PID strategy and performance at LHCb in Run 2

Carla Marin Benito[*][†]

LAL, Univ. Paris-Sud, CNRS/IN2P3, Université Paris-Saclay, Orsay, France

E-mail: carla.marin.benito@cern.ch

The LHCb particle identification (PID) system is composed of two ring-imaging Cherenkov detectors, a series of muon chambers and a calorimeter system. A novel strategy has been introduced in Run 2, where the selection of PID calibration samples for charged particles and neutrals is implemented in the LHCb software trigger. A further processing of the data is required in order to provide samples for the determination of PID performance, which is achieved through a centralised production that makes highly efficient use of computing resources. This talk covers the major steps of the implementation, and highlights the PID performance achieved in Run 2.

The 39th International Conference on High Energy Physics (ICHEP2018)
4-11 July, 2018
Seoul, Korea

---

[*]Speaker.

[†]on behalf of the LHCb collaboration.

1. Particle Identification at LHCb

LHCb is a dedicated heavy-flavour physics experiment at the LHC. Its core program is focused on the study of exclusive final states from the decays of *b*- and *c*-hadrons. In order to achieve world-leading results in this area, a high PID performance is crucial to disentangle different final states. The LHCb detector is a single-arm forward spectrometer composed of various sub-systems. The PID system is formed by two ring-imaging Cherenkov detectors (RICH), a series of muon chambers and a calorimeter system. A detailed description of the detector can be found in Ref. [1].

The PID strategy at LHCb is based on the combination of the information from the various PID sub-systems into high level variables, providing excellent separation power for the different particle species. Two types of variables are used for charged particles: the $DLL_{X\pi}$ are the combined likelihood difference between the $X \in \{K, p, \mu, e\}$ and π hypothesis, charged pions being the most abundant species at LHCb; the ProbNNX are the output of neural networks (NN), exploiting also tracking information, trained to separate species X from any other. Dedicated tools based on NNs are built from the calorimeter information to distinguish neutral species. In particular, the isNotE and isNotH variables are trained to separate photons from electrons and hadrons, respectively. The performance of these tools is measured by means of data-driven calibration techniques. Details about the latest training and calibration methods can be found in Ref. [2].

2. Computing strategy in Run 2

In Run 2 LHCb introduced a new trigger strategy based on the complete separation of the two software levels that allows a full alignment and calibration of the detector to be performed in between, providing offline quality reconstruction in the trigger [3]. As part of this model, a dedicated calibration stream, namely Turbo Calibration (TurboCalib), was introduced. It permits the selection of large calibration samples as required by high precision analyses together with the storage of the full raw event information as needed to exploit offline re-calibrations and new tunings of the PID variables [2]. The charged PID calibration samples are selected as part of the TurboCalib stream since the beginning of 2015, profiting from all of its benefits and contributing to its full development. Other species have been added more recently, as detailed in Sec. 3.

The large PID calibration samples selected in the TurboCalib stream are processed centrally and provided to analysts in the format of ROOT files including the necessary information to extract PID efficiencies for each analysis. This poses a better resource optimisation with respect to the do-it-yourself approach and provides control on the systematic uncertainties by using a common approach. Several steps are followed in the central processing of the data: first the offline and trigger candidates are matched to be able to properly evaluate efficiencies when PID requirements have been imposed at both stages; then an invariant mass fit to the candidates is performed and finally background subtracted samples are obtained using sWeights [4]. This information is provided to analysts to evaluate their particular PID selection. Moreover, the full raw event information is stored separately and can be used for more detailed PID studies, e.g., development of new variables.

3. PID calibration samples

Most LHCb analyses involve decays with charged particles in the final state, namely electrons,

Species	Low momentum	High momentum
γ	$D^+ \to [\rho^0 \gamma]_{\eta'} \pi^+$	$B^0 \to K^{*0} \gamma$
	$D_s^{*+} \to D_s^+ \gamma$	
	$\eta \to \mu^+ \mu^- \gamma$	
π^0	$D^0 \to K^+ \pi^- \pi^0$ (resolved)	$D^0 \to K^+ \pi^- \pi^0$ (merged)

Table 1: Decay modes used to obtain calibration samples of neutral species. Various channels covering different phase-space regions are used.

muons, pions, kaons and protons. High rate and very pure modes involving these species, which can be selected without the use of PID information, are used as calibration samples. They are included in TurboCalib since the beginning of Run 2 and documented in Ref. [2].

With the expansion of the LHCb physics program more analyses involve long-lived particles, such as Λ and K_S^0, and neutral objects. The first decay mostly outside the LHCb vertex locator (VELO) and might not have information from the first RICH detector. The second require specific calorimeter information for their identification. Consequently, dedicated calibration samples are selected in TurboCalib since the beginning of the 2017 data-taking. The decays $K_S^0 \to \pi^+ \pi^-$, $\Lambda \to p\pi^-$ and $\Omega \to \Lambda K^+$, with the parent decaying outside the VELO, are exploited to obtain pure samples of these types of pions, protons and kaons, respectively. Various decays involving photons and neutral pions are exploited to cover the full phase-space used in analyses, as listed in Tab. 1. Examples of invariant mass fits to 2017 data candidates, from which the background subtracted samples are obtained, are shown in Fig. 1.
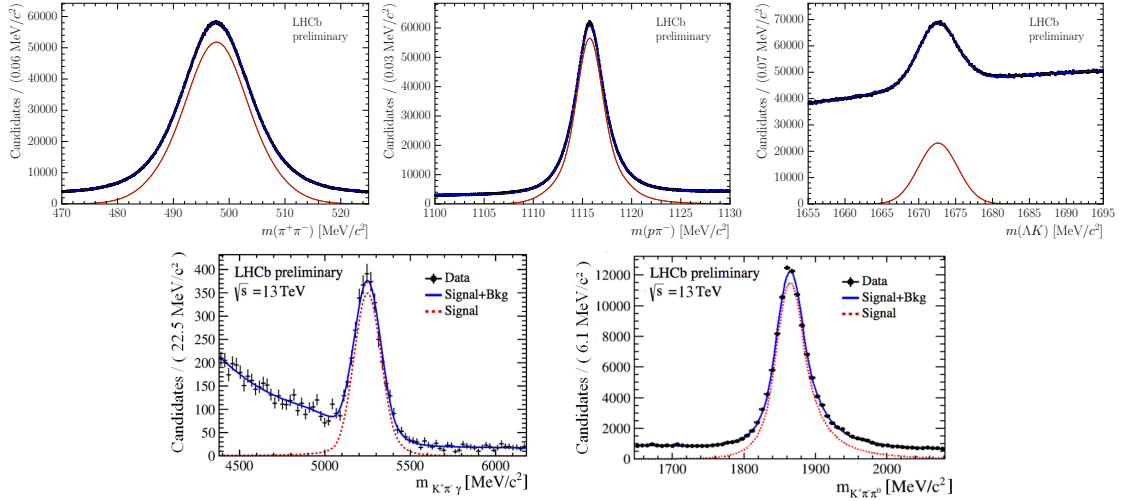

Figure 1: Invariant mass fits of (top-left to bottom-right) $K_S^0 \to \pi^+ \pi^-$, $\Lambda \to p\pi^-$ and $\Omega \to \Lambda K^+$ candidates reconstructed from tracks without VELO hits and $B^0 \to K^{*0} \gamma$ and $D^0 \to K^+ \pi^- \pi^0$ candidates. The data is represented by black dots while the blue and red curves describe the total and signal distributions, respectively.

4. PID performance

The performance of the charged PID high level variables is evaluated by comparing the efficiency and mis-identification probability of a given species, extracted from the calibration data, as a function of momentum. Fig. 2 shows the performance plots for the different charged species in the two top rows. An excellent separation is obtained in all cases in a wide momentum range. The performance of the neutral PID variables is presented in the bottom row, where the response of the isNotE (isNotH) NN is shown for signal photons and background electrons (neutral pions) in the left (right) plot. A good separation is achieved in both cases.
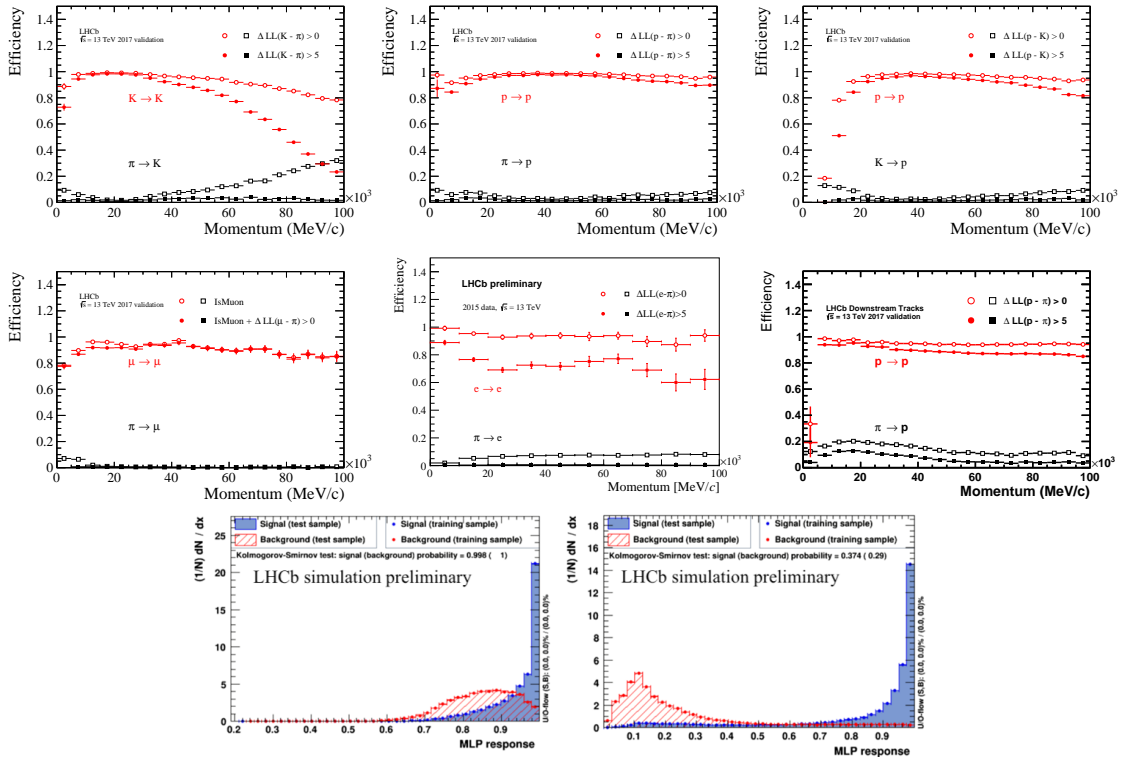

Figure 2: Top: efficiency and mis-identification probability for the charged PID high level variables as a function of momentum, for different species. Bottom: response of the isNotE (left) and isNotH (right) classifiers for simulated signal photons (blue) and background electrons and neutral pions (red), respectively.

References

[1] A. A. Alves Jr. et al. The LHCb detector at the LHC. *JINST*, 3:S08005, 2008.

[2] R. Aaij et al. Selection and processing of calibration samples to measure the particle identification performance of the LHCb experiment in Run 2. Technical Report LHCb-DP-2018-001, Mar 2018.

[3] R. Aaij et al. Tesla: an application for real-time data analysis in High Energy Physics. *Comput. Phys. Commun.*, 208:35–42, 2016.

[4] Muriel Pivk and Francois R. Le Diberder. sPlot: A statistical tool to unfold data distributions. *Nucl. Instrum. Meth.*, A555:356–369, 2005.