Search for the rare decays $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ with the LHCb experiment

G. Lanfranchi([1])
on behalf of the LHCb Collaboration
([1]) *Laboratori Nazionali di Frascati dell'INFN*

Summary. — A search for the decays $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ is performed with about 37 pb^{-1} of pp collisions at $\sqrt{s} = 7$ TeV collected by the LHCb experiment at the Large Hadron Collider at CERN. The observed numbers of events are consistent with the background expectations. The resulting upper limits on the branching ratios are BR$(B_s^0 \to \mu^+\mu^-) < 5.6 \times 10^{-8}$ and BR$(B^0 \to \mu^+\mu^-) < 1.5 \times 10^{-8}$ at 95% confidence level.

1. – Introduction

Measurements at low energies may provide interesting indirect constraints on the masses of particles that are too heavy to be produced directly. This is particularly true for Flavour Changing Neutral Currents (FCNC) processes which are highly suppressed in the Standard Model (SM) and can only occur through higher order diagrams.

The SM prediction for the Branching Ratios (BR) of the FCNC decays $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ have been computed [1] to be $BR(B_s^0 \to \mu^+\mu^-) = (3.2 \pm 0.2) \times 10^{-9}$ and $BR(B^0 \to \mu^+\mu^-) = (0.10 \pm 0.01) \times 10^{-9}$.

However New Physics (NP) contributions can significantly enhance this value. For example, within Minimal Supersymmetric extensions of the SM (MSSM), in the large $\tan\beta$ approximation [2], the $BR(B_s \to \mu^+\mu^-)$ is found to be proportional to $\sim \tan^6\beta$, where $\tan\beta$ is the ratio of vacuum expectation values of the two neutral CP-even Higgs fields. Therefore it could be strongly enhanced for large values of $\tan\beta$.

The most restrictive limits on the search for $B_{(s)}^0 \to \mu^+\mu^-$ have so far been achieved at the Tevatron, due to the large $b\bar{b}$ cross-section at hadron colliders. The best limits at 95% CL published so far are obtained using 6.1 fb^{-1} by the D0 collaboration [3], BR$(B_s^0 \to \mu^+\mu^-) < 5.1 \times 10^{-8}$ and using 2 fb^{-1} by the CDF collaboration [4], BR$(B_s^0 \to \mu^+\mu^-) < 5.8 \times 10^{-8}$ and BR$(B^0 \to \mu^+\mu^-) < 1.8 \times 10^{-8}$. The CDF collaboration has also presented preliminary results [5] with 3.7 fb^{-1} that lower the limits to BR$(B_s^0 \to \mu^+\mu^-) < 4.3 \times 10^{-8}$ and BR$(B^0 \to \mu^+\mu^-) < 0.76 \times 10^{-8}$.

The LHCb experiment is well suited for such searches due to its good invariant mass resolution, vertex resolution, muon identification and trigger acceptance.

In addition, LHCb has a hadronic trigger capability which provides large samples of $B^0_{s,d} \to h^+h'^-$ decays, where h and h' stand for a hadron (kaon or pion). These are used as control samples in order to reduce the dependence of the results on the simulation.

The measurements presented in this document use about 37 pb^{-1} of integrated luminosity collected by LHCb between July and October 2010 at $\sqrt{s} = 7$ TeV. Assuming the SM branching ratio, about 0.7 (0.08) $B^0_{(s)} \to \mu^+\mu^-$ ($B^0 \to \mu^+\mu^-$) are expected to be reconstructed using the $b\bar{b}$ cross-section, measured within the LHCb acceptance, of $75 \pm 14\,\mu b$ [6].

2. – The LHCb detector

The LHCb detector [7] is a single-arm forward spectrometer with an angular coverage from approximately 10 mrad to 300 (250) mrad in the bending (non-bending) plane.

The detector consists of a vertex locator (VELO), a warm dipole magnet with a bending power of $\int Bdl = 4$ T m, a tracking system, two ring-imaging Cherenkov detectors (RICH), a calorimeter system and a muon system.

Track momenta are measured with a precision between $\delta p/p = 0.35\%$ at 5 GeV/c and $\delta p/p = 0.5\%$ at 100 GeV/c. The RICH system provides charged hadron identification in a momentum range 2–100 GeV/c. Typically kaon identification efficiencies of over 90% can be attained for a $\pi \to K$ fake rate below 10%. The calorimeter system consists of a preshower, a scintillating pad detector, an electromagnetic calorimeter and a hadronic calorimeter. It identifies high transverse energy ($E_{\rm T}$) hadron, electron and photon candidates and provides information for the trigger. Five muon stations provide fast information for the trigger and muon identification capability: a muon identification efficiency of $\sim 95\%$ is obtained for a misidentification rate of about 1–2 % for momenta above 10 GeV/c.

LHCb has a two-level flexible and efficient trigger system both for leptonic and purely hadronic B decays. It exploits the finite lifetime and relatively large mass of charm and beauty hadrons to distinguish heavy flavour decays from the dominant light quark processes. The first trigger level (L0) is implemented in hardware and reduces the rate to a maximum of 1 MHz, the read-out rate of the whole detector. The second trigger level (High Level Trigger, HLT) is implemented in software running on an event filter CPU farm. The forward geometry allows the LHCb first level trigger to collect events with one or two muons with p_T values as low as 1.4 GeV/c for single muon and $p_{\rm T}(\mu_1) > 0.48$ GeV/c and $p_{\rm T}(\mu_2) > 0.56$ GeV/c for dimuon triggers. The $E_{\rm T}$ threshold for the hadron trigger varied in the range 2.6 to 3.6 GeV.

The dimuon trigger line requires muon pairs of opposite charge forming a common vertex and an invariant mass $M_{\mu\mu} > 4.7$ GeV/c^2. A second trigger line, primarily to select $J/\psi \to \mu\mu$ events, requires $2.97 < M_{\mu\mu} < 3.21$ GeV/c^2. The remaining region of the dimuon invariant mass is also covered by trigger lines that in addition require the dimuon secondary vertex to be well separated from the primary vertex. Other HLT trigger lines select generic displaced vertices, providing a high efficiency for purely hadronic decays (for instance $B^0_{s,d} \to h^+h'^-$).

3. – Analysis Strategy

The analysis for the $B^0_{(s)} \to \mu^+\mu^-$ search at LHCb is described in detail in [8]. It is done in two steps: first a very efficient selection removes the biggest amount of the background while keeping most of the signal within the LHCb acceptance. Then each event is given a probability to be signal or background in a two-dimensional probability space defined by the dimuon invariant mass and a multivariate analysis discriminant likelihood, the *Geometrical Likelihood* (GL). The compatibility of the observed distribution of events in the GL *vs* invariant mass plane with a given branching ratio hypothesis is evaluated using the $\mathrm{CL_s}$ method [9].

The number of expected signal events is evaluated by normalizing with channels of known branching ratios: $B^+ \to J/\psi K^+$, $B^0_s \to J/\psi\phi$ and $B^0 \to K^+\pi^-$, where $J/\Psi \to \mu^+\mu^-$ and $\phi \to K^+K^-$([1]). This normalization ensures that knowledge of the absolute luminosity and $b\bar{b}$ production cross-section are not needed, and that many systematic uncertainties cancel in the ratio of the efficiencies.

An important feature of this analysis is to rely as much as possible on data and to restrict to a minimum the use of simulation.

3˙1. *Event selection.* – The selection requires two muon candidates of opposite charge, forming a vertex with a $\chi^2/\mathrm{ndf} < 14$. Tracks are first required to be of good quality and to be displaced with respect to the closest primary vertex. To reject bad combinations before performing the vertex fit, the two tracks are required to have a distance of closest approach of less than 0.3 mm. The secondary vertex is required to be well fitted and must be clearly separated from the primary in the forward direction. When more than one primary vertex is reconstructed, the one that gives the minimum impact parameter significance for the candidate is chosen. The reconstructed candidate has to point to the the primary vertex.

Events passing the selection are considered $B^0_{(s)} \to \mu^+\mu^-$ candidates if their invariant mass lies within $60\,\mathrm{MeV/c^2}$ of the nominal $B^0_{(s)}$ mass. Assuming the SM branching ratio, 0.3 $B^0_s \to \mu^+\mu^-$ and 0.04 $B^0 \to \mu^+\mu^-$ events are expected after all selection requirements. There are 343 (342) $B^0_{(s)} \to \mu^+\mu^-$ candidates selected from data in the $B^0_s(B^0)$ mass window.

The dominant background after the $B^0_{(s)} \to \mu^+\mu^-$ selection is expected to be $b\bar{b} \to \mu\mu X$ [10]. This is confirmed by comparing the kinematical distributions of the sideband data with a $b\bar{b} \to \mu\mu X$ MC sample.

The muon misidentification probability as a function of momentum obtained from data using $K^0_S \to \pi^+\pi^-$, $\Lambda \to p\pi^-$ and $\phi \to K^+K^-$ decays is in good agreement with MC expectations. An estimate of the background coming from misidentified hadrons is obtained by reweighting the hadron misidentification probability using the momentum spectrum of the background in the invariant mass sidebands. The single hadron average misidentification probability is measured to be $(7.1 \pm 0.5) \times 10^{-3}$ and the double hadron misidentification probability is $(3.5 \pm 0.9) \times 10^{-5}$, where the correlation between the momenta of the two hadrons is taken into account.

The same selection without the muon identification requirement is applied to the control channel $B^0_{s,d} \to h^+h^{'-}$ and, with a softer pointing requirement, to the $J/\psi \to \mu\mu$

([1]) When the $B^0_s \to J/\psi\phi$ and $B^+ \to J/\psi K^+$ we always assume the decays $J/\psi \to \mu^+\mu^-$ and $\phi \to K^+K^-$.

decay of the control and normalization channels containing J/ψ in the final state.

3˙2. *Signal and background likelihoods.* – After the selection the signal purity is still about 10^{-3} for $B_s^0 \to \mu^+\mu^-$ and 10^{-4} for $B^0 \to \mu^+\mu^-$ assuming the SM branching ratios. Further discrimination is achieved through the combination of two independent variables: the multivariate analysis discriminant likelihood, GL, and the invariant mass. The GL combines information related with the topology and kinematics of the event as the $B^0_{(s)}$ lifetime, the minimum impact parameter of the two muons, the distance of closest approach of the two tracks, the $B^0_{(s)}$ impact parameter and p_T and the isolation of the muons with respect to the other tracks of the event. These variables are combined in an optimal way by taking their correlations properly into account [11, **?**].

The analysis is performed in two-dimensional bins of invariant mass and GL. The invariant mass in the signal regions (± 60 MeV/c^2 around the B_s^0 and the B^0 masses) is divided into six bins of equal width, and the GL into four bins of equal width distributed between zero and one. A probability to be signal or background is assigned to events falling in each bin.

The GL variable is defined using MC events but calibrated with data using $B^0_{s,d} \to h^+h^{'-}$ selected as the signal events and triggered independently on the signal (TIS events) in order to minimize the bias introduced by the hadronic trigger lines [8].

The number of $B^0_{s,d} \to h^+h^{'-}$ events in each GL bin is obtained from a fit to the inclusive mass distribution [13] assigning the muon mass to the two particles. The measured fractions in each GL bin can be seen in Fig. 1 and are quoted in Table I. The systematic uncertainties are included, estimated by comparing the results from the inclusive $B^0_{s,d} \to h^+h^{'-}$ fit model with those obtained using a double Crystal Ball function [14] and a simple background subtraction.

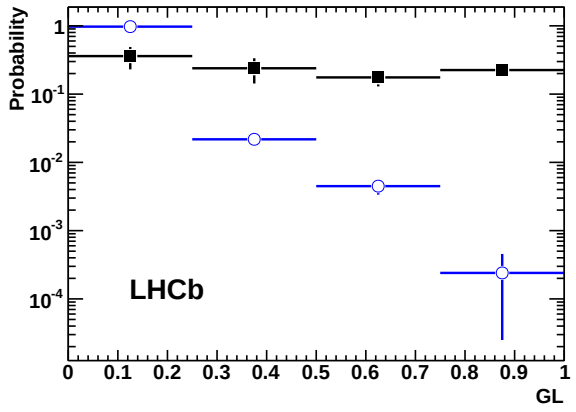


Fig. 1. – Probability of signal events in bins of GL obtained from the inclusive sample of $B^0_{s,d} \to h^+h^{'-}$ events (solid squares). The background probability (open circles) is obtained from the events in the sidebands of the $\mu\mu$ invariant mass distribution in the $B^0_{(s)}$ mass window.

Two methods have been used to estimate the $B^0_{(s)} \to \mu^+\mu^-$ mass resolution from data. The first method uses an interpolation between the measured resolutions for $c\bar{c}$

TABLE I. – *Probability of signal events in bins of GL obtained from the inclusive sample of $B_q^0 \to h^+h^{'-}$ events. The background probability in the B_s^0 mass window is obtained from the events in the sidebands of the dimuon invariant mass distribution.*

GL bin	Signal prob.	Background prob.
$0.0 - 0.25$	0.360 ± 0.130	$0.9735^{+0.0030}_{-0.0032}$
$0.25 - 0.5$	0.239 ± 0.096	$0.0218^{+0.0030}_{-0.0028}$
$0.5 - 0.75$	0.176 ± 0.046	$0.0045^{+0.0012}_{-0.0010}$
$0.75 - 1.0$	0.225 ± 0.036	$0.00024^{+0.00031}_{-0.00015}$

resonances (J/ψ, ψ(2S)) and $b\bar{b}$ resonances ($\Upsilon(1S)$, $\Upsilon(2S)$, $\Upsilon(3S)$) decaying into two muons. Interpolating linearly between the five fitted resolutions to $M_{B_s^0}$ an invariant mass resolution of $\sigma = 26.83 \pm 0.14$ MeV/c^2 was estimated.

The systematic uncertainty is estimated to be 1 MeV/c^2 mainly due to the reweighting of the momentum spectrum of the dimuon resonances and the variation of the resolution over the width of the $B^0_{(s)} \to \mu^+\mu^-$ signal region. The second method that was used to estimate the invariant mass resolution from data is to use the inclusive $B^0_{s,d} \to h^+h^{'-}$ sample. The particle identification requirement would modify the momentum and transverse momentum spectrum of pions and kaons, and thus the mass resolution. Therefore, the fit is performed to the inclusive $B^0_{s,d} \to h^+h^{'-}$ sample without requiring particle identification and assigning the muon mass to the decay products.

The fit has been performed in the GL range [0.25,1.0] and fitted parameters are the mass resolution, the B^0 and B_s^0 masses, the signal yield, the combinatorial background yields, as well as the fraction of radiative tail and the parameters that describe the combinatorial background. The relative contributions of B^0 and B_s^0 decays are fixed to their known values. The result of the fit for the mass resolution, $\sigma = 25.8 \pm 1.0 \pm 2.3$ MeV/c^2, is consistent with the value obtained from the interpolation method.

The weighted average of the two methods, $\sigma = 26.7 \pm 0.9$ MeV/c^2, is taken as the invariant mass resolution and considered to be the same for B^0 and $B^0_{(s)}$ decays.

The prediction of the number of background events in the signal regions is obtained by fitting with an exponential function the $\mu\mu$ mass sidebands independently in each GL bin in order to account for potentially different background compositions. The mass sidebands are defined in the range between $M_{B^0_{(s)}} \pm 600$ (1200) MeV/c^2 for the lower (upper) two GL bins, excluding the two search windows ($M_{B^0_{(s)}} \pm 60$ MeV/c^2).

The distribution of the invariant mass for each GL bin is shown in Fig. 2.

4. – Normalization factors

As already mentioned, the number of expected signal events is evaluated by normalizing with channels of known branching ratios, $B^+ \to J/\psi K^+$, $B_s^0 \to J/\psi\phi$ and $B^0 \to K^+\pi^-$, as shown in Table II, first column.

The first two decays have similar trigger and muon identification efficiency to the signal but a different number of particles in the final state, while the third channel has the same two-body topology but is selected with the hadronic trigger. The branching ratio of the $B_s^0 \to J/\psi\phi$ decay is not known precisely ($\sim 25\%$) but has the advantage

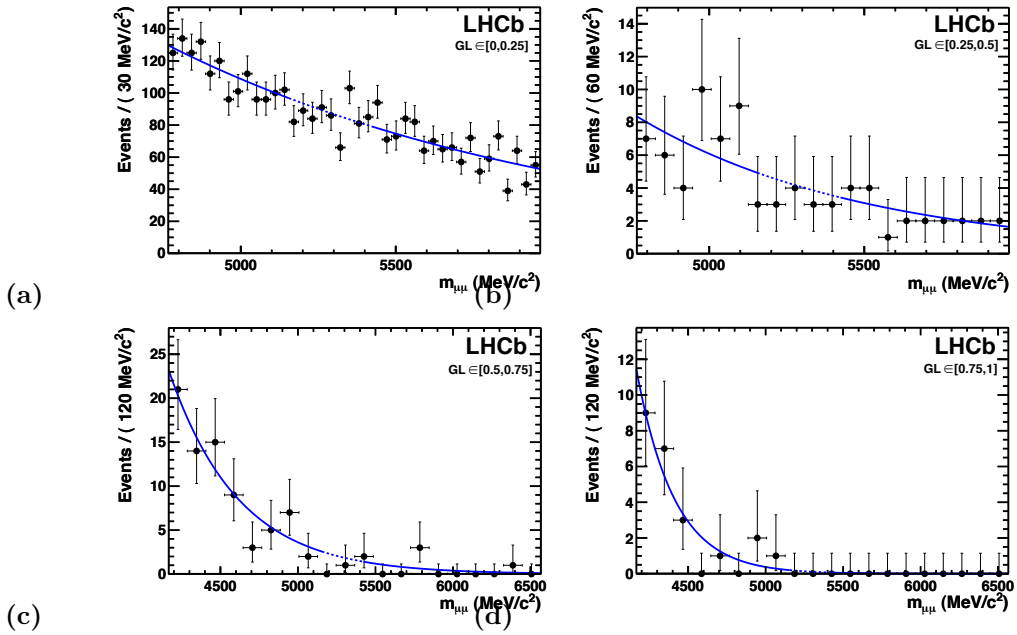


Fig. 2. – Distribution of the $\mu\mu$ invariant mass for different GL bins: (a) $[0, 0.25]$, (b) $[0.25, 0.5]$, (c) $[0.5, 0.75]$, (d) $[0.75, 1.0]$. The blue solid lines show the interpolation model used and the dashed line shows the result of the interpolation in the search windows.

that the normalization of $B^0_{(s)} \to \mu\mu$ with a B^0_s decay does not require the knowledge of the ratio of fragmentation fractions, which has an uncertainty of $\sim 13\%$ [16].

Using each of these normalization channels, $\mathrm{BR}(B^0_{(s)} \to \mu\mu)$ can be calculated as:

$$\begin{aligned} BR(B^0_{(s)} \to \mu\mu) &= BR_{\mathrm{norm}} \times \frac{\epsilon^{\mathrm{REC}}_{\mathrm{norm}} \epsilon^{\mathrm{SEL|REC}}_{\mathrm{norm}} \epsilon^{\mathrm{TRIG|SEL}}_{\mathrm{norm}}}{\epsilon^{\mathrm{REC}}_{\mathrm{sig}} \epsilon^{\mathrm{SEL|REC}}_{\mathrm{sig}} \epsilon^{\mathrm{TRIG|SEL}}_{\mathrm{sig}}} \times \frac{f_{\mathrm{norm}}}{f_{B^0_{(s)}}} \times \frac{N_{B^0_{(s)} \to \mu\mu}}{N_{\mathrm{norm}}} \\ &= \alpha_{B^0_{(s)} \to \mu\mu} \times N_{B^0_{(s)} \to \mu\mu} \,, \end{aligned} \tag{1}$$

where $\alpha_{B^0_{(s)} \to \mu\mu}$ denotes the normalization factor, $f_{B^0_{(s)}}$ denotes the probability that a b-quark fragments into a $B^0_{(s)}$ and f_{norm} denotes the probability that a b-quark fragments into the b-hadron relevant for the chosen normalization channel with branching fraction $\mathrm{BR}_{\mathrm{norm}}$. The reconstruction efficiency (ϵ^{REC}) includes the acceptance and particle identification, while $\epsilon^{\mathrm{SEL|REC}}$ denotes the selection efficiency on reconstructed events. The trigger efficiency on selected events is denoted by $\epsilon^{\mathrm{TRIG|SEL}}$.

The ratios of reconstruction and selection efficiencies are estimated from the simulation and checked on data, while the ratios of trigger efficiencies on selected events are determined from data [8].

The yields needed to evaluate the normalization factor are shown in Table II, where the uncertainty is dominated by the differences observed using different models in fitting the invariant mass lineshape.

As can be seen in Table II, the normalization factors calculated using the three com-

TABLE II. – *Summary of the factors and their uncertainties needed to calculate the normalization factors ($\alpha_{B^0_{(s)}\to\mu^+\mu^-}$) for the three normalization channels considered. The branching ratios are taken from Refs. [15, 17] and includes also the BR($J/\psi \to \mu^+\mu^-$) and BR($\phi \to K^+K^-$). The trigger efficiency and number of $B^0 \to K^+\pi^-$ candidates correspond to only TIS events, as described in the text.*

	BR ($\times 10^{-5}$)	$\frac{\epsilon^{\rm REC}_{\rm norm}\epsilon^{\rm SEL|REC}_{\rm norm}}{\epsilon^{\rm REC}_{\rm sig}\epsilon^{\rm SEL|REC}_{\rm sig}}$	$\frac{\epsilon^{\rm TRIG|SEL}_{\rm norm}}{\epsilon^{\rm TRIG|SEL}_{\rm sig}}$	$N_{\rm norm}$	$\alpha_{B^0_{(s)}\to\mu^+\mu^-}$ ($\times 10^{-9}$)	$\alpha_{B^0\to\mu^+\mu^-}$ ($\times 10^{-9}$)
$B^+ \to J/\psi K^+$	5.98 ± 0.22	0.49 ± 0.02	0.96 ± 0.05	$12,366 \pm 403$	8.4 ± 1.3	2.27 ± 0.18
$B_s^0 \to J/\psi\phi$	3.4 ± 0.9	0.25 ± 0.02	0.96 ± 0.05	760 ± 71	10.5 ± 2.9	2.83 ± 0.86
$B^0 \to K^+\pi^-$	1.94 ± 0.06	0.82 ± 0.06	0.072 ± 0.010	578 ± 74	7.3 ± 1.8	1.99 ± 0.40

plementary channels give compatible results. The final normalization factor is a weighted average which takes into account all the sources of correlations, in particular the dominant one coming from the uncertainty on $f_d/f_s = 3.71 \pm 0.47$ [16], with the result:

$$\alpha_{B^0_{(s)}\to\mu\mu} = (8.6 \pm 1.1) \times 10^{-9}\,,$$
$$\alpha_{B^0\to\mu\mu} = (2.24 \pm 0.16) \times 10^{-9}\,.$$

5. – Results

For each of the 24 bins (4 bins in GL and 6 bins in mass) the expected number of background events is computed from the fits to the invariant mass sidebands described in Sect. **3**˙2. The expected numbers of signal events are computed using the normalization factors from Sect. **4**, and the signal likelihoods computed in Section **3**˙2. The distribution of observed events in the GL *vs* invariant mass plane can be seen in Fig. 3.

The compatibility of the observed distribution of events in the GL *vs* invariant mass plane with a given branching ratio hypothesis is evaluated using the CL_s method [9]. The observed distribution of CL_s as a function of the assumed branching ratio can be seen in Fig. 4.

The expected distributions of possible values of CL_s assuming the background-only hypothesis are also shown in the same figure as a green shaded area that covers the region of $\pm 1\sigma$ of background compatible observations. The uncertainties in the signal and background likelihoods and normalization factors are used to compute the uncertainties in the background and signal predictions. These uncertainties are the only source of systematic uncertainty and they are included in the CL_s using the techniques described in Ref. [9]. Given the specific pattern of the observed events, the systematic uncertainty on the background prediction has a negligible effect on the quoted limit. The effect of the uncertainty on the signal prediction increases the quoted limits by less than 3%.

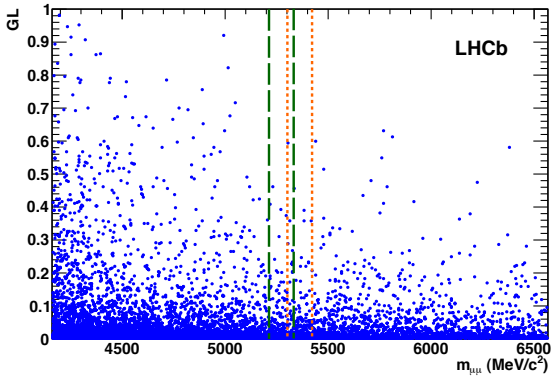
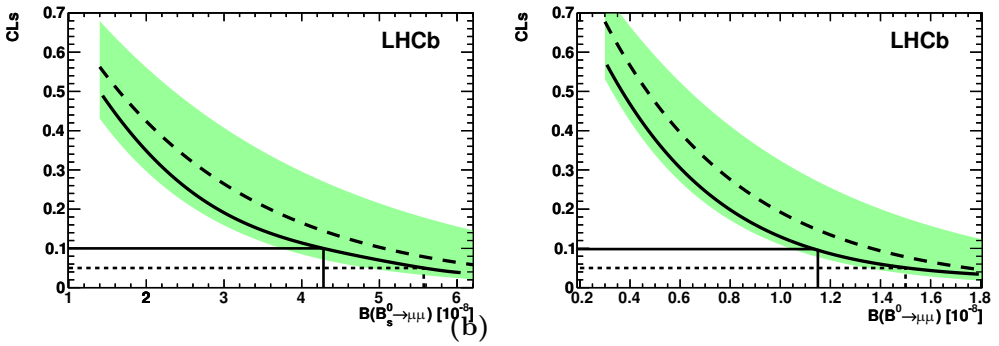


Fig. 3. – Observed distribution of selected dimuon events in the GL *vs* invariant mass plane. The orange short-dashed (green long-dashed) lines indicate the ± 60 MeV/c^2 search window around the $B_s^0(B^0)$.

The upper limits are computed using the CL_s distributions in Fig. 4 with the results:

$$\mathrm{BR}B^0_{(s)} \to \mu^+\mu^- < 4.3\ (5.6) \times 10^{-8} \text{ at } 90\,\%\ (95\,\%) \text{ C.L.},$$
$$\mathrm{BR}B^0 \to \mu^+\mu^- < 1.2\ (1.5) \times 10^{-8} \text{ at } 90\,\%\ (95\,\%) \text{ C.L.},$$

while the expected values of the limits are $\mathrm{BR}(B_s^0 \to \mu^+\mu^-) < 5.1\ (6.5) \times 10^{-8}$ and $\mathrm{BR}(B^0 \to \mu^+\mu^-) < 1.4\ (1.8) \times 10^{-8}$ at 90 % (95 %) CL The limits observed are similar to the best published limits [3] for the decay $B_s^0 \to \mu^+\mu^-$ and more restrictive for $B^0 \to \mu^+\mu^-$ the decay [4].



(a) (b)

Fig. 4. – (a) Observed (solid curve) and expected (dashed curve) CL_s values as a function of $\mathrm{BR}(B_s^0 \to \mu^+\mu^-)$. The green shaded area contains the $\pm 1\sigma$ interval of possible results compatible with the expected value when only background is observed. The 90 % (95 %) CL observed value is identified by the solid (dashed) line. (b) the same for $\mathrm{BR}(B^0 \to \mu^+\mu^-)$.

REFERENCES

[1] A.J. Buras, "Minimal flavour violation and beyond: Towards a flavour code for short distance dynamics", arXiv:1012.1447;
E. Gamiz *et al*, "Neutral B meson mixing in unquenched lattice QCD", Phys. Rev. D **80** (2009) 014503;
A.J. Buras, "Relations between $\Delta M_{s,d}$ and $B_{s,d} \to \mu^+\mu^-$ in models with Minimal Flavour Violation", Phys. Lett. B **566** (2003) 115.

[2] L. J. Hall, R. Rattazzi and U. Sarid, "The top quark mass in supersymmetric SO(10) unification" Phys. Rev. D **50** (1994) 7048;
C. Hamzaoui, M. Pospelov and M. Toharia, "Higgs-mediated FCNC in supersymmetric models with large $\tan\beta$", Phys. Rev. D **59** (1999) 095005;
K.S. Babu and C.F. Kolda, "Higgs-mediated $B_{s,d} \to \mu^+\mu^-$ in minimal supersymmetry", Phys. Rev. Lett. **84** (2000) 228.

[3] V. Abazov *et al.* [D0 Collaboration], "Search for the rare decay $B_s^0 \to \mu^+\mu^-$", Phys. Lett. B **693** (2010) 539.

[4] T. Aaltonen *et al.* [CDF Collaboration], "Search for $B_s^0 \to \mu^+\mu^-$ and $B_d^0 \to \mu^+\mu^-$ decays with 2 fb^{-1} of $p\bar{p}$ Collisions", Phys. Rev. Lett. **100** (2008) 101802.

[5] T. Aaltonen *et al.* [CDF Collaboration], "Search for $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ decays in 3.7 fb^{-1} of $p\bar{p}$ collisions with CDF II", CDF Public Note 9892.

[6] R. Aaij *et al.* [LHCb Collaboration], "Measurement of $\sigma(pp \to b\bar{b}X)$ at $\sqrt{s} = 7$ TeV in the forward region", Phys. Lett. B **694** (2010) 209.

[7] A.A. Alves *et al.* [LHCb Collaboration], "The LHCb detector at the LHC", JINST **3** (2008) S08005, and references therein.

[8] R. Aaji *et al.* [LHCb Collaboration], "Search for the rare decays $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$", Phys. Lett. B 699 (2011), 330-340.

[9] A.L. Read, "Presentation of search results: the CL_s technique", J. Phys. G **28** (2002) 2693;
T. Junk, "Confidence level computation for combining searches with small statistics", Nucl. Instrum. Meth. A **434** (1999) 435.

[10] B. Adeva *et al.* [LHCb Collaboration], "Roadmap for selected key measurements of LHCb", Chapter 5, v:0912.4179v2.

[11] D. Karlen, "Using projections and correlations to approximate probability distributions", Comp. Phys. **12** (1998) 380.

[12] D. Martinez Santos, "Study of the very rare decay $B_s^0 \to \mu^+\mu^-$ in LHCb", CERN-THESIS-2010-068.

[13] B. Adeva *et al.* [LHCb Collaboration], "Roadmap for selected key measurements of LHCb", Chapter 3, v:0912.4179v2.

[14] J.E. Gaiser, "Charmonium spectroscopy from radiative decays of the J/ψ and Ψ' ", PhD thesis, SLAC-R-255 (1982), Appendix F;
T. Skwarnicki, "A study of the radiative cascade transitions between the Υ and Υ' resonances", PhD thesis, DESY F31-86-02 (1986), Appendix E.

[15] K. Nakamura *et al.* [Particle Data Group], "Review of particle physics", J. Phys. G **37** (2010) 075021.

[16] D. Asner *et al.* [Heavy Flavour Averaging Group], "Averages of b-hadron, c-hadron, and τ-lepton properties", arXiv:1010.1589. Updated values for f_d/f_s available at http://www.slac.stanford.edu/xorg/hfag/osc/end_2009/ have been used for this document, as they include also pre-summer 2010 results which are not contained in [15].

[17] R. Louvot, "Υ(5S) Results at BELLE", arXiv:0905.4345v2.