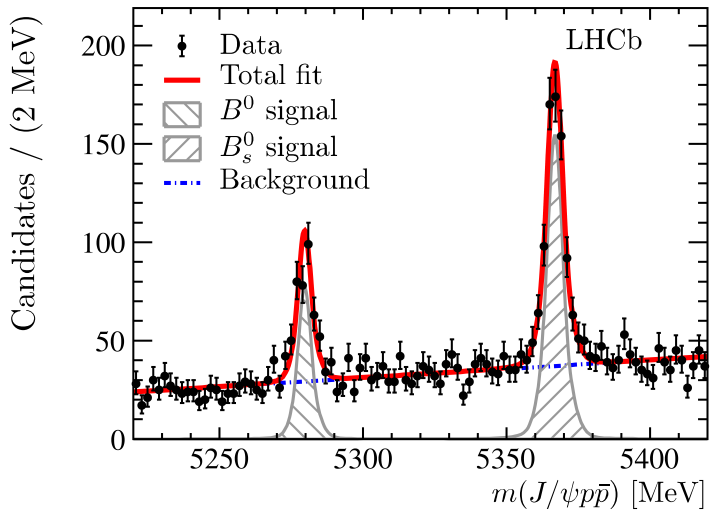


Figure 4.27 Fit to the $J/\psi p\bar{p}$ invariant mass distribution of the $B^0_{(s)}$ signal modes.

Table 4.17 Signal yields, statistical significance and masses for B^0 and B^0_s mesons.

Mode	Yield	Significance	$B^0_{(s)}$ mass [MeV]
$B^0 \to J/\psi p\bar{p}$	256 ± 22	14.6σ	5279.74 ± 0.30
$B^0_s \to J/\psi p\bar{p}$	609 ± 31	26.6σ	5366.85 ± 0.19

Table 4.18 Signal yields, statistical significance and masses for B^0 and B^0_s mesons for Run 1 and Run 2 separately.

Year	Mode	Yield	$B^0_{(s)}$ mass [MeV]
Run 1	$B^0 \to J/\psi p\bar{p}$	111 ± 14	5279.75 ± 0.40
	$B^0_s \to J/\psi p\bar{p}$	240 ± 20	5366.93 ± 0.31
Run 2	$B^0 \to J/\psi p\bar{p}$	150 ± 17	5279.70 ± 0.40
	$B^0_s \to J/\psi p\bar{p}$	375 ± 24	5366.78 ± 0.24

The fit to the $J/\psi K^+K^-$ invariant-mass distribution for the control mode is also shown in Fig. 4.28. The total yields for $B^0_s \to J/\psi\phi(\to K^+K^-)$ are 59088 ± 261 and 77879 ± 308 for Run 1 and Run 2, respectively.

4.6 Branching ratio measurements

The following branching fractions are measured with respect to the $B^0_s \to J/\psi\phi(\to K^+K^-)$ control mode

$$\begin{aligned}\frac{\mathcal{B}(B^0 \to J/\psi p\bar{p})}{\mathcal{B}(B^0_s \to J/\psi\phi) \times \mathcal{B}(\phi \to K^+K^-)) \times f_s/f_d} &= \frac{N^{\rm corr}_{B^0 \to J/\psi p\bar{p}}}{N^{\rm corr}_{B^0_s \to J/\psi K^+K^-}} \\ \frac{\mathcal{B}(B^0_s \to J/\psi p\bar{p})}{\mathcal{B}(B^0_s \to J/\psi\phi) \times \mathcal{B}(\phi \to K^+K^-))} &= \frac{N^{\rm corr}_{B^0_s \to J/\psi p\bar{p}}}{N^{\rm corr}_{B^0_s \to J/\psi K^+K^-}},\end{aligned} \tag{4.20}$$

where f_s/f_d is the ratio of the b-quark hadronization probabilities into B^0_s and B^0 mesons, and $N^{\rm corr}$ denotes efficiency-corrected signal yields, as described in Sec. 4.4.5.

For the absolute branching-fraction determination, the denominators need to be carefully evaluated. The value of

$$\mathcal{B}(B^0_s \to J/\psi\phi) \times \mathcal{B}(\phi \to K^+K^-)) \times f_s/f_d = (1.314 \pm 0.016 \pm 0.079) \times 10^{-4} \tag{4.21}$$

is obtained from Ref. [127] as the product of the two branching ratios $\mathcal{B}(B^0_s \to J/\psi\phi) = (10.50 \pm 0.13 \pm 0.64) \times 10^{-4}$ and $\mathcal{B}(\phi \to K^+K^-) = 0.489 \pm 0.005$, and the old value of the fragmentation factor $f_s/f_d = 0.256 \pm 0.020$ [128] for Run 1, while for Run 2 the ratio f_s/f_d has been multiplied by an additional scale factor of 1.068 ± 0.046 [91] to take into account the dependence on the center of mass energy. For the B^0_s-meson normalization, the B^0 normalization is further divided by the updated ratio $f_s/f_d = 0.259 \pm 0.015$ [128].

The small S-wave K^+K^- fraction under the $\phi(1020)$ resonance, $F_S = 0.0070 \pm 0.0005$ [127], is accounted for as a correction. This is due to the fact that the $B^0_s \to J/\psi\phi$ decay rate includes the S-wave of the $\phi(1020)$ resonance, while the PDG value of the

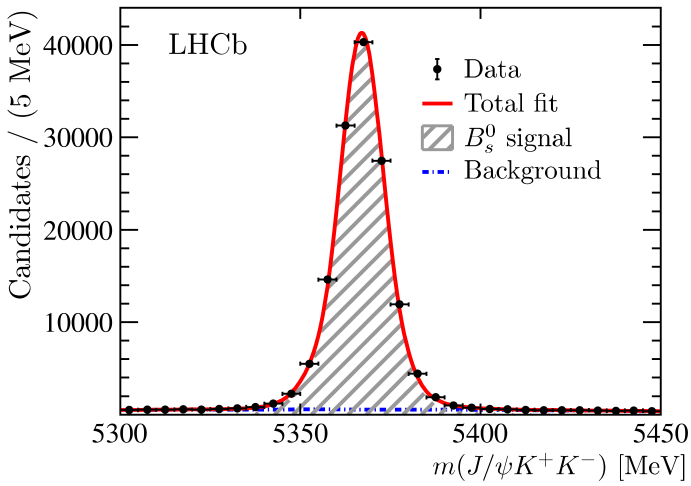


Figure 4.28 Fit to $m(J/\psi\, K^+K^-)$ invariant mass for Run 1+Run 2.

branching fraction is only for the P-wave contribution. Indeed, the S-wave fraction needs to be subtracted from our measured rate, using $(1 - F_s)$. The value of F_S is obtained from the measurement in Ref. [127], by interpolating the value of it in the exact $|m_{KK} - 1020| \leq 5$ MeV mass window, from Fig. 19 in Ref. [127].

4.6.1 Efficiency corrected yields

The efficiency corrected signal yields, $N^{\rm corr}$, are calculated in a different fashion for the signal and control mode.

For the control mode, where the MC generator model is a good representation of the data, they are computed considering the average efficiency, as

$$N^{\rm corr}_{B^0_s\to J/\psi\, K^+K^-} = \frac{N^{\rm meas}_{B^0_s\to J/\psi\, K^+K^-}}{\bar{\varepsilon}_{B^0_s\to J/\psi\, K^+K^-}}. \tag{4.22}$$

The estimated number of events for the control mode after the full selection yields to $N^{\rm meas}_{B^0_s\to J/\psi\, K^+K^-} = 59088 \pm 261$ and 77879 ± 308 events in Run 1 and Run 2, respectively, while the efficiencies, calculated as $N^{\text{recon. MC}}_{B^0_s\to J/\psi\, K^+K^-}/N^{\text{gen MC}}_{B^0_s\to J/\psi\, K^+K^-}$, correspond to the values

$$\bar{\varepsilon}^{\rm RunI}_{B^0_s\to J/\psi\, K^+K^-} = (1.134 \pm 0.003)\% \tag{4.23}$$

$$\bar{\varepsilon}^{\rm RunII}_{B^0_s\to J/\psi\, K^+K^-} = (1.484 \pm 0.003)\%. \tag{4.24}$$

Therefore, the corrected yields amount to

$$N^{\rm corr,RunI}_{B^0_s\to J/\psi\, K^+K^-} = (5.21 \pm 0.03)\cdot 10^5 \tag{4.25}$$

$$N^{\rm corr,RunII}_{B^0_s\to J/\psi\, K^+K^-} = (5.25 \pm 0.02)\cdot 10^5 \tag{4.26}$$

On the other hand, since the $B^0_{(s)} \to J/\psi\, p\bar{p}$ signal model is not known *a priori*, an event-by-event efficiency correction is considered, as described in Sec. 4.4.5. Each

background subtracted event i is weighted by the inverse of the efficiency parametrization, $1/\varepsilon(\vec{x}_i)$, from Eq. 4.16. The signal yields are extracted from the $m(J/\psi p\bar{p})$ invariant mass fit, as shown in Fig. 4.27 and with the results reported in Table 4.18. Background-subtracted events are then computed using the `sPlot` technique [118]. Given the large amount of background and the narrow resolution on the $B^0_{(s)}$ masses, the restricted ranges $m(J/\psi p\bar{p}) \in [5250, 5310]$ MeV and $m(J/\psi p\bar{p}) \in [5337, 5397]$ MeV are considered for the B^0 and B^0_s, respectively.

The background subtracted and efficiency corrected yields are then calculated as $\sum_i \mathcal{W}_i/\varepsilon(\vec{x}_i)$, where $\mathcal{W}_i$ are the $sWeights$, and amount to

$$N^{\text{corr,Run 1}}_{B^0\to J/\psi p\bar{p}} = (17.9 \pm 2.4)\cdot 10^3 \tag{4.27}$$

$$N^{\text{corr,Run 1}}_{B^0_s\to J/\psi p\bar{p}} = (35.2 \pm 3.2)\cdot 10^3 \tag{4.28}$$

$$N^{\text{corr,Run 2}}_{B^0\to J/\psi p\bar{p}} = (16.7 \pm 1.9)\cdot 10^3 \tag{4.29}$$

$$N^{\text{corr,Run 2}}_{B^0_s\to J/\psi p\bar{p}} = (37.6 \pm 2.5)\cdot 10^3 \tag{4.30}$$

where the errors include the statistical uncertainties only and are evaluated as:

$$\sigma_{N^{\text{corr}}} = \sqrt{\sum_i \frac{\mathcal{W}_i}{\epsilon(\vec{x}_i)}}. \tag{4.31}$$

In order to evaluate what is the effect of the 4D efficiency parametrization, average efficiencies after the event-by-event phase-space correction are computed as N^{fit}/N^{corr} (where the yields are taken from Table 4.18 and Eq. 4.27). The values reported below

$$\begin{aligned}
\bar{\varepsilon}^{\text{Run 1}}_{B^0\to J/\psi p\bar{p}} &\sim 0.621\%,\\
\bar{\varepsilon}^{\text{Run 2}}_{B^0\to J/\psi p\bar{p}} &\sim 0.899\%,\\
\bar{\varepsilon}^{\text{Run 1}}_{B^0_s\to J/\psi p\bar{p}} &\sim 0.682\%,\\
\bar{\varepsilon}^{\text{Run 2}}_{B^0_s\to J/\psi p\bar{p}} &\sim 0.996\%,
\end{aligned} \tag{4.32}$$

differ from the flat phase-space efficiencies (Table 4.15) by about 1%.

4.6.2 Branching fraction results

The ratio of branching fractions defined in Eq. 4.20 is reported here separately, for Run 1 and Run 2. For Run 1,

$$\frac{\mathcal{B}(B^0_s \to J/\psi p\bar{p})}{\mathcal{B}(B^0_s \to J/\psi\,\phi(\to K^+K^-))} = 0.68\% \pm 0.06\%(\text{stat}) \pm 0.05\%(\text{sys}), \tag{4.33}$$

$$\frac{\mathcal{B}(B_d \to J/\psi p\bar{p})}{\mathcal{B}(B^0_s \to J/\psi\,\phi(\to K^+K^-)) \times \frac{f_s}{f_d}} = 0.35\% \pm 0.05\%(\text{stat}) \pm 0.03\%(\text{sys}), \tag{4.34}$$

and for Run 2,

$$\frac{\mathcal{B}(B^0_s \to J/\psi p\bar{p})}{\mathcal{B}(B^0_s \to J/\psi\,\phi(\to K^+K^-))} = 0.72\% \pm 0.05\%(\text{stat}) \pm 0.05\%(\text{sys}), \tag{4.35}$$

$$\frac{\mathcal{B}(B_d \to J/\psi p\bar{p})}{\mathcal{B}(B^0_s \to J/\psi\,\phi(\to K^+K^-)) \times \frac{f_s}{f_d}} = 0.32\% \pm 0.04\%(\text{stat}) \pm 0.02\%(\text{sys}). \tag{4.36}$$

The absolute branching fractions obtained from our present knowledge of $\mathcal{B}(B_s^0 \to J/\psi\,\phi(\to K^+K^-))$ and f_s/f_d as described in Sec. 4.6 are, for Run 1,

$$\mathcal{B}(B^0 \to J/\psi\, p\overline{p}) = (4.54 \pm 0.62(\text{stat}) \pm 0.34(\text{syst}) \pm 0.28(\text{norm})) \times 10^{-7},$$
$$\mathcal{B}(B_s^0 \to J/\psi\, p\overline{p}) = (3.45 \pm 0.31(\text{stat}) \pm 0.26(\text{syst}) \pm 0.21(\text{norm}) \pm 0.20(f_s/f_d)) \times 10^{-6},$$

and for Run 2,

$$\mathcal{B}(B^0 \to J/\psi\, p\overline{p}) = (4.49 \pm 0.52(\text{stat}) \pm 0.31(\text{syst}) \pm 0.28(\text{norm}) \pm 0.19(f_s/f_d)) \times 10^{-7},$$
$$\mathcal{B}(B_s^0 \to J/\psi\, p\overline{p}) = (3.66 \pm 0.25(\text{stat}) \pm 0.25(\text{syst}) \pm 0.22(\text{norm}) \pm 0.21(f_s/f_d)) \times 10^{-6}.$$

where the first uncertainty is statistical, the second is systematic and will be evaluated in Sec. 4.7, the third comes from the normalization branching fraction and the last one from the uncertainty on f_s/f_d. For the B^0 mode in Run 1, the latter uncertainty is not present since it is already considered in the normalization, $\mathcal{B}(B_s^0 \to J/\psi\,\phi) \times \mathcal{B}(\phi \to K^+K^-)) \times f_s/f_d$ [127], while in Run 2, the additional energy-dependent correction on f_s/f_d has an uncertainty of 4.3% [91].

To combine the results from the Run 1 and Run 2 data samples, the systematic uncertainties are taken as fully correlated between the run periods, since the data are collected employing the same detector and analysis techniques. The statistical uncertainties are considered uncorrelated since the datasets are disjoint. The covariance matrices are constructed as

$$V_{ij} = \begin{pmatrix} \sigma^2_{\text{stat, Run 1}} + \sigma^2_{\text{syst, Run 1}} & \sigma_{\text{syst, Run 1}} \times \sigma_{\text{syst, Run 2}} \\ \sigma_{\text{syst, Run 1}} \times \sigma_{\text{syst, Run 2}} & \sigma^2_{\text{stat, Run 2}} + \sigma^2_{\text{syst, Run 2}} \end{pmatrix},$$

with the numerical values for the absolute branching fraction combination as

$$V_{B^0} = \begin{pmatrix} 0.5754 & 0.2022 \\ 0.2022 & 0.4849 \end{pmatrix} \times 10^{-14},$$
$$V_{B_s^0} = \begin{pmatrix} 0.2057 & 0.1123 \\ 0.1123 & 0.1727 \end{pmatrix} \times 10^{-12}.$$

The weighted mean value and uncertainty are then calculated as

$$\overline{x} = \sum_{i\in\{1,2\}} w_i x_i,$$
$$\sigma_x^2 = \sum_{i,k\in\{1,2\}} w_i w_k V_{ik},$$

respectively, where the variable x denotes the branching fraction and the weights are extracted from the inverse of the aforementioned covariance matrix as

$$w_i = \frac{\sum\limits_{k\in\{1,2\}} (V^{-1})_{ik}}{\sum\limits_{j,k\in\{1,2\}} (V^{-1})_{jk}}.$$

The combination of the measurements is taken as a weighted mean to give the branching fraction ratios

$$\frac{\mathcal{B}(B^0 \to J/\psi p\bar{p})}{\mathcal{B}(B^0_s \to J/\psi\,\phi(\to K^+K^-)) \times f_s/f_d} = (0.329 \pm 0.029\ (\text{stat}) \pm 0.022\ (\text{syst})) \times 10^{-2},$$
$$\frac{\mathcal{B}(B^0_s \to J/\psi p\bar{p})}{\mathcal{B}(B^0_s \to J/\psi\,\phi(\to K^+K^-))} = (0.706 \pm 0.037\ (\text{stat}) \pm 0.048\ (\text{syst})) \times 10^{-2},$$

where the first uncertainty is statistical and the second is systematic.

The absolute branching fractions are then combined to give

$$\mathcal{B}(B^0 \to J/\psi p\bar{p}) = (4.51 \pm 0.40\ (\text{stat}) \pm 0.44\ (\text{syst})) \times 10^{-7},$$
$$\mathcal{B}(B^0_s \to J/\psi p\bar{p}) = (3.58 \pm 0.19\ (\text{stat}) \pm 0.33\ (\text{syst})) \times 10^{-6},$$

where the systematic uncertainty is the sum in quadrature of the overall systematic contribution on the ratio of branching fractions, the normalization mode uncertainty and the f_s/f_d uncertainty for the B^0_s signal, reported in Table 4.19.

Table 4.19 Systematic uncertainties on the branching fraction measurements for Run 1 and Run 2. The total uncertainties on the branching fraction ratios (BFR) are the sum of the systematic uncertainties, added in quadrature. The total uncertainties on the absolute branching fractions (BF) include the normalization and the uncertainties on the ratio f_s/f_d from external measurements as well.

	$\mathcal{B}(B^0 \to J/\psi p\bar{p})$	$\mathcal{B}(B_s^0 \to J/\psi p\bar{p})$
	Run 1 (Run 2)	Run 1 (Run 2)
Fit model	1.0 (0.5)%	1.0 (0.9)%
Detector resolution	0.6 (0.5)%	0.4 (0.6)%
PID efficiency	5.0 (4.0)%	5.0 (4.0)%
Trigger	1.0 (1.0)%	1.0 (1.0)%
Tracking	5.0 (5.0)%	5.0 (5.0)%
Simulation reweighting	0.4 (0.4)%	0.3 (0.3)%
Multiple candidates	0.1 (0.1)%	0.1 (0.1)%
Total on BFR	7.2 (6.5)%	7.2 (6.6)%
Normalization	6.1 (6.1)%	6.1 (6.1)%
f_s/f_d	- (4.3)%	5.8 (5.8)%
Total on BF	9.4 (10.1)%	11.1 (10.7)%

4.7 Systematic uncertainties

By employing a control mode with topology very similar to that of the signal mode, most of the systematic uncertainties cancel in the branching fraction measurement. However, residual systematic effects still remain, which are summarized in Table 4.19 and discussed in more detail below.

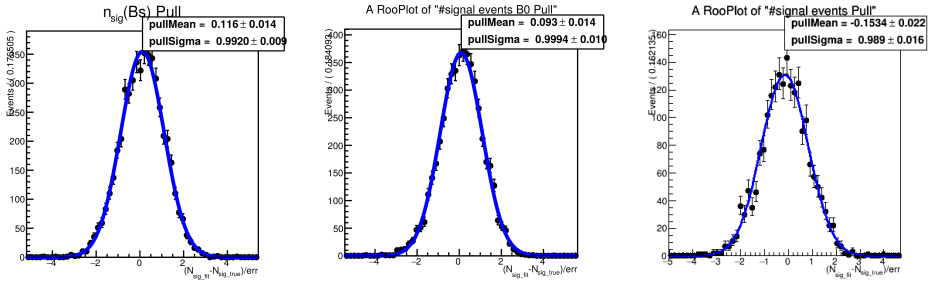
4.7.1 Fit model

To evaluate systematic uncertainties related to our choice of fit model, we vary the nominal model and take the change in the fit value as the associated systematic uncertainty. For the $B^0_{(s)} \to J/\psi p\bar{p}$ signal modes, the model is replaced by the convolution of three Gaussian functions for the signal and a sigmoidea function, defined as $1/(1+\exp(-b_0(m-b_1)))$, for the background. For this purpose, 5000 samples of pseudo-experiments were generated using the nominal model for generation and the alternative model for fitting.

The normalized residuals or "pull" of the yields are determined as in Fig. 4.30. The pull is defined as the difference between the number of fitted events and the true value (the mean of the generated Poisson distribution) over the error of the fit:

$$\text{Pull} = \frac{N_{fit} - N_{true}}{\sigma}. \tag{4.37}$$

One expects the pull to follow a normal Gaussian, so from its observed distribution one can draw conclusions about whether or not the fit reports unbiased central values and errors of correct coverage. If the pull distribution has mean $\mu_p \pm \sigma_{\mu_p}$ and width $\omega_p \pm \sigma_{\omega_p}$ that are not equal to 0 and 1, respectively, one can correct the yields extracted



(a) (b) (c)

Figure 4.29 Run 1 Fit model systematic uncertainty in Sec. 4.7.1: pulls of signal yields for (a) B^0_s (b) B^0, and (c) the control sample $B^0_s \to J/\psi\,\phi$.

from the nominal fit N_{fit} and the associated uncertainty σ_{fit} in this way:

$$N^c_{fit} = N_{fit} - \mu_p \sigma_{N_{fit}}, \tag{4.38}$$

$$\sigma^c_{N_{fit}} = \omega_p \cdot \sigma_{N_{fit}}. \tag{4.39}$$

The corrected yields are used to compute a new branching ratio measurement ($\mathcal{B}$) and the relative systematic uncertainty is taken as the difference between the nominal value $\mathcal{B}$ and the corrected one $\mathcal{B}^c$:

$$\sigma^{sys}/\mathcal{B} = (\mathcal{B}^c - \mathcal{B})/\mathcal{B}. \tag{4.40}$$

The associated uncertainty on the B^0_s signal yield is 0.9%; for B^0 a value of 1.0% uncertainty is assigned instead.

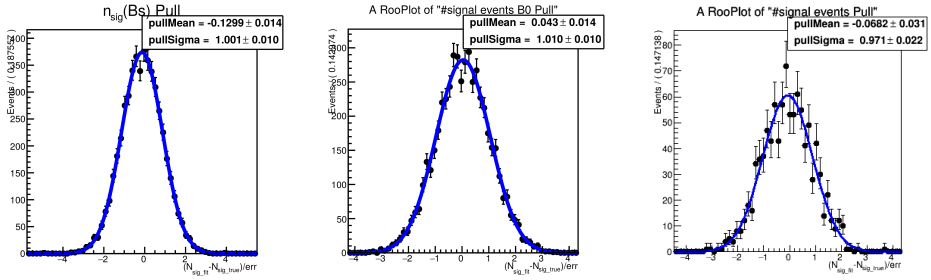
The same procedure is followed for the control mode, $B^0_s \to J/\psi\,\phi$, where the alternative model is built replacing the signal shape with three Gaussian functions and the background with an exponential. The uncertainty on the B^0_s yield is found to be 0.06% for Run 1 and 0.1% for Run 2. Figure 4.30 shows the pulls on the yields for both the signal and control modes in Run 1 and Run 2, respectively. The propagation of these uncertainties on the average branching fraction, using Eq. 4.40, is estimated to be 1% for both B^0_s and B^0 in Run 1 and 0.9% (0.5%) for B^0_s (B^0) in Run 2.

4.7.2 Detector resolution

The stability of the fit model is validated by changing the width parameters to account for possible systematic effects due to detector resolution. Initially, in the nominal model, they were fixed to the MC values, as described in Sec. 4.1.3. However, the widths extracted from data could vary due to the momentum resolution which is not perfectly reproduced in the simulations. As a result the width is parametrized as:

$$\sigma_{data} = \sigma_{MC} \cdot (1 + \alpha), \tag{4.41}$$

where α is the scale factor extracted from the fit to the data in the control sample $B^0_s \to J/\psi\,\phi$. The value of α is found to be 0.007 ± 0.004 for Run 1 and 0.010 ± 0.004 for Run 2 as shown in Fig. 4.31. To estimate the systematic effect due to this approximation, toys simulations are performed by generating with the nominal model and fitting with the inclusion of this scale factor. Figure 4.32 report the pull distributions of the signal yields for



(a) (b) (c)

Figure 4.30 Run 2 Fit model systematic uncertainty in Sec. 4.7.1: pulls of signal yields for (a) B_s^0 (b) B^0, and (c) the control sample $B_s^0 \to J/\psi\,\phi$.

both the modes. The resulting systematic uncertianties are 0.6%(0.5%) and 0.4%(0.6%) for Run 1(Run 2) for B_s^0 and B^0, respectively.

4.7.3 PID efficiency

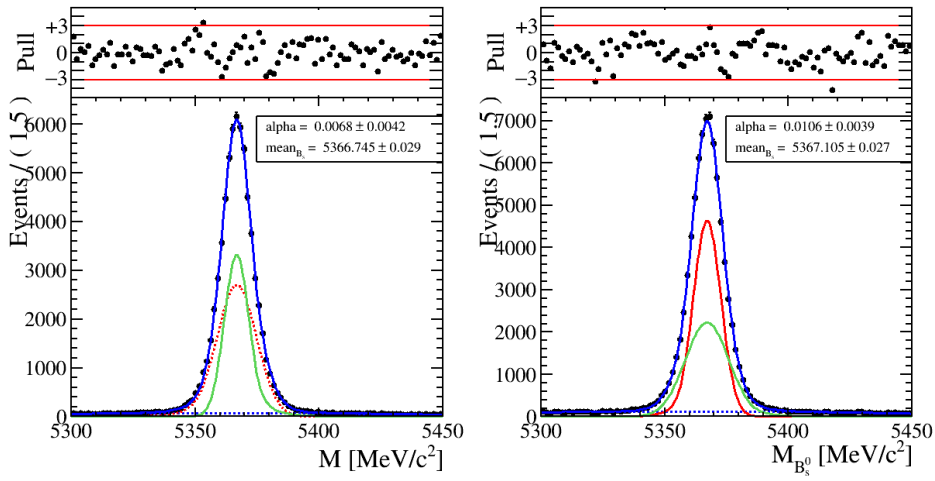
Given that the usage of the `ProbNNp` variables in the BDT plays a central role in this analysis, the PID systematic from the resampling has to be estimated carefully. Possible sources of systematic uncertainties may arise from the limited calibration sample statistics and the coverage of the calibration samples in the space of p, p_{T} and track multiplicity N_{tr}. For this reason, different calibration samples are considered for the systematic evaluation.

As described in Sec. 4.2.2 the resampling is performed with calibration samples taken from $\Lambda_c^+ \to pK^-\pi^+$ decays. As alternative calibration sample, the decay of $\Lambda^0 \to p\pi^-$ can be exploited, which is larger in size. However, since the Λ^0 is a long-lived particle, the `ProbNNp` variable can be distorted for tracks coming from the PV. Indeed, it seems to be correlated with the track displacement from the PV, or other related variables as the `MINIPCHI2`, which is defined as the minimum χ^2 of the fit to the PV adding/removing a final track. For this reason, the Λ^0 calibration sample is not used in the nominal PID correction. However, it can be used to assign a conservative systematic uncertainty.

The systematic is then assigned as the difference between the average efficiencies using as calibration samples the Λ^0 decays and Λ_c^+ decays. The results are reported in Table 4.20. In order to be conservative, we decided to assign an average 5% uncertainty in Run 1 and 4% in Run 2 for both B_s^0 and B^0.

4.7.4 Trigger

The generator-level p_T distributions of the final state particles in the $B_{(s)}^0 \to J/\psi\, p\overline{p}$ and $B_s^0 \to J/\psi\,\phi$ modes are very similar as expected. Since the triggers used are based on the (di-)muons, any systematic uncertainty related to the ratio of efficiencies will be small. The systematic uncertainty related to the trigger efficiency is taken from the study performed in Ref. [129] that compared trigger efficiency ratios in data and simulation for the high yield decay modes $B \to J/\psi\, K^+$ and $B \to \psi(2S)K^+$ and found agreement within 1%. These channels have similar kinematic properties to the decay modes considered here, so a 1% uncertainty is assigned.



(a) (b)

Figure 4.31 Mass resolution systematic in Sec. 4.7.2: invariant mass distribution of $B^0_s \to J/\psi\,\phi$ for Run 1(a) and Run 2(b) dataset after the full selection with the inclusion of the detector resolution, α.

Table 4.20 Total average efficiency evaluated using the alternative resampling tecnique for signal and control sample and the corresponding systematic uncertainty associated to the Branching Ratio.

Run I		
Mode	Total efficiency	Systematic on BR
$B^0_s \to J/\psi p\bar{p}$	0.00652 ± 0.00001	4.6%
$B^0 \to J/\psi p\bar{p}$	0.00595 ± 0.00001	5.3%
$B^0_s \to J/\psi\,\phi$	0.01139 ± 0.00001	
Run II		
Mode	Total efficiency	Systematic on BR
$B^0_s \to J/\psi p\bar{p}$	0.01041 ± 0.00002	4.4%
$B^0 \to J/\psi p\bar{p}$	0.00949 ± 0.00001	2.6%
$B^0_s \to J/\psi\,\phi$	0.01503 ± 0.00002	

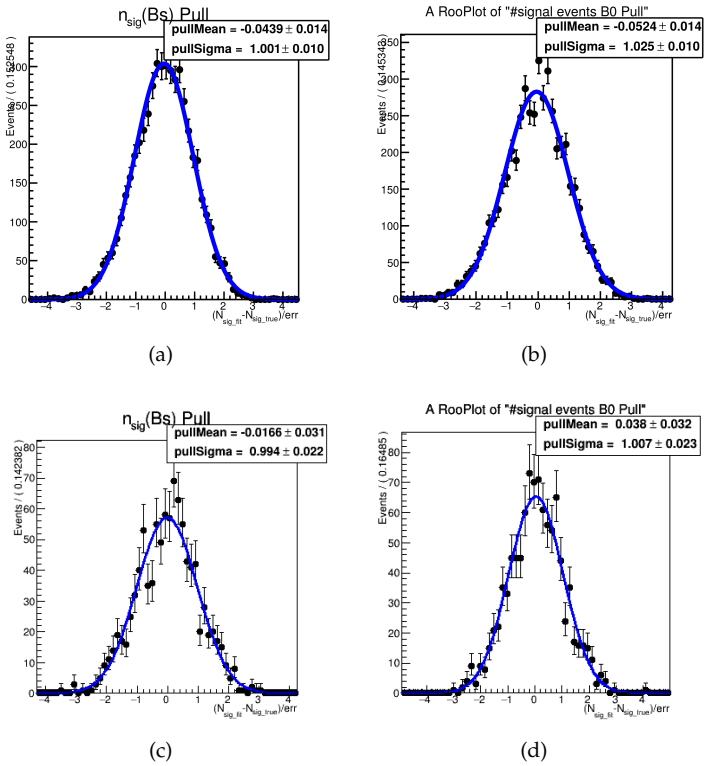


Figure 4.32 Pulls of signal yields for B_s^0 (a) and B^0 (b) in Run I and for B_s^0 (c) and B^0 (d) in Run II including into the fit model the scale factor α to account for detector resolution.

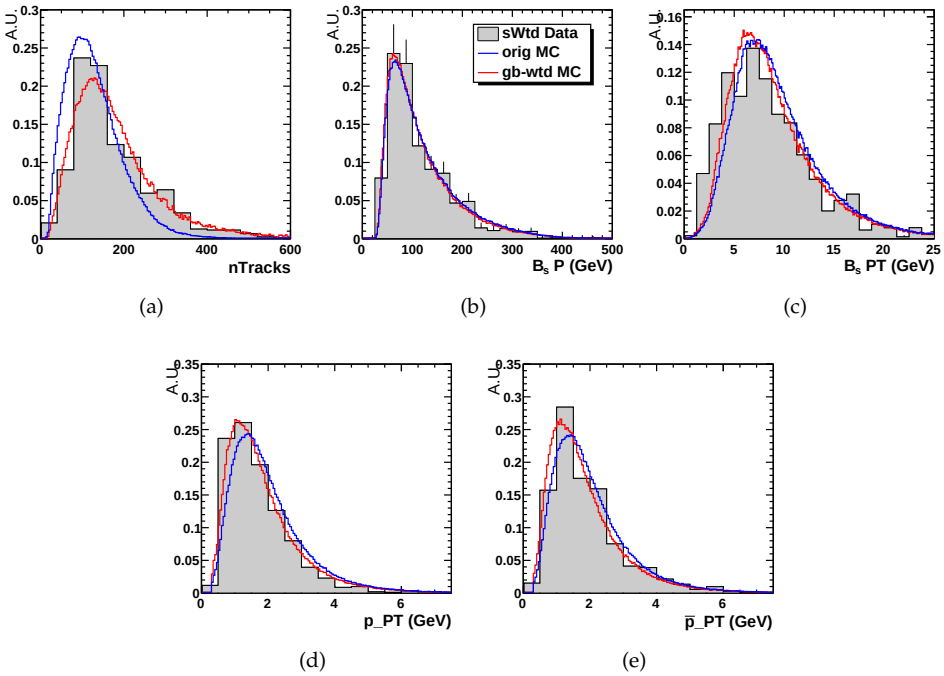
4.7.5 Tracking efficiency

The reconstruction efficiencies are determing using MC samples. A potential systematic effect is introduced if the tracking efficiencies present discrepancies between data and MC.

LHCb has studied and published $absolute$ `Long` tracking efficiencies for muons, pions and kaons tracks with a "probe-and-tag" method based on the decay $J/\psi \to \mu^+\mu^-$ [130]. This method uses two-prong decays, where one of the decay products, the "tag", is fully reconstructed as a `Long` track, while the other particle, the "probe", is only partially reconstructed, which means that the information related to one subdetector misses. The track reconstruction efficiency for `Long` tracks is then obtained by matching the partially reconstructed probe track to a long track. The ratio of the `Long` track reconstruction efficiency of muons in data and simulation was found to be compatible with unity and a systematic uncertainty of 0.4% for data taken in Run 1 is assigned. This uncertainty is taken as the difference of efficiency when reweighting the simulated samples to the data in different parameters such as the number of primary vertices, or the number of hits or tracks in different subdetectors.

- The first potential systematic effect is due to the difference in kinematic between the signal and the control mode. To determine the branching ratio only ratios of efficiencies are relevant, and mismodelling effects cancel to first order. In the case of our analysis, the difference between the control sample and the $B^0_{(s)} \to J/\psi p\bar{p}$ decays are two final kaons instead of two protons. To evaluate the difference in kinematic between two kaons and two protons we simulate 100 toy experiments varying the bin content of the correction tables randomly within statistical uncertainties. For each new efficiency table an average correction ratio is estimated by weighting each event for the efficiency correction. The systematic is taken as the width of the distribution of the average correction ratio and it is found to be 0.03% of the nominal correction ratio. Therefore it is negligible.

- The second systematic uncertainty regards the different hadronic interaction of the particles in the detector. The correction factors obtained from the muons of the $J/\psi \to \mu^+\mu^-$ decay do not take into account hadronic interactions. Following the work presented in the analysis of the $B^+ \to p\bar{\Lambda}$ decay [131] we assume that the corrections could vary by 1.1% for kaons, 1.4% for pions and 4% for protons with respect to the values obtained from muons. The uncertainty for protons is taken as the product of the uncertainty on the material budget times the number of protons that cannot be reconstructed due to hadronic interactions. The uncertainty on the material budget is 10% of the radiation lenght and the number of reconstructed protons is 20% as evaluated by simulations using Al, Be and Si as materials. The final uncertainty is multiplied by a genereous factor of 2 taking into account interactions with different materials which are not simulated ($2 \times 10\% \times 20\% = 4\%$). In order to consider this effect, 100 toy simulations are performed varying the bin content of the efficiency tables within 4% of the efficiency ratio for protons and 1.1% for kaons. The correction are considered as fully correlated along the phase-space. As previously, an average correction is computed multiplying each event for the ratio in tracking efficiency between data and Monte Carlo. The systematic is the spread of the average correction distribution and it is found to be 5% of the efficiency ratio. The same systematic is considered for Run 2.



(a) (b) (c)

(d) (e)

Figure 4.33 Comparisons between final sWeighted $B_s^0 \to J/\psi p\bar{p}$ data (combined Run I+Run II in filled grey) and signal MC. In red are the `gb-reweighted` normalized distributions and blue are the non-reweighted versions.

4.7.6 Simulation weighting

An important systematic is due to the reweighting procedure. As described in Sec. 4.2.1, the $B^0_{(s)} \to J/\psi p\bar{p}$ signal MC is reweighted using the weights extracted from the $B^0_s \to J/\psi\phi$ data-MC reweighting, normalized to the number of events in the $B^0_{(s)} \to J/\psi p\bar{p}$ signal MC sample. This procedure is based on the assumption that the relative difference between data and MC is the same between the signal and control modes in the most discriminating variables (`B_P`, `B_PT`, `nTracks`). To validate this assumption, another control mode, $B^0 \to J/\psi K^*$, is considered, which has the same 4-body topology; the weights are extracted from the reweighting of MC to data. The difference in the efficiency between using the weights from $B^0_s \to J/\psi\phi$ and $B^0 \to J/\psi K^*$ control modes, is assigned as the systematic. This is found to be 0.3% of the total average efficiency for B^0_s and 0.4% for B^0.

The small systematic validates the assumption that the relative data/MC corrections are reasonably consistent among different "J/ψ+di-hadron" topologies.

As an additional cross-check, comparisons between the final $sWeighted\ B^0_s \to J/\psi p\bar{p}$ distributions and the signal MC, after all cuts, are performed, as shown in Fig. 4.33. The weighted MC and data are in good agreement within statistical uncertainties, ensuring that the reweighting procedure is performing well.

Table 4.21 Cross-checks for $B^0_s \to J/\psi\, p\bar{p}$ and $B^0 \to J/\psi\, p\bar{p}$ branching fractions.

Type	$\mathcal{B}(B^0 \to J/\psi\, p\bar{p})$	$\mathcal{B}(B^0_s \to J/\psi\, p\bar{p})$
2011 Only	$(3.4 \pm 1.5) \times 10^{-7}$	$(4.1 \pm 0.8) \times 10^{-6}$
2012 Only	$(5.4 \pm 1.1) \times 10^{-7}$	$(3.8 \pm 0.6) \times 10^{-6}$
2016 Only	$(4.0 \pm 0.7) \times 10^{-7}$	$(3.7 \pm 0.3) \times 10^{-6}$
MU Only	$(5.4 \pm 1.0) \times 10^{-7}$	$(4.0 \pm 0.5) \times 10^{-6}$
MD Only	$(3.6 \pm 1.0) \times 10^{-7}$	$(3.9 \pm 0.5) \times 10^{-6}$
Average over phsp Run 1	$(4.5 \pm 0.7) \times 10^{-7}$	$(3.5 \pm 0.3) \times 10^{-6}$
Average over phsp Run 2	$(4.1 \pm 0.8) \times 10^{-7}$	$(3.7 \pm 0.3) \times 10^{-6}$

4.7.7 Multiple candidates

As described in Sec. 4.3.2, the multiple candidates for the signal modes are removed with a random choice. In order not to introduce a bias we decided to perform the selection with a different criteria and assign the shift in the yield as a systematic uncertainty. Since the BDT is uncorrelated to the mass the alternative criteria could be to select the candidate with highest BDT value. In this way the final yields for B^0_s and B^0 are shifted by 0.1%, proving that the effect is small.

4.7.8 Fragmentation fraction

The uncertainty on the fragmentation factor is considered as follows.

For the B^0_s meson, the uncertainty of 5.8%, taken from $f_s/f_d = 0.259 \pm 0.015$ [128], is considered, since the external normalization measurement for the B^0 meson, $\mathcal{B}(B^0_s \to J/\psi\,\phi) \times f_s/f_d$, is divided by the f_s/f_d ratio to obtain the B^0_s normalization, $\mathcal{B}(B^0_s \to J/\psi\,\phi)$. While the uncertainty on the f_s/f_d scale factor for Run 2 of 4.3% is applied to the B^0 meson because the B^0 normalization is measured using Run 1 data and therefore it has to be multiplied by the additional scale factor in Run 2.

4.7.9 Further run-condition cross-checks

We perform the analysis separately to 2011/2012 data and to magnet down/up data in Run1 and a check with only 2016 data for Run 2. The branching ratios are also calculated using phase-space-averaged efficiencies to validate the event by event efficiency procedure. The results of these cross-checks are listed in Table 4.21 and are found to be consistent within uncertainties.

4.8 $B^0_{(s)}$ mass measurements

Table 4.22 Systematic uncertainties of B^0 and B^0_s mass measurements.

	B^0	B^0_s
	[MeV]	[MeV]
Momentum scale	0.097	0.124
Mass fit model	0.020	0.020
Energy loss correction	0.030	0.030
Total	0.103	0.129

The small Q-values of the $B^0_{(s)} \to J/\psi p\overline{p}$ decays, corresponding to 306 (393) MeV for B^0 (B^0_s), allow for precise measurements of the B^0 and B^0_s masses, with a resolution of 3.3 MeV (3.8 MeV) for the B^0 (B^0_s) meson. The final results are

$$m_{B^0} = 5279.74 \pm 0.30\ (\text{stat}) \pm 0.10\ (\text{syst})\,\text{MeV},$$
$$m_{B^0_s} = 5366.85 \pm 0.19\ (\text{stat}) \pm 0.13\ (\text{syst})\,\text{MeV},$$

with a correlation of 3.813×10^{-4}. These represented the most precise single measurements for the B^0 and B^0_s masses at the time of the publication.

The sources of systematic uncertainties include uncertainties on particle interactions with the detector material, momentum scaling due to imperfections in the magnetic-field mapping and the choice of the signal model, as reported in Table 4.22. The systematic contributions are described in detail in the following.

Energy loss

We assign a 0.03 MeV systematic from Ref. [132] in the energy loss correction due to uncertainty of the material interaction lengths in the simulation. This is considered the same for both particles.

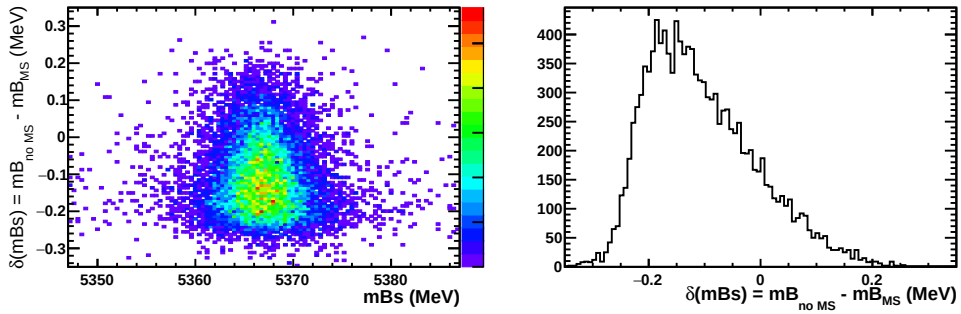
Momentum scale

To correct for various imperfections in the understanding of the detector (magnetic field maps, material interactions, *et al.*) the measured momenta of charged tracks require additional scaling [133]. Using the well-known masses of particles, such as B^+, K^0_S, J/ψ, $\Upsilon(2S)$, *et al.*, as "standard candles" for calibration, the shift from the PDG mass value, δm, is found to roughly scale with the Q-value (difference in mass between mother and daughters) of the decay as

$$|\delta m/Q| < \alpha \sim 0.0003, \tag{4.42}$$

where $\alpha \sim 0.0003$ is the upper limit of the scaling factor. This behavior is also found to be consistent between Run 1 and Run 2 of data taking.

To check the effect, a small sample of 2011 $B^0_s \to J/\psi p\overline{p}$ MC is processed with and without the momentum scaling (MS) of $\alpha = +0.03\%$, by correcting the final state momenta p, as $(1+\alpha)p$. The B^0_s mass is reconstructed from `DTF` with the J/ψ mass-constrained. Figure 4.34 shows the shift due to the scaling. The value of $\langle|\delta m|\rangle = 0.124$ MeV is consistent with the expectation corresponding to a $Q \sim 393$ MeV for the B^0_s decay and it is, therefore, taken as the upper limit of the expected systematic due to



(a) (b)

Figure 4.34 Effect of momentum scaling (MS) with $\alpha = +0.0003$ on truth-matched MC. The shift, δm, in the reconstructed B^0_s mass depends on the sign of α. The average shift from (b) is $|\langle \delta m \rangle| = 0.105$ MeV, while the average of the absolute shift is $\langle |\delta m| \rangle = 0.124$ MeV.

momentum scaling. For the B^0 mass, the B^0_s systematic is scaled by the ratio of Q-values of the B^0 and B^0_s decays ($B^0{}_{Q\text{-value}} \sim 306$ MeV). The resulting systematics is 0.097.

Fit model

As described in Sec. 4.7.1, 1000 toy experiments are performed generating the mass spectrum with the nominal PDF and fitting it with the alternative model. The pull of the mean of B^0_s mass is assigned as the systematic value. The corrected mean value is estimated as follows

$$m^c_{fit} = m_{fit} - \mu_p \sigma_{m_{fit}} \tag{4.43}$$

where m is the mean of the mass and μ_p is the mean of the pull. The uncertainty is the difference with respect to the nominal value of the mass and it is found to be 0.02 MeV. The same uncertainty is assigned to the B^0 mass.

4.9 Results

In conclusion, using the data sample collected during Run 1 + Run 2 (till 2016) by the LHCb experiment corresponding to an integrated luminosity of 3 fb^{-1}+ 2 fb^{-1}, we obtained the first observation of the $B^0_{(s)} \to J/\psi\, p\overline{p}$ rare decays.

The combination of the measurements is taken as a weighted mean to give the branching fraction ratios

$$\frac{\mathcal{B}(B^0 \to J/\psi\, p\overline{p})}{\mathcal{B}(B^0_s \to J/\psi\, \phi(\to K^+K^-)) \times f_s/f_d} = (0.329 \pm 0.029\ (\text{stat}) \pm 0.022\ (\text{syst})) \times 10^{-2},$$
$$\frac{\mathcal{B}(B^0_s \to J/\psi\, p\overline{p})}{\mathcal{B}(B^0_s \to J/\psi\, \phi(\to K^+K^-))} = (0.706 \pm 0.037\ (\text{stat}) \pm 0.048\ (\text{syst})) \times 10^{-2},$$

where the first uncertainty is statistical and the second is systematic. The results are dominated by the statistical uncertainties, while the most significant systematics come from the PID correction and the tracking efficiency.

The absolute branching fractions are then combined to give

$$\mathcal{B}(B^0 \to J/\psi\, p\overline{p}) = (4.51 \pm 0.40\ (\text{stat}) \pm 0.44\ (\text{syst})) \times 10^{-7},$$
$$\mathcal{B}(B^0_s \to J/\psi\, p\overline{p}) = (3.58 \pm 0.19\ (\text{stat}) \pm 0.33\ (\text{syst})) \times 10^{-6},$$

where the systematic uncertainty is the sum in quadrature of the overall systematic contribution on the ratio of branching fractions, the normalization mode uncertainty and the f_s/f_d uncertainty for the B^0_s signal, reported in Table 4.19.

Our results are compatible with the previous LHCb Run 1 (2011) analysis [53] with 2011 data, where only upper limits were placed on the $B^0_{(s)} \to J/\psi\, p\overline{p}$ branching fractions. We also confirm the trend that while the branching fraction for $B^0 \to J/\psi\, p\overline{p}$ is consistent with naive theory expectations [2], for $B^0_s \to J/\psi\, p\overline{p}$ is around two orders of magnitude larger. An interesting possibility behind this enhancement could be presence of glueballs in the $[p\overline{p}]$ system or pentaquarks in the $[J/\psi\, p, J/\psi\, \overline{p}]$ systems, which motivate more advanced studies reported in Chapter 5.

We also obtained the most precise single measurement of B^0_s and B^0 masses due to the low Q-value of the reactions. The masses are measured to be

$$m(B^0) = (5279.74 \pm 0.30\ (\text{stat}) \pm 0.10\ (\text{sys}))\ \text{MeV},$$
$$m(B^0_s) = (5366.85 \pm 0.19\ (\text{stat}) \pm 0.13\ (\text{sys}))\ \text{MeV}.$$

CHAPTER 5

Amplitude analysis of flavour untagged $B_s^0 \to J/\psi\, p\bar{p}$ decays

In recent years, the field of exotic states has gained an increasing interest, thanks especially to the discovery of pentaquark states made by LHCb in the decay of $\Lambda_b \to J/\psi\, pK^-$, as extensively detailed in Sec. 1.4. The nature of these states is still unknown, while there are many theoretical interpretations due to the complexity of QCD-inspired models. For these reasons, new confirmation of pentaquark states in independent channels are needed to shed light on their nature and help discriminate among different models.

In this framework, the $B^0_{(s)} \to J/\psi\, p\bar{p}$ decays are particularly attractive for pentaquarks decaying to the $J/\psi\, p$ final state. The B_s^0 mode is more interesting for this study with respect to the B^0 mode since it has a wider phase space for final state particles. Contrarily to the Λ_b decay, where plenty of Λ^* resonances occur in the Kp final states, this channel offers a "cleaner" environment, due to the absence of known resonances in both $p\bar{p}$ and $J/\psi\, p(\bar{p})$ final states. In particular, the B_s^0 decay could be sensitive to the ground state (spin $\frac{1}{2}$) pentaquark, $P_c^+(4312)$ [32], and the broad $P_c^+(4380)$, which both lie in the allowed $m(J/\psi\, p\bar{p})$ mass range. In addition, for the B_s^0 mode a possible intermediate glueball resonance $f_J(2220)$ decaying to $p\bar{p}$ final state can contribute, as predicted by Hsiao [2]. This is the equivalent baryonic decay of $B_s^0 \to J/\psi\, f_0(980)(\to \pi^+\pi^-)$ that has been studied at LHCb [56] earlier.

After the observation of the $B_s^0 \to J/\psi\, p\bar{p}$ decays reported in Chapter 4, a full amplitude analysis is performed in order to search for pentaquark and glueball existence and access their masses, widths and quantum numbers.

The Chapter is organized as following: in Sec. 5.1, the implementation of the amplitude formula to describe the $B_s^0 \to J/\psi\, p\bar{p}$ decay is given; in Sec. 5.5 the fit strategy is described, followed by Sec. 5.6 and 5.7, where the fit to data is performed. Then, the statistical analysis needed for the significance determination is detailed in Sec. 5.8. Finally, systematic uncertainties and results are summarized in Sec. 5.9 and 5.10.

5.1 Amplitude model

5.1.1 Four-body phase space

First of all, the $B_s^0 \to J/\psi\, p\bar{p}$ decay can be described by four independent variables, which are calculated assuming the conservation of momenta and the energy-mass relation. In general, a four body decay has 5 degrees of freedom (dof), resulting from 16 4-momentum components constrained by 4 mass requirements, 4 energy-momentum conservation relations and three Euler angles, that define the mother particle decay plane.

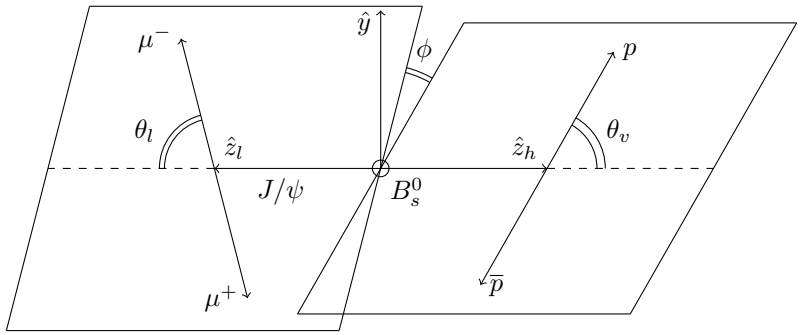


Figure 5.1 Definition of the $B_s^0 \to J/\psi\, p\overline{p}$ phase space variables.

The J/ψ mass constraint is then subtracted to them to end up with 4 degrees of freedom. Since the B_s^0 is a pseudoscalar, *i.e.* with spin zero, there is no preferred direction in space. Therefore, the initial reference frame is chosen as the B_s^0 rest system, directly.

The decay plane reference frame is chosen following the scheme in Fig. 5.1: the z-azis is aligned to the quantisation axis of $p\overline{p}$ system, while the $\hat{x}$ axis is defined in the $p\overline{p}$ plane by the direction orthogonal to the $p\overline{p}$ momentum defined by using the proton momentum,

$$\begin{aligned}
\hat{z}_B &= \hat{p}_X, \\
\hat{x}_B &= \frac{\hat{p}_p - [\vec{p}_p \cdot \vec{p}_X]\vec{p}_X}{|\vec{p}_p - [\vec{p}_p \cdot \vec{p}_X]\vec{p}_X|} \\
&= \frac{\vec{p}_X \times \vec{p}_p}{|\vec{p}_X \times \vec{p}_p|} \times \vec{p}_X, \\
\vec{y}_B &= \hat{z}_B \times \hat{x}_B \\
&= \frac{\vec{p}_X \times \vec{p}_p}{|\vec{p}_X \times \vec{p}_p|},
\end{aligned} \tag{5.1}$$

where the momenta are computed in the B_s^0 rest frame. The $p\overline{p}$ momentum is defined as: $\hat{p}_X = \frac{\vec{p}_X}{|\vec{p}_X|} = \frac{\vec{p}_p + \vec{p}_{\overline{p}}}{|p_p + p_{\overline{p}}|}$ [1].

With this definition, the phase space angles describe the rotation from the B_s^0 reference frame to the dihadron and dilepton helicity frames. The phase space is defined by: an azimuthal angle between the decay planes, $\phi \in (-\pi, \pi]$, and two helicity angles, θ_v and θ_l, computed in the dihadron and dilepton rest frames, respectively, as:

$$\begin{aligned}
\theta_v &= \text{acos}(\hat{p}_p^{\{X\}} \cdot \hat{z}_B), \\
\theta_l &= \text{acos}(\hat{p}_{\mu^-}^{\{\psi\}} \cdot \hat{z}_B), \\
\phi &= \text{atan2}(\hat{p}_{\mu^-}^{\{\psi\}} \cdot \hat{y}_B, \hat{p}_{\mu^-}^{\{\psi\}} \cdot \hat{x}_B).
\end{aligned} \tag{5.2}$$

where $\hat{p}_p^{\{X\}}$ and $\hat{p}_{\mu^-}^{\{\psi\}}$ are the proton and the μ^- momenta computed in the X and J/ψ rest frame, respectively. The J/ψ will be references as ψ throughout the thesis.

[1] The symbol $\hat{p}$ refers to the normalized momentum: $\hat{p} = \frac{\vec{p}}{|\vec{p}|}$

The four phase space variables are therefore chosen to be:

$$\Phi = (\cos\theta_v, \cos\theta_l, \phi, m_{p\bar{p}}) \tag{5.3}$$

so that the phase space is flat in all variables. In order to render the $m_{p\bar{p}}$ invariant mass flat, it is weighted by the inverse of the phase-space element, $1/d\Phi$, defined in Eq. 3.31.

5.1.2 Helicity amplitudes for $B_s^0 \to J/\psi\, p\bar{p}$ decay

The amplitude model is built using the helicity formalism described in Chapter 3. For the $B_s^0 \to J/\psi p\bar{p}$ decay, intermediate states can contribute in all the three two-particle systems: the $B_s^0 \to J/\psi\, X(\to)p\bar{p}$ decay chain, where X could be either a resonant or non-resonant contribution, the $B_s^0 \to P_c^+(\to J/\psi p)\bar{p}$ and $B_s^0 \to pP_c^-(\to J/\psi\bar{p})$ decay chains, where $P_c^\pm$ are the positive and negative pentaquarks. Therefore, the three possibilities are explained in details. Afterwards, a generalized formula for the total amplitude considering all contributions is given.

In a preliminary version of this analysis, a wrong implementation of the helicity formalism has been made, by following the work presented in the pentaquark analysis of $\Lambda_b \to J/\psi pK^-$ decays published in 2015 [21]. The issue was related to an incorrect alignment of the helicity state of the proton between the X and the Λ^* chains, which made the interference between resonances in the two chains equal to zero, which is unphysical. The same cancellation occurred also in the present analysis between the X and P_c^- chains. The origin of it is explained in the following.

Different solutions have been proposed to solve this issue [112, 134]. In particular, we decided to write the amplitudes in a similar fashion with respect to the pentaquark analysis, but following the prescription used in the "Dalitz-plot decomposition for three-body decays" (DPD) method presented in Ref. [112], where it was suggested to use the cyclic ordering of the final particles to define the helicity angles. The ordering is important in order to guarantee a proper spin matching of final particles, having spin of half-integer, between different decay chains, as discussed in Sec. 3.3.1. This ensures that all decay sequences share the same y-axis instead of the opposite, simplifying the alignment of the final particle spin states.

Hence, for the decay $B_s^0 \to J/\psi(1)p(2)\bar{p}(3)$, the angles need to be defined with respect to the first final particle in the 2-body decay:

- particle 1 for the decay of (12)
- particle 2 for the decay of (23)
- particle 3 for the decay of (31).

Therefore, the $p\bar{p}$ resonance uses the p direction, the P_c^+ chain the J/ψ direction, while the P_c^- the $\bar{p}$ direction. The same care must be taken for the J/ψ decay in each chain: the angles need to be properly defined in order to align the y-axis in different chains.

In order to show why the interference between different chains was vanishing in the first implementation of this analysis as well as in the $\Lambda_b \to J/\psi pK$ pentaquark analysis, we can simply relate the amplitude defined with the usage of the cyclic particle ordering to the amplitude without it. Let's consider the case of the P_c^- chain and relate the helicity angles computed using the ψ direction to the $\bar{p}$ ones. The polar angle changes as $\theta_\psi = \pi - \theta_{\bar{p}}$, contributing to the amplitude a term proportional to the Wigner-d, $d^{J_{P_c}}_{\lambda_{P_c},\lambda_\psi-\lambda_{\bar{p}}}(\pi-\theta_{\bar{p}})$,

for all values of $\phi_{\overline{p}}$. While, in order to satisfy ϕ belonging to $[-\pi, \pi]$, the azimuthal angle ϕ_ψ is defined as follows:

$$\phi_\psi = \begin{cases} \phi_{\overline{p}} + \pi, & \text{if} \quad \phi_{\overline{p}} < 0 \\ \phi_{\overline{p}} + \pi - 2\pi, & \text{if} \quad \phi_{\overline{p}} > 0. \end{cases} \tag{5.4}$$

contributing to the amplitude a term proportional to $e^{-i\phi_\psi \lambda_{P_c}}$. The difference of -2π between positive and negative $\phi_{\overline{p}}$ angles results in a negative sign for half-integer spin particles, due to the phase factor $e^{-i2\pi\lambda_{P_c}} = -1$. Since the interference is an odd function of the ϕ_ψ angle due to the negative sign, it cancels integrating over ϕ. Whereas, if we adopt the cyclic particle ordering the interference is recovered.

In addition, as mentioned in Sec. 3.2.3, in order to align different chains the particle-1 convention needs to be restored for the state of particle 2 in each two-body decay, by inserting the phase factor $(-1)^{J_2 - \lambda_2}$.

Adding these prescriptions, the amplitude model presented below is found to be equivalent to the "Dalitz-plot decomposition" method, which amplitude formula is given in Sec. A.2.1. Another check is performed for a simplified model of a stable J/ψ (without the decay into $\mu^+\mu^-$): the amplitude model built with the helicity formalism is compared to the covariant tensor formalism [105, 106]. Also in this case, perfect agreement is found, as shown in Sec. A.2.2.

Let's start from the decay chain of $B_s^0 \to J/\psi X$. Denoting the J/ψ as ψ, the matrix element is obtained as the product of the amplitude for the weak decay of the B_s^0, the strong decay of $X \to p\overline{p}$ and the J/ψ electromagnetic decay as,

$$\mathcal{M}^X_{\lambda_p, \lambda_{\overline{p}}, \eta} = \sum_{\lambda_\psi = 0, \pm 1} \tilde{\mathcal{H}}^{B_s^0}_{\lambda_\psi, \lambda_X} \mathcal{R}(m^2_{p\overline{p}}) D^{*J_X}_{\lambda_X, \lambda_p - \lambda_{\overline{p}}}(\phi_p^{\{X\}}, \theta_p^{\{X\}}, 0) \tilde{\mathcal{H}}^X_{\lambda_p, \lambda_{\overline{p}}} D^{*1}_{\lambda_\psi, \eta}(\phi_\mu^{\{\psi\}}, \theta_\mu^{\{\psi\}}, 0), \tag{5.5}$$

where η is defined as the difference between the muon helicities $\lambda_{\mu^-} - \lambda_{\mu^+}$. The first Wigner-D matrix of the B_s^0 decay is omitted because the B_s^0 has spin zero, implying that, for conservation of total angular momentum, the helicity of the ψ is equal to the helicity of X ($|\lambda_X - \lambda_\psi| \leq 0$). Hence, for the B_s^0 weak decay only the helicity couplings $\tilde{\mathcal{H}}^{B_s^0}_{\lambda_\psi, \lambda_\psi}$ enters in the formula. As pointed out in Sec. 3.2.3, the helicity couplings defined here contain the particle-2 phase factor, $\tilde{H}_{\lambda_1, \lambda_2} = (-1)^{s_2 - \lambda_2} H_{\lambda_1, \lambda_2}$.

The weak decay of $X \to p\overline{p}$ is expressed by the term,

$$\mathcal{R}(m^2_{p\overline{p}}) D^{*J_X}_{\lambda_X, \lambda_p - \lambda_{\overline{p}}}(\phi_p^{\{X\}}, \theta_p^{\{X\}}, 0) \tilde{\mathcal{H}}^X_{\lambda_p, \lambda_{\overline{p}}}$$

where the $\phi_p^{\{X\}}$ and $\theta_p^{\{X\}}$ angles are the azimuthal and polar angles of the proton momentum in the X helicity frame. The x-axis of the B_s^0 RF is along the p direction and is chosen to satisfy $\phi_p^{\{X\}} = 0$. If a resonance is present, its energy dependence can enter through the $\mathcal{R}(m_{p\overline{p}})$ lineshape. Since the strong decay conserves parity, the helicity couplings are reduced to 2 independent couplings following relations in Eq. 3.23. For $J^P(X) = 1^-$, which is the best choice to fit the data as demonstrated in Sec. 5.6.1, there are 3 (2) independent couplings $\tilde{\mathcal{H}}^{B_s^0}_{\lambda_\psi, \lambda_X}$ ($\tilde{\mathcal{H}}^X_{\lambda_p, \lambda_{\overline{p}}}$) to fit, already reduced by the parity conservation in the strong decay of X. Those couplings can be related to the LS couplings

$(B_{L,S})$ using the relation derived in Eq. 3.16:

$$\mathcal{H}^{A\to BC}_{\lambda_B,\lambda_C} = \sum_L \sum_S \sqrt{\frac{2L+1}{2J_A+1}} B_{L,S} \left(\begin{array}{cc|c} J_B & J_C & S \\ \lambda_B & -\lambda_C & \lambda_B - \lambda_C \end{array} \right) \times \left(\begin{array}{cc|c} L & S & J_A \\ 0 & \lambda_B - \lambda_C & \lambda_B - \lambda_C \end{array} \right) \tag{5.6}$$

where the terms in parenthesis are the Clebsh-Gordan coefficients. Here L is the orbital angular momentum in the decay, and S is the total spin of the daughters, $S = J_B + J_C$ ($|J_B - J_C| \leq S \leq J_B + J_C$). If the energy release in the decay, $Q = M_A - M_B - M_C$, is small, $Q/M_A << 1$, then higher values of L should be suppressed; this effect is usually called "the angular momentum barrier". The lowest L and S of the B_s^0 decay, $B^{B_s^0}_{L_{min}S_{min}}$, is always fixed to $(1, 0)$ for each contribution.

The electromagnetic decay contributes a term proportional to

$$D^{*1}_{\lambda_\psi,\eta}(\phi_\mu^{\{\psi\}}, \theta_\mu^{\{\psi\}}, 0),$$

where the y-axis of the ψ RF is chosen to be parallel to the y-axis in the B_s^0 RF to ensure a correct alignment; the angles $\theta_\mu^{\{\psi\}}$ and $\phi_\mu^{\{\psi\}}$ are the polar and azimuthal angles of the μ^- momentum in the ψ helicity frame. Since the ψ decay occurs through an electromagnetic interaction, the difference of the muon helicities restricts to the values: $\eta = \lambda_{\mu^-} - \lambda_{\mu^+} = \pm 1$. As the value of $\eta = 0$ is suppressed by m_μ/m_ψ for the electromagnetic transition, we omit it in the summation. The helicity coupling of the ψ decay is, therefore, ignored because it can be incorporated in the B_s^0 couplings. Indeed, the fit is not sensitive to a global factor.

Similarly, the matrix elements for the P_c^+ and P_c^- decay chains are given by

$$\mathcal{M}^{P_c^+}_{\lambda_p,\lambda_{\overline{p}},\eta} = \sum_{\lambda_\psi=0,\pm1} \tilde{\mathcal{H}}^{B_s^0\to P_c^+\overline{p}}_{\lambda_{P_c},\lambda_{\overline{p}}} \mathcal{R}(m^2_{J/\psi p}) D^{*J_{P_c}}_{\lambda_{P_c},\lambda_\psi-\lambda_p}(\phi_\psi^{\{P_c^+\}}, \theta_\psi^{\{P_c^+\}}, 0) \tilde{\mathcal{H}}^{P_c^+\to\psi p}_{\lambda_\psi,\lambda_p}$$
$$\cdot D^{*1}_{\lambda_\psi,\eta}(\phi_\mu^{\{P_c^+\}}, \theta_\mu^{\{P_c^+\}}, 0) \tag{5.7}$$

$$\mathcal{M}^{P_c^-}_{\lambda_p,\lambda_{\overline{p}},\eta} = \sum_{\lambda_\psi=0,\pm1} \tilde{\mathcal{H}}^{B_s^0\to P_c^- p}_{\lambda_{P_c},\lambda_p} \mathcal{R}(m^2_{J/\psi\overline{p}}) D^{*J_{P_c}}_{\lambda_{P_c},\lambda_{\overline{p}}-\lambda_\psi}(\phi_{\overline{p}}^{\{P_c^-\}}, \theta_{\overline{p}}^{\{P_c^-\}}, 0) \tilde{\mathcal{H}}^{P_c^-\to\overline{p}\psi}_{\lambda_{\overline{p}},\lambda_\psi}$$
$$\cdot D^{*1}_{\lambda_\psi,\eta}(\phi_\mu^{\{P_c^-\}}, \theta_\mu^{\{P_c^-\}}, 0) \tag{5.8}$$

where, as above, for angular momentum conservation, the following relations between helicities hold: $\lambda_{P_c} = \lambda_{\overline{p}}$ and $\lambda_{P_c} = \lambda_p$ for the P_c^+ and P_c^- chains, respectively. The x-axes are chosen in order to have $\phi_\psi^{\{P_c^+\}} = 0$ and $\phi_{\overline{p}}^{\{P_c^-\}} = 0$. The angle $\theta_\psi^{\{P_c^+\}}$ ($\theta_{\overline{p}}^{\{P_c^-\}}$) is the polar angle of ψ ($\overline{p}$) momentum in the P_c^+ (P_c^-) helicity frame. While $\theta_\mu^{\{P_c^\pm\}}$ and $\phi_\mu^{\{P_c^\pm\}}$ are the polar and azimuthal angles of the μ^- momentum in the ψ helicity frame in the two decay chains. The helicity couplings of the strong decay, $\mathcal{H}^{B_s^0\to P_c^+\overline{p}}_{\lambda_\psi,\lambda_p}$, are related by the following parity relations:

$$\mathcal{H}^{B_s^0\to P_c^+\overline{p}}_{-\lambda_\psi,-\lambda_p} = \eta_\psi\eta_p\eta_{P_c}(-1)^{J_{P_c}-(s_\psi+s_p)} \cdot \mathcal{H}^{B_s^0\to P_c^+\overline{p}}_{\lambda_\psi,\lambda_p} \tag{5.9}$$

$$= -\eta_{P_c}(-1)^{J_{P_c}-1/2} \cdot \mathcal{H}^{B_s^0\to P_c^+\overline{p}}_{\lambda_\psi,\lambda_p}, \tag{5.10}$$

where η_{P_c} is the intrinsic parity of the P_c state and η_ψ, η_p are the ψ and proton parities, equal to -1 and 1 respectively, which reduce the number of independent couplings by a factor of 2.

Then, the helicity couplings expressed in the LS basis are reduced to the lowest L values allowed, through Eq. 5.6. So, we end up with 1 (2) independent coupling to fit for $J^P = 1/2^-, 3/2^\pm$ ($J^P = 1/2^+$) of the quantum numbers of P_c^+. Due to the limited statistics, the couplings of P_c^+ and P_c^- resonances are fixed to the same values, as explained in Sec. 5.1.4.

5.1.3 Lineshape for P_c states

A single resonant contribution in the P_c decay chains is parameterized by,

$$R(m) = B'_L\left(p, p_0, d\right)\left(\frac{p}{M_{B_s^0}}\right)^{L_{B_s^0}} \mathrm{BW}\left(m|M_0,\Gamma_0\right) B'_{L_{P_{cj}}}\left(q, q_0, d\right)\left(\frac{q}{M_0}\right)^{L_{P_{cj}}}, \quad (5.11)$$

where the relativistic Breit-Wigner (BW) formula, commonly used to describe the lineshape of a resonance, assuming that $J/\psi\, p(\overline{p})$ is the only decay channel of the pentaquark candidate, is the following

$$\mathrm{BW}\left(m|M_0,\Gamma_0\right) = \frac{1}{M_0^2 - m^2 - iM_0\Gamma(m)}, \quad (5.12)$$

which is a function of the invariant mass m and is dependent on the mass M_0 and width Γ_0 of the resonance. The form of the relativistic BW arises from the propagator of an unstable particle, which has a denominator of the form $p^2 - M_0^2 + iM_0\Gamma$ (where p is the four-momentum vector). It can be obtained by the non-relativistic distribution, $1/((E - E_0) - i\Gamma/2)$, assuming that $(E + E_0) \sim 2M_0 \sim 2m$ in the rest frame of the resonance:

$$\frac{1}{(E - E_0) - i\Gamma/2} = \frac{E + E_0}{E + E_0}\frac{1}{(E - E_0) - i\Gamma/2} \sim \frac{m}{(m^2 - M_0^2) - iM_0\Gamma} \quad (5.13)$$

where E and E_0 are the center-of-mass energy of the system and of the resonance, respectively.

The BW in Eq. 5.12 also includes the mass dependent width

$$\Gamma(m) = \Gamma_0\left(\frac{q}{q_0}\right)^{2L_{P_{cj}}+1}\frac{M_0}{m}B'_{L_{P_{cj}}}\left(q, q_0, d\right)^2. \quad (5.14)$$

Here, p is the momentum of the P_c resonance in the B_s^0 rest frame and q is the momentum of one of the decay product in the P_c rest frame. While the momenta p_0 and q_0 denote values of them at the resonance peak ($m = M_0$). The orbital momentum of the B_s^0 decay is denoted as $L_{B_s^0}$ and that of the P_c decay as $L_{P_{cj}}$. In order to properly account for the suppression due to higher values of L, the Blatt-Weisskopf coefficients, $B_L(p, p_0, d)$, are used together with the threshold barrier factor, p^L, where d is the hadron radius size fixed to $3\,\mathrm{GeV}^{-1} = 0.6$ fm in the nominal fit, as in Ref. [21], and varied in the systematic evaluation. The Blatt-Weisskopf coefficients arise from centrifugal barrier effects and are computed using semiclassical theory. They depend on the orbital momentum L and on the interacting radius, that ranges from 0.25 to 0.75 fm for the meson resonance decays and from 0.1 to 1 fm for the baryon decays. The most frequently used coefficients up to

order 5 are listed below:

$$B'_0\left(p,p_0,d\right) = 1 \tag{5.15}$$

$$B'_1\left(p,p_0,d\right) = \sqrt{\frac{1+\left(p_0 d\right)^2}{1+(pd)^2}}$$

$$B'_2\left(p,p_0,d\right) = \sqrt{\frac{9+3\left(p_0 d\right)^2+\left(p_0 d\right)^4}{9+3(pd)^2+(pd)^4}}$$

$$B'_3\left(p,p_0,d\right) = \sqrt{\frac{225+45\left(p_0 d\right)^2+6\left(p_0 d\right)^4+\left(p_0 d\right)^6}{225+45(pd)^2+6(pd)^4+(pd)^6}}$$

$$B'_4\left(p,p_0,d\right) = \sqrt{\frac{11025+1575\left(p_0 d\right)^2+135\left(p_0 d\right)^4+10\left(p_0 d\right)^6+\left(p_0 d\right)^8}{11025+1575(pd)^2+135(pd)^4+10(pd)^6+(pd)^8}}$$

$$B'_5\left(p,p_0,d\right) = \sqrt{\frac{893025+99225\left(p_0 d\right)^2+6300\left(p_0 d\right)^4+315\left(p_0 d\right)^6+15\left(p_0 d\right)^8+\left(p_0 d\right)^{10}}{893025+99225(pd)^2+6300(pd)^4+315(pd)^6+15(pd)^8+(pd)^{10}}},$$

If $L = 0$, the coefficient is equal to 1, $B'_0\left(p,p_0,d\right) = 1$, which means that the lineshape of a scalar resonance is not affected by the centrifugal barrier effects. These coefficients become relevant if the momentum of the daughter particle p in the center of mass system of the mother particle is in the range $0-2$ GeV, or equivalently for resonances close to threshold. While they become negligible ($B' \to 1$) for p greater than 3 GeV, as shown in Ref. [135].

For nonresonant (NR) contributions in the X chain, we set BW$(m) = 1$ and $M_{0(NR)}$ to the midrange mass.

5.1.4 Relation between P_c couplings

Since we are using an untagged sample, as will be discussed in Sec. 5.1.6, our model is symmetric in $J/\psi p$ and $J/\psi \overline{p}$ distribution. In principle, in case of high statistics sample, even if we cannot distinguish between P_c^+ and P_c^- resonances, we could be able to extract the ratio between the two states, as well as their relative phase.

However, due to the limited statistics, from the fit to data we can extract neither two different couplings for the P_c states nor a relative phase between them. A relation between the helicity couplings of P_c^+ and P_c^- is thus imposed

$$\begin{aligned}\mathcal{H}^{B_s^0\to P_c^- p}_{\lambda_p,\lambda_p} &= -\mathcal{H}^{B_s^0\to P_c^+ \overline{p}}_{\lambda_{\overline{p}},\lambda_{\overline{p}}} \\ \mathcal{H}^{P_c^-\to \overline{p}\psi}_{\lambda_{\overline{p}},\lambda_\psi} &= (-1)^{J_{P_c}-s_p-s_\psi}\mathcal{H}^{P_c^+\to \psi p}_{\lambda_\psi,\lambda_p}\end{aligned} \tag{5.16}$$

where the sign in the first line is an arbitrary phase of the weak B_s^0 decay, while the phase factor in the second raw comes from the permutation of the final helicities, as discussed in Ref. [104] (Eq. 5.53).

This choice is arbitrary and is made in order to simplify the model and obtain the same interference pattern for all J^P hypotheses, as discussed below. In order to achieve that, we project the flavour eigenstates onto the CP basis of B_L and B_H. For CP conservation, the B_s^0 and $\bar{B}_s^0$ can be written in terms of CP eigenstates, as

$$|B_s^0\rangle = |B_H\rangle + |B_L\rangle \tag{5.17}$$

$$|\bar{B}_s^0\rangle = |B_H\rangle - |B_L\rangle, \tag{5.18}$$

as well as the final states the mother B_s^0 decays into. Applying CP to the final states, we obtain:

$$\begin{aligned} CP|X\rangle &= +|X\rangle, \\ CP|P_c^+\rangle &= e^{i\phi}|P_c^-\rangle, \\ CP|P_c^-\rangle &= e^{i\phi}|P_c^+\rangle, \end{aligned} \tag{5.19}$$

where the notation used for the states represent the decay through the specific resonance; therefore, in terms of the final state, $|X\rangle = |J/\psi\, X\rangle$, $|P_c^+\rangle = |P_c^+\overline{p}\rangle$ and $|P_c^-\rangle = |P_c^- p\rangle$. Indeed, the contribution in the X chain is already a CP eigenstate, and in particular for $J^P = 1^{--}$ a CP-even one, while the two single P_c are not. However, a combination of P_c^+ and P_c^- contributions can be projected onto a CP eigenstate orthogonal to the X contribution (CP-odd) by choosing the specific phase convention: $e^{i\phi} = -1(-1)^{J_{P_c}-3/2}$, which is the product of the phases in Eq. 5.16. Hence, the application of CP results in:

$$CP(|P_c^+\rangle - (-1)^{J_{P_c}-3/2}|P_c^-\rangle) = -(|P_c^+\rangle - (-1)^{J_{P_c}-3/2}|P_c^-\rangle) \tag{5.20}$$

which spin dependent factor follows from Eq. 5.16. Therefore, the interference between the X and the combination of the P_c contributions cancels out once integrating over the Dalitz, as they are ortoghonal CP eigenstates, as demonstrated in Appendix A.4.2. The cancellation happens only after integration because of the orthogonality of the Wigner-D functions for opposite quantum numbers.

In conclusion, by projecting the states onto the CP basis, we have reduced the dof of the model by imposing a relation between the $P_c^\pm$ helicity couplings (Eq. 5.16), which allows to obtain the same interference pattern for all J^P hypotheses (for instance, zero interference after integration over the Dalitz). This assumption has been verified not to introduce any bias in the results, as reported in Appendix A.4.3.

5.1.5 Total amplitude

Before summing up the amplitudes in the three decay chains coherently, a rotation to a reference spin state in which the final particle helicities are preserved is needed. Indeed, the final state helicities are computed in different frames for each decay sequence. The spin rotation is introduced in Sec. 3.3.1 for a general three-body decay. Here, we analize in detail the decay of interest.

For simplicity, the main reference frames with respect to which the rotations are performed are assumed to be the ones reached in the X chain. For each final particle with spin, the reference spin frame is simply obtained by doing a pure boost from the B_s^0 to the particle rest frame. The final state helicity frames of $p, \overline{p}$ and μ in the P_c chains must thus be rotated to the reference spin states in the X chain.

In the following, we treat the case of the P_c^+, considering that the same expressions are valid for the P_c^-, after the transformation $p \leftrightarrow \bar{p}$.

Regarding the proton, the quantization axes are along the proton direction in the X and the P_c rest frames, thus anti-parallel to the particles recoiling against the proton: the $\overline{p}$ and ψ, respectively. These directions are preserved when boosting to the proton rest frame (see Fig. 5.2). Hence, the polar angle between the two proton quantization axes (θ_p) can be determined from the opening angle between the $\overline{p}$ and ψ mesons in the p rest frame:

$$cos\theta_p = \hat{p}_{\overline{p}}^{\{p\}} \cdot \hat{p}_{\psi}^{\{p\}} \tag{5.21}$$

Since the $\overline{p}$ and the ψ are in the same plane, the azimuthal angle is set to 0. The spin state of the proton can be expressed as a rotation of the old axis, as in Eq. 3.28:

$$|\lambda_p\rangle = \sum_{\lambda_p^{P_c}} d^{J_p}_{\lambda_p^{P_c},\lambda_p}(\theta_p)|\lambda_p^{P_c}\rangle \tag{5.22}$$

The same procedure can be followed for the $\overline{p}$: the alignment of the helicity axes occurs when rotating clockwise by the angle between the p and the P_c momentum in the $\overline{p}$ spin frame (Fig. 5.3):

$$cos\theta_{\overline{p}} = \hat{p}_p^{\{\overline{p}\}} \cdot \hat{p}_{P_c}^{\{\overline{p}\}}. \tag{5.23}$$

The rotation of the $\overline{p}$ helicity state is the following:

$$|\lambda_{\overline{p}}\rangle = \sum_{\lambda_{\overline{p}}^{P_c}} d^{J_{\overline{p}}}_{\lambda_{\overline{p}}^{P_c},\lambda_{\overline{p}}}(-\theta_{\overline{p}})|\lambda_{\overline{p}}^{P_c}\rangle \tag{5.24}$$

The sign of the rotation needs also to be accounted properly, as discussed in Sec. 3.3.1 and explained through Fig. 3.2. Here, we assume that the y-axis is outgoing from the decay plane. As done in Fig. 3.2, a simple way to visualize it is to draw the momenta in the RF where the rotation angles are computed, the proton RF and $\overline{p}$ RF, respectively. As shown in Fig. 5.4 for the P_c^+ chain, in the proton RF the momentum of ψ is parallel to the P_c^+ direction (hence opposite to the quantization axis in P_c^+ decay) and the momentum of $\overline{p}$ is parallel to that of X (opposite to the quantization of X chain). Since the y-axis is outgoing from the plane, the rotation to align the P_c^+ decay chain to the X decay chain is counterclockwise between the ψ and the $\overline{p}$ momenta. While for the $\overline{p}$ RF, the rotation from the direction of P_c^+ to that of p is clockwise, hence resulting in a negative sign. Following the same procedure, you find that for the P_c^- chain the rotation is counterclockwise in the proton RF and clockwise in the $\overline{p}$ RF.

For the J/ψ decay, since the muons come from the J/ψ for both the decay chain, the polar angle is 0, implying the following: $d^{1/2}_{\lambda_\mu^{P_c},\lambda_\mu} = \delta_{\lambda_\mu^{P_c},\lambda_\mu}$. However, there is an azimuthal angle, α_μ, because of the offset in the x axis. Since the boost to the μ rest frame is the same for both decay chains (i.e.always from the ψ rest frame), α_μ can be determined in the ψ rest frame

$$\alpha_\mu = \text{atan2}((\hat{z}_3^{\{\psi\}P_c} \times \hat{x}_3^{\{\psi\}P_c}) \cdot \hat{x}_3^{\{\psi\}X}, \hat{x}_3^{\{\psi\}P_c} \cdot \hat{x}_3^{\{\psi\}X}), \tag{5.25}$$

where the index 3 refers to the RF after rotations; $\hat{z}_3^{\psi} = \hat{p}_\mu^{\psi}$ and the $\hat{x}_3$ axis can be derived as:

$$\hat{x}_3^{\{\psi\}P_c} = -a_{z_0\perp\mu} \tag{5.26}$$

$$a_{z_0\perp\mu} = -\hat{p}_p^{\{\psi\}} + (\hat{p}_p^{\{\psi\}} \cdot \hat{p}_\mu^{\{\psi\}})\hat{p}_\mu^{\{\psi\}} \tag{5.27}$$

as well as

$$\hat{x}_3^{\{\psi\}X} = -a_{z_0\perp\mu} \tag{5.28}$$

$$a_{z_0\perp\mu} = -\hat{p}_X^{\{\psi\}} + (\hat{p}_X^{\{\psi\}} \cdot \hat{p}_\mu^{\{\psi\}})\hat{p}_\mu^{\{\psi\}} \tag{5.29}$$

The term aligning the muon helicity states between the two reference frames is given by

$$\sum_{\lambda_\mu} D^{J_\mu}_{\lambda_\mu^{P_c},\lambda_\mu}(\alpha_\mu, 0, 0) = \sum_{\lambda_\mu} e^{i\lambda_\mu^{P_c}\alpha_\mu}\delta_{\lambda_\mu^{P_c},\lambda_\mu} = e^{i\lambda_\mu\alpha_\mu} \tag{5.30}$$

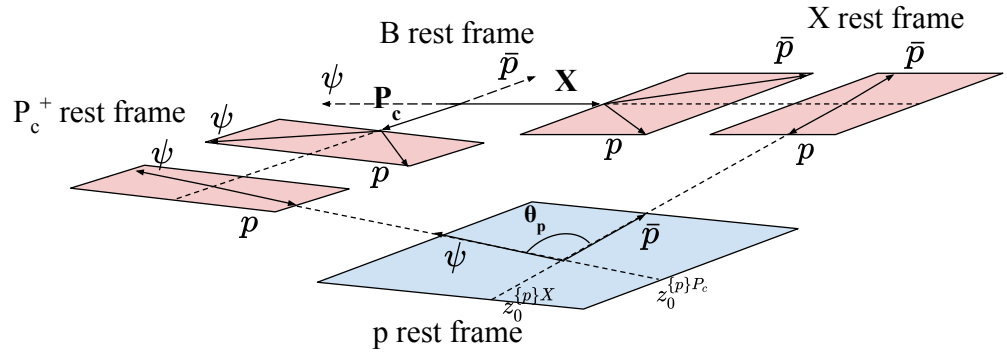


Figure 5.2 Definition of proton helicity angle θ_p.

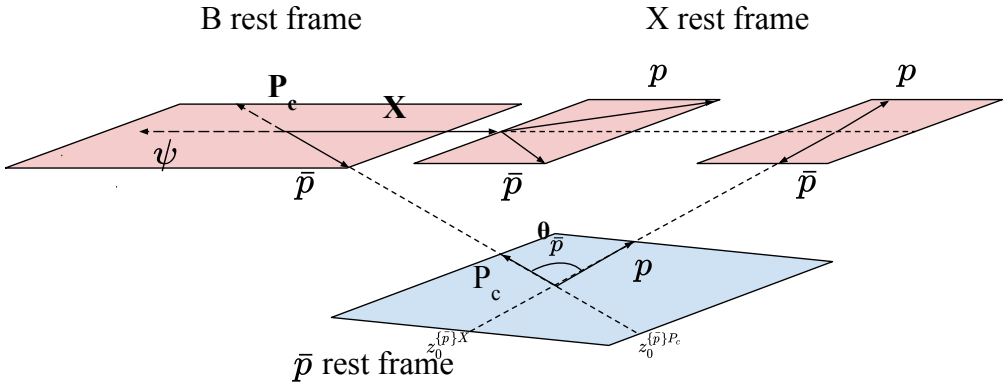


Figure 5.3 Definition of $\bar{p}$ helicity angle $\theta_{\bar{p}}$.

The transformation for the μ^- is equal to that of the μ^+ where the azimuthal angle takes a negative sign: $\alpha_{\mu^-} = -\alpha_{\mu^+}$. Considering the transformations of both muons, the final rotation is the multiplication of the two exponential

$$e^{i\lambda_\mu \alpha_\mu} e^{i\lambda_{\bar{\mu}} \alpha_{\bar{\mu}}} = e^{i\alpha_\mu(\lambda_\mu - \lambda_{\bar{\mu}})} = e^{i\alpha_\mu \Delta_{\lambda_\mu}}. \tag{5.31}$$

After the rotation of the final spin state, the total amplitude squared is obtained as the incoherent sum over the final helicities, as in Eq. 3.29:

$$|\mathcal{M}(\Phi)|^2 = \sum_{\lambda_p}\sum_{\lambda_{\overline{p}}}\sum_{\eta} | \mathcal{M}^X + e^{i\eta\cdot\alpha_\mu} \sum_{\lambda_p^{P_c}}\sum_{\lambda_{\overline{p}}^{P_c}} d^{1/2}_{\lambda_{\overline{p}}^{P_c},\lambda_{\overline{p}}}(-\theta_{\overline{p}}^{P_c^+}) d^{1/2}_{\lambda_p^{P_c},\lambda_p}(\theta_p^{P_c^+}) \mathcal{M}^{P_c^+}(\lambda_p^{P_c}, \lambda_{\overline{p}}^{P_c}, \eta) +$$
$$+ e^{i\eta\cdot\alpha_{\bar{\mu}}} \sum_{\lambda_p^{P_c}}\sum_{\lambda_{\overline{p}}^{P_c}} d^{1/2}_{\lambda_p^{P_c},\lambda_p}(\theta_p^{P_c^-}) d^{1/2}_{\lambda_{\overline{p}}^{P_c},\lambda_{\overline{p}}}(-\theta_{\overline{p}}^{P_c^-}) \mathcal{M}^{P_c^-}(\lambda_{\overline{p}}^{P_c}, \lambda_p^{P_c}, \eta) |^2 \tag{5.32}$$

where $\eta = \Delta\lambda_\mu^{P_c}$ and $\lambda_p, \lambda_{\overline{p}}$ stand for the p and $\bar{p}$ helicities.

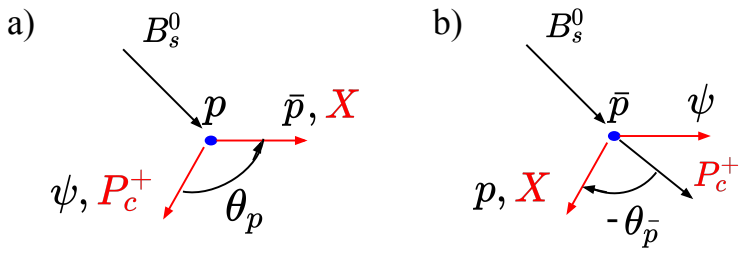


Figure 5.4 Sign of the rotation from the P_c^+ to the X decay chain in p RF (a) and $\overline{p}$ RF (b).

5.1.6 Flavour untagged sample

As explained in the introduction to measurements in the B_s^0 sector (Sec. 1.2), CP is a good symmetry for the decay under study. Indeed, for neutral B_s^0 decays where direct CP asymmetry cannot be observed and for untagged decay rate integrated in time, we have proved that no CP violation can be observed neither in mixing nor in the interference between mixing and decay, as derived in Sec. 1.2.3. Therefore, the total amplitude of B_s^0 meson should be equal to that of $\bar{B}_s^0$, apart from a relative phase, as derived in Eq. 5.32:

$$|\mathcal{M}^{\bar{B}_s^0}(\bar{\Phi})| = |\mathcal{M}^{B_s^0}(\Phi)|, \tag{5.33}$$

where $\mathcal{M}^{B_s^0}(\Phi)$ is the total amplitude reported in Eq. 3.29.

The $\bar{B}_s^0$ amplitude is obtained inverting particle with antiparticle, *i.e.* $p \leftrightarrow \overline{p}$, $\mu^- \leftrightarrow \mu^+$ due to C conjugation and inverting all azimuthal angles, $\phi \leftrightarrow -\phi$ for P inversion. In addition, under P all particle's helicities are inverted. Assuming these transformations, the expressions for the angles and the amplitudes in each of the three decay chains are reported below, to ensure that Eq. 5.33 is preserved.

First of all, in order to guarantee a proper spin matching the angles are defined using the following particle ordering: $J/\psi(1)p(2)\overline{p}(3)$. Hence, the y-axis is outgoing from the decay plane. Starting from the X chain, the helicity angles with respect to the B_s^0 ones are defined as:

$$\begin{aligned}
\theta_p^{\{X\}\bar{B}} &= \pi - \theta_p^{\{X\}B} \\
\theta_{\mu^+}^{\{X\}\bar{B}} &= \pi - \theta_{\mu^-}^{\{X\}B} \\
\phi_{\mu^+}^{\{X\}\bar{B}} &= -\phi_{\mu^-}^{\{X\}B}
\end{aligned} \tag{5.34}$$

and, knowing that P reverts the helicities, the amplitude becomes:

$$\begin{aligned}
A_X^{\bar{B}}(-\lambda_\psi, -\lambda_{\bar{p}}, -\lambda_p) &= (-1)^{J_X} \tilde{H}_{-\lambda_X,-\lambda_X} \cdot \tilde{H}_{-\lambda_p,-\lambda_{\overline{p}}} \cdot d^{J_X}_{-\lambda_X,-\lambda_{\bar{p}}-\lambda_p}(\theta_p^{\{X\}\bar{B}}) \\
&\quad \cdot d^1_{\lambda_\psi,\eta}(\theta_\mu) e^{i\lambda_\psi \phi_\mu} \\
&= (-1)^{J_X} A_X^B(\lambda_\psi, \lambda_p, \lambda_{\bar{p}}),
\end{aligned} \tag{5.35}$$

which is equal to Eq. 5.5. The factor $(-1)^{J_X}$ is the eigenvalue of the charge conjugation operator C on the X C-eigenstate and $\tilde{H}$ are the helicity couplings with the inclusion of the particle-2 phase-factor.

For the P_c chains, the angles of the P_c^+ are inverted with those of P_c^-, as:

$$\begin{aligned}\theta_{\psi}^{\{P_c^+\}\bar{B}} &= \pi - \theta_{\bar{p}}^{\{P_c^-\}B},\\ \theta_{\mu^+}^{\{P_c^+\}\bar{B}} &= \pi - \theta_{\mu^-}^{\{P_c^-\}B},\\ \phi_{\mu^+}^{\{P_c^+\}\bar{B}} &= -\phi_{\mu^-}^{\{P_c^-\}B},\end{aligned} \tag{5.36}$$

and analogously for the P_c^- chain:

$$\begin{aligned}\theta_{\bar{p}}^{\{P_c^-\}\bar{B}} &= \pi - \theta_{\psi}^{\{P_c^+\}B},\\ \theta_{\mu^+}^{\{P_c^-\}\bar{B}} &= \pi - \theta_{\mu^-}^{\{P_c^+\}B},\\ \phi_{\mu^+}^{\{P_c^-\}\bar{B}} &= -\phi_{\mu^-}^{\{P_c^+\}B}.\end{aligned} \tag{5.37}$$

Once the angles are properly defined the two amplitudes assume the same shape. Therefore, for the P_c^+ chain, the amplitude without the final rotation is:

$$\begin{aligned}A^{\{P_c^+\}\bar{B}}(-\lambda_\psi, -\lambda_{\bar{p}}, -\lambda_p) &= \tilde{H}_{-\lambda_p,-\lambda_p} \cdot \tilde{H}_{-\lambda_\psi,-\lambda_{\bar{p}}} \cdot d^{J_{Pc}}_{-\lambda_{Pc},-\lambda_\psi+\lambda_{\bar{p}}}(\theta_{\bar{p}}^{\{P_c^-\}\bar{B}})\\ &\cdot d^1_{\lambda_\psi,\eta}(\theta_\mu)\exp^{i\lambda_\psi\phi_\mu} =\\ &= A^{\{P_c^-\}B}(\lambda_\psi, \lambda_p, \lambda_{\bar{p}})\end{aligned} \tag{5.38}$$

which is equal to Eq. 5.8.

Regarding the final spin rotation, they remain the same as in Eq. 5.32, once the angles are defined as:

$$\begin{aligned}\theta_{p}^{\{P_c^+\}\bar{B}} &= \theta_{\bar{p}}^{\{P_c^-\}B},\\ \theta_{\bar{p}}^{\{P_c^+\}\bar{B}} &= \theta_{p}^{\{P_c^-\}B},\\ \alpha_\mu &= -\alpha_\mu\end{aligned} \tag{5.39}$$

In this way, the following relations between single amplitudes are obtained:

$$\begin{aligned}\mathcal{A}_{NR} &= \bar{\mathcal{A}}_{NR},\\ \mathcal{A}_+ &= \bar{\mathcal{A}}_-,\\ \mathcal{A}_- &= \bar{\mathcal{A}}_+,\end{aligned} \tag{5.40}$$

where $\mathcal{A}_+(\mathcal{A}_-)$ are the amplitudes of $B_s^0 \to P_c^+\bar{p}$ ($B_s^0 \to P_c^-p$), while $\bar{\mathcal{A}}_-$ and $\bar{\mathcal{A}}_+$ the ones for the $\bar{B}_s^0$ decay, $\bar{B}_s^0 \to P_c^-p$ and $\bar{B}_s^0 \to P_c^+\bar{p}$.

Finally, for untagged B_s^0 meson and time-integrated measurements, the total amplitude can be written as the average of the B_s^0 and $\bar{B}_s^0$ amplitudes since there is no interference between the two, as demonstrated in Sec. 1.2.3:

$$\frac{dN}{d\Phi} = \frac{1}{2}\left(\mid \mathcal{M}^{B_s^0}(\Phi) \mid^2 + \mid \mathcal{M}^{\bar{B}_s^0}(\Phi) \mid^2\right). \tag{5.41}$$

Following from Eq. 5.40 and 5.41, we can derive that our model is symmetric between $J/\psi\, p$ and $J/\psi\, \overline{p}$, because the contribution to each final state is given by the sum of the two P_c amplitudes, $\mathcal{A}_+ + \mathcal{A}_-$ (since $\mathcal{A}_+ + \bar{\mathcal{A}}_+ = \mathcal{A}_+ + \mathcal{A}_-$).

Finally, the differential decay rate can be expressed as:

$$d\Gamma = \frac{\frac{dN}{d\Phi}}{2m_B} d\Phi, \tag{5.42}$$

where $2m_B$ is the normalization factor, where m_B is the B_s^0 mass, and $d\Phi$ is the phase space element for a three body decay, defined as

$$d\Phi = pk \cdot dm_{p\bar{p}} \cdot d\cos\theta_l \cdot d\cos\theta_v \cdot d\phi \tag{5.43}$$

where p and k are the momenta of P_c and J/ψ in their mother particle RF, respectively.

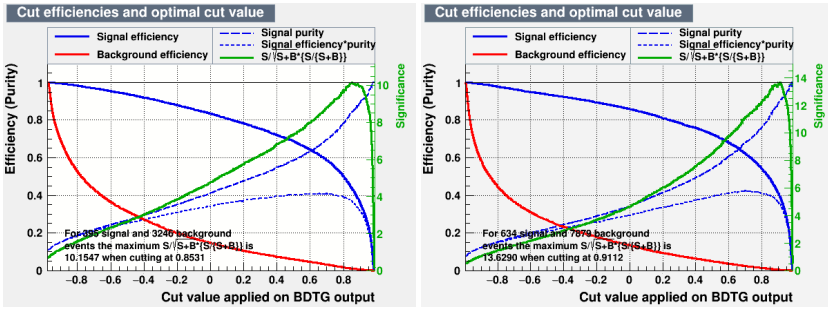


Figure 5.5 Signal efficiency and optimal cut value for Run 1 (left) and Run 2 (right).

5.2 Dataset and event selection

This analysis is based on 3 fb^{-1} data collected in Run 1 during 2011-2012 at the collision energy of 7 TeV and 8 TeV, respectively, and 6 fb^{-1} data collected in Run 2 during 2015-2018 at the energy of 13 TeV. The dataset of the analysis detailed in the first part of this thesis (Ch. 4) has therefore been expanded with the 2017 and 2018 data. The Monte Carlo samples used in the analysis are the same listed in Table 4.1.

The amplitude analysis exploits the same trigger and pre-selection requirements used for the observation of $B^0_{(s)} \to J/\psi p\bar{p}$ decays, described in Sec. 4.1.1. The optimization of the selection is also performed following the same strategy as in Sec. 4.3, using a BDT algorithm in two stages. A preliminary stage, which enables to see a signal peak is trained on the $B^0_s \to J/\psi\,\phi$ control sample, with kinematic variables in common between signal and control samples (Sec. 4.3.1). A second BDT exploits the PID information of the proton to obtain a more powerful selection and remove combinatorial background (Sec. 4.3.2). The BDT response trained on 2016 MC and sideband data is applied to the dataset collected during the latest years of data taking. Instead of maximising the significance to choose the optimal BDT 2 cut, as done for the observation of these modes, for the amplitude analysis a tighter selection has been exploited in order to further reduce the background. The significance times purity ($\mathcal{S}/\sqrt{\mathcal{S}+\mathcal{B}} \cdot \mathcal{S}/(\mathcal{S}+\mathcal{B})$) is used as Figure of Merit, where $\mathcal{S}$ and $\mathcal{B}$ are the signal and background yields extracted from a fit to the $m(J/\psi p\bar{p})$ distribution after the BDT 1 selection.

The final cuts on BDT variables are 0.4 (0.7) for BDT 1 and 0.853 (0.9112) for BDT 2 in Run 1 (Run 2). The optimization of the second BDT is shown in Fig. 5.5. If after the selection multiple candidates are still present, a single candidate is retained with a random selection.

5.3 Fit to $B^0_s \to J/\psi p\bar{p}$ data

The $m(J/\psi p\bar{p})$ invariant mass distribution for the events passing the full selection is shown in Fig. 5.6. A maximum likelihood fit is performed on data using a model composed by 2 Crystal Balls functions for the signal and a first order polynomial for the background. The fit returns the results summarized in Tab. 5.1. The resolution on the invariant masses is extracted from MC as shown in Fig. 5.8 and it is found to be 3.5 MeV for the B^0_s signal. The purity, computed in a region of 3σ around the B^0_s peak, reaches the level of 85%. In the full amplitude analysis, the signal is selected in the region of

3σ around the B_s^0 peak and the background is parametrized from the right sideband, as shown in Fig. 5.7. The fraction of background below the B_s^0 peak is fixed to 15%.

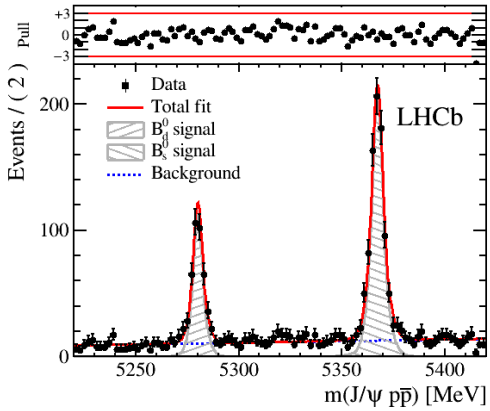


Figure 5.6 One-dimensional fit of the $m(J/\psi p\bar{p})$ invariant mass with 9 fb^{-1} datasample.

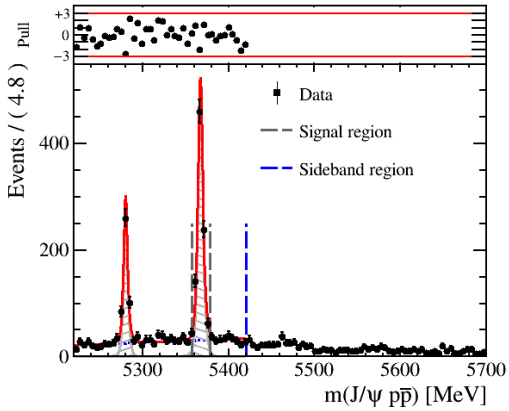


Figure 5.7 B_s^0 signal region delimited by grey dashed lines ($m_{B_s^0} = [5357, 5378]$ MeV) and sideband region above 5420 MeV.

5.4 Preliminary consideration

The invariant mass projections in Fig. 5.9 all indicate that there is some structure not described by the phase space distribution taken from MC sample. These structures are visible either in $p\bar{p}$ or in $J/\psi p$ and $J/\psi \bar{p}$ systems. In particular, the structures in $J/\psi p$ and $J/\psi \bar{p}$, if found to be significant, could be interpretated as pentaquark-like states. Indeed, LHCb discovered states compatible with pentaquarks in the $J/\psi p$ invariant mass in the decay of $\Lambda_b \to J/\psi K^- p$. So a new discovery could shed light on the theoretical interpretation for these states.

From 1D projections, it is not clear whether these structures are real states or are due to reflections from other resonances. For this reason, we will perform an amplitude analysis in order also to extract quantum numbers.

Table 5.1 Floating parameters of the fit on B_s^0 and B^0 reported in Fig. 5.6. $\mu_{B^0_{(s)}\to J/\psi\, p\overline{p}}$ are the mean values of B_s^0 and B^0 mass, while a_0 the parameter of the Chebyshev polynomial for background. $n(B_s^0)$ and $n(B^0)$ are the total yields for $B_s^0 \to J/\psi\, p\overline{p}$ and $B^0_{(s)} \to J/\psi\, p\overline{p}$ signal events.

Parameter	Value
$\mu_{B_s^0\to J/\psi\, p\overline{p}}$	5367.25 ± 0.14 MeV
$\mu_{B^0\to J/\psi\, p\overline{p}}$	5280.03 ± 0.18 MeV
a_0	0.22 ± 0.05
$n(B_s^0)$	797 ± 31
$n(B^0)$	371 ± 22

Before performing such analysis, it is checked that this peaking structure is not due to reflections from other b-decays, like $B_s^0 \to J/\psi\, KK$ or $B_s^0 \to J/\psi\, Kp$ and similar. After the selection, the presence of these mis-recontructed events is suppressed, as described in Sec. 4.1.4. In addition, the peaks cannot be due to efficiency effects related to the selection procedure, as clearly seen in the MC sample after selection, as reported in the efficiency plots in Fig. 5.11.

The resolution on the invariant mass of $J/\psi\, p$ and $J/\psi\, \overline{p}$ is computed from MC as the difference between reconstructed and generated invariant mass values as shown in Fig. 5.8, as an example. The resolution is computed as the RMS of the distribution and it is equal to 3.5 MeV. The resolution in Run 2 is compatible and it is measured to be 3.3 MeV. We also checked that the B_s^0 mass constraint does not improve the resolution significantly, so we do not apply it in the selection.

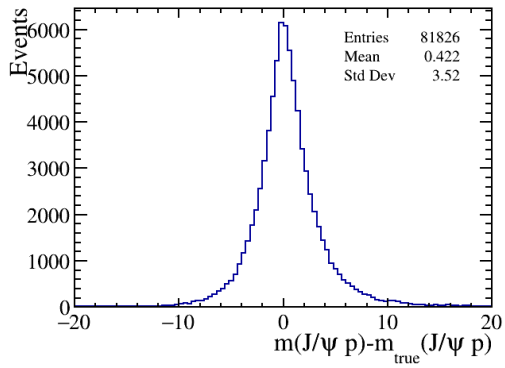


Figure 5.8 Resolution on the $J/\psi\, p$ invariant mass from $B_s^0 \to J/\psi\, p\overline{p}$ decays in Run I MC.

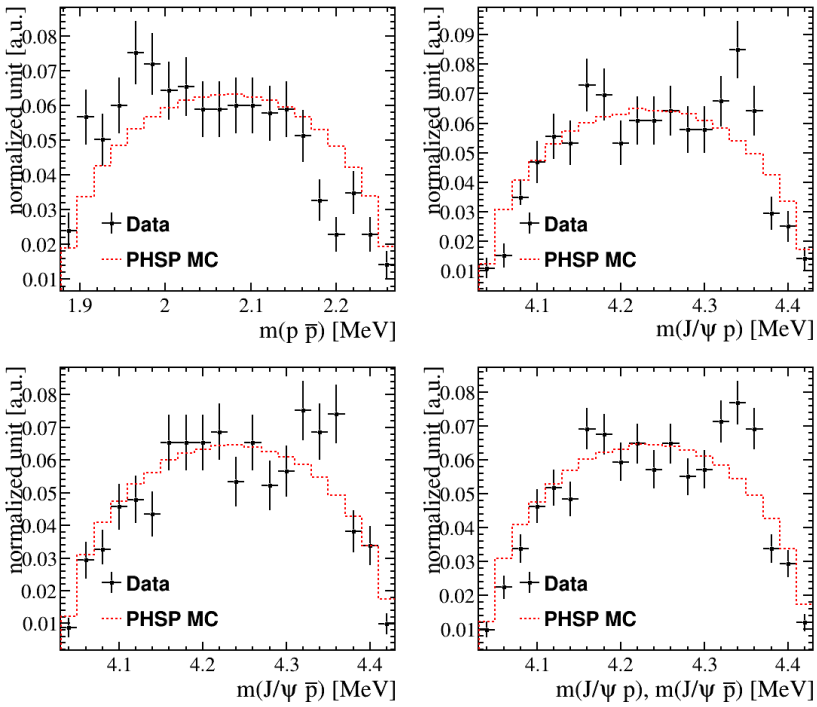


Figure 5.9 Invariant mass distribution for $p\bar{p}$, $J/\psi\, p$, $J/\psi\, \bar{p}$ and the sum of the two, with all dataset in a mass region of 3σ around the peak. The red curves represent the expected phsp distribution taken from MC.

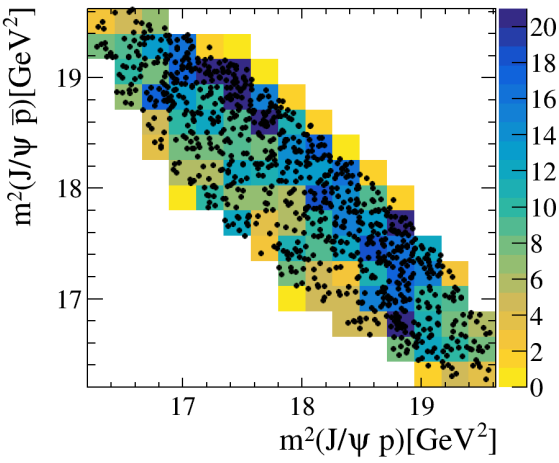


Figure 5.10 Dalitz-plot distribution in $m^2(J/\psi\, p)$ vs $m^2(J/\psi\, \overline{p})$ invariant masses.

5.5 Fit strategy

The fit is an unbinned maximum likelihood fit. Since the sample is composed by signal plus background, the likelihood must consider both contributions and the expression is the following:

$$-2\ln\mathcal{L}(\vec{\omega}) = -2\sum_i \ln\mathcal{P}\left(m_{p\bar{p}i}, \Omega_i|\vec{\omega}\right) \tag{5.44}$$

The pdf of the signal is defined as,

$$\mathcal{P}_{\rm sig}\left(m_{p\bar{p}}, \Omega|\vec{\omega}\right) \equiv \frac{1}{I(\vec{w})}\left|\mathcal{M}\left(m_{p\bar{p}}, \Omega|\vec{\omega}\right)\right|^2 \Phi\left(m_{pp}\right)\epsilon\left(m_{p\bar{p}}, \Omega\right), \tag{5.45}$$

where $I(\vec{\omega})$ is the normalization integral and is extracted from a flat distribution over the phsp variables:

$$I(\vec{\omega}) \equiv \int \mathcal{P}_{\rm sig}\left(m_{p\bar{p}}, \Omega|\vec{\omega}\right) dm_{p\bar{p}} d\Omega \tag{5.46}$$

and $\epsilon\left(m_{p\bar{p}}, \Omega\right)$ is the efficiency map that describes the dependence of the efficency over the phsp. The efficiency map is extracted from a MC sample as described below in Sec. 5.5.1.

Fixing the percentage of background to β, the total pdf can be written as:

$$\mathcal{P} = (1-\beta)\mathcal{P}_{\rm sig}\ \left(m_{p\bar{p}}, \Omega|\vec{\omega}\right) + \beta\mathcal{P}_{\rm bkg}\ \left(m_{p\bar{p}}, \Omega\right)$$

where the background is also normalized as in Eq. 5.46 and its parametrization is described in Sec. 5.5.2.

The pdf of the signal contains different contributions, that can be either non-resonant or resonant terms. The fit fraction associated to each of them is estimated as the integral of that specific component with respect to the total signal integral as,

$$FF_i = \frac{\int \mathcal{P}^i_{sig}(m_{p\bar{p}}, \Omega|\vec{\omega}) dm_{p\bar{p}} d\Omega}{I(\vec{\omega})} \tag{5.47}$$

where $\mathcal{P}^i_{sig}$ is the pdf of the i-th signal component.

5.5.1 Efficiency

Since the physics models for the $B^0_s \to J/\psi p\bar{p}$ decay process is unknown, for the acceptance correction, we employ an event-by-event correction to the data using an efficiency parametrization based on the decay kinematics. The technique is the same used in the branching ratio measurement, which is described in Sec. 4.4.5. The efficiency is parametrized as the product of Legendre polynomials:

$$\varepsilon(\vec{x}) = \sum_{i,j,k,l} c_{i,j,k,l}\ P(\cos\theta_\ell, i)P(\cos\theta_v, j)P(\phi', k)P(m'_{p\bar{p}}, l) \tag{5.48}$$

where $P(x, a)$ are the Legendre polynomials in x of order a and $-1 < x < 1$ and $\vec{x}$ is defined as $\{\cos\theta_\ell, \cos\theta_h, \phi', m'_{p\bar{p}}\}$, as in Sec. 5.1.1. The coefficients, $c_{h,i,j,k,l}$, are determined from PHSP MC using a moments method, employing the orthogonality of Legendre

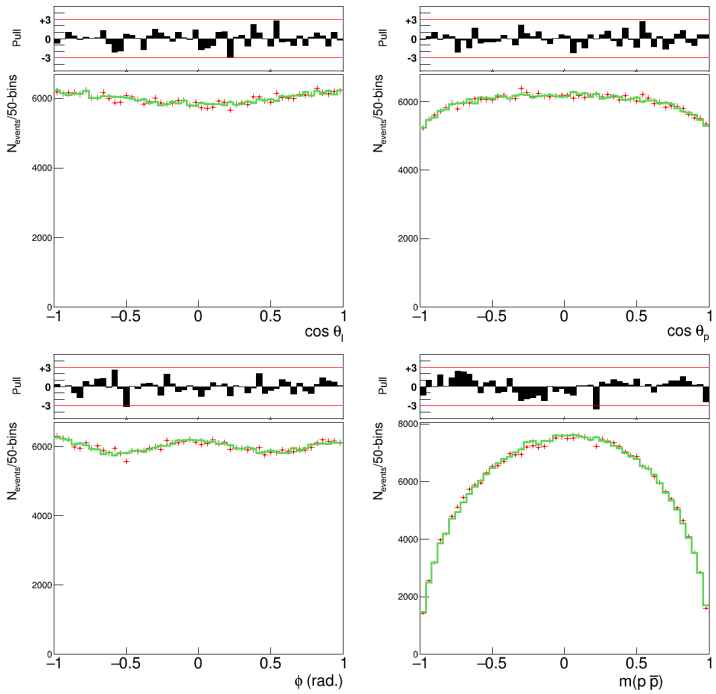


Figure 5.11 Run I $B_s^0 \to J/\psi p\bar{p}$: 1-d comparions in the kinematic variables between the reconstructed MC distributions (red) and the efficiency parameterization (green curve).

polynomials

$$c_{i,j,k,l} = \frac{C}{\sum w_n} \sum_{n=0}^{N_{\rm recon}} w_n \left(\frac{2i+1}{2}\right)\left(\frac{2j+1}{2}\right)\left(\frac{2k+1}{2}\right)\left(\frac{2l+1}{2}\right) \tag{5.49}$$

$$\times P(\cos\theta_\ell, i) P(\cos\theta_h, j) P(\phi', k) P(m', l), \tag{5.50}$$

which takes into account correlation among the phase space variables. The efficiency is extracted from MC after the trigger, preselection and selection requirements.

Figures 5.11 show the comparisons in 1D in the kinematic variables between the reconstructed MC and its parametrization for Run 1. The order of Legendre polynomial is determined as the lowest order possible that gives good agreement between the parameterization and the simulated PHSP events. These are: $n_{\cos\theta_\ell} = 5$, $n_{\cos\theta_h} = 7$, $n_{\phi'} = 7$, $n_{m'} = 3$. Up to statistical fluctuations, the parametrization follows the simulation data in all the distributions. In Fig. 5.12, the distributions for Run 1 and Run 2 are compared, they are statistically compatible and agree all over the phase space. Therefore, the parametrization extracted from Run 1 is used throughout and no systematic uncertainties have been evaluated due to the compatibility between the two runs.

5.5.2 Background

For a good modelling of the background, we parametrize it from the sideband after applying the same selection as for the signal candidates. Due to the quite limited statistics

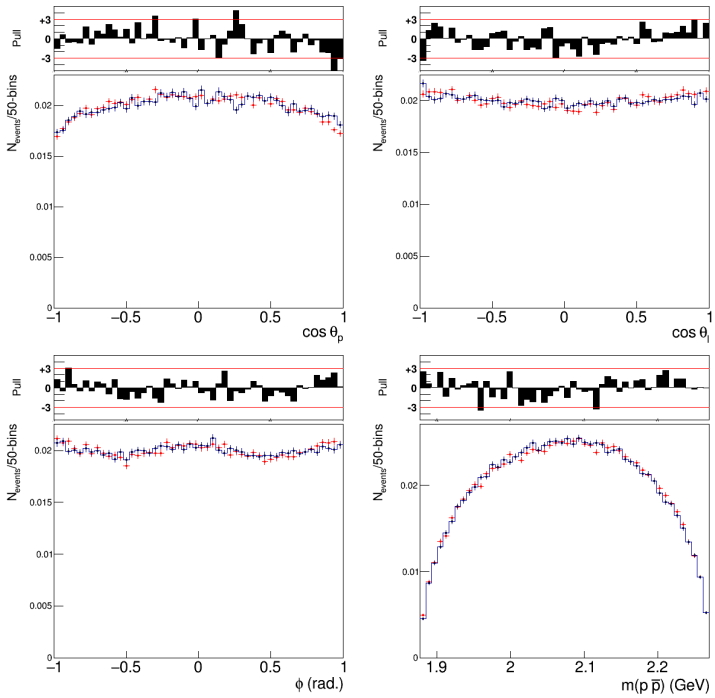


Figure 5.12 Comparisons in the kinematic variables between the reconstructed MC distributions in Run 1 (red) and Run 2 (blue).

in the sideband, we cannot perform a 4D parametrization. However, since the linear correlation is low ($< 5\%$), we decided to proceed doing a 1D parametrization for each variable in the phsp. The parametrization is achieved in a similar fashion with respect to the efficiency map done on signal and described in sec. 5.5.1. Each variable is parametrized doing an expansion on the Legendre polynomials basis and the total function can be written as:

$$f(m'_{p\bar{p}}, \phi', \cos\theta_l, \cos\theta_v) = \sum_{ijkl} c_i c_j c_k c_l P(m'_{p\bar{p}}, i) P(\phi', j) P(\cos\theta_l, k) P(\cos\theta_v, l), \quad (5.51)$$

where P is a legendre polynomial of a fixed order and c_i, c_j, c_k, c_l are coefficients extracted from the data sideband and computed independently from each other as:

$$c_i = \sum_{n=0}^{N_{\text{sideband}}} \left(\frac{2i+1}{2}\right) P\left(x, i\right), \quad (5.52)$$

where x is one of the four phase space variables. The sideband used in the nominal fit is taken from the right region $[5420, 5700]$ MeV and amounts to ~ 850 events, which is roughly 6 times the background below the peak. However, a cross-check is performed using the left plus right sideband in a region around the peak (m_B in $[5300, 5350]$ MeV and $[5420, 5460]$ MeV) since it is more likely to be a good representation of the background below the peak. However, due to the limited statistics, ~ 550 events, the first parametrization is used as nominal one. A systematic uncertainty due to it is assigned as described in Sec. 5.9.2. The corresponding parametrizations are shown in Fig. 5.13 and Fig. 5.14. The percentage of background events is fixed to 15% from the fit to data in the 3σ window around the B_s^0 peak.

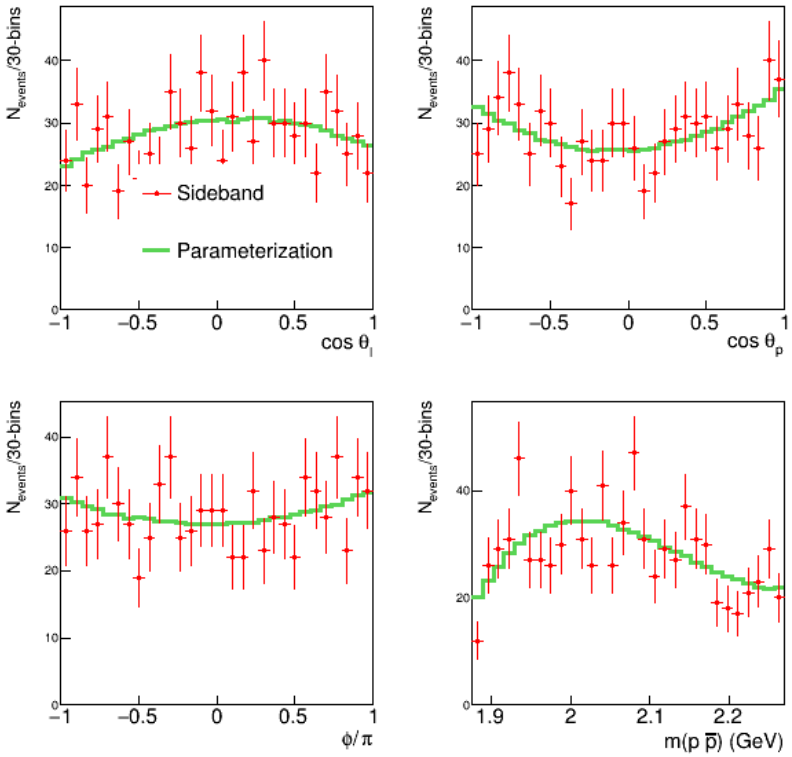


Figure 5.13 Parametrization for background using sideband data in the right region ($[5420, 5700]$ MeV). The Legendre polynomials are fixed to the orders: $n_{m_{pp}} = 5, n_{chi} = 3, n_{ctl} = 3, n_{ctv} = 3$.

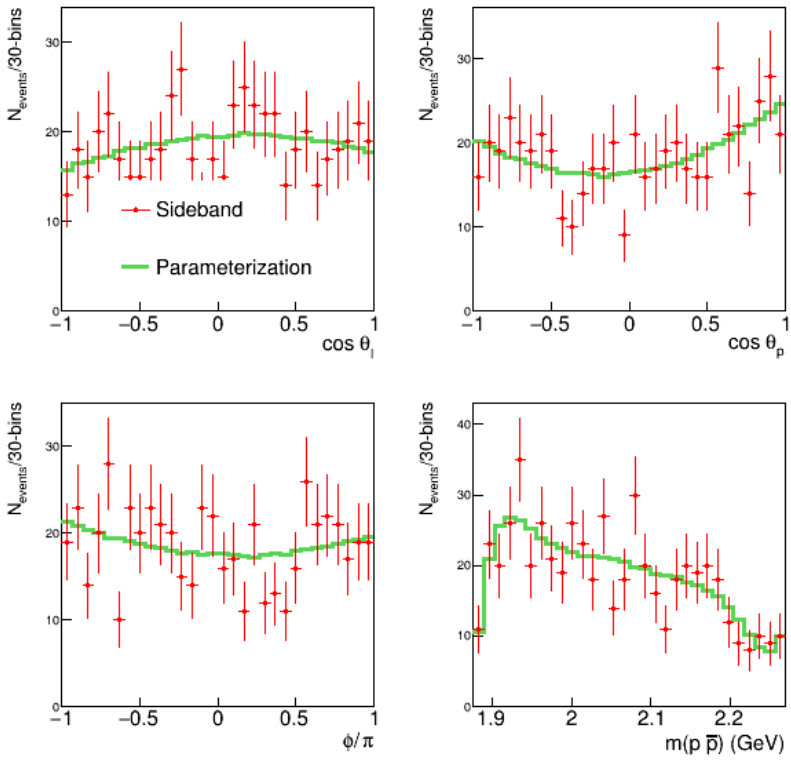


Figure 5.14 Parametrization for background using sideband data in left and right region close to the peak ($[5300, 5350]$ MeV and $[5420, 5460]$ MeV). The Legendre polynomials are fixed to the orders: $n_{m_{pp}} = 8, n_{chi} = 3, n_{ctl} = 3, n_{ctv} = 3$.

5.6 Null hypothesis fit based on NR contribution

To perform a good nominal fit many possibilities have been investigated. Looking at data (Fig. 5.15) it is evident that a simple non-resonant contribution in S-wave cannot describe well the spectrum in the invariant mass of $p\bar{p}$ system. Indeed, we expect an S-wave NR term to be flat in the angular variables.

So, a configuration able to describe the enhancement at low $m(p\bar{p})$ needs to be found. This enhancement could be treated as

1. a non-resonant (NR) contribution with high values of orbital momentum, as described in Sec. 5.6.1
2. a resonant term, to model the near threshold enhancement in $p\bar{p}$, as described in Sec. 5.6.2.

Since the statistics is limited, our best fit will be taken as the simplest model as possible. In addition, the model with a resonant term near threshold (model 2) does not improve the phase space description and results in comparable values of $-\log\mathcal{L}$. For this reason, the model 1 will be taken as the null hypothesis model and the model 2 as alternative one. The difference between the two will be considered in the systematic uncertainty.

5.6.1 NR parameterisation (model 1)

Indeed, the best null hypothesis model is found to be a non-resonant contribution in the decay chain of $B \to J/\psi X(\to p\bar{p})$ with $J^{PC} = 1^{--}$. Other NR models with different J^P hypotheses have been investigated, but none of them can successfully fit the data, as described in Appendix A.4.1.

The first fit is performed including all possible values of the orbital momentum.

The LS couplings allowed for the X decay are:

- for production ($B \to J/\psi X$): $(L,S) = (0,0),(1,1),(2,2)$.
- for decay ($X \to p\bar{p}$): $(l,s) = (0,1),(2,1)$, assuming P conservation

where uppercase (L,S) are the angular momentum and spin of the production and lowercase (l,s) are the ones of the decay. In this case, there are three complex couplings for the production and two for the decay. Since the fit is invariant under a normalization factor, we choose to put a product of two of them real and equal to 1, *i.e.* $A(L=0,S=0) \times A(l=0,s=1) = (1,0)$. In the following, we will call (A, ϕ) the modulus and the phase of each helicity coupling, where A is positively defined and ϕ can assume values between $-\pi$ and π.

The threshold factors suppress contributions from high partial waves, while the Blatt-Weisskopf coefficients are introduced in the model to remove the effect of the rising threshold factors at high masses.

Since some couplings assume values compatible with 0 as reported in Tab. 5.2, for instance $L = 1$ for the production and the phase ϕ of $l = 2$ for the decay, we decided to fix them in order to reduce the number of parameters and obtain a more stable fit. The results, reported in Table 5.3, correspond to the best fit among 100 fits, which are obtained varying the initial parameter values.

In Fig. 5.15, the phase space variables described in Sec. 5.1.1 and invariant masses of $J/\psi\, p(\bar{p})$ are reported for the final configuration.

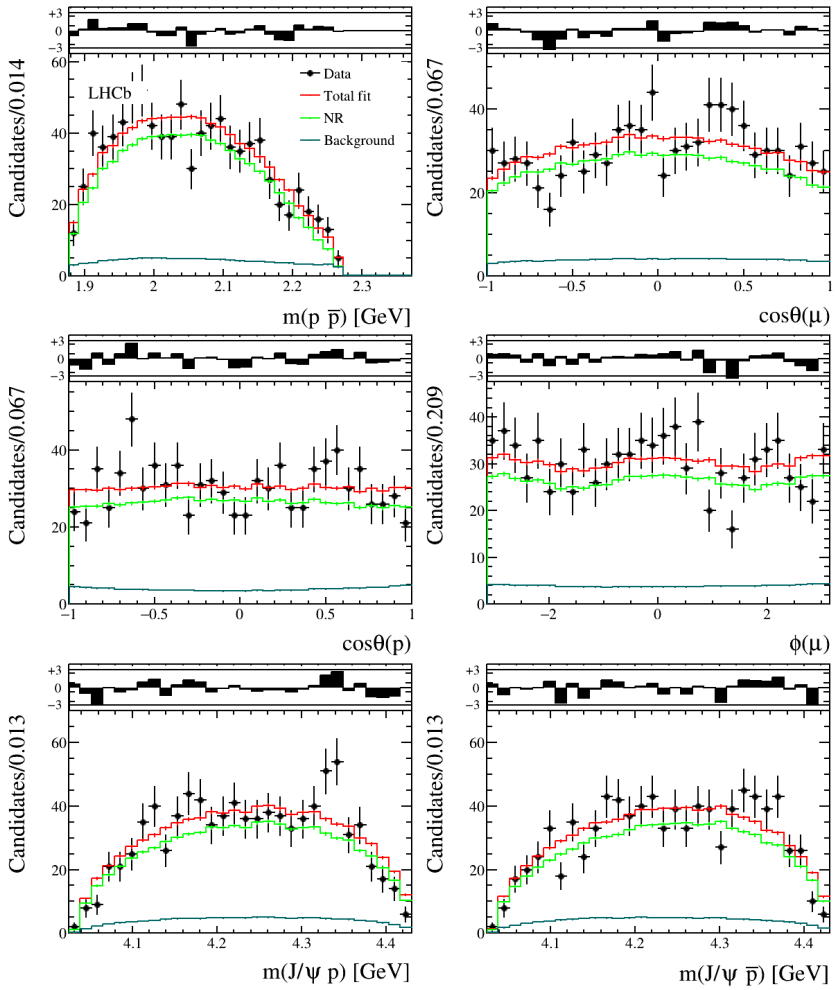


Figure 5.15 One dimensional distributions of kinematic variables in X decay chain for the fit with all possible values of LS couplings released.

Fit stability

In order to prove the stability of the fit, 500 toys are performed by generating and fitting with the same model. Each sample contains the same number of events as in data. As can be seen in Fig. 5.16, the pulls for each parameter are compatible with a standard Gaussian. So, no corrections on the parameters values will be implemented.

5.6.2 Fit with $p\bar{p}$ enhancement (model 2)

The excess of events at low $p\bar{p}$ invariant mass can also be motivated by a near "threshold enhancement", as has been seen in many B decays, as reviewed in Sec. 1.5.1. While the enhancement is explicitly predicted for the decay of $B^0 \to J/\psi p\bar{p}$, this is not the case for the B_s^0 decay since due to the OZI suppression hard gluons are responsible for the baryon-baryon production instead of soft ones (see Sec. 1.5.2). However, since

Parameter	Value	Error
A(NR) L=1	0.002	1.963
ϕ(NR) L=1	-0.4	2.6
A(NR) L=2	1.21	0.18
ϕ(NR) L=2	-1.47	0.13
A(NR) l=2	0.165	0.08
ϕ(NR) l=2	0.03	0.87
$-\log\mathcal{L}$	-71.4	

Table 5.2 Fit results for null hypothesis fit with all LS values.

Parameter	Value	Error
A(NR) L=2	1.26	0.20
ϕ(NR) L=2	-1.48	0.13
A(NR) l=2	0.17	0.08
$-\log\mathcal{L}$	-72.2	

Table 5.3 Fit results for null hypothesis fit for the default configuration.

the theoretical interpretation of this effect is not completely understood, we decided to perform a fit varying the null hypothesis model in order to account for the possible $p\bar{p}$ enhancement. This fit will be used to assign a systematic uncertainty as described in Sec. 5.9.1.

The null hypothesis model is modified in order to contain:

- an S-wave non-resonant contribution, $J^P = 1^-$, with lowest values of L, *i.e.* L,S = (0,0) for the production and l,s = (0,1) for the decay.
- a P-wave component to describe the $p\bar{p}$ enhancement, with $J^P = 0^-$ and a relativistic Breit-Wigner lineshape.

In the $p\bar{p}$ enhancement term, only one coupling contributes to the decay, since for $J^P = 0^-$, the only possibility for L is $L_B^X = 1$ for production and $L_X^p = 0$ for the decay.

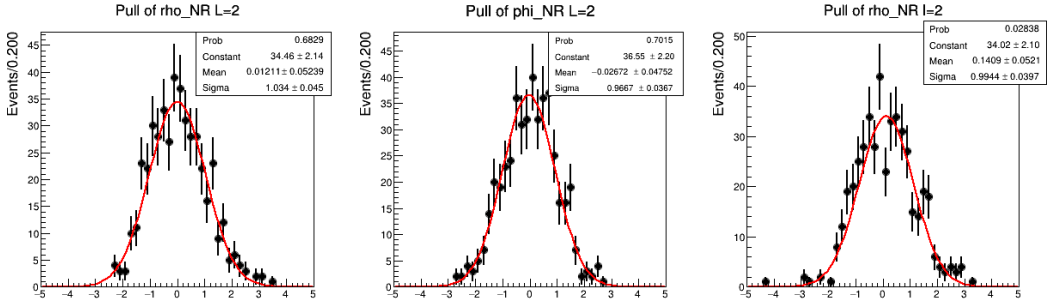
The results of the fit to data considering only the null hypothesis model is reported in Tab. 5.4 and in Fig. 5.17.

Parameter	Value	Error
mass(X)	1.93	0.03
width(X)	0.21	0.07
A(X)	0.42	0.15
$-\log\mathcal{L}$	-71.2	

Table 5.4 Fit results for alternative model of the null hypothesis.

5.7 Fit with P_c resonances

The fit with a non-resonant and a $p\bar{p}$ enhancement described in the previous section does not reproduce well the data in some angular distributions, like in $\cos(\theta_v)$, and in the invariant masses $m(J/\psi p)$ and $m(J/\psi\overline{p})$. Since contributions in the X chain cannot



(a) (b) (c)

Figure 5.16 Pulls for null hypothesis model for the three fitted parameters.

J^P	LS couplings	$c = (A, \phi)$
$1/2^-$	$(L,S) = (0,0), (l,s) = (0,1/2)$	(yes, yes)
$1/2^+$	$(L,S) = (0,0), (l,s) = (1,1/2)$	(yes, no)
$3/2^-$	$(L,S) = (1,1), (l,s) = (0,3/2)$	(yes, no)
$3/2^+$	$(L,S) = (1,1), (l,s) = (1,1/2)$	(yes, yes)

Table 5.5 Reduced number of LS couplings and fit parameters for different J^P hypotheses. Uppercase (L, S) stands for the production mechanism, while lowercase (l,s) for the decay. $c = (A, \phi)$ is coupling, where A is the magnitude and ϕ the phase.

compensate for the enhancement in $J/\psi p/J/\psi \bar{p}$, P_c^+ and P_c^- resonances are added to the model. Since we do not tag the flavour of the initial B meson, our model is symmetric in $J/\psi p$ and $J/\psi \bar{p}$ distributions and, due to the limited statistics, we would fit the same percentage of events for the two charged P_c states. This corresponds to reducing the couplings of P_c^+ and P_c^- to a single value, following the relation in Eq. 5.16, as explained in Sec. 5.1.4 and verified in Appendix A.4.3.

Different spin-parity hypotheses of the P_c^+ have been investigated; they correspond to spin $1/2$ and $3/2$, with parity $\eta_{P_c^+} = \pm 1$. The number of couplings in the model depends on the LS pairs we consider. Due to limited statistics the couplings are reduced to the lowest values of L allowed.

In Table 5.5, the lowest LS values for different J^P hypotheses have been summarized, together with the number of P_c couplings we are sensitive to. As said above, due to the limited statistics, there is only one coupling for each J^P that we are able to fit, which corresponds to the lowest values of L allowed both in production and in decay. In addition, the phase ϕ of the P_c coupling is extracted only for the spin-parity cases of P_c^+ equal to $J^P = 1/2^-$ and $J^P = 3/2^+$. When projecting the P_c states onto a CP eigenstate orthogonal to the X contribution, one would expect to be sensitive to a relative phase between the X and P_c states, as explained in detail in Sec. 5.1.4. In fact, the interference only cancels out after integration on the Dalitz, so we might be sensitive to an event-by-event interference (and thus fit a non-zero phase).

However, for the specific spin-parity assumptions $1/2^+$ and $3/2^-$, a more stringent relation due to parity applies. Indeed, in the case of a stable J/ψ, the interference always vanishes. As a consequence of the choice of LS couplings, the contribution in X is parity violating (PV) for the B_s^0 decay ($P_B = P_X P_\psi (-1)^L$ with $L = 0, 2$ and P being the intrinsic

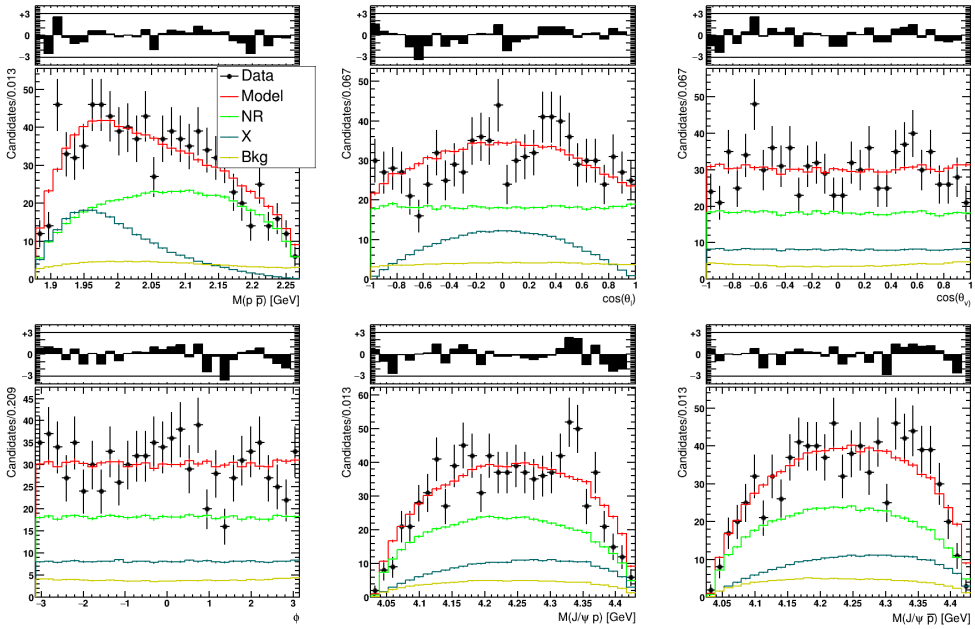


Figure 5.17 One dimensional distributions of kinematic variables of phase-space with $p\bar{p}$ enhancement and projections over $J/\psi p$ and $J/\psi \bar{p}$ invariant masses.

parities), while the P_c states are parity conserving (PC), ($P_B = P_{P_c} P_{\bar{p}} (-1)^L$ with $L = 0(1)$ for $1/2$ $(3/2)$). In Appendix A.3 we have shown that in these cases the interference cancels point by point in the phase space. This cancellation is only exact for a stable J/ψ, however the contribution to the interference is found to be small even in the case where the J/ψ decays and, therefore, we are unable to fit it with the current statistics.

In the following, we show the fit with the nominal model for the NR contribution, described in Sec. 5.6.1, while the alternative one, exploited for systematic studies, will be given in Appendix A.4.5.

5.7.1 Nominal fit with $J^P = 1/2^+$

Regarding the nominal model, in this section we report the results for the best fit to data, which corresponds to quantum numbers $J^P = 1/2^+$ for P_c^+ and $J^P = 1/2^-$ for P_c^-. The results for other spin-parity hypotheses are given in Table 5.7.

The couplings of the nominal model correspond to the orbital momentum $L = 0$ for production and $l = 1$ for decay. The highest value of L, for instance $(L, S) = (1, 1)$ and $(l, s) = (1, 3/2)$, are consistent with zero and, therefore, are fixed in the nominal model.

The result, reported in Fig. 5.18 and in Table 5.6, is the best one among 200 alternative fits varying the initial P_c parameters. The width is allowed to vary between $[5, 100]$ MeV, while the mass within $[4300, 4500]$ MeV. For most of the spin-parity hypotheses, there is only one local minimum. If more minima are present, they are well separated in $-\log \mathcal{L}$.

As an estimate of the goodness-of-fit, the χ^2 test is exploited on the 2d Dalitz plot (DP) variables, $m(J/\psi p)$ and $m(J/\psi \bar{p})$, assuming that the number of parameters is equal to: $\#bins - \#free\ parameters - 1$, where the number of bins, $\#bins = 50$, is chosen to contain at least 10 events per bin; $\#free\ parameters = 3$ is the number of free pa-

rameters of the fit divided by 2, which is the fraction of the DP variables over the phsp variables; and -1 is the degree associated to the normalization. In Fig. 5.20, the residual between the data and the fit model is shown, overlapped with the data events. For $J^P = 1/2^+$, a χ^2/ndf equal to 1.04 is estimated from the goodness-of-fit test.

Even if the fit quality adding the P_c states increases, the existence of these new resonances is not obvious. For this reason, some dedicated studies about significance will be performed in Sec. 5.8.4. In Table 5.6, the fit fractions are also reported; the values are extracted from toys as explained below.

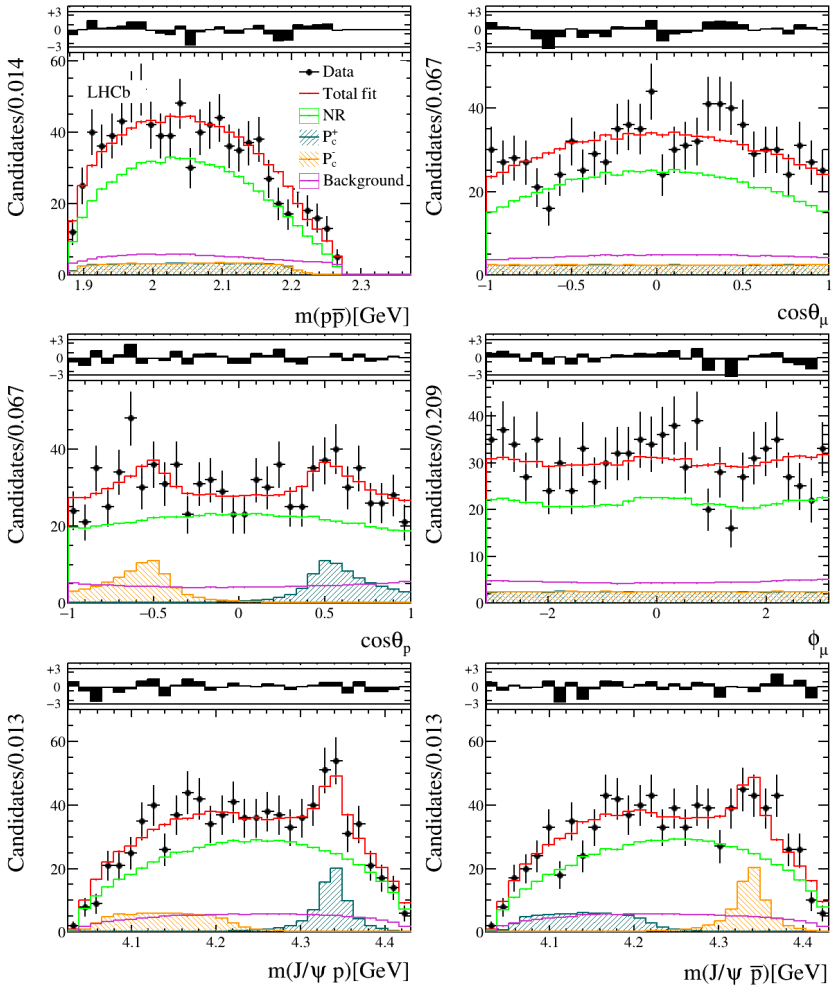


Figure 5.18 One dimensional distributions of kinematic variables of phase-space, plus invariant masses, as described in Sec. 5.1.2.

Fit stability

Also in this case, 1000 toys are performed to investigate the stability of the fit. The initial parameters are set to the best values extracted from data and a different seed is used

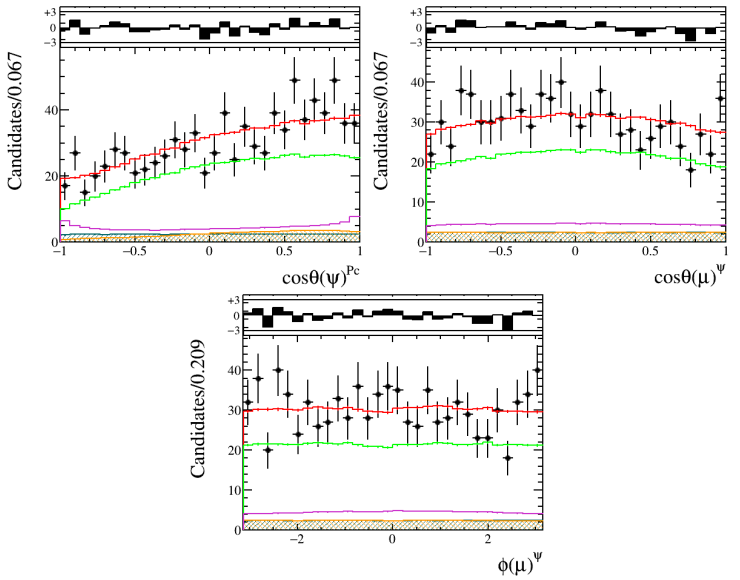


Figure 5.19 Angular distributions in the P_c decay chain, as described in Sec. 5.1.2.

for each toy generation. The same parametrization as in data is exploited both for the background and the efficiency.

Since the number of parameters increases with respect to the previous fit and we are dominated by low statistics, not all parameters have Gaussian pulls. Pulls are reported in Fig. 5.21.

It was also tested that the pulls tend to a standard Gaussian for high statistics samples ($\sim$ 10000 events), as shown in Fig. A.15 for spin $1/2^+$.

Biases of the pulls are taken into account in the systematic uncertainties as following. If the pulls have mean $\mu_p \pm \sigma_p$ and width $\omega_p \pm \sigma_\omega$, the corrected values can be computed as:

$$\mu_{fit}^{corr} = \mu_{fit} - \mu_p \cdot \sigma_{fit}, \tag{5.53}$$

$$\sigma_{fit}^{corr} = \sigma_{fit} \cdot \omega_p. \tag{5.54}$$

The fit fractions of all three components are also evaluated from toys, in order to have a better estimate of the confidence interval with a frequentistic approach. Toys are performed by generating with the nominal model and fitting with free parameters. The mean values and the corresponding uncertainties are then extracted from a fit to the fit fraction distributions with a bifurcated Gaussian in order to have access to the right and left uncertainties. Plots are reported in Fig. 5.21.

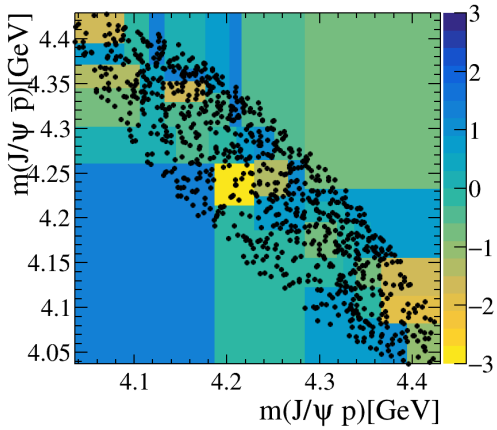


Figure 5.20 Residual between the data and the fit model in the Dalitz-plot variables, $m(J/\psi p)$ vs $m(J/\psi \bar{p})$ invariant masses with a $\chi^2/ndf = 1.04$.

Parameter	Value$^{\sigma+}_{\sigma-}$
A(NR) L=2	$1.25^{+0.27}_{-0.23}$
ϕ(NR) L=2	$1.62^{+0.21}_{-0.17}$
A(NR) l=2	$0.19^{+0.10}_{-0.09}$
A(Pc) l=1, s=1/2	$0.19^{+0.19}_{-0.07}$
A(Pc) l=1, s=3/2	$0.02^{+0.070}_{-0.06}$
Mass(Pc)	$4.337^{+0.006}_{-0.004}$
Width(Pc)	$0.029^{+0.026}_{-0.012}$
$-\log\mathcal{L}$	-85.6
Fit Fractions	
NR	$0.849^{+0.05}_{-0.08}$
P_c^+	$0.072^{+0.042}_{-0.020}$
P_c^-	$0.072^{+0.045}_{-0.021}$

Table 5.6 Fit results and fit fractions for the fit with a P_c^+ and a P_c^- state.

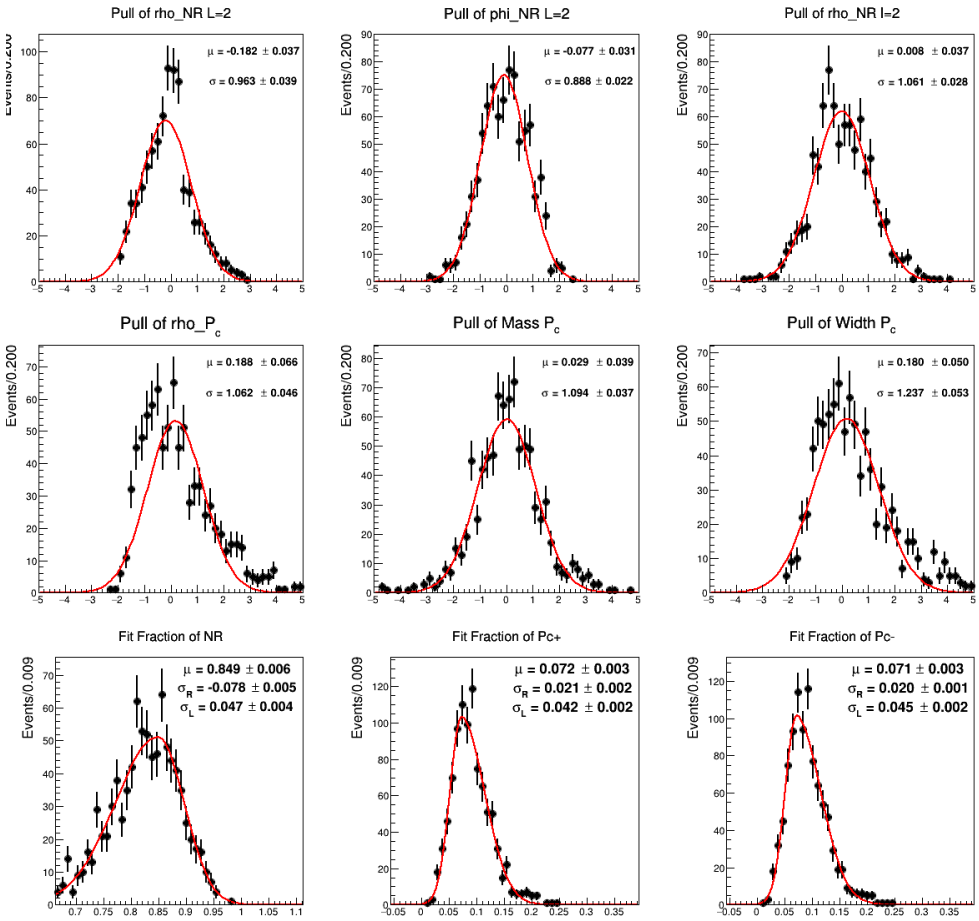


Figure 5.21 Pulls for alternative hypothesis model for all parameters and fit fraction distributions for the three fit components: NR, P_c^+, P_c^-.

5.8 Statistical analysis

The results presented in Sec. 5.8.4 and in Sec. 5.8.5 are obtained exploiting statistical techniques widely used in High Energy Particle Physics. Therefore, an introduction to the main techniques is provided below, before giving a detailed description of the application.

5.8.1 Hypothesis testing

The search for new phenomena can be performed in the context of hypothesis testing using likelihood-ratio based tests [136]. In general, one defines a null hypothesis H_0 and an alternative one, H_1, to be tested against the null. The null hypothesis is commonly a subset of a more general alternative hypothesis. For the purpose of discovering a new process, one wants to discard the null hypothesis, made up by only-background events, in favor of the alternative one, composed of signal plus background events.

The goal is to calculate what is the probability, under the assumption of the null hypothesis, of having a fluctuation of the null hypothesis that is equal or bigger than the excess observed in data. This quantity is called the p-value and is usually converted into Gaussian standard deviations, n_σ, or equivalently significances, through the equation:

$$n_\sigma = \Phi^{-1}(1-p) \tag{5.55}$$

where Φ^{-1} is the cumulative function, or equivalently the quantile, of the standard Gaussian. This corresponds to finding n_σ standard deviations above the mean of a Gaussian distributed variable with an upper-tail integral equal to the p-value. This corresponds to consider the one-sided interval definition [137]. It is also possible to define a two-sided fluctuation of a Gaussian variable. The method is referred to as two-sided interval.

The hypothesis under test, H_0, can be excluded above a certain threshold. The evidence for a new signal is claimed at the threshold of $n_\sigma = 3$, while the discovery is claimed at the threshold of $n_\sigma = 5$, corresponding to a p-value of $2.87 \cdot 10^{-7}$ in the one-sided definition. In the two-sided interval definition, the significance of 5 corresponds to a p-value of $5.7 \cdot 10^{-7}$. On the contrary, for the exclusion of a signal hypothesis, a threshold in p-value of 0.05 is commonly used, which corresponds to a confidence level (C.L.) of 95%.

In order to assign a level of significance, the likelihood ratio is usually exploited as test statistic. For the Neyman-Pearson lemma [138], the likelihood ratio, defined as,

$$\lambda = \frac{\mathcal{L}(H_1|x)}{\mathcal{L}(H_0|x)} \tag{5.56}$$

is the most powerful statistical test in discriminating between simple hypotheses, H_0 and H_1. The likelihood $\mathcal{L}(H|x)$ is a function of the data x and depends on the set of parameters the hypothesis belongs to. The parameters can be divided into "parameters of interest", like the rate of the signal process, and "nuisance parameters", which are not only given $a\ priori$, but can be determined from the fit. The likelihood is, therefore, a measure of compatibility between the data and the given hypothesis. The parameters are determined by the fit to data by maximizing the likelihood.

Instead of λ, it is convenient to choose as test statistic the q-ratio,

$$q = -2\log\lambda, \tag{5.57}$$

which converts the maximization of the likelihood ratio into a minimization and the logarithm transforms the product into sum of terms. Higher values of q correspond to higher disagreement between the data and the hypothesis under study.

Once the test-statistic is defined, we compute the p-value

$$p = \int_{q_{obs}}^{\infty} f(q|H)dq \tag{5.58}$$

where q_{obs} is the value of the test statistics observed from the data and $f(q|H)$ is the pdf of q under the assumption of the H hypothesis. The p-value is then converted into a significance, following Eq. 5.55.

For the Wilk's theorem [126], in case of high statistic samples, the distribution of the likelihood ratio asymptotically follows a χ^2 distribution with degrees of freedom equal to the difference between the free parameters in the null and alternative hypothesis. The p-value is then computed as

$$p = \int_{-2\Delta \log \mathcal{L}}^{\infty} \chi^2_k(q)dq. \tag{5.59}$$

where k being the number of degrees of freedom.

5.8.2 Look elsewhere effect

When performing a search for a new signal, the value of its mass is usually not known $a\ priori$. Therefore, the significance must take into account the probability that a fluctuation can happen everywhere in the mass range. This problem is commonly referred to as "look elsewhere effect". From a statistical point of view, it occurs when an hypothesis test depends on nuisance parameters, like the mass and width of a resonance, that are not defined under the null hypothesis, but only under the alternative one.

In such cases, the regularity of the likelihood is not satisfied and the Wilk's theorem does not hold anymore. Therefore, an effective number of degrees of freedom (ndf$_{eff}$) needs to be evaluated.

In order to consider this effect in the significance estimate, a sliding window mass analysis can be performed. The mass interval is divided in slices, for each of them the likelihood ratio is considered as test statistics. The fit is then repeated many times as the number of slices and a global test statistic is defined as the maximum of the likelihood ratio test over the entire mass range:

$$q_{max} = \max_{m} q(\hat{m}) \tag{5.60}$$

Details about the implementation of the method will be given in Sec. 5.8.4

5.8.3 Exclusion with CL_s method

In the case of an exclusion, the signal hypothesis has to be rejected against the null-hypothesis, and as said above, the exclusion principle is set once a C.L. of 95% is reached.

If two hypotheses exhibit overlapping distributions of the test statistics, it is possible to use the CL_s quantity [139], defined as

$$CL_s = \frac{p_s(H_1)}{1 - p_0(H_0)} \tag{5.61}$$

to exclude the H_1 hypothesis against H_0. The H_1 hypothesis is rejected with a confidence level of $(1 - \alpha)$, if a $CL_s \leq \alpha$ is found.

5.8.4 Significance of the P_c states

In order to assign a significance to the P_c resonances, an hypothesis test is performed, to evaluate whether the null hypothesis can be discarded in favor of an alternative hypothesis, as described in Sec. 5.8.1. The null hypothesis, H_0, is defined as composed by the NR contribution only, and it is described in Sec. 5.6.1, while the alternative one, H_1, contains the P_c states and it is provided in Sec. 5.7.1.

The distribution of the q-ratio, defined in Eq. 5.57, as the likelihood difference between the null and the alternative hypothesis, is exploited as test statistic. Therefore, the Wilk's theorem could be used in the asymptotic limit of high statistic sample. However, as mentioned in Sec. 5.8.2, we need to consider the "look elsewhere effect", since the present problem depends on nuisance parameters, like the mass and width of the resonance, that are not present under the null hypothesis.

For this reason, we follow a frequentistic approach to determine the effective number of degrees of freedom, ndf_{eff}.

Pseudo-experiments of the null hypothesis are then generated and fitted with both the null and the alternative hypotheses in order to draw the distribution of the q-ratio. For each pseudo-experiment, the sliding window mass analysis, anticipated in Sec. 5.8.2, is performed, where, instead of a single alternative fit, many fits to the same pseudo-data are made scanning the initial mass value in bins of size 50 MeV. Then, the test statistic is defined as the maximum q in the entire mass range, q_{max}, as in Eq. 5.60, to take into account the global significance. The scanning size is chosen to be at least 10 times higher than the mass resolution, in order not to fit unphysical fluctuations. A lower size of 35 MeV has also been used but no difference in $\Delta(-2\log\mathcal{L})$ has been found.

Two sets of simulations have been generated for the nominal models with different number of free parameters. The first one is generated with the hypothesis of spin-parity for the P_c^+ state equal to $1/2^+$, corresponding to the model where ϕ of the P_c is set to zero. The same simulations are also valid for the case of spin $3/2^-$. The second one is generated with spin $1/2^-$, which corresponds to the model with floating ϕ and applies to the case of spin $3/2^+$ as well. The $\Delta(-2\log\mathcal{L})$ distributions are reported in Fig. 5.22 in the right and left plot for the two sets, respectively.

The distribution of $\Delta(-2\log\mathcal{L})$ extracted from these pseudo-experiments does not follow a χ^2 distribution for all values of the test statistic, but it can be approximated to a χ^2 in the tail of it.

For this reason, we are following the same procedure as in the latest pentaquark analysis [32], where the idea is to fit the distribution of the test statistic only in the tail above a chosen threshold $\Delta(-2\log\mathcal{L}^0)$ and extract the significance using the formula:

$$n_\sigma = \sqrt{2}\text{erfc}^{-1}\left(f \cdot \frac{\int_{\Delta(-2\log\mathcal{L})_{data}}^{+\infty} \chi^2_{ndf_{eff}}(q)dq}{\int_{\Delta(-2\log\mathcal{L}^0)}^{+\infty} \chi^2_{ndf_{eff}}(q)dq}, \right) \tag{5.62}$$

where f is the fraction of events in the tail and is extracted counting the number of events above $\Delta(-2\log\mathcal{L}^0)$. The integral in the numerator is the p-value for a χ^2 distribution with ndf_{eff}, while the denominator is the integral of the same χ^2 distribution above the threshold, $\Delta(-2\log\mathcal{L}^0)$. It represents a normalization factor to account for the fact that only a fraction f of the pseudo-experiments is fitted with a χ^2 distribution with ndf_{eff} degrees of freedom.

The fit is performed above the value of $\Delta(-2\log\mathcal{L})^0 = 9$, which corresponds to a fraction of events f equal to $f = 0.27 \pm 0.05$ ($f = 0.36 \pm 0.03$) for the two distributions.

The number of ndf_{eff} is extracted from the fits in Fig. 5.22, and is measured to be

$$\begin{aligned} \mathrm{ndf}_{eff} &= 6.0 \pm 0.5, \\ \mathrm{ndf}_{eff} &= 7.2 \pm 0.3, \end{aligned} \tag{5.63}$$

for the spin-parity $J^P = (1/2^+, 3/2^-)$ and $(1/2^-, 3/2^+)$ of the P_c^+, respectively. The uncertainty on the significance is then assigned changing the central values by $+\sigma$ both on f and on ndf_{eff}. The significance of the best hypothesis (with $J^P = 1/2^+$) is measured to be 3.66 ± 0.10.

This procedure has also been validated with the frequentistic approach for the case of spin $3/2^+$, computing the p-value, as in Eq. 5.58, by counting the number of events above the value of the test-statistic observed in data ($\Delta(-2\log\mathcal{L})_{data} = 23.6$). The p-value is calculated to be 0.00178 ± 0.0008, which results in a significance of 3.13 ± 0.11, compatible with the result obtained with the previous method and reported in Table 5.7.

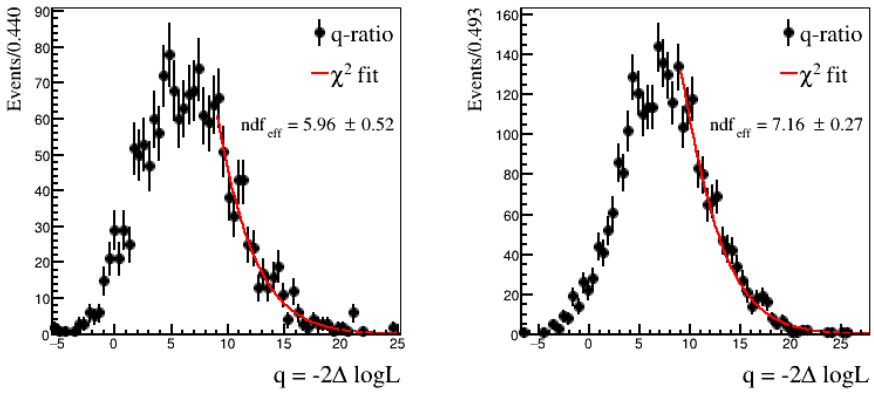


Figure 5.22 Test statistic distributions of the q-ratio under the assumption of the null hypothesis, fitted with a χ^2 distribution for $\Delta(-2\log\mathcal{L}) > 9$: for $J^P = (1/2^+, 3/2^-)$ on the left and for $J^P = (1/2^-, 3/2^+)$ on the right. The ndf_{eff} is compatible with 6 and 7 for the two cases.

5.8.5 Comparison of different spin-parity hypothesis

In order to access the quantum number of the P_c states in a reliable way, different spin hypotheses are tested. In Table 5.7, the results of the fit to data and their significance are summarized. For different J^P hypotheses, the significance is extracted as Eq. 5.62, assuming that the ndf_{eff} for $J^P = (1/2^+, 3/2^-)$ is approximately equal to 6, while for $(1/2^-, 3/2^+)$ is equal to 7.

To estimate if the alternative J^P hypothesis can be rejected, it is assumed as test statistic the likelihood ratio:

$$q = -2\log\left(\frac{\mathcal{L}(J^P_{alt})}{\mathcal{L}(J^P = 1/2^+)}\right) \tag{5.64}$$

where $\mathcal{L}(J^P_{alt})$ is the likelihood of the disfavored hypothesis, while $\mathcal{L}(J^P = 1/2^+)$ is the likelihood of the favored one. The distribution of the test statistics for each hypothesis

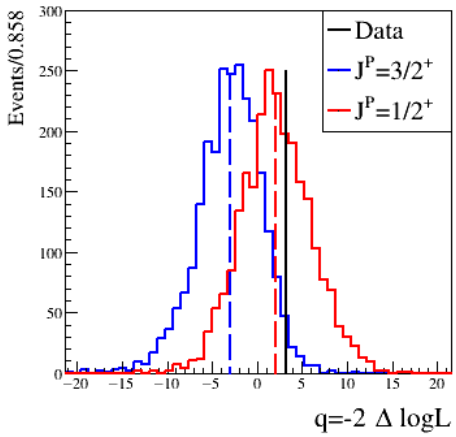


Figure 5.23 Distribution of the test statistics under the disfavored (blue) and the favored hypothesis (red) for $J^P_{alt} = 3/2^+$. The observed value is indicated by the vertical solid line and the expected medians by the dashed lines.

is obtained using pseudo-experiments. The toys are generated for the hypothesis with highest separation from the nominal one, for instance $J^P = 3/2^+$. The distributions of q under the assumption of the alternative J^P_{alt} hypothesis and the favored hypothesis $J^P = 1/2^+$ are drawn, in order to access which is the separation between them. The p-values, p_0, for each tested hypothesis is obtained, following Eq. 5.58, by integrating the test statistic distribution above (below) the value observed in data for the alternative hypothesis J^P_{alt} (for the favored hypothesis $J^P = 1/2^+$). Since the test-statistic distributions under the J^P_{alt} and $J^P = 1/2^+$ hypotheses overlap, as exclusion principle the CL_s method is exploited, as described in Sec. 5.8.3, where CL_s is defined as:

$$CL_s = \frac{p_0(J^P_{alt})}{1 - p_0(J^P = 1/2^+)}. \tag{5.65}$$

The hypothesis J^P_{alt} is assumed to be excluded if the CL_s value is lower than 5%.

In Fig. 5.23, the q-ratio distributions for the disfavored hypothesis with $J^P = 3/2^+$ is reported.

This hypothesis cannot be rejected at the 95% of CL because the CL_s is larger than 5%, as reported in Table 5.8, together with the associated value of $\Delta(-2\log\mathcal{L})$ and the p-values.

Summarizing, none of the J^P hypotheses can be excluded at the 95% of confidence level.

5.9 Systematic uncertainties

In this section, the systematic uncertainties on the mass, width and fit fractions of the P_c states are discussed. The main contributions are due to different modeling of the NR term and of the choice of the J^P assignment of the P_c resonances. On top of them, other systematic uncertainties can be assigned due the background and the efficiency

J^P	$\Delta(-2\log\mathcal{L})$	ndf	σ	M_0 [MeV]	Γ_0 [MeV]	A(P_c)	$\phi(P_c)$
$1/2^-$	26.2	7	3.40 ± 0.06	$4335\ ^{+3}_{-3}$	$23\ ^{+11}_{-8}$	$0.15\ ^{+0.07}_{-0.05}$	$2.77^{+1.3}_{-1.4}$
$1/2^+$	26.8	6	3.66 ± 0.10	$4337\ ^{+7}_{-4}$	$29\ ^{+26}_{-12}$	$0.19\ ^{+0.19}_{-0.08}$	$-0.6^{+2.44}_{-2.95}$
$3/2^-$	25.8	6	3.55 ± 0.10	$4337\ ^{+5}_{-3}$	$23\ ^{+16}_{-9}$	$0.14\ ^{+0.08}_{-0.05}$	$-1.26^{+1.91}_{-1.98}$
$3/2^+$	23.6	7	3.10 ± 0.06	$4336\ ^{+3}_{-2}$	$15\ ^{+9}_{-6}$	$0.10\ ^{+0.05}_{-0.03}$	$-3.13^{+0.6}_{-0.6}$

Table 5.7 Results for all spin-parity hypotheses. Values of $\Delta(-2\log\mathcal{L})$, ndf extracted from pseudo-experiments and corresponding significance, together with mass, width and couplings of the P_c states are reported.

J^P	$-2\Delta(\log\mathcal{L})_{data}$	$p_0(J^P_{alt})$	$1-p_0(J^P=1/2^+)$	CL_s
$3/2^+$	3.2	0.029	0.971	7.5%

Table 5.8 List of values of $-2\Delta(\log\mathcal{L})_{data}$ measured on data, p_0-values and CL_s extracted from pseudo-experiments reported in Fig. 5.23.

parametrization and the choice of the hadron radius size of the Blatt-Weisskopf coefficients of the P_c states. The systematic uncertainties are evaluated from the residual distributions between the alternative and the nominal model as the bias in the mean values. Indeed, the RMS of these distributions reflects the statistical uncertainty (comparable RMS are obtained for the toys of the nominal model, as reported in Fig. A.16). So in order not to account for a statistical effect many times, it is taken as a systematic estimate the shift with respect to the mean values.

5.9.1 Modeling of NR contribution

In order to evaluate possible biases due to the null hypothesis model, pseudo-experiments are performed generating samples with an alternative model and fitting them with the nominal one. The alternative model is chosen as a NR contribution with $J^P = 1^-$ plus a resonant term that describes the $p\overline{p}$ enhancement with total angular momentum equal to 0^-. The model is described in Sec. 5.6.2 and the full fit to data is shown in Sec. A.4.5. The samples are generated with the same statistics as in data.

The distribution of the residuals for each fitted parameter is drawn, as shown in Fig. 5.24a, and the mean of this distribution is taken as systematic uncertainty, as reported in Tab 5.9. While, for the fit fractions, the same procedure as in the nominal model is exploited. A fit with a bifurcated Gaussian is performed on the fit fraction distribution and the difference in the mean values between the nominal and the alternative model is assigned as systematic uncertainty, as shown in Fig. 5.26a.

5.9.2 Background studies

The B^0_s sidebands are used to model the background in the signal mass region as described in Sec. 5.5. This is based on the assumption that the background in the sidebands is the same as the one below the peak. To validate this assumption, the right sideband used in the nominal fit is replaced by the sum of left and right sidebands, see Fig. 5.14, which covers the mass region between $[5.30, 5.35]$ GeV and $[5.42, 5.46]$ GeV for a total of around 550 events. To better extract the systematic effects related to the difference between the two background parametrization, pseudo-experiments are performed generating the phase space distributions with the nominal model and fitting with the

alternative one. The corrected values for mass, width and fit fraction of the P_c are extracted from the mean of the residuals distribution, Fig. 5.24b. The fit fraction systematic is assigned as above. The results are reported in Fig. 5.26b.

Another source of systematic comes from the fixed percentage of background events. In the nominal fit the percentage of background is 15% of the total number of events. For estimating the uncertainty on it, the fraction is varied by 1σ, where σ corresponds to the uncertainty propagated from the signal and background yield uncertainties of the $J/\psi p\bar{p}$ invariant mass fit and it is estimated to be 0.5%. Toys are generated with the new fraction and fitted with the nominal one. Results of the residual distribution are reported in Fig. 5.24c and 5.26c.

At last, the two sources of systematic related to the background are summed in quadrature.

5.9.3 Efficiency

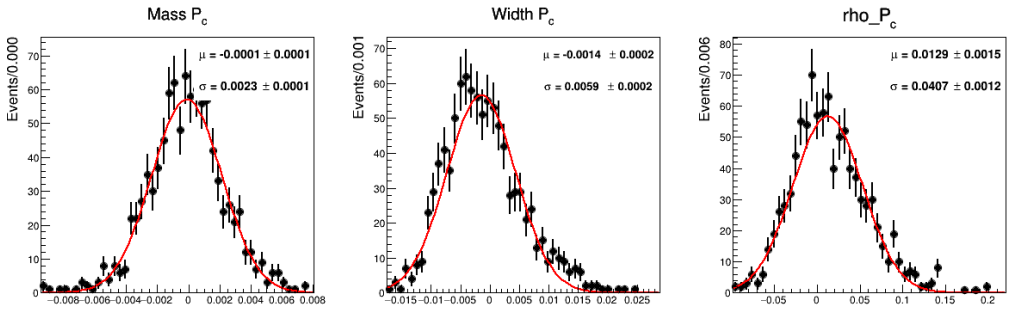
Another source of systematic is related to the modeling of the efficiency extracted from MC. We are interested to see which is the impact on the results in an extreme case of a parametrization without correlations. Therefore, we exploited a 1D Legendre parametrization, as done for the background presented in Sec. 5.5. The technique is the same as aforementioned: toys of the same statistics as in data are generated with the nominal efficiency parametrization and fitted with the alternative one. The systematic uncertainties on parameters are extracted from the mean of the residual distributions. In Fig. 5.28, comparisons between the 1D and 4D parametrizations are shown.

5.9.4 J^P assignment

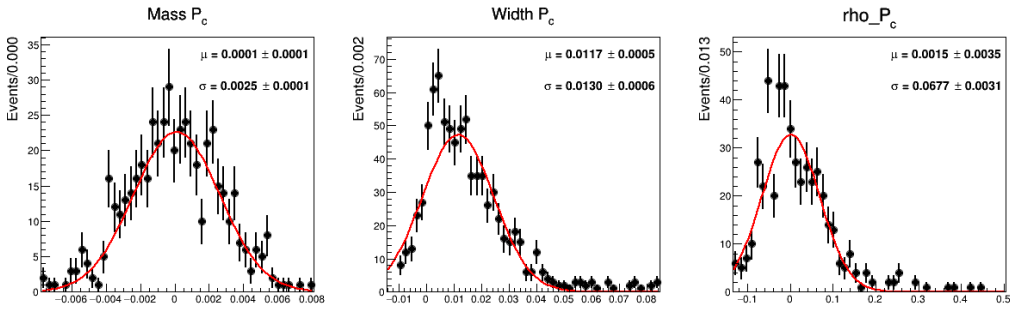
Since we are dominated by low statistics we are not able to assign the quantum numbers to the resonant state: all the J^P hypotheses are compatible within 2 standard deviations. For this reason, we assign a systematic uncertainty due to the difference between the nominal hypothesis $J^P = 1/2^+$ and the one with the largest discrepancies in the fitted parameters, which corresponds to $J^P = 3/2^+$. In order to do this, we performed pseudo-experiments of the nominal hypothesis with $J^P = 1/2^+$ and repeat the fit with the alternative hypothesis $J^P = 3/2^+$. The systematic is then assigned as the mean of the residuals between the best and the alternative model for each fitted parameter.

5.9.5 Hadron radius size

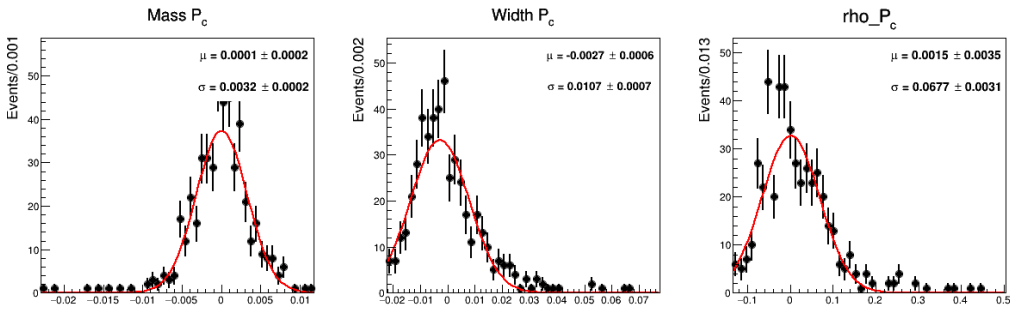
Another systematic is assigned related to the hadron radius d used in the Blatt-Weisskopf coefficients, described in Sec. 5.1.3, which is assumed to be $3\,\text{GeV}^{-1}$ in the nominal model. The alternative value for d is chosen to be $1.5\,\text{GeV}^{-1}$. Systematic uncertainties are assigned as the mean of the residual distributions. This contribution is found to be small because it depends on the dynamics of the decay and the L values considered. In particular, it becomes relevant for resonances closed to threshold and which are not in S-wave. For the best J^P hypothesis $1/2^+$, the systematic is negligible for production, because the resonances occur to be far from the production threshold and are in S-wave, for which the Blatt-Weisskopf coefficient becomes equal to 1. While for the decay, the value of the hadron radius can be relevant because the resonances are closer to the end point, and are in P-wave. However, this effect is found to give a smaller contribution than the other systematic effects.



(a)

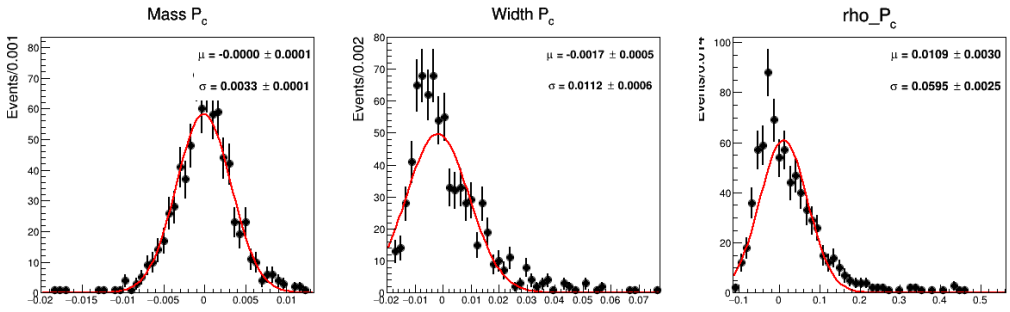


(b)

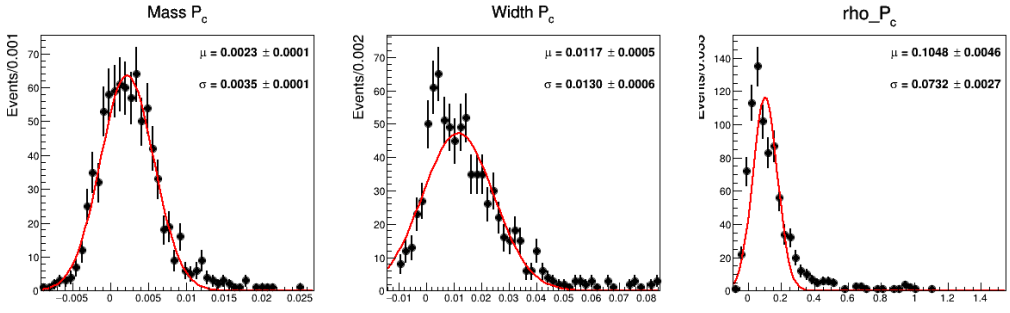


(c)

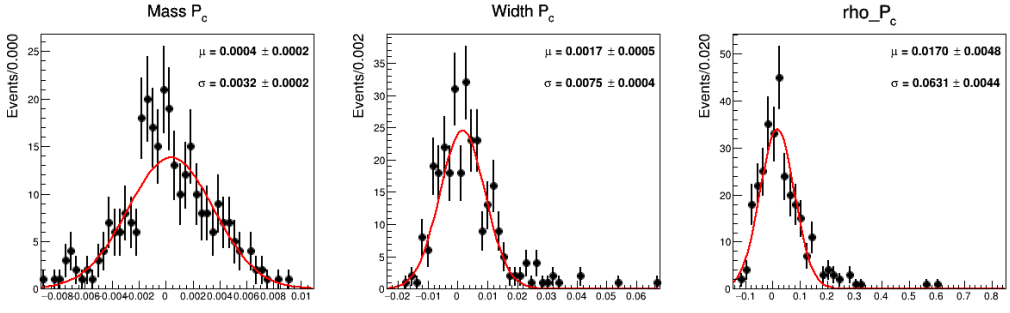
Figure 5.24 Residuals for mass, width and coupling of the P_c state for the modeling of NR term 5.24a, for background systematic 5.24b and for the fixed percentage of background events 5.24c.



(a)

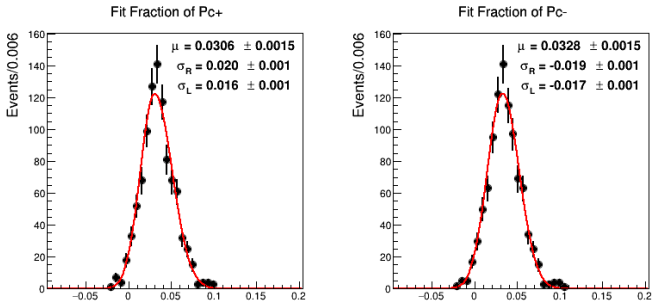


(b)

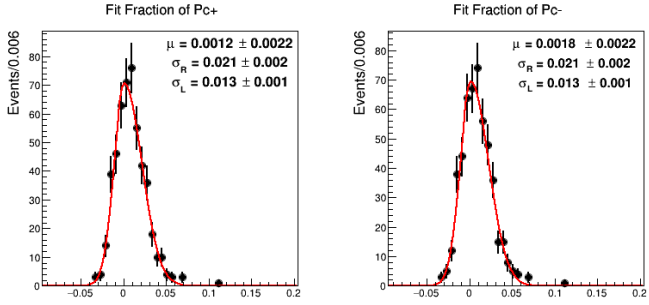


(c)

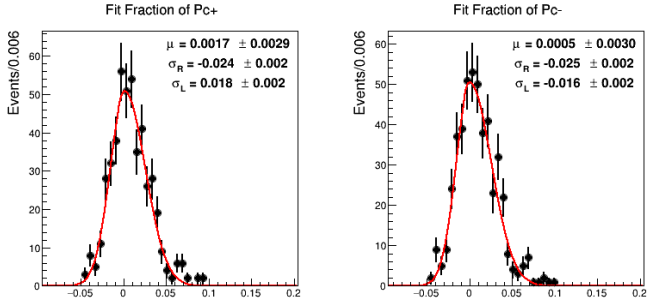
Figure 5.25 Residuals for mass, width and coupling of the P_c state for the efficiency 5.25a, for the choice of J^P assignment 5.25b and for the hadron radius size 5.25c.



(a)

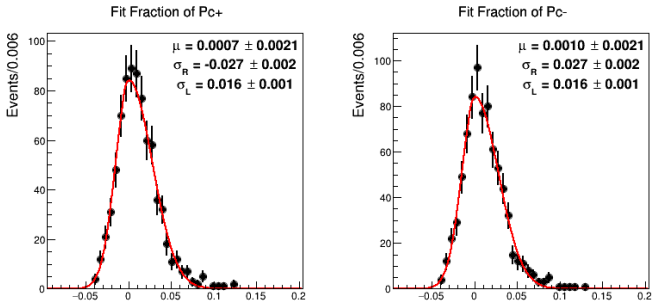


(b)

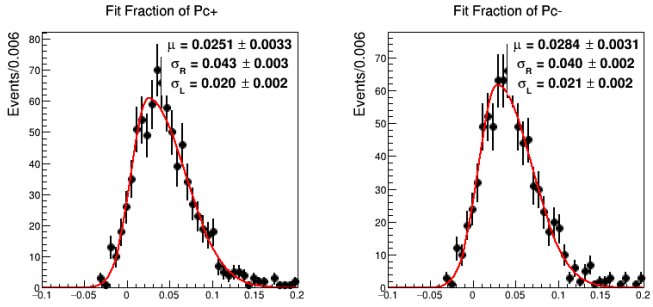


(c)

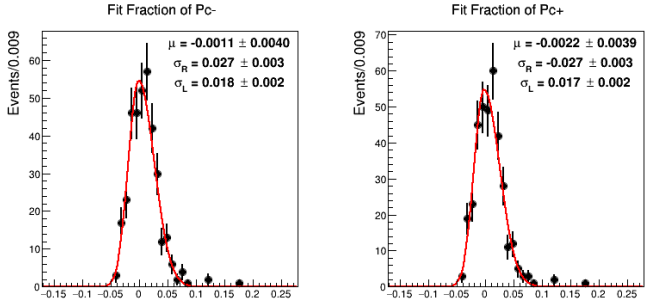
Figure 5.26 Residuals for fit fractions of P_c^+ and P_c^- for fit model systematic 5.26a, for background systematic 5.26b and for the fixed percentage of background 5.26c.



(a)



(b)



(c)

Figure 5.27 Residuals for fit fractions of P_c^+ and P_c^- for the efficiency 5.27a, for the J^P assignment 5.27b and for the hadron radius size 5.27c.

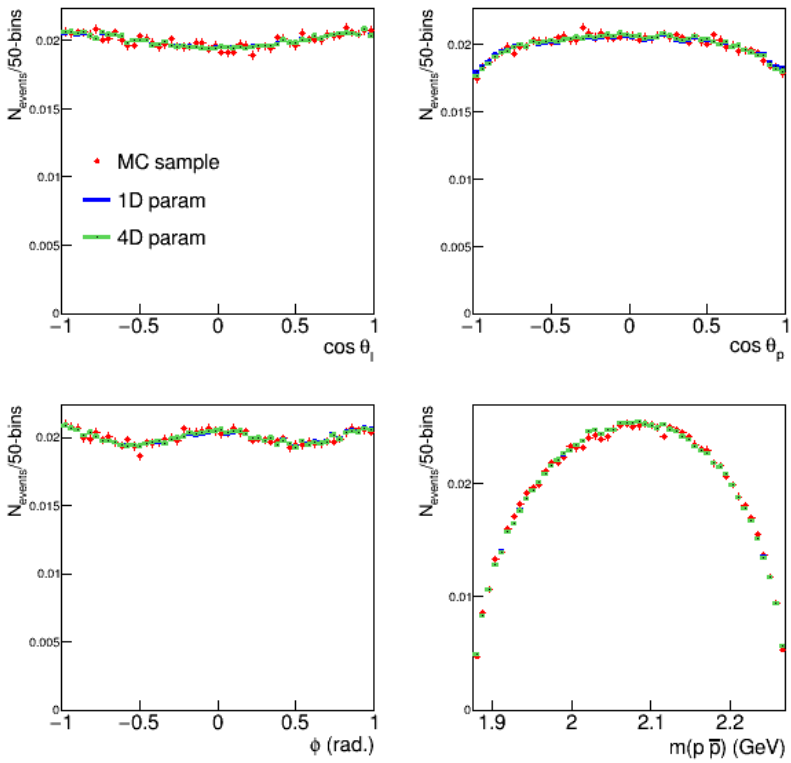


Figure 5.28 Comparisons between the 4D (green) and 1D efficiency parametrization (blue) in the phase space variables.

5.9.6 Fit bias

After the fit stability test performed in sec. 5.7 and reported in Fig. 5.21, where toys are generated and fitted with the nominal model, we decided to assign a systematic to account for the bias in the pull distributions. The systematic is assigned as the bias between the corrected and the nominal fit values and is equal to:

$$\sigma^{\pm}_{sys} = \mu^{corr}_{fit} - \mu_{fit} = \mu_{pull} \cdot \sigma^{\pm}_{fit}, \tag{5.66}$$

where μ_{fit} and $\sigma^{\pm}_{fit}$ are the mean and the statistical uncertainty of the nominal fit to data, respectively, μ^{corr}_{fit} is the corrected fit value and μ_{pull} is the central value of the pull distribution. The positive and negative systematic uncertainties reflects the asymmetric statistical errors.

5.9.7 Total systematics

All systematic uncertainties are summarized in Table 5.9, where the total systematic is obtained as the sum in quadrature of each component. The dominant systematic uncertainty is the assumption on the J^P value.

We have also verified that the assumption of a single coupling for the P_c states, together with the choice of the relative phase between P^+_c and P^-_c couplings in Eq. 5.16,

does not represent a systematic bias, as verified in Appendix A.4.3 and shown in the residual distributions in Fig. A.14. In addition, we have checked that we are not sensitive to higher values of the orbital momentum for the P_c states, as reported in Appendix A.4.6. Therefore, we do not consider as a possible systematic effect the inclusion of higher L in the P_c model. Finally, also the systematic on the invariant mass resolution is considered negligible with respect to the fitted width of the P_c states.

Source	M_0(GeV)	Γ_0(GeV)	A(P_c)	$f(P_c^+)$	$f(P_c^-)$
Fit model NR	0.0001	0.0014	0.013	0.031	0.031
Background	0.0001	0.003	0.016	0.002	0.002
Efficiency	0.0001	0.002	0.011	0.001	0.001
Fit Bias	$^{+0.0002}_{-0.0001}$	$^{+0.005}_{-0.002}$	$^{+0.04}_{-0.02}$	–	–
J^P assignment	0.002	0.012	0.10	0.027	0.027
Radius d	0.0004	0.002	0.017	0.002	0.002
Total	0.003	$^{+0.014}_{-0.013}$	$^{+0.11}_{-0.11}$	0.041	0.041

Table 5.9 Total systematic uncertainties associated to mass, width and couplings of the P_c.

5.9.8 Significance

As an estimate of the significance with systematic uncertainty, the fit on data is repeated with the alternative models used to evaluate the systematic uncertainties. The significance is computed with the same procedure as discussed in Sec. 5.8.4. The lowest significance among these fits is chosen as the nominal, which corresponds to $3.10 \pm 0.06\sigma$.

In Table 5.10, the value of $-2\Delta \log \mathcal{L}$ and the significance for all hypothesis are reported.

Syst	$-2\Delta(\log \mathcal{L})$	NDF	Significance σ
Fit model NR	30.8	6	4.07 ± 0.10
Background	33.2	6	4.33 ± 0.10
Efficiency	33.6	6	4.37 ± 0.10
Radius d	26.8	6	3.66 ± 0.10
$J^P = 1/2^+$	26.8	6	3.66 ± 0.10
$J^P = 1/2^-$	26.2	7	3.40 ± 0.06
$J^P = 3/2^-$	25.8	6	3.55 ± 0.10
$J^P = 3/2^+$	23.6	7	3.10 ± 0.06

Table 5.10 Significance for alternative models.

5.10 Results

The amplitude analysis of the $B_s^0 \to J/\psi p\bar{p}$ decay is performed with the data sample collected by the LHCb experiment during Run 1 and Run 2 and corresponding to 9 fb^{-1} of luminosity. An enhancement in the invariant mass $J/\psi p$ and $J/\psi \bar{p}$ is found with a statistical significance of 3.1σ after including systematic uncertainties. The evidence for a state can therefore be claimed. The excess can be interpreted as a resonance with mass and width equal to:

$$\begin{aligned} \mathrm{M}_{P_c} &= 4337^{+7}_{-4}(\mathrm{stat})^{+3}_{-3}(\mathrm{syst})\,\mathrm{MeV}, \\ \Gamma_{P_c} &= 29^{+26}_{-12}(\mathrm{stat})^{+14}_{-13}(\mathrm{syst})\,\mathrm{MeV}, \end{aligned} \tag{5.67}$$

which is not compatible with the results of the pentaquark states previously observed by LHCb in the $\Lambda_b \to J/\psi pK$ analysis [32], and in particular with the states $P_c(4312)$, $P_c(4380)$, $P_c(4440)$ and $P_c(4457)$. Therefore, it represents a new pentaquark candidate, denoted as $P_c(4340)$.

Since the untagged B_s^0 model is not sensitive to the P_c^+ and P_c^- contributions separately, but only to the sum of them, a fit fraction equivalent to the sum is computed:

$$f(P_c) = 14.4^{+8.0}_{-4.0}(\mathrm{stat}) \pm 8.2(\mathrm{syst})\%, \tag{5.68}$$

which corresponds to around 110 signal events.

The most favorite hypothesis corresponds to quantum numbers for the P_c^+ candidate equal to $J^P = 1/2^+$, although, due to low statistics, other hypotheses cannot be excluded and represent the main source of systematic uncertainties.

From the latest results on pentaquarks in $\Lambda_b \to J/\psi pK^-$ [32] decays, it seems that the molecular interpretation of these states is becoming the most favored one. In the molecular picture, the resonance should occur few MeV below the threshold of a molecule with same quark content and predominantly in S-wave. It is also probable to find a doublet of spin states, like in the neutron-proton binding scheme.

For these reasons, we have searched for a molecule of threshold above $4337\,\mathrm{MeV}$, in S-wave and with quark content equal to $uudc\bar{c}$. The main candidate is an hidden-charm molecule, for instance $\chi_{c0}(1P)p$, with threshold at $4352\,\mathrm{MeV}$. For angular momentum conservation, assuming the molecule to be in S-wave, the quantum number of the P_c state should only be $1/2^+$, which is our most favored hypothesis. Therefore, it will be very promising to see whether the excess will be present in future, with increased statistics, and whether the quantum number $1/2^+$ will be confirmed.

In addition, no evidence for a glueball state with mass equal to $2230\,\mathrm{MeV}$ and width around $20\,\mathrm{MeV}$, as predicted in Ref. [2], is observed. Indeed, adding a resonance in the X decay chain with fixed mass and width does not improve the negative log-likelihood of the nominal model.

CHAPTER 6

Front-end electronic tests for the UT upgrade

As explained in Sec. 2.4, the LHCb detector is undergoing a major upgrade, which installation will be finalized in 2021. The goal of this upgrade is to reach an expected integrated luminosity of $50\,\text{fb}^{-1}$ after Run 3 and Run 4. The main changes in the LHCb design regard the substitution of the front-end electronics to read out the detectors at the pp collision frequency, an improved detector granularity and radiation hardness.

In this section, I will detail my personal contribution to the upgrade, and in particular to the Upstream Tracker (UT), which will replace the current Tracker Turicensis (described in Sec. 2.2.1). In 2018, I have been involved in the tests of the front-end chip of UT, held in Milano, in collaboration with Syracuse and Kracov University.

A description of UT will be given in Sec. 6.1, highlighting the expected performance and the most relevant changes to the sensors and the read-out chip required in the upgrade. Afterwards, I will summarize the tests performed on the front-end electronics, namely the SALT Chip, version 2.5. The tests have revealed a malfunctioning of the chip itself, which could not allow to reach optimal performances in terms of signal over noise ratio. A new design was, therefore, needed to match the expected requirements.

6.1 The Upstream Tracker

Although the TT detector performed very well during Run 1 and Run 2, it has to be replaced by a new detector, called Upstream Tracker [98], for the following reasons: it is not sufficiently radiation hard to cope with the expected flux; the current strip geometry would lead to very high occupancy; the previous chip, called Beetle chip, was not designed for the 40 MHz read-out requirement. Furthermore, the distance between the detectors and the beam pipe have been reduced to obtain an improved acceptance in the very forward region.

6.1.1 Physics performance

The presence of the UT detector would be essential for the physics goals of LHCb, as well as the TT was. Indeed, thanks to the presence of this tracking system between VELO and the downstream tracker, it is possible not only to reconstruct long-lived particle decays, such as $K_s^0 \to \pi^+\pi^-$ and $\Lambda \to p\pi^-$ decays, which decay after VELO, but also to significantly improve the momentum resolution, as shown in Fig. 6.1 for the invariant mass of di-muon pairs. The increase of resolution adding the TT hits to the tracks is of $\sim 25\%$. In addition, due to the UT stations, it will be possible to reduce the "ghost tracks", which are fake tracks made up by wrong combinations of track segments, by a factor of three, according to simulations of the upgrade conditions. This helps in speeding up trigger decisions as well as reducing background events.

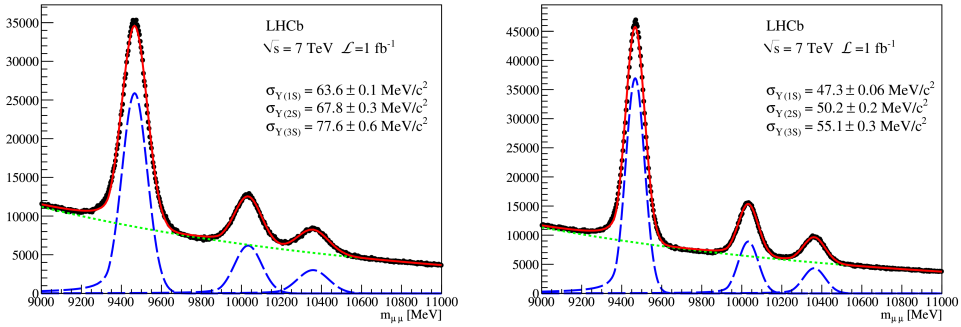


Figure 6.1 Invariant di-muon mass resolution with (right) and without TT hits (left). The (green) dotted line shows the background, the (blue) dashed line the signals and the solid curves the sum.

Finally, UT will be an essential part of the LHCb trigger, thanks to its improved acceptance and its higher reconstruction efficiency, which will be larger than 99.7% with the requirement of three hits out of four layers, and a single hit efficiency of 99%. In particular, the UT role in the software trigger is to reduce the timing of the forward tracking algorithm by measuring the momentum of the slow tracks that do not point to the primary vertex in VELO. Indeed, for the current tracking algorithm, the momentum is measured projecting all VELO tracks through the detector and matching them with the downstream tracker. This is very time consuming and it is simulated that, adding the UT information, the reconstruction time can be improved by a factor of three [140]. This is achievable thanks to a transverse momentum resolution, σ_{p_T}/p_T, of $\sim 15\%$, which is enough to assign the charge of the particle and exclude slow tracks ($p_T < 400\,\text{MeV}$) from reconstruction.

6.1.2 UT geometry

The UT geometry is very similar to the TT one. It consists of 4 layers of microstrip silicon sensors, placed at around 2.27-2.7 m from the pp interaction region, before the magnet. The layers are displaced with a $x-u-v-x$ configuration, as shown in Fig. 6.2, where the x planes measures the x coordinate with strips along the y axis, while the u and v planes are titled by a stereo angle of $\pm 5^\circ$ with respect to the y axis to measure the y coordinate. The acceptance is within 10 and ± 300 mrad in the xz plane and ± 250 mrad in the perpendicular plane.

The silicon sensors are built on a vertical structure, named stave, which is the physical support and it is used for the cooling of the front-end electronics and of the sensors. Each stave is 10 cm wide and 134 cm long and it consists of 14 or 16 sensors, which are mounted on both sides, in order to overlap in the y coordinate.

Despite TT, each UT sensor is composed of 250 μm thick silicon instead of 500 μm, which reduces the multiple scattering and increases the track reconstruction. The sensors are built with different technologies and different characteristics, depending on their position with respect to the beam pipe. The sensors positions are shown as coloured squares in Fig. 6.2: the yellow ones are the nominal, while the red and the green are the innermost and the outermost sensors, respectively. Different type of sensors are referred to as, A, B, C and D sensors, as shown in Fig. 6.3.

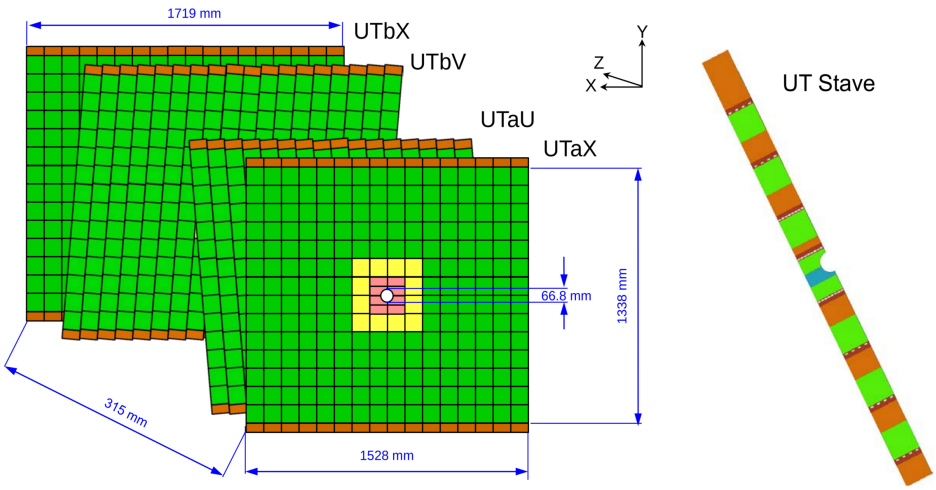


Figure 6.2 UT geometry: 4 layers of silicon strip sensors and vertical structure (*stave*).

The silicon sensors also differ for the technology being exploited: the B, C and D sensors use the n^+-in-p technology for the high dose of radiation of 40 MRad (considering a safety factor of four), while the A sensors are p^+-in-n for the modest radiation of 100-300 kRad, depending upon the location. In Table 6.1, the most important parameters of the sensor are reported.

The sensors are then read-out by an innovative front-end electronics, using an ASIC chip, called SALT (Silicon ASIC for LHCb Tracker). Each chip reads out 128 channels and therefore, 4 or 8 chips are needed for each sensor.

The front end electronics (SALT ASICs) will be mounted on hybrids, which are low mass flexible circuits that provide electrical connections of the ASICs to the data and power flexes. Each hybrid module is instrumented with 8 (4) SALT ASICs, depending on the type of sensor.

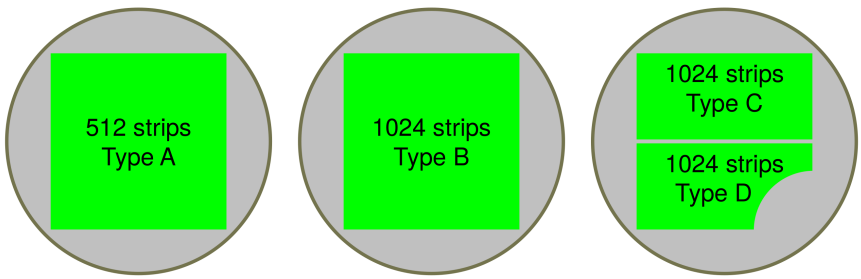


Figure 6.3 Sketch of the three sensor designs for the UT upgrade.

Property	Sensors B,(C,D)	Sensors A
Technology	n^+-in-p	p^+-in-n
Thickness	250 µm	320 µm
Physical dimensions	99.5 (50) $\times$ 97.5 mm^2	99.5 $\times$ 97.5 mm^2
Length of read-out strip	99.5 (50) mm	99.5 mm
Number of read-out strips	1024	512
Read-Out strip pitch	90 µm	195 µm
Sensor number (needed)	48 (16,16)	888

Table 6.1 Basic parameters of the silicon sensors.

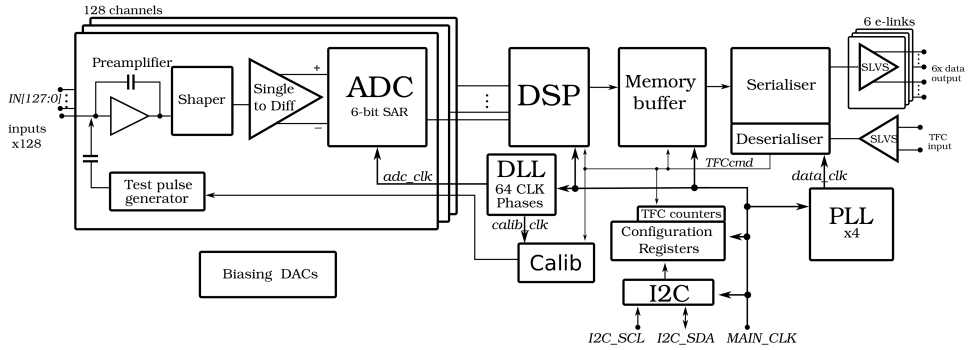


Figure 6.4 Block diagram of the 128-channel SALT version 2.5.

6.1.3 The SALT ASIC

The SALT [141] extracts and digitises analog signals from the sensor, performs digital processing and transmits serial output data. It is designed in CMOS 130 nm technology and uses a novel architecture comprising an analog front-end and 6-bit ADC in each channel. A block diagram of the SALT (version 2.5) is shown in Fig. 6.4. This is a preliminary version of the chip, which will not be installed in the detector. Future versions (*i.e.* v3.5, v3.8, v3.9), which meets the design requirements, are employed in the UT construction.

On the top left, it is shown the front-end scheme for each of the 128 channels. It comprises an analog block with a preamplifier, a shaper and a single-ended to differential converter. The shaper is designed to process signals every 25 ns, thanks to its peaking time, $T_{peak} = 25$ ns, and fast recovery. The differential analog signal is then sent to an Analog to Digital Converter (ADC). The ADC sampling is synchronized with the beam collisions through a dedicated ultra-low power (≤ 1 mW) DLL (Delay locked-loop) which takes an external clock and adjust its phase.

The digitised data from each ADC channel are processed by a Digital Signal Processing (DSP), which subtracts pedestal and mean common mode (MCM) noise. The pedestal is a constant offset for each read-out channel, induced by a DC current, while the MCM is a fluctuation in time of the signal baseline and is computed as an average noise over channels without signal. After the noise correction, the data are buffered in a dedicated RAM and trasmitted to the acquisition system (DAQ) via serial e-links at 320

MBit/s.

The ASIC is controlled via the LHCb common protocol consisting of two interfaces: the Timing and Fast Control (TFC) and the Experiment Control System (ECS) [142, 143]. The TFC interface delivers the 40 MHz clock and other crucial information and commands, synchronised with the experiment clock, while the ECS serves to configure and monitor the ASIC and it is realised through the Inter-Integrated Circuit (I2C) interface.

A complete description of the ASIC circuit is given in Ref. [141]. Here, we describe the analog front-end which is useful to perform the tests.

Analog front-end

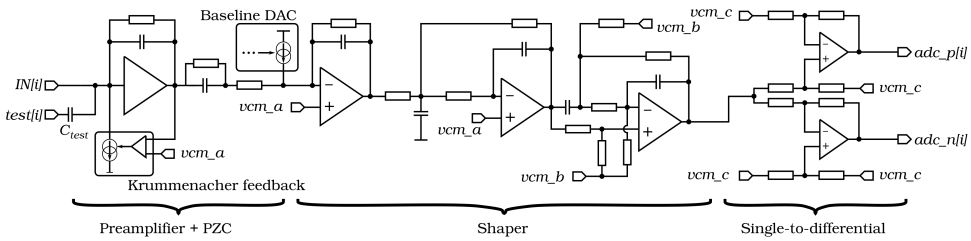


Figure 6.5 Analog front-end diagram block for each read-out channel.

The analog front-end is built with a peaking time of ∼25 ns and with fast recovery, in order to minimize pile-up and spill-over into the next bunch crossing. The chip should run at a low power consumption (1-2 mW/channel) and with different strip sensor capacitance (5-20 pF). It is also required to have a good enough signal to noise ratio (S/N >10), even in the worst operation conditions, for minimum ionizing particles (MIP). Since a MIP generates around 80 e-h couples per µm of silicon, we expect to collect a charge of:

$$Q = 80\text{e-h/ µm} \times 320\,\text{µm} \times q_e \sim 4fC. \tag{6.1}$$

A simplified block diagram of its structure is shown in Fig. 6.5. It consists of a preamplifier, which amplifies the signal, and a shaper in three stages, which provides the requested pulse shape.

A reference voltage at the entrance of each stage, vcm, should be set to half power supply for better performances.

In order to use the single-ended output as input in the ADC, it must be converted in a differential signal by a single-to-differential converter.

The analog front-end has four common parameters which are adjusted through internal DACs (Digital to Analog Converter). These are: preamplifier bias current I_{pre}; shaper bias current I_{sh}; and single-to-differential converter bias current I_{s2d}. The typical values for those currents are: 25 μA for I_{pre} and 3 μA for I_{sh} and I_{s2d}.

In order to control the entire SALT chain either without radiation or without the sensor, a calibration pulse can be generated, which mimics a real signal at the entrance of the detector. The current pulse has a shape of Dirac's delta and it is created by generating a voltage step on a test capacitor, C_{test}, at the preamplifier input. The input charge can then be obtained as:

$$Q_{in} = C_{test}V_{test} \tag{6.2}$$

where $C_{test} = 100$ fF, while V_{test} is between 0-170 mV, which results in an input charge of 0-17 fC. The characteristics of the pulse in terms of width, polarity and height are regulated by specific SALT registers. The formed pulse can be sent to some selected channels only.

6.2 Front-end chip tests

This section is dedicated to the quality tests of the SALT chip (version 2.5), performed in Milan between May and September 2018 and in which I was directly involved.

The experimental setup is described in Sec. 6.2.1, which consists of two boards, a "daughter board" on which the chip and the sensor are mounted and a "mother board", which regulates the transmission of the data between the chip and the acquisition system.

The most relevant tests are reported in Sec. 6.2.2 and are divided into two sets: test without any input load and tests with a full size sensor wire-bonded to the chip. A summary of the results is given in Sec. 6.3.

6.2.1 Setup

The MiniDAQ [144] is a standalone system for data acquisition used to test and debug the Front-End electronics. It is developed for being used in all LHCb sub-detectors. It is essentially composed by two parts, an acquisition board and a PC-server. The board, called PCIe40, is exploited to achieve three basic tasks: i) the data acquisition; ii) the slow control of the front-end (communication with the chip); iii) the distribution of the Timing and Fast Control (TFC). Its main component is an FPGA chip (Field Programmable Gate Array). The PCIe40 board makes also the interface between the FE electronics and the PC-server.

The MiniDAQ is connected to a "mother board", called VLDB (Versatile Link Demo Board) [145], through electrical link (e-link) ports, which offer line rates from 80 to 320 Mb/s. The VLDS board is a general kit developed for all LHC experiments for communications between detector front-ends and counting rooms. It is designed, therefore, to be radiation hard and magnetic-field tolerant. It has multiple functionalities: it can handle timing signal broadcast, trigger transmission, slow control, status monitoring and readout data. It is essentially composed by three ASICs, which regulate the aforementioned functionalities. They are called GBTx [146] (GigaBit Transceiver ASIC dedicated to serialization, deserialization and data and clock recovery), GBT-SCA [147] (Slow Control Adapter, ASIC dedicated to slow control and status monitoring) and VTRx/VTTx [148] (Versatile optical transceiver/dual transmitter and equipped with custom designed ASICs and qualified optical sub-assemblies).

The VLDB is then connected to a "daughter board", on which the chip and the sensor are mounted. The purpose of this board is to provide a connection between the VLDB and the hybrid with four or eight ASICs and eventually the sensor. The transfer of the data, the clock and the TFC is gathered by HDMI cables, while the SALT control is possible thanks to an I2C protocol. The board is equipped with a high-voltage power supply system.

These are its specifications:

- V_{in} = 3 V
- V_{out} = 1,2V

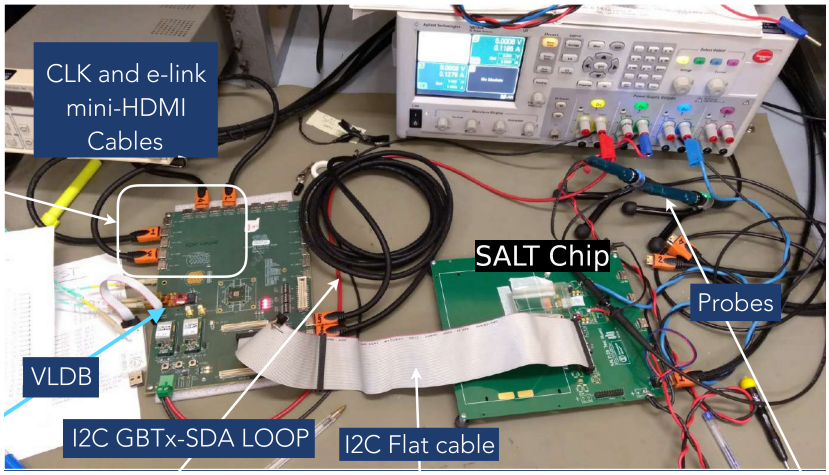


Figure 6.6 The test stand used to characterize the SALT ASIC in Milan. In the picture the two boards are shown: VLDB on the left, the mother board with the SALT on it on the right.

- I_{out} = 1,6 A
- Pd (Dissipated Power) = 3,6 W.

In addition, for the analog output an oscilloscope is exploited (LeCroy, SDA 808Zi-A), to test the signal shapes at the preamplifier and shaper output.

An example of the setup is shown in Fig. 6.6.

6.2.2 SALT ASIC tests

In order to establish a baseline performance that we can use as reference in assessing and diagnosing the problematic aspects of the chip, a preliminary set of tests without any input load is performed. Only in a second stage, the chip is bonded to a type-A sensor, to evaluate the performance with the sensor input load, approximately 10 pF of the strip capacitance.

Moreover, different chip configurations are tested by changing some crucial chip parameters, such as the number of enabled channels and the current in input at the preamplifier, I_{pre}.

As described in Sec. 6.1.3, a test capacitance is located at the input of the preamplifier for each ADC channel in order to generate a test pulse charge of about 1 MIP (Minimum Ionizing Particle), which emulates a real signal. The injection will be activated for the analog signal analysis. The nominal preamplifier current value of 24 μA is also changed to a lower value of 6 μA to verify signal saturation.

In the following, the two sets of tests are detailed with their specifications.

Test without any input load

For the measurement without the sensor, the assessment of the ASIC performance is based either on the analog signal at the preamplifier and shaper outputs or the digitized signal after the ADC processing.

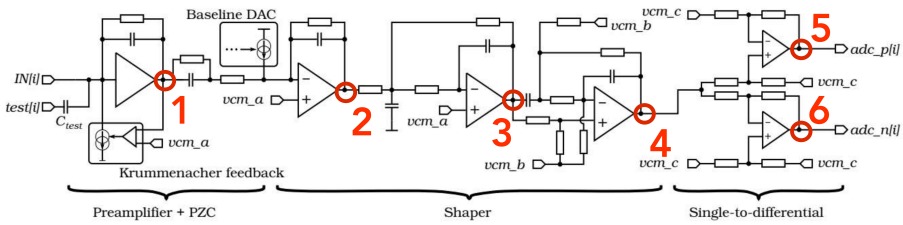


Figure 6.7 Simplified block diagrams of preamplifier and shaper.

First of all, we report the analog signal measurements. In order to perform them, probes are connected in specific points of the analog front-end chain of a single channel, as shown in Fig. 6.7. In particular, we are interested in: 1) preamplifier output and 2)-4) shaper outputs of the first to third shaper stages, when a charge is injected into the chain.

The first set of measurements is shown in Fig. 6.8, where the shape of a single signal peak in time (top) is overlapped to its fourier transform (FFT) (bottom), in the four points described above. In order, we have: 1) preamplifier output, 2) first stage output, second and third shaper outputs. All the 128 chip channels are activated simultaneously.

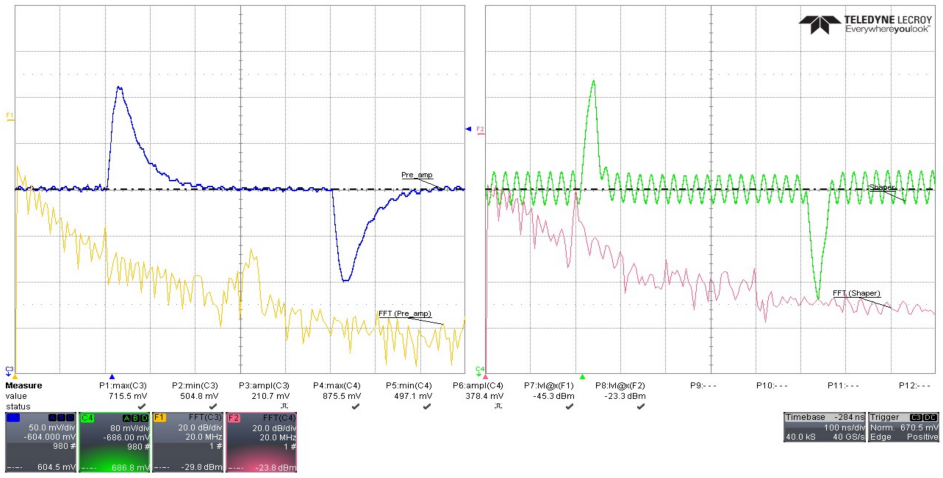
As can be observed, at each trigger signal the test pulse is splitted into a negative and a positive signal. The sign of the peak is also inverted by the first two stages of the shaper because they are inverting amplifiers. By examining the signal pulse shape, it matches the SALT requirements in terms of peaking time, consistent with 25 ns, while the signal height is decreased with respect to the input voltage value. Indeed, a value of around 170 mV is expected, in accordance to the value set as input, while an height of $\sim$ 100 mV is observed instead. This behaviour has been found to be related with a saturation problem of the preamplifier. Indeed, decreasing the input voltage to a lower value (30 mV) removes the disagreement between the input and the output.

The FFT of the signal is also overlapped to the signal pulse in order to assess anomalies in the frequency spectrum. As can be clearly seen at the shaper outputs, an oscillation of the baseline is present which is synchronous with a frequency of 40MHz, as seen by the peak in the FFT. This frequency corresponds to the LHC clock frequency, which is induced to the output lines.

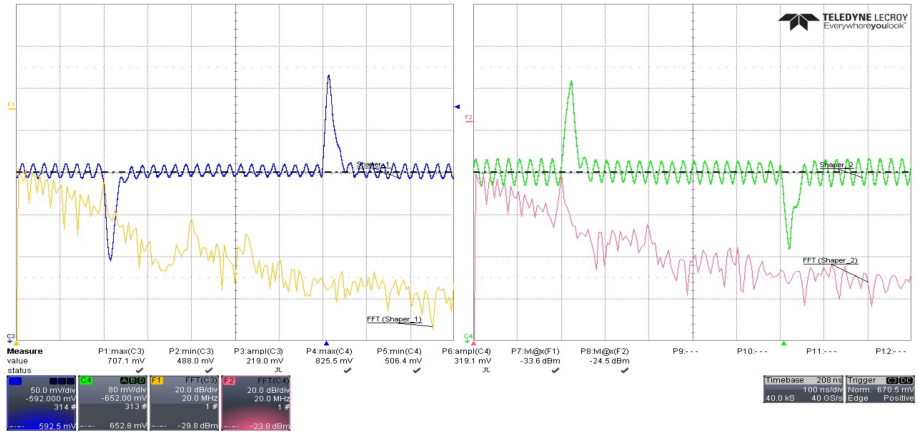
In order to verify its presence even in different configurations, the number of active channels is decreased to the minimum possible equal to 4 channels. In Fig. 6.9, the pulse shapes are reported for this configuration. As can be seen, the baseline oscillation is significantly reduced. We can therefore stated that with an increased number of activated channels, the signal over noise ratio (S/N) is deteriorated, going from a value of $S/N \sim 24$ for the latter configuration to a value of $S/N \sim 8$ for the former. Since the SALT is designed to run with all 128 channels enabled, the chip does not achieve the required performance of $S/N > 10$.

In addition to the analog measurements, tests on the digitized signal have been performed. In the following, we report the analysis of the baseline noise at the digital output, as a function of the number of channels that are activated. For performing this measurement, no test pulse is generated, in order to study only the noise in the ADC channels.

The output of the analog chain is converted into a digital signal from the ADC, as described in Sec. 6.1.3. In Fig. 6.10, the values of the digitized signal at the output of the SALT versus ADC channels are reported. They are measured in LSB, where 1 LSB is

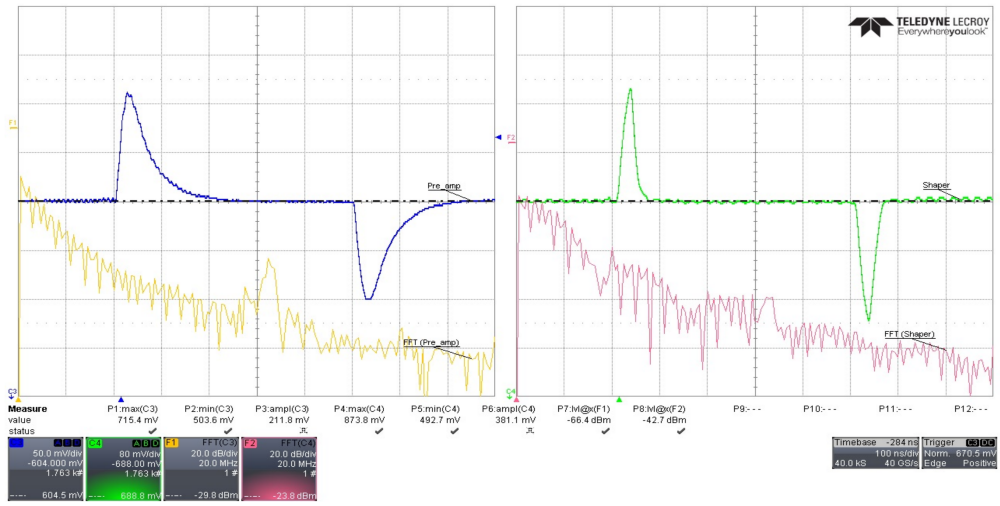


(a) Analog output at the preamplifier (left) and first shaper stage (right) with FFT distribution overlaid in yellow (left) and red (right).

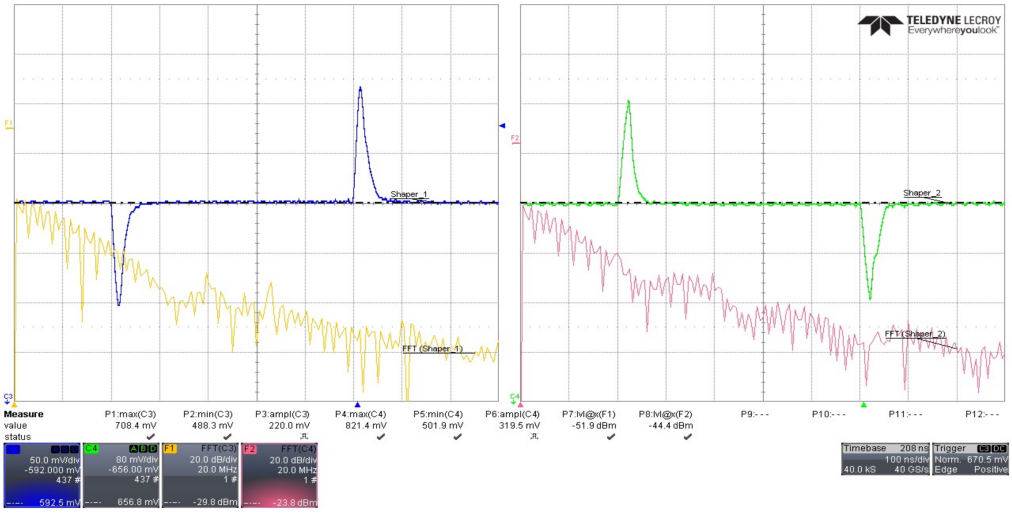


(b) Analog output at the second (left) and third shaper stage (right) with FFT distribution overlaid in yellow (left) and red (right).

Figure 6.8 Analog output with all 128 channels of the SALT chip enabled.



(a) Analog output at the preamplifier (left) and first shaper stage (right) with FFT distribution overlaid in yellow (left) and red (right).



(b) Analog output at the second (left) and third shaper stage (right) with FFT distribution overlaid in yellow (left) and red (right).

Figure 6.9 Analog output with only 4 channels enabled.

equal to 30 mV, and the zero in the LSB scale is equal to 600 mV, which is the reference voltage, vcm, of the analog chain. The maximum value of 30 LSB corresponds to the chip power supply, equal to 1.2 V.

From the baseline distribution in Fig. 6.10, we can evaluate the pedestal noise among channels, which is a voltage offset (or equivalently an ADC offset) with respect to 0, and the common mode noise, measured as the RMS of the pedestal distribution. Without any input load, the digital part of the chip performs as designed: neither a significant offset is observed nor a large RMS. A good behaviour is also found as a function of the number of enabled channels.

Test with full-size sensor

In this section, we will summarize the studies performed on the SALT v2.5 ASIC wire-bonded to a type-A sensor, which is described in Sec. 6.1.2 and has an area of about 10x10 cm^2. In Fig. 6.11, a picture of the daughter board with the sensor attached on it is shown. The sensor is covered by a black box to not allow light filtering and is biased with an High Voltage (HV) supply at 200V.

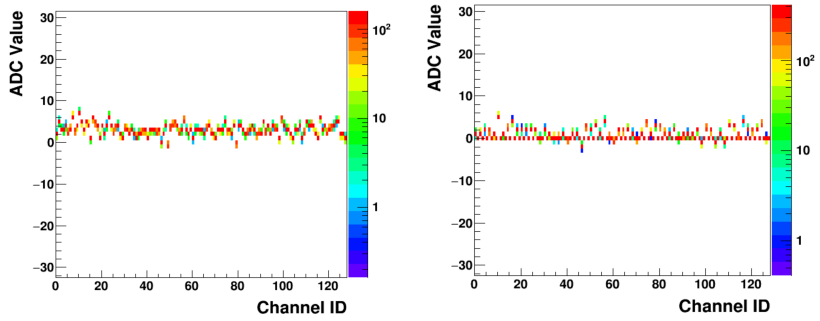
Due to problems when mounting the board, we were not able to control the analog signals, as presented for the case without the sensor. Therefore, the measurements presented here are related only to the digital output, particularly focusing on the ADC baseline analysis.

The measurements are performed varying the activated channels from 4 to 128 and with an increased value of the preamplifier bias current, in order to test how the bias configuration affects the chip performance. In the nominal configuration, the bias is set to the default value, 26 μA, and subsequently, is varied to a considerably lower value, equal to 6 μA. The ADC baseline is shown in Fig. 6.12 for $I_{pre} = 26$ μA and in Fig. 6.13 for $I_{pre} = 6$ μA.

As can be observed in Fig. 6.13a, when the chip is bonded to the sensor, a voltage drop in the ADC values is generated. This drop is predominant for the most central channels, where the values become closer to -30 LSB, or equivalently to 0 V.

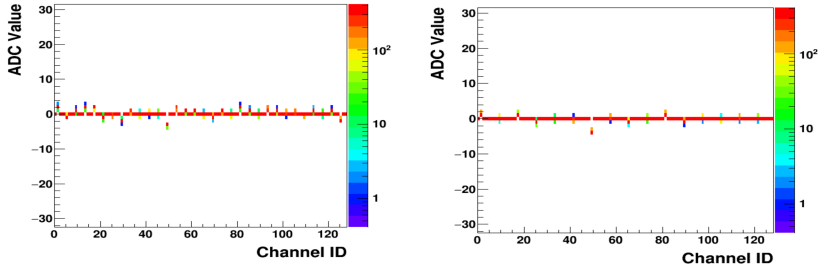
This behaviour is unexpected and is not acceptable in terms of the chip performance. The ADC values should be closer to 0 LSB for all channels. For such big shifts, indeed, it is not even possible to correct the baseline to obtain values compatible with zero. When discrepancy from 0 are observed, it is usually possible to tune the baseline of each channel by injecting a tunable DC current (using a TRIMDAC circuit) in one of the shaper stages. The measurement, called "TRIMDAC scan", consists in scanning the TrimDAC values (associated to the DC current) from 0 to 255 and measuring the corresponding pedestals. Since the change in signal baseline versus TrimDAC value is linear (as shown Fig. 6.14a), a correction may be applied by loading a value of TrimDAC that brings the baseline to 0 or as close as possible to 0 when large baseline shifts are observed. However, in the present configuration, these correction cannot be applied to all channels, as can be seen from Fig. 6.14b: for the most central channels the TrimDAC lines do not cross 0 and, therefore, a zero correction is applied. As a consequence, the corrected baseline does not improve with respect to the nominal configuration, as shown in Fig. 6.14c.

Studies performed on the sensors have shown that the capacitance seen by the chip wire-bonded to the input channel is about 11 pF. Indeed, the coupling between the chip and the heavy load of the sensor could be the source of the problem. This hypothesis was also verified using a mini-sensor, where the capacitance seen by the SALT is about a factor of 5 smaller than the one of the full size detector. Performances in between the two cases, without any load and with a full size sensor, are found.



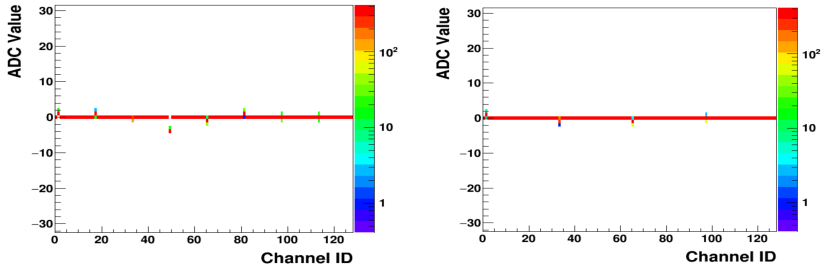
(a) Baseline with 128 channels enabled

(b) Baseline with 64 channels enabled



(c) Baseline with 32 channels enabled

(d) Baseline with 16 channels enabled



(e) Baseline with 8 channels enabled

(f) Baseline with 4 channels enabled

Figure 6.10 ADC baseline of the output channels of the ASIC without any input load, as a function of the number of activated channels.

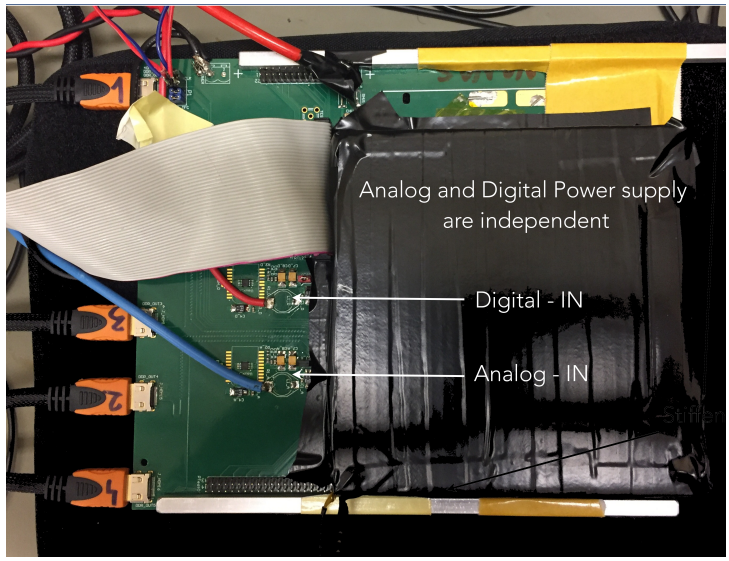


Figure 6.11 The board with the sensor attached and covered to not allow light filters. Digital and analog power supply are independent.

Besides, the voltage drop of the baseline decreases by switching off the channels. This effect could be related to the induction of the clock frequency to the ADC lines as observed for the configuration without any input load. The more channels are activated, the greater the induction on the ADC lines becomes. Therefore, it confirms the hypothesis that the large ADC shift is linked to the ADC activity.

We have also performed studies varying the preamplifier bias current to a lower value of 6 μA, as shown in Fig. 6.13. Also in this configuration, a voltage drop is observed, which causes not acceptable performances.

In addition, the chip is designed to run with all 128 channels enabled and with default value of the bias current. Hence, these alternative configurations cannot be taken as the default ones.

6.3 Results

The measurements reported in this thesis have shown critical performances of the SALT ASIC version 2.5. The main issues are related to:

- an oscillation synchronous with the 40 MHz clock, which is induced to the ADC read-out lines and deteriorates the S/N ratio. Even for the configuration without any input load, S/N is measured to be lower than 10 and, hence, is not acceptable compared to the expected performances. This value is expected to be even worse with a capacitance load attached to the chip.

- a voltage drop in the pedestal distribution for the most central ADC channels. In absence of any test pulse, the pedestal distribution should be close to 0, while a large shift is observed, which cannot be corrected even with the TRIMDAC scan. This effect is due to the high capacitance (of around 11 pF for each channel) seen by the SALT when connected to the sensor.

Alternative configurations have been tested by changing the bias currents and the ground connection of the sensors (which are not reported here) but none of them was found to satisfy the chip requirements.

These results have led to redesign the chip in order to eliminate background noise oscillation, in optimal operating conditions of the SALT.

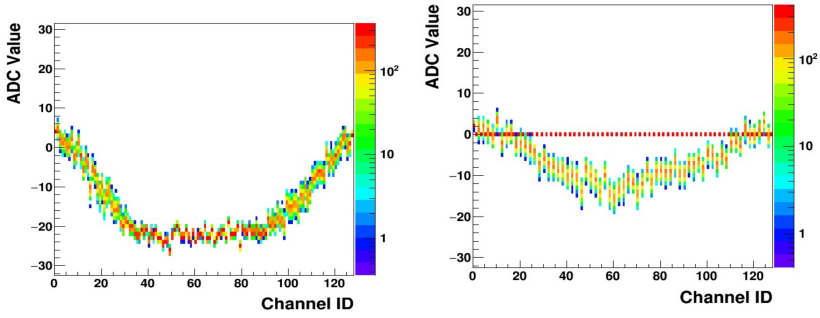
The new SALT chips (v3.5, v3.8 and v3.9) have been produced with the following main modifications:

- the preamplifier has been redesigned to improve the gain. The Krummenacher circuit has been removed and the components have been resized,
- the I_{adc} fluctuations, found as source of the 40 MHz oscillation, have been decreased by adding a dummy current between conversions,
- a new power distribution network has been designed in order to minimize the inductance between digital and analog side, making a better shielding and separation between the supply lines of the two parts.

These versions have been tested, by following a similar testing procedure with respect to what is reported here. Better performances have been obtained: the oscillation at 40 MHz has been reduced resulting in a signal over noise level larger than 10. Therefore, they are found to be within the design specifications.

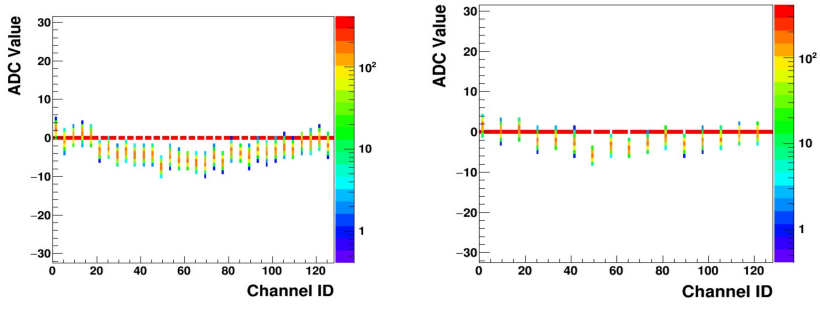
An example of ADC baseline of SALT v3.8 is reported in Fig. 6.15, where the chip is attached to a type-A sensor and is bonded on a 4-chips hybrid. The ADC baseline is constant, around 0 LSB, with a total noise per channel smaller than 1 LSB. A pulse shape scan is also performed by injecting a charge in all ADC channels that mimic a real signal. A signal over noise ratio larger than 13 can be extracted, as shown in Fig. 6.16.

The full production is ongoing and UT is scheduled to be installed in the LHCb cavern in 2021.



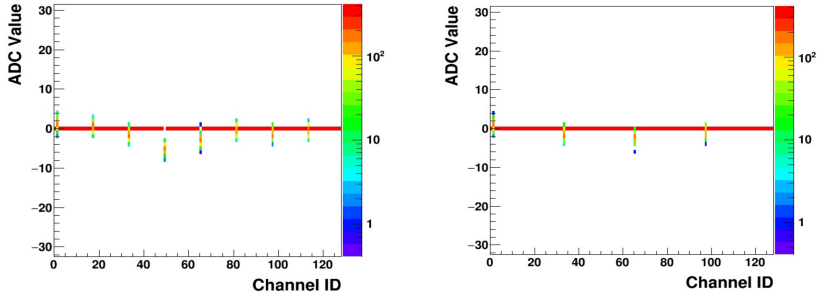
(a) Baseline with 128 channels enabled

(b) Baseline with 64 channels enabled



(c) Baseline with 32 channels enabled

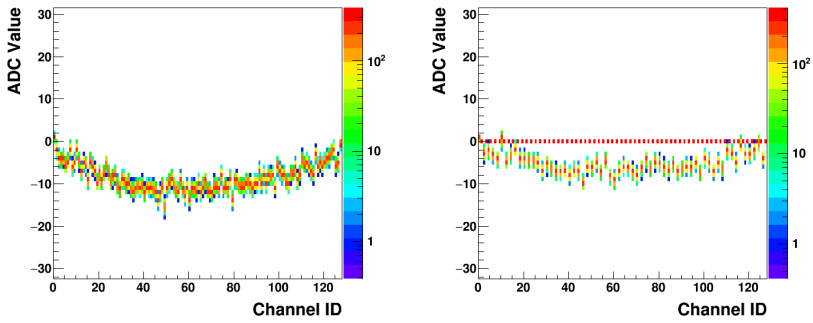
(d) Baseline with 16 channels enabled



(e) Baseline with 8 channels enabled

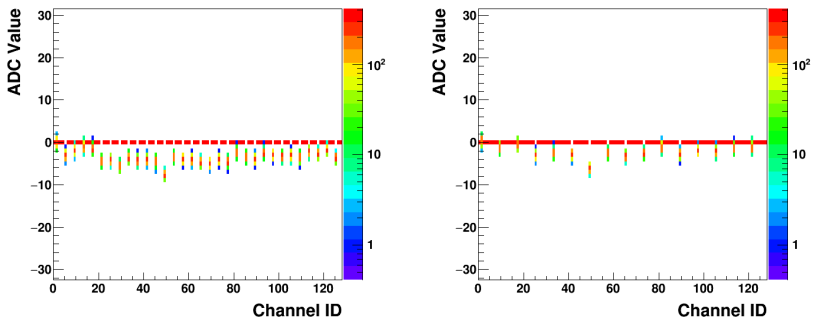
(f) Baseline with 4 channels enabled

Figure 6.12 ADC baseline of the output channels of the ASIC connected to a type-A sensor, as a function of the number of activated channels.



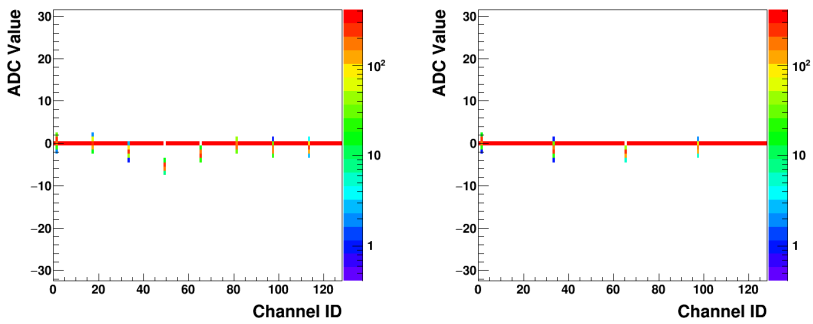
(a) Baseline with 128 channels enabled

(b) Baseline with 64 channels enabled



(c) Baseline with 32 channels enabled

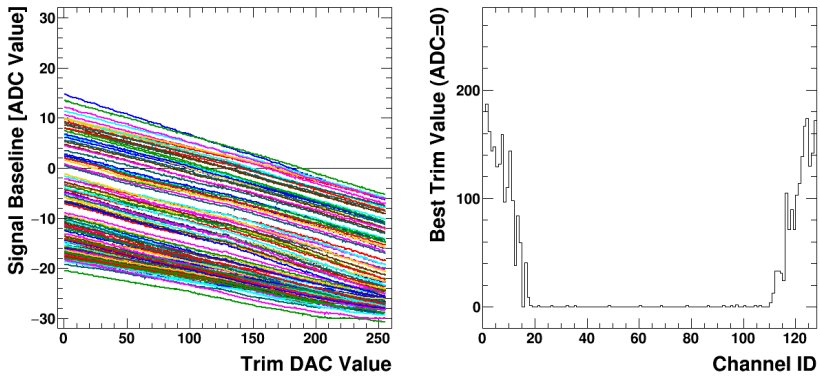
(d) Baseline with 16 channels enabled



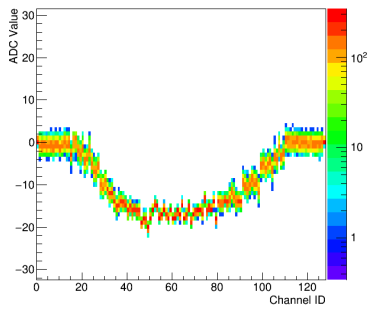
(e) Baseline with 8 channels enabled

(f) Baseline with 4 channels enabled

Figure 6.13 ADC baseline of the output channels of the ASIC connected to a type-A sensor, as a function of the number of activated channels with a lower value of the bias current, equal to 6 μA.



(a) (b)



(c)

Figure 6.14 Pedestal baseline correction. Fig. 6.14a: baseline of the 128 SALT channels as a function of the applied value of TRIMDAC, Fig. 6.14b: best trim DAC value to obtain a zero baseline, Fig. 6.14c: corrected ADC baseline after TRIMDAC subtraction.

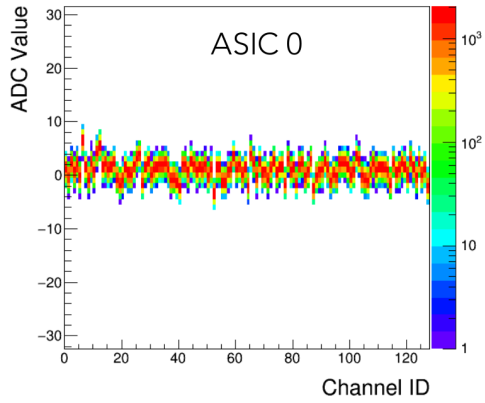


Figure 6.15 ADC baseline for SALT chip v3.8 attached to a 4-ASIC hybrid, tested in Milan in May 2019. Courtesy of the UT-Milano group.

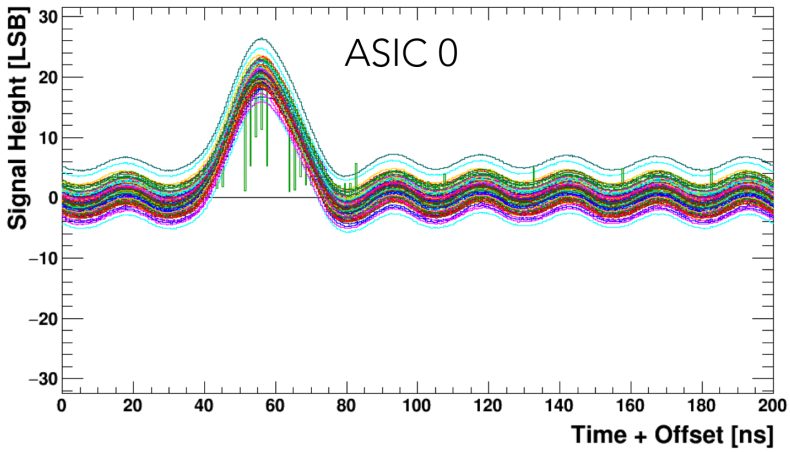


Figure 6.16 Pulse shape as a function of time tested with SALT chip v3.8, Milan May 2019. Courtesy of the UT-Milano group.

CHAPTER 7

Conclusions

The main topic of this thesis is the search for new pentaquark candidates in the $B^0_{(s)} \to J/\psi p\overline{p}$ decays. First of all, we report the observation of $B^0_{(s)} \to J/\psi p\overline{p}$ decays [1], analysed with the data collected by LHCb between 2011 and 2016 and corresponding to a luminosity of 5.2 fb^{-1}. The $B^0_{(s)} \to J/\psi p\overline{p}$ branching ratios are measured to be:

$$\mathcal{B}(B^0 \to J/\psi p\overline{p}) = (4.51 \pm 0.40\ (\text{stat}) \pm 0.44\ (\text{syst})) \times 10^{-7},$$
$$\mathcal{B}(B^0_s \to J/\psi p\overline{p}) = (3.58 \pm 0.19\ (\text{stat}) \pm 0.33\ (\text{syst})) \times 10^{-6},$$

where the first uncertainty is statistical and the second one is systematic. Thanks to the limited available phase space of 393 MeV and 306 MeV for the B^0_s and B^0 mode, respectively, we were able to perform a very precise measurement of their masses. The $B^0_{(s)}$ masses are measured to be $5366.85 \pm 0.19\ (\text{stat}) \pm 0.13\ (\text{sys})$ MeV and $5279.74 \pm 0.30\ (\text{stat}) \pm 0.10\ (\text{sys})$ MeV for the B^0_s and B^0, respectively.

In order to investigate the presence of exotic resonances, we perform a full amplitude analysis of the $B^0_s \to J/\psi p\overline{p}$ decay, using the dataset collected by LHCb during Run 1 and Run 2 and corresponding to an integrated luminosity of 9 fb^{-1}. The analysis revealed an evidence of a new state decaying to $J/\psi p$ and $J/\psi \overline{p}$ not compatible with the previous states observed by LHCb, that can be interpreted as a pentaquark candidate with quark content $c\bar{c}uud$. The statistical significance associated to this excess, using likelihood-ratio based frequentistic tests, is 3.1σ including systematic uncertianties. Its mass, width and fit fraction are measured to be:

$$\mathrm{M}_{P_c} = 4337^{+7}_{-4}(\text{stat})^{+3}_{-3}(\text{syst})\,\text{MeV},$$
$$\Gamma_{P_c} = 29^{+26}_{-12}(\text{stat})^{+14}_{-13}(\text{syst})\,\text{MeV},$$
$$f(P_c) = 14.4^{+8.0}_{-4.0}(\text{stat}) \pm 8.2(\text{syst})\%,$$

where the first uncertainty is statistical and the second one is systematic. The preferred spin hypothesis for $J/\psi p$ corresponds to quantum numbers equal to $J^P = 1/2^+$, even though other J^P assignments cannot be excluded and represents the main source of systematic uncertainty.

Finally, we report the tests and characterization of prototypes of the front-end electronic chip (SALT v2.5) of the Upstream Tracker detector. These tests gave fundamental feedback to identify problems and indicate areas for improvement in the SALT chip design. Subsequent versions of the SALT were found to be within specification, *i.e.* v3.5, v3.8, and v3.9, and are used in the construction of the detector.

APPENDIX A

Supporting material for the analyses of $B^0_{(s)} \to J/\psi\, p\overline{p}$ decays

A.1 Efficiency map plots

In the following, the comparisons between MC samples and the efficiency parametrization are given to show the results for all modes, $B^0_{(s)}$, in Run 1 and Run 2, as discussed in Sec. 5.5.1.

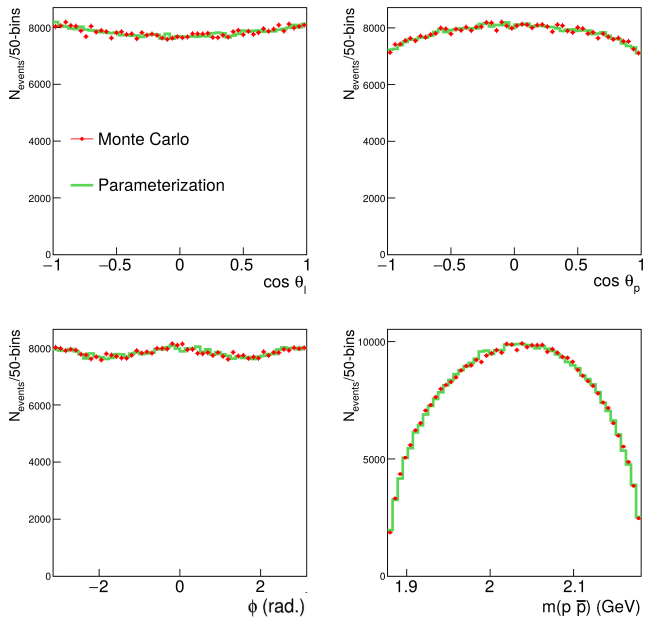
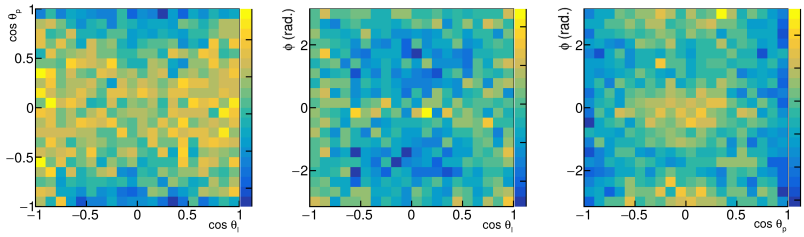
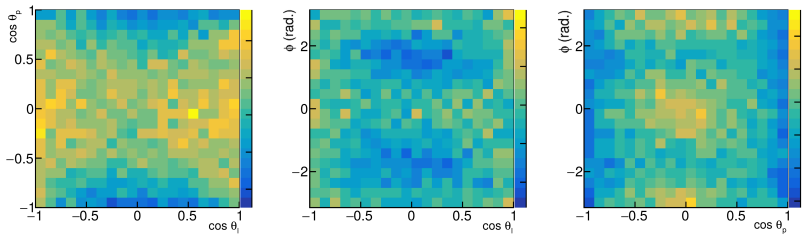


Figure A.1 Run 1 $B^0 \to J/\psi p\overline{p}$: 1-d comparions in the kinematic variables between the reconstructed MC distributions and the efficiency parametrization.



(a)



(b)

Figure A.2 Run 1 $B^0 \to J/\psi p\overline{p}$: 2-d scatter plots in the angular variables for the (a) reconstructed MC distributions and the (2) efficiency parametrization. The features in (a) are reproduced in (b), up to statistical fluctuations.

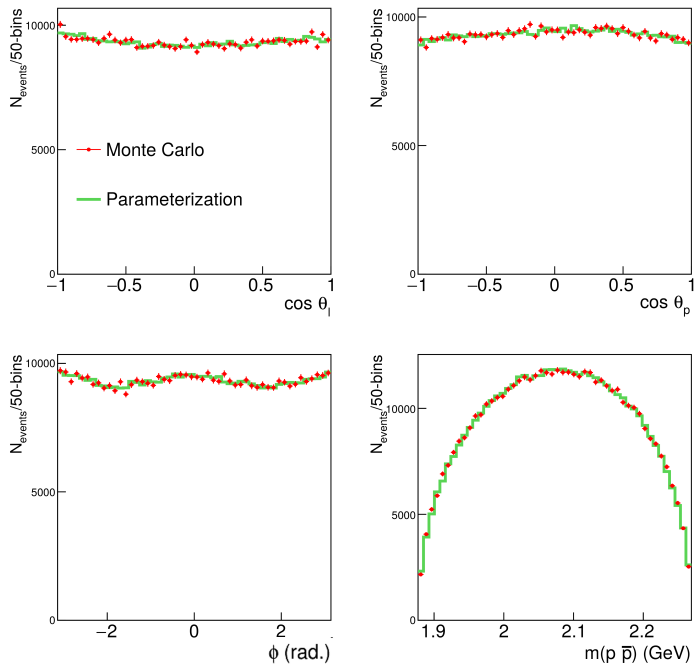
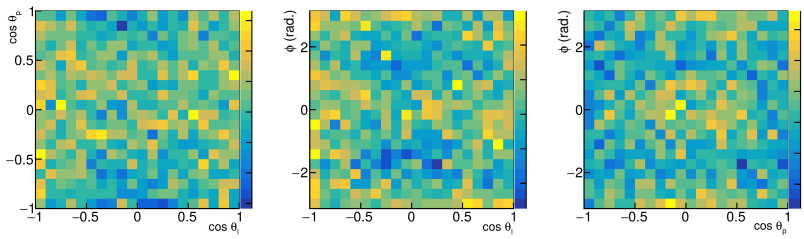
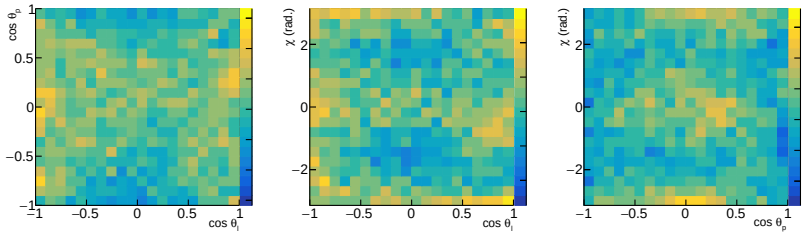


Figure A.3 Run 2 $B^0_s \to J/\psi p\bar{p}$: 1-d comparisons in the kinematic variables between the reconstructed MC distributions and the efficiency parametrization.



(a)



(b)

Figure A.4 Run 2 $B^0_s \to J/\psi p\bar{p}$: 2-d scatter plots in the angular variables for the (a) reconstructed MC distributions and the (b) efficiency parametrization. The features in (a) are reproduced in (b), up to statistical fluctuations.

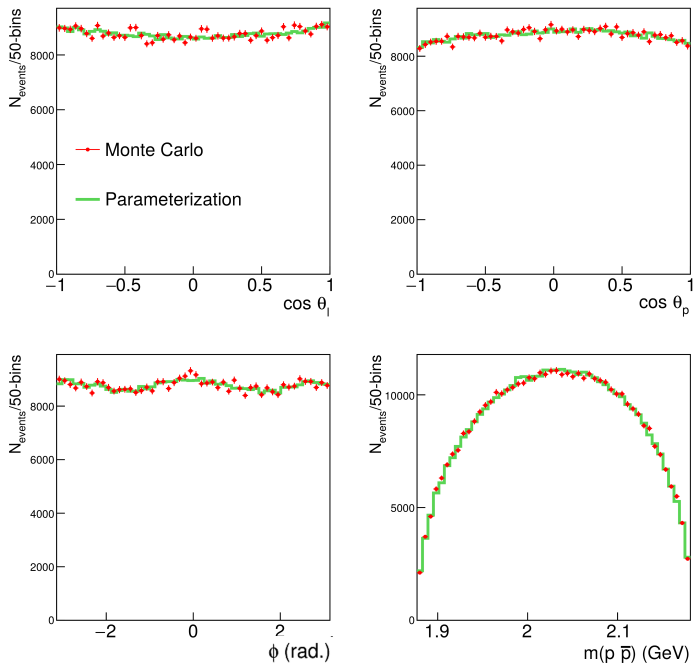


Figure A.5 Run 2 $B^0 \to J/\psi p\overline{p}$: 1-d comparions in the kinematic variables between the reconstructed MC distributions and the efficiency parameterization.

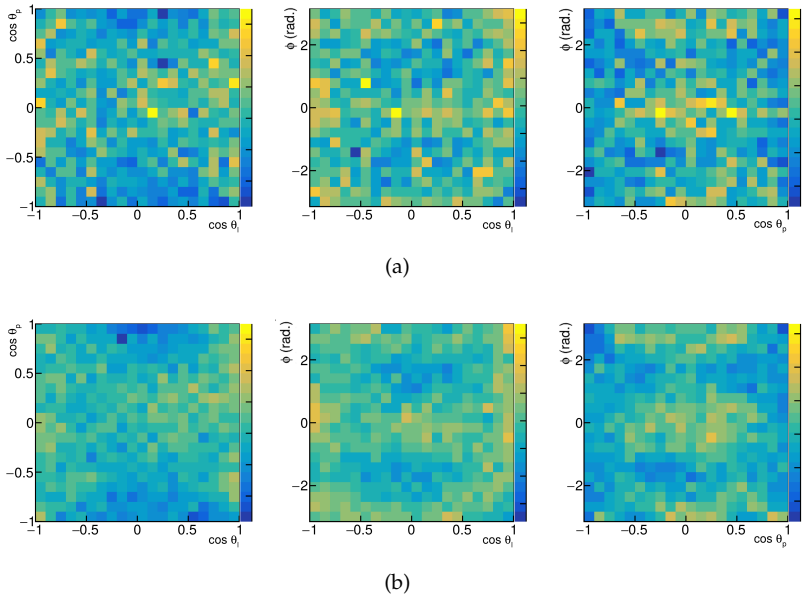


Figure A.6 Run 2 $B^0 \to J/\psi p\overline{p}$: 2-d scatter plots in the angular variables for the (a) reconstructed MC distributions and the (b) efficiency parametrization. The features in (a) are reproduced in (b), up to statistical fluctuations.

A.2 Amplitude definition

A.2.1 Dalitz-plot decomposition amplitude

An alternative amplitude formula is given in this section, adapting the work presented in Ref. [112], called Dalitz-plot decomposition, to the $B^0_s \to J/\psi p\overline{p}$ decay. The amplitude of the $B^0_s \to J/\psi p\overline{p}$ decay can be decomposed into two terms, called A and B, related to $B^0_s \to J/\psi p\overline{p}$ and $J/\psi \to \mu^+\mu^-$ decays respectively, as:

$$\begin{aligned} M_{\lambda,\sigma,\lambda_+,\lambda_-} &= \sum_{\lambda_\psi} A_{\lambda,\sigma,\lambda_\psi} B^{\lambda_\psi}_{\lambda_+\lambda_-} \\ &= \sum_{\lambda_\psi} O_{\lambda,\sigma,\lambda_\psi}(\{\sigma\}) \left[\sqrt{3} D^{1*}_{\lambda_\psi,\lambda_+-\lambda_-}(\phi_\mu,\theta_\mu,0)\, H^{1\to\mu^+,\mu^-}_{\lambda_+,\lambda_-}\right] \end{aligned} \tag{A.2.1}$$

where $\{\lambda,\sigma\}$ are the proton and $\overline{p}$ final particles helicities, $\{\lambda_p,\lambda_{\overline{p}}\}$, λ_ψ is the intermediate helicity of the J/ψ and λ_+,λ_- are the muon helicities, which are defined in the J/ψ rest frame. The angles ϕ_μ,θ_μ of the J/ψ decay are the same defined in Eq. 5.5. The overall D function that rotates the canonical configuration to the actual one is absent because the B has spin zero. Due to parity conservation in the electromagnetic decay, the helicity couplings $H^{1\to\mu^+,\mu^-}_{\lambda_+,\lambda_-}$ are reduced to a single term because the contribution with $\lambda_+ - \lambda_- = 0$ can be neglected since $m_\psi \gg m_\mu$.

The Dalitz-plot function $O_{\lambda,\sigma,\lambda_\psi}(\{\sigma\})$ can be expressed as

$$\begin{aligned} O_{\lambda,\sigma,\lambda_\psi}(\{\sigma\}) = &\sum_s^{X\to p\overline{p}} \sum_\tau \sqrt{2} n_s \delta_{0,\tau-\lambda_\psi} H^{0\to(23),1}_{\tau,\lambda_\psi} X_s(\sigma_1)\, d^s_{\tau,\lambda-\sigma}(\theta_{23})\, H^{(23)\to 2,3}_{\lambda,\sigma} \\ &+ \sum_s^{P_c^+\to J/\psi p} \sum_{\tau,\mu',\lambda',\sigma'} \sqrt{2} n_s H^{0\to(12),3}_{\tau,\tau} X_s(\sigma_3)\, d^s_{\tau,\mu'-\lambda'}(\theta_{12})\, H^{(12)\to 1,2}_{\mu',\lambda'} \qquad \text{(A.2.2)} \\ &\quad \cdot d^1_{\mu',\lambda_\psi}\left(\zeta^1_{3(1)}\right) d^{1/2}_{\lambda',\lambda}\left(\zeta^2_{2(1)}\right) d^{1/2}_{\sigma',\sigma}\left(\zeta^3_{2(1)}\right) \\ &+ \sum_s^{P_c^-\to \overline{p} J/\psi} \sum_{\tau,\mu',\lambda',\sigma'} \sqrt{2} n_s H^{0\to(31),2}_{\tau,\tau} X_s(\sigma_2)\, d^s_{\tau,\sigma'-\mu'}(\theta_{31})\, H^{(31)\to 3,1}_{\sigma',\mu'} \qquad \text{(A.2.3)} \\ &\quad \cdot d^1_{\mu',\lambda_\psi}\left(\zeta^1_{3(1)}\right) d^{1/2}_{\lambda',\lambda}\left(\zeta^2_{3(1)}\right) d^{1/2}_{\sigma',\sigma}\left(\zeta^3_{3(1)}\right) \qquad \text{(A.2.4)} \end{aligned}$$

where the first sum is over the spin of the intermediate resonance s and the second one over the helicities in the isobar helicity frame, where the letters $\tau,\mu',\lambda',\sigma'$ refer to the intermediate resonance (P_c or X), the J/ψ, the p and the $\overline{p}$ helicities, respectively. While the indexes 1,2,3 refer to J/ψ, p and $\overline{p}$. The helicity couplings enter the equation with the symbol H and contains the particle-2 factor, $H_{\lambda_1,\lambda_2} = (-1)^{J_2-\lambda_2} h_{\lambda_1,\lambda_2}$. The helicity of P_c^+ and P_c^- are assumed to be equal except for the particle-2 factor and the permutation factor as in Eq. 5.16: $h^{(31)\to 3,1}_{\sigma',\mu'} = (-1)^{J_{P_c}-3/2} h^{(12)\to 1,2}_{\mu',\lambda'}$ and $h^{0\to(12),3}_{\tau,\tau} = -h^{0\to(31),2}_{\tau,\tau}$. The lineshape dependence is contained in the term $X_s(\sigma)$, where σ is the Mandelstam variable. The polar rotations of the angles $\theta_{23},\theta_{12},\theta_{31}$ are performed in the isobar helicity frames of X, P_c^+, P_c^-, respectively, and are equal to the angles $\theta_p, \theta^{\{P_c^+\}}_{P_c}$ and $\theta^{\{P_c^-\}}_{P_c}$, entering Eq. 5.5, 5.7, 5.8. Instead, the final Wigner rotations align the helicities

of the J/ψ, p and $\overline{p}$ from the isobar helicity frame to the aligned configuration. The angle $\zeta^i_{k(1)}$ is defined in the i reference frame and is equivalent to the rotation between the isobar k and the isobar 1. The Wigner rotations of p and $\overline{p}$ helicities are identical to the ones used in Eq. 5.32, while the angle of J/ψ is not needed in the standard helicity formalism because the J/ψ is treated as an intermediate particle.

A.2.2 Covariant formalism

The model has also been validated making a comparison with the covariant approach [105] for the simplified model of a stable J/ψ.

The two formalism are equivalent, even if they have pros and cons. The helicity formalism is more efficient in determining the quantum numbers associated to resonances because it can be decomposed in the LS basis and, hence, be associated to a specific spin of the resonance. While, this is not the case for the covariant. However, the covariant formalism can be used as a cross-check for the final spin rotations, because they are already embedded into the model in the spinor definition. So no care must be taken to align the spin of final particles in different chains.

In the following, we report the amplitude definition chain by chain. We consider the lagrangian

$$\mathcal{L} = B\bar{p}\gamma^\mu p\psi_\mu \tag{A.2.5}$$

with the B representing a field with (provisionally) quantum numbers 0^{++}, so that we can enforce C and P conservation in the decay. In the X decay chain, the X is forced to have 1^{--} quantum numbers. One gets the following decay amplitude:

$$\mathcal{A}_{\lambda_\psi,\lambda_p,\lambda_{\bar{p}}} = \frac{1}{2m_p}\epsilon^*_\mu(\lambda_\psi)\bar{u}(\lambda_p)\gamma^\mu v(\lambda_{\bar{p}}) \tag{A.2.6}$$

$$\tag{A.2.7}$$

which in terms of helicity couplings becomes:

$$\mathcal{A}_{\lambda_\psi,\lambda_p,\lambda_{\bar{p}}} = H_{\lambda_\psi}H_{\lambda_p,\lambda_{\bar{p}}}d^1_{\lambda_\psi,\lambda_p-\lambda_{\bar{p}}}(\theta_p) \tag{A.2.8}$$

where H are the helicity couplings defined in terms of the energy, namely $H_1 = H_{1/2,1/2} = 1$, $H_0 = -E^X_\psi/m_\psi$, $H_{1/2,-1/2} = \sqrt{2}E^X_p/m_p$.

For the P_c^+ decay chain, the same lagrangian corresponds to a P_c^+ with quantum numbers $1/2^-$. A projector $(\gamma^0 - 1)/2$ is needed to have a definite parity state. We get

$$\mathcal{A}_{\lambda_\psi,\lambda_p,\lambda_{\bar{p}}} = \frac{1}{2m_p}\left(\epsilon^*_\mu(\lambda_\psi)\bar{u}(\lambda_p)\gamma^\mu\frac{\gamma^0-1}{2}\right)v(\lambda_{\bar{p}}) \tag{A.2.9}$$

$$\mathcal{A}_{\lambda_\psi,\lambda_p,\lambda_{\bar{p}}} = H_{\lambda_{\bar{p}}}H_{\lambda_\psi,\lambda_p}d^{1/2}_{\lambda_{\bar{p}},\lambda_\psi-\lambda_p}(\theta_\psi) \tag{A.2.10}$$

with $H_{1/2} = -\sqrt{E^{P_c^+}_{\bar{p}} + m_p}/\sqrt{2}m_p$, $H_{1,1/2} = \sqrt{E^{P_c^+}_p + m_p}$, $H_{1,-1/2} = 0$, and $H_{0,1/2} = \left[\sqrt{E^{P_c^+}_p + m_p}E^{P_c^+}_\psi + \sqrt{E^{P_c^+}_p - m_p}p^{P_c^+}_\psi\right]\Big/\sqrt{2}m_\psi$. Similarly for the P_c^- with $J^P = 1/2^+$,

$$\mathcal{A}_{\lambda_\psi,\lambda_p,\lambda_{\bar{p}}} = \frac{1}{2m_p}\bar{u}(\lambda_p)\left(\frac{\gamma^0+1}{2}\gamma^\mu v(\lambda_{\bar{p}})\epsilon^*_\mu(\lambda_\psi)\right) \tag{A.2.11}$$

$$\mathcal{A}_{\lambda_\psi,\lambda_p,\lambda_{\bar{p}}} = H_{\lambda_p}H_{\lambda_{\bar{p}},\lambda_\psi}d^{1/2}_{\lambda_p,\lambda_{\bar{p}}-\lambda_\psi}(\theta_{\bar{p}}) \tag{A.2.12}$$

with $H_{1/2} = -\sqrt{E_p^{P_c^-} + m_p}/\sqrt{2}m_p$, $H_{1/2,1} = \sqrt{E_{\bar{p}}^{P_c^-} + m_p}$, $H_{1/2,-1} = 0$, and $H_{1/2,0} = \left[\sqrt{E_{\bar{p}}^{P_c^-} + m_p}E_\psi^{P_c^-} + \sqrt{E_{\bar{p}}^{P_c^-} - m_p}p_{\bar{p}}^{P_c^-}\right]\Big/\sqrt{2}m_\psi$.

The numerical results of the two models for a stable J/ψ agree perfectly.

A.3 Relation between amplitudes with opposite P eigenvalues

In order to prove that the interference cancels between orthogonal P eigenstates for the simplified case of a stable J/ψ, we consider the term that enters in the interference between 2 states with quantum numbers J_1 and J_2 in the decay chain of $B \to J/\psi X(\to p\bar{p})$:

$$\mathcal{I}^X = \sum_{\lambda_p,\lambda_{\bar{p}},\lambda_\psi} \mathcal{H}_{\lambda_\psi}\mathcal{H}_{\lambda_p,\lambda_{\bar{p}}} \cdot d^{J_1}_{\lambda_\psi,\lambda_p-\lambda_{\bar{p}}}(\theta_p)\mathcal{H'}_{\lambda_\psi}\mathcal{H'}_{\lambda_p,\lambda_{\bar{p}}} \cdot d^{J_2}_{\lambda_\psi,\lambda_p-\lambda_{\bar{p}}}(\theta_p). \tag{A.3.1}$$

Since we sum over the helicities, they are mute indexes. Therefore, in order to see what happens under P, we replace the helicities $\{\lambda_\psi \to -\lambda_\psi, \lambda_p \to -\lambda_p, \lambda_{\bar{p}} \to -\lambda_{\bar{p}}\}$ because we know that this is equivalent to apply the parity P.

Eq. A.3.1 becomes:

$$\begin{aligned}
\mathcal{I}^X &= \sum_{\lambda_p,\lambda_{\bar{p}},\lambda_\psi} \mathcal{H}_{-\lambda_\psi}\mathcal{H}_{-\lambda_p,-\lambda_{\bar{p}}} \cdot d^{J_1}_{-\lambda_\psi,-\lambda_p+\lambda_{\bar{p}}}(\theta_p)\mathcal{H'}_{-\lambda_\psi}\mathcal{H'}_{-\lambda_p,-\lambda_{\bar{p}}} \cdot d^{J_2}_{-\lambda_\psi,-\lambda_p+\lambda_{\bar{p}}}(\theta_p) \\
&= \sum_{\lambda_p,\lambda_{\bar{p}},\lambda_\psi} \eta_B\eta_\psi\eta_X^2\eta_p\eta_{\bar{p}}\mathcal{H}_{\lambda_\psi}\mathcal{H}_{\lambda_p,\lambda_{\bar{p}}} \cdot (-)^{\lambda_\psi-\lambda_p+\lambda_{\bar{p}}} d^{J_1}_{\lambda_\psi,\lambda_p-\lambda_{\bar{p}}}(\theta_p) \\
&\qquad \times \eta'_B\eta'_\psi(\eta'_X)^2\eta_p\eta_{\bar{p}}\mathcal{H'}_{\lambda_\psi}\mathcal{H'}_{\lambda_p,\lambda_{\bar{p}}} \cdot (-)^{\lambda_\psi-\lambda_p+\lambda_{\bar{p}}} d^{J_2}_{\lambda_\psi,\lambda_p-\lambda_{\bar{p}}}(\theta_p) \\
&= P_B P'_B \mathcal{I}^X
\end{aligned} \tag{A.3.2}$$

where we have substituted the following relations:

- for the properties of the Wigner-d matrixes:

$$d^{J_2}_{-\lambda_\psi,-\lambda_p+\lambda_{\bar{p}}}(\theta_p) = (-)^{\lambda_\psi-\lambda_p+\lambda_{\bar{p}}} d^{J_2}_{\lambda_\psi,\lambda_p-\lambda_{\bar{p}}}(\theta_p), \tag{A.3.3}$$

- for the helicity couplings, the product of the naturalities of the initial and final states:

$$\mathcal{H}_{-\lambda_\psi} = \eta_B\eta_\psi\eta_X\mathcal{H}_{\lambda_\psi} \tag{A.3.4}$$

$$\mathcal{H}_{-\lambda_p,-\lambda_{\bar{p}}} = \eta_X\eta_p\eta_{\bar{p}}\mathcal{H}_{\lambda_p,\lambda_{\bar{p}}} \tag{A.3.5}$$

  where η are the naturalities.

Knowing that the naturality of the B is equal to $\eta_B = P_B(-1)^{J=0}$, we obtain the last equality.

Therefore, the interference between 2 X states with J_1 and J_2 quantum numbers are equal except for the product of the intrinsic parities of the B. Therefore, if $P_B P_{B'} = -1$ (or equivalently if they have opposite P eigenvalue), the interference cancels. Since this results is achieved for a specific value of λ_ψ, considering also the helicity term $\mathcal{H}_{\lambda_\psi}$, the cancellation is valid only for a stable J/ψ. However, it is expected to be negligible when the J/ψ decay is considered. The same reasoning can be applied to the interference between different chains, because it always depends on the intrinsic parities of the B decay.

A.4 Cross-checks for amplitude analysis

A.4.1 Alternative J^P for null-hypothesis model

Other possibilities of J^P for the null-hypothesis model have also been investigated but none of them is found to be a good alternative model for it. We present them below:

1. Model with $J^P = 1^+$: it cannot fit the data especially the enhancement at low $p\overline{p}$ values due to few values of L allowed. Indeed, the dependence to L appears in the threshold factor p^L which suppresses the higher partial waves. In fact, only if angular momentum l of the decay, $X \to p\overline{p}$, assumes value equal to 0, the enhancement at low $p\overline{p}$ can be fitted. The discrepancy with the nominal fit is also evident from the $\Delta \log \mathcal{L}$ values, which is equal to -2 for $J^P = 1^+$ instead of -70.2. Hence this model can be excluded (see fig. A.7 for $J^P = 1^+$).

2. Models with $J^P = 0^{+/-}$ do not represent the data especially in $cos(\theta_l)$ distribution, as can be seen in Fig. A.8, A.9. The $\Delta \log \mathcal{L}$ of $J^P = 0^-$ with these model is larger than zero.

3. Alternative model: NR $J^P = 0^{+/-}$ + Resonant 0^-. The phase space is fitted only by the pp enhancement, which is unphysical, and the NR fraction is 0. See fig. A.10.

4. Alternative model: NR $J^P = 1^+$ + Resonant 0^-. It is dominated by the $p\overline{p}$ enhancement and the mass of it goes to the edge of the allowed range. The $\log \mathcal{L}$ is also much smaller than in the nominal fit (~ -30) because other phase space variables are not well reproduced. For these reasons is not considered as a good alternative model. See fig. A.11.

To conclude, no other null hypothesis models with different quantum numbers have been found, apart from the nominal (Sec. 5.6.1) and its alternative one (Sec. 5.6.2), used to evaluate systematic uncertainty.

A.4.2 Interference

In order to validate the model with the new prescriptions as described in Sec. 5.1.2, the interference is studied in the CP basis where it is known that it cancels out for orthogonal CP eigenstates.

As discussed in Sec. 5.1.4, the state in the X chain with $J^P = 1^{--}$ is already a CP-even eigenstate, while the two P_c states are not. However, a combination of P_c^+ and P_c^- amplitudes with the same couplings becomes an eigenstate. If the sum was an eigenstate with negative eigenvalue, the difference would be orthogonal. For our model, it is found that the following holds:

$$\begin{aligned} CP(|P_c^+\rangle + (-1)^{J_{P_c}-3/2}|P_c^-\rangle) &= +(|P_c^+\rangle + (-1)^{J_{P_c}-3/2}|P_c^-\rangle), \\ CP(|P_c^+\rangle - (-1)^{J_{P_c}-3/2}|P_c^-\rangle) &= -(|P_c^+\rangle - (-1)^{J_{P_c}-3/2}|P_c^-\rangle). \end{aligned} \tag{A.4.1}$$

Indeed, zero interference is present between the two orthogonal states: $|X>$ and $|P_c^+\rangle - (-1)^{J_{P_c}-3/2}|P_c^-\rangle$. As shown in Fig. A.12 for spin $J^P = 1/2^-$, the sum of the two P_c amplitudes is non-zero in bins of the $m_{p\overline{p}}$ invariant mass, while the interference is exactly zero. In addition, it is also checked that the interference is recovered if the difference between the P_c is used.

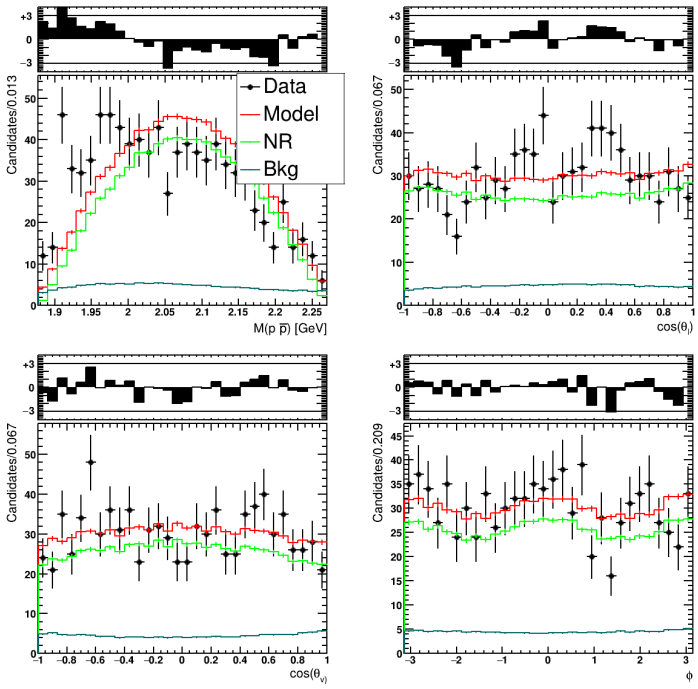


Figure A.7 Phase space distributions for a NR model with $J^P = 1^+$

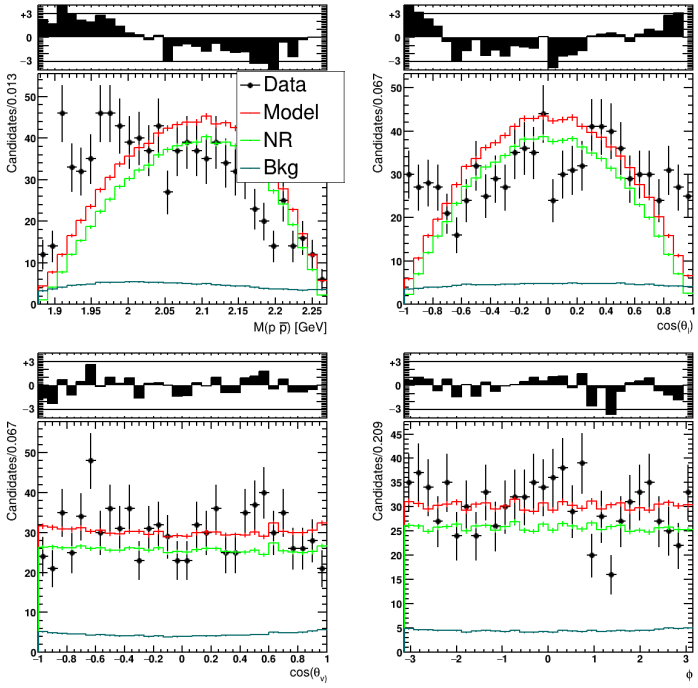


Figure A.8 Phase space distributions for a NR model with $J^P = 0^+$

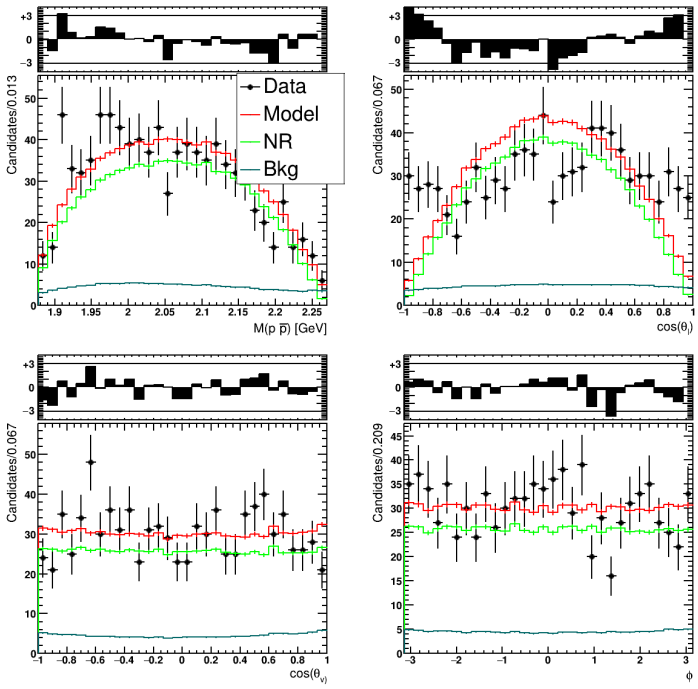


Figure A.9 Phase space distributions for a NR model with $J^P = 0^-$

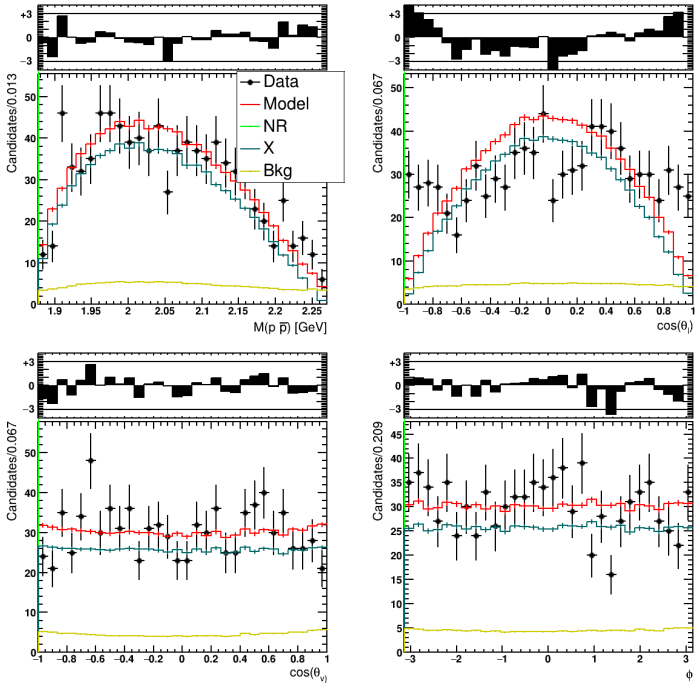


Figure A.10 Phase space fit with alternative model described in item 3.

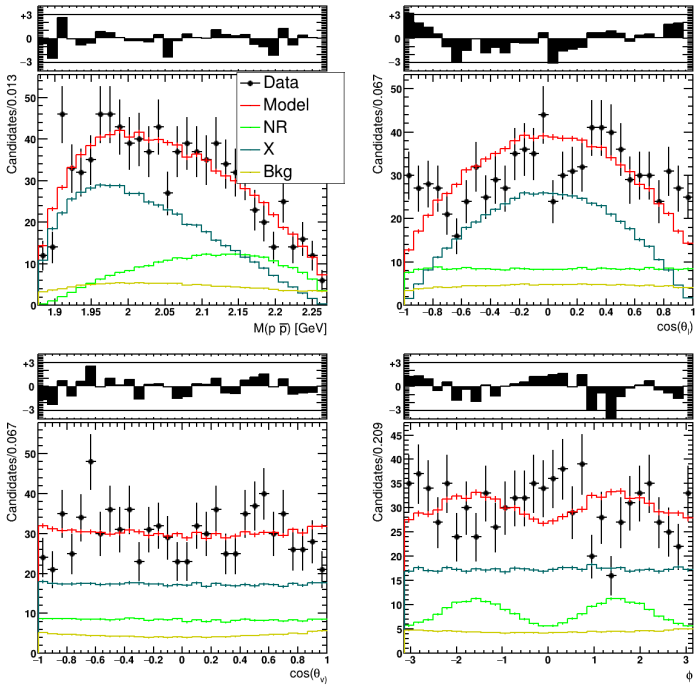


Figure A.11 Phase space fit with alternative model described in item 4.

A.4.3 Number of P_c couplings

In this section, we want to prove that we are sensitive only to a single effective coupling for the opposite charged states, P_c^+ and P_c^-. In order to achieve that, we generate a toy with the same statistics as in data with two different values of the P_c couplings ($(A,\phi)_{P_c^+} = (0.15, 2.8)$ and $(A,\phi)_{P_c^-} = (0.3, 1.4)$) and then fit it to see which are the parameters that can be extracted. In Fig. A.13 and in Tab. A.1, the fit results are reported for $J^P = 1/2^-$. The values of the couplings are compatible between the two. The mass projection of $J/\psi p$ and $J/\psi \bar{p}$ are also overlapped to show that the two distributions are identical within statistical fluctuations. It is also tested that the relative phase between the two couplings cannot be extracted from the fit, indeed a value different from generation is fitted, as reported in Tab. A.2. For these reasons, the two couplings have been fixed equal with a relative phase, following Eq. 5.16.

In order to test the effect of this choice on the mass, width and fit fractions of the P_c states, toys are performed, by generating with the nominal model and fitting with an alternative one, where two couplings for the P_c states are allowed to vary. No bias of the residual distributions is observed as shown in Fig. A.14. Therefore, the systematic effect is considered negligible.

A.4.4 Pulls

In this section, pulls with 10 times the statistics of data are shown in Fig. A.15, in order to prove that the asymmetric shape of some parameters is due to the low statistics.

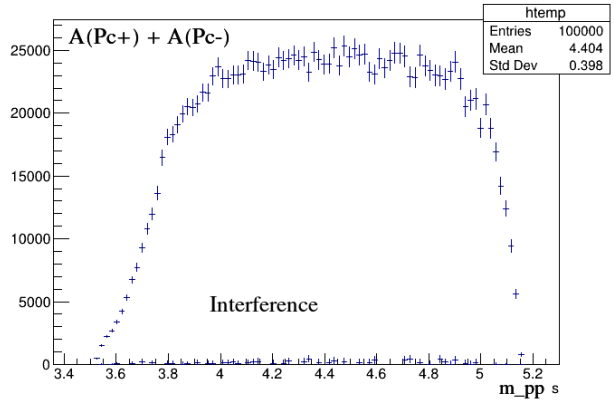


Figure A.12 Sum of P_c^+ and P_c^- amplitudes and interference with the X chain in bins of $m_{p\bar{p}}$ for $J^P = 1/2^-$.

Parameter	Initial value	Fitted value	Error
A(NR) L=2	1.23	0.86	0.19
ϕ(NR) L=2	1.6	1.45	0.16
A(NR) l=2	0.189	0.005	0.54
A(P_c^+)	0.15	0.10	0.15
ϕ(P_c^+)	2.8	2.96	1.79
A(P_c^-)	0.3	0.06	0.16
ϕ(P_c^-)	1.4	-0.4	2.8
Mass(P_c)	4.338	4.343	0.004
Width(P_c)	0.023	0.0234	0.007

Table A.1 Fit results for toy with 2 different P_c initial values of couplings: A(P_c^+) $= 0.14$ and A(P_c^-) $= 0.3$.

A.4.5 Alternative fit with $p\bar{p}$ enhancement

The alternative hypothesis fit with the $p\bar{p}$ enhancement is reported here. On top of the model described in Sec. 5.6.2, two P_c states are considered.

The fit results with the full model are gathered in Table A.3 and the projections over the phase-space variables are reported in Fig. A.17. The angular distributions in the P_c^+ chain are also shown in Fig. A.18.

The significance of different spin hypothesis of the P_c are tested to prove that the choice of the null hypothesis does not influence them. Indeed, this alternative model gives the same ranking for J^P of the P_c states as the nominal one: the preferred hypothesis is $1/2^+$ and the others have lower significance, compatible with the nominal model. The slightly higher significance is due to a worst $\log \mathcal{L}$ of the null-hypothesis. The values are also summarised in Table A.4.

A.4.6 Fit with all LS couplings

The nominal fit with P_c states is performed reducing the orbital momentum to the lowest value possible. In this check, we want to verify if we are sensitive to the inclusion of higher L in the P_c model. In Tab.A.5 the fit results with $J^P = 1/2^+$ are listed. Upper-

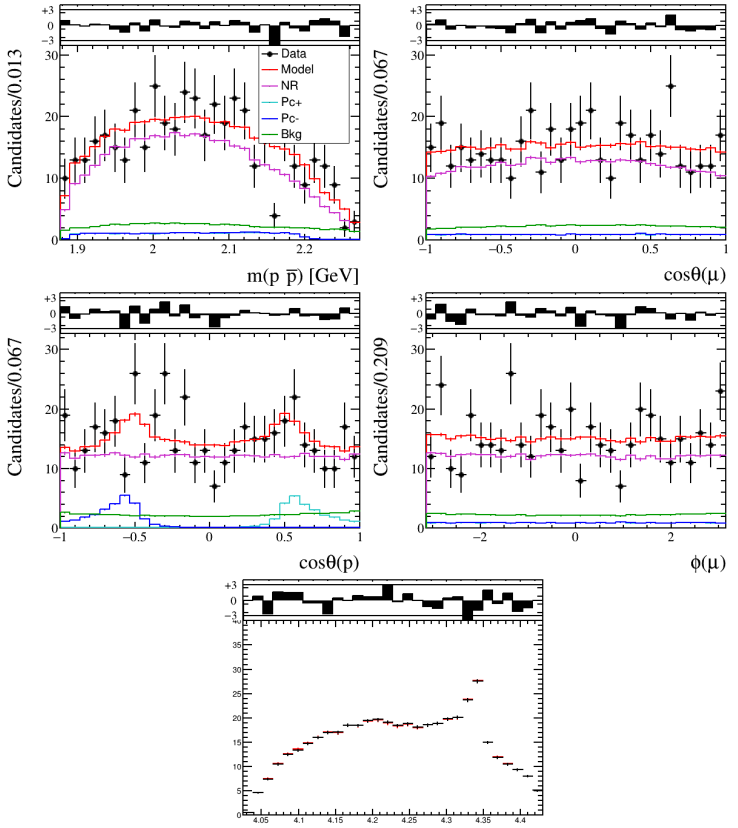


Figure A.13 Phase-space distribution and plot of $J/\psi\, p$ and $J/\psi\, \overline{p}$ invariant masses superimposed.

case L refers to the orbital momentum in the B^0_s RF, while lowercase l to the one in the P_c RF. Since the couplings associated to higher orbital momentum are compatible with zero, we may draw the conclusion that we are not able to extract this information, probably due to the low statistics.

As done for the case of two P_c couplings, we have performed toys to evaluate whether the choice of reducing the number of LS affects our results and whether a systematic needs to be assigned. Toys are performed by generating with the nominal model and fitting with the alternative one. No appreciable bias is found, as shown in the residual distribution in Fig. A.19.

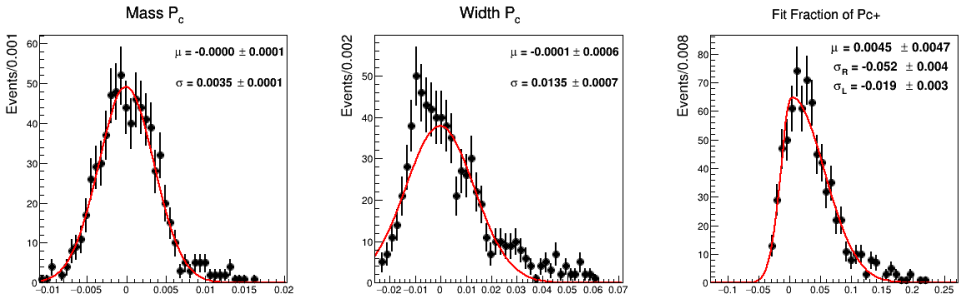


Figure A.14 Residuals for mass, width and fit fraction of the P_c states for the alternative model with two P_c couplings, described in Sec. A.4.3. No systematic effect is assigned.

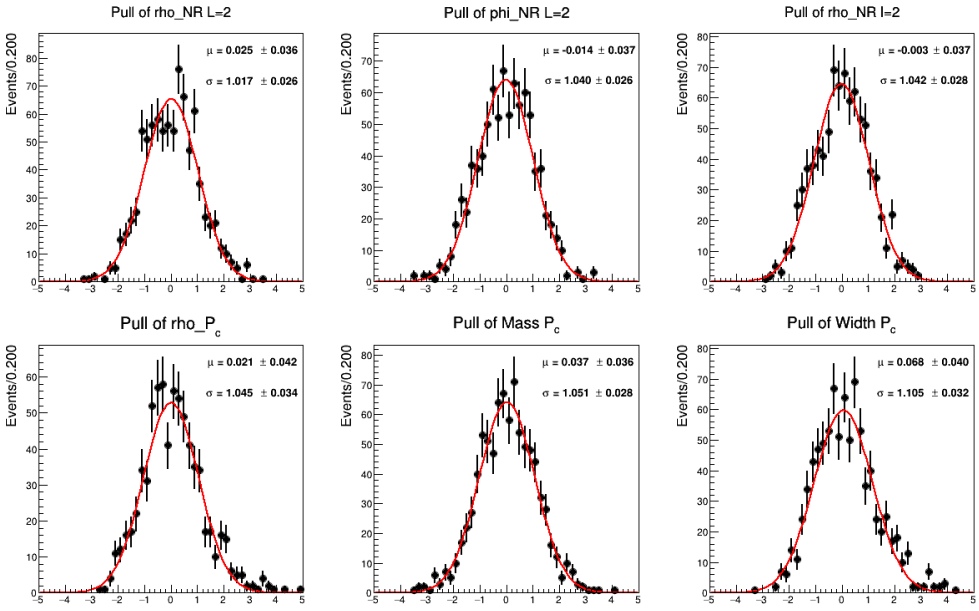


Figure A.15 Pulls for $J^P = 1/2^+$ with a statistics 10 times larger than the one on data.

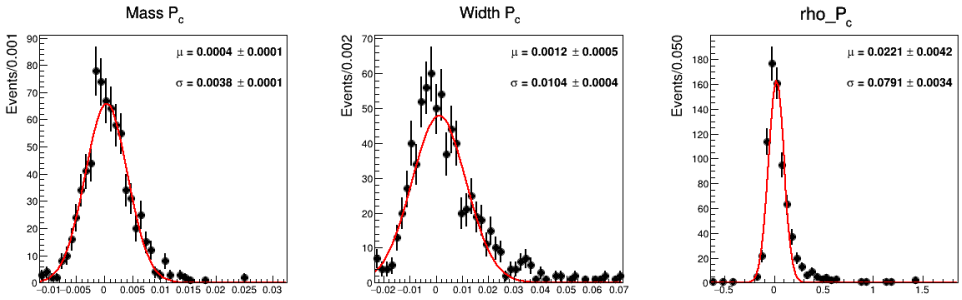


Figure A.16 Residual distributions for fit parameters of the P_c states.

Parameter	Initial value	Fitted value	Error
A(NR) L=2	1.23	1.038	0.22
ϕ(NR) L=2	1.6	1.54	0.14
A(NR) l=2	0.189	0.05	0.11
A(P_c)	0.15	0.10	0.03
$\phi(P_c)$	2.8	2.23	0.69
$\phi\ rel$	1.4	3.12	1.2
Mass(P_c)	4.338	4.343	0.003
Width(P_c)	0.023	0.022	0.005

Table A.2 Fit results for toy with a relative phase between the two, with an initial value of 1.4.

Parameter	Value	Error
mass(X)	1.93	0.03
width(X)	0.21	0.07
A(X)	0.45	0.16
A(P_c)	0.17	0.13
Mass(P_c)	4.334	0.003
Width(P_c)	0.025	0.019
loglh	-85.4	

Table A.3 Fit results for the alternative hypothesis fit with $p\bar{p}$ enhancement.

J^P	-2 ($\Delta \log \mathcal{L}$)	Significance σ	M_0 [MeV]	Γ_0 [MeV]	A(P_c)
$1/2^-$	28.6	3.86 ± 0.10	$4335\ ^{+3}_{-3}$	$24\ ^{+10}_{-8}$	$0.12\ ^{+0.05}_{-0.04}$
$1/2^+$	30.8	4.07 ± 0.10	$4337\ ^{+4}_{-3}$	$33\ ^{+18}_{-12}$	$0.18\ ^{+0.10}_{-0.06}$
$3/2^-$	19.7	2.80 ± 0.10	$4336\ ^{+3}_{-3}$	$18\ ^{+9}_{-7}$	$0.09\ ^{+0.04}_{-0.03}$
$3/2^+$	15.2	2.18 ± 0.10	$4333\ ^{+2}_{-3}$	$12\ ^{+7}_{-5}$	$0.06\ ^{+0.03}_{-0.02}$

Table A.4 Comparison among different J^P hypothesis for the alternative model. Mass and width of the P_c of all hypothesis are also reported.

Parameter	Value	Error
A(NR) L=2	1.15	0.21
ϕ(NR) L=2	1.63	0.17
A(NR) l=2	0.18	0.09
A(P_c^+) L=1	0.22	0.52
A(P_c^+) l=1, s=1/2	0.18	0.11
A(P_c^+) l=1, s=3/2	0.03	0.06
Mass(P_c)	4.338	0.004
Width(P_c)	0.030	0.017

Table A.5 Fit results with all couplings.

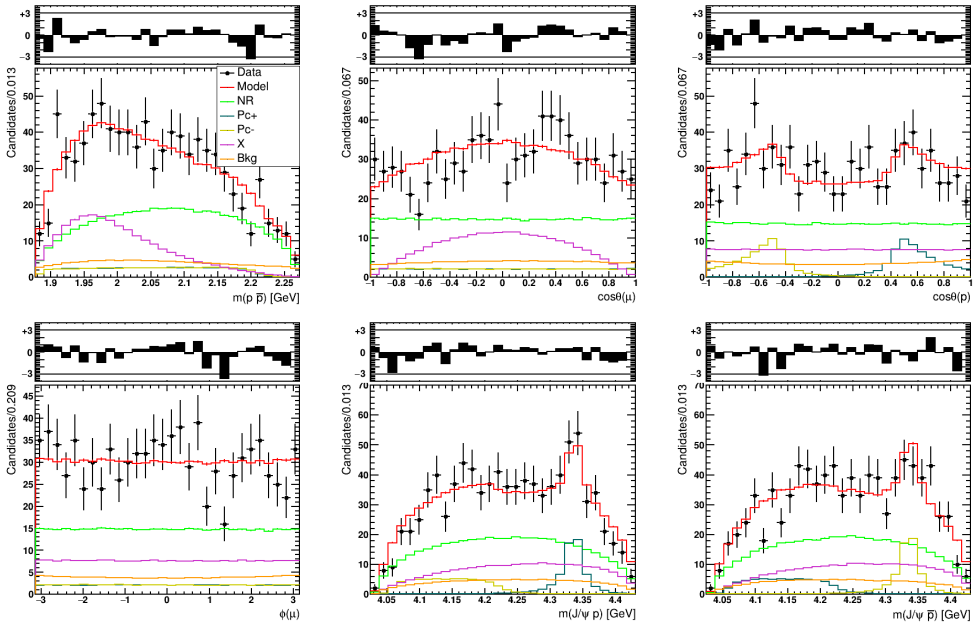


Figure A.17 One dimensional distributions of kinematic variables of phase-space with P_c states and $J/\psi\, p$, $J/\psi\, \overline{p}$ invariant mass projections.

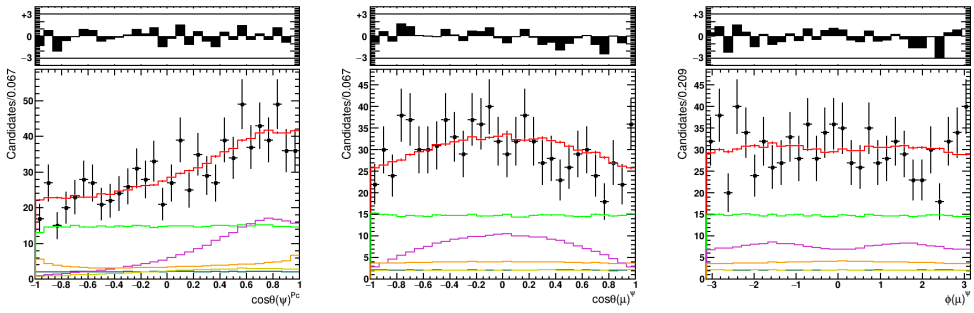


Figure A.18 Angular distributions in the P_c decay chain, as described in sec. 5.1.4.

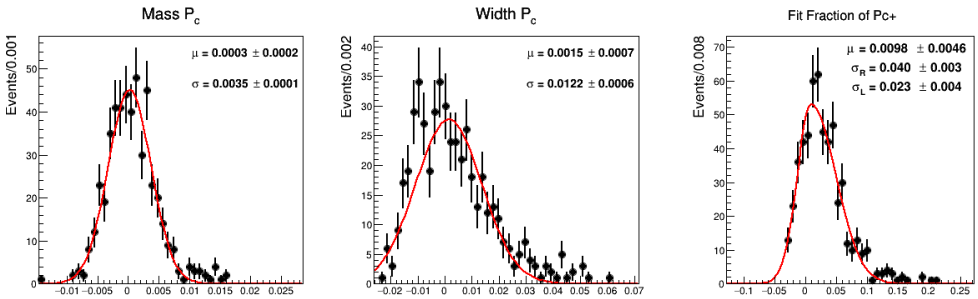


Figure A.19 Residuals for mass, width and fit fraction of the P_c states for the alternative model with all LS values for the P_c couplings allowed to vary, as described in Sec. A.4.6. No systematic effect is assigned because it is negligible contribution with respect to other sources.

Bibliography

[1] LHCb collaboration, R. Aaij *et al.*, *Observation of $B^0_{(s)} \to J/\psi p\overline{p}$ decays and precision measurements of the $B^0_{(s)}$ masses*, Phys. Rev. Lett. **122** (2019) 191804, arXiv:1902.05588.

[2] Y. K. Hsiao and C. Q. Geng, *$f_J(2220)$ and hadronic $\bar{B}^0_s$ decays*, Eur. Phys. J. **C75** (2015), no. 3 101, arXiv:1412.4900.

[3] S. L. Glashow, *Partial-symmetries of weak interactions*, Nuclear Physics **22** (1961), no. 4 579 .

[4] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264.

[5] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. C **680519** (1968) 367.

[6] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **13** (1964) 508.

[7] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa matrix*, Phys. Rev. Lett. **51** (1983) 1945.

[8] A. Höcker, H. Lacker, S. Laplace *et al.*, *A new approach to a global fit of the CKM matrix*, Eur. Phys. J. **C 21** (2001) 225–259, and updates at http://ckmfitter.in2p3.fr/.

[9] P. D. Group *et al.*, *Review of Particle Physics*, Progress of Theoretical and Experimental Physics **2020** (2020) , 083C01.

[10] I. e. a. Dunietz, *In Pursuit of New Physics with B_s Decays*, JINST (2016) 11 C11040 .

[11] M. Gell-Mann, *A schematic model of baryons and mesons*, Phys. Rev. Lett. **8** (1964) 214.

[12] G. Zweig, *An SU(3) model for strong interaction symmetry and its breaking*, 1964.

[13] D. J. Gross and F. Wilczek, *Ultraviolet behavior of non-abelian gauge theories*, Phys. Rev. Lett. **30** (1973) 1343.

[14] S. L. Olsen, T. Skwarnicki, and D. Zieminska, *Nonstandard heavy mesons and baryons: Experimental evidence*, Rev. Mod. Phys. **90** (2018) 015003.

[15] N. N. Achasov, A. V. Kiselev, and G. N. Shestakov, *Theory of scalars*, Nuclear Physics B - Proceedings Supplements **181-182** (2008) 169, Proceedings of the International Workshop on e^+e^- Collisions from to .

[16] JPAC, A. Rodas *et al.*, *Determination of the pole position of the lightest hybrid meson candidate*, Phys. Rev. Lett. **122** (2019), no. 4 042002, arXiv:1810.04171.

[17] E. Eichten *et al.*, *Charmonium: Comparison with experiment*, Phys. Rev. D **21** (1980) 203.

[18] Belle Collaboration, S.-K. e. a. Choi, *Observation of a narrow charmoniumlike state in exclusive $B^{\pm} \to K^{\pm}\pi^+\pi^- J/\psi$ decays*, Phys. Rev. Lett. **91** (2003) 262001.

[19] Y. Dong, A. Faessler, T. Gutsche, and V. E. Lyubovitskij, *$J/\psi\gamma$ and $\psi(2S)\gamma$ decay modes of the $X(3872)$*, Journal of Physics G: Nuclear and Particle Physics **38** (2010) 015001.

[20] LHCb collaboration, R. Aaij *et al.*, *Observation of the resonant character of the $Z(4430)^-$ state*, Phys. Rev. Lett. **112** (2014) 222002, arXiv:1404.1903.

[21] LHCb collaboration, R. Aaij *et al.*, *Observation of $J/\psi p$ resonances consistent with pentaquark states in $\Lambda_b^0 \to J/\psi p K^-$ decays*, Phys. Rev. Lett. **115** (2015) 072001, arXiv:1507.03414.

[22] A. Esposito, A. Pilloni, and A. D. Polosa, *Multiquark resonances*, Physics Reports **668** (2017) 1 .

[23] T. Blum *et al.*, *Report of the Snowmass 2013 Computing Frontier working group on Lattice Field Theory – lattice field theory for the energy and intensity frontiers: Scientific goals and computing needs*, arXiv:1310.6087.

[24] K. Orginos and D. Richards, *Improved methods for the study of hadronic physics from lattice QCD*, Journal of Physics G: Nuclear and Particle Physics **42** (2015) 034011.

[25] Hadron Spectrum, G. K. C. Cheung *et al.*, *DK I = 0,$D\overline{K}$ I = 0, 1 scattering and the $D_{s0}^*(2317)$ from lattice QCD*, JHEP **02** (2021) 100, arXiv:2008.06432.

[26] L. Gayer *et al.*, *Isospin-1/2 $D\pi$ scattering and the lightest D_0^* resonance from lattice QCD*, arXiv:2102.04973.

[27] T. Nakano *et al.*, *Evidence for a narrow $s = +1$ baryon resonance in photoproduction from the neutron*, Phys. Rev. Lett. **91** (2003) 012002.

[28] CLAS Collaboration, S. Stepanyan, K. Hicks, and D. S. Carman, *Observation of an Exotic $S = +1$ Baryon in Exclusive Photoproduction from the Deuteron*, Phys. Rev. Lett. **91** (2003) 252001.

[29] CLAS Collaboration, B. McKinnon *et al.*, *Search for the Θ^+ pentaquark in the reaction $\gamma d \to p K^- K^+ n$*, Phys. Rev. Lett. **96** (2006) 212001, arXiv:hep-ex/0603028.

[30] Particle Data Group, K. A. Olive *et al.*, *Review of particle physics*, Chin. Phys. **C38** (2014) 090001.

[31] LHCb collaboration, R. Aaij *et al.*, *Model-independent evidence for $J/\psi p$ contributions to $\Lambda_b^0 \to J/\psi p K^-$ decays*, Phys. Rev. Lett. **117** (2016) 082002, arXiv:1604.05708.

[32] LHCb collaboration, R. Aaij *et al.*, *Observation of a narrow $P_c(4312)^+$ state, and of two-peak structure of the $P_c(4450)^+$*, Phys. Rev. Lett. **122** (2019) 222001, arXiv:1904.03947.

[33] Belle collaboration, K. Abe *et al.*, *Observation of $\overline{B}^0 \to D^{(*)0} p\overline{p}$*, Phys. Rev. Lett. **89** (2002) 151802, arXiv:hep-ex/0205083.

[34] Belle, M. Z. Wang *et al.*, *Observation of $B^0 \to p\overline{\Lambda}\pi^-$*, Phys. Rev. Lett. **90** (2003) 201802, arXiv:hep-ex/0302024.

[35] CLEO, X. Fu *et al.*, *Observation of exclusive B decays to final states containing a charmed baryon*, Phys. Rev. Lett. **79** (1997) 3125.

[36] LHCb, R. Aaij *et al.*, *Evidence for CP Violation in $B^+ \to p\bar{p}K^+$ Decays*, Phys. Rev. Lett. **113** (2014), no. 14 141801, arXiv:1407.5907.

[37] LHCb collaboration, *First observation of a baryonic B_s^0 decay*, LHCb-CONF-2016-016. Retired. Please do not cite it.

[38] LHCb, R. Aaij *et al.*, *First observation of a baryonic B_c^+ decay*, Phys. Rev. Lett. **113** (2014), no. 15 152003, arXiv:1408.0971.

[39] BaBar collaboration, P. del Amo Sanchez *et al.*, *Observation and study of the baryonic B-meson decays $B \to D^{(*)} p\overline{p}(\pi)(\pi)$*, Phys. Rev. **D85** (2012) 092017, arXiv:1111.4387.

[40] Belle collaboration, J.-T. Wei *et al.*, *Study of $B^+ \to p\bar{p}K^+$ and $B^+ \to p\bar{p}\pi^+$*, Phys. Lett. **B659** (2008) 80, arXiv:0706.4167.

[41] LHCb collaboration, R. Aaij *et al.*, *Measurements of the branching fractions of $B^+ \to p\bar{p}K^+$ decays*, Eur. Phys. J. **C73** (2013) 2462, arXiv:1303.7133.

[42] BaBar collaboration, B. Aubert *et al.*, *Evidence for the $B^0 \to p\bar{p}K^{*0}$ and $B^+ \to \eta_c K^{*+}$ decays and study of the decay dynamics of B meson decays into $p\bar{p}h$ final states*, Phys. Rev. **D76** (2007) 092004, arXiv:0707.1648.

[43] Belle collaboration, J.-H. Chen *et al.*, *Observation of $B^0 \to p\bar{p}K^{*0}$ with a large K^{*0} polarization*, Phys. Rev. Lett. **100** (2008) 251801, arXiv:0802.0336.

[44] C. Q. Geng and Y. K. Hsiao, *Angular distributions in three-body baryonic B decays*, Phys. Rev. **D74** (2006) 094023, arXiv:hep-ph/0606141.

[45] BES Collaboration, J. Z. Bai, Y. Ban, Bian *et al.*, *Observation of a near-threshold enhancement in the $p\overline{p}$ mass spectrum from radiative $j/\psi \to \gamma p\overline{p}$ decays*, Phys. Rev. Lett. **91** (2003) 022001.

[46] BES Collaboration, M. Ablikim and J. Z. Bai, *Observation of a threshold enhancement in the $p\overline{\Lambda}$ invariant-mass spectrum*, Phys. Rev. Lett. **93** (2004) 112002.

[47] Belle Collaboration, Q. L. Xie and K. Abe, *Observation of $B^- \to J/\psi\Lambda\overline{p}$ and searches for $B^- \to J/\psi\Sigma^0\overline{p}$ and $B^0 \to J/\psi p\overline{p}$ decays*, Phys. Rev. D **72** (2005) 051105.

[48] M. Suzuki, *Partial waves of baryon-antibaryon in three-body B meson decay*, J. Phys. G **34** (2006) 283.

[49] C.-K. Chua, W.-S. Hou, and S.-Y. Tsai, *Charmless three-body baryonic B decays*, Phys. Rev. D **66** (2002) 054004.

[50] G. W. S. Hou, *Possible hints and search for glueballs in charmless rare B decays*, Nuclear Physics B - Proceedings Supplements **117** (2003) 681 .

[51] S. Okubo, *ϕ-meson and unitary symmetry model*, Physics Letters **5** (1963), no. 2 165.

[52] J. Iizuka, *A Systematics and Phenomenology of Meson Family*, Progress of Theoretical Physics Supplement **37-38** (1966) 21.

[53] LHCb, R. Aaij *et al.*, *Searches for $B^0_{(s)} \to J/\psi p\bar{p}$ and $B^+ \to J/\psi p\bar{p}\pi^+$ decays*, JHEP **09** (2013) 006, arXiv:1306.4489.

[54] C. Q. G. Chun-Hung Chen, Hai-Yang Cheng and Y. K. Hsiao, *Charmful three-body baryonic B decays*, Phys. Rev. **D78** (2008) 054016.

[55] C. J. Morningstar and M. Peardon, *Glueball spectrum from an anisotropic lattice study*, Phys. Rev. D **60** (1999) 034509.

[56] LHCb, R. Aaij *et al.*, *Analysis of the resonant components in $B_s \to J/\psi\pi^+\pi^-$*, Phys. Rev. **D86** (2012) 052006, arXiv:1204.5643.

[57] V. Crede and C. A. Meyer, *The experimental status of glueballs*, Progress in Particle and Nuclear Physics **63** (2009), no. 1 74 .

[58] R. M. Baltrusaitis *et al.*, *Observation of a narrow KK^- state in J/ψ radiative decays*, Phys. Rev. Lett. **56** (1986) 107.

[59] BES Collaboration, J. Z. Bai *et al.*, *Studies of ξ(2230) in J/ ψ radiative decays*, Phys. Rev. Lett. **76** (1996) 3502.

[60] JETSET, C. Evangelista *et al.*, *Study of the reaction $\overline{p}p \to \phi\phi$ from 1.10 GeV/c to 2.0 GeV/c*, Phys. Rev. D **57** (1998) 5370, arXiv:hep-ex/9802016.

[61] L3, M. Acciarri *et al.*, *$K^0_S K^0_S$ final state in two photon collisions and implications for glueballs*, Phys. Lett. B **501** (2001) 173, arXiv:hep-ex/0011037.

[62] O. S. Brüning *et al.*, *LHC Design Report*, CERN Yellow Reports: Monographs, CERN, Geneva, 2004. doi: 10.5170/CERN-2004-003-V-1.

[63] T. L. Collaboration *et al.*, *The LHCb detector at the LHC*, Journal of Instrumentation **3** (2008) S08005.

[64] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, arXiv:1412.6352.

[65] LHCb collaboration, R. Aaij *et al.*, *Measurement of the b-quark production cross-section in 7 and 13 TeVpp collisions*, Phys. Rev. Lett. **118** (2017) 052002, Erratum ibid. **119** (2017) 169901, arXiv:1612.05140.

[66] Y. A. e. a. Heavy Flavor Averaging Group (HFAG), *Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2016*, arXiv:1612.07233.

[67] A. A. Alves Jr. *et al.*, *Performance of the LHCb muon system*, JINST **8** (2013) P02022, arXiv:1211.1346.

[68] LHCb collaboration, *LHCb VELO (VErtex LOcator): Technical Design Report*, CERN-LHCC-2001-011.

[69] R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, JINST **9** (2014) P09007, arXiv:1405.7808.

[70] LHCb collaboration, *LHCb inner tracker: Technical Design Report*, CERN-LHCC-2002-029.

[71] LHCb collaboration, *LHCb outer tracker: Technical Design Report*, CERN-LHCC-2001-024.

[72] LHCb collaboration, *LHCb online system, data acquisition and experiment control: Technical Design Report*, CERN-LHCC-2001-040.

[73] LHCb collaboration, *LHCb magnet: Technical Design Report*, CERN-LHCC-2000-007.

[74] R. Frühwirth, *Application of kalman filtering to track and vertex fitting*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **262** (1987), no. 2 444 .

[75] LHCb collaboration, *LHCb RICH: Technical Design Report*, CERN-LHCC-2000-037.

[76] M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. **C73** (2013) 2431, arXiv:1211.6759.

[77] LHCb collaboration, *LHCb calorimeters: Technical Design Report*, CERN-LHCC-2000-036.

[78] LHCb collaboration, *LHCb muon system: Technical Design Report*, CERN-LHCC-2001-010.

[79] F. Archilli *et al.*, *Performance of the muon identification at LHCb*, JINST **8** (2013) P10020, arXiv:1306.0249.

[80] R. Aaij *et al.*, *The LHCb trigger and its performance in 2011*, JINST **8** (2013) P04022, arXiv:1211.3055.

[81] R. Aaij *et al.*, *Performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC*, JINST **14** (2019), no. LHCb-DP-2019-001 P04013, arXiv:1812.10790.

[82] G. Barrand *et al.*, *GAUDI - A software architecture and framework for building HEP data processing applications*, Comput. Phys. Commun. **140** (2001) 45.

[83] LHCb Collaboration, *The Moore project*, arXiv:http://lhcbdoc.web.cern.ch/lhcbdoc/moore/.

[84] LHCb Collaboration, *The Brunel project*, arXiv:http://lhcbdoc.web.cern.ch/lhcbdoc/brunel/.

[85] LHCb Collaboration, *The DaVinci project*, arXiv:http://lhcbdoc.web.cern.ch/lhcbdoc/davinci/.

[86] LHCb Collaboration, *The Gauss project*, arXiv:http://lhcbdoc.web.cern.ch/lhcbdoc/gauss/.

[87] T. Sjöstrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, JHEP **05** (2006) 026, arXiv:hep-ph/0603175.

[88] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.

[89] Geant4 collaboration, J. Allison *et al.*, *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. **53** (2006) 270.

[90] LHCb Collaboration, *The Boole project*, arXiv:http://lhcbdoc.web.cern.ch/lhcbdoc/boole/.

[91] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $B^0_s \to \mu^+\mu^-$ branching fraction and effective lifetime and search for $B^0 \to \mu^+\mu^-$ decays*, Phys. Rev. Lett. **118** (2017) 191801, arXiv:1703.05747.

[92] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP-averaged observables in the $B^0 \to K^{*0}\mu^+\mu^-$ decay*, Phys. Rev. Lett. **125** (2020) 011802, arXiv:2003.04831.

[93] LHCb collaboration, R. Aaij *et al.*, *Observation of CP violation in two-body B^0_s-meson decays to charged pions and kaons*, arXiv:2012.05319, submitted to JHEP.

[94] LHCb collaboration, R. Aaij *et al.* , and A. Bharucha *et al.*, *Implications of LHCb measurements and future prospects*, Eur. Phys. J. **C73** (2013) 2373, arXiv:1208.3355.

[95] LHCb collaboration, *Framework TDR for the LHCb Upgrade: Technical Design Report*, CERN-LHCC-2012-007.

[96] LHCb collaboration, *LHCb Trigger and Online Technical Design Report*, CERN-LHCC-2014-016.

[97] LHCb collaboration, *LHCb VELO Upgrade Technical Design Report*, CERN-LHCC-2013-021.

[98] LHCb collaboration, *LHCb Tracker Upgrade Technical Design Report*, CERN-LHCC-2014-001.

[99] LHCb collaboration, *LHCb PID Upgrade Technical Design Report*, CERN-LHCC-2013-022.

[100] LHCb collaboration, *Expression of Interest for a Phase-II LHCb Upgrade: Opportunities in flavour physics, and beyond, in the HL-LHC era*, CERN-LHCC-2017-003.

[101] I. Bejar Alonso, L. Rossi *et al.*, *HiLumi LHC Technical Design Report: Deliverable: D1.10*, Tech. Rep. CERN-ACC-2015-0140, Nov, 2015.

[102] Belle Collaboration, T. Abe *et al.*, *Belle II Technical Design Report*, arXiv:1011.0352.

[103] M. Jacob and G. C. Wick, *On the general theory of collisions for particles with spin*, Annals of Physics **7** (1959), no. 4 404 .

[104] M. J. Moravcsik, *Elementary particle theory. a. d. martin and t. d. spearman.*, Science **170** (1970), no. 3964 1295.

[105] C. Zemach, *Use of angular-momentum tensors*, Phys. Rev. **140** (1965) B97.

[106] S. U. Chung, *Helicity-coupling amplitudes in tensor formalism*, Phys. Rev. D **48** (1993) 1225.

[107] S.-U. Chung and J. M. Friedrich, *Covariant helicity-coupling amplitudes: A new formulation*, Phys. Rev. D **78** (2008) 074027.

[108] J. D. Richman, *An experimenter's guide to the helicity formalism*, Tech. Rep. CALT-68-1148, Calif. Inst. Technol., Pasadena, CA, Jun, 1984.

[109] LHCb collaboration, R. Aaij *et al.*, *Search for the $\Lambda_b^0 \to \Lambda\eta$ and $\Lambda_b^0 \to \Lambda\eta'$ decays with the LHCb detector*, JHEP **09** (2015) 006, arXiv:1505.03295.

[110] E. Leader, *Spin in particle physics*, vol. 15, 2, 2011.

[111] JPAC Collaboration, M. Mikhasenko *et al.*, *What is the right formalism to search for resonances?*, arXiv:1712.02815.

[112] M. Mikhasenko *et al.*, *Dalitz-plot decomposition for three-body decays*, Phys. Rev. D **101** (2020) 034033.

[113] W. D. Hulsbergen, *Decay chain fitting with a kalman filter*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **552** (2005), no. 3 566 .

[114] Particle Data Group, C. Patrignani *et al.*, *Review of Particle Physics*, Chin. Phys. **C40** (2016), no. 10 100001.

[115] G. Passaleva, *A recurrent neural network for track reconstruction in the LHCb Muon System*, Nuclear Science Symposium Conference Record, 2008. NSS '08. IEEE .

[116] V. V. Gligorov and M. Williams, *Efficient, reliable and fast high-level triggering using a bonsai boosted decision tree*, Journal of Instrumentation **8** (2013) P02013.

[117] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances*, PhD thesis, Institute of Nuclear Physics, Krakow, 1986, DESY-F31-86-02.

[118] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, arXiv:physics/0402083.

[119] LHCb collaboration, R. Aaij *et al.*, *Observation of exotic $J/\psi\phi$ structures from amplitude analysis of $B^+ \to J/\psi\phi K^+$ decays*, Phys. Rev. Lett. **118** (2016) 022003, arXiv:1606.07895.

[120] A. Rogozhnikov, *Reweighting with Boosted Decision Trees*, J. Phys. Conf. Ser. **762** (2016), no. 1 012036, arXiv:1608.05806.

[121] LHCb collaboration, R. Aaij *et al.*, *Searches for $B^0_{(s)} \to J/\psi p\bar{p}$ and $B^+ \to J/\psi p\bar{p}\pi^+$ decays*, JHEP **09** (2013) 006, arXiv:1306.4489.

[122] T. Sjöstrand *et al.*, *An introduction to pythia 8.2*, Computer Physics Communications **191** (2015) 159–177.

[123] A. Poluektov, *Kernel density estimation of a multidimensional efficiency profile*, JINST **10** (2015), no. 02 P02011, arXiv:1411.5528.

[124] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees*, Wadsworth international group, Belmont, California, USA, 1984.

[125] P. Speckmayer, A. Hocker, J. Stelzer *et al.*, *The toolkit for multivariate data analysis, TMVA 4*, J. Phys. Conf. Ser. **219** (2010) 032057.

[126] S. S. Wilks, *The large-sample distribution of the likelihood ratio for testing composite hypotheses*, Annals Math. Statist. **9** (1938) 60.

[127] LHCb collaboration, R. Aaij *et al.*, *Amplitude analysis and branching fraction measurement of $\overline{B}^0_s \to J/\psi K^+K^-$*, Phys. Rev. **D87** (2013) 072004, arXiv:1302.1213.

[128] LHCb collaboration, *Updated average f_s/f_d b-hadron production fraction ratio for 7 TeV pp collisions*, LHCb-CONF-2013-011.

[129] LHCb collaboration, R. Aaij *et al.*, *Measurement of relative branching fractions of B decays to $\psi(2S)$ and J/ψ mesons*, Eur. Phys. J. **C72** (2012) 2118, arXiv:1205.0918.

[130] R. A. et al. *Measurement of the track reconstruction efficiency at lhcb*, JINST **10** (2015) P02007, arXiv:1408.1251.

[131] J. B. et al. *Search for the Rare Two-body Charmless Baryonic decay $B^+ \to p\bar{\Lambda}$*, LHCb-ANA-2014-094.

[132] LHCb, R. Aaij *et al.*, *Measurement of b-hadron masses*, Phys. Lett. **B708** (2012) 241, arXiv:1112.4896.

[133] M. D. Needham, *Momentum scale calibration of the LHCb spectrometer*, Tech. Rep. LHCb-INT-2017-008. CERN-LHCb-INT-2017-008, CERN, Geneva, Apr, 2017.

[134] D. Marangotto, *Helicity amplitudes for generic multi-body particle decays featuring multiple decay chains*, arXiv:1911.10025.

[135] N. Wu, *Centrifugal—barrier effects and determination of interaction radius*, Communications in Theoretical Physics **61** (2014) 89–94.

[136] F. James, *Statistical Methods in Experimental Physics*, WORLD SCIENTIFIC, 2nd ed., 2006. doi: 10.1142/6096.

[137] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, The European Physical Journal C **71** (2011) .

[138] J. Neyman and E. S. Pearson, *On the Problem of the Most Efficient Tests of Statistical Hypotheses*, Phil. Trans. Roy. Soc. Lond. A **231** (1933), no. 694-706 289.

[139] A. L. Read, *Presentation of search results: the CL_s technique*, Journal of Physics G: Nuclear and Particle Physics **28** (2002) 2693.

[140] E. Bowen and B. Storaci, *VeloUT tracking for the LHCb Upgrade*, LHCb-PUB-2013-023. CERN-LHCb-PUB-2013-023. LHCb-INT-2013-056.

[141] S. Bugiel *et al.*, *SALT, a dedicated readout chip for high precision tracking silicon strip detectors at the LHCb Upgrade*, Journal of Instrumentation **11** (2016) C02028.

[142] K. Wyllie *et al.*, *Electronics Architecture of the LHCb Upgrade*, Tech. Rep. LHCb-PUB-2011-011. CERN-LHCb-PUB-2011-011., CERN, Geneva, 2011.

[143] F. Alessio and R. Jacobsson, *Readout Control Specifications for the Front-End and Back-End of the LHCb Upgrade*, Tech. Rep. LHCb-PUB-2012-017. CERN-LHCb-PUB-2012-017. LHCb-INT-2012-018, CERN, Geneva, Feb, 2014.

[144] L. Cardoso *et al.*, *LHCb MiniDAQ control system*, EPJ Web of Conferences **214** (2019) 01005.

[145] R. M. Lesma *et al.*, *The Versatile Link Demo Board (VLDB)*, Journal of Instrumentation **12** (2017) C02020.

[146] P. Moreira *et al.*, *The GBT-SerDes ASIC prototype*, Journal of Instrumentation **5** (2010) C11022.

[147] A. Caratelli *et al.*, *The GBT-SCA, a radiation tolerant ASIC for detector control and monitoring applications in HEP experiments*, Journal of Instrumentation **10** (2015) C03034.

[148] C. Soós *et al.*, *The Versatile Transceiver: towards production readiness*, Journal of Instrumentation **8** (2013) C03004.

Acknowledgments

At the end of these intense years as a PhD student, there are many people I would like to thank. Firstly, I would like to thank my supervisor, Prof. Nicola Neri, for his constant support and guidance during this work. He taught me how to conduct research and through his precious advice made my work richer and fascinating.

Besides my supervisor, I would like to thank Dr. Alessandro Pilloni, who supervised and helped me to develop the amplitude model from a theoretical point of view. I am grateful for his patience and constructive discussions that helped me deepen my knowledge and achieve the results presented in this thesis.

I would like to express my sincere gratitude to the Fondazione Fratelli Confalonieri for funding my PhD research during these three years, for their support and generosity.

Special thanks to the thesis referees, Dr. Mikhail Mikhashenko and Prof. Adam Szczepaniak, for their careful revision.

I would like to extend my acknowledgments to my labmates, starting with the LHCb group in Milan: Paolo, Andrea, Daniele, Salvatore and Jinlin for the daily discussions and for sharing their expertise and knowledge. Also, I would like to thank the UT-Milan group for the time we spent in the lab working on the SALT chip and for the resources they provided me, especially Antonio, Nadim, Alessandro, Federico and Marco.

I want to share this achievement with all the friends who have accompanied me in these years, especially those who shared the life at the university during the PhD: Marta, Benny, Alessia, Elia and Nicolas. Thanks also to the friends who always accompany me: Nicol and Simone, Francesca and Davide with little Marti, Samuele and Chiara with little Filippo Leo, Francesco and all the friends I met at university. A special acknowledgment also to the friends we met in Como in the last years, because they made us feel at home.

Finally, I would like to thank my family for their support during the studies and for always believing in me even when it was difficult for me to do so. A special thanks to my husband, Luigi, because his lovely presence makes it easier to face everything.