Very Rare Decays at LHCb

Serena Oggero, on behalf of the LHCb Collaboration
Nikhef, Science Park 105 1098 XG Amsterdam, The Netherlands
E-mail: soggero@nikhef.nl

Abstract. Rare leptonic decays of the $B^0_s$, $D$ and $K$ mesons are sensitive probes of physics beyond the Standard Model. In particular, the search for the decays $B^0_s \rightarrow \mu^+\mu^-$ provides information on the presence of new (pseudo-)scalar particles. LHCb is well suited for these analyses due to its large acceptance for $B$-decays and trigger efficiency, as well as its excellent invariant mass resolution and electron and muon identification capabilities. Results on these searches using data collected by LHCb in 2011 and 2012 are presented.

1. Theoretical motivations

In the Standard Model (SM) of particle physics, the weak interaction eigenstates and corresponding mass eigenstates are connected through the Cabibbo-Kobayashi-Maskawa (CKM) matrix. The CKM matrix is not diagonal, therefore transitions between different quark generations (Flavour Changing Currents) are allowed. Among these, processes in which the quark changes flavour but not charge are named Flavour Changing Neutral Currents (FCNC) and are suppressed at tree level in the SM. Charged currents mediated by $W^\pm$ bosons can instead violate flavour, therefore FCNC processes are possible at one-loop and higher order diagrams. At the one-loop level, they can be described by a set of triple effective vertices named “penguin” diagrams, and by a set of quartic ones named “box” diagrams.

The FCNCs suppression is experimentally established and finds a natural explanation in the Glashow-Iliopoulos-Maiani (GIM) mechanism [1]. The hierarchy of FCNC transitions depends on the masses $m_i$ involved in the internal loops. For instance, the GIM mechanism is less effective in suppressing processes governed by top quarks contributions ($K$ and $B$ decays) than those in which only $d$, $s$ and $b$ quarks enter internal loops ($D$ decays). The size of FCNC transitions also depends on the concerned $V_{\text{CKM}}$ elements, e.g. $B_s$ decays are less suppressed than $B_d$ decays, due to the different sizes of the CKM matrix element involved.

In any case, the fact that FCNCs take place only at the higher orders makes them useful to test the SM and particularly fruitful to probe possible models of New Physics (NP). The rare decay branching fraction ($\mathcal{B}$) may receive large enhancements within specific NP models. Moreover, a purely leptonic final state can lead to a clean theoretical
computation of the SM effective Hamiltonian. This motivates the strong interest shown by the physics community towards very rare leptonic decays, despite their challenging analysis aspects. In particular, the LHCb experiment has a very rich rare decays search program, which aims at searching for New Physics with an “indirect” approach, mostly complementary to the direct NP searches conducted at Atlas and CMS.

2. LHCb analyses

The focus of this talk is towards purely leptonic very rare decays investigated at LHCb; in particular, the search for $D^0 \rightarrow \mu^+\mu^-$, $K_S^0 \rightarrow \mu^+\mu^-$ and $B^0 \rightarrow \mu^+\mu^-$ decays will be presented. All these processes have a very clean signature in the LHCb experiment (see Sec. 2.1) and calls for a common general analysis strategy (see Sec. 2.2).

2.1. LHCb experiment

LHCb [2] is a dedicated experiment for the study of heavy flavour physics at the Large Hadron Collider (LHC). Taking advantage of the correlated $b\bar{b}$ pair production at high energies and small angles with respect to the beam axis, the LHCb detector is designed as a single arm spectrometer, with a unique pseudorapidity acceptance of $1.9 < \eta < 4.9$ and a forward angular coverage from approximately 10 mrad to 300 mrad.

The LHCb experiment performs precision measurements in flavour physics thanks to an efficient tracking system, an extensive Particle Identification (PID) system and a highly adaptable trigger. The tracking system consists of a silicon-strip vertex detector (VELO), surrounding the $pp$ interaction region, a large-area silicon-strip detector (Tracker Turicensis, TT) located upstream of a warm dipole magnet with a bending power of about 4 Tm, and three stations of silicon-strip detectors (Inner Tracker, IT) and straw drift tubes (Outer Tracker, OT), placed downstream.

The LHCb PID is based on two aerogel and gas Ring Imaging Cherenkov counters (RICH) for the charged hadrons, a hadron (HCAL) and electromagnetic (ECAL) calorimeter for photon, electron and hadron candidates, and finally a set of muon chambers, composed of alternating layers of iron and multiwire proportional chambers.

The trigger consists of a hardware stage (L0), mainly based on information from the calorimeter and the muon system, and a software stage (HLT), running a full event reconstruction on a farm of parallel-processing CPUs. The L0 takes events at an input frequency rate of 40 MHz, while the HLT reduces the event rate from an input of 1 MHz to an output rate of about 3 kHz.

2.2. Analysis strategy

All rare leptonic decay searches have a common general analysis strategy that can be illustrated as follows.

Selection. At the L0 level, the single muon and di-muon trigger lines are mostly involved; the HLT trigger applies impact parameter (IP) and invariant mass requirements, leading to an overall $\sim 90\%$ trigger efficiency for high-pt muons in the final state. Afterwards, the first selection of events requires daughter particles to make a displaced vertex, by applying cuts on the IP $\chi^2$ of each particle with respect to the primary vertex. All daughter tracks are also required to originate from the same secondary vertex, by requiring a good vertex fit with a cut on its $\chi^2$. Similarly, the
mother hadron has to be consistent with originating from a primary vertex (PV); this is
guaranteed by a cut on the hadron IP $\chi^2$. Other variables typically entering the selection
are the transverse momentum and the tracks quality of the involved particles, the hadron
flight distance and its angle of pointing with respect to the primary vertex. The invariant
mass of the mother hadron is required to fall in a certain range optimised for the
analysis. In addition, particle identification criteria are used to separate between muons
and hadrons. The muon identification efficiency can reach very high values ($\sim 98\%$ for
the $B^{0}_{s} \rightarrow \mu^+ \mu^-$ analysis), with corresponding very low pion and kaon misidentification
rates (less than about $1\%$).

The same selection is applied to both signal and control channels, with the exception
of the particle identification requirements and the invariant mass cuts.

**Background rejection.** To further increase the signal over background separation,
MVA classifiers are used, typically the output of a Boosted Decision Tree (BDT). The
BDT operator is built from kinematic and topological variables, combined on the base of
their performance during training. The BDT is trained on MC events for the signal and
data events from the mass sidebands for the background; the operator optimisation aims
at the maximum separation power between signal and background events for high BDT
values. The signal BDT distribution is re-calibrated on data (typically using hadronic
exclusive decays).

The use of a multivariate selection helps to reduce the combinatorial background,
where a candidate meson-vertex is formed from daughter particles that did not originate
from a single meson. Peaking background components – typically decays with final
state hadrons ($h$) misidentified as muons (“misId”) or exclusive semileptonic decays–
might need specific studies. They are usually statistically separated from the signal by
mean of particle identification criteria. The misidentification rates, estimated with data
calibration samples, are thus a key ingredient of the analysis.

**Normalisation.** The number of signal events is normalised to one or more channels
($norm$), sharing the same kinematics. In this way the branching fraction can be obtained
as

$$B(signal) = \frac{N_{signal}\epsilon_{signal}}{N_{norm}\epsilon_{norm}} \cdot B(norm) \tag{1}$$

where $\epsilon$ represents the efficiency of reconstruction, trigger and selection on the signal
or normalisation channel. The use of a normalisation channel allows for cancellation of
uncertainties in the determination of the investigated branching fration, for instance due
to the integrated luminosity and the meson production cross section. The normalisation
channel is selected with a strategy very similar to that used to select signal events;
this leads to a minimisation of the systematic effects entering the ratio. The remaining
needed efficiency terms are estimated using MC simulations corrected with data-driven
methods.

Equation (1) can also be used to estimate the number of expected signal events for a
given branching fraction hypothesis, needed for the limit computation (see later in this
section).

**Blind approach.** The data in the signal region is only analysed once the full analysis
strategy is defined. This means that a certain region in invariant mass of the daughter
particles is blinded during the analysis optimisation and only used (“unblinded”) for the
computation of the final results.

**Branching fraction estimation and limit determination.** Different statistical techniques are used to assess the result of the rare decay search. To estimate the signal yield and branching fraction, typically an unbinned likelihood fit of the invariant mass spectra is used, where the different background components enter as separate PDF’s of the fit. All the uncertainties on the separate variables are implemented as parameter constraints and contribute to the final uncertainty.

The compatibility of the observation with expectations (significance of the result) and the limit determination in case of no significant excess are computed with the $CL_s$ method. The following sections give an insight on some leptonic rare decays searches conducted at LHCb, with the last published results.

### 3. Search for the $D^0 \rightarrow \mu^+\mu^-$ decay

As already mentioned in Sec. 1, charm hadron decays are most effectively suppressed by the GIM mechanism and in particular the $D^0 \rightarrow \mu^+\mu^-$ decay is expected to be very rare in the SM. In fact, the long distance contribution to $D^0 \rightarrow \mu^+\mu^-$, which dominates its branching fraction, is predicted to have an upper limit of about $6 \times 10^{-11}$ at 90% C.L. [6]. However, sizeable enhancements are expected in several NP models, for instance via tree-level contributions in R-parity violating scenarios.

The analysis conducted at LHCb on 0.9 fb$^{-1}$ collected in 2011 follows the strategy flow illustrated in Sec. 2.2. Signal candidates are determined on a sample of selected $D^{*+} \rightarrow D^0(\rightarrow \mu^+\mu^-)\pi^+$ decays. The most dangerous peaking background after selection is formed by $D^0 \rightarrow \pi^+\pi^-$ decays with the two pions misidentified as muons. The yield of signal candidates observed after selection is determined with a 2 dimensional fit in the $D^0$ mass ($\mu\mu$ invariant mass) and $D^0 - D^{*0}$ mass difference $\Delta m (m_{\mu\mu\pi} - m_{\mu\mu})$. The fit performed is an unbinned extended maximum likelihood fit and comprises a signal PDF, a combinatorial background PDF and two PDF’s describing the misidentified $D^0 \rightarrow \pi^+\pi^-$ and $D^0 \rightarrow K^-\pi^+$ events. The latter does not peak in the signal window but its shape is needed to correctly describe the data sidebands.

Since the fit to the data does not reveal any excess of signal events with respect to the expected background, the $CL_s$ technique is used to determine a limit on $\mathcal{B}(D^0 \rightarrow \mu^+\mu^-)$. The yield of expected signal events is determined by mean of the normalisation channel $D^0 \rightarrow \pi\pi$ (see Eq. 1). The result obtained, $\mathcal{B}(D^0 \rightarrow \mu^+\mu^-) < 1.3(1.1) \times 10^{-8}$ at 95(90)% C.L. [6], improves the existing upper limit set by the Belle collaboration ($\mathcal{B}(D^0 \rightarrow \mu^+\mu^-) < 1.4 \times 10^{-7}$ at 90% C.L. [7]) by more than one order of magnitude. Nevertheless, it is still well above the current SM theoretical prediction, leaving space for future improvements with the full LHCb available dataset.

### 4. Search for the $K_S^0 \rightarrow \mu^+\mu^-$ decay

The expected SM branching fraction for $K_S^0 \rightarrow \mu^+\mu^-$ decays is $\mathcal{B}_{SM}(K_S^0 \rightarrow \mu^+\mu^-) = (5.0 \pm 1.5) \times 10^{-12}$ [8], but NP contributions can enhance the SM expectation of up to one order of magnitude [9].

LHCb is a fruitful place for searching rare kaon decays thanks to the high yields of $K_S^0$ ($\sim 10^{13}$ per fb$^{-1}$) in the detector acceptance. The analysis conducted on 1 fb$^{-1}$ of $pp$ collisions collected in 2011 follows the strategy illustrated in Sec. 2.2. After selection of
di-muon candidates, the invariant mass range \([492-504]\,MeV/c^2\) is defined as the search region. The background level in this region is calibrated by interpolating the observed yield from mass sidebands with an unbinned maximum likelihood fit, accounting for the combinatorial component (described by an exponential function) and the peaking component. The latter is mainly composed by \(K^0_S \to \pi^+ \pi^-\) misidentified events and it is modeled by a power law. For the signal yield determination, \(K^0_S \to \pi^+ \pi^-\) decays are used as normalisation channel.

The observed number of events is found to be compatible with the background expectations and the upper limit is set at \(B(K^0_S \to \mu^+ \mu^-) < 11(9) \times 10^{-9}\) at 95(90)\% [10], improving the previous measurements by a factor of thirty.

5. Search for the \(B^{(s)}_0 \to \mu^+ \mu^-\) decays

A recent evaluation of the branching fraction for \(B^{(s)}_0 \to \mu^+ \mu^-\) decays, and accounting for electroweak calculations up to the leading order and QCD calculations at NLO, predicts [11]

\[
B_{SM}^{(0)}(B^{(s)}_s \to \mu^+ \mu^-) = (3.23 \pm 0.15 \pm 0.23 f_{B^0_s}) \times 10^{-9},
\]  

(2)

where \(f_{B^0_s}\) is the meson decay constant and its theoretical calculation represents the dominant source of systematic uncertainty in the SM prediction. This result needs to be corrected to allow for a comparison with the experimental detection and account for the sizeable decay width difference in the \(B^0_s\) system [12]. The corrected SM prediction becomes

\[
B_{SM}^{(0)}(B^{(s)}_s \to \mu^+ \mu^-) = (3.54 \pm 0.30) \times 10^{-9}.
\]  

(3)

This correction is negligible for the \(B^0 \to \mu^+ \mu^-\) decay, since the decay width difference is \(\sim 10^{-3}\) in the \(B^0\) system. The current best theoretical prediction for the SM \(B^0 \to \mu^+ \mu^-\)-branching fraction is the CP-averaged calculation performed in [11]:

\[
B_{SM}^{(0)}(B^0 \to \mu^+ \mu^-) = (0.107 \pm 0.010) \times 10^{-9}.
\]  

(4)

The \(B^{(s)}_0 \to \mu^+ \mu^-\) decay may receive large enhancements within specific New Physics models. For instance, the expected branching ratio in the framework of certain MSSM models can be very much enhanced due to a dependence with tan\(\beta\) [13]. Nevertheless, it is important to notice that the SUSY impact on the \(B^{(s)}_0 \to \mu^+ \mu^-\) decay can also be “hidden” and the MSSM contribution may well be suppressed by moving to regions of intermediate tan\(\beta\) values or large masses of the physical Higgs boson \(m_A\). The strong interest shown by the physics community towards \(B^{(s)}_0 \to \mu^+ \mu^-\) decays is motivated as this is a particularly fruitful probe of NP in the scalar and pseudoscalar sectors.

The search for a \(B^{(s)}_0 \to \mu^+ \mu^-\) decay at LHCb is the search for two very specific muonic tracks in the detector, but the very high background makes the analysis extremely challenging. After selection, the comparison between number of candidates, number of events expected for a given branching fraction hypothesis and number of expected background events is done in bins of two variables, independent from each other. They are the invariant mass (IM) and the output of the BDT. The signal IM is described using
a Crystal Ball function, whose parameters are mainly obtained from data. The yield of expected signal is then obtained using $B^+ \rightarrow J/\psi(\rightarrow \mu^+\mu^-)K^+$ and $B^0 \rightarrow K^+\pi^-$ decays as normalisation channels.

Concerning the number of expected background events, the combinatorial background is evaluated from the IM sidebands, defined as $[4900, 5224]$ MeV/$c^2$ for the left sideband and $[5432, 6000]$ MeV/$c^2$ for the right one. The BDT and IM PDF’s of the combinatorial background are then obtained from an extended likelihood fit to the sidebands, performed with a single exponential function, and the corresponding yields are extrapolated to the search mass windows. Among the peaking backgrounds, $B^0_{(s)} \rightarrow h^+h^-$ decays where both hadrons $h$ are misidentified as muons are particularly dangerous. In addition, $B^0_d \rightarrow \pi^-\mu^+\nu\mu$ and $B^{+ (0)} \rightarrow \pi^{+ (0)}\mu^+\mu^-$ decays are found to affect not negligibly the combinatorial background determination from the sidebands, at BDT > 0.8 (sensitive region). For this reason, the fit to the background accounts for the $B^0_d \rightarrow \pi^-\mu^+\nu\mu$ and $B^{+ (0)} \rightarrow \pi^{+ (0)}\mu^+\mu^-$ contributions as well, by adding their PDF’s to the exponential PDF describing the combinatorial component.

![Figure 1: $B^0_{(s)} \rightarrow \mu^+\mu^-$ analysis, unblinded 2012 data: 2D plot of the di-muon candidates invariant mass versus the event BDT. The orange-dashed (green-dashed) lines indicate the $\pm 60 MeV/c^2$ $B^0_s(B^0_d)$ search window.](image)

The distribution of events from the 2012 dataset in the (IM, BDT) plane, after the "unblinding" of the signal search regions, is shown in Figure 1. The orange-dashed and green-dashed lines indicate the $\pm 60 MeV/c^2$ signal search windows around $m(B^0_s)$ and $m(B^0_d)$, respectively.

The significance of the observation in each (IM, BDT) bin is assessed by applying the $CL_s$ technique both to the 2012 dataset alone, and to the 2012 and 2011 datasets combined (corresponding to 2.1 fb$^{-1}$): the latter result is here presented. The combination is performed by computing the $CL_s$ simultaneously for all the bins of the 2011 and 2012 data samples. The distributions of the expected value of $CL_s$
for the $B_0^s \rightarrow \mu^+\mu^-$ and $B^0 \rightarrow \mu^+\mu^-$ decays, versus the various branching fractions hypotheses, are shown in Fig. 2 (a) and (b) respectively, for the 2011 and 2012 datasets combined. The $p$-value measured for the $B^0 \rightarrow \mu^+\mu^-$ decay is 11%. For the

![Distribution of the expected value of $CL_s$ for the $B_0^s \rightarrow \mu^+\mu^-$ (on top) and $B^0 \rightarrow \mu^+\mu^-$ (below) versus the $B^0$ hypothesis, computed for the 2011 and 2012 datasets. The grey (purple) curve is the median of the expected $CL_s$ in the signal-plus-background (background-only) hypothesis, while the yellow (green) band covers 34% of the distribution on each side of the median. The red curve is the $CL_s$ measured on data.](image)

$B_0^s \rightarrow \mu^+\mu^-$ instead, the $p$-value is computed to be $6 \times 10^{-4}$. This corresponds to an excess of $B_0^s \rightarrow \mu^+\mu^-$ candidates with respect to background expectations with a significance of about 3.5$\sigma$, corresponding to the first signal evidence for this decay[14].

Given the high signal significance obtained, a double-sided limit on the branching fraction is also estimated by computing the lower and upper limits at $CL_{s+b} = 0.95(0.975)$ and $CL_{s+b} = 0.05(0.025)$ respectively, for a 90% (95%) C.L. The obtained result is $B(B_0^s \rightarrow \mu^+\mu^-) \in \left[1.3, 5.8\right] \times 10^{-9} \left[1.1, 6.4\right] \times 10^{-9}$, at 90% (95%) C.L.. Moreover, a branching ratio measurement can be obtained by means of a global fit to the di-muon invariant mass, with the input $p.d.f$’s already used for the fit to the invariant mass sidebands previously described. The value of the $B_0^s \rightarrow \mu^+\mu^-$ branching fraction determined from the fit is

$$B(B_0^s \rightarrow \mu^+\mu^-) = (3.2^{+1.4}_{-1.2} \text{(stat)}^{+0.5}_{-0.3} \text{(syst)}) \times 10^{-9}.$$  

(5)

The signal significance obtained from the difference in log-likelihood between signal and null hypothesis corresponds to a 3.5 standard deviations evidence of the $B_0^s \rightarrow \mu^+\mu^-$ decay. The result in Eq.(5) is in agreement with the SM expectations (Eq. 3), but leaves still place for New Physics scenarios.
6. Conclusions

FCNC processes are expected to be extremely rare in the SM but provide a fruitful probe for NP scenarios. In this context, the LHCb experiment has a rich program of rare leptonic decay searches and the 2012 analysis campaign has produced several important results. For instance, the search for $D^0 \rightarrow \mu^+\mu^-$ decay with $0.9 \text{ fb}^{-1}$ data leads to the new world’s best limit: $\mathcal{B}(D^0 \rightarrow \mu^+\mu^-) < 1.3(1.1) \times 10^{-8}$ at 95(90)\% C.L. [6]. Similarly, $K^0_S \rightarrow \mu^+\mu^-$ decays have been searched at LHCb with 1 fb$^{-1}$ data; the new limit is set at $\mathcal{B}(K^0_S \rightarrow \mu^+\mu^-) < 11(9) \times 10^{-9}$ at 95(90)\% [10] and improves the previous measurements by a factor of thirty. Finally, the search for $B^0(\kappa) \rightarrow \mu^+\mu^-$ decays with 2.1 fb$^{-1}$ data leads to the first 3.5$\sigma$ evidence for the $B^0(\kappa) \rightarrow \mu^+\mu^-$ decay [14]. The obtained branching fraction $\mathcal{B}(B^0(\kappa) \rightarrow \mu^+\mu^-) = (3.2^{+1.4}_{-1.2}(\text{stat})^{+0.5}_{-0.3}(\text{syst})) \times 10^{-9}$ is compatible with the SM expectations.

References

[1] Glashow S L, Iliopoulos J and Maiani L 1970 Phys. Rev. D 2(7) 1285–1292 URL http://link.aps.org/doi/10.1103/PhysRevD.2.1285
[2] The LHCb collaboration 2008 JINST 3 S08005
[3] The LHCb collaboration 2003 LHCb reoptimized detector design and performance: Technical Design Report Technical Design Report LHCb (Geneva: CERN)
[4] Read A L 2002 J. Phys. G 28 2693 URL http://stacks.iop.org/0954-3899/28/i=10/a=313
[5] Junk T 1999 Nucl. Instrum. Meth. A 434 435 – 443 URL http://www.sciencedirect.com/science/article/pii/S0168900299004982
[6] The LHCb Collaboration 2012 Subm. to 47th Rencontres de Moriond on Electroweak Interactions and Unified Theories LHCb–CONF–2012–005
[7] Petric M and others for the Belle Collaboration 2010 Phys. Rev. D 81 091102
[8] Ecker G and Pich A 1991 Nucl. Phys. B 366 189 – 205 URL http://www.sciencedirect.com/science/article/pii/0550321391900564
[9] Isidori G and Unterdorfer R 2004 J. Phys. 2004 009 URL http://stacks.iop.org/1126-6708/2004/i=01/a=009
[10] The LHCb Collaboration 2013 J. Phys. 2013(1) 1–18 URL http://dx.doi.org/10.1007/JHEP01(2013)29090
[11] Buras A J, Girrbach J, Guadagnoli D and Isidori G 2012 Eur. Phys. J. C72 2172 (Preprint 1208.0934)
[12] De Bruyn K, Fleischer R, Kneegjens R, Koppenburg P, Merk M and Tuning N 2012 Phys. Rev. D 86(1) 014027 URL http://link.aps.org/doi/10.1103/PhysRevD.86.014027
[13] Bobeth C, Buras A J, Krger F and Urban J 2002 Nucl. Phys. B 630 87 – 131 URL http://www.sciencedirect.com/science/article/pii/S0550321300001444
[14] The LHCb Collaboration 2013 Phys. Rev. Lett. 110(2) 021801 URL http://link.aps.org/doi/10.1103/PhysRevLett.110.021801