Universität Zürich[UZH]

Measurement of Asymmetries in the Decay $B^0 \to K^{*0}\mu^+\mu^-$ at the LHCb Experiment

---

Master Thesis

of

Marco Tresch

Mathematisch-naturwissenschaftliche Fakultät

der

Universität Zürich

Prof. Dr. U. Straumann

Dr. N. Serra

M. De Cian

Zürich

2012

Abstract

The decay $B^0 \to K^{*0}\mu^+\mu^-$ is one of the most promising for the search of New Physics at LHCb. It has several observables which are sensitive to physics beyond the Standard Model and where theoretical uncertainties are under control. One of them is the zero-crossing point of the forward-backward asymmetry, A_{FB}. Early measurements of the A_{FB} by the B-factories and CDF seemed to hint to a deviation from Standard Model predictions at low q^2, in particular to an opposite sign of the $\mathcal{C}_7$ Wilson coefficient. This would have implied no zero-crossing in the A_{FB}. This thesis discusses the extraction of the zero-crossing point of A_{FB}, which uses an empirical model to fit the data. A crucial aspect in this measurement is the validation of this fitting model. This is done by using a multidimensional unbinned goodness of fit method, which is described in details. Measurements of other asymmetries sensitive to New Physics, that can be extracted with counting experiment techniques, are also studied.

Contents

1 The LHCb Experiment **8**

1.1 The LHCb Detector 8

2 Theory **11**

2.1 The Standard Model 11

2.2 New Physics search at LHCb 12

2.3 Decay $B^0 \to K^{*0}\mu^+\mu^-$ 14

2.3.1 Kinematics 14

2.3.2 Observables 15

3 The Unbinned Counting Method **21**

3.1 Forward-Backward Asymmetry 22

3.2 S_i Asymmetries 22

3.2.1 Consistency Test 25

3.2.2 Zero-Crossing Point Extraction 30

4 Goodness of Fit Test **32**

4.1 Terminology 32

4.2 Point-to-Point Dissimilarity Method 33

4.2.1 Physical Motivation 33

4.2.2 Goodness of Fit Implementation 34

4.3 Measurement 37

4.3.1 Fit Models 37

4.3.2 p-Value Distributions 37

5 Measurement **44**

5.1 Selection 44

5.2 $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to J/\psi K^{*0}$ invariant mass distributions 45

5.3 Zero-Crossing Point of A_{FB} 46

5.3.1 Result 46

6 Summary **48**

Appendix **52**

A RootFinder Method **52**

Introduction

Since 2010 the LHC is running at energy of $\sqrt{s} = 7$ TeV and the experiments are collecting data. One of the main goals of the LHC is the search for physics beyond the Standard Model (SM), often referred to as New Physics (NP). The LHCb experiment is investigating NP by performing measurements of Heavy Flavor particles, hadrons containing charm and beauty quarks. One of the most promising decay for these searches is the channel $B^0 \to K^{*0}\mu^+\mu^-$, which proceeds through a flavour changing neutral current (FCNC). As FCNC transitions are only possible in the SM trough loop or box diagrams, NP processes can enter at the same level as SM processes. In the $B^0 \to K^{*0}\mu^+\mu^-$ decay there are several angular observables, where theoretical uncertainties are under control, and which are sensitive to NP. Specially the zero-crossing point of the forward-backward asymmetry A_{FB}, where form factor uncertainties cancel out (Ref. [23]). Other observables sensitive to NP are the S_i asymmetries [23], in particular the asymmetries S_4, S_5, S_7 and S_8 have never been measured.

This thesis describes the method used for the extraction of the zero-crossing point of A_{FB}, known as unbinned counting method [8]. In particular it focuses on the validation of the empirical model used to fit the data. The goodness of the fit is measured with a point-to-point dissimilarity technique. Sensitivity studies for the measurement of the observables S_4, S_5, S_7 and S_8 with counting experiment techniques are performed. The extraction of the zero-crossing point of the asymmetries S_4 and S_5, where form factor uncertainties cancel out (in the same way as for A_{FB}), is also studied.

Outline: Chapter one gives a short overview of the LHCb experiment. In chapter two the Standard Model is summarised and the decay $B^0 \to K^{*0}\mu^+\mu^-$ is discussed from the theoretical point of view. In chapter three the unbinned counting method, used to extract the zero-crossing point of A_{FB} is described. Sensitivity studies for the measurement of S_4, S_5, S_7 and S_8 are performed and discussed. In chapter four is described the multidimensional goodness of fit method. In chapter five the method is applied to data and the results are discussed. In the last chapter the results of this thesis are summarised. short summary and an outlook.

1. The LHCb Experiment

The Large Hadron Collider (LHC) is a proton-proton accelerator and collider installed in a tunnel with 26.7 km circumference at CERN, in Geneva. The LHC was operating at the center-of-mass energy of 7 TeV during the years 2010 and 2011. The collider has four collision points, where the four main experiments are located: ATLAS, ALICE, CMS and LHCb. For this thesis 1 fb^{-1} of data, collected at LHCb during 2011 are used.

The CMS and ATLAS experiments investigate a wide range of physics. One of their main goals is the search for the Higgs boson, which was the only yet unobserved particle of the SM. In July 2012, CMS and ATLAS observed a new particle with mass at 125 GeV/c^2, compatible with being the Higgs boson (Ref. [9]). The ALICE experiment investigates the quark-gluon plasma in heavy ion collisions.

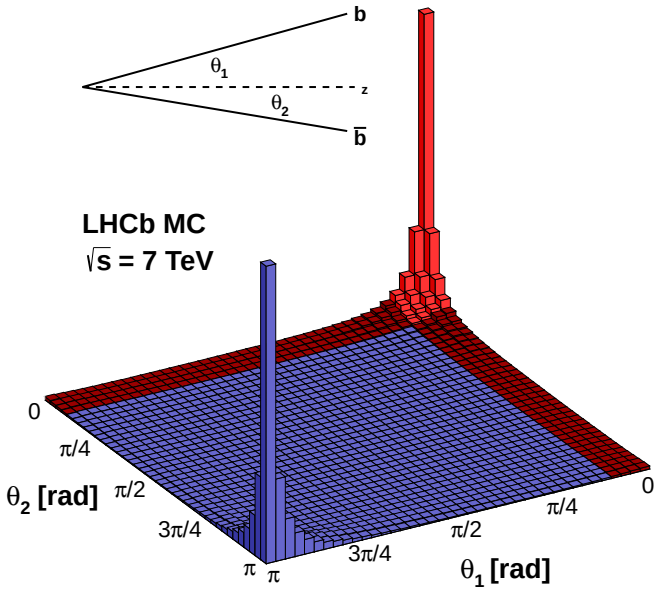


Figure 1: Correlation of the polar angle of the b and $\bar{b}$ quarks at LHC.

LHCb is the only LHC experiment dedicated to flavour physics. Its primary goal is to search for NP by doing precise measurements of rare decays and *CP* violation in beauty and charm hadrons. In the following a short description of the LHCb detector is given.

1.1. The LHCb Detector

In proton-proton collisions beauty hadrons are mainly produced at small angles with respect to the beam pipe (see Fig. 1). This justifies the geometry of the LHCb detector,

which is a single arm forward spectrometer with an acceptance of $\eta = 1.6 - 4.9$ [1] (For more detail see Ref. [1]).

Tracking System: The tracking system of LHCb is split up in stations upstream and downstream of the dipole magnet. The vertex locator (VELO) and the Tracker Turicensis (TT) are located upstream of the magnet, whereas three planar tracking stations (T1-T3) are placed downstream of the magnet. VELO and TT are build on silicon microstrip detectors. For the region close to the beam pipe of T1-T3 (Inner Tracker, IT) silicon microstrip detectors are used. In the outer region of T1-T3 straw-tubes were installed. The tracking system combined with the magnet is able to measure the momentum and the flight direction of the charged particles with a resolution of $\frac{\Delta p}{p} \sim 0.5\%$ for particles in the momentum range 5$\sim$100 GeV/c^2.

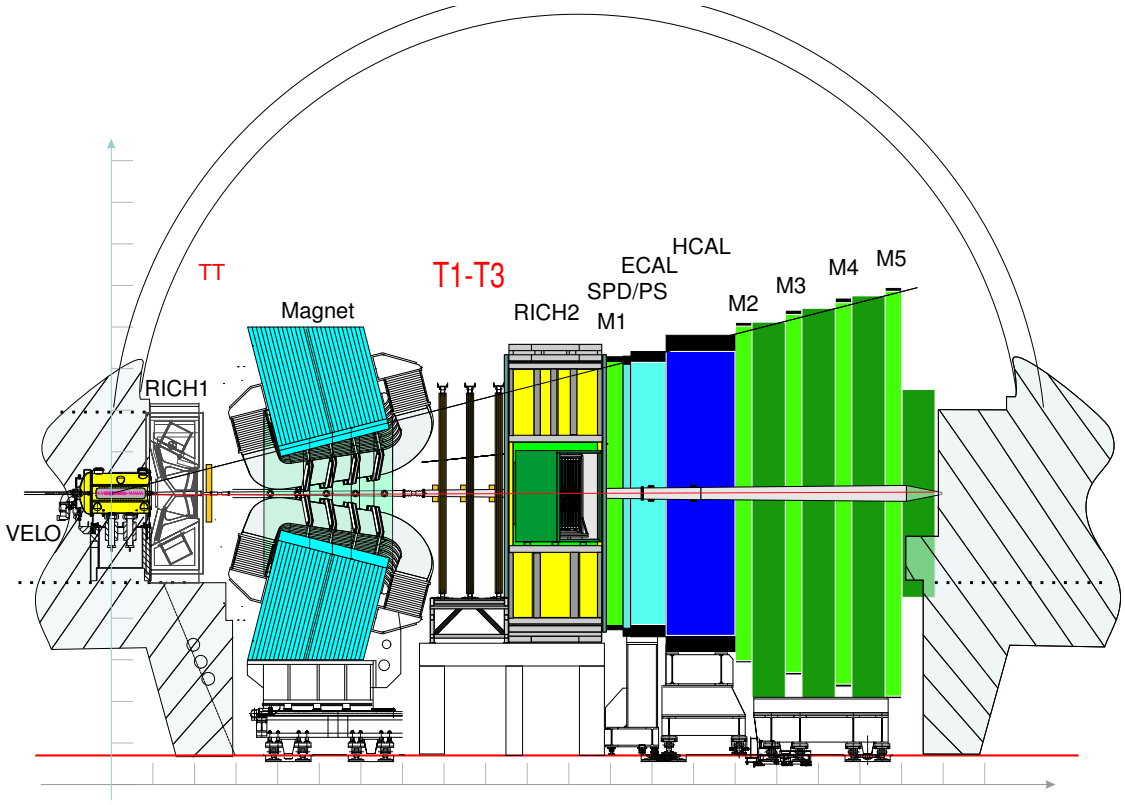


Figure 2: Schematic overview of the LHCb detector. In the figure the different subdetectors are shown: The vertex locator (VELO), the two Ring Imaging Cherenkov Detectors (RICH1 and RICH2), the tracking stations (TT and T1-T3), the scintillator pad detector (SPD), the preshower (PS), the electromagnetic calorimeter (ECAL), the hadronic calorimeter (HCAL) and the muon stations (M1-M5).[1]

Hadronic Particle Identification: For the separation of pions from kaons in selected B hadron decays, the particle identification (PID) is fundamental. The muon system, the calorimeters and the two Ring Imaging Cherenkov detectors (RICH) provide PID

[1] η = -ln(tan(θ/2)), where θ is the angle between beam-pipe and momentum direction, [27]

information. The muon system and the calorimeters are described in separate paragraphs.

The RICH1, situated between the VELO and the TT, covers the low momentum charged particles (p = 1-60 GeV/c) using aerogel and C_4F_{10} as Cherenkov active material (radiator). This means the light velocity in the material is lower then the velocity of the particle which flight trough the material ($v < c/n$, n refraction index of the material). The RICH2, situated between T3 and M1, covers the high momentum charged particles (p = 15-100 GeV/c) using a CF_4 radiator.

Both RICH detectors focus the Cherenkov light using a combination of spherical and flat mirrors, which reflect the image out of the LHCb acceptance to the photomultipliers.

Calorimeters: The LHCb calorimeter system has a classical structure of an electromagnetic calorimeter (ECAL) followed by a hadronic calorimeter (HCAL). The ECAL is a sampling scintillator/lead structure read out by plastic wave length shifting fibers. The HCAL is a sampling device consisting of iron and scintillating tiles, as absorber and active material respectively. To reject background of soft neutral and charged pions Scintillating Pad Detector (SPD) and a Preshower (PS) are placed in front of the ECAL.

Muon System: The muon system consists of five stations (M1-M5) placed along the beam axis. Stations M2 to M5 are placed downstream of the calorimeters and are interleaved with iron absorbers to select muons. M1 is placed between RICH2 and SPD/PS to improve the p_T measurement in the trigger. The detectors are based on Multi-wire proportional chambers (MWPC), except for the inner region of station M1, where the particle rate exceeds safety limits for ageing. In this region triple-GEM detectors are used.

Trigger: The LHCb trigger system consists of three stages. The first stage is the L0-trigger which is implemented in hardware and reduces the visible event rate from 12-15 MHz to 1 MHz. It selects events with large p_T muons in the muon chambers and large E_T hadron, electron and photon clusters in the calorimeters.

The second stage (HLT1-trigger) performs a partial reconstruction of the event and searches for a single track with high momentum, a large impact parameter with respect to all primary vertices in the event, and a good track quality. In addition it uses lifetime information to reduce the rate by a factor of ~ 20.

The third stage (HLT2-trigger) uses the information of fully reconstructed events to reduce the rate to ~ 3 kHz. All events passing the HLT2-trigger are saved to tape [31].

2. Theory

This chapter gives an overview of the SM of particle physics. Furthermore it contains the theory description of the decay $B^0 \to K^{*0}\mu^+\mu^-$ and of the observables sensitive to NP: A_{FB}, S_4, S_5, S_7 and S_8.

2.1. The Standard Model

The SM of particle physics describes the elementary particles and their interactions (electromagnetic, weak and strong). The particles consist of fermions (spin-$\frac{1}{2}$), gauge bosons (spin-1) and the Higgs boson (spin-0).

The elementary fermions in the SM are the 6 leptons and the 6 quarks, each grouped in 3 families (see Fig. 3). Non-elementary fermions and bosons are made of the 6 quarks, for example the proton and the neutron. Both are made of up- and down-quarks. The proton, for instance, is composed by up-up-down.

The gauge bosons are the gauge mediators of the electroweak and strong forces; for example, the photon transmits the electromagnetic force. The SM has 4 gauge bosons, two are massless (photon and gluon) and two are massive (W and Z). The Higgs boson is needed in the SM to break the electroweak symmetry and give mass to SM particles, as observed experimentally.

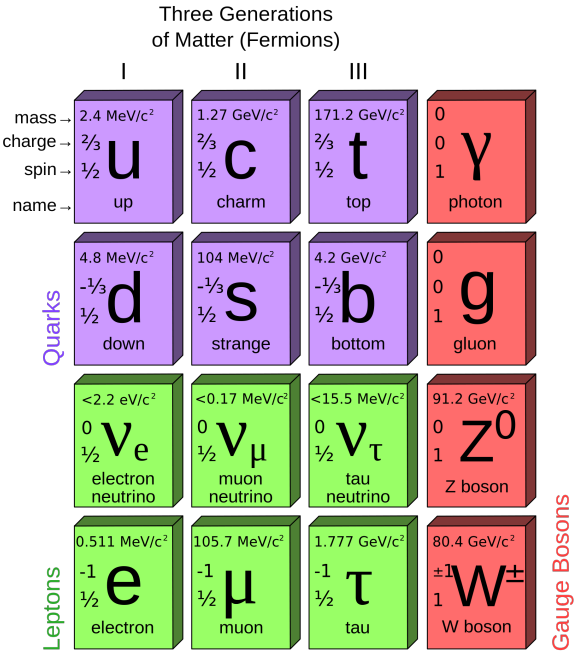


Figure 3: Elementary fermions and gauge mediators of electroweak and strong force in the SM [36].

The electromagnetic force interacts via exchange of a photon, whereas the strong force interacts via exchange of a gluon, and the weak force interacts via exchange of W or Z bosons. The electromagnetic force describes the interaction between electrically charged particles (charged leptons and quarks). The strong force is responsible for holding composite fermions and bosons together. Only quarks interact strongly.

The weak force couples to quarks and leptons, allowing changes of flavours: $\ell^- \to W^- \nu_\ell$ and $d \to W^- u$. Since the quark mass eigenstates are not flavour eigenstates, the weak interaction also allows changes between different quark families (e.g. $b \to W^- u$). The transition amplitude, in the quark sector, is proportional to the CKM matrix element $|V_{ij}|^2$. In the SM these transitions involve a change in the electrical charge. Flavour transitions between quarks of the same electrical charge are only allowed in the SM via loop processes (penguin and W box diagrams, examples are shown in Fig. 4). In the SM the transition involves two flavour changings, one between a family and one inside the family. These processes are named FCNC, because the electrical charge of the initial and final quark is the same. The CKM matrix and the measured magnitudes are as follows [27]:

$$
\begin{aligned}
V_{CKM} &= \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \\
&= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} \qquad (1) \\
&= \begin{pmatrix} 0.97428 \pm 0.00015 & 0.2253 \pm 0.0007 & 0.00347^{+0.00016}_{-0.00012} \\ 0.2252 \pm 0.0007 & 0.97345^{+0.00015}_{-0.00016} & 0.0410^{+0.0011}_{-0.0007} \\ 0.00862^{+0.00026}_{-0.00020} & 0.0403^{+0.0011}_{-0.0007} & 0.999152^{+0.000030}_{-0.000045} \end{pmatrix}
\end{aligned}
$$

where $s_{ij} = \sin\theta_{ij}$, $c_{ij} = \cos\theta_{ij}$ and δ is the phase responsible for all *CP*-violating phenomena in flavour changing processes in the SM. The *CP* transformation combines charge conjugation C with parity conjugation P. Gravitational, electromagnetic and strong forces[2] for example are *CP* conserving. The weak interaction violates C and P, but *CP* is still preserved in most weak interaction processes. However the *CP* symmetry is violated in certain rare processes as in neutral K decays and in B decays. The phase δ is responsible for such violation.

From the CKM matrix it can be seen that transitions between families are suppressed, most strongly transitions involving the third family.

2.2. New Physics search at LHCb

The observed Universe is composed almost entirely of matter and little or no primordial antimatter. The symmetry between particles and antiparticles, firmly established from collider physics, leads to the question why in the Universe this symmetry is broken. Several conditions, called Sakharov conditions, must be fulfilled to explain this asymmetry, one of these conditions is *CP*-violation. However the *CP*-violation present in the SM seems to be insufficient to explain the matter-antimatter asymmetry in our Universe [32, 33].

[2] *CP* conservation in strong interactions is known as the "strong *CP* puzzle", since the most general Lagrangian in strong interaction contains a term that violates *CP*. This term is experimentally compatible with zero.

NP models have in general additional sources of *CP*-violation, which could have played a fundamental role in the early phase of the Universe, generating the over-abundance of matter with respect to anti-matter in today's Universe. Therefore measurements of *CP*-violation are promising probes for NP. In addition, over-constraining the magnitudes and the phases of the CKM elements provides excellent sensitivity to NP. Further measurements of rare decays, which are suppressed in the SM, can probe NP since new heavy particles inside the loops can enter in competitions with the SM contributions.

The search for physics beyond the SM can be conducted with two complementary approaches: direct and indirect. Direct search means that new particles are looked for as real particles, as mainly done by ATLAS and CMS. LHCb, instead, is well suited for indirect searches. There, the experimental observations for specific decays are compared to SM predictions, for example additional contribution from NP can increase branching fractions with respect to SM predictions. In general the following requirements are needed for indirect searches:

– Theoretical errors on the observables need to be small in comparison to the variation among the different models.

– The variation between different theoretical models should be large in comparison to the statistical error for a dataset of realistic size.

Particularly interesting are processes where the SM contribution is suppressed. One of the advantages of indirect searches is that they are sensitive to energy scales of about 200 TeV, which is much higher than the scale for direct searches [23]. With this indirect approach LHCb has already constrained many NP scenarios (see Ref. [7, 6, 4, 11, 5]).

2.3. Decay $B^0 \to K^{*0}\mu^+\mu^-$

The decay $B^0 \to K^{*0}\mu^+\mu^-$ is a FCNC and has several observables that fulfill the previous conditions. Figure 4 shows the corresponding SM lowest order Feynman diagrams. The decay has four charged particles in the final state, where the K^{*0} decays into K^+ and π^-. The two leptons arise from loop diagrams via virtual photons or Z bosons, or via the box diagram.

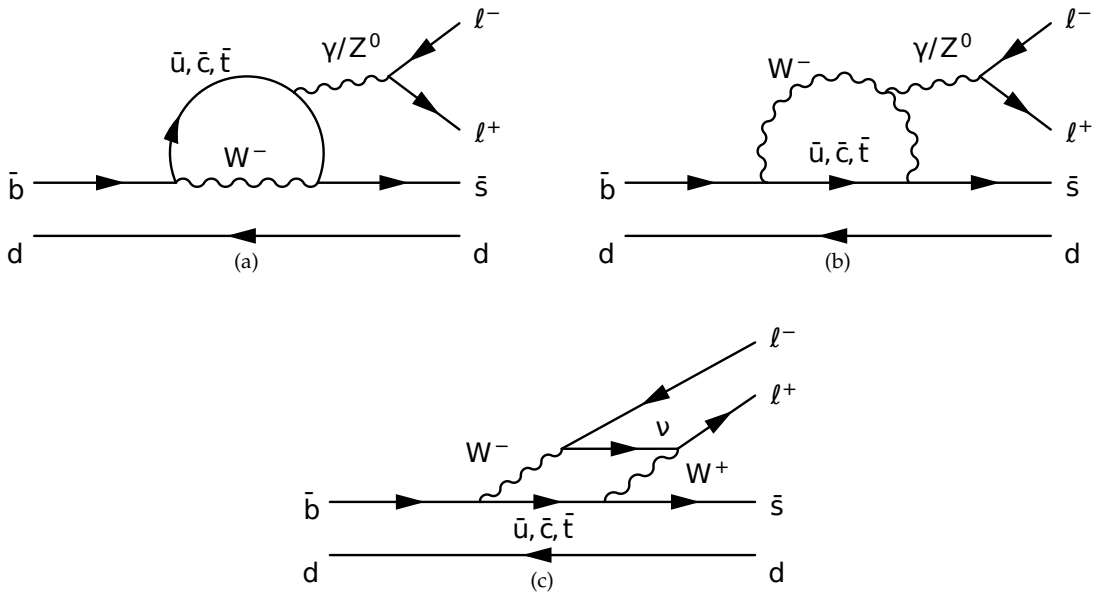


Figure 4: Feynman diagrams for the decay $B_d \to K^{*0}\mu^+\mu^-$ at lowest order in the SM.[25]

2.3.1. Kinematics

The $B^0 \to K^{*0}\mu^+\mu^-$ decay is completely described by the three angles θ_l, θ_k, ϕ which are illustrated in Fig. 5, and the invariant mass squared of the di-muon system, q^2. Let $\mathbf{p}$ denote the momentum vectors in the B^0 rest frame, $\mathbf{q}$ the same in the di-muon rest frame, and $\mathbf{r}$ in the K^{*0} rest frame, where the K^{*0} decays into K^+ π^-. Three unit vectors are defined as follows:

$$\mathbf{e}_z = \frac{\mathbf{p}_{K^+} + \mathbf{p}_{\pi^-}}{|\mathbf{p}_{K^+} + \mathbf{p}_{\pi^-}|}, \quad \mathbf{e}_l = \frac{\mathbf{p}_{\mu^+} \times \mathbf{p}_{\mu^-}}{|\mathbf{p}_{\mu^+} \times \mathbf{p}_{\mu^-}|}, \quad \mathbf{e}_K = \frac{\mathbf{p}_{K^+} \times \mathbf{p}_{\pi^-}}{|\mathbf{p}_{K^+} \times \mathbf{p}_{\pi^-}|} \tag{2}$$

The angles have the following definitions:

$$\cos\theta_l = \frac{\mathbf{q}_{\mu^+} \cdot \mathbf{e}_z}{|\mathbf{q}_{\mu^+}|}, \quad \cos\theta_K = \frac{\mathbf{r}_{K^+} \cdot \mathbf{e}_z}{|\mathbf{r}_{K^+}|} \tag{3}$$

and

$$\sin\phi = (\mathbf{e}_{\mu^+} \times \mathbf{e}_{K^+}) \cdot \mathbf{e}_z, \quad \cos\phi = \mathbf{e}_{K^+} \cdot \mathbf{e}_{\mu^+} \tag{4}$$

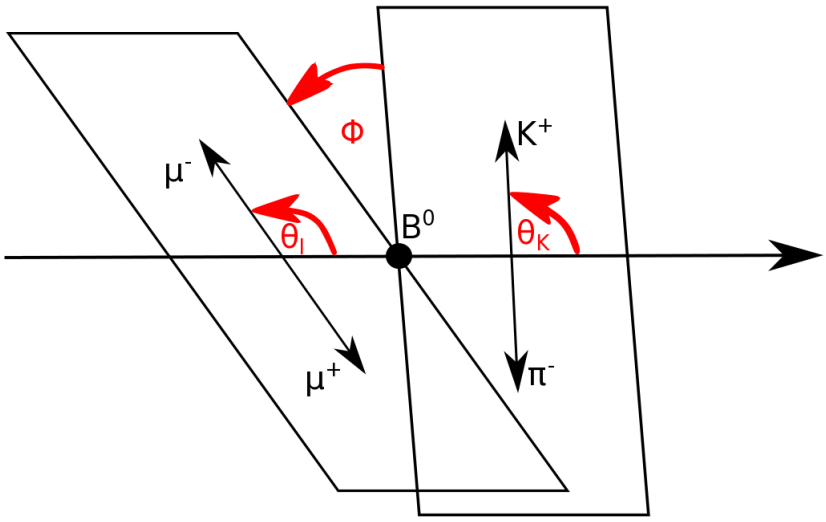


Figure 5: Definition of the kinematic variables in the decay $B^0 \to K^{*0}\mu^+\mu^-$. For details see the text.

The angles are defined in the intervals

$$-1 \leq \cos\theta_l \leq 1, \quad -1 \leq \cos\theta_K \leq 1, \quad -\pi \leq \phi \leq \pi \tag{5}$$

For the $\bar{B}^0$ meson decay all the charges are conjugated. The kinematically allowed region for the dilepton invariant mass squared is

$$4m_\mu^2 \leq q \leq (m_B - m_{K^*})^2. \tag{6}$$

In words: for the B^0, the angle θ_l is measured between the μ^+ and the flight direction of the B^0 in the di-muon rest frame. The angle θ_k is measured between the K^+ and the flight direction of the B^0 in the K^{*0} rest frame. Finally the angle ϕ is measured between the decay plane of $K^{*0} \to K^+ \pi^-$ and the $\mu^+ \mu^-$ plane in the B^0 rest frame [25].

2.3.2. Observables

In this short section the theoretical description of the measured observables is given, for further details see Ref. [17]. As introduced in the previous section, the decay $B^0 \to K^{*0}\mu^+\mu^-$ is completely described by the angles θ_l, θ_K, ϕ and q^2. One obtains the following differential decay rate for the B^0:

$$\frac{\mathrm{d}^4\Gamma}{\mathrm{d}\cos\theta_\ell \, \mathrm{d}\cos\theta_K \, \mathrm{d}\phi \, \mathrm{d}q^2} = \frac{9}{32\pi} I(q^2, \theta_l, \theta_K, \phi) \tag{7}$$

with

$$\begin{aligned} I(q^2, \theta_l, \theta_K, \phi) = {} & I_1^s \sin^2\theta_K + I_1^c \cos^2\theta_K + \left(I_2^s \sin^2\theta_K + I_2^c \cos^2\theta_K\right) \cos 2\theta_l \\ & + I_3 \sin^2\theta_K \sin^2\theta_l \cos 2\phi + I_4 \sin 2\theta_K \sin 2\theta_l \cos\phi \\ & + I_5 \sin 2\theta_K \sin\theta_l \cos\phi \\ & + \left(I_6^s \sin^2\theta_K + I_6^c \cos^2\theta_K\right) \cos\theta_l + I_7 \sin 2\theta_K \sin\theta_l \sin\phi \\ & + I_8 \sin 2\theta_K \sin 2\theta_l \sin\phi + I_9 \sin^2\theta_K \sin^2\theta_l \sin 2\phi, \end{aligned} \tag{8}$$

where the angular coefficients I_i are functions of q^2. The lepton mass is neglected in this expression. For the $\bar{B}^0$ a similar expression involving $\bar{I}_i$ coefficients can be written. This allows us to build *CP*-averaged (S_i) and *CP*-violating (A_i) observables:

$$S_i = (I_i + \bar{I}_i) / \frac{d(\Gamma + \bar{\Gamma})}{dq^2} \tag{9}$$

and

$$A_i = (I_i - \bar{I}_i) / \frac{d(\Gamma + \bar{\Gamma})}{dq^2}. \tag{10}$$

This thesis will concentrate on *CP* averaged quantities. Measuring the corresponding *CP*-asymmetries is experimentally more challenging, for example the detector- and production asymmetry for particles and antiparticles must be determined.
One of the popular observables in the $B^0 \to K^{*0}\mu^+\mu^-$ decay is the **forward-backward asymmetry,** A_{FB}. It is defined as:

$$A_{FB} = \left(\int_0^1 - \int_{-1}^0 \right) d\cos\theta_l \frac{d^2(\Gamma + \bar{\Gamma})}{d\cos\theta_l \, dq^2} / \frac{d(\Gamma + \bar{\Gamma})}{dq^2} = \frac{3}{8}(2S_6^s + S_6^c). \tag{11}$$

In words: A_{FB} can be defined as the difference between the number of events in the forward direction (μ^+ is emitted in flight direction of the B^0 ($\cos\theta_l > 0$)) and the number of events in the backward direction (μ^+ is emitted in the opposite direction of the flight direction of B^0 ($\cos\theta_l < 0$)) normalized over the total number of events. Figure 6 shows the prediction for the SM and different NP models. The asymmetry vanishes at a well defined value of q^2. This zero-crossing point, q_0^2, is sensitive to NP. At leading order the dominant hadronic uncertainties, coming from the form factors, cancels out leading to a small theoretical uncertainty. In the SM, the zero-crossing point is predicted to be: $q_0^2 = 4.97^{+0.03}_{-0.03}|_{FF}{}^{+0.09}_{-0.09}|_{SL}{}^{+0.29}_{-0.27}|_{SD}\,\mathrm{GeV}^2/c^4$[22][3].

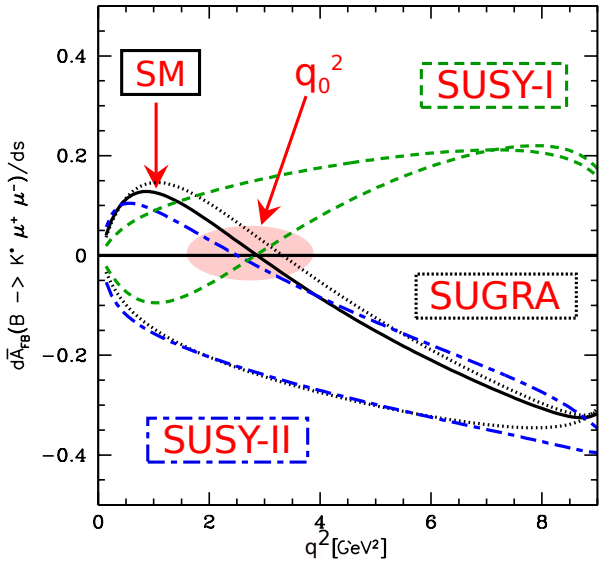


Figure 6: Theoretical expectation for A_{FB} for the SM and different SUSY models, taken from [25]. In the figure are shown the curves for leading order calculation. The zero-crossing point prediction cited in the text corresponds to next-to-leading order calculation.

[3] (FF) are the form factor uncertainties, (SL) are the uncertainties from the Λ/m_b corrections and (SD) are the uncertainty from short distance parameters (m_t, m_W and the μ-scale)

A_{FB} was first measured at Belle, BaBar and CDF. These measurements seemed to hint to a deviation from the SM at low q^2, in particular to a possible flipped C_7 Wilson coefficient with respect to SM predictions. This would have implied no zero-crossing in the A_{FB} [12, 19, 35]. The large increase in statistic at LHCb compared to Belle, BaBar and CDF allows measurements with higher precision, and also measurements of additional observables.

The twelve angular coefficients of Eq. 7 reduce to eight independent coefficients. In addition to A_{FB} and its zero-crossing point, the thesis concentrate on S_4, S_5, S_7, S_8 and on their zero-crossing point [15, 20]. These observables can be measured by simple counting experiments, using the following formulas:

$$S_4 = \frac{\pi}{2}\left(\int_{-\pi/2}^{\pi/2} - \left(\int_{-\pi}^{-\pi/2} + \int_{\pi/2}^{\pi}\right) d\phi\right)\left(\int_0^{\pi/2} - \int_{\pi/2}^{\pi} d\theta_l\right)\left(\int_0^{\pi/2} - \int_{\pi/2}^{\pi} d\theta_K\right) \frac{d^4(\Gamma+\bar{\Gamma})}{d\cos\theta_l\, d\cos\theta_K\, d\phi\, dq^2} \Big/ \frac{d(\Gamma+\bar{\Gamma})}{dq^2} \tag{12}$$

$$S_5 = \frac{4}{3}\left(\int_{-\pi/2}^{\pi/2} - \left(\int_{-\pi}^{-\pi/2} + \int_{\pi/2}^{\pi}\right) d\phi\right)\left(\int_0^{\pi} d\theta_l\right)\left(\int_0^{\pi/2} - \int_{\pi/2}^{\pi} d\theta_K\right) \frac{d^4(\Gamma+\bar{\Gamma})}{d\cos\theta_l\, d\cos\theta_K\, d\phi\, dq^2} \Big/ \frac{d(\Gamma+\bar{\Gamma})}{dq^2} \tag{13}$$

$$S_7 = \frac{4}{3}\left(\int_{-\pi}^{0} - \int_0^{\pi} d\phi\right)\left(\int_0^{\pi} d\theta_l\right)\left(\int_0^{\pi/2} - \int_{\pi/2}^{\pi} d\theta_K\right) \frac{d^4(\Gamma+\bar{\Gamma})}{d\cos\theta_l\, d\cos\theta_K\, d\phi\, dq^2} \Big/ \frac{d(\Gamma+\bar{\Gamma})}{dq^2} \tag{14}$$

$$S_8 = \frac{\pi}{2}\left(\int_{-\pi}^{0} - \int_0^{\pi} d\phi\right)\left(\int_0^{\pi/2} - \int_{\pi/2}^{\pi} d\theta_l\right)\left(\int_0^{\pi/2} - \int_{\pi/2}^{\pi} d\theta_K\right) \frac{d^4(\Gamma+\bar{\Gamma})}{d\cos\theta_l\, d\cos\theta_K\, d\phi\, dq^2} \Big/ \frac{d(\Gamma+\bar{\Gamma})}{dq^2} \tag{15}$$

The SM predictions for the different S_i are shown in Fig. 7. The observables S_4 and S_5 are particularly interesting since they both have a zero-crossing point, where form factor uncertainties cancel out, similar to what happens with A_{FB}. Figure 8 and 9 show the predictions of various NP models for the different S_i and A_i.

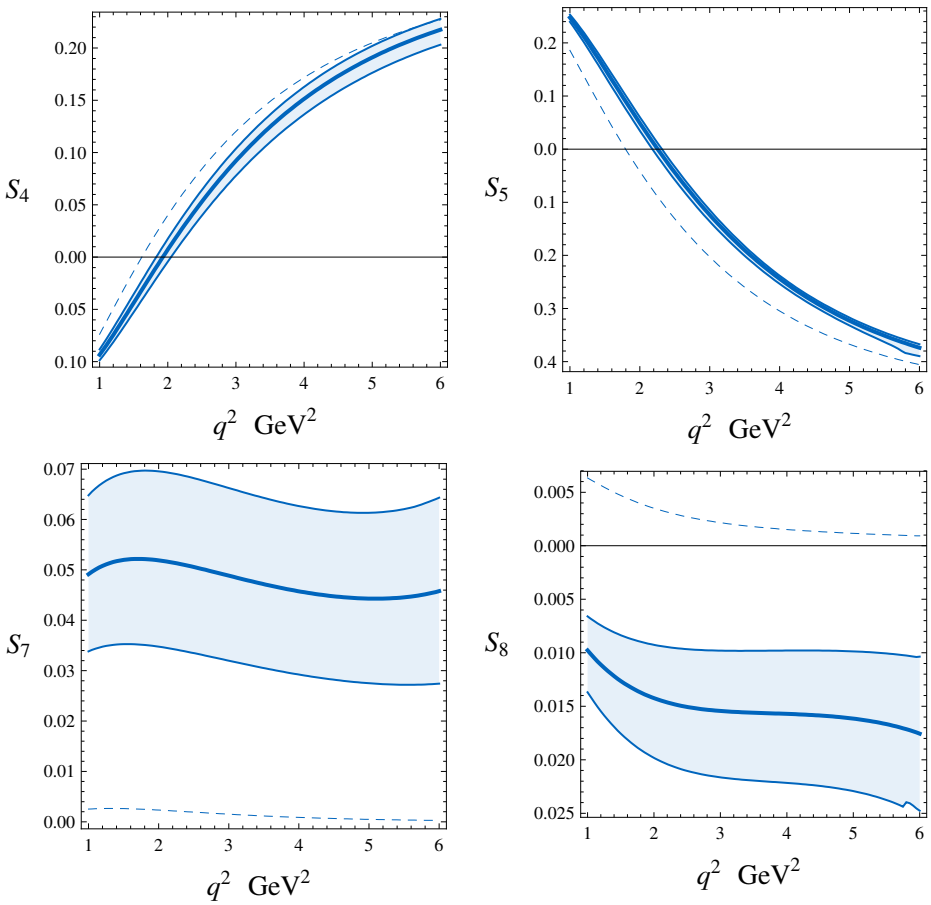


Figure 7: The angular observables S_4, S_5, S_7 and S_8 in the SM as function of q^2. The dashed lines are the leading-order (LO) contributions and the thick solid lines are the next-to-leading order (NLO) predictions. The blue band defines the total error for the NLO results, for details see [17].

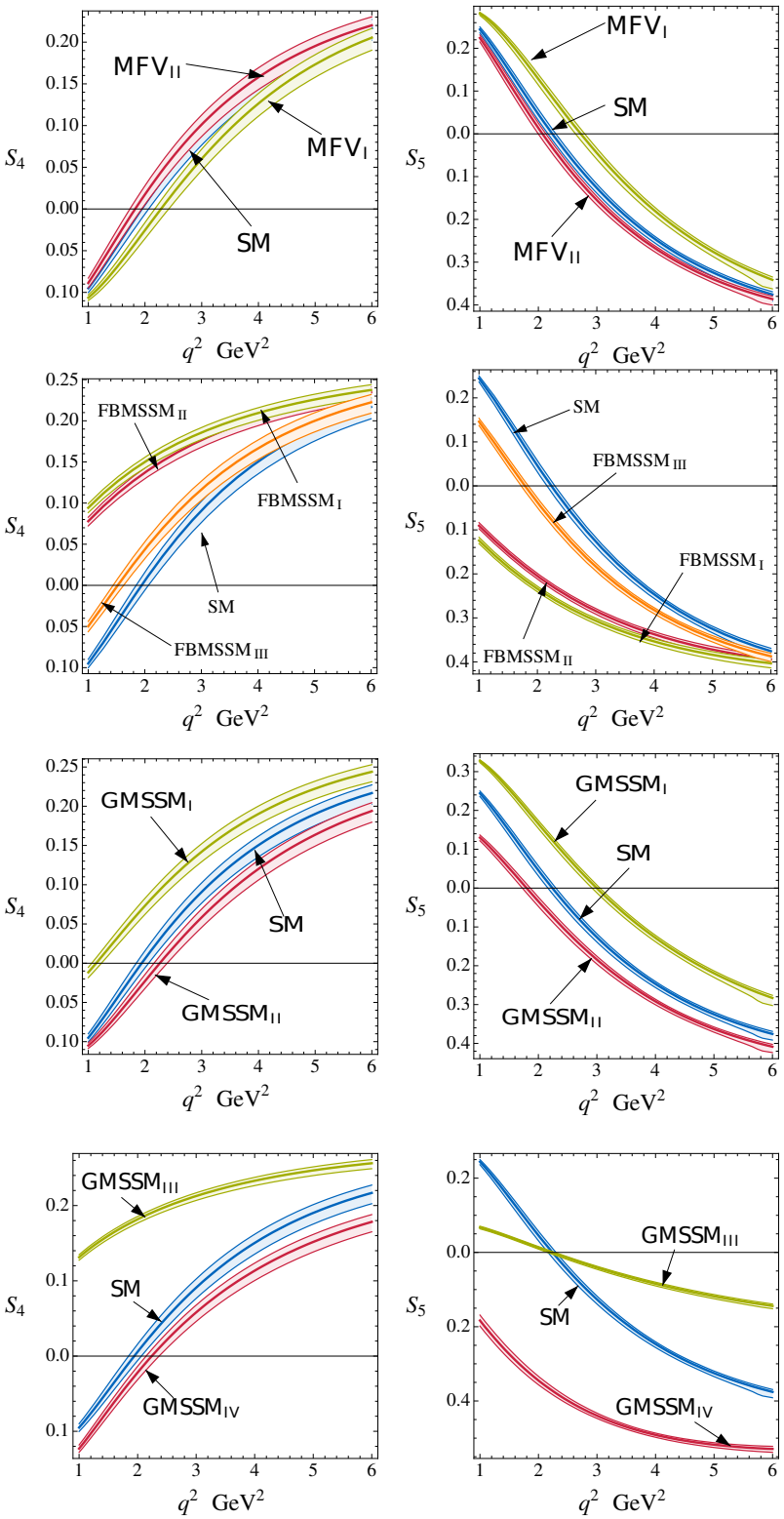


Figure 8: The theoretical prediction for S_4 and S_5, the blue band corresponds to the prediction of S_i in the SM. For more details see [17].

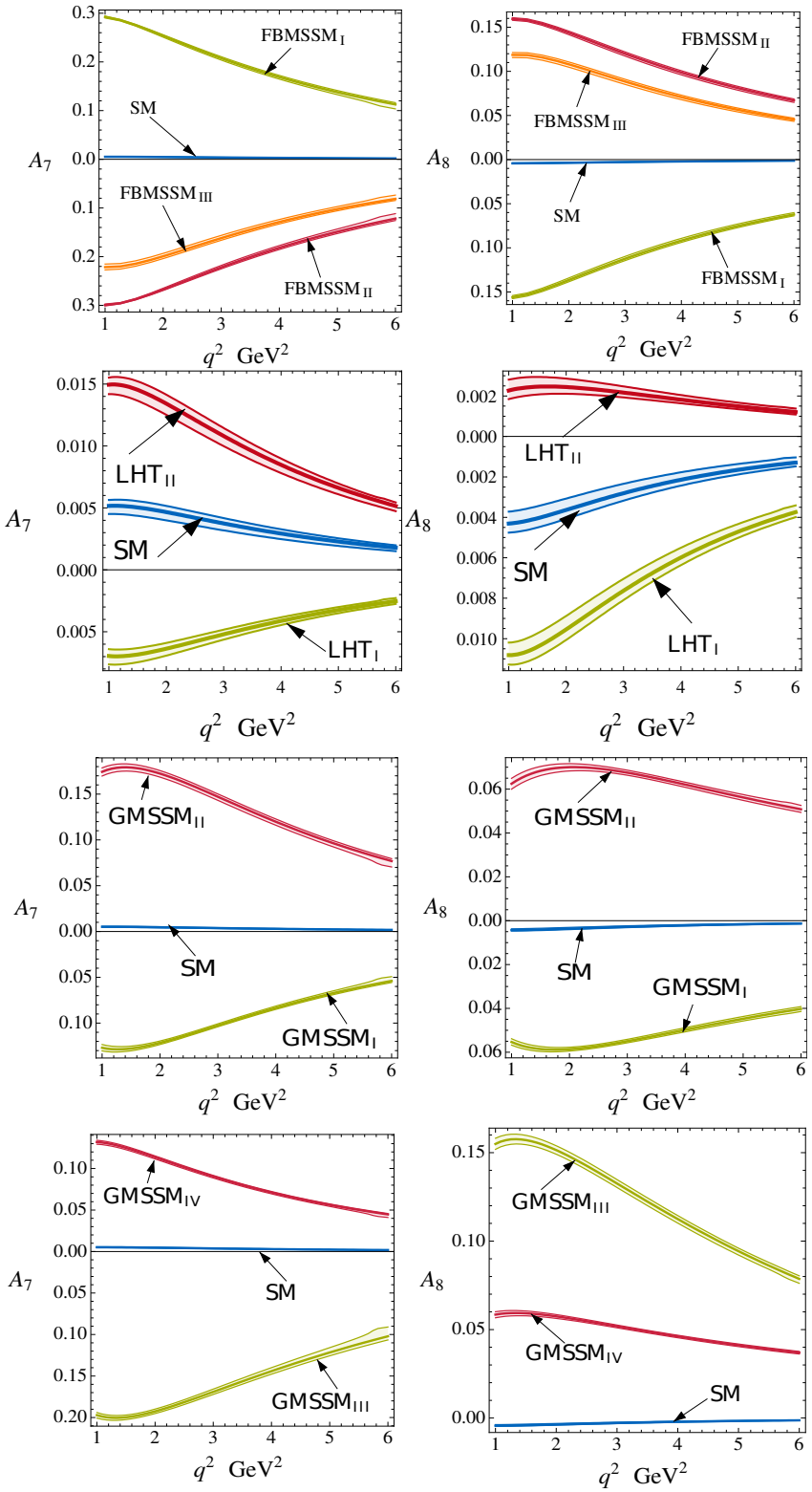


Figure 9: The theoretical prediction for A_7 and A_8, the blue band corresponds to the prediction of A_i in the SM. For more details see [17].

3. The Unbinned Counting Method

Several observables in the decay $B^0 \to K^{*0}\mu^+\mu^-$ are sensitive to NP and theoretically clean. In this chapter the methods for the measurement of A_{FB}, S_i and their the zero-crossing points are presented.

The **full angular fit** of the $B^0 \to K^{*0}\mu^+\mu^-$ decay rate is done in six q^2 bins. To compare the results with earlier measurements from Belle, BaBar and CDF the same binning scheme is used (Ref. [12, 19, 35]). The region 1-6 GeV$^2/c^4$ is favoured by theorists, as it is far away from the photon pole at $q^2 \to 0$ and from the contributions from $c\,\bar{c}$ resonances that would lead to large theoretical errors. In the SM the zero-crossing point of A_{FB} lays in the region 1-6 GeV$^2/c^4$. In Fig. 10 it is shown the result of A_{FB} from the full angular fit. For this result Eq. 7 was reformulated with the transformation, $\phi \to \phi + \pi$ when $\phi < 0$. This cancels out the terms depending on cosϕ and sinϕ, reducing the number of free parameters in the fit. For additional details see Ref. [10].

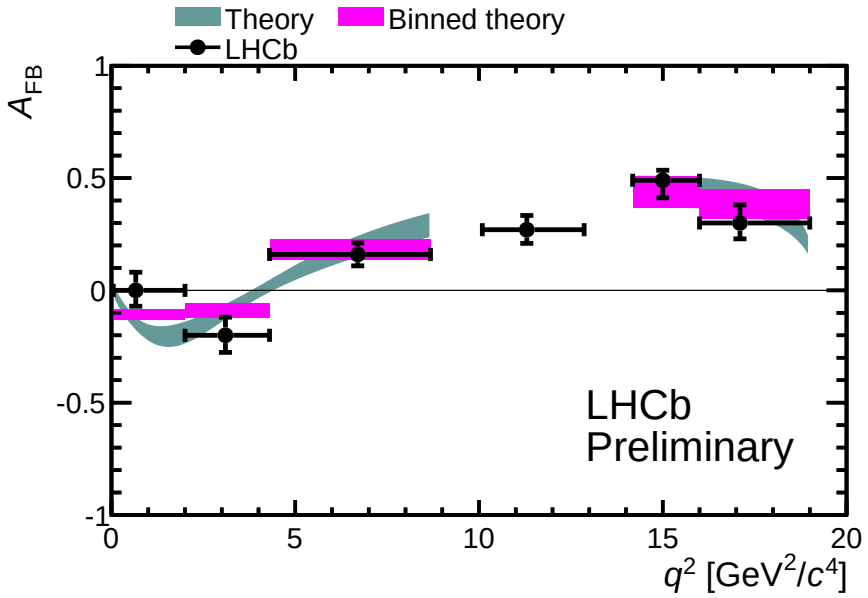


Figure 10: Figure reproduced from Ref. [8], it shows the preliminary results of A_{FB} from the angular fit, as a function of q^2. The data points include both statistical and systematic uncertainties. The theory prediction is described in Ref. [21].

The zero-crossing point cannot simply be extracted by fitting the A_{FB} curve since the functional form of this curve is not exactly known and it depends largely on form factors. Moreover this would be in contrast with the *blinding* procedure used for this analysis. Therefore an alternative method was used: the unbinned counting method.

$B^0 \to K^{*0}\mu^+\mu^-$ events are split up in two categories: forward ($\cos\theta_l > 0$) and backward ($\cos\theta_l < 0$). The difference of these two terms, properly normalized is the A_{FB}. For experimental purpose Eq. 11 can be reformulated as:

$$A_{FB}(q^2) = \frac{\int_0^1 \frac{\partial^2\Gamma}{\partial q^2 \partial \cos\theta_l} d\cos\theta_l - \int_{-1}^0 \frac{\partial^2\Gamma}{\partial q^2 \partial \cos\theta_l} d\cos\theta_l}{\int_0^1 \frac{\partial^2\Gamma}{\partial q^2 \partial \cos\theta_l} d\cos\theta_l + \int_{-1}^0 \frac{\partial^2\Gamma}{\partial q^2 \partial \cos\theta_l} d\cos\theta_l}. \quad (16)$$

An analogous strategy can be adopted for S_i. While for A_{FB} the dataset is split up in the two categories forward and backward, according with Eq. 16, for the S_i the expressions are more complicated, but events can be classified according with their sign in Eq. 12, 13, 14 and 15. The S_i observables become:

$$S_i(q^2) = \frac{Int_P(q^2) - Int_N(q^2)}{Int_P(q^2) + Int_N(q^2)}, \tag{17}$$

where the function $\text{Int}_P(q^2)$ ($\text{Int}_N(q^2)$) is the event distributions for the positive (negative) category as a function of q^2.

Summarizing: The q^2 dependence is fitted separately for the two event categories. The fit is performed in two dimensions: q^2 and invariant mass of the B^0; to separate signal from background. Then the observables S_i and A_{FB} are calculated. This procedure is unbiased and allows to verify the goodness of fit (see Sec. 4) before actually computing the quantity of interest.

3.1. Forward-Backward Asymmetry

The procedure for the measurement of A_{FB} and the S_i is identical. The only difference is in the definition of the positive and negative categories. In the next section the method to extract the zero-crossing point of S_4 and S_5 is described in details. The same method has been applied to the zero-crossing point of A_{FB} (see Ref. [8]) and the results are presented in Sec. 5.

3.2. S_i Asymmetries

Here the application of the unbinned counting method to simulated events[4] is described step-by-step:

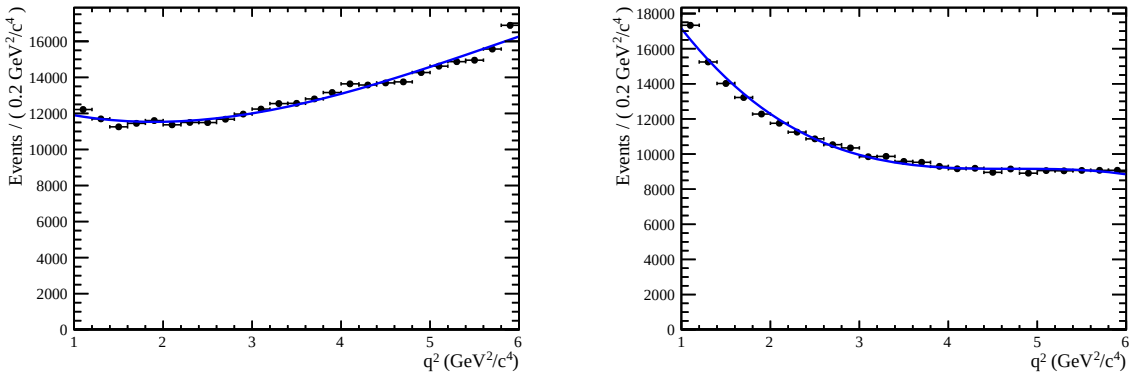


Figure 11: **Left:** Simulated data for the positive category with fit superimposed. **Right:** Simulated data for the negative category with fit superimposed.

[4] LHCb simulation, using the package EvtGen.

1. **Split up:** Events are split up in positive and negative category according to their sign in the S_i expression. For S_5 this corresponds to:

$$\begin{aligned} Int_P(q^2) = \frac{4}{3}\Bigg(&\int_{-\pi/2}^{\pi/2} \mathrm{d}\phi \int_0^{\pi/2} \mathrm{d}\theta_\ell \int_0^{\pi/2} \mathrm{d}\theta_K + \int_{-\pi/2}^{\pi/2} \mathrm{d}\phi \int_{\pi/2}^{\pi} \mathrm{d}\theta_\ell \int_{\pi/2}^{\pi} \mathrm{d}\theta_K \\ &+ \left(\int_{-\pi}^{-\pi/2} + \int_{\pi/2}^{\pi}\right) \mathrm{d}\phi \int_0^{\pi/2} \mathrm{d}\theta_\ell \int_{\pi/2}^{\pi} \mathrm{d}\theta_K \\ &+ \left(\int_{-\pi}^{-\pi/2} + \int_{\pi/2}^{\pi}\right) \mathrm{d}\phi \int_{\pi/2}^{\pi} \mathrm{d}\theta_\ell \int_0^{\pi/2} \mathrm{d}\theta_K \Bigg) \frac{\mathrm{d}^4\Gamma}{\mathrm{d}\cos\theta_\ell \, \mathrm{d}\cos\theta_K \, \mathrm{d}\phi \, \mathrm{d}q^2} \end{aligned} \quad (18)$$

$$\begin{aligned} Int_N(q^2) = \frac{4}{3}\Bigg(&\int_{-\pi/2}^{\pi/2} \mathrm{d}\phi \int_0^{\pi/2} \mathrm{d}\theta_\ell \int_{\pi/2}^{\pi} \mathrm{d}\theta_K + \int_{-\pi/2}^{\pi/2} \mathrm{d}\phi \int_{\pi/2}^{\pi} \mathrm{d}\theta_\ell \int_0^{\pi/2} \mathrm{d}\theta_K \\ &+ \left(\int_{-\pi}^{-\pi/2} + \int_{\pi/2}^{\pi}\right) \mathrm{d}\phi \int_0^{\pi/2} \mathrm{d}\theta_\ell \int_0^{\pi/2} \mathrm{d}\theta_K \\ &+ \left(\int_{-\pi}^{-\pi/2} + \int_{\pi/2}^{\pi}\right) \mathrm{d}\phi \int_{\pi/2}^{\pi} \mathrm{d}\theta_\ell \int_{\pi/2}^{\pi} \mathrm{d}\theta_K \Bigg) \frac{\mathrm{d}^4\Gamma}{\mathrm{d}\cos\theta_\ell \, \mathrm{d}\cos\theta_K \, \mathrm{d}\phi \, \mathrm{d}q^2} \end{aligned} \quad (19)$$

2. **Fit:** The second step is the extended maximum likelihood fit for the q^2 dependence of each data category, positive and negative. The integrals above are performed over the angles, while the two functions Int_P and Int_N sill depend on q^2. These functions are fitted with Chebychev polynomials. It is empirically found that the third order is the lowest that allows to describe the data. This has been verified in Monte Carlo (MC) and data with the goodness of fit test discussed in Sec. 4. The fits are shown in Fig. 11.

3. **Combine:** The last step consists of the combination of the functions Int_P and Int_N. Since the result of the fit is a probability density function (p.d.f.), they need to be properly normalized, according with the formula:

$$S_i = \frac{N_P \cdot PDF_P(q^2) - N_N \cdot PDF_N(q^2)}{N_P \cdot PDF_P(q^2) + N_N \cdot PDF_N(q^2)} \quad (20)$$

where the PDF_P (PDF_N) is the fit p.d.f. for the positive (negative) category and the N_P (N_N) is the corresponding number of events. Fig. 12 shows the result of this procedure which can be compared with the prediction in Fig. 7. S_7 and S_8 are indeed very small and the small bias observed is much smaller than the present or near-future experimental sensitivity.

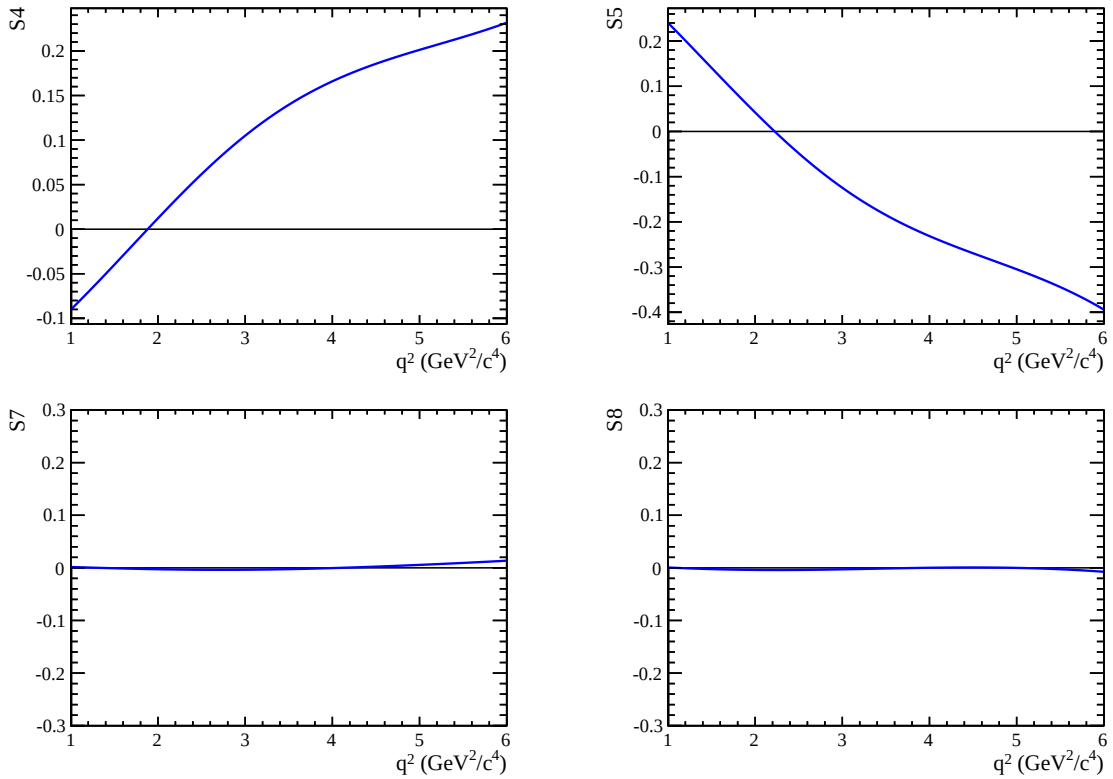


Figure 12: The figures show the S_i curves for simulated events. S_4 and S_5 are in good agreement with SM predictions. S_7 and S_8 wiggle around zero, but the small bias is negligible compared to experimental sensitivity. For more details see the text.

3.2.1. Consistency Test

In this section a consistency check is described. The goal of this test is to check if the method works for smaller samples with a number of events corresponding to what expected for 1 fb^{-1} at LHCb. At LHCb 1 fb^{-1} corresponds roughly to 1000 reconstructed $B^0 \to K^{*0}\mu^+\mu^-$ signal events. In other words the average of the results of several 1 fb^{-1} pseudo-experiments should be compatible with the input curve. This test was performed with and without simulated background.

For this study "toy"-MC samples, generated according with SM expectations are used. The signal mass distribution is generated using a Gaussian p.d.f., while the q^2 distributions (for positive and negative categories) are generated with third order Chebychev polynomials. The background is generated with an exponential p.d.f. for the mass distribution and second order Chebychev polynomials for the q^2 distributions. The coefficients of the polynomials for the signal and background q^2 distributions are extracted form the LHCb simulation, that uses the package EvtGen. The mean and the sigma of the signal mass distribution was fixed using the full LHCb simulation, that uses GEANT4 to simulate the interaction with the detector material. The exponential p.d.f. for the invariant mass distribution for the background is extracted using the sidebands of the B^0 invariant mass, which consist of the region 5150-5220 MeV/c^2 and 5350-5800 MeV/c^2.

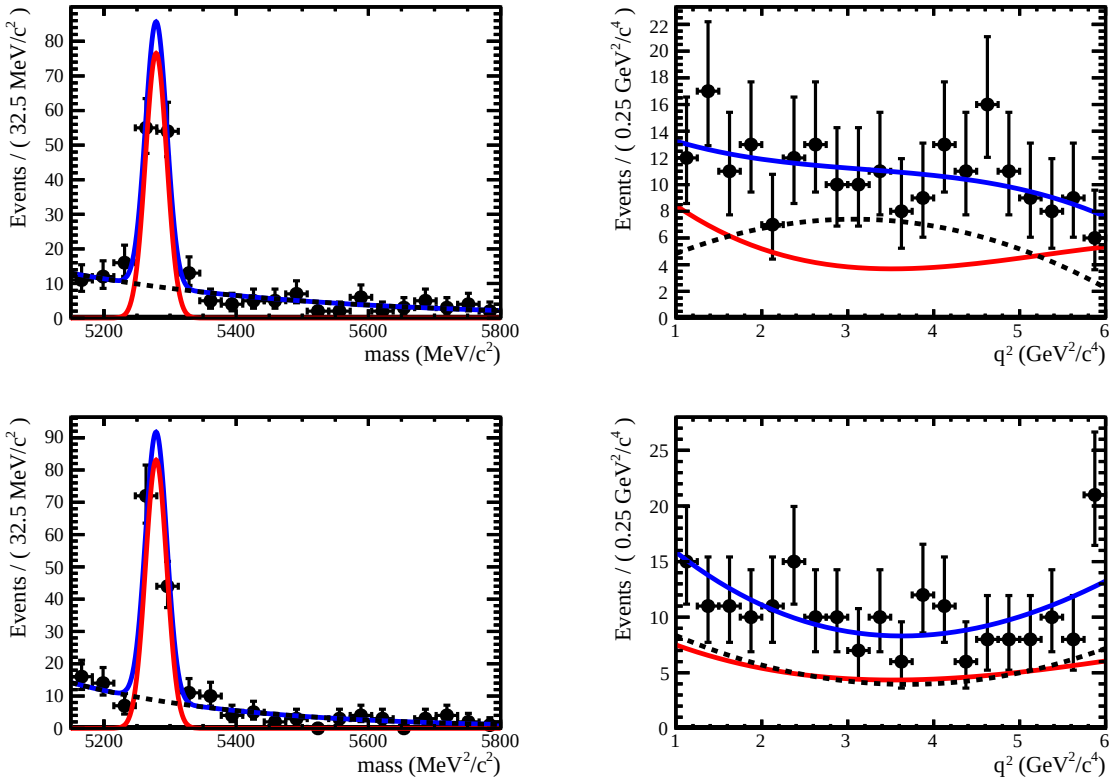


Figure 13: B^0 invariant mass and q^2 distributions for the positive category (top) and for the negative category (bottom). Fit example for one 1 fb^{-1} sample. The blue curve is the total signal and background. The red curve is only signal and the black dotted line is the background only.

The procedure for the consistency test is as follows:

1. 10000 datasets are generated for signal only and including background.

2. The unbinned counting method is applied, fitting the q^2 distributions and the invariant mass. Fig. 13 shows the fit result for one particular pseudo-experiment.

3. The zero-crossing point is calculated in the region 1-6 GeV^2/c^4 (see Sec. 3.2.2).

4. For a given q^2 value an histogram is filled with the measured S_i for the 10000 pseudo experiments (an example for $q^2 = 3\ \mathrm{GeV}^2/c^4$ is shown in Fig. 14). This is repeated for several q^2 values in the range 1-6 GeV^2/c^4. The histograms are fitted with a Gaussian function. The sigma and the mean of these Gaussians are plotted for the different q^2 values in Fig. 16 and Fig. 17 (right). The mean of the Gaussian is compared to the input value in Fig. 16 and Fig. 17 (left).

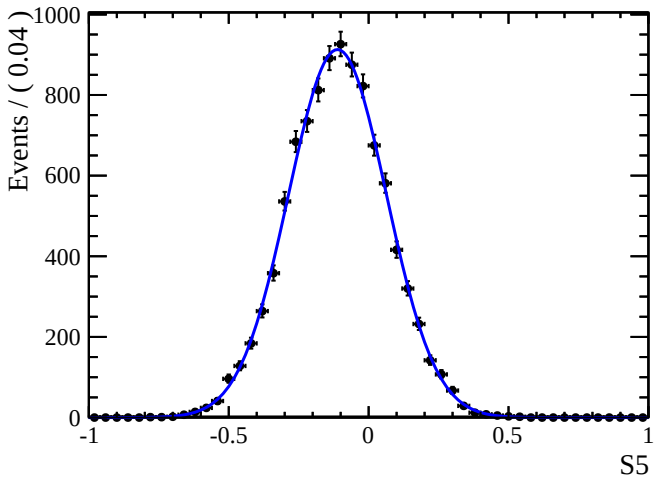


Figure 14: Gaussian fit to the output of 10000 toy experiments for S_5, for $q^2 = 3.0\ \mathrm{GeV}^2/c^4$.

Assuming the SM, the expected experimental resolution will be enough to have a measurement of S_5 not compatible with zero in the q^2 range 1-6 GeV^2/c^4, whereas this is not the case for S_4, S_7 and S_8. In conclusion the method is consistent and unbiased in all cases, with and without background. The sensitivity study of the zero-crossing point of the observable S_4 and S_5 is presented in the next section.

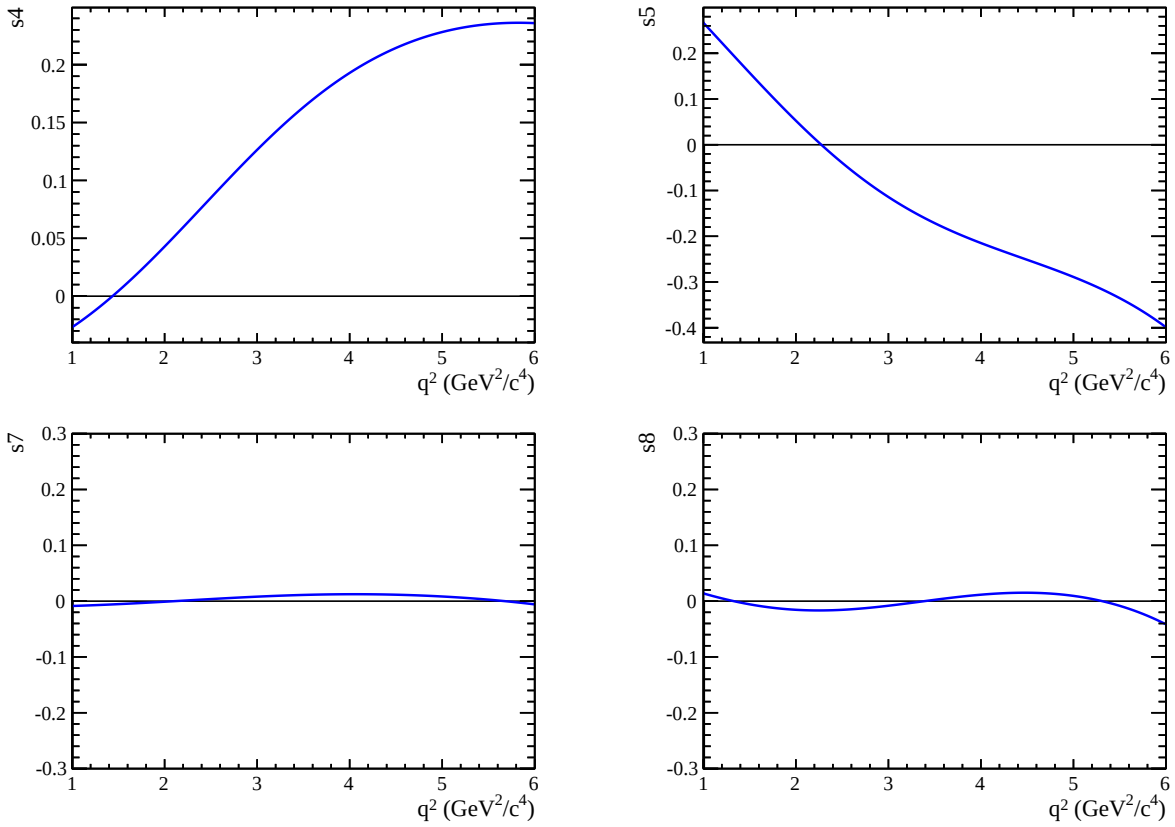


Figure 15: Input curves for S_4, S_5, S_7 and S_8 used for the consistency test.

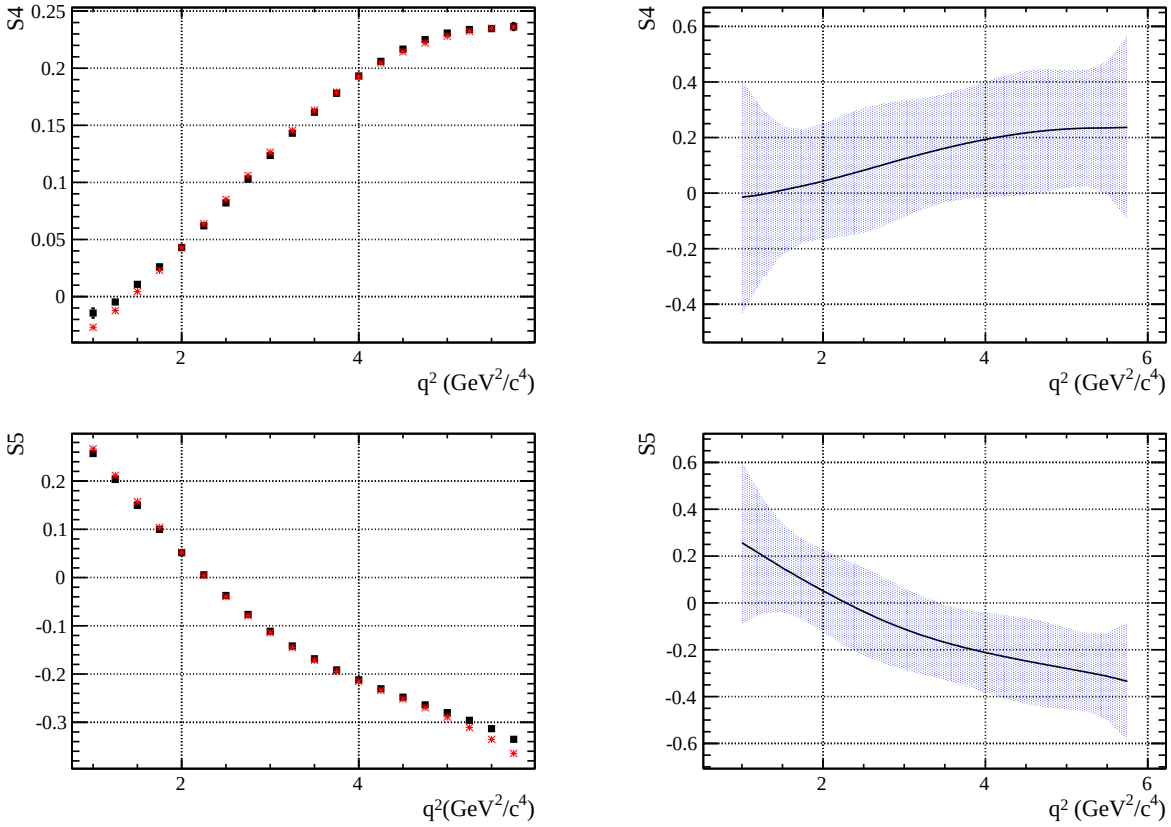


Figure 16: **Left:** Comparison between the mean of 10000 toy experiments and the input value for several q^2 values, for the observables S_4 and S_5. **Right:** The expected 68% confidence level and the most probable value for 10000 toy experiments is shown for S_4 and S_5.
Discussion: From the plots on the left it can be seen how the mean of the 10000 toy experiments is consistent with the input curve. The only small discrepancy can be seen close to the boundaries. From the corresponding plots on the right it can be seen that the fit to the polynomials becomes unstable close to the boundaries. Since the S_4 zero-crossing is close to the lower boundary, we expect difficulties in the extraction of this observable. See next section.

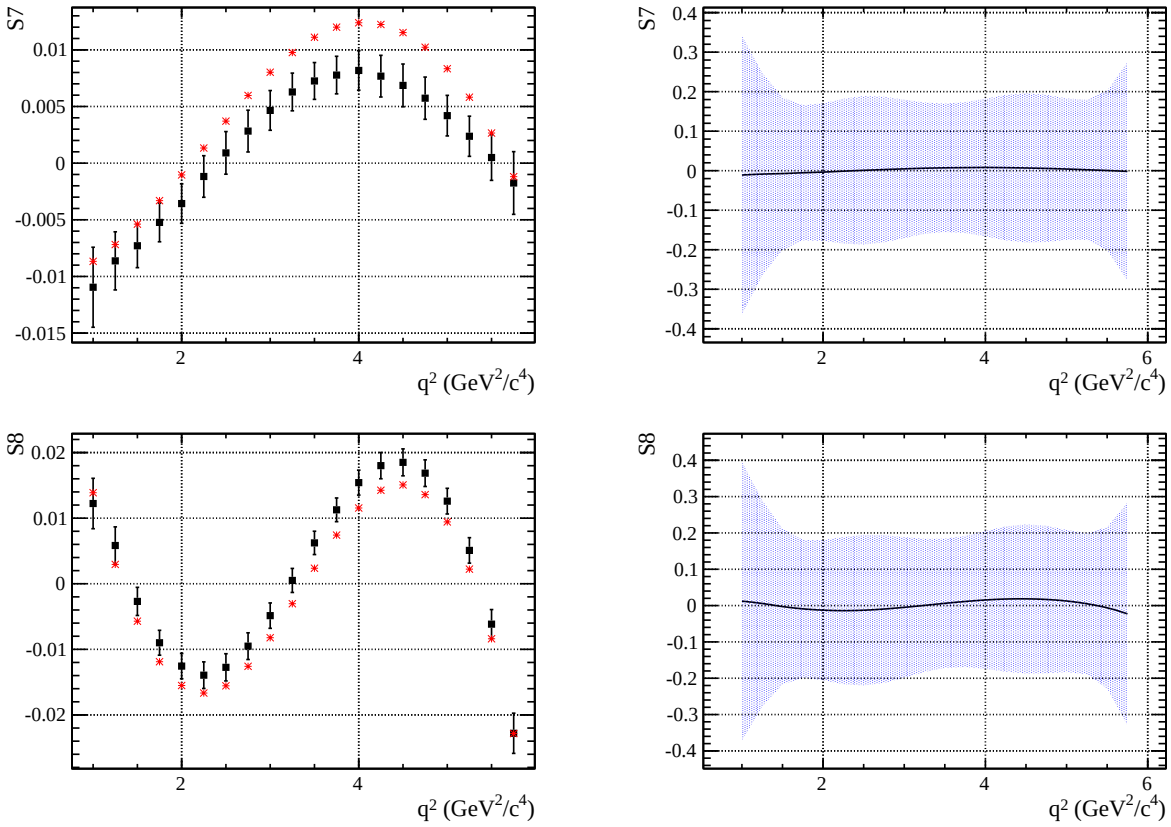


Figure 17: **Left:** Comparison between the mean of 10000 toy experiments and the input value for several q^2 values, for the observables S_7 and S_8. **Right:** The expected 68% confidence level and the most probable value for 10000 toy experiments is shown for S_7 and S_8.
Discussion: The plots on the left show a small bias in the measurement of the observables S_7 and S_8. However as can be seen from the sensitivity plot on the right, this bias is negligibly small compared to the expected experimental resolution.

3.2.2. Zero-Crossing Point Extraction

The zero-crossing point is calculated with the *rootFinder* algorithm, which uses a bisection method (see App. A). The two observable S_4 and S_5, which have a zero-crossing point in the SM, are considered in this section.

In Fig. 15 it can be seen that the zero-crossing points for S_4 and S_5 are close to the low q^2 boundary. From the sensitivity study, as shown in Fig. 16, it can be expected that some of the pseudo-experiments will show no or multiple zero-crossing points. In the case where multiple zero-crossing points are found, the lowest in q^2 is chosen. The additional zero-crossing points, at high q^2, come from fluctuations of the polynomials, as the fit is not stable close to the boundaries.

Based on these issues the *rootFinder* method is built as follows: First it only allows curves which have an odd number of zero-crossing points. Then the whole q^2 range is splitted up in subranges with an odd number of zero-crossing points. For the lowest subrange the bisection method is used to compute the zero-crossing point. This procedure is tested with the same data samples used for the consistency test in Sec. 3.2.1.

Probability of finding one zero-crossing point:

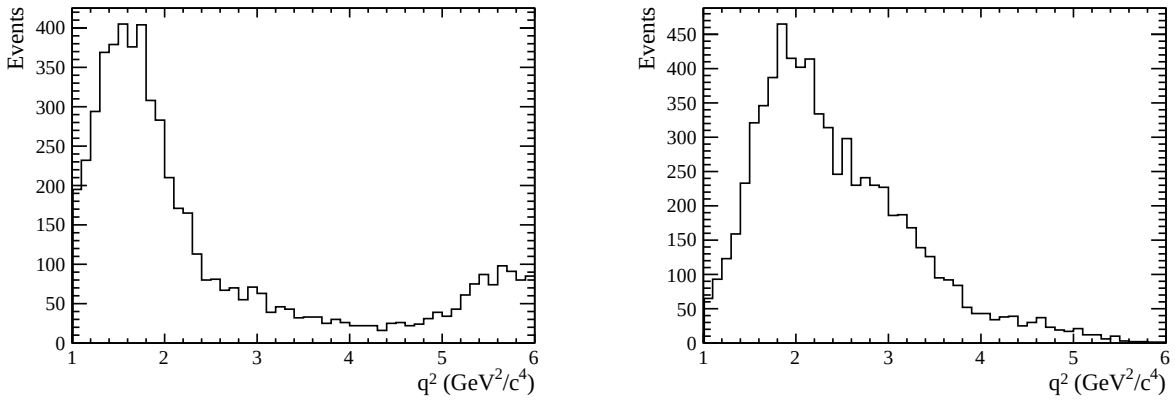


Figure 18: Zero-crossing point distribution for S_4 (left) and S_5 (right), from 10000 toy experiments. For S_4 approximately 5600 toy experiments have only one zero-crossing point and for S_5 approximately 7200 toy experiments have only one zero-crossing point.

For each toy-experiment the zero-crossing point is calculated. The distribution of zero-crossing points for the different toy experiments is shown in Fig. 18. Assuming the SM, the probability to measure only one zero-crossing point is about 72 % for the S_5 observable, for a dataset consisting of 1000 events. For S_4 the probability is 56%. The second peak in Fig. 18 (left) comes from the worse sensitivity of S_4 at the edge of the range. This means, in some cases the S_4-curve has only one zero-crossing at high q^2 values. This was expected from the sensitivity study (see Fig. 16).

This study shows that the zero-crossing point of the asymmetry S_5 can be measured at LHCb with a dataset corresponding to roughly 1 fb^{-1}, while the extraction of the zero-crossing point of the asymmetry S_4 requires larger statistics. The method here described and validated can be applied to data. It is foreseen that the measurement of

the S_i observables and of the zero-crossing point of S_5 will be performed at LHCb in the near future. At time of writing all these observables are still covered by the blinding procedure adopted in the experiment and not yet accessible in the signal region.

Error estimation for the zero-crossing point:
As the zero-crossing point is measured with the bisection method it is complicate to calculate the statistical uncertainty. Therefore the bootstrapping method is used [24]. This method uses a re-sampling technique to estimate the 68% confidence interval. Schematically, first is taken the dataset of N events,

$$d = \{\overrightarrow{\Omega}_0, \overrightarrow{\Omega}_1, ..., \overrightarrow{\Omega}_{N-2}, \overrightarrow{\Omega}_{N-1}\}$$

and then create a new, re-sampled dataset form it of the size, d_1. The total number of events are varied according with a Poisson distribution. For the re-sampling duplicated events are allowed, e.g.:

$$d_1 = \{\overrightarrow{\Omega}_0, \overrightarrow{\Omega}_0, ..., \overrightarrow{\Omega}_{N-2}, \overrightarrow{\Omega}_{N-1}\}$$

where event '0' appears twice and event '1' is omitted from d_1. The unbinned counting method is performed on each re-sampled dataset to extract the zero-crossing point, leading to a distribution of zero-crossing points. The resulting distribution is used to estimate the 68 % confidence interval on q_0^2.

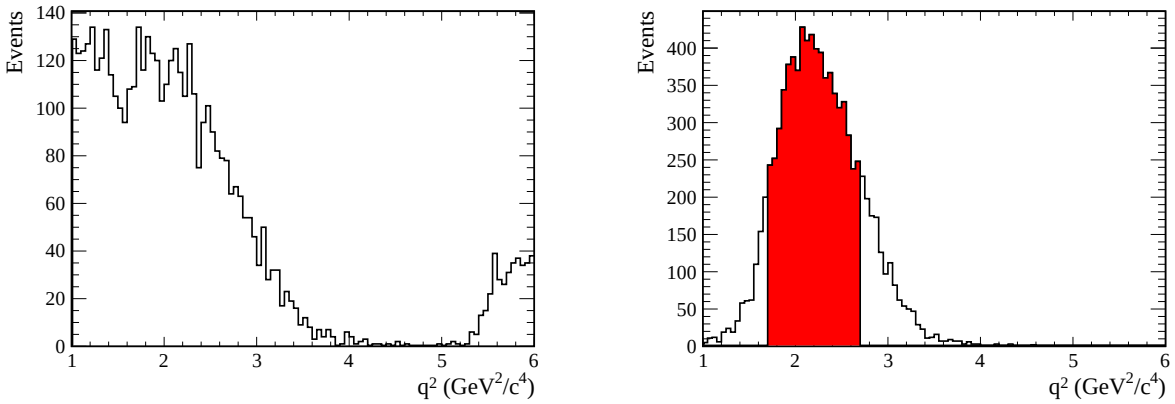


Figure 19: Zero-crossing point distribution for S_4 (left) and S_5 (right) for one particular experiment, using the bootstrapping method. The 68% confidence level is indicated by the red region for S_5.

The result of the bootstrapping for S_4 and S_5 is shown in figure 19 . It could be seen that for S_5 the crossing is well separated from the lower end of the q^2 range. This is not the case for the observable S_4, as can be seen in Fig. 19 left. Moreover the distribution shows a second peak around the q^2 upper bound, due to a bad behavior of the polynomial near boundaries.

4. Goodness of Fit Test

Multidimensional analysis, as used in this thesis, often involve performing an unbinned maximum likelihood fit of a p.d.f. to data. This method is powerful but the maximum likelihood value cannot be used to determine the goodness of fit (g.o.f).

Instead a common practice in physics is to bin the data and compute a χ^2 value. If the bin size is too small the goodness of fit is overestimated with the χ^2 method. If the bin size is too large it would be impossible to compare the finer structure of the fit p.d.f with the data [39]. Furthermore a binning of data always results in a loss of information. Therefore unbinned g.o.f methods are preferred for multidimensional problems.

This chapter first gives an overview on the terminology of g.o.f methods and afterward describes the g.o.f method that was used for the analysis of the S_i and A_{FB}: the point-to-point dissimilarity. Furthermore the method is tested with different fit p.d.f.

4.1. Terminology

Goodness of fit tests are basically hypotheses tests. The H_0 hypothesis is then the case where the fit p.d.f. describes the given data sample and the H_1 hypothesis is the case where the fit p.d.f. does not describe the data sample. The fit p.d.f. is then rejected if the calculated p-value is smaller then the significance level[5] α, which is set in most cases to $\alpha = 0.05$. For this α the hypothesis H_0 is rejected at 95 % confidence level. Two terms are needed for the hypothesis test:

– **Test statistic:** For a g.o.f method a test statistic, T, quantifies the agreement between data and the test p.d.f. For the most common definition a large value of T corresponds to a worse level of agreement, however this is not a universal property. The most familiar test statistic is the χ^2-test.

– **p-value:** For the case where larger T values corresponds to worse level of agreement the p-value is defined as follows:

$$p = \int_T^\infty g_{f_0}(T')dT'. \tag{21}$$

f_0 denotes the true p.d.f. of the data and $g_{f_0}(T')$ is the probability density function for the test statistic under the assumption that f_0 corresponds to the fitted p.d.f. f. Thus the p-value is the probability of finding a T-value corresponding to a worse agreement of f to f_0 then the observed T-value.
For the case where f_0 corresponds exactly to the fit of the data, $g_{f_0}(T')$ is uniform distributed between zero and one.

[5] The significance level indicates Type I error: the probability that a decision to reject the null hypothesis will be made when it is in fact true and should not have been rejected.

It should be mentioned that a p-value greater than the significance level does not imply that the H_0 hypothesis is true, it only states that the fit p.d.f. is not distinguishable from the true p.d.f. with the given amount of data [39, 40].

4.2. Point-to-Point Dissimilarity Method

The electrostatic energy built from a positive and a negative charge distribution is minimal if the two charge distributions coincide. This behaviour can be used to formulate a statistical expression which is minimum if the two statistical distributions agree. If one of the samples is a generated MC sample from a fit p.d.f. and the other sample are the data points, then this method can be used to build a goodness-of-fit test.

4.2.1. Physical Motivation

The potential energy ϕ of a continuous charge distribution $\rho(\mathbf{x})$ is described as follows:

$$\phi = \frac{1}{2}\int d\mathbf{x}\int d\mathbf{y}\frac{\rho(\mathbf{x})\rho(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}. \tag{22}$$

This quantity will be positive unless $\rho(\mathbf{x}) \equiv 0$ with $\phi = 0$, whereas the total charge is fixed to zero

$$\int d\mathbf{x}\rho(\mathbf{x}) = 0. \tag{23}$$

The property $\phi \geq 0$ follows from the property of the distance function $R(|\mathbf{x}-\mathbf{y}|) = 1/|\mathbf{x}-\mathbf{y}|$.

The charge distribution can be split up into two components, a positively charged distribution $\rho_+(\mathbf{x})$ and a negatively charged distribution $-\rho_-(\mathbf{x})$ of opposite total charge as follows:

$$\rho(\mathbf{x}) = \rho_+(\mathbf{x}) - \rho_-(\mathbf{x}), \tag{24}$$

$$\int d\mathbf{x}[\rho_+(\mathbf{x}) - \rho_-(\mathbf{x})] = 0. \tag{25}$$

As mentioned before, if the two charge distributions are equal ($\rho_+(\mathbf{x}) \equiv \rho_-(\mathbf{x})$), then the potential energy is minimal. This property can be used to compare statistical distributions. Therefore the charges are replaced by observations (each observation has the same charge, positive for ρ_+ and negative for ρ_- and Eq. 25 must be fulfilled). After the substitution of ρ from Eq 25 into Eq 22

$$\begin{aligned}\phi &= \tfrac{1}{2}\int d\mathbf{x}\int d\mathbf{y}[\rho_+(\mathbf{x}) - \rho_-(\mathbf{x})] \times [\rho_+(\mathbf{y}) - \rho_-(\mathbf{y})]R(|\mathbf{x}-\mathbf{y}|) \\ &= \tfrac{1}{2}\int d\mathbf{x}\int d\mathbf{y}[\rho_+(\mathbf{x})\rho_+(\mathbf{y}) - 2\rho_-(\mathbf{x})\rho_+(\mathbf{y}) + \rho_-(\mathbf{x})\rho_-(\mathbf{y})]R(|\mathbf{x}-\mathbf{y}|)\end{aligned} \tag{26}$$

and the replacement of the integrals by corresponding sample means, the charge distribution results in:

$$\phi_{NM} = \frac{1}{M(M-1)}\sum_{j>i}R(|\mathbf{y}_i-\mathbf{y}_j|) - \frac{1}{NM}\sum_{i,j}R(|\mathbf{x}_i-\mathbf{y}_j|) + \frac{1}{N(N-1)}\sum_{j>i}R(|\mathbf{x}_i-\mathbf{x}_j|). \tag{27}$$

M and N correspond to the number of observations for ρ_+ and ρ_-. The x_i therefore correspond to particles with a charge of 1/N and the y_i to particles with a charge of -1/M. With this definition Eq. 25 is still fulfilled. The $1/r$ power law is replaced with a more general distance function (more on this later). For large numbers of M and N it follows that the sample mean converge to the integral due to the law of large numbers. For more details, see the Ref. [16, 18, 40] .

4.2.2. Goodness of Fit Implementation

It was mentioned before that goodness of fit tests are hypothesis tests with the following definition:

$$H_0: \quad f_0 \quad = \quad f \tag{28}$$

$$H_1: \quad f_0 \quad \neq \quad f \tag{29}$$

where f_0 is the true p.d.f. of the data and f is the p.d.f. from the fit. H_0 is the null hypothesis and H_1 is the alternative hypothesis. For the test of these hypotheses a test statistic is needed. Eq. 26 is a possible equation for a test statistic. In the case of goodness of fit tests f_0 corresponds to the ρ_- and f corresponds to ρ_+.

$$\begin{aligned} \phi &= \tfrac{1}{2}\int d\mathbf{x}\int d\mathbf{y}[f_0(\mathbf{x}) - f(\mathbf{x})] \times [f_0(\mathbf{y}) - f(\mathbf{y})]R(|\mathbf{x}-\mathbf{y}|) \\ &= \tfrac{1}{2}\int d\mathbf{x}\int d\mathbf{y}[f_0(\mathbf{x})f_0(\mathbf{y}) - 2f_0(\mathbf{x})f(\mathbf{y}) + f(\mathbf{x})f(\mathbf{y})]R(|\mathbf{x}-\mathbf{y}|) \end{aligned} \tag{30}$$

As mentioned before the statistical energy is minimal if the two distributions fully overlap. As the distribution f_0 is unknown for the data points, otherwise no fit is needed, Eq. 26 is useless as a test statistic. However Eq. 27 can be used, as it only needs the observed data (data points) and the fit observation (fit data points). The points from the fit are generated from the fit p.d.f. with a MC technique. The number of measured data points are fixed by the experiment and the number of MC data points can be chosen. For the case where $N \ll M$, Eq. 27 can be simplified:

$$\phi_{NM} = +\frac{1}{N^2}\sum_{j>i} R(|\mathbf{x}_i - \mathbf{x}_j|) - \frac{1}{NM}\sum_{i,j} R(|\mathbf{x}_i - \mathbf{y}_j|). \tag{31}$$

As Eq. 31 shows, the energy term with the generated data points in Eq. 27 cancels out. Under the assumption that N is large, the $N(N-1)$ simplifies to N^2. Equation 31 is the test statistic for the point-to-point dissimilarity method, but two things are still needed. First a reasonable choice of the distance function is needed and second the procedure to calculate the p-value with the test statistic from Eq. 31. This is described in the following sections.

Distance Function: The distance function $R(r)$ is $1/r$ for the electrostatic potential, in this context different functions can be used, with the only condition that they have to decrease monotonically with r.

$$R_{pow}(r) = \frac{1}{r^\kappa + \varepsilon} \tag{32}$$

$$R_{log}(r) = -\ln(r+\varepsilon) \tag{33}$$

$$R_G(r) = e^{-r^2/(2s^2)} \tag{34}$$

The three most commonly used distance functions are the inverse power laws, the logarithmic function and the Gaussian function. The first two have poles at r equals to zero. This could produce infinities in the test statistic. Therefore these poles are suppressed by a small constant cutoff ε. This constant, introduced for numerical reasons, will improve the power of the test because resolving extremely small distances, perhaps smaller then the experimental resolution, could introduce large accidental contributions to the statistical energy. In Ref. [40] is proposed to set

$$\varepsilon = \frac{1}{Mf_{0max}d} \tag{35}$$

where f_{0max} is the maximum of the d dimensional probability density function f_0. Further in Ref. [40] it has been found that the precise value of ε has a negligible effect on the p-value calculation. The Gaussian function has a similar coefficient, s, which is the resolution of the distance function. If s is too small then the fit p.d.f. is approximately constant in the Gaussian region around the events. On the other side if s is to big, the finer structure in the fit p.d.f. is lost [39]. In Ref. [40] it is recommended to use either the logarithmic or the Gaussian distance function. The logarithmic function is long range, the p-values are invariant under scaling ($r \to ar$) and it does not contain a parameter which has to be adjusted. On the other side the Gaussian function has a limited range for the correlation between the data values. Furthermore the parameter s in Eq. 34 should be adjusted to the density of observations and to the shape of f_0. In section 4.3 a comparison of these two distance functions for a specific p.d.f. is shown [16, 18, 40].

p-Value: Now the test statistic T can be calculated, but the distribution of T for the case $f_0 = f$ is not known. This means that the p-value cannot be calculated directly. For an estimation of the p-value a re-sampling method known as the *permutation test* can be used. This test involves the following steps:

1. Combine data and MC data (from the fit p.d.f) into a pooled sample of size $n = n_d + n_{MC}$.
2. Randomly select n_d events from the pooled sample and label these "data" (duplicates are not allowed).
3. Label the remaining events with "MC".
4. calculate the test statistic value for this data; MC combination and name it, T_{perm}.
5. repeat steps 1 to 4 until the number of repetition reach n_{perm}, which is maximum $n!/(n_d!n_{MC}!)$.

This results then in a set of T-values $\{T^1_{perm}, ..., T^{n_{perm}}_{perm}\}$, where the p value is simply the fraction of times for which $T < T_{perm}$ is valid. To avoid large processing time, a random subset of combinations may be used and not all permutations.

This technique works because, if the test p.d.f f and the parent p.d.f f_0 are the same, the assignment "data" and "MC" are just labels. This means reassigning these should

have no effect on the mean of T. Furthermore each of the $n!/(n_d!n_{MC}!)$ events combinations could be observed from an experiment. Thus, the technique can be used to estimate the p.d.f. of T and, in turn, obtain an estimate for the p-value [39].

Additional Issues: In physics analysis events are in many cases weighted, for example to correct for the detector acceptance. As the data points correspond to particles with charges 1/N and the MC data points to particles with charges 1/M, the weights can be used to rescale the charges for the data points:

$$\frac{1}{N_w} \cdot w_1 + ... + \frac{1}{N_w} \cdot w_N = \frac{1}{N_w} \cdot (w_1 + ... + w_N) = 1 \tag{36}$$

where $(w_1 + ... + w_N) = N_w$ by definition. N_w is the number of events after weighting and the w_i are the weights.

The second issues is, that in multidimensional analyses not all dimensions describe the same quantities. In Ref.[39] it is proposed to do a rescaling of each variable x_i lest one dimension dominate the test function. In Ref. [40] it is proposed to scale each x_i relative to its mean value $\bar{x_i}$ by the inverse of the square root of the variance σ_i^2:

$$x_i \to x_i' = \frac{x_i - \bar{x_i}}{\sigma_i}. \tag{37}$$

This rescaling method however has a simple problem. Not each dimension variable x_i has a Gaussian dependence. This rescaling influence then the dependence of the data on each x_i. In general a simple rescaling is used:

$$x_i \to x_i' = \frac{x_i - Min(x_i)}{Max(x_i) - Min(x_i)}. \tag{38}$$

This method rescales the original dependence for each x_i to the range 0 to 1.

4.3. Measurement

In this section the point-to-point dissimilarity method is tested and applied to data. For the experimental use it is necessary to know how well the method performs. In Ref. [39] the *point-to-point dissimilarity method* is compared to other g.o.f. methods, for example the χ^2 method. From the comparison in Ref. [39] it follows that the dissimilarity method has the best behaviour, also for small sample sizes. Based on this paper [39] the section focus on the fit model of the S_i measurement and studies at the influence of different Chebychev polynomials on the g.o.f. test.

4.3.1. Fit Models

For the goodness of fit test two p.d.f.s are needed. One is the parent distribution of the data and the other is the fit distribution. To test the method similar p.d.f.s as used for the toy-study of Sec. 3 are used.

1. **Parent Distribution:** The p.d.f consists of a signal and a background components. For the signal a Gaussian distribution in mass and a third order Chebychev polynomial in q^2 is used, with the same mass and q^2 ranges as in the measurements of the S_i. For the background an exponential distribution in mass and a third order Chebychev polynomial in q^2 is used.

2. **first order Chebychev:** This p.d.f. consists, in contrast to the parent distribution, of a first order Chebychev polynomial as signal p.d.f. in q^2.

3. **second order Chebychev:** This p.d.f. consists, in contrast to the parent distribution, of a second order Chebychev polynomial as signal p.d.f. in q^2.

4. **fourth order Chebychev:** This p.d.f. consists, in contrast to the parent distribution, of a fourth order Chebychev polynomial as signal p.d.f. in q^2.

5. **fifth order Chebychev:** This p.d.f. consists, in contrast to the parent distribution, of a fifth order Chebychev polynomial as signal p.d.f. in q^2.

6. **Case 2-5 extended:** In addition to the four previous cases, the q^2 distribution for the background is described by a Chebychev polynomial with the same order as the signal p.d.f in q^2, instead of a third order Chebychev polynomial as in the parent distribution.

For the test **1000 data samples** each with 500 events are generated from the parent distribution.

4.3.2. p-Value Distributions

The p-value distribution should be flat for the case where fit and parent distribution are the same. If this is not the case, the p-values should all have small values. The procedure to generate the p-value distribution is to fit the generated 1000 data samples from the parent p.d.f. with the fit p.d.f. Then use the fitted p.d.f. to generate one MC data sample and calculate the p-Value for each of the 1000 data samples. These 1000

p-values give then the p-value distribution. The whole procedure is repeated for each different fit p.d.f. (cases 1-6).

For a numerical comparison of the p-value distribution the "power" of the g.o.f. test is calculated. It is defined as the percentage of p-values which are lower than significance level 0.05. In case the fit p.d.f. is equal to the parent p.d.f., the power of the g.o.f. should be 5%. For all other cases the power should be higher.

Distance Function logarithmic: First the logarithmic function was tested since the coefficient ϵ is clearly defined. From Fig. 20 it can be seen that f_{0max} = 95. The dimension of the fit is d = 2 and the number of MC data points is M = 10000 which gives $\epsilon = 5{\cdot}10^{-7}$.

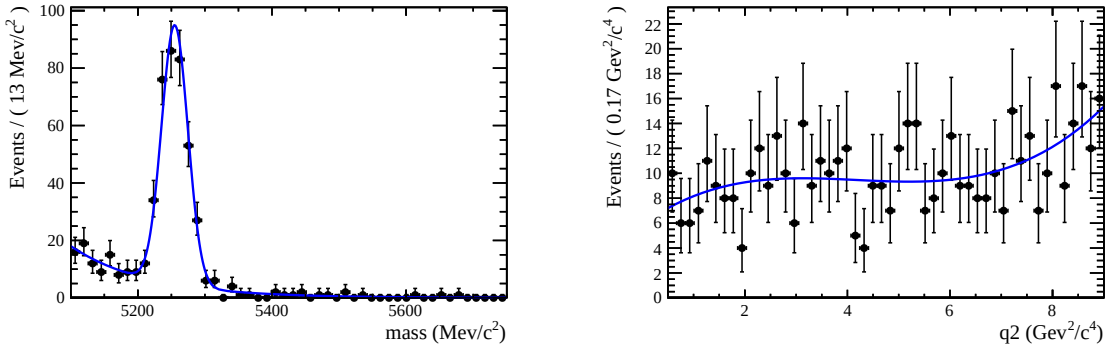


Figure 20: Invariant mass (left) and q^2 distribution (right) for the parent distribution.

In figure 21, on the left, it is shown the distribution for the case when the fit p.d.f. is the same as the parent p.d.f. with fixed parameter values. It is clearly flat as expected for f = f_0. Figure 21 on the right side shows the p-value distribution but when the parameters of the fit p.d.f are not fixed. The distribution is clearly not flat and has a bias to higher p-values. This is an indication, that the logarithmic distance function is unable to reproduce the fast fluctuations in data.

As conclusion it follows that the logarithmic distance function is not the right choice for the fit distribution used in this thesis. A more promising distance function is the Gaussian, as it has a limited range for the correlation and can be adjusted to faster fluctuations.

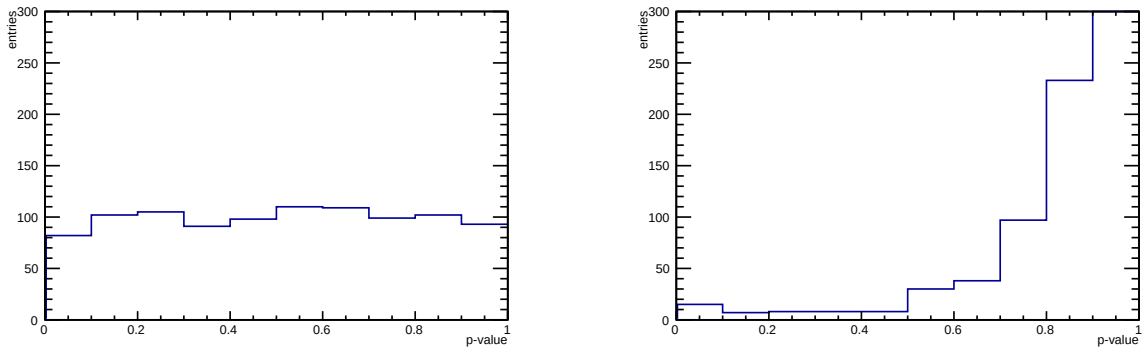


Figure 21: **Left:** p-value distribution for the case when the fit p.d.f. is fixed to the parameter values of the the parent p.d.f. The p-values are calculated with the logarithmic distance function. The distribution is flat as expected. **Right:** p-value distribution where the fit p.d.f. is fitted to the data and the parameters are not fixed. The p-values are calculated with the logarithmic distance function. The figure shows a not flat distribution.

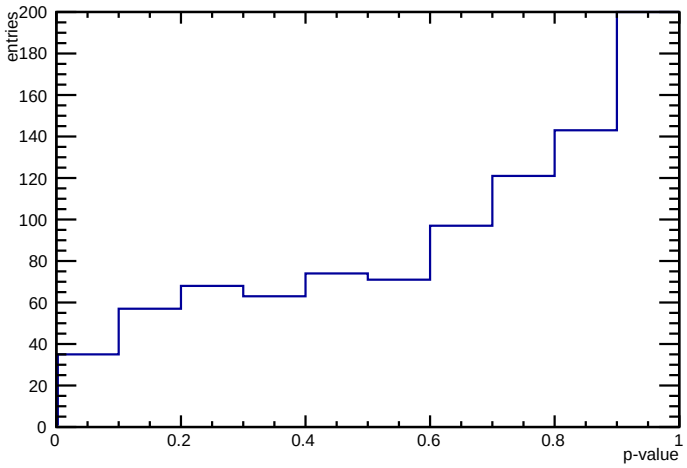
Distance Function Gaussian: As for the logarithmic distance, also the Gaussian distance has one free parameter. In contrast to the logarithmic case, it is not a cutoff to avoid the pole, but it defines the width of the Gaussian. For simplicity the coefficient s is constant over all x_i.

As shown in Ref. [40] for large values of s the finer structure of the fit p.d.f. is lost. Therefore the Gaussian distance function can have the same problem as the logarithmic distance function, the p-value distribution can have a bias to higher p-values. If the parameter s is too small, as shown in Ref. [39], the fit p.d.f. will be constant in the Gaussian region around each data point. Therefore the p-value distribution will always be flat.

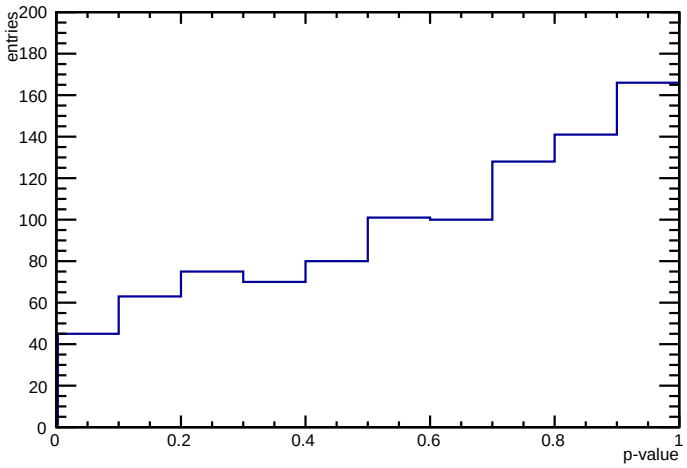
To test this behaviour, p-value distributions for a wide range of s are produced. From Ref. [39] it follows that the most promising s-values for sample with 500 events have small almost vanishing bias to higher p-values. According with this conditions, the best p-value was found to be s = 0.0001. The value distribution for the best s- and for two other examples is shown in Fig. 22 and in Fig. 23. Table 1 shows the "power" for each case.

	s= 0.001	s = 0.0001	s = 0.0005
parent p.d.f.	1.4	3.3	1.5
1st Chebychev signal+background	14.6	7.9	10.5
1st Chebychev signal	15.9	7.4	11.7
2nd Chebychev signal+background	3.5	3.7	3.8
2nd Chebychev signal	5.5	4.3	4.2
4th Chebychev signal+background	2.7	3.9	2.8
4th Chebychev signal	0.7	2.1	1.3
5th Chebychev signal+background	8.1	4.1	7.4
5th Chebychev signal	4.8	3.1	3.5

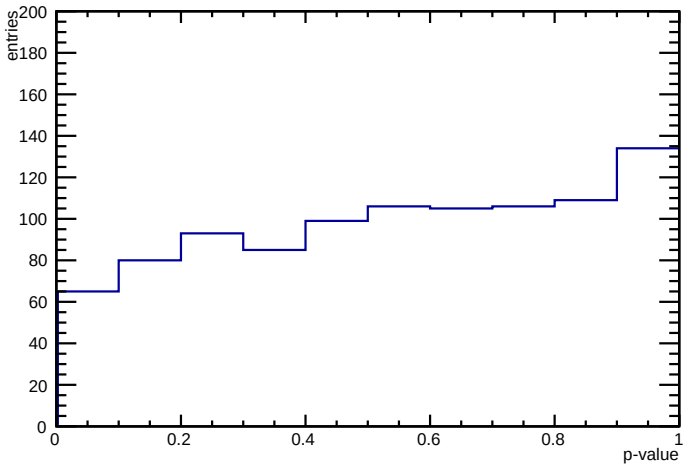
Table 1: The power of the goodness of fit test for the different cases of fit p.d.f.



(a) $s = 0.001$

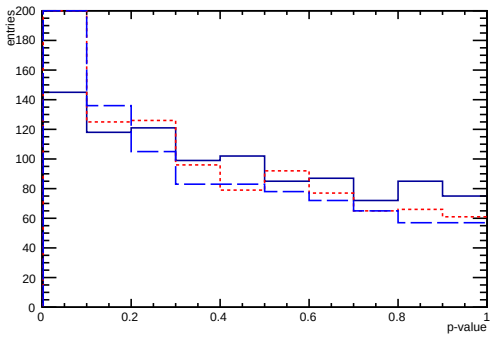


(b) $s = 0.0005$

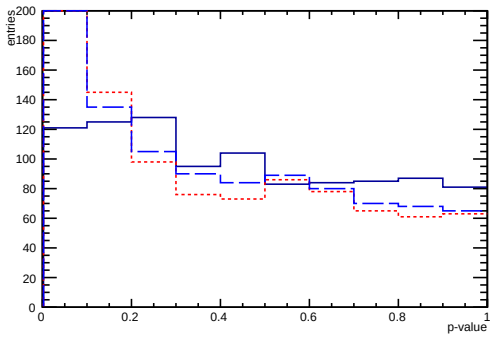


(c) $s = 0.0001$

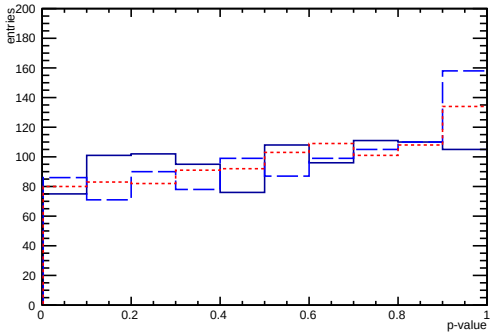
Figure 22: p-value distribution for the case where parent p.d.f. is fitted to data samples.



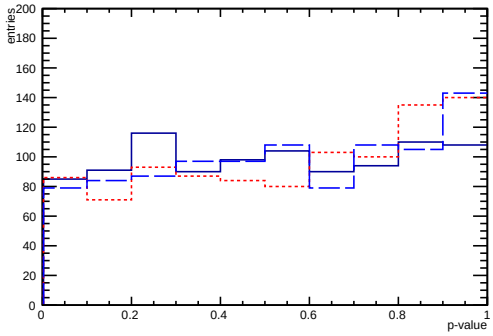
(a) 1st order Chebychev polynomial



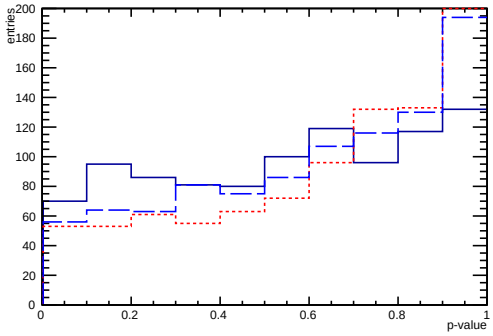
(b) 1st order Chebychev polynomial



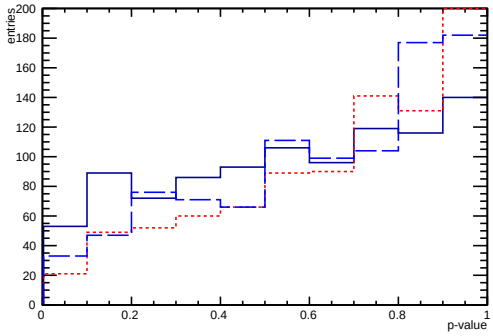
(c) 2nd order Chebychev polynomial



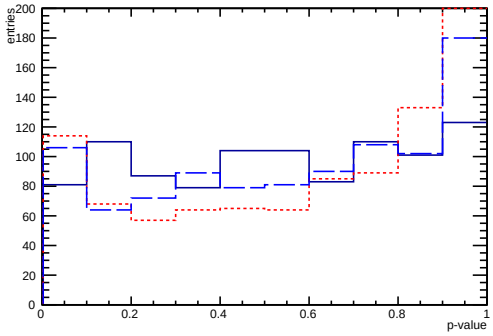
(d) 2nd order Chebychev polynomial



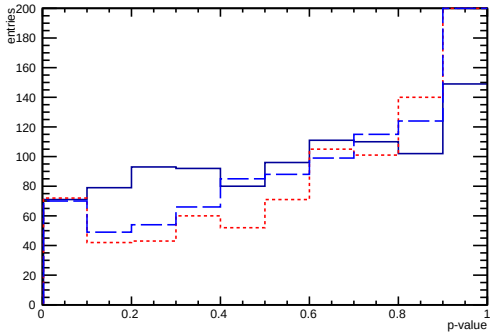
(e) 4th order Chebychev polynomial



(f) 4th order Chebychev polynomial



(g) 5th order Chebychev polynomial



(h) 5th order Chebychev polynomial

Figure 23: p-value distribution for different fit p.d.f. fitted to data samples. The left figures show the distribution when both signal and background have a different Chebychev polynomial then the parent p.d.f. The right figures show the distribution when the signal only has a different Chebychev polynomial then the parent p.d.f. The blue straight line corresponds to $s = 0.0001$, the red dotted line corresponds to $s = 0.001$ and blue dotted is $s = 0.0005$.

Discussion: From the p-value distributions in Fig. 22 it follows that $s = 0.0001$ is the best choice, because it has the smallest bias. As mentioned before s could be too small and therefore the p-value distribution would be always flat. The p-value distributions in Fig. 23 shows that this is not the case for $s = 0.0001$. Furthermore it has the smallest bias to higher p-values. This bias comes from the behaviour of the Chebychev polynomials. Higher order polynomials will fit the small fluctuations better and therefore the s parameter is too large to see the finer structure. Thus the p-value distribution has a bias to higher p-values. This bias vanish for Chebychev polynomials with one order lower than the true polynomial. For a difference of two order of the Chebychev polynomials the p-values peaks to smaller values.

Testing the invariant mass distribution This procedure was used to look for the behaviour of the p-value distribution for two cases of peaking background: first an additional small peak in the mass distribution of $B^0 \to K^{*0}\mu^+\mu^-$ and second additional background contributions in the $B^0 \to J/\psi K^{*0}$ mass. In the first case the additional peak influences directly the $B^0 \to K^{*0}\mu^+\mu^-$ mass fit and in the second case additional background contributions in the $B^0 \to J/\psi K^{*0}$ mass fit influence the fixed mean of the B^0 mass peak. Figure 24 shows p-value distribution for $s = 0.0001$ of the first case of peaking background.

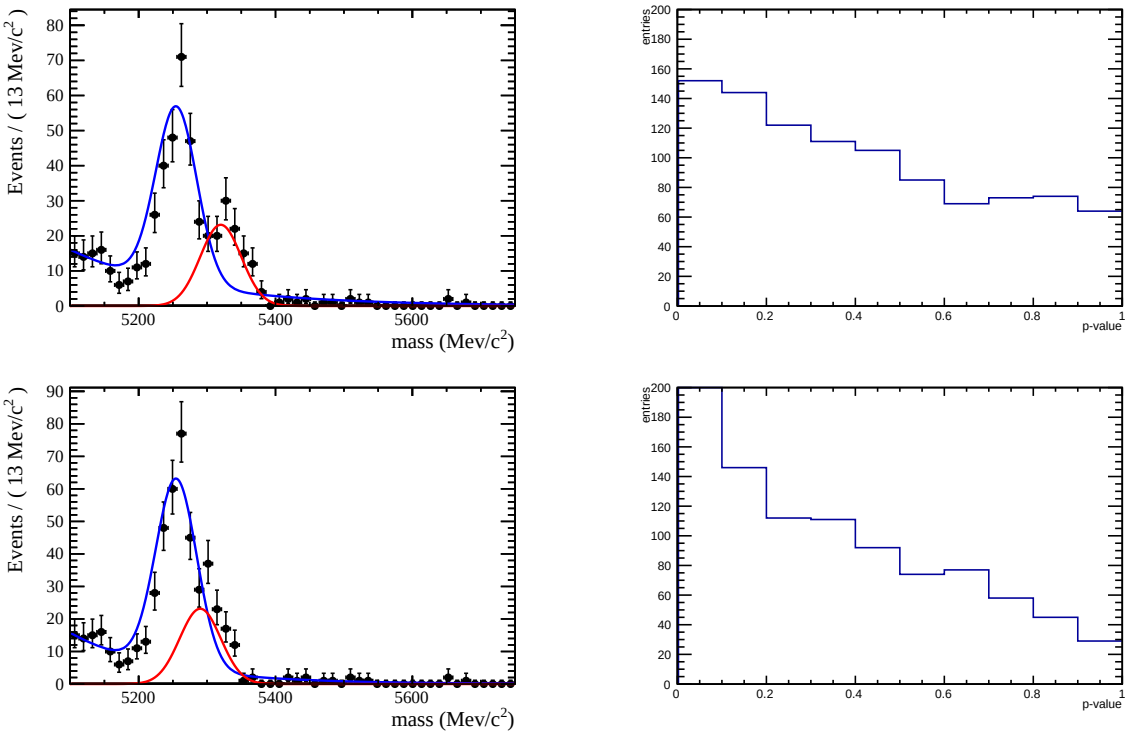


Figure 24: **Left:** The B^0 invariant mass distribution is shown for a toy experiment, when an additional peaking component, modeled with a Gaussian function with mean 5320 MeV/c^2 (top), 5290 MeV/c^2 (bottom) and sigma 30 MeV/c^2, is included. The corresponding p-value distribution is shown on the right.

Figure 25 shows the p-value distribution for different fixed mean values of the fit p.d.f. From shifts of 5 MeV/c^2 on the p-value distribution changes. The p-value distribution go to lower p-values and for a difference of 20 GeV^2/c^4 (width of the peak =

20 GeV$^2/c^4$) all p-values are below 0.05.

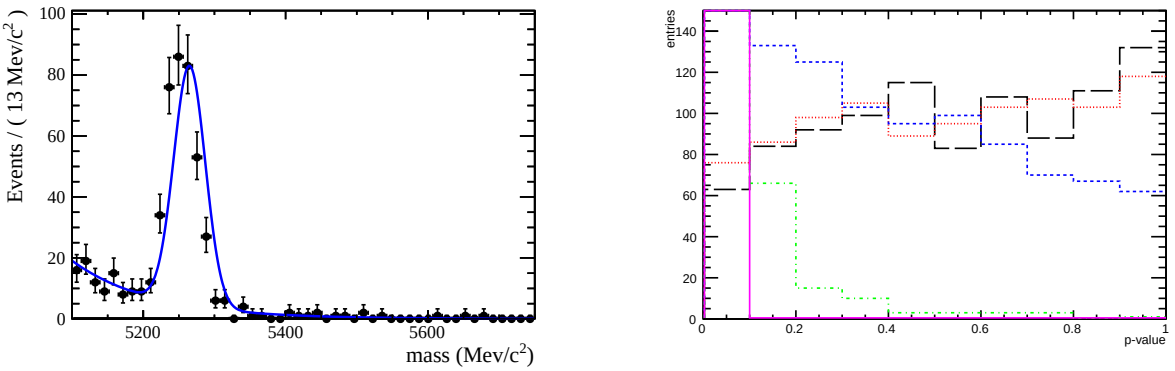


Figure 25: The B^0 invariant mass distribution is shown for a toy experiment (left), when the fixed Gaussian mean of the fit function is shifted by 10 MeV/c^2. The corresponding p-value distribution is shown on the right. The black line corresponds to a shift of 1 MeV/c^2, the blue line corresponds to 2 MeV/c^2, the red line correspond to 5 MeV/c^2, the green line corresponds to 10 MeV/c^2 and the purple line corresponds to 20 MeV/c^2.

In conclusion: additional peaks which are not included in the fit p.d.f., shift the p-value distribution to lower p-values. As it should be when the fit p.d.f. is not equal to the true p.d.f. of the data.

Remark

The point-to-point dissimilarity method, which was implemented and tested in this sections, has been applied to the distribution for forward and backward events before computing of the zero-crossing point of A_{FB}, to decide if the fitted p.d.f. well describes the data. This allowed us to model the q^2 distributions for the signal (for forward and backward category) keeping blind the A_{FB}. Avoiding the risk of involuntarily biasing the result. The choice of the order of the polynomial was only driven by the goodness of fit test, and the y-axis (i.e. the number of events for the two categories) was kept blind, such that it was impossible to compute the A_{FB} before choosing the p.d.f. used for fitting the signal.

5. Measurement

In Sec. 3 the unbinned counting method was described and tested. In this section the signal selection and the application of the method to data are described.

5.1. Selection

Here the selection of the signal events is shortly described. For more information about the selection see Ref. [10]. The signal candidates of $B^0 \to K^{*0}\mu^+\mu^-$ events must first pass a hardware trigger which selects muons with at transverse momentum, $p_T > 1.48\,\text{GeV}/c$. In the software trigger additional conditions are required to select the signal events (Refs. [29, 38]). The residual background can be classify according with the three categories: combinatorial background, partially reconstructed background and peaking background.

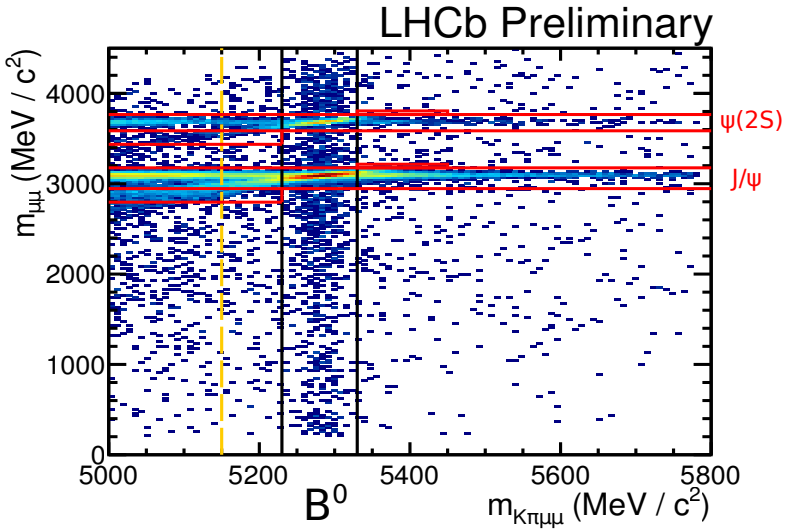


Figure 26: The $K^+\,\pi^-\,\mu^+\,\mu^-$ invariant mass versus the $\mu^+\,\mu^-$ invariant mass distribution for candidates in data. Plot reproduced from Ref. [8]. The red solid lines are the J/ψ and $\psi(2S)$ resonance regions. They are removed from the analysis. The yellow dashed line is the $K^+\,\pi^-\,\mu^+\,\mu^-$ invariant mass cut to remove partially reconstructed backgrounds. The black lines show a $\pm 50\,\text{MeV}/c^2$ window around the reconstructed B^0 mass.

The combinatorial background consists of particles from different decay trees which are randomly combined. This component is described with an exponential in the B-invariant mass fit and is reduced using a Boosted Decision Tree (BDT). The BDT uses information about the event kinematics, vertex and track quality, impact parameter and particle identification information.

The peaking background consists of the following decays: $B^0 \to J/\psi\,K^{*0}$, $B^0 \to \psi(2S)\,K^{*0}$ and $B^0 \to \phi\,\mu^+\,\mu^-$, where at least one particle is misidentified. Vetoes have been applied in order to reduce this background to a negligible level.

The partially reconstructed background consists of particles coming from the same B-meson decay where some particles from the same decay tree are missing. As a consequence this background is on the left of the B^0 mass.

5.2. $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to J/\psi K^{*0}$ invariant mass distributions

The mass model used for the signal and background is tested using $B^0 \to J/\psi K^{*0}$ events and $B^0 \to K^{*0}\mu^+\mu^-$ MC. The combinatorial background is modeled with an exponential function, while the partially reconstructed background is modeled with an empirical model, known as *RooExpAndGauss*. The *RooExpAndGauss* models an exponential rise to a threshold and a Gaussian fall off above the threshold.

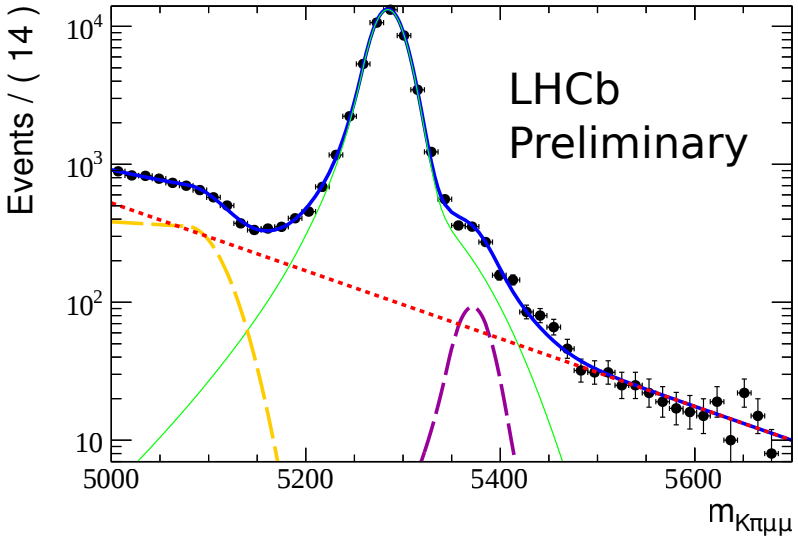


Figure 27: The $K^+\pi^-\mu^+\mu^-$ invariant mass of $B^0 \to J/\psi K^{*0}$ is shown. Plot reproduced from Ref. [8]. The red dashed line is the exponential background, the yellow long dashed line is the *RooExpAndGauss* contribution, the long dashed purple line is the additional signal component for $B_s^0 \to \bar{K}^{*0}\, J/\psi$ and the green line is the signal double Crystal Ball shape.

For the signal mass distribution of $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to J/\psi K^{*0}$ a double Crystal Ball shape [34] is used with both tails on the left-hand side of the mean. It is assumed that the B^0 mass, μ_{B^0}, and shape parameters α and n are identical for $B^0 \to J/\psi K^{*0}$ and $B^0 \to K^{*0}\mu^+\mu^-$. This has been verified using the MC simulation. Figure 27 shows the mass distribution for $B^0 \to J/\psi K^{*0}$ in the J/ψ mass window. In Fig. 28 the mass distribution for $B^0 \to K^{*0}\mu^+\mu^-$ is shown. The mean, μ_{B^0}, and the shape parameters α and n are fixed from the $B^0 \to J/\psi K^{*0}$ measurement, but the widths σ_1 and σ_2 are left floating in the fit.

In the $B^0 \to J/\psi K^{*0}$ fit an additional component is included for $B_s^0 \to \bar{K}^{*0}\, J/\psi$ decays, which is suppressed by f_s/f_d and a CKM factor. The fraction of B_s^0 decays is constrained from Ref. [3] to be $0.7 \pm 0.2\%$. In the fit for $B^0 \to K^{*0}\mu^+\mu^-$ the B_s^0 contribution is neglected, which result in a small additional systematic uncertainty.

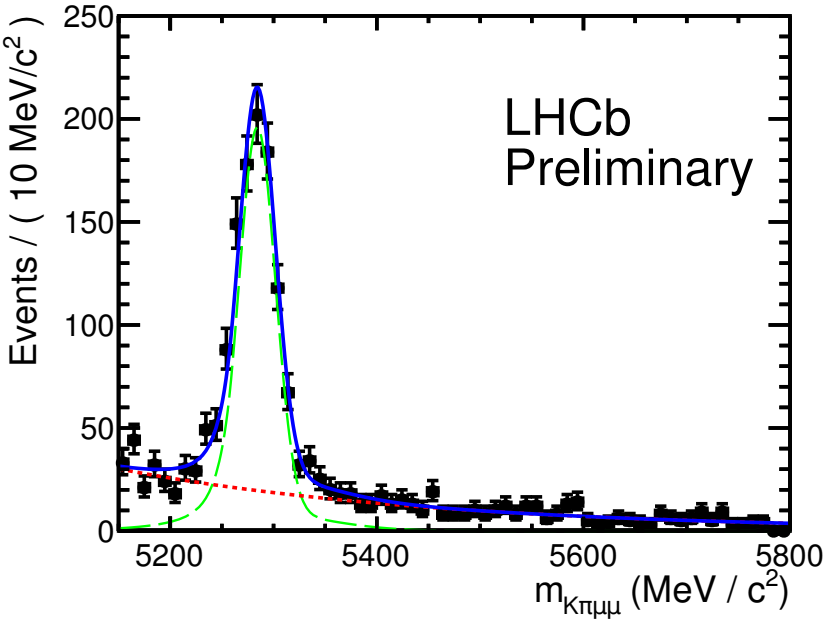


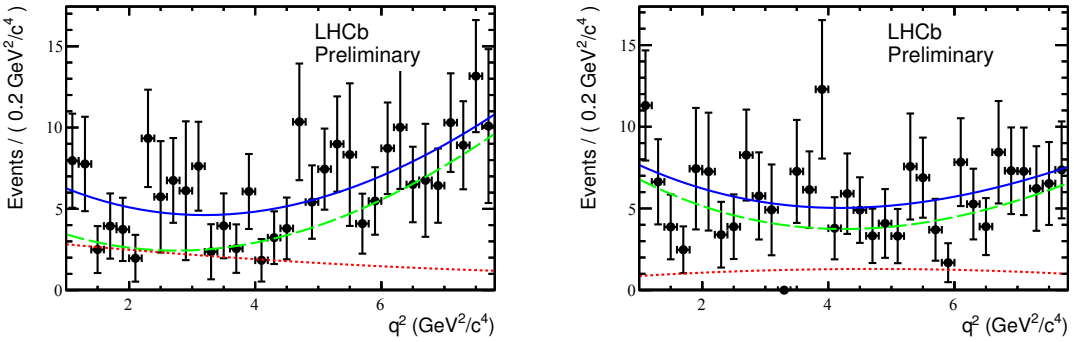
Figure 28: $K^+\pi^-\mu^+\mu^-$ invariant mass distribution of $B^0 \to K^{*0}\mu^+\mu^-$ candidates, in the range $4m^2 < q^2 < 19$ GeV$^2/c^4$, after the full selection has been applied. The signal mass model (green line) and the background mass model (red line) are described in the text. Figure taken from Ref. [8].

5.3. Zero-Crossing Point of A_{FB}

In this section the result of the zero-crossing point of A_{FB} is shown. For more details see Ref.[8]. The p.d.f. for forward- and backward events are expected to be smooth in q^2 in the range $1 - 7.8$ GeV$^2/c^4$. Above 7.8 GeV$^2/c^4$ there is a non-negligible contribution from the radiative tails of the J/ψ. Below 1 GeV$^2/c^4$ there is the photon pole, which is difficult to parametrize with a smooth polynomial. For the case $q^2 \to 0$ the amplitude of the penguin diagram, which includes the photon (see Fig. 4), increases rapidly, resulting in an infrared divergence. The mass model described in section 5.2 is used for the signal mass distribution and a third order Chebychev polynomial is used for the signal q^2 distribution. For the background q^2 distribution a second order Chebychev polynomial is used.

5.3.1. Result

The unbinned counting method from Sec. 3 was applied to data. Before calculating the A_{FB}, the fit of the forward and backward category was tested with the point-to-point dissimilarity method described in Sec. 4. Figure 29 shows the forward and backward going candidates.



(a) forward-going candidates (b) backward-going candidates

Figure 29: The q^2 distribution for the forward category (left) and for the backward category (right) for the signal region. The fit to data is superimposed. The dashed red line is the background, the dashed green line is the signal and the blue line is the sum of both. Figure taken from Ref. [8].

The fit for the forward category has a p-value of 0.6 and the fit for the backward category has a p-value of 0.9. Therefore the fits are in good agreement with data. After the goodness of fit test, the zero-crossing point of the A_{FB} and the 68% confidence level were calculated as shown in section 3. The results are shown in Fig. 30. The extracted zero-crossing point for $B^0 \to K^{*0}\mu^+\mu^-$ is:

$$q_0 = 4.9^{+1.1}_{-1.3}\,\mathrm{GeV}^2/c^4 [8].$$

This result is compatible with SM predictions and strongly disfavours models with flipped $\mathcal{C}_7$ Wilson coefficient with respect to SM.

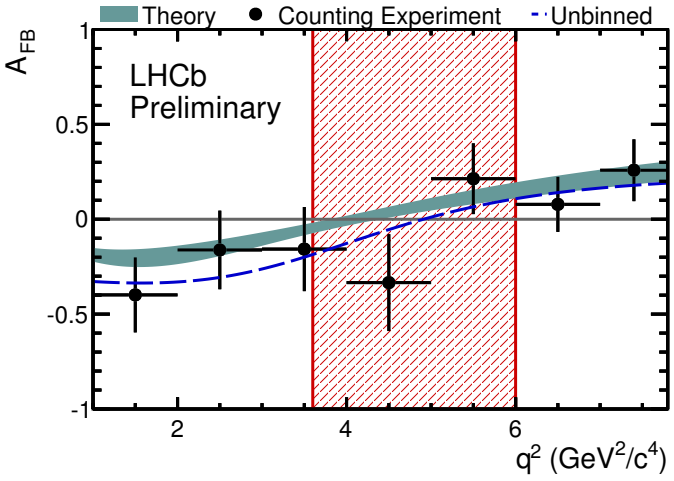


Figure 30: The A_{FB} as a function of q^2. The blue dashed line comes from the unbinned counting experiment. The green band is the theory prediction from Ref. [22]. The data-points are the result of counting forward- and backward-going events in $1\,\mathrm{GeV}^2/c^4$ bins of q^2. The uncertainty on the data-points is statistical only. The red-hatched region is the 68% confidence interval on the zero-crossing point observed in the data. Figure taken from Ref. [8].

6. Summary

For the analysis of the zero-crossing point of $A_{\rm FB}$ the unbinned counting method was used. The method splits the data into forward and backward events. Afterward the two categories are fitted in mass and q^2, and combined together to form the $A_{\rm FB}$ as a function of q^2. This method was tested on MC and later applied to data. The result for the zero-crossing point of $A_{\rm FB}$ is

$$q_0 = 4.9^{+1.1}_{-1.3}\,\mathrm{GeV}^2/c^4 [8]$$

and the SM prediction is

$$q_0^2 = 4.97^{+0.03}_{-0.03}|_{FF}{}^{+0.09}_{-0.09}|_{SL}{}^{+0.29}_{-0.27}|_{SD}\,\mathrm{GeV}^2/c^4 [22],$$

where (FF) are the form factor uncertainties, (SL) are the uncertainties from the Λ/m_b corrections and (SD) are the uncertainty from short distance parameters (m_t, m_W and the μ-scale). The measured $A_{\rm FB}$ is compatible with the SM prediction. As mentioned in section 2, earlier measurements seemed to prefer a flipped $\mathcal{C}_7$ Wilson coefficient with respect to the SM. This would have implied no zero-crossing in the $A_{\rm FB}$. The present analysis does not confirm this observations.

A goodness of fit method for multidimensional analysis was implemented: the point-to-point dissimilarity method. This has been a crucial aspect to test and validate the empirical model used to fit the data.

In addition a sensitivity study for the measurement of the observables S_4, S_5, S_7 and S_8 has been performed. It has been found that the zero-crossing point of the S_5 asymmetry can be measured by LHCb with a dataset corresponding to 1 fb^{-1}. These measurements have been validated using the LHCb simulation. The background model has been studied by using the B^0-mass sidebands in data. At the moment these observables are still "blind", but a measurement of them is foreseen for the near future.

References

[1] *The LHCb Detector at the LHC*. *Journal of Instrumentation*, 3(08):S08005, 2008.

[2] *Numerical Methods with Applications: Abridged*. Autar Kaw, 2011.

[3] *Evidence for the decay* $B_s^0 \to J/\psi\overline{K}^{*0}$. Jun 2011. Linked to LHCb-ANA-2011-025, LHCb-ANA-2011-071.

[4] *Tagged time-dependent angular analysis of* $B_s^0 \to J/\psi\phi$ *decays at LHCb*. Mar 2012. Linked to LHCb-ANA-2012-004.

[5] *Measurement of the direct CP asymmetry in the* $B_d^0 \to K^{*0}\gamma$ *decay*. Mar 2012. Linked to LHCb-ANA-2012-014.

[6] *Search for the* $D^0 \to \mu^+\mu^-$ *decay with 0.9* fb^{-1} *at LHCb*. Feb 2012. Linked to LHCb-ANA-2011-017.

[7] *Measurement of time-dependent CP violation in charmless two-body B decays*. Feb 2012. Linked to LHCb-ANA-2012-020.

[8] *Differential branching fraction and angular analysis of the* $B^0 \to K^{*0}\mu^+\mu^-$ *decay*. Jun 2012. Linked to LHCb-ANA-2011-089.

[9] *Latest update in the search for the Higgs boson. Latest update in the search for the Higgs boson*. Jul 2012.

[10] R. Aaij et al. *Differential branching fraction and angular analysis of the decay* $B^0 \to K^{*0}\mu^+\mu^-$. *Phys.Rev.Lett.*, 108:181806, 2012.

[11] R. Aaij et al. *Strong constraints on the rare decays* $B_s \to \mu^+\mu^-$ *and* $B^0 \to \mu^+\mu^-$. *Phys. Rev. Lett. 108,*, 231801, 2012. doi: 10.1103/PhysRevLett.108.231801.

[12] T. Aaltonen et al. *Measurements of the Angular Distributions in the Decays* $B \to K^{(*)}\mu^+\mu^-$ *at CDF*. *Phys.Rev.Lett.*, 108:081807, 2012.

[13] I. Adachi et al. *Measurement of the Differential Branching Fraction and Forward-Backward Asymmetry for* $B \to K^{(*)}\ell^+\ell^-$. 2008. arXiv:0810.0335v1 [hep-ex].

[14] A. Ali, G. Kramer, and Guo-huai Zhu. $B \to K^*\ell^+\ell^-$ *decay in soft-collinear effective theory*. *Eur.Phys.J.*, C47:625–641, 2006. doi: 10.1140/epjc/s2006-02596-4.

[15] Ahmed Ali, Patricia Ball, L.T. Handoko, and G. Hiller. *A Comparative study of the decays* $\bar{B} \to (K, K^*)\ell^+\ell^-$ *in standard model and supersymmetric theories*. *Phys.Rev.*, D61:074024, 2000. doi: 10.1103/PhysRevD.61.074024.

[16] Berkan Alsan. *The concept of energy in nonparametric statistics- Goodness-of-Fit problems and deconvolution*. PhD thesis, Universität Siegen, 2004.

[17] Wolfgang Altmannshofer, Patricia Ball, Aoife Bharucha, Andrzej J. Buras, David M. Straub, et al. *Symmetries and Asymmetries of* $B \to K^*\mu^+\mu^-$ *Decays in the Standard Model and Beyond*. *JHEP*, 0901:019, 2009. doi: 10.1088/1126-6708/2009/01/019.

[18] B. Aslan and G. Zech. *Statistical energy as a tool for binning-free, multivariate goodness-of-fit tests, two-sample comparison and unfolding*. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 537(3):626 – 636, 2005. ISSN 0168-9002. doi: 10.1016/j.nima.2004.08.071.

[19] Bernard Aubert et al. *Angular Distributions in the Decays $B \to K^{(*)}\mu^+\mu^-$*. *Phys.Rev.*, D79:031102, 2009. doi: 10.1103/PhysRevD.79.031102.

[20] M. Beneke, Th. Feldmann, and D. Seidel. *Exclusive radiative and electroweak $b \to d$ and $b \to s$ penguin decays at NLO*. *Eur.Phys.J.*, C41:173–188, 2005. doi: 10.1140/epjc/s2005-02181-5.

[21] Christoph Bobeth, Gudrun Hiller, and Danny van Dyk. *More Benefits of Semileptonic Rare B Decays at Low Recoil: CP Violation*. *JHEP*, 1107:067, 2011. doi: 10.1007/JHEP07(2011)067.

[22] Christoph Bobeth, Gudrun Hiller, Danny van Dyk, and Christian Wacker. *The Decay $\bar{B} \to \bar{K}\ell^+\ell^-$ at Low Hadronic Recoil and Model-Independent Delta B = 1 Constraints*. *JHEP*, 1201:107, 2012. doi: 10.1007/JHEP01(2012)107.

[23] Andrzej J. Buras. *Flavour Theory: 2009*. *PoS*, EPS-HEP2009:024, 2009.

[24] B. Efron. *Bootstrap methods: another look at the jackknife*. *Ann. Statistics*, 7, 1979.

[25] U Egede. *Angular correlations in the $\overline{B}_d \to \overline{K}^{*0}\mu + \mu-$ decay*. Technical Report LHCb-2007-057. CERN-LHCb-2007-057, CERN, Geneva, Jul 2007.

[26] U. Egede, T. Hurth, J. Matias, M. Ramon, and W. Reece. *New observables in the decay mode $\bar{B}_d \to \bar{K^{*0}}\ell^+\ell^-$*. *JHEP*, 0811:032, 2008. doi: 10.1088/1126-6708/2008/11/032.

[27] K. Nakamura et al. (Particle Data Group). *Particle Data Booklet*. *J. Phys. G 37*, 075021 (2010).

[28] Kevin T. Flood. *Angular and Rate Asymmetries in the Decays $B \to K^{(*)}\ell^+\ell^-$*. 2008. arXiv:0810.0837v2 [hep-ex].

[29] Vladimir V Gligorov. *A single track HLT1 trigger*. Technical Report LHCb-PUB-2011-003. CERN-LHCb-PUB-2011-003. LHCb-INT-2010-053, CERN, Geneva, Jan 2011.

[30] F. Jansen, N. Serra, G. Y. Smit, and N. Tuning. *Determination of the forward-backward asymmetry in the decay $B^0 \to K^{*0}\mu^+mu^-$ with an unbinned counting analysis*. Technical Report LHCb-2009-003. CERN-LHCb-2009-003, CERN, Geneva, May 2009.

[31] Albert Puig Navarro and Ricardo Graciani Diaz. *First measurements of radiative B decays in LHCb. oai:cds.cern.ch:1434459*. PhD thesis, Barcelona U., 2012. Presented 09 Mar 2012.

[32] Antonio Riotto and Mark Trodden. *Recent progress in baryogenesis*. *Ann.Rev.Nucl.Part.Sci.*, 49:35–75, 1999.

[33] Andrei D Sakharov. *Violation of CP in variance, C asymmetry, and baryon asymmetry of the universe*. *Soviet Physics Uspekhi*, 34(5):392, 1991.

[34] Tomasz Skwarnicki. *A study of the radiative CASCADE transitions between the Upsilon-Prime and Upsilon resonances*. 1986. DESY-F31-86-02.

[35] J.-T. Wei et al. *Measurement of the Differential Branching Fraction and Forward-Backward Asymmetry for* $B \to K(*)l + l-$. *Phys.Rev.Lett.*, 103:171801, 2009. doi: 10.1103/PhysRevLett.103.171801.

[36] Wikipedia. *Standard Model of Elementary Particles*, 2012. URL *http://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg*.

[37] Wikipedia. *Bisection Method*, 2012. URL *http://en.wikipedia.org/wiki/Bisection_method*.

[38] M. Williams, V. Gligorov, C. Thomas, H. Dijkstra, J. Nardulli, and P. Spradlin. *The HLT2 Topological Lines*. Technical Report LHCb-PUB-2011-002. CERN-LHCb-PUB-2011-002, CERN, Geneva, Jan 2011.

[39] Mike Williams. *How good are your fits? Unbinned multivariate goodness-of-fit tests in high energy physics*. *JINST*, 5:P09004, 2010. doi: 10.1088/1748-0221/5/09/P09004.

[40] G. Zech and B. Aslan. *A multivariate two-sample test based on the concept of minimum energy*. In *PHYSTAT*, 2003.

Appendix

A. RootFinder Method

The *RootFinder* method is described in section 3.2.2. It is based on the bisection method to find the root of a function.[2]

Algorithm for the bisection method

1. Choose x_l and x_k such that $f(x_l)f(x_k) < 0$, or in other words, $f(x)$ changes sign between x_l and x_k.

2. The first estimation for the root, x_m is the mid-point between x_l and x_k as

$$x_m = \frac{x_l + x_k}{2} \tag{39}$$

3. Now different cases must be checked:

   a) If $f(x_l)f(x_m) < 0$, then the root lies between x_l and x_m; then $x_l = x_l$ and $x_k = x_m$.
   b) If $f(x_l)f(x_m) > 0$, then the root lies between x_m and x_k; then $x_l = x_m$ and $x_k = x_k$.
   c) If $f(x_l)f(x_m) = 0$, then the root is x_m. Stop the algorithm if it is true.

4. Repeat step 2-3 until the number of iteration is greater then a pre-specified number of iteration N_{max} or $\frac{x_k - x_l}{2} < \varepsilon$, where ε is the pre-specified error tolerance.

Figure 31 shows an iteration example for the described bisection method.

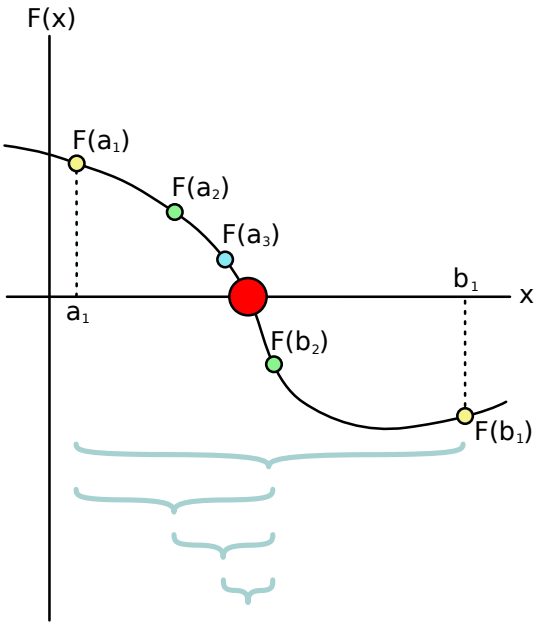


Figure 31: The figure shows the iteration for finding the root. The gray brackets show the different iteration steps from top to down.[37]