Belle II Reconstruction Software and Performance

J. Bennett, T. Bilka, G. Casarosa, T. Ferber, J-F. Krohn, L. Piilonen, L. Santelj, M. Staric

Abstract

An overview of the reconstruction software and performance is presented. This document is a draft chapter of the B2TiP book.

1 Reconstruction Software

Section author(s): J. Bennett, T. Bilka, G. Casarosa, T. Ferber, L. Piilonen, L. Santelj, M. Staric

1.1 Introduction

Online and offline data handling is performed by the Belle II analysis software framework (basf2). The framework is designed to allow independent processing blocks called modules to perform relatively small tasks, which are executed linearly within a defined path. The configuration of modules for a specific purpose is defined using steering files. Modules communicate by passing information to and from a common storage element, which also keeps track of relationships between data objects.

Given the enormous data output rate at Belle II, a robust and efficient framework for data analysis is vital. Data sets will be processed in several phases, with a reduction and enhancement occurring in each phase. The raw data is reconstructed to provide physical quantities from detector information like track hits and calorimeter clusters. The hit and cell level information is then discarded and the data is reduced by approximately a factor of 40. This reduced information is then used to determine particle level information such as four-momentum and event shape variables.

Unless otherwise noted, the performance plots and reconstruction algorithms described in this chapter are based on basf2 release-00-05-03, which was used in the fifth MC campaign (MC5).

1.2 Tracking

Charged particle tracking is achieved using the two innermost Belle II detectors, the VXD (which consists of two components, the SVD and the PXD) and the CDC. The silicon detector provides essential information in order to precisely determine the track impact parameters, and therefore the vertex position of the decaying particles. The Drift chamber provides an excellent measurement of the momentum of the charged tracks.

Tracking is divided into three main stages.

- *Pattern recognition*: Detector hits belonging to a single track are collected together into a track candidate.

- *Track fitting*: A helical fit is applied to the track candidate, taking into account the multiple scattering and the energy loss with a Kalman Fit. The track is extrapolated to the Point of Closest Approach (POCA) and the five helix parameters are stored for offline reconstruction.

- *V0-like particle reconstruction*: Charged tracks are fit to a common vertex in order

to find long-lived particles, such as K_S^0, Λ, and converted photons, that most likely decay outside the beampipe but inside the tracking volume. The track pair is stored in the V0 object.

1.2.1 Pattern recognition algorithms

The features of detector hits in the CDC or the VXD are completely different. Therefore, dedicated pattern recognition (PR) algorithms for the two detectors have been developed. The PR modules provide track candidates, *i.e.* collections of hits belonging to the same track, and are described below. The track candidates from the PR of the two detectors are then merged together according to the distance between the VXD and CDC track candidates extrapolated to the CDC outer wall. In the future, there will be also modules to extrapolate CDC track candidates toward the VXD detector planes in order to add VXD hits to CDC track candidates. The same will be done for the addition of CDC hits to the VXD track candidates, extrapolating the track outward. These are not presently included in the tracking package.

Vertex Detector

*** brief description VXD pattern recognition, no plots - or maybe one ***
No background cluster filter is applied yet.

Drift chamber

*** brief description CDC pattern recognition, no plots - or maybe one ***
First, a background hit filter is applied. Then the PR algorithm is performed with two complimentary approaches.

- *The Legendre Finder* is a global approach with hit information transformed to Legendre-space and a fast, binned quad-tree search.

- *The Cellular Automaton Finder* is a local approach in which segments of nearby hits are combined via hit following.

Finally, the track candidates are merged.

1.2.2 Track fitting

*** brief description of the helix parameters and the Kalman Fit ***
The track candidates are then fitted to tracks with the mass hypothesis of the pion.

1.2.3 Combined performance

*** tracking performance plots about efficiency and quality of the tracks. The plots go here. ***

1.2.4 V^0 reconstruction

Long-lived neutral particles that decay into two charged particles at some distance away from the interaction point are reconstructed using a dedicated basf2 module. This V^0 reconstruction takes place after the reconstruction of charged particles and is intended to avoid extrapolation through material on the analysis level, where the actual V^0 selection takes place. This is in accordance with the design goal of removing dependence of analysis level information on knowledge about the detector material (***not only, the daughters have not actually gone through that material, they didn't exist yet***).

The goal of V^0 reconstruction is to keep all reasonably accurate V^0 vertices outside the beam pipe as well as those inside the beam pipe whose reconstructed mass is reasonably close to the K_S mass. Unlikely track combinations may be suppressed by restricting the χ^2 from the vertex fit or the radius of the V^0 vertex. For vertices inside the beam pipe, the reconstructed invariant mass, which will depend on the particle hypothesis made for the tracks and thus the hypothesis for the identity of the V^0, can also be restricted.

The V^0 reconstruction algorithm pairs all oppositely charged tracks and extrapolates each to the innermost hit of either track. If the extrapolation fails, the combination is rejected. Studies show that this restriction has no effect on efficiency. Each accepted combination is processed by the vertexing package RAVE. If the vertex fit fails, the
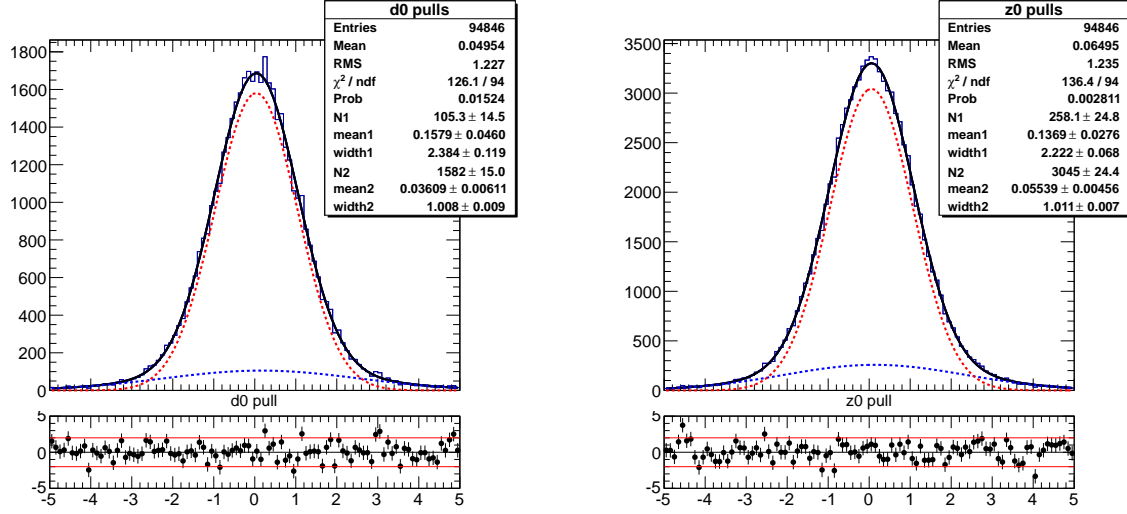
Figure 1.1: Transverse (left) and longitudinal (right) Impact parameters pulls fitted with the sum of two Gaussians.
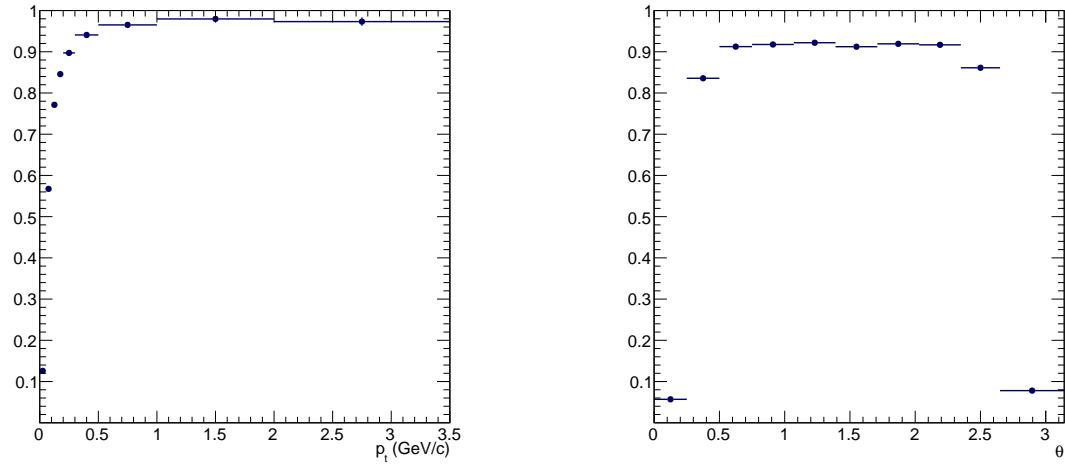

Figure 1.2: Tracking Efficiency as a function of transverse momentum (left) and polar angle (right).

3

combination is rejected. Each surviving combination is then subjected to default cuts on the radius (less than 1 cm inside the beam pipe and greater than 1 cm outside the beam pipe), the vertex fit χ^2 (less than 50), and, for vertices inside the beam pipe, the mass window (within 30 MeV/c^2 of the nominal K_S mass).

1.2.5 Alignment

To reach the design performance of the detector, various calibration constants must be determined. For the VXD, many of these constants describe the position and orientation of the silicon sensors. This calibration is commonly referred to as alignment. To determine the alignment constants, the so called global approach with the Millepede II tool [1, 2] has been chosen for use at Belle II. The alignment is computed through minimization of track-to-hit residuals by means of a linear least squares method. Because Millepede fits all track and alignment parameters simultaneously, all correlations are kept in the solution. Therefore it is desirable to determine as many constants simultaneously as possible. For this reason, the CDC is also integrated into the procedure and its alignment and particular calibration constants can be determined together with the VXD alignment. Investigation of the possibility to integrate other sub-detectors into the procedure, for example the alignment of the muon system, is ongoing.

Track Parametrization

Reconstructed tracks and decays, as well as cosmic muons, can be input of the alignment procedure. All such tracks are first re-fitted by the General Broken Lines algorithm (GBL) [3], integrated into the GENFIT toolkit [4] and basf2. The GBL parametrization allows a proper treatment of multiple scattering effects, adding additional fit parameters (kink angles) to an initial reference trajectory derived from the result of the standard reconstruction output. The additional degrees of freedom are removed by constraining the variance of multiple scattering angles from the moments of the radiation length distribution along the particle reference trajectory.

The track is locally parametrised by five or four parameters at each measurement plane depending on the presence or absence of a magnetic field, respectively. For the drift chamber, a virtual measurement plane is constructed by means of the GENFIT formalism.

Combined objects, composed of multiple particle tracks constrained to originate from a common vertex can be an input of the alignment as well. Optionally, a beam constraint can be added for decays originating from the primary interaction point, like di-muon events. In a similar manner, e.g. two body decays with an invariant mass constraint can be introduced in the procedure.

Alignment Parametrization

In the VXD, the sensors are parametrized as planes with six rigid body alignment parameters. For the 212 sensors, this means 1276 parameters. Sensor deformations or additional calibrations of the Lorentz angle in the magnetic field can be included in the procedure as well. For the treatment of correlated movements of sensors, a hierarchy of alignment objects might be defined. This allows the treatment of time dependence of larger structures, different from internal sensor alignments during the simultaneous minimization.

For the CDC, the alignment of the layers and larger structures, e.g. end plates is considered. The x-t relation can be calibrated, as well as channel timing offsets or time walk corrections. For the muon system, the modules are treated as rigid planar bodies in the initial stage.

Beam and vertex constrained decays rely on the estimation of the primary vertex position and corrections to it can be determined as well during the simultaneous minimization. Only the position of the primary vertex is calibrated. The covariance matrix of the primary beam spot is an input of the procedure and will be calibrated by other means.

Alignment Workflow

Millepede II is integrated into the common calibration framework, which makes use of dedicated basf2 modules to collect samples and runs calibration algorithms. During the collection step for alignment, reconstructed tracks are re-fitted using the nominal detector positions corrected with previously determined alignment constants. Each detector interfaces this procedure via a special class representing the local-to-global transformation. This class also provides the derivatives of local residuals with respect to its assigned calibration parameters.

Due to the nature of the global approach, various track samples (primary decays, background, cosmic rays, etc.) from different operating conditions (cosmics without magnetic field) should be included in the procedure. Various samples and their importance are under investigation.

Some constants might be allowed to be time-dependent, especially those affecting many measurements, like positions of large structures, while keeping the procedure computationally manageable. The procedure as a whole can also be applied locally, for example only for the PXD alignment or to determine the relative alignment of the PXD and SVD. If such corrections are determined on-line, they will serve as initial values for the global procedure, when enough data and track samples are accumulated.

1.2.6 Beamspot

*** do we need this? there is no beam spot, in the sense that events are generated at 0,0,0 - CHECK THIS for rel 5.3 ***

1.3 Calorimeter reconstruction

The electromagnetic calorimeter is used to reconstruct the energy and position of depositions from neutral and charged particles with the best possible resolution and efficiency. While the energy and position reconstruction is primarily needed for photons and neutral hadrons, it may also aid the electron and charged hadron reconstruction in regions without or with limited tracking coverage. The sum of all reconstructed showers is used to constrain the missing energy in decays involving neutrinos. A special case is the reconstruction of highly energetic $\pi^0 \to \gamma\gamma$ decays where the two photon showers overlap.

The second task of the calorimeter is particle identification for electrons, muons, charged hadrons, neutral hadrons and photons based on shower shape variables and tracks matched to clusters.

The current clustering is an incomplete adaptation of the Belle clustering code which was developed for a low background environment and does not use timing information for each crystal during clustering. It starts from an initial list of crystals with energy deposits above some threshold, nominally 0.5 MeV, which is about twice the expected level of noise equivalent energy from electronics noise. To obtain some robustness against high beam backgrounds, the energy threshold was raised as function of crystal polar angle to between 1.28 MeV (barrel) and 2.5 MeV (outer endcap rings). A cluster starts with a seed crystal with at least 10 MeV that is a local energy maximum amongst its nearest neighbour crystals. A nearest neighbour touches either the side or the corner of the crystal and a local maximum is a crystal whose energy exceeds that of its next neighbours. All crystals from the initial list that are nearest or next-to-nearest neighbours of the seed crystal are added to the cluster. In the barrel, the size of a cluster is thus limited to symmetric 5×5 crystals. If clusters share crystals after this step, their energies are split according to the ratio of energy of each cluster to the sum of energies of all crystals in the overlapping clusters. This energy splitting does not provide the correct position nor the correct weighted list of crystals for subsequent shower shape calculations. The centroid $\vec{x}$ of each cluster is calculated by using linear weights of all

crystals in a cluster,

$$\vec{x} = \frac{\sum_i E_i \vec{x}_i}{\sum_i E_i}, \qquad (1.1)$$

where E_i is the energy of the i–th crystal and x_i is the center of the i–th crystal. It should be noted that this position reconstruction is known to be biased towards the crystal center of the highest energy crystal in the shower. The cluster energy is reconstructed as the linear sum over all included crystals. The peak position of the reconstructed photon energy is corrected to the true value in a subsequent step as a function of reconstructed polar angle and energy. The cluster time t_{cluster} is the time of the highest energetic crystal in the cluster. Clusters with $|t_{\text{cluster}}| < 125\,\text{ns}$ are rejected. Clusters are matched with tracks using a GEANT based extrapolation routine to determine whether a cluster is generated by a charged or a neutral particle. A cluster that contains a crystal hit by the extrapolated track is matched to that track.

The existing calorimeter reconstruction does not perform optimally in a high background environment and has various shortcomings (e.g. biased position reconstruction and oversimplified cluster splitting). The calorimeter reconstruction is currently being rewritten. The new cluster algorithm will reconstruct connected regions (CR) starting with single crystals with an energy of at least $10.0\,\text{MeV}$ (tbc) as seeds, as before. Surrounding crystals are added if their energy is above $0.5\,\text{MeV}$ (tbc). This procedure is continued if the added crystal energy is at least $1.5\,\text{MeV}$ (tbc). If two CRs share a crystal, they are merged. The optimal CR contains all deposited energy for a particle and merges CRs from different particles only if different particles deposit energy in the shared crystals. Each CR is then split into one or more clusters using different hypothesis-dependent algorithms.

In the following, the photon hypothesis – most important for the ECL reconstruction – is described. A boosted decision tree (tbc) is used to find photon cluster center crystals based on the nearest neighbours of each local maximum in the CR. Each CR with more than one local maximum is split into as many clusters as there are local maxima identified as photons. The crystals of each cluster are weighted assuming the lateral distribution of an electromagnetic shower which decreases exponentially

$$w_i = \frac{E_i e^{(-Cd_i)}}{\sum E_k e^{(-Cd_k)}}, \qquad (1.2)$$

where $C = 0.7$ (tbc) is a constant determined from MC and r_i is the distance between the i–th crystal and the cluster centroid. The centroid $\vec{x}$ of each cluster is calculated by using logarithmic weights of all crystals with positive weight w_i in a cluster,

$$\vec{x} = \frac{\sum_i w_i \vec{x}_i}{\sum_i w_i}, \qquad (1.3)$$

where $w_i = 4.0 + \log(E_i/E_{cluster})$ (tbc). This procedure is iterated until the average centroid positions of all clusters in the CR are stable within $1\,\text{mm}$ (tbc). Based on the measured background per event and the energy of the 8 nearest neighbours of the local maximum and the local maximum itself, an optimal number of highest energetic crystals among the 21 nearest neighbours - between 2 for low energy clusters at high background and 21 for no backgrounds - of the local maximum is chosen to determine the energy estimate of the cluster. A comparison of the energy resolution obtained using the current clustering code and the new clustering code (photon hypothesis) is shown in Figure 1.3. Other hypotheses include neutral hadrons, electrons and bremsstrahlung photons, charged hadrons and merged π^0 where the two photons cannot be separated into two different clusters. Track matching will be based on a likelihood of the nearest track to a cluster using the covariance matrix of the track fit.

1.4 Charged particle identification

Effective and efficient charged particle identification (PID) is vital to the physics goals of the Belle II experiment. Good PID information is
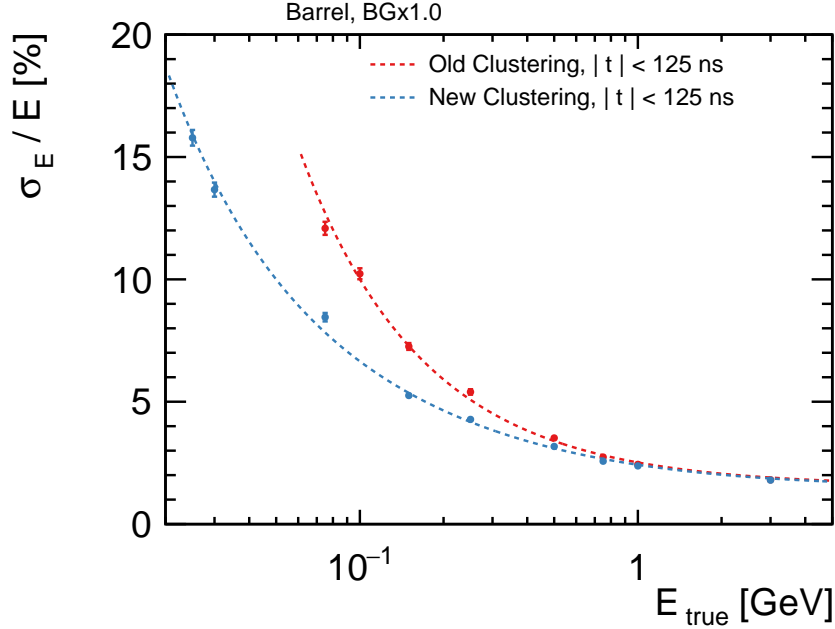
Figure 1.3: Peak energy resolution in the ECL barrel as function of true photon energy for full background.

necessary to isolate hadronic final states, reduce backgrounds and enable flavor-tagging techniques. The Belle II detector **??** contains an upgraded PID system, including a Time-Of-Propagation (TOP) counter in the barrel region of the detector and a proximity-focusing Aerogel Ring-Imaging Cherenkov (ARICH) detector in the forward end-cap region, to provide information on charged particles over the full kinematic range. The information from these detector systems is combined with that from specific ionization (dE/dx) measurements from the SVD and CDC to act as the primary sources of information for charged hadron PID. In a similar way, the ECL provides the primary information for use in electron identification and the KLM provides that for muon identification. Charged hadron and lepton PID is described in more detail in the following sections.

Charged particle identification at Belle II relies on likelihood based selectors. Information from each PID system is analyzed independently to determine a likelihood for each charged particle hypothesis. These likelihoods may then be used to construct a combined likelihood ratio. Analysis specific criteria may be used to construct prior probabilities. When combined with the likelihoods, the priors allow for the construction of the probability for a charged track to have a particular identity. This provides the optimal PID performance, but comes at the cost of requiring analysis specific optimization. The uncertainty on the selection efficiency cannot be pre-determined using this method.

The likelihood selectors rely on likelihood ratios constructed in the following way. First, the PID log likelihoods from each detector are summed to create a combined PID likelihood for each of six long-lived charged particle hypotheses: electron, muon, pion, kaon, proton and deuteron. Next, the difference in log likelihood between two particle hypotheses is used to construct a PID value $L(\alpha : \beta)$ according to

$$\mathcal{L}(\alpha : \beta) = \frac{1}{1 + e^{\ln \mathcal{L}_\alpha - \ln \mathcal{L}_\beta}}$$
$$= \frac{\prod_{\text{det}} \mathcal{L}(\alpha)}{\prod_{\text{det}} \mathcal{L}_\alpha + \prod_{\text{det}} \mathcal{L}_\beta}, \qquad (1.4)$$

where α and β represent two different particle types and the product is over the active detectors for the PID type of interest. The value $\mathcal{L}(\alpha : \beta)$ is greater than 0.5 for a charged track that more closely resembles a particle of type α than one of type β and is less than 0.5 otherwise. More details on the PID types are given in the following sections.

The performance plots included in this section were all generated from an inclusive sample of 10^6 $c\bar{c}$ events generated during the fifth MC campaign. These samples were reconstructed with release-00-05-03 of the Belle II software.

1.4.1 dE/dx measurements

The ionization energy loss, dE/dx, of a particle travelling through the Belle II detector is determined from measurements in the VXD and CDC. It is expected that the dE/dx measurement should depend only on the particle velocity, β, or equivalently $\beta\gamma = p/m$. Of course, a careful calibration is required to avoid systematic effects that break this dependece. The $\beta\gamma$ universality of dE/dx response for pions and kaons at Belle II is displayed in Fig. 1.4. In general, dE/dx information is of most use for particle momenta below about 1 GeV/c (Fig. 1.5).

Determination of Likelihoods

As dE/dx in the VXD and CDC depends on different physical processes, the two detectors make independent measurements and calibrations. At the time this document is written, the dE/dx reconstruction algorithms in both subsystems construct likelihood values using information from individual hits. A likelihood value is determined for each particle hypotheses, including pion, kaon, proton, muon, electron, and deuteron, using a lookup table constructed from large MC samples. To reduce the effect of outliers, the lowest 5% and highest 25% dE/dx measurements of each track are not used in the likelihood determination. It is also possible to calculate the likelihood using the truncated mean of dE/dx via a basf2 module option.

Future versions of the software will use a parameterization of the truncated mean and resolution to determine PID variables. A PID variable is determined by comparing the measured dE/dx truncated mean to a predicted value and resolution. The predicted values are calculated from a parameterization of dE/dx as a function of $\beta\gamma$. The predicted resolutions depend on the dE/dx measurement, the number of hits on the track, and the polar angle of the track. After determining the parameterization for the predicted means and resolutions, a PID variable χ is determined according to

$$\chi_h = \frac{I_{\mathrm{meas}} - I_{\mathrm{pred},h}}{\sigma_{\mathrm{pred},h}} \qquad (1.5)$$

where h is the particle type, I is the dE/dx truncated mean, and σ is the resolution for the given particle type. As the distributions of this χ variable are approximately Gaussian, it may be converted to a likelihood and combined with the output of other PID systems. The performance of such an algorithm is generally similar to the current method, but will enable a better charaterization of the resolution.

Performance

Defining the signal efficiency as the fraction of events relative to the generated quantity that have a likelihood of being identified as the true particle type greater than that of being identified as another particle type (e.g., $L_K>L_\pi$), the average kaon efficiency in the SVD below 700 MeV/c is about 96%. The comparable value in the CDC is nearly identical. The fraction of pions misidentified as kaons under the same criteria is about 3.1% in the SVD and about 1.1% in the CDC. Combining the information from these two detectors yields an average kaon efficiency of about 99.5% below 700 MeV/c, with a fake rate of about 0.2%. The kaon efficiency using dE/dx information is given in Fig. 1.6.

1.4.2   Charged hadron identification

Particle identification for charged hadrons, which in this context include pions, kaons, protons and deuterons, $\{\pi, K, p, d\}$, depends primarily on likelihood information from the CDC, TOP, and ARICH
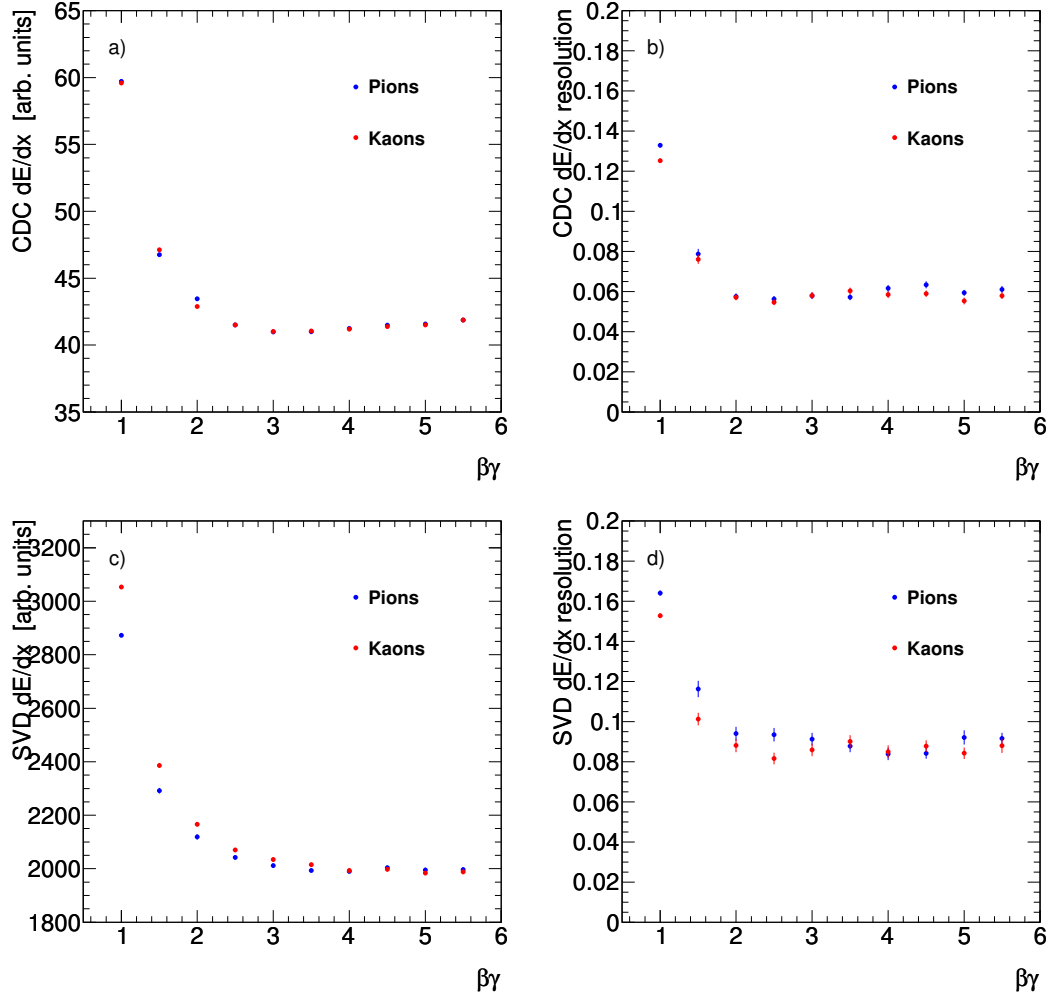
Figure 1.4: Truncated dE/dx means (a,c) and resolutions (b,d) for pion and kaon samples generated at specific values of $\beta\gamma$. The residual non-universality in the SVD is due to the fact that the measured momentum at the IP is different than the momentum in each SVD layer.
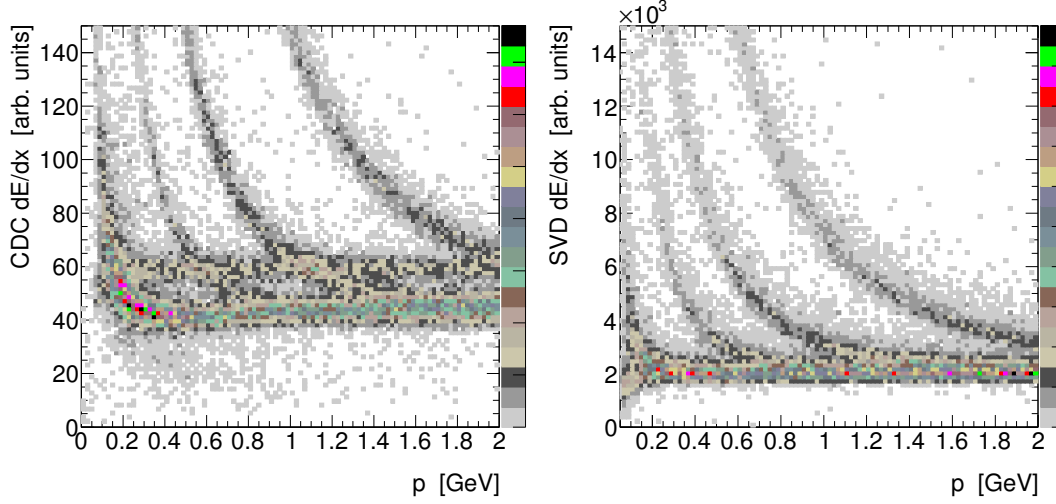
Figure 1.5: Truncated dE/dx means as a function of momentum for the CDC (left) and SVD (right). Distinct bands are evident for the various particle species below about 1 GeV/c.

detectors. These detectors also contribute to the particle identification of charged leptons, $\{e, \mu\}$. The methods to construct the likelihoods for each of these detector systems are briefly described here.

TOP Likelihoods

The TOP counter is a novel type of PID device that combines time-of-flight measurements with the Cherenkov ring imaging technique **??**. The dominant contribution to the resolution of this detector is the dispersion of light while propagating in the quartz bar. This effect is mitigated by focusing the Cherenkov light onto the photon detector with a spherical mirror and measuring a second coordinate of the photon impact position. To further improve the resolution, an expansion prism is added at the bar exit window.

The TOP counter consists of sixteen 2.7 m long modules positioned in the space between the CDC and the ECL and covers the polar angles from $32°$ to $120°$. The gaps between the modules account for about 5% of uncovered area.

An extended likelihood method is used to determine log likelihoods for the six long-lived charged particle types. The extended log likelihood probability for a given charged particle hypothesis h is

defined as

$$\ln \mathcal{L}_h = \sum_{i=1}^{N} \ln\Big(\frac{S_h(x_i, y_i, t_i) + B(x_i, y_i, t_i)}{N_e}\Big) + \\ \ln P_N(N_e), (1.6)$$

where $S_h(x, y, t)$ is the signal distribution for the hypothesis h, $B(x, y, t)$ is the distribution of background and $N_e = N_h + N_B$ is the expected number of detected photons, being a sum of the expected number of signal photons N_h for hypothesis h and the expected number of background photons, N_B. The channel coordinates are given by x and y, and the integration is performed over the full range t of the time-of-arrival measurement. The second term in Eq. 1.6 is the Poisson probability to obtain N photons if the mean is N_e.

The normalizations of $S_h(x, y, t)$ and $B(x, y, t)$ are:

$$\sum_{j=1}^{n_{ch}} \int_0^{t_m} S_h(x_j, y_j, t)dt = N_h, \qquad (1.7)$$

$$\sum_{j=1}^{n_{ch}} \int_0^{t_m} B(x_j, y_j, t)dt = N_B, \qquad (1.8)$$

where the sum runs over all channels n_{ch} of the photon detector array.
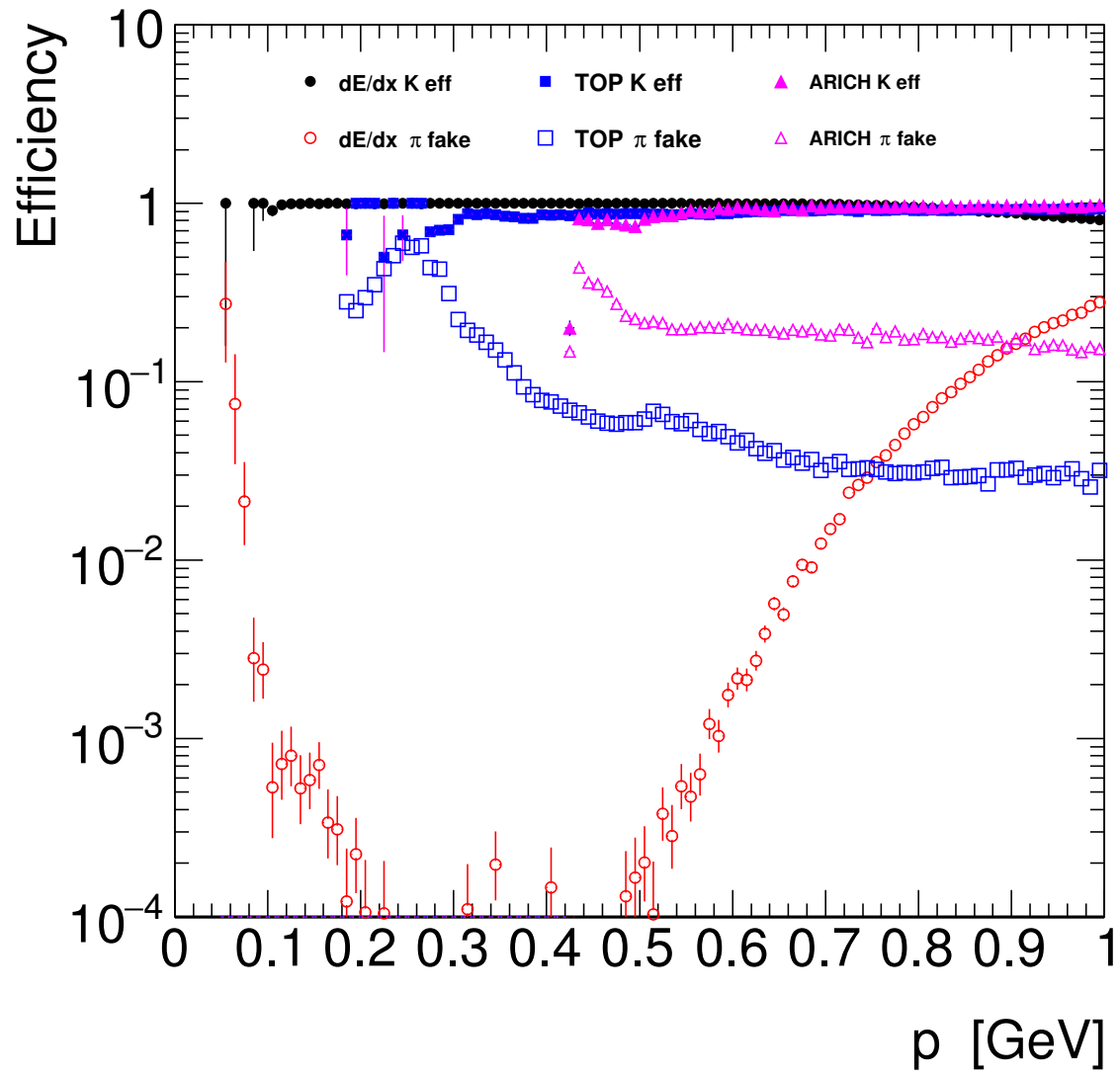
10

Figure 1.6: Kaon detection efficiency and pion fake rate using dE/dx information only.

The ring image of the TOP counter is a complicated pattern, which, besides the Cherenkov angle, also depends on the particle impact position and the angles with respect to the quartz bar. The distribution for a particular detection channel j can be parametrized as a sum of Gaussian distributions

$$S_h(x_j, y_j, t) = \sum_{k=1}^{m_j} n_{kj} g(t - t_{kj}; \sigma_{kj}), \qquad (1.9)$$

where n_{kj} is the number of photons in the k-th peak of channel j; t_{kj} is the position and σ_{kj} the width of the peak, and $g(t - t_{kj}; \sigma_{kj})$ is the normalized Gaussian distribution; and m_j counts the number of peaks in the channel j for $t < t_m$.

The quantities n_{kj}, t_{kj} and σ_{kj} are functions of the Cherenkov angle θ_c, the photon emission point (x_0, y_0, z_0) given by the particle impact position, the particle impact angles (θ, ϕ), and the unfolded channel coordinate $x_D^{kj} = ka \pm x_j$, where k represents the number of internal reflections at the side walls and a the width of the quartz bar. Using the above input data it is possible to solve for the unknown Cherenkov azimuthal angle ϕ_c^{kj} and thus determine the photon directional vector [5, 6].

Once the photon direction is known, t_{kj} is obtained by ray-tracing. The number of photons in the peak is calculated with

$$n_{kj} = N_0 \ell \sin^2 \theta_c \frac{\Delta \phi_c^{kj}}{2\pi}, \qquad (1.10)$$

where N_0 is the figure of merit of the Cherenkov counter, ℓ is the length of the particle trajectory in the quartz bar and $\Delta \phi_c^{kj}$ is the range of the Cherenkov azimuthal angle covered by the measuring channel j. The peak width σ_{kj} is obtained by summing in quadrature the following contributions:

- photon emission point spread (parallax error), $\sigma_\ell = \frac{dt_{kj}}{d\lambda} \frac{\ell}{\sqrt{12}}$, where λ is the running parameter of the particle trajectory inside the quartz bar $(0 \le \lambda \le \ell)$,

- multiple scattering of the particle in the quartz, $\sigma_{\text{scat}} = \frac{dt_{kj}}{d\theta_c} \theta_0(\ell/2)$, where $\theta_0(\ell/2)$ is calculated with the well known multiple scattering approximation [7],

- dispersion (chromatic error), $\sigma_{\text{disp}} = \frac{dt_{kj}}{de} \sigma_e$, where σ_e is the r.m.s of the energy distribution of detected photons in the channel,

- channel size $\sigma_{\text{ch}} = \frac{dt_{kj}}{dx_D} \frac{\Delta x_j}{\sqrt{12}}$, where Δx_j is the channel width,

- transit time spread of the photon detector, σ_{TTS}.

The derivatives $\frac{dt_{kj}}{d\lambda}$, $\frac{dt_{kj}}{d\theta_c}$, $\frac{dt_{kj}}{de}$ and $\frac{dt_{kj}}{dx_D}$ are calculated numerically according to the method described in detail in Ref. [5, 6].

Identification and mis-identification efficiencies have been studied with a MC simulation. Using generic $c\bar{c}$ samples we obtain the performance shown in Fig. 1.7. The efficiency is defined as the proportion of tracks that are properly identified according to the generated information for all tracks that fall within the TOP acceptance. In the momentum region below 2 GeV/c the efficiency to identify a kaon is about 94% with about a 4% fake rate to be mis-identified as a pion. Above 2 GeV/c the performance slowly decreases and gives about 85% efficiency with a 15% fake rate at 3 GeV/c. Fig. 1.7 also shows that when the nominal beam background is included the performance of the counter is not appreciably degraded. Other interesting studies are discussed in Ref. [8].

ARICH Likelihoods

In the Belle II detector, PID in the forward endcap is achieved with the aerogel ring imaging Cherenkov counter (ARICH). The ARICH covers the polar angle range from $17°$ to $35°$. Reconstructed tracks from the CDC are extrapolated to the ARICH detector volume and a likelihood function is constructed for each of the six different particle type hypotheses for tracks that pass through the aerogel layer. The likelihood function is based on a comparison of the observed spatial distribution of Cherenkov photons on the photodetector plane with the expected distribution for the given track parameters (position and momentum vector on the aerogel plane) for a given particle type.
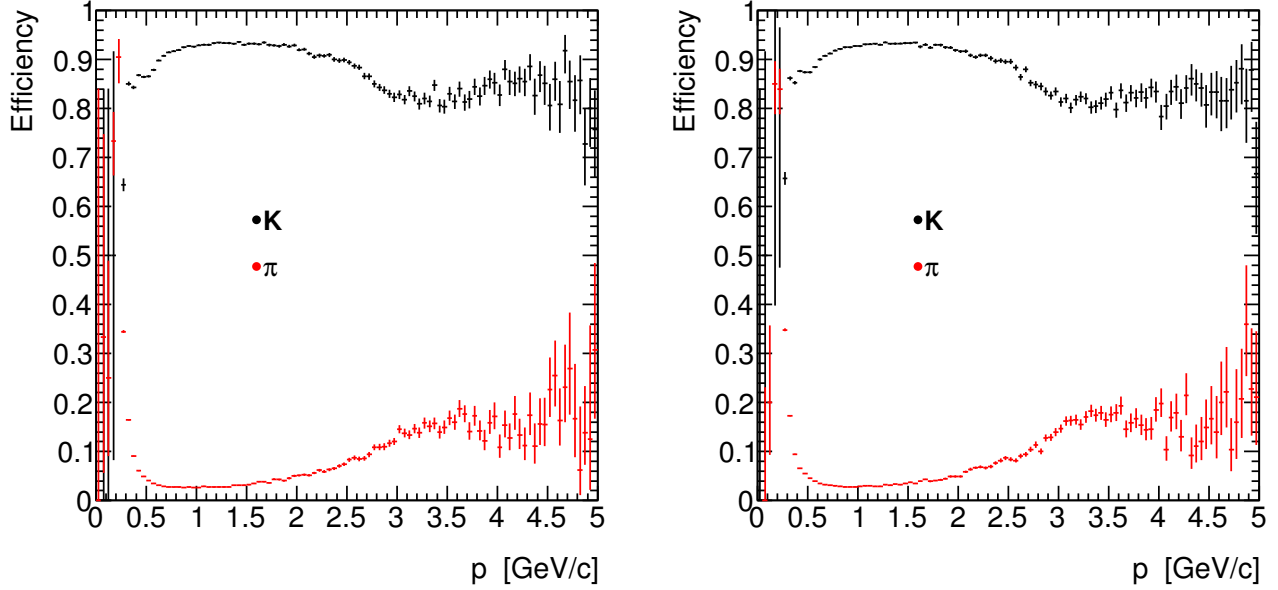
Figure 1.7: TOP counter kaon identification efficiency (black markers) and pion fake rate (red markers) as a function of momentum for $\ln \mathcal{L}_\mathcal{K} > \ln \mathcal{L}_\pi$, obtained with a MC simulation without beam background (left) and with beam background (right). Only tracks that fall within the TOP acceptance are considered.

The ARICH likelihood functions are constructed based on the method described in Ref. [9, 10]. For each of the particle type hypotheses (h) a likelihood function is calculated as $\mathcal{L}_h = \prod_i p_i^h$, where i runs over all pixels of the detector and p_i^h is the probability for pixel i to record the observed number of hits (1 or 0) assuming particle type h. As the p_i^h is a Poissonian distribution, one can show that $\mathcal{L}_h$ can be rewritten as $\ln \mathcal{L} = -N + \sum_i (n_i + \ln(1 - e^{-n_i}))$, where N is the expected total number of hits, n_i is the expected (calculated) average number of hits on pixel i, and the sum runs only over the pixels that were hit in an event[1].

The expected average number of hits on pixel i, n_i, is obtained as a sum of contributions from signal and background hits, $n_i = n_i^s + n_i^b$, where the signal contribution is divided into contributions from the first and second aerogel layers, $n_i^s = n_i^{s,1} + n_i^{s,2}$. The contribution of each aerogel layer (r) is calculated

---

[1] For transparency index h is omitted, but note that N and n_i depend on h.

as

$$n_i^{s,r} = \epsilon_{det} N^{s,r} \int_{\Omega_i} \frac{1}{2\pi} G(\theta, \theta_h^r, \sigma_h^r) d\theta d\phi \quad (1.11)$$

where ϵ_{det} is the photon detection efficiency and $N^{s,r}$ is the number of photons emmitted from aerogel layer r (theoretically calculated). The integral gives the probability for a Cherenkov photon being emmitted by particle type h from aerogel layer r into the solid angle covered by pixel i (θ and ϕ are the polar and azimuthal angles with respect to the track direction). A Gaussian function G with a mean at the expected Cherenkov angle (θ_h^r) and width σ_h^r (due to uncertainty in photon emmission position) for a track of particle type h is used to describe Cherenkov angle distribution (i.e. θ). To obtain the number of photons emited from the aerogel layer ($N^{s,r}$) a general expression is used for the Cherenkov photon yield, where the track path length in the aerogel, Rayleigh scattering, and the photon loss on the edges of aerogel tiles are taken into account. A constant (pixel-independent) value

is assumed for the background contribution n_i^b, set to correctly describe the observed distribution.

The expected total number of hits, N, is obtained as $\epsilon_{det}\epsilon_{acc}(N^{s,1} + N^{s,2})$, where ϵ_{acc} is the geometrical acceptance correction factor (i.e. what fraction of the Cherenkov ring falls on the photo-sensitive surface). The acceptance correction factor is calculated using a simple ray tracing simulation in which 200 rays, uniformly distributed in ϕ and at the expected Cherenkov angle θ_h^r, are propagated from the mean emission point in the aerogel to the detector plane. The number of track lines that hit the photo-sensitive surface is used to determine the correction factor.

The above procedure is carried out for all 6 particle hypotheses. The log-likelihood difference, $\ln \mathcal{L}_{h_1} - \ln \mathcal{L}_{h_2}$, is used to distinguish h_1 and h_2 particle types. On average about 12 signal photons are detected per saturated track. The PID performance is mainly degraded by tracks with poor track information (position and direction on the aerogel plane), which result either from poor reconstruction or rescattering in the CDC aluminum endplate, and by tracks that produce a very low number of photons. The latter are mainly tracks that pass through the gap between two adjacent aerogel tiles, or tracks producing a Cherenkov ring that largely misses the photosensitive area.

The PID performance of the ARICH detector for K/π separation is depicted in Fig. 1.8.

1.4.3 Muon identification

Muon identification in the KLM uses the differences in longitudinal penetration depth and transverse scattering of the extrapolated track. The likelihoods for each of the six particle hypotheses are assigned to a given track and stored as log-likelihood values in the `Muid` data-object. In the post-reconstruction analysis, the log-likelihood differences may be used to select or reject the muon hypothesis for a given track.

The KLM geometry (Fig. 1.9) exhibits several features: the barrel has 15 detector layers with no iron before the innermost layer; the forward (backward) endcap has 14 (12) detector layers with iron before the innermost layer. The iron plates are about 4.7 cm thick and are separated by detector-filled 4.4-cm gaps. The KLM has less iron and detector coverage in the forward and backward overlap regions since the endcaps' outer radius is about 310 cm: there may be as few as 8 detector layers for some polar angles. Thus, the separation power between muons and non-muons is weaker here.

Track Extrapolation

Each charged track that is reconstructed in the tracking detectors (CDC, SVD and PXD) is extrapolated outward using `geant4e` [11], starting at the last point on the reconstructed track in the CDC, assuming the muon hypothesis. During this extrapolation, `geant4e` reduces the track's momentum by the mean specific-ionization (dE/dx) energy loss in the intervening material and inflates the elements of the phase-space covariance matrix due to (elastic) multiple scattering and fluctuations in dE/dx. Particles do not decay during this extrapolation.

Extrapolation through the non-KLM sections by `geant4e` does not consider actual hits in any of the sensitive elements. In contrast, extrapolation through the KLM uses each matching hit in a Kalman-filtering adjustment of the track parameters and covariance matrix [12].

The `muid` reconstruction module in the tracking package of `basf2` proceeds in two steps: (1) track extrapolation using the muon hypothesis and (2) likelihood extraction for each of the six particle hypotheses.

The extrapolation proceeds step by step through the detector geometry, starting at the outermost point on the reconstructed-track's trajectory (usually in the CDC) and with the reconstructed track's phase-space coordinates and covariance matrix. Upon crossing a KLM detector layer, the nearest two-dimensional hit—if any—in that layer is considered for association with the track. If the hit is within about 3σ in either of the two local-coordinate directions (where σ is the sum in quadrature of the extrapolation's position uncertainty
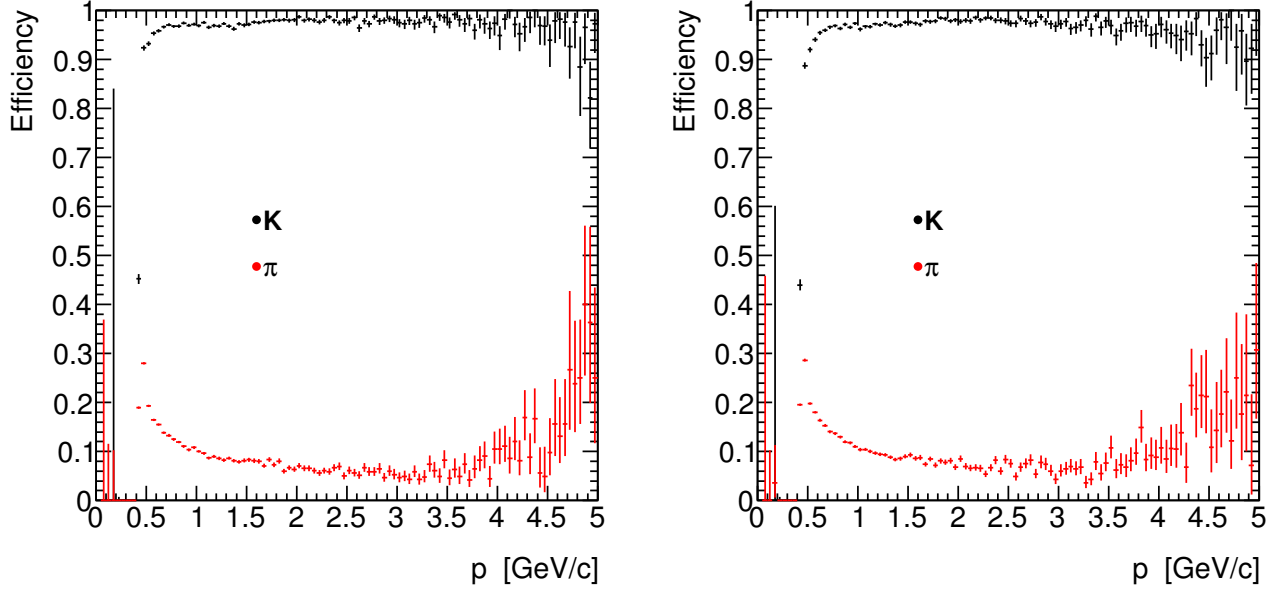
Figure 1.8: ARICH kaon identification efficiency (black markers) and pion fake rate (red markers) as a function of momentum for $\ln \mathcal{L}_\mathcal{K} > \ln \mathcal{L}_\pi$, obtained with a MC simulation without beam background (left) and with beam background (right). Only tracks that fall within the ARICH acceptance are considered.

and the hit-measurement uncertainty) then it is declared a matching hit and the Kalman filter uses it to adjust the track properties before the next step in the extrapolation. At the same time, the Kalman filter's fit quality (χ^2) is accumulated for the track. A given 2D hit may be associated with more than one track. The hit-matching algorithm ratchets outward so that a given KLM detector layer is examined at most once for a given track. The extrapolation ends when the kinetic energy falls below a user-defined threshold (nominally 2 MeV) or the track curls inward to a cylindrical radius below 40 cm or the track escapes from the KLM.

Determination of Likelihoods

If the track reached the KLM, it is classified according to how and where the extrapolation ended (stopped or exited and in the barrel or the end-cap). The likelihood of having the matched-hit range and transverse-scattering χ^2 distribution is obtained from pre-calculated probability density functions (PDFs). There are separate PDFs for each charged-particle hypothesis and charge and for each extrapolation outcome.

The longitudinal-profile PDF value $\mathcal{P}_L(\vec{x};\ O,\ \ell,\ H)$ for extrapolation-ending outcome O and outermost layer ℓ and for particle hypothesis H is sampled according to the measurement vector $\vec{x}$ given by (a) the pattern of all KLM layers touched during the extrapolation—not just the outermost such layer—and (b) the pattern of matched hits in the touched layers. Note that the momentum and direction are not included in the measurement vector. If $p_k(O,\ \ell,\ H)$ is the probability of a matching hit in touched layer k, then

$$\mathcal{P}_L(\vec{x};\ O,\ \ell,\ H) =$$

$$\prod_k \begin{cases} p_k(O,\ \ell,\ H) & \text{matching hit in layer } \ell \\ 1 - p_k(O,\ \ell,\ H) & \text{otherwise} \end{cases}$$

$$(1.12)$$

The extrapolation outcome O accounts for the KLM geometry by classifying a track as stopping in or exiting (1) the barrel only, (2) the forward endcap only, (3) the backward endcap only, (4) the overlap region between the forward barrel and endcap, and (5) the overlap region between the backward barrel and endcap. The overlap regions are fine-grained, with a distinct outcome assigned for a given barrel layer that is the outermost before the transition to the endcap. There are 66 possible outcome values: (2 for stop vs exit) $\times$ (15 $\times$ 2 for the overlap regions + 3 for pure barrel/forward/backward). The outermost extrapolation layer ℓ is an indirect proxy for the track's momentum and polar angle. Sample PDFs for exiting tracks are shown in Fig. 1.10 for muons and pions. To avoid artifacts due to statistical fluctuations, the raw histograms (points with error bars) are fitted with higher-statistics proxies; the purple histograms represent the actual PDFs. There are 990 such PDFs tabulated for each particle-and-charge hypothesis.

The transverse-scattering probability density function $\mathcal{P}_T(\chi^2, n; D, H)$ for KLM region D (barrel-only, endcap-only, or overlap) and particle hypothesis H is sampled according to the measurements of χ^2 from the Kalman filter and the number $n \in \{2, 4, ..., 36\}$ of degrees of freedom—twice the count of matching-hit layers. The muon-hypothesis PDF is very close to the ideal χ^2 distribution for the given number of degrees of freedom while the non-muon-hypothesis PDFs are considerably broader for low degrees of freedom—the most likely scenario for a true non-muon. The probability density function is tabulated as a reduced-χ^2 histogram up to $\chi^2 = 10$ in bins of 0.1. For χ_r^2 values below a tabulated cutoff for a given PDF, a spline interpolation of the histogram is used for the PDF value. For χ_r^2 values above this cutoff, an ad-hoc fitted function is used: this is the ideal χ^2 distribution for n degrees of freedom but with a floating horizontal-axis scaling factor and a floating normalization. The function's parameters are determined from the histogram's upper tail so that the histogram and ad-hoc function join smoothly at the cutoff. Sample PDFs are shown in Fig. 1.11 for muons

and pions. There are 54 such PDFs tabulated for each hypothesis H: (3 detector regions) $\times$ (18 d.o.f. values).

For each track, the likelihood for a given particle hypothesis is the product of the corresponding longitudinal-profile and transverse-scattering PDF values:

$$
\begin{aligned}
\mathcal{L}(H;\, O,\, \ell,\, D,\, \vec{x},\, \chi^2,\, n) = \\
\mathcal{P}_L(\vec{x};\, O,\, \ell,\, H) \cdot \mathcal{P}_T(\chi^2,\, n;\, D,\, H)\,.
\end{aligned}
\tag{1.13}
$$

The natural logarithm of this value is stored in the `Muid` data-object. Then, the six likelihood values are normalized by dividing by their sum and stored in the `Muid` data-object.

Presently, significance levels are not available. Such values might be used, for example, to remove tracks that are not consistent with *any* hypothesis by requiring $\mathcal{S} > \mathcal{S}_{\min}$. This feature will be added in a future release.

Muon Efficiency and Pion Fake Rate

The log-likelihood difference

$$
\begin{aligned}
\Delta \equiv{}& \log(\mathcal{L}(\mu^+; O,\, \ell,\, D,\, \vec{x},\, \chi^2,\, n)) \\
&- \log(\mathcal{L}(\pi^+; O,\, \ell,\, D,\, \vec{x},\, \chi^2,\, n))
\end{aligned}
\tag{1.14}
$$

is the most powerful discriminator between the competing hypotheses. The requirement $\Delta > \Delta_{\min}$ for a user-selected $\Delta_{\min}$ provides the best signal efficiency for the selected background rejection.

Log-likelihood differences for true muons and pions are shown in Fig. 1.12 as a function of the track momentum. Clearly, choosing a momentum-independent cut $\Delta_{\min}$ that is non-zero (and positive) will reject soft muons preferentially. Similar behavior is seen when choosing a cut that is independent of the polar or azimuthal angles since the log-likelihood differences are softer in the uninstrumented azimuthal cracks between sectors and in the barrel-endcap overlap regions where the KLM is thinner (with only 8 or so detector and iron layers).

Muon efficiency and pion fake rate (scaled up by 10) are shown in Fig. 1.13 as functions of momentum, polar angle, and azimuthal angle for three values of the log-likelihood-difference threshold. The

black curves exhibit the behavior for the nominal cut of $\Delta > 0$: the muon efficiency is 90–98% for momenta above 1.0 GeV/c while the pion fake rate is 2.5–6%; the muon efficiency is flat at 96% in θ while the pion fake rate is 2–6%; the muon efficiency is 92–98% in ϕ (with dips at each octant boundary and at the solenoid chimney) while the pion fake rate is roughly flat at 3.5% (or 4% at the chimney). The red curves exhibit more pronounced differences as a function of p, θ and ϕ for the much tighter cut of $\Delta > 20$, where muon efficiency is sacrificed somewhat—and unevenly in each of these variables—for much better purity.

1.4.4 Electron identification

E/p

1.4.5 Combined PID performance

The performance of Belle II PID is estimated using inclusive $c\bar{c}$ MC samples. Minimal track quality restrictions are applied. Using the generated information, a sample of each particle type is constructed. The PID efficiency for a sample of particles of type α is determined by taking the ratio of events that have $L(\alpha : \beta) > 0.5$ to the total sample size, for a given β. For example, the K/π selection efficiency is given by the fraction of a sample of true kaon tracks that have $L(K : \pi) > 0.5$. In a similar fashion, the pion fake rate is the fraction of a sample of true pion tracks that have $L(K : \pi) > 0.5$. The selection efficiency for various pairs of particle types are given in Fig. 1.14. Only tracks that fall within the acceptance of at least one of the PID detectors or the CDC are considered.

1.5 Neutral particle identification

1.5.1 Photon and π^0 identification

The identification of photons in the ECL is based on parameters that describe the shower shape of ECL clusters that are not matched to a reconstructed track. The identification relies on the fact that electromagnetic showers caused by an incident photon is cylindrical symmetric in the lateral direction and the energy deposition decreases exponentially with distance from the incident axis. The original ECL reconstruction provides only the energy ratio of the nearest 3×3 to the nearest 5×5 crystals around a local maximum which is close to unity for true photons. The main background for photon cluster comes from neutral or charged hadron interactions. These interactions create asymmetric shower shapes and often result in more than one ECL cluster that are not matched to charged tracks, so called hadronic splitoffs, which yield a large number of fake photon candidates if not identified. The minimal energy of an ECL cluster for physics studies in the presence of nominal backgrounds is 100 MeV in the forward endcap, 90 MeV in the barrel and 160 MeV in the backward endcap for the original ECL reconstruction. The rewritten ECL reconstruction will also provide additional shower shape parameters and the improved clustering algorithm will allow to lower the energy threshold to (30–50) MeV.

The reconstruction of π^0s from $\pi^0 \to \gamma\gamma$ is based on the combination of two photon candidates. For π^0 energies below about 1 GeV the angular separation between the two photons is usually large enough to produce two non–overlapping ECL clusters. For π^0 energies above about 1 GeV but below about 2.5 GeV, the ECL clusters from the two photons overlap but can still be reconstructed as two separate photon candidates in the ECL. The π^0 energy can be directly reconstructed from the photon 4-momenta. The π^0 energy resolution is improved by performing a mass constrained fit of the two photon candidates to the nominal π^0 mass. A low photon energy threshold is mandatory to obtain a high π^0 efficiency for generic B decays: A 50 MeV threshold for both photons results in a π^0 efficiency of 76 %, 30 MeV in 93 % and 20 MeV in 98 %.

For π^0 energies above about 2.5 GeV, e.g. from $B \to \pi^0\pi^0$, the two photon induced showers often do not have separate local maxima anymore and are reconstructed as one photon candidate. The
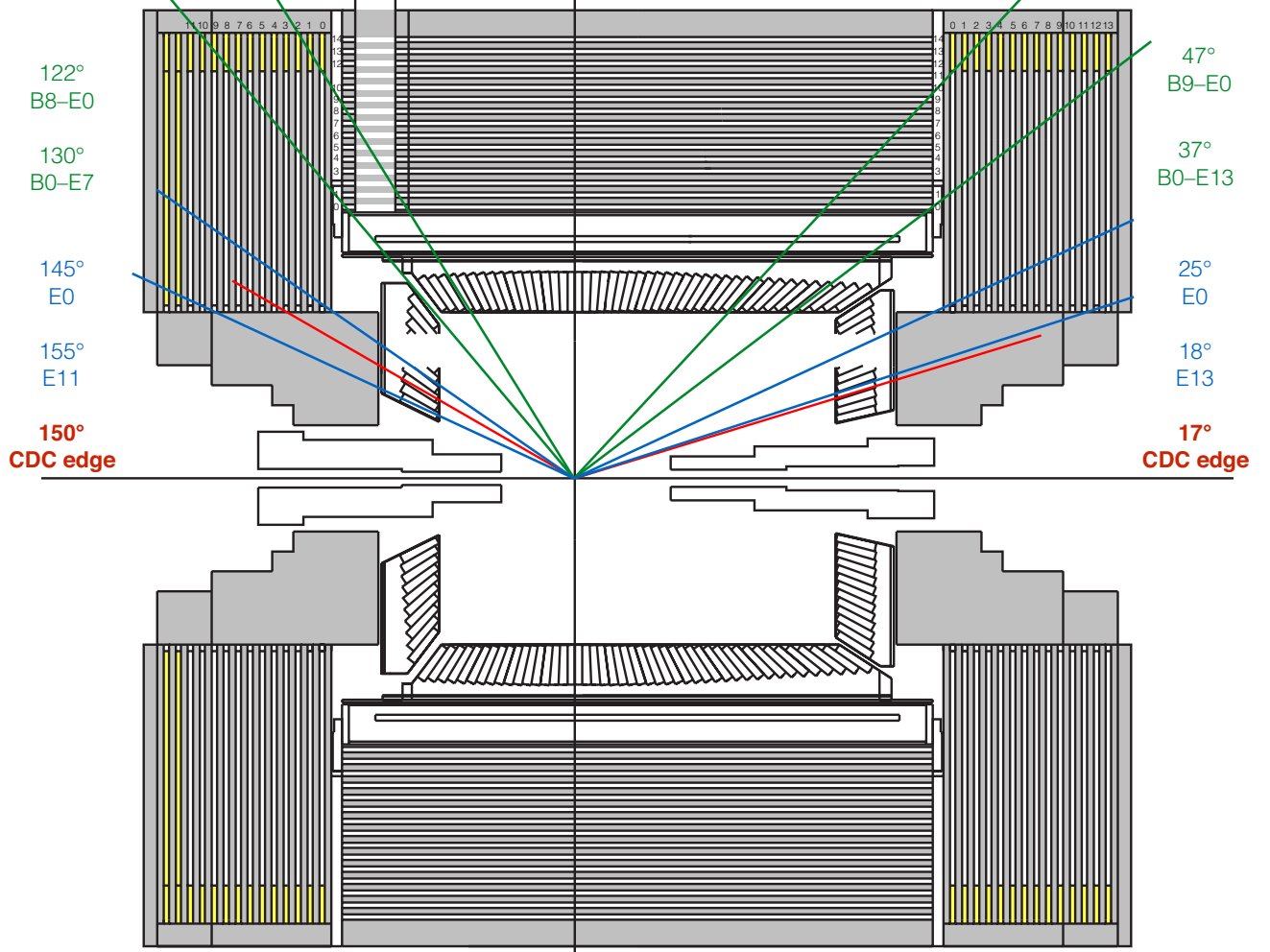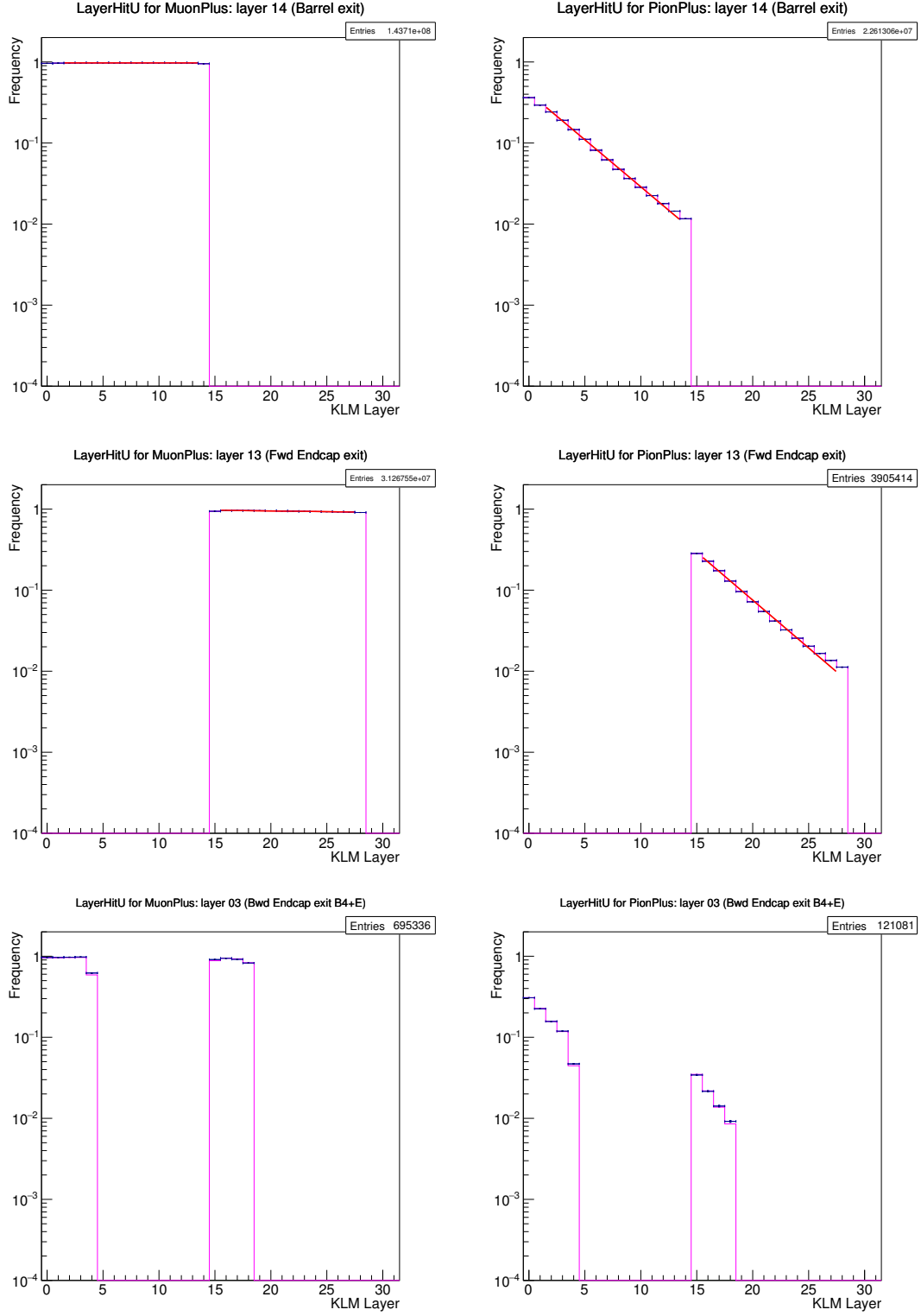
Figure 1.9: *Side view of the Belle II KLM. Within the CDC polar angular coverage of $17° < θ < 150°$, the KLM is subdivided into several angular ranges: $18° < θ < 25°$ and $145° < θ < 150°$ for partial (outer layers only) endcap coverage; $25° < θ < 37°$ and $130° < θ < 145°$ for full endcap coverage; $37° < θ < 47°$ and $122° < θ < 130°$ for barrel-and-endcap overlap; and $47° < θ < 122°$ for barrel-only coverage. The yellow-shaded parts of the endcaps are not instrumented.*

Figure 1.10: *Sample longitudinal-profile PDFs for energetic positively-charged muons (left) and pions (right) for the barrel (top), the forward endcap (middle) and a selected barrel-endcap overlap (bottom). The purple histogram represents the PDF. Barrel (endcap) layers are numbered 0–14 (15–28).*
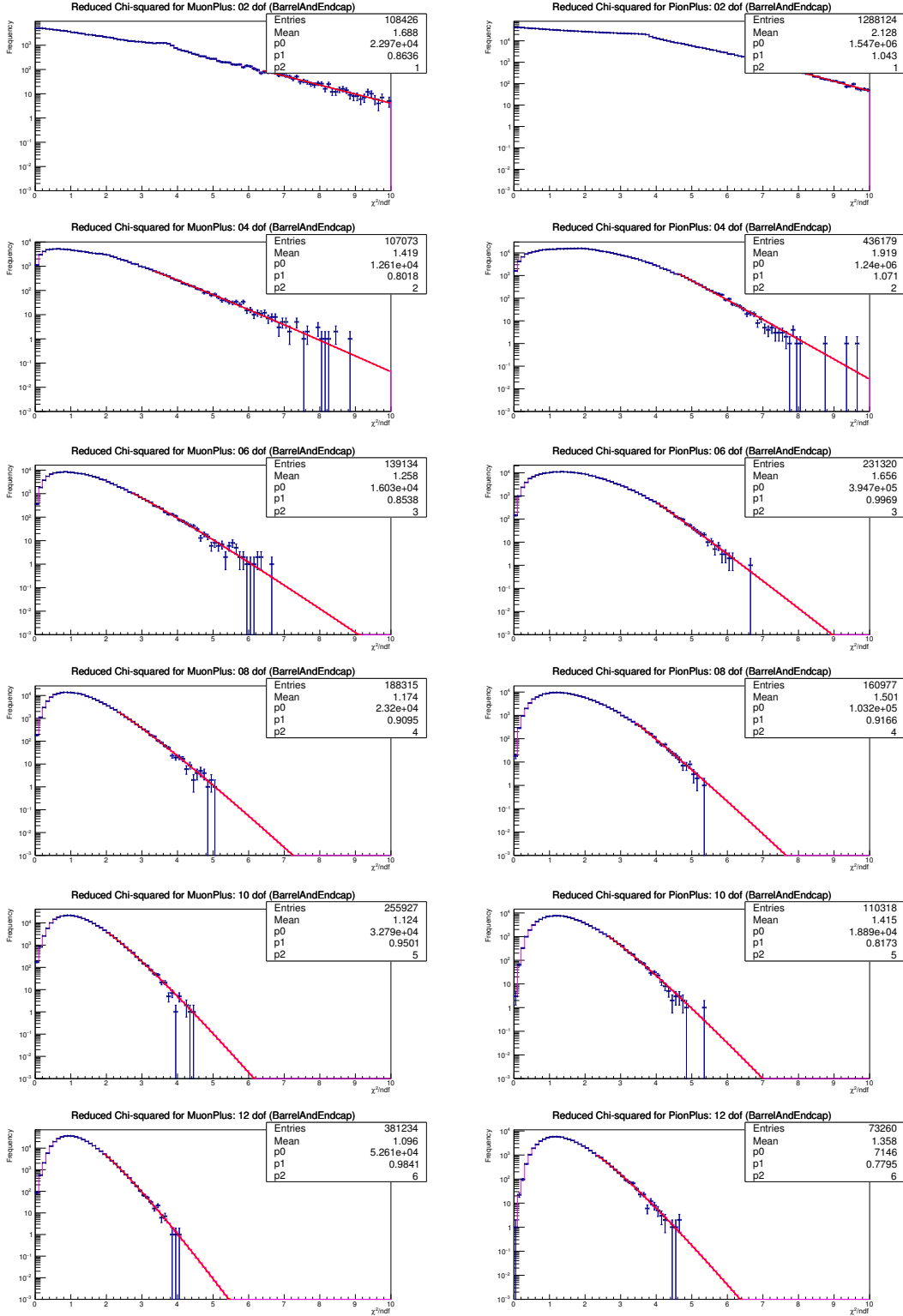
Figure 1.11: *Sample tranverse-profile (reduced χ^2) distributions for positively charged muons (left) and pions (right) for 2–12 degrees of freedom. In each panel, the red curve is the fit to the upper tail of the histogram, starting at the given cutoff (e.g., 6.4 for 2 degrees of freedom).*
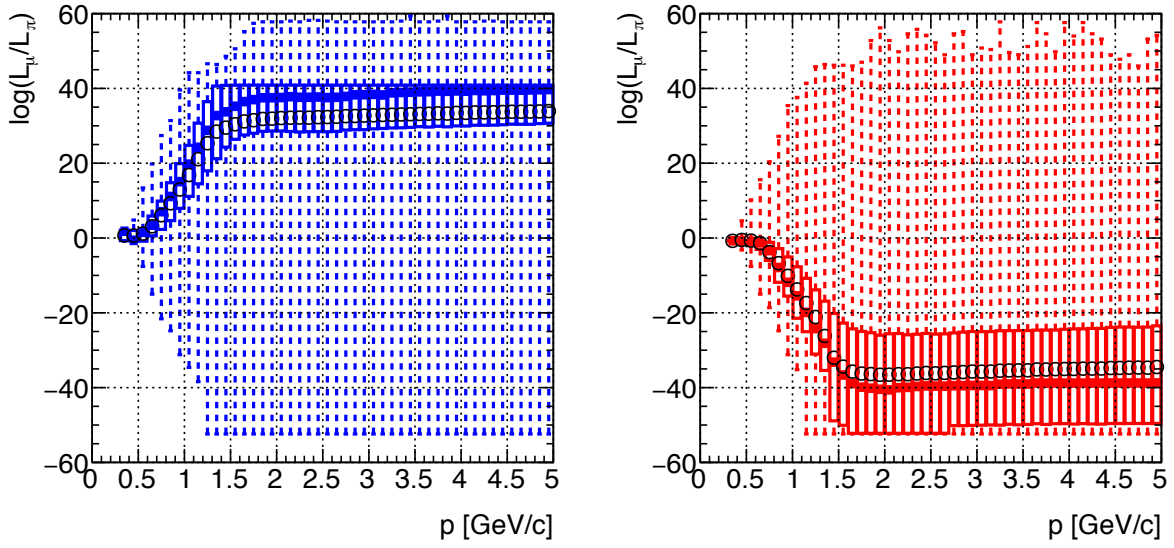
Figure 1.12: *Log-likelihood difference between muon and pion hypotheses for true muons (left) and true pions (right) as a function of the track momentum in GeV/c. In each bin, the "candle plot" shows five features: minimum and maximum values (bounded by the dashed vertical line); the lower and upper quartiles (below or above the rectangular box); the median (the thick horizontal line segment); and the mean (the circle). The mean [or median] value deviates increasingly from zero with rising momentum (i.e., with increasing number of crossed KLM layers) and then saturates at about ± 40 for exiting tracks.*
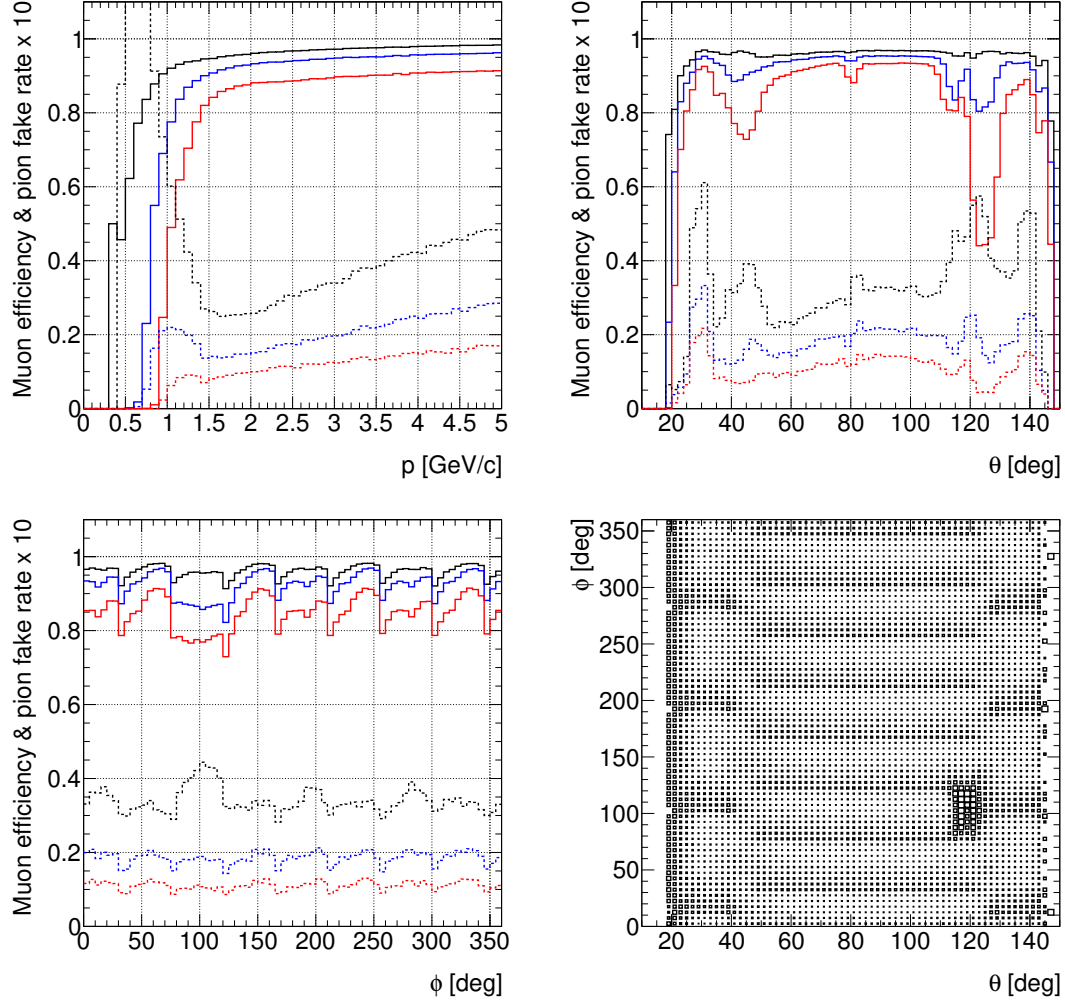
Figure 1.13: *Muon efficiency (solid) and pion fake rate scaled by 10 (dashed) for three values of the log-likelihood-difference cut:* $\Delta_{\min} = 0$ *(black)*, *10 (blue)*, *and 20 (red) as a function of momentum (top left), polar angle (top right), and azimuthal angle (bottom left). Muon inefficiency as a function of ϕ vs θ (bottom right), illustrating the geometric inefficiencies at the sector boundaries (8 horizontal enhancements in the barrel; 4 horizontal enhancements in each endcap) and in the vicinity of the solenoid chimney.*
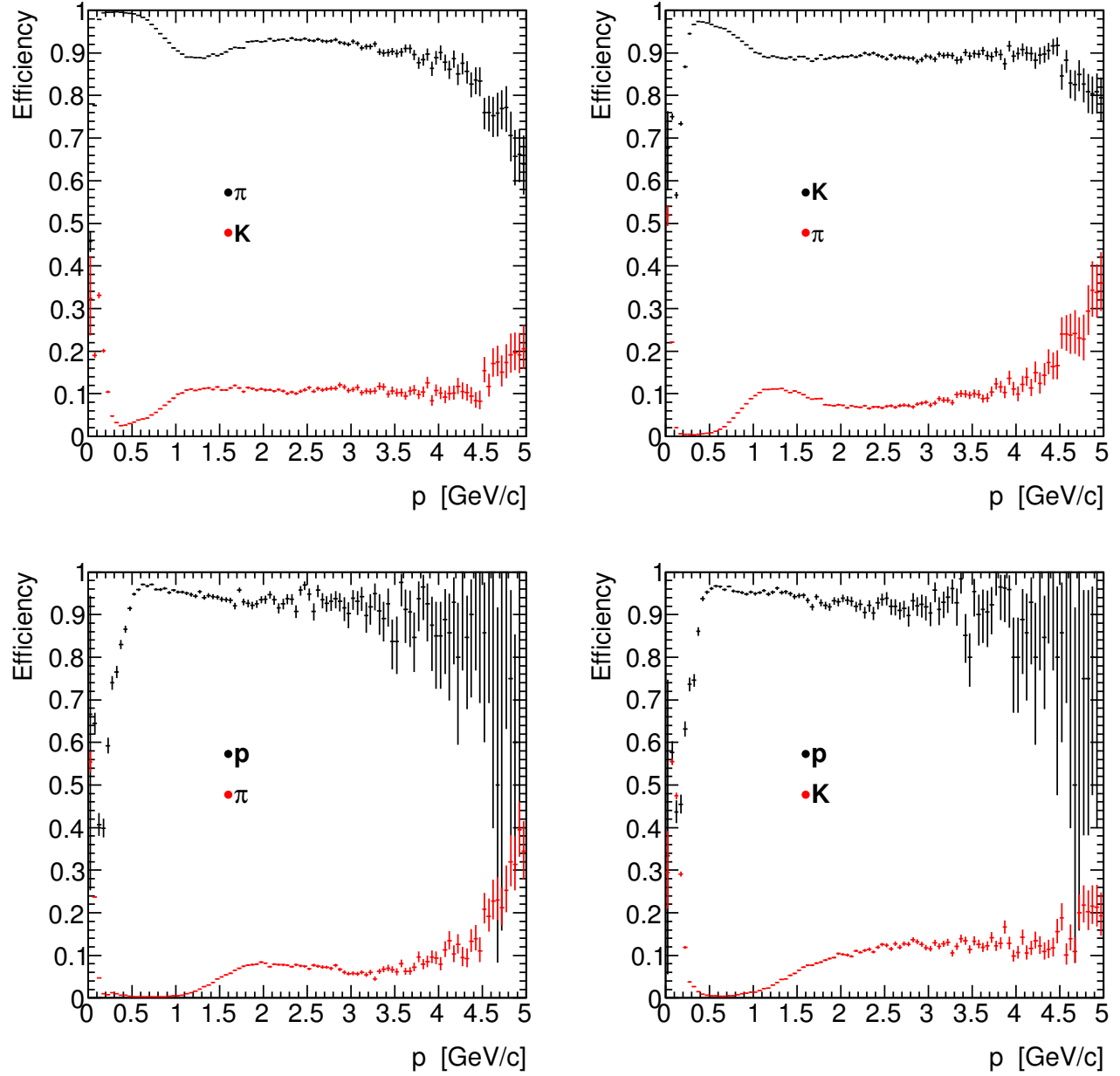
Figure 1.14: Charged particle identification selection efficiency for various pairs of particle types as a function of momentum. The black markers show the selection efficiency as determined from an inclusive MC sample without beam backgrounds, while the red markers show the fake rate. Only tracks that fall within the acceptance of at least one of the PID detectors or the CDC are considered.

π^0 energy can still be deferred from shower shape variables that will be included in the new ECL reconstruction. Add plot: pi0 purity vs efficiency for new clustering.

1.5.2 K_L^0 identification

The identification of K_L^0 mesons is based on information collected by the KLM and ECL detector. The detector material of the KLM provides > 3.9 hadronic inter action lengths λ_0 and the ECL provides $\approx 0.8\ \lambda_0$.

Multivariate methods are used to classify ECLClusters and KLMClusters according to their probability to originate from a K_L^0.

Classifiers are constructed from Stochastically Gradient Boosted Decision Trees (BDT), implemented as described in [13] and [14]. The classification is performed separately for the KLM and ECL in the reconstruction package of basf2. The resulting probabilities are saved in a KlId data-object.

The classifiers are fitted on a K_L^0 target and their output is normalized to $x \in (0, 1)$. The purity of the classification is low due to high background levels and weak K_L^0 signals, which are not significantly different than those caused by other neutral particles like neutrons, photons (in the ECL) or π^0s. In both detectors, more than 30% of the clusters originated from beam background, making it the dominant background contribution. The quality of the beam background simulation is unknown at the time of writing.

Variables used in the classifications

The classifiers are fed with all information that is available, including cluster shapes, kinematic variables, and information gained from other detectors and algorithms. There is no single variable that provides significant separation power alone. The three most important variables that have the best proven separation in the KLM are:

- **distance to the next track extrapolation:** This variable is calculated in the muid module. Neutral clusters are not likely to have a nearby track.

- **cluster timing:** Fake clusters from beam backgrounds are not in time with the collision of interest.

- **number of inner most layers hit:** Hadronic clusters are likely to have a wider radius than electromagnetic clusters.

In total, 23 variables are fed to the classifier.

In the ECL the most significant variables are:

- **distance to the next track extrapolation:** Neutral clusters should be unlikely to originate from a track.

- **E9oE25:** Energy in the central 9 crystals divide by the energy in the outer 25 crystals of a cluster. The shape can tell if its a hadronic or electromagnetic cluster.

- **energy deposition in the cluster:** Each K_L^0 deposits very little energy in the clusters, typically in the < 50 MeV range.

The ECL software is being re-factored at the time of writing. The objective of this is to perform clustering under different particle hypotheses. Therefore the architecture of the ECL classifier might change. It may be possible to implement multiple parallel or nested classifiers, depending on what gives the best efficiency-purity behavior.

Performance of the classifications

The classifier performances are evaluated on generic $\Upsilon(4S)$ events using a version of software more recent than release-00-07-00 (June 2016). The purity on predetermined selections therefore might differ. The efficiency-purity behavior of the classifiers is depicted in Fig. 1.15.

Fake rates and efficiencies of the classifications depend on the chosen working point (cut). The cut is arbitrarily chosen to yield $\approx 10\%$ purity in the ECL and $\approx 30\% - 40\%$ purity in the KLM. The optimal cut value depends on the desired performance and type or size of background. The fake rates and

efficiencies for the arbitrary cut described above are depicted in Fig. 1.16. The fake rate (1−purity) and efficiency are correlated by the shape of the ROC curve (Fig. 1.15).

In general the K_L^0 classification performance depends on the background level and composition, the magnetic field map, and the tracking performance, which might still change compared to the current estimates.

Bibliography

[1] V. Blobel and C. Kleinwort, *A New Method for the High-Precision Alignment of Track Detectors*, arXiv:0208021 [hep-ex].

[2] V. Blobel, *Software alignment for Tracking Detectors*, Nucl. Instrum. Meth. **A566** (2006) 5–13.

[3] C. Kleinwort, *General Broken Lines as advanced track fitting method*, Nucl. Instrum. Meth. **A673** (2012) 107–110.

[4] J. R. et al., *GENFIT - a Generic Track-Fitting Toolkit*, arXiv:1410.3698 [hep-ex].

[5] M. Staric, K. Inami, P. Krizan, and T. Iijima, *Likelihood analysis of patterns in a time-of-propagation (TOP) counter*, Nucl. Instrum. Meth. **A595** (2008) 252–255.

[6] M. Staric, *Pattern recognition for the time-of-propagation counter*, Nucl. Instrum. Meth. **A639** (2011) 252–255.

[7] K. e. a. Nakamura, Particle Data Group J. Phys. G **37** (2010) 075021.

[8] M. Staric, *Performance studies of the Belle II TOP counter*, Nucl. Instrum. Meth. **A766** (2014) 237–240.

[9] P. Baillon, *Cherenkov ring search using a maximum likelihood technique*, Nucl. Instrum. Meth. **A238** (1985) 341–346.

[10] R. Forty, LHCb, *RICH pattern recognition for LHCb*, Nucl. Instrum. Meth. **A433** (1999) 257–261.

[11] S. Agostinelli et al., GEANT4, *GEANT4: A Simulation toolkit*, Nucl.Instrum.Meth. **A506** (2003) 250–303.

[12] R. Frühwirth, *Application of Kalman Filtering to Track and Vertex Fitting*, Nucl.Instrum.Meth. **262** (1987) 444.

[13] J. H. Friedman, *Greedy function approximation: a gradient boosting machine*, Annals of statistics (2001) 11891232.

[14] J. H. Friedman, *Stochastic gradient boosting*, Computational Statistics & Data Analysis **38** (2002) no. 4, 367378.
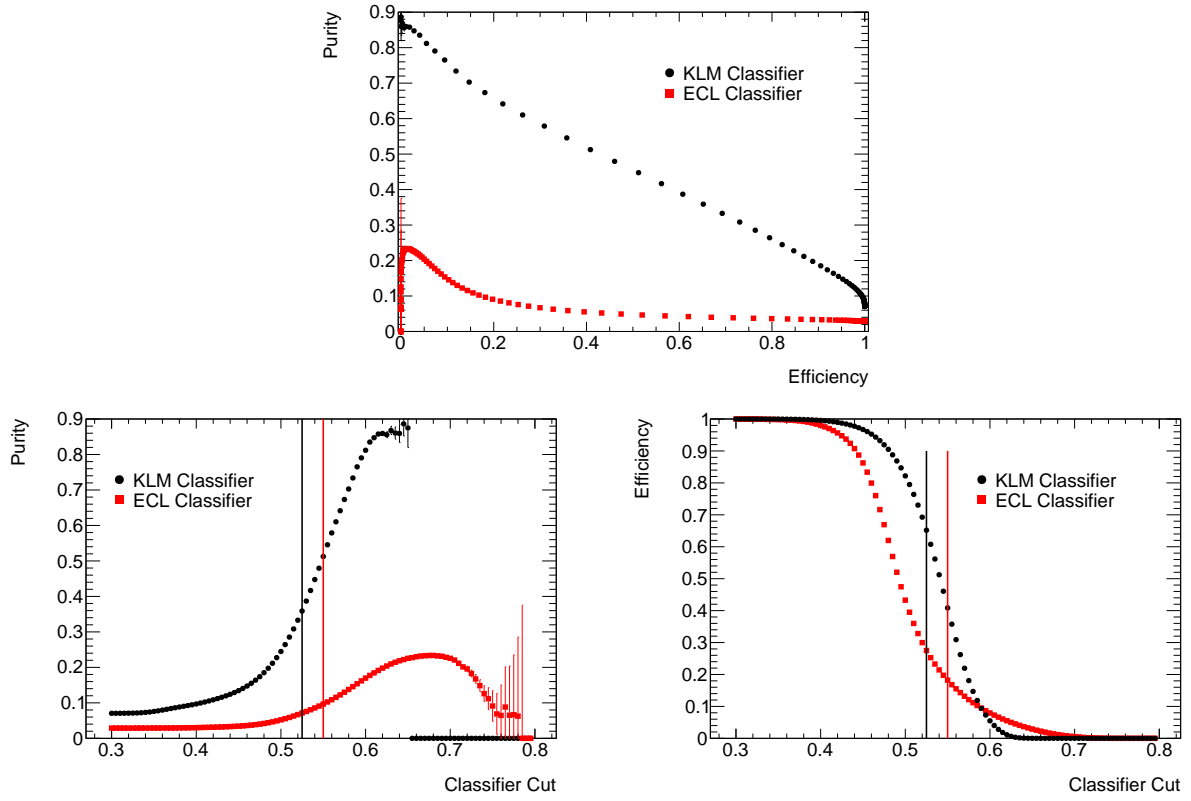
Figure 1.15: *a) Receiver operating characteristics (ROC) of the KLM (black) and ECL (red) classifier. b) Cut value vs. purity. The cuts chosen for efficiency and fake rate analysis are depicted with straight lines. c) Cut value vs. efficiency.*
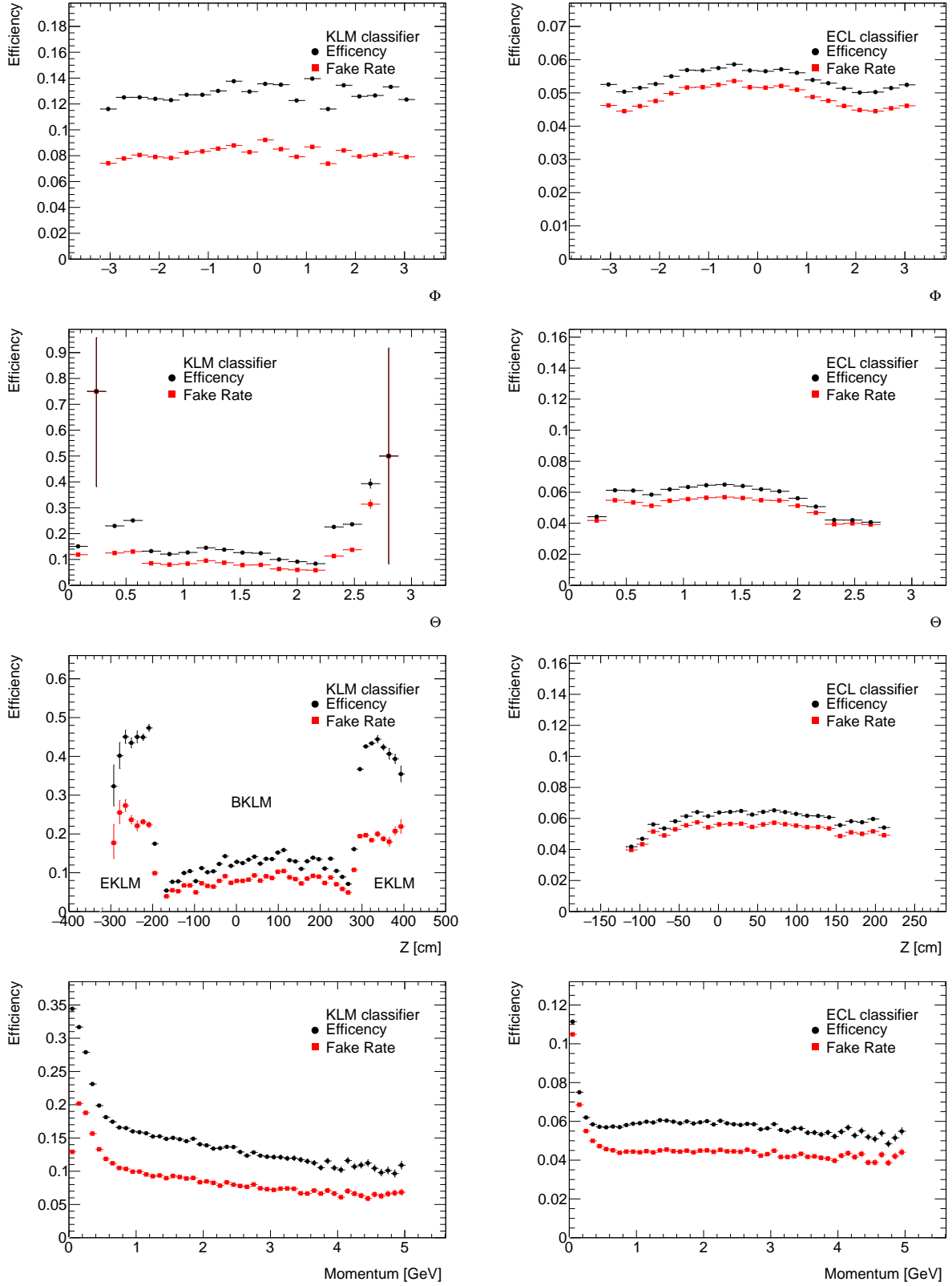
Figure 1.16: K_L^0 efficiencies (black) and fake rates (red). For a fixed cut value on the classifier. The left side depicts the KLM classifier. The right side shows the ECL classifier. The cut values are arbitrarily chosen, see Fig. 1.15. The first two rows show efficiency and fake rate in ϕ and θ, the last two rows in Z and momentum.