

УДК 519.6:539.1.07

A STUDY OF MILLEPEDE ALIGNMENT ALGORITHM ON MONTE CARLO MODEL OF HERA-B OUTER TRACKER

*I. Belotelov*¹, *A. Lanyov*², *G. Ososkov*³

Joint Institute of Nuclear Research, Dubna

In this paper a method of internal alignment of HERA-B OTR PC chambers is discussed. The method is based on simultaneous fit of the track and alignment parameters using Millepede matrix reduction and singular value decomposition. Software which implements this idea has been developed, the method has been studied on Monte Carlo models with different levels of simulation. A method generalization for the case of track nonlinear model has been proposed.

В данной работе обсуждается метод внутреннего геометрического выравнивания внешней трекинговой системы эксперимента HERA-B. Метод основан на одновременном определении параметров треков и геометрических поправок с использованием техники редукции матрицы нормальных уравнений Миллепед и сингулярного разложения. Разработано программное обеспечение, воплощающее эти идеи, метод протестирован на моделях Монте-Карло различного уровня сложности и реалистичности. Предложено обобщение метода на случай нелинейной модели трека.

INTRODUCTION

HERA-B is the fixed target experiment at the 920 GeV proton storage ring of HERA at DESY [1]. The tracking system of HERA-B consists of a vertex detector system (VDS) and a main tracker system. The latter is separated into an inner tracker (ITR) close to the proton beam pipe and outer tracker (OTR) farther out. The HERA-B Outer Tracker uses honeycomb drift chambers [2]. Pattern recognition Chambers (PC) of OTR are made of 998 separate sensitive planes. These planes are grouped in 24 stereolayers, the layers are grouped in four superlayers. Some layers are rotated by ± 80 mrad. The superlayers are divided into two parts: $(+X)$ and $(-X)$ halves wrapping proton beam pipe. Since OTR PC geometry is rather complicated and track multiplicity is high, in order to provide accurate results all the internal parameter correlations should be taken into account, external degrees of freedom should be known and properly fixed. Several techniques have been used to fulfill these requirements.

¹E-mail: Ivan.Belotelov@jinr.ru

²E-mail: lanyov@cv.jinr.ru

³E-mail: Gennadi.Ososkov@jinr.ru

1. DATA-DRIVEN ALIGNMENT

1.1. Internal Alignment as Functional Minimization Problem. Any tracking system can be considered as an assembly of many separate modules, the positions of which can be displaced during assembling and the detector maintenance despite of any preventive means. Therefore, an alignment procedure is needed to detect possible distortions. We follow here a concept of a data-driven alignment, i.e., the use of data sets of real measurements produced by the tracking system. Misalignment of detector modules can be found by analyzing residuals between the measured values and fitted track coordinate. These residuals are functionally dependent on the two types of parameters: track model parameters and alignment parameters. Therefore, the alignment problem can be formulated as a mathematical problem of minimizing a functional summing the squares of all residuals by both types of parameters:

$$\chi^2 = \sum_{\text{events}} \sum_{\text{tracks}} \sum_{\text{hits}} \Delta_i^2 / \sigma_i^2, \quad (1)$$

where Δ_i is hit residual — the difference between the measured and fitted track position:

$$\Delta_i = x_{\text{fit}} - x_{\text{meas}} \quad (2)$$

and σ_i is accuracy of hit coordinate measurement. Fitted track position x_{fit} is defined by so-called *track model*:

$$x_{\text{fit}} = \sum_{i=1}^n a_i \cdot d_i + \sum_{j=1}^{\nu} \alpha_j \cdot \delta_j. \quad (3)$$

The simplest track model in misaligned detector is linear and includes the estimated track and alignment parameters ($\{a_i\}$ and $\{\alpha_i\}$) and their derivatives ($\{d_i\}$ and $\{\delta_i\}$). Solution (the set of parameters values, minimizing the functional) is obtained from a set of linear equations (so-called normal equations of least squares):

$$Ca = b, \quad (4)$$

where C is matrix of derivatives with dimension equal to the number of estimated parameters. The solution can be found as

$$a = C^{-1}b. \quad (5)$$

Here some difficulties appear: extreme parameter multiplicity and unconstrained external degrees of freedom. The number of tracks necessary to use to get reasonable alignment precision is in the order of 10^4 – 10^6 , and the number of alignment parameters for modern detectors is in the order of 10^3 . Therefore, the matrix of normal equation is not invertible using the common techniques. Internal alignment, based on residual functional minimization, is blind to the effects of geometrical transformations which project tracks into themselves. This leads to singular normal equation matrix. Singular matrix does not have an inverse.

1.2. Minimization of Alignment Functional. In order to eliminate the above-mentioned difficulties, one could either simplify the initial task or explore its special properties and apply special numerical algorithms.

Iterating Biased Fit. To deal with extremal parameter multiplicity, some alignment procedures attempt to find the alignment parameters by performing a series of track fits which initially ignore the misalignments [3,4]. From the distribution of residuals the alignment parameters are then deduced. But the fitted track parameters are a function of misalignment, and dividing initial functional Eq. (1) in sets of separate track parameters fits and alignment parameters fit ignores all correlations between the track and alignment parameters existing in the initial functional. This approach under some circumstances allows one to reduce the remaining residuals, but it does not allow one to control external degrees of freedom without applying any special requirements.

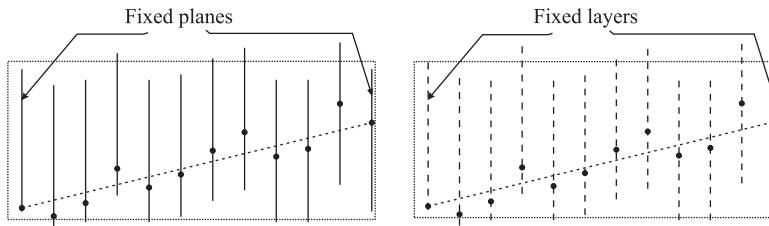


Fig. 1. Fixing selected planes/layers of planes

For simple telescopic detector geometries we can choose two fixed planes (or four planes for three-dimensional tracks) and align other detector elements in respect to these selected planes. This is achieved by using tracks passing these two fixed planes and also all other planes being aligned. For large complex setups detector layer consists of several planes; therefore, all the planes in the layer have to be fixed. But the planes in a layer themselves are unaligned in respect to each other.

Using External Tracks. Another way to reduce the dimension of the problem is using tracks extrapolated from other detectors if it is possible. In this case misalignment in the examined part of the tracking system does not affect the track parameters anymore. It allows one to align the given subdetector in respect to the other parts of the system. But because of long extrapolation length, reference detector misalignment, multiple scattering, etc., precision of the extrapolated track parameters is much worse than that for the tracks reconstructed in the given subdetector. This method provides the ability to avoid displacement of the whole detector with high precision, but accuracy for determining individual modules' position is less than desirable.

Exploring Special Matrix Structure. In contrast to these two methods, an approach suggested by V. Blobel [5] makes one fit to all track and alignment parameters simultaneously using one huge error matrix which precisely describes the relationship between track (*local*) parameters α_i and alignment (*global*) parameters a_i . This method gets around the problem's dimension using matrix inversion by partitioning technique. It splits the large inversion into set of smaller inversions. The next section gives a detailed description of this idea.

1.3. Idea of Millepede Alignment Algorithm. Matrix of normal equation C is very sparse and has regular block structure, with many vanishing submatrices. It has three types of contributions from every particular measurement. The first part is a contribution of a symmetric matrix C_i of dimension n (number of global parameters):

$$C_{ij} = \sum_k w_k d_{ik} d_{jk}. \quad (6)$$

The second part is symmetric matrix Γ_i defined as

$$\Gamma_{ij} = \sum_k w_k \delta_{ik} \delta_{jk}, \quad (7)$$

which gives the contribution to the big matrix on the diagonal and depends only on the i th partial track. The third contribution is a rectangular matrix G_i which contains the correlation between the global and local parameters:

$$G_{ij} = \sum_k w_k d_{ik} \delta_{jk}. \quad (8)$$

There are two contributions to the vector of normal equation, b_i for the global and β_j for the local parameters:

$$\beta_j = \sum_k w_k x_k \delta_{jk}, \quad b_j = \sum_k w_k x_k d_{jk}. \quad (9)$$

In terms of these submatrices the whole matrix of normal equation looks like

$$\begin{pmatrix} \sum C_l & \dots & G_l & \dots \\ \vdots & \ddots & 0 & 0 \\ G_l^T & 0 & \Gamma_l & 0 \\ \vdots & 0 & 0 & \ddots \end{pmatrix} \cdot \begin{pmatrix} a \\ \vdots \\ \alpha_l \\ \vdots \end{pmatrix} = \begin{pmatrix} \sum b_l \\ \vdots \\ \beta_l \\ \vdots \end{pmatrix}. \quad (10)$$

One can solve the local equation $\Gamma\alpha = \beta$ for each track $\alpha = \Gamma^{-1}\beta$. Using matrix inversion by partitioning, the huge matrix could be reduced to n normal equations for global parameters

$$C'a = b' \quad (11)$$

with modified matrices C' and b' ,

$$C' = \sum_i C_i - \sum_i G_i \Gamma_i^{-1} G_i^T, \quad b' = \sum_i b_i - \sum_i G_i (\Gamma_i^{-1} \beta_i^T). \quad (12)$$

All information about detector geometry, tracks, all parameters correlation which were encoded in the initial huge matrix C now is kept in reduced C' . The direct solution $a = C'^{-1}b'$ represents solution vector a with covariance matrix C'^{-1} .

The program Millepede written in Fortran by V. Blobel implements this technique in code.

1.4. External Degrees of Freedom. For tracks parameterized by a straight line, from the point of view of χ^2 value the track model is invariant to any linear coordinate transformation. In three-dimensional space these transformations can be described by 12 parameters:

- axis scalings F_x, F_y, F_z ;
- translations D_x, D_y, D_z ;
- rotations R_x, R_y, R_z ;
- shearings S_{xz}, S_{yz}, S_{xy} .

For detectors used in modern high energy physics not all of these degrees of freedom are valid. Let us assume the simplest detector having some sensitive planes with strips or wires along Y axis, which measure only one x coordinate transverse to Z axis. Let us assume the

detector planes to be rigid bodies, where the strip or wire pitch within the modules cannot be changed, thus scalings in x and y are forbidden. The remaining external degrees of freedom depend on the track model. Let us analyze some of them which we will use in further study.

First, let us suppose that only possible misalignment is parallel to the measured coordinate x . For such a detector and track model an overall detector shift D_x and shearing S_{xz} are not fixed.

The other interesting case is the track model including shifts in X and Z and rotations around Z . This is a more or less realistic scenario for our detectors, since such a detector is not sensitive to shifts in Y and the effect of rotations around X and Y is much smaller than that around Z axis. For such a track model we should fix five external degrees of freedom: overall detector shifts D_x and D_z , shearing S_{xz} , scaling in Z and overall rotation around Z .

If we are measuring both x and y coordinates (i.e., our planes are rotated relative to each other), then we also should fix overall shift D_y and shearing S_{yz} .

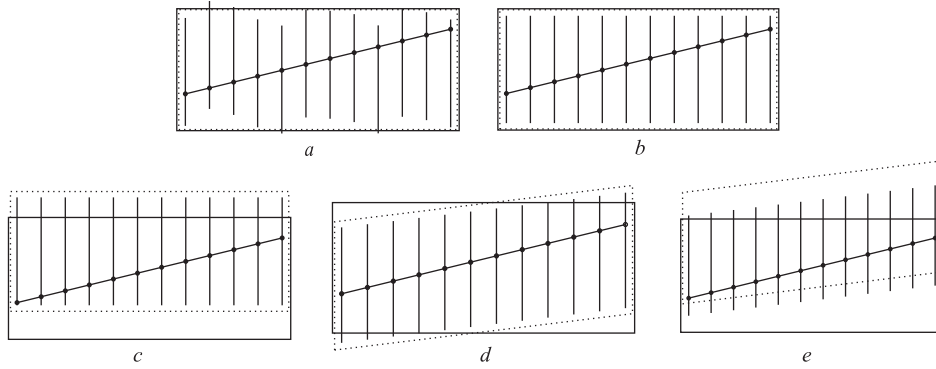


Fig. 2. Examples of misaligned geometry

Plots in Fig. 2 illustrate the meaning of these unconstrained degrees of freedom. Figure 2, *a* shows initial state of misaligned detector; Fig. 2, *b* is one of the possible solutions of internal alignment. But the situation represented in Figs. 2, *c–e* also corresponds to the minimum of alignment functional. We have an infinite number of possible solutions and we need to be able to choose the desired one. It could be done in several ways [6].

Let us assume that we know the position of some modules in our detector, e.g., from survey measurements. Then we prefer to fix these modules and keep them at the same known position.

In case the number of these modules is exactly equal to the number of unconstrained degrees of freedom, in order to constraint our problem we need to add such a term to our initial residual functional (i corresponds to the number of fixed modules):

$$a_i = 0. \quad (13)$$

In Blobel's notation it corresponds to keeping measurements in fixed planes during local fits, but excluding corresponding alignment parameters a_i from global fit. It allows us to choose exactly one solution, which puts our functional to its minimum and globally aligns the detector in respect to those fixed planes.

If the number of planes with more or less known positions is greater than the number of external degrees of freedom, simple fixing of all these planes produces overconstraining. It is better to allow these planes with parameters a_i some movements inside some ranges Δ_i , minimizing this functional:

$$\chi_{\text{extra}}^2 = \sum_i a_i^2 / \Delta_i^2. \quad (14)$$

In matrix form it corresponds to adding the term $1/\sigma_i^2$ to the diagonal elements of the matrix of normal equation, corresponding to the i th plane.

Sometimes it is not possible to determine the position of any particular plane in global frame, but it is known that average overall misalignment is equal to zero. Usually it is done by a global alignment procedure. And the only question is how to keep the system globally fixed during the detector internal alignment.

We can achieve it by applying some limitation on all fitted alignment parameters:

$$\sum_i w_i \cdot a_i = 0. \quad (15)$$

Choosing proper weights, we could fix overall shifts, shearings, rotations and axis rescalings. Taking these extra conditions into account can be done using Lagrange multipliers; in matrix form it corresponds to adding extra lines (rows and columns) holding weights w_i to matrix of normal equation C or to reduced matrix C' .

1.5. Singular Value Decomposition. Matrix C of normal equation for internal alignment is singular. It has *rank defect*. There exists a very powerful set of techniques, known as *singular value decomposition* for dealing with sets of equations or matrices that are either singular or else numerically very close to singular [7, 8]. In our case standard technique, which is Gaussian elimination, fails, but SVD could diagnose precisely what the problem is. It also could be used not only to diagnose the problem, but also to solve it. SVD is defined by the following theorem of linear algebra:

If A is real m -by- n matrix, then there exist orthogonal matrices

$$U = [u_1, \dots, u_m] \in R^{m \times m} \quad \text{and} \quad V = [v_1, \dots, v_n] \in R^{n \times n} \quad (16)$$

such that

$$U^T A V = \text{diag}(\sigma_1, \dots, \sigma_p) \in R^{m \times n}, \quad p = \min\{m, n\}, \quad (17)$$

where

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_p \geq 0, \quad (18)$$

σ_i are the *singular values*, columns of U and rows of V are the *singular vectors*. If r is defined by

$$\sigma_1 \geq \dots \geq \sigma_r \geq \sigma_{r+1} = \dots = \sigma_p = 0, \quad (19)$$

then $\text{rank}(A) = r$. Now we have SVD *expansion* for matrix A :

$$A = \sum_{i=1}^r \sigma_i u_i v_i^T. \quad (20)$$

Condition number κ is defined by

$$\kappa(A) = \frac{\sigma_{\max}(A)}{\sigma_{\min}(A)}. \quad (21)$$

Lower κ here means better conditioned problem. For *singular* matrix the condition number κ is infinite and for *ill-conditioned* matrices it is too large (approaches the machine's floating-point precision).

From (17) and (20) it follows immediately that the inverse of matrix A is

$$A^{-1} = V \left[\text{diag} \left(\frac{1}{\sigma_i} \right) \right] U^T. \quad (22)$$

Let us consider the set of simultaneous equations

$$A \cdot x = b. \quad (23)$$

Here A works as a linear mapping from vector space x to vector space b . If A is singular, then there is some subspace of x , called the *nullspace*, that is mapped to zero: $A \cdot x = 0$. There is also some subspace of x that can be mapped to b . This subspace of b is *range* of A . The dimension of the range is *rank* of A .

SVD explicitly constructs orthonormal bases for the nullspace and range of a matrix. The columns of U whose same-numbered elements σ_j are nonzero are an orthonormal set of basis vectors that span the range; the columns of V whose same-numbered elements σ_j are zero are an orthonormal basis for the nullspace. If vector b lies in the range of A , singular set of equations has more than one solution, since any vector in the nullspace can be added to x in any linear combination. If we want to get one particular solution, we can replace $1/\sigma_i = 0$ when $\sigma_i = 0$, then compute:

$$x = V \left[\text{diag} \left(\frac{1}{\sigma_i} \right) \right] (U^T \cdot b). \quad (24)$$

This will be the solution vector with the smallest length $|x|^2$.

2. EXPLORING EXTERNAL DEGREES OF FREEDOM FOR SIMPLE MONTE CARLO DETECTOR MODEL USING SVD

A simple detector model with 12 planes simulated by Monte Carlo methods was used. Straight line tracks pass through all the planes. Each plane measures one coordinate x , transverse to the axis of the detector. Each plane may be shifted in the direction of measured coordinate.

A Millepede-like code rewritten in C++ was used to solve the alignment problem. Matrix C was extracted and passed to singular value decomposition routine, provided by LAPACK library [9].

The shape of singular value spectrum shown in Fig. 3 has two zero singular values. One can deduce that for this problem two degrees of freedom are not constrained. Two singular vectors corresponding to these singular values construct the basis for all transformations of

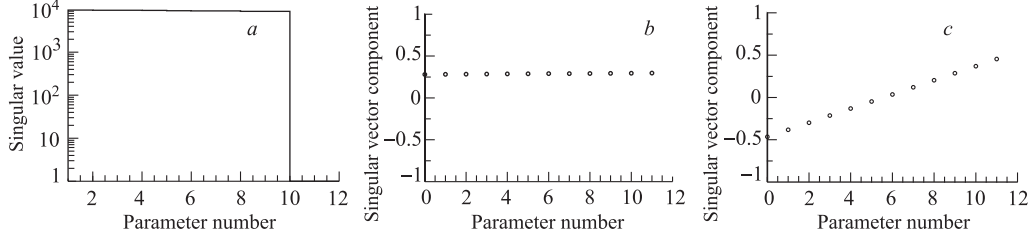


Fig. 3. Shape of singular value spectrum (a) and two singular vectors, corresponding to zero singular values (b, c) for unconstrained alignment problem

our alignment parameters which keep hit residuals the same. The shape of these vectors is defined by unconstrained degrees of freedom.

In this case one can show that these transformations are overall shift and shearing. To fix the overall shift, one can apply the following requirements on estimated alignment corrections a_i :

$$\sum a_i = 0. \tag{25}$$

Now the «center of gravity» of our model detector is fixed, but still some external degrees of freedom are not yet under control, namely each plane can be shifted by a value proportional to its distance from the center of gravity along Z axis, and these movements are not forbidden neither by initial residual functional Eq. (1) nor by Eq. (25). To fix these transformations, one can apply constraint with weights, also proportional to Z position of the corresponding plane:

$$\sum \Delta z_i a_i = 0. \tag{26}$$

Applying linear constraints in form of Lagrange multipliers Eq. (15), these external parameters can be fixed.

The situation after constraining is shown in Fig. 4: all the singular values are nonzero — the spectrum is flat.

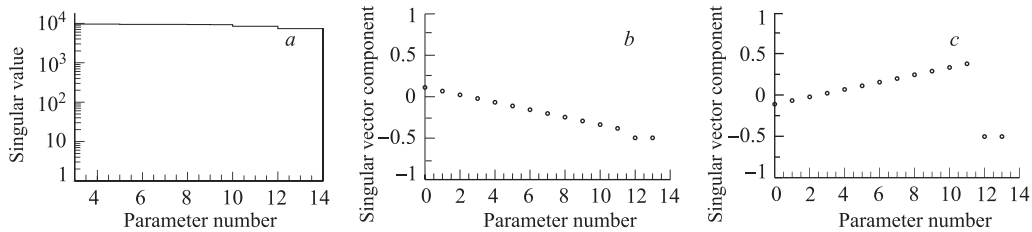


Fig. 4. Shape of singular value spectrum (a) and two singular vectors, corresponding to zero singular values (b, c) after linear constraints are applied

The two singular vectors shown in Fig.4 still construct a basis for overall shift and shearing, but these transformations are not unconstrained — the corresponding singular values are nonzero. If external degrees of freedom are not properly fixed, the detector could get significant overall displacement.

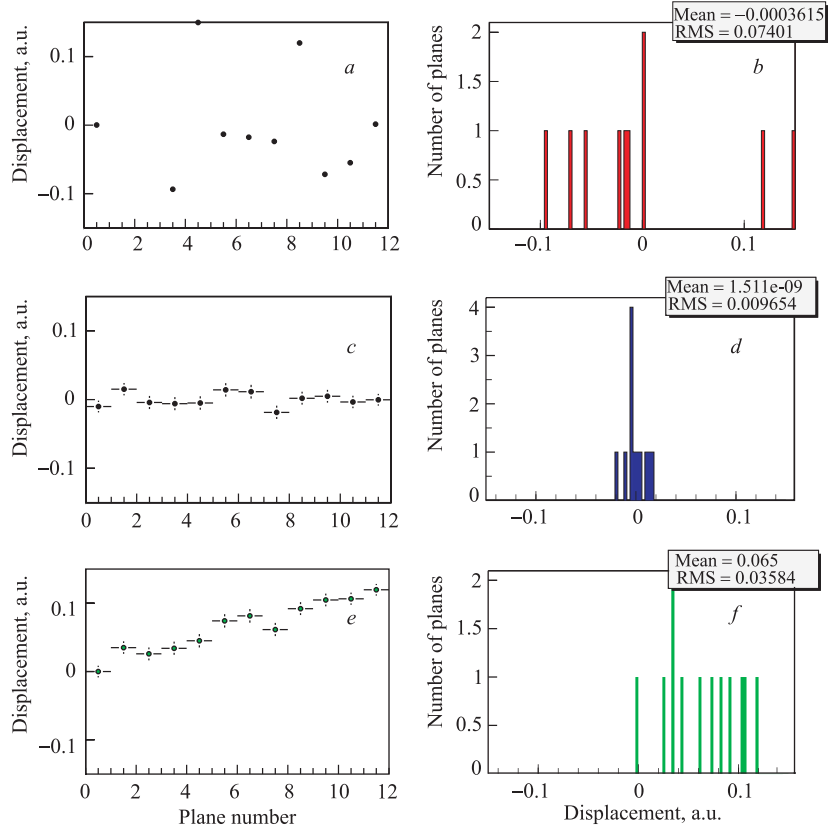


Fig. 5. Simple Monte Carlo test of constraining: *a, b*) initial misalignment; *c, d*) after alignment with constraints; *e, f*) after unconstrained alignment

Figures 5, *a, b* show the initial misaligned detector. Figures 5, *c, d* show this detector after the alignment with properly constrained matrix C . In case the matrix C remains unconstrained, the solution could be biased, as is shown in Figs. 5, *e, f*.

3. STUDY OF OTR ALIGNMENT IN MONTE CARLO

3.1. Monte Carlo Model of HERA-B OTR. The outer part of HERA-B main tracker (OTR) is designed to detect charged particle tracks with distance from proton beam $r \geq 20$ cm [2].

The layout of OTR superlayers used for 2002–2003 data taking is shown in Fig. 6 (without the magnet superlayer). The proton beam comes from the upper right corner of this picture.

The pattern recognition chambers of the outer tracker (PC01–PC04) are installed in the field-free region between the magnet and the RICH. It is about 1.5 m long, the size of superlayers is about 3×4 m. PC chambers play the key role in track reconstruction chain and particle momentum estimation, that is why it is necessary to align this subsystem.

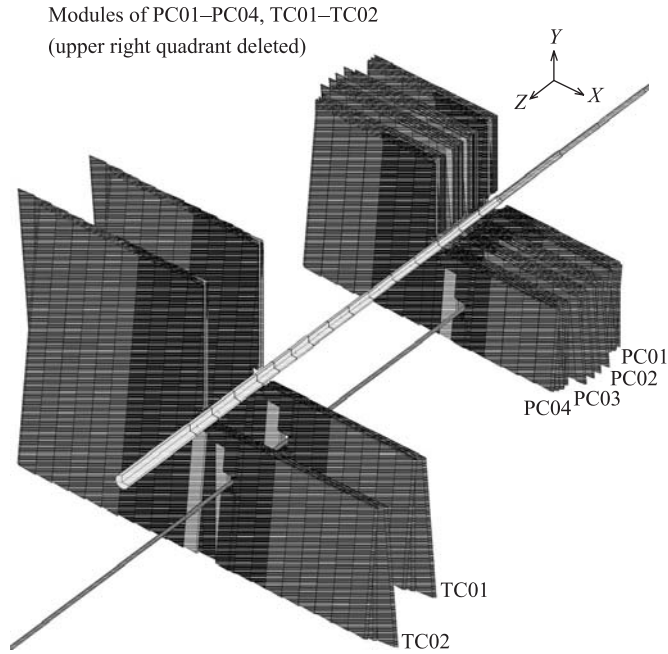


Fig. 6. GEANT model of OTR geometry

HERA-B OTR PC has 998 separate sensitive planes — GEDEs. Some GEDEs are combined in modules. The OTR layers are built from honeycomb drift cells. In the outer part the size of a drift tube cell is 10 mm, while in the inner part, where occupancy is higher, the size of a cell is 5 mm. The distances between wires in a monolayer (the so-called «itches») are 17.3 and 8.7 mm, for 10- and 5-mm cells, respectively. In our study each separate GEDE can be considered as a rigid body and be shifted in two directions along X and Z axes (coordinate y along the wire is not measured in the drift tubes) and rotated around three axes X, Y, Z . Therefore, the most general problem has ≈ 5000 alignment parameters.

The effects of magnetic field in the PC region are small enough to be not considered. Applying momentum cut during track selection, we could also neglect the multiple scattering in PC.

The interaction of particles passing through the detector is simulated using GEANT 3.21. Simulation includes all known effects, such as bremsstrahlung, Coulomb scattering, etc. Particles hitting the detector produce Monte-Carlo Impact Points (MIMP).

The next step is simulation of detector response. There exist several modes of digitization in OTR. The simplest digitization includes MIMP smearing with some predefined width, while the most advanced version implements simulation for all processes in drift cell, such as noise and inefficiency, cross-talk. The output of digitization is stored in HITB table in the same form as real data after hit preparation. Detector response and reconstruction are done within ARTE package.

3.2. Misalignment Simulation. To simulate misaligned detector within ARTE environment, the following procedure is used:

- MIMP data are acquired using HBGEAN assuming nominal geometry without any misalignment applied;
- misalignment is introduced in detector model before the digitization step — shifting the local coordinate plane of every GEDE, we will get all digitized information also shifted; digitized hit information is stored in HITB;
- track building should be done assuming nominal geometry, before reconstruction geometry is restored to normal, but HITB remains shifted;
- hit to track association was done using Monte Carlo truth — we just follow the MC track looking for digitized hits, close to it.

In order to test model, algorithm and software, we assume misalignment only in the direction transverse to the detector axis and GEDE wires. We could use the following straight-line track model in misaligned detector:

$$u_i^j = \mathbf{a}_i \mathbf{p}^j = x_0^j \cos \alpha_i + t_x^j z \cos \alpha_i + y_0^j \sin \alpha_i + t_y^j z \sin \alpha_i + \Delta u_i, \quad (27)$$

where $\mathbf{a}_i = \{x_0, t_x, y_0, t_y, \Delta u_i\}$ is the row of track and alignment parameters and $\mathbf{p} = \{\cos \alpha, z \cos \alpha, \sin \alpha, z \sin \alpha, 1\}^T$ is the column of parameters derivatives; x_0, t_x, y_0, t_y are the track shifts and slopes in XY and YZ planes; Δu_i is the misalignment transverse to the detector axis and wires; α_i is the stereo angle on the i th plane. This assumption allows us to keep the problem linear. In order to reach reasonable alignment precision, we have to use $\approx 10^5$ tracks, so the number of track parameters $\approx 10^5 - 10^6$. After applying Blobel's reduction we have linear least squares problem for alignment parameters only.

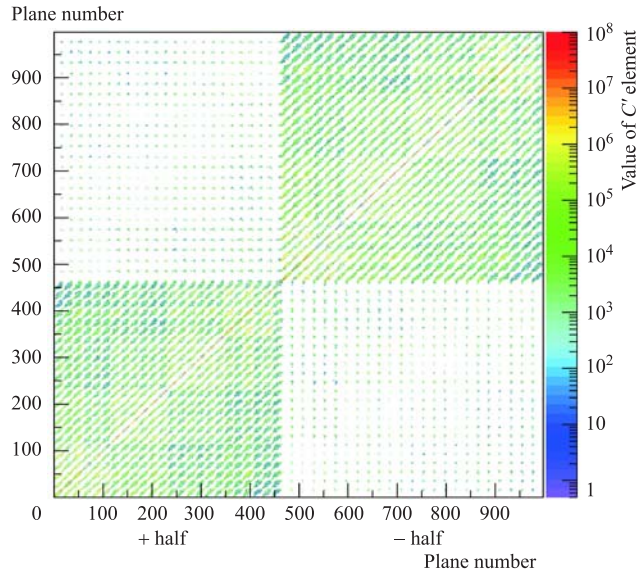


Fig. 7. So-called «floor plan» of the normal equation matrix of OTR alignment problem after dimension reduction

The matrix shown in Fig. 7 is obtained assuming only Δu shifts are included, in this case the matrix dimension is equal to the number of aligned units. HERA-B OTR consists of two halves, divided approximately by plane $x = 20$ cm. The majority of tracks pass through one or the other half, and only a small number of tracks go through both halves. The floor plan shows that the correlation between the parameters coming from the same half is strong, while the correlation between different halves is weak. If this second correlation would be exactly zero, it means that the two halves of OTR are completely independent and not connected by tracks to each other. In this case we need to constrain external degrees of freedom for both parts separately. If this connection exists, even if small, we need to estimate if it is strong enough to glue the halves together.

Analyzing nonzero elements of the cross-halves part of C' , we can conclude that the connection between the halves is realized mostly via the overlapped regions of the rotated layer and not only via low-momentum tracks. That means that even after momentum cut (needed for good track selection) we have OTR halves connected by tracks passing through both halves.

3.3. SVD Analysis. In order to estimate the number of external degrees of freedom, one could apply singular value decomposition to the reduced matrix of normal equation C' .

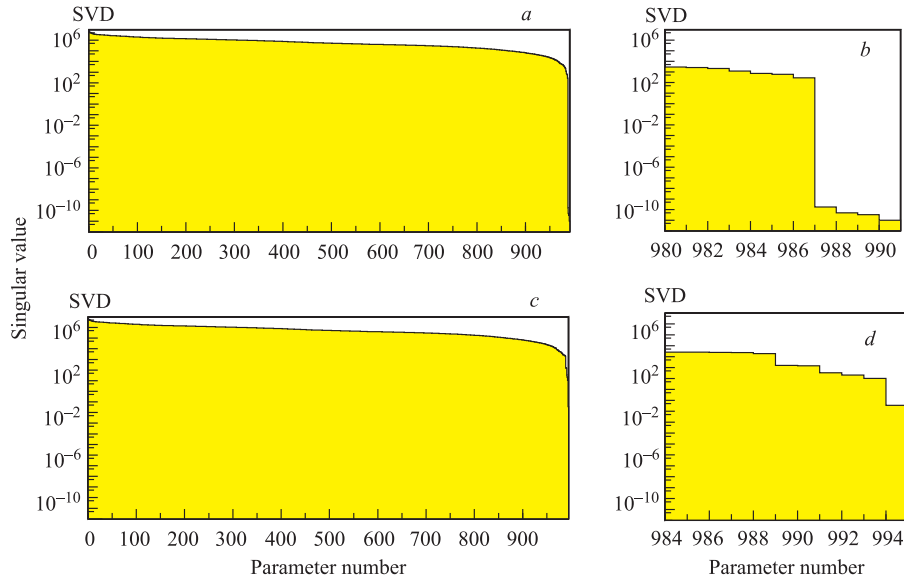


Fig. 8. Singular value spectra for unconstrained problem (a, b) and the problem constrained by Lagrange multipliers (c, d)

Figures 8, a, b show singular value spectrum for the initial unconstrained matrix C' . It is a decreasing sequence with a sharp drop on its tail. This drop represents singularity of C' , or in other words the presence of unconstrained degrees of freedom. The number of values in the spectrum after the drop, or the number of singular values equal to zero within computer precision yields the number of unconstrained degrees of freedom. In our case $N_{\text{unconstr}} = 4$. From the study with simple geometry, one can conclude that the unconstrained degrees of

freedom are two shifts (translations in x_0 and y_0) and two shearings (translations in t_x and t_y). SVD shows that from the point of view of external degrees of freedom the OTR is a rigid body but not the two separate halves, otherwise N_{unconstr} would be 8 (4 for every part).

To fix these degrees of freedom for HERA-B OTR case, there are two solutions: fix some alignment parameters Eq. (13) or apply some constraint to the whole set of alignment parameters Eq. (15). In order to apply the first solution, one should fix the position of four planes — two pairs of GEDE with nonzero relative stereo angle (the so-called «crosses»). Longer distance between crosses gives better fixing. The condition number κ of C' matrix is the estimation of fixing efficiency: a lower κ means a better conditioned problem (Fig. 9).

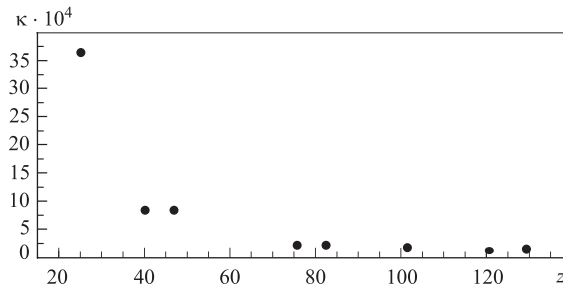


Fig. 9. Condition number κ decreasing with increasing distance between fixed «crosses»

To fix external degrees of freedom constraining the whole set of parameters Eq. (15), one needs to properly choose values for w_i . Our case study with simplified geometry shows that to fix overall shift, w_i should be constant (see Eq. (25)), and to fix shearing, w_i should be proportional to Δz (see Eq. (26)). Taking planes rotation into account, the view of four weight vectors used to fix four degrees of freedom for HERA-B OTR looks like

$$\begin{aligned}
 w_i^1 &= \cos \alpha_i && \text{— shifts } (x_0), \\
 w_i^2 &= \cos \alpha_i z_i && \text{— shearings } (t_x), \\
 w_i^3 &= \sin \alpha_i && \text{— shifts } (y_0), \\
 w_i^4 &= \sin \alpha_i z_i && \text{— shearings } (t_y).
 \end{aligned} \tag{28}$$

3.4. Test of Constraining. To check different ways of fixing external degrees of freedom, we introduced some random shifts (normal distributed with width $300 \mu\text{m}$) in PC, with some number of tracks passing through it, then track information was used by alignment procedure.

If no constraints are used the matrix C' is singular and solution could not be obtained using normal Gauss elimination technique. But the Millepede algorithm uses some custom modification of this technique, it allows one to get solutions even if the matrix is singular [5]. As shown in Figs. 10, *c*, *d*, alignment in layers could be achieved, but since some parameters are not fixed, its values are unpredictable and the solution is unusable.

Figures 10, *e*, *f* represent the situation when the linear constraints defined by Eq. (15) with weights calculated according to Eq. (28) are applied. These figures show that misalignment is reduced and the external degrees of freedom are fixed. The same is for the case when the alignment problem is constrained by fixing two «crosses» (Figs. 10, *g*, *h*). The remaining misalignment here is defined by the number of tracks used in alignment procedure. A detailed study of this dependence is presented in the next subsection.

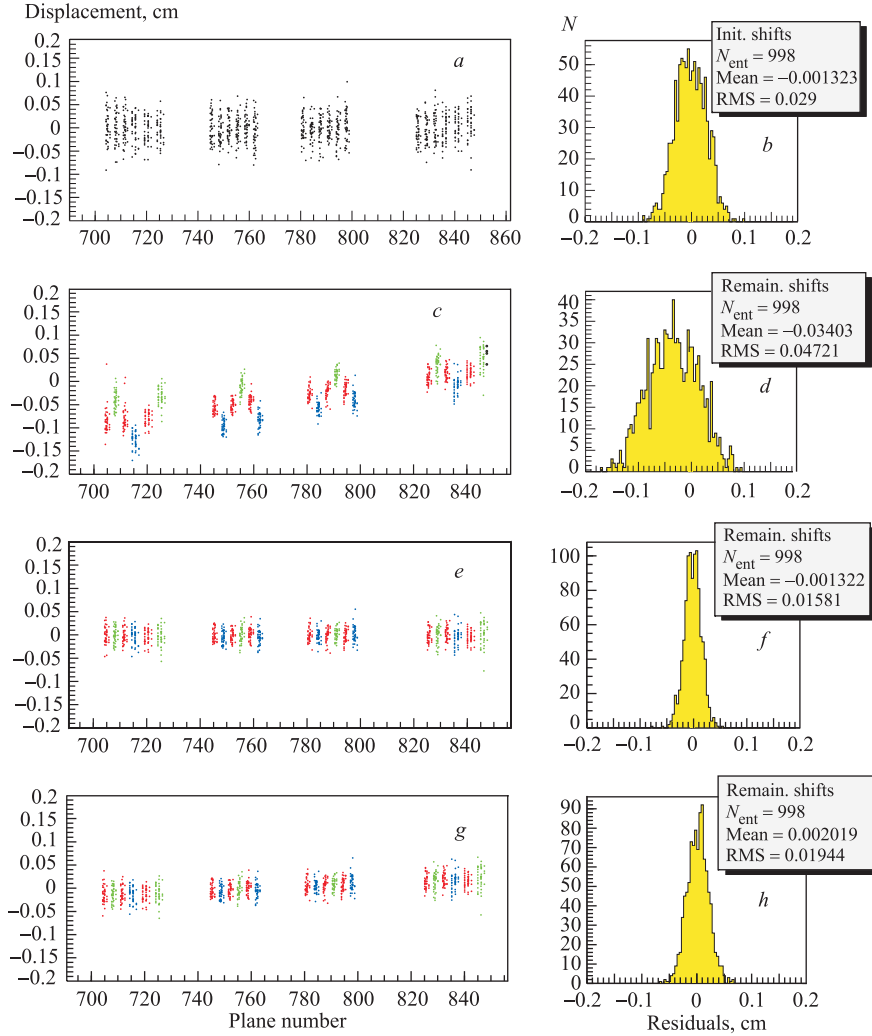


Fig. 10. Remaining misalignment distributions for different ways of constraining C'

3.5. Alignment Quality vs. Number of Tracks Used. The precision of determination of alignment parameters with Monte Carlo tests depends on some factors, including initial misalignment, accuracy of single hit measurement, number of tracks used and also correctness of algorithm implementation. Taking into account the computational complexity of the algorithm, some additional questions arise, for example, accumulating of round-off errors.

Absolute Remaining Misalignment. To check the algorithm, we run it over some data samples with different number of tracks and different introduced misalignments. The first value we watched to estimate the alignment quality on Monte Carlo was the width of the remaining misalignment distribution, and we also checked how it depends on the introduced misalignment and the number of tracks used (Fig. 11).

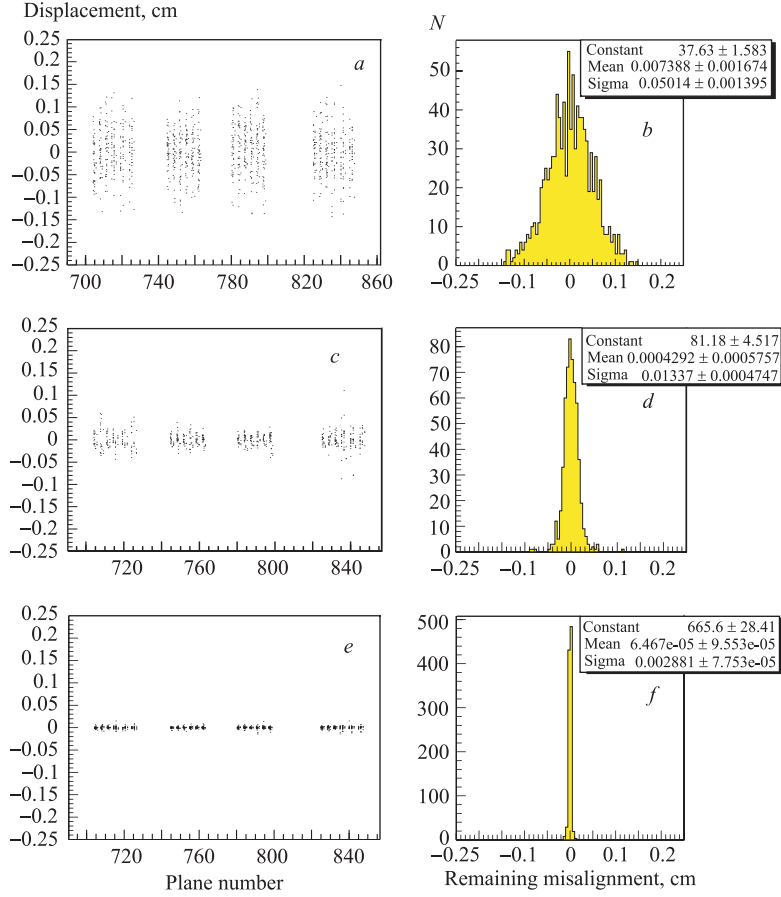


Fig. 11. Remaining misalignment vs. the number of tracks used (initial misalignment is $500 \mu\text{m}$): *a, b*) initial misalignment; *c, d*) remaining misalignment after 10000 tracks are used; *e, f*) remaining misalignment after 250000 tracks are used

Figure 12, *a* shows the dependence of the remaining misalignment on the number of tracks used, and we can conclude that it follows the expected $1/\sqrt{N_{\text{track}}}$ law. It also helps to estimate the number of tracks we should use $\approx 10^5$.

Normalized Residuals. One of the advantages of using linear least squares minimization is the ability to get error estimations using error propagation technique. In our case the error estimations for alignment parameters are presented by the diagonal elements of C'^{-1} .

Studying this alignment method with Monte Carlo model, it is worth checking the agreement between the error estimations σ_{a_i} predicted by error propagation and real residuals between the introduced misalignment a_i^0 and the estimated alignment parameters a_i .

It could be done analyzing the distribution of the so-called normalized residuals:

$$r_n = \frac{a_i^0 - a_i}{\sigma_{a_i}}.$$

We check the width of normalized residual distributions for some samples with different initial misalignment and number of tracks.

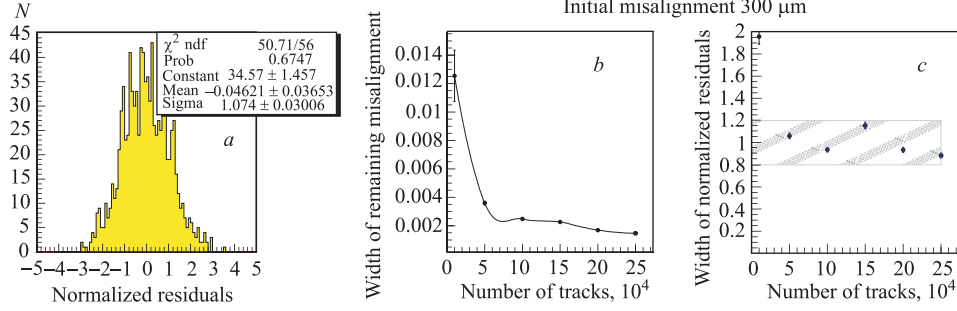


Fig. 12. a) Normalized residuals, initial misalignment 300 μm , 250000 tracks are used for alignment; b, c) remaining misalignment and width of normalized residuals vs. the number of tracks used for different initial misalignments

There are several physics effects (multiple scattering, remaining magnetic field, etc.) which exist but are not taken into account in our model. There are also some computational difficulties we face (extremely large number of parameters, matrix singularity, round-off error accumulations). Figure 12, b shows that the widths of these distributions are always close to 1, it is what one would expect if the problem is still linear, the algorithm is implemented correctly and it allows one to deal with the problem of such a dimension (10^6 parameters).

4. ALGORITHM GENERALIZATION FOR THE CASE OF NONLINEAR TRACK MODEL

The alignment parameters discussed above — shifts perpendicular to the detector axis and wires, are the good benchmarks to test the algorithm with Monte Carlo, since the track model remains linear after introducing alignment parameters:

$$u_i^j = \mathbf{a}_i \mathbf{p}^j = x_0^j \cos \alpha_i + t_x^j z \cos \alpha_i + y_0^j \sin \alpha_i + t_y^j z \sin \alpha_i + \Delta u_i. \quad (29)$$

But after applying Δu , $\Delta \alpha$ and Δz misalignments together:

$$\begin{aligned} u &= x_0 \cos(\alpha + \Delta \alpha) + t_x(z + \Delta z) \cos(\alpha + \Delta \alpha) + \\ &\quad + y_0 \sin(\alpha + \Delta \alpha) + t_y(z + \Delta z) \sin(\alpha + \Delta \alpha) + \Delta u = \\ &= x_0(\cos \alpha - \sin \alpha \Delta \alpha) + t_x(z + \Delta z)(\cos \alpha - \sin \alpha \Delta \alpha) + \\ &\quad + y_0(\sin \alpha + \cos \alpha \Delta \alpha) + t_y(z + \Delta z)(\sin \alpha + \cos \alpha \Delta \alpha) + \Delta u, \quad (30) \end{aligned}$$

the problem becomes nonlinear — the track model includes terms where the alignment parameters are multiplied by other alignment parameters and track parameters. Expanding this

equation to Taylor series and leaving only linear terms in misalignment corrections, we get

$$u = x_0^j \cos \alpha_i + t_x^j z \cos \alpha_i + y_0^j \sin \alpha_i + t_y^j z \sin \alpha_i + \Delta u_i + \\ + (-x_0 \sin \alpha - t_x z \sin \alpha + y_0 \cos \alpha + t_y z \cos \alpha) \Delta \alpha + (t_x \cos \alpha + t_y \sin \alpha) \Delta z. \quad (31)$$

The problem is still nonlinear — derivatives $\frac{\delta u}{\delta \alpha}$ and $\frac{\delta u}{\delta z}$ depend on local track parameters. To manage it we use two iterations. First we suppose that misalignment corrections are fixed and fit only local parameters $\{x_0, t_x, y_0, t_y\}$ for each track. Then the fitted local parameters are used as fixed values in these derivatives.

The more alignment corrections are accounted in the track model, the more detector geometry transformations are possible. For this set of alignment parameters, seven degrees of freedom are not constrained: overall detector shifts D_x , D_y and D_z , shearings S_{xz} and S_{yz} , scaling in Z and overall rotation around Z . Here, as in the simplest linear case, two strategies of constraining these transformations are applicable: fixing some planes or applying Lagrange multipliers.

For the first case it is enough to fix two pairs of detector planes with different stereo angles («crosses»), as was described before.

To constrain external transformations by Lagrange multipliers, one should use the weights calculated according to Eq. (28) and some additional weight vectors to fix new external degrees of freedom:

$$\begin{aligned} w_i^5 &= 1 \text{ — overall shift in } z, \\ w_i^6 &= z_i \text{ — scaling in } z, \\ w_i^7 &= 1 \text{ — overall rotation around } z. \end{aligned} \quad (32)$$

The first four constraints should be applied to Δx misalignment parameters set; the fifth and sixth, to Δz ; and the last weight vector, to rotational misalignment parameters $\Delta \alpha$.

CONCLUSION

The method of alignment parameters estimation, based on simultaneous fit of track and alignment parameters, has been applied to HERA-B OTR PC Monte Carlo model.

At first, the mathematical idea and developed software have been tested on a simple detector model, a special study using singular value decomposition has been done to explore external degrees of freedom and the ways of fixing them.

Then the same combination of techniques has been applied to the complex GEANT model of HERA-B OTR PC including the full geometry description, detector response, channel and chamber inefficiencies. The initial functional minimization problem with $\approx 10^6$ parameters was reduced to an affordable size $\approx 10^3$ parameters using Blobel's approach. The number of external degrees of freedom has been estimated using singular value decomposition. Several methods of constraining these degrees of freedom have been tested.

The full set of desired alignment parameters makes the functional a nonlinear one. Linearization of the problem was proposed.

The method appears to be suitable to solve the problem with $\approx 10^6$ track parameters and $\approx 10^3$ alignment parameters.

Acknowledgements. We would like to express our gratitude to A. Spiridonov for rising this problem, M. Bräuer for guiding through the internals of matrix computations, D. Emelianov, D. Krücker, F. Sanchez, T. Zeuner and other colleagues for many useful discussions.

REFERENCES

1. *Hartouni E. et al.* HERA-B Design Report. DESY PRC 95/01. 1995.
2. *Albrecht H. et al.* The Outer Tracker Detector of the HERA-B Experiment. Part I. Detector. Preprint DESY-05-099. physics/0507048. 2005.
3. *Krücker D.* OTR alignment method description. <http://www.desy.de/~kruecker/OTR.html>
4. HERA-B Alignment sub-group home page.
<http://www-hera-b.desy.de/subgroup/alignment/alignment.html>
5. *Blobel V., Kleinwort C.* A new method for the high-precision alignment of track detectors // Conf. on Advanced Statistical Techniques in Particle Physics, Durham, March 18–22, 2002. Preprint DESY-02-077. hep-ex/0208021.
6. *Belotelov I.I., Ososkov G.A.* How to constrain external degrees of freedom during alignment of modern detectors // Czech. J. Phys. 2003. V. 53. Suppl. B441.
7. *Press W.H. et al.* Numerical Recipes in C. The Art of Scientific Programming. 3rd ed. Cambridge University Press, 1992.
8. *Golub G.H., Van Loan C.F.* Matrix Computations. 2nd ed. John Hopkins University Press, 1989.
9. LAPACK: Program library for matrix computation. <http://www.netlib.org/lapack>

Received on October 26, 2005.