CERN-LHCb-PUB-2016-017
October 4, 2016

Sensitivity of LHCb and its upgrade in the measurement of $\mathcal{B}(K^0_S \to \pi^0\mu^+\mu^-)$

V. Chobanova[1], X. Cid Vidal[1], J. P. Dalseno[2], M. Lucio Martínez[1], D. Martínez Santos[1], V. Renaudin[3]

[1] *Universidade de Santiago de Compostela, Santiago de Compostela, Spain*
[2] *University of Bristol, Bristol, The United Kingdom*
[3] *Laboratoire de l'Accelerateur Lineaire (LAL), Orsay, France*

Abstract

The sensitivity of the LHCb experiment to $\mathcal{B}(K^0_S \to \pi^0\mu^+\mu^-)$ is analyzed in light of the 2011, 2012 and 2016 data and the oportunities the full software trigger of the LHCb upgrade provides. Two strategies are considered: the full reconstruction of the decay products and the partial reconstruction using only the dilepton pair and kinematic constraints. In both cases, the sensitivity achieved can surpass the world's current best. Both approaches could be statistically combined, further improving the result.

1 Introduction

The $s \to d$ decay processes (see Fig. 1) have the strongest CKM suppression factor of all quark transitions. Hence, they are particularly sensitive to sources of flavour violation different from those of the Standard Model (SM). Indeed, flavour violation can induce detectable effects at accessible energy in flavour-changing processes even if the scale of the new dynamics is heavy and well above their direct production at accelerators. Among these transitions, the decay $K^0_{\mathrm{L}} \to \pi^0\mu^+\mu^-$ has been shown to be sensitive to, for example, models with extra dimensions [1]. However, the potential for this decay to constrain scenarios beyond the Standard Model is limited by the large SM uncertainty on its branching fraction prediction [1],

$$\mathcal{B}(K^0_L \to \pi^0\mu^+\mu^-)_{\mathrm{SM}} = \{1.4 \pm 0.3; 0.9 \pm 0.2\} \times 10^{-11}. \quad (1)$$

The two numbers in the brackets correspond to two theoretical solutions, depending on whether constructive or destructive interference between the contributing waves is present. The reason for the large theoretical uncertainty on $\mathcal{B}(K^0_L \to \pi^0\mu^+\mu^-)_{\mathrm{SM}}$ is the limited precision on the chiral-perturbation-theory parameter $|a_S|$. An improved measurement of $\mathcal{B}(K^0_{\mathrm{S}} \to \pi^0\mu^+\mu^-)$ will reduce this uncertainty. The most precise measurement of $\mathcal{B}(K^0_{\mathrm{S}} \to \pi^0\mu^+\mu^-)$ was performed by the NA48 experiment at CERN [2], which obtained

$$\mathcal{B}(K^0_S \to \pi^0\mu^+\mu^-) = (2.9^{+1.5}_{-1.2}(\mathrm{stat}) \pm 0.2(\mathrm{syst})) \times 10^{-9}. \quad (2)$$

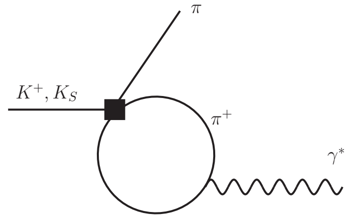

Figure 1: Feynman diagram of the process $K^0 \to \pi^0\mu^+\mu^-$.

The LHCb experiment [3] has demonstrated very good performance in the search for rare leptonic K^0_{S} decays [4]. In this note, we evaluate the potential sensitivity of LHCb to $\mathcal{B}(K^0_{\mathrm{S}} \to \pi^0\mu^+\mu^-)$ considering the data to be collected with the LHCb detector before and after its upgrade in 2018.

This document is organized as follows: in Sect. 2, the analysis strategy is summarized; in Sect. 3, details on the signal reconstruction and selection are given; in Sect. 4, the study on the expected background sources is presented; in Sect. 5, the likelihood fit is described; in Sect. 6, the sensitivity to $\mathcal{B}(K^0_{\mathrm{S}} \to \pi^0\mu^+\mu^-)$ is reported and finally, conclusions are drawn in Sect. 7.

2 Analysis strategy

Decays of the K^0_{S} in LHCb are characterized by decay vertices separated from the interaction point[1], and with tracks having an average transverse momentum significantly

[1]The K^0_{S} at LHC typically decays after traversing tens of centimeters to even several meters.

lower than those from b and c decays. The transverse momentum range is similar to typical tracks generated in the proton-proton collision and hence has almost no discriminating power.

Muon candidates are combined into $\mu^+\mu^-$ pairs. Then a π^0 can be added to the dimuon pair to make a fully reconstructed K^0_S decay. However, since the reconstruction efficiency of the π^0 is limited, events in which no π^0 is found are also considered, based only on the dimuon information. This leads to two independent analyses: one for the events in which all decay products are considered (hereafter FULL) and one in which only the dimuon pair is used (hereafter PARTIAL). The reconstructed candidates are then passed through a selection algorithm followed by a *Boosted Decision Tree* (BDT) classification, to reduce the high level of background.

The properties of the $K^0_S \to \pi^0\mu^+\mu^-$ decays are studied using simulated samples with a differential decay rate modeled according to Ref. [5]. The corresponding $\mu\mu$ mass distribution, $m_{\mu\mu}$, as well as the dependence of the (cosine of the) dimuon helicity angle, $\cos\theta_\mu$ (see the angle definitions in Fig. 2), on $m_{\mu\mu}$ are shown in Fig. 3.

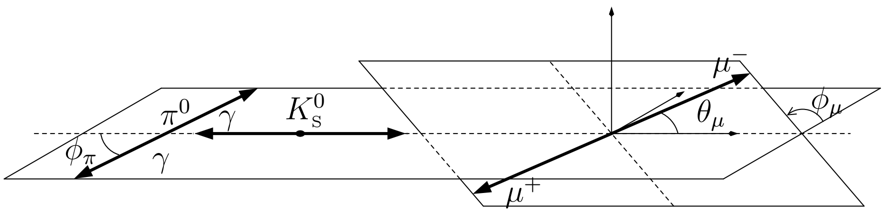

Figure 2: Definition of the helicity angles in the K^0_S rest frame.

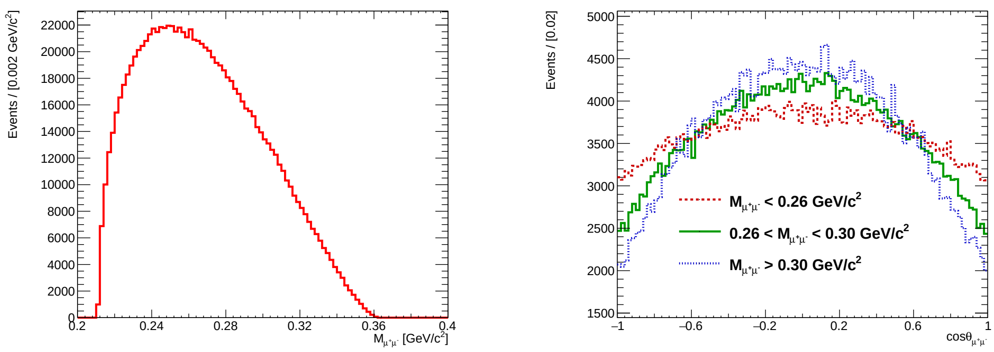

Figure 3: $m_{\mu\mu}$ distribution (left), and the dimuon helicity angle depending on $m_{\mu\mu}$ (right).

The BDT is trained with simulated signal events and combinatorial background events from the existing LHCb data. Since the main goal of this study is to evaluate the sensitivity for the LHCb upgrade, where the trigger efficiency is expected to be very high, trigger unbiased data samples are preferred. Therefore, the events are obtained from the *Trigger Independent of Signal* (TIS) [6] category of the LHCb trigger. This means that the tracks and clusters of the reconstructed candidate are not needed to fire the trigger at any level, because another object in the underlying event already fired it. This ensures an almost trigger unbiased data set, while still providing a sample much larger than random selection triggers.

The expected signal yield is obtained assuming the NA48 central value for $\mathcal{B}(K^0_S \to \pi^0\mu^+\mu^-)$, normalizing the signal yield with respect to $K^0_S \to \pi^+\pi^-$ as

$$\frac{N(K^0_S \to \pi^0\mu^+\mu^-)}{N(K^0_S \to \pi^+\pi^-)} = \frac{\mathcal{B}(K^0_S \to \pi^0\mu^+\mu^-)\epsilon_{K^0_S\to\pi^0\mu^+\mu^-}}{\mathcal{B}(K^0_S \to \pi^+\pi^-)\epsilon_{K^0_S\to\pi^+\pi^-}}, \tag{3}$$

where the observed $K^0_S \to \pi^+\pi^-$ yield is extracted from data and the efficiency ratio, $\frac{\epsilon_{K^0_S\to\pi^0\mu^+\mu^-}}{\epsilon_{K^0_S\to\pi^+\pi^-}}$, is obtained from simulation.

The $\mathcal{B}(K^0_S \to \pi^0\mu^+\mu^-)$ sensitivity is measured in a pseudo-experiment study. First, the signal and background yields are extrapolated for a desired expected luminosity and trigger efficiency, then pseudo-experiments are generated according to those yields. The $\mathcal{B}(K^0_S \to \pi^0\mu^+\mu^-)$ uncertainty is obtained from a fit to the K^0_S mass distribution of the pseudo-experiments, using the signal and background models obtained from MC and the fit to the available LHCb data, respectively. The mass fit range is $[420, 580]\,\mathrm{MeV}/c^2$.

3 Reconstruction and selection

Pairs of muon candidates are reconstructed combining opposite-charged tracks with hits in the vertex locator (VELO), trigger tracker, tracker stations, and muon chambers. In addition, the tracks are required to be separated by at least 6σ from any $p-p$ collision point in the event. Tracks with transverse momentum lower than $80\,\mathrm{MeV}/c$ are ignored. A dimuon candidate pair can be combined with a π^0 candidate to build a K^0_S candidate. The events in which the entire decay chain is used are classified as FULL. When only the dimuon information is used, they are clasified as PARTIAL.

Neutral pion candidates are reconstructed from γ candidate pairs that correspond to two independent clusters in the calorimeter. Each photon candidate is required to have a transverse momentum of at least $200\,\mathrm{MeV}/c$ and the pion candidate a mass within $30\,\mathrm{MeV}/c^2$ of the world average π^0 mass. The mass resolution is then improved by constraining the π^0 candidate mass to the world average π^0 mass, and by constraining the three-momentum vector of the K^0_S to point back to the production vertex. For the PARTIAL candidates, a momentum vector with an absolute value of $\approx 10\,\mathrm{GeV}/c$ is used as a representative of the π^0 momentum when calculating the invariant mass. As a consequence of these kinematic constraints, the K^0_S candidate mass resolution depends only weakly on the π^0 momentum.

Additional selection requirements are applied to reduce the amount of data to analyze, fulfil the rate requirements for LHCb offline processing and reduce the amount of background. These include a K^0_S candidate lifetime of at least 1 ps and removing events in the kinematic region of $\Lambda \to p\pi$ and $K^0_S \to \pi^+\pi^-$ in the Armenteros-Podolanski plane [7]. The total reconstruction and selection efficiency for the FULL channel is 5.47×10^{-4}.

Requiring a well-reconstructed π^0 implies an inefficiency penalty of a factor ten. Thus, a complementary strategy for the PARTIAL candidates is also investigated. Indeed, the constraints on the π^0 mass and the K^0_S momentum are sufficient to create a peaking distribution if there is an estimate of the typical value of the π^0 momentum ($\approx 10\,\mathrm{GeV}/c$), as shown in Fig. 4. A comparison of the reconstructed mass resolution between FULL and PARTIAL is difficult due to the asymmetric and non-Gaussian distribution of the PARTIAL case. To get an estimate, the corresponding FWHM values are calculated. In the FULL case, it is 23.3 MeV/c^2 and in the PARTIAL 40.6 MeV/c^2.

The PARTIAL selection does not require any information about a reconstructed π^0. Some requirements had to be tightened in order to keep the background at a manageable level. These include a lower distance of closest approach between the two muon tracks; a minimum requirement on the K^0_S vertex quality, $\chi^2/ndof = 9$; a higher minimum requirement on the K^0_S vertex detachment from the interaction point; and minimum radial, z- and absolute distance requirements between the K^0_S vertex and the interaction point. The total reconstruction and selection efficiency for the PARTIAL analysis is 3.0×10^{-3} , well above that of the FULL, but at a cost of an increased background yield.

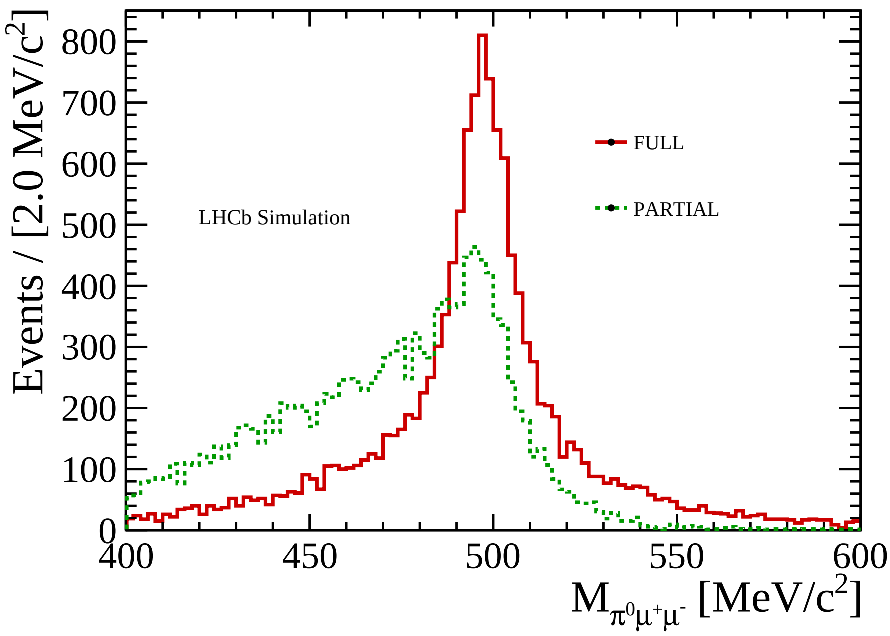

Figure 4: Comparison between the FULL (solid red) and PARTIAL (dashed green) kaon candidate mass distributions.

A BDT is used to separate signal from combinatorial background. It is trained with MC events (signal class) and a part of the data that is not used in the fit (combinatorial background class). The BDT uses information about the geometrical properties of the events, kinematics, track quality, and muon identification quality. The BDT response for signal and background for both FULL and PARTIAL is shown in Fig. 5.

The events are classified in four bins of the BDT response. The signal yields are obtained in a simultaneous fit of the mass distribution in each BDT bin, as described in the following sections.

4 Background sources

Several sources of background are investigated to assess their relevance for a measurement of $\mathcal{B}(K^0_S \to \pi^0\mu^+\mu^-)$:

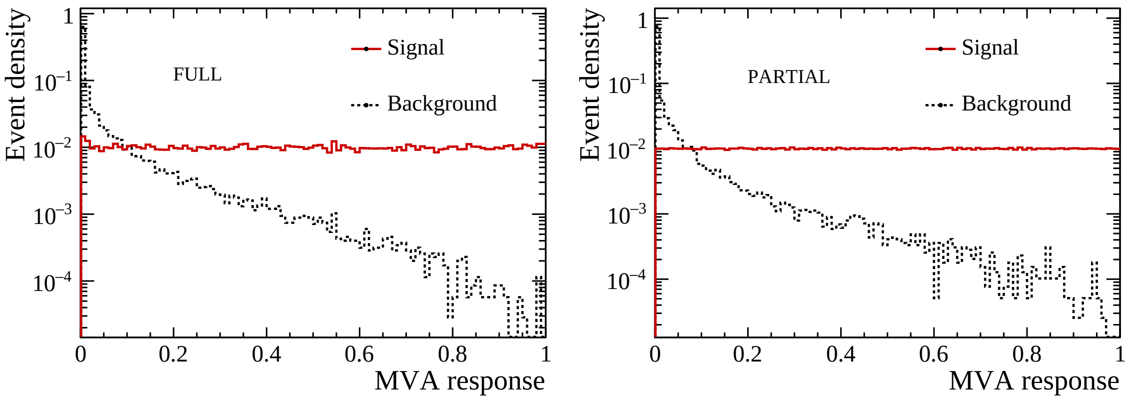

Figure 5: BDT response both for signal (solid red) and background (dashed black). Right: FULL channel. Left: PARTIAL channel. Signal and background are normalized to the same area.

- $K^0_{\rm S} \to \pi^+\pi^-$ decays, where both pions are misidentified as muons, and in the case of the FULL category, combined with a random π^0 from the underlying event. These decays have a mass larger than that of the $K^0_{\rm S}$ and do not enter the fit region, except for potential residual tails that effectively add up to the combinatorial background. No evidence for $K^0_{\rm S} \to \pi^+\pi^-$ background is seen for the BDT region being fitted.

- $K^0 \to \mu^+\mu^-\gamma\gamma$ decays. This background was considered in the NA48 analysis [2], However, its contribution at LHCb is found to be negligible: In the case of the $K^0_{\rm L}$ decay (with a branching fraction of $1.0^{+0.8}_{-0.6} \times 10^{-8}$ [8]) the upper decay time acceptance introduces an effective 10^{-3} reduction with respect to $K^0_{\rm S}$ and hence the effective $\mathcal{B}(K^0_{\rm L} \to \mu^+\mu^-\gamma\gamma)$ becomes as low as 10^{-11}. There is no experimental measurement of $\mathcal{B}(K^0_{\rm S} \to \mu^+\mu^-\gamma\gamma)$, however, since the process is dominated by the two-photon exchange[2], it can be estimated as:

$$\mathcal{B}(K^0_{\rm S} \to \mu^+\mu^-\gamma\gamma) = \frac{\mathcal{B}(K^0_{\rm S} \to \gamma\gamma)}{\mathcal{B}(K^0_{\rm L} \to \gamma\gamma)}\mathcal{B}(K^0_{\rm L} \to \mu^+\mu^-\gamma\gamma) \sim 4.8 \times 10^{-11} \quad (4)$$

  and is thus negligible.

- $K^0_{\rm L} \to \pi^0\pi^+\pi^-$ decays. The mass distribution of these decays is shown in Fig. 6 as obtained in simulation. Since there is no evidence of this background in the data, it is neglected. Including a $K^0_{\rm L} \to \pi^0\pi^+\pi^-$ component to the observed background does not change significantly the sensitivity estimates. The $K^0_{\rm S}$ counterpart has a branching fraction of 3.5×10^{-7} and thus is about four orders of magnitude smaller than $K^0_{\rm L} \to \pi^0\pi^+\pi^-$. In general, no sign of a resonant structure in the $\pi^+\pi^-\pi^0$ is seen on data.

- Combinatorial background. Combinatorial background is considered to be composed by random combination of tracks, including those generated by pseudo-random combinations of hits during the pattern recognition. It has a monotonic shape across the studied invariant mass range.

[2]Isidori and D'Ambrosio, private communication.

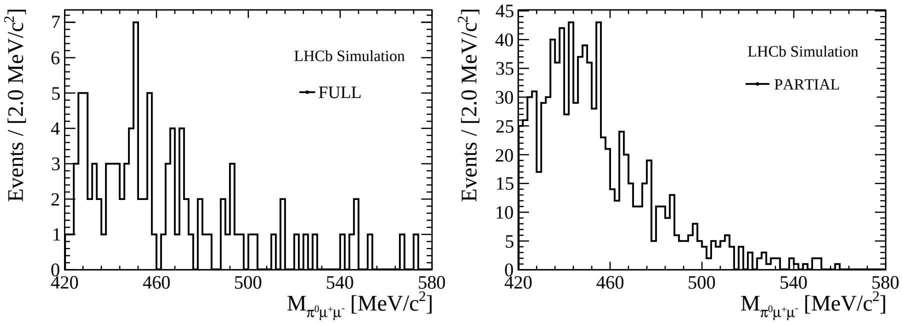

Figure 6: Invariant mass distribution of simulated $K^0 \to \pi^+\pi^-\pi^0$ decays selected in the FULL (left) and PARTIAL (right) categories.

5 Fit model

Only events in the BDT range [0.6,1] are considered in the fit to the data. A simultaneous fit to the mass distribution across four equally-sized independent bins of the BDT response is performed. The combinatorial background is described with an exponential PDF for both FULL and PARTIAL analysis, with independent floating yields and decay constants in each BDT bin. The signal model is an Hypathia distribution [9] with different configurations for FULL and PARTIAL (see Fig. 7). The signal model parameters are independent in each BDT bin and are obtained from simulation. The fractions of signal events allotted to each BDT bin are also fixed from values obtained from simulation, with a total signal yield remaining as the sole free parameter describing signal in the simultaneous fit. The signal yield is floated in the fit to the data. It is measured to be compatible with zero within one to two sigma. The fit projections to the FULL and PARTIAL data are shown in Fig. 8.

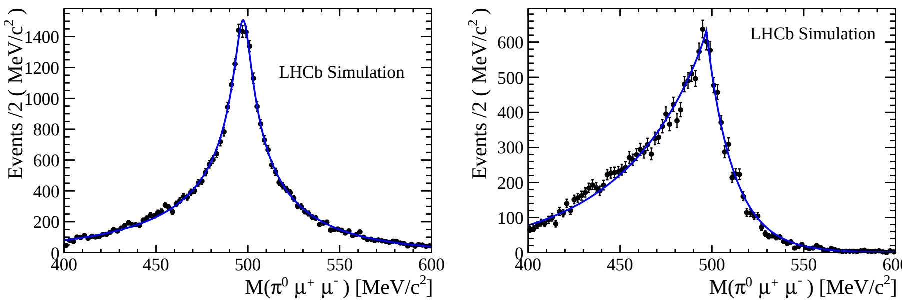

Figure 7: Signal fit using the Hypathia function for FULL (left) and PARTIAL (right) categories.

6 Expected sensitivity

The expected statistical precision on $\mathcal{B}(K^0_S \to \pi^0\mu^+\mu^-)$ for multiple values of the integrated luminosity up to 100 fb^{-1} is estimated in this section. The TIS samples used are equivalent

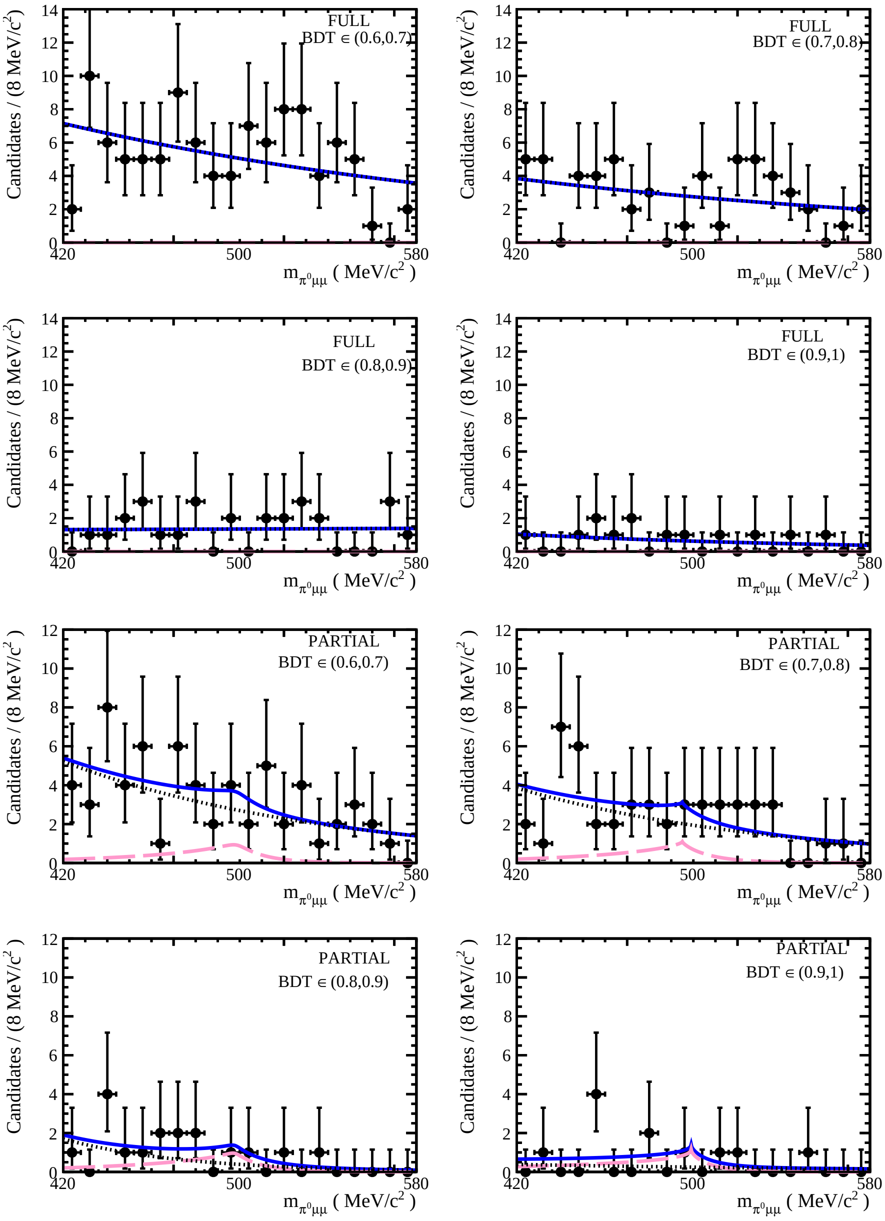

Figure 8: Fit to data for FULL (top) and PARTIAL (bottom) categories. The magenta dashed line shows the signal contribution, the dotted black line the background, and the solid blue line the prediction from the total fit model.

to a 100% trigger efficiency sample with an integrated luminosity of 4.9 and 0.77 pb^{-1} for the FULL and PARTIAL samples, respectively. The expected background yield is extrapolated from the current data fit result, where the signal yield is consitent with zero. The background yield is scaled linearly for larger integrated luminosities.

For each integrated luminosity in the studied range, sets of pseudo-experiments are generated with the above background expectations, and with a signal yield expectation of

$$N_{sig} = \frac{\mathcal{B}(K^0_\mathrm{S} \to \pi^0\mu^+\mu^-)}{\mathcal{B}(K^0_\mathrm{S} \to \pi^+\pi^-)} \frac{\epsilon_{K^0_\mathrm{S}\to\pi^0\mu^+\mu^-}}{\epsilon_{K^0_S\to\pi^+\pi^-}} N(K^0_\mathrm{S} \to \pi^+\pi^-) \times \frac{L_{fut}}{L_{curr}}, \tag{5}$$

where L_{fut} and L_{curr} are the future and current luminosities, respectively. The models described in Sect. 5 are fit to each pseudo-experiment with a floating $\mathcal{B}(K^0_S \to \pi^0\mu^+\mu^-)$. The background model parameters used are the ones obtained from the fit to the data Sect. 5. The statistical uncertainties are obtained as the variations of $\mathcal{B}(K^0_S \to \pi^0\mu^+\mu^-)$ that deviate from the minimum of the log-likelihood profile by half a unit. Finally, the uncertainties are averaged across the set of pseudo-experiments for a given integrated luminosity. The uncertainties on the background extrapolation are large and translate into large uncertainties on the luminosity needed for achieving a given sensitivity. The resulting sensitivity curves are shown in Fig. 9. It can be seen that the analyses of both PARTIAL and FULL categories can lead to a precision better than NA48 for the LHCb upgrade if a trigger efficiency above $\approx 50\%$ can be maintained. The K^0_S production cross-section increases by $\approx 20\%$ at 14 TeV compared to 8 TeV, but this increase is cancelled by a larger fraction of K^0_S decaying outside of the VELO volume. For this reason, no energy correction has been applied to the sensitivity estimate. Studies of $K^0_S \to \pi^0\mu^+\mu^-$ and minimum bias samples simulated with the LHCb upgrade detector and conditions show that the High Level Trigger rate can be kept low enough for a 100 % efficiency. Further timing studies are currently ongoing.

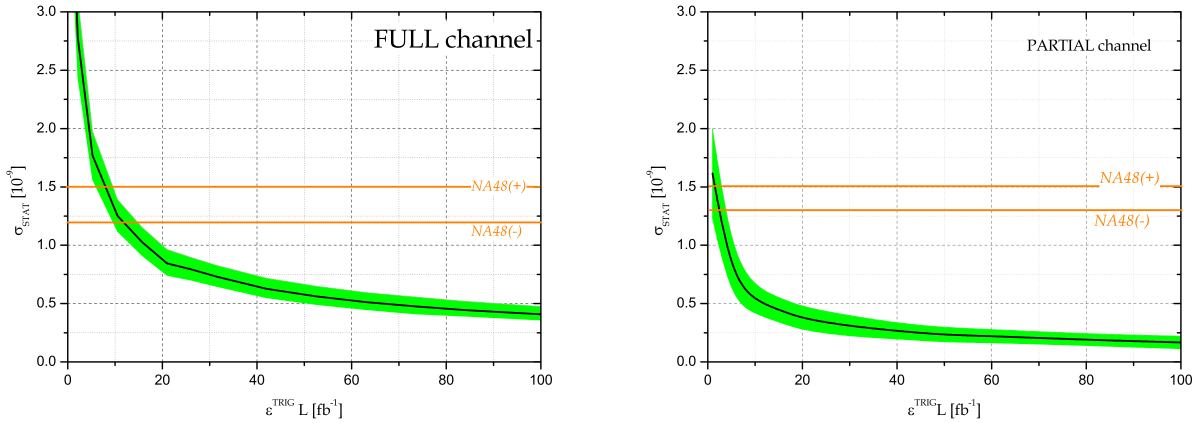

Figure 9: Expected precision on $\mathcal{B}(K^0_S \to \pi^0\mu^+\mu^-)$ for the FULL (top) and PARTIAL (bottom) channels, as a function of the integrated luminosity times trigger efficiency, $L \times \varepsilon^{TRIG/SEL}$.

7 Conclusions

A precise measurement of the $K^0_{\rm S} \to \pi^0\mu^+\mu^-$ branching fraction is crucial for a precise $\mathcal{B}(K^0_L \to \pi^0\mu^+\mu^-)$ SM theoretical prediction and the search for physics beyond the SM in $K^0_L \to \pi^0\mu^+\mu^-$. The sensitivity of the LHCb experiment to $\mathcal{B}(K^0_{\rm S} \to \pi^0\mu^+\mu^-)$ was studied based on 3 fb^{-1} of data recorded at 7 and 8 TeV center-of-mass energy during 2011 and 2012, and on 0.3 fb^{-1} of data recorded at 13 TeV center-of-mass energy during 2016. Full and partial decay reconstruction algorithms were considered, aiming at a high reconstruction efficiency. The sensitivity study was performed using pseudo-experiments by extrapolating signal yield results based on the currently available data to expected future integrated luminosities. If a trigger efficiency of at least 50% can be assured in the future, LHCb can determine $\mathcal{B}(K^0_{\rm S} \to \pi^0\mu^+\mu^-)$ with a precision significantly better than that of NA48.

Acknowledgements

We would like to thank Gino Isidori, Giancarlo D'Ambrosio, and Ikaros Bigi for their theoretical input. We would like to thank Teresa Fonseca for details on the NA48 analysis. We express our gratitude to our colleagues in the CERN accelerator departments for the excellent performance of the LHC. We thank the scientifical, technical and administrative staff at the LHCb institutes. We acknowledge support from ERC (grant number ERC-StG-639068), EPLANET, and Xunta de Galicia. We would like to dedicate this work to the memory of Pablo Rodríguez.

References

[1] M. Bauer, S. Casagrande, U. Haisch, and M. Neubert, *Flavor Physics in the Randall-Sundrum Model: II. Tree-Level Weak-Interaction Processes*, JHEP **09** (2010) 017, `arXiv:0912.1625`.

[2] NA48/1, J. R. Batley *et al.*, *Observation of the rare decay* $K_S \to \pi^0\mu^+\mu^-$, Phys. Lett. **B599** (2004) 197, `arXiv:hep-ex/0409011`.

[3] T. L. Collaboration, *The LHCb Detector at the LHC*, Journal of Instrumentation **3** (2008), no. 08 S08005.

[4] LHCb, R. Aaij *et al.*, *Search for the rare decay* $K_S \to \mu^+\mu^-$, JHEP **01** (2013) 090, `arXiv:1209.4029`.

[5] G. D'Ambrosio, G. Ecker, G. Isidori, and J. Portoles, *The Decays* $K \to \pi l^+ l^-$ *beyond leading order in the chiral expansion*, JHEP **08** (1998) 004, `arXiv:hep-ph/9808289`.

[6] E. L. A. *et al*, *Measurement of trigger efficiencies and biases*, LHCb public note CERN-LHCB-2008-073 (2008).

[7] R. Armenteros and J. Podolanski, *Analysis of V-events*, Phil. Mag. **45** (1954) 13.

[8] Particle Data Group, K. A. Olive *et al.*, *Review of particle physics*, Chin. Phys. **C38** (2014) 090001, and 2015 update.

[9] D. Martínez Santos and F. Dupertuis, *Mass distributions marginalized over per-event errors*, Nucl. Instrum. Meth. **A764** (2014) 150, `arXiv:1312.5000`.