Evidence for the rare decay $\Sigma^+ \to p\mu^+\mu^-$ at LHCb

To cite this article: Francesco Dettori 2017 *J. Phys.: Conf. Ser.* **800** 012021

View the article online for updates and enhancements.

Evidence for the rare decay $\Sigma^+ \to p\mu^+\mu^-$ at LHCb

Francesco Dettori

CERN European Organization for Nuclear Research, CH1211 Geneve 23, Switzerland

E-mail: `francesco.dettori@cern.ch`

Abstract. A search for the rare decay $\Sigma^+ \to p\mu^+\mu^-$ is performed using pp collision data recorded by the LHCb experiment at centre-of-mass energies $\sqrt{s} = 7$ and 8 TeV, corresponding to an integrated luminosity of $3\,\text{fb}^{-1}$. An excess of events is observed with respect to the background expectations with a signal significance of 4.0 standard deviations. No significant structure is observed in the dimuon invariant mass distribution.

1. Introduction

The $\Sigma^+ \to p\mu^+\mu^-$ decay[1] is a flavour changing neutral current process, allowed only at loop level in the standard model (SM). The process is dominated by long-distance contributions for a predicted branching fraction of $\mathcal{B}(\Sigma^+ \to p\mu^+\mu^-) \in [1.6, 9.0] \times 10^{-8}$ [1], while the short-distance SM contributions are suppressed at a branching fraction of about 10^{-12}. Evidence for this decay was seen by the HyperCP experiment [2] with a measured branching fraction $\mathcal{B}(\Sigma^+ \to p\mu^+\mu^-) = (8.6^{+6.6}_{-5.4} \pm 5.5) \times 10^{-8}$, which is compatible with the SM. Remarkably, the three observed decays have almost the same dimuon pair invariant mass of $m_{X^0} = 214.3 \pm 0.5\,\text{MeV}/c^2$, which lies close to the kinematic limit. Such a distribution, if confirmed, would point towards a decay with an intermediate particle coming from the Σ^+ baryon and decaying into two muons, *i.e.* a $\Sigma^+ \to pX^0(\to \mu^+\mu^-)$ decay, which would constitute evidence of physics beyond the SM (BSM). Various BSM explanations have been proposed to explain the HyperCP result. The intermediate X^0 particle could be for example a light pseudoscalar Higgs boson [3] [4] in a NMSSM or 2HDM model, or a sgoldstino [5] [6] in various supersymmetric models. Other interpretations and implications can be found in Ref. [7] [8] [9] [10] [11] [12] [13]; in general a pseudoscalar particle is favoured over a scalar one and a lifetime of order 10^{-14} s is estimated for the former case. This lifetime would correspond to a prompt X^0 signal, decaying in the same vertex as the Σ^+ baryon, in any present search for this particle. Attempts to confirm the presence of this X^0 particle have been made by different experiments in various initial and final states without finding any signal [14–21], and in LHCb through the final states $B^0_{(s)} \to \mu^+\mu^-\mu^+\mu^-$ [22] and $B^0 \to K^{*0}\mu^+\mu^-$ [23]. However, the search for the $\Sigma^+ \to p\mu^+\mu^-$ decay has not been repeated, mainly due to the absence of experiments with large hyperon production and due to experimental difficulties.

Hyperons are produced copiously in high energy proton-proton collisions at the Large Hadron Collider (LHC). A search for $\Sigma^+ \to p\mu^+\mu^-$ decays at the LHCb experiment could therefore confirm or disprove the HyperCP evidence, and measure its branching fraction. This search was also suggested in Ref. [24]. In this report, the search for the $\Sigma^+ \to p\mu^+\mu^-$ decay is presented as performed using pp collision data recorded by the LHCb experiment at centre-of-mass energies

---

[1] The inclusion of charge conjugated processes is implied throughout this note.

$\sqrt{s} = 7$ and $8\,\mathrm{TeV}$, corresponding to an integrated luminosity of $3\,\mathrm{fb}^{-1}$. The measurement described here is detailed in Ref. [25].

2. Detector and data samples

The LHCb detector is a single-arm forward spectrometer covering the pseudorapidity range $2 < \eta < 5$, designed for the study of particles containing b or c quarks, and is detailed elsewhere [26] [27]. The online event selection is performed by a trigger, which consists of a hardware stage, based on information from the calorimeter and muon systems, followed by two software stages, the first performing a preliminary event reconstruction based on partial information while the second applying a full event reconstruction. Each of the three trigger stages is divided into several parallel trigger lines dedicated to different kinds of signals. Decay particles involved in this analysis are often not able to activate one or more trigger stages, owing mainly to their soft transverse momenta. Nevertheless, since Σ^+ baryons are copiously produced in the pp collisions recorded by LHCb, the present search can be performed on the events already recorded. In the offline processing, trigger decisions are associated with reconstructed candidates. A trigger decision on a particular line or on the full trigger can thus be ascribed to the reconstructed candidate, the rest of the event or both; events triggered as such are defined respectively as triggered on signal (TOS), independently of signal (TIS), or triggered on both. The estimation of the trigger efficiency when no specific path is selected is difficult, therefore a different strategy is adopted here: all the candidates passing the selection are used in the search for $\Sigma^+ \to p\mu^+\mu^-$ decays, while only the TIS candidates are used to convert the event yield into a branching fraction. This ensures a partial cancellation of the TIS trigger efficiency between signal and normalisation channels. Furthermore, control channels with large statistics can be exploited to estimate the trigger efficiency directly on data by measuring the overlap of events which are TIS and TOS simultaneously [28].

3. Analysis strategy and selection

After the trigger, a loose selection is applied based on geometric and kinematic variables. Afterwards candidates are selected by means of a multivariate selection based on a boosted decision trees algorithm (BDT) [29] [30]. The final search datasets are obtained by rejecting the background with a cut on the BDT output and on particle identification variables. The signal yield is obtained from a fit to the $p\mu^+\mu^-$ invariant mass and is converted into a branching fraction by normalising to the $\Sigma^+ \to p\pi^0$ control channel. To avoid experimenter bias, candidates in the signal regions were not examined until the analysis procedure had been finalised.

The analysis is designed in order to search for possible peaks in the dimuon invariant mass, pointing towards unknown intermediate particles. The resolution on the dimuon invariant mass as a function of the mass itself is shown in Figure 1(a). The selection is devised such that no fake structures are induced in the dimuon invariant mass distribution of signal candidates. The signal efficiency varies as a function of the dimuon invariant mass and is shown in Figure 1(b). The efficiency is larger at lower dimuon mass owing to the larger recoil momentum against the proton, which gives larger minimum transverse momentum to the two muons, ensuring a better tracking efficiency.

Together with the signal $\Sigma^+ \to p\mu^+\mu^-$ decay, the following channels are also selected in data: the $\Sigma^+ \to p\pi^0$ decay as normalisation channel, the $\Sigma^+ \to \bar{p}\mu^+\mu^+$ candidate decay as control channel for the combinatorial background, and the $K^+ \to \pi^+\pi^-\pi^+$ decay as a control channel for different parts of the efficiency evaluation. Candidate $\Sigma^+ \to p\mu^+\mu^-$ decays are selected by combining two good quality oppositely charged tracks with muon identification with a third track with proton identification. The three tracks are required to form a good quality secondary vertex (SV), displaced from any pp interaction vertex (PV) requiring a measured Σ^+ lifetime greater than 6 ps. Only Σ^+ candidates with transverse momentum $p_{\mathrm{T}} > 0.5\,\mathrm{GeV}/c$, $\chi^2_{\mathrm{IP}} < 36$ and cosine of the angle between the flight direction and the reconstructed momentum larger than 0.9 are retained. Candidate $\Sigma^+ \to p\mu^+\mu^-$ decays are considered only if the invariant mass
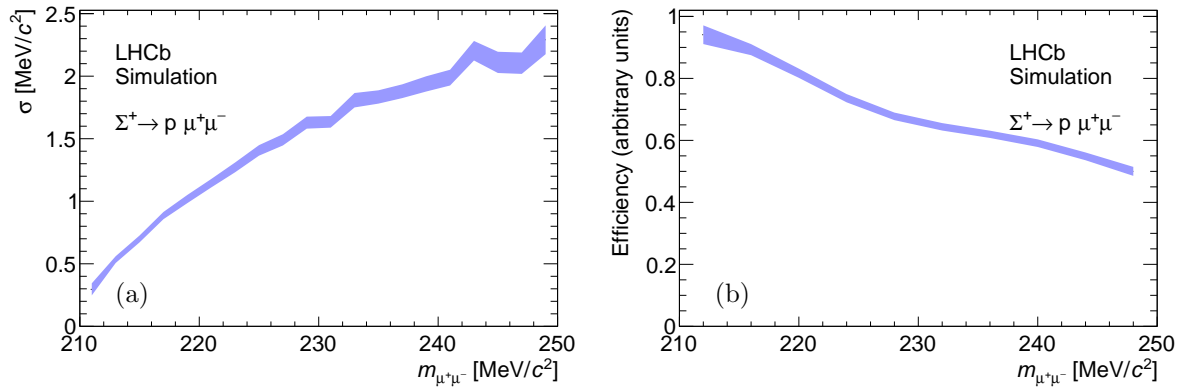
Figure 1: Resolution (a) on the dimuon invariant mass $m_{\mu^+\mu^-}$ as a function of the mass itself for simulated $\Sigma^+ \to p\mu^+\mu^-$ decays in LHCb. Efficiency (b) for simulated $\Sigma^+ \to p\mu^+\mu^-$ decays in LHCb as a function of the dimuon invariant mass.

satisfies $|m_{p\mu^+\mu^-} - m_{\Sigma^+}| < 500\,\mathrm{MeV}/c^2$, where m_{Σ^+} is the known mass of the Σ^+ particle [31]. A large background component is present in data due to $\Lambda^0 \to p\pi^-$ decays, where the pion is misidentified as a muon and the two tracks are combined with a third one. This is vetoed discarding candidates having a $p\mu^-$ pair mass within 5 MeV/c^2 from the Λ^0 known mass when calculated with the $p\pi^-$ mass hypothesis. Possible backgrounds from exclusive decays peaking in the $p\mu^+\mu^-$ invariant mass have been examined, including $K^+ \to \pi^+\pi^-\pi^+$ and $K^+ \to \pi^+\mu^-\mu^+$ decays and various hyperon decays, and none has been found to contribute. In case of multiple candidates in a single event, all are retained in the selection. Multiple candidates are present in 5% of the events after the initial selection, while no multiple candidate is present in the final selection. The selection for the control channel $\Sigma^+ \to \bar{p}\mu^+\mu^+$ is identical to that of the signal but considering same-sign dimuon pairs. Candidate $\Sigma^+ \to p\pi^0$ decays are selected by combining one good quality track with proton identification with a π^0 reconstructed in the $\pi^0 \to \gamma\gamma$ mode from two clusters in the calorimeter. The selection of this decay is similar to that of the signal, but places a tighter requirement on the proton identification, and on the transverse momenta of the daughters in order to reduce the high combinatorial background. Finally, candidate $K^+ \to \pi^+\pi^-\pi^+$ decays are selected from three good quality tracks, one of which is oppositely charged with respect to the other two; the selection is similar but tighter than that of the signal to cope with the high level of combinatorial background. See Ref. [25] for further details on the selection criteria of the different channels.

The sample of $\Sigma^+ \to p\mu^+\mu^-$ candidates in data after the selection is dominated by combinatorial background, part of which is due to misidentified particles. This is rejected by cuts on the BDT and on multivariate particle identification variables [27] on the muons and on the proton. The BDT variable combines geometric and kinematic variables chosen so that the dependence on the $p\mu^+\mu^-$ invariant mass and on the dimuon invariant mass is linear and small to avoid biases. The BDT is optimised using simulated samples of $\Sigma^+ \to p\mu^+\mu^-$ events for the signal and candidates from the sidebands of the $\Sigma^+ \to \bar{p}\mu^+\mu^+$ selection in data for the background. The final cut values were chosen in order to optimise the sensitivity to a signal evidence for the smallest possible branching fraction [32]. No BDT selection is applied to the normalisation and control channels.

X International Conference on Kaon Physics

IOP Publishing

IOP Conf. Series: Journal of Physics: Conf. Series **800** (2017) 012021

doi:10.1088/1742-6596/800/1/012021

4. Normalisation

The number of signal candidates in the TIS sample is converted into a branching fraction with the formula

$$\mathcal{B}(\Sigma^+ \to p\mu^+\mu^-) = \frac{\varepsilon_{\Sigma^+ \to p\pi^0}}{\varepsilon_{\Sigma^+ \to p\mu^+\mu^-}} \frac{N_{\Sigma^+ \to p\mu^+\mu^-}}{N_{\Sigma^+ \to p\pi^0}} \mathcal{B}(\Sigma^+ \to p\pi^0) = \alpha \cdot N_{\Sigma^+ \to p\mu^+\mu^-} \quad ,$$

where ε, N and $\mathcal{B}$ are the efficiency, candidate yield and branching fraction of the corresponding channel, respectively, and α is the single event sensitivity.

The ratio of signal and normalisation channel efficiencies, which includes the acceptance, the reconstruction efficiency of the final state particles and the selection efficiency, is computed with samples of simulated events corrected to take into account known differences between data and simulation. The reconstruction efficiency for the π^0 is calibrated using the ratio of $B^+ \to J/\psi K^{*+}(\to K^+\pi^0)$ and $B^+ \to J/\psi K^+$ decays reconstructed in data. The particle identification efficiency of protons and muons is calibrated exploiting control channels in data. Residual differences between data and simulation are treated as sources of systematic uncertainty. The signal and normalisation channels are required to be TIS at all trigger levels. The trigger efficiency ratio is thus expected to be unity. However, small differences in the average kinematic of the rest of the event in the two samples are present which cause the ratio to be different. The ratio of trigger efficiencies is thus evaluated with data-driven techniques [28] exploiting the large sample of $K^+ \to \pi^+\pi^-\pi^+$ decays. A systematic uncertainty is assigned for the applicability of this method to the relatively soft events of this analysis.

The invariant mass distribution of the $K^+ \to \pi^+\pi^-\pi^+$ control channel candidates in data is shown in Figure 2(a). A binned maximum likelihood extended fit is performed to the invariant mass distribution. The signal is described as an Hypatia function [33] while the background is described by a second-order polynomial. A total of $(966 \pm 2) \times 10^3$ $K^+ \to \pi^+\pi^-\pi^+$ candidates is measured.
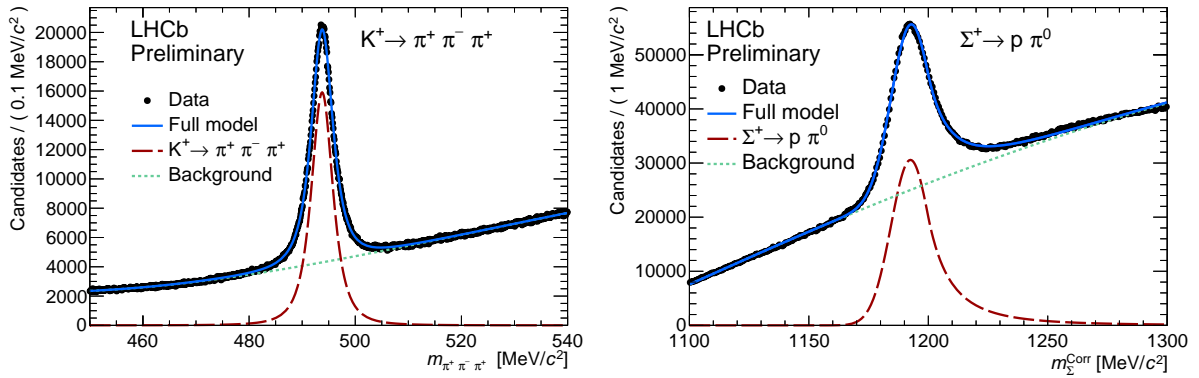

Figure 2: Invariant mass distribution of (a) $K^+ \to \pi^+\pi^-\pi^+$ candidates superimposed with the fit to data; (b) fit to the distribution of the corrected mass m_{Σ}^{corr} for $\Sigma^+ \to p\pi^0$ candidates superimposed with the fit to data.

The observed number of $\Sigma^+ \to p\pi^0$ candidates is $(1\,711 \pm 9) \times 10^3$ as obtained from a binned maximum likelihood extended fit to the corrected invariant mass distribution. The corrected invariant mass is defined as $m_{\Sigma}^{\text{corr}} = m_{p\pi^0} - m_{\pi^0} + m_{\pi^0}^{\text{PDG}}$ to correct for the π^0 mass reconstructed from the two photons. The $\Sigma^+ \to p\pi^0$ distribution is described as a Gaussian function with a power tail on the right side, while the background is described by a modified ARGUS function [34]. The invariant mass distribution is shown in Figure 2(b), superimposed with the fit. The single event sensitivity is $\alpha = (1.1 \pm 0.6) \times 10^{-8}$, where the uncertainty is

dominated by the aforementioned systematic uncertainties, and corresponding to about 4×10^{11} Σ^+ particles produced in the LHCb acceptance in the full dataset in TIS events. This corresponds to 4.6 ± 4.2 $\Sigma^+ \to p\mu^+\mu^-$ candidates assuming a branching fraction of $(5 \pm 4) \times 10^{-8}$, to cover the SM predicted range.

The observed number of signal $\Sigma^+ \to p\mu^+\mu^-$ events is obtained with a fit to the $p\mu^+\mu^-$ invariant mass distribution in the range $1149.6 < m_{p\mu^+\mu^-} < 1409.6 \, \mathrm{MeV}/c^2$. The signal distribution is described by an Hypatia function [33]. The mass resolution and scale are calibrated using the control channel $K^+ \to \pi^+\pi^-\pi^+$ and comparing data and simulation distributions. No bias is seen in the peak position, while a 25% correction to the resolution has to be applied to match the one observed in data. A resolution of $4.28 \pm 0.19 \, \mathrm{MeV}/c^2$ is used as width of the signal $\Sigma^+ \to p\mu^+\mu^-$ distribution. The resolution is allowed to vary in the fit but constrained to the central value with a Gaussian constraint. The combinatorial background is described as a modified ARGUS function, with all parameters left free with exception of the threshold which is fixed to the kinematic limit.

While the normalisation is available for the TIS sample, further studies are needed to calculate the trigger efficiency for the full sample. Since the TIS dataset is a sub-sample of the full sample, the single event sensitivity of the latter will be equal or better than the one of the former.

5. Results and discussion

The invariant mass distribution of the $\Sigma^+ \to p\mu^+\mu^-$ candidates in data is shown in Figure 3(a) and (b) for the full and TIS datasets, respectively. A significant signal is present in the full dataset. The significance of the signal is of $4.0\,\sigma$, obtained from the comparison of the likelihood value of the full fit with that of the background-only fit, and includes the relevant systematic uncertainties as gaussian constrains to the likelihood. A total of $12.9^{+5.1}_{-4.2}$ signal candidates is observed. The signal in the TIS sample is found not to be significant. This is due to the observed signal candidates being needed for the related event to switch on at least one of the LHCb trigger stages. From the absence of a significant signal in the TIS dataset an upper limit on the $\Sigma^+ \to p\mu^+\mu^-$ branching fraction is derived, using the CLs method [35], of $\mathcal{B}(\Sigma^+ \to p\mu^+\mu^-) < 6.3 \times 10^{-8}$ at 95% confidence level (CL).

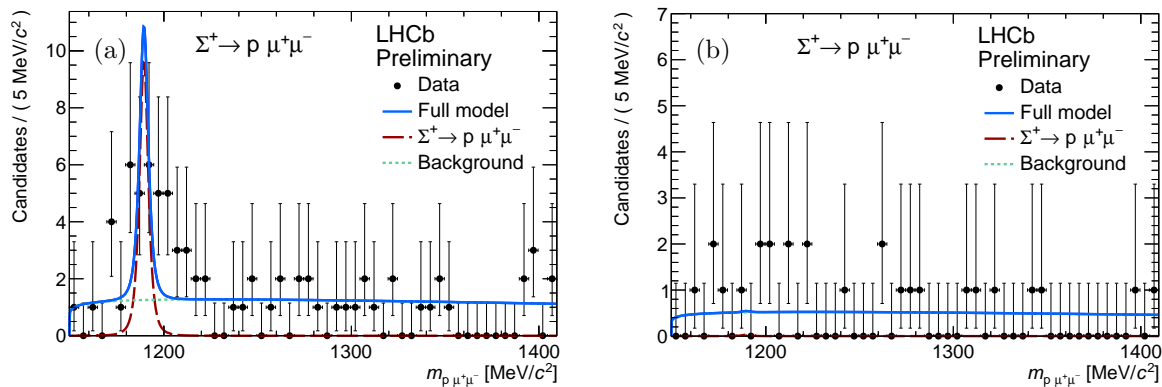
Figure 3: Invariant mass distribution for $\Sigma^+ \to p\mu^+\mu^-$ candidates: (a) all the candidates and (b) only the candidates satisfying the TIS requirement.

Considering candidates in the full selection, a scan for possible signals in the dimuon invariant mass is performed, restricted to within two times the resolution in the $\Sigma^+ \to p\mu^+\mu^-$ invariant mass around the known Σ^+ mass. The scan is performed considering a single Gaussian function for a putative X^0 resonance (signal) and a linear distribution for the remaining candidates (background). In Figure 4(a) the local p-value of the background-only hypothesis as a function

of the dimuon mass is shown. No significant signal is found. In Figure 4(b) the fit corresponding to a mass of $214.3\,\mathrm{MeV}/c^2$ is shown. The fitted number of events for this hypothesis is 1.6 ± 1.9, corresponding to a fraction 0.078 ± 0.092 of the considered candidates.
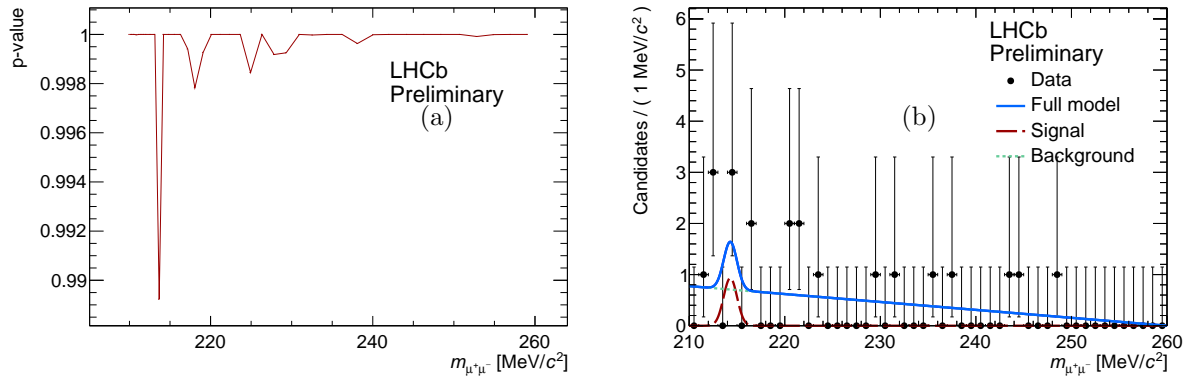

Figure 4: Plot of the local p-value as a function of the dimuon mass as obtained from the fits (see text) in (a) and (b) fit for a signal with mass $m_{X^0} = 214.3\,\mathrm{MeV}/c^2$.

In summary, a search for the $\Sigma^+ \to p\mu^+\mu^-$ rare decay is performed by the LHCb experiment at centre-of-mass energies $\sqrt{s} = 7$ and $8\,\mathrm{TeV}$, corresponding to an integrated luminosity of $3\,\mathrm{fb}^{-1}$. Evidence for the $\Sigma^+ \to p\mu^+\mu^-$ decay is found in the full dataset, albeit not found when requiring the events to be triggered independently of the signal decay. The observed signal candidates show a dimuon invariant mass distribution consistent with phase space; no significant peak consistent with an intermediate particle is found in the dimuon invariant mass distribution. The upper limit on the branching fraction of the $\Sigma^+ \to p\mu^+\mu^-$ decay is 6.3×10^{-8} at 95% CL for a SM-like signal.

References

[1] He X G, Tandean J and Valencia G 2005 *Phys.Rev.* **D72** 074003
[2] Park H *et al.* (HyperCP collaboration) 2005 *Phys.Rev.Lett.* **94** 021801
[3] He X G, Tandean J and Valencia G 2007 *Phys.Rev.Lett.* **98** 081802
[4] He X G, Tandean J and Valencia G 2006 *Phys. Rev.* **D74** 115015
[5] Gorbunov D and Rubakov V 2006 *Phys.Rev.* **D73** 035002
[6] Demidov S V and Gorbunov D S 2007 *JETP Lett.* **84** 479–484
[7] He X G, Tandean J and Valencia G 2005 *Phys.Lett.* **B631** 100–108
[8] Geng C Q and Hsiao Y K 2006 *Phys. Lett.* **B632** 215–218
[9] Deshpande N, Eilam G and Jiang J 2006 *Phys.Lett.* **B632** 212–214
[10] Chen C H, Geng C Q and Kao C W 2008 *Phys.Lett.* **B663** 400–404
[11] Xiangdong G, Li C S, Li Z and Zhang H 2008 *Eur.Phys.J.* **C55** 317–324
[12] Mangano M L and Nason P 2007 *Mod.Phys.Lett.* **A22** 1373–1380
[13] Pospelov M 2009 *Phys. Rev.* **D80** 095002
[14] Love W *et al.* (CLEO collaboration) 2008 *Phys. Rev. Lett.* **101** 151802
[15] Tung Y C *et al.* (E391a collaboration) 2009 *Phys. Rev. Lett.* **102** 051802
[16] Abazov V M *et al.* (D0 collaboration) 2009 *Phys. Rev. Lett.* **103** 061801
[17] Aubert B *et al.* (BaBar collaboration) 2009 *Phys. Rev. Lett.* **103** 081803
[18] Hyun H J *et al.* (Belle collaboration) 2010 *Phys. Rev. Lett.* **105** 091801
[19] Abouzaid E *et al.* (KTeV collaboration) 2011 *Phys. Rev. Lett.* **107** 201803
[20] Ablikim M *et al.* (BESIII collaboration) 2012 *Phys. Rev.* **D85** 092012
[21] Lees J P *et al.* (BaBar collaboration) 2014 *Phys. Rev. Lett.* **113** 201801
[22] Aaij R *et al.* (LHCb collaboration) 2013 *Phys. Rev. Lett.* **110** 211801
[23] Aaij R *et al.* (LHCb collaboration) 2015 *Phys. Rev. Lett.* **115** 161802
[24] Park H K 2010 *JHEP* **10** 052
[25] Alves Jr A A *et al.* (LHCb Collaboration) 2016 Evidence for the rare decay $\Sigma^+ \to p\mu^+\mu^-$ LHCb-CONF-2016-013 URL https://cds.cern.ch/record/2224468

[26] Alves Jr A A *et al.* (LHCb collaboration) 2008 *JINST* **3** S08005
[27] Aaij R *et al.* (LHCb collaboration) 2015 *Int. J. Mod. Phys.* **A30** 1530022
[28] Aaij R *et al.* 2013 *JINST* **8** P04022
[29] Breiman L, Friedman J H, Olshen R A and Stone C J 1984 *Classification and regression trees* (Belmont, California, USA: Wadsworth international group)
[30] Schapire R E and Freund Y 1997 *J. Comput. Syst. Sci.* **55** 119
[31] Olive K A *et al.* (Particle Data Group) 2014 *Chin. Phys.* **C38** 090001 and 2015 update
[32] Punzi G 2003 *Statistical Problems in Particle Physics, Astrophysics, and Cosmology* ed Lyons L, Mount R and Reitmeyer R p 79 (*Preprint* physics/0308063)
[33] Martinez Santos D and Dupertuis F 2014 *Nucl. Instrum. Meth.* **A764** 150–155
[34] Albrecht H *et al.* (ARGUS collaboration) 1990 *Phys. Lett.* **B241** 278–282
[35] Read A L 2002 *J. Phys.* **G28** 2693–2704