Internal Alignment of the SLD Vertex Detectors

David Jackson, Su Dong, Fred Wickens

(RAL, SLAC)

Overview

- Vertex Detectors at SLD
- Residuals used
- Alignment Matrix I
- Shape corrections
- Alignment Matrix II
- Results + History
- Comments

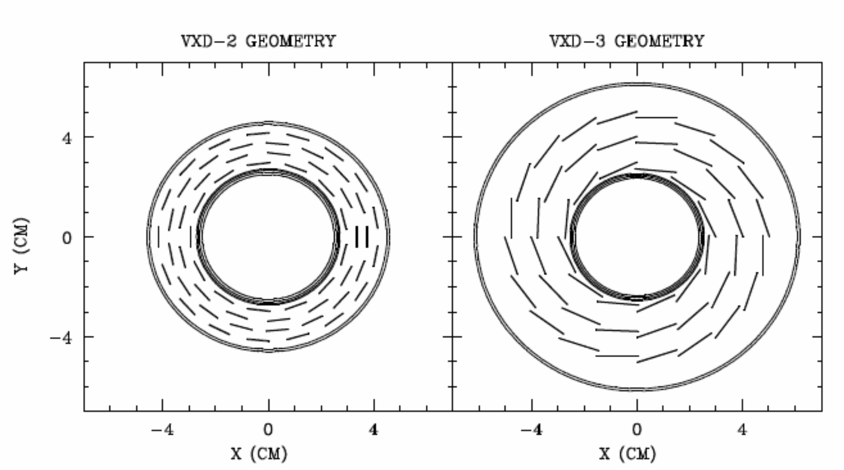
LHC Alignment Meeting
CERN
4th September 2006

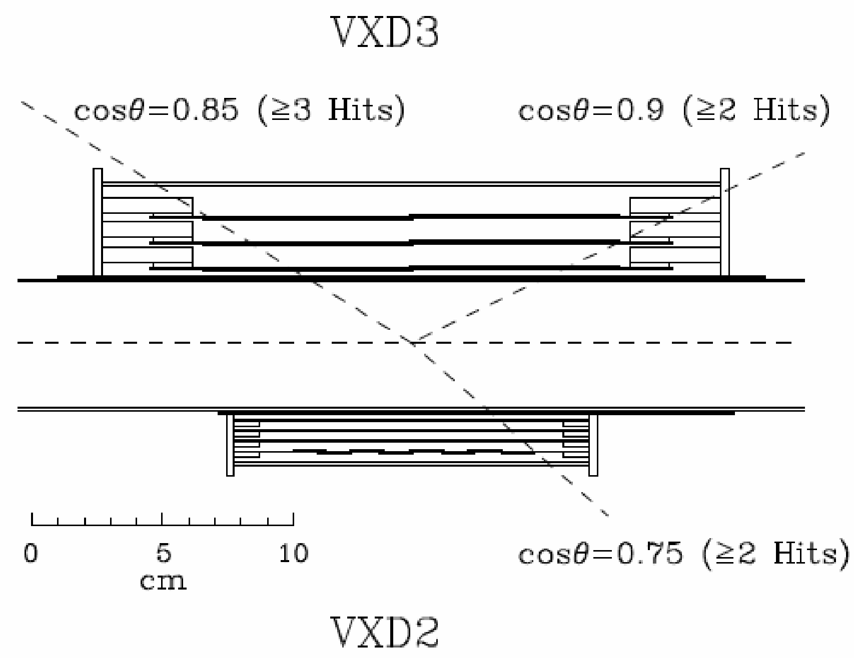
Health Warning

- Work described was done quite a while ago and over a long period (~1992 - 1999+)
- I was never familiar with all of the details, and many of those I did know I have forgotten
- It involves a lot of algebra which I will not attempt to duplicate here
- I will attempt to answer questions, but may need to pass them on to my co-authors
- For the full details see:
  - D.J.Jackson, D.Su, F.J.Wickens; NIM A510, 233 (2003)
  - Or in a slightly expanded form at
    - http://www-sldnt.slac.stanford.edu/sldbb/SLDNotes/sld-note-271.pdf
- This talk is based heavily on a talk by David Jackson given in early 2005 who I am sure got it right - I take responsibility for any errors introduced here

Vertex Detectors at SLD

- VXD2 - 1992 - 1995
  - 60 ladders each with 8 small CCDs (8 x 13 mm)
  - Typically 2 hits/track
  - Operating Temp. 190º K
  - ~150,000 hadronic Z decays
- VXD3 - 1996 - 1998
  - 48 ladders each with 2 large CCDs (16 x 80 mm)
  - Typically 3 hits/track
  - Operating Temp. 220º K (-1996)
185º K (1997-)
  - ~400,000 hadronic Z0 decays
- CCD hit resolution < 5 μm
- Optical surveys ~ 10 μm
  - VXD2 - cold, VXD3 - room temp.
- Stable mechanical support structure
  - Rigid external shell
  - Each ladder 1 fixed end, 1 sliding on precision ceramic blocks





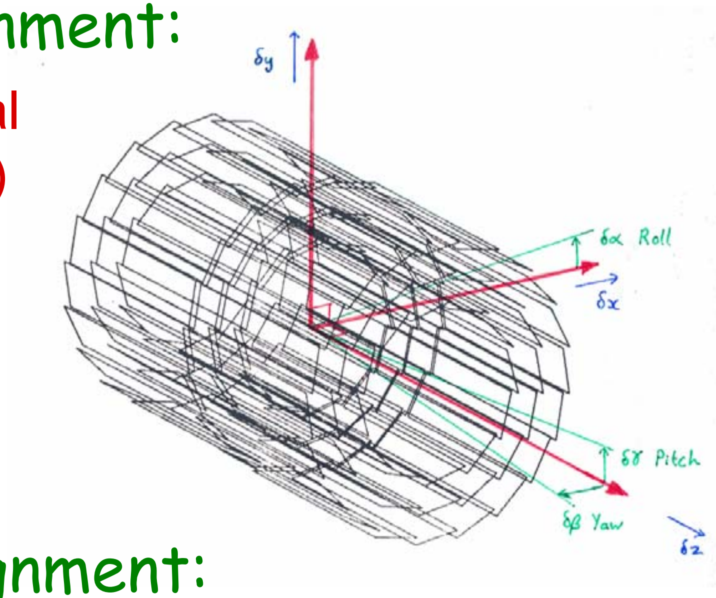
Rigid Body Alignment in 3D:

3 translation + 3 rotation parameters

Global Alignment:

(align to Central Drift Chamber)

1 x



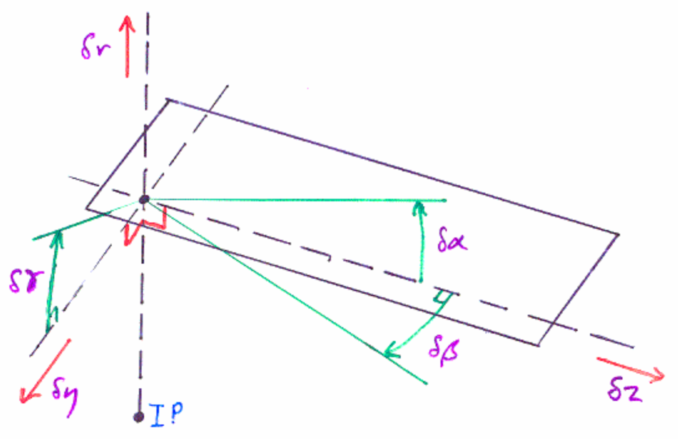
= 6 parameters

□x, □y, □z, pitch, yaw, roll

Internal Alignment:

(mainly internal to VXD3)

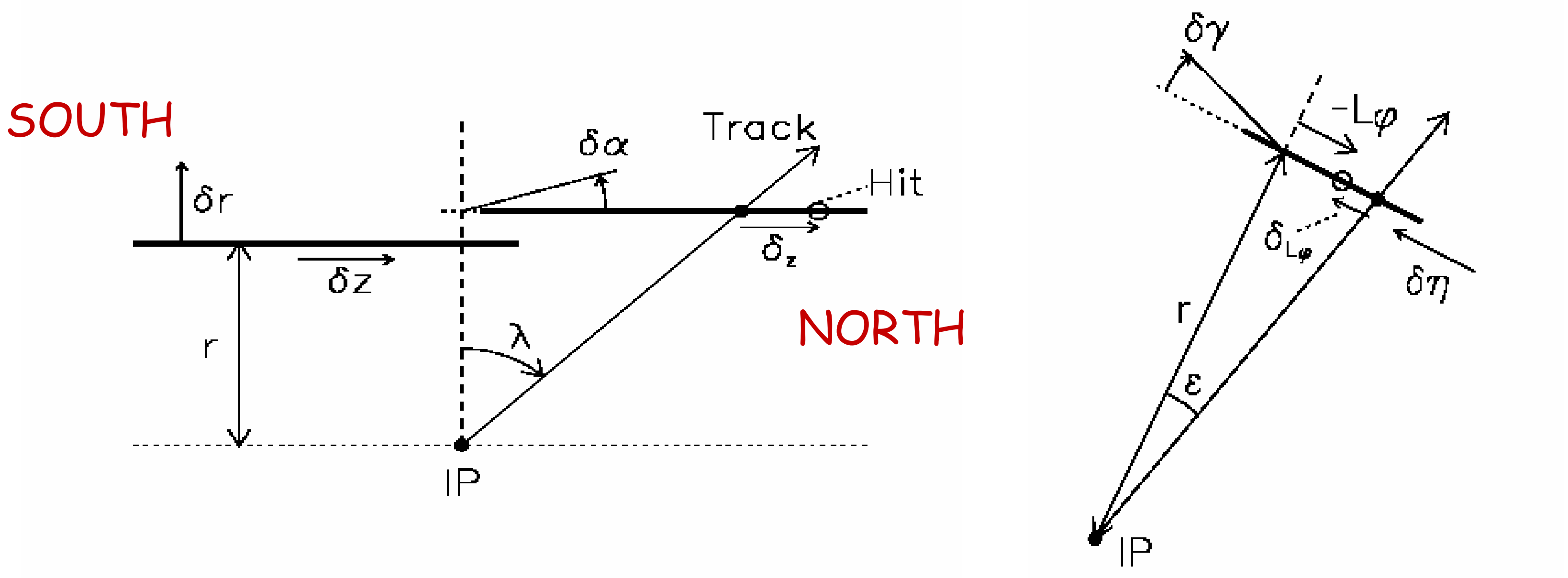
96 x



= 576 parameters

□r, □□, □z,
□□ (pitch) , □□ (yaw), □□ (roll)

Effect of CCD misalignments on the apparent hit position from a known track (from the IP)



(a) rz View

(b) rφ View

$$\delta_z = -\delta z + \delta r \tan\lambda + \delta\alpha\, r \tan^2\lambda + \delta\gamma L_\phi \tan\lambda + \delta\beta L_\phi$$

$$\delta_{L\phi} = -\delta\eta + \frac{\delta r}{r} L_\phi + \frac{\delta\gamma}{r} L_\phi^2 + \delta\alpha L_\phi \tan\lambda - \delta\beta r \tan\lambda$$

Note use of r, L_ϕ and λ, these are defined from the nominal geometry

General form for Residuals

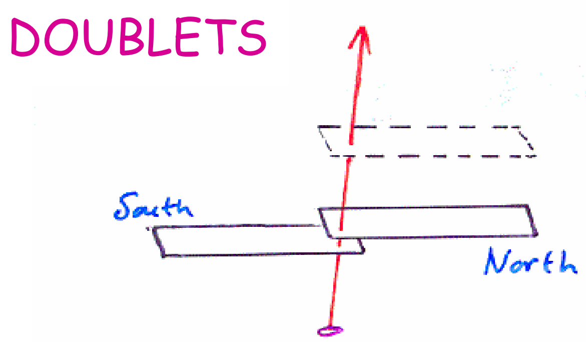
- The CCDs themselves provide the most precise measurements of the track trajectory
- Principal idea was to fix a track to two CCD hits and measure a 'residual' to a third CCD
- The 3 CCDs in each residual contribute to the residual in proportion to a lever-arm weight determined by their relative spacing
- e.g.

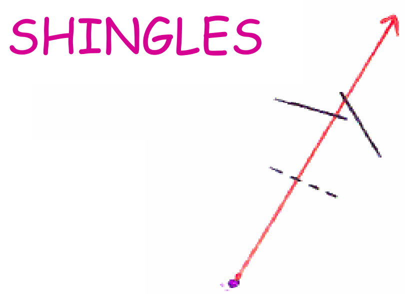
$$\begin{aligned}
\delta_z = \quad & \textstyle\sum_i w_i \big[(-\delta z_i + b_i \delta\beta_i) && 1 \\
& +(\delta r_i + b_i \delta\gamma_i) && \tan\lambda \\
& +\, (r_i \delta\alpha_i) && \tan^2\lambda \\
& +(a_i \delta\gamma_i) && L_\phi \tan\lambda \\
& +(a_i \delta\beta_i) && L_\phi \big]
\end{aligned}$$

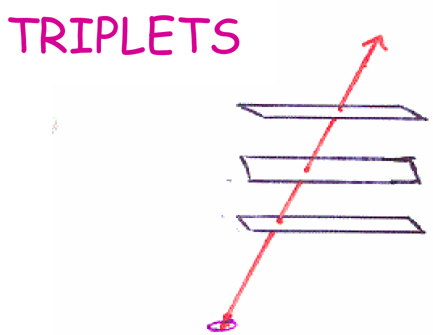
Where the sum is over the 3 CCDs used

Residuals types used

- ‘doublets’
  - use the small overlap region between the 2 CCDs on a ladder
  - connect the North/South halves
  - weight for 3rd CCD is very small
- ‘shingles’
  - use the overlap between adjacent CCDs in the same layer
  - connect the CCDs within each layer
  - weight for 3rd CCD is very small
- ‘triplets’
  - use CCDs from different layers
  - connect the three layers of the detector







…three further residual types were added

These: - are essential to fix opposite sides of the detector
- use layers 1 +3 only

PAIRS

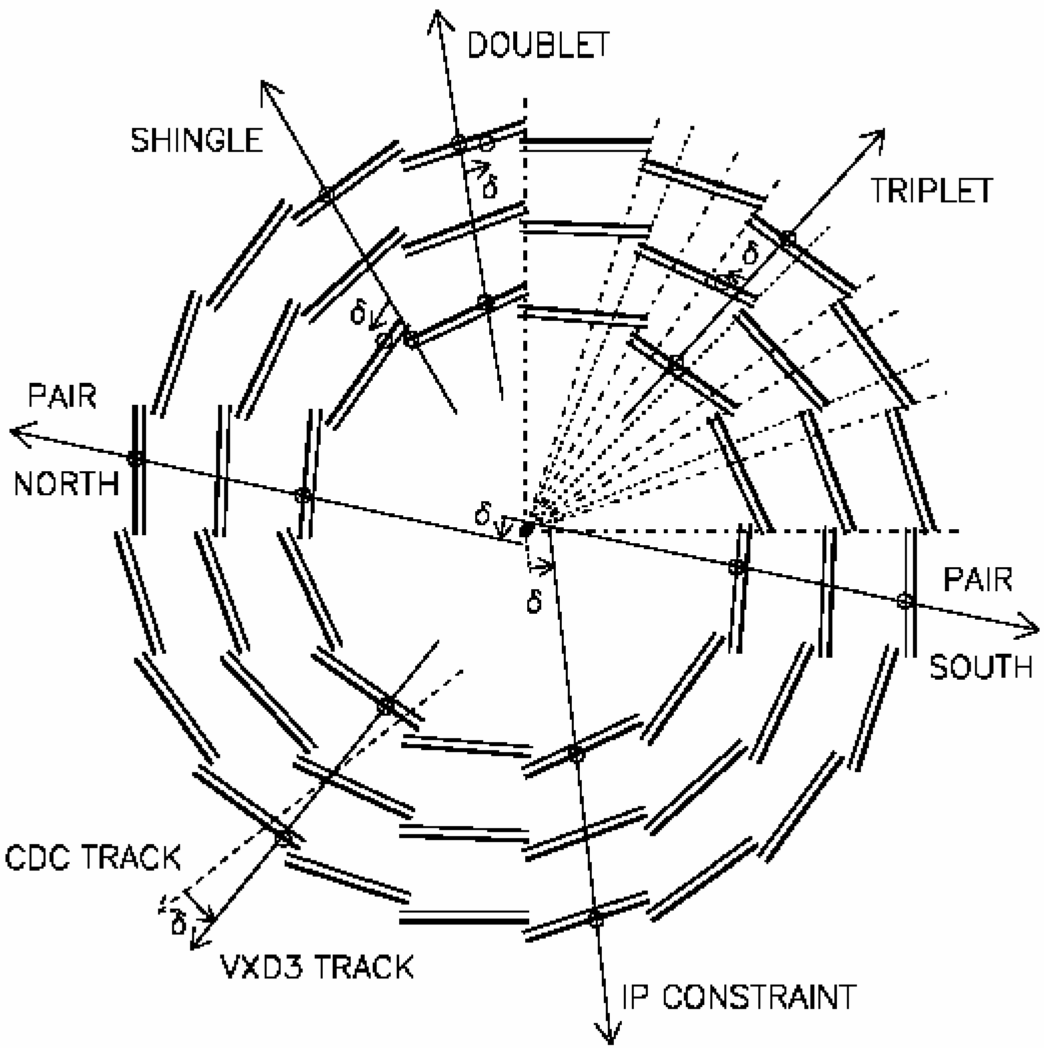
Back-to-back electrons or muons

VXD3 vs CDC Track angle

Should be same for high momentum tracks

IP Constraint

High momentum tracks from light flavours should point to IP



Functional forms of residual distributions

(treating each CCD as a rigid body)

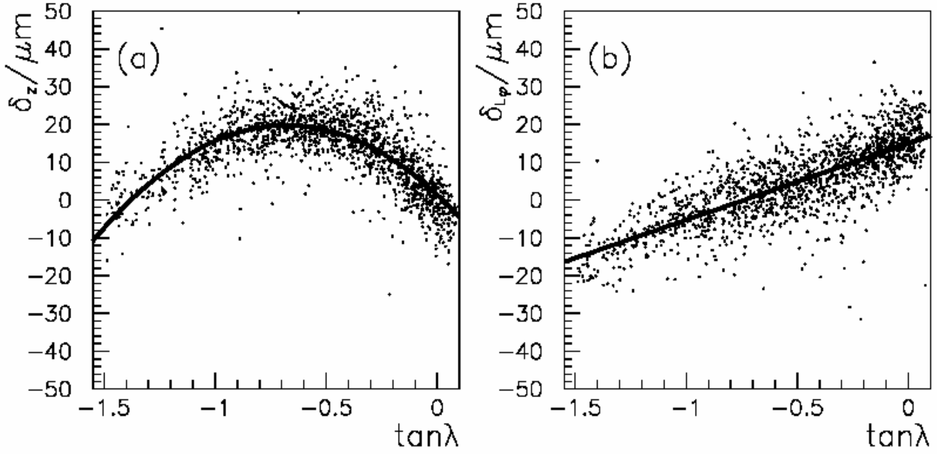
Type	Functional Form	N_I	N_C
Shingles	$\delta_z = s_1^{\parallel} + s_2^{\parallel} \tan\lambda + s_3^{\parallel} \tan^2\lambda$	96	288
	$\delta_{L\phi} = s_1^{\perp} + s_2^{\perp} \tan\lambda$	96	192
Doublets	$\delta_z = d_1^{\parallel} + d_2^{\parallel} \mathrm{L}_\phi$	48	96
	$\delta_{L\phi} = d_1^{\perp} + d_2^{\perp} \mathrm{L}_\phi + d_3^{\perp} \mathrm{L}_\phi^2$	48	144
Triplets	$\delta_z = t_1^{\parallel} + t_2^{\parallel} \tan\lambda + t_3^{\parallel} \tan^2\lambda + t_4^{\parallel} \mathrm{L}_\phi \tan\lambda + t_5^{\parallel} \mathrm{L}_\phi$	80	400
	$\delta_{L\phi} = t_1^{\perp} + t_2^{\perp} \mathrm{L}_\phi + t_3^{\perp} \mathrm{L}_\phi^2 + t_4^{\perp} \mathrm{L}_\phi \tan\lambda + t_5^{\perp} \tan\lambda$	80	400
Pairs	$\delta_{rz} = p_1^{\parallel} + p_2^{\parallel} \tan\lambda + p_3^{\parallel} \tan^2\lambda$	28	84
	$\delta_{r\phi} = p_1^{\perp} + p_2^{\perp} \tan\lambda$	28	56
	$\delta_\phi = p_1^{\phi} + p_2^{\phi} \tan\lambda$	28	56
CDC Angle	$\delta_\lambda = c_1^{\lambda} + c_2^{\lambda} \tan\lambda + c_3^{\lambda} \tan^2\lambda$	56	168
	$\delta_\phi = c_1^{\phi} + c_2^{\phi} \tan\lambda$	56	112
IP Constraint	$\delta_{r\phi} = i_1^{\perp} + i_2^{\perp} \tan\lambda$	56	112
	Total	700	2108

A total of 700 polynomial fits (with 2108 coefficients)

Residual fits

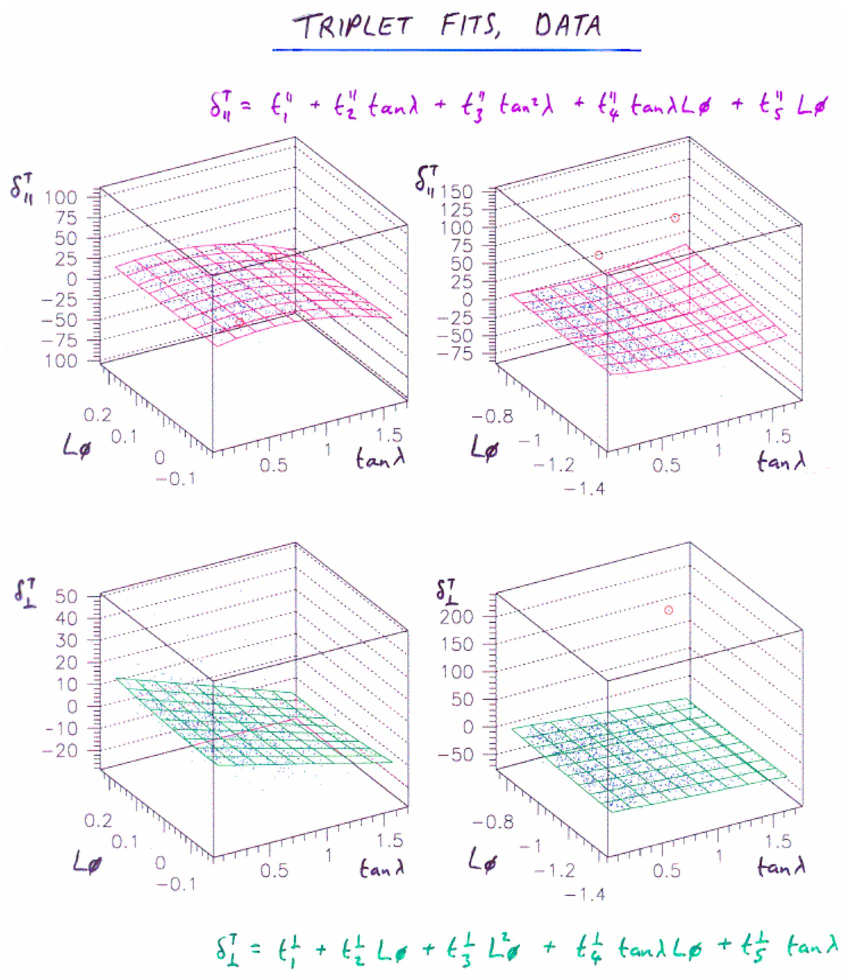
- For each type of residual n-tuples were accumulated for each unique combination of CCDs
- These were then fit to the appropriate functional forms with an automated procedure using MINUIT

- E.g. The two fits to one shingle region
- This shingle conforms very well to the predicted functional forms
- Vertical scatter is due to the intrinsic spatial hit resolution of the CCDs
- The procedure included automatic removal of outliers



Residual fits cont'd

- Examples of triplet fits
- The plots show the two fits to each of two triplet regions (one triplet on left, the other on right)



Internal Alignment Matrix Equation I

The coefficients obtained from the residual fits can also be expressed in terms of a large number of simultaneous linear equations relating them to the unknown alignment parameters; expressing these as a single matrix equation

$$\begin{pmatrix} \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \ldots\ldots\mathbf{A}(w_i, r_i, a_i, b_i, f_i)\ldots\ldots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \end{pmatrix} \begin{pmatrix} \delta \mathrm{z}_1 \\ \vdots \\ \delta \mathrm{z}_{96} \\ \delta\eta_j \\ \delta \mathrm{r}_j \\ \delta\alpha_j \\ \delta\beta_j \\ \delta\gamma_j \\ \delta \mathrm{x} \\ \delta \mathrm{y} \end{pmatrix} = \begin{pmatrix} s_1^{\parallel}(96) \\ s_2^{\parallel}(96) \\ s_3^{\parallel}(96) \\ s_1^{\perp}(96) \\ s_2^{\perp}(96) \\ d_1^{\parallel}(48) \\ \vdots \\ t_1^{\parallel}(80) \\ \vdots \\ \vdots \\ p_1^{\parallel}(28) \\ \vdots \\ c_1^{\lambda}(56) \\ \vdots \\ i_1^{\perp}(56) \\ i_2^{\perp}(56) \end{pmatrix}$$

Weight matrix – determined to an extremely good approximation from the known ideal geometry

Alignment corrections to be determined

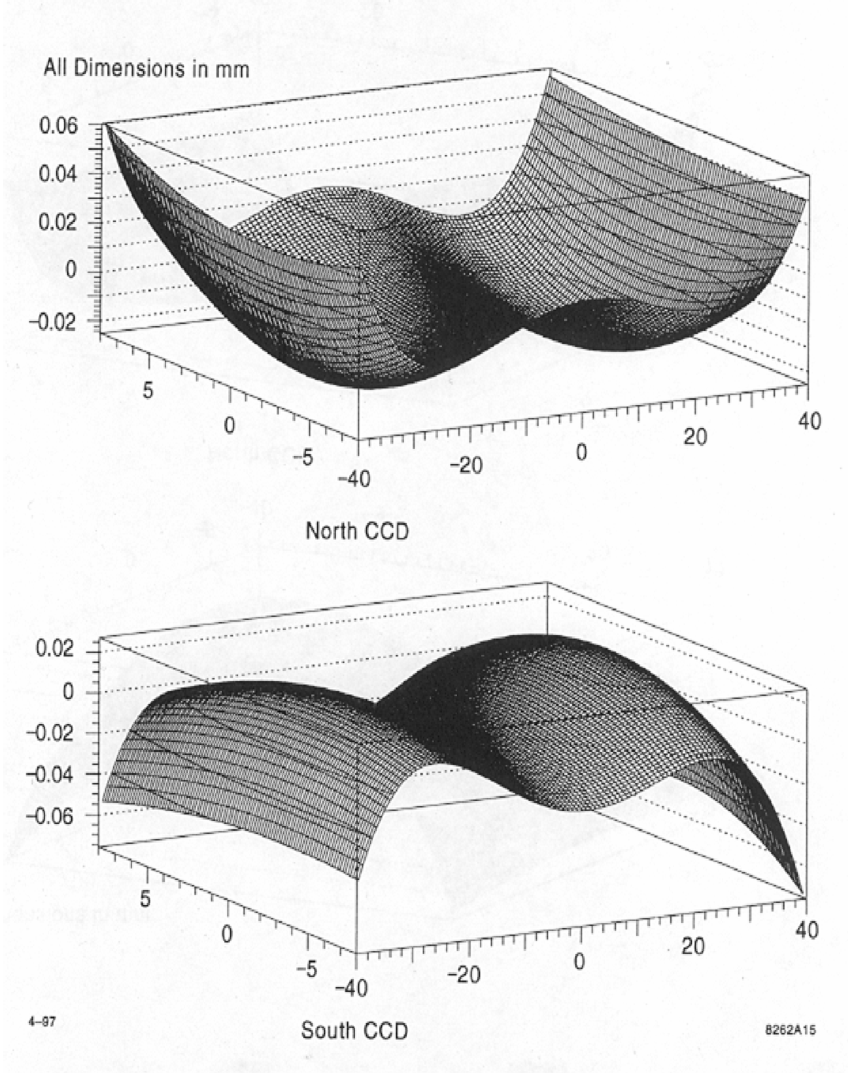
Coefficients measured in residual fits

Internal Alignment Matrix Equation I

- This matrix equation **A x = d** can now be solved, and our chosen method - using a Singular Value Decomposition (see further details in backup foils)
  - Is robust
  - Handles singularities
    - identifies any unknown parameters which are not constrained by the data
  - Provides a 'least squares' solution
- But note that this does not take into account the error and correlation information from the residual fits

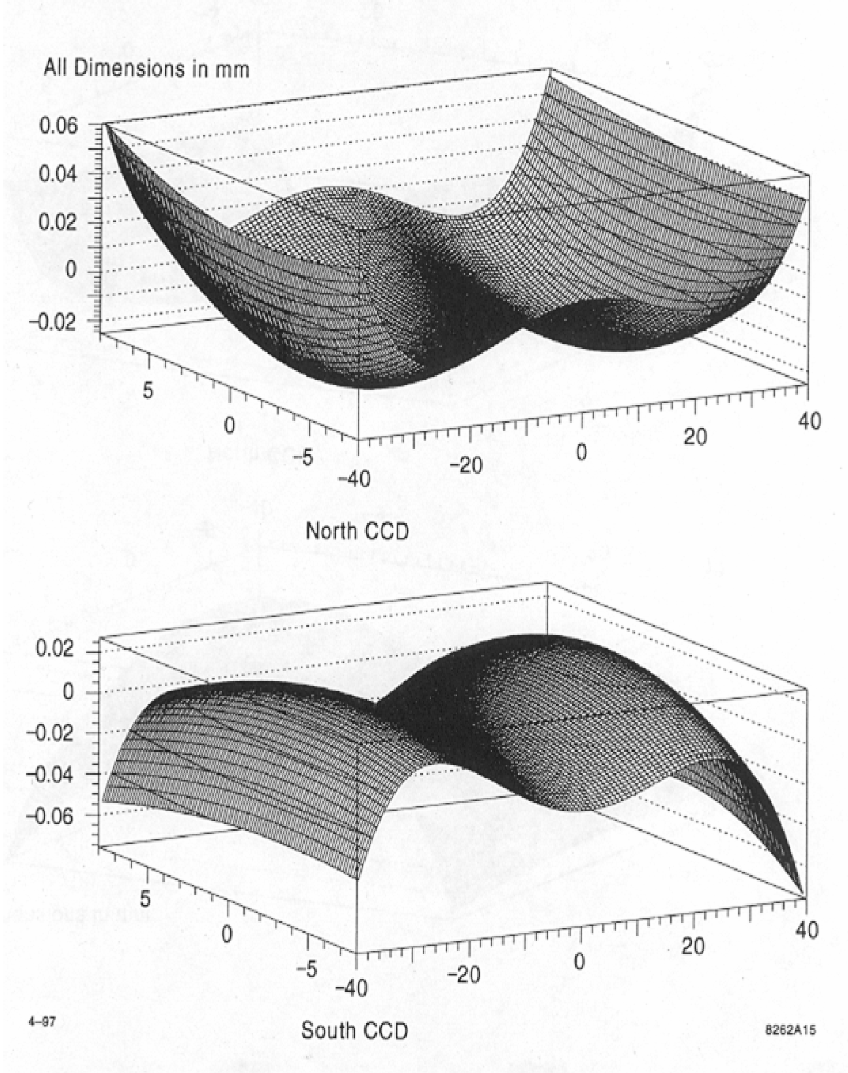
Corrections to CCD Shapes

- The rigid body used for each CCD in the initial internal alignment allowed for a 14-parameter Chebychev polynomial shape fitted to optical survey data



Corrections to CCD Shapes

- However, a large number of track residual distributions showed signs of the CCD shapes deviating from the optical survey data
- The biggest effects could be described by a 4th order polynomial as a function of the z axis
  - Consistent with the dominant "W" shape changing during cool down to operating temperature
- This required 3 extra parameters and introduced higher order terms in $\tan\lambda$



With shape parameters included the same residual distributions were fitted to extended higher order functional forms:

		1	$\tan\lambda$	$\tan^2\lambda$	$\tan^3\lambda$	$\tan^4\lambda$	$\tan^5\lambda$	$L\phi$	$L^2\phi$	$L\phi\tan\lambda$	$L\phi\tan^2\lambda$	$L\phi\tan^3\lambda$	$L\phi\tan^4\lambda$	# FITS	# PARAM.
TRIPLETS	$\delta_z^t =$	$t_1^{\parallel}$	$t_2^{\parallel}$	$t_3^{\parallel}$	$t_6^{\parallel}$ ▲	$t_7^{\parallel}$ ▲	$t_8^{\parallel}$ ▲	$t_5^{\parallel}$		$t_4^{\parallel}$				80	640
	$\delta_{L\phi}^t =$	$t_1^{\perp}$	$t_5^{\perp}$	$t_9^{\perp}$ ▲	$t_{10}^{\perp}$ ▲	$t_{11}^{\perp}$ ▲		$t_2^{\perp}$	$t_3^{\perp}$	$t_4^{\perp}$	$t_6^{\perp}$ ▲	$t_7^{\perp}$ ▲	$t_8^{\perp}$ ▲	80	880
SHINGLES	$\delta_z^s =$	$s_1^{\parallel}$	$s_2^{\parallel}$	$s_3^{\parallel}$	$s_4^{\parallel}$ ▲	$s_5^{\parallel}$ ▲	$s_6^{\parallel}$ ▲							96	576
	$\delta_{L\phi}^s =$	$s_1^{\perp}$	$s_2^{\perp}$	$s_3^{\perp}$ ▲	$s_4^{\perp}$ ▲	$s_5^{\perp}$ ▲								96	480
DOUBLETS	$\delta_z^d =$	$d_1^{\parallel}$						$d_2^{\parallel}$						48	96
	$\delta_{L\phi}^d =$	$d_1^{\perp}$						$d_2^{\perp}$	$d_3^{\perp}$					48	144
PAIRS	$\delta_{rz}^p =$	$p_1^{\parallel}$	$p_2^{\parallel}$	$p_3^{\parallel}$	$p_4^{\parallel}$ ▲	$p_5^{\parallel}$ ▲	$p_6^{\parallel}$ ▲							28	168
	$\delta_{xy}^p =$	$p_1^{\perp}$	$p_2^{\perp}$	$p_3^{\perp}$ ▲	$p_4^{\perp}$ ▲	$p_5^{\perp}$ ▲								28	140
	$\delta_{\phi}^p =$	p_1^{ϕ}	p_2^{ϕ}	p_3^{ϕ} ▲	p_4^{ϕ} ▲	p_5^{ϕ} ▲								28	140
CDC	$\delta_{\lambda}^c =$	c_1^{λ}	c_2^{λ}	c_3^{λ}	c_4^{λ} ▲	c_5^{λ} ▲	c_6^{λ} ▲							56	336
	$\delta_{\phi}^c =$	c_1^{ϕ}	c_2^{ϕ}	c_3^{ϕ} ▲	c_4^{ϕ} ▲	c_5^{ϕ} ▲								56	280
IP	$\delta_{\perp}^i =$	$i_1^{\perp}$	$i_2^{\perp}$	$i_3^{\perp}$ ▲	$i_4^{\perp}$ ▲	$i_5^{\perp}$ ▲								56	280
														700	4,160

The required new fit coefficients ▲ roughly doubling the total number to 4,160

Six examples of the 28 Pair δrz residual fits

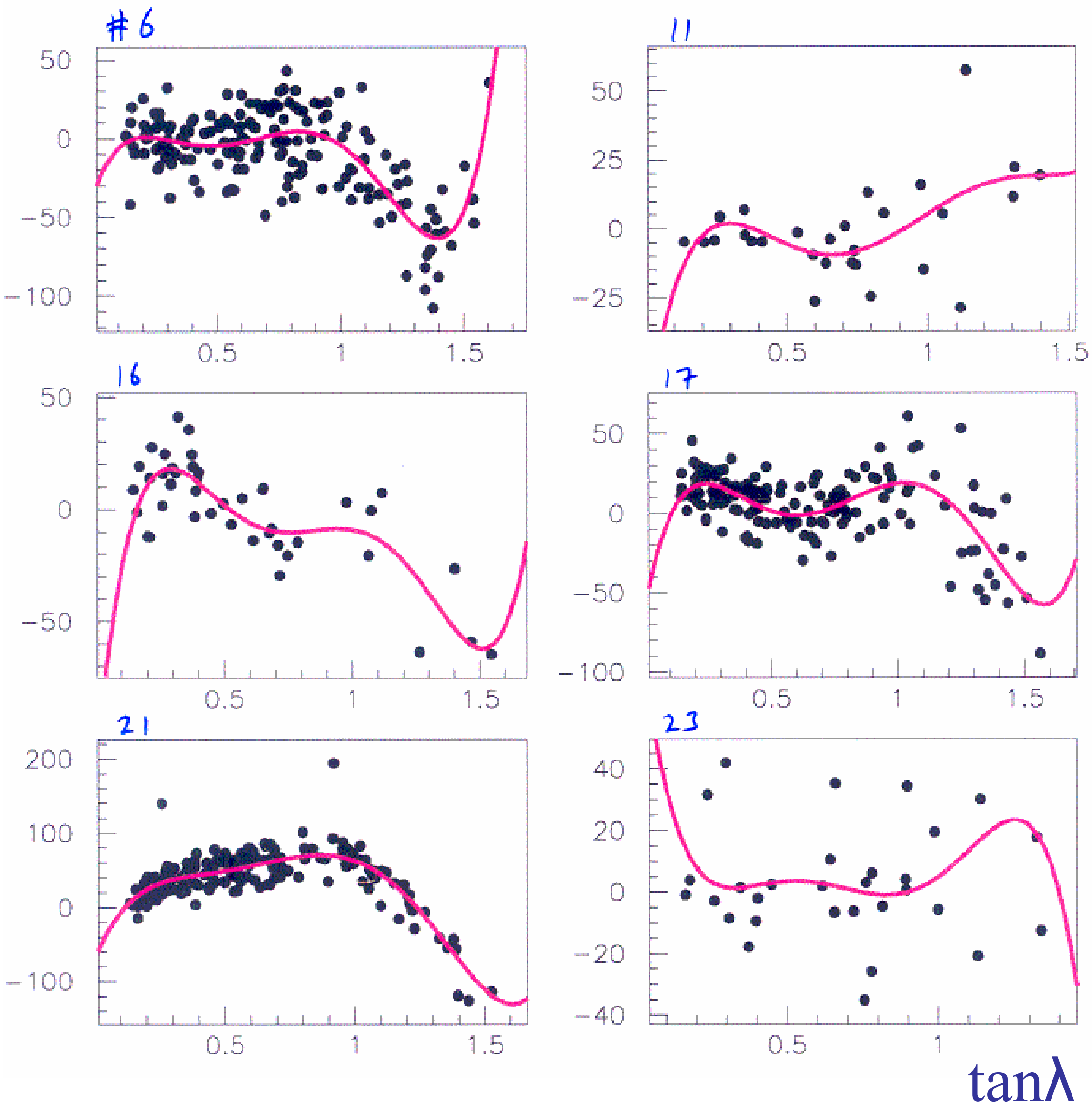
(would take quadratic form without shape corrections)

Pairs, using

$Z^0 \rightarrow \mu^+\mu^-$

$Z^0 \rightarrow e^+e^-$

events, were the most limited in statistics.



Important to correctly take into account correlations in each fit.

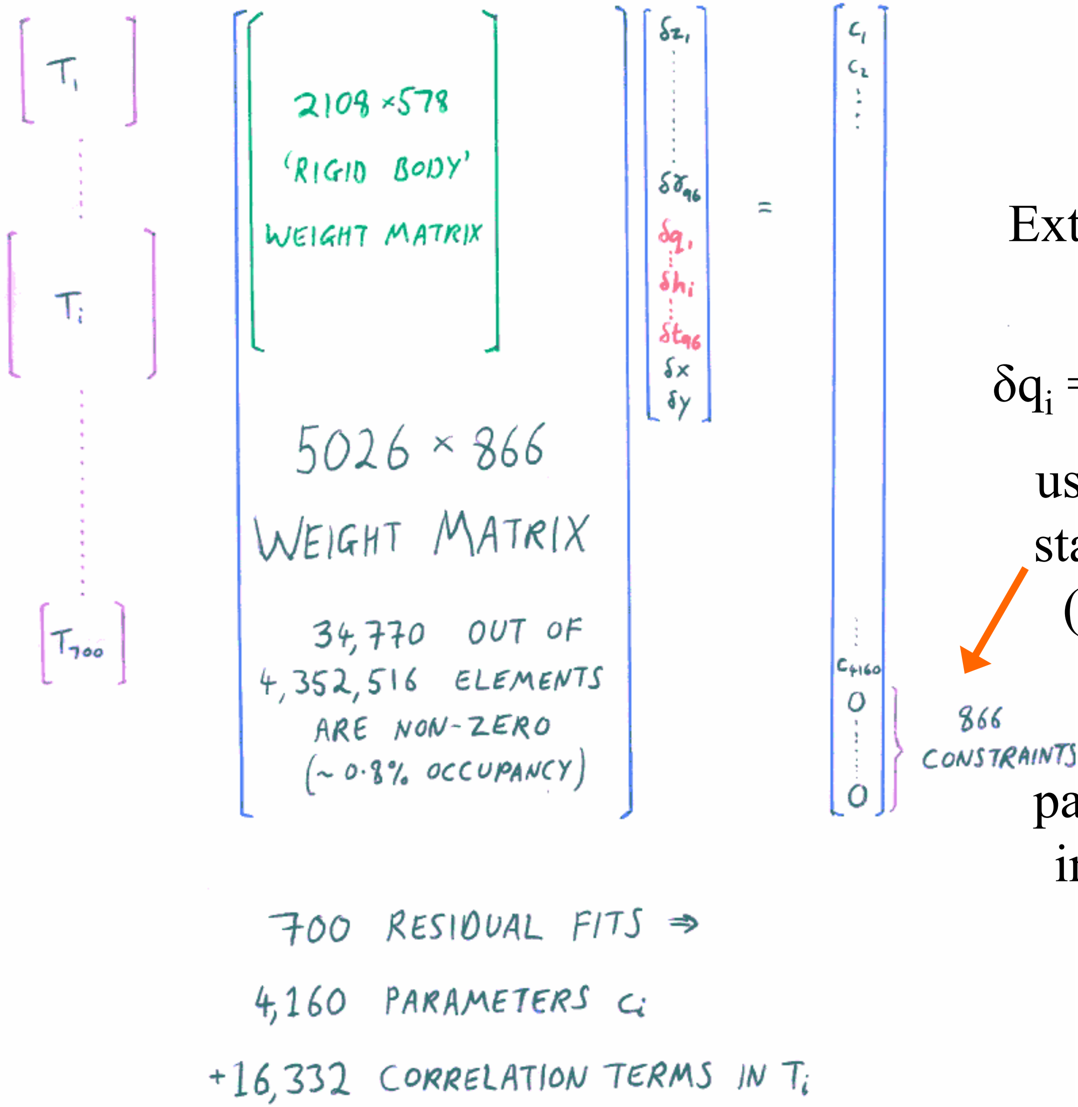
Internal Alignment Matrix Equation II

- To remove the correlations and to take account of the errors we redefined the basis of parameters for each residual fit
- Thus if each residual fit produces a vector **p** of *n* parameters with covariance matrix **W**
- **W** was decomposed to $\mathbf{HH}^{T}$, where **H** is a non-singular lower triangular matrix
- The parameter vector was transformed to $\mathbf{p}' = \mathbf{H}^{-1}\mathbf{p}$, which has a unit covariance matrix (thus removing correlations and including the quality of each fit)
- Gathering together all the elements of the many matrices $\mathbf{H}^{-1}$ from all of the residual fits can produce a single 4160 x 4160 matrix **T**
- Solving the matrix equation **TA x = Td = c** with SVD now gave a ^{2} minimization over all the track residuals for the alignment parameters

Internal Alignment Matrix Equation II

Each of 700 residual fit error matrices used to determine linearly independent basis in each case.

The SVD technique is improved from a 'least squares' to an optimal χ^2 fit.



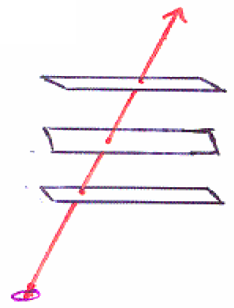
Extra constraints such as

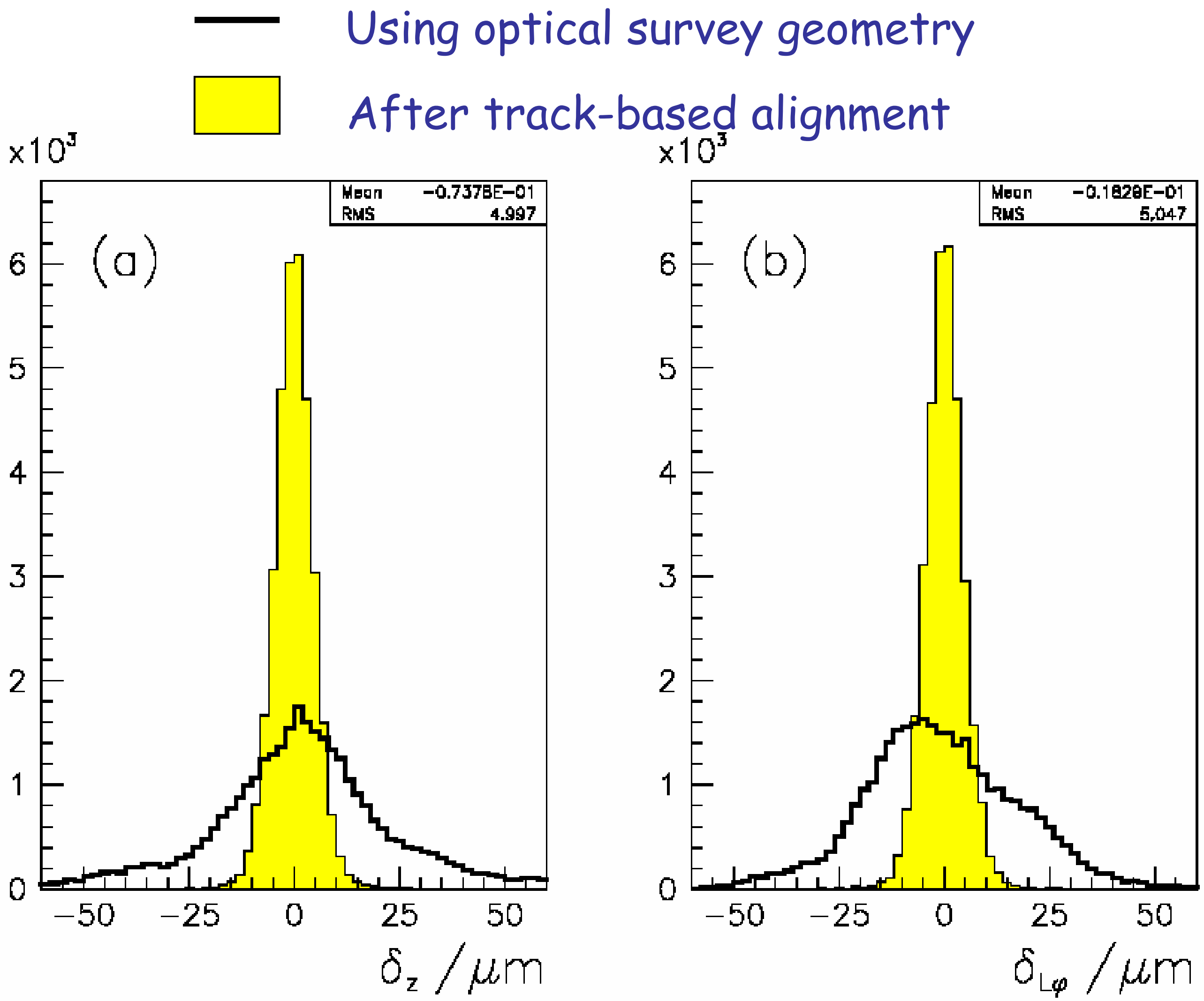
$\delta q_i = 0.0 \pm 5.0\ \mu m$

used to ensure stable solution (where data limited e.g. shape parameters for inner layers)

866 (9 x 96 + 2) alignment corrections to be determined

Results - Triplet Residuals

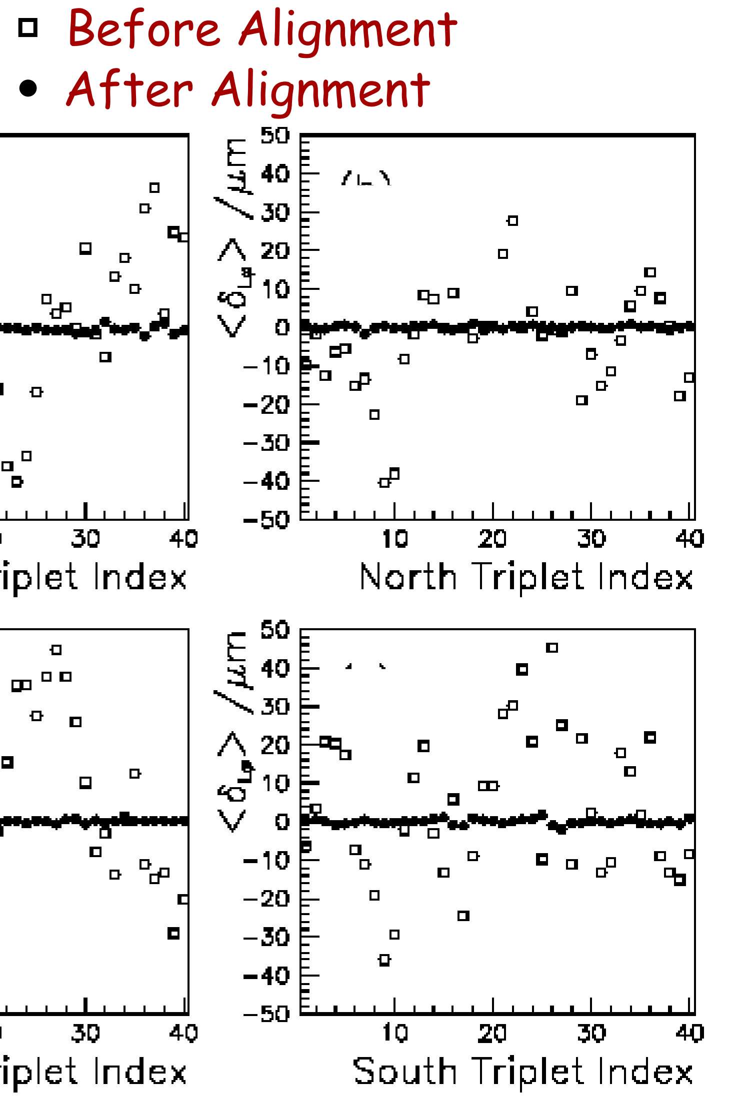




Tracks with
P > 5 GeV

Post-alignment single hit resolution ~ 3.6 μm

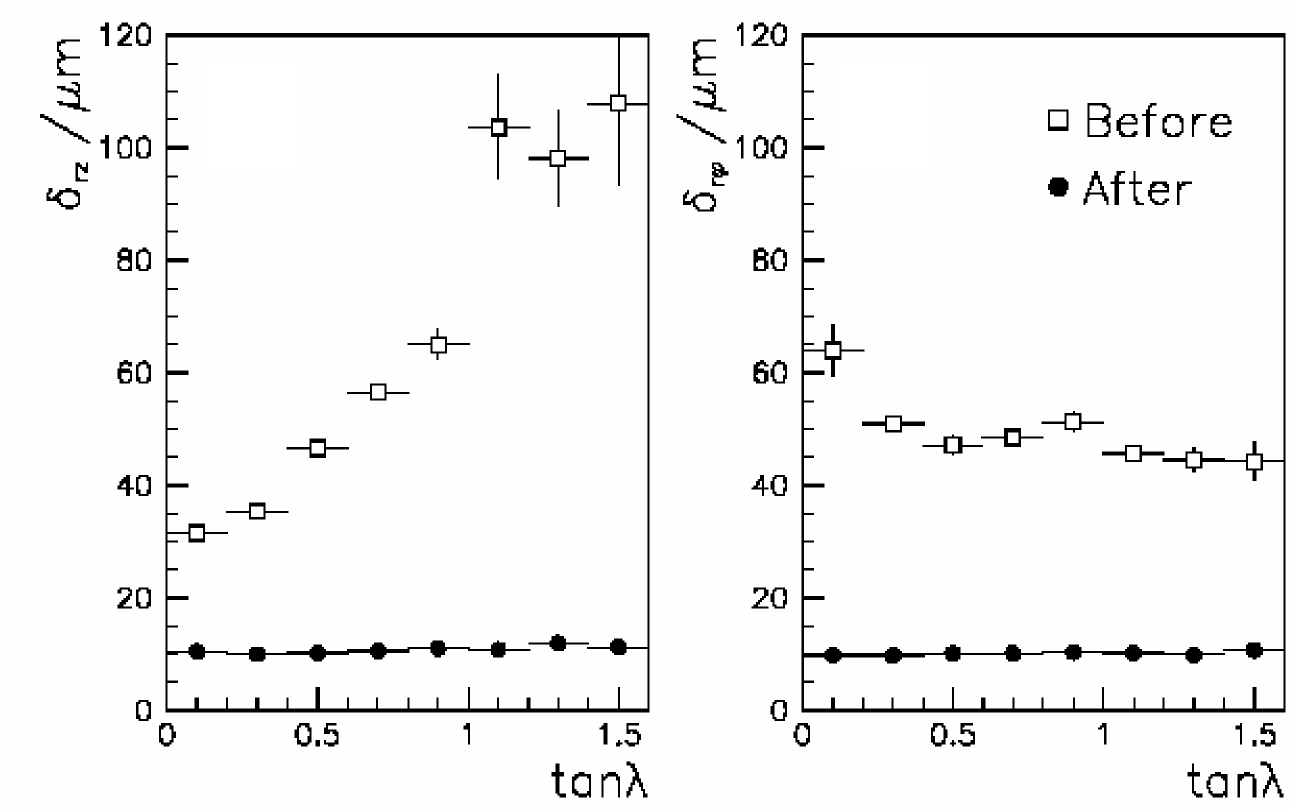
Triplet residual mean as function of φ-dependent index



Systematic effects ≲ 1μm level

Pair Residuals rms at Interaction Point

(divided by √2 to give single track contribution)



Impact Parameter resolution (for full track fit):

$$\sigma_{rz} = 9.7 \oplus \frac{33}{p \sin^{3/2}\theta} \mu\mathrm{m} \qquad \sigma_{r\phi} = 7.8 \oplus \frac{33}{p \sin^{3/2}\theta} \mu\mathrm{m}$$

...design performance achieved

History - VXD2

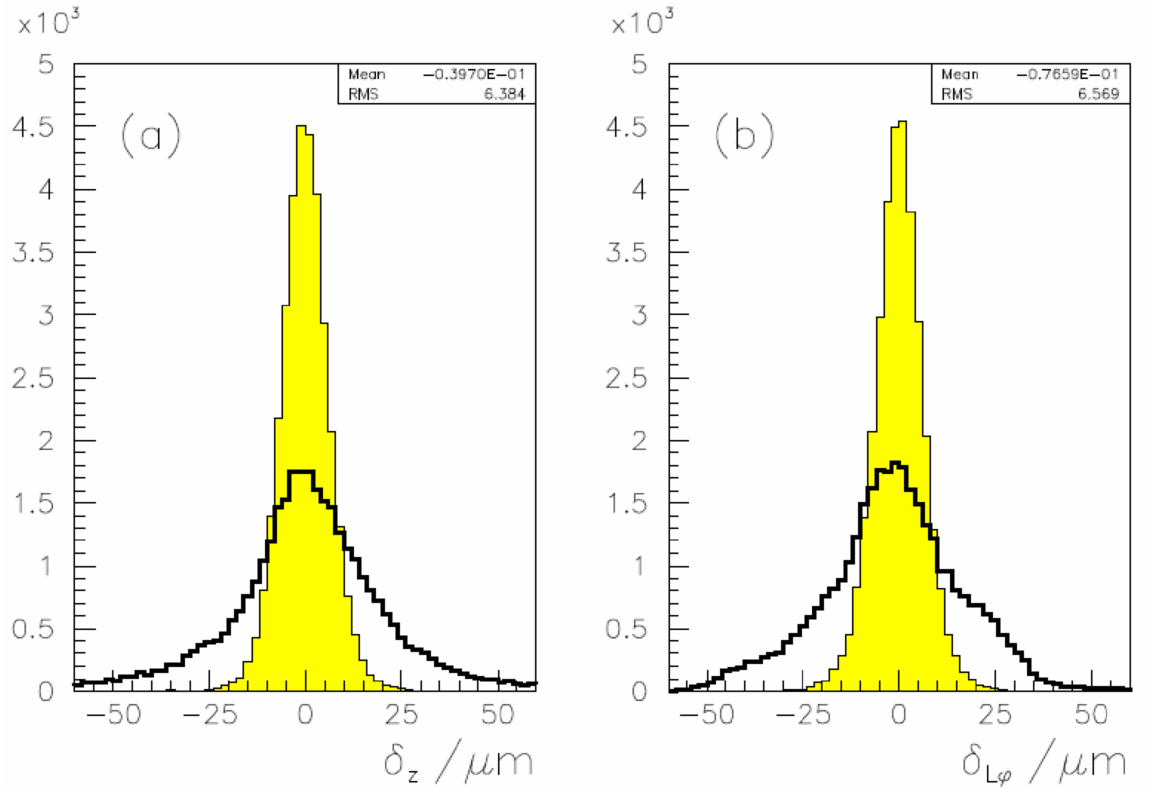
- 480 CCDs on 60 ladders
- Each ladder treated as a rigid body, apart from a bow
  - cold optical survey of each ladder during assembly
  - “hand” corrections for some CCDs based on residual distributions
- Residual fits made to doublets and triplets, similar to those described for VXD3
- Then fitted coefficients to determine 6 degrees of freedom per ladder (□x, □y, □z, pitch, yaw, bow)
  - did not include roll as ladders were narrow

History - VXD2

- Initially two separate matrices
  – Essentially // and perpendicular to length of ladder
  – Coefficients from residual fits plus constraints
  – Used SVD technique to solve (and identify under-constrained parameters)
- However:
  – did not take residual fit errors & correlations into account (I.e. “least sq” c.f. Chi-sq)
  – two dead ladders meant barrel split into two parts
- Later used back-to-back pairs to join the two parts - in separate Minuit fit
- Process was less polished
- But obtained ~5.5 micron hit resolution

History - VXD3

- 1996 did not include IP and limited data to tan□ < 1
  - Knowledge of IP required a well aligned VXD
  - Relatively less data at larger angles and CCD shape uncertainties degrade data at large tan□
  - Including transformation matrix **T** made a major impact
  - Single hit resolutions of ~ 4.9 (rz) and 5.2 (rφ) μm

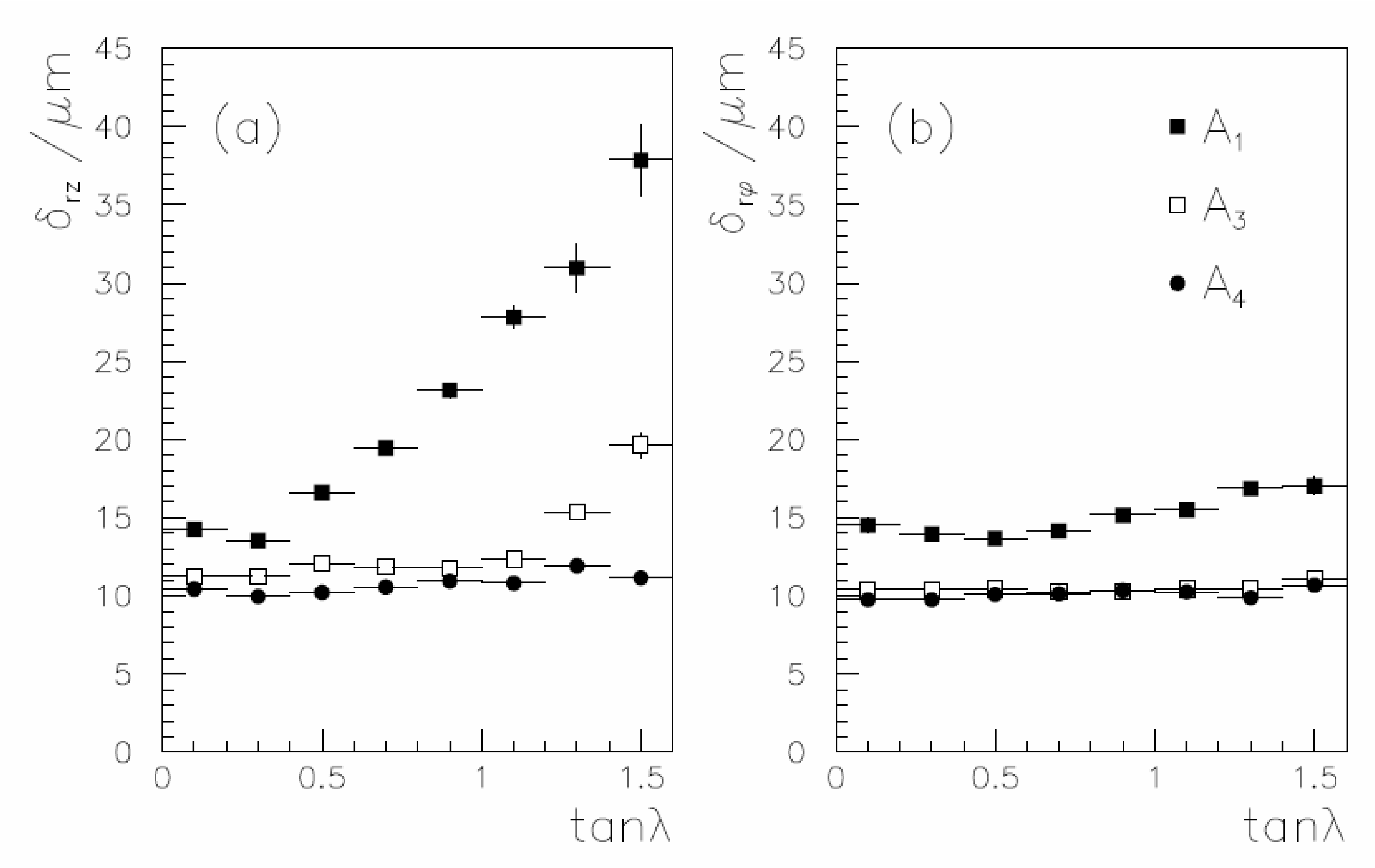


History - VXD3

- 1997 initial fit followed same procedure as 1996
  - Single hit resolutions of ~ 4.9 and 4.6 μm ($r\phi$ improved since 1996 due to removal of some electronic smearing in the detector)
- Used this to find average IP over each 30 events, then refitted geometry including IP
- 1998 added the shape corrections and extended fits to include data over all $\tan\lambda$

History - VXD3

- Single track contributions to pair residuals as a function of track angle show the improving knowledge of the geometry



A_1=1996

A_3=1997

A_4=1997
+1998

History - VXD3

- Procedure was checked
  - With Monte-Carlo studies
  - With 2nd iteration (made negligible difference)
- Examination of residual distributions was important to check for deviations from the assumed functional forms
- To understand details of the problem it was found useful to vary the constraints applied to limit the variations of the corrections (e.g. $\delta q_i = 0.0 \pm 5.0\ \mu m$)
  - But this tuning was not significant to the final results
- Some small effects remained in the final residual distributions (e.g. due to slight bow across some doublets) - but were considered too minor to extend the algorithm further

Comments for other trackers I

- **The technique -** could be used for any system
  - where the required solution takes the form of a perturbation described by O(1000) parameters which are small compared to the dimensions of the system
  - and for which constraining data exists that can be expressed in terms of a set of simultaneous equations for the parameters.
- **Practicalities** - we have demonstrated that it was possible (in 1999) to handle simply and reliably the matrices required for the VXD3 alignment (inversion of sparse matrices of order of 5000 × 1000 elements) using double precision arithmetic in modest times on a standard workstation. Only ~1% or ~35, 000 elements of the final 5026 × 866 design matrix A were given non-zero values

Comments for other trackers II

- **Singular Value Decomposition** – this alignment technique allowed a robust unbiased solution for SLD; but the method is somewhat secondary in that any technique will have similar statistical dependence on the data and geometry.

Alignment is aided by:

- **Symmetry of the detector** – greatly assists book-keeping and allows comparison of different parts of the detector.
- **Overlap regions** – allows devices to be stitched together with favourable lever arm (data α area of overlap).
- **Large devices** – obviously better to have a single element than two with an overlap.

Comments for other trackers III

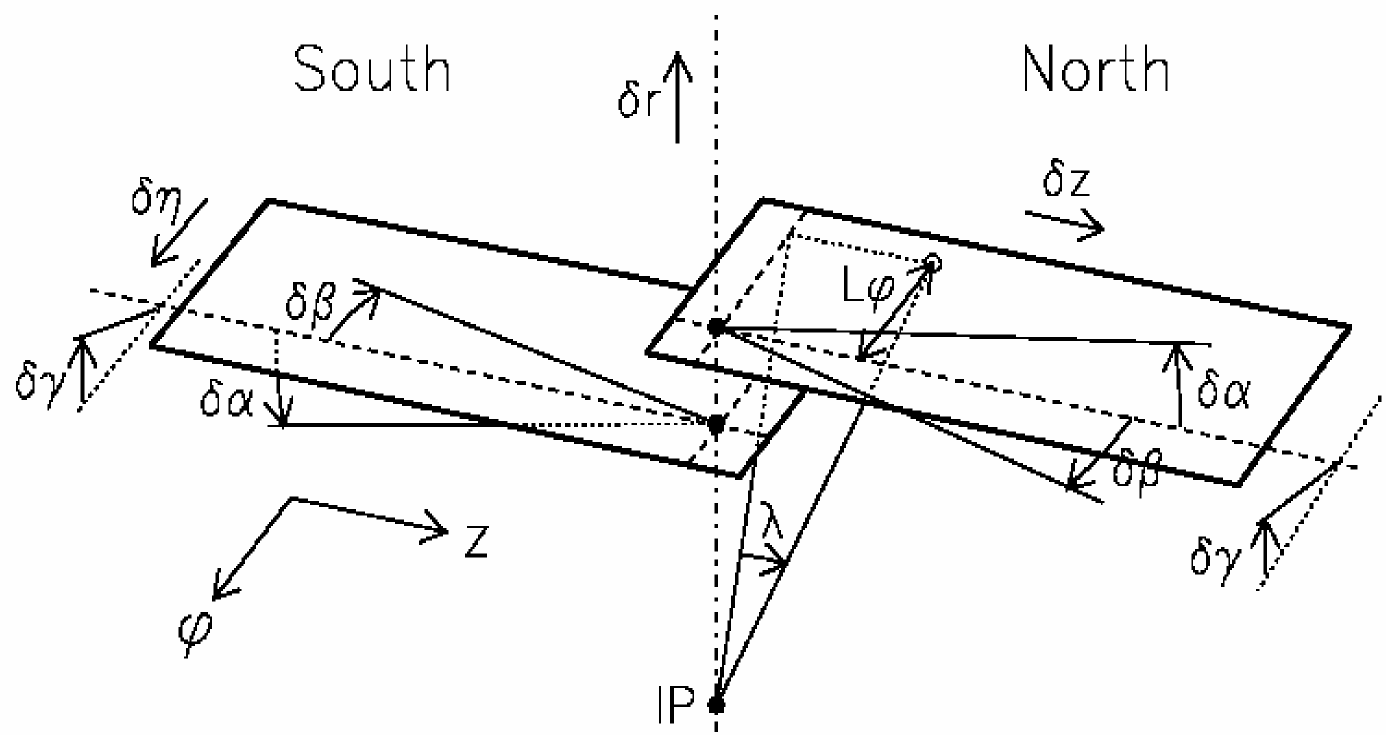
- **Stability** - the geometry (devices and support structure) should be stable with respect to time. Changes due to temperature fluctuations, cycling of magnetic field, ageing under gravity/elastic forces, should be ‘small’; at least over a period of time long enough to collect sufficient track data for alignment.

- **Shape** - within reason the shape of the device is irrelevant; only the uncertainty in the shape is important and the ability to describe the shape correction with as few parameters as possible. Making the devices ‘flat’ is somewhat arbitrary; introducing a deliberate bow of around 1% could greatly increase mechanical stability and decrease shape uncertainty without effecting tracking performance.

VXD3 alignment: D.J.Jackson, D.Su, F.J.Wickens; NIM **A510**, 233 (2003)
Also in an expanded form at http://www-sldnt.slac.stanford.edu/sldbb/SLDNotes/sld-note-271.pdf

Back-up slides

Definition of Parameters

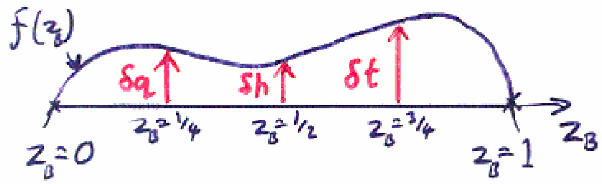


An arbitrary surface shape can be introduced by setting:

$$\delta r \rightarrow \delta r + f(z)$$

4TH ORDER POLYNOMIAL 'FIXED' AT EACH END

(RIGID BODY δr, $\delta\alpha$ CORRECTIONS ALLOW ENDS TO MOVE)



$$\delta q = f(1/4)$$

$$\delta h = f(1/2)$$

$$\delta t = f(3/4)$$

$$f(z_B) = c_1 z_B + c_2 z_B^2 + c_3 z_B^3 + c_4 z_B^4$$

.... A LITTLE ALGEBRA ....

$$c_1 = 16\,\delta q - 12\,\delta h + \frac{16}{3}\,\delta t$$

$$c_2 = -\frac{208}{3}\,\delta q + 76\,\delta h - \frac{112}{3}\,\delta t$$

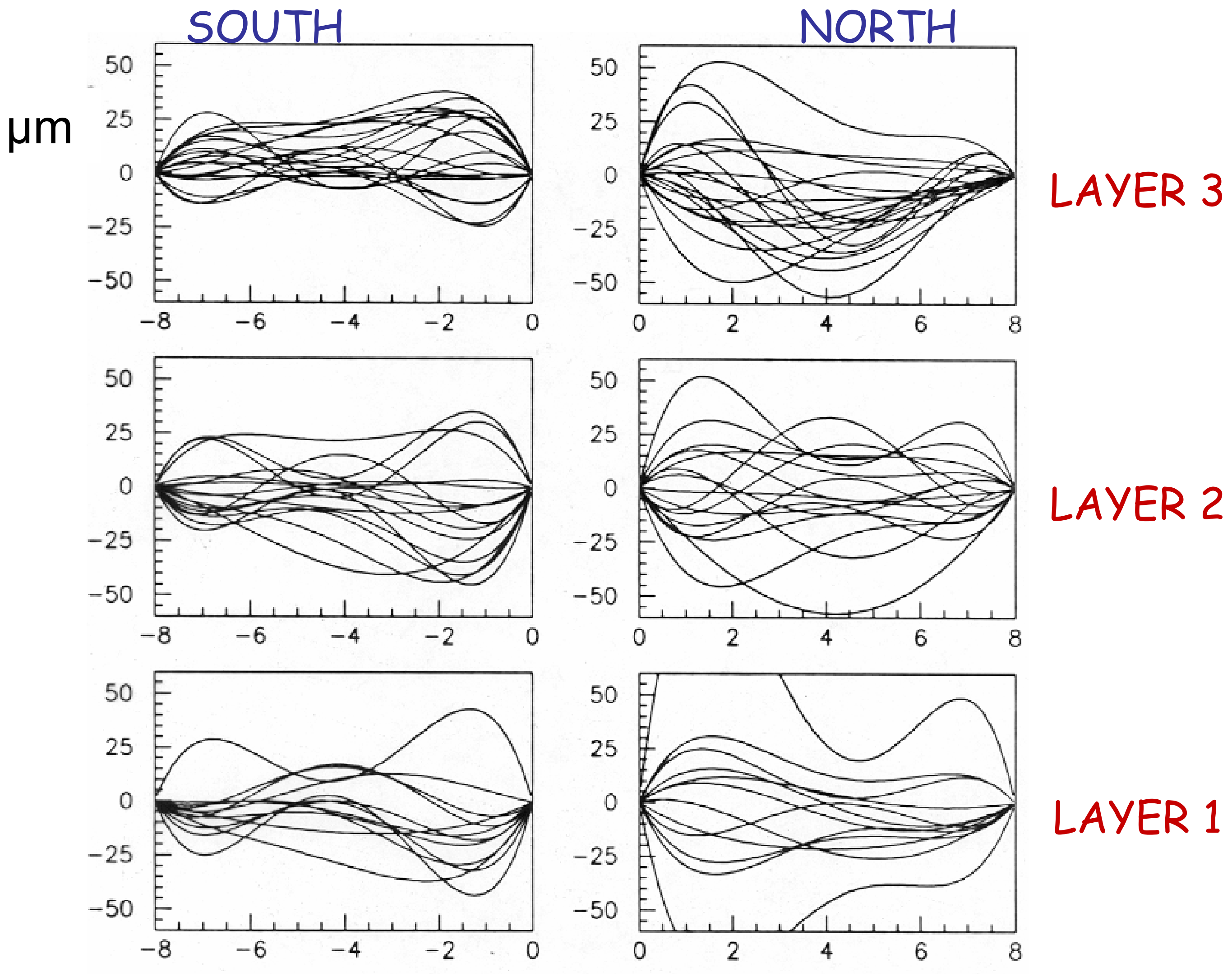
$$c_3 = 96\,\delta q - 128\,\delta h + \frac{224}{3}\,\delta t$$

$$c_4 = -\frac{128}{3}\,\delta q + 64\,\delta h - \frac{128}{3}\,\delta t$$

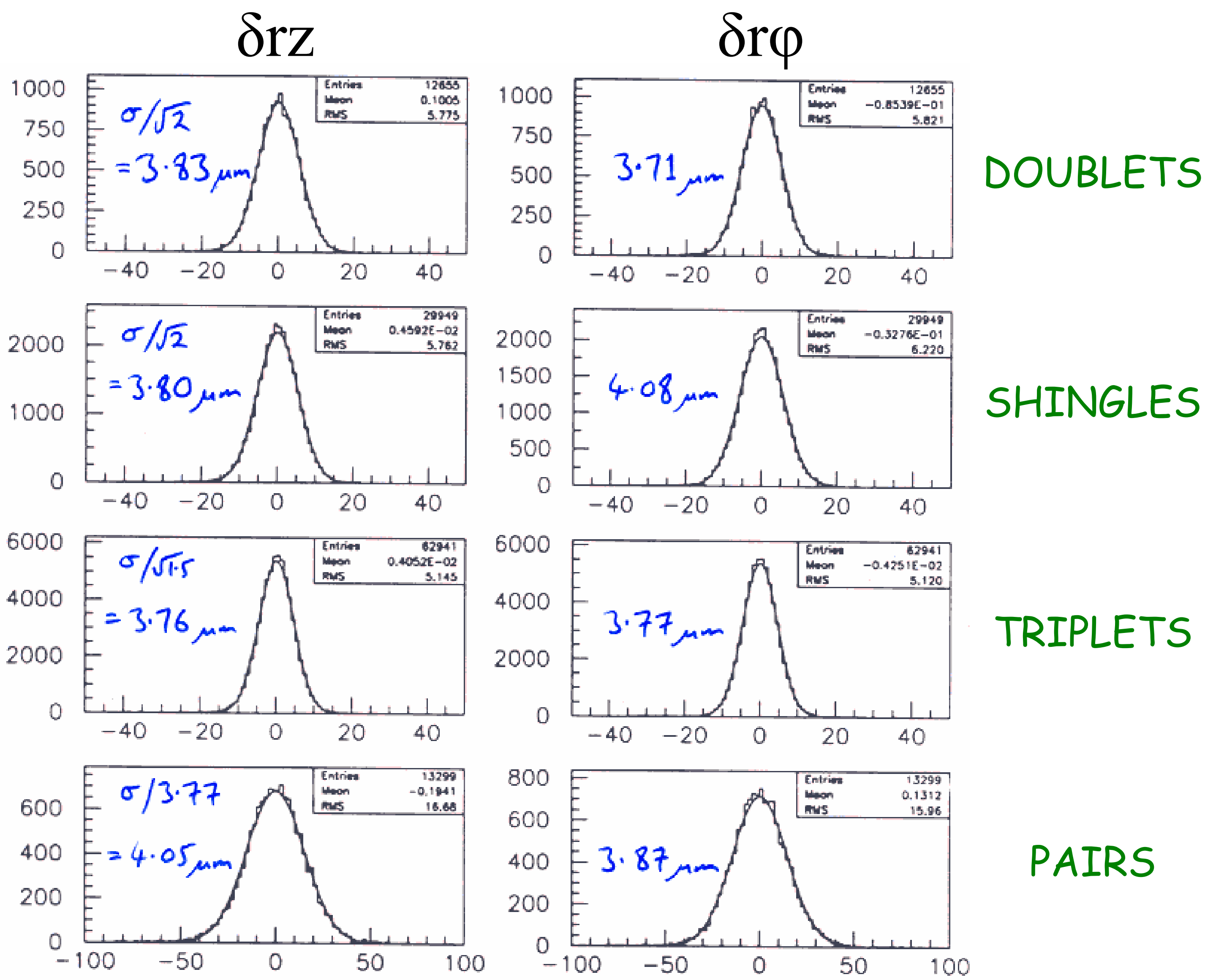
For convenience the base of the CCDs (each 8cm in length) was taken as:

$$z_B = (r \tan\lambda)/8$$

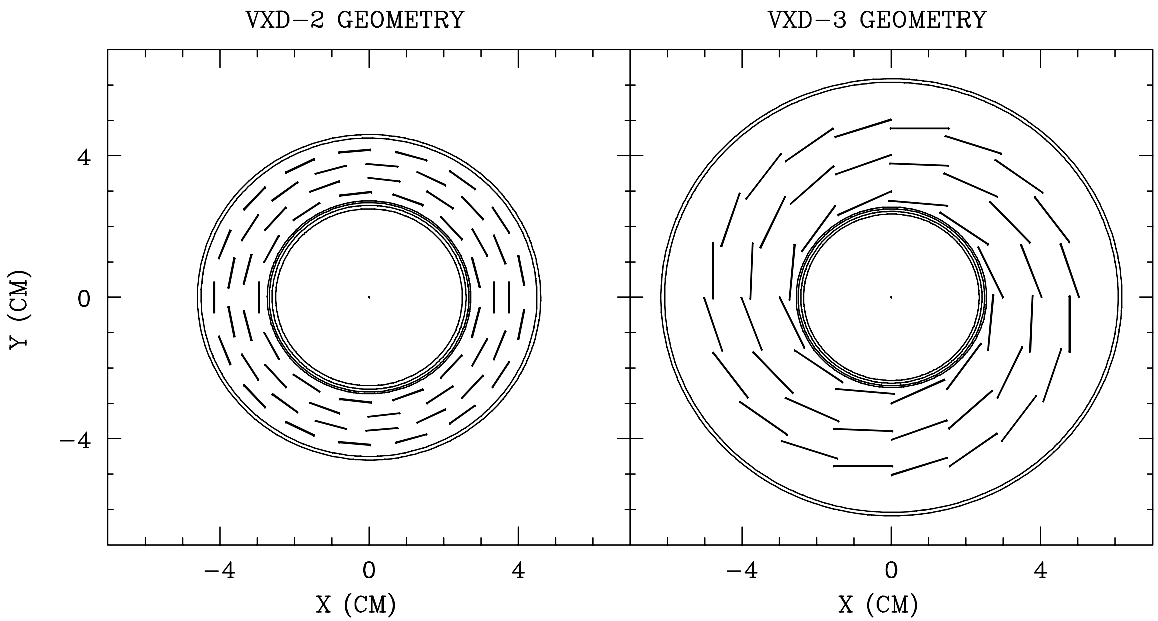
Alignment Shape Corrections



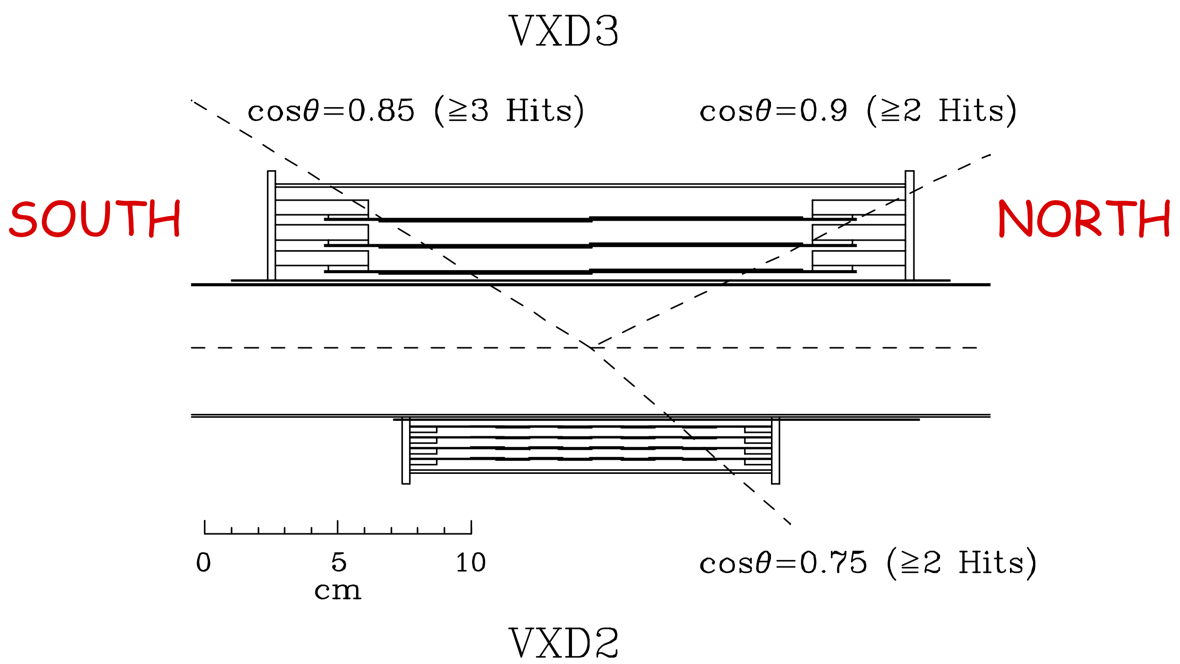
Single hit resolution



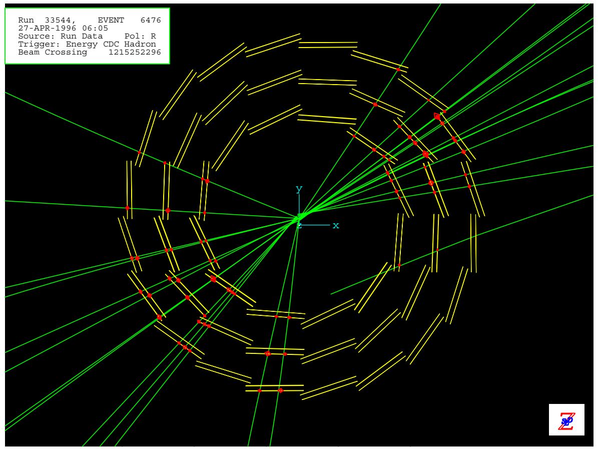
hit resolution consistently ~ 3.8 μm

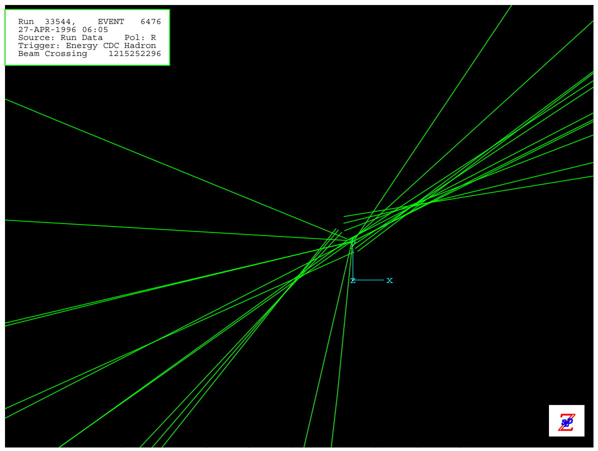


VXD−2 GEOMETRY
VXD−3 GEOMETRY
Y (CM)
X (CM)
X (CM)
4
0
−4



VXD3
cosθ=0.85 (≧3 Hits)
cosθ=0.9 (≧2 Hits)
SOUTH
NORTH
0
5
10
cm
cosθ=0.75 (≧2 Hits)
VXD2





Singular Value Decomposition

$$\underset{m \times n}{\begin{pmatrix} \\ \mathbf{A} \\ \\ \end{pmatrix}} = \underset{\substack{m \times m \\ \text{othogonal}}}{\begin{pmatrix} \\ \mathbf{U} \\ \\ \end{pmatrix}} \underset{\substack{m \times n \\ \text{diag.}}}{\begin{pmatrix} s_1 & & & \\ & s_2 & & \\ & & \ddots & \\ & & & s_r \end{pmatrix}} \underset{\substack{n \times n \\ \text{orthogonal}}}{\begin{pmatrix} \mathbf{V}^T \end{pmatrix}}$$

$s_1 \ldots s_r$ are called the 'singular values' of matrix **A**; $s_i \sim 0$ corresponds to a singularity of **A**

Here's the SVD trick:

define the inverse $\mathbf{A}^+ = \mathbf{V}\mathbf{S}^+\mathbf{U}^T$ with $\mathbf{S}^+ = \begin{pmatrix} 1/s_1 & & & \\ & 1/s_2 & & \\ & & \ddots & \\ & & & 1/s_r \end{pmatrix}$ with $1/s_i = 0$ if $s_i \sim 0$

Then if $\mathbf{A}x = b$ (for vectors x,b)

The solution $x_0 = \mathbf{A}^+ b$ is such that: $|\,\mathbf{A}x_0 - b\,|$ has minimum length

That is, the SVD technique gives the closest 'least squares' solution for an over-constrained (and possibly singular) system