ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE

TEZE K DISERTAČNÍ PRÁCI

České vysoké učení technické v Praze
Fakulta jaderná a fyzikálně inženýrská
Katedra Fyziky

Jaroslav Günther

Search for $B \rightarrow \mu^+\mu^-$ Decays with the Full Run I Data of The ATLAS Experiment

Doktorský studijní program: Aplikace přírodních věd

Studijní obor: Jaderné inženýrství

Teze disertace k získání akademického titulu "doktor", ve zkratce "Ph.D."

Praha, červen 2015

Disertační práce byla vypracována v prezenční/distanční/kombinované* formě doktorského studia na katedře fyziky Fakulty jaderné a fyzikálně inženýrské ČVUT v Praze.

Uchazeč: Ing. Jaroslav Günther
Fakulta jaderná a fyzikálně inženýrská ČVUT
Břehova 7, Praha 1

Školitel: prom.fyz. Václav Vrba, CSc.
Katedra Fyziky
Fakulta jaderná a fyzikálně inženýrská ČVUT
Břehova 7, Praha 1

Oponenti: ....

....

....

Teze byly rozeslány dne: ....

Obhajoba disertace se koná dne .... v ……… hod. před komisí pro obhajobu disertační práce ve studijním oboru *Jaderné inženýrství* v zasedací místnosti č .... Fakulty jaderné a fyzikálně inženýrské ČVUT v Praze

S disertací je možno se seznámit na děkanátě Fakulty jaderné a fyzikálně inženýrské ČVUT v Praze, na oddělení pro vědeckou a výzkumnou činnost, Břehová 7, Praha 1.

předseda komise pro obhajobu disertační práce
ve studijním oboru

Jaderné inženýrství

Fakulta jaderná a fyzikálně inženýrská ČVUT, Břehová 7, Praha 1

Contents

This thesis statement is an excerpt from the doctoral thesis submitted to the Czech Technical University in Prague, Faculty of Nuclear Science and Physical Engineering in preparation for the defense of a doctoral degree. Please refer to the full document for a detailed explanation of the measurement herein. The full thesis has 208 pages and contains also a detailed overview of all experimental studies and many consistency cross-checks performed.

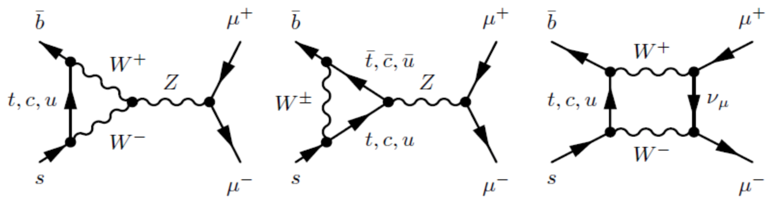
1 Introduction

The huge data analysed by LHC experiments suggest that the "Higgs" discovered is actually very close to the Standard Model (SM) Higgs boson with all properties and new hints of new physics particles have not been detected yet. The B_s and B_d meson decays into two muons are very sensitive to physics beyond the equisitely tested (e.g. Ref. [1]) Standard Model. $B^0_{(s)} \to \mu^+\mu^-$ purely muonic decays are forbidden at the tree-level of the Standard Model (SM). $B^0_s \to \mu^+\mu^-$ is therefore a very rare subatomic decay which happens about four times out of one billion decays and $B^0 \to \mu^+\mu^-$ decay is estimated to be about $40\times$ less frequent. Helicity suppressed flavour changing neutral currents contribute to these processes. An example of Feynman diagrams can be seen in Figure 1[1]. Flavour structure of the SM is very important to be investigated for its own sake. CKM matrix elements are being determined by combining heavy-light pseudoscalar meson decay constants from theory (lattice QCD) and decay rates from experi-

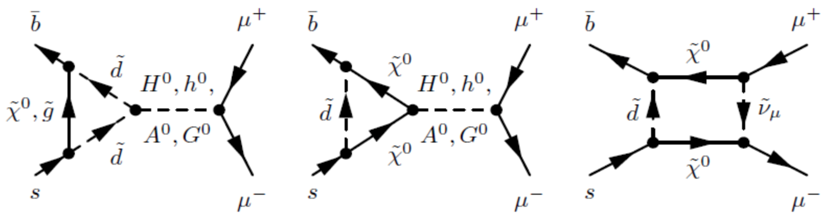
[1]Flavour changing neutral current (FCNC) transitions $b \to s$ or d are forbidden at tree-level in the SM.

ments. Being purely muonic, $B^0_{(s)} \to \mu^+\mu^-$ decays constitute very promising field to study since they offer the possibility of precise and rigorous theoretical predictions (mostly QCD-free constraint) to be compared to clean experimental signature. In particular, these decays are studied as they could open a window to theories extending the Standard Model to regions, where it does not cover for a satisfactory answer to the observed phenomena. In these various extended theoretical scenarios new (pseudo-)scalar operators could lift the strong SM helicity suppression of these FCNC, or the branching ratio could be suppressed by destructive interference between new physics operators with the ones already implemented in our SM. Thus, any deviation from the SM predictions on the branching ratios of $B^0_{(s)} \to \mu^+\mu^-$ could indicate unknown non-SM processes (involving new particle species) to contribute. On the other hand these decays serve to perform genuine probe of Yukawa interactions or to an Electroweak precision test (with respect to the Z penguin diagram). $B^0_s \to \mu^+\mu^-$ decays have been discovered only very recently by CMS and LHCb which collaboratively analyzed their collected data together. A small hint of a deviation from the SM observed in the recent experimental measurements of the branching ratios of these decays $\mathcal{BR}(B^0_{(s)} \to \mu^+\mu^-)$ triggered a lot of activity on both the experimental and theoretical fields. The ATLAS Collaboration has been searching for $B^0_{(s)} \to \mu^+\mu^-$ decays using merged 2011 $\sqrt{s} = 7$ TeV and 2012 $\sqrt{s} = 8$ TeV Full Run I Data sample ($\approx 25 fb^{-1}$). The analysis procedure has been firmly established and so called "unblinding" of the search region of $B^0_{(s)} \to \mu^+\mu^-$ is imminent and paper shall be published very soon.

The events for ATLAS analysis are selected by di-muon triggers and passed over to the reconstruction at Tier-0 after which the obtained data are analysed with the help of extensive Monte Carlo simulations. The parts of the ATLAS detector, which these analyses make an extensive use of, are the Inner Detector and the Muon Spectrometer with its dedicated tracking chambers. The ATLAS Collaboration uses by now well established strategy of a blind analysis technique excluding the signal region of B_s invariant mass distribution from the analysis data until the full analysis procedure has been firmly settled. Sideband events in data are split to allow for the following two procedures to proceed unbiased: the unbiased interpolation of the background into the signal region (1) and selection optimization (2). Better accuracy is achieved by performing the measurement of the branching ratio ($\mathcal{BR}$ ($B^0_{(s)} \to \mu^+\mu^-$)) with respect to a reference signal decay $B^{\pm} \to J/\psi K^{\pm}$. Another common feature to all ATLAS searches for $B^0_{(s)} \to \mu^+\mu^-$decays is a use of a multivariate analysis (MVA) classifier for signal-background separation. The analysis flow on Full Run I dataset is in many aspects revised and has several significant differences from the previous two analysis versions. This is predominantly because the 2012 dataset has different characteristics than 2011 one and represents different challenges. In this analysis we concentrated on the possibility to measure the actual $\mathcal{BR}$ ($B^0_{(s)} \to \mu^+\mu^-$) branching fraction since the recent evidences from CMS and LHCb experiments has shown 4σ effects for the $B^0_s \to \mu^+\mu^-$ final state resulting in a combined average $\mathcal{BR}$ ($B^0_s \to \mu^+\mu^-$) $= (2.9 \pm 0.7) \times 10^{-9}$ (July 2013). Only very recently CMS and LHCb published a combined analysis results in which the data from both experiments were analysed together $\mathcal{BR}$ $(B_s^0 \to \mu^+\mu^-) = (2.8^{+0.7}_{-0.6}) \times 10^{-9}$ and $\mathcal{BR}$ $(B^0 \to \mu^+\mu^-) = (3.9^{+1.6}_{-1.4}) \times 10^{-10}$. Their results were in excellent agreement and both fell just below the 5 sigma statistical precision historically needed to claim an observation of the $B_s^0 \to \mu^+\mu^-$channel. The combined analysis easily exceeded this requirement, reaching 6.2 sigma for the $B_s^0 \to \mu^+\mu^-$(3.2 sigma for $B^0 \to \mu^+\mu^-$). $\mathcal{BR}(B_s^0 \to \mu^+\mu^-) = (3.66 \pm 0.23) \times 10^{-9}$ and $\mathcal{BR}(B^0 \to \mu^+\mu^-) = (1.06 \pm 0.09) \times 10^{-10}$ Ref. [2] Ref. [3] Ref. [4] Ref. [5] are the latest theoretical predictions. Due to limited trigger efficiency and mass resolution of the ATLAS detector, we are less sensitive to these decays apriori. A sensitivity of the analysis to the $B_s^0 \to \mu^+\mu^-$ signal is estimated to be $4.7 \pm 1.0\sigma$. The description of the ATLAS $B_{(s)}^0 \to \mu^+\mu^-$ analysis parts to which the doctoral thesis contributed is given in brief in this thesis statement together with a parallel measurement and the results obtained are summarised.



(a) Penguin. **(b)** Penguin. **(c)** Box.



(a) Penguin. **(b)** Penguin. **(c)** Box.

Figure 1: Leading Order (LO) Feynman diagrams *Top Row:* of processes contributing to $B_s^0 \to \mu^+\mu^-$ decay in the Standard Model (SM). *Bottom Row:* Feynamn diagrams of possible processes contributing to $B_s^0 \to \mu^+\mu^-$ decay in SM extensions such as the Minimal Super-Symmetric Model (MSSM). H^0, h^0, A^0 and G^0 are the neutral Higgs and would-be Goldstone bosons,$\tilde{\nu_\mu}$ is the sneutrino, $\tilde{d}$ denotes the down-type superpartners of the quarks (squarks), $\tilde{\chi^0}$ are a neutralinos (Higgs and EW superpartners) and $\tilde{g}$ is gluino.

2 The ATLAS $B^0_{(s)} \to \mu^+\mu^-$ Analysis Strategy

This Full Run I analysis on $\approx 25 fb^{-1}$ is again a "blind" analysis which excludes the invariant mass window of 360 MeV width around $B^0_{d,s}$ mass (5166 – 5526 MeV) from the analysis development to avoid biasing the analysis optimization. Another common feature of this analysis with the previous two rounds is the dependency of on precise extraction of the reference channel $B^{\pm} \to J/\psi K^{\pm}$ yield to achieve the best possible accuracy on the measured branching ratio $\mathcal{BR}$ $(B^0_{(s)} \to \mu^+\mu^-)$. In general, the main $B^0_s \to \mu^+\mu^-$analysis tasks could be roughly subdivided into several main steps which are discussed further in the next sub-sections :

1. **Monte Carlo Production** of simulated samples for signals and backgrounds involved
2. ***Calibration*** of the main discrepancies between the Monte Carlo Samples and the Data
3. **Candidate Preselection** described in Section 3.
4. **Development of MVA classifiers** to discriminate against various backgrounds, the BDT for peaking background rejection decribed in Sections 4.
5. **Extraction of Signal and Reference Channel Yields** from the Data (the former only after unblinding)
6. **Assessment of relative efficiencies and detector acceptances** - Monte Carlo derived evaluation of the relative $B^0_{(s)} \to \mu^+\mu^-$vs $B^{\pm} \to J/\psi K^{\pm}$ Efficiency×Acceptance Ratio
7. **Evaluation of systematic uncertainties** on all measurement ingredients
8. **Branching Ratio Extraction** described in Section 6.

Background Sources

In our analysis, one needs to identify all dangerous background processes to have a chance to uncover the very rare $B^0_{(s)} \rightarrow \mu^+\mu^-$signal. From topological point of view, signal decay is reconstructed in ATLAS detector as two oppositely charged muon tracks using information from the Muon Spectrometer and Inner Detector. Both such muon candidates are fitted Ref. [6] in a common decay vertex of the $B^0_{d,s}$meson. We benefit from the long lifetime of a B^0_smeson ($\tau = 1.47 \pm 0.03\ ps$) which allows us to detect the displacement of the B meson's production vertex from its the decay vertex. There are several useful quantities that can be built by using this information and help us quite easily discriminate against dominant (prompt) Drell-Yan pairs ($pp \rightarrow \mu^+\mu^-$). As for the non-prompt background, we categorize our background sources based on various topologies described in the following paragraphs. A sketch of the signal topology and various same vertex decay topologies can be found in Figure 2.

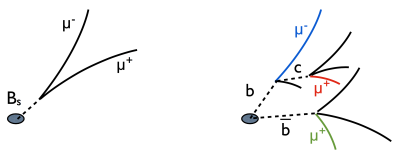


Figure 2: Topology of a signal decay (left) and various topology options for background (right). The double semileptonic $\mu^+\mu^-$ (green, blue), opposite-side sequential-semileptonic constributions (red, green) and negligible contribution of sequential (same-side) decays (red, blue).

Peaking Background has exactly the same topology as the main decay is composed of $B \to hh'$, mainly $B_s^0 \to K^+K^-$and $B^0 \to K^\pm\pi^\mp$, in which both hadrons are misidentified as muons (*fake-muons*). It is a very dangerous background source overlapping with our the signal peaks. I have invested an extra effort and reduced this background by suppressing the *fake-muons* fractions. The peaking background mass shapes can be seen in Figure 3. Thanks to my specific multi-variate selection for *fake-muons* (BDT), the amount of this background is expected to be ≈ 2.5 % of the B_s^0 signal, ≈ 20 % of the expected B^0 event yield.

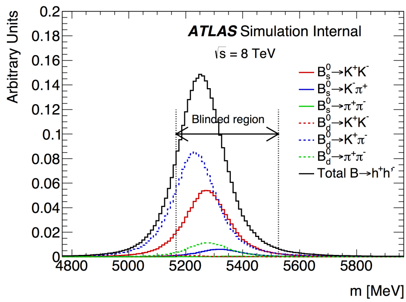


Figure 3: Invariant mass distribution of the peaking background components $B \to hh'$.

Combinatorial Background due to *opposite-side*[2] semileptonic decays is very large background source in this analysis and was very well reduced by an MVA classifier (continuum-

[2]Each of the two muon candidates originates from one of b-hadron flavours in the event $(b/\bar{b})$ - opposite-side. If they both originate from the same b-hadron decay we call such decays same-side decay

BDT). After all selection cuts were applied, this background source dominates the high-mass sideband.

Misreconstructed decays dominate, after a final selection is applied, the low-mass region. These candidates originate in misreconstructed semileptonic b-decays and can be categorized as follows:

- B_c background composes effectively small contribution in which B_c decays into $B_c \to J/\psi\ \mu^+\nu \to \mu^+\mu^-\mu^+\nu$. The MVA classifier values are distributed between the signal-like and background-like values and the mass shape is smoothly decreasing towards the signal region.
- same-vertex (SV) background, due to partially reconstructed B^0 and B_s^0 events containing a muon pair, such as $B^0 \to K\,\mu^+\mu^-$; where both muons come form the same vertex;
- same-side (SS) background, due to *same-side* combinatorial background from cascades $b \to c\,\mu^-\nu \to s(c)\,\mu^+\mu^-\nu$; where the two muons do not originate from the same vertex;

The same-side and same-vertex (SS-SV) background includes double semileptonic cascade events (e.g., $B \to D\mu X \to \mu\mu X'$), which we call SS, where the muons do not originate from the same vertex, and events where the muons come from the same vertex (e.g., $B \to K\mu\mu$) , which we call SV. In both cases, the mass distribution of the two muons is peaked far below the signal region, and we are sensitive to a tail of the distribution determined by kinematic limits and detector resolution effects. Given the relatively smaller amplitude, these events are expected to feed into the continuum and partially in the SS-SV events.

Semileptonic Background is due few-body semileptonic b-decays feeding into our final selections though a misidentification $h \to \mu$, in the limit of low energy neutrinos. In particular $B^0 \to \pi\mu\nu$ and $B_s^0 \to K\mu\nu$ can contribute, together $\Lambda_b \to p\mu\nu$. The mass distribution for the last process extends closer to the signal region, but is highly suppressed because of a very low probability of misidentifying the proton as muon in ATLAS. Despite the very low fraction for $h \to \mu$ misidentification achieved, the last background contribution was explicitly tested [3]. It was not found significant with the ATLAS detector, presumably because of a reduced fraction of misidentification via punch-through (Section 4).

In the Full Run I analysis all relevant background sources were studied in depth and their characteristics assessed with the help of the largest Monte Carlo production ever made for a single analysis - the *four-corners* production sample. Background sources discussed and can be seen in Figure 4 where the 4-corner MC is compared to the Data mass sidebands and in Figure 3 where the $B \to hh'$ MC mass contributions are displayed. In both cases the comparisons are made with selections as close as possible to the final selection (unless stated otherwise explicitly). The background sources of

[3] it was defined as significant for the LHCb and CMS analyses

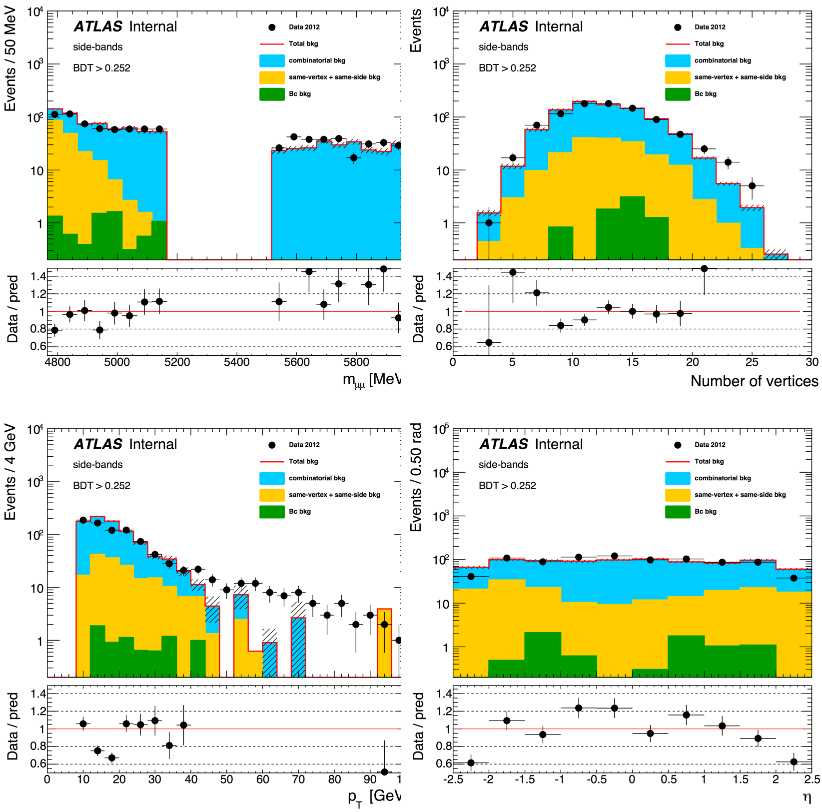


Figure 4: Distribution comparisons of data mass side-bands and the 4-corner MC sample after all selection cuts,and after requiring continuum-BDT > 0.252. The normalization of the 4-corners sample is done after the cut on continuum-BDT. *From left to right, from top to bottom*: invariant mass, number of primary vertices, B meson p_T and η.

the reference channel $B^{\pm} \to J/\psi K^{\pm}$ signal analysis are treated in a separate Section 5.

The Branching Ratio Measurement Strategy

Looking at the task list (described in the former section) from a practical point of view of the branching ratio measurement, the master formula is based on the idea of performing this study relative to a similar decay with sufficient statistics observed. The analysis is thus analogous to the measurement of a relative branching ratio with respect to a well established reference signal. There are several candidates on such reference signal decay $B_s^0 \to K^+K^-$, $B_s^0 \to J/\psi\phi$ and $B^{\pm} \to J/\psi K^{\pm}$. The decays having B_s^0 in the initial state would bring the advantage in canceling out the f_u/f_s ratio (and related uncertainty) in Equation (1). The first candidate is difficult to handle since ATLAS does not have means to identify kaons and there is no hadronic trigger present. Such otherwise tempting option for a reference channel is therefore ruled out. Despite having only twice lower branching fraction compared to the third candidate decay $B^{\pm} \to J/\psi K^{\pm}$ the $B_s^0 \to J/\psi\phi$ channel still has a disadvantage in reconstructing 4-tracks to be fitted which results in additional uncertainty term with respect to $B^{\pm} \to J/\psi K^{\pm}$ reconstruction. Thus, $B^{\pm} \to J/\psi K^{\pm}$ has been chosen as the best reference channel candidate and $B_s^0 \to J/\psi\phi$ has been deployed as our control channel to test the general B_s^0meson kinematic (and other background discriminating) variables in the Data-MC comparison studies. This way we achieve a substantial reduction of the production, luminosity and efficiency uncertainties.

The $\mathcal{BR}$ formula as represented in Equation (1) consists of 3 main inputs, observed event counts ($N_{\mu^+\mu^-}$, $N^k_{J/\psi K^\pm}$) of the extracted signals from the data (1), the efficiency $\times$ acceptance ratios $\frac{(A\varepsilon)^k_{\mu^+\mu^-}}{(A\varepsilon)^k_{J/\psi K^\pm}}$ derived based on MC simulation (2) and finally from the $\mathcal{BR}$ ($B^\pm \to J/\psi K^\pm$) and relative pp production rates of B^+/B_s mesons ($\frac{f_u}{f_s}$) (3) taken from the latest experimental results (LHCb). This formula takes also into account the use of different triggers in our analysis (see Section 3).

$$\mathcal{BR}(B^0_{(s)} \to \mu^+\mu^-) = \mathcal{BR}(B^\pm \to J/\psi K^\pm \to \mu^+\mu^- K^\pm) \times \frac{f_u}{f_s} \times \\ \times N_{\mu^+\mu^-} \times \left(\sum_k N^k_{J/\psi K^\pm} \alpha_k \frac{(A\varepsilon)^k_{\mu^+\mu^-}}{(A\varepsilon)^k_{J/\psi K^\pm}} \right)^{-1} \tag{1}$$

In Equation (1) the index k runs on the trigger categories used in the analysis (see Section 3). The α_k parameter takes into account the prescaling factor applied to $B^\pm \to J/\psi K^\pm$ events data.

Basic difference with respect to the previous ATLAS analysis is performing a mass fit to the signal $B^0_{(s)} \to \mu^+\mu^-$ (rather than multi-bin "cut & count" on the widest possible set of events to increase the signal sensitivity. Maximum-likelihood fit to the signal $B^0_{(s)} \to \mu^+\mu^-$ invariant mass distribution makes the quantitative conclusion on the amount in which each type of background and signal is present in real data. Therefore, a loose selection is applied to retain a maximum of signal events. After applying all preselection, additional cuts and *fake-muon* rejection cut (discussed in

Section 3 and Section 4) also a cut is made on the MVA classifier for reduction of a continuum background, the remaining events are fit in three intervals of this variable. CMS and LHCb used a similar approach.

3 Candidate Preselection

The requirements on the selection of all analysis decay candidates $B^{\pm} \to J/\psi K^{\pm}$, $B_s^0 \to J/\psi\phi$ and $B_{(s)}^0 \to \mu^+\mu^-$ are kept as consistent as possible with the signal decay $B_{(s)}^0 \to \mu^+\mu^-$. In this Section I intend to discuss only the main signal channel candidate selection. The particularities of candidate preselection for the $B^{\pm} \to J/\psi K^{\pm}$ signal analysis is treated in the appendix of the thesis.

To select a collision event and filter out cosmic muons we require an event to contain at least one reconstructed primary vertex with at least three associated Inner Detector tracks. Further, following the reconstruction guidelines of the corresponding ATLAS performance subgroups, we fit to a common vertex any two tracks associated with oppositely charged muon candidates using an algorithm described in Ref. [6]. The calculation of the di-muon invariant mass has been performed using the whole *combined muon* track candidate information, i.e. involving both Inner Detector and Muon Spectrometer tracking information. We do so, because including the latter improves the invariant mass resolution (especially in the end-caps). Reconstructed events are considered to contain a $B_{(s)}^0 \to \mu^+\mu^-$candidate if the following criteria are satisfied.

- Each muon must be a *combined muon* [4].
- Each muon must have $p_T > 4$ GeV and lie within $|\eta| < 2.5$.
- Both muon candidates match in associated secondary vertex fit with $\chi^2/NDF < 6$.
- The $B^0_{d,s}$ invariant mass is in the range [4766 – 5966] MeV.
- All B candidates passed also $p_T^B > 8.0$ GeV and $\left|\eta^B\right| < 2.5$ cuts.

For the selection of combined muon candidates mentioned above, the Muon Combined Performance (MCP) group recommendations are used:

- >0 (Pixel hits + crossed dead Pixel sensors)
- >4 (SCT hits + crossed dead SCT sensors)
- if $0.1 < |\eta| < 1.9$:

  (TRT hits + TRT outliers)>5
  and (TRT outliers) < 0.9*(TRT hits + TRT outliers)

- <3 (Pixel + SCT) holes

We also apply *additional cuts* to reduce the background without cutting away signal candidates based on known appreciable signal features such as long $B^0_{d,s}$ lifetime or the good isolation of the emitted muons which have a sum of their transverse momenta pointing at a small angle with respect to the $B^0_{d,s}$ p_T^B. Thus, ΔR cone of the two muons is required $\Delta R < 1.5$, B^0_smeson momentum pointing angle to the primary vertex in 2D is requested $|\alpha_{2D}| < 1.0$

[4] Having both ID and MS track segments.

and a transverse decay length to satisfy $L_{xy} > 0$ [5]. These additional cuts combined reduce the background by a factor of 2.5 without cutting into the signal.

Trigger Selection

The high level trigger rate bandwidths are shared among ATLAS physics groups and a careful tuning was made before the data aquisition to satisfy the needs of each analysis group as democratically as possible. For 2011 dataset two different trigger chain algorithms (*2mu4T* and *2mu4*) are used to identify and record high-quality di-muon event (p_T cut at 4 GeV). The *2mu4* trigger was seeded at level 1 trigger with no p_T cut and requested to fire on the first half of 2011 data-taking. For the second half of 2011 the *2mu4T* was seeded already at level 1 with with $p_T$cut of 4 GeV. In general many physics analysis including ours confirmed that the effect of this change during 2011 data-taking is negligible for the analysis flow. Thus, entire 2011 data was considered as a whole and consistent dataset in the Full Run I analysis.

In 2012 data-taking the pile-up and higher luminosity conditions made it impossible to keep the *2mu4T*-like triggers unprescaled[6] while fitting into a sustainable bandwidth thresholds. The introduced prescaling reduces significantly the amount of data we can use for our measurement from about 20.3 fb^{-1} to effective 16.2 fb^{-1} coming from the *2mu4T* trigger.[7] Therefore an effort

[5] $c\tau = \overrightarrow{L_{xy}} \times \overrightarrow{p_T} / M_{B_s^0}$

[6] Throwing away every e.g. 10th interesting event.

[7] This prescaling was marginally present already for the second half of 2011

has been invested into studying how we could partially recover the 20% loss by using signal events fired by other triggers (e.g. *2mu4Tmu6*-like) and include those in our selection.

It was found that nearly all of the events (98%) pass either one of the following three high level trigger chains: EF_2mu4T_Bmumu, EF_2mu4T_Bmumu_Barrel, or EF_mu4Tmu6_Bmumu. An equivalent set of triggers has been settled upon for the $B^+ \to J/\psi K^+$ channel labeled *Jpsimumu* instead of *Bmumu*. The *Jpsimumu* triggers were prescaled by a (different) factor of ≈ 10 [8]. As a consequence, we are introducing the extra factor α_k in Equation (1) on the efficiency ratio between the signal and the reference channel.

Taking into account the 3 triggers selected in the paragraph above we defined 3 mutually exclusive trigger selection categories (N_1, N_2, and N_3) to better isolate the specific topological differences.

- N_1: EF_2mu4T_Bmumu && !(EF_2mu4T_Bmumu_Barrel || EF_mu4Tmu6_Bmumu)
- N_2: EF_2mu4T_Bmumu_Barrel && !(EF_mu4Tmu6_Bmumu)
- N_3: EF_mu4Tmu6_Bmumu

N_1 are low-p_T end-cap events, N_2 are low-p_T barrel events and N_3 are events not restricted to a specific pseudorapidity region having 4 and 6 GeV p_T thresholds set for each of the two muons respectively. The relative efficiencies of these trigger categories

data-taking.

[8] The actual number is calculated from the accurate luminosity information.

were also compared both in the data sidebands and in the exclusive MC samples (both for the signal and reference channels). As a result of this comparison a weight factors correcting the MC to the data relative abundances are then applied to our MC samples when necessary.

4 Peaking Background Discrimination

The $B \to hh'$ charmless two body B decays (h being a charged K or π) present a big threat to the $B^0_{(s)} \to \mu^+\mu^-$ measurement since these resonant background sources peak under our $B^0_{(s)} \to \mu^+\mu^-$ signal and are topologically identical with it. The only way this background enters our analysis is obviously by misreconstructing charged hadronic candidate as a muon candidate. Extreme reliability of ATLAS muon identification capability is necessary for this purpose. Thus, a dedicated study has been performed to suppress the fractions of misidentified hadrons as muons (further also termed as *fake muon rates*).

Since this section is intended to be rather a brief overview of the analysis flow and this Subsection 4 represents one of my analysis contribution outcomes, I would like to point the reader to the full thesis, where one can find elaborate description and justifications to all results summarized in this chapter as well as a complete description of how a tool to discriminate against the dangerous $B \to hh'$ background source was developed. Good baseline information about machine learning and the ROOT TMVA framework used for this work can be found in Ref. [7], or Ref. [8] and references therein.

Reduction of ATLAS muon fake fractions.

The study of fraction of misidentified hadrons as muons (*fake-muon rates*) has been performed using 4 MC samples: signal $B^0_s \to$

$\mu^+\mu^-$ (1), default sample of charmless two-body decays $B \to hh'$ (2), similarly produced additional calo-sample of $B \to hh'$ (3) and finally $\Lambda_b \to ph$ sample (4). These samples have been produced with full GEANT simulation in order to accurately describe the hadrons after they leave the Inner Detector. To follow the recommendations of the ATLAS muon performance subgroup (MCP) and stay in consistent kinematic regime to the main signal, the same preliminary cuts (preselection and additional cuts) have been applied on the reconstructed MC events as the ones listed in Section 3 for $B_s^0 \to \mu^+\mu^-$ channel. The basic *2mu4T* trigger request has been applied to signal events. Preselected hadrons (h-legs) from $B \to hh'$ that were misidentified as *combined muons* are counted as *fake muons*. The *fake rate* is therefore defined as:

$$muon\ fake\ rate = \frac{\#F}{\#P} = \frac{N^{\text{combined muons}}_{\text{kinematic+MCP cuts}}}{N^{\text{no MS requirement}}_{\text{kinematic+MCP cuts}}} \tag{2}$$

, where $\#F$ is the number of single h-legs passing all selections described above including the h leg being tagged as a combined muon. $\#P$ is the number of single h-legs passing the preselection and additional cuts without applying any Muon Spectrometer related cut.

Table 1 displays the misidentification fraction for protons, kaons, and pions after the preliminary cuts as measured through the full simulation of the decays of b-hadrons to pairs of charged, long lived hadrons, in which one of the hadron is misidentified as a combined muon (see Section 3).

Specialized study with a dedicated MC production[9] of the above mentioned calo-sample of $B \to hh'$ has shown that kaons 97% and 92% of pions are flagged as fake muons due to decays in flight. The remaining cases correspond to a negligible amount of punch-through (when hadron does not decay into muon but reaches the muon spectrometer leaving a legitimate track trace). The small fraction of such events explains the negligible contribution of protons and antiprotons to the total hadron misidentification as muon.

STACO muons	$\Lambda_b \to ph$	$B \to hh$	$\Lambda_b \to ph$	$B \to hh$	$\Lambda_b \to ph$	$B \to hh$	$\Lambda_b \to ph$
	$p(\bar{p})$	$K^\pm$		$\pi^\pm$		global K/π	
#P : $K/\pi/p$	776916	1634183	481597	1719861	295319	3354044	776916
#F : $K/\pi/p$	26	6458	1894	3651	678	10109	2572
fake rate (after preliminary cuts)	$(3.3\pm0.7)\times10^{-5}$	$(3.95\pm0.05)\times10^{-3}$	$(3.93\pm0.09)\times10^{-3}$	$(2.12\pm0.04)\times10^{-3}$	$(2.30\pm0.09)\times10^{-3}$	$(3.01\pm0.03)\times10^{-3}$	$(3.31\pm0.07)\times10^{-3}$

Table 1: MC samples (with full simulation) used for the studies of fake muons (first line), misidentified hadron (second line), number of generated hadron (#P) and number misidentified as muon (#F) after the preliminary cuts (third and fourth line), fraction of hadron identification as muon (#F/#P) after the preliminary cuts (last line).

When performing approximate extrapolation from 2011 analysis results to the Full RUN I analysis conditions, the expected number of signal events in the Full Run I analysis is ≈ 35 and the number of $B \to hh'$ events with the fake rates as observed in Table 1 is expected to be ≈ 10. From these rough estimates, it is clear that this background needs further rejection.

[9]MC production in which the propagation and interactions of the hadrons using GEANT simulation are recorded in finer detail also from the calorimeter.

≈ 36 properties of misreconstructed h-legs (as muons) and signal muons has been studied in detail and it was concluded that no efficient separation between them can be performed by simple cut and count approach. Inspired by work in the Muon Combined Performance subgroup (exploring Multivariate analysis possibilities), we were suggested to use a method of Boosted Decision Trees for developing a BDT classifier separating such misidentified hadrons from true muons. After a detailed analysis described in Section 4, we have selected a specific number of discriminating variables tailored to our needs to discriminate between true and fake muons.

In order to optimize the training of our BDT to be sensitive to the data-like h-legs misidentified as muons a single muon trigger object match (*mu4T*) has been requested. The requirement of matching such muon candidate to a muon trigger object reduces further the number of fake muons (mis-reconstructed h-legs) by a factor equal to 0.582 ± 0.015, the same for kaons and pions. The association is based on the angular separation $\Delta R = \sqrt{(\Delta\phi^2 + \Delta\eta^2)}$. When a matching trigger object is found, the distributions of ΔR are very similar between real and fake muons (peaking at 0.0002 and extending up to about 0.0025). Not requiring trigger conditions on the tracks in the BDT training, or separate training for kaon/pion discrimination showed only marginal impact on the final BDT performance. These are one of large number of variations tested during the BDT training phase discussed in Section 4. It has also been explored that BDT training dependency on the track kinematics (p_T , η) is not significantly improving the final performance. Reconstructed and *truth-level* (MC generated) values of the momentum of fake muons have been compared, and differences in the sample of fakes retained after the BDT selection are negligible.

As it is discussed in Section 4, a study has shown that a muon reconstruction algorithm STACO can be safely chosen for this analysis instead of MUONS. Both containers STACO and MUONS featured very consistent selections of fake muons after the BDT classifier cut (at 90% single signal muon efficiency or higher).

AbsEta	*nTrees* = 800
EtCore	*BoostGrad* = *Grad*
fitChi2	*MinNodeSize* = 0.5%
match quality	*nCuts* = 100
MDT Hits	*maxDepth* = 4
Scattering Curvature Significance	*Shrinkage* = 0.1
qoverpME/qoverpID	*BaggedSampleFraction* = 0.6
qoverpMS	*NormMode* = *NumEvents*

Table 2: List discriminating variables (left), and configuration parameters of BDT used for rejection of fake muons (right).

The list of the variables used in the final selection and the parameters describing the BDT configuration are shown in Table 2. These variables have the following definitions:

- *AbsEta*: absolute value of η of the ID track.
- *EtCore*: energy deposited in the calorimeter around track passing through. More energy is expected from K or π.
- *fitChi2*: muon track fit χ^2.
- *match quality*: χ^2/n.d.f. of the match between ID and MS

tracks.

- *MDTHits*: number of MDT hits.
- *Scattering Curvature Significance*: Scattering curvature significance is related to the difference in track curvatures computed upstream or downstream of a detection plane of the ID. Fitted track curvature from one side minus from the other side of the ID measuring surface. The maximum among all surfaces taken. Plus sign indicates increase in curvature while minus sign decrease in momentum.
- $qoverpME/qoverpID = \frac{trkMuonExtr\_qoverp}{trkID\_qoverp}$: where $trkID\_qoverp$ is q/p_{ID}, p_{ID} = total momentum of ID track; $trkMuonExtr\_qoverp$ represents q/p_{ME}, where $p = p_{ME}$ = total momenta of the track extrapolated to the ID perigee = p_{MS} + energy loss (parametrised). The energy loss contains the amount of energy lost in the material between ID and MS.
- *qoverpMS*: q/p_{MS}, p_{MS} = total momentum of MS track

Final BDT Evaluation

For the final performance evaluation of the BDT an independently generated and simulated *calo*-sample of 4M *Bhh* events was chosen together with the previously used $B_s^0 \to \mu^+\mu^-$ sample with the same trigger selection as used in the final analysis (events in either of the 3 trigger categories N1 or N2 or N3 are accepted). The final BDT cut value for 95 % single signal muon efficiency has been measured to be -0.458. In 2011 Data the SCSig variable was not accessible, therefore a second 2011-conditions-specific BDT was retrained without it and the corresponding cut value was determined to be -0.435 (used for 2011 data/MC studies). A good agreement was concluded between the ROC curve from the TE and EVAL samples as seen in Figure 5. These ROC curves were evaluated using h-leg weighting (as it was the case for the training). The *fake* rates as seen from weighted events in the EVAL sample before and after trigger matching and BDT cut can be found in Table 3. This table shows the overall performance of the fake-muons reduction procedure, showing the fraction of misidentified hadrons after preliminary cuts, then after adding the trigger match and also after the final BDT selection.

The amount of decays in flight has been studied by looking at the hadron decay vertex position for those that decay into a μ, this information can be seen in $r - z$ plane in Figure 6. It was found that only 8 % of kaon *fakes* and about 3 % of pion *fakes* appear to be punch-through *fake* muons (traversing the detector without decaying into muons). Also pion decays in flight (DIF) are rejected in 66 % of cases and kaon DIF in 64 % (after the trigger match) which seems to agree with the expectation that the punch-through fakes is difficult to discriminate even with our developed multivariate classifier.

Table 14 shows the reduction in the fraction of fake hadrons obtained with a BDT threshold corresponding to a single signal muon selection efficiency equal to 95%. The fake fraction would be further reduced by a factor $\simeq 0.8$ if the selection would be tuned to 90% single signal muon efficiency. The errors shown in Table 14 are related to statistical fluctuation in the evaluation sample.

particle type	after preliminary selection	adding trigger match	adding BDT selection
K^-	0.00360	0.00207	0.00076 ± 0.00005
K^+	0.00440	0.00263	0.00101 ± 0.00005
π^-	0.00202	0.00116	0.00044 ± 0.00004
π^+	0.00206	0.00121	0.00042 ± 0.00004
average	0.00309	0.00181	0.00067 ± 0.00002

Table 3: Cut flow of hadron misidentification fraction. The BDT selection is tuned for 95% muon efficiency, and the error is the statistical uncertainty after all cuts.

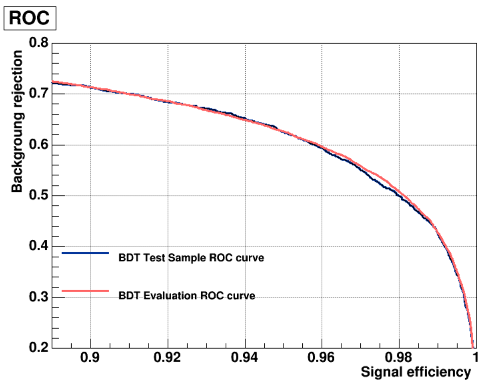


Figure 5: Final BDT ROC curve evaluated from the TE sample (blue) and EVAL sample (red), in both cases weighted events were used.

Double *Fake* Fractions

The relative contribution of the peaking background channels can be seen in Figure 3. Table 5 summarizes the most relevant decay modes, with the corresponding branching fractions (uncertainties 5–7% for the two main channels), and the total rejection factors obtained when the dimuon selection efficiency is equal to 90% (sta-

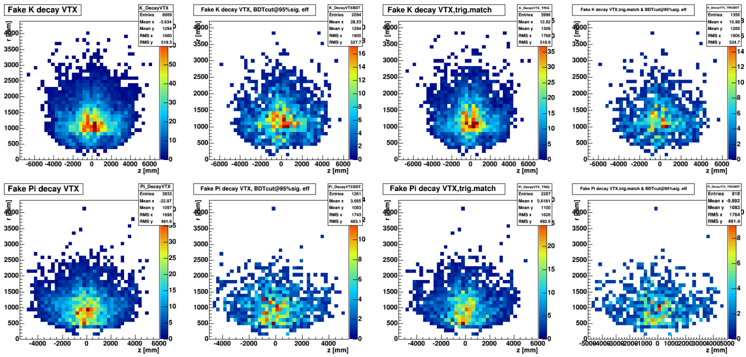


Figure 6: From the *fake* muons in the EVAL sample, those decaying into muon (decays in flight - DIF) were searched and their decay vertex position in the ATLAS detector ($r-z$ plane) is plotted above. Top row shows Kaon DIF, bottom row are pion DIF. From left to right the plots show: *fakes'* DIF VTX position without any trigger match or BDT cut requirement; *fakes'* DIF VTX position after BDT cut at 95 % signal efficiency; *fakes'* DIF VTX position after baseline selection and trigger match; *fakes'* DIF VTX position after trigger match and BDT cut at 95 % signal efficiency

$K^{\pm}$	0.376±0.007
$\pi^{\pm}$	0.366±0.010

Table 4: Fake muon reduction factors obtained with the BDT selection, for 95% muon efficiency, with statistical uncertainty.

tistical errors of about 5%). Including the factor of 3.9 between B^0 and B_s^0 production cross sections (f_d/f_s), the total background corresponds to an effective branching fraction for B_s^0 equal to about 6×10^{-11}.

peaking bkg. channel	branching fraction	fake rejection factor
$B_s^0 \to K^+K^-$	25×10^{-6}	7.6×10^{-7}
$B^0 \to K^{\pm}\pi^{\mp}$	25×10^{-6}	3.8×10^{-7}
$B^0 \to \pi^+\pi^-$	5.1×10^{-6}	1.9×10^{-7}

Table 5: Fake rejection factor denotes the double *fake* fraction for main channels of peaking background, after all selection requirements. The branching fractions refer to the PDG [9] values.

The estimated systematic uncertainty on the contamination fractions were estimated with checks against the Data. The fraction of fakes after preliminary cuts has been tested looking for tracks identified as muons in the resonance peaks of $K_s \to \pi^+\pi^-$ and $\phi \to K^+K^-$. This method was already used for the analyses of data collected in 2011. In addition, the channel $B^+ \to J/\psi K^+$ has been used for studying kaon misidentification in both data and MC, with full GEANT simulation. The simulation shows a fraction of fake muons significantly lower (by a factor $\simeq$4) than the one observed for $B^0 \to hh'$. This effect appears to originate from a tighter event selection at the level of $B^+ \to J/\psi K^+$ reconstruction with vertex constraint. The simulation of single kaons provides a fraction of fake-muons in agreement with the simulation of $B^0 \to hh'$. We have used the same fit procedure developed for the reference channel yield extraction and checked the kaon fake rate at B^+ in real data. A confirmation of the result with the simulation was found within a factor of 0.9$\pm$0.3. The peaking background after all selection cuts (including the cut on the BDT for *fake* muons rejection) was estimated to be $1.0^{+0.8}_{-0.5}$ signal candidate. The negative uncertainty is a -0.5 conservative estimate of possible uncertainties in the MC modelling of the selection against fake muons, and the positive one is extracted from the analysis on additional studies performed on both the $B^\pm \to J/\psi K^\pm$ Data and on $B^0_{(s)} \to \mu^+\mu^-$ sideband Data. Given the significant reduction in the size of the peaking background, the estimated uncertainty has a very small effect on the expected sensitivity for $B^0 \to \mu^+\mu^-$ signal.

5 Reference Channel Yield Extraction

In this section I will only shortly summarize the main points and results of the extraction of the reference channel yield. Please refer to the full description in the thesis for more information.

The $\mathcal{BR}(B^0_{(s)} \to \mu^+\mu^-)$ measurement master formula in Equation (1) has to have precisely measured yield of $B^\pm \to J/\psi K^\pm$ from data in all 4 analysis categories (3 trigger categories of 2012 and 1 category for 2011 data). All datasets used and event selection cuts are summarized in the main document. For this measurement a simultaneous unbinned extended maximum likelihood fit in mass was developed. This fit is simultaneously fitting at the same time all Monte Carlo models and Data which allows for automatic data-MC cross constraint. A brief description of the background and fit structure (in N3 trigger category) will be given targeting the fit results in the following text.

In Figure 7 on the left the invariant mass distribution of $B^\pm$ candidates in N3 trigger category is shown with a clear signal peak in the middle. The invariant mass composition can be sorted into the following sources:

- $N_{J/\psi K^\pm}$: number of $B^\pm \to J/\psi K^\pm$ signal events
- $N_{J/\psi \pi^\pm}$: number of $B^+ \to J/\psi \pi^+$ exclusive background events which form a small contribution under the signal peak
- N_{pr}: number of partially reconstructed background events (PRD) which form evident step-like structure

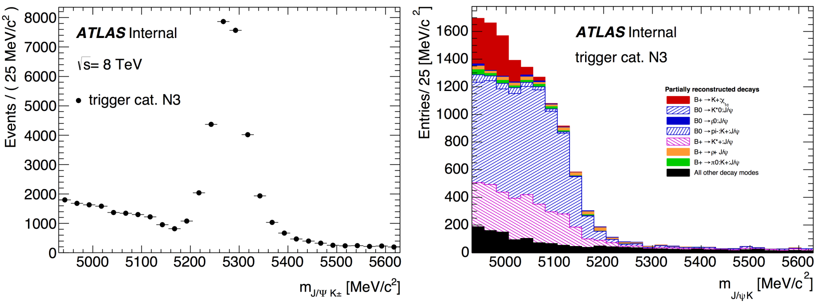


Figure 7: *Left*: $J/\psi K^{\pm}$ invariant mass distribution for all $B^{\pm}$ candidates in the trigger category N3 in 2012 data. *Right*: Partially reconstructed B decays contributing to the background as described by Monte Carlo.

- N_{comb}: number of combinatorial background events which are smoothly crossing our mass window

The mis-reconstructed decays (right plot in Figure 7) forming the evident structure of the left sideband are decays such as $B^{+/0} \to K^{*+/0} J/\psi$, $B^{+} \to K^{+}\chi_{c1,2}$ and similar, where one or more of the final state particles are missed in the reconstruction (or mis-reconstructed). Slightly right of the signal peak mean, we can find a small contribution from the reflection of the Cabibbo-suppressed $B^{\pm} \to J/\psi\pi^{\pm}$ decay with the assignment of the kaon mass to the final state pion. The last background source - combinatorial background - is composed mostly by $b\bar{b} \to J/\psi X$ events[10] and continuously spans entire fit window.

[10] Random combination of J/ψ (produced promptly in pp collisions or in feed-down from B-decays) with a track.

Four mass fits are performed, one for each of the 4 categories (N1-3 trigger categories in 2012, plus 2011 data category). The simultaneous inclusion of the three MC samples ($B^{\pm} \to J/\psi K^{\pm}$,$B^{\pm} \to J/\psi \pi^{\pm}$ and $b\bar{b} \to J/\psi X$ MCs) allows to guide the modeling of the most critical fit components in data. The data model fit components' shapes related parameters are tied to the corresponding Monte Carlo model parameters. This results in a "MC assisted" determination of the most critical fit component shapes in data, while automatically accounting for the statistical uncertainties of the MC. The combinatorial shape is not Monte Carlo driven in this sense and is added as an unconstrained fit component to fit the Data mass distribution. In addition, two additional parameters are added to all Data component models to allow the fit to adjust to residual data-MC discrepancies. These are the mass scale and the other for the mass resolution which are both extracted as mainly driven by the $B^{\pm} \to J/\psi K^{\pm}$ peak in data, but included also in all other data components consistently (PRDs and $J/\psi\pi^{\pm}$).

For more elaborate description of the fit likelihood function and PDFs used I would point the reader to the thesis. One of the several novelties of this fit is a use of Johnson S_U distribution which accommodates more than 95 % of the signal events. Johnson S_U PDF description can be found in Ref. [10]. The family of these PDFs is described in the following technical report in Ref. [11]. Furthermore the PRD background component MC sample has been split in 3 sub-components based on a study of ranking individual decay modes by their relative abundances (as seen in the fit mass window) and shape consistency (assessed by a χ^2 test). These 3 PRD1-3 subsamples allow for more accurate modeling of the PRD shape in data. Similarly $B^{\pm} \to J/\psi K^{\pm}$ signal MC sample has been split in two fit sub-samples. These are phrased as radiative and non-radiative signal components. The radiative contribution to $B^{\pm} \to J/\psi K^{\pm}$ decays is formed by cases when the B radiates a γ. Such radiative shape is skewed on the left and to have an accurate fit to the signal, we needed to consider this component separately. All other signal decays are falling into the non-radiative signal decay category. In N3 trigger category we have then 2 $B^{\pm} \to J/\psi K^{\pm}$ signal, 1 $J/\psi\pi^{+}$ and 3 PRD1-3 Monte Carlo samples which are feeded into the fitter together with the data sample resulting in this case in seven dimensional unbinned maximum likelihood fit. Figures 9 and 10 show the data projections of the fit result in the different data categories respectively. The projections on each MC fit sample for N3 trigger category are shown as an example in Figure 8 and the resulting parameters are shown in Table 7.

Systematic Uncertainties

The systematic uncertainties were assessed as coming from two main sources. (1) from the fit assumptions on the chosen fit models and (2) from the data-MC discrepancies and detector effects. Some of the systematic effects of type (2) are taken care of automatically in the fit. The MC effect of limited statistics, for example, is included in the statistical fit error from the simultaneous fit.

In addition the data-MC discrepancy in the mass scale and resolution are extracted as additional fit parameter values and hence are included in the fit in all data models consistently. All other systematic uncertainties are evaluated repeating the whole fit procedure again for each systematic effect. The variation of each such separate systematic fit result from the measured default fit result is then taken as systematic uncertainty. In brief, the data-MC residual discrepancies are assessed by MC signal and MC $J/\psi\pi^+$ sample reweighting using GLC and DDW weights. The PRD decay mode composition is also reweighted to the PDG expected relative abundances in one of the systematic studies. Finally there is many fit shape assumption variations for which systematic effects were evaluated by repeating the fit varying the fit models. As an example of the breakdown of the the systematic uncertainties for category N3 can be found in Table 6.

Systematic uncertainties in trigger category N3				
	Systematic	**Signal**	$J/\psi\pi^{\pm}$	**ratio**
nr.1	MC reweighting (QLC&DDW)	0.41%	1.42%	1.84%
nr.2	PRD re-weighting	0.63%	9.50%	8.93%
nr.3	PRD3 alternate model	0.07%	0.55%	0.48%
nr.4	Combinatorial alternate model	0.09%	12.50%	12.57%
nr.5	Signal peak charge asymmetry	0.29%	7.13%	7.40%
nr.6	PRD1&2 alternate models	0.03%	1.02%	1.05%
	Total	0.81%	17.34%	17.24%

Table 6: Relative changes with respect to the default fit obtained with each systematic check described in the text. The total effect is given both in relative effect and in the absolute number of events (or value of pi/K ratio). Systematic uncertainties in N3 category is shown.

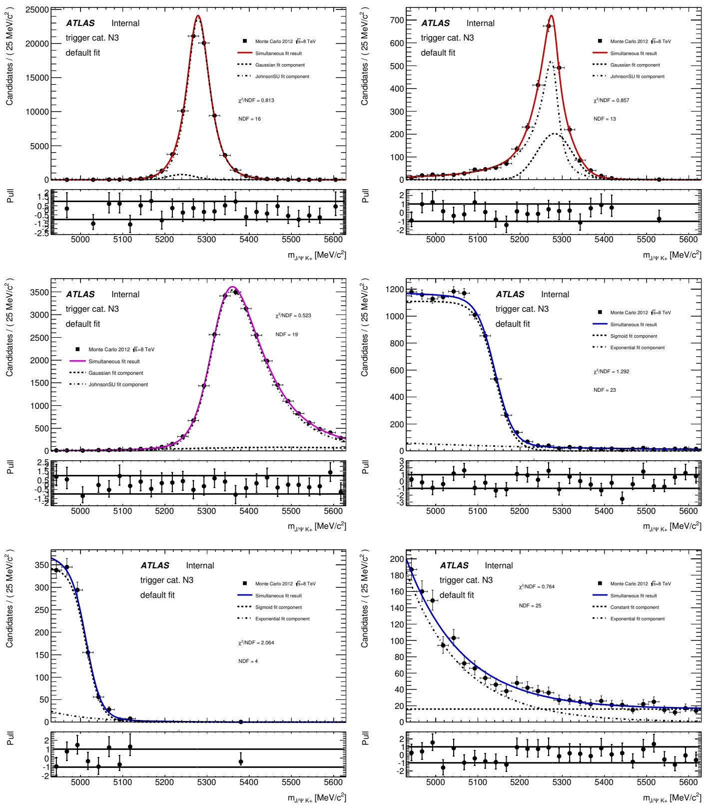


Figure 8: Fit projections on the MC samples simultaneous fitted with the data for N3 trigger category. From left to right, from top to bottom: non-radiative $B^{\pm} \to J/\psi K^{\pm}$ signal, radiative $B^{\pm} \to J/\psi K^{\pm}$ signal, peaking background $J/\psi\pi$, PRD1, PRD2 and PRD3. The red line is used for the $B^{\pm} \to J/\psi K^{\pm}$ signal (including both radiative and non-radiative components), while the magenta line is for the $J/\psi\pi$ peaking component. The blue lines refer to all the three partially reconstructed contributions. In each plot, the black non-continuous lines show the single functions of the total PDF used to model the given component.

N_{Signal}	27272 ± 231
N_{JPSIPI}	1145 ± 173
N_{PRD}	8878 ± 674
N_{Comb}	9568 ± 885
s_{μ}	$-0.0 \pm 0.3 MeV$
s_{σ}	$9.6 \pm 0.8 MeV$

Signal
Johnson S_U + Gaussian

ξ	5276.3 ± 0.5
λ	39.4 ± 0.7
δ	1.795 ± 0.043
γ	-0.183 ± 0.022
μ	5238.8 ± 10.5
σ	32.4 ± 5.0

Radiative signal
Johnson S_U + Gaussian

ξ	5280.6 ± 2.1
λ	18.6 ± 2.7
δ	0.381 ± 0.080
γ	0.483 ± 0.057
μ	5281.6 ± 5.3
σ	41.3 ± 2.7

Signal pdf fractions

Signal f_A	0.9747 ± 0.0027
Signal f_B	0.0115 ± 0.0027
Signal f_C	0.0333 ± 0.0148

$J/\psi\pi$
Johnson S_U + Gaussian

ξ	5313.8 ± 4.2
λ	80.5 ± 5.2
δ	2.031 ± 0.317
γ	-1.607 ± 0.216
μ	5276.3 ± 0.5
σ	248.8 ± 38.7
Gaussian frac	0.057 ± 0.020

PDR1 Fermi-Dirac + Exponential

μ_{FD}	5140.4 ± 1.3
α_{FD}	21.4 ± 1.0
FD frac	0.920 ± 0.011
a	-0.0023 ± 0.0006

PDR2 Fermi-Dirac + Exponential

μ_{FD}	5013.0 ± 3.3
α_{FD}	17.7 ± 2.5
FD frac	0.925 ± 0.056
a	-0.0103 ± 0.0032

PDR3 Exponential + constant

a	-0.0075 ± 0.0006
Expo frac	0.68 ± 0.04

PRD fractions

PRD f_a	0.892 ± 0.003
PRD f_b	0.111 ± 0.003

Combinatorial Exponential

a	-0.00206 ± 0.00023

Yields of control samples

$N^{\text{ctl}}_{\text{NON-RadiativeSignal}}$	72654 ± 270
$N^{\text{ctl}}_{\text{JPSIPI}}$	25204 ± 159
$N^{\text{ctl}}_{\text{PRD}_{\text{tot}}}$	12810 ± 113

Table 7: Results for the parameters of the fit to the N3 category.

.

Result of the $B^\pm \to J/\psi K^\pm$ Yield Extraction

The $B^\pm \to J\psi K^\pm$ reference channel yield has been measured with full systematic uncertainty evaluation in all 4 measurement categories with a result in Table 8. Figures 9 and 10 show the data projections of the fit result in the different data categories respectively.

Measured reference channel yield			
Trigger Category	**Yield**	**stat.uncert.**	**syst.uncert**
N1	1237	±50	±12
N2	2481	±63	±23
N3	27272	±231	±221
N2011	61507	±346	±519

Table 8: Result of the reference channel yield measurement in the three trigger categories.

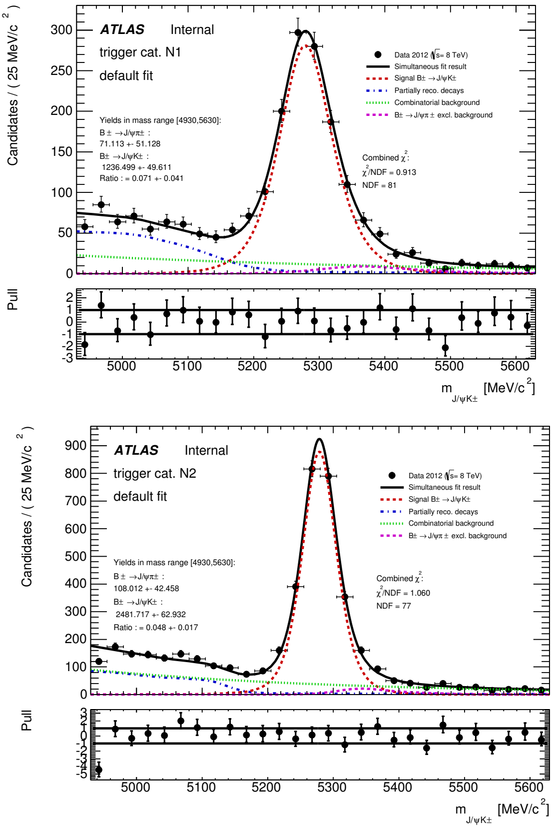


Figure 9: Fit projection on data for N1 trigger category (top) and for the N2 category (bottom). The red line represents the $B^{\pm} \to J/\psi K^{\pm}$ signal (including both radiative and non-radiative components), while the magenta line represents the $J/\psi\pi$ peaking component. The blue line shows all the three partially reconstructed contributions and the green line represents the combinatorial background. The total of all functions is presented with the black line.

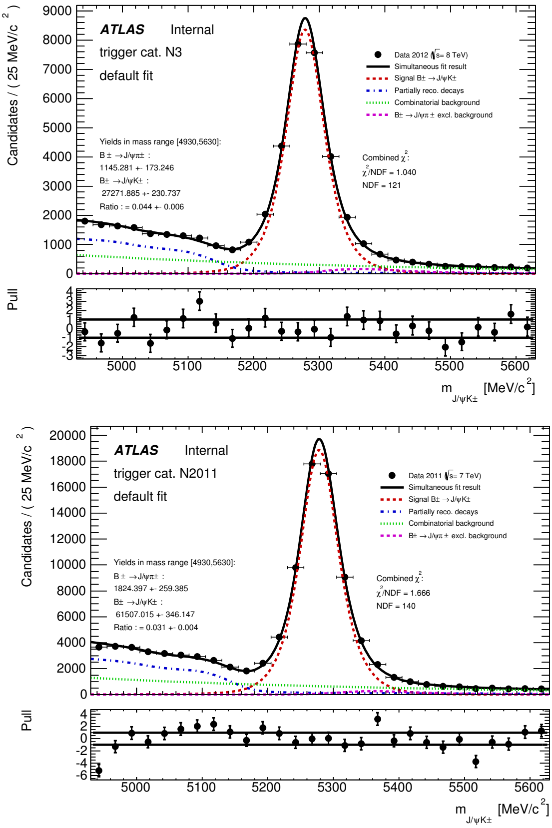


Figure 10: Fit projection on data for N3 trigger category (top) and for the 2011 data (bottom). The red line represents the $B^{\pm} \to J/\psi K^{\pm}$ signal (including both radiative and non-radiative components), while the magenta line represents the $J/\psi\pi$ peaking component. The blue line shows all the three partially reconstructed contributions and the green line represents the combinatorial background. The total of all functions is presented with the black line.

6 $\mathcal{BR}$ $(B^0_{(s)} \to \mu^+\mu^-)$ Extraction

The master equation for the $B^0_s \to \mu^+\mu^-$ Branching Ratio measurement in Equation (1) gives a straight recipe on how to extract the actual measured value from all inputs presented in the previous chapters. This is happening once the analysis has unblinded the signal mass window and measured the signal yield $N_{\mu^+\mu^-}$ on the data. At the time of writing up this thesis the unblinding was not yet performed and by easily reverting the Equation (1) one can estimate the number of expected signal events assuming the SM branching ratio of the signal. This yields 54 $B^0_s \to \mu^+\mu^-$ events, which will be assumed as a test number extracted from Full Run I Data for this exercise. A relative error 26% corresponding to the expected $B^0_s \to \mu^+\mu^-$ signal fit yield error is assigned to $N_{\mu^+\mu^-}$ giving an estimate of 54 ± 14 events. The branching ratio formula can be rewritten to the following form:

$$\mathcal{BR}(B^0_{(s)} \to \mu^+\mu^-) = \frac{\mathcal{F}_{\text{ext}} \times N_{\mu^+\mu^-}}{\mathcal{D}_{\text{norm}}} \tag{3}$$

Apart of the signal yield, there are several external inputs, product of which $\mathcal{F}_{\text{ext}}$ enters the Equation (1):

- **Reference channel**
  $\mathcal{BR}(B^\pm \to J/\psi K^\pm) = (1.027 \pm 0.031) \times 10^{-3}$
  and $\mathcal{BR}(J/\psi \to \mu^+\mu^-) = (5.961 \pm 0.033) \times 10^{-2}$
  both from PDG Ref. [9]

- **Relative hadronisation probability** f_u/f_s[11] $= 0.259 \pm 0.015$ from LHCb experiment Ref. [13] assuming $f_d/f_u = 1$

The external term $\mathcal{F}_{\mathrm{ext}}$ carries in total 6.6% of relative uncertainty as evaluated below:

$$\begin{aligned} \mathcal{F}_{\mathrm{ext}} &= \mathcal{BR}(B^\pm \to J/\psi K^\pm \to \mu^+\mu^- K^\pm) \times \tfrac{f_u}{f_s} = \\ &= (2.36 \pm 0.15) \times 10^{-4} \end{aligned} \tag{4}$$

In the denominator of the master Equation (1) the $\mathcal{D}_{\mathrm{norm}}$ term contains efficiency, acceptance and luminosity weighted number of events extracted for the reference channel.

$$\mathcal{D}_{\mathrm{norm}} = \sum_k N^k_{J/\psi K^\pm} \alpha_k \left(\frac{A_{\mu^+\mu^-}}{A_{J/\psi K^\pm}} \frac{\epsilon_{\mu^+\mu^-}}{\epsilon_{J/\psi K^\pm}} \right)^k = \sum_k \frac{N^k_{J/\psi K^\pm} \alpha_k}{(R_{A\epsilon})^k} \tag{5}$$

The sum in $k = N1, N2, N3, 2011$ spans the four measurement categories. Note that the $R_{A\epsilon}$, measured as described in the thesis, is the inverse of what enters the first equality. The inputs for this denominator collected from the corresponding sections in the thesis are summarized in Table 9. The total relative uncertainty is about 10% on $\mathcal{D}_{\mathrm{norm}}$ which has the following value:

$$\mathcal{D}_{\mathrm{norm}} = \sum_k \frac{N^k_{J/\psi K^\pm} \alpha_k}{(R_{A\epsilon})^k} = (3.51 \pm 0.34) \times 10^6$$

The final result then matches the SM expectation ($\mathcal{BR}(B^0_{(s)} \to \mu^+\mu^-) = 3.66 \times 10^{-9}$) with which we started this exercise when

[11] The dependence of the f_u/f_s ratio on the decay kinematic is found to be negligible for this analysis Ref. [12].

Input values to formula 5.			
Category	Yield	α_k	$(\mathcal{A} \times \epsilon)$ ratio
N1	1237 ± 51	7.23	0.073 ± 0.017
N2	2481 ± 67	7.28	0.089 ± 0.012
N3	27272 ± 320	7.29	0.078 ± 0.010
2011	61507 ± 624	1	0.097 ± 0.013

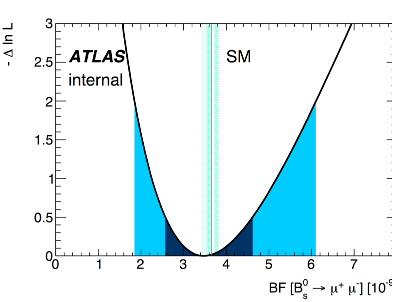
Table 9: Inputs needed for formula 5: the $B^{\pm}$ yields from Table 8 with the statistical and systematic errors summed in quadrature, α_k (see Section 2) factors coming from the category-by-category ratios of the total luminosities.$(\mathcal{A} \times \epsilon)$ ratio is the efficiency times acceptance ratio $(1/R_{A\epsilon})$

estimating the number of expected $N_{\mu^+\mu^-} = 54 \pm 14$ events and is obtained as:

$$\mathcal{BR}(B^0_{(s)} \to \mu^+\mu^-) = \frac{\mathcal{F}_{\rm ext} \times N_{\mu^+\mu^-}}{\mathcal{D}_{\rm norm}} = (3.63 \pm 1.04) \times 10^{-9}$$

where the relative uncertainty is about 29%.

Scanning the signal yield parameter of the mass fit likelihood used to extract the branching ratio, one can obtain also the 1-D likelihood shape which was tested on a mock dataset and the result can be seen in Figure 11. The right plot of in Figure 11 shows the 2-dimensional likelihood scan (B^0_s vs B^0) on another mock dataset. The contours correspond to the 68/95/99.7% 2D probabilities. These plots are only illustration of the final result of the analysis.



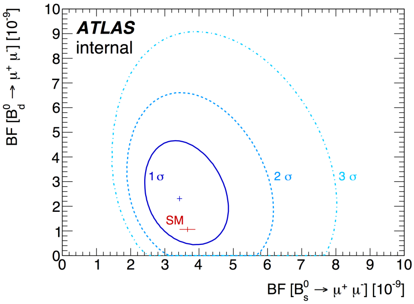


Figure 11: *Left:* 1D likelihood scan on a mock dataset. *Right:* 2D likelihood scan on another mock dataset.

7 $\frac{\Gamma(B^{\pm}\to J/\psi\pi^{\pm})}{\Gamma(B^{\pm}\to J/\psi K^{\pm})}$ ratio measurement

From the fit described above, we extract both the yields for $B^{\pm} \to J/\psi K^{\pm}$ and $B^{\pm} \to J/\psi\pi^{\pm}$ and the systematic checks record the variation on both yields - as in Table 6 for N3 trigger category. In each data category we take the ratio of the yields:

$$\mathrm{R}_{\pi/\mathrm{K}} = \frac{N_{J/\psi\pi^{\pm}} \times I_{ext}^{J/\psi\pi}}{N_{J/\psi K^{\pm}}} \tag{6}$$

where the $J/\psi\pi^{\pm}$ yield is corrected by the factor $I_{ext}^{J/\psi\pi}$ explained below. Then we define the ratio:

$$\rho_{\pi/K} = \frac{\mathcal{BR}(B^{\pm} \to J/\psi\pi^{\pm})}{\mathcal{BR}(B^{\pm} \to J/\psi K^{\pm})} = \frac{N_{J/\psi\pi^{\pm}} \times I_{ext}^{J/\psi\pi}}{N_{J/\psi K^{\pm}}} \times \frac{\epsilon_{J/\psi K^{\pm}}}{\epsilon_{J/\psi\pi^{\pm}}} =$$
$$= \mathrm{R}_{\pi/\mathrm{K}} \times \left[\frac{\epsilon_{K+}}{\epsilon_{\pi+}} \times \frac{1 + \frac{\epsilon_{K-}}{\epsilon_{K+}}}{1 + \frac{\epsilon_{\pi-}}{\epsilon_{\pi+}}} \right] \tag{7}$$

where N_X is the yield for channel X ($J/\psi\pi^{\pm}$, $J/\psi K^{\pm}$), and ϵ_X is the efficiency-times-acceptance product for channel X. In the last equality we have used the asymptotic relation $\epsilon_h^{\pm} = \frac{\epsilon_h^+ + \epsilon_h^-}{2}$, and $\frac{\epsilon_{K-}}{\epsilon_{K+}}$ and $\frac{\epsilon_{\pi-}}{\epsilon_{\pi+}}$ are the kaon and pion charge asymmetries, respectively.

The fit for the reference channel of Section 5 extracts both yields $J/\psi K^{\pm}$ and $J/\psi\pi^{\pm}$ from the fit mass window $[4930., 5630.]$ MeV/c^2. While the $J/\psi K^{\pm}$ component is accommodated entirely within our mass fit window, the $J/\psi\pi^{\pm}$ component extends by a small fraction outside the right boundary. The fraction $I_{ext.}^{J/\psi\pi}$ of $J/\psi\pi^{\pm}$ candidates counted in the extended region $[3500., 7000.]$

MeV/c^2 with respect to the candidates counted in the default fit mass window is taken as a correction factor of the yield extracted from the fit. The result of such calculation $I^{J/\psi\pi}_{ext.}$ together with the corrected yield is summarized in Table 10.

	$I^{J/\psi\pi}_{ext.}$	$N_{J/\psi\pi\pm} \times I^{J/\psi\pi}_{ext.}$
N1	1.039 ± 0.135	74 ± 53
N2	1.006 ± 0.290	108 ± 43
N3	1.043 ± 0.032	1195 ± 181
2011	1.039 ± 0.031	1896 ± 270

Table 10: Values for the $I^{J/\psi\pi}_{ext.}$ correction factor and the corrected yields (stat. uncertainties only) for the $B^{\pm} \to J/\psi\pi\pm$ as used in the ratio.

Regarding the $\frac{\epsilon_{K\pm}}{\epsilon_{\pi\pm}}$ ratio, three are the factors contributing: the kaon and pion charge asymmetries ($\frac{\epsilon_{K-}}{\epsilon_{K+}}$, $\frac{\epsilon_{\pi-}}{\epsilon_{\pi+}}$) and the relative K^+/π^+ efficiency $\frac{\epsilon_{K+}}{\epsilon_{\pi+}}$. For the pion charge asymmetry we assume the central value to be $\frac{\epsilon_{\pi-}}{\epsilon_{\pi+}} = 1$ and we assign an uncertainty to it as discussed below. We evaluate the remaining two factors using Monte Carlo and estimate systematic uncertainties as described below.

The four measurements of $\rho_{\pi/K}$ in the separate data categories are then combined to minimise the measurement uncertainty.

Most systematic effects (luminosity, trigger, reconstruction efficiencies) cancel in the measurement of this ratio due to the almost identical topology and kinematics of the two decay channels. Residual systematic uncertainties on the ratio of branching fractions come from uncertainties on the parametrisation of the fit PDFs, data-MC discrepancies, the K^-/K^+ and π^-/π^+ charge asymmetries, and the K^+/π^+ relative efficiency. Introducing scaling factors $I_{ext.}^{J/\psi\pi}$ to account for the (1−4%) of $J/\psi\pi^\pm$ events falling outside of the fit window does not introduce significant source of systematic uncertainties.

Table 11 contains the systematic contributions to the final averaged $\overline{\rho_{\pi/K}}$ ratio from the study performed for the reference channel yield extraction described in Section 5. In addition, a number of ratio-specific systematic checks need to be performed on the kaon and pion charge asymmetries and the relative K^+/π^+ efficiency:

K^-/K^+ **charge asymmetry** (0.08%) The K^-/K^+ charge asymmetry is measured on $B^\pm \to J/\psi K^\pm$ MC. The uncertainty on the charge asymmetry is driven by the data-MC discrepancies on the kaon kinematics which are connected in the $B^\pm \to J/\psi K^\pm$ decay by the data-MC discrepancies on the B meson kinematics. By reweighting our MC sample using GLC and DDW weights (described in the thesis) we re-evaluate the charge asymmetry and take the resulting variation on $\epsilon_{K^-}/\epsilon_{K^+}$ as systematic uncertainty.

π^-/π^+ **charge asymmetry** (0.8%) The π^-/π^+ charge asymmetry is assumed to be $= 1$, compatible within statistical uncertainty with what predicted from MC. We estimate the systematic uncertainty on $\epsilon_{\pi^-}/\epsilon_{\pi^+}$ comparing the central value of the $\epsilon_{K^-}/\epsilon_{K^+}$ (described above) with $\frac{\epsilon_{K^-}/\epsilon_{K^+}}{\epsilon_{\pi^-}/\epsilon_{\pi^+}}$ obtained from the $B \to h^-h^+$ MC12 sample (the highest statistics π/K simulation we have with hadron spectra similar to $J/\psi\pi^\pm/J/\psi K^\pm$). Hadron selection on $B \to h^-h^+$ events are kept as close as possible to those on the B^+ signal selection. We then assume $\frac{\epsilon_{B_d \to K^- \pi^+}}{\epsilon_{B_d \to K^+ \pi^-}} \approx \frac{\epsilon_{K^-}/\epsilon_{K^+}}{\epsilon_{\pi^-}/\epsilon_{\pi^+}}$. The π^-/π^+ charge asymmetry is then estimated as :

$$\epsilon_{\pi^-}/\epsilon_{\pi^+} = \frac{\epsilon_{B_d \to K^+ \pi^-}}{\epsilon_{B_d \to K^- \pi^+}} \times \frac{\epsilon_{K^-}}{\epsilon_{K^+}} = 1.008 \pm 0.005 \qquad (8)$$

The full difference from 1 is taken as our systematic uncertainty on $\epsilon_{\pi^-}/\epsilon_{\pi^+}$.

K^+/π^+ **relative efficiency** (3.24%) $\epsilon_{K^+}/\epsilon_{\pi^+}$ is measured on GLC- and DDW-weighted $B^\pm \to J/\psi K^\pm$ MC sample using the same machinery as $\frac{\epsilon_{B^+ \to J/\psi K^+}}{\epsilon_{B_s \to \mu^+ \mu^-}}$ in the main analysis. Discrepancies on this parameter arise predominantly from residual data-MC discrepancies in the B spectrum model.

The final efficiency ratios entering in all categories can be found in Table 11. For comparison and crosscheck, we report also the default fit results for $\frac{N_{J/\psi\pi^-}}{N_{J/\psi\pi^+}}$ and $\frac{N_{J/\psi K^-}}{N_{J/\psi K^+}}$.

Table 12 reports the yield ratio $\mathrm{R}_{\pi/\mathrm{K}}$ with its statistical uncertainty for each data category. After the efficiency correction, the measurement of BR ratio $\rho_{\pi/K}$ with correctly propagated statistical uncertainties can be found in last two columns of the same Table 12. To combine the measurements, we use the squared inverse value of statistical uncertainty on each measurement as a weight and calculate the weighted mean $\overline{\rho_{\pi/K}}$. To evaluate the systematic uncertainty on the result $\overline{\rho_{\pi/K}}$ we re-evaluate the combination for each systematic variation, therefore accounting for correlated effects. The difference with the default value $\overline{\rho_{\pi/K}}$ is taken as the combined systematic uncertainty for each effect. Systematic un-

MC	N1	N2	N3	N2011
$\frac{\epsilon_{J/\psi K+}}{\epsilon_{J/\psi\pi+}}$	$1.109 \pm 0.038 \pm 0.021$	$1.024 \pm 0.023 \pm 0.019$	$1.141 \pm 0.009 \pm 0.005$	$1.130 \pm 0.008 \pm 0.006$
$\frac{\epsilon_{J/\psi K-}}{\epsilon_{J/\psi K+}}$	$0.966 \pm 0.020 \pm 0.001$	$0.973 \pm 0.015 \pm 0.001$	$0.975 \pm 0.005 \pm 0.002$	$0.974 \pm 0.005 \pm 0.002$
$\frac{\epsilon_{J/\psi\pi-}}{\epsilon_{J/\psi\pi+}}$	$1. \pm 0.005 \pm 0.008$			

DATA	N1	N2	N3	N2011
$\frac{N_{J/\psi K-}}{N_{J/\psi K+}}$	1.07 ± 0.09	1.006 ± 0.052	0.964 ± 0.016	0.966 ± 0.011
$\frac{N_{J/\psi\pi-}}{N_{J/\psi\pi+}}$	0.20 ± 0.65	1.95 ± 1.82	0.79 ± 0.24	0.69 ± 0.20

Table 11: Kaon charge asymmetry $\frac{\epsilon_{K-}}{\epsilon_{K+}}$ and the relative K^+/π^+ efficiency $\frac{\epsilon_{K+}}{\epsilon_{\pi+}}$ measured on MC as described in the text. The bottom part of the table shows for reference the $\frac{N_{J/\psi\pi-}}{N_{J/\psi\pi+}}$ and $\frac{N_{J/\psi K-}}{N_{J/\psi K+}}$ ratios as extracted from the data fit. First uncertainty is statistical, second (if reported) is systematic.

certainties obtained this way are summarised in Table 13 and they are summed in quadrature to obtain the combined uncertainty.

trig.cat.	$\mathrm{R}_{\pi/\mathrm{K}}$	$\sigma^{stat.}_{\mathrm{R}_{\pi/\mathrm{K}}}$	$\rho_{\pi/K}$	$\sigma^{stat.}_{\rho_{\pi/K}}$
N1	0.0598	$\pm$ 0.0430	0.0652	$\pm$ 0.0469
N2	0.0437	$\pm$ 0.0174	0.0441	$\pm$ 0.0176
N3	0.0438	$\pm$ 0.0066	0.0494	$\pm$ 0.0075
2011	0.0308	$\pm$ 0.0044	0.0344	$\pm$ 0.0049
Weighted average	–	–	0.0393	$\pm$ 0.0040

Table 12: Second and third columns: uncorrected relative $J/\psi\pi$ / $J/\psi K$ yield measured in the four data categories (statistical errors only). Fourth and fifth columns: fit result for $\rho_{\pi/\mathrm{K}}$ in all categories with statistical errors.

The largest systematic uncertainty on the measured ratio comes from the combinatorial background model parametrisation ($\approx 21\%$), followed by the effect of PRD reweighting ($\approx 14\%$), by the $B^+ \to J/\psi K^+$ signal peak shape charge asymmetry ($\approx 5\%$) and by the effect of the radiative tails in the signal models ($\approx 5\%$). All other systematic sources have minor effects ($\approx 3\%$ or less).

systematic effect	$\sigma^{syst.}_{\overline{\rho_{\pi/\mathrm{K}}}}$
Combinatorial model	20.62%
PRD reweighting	14.52%
Signal peak charge asymmetry	4.94 %
RAD tails in signal models	4.61%
K^+/π^+ **relative efficiency**	3.24%
PRD1&2 parametrisation	2.26 %
MC reweighting (QLC&DDW)	1.68%
PRD3 parametrisation	1.08%
π^-/π^+ **charge asymmetry**	0.82 %
K^-/K^+ **charge asymmetry**	0.08 %
Total	26.5%

Table 13: Relative systematic uncertainties on the $\overline{\rho_{\pi/\mathrm{K}}}$ measurement after combination of the four data categories. The middle column corresponds to considered systematic effect and the right column reports the systematic uncertainty.

The final result on the ratio of branching fractions $\frac{\mathcal{BR}(B^\pm \to J/\psi\pi^\pm)}{\mathcal{BR}(B^\pm \to J/\psi K^\pm)}$ is:

$$\overline{\rho_{\pi/K}} = (3.9 \pm 0.4^{stat.} \pm 1.0^{syst.})\% \tag{9}$$

8 Summary

Two of the main inputs to the $\mathcal{BR}(B^0_{(s)} \to \mu^+\mu^-)$ measurement formula (see Equation (10)) contain my direct contribution. The extraction of the $B^0_{(s)} \to \mu^+\mu^-$ candidate event count $N_{\mu^+\mu^-}$ was cleaned from the dangerouns peaking background contribution and the yield of the $N^k_{J/\psi K^\pm}$ reference channel was measured on the Full Run I dataset in all 4 trigger and data categories.

$$\mathcal{BR}(B^0_{(s)} \to \mu^+\mu^-) = \mathcal{BR}(B^\pm \to J/\psi K^\pm \to \mu^+\mu^- K^\pm) \times \frac{f_u}{f_s} \times \\ \times N_{\mu^+\mu^-} \times \left(\sum_k N^k_{J/\psi K^\pm} \alpha_k \frac{(A\varepsilon)^k_{\mu^+\mu^-}}{(A\varepsilon)^k_{J/\psi K^\pm}} \right)^{-1} \tag{10}$$

The developed unbinned simultaneous maximum likelihood fit to extract the reference channel yield would have been unstable and less accurate without the model for the $J/\psi\pi^\pm$ component. In addition the fit has shown high sensitivity to this contribution which is extracted in paralel to the main reference channel yield as a part of the fit result. Thus, $\frac{\Gamma(B^\pm \to J/\psi\pi^\pm)}{\Gamma(B^\pm \to J/\psi K^\pm)}$ has been measured and shall be published in parallel to the $B^0_{(s)} \to \mu^+\mu^-$ main analysis results in the same paper. The purpose of this measurement is also to provide a sanity check of the reference channel fit by comparison against the PDG average (consistency found with all measurements therein). As a final word of conclusion let me note that the BDT for peaking background rejection as well as the reference channel yield fit have both found applications in other physics analyses not described in the thesis. In the following the results are briefly summarised.

Result of Peaking Background Rejection

Table 14 shows the reduction in the fraction of fake hadrons obtained with a BDT threshold corresponding to a single signal muon selection efficiency equal to 95%. The fake fraction would be further reduced by a factor $\simeq 0.8$ if the selection would be tuned to 90% single signal muon efficiency. The errors shown in Table 14 are related to statistical fluctuation in the evaluation sample.

$K^{\pm}$	0.376±0.007
$\pi^{\pm}$	0.366±0.010

Table 14: Fake muon reduction factors obtained with the BDT selection, for 95% muon efficiency, with statistical uncertainty.

Including the factor of 3.9 between B^0 and B_s^0 production cross sections (f_d/f_s), the total peaking background corresponds to an effective branching fraction for B_s^0 equal to about 6 $\times 10^{-11}$. The peaking background after all selection cuts (including the cut on the BDT for *fake* muons rejection) was estimated to be $1.0^{+0.8}_{-0.5}$ signal candidate. The negative uncertainty is a -0.5 conservative estimate of possible uncertainties in the MC modelling of the selection against fake muons, and the positive one is extracted from the analysis on additional studies performed on both the $B^{\pm} \to J/\psi K^{\pm}$ Data and on $B^0_{(s)} \to \mu^+\mu^-$ sideband Data.

Result of Reference Channel Yield

The $B^{\pm} \to J\psi K^{\pm}$ reference channel yield as it enters the $B^0_{(s)} \to \mu^+\mu^-$ branching ratio master formula, was measured with full systematic uncertainty evaluation in all 4 measurement categories with a result in Table 15. The fit result projected on the data mass distributions can be seen in Figure 9 for N1 and N2 and Figure 10 for N3 and 2011 measurement categories.

Measured reference channel yield			
Trigger Category	**Yield**	**stat.uncert.**	**syst.uncert**
N1	1237	±50	±12
N2	2481	±63	±23
N3	27272	±231	±221
N2011	61507	±346	±519

Table 15: Result of the reference channel yield measurement in the three trigger categories.

Result of $\frac{\Gamma(B^{\pm}\to J/\psi\pi^{\pm})}{\Gamma(B^{\pm}\to J/\psi K^{\pm})}$ measurement

The ratio of branching fractions $\frac{\mathcal{BR}(B^{\pm}\to J/\psi\pi^{\pm})}{\mathcal{BR}(B^{\pm}\to J/\psi K^{\pm})}$ was measured to be:

$$\overline{\rho_{\pi/K}} = (3.9 \pm 0.4^{stat.} \pm 1.0^{syst.})\% \tag{11}$$

The PDG [9] value is 4.0 ± 0.4 %.

Personal Contributions

[1] *Limit on $B_s \to \mu\mu$ based on 2.4 fb^{-1} of integrated luminosity*, Tech. Rep. ATL-COM-PHYS-2011-1619, CERN, Geneva, Nov, 2011.

[2] ATLAS Collaboration, G. Aad et al., *Search for the decay Bs0 -> mu mu with the ATLAS detector*, Phys.Lett. **B713** (2012) 387–407, `arXiv:1204.0735 [hep-ex]`.

References

[1] D. Hanneke, S. Fogwell, and G. Gabrielse, *New Measurement of the Electron Magnetic Moment and the Fine Structure Constant*, Phys. Rev. Lett. **100** (Mar, 2008) 120801. `http://link.aps.org/doi/10.1103/PhysRevLett.100.120801`.

[2] C. Bobeth, M. Gorbahn, T. Hermann, M. Misiak, E. Stamou, et al., *$B_{s,d} \to l^+l^-$ in the Standard Model with Reduced Theoretical Uncertainty*, Phys.Rev.Lett. **112** (2014) 101801, `arXiv:1311.0903 [hep-ph]`.

[3] O. Witzel, *B-meson decay constants with domain-wall light quarks and nonperturbatively tuned relativistic b-quarks*, PoS **LATTICE2013** (2014) 377, `arXiv:1311.0276 [hep-lat]`.

[4] H. Na, C. Monahan, C. Davies, E. Follana, R. Horgan, et al., *Precise Determinations of the Decay Constants of B and D mesons*, PoS **LATTICE2012** (2012) 102, `arXiv:1212.0586 [hep-lat]`.

[5] A. Bazavov, C. Bernard, C. M. Bouchard, C. DeTar, M. Di Pierro, A. X. El-Khadra, R. T. Evans, E. D. Freeland, E. Gámiz, S. Gottlieb, U. M. Heller, J. E. Hetrick, R. Jain, A. S. Kronfeld, J. Laiho, L. Levkova, P. B. Mackenzie, E. T. Neil, M. B. Oktay, J. N. Simone, R. Sugar, D. Toussaint, and R. S. Van de Water, *B- and D-meson decay constants from three-flavor lattice QCD*, Phys. Rev. D **85** (Jun, 2012) 114506. `http://link.aps.org/doi/10.1103/PhysRevD.85.114506`.

[6] V. Kostyukhin, *VKalVrt - package for vertex reconstruction in ATLAS*, ATLAS Note ATL-PHYS-2003-031, 2003.

[7] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss, *TMVA 4: Toolkit for Multivariate Data Analysis*, PoS ACAT **040** (2007) , `arXiv:physics/0703039`.

[8] P. C. Bhat, *Multivariate Analysis Methods in Particle Physics*, Annual Re-
view of Nuclear and Particle Science **61** (2011) no. 1, 281–309,
`http://dx.doi.org/10.1146/annurev.nucl.012809.104427`.
`http:`
`//dx.doi.org/10.1146/annurev.nucl.012809.104427`.

[9] Particle Data Group Collaboration, K. Olive et al., *Review of Particle Physics*, Chin.Phys. **C38** (2014) 090001.

[10] *A User's Guide to the RooFitTools Package for Unbinned Maximum Likelihood Fitting*,
https://dl.dropboxusercontent.com/u/14841111/vertexingMuons.pdf,
2001.

[11] M. Jones and A. Pewsey, *Sinh-arcsinh distributions: a broad family giving rise to powerful tests of normality and symmetrySinh-arcsinh distributions: a broad family giving rise to powerful tests of normality and symmetry*, Tech. Rep. ATL-COM-MUON-2014-001, The Open University, UK; University of Extremadura, Spain, 2008.

[12] LHCb Collaboration, R. Aaij et al., *Measurement of the fragmentation fraction ratio f_s/f_d and its dependence on B meson kinematics*, `arXiv:1301.5286 [hep-ex]`.

[13] T. L. Collaboration, *Updated average f_s/f_d b−hadron production fraction ratio for 7 TeV pp collisions*, `LHCb-CONF-2013-011,CERN-LHCb-CONF-2013-011`.

9 Resume

The ATLAS Collaboration at CERN has been searching for rare B_s and B_d meson decays into two muons and a paper shall be published very soon. These decays are forbidden at the tree level of the Standard Model. They offer the opportunity to perform genuine probes of Yukawa interactions or Electroweak precision tests and play very important role to find signatures of physics beyond the Standard Model. A brief insight into the theoretical foundations of these decays is given in the introductory section of the presented thesis, followed by a section with the ATLAS experiment description. The ATLAS Collaboration has been searching for $B^0_{(s)} \to \mu^+\mu^-$ decays using merged 2011 $\sqrt{s} = 7$ TeV and 2012 $\sqrt{s} = 8$ TeV Full Run I Data sample ($\approx 25 fb^{-1}$). The analysis procedure has been firmly established and unblinding of the search region of $B^0_{(s)} \to \mu^+\mu^-$ is imminent. A sensitivity of the analysis to the $B^0_s \to \mu^+\mu^-$ signal is estimated to be $4.7 \pm 1.0\sigma$. The description of the whole ATLAS $B^0_{(s)} \to \mu^+\mu^-$ analysis procedure is given in the third section, where a summary of author's contributions is described as well. The final three sections of the thesis describe in great detail explicitly author's contributions to the analysis and beyond. The algorithm for the rejection of the ATLAS mis-identified hadrons (as muons) has proven to be very useful not only to separate the signal from the almost indistinguishable peaking background for the $B^0_{(s)} \to \mu^+\mu^-$ analysis, but found use also in other physics analysis. Secondly, the branching ratio $\mathcal{BR}(B^0_{(s)} \to \mu^+\mu^-)$ measurement on ATLAS is performed with respect to a reference channel decay $B^\pm \to J/\psi K^\pm$, the yield of which was extracted with a very good accuracy. Finally, as a natural outcome from the reference channel yield extraction, a measurement of $\Gamma(B^\pm \to J/\psi\pi^\pm)/\Gamma(B^\pm \to J/\psi K^\pm)$ has been found competitive with other measurements and performed in parallel to the main $B^0_{(s)} \to \mu^+\mu^-$ analysis on the Full Run I Data with the result of $3.9 \pm 0.4^{stat.} \pm 1.0^{syst.}\%$.