December 6, 2013

The search for $D^0 \to e^\pm \mu^\mp$

THOMAS BIRD

ON BEHALF OF THE LHCB COLLABORATION

School of Physics and Astronomy
The University of Manchester, Manchester, UK

An overview is presented of a method to search for $D^0 \to e^\pm \mu^\mp$ with LHCb data. In order to reduce combinatorial backgrounds, tagged D^0 candidates from the decay $D^{*+} \to D^0 \pi^+$ are used. This measurement is performed with respect to $\mathcal{B}\left(D^0 \to \pi^+ \pi^-\right)$, which cancels uncertainties in the luminosity and D^{*+} production cross-section. It is estimated that using $3\,\mathrm{fb}^{-1}$ of LHCb data an upper limit can be attained of $\mathcal{O}\left(10^{-7}\right)$ at a 90% confidence level.

PRESENTED AT

The 6th International Workshop on Charm Physics
(CHARM 2013)
Manchester, UK, 31 August – 4 September, 2013

1 Overview

The lepton flavour violating decay $D^0 \to e^\pm \mu^\mp$ * is forbidden in the Standard Model (SM) and so its observation would be a clear sign of new physics. This decay is predicted to occur by several different SM extensions, with predicted rates varying by eight orders of magnitude. An R-parity violating minimal supersymmetric SM predicts $\mathcal{B}(D^0 \to e^\pm \mu^\mp) < 1.0 \times 10^{-6}$ [1], while theories with multiple Higgs doublets predict rates 4 orders of magnitude lower, $\mathcal{B}(D^0 \to e^\pm \mu^\mp) \sim 7 \times 10^{-10}$ [1]. The lowest branching fraction predictions are given by SM extensions with extra fermions, which predict $\mathcal{B}(D^0 \to e^\pm \mu^\mp) < 1.0 \times 10^{-14}$ [1].

In 2011 and 2012 the LHCb experiment collected $3\,\text{fb}^{-1}$ of pp collisions produced by the LHC at center-of-mass energies $\sqrt{s} = 7\,\text{TeV}$ and $8\,\text{TeV}$. The analysis procedure is presented to measure $\mathcal{B}(D^0 \to e^\pm \mu^\mp)$ with the $3\,\text{fb}^{-1}$ dataset. In order to reduce the combinatorial background, tagged D^0 candidates are used from the decay $D^{*+} \to D^0 \pi^+$. The $\mathcal{B}(D^0 \to e^\pm \mu^\mp)$ measurement will be performed with respect to $\mathcal{B}(D^0 \to \pi^+ \pi^-)$, in order to cancel uncertainties in the luminosity and D^{*+} production cross-section.

Candidate $D^0 \to e^\pm \mu^\mp$ decays are selected in bins of Boosted Decision Tree (BDT) output; this helps to separate signal from background events. The number of signal and background events are fitted in three BDT bins and combined with the PDG value of $\mathcal{B}(D^0 \to \pi^+ \pi^-)$, which is currently $(1.402 \pm 0.026) \times 10^{-3}$ [2], to produce a limit on $\mathcal{B}(D^0 \to e^\pm \mu^\mp)$.

2 Selection efficiency

Before the total number of $D^0 \to e^\pm \mu^\mp$ or $D^0 \to \pi^+ \pi^-$ events can be measured the selection efficiency must be measured; this is performed on simulated events. PYTHIA 6.4 [3] is used for the simulation, with a specific LHCb configuration [4]. Decays of hadronic particles are described by EVTGEN [5], in which final-state radiation is generated using PHOTOS [6]. Particle interactions with the detector and the detector response are implemented using the GEANT4 toolkit [7, 8] as described in Ref. [9]. Identical reconstruction and selection algorithms can be applied to both the data and simulation.

The selection efficiency, ε_{sel}, can be broken down into several parts,

$$\varepsilon_{\text{sel}} = \varepsilon_{\text{gen}} \; \varepsilon_{\text{off|gen}} \; \varepsilon_{\text{trig|off}},$$

where ε_{gen} is the fraction of generated events that pass requirements on the muon and electron momenta during the simulation (these requirements are not applied to the $D^0 \to \pi^+ \pi^-$ sample); $\varepsilon_{\text{off|gen}}$ is the fraction of generated events that pass the offline

---

*In this paper, charge conjugate modes are always implied.

Table 1: Generation, offline and trigger efficiencies measured on simulated LHCb data for the decays $D^0 \to e^\pm \mu^\mp$ and $D^0 \to \pi^+\pi^-$.

Decay	Generation efficiency, ε_{gen}	Offline efficiency, $\varepsilon_{\text{off}|\text{gen}}$	Trigger efficiency, $\varepsilon_{\text{trg}|\text{off}}$
$D^0 \to e^\pm \mu^\mp$	14%	1.6%	19%
$D^0 \to \pi^+\pi^-$	100%	1.8%	20%

selection and $\varepsilon_{\text{trig}|\text{off}}$ is the fraction of simulated events that are triggered given they passed the offline selection. The measured values of these efficiencies are given in Table 1.

3 Signal-background separation

To improve separation between signal and background events, a BDT is trained using the TMVA package [10]. The BDT is trained on several kinematic and geometrical variables using data from the $e\mu$-combination mass sidebands as a background sample and the simulated signal dataset as the signal sample; its separation power can be seen in Fig. 1, where the BDT output distribution is shown for both signal and background samples. The BDT output distribution from the signal sample is divided into three bins (< 0, 0–0.08 and > 0.08) with an approximately equal number of events in each. In each bin a fit is performed to two variables, the mass of the $e\mu$-combination and the mass difference between the $e\mu\pi$-combination and the $e\mu$-combination. When fitting data, the second variable helps to separate out combinatorial background events, where a real π^+ is randomly combined with a real D^0 to form a fake D^{*+}. The plots
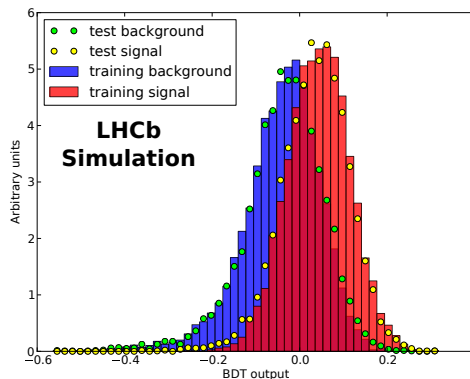

Figure 1: BDT output for background (blue) and signal (red) events. Events to the left are more likely background and events to the right are more likely signal. Independent test samples of signal and background events were used to check the BDT was not overtrained; the output of a signal (background) test sample can be seen as yellow (green) circles.

in Fig. 2 show the mass distribution of the $e\mu$-combination in simulated $D^0 \to e^{\pm}\mu^{\mp}$ events for each bin of the BDT output.
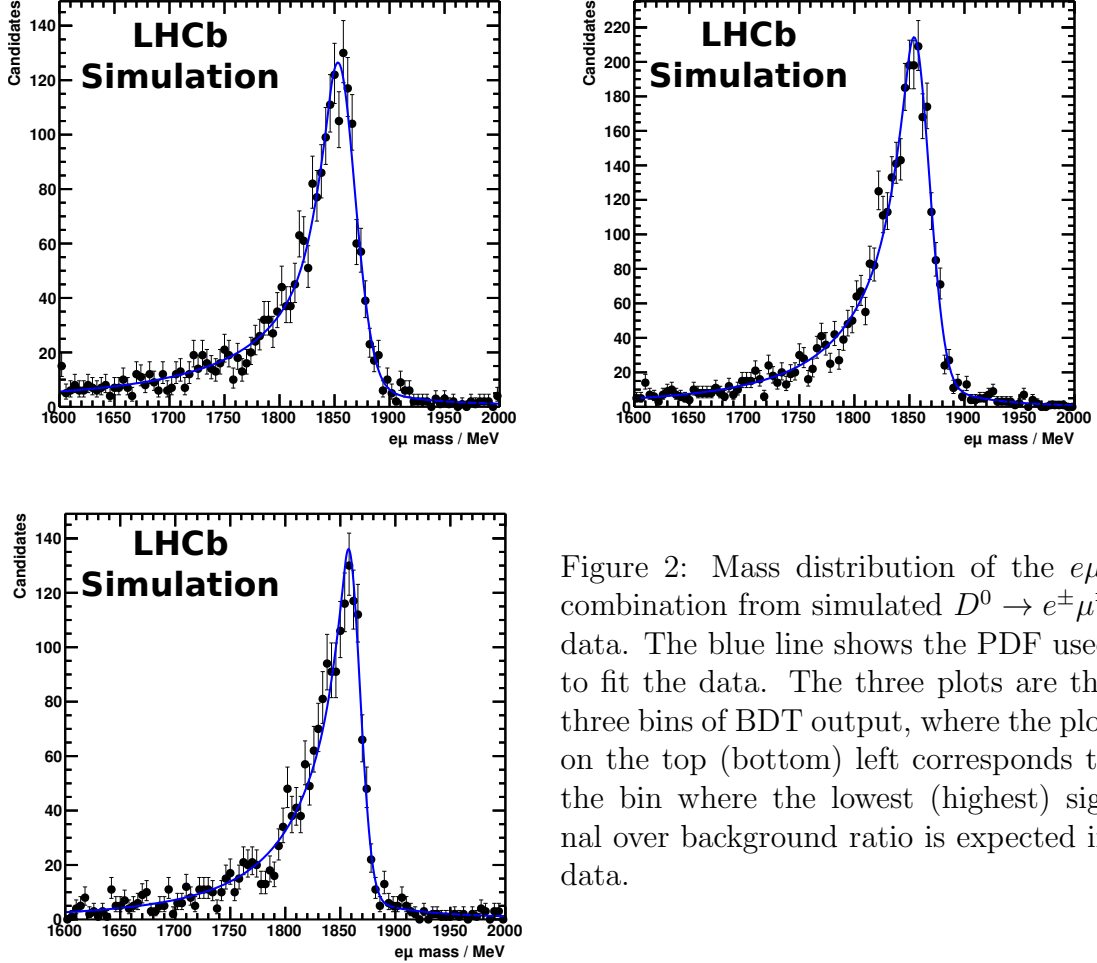

Figure 2: Mass distribution of the $e\mu$-combination from simulated $D^0 \to e^{\pm}\mu^{\mp}$ data. The blue line shows the PDF used to fit the data. The three plots are the three bins of BDT output, where the plot on the top (bottom) left corresponds to the bin where the lowest (highest) signal over background ratio is expected in data.

4   Estimation of expected limit

In order to estimate the possible limit that could be achieved on $\mathcal{B}\left(D^0 \to e^{\pm}\mu^{\mp}\right)$ a toy study is performed. In this toy study the $D^0 \to \pi^+\pi^-$ events are simultaneously fitted along with $D^0 \to e^{\pm}\mu^{\mp}$ events split into the three BDT bins. The number of expected signal $D^0 \to \pi^+\pi^-$ events after selection is approximated using the known D^{*+} cross-section, branching ratios, integrated luminosity and measured efficiencies; it is calculated approximately 51000 $D^0 \to \pi^+\pi^-$ events will be selected in $3\,\text{fb}^{-1}$ of LHCb data. No signal $D^0 \to e^{\pm}\mu^{\mp}$ events are simulated in this study, but they are fitted for nevertheless to evaluate the expected sensitivity in data. The signal $D^0 \to e^{\pm}\mu^{\mp}$ shape is taken and fixed from the fit to the simulated data, while for
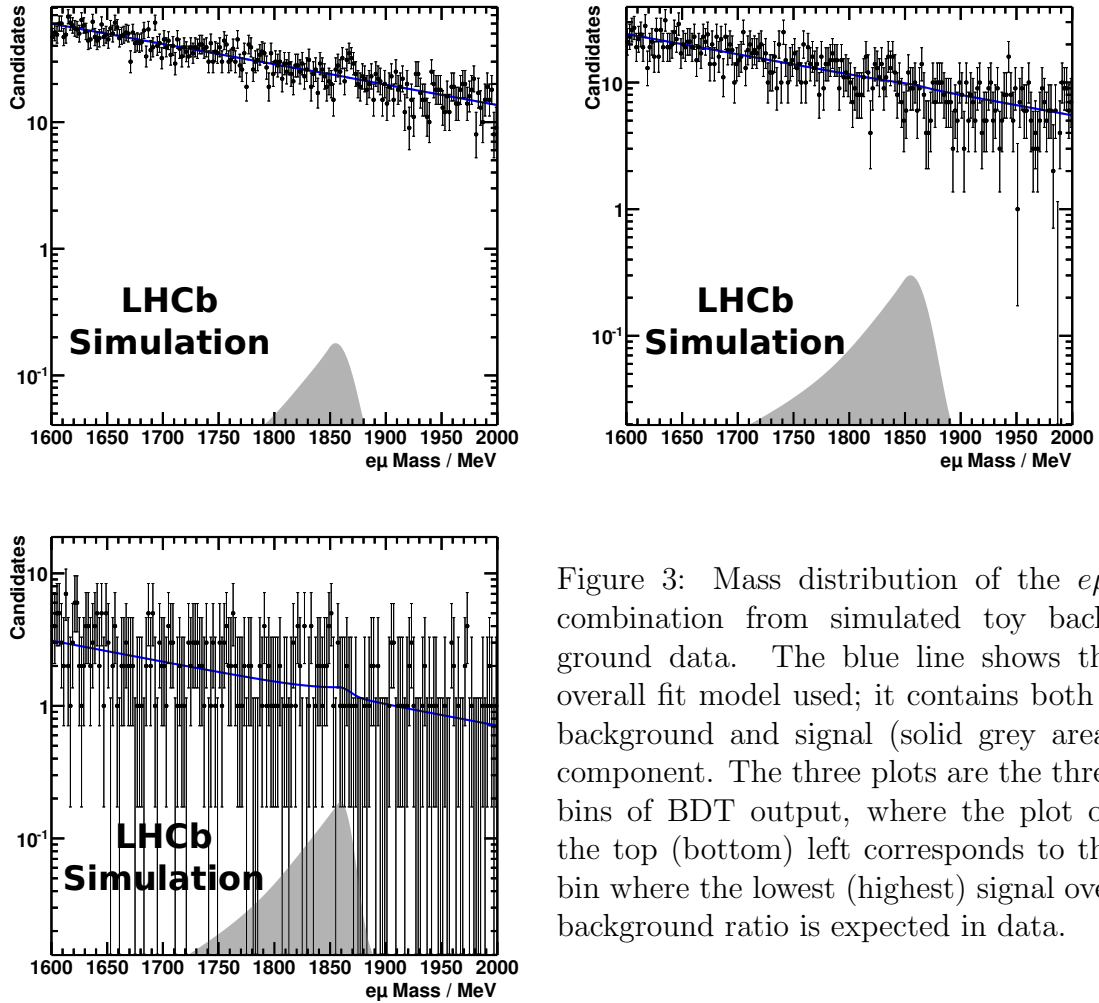
3

Figure 3: Mass distribution of the $e\mu$-combination from simulated toy background data. The blue line shows the overall fit model used; it contains both a background and signal (solid grey area) component. The three plots are the three bins of BDT output, where the plot on the top (bottom) left corresponds to the bin where the lowest (highest) signal over background ratio is expected in data.

the background shape various assumptions are made about the number and distribution of events. Figure 3 shows the $e\mu$-combination mass distribution in the three bins of BDT output. The toy study is able to set an upper limit on the branching fraction $\mathcal{B}\left(D^0 \to e^{\pm}\mu^{\mp}\right) < \mathcal{O}\left(10^{-7}\right)$ at a 90% confidence level. The upper limit is calculated by taking the ratio of the signal-plus-background and the background-only profile likelihoods and assumes Wilks' theorem [11] to calculate a p-value and thus the confidence interval.

Belle measured the most stringent limit on the branching fraction $\mathcal{B}\left(D^0 \to e^{\pm}\mu^{\mp}\right)$, achieving an upper limit of 2.6×10^{-7} at a 90% confidence level [12]. It is likely LHCb will set a comparable limit with the $3\,\text{fb}^{-1}$ dataset.

References

[1] G. Burdman, E. Golowich, J. L. Hewett and S. Pakvasa, Phys. Rev. D **66** (2002) 014009.

[2] J. Beringer *et al.* [Particle Data Group Collaboration], Phys. Rev. D **86** (2012) 010001.

[3] T. Sjostrand, S. Mrenna and P. Z. Skands, JHEP **0605** (2006) 026.

[4] I. Belyaev *et al.* [LHCb Collaboration], IEEE Nucl. Sci. Symp. Conf. Rec. **2010** (2010) 1155.

[5] D. J. Lange, Nucl. Instrum. Meth. A **462** (2001) 152.

[6] P. Golonka and Z. Was, Eur. Phys. J. C **45** (2006) 97.

[7] J. Allison, K. Amako, J. Apostolakis, H. Araujo, P. A. Dubois, M. Asai, G. Barrand and R. Capra *et al.*, IEEE Trans. Nucl. Sci. **53** (2006) 270.

[8] S. Agostinelli *et al.* [GEANT4 Collaboration], Nucl. Instrum. Meth. A **506** (2003) 250.

[9] M. Nicol, [LHCb Collaboration], EPJ Web Conf. **28** (2012) 12028.

[10] A. Hocker, J. Stelzer, F. Tegenfeldt, H. Voss, K. Voss, A. Christov, S. Henrot-Versille and M. Jachowski *et al.*, PoS ACAT (2007) 040.

[11] S. S. Wilks, Annals Math. Statist. **9** (1938) 1, 60.

[12] M. Petric *et al.* [Belle Collaboration], Phys. Rev. D **81** (2010) 091102.