

**Department of Physics and Astronomy
University of Heidelberg**

Bachelor Thesis in Physics
submitted by

Philipp Kühnl

born in Bruchsal (Germany)

2015

Determination of the branching ratio of the decay

$$B^+ \rightarrow \mu\mu K^{*+} (\rightarrow K_S^0 \pi^+)$$

This Bachelor Thesis has been carried out by Philipp Kühnl at the
Physikalisches Institut, LHCb group
under the supervision of
Prof. Dr. Ulrich Uwer

Abstract

This thesis presents a measurement of the branching ratio \mathcal{B} of the rare B meson decay mode $B^+ \rightarrow \mu\mu K^{*+}$ relative to the decay mode $B^+ \rightarrow J/\Psi(\rightarrow \mu\mu)K^{*+}$. The used data has been taken at the LHCb detector and represents the entire data taken within the first run, that is the years 2011 and 2012. The LHC was operating at center-of-mass energies of the pp -collisions of 7 TeV (2011) and 8 TeV (2012), respectively. An amount of $N = 155 \pm 16$ events decaying as $B^+ \rightarrow \mu\mu K^{*+}$ has been reconstructed, translating into a relative branching ratio of

$$\frac{\mathcal{B}(B^+ \rightarrow \mu\mu K^{*+})}{\mathcal{B}(B^+ \rightarrow J/\Psi(\rightarrow \mu\mu)K^{*+})} = (1.23 \pm 0.13_{stat}) \times 10^{-2},$$

where only statistical errors have been quantified. Utilizing the known branching ratio of the reference channel [1] the absolute branching ratio can be determined as

$$\mathcal{B}(B^+ \rightarrow \mu\mu K^{*+}) = (1.05 \pm 0.11_{stat} \pm 0.06_{norm}) \times 10^{-6}.$$

Kurzfassung

In dieser Arbeit wird das Verzweigungsverhältnis \mathcal{B} des seltenen Zerfalls eines B -Mesons gemäß $B^+ \rightarrow \mu\mu K^{*+}$ relativ zum Zerfallskanal $B^+ \rightarrow J/\Psi(\rightarrow \mu\mu)K^{*+}$ bestimmt. Die hierfür verwendeten Daten stammen aus pp -Kollisionen am LHCb-Detektor und wurden in den Jahren 2011 und 2012 bei Schwerpunktsenergien von 7 TeV bzw. 8 TeV aufgezeichnet. Insgesamt konnte eine Anzahl von $N = 155 \pm 16$ Zerfälle gem. $B^+ \rightarrow \mu\mu K^{*+}$ rekonstruiert werden, was einem relativen Verzweigungsverhältnis von

$$\frac{\mathcal{B}(B^+ \rightarrow \mu\mu K^{*+})}{\mathcal{B}(B^+ \rightarrow J/\Psi(\rightarrow \mu\mu)K^{*+})} = (1.23 \pm 0.13_{stat}) \times 10^{-2}$$

entspricht. Dabei wurden nur statistische Unsicherheiten quantitativ bestimmt. Unter Hinzunahme des bekannten Verzweigungsverhältnisses des Referenzkanals [1] erhält man das absolute Verzweigungsverhältnis von

$$\mathcal{B}(B^+ \rightarrow \mu\mu K^{*+}) = (1.05 \pm 0.11_{stat} \pm 0.06_{norm}) \times 10^{-6}.$$

Contents

1	Introduction	6
2	Theory	8
2.1	Particles of the Standard Model	8
2.2	Quantum flavor dynamics	10
2.3	The decay $B^+ \rightarrow \mu\mu K^{*+}$	11
3	Experimental setup	14
3.1	The Large Hadron Collider	14
3.2	The LHCb detector	15
3.2.1	Tracking	16
3.2.2	Particle identification	17
3.2.3	Trigger system	18
4	Particle reconstruction and used data sets	20
4.1	Reconstruction of decayed particles	20
4.2	Data set from the LHCb detector	20
4.3	Mass distributions after stripping	21
4.4	Simulated data	23
5	Selection Process	24
5.1	Summary of the analysis strategy	24
5.2	Selection of $\mu\mu$ -resonances	24
5.2.1	Number of $B^+ \rightarrow J/\Psi(\rightarrow \mu\mu)K^{*+}$ events	25
5.2.2	Values for J/Ψ and $\Psi(2S)$ vetoes	27
5.3	Reduction of combinatorial background	29
5.3.1	Definition of the used variables	29
5.3.2	Multivariate Analysis	30
5.4	Results of the selection process	34
6	Determination of the branching ratio	38
7	Discussion of uncertainties	43
7.1	Size and effects of statistical errors	43
7.2	Systematic uncertainties	43
8	Summary	47
	References	49

1 Introduction

Since the discovery of the electron and its first description as a particle in 1897 [2], the physicists' models about the fundamental building blocks of matter have been modified and replaced many times. New experimental discoveries incompatible with or insufficiently described by these models were the impetus of these changes. Today, the best experimentally validated theory is the Standard Model (SM) of particle physics. Its predictive power has been tested and approved many times, the most recent success being the discovery of the Higgs boson in 2013. Nonetheless, it cannot explain every observed phenomenon like the character of so-called dark matter and dark energy, which (according to current understanding and theories) make up the most part of the universes energy. Also, the SM description contains only massless neutrinos, which contradicts experiments that indicate neutrino oscillations, which can only occur if these particles were massive. Another striking topic is the matter-antimatter asymmetry in the observed universe. Although the Standard Model gives explanations for asymmetries, these cannot be solely responsible for the observed asymmetry that is our universe being made of matter, not antimatter. Additionally, the SM does not account for the effects of gravity, one of the four fundamental forces in modern physics. A final and unified theory is yet to be found, and there exist a variety of models waiting to be tested which could maybe become a step into that direction.

To further study and (possibly) answer these and many more questions, the Large Hadron Collider (LHC) and its four major detectors ATLAS, CMS, ALICE and LHCb have been build at CERN. Among these four experiments, the LHCb detector focuses especially on the observation of B and D mesons. Its main goal is to investigate the matter-antimatter asymmetry, whose origin lies in CP-violation [3]. Rare decays of B mesons are also especially suited for the indirect search of possible particles beyond the Standard Model, which could account for deviations from the SM predictions in the branching ratio of these decays. Recent results hint in that direction, yet await to be further studied ([4], [5], [6]).

The data used in this analysis has been recorded with the LHCb detector in the years 2011 and 2012. During this runtime, the proton-proton collisions took place at center-of-mass energies of 7 TeV (2011) and 8 TeV (2012), respectively, corresponding to a total integrated luminosity of $\mathcal{L}_{int} \approx 3\text{fb}^{-1}$.

This bachelor thesis is structured as follows: It will start with a brief theoretical introduction into the Standard model in Section 2.1 and especially discuss flavour physics in some more detail in the separate Section 2.2. Subsequently, Section 2.3 will then focus on the specific discussion of the decay $B^+ \rightarrow \mu\mu K^{*+}$.

To realize an experimental examination of the theoretical models one needs huge and yet precise experimental setups. The data used within this thesis has been collected at the LHCb detector, which will be introduced in Section 3. Knowing now the theoretical

and experimental context, [Section 4](#) will give record of the used data sets (consisting of real detector data as well as simulated events). In [Section 5](#) the different steps of the selection process of the events decaying as $B^+ \rightarrow \mu\mu K^{*+}$ will be explained, likewise for the reference channel $B^+ \rightarrow J/\Psi K^{*+}$. The final determination of the branching ratio will then be done in [Section 6](#). For the error estimation only statistical uncertainties will be considered. Although this thesis cannot provide a detailed analysis of all systematic effects, these shall be discussed to some extent together with the relevance of the statistical errors in [Section 7](#) before [Section 8](#) gives a concluding summary.

2 Theory

This section gives a brief theoretical overview of the essentials needed to understand the analysis presented within this thesis. It will explain the basic concepts of the currently used and to wide parts experimentally validated theory of particle physics: the so-called Standard Model (SM) of particle physics. Of the three fundamental forces of the SM the focus will be laid on the weak force and its theory of quantum flavor dynamics, since this is the force responsible for the decay of the B meson. The end of this section is dedicated to a little deeper look into B meson physics, especially the decay $B^+ \rightarrow \mu\mu K^{*+}$ investigated in this thesis. Further and more extensive information can be found in any introductory particle physics or SM textbook, e.g. [8], [9] or [10]. It shall be noted that throughout this thesis natural units will be used (speed of light set to $c = 1$), resulting in energy, momentum and mass share the unit of measure (MeV).

2.1 Particles of the Standard Model

The Standard Model is a quantum field theory containing the current knowledge of fundamental particles and the interactions between them. It has derived over the second half of the 20th century and has been widely and successfully tested in various experiments ever since. Its latest success has been the discovery of the Higgs boson in 2013, experimentally validating the Higgs mechanism incorporated in the SM. Despite its predictive power, the SM is still an incomplete theory. It cannot explain all observed phenomena and needs to be extended to (or replaced by) a more complex theory someday.

The fundamental constituents of matter within the SM are twelve fermions with spin $\frac{1}{2}$ which can be grouped into leptons and quarks and further divided into three generations each, as shown in Table 1 and Table 2. The unique type of each of these particles is called

Gen.	Type	Symbol	El. charge [e]	Mass
I	up	u	+2/3	$2.3^{+0.7}_{-0.5}$ MeV
	down	d	-1/3	$4.8^{+0.7}_{-0.3}$ MeV
II	charm	c	+2/3	1.275 ± 0.025 GeV
	strange	s	-1/3	95 ± 5 MeV
III	top	t	+2/3	$173.21 \pm 0.51 \pm 0.71$ GeV
	bottom ¹	b	-1/3	4.18 ± 0.03 GeV

Table 1: List of quarks, masses taken from [1]

flavor. In addition for each particle there is its corresponding antiparticle ², adding up to a number of 24 elementary fermions.

The fundamental forces of physics (with the exception of gravity, whose strength is negligible at the considered scales and not included in the SM) are represented in form of interactions between these particles, mediated by gauge bosons. Each force has its corresponding charge ³, and only particles carrying this type of charge can interact with the respective boson. The most important properties of the gauge bosons are summarized in Table 3. The gauge boson of the electromagnetic force, the photon, couples to all particles carrying an electric charge. Its range is infinite, yet decreasing with distance.

Gen.	Type	Symbol	El. charge [e]	Mass
I	electron	e^-	-1	0.511 MeV
	electron neutrino	ν_e	0	< 2 eV
II	muon	μ^-	-1	105.66 MeV
	muon neutrino	ν_μ	0	< 2 eV
III	tauon	τ^-	-1	1776.82 ± 0.16 MeV
	tau neutrino	ν_τ	0	< 2 eV

Table 2: List of leptons, masses taken from [1]. For e^- and μ^- the measured uncertainty is several orders of magnitude below the rounded values shown here. For the neutrino masses only upper border values exist so far.

Strong force As the name suspects, it is the strongest of the fundamental forces, but its range is very short (of about up to 1 fm). The only particles affected by it are quarks (and the gluons themselves), since they are the only ones carrying color charge. The fact that the gauge bosons also carry charge allows gluon-gluon coupling. In contrast to the electric charge color comes in three different types, labeled as red, green and blue. The so-called *confinement* describes the impossibility of having a free quark: The further two quarks get separated from each other, the larger the potential energy between them, giving rise to the creation of quark-antiquark pairs. Hence, quarks never appear alone and are forced to form bound states. These bound states must be neutral in color and therefore can consist of either three quarks (or three antiquarks), called (anti-) baryon, or out of a quark-antiquark pair, called meson. The terms meson and baryon can be combined to the term hadron, which labels any particle made of quarks.

²The antiparticle has the reverse electrical charge but shares all other physical properties. Antiparticles are mostly labeled by a bar on top of the corresponding particles symbol (\bar{u} instead of u), or by referring to the charge (e^+ instead of e^-).

³This expression is chosen in analogy of the electric charge of the electromagnetic force.

Weak force and electroweak unification The unique property of the weak force in contrast to the other two is the fact that the gauge bosons (W^\pm and Z) have a mass. This is the reason this force has been labeled *weak*, as these masses suppress interactions via these bosons in typical energy ranges of nuclear physics (around a few GeV). For much higher particle momenta this suppression gets minuscule and the strength of the weak force exceeds the electromagnetic interaction. The short range of the weak force, which is about 10^{-3} fm can be explained by the lifetimes of the bosons, which are of order 10^{-25} seconds. Since the weak force is the only one that can change the flavor of a particle, it is the one of interest for the decay investigated in this thesis and will be further discussed throughout this section. For completeness, it should be mentioned that it has been achieved to unify the theories of electromagnetic and weak interactions to the so-called electroweak interaction, which is theoretically described in the Glashow-Salam-Weinberg (GSW) model embedded in the SM. Since this distinction is not of further interest for the goals of this thesis it is sufficient to solely refer to the weak interaction theory from now on.

Interaction	Couples to	Boson	Spin	Mass
electromagnetic	electric charge	photon (γ)	1	0
strong	colour charge	8 gluons (g)	1	0
weak	weak charge	W^\pm	1	80.4 GeV
		Z	1	91.2 GeV
-	Mass	Higgs (H^0)	0	125.7 ± 0.4 GeV

Table 3: List of the fundamental gauge bosons plus the Higgs. Masses taken from [1].

2.2 Quantum flavor dynamics

As already mentioned, the quark flavor is conserved in all interactions except weak processes. The change of flavor via weak interactions is described by the Cabibbo-Kobayashi-Maskawa (CKM) mechanism as follows: Since the weak eigenstates (d' , s' , b') differ from the physical eigenstates, also called mass eigenstates (d , s , c), it is possible for a quark to change its flavor. The two types of eigenstates are linked via the unitary CKM matrix as

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \quad (1)$$

The probability of a quark of flavor i to decay to flavor j is proportional to $|V_{ij}|^2$. There exist different ways of parametrizing this matrix which will not be introduced here.

Instead, the measured numerical values for V_{ij} shall be given (values taken from [8]):

$$\begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} \approx \begin{pmatrix} 0.974 & 0.225 & 0.004 \\ 0.225 & 0.973 & 0.041 \\ 0.009 & 0.040 & 0.999 \end{pmatrix}. \quad (2)$$

As can be seen from these values, it is most likely for the flavor change to happen within a generation (diagonal values close to 1), yet to much smaller amounts the decay into another generation is possible. A direct decay between an up-type quark and a down-type quark ⁴ is realized by emission of a W^\pm as indicated in Figure 1. In the chosen example a strange-quark ($-1/3$) changes flavor to any of the up-type quarks ($+2/3$). Due to charge conservation, the emitted W boson needs to be negatively charged. Such a direct decay is called to be on *tree level*, the fact that the W is charged gives rise to the expression *flavor changing charged current*. Decays of higher order, involving more than one gauge boson are possible, however the probability of this to happen is much lower. In particular, decays characterized by no change of net charge (flavor changing *neutral currents*) are only possible in higher order processes, hence such decay modes are rare.

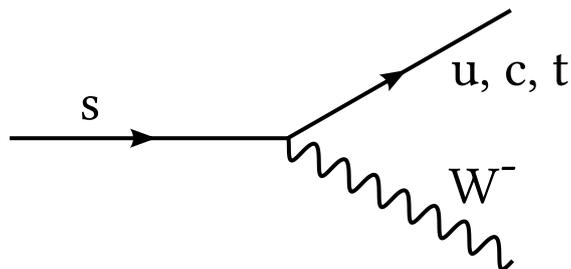


Figure 1: Feynman diagram of a weak process, where a s decays to any up-type quark. This diagram only shows the part of interest, which is the vertex of the weak interaction. It is incomplete in the way that in nature the decaying quark is in a bound state, whose other components are missing here. Also, the further interaction of the W is not shown.

2.3 The decay $B^+ \rightarrow \mu\mu K^{*+}$

The investigated decay channel in this thesis originates from a B^+ meson, which is a bound $u\bar{b}$ -state. Of these two quarks, the \bar{b} is the one that will decay weakly, whilst the up-quark has merely the role of a spectator. In this process the \bar{b} transforms into a \bar{s} .

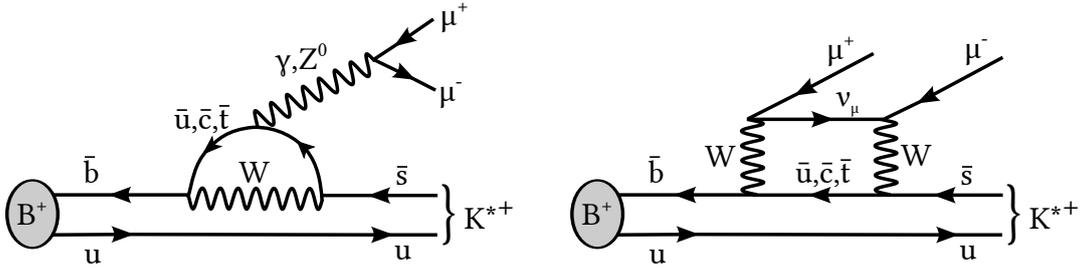
⁴*Up-type* labels the quarks with an el. charge of $+2/3$, like the up-quark, whereas *down-type* refers to the ones with $-1/3$.

Hence, this decay is a representative of a flavor changing neutral current, as both have a charge of $+1/3$. As mentioned before this means tree level diagrams can be excluded and the leading order Feynman diagrams are so-called penguin or box diagrams. These are shown in [Figure 2](#).

In these diagrams the inner solid lines, representing so-called *virtual* particles, can be either a \bar{u} , \bar{c} or \bar{t} . For each of these three possibilities i the transition amplitudes at the two vertices with the W are proportional to the CKM matrix elements (V_{ib} for the first and V_{is} for the second vertex). If compared to the numbers of [Equation \(2\)](#) one can see that these processes get dominated by virtual (anti-) top quarks. Nonetheless, in physics beyond the Standard Model it could be possible for some yet unknown particles to contribute to the total transition amplitude at exact this point. Measuring branching ratios of rare B meson decays is one method to get information about possible new particles. Deviations of SM predictions for such rare decay modes can therefore be a hint to yet unknown physics. The so far most precisely measured value for the branching ratio is

$$\mathcal{B}(B^+ \rightarrow \mu\mu K^{*+}) = (1.12 \pm 0.15) \times 10^{-6} \text{ [1]}.$$

The two final state quarks in the Feynman diagrams of [Figure 2](#) form the K^{*+} , which decays further into a K_s^0 and a π^+ . [Figure 3](#) sketches the spatial topology of the decay.



(a) Penguin diagram

(b) Box diagram

Figure 2: Feynman diagrams of the decay $B^+ \rightarrow \mu\mu K^{*+}$.

Here, the primary vertex (PV) labels the point where the B meson gets created. After some time of flight, at the secondary vertex (SV), it decays to the K^{*+} and the two muons. Since the lifetime of the K^{*+} is of order 10^{-23} seconds it cannot be spatially resolved and its decay products π^+ and K_s^0 seem to emerge from the SV, too. The K_s^0 carries no electric charge, hence is indicated as a dotted line illustrating the inability of detection within the charged particle subdetectors. Finally, at the tertiary vertex (TV), the K_s^0 decays into two pions.

Furthermore, the decay $B^+ \rightarrow J/\Psi K^{*+}$ will be used as reference channel. The J/Ψ is a $c\bar{c}$ -resonance, which can decay to two muons, in which case the final decay products are identical to the ones of the investigated decay. The leading order Feynman diagram

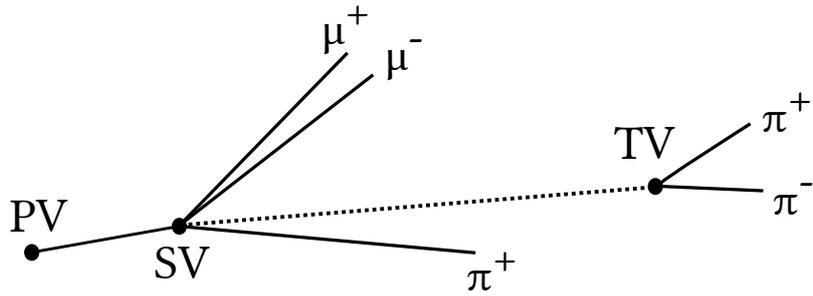


Figure 3: Schematic representation of the spatial occurrence of the decay $B^+ \rightarrow \mu\mu K^{*+}$. Labeled are only the final state particles (π^\pm, μ^\pm).

(tree level) of this process is shown in [Figure 4](#). Higher orders exist but contribute to a much smaller amount. The branching ratio of this reference channel has been measured to high accuracy as

$$\mathcal{B}(B^+ \rightarrow J/\Psi K^{*+}) = (1.44 \pm 0.08) \times 10^{-3} [1].$$

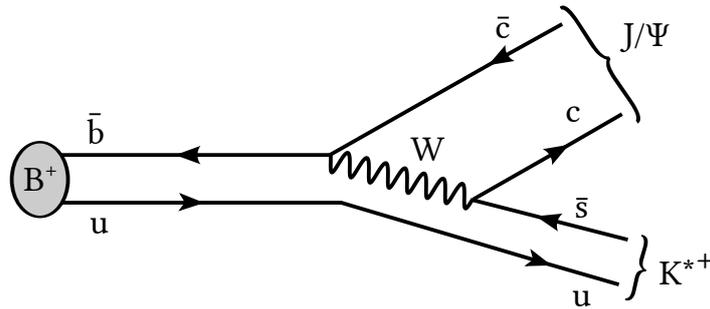


Figure 4: Feynman diagram (tree level) of the decay $B^+ \rightarrow J/\Psi K^{*+}$

3 Experimental setup

This sections' aim is a short presentation of the Large Hadron Collider beauty (LHCb) experiment, which is one of the four major detectors located at the LHC at CERN ⁵ near Geneva, Switzerland. The two next subsections will introduce first the LHC, followed by the LHCb detector.

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is a proton-proton (pp) collider and so far the worlds largest particle accelerator. It is located at CERN in Geneva (Switzerland) in an underground tunnel with circumference of about 27 km. An illustration of the structure of the collider can be seen in [Figure 5](#). Inside the tube two proton beams (squeezed into separate bunches) get accelerated in opposite directions. At four points, around which the major detectors are built, the beams cross to force pp -collisions. The LHC was designed to operate at center-of-mass energies of up to $\sqrt{s} = 14$ TeV, yet this value will be achieved gradually. At its first run, the LHC was operating at 7 TeV (2011) and 8 TeV (2012). After a shutdown phase used to optimize and replace parts of the collider and the detectors, the second run is about to start in spring 2015 with an aimed center-of-mass energy of $\sqrt{s} = 13$ TeV.

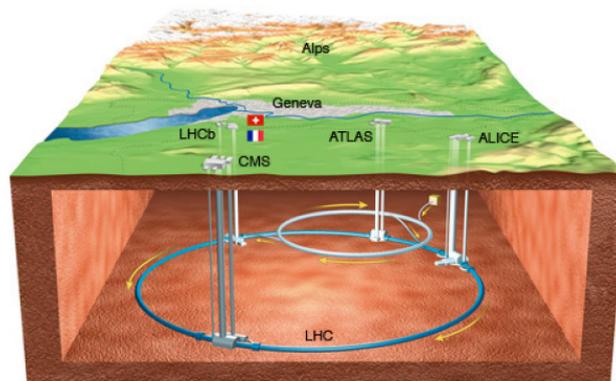


Figure 5: Schematic illustration of the LHC with its four major experiments, among them the LHCb detector [11].

⁵Conseil Européen pour la Recherche Nucléaire, the European organization for nuclear research

Of the four major experiments, ATLAS (A Toroidal LHC Apparatus) and CMS (Compact Muon Solenoid) are designed as general purpose detectors covering a large solid angle to detect very heavy particles. These are also the two experiments which first discovered the Higgs boson in 2013. The ALICE (A Large Ion Collider Experiment) detector has been built to study the quark-gluon plasma in heavy ion collisions ⁶, while the LHCb experiment focuses on precision measurements of hadrons containing *b*- and *c*-quarks.

3.2 The LHCb detector

This section is meant to give an overview about the structure and functioning of the LHCb ⁷ detector, which provided the data used within this thesis. All information displayed here has been gathered from [12], to which is referred to in case of additional interest.

The purpose of this detector, precision measurements of *B* and *D* decays, makes its setup quite different from the other three major experiments. Figure 6 shows a cut of the detector along the z-y-plane, where the z-axis is chosen along axis of the beam and the y-axis represents the vertical axis. One major difference to the other detectors is the covered solid angle, as LHCb only covers a cone around the beam axis. As the momenta of most of the colliding partons ⁸ in z-direction exceed their transverse momenta by far, this leads to the fact that the largest amount of *B* mesons produced at the *pp*-collisions are boosted in either $\pm z$ direction and therefore appear at angles close to the beam axis. This gets quantified in a variable called pseudo-rapidity η ⁹. The angular coverage of the detector spreads from 10 to 250 mrad for the y-z-plane and from 10 to 300 mrad for the x-z-plane, which corresponds to a pseudo-rapidity window of approximately $1.6 < \eta < 4.9$. Of course, since the detector setup is oriented in one direction ($+z$), all produced particles with a boost in the other direction ($-z$) cannot be detected. Being able to detect all different types of particles requires different subdetectors which are labeled in Figure 6. These parts can be divided into two categories relevant for tracking or particle identification purposes, as well as the trigger system. The following subsections will cover each of these separately.

⁶In addition to protons, the LHC gets also used to accelerate ionized Pb atoms.

⁷The b stands for *beauty*, which is used in parallel to the term *bottom* for the *b*-quark

⁸The term *parton* refers to any particle inside the bound nucleon, hence (valence-/sea-)quarks and gluons.

⁹ The pseudo-rapidity is defined as $\eta = -\ln[\tan(\frac{\theta}{2})]$, where θ represents the polar angle with respect to the z-axis.

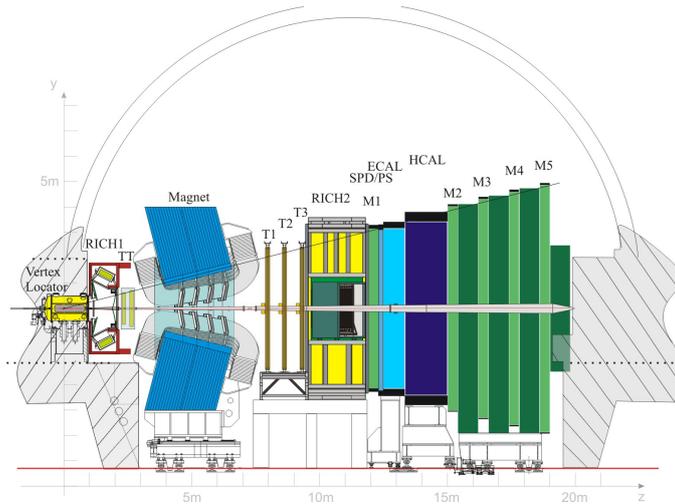


Figure 6: View of the LHCb detector as a lateral cut [12].

3.2.1 Tracking

The parts associated with tracking and track reconstruction are the Vertex Locator (VELO), the Tracker Turicensis ¹⁰ (TT) as well as the three tracking stations (T1 - T3). The detectors dipole magnet (not superconducting) is oriented such that charged particles get deflected in $\pm x$ -direction. This allows to assign the charge of the particle. Also it is used for measuring the particles momenta in alliance with the tracking detectors. The homogeneous magnetic field has an magnetic flux of 4 Tm, integrated over a distance of 10 m. Furthermore it is possible to switch the polarization of the magnet, which allows to correct for possible charge asymmetries of the detector.

VELO The Vertex Locator is constructed directly around the point of pp -collisions. Its purpose is the precise tracking of charged particles emerging of a decay vertex, hence it provides data to spatially reconstruct these vertices. Due to its high resolution it is possible to determine at which vertex a particle was produced, which is essential to reconstruct decay chains involving more than one vertex. Since the lifetimes of the decaying particles of interest are very short, these vertices are spatially close. The VELO consists of 21 circular slices (which are split into two half-circles each) of silicon strip detectors arranged perpendicular to the z -axis at different locations of z . The tracks of charged particles can then be reconstructed from the various interaction points on the detector slices, each of which can be determined with a precision of up to $4 \mu\text{m}$.

¹⁰Formerly also known as Trigger Tracker

TT and tracking stations The three tracking stations (T1 - T3) consist of an inner tracker (IT) close to the beam axis, where the density of particle tracks is much higher, and an outer tracker (OT). Both the Tracker Turicensis (TT) and the IT are again silicon strip detectors functioning similar to the VELO, yet with a lower resolution. The IT for example allows for a spatial resolution of about $50 \mu\text{m}$. They are made out of four layers of silicon strips each, which are alternately rotated by an angle of 10° within the x-y-plane to allow for a three-dimensional positioning. The OT is made of gas-filled drift chambers, where hits of ionizing particles can be used to reconstruct their tracks. Here, the spatial resolution is again lower and of about $200 \mu\text{m}$. Accounting for the bending of the particles tracks due to the dipole magnet, the trackers can be used for a measurement of charged particle momenta. Combining all interaction points within the different tracking systems yields a full reconstruction of the particle tracks.

3.2.2 Particle identification

After having collected information about the track and the momentum of a particle the remaining interest lies in identifying the exact type of the detected particle. This process is also called particle identification (PID). The components fulfilling this task are the RICH-detectors, the electromagnetic and hadronic calorimeter (ECAL and HCAL) as well as the muon chambers (M1 - M5).

RICH-detectors These RICH (Ring-imaging Cherenkov)-detectors utilize the effect of Cherenkov radiation to obtain information about a particle passing through. This radiation is emitted if a charged particle traverses a dielectric medium (with refraction index n) with a velocity v greater than the speed of light $c' = \frac{c}{n}$ in that medium. This leads to the emission of photons along a cone whose opening angle θ is defined as

$$\cos \theta = \frac{1}{n\beta}, \quad (3)$$

with $\beta = \frac{v}{c}$. Hence the radiation angle is a direct measure of the particles speed. If combined with the momentum information of the tracker systems this allows for the calculation of the particles mass. The measured mass can then be compared to the existing particle listing of the Particle Data Group (PDG, see [1] or [7]) to identify the particles type. Within the LHCb detector two RICH detectors containing different radiator materials get used, covering different particle momenta between about 1 and 60 GeV (RICH1) and about 15 to 100 GeV (RICH2).

Calorimeter The whole calorimeter system consists of an electromagnetic (ECAL) and an hadronic (HCAL) calorimeter, as well as the built-in Scintillator Pad Detector (SPD) and Preshower detector (PS). They have the purpose of identifying photons, electrons (or positrons) and different types of hadrons. A trespassing particle deposits its energy within the calorimeter in a cascading shower of secondary particles producing scintillation light amplified and detected by photomultipliers. The amount of produced scintillation light can be set in relation to the incoming particles energy deposit. The energy resolution of the used calorimeter is (E measured in GeV):

$$\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E}} \oplus 1\% \quad (\text{ECAL}), \quad (4)$$

$$\frac{\sigma_E}{E} = \frac{(69 \pm 5)\%}{\sqrt{E}} \oplus (9 \pm 2)\% \quad (\text{HCAL}). \quad (5)$$

In addition, the hit position of the calorimeter can help in track reconstruction.

Muon chambers The muon chambers M1 - M5 are, as the name suspects, responsible for the detection of muons. The first one is located in front of, the remaining four behind the calorimeter (defined along +z direction). The latter four are separated by 80 cm thick iron absorbers to prevent the passage of any high-energetic hadron.

Particle hypothesis Combining now all information of the different subdetectors, a hypothesis for the type x of each detected particle can be made. This is done by using the (logarithmic) Likelihood method. Each particle track is associated with a likelihood value \mathcal{L}_x which can be compared to different particle hypotheses (like e.g. being a pion, kaon or proton). Conventionally it is chosen to be set in relation to a π . The hypothesis is then quantified by the so-called DLL variables:

$$\text{DLL}_{x\pi} = \Delta \ln \mathcal{L}(x - \pi) = \ln \mathcal{L}(x) - \ln \mathcal{L}(\pi) = \ln\left(\frac{\mathcal{L}(x)}{\mathcal{L}(\pi)}\right), \quad (6)$$

where mostly a π is chosen as reference, being the particle type produced in largest amount.

3.2.3 Trigger system

The trigger systems duty is to reduce the collected data to an amount that is manageable to store on disk and contains all possibly interesting events. It consists of two stages:

The hardware-based Level-0 (L0) triggers aim is to reduce the nominal beam crossing rate of up to 40 MHz (when all bunches are filled, in 2011 and 2012 the LHC operated at lower rates) to a rate below 1 MHz, at which the entire detector can be read out. This is achieved by filtering the events for possible B meson candidates which, due to their large mass, preferably decay into particles of high transverse energy (E_T) and momentum (p_T). The L0 trigger therefore concentrates on calorimeter events with high E_T as well as high p_T muon chamber hits. In the next stage, the software-based High-Level-Trigger (HLT) further reduces this rate in two steps (HLT1 and HLT2) to about 2 kHz which is then stored on disk. Since at LHCb a variety of analyses is done, each of which may have different requirements on these trigger selections, the algorithms run with different so-called trigger lines on which selections can be made to retain the suited criteria for each offline analysis individually.

4 Particle reconstruction and used data sets

This section shall give a short overview about the procedure of reconstructing decayed particles as well as introducing the used sets of real data and its pre-selection process as well as simulated data.

4.1 Reconstruction of decayed particles

The particles which are actually detected as tracks within the detector system are the final state particles after the decay process, as has been outlined in [Section 2.3](#) and [3](#). In case of this particular decay the final state particles are pions and muons. Using the information of each particles traveled path one can reconstruct which particles originate from a common vertex, and with the measured four-momentum of each final state particle p_i one can infer the four-momenta of the the decaying particle p at this vertex as the sum of its decay products and therefore its invariant mass m via

$$p = \sum_i p_i \quad , \quad m^2 = p^2 = \left(\sum_i p_i \right)^2.$$

Reconstructing the K_s , for example, is done using a common vertex of a π^+ and a π^- and their four-momenta:

$$p_{K_s} = p_{\pi^+} + p_{\pi^-} \quad , \quad m_{K_s}^2 = (p_{\pi^+} + p_{\pi^-})^2.$$

Similarly, all other particles of the decay mode up to the B^+ are reconstructed via their daughter particles, which may be reconstructed themselves (like the K_s). Due to detector resolution and efficiency, as well as the lifetime of the decaying particle, the resulting invariant mass shows a broadened distribution, not a discrete value. If sth. is hereinafter referred to as K_s mass, it is meant (if not stated otherwise) exactly this reconstructed mass distribution as derived above. The same is applied to the terms dimuon mass (reconstructed mass of a μ^+ and a μ^-) as well as B^+ mass (reconstructed throughout the whole decay chain). The important mass distributions for this analysis are shown in [Section 4.3](#) and will be further utilized throughout the main analysis in [Section 5](#).

4.2 Data set from the LHCb detector

The data used for the following analysis was taken at the LHCb detector in the years 2011 and 2012 and therefore covers the entire first run of the LHC. Since the collider

was operating at different center-of-mass energies \sqrt{s} in these two years, it should be mentioned that the data represents collisions at $\sqrt{s} = 7$ TeV for 2011 and $\sqrt{s} = 8$ TeV for 2012, respectively. The overall integrated luminosity over the two years adds up to about $\mathcal{L}_{int} = 3fb^{-1}$.

Since only a small part of the entire collected data is particularly interesting for the decay channel investigated in this thesis, a pre-selection is needed. The so-called stripping reduces the data into a set which is easier manageable to work with and still holds all information needed for this analysis. The cuts used in the stripping are summarized in [Table 4](#). No further HLT trigger requirements have been used. After the stripping process the remaining data sample then consists of about 2.7 million recorded events, out of which the interesting ones need to be found via the selection process of the further analysis.

Category	selection criteria
Cuts on B^+ meson	$IP_{\chi^2} < 16$
	Flight distance $\chi^2 > 121$
	Vertex $\chi^2/ndf < 8$
	$DIRA > 0.9999$
	$4600 \text{ MeV} < B \text{ mass} < 7000 \text{ MeV}$
Cuts on DiMuon	Vertex $\chi^2/ndf < 8$
	$\mu\mu\text{-mass} < 7100 \text{ MeV}$
Cuts on muons	$IP_{\chi^2}^2/ndf < 9$
	isMuon == TRUE
	$DLL_{\mu\pi} > -3$
Cuts on K_s^0	$\tau > 2 \text{ ps}$
	$467 \text{ MeV} < K_s^0\text{-mass} < 527 \text{ MeV}$
K/ π tracks from B vertex	$\min IP_{\chi^2}^2 > 9$

Table 4: Cut values used in the stripping process. isMuon labels a binary decision within the PID of the muon, τ the lifetime of a particle. All further variables will be explained in more detail in [Section 5.3.1](#).

4.3 Mass distributions after stripping

This section shall introduce the most important mass distributions, which will be further utilized during the main analysis in [Section 5](#). At this point, the distributions show the data after the stripping process without further cuts and improvements, which will be carried out at a later point.

The B^+ mass spectrum is shown in [Figure 7](#). Beside the large amount of combinatorial background in shape of a decreasing exponential function, one can see a clear structure peaking around 5280 MeV, which represents the decaying B^+ into the final state particles $\mu^+\mu^-\pi^+\pi^+\pi^-$. However, only a small part of the events of that peak represent the desired decay

$$B^+ \rightarrow \mu^+\mu^-K^*(\rightarrow K_s\pi^+) \quad (7)$$

with two non-resonant muons. The majority of events in this peak is from one of the two decay modes

$$B^+ \rightarrow J/\Psi(\rightarrow \mu^+\mu^-)K^*(\rightarrow K_s\pi^+), \quad (8)$$

$$B^+ \rightarrow \Psi(2S)(\rightarrow \mu^+\mu^-)K^*(\rightarrow K_s\pi^+), \quad (9)$$

which share the same final state particles as [Equation \(7\)](#), where in contrast the two muons originate not directly from the B meson but from a resonance - either a J/Ψ or a $\Psi(2S)$. This can be seen in more detail when plotting the spectrum of the dimuon

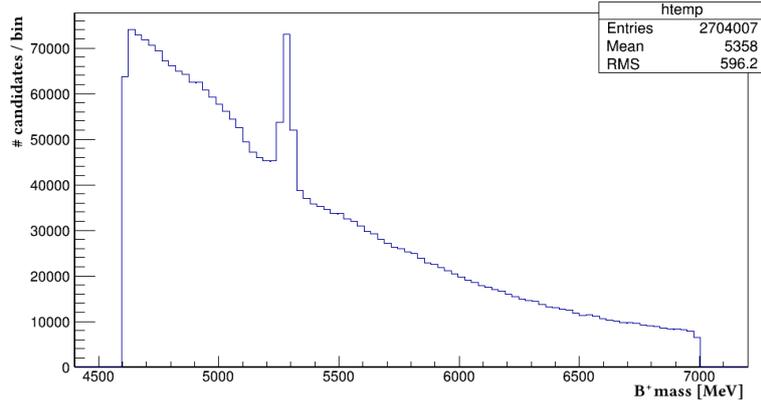


Figure 7: Reconstructed B mass distribution after stripping

mass, as shown in [Figure 8](#), again taken from the raw data. In addition to non-resonant decays which result in a broad dimuon mass distribution, there are two visible peaks close to 3100 MeV and 3700 MeV representing the J/Ψ and the $\Psi(2S)$ respectively ¹¹. These resonant decay modes need to be cut out later in the analysis.

¹¹The precise values for J/Ψ and $\Psi(2S)$ are 3096.916 ± 0.011 MeV and $3686.109^{+0.012}_{-0.014}$ MeV respectively [1]

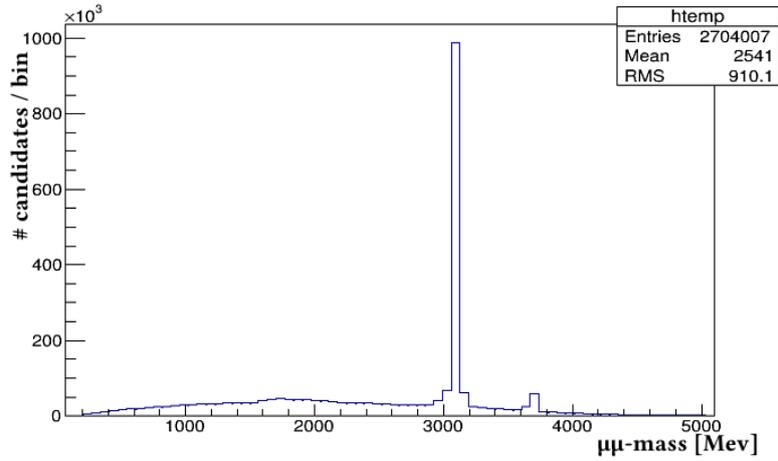


Figure 8: Reconstructed $\mu\mu$ -mass distribution after stripping.

4.4 Simulated data

For reasons that will be explained later in more detail, this analysis needs sets of simulated data-like events in addition to the real LHCb data. These simulated event samples were created via Monte Carlo (MC) simulation methods. Information about the decay mode and the number of events of the two MC samples used in this analysis is shown in [Table 5](#).

Simulated decay channel	Number of events
$B^+ \rightarrow \mu\mu K^{*+}$	20430
$B^+ \rightarrow J/\Psi(\rightarrow \mu\mu)K^{*+}$	7666

Table 5: Summary of the used MC samples.

5 Selection Process

This section contains the selection of the events for the decay $B^+ \rightarrow \mu\mu K^{*+}$ as well as the reference channel and the used MC samples. This section starts with a short description of the analysis strategy, followed by the detailed description of these analysis steps.

5.1 Summary of the analysis strategy

The goal of this thesis is to derive the branching ratio \mathcal{B} of the B^+ meson decay mode

$$B^+ \rightarrow \mu\mu K^{*+},$$

hereinafter called decay channel.

However, this value can only be an estimate. It will include statistical errors, however systematic influences and uncertainties will not be covered in full complexity but nevertheless will be discussed in [Section 7.2](#) as far as they could have been identified. The branching ratio will be initially determined as relative value to the decay mode

$$B^+ \rightarrow J/\Psi(\rightarrow \mu\mu)K^{*+},$$

hereinafter called reference channel.

This relative branching ratio can then be calculated using

$$\frac{\mathcal{B}(B^+ \rightarrow \mu\mu K^{*+})}{\mathcal{B}(B^+ \rightarrow J/\Psi(\rightarrow \mu\mu)K^{*+})} = \frac{N(B^+ \rightarrow \mu\mu K^{*+})}{N(B^+ \rightarrow J/\Psi(\rightarrow \mu\mu)K^{*+})} \cdot \varepsilon \cdot \gamma. \quad (10)$$

$N(X \rightarrow Y)$ represents the number of selected events of the decay mode $X \rightarrow Y$, which will be derived (with statistical errors) out of fits in the following subsections. ε is the relative efficiency of the selection process, γ the relation of the Monte Carlo generation acceptances, which will be further explained and calculated in [Section 6](#). Finally, using the well known value (literature value) for the reference channel the absolute value for the decay channels branching ratio will be determined.

5.2 Selection of $\mu\mu$ -resonances

The proper selection and identification of all decays via a $\mu\mu$ -resonance is important for mainly two reasons. Firstly, as the expected branching ration for the decay channel

$B^+ \rightarrow \mu\mu K^{*+}$ is of about three orders of magnitude smaller than the resonant decay via a J/Ψ , one must ensure that no resonant decay events remain within the selected data, which could severely distort the results. The method and the results are discussed in [Section 5.2.2](#). Secondly, the amount of events decaying as $B^+ \rightarrow J/\Psi(\rightarrow \mu\mu)K^{*+}$ in the used data set needs to be known, as the branching ratio will be determined as value relative to this very decay channel, as indicated in [Equation \(10\)](#). This will be further treated in the following [Section 5.2.1](#).

5.2.1 Number of $B^+ \rightarrow J/\Psi(\rightarrow \mu\mu)K^{*+}$ events

To achieve this, the dimuon mass spectrum - as introduced for the first time in [Figure 8](#) - will be used. Further constraints after the stripping are an acceptance window for the K^* mass of ± 150 MeV around the reference value obtained from PDG [7], resulting in the window [742 MeV, 1042 MeV]. The invariant mass region around the peaking J/Ψ -resonance is shown for the resonant MC sample in [Figure 9](#) and for data in [Figure 10](#). On both distributions a fit has been applied, whose details shall be further discussed at this point:

At a first look, the shape of the peaking structure resembles a Gaussian, but shows a tail at its left flank (better visible in the MC distribution of [Figure 9](#) in the area of 3040 – 3080 MeV). The physical reason for this tail are radiation losses of the muons in the detector due to Bremsstrahlung effects, which *lowers* the reconstructed $\mu\mu$ -mass. This can be modeled by using the so-called Crystal-Ball (CB) function ¹², a probability density function which modifies a Gaussian by a tail below a certain threshold. The mathematical expression of the four-parametric CB function, here as a function of the mass m is

$$CB(m; \alpha, n, \mu, \sigma) = N \cdot \begin{cases} e^{-\frac{(m-\mu)^2}{2\sigma^2}}, & \text{for } \frac{m-\mu}{\sigma} > -\alpha \\ A \cdot (B - \frac{m-\mu}{\sigma})^{-n}, & \text{for } \frac{m-\mu}{\sigma} \leq -\alpha \end{cases}, \quad (11)$$

where by definition

$$A = \left(\frac{n}{|\alpha|}\right)^n \cdot e^{\frac{|\alpha|^2}{2}}, \quad B = \frac{n}{|\alpha|} - |\alpha|.$$

¹²Named after the Crystal Ball collaboration, which operated the eponymous detector at the Stanford Linear Accelerator Center (SLAC) starting in 1979.

As in a Gaussian, the parameters μ and σ label the position of the peak and its Gaussian width, whereas n and α are used to characterize the form and the threshold value of the tail, respectively. N is a normalization constant.

For the total fit function, a sum of two CB functions sharing the same value for μ is used for the signal part. The background part¹³ can be equally well described by either an exponential function whose decay constant in this case would be very close to zero, or (therefore by a good approximation) by a constant value c . Here, the latter option was chosen.

The overall fit function F is described by

$$F = N_{sig} \cdot (f_1 \cdot CB_1 + (1 - f_1) \cdot CB_2) + N_{bgr} \cdot c, \quad (12)$$

where f_1 is the first CB functions fraction of the signal part, N_{sig} and N_{bgr} label the amount of signal and background events, respectively. It should be mentioned that, even though the data points are presented in bins for graphical reasons, this fit is carried out on the unbinned data.

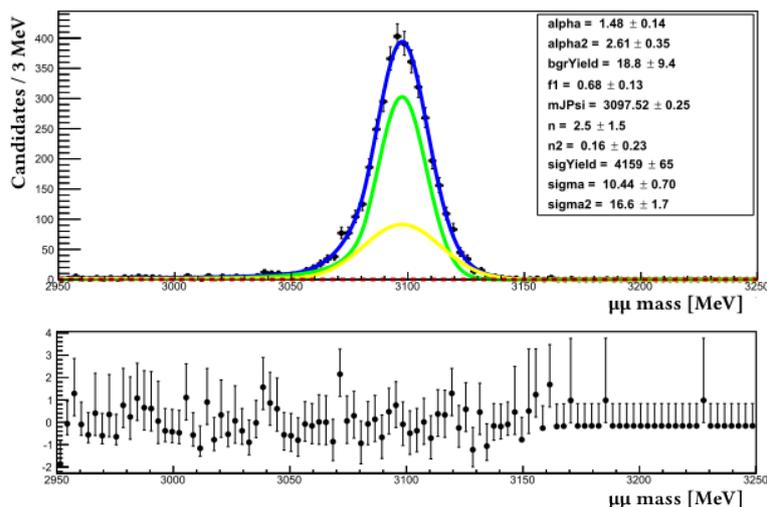


Figure 9: Top: Fit of the function F (blue) to the resonant MC sample. The fit function is composed of two Crystal Ball functions (green and yellow) plus a constant value for the background (red).

Bottom: Pull histogram, showing the deviation of the fit from the data in units of standard deviations. The fit function describes the distribution well.

The same applies to all following invariant mass fits in this thesis. To avoid the possibility of the tail describing not only signal, but also a part of the background events, the

¹³The term *background* is used in this specific case as contrast to the term *signal* for the peaking structure. It consists of non-resonant events and therefore holds the later identified desired non-resonant *signal* events

parameters n and α get determined in a first step by fitting F to a MC sample of the pure decay $B^+ \rightarrow J/\Psi K^{*+}$ (Figure 9) and fixing these values in the fit to the real data. As can be seen in the pull histogram of Figure 10, the fit function shows deviations from the data points as can be seen in the pull histogram, especially in the area of 3050 - 3080 MeV. Nonetheless, the overall fit quality is sufficient for the further analysis, therefore the fit model F will get used again at later points. The choice of this fit model introduces systematic uncertainties, which will be further discussed in Section 7. For Figure 10, the fit procedure yields the value

$$N_{sig}(B^+ \rightarrow J/\Psi K^{*+}) = 88175 \pm 310.$$

events of type $B^+ \rightarrow J/\Psi K^{*+}$ within the data sample.

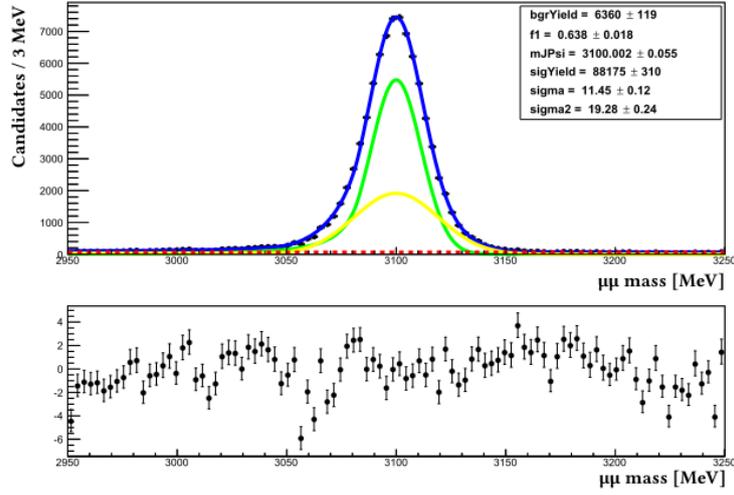


Figure 10: Top: Fit of the function F to the data set. The parameters n and α get fixed to the values determined from the MC sample in Figure 9, also the identical colors get used. Bottom: Pull histogram. Here, the fit function does not optimally describe the data in the area of 3050 - 3080 MeV.

5.2.2 Values for J/Ψ and $\Psi(2S)$ vetoes

To sort out all events containing muons originating from a resonance, the two regions around the invariant dimuon masses showing the peaking structures of the J/Ψ and $\Psi(2S)$ will be cut out of the data set. The values for these veto cuts must be determined carefully to ensure the exclusion of all resonant events in the further used data set. Within the excluded mass windows there may also be a few signal events with non-resonant muons, which get lost for the further analysis by placing these cuts.

Therefore the goal should be to narrow the vetoed mass windows as much as possible, to maintain the maximum amount of possibly interesting events, while getting rid of all resonant events. To get the values for the veto cuts, a fit of the $\mu\mu$ -mass distribution

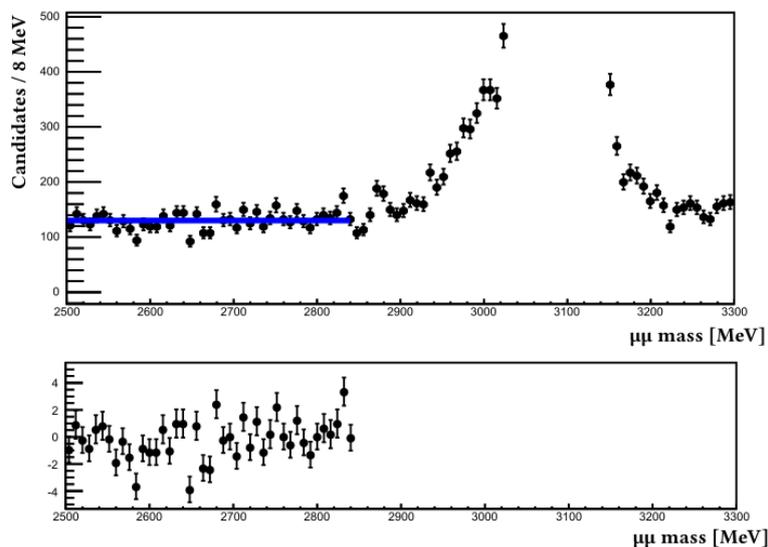


Figure 11: Top: Fit of a constant to the background of the peaking J/Ψ . This fit is done multiply times, while varying the right border value to get the best possible cut value.

Bottom: Pull histogram in units of standard deviations σ . The first deviation of more than 3σ can be seen between 2820 and 2840 MeV.

of non-resonant muons will get used. The chosen fit function is a constant line, which is a good approximation to the invariant mass distributions in case of a sufficiently small range (here about 300 MeV). The used procedure is exemplary shown for the left value of the J/Ψ veto window in the following and gets applied analogously for the remaining three cut values. As shown in Figure 11 the background gets fitted by a constant which is repeated multiple times with increasing fit range range (in this case meaning increasing right border value). At some value, the deviation of the data (lower part of Figure 11) from this constant value exceeds a selected border value, here chosen to be 3σ ¹⁴. This bin is then chosen as the first one¹⁵ to be excluded by the veto cut. The values for the veto windows are summarized in Table 6.

¹⁴In case of the value for the right border of the $\Psi(2S)$ veto window it is chosen to use a bin with a deviation of 2.6σ instead.

¹⁵For the right side of the peak the last one, respectively.

Description	veto window [MeV]
J/Ψ veto	[2840, 3210]
$\Psi(2S)$ veto	[3560, 3800]

Table 6: Summary of the used veto cuts on the $\mu\mu$ -mass

5.3 Reduction of combinatorial background

Having determined the resonance veto windows, one is able to eliminate all unwanted events of the peaking structure in the reconstructed B mass distribution, so far only shown in raw form in [Figure 7](#). The remainders are a small amount of nonresonant events overlaid by a much larger amount of combinatorial background. The next goal is therefore to separate the signal from these combinatorial background events, from now on referred to simply as background. Achieving a good separation between signal and background needs knowledge about the different experimental signatures of these two kinds. For this purpose, all event reconstruction variables used to realize this separation shall be defined at this point.

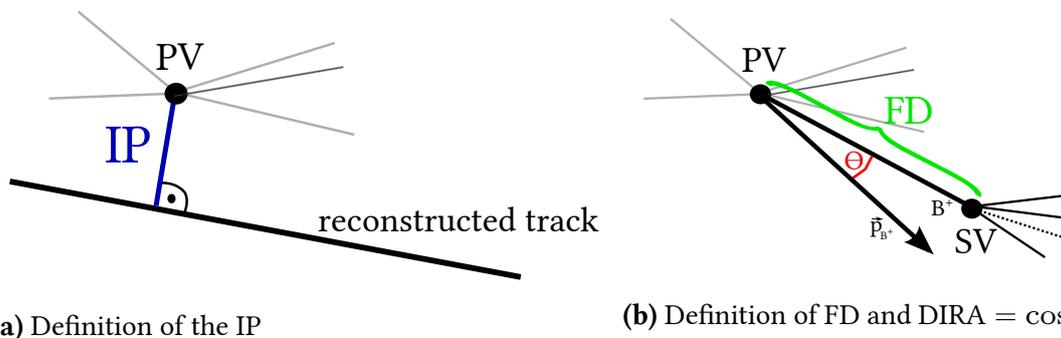
5.3.1 Definition of the used variables

Here, all variables that are used to separate the signal events from the background will be introduced. First, the so-called impact parameter (IP) provides information about the originating vertex of a particle. It is defined as the minimal distance of the reconstructed track to the primary vertex (PV), see also [Figure 12a](#). The mother particle (here B^+) which originates from the PV should therefore be recognizable by a small IP, whereas the daughter particles tend to have a higher IP. Combinatorial background has sources in many different directions, therefore can have very different values for the IP. All too large values are with high probability due to background events. The χ^2 of the impact parameter is defined analogously.

In [Figure 12b](#), the so-called direction angle (DIRA) and the flight distance (FD) are shown schematically. The DIRA is defined as the cosine of the angle θ between the path of a particle (the direct line between creation vertex and decay vertex) and its reconstructed momentum vector. Assuming a good reconstruction of all participating particles, vertices and particle momenta this angle should be small, hence the DIRA close to one. The flight distance simply is the distance between creation and decay vertex. This variable is of course correlated to the lifetime τ of a particle.

Further used variables include the transverse momentum p_T , that is the momentum in the plane perpendicular to the beam axis (x-y-plane) and the vertex χ^2 , which is a

measure of the accuracy of the vertex reconstruction. In addition, as already mentioned in [Section 3](#) the DLL variables get used for particle identification.



(a) Definition of the IP

(b) Definition of FD and DIRA = $\cos \theta$

Figure 12: Visualization of the variables impact parameter (IP), flight distance (FD) and direction angle (DIRA).

5.3.2 Multivariate Analysis

To distinguish between signal and background events, the variables introduced above need to be used. For example, the distributions of the mother particles (B^+) Vertex χ^2 or IP_{χ^2} can be used to achieve a more precise vertex reconstruction. For a signal event, it is expected to have small values for both of these variables, while for the background a much more widespread range is possible. The angle θ is also to be expected small for signal events, resulting in $DIRA \approx 1$. For combinatorial background, lower values of the DIRA are also possible.

The transverse momenta of daughter particles give information about the invariant mass of their mother. Since this analysis requires a reconstructed B meson, the sum of the momenta of its daughter particles is restricted to an area around the invariant B mass. Also, the track reconstruction via IP_{χ^2} can be used to prefer daughter particles originating from the SV, while suppressing the ones from the PV. To further ensure the correct SV is chosen, one can have a look at the variable (B^+) Vertex χ^2 . The majority of signal events should have been reconstructed with lower values for this variable, while combinatorial background can also have much higher values. To see examples of the distributions just described, one can look ahead at [Figure 14](#), containing some of the distributions that will later be used in this section.

Having therefore identified the key variables to separate data from background events, one needs to place cuts on these variables. As a first attempt, one could now for example place simple linear cuts on each one of these variables manually. This would imply the assumption that all used variables are independent of each other, otherwise more

complicated non-linear cuts would be more efficient. Since the correlations between variables can be quite different and are not necessarily known a priori, placing all the right cuts can be a very complicated matter. In addition, the usage of a variety of cuts further and further decreases the amount of passing events. Especially for rare decays, like it is the case for $B^+ \rightarrow \mu\mu K^{*+}$, this can reduce the remaining events significantly.

A solution to this problem is provided by the Toolkit for Multivariate Data Analysis with ROOT (TMVA) [13]. TMVA comes with different implemented methods and algorithms suited for a variety of analysis strategies. The aim of this toolkit is to use the information of the distributions of the identified separation variables for each signal and background events individually to classify all recorded events by a single variable. Hence only a single cut would be needed instead of multiple different ones. For this thesis, the so-called Boosted Decision Tree (BDT) ¹⁶ algorithm gets used. Before the algorithm gets capable of classifying the real data, it must be trained how to separate between signal and background events. Therefore, it is provided with two samples of a priori distinguishable signal and background events, respectively. The events in these samples get split in two parts. The first one gets used for this training phase, the second part is hold back for the subsequent testing phase. A single decision tree, as can be schematically seen in Figure 13, is a sequence of binary decisions made on criteria of the separation variables. The sequence terminates in case of an event being

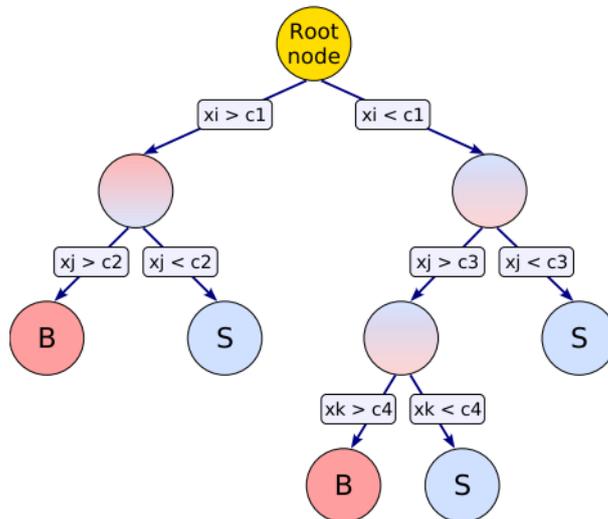
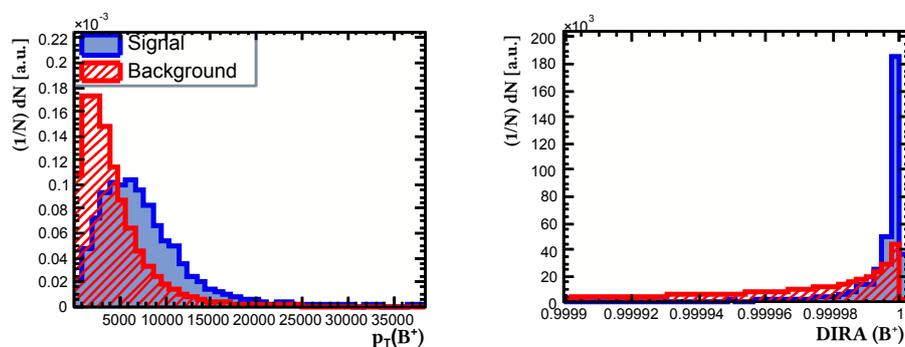


Figure 13: Schematic illustration of a decision tree [13].

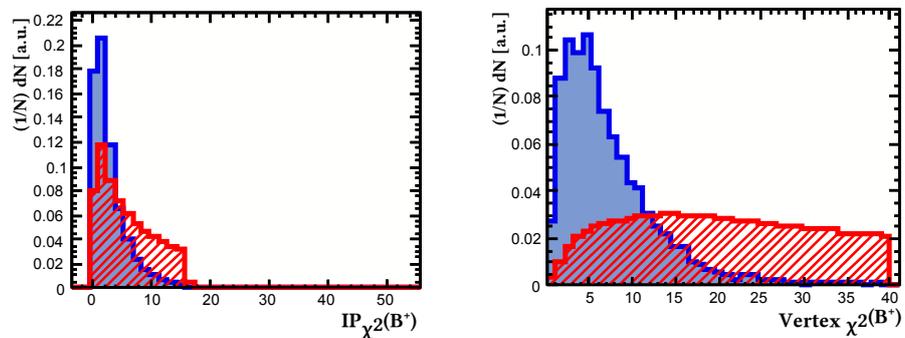
¹⁶The BDT algorithm itself comes in different variants. Here, the *Adaptive Boost* technology is used. More information in [13].

classified as signal-like (S) or background-like (B) or if the statistics get too low after several decisions. The output of the tree allocates the value of $+1$ for signal-like, and -1 for background-like events. To further improve the algorithm, the (adaptive) boost technology gets used. It reruns the decision tree and all events that have been falsely classified get assigned a larger weight to account for this wrong estimate. This process gets repeated iteratively. This way, a large amount of decision trees gets created. The final BDT response for each event is then calculated as the sum over all single tree responses normalized by the number of decision trees and results in a distribution between -1 and $+1$. After the training is complete, the TMVA will run a testing phase with the remaining events. This testing phase is necessary to check for so-called overtraining effects. An overtraining can be possible if the amount of degrees of freedom of the used BDT algorithm gets too high, which has to be avoided.

In this particular analysis, a MC sample of the decay channel is used for the signal



(a) Transverse momentum of the B^+ . (b) DIRA of the B^+ .



(c) $IP\chi^2$ of the B^+ . Due to the stripping criteria only values up to 16 are within the data sample. (d) The distribution for χ^2 of the vertex reconstruction for the decay vertex of the B^+ .

Figure 14: Selection of four of the most important distributions of variables used for the BDT algorithm. The signal distributions are shown in blue, while the combinatorial background ones are in red. All further variables used are listed in [Table 7](#).

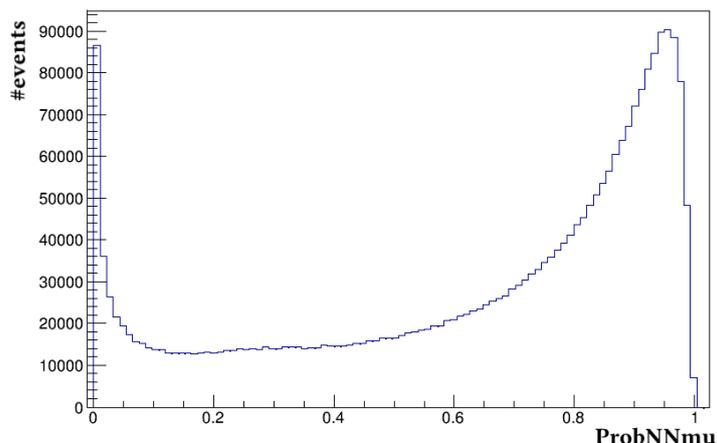


Figure 15: Distribution of the variable ProbNNmu, which quantifies the likeliness of a particle reconstructed as muon to be of its type. For low values there is a large amount of events corresponding to falsely reconstructed muons (fake muons).

component. The background sample consists of all sideband events of the real data set with a reconstructed B mass > 5700 MeV, which will not be used in the rest of the analysis. A list of all separation variables used is provided in Table 7. In Figure 14 the signal and background distributions of a selection of these variables are exemplary shown. Similar distributions of all the other variables of Table 7 exist, whereas these four are exemplary chosen because of their stronger and clearly visible discrimination power. Before starting the BDT, two more cuts have been placed. The first one is again an acceptance window for the K^* mass of ± 150 MeV around the literature value, the second one is made on a variable used in the particle identification (PID) process. It quantifies the likeliness of a particle identified as a muon to be a real muon¹⁷. Exemplarily, its distribution is shown in Figure 15 (for μ^+ only). As can be seen in this variable's distribution a significant amount of fake muons, being possibly hadrons and therefore mainly pions, is present (located at low values of the muon identification variable). These events are sorted out by a cut requiring ProbNNmu to be greater than 0.2 for reconstructed μ^+ as well as μ^- .

particle	B^+	K_s^0	π^\pm
variables	p_T , FD, DIRA vertex χ^2 , IP, IP χ^2	p_T , IP χ^2	p_T , IP χ^2

Table 7: Summary of all variables used within the BDT analysis.

¹⁷The variable is labeled as ProbNNmu.

The results of the BDT training and testing phase for these data and MC samples is shown in Figure 16. One can clearly see the discrimination power of the BDT algorithm, as events that have been classified as combinatorial background get assigned much lower BDT response values as those classified as signal. However, the discrimination power is not strong enough to totally separate the two distributions from each other, there remains an overlap. The plot shows also good agreement of both distributions for training (points with error bars) and testing phase (filled areas), indicating that there is no large effect of overtraining present.

The trained decision trees and their weight are now ready to be used to classify the real data. After the application, a proper cut value on the BDT response needs to be found to maintain most of the signal while eliminating the majority of background. To quantify this, the remaining events after the cut should maximize the function $\frac{S}{\sqrt{S+B}}$, where S and B label the number of signal respectively background events. The TMVA analysis would internally calculate the best cut value if provided with an estimate of the number of expected signal and background events. Since the number of signal events is not known at this point, the optimal cut value will be determined manually in Section 5.4.

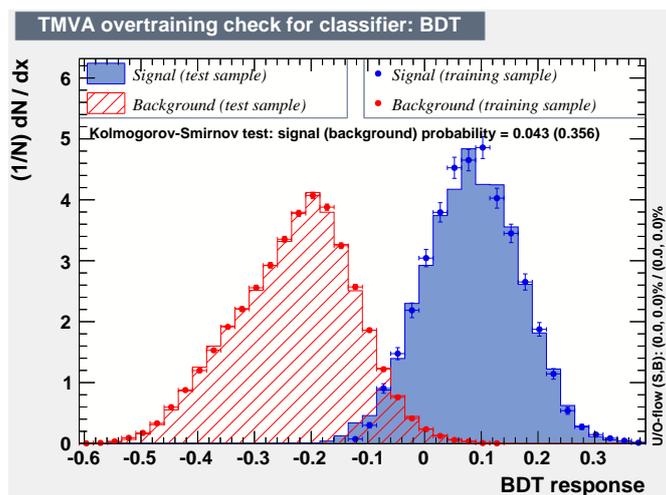


Figure 16: Response of the BDT training and testing phase. One can see the achieved separation for classifying the events as signal (blue) or background (red). The filled areas represent the testing phase, while the distributions marked by single points are the results of the training.

5.4 Results of the selection process

After having assigned a BDT response value to each data event, the optimal value for a cut on this BDT response needs to be found. To achieve this, the reconstructed B mass

distribution gets plotted and fitted multiple times, each with a different BDT response cut value. The chosen fit function consists of a Gaussian¹⁸ for the peaking B^+ and an exponential function for the remaining background. In this way, the number of remaining signal (N_{sig}) as well as background (N_{bgr}) events can be determined. As already mentioned in Section 5.3.2, the optimal cut value on the BDT response is chosen as the one maximizing $\frac{N_{sig}}{\sqrt{N_{sig}+N_{bgr}}}$. In this process, N_{bgr} is calculated in a 5σ area

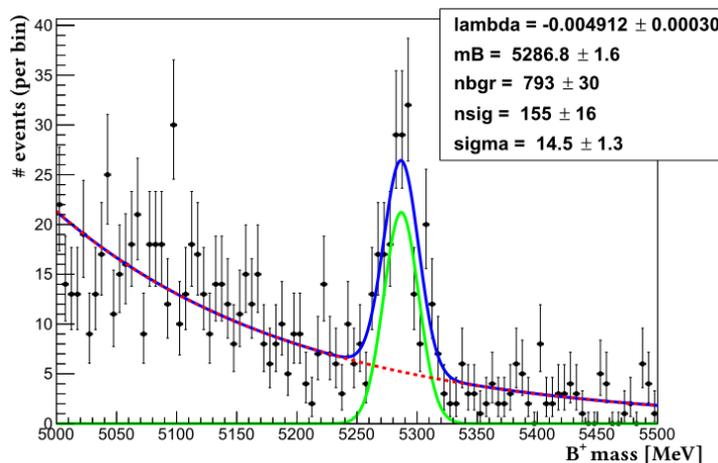


Figure 17: The resulting distribution for the reconstructed B mass after all placed cuts, including the one on the BDT response. The remaining events show a clear peak of B mesons on top of the remaining exponential background. The total fit function (blue) is composed of a Gaussian for the signal peak (green) and an exponential for the background (red). Parameters are shown in Table 8.

Parameter	Value
μ	5286.8 ± 1.6
σ	14.5 ± 1.3
λ	$(-4.912 \pm 0.300) \times 10^{-3}$
N_{sig}	155 ± 16
N_{bgr}	793 ± 30

Table 8: Parameters of the fit in Figure 17. The value for N_{bgr} is accumulated over the entire fit range, not the 5σ area.

around the B^+ peak. The best cut value on the BDT response is found to be -0.03 , the corresponding plot of the remaining data events is shown in Figure 17, its fit parameters

¹⁸For such low statistics, a Crystal Ball is not suited, since the tail parameters could vary a lot and tend to describe part of the background.

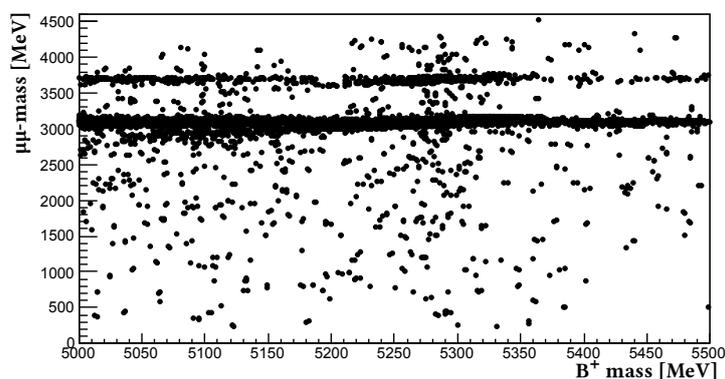
are again given in [Table 8](#). As a result, the amount of events of the decay channel $B^+ \rightarrow \mu\mu K^{*+}$ with its corresponding statistical error is determined as

$$N_{sig}(B^+ \rightarrow \mu\mu K^{*+}) = 155 \pm 16,$$

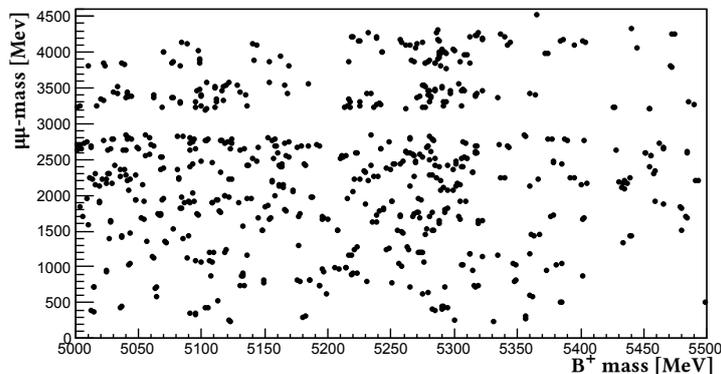
with a statistical significance of

$$s_{stat} = \frac{N_{sig}}{\sqrt{N_{sig} + N_{bgr}}} = 8.9\sigma.$$

To further visualize the selected events and the application of the veto cuts, [Figure 18](#)



(a) Only BDT response cut gets applied. The two horizontal accumulations correspond to the $\Psi(2S)$ -resonance (around 3700 MeV) and the J/Ψ -resonance (around 3100 MeV).



(b) In addition to the BDT response cut, the resonance vetoes get applied. As a result, both accumulations have been sorted out from the data set.

Figure 18: 2-dimensional plots showing both B^+ -mass and $\mu\mu$ -mass of the remaining events after the given cuts.

shows two-dimensional plots of the B^+ -mass vs. the $\mu\mu$ -mass. In [Figure 18a](#), all events that remain after only the BDT response cut are shown. One can clearly see the two clusters of events representing the J/Ψ and $\Psi(2S)$ resonances. These are sorted out by applying the veto cuts from [Table 6](#), resulting in the distribution shown in [Figure 18b](#). All events remaining within this plot correspond to the ones displayed and fitted in [Figure 17](#).

In addition to the data, the identical BDT cut value is applied to both MC samples, which will be needed for the determination of the selection efficiency following in [Section 6](#).

6 Determination of the branching ratio

In this section, the selected data and MC events shall be used to determine the branching ratio of the decay $B^+ \rightarrow \mu\mu K^{*+}$ relative to the reference channel. Finally, the literature value of the reference channels branching ratio will be used to calculate the absolute value for $\mathcal{B}(B^+ \rightarrow \mu\mu K^{*+})$.

At this point the formula used for the determination, already mentioned as [Equation \(10\)](#) in [Section 5.1](#) shall be discussed in more detail. It was introduced as

$$\frac{\mathcal{B}(B^+ \rightarrow \mu\mu K^{*+})}{\mathcal{B}(B^+ \rightarrow J/\Psi(\rightarrow \mu\mu)K^{*+})} = \frac{N(B^+ \rightarrow \mu\mu K^{*+})}{N(B^+ \rightarrow J/\Psi(\rightarrow \mu\mu)K^{*+})} \cdot \varepsilon \cdot \gamma. \quad (13)$$

Here, $N(B^+ \rightarrow \mu\mu K^{*+})$ labels the amount of events of the decay channel that could be reconstructed. Its value has been determined at the end of [Section 5.4](#). Similarly, $N(B^+ \rightarrow J/\Psi(\rightarrow \mu\mu)K^{*+})$ is the amount of reference channel events, which will be determined later within this section. Further, it has to be taken into account that the actual efficiency of the event selections for the two channels are not equal. This is corrected via the relative selection efficiency ε ,

$$\varepsilon = \frac{\varepsilon_{J/\Psi K^*}}{\varepsilon_{\mu\mu K^*}}, \quad (14)$$

the quotient of the two channels individual selection efficiencies

$$\varepsilon_{J/\Psi K^*} = \frac{N_{MC}^{sel}(B^+ \rightarrow J/\Psi K^{*+})}{N_{MC}^{gen}(B^+ \rightarrow J/\Psi K^{*+})}, \quad (15)$$

$$\varepsilon_{\mu\mu K^*} = \frac{N_{MC}^{sel}(B^+ \rightarrow \mu\mu K^{*+})}{N_{MC}^{gen}(B^+ \rightarrow \mu\mu K^{*+})}. \quad (16)$$

These individual selection efficiencies will get derived out of the MC samples, as indicated by the lower index MC. The upper indices refer to the amount of MC events generated (the sample size) respectively selected (passing the selection procedure described in [Section 5](#)). Also, the generation processes of the two MC samples use slightly different conditions for the detectors regions of acceptance ¹⁹, which gets accounted

¹⁹Since the LHCb detector covers only a small part of the full solid angle, the simulations are restricted to the angular acceptance of the detector, as this saves a lot of calculation time. However, the two decay modes use different simulation models, as simulating the reference channel needs to account for the appearance of a J/Ψ resonance. Within these different simulation packages the angular acceptances might differ, which is why this correction is used.

for by the relative generation acceptance (here called γ)

$$\gamma = \frac{\gamma_{J/\Psi K^*}}{\gamma_{\mu\mu K^*}}, \quad (17)$$

again defined by the decay modes individual acceptance efficiencies, which are summarized together with the other determined values in [Table 11](#).

Number of resonant events After determining the best BDT cut value in [Section 5.4](#) and applying it (as well as the other cuts) on the data set, the amount of nonresonant events (decay channel) $N(B^+ \rightarrow \mu\mu K^{*+})$ has been determined. Calculating the amount of resonant events (reference channel) can be done similarly. The cuts on the K^* mass of ± 150 MeV around the PDG-value as well as the muon identification are used identically to the nonresonant channel. To ensure retaining all resonant J/Ψ events, the $\mu\mu$ -mass window is set to the values determined for the J/Ψ veto cut in [Table 6](#), excluding all $\mu\mu$ -masses outside this range. The distribution of all events

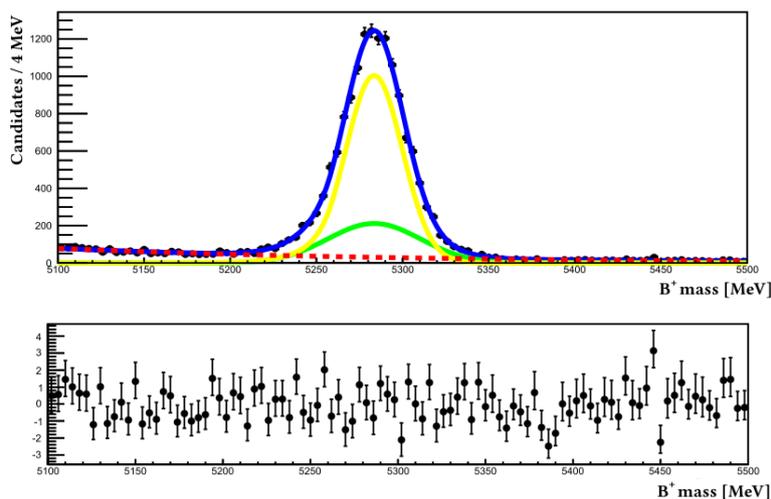


Figure 19: Top: Fitted B mass distribution of all resonant decays after all cuts and restrictions. The total fit function (blue) consists of two Crystal Ball functions (green, yellow) and an exponential for the background (red). Parameters are shown in [Table 9](#). Bottom: Pull histogram, showing a good agreement between fit and data.

passing these criteria as well as the BDT cut are shown in [Figure 19](#). Again, the fit function consists of a double Crystal Ball (compare to F in [Equation \(12\)](#)) for the signal part, yet the background is described by a decreasing exponential function $\exp^{-\lambda m}$. The parameters of the fit, including the desired value for $N(B^+ \rightarrow J/\Psi K^{*+})$ are displayed in [Table 9](#).

Parameter	Value
μ	5283.51 ± 0.16
σ_1	27.1 ± 1.8
σ_2	16.36 ± 0.19
α_1	3.80 ± 0.15
α_2	1.50 ± 0.56
n_1	0.1 ± 9.7
n_2	9.9 ± 7.9
λ	$(-5.209 \pm 0.240) \times 10^{-3}$
f_1^{CB}	0.253 ± 0.027
N_{sig}	14182 ± 194
N_{bgr}	3286 ± 185

Table 9: Parameters of the fit function for [Figure 19](#).

Number of selected MC events The last missing values to ascertain the relative branching ratio are the amount of MC events passing the selection criteria, both needed to calculate the relative selection efficiency ε ([Equation \(14\)](#)). Therefore the identical cuts and restrictions made for nonresonant and resonant data are applied on their respective MC counterpart. The results are displayed in [Figure 20](#) and the related [Table 10](#).

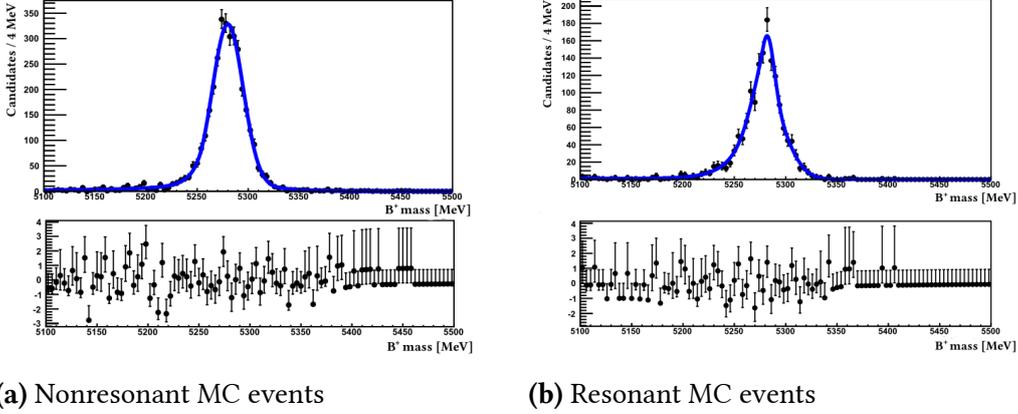


Figure 20: Top: Fits to the B mass distributions of the remaining events of the MC samples. The fit function is identical as described for [Figure 19](#), yet only the overall fit function drawn for better visibility. Bottom: Pull histograms. Again, the left tails are not described optimally, for the resonant events there is also some deviation at the center of the peak.

Parameter	Value	Parameter	Value
μ	5280.08 ± 0.0034	μ	5281.9 ± 1.0
σ_1	14.99 ± 0.35	σ_1	9.3 ± 1.5
σ_2	40.0 ± 8.0	σ_2	18.5 ± 1.1
α_1	1.60 ± 0.17	α_1	0.55 ± 0.22
α_2	0.53 ± 0.12	α_2	1.20 ± 0.19
λ	$(-1.75 \pm 3.40) \times 10^{-3}$	λ	$(-7.20 \pm 1.80) \times 10^{-3}$
f_1^{CB}	0.882 ± 0.021	f_1^{CB}	0.324 ± 0.099
N_{sig}	3434 ± 66	N_{sig}	1526 ± 40
N_{bgr}	46 ± 30	N_{bgr}	42.4 ± 9.6

(a) Nonresonant MC events.

(b) Resonant MC events.

Table 10: Parameters for the fits of [Figure 20](#). The values for n_1 and n_2 have been fixed.

Sample	Quantity	Value
MC nonresonant	$N_{MC}^{gen}(B^+ \rightarrow \mu\mu K^{*+})$	20430
	$N_{MC}^{sel}(B^+ \rightarrow \mu\mu K^{*+})$	3434 ± 66
	$\gamma_{\mu\mu K^*}$	0.0711
MC resonant	$N_{MC}^{gen}(B^+ \rightarrow J/\Psi(\rightarrow \mu\mu)K^{*+})$	7666
	$N_{MC}^{sel}(B^+ \rightarrow J/\Psi(\rightarrow \mu\mu)K^{*+})$	1526 ± 40
	$\gamma_{J/\Psi K^*}$	0.0674
Data	$N_{sig}(B^+ \rightarrow \mu\mu K^{*+})$	155 ± 16
	$N_{sig}(B^+ \rightarrow J/\Psi(\rightarrow \mu\mu)K^{*+})$	14182 ± 194

Table 11: Summary of all values needed for the calculation of the branching ratio.

The Branching Ratio Having now available all required values, which are summarized in [Table 11](#), the efficiencies are calculated as

$$\begin{aligned}\varepsilon_{J/\Psi K^*} &= (19.9 \pm 0.5)\%, \\ \varepsilon_{\mu\mu K^*} &= (16.8 \pm 0.3)\%.\end{aligned}$$

The values of the efficiencies differ for the two decay modes. A possible reason for this could be the threshold value for transverse momentum/ energy, below which the events do not pass the trigger system. For the decaying resonances, the daughter particles get enough transverse energy to fulfill the requirements. But since for the nonresonant decays a much wider range for the dimuon mass is acceptable, especially in case of low values for $m_{\mu\mu}$ some events may not pass the trigger, and are therefore lost. Also for high values of the $\mu\mu$ -mass, some tracks may leave the covered solid angle of the detector system and therefore do not contribute.

Inserting the results for the efficiencies according to [Equation \(13\)](#), the branching ratio of the decay mode $B^+ \rightarrow \mu\mu K^{*+}$ relative to the reference channel $B^+ \rightarrow J/\Psi(\rightarrow \mu\mu)K^{*+}$, both with the subsequent decay $K^{*+} \rightarrow K_s^0\pi^+$, is calculated as

$$\frac{\mathcal{B}(B^+ \rightarrow \mu\mu K^{*+})}{\mathcal{B}(B^+ \rightarrow J/\Psi(\rightarrow \mu\mu)K^{*+})} = (1.23 \pm 0.13_{stat}) \times 10^{-3}.$$

Utilizing the known value for the branching ratio of the reference channel

$$\mathcal{B}(B^+ \rightarrow J/\Psi K^{*+}) = (1.44 \pm 0.08) \times 10^{-3} \text{ [1]}.$$

as well as the subsequent decay mode

$$\mathcal{B}(J/\Psi \rightarrow \mu\mu) = (5.93 \pm 0.06) \times 10^{-2} \text{ [1]}.$$

the absolute branching ratio of the decay $B^+ \rightarrow \mu\mu K^{*+}$ is calculated to

$$\mathcal{B}(B^+ \rightarrow \mu\mu K^{*+}) = (1.05 \pm 0.11_{stat} \pm 0.06_{norm}) \times 10^{-6}.$$

Hereby, the uncertainty of the reference branching ratio is accounted by introducing an additional error. A discussion of the statistical error as well as some systematic uncertainties will follow in [Section 7](#).

The comparison to the value of previous measurements, which is $\mathcal{B}(B^+ \rightarrow \mu\mu K^{*+}) = (1.12 \pm 0.15) \times 10^{-6}$ [1] indicates that in fact the relative statistical error has decreased due to a higher event count. The value determined in this thesis is 6% lower than in the previous measurement, and also the relative statistical ²⁰ error has decreased. However, both measurements are still consistent with each other, as they have the respective other value within their 1σ -uncertainty.

²⁰The value from [1] includes both statistical and systematic errors, yet the statistical one is dominant.

7 Discussion of uncertainties

In this section the measured statistical errors will be set into context, followed by a qualitative discussion of identified and further possible systematic uncertainties.

7.1 Size and effects of statistical errors

The statistical uncertainty of $\mathcal{B}(B^+ \rightarrow \mu\mu K^{*+})$ in this thesis has been subdivided into the purely statistical error due to the reconstructed number of events as well as an error that occurs because of the uncertainties of the literature value of the reference channel, hence called normalization error. Their absolute values are 1.1×10^{-7} and 0.6×10^{-7} , respectively. Therefore the relative statistical error is given as

$$\frac{\Delta\mathcal{B}_{stat}}{\mathcal{B}} \approx 10.9\%,$$

for the normalization error this value is about the half. Contributions to the statistical error are made by all values of [Table 11](#) having errors, while its majority is clearly due to the error of the determined amount of events in the decay channel $N_{sig}(B^+ \rightarrow \mu\mu K^{*+})$, which contributes about 95% of the entire statistical error. Hence this is the error that has to be decreased in order to measure the branching ratio with higher accuracy. Since the size of this error is of statistical nature, it can be mainly decreased by accumulating more data ²¹, but also further improving the selection process could possibly retain a few more events that have been cut away within the treatment in this thesis. Since the veto cuts are necessary and cannot change by a larger amount, the main process available for improvement is the BDT analysis. Although the main separation variables have been identified some additional variables not considered here could improve the discrimination power. The size of the statistical error will therefore decrease with future measurements on this channel as LHCb (and possible successor experiments) will collect more statistics.

7.2 Systematic uncertainties

As the realization of this analysis was limited in time, a comprehensive quantification or even a removal of some systematical uncertainties was not possible. Nonetheless

²¹The same goes for the normalization error, as the utilized value was mainly statistically limited.

this section shall give a brief overview about (un-/)identified systematic effects that have or could have influenced the results. They are summarized into three categories treated in the following paragraphs.

Veto of $\mu\mu$ -resonances In this analysis the largest amount of events containing the required final state particles represented muons originating from a resonance and therefore had to be cut out. In doing so it was payed attention to cut off all resonant events, yet still a very tiny portion could remain, especially if the radiation loss was very large and therefore the $\mu\mu$ -mass was at the far left end of the tail, indistinguishable from nonresonant muons. In this case the branching ratio would have been guessed a little too high.

Also, in widening the veto windows nonresonant events have been cut away, lowering the amount of events left for the further analysis. This results of course in lower statistics, but also could lower the resulting value for the branching ratio. Overall, these effects are considered minor.

Discrepancy of Monte Carlo events Because this analysis was treating a very rare decay the actual amount of events in the decay channel was quite small. Thus, it heavily relied on simulated MC events to have a sample of events large enough to study the properties of some variables and ultimately using it to train the BDTs in the multivariate analysis. The MC samples were also used to estimate the selection efficiency correction, having again an influence on the resulting calculation without being quantified. Even though the simulation algorithms have improved over the past, there is still a discrepancy between the distributions of some real data variables and their simulated replicas. To account for this, a reweighting of some key variables would be necessary before further utilizing the MC samples. This has been started but could not be fully completed in time, therefore has not been included in the analysis steps presented within this thesis. Therefore it shall be shortly mentioned and outlined at this point.

Exemplarily, the comparison of the distributions of the B mesons transverse momentum p_T ²² as shown in [Figure 21](#) shall be used. For this comparison the reference channel has been used (both for data and MC sample), since this should take place *before* the multivariate analysis, which will provide the decay channel events at first. The extraction of the signal distribution from the background parts has been done via the *sPlot*²³ (or *sWeights*) technique [15], where as before the fit models have been chosen as

²²Other distributions showing discrepancies are e.g. Vertex χ^2 and IP_χ^2 of the B meson.

²³*sPlot* is a statistical tool designed to investigate data sets containing distributions from different sources, like various overlaying decays.

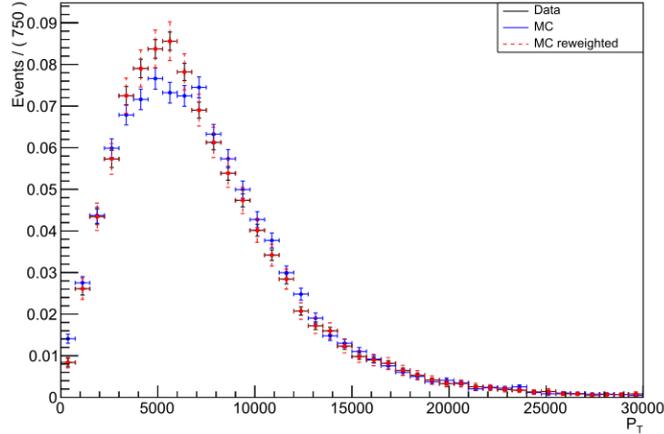


Figure 21: Normalized distributions of the B mesons transverse momentum for real data (black) and MC (blue) events. The MC distribution shows a dip around the peaking region.

a double CB for the J/Ψ peak and an exponential for the background. In the opinion of the author this correction contributes the largest systematic uncertainty.

On the subject of Monte Carlo there are two more sources of systematic errors, first being errors in the used generation efficiencies γ . These are not quantified for the values used here, yet expected to be very small and therefore should only give rise to negligible contributions. Secondly, efficiency corrections have only been made for the selection process. Discrepancies of tracking efficiency or trigger efficiency between the simulated and the real data are also possible.

Fit models As in every analysis that requires fitting, the used models must be chosen carefully. In case of this thesis, the (double) Crystal Ball function was used in many fits. It turned out its parameters α and especially n are difficult to choose and can vary to some extent, also when a *sum* of two CB was used. To prevent any inclusion of background events into the tails, these parameters have been determined by fitting a MC sample in [Section 5.2.1](#)²⁴. As an example, a free-parameter fit of the data and a subsequent comparison with the method used here could be a way to quantify the systematic uncertainties in this case. Also the choice of other fit models could be valid, for the signal part e.g. a double Gaussian, only a single Crystal Ball or even some other modification of a Gaussian could describe the data equally well. The same applies for the background part, where in this thesis either a constant value or an exponential function got used. Other choices could have been a hyperbola or a polynomial. Hence, the choice of the fit function represents a systematic uncertainty.

²⁴In the fits of [Section 6](#) this was not done, since the background was very low and it was not of importance to get every single event as for the veto cuts.

In [Figure 20b](#), the lack of a comprehensive systematic study is clearly visible. Comparing this distribution with other similar ones, it is apparent that this distributions shape differs quite a lot (it has a sharper pike). This could be due to biased variables in the MC files not resembling the data close enough, leading to distorted classification in the BDT analysis. The reweighting of the MC samples could help with that, yet this is just an estimation and would need to be properly tested. Also the fit does not describe the distribution as well as in the other cases, raising the question if in this case the semi-phenomenological Crystal Ball model was the best choice. Here it would be necessary to study the systematic effects in more detail.

Finally, it can be said that in the authors opinion the systematic errors, even though not been quantified and therefore widening the measured value by some extent, are *yet* still of minor size compared to the statistical uncertainties. Nonetheless it is important to further analyze them, especially as future measurements will get into regions where the ratio of uncertainties will shift towards the systematic one. For the data used in this thesis, the measured error is statistically dominated.

8 Summary

In this thesis, the branching ratio \mathcal{B} of the rare B meson decay $B^+ \rightarrow \mu\mu K^{*+}$ relative to the decay mode $B^+ \rightarrow J/\Psi(\rightarrow \mu\mu)K^{*+}$ has been determined. The used data was recorded at the LHCb detector at CERN in the years 2011 and 2012, during which the LHC was conducting pp -collisions at center-of-mass energies of 7 TeV and 8 TeV, respectively. The data set represents a total integrated luminosity of $\mathcal{L}_{int} = 3fb^{-1}$. In addition to the data set, for both decay modes a separate Monte Carlo sample was used. During the analysis, an event selection based on a multivariate analysis as well as additional linear cuts on the muon probability and the reconstructed K^* mass was performed. Additional focus was put on the identification of all events decaying via a resonance. Since these events overlay the decay mode $B^+ \rightarrow \mu\mu K^{*+}$, they needed to be filtered out. Also those events decaying via a J/Ψ have been used as reference channel to determine the relative branching ratio as

$$\frac{\mathcal{B}(B^+ \rightarrow \mu\mu K^{*+})}{\mathcal{B}(B^+ \rightarrow J/\Psi(\rightarrow \mu\mu)K^{*+})} = (1.23 \pm 0.13_{stat}) \times 10^{-3},$$

with a total amount of 155 ± 16 reconstructed nonresonant events. Utilizing the known value for the branching ratio of the reference channel and its subsequent decay mode (both taken from [1])

$$\begin{aligned}\mathcal{B}(B^+ \rightarrow J/\Psi K^{*+}) &= (1.44 \pm 0.08) \times 10^{-3}, \\ \mathcal{B}(J/\Psi \rightarrow \mu\mu) &= (5.93 \pm 0.06) \times 10^{-2},\end{aligned}$$

the absolute branching ratio of the decay $B^+ \rightarrow \mu\mu K^{*+}$ is calculated to

$$\mathcal{B}(B^+ \rightarrow \mu\mu K^{*+}) = (1.05 \pm 0.11_{stat} \pm 0.06_{norm}) \times 10^{-6}.$$

The relative statistical error is of about 10.9%, while the size of the error due to the normalization with the reference channel is about 5.6%. Systematic uncertainties have been discussed qualitatively, and some suggestions for further studies have been made (as far as the author was able to identify these). Although the systematic error was not quantified, the total error seems to be statistically dominated.

While the performed analysis gives a glimpse on the branching ratio of the decay $B^+ \rightarrow \mu\mu K^{*+}$, a full angular analysis of this decay mode using the combined data of the years 2011 and 2012 has not been done so far. The obtained value is in agreement with previous measurements, which determined $\mathcal{B}(B^+ \rightarrow \mu\mu K^{*+}) = (1.12 \pm 0.15) \times 10^{-6}$ [1]. Both values are consistent within the 1σ -uncertainties, while the relative (but only

statistical) error has decreased with this measurement.

The LHC is currently in its preparation and testing phase for the second run at higher energies, which is scheduled to launch within this quarter year. This will give rise to even more data available for future analyses.

References

- [1] K.A. Olive (Particle Data Group). *Review of Particle Physics*. Chinese Physics C38, 090001, 2014. DOI: [10.1088/1674-1137/38/9/090001](https://doi.org/10.1088/1674-1137/38/9/090001)
- [2] J. J. Thomson. *Cathode Rays*. Philosophical Magazine 44 (269):293, 1897.
- [3] A. Sakharov, *Violation of CP invariance, C asymmetry, and baryon asymmetry of the universe*. Pisma Zh. Eksp. Teor. Fiz. 5 (1967) 32.
- [4] S. Descotes-Genon, J. Matias, J. Virto. *Understanding the $B \rightarrow K^* \mu^+ \mu^-$ - Anomaly*. [arXiv:hep-ph/1307.5683](https://arxiv.org/abs/hep-ph/1307.5683).
- [5] W. Altmannshofer, D. M. Straub. *New physics in $B \rightarrow K^* \mu \mu$?* [arXiv:hep-ph/1308.1501](https://arxiv.org/abs/hep-ph/1308.1501)
- [6] S. L. Glashow, D. Guadagnoli, K. Lane. *Lepton flavor violation in B decays?* [arXiv:hep-ph/1411.0565](https://arxiv.org/abs/hep-ph/1411.0565).
- [7] Particle Data Group webpage <http://pdg.lbl.gov>.
- [8] M. Thomson. *Modern Particle Physics*. Cambridge: Cambridge University Press, 2013.
- [9] D. J. Griffiths. *Introduction to Elementary Particles*. Weinheim: Wiley-VCH, 2008.
- [10] W.N. Cottingham. *An Introduction to the Standard Model of Particle Physics*. Cambridge: Cambridge University Press, 2007.
- [11] Schematic map of the LHC <http://lhcb-public.web.cern.ch/lhcb-public/en/detector/Detector-en.html>.
- [12] The LHCb Collaboration. *The LHCb detector at the LHC*. JINST 3 S08005, 2008. doi:[10.1088/1748-0221/3/08/S08005](https://doi.org/10.1088/1748-0221/3/08/S08005)
- [13] A. et al. Hoecker. *TMVA - Toolkit for Multivariate Data Analysis*. [arXiv:physics/0703039](https://arxiv.org/abs/physics/0703039).
- [14] S. Playfer and S. Stone. *Rare b Decays*. [arXiv:hep-ph/9505392](https://arxiv.org/abs/hep-ph/9505392).

- [15] M. Pivk, F. R. Le Diberder. *sPlot: a statistical tool to unfold data distributions*.
[arXiv:physics/0402083](https://arxiv.org/abs/physics/0402083).

Declaration

Ich versichere, dass ich diese Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Heidelberg, den 22.04.2015,