Dissertation

submitted to the
Combined Faculty of the Natural Sciences and Mathematics
of Heidelberg University, Germany
for the degree of
Doctor of Natural Sciences

Put forward by

Jascha Grabowski

born in: Coburg

Oral examination: 13.12.2021

Amplitude Analysis of the Decay $B_s^0 \to \psi(2S)\pi^+\pi^-$ with the LHCb Experiment

Referees: Prof. Dr. Ulrich Uwer
Prof. Dr. Klaus Reygers

"The best that most of us can hope to achieve in physics is simply to misunderstand at a deeper level."

Wolfgang Pauli

Abstract

Abstract This thesis presents the first analysis of the hadronic amplitudes in the decay $B_s^0 \to \psi(2S)\pi^+\pi^-$. The pions originate from the $s\bar{s}$ quark system of the weak B_s^0 decay, which removes contributions from Isospin 1 resonances. This allows to test a novel dispersive approach for a process-independent description of the dipion isoscalar S-wave. Additionally, the tetraquark candidate $Z(4430)^-$ is known to contribute to $\psi(2S)\pi^-$ in $B^0 \to \psi(2S)K^+\pi^-$. An observation of this state in the $\psi(2S)\pi^\pm$ systems in $B_s^0 \to \psi(2S)\pi^+\pi^-$ would give insights into its production mechanism and internal structure. The search for this state among the intermediate hadronic states of the $B_s^0 \to \psi(2S)\pi^+\pi^-$ decay is thus also part of the thesis.

Using proton-proton collision data collected by the LHCb experiment in the years 2011, 2012, 2015-2018 at 7, 8, and 13 TeV corresponding to integrated luminosities of 1, 2, and 5 fb^{-1}, the fit fractions of S-wave and the individual f_2 states are determined and the process-independence of the S-wave parametrisation is shown. Furthermore, a search for the exotic tetraquark candidate $Z(4430)^\pm$ in $B_s^0 \to Z(4430)^\pm\pi^\mp$ with $Z(4430)^\pm \to \psi(2S)\pi^\pm$ is performed. No significant exotic contribution is measured and an upper limit on the fraction of decays proceeding via $Z(4430)^\pm$ is found to be $\mathcal{F}_Z < 1.0\,\%$ at 95 % confidence level.

Zusammenfassung Diese Dissertation stellt die erste Analyse der hadronischen Amplituden im Zerfall von $B_s^0 \to \psi(2S)\pi^+\pi^-$ dar. Die Pionen entstammen einem $s\bar{s}$ Zustand aus dem schwachen Zerfall von $B_s^0 \to \psi(2S)\pi^+\pi^-$, wodurch Beiträge von Isospin 1 Resonanzen verhindert werden. Dies ermöglicht das Testen eines neuartigen dispersiven Ansatzes für die prozessunabhängige Beschreibung der isoskalaren Pion-Pion S-Welle. Außerdem ist bekannt, dass der Tetraquark Kandidat $Z(4430)^-$ zur $\psi(2S)\pi^-$ Verteilung von $B^0 \to \psi(2S)K^+\pi^-$ beiträgt. Ein Nachweis dieses Zustandes in den $\psi(2S)\pi^\pm$ Verteilungen von $B_s^0 \to \psi(2S)\pi^+\pi^-$ würde Rückschlüsse auf seinen Produktionsmechanismus und Aufbau erlauben. Die Suche nach diesem exotischen Zustand innerhalb der hadronischen Zwischenzustände im $B_s^0 \to \psi(2S)\pi^+\pi^-$ Zerfall ist deshalb auch Teil dieser Dissertation.

Es werden Daten aus Proton-Proton Kollisionen verwendet, die in den Jahren 2011, 2012 und 2015 bis 2018 am LHCb Detektor aufgezeichnet wurden. Sie entsprechen einer integrierten Luminosität von 1, 2 und 5 fb^{-1} bei Kollisionsenergien von 7, 8 und 13 TeV. Die Anteile der S-Welle und der einzelnen f_2 Resonanzen werden bestimmt und die Prozessunabhängigkeit der S-Wellenparametrisierung überprüft. Desweiteren wird eine Suche nach dem exotischen Tetraquark Kandidaten $Z(4430)^\pm$ in $B_s^0 \to Z(4430)^\pm\pi^\mp$ mit $Z(4430)^\pm \to \psi(2S)\pi^\pm$ durchgeführt. Es werden keine signifikanten exotischen Beiträge beobachtet und das 95 % Konfidenzintervall für den Anteil der Zerfälle über $Z(4430)^\pm$ zu $< 1.0\,\%$ bestimmt.

Preface

The measurement presented in this work is not yet published in a peer-reviewed journal. It is currently undergoing the internal review process of the LHCb collaboration, before it can be submitted to a journal.

The author of this work is the main person responsible for the analysis work presented here and all presented analysis steps were performed by him. The measurement has been carried out within the LHCb collaboration, which consists of more than 1400 scientists and engineers from various countries. The data used in this work was collected at the LHCb detector at the LHC and reconstructed and analysed using software tools provided by the collaboration. Both detector and software tools are the product of the work of many current and former collaboration members.

The description of the dipion spectrum using a dispersive approach developed in [1] was carried out in collaboration with the theorists Christoph Hanhart and Bastian Kubis, who were affiliated to the LHCb collaboration in order to share data and analysis software.

Parts of the limit-setting procedure described in Sec. 10 were developed by an intern under the supervision of the author.

A more technical description of the analysis is available as internal LHCb note in [2].

Besides the analysis presented in this work, the author was also the main contributor to another publication, which resulted from work during his master project and the beginning of his PhD. This measurement of the branching fraction and *CP* asymmetry in $B^+ \to J/\psi \rho^+$ decays was published in Ref. [3].

Contents

1 Introduction

One of the major triumphs of modern physics is the unification of three of the four known fundamental forces of nature into the so-called Standard Model (SM) of Particle Physics during the 1960s.[1] Since then the model withstood essentially all tests of its predictions.

But since its conception it is also clear that the SM needs to be extended: Not only does it not account for the gravitational force, it is not able to describe the observed baryon asymmetry in the universe [6, 7] or those 95 % of its energy content, which are attributed to dark matter and dark energy. Another shortcoming is its lack of justifications for *e.g.* the number of fundamental fermion generations or the fact that no *CP* violation is observed in the strong interaction. Over the decades different expansions have sprung up, but new experimental results have driven them since into less favoured regions of parameter space, diminishing their popularity again.

In recent years apparent experimental discrepancies in the lepton sector have been observed, *e.g.* in the value of $g-2$ of muons [8] or the lepton universality parameter R_K [9] with increasing significances too, but it remains to be seen, whether they prove to be the guiding light to a unified model of fundamental particle physics.

But even within the region of its validity, some processes are very hard to access both from the experimental and the theoretical side: The 5 % of hadronic matter in the universe consist of quarks bound together by gluons, both of which carry the colour charge of strong interaction. Due to confinement, no coloured objects can exist freely though, therefore measurements of the strong interaction (at low energies) cannot probe the individual constituents of hadrons directly, but rather measure the properties of bound states and their interactions. From the theory perspective the non-Abelian structure of the theory does not allow a perturbative approach to describe the interactions of quarks and gluons either.

Despite these difficulties, the theory of strong interaction has come far in describing the spectrum of the bound states of hadrons and their interactions.

[1] With the later addition of neutrino masses after the observation of neutrino oscillations [4, 5].

But much is still to be learned about *e.g.* the formation process of these states and their internal structure. Open questions include the nature of the recently observed pentaquark candidates [10, 11], tetraquark candidates with two bottom [12], two charm [13–15] or even four charm quarks [16, 17] and the existence of the predicted glueball and hybrid states consisting solely or partly of excited gluons.

The description of the strong interaction is not only an interesting subject in itself though, it also has consequences on other fields as well. For example the precision achieved for the theoretical prediction in tests of other sectors of the SM like the electroweak sector are often limited by the current knowledge of the hadronic part of the interactions. The aforementioned value of $g-2$ of the muon is an example. The hadronic vacuum polarisation and hadronic light-by-light scattering are the largest sources of uncertainty for the standard model prediction [18] of $g-2$.

In close cooperation between experiment and theory techniques have been refined to identify observables that are accessible from both theoretical and experimental side. This work tests one of these techniques, which uses a dispersive formalism to describe the strong interaction between two pions in the low-energy region, as detailed in Sec. 2 (from here on referred to as Bonn-Jülich model). It is better motivated theoretically than other parametrisations conventionally used in experiments and should have the advantage that most of its parameters are process independent.

This independence is tested in this thesis on the dipion spectrum of $B_s^0 \to \psi(2S)\pi^+\pi^-$[2] decays against a less sophisticated, but commonly used, model. The decays were measured with the LHCb experiment in the years 2011-2018 at the LHC collider at CERN, which is introduced in Sec. 3. Since the data is collected inclusively on the signature of $\psi(2S) \to \mu^+\mu^-$, it also contains a large fraction of background and the signal events first need to be selected from it, as described in Sec. 4.

Another goal of this thesis is the search for exotic contributions from the tetraquark candidate $Z(4430)^\pm$ to the spectra of $\psi(2S)\pi^+$ and $\psi(2S)\pi^-$ in $B_s^0 \to \psi(2S)\pi^+\pi^-$. Their occurrence in this decay would contradict a popular model describing these structures as triangle singularities and favour true resonant descriptions via tetraquark or molecular compositions.

[2] The charge conjugated process is implied as well throughout this thesis, unless stated otherwise.

Section 5 introduces the helicity amplitude model, which is used to describe the full decay chain. This also includes the descriptions of the different parametrisations of the dipion spectrum and the treatment of possible $Z(4430)^{\pm}$ contribution.

The practical application of the helicity amplitude framework in an extended unbinned maximum likelihood amplitude fit to the data is shown in Sec. 6, while the test of the process independence of the Bonn-Jülich model is given in Sec. 7. The corresponding statistical and systematic uncertainties are addressed in Sec. 8 and 9.

The search for decays via exotic intermediate states $B_s^0 \to Z(4430)^- \pi^+$ and $B_s^0 \to Z(4430)^+ \pi^-$ in $B_s^0 \to \psi(2S)\pi^+\pi^-$ is documented in Sec. 10 and concluding remarks are given in Sec. 11.

2 Theoretical Background

After a short introduction into the structure of the SM in Sec. 2.1 and a sketch of the relevant phenomena of weak B-meson decays in Sec. 2.2, some aspects of the strong interaction are given in Sec. 2.3. In Sec. 2.5 the signal decay channel $B_s^0 \to \psi(2S)\pi^+\pi^-$ is introduced. The hadronic phenomena, which can be studied in this channel, are then detailed in the following sections: Exotic hadronic states are described in Sec. 2.6 with special attention to the tetraquark candidate $Z(4430)^\pm$ that could occur as intermediate state in the $\psi(2S)\pi^+$ or $\psi(2S)\pi^-$ mass spectra. Its observation in this channel can give new insights into its nature. A dispersive ansatz for low-energy description of QCD is outlined in Sec. 2.4, while Sec. 2.4.2 details a specific extension first introduced in [1], which is tested in this work using the $\pi^+\pi^-$ mass spectrum of $B_s^0 \to \psi(2S)\pi^+\pi^-$.

2.1 Standard Model of Particle Physics

A brief overview of the fundamental particles and interactions of the SM is given, a more detailed introduction from an experimental perspective can be found in textbooks, *e.g.* [19], while more involved concepts can be found in [20].

2.1.1 Particles

All visible matter in the universe is composed of the 12 fundamental spin-$\frac{1}{2}$ particles of the SM. Their properties are shown in Tab. 2.1[1]. They can be grouped into quarks, which are subject to the strong force, and leptons, which are not. The quarks can have electric charges of either $+\frac{2}{3}\,e$ (up-type) or $-\frac{1}{3}\,e$ (down-type), where e denotes the elementary charge. Leptons can have electric charge $-1\,e$ (e^-, μ^-, τ^-) or $0\,e$ (neutrinos). Each charged lepton is paired with one neutrino, while each down-type quark is paired with one up-type quark. These pairs are grouped into three generations with ascending mass, which is the only difference between the generations.

[1] Natural units are used throughout the thesis, implying $\hbar = c = 1$.

	Leptons				Quarks			
	Particle		Q	mass/ GeV	Particle		Q	mass/ GeV
First	electron	(e^-)	-1	0.0005	down	(d)	$-1/3$	0.003
generation	neutrino	(ν_e)	0	$< 10^{-9}$	up	(u)	$+2/3$	0.005
Second	muon	(μ^-)	-1	0.106	strange	(s)	$-1/3$	0.1
generation	neutrino	(ν_μ)	0	$< 0.2 \times 10^{-3}$	charm	(c)	$+2/3$	1.3
Third	tau	(τ^-)	-1	1.78	bottom	(b)	$-1/3$	4.5
generation	neutrino	(ν_τ)	0	$< 20 \times 10^{-3}$	top	(t)	$+2/3$	174.

Table 2.1: Fundamental fermions of the SM model with their masses and electric charges. Values taken from [21].

2.1.2 Interactions

The interactions of the SM particles are described via a $SU(3)_c \times SU(2)_L \times U(1)_Y$ gauge theory ($U(i)$ and $SU(i)$ stand for the i-dimensional unitary and special unitary group respectively). The strong interaction corresponds to the $SU(3)_c$ color group, while the electroweak interaction stems from its $SU(2)_L \times U(1)_Y$ part. Similar to the Noether theorem of classical physics, the symmetries of these groups introduce conserved charges into the SM: The electrical charge (Ward–Takahashi [22,23]), the 3 color charges, and the weak isospin (Slavnov–Taylor [24]).

The group structure also gives rise to the spin-1 gauge bosons, which mediate the interactions between particles that carry the corresponding charges: The photon (electric), 8 gluons (color), the $W^\pm$ and Z (weak isospin).

Unlike the photon and the gluons, the $W^\pm$ and Z bosons are massive, but their mass terms would violate the underlying $SU(2)$ symmetry. This is avoided via the Higgs mechanism, which generates their masses dynamically. It starts with a high-energy theory with massless particles and with an additional scalar field, whose symmetry is spontaneously broken at lower energies, the "electroweak symmetry breaking" (EWSB), giving rise to the mass terms of the $W^\pm$ and Z and a massive scalar particle called Higgs-boson [25,26]. This Higgs field can also explain the mass terms of the quarks and the charged leptons via Yukawa couplings [27,28]. From neutrino oscillation measurements it is established that neutrinos of different generations have non-zero mass differences, implying a non-zero mass for at least two of the neutrino species. Since possible right-handed neutrinos would be singlets under all SM interactions, it is not possible to generate neutrino mass terms inside the SM. The origin and first direct measurement of the neutrino masses are still open tasks for theory and experiment.

2.2 Weak Interaction and Quark Flavour Transitions

For quarks the mass eigenstates and the eigenstates of the weak charged current aren't the same. By convention the rotation between them is described via the Cabibbo-Kobayashi-Maskawa (CKM) mixing matrix V_{CKM} [29,30]. The components V_{ij} describe the transition of a quark of type i into a quark of type j. The corresponding weak charged-current vertices for quarks are given by

$$-i\frac{g_{\mathrm{W}}}{\sqrt{2}}(\overline{u},\overline{c},\overline{t})\gamma^{\mu}\frac{1}{2}\left(1-\gamma^{5}\right)\begin{pmatrix} V_{\mathrm{ud}} & V_{\mathrm{us}} & V_{\mathrm{ub}} \\ V_{\mathrm{cd}} & V_{\mathrm{cs}} & V_{\mathrm{cb}} \\ V_{\mathrm{td}} & V_{\mathrm{ts}} & V_{\mathrm{tb}} \end{pmatrix}\begin{pmatrix} d \\ s \\ b \end{pmatrix}, \tag{2.1}$$

where g_{W} is the weak coupling constant, γ^{μ} are the gamma matrices and $\frac{1}{2}(1-\gamma^{5})$ is the left-handed chiral projection operator (which is the reason for the parity violation in the weak interaction). d, s, b are the down-type quark spinors and $\overline{u}$, $\overline{c}$, $\overline{t}$ are the adjoint representations of the up-type quark spinors. If instead up-type quarks enter the interaction as spinors and the down-type quarks as adjoint spinors, the complex conjugate of the matrix is used.

The matrix is unitary, reducing the number of free parameters to three rotation angles and a complex phase. Since transitions between quark generations are suppressed, V_{CKM} is almost diagonal, which can be exploited in the Wolfenstein parametrisation [31], where the elements are expanded in the small parameter $\lambda \approx 0.23$. In addition the real parameters $A \approx 0.8$, $\rho \approx 0.13$, and $\eta \approx 0.35$ are needed to describe the degrees of freedom of the rotation:

$$\begin{pmatrix} V_{\mathrm{ud}} & V_{\mathrm{us}} & V_{\mathrm{ub}} \\ V_{\mathrm{cd}} & V_{\mathrm{cs}} & V_{\mathrm{cb}} \\ V_{\mathrm{td}} & V_{\mathrm{ts}} & V_{\mathrm{tb}} \end{pmatrix} = \begin{pmatrix} 1-\lambda^{2}/2 & \lambda & A\lambda^{3}(\rho-i\eta) \\ -\lambda & 1-\lambda^{2}/2 & A\lambda^{2} \\ A\lambda^{3}(1-\rho-i\eta) & -A\lambda^{2} & 1 \end{pmatrix} + O\left(\lambda^{4}\right) \tag{2.2}$$

The complex phase introduced by $i\eta$ is the source of *CP* violation in the SM [2], since it changes sign under *CP* transformations, while the phases introduced in the strong and electromagnetic interaction do not.

[2] with the possible exception of *CP* violation in the neutrino sector, where a similar matrix V_{PMNS} governs the transition between neutrinos of different generations allowing for a separate *CP* violating phase [32]. Within the current experimental precision it is still compatible with 0 [33]. There is no symmetry in the SM, which prohibits a *CP* violating term in the strong sector. Nevertheless it is experimentally constrained to negligible values. This is referred to as strong *CP* problem.

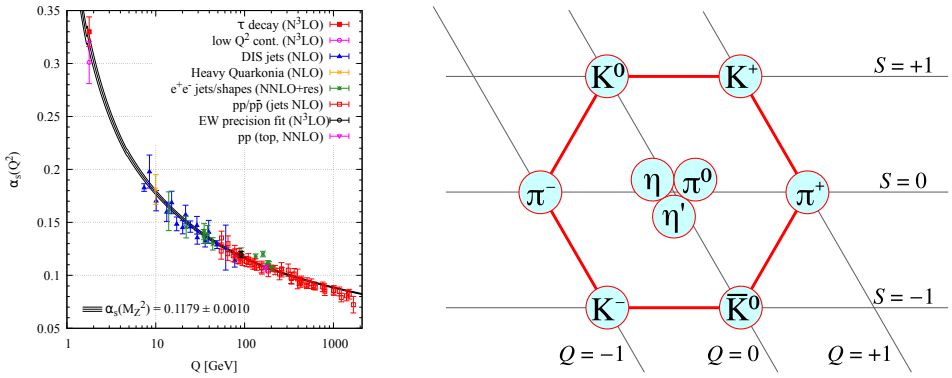

Figure 2.1: Left: Running coupling of QCD as function of momentum transfer Q in data (points) and from theory (line). Taken from [21]. Right: Nonet of pseudo-scalar mesons with spin 0 grouped via strangeness S and electric charge Q. Taken from [36]

2.3 QCD

Quantum Chromodynamics (QCD) is the theory underlying the strong interaction [34,35]. It describes the action of the fields of quarks q_f, antiquarks $\bar{q}_f$ and 8 gluons A^c_μ ($c = 1...8$) in the Lagrange density

$$\mathcal{L}_{\text{QCD}} = \sum_{f=u,d,s,c,b,t} \bar{q}^a_f \left(i D^{ab}_\mu \gamma^\mu - m_f \delta^{ab} \right) q^b_f - \frac{1}{4} \left(G^a \right)_{\mu\nu} \left(G^a \right)^{\mu\nu}, \tag{2.3}$$

where

$$D^{ab}_\mu = \partial_\mu \delta^{ab} - i g_s \lambda^{ab}_c A^c_\mu \tag{2.4}$$

is the covariant derivative with λ^{ab}_c ($a, b = 1...3$ running over the three colour charges) the adjoint representations of the generators of the $SU(3)$ algebra. The gluonic field strength tensor is given by

$$\left(G^a \right)_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g_s f^{abc} A^b_\mu A^c_\nu. \tag{2.5}$$

Here f^{abc} are the structure constants defined as

$$[\lambda_a, \lambda_b] = i f^{abc} \lambda_c \tag{2.6}$$

and g_s is the strong gauge coupling constant related to the coupling strength of the strong interaction

$$\alpha_s = \frac{g_s^2}{4\pi^2}. \tag{2.7}$$

The value of α_s at a scale Q^2 can be inferred from its value at a renormalisation scale Λ via

$$\alpha_s(Q^2) = \frac{\alpha_s(\Lambda^2)}{1 + B\alpha_s(\Lambda^2)\log\left(\frac{Q^2}{\Lambda^2}\right)}, \tag{2.8}$$

where B depends on the numbers of fermionic (quarks) and bosonic (gluons) objects with colour charge N_f and N_c in the theory

$$B = \frac{11N_c - 2N_f}{12\pi}. \tag{2.9}$$

This so-called *running* of the coupling strength is an effect of higher order quantum corrections to the interaction. It is closely related to the renormalisation of the theory, which is necessary to treat singularities related to self-interactions of quarks and gluons.

Fig. 2.1 shows the measured values for α_s at different momentum transfers. It decreases with growing Q, approaching 0 for large momenta (small distance scales, weak coupling regime), where quarks and gluons act like free particles. This phenomenon is known as *asymptotic freedom* and allows to study individual quarks and gluons in high-energy processes. In this regime it is also possible to describe the strong interaction perturbatively via Feynman diagrams, which expand the interaction in powers of α_s.

The figure also shows that for low energies α_s becomes very large (strong coupling regime), which makes it impossible to describe the strong interaction with the usual perturbative approaches. The energy scale, at which perturbation theory breaks down in QCD, is $\Lambda_{\mathrm{QCD}} \approx 200 - 400\,\mathrm{MeV}$.

Another phenomenon is *colour confinement*, which means that no free coloured object is found in nature. Although no mathematical proof for this behaviour of QCD has been given yet, decades of observations have confirmed this property. A conventional graphical explanation is that the gluon field between two coloured objects forms a flux tube, which grows longer as the two objects are moved further apart. In this picture the force between them does not decrease with distance (like it would in the case of electrically charged objects) so the energy content of the system increases until it becomes energetically favourable to form a $q\bar{q}$ pair with matching colour and anti-colour in between. The new quark and antiquark pair up with each of the initial objects creating two uncoloured hadronic objects. This process is called hadronisation.

The combinations of 3 quarks $qq'q''$ into a hadron are called baryons, while combinations of quark-antiquark $q\bar{q}'$ are called mesons.

2.3.1 Chiral Theory

As shown in the previous section, low-energy QCD processes like the description of the hadronic bound states are hard to access at the quark-level, both from the theoretical side, due to the running of the strong coupling α_s, and from the experimental side, due to the confinement of quarks and gluons. In the following, chiral theory is briefly introduced, which provides an effective description of low-energy QCD phenomena like the observed spectrum of the light hadronic states, which lie at the center of this thesis. A detailed introduction to the application of chiral symmetry in nuclear physics is given in [37].

Since strong interactions behave very differently above and below Λ_{QCD}, the quarks in Tab. 2.1 can be grouped into heavy quarks with masses above (c, b, t) and light quarks with masses below Λ_{QCD} (u, d, s). In the following we only focus on the light quarks and the hadronic states that are made up from them.

At energies below Λ_{QCD} the mass terms of the light quarks in eq. 2.3 can be ignored. This massless $\mathcal{L}_{\mathrm{QCD}}$ then exhibits a global $SU(3)_f$ flavour symmetry for the left-handed and right-handed quarks separately, corresponding to a $SU(3)_{\mathrm{L}} \times SU(3)_{\mathrm{R}}$ symmetry group.

This symmetry is spontaneously broken through a non-zero vacuum expectation value for the scalar quark condensate $\langle q\bar{q}\rangle \neq 0$, which reduces the symmetry of the Lagrangian to the $SU(3)_V$ flavour symmetry. This process generates mass terms for all hadrons consisting of light quarks, except for the resulting eight Nambu-Goldstone bosons from the generators of the broken symmetries, which correspond to the pseudoscalar meson multiplet shown in Fig. 2.1 (except for the η', which corresponds to a $SU(3)_V$ singlet state).

Since the chiral symmetry is explicitly broken by the small quark mass terms, the pseudoscalar mesons are only pseudo Nambu-Goldstone bosons and therefore not massless. Instead the pions (π^+, π^-, π^0) acquire a mass of ≈ 130 MeV, while the Kaons (K^+, K^-, K^0) and the η have masses of about 500 MeV[3]. This difference can be explained by the strange quark mass being much larger than the up- and down-quark mass $m_s \gg (m_d + m_u)/2$. The

[3] The η', which is not protected by the symmetry has a mass of 957 MeV.

chiral symmetry is therefore broken to a larger degree for mesons with a significant strange quark contribution to their wavefunction.

The so-called *isospin* symmetry holds to a much better degree, because its $SU(2)_f$ group leaves out the strange quark and the masses of u and d quarks are much smaller than Λ_{QCD}. Here u and d quarks have an isospin quantum number of $I = 1/2$, while the third component I_3 is $+1/2$ for u quark and $-1/2$ for d quark. For $\overline{u}$ and $\overline{d}$ the I_3 component switches sign. Other quarks have $I = 0$ and therefore $I_3 = 0$.

Note that for high enough densities and temperatures the chiral symmetry is restored in this model and the masses of the pseudoscalar mesons should increase to values similar to those of their scalar counterparts like the $f_0(500)$.

In chiral theory the scalar mesons also form a nonet of states, but unlike the pseudoscalars they are much harder to access due to their large widths and the opening of additional decay channels near to their pole positions. This makes a description and interpretation of their dynamics in scattering experiments very challenging both experimentally and theoretically. Describing the dynamics of these scalar mesons in their coupling to two pions is a central part of this work.

An additional complication in the classification of these states lies outside of the simple chiral theory. Exotic hadronic states (introduced in Sec. 2.6) with the same quantum numbers are predicted for the same energy region. The observed states could even be superpositions of classical $q\overline{q}$ configurations with significant exotic contributions. Therefore, establishing the quark (and gluon) content of the scalar mesons is still an active field of research (an overview can be found in the review on scalar mesons in Ref. [21]).

2.4 Dispersion Theory

A brief overview of the properties of the low-energy dispersive scattering theory are given following the review on resonances in [21] and descriptions in [38, 39], where further information about the subject is available. Even though it can be applied to other interactions and the high-energy regime as well, the focus lies on the description of hadronic phenomena.

It was shown in Sec. 2.3 that the strong coupling constant grows towards lower energy scales making a perturbative description impossible. S-matrix theory represents a valid alternative in these energy regions.

Its center is the Lorentz-invariant unitary operator called S-matrix S, which describes the transition amplitude from an initial state $|i\rangle$ to the final state $\langle f|$:

$$S_{fi} = \langle f|S|i\rangle \tag{2.10}$$

with

$$SS^{\dagger} = S^{\dagger}S = I. \tag{2.11}$$

The T matrix corresponds to those transitions, where an interaction has taken place:

$$S = I + iT. \tag{2.12}$$

Since PT, the combined application of the parity operator P and time operator T, is an exact symmetry in the strong interaction,

$$\langle i|S|f\rangle = \langle f|S|i\rangle. \tag{2.13}$$

Eq. 2.13 can be combined with the unitarity condition in eq. 2.11 and after some algebra (see Ref. [38]) results in

$$\Im\langle f|T|i\rangle = \frac{1}{2}\int \sum_m (2\pi)^4 \delta^4\left(p_i - p_m\right) \langle f|T|m\rangle^{\star}\langle m|T|i\rangle, \tag{2.14}$$

where $\Im(x)$ corresponds to the imaginary part of x, the sum over m runs over all intermediate states, which are allowed by symmetries and kinematic constraints. The delta function $(2\pi)^4\delta^4\left(p_i - p_m\right)$ enforces properly normalised 4-momentum conservation. For $\langle f| = \langle i|$ eq. 2.14 results in the optical theorem

$$\begin{aligned} \Im\left\langle i|T|i\right\rangle =& \frac{1}{2}\int \sum_m (2\pi)^4 \delta^4\left(p_i - p_m\right) \left|\langle m|T|i\rangle\right|^2, \\ \Rightarrow \sigma_{elastic}|_{\text{forward}} =& \sigma_{\text{tot}} \end{aligned} \tag{2.15}$$

which relates the imaginary part of the elastic forward scattering amplitude to the total cross section of the process.

Further relations can be explored exemplarily using 2 by 2 scattering of scalar particles with momenta p_i and masses m_i

$$\phi_1(p_1) + \phi_2(p_2) \rightarrow \phi_3(p_3) + \phi_4(p_4). \tag{2.16}$$

Since the particles are spinless, the process can be fully described with the Mandelstam [40] variables

$$s = (p_1 + p_2)^2, t = (p_1 - p_3)^2, u = (p_1 - p_4)^2, \tag{2.17}$$

of which only two variables are independent, since

$$s + t + u = m_1^2 + m_2^2 + m_3^2 + m_4^2. \tag{2.18}$$

Alternatively the scattering process can be described using only s and the scattering angle θ_s, which describes the angle between $\vec{p}_1$ and $\vec{p}_3$ in the center of mass system (cms). This makes it possible to write the transition amplitude as partial-wave projection

$$\langle f|T|i\rangle = T^{fi}(s,t,u) = T^{fi}(s,\cos\theta_s) = 16\pi \sum_{\ell=0}^{\infty}(2\ell+1)P_\ell(\cos\theta_s)t_\ell^{fi}(s), \tag{2.19}$$

where $P_\ell(\cos\theta_s)$ denote the Legendre-polynomials and $t_\ell^{fi}(s)$ the partial wave amplitude. This result is very powerful: Since the Legendre polynomials are orthogonal, transitions corresponding to different angular momenta ℓ can be treated separately. This is one of the cornerstones of the helicity formalism used in Sec. 5.2. A thorough description of the method – also for states with spin – can be found in [41, 42].

Combining eq. 2.19 with eq. 2.14 and considering only two-particle intermediate states a and b for simplicity, results in

$$\Im\, t_\ell^{fi} = \sum_{m=\{m_a,m_b\}} (t_\ell^{fm}(s))^* \sigma(s,m_a,m_b) t_\ell^{mi}(s), \tag{2.20}$$

where $\{m_a, m_b\}$ stands for the intermediate state with particle a and particle b and $\sigma(s,m_a,m_b)$ is the two-body phase space

$$\sigma(s,m_a,m_b) = \frac{\sqrt{\lambda(s,m_a^2,m_b^2)}}{s}\theta(s-(m_a+m_b)^2), \tag{2.21}$$

where $\lambda(s,m_a^2,m_b^2)$ is the Källen function

$$\lambda(a,b,c) = \left(a - \left(\sqrt{b}+\sqrt{c}\right)^2\right)\left(a - \left(\sqrt{b}-\sqrt{c}\right)^2\right) \tag{2.22}$$

and $\theta(s,m_a,m_b)$ is the Heaviside step function.

CPT invariance allows to relate the 2 by 2 scattering described above (the

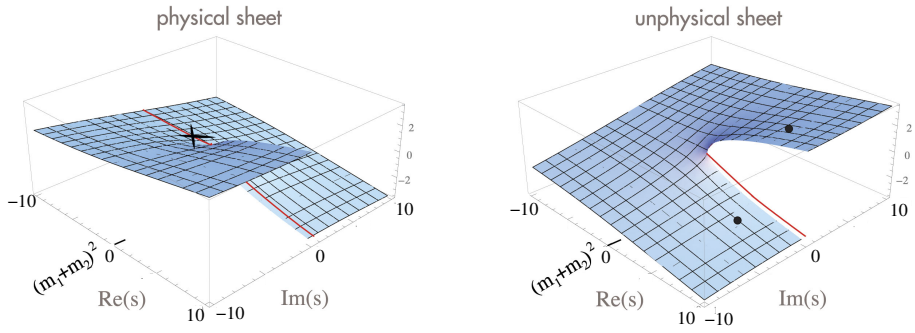

Figure 2.2: Sketch of imaginary part of a single channel amplitude in the complex s-plane. Solid dots on unphysical sheet show allowed positions for resonance poles, the cross on the physical sheet the position for a bound state. The red solid line denotes the physical axis (shifted into the physical sheet). The two sheets are connected smoothly along their discontinuities. Taken from review on resonances in [43].

so called s-channel) to other processes via replacing particles with momentum p with antiparticles with momentum $-p$

$$\begin{aligned} \phi_1(p_1) + \overline{\phi}_3(-p_3) &\to \overline{\phi}_2(-p_2) + \phi_4(p_4), \\ \phi_1(p_1) + \overline{\phi}_4(-p_4) &\to \overline{\phi}_3(-p_3) + \phi_2(p_2), \end{aligned} \tag{2.23}$$

which are called t-channel and u-channel respectively. The corresponding Mandelstam variables are related as well. The Mandelstam variables s_t, t_t, and u_t of the t-channel process are given by

$$\begin{aligned} s_t &= (p_1 - p_3)^2 = t, \\ t_t &= (p_1 + p_2)^2 = s, \\ u_t &= (p_1 - p_4)^2 = u. \end{aligned} \tag{2.24}$$

This behaviour is called crossing symmetry. The s, t, and u-channels are all described by the same amplitude $T(s,t,u)$, just in different regions of the $(s, t-u)$-plane.

Because of energy conservation the right hand side of eq. 2.14 only evaluates to non-zero values above the energy threshold for the initial and final state particles. In the two-body scattering case this corresponds to

$$s > \max\{(m_1+m_2)^2, (m_3+m_4)^2\} = s_{\text{thr}}. \tag{2.25}$$

The effect of the onset of an imaginary part is shown in Fig. 2.2.

Above s_{thr} the scattering amplitude has an imaginary part from eq. 2.14. The point s_{thr} is called a branch point, because the two solutions of the root function in eq. 2.21 branch off from this point, while below there is only one solution. The two solutions correspond to a so-called right-hand cut along the real axis of the complex s-plane [21], which separates the two solutions of the root function. It is called right-hand, because it starts at a finite value s_{thr} and extends to $+\infty$. The two solutions define two Riemann surfaces in the complex plane, which are called physical sheet and unphysical sheet. Branch points appear, whenever a channel opens up. This means that at each threshold for massive particles, the number of Riemann sheets doubles [21].

Due to the crossing symmetry, channel openings in the crossed channels t and u also introduce cuts, *e.g.* for

$$t > \max\{(m_1 + m_3)^2, (m_2 + m_4)^2\} = t_{\text{thr}}. \tag{2.26}$$

the t-channel gets imaginary contributions (see eq. 2.24). This leads to a so-called left-hand cut in the s-channel, because this channel is only kinematically allowed for large values of t and kinematically large values of t coincide with low values of s. The physical region for s-channel interactions therefore corresponds to the region, where the amplitude has a right hand cut, while the physical region for t-channel interactions lies, where the s-channel has a right-hand cut. This is illustrated in Fig. 2.3

Bound states of the interaction are defined as poles, which lie on the physical axis below s_{thr}, while resonances are defined as poles on the unphysical sheet (see Fig. 2.2, noting that the physical axis does not extend below s_{thr} in the unphysical sheet.) [21].

Using $S_{ba} = \delta_{ba} + 2if_{ba}$, where $f_{ba} = t_\ell^{ba}(s)/\sigma(s, m_a, m_b)$ (the index ℓ was dropped for convenience here) one can use the unitarity of the S-matrix to get

$$f_{bb} = (\eta \exp(2i\delta_b) - 1)/2i, \tag{2.27}$$

where δ_b denotes the phase shift for scattering of channel b to b, and η the elasticity parameter (also called inelasticity) [21]. For $\eta = 1$, the scattering is purely elastic, while in general $0 \leq \eta \leq 1$. The trajectory of f_{bb} for varying s is displayed in a so-called Argand plot (see Fig. 2.3 on the right for an example).

Often the scattering amplitude T is decomposed into two parts. One which contains all the poles, called T^{pole}, and T^{bg}, the background part, which does not contain any poles.

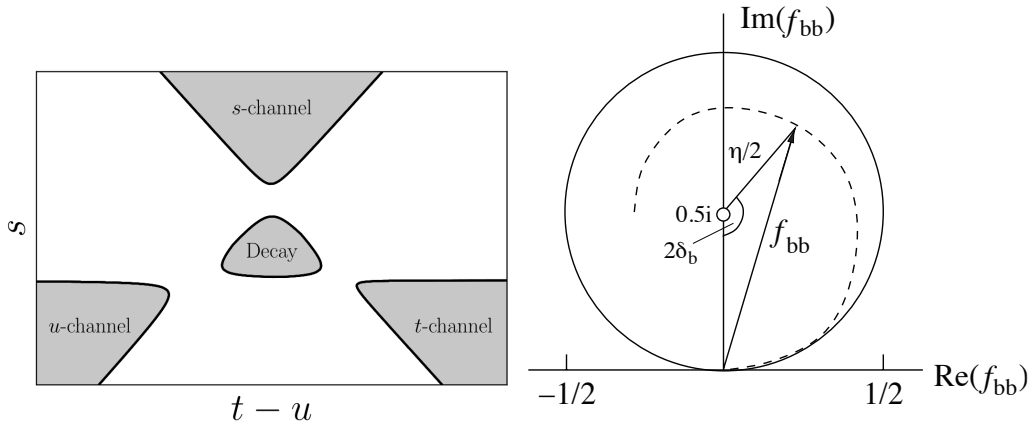

Figure 2.3: Left: Sketch of Mandelstam plane for an arbitrary scattering process. The shaded regions are the physically allowed regions for interactions. Since one of the particles is heavy enough to decay to the three others, a decay region is also included. Taken from [1]. Right: Argand plot showing diagonal element of partial-wave amplitude as function of energy. For elastic processes, the trajectory follows the unit circle, but leaves it as soon as inelasticity sets in and $\eta < 1$, *e.g.* at a threshold, where an additional decay channel opens up. Taken from [21].

$$T = T^{\text{bg}} + T^{\text{pole}}. \tag{2.28}$$

This choice is not unique, but the position of the poles in the complex s-plane is independent of it. One usually assumes that T^{bg} is unitary and that its phase motion is much smaller than that of T^{pole}.

The poles are defined through their mass M_{R} and width Γ_{R} as

$$\sqrt{s_{\text{R}}} = M_{\text{R}} \pm i\Gamma_{\text{R}}/2. \tag{2.29}$$

Note that these values generally do not correspond to the mass and width parameters M_0 and Γ_0 of the Breit-Wigner parametrization of resonances introduced in Sec. 2.4.1.

For N resonances in a particular channel the pole part can be written in the general form

$$T_{fi}^{\text{pole}}(s) = \Omega_f(s) \left[1 - V^{\text{R}}(s)\Sigma(s)\right]^{-1}_{fc} V^{\text{R}}_{ci}(s)\Omega_i(s), \tag{2.30}$$

where $\Omega_i, \Sigma(s)_{ij}$ and $V^{\text{R}}_{ij}(s)$ are matrices in channel space. Ω_i is the normalized vertex function and $\Sigma(s)_{ij}$ corresponds to the self-energy. The resonance scattering kernel $V^{\text{R}}_{ij}(s)$ is defined as

$$V_{fi}^{\mathrm{R}}(s) = -\sum_{n=1}^{N} \frac{g_{nf}\, g_{ni}}{s - M_n^2}, \tag{2.31}$$

where g_{na} denotes the coupling of resonance R_n to channel a and M_n is its mass parameter (not necessarily the real part of the pole position).

From the experimental point of view the coupling to a channel f of a resonance produced in channel i is defined via the residue $\mathcal{R}^{fi}$ of the corresponding pole, which can be calculated from a contour integral around that pole in the transition amplitude

$$\mathcal{R}^{fi} = \frac{1}{2\pi i} \oint ds T^{fi}. \tag{2.32}$$

It can be corrected for the coupling to the production channel i via defining the pole coupling

$$\tilde{g}_f = \mathcal{R}^{fi} / \sqrt{\mathcal{R}^{ii}}. \tag{2.33}$$

Using $\tilde{g}_f$ a partial width can be defined as

$$\Gamma_f = \frac{|\tilde{g}_f|^2}{M_{\mathrm{R}}} \sigma_f(M_{\mathrm{R}}^2), \tag{2.34}$$

where σ is the phase space of the channel f [43].

The pole positions and their pole couplings are process independent and fully describe a resonance. Process independence means here, that these values are the same for a given resonance, independent, whether they are measured in a channel a or a channel b, via a scattering experiment or in a particle decay.

The physical quantities of interest for amplitude analyses for hadron spectroscopy are therefore the pole parameters. In order to extract the correct values for these, a solution for S needs to be found, which fulfils unitarity and analyticity and describes the data, which is measured on the real axis. Then the pole positions and their residues are determined by analytic continuation of the amplitudes measured on the physical sheet, to the corresponding unphysical sheet, where the the pole lies.[4]

[4] This last step is beyond the scope of this thesis and. More on this can be found *e.g.* in [38].

2.4.1 Breit-Wigner Model

Unfortunately, fully model-independent dispersive approaches can not be used for all energy regions and processes, *e.g.* with the onset of inelasticities at high energies, an explicit model needs to be introduced to describe them effectively.

All phenomena, which cannot be described purely dispersively, therefore need to be modelled with some assumptions. One popular simplification is the Breit-Wigner parametrisation. In the following its derivation is shown. Afterwards it is explained, why this parametrisation in general does not fulfil the unitarity requirement on amplitudes and cannot be used to extract pole parameters or couplings, unless the corresponding pole is well separated from other poles and thresholds.

For a single resonance eq. 2.30 can be simplified to [43]

$$T_{ba}^{\text{pole}}(s) = \Omega_b(s)\frac{g_b g_a}{s - \hat{T}^{\text{R}}(s)^2 + i\sqrt{s}\Gamma^{\text{R}}(s)_{\text{tot}}}\Omega_a(s), \tag{2.35}$$

with mass function $\hat{T}^{\text{R}}(s) = M^2 + \sum_c g_c^2 \Re(\Sigma_c(s))$, where the sum runs over all channels c. The total width is related to the imaginary part of the self energy via

$$\Gamma^{\text{R}}(s)_{\text{tot}} = \sum_c \Gamma_c^{\text{R}}(s) \tag{2.36}$$

with

$$\Gamma_c^{\text{R}}(s) = \frac{(2\pi)^4}{2\sqrt{s}} g_c^2 \text{Disc}\,\Sigma_c(s), \tag{2.37}$$

and the discontinuity of the self energy

$$\text{Disc}\,\Sigma_a(s) = [\Sigma_a(s) - \Sigma_a^*(s)] = 2i\sum_c \sigma_c(s)\Omega_{ca}^*(s)\Omega_{ca}(s). \tag{2.38}$$

If all relevant thresholds are far away from the observed mass region ($2(M^2 - s_{\text{thr}}) \gg \Gamma_{\text{tot}}^{\text{R}}$), then $\Gamma^{\text{R}}(s)_{\text{tot}} \approx \Gamma_0$ can be considered constant. In this case also $\Re(\Sigma)$ is constant and $\hat{T}^{\text{R}}$ can be replaced with M_0 [43]. For narrow resonances also $\sqrt{s} \approx M_0$ can be used, which simplifies eq. 2.35 to the Breit-Wigner parametrisation

$$T_{ba}^{\text{pole}} = -\frac{g_b\, g_a}{s - M_0^2 + iM_0\Gamma_0}. \tag{2.39}$$

This description contains an additional simplification: For $T^{\text{bg}} = 0$, the vertex functions can be written as

$$\Omega_a(s) = p_a^{\ell_a} F_{\ell_a}(p_a, p_0), \tag{2.40}$$

where p_a is the momentum of one of the particles in state a in their cms and ℓ_a is the angular momentum of the particles in channel a. Usually F_{ℓ_a} is included in form of the Blatt-Weisskopf phenomenological form factor [44], which depends also on an intrinsic scale p_0, which could be the value of p_a at M_0 for Breit-Wigner parametrisations. The form factors scale with $p_a^{-\ell_a}$ for large energies. This counteracts the contribution of the threshold terms $p_a^{\ell_a}$ to the amplitude, which diverge for large energies, making the combined amplitude normalisable.

In eq. 2.39 an $\ell = 0$ (S-wave) partial wave is assumed, which reduces the vertex functions to $\Omega_a(s) = 1$.

Only if all these conditions are met, the parameters M_0 and Γ_0 of eq. 2.39 correspond to the pole parameters M_R and Γ_R.

Even if only one resonance is present in the studied partial wave, its parameters can be different from the real pole parameters, if it lies too close to a threshold, where additional coupling channels open up. Depending on the coupling strength to the opening channel, the width can change drastically at the threshold. In that case the width parameter is energy-dependent and has to be described with eq. 2.36. This is done for the so-called Flatté amplitude introduced in Sec. 5.5, which is able to describe the $f_0(980)$ shape correctly, even though this resonance lies in the direct vicinity of the $K\bar{K}$ threshold.

The sum of two Breit-Wigner amplitudes is generally not an adequate description of two resonances, if they are not well separated: Consider a single channel with two poles, that are supposed to be described with two Breit-Wigner amplitudes from eq. 2.39

$$T(s) = -\frac{\tilde{g}_1}{s - M_1^2 + iM_1\Gamma_1} - \frac{\tilde{g}_2}{s - M_2^2 + iM_2\Gamma_2}. \tag{2.41}$$

A mass dependence of the width $\Gamma(s)$ does not change the outcome, but is left out for simplicity. Setting i to f in eq. 2.20 results in the two-body version of the optical theorem from eq. 2.15

$$\Im\, T = \sigma |T|^2, \tag{2.42}$$

where σ is the two-body phase space defined in eq. 2.21. This results directly from the unitarity condition on the S-matrix. If it is not fulfilled, this means the unitarity condition is violated. Using now eq. 2.41 one can rewrite the left

side of eq. 2.42 to

$$\Im T = \frac{\tilde{g}_1^2 \Gamma_1 M_1}{(s - M_1^2)^2 + i M_1^2 \Gamma_1^2} + \frac{\tilde{g}_2^2 \Gamma_2 M_2}{(s - M_2^2)^2 + i M_2^2 \Gamma_2^2}. \tag{2.43}$$

Using eq. 2.34 the right side of eq.2.42 can be evaluated to

$$\begin{aligned} |T|^2 = & \frac{\tilde{g}_1^2 \Gamma_1 M_1}{(s - M_1^2)^2 + i M_1^2 \Gamma_1^2} + \frac{\tilde{g}_2^2 \Gamma_2 M_2}{(s - M_2^2)^2 + i M_2^2 \Gamma_2^2} \\ & + 2\sigma \mathcal{R}e \left[\frac{\tilde{g}_1}{s - M_1^2 + i M_1 \Gamma_1} \frac{\tilde{g}_2}{s - M_2^2 + i M_2 \Gamma_2} \right]. \end{aligned} \tag{2.44}$$

Eq. 2.42 holds only (and therefore unitarity is only conserved), when the interference term is zero. This is only the case, if the two resonances are sufficiently narrow and separated so that they have no significant overlap. But this is *e.g.* not the case for the $f_0(980)$ and $f_0(1500)$ resonances observed in the decay $B_s^0 \to \psi(2S)\pi^+\pi^-$.

It might still be possible to describe the experimental data with high precision using sums of overlapping Breit-Wigner amplitudes, even if not all conditions above are met. But due to unitarity violation, the corresponding Breit-Wigner parameters are process dependent, which means that, while they are able to describe one process, they cannot be used to adequately describe a different process. A Breit-Wigner parametrisation of a resonance extracted from a decay measurement is then not be suitable to adequately describe the same resonance in a scattering experiment. This is different for the pole parameters of resonances, which are fixed values on the complex s-plane, independent of the involved processes.

2.4.2 Bonn-Jülich Model

The Flatté parametrisation of the $f_0(980)$ describes the dipion spectrum at the $K\bar{K}$ threshold correctly. If the studied mass region is extended to higher values though, the presence of further, overlapping S-wave states, mostly the $f_0(1500)$, results in a combined amplitude, which does not fulfil the analyticity and unitarity requirements. Ultimately this means that the extracted pole positions from the fit are not universal and are process-dependent. Also the couplings, fit fractions and relative phases are not reliable, even if the fit can describe the data well.

This calls for a more sophisticated approach to describe the dipion spectrum beyond the $K\bar{K}$ threshold correctly. The presented analysis tests a specific model for the scalar dipion S-wave component in $B_s^0 \to \psi(2S)\pi^+\pi^-$, which was described in [1, 45, 46]. The model was implemented into the analysis of $B_s^0 \to \psi(2S)\pi^+\pi^-$ in close collaboration with the authors of these publications.

The model provides a robust, process independent parametrisation of the isoscalar S-wave of the dipion system. Its parameters were determined from low-energy pion-pion scattering data and data from $B_s^0 \to J/\psi\,\pi^+\pi^-$ and $B_s^0 \to J/\psi\,K^+K^-$ decays (note that here the 1S charmonium state and not the 2S state was used). The model can be used to describe the dipion S-wave spectrum in other decay processes, where a direct determination of the dipion S-wave component would increase the uncertainties on the actual parameters of interest. An example would be measurements of time-dependent *CP* violating phase differences, where the uncertainty of the model of the hadronic interactions can make up a large fraction of the total systematic uncertainty. These analyses sometimes use simple models for the parametrisation of the S-wave component, which do not obey unitarity and analyticity constraints. As shown in Sec. 2.4.1, these parametrisations are process-dependent and can therefore not easily be transferred from one measurement to another. Consequently, the description of the strong phase motion of the amplitude is not necessarily determined correctly by this approach either. Several solutions might exist for naive parametrisations. Since the extracted *CP* violating parameters depend on both strong and weak phases, the large uncertainty on the strong phase motions directly degrades the accuracy of the measurement of the *CP* violating parameters.

One such case is the decay $B_s^0 \to J/\psi\,\pi^+\pi^-$, where two S-wave resonances $f_0(980)$ and $f_0(1500)$ were found to overlap significantly in the dipion spectrum [47]. A process-independent description of the dipion S-wave component could be used as external input to the amplitude fit and reduce the uncertainty on the extracted *CP* parameters.

In the following only the main points of the method are touched upon; further information can be found in [1, 45, 46] and a more thorough introduction is given in [39].

The method builds upon precise measurements available for the low-energy scattering behaviour of light mesons, which has been measured experimentally over the last decades. These measurements were implemented into an S-matrix description in a fully model-independent, dispersive way.

This scattering amplitude is able to describe the $\pi\pi$ scattering data excellently up to energies of 1 GeV, including the non-Breit-Wigner shapes of the $f_0(500)$ and the $f_0(980)$ resonances. Above this value additional inelastic channels open up making a model-dependent approach necessary. Since couplings of the additional channels are dominated by resonances, this high-energy part is described via resonance exchange [48]. It is known from scattering experiments that 3 channels are open and coupling to the resonances: $\pi\pi$, KK, and $\pi\pi\pi\pi$, where the latter is modelled via either $\rho\rho$ or $\sigma\sigma$ resonance exchange.

The scattering matrix is modelled with a recursive Bethe-Salpeter equation [49], which correctly implements unitarity and analyticity

$$T_{if} = V_{if} + V_{im} G_{mm} T_{mf}, \tag{2.45}$$

where V_{if} is the scattering kernel of the channel i into channel f. Multiply occurring indices are being summed over. G is the diagonal loop operator, which describes the free propagation of the particles in channel $m \in \{\pi\pi, KK, 4\pi\}$. The finite width of $\rho\rho$ and $\sigma\sigma$ are taken care of with an effective description in the propagator. More details on this can be found in [39].

The two-potential decomposition introduced in eq. 2.28 is used to separate the well-described low-energy part of the scattering amplitude, which is taken as external input, from the high-energy part, which is dominated by resonance exchange [48]. This is achieved via splitting the scattering kernel $V = V_0 + V_\mathrm{R}$ and the scattering matrix $T = T_0 + T_\mathrm{R}$. The term V_0 is fixed from the external input from the low-energy part including the effects of the $f_0(980)$ resonance, while heavier resonances enter the amplitude via V_R.

Since V_0 cannot be assumed constant (in contrast to the case for the derivation for the Breit-Wigner amplitude in eq. 2.39), the vertex function $\Omega = 1 + T_0 G$ is introduced, which can be identified as an Omnes matrix and dispersively calculated [38].

The full solution for the scattering matrix can be found from eq. 2.45 [39]

$$T(s) = T_0 + \Omega[1 - V_\mathrm{R}\Sigma]^{-1} V_\mathrm{R} \Omega^t, \tag{2.46}$$

where Σ is the self-energy matrix $\Sigma = G\Omega$ with a discontinuity

$$\mathrm{disc}\,\Sigma_{ij}(s) = 2i\Omega^\dagger_{im}(s)\mathrm{disc}\,G_{mm}(s)\Omega_{mj}(s). \tag{2.47}$$

It is possible to write Σ as once-subtracted dispersion integral [39]

$$\Sigma_{ij}(s) = \Sigma_{ij}(0) + \frac{s}{\pi} \int \frac{dz}{z} \frac{\text{disc}\,\Sigma_{ij}(z)}{z - s - i\epsilon}, \tag{2.48}$$

The subtraction constants $\Sigma_{ij}(0)$ are an artefact of so-called subtractions, which divide the amplitude $\Sigma_{ij}(z)$ by a polynomial of order n $P_n(z)$, if the simple dispersion integral diverges. The polynomial does not contain any discontinuities, which means that the analytic structure of

$$g(z) = \Sigma_{ij}(z)/P_n(z), \tag{2.49}$$

is the same as that of the original $\Sigma_{ij}(z)$. The subtraction constants contain information about the low-energy regime, which cannot be constrained by the dispersion integral. In the case of the Bonn-Jülich model they are absorbed into V_R in eq. 2.46 [39].

After some algebra, eq. 2.48 can be used to calculate the form factors for the three channels

$$\Gamma_i^s = \Omega_{im}[1 - V_\text{R}\Sigma]^{-1}_{mn} M_n \tag{2.50}$$

V_R does not have a strong impact for low energies, where the dispersive results contained in T_0 (which itself contributes to $\Omega_{im}(z)$ and $\Sigma_{ij}(z)$) describe the data very well. This is why the subtraction constant in eq. 2.48 can be absorbed into V_R leading to

$$(V_\text{R})_{ij} = \sum_\text{R} g_i^\text{R} \frac{s}{m_\text{R}^2(m_\text{R}^2 - s)} g_j^\text{R}, \tag{2.51}$$

where the subtraction constants were already absorbed into T_0 in eq. 2.46. The sum runs over the additional resonances above the $f_0(980)$.

The bare resonance-channel couplings g_i^R and bare resonance masses m_R are free parameters, which need to be determined from data. They do not coincide with the physical values for couplings and masses, but they were related with each other in [39] via the method of Padé approximants [50–52].

The transition matrix M_i from source to channel i uses the following parametrisation in [39]

$$M_i = c_i + l_i s + \ldots - \sum_\text{R} g_i^\text{R} \frac{s}{m_\text{R}^2 - s} \alpha_\text{R}. \tag{2.52}$$

The bare masses and resonance-channel couplings are the same as in eq. 2.51. The M_i and therefore the form factors Γ_i are fixed up to a real polynomial in s,

where the 0th order are the c_i, which are the normalisations of the contributing form factors at $s = 0$. Its 1st order is denoted by the l_i. They are not constrained by theory considerations, but in practice these linear coefficients turn out to be small. The real parameters α_R denote the resonance-source couplings, where R are the high-energy resonances.

The analysis in [1, 39] determined the form factors from eq. 2.50 and the physical couplings and pole parameters in a fit to $B_s^0 \to J/\psi\pi^+\pi^-$ and $B_s^0 \to J/\psi K^+K^-$ data from LHCb. They use one source (B_s^0) and the three channels defined above ($\pi\pi$, KK, and 4π). Two high-energy resonances were found to be sufficient to describe the data. Their poles were determined and identified with the $f_0(1500)$ and the $f_0(2020)$ resonances.

The full dipion form factor is the sum of the three factors described in eq. 2.50 and can be be rewritten to separate the the external inputs more clearly from the process-dependent parameters:

$$F_\pi(s) = \sum_{i=\pi\pi,KK,4\pi} \tilde{F}_i(s)(c_i + l_i s) + \sum_{R=f_0(1500),f_0(2020)} \bar{F}_R(s)\alpha_R, \tag{2.53}$$

where

$$\bar{F}_R(s) = \tilde{F}_i(s) g_i^R \frac{s}{m_R^2 - s} \tag{2.54}$$

and $\tilde{F}_i(s)$ corresponds to the remaining terms of Γ_i^s from eq. 2.50.

Since the weak decay of $B_s^0 \to \psi(2S)\pi^+\pi^-$ contains to leading order a pure pure $s\bar{s}$ source (see Fig. 2.4) c_π and $c_{4\pi}$ are negligibly small and only c_K is non-zero. The value of c_k is process-dependent and can therefore not be transferred to different processes. The same is true for the parameters α_R and l_i. All other parameters of the model, including the bare masses and couplings of eq. 2.54 can be simply taken from the high-statistics results in [1].

The only process-dependent parameters in this model are therefore real and linear. This means that the shape and phase motion of the S-wave is fully fixed by the Bonn-Jülich model and only the relative contributions of the base functions $\tilde{F}_i(s)$ and $\bar{F}_R(s)$ need to be fitted, if it is applied in a different process. This is an advantage over a parametrisation with sums of Breit-Wigner amplitudes, where phases and shape parameters need to be determined for every process separately.

The corresponding line-shape in $s = M^2_{\pi^+\pi^-}$ is then given by

$$R(M^2_{\pi^+\pi^-}) = p_\psi m_B F_\pi(M^2_{\pi^+\pi^-}), \tag{2.55}$$

with p_ψ being the $\psi(2S)$ momentum in the B_s^0 rest frame, and m_B the B_s^0 mass.

This analysis tests the Bonn-Jülich approach by fitting eq. 2.53 to the S-wave in the decay $B_s^0 \to \psi(2S)\pi^+\pi^-$.

2.5 The Decay $B_s^0 \to \psi(2S)\pi^+\pi^-$

The phenomena and methods described in the previous sections were tailored to the properties of the signal decay of this analysis, $B_s^0 \to \psi(2S)\pi^+\pi^-$. The two-pion system in the final state almost always has angular momentum $\ell = 0$ and positive parity, which only allows scalar intermediate hadronic resonances in the decay. These can be studied here without background from pseudoscalar or vector states. In addition, the exotic tetraquark candidate $Z(4430)^\pm$ could couple to the $\psi(2S)\pi^\pm$ systems in this channel.

The branching ratio of $B_s^0 \to \psi(2S)\pi^+\pi^-$ has already been measured to $\mathcal{B} = (6.87 \pm 0.81 \pm 0.65 \pm 0.39) \times 10^{-5}$ in [53], where the uncertainties are statistical, systematic and due to the normalisation. The latter includes both the uncertainty due to f_s/f_d (the ratio of B_s^0 mesons over B^0 mesons produced inside the LHCb acceptance), and due to the branching fraction of the normalisation channel. The subsequent decay of $\psi(2S)$ into two muons has a branching ratio of $\mathcal{B} = (0.77 \pm 0.08)\,\%$ [21]. The leading order Feynman

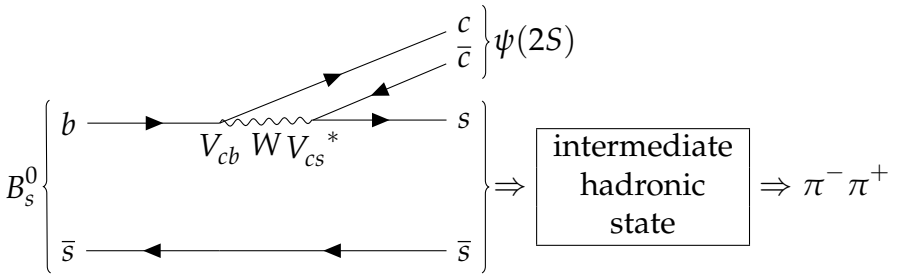

Figure 2.4: Leading Order Feynman diagram of the process $B_s^0 \to \psi(2S)\pi^+\pi^-$.

diagram of $B_s^0 \to \psi(2S)\pi^+\pi^-$ is shown in Fig. 2.4. The b quark in the initial state transforms via the weak charged-current introduced in eq. 2.1 into a $c\bar{c}$, (which then form a $\psi(2S)$ resonance) and an s quark, which pairs up with the $\bar{s}$ spectator quark. This $s\bar{s}$ configuration cannot directly form two pions (their production would be OZI suppressed [54, 55]). Instead they form an intermediate hadronic state, which then strongly decays into two charged pions. However, due to symmetry considerations only specific states are possible:

Because the $s\bar{s}$ pair does not contain u and d quarks, it is in an isospin singlet configuration with $I = 0$. Isospin is conserved in strong interactions (only broken by the current quark masses of u and d), so the isospin of the intermediate state needs to be 0 as well. Under isospin the three pion states π^+, π^0 and π^- form a triplet with $I = 1$ and $I_3 = 1(\pi^+), 0(\pi^0), -1)(\pi^-)$. All strong interactions are therefore invariant under rotations in isospin space (*e.g.* replacing a π^+ with a π^-).

Another symmetry that is conserved in the strong interaction is parity. In this case the parity of the hadronic state has to match that of the two-pion system

$$P(\pi^+\pi^-) = (-1)^{\ell_\pi} P(\pi^+)P(\pi^-) = (-1)^{\ell_\pi}, \tag{2.56}$$

since the parity of pions is -1. ℓ_π is the relative angular momentum between the two pions. Since the two pions are spinless, ℓ_π has to match the spin J of the intermediate hadronic state to conserve angular momentum.

The full wave-function of the two-pion system is

$$\psi = \phi_{\text{isospin}} \eta_{\text{space}}, \tag{2.57}$$

and the symmetry of the spatial wave-function η_{space} corresponds to the parity of the system. From eq. 2.56 one sees that it is determined from ℓ. For even ℓ_π, η_{space} is symmetric, otherwise it is antisymmetric.

The only relevant interaction in the decays of the intermediate hadronic states is the strong interaction, which conserves isospin. The two pions can therefore be considered indistinguishable and, because they are bosons, their combined wave-function ψ needs to be symmetric under exchange of the two.

The isospin wave-function ϕ_{isospin} is symmetric (antisymmetric), if I is even (uneven). Because $I = 0$ in the initial state and isospin is conserved in strong interactions, ϕ_{isospin} is symmetric. This only leaves a positive parity for the hadronic state, which is only possible for even $\ell_\pi = J$.

In summary, only states with $J^P = 0^+, 2^+, 4^+, \ldots$ are allowed due to the symmetry constraints of the strong interaction.

This scheme applies to the decay $B_s^0 \to J/\psi\pi^+\pi^-$ as well: It has the same leading order Feynman diagram as $B_s^0 \to \psi(2S)\pi^+\pi^-$, only the $c\bar{c}$ form a different charmonium resonance. Its $\pi^+\pi^-$ spectrum is therefore expected to consist of the same partial waves and the same resonances should occur there. Its resonance content was established in a previous analysis [56] to be

$f_0(980)$, $f_0(1500)$, $f_0(2020)$[5], $f_2(1270)$, and $f_2(1525)$. They all belong to the allowed J^P assignments derived above and are expected to contribute to the decay $B_s^0 \to \psi(2S)\pi^+\pi^-$ as well.

2.6 Exotic Hadronic Resonances

Apart from the configurations introduced in Sec. 2.3.1 (baryons and mesons), QCD in principal also allows for hadrons that are built from different constituents. Since these configurations could not be established in experiments for decades, they are called "exotic" nowadays. The review on non-$q\bar{q}$ mesons and the review on pentaquarks in [21] give a detailed overview of the current status, while here only some aspects are introduced.

Since gluons can couple not only to quarks, but also to themselves, a bound state formed purely from excited gluons, a so-called glueball, should be possible. The $f_0(1500)$ state that is observed in the studied decay $B_s^0 \to \psi(2S)\pi^+\pi^-$ is actually a long-standing candidate for a glueball configuration [57]. Also configurations, where an excited gluonic component forms a bound state with a $q\bar{q}$ pair is possible. In this case, the quantum numbers of the combined state can have different values than for a classical meson.

As long as the states are colour singlets, it is also possible to combine more than $q\bar{q}$ or qqq into a hadron. This allows for $q\bar{q}q\bar{q}$ states, which can either form compact diquark-diantiquark states (tetraquarks), or molecules, where two mesons are bound together. A first confirmed example of this species is the $\chi_c(3872)$, which was first observed by the Belle collaboration in 2003 [58] and whose quantum numbers were later established by the LHCb collaboration to be 1^{++} [59]. Its exotic quark content is determined from its mass and decay properties, that cannot be explained with conventional quark configurations.

For baryons the exotic combination $qqqq\bar{q}$ has been observed in the form of the pentaquark states decaying into $J/\psi\, p$ in the $\Lambda_b^0 \to J/\psi\, pK^-$ channel by the LHCb experiment in 2015 [11, 60]. Since the pentaquark states are electrically charged, but also decay strongly to charmonia, their quark content has to be $c\bar{c}uud$.

[5] The other $I = 0, J^P = 0^+$ state $f_0(500)$ (the sigma meson mentioned in Sec. 2.3.1) has a negligible $s\bar{s}$ content and therefore does not contribute significantly to that decay.

2.6.1 $Z(4430)^+$

The $Z(4430)^+$ state is another tetraquark candidate, which has first been observed by the Belle collaboration in the $\psi(2S)\pi^+$ mass spectrum of $B^0 \to \psi(2S)K^+\pi^-$ decays [61, 62] and was later confirmed and its quantum numbers measured to 1^+ by the LHCb collaboration [63–65]. Due to its large width of 172 MeV (see Fig. 2.5 left) and electric charge, it is considered a smoking gun signal for a state with quark content $c\bar{c}u\bar{d}$. Its measured phase motion is shown to be compatible with a Breit-Wigner resonance on the right side of Fig. 2.5, which further substantiates its resonant character. A simple Breit-Wigner parametrisation creates a circle in the Argand diagram, since it does not account for any inelasticities (see the introduction of the Argand diagram in Sec. 2.4).

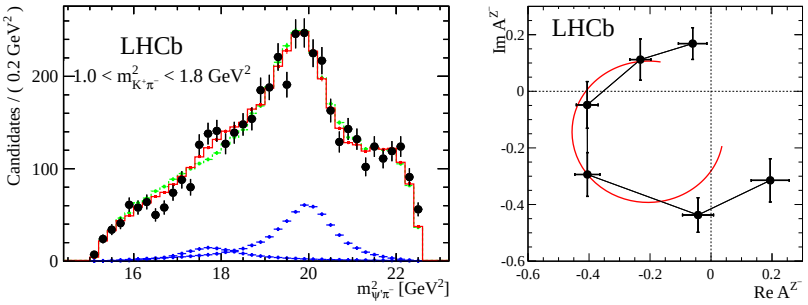

Figure 2.5: Left: $m^2_{\psi\pi^+}$ distribution of $B^0 \to \psi(2S)K^+\pi^-$ data from LHCb Run I data inside $m^2_{K^-\pi^+}$ mass range, where no strong K^{*0} contribution is expected. The green (red) curve is the projection of a 4-dimensional amplitude fit to the data including the known K^* resonances and one (two) additional Z^- resonances. The blue curves show the Z^- contributions from the fit with two additional resonances, where the high-mass one corresponds to the $Z(4430)^+$. Taken from [63]. Right: Argand diagram of the fitted values of $Z(4430)^+$ amplitude in six bins of $M(\psi(2S)\pi)$ increasing counter-clockwise shown as connected error bars. Breit-Wigner prediction for masses and widths as extracted for the $Z(4430)^+$ shown as red circle. Taken from [63].

Its interpretations range from hadrocharmonium ($c\bar{c}$ core with $u\bar{d}$ "shell") to a $\bar{D}_1 D^*$ molecule or compact tetraquark. Another approach explains it with a non-resonant cross-channel effect of triangle singularities from open charm states [66] or a triangle with $K^*\pi^+$ and Y(4260) (see Fig. 2.6 left) [67].

Triangle singularities occur in a reaction, if all three intermediate particles are on-shell. Even though the shape and phase motion of the corresponding amplitude can mimic that of a resonance (see Fig. 2.6 right), in the triangle scenario the observed peak in the $\psi(2S)\pi^+$ mass would not correspond to a

genuine resonance, but a kinematic effect: While real resonances are related to poles of the S-matrix, the singularity from a triangle diagram is logarithmic [68]. By definition it is therefore only a kinematic effect and not a genuine resonance. More on resonances can be found in Sec. 2.4.

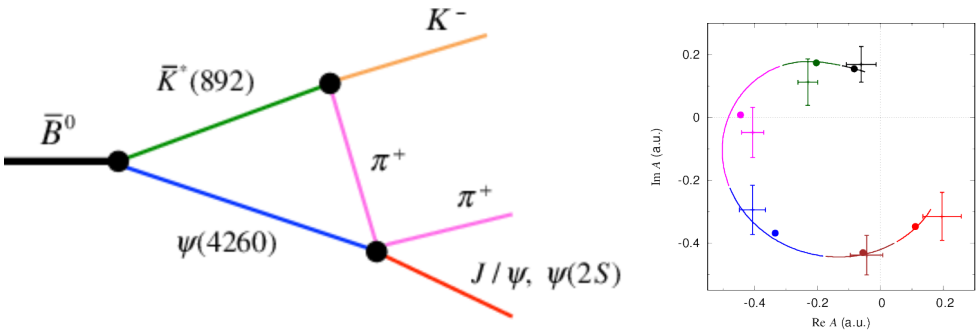

Figure 2.6: Left: Triangle diagram generating a singularity, which could be responsible for the observed shape and phase motion of the tetraquark candidate $Z(4430)^+$ in $B^0 \to \psi(2S)K^+\pi^-$. Taken from [67]. Right: Argand diagram of triangle singularity from [67]. The points with error bars correspond to the phase motion measure by the LHCb collaboration in [63]. The coloured lines show the phase motion of the triangle singularity through the $\psi(2S)\pi^+$ mass. The full dots show the average values of the line segment of the same colour. Taken from [67].

The triangle singularities are very sensitive to the masses of the intermediate states in the triangle and also rely on strong couplings between the intermediate and outgoing states. The diagram in Fig. 2.6 does not correspond to a usual Feynman diagram, it merely depicts the effective hadronic interactions taking place. It merely depicts a rescattering process, where the B^0 decays into a K^* and $Y(4260)$, which then decay further, $K^{*0} \to K^-\pi^+$ and $Y(4260) \to \psi(2S)\pi^+\pi^-$ exchanging one pion between them. The masses of the involved particles make it possible that all intermediate particles are on-shell, which leads to a logarithmic singularity in the scattering amplitude. The size of this effect depends on the values of the couplings at each vertex though. Their values are not well constrained currently, which makes it impossible to predict the effect size accurately.

Since this kinematical singularity only occurs for this process, where the masses have exactly the correct values, the triangle hypothesis can be tested by searching for a peak in the $\psi(2S)\pi^+$ spectrum in other channels and processes, where no triangle singularity can occur.

This is the case in $B_s^0 \to \psi(2S)\pi^+\pi^-$, since there is no K^- in the final state and K^* does not couple to $\pi^+\pi^-$. Exchanging the K^- and π^+ in the

diagram of Fig. 2.6 does not work either, since the Y(4260) does not couple to $\psi(2S)K^-\pi^+$. Even if other triangle diagrams with different intermediate states were possible, the shape in the $\psi(2S)\pi^+$ spectrum would not coincide with the one in $B^0 \to \psi(2S)K^+\pi^-$. An observation of $Z(4430)^+$ in $B_s^0 \to \psi(2S)\pi^+\pi^-$ would therefore be a strong sign against the triangle interpretation and in favour of a genuine resonant state.

A non-observation of the $Z(4430)^+$ in the decay $B_s^0 \to \psi(2S)\pi^+\pi^-$ on the other hand could not exclude a particular model, because the necessary hadronisation of the $s\bar{s}$ into an isoscalar resonance (see Fig. 2.4 in Sec. 2.5) could also suppress the coupling to the $Z(4430)^+$. Figuratively the π^+ and the $\psi(2S)$ can only form a resonance after the hadronic intermediate state formed from the $s\bar{s}$ source has decayed into two pions. Before that the necessary valence quark content for the $Z(4430)^+$ is just not available. A re-scattering into π^- + $Z(4430)^+$ could then be suppressed significantly with respect to the conventional $\psi(2S)$ + isoscalar modes and it would be hard to detect given the current data sample size.

Due to the complicated production process, no explicit predictions exist for the expected fractions of $Z(4430)^+$ in $B_s^0 \to \psi(2S)\pi^+\pi^-$ so far for the molecular or tetraquark models. There are ongoing efforts on the theory side though, which should result in publications on the expected fraction in the near future.

3 The LHCb experiment

This chapter introduces the LHCb detector at the Large Hadron Collider (LHC) located at the research facility CERN. In Sec. 3.1 the proton collider ring is introduced, while the LHCb detector and its subcomponents are described in Sec. 3.2. The chapter focuses on the components relevant for the analysis of the B mesons that are studied in this work.

3.1 The LHC Machine

The LHC is the most powerful proton collider ring to date with a circumference of 27 km located 50 to 175 m underground in the Geneva area. It receives protons with energies of 450 GeV from the SPS pre-accelerator and accelerates them to up to 6.5 TeV, while bending them inside the ring using 1232 superconducting dipole magnets. The protons are travelling in (up to) 2×2808 bunches with 1.15×10^{11} particles each, one half travelling clockwise, the other counter-clockwise in the ring. They can collide in 4 interaction regions, where detectors are situated in order to measure the particles that were produced in the collisions [69].

The ATLAS [70], CMS [71], ALICE [72] and LHCb [73] experiments are located directly around these interaction points, while other, smaller experiments lie further away from them (LHCf [74], MOEDAL [75], TOTEM [76], FASER [77]). These specialised experiments try to measure those particles, that were produced almost collinear with the beam or those hypothetical, weakly interacting particles, which only decay after a certain distance into particles that can actually interact with the detector.

The machine is also used to accelerate bunches of lead ions to collide them with either protons or other lead ions.

The collider is currently upgraded and will resume collisions in 2022 with an increased luminosity and a possible increase in proton energy to 7 TeV.

3.2 The LHCb Detector

The LHCb detector is a single arm forward spectrometer specifically designed to reconstruct decays of hadrons with heavy quarks with a high precision [73]. At the pp collision energies of the LHC these particles are mainly produced via gluon-gluon fusion, where the two gluons usually carry largely different fractions of the proton momentum. Due to this asymmetry the newly formed particles are usually boosted along the beam direction. The LHCb detector is therefore built to cover exactly this region. A quarter of the $b\bar{b}$ pairs created in the interaction are produced within its acceptance of $1.8 < \eta < 4.9$, which only corresponds to 4 % of solid angle.

Fig. 3.1 shows a schematic of the detector. Even though it covers a small solid angle, it follows the structure of general purpose particle detectors: A vertex detector sits around the interaction region of the beams to reconstruct the interaction vertex via the tracks of charged particles produced in the collision (VELO). These fly further into the detector, where they are bent horizontally with a strong magnet. Their paths are reconstructed with tracking detectors (TT, T1-T3) and from the kink they get from the magnetic field, their momenta are inferred. The particle hypotheses of the tracks are determined with Cherenkov detectors (RICH1, RICH2). Further outside are the electromagnetic calorimeter (ECAL), which measures the energy of electrons and photons (stopping photons and electrons in the process), then the hadronic calorimeter, which measures the energy of all hadronically interacting particles (HCAL), which were not stopped in the ECAL. The remaining particles behind the calorimeters are then neutrinos and muons. The former have negligible interaction probability with the detector and traverse it unseen. The latter are measured with dedicated muon chambers (M1-M5).

In each proton-proton collision at $\sqrt{s} = 13$ TeV hundreds of charged particles are produced inside the LHCb acceptance. With the maximum luminosity the LHC can provide, tens of collisions take place at each bunch crossing. The resulting flux of particles would lead to large occupancies in the detector components and degrade the performance of the track reconstruction. The instantaneous luminosity $\mathcal{L}$ in the LHCb interaction region is therefore 20 times lower than the maximum luminosity the LHC can achieve. This reduces the average number of simultaneous interactions at each bunch crossing to approximately 1.6 in typical running conditions.

The reduced $\mathcal{L}$ has an additional advantage: During the $\approx$12 h that the proton bunches circulate in the LHC, the peak $\mathcal{L}$ decreases, because the bunches

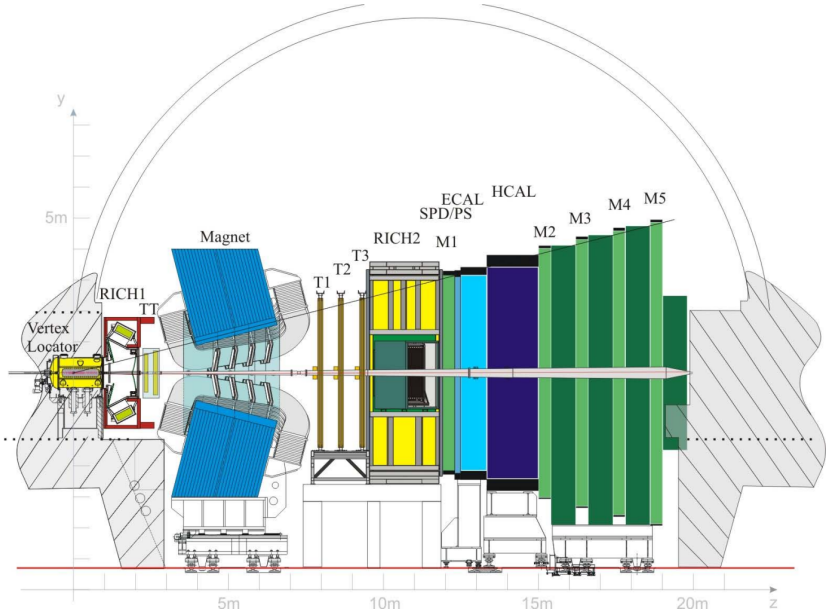

Figure 3.1: Schematic of the LHCb detector. Taken from [78].

become smaller and their shape gets distorted from the repeated interactions. A detector measuring at peak $\mathcal{L}$ has therefore a large occupancy at the beginning of fills, when freshly prepared bunches circulate in the LHC, and a much lower towards the end. With the reduced instantaneous luminosity in the LHCb interaction region, it is possible to keep the measurement conditions constant throughout the full fill, which makes the data easier to calibrate.

3.2.1 Vertexing and Tracking Detectors

B-mesons have a lifetime of around 10^{-12}s before they decay weakly. Due to their large boost inside the LHCb acceptance, this means that most of them travel a measurable distance of a few mm inside the detector. Since lifetime measurements are a core part of the physics programme at LHCb, the distance between the primary vertex and the secondary vertex (where the B-meson decayed) need to be determined very precisely. The decay time of the B-meson can then be determined from the measured flight distance and the reconstructed 4-momentum.

The **VELO** depicted in Fig. 3.2 is a silicon detector in $R-\phi$ geometry installed in 2×21 half-circle segments around the interaction point. Charged particles flying through silicon detectors create electron-hole pairs, which are amplified and read out. Since the position of each detector channel is known, the flight path of each particle can be reconstructed from the channels that measured a signal. After the beam is ramped up to the full energy and

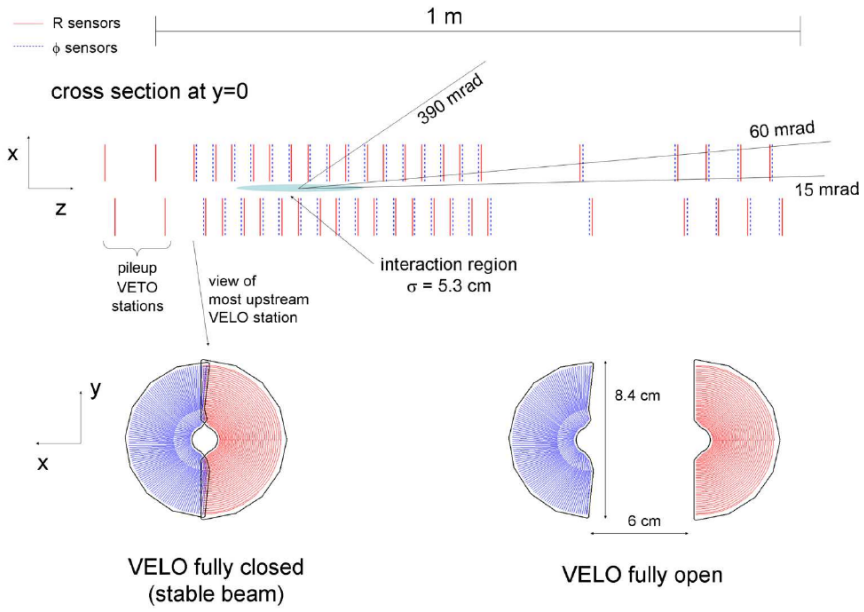

Figure 3.2: Schematic of the VELO detector. Taken from [73].

stable, the segments are moved within 8 mm of the beam lines in a staggered configuration to measure the trajectory of the charged particles close to the primary vertex (PV) and decay vertices. The sensors are 220 µm thin to reduce scattering inside the detector material, which would deteriorate the momentum measurement.

The spatial resolution on the measurement of the PV transverse (parallel) to the beam direction varies from 35 (270) µm for interactions with 5 tracks to 10 (60) µm for those with 40 tracks. The resolution in beam-direction is worse, because the detector layout makes it insensitive to particles emitted under a large angle relative to the beam direction.

The **Silicon Tracker** (ST) consists of the **Tracker Turicensis** (TT) and the **Inner Tracker** (IT). Their sensors use the same silicon microstrip technology to reconstruct track stubs near the beam-pipe, where the particle density is high and a high granularity is needed. Their strips are 183 µm wide and only 300 µm thick to reduce multiple scattering.

The TT consists of four layers of sensors as shown in Fig. 3.3. The two outer ones are vertically orientated, while the middle ones are tilted ±5 to achieve resolution in the vertical direction as well. This subdetector allows to reconstruct low-momentum tracks ($p_{\mathrm{T}} < 2\,\mathrm{GeV}$), which are bent out of the acceptance by the magnetic field, before they reach the tracking stations further downstream. Additionally it makes reconstruction of long-lived particles like the K^0_{S} possible ($\tau = (8.954 \pm 0.004) \times 10^{-11} s$), which are boosted enough so that a large portion of them just traverse the VELO, leaving no reconstructible

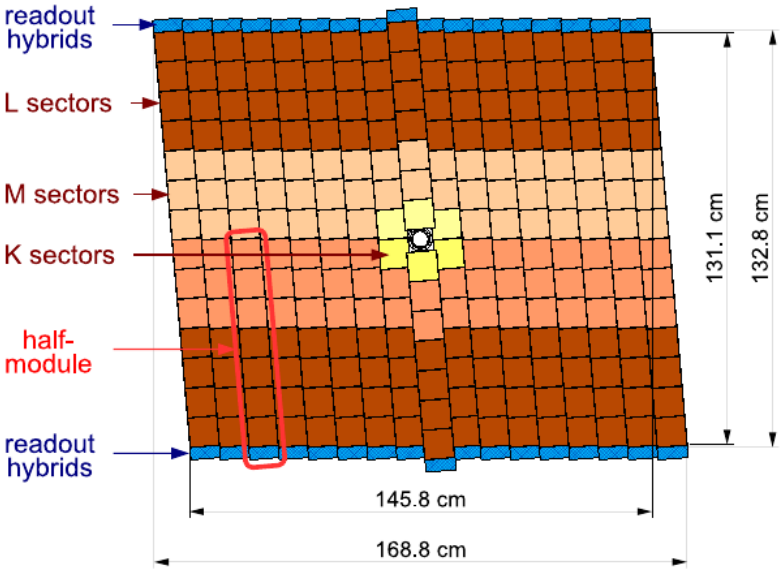

Figure 3.3: Schematic of a TT detector layer. Taken from [73].

trace, before they decay into pions.

The last part of the tracking system consists of three tracking stations behind the magnet shown in Fig. 3.4. Each of the stations consists of 4 layers in the same geometry as the TT. The IT covers the inner part of these layers, while the outer part, where particle fluxes are lower, is covered with a straw gas tube detector called **Outer Tracker** (OT). It is made up from cylindrical tubes with diameters of 4.9 mm and a concentric anode wire inside. Each tube acts as drift tube: A high voltage is put between wire and the tube's inner surface. When a charged particle flies through, it produces ions and electrons in the gas mixture, which get accelerated in the electric field. From the electron drift time towards the anode the hit position in a circle around the wire can be determined with a precision of 200 µm and ambiguities are resolved by combination of multiple hits.

The magnetic field for the measurement of the particle momenta is provided by a normal-conducting dipole magnet with an integrated field strength $\int \vec{B} d\vec{s} = 4\text{Tm}$ bending charged particles in the horizontal plane. Inhomogeneities in the detector material and the magnetic field could create asymmetries in the track reconstruction efficiency that mimic asymmetries in the measured particle decays. To determine these artificial asymmetries, the magnet polarity is switched regularly.

If a charged particle leaves hits in all sub-detectors of the tracking system (LHCb jargon: Long Track) the momenta are determined using a Kalman filter approach, which achieves a relative uncertainty between 0.5 % for momenta

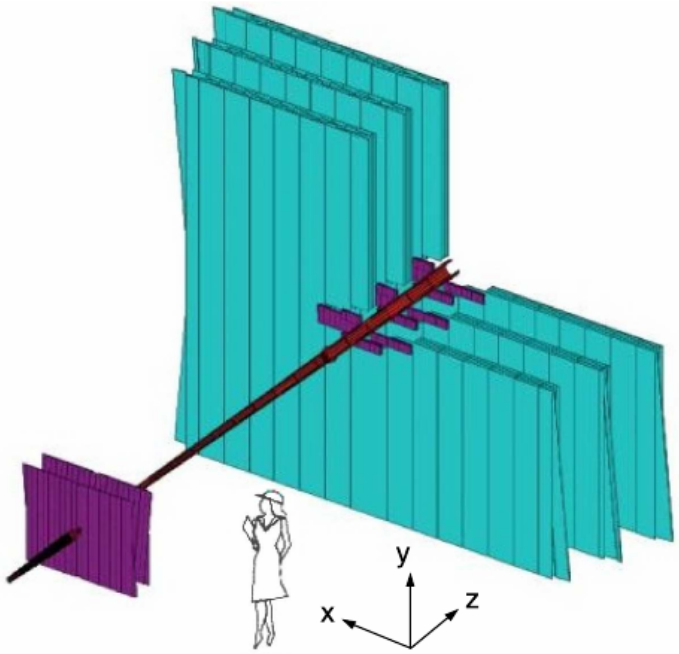

Figure 3.4: Schematic of the tracking stations with the TT in front. Purple for the silicon detectors TT and IT, cyan for OT. Taken from [73].

below 20 GeV and 1.0 % for momenta around 200 GeV. All tracks used in this analysis belong to this category.

3.2.2 Particle Identification

Particle identification (PID) information is provided by several sub-detectors at LHCb: Two **Ring Imaging Cherenkov** (RICH) detectors are in front and behind the magnet. They use the Cherenkov light emitted by charged particles that travel in a gas mixture of refractive index n with a velocity v above the local light speed. From the the angle θ_c under which the photons are emitted, the speed of the particle is calculated from

$$\cos(\theta_c) = \frac{1}{nv/c}, \tag{3.1}$$

where c stands for the speed of light. Together with the momentum of the corresponding track, a mass hypothesis is deduced. If the velocity is too small, no Cherenkov light is emitted, if it is too large, the angles for different mass hypotheses do not differ enough to be distinguishable. To provide good separation between particle species over a large momentum range, the refractive index in RICH1 is 1.0015 for low momentum particle separation

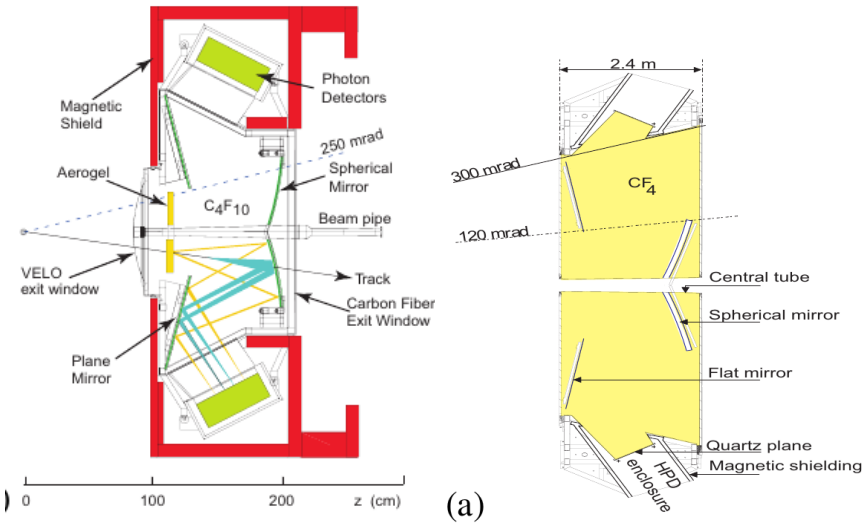

Figure 3.5: Left: Schematic of the RICH1 detector. The aerogel depicted here was only used during Run 1 of the LHC. Right: Schematic of the RICH2 detector. Taken from [73].

with $2\,\text{GeV} < p < 60\,\text{GeV}$, while in RICH2 $n = 1.0005$, which covers particles with momenta $15\,\text{GeV} < p < 100\,\text{GeV}$ [79]. The photons are reflected by a mirror system and detected with arrays of Pixel Hybrid Photon Detectors (HPDs) as shown in Fig. 3.5.

The calorimetry system is used to distinguish between high-p_T hadrons, photons and electrons and to measure their energy. It consists of three parts: The **Scintillating Pad** and **Preshower** detectors (SPD/PS), which consist of scintillators and lead absorbers, the **Electromagnetic Calorimeter** (ECAL) and the **Hadronic Calorimeter** (HCAL). All are segmented into regions with higher granularity towards the beam-pipe, where the particle flux is two orders of magnitude larger than at the outskirts of the LHCb acceptance.

The SPD is used to discriminate charged particles against photons, because only charged particles deposit energy in it. The PS detector, which lies behind a layer of lead that induces electromagnetic showers, uses the fact that shower shapes for electrons and light hadrons differ.

The ECAL uses a shashlik structure of 2 mm lead and 4 mm plastic scintillator layers corresponding to 25 radiation lengths. The scintillator photons from electromagnetic showers are detected with photomultiplier tubes and the shower energies are measured with a resolution of $\sigma(E)/E = 10\%/\sqrt{E[\,\text{GeV}} \oplus 1\%$.

The HCAL measures the hadronic showers of the remaining particles using 16 mm iron layers interleaved with 4 mm of scintillators summing up to 5.6 nuclear interaction lengths. The energy resolution for hadronic showers

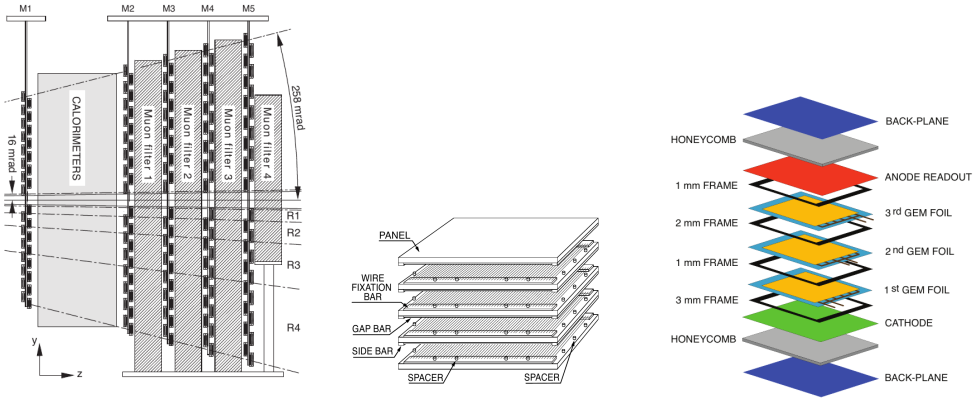

Figure 3.6: Left: Schematic of the Muon stations. Middle: Sketch of MWPC. Right: Sketch of GEM. Taken from [73].

is $\sigma(E)/E = 69\%/\sqrt{E}/\,\text{GeV} \oplus 9\%$.

The muon identification is done with 5 **Muon Stations**, one in front of the calorimeters, the others behind. M2-M5 are separated by 80 cm thick iron walls, which are supposed to absorb all particles, that are not minimum ionizing muons (see Fig. 3.6 left). Additionally, M1-M3 can measure p_T of muon candidates to 20 % resolution without using the tracking system. This forms the basis of the L0 muon trigger described in Sec. 3.2.3.

The stations are instrumented with multi-wire proportional chambers (MWPC), where charged particles ionize the gas mixture around anode wires. Due to high electric fields this sets off a cascade of charged particles, which are collected on the nearest wire (see Fig. 3.6 middle). In the most central region of M1 triple gas electron multiplier (GEM) detectors are used to cope with the large particle flux of 500 kHz/ cm^2. A GEM uses planar anode and cathode layers with 3 separated layers of copper-clad dielectric Kapton foils in between. Small holes are etched into the Kapton and the inside of the detector is filled with a mixture of Argon, CO_2 and CF_4 (see Fig. 3.6 right). Electrons from ionisation of the gas are accelerated by the strong fields inside the Kapton foil leading to a localized avalanche that is amplified through the next foils and can be read out on the anode.

3.2.3 Trigger System

The full detector cannot be read out at the 40 MHz bunch crossing rate of the LHC. In order to select interesting events (*e.g.* those containing decays of heavy quarks), a hardware trigger (L0) prompts the readout for those events with either high transverse momentum tracks in the muon stations or tracks that deposit large transverse energy in the calorimeters indicating decays of

heavy hadrons, where large p_T is expected. This reduces the data rate to about 1 MHz, which is analysed further in the two stages of the software trigger HLT1 and HLT2 [80].

HLT1 performs a partial event reconstruction and selects events based on displaced tracks and vertices, which can come from the weak decay of heavy particles. Additionally events with dimuons are triggered, because they often originate from heavy neutral resonances and have very little background from misidentified pions. Since Run 2 the full detector information for these events is buffered to 10 PB of disk space. The buffer is used to perform an online detector calibration and alignment between LHC fills, so that in the HLT2 stage an offline-like resolution is available [80]. Here hundreds of exclusive as well as inclusive lines are selecting *e.g.* prompt or detached muons or generic n-body decays into charged particles. Its output was 3.5 kHz in Run 1 and was extended to 12.5 kHz in Run 2. The data of this output are recorded.

3.2.4 LHCb Data Sample

The LHC operated at $\sqrt{s}$ of 7 TeV, 8 TeV, and 13 TeV in 2011, 2012, and 2015-2018 respectively. At these energies the LHCb experiment accumulated pp collision data corresponding to integrated luminosities of 1 fb^{-1}, 2 fb^{-1} and 6 fb^{-1}.

The data reconstruction software is based on the GAUDI software [81], which was developed as general purpose data processing framework for HEP experiments. It steers all reconstruction steps that reside as individual projects within the LHCb software.

Simulated samples for the channels $B_s^0 \to \psi(2S)\pi^+\pi^-$ and $B^0 \to \psi(2S)K^+\pi^-$ are needed to describe the efficiency of the selection and the effects of the detector resolution on the measured data. The proton-proton collisions are simulated for each beam condition using PYTHIA [82], [83]. EVTGEN [84] describes the decay of the B-mesons from the collision, while PHOTOS [85] adds the final state radiation of the daughter particles. GEANT4 [86], [87] describes the response of the detector to the simulated particles, which is then digitized using BOOLE [88]. After this step the simulated data resembles the output of the real LHCb detector and is processed from the trigger step onwards just like real data. The simulated data samples do not describe the data accurately in all important variables though. In particular the overall event multiplicity is underestimated in simulated events. In order to decrease the discrepancy between data and

simulation, the latter is therefore weighted in the per-event track multiplicity and the transverse momentum of the B_s^0 meson using data from the channel $B^0 \to \psi(2S)K^+\pi^-$. Additional weights are applied in order to correct the mismodelling of the tracking efficiency [89] and the PID response is corrected with a data-driven approach [90]. After these steps simulated events can directly be used for the estimation of efficiency and acceptance effects.

4 Selection of $B_s^0 \to \psi(2S)\pi^+\pi^-$ decays

This chapter describes the selection steps that were applied to the reconstructed LHCb data (see Sec. 3.2.4) to select the signal decays $B_s^0 \to \psi(2S)\pi^+\pi^-$ and the decays of $B^0 \to \psi(2S)K^+\pi^-$, which serves as control channel. The absolute values for the cuts are listed in App. A.

4.1 Trigger Selection

The $\psi(2S)$ in the signal decay $B_s^0 \to \psi(2S)\pi^+\pi^-$ is reconstructed from its decay into two muons, which has a branching ratio of $\mathcal{B} = (0.77 \pm 0.08)\,\%$ [21]. Even though it is much smaller than *e.g.* for the decay into $\chi_c\gamma$, the final state with just two muons is superior, because the muons leave a clean signature in the detector, which can be easily reconstructed and distinguished from possible backgrounds.

This muon signature is also the center of the trigger strategy. The Muon chambers introduced in Sec. 3.2.2 can almost exclusively be reached by minimum ionizing muons. If they detect a particle with a large transverse momentum $p_\mathrm{T} > 1.5\,\mathrm{GeV}$ (single muon), or two particles with at least 1.3 GeV $< \sqrt{p_{T1} \times p_{T2}}$ (double muon), the full detector is read out. The exact values for the thresholds were changed throughout the years depending on the LHC running conditions and free capacity of the disk buffer.

The lifetimes of B_s^0 and B^0 are approximately $1.5 \times 10^{-12}\,s$ each [21]. They are produced with a large boost in beam direction, which allows them to travel an average distance of 1-10 mm inside the VELO, before they decay. On the other hand each proton-proton collision produces dozens of particles in strong and electromagnetic interactions, which can decay promptly into muons. These muons need to be filtered out in order to keep the output rate of the detector low enough to process the events before the disk buffer fills up (see Sec. 3.2.3). The selection exploits the different origins of the muons at the

HLT1 stage via requiring that the muon tracks have a significant separation from the PV.

At the HLT2 level the full reconstruction can be run and the trigger has access to higher level features from the fits of secondary vertices. At this stage the two muons are combined into a $\psi(2S)$ candidate. The event is kept, if they form a good vertex, that is well separated from the PV and if additionally the invariant mass of the $\psi(2S)$ candidate is within 100 MeV of the PDG average.

4.2 Cut Based Selection

In the next stage, $\psi(2S)$ candidates are combined with two additional tracks and the quality of the combined vertex is tested. Each of these tracks needs to have a low probability to be constructed from a combination of random hits in the detector. This is done with a cut on the output of a shallow neural network[1], which was trained with variables related to the quality of the track fit. Additionally they need to be separated from the PV to avoid combinations of randomly associated tracks coming directly from the proton-proton interaction. This is accomplished with a cut on the impact parameter significance IP_{χ^2}. IP is the closest distance from the track to the PV and for IP_{χ^2} that value is divided by the χ^2 of a fit of the PV with the track included. For tracks that originate from the PV this ratio is typically small.

The two non-muonic tracks need to have signals in the RICH detectors and their signatures in the detector need to be compatible with pions. Pions from the underlying event have on average low transverse momenta, because they are produced in strong interactions, whose cross sections decrease rapidly for large momentum transfers (see Sec. 2.3). The final state particles of B meson decays on the other hand have on average a larger p_T due to the large rest-mass of the B meson of around 5 GeV. This is exploited allowing only pion candidates with $p_\mathrm{T} > 200$ MeV in the further selection.

The decay vertex position (secondary vertex SV) is extracted from a fit of the intersection of the four tracks and only candidates with a good fit quality are kept. The so-formed B meson candidate must have a momentum vector, which agrees within 1.4° to the vector from the PV to the SV, and a fit for the

[1] A shallow network only has a few layers between input and output layers. Usually the nodes are fully connected in this case. The term is used to distinguish them from the deep neural networks used *e.g.* in image recognition or fast simulation approaches.

intersection of the PV with the 4 signal candidate tracks must have a large χ^2[22]. Additionally, the IP_{χ^2} of the B candidate is required to be small.

Its invariant mass must lie between 5150 and 5500 MeV (the PDG average for the B_s^0 meson mass is 5368 MeV). The invariant mass of the B candidate $M(\psi(2S)\pi^+\pi^-)$ is not calculated directly from the 4-momenta of the final state particles, instead the full decay chain is fitted to improve the resolution [91]. The fit recalculates the 4-momenta of the final state tracks with external constraints from the position of the primary vertex and secondary vertex, and the mass hypothesis for the narrow $\psi(2S)$.

4.3 Vetos for Physical Backgrounds

Several known B decays have a topology similar to the signal decay.

If *e.g.* a Kaon is misidentified as a pion $K^+ \to \pi^+$, then the decay $B^0 \to \psi(2S)K^+\pi^-$ can be reconstructed as $B_s^0 \to \psi(2S)\pi^+\pi^-$. Both decays have similar branching ratios, but $B^0 \to \psi(2S)K^+\pi^-$ is much more abundant than the signal decay (the chance of a B_s^0 being produced in the pp collision is about $1/4$ of a B^0). This background can therefore obscure the signal component in data, even if the misidentification probability is small.

Another physical background comes from $B^+ \to \psi(2S)\pi^+$, where the three tracks by chance form a good vertex with a random pion from the underlying event (everything produced in the pp collision that is not part of the decay chain in question). This decay is Cabibbo suppressed (its leading order Feynman diagram contains the off-diagonal V_{CKM} matrix elements V_{cd} and V_{cb}), but due to the larger production rate of B^+ mesons it needs to be addressed.

The third background in this category comes from $B^+ \to \psi(2S)K^+$, where an unrelated pion is added like for $B^+ \to \psi(2S)\pi^+$, but additionally the Kaon is misidentified as a pion $K^+ \to \pi^+$. This decay is Cabibbo favoured (it only contains one off-diagonal V_{CKM} element from the b-quark decay), which outweighs the suppression from the misidentification.

No backgrounds from muon misidentification are considered, because very few other particles can penetrate the detector into the muon stations and create a muon-like signal [3]. The additional cut on the dimuon candidate mass around the $\psi(2S)$ mass reduces these backgrounds to a negligible level.

[2] This cut is correlated to the aforementioned angular cut for signal decays, but not so much for combinatoric background making the combination of both in the MVA very efficient.

[3] Decays of pions and kaons in flight into muons are the most dangerous ones, but most of them are vetoed by the signature they have left in the RICH detectors before their decay.

The physical backgrounds from misidentified pions are vetoed explicitly via recalculating the invariant mass of the track combinations under the hypothesis that they actually come from the wrong decay channel. To veto $B^0 \to \psi(2S)K^+\pi^-$ decays, the invariant mass of the B candidate is recalculated with a Kaon mass hypothesis for the track associated to a π^+. If the B candidate mass lies within 25 MeV of the nominal B^0 mass for one of the hypotheses, the event is vetoed. Fig. 4.1 shows the mass distribution of the B candidate under the K^- and π^- hypotheses. It shows the data sample after the full selection including the MVA cut introduced in Sec. 4.4 to make the effect of the veto more visible.

The same procedure is applied for $B^+ \to \psi(2S)K^+$: The mass hypothesis of the π^+ track is changed to K^+ and the invariant mass of the $\psi(2S)$ and the K^+ candidate is calculated. If the mass lies within 30 MeV of the B^+ nominal mass, a veto is applied. To veto $B^+ \to \psi(2S)\pi^+$, the invariant mass of $\psi(2S)$ and π^+ is calculated events within a 30 MeV region around the B^+ nominal mass removed. The same vetos are applied to the charge-conjugated modes as well.

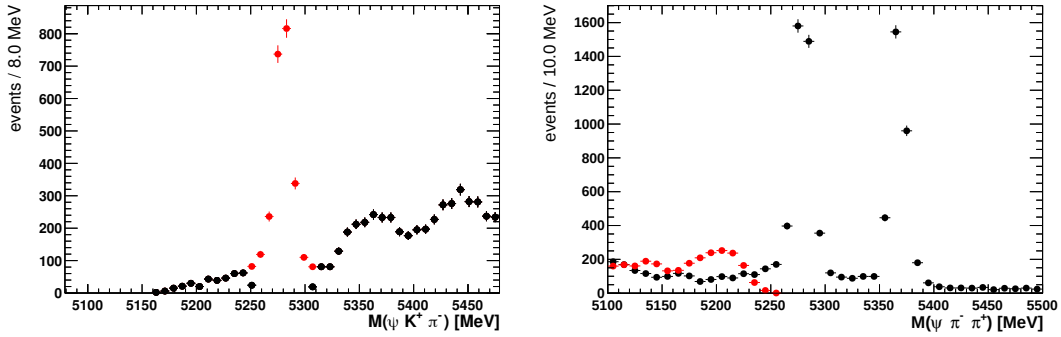

Figure 4.1: Data after full selection. Left: invariant mass of B candidate under $\psi(2S)K^+\pi^-$-hypothesis. Right: invariant mass of B candidate under $\psi(2S)\pi^+\pi^-$-hypothesis. The two peaks in the black histogram correspond to $B^0 \to \psi(2S)\pi^+\pi^-$ (left) and $B_s^0 \to \psi(2S)\pi^+\pi^-$ (right). Red Histograms: Events, which are vetoed to suppress contributions from the decay $B^0 \to \psi(2S)K^+\pi^-$. Black Histograms: Remaining events after veto. The red peak consists mainly of $B^0 \to \psi(2S)K^+\pi^-$ events.

4.4 Multivariate Selection

After the trigger stage and the suppression of physical backgrounds, the remaining background events in the data sample are mainly combinatoric

background consisting of a real $\psi(2S)$ decay into two muons that is associated with two unrelated tracks identified as pions. The final selection step discriminates the signal decay $B_s^0 \to \psi(2S)\pi^+\pi^-$ from these events with a multivariate analysis (MVA) classifier. This classifier "learns" the multidimensional distribution of input features for signal and background events and transforms a point in feature-space into a single output value, which represents, how signal-like that point is. In contrast to a simple cut-based selection, MVA classifiers can exploit nonlinear correlations between variables to achieve a better separation of signal and background.

A Boosted Decision Tree (BDT) [92] from the TMVA package [93] is chosen for this step, because it is a well-tested method in HEP, provides good classification performance without extensive tuning of modelling parameters, and can exploit non-linear correlations of input parameters. Two flavours of boosting are tested: Adaptive Boosting [94] and Gradient Boosting [95]. A shallow artificial neural network (ANN) is also tested, but its performance is slightly worse.

A BDT consists of an ensemble of decision trees. Each of these trees is a set of nodes, each of which represents a rectangular cut in one variable. For each node the optimal cut value is determined with a metric, often the Gini-index [96], which measures the separation $p(1-p)$ for p the purity of the partition after the cut. Ideally this is purely signal or purely background. Since each node is optimized separately, it only optimizes the classification locally, making it a "weak learner". In order to create a "strong learner", so-called boosting can be applied. First a decision tree is optimized (vulgo "trained") on a labelled data set, where for each event it is known, whether it belongs to the signal or background category. The events are given weights according to the correctness of their classification by the initial tree and misidentified events are given a larger weight in the optimisation of the next tree. Repeatedly misidentified events get larger and larger weights making their correct classification more favourable for the minimisation of the separation metric. This approach is taken for the BDT with Adaptive boost. Outlying events are problematic in this approach, because their weights can grow very large, which can lead to a reduced performance of the general classification, because the classifier focuses on the outliers.

The other method used here extends this approach to a gradient method, which is less affected by outliers in training data [97]. It uses a different loss function and applies a stochastic gradient descent algorithm to optimize the loss for each tree. In order to reduce overfitting of the classifier, a different

random subsample of the training events is used at each iteration.

The MVA methods are trained with simulated $B_s^0 \to \psi(2S)\pi^+\pi^-$ and $B^0 \to \psi(2S)\pi^+\pi^-$ events as signal proxy. The two decay channels have the same final state and very similar kinematic distributions, since the mass difference between B_s^0 and B^0 is only around 87 MeV. Using both simulated samples doubles the number of training events and makes the outcome of the classifier more stable. The background proxy consists of measured events from the upper sideband, where $M(\psi(2S)\pi^+\pi^-) \in [5450, 6000]$. This region is chosen, because it does not contain real B hadron decays, but only combinatoric background. In addition it does not overlap with the mass range, that is later used in the fit to extract the signal component in Sec. 6. This avoids possible biases in the MVA output cut optimisation [4].

The variables, which are used as input for the MVA are:

- p_T (B): The transverse component of the momentum of the B candidate. Combinatoric background events are mostly random combinations of $\psi(2S)$ decays with tracks from soft processes in the pp interaction. The reconstructed p_T of these combinations averages out due to their missing correlation with each other. It is therefore smaller than for signal, where all tracks come from the same B-meson that has momentum in a particular direction.

- FD $\chi^2(B^0_{(s)})$: The flight distance (FD) χ^2 is the significance of a non-zero separation between the PV and SV: $FD\chi^2(B^0_{(s)}) = \text{FD}/\sigma(\text{FD})$. Combinatoric background comes mainly from tracks created in the PV, while the B mesons fly on average a few mm before they decay and produce charged tracks.

- Vertex χ^2: The χ^2 of the fit of the combined vertex of the 4 tracks is larger for randomly combined tracks than it is for signal, where they really come from the same spot

- IPχ^2(B): IPχ^2 is the significance of the impact parameter of the B flight direction with the PV. Since the B meson comes from the PV, it peaks around 0 for signal, while for combinatoric background events the combined track momenta do not point somewhere particular

[4] In the training the classifier could "learn" the statistical fluctuations in the distributions of its input variables and show a better separation power on those events (this effect is called overtraining/overfitting). It would distort the shape of the combinatoric background distribution in the later fit to $M(\psi(2S)\pi^+\pi^-)$, which could bias the extracted signal sample.

- $\min(IP_{\chi^2}(X^+), IP_{\chi^2}(X^-))$: $IP_{\chi^2}(X)$ is the significance of a non-zero value of IP for a track X. In combinatoric background typically at least one particle originates from the V and not from the B decay vertex and has therefore a small value for the impact parameter IP.

- DTFχ^2: A fit to the full decay chain is performed. It uses the final state tracks as input to reconstruct the B candidate mass, decay vertex and momentum. In addition it also constrains the mass of the dimuon combination to the $\psi(2S)$ mass and the origin of the B candidate to the V in the event[5]. The additional constraints increase the precision on the reconstructed quantities. The fit is done with the Decay Tree Fit approach first described in [91]. The χ^2 of this fit is smaller for signal than for random track combinations, because the B candidates from combinations of random tracks rarely form a good vertex or have flight directions compatible with a V.

- vertex $\Delta\chi^2$(AddTrack): The vertex fit is repeated with additional tracks from the event. An additional track always increases the χ^2 of the vertex fit. Vertex $\Delta\chi^2$(AddTrack) corresponds to the vertex fit with the smallest increase. In signal it is usually large, because no additional track is nearby, while for partially reconstructed background it tends to lower values, if the missing particle created a track that was reconstructed. Combinatoric background comes from near the PV where the density of tracks is large. An additional track can be found more easily in these regions. Signal decays are less correlated to this high-density region.

- `DIRA`: cosine of the angle between the reconstructed B candidate momentum and the direction from PV to the reconstructed decay vertex. For signal the two vectors should be aligned and `DIRA` should be 1. For background with random track combinations (or where daughter particles are not reconstructed), the two vectors are not aligned.

The distributions of these variables for background and signal proxy are shown in Fig. 4.2 for Run 1. Since the the input variable distributions show large spikes and their values range over several decades, they are transformed using $\log x$ and $\log(1-x)$ to smooth out their distributions. This reduces the number of trees that are needed in the BDTs to achieve a good classification. Due to slightly different input variable distributions in Run 1 and Run 2 separate BDTs are trained for the two data-taking periods.

[5] If several V are present in the event, that V is chosen, for which DTFχ^2 is minimal.

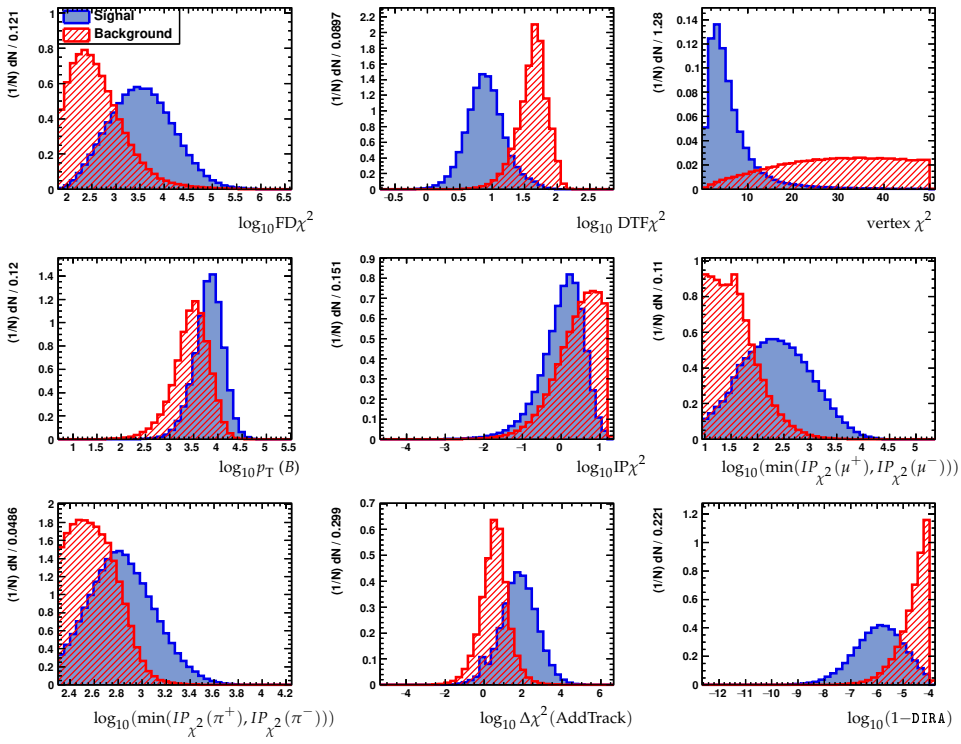

Figure 4.2: Input variables for the MVA for Run 2. Simulated truthmatched events from $B_s^0 \to \psi(2S)\pi^+\pi^-$ decays from Run 2 conditions in Blue and Run 2 data from the uppper sideband in red. First Row: FDχ^2, DTFχ^2, vertex χ^2. Second Row: p_{T} (B), IPχ^2, $\min(IP_{\chi^2}(\mu^+), IP_{\chi^2}(\mu^-))$. Third Row: $\min(IP_{\chi^2}(\pi^+), IP_{\chi^2}(\pi^-))$, $\Delta\chi^2$(AddTrack), DIRA. Some variables are transformed using $\log x$ and $\log(1-x)$ to smoothen them.

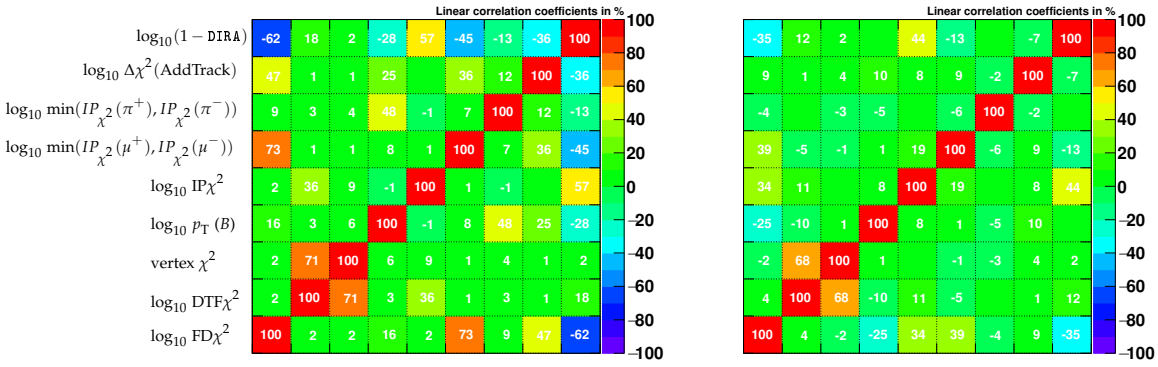

Figure 4.3: Linear correlation coefficients of Run 2 MVA input variables for signal (left) and background (right). The variables have the same order as in Fig. 4.2.

In Fig. 4.3 the linear correlation coefficients are shown for signal and background samples of the MVA trained for Run 2. The coefficients differ between signal and background samples, which justifies the use of BDTs and ANN, which can exploit these correlations for the classification.

The background rejection and signal efficiencies are compared in Fig. 4.4. The gradient BDT shows the largest background rejection in the region of interest above 50 % signal efficiency. Its output distribution is shown there as well. It shows much larger separation power than the individual input variables.

4.4.1 Cut Optimisation

The particle identification (PID) variables are correlated to the multiplicity of the underlying event, the position of the track inside the detector and the kinematic of the particles (see Sec. 3.2). These correlations are not perfectly described in the simulated samples, even after a correction step that corrects their dependence on momenta and the track multiplicity in the event.

Since the MVA should not be trained on these potentially mismodelled correlations, the PID variables are not used as inputs. They show good background rejection nevertheless, since a sizeable fraction of the pion candidates in the combinatoric background consists of misidentified Kaons. As a compromise the MVA output and the PID cuts are varied independently, but are optimised simultaneously in a grid search.

For each cut combination the data set is fitted in the B candidate mass $M(\psi(2S)\pi^+\pi^-)$ from 5150 to 5450 MeV. In this range 3 processes contribute

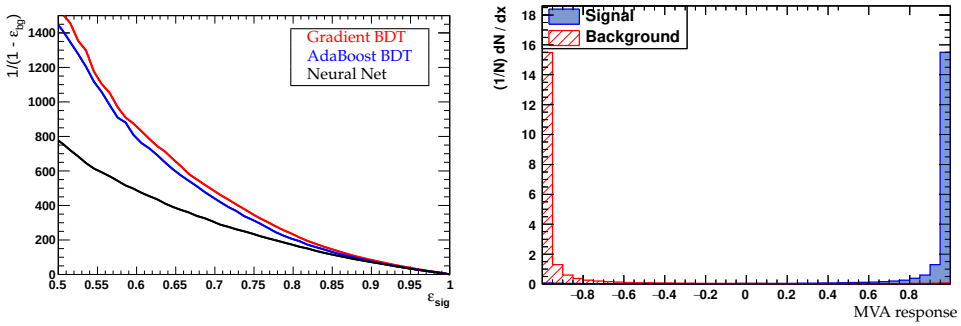

Figure 4.4: Left: 1/Background rejection $1/(1-\epsilon_{\mathrm{bg}})$ over signal efficiency ϵ_{sig} for three Classification algorithms for Run 2. Black: Neural net with 5 hidden layers and 30 nodes each. Blue: BDT with adaptive boost. Red: BDT with gradient boost. The BDT with gradient boost performs visibly better in the region of interest between 50 % and 70 % signal efficiency. Right: Output distribution of the gradient BDT for signal in blue and background in red for test samples.

to the data: The decays $B_s^0 \to \psi(2S)\pi^+\pi^-$, $B^0 \to \psi(2S)\pi^+\pi^-$ and combinatoric background[6].

The combinatoric background component in the data sample is approximately exponentially distributed in this mass range [98]. The decay constant and yield of the background pdf is left free in the fit.

The B_s^0 component is not Gaussian distributed in $M(\psi(2S)\pi^+\pi^-)$, because its charged daughter particles emit bremsstrahlung, when they traverse the detector and thereby lose energy. The invariant mass calculated from the measured 4-momenta of the 4 final state tracks therefore has a tail towards lower energies. This effect is partially compensated in the invariant mass that results from the fit of the full decay tree [91] using the external constraints on the vertices and the $\psi(2S)$ mass (see Sec. 4). The constraints lead to a reduced low-mass tail of the distribution and also reduce the width of the core-component. This does not work for every event though: The decay tree fit results in too large mass values for some events, which create a small contribution at higher masses.

The Crystal ball (CB) probability density function (pdf) [99] consists of a Gaussian core with mean $\bar{x}$ and width σ and a power-law tail on one side

[6] There is also a contribution from $B^0 \to \psi(2S)K^0$, where $K^0 \to \pi^+\pi^-$, which peaks at the same position as $B^0 \to \psi(2S)\pi^+\pi^-$. These additional events would be wrongly attributed to the yield of $B^0 \to \psi(2S)\pi^+\pi^-$ in the fit. Since that would bias the estimate for the signal yield, events are removed in this study, if their dipion mass lies around the K^0 mass. The s-weights for the amplitude fit are extracted in a reduced mass window, which does not contain events from $B^0 \to \psi(2S)K^0$. They are therefore not listed in Sec. 4.3.

with power n that starts at a distance α from the peak of the Gaussian. Its functional form is given in the following

$$f(x;\alpha,n,\bar{x},\sigma) = N \cdot \begin{cases} \exp(-\frac{(x-\bar{x})^2}{2\sigma^2}), & \text{for } \frac{x-\bar{x}}{\sigma} > -\alpha \\ A \cdot (B - \frac{x-\bar{x}}{\sigma})^{-n}, & \text{for } \frac{x-\bar{x}}{\sigma} \leqslant -\alpha \end{cases}, \tag{4.1}$$

where

$$\begin{aligned} A =& \left(\frac{n}{|\alpha|}\right)^n \cdot \exp\left(-\frac{|\alpha|^2}{2}\right), \\ B =& \frac{n}{|\alpha|} - |\alpha|, \\ N =& \frac{1}{\sigma(C+D)}, \\ C =& \frac{n}{|\alpha|} \cdot \frac{1}{n-1} \cdot \exp\left(-\frac{|\alpha|^2}{2}\right), \\ D =& \sqrt{\frac{\pi}{2}}\left(1 + \operatorname{erf}\left(\frac{|\alpha|}{\sqrt{2}}\right)\right). \end{aligned} \tag{4.2}$$

The sum of two Crystal ball pdfs is chosen to describe B_s^0 in $M(\psi(2S)\pi^+\pi^-)$, one with a tail to the left, one with a tail to the right. Since the core of the distribution is not perfectly Gaussian-shaped due to the use of external constraints in the fit, the two pdfs only share their parameter $\bar{x}$, but all other parameters, including the widths and the scale of the pdfs are allowed to differ.

Since the resolution is slightly too small in simulation, the total width of the Gaussian core is allowed to float in the fit to data. All other shape parameters (the tail parameters, the relative widths, and the relative yields of the two CB) are fixed to the values extracted from a fit to simulated $B_s^0 \to \psi(2S)\pi^+\pi^-$ events shown in Fig. 4.5. The tails differ significantly for the two sides, additionally a small bump can be seen below 5200 MeV, where the mass vetos introduced in Sec. 4.3 cut also some signal component away. This effect is negligibly small (note the logarithmic y-scale), and has little influence on the cut optimisation step. The actual fit for the signal extraction in Sec. 4.6 is performed in a smaller mass window, in order to avoid this bias.

The considerations for B_s^0 apply for the B^0 peak as well. It is therefore also fitted with a sum of two Crystal Ball functions with common mode position and its tail parameters are extracted from a fit to simulated events. The corresponding distribution in the simulated sample together with the fit is shown in Fig. 4.5.

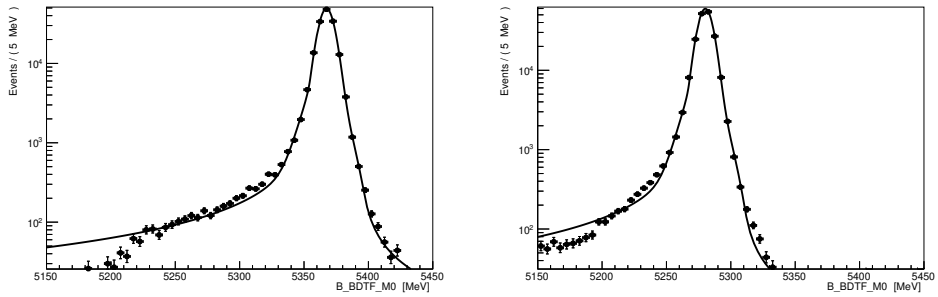

Figure 4.5: Fits to $M(\psi(2S)\pi^+\pi^-)$ after full selection. Left: Simulated $B_s^0 \to \psi(2S)\pi^+\pi^-$ events for RunII conditions. Fit model in black. Right: Simulated $B^0 \to \psi(2S)\pi^+\pi^-$ events for RunII conditions. Fit model in black.

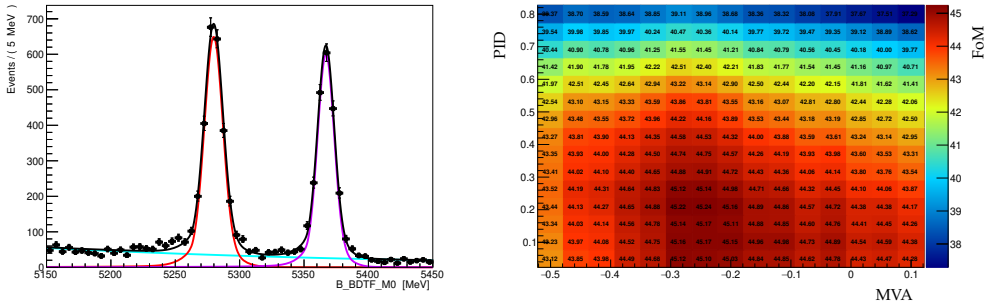

Figure 4.6: Left: Measured events from RunII data taking after all selection cuts. Full fit pdf in black, $B_s^0 \to \psi(2S)\pi^+\pi^-$ component in violet, $B^0 \to \psi(2S)\pi^+\pi^-$ component in red, combinatoric background component in cyan. Right: FoM on RunII data for $B_s^0 \to \psi(2S)\pi^+\pi^-$ for different cuts in MVA output (x-axis) and PID on the pion candidates (y-axis).

The fit to data is shown in Fig. 4.6. It uses data where the optimized cut values for the MVA output and the PID are already applied. Where these values come from is explained in the following.

If the shapes of all fit components are fixed in the fit, the fit reduces to the description of Poisson counting experiment. In this case the optimal cut with the highest significance for the signal yield S is found, if $S/\sqrt{S+B}$ is maximized, where B stands for the background yield inside the signal region (for this see *e.g.* the review on statistics in [21]). This is not the case for the fit that is performed on the data sample described in Sec. 4.6, because the signal and background shape parameters are partly left free in the fit. The additional uncertainty introduced by the free shape parameters reduce the signal significance for a given cut. This not only increases the statistical uncertainty on the results on the final amplitude model, the increased uncertainty

	Run 1	Run 2
$N_{B_s^0}$	737 ± 31	3062 ± 68
N_{B^0}	816 ± 33	3631 ± 62
$N_{\text{background}}$	780 ± 36	2786 ± 71

Table 4.1: Yields from fit to $M(\psi(2S)\pi^+\pi^-)$ in range $[5150, 5450]$ MeV.

on the background component's yield and shape can also lead to larger systematic uncertainties. In order to remedy this, the "classical" figure of merit is multiplied with an empirical penalty factor corresponding to the signal purity:

$$FoM = \frac{S}{S+B}\frac{S}{\sqrt{S+B}}. \tag{4.3}$$

This is a common choice for amplitude analyses a LHCb, *e.g.* [10].

In order to avoid a bias from optimizing the cut on the data itself, the FoM is not calculated with the signal yield directly. The branching ratios $\mathcal{B}(B_s^0 \to \psi(2S)\pi^+\pi^-)$ and $\mathcal{B}(B^0 \to \psi(2S)\pi^+\pi^-)$ are known, as well as the ratio of fragmentation fractions f_s/f_d, which describes how many B_s^0 are produced for every B^0 in the pp collision. Also acceptance effects and efficiencies are almost identical for both channels. This allows to easily calculate the expected B_s^0 yield for a given cut from the fit result for the B^0 yield, since efficiencies and acceptances mostly cancel in the ratio:

$$N_{B_s^0} = f_s/f_d \times \frac{\mathcal{B}(B_s^0 \to \psi(2S)\pi^+\pi^-)}{\mathcal{B}(B^0 \to \psi(2S)\pi^+\pi^-)} \times N_{B^0} \tag{4.4}$$

The 2D FoM distribution as function of the cut values on PID and BDT output is shown in Fig. 4.6 on the right. The final cuts are chosen to lie on the plateau with the largest values for the FoM: BDT($> 0.1(-0.2)$ and PID$> 0.25(0.25)$ for Run 1 (Run 2). For these cuts the fitted yields are given in Tab. 4.1.

4.5 Removal of Multiple Candidates

For some events two or more signal candidates are found that fulfil all the selection requirements. Since the branching ratio of the decay is very small, it is statistically unlikely that both come from signal decays. These additional candidates can lead to biased estimates of efficiencies [100] and are therefore

removed from the sample: From each event with multiple candidates, only one candidate is kept at random. This additional cut removes around 0.7 % of candidates on simulated $B_s^0 \to \psi(2S)\pi^+\pi^-$ events after all other selection steps. Since its effect on the FoM is small, the removal was not applied in the cut optimisation.

4.6 Signal Extraction

In the next step a "pure" $B_s^0 \to \psi(2S)\pi^+\pi^-$ sample is statistically extracted from the dataset via the `sPlot` technique [101], which calculates event-wise weights from a fit to the data. The weighted data set can then be used like a pure signal sample for all those distributions, which are not correlated to the fitted ones.

The `sPlot` corresponds to an optimized sideband subtraction technique. If the distribution for all components in a discriminant variable m are known (in this analysis these are the components of the fit to $M(\psi(2S)\pi^+\pi^-)$) the approach can be used to obtain the signal distributions in a control variable t (in this analysis the variables of the amplitude fit), if the pdfs of the individual components are not correlated between discriminant and control variable.

The s-weights $w_{\text{sig}}(m)$ are implemented as orthogonal functions on signal and background via [101]

$$\begin{aligned} \int \mathrm{d}m w_{\text{sig}}(m)\mathcal{P}_{\text{sig}}(m) &= 1 \\ \int \mathrm{d}m w_{\text{sig}}(m)\mathcal{P}_{\text{bg}}(m) &= 0, \end{aligned} \tag{4.5}$$

where $\mathcal{P}_{\text{sig}}$ and $\mathcal{P}_{\text{bg}}$ stand for the signal or background pdf respectively.

To obtain the s-weights, a fit of the mass distribution in $M(\psi(2S)\pi^+\pi^-)$ is used. The amplitude fit in Sec. 5 then uses the weighted data set and does not need to contain an explicit background component, because this component vanishes with application of the weights according to eq. 4.5. In order to remove possible remaining physical backgrounds, which could bias the the amplitude fit distributions due to their correlation to $M(\psi(2S)\pi^+\pi^-)$, the fit introduced in Sec. 4.4.1 is repeated in a narrower mass range around the B_s^0 peak. It is shown in Fig. 4.7 for simulated events and measured data from Run 2. This narrower range also excludes the bump from the mass vetos that is visible in Fig. 4.5. The narrower range also excludes the region, where

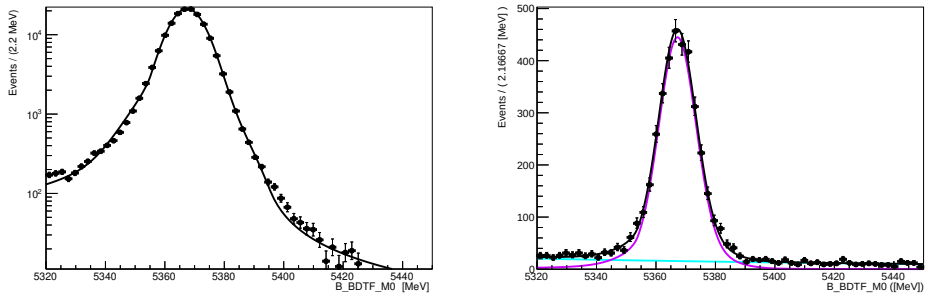

Figure 4.7: Fits to $M(\psi(2S)\pi^+\pi^-)$ in $5320, 5450]$ after full selection. Left: Simulated $B_s^0 \to \psi(2S)\pi^+\pi^-$ events for Run 2 conditions. Fit model in black. Right: Measured events from Run 2 data taking after all selection cuts. Full fit pdf in black, $B_s^0 \to \psi(2S)\pi^+\pi^-$ component in violet, combinatoric background component in cyan.

$B^0 \to \psi(2S)\pi^+\pi^-$ decays make a measurable contribution to the sample, allowing to fit the mass distribution with just the signal component and the combinatoric background.

In order to simplify the further analysis steps, the two data samples of Run 1 and Run 2 are combined in the fit to extract the *s*-weights. It has been tested, that the fit parameters are compatible between the data samples. In addition the systematic uncertainty from this choice is estimated in Sec. 9. The fitted yields are $N_{B_s^0} = 3683 \pm 70$ and $N_{\text{background}} = 1164 \pm 49$. The signal yield does not add up with the numbers obtained in Sec. 4.4.1, because 2 % of signal events lie outside the narrow mass window and multiple candidates are removed only after the MVA cut optimisation step.

5 Amplitude Model

One goal of this analysis is the test of the Bonn-Jülich parametrisation of the S-wave of the di-pion system [1, 46, 102] in $B_s^0 \to \psi(2S)\pi^+\pi^-$ using an amplitude fit to the data. It is tested against a more naive model using Flattè and Breit-Wigner (see PDG review on resonances in [21]) amplitudes to describe the S-wave. The amplitudes are implemented in the so-called helicity framework [41, 42], where they factorize into an angular part and a mass-dependent part. In the following, first the kinematic observables are introduced, then the helicity formalism is

described in Sec. 5.2. Its angular description and the necessary rotations for the application to $B_s^0 \to \psi(2S)\pi^+\pi^-$ are given in Sec.5.3. The different models for the invariant mass distribution of the mass-dependent dynamic part are given in Sec. 5.4 and 5.6. The full amplitude squared as a combination of all decay chains is derived in Sec. 5.7. Possible effects from mixing between the decays of B_s^0 and $\bar{B}_s^0$ are also covered there.

5.1 Kinematic Variables

Kinematic variables used in the fit follow the angular analyses of $B^0 \to \psi(2S)K^+\pi^-$ and $B_s^0 \to J/\psi\,\pi^+\pi^-$ in Belle and LHCb [14, 15, 103]. The decay amplitudes are constructed from subsequent two-body decays in the isobar model: The default decay chain is $B_s^0 \to \psi(2S)f_k$ with $\psi(2S) \to \mu^+\mu^+$ and the intermediate hadronic state f_k decays further via $f_k \to \pi^+\pi^-$. Since the $\psi(2S)$ intermediate state is very narrow (it decays predominantly electromagnetically), its mass can be assumed to be constant in the amplitude fit. Since a narrow state cannot interfere significantly with other states, this reduces the number of degrees of freedom that need to be described in the fit: For a full description of the $B_s^0 \to \psi(2S)\pi^+\pi^-$ decay chain the following 4 variables are sufficient[1]:

- $M_{\pi\pi}$: Invariant mass of the di-pion system

[1] Four final state particles have 16 degrees of freedom (dof) from their 4-vector components, but 4-momentum conservation removes 4 dof from that. Since the B_s^0 is spinless, the orientation in space is arbitrary in its center of mass system, which removes 3 dof. The masses

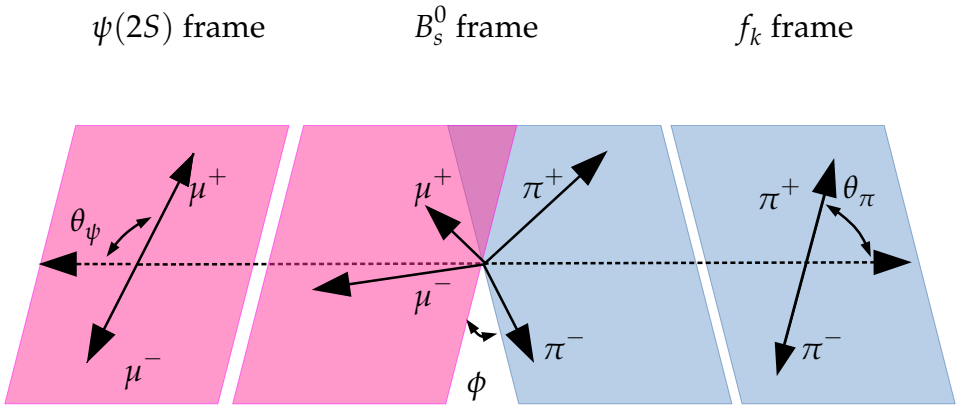

Figure 5.1: Definition of the helicity angles for the f_k chain.

- $\cos\theta_\pi$: cosine of the $\pi\pi$ helicity angle, which is the angle between the the π^+ momentum and the negative B_s^0 direction in the $\pi^+\pi^-$ rest frame.
- $\cos\theta_\psi$: cosine of the $\psi(2S)$ helicity angle, which is the angle between the μ^+ direction and the negative B_s^0 direction in the $\psi(2S)$ rest frame
- ϕ: angle between the decay planes of $f_k \to \pi^+\pi^-$ and the $\psi(2S) \to \mu^+\mu^-$ in the B_s^0 rest frame[2]

The angles θ_π,θ_ψ and ϕ are sketched in Fig. 5.1.

For the parametrisation of a decay via a possible $Z(4430)^\pm$ resonance, where $B_s^0 \to Z(4430)^\pm\pi^\mp$ with $Z(4430)^\pm \to \psi(2S)\pi^\pm$, the following four variables are chosen

- $M(\psi(2S)\pi^+)$, $M(\psi(2S)\pi^-)$: the invariant mass of the $\psi(2S)\pi^+$ and $\psi(2S)\pi^-$ system respectively

of the 4 final state particles and $\psi(2S)$ are fixed, which removes another 5 dof and brings the total number of dof down to 4.

[2] Alternatively it can be defined as the sum of the azimuthal angles of π^+ and μ^+ in the B_s^0 rest frame: $\phi = \phi_{\pi^+} + \phi_{\mu^+}$, which is used in the derivation of the angular description in eq. 5.13. For this to work, the coordinate systems of f_k and $\psi(2S)$ in the B_s^0 frame need to be chosen properly, which is achieved by aligning their x- and z-axes.

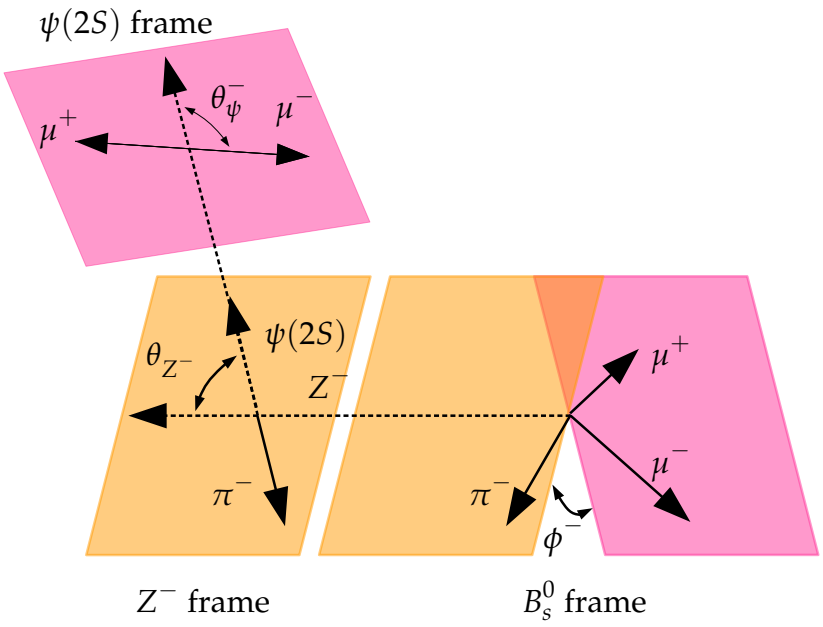

Figure 5.2: Definition of the helicity angles for the Z^- chain.

- $\cos\theta_{Z^+}$, $\cos\theta_{Z^-}$: cosine of the $Z(4430)^+$ and $Z(4430)^-$ helicity angle respectively. θ_{Z^+} is the angle between the $\psi(2S)$ and the negative B_s^0 direction in the $Z(4430)^+$ rest frame, while θ_{Z^-} is defined accordingly in the $Z(4430)^-$ rest frame.
- $\cos\theta_\psi^+$,$\cos\theta_\psi^-$: cosine of the $\psi(2S)$ helicity angle, which is the angle between the μ^+ (μ^-) direction and the negative $Z(4430)^+$ ($Z(4430)^-$) direction in the $\psi(2S)$ rest frame
- ϕ^+,ϕ^-: angle between the decay planes of the $Z(4430)^+$ ($Z(4430)^-$) and that of the $\psi(2S)$ in the B_s^0 rest frame

They are sketched in Fig. 5.2.

In order to correctly describe the interference between the amplitudes of the the decay via intermediate $\pi^+\pi^-$ resonances and that of decays via intermediate $\psi(2S)\pi^+$ resonances, an additional rotation is needed, which correctly aligns the helicity quantisation axis of the muons for the two decay chains. This angle $\alpha^\pm$ does not represent an additional degree of freedom, but depends on the other introduced kinematic variables. It is defined as the angle between the planes spanned by (μ^+,π^+) and (μ^+, $\pi^+\pi^-$) momenta in the $\psi(2S)$ rest frame for α^+. For the angle α^- the (μ^+,π^-) plane is used instead. It is depicted in Fig. 5.3.

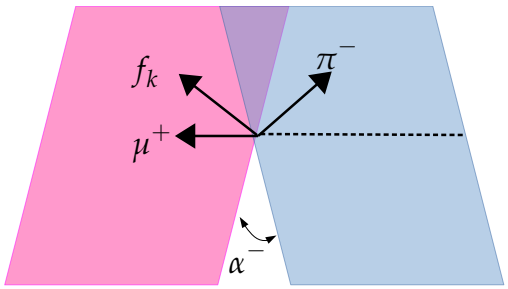

Figure 5.3: Definition of the angle α^-.

There is a one-to-one correspondence between the 4 variables used in the $\pi^+\pi^-$ decay chain and those of the $Z(4430)^+$ and the $Z(4430)^-$ chain. The mapping between the frames is detailed in the appendix of [15]. A different, more generic, approach is derived in [104]. It is shown there that both approaches give identical results.

5.2 Helicity Formalism

Different frameworks exist for the description of the decay amplitudes of relativistic particles with spin. Among others are the helicity, canonical or transversity basis, as well as different tensor formalisms (Zemach, covariant or Lorentz-invariant), see *e.g.* [105, 106]. Depending on the studied process, one is preferred over the other.

In contrast to the amplitudes in the tensor formalism, the helicity amplitudes have only dynamical singularities, which makes them favourable for the study of resonance properties, as explained in [107]. If a formalism has additional singularities, these can mimic a resonant behaviour in the amplitude and much care is needed to distinguish these from the real resonances and their extracted pole parameters. [107] also finds that the common covariant tensor approaches violate crossing symmetry introduced in eq. 2.23:

It is then possible that resonance parameters extracted from scattering experiments do not correspond to those extracted from decays, making these approaches less favourable for eventual future coupled channel approaches. It was therefore decided against a tensor formalism for the description of the decay amplitudes.

The available non-tensor formalisms can easily be transformed into each other via orthogonal functions. Against popular belief they all are relativistic in the sense that they transform correctly under Lorentz boosts, even though the involved wave functions include both covariant and non-covariant indices [107].

Since parity conservation can be introduced straightforwardly in the canonical basis using partial waves with fixed angular momenta ℓs (these partial waves with fixed spin and angular momentum being parity eigenstates), this formalism is chosen as basis for this work. In the descriptions of the implementations of the individual decay chains both helicity couplings $H^J_{\lambda_a\lambda_b}$ and ℓs couplings and their translation to ℓs couplings $B^J_{\ell s}$ are introduced, since some relations might be more familiar in the helicity framework.

In the following the derivation of the angular description of the decay is sketched following [42].

A spin state in the helicity basis $|j\lambda\rangle$ with spin j and helicity λ of a particle A in its rest frame $\vec{r}^A$ can be expressed in a different basis $\vec{r}^B$ with momentum $\vec{p} = (p, \theta, \phi)$ by application of a rotation $R(\phi, \theta, 0)$ and a Lorentz boost $L(\vec{p})$ [3] :

$$|\vec{p}, j\lambda\rangle_B = |\phi, \theta, p, j\lambda\rangle = L(\vec{p})R(\phi, \theta, 0)|\vec{0}, j\lambda\rangle_A. \tag{5.1}$$

From eq. 5.1 two-particle states can be built from the helicity basis vectors $|\vec{p}, j\lambda\rangle$

$$\begin{aligned} |\phi\theta\lambda_1\lambda_2\rangle &= aR(\phi, \theta, 0)\left\{L_z(p)|s_1\lambda_1\rangle L_{-z}(p)|s_2\ -\lambda_2\rangle\right\} \\ &= R(\phi, \theta, 0)|00\lambda_1\lambda_2\rangle), \end{aligned} \tag{5.2}$$

where $|s_i\lambda_i\rangle$ are the rest states of the individual particles and a is a normalisation. $L_{\pm z}$ is a Lorentz boost along the positive and negative z direction respectively.

[3] The helicity in the particle rest frame is 0, but is meant here as the helicity of the particle having an infinitesimal boost in the direction of $\vec{p}$.

A state with definite angular momentum J and z-component M can be derived from eq. 5.2 after some algebra, which is omitted here.

$$|JM\lambda_1\lambda_2\rangle = \sqrt{\frac{2J+1}{4\pi}} \int \mathrm{d}\cos\theta\, \mathrm{d}\phi\, D^{J*}_{M\lambda}(\phi,\theta,0)|\phi\theta\lambda_1\lambda_2\rangle, \tag{5.3}$$

where θ, ϕ are taken from particle-state 1 by convention. $\lambda = \lambda_1 - \lambda_2$ and $D^{J*}_{M\lambda}$ is the Wigner D-matrix

$$D^{J*}_{M\lambda}(\phi,\theta,0) = e^{-iM\phi} d^J_{M\lambda}(\theta), \tag{5.4}$$

where $d^J_{M\lambda}$ are the Wigner (small) d-matrices, which are tabulated.

Expanding the state $|JM\lambda_1\lambda_2\rangle$ in partial waves then gives

$$|JM\lambda_1\lambda_2\rangle = \sum_{\ell s} \left(\frac{2\ell+1}{2J+1}\right)^{\frac{1}{2}} (\ell 0 s\lambda|J\lambda)(s_1\lambda_1 s_2 - \lambda_2|s\lambda)|JM\ell s\rangle, \tag{5.5}$$

where $(\ell 0 s\lambda|J\lambda)(s_1\lambda_1 s_2 - \lambda_2|s\lambda)$ are the two Clebsch-Gordan coefficients, which project out the spin states. $\sum_s$ runs over all allowed values for the total spin $\vec{s_1} + \vec{s_2} = |s_1 - s_2|, ..., s_1 + s_2$, while $\sum_\ell$ runs over $\vec{s_1} + \vec{s} = |s_1 - s|, ..., s_1 + s$.

After some further manipulations the two-body decay amplitude A of a resonance with spin J can then be written as

$$\begin{aligned} A &= \langle \vec{p}\lambda_1'; -\vec{p}\lambda_2'|\mathcal{T}|JM\rangle \\ &= 4\pi\sqrt{\frac{m_0}{p}}\langle\phi,\theta,\lambda_1'\lambda_2'|JM\lambda_1\lambda_2\rangle\langle JM\lambda_1\lambda_2|\mathcal{T}|JM\rangle' \end{aligned} \tag{5.6}$$

where m_0 corresponds to the characteristic mass scale of the decaying state. $\mathcal{T}$ is the transition matrix element between the states. Since $\mathcal{T}$ is rotationally invariant, $\langle JM\lambda_1\lambda_2|\mathcal{T}|JM\rangle$ needs to be rotationally invariant as well, which means it can only depend on the rotationally invariant quantities J, λ_1 and λ_2 and the value of the daughter momenta p [105]. Using

$$\langle\phi,\theta,\lambda_1'\lambda_2'|JM\lambda_1\lambda_2\rangle = \sqrt{\frac{2J+1}{4\pi}} D^J_{M\lambda}(\phi,\theta,0)\delta_{\lambda_1'\lambda_1}\delta_{\lambda_2'\lambda_2} \tag{5.7}$$

from the normalisation and completeness relations of the helicity amplitudes, it is then possible to rewrite eq. 5.6 to

$$A = F^J_{\lambda_1\lambda_2}(p) D^{J*}_{M\lambda}(\phi,\theta,0), \tag{5.8}$$

The decay amplitude therefore factorises into an angular part (depending on the angles θ and ϕ), that is determined from first principles via the Wigner D-matrix and a dynamic part $F^J_{\lambda_1\lambda_2}(p)$, which describes the dependence on the daughter momenta.

In practice $F^J_{\lambda_1\lambda_2}(p)$ is usually split further into a (normalised) lineshape $R^J_{\lambda_1\lambda_2}(p)$, which describes the momentum-dependence of the term and a complex number $H_{\lambda_1\lambda_2}$, which describes the strength of each helicity amplitude. Different parametrisations exist for $R^J_{\lambda_1\lambda_2}(p)$, the best known being the Breit-Wigner amplitude. The parametrisations used in this analysis are covered in Sec. 5.4 and 5.6.

It is possible to expand the helicity decay amplitude F in terms of partial-wave amplitudes, which have fixed orbital angular momentum ℓ (and therefore fixed parity eigenvalues) via

$$\langle JM\lambda_1\lambda_2|\mathcal{T}|JM\rangle = \sum_{\ell s}\left(\frac{2\ell+1}{2J+1}\right)^{\frac{1}{2}}(\ell 0 s\lambda|J\lambda)(s_1\lambda_1 s_2-\lambda_2|s\lambda)\langle JM\ell s|\mathcal{T}|JM\rangle, \tag{5.9}$$

where the quantities in the brackets are the Clebsch-Gordan coefficients again.

This allows to relate the couplings $H_{\lambda_1\lambda_2}$ directly to their partial-wave counterparts $B^J_{\ell s}$ (called $a^J_{\ell s}$ in [105]) with the following relation:

$$H^J_{\lambda_1\lambda_2} = \sum_{\ell s} B^J_{\ell s}\left(\frac{2\ell+1}{2J+1}\right)^{\frac{1}{2}}(\ell 0 s\lambda|J\lambda)(s_1\lambda_1 s_2-\lambda_2|s\lambda). \tag{5.10}$$

5.3 Angular Description

In the following the expressions for the decay amplitudes for the different possible decay chains in $B^0_s\to\psi(2S)\pi^+\pi^-$ are derived. The focus here lies on the angular descriptions, while the dynamic part is covered in Sec. 5.4 and 5.6.

5.3.1 Decay via intermediate f_k resonance

If the decay of B^0_s proceeds via a partial wave f_k in the dipion spectrum, in principle three decays need to be considered: $B^0_s\to\psi(2S)f_k$, $\psi(2S)\to\mu^+\mu^-$, and $f_k\to\pi^+\pi^-$. Since the B^0_s meson does not have a spin, it decays isotropically and does not imprint an angular structure on the daughter particles (the Wigner D matrix is unity). Its decay amplitude therefore does not need to be considered here. This leaves the descriptions of $\psi(2S)\to\mu^+\mu^-$ and $f_k\to\pi^+\pi^-$, whose angular parts are taken from eq. 5.6:

The decay $f_k \to \pi^+\pi^-$ is described with

$$\begin{aligned} A^{f_k}_{\lambda_{f_k}}(\theta_\pi, \theta_\psi, \phi_{\pi^+}) &= H^{f_k}_{\lambda_{f_k}\lambda_\pi} D^{J_{f_k}*}_{\lambda_{f_k}\lambda_\pi}(\phi_{\pi^+}, \theta_\pi, 0) \\ &= H^{f_k}_{\lambda_{f_k}} e^{i\lambda_{f_k}\phi_{\pi^+}} d^{J_{f_k}}_{\lambda_{f_k}0}(\theta_\pi), \end{aligned} \tag{5.11}$$

where eq. 5.4 was used. λ_{f_k} is the helicity of the f_k partial wave with spin k, while $\lambda_\pi = \lambda_{\pi^+} - \lambda_{\pi^-} = 0$ is the helicity difference between the pions, which is 0, since both are spinless. The index is therefore dropped for the second line and the coupling strength $H^{f_k}_{\lambda_{f_k}}$ contains only one index.

The decay $\psi(2S) \to \mu^+\mu^-$ can be described with

$$\begin{aligned} A^{\psi}_{\lambda_\psi}(\theta_\pi, \theta_\psi, \phi_{\mu^+}) &= H^{\psi}_{\lambda_\psi\xi} D^{1*}_{\lambda_\psi\xi}(\phi_{\mu^+}, \theta_\psi, 0) \\ &= H^{\psi}_{\lambda} e^{i\lambda_\psi\phi_{\mu^+}} d^{1}_{\lambda_\psi\xi}(\theta_\psi), \end{aligned} \tag{5.12}$$

where λ_ψ is the helicity of the $\psi(2S)$ with the quantisation axis parallel to the B^0_s momentum in the $\psi(2S)$ rest frame. $\xi = \lambda_{\mu^+} - \lambda_{\mu^-}$ is the helicity of the muon pair with its quantisation axis the f_k momentum in the $\psi(2S)$ frame. $\xi = 0$ is helicity-suppressed with $m_\mu / \mathrm{m}_{\psi(2S)}$ and can therefore be neglected leaving for ξ therefore only the values -1 and 1. Due to parity conservation in the electromagnetic interaction, the corresponding coupling strengths are identical (which is why the index ξ is dropped in H_λ).

Since $J_{B^0_s} = 0$, $\lambda_\psi = \lambda_{f_k}$ is required from angular momentum conservation in the decay $B^0_s \to \psi(2S) f_k$. The helicities of f_k and $\psi(2S)$ λ_{f_k} and λ_ψ can therefore be replaced with λ everywhere, which allows to merge the helicity couplings of the $\psi(2S)$ decay with that of the f_k decay $H^{f_k}_{\lambda_{f_k}} H^{\psi}_{\lambda_\psi} = H^{f_k}_{\lambda}$ and one can combine eq. 5.11 and 5.12 to the full angular expression for the f_k decay chain

$$A^{f_k}_{\lambda\xi}(\theta_\pi, \theta_\psi, \phi) = H^{f_k}_{\lambda} d^{J_{f_k}}_{\lambda 0}(\theta_\pi) e^{i\lambda\phi} d^{1}_{\lambda\xi}(\theta_\psi), \tag{5.13}$$

where $\phi = \phi_{\pi^+} + \phi_{\mu^+}$ was used.

The complex helicity couplings $H^{f_k}_{\lambda}$ can be written as

$$H^{f_k}_{\lambda} = a_\lambda e^{i\beta_\lambda} = \Re\left(H^{f_k}_{\lambda}\right) + i\Im\left(H^{f_k}_{\lambda}\right), \tag{5.14}$$

where a_λ and β_λ represent its magnitude and phase. In the fit the real and imaginary part $\Re\left(H^{f_k}_{\lambda}\right)$ and $\Im\left(H^{f_k}_{\lambda}\right)$ are used as free real parameters to avoid the cyclic parameter β_λ. This improves the convergence of the fits.

Using the ℓs decomposition given in eq. 5.6 the amplitudes can be rewritten

as a sum of individual contributions with well-defined spin and parity. In this basis parity conservation in the strong and electromagnetic decays of the intermediate resonances can be enforced, which reduces the number of free helicity couplings in the fit. In addition, angular momentum barrier effects are easily incorporated as explained below.

In the decay $B_s^0 \to \psi(2S) f_k$ the total spin of the $\psi(2S)$ and f_k is $\vec{s} = \vec{J}_\psi + \vec{J}_{f_k}$ and the helicity couplings can be translated to the ℓs couplings applying eq. 5.10

$$H^{f_k}_{\lambda_{f_k}\lambda} = H^{f_k}_{\lambda\lambda} = \sum_{\ell_{f_k}} \sum_{s} \left(\frac{2\ell_{f_k}+1}{2J+1} \right)^{\frac{1}{2}} B^{f_k}_{\ell_{f_k} s} \left(\begin{matrix} J_\psi & J_{f_k} \\ \lambda & -\lambda \end{matrix} \middle| \begin{matrix} s \\ \lambda - \lambda \end{matrix} \right) \left(\begin{matrix} \ell_{f_k} & s \\ 0 & \lambda - \lambda \end{matrix} \middle| \begin{matrix} J_{B_s^0} \\ \lambda - \lambda \end{matrix} \right), \tag{5.15}$$

where $|J_{f_k} - J_\psi| \leq \ell_{f_k} \leq J_{f_k} + J_\psi$. The quantities in the brackets to the right describe the Clebsch Gordan coefficients in a different nomenclature, which shows the angular momentum quantum numbers on top and their projections on the bottom. Eq. 5.15 reduces to

$$H^{f_k}_{\lambda} = \sum_{\ell_{f_k}} \sum_{s} \left(\frac{2\ell_{f_k}+1}{2J+1} \right)^{\frac{1}{2}} B^{f_k}_{\ell_{f_k} s} \left(\begin{matrix} 1 & J_{f_k} \\ \lambda & -\lambda \end{matrix} \middle| \begin{matrix} s \\ 0 \end{matrix} \right) \left(\begin{matrix} \ell_{f_k} & s \\ 0 & 0 \end{matrix} \middle| \begin{matrix} 0 \\ 0 \end{matrix} \right) \tag{5.16}$$

and one can read off the relations for different f_k states.

For $J_{f_k} = 0$ the only ℓs combination with non-zero Clebsch-Gordan coefficients is 1 1, which corresponds to the complex parameter $B^{f_0}_{11}$. The only allowed helicity is $\lambda = 0$, the Clebsch-Gordan coefficients vanish for the other values and therefore

$$H^{f_0}_0 = B^{f_0}_{11} \tag{5.17}$$

In Sec. 2.5 it is shown that $J_{f_k} = 1$ states are forbidden due to isospin symmetry considerations in this channel. Assuming isosping breaking, the alowed values for ℓ_{f_k} would be $0 \leq \ell_{f_k} \leq 2$.

For $J_{f_k} = 2$ the allowed values are $1 \leq \ell_{f_2} \leq 3$, which correspond in principle to 3 complex parameters $B^{f_2}_{11}$, $B^{f_2}_{22}$, $B^{f_2}_{33}$. Due to the small energy release in the decay, configurations with large angular momenta, like the ones described by $B^{f_2}_{22}$ and $B^{f_2}_{33}$, are suppressed heavily. This is called angular momentum barrier effect. Therefore, in the default configuration only $B^{f_2}_{11}$ is considered, which results in

$$H^{f_2}_1 = H^{f_2}_{-1} = \frac{3}{5\sqrt{2}} B^{f_2}_{11} \tag{5.18}$$

and

$$H_0^{f_2} = -\frac{\sqrt{6}}{5} B_{11}^{f_2}. \tag{5.19}$$

The perpendicular polarisation of the decays via f_2 states is therefore $H_\perp^{f_2} = H_1^{f_2} - H_{-1}^{f_2} = 0$ for decays via f_2 states in the default fit configuration. In order to estimate the effect of the angular momentum barrier effect, $B_{22}^{f_2}$ couplings are included as well for a systematic study in Sec. 9.

5.3.2 Decays via intermediate $Z(4430)^-$ and $Z(4430)^+$

In the amplitude for a decay via an intermediate $Z(4430)^-$ the B_s^0-decay decays isotropically as well. Similar to the f_k decay chain one needs to describe here the decays $Z(4430)^- \to \psi(2S)\pi^-$ and $\psi(2S) \to \mu^+\mu^-$:

$$\begin{aligned} A^{Z^-}_{\lambda^-\xi}(\theta_{Z^-}, \phi^-, \theta^-_\psi, \alpha^-) &= H^{Z^-}_{\lambda^- 0} D^{J_Z *}_{0\lambda^-}(0, \theta_{Z^-}, 0) D^{1*}_{\lambda^-\xi}(\phi^-, \theta^-_\psi, 0) e^{i\xi\alpha^-} \\ &= H^{Z^-}_{\lambda^-} d^{Z^-}_{\lambda^- 0}(\theta_{Z^-}) e^{i\lambda^-\tilde{\phi}^-} d^1_{\lambda^-\xi}(\theta^-_\psi) \end{aligned} \tag{5.20}$$

This equation is derived similar to eq. 5.13 from eq. 5.6 as two subsequent decays. The coordinate system is chosen such that the azimuthal angle in the Z^- decay is 0.

The quantisation axis for the $\psi(2S)$ helicity λ^- in this decay chain is parallel to the π^- in the $\psi(2S)$ rest frame, which is a different axis than in the f_k chain, where the axis is parallel to the f_k flight direction instead. In order to correctly describe interference effects between the two chains, it is necessary to align the two chains with an additional rotation of α^-.

Following the considerations in [62], the helicity amplitudes are mapped to $\ell_{Z^-}s$ amplitudes using Clebsch-Gordan coefficients with orbital angular momentum L_Z and total spin of the decay products $\vec{s} = \vec{J}_\psi + \vec{J}_\pi = \vec{J}_\psi = 1$:

$$H^{Z^-}_{\lambda^-} = \sum_{\ell_{Z^-}} \sum_s \left(\frac{2\ell^{Z^-}+1}{2J+1}\right) B^{Z^-}_{\ell_{Z^-}s} \left(\begin{array}{cc|c} J_\psi & J_\pi & s \\ \lambda^- & -\lambda_\pi & \lambda^- - \lambda_\pi \end{array}\right) \left(\begin{array}{cc|c} \ell_{Z^-} & s & J_Z \\ 0 & \lambda^- - \lambda_\pi & \lambda^- - \lambda_\pi \end{array}\right) \tag{5.21}$$

From $\lambda_\pi = J_\pi = 0$ this reduces to

$$H^{Z^-}_{\lambda^-} = \sum_{\ell_{Z^-}} \left(\frac{2\ell^{Z^-}+1}{2J+1}\right) B^{Z^-}_{\ell_{Z^-}1} \left(\begin{array}{cc|c} \ell_{Z^-} & 1 & 1 \\ 0 & \lambda^- & \lambda^- \end{array}\right) \tag{5.22}$$

and one gets $|J_Z - 1| \leq \ell_{Z^-} \leq J_Z + 1$

Since the $\ell_{Z^-}s$ states are eigenstates of the parity operator, parity conservation in the strong decay $Z(4430)^- \to \psi(2S)\pi^-$ can be enforced via considering only ℓ_{Z^-} values, which satisfy[4]

$$P(\psi)P(\pi)(-1)^{\ell_{Z^-}} \equiv P(Z^-). \tag{5.23}$$

For $Z(4430)^-$, which has $J_Z^P = 1^+$ this allows only $\ell_{Z^-} \in [0,2]$, but due to the very small relative energy release in $Z(4430)^- \to \psi(2S)\pi^+$, higher orbital angular momentum states are strongly suppressed, leaving only $\ell_{Z^-} = 0$. Inserted into eq. 5.22, this gives $B_{01}^{Z^-} = H_{-1}^{Z^-} = H_1^{Z^-} = H_0^{Z^-}$, which leaves only one complex parameter free for the decay via $Z(4430)^-$.

Interpreted in the transversity basis [108], the perpendicular polarisation of the decays via $Z(4430)^\pm$ states would therefore be $H_\perp^{f_2} = H_1^{f_2} - H_{-1}^{f_2} = 0$.

The angular description in the $Z(4430)^+$ chain is the same as for the $Z(4430)^-$ chain, but all π^+ and π^- occurrences are exchanged, since they switch their roles from spectator to daughter and vice versa:

$$A_{\lambda^+\xi}^{Z^+}(\theta_{Z^+}, \phi^+, \theta_\psi^+, \alpha^+) = H_{\lambda^+}^{Z^+} d_{\lambda^+ 0}^{Z^+}(\theta_{Z^+}) e^{i\lambda^+\tilde{\phi}^+} d_{\lambda^+\xi}^{1}(\theta_\psi^+) e^{i\xi\alpha^+}. \tag{5.24}$$

5.4 Breit-Wigner Parametrisation

A simple way to parametrize a resonance of a given partial wave is the Breit-Wigner shape. This ansatz does not preserve unitarity and analyticity of the resulting amplitudes in cases, where resonances overlap and in the vicinity of thresholds (see Sec. 2.4.1). Nevertheless it is commonly used to parametrize experimental results. It serves as a benchmark for the alternative description provided by the Bonn-Jülich model, which is derived in Sec. 2.4.

The Breit-Wigner description of a resonance f_k in $B_s^0 \to \psi(2S) f_k$ with $f_k \to \pi^+\pi^-$ has the following form:

$$R(M_{\pi\pi}) = B'_{\ell_{f_k}}(p, p_0, d_{B_s^0}) \left(\frac{p}{p_0}\right)^{\ell_{f_k}} \cdot BW(M_{\pi\pi}|m_{0f_k}, \Gamma_{0f_k}) B'_{\ell_\pi}(q, q_0, d_{f_k}) \left(\frac{q}{q_0}\right)^{\ell_\pi} \tag{5.25}$$

Here $M_{\pi\pi}$ is the invariant mass of the dipion system introduced in Sec. 5.1, p is the momentum of f_k in the B rest frame, q is the momentum of the π^+ in the f_k frame. p_0 and q_0 represent their value at $m_{0f_k} = M_{\pi\pi}$, which

[4] In the helicity basis this requirement would be enforced via $H_{\lambda^-}^{Z^-} = P(\psi)P(\pi)P(Z^-)(-1)^{J(\psi)+J(\pi)-J(Z^-)} H_{-\lambda^-}^{Z^-}$.

is the mass parameter of the Breit-Wigner distribution mentioned earlier in Sec. 2.4.1. The orbital angular momentum between the f_k and the $\psi(2S)$ is ℓ_{f_k}, while ℓ_π stands for the orbital angular momentum between the two pions in the f_k decay. Since the pions don't have spin, $\ell_\pi = J_{f_k}$ is required from angular momentum conservation. The two fractions are needed to describe the turn-on behaviour at the thresholds of each decay correctly. Since these contributions grow rapidly for $\ell > 0$, the non-relativistically motivated, but purely phenomenological, form factors $B'_\ell(p, p_0, d)$ called Blatt-Weisskopf functions are introduced [44]. These cancel out the threshold factors at high momenta [43]. The two terms in eq. 5.25 correspond to the decay of the B_s^0 and f_k respectively. Even though the B_s^0 line-shape cannot be resolved experimentally (its width is much smaller than the detector resolution in $M(\psi(2S)\pi^+\pi^-)$), it is imprinted on the decay amplitudes of the dipion resonances via the Blatt-Weisskopf term, which depends on the B_s^0 mass and effective size $d_{B_s^0}$ of the B_s^0 meson.

The Breit-Wigner shape is also used for the description of decays via intermediate $Z(4430)^\pm$ resonances. In that case the corresponding values of the $B_s^0 \to Z(4430)^+\pi^-$ and $B_s^0 \to Z(4430)^-\pi^+$ decay chains are used in eq. 5.25. ℓ_{f_k} gets replaced by the angular momentum $\ell_{Z^\pm}$, ℓ_π with ℓ_ψ the angular momentum in the decay $Z^\pm \to \psi(2S)\pi^\pm$, d_{f_k} with $d_{Z^\pm}$ the effective size of the $Z^\pm$ state, m_{0f_k}, Γ_{0f_k} with $m_{0Z^\pm}$, $\Gamma_{0Z^\pm}$ the mass and width parameters of the Breit-Wigner parametrisation of the Z state.

The effective size of the mother system d is set to $1.6\,\mathrm{GeV}^{-1}$ for B_s^0, f_k and possible $Z(4430)^+$. Different values are tested for studies of systematic uncertainties in Sec. 9.6.

The relativistic Breit-Wigner function itself is given by

$$BW(M_{\pi\pi}|m_{0f_k}^2, \Gamma_{0f_k}) = \frac{1}{m_{0f_k}^2 - M_{\pi^+\pi^-}^2 - i m_{0f_k}\Gamma(M_{\pi\pi}, \Gamma_{0f_k})}, \tag{5.26}$$

where in contrast to the form used in eq. 2.41, here a mass-dependent width is used:

$$\Gamma(M_{\pi\pi}, \Gamma_{0f_k}) = \Gamma_{0f_k} \cdot \left(\frac{q}{q_0}\right)^{2\ell_{f_k}+1} \cdot \frac{m_{0f_k}}{M_{\pi\pi}} \cdot B'_{\ell_{f_k}}(q, q_0, d_{f_k})^2. \tag{5.27}$$

Note that this mass dependent width is not as general as the one introduced in eq. 2.36. It only accounts for the threshold of the dipion channel itself, not for other possible thresholds near m_{0f_k}.

The Breit-Wigner parameters m_{0f_k} and Γ_{0f_k} do not have a strict physical

meaning in general, but can be interpreted as the mass and width of the resonant state, if there are no other, overlapping resonances contributing to the decay and no thresholds to additional decay channels lie within the vicinity of its mass (see Sec. 2.4.1).

5.5 Flatté Parametrisation

The $f_0(980)$ is the dominating resonance in the $M_{\pi\pi}$ spectrum in $B_s^0 \to \psi(2S)\pi^+\pi^-$. It can not be described with a simple Breit-Wigner amplitude, because it lies directly at the $K\bar{K}$ threshold and couples strongly to this system. This makes a coupled channel treatment necessary, which can be approximated with a Flatté distribution using a two-width parametrisation (see the review on resonances in [21]). Here Γ in the relativistic Breit-Wigner amplitude from Eq. 5.26 is replaced with $\tilde{\Gamma}$:

$$\tilde{\Gamma} = \frac{2\rho_1\Gamma_1 + 2\rho_2\Gamma_2}{m_{0f_k}} \tag{5.28}$$

with $\Gamma_i = g_i^2$ acting as partial decay widths to the two channels ($\pi\pi$ and KK) as defined in eq. 2.34, and ρ_1, ρ_2 the phase space terms for two-body decays (with symmetric final states) given by:

$$\begin{aligned} \rho_1 &= \frac{2q}{M_{\pi\pi}}, \\ \rho_2 &= \frac{2q_K}{M_{\pi\pi}}, \end{aligned} \tag{5.29}$$

where q is the momentum of the π in the $f_0(980)$ frame in the decay via $f_0(980) \to \pi\pi$ (as above) and q_K is the momentum of the K in the $f_0(980)$ frame in the decay $f_0(980) \to KK$. The $f_0(980)$ also couples to the neutral channels $\pi^0\pi^0$ and K^0K^0, whose thresholds differ by a few MeV due to the different masses of the daughters from isospin breaking. The peak shape is sensitive to the threshold, so effectively 4 channels need to be considered in the Flatté description.

The couplings g_i on the other hand are assumed to obey Isospin symmetry, since the shape is not as dependent on them. In summary the neutral isospin

partner channels can be incorporated into eq. 5.28 via:

$$\begin{aligned}\rho_1\Gamma_1 &\to \left(\sqrt{\frac{2}{3}}\frac{2q}{M_{\pi\pi}} + \sqrt{\frac{1}{3}}\frac{2q_{\pi^0}}{M_{\pi\pi}}\right)\Gamma_1 \\ \rho_2\Gamma_2 &\to \left(\sqrt{\frac{1}{2}}\frac{2q_{K^+}}{M_{\pi\pi}} + \sqrt{\frac{1}{2}}\frac{2q_{K^0}}{M_{\pi\pi}}\right)\Gamma_2,\end{aligned} \tag{5.30}$$

where q_{π^0}, q_{K^0}, and q_{K^+} are the momenta of one of the final state particles of the corresponding channel in the rest frame of the $f_0(980)$ and the prefactors are derived from isospin symmetry.

The Flatté line shape is then given by

$$R(M_{\pi\pi}) = B'_{\ell_{f_0}}(p, p_0, d_{B_s^0})\left(\frac{p}{p_0}\right)^{\ell_{f_0}} \cdot BW(M_{\pi\pi}|m_{0f_k}, \tilde{\Gamma}')B'_{\ell_\pi}(q, q_0, d_{f_k})\left(\frac{q}{q_0}\right)^{\ell_\pi} \tag{5.31}$$

As derived in Sec. 5.3, the only allowed value for ℓ_{f_0} is 1. In addition, since $f_0(980)$ has $J = 0$, ℓ_π is 0 too and the factors B'_{ℓ_π} and $(q/q_0)^{\ell_\pi}$ simplify to unity, which reduces eq. 5.31 to

$$R(M_{\pi\pi}|m^2_{0f_0}, \Gamma_{0f_0}) = B'_1(p, p_0, d_{B_s^0})\left(\frac{p}{p_0}\right) BW(M_{\pi\pi}|m_{0f_0}, \tilde{\Gamma}'). \tag{5.32}$$

Note that even though this parametrisation can correctly describe the $M_{\pi\pi}$ shape around the threshold, it has significant overlap with the $f_0(1500)$ resonance, which leads to an overall non-unitary amplitude.

5.6 Bonn-Jülich S-wave description

The Flatté description of the $f_0(980)$ describes the phenomenology at the $K\bar{K}$ threshold, but due to the presence of further, overlapping S-wave states, mostly the $f_0(1500)$, the resulting combined amplitude does not fulfil the analyticity or unitarity requirements. Ultimately this means that the extracted pole positions from the fit are not universal and will depend on the reaction in which they are measured. Also the couplings, fit fractions and relative phases are not reliable, even if the fit can describe the data reasonably well.

The Bonn-Jülich model [1, 46, 102] avoids this problem with a new parametrisation of the scalar pion form factors. It is introduced in Sec. 2.4.2, while here the technical implementation into the amplitude fit

of $B_s^0 \to \psi(2S)\pi^+\pi^-$ is described. The form factor derived in eq. 2.53 is

$$F_\pi(M^2_{\pi^+\pi^-}) = \sum_{i=\pi,K,4\pi} \tilde{F}_i(M^2_{\pi^+\pi^-})(c_i + l_i M^2_{\pi^+\pi^-}) + \sum_{r=f_0(1500)f_0(2020)} \bar{F}_r(M^2_{\pi^+\pi^-})\alpha_r. \tag{5.33}$$

The base functions $\tilde{F}_i(s)$ and $\bar{F}_r(s)$ are process independent and have been determined by the Bonn-Jülich group numerically using different $\pi\pi$ data from $B_s^0 \to J/\psi\pi^+\pi^-$ decays [1]. These base functions were evaluated on approximately 1300 points for $M_{\pi\pi} \in [0, 2.3\,\text{GeV}]$ and the resulting values provided via .csv files. This resolution is sufficiently better than the experimental resolution on $M_{\pi\pi}$ (determined in Sec. 6.2.1). They are interpolated with cubic splines in the amplitude fit. These interpolations are shown in Fig. 5.4. The normalisation constants c_i are fixed from quark-level considerations: The pure $s\bar{s}$ source cannot directly form $\pi\pi$ or 4π states at small masses due to OZI suppression. This means only c_k is non-zero. The parameters l_i are known to only give small corrections to the form factor in other processes. They are therefore set to 0 for the nominal fit, which also means that the first sum in eq. 5.33 reduces to the KK component. The only parameters that need to be determined from the fit to $B_s^0 \to \psi(2S)\pi^+\pi^-$ in the Bonn-Jülich model are therefore $\alpha_{f_0(1500)}$ and $\alpha_{f_0(2020)}$, which describe the coupling to the non-dispersively introduced resonances. The corresponding line-shape in $M^2_{\pi^+\pi^-}$ is then given by eq. 2.55 from Sec. 2.4.2

$$R(M^2_{\pi^+\pi^-}) = p_\psi m_B F_\pi(M^2_{\pi^+\pi^-}), \tag{5.34}$$

with p_ψ the $\psi(2S)$ momentum in the B_s^0 rest frame, and m_B the B_s^0 mass.

5.7 $B_s^0 \to \psi(2S)\pi^+\pi^-$ Amplitude via intermediate f_k and $Z(4430)^\pm$

The full time-independent part of the amplitude squared of $B_s^0 \to \psi(2S)\pi^+\pi^-$ is the combination of 5.6, 5.13, 5.20, and 5.24. As for all quantum mechanical measurements, only the squared amplitude is accessible to the experiment:

$$\mathcal{A}(\Phi) = \sum_{\xi=-1,1} \left| \sum_{f_k} \sum_{\lambda=-1,0,1} A^{f_k}_{\lambda\xi} + \sum_{\lambda^-=-1,0,1} A^{Z^-}_{\lambda^-\xi} + \sum_{\lambda^+=-1,0,1} A^{Z^+}_{\lambda^+\xi} \right|^2. \tag{5.35}$$

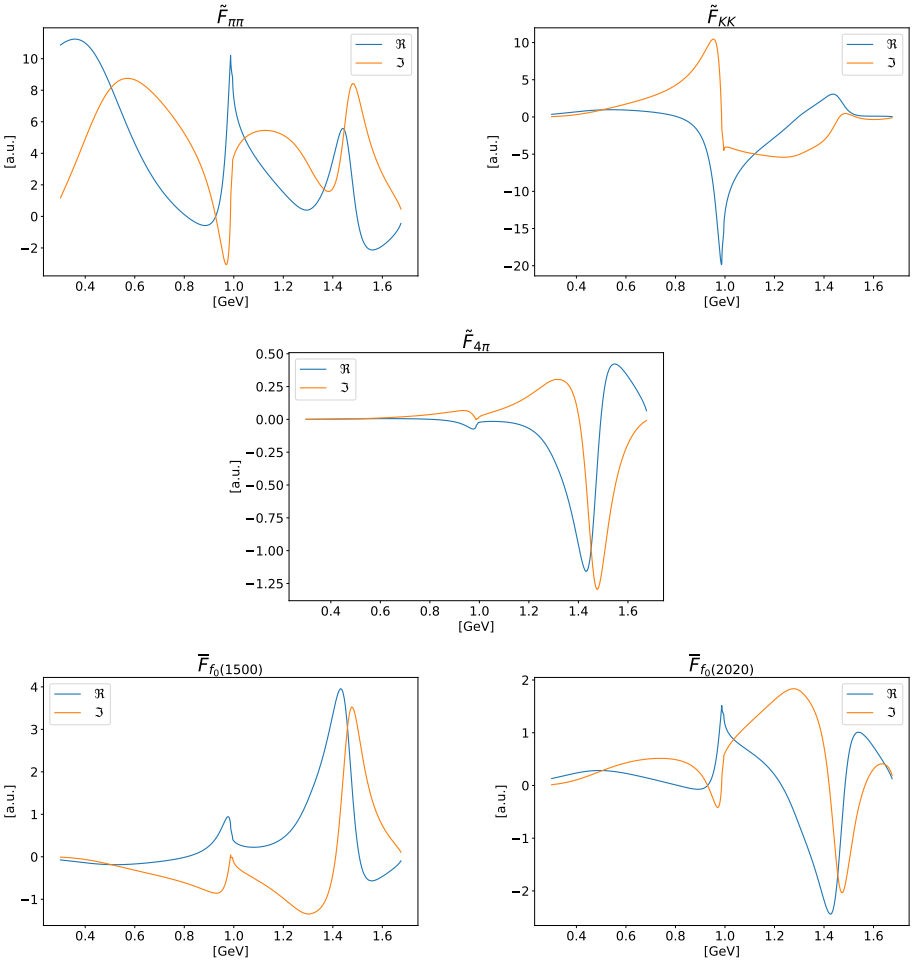

Figure 5.4: Real and Imaginary parts of base functions based on Bonn-Jülich model versus energy in GeV [1]. Top left $\tilde{F}_{\pi}(M_{\pi\pi})$, Top Right: $\tilde{F}_{K}(M_{\pi\pi})$. Middle: $\tilde{F}_{4\pi}(M_{\pi\pi})$. Bottom Left: $\bar{F}_{f_0(1500)}(M_{\pi\pi})$ Bottom Right: $\bar{F}_{f_0(2020)}(M_{\pi\pi})$. $\tilde{F}_{\pi}(M_{\pi\pi})$ and $\tilde{F}_{4\pi}(M_{\pi\pi})$ are only used for systematic studies where linear terms $l_i M^2_{\pi^+\pi^-}\tilde{F}_i(M_{\pi\pi})$ are allowed in the fit.

All internal helicity states are summed up coherently, the same goes for the hadronic intermediate states from the $\pi^+\pi^-$, $Z(4430)^+$, and $Z(4430)^-$ decay chains. The summation over the muon helicity ξ is not coherent, since quantum mechanically the measurement of the muon helicities is possible, even though it is not feasible with the LHCb detector.

But the amplitude $\mathcal{A}(\Phi)$ only describes the decay of a pure B_s^0 state, while in the experiment both B_s^0 and $\bar{B}_s^0$ mesons are produced and both can decay into the common final state $\psi(2S)\pi^+\pi^-$. In order to correctly describe the decays of both mesons and their mixing, the approach detailed in [108] is taken up. This approach is briefly sketched in the following.

The fact that the quark mass eigenstates are not aligned with the quark flavour eigenstates was introduced in Sec. 2.2. Similarly the mass eigenstates $|B_L\rangle$ and $|B_H\rangle$ are not eigenstates of the quark flavour basis $\left|B_s^0\right\rangle$ ($\bar{b}s$) and $\left|\bar{B}_s^0\right\rangle$ ($b\bar{s}$) either, but can be expressed as superpositions

$$\begin{aligned} |B_L\rangle &= p\left|B_s^0\right\rangle + q\left|\bar{B}_s^0\right\rangle \\ |B_H\rangle &= p\left|B_s^0\right\rangle - q\left|\bar{B}_s^0\right\rangle, \end{aligned} \tag{5.36}$$

where $|p|^2 + |q|^2 = 1$ provides the proper normalisation of the wavefunction. The time evolution of a $B_s^0\bar{B}_s^0$ system can be described with the Schrödinger equation

$$\mathbf{H}\begin{pmatrix} B_s^0(t) \\ \bar{B}_s^0(t) \end{pmatrix} = i\hbar\frac{\partial}{\partial t}\begin{pmatrix} B_s^0(t) \\ \bar{B}_s^0(t) \end{pmatrix}, \tag{5.37}$$

where $\hbar$ is the Planck constant and $\mathbf{H}$ is the system's Hamiltonian, , which can be written as

$$\mathbf{H} = \mathbf{M} - \frac{i}{2}\mathbf{\Gamma} = \begin{pmatrix} M_{11} & M_{12} \\ M_{12}^* & M_{11} \end{pmatrix} - \frac{i}{2}\begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{12}^* & \Gamma_{11} \end{pmatrix}. \tag{5.38}$$

The matrices $\mathbf{M}$ and $\mathbf{\Gamma}$ describe the off-shell and on-shell contributions to the B_s^0-$\bar{B}_s^0$ mixing respectively. Their diagonal elements are constrained by *CPT* symmetry to $M_{11} = M_{22}$ and $\Gamma_{11} = \Gamma_{22}$ respectively. In addition M and Γ are Hermitian, which implies $M_{21} = M_{12}^*$ and $\Gamma_{21} = \Gamma_{12}^*$.

The solution of this equation for an initial state B_s^0 or $\bar{B}_s^0$ is

$$\begin{aligned}\left|B_s^0(t)\right\rangle =& g_+(t)\left|B_s^0\right\rangle - \frac{q}{p}g_-(t)\left|\bar{B}_s^0\right\rangle \\ \left|\bar{B}_s^0(t)\right\rangle =& g_+(t)\left|\bar{B}_s^0\right\rangle - \frac{p}{q}g_-(t)\left|B_s^0\right\rangle \\ &,\end{aligned} \tag{5.39}$$

where g_+ and g_- describe the flavour-conserving and flavour-changing part of the mixing expressed by

$$\begin{aligned}g_+(t) =& e^{-imt}e^{-\Gamma t/2}\left(\cosh\frac{\Delta\Gamma t}{4}\cos\frac{\Delta m t}{2} - i\sinh\frac{\Delta\Gamma t}{4}\sin\frac{\Delta m t}{2}\right) \\ g_-(t) =& e^{-imt}e^{-\Gamma t/2}\left(-\sinh\frac{\Delta\Gamma t}{4}\cos\frac{\Delta m t}{2} + i\cosh\frac{\Delta\Gamma t}{4}\sin\frac{\Delta m t}{2}\right),\end{aligned} \tag{5.40}$$

where $m = \frac{1}{2}(m_H + m_L)$, $\Delta m = m_H - m_L$, $\Gamma = \frac{1}{2}(\Gamma_H + \Gamma_L)$ and $\Delta\Gamma = \Gamma_L - \Gamma_H$. $m_{H/L}$, $\Gamma_{H/L}$ are the masses and widths of $|B_H\rangle$ and $|B_L\rangle$.

To calculate the time-dependent decay amplitudes of a flavour state B_s^0 and its *CP* conjugate state $\bar{B}_s^0$ into a common final state f, we first define the full time-dependent amplitudes

$$\begin{aligned}\mathcal{A} =& \left\langle f|B_s^0\right\rangle \\ \overline{\mathcal{A}}_f =& \left\langle f|\bar{B}_s^0\right\rangle,\end{aligned} \tag{5.41}$$

where we dropped the t and Φ dependence. Additionally

$$\begin{aligned}\mathcal{A} =& \mathcal{A}_f \\ \overline{\mathcal{A}} =& \frac{q}{p}\overline{\mathcal{A}}_f.\end{aligned} \tag{5.42}$$

Following [109] these definitions allow to rewrite eq. 5.39 and build the time dependent decay rate

$$\begin{aligned}\Gamma(t) =& Ne^{-\Gamma t}\left[\frac{|\mathcal{A}|^2+|\overline{\mathcal{A}}|^2}{2}\cosh\left(\frac{\Delta\Gamma t}{2}\right)+\frac{|\mathcal{A}|^2-|\overline{\mathcal{A}}|^2}{2}\cos(\Delta m t)\right.\\ &\left.-\Re\left(\frac{q}{p}\mathcal{A}^*\overline{\mathcal{A}}\right)\sinh\left(\frac{\Delta\Gamma t}{2}\right)-\Im\left(\frac{q}{p}\mathcal{A}^*\overline{\mathcal{A}}\right)\sin(\Delta m t)\right],\\ \overline{\Gamma}(t) =& \left|\frac{p}{q}\right|^2 Ne^{-\Gamma t}\left[\frac{|\mathcal{A}|^2+|\overline{\mathcal{A}}|^2}{2}\cosh\left(\frac{\Delta\Gamma t}{2}\right)-\frac{|\mathcal{A}|^2-|\overline{\mathcal{A}}|^2}{2}\cos(\Delta m t)\right.\\ &\left.-\Re\left(\frac{q}{p}\mathcal{A}^*\overline{\mathcal{A}}\right)\sinh\left(\frac{\Delta\Gamma t}{2}\right)+\Im\left(\frac{q}{p}\mathcal{A}^*\overline{\mathcal{A}}\right)\sin(\Delta m t)\right],\end{aligned} \tag{5.43}$$

where N is a normalisation constant.

In the data sample used for this analysis the initial B meson flavour cannot be determined from the final state, since both B_s^0 and $\overline{B}_s^0$ can decay into it. A flavour-tagging from the presence of flavour-specific by-products of the hadronisation of the B mesons is infeasible either, because the efficiency of this approach at LHCb is not very high. In the analysis of $B_s^0 \to J/\psi\,\pi^+\pi^-$ the effective sample size decreased to only $(3.89 \pm 0.25)\,\%$ of its original size due to the flavour-tagging inefficiency [110]. The analysis in [110] was designed to extract the CP violating phase ϕ_s, for which the knowledge of the initial flavour is essential. The analysis of $B_s^0 \to \psi(2S)\pi^+\pi^-$ presented here on the other hand, does not try to extract weak phases though. Therefore a flavour-tagging is not necessary here and only the sum of the two decay rates of B_s^0 and $\overline{B}_s^0$ are measured

$$\Gamma(t) + \overline{\Gamma}(t). \tag{5.44}$$

Here it is assumed that there are negligible asymmetries in the production of B_s^0 mesons, which is supported by [111], where no significant asymmetry was found. Additionally, it assumes $\left|\frac{q}{p}\right|^2 \approx 1$, which was measured in [112] to 1.0039 ± 0.0033.

Inserting eq. 5.43 into eq. 5.44 two of the terms disappear, leaving only

$$\Gamma(t) + \overline{\Gamma}(t) \propto e^{-\Gamma t}\left[\frac{|\mathcal{A}|^2+|\overline{\mathcal{A}}|^2}{2}\cosh\left(\frac{\Delta\Gamma t}{2}\right)-\Re\left(\frac{q}{p}\mathcal{A}^*\overline{\mathcal{A}}\right)\sinh\left(\frac{\Delta\Gamma t}{2}\right)\right]. \tag{5.45}$$

Additionally, the amplitude measurement is performed time-independently, therefore the time-dependence of eq. 5.45 needs to be integrated over via

$$\int_0^\infty \Gamma(t) + \overline{\Gamma}(t) dt, \tag{5.46}$$

which assumes a flat acceptance and reconstruction efficiency as function of proper decay time. This approximation is motivated in Sec. 5.7.1. Eq. 5.46 can be solved analytically via integration by parts to

$$\int_0^\infty \Gamma(t) + \overline{\Gamma}(t) dt \propto \frac{|\mathcal{A}|^2 + |\overline{\mathcal{A}}|^2}{2} \left(\frac{1}{\Gamma_L} + \frac{1}{\Gamma_H} \right) + \Re \left(\frac{q}{p} \mathcal{A}^* \overline{\mathcal{A}} \right) \left(\frac{1}{\Gamma_L} - \frac{1}{\Gamma_H} \right). \tag{5.47}$$

Even though the time-dependence is integrated out, the effect of the mixing is therefore still measurable through the second term in eq. 5.47. The values for $\frac{1}{\Gamma_L} = 1.414 \pm 0.006\,\text{ps}$ and $\frac{1}{\Gamma_H} = 1.619 \pm 0.009\,\text{ps}$ are taken from [113]. Inserted into 5.47, this dilutes the mixing contribution from the second term with respect to the incoherent contribution from the first term by

$$D = \left(\frac{1}{\Gamma_L} - \frac{1}{\Gamma_H} \right) / \left(\frac{1}{\Gamma_L} + \frac{1}{\Gamma_H} \right) \approx 0.068 \tag{5.48}$$

In order to make use of eq. 5.47 one still needs an explicit expression for $\overline{\mathcal{A}}$, describing the $\bar{B}_s^0 \to \psi(2S)\pi^+\pi^-$. The angle definitions are the same for decays of B_s^0 and those of $\bar{B}_s^0$, but the ℓs $B_{\ell s}$ couplings need to be replaced by $\overline{B}_{\ell s}$, which are related by the *CP* parities η_i of the decay amplitude $\mathcal{A}_i$ [108]:

$$\frac{q}{p} \frac{\overline{\mathcal{A}}_i}{\mathcal{A}_i} = \eta_i \exp i\phi_s. \tag{5.49}$$

Here $\phi_s = -0.050 \pm 0.019\,\text{rad}$ describes the weak CKM phase responsible for *CP* violation in mixing. Assuming no direct *CP* violation in the decay, $\frac{q}{p} = \exp i\phi_s$. This is satisfied for the decay $B_s^0 \to \psi(2S)\pi^+\pi^-$, because the leading Feynman diagram shown in Fig. 2.4 does not contain a weak phase up to order λ^3 in the Wolfenstein parametrisation (see eq. 2.2) and next to leading order diagrams have negligibly small amplitudes. This allows to rewrite eq. 5.49 to

$$\overline{\mathcal{A}}_i = \eta_i \mathcal{A}_i \tag{5.50}$$

The values of η_i can be inferred from the application of *CP* on the individual systems. The $\psi(2S)$ is a *CP* eigenstate with $P = C = -1$, its combined *CP* parity is therefore 1. Applying C to the dipion system exchanges the π^+ with the π^-, which has the same effect as a parity transformation P. The parity of

the two-pion system is given by

$$P(\pi^+)P(\pi^-)(-1)^{\ell_\pi} = (-1)^{\ell_\pi}, \tag{5.51}$$

where $P(\pi^+) = P(\pi^-) = -1$ and ℓ_π is the relative angular momentum between the two pions introduced in Sec. 2.5 (in the f_k chain, this is given by the spin of the f_k state). Since $C = P = (-1)^\ell$ for the dipion system, the combined CP parity is always even, which means that the final state $\psi(2S)\pi^+\pi^-$ has even CP parity.

If the two terms in eq. 5.47 are proportional, the mixing only introduces an overall scale factor for the amplitudes and therefore the pdf. This scales the $B_{\ell s}$ coupling parameters in the fit, but their absolute sizes are not relevant for the determination of the fit fractions, interferences and polarisations, which are the typical parameters of interest of amplitude analyses. Only if the two terms are not proportional, the mixing changes the shape of the amplitudes and therefore the pdf. In that case the mixing needs to be considered explicitly in the amplitude fit. In the following it is shown that this is the case, if amplitudes with different CP parities are present for the intermediate states.

If two decay paths with different CP parities A^+ and A^- with $CP(A^\pm) = \pm 1$ and $\mathcal{A} = A^+ + A^-$ are possible, the second term of eq 5.47 evaluates after some algebra to

$$\begin{aligned} \Re \frac{q}{p} A^* \overline{A} &= \Re\left(\frac{q}{p}\left(\left|A^+\right|^2 - \left|A^-\right|^2 + 2\Im\left(A^{-*}A^+\right)\right)\right) \\ &= \left|A^+\right|^2 - \left|A^-\right|^2 + 2\Im\left(A^{-*}A^+\right), \end{aligned} \tag{5.52}$$

where $\frac{q}{p} = \exp i\phi_s \approx 1$ was used. The interference term between amplitudes of different CP eigenvalues $\left(A^{-*}A^+\right)$ has different angular dependencies than the other components of eq. 5.47, which needs to be taken into account, if more than one CP eigenstate is present in the decay into a common final state.

In the following, the CP conjugate decay amplitudes of the different decay chains of $B_s^0 \to \psi(2S)\pi^+\pi^-$ are derived and their CP eigenvalues determined.

5.7.1 CP conjugate Amplitude for $B_s^0 \to f_k\psi(2S)$

In the decay chain involving intermediate f_k states, the f_k states decay strongly via $f_k \to \pi^+\pi^-$. This decay conserves CP and the $\pi^+\pi^-$ system is an even eigenstate of CP, as shown in eq. 5.51. Since visible mixing effects only occur

for processes, which involve amplitudes with different CP parities, for the following considerations on CP eigenvalues this decay does not play a role. The same considerations hold for the $\psi(2S)$, which is also an even eigenstate of CP.

The CP parity of the intermediate state $\langle\psi(2S)f_k|$ can be determined via

$$CP(\langle\psi(2S)f_k|) = CP(f_k)CP(\psi(2S))(-1)^{\ell_{f_k}}\langle\psi(2S)f_k| = (-1)^{\ell_{f_k}}\langle\psi(2S)f_k|\,, \tag{5.53}$$

where ℓ_{f_k} describes the relative angular momentum between the $\psi(2S)$ and f_k states introduced in Sec. 5.3.1. For the S-wave states with $J = 0$, the only allowed value is $\ell_{f_k} = 1$ resulting in an odd CP parity. For the f_2 states the values $1, 2, 3$ are allowed though. These correspond to CP parities $-1, 1, -1$ respectively. In the default fit only the lowest $B_{\ell_{f_k}s}$ couplings are taken into account though. In this case the interference term in eq. 5.47 does not change the shape of the pdf, because all amplitudes have the same CP parity. The mixing effect in the decay chain with intermediate f_k therefore only needs to be taken into account in the systematic study, where higher ℓs couplings are allowed in Sec. 9.

In order to determine the amplitude for the full decay chain $\overline{B}{}^0_s \to \psi(2S)f_k$ with $f_k \to \pi^+\pi^-$ the effect of CP on $\left|B^0_s\right\rangle$

$$CP\left|B^0_s\right\rangle = P\left|\overline{B}{}^0_s\right\rangle = -\left|\overline{B}{}^0_s\right\rangle \tag{5.54}$$

needs to be considered as well. This can be combined with eq. 5.53 into

$$CP\langle\psi(2S)f_k|B^0_s\rangle = -\langle\psi(2S)f_k\left|\overline{B}{}^0_s\right\rangle, \tag{5.55}$$

or in the notation used in Sec. 5.3.1

$$A^{f_k}_{\lambda\xi}(M_{\pi\pi}, \cos\theta_\pi, \theta_\psi, \phi) = -\overline{A}{}^{f_k}_{\lambda\xi}(M_{\pi\pi}, \cos\theta_\pi, \theta_\psi, \phi) \tag{5.56}$$

which holds for both f_0 and f_2 intermediate states. At the level of the $B_{\ell_{f_k}s}$ couplings this means that

$$\overline{B}_{11} = -B_{11}, \tag{5.57}$$

if only the lowest $\ell_{f_k} s$ states are taken into account. If higher order couplings are allowed as well, the following relation holds

$$\begin{aligned} \overline{B}_{11} &= -B_{11} \\ \overline{B}_{22} &= B_{22} \\ \overline{B}_{33} &= -B_{33}. \end{aligned} \tag{5.58}$$

5.7.2 CP conjugate Amplitude for $B_s^0 \to Z(4430)^\pm \pi^\mp$

The determination of the *CP* conjugated decay amplitudes for the decays via intermediate $Z(4430)^+$ and $Z(4430)^-$ is less straightforward, because the intermediate states are not *CP* eigenstates.

For the decay amplitude of $Z(4430)^\pm \to \psi(2S)\pi^\pm$ *CP* transformation gives

$$CP\langle\psi(2S)\pi^-|Z^-\rangle = CP(\psi(2S))P(\pi)P(Z^+)(-1)^{\ell^Z} = (-1)^{\ell^Z+1}\langle\psi(2S)\pi^+|Z^+\rangle, \tag{5.59}$$

where $P(Z) = +1$ was used.

Since both allowed values for $\ell^Z \in [0,2]$ are even (see Sec. 5.3.2)

$$CP\langle\psi(2S)\pi^-|Z^-\rangle = -\langle\psi(2S)\pi^+|Z^+\rangle. \tag{5.60}$$

Assuming no direct *CP* violation in the decay (using the same argument as for eq. 5.49), the decay into the intermediate state $B_s^0 \to Z(4430)^+\pi^-$ $\langle Z^+\pi^-|B_s^0\rangle$ can be related to its *CP* conjugate via

$$\begin{aligned} CP\langle Z^+\pi^-|B_s^0\rangle =& P(\pi)P(Z)P(B_s^0)(-1)^{\ell^{B_s^0\to Z^+\pi^-}}\langle Z^-\pi^+|\bar{B}_s^0\rangle \\ =& (-1)^{\ell^{B_s^0\to Z^+\pi^-}}\langle Z^-\pi^+|\bar{B}_s^0\rangle. \end{aligned} \tag{5.61}$$

The only allowed value for the angular momentum $\ell^{B_s^0\to Z^+\pi^-}$ between the Z^+ and π^- in the B_s^0 decay is $\ell^{B_s^0\to Z^+\pi^-} = \vec{J}(B_s^0) - \vec{J}(\pi) - \vec{J}(Z) = 1$, since $J(B_s^0) = J(\pi) = 0$ and $J(Z) = 1$. Therefore

$$CP\langle Z^+\pi^-|B_s^0\rangle = -\langle Z^-\pi^+|\bar{B}_s^0\rangle. \tag{5.62}$$

The same consideration holds for the decay to $Z^-\pi^+$

$$CP\langle Z^-\pi^+|B_s^0\rangle = -\langle Z^+\pi^-|\bar{B}_s^0\rangle. \tag{5.63}$$

Eq. 5.63 can now be combined with eq. 5.60 to relate the full decay amplitudes from Sec. 5.3.2 to their *CP* conjugates

$$\begin{aligned} CPA^{Z^-}_{\lambda^-\xi} &= CP\langle Z^-\pi^+|B^0_s\rangle\langle\psi(2S)\pi^-|Z^-\rangle \\ &= (-1)^2\langle Z^+\pi^-|\bar{B}^0_s\rangle\langle\psi(2S)\pi^+|Z^+\rangle \\ &= \overline{A}^{Z^+}_{\lambda^+\xi} \end{aligned} \tag{5.64}$$

and for eq. 5.61

$$CPA^{Z^+}_{\lambda^+\xi} = \overline{A}^{Z^-}_{\lambda^-\xi}. \tag{5.65}$$

At the level of the ℓs couplings derived in Sec. 5.3.2

$$\begin{aligned} B^{Z^+}_{01} &= \overline{B}^{Z^-}_{01} \\ B^{Z^-}_{01} &= \overline{B}^{Z^+}_{01}, \end{aligned} \tag{5.66}$$

which means that an exotic contribution from $B^0_s \to Z(4430)^-\pi^+$ and $B^0_s \to Z(4430)^+\pi^-$ can be described with two independent complex couplings, which need to be fitted in the amplitude fit.

Given the limited data sample size, an additional constraint needs to be applied to the decay amplitudes in order to increase the fit stability. The limit-setting procedure in Sec. 10 is performed twice with different relative phases between the exotic couplings, corresponding to minimal mixing terms

$$\left\langle Z^-\pi^+|B^0_s\right\rangle = \left\langle Z^+\pi^-|B^0_s\right\rangle \tag{5.67}$$

or maximal mixing terms

$$\left\langle Z^-\pi^+|B^0_s\right\rangle = -\left\langle Z^+\pi^-|B^0_s\right\rangle. \tag{5.68}$$

Note that on both sides of the equation here is an initial B^0_s state.

For the $B_{\ell s}$ couplings this corresponds to the following relations

$$B^{Z^+}_{01} = B^{Z^-}_{01} = \overline{B}^{Z^-}_{01} = \overline{B}^{Z^+}_{01} \tag{5.69}$$

for minimal mixing or

$$B^{Z^+}_{01} = -B^{Z^-}_{01} = \overline{B}^{Z^-}_{01} = -\overline{B}^{Z^+}_{01} \tag{5.70}$$

for maximal mixing effects.

The addtional constraint decreases the number of free couplings related to exotic contributions to just one. These choices are similar to the choice made for the coupling parameters in the analysis of $B_s^0 \to J/\psi\, p\overline{p}$ in [10]. While the choice for the relative phase of the couplings is arbitrary, the assumption of the same magnitude for the couplings is motivated by isospin symmetry between $Z(4430)^+$ ($c\bar{c}u\bar{d}$) and $Z(4430)^-$ ($c\bar{c}\bar{u}d$) in the tetraquark picture. A priori there is no reason, why the six quarks $c\bar{c}\bar{u}du\bar{d}$ in the potentially exotic final state should preferably form a $Z(4430)^+\pi^-$ configuration over a $Z(4430)^-\pi^+$ configuration. This assumption can be relaxed, once a larger data sample is available for study.

5.7.3 Fit Fractions And Interferences

In a standard fit with normalised pdfs the contribution to the decay width via intermediate states would be described by the yield parameters, which are usually realized as prefactors of the normalised pdfs of the individual components. In an amplitude fit this is not possible directly, because the free parameters of the amplitude fit are not yields or fit fractions, but the ℓs-couplings $B^J_{\ell s}$. The contribution of decays via each individual intermediate hadronic amplitude therefore needs to be calculated separately.

For each intermediate state i that is allowed in the fit, the fit fraction is defined as

$$\mathcal{F}_i = \frac{\int d\Phi \sum_\xi \left|\sum_{\lambda'} A^i_{\lambda',\xi}\right|^2 + \left|\sum_{\lambda'} \overline{A}^i_{\lambda',\xi}\right|^2}{\int |\mathcal{A}|^2 + |\overline{\mathcal{A}}|^2}, \tag{5.71}$$

where λ' runs over the appropriate helicities for the decay chain and $A^i_{\lambda',\xi}$ and $\overline{A}^i_{\lambda',\xi}$ are the individual amplitudes for the intermediate state. The squared amplitudes $\mathcal{A}$ and $\overline{\mathcal{A}}$ are defined in eq. 5.35. The small correction from the oscillation effect described in eq. 5.47 is negligible with respect to the statistical uncertainties on the fit fractions and therefore not included here.

The integrals are not evaluated analytically, but via MC integration with uniform sampling (for an introduction to MC methods see *e.g.* [114]). A simulated sample of $B_s^0 \to \psi(2S)\pi^+\pi^-$ decays is used, which is generated without intermediate hadronic resonances in $\pi^+\pi^-$, $\psi(2S)\pi^+$ or $\psi(2S)\pi^-$. This sample is therefore distributed uniformly in phase space Φ. The integral can then be approximated with the sum over the integrand evaluated on the simulated events. The size of the simulated sample is > 15 times larger than the data sample size, thus the additional statistical uncertainty introduced by

this procedure is negligible compared to the statistical uncertainty from the fit to data.

The sum of these contributions is not necessarily unity, because of potential interference between the states, but in the case of negligible interference, the fit fractions $\mathcal{F}_i$ can be identified with the relative decay widths $\Gamma_i/\Gamma_{\text{tot}}$ of the decay $B_s^0 \to \psi(2S)\pi^+\pi^-$.

The interferences between states i and j can be quantified using

$$\mathcal{F}_{ij} = \frac{\int d\Phi \sum_\xi 2\mathcal{R}e\left(\sum_{\lambda'} A^i_{\lambda',\xi} A^{j*}_{\lambda',\xi}\right) + 2\mathcal{R}e\left(\sum_{\lambda'} \overline{A}^i_{\lambda',\xi} \overline{A}^{j*}_{\lambda',\xi}\right)}{\int |\mathcal{A}|^2 + |\overline{\mathcal{A}}|^2}. \tag{5.72}$$

Uncertainties on these values need to be propagated from the uncertainty estimates on the ℓs-couplings. In principle this could be done with Gaussian uncertainty propagation. But this approach does not take into account the additional statistical uncertainties introduced in the s-weighting process described in Sec. 4.6 and also does not correctly take into account the non-linear correlations between the fit parameters, which are pronounced for small fit fractions. In order to estimate the full statistical uncertainty on the fractions and interferences the bootstrapping approach is applied. It is described in Sec. 8.

6 Amplitude Fit to Data

In the following the fit of the amplitude model (derived in Sec. 5.7) to the s-weighted data (described in Sec. 4) is introduced. First the signal pdf and likelihood are explained in Sec. 6.1, the concrete realisation of the amplitude model is given in Sec. 6.3 and the fit results are given in Sec. 6.4, 6.5 and 6.6.

6.1 Signal pdf and extended likelihood

Signal pdf and likelihood are constructed following the procedure in [115]. After the s-weights are applied to the data set the sample corresponds to a pure signal sample and a parametrisation of a background component in the amplitude variables is not necessary any more. In general, the signal pdf $\mathcal{P}$ depends on the invariant dipion mass $M_{\pi\pi}$ and on the angular variables of the decay chain θ_π, θ_ψ, and ϕ. In the following these variables are summarised as ξ with the pdf taking the form:

$$\mathcal{P}(\xi|\boldsymbol{w}) = \frac{1}{I(\boldsymbol{w})}|M(\xi|\boldsymbol{w})|^2 \cdot \epsilon_{sig}(\xi) \cdot \phi(\xi) \tag{6.1}$$

$M(\xi)$ is the matrix element describing the complete decay process from the B_s^0 to the final-state particles π^+, π^-, μ^+, μ^-. ϵ_{sig} is the efficiency of selecting signal events, ϕ is the phase-space element and $\boldsymbol{w}$ is the set of model parameters. $I(\boldsymbol{w})$ is a normalisation factor:

$$I(\boldsymbol{w}) = \int |M(\xi|\boldsymbol{w})|^2 \cdot \epsilon_{sig}(\xi) \cdot \phi(\xi) \cdot d\xi \tag{6.2}$$

The amplitude fit minimises a weighted unbinned extended likelihood $\mathcal{L}_{ext}$, where the weights are the s-weights extracted from a fit to the B_s^0 candidate mass in Sec. 4.6. The following derivation comes from [115], which itself is based on [116–118]:

$$\mathcal{L}(\boldsymbol{w}, \nu)_{ext} \equiv \frac{\nu^N e^{-\nu}}{N!} \prod_{i=1}^{N} \mathcal{P}(\xi_i|\boldsymbol{w})^{s_i}, \tag{6.3}$$

where N is the number of observed events, a Poisson random variable with mean value ν and s_i is the s-weight associated to the ith candidate.

Using Equation 6.1, $\mathcal{L}(\boldsymbol{w},\nu)_{ext}$ takes the form:

$$\mathcal{L}(\boldsymbol{w},\nu)_{ext} = \frac{\nu^N e^{-\nu}}{N!}\prod_{i=1}^{N}\left[\frac{|M(\xi_i|\boldsymbol{w})|^2\cdot\epsilon_{sig}(\xi_i)\cdot\phi(\xi_i)}{I(\boldsymbol{w})}\right]^{s_i} \tag{6.4}$$

Taking the logarithm of Equation 6.4 and using Stirling's approximation [1], the extended weighted log-likelihood is given by:

$$\begin{aligned}\ln\mathcal{L}(\boldsymbol{w},\nu)_{ext} &= -N\ln N-\nu+\sum_i^N\ln\nu+\sum_i^N\ln\left[\frac{|M(\xi_i|\boldsymbol{w})|^2\epsilon_{sig}(\xi_i)\phi(\xi_i)}{I(\boldsymbol{w})}\right]^{s_i}\\ &= -N\ln N-\nu+\sum_i^N\ln\nu+\sum_i^N s_i\ln|M(\xi_i|\boldsymbol{w})|^2\epsilon_{sig}(\xi_i)\phi(\xi_i)\\ &\quad-\sum_i^N s_i\ln I(\boldsymbol{w})\end{aligned} \tag{6.5}$$

The expected number of total events ν is given by the sum of expected number of signal and background events:

$$\nu = \bar{n}_s+\bar{n}_b \tag{6.6}$$

with $\bar{n}_s$ being given by the normalisation integral:

$$\bar{n}_s = I(\boldsymbol{w}) = \int|M(\xi|\boldsymbol{w})|^2\epsilon_{sig}(\xi)\phi(\xi)d\xi \tag{6.7}$$

The term $\ln\nu$ in Equation 6.5 can be rewritten as:

$$\begin{aligned}\ln\nu &= \ln(\bar{n}_s+\bar{n}_b)\\ &= \ln\left[\bar{n}_s\cdot\left(1+\frac{\bar{n}_b}{\bar{n}_s}\right)\right]\\ &= \ln\bar{n}_s+\ln\left(1+\frac{\bar{n}_b}{\bar{n}_s}\right)\end{aligned} \tag{6.8}$$

Assuming that the ratio of the expectation values can be replaced by the ratio of observed events, as determined by the mass fit that produced the s-weights, *i.e.*

[1] $\ln 1/N! = -N\ln N+N$

$$\frac{\bar{n}_b}{\bar{n}_s} = \frac{n_b}{n_s} \tag{6.9}$$

and inserting Equation 6.8 in Equation 6.5, the extended weighted log-likelihood is given by

$$\begin{aligned} \ln \mathcal{L}(\boldsymbol{w}, \nu)_{ext} = & - N \ln N - \nu + \sum_i^N \left[\ln \bar{n}_s + \ln(1 + \frac{n_b}{n_s}) \right] \\ & - \sum_i^N s_i \ln I(\boldsymbol{w}) + \sum_i^N s_i \ln |M(\xi_i|\boldsymbol{w})|^2 \epsilon_{sig}(\xi_i) \phi(\xi_i). \end{aligned} \tag{6.10}$$

Using eq. 6.7 this expression can be reorganised into

$$\begin{aligned} \ln \mathcal{L}_{ext}(\boldsymbol{w}, \nu) = & -N \ln N - \nu + \sum_i^N \ln(1 + \frac{n_b}{n_s}) \\ & + \sum_i^N (1 - s_i) \ln I(\boldsymbol{w}) \\ & + \sum_i^N s_i \ln |M(\xi_i|\boldsymbol{w})|^2 \epsilon_{sig}(\xi_i) \phi(\xi_i) \\ = & -N \ln N - \nu + N \ln(1 + \frac{n_b}{n_s}) \\ & + \sum_i^N (1 - s_i) \ln I(\boldsymbol{w}) \\ & + \sum_i^N s_i \ln |M(\xi_i|\boldsymbol{w})|^2 + \sum_i^N s_i \, \epsilon_{sig}(\xi_i) \phi(\xi_i) \end{aligned} \tag{6.11}$$

It is important to note that the efficiency and phase-space terms do not depend on the model parameters, which all go into the matrix element $M(\xi_i|\boldsymbol{w})$. The terms which do not depend on $\boldsymbol{w}$ represent only an offset of the likelihood and do not change the position of the maximum of the log-likelihood in the parameter space. Thus they can be removed from Equation 6.11, leading to

$$\begin{aligned} - \ln \mathcal{L}_{ext}(\boldsymbol{w}, \nu) &= \nu - \sum_i^N (1 - s_i) \ln I(\boldsymbol{w}) - \sum_i^N s_i \ln |M(\xi_i|\boldsymbol{w})|^2 \\ &= \bar{n}_b + I(\boldsymbol{w}) - \sum_i^N (1 - s_i) \ln I(\boldsymbol{w}) \quad - \sum_i^N s_i \ln |M(\xi_i|\boldsymbol{w})|^2 \end{aligned} \tag{6.12}$$

where Equations 6.6-6.7 have been used. The expected number of background events $\bar{n}_b$ does not depend on the model parameters, which allows to use the observed number of background events n_b as an estimate ($\bar{n}_b = n_b$). The latter can be rewritten as

$$n_b = N - n_s = \sum_i^N (1 - s_i). \tag{6.13}$$

This can be brought to the final form of the negative extended and weighted log-likelihood, where we neglect the constant contribution n_b

$$\begin{aligned} -\ln \mathcal{L}_{ext}(\boldsymbol{w}, \nu) = I(\boldsymbol{w}) \quad & - \sum_i^N (1 - s_i) \ln I(\boldsymbol{w}) \\ & - \sum_i^N s_i \ln |M(\xi_i|\boldsymbol{w})|^2 \end{aligned} \tag{6.14}$$

In order to properly take into account the statistical uncertainty introduced by the weighting procedure, the negative log-likelihood is multiplied by a factor α, which scales it to the *effective sample size* [119]

$$-\ln \widehat{\mathcal{L}}_{ext}(\boldsymbol{w}, \nu) = -\alpha \ln \mathcal{L}(\boldsymbol{w}, \nu)_{ext} = -\frac{\sum_i^N s_i}{\sum_i^N s_i^2} \ln \mathcal{L}(\boldsymbol{w}, \nu)_{ext}, \tag{6.15}$$

that leads to larger uncertainties for non-uniform weights. This factor only holds for linear estimators, but is used as an approximation here. A more correct uncertainty estimate for the s-weighting procedure is addressed via bootstrapping in Sec. 8.

Now the normalisation integral $I(\boldsymbol{w})$ needs to be addressed. In general the efficiency $\epsilon(\xi)$ is not known in closed form, making it necessary to approximate it with *e.g.* polynomial shapes, which are fitted to the efficiency distribution of a simulated sample or a control data sample. Using the approximated efficiency, it is then possible to perform the integral $I(\boldsymbol{w})$. Since the matrix element $M(\xi|\boldsymbol{w})$ usually cannot be integrated analytically, a numeric approximation like Gaussian quadrature is used for the evaluation of $I(\boldsymbol{w})$ in that case. This needs to be repeated for each step of the optimisation process, because the matrix element also depends on the model parameters $\mathbf{w}$, part of which are updated at each fit iteration. Since the fit might need hundreds of iterations to converge to a minimum, the numeric integration needs to performed as often, slowing down the fit process.

In this analysis a different approach is followed. Instead of first approximating $\epsilon(\xi)$ and then performing numerical integrations of the integral $I(\boldsymbol{w})$, $I(\boldsymbol{w})$ is evaluated directly using Monte Carlo integration with uniform sampling. An introductory lecture to Monte Carlo sampling techniques can be found *e.g.* in [114].

An integral $I = \int dx f(x)$, where $x = (u_1, ...u_d)$ is a point in a d-dimensional space can be approximated with

$$E = \frac{1}{N} \sum_{n=1}^{N} f(x_n), \tag{6.16}$$

where x_n is sampled uniformly from the d-dimensional hypercube. In the limit $N \to \infty$ it becomes identical to I.

It is therefore possible to approximate

$$I(\boldsymbol{w}) = \int |M(\xi|\boldsymbol{w})|^2 \cdot \epsilon_{sig}(\xi) \cdot \phi(\xi) \cdot d\xi \approx 1/N_{gen} \sum_{n}^{N_{\text{gen}}} |M(\xi_n|\boldsymbol{w})|^2 \cdot \epsilon_{sig}(\xi_n) \tag{6.17}$$

via sampling events uniformly from the phase-space distribution $\Phi(\xi)$.

The efficiency $\epsilon_{sig}(\xi)$ is approximated by the simulated events used in Sec. 4 as signal proxy to optimize the BDT. They were sampled uniformly from the allowed phase space $\Phi(\xi)$ and then treated the same way as measured data, which means that detector acceptance effects and kinematic and topological cuts are the same as for measured signal data. The efficiency at each generated point ξ_n can then be approximated as a binary expression

$$\epsilon_{sig}(\xi_n) \approx \sum_{i}^{N_{\text{acc}}} [\xi_n - \xi_i \equiv 0], \tag{6.18}$$

where the sum runs over all those generated events, which matched all selection criteria and the Iverson bracket $[\xi_n - \xi_i \equiv 0]$ evaluates to 1, if the condition in it is met (the generated event was accepted by the selection criteria), otherwise it evaluates to 0 (the generated event was not selected).

This approximation simplifies eq. 6.17 to

$$I(\mathbf{w}) \approx \frac{1}{N_{gen}} \sum_{n}^{N_{\text{acc}}} |M(\xi_n|\boldsymbol{w})|^2, \tag{6.19}$$

where the evaluation now only needs to be performed on the accepted simulated events, instead of all generated ones. In addition, no parametric description of the 4D efficiency distribution over the phase space needs to be introduced.

Instead of one computing-intensive numerical integration of the matrix element and efficiency distribution in every iteration of the fit, now only the matrix element needs to be evaluated at N_{acc} points in every iteration, which is significantly faster.

The simulated samples correspond to the data-taking conditions in each year and each magnet polarity. The relative sample sizes do not exactly correspond to the integrated luminosity, which was measured in data. Different data taking conditions can result in slightly different acceptances and efficiencies. In the amplitude fit's variables, which are measured in the rest frames of the particles, these effects are largely washed out, because of the different orientations of the decays. They still need to be taken into account. Since the pdf takes the acceptance and efficiency distributions directly from simulated samples, a wrong proportion of the individual subsamples from different run conditions could lead to a bias in the fit result. Additional weights are introduced to scale each simulated sample i with generated events N_i by the integrated luminosity $\mathcal{L}_i$ that was measured in these conditions

$$w_{i\text{lum}} = \mathcal{L}_i / \sum_j \mathcal{L}_j \frac{\sum_j N_j}{N_i}, \tag{6.20}$$

where the sum runs over all generated samples for the 6 years of data taking.

An additional effect, which needs to be taken into account, when scaling the simulated samples, is the production cross section of B_s^0 mesons, which was higher in Run 2 than in Run 1 due to the higher proton-proton collision energies. Because no direct measurement is available for the B_s^0 production cross section at these two energies, the measured B^+ production cross section ratios for 7 over 13 TeV are used as estimate for the B_s^0 cross section correction [120]. This assumption is justified by the results in [53], where the relative production cross sections of B_s^0 and B^0 were compared for 8 and 13 TeV respectively. These ratios $\frac{f_s}{f_d}(8\,\text{TeV})$ and $\frac{f_s}{f_d}(13\,\text{TeV})$ are compatible with each other within $1.5\,\sigma$. Since isospin is conserved to a good degree in the hadronisation, it can therefore be expected that the relative production ratios of B_s^0 and B^+ are compatible as well between 8 and 13 TeV. The correction is

implemented as a weight

$$w_{i\text{prod}} = \frac{f_s}{f_d}(i\,\text{TeV}) \cdot \frac{f_s}{f_d}(8\,\text{TeV}) / \frac{f_s}{f_d}(13\,\text{TeV}), \tag{6.21}$$

where i represents the beam collision energies, which were used in the generation of the sample.

The weights, which scale the simulated samples by the correct luminosity and production cross section can be multiplied to the events, which were used to correct the simulated distributions of event multiplicity and B_s^0 p_T w_{kin} (introduced in Sec. 3.2.4) via

$$w_\text{tot} = w_{kin} w_\text{lum} w_\text{prod} \tag{6.22}$$

Since they give a weight to each simulated event, they can be included straightforwardly into the normalisation integral via

$$I(\mathbf{w}) \approx \frac{1}{N_{gen}} \sum_n^{N_\text{acc}} |M(\xi_n|\boldsymbol{w})|^2 w_{n\text{tot}}. \tag{6.23}$$

6.2 Data and Efficiency Distributions

In order to correctly describe the data with the pdf of the amplitude fit, the acceptances and efficiencies need to be well-described. Sharp features in the efficiency distributions could lead to biases in the fit result, if they are not described appropriately in the efficiency model. In the chosen fit procedure this is done via directly including a fully simulated sample into the likelihood to avoid the problem of "too smooth" or inappropriate parametrisations. Nevertheless the distributions need to be verified and understood to check for problematic regions in the phase-space, where a small mismodelling could lead to drastically different shapes. This would especially be true, if it occurred in a region, where also the fit model shows large variations, *e.g.* around the $f_0(980)$ peak in $M_{\pi\pi}$. If such a region is identified, a different selection might be necessary, which results in a smaller dependence on the variable, reducing the effect on the result.

For a visual inspection, the combined efficiency and acceptance projections to the amplitude variables are shown in Fig. 6.1. They are produced using fully simulated, phase space distributed $B_s^0 \to \psi(2S)\pi^+\pi^-$ events introduced in Sec. 3.2.4.

The reference distribution is given by a sample uses a sample that is generated flat in phase space as well, but without any acceptance or detector effects. Since their underlying models are the same, the combined efficiency and acceptance effects in Fig. 6.1 can be calculated directly from the bin-wise ratio of the distribution from the fully simulated sample and the reference distribution.

The efficiency shows a steep drop around $\cos\theta_\pi = \pm 1$. This can be explained from the configuration of the momenta in these extreme cases: Since there are a large number of low-p_T pions coming from the primary interaction forming the combinatoric background (see Sec. 4), the cuts on the pion p_T in the selection are tight, in order to have a high signal purity. For $\cos\theta_\pi = \pm 1$ one of the pions flies in the opposite direction of the flight direction of the $\pi\pi$ system, which means that this pion has a low momentum in the B_s^0 system and therefore in the lab frame as well. If that is the case, it is more likely that the pion does not fulfil the p_T requirement of the selection and therefore the efficiency is lower in that region.

The $\psi(2S)$ part of the decay is much cleaner and very few dimuon combinations are formed from random muon tracks. Because of this the p_T cuts on the muons are much looser, which makes the efficiency in $\cos\theta_\psi$ much flatter than its pion counterpart.

The efficiency distribution in $M_{\pi\pi}$ can be explained as well by the p_T cuts on the pions: For very low values of $M_{\pi\pi}$ the pions are emitted with low momenta in the $\pi\pi$ system and no cancellation of momenta occurs. Their momenta in the B_s^0 system are rather high though, because $\psi(2S)$ and $\pi\pi$ are emitted with maximum momentum. For very large values of $M_{\pi\pi}$ on the other hand the $\pi\pi$ system has only a small momentum in the B_s^0 system, while in the $\pi\pi$ system the pions are emitted with large momenta. In both these configurations, low and high $M_{\pi\pi}$, the pion momenta are high in the B_s^0 frame, which increases the chance of large p_T in the lab frame leading to a larger efficiency.

The distribution in ϕ can be explained via the acceptance of the detector: If the decay planes of $\psi(2S)$ and the $\pi\pi$ system lie perpendicular ($\phi = \pm\pi/2$), it is more likely that one of the particles leaves the instrumented region of the detector, because the 4 tracks extend over a larger solid angle of the detector.

For very small values of $M(\psi(2S)\pi^+)$, the π^+ is almost at rest in the B_s^0 frame, reducing its possible momentum in the lab frame. For very large $M(\psi(2S)\pi^+)$, the spectator π^- is the one at rest. For $M(\psi(2S)\pi^-)$ the opposite holds. These events with low pion momenta are more likely to be

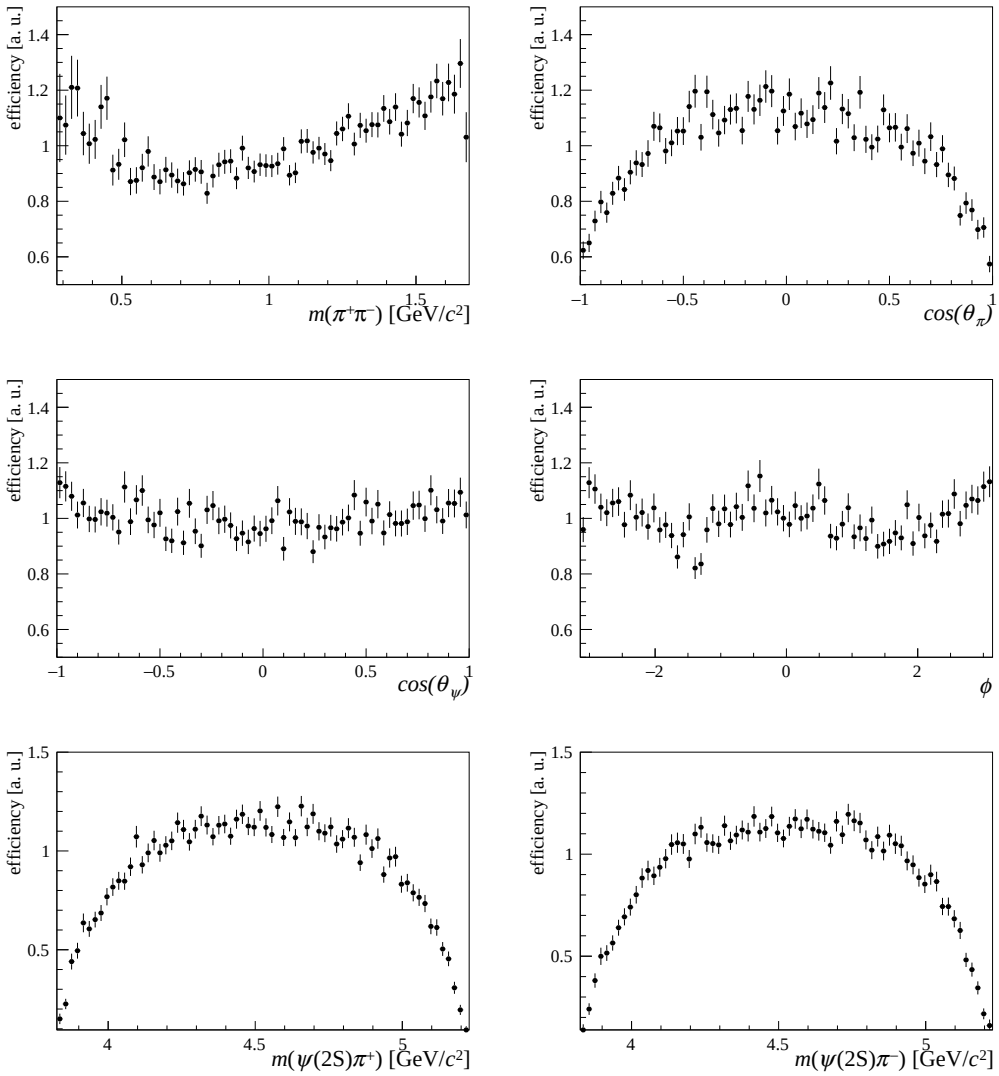

Figure 6.1: Combined efficiency and acceptance distributions for amplitude fit variables. Obtained from fully simulated $B_s^0 \to \psi(2S)\pi^+\pi^-$ events.

removed by the pion p_T cut in the selection, reducing the efficiency in this phase-space region.

The signal component of the data obtained with the *s*-weight technique is shown in Fig. 6.2 projected into the amplitude fit variables $M_{\pi\pi}$, θ_π, θ_ψ and ϕ, as well as the dependent variables $M(\psi(2S)\pi^+)$ and $M(\psi(2S)\pi^-)$. The data is dominated by the $f_0(980)$ resonance and to a lesser extent the $f_0(1500)$ resonance. No peaking structures are visible by eye in the $\psi(2S)$ π masses, where exotic contributions would occur.

This is supported by the Dalitz plot in Fig. 6.3, where only bands in $M^2_{\pi^+\pi^-}$ are visible. The $Z(4430)^\pm$ would be visible as a horizontal band around 20 GeV in $M^2(\psi(2S)\pi)$. A toy sample with a large $Z(4430)^\pm$ fraction is shown in

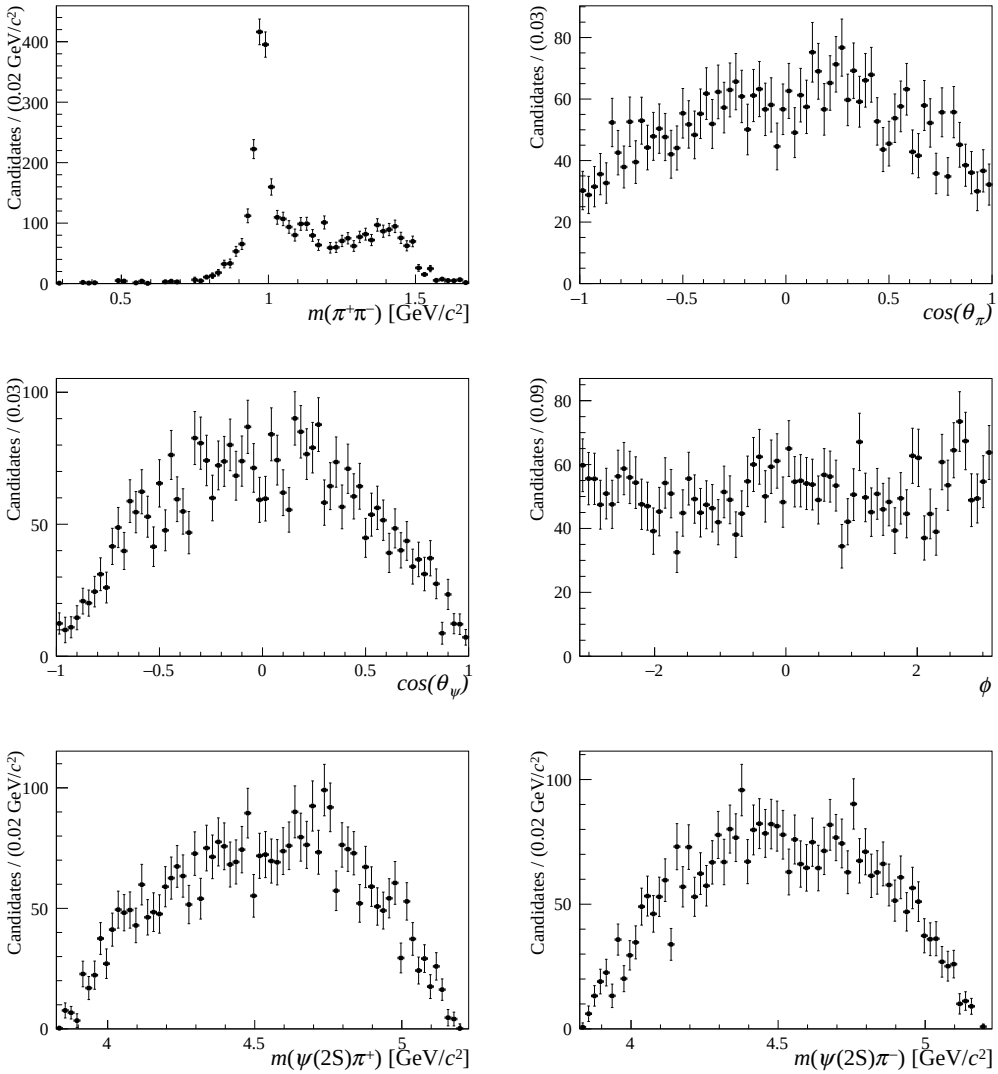

Figure 6.2: Signal component of the $B_s^0 \to \psi(2S)\pi^+\pi^-$ data obtained with the *s*-weight technique for amplitude fit variables. These are not efficiency or acceptance-corrected.

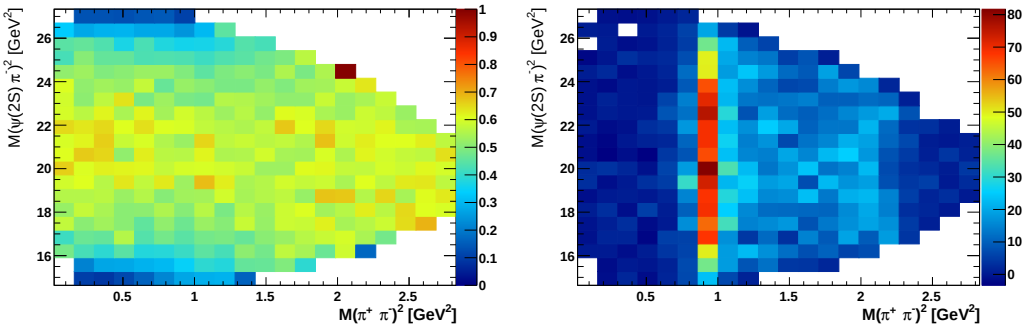

Figure 6.3: Left: Relative acceptance and efficiency as function of Dalitz plane (arbitrary scale). Right: signal component of the $B_s^0 \to \psi(2S)\pi^+\pi^-$ data obtained with the *s*-weight technique in Dalitz plane.

Fig. 10.1 in Sec. 10 Its absence in Fig 6.3 cannot be explained with the efficiency, which is flat in this region of the Dalitz plane, as can be seen in the efficiency map on the left of Fig. 6.3. This is quantified with a dedicated search for the $Z(4430)^{\pm}$ in these spectra in Sec. 10.

6.2.1 Resolution Effects

Due to the finite detector resolution and the reconstruction effects the amplitude fit variables can only be measured with a limited precision. These resolution effects can distort the shapes of the distributions that were introduced in Sec. 5 and must therefore be studied and (if necessary) described by the fit model as well.

Since the simulated sample includes all detector effects, it can be used to directly determine the distribution of resolution effects by comparing the simulated values at generator level with the reconstructed values for the fit variables of concern. The distributions of these differences are shown in Fig. 6.5 and can be used to directly estimate the size of the resolution effects that need to be considered. The reconstructed variables are the output of the decay tree fit as described for $M(\psi(2S)\pi^+\pi^-)$ in Sec. 4.4.1: It constrains the B_s^0 to originate from the PV, fix the dimuon invariant mass to the PDG value of the $\psi(2S)$ mass. In addition here the B_s^0 mass is fixed to its PDG value as well. This improves the resolution significantly towards the edges of the phase-space of the decay, *e.g.* where $M_{\pi\pi} \approx m(B_s^0) - m(\psi(2S))$.

The angular resolution is better than 8 mradfor all angles, far below the visible effect of possible structures, making a resolution treatment unnecessary[2].

[2]The highest spin states considered in the fit have Spin 2, which leads to Legendre polynomials of up to order 4 in the angles. This can be seen from the Wigner-d matrix in eq. 5.13,

The resolution of $m_{\psi\pi}$ is with around 2.5 MeV also well below the measured width of the $Z(4430)^+$, which is around 180 MeV.

For $M_{\pi\pi}$ a small effect is possible for the core region of the $f_0(980)$ resonance, where the resolution is around 3 MeV. Its effect is treated as a systematic uncertainty in Sec. 9. This effect is evaluated using a convolution integral in the pdf

$$\mathcal{P}_R(\xi|\boldsymbol{w}) = \int \mathcal{P}(\xi|\boldsymbol{w}')R(\boldsymbol{w}-\boldsymbol{w}',\boldsymbol{w})d\boldsymbol{w}'. \tag{6.24}$$

The resolution R is modelled using the simulated sample following the procedure in [121]. The convolution integral is numerically approximated using the approach detailed in [115]: Each data point l in $M_{\pi\pi}$ is replaced with $n_{sam} = 19$ points with equidistant mass values within $\pm 3\sigma$ around it:

$$\overline{M}_{\pi\pi l,j} = M_{\pi\pi l} + j \cdot \frac{3\sigma(M_{\pi\pi l})}{n_{sam}-1} \tag{6.25}$$

with $j = -\frac{n_{sam}-1}{2}, .., 0, .., \frac{n_{sam}-1}{2}$ and $\sigma(M_{\pi\pi l})$ being the mass resolution value at mass $M_{\pi\pi}$ extracted from simulation.

Each sampled point is assigned a Gaussian weight $w_{l,j}$:

$$w_{l,j} = C\Delta_j G(\overline{M}_{\pi\pi l,j}|M_{\pi\pi l},\sigma(M_{\pi\pi l})) \tag{6.26}$$

where C provides the correct normalisation of the weights and Δ_j corresponds to the step size between sampling points.

$\sigma(M_{\pi\pi l})$ gets smaller towards the edges of the phase-space due to kinematic effects and the mass constraints from the B_s^0, $\psi(2S)$, π^+, and π^-. This is accounted for via extracting the resolution in 30 bins over the allowed $M_{\pi\pi}$ range. From each bin the standard deviation of the distribution $M_{\pi\pi\text{true}} - M_{\pi\pi\text{measured}}$ is extracted via a Gaussian fit. The resulting resolution points are fitted with a fourth order polynomial, in order to estimate the resolution for arbitrary points. Both are shown in Fig. 6.4. For the determination of $\sigma(M_{\pi\pi l})$ the value given by the fit function is used.

For constant step sizes, Δ_j could be absorbed into C, but near the phase-space boundary the 3σ range can stretch outside the allowed mass range. When that is the case, the step size is reduced on one side such that the furthest point sits on the boundary. The resulting imbalance (on the side near

which is of the form $d^2_{\lambda 0}(\theta_\pi) = \frac{1}{2}\left(\cos^2\theta_\pi - 1\right)$ for $\lambda = 0$, which creates terms of up to order 4 in the squared amplitude. These are steepest around $\theta_\pi = \pm 1$, but even there the value of the polynomial changes by less than 5 % within 8 mrad. Given that this only affects contributions from f_2 states, which make up less than 5 % of the full decay width of $B_s^0 \to \psi(2S)\pi^+\pi^-$, this is deemed negligible.

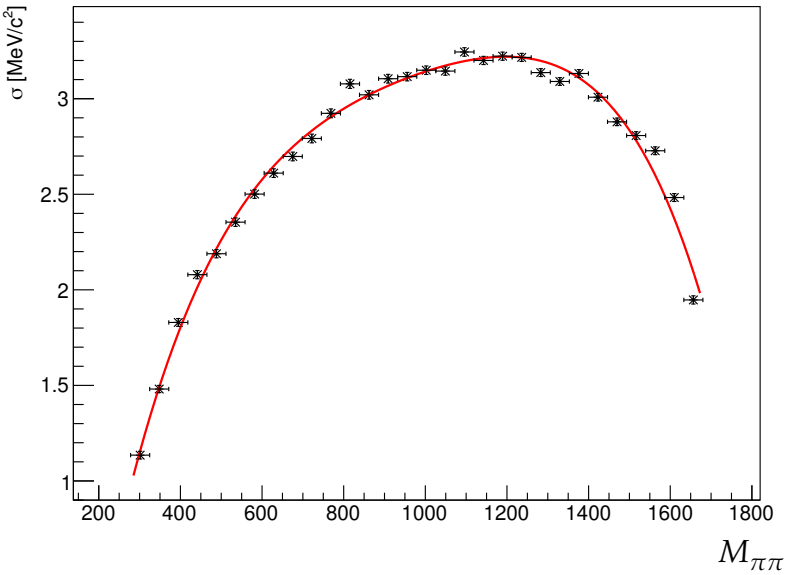

Figure 6.4: Resolution curve of $M_{\pi\pi}$ distribution evaluated from simulation, fitted with a fourth grad polynomial.

the boundary the sampling points lie nearer to the mode of the Gaussian) is countered with the factor Δ_j to avoid a bias towards the phase-space edge.

The resulting weighted data sample replaces the original data sample in the amplitude fit and thereby folds the resolution directly into the likelihood calculation given in eq. 6.12:

$$-\ln \mathcal{L}(\boldsymbol{w},\nu)_{ext} = \int |M(\xi|\boldsymbol{w})|^2 \epsilon_{sig}(\xi)\phi(\xi)d\xi - \sum_l^N (1-s_l)\ln \int |M(\xi|\boldsymbol{w})|^2 \epsilon_{sig}(\xi)\phi(\xi)d\xi - \sum_l^N s_l \ln \sum_j^{n_{sam}} w_{l,j}|M(\xi_{l,j}|\boldsymbol{w})|^2 \quad (6.27)$$

It is important to note that the s-weight s_l and the resolution weights $w_{l,j}$ cannot just be combined into a single weight, since they enter the expression at different places.

Due to the larger data sample the computing time for fits increases significantly, when the resolution effect is taken into account. On the other hand the inclusion of resolution effects has marginal impact on the extracted quantities. Therefore the nominal fit does not include these resolution effects, but the systematic uncertainty due to their neglect is determined in Sec. 9.

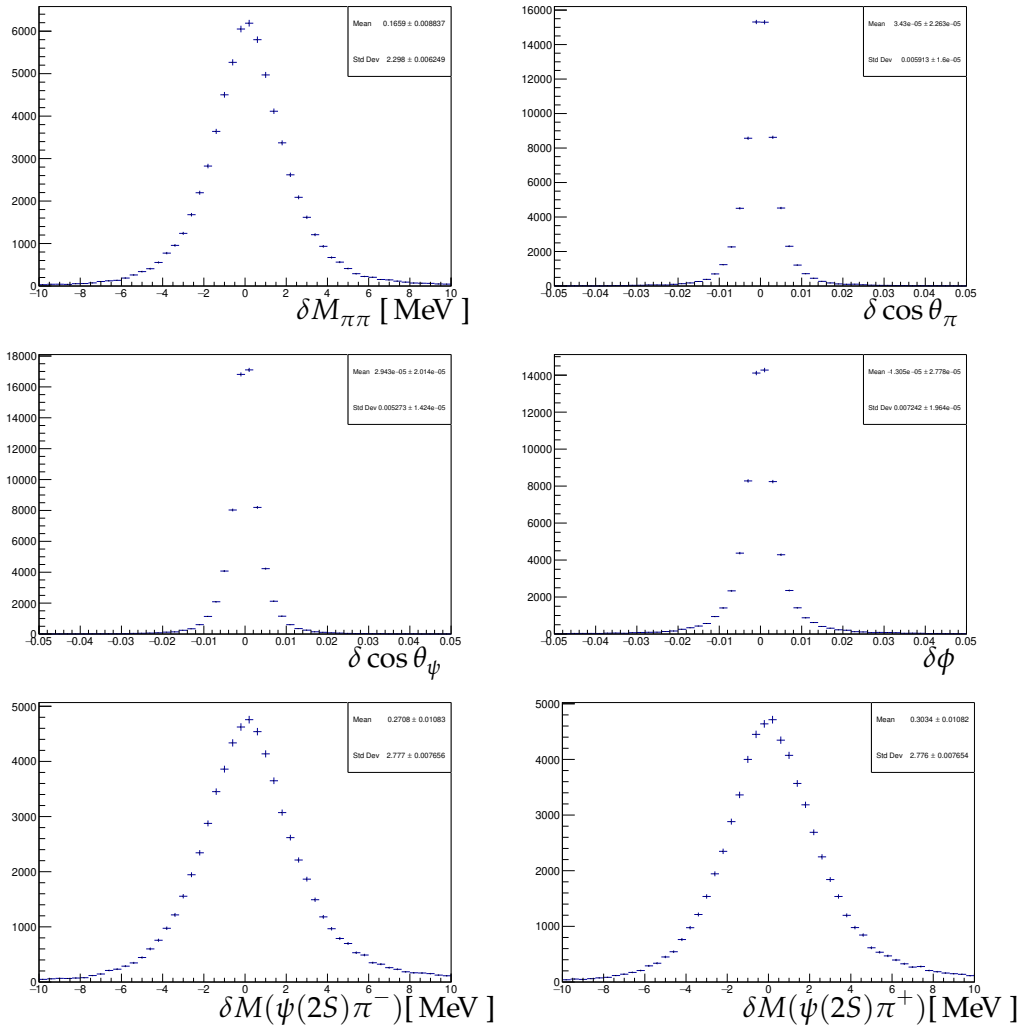

Figure 6.5: Distribution of resolution in $B_s^0 \to \psi(2S)\pi^+\pi^-$ for amplitude fit variables taken from simulation of Run II conditions. Top Left: $\delta M_{\pi\pi}$ in MeV. Top Right: $\delta\cos\theta_{\pi}$. Middle Left: $\delta\cos\theta_{\psi}$. Middle Right: ϕ. Bottom Left: $\delta M(\psi(2S)\pi^-)$ in MeV. Bottom Right: $\delta M(\psi(2S)\pi^+)$ in MeV.

6.3 Fit Model

The known resonances included into the fit are summarized in Tabs. 6.1, 6.2, 6.3 for each fit configuration.

Resonance	Spin	description	Mass (GeV)	Width(s) (GeV)	Source
$f_0(980)$	0	Flatté	0.9454	0.167/0.579 [3]	LHCb [56]
$f_0(1500)$	0	BW	1.4609	0.124	LHCb [56]
$f_0(2020)$	0	BW	1.814	0.328	LHCb [56]
$f_2(1270)$	2	BW	1.2751	0.1851	LHCb [56]
$f_2(1525)$	2	BW	1.522	0.0814	LHCb [56]

Table 6.1: Resonances used in the amplitude fit to $B_s^0 \to \psi(2S)\pi^+\pi^-$ with shape parameters fixed to results from [56] Solution I.

Resonance	Spin	description	Mass (GeV)	Width (GeV)	Source
$f_0(980)$	0	Flatté	0.9454	free	
$f_0(1500)$	0	BW	free	free	
$f_2(1270)$	2	BW	1.2751	0.1851	LHCb [56]
$f_2(1525)$	2	BW	1.522	0.0814	LHCb [56]

Table 6.2: Resonances used in the amplitude fit to $B_s^0 \to \psi(2S)\pi^+\pi^-$ and their parameters as used in the approach with Flatté and Breit-Wigner amplitudes for the S-wave description with freely floating shape parameters. The mass parameter of the Flatté is fixed, because the Flatté amplitudes are ambiguous, if both mass and width parameters are left free in the fit, resulting in a underdetermined covariance matrix of the fit as demonstrated in [122].

In the ansatz using Flatté/Breit-Wigner amplitudes for the description of the S-wave, the $f_0(980)$ is described with a Flatté shape (introduced in Sec. 5.4). In the first fit configuration, all shape parameters are fixed to the values obtained in [56] Solution I. These values are given in Tab. 6.1. In the second fit configuration, the two widths are left floating. The observed mass parameter lies within the region of approximate ambiguity that is demonstrated in [122]. In this region of parameter space, the 3 shape parameters $m_{0f_0(980)}, \Gamma_1, \Gamma_2$ are not independent and different combinations of the three result in very similar amplitudes. In order to resolve this ambiguity, the mass parameter is fixed in the fit.

[3] corresponding to Γ_1 and Γ_2 of the Flatté parametrisation given in eq. 5.32.

Resonance	Spin	description	Mass / pole position (GeV)	Width (GeV)	Source
$f_0(980)$	0	Bonn-Jülich	$0.996 + 0.057i$		[1]
$f_0(1500)$	0	Bonn-Jülich	$1.465 + 0.100i$		[1]
$f_0(2020)$	0	Bonn-Jülich	$1.910 + 0.398i$		[1]
$f_2(1270)$	2	BW	1.2751	0.1851	LHCb [56]
$f_2(1525)$	2	BW	1.522	0.0814	LHCb [56]

Table 6.3: Resonances used in the amplitude fit to $B_s^0 \to \psi(2S)\pi^+\pi^-$ and their parameters as used in the approach with Bonn-Jülich model for the S-wave description [1].

The shape of the $f_0(1500)$ is parametrized with a Breit-Wigner amplitude. In the first fit its shape parameters are fixed to the values obtained in [56] Solution I. In the second fit configuration the shape parameters $\Gamma_{0f_0(1500)}$ and $m_{0f_0(1500)}$ are floating freely.

The $f_0(2020)$ amplitude is only included in the first fit with Flatté/Breit-Wigner amplitudes. Its shape parameters are fixed there to the values obtained in [56] Solution I. In the second fit configuration is was not possible to include the $f_0(2020)$ component as well and obtain a meaningful fit result, even when keeping the shape parameters of the $f_0(2020)$ fixed while only floating those of the $f_0(1500)$ and $f_0(980)$.

In the fit with the Bonn-Jülich model, the S-wave components are all described as part of the model introduced in Sec. 5.6, where their shapes and relative phases are fixed from a fit to $B_s^0 \to J/\psi\,\pi^+\pi^-$ and $B_s^0 \to J/\psi\,K^+K^-$ [1]. The pole positions are given in Tab. 6.3 .

The decays via the two intermediate f_2 states contribute less than 5 % of the decay width of $B_s^0 \to \psi(2S)\pi^+\pi^-$. Since the $f_2(1270)$ and $f_2(1525)$ are separated from each other and do not lie directly at thresholds (see Sec. 2.4.1 for more about the shortcomings of Breit-Wigner parametrisations), their shapes can be approximated with Breit-Wigner functions and their shape parameters are fixed to those obtained in [56]. Their parametrisation is the same for all amplitude fits described in this section.

As for all quantum mechanical measurements, the amplitude fit is not sensitive to a global phase. This ambiguity in the phase would make the fit underdetermined, therefore an arbitrary global phase is chosen by setting the imaginary part of $B_{11}^{f_0(980)}$ to 0 (for the Bonn-Jülich model this is done to the B_{11}^{S} coupling of the S-wave component).

Three different fit configurations are presented in the following differing by their dipion S-wave description. The first uses the shape parameters

extracted in [56]. The second lets the shape parameters float freely in the fit. The third describes the S-wave with the model introduced in [1], where all shape parameters are fixed as described in Sec. 5.6.

The total number of real parameters that describe the S-wave in these fit configurations are therefore

- 5 for the fit with the Flatté/BW with fixed shape parameters: The real part of $B_{11}^{f_0(980)}$ and the complex parameters $B_{11}^{f_0(1500)}$ and $B_{11}^{f_0(2020)}$. It uses the shape parameters extracted from $B_s^0 \to J/\psi\,\pi^+\pi^-$ in [56]. This fit is only performed to display the process-dependence of Breit-Wigner and Flatté parameters for overlapping resonances. It is detailed in Sec. 6.4. The fixed parameters are shown in Tab. 6.1. They use Solution I from [56], but similar results are obtained with Solution II.

- 7 for the fit with the Flatté/BW model with no external shape parameters: The real part of $B_{11}^{f_0(980)}$, $\Gamma_{KK}^{f_0(980)}$, $\Gamma_{\pi\pi}^{f_0(980)}$, the complex $B_{11}^{f_0(1500)}$, $\Gamma^{f_0(1500)}$, and $m_0^{f_0(1500)}$. The results are shown in Sec. 6.5.[4] The fixed parameters are shown in Tab. 6.2.

- 3 for the fit with the Bonn-Jülich model: The real part of B_{11}^{S} and the resonance-source couplings $\alpha_{f_0(1500)}$ and $\alpha_{f_0(2020)}$. The results are shown in Sec. 6.6. The fixed parameters are shown in Tab. 6.3.

The Flatté/BW model with free shape parameters has 4 nonlinear parameters that describe the resonance shapes, which are highly correlated among each other and make the fit less stable. These are model-inherent and cannot be circumvented, unless the data sample size is increased. Due to the correctly propagated external inputs into the Bonn-Jülich model, it needs fewer fit parameters (the same number as the BW/Flatté model with fixed resonance shapes). And those parameters that need to float are only real couplings, since all relative phases are already fixed. This results in a better convergence behaviour of the fit.

All fits described in the following are performed using algorithms from the `Minuit` package [123]. `MIGRAD` is used to optimize the fitted likelihood and `HESSE` to estimate the parameter uncertainties. In complex fits there can be several local minima of the likelihood. The algorithm `MIGRAD` cannot find global minima by itself (it only follows the local gradient of the likelihood as function of the fit parameters) and therefore sometimes converges to a

[4] It is not possible to include the $f_0(2020)$ resonance as well, because the fit parameters do not converge to sensible values in that case, even if the shape parameters of the $f_0(2020)$ are fixed to the values obtained in [56].

non-optimal parameter-configuration. Each fit is therefore repeated with random start values and step sizes and the result with the best likelihood and fit status is taken as the result. Because of the different start values for the optimization procedure, the fit sometimes finds a different, better minimum of the likelihood.

6.4 Amplitude Fit with Breit-Wigner/Flatté amplitudes with Fixed Shape Parameters

As a first step, the amplitude fit is performed with all shape parameters fixed to those from a measurement of the resonant components in $B_s^0 \to J/\psi\,\pi^+\pi^-$ [56] as given in Tab. 6.1. The different charmonia resonances have the same quantum numbers, so the two channels only differ by the size of the phase-space corresponding to the different masses of the charmonia ($m_{\psi(2S)} = 3686\,\text{MeV}$, $m_{J/\psi} = 3096\,\text{MeV}$). If the shape parameters from fits with Breit-Wigner and Flatté amplitudes to $B_s^0 \to J/\psi\,\pi^+\pi^-$ were process-independent, it should therefore be possible to use them to describe the decay $B_s^0 \to \psi(2S)\pi^+\pi^-$. It is shown in the following that this is not the case due to the large overlap between the S-wave states in these channels.

The results for the 9 free parameters and their uncertainty estimates from the `HESSE` algorithm are shown in Tab. 6.4. They correspond to the complex ℓs-couplings of $f_0(1500)$, $f_0(2020)$, $f_2(1270)$, and $f_2(1525)$ in addition to the real part of the ℓs-coupling of $f_0(980)$[5].

The fit converged and the corresponding fractions and interferences as defined in Sec. 5.7 are shown in Tab. 6.5, 6.6. The sum of the fit fractions show a large S-wave component and a significant contribution from the $f_0(2020)$, whose pole lies significantly outside the allowed mass region in $M_{\pi\pi}$. Only its low-mass tail can therefore contribute to the decay amplitude. Large relative interferences are observed between all fit components. As shown in Sec. 2.4.1 interferences between Breit-Wigner amplitudes conflict with the unitarity condition on the decay amplitudes and implies violation of probability conservation. The measured fit fractions can therefore not be interpreted as relative branching fractions of $B_s^0 \to \psi(2S)\pi^+\pi^-$. The violation

[5] Note that the individual components are not normalised, which means that the couplings of resonances with different shape parameters differ by an arbitrary factor (see Sec. 6.1, [6]. The coupling parameters for the same resonance have the same scaling factor though, allowing to determine the resonance's polarisation from their relative sizes.

of unitarity is also a sign for the process dependence of the mass and width parameters of the parametrisations.

While the fitted fractions for the percent-level contributions of the f_2 states are compatible with those obtained in $B_s^0 \to J/\psi\,\pi^+\pi^-$ in [56], this is not the case for the S-wave fractions, which differ by factors.

Projections of data and fit model are shown in Fig. 6.6. The fit model histograms are sampled from the optimized pdf using fully simulated $B_s^0 \to \psi(2S)\pi^+\pi^-$ events in order to automatically describe the efficiencies and acceptances correctly. Due to the limited size of the simulated sample, the statistical uncertainty of the bin contents of the model histogram is shown as error bars. The histogram below the plot shows the pull distribution g, which describes the difference between the value x_i of the data in bin i and the value of the model bin $\overline{x}_i$ in the corresponding bin in multiples of statistical uncertainty on the data $\sigma(x_i)$ and the model $\sigma(\overline{x}_i)$:

$$g = \frac{x_i - \overline{x}_i}{\sqrt{\sigma(x_i)^2 + \sigma(\overline{x}_i)^2}}. \tag{6.28}$$

For large number of entries in the bins and if both the data and the model are drawn from the same underlying distribution (*i.e.* if the fit model is correct), the pulls are Gaussian distributed with mean 0 and width 1 [124].

The projections of the angles can be excellently described by the model, as can be seen by the small values for the pull distributions. It is obvious though from the projection in $M_{\pi\pi}$ that the model is not able to capture this spectrum correctly. There are large deviations around the $f_0(980)$ peak, while the model gives systematically too large contributions at low values. This means that the parameters extracted from $B_s^0 \to J/\psi\,\pi^+\pi^-$ cannot be used to describe the the S-wave in $B_s^0 \to \psi(2S)\pi^+\pi^-$, which proves their process-dependence, as explained in Sec. 2.4.1.

6.5 Amplitude Fit with Breit-Wigner/Flattè with Free Shape Parameters

As demonstrated in Sec. 6.4 the shape parameters of the Breit-Wigner/Flatté $\pi\pi$ S-wave description are process-dependent. In order to describe the data

[7] The extended log likelihood described in Sec. 6.1 allows for the sum of the pdf components to be different from the fitted sample size, even when no interference effects are taken into account.

name	value
$\Im\left(B_{11}^{f_0(1500)}\right)$	$(4.02 \pm 0.33) \times 10^3$
$\Re\left(B_{11}^{f_0(1500)}\right)$	$(-6.3 \pm 0.4) \times 10^3$
$\Im\left(B_{11}^{f_0(2020)}\right)$	$(34.0 \pm 2.4) \times 10^3$
$\Re\left(B_{11}^{f_0(2020)}\right)$	$(-41 \pm 2) \times 10^3$
$\Re\left(B_{11}^{f_0(980)}\right)$	3650 ± 80
$\Im\left(B_{11}^{f_2(1270)}\right)$	$(3.6 \pm 0.7) \times 10^3$
$\Re\left(B_{11}^{f_2(1270)}\right)$	$(1.1 \pm 0.7) \times 10^3$
$\Im\left(B_{11}^{f_2(1525)}\right)$	$(-2.8 \pm 0.6) \times 10^3$
$\Re\left(B_{11}^{f_2(1525)}\right)$	$(2.8 \pm 0.4) \times 10^3$
$\log \mathcal{L}$	33842.07 converged

Table 6.4: Amplitude fit result for model with Flatté and Breit Wigners for the description of the S-wave component with shape parameters fixed to the values obtained in [56]. Coupling terms contain arbitrary scale factor, see footnote 5.

name	value
$f_0(1500)$	0.465
f0_2020	0.328
$f_0(980)$	0.316
$f_2(1270)$	0.018
$f_2(1525)$	0.03
Sum	1.157

Table 6.5: Amplitude fit fractions for model with Flatté and Breit Wigners for the description of the S-wave component with fixed shape parameters obtained in [56]. Since the fit describes the data poorly, the uncertainties are not propagated into the fraction estimates. The Sum over all amplitudes exceed 1 due to interference effects and the bad description of the data[7]

	$f2\_1270$	$f2\_1525$	$f0\_980$	$f0\_2020$	$f0\_1370$
$f2\_1270$	/	/	/	/	/
$f2\_1525$	-0.0141	/	/	/	/
$f0\_980$	0.002	-0.0005	/	/	/
$f0\_2020$	0.0021	-0.0002	0.0923	/	/
$f0\_1370$	0.0021	-0.0016	0.05	-0.243	/

Table 6.6: Amplitude fit interferences for model with Flatté and Breit Wigner functions for the description of the S-wave component with shape parameters fixed to the values obtained in [56]. Very large interferences between the resonances are observed.

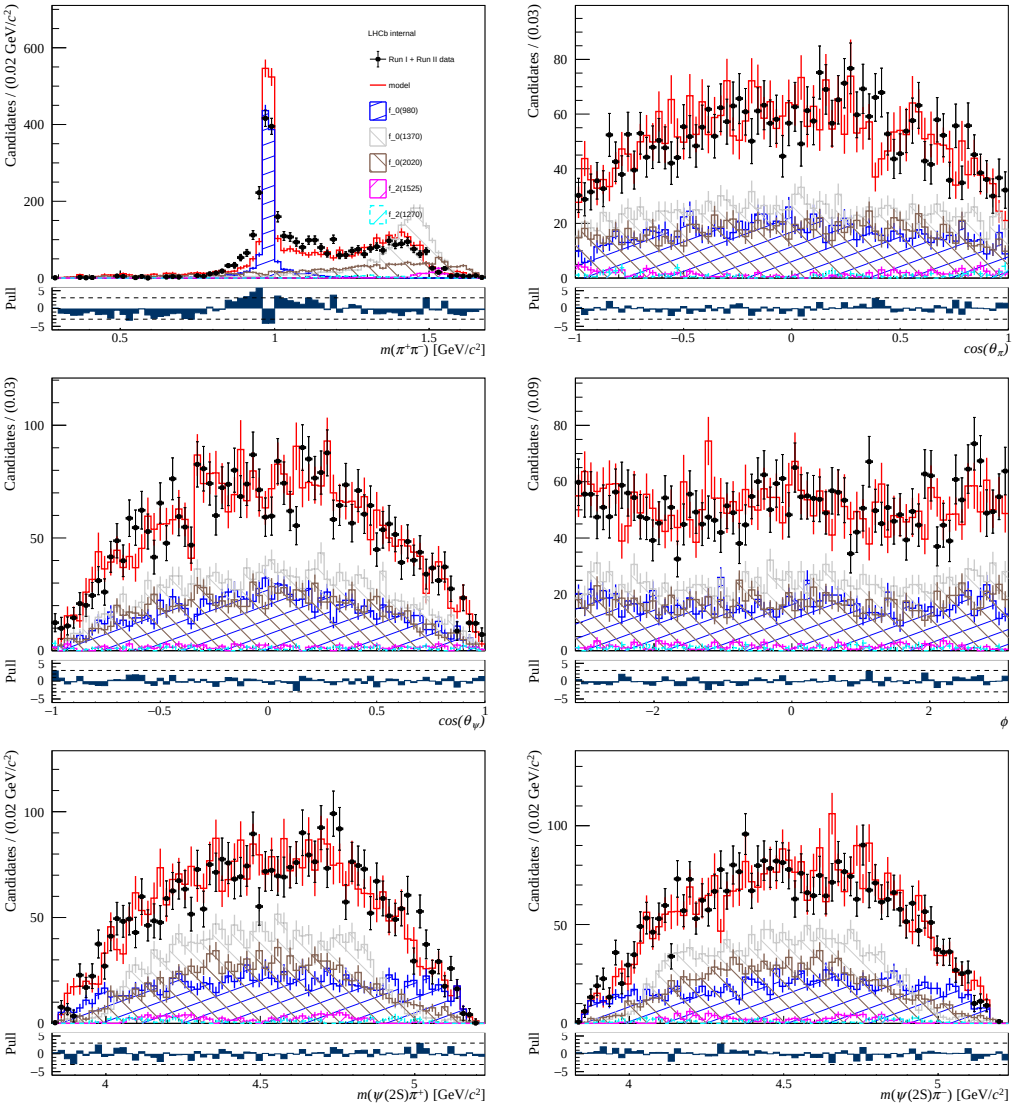

Figure 6.6: Amplitude fit projections using Flatté for the description of $f_0(980)$ and Breit-Wigner for $f_0(1500)$ with shape parameters fixed to the values obtained in [56].

sufficiently well with Breit-Wigner/Flatté amplitudes, these parameters are kept floating in this section. This comes with the price of reduced fit stability and larger uncertainties on the fit fractions, which is shown in the following.

The amplitude fit is repeated with free shape and phase parameters for the S-wave amplitude as described in Tab. 6.2. It was observed that fits including $f_0(2020)$ did not improve the fit quality, even when fixing the shape parameters of $f_0(2020)$ to the values obtained in [56] . The large number of free shape parameters and couplings could not be constrained sufficiently by the data making the fit unstable. Therefore the resonance $f_0(2020)$ was removed from the fit in order to increase the stability. This is sensible, because its pole lies significantly outside the allowed region for $M_{\pi\pi}$ in $B_s^0 \to \psi(2S)\pi^+\pi^-$. The results of the fit are shown in Tab. 6.7, fit fractions in Tab. 6.8. Randomized starting values, different step sizes and $f_0(980)$ mass parameter values have been tried, but due to the large number of free parameters and the small data sample size, no configuration was found that leads to a converging fit with positive definite covariance matrix. This means the quoted uncertainties on the fit parameters are incorrect.

The large number of free parameters allows for a good description of the data. But this large flexibility comes with the additional drawback of an ambiguity in the fit model: Two configurations can describe the data equally well. Solution 1 has the mass parameter $m^{s1}_{0f_0(1500)} = 1461.2\,\text{MeV}$ and the width parameter $\Gamma^{s1}_{0f_0(1500)} = 219\,\text{MeV}$, while solution 2 has $m^{s2}_{0f_0(1500)} = 1452\,\text{MeV}$ and $\Gamma^{s2}_{0f_0(1500)} = 186\,\text{MeV}$.

This additional solution can be a source of significant systematic uncertainty, not only on the fit fractions themselves, but also on other quantities, which should be determined from the data. A recent example are the *CP* observables in $B_s^0 \to J/\psi\,\pi^+\pi^-$, where the uncertainty from different fit solutions is one of the leading systematic uncertainties [125].

Both solutions show significant interference effects between the S-wave resonances, which violate the unitarity condition for decay amplitudes (see Sec. 2.4.1). The fit fractions can therefore not be interpreted as branching fractions and the shape parameters are process-dependent and cannot be interpreted as pole positions either.

While the mass parameters for $f_0(1500)$ in Solution 1 is compatible with that extracted in $B_s^0 \to J/\psi\,\pi^+\pi^-$ in [56] Solution II, this does not hold for the width parameters of the $f_0(980)$ and $f_0(1500)$ (compare Tabs. 6.1 and 6.7). This shows again the process dependence of Breit-Wigner and Flatté parametrisations for overlapping resonances. The fit fractions of $f_0(980)$,

name	Solution 1	Solution 2
$\Im\left(B_{11}^{f_0(1500)}\right)$	233 ± 11	459 ± 9
$\Re\left(B_{11}^{f_0(1500)}\right)$	-7800 ± 20	-8632 ± 16
$m_{0f_0(1500)}$	$(1452 \pm 4) \times 10^{-3}$	$(1461.2 \pm 2.5) \times 10^{-3}$
$\Gamma_{0f_0(1500)}$	$(186 \pm 6) \times 10^{-3}$	$(219.4 \pm 2.3) \times 10^{-3}$
$\Re\left(B_{11}^{f_0(980)}\right)$	12680 ± 40	12266 ± 27
$\Gamma_{1f_0(980)}$	$(368.0 \pm 3.5) \times 10^{-3}$	$(362.4 \pm 2.3) \times 10^{-3}$
$\Gamma_{2f_0(980)}$	$(666.5 \pm 0.9) \times 10^{-3}$	$(652.0 \pm 2.2) \times 10^{-3}$
$\Im\left(B_{11}^{f_2(1270)}\right)$	$(0.8 \pm 0.4) \times 10^{3}$	$(1.02 \pm 0.35) \times 10^{3}$
$\Re\left(B_{11}^{f_2(1270)}\right)$	$(-1.6 \pm 1.6) \times 10^{3}$	$(-3.0 \pm 0.8) \times 10^{3}$
$\Im\left(B_{11}^{f_2(1525)}\right)$	$(2.4 \pm 0.5) \times 10^{3}$	$(-0.3 \pm 0.5) \times 10^{3}$
$\Re\left(B_{11}^{f_2(1525)}\right)$	$(-0.3 \pm 0.5) \times 10^{3}$	$(1.36 \pm 0.32) \times 10^{3}$
$\log \mathcal{L}$	34002.66	34002.72

Table 6.7: Amplitude fit result for model with Flatté and Breit Wigners for the description of the S-wave component. The last row contains the $\log \mathcal{L}$ value. The fits did not result in positive-definite correlation matrix, which leads to incorrect uncertainty estimates. $B_{\ell s}$ terms contain arbitrary scale factor, see footnote 5.

$f_2(1270)$ and $f_2(1525)$ are comparable to the values obtained in [56] Solution I, but a quantitative statement is not possible without a reliable fit.

Apart from the conceptional problems from the violation of unitarity and the corresponding process dependence of the shape parameters, the ansatz using Flatté and Breit-Wigner amplitudes for the description of the dipion S-wave has also technical flaws. These flaws stem from the large number of free, non-linear parameters in the fit. The fits become prone to overfitting and can have unstable fit results, which is the reason for the non-positive definite covariance matrix in both fit solutions.

Even though the model can effectively describe the data, it is not possible to extract meaningful pole positions of the resonances or branching fractions for the individual resonances, since these are different for each measured process. In addition, the large number of non-linear, free parameters make the fit unstable for small datasets. It is therefore not a reasonable choice for the description of the dipion spectrum in $B_s^0 \to \psi(2S)\pi^+\pi^-$.

name	Solution 1	Solution 2
$f_0(1500)$	0.331	0.33
$f_0(980)$	0.752	0.755
$f_2(1270)$	0.013	0.004
$f_2(1525)$	0.004	0.012
Sum	1.099	1.101

Table 6.8: Amplitude fit fractions for model with Flatté and Breit Wigners for the description of the S-wave component. No uncertainties are quoted here for the fractions, since the uncertainty estimates on the corresponding fit parameters are incorrect.

		$f2$_1270	$f2$_1525	$f0$_980	$f0$_1370
Solution 1	$f2$_1270	/	/	/	/
	$f2$_1525	-0.0012	/	/	/
	$f0$_980	-0.0002	-0.0001	/	/
	$f0$_1370	0.0025	-0.0006	-0.0707	/

		$f2$_1270	$f2$_1525	$f0$_980	$f0$_1370
Solution 2	$f2$_1270	/	/	/	/
	$f2$_1525	-0.0033	/	/	/
	$f0$_980	0.0002	0.0013	/	/
	$f0$_1370	0.0014	-0.0005	-0.0702	/

Table 6.9: Amplitude fit interferences for model with Flatté and Breit Wigners for the description of the S-wave component. Top: Solution 1. Bottom: Solution 2.

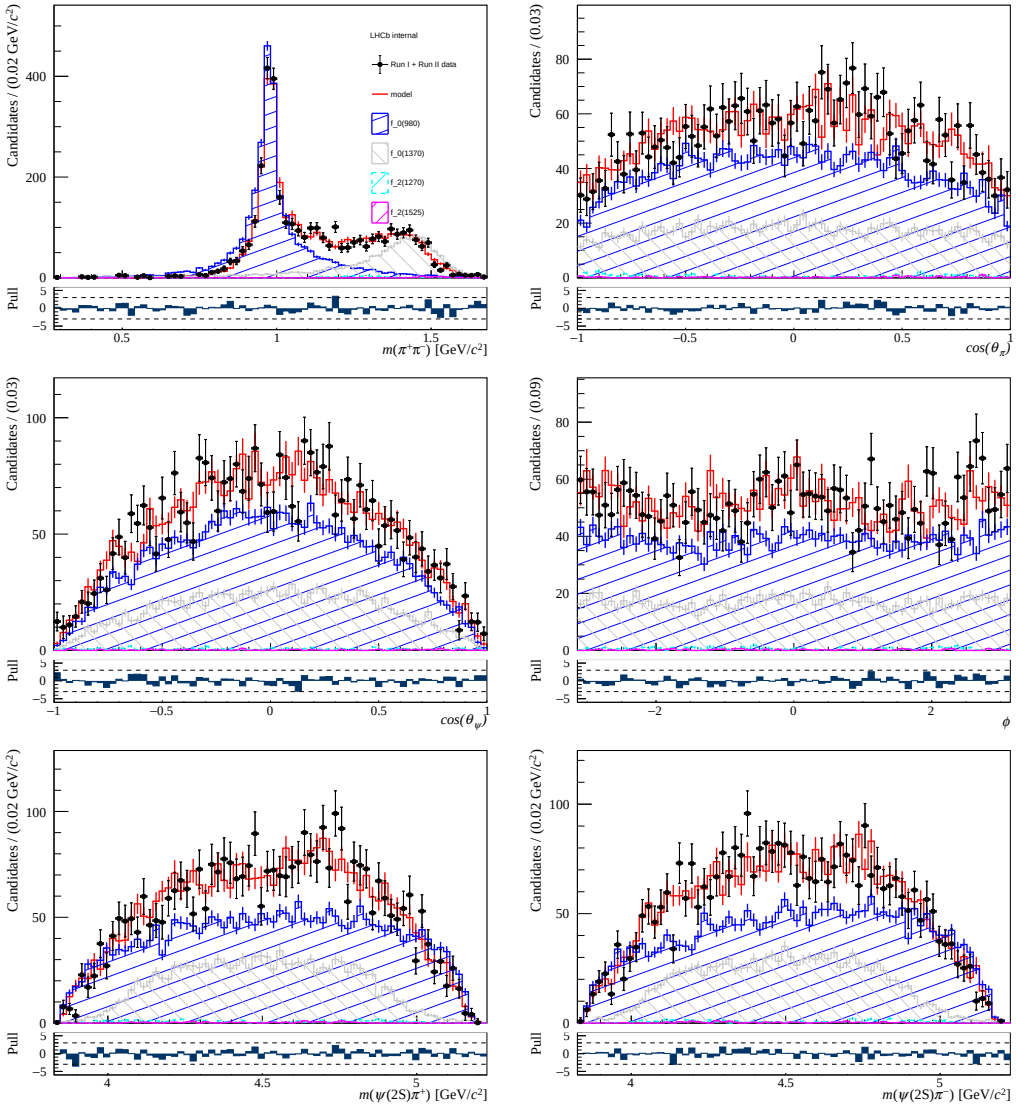

Figure 6.7: Amplitude fit projections using Flatté for the description of $f_0(980)$ and Breit-Wigner for $f_0(1500)$ (Solution 1). Pull distributions include statistical uncertainties on both data and model bins.

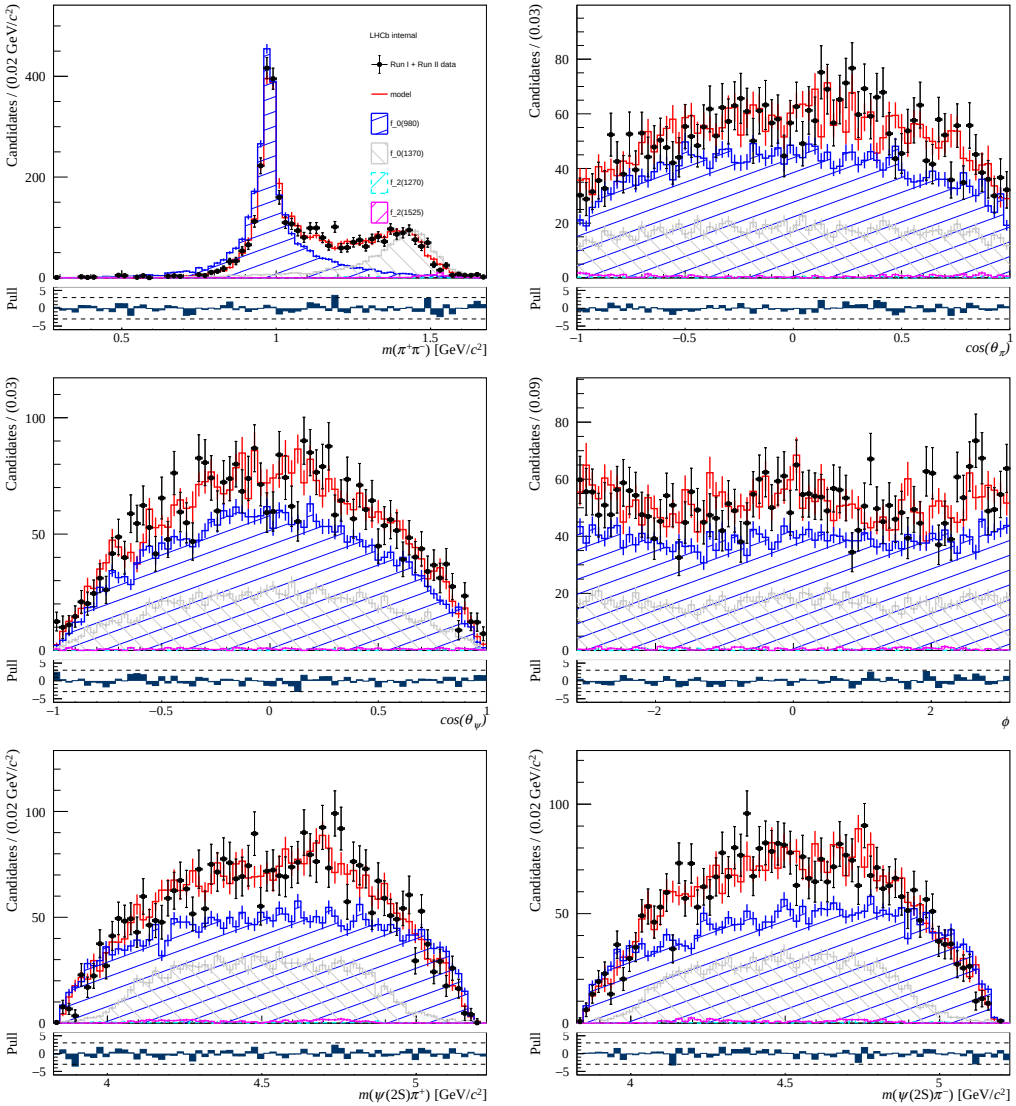

Figure 6.8: Amplitude fit projections using Flatté for the description of $f_0(980)$ and Breit-Wigner for $f_0(1500)$ (Solution 2). Pull distributions include statistical uncertainties on both data and model bins.

6.6 Amplitude Fit with Bonn-Jülich Model

The results of the fit with the Bonn-Jülich model for the S-wave component are shown in Tab. 6.10. In contrast to the model with Flatté and Breit-Wigner resonance parametrisations, it converged and the uncertainty estimates are reliable. The amplitude model has two solutions with similar likelihood values. These do not stem directly from different solutions for the S-wave amplitude, but they can be attributed to a phase flip in the couplings of the f_2 states to which the fit is not sensitive, the imaginary parts $\Im\left(B_{11}^{f_2(1270)}\right)$ and $\Im\left(B_{11}^{f_2(1525)}\right)$ have different signs in the two solutions.

The corresponding fractions and interferences are shown in Tab. 6.11 and 6.12. They are not listed for each S-wave pole separately, because there is no one-to-one correspondence between the underlying base functions and the resonance content of the dipion spectrum. This can be seen from their distributions in Fig. 5.4. The uncertainties on the fractions are not shown here, because they are determined separately in a bootstrapping approach in Sec. 8.

Projections of data and fit model are shown in Figs. 6.9 and 6.10 . The fit model histograms are sampled from the pdf using fully simulated $B_s^0 \to \psi(2S)\pi^+\pi^-$ events in order to automatically describe the efficiencies and acceptances correctly.

All distributions can be described well, except for the $f_0(980)$ region in $M_{\pi\pi}$, which seems to be slightly narrower in data than in the model. The overall quality of the fit is good nevertheless. This holds especially for the $M(\psi(2S)\pi)$ projections, where one would expect to observe $Z(4430)^{\pm}$ contributions as systematic variations in the pulls around 4.5 GeV. An upper limit on the fit fraction of this exotic tetraquark candidate is determined in Sec. 10.

The main difference at the amplitude level between the two solutions is the different interference between the f_2 states. For Solution 1 there is little interference, while in Solution 2 the negative interference between the states is almost as large as their corresponding fit fractions. Since large interference terms between Breit-Wigner amplitudes come with unitarity violation (see Sec. 2.4.1), the fractions for the f_2 states in Solution 2 are not process independent. Additionally, the $\log\mathcal{L}$ of the amplitude fit is better by 1.5 points for Solution 1. The favoured solution is therefore Solution 1, while Solution 2 is only considered for the sake of completeness in the following.

The existence two solutions for the f_2 states with large interference terms indicates that the two poles are not separated enough to describe these states

name	Solution 1	Solution 2
$\alpha_{f_0(1500)}$	$(-2.6 \pm 0.8) \times 10^3$	$(-2.1 \pm 0.8) \times 10^3$
$\alpha_{f_0(2020)}$	$(-6.4 \pm 0.5) \times 10^3$	$(-6.2 \pm 0.5) \times 10^3$
c_K	3310 ± 50	3350 ± 50
$\Im\left(B_{11}^{f_2(1270)}\right)$	$(3.6 \pm 1.0) \times 10^3$	$(-2.0 \pm 1.4) \times 10^3$
$\Re\left(B_{11}^{f_2(1270)}\right)$	$(-1.2 \pm 0.5) \times 10^3$	$(-1.3 \pm 0.5) \times 10^3$
$\Im\left(B_{11}^{f_2(1525)}\right)$	$(-1.2 \pm 0.6) \times 10^3$	$(1.5 \pm 0.8) \times 10^3$
$\Re\left(B_{11}^{f_2(1525)}\right)$	$(0.1 \pm 0.4) \times 10^3$	$(-1.05 \pm 0.35) \times 10^3$
$\log \mathcal{L}$	33989.56	33988.1

Table 6.10: Amplitude fit result for Bonn-Jülich model for the description of the S-wave component. The last row contains the $\log \mathcal{L}$ value. Coupling terms contain arbitrary scale factor, see footnote 5.

name	Solution 1	Solution 2
S-wave	0.975	0.989
$f_2(1270)$	0.019	0.007
$f_2(1525)$	0.003	0.007
Sum	0.996	1.003

Table 6.11: Amplitude fit fractions for Bonn-Jülich for the description of the S-wave component. The corresponding uncertainties are determined in Sec. 8.

	$f_2(1270)$	$f_2(1525)$	S-wave
$f_2(1270)$	/	/	/
$f_2(1525)$	0.0005	/	/
S-wave	0.0021	0.0002	/

Solution 2

	$f_2(1270)$	$f_2(1525)$	S-wave
$f_2(1270)$	/	/	/
$f_2(1525)$	-0.0043	/	/
S-wave	0.0021	-0.0	/

Table 6.12: Amplitude fit interferences for Bonn-Jülich model for the description of the S-wave component.

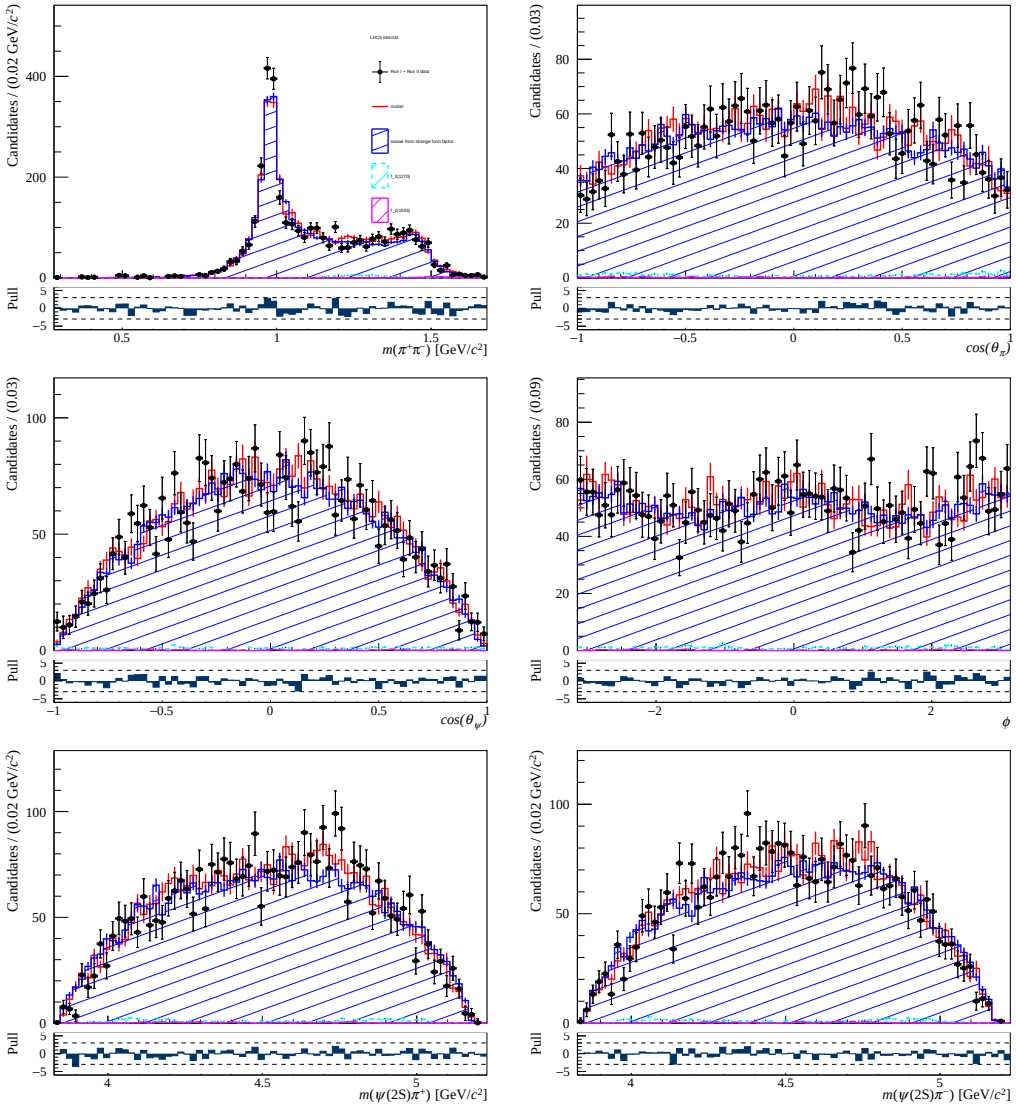

Figure 6.9: Amplitude fit projections Bonn-Jülich model for the description the S-wave. Shown for Solution 1. Pull distributions include statistical uncertainties on both data and model bins.

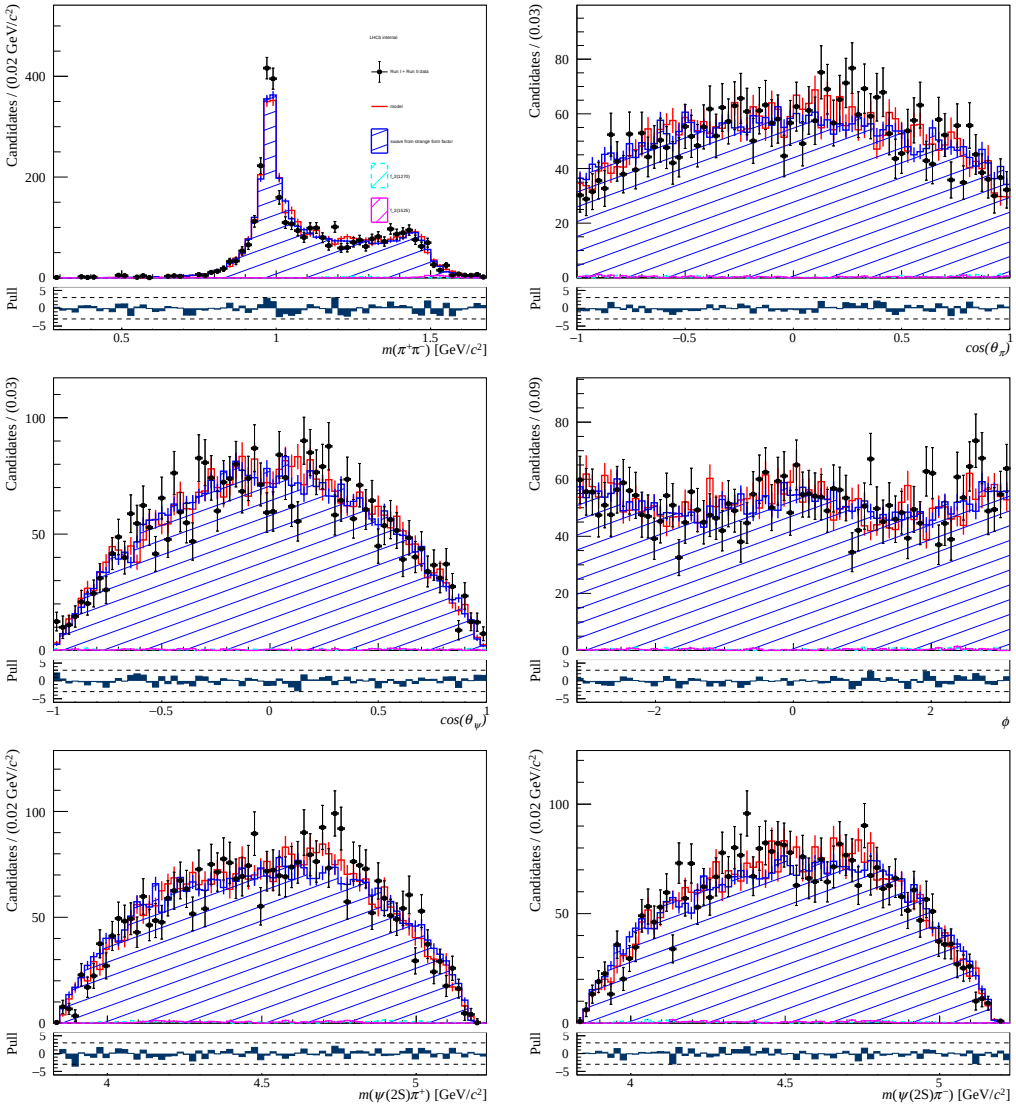

Figure 6.10: Amplitude fit projections Bonn-Jülich model for the description the S-wave. Shown for Solution 2. The contribution of the f_2 states is small for this solution. Pull distributions include statistical uncertainties on both data and model bins.

perfectly with Breit-Wigner parametrisations. Solution 1 shows little interference between these states though and is therefore considered reliable. Since their contribution to the overall fit is small, and the focus of this analysis lies on the dominant S-wave contribution, no other parametrisations for the f_2 states are considered. With a larger data sample, a more sophisticated parametrisation should be tested.

The qualitative result of the amplitude fit with the Bonn-Jülich model for the S-wave is comparable to the values obtained for $B_s^0 \to J/\psi\,\pi^+\pi^-$. The decay is dominated by the S-wave component, while the f_2 states make up less than 3 % of the fit fraction. The $f_2(1270)$ has a fit fraction of 1–2 % and the $f_2(1525)$ contributes even less than that.

It has been shown that a description using Flatté and Breit-Wigner amplitudes for the dipion S-wave spectrum is not appropriate due to violation of unitarity and technical problems in the fit from large numbers of free non-linear parameters. The Bonn-Jülich model on the other hand provides a stable, unitarity-conserving parametrisation of the S-wave amplitude, which is able to describe the data sufficiently well in $B_s^0 \to \psi(2S)\pi^+\pi^-$. Possible unitarity-violations in the Breit-Wigner parametrisation of the f_2 states could be remedied in a future analysis using a K-matrix parametrisation for these states (see review on resonances in [21]). The additional free parameters introduced with a K-matrix model cannot be constrained sufficiently from the current data. The use of this model in an amplitude fit of $B_s^0 \to \psi(2S)\pi^+\pi^-$ therefore requires a larger sample, which should be available after the next data taking period of the LHCb detector.

7 Process-Independence of Bonn-Jülich Model

In Sec. 6.6 it is shown that the Bonn-Jülich model, which was extracted from $B_s^0 \to J/\psi K^+K^-$ and $B_s^0 \to J/\psi \pi^+\pi^-$, is able to describe the data in $B_s^0 \to \psi(2S)\pi^+\pi^-$. It uses external parameters for the description of the high-mass resonances, which are fixed in the fit to $B_s^0 \to \psi(2S)\pi^+\pi^-$. These external parameters $\boldsymbol{w}_{\text{ext}}$ are the bare resonance-channel couplings g_i^{R} and bare resonance masses m_{R} introduced in Sec. 2.4.2. They are listed in Tab. 7.1. Even though the data is described well, it is still possible that these parameters are process-dependent and a different configuration of $\boldsymbol{w}_{\text{ext}}$ could describe the S-wave in $B_s^0 \to \psi(2S)\pi^+\pi^-$ even better. This is tested in the following.

Since a combined fit of all 4 decay channels is beyond the scope of this work, a simpler approach is taken: The likelihood of $B_s^0 \to \psi(2S)\pi^+\pi^-$ $\mathcal{L}(\boldsymbol{w}, \boldsymbol{w}_{\text{ext}})$ is a function of both the $\boldsymbol{w}_{\text{ext}}$ and the internal parameters $\boldsymbol{w}$, but in Sec. 6.6 only $\mathcal{L}(\boldsymbol{w}|\boldsymbol{w}_{\text{ext}})$ was optimized. If the external parameters are really process-independent, then $\boldsymbol{w}_{\text{ext}} \approx \boldsymbol{w}_{\text{ext}}^{\text{opt}}$ and the fit results are already optimal:

$$\mathcal{L}_{\text{max}}(\boldsymbol{w}^{\text{opt}}|\boldsymbol{w}_{\text{ext}}) = \mathcal{L}_{\text{max}}(\boldsymbol{w}^{\text{opt}}, \boldsymbol{w}_{\text{ext}}^{\text{opt}}). \tag{7.1}$$

If the $\boldsymbol{w}_{\text{ext}}$ are not process-independent on the other hand, then $\mathcal{L}_{\text{max}}(\boldsymbol{w}^{\text{opt}}|\boldsymbol{w}_{\text{ext}})$ is only a saddle point of $\mathcal{L}(\boldsymbol{w}, \boldsymbol{w}_{\text{ext}})$. This is checked in the following with parameter configurations $\boldsymbol{w}'_{\text{ext}}$ close to the nominal $\boldsymbol{w}_{\text{ext}}$ obtained by sampling from the covariance matrix of the fit to $B_s^0 \to J/\psi \pi^+\pi^-$ and $B_s^0 \to J/\psi K^+K^-$ in [1]. For each configuration $\boldsymbol{w}'_{\text{ext}}$ the fit to $B_s^0 \to \psi(2S)\pi^+\pi^-$ as described in Sec. 6.6 is then repeated resulting in a maximized likelihood $\mathcal{L}_{\text{max}}(\boldsymbol{w}'^{\text{opt}}|\boldsymbol{w}'_{\text{ext}})$. Like this the internal parameters $\boldsymbol{w}$ are profiled out and one can focus on $\mathcal{L}(\boldsymbol{w}_{\text{ext}})$. The distribution of this likelihood against the external parameters is visualized in Fig. 7.1 as scatter plots.

A process-dependence of one parameter $\boldsymbol{w}_{\text{ext}\,i}$ would be visible here as a parabolic shape [1] in the scatter plot between $\boldsymbol{w}_{\text{ext}\,i}$ and the $\log \mathcal{L}$. If the external parameters are optimal, the likelihood is uncorrelated to the $\boldsymbol{w}'_{\text{ext}}$.

[1] Or a strong linear correlation, in case the external parameter is far from optimal.

name	value [GeV]
$m_{f_0(1500)}$	1.2941 ± 0.020
$m_{f_0(2020)}$	3.26873 ± 0.401
$g_\pi^{f_0(1500)}$	0.4149 ± 0.0097
$g_\pi^{f_0(2020)}$	0.4257 ± 0.17
$g_K^{f_0(1500)}$	0.358 ± 0.015
$g_K^{f_0(2020)}$	-0.03 ± 0.11
$g_{4\pi}^{f_0(1500)}$	0.52 ± 0.12
$g_{4\pi}^{f_0(2020)}$	4.6 ± 1.3

Table 7.1: External parameters of the Bonn-Jülich model. $m_{f_0(1500)}$ and $m_{f_0(2020)}$ correspond to the bare masses, not the physical masses of the resonance poles.

The scatter plots in Fig. 7.1 show the result for $N_{\text{gen}} = 200$ external parameter configurations for Solution 1 of the amplitude fit. They do not show any significant correlations between the $\log \mathcal{L}$ and the external parameters $\boldsymbol{w}_{\text{ext}}$.

This means that within the sensitivity of the amplitude fit to $B_s^0 \to \psi(2S)\pi^+\pi^-$[2] the external parameters can be considered to be already at the optimal values, which supports the hypothesis of process independence.

[2]The sensitivity to the exact value of the bare mass of the high-lying resonance is limited in $B_s^0 \to \psi(2S)\pi^+\pi^-$ with respect to the data studied in [1] due to its smaller phase space. A process dependence at higher energies would need to be checked in a different channel, *e.g.* $J/\psi \to \pi^+\pi^-\gamma$.

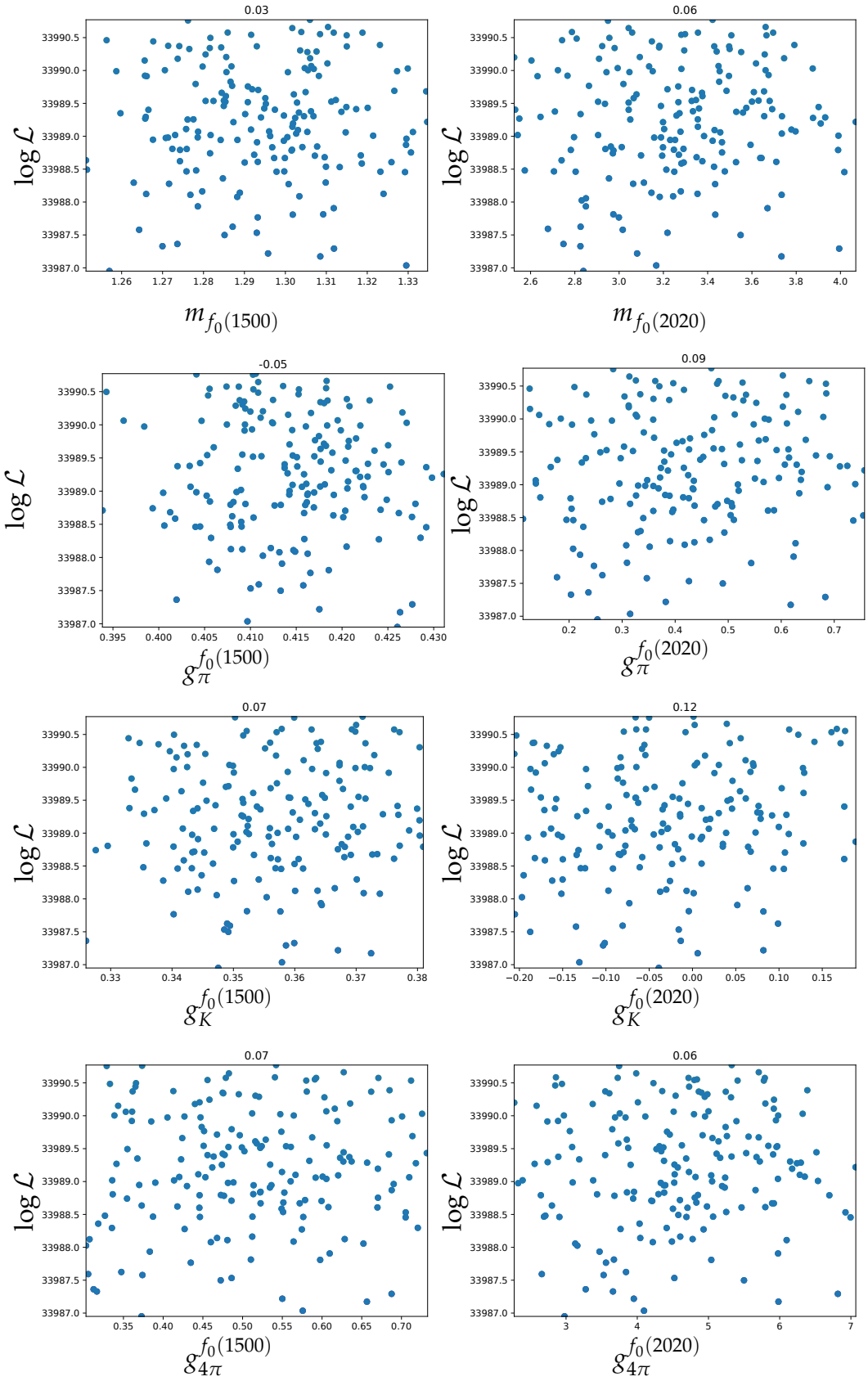

Figure 7.1: Scatter plots of external parameters of Bonn-Jülich model sampled from their covariance matrix and corresponding likelihoods in fit to $B_s^0 \to \psi(2S)\pi^+\pi^-$. Titles show linear correlation coefficients. No significant correlations are observed.

8 Statistical Uncertainties

The results of the amplitude fit might be biased, *e.g.* due to non-parabolic likelihood shapes because of the small sample size or because of non-linear dependencies between fit parameters. In this case, the resulting uncertainty estimates from `MIGRAD` and `HESSE` are not reliable, because they assume linear correlations between the parameters [123]. Additionally, the statistical uncertainty that stems from the extraction of the s-weights (see Sec. 4.6) is not fully reflected in the results: A factor that describes the reduction of effective sample size due to the non-uniform weights is introduced in Sec. 6.1, but this procedure does not account for non-Poissonian uncertainties in the fit to $M(\psi(2S)\pi^+\pi^-)$. These additional uncertainties come from the floating shape parameters of the fit components like the width of the signal distribution or the decay constant of the exponential distribution describing the combinatoric background. The uncertainties are underestimated by up to 15 % for the $B_{\ell s}$ couplings of the f_2 states.

An additional complication arises with the fit fractions, which are important physics results of this analysis, but are not fitted directly. Their dependence on the fit parameters is not linear, since they depend on the squared amplitudes. This non-linearity is especially pronounced for amplitudes with small fit fractions, like the decays via f_2 states, which lie near to their physical boundaries. For these the uncertainties on the fractions are not symmetric, because fractions are positive by definition. Gaussian uncertainty propagation from the fit parameters is therefore inaccurate in that case.

The correct statistical uncertainties of the fit parameters and derived quantities can be determined using the bootstrapping method [126]. The approach suggested in [127] is followed here: A new sample (called resample) of similar size is drawn from the original data sample; in this procedure individual events can be drawn more than once. All resamples are therefore drawn from the same underlying distribution, which is in this case the original sample. This is analogue to performing repeated measurements of a physical quantity, which follows a random distribution. Since the resamples have the same size as the original sample, the quantities derived from the resamples like mean, standard deviation, but also the maximum likelihood estimators have

name	Solution 1	Solution 2
${}^{\alpha}f_0(1500)$	$(-2.6 \pm 0.8) \times 10^3$	$(-2.1 \pm 0.8) \times 10^3$
${}^{\alpha}f_0(2020)$	$(-6.4 \pm 0.5) \times 10^3$	$(-6.1 \pm 0.6) \times 10^3$
c_K	3310 ± 40	3350 ± 40
$\Im\left(B_{11}^{f_2(1270)}\right)$	$(3.6 \pm 1.2) \times 10^3$	$(-2.7 \pm 1.4) \times 10^3$
$\Re\left(B_{11}^{f_2(1270)}\right)$	$(-1.3 \pm 0.6) \times 10^3$	$(-1.1 \pm 0.5) \times 10^3$
$\Im\left(B_{11}^{f_2(1525)}\right)$	$(-1.3 \pm 0.5) \times 10^3$	$(1.8 \pm 0.9) \times 10^3$
$\Re\left(B_{11}^{f_2(1525)}\right)$	$(0.0 \pm 0.5) \times 10^3$	$(-1.0 \pm 0.4) \times 10^3$

Table 8.1: Results for amplitude fit variables from full bootstrapping with Bonn-Jülich model.

the same fluctuations as for the original sample. The confidence levels of an estimator on the original sample can then be estimated from the width of the distribution of estimator results from the resamples.

A drawback of this approach is the computational cost, since the numerically expensive amplitude fit needs to be repeated for every resample and cannot be performed an arbitrary number of times. As a compromise between cost and accuracy, $N_{toy} = 600$ resamples are drawn. For each the fit to $M(\psi(2S)\pi^+\pi^-)$ and subsequent determination of s-weights is repeated. The amplitude fit is then repeated with the resamples and their new weights [1]. The distribution of the resulting parameters x_{toy} for the N_{toy} fits can then be used to determine the uncertainty of the parameters for the original sample with correct coverage.

Fig. 8.1 shows the bootstrapping distributions for the fit variables. The distribution for each variable is fitted with a Gaussian pdf, whose width parameter is taken as uncertainty for that variable. Note that only the plots for Solution 1 are shown, since Solution 2 is disfavoured due to its smaller likelihood and the large unitarity-violating interference terms between the f_2 states (see Sec. 6.6).

The results for mean and width of the Gaussians for the fit variables from the bootstrapping approach are given in Tab. 8.1. They can be compared with the values in Tab. 6.10. Due to the limited number of resamples, the mean of the Gaussians does not coincide exactly with the estimate from the nominal fit, but the remaining difference is much smaller than the statistical uncertainty. For the $B_{\ell s}$ couplings of the f_2 states, the uncertainties increase noticeably with respect to the estimates from the amplitude fit, while for the other parameters the differences are smaller.

[1] The calculation of the s-weights for each resample is necessary to correctly propagate the statistical uncertainty of this technique to the results, as shown in [127].

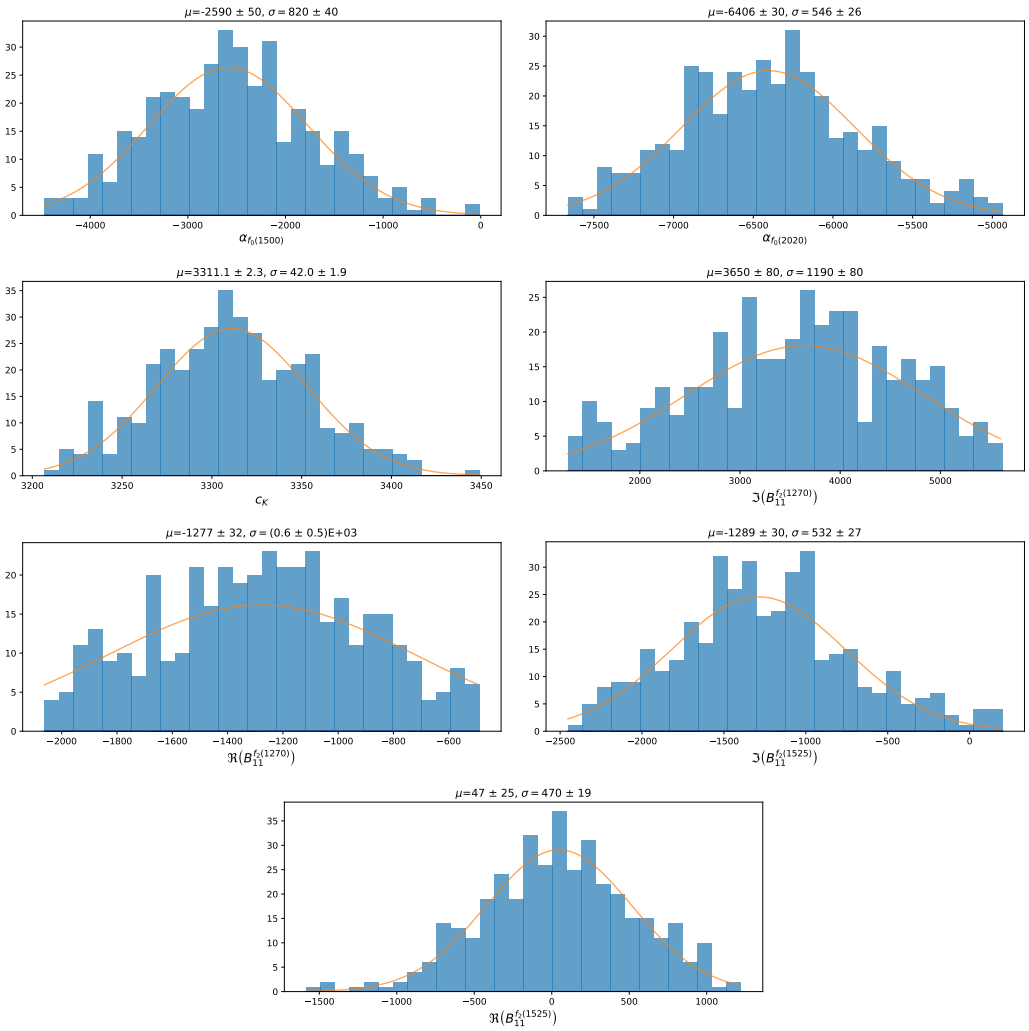

Figure 8.1: Distributions of amplitude fit parameters from full bootstrapping for the amplitude fit with Bonn-Jülich model. Shown for Solution 1. Overlaid with a fit of a Gaussian distribution. Parameter mean and width of the fit result are shown in the title together with their uncertainties.

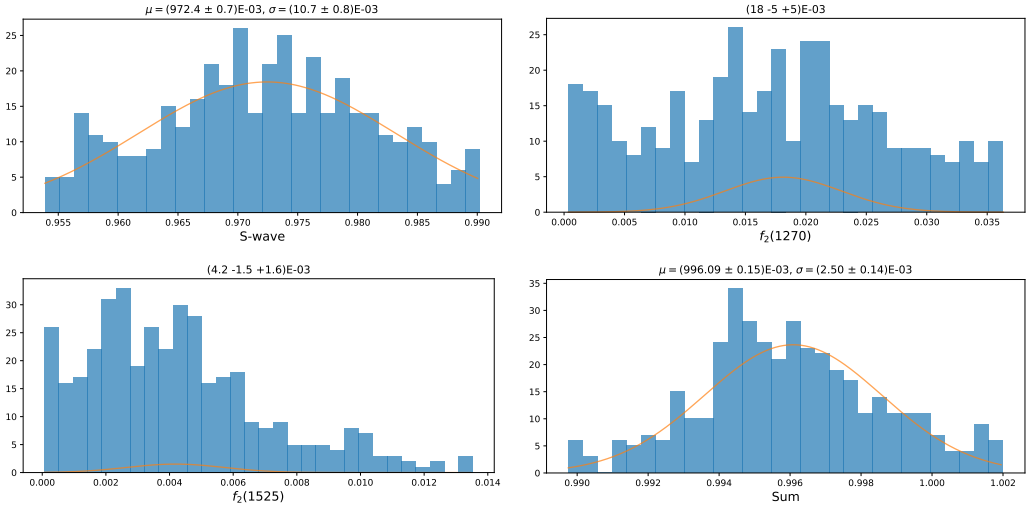

Figure 8.2: Distributions of amplitude fit fractions from full bootstrapping for amplitude fit variables from full bootstrapping with Bonn-Jülich model. Orange: Fit with a Gaussian distribution. Parameter name, mean and width of the fit result are shown in the title. In case of skewed distributions asymmetric uncertainties are determined from percentiles. The overlaid Gaussian then depicts the symmetrized interval for visual purposes only. Shown for Solution 1.

The uncertainties on fit fractions and interferences are extracted from the bootstrapping directly and not from the estimates provided by the fit. The corresponding distributions are shown in Fig. 8.2 for the fractions and Fig. 8.3 for the interferences. Where the distribution cannot be approximated with a simple Gaussian, the median and asymmetric 68 % intervals of the distribution are calculated directly from the distribution. The asymmetric intervals are then symmetrized for simplicity. This is the case for the fractions of the f_2 states, which are skewed towards 0 due to their small contribution to the fit.

The results for the fractions are combined into Tab. 8.2 and can be compared to the values obtained from the nominal fit in Tab. 6.11. The central values from the bootstrapping are included as a consistency check here. With increasing number of resamples the peak of the bootstrapping distribution should approach the central value obtained from the nominal fit [128], which is is the case here. The nominal central values of the variables, fractions and interferences, which will be quoted as final result, are the values obtained from the nominal fit. The fact that the central values from the bootstrapping agree with the fit estimates indicates that enough resamples were used and also the estimate for the uncertainties from the bootstrapping is sound.

The interference terms between the fit components are combined into Tab. 8.3. They are comparable to the values obtained from the nominal fit

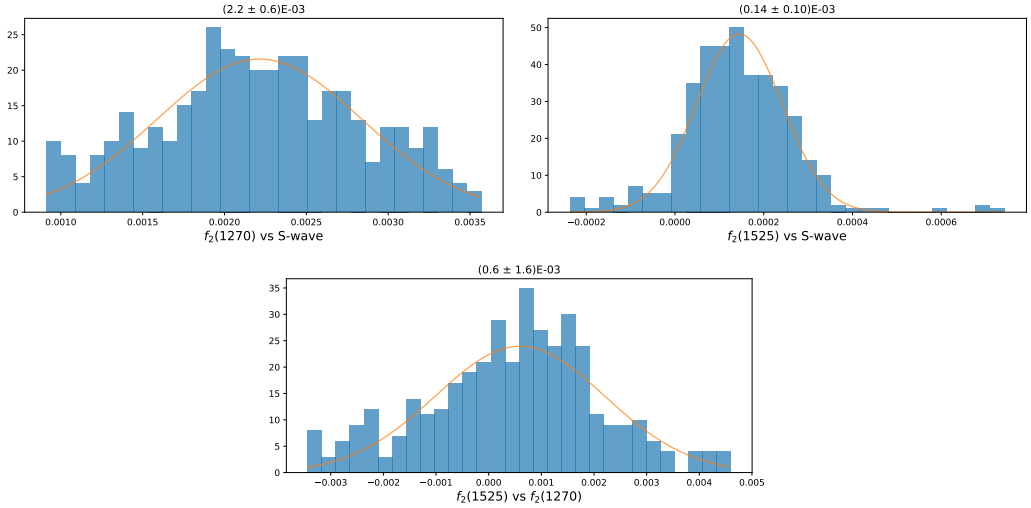

Figure 8.3: Distributions of amplitude fit interferences from full bootstrapping for the amplitude fit with Bonn-Jülich model. Overlaid with a fit of a Gaussian distribution. Parameter mean and width of the fit result are shown in the title together with their uncertainties. Top Left: S-wave – $f_2(1270)$. Top Right: S-wave – $f_2(1525)$. Bottom: $f_2(1270)$– $f_2(1525)$. Shown for Solution 1.

name	Solution 1	Solution 2
S-wave	0.972 ± 0.011	$(987 \pm 9) \times 10^{-3}$
$f_2(1270)$	$(18 \pm 5) \times 10^{-3}$	$(9 \pm 4) \times 10^{-3}$
$f_2(1525)$	$(4.2 \pm 1.5) \times 10^{-3}$	$(9.9 \pm 3.4) \times 10^{-3}$
Sum	$(996.1 \pm 2.5) \times 10^{-3}$	$(1005.0 \pm 3.1) \times 10^{-3}$

Table 8.2: Results for amplitude fit fractions for amplitude fit variables from full bootstrapping with Bonn-Jülich model.

given in Tab. 6.12. Taking into account the statistical uncertainties, a significant interference between the S-wave and the $f_2(1270)$ is found for both solutions. With only 0.2 % it is small though compared to the fit fraction of the involved S-wave amplitude. The other interference terms are not significant for Solution 1. In Solution 2 the interference between the f_2 states is significant though. The large interference between the Breit-Wigner amplitudes violates unitarity and the resulting fit fractions cannot be interpreted as physical quantities, which supports the reasoning in Sec. 6.6 to disfavour this solution.

Solution 1

	S-wave	$f_2(1270)$	$f_2(1525)$
S-wave	1	$(2.2 \pm 0.6) \times 10^{-3}$	$(0.14 \pm 0.10) \times 10^{-3}$
$f_2(1270)$		1	$(0.6 \pm 1.6) \times 10^{-3}$

Solution 2

	S-wave	$f_2(1270)$	$f_2(1525)$
S-wave	1	$(1.7 \pm 0.8) \times 10^{-3}$	$(-0.05 \pm 0.19) \times 10^{-3}$
$f_2(1270)$		1	$(-5.2 \pm 2.7) \times 10^{-3}$

Table 8.3: Interferences for amplitude fit amplitudes from full bootstrapping with Bonn-Jülich model.

9 Systematic Uncertainties

In addition to the statistical uncertainties determined in Sec. 8 also systematic uncertainties need to be estimated for the amplitude fit results. The necessary studies are detailed in the following paragraphs and summarized in Tab. 9.4 and 9.3.

Most systematic studies presented here deal with choices in the fit models, where it is not clear, which one is "better". One approach to estimate the related uncertainties is to repeat the fit to the data sample with modified models and take the difference to the nominal result as systematic uncertainty from the model choice. This approach is prone to be biased by statistical fluctuations in the data sample and can therefore over- or under-estimate the effect size.

Since the data sample is the same for all model choices, bootstrapping can be used to reduce the statistical fluctuations compared to the one-fit approach described above. The data sample is bootstrapped as explained in Sec. 8 and the fit procedure is performed twice on each resample i, once with the nominal fit configuration resulting in parameters $\mathbf{w}_i$ and once with the modified configuration resulting in parameters $\mathbf{w}'_i$. This is repeated N times and the difference $\Delta\mathbf{w}_i$ between nominal and modified results is taken for each resample. The width of the distribution of $\Delta\mathbf{w}_i$ corresponds to the systematic uncertainty on the parameter $\mathbf{w}$. The peak of the distribution of $\Delta\mathbf{w}_i$ is an estimate of the systematic bias between the two models. In the case that one model can be preferred over the other, this bias can be corrected for. In most cases there is no clear preference of one model over the other. In that case the quadratic sum of peak position and width of the distribution of $\Delta\mathbf{w}_i$ is taken as systematic uncertainty on $\mathbf{w}$. This procedure is referred to as approach 1.

The peak position and width are determined with a fit of a Gaussian pdf and the mean of the Gaussian is taken as systematic uncertainty.

For some systematic studies this approach is not applicable, because the data sample is different in the nominal fit and in the systematic study. This is *e.g.* the case, when additional cuts are introduced in the selection. When that is the case, the systematic uncertainty is taken directly as the difference between

the nominal fit and the fit to the modified data sample. This is referred to as approach 2.

As discussed in Sec. 6.6, Solution 2 of the amplitude fit shows significant interferences between the f_2 states. It is therefore not considered as viable result. Systematic uncertainties for Solution 2 are only included for the sake of completeness.

9.1 Quality of Simulation

Mismodelling in Simulation The simulated data samples are reweighted in the track-multiplicity and $p_{\mathrm{T}}(B_s^0)$ in order to make them resemble the data better (see Sec. 3.2.4). Other variables, which are used as inputs to the MVA in Sec. 4, are slightly mismodelled as well though. The cut on the MVA output in the selection could therefore affect the efficiency distribution as function of the amplitude variables differently for the signal in data than for the simulated signal sample, which could bias the fit results. On the other hand, weighting in too many variables leads to a large spread of weights, which reduces the effective size of the weighted sample (see eq. 6.15). This can then lead to a bias in the amplitude fit, because the statistical fluctuations in the normalisation sample in eq. 6.14 distort the likelihood distribution. Therefore, the nominal fit uses a reweighting in just 2 variables. To study the effect of ignoring the remaining mismodelling, an alternative reweighting is performed on the simulated sample in 9 variables. Since the data sample is not affected by the reweighting, the systematic uncertainty can be evaluated with approach 1. After the systematic uncertainty from the amplitude model variation this is the largest source of systematic uncertainty.

Mismodelling of VELO Errors The VELO hit uncertainty estimates are known to be mismodelled in the simulated samples of 2018 and 2017 data, which results in differences in the χ^2_{IP} distributions. In order to test the effect, the selection in these variables is tightened, reducing the signal efficiency by 10 % and the difference to the nominal result is estimated with approach 2. Since the fit uses an extended likelihood, the coupling variables need to be scaled for the modified data samples by a factor $\sqrt{\epsilon_{\mathrm{extracuts}}}$.

Tracking Efficiency Corrections Simulated tracking efficiencies are corrected in a data-driven way (see Sec. 3.2.4). The track-wise correction factors are provided in a 2-dimensional table in track momentum and track η and the

4 factors from the final state tracks are multiplied together for the correction-weight of each simulated event. These corrections are mostly below 5 %. For the amplitude variables this effect is smaller, because they are measured in the B_s^0's frame and its daughters' and the boost into these systems smears out the corrections (note that the fit is only sensitive to relative efficiency variations in the variables). In order to estimate the uncertainty on this correction, the table is varied within its respective statistical and (correlated) systematic uncertainties 200 times and the fit to the nominal data sample is repeated for each correction. The difference to the fit result with the nominal correction is taken for each fit and the distribution of these differences is fitted with a Gaussian pdf. The width parameter of this pdf corresponds to the systematic uncertainty of the tracking efficiency corrections. It is found to be negligibly small.

Difference between Runs in $M(\psi(2S)\pi^+\pi^-)$ The nominal fit extracts the *s*-weights from the combined Run 1 and Run 2 LHCb data samples. This is justified, because the shape parameters of the signal and background components are compatible for both data taking conditions within uncertainties. Nevertheless, the assumption that the shape parameters are the same, is addressed with a systematic uncertainty. The Run 1 and Run 2 samples are combined after *s*-weighting them separately and the difference to the nominal result is measured with approach 2.

Difference between Runs Efficiency variations over the Dalitz plane could be modelled with different precision for Run 1 and Run 2 simulation. In order to estimate this effect, the bootstrapping, *s*-weighting and amplitude fit are performed once with only Run 2 data and once with both Run 1 and Run 2 data. For the latter the data samples are combined only after the *s*-weighting in order to not mix the effect with the difference between Runs in $M(\psi(2S)\pi^+\pi^-)$. The mode of the differences is estimated with a Gaussian fit and taken as systematic uncertainty. Since the fit uses an extended likelihood, the coupling variables are corrected for the Run 2 only fits by a factor $\sqrt{\frac{N_{\text{Run2}}}{N_{\text{Run1}}+N_{\text{Run2}}}}$. Additionally, the Bonn-Jülich parameters c_K, $\alpha_{f_0(2020)}$, and $\alpha_{f_0(1500)}$ have a different configuration, which increases the measured difference for these parameters. Since the fit fractions are the values of interest in this analysis and not the couplings, this is not a problem, because the effect on the fractions is small.

9.2 Fit Model in $M(\psi(2S)\pi^+\pi^-)$

The choice of parametrisation of the distributions in $M(\psi(2S)\pi^+\pi^-)$ (see Sec. 4) is not unique. To test the systematic uncertainty related to the choice of background model, the exponential pdf is exchanged with a first order Chebychev polynomial and approach 1 is followed. This is repeated for the signal component, where the double Crystal Ball pdf is replaced with a Hypatia pdf [129].

9.3 Resolution Effects

The resolution in the amplitude variables is well below the size of the expected structures (see Sec. 6.2.1), which is why resolution effects are not part of the nominal fit. To estimate the systematic uncertainty related to neglecting resolution effects, approach 2 is used with the modified likelihood from eq. 6.27. The effect is found to be negligibly small.

9.4 Time Acceptance

The derivation of eq. 5.47 does not take into account the non-uniform decay time acceptance at LHCb, which is shaped by *e.g.* the PV-detachment criteria and other selection cuts introduced in Sec. 4. The size of this mismodelling effect is estimated using approach 1, where the interference term in eq. 5.47 is removed, neglecting time-dependent effects, that could be influenced by the time acceptance. Since in the nominal fit all components have odd *CP* parity, which removes any interference effects, this is done with the model, which includes higher angular moments in the B_s^0 decay for the f_2 states as explained in Sec. 5.3.1 and 9.6.

9.5 Theory Uncertainty

The process-independent parameters used in the Bonn-Jülich model have their own uncertainties. They need to be propagated into the fit results. The study performed in Sec. 7 is used for this purpose: The distribution of results from fits with randomly sampled values for the external parameters is fitted with a Gaussian pdf, whose standard deviation is taken as systematic uncertainty.

9.6 Amplitude Fit Parametrisation

Apart from the uncertainty introduced from the statistical uncertainty of the external Bonn-Jülich model parameters described above, the nominal fit model makes assumptions on the size of other external parameters as well. The number of participating resonances and additional process-dependent parameters of the Bonn-Jülich model need to be considered. Since the data sample does not change for model variations, it is possible to use approach 1 for these studies. The following changes were made to the model to estimate their corresponding systematic uncertainty:

- In the nominal fit the mass and width parameters of the Breit-Wigner description of the f_2 states are fixed to the central values taken from PDG [21]. In the modified model these parameters are allowed to vary. Their values are constrained by the uncertainties quoted by the PDG through a parabolic term, added to the $\log \mathcal{L}$.

- The nominal fit allows only the lowest angular momentum couplings, because higher angular momentum states should be suppressed by the momentum barrier effect (see Sec. 5.3). This assumption is tested allowing $L_{\text{min}} + 1$ for the f_2 states, which adds 4 additional real fit parameters, corresponding to two complex couplings $B_{22}^{f_2(1270)}$ and $B_{22}^{f_2(1525)}$.

- The hadronic scale of the empirical Blatt-Weisskopf barrier factors used in the Breit-Wigner descriptions of the f_2 states (see Sec. 5.4) is set to $1.5\,\text{fm}^{-1}$ for the nominal fit. It describes the effective size of the described system, which is not known precisely. In order to estimate its effect, alternative models are fitted with $0.5\,\text{fm}^{-1}$ and $2\,\text{fm}^{-1}$ for the scale.

- In order to test, if all necessary resonances are included in the model, a Breit-Wigner parametrisation of the $\rho_3(1690)$ resonance is added with mass and width parameter fixed to PDG values. This resonance is prohibited by isospin symmetry. It is added here to estimate the effects of isospin symmetry breaking in the decay.

- In the Bonn-Jülich model described in eq. 5.4 additional linear polynomial terms in $M^2_{\pi^+\pi^-}$ are in principle allowed. Since they are highly correlated to the 0th order coefficients, they are not included in the nominal fit to increase fit stability. Their effect is tested by allowing 2 or all 3 of them to float in the model. This corresponds to up to 3 additional real couplings l_π, l_K, $l_{4\pi}$. Since the allowed phase space for

	Sum × 0.001	S-wave × 0.001	$f_2(1270)$ × 0.001	$f_2(1525)$ × 0.001
varied PDG inputs	0.60	2.8	2.1	3.22
$\rho(1690)$	0.45	1.2	0.7	0.67
nr component	0.16	7.8	5.2	0.11
larger L	**4.58**	**8.9**	5.8	2.39
larger had scale	0.60	1.2	1.7	0.28
smaller had scale	0.38	0.9	1.1	0.23
3 more S-wave terms	3.91	8.0	5.6	4.99
$l_{pi}l_K$ S-wave terms	3.03	6.8	7.2	**6.50**
$l_{pi}l_{4pi}$ S-wave terms	0.41	3.7	4.8	0.23
l_Kl_{4pi} S-wave terms	0.59	3.8	**8.8**	1.18
quadratic sum	6.10	14.6	12.1	7.68

	Sum × 0.001	S-wave × 1e-06	$f_2(1270)$ × 1e-06	$f_2(1525)$ × 1e-06
varied PDG inputs	3.06	4379	3438	4360
$\rho(1690)$	0.38	2310	657	626
nr component	0.15	4	5	5
larger L	**7.23**	**11140**	6662	**4962**
larger had scale	0.51	515	782	306
smaller had scale	0.22	370	430	310
3 more S-wave terms	3.64	5186	**8781**	4721
$l_{pi}l_K$ S-wave terms	3.16	5983	7053	2568
$l_{pi}l_{4pi}$ S-wave terms	4.29	2883	7618	2627
l_Kl_{4pi} S-wave terms	1.96	2563	3284	2383
quadratic sum	8.97	13589	11591	8149

Table 9.1: Systematic Uncertainties on fit fractions from sources related to the amplitude model. Bold for largest contributions. Top: Solution 1. Bottom: Solution 2.

$M_{\pi\pi}$ is only within $[2m_\pi, m_{B_s^0} - m_{\psi(2S)}] = [280, 1680]$ MeV, and the dominant base function $\tilde{F}_{KK}$ is very small for low masses (see Fig 5.4), the fit can hardly distinguish the contributions from the 0th order and the 1st order terms in the polynomial in eq. 5.4. These terms are therefore highly anti-correlated with each other, which leads to a large uncertainty and for some resamples also sign flips for some of the parameters. The effect on the full S-wave amplitude is mostly cancelled out between the 0th and 1st order components of the polynomial of eq. 5.4. The overall shape of the amplitude is therefore barely affected by the addition of the linear terms. This can be seen from the fact that the fit parameters of the Bonn-Jülich model have significantly different values, when the linear terms are included, while the fit fractions of the S-wave and the f_2 states change by less than 0.01.

The results for the fractions and fit parameters are shown in Tab. 9.1 and 9.2. The largest value in each group is added in quadrature for the quadratic sum in the last line. The largest uncertainties on the fit parameters are introduced via inclusion of 1st order polynomial terms in the S-wave model.

	$\alpha_{f_0(1500)}$	$\alpha_{f_0(2020)}$	c_K	$\Im\left(B_{11}^{f_2(1270)}\right)$	$\Re\left(B_{11}^{f_2(1270)}\right)$	$\Im\left(B_{11}^{f_2(1525)}\right)$	$\Re\left(B_{11}^{f_2(1525)}\right)$
varied PDG inputs	38	16	1.5	80	50	177	140
$\rho(1690)$	79	76	5.2	250	50	125	70
nr component	939	363	16.6	870	180	69	170
larger L	170	59	8.0	2380	130	**820**	370
larger had scale	40	15	2.6	180	50	51	100
smaller had scale	26	10	1.9	120	50	32	70
3 more S-wave terms	2226	**11891**	**2912.2**	**2620**	**680**	357	1480
$l_{pi}l_K$ S-wave terms	**13993**	9948	2890.3	2250	430	481	**1960**
$l_{pi}l_{4pi}$ S-wave terms	2382	11811	403.5	120	110	107	60
$l_K l_{4pi}$ S-wave terms	5130	1781	1628.4	360	430	153	160
quadratic sum	14026	11897	2912.2	3660	720	978	2010

	$\alpha_{f_0(1500)}$	$\alpha_{f_0(2020)}$	c_K	$\Im\left(B_{11}^{f_2(1270)}\right)$	$\Re\left(B_{11}^{f_2(1270)}\right)$	$\Im\left(B_{11}^{f_2(1525)}\right)$	$\Re\left(B_{11}^{f_2(1525)}\right)$
varied PDG inputs	66.03	113.44	2.769	307.135	300.31	661.01	171.33
$\rho(1690)$	124.56	120.47	3.930	78.833	39.88	127.23	26.85
nr component	0.13	0.20	0.011	0.018	0.27	0.26	0.16
larger L	420.88	139.69	29.755	2186.240	236.96	1591.38	461.38
larger had scale	21.71	13.38	1.184	65.653	37.27	52.77	19.35
smaller had scale	13.29	33.27	0.761	43.892	23.10	53.43	17.48
3 more S-wave terms	**10347.80**	**18992.70**	**3343.458**	**3142.392**	**2416.43**	**3993.67**	**482.62**
$l_{pi}l_K$ S-wave terms	5788.89	861.81	2564.790	1085.121	326.28	743.97	132.51
$l_{pi}l_{4pi}$ S-wave terms	1483.55	11753.40	532.014	836.298	113.71	146.79	132.53
$l_K l_{4pi}$ S-wave terms	4779.81	1706.23	1759.230	809.355	260.45	287.55	137.01
quadratic sum	10357.34	18993.96	3343.594	3841.761	2447.13	4351.76	690.11

Table 9.2: Systematic Uncertainties on fit parameters from sources related to the amplitude model. Bold for largest contributions. Top: Solution 1. Bottom: Solution 2.

9.7 Combination of Systematic Uncertainties

The different systematic uncertainties described above are added in quadrature in Tabs. 9.4 and 9.3. The systematic uncertainties are significantly larger than the statistical uncertainties, not only for the fit parameters, but also for the extracted fit fractions and therefore the partial branching ratios. The second largest systematic uncertainty is related to the quality of the simulation of the signal decays.

The largest systematic uncertainties are related to variations for the amplitude model and within these choices those related to the Bonn-Jülich model contribute most. This gives the impression that the Bonn-Jülich model might not be a good parametrisation of the S-wave component of the dipion spectrum. One needs to consider though that the main reason, why the statistical uncertainties are so small in the first place, is the fact that the shape of the S-wave amplitude is fixed by the Bonn-Jülich model. The Breit-Wigner and Flatté amplitudes, which require floating shape parameters, did not result in a converging fit and the fit parameters and fit fractions could not be constrained effectively by the data (see Sec. 6.5).

The systematic uncertainties tend to be larger for the disfavoured (see Sec. 6.6) Solution 2.

The results for fit parameters together with their statistical and systematic

Solution 1

	$\alpha_{f_0(1500)}$ $\times$ 1000.0	$\alpha_{f_0(2020)}$ $\times$ 1000.0	c_K	$\Im\left(B_{11}^{f_2(1270)}\right)$ $\times$ 1000.0	$\Re\left(B_{11}^{f_2(1270)}\right)$	$\Im\left(B_{11}^{f_2(1525)}\right)$	$\Re\left(B_{11}^{f_2(1525)}\right)$
MC weights	0.33	0.17	60	0.48	200	140	24
VELO uncertainties	0.150	0.300	178.0	0.014	80	200	170
Run differences $M(\psi(2S)\pi^+\pi^-)$	0.050	0.050	0.4	0.190	90	80	80
Run differences amplitude variables	0.19	1.44	920	1.34	410	300	82
resolution effects	0.041	0.039	5.0	0.009	8	17	4
different background model	0.011	0.018	2.5	0.024	7	21	7
different signal model	200	174	16.1	0.084	43	59	37
time acceptance	0.534	0.1233	625.0	0.701	256	268	85
theory input	0.70	0.52	70	0.17	60	80	101
amplitude model	14.03	11.90	2910	3.66	720	980	2005
quadratic sum	14.1	12.1	3120	4.0	900	1100	2000
statistical uncertainty	0.7	0.6	50	1.2	600	500	500

Solution 2

	$\alpha_{f_0(1500)}$	$\alpha_{f_0(2020)}$ $\times$ 1000.0	c_K	$\Im\left(B_{11}^{f_2(1270)}\right)$ $\times$ 1000.0	$\Re\left(B_{11}^{f_2(1270)}\right)$	$\Im\left(B_{11}^{f_2(1525)}\right)$	$\Re\left(B_{11}^{f_2(1525)}\right)$
MC weights	208	0.6	110	6.74	60	3040	1140
VELO uncertainties	210	0.39	183	0.360	420	800	14
Run differences $M(\psi(2S)\pi^+\pi^-)$	50	0.21	29	40	0.230	700	113
Run differences amplitude variables	8	1.3	940	0.81	270	510	290
resolution effects	43	0.042	5.0	0.008	9	21	2.9
different background model	17	0.026	3.3	0.011	20	41	7.9
different signal model	254	0.168	15.8	0.074	37	119	26.8
time acceptance	427	1.139	634.1	0.538	281	399	196.5
theory input	620	0.4	70	0.15	100	90	50
amplitude model	10357	19.0	3340	3.84	2450	4350	690
quadratic sum	+10400	19.1	+3540	7.8	2500	+5500	1400
statistical uncertainty	700	0.8	50	1.2	700	1000	500

Table 9.3: Systematic Uncertainties on fit parameters compared to statistical uncertainty. Top: Solution 1. Bottom: Solution 2.

Solution 1

	Sum $\times$ 0.001	S-wave $\times$ 0.001	$f_2(1270)$ $\times$ 0.001	$f_2(1525)$ $\times$ 0.001
MC weights	0.83	3.9	5.6	1.00
VELO uncertainties	1.3	2.5	0.26	0.34
Run differences $M(\psi(2S)\pi^+\pi^-)$	0.5	1.2	1.65	0.11
Run differences amplitude variables	0.04	2.6	3.1	0.43
resolution effects	0.029	0.070	0.062	61
different background model	0.043	0.284	0.171	0.104
different signal model	0.292	0.824	0.812	0.322
time acceptance	0.017	0.005	0.012	0.005
theory input	0.41	1.8	1.8	0.40
amplitude model	8.89	13.5	13.3	7.93
quadratic sum	9.0	15	15	8.0
statistical uncertainty	2.6	11	5	1.3

Solution 2

	Sum $\times$ 0.001	S-wave $\times$ 0.001	$f_2(1270)$ $\times$ 0.001	$f_2(1525)$ $\times$ 0.001
MC weights	10.14	12.0	9.90	6.9
VELO uncertainties	1.7	1.3	3.18	5.01
Run differences $M(\psi(2S)\pi^+\pi^-)$	1.1	1.4	0.18	0.11
Run differences amplitude variables	0.27	0.7	0.12	0.6
resolution effects	0.058	0.129	0.068	0.152
different background model	0.165	0.311	0.139	0.341
different signal model	0.658	0.961	0.492	0.858
time acceptance	0.003	0.007	0.005	0.004
theory input	0.30	0.8	1.03	0.7
amplitude model	8.97	13.6	11.59	8.1
quadratic sum	13.7	18	16	11.9
statistical uncertainty	3.4	10	4	3.3

Table 9.4: Systematic Uncertainties on fit fractions compared to statistical uncertainty. Top: Solution 1. Bottom: Solution 2.

uncertainties are given below. The systematic uncertainties are dominated by the effect from the amplitude model choice.

Solution 1

$$\begin{array}{cc}
\alpha_{f_0(1500)} & (-2.6 \pm 0.7 \pm 14.1) \times 10^3 \\
\alpha_{f_0(2020)} & (-6.4 \pm 0.6 \pm 12.1) \times 10^3 \\
c_K & 3310 \pm 50 \pm 3120 \\
\Im\left(B_{11}^{f_2(1270)}\right) & (3.6 \pm 1.2 \pm 4.0) \times 10^3 \\
\Re\left(B_{11}^{f_2(1270)}\right) & (-1.2 \pm 0.6 \pm 0.9) \times 10^3 \\
\Im\left(B_{11}^{f_2(1525)}\right) & (-1.2 \pm 0.5 \pm 1.1) \times 10^3 \\
\Re\left(B_{11}^{f_2(1525)}\right) & (0.1 \pm 0.5 \pm 2.0) \times 10^3
\end{array}$$

Solution 2

$$\begin{array}{cc}
\alpha_{f_0(1500)} & (-2.1 \pm 0.7 \pm 10.4) \times 10^3 \\
\alpha_{f_0(2020)} & (-6.2 \pm 0.8 \pm 19.1) \times 10^3 \\
c_K & 3350 \pm 50 \pm 3540 \\
\Im\left(B_{11}^{f_2(1270)}\right) & (-2.0 \pm 1.2 \pm 7.8) \times 10^3 \\
\Re\left(B_{11}^{f_2(1270)}\right) & (-1.3 \pm 0.7 \pm 2.5) \times 10^3 \\
\Im\left(B_{11}^{f_2(1525)}\right) & (1.5 \pm 1.0 \pm 5.5) \times 10^3 \\
\Re\left(B_{11}^{f_2(1525)}\right) & (-1.1 \pm 0.5 \pm 1.4) \times 10^3
\end{array}$$

The results for the fit fractions are the following. (Note again that Solution 2 is only shown here for the sake of completeness)

Solution 1

$$\begin{array}{cc}
\text{S-wave} & 0.975 \pm 0.011 \pm 0.015 \\
f_2(1270) & (19 \pm 5 \pm 15) \times 10^{-3} \\
f_2(1525) & (3.1 \pm 1.3 \pm 8.0) \times 10^{-3} \\
Sum & (996.1 \pm 2.6 \pm 9.0) \times 10^{-3}
\end{array}$$

Solution 2

$$\begin{array}{cc}
\text{S-wave} & 0.989 \pm 0.010 \pm 0.018 \\
f_2(1270) & (7 \pm 4 \pm 16) \times 10^{-3} \\
f_2(1525) & (7.0 \pm 3.3 \pm 11.9) \times 10^{-3} \\
Sum & (1003.4 \pm 3.4 \pm 13.7) \times 10^{-3}
\end{array}$$

The absolute values for the systematic uncertainties of the fit fractions are of similar size for all fit components, for the f_2 states these systematic uncertainties are of the same order as their absolute contribution to the nominal fit though, resulting in a large relative uncertainty.

10 Search for $Z(4430)^{\pm}$

In the previous sections only conventional hadronic states were considered in the amplitude fits. Since no bands were visible in the exotic $\psi(2S)\pi^+$ and $\psi(2S)\pi^-$ axes of the Dalitz plot of data, this was justified (see Fig. 6.3). In the following section an upper limit on the possible contribution from the exotic $Z(4430)^{\pm}$ to $B_s^0 \to \psi(2S)\pi^+\pi^-$ is quantified. This tetraquark candidate is introduced in Sec. 2.6.1 and its description within the amplitude model can be found in Sec. 5.3.2. Its Breit-Wigner parameters are fixed to the values obtained in the analysis of $B^0 \to \psi(2S)K^+\pi^-$ [63]. A possible $Z^+(4200)$ state was observed there as well, but its contribution to the decay was measured to be much smaller than that of the $Z(4430)^{\pm}$, making a search for this state infeasible given the current sample size.

In Fig. 10.1 fits to toy samples are shown, where about 40 % of the decays proceed via either $B_s^0 \to Z(4430)^+\pi^-$ or $B_s^0 \to Z(4430)^-\pi^+$. The two configurations use the same coupling strength, but to demonstrate the effect of interferences, their phase relative to the conventional system from the dipion states differ by π. While the effect is less pronounced in the projection to $M(\psi(2S)\pi^+\pi^-)$, the interference pattern drastically influences the distribution in $M_{\pi\pi}$ (compared *e.g.* to Fig 6.9). In both configurations an enhancement is visible for low values of $M_{\pi\pi}$ though, which is not observed in data. The interference effects also drastically change the pattern in the Dalitz plane.

As shown in Sec. 5.7 the coupling strengths of the decays $B_s^0 \to Z(4430)^+\pi^-$ and $B_s^0 \to Z(4430)^-\pi^+$ can be different in principle. Due to the limited data sample size, their magnitudes are set equal and their relative phases are fixed in the amplitude fit. Two choices of the phase are considered, which result in the following relations between the ls coupling terms

$$B_{01}^{Z^+} = B_{01}^{Z^-} = \overline{B}_{01}^{Z^-} = \overline{B}_{01}^{Z^+} \tag{10.1}$$

or

$$B_{01}^{Z^+} = -B_{01}^{Z^-} = \overline{B}_{01}^{Z^-} = -\overline{B}_{01}^{Z^+}. \tag{10.2}$$

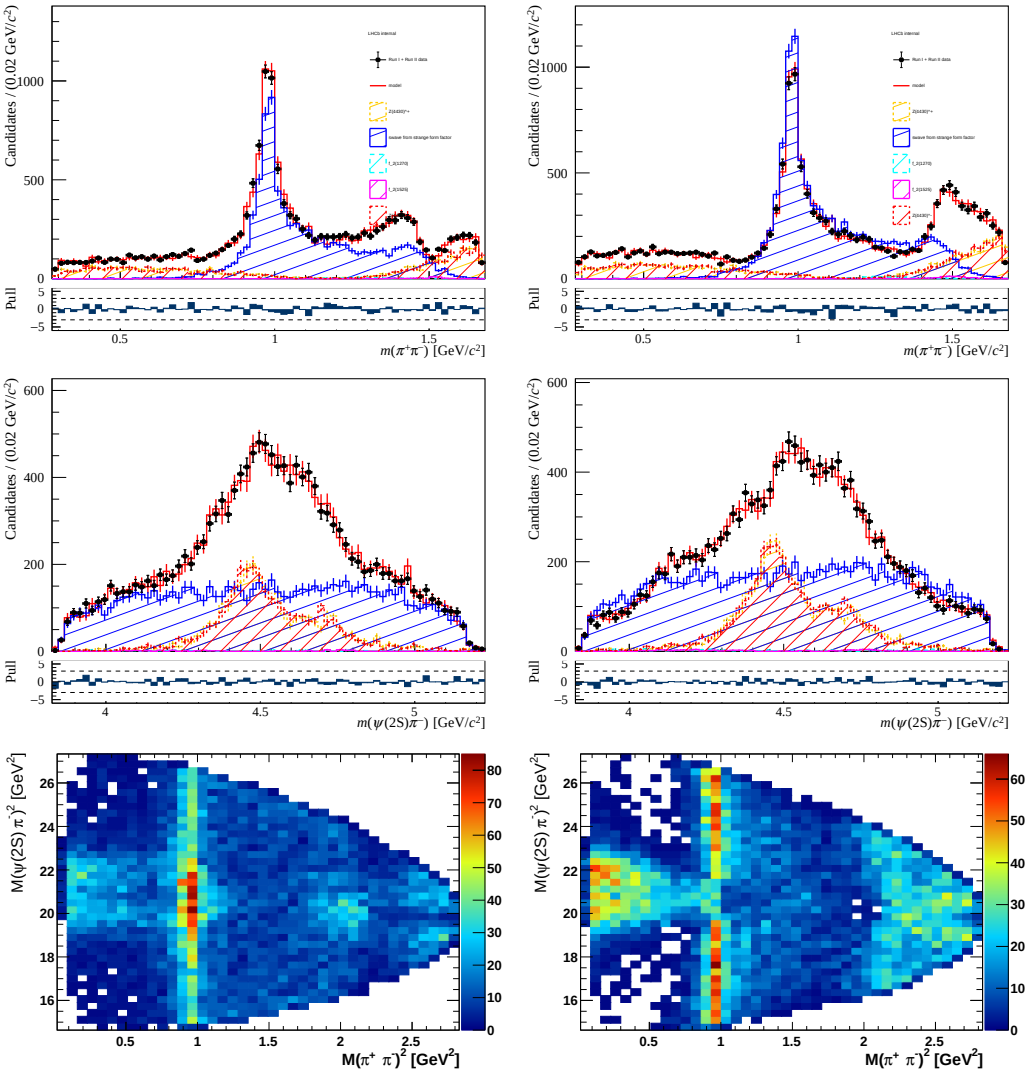

Figure 10.1: Projections of 15000 toy events with 20 % $B_s^0 \to Z(4430)^+\pi^-$ and 20 % $B_s^0 \to Z(4430)^-\pi^+$ fit fractions. Left: Model with relative phase 0.4π between S-wave coupling $B^{\mathrm{S-wave}_{00}}$ and Z^- coupling $B_{01}^{Z^-}$. Right: Model with relative phase -0.6π between S-wave coupling $B^{\mathrm{S-wave}_{00}}$ and Z^- coupling $B_{01}^{Z^-}$. Top: $M_{\pi\pi}$ projection. Middle: $M(\psi(2S)\pi)$ projection. Bottom: Dalitz projection in $M^2(\psi(2S)\pi)$ versus $M^2_{\pi^+\pi^-}$.

The limit-setting procedure is performed once for each choice of the phase and the larger of the two resulting limits is chosen as upper limit.

A common nuisance for searches in particle physics is, that before the measurement it is not clear, whether the final result will be a two sided interval (value and confidence interval) or a one-sided interval (upper limit), since this depends on the significance of a potential signal. If the method is only chosen after the significance is known, the correct coverage of the extracted intervals can be incorrect. This effect is known as "flip-flopping". The approach of Feldman-Cousins unifies the procedure providing both upper limits and two-sided confidence intervals with correct coverage [130] and is therefore chosen here.

The unified Feldman-Cousins approach is itself based on Neyman's construction of confidence intervals, which is introduced here following the description in the review on statistics in [21]. Consider a pdf $f(x|\theta)$, where x represents the measured estimate for coupling strength of the exotic contribution to the decay $B_s^0 \to \psi(2S)\pi^+\pi^-$ and θ is the true value for this coupling strength[1]. For every allowed value of θ and for a given probability $1-\alpha$ one can find intervals $[x_1(\theta,\alpha), x_2(\theta,\alpha)]$ such that

$$P(x_1 < x < x_2;\theta) = \int_{x_1(\theta)}^{x_2(\theta)} f(x;\theta)\mathrm{d}x \leq 1-\alpha. \tag{10.3}$$

For repeated measurements and a true value θ, this interval then contains a fraction of $1-\alpha$ of measurements of x. For continuous x and θ the union of the intervals $[x_1(\theta,\alpha), x_2(\theta\alpha)]$ forms an area in the $x-\theta$ plane, called confidence belt. An example of a confidence belt is shown in Fig. 10.2.

In the measurement a single value of x is determined (in this work the fitted value of the coupling), which singles out a vertical slice $[\theta_1(x), \theta_2(x)]$ of the confidence belt. The so-called confidence interval (CI) for θ from this measurement corresponds to all values of θ, which lie in this interval.

For repeated measurements x, the interval $[\theta_1(x), \theta_2(x)]$ varies. It can be shown that for θ_0 being the true value, a fraction $1-\alpha$ of experiments will contain θ_0. Confidence intervals constructed in this way are therefore said to have a confidence level (CL) of $1-\alpha$.

The interval $[x_1(\theta,\alpha), x_2(\theta,\alpha)]$ above is not uniquely determined. It could could be any interval or even a set of distinct intervals within the allowed

[1] In Sec. 5 the estimate for the coupling strength for decays via $Z(4430)^+$ is denominated $B_{01}^{Z^+}$. In the derivation, the nomenclature from [21] is used instead, which is common in the field.

region for x. The unified approach by Feldman and Cousins provides a method to define unique intervals $[x_1(\theta,\alpha), x_2(\theta\alpha)]$, while also providing the correct coverage.

It introduces a likelihood ratio

$$\lambda(\theta) = \frac{f(x;\theta)}{f(x;\hat{\theta})}, \tag{10.4}$$

where $\hat{\theta}$ is the value for θ, which maximizes $f(x;\theta)$.

From eq. 10.4 one can form the test statistic

$$t(\theta) = -2\log\lambda(\theta). \tag{10.5}$$

The intervals in x are then chosen as the ones, which minimize $t(\theta)$.

The particular formulation of the test statistic in eq. 10.5 was chosen, because according to Wilk's theorem [131] for large sample sizes, $t(\theta)$ asymptotically follows a χ^2 distribution with degrees of freedom equal to the dimensionality of θ, if the null hypothesis $H_0 : \theta = \hat{\theta}$ is true.

Since the χ^2 distribution is well studied, the corresponding values for k_α are known and tabulated (see *e.g.* the review on statistics in [21]). In case of a one-dimensional θ, the null hypothesis H_0 is accepted at significance $\alpha = 0.95$, if [132]

$$t(\theta) < k_\alpha \approx \chi^2_{1-\alpha}(\text{ndof} = 1) = 3.84. \tag{10.6}$$

Determining the interval from this expression is not straightforward, because $f(x;\theta)$ is determined from optimizing the likelihood of the amplitude fit, which contains not only the coupling strength θ, but also nuisance parameters $\boldsymbol{w}'$. In order to make the test statistic independent of these nuisance parameters, the profile likelihood ratio is introduced

$$\lambda_p(\theta) = \frac{\mathcal{L}(\theta, \hat{\hat{\boldsymbol{w}}}'(\theta))}{\mathcal{L}(\hat{\theta}, \hat{\boldsymbol{w}}')}, \tag{10.7}$$

where $\hat{\theta}$ and $\hat{\boldsymbol{w}}'$ are the maximum likelihood estimators from the amplitude fit, while $\hat{\hat{\boldsymbol{w}}}'(\theta)$ is the value for the nuisance parameters, which optimize the likelihood for a fixed value of θ. In practice the likelihoods are optimized on pseudo data generated with the MC technique from the pdf of the amplitude fit. For each tested value of θ_0 a new sample needs to be generated.

In order to find the interval, where relation eq. 10.6 holds, the likelihood in the numerator of eq. 10.4 is optimized for several values of θ around the value

$\hat{\theta}$, as shown in Fig. 10.2. This is only feasible for low-dimensional θ, where only a few values need be scanned.

In order to apply this procedure to the contribution of $Z(4430)^{\pm}$ in $B_s^0 \to \psi(2S)\pi^+\pi^-$ three technicalities need to be addressed:

1. The parameters in the likelihood of the amplitude fit are not directly the fit fraction of $Z(4430)^{\pm}$, but the ℓs couplings $B_{01}^{Z^+}$. The relation between coupling and fit fraction is not straightforward (see eq. 5.71). Therefore the confidence intervals are first determined in the couplings and only afterwards translated into fit fractions.

2. The couplings are complex parameters, which would result in two real parameters for θ, but the described process for the confidence interval estimation is computationally very expensive in more than one dimension. Therefore the 2D scan procedure is only approximated. The phase of the Z coupling is fixed to the value obtained from a fit to data. This reduces the limit-setting to one-dimensional tests. To estimate the effect from fixing the phase, the limit-setting procedure is performed 4 times at the values $\beta' = \beta, \beta + \frac{1}{2}\pi, \beta + \pi, \beta + \frac{3}{2}\pi$, where $\hat{\beta}$ is the value obtained from the fit to data. This still reduces the computational cost significantly compared to a full 2D scan.

3. The s-weighting procedure reduces the effective data sample size. The number of generated pseudo-data events for the profile likelihood test therefore needs to be scaled by the factor α from eq. 6.15 in order to have the same statistical power as the data sample. If this would not be taken into account, the derived confidence intervals would be too small.

   This brings another complication though. Since the sample size of data and that of the generated pseudo data is different, the value of the couplings is different as well. The reason for this is that the likelihood, which is optimized, is extended, so the values of all the couplings in the amplitude fit scale with the number of events in the data sample, on which the likelihood is optimized. In the implementation of the likelihood defined in eq. 6.14 the couplings in the fit to data correspond to a pure signal sample of size $N_{\text{tot}} = N_{\text{sig}} + N_{\text{bg}}$. The couplings extracted from a fit to pseudo-data on the other hand only correspond to a pure signal sample of size N_{sig}, since no background events are simulated. The couplings in the fit to pseudo data are therefore smaller by a factor

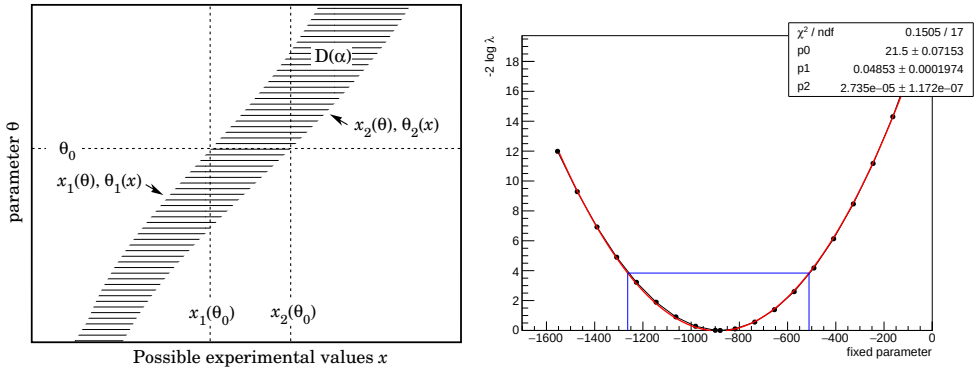

Figure 10.2: Left: Confidence belt example. Taken from [21]. Right: Likelihood for scan of Z coupling magnitude in toy sample generated with a parameter value of -1170. Overlaid with a fit of a second order polynomial. The blue box corresponds to the interval $[x_1(\theta,\alpha), x_2(\theta,\alpha)]$ for $\alpha = 95\,\%$ CL region from Wilk's theorem.

of $\sqrt{\frac{N_{\text{tot}}}{\alpha N_{\text{sig}}}} = 0.8$.[2] This is the reason, why the CI of the fitted value for the parameter is offset by 25 % in Fig. 10.3 from the true parameter on the y-axis. This is only a technical detail though, as the fit fractions $\mathcal{F}_Z$ are unaffected, because a scale factor cancels out in the ratio in eq. 5.71.

The confidence level is therefore determined in the following way:

- Generate a pseudo data sample from the amplitude fit model with an exotic contribution with coupling strength $B_{\text{gen}} = |B|_{\text{gen}} \exp i\beta_{\text{gen}}$. The sample size of the pseudo data sample is $N_{\text{gen}} = 3000$ events, corresponding to the statistical power of the data sample.
- Calculate the profile likelihood ratio for 20 values of $|B|$ around $|B_{\text{gen}}|$.
- Fit $t(\theta)$ with a second order polynomial to determine the region, where $t(\theta)$ fulfils eq.10.2 (see Fig. 10.2).
- Repeat the previous steps for different values of $|B|_{\text{gen}}$ and build the confidence belt for the coupling strength and corresponding fit fractions $\mathcal{F}_Z$ (see example in Fig. 10.3).
- Overlay the confidence belt with the measured value for the fit fraction and read off the confidence limit (see Fig. 10.3).
- Repeat the previous steps for different values of β_{gen}.

[2] The square root here comes from the fact that the pdf is proportional to the squared amplitudes, which contain the couplings as linear parameters.

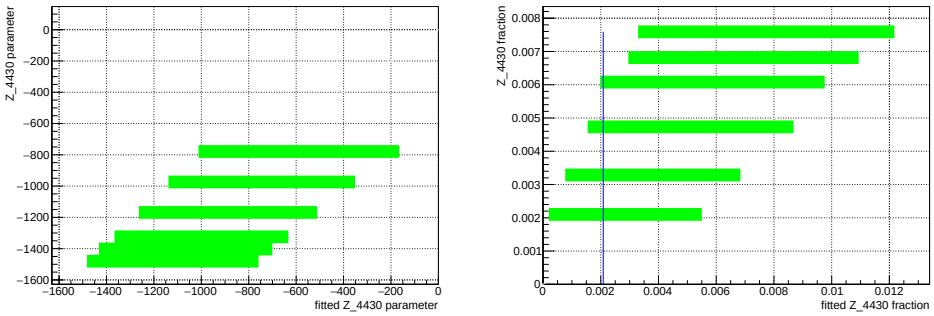

Figure 10.3: Left: Confidence belt on Z coupling magnitude evaluated on toys. Right: Confidence belt on fit fraction $\mathcal{F}_Z$ determined from confidence regions on Z coupling magnitudes evaluated on toys. The vertical line shows the result from a fit to data.

- Perform all previous steps for each of the two hypotheses for the phase between $B_s^0 \to Z(4430)^+\pi^-$ and $B_s^0 \to Z(4430)^-\pi^+$ given in eqs.10.1 and 10.2.

The procedure above results in 8 different limits on the fit fraction $\mathcal{F}_Z$, two from the choice of phase between the $B_s^0 \to Z(4430)^+\pi^-$ and $B_s^0 \to Z(4430)^-\pi^+$ amplitude and four from the overall phase of the exotic amplitudes with respect to the amplitudes of decays via intermediate dipion resonances. These results range from 0.2 % to 1.0 %. As a conservative estimate for the upper limit on the fit fraction, the largest of these values is taken and therefore

$$\mathcal{F}_Z < 1.0\,\% \tag{10.8}$$

at 95 % CL. Note that $\mathcal{F}_Z = \mathcal{F}_{Z^+} = \mathcal{F}_{Z^-}$, since the magnitudes of the couplings of the exotic amplitudes are connected. The upper limit on the fit fraction from either $B_s^0 \to Z(4430)^+\pi^-$ or $B_s^0 \to Z(4430)^-\pi^+$ is therefore $\approx 2\mathcal{F}_Z$, where the $\approx$-sign stems from possible interference effects between the exotic amplitudes from the two decay chains.

As described in Sec. 2.6.1 it is hard to determine theoretical predictions for the fraction in $B_s^0 \to \psi(2S)\pi^+\pi^-$ in the tetraquark and molecule picture. Considering the Feynman diagram in Fig. 2.4, not all the necessary valence quarks are available right after the weak decay process. They only occur after hadronic interactions in the $s\bar{s}$ system, which could reduce the probability of the formation of a $Z(4430)^{\pm}$ drastically.

Since no explicit theory predictions exist for the fraction in $B_s^0 \to \psi(2S)\pi^+\pi^-$, it is compared to the value in $B^0 \to \psi(2S)K^+\pi^-$ [63], where $\mathcal{F}_Z = (5.9 \pm 0.9^{+1.5}_{-3.3})$ % was measured. A fit fraction of that size can be

excluded in in $B_s^0 \to \psi(2S)\pi^+\pi^-$. The limit lies in the region of the observed fraction in $B^0 \to J/\psi K^+\pi^-$ [15], where the $Z(4430)^+$ fraction was measured to $(0.5^{+0.4}_{-0.1})$ %.

The description with a triangle singularity implicitly excludes any contribution in this channel though (see Sec. 2.6.1), an observation here would have therefore strongly disfavoured that model. Since this is not the case, no final conclusion can be drawn and the nature of the $Z(4430)^{\pm}$ remains unknown for now.

11 Conclusions and Outlook

The hadronic partial waves in the decay $B_s^0 \to \psi(2S)\pi^+\pi^-$ are studied in this thesis using an amplitude analysis of the full decay chain. The analysis focuses on a novel description of the dipion S-wave, which uses information extracted from the channels $B_s^0 \to J/\psi\,\pi^+\pi^-$ and $B_s^0 \to J/\psi\,K^+K^-$ [1,46,102], applying it for the first time to a different process. In contrast to the conventional description with Breit-Wigner and Flatté parametrisations, this description of the dipion interaction conserves unitarity and crossing symmetry, allowing to transfer most of its parameters to other processes.

The process-independence of the S-wave description is proven in this dissertation by its successful application to $B_s^0 \to \psi(2S)\pi^+\pi^-$ opening the way to its use in other amplitude analyses with dipion S-wave component.

The Bonn-Jülich parametrisation is compared to a more naive approach describing the dipion S-wave as sum of overlapping Flatté and Breit-Wigner amplitudes. Even though the naive approach can effectively describe the data, it was shown that it cannot be used to extract meaningful estimates of the partial branching fractions for the intermediate hadronic partial waves of the dipion system. This is because the overlapping Flatté and Breit-Wigner resonances produce significant interference terms, which violate unitarity and analyticity making an interpretation of the fit fractions and shape parameters of the amplitudes as partial branching fractions and pole parameters impossible.

The amplitude fit using a Bonn-Jülich parametrisation for the description of the dipion S-wave on the other hand shows much smaller interference effects and the extracted fit fractions shown below can be interpreted as partial branching fractions

$$\begin{array}{cc} \text{S-wave} & 0.974 \pm 0.011 \pm 0.015 \\ f_2(1270) & (17 \pm 5 \pm 15) \times 10^{-3} \\ f_2(1525) & (4.1 \pm 1.3 \pm 8.0) \times 10^{-3} \end{array}$$

After the process independence of the Bonn-Jülich parametrisation is now shown in $B_s^0 \to \psi(2S)\pi^+\pi^-$ the parametrisation can be used to reduce the systematic uncertainty from the modelling of the dipion system in other

channels as well. At the current experimental precision the dipion description is already the leading systematic uncertainty in the measurement of the *CP* violating phase ϕ_s in $B_s^0 \to J/\psi\,\pi^+\pi^-$ [125]. With the 50 fb^{-1} of data from the upcoming runs of the LHCb detector the statistical precision on this observable will decrease significantly and with the improved model it should be possible to keep the systematic uncertainty low enough to make full use of the additional data.

This additional data will also open up the possibility of coupled channel analyses with scalar isospin singlet sources including $B_s^0 \to J/\psi\,\pi^+\pi^-$, $B_s^0 \to J/\psi\,K^+K^-$, $B_s^0 \to \psi(2S)\pi^+\pi^-$, $B_s^0 \to J/\psi\,K^+K^-$, and $J/\psi \to \pi^+\pi^-\gamma$, which should make it possible to increase the precision on the extracted pole positions further, including higher mass resonances within the large phase space in $J/\psi \to \pi^+\pi^-\gamma$. This might also shed light on the possible glueball nature of the high-mass resonances like $f_0(1500)$ and $f_0(2020)$ through the comparison of the relation between the spectra and their production processes described in [133].

The presence of exotic contributions in the form of the $Z(4430)^{\pm}$ in the spectra of $M(\psi(2S)\pi^+)$ and $M(\psi(2S)\pi^-)$ is tested as well, but since no significant signal is seen, an upper limit on its fit fraction in $B_s^0 \to \psi(2S)\pi^+\pi^-$ is determined to $\mathcal{F}_Z < 1.0\,\%$ at 95 % confidence level in agreement with the interpretation of the $Z(4430)^{\pm}$ as triangle singularity. No theoretical predictions exist to date for this value in the tetraquark and molecular models, but they are being worked on and should be available soon.

The next natural step after the presented analysis would be the exploration of the dipion spectrum and possible exotic contributions in $B^0 \to \psi(2S)\pi^+\pi^-$, where the dipion system is produced in an isospin multiplet, built from a $d\bar{d}$ configuration, allowing to test the corresponding model from [46], while the final state quark content would be more favourable for the production of possible $Z(4430)^{\pm}$ under the assumption of a true tetraquark state.

A Preselection Cuts

The first selection step after reconstruction of the events is called Stripping at LHCb. The cuts given in Tab. A.1 are used to select those events with two muons, which form $\psi(2S)$ candidates, which originate in a distance from the V. The Strippings are performed centrally. $DLL_{\mu\pi}$ corresponds to the likelihood difference for a muon over a pion hypothesis of the RICH response.

The trigger requirements for this analysis are given in Tab. A.2. The corresponding cuts can be found in [134].

The other preselection cuts are given in Tab. A.3. The definitions of the abbreviations are given in Sec. 4.

Candidate	Variable	Cut
$\mu^{\pm}$	p_{T}	> 550
	$\mathrm{DLL}_{\mu\pi}$	> 0
	track χ^2/ndf	< 5.0
	global muon PID decision	True
di–μ	Decay vertex χ^2/ndf	< 20.0
	Mass window	[3586.109, 3786.109] MeV
	Decay length significance	> 3

Table A.1: Stripping line selection requirements.

L0	L0Muon or L0DiMuon
HLT1	TrackMuon
HLT2	DiMuonDetachedPsi2S or DiMuonDetachedHeavy

Table A.2: Trigger Requirements

Candidate	Variable	Cut
$\mu^{\pm}$	ghost probability	< 0.5
	IP χ^2	> 9
$\pi^{\pm}$	ghost probability	< 0.5
	p_{T}	> 200
	$\mathrm{DLL}_{K\pi}$	< 5
	IP χ^2	> 6.0
	HASRICH	True
B	DIRA	> 0.9999
	PV χ^2 distance	> 64
	IP χ^2	< 16
	vertex χ^2	< 50
	DTF Mass window	[5150, 5500] MeV
	DTF $\lvert M_{\psi\pi^{\pm}(\pi^{\mp}\to K^{\mp})} - M_{\mathrm{B}^0,\mathrm{PDG}}\rvert$	> 25
	DTF $\lvert M_{\psi(\pi^{\pm}\to K^{\pm})} - M_{\mathrm{B}^0,\mathrm{PDG}}\rvert$	> 30
	DTF $\lvert M_{\psi\pi^{\pm}} - M_{\mathrm{B}^0,\mathrm{PDG}}\rvert$	> 30
	vertex $\Delta\chi^2_{add-track}$	> 10

Table A.3: Preselection requirements.

B Toy study on Oscillation Effects

To test the effect of oscillations, toy samples are generated using eq. 5.47 and compared with toy samples, where $\tau_s = \tau_l$ and the mixing term vanishes. 40k events each are simulated, both samples use the same seed in the generation to reduce statistical noise. Two models are used: a realistic one, with about 96 % S-wave contribution and 1-2 % contribution from f2(1270) and f2(1525) each. Like in the nominal fit only the smallest allowed values for ℓ are considered for the partial waves in the deceay. The comparisons are shown in Fig. B.1. No differences are visible. This is expected, since all involved decay amplitudes have the same *CP* eigenvalues and therefore the mixing term in eq. 5.47 is proportional to the first two terms in that equation.

A second model is generated, which only contains f2(1270) with a significant component of *CP* even and *CP* odd partial waves. In addition to the coupling $B_{11}^{f_2(1270)}$, which describes the strength of the *CP* odd amplitude, $B_{22}^{f_2(1270)}$ is set to finite values as well. The corresponding amplitudes of the higher ℓ state is *CP* even. Since now the decay contains both *CP* even and *CP* odd amplitudes, the mixing term in eq. 5.47 becomes nontrivial and distortions of several percent can be seen in the angular distributions shown in Fig. B.2. The $M_{\pi\pi}$ distributions for both models including B_s^0 oscillation effects are shown in Fig. B.3.

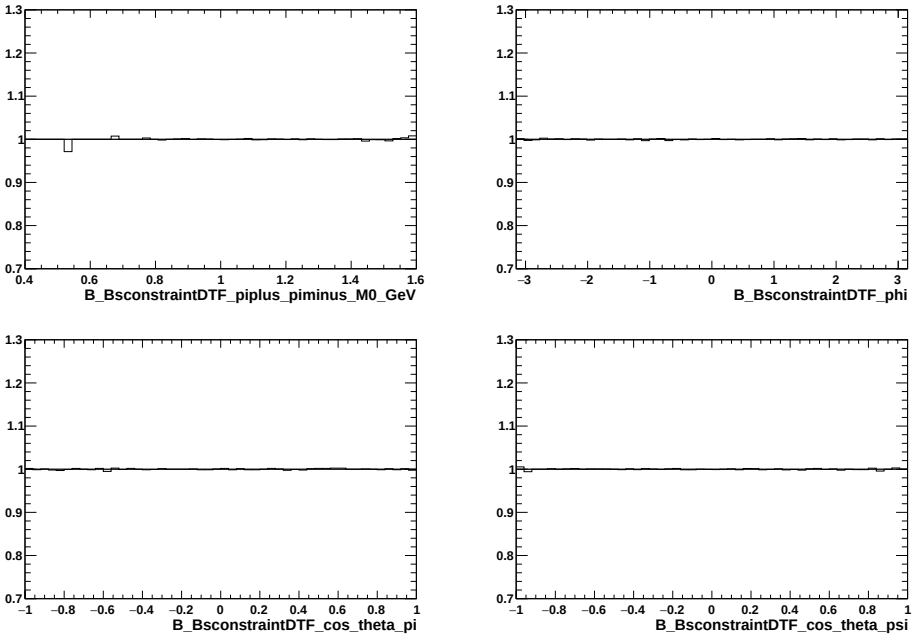

Figure B.1: Ratio of amplitude fit variable distributions for realistic model with small contributions from f2 states. The model in the denominator includes the $B_s^0\ \overline{B_s^0}$ interference term, the one in the nominator does not account for it. The size of the effect is small.

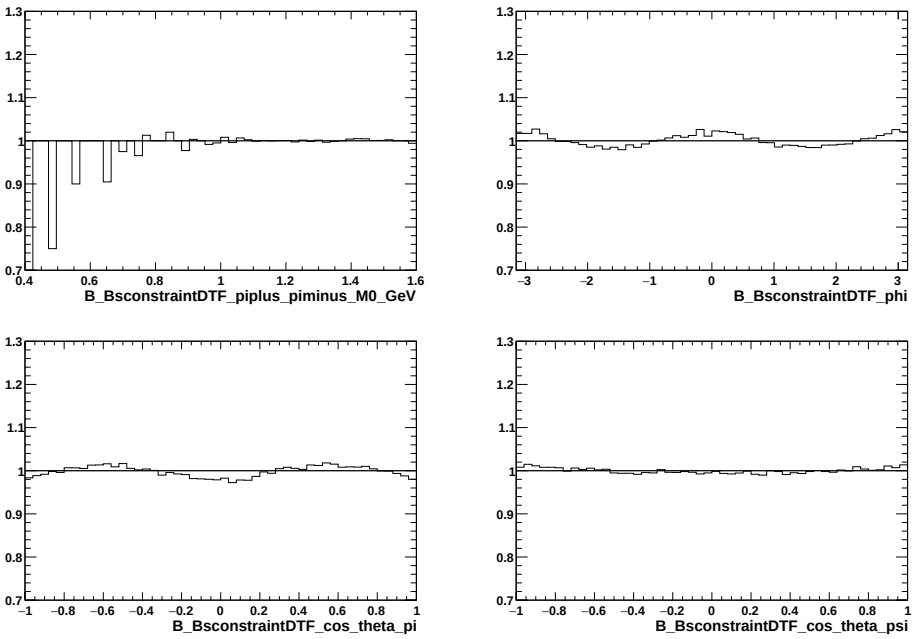

Figure B.2: Ratio of amplitude fit variable distributions for model with only contributions from $f_2(1270)$. The model in the denominator includes the $B_s^0\ \overline{B_s^0}$ interference term, the one in the nominator does not account for it. The effect is sizable, especially in the angle ϕ

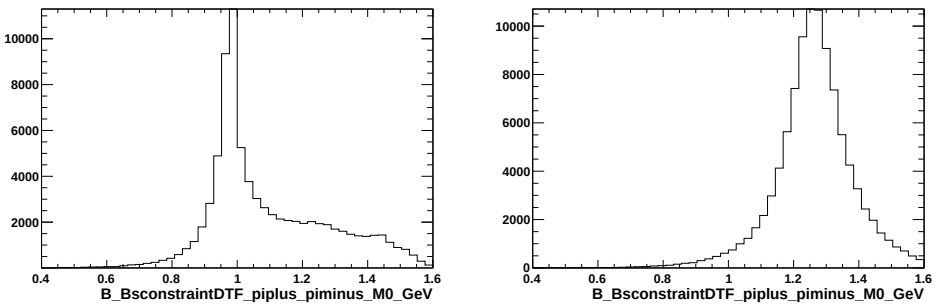

Figure B.3: $M_{\pi\pi}$ distribution of the toy models for the "realistic" model (Left) and the omdel with only contributions from $f_2(1270)$ (Right). Both are simulated taking B_s^0 oscillation effects into account.

Bibliography

[1] S. Ropertz, C. Hanhart, and B. Kubis, *A new parametrization for the scalar pion form factors*, The European Physical Journal C **78** (2018) .

[2] J. Grabowski *et al.*, *Amplitude Analysis of* $B_s \to \psi(2S)\pi^+\pi^-$., , "https://cds.cern.ch/record/2764397".

[3] LHCb collaboration, R. Aaij *et al.*, *Measurement of the branching fraction and CP asymmetry in* $B^+ \to J/\psi\rho^+$ *decays*, Eur. Phys. J. **C79** (2019) 537, arXiv:1812.07041.

[4] R. Davis, D. S. Harmer, and K. C. Hoffman, *Search for Neutrinos from the Sun*, "Phys. Rev. Lett. " **20** (1968) 1205.

[5] Q. R. Ahmad *et al.*, *Measurement of the rate of* $\nu_e + d \to p + p + e-$ *interactions produced by b8 solar neutrinos at the sudbury neutrino observatory*, Physical Review Letters **87** (2001) .

[6] M. B. Gavela, P. Hernandez, J. Orloff, and O. Pène, *Standard model CP violation and baryon asymmetry*, Mod. Phys. Lett. **A9** (1994) 795, arXiv:hep-ph/9312215.

[7] P. Huet and E. Sather, *Electroweak baryogenesis and standard model CP violation*, Phys. Rev. **D51** (1995) 379, arXiv:hep-ph/9404302.

[8] B. Abi *et al.*, *Measurement of the positive muon anomalous magnetic moment to 0.46 ppm*, Physical Review Letters **126** (2021) .

[9] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality in beauty-quark decays*, arXiv:2103.11769, to appear in Nature Physics.

[10] LHCb collaboration, R. Aaij *et al.*, *Evidence for a new structure in the* $J/\psi p$ *and* $J/\psi\overline{p}$ *systems in* $B^0_s \to J/\psi p\overline{p}$ *decays*, arXiv:2108.04720, submitted to PRL.

[11] LHCb collaboration, R. Aaij *et al.*, *Observation of a narrow pentaquark state,* $P_c(4312)^+$*, and of two-peak structure of the* $P_c(4450)^+$, Phys. Rev. Lett. **122** (2019) 222001, arXiv:1904.03947.

[12] Belle Collaboration, A. Bondar *et al.*, *Observation of two charged bottomoniumlike resonances in* $\Upsilon(5s)$ *decays*, Phys. Rev. Lett. **108** (2012) 122001.

[13] LHCb collaboration, R. Aaij *et al.*, *Observation of new resonances decaying to* $J/\psi K^+$ *and* $J/\psi\phi$, Phys. Rev. Lett. **127** (2021) 082001, arXiv:2103.01803.

[14] LHCb, R. Aaij *et al.*, *Observation of the resonant character of the* $Z(4430)^-$ *state*, Phys. Rev. Lett. **112** (2014) 222002, arXiv:1404.1903.

[15] Belle, K. Chilikin *et al.*, *Observation of a new charged charmonium-like state in* $\overline{B}^0 \to J/\psi K^- \pi^+$ *decays*, Phys. Rev. **D90** (2014) 112009, arXiv:1408.6457.

[16] LHCb collaboration, R. Aaij *et al.*, *Observation of an exotic narrow doubly charmed tetraquark*, arXiv:2109.01038, submitted to journal.

[17] LHCb collaboration, R. Aaij *et al.*, *Observation of structure in the* J/ψ*-pair mass spectrum*, Science Bulletin **65** (2020) 1983, arXiv:2006.16957.

[18] T. Aoyama *et al.*, *The anomalous magnetic moment of the muon in the standard model*, Physics Reports **887** (2020) 1–166.

[19] M. Thomson, *Modern particle physics*, Cambridge University Press, New York, 2013.

[20] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*, Addison-Wesley, Reading, USA, 1995.

[21] Particle Data Group, P. A. Zyla *et al.*, *Review of Particle Physics*, PTEP **2020** (2020) 083C01.

[22] J. C. Ward, *An identity in quantum electrodynamics*, Phys. Rev. **78** (1950) 182.

[23] Y. Takahashi, *On the generalized ward identity*, Il Nuovo Cimento **6** (1957) 371.

[24] G. 't Hooft, *Renormalization of massless Yang-Mills fields*, Nuclear Physics B **33** (1971) 173.

[25] S. Weinberg, *A model of leptons*, Phys. Rev. Lett. **19** (1967) 1264.

[26] S. Weinberg, *Implications of dynamical symmetry breaking*, Phys. Rev. D **13** (1976) 974.

[27] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, Phys. Rev. Lett. **13** (1964) 321.

[28] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys. Lett. **12** (1964) 132.

[29] N. Cabibbo, *Unitary symmetry and leptonic decays*, Phys. Rev. Lett. **10** (1963) 531.

[30] M. Kobayashi and T. Maskawa, *CP-Violation in the renormalizable theory of weak interaction*, Prog. Theor. Phys. **49** (1973) 652.

[31] L. Wolfenstein, *Parametrization of the kobayashi-maskawa matrix*, Phys. Rev. Lett. **51** (1983) 1945.

[32] Z. Maki, M. Nakagawa, and S. Sakata, *Remarks on the Unified Model of Elementary Particles*, Progress of Theoretical Physics **28** (1962) 870.

[33] *Constraint on the matter–antimatter symmetry-violating phase in neutrino oscillations*, Nature **580** (2020) 339–344.

[34] G. Zweig, *An SU_3 model for strong interaction symmetry and its breaking; Version 1*, Tech. Rep. CERN-TH-401, CERN, Geneva, Jan, 1964.

[35] M. Gell-Mann, *A Schematic Model of Baryons and Mesons*, Phys. Lett. **8** (1964) 214.

[36] W. Commons, *File:meson nonet - spin 0.svg — wikimedia commons, the free media repository*, 2020. [Online; accessed 12-May-2021] https://commons.wikimedia.org/w/index.php?title=File:Meson_nonet_-_spin_0.svg&oldid=458584469.

[37] V. Koch, *Aspects of chiral symmetry*, International Journal of Modern Physics E **06** (1997) 203–249.

[38] J. A. Oller, *A Brief Introduction to Dispersion Relations*, SpringerBriefs in Physics, Springer, 2019.

[39] Stefan Ropertz, *A formalism for a consistent treatment of two-pion interactions in heavy meson decays*, Oct., 2020. http://hdl.handle.net/20.500.11811/8670.

[40] S. Mandelstam, *Determination of the pion - nucleon scattering amplitude from dispersion relations and unitarity. General theory*, Phys. Rev. **112** (1958) 1344.

[41] M. Jacob and G. C. Wick, *On the General Theory of Collisions for Particles with Spin*, Annals Phys. **7** (1959) 404.

[42] S. U. Chung, *Spin formalisms*, 3, 1971. doi: 10.5170/CERN-1971-008.

[43] Particle Data Group, M. Tanabashi *et al.*, *Review of Particle Physics*, Phys. Rev. D **98** (2018) 030001.

[44] J. M. Blatt and V. F. Weisskopf, *Theoretical nuclear physics*, Springer, New York, 1952.

[45] C. Hanhart, *A new parameterization for the pion vector form factor*, Physics Letters B **715** (2012) 170–177.

[46] J. T. Daub, C. Hanhart, and B. Kubis, *A model-independent analysis of final-state interactions in* $\overline{B}^0_{d/s} \to J/\psi\pi\pi$, JHEP **02** (2016) 009, arXiv:1508.06841.

[47] LHCb collaboration, R. Aaij *et al.*, *Measurement of the resonant and CP components in* $\overline{B}^0 \to J/\psi\,\pi^+\pi^-$ *decays*, Phys. Rev. **D90** (2014) 012003, arXiv:1404.5673.

[48] K. Nakano, *Two-potential formalisms and the Coulomb-nuclear interference*, Phys. Rev. C **26** (1982) 1123.

[49] Y. Nambu, *Force Potentials in Quantum Field Theory**, Progress of Theoretical Physics **5** (1950) 614.

[50] P. Masjuan and J. J. Sanz-Cillero, *Padé approximants and resonance poles*, The European Physical Journal C **73** (2013) .

[51] P. Masjuan, J. Ruiz de Elvira, and J. J. Sanz-Cillero, *Precise determination of resonance pole parameters through Padé approximants*, Physical Review D **90** (2014) .

[52] I. Caprini, P. Masjuan, J. Ruiz de Elvira, and J. J. Sanz-Cillero, *Uncertainty estimates of the* σ*-pole determination by padé approximants*, Physical Review D **93** (2016) .

[53] LHCb collaboration, R. Aaij *et al.*, *Precise measurement of the* f_s/f_d *ratio of fragmentation fractions and of* B^0_s *decay branching fractions*, arXiv:2103.06810, to appear in PRD.

[54] J. Iizuka, *A Systematics and Phenomenology of Meson Family**, Progress of Theoretical Physics Supplement **37-38** (1966) 21.

[55] S. Okubo, *ϕ-meson and unitary symmetry model*, Physics Letters **5** (1963) 165.

[56] LHCb collaboration, R. Aaij *et al.*, *Measurement of resonant and CP components in $\bar{B}^0_s \to J/\psi\,\pi^+\pi^-$ decays*, Phys. Rev. **D89** (2014) 092006, arXiv:1402.6248.

[57] Z. Zhao *et al.*, *Study of light tetraquark spectroscopy*, arXiv:2108.06155.

[58] S.-K. Choi *et al.*, *Observation of a narrow charmoniumlike state in exclusive $B^\pm \to K^\pm \pi^+\pi^- J/\psi$ decays*, Physical Review Letters **91** (2003) .

[59] LHCb collaboration, R. Aaij *et al.*, *Quantum numbers of the $X(3872)$ state and orbital angular momentum in its $\rho^0 J/\psi$ decays*, Phys. Rev. **D92** (2015) 011102(R), arXiv:1504.06339.

[60] LHCb collaboration, R. Aaij *et al.*, *Observation of $J/\psi\,p$ resonances consistent with pentaquark states in $\Lambda_b^0 \to J/\psi\,pK^-$ decays*, Phys. Rev. Lett. **115** (2015) 072001, arXiv:1507.03414.

[61] Belle, S. K. Choi *et al.*, *Observation of a resonance-like structure in the pi+- psi-prime mass distribution in exclusive B —> K pi+- psi-prime decays*, Phys. Rev. Lett. **100** (2008) 142001, arXiv:0708.1790.

[62] J. Bressieux *et al.*, *Evidence for the resonant character of the $Z(4430)^- \to \psi(2S)\pi^-$ mass peak observed in $B^0 \to \psi(2S)K^+\pi^-$ decays, and determination of the $Z(4430)^-$ spin-parity*, , https://cds.cern.ch/record/1553766.

[63] 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.

[64] LHCb collaboration, R. Aaij *et al.*, *Study of Υ production and cold nuclear matter effects in pPb collisions at $\sqrt{s_{NN}}$ =5 TeV*, JHEP **07** (2014) 094, arXiv:1405.5152.

[65] LHCb collaboration, R. Aaij *et al.*, *Model-independent confirmation of the $Z(4430)^-$ state*, Phys. Rev. **D92** (2015) 112009, arXiv:1510.01951.

[66] P. Pakhlov and T. Uglov, *Charged charmonium-likez+(4430)from rescattering in conventional b decays*, Physics Letters B **748** (2015) 183–186.

[67] S. X. Nakamura and K. Tsushima, *$Z_c(4430)$ and $Z_c(4200)$ as triangle singularities*, Phys. Rev. D **100** (2019) 051502, arXiv:1901.07385.

[68] V. Gribov, *Strong Interactions of Hadrons at High Energies: Gribov Lectures on Theoretical Physics*, Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology, Cambridge University Press, 2008.

[69] *LHC Machine*, JINST **3** (2008) S08001.

[70] ATLAS, A. Airapetian *et al.*, *ATLAS: Detector and physics performance technical design report. Volume 2*, .

[71] CMS, G. L. Bayatian *et al.*, *CMS Physics: Technical Design Report Volume 1: Detector Performance and Software*, .

[72] ALICE, N. Antoniou *et al.*, *Letter of Intent for A Large Ion Collider Experiment*, .

[73] LHCb collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008) S08005.

[74] T. L. Collaboration *et al.*, *The LHCf detector at the CERN large hadron collider*, Journal of Instrumentation **3** (2008) S08006.

[75] MoEDAL Collaboration, J. Pinfold *et al.*, *Technical Design Report of the MoEDAL Experiment*, tech. rep., Jun, 2009.

[76] TOTEM, G. Anelli *et al.*, *The TOTEM experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08007.

[77] F. Collaboration *et al.*, *Technical proposal for faser: Forward search experiment at the lhc*, arXiv:1812.09139.

[78] R. Lindner, *LHCb layout_2. LHCb schema_2*, LHCb Collection https://cds.cern.ch/record/1087860, Feb, 2008.

[79] A. Powell *et al.*, *Particle identification at LHCb*, PoS **ICHEP2010** (2010) 020, LHCb-PROC-2011-008.

[80] G. Dujany and B. Storaci, *Real-time alignment and calibration of the LHCb Detector in Run II*, J. Phys. Conf. Ser. **664** (2015) 082010.

[81] G. Barrand *et al.*, *Gaudi — a software architecture and framework for building hep data processing applications*, Computer Physics Communications **140** (2001) 45, CHEP2000.

[82] M. Clemencic *et al.*, *The LHCb simulation application, Gauss: Design, evolution and experience*, J. Phys. Conf. Ser. **331** (2011) 032023.

[83] I. Belyaev *et al.*, *Handling of the generation of primary events in Gauss, the LHCb simulation framework*, J. Phys. Conf. Ser. **331** (2011) 032047.

[84] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.

[85] P. Golonka and Z. Was, *PHOTOS Monte Carlo: A precision tool for QED corrections in Z and W decays*, Eur. Phys. J. **C45** (2006) 97, arXiv:hep-ph/0506026.

[86] Geant4 collaboration, S. Agostinelli *et al.*, *Geant4: A simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.

[87] Geant4 collaboration, J. Allison *et al.*, *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. **53** (2006) 270.

[88] LHCb collaboration, *The BOOLE Project*, http://lhcbdoc.web.cern.ch/lhcbdoc/boole/.

[89] LHCb collaboration, R. Aaij *et al.*, *Measurement of the track reconstruction efficiency at LHCb*, JINST **10** (2015) P02007, arXiv:1408.1251.

[90] L. Anderlini *et al.*, *The PIDCalib package*, LHCb-PUB-2016-021.

[91] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, Nucl. Instrum. Meth. **A552** (2005) 566, arXiv:physics/0503191.

[92] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees*, Wadsworth international group, Belmont, California, USA, 1984.

[93] A. Hoecker *et al.*, *TMVA 4 — Toolkit for Multivariate Data Analysis. Users Guide.*, arXiv:physics/0703039.

[94] Y. Freund and R. E. Schapire, *A decision-theoretic generalization of on-line learning and an application to boosting*, J. Comput. Syst. Sci. **55** (1997) 119.

[95] J. H. Friedman, *Greedy function approximation: A gradient boosting machine.*, The Annals of Statistics **29** (2001) 1189 .

[96] C. Gini, *Measurement of inequality of incomes*, The Economic Journal **31** (1921) 124.

[97] H. Voss, A. Hoecker, J. Stelzer, and F. Tegenfeldt, *TMVA - Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040.

[98] *Approximate pdf for the invariant mass distribution of combinatorial background*, https://github.com/HDembinski/essays/blob/master/invariant%20mass%20combinatorial%20background.ipynb, 2021. accessed 21-06-11.

[99] 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.

[100] P. Koppenburg, *Statistical biases in measurements with multiple candidates*, arXiv:1703.01128.

[101] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, arXiv:physics/0402083.

[102] S. Ropertz, C. Hanhart, and B. Kubis, *Scalar isoscalar form factors for energies above 1 GeV*, EPJ Web Conf. **202** (2019) 06002.

[103] LHCb, R. Aaij *et al.*, *Measurement of resonant and CP components in* $\bar{B}^0_s \to J/\psi\pi^+\pi^-$ *decays*, Phys. Rev. **D89** (2014) 092006, arXiv:1402.6248.

[104] M. Mikhasenko *et al.*, *Dalitz-plot decomposition for three-body decays*, Physical Review D **101** (2020) .

[105] S.-U. Chung and J. M. Friedrich, *Covariant helicity-coupling amplitudes: A new formulation*, Physical Review D **78** (2008) .

[106] V. Filippini, A. Fontana, and A. Rotondi, *Covariant spin tensors in meson spectroscopy*, Phys. Rev. D **51** (1995) 2247.

[107] J. Collaboration *et al.*, *What is the right formalism to search for resonances?*, arXiv:1712.02815.

[108] L. Zhang and S. Stone, *Time-dependent Dalitz-plot formalism for* ${B_q}^0 \to J/\psi h^+ h^-$, Physics Letters B **719** (2013) 383, arXiv:1212.6434.

[109] U. Nierste, *Three Lectures on Meson Mixing and CKM phenomenology*, 2009.

[110] LHCb collaboration, R. Aaij *et al.*, *Measurement of the CP-violating phase* ϕ_s *in* $\bar{B}^0_s \to J/\psi\,\pi^+\pi^-$ *decays*, Phys. Lett. **B736** (2014) 186, arXiv:1405.4140.

[111] LHCb collaboration, R. Aaij *et al.*, *Measurement of* B^0, B^0_s, B^+ *and* Λ^0_b *production asymmetries in 7 and 8 TeV proton-proton collisions*, Phys. Lett. **B774** (2017) 139, arXiv:1703.08464.

[112] LHCb collaboration, R. Aaij *et al.*, *Measurement of the CP asymmetry in B_s^0–$\bar{B}_s^0$ mixing*, Phys. Rev. Lett. **117** (2016) 061803, arXiv:1605.09768.

[113] HFLAV, Y. S. Amhis *et al.*, *Averages of b-hadron, c-hadron, and τ-lepton properties as of 2018*, Eur. Phys. J. **C81** (2021) 226, arXiv:1909.12524, updated results and plots available at https://hflav.web.cern.ch/.

[114] H. G. Katzgraber, *Introduction to monte carlo methods*, arXiv:0905.1629.

[115] S. Neubert, A. Piucci, and N. A. Skidmore, *Amplitude analysis of $\Lambda_b^0 \to \Lambda_c^+ \overline{D}^0 K^-$ decays and search for pentaquarks in the $\Lambda_c^+ \overline{D}^0$ system* , , "https://cds.cern.ch/record/2680864".

[116] F. Hu and J. Zidek, *The Weighted Likelihood*, The Canadian Journal of Statistics **30** (2002) 347.

[117] X. Wang and J. Zidek, *Selecting likelihood weights by cross-validation*, Annals of Statistics **33** (2005) 463, arXiv:math/0505599.

[118] Y. Xie, *sFit: a method for background subtraction in maximum likelihood fit*, arXiv:0905.0724.

[119] H. J. Thiébaux and F. W. Zwiers, *The interpretation and estimation of effective sample size*, Journal of Climate and Applied Meteorology **23** (1984) 800.

[120] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $B^\pm$ production cross-section in pp collisions at $\sqrt{s}$ =7 and 13 TeV*, JHEP **12** (2017) 026, arXiv:1710.04921.

[121] LHCb Collaboration, R. Aaij *et al.*, *Study of the D^0 p amplitude in Λ_b^0 to D^0 $p\pi^-$ decays*, J. High Energy Phys. **2017** (2017) .

[122] V. Baru *et al.*, *Flatte-like distributions and the a(0)(980) / f(0)(980) mesons*, Eur. Phys. J. A **23** (2005) 523, arXiv:nucl-th/0410099.

[123] *Website of the ROOT implementation of Minuit*, https://root.cern.ch/doc/master/classTMinuit.html.

[124] L. Demortier and L. Lyons, *Everything you always wanted to know about pulls*, , https://lucdemortier.github.io/assets/papers/cdf5776_pulls.pdf.

[125] LHCb collaboration, R. Aaij *et al.*, *Measurement of the CP-violating phase ϕ_s from $B_s^0 \to J/\psi \pi^+ \pi^-$ decays in 13 TeV pp collisions*, Phys. Lett. **B797** (2019) 134789, arXiv:1903.05530.

[126] B. Efron, *Bootstrap methods: Another look at the jackknife*, Ann. Statist. **7** (1979) 1.

[127] C. Langenbruch, *Parameter uncertainties in weighted unbinned maximum likelihood fits*, arXiv:1911.01303.

[128] B. Efron, *Bootstrap methods: Another look at the jackknife*, Ann. Statist. **7** (1979) 1.

[129] D. Martínez Santos and F. Dupertuis, *Mass distributions marginalized over per-event errors*, Nucl. Instrum. Meth. **A764** (2014) 150, arXiv:1312.5000.

[130] G. J. Feldman and R. D. Cousins, *Unified approach to the classical statistical analysis of small signals*, Physical Review D **57** (1998) 3873–3889.

[131] S. S. Wilks, *The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses*, Ann. Math. Statis. **9** (1938) 60.

[132] 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) .

[133] E. Klempt, *Scalar mesons and the fragmented glueball*, arXiv:2104.09922.

[134] V. Gligorov, C. Thomas, and M. Williams, *The HLT inclusive B triggers*, Tech. Rep. LHCb-PUB-2011-016, 2011. https://cds.cern.ch/record/1384380.

Acknowledgements

This thesis would not have been possible without the help of numerous people, who I would like to thank.

My supervisor Ulrich Uwer, who gave me the opportunity to join his group for this thesis and for all his advice during the last years.

Michel De Cian and Sevda Esen, who I worked with during my first analysis at LHCb during my Master thesis and at the beginning of my PhD.

Sebastian Neubert and his spectroscopy group, who introduced me into the colourful world of hadron physics in the first place and worked with me on the topics presented in this thesis.

Christoph Hanhart, Bastian Kubis, and Stefan Ropertz, whose dispersive model for the description of the dipion S-wave is a cornerstone of this work and whose help was invaluable, when I implemented their model into my amplitude fit.

Thomas Nikodem, Marian Stahl and Alessio Piucci, from whose expertise in C++, ROOT, Tensorflow, Tiramisu, and Nussecken I profited very much.

The whole LHCb Heidelberg group, for the great working environment, all the shared cakes and coffee breaks.

My family and friends, both from Heidelberg and back home in Coburg, who have given me their support and believed in me through the years.