The State Scientific Center of Russia
The Budker Institute of Nuclear Physics
SB RAS

A.D. Bukin

OPTIMAL ROTATION PROCEDURE

Budker INP 97-50

НОВОСИБИРСК

Optimal rotation procedure

A.D.Bukin

Budker Institute of Nuclear Physics SB RAS
630090 Novosibirsk, Russia

Abstract

One of the possible procedures for the kinematic reconstruction of multiparticle events is considered. For a suggested approximate angular part of the likelihood function the optimal rotation procedure is derived and implemented in Fortran-77.

1 Introduction

The task of kinematic reconstruction of an event is typical of high energy physics. Usually the solution is found by numerical minimization of the log-likelihood function. In this paper we consider one possible parametrization of the event. The main guidelines of this approach were reported at CHEP-97 [1]

For conveniency let us consider the particular process $e^+e^- \to \phi \to \eta\gamma \to \pi^+\pi^-\pi^0\gamma \to \pi^+\pi^-\gamma\gamma\gamma$. The final state contains five particles and is described by 15 parameters: 5 momenta and 10 angles (θ and φ angles in the polar system). But these parameters are not independent. There are 5 constraints: 4 energy-momentum conservation laws and a known value of the π^0 mass. So there are only 10 free parameters. In general, if a detector measures the angles of all 5 particles, one can reconstruct momenta of all particles. Of course the accuracy of angular measuments will determine the accuracy of the reconstructed η-meson mass. If there are some additional measured parameters, then the accuracy of η meson mass will increase and the minimum value of the log-likelihood function can be used for rejection of background events.

The possible set of 10 free parameters is the following. The first parameter is an invariant mass M_η of the η-meson. For an e^+e^- collider with equal energies of the initial electron and positron E_0, the total momentum equals zero and the total energy equals $2E_0$. Using this information one can easily derive the energy of the first photon and momentum of the η-meson. Let η-

meson move along the Z-axis in some system of reference and π^0 move in the XZ-plane in this system. In the system of η-meson we can define two more free parameters: momentum p_0 of π_0 and angle θ_0 between this momentum and the Z-axis. Now in the system of π^0-meson we can define the angle θ_γ between one of the photons and π^0 momentum and axial angle φ_γ, which determines the rotation of $\pi^0 \to \gamma\gamma$ decay plane with respect to the XZ-plane. The corresponding angles θ_+ and φ_+ define the π^+ momentum vector in the center-of-mass system of $\pi^+\pi^-$. At last the rotation angles ψ_1, ψ_2, ψ_3 allow an arbitrary rotation of this 5-particles construction as a whole.

Let us represent the log-likelihood function L as a sum $L = L_E + L_A$, where L_E takes into account the deviations of the particles energies from the measured values (for those particles, whose energies are measured) and L_A takes into account the deviations of particle directions from the measured ones:

$$L_A = \sum_i \frac{\Delta\alpha_i^2}{2\sigma_i^2} \tag{1}$$

Here $\Delta\alpha_i$ is the angle between the direction of the i-th particle momentum in our model and experimentally measured one, σ_i is an estimation of experimental angular accuracy.

Obviously the energies of particles in our model do not depend on the angles ψ_1, ψ_2, ψ_3, hence the energy part of likelihood function L_E does not depend on these angles. There is no profit to minimize separately L_A by numerical methods over $\psi_{1,2,3}$ for every set of 7 parameters M_η, p_0, θ_0, θ_γ, φ_γ, θ_+, φ_+. On opposite there will be great increase of CPU time consumption. In this paper the analytic solution for minimization of the approximate form of L_A over $\psi_{1,2,3}$ is presented.

2 Approximate form of likelihood function

For small deviations $\Delta\alpha_i$ in (1) the likelihood function can be approximated by

$$L_A \approx \sum_i \frac{1-\cos\Delta\alpha_i}{\sigma_i^2} = \sum_i \frac{1-\mathrm{e}_i\cdot\mathrm{n}_i}{\sigma_i^2} = \sum_i \frac{1}{\sigma_i^2} - \sum_i \frac{\mathrm{e}_i\cdot\mathrm{n}_i}{\sigma_i^2} \tag{2}$$

Here vector e_i is the unit vector, defining the measured direction of the i-th particle, vector n_i is the theoretical unit vector for the i-th particle, defined in our model by 10 variable parameters.

The first sum in (2) is constant and can be omitted. The likelihood function L_A reaches its minimum value, when the second sum in (2)

$$L_M = \sum_i \frac{\mathrm{e}_i\cdot\mathrm{n}_i}{\sigma_i^2} \tag{3}$$

reaches its maximum value. Vectors n_i can be presented as unit vectors s_i, depending in our model only on the first 7 parameters, transformed with a rotation matrix T, depending only on the angles $\psi_{1,2,3}$

$$\mathrm{n}_i = \mathrm{T}\cdot\mathrm{s}_i \tag{4}$$

Now we can repeat the task in another way: it is necessary to find the rotation matrix T such that

$$L_M = \sum_i \frac{\mathrm{e}_i^T\mathrm{T}\mathrm{s}_i}{\sigma_i^2} = Tr\left(\mathrm{T}\sum_i \frac{\mathrm{s}_i\mathrm{e}_i^T}{\sigma_i^2}\right) = Tr(\mathrm{T\,V}) \tag{5}$$

reaches its maximum value, where superscript "b^T" means transposed matrix b.

3 Solution for optimal rotation

Let us introduce matrix V

$$\mathrm{V} = \sum_i \frac{\mathrm{s}_i\mathrm{e}_i^T}{\sigma_i^2} \tag{6}$$

and let vectors $\mathrm{v}_1, \mathrm{v}_2, \mathrm{v}_3$ be the columns of the matrix V, vectors $\mathrm{u}_1, \mathrm{u}_2, \mathrm{u}_3$ be the rows of the matrix T. Then the maximized function

$$L_M = \sum_{k=1}^{3} \mathrm{u}_k\cdot\mathrm{v}_k \tag{7}$$

Being the rows of rotation matrix, vectors u_i satisfy the following constraints

$$\mathrm{u}_1^2 = \mathrm{u}_2^2 = 1, \quad \mathrm{u}_1\cdot\mathrm{u}_2 = 0, \quad \mathrm{u}_3 = \mathrm{u}_1\times\mathrm{u}_2 \tag{8}$$

Now the task is transformed to the following: for given vectors $\mathrm{v}_k, k = 1,2,3$ it is necessary to find vectors u_1 and u_2, maximizing the function L_M (7) and satisfying the constraints (8).

Further on we shall use the following variables:

$$V_{ij} = \mathbf{v}_i \cdot \mathbf{v}_j, \quad V_{123} = \mathbf{v}_1 \cdot [\mathbf{v}_2 \times \mathbf{v}_3] \tag{9}$$

The searched rotation matrix is derived by the well-known method of Lagrange uncertain coefficients (for example [2]). Auxilliary transformation of expressions was performed by REDUCE code [3].

There are several classes of solutions depending on the rank of the matrix V.

3.1 Non-zero determinant

For the case of the non-zero determinant $V_{123} \neq 0$ the searched vectors $\mathbf{u}_{1,2}$ can be written as

$$\mathbf{u}_1 = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3, \quad \mathbf{u}_2 = x_7\mathbf{v}_1 + x_8\mathbf{v}_2 + x_9\mathbf{v}_3 \tag{10}$$

The searched coefficients $x_1, x_2, x_3, x_7, x_8, x_9$ were found to be equal to

$$\begin{aligned} x_1 &= \frac{2\lambda_3 V_{123} - V_{22}V_{33} + V_{23}^2}{z_1 V_{123}} \\ x_2 &= \frac{V_{33}V_{12} - V_{13}V_{23} - \lambda_1 V_{123}}{z_1 V_{123}} \\ x_3 &= \frac{V_{123}\cdot(2V_{13}\lambda_3 - V_{23}\lambda_1) + (V_{22}V_{13} - V_{12}V_{23})(z_1 - V_{33})}{z_1 V_{123}\cdot(z_1 - V_{33})} \\ x_7 &= \frac{-\lambda_1 V_{123} + V_{12}V_{33} - V_{13}V_{23}}{z_1 V_{123}} \\ x_8 &= \frac{2\lambda_2 V_{123} - V_{11}V_{33} + V_{13}^2}{z_1 V_{123}} \\ x_9 &= \frac{V_{123}\cdot(2V_{23}\lambda_2 - V_{13}\lambda_1) + (V_{23}V_{11} - V_{12}V_{13})(z_1 - V_{33})}{z_1 V_{123}\cdot(z_1 - V_{33})} \end{aligned} \tag{11}$$

where

$$\begin{aligned} \lambda_1 &= \frac{2\lambda_2[V_{12}\cdot(z_1 - V_{33}) + V_{13}V_{23}]}{V_{13}^2 + (V_{11} - z_1)(z_1 - V_{33})}, \quad \lambda_2 = \frac{\lambda_3[(V_{11} - z_1)(z_1 - V_{33}) + V_{13}^2]}{V_{23}^2 + (V_{22} - z_1)(z_1 - V_{33})} \\ \lambda_3 &= \frac{V_{123}\cdot\left[z_1^2 - (V_{22} + V_{33})z_1 + V_{22}V_{33} - V_{23}^2\right]}{z_1^3 + b\cdot z_1/2 + 2\left(V_{11}V_{22}V_{33} - V_{11}V_{23}^2 - V_{22}V_{13}^2 - V_{33}V_{12}^2\right) + 4V_{12}V_{13}V_{23}} \end{aligned} \tag{12}$$

and z_1 is a root of the following equation

$$z_1^4 + bz_1^2 + cz_1 + d = 0 \tag{13}$$

where

$$\begin{aligned} b =& \; 2\left(V_{12}^2 + V_{13}^2 + V_{23}^2 - V_{11}V_{22} - V_{11}V_{33} - V_{22}V_{33}\right) \\ c =& \; 8V_{123}^2 \\ d =& \; \left(V_{22}^2 + V_{33}^2 + 2V_{23}^2\right)V_{11}^2 - 2\left(V_{11} + V_{22} + V_{33}\right)V_{123}^2 \\ & -2\left(V_{22}V_{12}^2 + V_{33}V_{13}^2 + 2V_{12}V_{13}V_{23}\right)V_{11} \\ & +\left(V_{33}^2 + 2V_{13}^2\right)V_{22}^2 - 2\left(V_{33}V_{23} + 2V_{12}V_{13}\right)V_{22}V_{23} + 2\left(V_{13}^2 + V_{23}^2\right)V_{12}^2 \\ & +2V_{33}^2V_{12}^2 - 4V_{33}V_{12}V_{13}V_{23} + V_{12}^4 + V_{13}^4 + V_{23}^4 + 2V_{13}^2V_{23}^2 \end{aligned}$$

Usually there are four real roots of this equation. For all real values of z_1 the value of the goal function L_M is evaluated, and that value of z_1 is chosen, which provides the maximum value of L_M.

3.2 Zero determinant: matrix V of rank 2

There can be a case when $V_{123} \approx 0$, but there are at least two acolinear vectors $\mathbf{v}_k$. For conveniency let us consider the case when these are the vectors $\mathbf{v}_1$ and $\mathbf{v}_2$.

Let us introduce auxilliary variables

$$\begin{aligned} D &= V_{11}V_{22} - V_{12}^2 > 0, \\ r_1 &= fracV_{22}V_{13} - V_{12}V_{23}D, \\ r_2 &= \frac{V_{11}V_{23} - V_{12}V_{13}}{D} \end{aligned} \tag{14}$$

Now the solution is represented as

$$\begin{aligned} \mathbf{u}_1 &= x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_4[\mathbf{v}_1 \times \mathbf{v}_2], \\ \mathbf{u}_2 &= x_7\mathbf{v}_1 + x_8\mathbf{v}_2 + x_{10}[\mathbf{v}_1 \times \mathbf{v}_2], \\ \mathbf{v}_3 &= r_1\mathbf{v}_1 + r_2\mathbf{v}_2 \end{aligned} \tag{15}$$

where

$$
\begin{aligned}
x_1 &= \frac{(V_{22}+2V_{12}\lambda_3)r_1r_2+r_1^2V_{12}+2\left(r_2^2V_{22}-z_1\right)\lambda_3}{(4\lambda_2\lambda_3-1)\lambda_1 z_1},\\
x_2 &= -\frac{(V_{12}+2V_{11}\lambda_3)r_1r_2+r_1^2V_{11}+2r_2^2V_{12}\lambda_3-z_1}{(4\lambda_2\lambda_3-1)\lambda_1 z_1},\\
x_7 &= -\frac{(V_{12}+2V_{22}\lambda_2)r_1r_2+r_2^2V_{22}+2r_1^2V_{12}\lambda_2-z_1}{(4\lambda_2\lambda_3-1)\lambda_1 z_1},\\
x_8 &= \frac{(V_{11}+2V_{12}\lambda_2)r_1r_2+r_2^2V_{12}+2\left(r_1^2V_{11}-z_1\right)\lambda_2}{(4\lambda_2\lambda_3-1)\lambda_1 z_1},\\
x_4 &= \frac{r_1}{z_1},\\
x_{10} &= \frac{r_2}{z_1},\\
z_1 &= \pm\sqrt{D\cdot(1+r_1^2+r_2^2)},\\
\lambda_2 &= \frac{(V_{11}z_1-D)\left(1+r_1^2\right)+2r_1r_2V_{12}z_1+r_2^2(V_{22}z_1-2D)}{2(r_1r_2D+V_{12}z_1)},\\
\lambda_1 &= \pm\sqrt{\frac{z_1-V_{33}}{1-4\lambda_2\lambda_3}},\\
\lambda_3 &= \frac{(V_{22}z_1-D)\left(1+r_2^2\right)+2r_1r_2V_{12}z_1+r_1^2(V_{11}z_1-2D)}{2(r_1r_2D+V_{12}z_1)}
\end{aligned}
\qquad (16)
$$

3.3 Zero determinant: matrix V of rank 1

In this case all vectors v_k are proportional to one vector p

$$\mathrm{v}_1 = a\mathrm{p},\quad \mathrm{v}_2 = b\mathrm{p},\quad \mathrm{v}_3 = c\mathrm{p},\quad \mathrm{p}^2 = 1 \qquad (17)$$

Maximum value of L_M is

$$f_M = \max L_M = \max Tr\,(\mathrm{TV}) = \sqrt{a^2+b^2+c^2} \qquad (18)$$

The rotation matrix T is not unique in this case. One of the possible solutions is

$$\mathrm{T} = \mathrm{T_2T_1},\quad \text{where Tbnldmath } \mathrm{T}_2 = \begin{pmatrix} 0 & \frac{\sqrt{b^2+c^2}}{f_M} & \frac{a}{f_M} \\ -\frac{c}{\sqrt{b^2+c^2}} & -\frac{ab}{\sqrt{b^2+c^2}f_M} & \frac{b}{f_M} \\ \frac{b}{\sqrt{b^2+c^2}} & -\frac{ac}{\sqrt{b^2+c^2}f_M} & \frac{c}{f_M} \end{pmatrix} \qquad (19)$$

and the third line of matrix T_1 is vector p. The first two lines of T_1 are any two unit vectors, which form with p a right-hand coordinate system.

4 Description of input and output arguments

Following the described algorithm there was written the subroutine BUROTAT in Fortran-77 with typical running time about 1.1 ms per call on VAX-3600 computer.

Subroutine BUROTAT is called with the following arguments:

```
call BUROTAT(N,E,V,SIG,TT,FL)
```

where

```
integer *4 N        ! Number of particles
real *8 E(3,N)      ! Unit vectors of "experimental particles"
real *8 V(3,N)      ! Unit vectors of "theoretical particles"
real *8 SIG(N)      ! Experimental angular accuracy (radian)
real *8 TT(3,3)     ! Searched rotation matrix
real *8 FL          ! Searched minimum of
                    ! angular likelihood function L_A
```

Input: N, E, V, SIG.
Output: TT, FL.

5 Conclusion

One of the possible procedures for the kinematic reconstruction of multi-particle events is considered. For special parametrization of an event and suggested approximate angular part of the likelihood function the optimal rotation procedure is derived and implemented in Fortran-77.

This procedure allows to decrease the number of free optimized parameters by three, and hence decrease the CPU consumption by more than two times for the sample process $e^+e^- \to \eta\gamma \to \pi^+\pi^-\gamma\gamma\gamma$ (decrease from 10 to 7 optimized parameters). Even more important there is an increase of a detection efficiency for this process by several percents, that characterises the simplification of the likelihood function profile and hence the smaller probability of finding the false minimum (minimization was performed by MINUIT code).

References

[1] *A.D.Bukin.* On the Kinematic Reconstruction of Multiparticle Events. Reported at the "Computing in High Energy Physics" Conference, April 7–11, 1977, Berlin

[2] *G.A.Korn and T.M.Korn.* Mathematical handbook for scientists and engineers. McGraw-Hill, 1968.

[3] *Anthony C. Hearn.* REDUCE user's manual. Rand, 1987.

Addendum. Output of the test program

For the purpose of control of fatal corruption of subroutine BUROTAT text, a simple test program TESBURO has been written. The construction of four particles with some directions of movement e_k is transformed by rotation transformation in order to obtain the "theoretical" directions s_k. Then both sets of unit vectors are processed by BUROTAT subroutine to find the inverse transformation and minimum value of angular part of the likelihood function L_A. The minimum value L_A in this case must be equal to zero. The following is the output listing of this test program:

```
                Test of BUROTAT subroutine

   4 initial unit "experimental" vectors:
E1=    0.42426    0.56569    0.70711
E2=    0.09759    0.19518    0.97590
E3=   -0.53452    0.80178    0.26726
E4=   -0.09950   -0.99504    0.00000

       were rotated by orthogonal transformation T3:

          0.36851    0.92077    0.12799
T3=      -0.92473    0.34898    0.15195
          0.09525   -0.17435    0.98007

       and so the "theoretical" vectors obtained:
V1=    0.76771   -0.08747    0.63480
V2=    0.34058    0.12616    0.93171
V3=    0.57548    0.81470    0.07123
V4=   -0.95287   -0.25523    0.16401

       Angular errors were put equal to  0.100  0.050  0.080  0.150

       BUROTAT subroutine has searched inverse transformation TT:

          0.36851   -0.92473    0.09525
TT=       0.92077    0.34898   -0.17435
          0.12799    0.15195    0.98007

       and calculated Likelihood function FL=  0.000000
       (which should be equal to zero here)

       Product T3*TT must be equal to unit matrix:
```

```
         1.000000  0.000000  0.000000
T3*TT=   0.000000  1.000000  0.000000
         0.000000  0.000000  1.000000
```

Besides this simple check a more sophisticated test with 10^4 Monte Carlo events was carried out. Test program TESBURO1 had the following algorithm of generating each event:

1. Five unit vectors u_i are generated, isotropic and independant from each other. These vectors are treated as theoretical vectors.

2. Random rotation matrix R is generated.

3. Five "experimental" vectors e_i are obtained: $e_i = R \cdot u_i$.

4. These sets of "theoretical" and "experimental" vectors are offered to subroutine BUROTAT to find the rotation matrix T and minimum value of likelihood function. Angular accuracy is assigned to the first two particles equal to 2° and to the last three particles equal to 5° (of course, recalculated to radians).

5. For this case we know the exact result: likelihood function should be equal to zero and rotation matrix T should be equal to the matrix R. Moreover we can calculate ourselves the likelihood function and compare it with BUROTAT value.

6. If loop on the events is not over, then proceed with the point 1.

Over all statistics the maximum deviation of the returned from BUROTAT rotation matrix from the known original rotation matrix was found to be $3.4 \cdot 10^{-9}$ (all calculations were made with double precision). Maximum value of likelihood function is equal to $2 \cdot 10^{-5}$ and matches with maximum deviation of likelihood function, calculated via rotation matrix T, from that returned from BUROTAT subroutine. Obviously there is a great loss of accuracy because of many arithmetic operations, but it seems suitable for most realistic cases.

Randomness of the rotation matrix R is assured by the distributions of its diagonal elements (Fig. 1), which are the cosines of rotation angles of the three coordinate vectors. Now let us "spoil" the directions of the experimental vectors in accordance with angular accuracy, assigned to each particle. Then the likelihood function must not be equal to zero. The distribution of

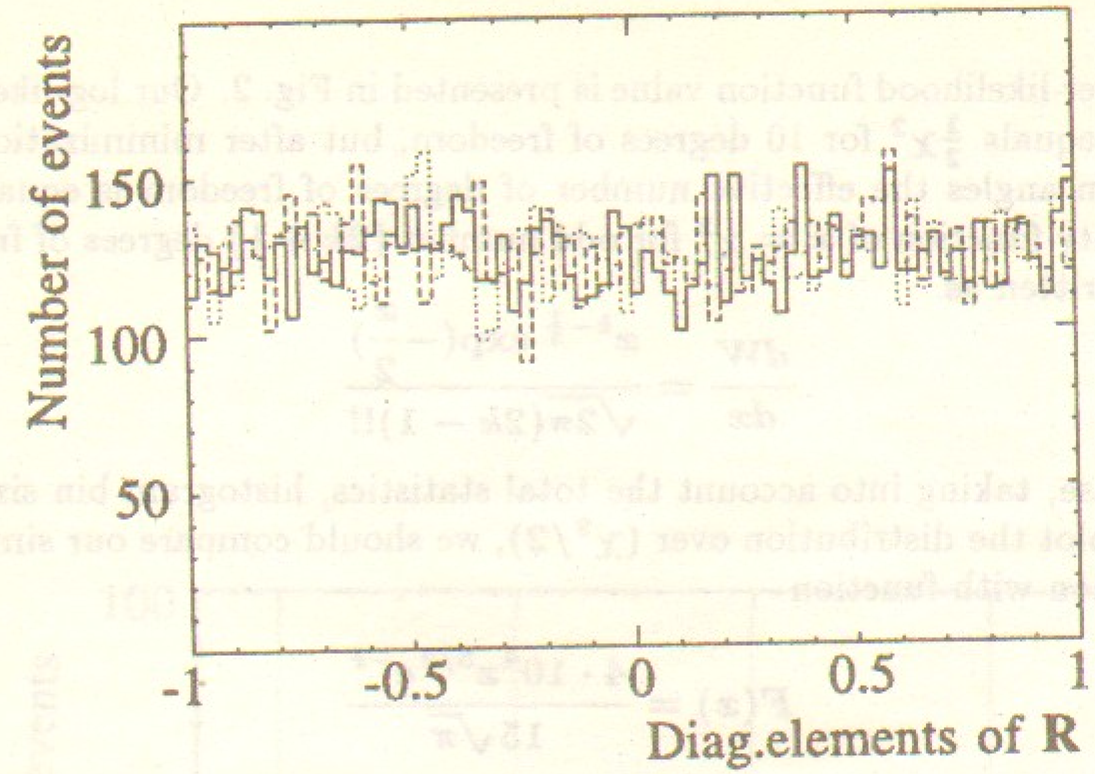

Figure 1: Distributions of the events over the diagonal elements of rotation matrix R.

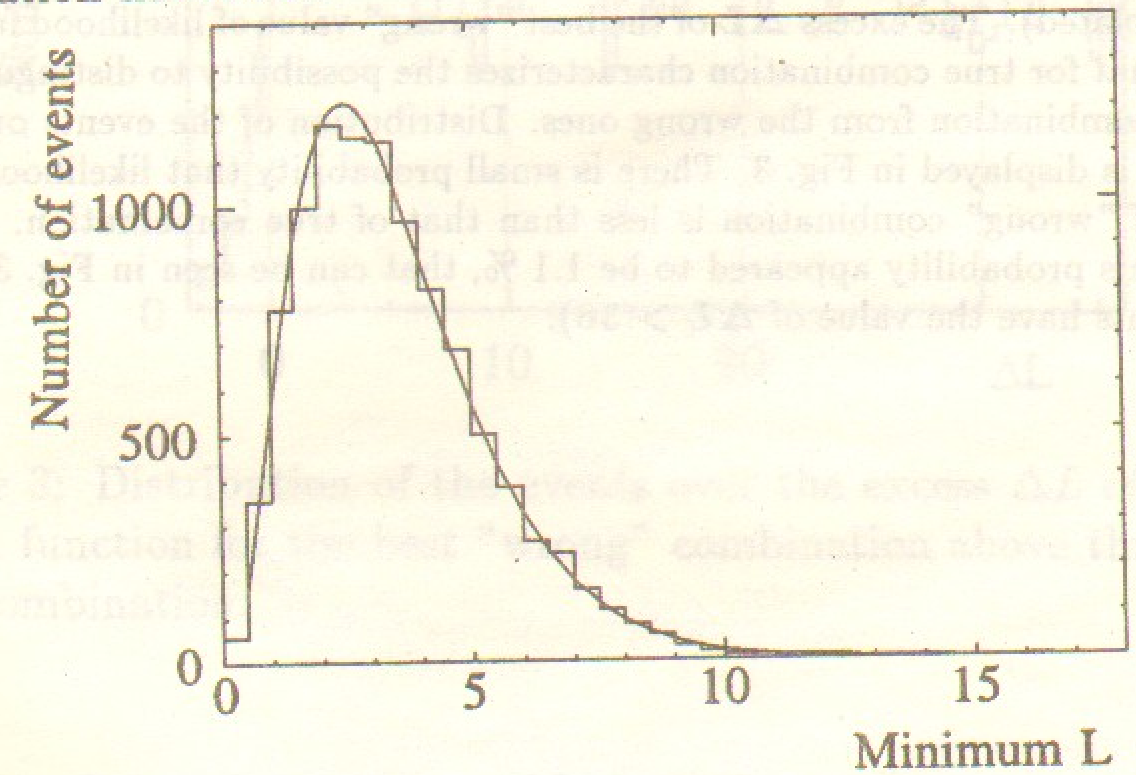

Figure 2: Distribution of the events over the minimum value of likelihood function. Smooth curve shows the distribution function over $(\chi^2/2)$ for 7 degrees of freedom.

events over likelihood function value is presented in Fig. 2. Our log-likelihood function equals $\frac{1}{2}\chi^2$ for 10 degrees of freedom, but after minimization over 3 rotation angles the effective number of degrees of freedom is equal to 7. Probability function of $x = \chi^2$ for odd number $(2k+1)$ degrees of freedom can be written as

$$\frac{dW}{dx} = \frac{x^{k-\frac{1}{2}} \exp(-\frac{x}{2})}{\sqrt{2\pi}(2k-1)!!} \qquad (20)$$

In our case, taking into account the total statistics, histogram bin size, and that we plot the distribution over $(\chi^2/2)$, we should compare our simulated distribution with function

$$F(x) = \frac{4 \cdot 10^4 x^{5/2} e^{-x}}{15\sqrt{\pi}} \qquad (21)$$

(smooth curve in Fig. 2).

For every event we can calculate the value of likelihood function for all "wrong" combinations (correspondence of theoretical and experimental particles violated). The excess ΔL of the best "wrong" value of likelihood function over that for true combination characterizes the possibility to distinguish the right combination from the wrong ones. Distribution of the events over this excess is displayed in Fig. 3. There is small probability that likelihood function of "wrong" combination is less than that of true combination. In our case this probability appeared to be 1.1 %, that can be seen in Fig. 3 (62 % of events have the value of $\Delta L > 36$).

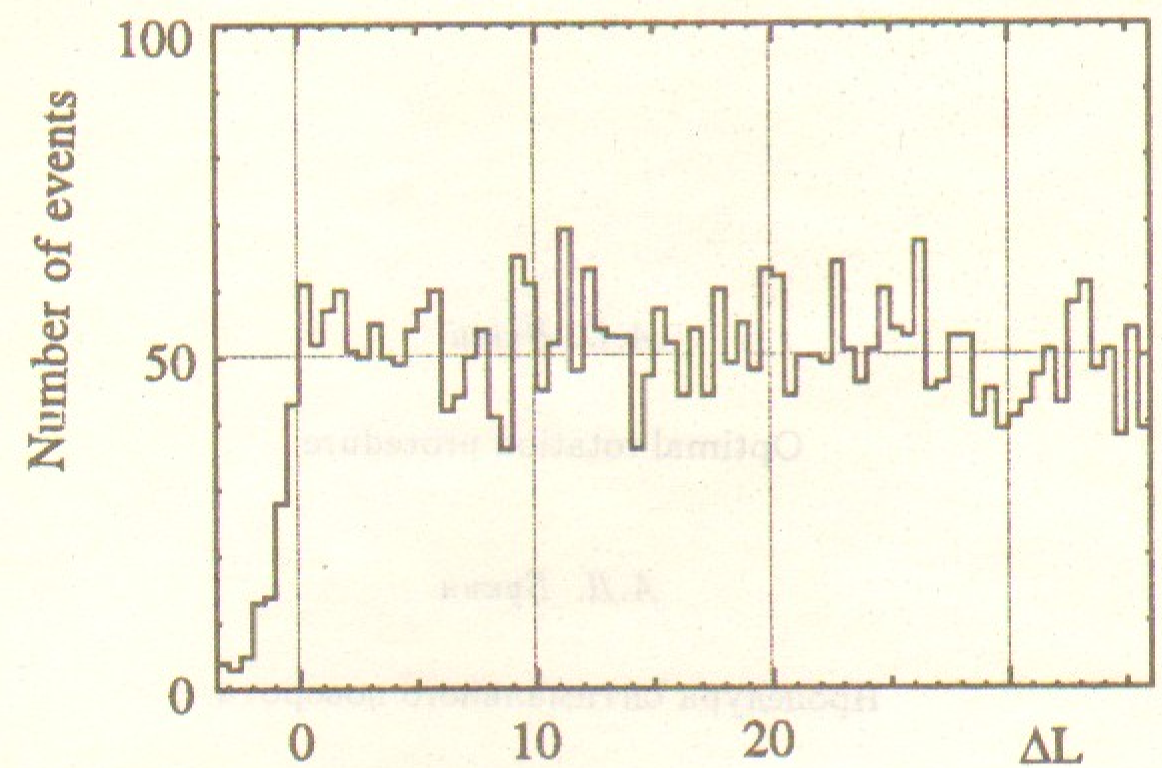

Figure 3: Distribution of the events over the excess ΔL of likelihood function for the best "wrong" combination above that for true combination.

A.D. Bukin

Optimal rotation procedure

А.Д. Букин

Процедура оптимального поворота

Budker INP 97-50

Ответственный за выпуск А.М. Кудрявцев
Работа поступила 29.01.1997 г.

Сдано в набор 26.05.1997 г.
Подписано в печать 2.06.1997 г.
Формат бумаги 60×90 1/16 Объем 0.9 печ.л., 0.8 уч.-изд.л.
Тираж 150 экз. Бесплатно. Заказ N 50.

Обработано на IBM PC и отпечатано на
ротапринте ГНЦ РФ "ИЯФ им. Г.И. Будкера СО РАН",
Новосибирск, 630090, пр. академика Лаврентьева, 11.