IMPROVEMENT OF PARTICLE IDENTIFICATION BY ENERGY LOSS IN A STACK OF SILICON DETECTORS

M. REGLER, R. FRÜHWIRTH AND M. FRIEDL

Institute of High Energy Physics of the Austrian Academy of Sciences,
Nikolsdorfer Gasse 18,
1050 Wien, Austria
E-mail: fru@hephy.oeaw.ac.at

We have studied the optimal use of information from silicon track detectors for particle identification, using a recent measurement with 4 GeV/c pions with low electronic noise. The raw data were modeled by the convolution of a density obtained by detailed simulation with a Gaussian that accounts for the measurement noise only. We observe excellent agreement with the data. Then a stack of fifteen detectors has been simulated from the experimental distribution. Using these data, the performance and robustness of some location estimators have been compared, including traditional methods such as truncated means and the Maximum Likelihood estimator, as well as a novel method, the optimal L-estimator.

1. Introduction

The study of energy loss in silicon has a long history. Due to the complex structure of the cross section (thresholds and multiple modes in the low energy domain, and the possibility of very large energy transfer in a single collision leading to a large skewness), the number of collisions is too small for the Central Limit Theorem. Many attempts for practicable approximations have been made with some well-known milestones like the work of Landau[1], Vavilov[2], and Shulek[3], along with the earlier work of Blunck and Leisegang[4] who used a Gaussian sum approximation.

The model used in this contribution is a probability density from a more refined simulation (without experimental noise), followed up by a convolution procedure, performed for us by Hans Bichsel[5,6].

2. The experimental data

2.1. *Detector and experimental set-up*

A prototype silicon detector module with double sided readout was built for an upgrade of the BELLE experiment[7] at KEK (Tsukuba, Japan) and studied in a beam test with $4\,\mathrm{GeV}/c$ pions in April 2005. The sensor is $300\,\mu\mathrm{m}$ thick and has AC-coupled strips with a readout pitch of $51\,\mu\mathrm{m}$ on either side. The strip signals are amplified by the low-noise APV25 front-end chip[8], which was originally developed for the CMS experiment at CERN (Geneva, Switzerland).

2.2. *Raw data analysis*

Out of the various measurements performed in that beam test, we will only show data of the p-side measured at “standard” conditions like perpendicular incidence and moderate over-depletion.

The raw strip data was processed with the usual pedestal subtraction and common mode correction algorithms, followed by a cluster finding procedure with two different thresholds: Once a seed strip with a signal above 5 times the RMS of the noise is found, neighbouring strips are added to the cluster as long as their signals exceed 3 times the RMS of the noise.

Statistical tests were carried out to ensure that the noise of individual strips is largely uncorrelated as expected. Consequently, the noise of a cluster signal depends on the number c of strips involved. Assuming identical RMS noise n_s on each strip (which is a good approximation of the real system), we obtain the RMS cluster noise n_c by $n_c = \sqrt{c}\,n_s$. In the data presented here, we found an average RMS strip noise of 522 e-h pairs and a mean cluster size of 2.49. Hence, the cluster noise is expected to be 824 e-h pairs in average. By modifying the cluster finding for noise evaluation such that the same number of strips are summed up but displaced from the signal cluster, we found an RMS cluster noise of 739 e-h pairs. Note that the data has been taken in a clean test-beam environment.

3. The electron-hole pair distribution

3.1. *The Bichsel model of the energy loss*

We thank H. Bichsel for providing us with a (pointwise) probability density function of the energy loss corresponding to our experimental setup (4 GeV/c pions, detector thickness of $300\,\mu\mathrm{m}$). The mode (most probable value) of the Bichsel density ist at $E_{\mathrm{mode}} = 82.64\,\mathrm{keV}$, corresponding to

22637 e-h pairs. We have "standardized" the Bichsel PDF by an affine transformation such that its location (mode) b is at zero and its scale (one quarter of the full width at half maximum) a is equal to 1. The standardized Bichsel density can be approximated quite well by the convolution of a Landau density (with mode $b = -0.26$ and scale $a = 0.791$) with a Gaussian density with $\mu = 0$ and $\sigma = 0.7724$. This explains the apparent success of the Landau-Gaussian-convolution in modeling energy loss data.

3.2. *Modeling the data by the Bichsel and by the Landau model*

The observed e-h pair counts are corrupted by the electronic noise. We have modeled the noise by Gaussian fluctuations of the counts with zero mean and a width to be determined from the data. The convolution of the Bichsel density with a Gaussian gives an excellent fit (Figure 1). The Bichsel density has mode $b = 19785$ and scale $a = 2058$. The standard deviation of the Gaussian is equal to $\sigma = 723$, which is in very good agreement with the estimated noise width of 739 e-h pairs. The mode of the signal distribution is approximately 10% lower than the theoretical expectation due to losses in the actual silicon sensor readout.

A Landau-Gaussian convolution gives an equally good fit. In this case the Landau density has location $b = 19256$, scale $a = 1610$, and the standard deviation of the Gaussian is equal to $\sigma = 1783$. Clearly the Bichsel

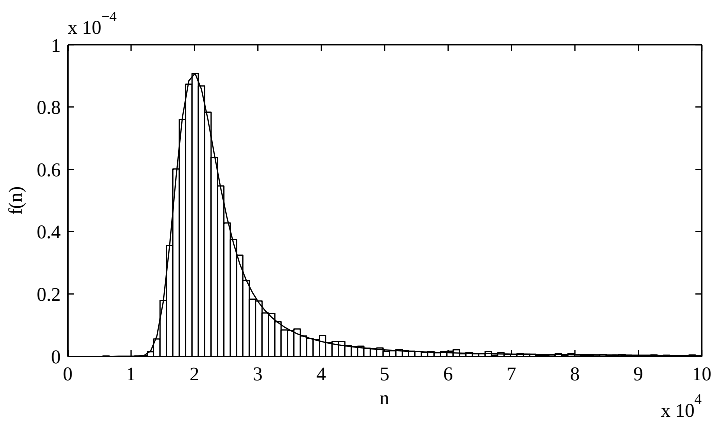

Figure 1. The freqency distribution of the experimental e-h pair counts, normalized to 1, and the density of the fitted Bichsel-Gaussian convolution.

model describes the data much better than the Landau model which requires the convolution of a Gaussian whose width is far beyond the estimated noise width in order to account for the "Shulek braodening". This makes the estimation of low noise from the data difficult.

4. Comparison of analysis methods

In this section we investigate some methods of estimating the location parameter of the parent distribution from a small sample. As we have at our disposal only observations by a single detector, we have artificially generated 50000 random samples by drawing from the observed distribution of *e*-*h* pair counts. We use a sample size of fifteen, which would be a typical sample size from a silicon tracker. Following the work of Talman[9], we concentrate on L-estimators, i.e. linear functions of the ordered sample values. Truncated means are special cases of the general L-estimator.

4.1. *Equivariant L-estimators of location*

Let $\vec{x} = (x_1 < \ldots < x_m)$ denote an ordered random sample of size m. An L-estimator of location $\mathcal{L}(\vec{x}) = \vec{w}^{\mathrm{T}}\vec{x}$ is equivariant[a] if and only if $\sum w_i = 1$. The mean of any subset of the ordered sample $\vec{x}$ and every truncated mean is equivariant. It should be noted that truncation is not tantamount to simply discarding part of the information — the order of the observations depends on all values in the sample. The mean of observations x_i to x_j is denoted by $\mathcal{L}_{(i,j)}$.

In the class of equivariant L-estimators of location with given expectation t there is one, denoted by $\mathcal{L}_{\mathrm{opt}}$, that has the smallest variance. Under the constraints $E(\mathcal{L}_{\mathrm{opt}}) = t$ and $\sum w_i = 1$ the optimal weights are given by $\vec{w}_{\mathrm{opt}} = \Delta^{-1} B(t\vec{e} - \vec{\mu})$, where $\vec{\mu} = E(\vec{x})$, $B = G(\vec{\mu}\vec{e}^{\mathrm{T}} - \vec{e}\vec{\mu}^{\mathrm{T}})G$, $G = [\mathrm{cov}(\vec{x})]^{-1}$, $\Delta = \vec{\mu}^{\mathrm{T}} B\vec{e}$ and $\vec{e}$ is a vector of ones. The optimal weights can be estimated from the data.

4.2. *Results from experimental data*

The separation of different parent distributions (particle types) is optimal if the ratio $\rho = \mathrm{mean}/\mathrm{std}$ of the location estimator is maximal. We have investigated the robustness of various truncated mean estimators and of the optimal L-estimator by contaminating the samples with various amounts

[a] A location estimator $l(\vec{x})$ is equivariant iff $l(a\vec{x} + b) = a\, l(\vec{x}) + b$.

of noise, uniformly distributed between 0 and 100,000. The ratio ρ drops by only about 5 percent with a 2%-contamination of noise.

Figure 2 shows the results for 2% contamination. The separation is optimal when the expectation of the L-estimator is equal to about 19200, somewhat below the mode of the parent distribution. The optimal L-estimator performs only slightly better than the best truncated mean estimators ($\mathcal{L}_{(2,5)}$ and $\mathcal{L}_{(2,7)}$). Besides being not equivariant, the maximum likelihood estimator is slightly worse than the best L-estimators and much slower to compute.

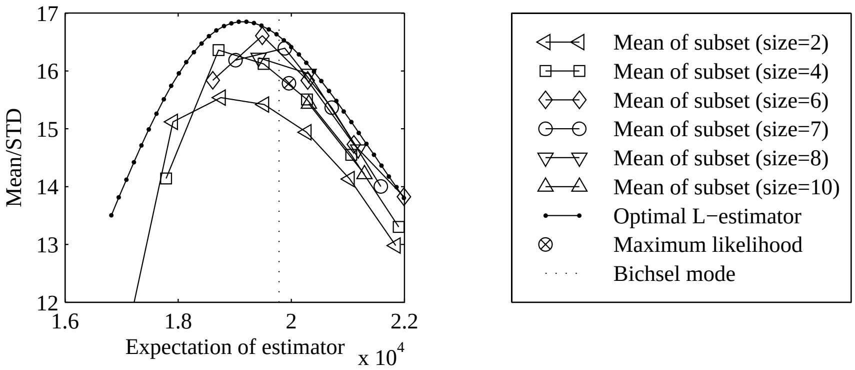

Figure 2. The ratio mean/std of truncated mean estimators,the optimal L-estimator and the maximum likelihood estimator. The abscissa is the expectation of the estimators. The truncated mean estimators of size m are computed by taking the average of observations k to $k+m-1$, starting at $k=1$. The dotted vertical line is at the mode of the Bichsel PDF.

References

1. L. Landau, *J. Phys. (USSR)* **VIII** , 201 (1944).
2. P.V. Vavilov, *Sov. Phys.-JETP* **5**, 749 (1957).
3. P. Shoulek et al., *J. Nucl. Phys. (USSR)* **IV**, 564 (1966).
4. O. Blunck und S. Leisegang, *Z. Phys.* **128**, 500 (1950).
5. H. Bichsel, *Rev. Mod. Phys.* **60**, 663 (1988).
6. H. Bichsel, A method to improve particle identification in TPCs and silicon detectors (in preparation).
7. A.Abashian *et al.*, *Nucl. Instr. Meth.* **A 479**, 117 (2002).
8. L.L.Jones *et al.*, *The APV25 Deep Submicron Chip for CMS Detectors*, CERN-99-09, p.162-166, CERN, Geneva (1999)
9. R. Talman, *Nucl. Instr. Meth.* **159**, 189 (1979).