PROCEEDINGS OF SCIENCE

Measurement of rare $B \to \mu\mu$ decays and search for $\tau \to 3\mu$ decays at CMS

Kai-Feng Chen*
National Taiwan University
E-mail: kfjack@phys.ntu.edu.tw

The CMS experiment is a privileged test bench for rare decays involving heavy flavors, based on the highly efficient dimuon triggers and the large integrated luminosity provided by the LHC collider. The most recent studies from CMS collaboration based on the data collected during LHC Run-1 and 2016 are presented, including the measurement of B_s^0 and B^0 mesons decaying into dimuons and the search for τ lepton decaying into three muons.

18th International Conference on B-Physics at Frontier Machines - Beauty2019 -
29 September / 4 October, 2019
Ljubljana, Slovenia

*Speaker.

PoS(Beauty2019)018

1. Measurement of rare $B \to \mu\mu$ decays

The $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ decays can only proceed through the higher-order flavor-changing-neutral-current processes, and are highly suppressed in the standard model (SM). However, since these processes are theoretically clean, they are excellent places to look for new physics. Key observables include the decay branching fractions for B_s^0 and $B^0 \to \mu^+\mu^-$, as well as the effective lifetime measured from $B_s^0 \to \mu^+\mu^-$ candidates. The decay of $B_s^0 \to \mu^+\mu^-$ was first observed based on the combined analysis of CMS and LHCb data [1], and its study now starts to enter the precision regime; furthermore, there are hopes that the first evidence of $B^0 \to \mu^+\mu^-$ decay might soon emerge. The effective lifetime is defined by

$$\tau_{\mu\mu} = \frac{\int t\Gamma(B_s(t) \to \mu\mu)dt}{\int \Gamma(B_s(t) \to \mu\mu)dt} = \frac{\tau_{B_s^0}}{1-y_s^2}\left(\frac{1+2\mathscr{A}_{\Delta\Gamma}^{\mu\mu} y_s + y_s^2}{1+\mathscr{A}_{\Delta\Gamma}^{\mu\mu} y_s}\right), \tag{1.1}$$

where $\mathscr{A}_{\Delta\Gamma}^{\mu\mu}$ and y_s are defined by $-\Re(\lambda)/(1+|\lambda|^2)$ and $\tau_{B_s^0}\Delta\Gamma_s/2$, respectively. In the SM, only the heavy B_s state can decay into dimuons, but different composition of CP states may be allowed by extended models. The SM predicts [2, 3] $\mathscr{B}(B_s^0 \to \mu^+\mu^-) = 3.57 \pm 0.17 \times 10^{-9}$, $\mathscr{B}(B^0 \to \mu^+\mu^-) = 1.06 \pm 0.09 \times 10^{-10}$, and $\tau_{\mu\mu} = 1.615$ ps. Details of the analysis is documented in Ref. [4].

1.1 Analysis Aspects

The analysis of $B \to \mu\mu$ relies on the distinct signal signatures of two muons, an associated single displaced vertex, which is isolated from other activities. The momentum of B candidate should align with its flight direction, and the invariant mass should peak at the expected values of $M(B_s)$ and $M(B_d)$. The major background sources include combinatorial events and other rare B decay processes. The combinatorial background consists of a combination of the muons from two semileptonic B decays, or one muon from a semileptonic B decay with a misidentified hadron. The rare B decay background includes both three-body decays such as $B \to K(\pi)\mu\nu$ or $K(\pi)\mu\mu$, and two-body decays such as $B \to K\pi$, $\pi\pi$, or KK events. The background from two-body decays forms a similar peaking structure very close to the expected signal distributions.

Strong background suppression is achieved by a strict muon identification requirement, which is based on a Boosted Decision Tree (BDT) and includes tracking and muon related detector information. Topological and kinematical variables are introduced in the BDT discriminant for event classification. The analysis procedures are validated and calibrated using the $B^+ \to J/\psi K^+$ and $B_s^0 \to J/\psi\phi$ events, while the branching fraction is normalized with respect to $B^+ \to J/\psi K^+$ decays. The analysis benefits a similar trigger as well as selection criteria between signal and normalization decays, allowing the main sources of the systematic uncertainties to be partially cancelled.

Background suppression variables are categorized into those associated with the B candidate (kinematic variables p_T and η, and the minimal distance between two trajectories), those associated with the secondary vertex (impact parameters, flight length, pointing angle, and the quality of the vertex fit), and those associated with isolation (tracking activities around the B candidate and daughter muons). Figure 1 shows the comparison between a simulated sample and data events for the top-three discriminating variables: pointing angle, flight length significance, and the number of tracks around the candidate. The background subtracted distributions from $B^+ \to J/\psi K^+$ and

$B_s^0 \to J/\psi\phi$ candidates show good agreement, while the difference is accounted for as a systematic uncertainty.

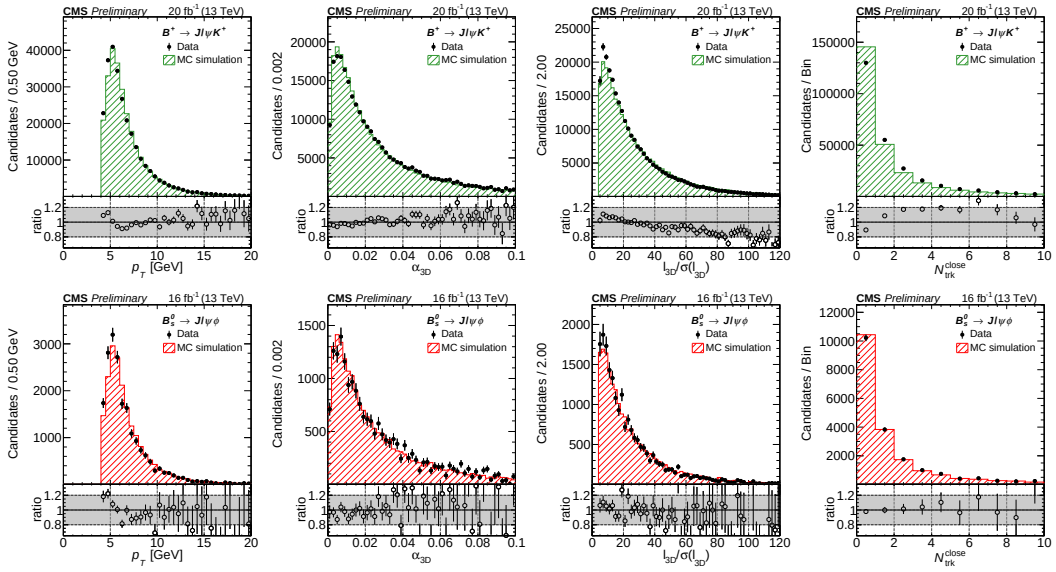

Figure 1: From left to right, comparison of measured and simulated distributions for the B candidate transverse momentum p_t, pointing angle, flight-length significance, and number of tracks around the candidate, respectively, in central-region channel for $B^+ \to J/\psi K^+$ (top) and $B_s^0 \to J/\psi\phi$ (bottom) candidates. Details can be found in Ref. [4].

Categories based on the background suppression BDT distribution are introduced and optimized according to largest significance or smallest lifetime uncertainty of the $B_s \to \mu^+\mu^-$ decay. The significance value and the lifetime uncertainty are calculated using a maximum likelihood estimator based on an asymptotic method. The run-1 categories have been re-optimized. There are 14 categories introduced for the branching fraction measurement and 8 categories for the effective lifetime study.

The yields of the processes $B^+ \to J/\psi K^+$ and $B_s^0 \to J/\psi\phi$ are extracted with fits to the invariant-mass distributions. The signal (combinatorial background) shape is modelled with a sum of two Gaussians (exponential function). The partially reconstructed J/ψ+X events are described with an error function. The $B^+ \to J/\psi\pi^+$ ($B^0 \to J/\psi K^*$) background process is modelled as a sum of three Gaussians (sum of three Gaussians and a linear function). The systematic uncertainty on the yield of the normalization channel $J/\psi K^+$ is estimated to be 4%, mainly derived from the difference of the fits with and without constraining the J/ψ invariant mass.

The branching fraction of $B_s^0 \to \mu^+\mu^-$ is derived according to

$$\mathscr{B}(B_s^0 \to \mu^+\mu^-) = \frac{N_s}{N(B^+ \to J/\psi K^+)} \times \mathscr{B}(B^+ \to J/\psi K^+) \times \frac{A \cdot \varepsilon(B^+)}{A \cdot \varepsilon(B_s^0)} \times \frac{f_u}{f_s}, \qquad (1.2)$$

where N_s and $N(B^+ \to J/\psi K^+)$ are the yields of signal and normalization $B^+ \to J/\psi K^+$ decays, respectively; the $A \cdot \varepsilon(B^+)$ and $A \cdot \varepsilon(B_s^0)$ are the acceptance-time efficiencies for $B^+ \to J/\psi K^+$ and $B_s^0 \to \mu^+\mu^-$; the ratio, f_u/f_s, accounts for the b-quark fragmentation fractions into B^+ and B_s^0

mesons. The branching fraction of $B^0 \to \mu^+\mu^-$ is calculated with a very similar formula, except for the correction for the fragmentation fractions.

The ratio of fragmentation fraction, f_s/f_u, is an external input to this measurement. The PDG value [5], which is an average of LHCb [6] and ATLAS [7] 7 TeV measurements, is introduced, with an additional uncertainty estimated from the LHCb [8] 13 TeV result:

$$f_s/f_u = 0.252 \pm 0.012\,(\text{PDG}) \pm 0.015\,(\text{energy and } \mathrm{p_T} \text{ dependence}). \tag{1.3}$$

The resulting branching fraction can be rescaled with a different value, as knowledge of this ratio improves in future.

1.2 Results

The branching fractions are extracted with a three-dimensional unbinned maximum likelihood fit to the dimuon invariant mass $m_{\mu\mu}$, the per-event mass resolution $\sigma(m_{\mu\mu})$, and a dimuon bending configuration $\mathscr{C}$ (either bending towards or away from each other). The signal B_s^0 and $B^0 \to \mu^+\mu^-$ are modelled by Crystal-ball functions, where the Gaussian width is scaled according to $\sigma(m_{\mu\mu})$. The peaking $B \to hh$ (non-peaking $B \to h\mu\nu, h\mu\mu$) background, where h stands for a charged kaon, pion, or proton, is described by a sum of Crystal-ball function and a Gaussian (a non-parametric keys model with Gaussian kernel). The combinatorial background is fitted by a first order Bernstein polynomial.

The fitted branching fractions are

$$\mathscr{B}(B_s^0 \to \mu^+\mu^-) = [2.9^{+0.7}_{-0.6}\,(\text{exp}) \pm 0.2\,(\mathrm{f_s/f_u})] \times 10^{-9}, \tag{1.4}$$

$$\mathscr{B}(B^0 \to \mu^+\mu^-) = (0.8^{+1.4}_{-1.3}) \times 10^{-10}. \tag{1.5}$$

The observed (expected) significances for B_s^0 and B^0 decays are 5.6σ and 0.6σ (6.5σ and 0.8σ), respectively. The invariant mass distributions with the fit projection overlays, as well as the likelihood contours, are shown in Fig. 2. The results are consistent with the prediction of SM, and are in agreement with the CMS results with run-1 data [9]. Given the significance for $B^0 \to \mu^+\mu^-$ is small, the upper limit based on the full CL_s prescription has been evaluated to be $< 3.6 \times 10^{-10}$ ($< 3.1 \times 10^{-10}$) at the 95% (90%) confidence interval.

The effective lifetime measurement is carried out with two methods, a primary method which introduces a two-dimensional likelihood fit to the proper decay time and dimuon invariant mass, and a second independent analysis based on a one-dimensional binned likelihood fit to the sPlot [10] of the decay time distribution. The model introduced in the 2D likelihood fit adopts the per-event decay time resolution as a conditional parameter in the resolution model, and an efficiency correction is applied to the decay-time line shape. The sPlot method uses the weights derived from the model used in the branching fraction measurement, and a custom implementation of the binned likelihood fit is introduced with proper asymmetric uncertainties. The results, as shown in Fig. 3, from the two methods are found to be consistent with each other, as well as the expectation of the SM:

$$\tau_{\mu\mu}(\text{2D fit}) = 1.70^{+0.61}_{-0.44}\ \text{ps}, \ \text{ and } \ \tau_{\mu\mu}(\text{sPlot}) = 1.55^{+0.52}_{-0.33}\ \text{ps}. \tag{1.6}$$

The systematic uncertainties are smaller than the statistical uncertainties.

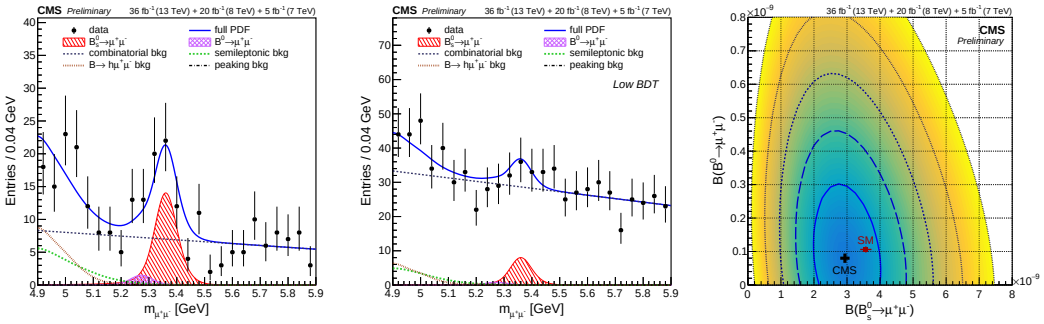

Figure 2: Invariant mass distributions with the fit projection overlays for the branching fraction results, where the left (middle) plot shows the combined results from the high-(low-)range analysis BDT categories. The fit results are shown with the solid lines, and the different background components are shown with the broken lines. The signal components are shown by the hatched distributions. Right plot shows the likelihood contours for the fit to the branching fractions together with the best-fit value (cross) and the SM expectation (solid square). Details can be found in Ref. [4].

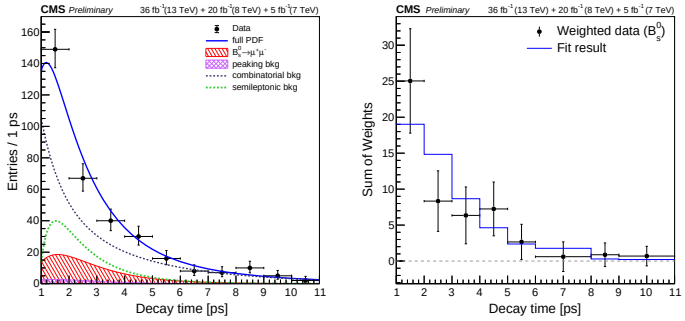

Figure 3: Left plot shows the proper decay time distribution, with the 2D UML fit projections overlaid. The fit results are shown with the solid line, and the different background components are shown with the broken lines and cross-hatched distributions. The signal component is shown by the single-hatched distribution. Right plot presents the proper decay time distribution, with the sPlot fit projections overlaid. Details can be found in Ref. [4].

The rare decay $B_s^0 \to \mu^+\mu^-$ has been measured and the branching fractions have been updated; more data are required for probing $B^0 \to \mu^+\mu^-$ decays. An effective lifetime measurement with $B_s^0 \to \mu^+\mu^-$ events has been carried out for the first time at the CMS experiment. All experimental results [**?**] are consistent with the expectations from the SM.

2. Search for $\tau \to 3\mu$ decays

The decay, $\tau \to 3\mu$, is a charged lepton-flavor violating decay of the τ lepton with no missing neutrino. In the SM, the process is allowed by neutrino oscillations, but the branching fraction is extraordinarily tiny and is beyond experimental accessibility. The decay rate can be strongly enhanced when new physics scenarios are introduced; experimentally the three-muon final state is accessible and clean. Searches have been performed by experiments and no hint of signal has been observed yet; the best limit, $\mathscr{B} < 2.1 \times 10^{-8}$ at the 90% confidence level, has been set by the

Belle [11] experiment. CMS [12] performed a search for the $\tau \to 3\mu$ decay, employing τ leptons produced in D and B hadron decays, using the 13 TeV data set collected in 2016 of 33 fb^{-1}.

A τ candidate is triggered with two muons plus a track, with vertex and mass requirements. At the offline stage three muons are required and the sum of charges should be ± 1. The normalization channel, $D_s^+ \to \phi\pi^+ \to \mu^+\mu^-\pi^+$, is triggered with the same criteria and selected with very similar thresholds. The fraction of non-prompt D_s events is estimated from a fit to the proper decay length distribution. Table 1 summarizes the expected inclusive yields of τ leptons based on the assumption of $\mathcal{B}(\tau \to 3\mu) = 10^{-7}$; the $\mu\mu\pi$ invariant mass distribution and the proper decay length distribution for the τ candidates are shown in Fig. 4.

Table 1: The expected inclusive number of τ leptons produced in D and B meson decays.

	$D \to \tau$ Signal	$B \to \tau$ Signal	Data
Production	4.4×10^5	1.5×10^5	
3τ in fiducial volume	6.6×10^3	2.3×10^3	
Trigger	214	114	
3μ $p_t > 2$ GeV	88	47	1×10^7
3μ candidate	64	29	1×10^5

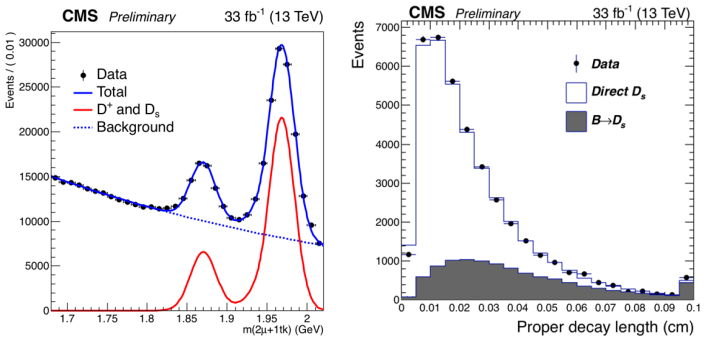

Figure 4: Left plot shows the invariant mass distribution for two muons and a pion after applying signal-like kinematic cuts on two muons and a pion, after requiring that the two muons have opposite signs and their invariant mass is consistent the ϕ meson mass. Right plot presents the fit of the prompt and non-prompt D_s contributions to data. The histograms are from $D_s \to \phi(\mu\mu)\pi$ MC. Details can be found in Ref. [12].

Background suppression relies on the variables associated with the three-μ vertex and those associated with fake muons from hadron decays. The most discriminating variables are the normalized χ^2 from the three-μ vertex fit, significance of the three-μ vertex displacement w.r.t. the interaction point, pointing angle, and the kink parameter for the tracker track. The variables are included in the BDT discriminant trained with signal MC and sideband data as background. The three-μ mass resolution is evaluated event-by-event, ranging from 0.4 to 1.5%, depending on the muon rapidity. The events are categorized based on the mass resolution and the BDT score, resulting in six categories in total.

A simultaneous maximum likelihood fit to the three-μ invariant mass, is performed with all six categories, as shown in Fig. 5. The signal model is parametrized with Crystal-Ball functions, while the background component is modeled with an exponential plus a polynomial. No hint of signal

is found, and the observed (expected) limit is evaluated using the CL_s method, yielding $\mathscr{B}(\tau \to 3\mu) < 1.1 \times 10^{-7}$ (1.2×10^{-7}) at the 95% confidence level. In the limit calculation, systematic uncertainties are incorporated via nuisance parameters. Statistical uncertainties are dominant over the systematic ones.

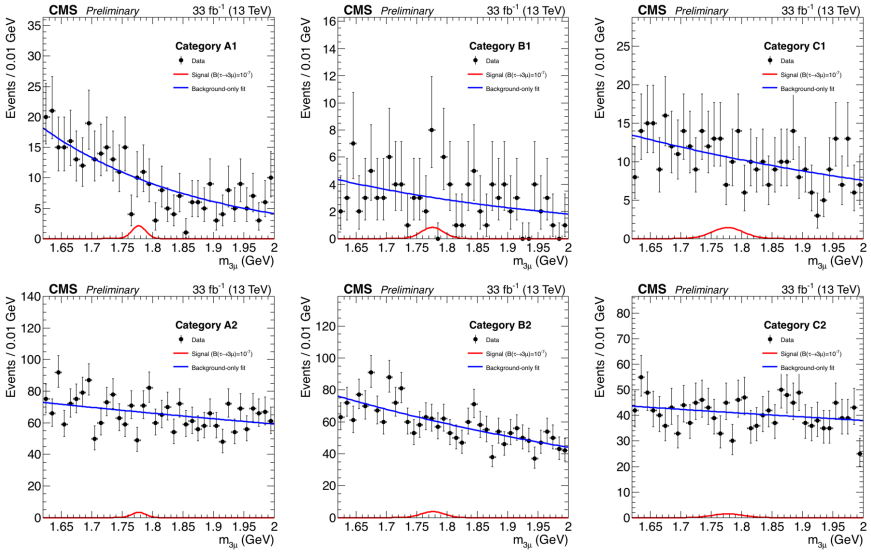

Figure 5: Trimuon mass distributions in the six independent event categories used in the analysis [12]. Data are shown with points. The background-only fit and the expected signal for $\mathscr{B}(\tau \to 3\mu) = 10^{-7}$ are shown.

References

[1] V. Khachatryan *et al.* [CMS and LHCb Collaborations], Nature **522**, 68 (2015) doi:10.1038/nature14474 [arXiv:1411.4413 [hep-ex]].

[2] C. Bobeth, M. Gorbahn, T. Hermann, M. Misiak, E. Stamou and M. Steinhauser, Phys. Rev. Lett. **112**, 101801 (2014) doi:10.1103/PhysRevLett.112.101801 [arXiv:1311.0903 [hep-ph]].

[3] M. Beneke, C. Bobeth and R. Szafron, Phys. Rev. Lett. **120**, no. 1, 011801 (2018) doi:10.1103/PhysRevLett.120.011801 [arXiv:1708.09152 [hep-ph]].

[4] A. M. Sirunyan *et al.* [CMS Collaboration], [arXiv:1910.12127 [hep-ex]].

[5] M. Tanabashi *et al.* [Particle Data Group], Phys. Rev. D 98, 030001 (2018) doi:10.1103/PhysRevD.98.030001.

[6] R. Aaij *et al.* [LHCb Collaboration], JHEP **1304**, 001 (2013) doi:10.1007/JHEP04(2013)001 [arXiv:1301.5286 [hep-ex]].

[7] G. Aad *et al.* [ATLAS Collaboration], Phys. Rev. Lett. **115**, no. 26, 262001 (2015) doi:10.1103/PhysRevLett.115.262001 [arXiv:1507.08925 [hep-ex]].

[8] R. Aaij *et al.* [LHCb Collaboration], Phys. Rev. D **100**, no. 3, 031102 (2019) doi:10.1103/PhysRevD.100.031102 [arXiv:1902.06794 [hep-ex]].

[9] S. Chatrchyan *et al.* [CMS Collaboration], Phys. Rev. Lett. **111**, 101804 (2013) doi:10.1103/PhysRevLett.111.101804 [arXiv:1307.5025 [hep-ex]].

[10] M. Pivk and F. R. Le Diberder, Nucl. Instrum. Meth. A **555**, 356 (2005) doi:10.1016/j.nima.2005.08.106 [physics/0402083 [physics.data-an]].

[11] R. Aaij *et al.* [LHCb Collaboration], Phys. Rev. Lett. **118**, no. 19, 191801 (2017) doi:10.1103/PhysRevLett.118.191801 [arXiv:1703.05747 [hep-ex]].

[12] K. Hayasaka *et al.*, Phys. Lett. B **687**, 139 (2010) doi:10.1016/j.physletb.2010.03.037 [arXiv:1001.3221 [hep-ex]].

[13] CMS Collaboration, CMS-PAS-BPH-17-004 [http://cds.cern.ch/record/2668282].