CDF/ANAL/BOTTOM/CDFR/9592
Version 1.1
May 27, 2009

Extended and improved track-based dE/dx calibration for 3/fb analyses

F. Azfar, L. Oakes
Oxford University

D. Tonelli
Fermilab

Abstract

We improved and extended to the 3/fb sample the calibration of the ionization energy-loss for unit length in the COT. Using pure samples of charged kaons and pions from D^{*+}–tagged $D^0 \to K^-\pi^+$ decays, we calibrated the dE/dx gain against variations depending on the position of the particle's trajectory in the tracking volume and environmental conditions.

1 Introduction

Several key flavor-physics analyses at CDF exploit the particle identification capabilities of the detector. Examples include the analysis of charmless two body b–hadron decays, the searches for FCNC in $B^0_{(s)} \to \ell^+\ell^-$ and $B^0_{(s)} \to e\mu$ decays, the measurement of charm mixing, the B^0_s mixing measurement, and the measurement of mixing-induced CP-violation in $B^0_s \to J/\psi\phi$ decays.

However, identification of charged hadrons in CDF is difficult, since the detector was mostly designed for high-p_{T} physics measurements and not optimized for these capabilities. The TOF is the only subdetector entirely devoted to this function, but its performances are marginal for particles with momenta greater than 2 GeV/c and in high-luminosity collisions. Similarly, specific ionization from the silicon tracker is of little help, because its identification power is only effective for particles with $p_{\mathrm{T}} \lesssim 800$ MeV/c. Conversely, for charged particles of $p_{\mathrm{T}} > 2$ GeV/c, the identification from the rate of energy loss through ionization (dE/dx) in the gas that fills the active volume of the drift chamber is effective. To a good approximation, the most probable dE/dx value for a charged particle (other than electron) is a function of its velocity [1]; thus, if the momentum of the particle is measured, the mass can also be determined. In the COT, the signal induced on each sense-wire depends on the amount of ionization charge produced by the passage of the charged particle near the wire. It is measured in nanoseconds because it is (logarithmically) encoded as the digital pulse-width between the leading- and the trailing-edge times of the hit. Multiple samplings along the trajectory of the charged particle allow a more reliable estimation of dE/dx, which has usually a broad distribution. The COT samples a maximum of 96 dE/dx measurements per track, from which a 80% truncated mean is calculated to avoid the adverse effect of long positive tails in the estimation of the average dE/dx.

In the first two years of data-taking, an unexpected wire-aging was observed, and the effect was severe enough to significantly degrade the performance of the chamber. Gain losses of up to 50% were observed as a function of time and location in the chamber. The aging was attributed to polymer build-up in the avalanche, which grew in a strong radiation environment and deposited on sense wires. The importance of the effect depended on local temperature and gas-flow patterns. The wire coating was identified as being mostly composed of carbon, hydrogen and a small amount of oxygen. Eventually, it was found that operating the chamber with about 100 ppm of O_2 added to the gas admixture (compared with a standard level of less than 12 ppm), reversed two years of gain loss in less than ten days [2].

Although the data collected when the COT performance was maximally compromised were excluded from physics analyses, smaller gain variations were observed throughout all data. In addition a major change in the COT read-out configuration occurred in the middle of Run II. Effective from November 28, 2006 (run 228819) the dE/dx encoding was turned-off from the two innermost superlayers (SL0 and SL1) of the COT. This aimed at increasing the number of high-quality hits attached to tracks. In fact, the dE/dx circuitry encodes the charge of the pulse in the pulse-width, yielding longer pulses when dE/dx is enabled. The longer the pulses, the higher the probability for hits of following/neighboring tracks to pile-up with the previously generated pulse, so that their leading-edges are obscured and the hits lost. This inefficiency was negligible in the lower-luminosity runs at the beginning of Run II. But starting in 2005, significantly increased initial store luminosities were exploited, and hits pile-up become a significant source of tracking inefficiency, especially in the inner layers, where higher track density is experienced.

Because of the above considerations, an accurate calibration of the uniformity of the dE/dx response in time and over the chamber volume is required. The goal of the calibration is to remove any spurious dE/dx dependence on local properties of the chamber or on the environmental conditions in which data are collected. This has the effect of optimizing the PID performance in terms of *separation power* and *correlations*.

Separation power – the capability of discriminating among different classes of charged particles is crucial because the larger the separation, the smaller the statistical uncertainty on measurements of sample composition.

Correlations – Even more critical, although often overlooked, are the correlations between dE/dx measurements. Residual gain variations cause non-zero correlations between observed dE/dx of particles in the same event, yielding biased estimates of the sample composition, if the dE/dx is applied to multiple particles per event.

This work extends to a larger sample and improves the last dE/dx calibrations, carried out in 2004 [3, 4] and updated for the B_s^0-mixing measurement. This document reports the details of the calibration procedure, the results, and the information needed by dE/dx users. Code and up-to-date information is available at http://www-cdf.fnal.gov/htbin/twiki/bin/view/TOFandDeDx/DeDxCalibration.

2 Dataset and tools

To calibrate the response to charged kaons and pions, we used a high-statistics and very pure sample of XFT-triggered kaons and pions from $D^{\star}$-tagged D^0 decays. The strong D^{*+} decay unambiguously identifies the flavor of the D^0 meson, which is reconstructed in its Cabibbo-favored decay $D^0 \to K^-\pi^+$.[1] Thus, final states always contain two like-sign charged pions and one oppositely-charged kaon, allowing collection of high-purity samples of kaons and pions from the D^0 decay. The pion originated from the D^{*+} decay is denoted as "soft pion", because it carries a much lower transverse momentum than all other particles in the decay. The low Q-value of the $D^{*+} \to D^0\pi^+$ decay provides a powerful discriminating quantity against background: $m(D^{*+}) - m(D^0) - m_\pi$. Indeed, the distribution of this variable for signal decays is much narrower (6 MeV/c^2) than the $m(D^{*+})$ distribution, because correlated systematic effects between the reconstucted masses $m(D^{*+})$ and $m(D^0)$ cancel in the difference.

We skimmed B_CHARM trigger data (in its three variants, standard, low–, and hight–p_T) from Bsntuples corresponding to the datasets xbhdid (period 0), h77jg0 (1–7), xbhdii (8–10), xbhdij (11–13) and xbhdik (14–17). The corresponding run range is 138425–261005. In sec. 10.1 is discussed the validity of the calibrations also for periods 18–21. We used the goodrun list VXXX. We apply a first basic reconstruction of signal solely based on tracking and on the information on the identity of D^0 decay-products provided by the charge of the soft pion. Some basic quality requirements were applied to all tracks of the candidate decay. One $D^0 \to K^-\pi^+$ and one $\overline{D}^0 \to K^+\pi^-$ candidate are formed for each pair of oppositely-curved tracks found in the XFT fiducial region ($|\eta| < 1$) and a kinematic fit to a common space-point (D^0 vertex) is applied. Requirements on the product of the signs of impact parameters of the tracks ($\mathrm{sgn}(d_0(K) \times d_0(\pi)) < 0$), on minimum candidate transverse decay-length ($L_T(D^0) > 300$ μm),and on its impact parameter ($|d_0(D^0)| < 140$ μm) are applied to reject $\mathcal{O}(10\%)$ contributions from background. Candidates D^0 decays are combined with a third charged particle with $p_T > 0.4$ GeV/c (soft pion candidate) and impact parameter $|d_0(\pi_s)| < 500$ μm to form a $D^{*+} \to D^0\pi^+$ candidate. The charged pion mass is assigned to the like-sign pair of particles, thus adopting the "right-sign" (Cabibbo-favored) mass hypothesis. No request of common vertex between the soft pion candidate track and the D^0 candidate tracks is applied. Due to the low resolution of soft pion's track parameters and to its trajectory being often approximately collinear with one of the other two tracks, the effect of this choice has little effect for the purpose of this work. The Q–value, i. e., the difference between the reconstructed D^{*+} mass and the sum of the nominal pion mass and reconstructed D^0 mass, is required to be within 1.5 MeV/c^2 from the nominal value to reduce backgrounds (combinations of real D^0 decays with random tracks and, to a lesser extent, random three-track combinations)

The resulting distributions in $m(K\pi)$ (which is the $K^+\pi^-$ mass of the trigger tracks

[1] The $\mathcal{O}(10^{-3})$ contamination of the doubly Cabibbo-suppressed mode $D^0 \to K^+\pi^-$ is negligible for our purposes.

when the soft pion has negative charge and their $K^-\pi^+$ mass when the soft pion has positive charge) and $Q = m(K\pi\pi) - m(K\pi) - m_\pi$, where $m(K\pi\pi)$ is the measured $K\pi\pi$ mass (with same assignments as for the measured $K\pi$ mass, $m(K\pi)$) and m_π is the known mass of the charged pion. The final calibration sample is restricted to candidates found

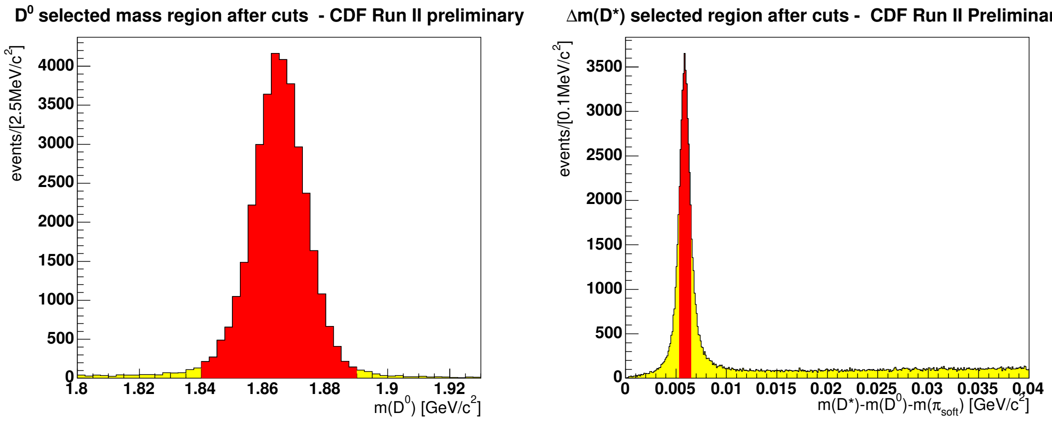

Figure 1: Invariant $K\pi$-mass distribution (a) and distribution of the difference between the reconstructed D^{*+} mass and the sum of the reconstructed D^0 mass and the nominal pion mass, $m(D^{*+}) - m(D^0) - m_\pi$ (b).

within 25 MeV/c^2 from the world average D^0 mass [5], and with $5.3 < Q < 6.5$ MeV/c^2.

It consists of XXXX $D^{*-} \to \overline{D}^0\pi^- \to [K^+\pi^-]\pi^-$ and XXXX $D^{*+} \to D^0\pi^+ \to [K^-\pi^+]\pi^+$ candidates. A D^0 signal with about 8.5 MeV/c^2 r.m.s. standard deviation is visible over a low background that includes random triplets of tracks satisfying accidentally the selection requirements (combinatorial, a smooth flat-ish background in both distributions) random tracks combined with a real D^0 decay ('fake D^{*+}', which peaks in the D^0 mass, but is a smooth structure-less background in the mass-difference distribution), $D^0 \to K^+K^-$ decays in which a kaon is mis-assigned the pion mass, and $D^0 \to K^-\pi^+\pi^0$ and $D^0 \to h^+\ell^-\bar{\nu}_\ell$ ($h = \pi$ or K, $\ell = e$ or μ) decays in which a π^0 or the (anti)neutrinos are not reconstructed, which both contaminate the lower mass-side of the signal, and $D^0 \to \pi^+\pi^-$ decays, in which a pion is mis-assigned the kaon mass in the high-mass side.

In the BSTntuples the dE/dx information is stored in the following variables:

- `double DeDxRaw()`, which is the raw dE/dx retrieved from the production bank;
- `int DeDxNHits()`, which is the number of COT hits with dE/dx information associated.

The raw dE/dx contains the value of the measured dE/dx after the truncated mean and the hit-level corrections. These are low-level calibrations performed at Production stage and correct for cell-geometry, gas pressure, and some other effects. The hit level correction constants currently present in the Production executable are based on work done only once at the beginning of Run II based on a very small portion of the data sample [6]. Because of this and of a bug in the Gen6 production that left uncorrected a large fraction of tracks, we expect to find effects caused by the incomplete correction at the hit-level. While the ideal goal is to finalize and update the hit-level calibrations before passing to the higher-level track-based work, it was decided to prioritize the development of an usable dE/dx tool for the Summer 2009 analyses and to postpone the extension of the hit-level correction to when we will move to Gen7 tracking. The BStntuples contain also the variables `CorrectedDeDx`, `PredictedDeDx`, `PredictedDeDxVSP`, `SigmaDeDx`,`PredictedZ`, `SigmaZ`, which are higher-level quantities derived from the raw ones after the old dE/dx calibrations. Since our calibrations are meant to improve the old ones using increased information and an improved procedure we do *not* use these higher-level quantities in pur corrections. We decided to determine the new set of correction constants from scratch, by starting from the raw quantities.

To simplify the study of the dE/dx dependencies we did not subtract the background, whose contamination is already reduced to $\mathcal{O}(1\%)$. Kaons (or pions) with positive and negative charges were studied separately to properly account for the effects due to the asymmetry of the COT layout, which makes the efficiency of charge-collection dependent on the sign of the track curvature.

3 Calibration procedure

We studied the observed dE/dx response as a function of a set of "macroscopic" track- or environment-related parameters such as time, track azimuthal angle, pseudo-rapidity, etc. We did not equalize the response based on lower-level COT structures such as super-layers, cells, or hit widths, as it would be optimal because this information is not available at the BStntuples level. This further step will hopefully be carried out with the next generation of tracking code. Our calibration procedure can be schematized in the following steps:

1. extensive search for possible dE/dx gain variations as functions of several environmental- and track-related quantities
2. identification of the the subset of variables $\{x_i\}$ showing the most relevant effects;
3. removal of the dE/dx dependence on each x_j variable while taking into account dependencies between variables, which may require joint correction of multiple effects;
4. extraction of the ionization curve, i. e., the function that describes the expected average dE/dx for a charged particle as a function of its Lorentz boost ($\beta\gamma$);
5. determination of the dE/dx resolution;
6. evaluation of the separation between kaons and pions;
7. extraction of the dE/dx correlations.

The crucial point is the correction (third step). The simplest case is when the gain variation is function of just one variable, say x, whose distribution in the sample is independent of the momentum distribution of the sample itself. The corrected dE/dx is simply obtained by parameterizing the dE/dx-dependence on that variable and applying the resulting 'correction function' to the uncorrected dE/dx. A minor complication is encountered when the distributions of x and momentum (p) do not factorize in the sample. In this case, if dE/dx versus x is flattened, there will be an undesired sculpting of the dE/dx versus p distribution, which should *not* be modified because this one is the desired, physical dependence that provides PID information. In this case, one flattens the x-dependent gain variation using a subsample restricted to a limited domain in momentum (a 'slice' in p), and applies the resulting correction to the whole sample.

If the dE/dx depends on multiple quantities, e. g., x_1 and x_2, the situation gets more complicated. If the dependencies are factorizable, i. e., if the uncorrected ionization can be written as $\frac{dE}{dx}(x_1, x_2) = f(x_1) \times g(x_2)$, with f and g being generic functions, then the corrected dE/dx is easily determined by first correcting the dependence on one variable and then on the other one. The worst case is when multiple dependencies can not be factorized, i. e., $\frac{dE}{dx}(x_1, x_2) = h(x_1, x_2)$. In principle, the multidimensional joint distribution of dE/dx as a function of all quantities provides all the information needed to simultaneously correct all effects. In practice, the size of the sample needed to obtain reliable corrections increases significantly with the number of effects to be corrected for. Therefore this approach is often combined with iterative corrections in a compromise between the conflicting needs for a simultaneous correction of correlated dependencies, and samples with adequate statistics. In the present case, the observed dE/dx was affected by multiple effects correlated in a non-factorizable way between them and with momentum.

We applied multiplicative, rather than additive, corrections because most dE/dx dependencies are expected to originate from geometric or gas-related effects of the chamber. They are therefore expected to influence the gain, rather than the offset, of the response.

To evaluate the performance of the corrections in an unbiased way, we divided the sample in two mutually exclusive subsamples: the *calibration* sample (90% of the events) and the *validation* sample (10% of the events). The events contained in these subsamples are uniformly sampled in time by picking one out of ten events across the whole run range. We studied the dependences, and extracted the calibrations constants, in the calibration sample and validated their effect on the validation sample.

4 Corrections

4.1 Search for effects

After an extensive scan of track- and environment-related quantities that can show abnormal effects on the dE/dx we identified the worst offenders as follows (see fig. 2).

1. A ≈ 5 ns dE/dx variation as a function of time (run number);
2. A ≈ 4 ns variation as a function of instantaneous luminosity;

3. A $\approx$ 4 ns variation as a function of the number of associated COT hits to the track;
4. A $\approx$ 3 ns variation as a function of track-density around the probed track (secance);
5. A $\approx$ 2.5 ns variation as a function of track pseudorapidity;
6. A $\approx$ 2 ns variation as a function of the azimuthal emission angle of the track (ϕ_0);

A quantitative idea of the severity of the effects can be obtained by comparing their size with 1.5 ns of typical K–π expected separation.

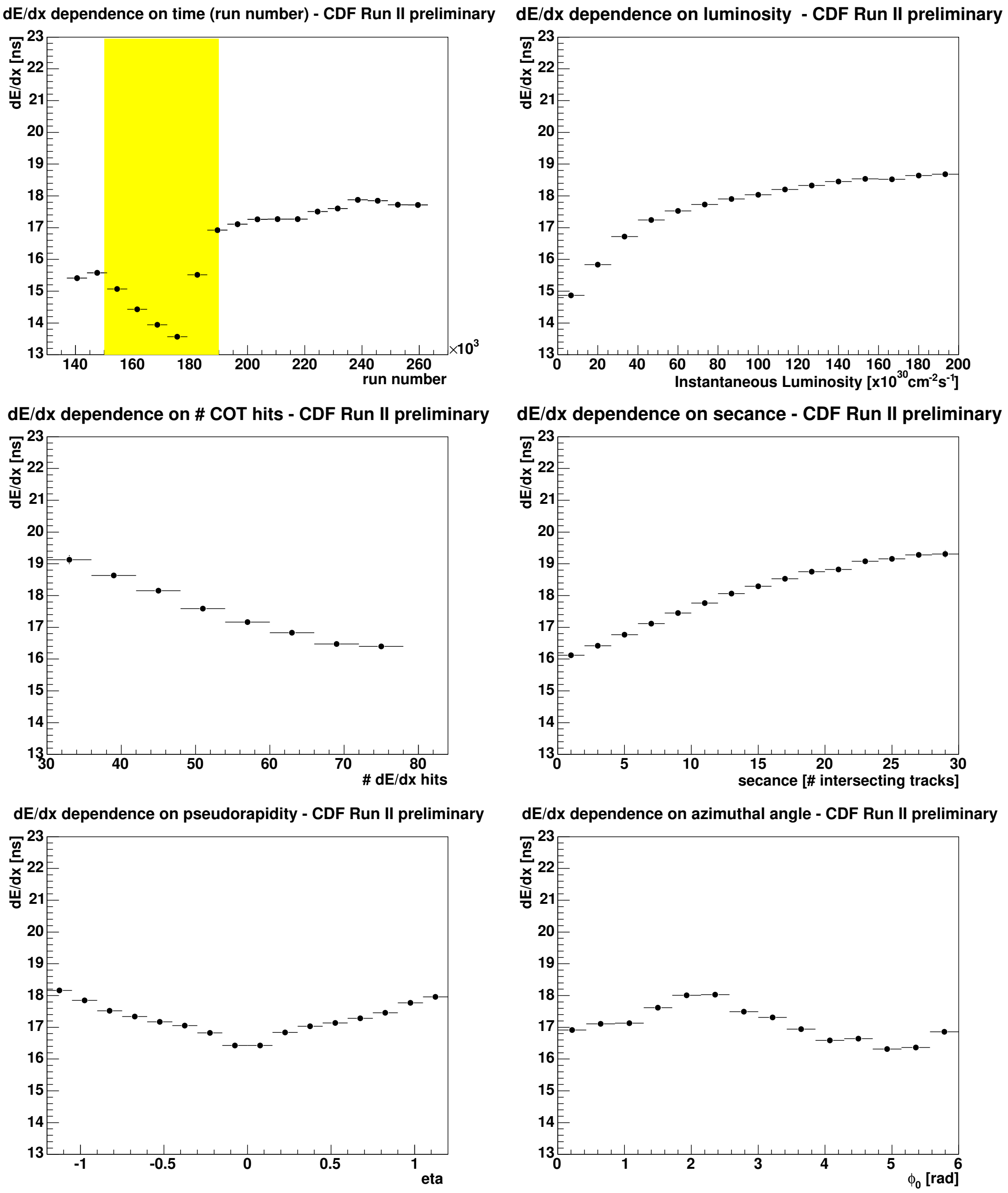

Figure 2: Profile plots of dE/dx as a function of time, instantaneous luminosity, number of associated COT hits, secance, pseudorapidity, and azimuthal angle.

Several tests were done to understand (a) which gain-variations are induced by, or correlated with, the expected momentum dependence and (b) which gain-variations can be corrected individually and which need multidimensional treatment.

4.2 Gain variations correlated with momentum

The first step is to check which of the above effects could be (at least partially) induced by the natural dependence of dE/dx on momentum. For each quantity (time, luminosity, etc.), we sliced its distribution obtaining two mutually exclusive subsamples of tracks. Then we compared the momentum spectra (fig. **??**) of the two subsamples. Any difference between momentum spectra of the two subsample would indicate a dependence between momentum and the considered variable in the sample at hand. All variables but pseudorapidity show minimal to no dependences on the momentum distribution. The observed pseudorapidty dependence on momentum is expected since in sample of particles selected by the trigger to have $p_{\mathrm{T}} > 2$ GeV/c, any particle emitted at $\eta \neq 0$ should have a p_z component that increases its momentum. The small dependences on luminosity, time, and secance are also qualitatively understood in terms of increasing contributions of triggers with higher p_{T} thresholds at increasing luminosities (and track densities).

4.2.1 Dependence on pseudorapidity (at given momentum)

It is known from the past that pseudorapidity-dependent gain variations are present. This effect is partially explained with the pseudorapidity-momentum correlation present in our calibration sample (the larger $|\eta|$, the higher the momentum). However, we observed that even by looking at dE/dx in a thin slice of momentum (see fig. **??**), a residual gain variation as a function of η is present. After checking that the pseudorapidity effect is independent of time, luminosity, secance and hits-multiplicity, we extracted a correction using the tracks with $2 < p < 3$ GeV/c and applied the resulting correction to the whole sample. The above momentum slice was chosen to be the most populated.

4.3 Gain variations independent of momentum

The dE/dx gain variations as a function of time, hit-density, hit-multiplicity and luminosity were found to be non-factorizable: the response depends jointly on all these quantities in a non trivial way. Therefore, a simultaneous multi-dimensional correction is in order to properly reduce the observed dependences. This finding is the result of a number of checks done by studying the dE/dx vs x_j distribution after having 'sliced' the sample in small intervals of all other x_i ($i \neq j$) quantities. Although a number of studies qualitatively suggested some possible causes for this complicated situation (see discussion in the following sections), further insight would call for updated hit/cell-level calibration. We therefore proceeded towards a multidimensional correction within the statistical limitation of the finite sample size we had.

4.3.1 Dependence on time

The most severe observed effect is the gain variation as a function of time. Actual variations of the working-point of the chamber (temperature, pressure...) across the data-taking period are known to impact the gain of the chamber and may be responsible for the observed drift as a function of time. For instance, starting from run 228819 the dE/dx read-out was disabled — by reducing the charge-integration time — from the two inner COT superlayers (SL1 and SL2) to facilitate tracking in high-luminosity events. The effect on the dE/dx of this change in configuration was unknown and has been one of the main motivations for this updated set of calibrations. On the other hand, time-dependent gain variations may be induced by actual dependences on third quantities (e. g., luminosity) that are themselves dependent on time. It is therefore expected that part of the time-variation will not factorize with other global effects and before correcting the time-drift we need to check its correlation with other gain-variations.

4.3.2 Dependence on hit-density

Since the COT wires have no longitudinal segmentation, any extra hit that is axially close to the track of interest can contribute to the dE/dx of the track, introducing a dependence on hit-density. These hits may or may not be associated to reconstructed tracks, but since BStntuples have only information on reconstructed tracks, we will use the density of tracks close, in the $r - \phi$ plane, to the track of interest as an estimator of hit density. This assumptions has limited validity, since it is known that hits associated to tracks can be a fraction as low as 10% of the total number of COT hits in an event, and, moreover, this fraction may depend on luminosity and sample. However, we chose to implement a correction based on track-density, and further residual effects may be accounted for by an additional, luminosity-based correction. Track-density is estimated using the secance ($\mathcal{S}$) observable, defined, for each track, as the number of axially-intersecting tracks within the COT volume (see Appendix for details). We therefore study the variations of dE/dx as a function of secance.

4.3.3 Dependence on luminosity

Variations in luminosity and consequent variations in COT occupancy are known to affect the dE/dx. The higher the charged particle multiplicity in the event, the higher the ionized charge in the chamber. Among others, this has two possible consequences: one is the formation of space charge in the gas volume, which may degrade the omogeneity of the drift field and therefore the gain of the chamber, the other is the increased probability of hits piling-up in the read-out due to high density of neighboring tracks close to the track of interest. While the latter effect was (at least partially) taken into account by the secance correction, the former effect and residual dependencies on hit-density not corrected by the secance can be corrected by studying the dE/dx as a function of luminosity.

4.3.4 Dependence on the multiplicity of associated hits

Even in the past incarnation of the calibrations, a dependence of the gain on the number of associated COT hits was found to be significant. We observed a similar behavior on the 3 fb^{-1} sample and we corrected it along with the other effects.

4.4 Dependence on azimuthal angle

Gain-variations depending on the azimuthal sector traversed by the track in the chamber have been observed since the beginning of Run II. The observed different gains for particles traversing different sectors are probably due to the interplay between the flowing-pattern of the gas within the chamber and the 2–4°C temperature gradient between the inner and outer parts of the chamber [7]. The former is adjacent to the ISL cooling system, which is kept at -10–0°C while the latter typically experiences higher temperatures. This induces a net gain variation as a function of ϕ_0 for $p > 2$ GeV/c particles, due to the small curvature of their tracks, which do not vary a lot their azimuthal direction along their path in the chamber. Indeed, we chose ϕ_0 as the quantity to investigate this effect. Although this effect is quite severe in size, it was found to be independent, modulo a global shift, of all other gain variations, and easily factorizable from the rest of the effects. This is shown in fig. **??** and fig. **??**.

5 Correction function

The final correction function has therefore the following symbolic form:

$$\boldsymbol{dE/dx = f(\phi_0) \times h(\eta|p) \times g(\mathcal{H}, \mathcal{L}, t, \mathcal{S}),} \tag{1}$$

where the ϕ_0-dependence is independent of all others (modulo an overall shift), the η-dependence is treated as conditional to the momentum and the dependences on luminosity ($\mathcal{L}$), number of COT hits ($\mathcal{H}$), track density (secance, $\mathcal{S}$), and time (t) were treated jointly.

To obtain the multiplicative correction factors, we divided in bins the distribution of dE/dx as a function of each of the above quantitities. The uniform binning pitch was chosen to allow reasonable statistics in all bins. Instead of fitting these distributions with smooth functions, we reweighted the content $\langle dE/dx\rangle_i$ of each bin i by $w_i = 17.7[\mathrm{ns}]/\langle dE/dx\rangle_i$ so that all histograms are flattened to the arbitrarily chosen value of 17.7 ns. This procedure was chosen because statistics is enough to make the effect of bin-by-bin statistical fluctuations very small with respect to the size of effects to be corrected. So the small loss in statistical power was largely recovered by the benefit of having a simpler and faster procedure.

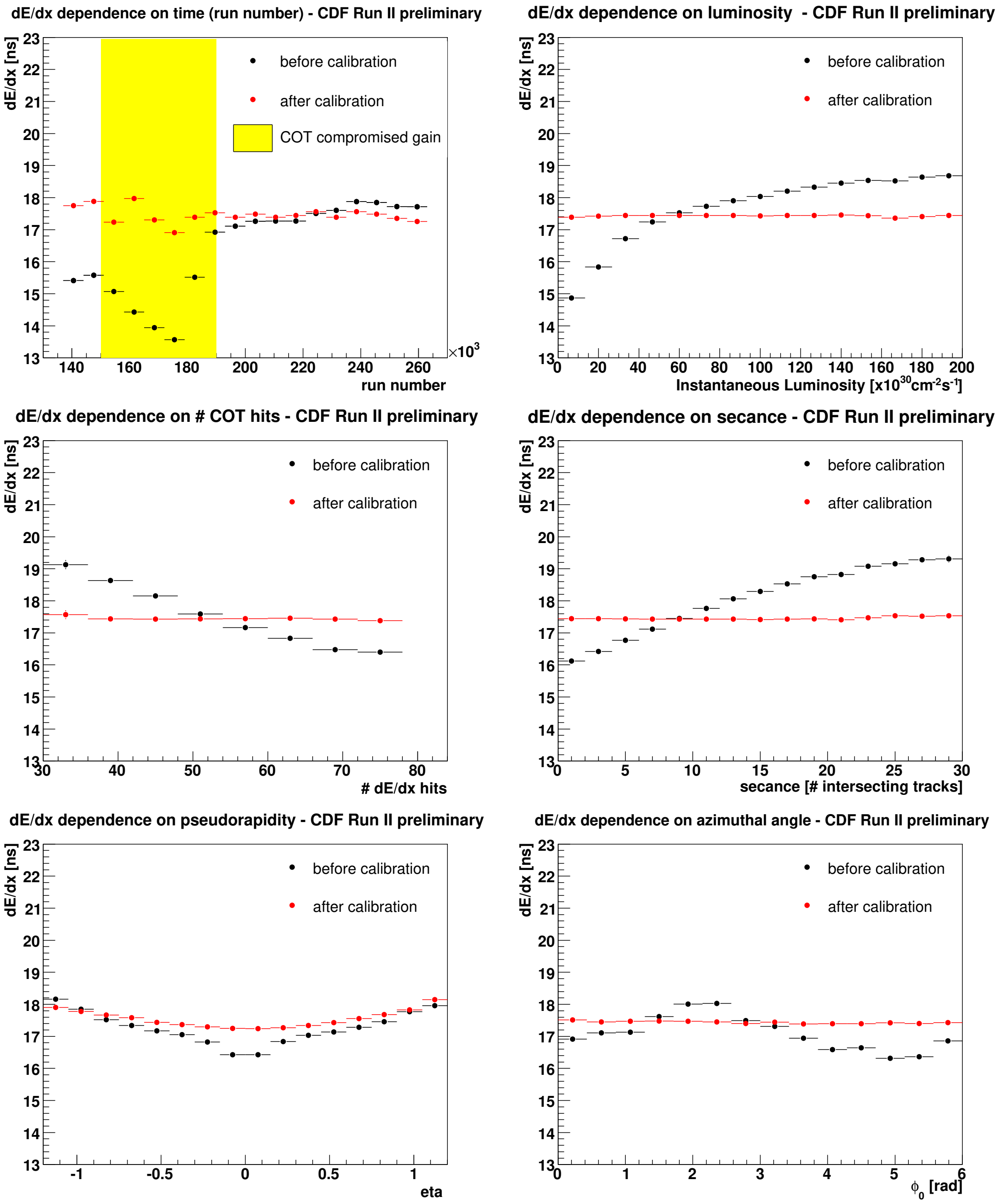

Figure 3: Distribution of $\langle dE/dx \rangle$ as a function of the investigated quantities before (black) and after (red) the corrections.

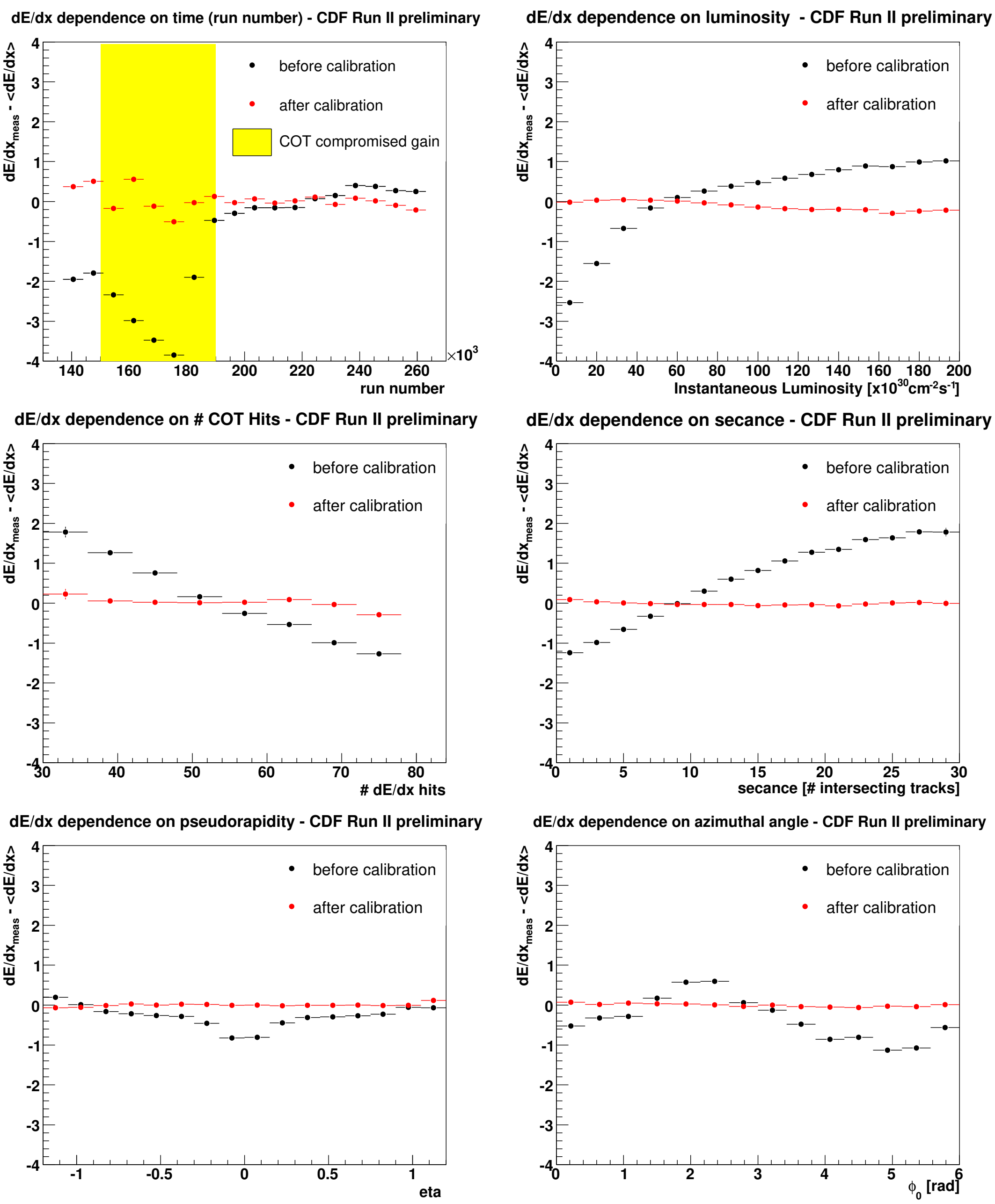

Figure 4: Distribution of dE/dx residuals as a function of the investigated quantities before the corrections (black) and after the corrections (red).

6 Results of corrections

Figures 3 and 4 show the impact of corrections. Figure 3 shows the average dE/dx response as a function of each variable before (black) and after (red) the corrections. Figure 4 shows the average *dE/dx residual.* The residual is defined as the difference between the observed dE/dx for a particle and its expected value (function of momentum and mass hypothesis). The advantage of the residual with respect to the mere dE/dx is that it subtracts the natural momentum dependence thus decoupling the effects induced by momentum from the undesired ones induced by environmental dependences. As an example one can compare the bottom-left panel in fig. 3 and fig. 4. The average dE/dx still shows an η–dependence even after the corrections (red points), but the average residual is instead flat, showing that the observed dependence was induced by the pseudorapidity–momentum correlation rather than an poor calibration. All observed effects are significantly reduced, down to the < 0.5 ns level.

7 Extraction of the curve of expected dE/dx

After all major gain variations have been corrected we extracted the curve of the expected dE/dx has a function of momentum. In order to minimize possible biases of this curve due to uncontrolled background we further tightened the selection to increase the purity of kaons and pions being used. We tightened the mass cuts to $XXX < m(K\pi) < XXX$ and $XXX < Q < XXX$ (see fig. 5) and obtain an extremely pure sample of pions and kaons.

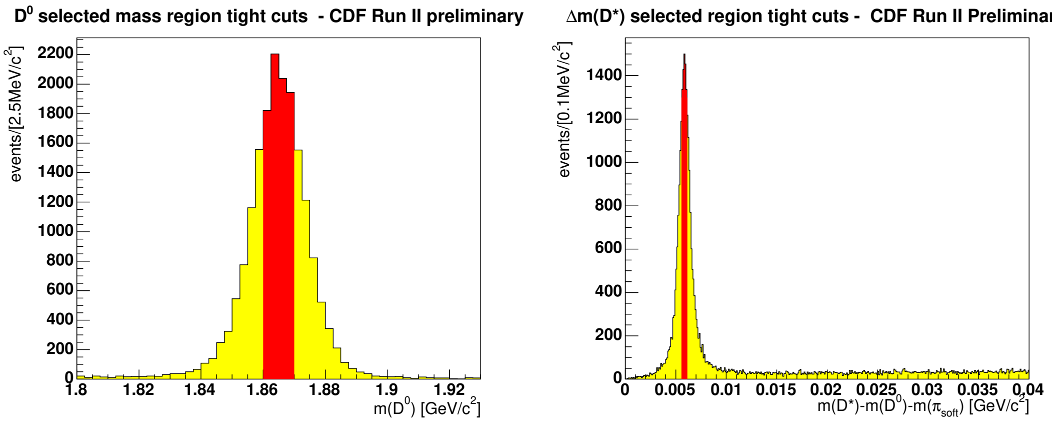

Figure 5: Invariant $K\pi$-mass distribution (a) and distribution of the mass difference between D^{*+} mass and the sum of D^0 and pion masses, $m(K\pi\pi) - m(K\pi) - m_\pi$ (b) with tightened selection, for the extraction of the predicted dE/dx curves.

Figure 6 shows average values of the dE/dx as a function of momentum, separately for pions/kaons (blue/green) and positively/negatively-charged particles. Fit of these curves (with a_i, b, c_j, and C free parameters) with the usual, empiric modification of the Bethe-Bloch curve traditionally used at CDF since Run I

$$\left\langle \frac{dE}{dx} \right\rangle = \frac{1}{\beta^2} \left[c_1 \ln \left(\frac{\beta\gamma}{b + \beta\gamma} \right) + c_0 \right] + a_1(\beta - 1) + a_2(\beta - 1)^2 + C, \tag{2}$$

are overlaid, and the fit residuals are shown below each plot. The fits are satisfactory across the whole momentum range for all four subsamples, with departures from data not exceeding the 2–3σ level for most of the spectrum. The figure of all four curves in the same plot shows some features that are worth discussing. A systematic, momentum-dependent difference between the average dE/dx response of positively- and negatively-charged particles is visible for both kaons and pions. The dE/dx of positive kaons and pions is larger than their negative counterparts, with a difference that is as large as 1 ns at low momenta and becomes negligible at momenta larger than 15–20 GeV/c. In addition, the difference between kaons and pions average dE/dx (distance between the two curves) decreases at higher momenta, indicating a poorer PID performance for high-momentum tracks. These momentum-dependent differences can be ascribed, at least partially, to the fact that incomplete hit-level calibrations are applied to the tracks in our sample. The hit level corrections, in particular the one that correct the dE/dx as a function of the aspect-angle of the track in the cell, are know to remove most of these differences, and the fact that these are not present in our sample qualitatively explains the larger severity of these effects with respect to old iterations of calibrations.

8 Performance

8.1 Separation

Any improvement in particle identification capability quantifies the impact of the dE/dx corrections. Given a PID-related observable, (dE/dx residual, for instance) the identification performance relies on the difference in the distributions of the chosen observable between the classes of events to be identified. Such difference is generally expressed in terms of a *separation* between those distributions.

The conventional way to quote the separation is to provide an estimation of the distance between the centers of the distributions in units of their standard deviations (σ). The reliability of this estimation degrades increasingly as the distributions deviate from the Gaussian shape. In presence of long tails or strongly asymmetric distributions, the separation estimated with this approach is not reliable. In addition, this choice has limited applicability: if the PID information is given in form of a multi-dimensional observable, no straightforward and unambiguous way to quote the separation is provided.

We therefore chose a different approach. Suppose that one uses PID information in a sample of N events, of which a fraction f are pions and a fraction $1 - f$ kaons (see fig. 8). An unambiguous estimator of the separation power is the precision ($1/\sigma_f$) in estimating the fraction f, where σ_f is the statistical uncertainty on the fraction. An upper bound to the precision is provided by the value $1/\sigma_f^{\text{best}} = \sqrt{N/f(1-f)}$ obtained in the ideal case of classes of events being totally separated, i. e., the distributions of the chosen PID

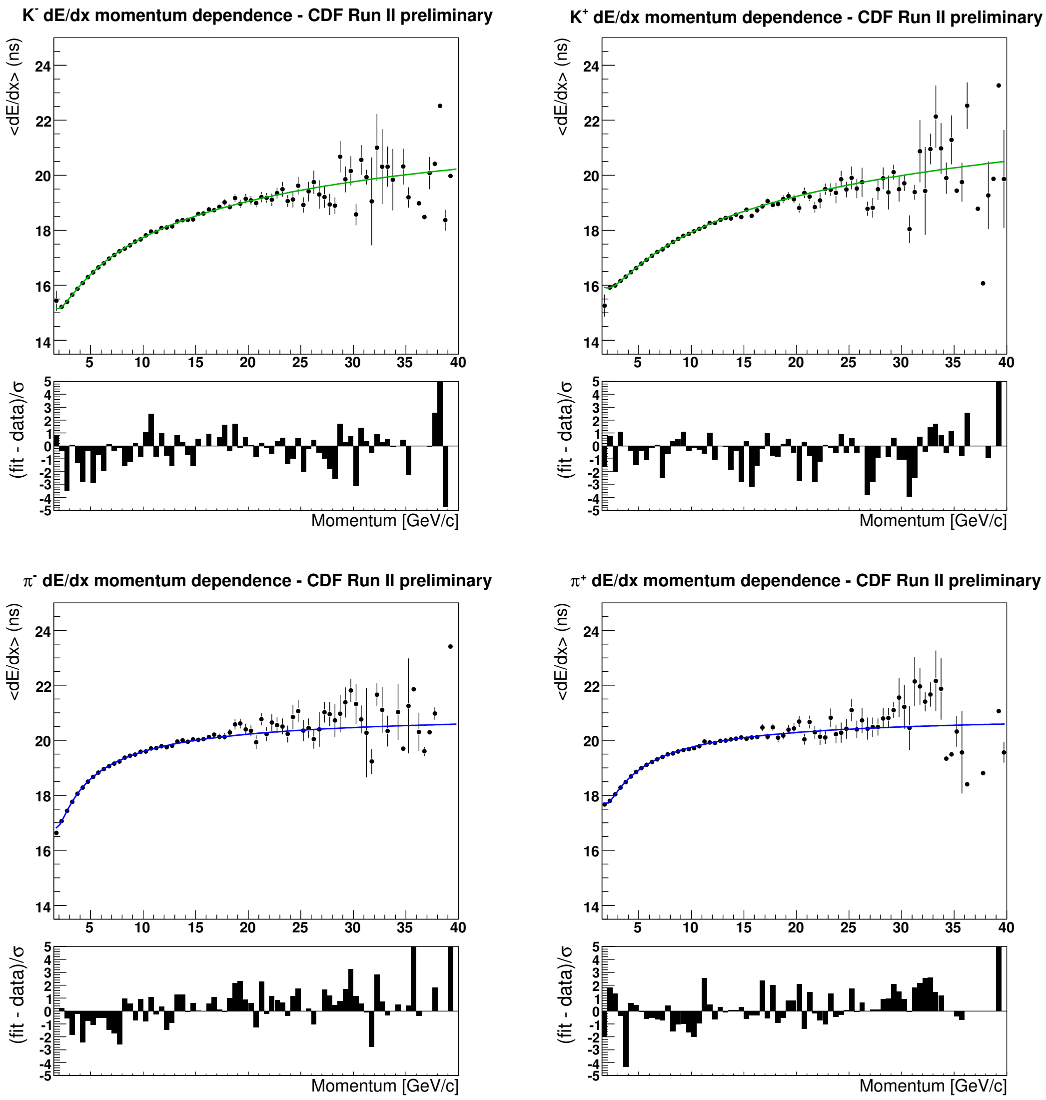

Figure 6: Average dE/dx response as a function of momentum for negative kaons, positive kaons, negative pions and positive pions. Fit results are overlaid and fit residuals are shown.

observable having zero overlap (see fig. 8(b)). In this limiting case, the only uncertainty in assigning an event to one of the two classes comes from the Binomial fluctuation due to the finite sample-size.

We quote the separation between kaons and pions as the relative precision with respect to the ideal case $s = \sigma_f^{\text{best}}/\sigma_f$. The resolution σ_f can in principle be determined by repeating a maximum likelihood fit of the fraction f on a sufficient number of pseudo-

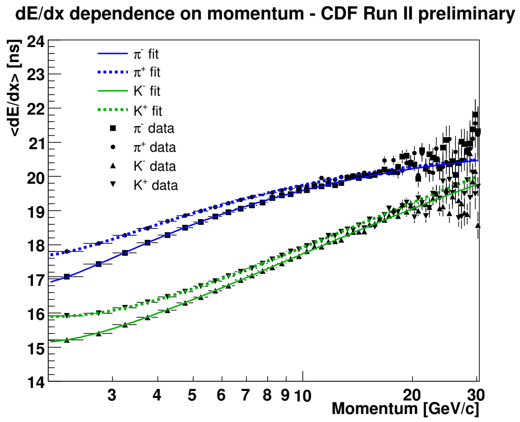

Figure 7: Average dE/dx response for pions and kaons split by charge overlaid in the same plot.

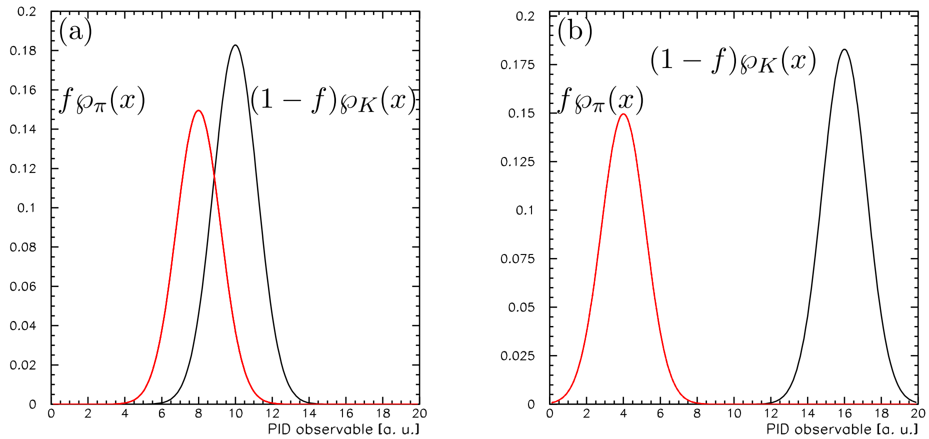

Figure 8: Illustration of the distribution of a simulated PID-observable (x, in the text) for pions (red curve) and kaons (black curve) in case of partial (a) and total (b) event separation.

experiments that simulate the experimental data, and evaluating the spread of results around the input values. We used, instead, the Minimum Variance Bound [8]. Choosing a variable x to separate the classes of events, in the simple case of two classes, the MVB is written as

$$\sigma_f^2 = \frac{1}{N}\left[\int \frac{(\wp_\pi(x) - \wp_K(x))^2}{f\wp_\pi(x) + (1-f)\wp_K(x)}dx\right]^{-1}, \tag{3}$$

where $\wp_\pi(x)$ and $\wp_K(x)$ are the probability distributions of x, normalized to unit area, for pions and kaons respectively (see fig. 8).[2] Following our approach, the separation power of the variable x in the given sample is determined by evaluating

$$s = \sigma_f^{\text{best}}/\sigma_f = \sqrt{f(1-f)\int \frac{(\wp_\pi(x) - \wp_K(x))^2}{f\wp_\pi(x) + (1-f)\wp_K(x)}dx}. \tag{4}$$

This quantity is independent of the sample size, but depends on the true values of fractions as generally happens for resolutions. This is also intuitive, since an easier separation is expected among populations similar in size. The separation ranges from $s = 0$, i. e., no separation corresponding to completely overlapping distributions, to $s = 1$, i. e., the maximum achievable separation in the given sample. The quantity s is well-defined whatever the shape and the dimensionality of the PID distributions involved. In addition, any s value can be analytically converted in an "equivalent $n\sigma$ separation" to quote, in a more conventional way, the separation in σ units one would have observed if the distributions of the chosen PID variable were Gaussian.

The achieved separation power depends on the variable (i. e., x) used to measure it, and a wise choice of the variable (or set of variables) may enhance the actual separation. In our case we used the dE/dx residual in pion hypothesis: $dE/dx_{obs} - dE/dx_\pi$. The separation determined with the uncorrected dE/dx was $s = 46.8\%$, while the value after the corrections increase to $s = 56.4\%$. Both are integrated in momentum. This is an 22% improvement, which is also visible qualitatively by comparing the width of the curves in fig. 9 before (left) and after (right) the corrections. These values correspond to approximately $XXX\sigma$ and $XXX\sigma$ Gaussian-equivalent separations, respectively. Since s depends on the specific proportions among classes of events present in the sample, the above values hold only for samples with approximately equal contributions from pions and kaons.

	Separation s	Equivalent separation [Gaussian σ]
Uncorrected	$\simeq 0.47$	$\simeq XXX$
φ_0 only	$\simeq 0.49$	$\simeq XXX$
$\eta\|p$ only	$\simeq 0.47$	$\simeq XXX$
$\mathcal{H}, \mathcal{L}, t, \mathcal{S}$ only	$\simeq 0.53$	$\simeq XXX$
Corrected	$\simeq 0.56$	$\simeq XXX$

Table 1: Observed separation between kaons and pions as a function of the corrections applied.

Table 1 contains the values of the separation evaluated after each step of the calibration.

[2] The symbol x may stand for a set of many variables, discrete or continuous, and the integral extends over the whole x domain.

	Uncorrected	φ_0 only	$\eta\|p$ only	$\mathcal{H}, \mathcal{L}, t, \mathcal{S}$ only	Corrected
$2 < p < 3$ GeV/c	$\simeq XXX$	$\simeq XXX$	$\simeq XXX$	$\simeq XXX$	$\simeq XXX$
$3 < p < 4$ GeV/c	$\simeq XXX$	$\simeq XXX$	$\simeq XXX$	$\simeq XXX$	$\simeq XXX$
$4 < p < 5$ GeV/c	$\simeq XXX$	$\simeq XXX$	$\simeq XXX$	$\simeq XXX$	$\simeq XXX$
$5 < p < 7$ GeV/c	$\simeq XXX$	$\simeq XXX$	$\simeq XXX$	$\simeq XXX$	$\simeq XXX$
$7 < p < 10$ GeV/c	$\simeq XXX$	$\simeq XXX$	$\simeq XXX$	$\simeq XXX$	$\simeq XXX$
$p > 10$ GeV/c	$\simeq XXX$	$\simeq XXX$	$\simeq XXX$	$\simeq XXX$	$\simeq XXX$

Table 2: Observed separation between kaons and pions as a function of the corrections applied in different momentum ranges.

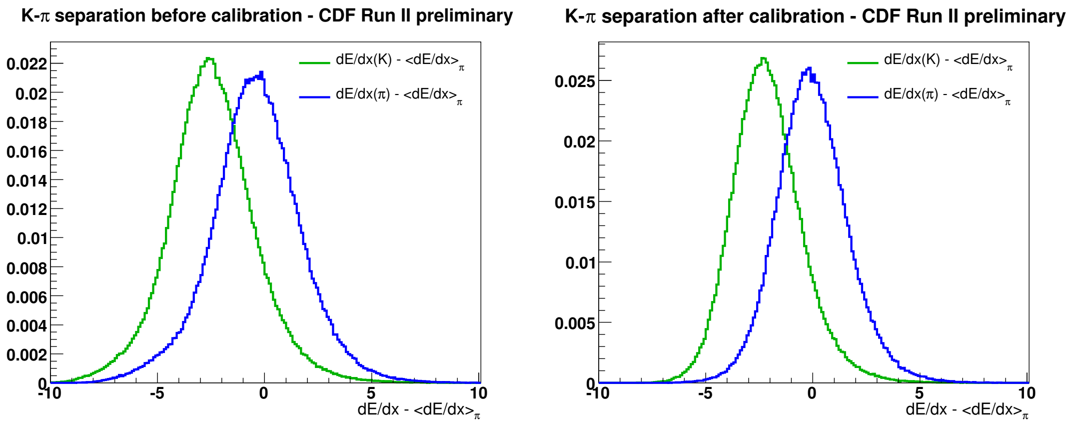

Figure 9: Distributions of the pion residual for pions (blue curve) and kaons (green) before (left) and after (right) the corrections.

8.2 Residual correlations

With an ideal PID detector, no correlation is expected between PID measurements on different particles. Non vanishing correlation indicates the presence of residual dE/dx gain variations. An uncorrected gain variation, in fact, would induce a correlation between the observed ionizations of distinct particles, through the inevitable correlations present in the calibration sample.

The mechanism through which the correlations of the calibration sample induce correlations in the dE/dx response is explained in mathematical terms. Suppose to use dE/dx information from two particles per event. One wants to properly write the distributions of dE/dx for the two particles, in general $\wp(dE'/dx', dE''/dx'')$. If the dE/dx shows a gain variation as a function of, say φ_0, the p.d.f. becomes conditional at given pair of azimuthal angles:

$$\wp\left(\frac{dE'}{dx'}, \frac{dE''}{dx''}\right) \quad \longrightarrow \quad \wp\left(\frac{dE'}{dx'}, \frac{dE''}{dx''}\middle|\varphi_0', \varphi_0''\right) \times \wp\left(\varphi_0', \varphi_0''\right). \tag{5}$$

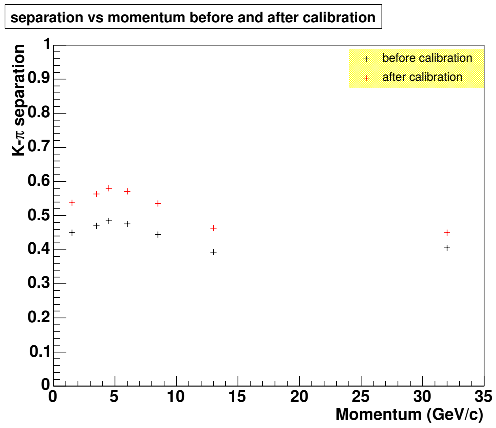

Figure 10: Separation as a function of momentum.

Hence, the desired distribution is obtained as

$$\begin{aligned}\wp\left(\frac{dE'}{dx'},\frac{dE''}{dx''}\right) &= \int \wp\left(\frac{dE'}{dx'},\frac{dE''}{dx''}\middle|\varphi_0',\varphi_0''\right)\times\wp\left(\varphi_0',\varphi_0''\right)d\varphi_0' d\varphi_0'' \\ &= \int \wp\left(\frac{dE'}{dx'}\middle|\varphi_0'\right)\wp\left(\frac{dE''}{dx''}\middle|\varphi_0''\right)\times\wp\left(\varphi_0',\varphi_0''\right)d\varphi_0' d\varphi_0''. \end{aligned} \tag{6}$$

If the azimuthal angles are independent observables,

$$\wp(\varphi_0',\varphi_0'') = \wp(\varphi_0')\times\wp(\varphi_0''), \tag{7}$$

and eq. (6) can be factorized as:

$$\begin{aligned}\wp\left(\frac{dE'}{dx'},\frac{dE''}{dx''}\right) &= \int \wp\left(\frac{dE'}{dx'}\middle|\varphi_0'\right)\wp\left(\varphi_0'\right)d\varphi_0' \int \wp\left(\frac{dE''}{dx''}\middle|\varphi_0''\right)\wp\left(\varphi_0''\right)d\varphi_0'' \\ &= \wp\left(\frac{dE'}{dx'}\right)\wp\left(\frac{dE''}{dx''}\right). \end{aligned} \tag{8}$$

However, eq. (7) does not usually hold for particles in a same event. In particular, if the particles originate from the same decay, their kinematic quantities (φ_0, η, ...) will be strongly dependent, preventing the factorization of the dE/dx distribution and yielding correlated values of the observed dE/dx. The sources of correlation can be divided into two types:

Environmental effects – these are all effects not-related to kinematics. Suppose that the dE/dx shows gain variations as a function of the instantaneous luminosity: $\frac{dE}{dx} = \frac{dE}{dx}(\beta\gamma,\mathcal{L})$. Then, since the kaon and the pion from a D^0 decay are reconstructed in the same event (i. e., in the same conditions of luminosity), their observed dE/dx would appear correlated by the common dependence on luminosity. This may apply to a variety of global variables, such as time, pressure or temperature of the gas, and so forth.

Kinematic effects – these are all effects related to kinematics. Suppose that the dE/dx shows gain variations as a function of the azimuthal angle of emission of the particle: $\frac{dE}{dx} = \frac{dE}{dx}(\beta\gamma, \varphi_0)$. Then, since the azimuthal angle of a kaon and a pion from a D^0 decay are correlated by the kinematic of the decay and the selection biases, their observed dE/dx would become correlated. This may apply to a variety of local variables, such as η, z_0, hit multiplicity, etc.

8.3 Impact of correlations

In presence of correlations between observables of a likelihood fit, not only the statistical precision is degraded, but biased estimates may occur. The correlation between observed values of dE/dx plays a role when dE/dx information of multiple particles per event is used. If the values dE'/dx', dE''/dx'',..., dE^n/dx^n are observed for n particles in an event, the presence of correlation imply that these random variables are not independently distributed, i. e., the probability of observing a set of dE/dx values is not factorizable in the single probabilities:

$$\wp\left(\frac{dE'}{dx'}, \frac{dE''}{dx''}, \ldots, \frac{dE^n}{dx^n}\right) \neq \wp'\left(\frac{dE'}{dx'}\right) \times \cdots \times \wp^n\left(\frac{dE^n}{dx^n}\right). \tag{9}$$

The joint p.d.f. in the left-hand side term of above relation needs to be used, instead of the factorized form in the right-hand side. Otherwise, when the dE/dx is used as observable in a multivariate discriminator (ML fit, ANN, BDT...), the multi-particle probability distribution for a dE/dx observable would be incorrect. Similarly, when an event-by-event dE/dx-based selection is applied (i. e., when "cutting" on the dE/dx), the expected overall efficiency corresponding to the chosen requirements would be wrong. Both cases may yield biased estimates of the sample composition.

Relation (9) may hold also in absence of correlation, since *dependence* between variates, a weaker condition than *correlation*, is the harmful feature.[3] However, following a common colloquial convention, here we refer to this dependence as correlation, because in our specific case (common-mode gain-variations, see next section) correlation and dependence coincide.

8.4 Study of correlations

We investigated the combined effect of any possible residual gain variation by allowing for a generic, time-dependent common-mode fluctuation $c(t)$ that affects and correlates the observed dE/dx values in the event. In particular, we extracted the variance (σ_c^2) of the distribution of the common mode, as an estimator of the size of the correlation. We denote the probability distributions of dE/dx residual for pions (with pion mass-hypothesis) as $\wp_\pi(\delta_\pi)$, with standard deviation σ_π. An analogous notation is adopted for kaons. If δ_π and δ_K were independent variables, the probability distribution of their sum ($\delta_K + \delta_\pi$)

[3] As an example, the variates x and $y = x^2$ with flat x-distribution in $[0, 1]$ are fully dependent but show zero correlation.

would satisfy

$$\wp(\delta_\pi + \delta_K) = \wp_\pi(\delta_\pi) * \wp_K(\delta_K), \tag{10}$$

in which the symbol $*$ indicates the Fourier convolution product.[4] Similarly, their difference $\delta_\pi - \delta_K$ would be distributed as

$$\wp(\delta_\pi - \delta_K) = \wp_\pi(\delta_\pi) * \wp_{-K}(-\delta_K), \tag{11}$$

where $\wp_{-K}(-\delta_K)$ is the distribution of the negative residual for kaons $(\frac{dE}{dx}_K - \frac{dE}{dx}_{\text{obs}})$, whose variance satisfies the condition $\sigma^2_K = \sigma^2_{-K}$. Since the variance of a convolution product is the sum of variances of the convoluted distributions, the standard deviations of the distributions of sum and difference are equal:

$$\sigma_{\pi+K} = \sigma_{\pi-K} = \sqrt{\sigma^2_\pi + \sigma^2_K}. \tag{12}$$

On the other hand, if the two residuals are correlated by a common-mode fluctuation, the observed residual (δ^{obs}) is written as the sum of the intrinsic, uncorrelated residual with the common-mode shift:

$$\delta^{\text{obs}}_\pi = \delta_\pi + c \quad \text{and} \quad \delta^{\text{obs}}_K = \delta_K + c. \tag{13}$$

Therefore, the sum of observed residuals, $\delta^{\text{obs}}_\pi + \delta^{\text{obs}}_K = \delta_\pi + \delta_K + 2c$, is distributed as

$$\wp(\delta^{\text{obs}}_\pi + \delta^{\text{obs}}_K) = \wp_\pi(\delta_\pi) * \wp_K(\delta_K) * \wp_c(2c), \tag{14}$$

whereas their difference, $\delta^{\text{obs}}_\pi - \delta^{\text{obs}}_K = \delta_\pi + c - \delta_K - c = \delta_\pi - \delta_K$, is distributed as

$$\wp(\delta^{\text{obs}}_\pi - \delta^{\text{obs}}_K) = \wp_\pi(\delta_\pi) * \wp_{-K}(-\delta_K). \tag{15}$$

Equations (14) and (15) show that, in presence of a common mode, the sum of residuals has greater variance than their difference, $\sigma^2_{\pi+K} > \sigma^2_{K-\pi}$. The standard deviation of the correlation is easily obtained:

$$\sigma_c = \frac{1}{2}\sqrt{\sigma^2_{\pi+K} - \sigma^2_{\pi-K}}. \tag{16}$$

Following eq. (16), we used the distributions of sum and difference of the observed residual to estimate the magnitude of time-dependent common modes. We compared the variances of the correlation extracted from two samples:

physical sample – the sample used in the calibration: each event contains a kaon and a pion produced in the decay of a single D^0 meson. In this sample, the common mode receives contributions from gain variations due to *environmental* effects (pressure, time, luminosity,...) and *kinematic* effects (φ_0, η, ...) since kaons and pions are kinematically correlated.

[4] Henceforth, "convolution" always denote the Fourier convolution product.

mixed-decay sample – an artificial sample containing events in which the kaon and the pion originate from decays of two distinct D^0 mesons, reconstructed in two consecutive events. In this sample, we expect the magnitude of the common mode to be smaller than in the physical sample. The contribution to the gain variations due to *kinematic* effects are reduced, since pion and kaon are kinematically uncorrelated. In addition, the larger the time scale of the gain variation with respect to the typical time interval between the collection of two D^0 decays, the smaller becomes the contribution of *environmental* effects.

This allowed studying what fraction of the correlation (σ_c) is enhanced by the kinematics of the decay and what is due to global effects.

		Physical sample			Mixed sample		
	Units	$\sigma_{\pi+K}$	$\sigma_{\pi-K}$	σ_c	$\sigma_{\pi+K}$	$\sigma_{\pi-K}$	σ_c
Uncorrected	ns	3.52	2.27	1.35	3.31	2.53	1.07
Corrected	ns	2.44	2.25	0.46	2.39	2.32	0.29

Table 3: Observed correlations in the physical and mixed-decay samples after each of the calibrations.

		Physical sample		Mixed sample	
	Units	Uncorr. σ_c	Corr. σ_c	Uncorr. σ_c	Corr. σ_c
$2 < p < 3$	ns	$\simeq XXX$	$\simeq XXX$	$\simeq XXX$	$\simeq XXX$
$3 < p < 4$	ns	$\simeq XXX$	$\simeq XXX$	$\simeq XXX$	$\simeq XXX$
$4 < p < 5$	ns	$\simeq XXX$	$\simeq XXX$	$\simeq XXX$	$\simeq XXX$
$5 < p < 7$	ns	$\simeq XXX$	$\simeq XXX$	$\simeq XXX$	$\simeq XXX$
$7 < p < 10$	ns	$\simeq XXX$	$\simeq XXX$	$\simeq XXX$	$\simeq XXX$
$p > 10$	ns	$\simeq XXX$	$\simeq XXX$	$\simeq XXX$	$\simeq XXX$

Table 4: Observed correlations in the physical and mixed-decay samples in different momentum intervals.

9 dE/dx distributions

Use of dE/dx information in a likelihood fit requires modeling the distributions of the desired observables. As a main dE/dx observable we arbitrarily chose the "pion residual", that is, the dE/dx measured with respect to the average pion response. For increased flexibility and cross-check we also provide the distributions of the "kaon residual", which is the observed dE/dx value around the average kaon response.

The distributions of residuals depend on momentum and on the associated number of COT hits. In order to keep into account these two dependencies we decided to provide

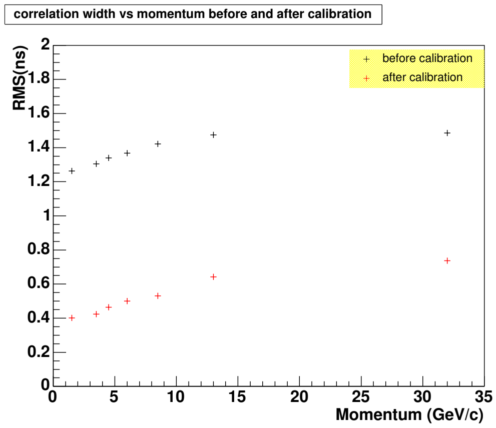

Figure 11: Magnitude of the correlation (i. e.,, σ_c) as a function of momentum.

independent residual parameterizations in six momentum bins: $2 < p < 3$ GeV/c, $3 < p < 4$ GeV/c, $4 < p < 5$ GeV/c, $5 < p < 7$ GeV/c, $7 < p < 10$ GeV/$c, p > 10$ GeV/c. We separately provide a function that models the width of the residual as a function of the number of associated hits. This way we are able to keep separated these two effects and leave to the user the choice of exploiting the maximum information (momentum and hits) or just the momentum-related information. For those analyses where the details of the dE/dx dependence on momentum are of marginal impact we also provide a single parameterization of the dE/dx integrated in momentum (see fig. 12).

We extracted a parameterization of the dE/dx resolution as a function of the number of associated COT hits by fitting the width of the bulk of the pion residual with a single-gaussian for samples of tracks with different number of associated COT hits. The resulting dependence is a linear improvement of the resolution as the number of sampling increases (see fig. **??**).

10 Extending the calibrations to other samples?

This set of calibration has been validated for SVT-triggered charged kaons and pions reconstructed in BSntuples in periods 0–17. However, many analyses for Summer 2009 aim at using data extending past period 17. In addition, analyses requiring use of production-flavor tagging algorithms would benefit from exploiting dE/dx of tracks with lower momenta than the 2 GeV/c required to SVT tracks. We have checked to which extent the present calibrations can be applied also to data of period 18–21 and to soft tracks.

10.1 Extension to post-period 17 data

We have applied the present calibration constants, obtained from data collected in period 0–17, to D^{*+}–tagged $D^0 \to K^-\pi^+$ decays collected in periods 18–21. This is clearly a

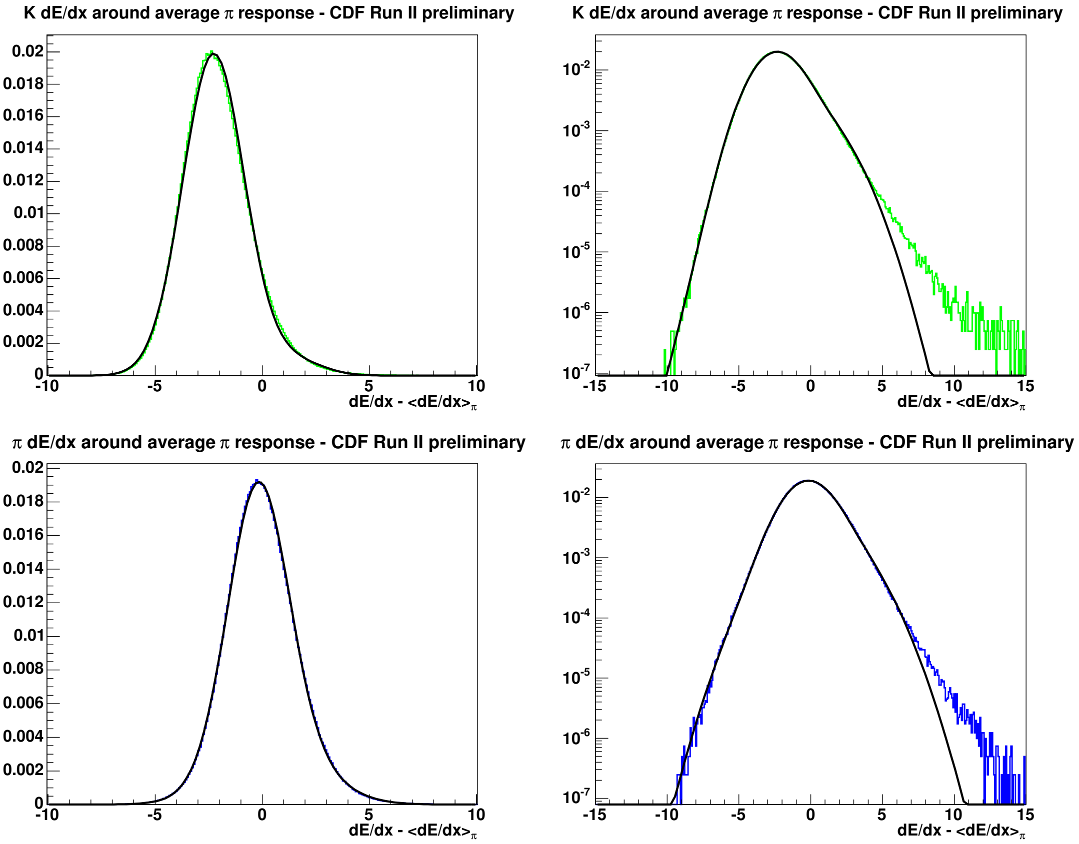

Figure 12: Distribution of observed dE/dx around the average kaon (top plots) or pion (bottom plots) response in linear (left) and log-scale (right). Results of fits with the sum of two gaussian functions are overlaid.

suboptimal approximation of the correct approach, which would imply to re-extract a new set of calibrations from the entire period 0–21 range. However, in the assumptions that no wild variations of performance occurred in the latest data one can try if the calibration constants can be extended reasonably to the period 18–21 data. In fact, by looking at the distributions of dE/dx as a function of the various calibration quantities (see fig. 13) we conclude that at first order, the calibrations can be safely used also for period 18–21 data. Of course the overall PID performance will be slightly degraded and when sufficient data will be collected after period 17 we will proceed to an overall recalibration of the whole sample.

10.2 Extension to soft tracks

The sample of D^{*+}–tagged $D^0 \to K^-\pi^+$ decays provides a straightforward probe of the dE/dx performance for low-momentum tracks: the soft pion originated from the D^{*+}

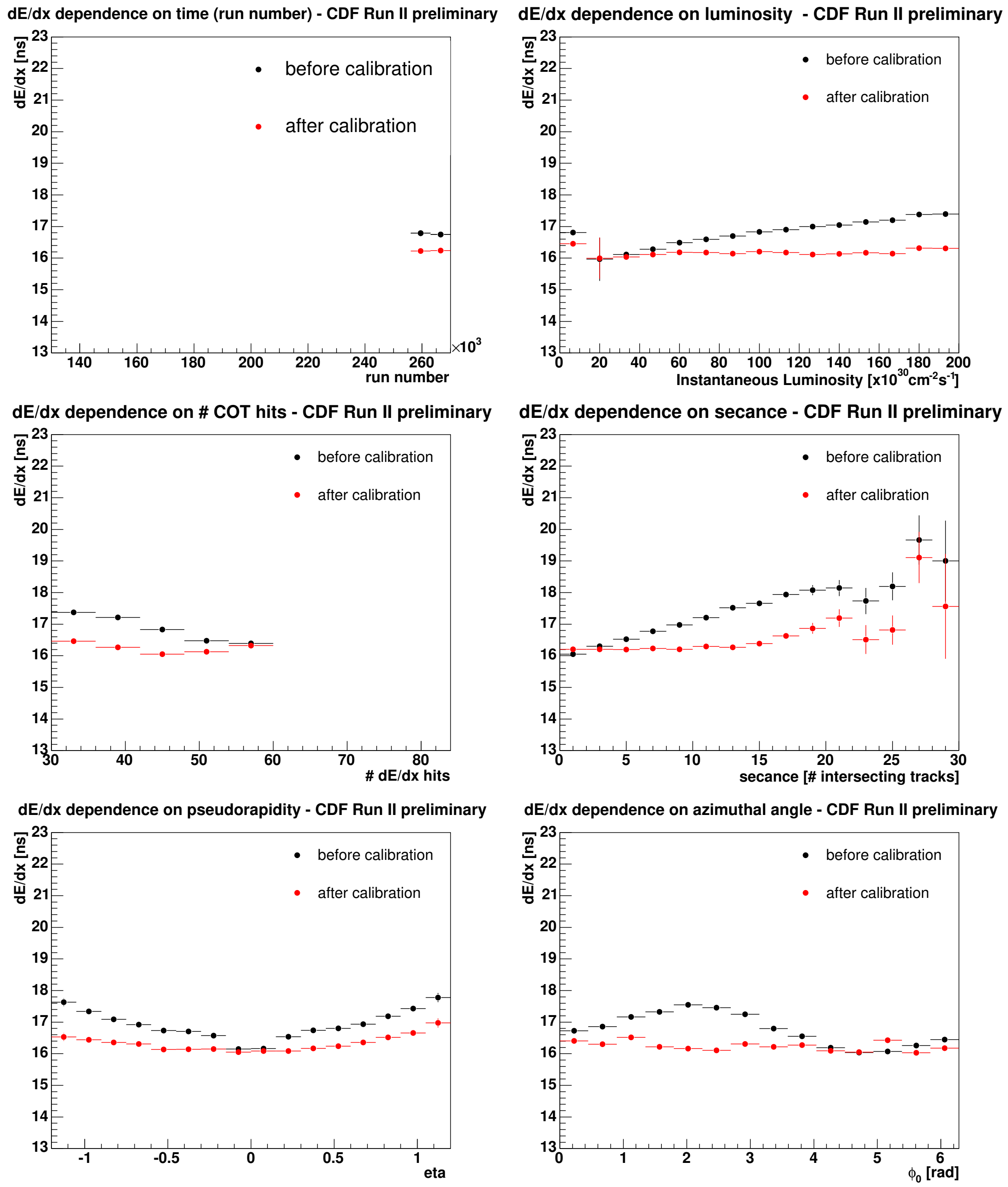

Figure 13: Distribution of $\langle dE/dx \rangle$ for data collected post period 17 as a function of the investigated quantities before (black) and after (red) the corrections.

decays. The typical momentum of this pion is of a few hundred MeV$/c$ that is a completely different region of the momentum spectrum of the sample used to extract the calibrations. In fig. 14 we show the distribution of average dE/dx as a function of Lorentz boost for positive and negative soft pions, overlaid to the curves of the calibrations samples ($p >$ 2 GeV/c). The average dE/dx for negative low-energy pions qualitatively follows the trend observed for negative kaons of the same $\beta\gamma$. However the curves for positive particles show a large discrepancy: the curve for soft pions departs from the kaon curve showing higher values of average dE/dx. The discrepancy between the two curves reaches 1 ns at very low momenta. While this trend needs further investigation, it is qualitatively understood by the lack of a complete set of hit-level corrections. These, especially the aspect-angle correction, are instrumental in removing any curvature-dependent effect, whose relevance is higher the lower the momentum probed. Due to this large effect we discourage the use of the present corrections for low-momentum tracks.

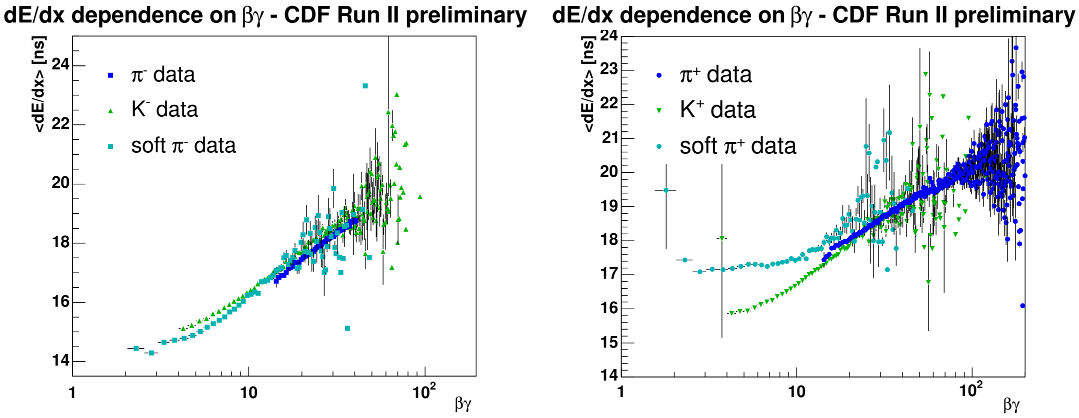

Figure 14: Average dE/dx for negative– (left) and positive–charged (right) soft pions as a function of the Lorentz boost, compared to the one of SVT-triggered pions and kaons.

11 Conclusions and summary

We have produced an updated set of dE/dx calibrations for the Summer 2009 analyses. Rather than extending the previous calibrations, we chose to redo from scratch the study of the effects that affect the dE/dx gain. This yielded several improvements: the present calibrations take into account, for the first time, gain variations depending on luminosity and track density; owing to the larger statistics of the calibration sample, we were able to increase the number of effects that are corrected simultaneously providing a better final result.

We obtained a 22% improvement in the dE/dx separation between charged kaons and pions with $p > 2$ GeV/c, that is found to be consistently above 1.3σ for the whole run and momentum range; we have also greatly reduced the track-to-track correlations induced by use of dE/dx for multiple tracks in the same event. We have studied the dependence of separation and correlations as a function of momentum. We have provided the distributions, and parameterized shapes, for several relevant dE/dx quantities for use in likelihood fits. We have studied and parameterized the dE/dx resolution as a function of the number of COT samplings.

The present calibrations are validated for pions and kaons with $p > 2$ GeV/c collected in period 0–17. The use of the present calibrations can be extended to the periods 18–21 taking into account that the performance will be slightly degraded. We discourage the use of the present calibrations for particles with momenta lower than 2 GeV/c due to the lack of complete low-level hit-based corrections. The calibrations (plots) were blessed in the B physics group meeting of Feb 5 (May 14), 2009.

12 User's interface

All the code and instructions needed to apply the calibrations and use the dE/dx calibrations in 3/fb analyses based on Bsntuples is made available to the user at the following twiki page: `http://www-cdf.fnal.gov/htbin/twiki/bin/view/TOFandDeDx/DeDxCalibration` which contains the most up-to-date information on the status of the calibrations and how to use them.

All needed code is provided as self-contained ROOT macros than can be used standalone after compilation. These are

1. `DeDx2008_funcs.C` (with associated header `DeDx2008_funcs.h`) that contains all functions for dE/dx corrections, predicted dE/dx and residuals:

   - `CorrectedDeDx(double fDeDxRaw, int fDeDxHits, int run, double lumi, int secance, double eta, double phi)` which applies the relevant correction factors.
   - `PredictedDeDx(int charge, double betagamma, double mass)` which produces the mean expected dE/dx for a particle of given mass hypothesis.
   - `SigmaDeDx(int hits, int run)` which produces the expected dE/dx resolution for a generic particle.
   - `SigmaDeDxPi(int hits, int run)` which produces the expected dE/dx resolution for a pion.
   - `SigmaDeDxK(int hits, int run)` which produces the expected dE/dx resolution for a kaon.
   - `DeDxResiduals(double x, double mass, double mom)` which produces the residual shape in *pion* hypothesis.
   - `DeDxResidualsK(double x, double mass, double mom)` which produces the residual shape in *kaon* hypothesis.

   The users simply needs to "include" this in their analysis code to use the functions;

2. `DeDx2008_secance.C` (`DeDx2008_secance.h`) that contains function to calculate the secance. This needs to be included, but the user also needs to make sure the function `Secance()` is called in a loop on all tracks containing COT hits from the event. An example program to demonstrate this as detailed below;

3. `TestSecance.C` (`TestSecance.h`): a simple ROOT program generated on one test ntuple. It calls the `Secance()` function in a loop over all tracks from each event, then compares the calculated secance to the stored value. It is designed to show the user how to correctly use this function, and check that it returns the same value as we have calibrated on.

4. `TestResiduals.C` (`TestResiduals.h`) that is a ROOT macro to plot the fitted residuals.

References

[1] R. Talman," On the Statistics of particle Identification Using Ionization" Nucl. Instrum. Methods 159 (1979); W.W.M. Allison and J.H. Cobb, "Relativistic charged particle identification by energy loss", Ann. Rev. Nucl. Part. Sci. 30, 253 (1980); F. Sauli, "Principles of operations of multiwire proportional and drift chambers", in *Experimental Techniques in High Energy Physics*, T. Ferbel (ed.), Menlo Park, Addison-Wesley (1987).

[2] M. Binkley *et al.*, "Aging in Large CDF Tracking Chambers", Nucl. Instrum. Methods, A515, 53 (2003); D. Allspach *et al.* "Aging in the large CDF axial drift chamber", IEEE Trans. Nucl. Sci. 52, 2956 (2005).

[3] D. Tonelli, P. Catastini, G. Punzi *et al.*, "Track-based calibration of the COT specific ionization", CDF Note 6932

[4] D. Tonelli, "First observation of the $B_s^0 \to K^+K^-$ decay-mode and measurements of the B^0 and B_s^0 mesons decay rates into two-body charmless final states", PhD thesis, CDF Note 8667

[5] C. Amsler *et al.* (Particle Data Group), "The Review of Particle Physics", PLB66712008.

[6] S.-S."Eiko"Yu *et al.*, "COT dE/dx Measurement and Corrections", CDF Note 6361

[7] Bob Wagner, private communication.

[8] F. James, "Statistical Methods in Experimental Physics" (2nd ed.) World Scientific.

[9] H.-C. Fang, M. Shapiro *et al.*, "dE/dx Templates for the Measurement of the Relative Branching Ratios BR($B \to DK$/BR($B \to D\pi$))", CDF Note 8716

Appendix A: adding luminosity information in BStntuples

Because of a known bug, the instantaneous luminosity information is not stored in the BStntuples: the `TeV_lumi` variable has meaningless values. Since this quantity can have a sizable effect on the dE/dx performance we had to attach luminosity information in our data. This was done by retrieving it from the database. We generated an appropriate `sql` query and executing it in the shell at the beginning of each run from the user's `"BeginRun"` method (or analogue) and removing the previous runs dump of the instantaneous luminosity vs run number. (`getrunlumi.sh` is called with argument run-number this modifies `getrunlumi_template.sql` — with a dummy string for a run-number — into `getrunlumi.sql` — which has the dummy string replaced by run-number). In the `Event` method we read in the information from the text file generated in iii) using `ifstream` and simply grab the lumi for the current run-section number. Below follow detailed instructions. Put the following script in the SAME directory that you are running your root code in, call it `getrunlumi.sh`

```
#!/bin/tcsh -f
# getrunlumi retrieves section number and instantaneous luminosity
# from the oracle database and dumps into a text file.
# getrunlumi_template MUST be present in the same directory
# where this code is run from. Extraneous output from the sql query
# is removed. Useage  "./getrunlumi.sh 257569 > /dev/null"
# The /dev/nul makes sure that there is no screen output.
# Author Farrukh Azfar, 8th February 2008.

#set up cdfsoftware to make sure Oracle is up and running.

source ~cdfsoft/cdf2.cshrc
setup cdfsoft2

#the run number is the first argument

set run_number = $1

#replace  the string THISRUN with run_number in the template file
#to make the new sql file to read the database

sed -e s/THISRUN/"$run_number"/g getrunlumi_template.sql > getrunlumi.sql

#run the sql query which will spool the info into a file
sqlplus cdf_reader/reader@cdfofprd @getrunlumi.sql

#run sed to replace all empty spaces
sed -e '/^ *$/d' lumi_run_$run_number.txt > lumi_run_$run_number.txt.tmp

#remove the original file
rm -f lumi_run_$run_number.txt

#now remove the last line (text from oracle) in the tmp file
```

```
sed '$d'  lumi_run_$run_number.txt.tmp > lumi_run_$run_number.txt.tmp.tmp

#rm the tmp file
rm -f lumi_run_$run_number.txt.tmp


mv -f lumi_run_$run_number.txt.tmp.tmp lumi_run_$run_number.txt
```

Put the following `sql` query into a file called `getrunlumi_template.sql` once again in the SAME directory that your root code runs in and call it

```
SET NEWPAGE NONE
SET HEADING OFF

SPOOL runSections_THISRUN.lis
SPOOL lumi_run_THISRUN.txt
SELECT SECTION_NUMBER,BOILUM FROM RUNSECTIONS WHERE RC_RUNNUMBER=THISRUN
  ORDER BY SECTION_NUMBER
/

SPOOL OFF
EXIT
```

Make sure that the very top of your root code has:

```
#include "Stntuple/obj/TStnHeaderBlock.hh"

   tempAna = GetAna();
   header = tempAna->GetHeaderBlock();
```

Next in your "BeginRun" method:

```
//remove the file from previous run...else you will have tons of them !
    char remove_file[300];
    sprintf (remove_file,"rm -f lumi_run_*.txt");
    system(remove_file);



    int run_number = header->RunNumber();



  //define a character string to pass to the system
   char systemcall[300];
```

```
  //puts everthing into a text file FORMAT is important: "lumi_run_RUNNUMBER.txt"
   sprintf (systemcall,"./getrunlumi.sh %d > /dev/null", run_number);
   system(systemcall);
```

and then finally in your "Event" method:

```
//define character strings for various things
  int section_number = header->SectionNumber();
  int runnumber = header->RunNumber();

  char lumifilename[300];
  string read_input;
  float read_lumi, read_section_num;
  float out_lumi;


  sprintf(lumifilename,"lumi_run_%d.txt",runnumber);
  ifstream lumifile(lumifilename);

  while(!lumifile.eof()) {

    lumifile >> read_section_num >> read_lumi;

    if(read_section_num == section_number) {
      out_lumi = read_lumi;
      break;
    }

  }

//outlumi is the variable representing the instantaneous luminosity for
//for section_number
```

Appendix B: Secance — an estimator of track density

To estimate the effect of possible pile-up of COT hits from tracks intersecting radially the pion or kaon of interest we used the secance quantity, as suggested by Aseet Mukherjee. This was first used by the LBL group [9], and building on their experience we refined its definition for a better estimation of density of hits from extra-tracks around the track of interest. Given a track whose dE/dx is measured, we define secance as the number axially intersecting tracks in the 30–140 cm radial range. After removing duplicate tracks ($\approx 0.1\%$), we considered all tracks with $p_{\mathrm{T}} > XXX$ GeV/c and with at least one hit associated in the COT. We found that the track multiplicity per event can be very large, sepcially at high luminosity, and so can be the number of intersection. We decided to limit the possible number of intersections to 300, thus neglecting a very small (XX%) fraction of events exceeding this threshold.

The intersection condition is imposed by the following code snippet:

```
In progress
```

which takes properly into account the flight-direction of the particle, considering only the section of the circular trajectory corresponding to $t > 0$. This is an improvement with respect to the original definition used by the LBL Group.

Acknowledgments

We wish to thank Mark Mattson and Luigi Calligaris for help with the code for skimming the Bsntuples, Angelo Di Canto for help in validating them, Aseet Mukherjee for suggesting the secance and useful discussion, Thomas Kuhr and Michal Kreps for cross-checking the secance and help with BottomMods framework. Michael J. Morello and Paola Squillacioti to volunteer as beta-testers of the calibrations.