LHCb-PUB-2014-038
May 21, 2014

PID Performance in the LHCb High Level Trigger

S. Benson[1], O. Lupton[2]

[1] *CERN, Geneva*
[2] *University of Oxford, United Kingdom*

Abstract

The LHCb High Level Trigger (HLT) incorporates RICH particle identification (PID) calculations that are designed to be faster than those used for the offline reconstruction, which is necessary due to the timing constraints imposed on the HLT. The performance of this faster calculation has been evaluated using $B^+ \to \overline{D}^0\pi^+$ events recorded during 2011 data taking and compared against the performance of the offline calculation. The time taken to calculate PID information in the HLT has been measured both for the configuration used in Run 1, and for configurations approaching that used offline. Similar overall efficiency and misidentification performance is seen in the configurations used offline and in the HLT, though the raw values of PID variables can differ substantially for a given candidate.

LHCb-PUB-2014-038
21/05/2014

Contents

1 Introduction

The LHCb detector [1] uses information from Ring Imaging Cherenkov (RICH) detectors extensively in physics analyses in the form of combined differences in the likelihoods between kaon and pion mass hypotheses ($\mathrm{DLL}_{K\pi}$). In the calculation of the particle identification (PID) variables, information about tracks is essential in matching Cherenkov photons to tracks [2]. The use of RICH PID information in the LHCb High Level Trigger (HLT) has not been so extensive at the time of writing. The reconstruction provided in the HLT, while being of high quality, is built with time constraints in mind and is therefore necessarily different from that applied offline.

The HLT operates as a software trigger that, in Run 1, was split into two levels, which were executed sequentially. The first level (HLT1) is designed as a single track and dimuon trigger, whereas the second level (HLT2) performs full reconstruction of heavy flavour decays [3, 4]. In future data taking periods, the HLT2 stage will be performed in a more separated fashion, with greater time allowed between the two. The extra time afforded by the so-called HLT splitting also allows for PID information to be calculated and used much more widely [5].

The additional time allotted to the HLT allows for a reconstruction much closer to that performed offline [5]. To achieve this, the split HLT will require offline quality calibrations to be included in the HLT reconstruction. This will permit the calculation of offline quality PID variables, in addition to allowing time for PID information to be calculated for more events. The offline quality calibrations required include a regularly updated alignment and the RICH image and refractive index calibrations. These require around $20,000$ $D^0 \to K^-\pi^+$ events and 500 Hz output rate respectively. At the time of writing, all calibrations are either completely automated or close to complete automation.

Before the PID information can be used more widely in the HLT, confidence is required that the PID online is performing in a similar manner as used in the offline reconstruction, and differences observed must be understood. The rest of this Note describes the data driven evaluation of the online PID performance, the comparison with offline PID performance, and the time taken to calculate PID information. Performance is evaluated using $B^+ \to \overline{D}^0\pi^+$ events [1]. In Section 2, the dataset used to determine the PID performance is defined. The strategy used to fit the $B^+ \to \overline{D}^0\pi^+$ events is given in Section 3. The performance of the $\mathrm{DLL}_{K\pi}$ variable as calculated in the HLT is described in Section 4, and compared against the offline value in Section 5. The $\mathrm{DLL}_{K\pi}$ agreement at the track level is compared in Section 6. Details of the time taken to calculate PID information are given in Section 7.

Studies presented in this document are intended as a baseline for future use of PID information in the HLT. Using such information will become increasingly necessary in the conditions of the LHCb Upgrade, where constraints on the HLT output rate will impact the HLT efficiency unless PID information can be used.

[1] Charge conjugation is assumed unless otherwise stated.

Variable	Requirement
B^+ vertex separation χ^2	> 36
B^+ $\cos\theta_{\rm DIRA}$	> 0.99989
$\lvert m_{B^+} - 5279.25\,{\rm MeV}/c^2\rvert$	$< 100.0\,{\rm MeV}/c^2$
$\lvert m_{D^0} - 1869.62\,{\rm MeV}/c^2\rvert$	$< 50.0\,{\rm MeV}/c^2$
B^+, $\overline{D}^0$ vertex χ^2/NDF	< 10
K,π $p_{\rm T}$	$> 100\,{\rm MeV}/c$
K,π track χ^2/NDF	< 4
K,π p	$> 1\,{\rm GeV}/c$
K,π DOCA	$< 0.5\,{\rm mm}$
K,π sum $p_{\rm T}$	$> 1.8\,{\rm GeV}/c$
Bachelor IP χ^2	> 21
BDT response	> 0.05

Table 1: Selection requirements imposed on triggered $B^+ \to \overline{D}^0\pi^+$ candidates, where $\theta_{\rm DIRA}$ is the angle between the B^+ momentum vector and the vector defined by the production vertex and decay vertex, and DOCA is the distance of closest approach to the nearest PV. The impact parameter (IP) χ^2 requirement ensures the bachelor pion is inconsistent with originating from the primary pp interaction vertex.

2 Dataset

The data driven method of determining the PID efficiencies of pions and kaons uses $B^+ \to \overline{D}^0\pi^+$ events. Candidates are required to have fired the generic HLT1 single track trigger and the 2 and 3 body topological HLT2 triggers. In the hardware trigger, the decision is required to either be independent of the selected candidate, or be due to the candidate firing the hadron trigger.

Triggered events are then subjected to a loose pre-selection on the particle kinematics, vertex and track quality, and displacement from the primary interaction. A classifier based on a boosted decision tree (BDT) is applied, which is trained on simulated events. The BDT uses the B^+ transverse momentum, $p_{\rm T}$, the B^+ hadron vertex separation χ^2 from the nearest primary vertex (PV), and the sum of the B^+ and $\overline{D}^0$ vertex χ^2 divided by the sum of the numbers of degrees of freedom. The chosen requirement on the BDT response is found to be greater than 99 % efficient. The selections are summarised in Table 1. For simplicity, the Cabibbo favoured decay $D^0 \to K^-\pi^+$ is chosen.

Run range	Moore version	TCK	CondDB tag	DDDB tag
101762-104486	v12r9p3	0x00790038	head-20110901	head-20110722
101122-101761	v12r9p3	0x00790037	head-20110901	head-20110722
101092-101121	v12r8g1	0x00760037	head-20110722	head-20110722
101012-101067	v12r8g1	0x00760037	head-20110622	head-20110302
95929-101011	v12r8g1	0x00760037	head-20110622	head-20110302
94013-94386	v12r6p4	0x00730035	head-20110622-Reco10	head-20110302
92929-94012	v12r6p4	0x00730035	head-20110524	head-20110302
92838-92906	v12r6p4	0x006D0032	head-20110524	head-20110302
92821-92826	v12r6p4	0x00710035	head-20110524	head-20110302
92317-92735	v12r6p4	0x006D0032	head-20110524	head-20110302
91631-92316	v12r6p4	0x006D0032	head-20110512	head-20110302

Table 2: Versions of Moore and database tags used as a function of run number.

2.1 PID Performance in 2011 Data Taking

To evaluate the performance of the trigger as it was run during 2011 data taking, the dataset is processed using the same versions and configurations of the trigger application, Moore, as were used during data taking. Small modifications are made such that the PID values calculated online are saved for later analysis. Table 2 shows the various configurations used during 2011 data taking. After the trigger is re-evaluated on the data, the online tracks are matched to the tracks of the offline candidates. To be considered a match, at least 70% of the tracker hits must be shared between the online and offline tracks. In all results shown, it has been required that all tracks possess both offline and online PID information. This requirement has the effect of imposing the cuts used in the RICH reconstruction on all candidates.

2.2 Configurations Approaching Offline PID Performance

Four cumulative iterations have been considered to approach the PID calculation used in the offline reconstruction. The first of these is the configuration used during Run 1.

2.2.1 Scenario 2

The modifications to the PID calculation in Scenario 2 consisted of:

- Increasing the number of mass hypotheses considered from the two used in the HLT (π or K) to the six used offline (π, K, μ, p, e or "belowThreshold").
- Changing the likelihood calculation from the "FastGlobal" setting to the "FullGlobal" setting used offline.
- Removal of RICH selection requirements (based on the transverse momentum, momentum and track fit quality), such that all forward reconstructed tracks were considered in the PID calculation.
- The context of the RICH calculation changed to "Offline" from "HLT". This had the effect of using all offline calibrations and ensuring all settings used for the PID variable calculation were the same as those used offline.
- The utilisation of the latest available database tags.

2.2.2 Scenario 3

The modifications to the PID calculation in Scenario 3 consisted of the modifications listed in Scenario 2 plus:

- Removal of the HLT forward reconstruction selection requirements, such that all HLT tracks were considered for the calculation of the PID variables.

2.2.3 Scenario 4

The modifications to the PID calculation in Scenario 4 consisted of the modifications listed in Scenario 3 plus:

- Inclusion of the Aerogel radiator in the calculation of the PID variables.

3 Fitting Strategy

In order to determine the PID performance it is necessary to first statistically subtract the background events from the $B^+ \to \overline{D}^0\pi^+$ dataset. We use the ${}_s\mathcal{P}lot$ method [6], with the B^+ and D^0 masses as control variables, to generate event-by-event sWeights which perform this subtraction.

There are three components included in the fit to the $B^+ \to (\overline{D}^0 \to K^+\pi^-)\pi^+$ candidates: the signal component, a partially reconstructed background containing a real D^0, and a combinatoric component. The presence of the partially reconstructed background requires a two-dimensional fit to the $KK\pi$ and $K\pi$ invariant masses.

3.1 Signal Model

The signal model consisted of a StudentT function [7] modelling the B^+ mass distribution:

$$F_{\rm ST}(m) = \frac{1}{\sqrt{\pi n s^2}}\left(1+\frac{(m-\mu)^2}{ns^2}\right)^{-0.5(n+1)}\left(\frac{\Gamma[0.5n]}{\Gamma[0.5(n+1)]}\right), \quad (1)$$

where Γ is the Euler Gamma function, and n, s, and μ are parameters to be fitted. The value of the tail parameter, n, was fixed to 4.56 from simulated events modelling the 2011 data conditions. A sum of two Gaussian functions was used to model the D^0 mass distribution.

3.2 Background Model

In order to determine the shape of the partially reconstructed background, simulated events containing a $B \to DX$ cocktail modelling 2010 data conditions were used. An ARGUS [8] function convolved with a Gaussian resolution function, with width $\sigma_{\rm ARG}$, was used to describe the $K\pi\pi$ invariant mass distribution, together with a sum of two Gaussian functions describing the $K\pi$ invariant mass distribution. The ARGUS function, $F_{\rm ARG}$, is of the form:

$$F_{\rm ARG}(m) = m\left(1-\left(\frac{m}{\mu}\right)^2\right)^p \cdot \exp\left(c\left(1-\left(\frac{m}{\mu}\right)^2\right)\right), \quad (2)$$

where $\mu_{\rm ARG}$, c and p are parameters to be fitted. The $K\pi\pi$ and $K\pi$ invariant mass distributions and fit curves for the simulated partially reconstructed events are shown in Figure 1, and the extracted ARGUS parameters are listed in Table 3. These parameters are then fixed in the fit to 2011 data.

3.3 2011 Data Results

In the fit to 2011 data, $B^+ \to \overline{D}^0\pi^+$ and $B^- \to D^0\pi^-$ events have been fitted separately to obtain the sWeights and have then been merged for the analysis of the PID performance.

Parameter	Fit result
p	4.3 ± 2.8
c	-80.6 ± 38.8
$\mu_{\rm ARG}$ [MeV/c^2]	5368 ± 31
$\sigma_{\rm ARG}$ [MeV/c^2]	0.26 ± 1.03

Table 3: Fitted parameters from the simulated $B \to DX$ cocktail.

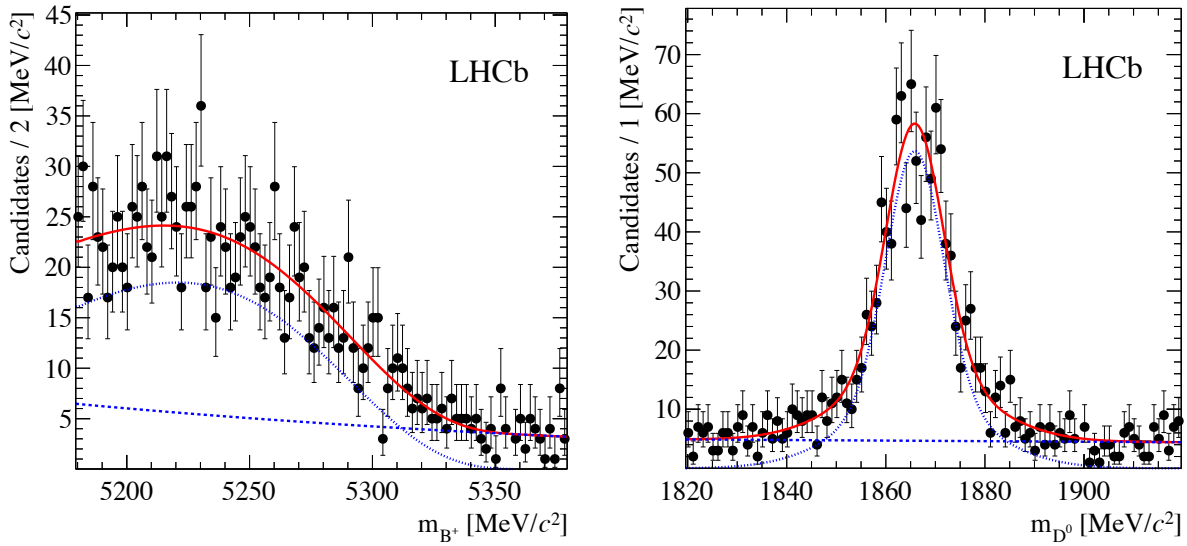

Figure 1: Projections on to the $K\pi\pi$ (left) and $K\pi$ (right) invariant masses from the fit to the simulated $B \to DX$ cocktail. The total PDF is given by the solid red line, the dashed blue line represents the combinatoric component and the dotted blue line represents the partially reconstructed background.

It should be noted that the D^0 parameters are shared between the partially reconstructed background and the $B^+ \to \overline{D}{}^0\pi^+$ signal. The projections on to the $K\pi\pi$ and $K\pi$ invariant masses are shown in Figure 2, with the extracted parameters and yields listed in Table 4.

Parameter	Fit result (B^+)	Fit result (B^-)
N_{sig}	38686 ± 251	37744 ± 247
N_{com}	12115 ± 169	12115 ± 167
N_{MR}	5701 ± 169	5299 ± 165
m_{B^+} [MeV/c^2]	5284.2 ± 0.1	5283.9 ± 0.1
σ_{B^+} [MeV/c^2]	18.3 ± 0.1	18.5 ± 0.1
m_{D^0} [MeV/c^2]	1866.05 ± 0.05	1866.03 ± 0.05
$\sigma^1_{D^0}$ [MeV/c^2]	10.8 ± 0.5	10.5 ± 0.5
$\sigma^2_{D^0}$ [MeV/c^2]	6.9 ± 0.3	6.8 ± 0.3
f_1	0.43 ± 0.09	0.47 ± 0.09
α_{B^+} [(MeV/c^2)$^{-1}$]	-0.0030 ± 0.0002	-0.0027 ± 0.0002
α_{D^0} [(MeV/c^2)$^{-1}$]	-0.0015 ± 0.0003	-0.0013 ± 0.0003

Table 4: Fitted parameters obtained from the 2011 data candidates.

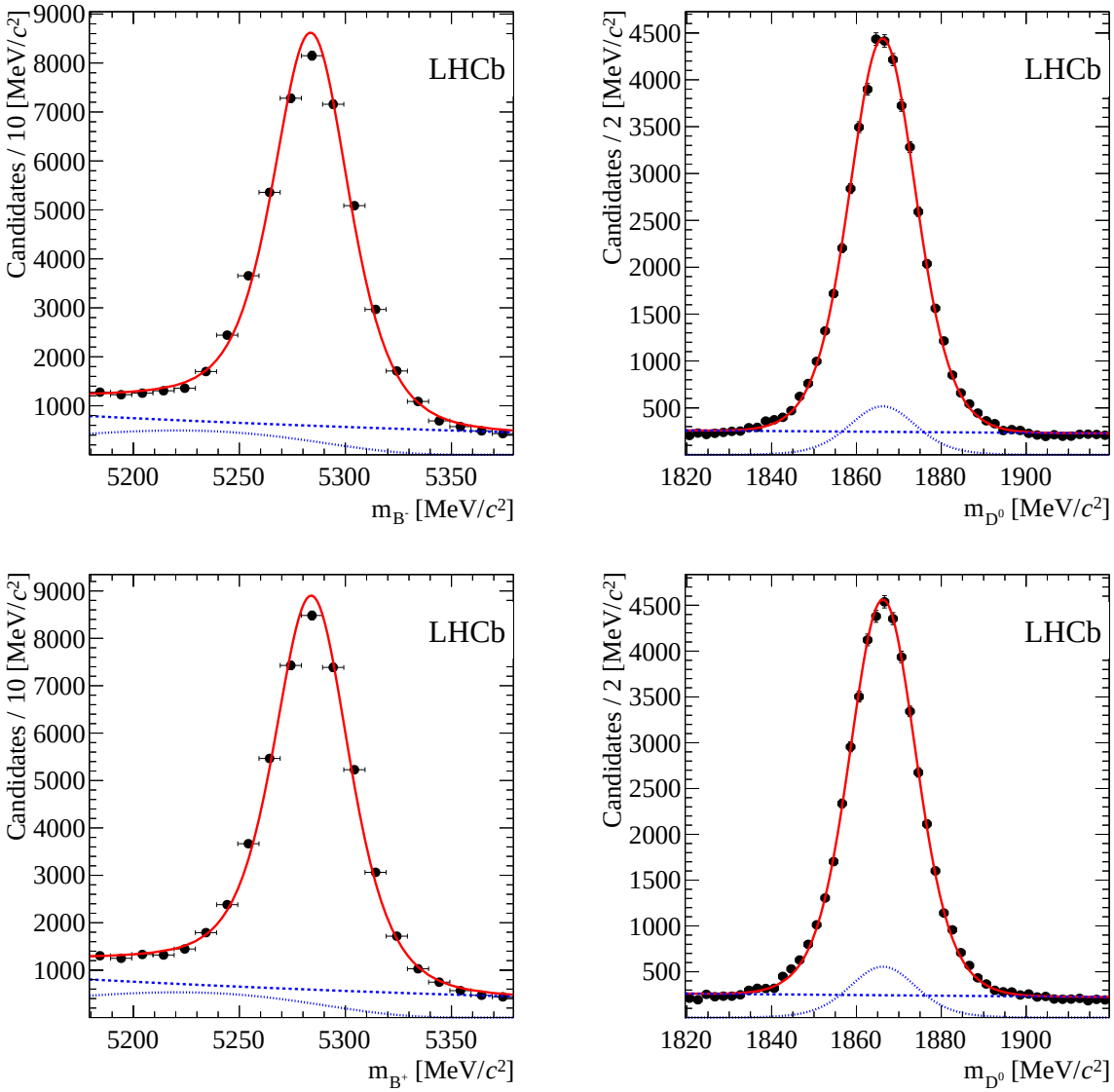

Figure 2: Projections on to the $K\pi\pi$ (left) and $K\pi$ (right) invariant masses from the fit to the $B^- \to D^0\pi^-$ (top) and $B^+ \to \overline{D}^0\pi^+$ (bottom) datasets. The total PDF is given by the solid red line, the dashed blue line represents the combinatoric component and the dotted blue line represents the partially reconstructed background.

4 PID Performance Evaluation

PID performance is evaluated by measuring the efficiency of a particular PID requirement on a given particle species in conjunction with the mis-identification probability when the requirement is applied to a species intended to be rejected. Therefore the performance of the $\mathrm{DLL}_{K\pi}$ variable can be evaluated in terms of the K efficiency and the $\pi \to K$ mis-identification probability of a given $\mathrm{DLL}_{K\pi} > x$ requirement using the D^0 daughter tracks. The markers in Figures 3, 5, 7 and 9 correspond to different x values, and a mis-alignment of the markers for the offline and online curves indicates the raw $\mathrm{DLL}_{K\pi}$ values differ in the two cases.

Standard tools are used to extract the online PID performance of the $\mathrm{DLL}_{K\pi}$ variable from the sWeighted $B^+ \to \overline{D}^0\pi^+$ datasets for each of the scenarios defined in Section 2. The performance of the offline calculated $\mathrm{DLL}_{K\pi}$ variable is extracted using prompt $D^{*+} \to (D^0 \to K^-\pi^+)\pi^+$ decays as the calibration sample, making use of the pion and kaon samples provided by the D^0 decay.

5 Online/Offline Performance of $\mathrm{DLL}_{K\pi}$

5.1 2011 HLT Conditions

The kaon efficiency versus the $\pi \to K$ mis-identification probability is shown in Figure 3, where similar performance can be seen, though the performance of online conditions is noticeably poorer than found in the offline calculation. It is also worth noting that while the disagreement between the online and offline calculation is not very large considering the sacrifices made to ensure a timely calculation, the decreased efficiency in the case of the online calculation poses difficulties in the exact knowledge of the efficiencies of PID cuts for analyses using candidates that have passed both the online and offline PID requirements.

Performance is shown as a function of momentum for cuts of $\mathrm{DLL}_{K\pi} > 0.0$ and $\mathrm{DLL}_{K\pi} > 5.0$ in Figure 4. As expected, efficiencies are consistently lower for the case of the online calculation. However for the mis-identification probability, the picture is mixed. For momentum values larger than $20\,\mathrm{GeV}/c$, the mis-identification probabilities are lower offline than in the HLT, however for momentum $< 20\,\mathrm{GeV}/c$, the mis-identification probabilities are found to be lower as calculated in the HLT.

The efficiency versus mis-identification probability curve in Figure 3 also shows that, while a given value of the $\mathrm{DLL}_{K\pi}$ variable has a different meaning when calculated in the HLT and offline, similar working points can be found using different cut values on the $\mathrm{DLL}_{K\pi}$ variable.

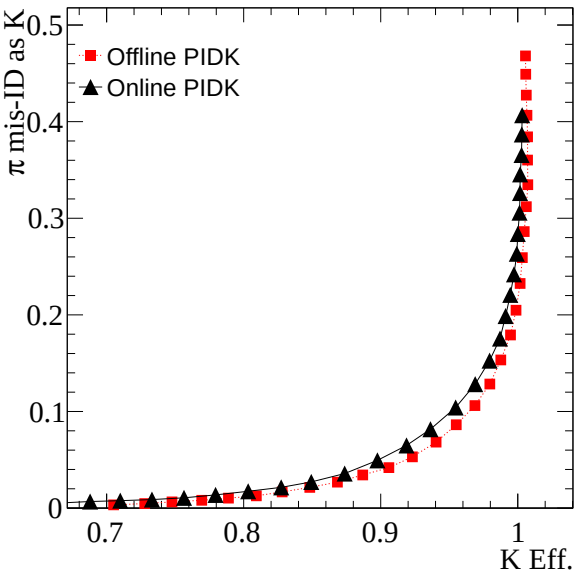

Figure 3: Efficiency with respect to kaons versus mis-identification probability with respect to pions of the offline $\mathrm{DLL}_{K\pi}$ variable (red squares) compared to the online $\mathrm{DLL}_{K\pi}$ variable (black triangles) as calculated by the HLT in 2011.

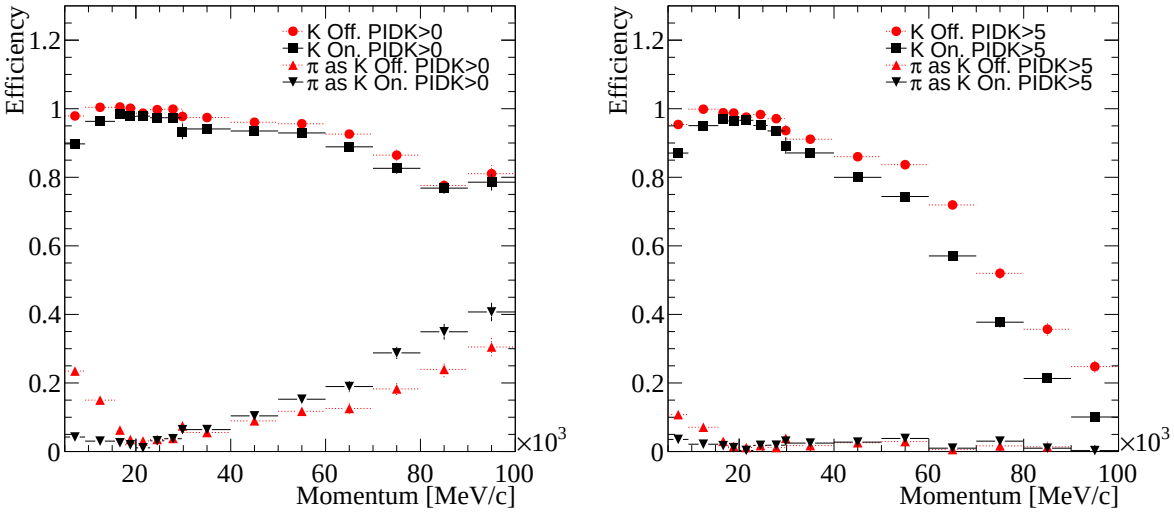

Figure 4: Kaon efficiency as a function of momentum for the $\mathrm{DLL}_{K\pi} > 0$ (left) and $\mathrm{DLL}_{K\pi} > 5$ (right) requirements on the offline calculation (red circles) compared to the online calculation (black squares). The pion mis-identification probability is given by the triangular markers.

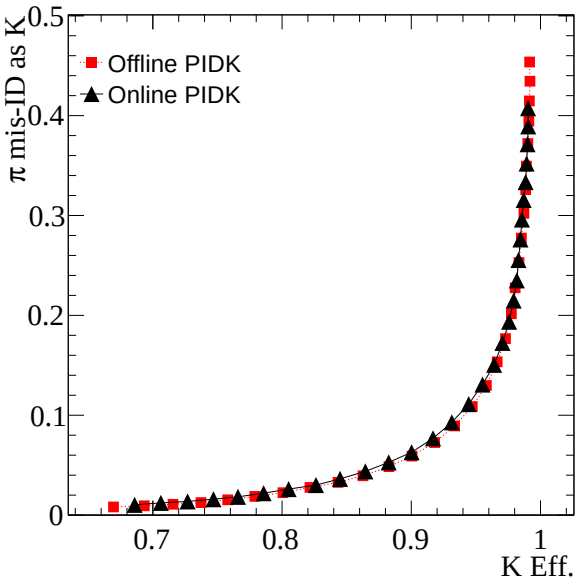

Figure 5: Efficiency with respect to kaons versus the mis-identification probability with respect to pions of the offline $\mathrm{DLL}_{K\pi}$ variable (red squares) compared to the online $\mathrm{DLL}_{K\pi}$ variable (black triangles) as calculated by the HLT in Scenario 2.

5.2 Scenario 2

The modification to the PID calculation to the "Offline" context ensures that the latest calibrations of the RICH mirror alignment and refractive index are used. In addition the inclusion of all available charged particle hypotheses ensures that the $\mathrm{DLL}_{K\pi}$ calculation has all the degrees of freedom that were available to the offline calculation.

The effect on the kaon efficiency against the $\pi \to K$ mis-identification probability can be seen in Figure 5, where the online and offline calculation can be seen to agree much more closely, though again there is an offset in the curves, showing that while a similar working point can be found for each calculation, the raw PID variable values differ.

The efficiencies and mis-identification probabilities as a function of momentum are shown in Figure 6, again for cuts of $\mathrm{DLL}_{K\pi} > 0.0$ and $\mathrm{DLL}_{K\pi} > 5.0$. The efficiency is then found to be in good agreement between the value calculated in the HLT and that found offline, though again efficiencies are consistently better in the offline case in all but the highest momentum bins. The $\pi \to K$ mis-identification probability is also very similar between the online and offline calculation, however the high momentum regions show an improved performance with the offline calculation, as was seen in the previous section.

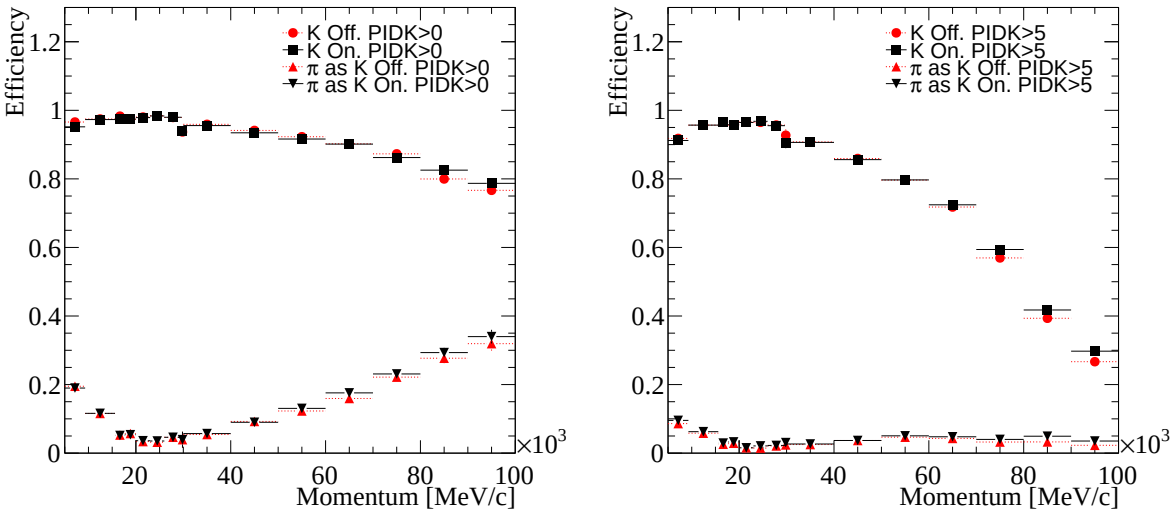

Figure 6: Kaon efficiency as a function of momentum for the $\mathrm{DLL}_{K\pi} > 0$ (left) and $\mathrm{DLL}_{K\pi} > 5$ (right) requirements on the offline calculation (red circles) compared to the online calculation in Scenario 2 (black squares). The pion mis-identification probability is given by the triangular markers.

5.3 Scenario 3

In Scenario 3, the kinematic requirements on the HLT tracks were removed. This then meant that all available tracks were considered for the assignment of a $\mathrm{DLL}_{K\pi}$ value. While the presence of these kinematic requirements should not have a big impact on performance, the absence of the tracks that do no meet the requirements could alter the $\mathrm{DLL}_{K\pi}$ value, altering the efficiency versus mis-identification probability working point for a given cut value.

Figure 7 shows the kaon efficiency versus the $\pi \to K$ mis-identification probability, where it can be seen that the difference in performance between the online and offline calculated values appears very similar to that reported from Scenario 2 in Figure 5. However, one important difference is that the offset of the curves has been reduced.

The performance as a function of momentum is shown in Figure 8, where again similar efficiencies and mis-identification probabilities can be seen between the online and offline calculations.

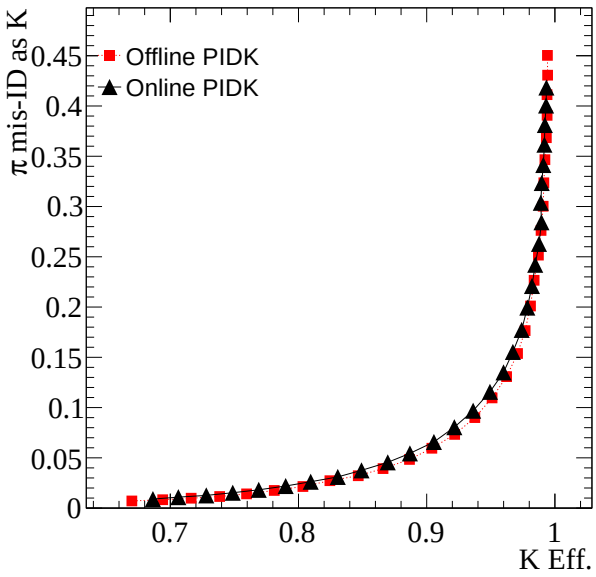

Figure 7: Efficiency with respect to kaons versus the mis-identification probability with respect to pions of the offline $\mathrm{DLL}_{K\pi}$ variable (red squares) compared to the online $\mathrm{DLL}_{K\pi}$ variable (black triangles) as calculated by the HLT in Scenario 3.

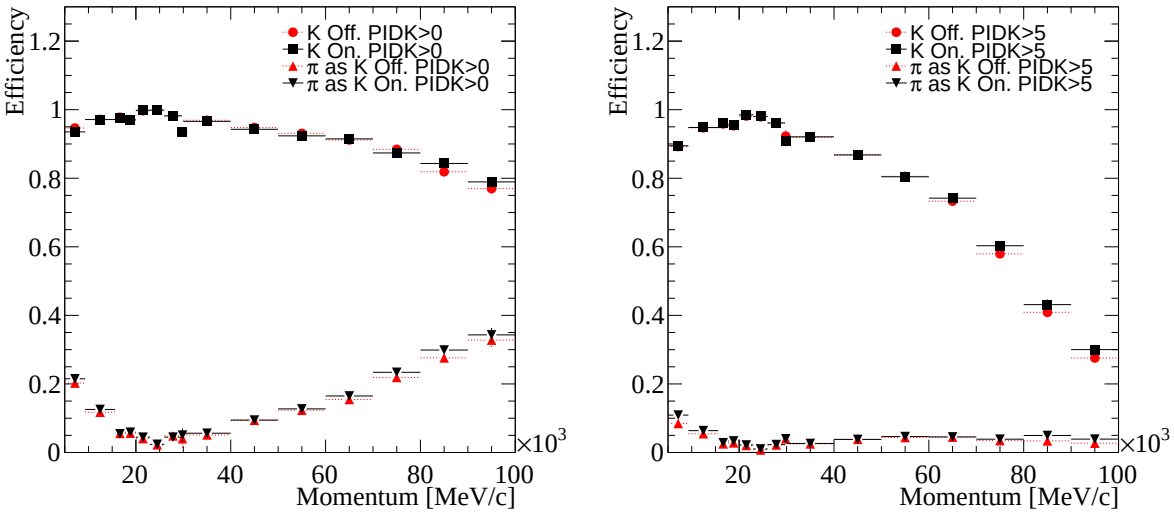

Figure 8: Kaon efficiency as a function of momentum for the $\mathrm{DLL}_{K\pi} > 0$ (left) and $\mathrm{DLL}_{K\pi} > 5$ (right) requirements on the offline calculation (red circles) compared to the online calculation in Scenario 3 (black squares). The pion mis-identification probability is given by the triangular markers.

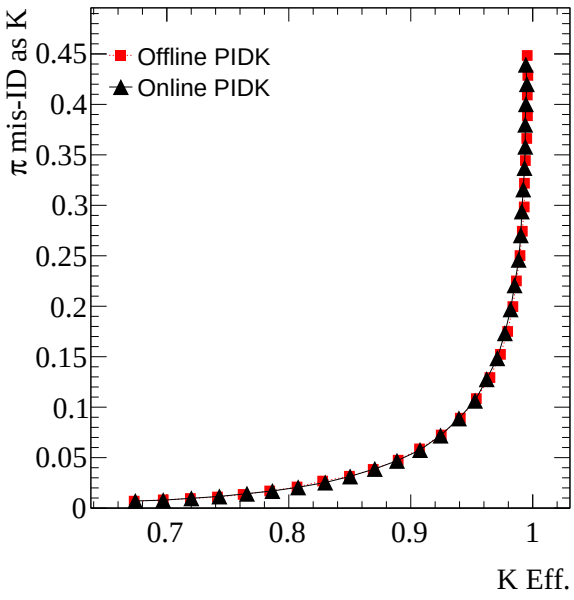

Figure 9: Efficiency with respect to kaons versus the mis-identification probability with respect to pions of the offline $\mathrm{DLL}_{K\pi}$ variable (red squares) compared to the online $\mathrm{DLL}_{K\pi}$ variable (black triangles) as calculated by the HLT in Scenario 4.

5.4 Scenario 4

The explicit addition of the Aerogel radiator to the online calculation can be seen to improve the performance to be nearly identical to that of the online calculation. It is worth noting that PID performance in the low momentum regions is mainly provided by the Aerogel radiator. The major difference between the online calculation used in Run 1 and the offline calculation is that the Aerogel is not included in the online calculation. It can be seen from Figure 9, that the agreement in terms of the efficiency versus the mis-identification probability shows good agreement between the offline and online calculations, with the offset between the curves reduced even more than seen for Scenario 3 in Figure 7. The performance as a function of momentum is shown in Figure 10, where near identical performance is seen between the online and offline calculation.

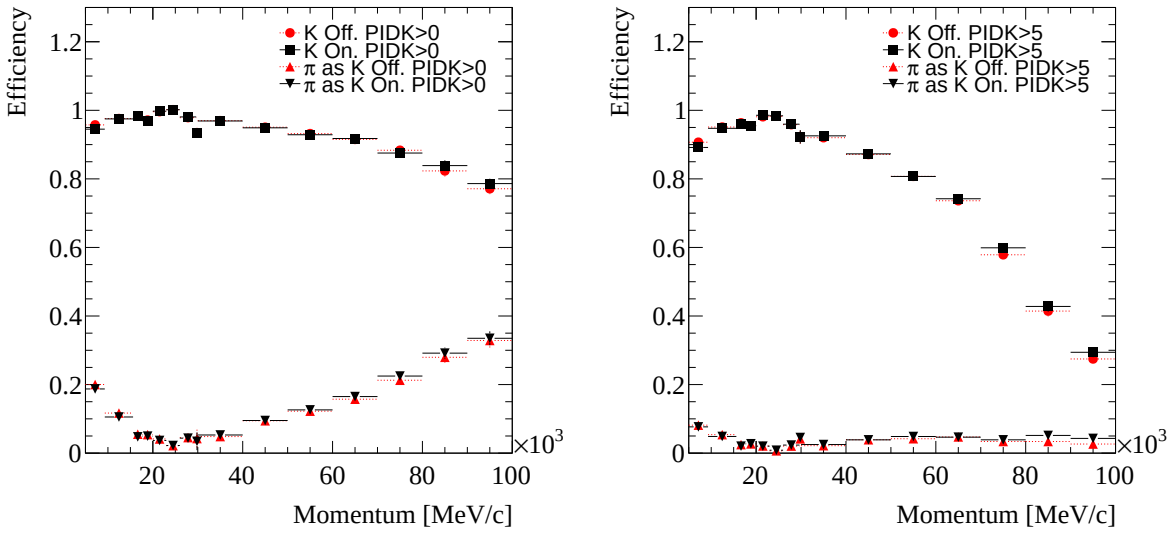

Figure 10: Kaon efficiency as a function of momentum for the $\mathrm{DLL}_{K\pi} > 0$ (left) and $\mathrm{DLL}_{K\pi} > 5$ (right) requirements on the offline calculation (red circles) compared to the online calculation in Scenario 4 (black squares). The pion mis-identification probability is given by the triangular markers.

6 Event Level $\mathrm{DLL}_{K\pi}$ Differences

In addition to measuring the efficiencies and mis-identification probabilities from the D^0 daughters, differences in the online and offline $\mathrm{DLL}_{K\pi}$ calculations can be studied in terms of the per track differences in PID variable values in each event. Figure 11 shows the distribution of the difference in the online and offline calculated $\mathrm{DLL}_{K\pi}$ variables for the bachelor pion in $B^+ \to \overline{D}^0\pi^+$ events.

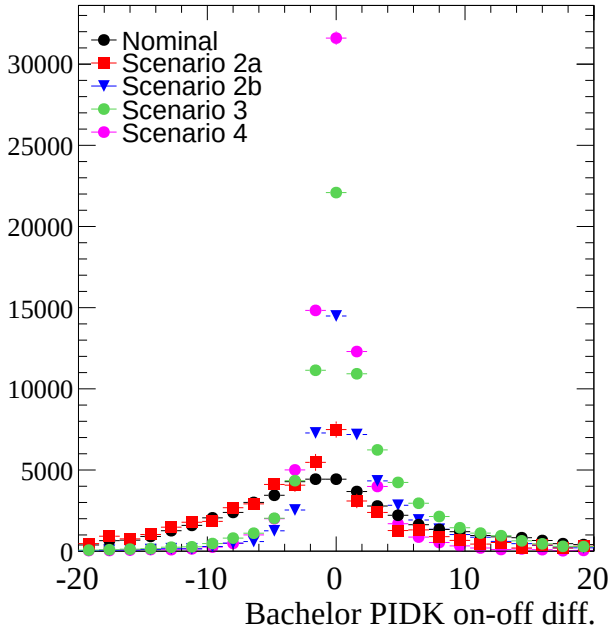

Figure 11: Distribution of the difference in the $\mathrm{DLL}_{K\pi}$ variable for the bachelor pion in $B^+ \to \overline{D}^0\pi^+$ candidates calculated for each event in Scenarios 2a, 2b, 3 and 4, compared to the calculation as performed in Run I data taking.

For the purposes of comparing the online and offline $\mathrm{DLL}_{K\pi}$ variable at the track level, Scenario 2 has been split in to two parts, 2a and 2b, where Scenario 2a corresponds to:

- Increasing the number of mass hypotheses considered from the two used in the HLT (π or K) to the six used offline (π, K, μ, p, e or "belowThreshold").
- Changing the likelihood calculation from the "FastGlobal" setting to the "FullGlobal" setting used offline.

Scenario 2b corresponds to the changes in Scenario 2a in addition to:

- Removal of RICH selection requirements (based on the transverse momentum, momentum and track fit quality), such that all forward reconstructed tracks are considered in the PID calculation.

- The context of the RICH calculation changed to “Offline” from “HLT”.
- The utilisation of the latest available database tags.

It can be seen from Figure 11, that the difference between the online and offline calculation peaks at zero though the peak is clearly wide and asymmetric for the calculation as performed in Run 1 data taking. The addition of all mass hypotheses and the modification of the PID configuration to the offline setting greatly improves the agreement, as does the removal of the RICH track cuts and the change in context to match the offline calculation.

The removal of the forward reconstruction selection requirements, allowing all tracks to be considered for the $\mathrm{DLL}_{K\pi}$ calculation understandably increases agreement as the tracks used for the offline calculation are subjected to looser kinematic and track fit quality requirements. However it can be seen that in Scenario 3, while a peak with smaller width is obtained, a shoulder on the right side of the maximum persists. The cause for this has been found to be due to the RICH radiators included in the $\mathrm{DLL}_{K\pi}$ calculation.

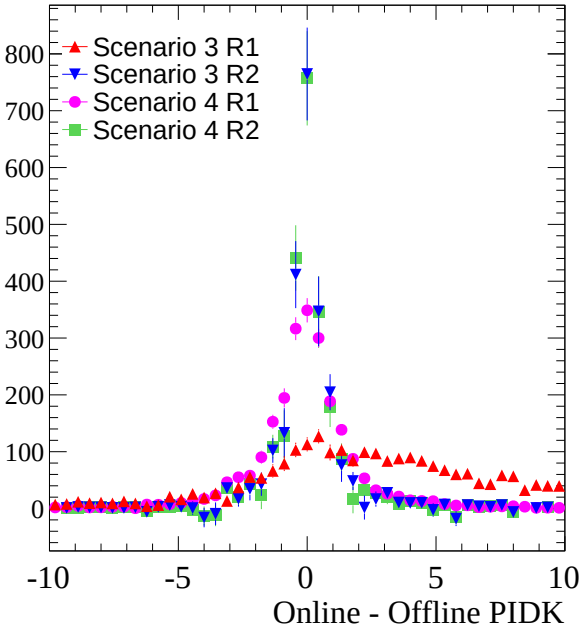

Figure 12: Distribution of the per-event difference in the $\mathrm{DLL}_{K\pi}$ variable for the bachelor pion in $B^+ \to \overline{D}^0\pi^+$ candidates. Online $\mathrm{DLL}_{K\pi}$ values are calculated using the configurations of Scenarios 3 and 4, and the distributions are shown separately for RICH1 and RICH2.

Figure 12 shows the difference in $\mathrm{DLL}_{K\pi}$ for bachelor tracks that use only the RICH1 and RICH2 radiators in Scenarios 3 and 4. For Scenario 3, it can be seen that the cause of the shoulder in the right hand side is the $\mathrm{DLL}_{K\pi}$ calculated in RICH1. This prompted the Aerogel radiator to be explicitly added to the online calculation. The effect of this can

Scenario	Time / event [ms]
1	15
2a	18
2b,3	76
4	100

Table 5: Timing measurements for different RICH configurations.

be seen in the change on the RICH1 radiator in Scenario 4, where much better agreement is seen and no shoulder is present in Figure 11.

7 CPU time requirements

The CPU requirements of the various configurations discussed in Section 2.2 have been measured using a machine typical of those used to run the HLT online. The results are listed in Table 5. These measurements use a much more modern version of the HLT application, Moore v22r0, than was used in 2011 data taking and for the performance studies in this Note. This is because the measurements are mostly relevant in the context of deciding what PID configurations can be used when data taking resumes in 2015. A sample of $10,000$ events recorded in 2012 based only on the hardware trigger decision was used. This sample has an average of 1.6–1.7 visible pp interactions per bunch crossing.

It should be noted that the configurations in Table 5 do not precisely correspond to those in Section 2.2: the forward reconstruction selection requirements are not removed in scenarios 3 and 4, leaving scenarios 2/2b and 3 identical to one another. This is because the default values, which were used in 2012, are thought to be representative of what is viable in 2015.

It is important to stress that HLT selections will significantly reduce the rate using kinematic and geometric preselections before the PID information is calculated and used to obtain a final reduction in output rate. This means that the PID calculation is only performed at a few times the output rate of the HLT selections which take advantage of it, and the average PID calculation time per event passed into the HLT will be much lower than those quoted in Table 5.

8 Summary and Conclusions

The PID performance as calculated in the HLT has been evaluated in terms of the $\mathrm{DLL}_{K\pi}$ variable in $B^+ \to \overline{D}^0\pi^+$ data events. The kaon efficiency versus the $\pi \to K$ mis-identification probability curves have been found to be similar between the online and offline calculations, though differences in raw $\mathrm{DLL}_{K\pi}$ variable values can be large at the per-track level.

Through the inclusion of all particle hypotheses and the modification of the PID calculation to use offline calibrations and settings, the PID performance, in addition to the track level $\mathrm{DLL}_{K\pi}$ value, has been found to show increasingly good agreement between the online and offline calculations. A residual difference remains, even when the calculation settings of the $\mathrm{DLL}_{K\pi}$ variable are the same in the HLT as calculated offline.

The large differences in raw $\mathrm{DLL}_{K\pi}$ values, at the per-track level, between the online and offline calculations have been shown to be related to the inclusion of the Aerogel radiator in the offline calculation. Adding the Aerogel radiator to the online calculation explicitly is seen to significantly reduce the effect.

The time taken to perform the online calculation used in Run 1 has been measured as 15 ms per event. When all settings are changed to match the offline calculation, the time taken rises to 100 ms per event, around 25% of which is due to explicit inclusion of the Aerogel.

Acknowledgements

We would like to thank both Vladimir V. Gligorov and Chris Jones for their guidance and many helpful discussions.

References

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

[2] M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, LHCb-DP-2012-003.

[3] V. V. Gligorov, *A single track HLT1 trigger*, LHCb-PUB-2011-003.

[4] R. Aaij *et al.*, *The LHCb Trigger and its Performance in 2011*, JINST **8** (2013) P04022, arXiv:1211.3055.

[5] M. Adinolfi *et al.*, *Online alignment and calibration of LHCb beyond LS1 in the context of a split HLT*, LHCb-INT-2013-021. CERN-LHCb-INT-2013-021.

[6] M. Pivk and F. R. Le Diberder, *sPlot: a statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, arXiv:physics/0402083.

[7] E. W. Weisstein, *Student's t-Distribution*, From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/Studentst-Distribution.html.

[8] ARGUS Collaboration, H. Albrecht *et al.*, *Search for Hadronic $b \to u$ Decays*, Phys. Lett. **B241** (1990) 278.