Searches for very rare decays to purely leptonic final states at LHCb

Mathieu Perrin-Terrin[*][†]
CPPM, Aix-Marseille Université, CNRS/IN2P3, Marseille, France
E-mail: perrin-terrin@cppm.in2p3.fr

We present a review of the searches for very rare decays to muonic final states performed at LHCb using 1.0 fb^{-1} of pp collisions at 7 TeV centre of mass energy. Flavour changing neutral current processes, such as $B^0_{(s)} \to \mu^+\mu^-$ and $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$ are highly suppressed in the Standard Model (SM). Such decays therefore allow contributions from new processes or new heavy particles to significantly modify the expected SM rates. Charged lepton flavour violating processes, such as the neutrino-less $\tau^- \to \mu^-\mu^-\mu^+$ decay, have vanishingly small decay rates in the SM, but can be significantly enhanced in extended models. We report the latest results on these channels from LHCb.

---

[*]Speaker.

[†]on behalf of the LHCb Collaboration.

1. Introduction

The searches for $B^0_{(s)} \to \mu^+\mu^-$, $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$ and $\tau^- \to \mu^-\mu^-\mu^+$ performed by analysing 1.0 fb^{-1} of pp collisions recorded in 2011 with the LHCb detector [1] are presented. The three decay modes being purely muonic, the analyses have similarities. They all take advantage of the efficient muonic trigger lines [2], that for instance allow to trigger events with a muon of low transverse momentum (down to 0.5 GeV for di-muons), and reach an efficiency of 90% for di-muon signal. In addition, high rejections of combinatorial backgrounds are achieved using the excellent tracking performances (e.g. the resolution on the momentum is $\frac{\delta p}{p} \in [0.4, 0.6]$ for $p \in [5; 100]$ GeV). Finally physical backgrounds, where hadrons are misidentified as muons, are reduced with the particles identification detectors that yield an efficiency for identifying muon of 97% for misidentification rate of pions (kaons) into muons below 3 (5)%. To avoid unconscious bias, the events for which the $B^0_{(s)}$ or τ mass candidates lie within three mass resolutions of the expected mean are kept hidden until all crucial choices for the analysis were made.

2. Searches for $B^0_{(s)} \to \mu^+\mu^-$

Within the Standard Model (SM), $B^0_{(s)} \to \mu^+\mu^-$ are rare processes as they occur only via loop diagrams and are helicity suppressed. The amplitudes contributing to the branching fraction ($\mathscr{B}$) can be expressed in terms of the scalar (c_S), pseudoscalar (c_P) and axial vector (c_A) Wilson coefficients in a general approach [3]. Within the SM, c_S and c_P contributions are negligible while c_A is calculated with a few percent accuracy [4] and leads [5] to a prediction of $\mathscr{B}(B^0_s \to \mu^+\mu^-)_{SM} = (3.2 \pm 0.2) \times 10^{-9}$ and $\mathscr{B}(B^0 \to \mu^+\mu^-)_{SM} = (1.0 \pm 0.1) \times 10^{-10}$. Models beyond the SM could contribute to these Wilson coefficients and change significantly the $\mathscr{B}$. For instance, within the Minimal Supersymmetric SM (MSSM) in the large $\tan\beta$ approximation [6], $c^{MSSM}_{S,P} \propto \tan^3\beta/M^2_A$, where M_A denotes the pseudoscalar Higgs mass and $\tan\beta$ the ratio of Higgs vacuum expectation values. The LHCb experiment sets the most restrictive upper limits [7], $\mathscr{B}(B^0_s \to \mu^+\mu^-) < 1.4 \times 10^{-8}$ and $\mathscr{B}(B^0 \to \mu^+\mu^-) < 3.2 \times 10^{-9}$ at 95% C.L. The present paper expounds the update of this analysis using the full 2011 data set.[1]

The strategy of the analysis is to derive the expected numbers of events of background and signal for a given $\mathscr{B}$ hypothesis and to compare these numbers to the observed ones with the CL$_S$ method [9]. The method provides CL$_{s+b}$(CL$_b$), a measure of the compatibility of the observed distribution with the signal plus background (background only) hypothesis, and CL$_S$ = CL$_{s+b}$/CL$_b$ which is used to set upper limits on the $\mathscr{B}$.

For a given $\mathscr{B}$, the number of events is obtained by scaling the yields observed in three controls channels, $B^0_s \to J/\psi\phi$, $B^0 \to K^+\pi^-$ and $B^+ \to J/\psi K^+$, by the ratio of the $\mathscr{B}$ hypothesis to the control channel $\mathscr{B}$. The scaling factor corrects also for the different efficiencies between signal and control channels and for different intial states using f_s/f_d measured at LHCb [10].

To improve the sensitivity of the analysis, a selection which reduces the backgrounds is applied. Backgrounds stem from combinatorics where the two muons come from two different b-mesons and from peaking background from $B^0_{(s)} \to h^+h'^-$ (h standing for K$^+$ or π^+) where hadrons

---

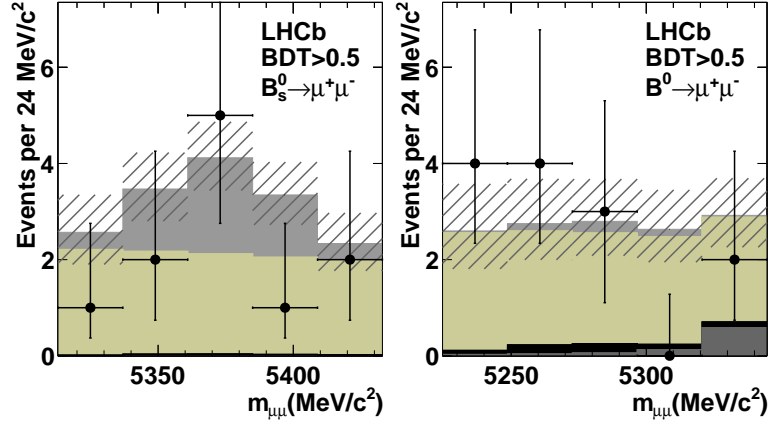[1]Since then, the analysis has been updated leading to the first evidence for the $B^0_s \to \mu^+\mu^-$ decay [8].

Figure 1: Distribution of selected candidates (black points) in the (left) $B_s^0 \to \mu^+\mu^-$ and (right) $B^0 \to \mu^+\mu^-$ mass window for BDT>0.5, and expectations for, from the top, $B_s^0 \to \mu^+\mu^-$ SM signal (gray), combinatorial background (light gray), $B_{(s)}^0 \to h^+ h'^-$ background (black), and cross-feed of the two modes (dark gray). The hatched area depicts the uncertainty on the sum of the expected contributions.

are identified as muons. The first type of background is reduced by cutting on topological and kinematical variables and on a combination of them obtained with a Boosted Decision Tree (BDT). This selection has an efficiency similar on the signal, normalisation and control channels. The peaking background is reduced using information from the particle identification detectors.

After this selection, events are classified in bins of the di-muon invariant mass and bins of a topological variable built with a second BDT. In each bin the expected numbers of signal and background events are derived. The signal mass shape is assumed to be a Crystal Ball shape, where the mean and the resolution are obtained on data while the transition point is derived from simulations. The BDT shape is obtained on data by extracting with a fit to the mass distribution, in each BDT bin, the yields of $B_{(s)}^0 \to h^+ h'^-$. For the combinatorial background, the mass and the BDT shape are obtained simultaneously by interpolating, in each BDT bin, the mass side-bands into the signal regions with an exponential function. Finally, the peaking background BDT shape is assumed to be the same as the signal one (correcting for trigger bias) and the mass shape is taken from simulations. The total number of peaking backgrounds is derived by a data driven method. The observed data and the expected distributions for the SM signal and background components are shown in Figure 1.

The comparison of the distributions of observed events and expected background events results in a p-value $(1 - CL_b)$ of 18 % (60 %) for the $B_s^0 \to \mu^+\mu^-$ ($B^0 \to \mu^+\mu^-$) decay, where the CL_b values are those corresponding to $CL_{s+b} = 0.5$. No excess being observed, uppers limits are set and reported in Table 1. The expected and observed CL_S values are shown in Fig. 2 for the $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ channels, each as a function of the assumed $\mathscr{B}$.

Assuming the same sensitivity for the analysis, prospect for the discovery of the $B_s^0 \to \mu^+\mu^-$ decay can be derived. Figure 3 shows the value of $\mathscr{B}(B_s^0 \to \mu^+\mu^-)$ for which a 3 σ evidence of the decay could be obtained as a function of the integrated luminosity. This extrapolation takes into account the increase of energy at the center of mass to 8 TeV and the already observed data pattern for 2011.
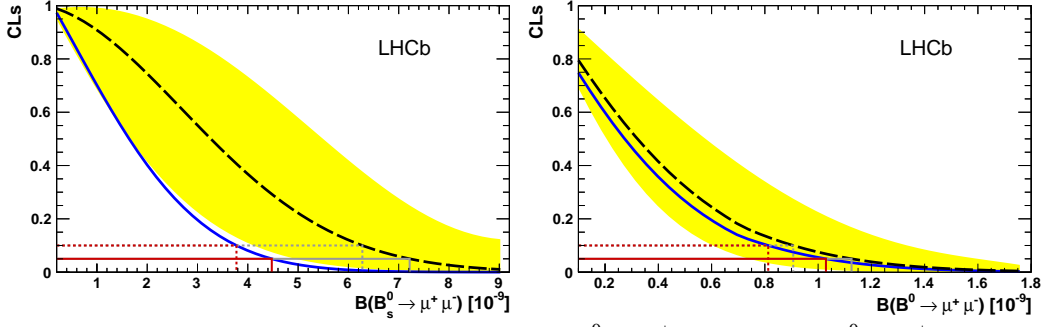
Figure 2: CL_S as a function of the assumed $\mathscr{B}$ for (left) $B_s^0 \to \mu^+\mu^-$ and (right) $B^0 \to \mu^+\mu^-$ decays. The long dashed black curves are the medians of the expected CL_S distributions for $B_s^0 \to \mu^+\mu^-$, if background and SM signal were observed, and for $B^0 \to \mu^+\mu^-$, if background only was observed. The yellow areas cover, for each $\mathscr{B}$, 34% of the expected CL_S distribution on each side of its median. The solid blue curves are the observed CL_S. The upper limits at 90 % (95 %) C.L. are indicated by the dotted (solid) horizontal lines in red (dark gray) for the observation and in gray for the expectation.

Mode	$B_s^0 \to \mu^+\mu^-$		$B^0 \to \mu^+\mu^-$	
Limit	at 90 % C.L.	at 95 % C.L.	at 90 % C.L.	at 95 % C.L.
Exp. bkg+SM	6.3×10^{-9}	7.2×10^{-9}		
Exp. bkg	2.8×10^{-9}	3.4×10^{-9}	0.91×10^{-9}	1.1×10^{-9}
Observed	3.8×10^{-9}	4.5×10^{-9}	0.81×10^{-9}	1.0×10^{-9}

Table 1: Expected and observed limits on $\mathscr{B}(B_{(s)}^0 \to \mu^+\mu^-)$.
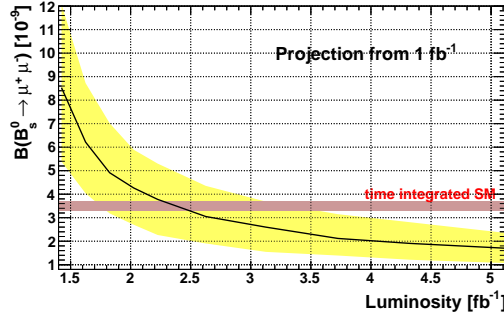

Figure 3: $\mathscr{B}$ at which the $B_s^0 \to \mu^+\mu^-$ decay would be observed at 3σ as a function of the integrated luminosity. The black plain line is the median of the $\mathscr{B}$ distribution giving a 3σ observation. The yellow areas cover, for each integrated luminosity, 34% of the $\mathscr{B}$ distribution on each side of its median. The red horizontal band represents the time integrated SM prediction for $\mathscr{B}(B_s^0 \to \mu^+\mu^-)$ [11].

3. Searches for $B_{(s)}^0 \to \mu^+\mu^-\mu^+\mu^-$

The decays of $B_{(s)}^0$ mesons to four muons are strongly suppressed in the SM and, for B_s^0, mainly occur via the resonant mode $B_s^0 \to J/\psi(\mu^+\mu^-)\Phi(\mu^+\mu^-)$ for which the $\mathscr{B}$ is estimated to be $(2.3 \pm 0.9) \times 10^{-9}$ [12]. In the SM, the $\mathscr{B}$ of the non-resonant modes are predicted not to exceed 10^{-10} [13], but could receive non negligible contributions from models beyond the SM. These decay modes have therefore potential for discovering physics beyond the SM and reassessing the hyper-CP anomaly [14].
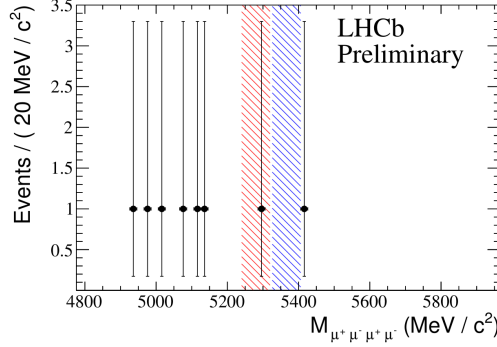
Figure 4: Observed invariant mass distribution of the $B_{(s)}^0 \to \mu^+\mu^-\mu^+\mu^-$ candidates. The blue (red) dashed area is the $B_s^0(B^0)$ mass window.

The searches are implemented with a cut and count analysis. Similarly to the $B_{(s)}^0 \to \mu^+\mu^-$ analysis, a control channel, $B^0 \to J/\psi(\mu\mu)\overline{K}^{*0}(K^+\pi^-)$, is used to derived the number of signal events corresponding to a $\mathscr{B}$ hypothesis. The selection is designed to be similar for the signal and the normalisation channel. The four muon candidates are required to be made of tracks with a good particle identification and to have an impact parameter significance (IPS) with respect to the primary vertex larger than 16. The candidate should also form a vertex with a χ^2 per degree of freedom smaller than 6 and have an IPS smaller than 9. Resonant candidates are vetoed when their di-muons invariant masses lie around the masses of the J/ψ and Φ.

The only relevant source of background is the combinatorial one. The expected number of these events is obtained by interpolating the side-bands and yields to $0.38^{+0.23}_{-0.17}$ ($0.30^{+0.22}_{-0.20}$) events in the $B_s^0(B^0)$ mass window. The observed pattern is shown in Figure 4 and is compatible with the background expectation. Using the CL_S method [9] the first limits on these two processes are set to $\mathscr{B}(B_s^0 \to \mu^+\mu^-\mu^+\mu^-) < 1.3 \times 10^{-8}$ and $\mathscr{B}(B^0 \to \mu^+\mu^-\mu^+\mu^-) < 5.4 \times 10^{-9}$ at 95% C.L. .

4. Search for $\tau^- \to \mu^-\mu^-\mu^+$

In the SM the lepton flavour violation (LFV) decay $\tau^- \to \mu^-\mu^-\mu^+$ is almost completely suppressed, while many theories beyond the SM [15] predict an enhanced LFV and in particular in the τ sector, raising in some cases $\mathscr{B}(\tau^- \to \mu^-\mu^-\mu^+)$ to an experimentally observable level.

The analysis strategy is similar to the one used for the searches of $B_{(s)}^0 \to \mu^+\mu^-$. The main background processes consist of long decay chains of heavy mesons with three real or misidentified muons in the final state. The dominant contribution in these processes is coming from $D_s^- \to \eta(\mu^+\mu^-\gamma)\mu^-\overline{\nu}_\mu$, and is treated separately.

After a loose selection based on the kinematics of the decay, the three muon candidates are classified in bins of invariant mass and of two multivariate operators combining information of the topology of the candidates and of the particle identification of the tracks used to form them. The three variables line shapes are determined for the signal with data driven methods and for the background by interpolating from the mass side-bands or with simulated data for $D_s^- \to \eta(\mu^+\mu^-\gamma)\mu^-\overline{\nu}_\mu$. The signal yield corresponding to a $\mathscr{B}$ hypothesis is obtained by normalising to $D_s^- \to \Phi(\mu^+\mu^-)\pi^-$.
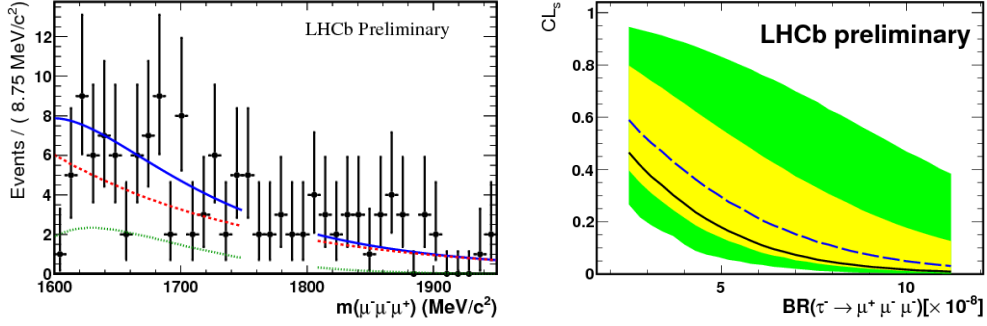
Figure 5: (left) Observed invariant mass distribution of the $\tau^- \to \mu^-\mu^-\mu^+$ candidates in the most signal like merged bins. The solid line is the fit to the data of the sum of the combinatorial exponential contribution (dashed red line) and of the $D_s^- \to \eta(\mu^+\mu^-\gamma)\mu^-\overline{\nu}_\mu$ one (dashed green line). (right) CL_S as a function of the assumed $\mathcal{B}$. The long dashed blue curves are the medians of the expected CL_S distributions for $\tau^- \to \mu^-\mu^-\mu^+$ if background only was observed. The yellow (green) areas cover, for each $\mathcal{B}$, 34% (48%) of the expected CL_S distribution on each side of its median. The solid black curves are the observed CL_S.

The observed data together with the background expectations are shown in Figure 5. The compatibility of the data to the background or background plus signal hypothesis are computed with the CL_S method [9]. The observed data are compatible with the background only hypothesis. Figure 5 shows the CL_S as a function of the assumed $\mathcal{B}$. An upper limit is set at $\mathcal{B}(\tau^- \to \mu^-\mu^-\mu^+) < 7.8(6.3) \times 10^{-8}$ at 95 (90)% C.L. which is the first limit of a LFV τ decay set at a hadron collider.

References

[1] LHCb Collaboration, JINST **3**, S08005 (2008).

[2] R. Aaij, *et al.*, arXiv:1211.3055

[3] C. Bobeth *et al.*, Phys. Rev. D **64** (2001) 074014.

[4] G. Buchalla and A.J. Buras, Nucl. Phys. B **548** (1999) 309.

[5] A.J. Buras, arXiv:1012.1447; E. Gamiz *et al.*, Phys. Rev. D **80** (2009) 014503; A.J. Buras, Phys. Lett. B **566** (2003) 115, A. J. Buras *et al.*, Eur. Phys. J. C **72**, 2172 (2012).

[6] L. J. Hall *et al.*, Phys. Rev. D **50** (1994) 7048; C. Hamzaoui *et al.*, Phys. Rev. D **59** (1999) 095005; C-S. Huang *et al.*, Phys. Rev. D **59** (1999) 11701;S. R. Choudhury and N. Gaur, Phys. Lett. B **451**, (1999) 86;K.S. Babu and C.F. Kolda, Phys. Rev. Lett. **84** (2000) 228.

[7] LHCb Collaboration,Phys. Rev. Lett.**108** 231801 (2012).

[8] LHCb Collaboration, arXiv:1211.2674

[9] A.L. Read, J. Phys. G **28** (2002) 2693; T. Junk, Nucl. Instrum. Meth. A **434** (1999) 435.

[10] LHCb Collaboration, Phys. Rev. D **85**, 032008 (2012)

[11] K. De Bruyn *et al.*, Phys. Rev. Lett. **109**, 041801 (2012)

[12] Particle Data Group Collaboration, J. Phys. G **37**, 075021 (2010).

[13] D. Melikhov and N. Nikitin, Phys. Rev. D **70**, 114028 (2004)

[14] HyperCP Collaboration, Phys. Rev. Lett. **94**, 021801 (2005)

[15] W. J. Marciano *et al.*, Ann. Rev. Nucl. Part. Sci. **58**, 315 (2008).