

Particle identification using energy loss in the pixel detector

Ferenc Siklér*

*KFKI Research Institute for Particle and Nuclear Physics
Budapest, Hungary*

June 22, 2006

Abstract

Contents

1	Introduction	2
2	Energy loss estimators	2
3	Energy loss model	3
3.1	Resonant part	3
3.2	Coulomb part	4
3.3	Electronic noise	4
3.4	Average and most probable energy loss	4
4	Approximate energy loss model	5
4.1	Determination of model parameters	6
5	Derivatives	6
6	Minimization	9
7	Extraction of yields	10
8	Results	10

*E-mail: <sikler@rmki.kfki.hu>. Twiki: CMS/TrackerSoftware.

1 Introduction

Identification of charged particles is crucial for hadron physics where measured particle yields, spectra and correlations have to be compared to model predictions. It is also very important for reducing the background of physics processes (e.g. in B-physics the decay $B_s^0 \rightarrow D_s^\mp K^\pm$; e/π separation [1]; use for stau in gauge-mediated SUSY breaking models [2]; cleaning candidates of resonance decays).

Silicon detectors can be employed for identification by proper use of energy deposit measurements along the trajectory of the particle. In silicon there is practically no logarithmic rise of specific energy loss, thus it can only be used below the minimum ionization region. Recent studies show that five or even four layers of silicon allow to reach 10% resolution, using the truncated mean method [2].

In CMS the high occupancy of silicon strips in central A+A collisions renders their inclusion into particle identification particularly difficult. The use of silicon pixels alone allows to use the same analysis for low multiplicity p+p, p+A and high multiplicity A+A events. (At the same time it enables the reconstruction of very low p_T particles, down to 200 MeV/c.) This is a challenge at the same time, because a charged particle, defined by pixel hit triplets, has only three hits. This fact prompts the development of a new method for the extraction of energy loss information.

2 Energy loss estimators

There are several estimators which can represent a measured data sample. The mean is not a robust one, because it is sensitive to large fluctuations. The median is already a better choice.

For probability density functions with long tails, like energy loss distributions, the method of truncated mean has been developed. Here some percentage of the highest (and sometimes lowest) E/x values are discarded and the remaining ones are averaged. While the method is easily applicable, it has some problems, as well:

- When many detector channels are in overflow, the truncation does not work. This way particles with low momentum are not separable (e.g. kaons and protons below 500 MeV/c).
- The distribution of the truncated mean estimator is not gaussian. It still has tails toward higher energy deposits, reflecting the original nature of the probability density function.
- Its mean and its standard deviation is not calculable, these have to be extracted from a multi parameter fit. It is disadvantageous when the distributions of the estimator for different particles overlap (e.g. pions and kaons above 600 MeV/c).
- A particle has energy loss measurements from different regimes of silicon thickness (x). At the same time the distribution E/x depends on x , so the truncation uses values belonging to different distributions. That is why the extracted mean and its standard deviation depends on kinematical variables (η and p_T), the resulting distributions have to be refitted for each phase space bin.

The method of maximum likelihood is a general framework for point estimation. It has beneficial asymptotic properties: unbiased, gaussian distributed; efficient, reaches the Cramer-Rao lower bound; it can be used for point and interval estimation; its distribution can be calculated with help of the Fisher information matrix. These benefits come at a price: the method needs a very good description of the probability density function (see Section 3). In case of energy loss measurement the knowledge (calculation or measurement) of the conditional probability distribution

$$P(y|\beta\gamma;x) \tag{1}$$

	K	L	M
E_{0n} [eV]	4033	241	17
F_n	2/14	8/14	4/14

Table 1: Effective binding energies E_{0n} and oscillator strengths F_n for the K, L and the M-shell in silicon, from [3].

is needed, where y is the measured ADC value, $\beta\gamma = p/m$ belonging to the particle, and x is the path of the particle inside the silicon. As it will be shown in Section 4, with some approximations this conditional probability can be reduced to a simpler form

$$P(y|m) \quad (2)$$

where m is the most probable value of the energy loss distribution.

3 Energy loss model

Here we follow the easily calculable model in [3]. The typical energy loss is

$$\xi(x, \beta) = \frac{K}{2} z^2 \frac{Z}{A} \rho \frac{x}{\beta^2} \quad (3)$$

where $K = 4\pi N_A r_e^2 m_e c^2 \approx 0.307\,075$ MeV cm²/mol; $Z = 14$ is the atomic number, $A = 28.0855$ is the atomic mass, $\rho = 2.33$ g/cm³ is the density, x is the thickness of the absorber; ze is the charge, β is the velocity of the incident particle. For $\beta = 1$ and $x = 0.03$ cm silicon, $\xi = 5.31$ keV. Note that 1 ADC $\approx$ 0.5 keV.

The $w(E)$ collision probability, the cross section, of the particle can be split in two parts: a δ -function term describing resonance excitations and a truncated $1/E^2$ term due to Coulomb excitations:

$$w(E) = \frac{1}{x} \left[\sum_n m_{rn} \delta(E - E_{0n}) + m_{Cn} E_{0n} H(E - E_{0n})/E^2 \right] \quad (4)$$

where H denotes the Heaviside function.

3.1 Resonant part

The average number of resonance collisions is

$$m_{rn} = \xi \frac{F_n}{E_{0n}} \left[\log \frac{2m_e c^2 \beta^2 \gamma^2 E_{0n}}{E_{0n}^2 + (\hbar l)^2} + \frac{l^2}{\gamma^2 \omega_p^2} - \beta^2 \right] \quad (5)$$

where l is defined by the equation

$$\omega_p^2 \sum_n \frac{F_n}{l^2 + \omega_n^2} = \frac{1}{\beta^2 \gamma^2}, \quad E_{0n}^2 = (\hbar \omega_n)^2 + (\hbar \omega_p)^2 F_n \quad (6)$$

and the plasma energy

$$\hbar \omega_p = 28.8 \sqrt{\frac{Z}{A}} \rho = 31 \text{ eV} \quad (7)$$

The equation can be solved for l by bracketing the root followed by bisection. If $\beta\gamma < \beta\gamma_{crit} \approx 1.58$ then $l = 0$ is taken. The distribution of energy loss for resonance excitations is

poissonian, which can be replaced by a gaussian if $m_{rn} \gg 1$. The gaussian has an average $\overline{\Delta}_r$ and sigma σ_r given as

$$\overline{\Delta}_r = \sum_n m_{rn} E_{0n} \approx 13.30\xi, \quad \sigma_r^2 = \sum_n m_{rn} E_{0n}^2 \approx 11.43\text{keV} \cdot \xi \quad (8)$$

The probability density function is

$$f_r(\Delta_r|\overline{\Delta}_r, \sigma_r) = \frac{1}{\sigma_r \sqrt{2\pi}} \exp \left[-\frac{(\Delta_r - \overline{\Delta}_r)^2}{2\sigma_r^2} \right] \quad (9)$$

3.2 Coulomb part

The average number of Coulomb-type collisions is

$$m_{Cn} = \xi \frac{F_n}{E_{0n}} \quad (10)$$

The distribution function is given by the Landau distribution if $m_{Cn} \gg 1$. The mean excitation energy $I = 169$ eV is the logarithmic weighted average of the binding energies

$$\log I = \sum_n F_n \log E_{0n} \quad (11)$$

The probability density function is

$$f_C(\Delta_C|\xi) = \frac{1}{\xi} \phi \left[\frac{\Delta_C}{\xi} - \left(\log \frac{\xi}{I} + 1 - C \right) \right] \quad (12)$$

where $C = 0.577215$ is the Euler-constant. The most probable value of the Landau distribution is $\lambda_{mp} = -0.225$. The full width half maximum of the Landau distribution is 4.02ξ . The Coulomb part cannot be reasonably approximated due to its long tail towards higher energy deposits.

3.3 Electronic noise

The noise has gaussian distribution, centered around zero, with sigma σ_e . The probability density function is

$$f_e(\Delta_e|\sigma_e) = \frac{1}{\sigma_e \sqrt{2\pi}} \exp \left[-\frac{\Delta_e^2}{2\sigma_e^2} \right] \quad (13)$$

3.4 Average and most probable energy loss

The average restricted energy loss is

$$\Delta_{restr} = \xi \left[\log \frac{2m_e c^2 \beta^2 \gamma^2 E_0}{I^2} - \frac{E_0}{2m_e c^2 \gamma^2} - \beta^2 - \delta \right] \quad (14)$$

where the energy transfers are restricted to $E < E_0 \approx \mathcal{O}(500 \text{ keV})$. The most probable energy loss is

$$\Delta_{mp} = \xi \left[\log \frac{2m_e c^2 \beta^2 \gamma^2 \xi}{I^2} + 0.198 - \beta^2 - \delta \right] \quad (15)$$

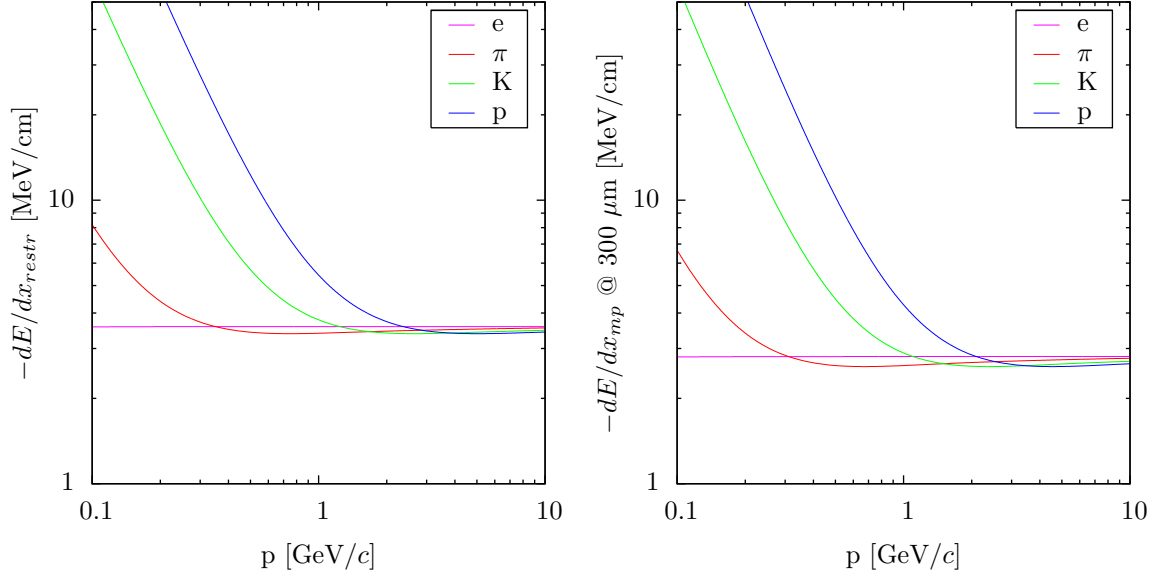


Figure 1: The average restricted and the most probable energy loss at 300 μm silicon thickness, for electrons, pions, kaons and protons.

4 Approximate energy loss model

The resulting energy loss distribution is the convolution of the three (resonant, Coulomb and noise) components

$$f(x, \beta) = f_r * f_C * f_e \quad (16)$$

which in the end is the convolution of the Landau distribution with gaussian. This can be calculated with difficulty and slowly because it needs numerical integration. It seems advantageous to approximate f by a combination of gaussian, exponential and polynomial functions

$$f(y) = \frac{N}{\sigma\sqrt{2\pi}} \cdot \begin{cases} \exp\left[-\frac{(y-m)^2}{2\sigma^2}\right] & \text{if } y < t_1 \\ \exp\left[-\frac{(t_1-m)(2y-t_1-m)}{2\sigma^2}\right] & \text{if } t_1 < y < t_2 \\ \exp\left[-\frac{(t_1-m)(2t_2-t_1-m)}{2\sigma^2}\right] & \text{if } t_2 < y \end{cases} \cdot \frac{1}{\left[1 + \frac{(x-t_2)(t_1-m)}{2\sigma^2}\right]^2} \quad (17)$$

The primitive functions are

$$F(y) = N \cdot \begin{cases} \frac{1}{2} \left[1 + \text{erf}\left(\frac{y-m}{\sigma\sqrt{2}}\right) \right] & \text{if } y < t_1 \\ -\frac{\sigma}{(t_1-m)\sqrt{2\pi}} \exp\left[-\frac{(t_1-m)(2y-t_1-m)}{2\sigma^2}\right] & \text{if } t_1 < y < t_2 \\ -\frac{2\sigma}{(t_1-m)\sqrt{2\pi}} \exp\left[-\frac{(t_1-m)(2t_2-t_1-m)}{2\sigma^2}\right] \frac{1}{1 + \frac{(x-t_2)(t_1-m)}{2\sigma^2}} & \text{if } t_2 < y \end{cases} \quad (18)$$

Note that $F_g(-\infty) = 0$ and $F_p(\infty) = 0$. The N normalization factor is such that

$$\int f(y)dy = [F_g(t_1) - 0] + [F_e(t_2) - F_e(t_1)] + [0 - F_p(t_2)] = 1 \quad (19)$$

The f approximating function was constructed such that it is gaussian below the t_1 turning point, exponential between points t_1 and t_2 and polynomial ($1/x^2$) above t_2 . Note that the

asymptotic behavior of the Landau distribution is $1/x^2$, as well. f and its derivative with respect to y is continuous at t_1 and t_2 . It can be shown that both the parameters σ , t_1 and t_2 are closely linear functions of the most probable value m and they do not depend directly on $\beta\gamma$. m is function of $\beta\gamma$ and x which can be approximately factorized.

$$m(\beta\gamma, x) = \varepsilon(\beta\gamma) \cdot x(1 + ax) \quad (20)$$

The probability of measuring a given ADC value y is thus

$$P(y < y' < y + 1) = \int_y^{y+1} f(y) dy \quad (21)$$

Similarly, the probability of underflow (below threshold y_0) and overflow (above maximum measurable ADC y_1) can be given

$$P(y' < y_0) = \int_{-\infty}^{y_0} f(y) dy \quad P(y_1 < y') = \int_{y_1}^{\infty} f(y) dy \quad (22)$$

4.1 Determination of model parameters

The parameters of the approximate models can be extracted from theory, simulated and real data. Here only the determination from theory is demonstrated.

For a series of $\beta\gamma$ and x values 10^5 energy deposits have been generated according to the model described in Section 3. The corresponding values are $\beta\gamma = 0.4, 0.8, 1.6, 3.2, 6.4$ and 12.8 ; $x = 20, 40, \dots, 500 \mu\text{m}$ in $20 \mu\text{m}$ steps. The resulting ADC distributions have been fitted using the approximate model, giving m , σ , t_1 and t_2 values. The figures in Fig. 3. show that σ , $t_1 - m$ and $t_2 - m$ are closely linear function of the most probable value m . At the same time $m(x)$ can be well approximated by a second order polynomial with constant coefficients.

The corresponding $-2\log P$ values as function of the most probable deposit m for several measured ys are shown in Fig. 4. These numbers are tabulated and can be used for minimization with help of natural cubic spline interpolation.

5 Derivatives

For the calculation of the derivative vector and the approximate Hessian matrix it is enough to compute the first derivatives of the mean m with respect to the parameters $\mathbf{a}$. It can be decomposed into two parts:

$$\frac{\partial m}{\partial \mathbf{a}} = \frac{\partial m}{\partial l} \frac{\partial l}{\partial \mathbf{a}} = \varepsilon \frac{\partial l}{\partial \mathbf{a}}, \quad \frac{\partial m}{\partial \varepsilon} = l \quad (23)$$

Thus the only partial derivative to calculate is the dependence of the path-length inside a pixel on the parameters $\mathbf{a}$. These parameters are the coordinates of the (center of the) hit $\mathbf{H}$ with the average energy loss ε .

The path-length in a pixel is given by

$$l^2 = (\mathbf{C}_1 - \mathbf{C}_2)^2 \quad (24)$$

and the differentials are

$$l\delta l = (\mathbf{C}_1 - \mathbf{C}_2)(\delta\mathbf{C}'_1 - \delta\mathbf{C}'_2) \quad (25)$$

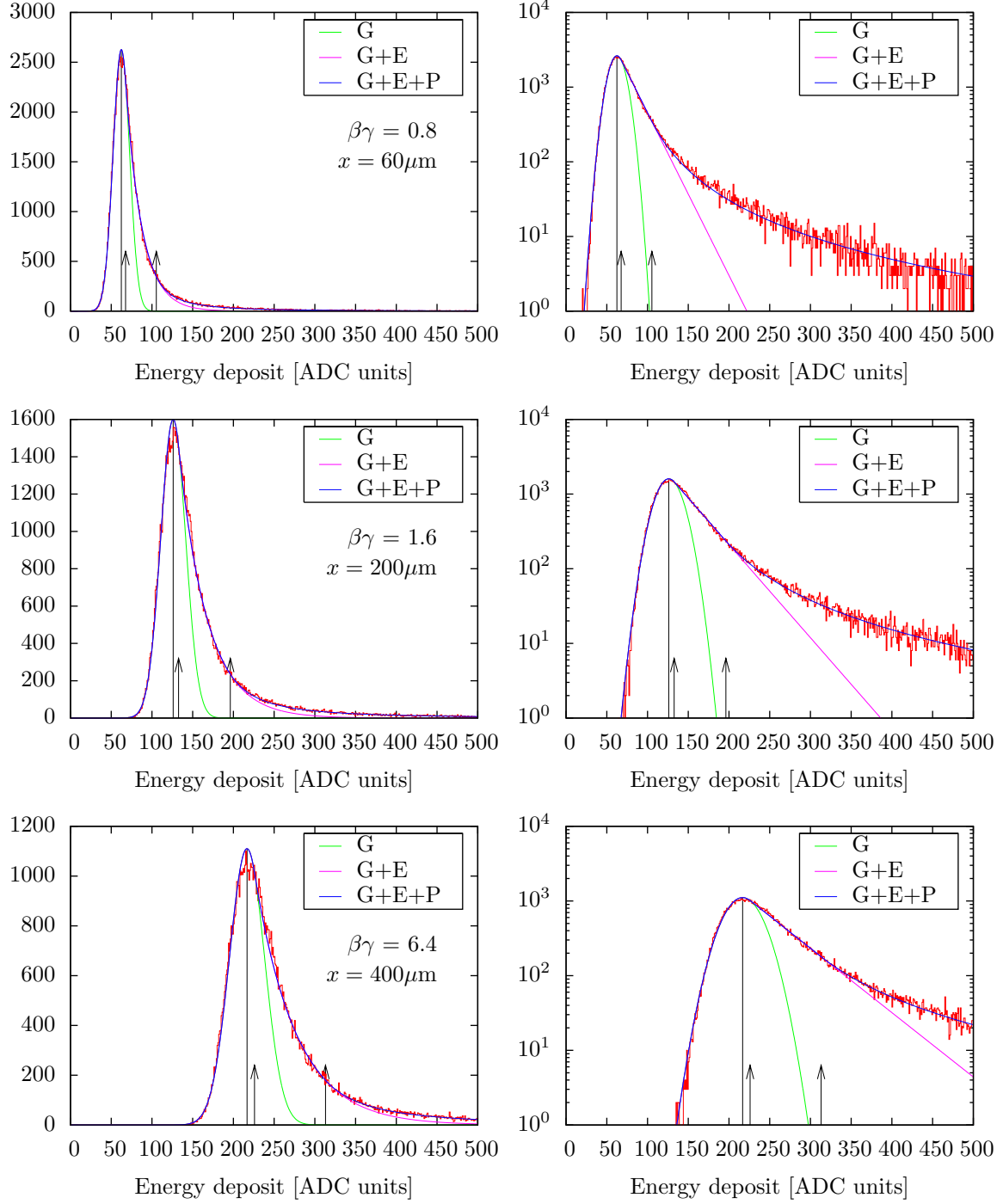


Figure 2: Energy loss distributions for several $\beta\gamma$ and x values. The approximating function is shown in blue (G+E+P), the gaussian only (G) and gaussian+exponential part (G+E) is also given.

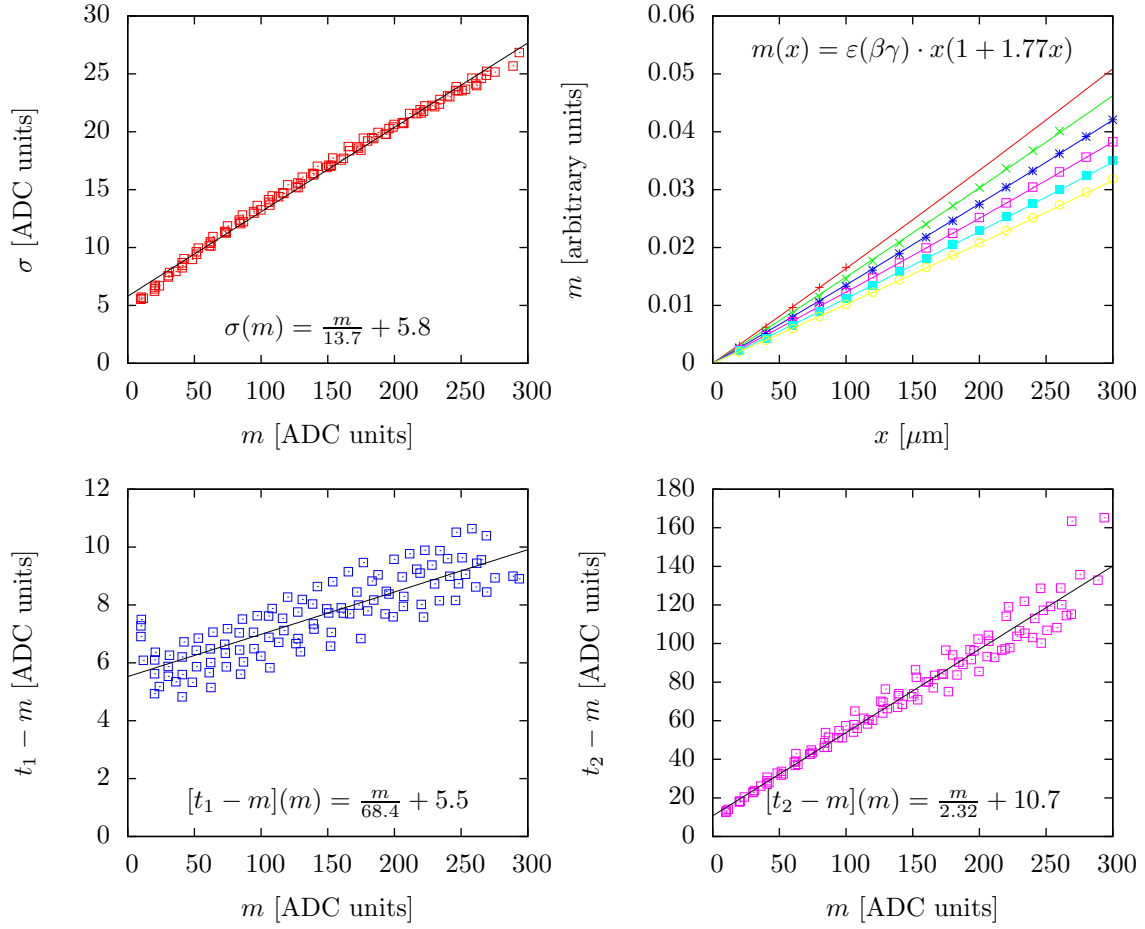


Figure 3: Dependence of parameters σ , the turning points t_1 and t_2 on the most probable deposit m and the dependence of m on the path x .

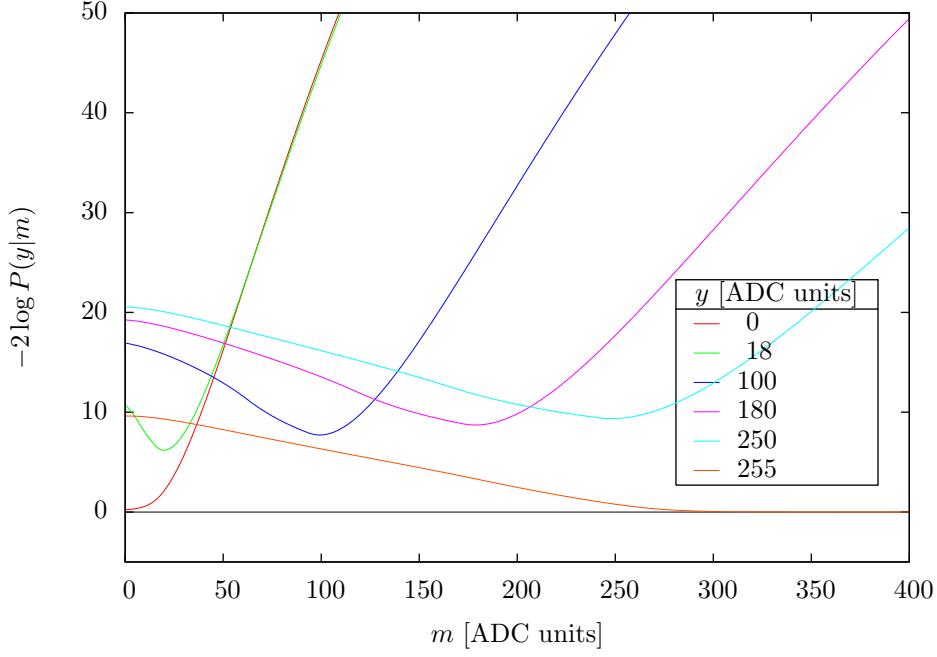


Figure 4: Log-likelihood values as function of the most probable deposit m for several measured y s.

By moving the hit, the endpoints of the segments are moved as well

$$\delta \mathbf{C}_i = \delta \mathbf{H} \quad (26)$$

The segment endpoints are not always movable in both directions. If the endpoint is a crossing point of the track and the cell boundary, it is only movable along the boundary. The segment endpoint is movable in all directions if it is itself a track endpoint. The partial derivative is

$$\frac{\partial l}{\partial \mathbf{H}} = \sum_i \frac{\mathbf{n}_i}{|\mathbf{n}_i \mathbf{e}|} \quad (27)$$

6 Minimization

The Levenberg-Marquardt method was modified in order to meet the needs of the log-likelihood estimation. The merit function is

$$''\chi^2'' = -2 \sum_i \log f(i; \mathbf{a}) \quad (28)$$

Here $\mathbf{a}$ denotes the list of parameters for the cluster (position, length, orientation, average energy loss). The index i runs for all the channels belonging to the cluster, f is the probability density function.

$$\beta_k \equiv -\frac{1}{2} \frac{\partial \chi^2}{\partial a_k} = \sum_i \frac{\partial \log f(i; \mathbf{a})}{\partial a_k} \quad (29)$$

$$\alpha_{kl} \equiv \frac{1}{2} \frac{\partial^2 \chi^2}{\partial a_k \partial a_l} = \sum_i \frac{\partial^2 \log f(i; \mathbf{a})}{\partial a_k \partial a_l} \quad (30)$$

The $\mathbf{R}$ Fisher information matrix is

$$R_{kl} = - \sum_i E \left[\frac{\partial^2 \log f(i; \mathbf{a})}{\partial a_k \partial a_l} \right] = - \sum_i \sum_{y_i} f(i; \mathbf{a}) \frac{\partial^2 \log f(i; \mathbf{a})}{\partial a_k \partial a_l} \quad (31)$$

$$\sigma_k^2 = R^{-1}_{kk} \quad (32)$$

The maximum likelihood estimator is consistent and asymptotically normally distributed with covariance matrix $\mathbf{R}^{-1}$. The diagonal elements simply give the corresponding σ_k^2 values.

7 Extraction of yields

With help of the known distribution functions $f_l(x)$ for different particle species, a fitting function $F(x_i)$ with weights a_l can be constructed. It is simply a linear combination.

$$F(x_i) = \sum_l a_l f_l(x_i) \quad (33)$$

$$(34)$$

The yields a_l of the particles can be extracted using an analytical solution of the least-squares method. The χ^2 is a sum for all the bins i :

$$\chi^2 = \sum_i \left[\frac{F(x_i) - y_i}{\sigma_i} \right]^2 \quad (35)$$

For many entries $\sigma_i^2 \approx y_i$. Its minimum can be determined by solving the equations

$$\sum_l \left[\sum_i \frac{f_l(x_i) f_l(x_i)}{y_i} \right] a_l = 1 \quad (36)$$

8 Results

Thousand special events have been generated with flat multiplicity distribution up to 500 particles. Each event has same number of π^+ , π^- , K^+ , K^- , p and $\bar{p}$. This choice enables to see the kaons and protons without having to generate a lot of minimum bias events where pions dominate. The particles have flat pseudo-rapidity distribution in the range $[-3, 3]$ and transverse momentum according to the $p_T \exp[-p_T^2/(2\sigma)^2]$ distribution with $\sigma = 400$ MeV/c. The reconstruction was based on pixel hit triplets. Only those pixel cluster have been used which do not lie on the boundaries of the silicon and has a clean cluster shape compatible with the predicted thrust of the track.

The estimated most probable energy loss in 300 μm path and the corresponding truncated mean energy loss as function of total momentum is given in Figures 5 and 6, both for positive and negative particles. The energy loss can be histogrammed and plotted in different momentum slices. Distributions for maximum likelihood and truncated mean methods at $p = 0.4, 0.8$ and 1.2 GeV/c are shown in Figures 7, 8 and 9, respectively. Separation powers for π -K and p -K are given which is defined as $2(m_2 - m_1)/(\sigma_1 + \sigma_2)$. The resolution of the pion peak is between 9 - 10%, it is narrower than the result of the truncated mean method by 2%.

Acknowledgement

The author wishes to thank to Sándor Hegyi and András László for helpful discussions. This work was supported by the Hungarian Scientific Research Fund (T 048898).

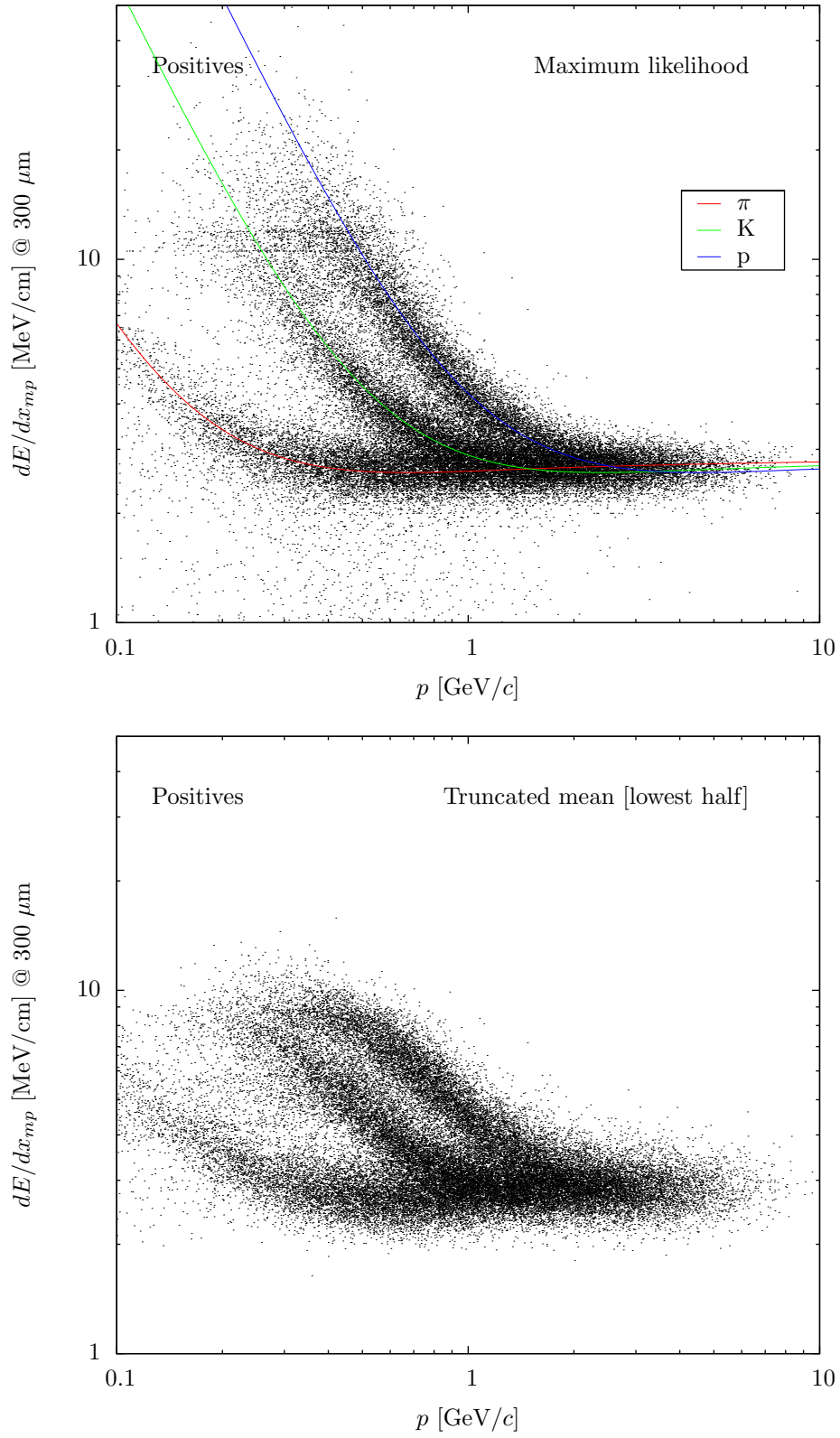


Figure 5: Most probable energy loss on $300\ \mu\text{m}$ path in silicon as function of total momentum p , for positive particles. Results with the maximum likelihood and truncated mean methods are shown for comparison.

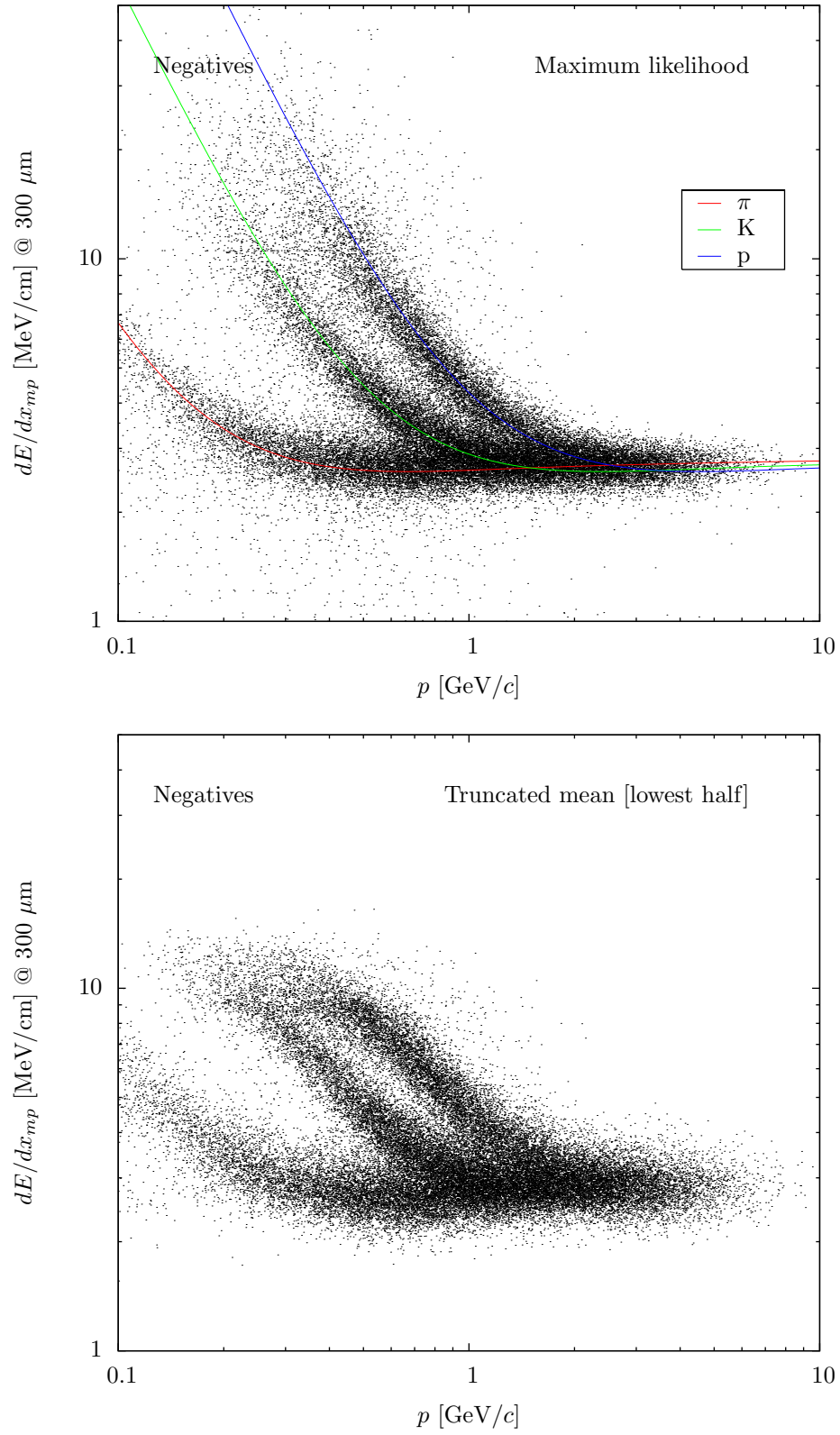


Figure 6: Most probable energy loss on 300 μm path in silicon as function of total momentum p , for negative particles. Results with the maximum likelihood and truncated mean methods are shown for comparison.

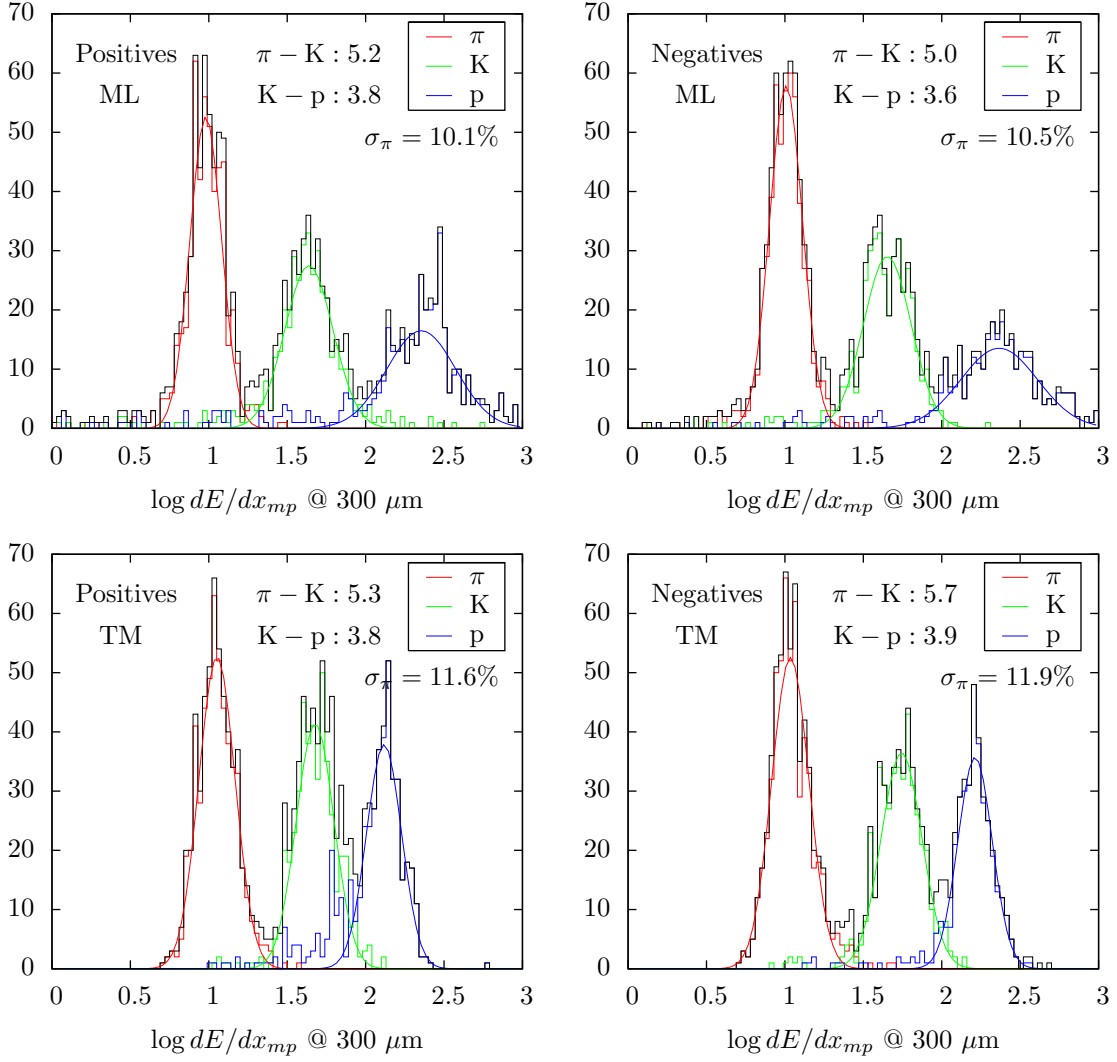


Figure 7: Distribution of $\log dE/dx_{mp}$ in momentum range $[0.40, 0.44] \text{ GeV}/c$ from maximum likelihood (ML) and truncated mean (TM) methods, for positive and negative particles. Separation powers are also given.

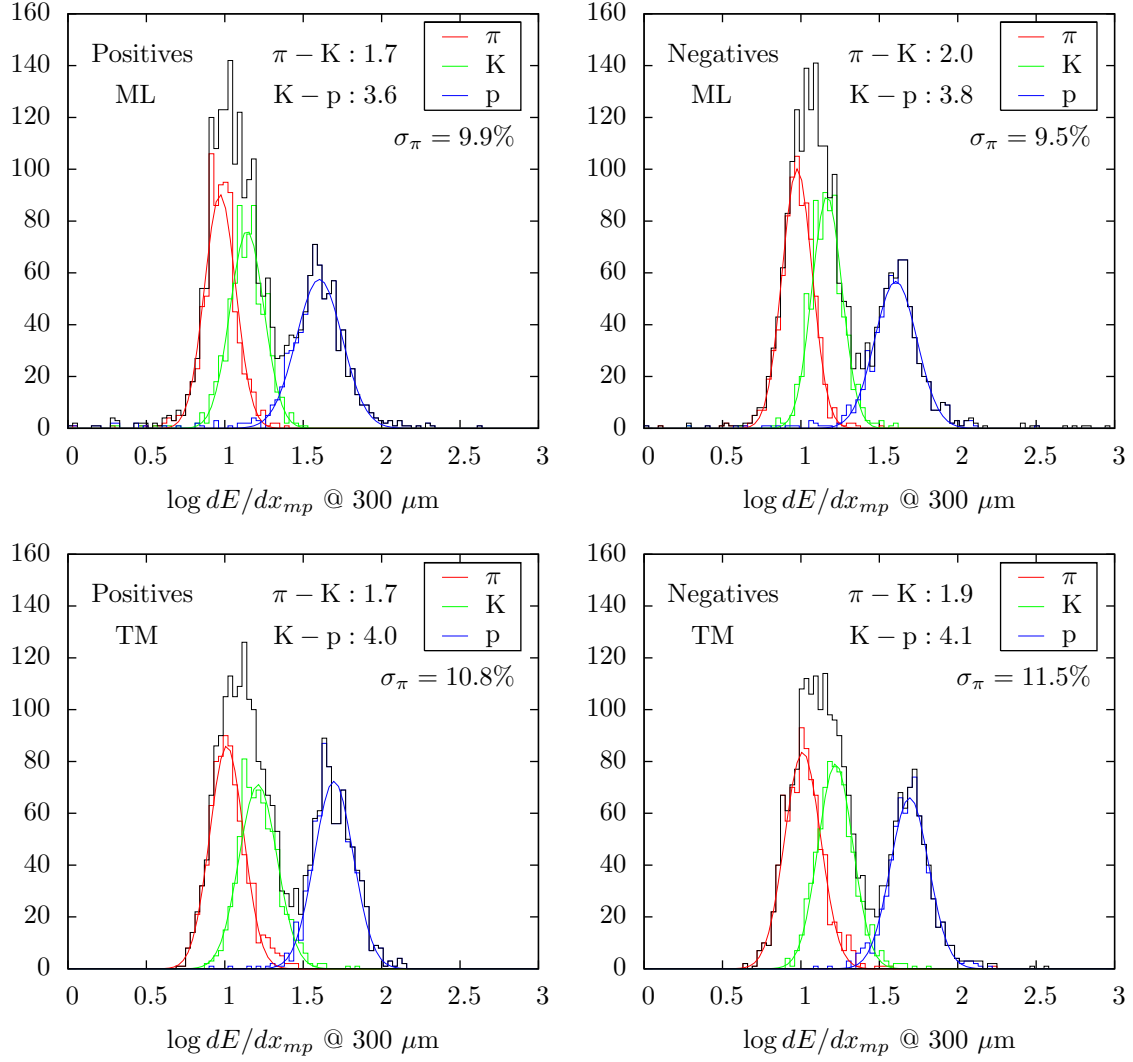


Figure 8: Distribution of $\log dE/dx_{mp}$ in momentum range $[0.80, 0.88] \text{ GeV}/c$ from maximum likelihood (ML) and truncated mean (TM) methods, for positive and negative particles. Separation powers are also given.

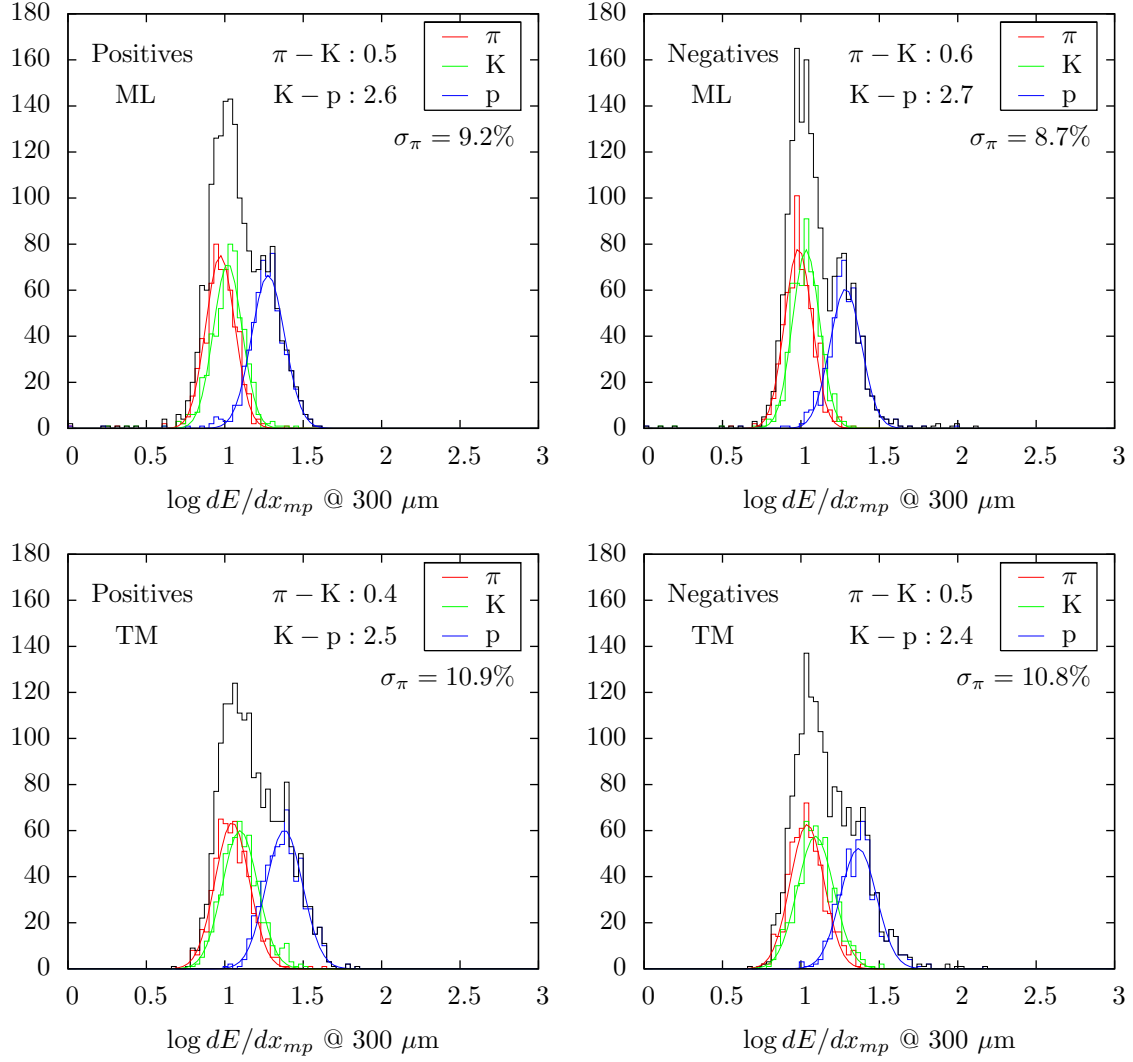


Figure 9: Distribution of $\log dE/dx_{mp}$ in momentum range $[1.20, 1.32] \text{ GeV}/c$ from maximum likelihood (ML) and truncated mean (TM) methods, for positive and negative particles. Separation powers are also given.

References

- [1] O. Ullaland, Nucl. Phys. Proc. Suppl. **125**, 90 (2003).
- [2] H. Yamamoto, hep-ex/9912024.
- [3] J. F. Bak *et al.*, Nucl. Phys. **B288**, 681 (1987).
- [4] S. Hancock, F. James, J. Movchet, P. G. Rancoita and L. VanRossum, Phys. Rev. A **28**, 615 (1983).
- [5] H. Bichsel, Rev. Mod. Phys. **60**, 663 (1988).
- [6] T. Anticic *et al.*, Nucl. Instrum. Meth. **A374**, 309 (1996).
- [7] S. Hancock, F. James, J. Movchet, P. G. Rancoita and L. Van Rossum, Nucl. Instr. Meth. **B1**, 16 (1984).
- [8] G. Hall, Nucl. Instr. Meth. **A220**, 356 (1984).