СООБЩЕНИЯ
ОБЪЕДИНЕННОГО
ИНСТИТУТА
ЯДЕРНЫХ
ИССЛЕДОВАНИЙ

Дубна

95-468

E10-95-468

N.Chernov[1], E.Kolganova, G.Ososkov

OPTIMAL WEIGHTS FOR CIRCLE FITTING WITH DISCRETE GRANULAR DATA

[1]University of Alabama, Birmingham, AL 35294, USA

1995

1 Introduction

The problem of approximating of measured data points by a circle is one of importance for many data handling problem in high energy physics [1, 2, 3]. This is getting more important with inventing such modern detectors, as for example, RICH (Ring Imaging CHerenkov) , requiring in each event the parameter determination of tens and hundreds rings formed by Cherenkov photons. The information of the RICH detectors is read out via two-dimensional arrays each of about 50000 pads (cells), allowing the unambiguous reconstruction of single-photon hits [4].

Therefore the problem formulation for circle fitting to real RICH raw data is considerably different from how it was formulated in the above cited works, where a circle was measured in separate points (x_i, y_i) $i = 1, \ldots, n$. While practically all modern detectors having the discrete cell structure register, in fact, the energy dissipation produced by a passing elementary particle not in a single point, but in several adjacent cells where all this energy is distributed as one can see in fig.1.

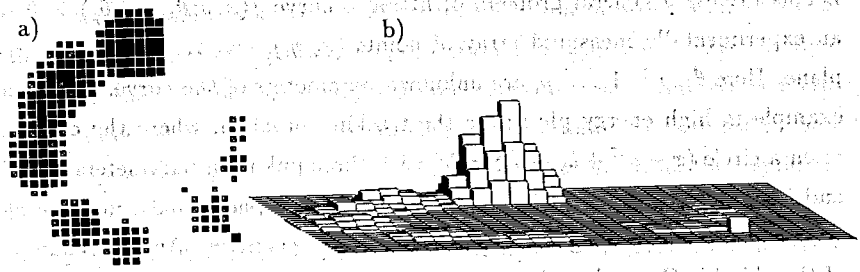

Figure 1: *2D (a) and 3D (b) images of simulated discretized signals of the same circle*

In this case a circle is to be fit not to separate points, but to clusters of adjacent cells. That is as a completely different problem. Besides of noted above the high occupancy of RICH detector, such its characteristics as the presence of background hits and appearance of several overlapping rings make inapplicable circle fitting methods cited above due to their noise sensitivity.

Объединенный институт
ядерных исследований

Even in the case of one circle one single point-outlier can distort considerably estimations of the circle center or radius.

In this paper it's shown that the efficient way to overcome these problems of the curve fitting is the *robust* fitting technique based on a reweighted least square method with optimally chosen weights.

Therefore two problems are to be solved:

- elaborate a new type of weights considered as optimal according to the maximum likelihood criterum;
- develop a robust version of one of effective algorithms for single circle fitting to granulated data;

The solution of above problems is presented here followed by results of comparative testing of corresponding algorithms.

2 Bimodal (bihorn) weight concept

Let us discuss a general problem of fitting a curve $f(x, y; \theta_1, \ldots, \theta_p) = 0$ to an experimentally measured array of points (x_i, y_i), $i = 1, \ldots, n$ on the xy-plane. Here θ_j, $j = 1, \ldots, p$, are unknown parameters of the curve. A typical example in high energy physics is the tracking problem, where the curve is often a circle $(x-a)^2 + (y-b)^2 = R^2$ with three unknown parameters a, b, R, and the experimental data (x_i, y_i) are bubbles on a photo taken in a bubble chamber, or electrical signals measured in a wire chamber. Another example of that kind is Cherenkov ring reconstruction in RICH detectors described in the next section.

The classical least square fit (LSF) is based on the minimization of the residual sum of squares (RSS):

$$\mathrm{RSS}(\theta_1, \ldots, \theta_p) = \sum_{i=1}^{n} d_i^2 \to \min \qquad (1)$$

Here d_i is the distance of the point (x_i, y_i) from the fitting curve. This method is optimal if the distances of the experimental points (x_i, y_i) from the actual

curve (track) are independent normally distributed random variables with a common variance.

In real experiments, however, the measured points are *not* normally distributed around the observed curve (track). There are two main reasons for the violation of the normality assumption in real experiments:

1. There are noisy points resulting from malfunctions of the detectors, side effects of the experiment and other tracks that happen to be close to the one we are estimating.

2. The technical limitations of the detectors, which prevent the possibility of measuring points right on the track or at arbitrary distance from it. Such are discrete (granular) detectors where the measured points are necessarily located at cites of a specific grid. In such detectors, often the amplitude of a signal is measured, too, so that the data are triples (x_i, y_i, a_i), where a_i stands for the amplitude of the signal at the point (x_i, y_i). Then the amplitude a_i is also discretized, e.g. by rounding to the nearest integer and cutting off too weak signal by a truncation rule $a_i \geq a_{\min} > 0$. A detector of that kind will be described in the next section.

These violations of the normality assumption often cause a complete breakdown of the LSF. There are modifications of the LSF which are less sensitive to changes in the distribution of points around the track (so called robust methods), or oriented to specific classes of such distributions. We will discuss here the most popular modifications of the LFS, which are based on the maximum likelihood estimates (M-estimates).

Assume that the distances of the measured points (x_i, y_i) from the track are independent identically distributed random variables, with common density $\rho(x)$. The maximum likelihood estimate requires the maximization of the so called logarithmic likelihood function

$$L(\theta_1, \ldots, \theta_p) = \sum_{i=1}^{n} \ln \rho(d_i) \to \max \qquad (2)$$

This function takes its maximum at a solution of the system of equations

$$0 = \frac{\partial L(\theta_1, \ldots, \theta_p)}{\partial \theta_j} = \sum_{i=1}^{n} \frac{\partial \ln \rho(d_i)}{\partial \theta_j}, \qquad j = 1, \ldots, p$$

This is equivalent to the weighted least square fit (WLSF)

$$\mathrm{WRSS}(\theta_1,\dots,\theta_p)=\sum_{i=1}^{n} w(d_i)\,d_i^2 \to \min \tag{3}$$

with the weight function $w(x)$ given by

$$w(x)=-\frac{1}{x}\frac{\partial}{\partial x}\ln\rho(x)=-\frac{\rho'(x)}{x\rho(x)} \tag{4}$$

Note that if $\rho(x)$ is a gaussian with zero mean, then $w(x)$ =const.

Typical density functions $\rho(x)$ for experiments with noisy data are mixtures of normal distributions (for signals produced by the track) and uniform distributions (for noisy signals):

$$\rho(x)=(1-c)\,g(x)+c\,u_a(x) \tag{5}$$

where $g(x)=(2\pi\sigma^2)^{-1/2}\,e^{-x^2/2\sigma^2}$ is a gaussian density with zero mean, c is the rate of noisy signals in the sample (in other words, $(1-c):c$ is the signal-noise ratio), and $u_a(x)$ is the uniform density on a large segment, $[-a,a]$, with $a\gg\sigma$, i.e. $u(x)=(2a)^{-1}$ for $|x|\le a$. We then conclude that the optimal weight function, according to (4), is

$$w(x)=\sigma^{-2}\left(1+\frac{c}{1-c}\cdot\frac{u_a(x)}{g(x)}\right)^{-1} \tag{6}$$

This is a unimodal plateau-like function, symmetric about $x=0$, nearly constant $\approx\sigma^{-2}$ for small x and rapidly decreasing to zero as $|x|$ becomes large. This kind of weight functions are widely used for fitting curves to contaminated data (i.e., with noisy signals). Some popular functions are mentioned below in Section 4.

We now turn to the main point of our discussion. The discretization of the measured signals with a cut-off rule for weak signals means that the main component $g(x)$ of the mixture (5) is no longer gaussian. Most importantly, due to the truncation rule, it will decrease to zero more abruptly for such values of x that the measured signals become weak enough to be cut off. One can think of a bell-shaped gaussian curve whose tails are chopped off.

Accordingly, we assume that

$$\rho_d(x)=(1-c)\,g_d(x)+c\,u_a(x) \tag{7}$$

where, as compared to (5), we have $g_d(x)=\Delta(x)\,g(x)$. Here $\Delta(x)$ is a plateau on an interval $|x|\le d\sigma$, rapidly decaying to zero for $x>d\sigma$ to chop off the tails of $g(x)$. The value of d is specified by the truncation rule. For instance,

$$\Delta(x)=\begin{cases}1 & \text{if } |x|\le d\sigma\\ (1+D((\sigma d)^{-1}x-1)^2)^{-1} & \text{if } |x|>d\sigma\end{cases}$$

with, say, $d=3$ and a large $D\ge 1$.

Steep descent of the function $g_d(x)$ to zero for the values of x just outside the segment $[-d\sigma,d\sigma]$ cause the derivative of $\rho(x)$ to grow for these values of x, and grow significantly. According to (4), the weight function $w(x)$ can grow for these values of x also. Of course, the weight function will then drop to zero, as the noisy component in (7) becomes dominant. Thus, the function $w(x)$ may have two pronounced peaks near $x=\pm d\sigma$, with a relatively high plateau $x\approx\sigma^{-2}$ in between.

As one can see, bimodal weight functions become optimal for least square fit to signals measured by discrete, granular detectors applying truncation rules for weak signals. This is a general concept, independent of the fitting curves (lines, circles, etc.) and technical specifications of detectors.

Intuitively, the value of the weight function $w(x)$ in (3) can be interpreted as a force by which the point at distance x from the fitting curve 'attracts' it. In the 'pure' LSF (1), all the data points are 'equally attractive'. The robust plateau-like function (6) makes all the points in a strip around the fitting curve nearly equally attractive, and the rest neutral. The logic is simple – the points in the strip belong to the track, and the rest are just noise.

Bimodal weight functions, as compared to plateau-like ones, make the points on the sides of the main strip *more attractive* than those in the middle of the strip, leaving the points outside the strip neutral. An additional attracting force assigned to the points on the sides of the main strip (poorly fitted by the curve) gives the fitting curve better chances to adjust itself. It becomes more flexible and less likely to fall into a wrong local minimum of the function (3). Our numerical experiment reported below in Section 4 demonstrates the advantages of bimodal weight functions over various unimodal and plateau-like ones.

3 Optimal weights for the LSF procedure

Here we deduce an analytical formula for the optimal weight function for the least-square fit (LSF) procedure, assuming the model described in the previous section.

To make our main argument simpler, we first discuss a one-dimensional analog of our cellular detector. This analog consists of identical bars ('one-dimensional cells') lined up in a row. A track hits this detector at some point, with coordinate a. This hit results in energy production in the neighboring cells. The total amount of released energy, A, is a random variable with some probability density $f(A)$, $0 < A < \infty$. The distribution of the released energy among cells is a gaussian centered at a with a constant variance σ_0^2.

For any particular cell (i.e., bar) with center x, the energy produced by the hit at point a is then

$$\bar{B} = \frac{A}{\sqrt{2\pi\sigma_0^2}} \exp\left[-\frac{(x-a)^2}{2\sigma_0^2}\right] \cdot dx \tag{8}$$

where $dx = 1$ is, as usual, the size of the cell. The energy actually measured in this cell differs from $\bar{B}$ by a small amount (error). Let us assume that the measured energy, B, is a normally distributed random variable with mean $\bar{B}$ and a constant standard deviation $\sigma_1 > 0$. Thus, the probability density of B is

$$\rho_{a,A}^{(0)}(x,B) = \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left[-\frac{(B-\bar{B})^2}{2\sigma_1^2}\right] \tag{9}$$

where $\bar{B}$ is given previously by (8).

In the density formula (9) we have an extra parameter, A, which is unknown in the experiment, so we have to get rid of it. We will 'integrate it away' as follows:

$$\rho_a^{(0)}(x,B) = \int_0^\infty \rho_{a,A}^{(0)}(x,B) f(A)\, dA \tag{10}$$

This is the probability density for measuring an energy of B in the cell with center x, given a hit at a. This density is based on idealized, noise-free model. The noisy signals can be recorded in any cell, so we have to add a uniform in x distribution:

$$\rho_a(x,B) = \text{const} \cdot \left(\rho_a^{(0)}(x,B) + \rho^{(n)}(B)\right) \tag{11}$$

where the constant factor is necessary for normalization. Here $\rho^{(n)}(B)$ is the probability of a noisy signal of amplitude B in an arbitrary cell.

Naturally, $\rho_a(x,B) = \rho(x-a,B)$, i.e. the function (11) depends only on the difference $x-a$. We now get the optimal weight formula for the least-square procedure:

$$w_B(x) = -\frac{\rho'(x,B)}{x\rho(x,B)} = -\frac{1}{x}\frac{d}{dx}\ln\rho(x,B) \tag{12}$$

Note that this function depends on B, the amplitude of the measured energy in the cell.

Before attempting a theoretical study of the function (12), we will find a good approximation to it. The integration in (10) obviously makes the exact value of $w_B(x)$ hardly available. Certainly, one does not want to evaluate the integral (10) for every signal in real mass data processing. Either we have to tabulate $w_B(x)$ for practical use, or we can find a satisfactory approximation. Fortunately, a good and simple approximation to $w_B(x)$ exists.

We observe that the factor $\rho_{a,A}^{(0)}(x,B)$ in (10), as a function of A, is a classical bell-shaped gaussian. Gaussian densities quickly converge to zero as the argument moves away from the mean value. Here the mean value, $\bar{A}$, can be found from the equation $B = \bar{B}$, which gives

$$\bar{A} = B \cdot \sqrt{2\pi\sigma_0^2} \exp\left[\frac{(x-a)^2}{2\sigma_0^2}\right] \tag{13}$$

We now can replace the function $f(A)$ in (10) by its value $f(\bar{A})$ at $\bar{A}$. The integral in (10) can be then computed approximately:

$$\begin{aligned} \rho_a^{(0)}(x,B) &\approx \int_0^\infty \rho_{a,A}^{(0)}(x,B) f(\bar{A})\, dA = \bar{A} f(\bar{A})/B \\ &= f\left(B \cdot \sqrt{2\pi\sigma_0^2} \cdot e^{(x-a)^2/2\sigma_0^2}\right) \cdot \sqrt{2\pi\sigma_0^2} \cdot e^{(x-a)^2/2\sigma_0^2} \end{aligned} \tag{14}$$

(Note that the parameter σ_1 is missing from this formula, so we no longer need its value!)

A direct numerical test, with an exponential density $f(A)$, shows that the approximation (14) is very accurate, see also below. It makes it possible to express the weight function $w_B(x)$ explicitly through the density $f(A)$ and its

first derivative $f'(A)$:

$$w_B(x) = \frac{E\,f(EB) + E^2 B\,f'(EB)}{E\,f(EB) + \rho^{(n)}(B)}$$

where $E = \sqrt{2\pi\sigma_0^2}\cdot e^{x^2/2\sigma_0^2}$.

We will predict the shape of the graph of $w_B(x)$. For concreteness, we now assume that the distribution of the total energy released by a hit, $f(A)$, is an exponential one with mean A_0:

$$f(A) = A_0^{-1}\cdot e^{-A/A_0}, \qquad A > 0 \tag{15}$$

Let us first set $\rho^{(n)}(B) = 0$ in (11), so that to neglect noisy signals. Then the weight function is

$$w_B(x) = BA_0^{-1}\sqrt{2\pi\sigma_0^2}\cdot e^{x^2/2\sigma_0^2} - 1$$

This function rapidly grows as x increases. It is positive for moderately large values of B. It has a local minimum at $x = 0$. For small values of x, the function $w_B(x)$ may drop below zero if B is small enough. These phenomena have clear interpretations. Indeed, if the measured signal B is large, that cell is likely to be very close to the hit, and then positive weight function 'attracts' the estimated coordinate of the hit, a, to the cell x. If B is small, it is then unlikely that the cell with the measured energy B is very close to the hit. Instead, it is likely to be at distance σ_0 to $3\sigma_0$ from the hit. Thus, low signals have to 'repulse' the estimated coordinate of the hit from their immediate neighborhoods, but still attract it outside of those neighborhoods. This makes $w_B(x)$ negative for small x and positive for large x.

Taking into account noisy signals, i.e. assuming $\rho^{(n)}(B) > 0$, will hardly affect the shape of $w_B(x)$ for small x, where the 'pure' component $\rho^{(0)}$ dominates over the noisy one $\rho^{(n)}$ in (11). So, $w_B(x)$ still can have a local minimum at $x = 0$. On the contrary, for large x the noisy component becomes dominant in (11). It is a constant function in x, so it forces $w_B(x)$ to vanish, according to (12). We emphasize these two observations: $w_B(x)$ has a local minimum at $x = 0$ and decreases approaching zero as x increases. Of course, $w_B(x)$ must be an even function, symmetric about $x = 0$. The simplest shape meeting these criteria is a bimodal (bihorn) one discussed in Section 2.

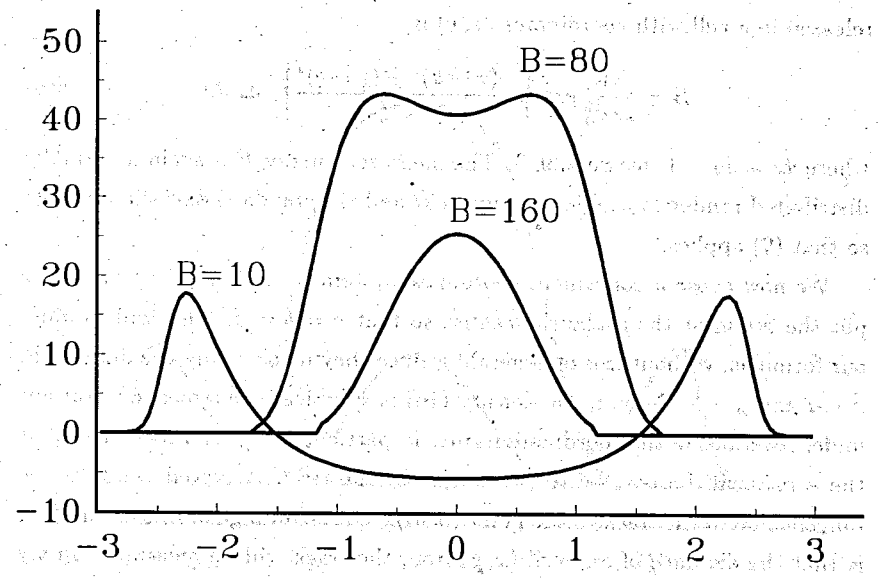

Figure 2: *The weight function $w_B(x)$ for different amplitude values, one-dimensional case.*

As an illustration, the graphs of the weight function $w_B(x)$ are shown on Fig.2. They are plotted by a computer assuming that $f(A)$ is an exponential function (15), with $A_0 = 200$, $\sigma_0 = 1$, $\rho^{(n)}(B) = 0.01$. Three tested values of B clearly show the tendency of $w_B(x)$ to increase and flatten around zero, as B goes up. These plots were calculated by both direct numerical integration in (11) and approximation (14). The difference was so small that the two curves practically coincided for every tested value of B.

We now turn to the real, two-dimensional cellular detectors described in the previous section. An avalanche of released energy can now occur at any point on the track. Let $x = a, y = b$ be the coordinates of the center of an avalanche. Again, we denote the total released energy by A. The value of A is random, with some probability density $f(A)$, $A > 0$. The released energy is distributed among the neighboring cells by a (two-dimensional) gaussian, which is spherically symmetric with covariance σ_0^2 in each coordinate. Therefore, the energy

released in a cell with coordinates (x, y) is

$$\bar{B} = \frac{A}{2\pi\sigma_0^2} \exp\left[-\frac{(x-a)^2 + (y-b)^2}{2\sigma_0^2}\right] \cdot dx\, dy \qquad (16)$$

where $dx = dy = 1$, see Section 3. The measured energy B is again a normally distributed random variable with mean $\bar{B}$ and the standard deviation $\sigma_1 > 0$, so that (9) applies.

We now chose a convenient coordinate system as follows. First of all, we put the origin at the avalanche center, so that $a = b = 0$. This will simplify our formulas, without loss of generality, since they involve only the differences $x - a$ and $y - b$. Second, the density (16) is spherically symmetric, invariant under rotations of the coordinate frame. In particular, we can rotate it so that the y axis will be parallel to the track. (If the track is curved, we take the tangent line to the track at the point $(0, 0)$). The advantage of this orientation is that the distance of any cell (x, y) from the track will be measured simply by $|x|$. The y coordinate will be completely irrelevant.

In the new coordinate system the equation (16) takes form

$$\bar{B} = \frac{A}{2\pi\sigma_0^2} \exp\left[-\frac{x^2 + y^2}{2\sigma_0^2}\right] \qquad (17)$$

Compared to the one-dimensional problem discussed before, we now have two extra parameters, A and y, which are unknown in the experiment (recall, that y represents a coordinate of the cell in the *new* coordinate system attached to the unknown track!). So we have to get rid of both y and A. We will again 'integrate them away' as follows:

$$\rho^{(0)}(x, B) = \int_{-\infty}^{\infty} \int_0^{\infty} f(A)\rho_A^{(0)}(x, B)\, dA\, dy \qquad (18)$$

where $\rho_A^{(0)}(x, B)$ is given by (9) with $\bar{B}$ given by (17). This is the probability density for measuring an energy of B in a cell whose distance from the track is x. The inner integral can be accurately approximated by the same trick as in (14). As a result, we get

$$\rho^{(0)}(x, B) \approx \int_{-\infty}^{\infty} Ef(EB)\, dy \quad \text{where} \quad E = 2\pi\sigma_0^2 \exp\left[\frac{x^2 + y^2}{2\sigma_0^2}\right] \qquad (19)$$

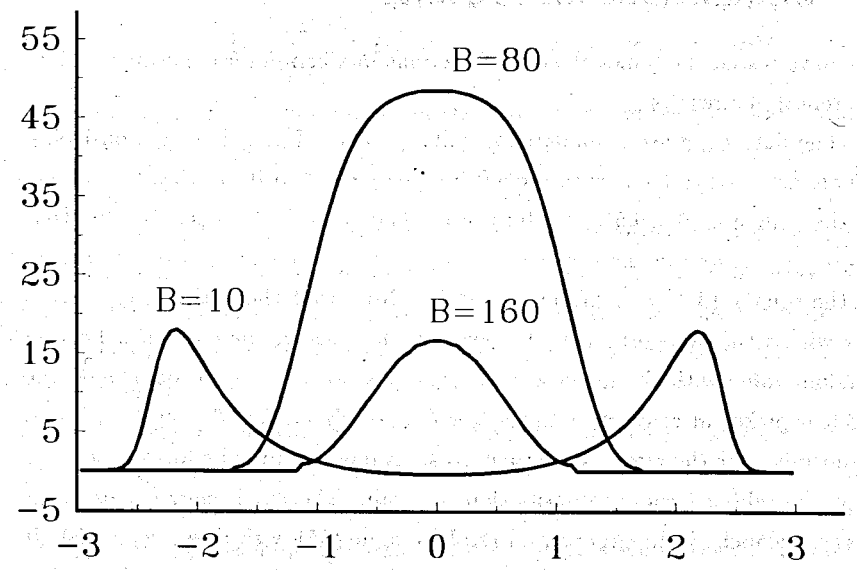

Figure 3: *The weight function $w_B(x)$ for different amplitude values, two-dimensional case.*

The remaining integral hardly admits a simple and accurate approximation, unfortunately.

As before, taking into account noisy signals results in

$$\rho(x, B) = \text{const} \cdot \left(\rho^{(0)}(x, B) + \rho^{(n)}(B)\right) \qquad (20)$$

const being the normalization factor. The optimal weight function $w_B(x)$ for the least-square procedure is then computed by the same formula (12). It still depends on the measured energy B in the cell and the distance x from the cell to the estimated track.

The graphs of the weight function $w_B(x)$ are shown on Fig. 3 for three values of B. As before, the density $f(A)$ was assumed to be exponential with $A_0 = 200$, $\sigma_0 = 1$, $\rho^{(n)}(B) = 0.01$. Compared to Fig. 2, one can notice that the drops below zero are now much shorter and less pronounced. This is a reassuring news, since negative weights can cause computational troubles at times.

4 Simulations and results

We have tested the bimodal weight functions in a computer experiment. Here we report the results.

The data have been simulated as follows. The cellular detector consists of square cells of size 1×1, with the left lower corner at $(0,0)$, so that the centers of the cells are $(0.5,0.5)$, $(1.5,0.5)$, etc. The track is a circle. Its radius is chosen at random in the range $8 < R < 12$, and its center is randomly placed in the square $13 < x < 17$, $13 < y < 17$. Note that the entire circle always lies within the detector. The number of avalanches on the circle is a Poisson random value with the mean $\lambda = 11$. The position of every avalanche on the circle is picked at random, with the angular coordinate $0 < \theta < 2\pi$ distributed uniformly over the circle. Of course, some avalanches may be located so close that the released energy distributions overlap. The total energy released at every avalanche is an exponential random value (15) with mean $A_0 = 200$. It is distributed among neighboring cells according to a two-dimensional gaussian (16) with $\sigma_0 = 1$. An energy released in any particular cell is only recorded if it exceeds 3 units, i.e. the truncation rule $B > 3$ is applied.

Some noisy signals are then added in the area of the detector $0 < x < 30$, $0 < y < 30$. The number of noisy signals was equal to the number of previously generated signals along the circle, so that the signal/noise ratio was 1. The positions of noisy signals were distributed uniformly within the above area, and the energy for every noisy signal was an exponential random value with mean 5. (It might seem too low, compared to the mean energy release $A_0 = 200$. Note, however, that the latter is typically spread over about $4 \times 4 = 16$ neighboring cells, so that the average signal per cell was about 12.) Again, the cut-off rule $B > 3$ is applied.

A typical data sample in this experiment is histogrammed and shown on Fig. 1. Here the height of a bar over a cell represents the amplitude of energy measured in it. One can barely see a circle fitting these data!

The simulated data are then fit by a circle through weighted least square procedure. Specifically, we used Crawford's algorithm based on minimizing

the function

$$L(a,b,R) = \sum_{i=1}^{n} w_i (x_i^2 + y_i^2 - 2ax_i - 2by_i + a^2 + b^2 - R^2)^2 \qquad (21)$$

where (x_i, y_i) are the coordinates of the signals, w_i are weights, and (a, b, R) are the circle parameters. The great advantage of this method is that, if we fix the weights w_i, it is a *linear* regression in three parameters: a, b, and $c = a^2 + b^2 - R^2$. Thus, the computation of a, b, R involves only a 3×3 system of linear equations, provided the weights are given. This method works very accurately if the data points (x_i, y_i) are spread over the entire circle (as opposed to tracking, where data points are normally located along very small arcs), see [2] for detail.

Of course, the weights w_i in (21) cannot be fixed, they depend on the distance of the points (x_i, y_i) from the track, and also have to depend significantly on the amplitudes of the signals B_i. A standard iterative procedure was used to compute w_i based on the circle found at the previous iteration. The initial circle is picked randomly, with center $13 \le a, b \le 17$ and radius $8 \le R \le 12$. This is a pretty bad approximation, since these values are typically about two cells off the actual values of a, b, R. Much more accurate initial approximations can be easily found, of course. However, we intentionally tested the algorithm by feeding it with poor initial values of a, b, R. Then the iterative procedure works until convergence, but not longer than 10 iterations.

The quality of the algorithm is characterized by three parameters. The first is the probability of a complete failure, P_f, which occurs if the estimated values of a and b are off by ≥ 1 (here 1 is the size of a cell) or the estimated value of R is off by ≥ 0.5. The other two characteristics are the root mean square error in the estimates of a and b, denoted by $D_{a,b}$, and the same error in R, denoted by D_R. The values of $D_{a,b}$ and D_R are computed based on non-failing runs only (in the above sense).

The weights w_i were computed by various methods. We tried both unimodal and bimodal weight functions. Two unimodal functions we tested were famous Huber's function [5]

$$w^H(x) = \begin{cases} 1 & \text{if } |x| < \sigma_0 \\ 2\sigma_0 x^{-1} - \sigma_0^2 x^{-2} & \text{otherwise} \end{cases}$$

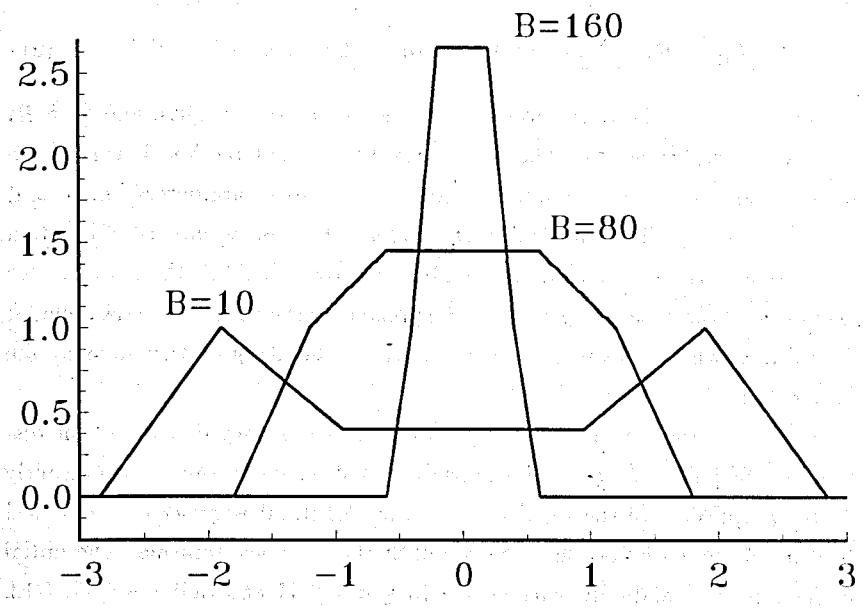

Figure 4: *The weight function $w_B^{(1)}(x)$ for different amplitude values.*

Tukey's biweight [6]

$$w^T(x) = \begin{cases} (1-(x/c\sigma_0)^2)^2 & \text{if } |x| < c\sigma_0 \\ 0 & \text{otherwise} \end{cases}$$

with $3 \le c \le 6$, and Andrews's sine [7]

$$w^A(x) = \begin{cases} \sin(x/c\sigma_0) & \text{if } |x| < c\sigma_0\pi \\ 0 & \text{otherwise} \end{cases}$$

with $c = 2$.

We tested various simple (piecewise linear) approximations to the bimodal optimal weight functions found in the previous section. The following one was the most successful:

$$w_B^{(1)}(x) = \begin{cases} 1 + 0.015 \cdot (B-50) & \text{if } |x| < r\sigma_0 \\ 1 - 0.015 \cdot (B-50) \cdot (x/r\sigma_0 - 2) & \text{if } r\sigma_0 \le |x| < 2r\sigma_0 \\ 3 - x/r\sigma_0 & \text{if } 2r\sigma_0 \le |x| < 3r\sigma_0 \\ 0 & \text{if } 3r\sigma_0 \le |x| \end{cases}$$

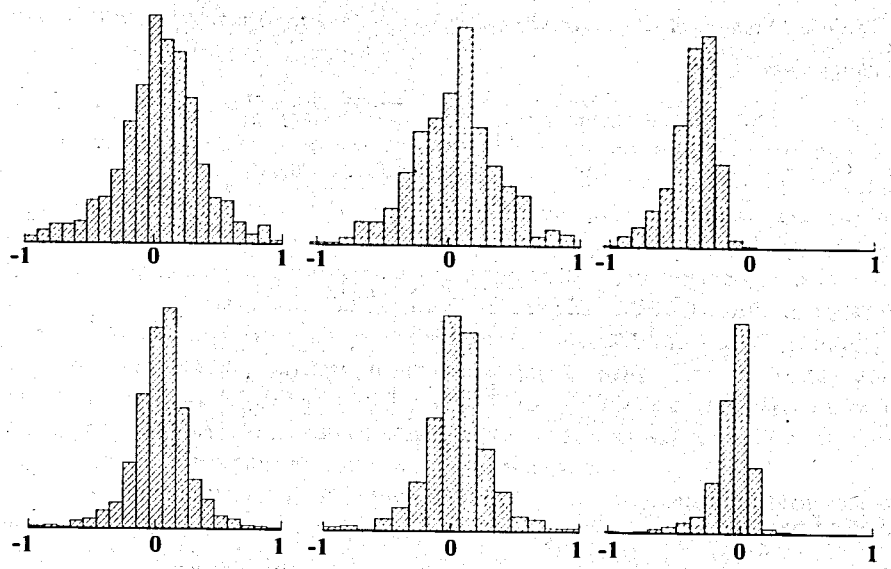

Figure 5: *The experimetal distributions of deviations between exact (modelled) values of parameters a, b, R and their estimations with Huber's weights (upper figures) and bimodal ones (bottom figures)*

were $r = 1 - B/200$. The graphs of this weight function $w_B^{(1)}(x)$ are shown on Fig.4.

The numerical results of our single circle experiments are summarized in the table 1.

Only Tukey's unimodal biweight with $c = 4$ stands the competition to some extent, other unimodal functions are clearly poorer than our bimodal one. The first line corresponding to the non-weighted least square fit (1) is included just to demonstrate the necessity of robust algorithms for accurate processing of noisy data. The experimetal distributions of deviations between exact (modelled) values of parameters a, b, R and their estimations with Huber's weights and bimodal ones are shown on fig. 5.

Table 1: *Numerical characteristics of four algorithms for circle fitting to simulated data.*

method	P_f	$D_{a,b}$	D_R
LSF	0.7570	0.231	0.420
Huber	0.1702	0.252	0.321
Andrew	0.0771	0.250	0.248
Tukey ($c = 6$)	0.0578	0.242	0.205
Tukey ($c = 4$)	0.0415	0.232	0.144
Tukey ($c = 3$)	0.0821	0.240	0.157
bimodal	0.0513	0.219	0.130

References

[1] J.F.Crawford, Mucl.Instr. and Meth. **211** (1983), 223.

[2] N.I.Chernov, G.A.Ososkov, Computer Phys. Comm. 33 (1984) 329.

[3] V.Karimaki, Mucl.Instr. and Meth. **A305** (1991) 185.

[4] T.Ullrich et al, *CERES/NA45 Contribution to NIM (Proceedings of RICH'95 in Uppsala)* to be published

[5] P.Huber. *Robust statistics*, J.Willey & Sons, N-Y (1981).

[6] F.Mosteller, W.Tukey. *Data analysis and regression: a second course in statistics*, Addison - Wesley (1977).

[7] D.F.Andrews, Technometrics, **16** (1974) 523.

[8] F.James and M.Roos MINUIT - Users Guide Program Library D506. CERN, 1989.

[9] G.Ososkov, G.A.Agakishiev, Yu.Panebratzev, Proceed. of the 6-th Intern. Conference on Phys.Comp., PC'94, European Phys. Society, Geneva, 1994, 361.

Received by Publishing Department
on November 17, 1995.