Alignment of the Inner Detector Alignment
Theory and Practice

Artemis School at MPI
16 September 2008

Andrea Bocci
bocci@fnal.gov
On behalf of the ATLAS ID alignment community

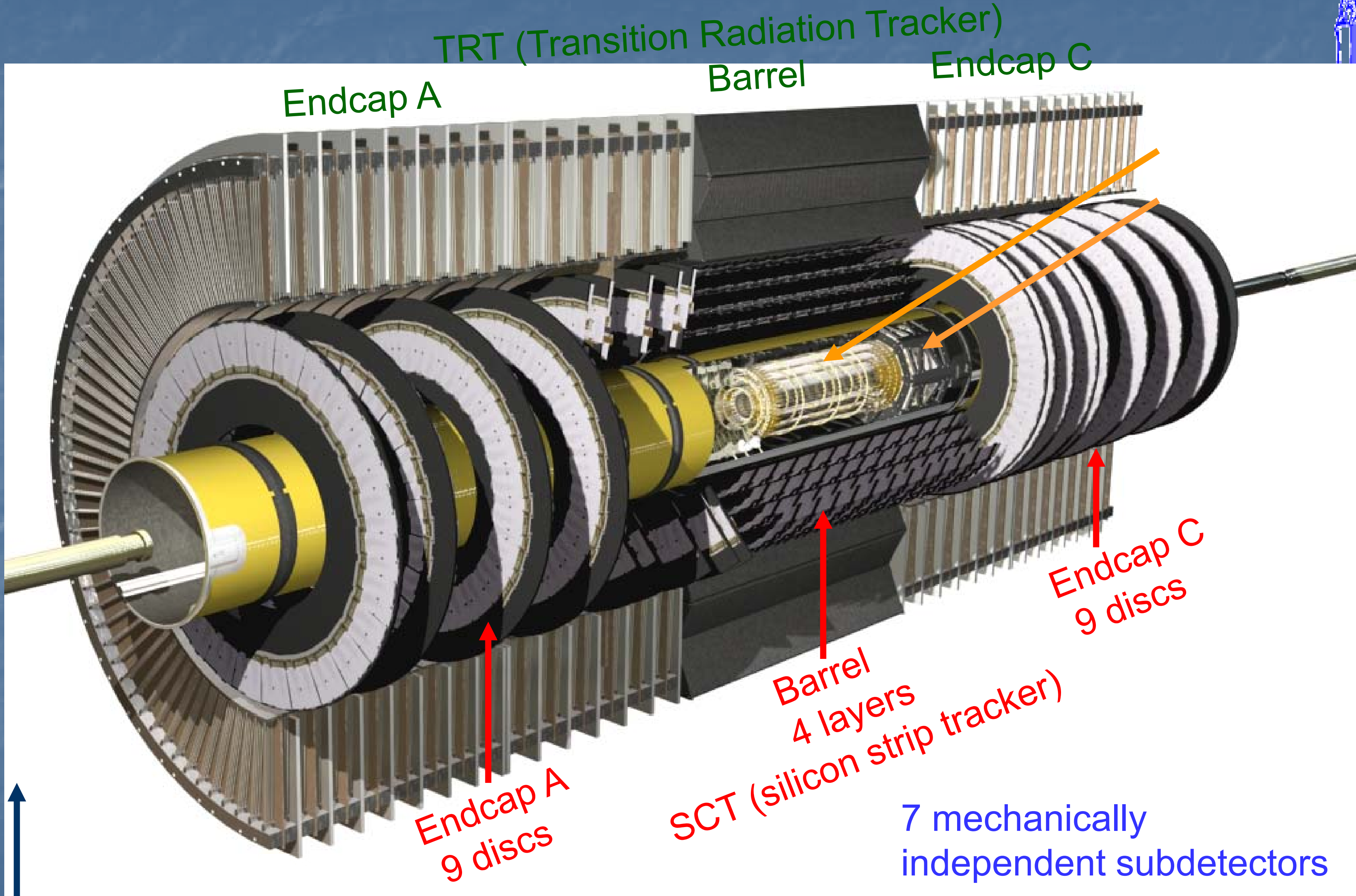
Introduction to the Inner Detector (ID)

The Global Chi2 algorithm:

- Theory
- Results With Data
- Full Scale Alignment
- Weak Modes

Disclaimer: Could not cover a lot of interesting stuff and I have tried to provide a reasonable pedagogical approach

The ATLAS Inner Detector



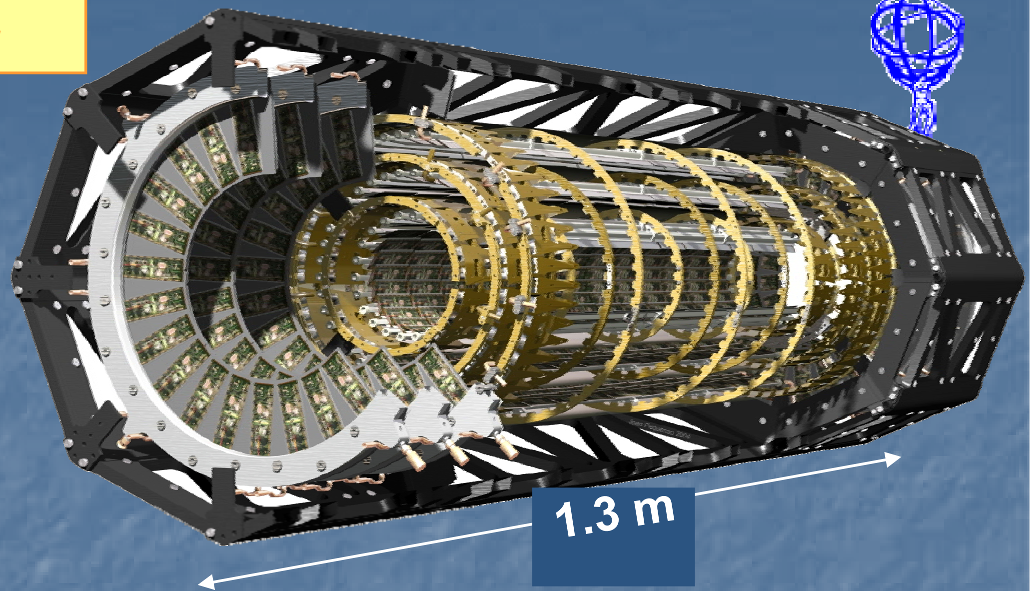
PIXEL

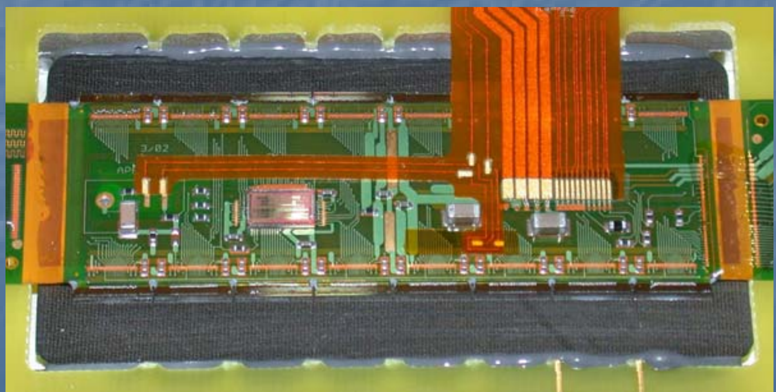
Barrel: 1456 modules
<u>Endcap: 2 x 144 modules</u>
1744 modules

One module: 46.080 pixels
Total: ~80.000.000 pixels

One pixel: 50µm x 400 µm
Resolution: 12µm x 60 µm

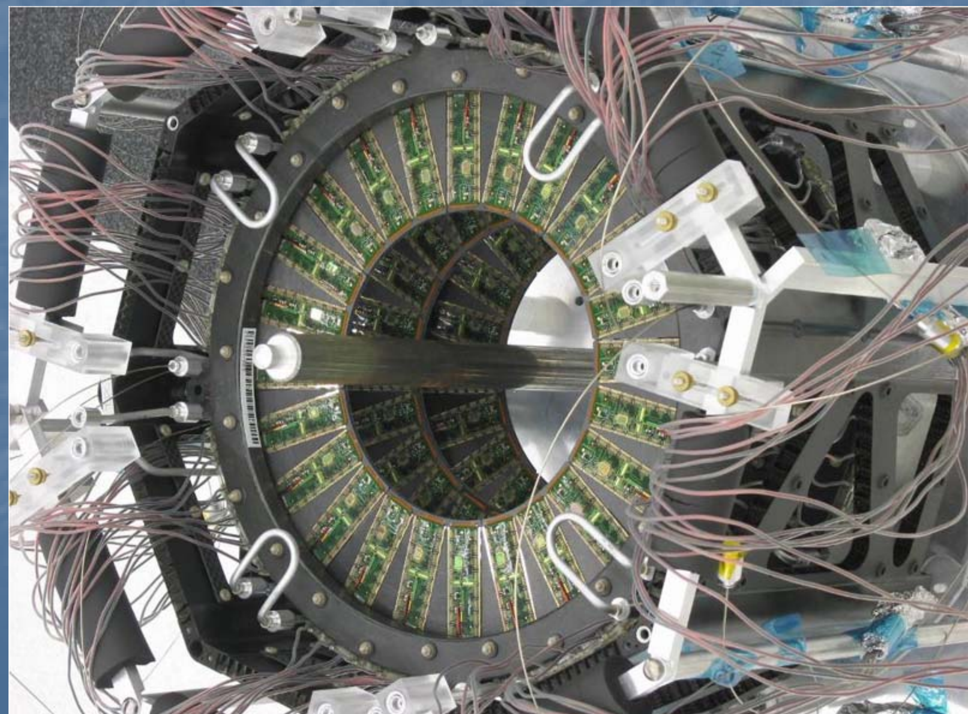
Hits per track: 3





Single Pixel module

All modules have same layout



A view of 3 completed discs of one Endcap

SCT (Silicon Strip)

Barrel: 2112 modules
<u>Endcap: 2 x 988 modules</u>
4088 modules

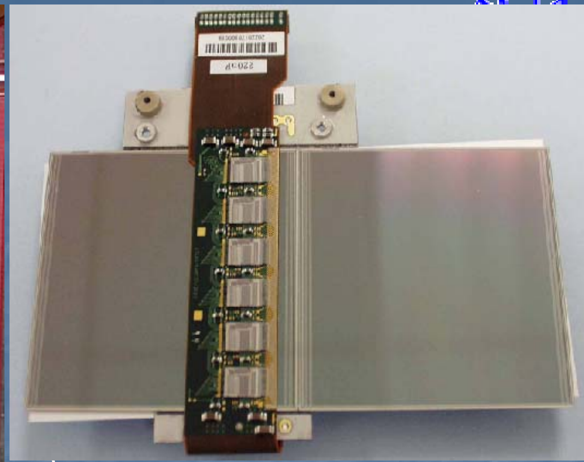
One module: 2 layers x 768 channels
Total: ~ 6.000.000 channels

Channel size: 80µm x 120 mm
Resolution: 16µm x 580 µm

Hits per track: Barrel: 4 Endcap: 9

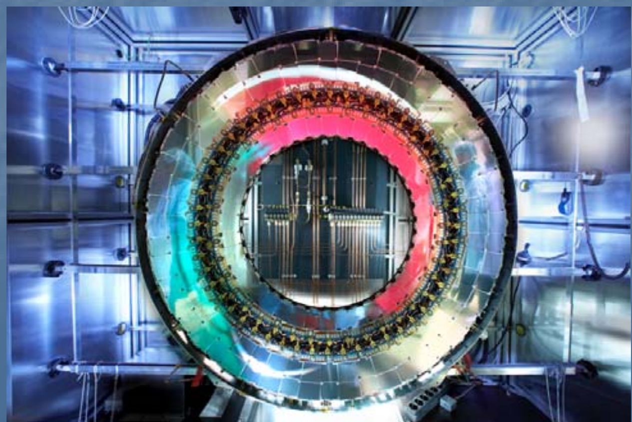
Endcap module

Barrel module



One module has 2 layers with 40 mrad stereo angle

4 different module layouts (3 endcap, 1 barrel)



SCT EndCap

Fully assembled SCT Barrel

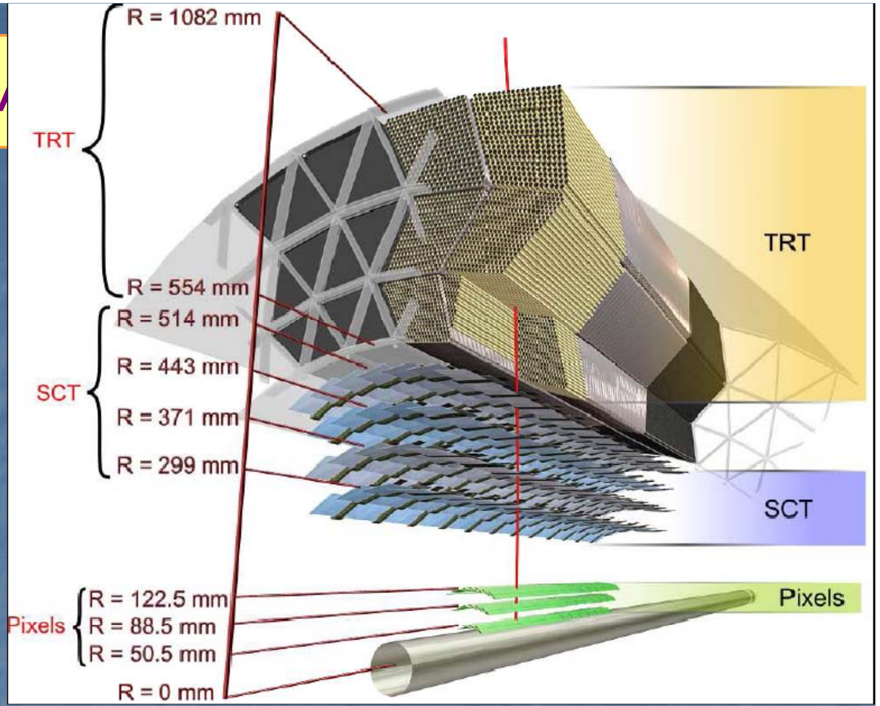
TRT (Transition Radiation using straw

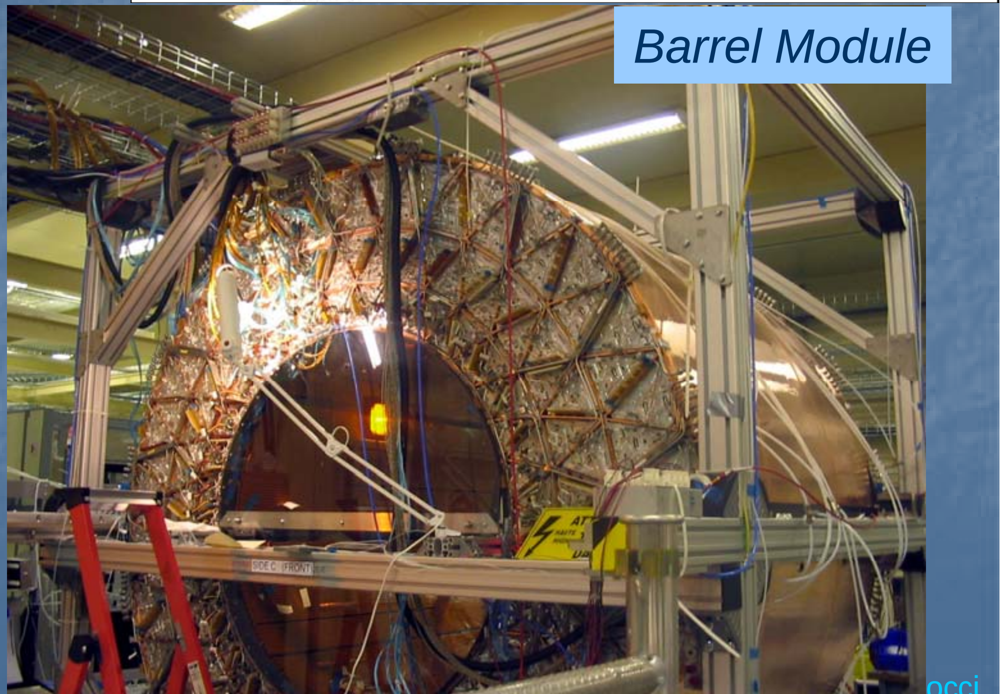
Barrel: 96 modules
Endcap: 28 modules 2

Total: 300.000 straw tubes

Channel size: 4mm x 740 mm
Resolution: 140 µm (perpendicular to wire)

Hits per track: 36





Barrel Module

radiator: poly-propylene
gas mixture: $XeCO_2O_2$ (70+27+3%)

Important:
RT relation has to be determined

The alignment challenge

Alignment is the determination of the position and orientation of the detector components (need to be not far from the intrinsic resolution)

Pixel Detector	Silicon Strip	Transition Radation Tracker
1744 Modules	4088 Modules	96+28 Modules
46K Pixel per Module	768x2 channels/module	300K straw tubes
Dim: 50μm x 400 μm	Dim: 80μm x 120 μm	Dim: 4mm radius
Res: 14 μm x 115 μm	Res: 23 μm (rφ)	Res: 170 μm

Source of knowledge for alignment:

- Assembly knowledge: construction precision and survey, for initial precision corrections and errors
- Online monitoring and alignment: lasers, cameras, before and during run
- **Offline track-based alignment: using physics and track residual information**
- Offline monitoring: using physics, tracks and particle ID parameters

The alignment challenge

Built in/placement precision for the subdetectors :
Translation → several mm
Rotation → several hundred μrad

Built in/placement precision for the modules within subdetectors:
Pixel → 10/100 μm in Rφ/Z
SCT → 50/250 μm in Rφ/Z
TRT → 100 μm (wire placement)

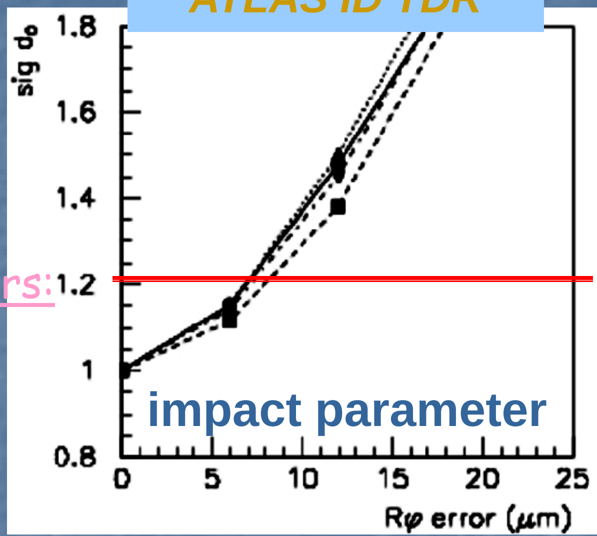
Degradation of the track parameters due to misalignment < 20%

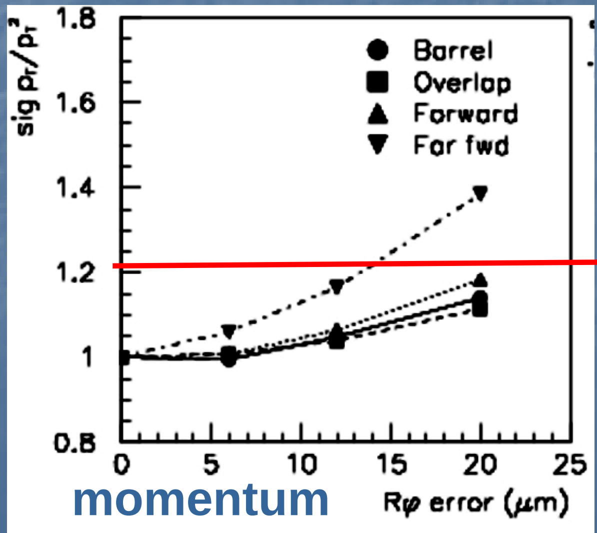
Studies of impact of SCT+Pixel random misalignment on B-Tagging abilities show:

light jet reduction get worse by 10% for $\sigma_{R\Phi}=10\mu m$
light jet reduction get worse by 30% for $\sigma_{R\Phi}=20\mu m$

S. Corréard et al, ATL-COM-PHYS-2003-049

ATLAS ID TDR





The alignment challenge

Built in/placement precision for the subdetectors :
Translation → several mm
Rotation → several hundred μrad

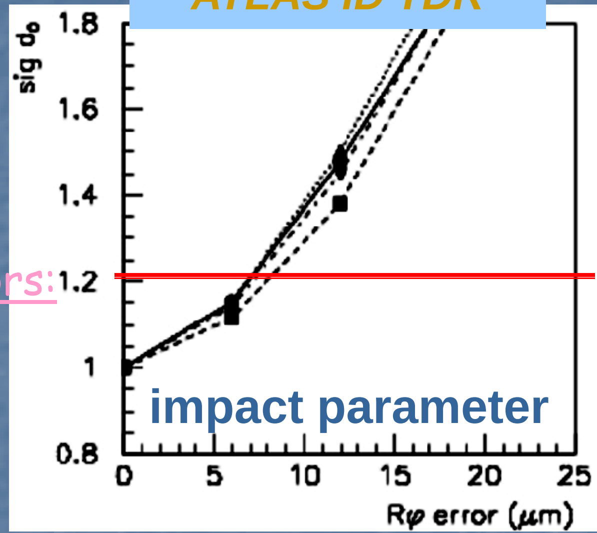
Built in/placement precision for the modules within subdetectors:
Pixel→ 10/100 μm in Rφ/Z
SCT→ 50/250 μm in Rφ/Z
TRT→ 100 μm (wire placement)

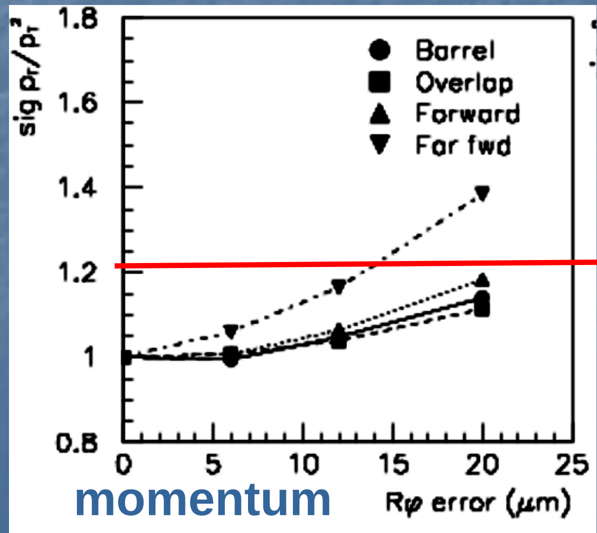
Degradation of the track parameters due to misalignment < 20%

Goal

Pixel → < 7 μm (Rφ)
SCT → < 12 μm (Rφ)
TRT → < 30 μm (Rφ)

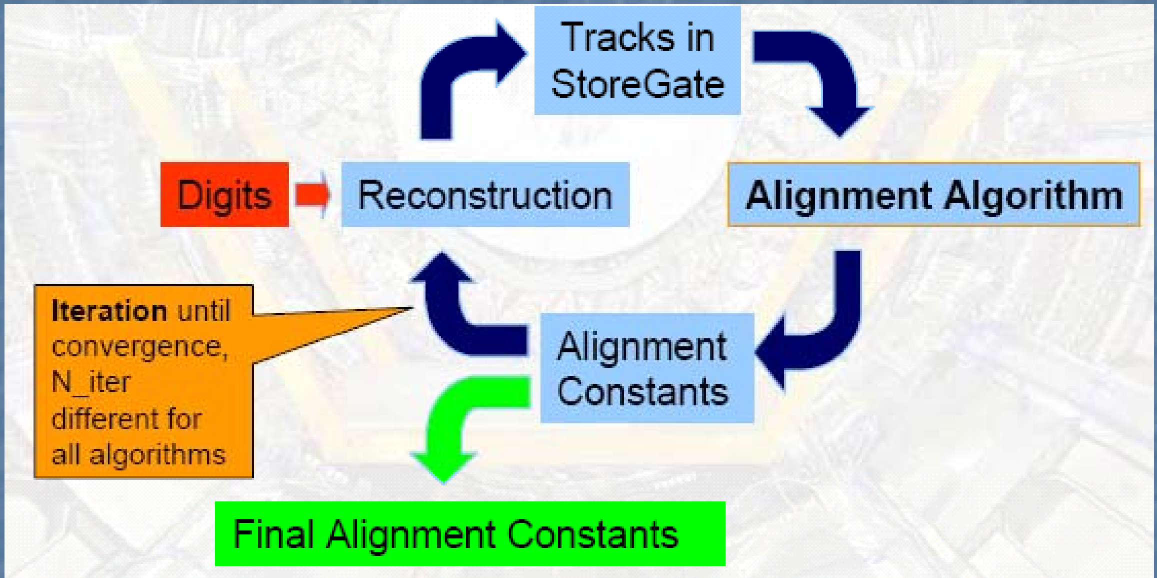
ATLAS ID TDR

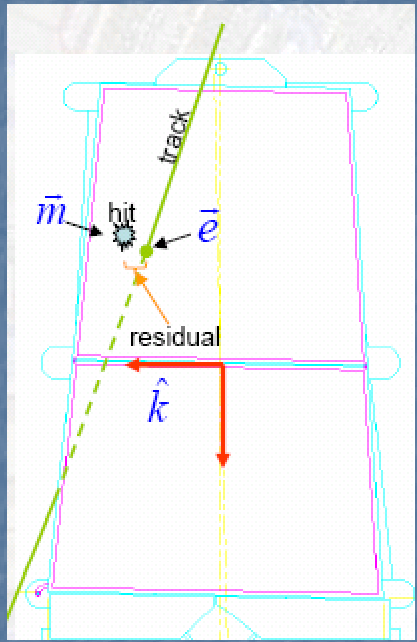




Track Based Alignment

Track based alignment works by minimizing residuals (hit to track distance)





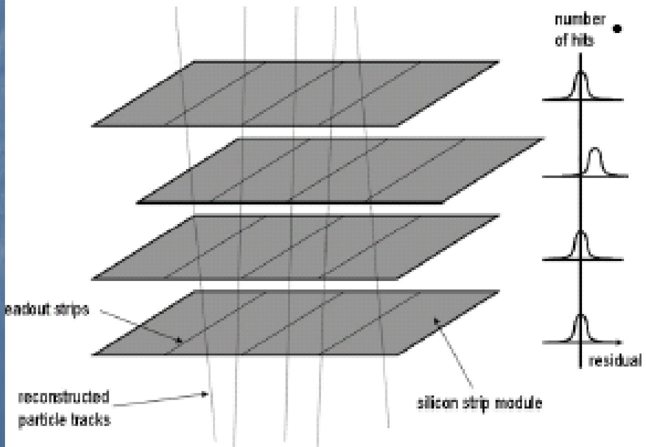
Typically it requires a number of iteration for the detector position to converge

Track Based Alignment

3 algorithms are fully implemented in the ATLAS offline framework

Robust method

- centres residual and overlap-residual distributions
- aligns 2-3 degrees of freedom (DoF)
- iterative procedure



χ^2-based methods $\chi^2 = \sum_{tracks} r_i^T V^{-1} r_i$

Localχ^2

- invert only 6x6 matrices per alignable structure
- iterate to get correlations back

Globalχ^2

solve full linear system of equations with 36k parameters

TRT Algorithm

- global and local χ^2 mode
- also has to calibrate R-t relation and T_0 of the tubes

The Global Chi2 Method

- Several ways to solve the same problem
- I followed the one reported in [1], but equivalent approaches can be find in [2], [3].
- Thanks to Pawel Bruckman and Wouter Hulsbergen for discussions and slides

Mean and Variance

- mean of P

$$\mu_x \equiv \langle x \rangle \equiv \int_{-\infty}^{\infty} x \mathcal{P}(x) \mathrm{d}x$$

- variance

$$\sigma_x^2 \equiv \mathrm{var}(x) \equiv \left\langle (x - \langle x \rangle)^2 \right\rangle = \left\langle x^2 \right\rangle - \langle x \rangle^2$$

- example, the gaussian distribution

$$\mathcal{P}(x)\, \mathrm{d}x = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] \mathrm{d}x$$

$$\langle x \rangle = \mu \qquad \mathrm{var}(x) = \sigma^2$$

Multi-Dimensional Case

- covariance
$$V_{xy} \equiv \mathrm{cov}(x, y) \equiv \langle (x - \langle x \rangle)(y - \langle y \rangle) \rangle$$
- correlation coefficient $\rho_{x,y} \equiv \dfrac{\mathrm{cov}(x,y)}{\sqrt{\mathrm{var}(x)\,\mathrm{var}(y)}}$
- note: $\mathrm{cov}(x, y) = \mathrm{cov}(y, x)$
$$\mathrm{var}(x) = \mathrm{cov}(x, x)$$
$$-1 \leq \rho_{x,y} \leq 1$$

Covariance Matrix

- covariance conveniently organized in matrix

$$V(x, y, z, \cdots) = \begin{pmatrix} V_{xx} & V_{xy} & V_{xz} & \cdots \\ V_{yx} & V_{yy} & V_{yz} & \cdots \\ V_{zx} & V_{zy} & V_{zz} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

- matrix V is symmetric and positive-definite (det(V)>=0)
- example: gaussian (normal) distribution in N dimensions

$$\mathcal{P}(x_1, \ldots, x_N)\, \mathrm{d}x_1 \cdots \mathrm{d}x_N \propto \exp\left[\frac{1}{2}\, x^T\, V^{-1}\, x\right]\, \mathrm{d}x_1 \cdots \mathrm{d}x_N$$

- where $x = (x_1, \cdots, x_N)$ and V as above

Usually diagonal

Linear Transformations

- if ***F*** a linear transformation such that

$$y = F\,x \qquad \text{for vectors } x \in R^n, y \in R^m \text{ and matrix } F \in R^m \times R^n$$

then

$$\langle y \rangle = F \; \langle x \rangle \qquad \boxed{\mathrm{var}\,(y) = F\;\mathrm{var}\,(x)\;F^T}$$

- this is the familiar 'error propagation'
- if the transformation is not linear, e.g. $y = f(x)$

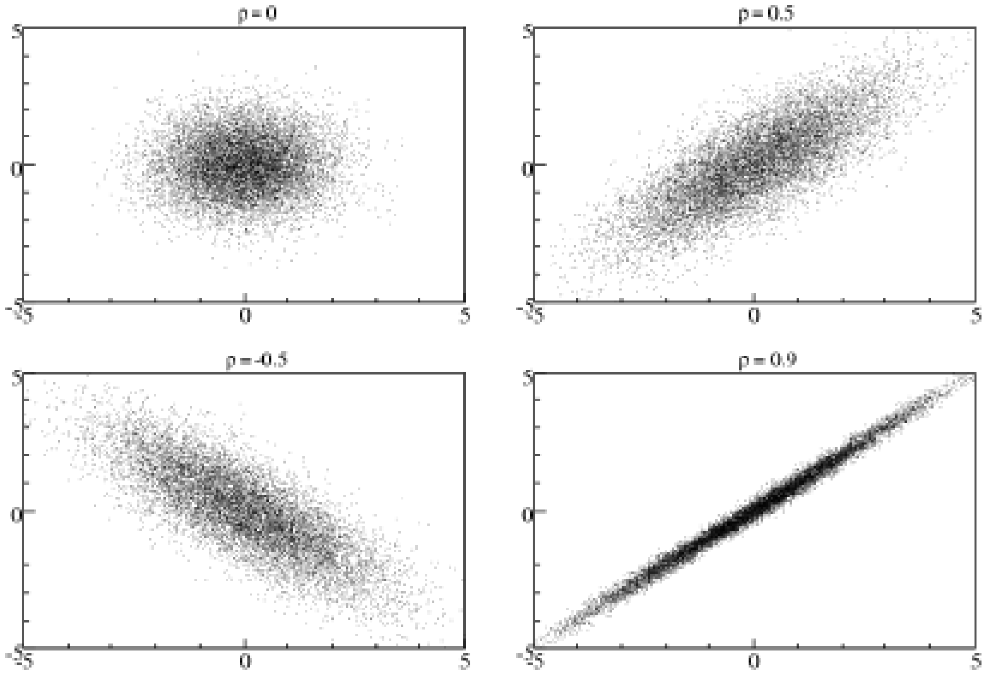
  the expressions above hold ***to first order*** *in* ***x*** with jacobian

$$F_{ij} = \frac{\partial y_i}{\partial x_j}$$

- this is just an approximation: if you want the true variance of y, you need to calculate ***var(f(x))***

Intermezzo

- example: x and y gaussian distributed with unit variance



- correlation tells about the sign of the *direction* of the slope and how *squeezed* the distribution is
- sizes of half the major and minor axis of the 'ellipse' correspond to eigenvalues of covariance matrix V

Estimators

- suppose we have
  - a data set $\{x_i\}$
  - a model with unknown parameters α
- a *statistic* is any function of the data that does not depend on α
- an *estimator* for α is a statistic whose value estimates α
- some important properties of estimators
  - **consistency**: estimator is consistent if it approaches true value with more data
  - **bias**: difference between expectation value of estimate and α
  - **efficiency**: ratio between variance of estimate and best possible variance of any estimate for α

Least Square Method

- consider N independent measurements with Gaussian PDF

measurement *i*

measurement model

$$\mathcal{P}_i\,(m_i\,;\,x) = \frac{1}{\sqrt{2\pi}} \exp\left[\frac{1}{2}\left(\frac{m_i - h_i(x)}{\sigma_i}\right)^2\right]$$

model parameters

uncertainty in measurement *i*

Least Square Method - II

- consider N independent measurements with Gaussian PDF

$$\mathcal{P}_i\left(m_i\,;\,x\right) = \frac{1}{\sqrt{2\pi}}\exp\left[\frac{1}{2}\left(\frac{m_i - h_i(x)}{\sigma_i}\right)^2\right]$$

- define the **chi-square**

$$\chi^2 \equiv \sum_i \left(\frac{m_i - h_i(x)}{\sigma_i}\right)^2$$

- the value x-hat for which the chi-square is minimum is called the **least squares estimator (LSE)**

Chi2 in Matrix Notation

- rewrite chi-square using covariance matrix for measurements

vector of measurements

measurement model

(diagonal) measurement covariance matrix

$$\chi^2 = \sum_i \left(\frac{m_i - h_i(x)}{\sigma_i} \right)^2 = \underbrace{(m - h(x))^T}_{\text{vector of 'residuals'}} V^{-1} \, (m - h(x))$$

- condition that chi-square is minimum, can now be written as

$$0 = \frac{\mathrm{d}\chi^2}{\mathrm{d}x} = -2 \left(\frac{\mathrm{d}h(x)}{\mathrm{d}x} \right)^T V^{-1} \, (m - h(x))$$

derivative matrix

X vector of M parameters

- for N measurements and M parameters, derivative is NxM matrix

Linear Case

- in many fit applications derivative of h(x) varies slowly with respect to measurement errors
- therefore, consider linear measurement model

$$h(x) = h_0 + H\,x$$

where the derivative matrix $H \equiv \frac{\mathrm{d}h(x)}{\mathrm{d}x}$ is constant

- condition that chi-square derivative vanishes, becomes

$$\frac{\mathrm{d}\chi^2}{\mathrm{d}x} = -2\,H^T V^{-1}\,(m - h_0 - Hx) = 0$$

which has a solution

$$\hat{x} = \left(H^T V^{-1} H\right)^{-1}\,H^T V^{-1}\,(m - h_0)$$

Bias and Variance of x-hat

- provided that the measurements are unbiased and have variance V

$$\langle m \rangle = m^{\text{true}} \equiv h_0 + H x^{\text{true}} \qquad \text{var}(m) \equiv V$$

- we find that the bias of the LSE is zero

$$\begin{aligned}\langle \hat{x} - x^{\text{true}} \rangle &= \left(H^T V^{-1} H\right)^{-1} H^T V^{-1} (\langle m \rangle - h_0 - H x^{\text{true}}) \\ &= 0\end{aligned}$$

- and that its variance is

$$\text{var}(\hat{x}) = \text{var}\left(\left(H^T V^{-1} H\right)^{-1} H^T V^{-1} (m - h_0)\right)$$

drop constants

$$= \text{var}\left(\left(H^T V^{-1} H\right)^{-1} H^T V^{-1} m\right)$$

var(Ax) = A var(x) A^T

$$= \left(H^T V^{-1} H\right)^{-1} H^T V^{-1} \,\text{var}(m)\, V^{-1} H \left(H^T V^{-1} H\right)^{-1}$$

var(m)=V

$$= \left(H^T V^{-1} H\right)^{-1}$$

Summary

- least squares ('minimum chi-square') estimator

$$\hat{x} = C H^T V^{-1} (m - h_0)$$

$$C \equiv \mathrm{var}(\hat{x}) = \left(H^T V^{-1} H\right)^{-1}$$

- simplest example: weighted mean
  - consider measurements m_i with known uncertainty σ_i
  - assuming they measure the same thing 'x', what value has 'x'?

$$h_i(x) = x \quad \Longrightarrow \quad H^T = (1, 1, 1, \cdots)$$

$$\hat{x} = \left(\sum \frac{1}{\sigma_i^2}\right)^{-1} \sum \frac{m_i}{\sigma_i^2} \qquad \mathrm{var}(\hat{x}) = \left(\sum \frac{1}{\sigma_i^2}\right)^{-1}$$

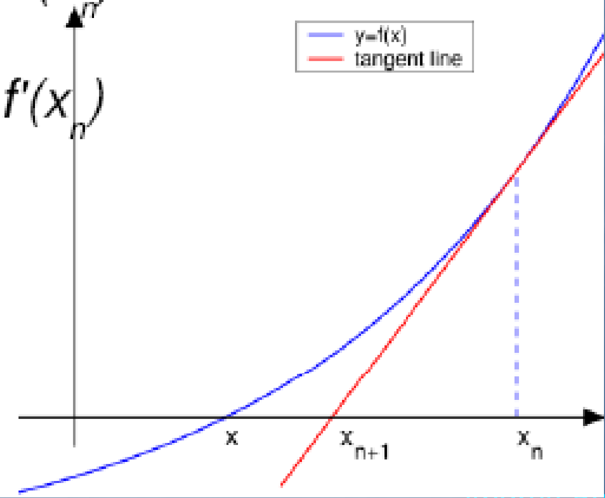
Non-linear Case

- what if the measurement model *h(x)* is not linear?
- first derivative of the chi-square now looks like

$$\frac{\mathrm{d}\chi^2}{\mathrm{d}x} = 2\,\frac{\mathrm{d}h(x)}{\mathrm{d}x}^T V^{-1}\,(h(x) - m)$$

  where the derivative *dh/dx* now depends on x as well

- use Newton-Raphson to find the zero-crossing '*x*' of a function *f(x)*
  - starting from an estimate x_n, evaluate $f(x_n)$ and $f'(x_n)$
  - estimate a better value as $x_{n+1} = x_n - f(x_n) / f'(x_n)$
  - iterate until you're happy with the value of *f(x)*



Linearization

- second derivative becomes

$$\frac{\mathrm{d}^2\chi^2}{\mathrm{d}x^2} = \underbrace{2\,\frac{\mathrm{d}h(x)}{\mathrm{d}x}^T V^{-1}\,\frac{\mathrm{d}h(x)}{\mathrm{d}x}}_{\text{this term also appears for a linear model}} + \underbrace{2\,\frac{\mathrm{d}^2h(x)}{\mathrm{d}x^2}^T V^{-1}\,(h(x)-m)}_{\text{this term is new}}$$

- the second term appears because the derivative is not constant

Since we don't know how to handle higher orders, we just drop the second term

Now the problem is linear again !

(usually second term much smaller that the first)

Summary of non-linear case

- summarizing: choosing a starting point x_0, we have

$$\frac{1}{2}\left.\frac{\mathrm{d}\chi^2}{\mathrm{d}x}\right|_{x_0} = -H^T\,V^{-1}\,(m - h(x_0))$$

$$\frac{1}{2}\left.\frac{\mathrm{d}^2\chi^2}{\mathrm{d}x^2}\right|_{x_0} = H^T\,V^{-1}\,H$$

$$H \equiv \left.\frac{\mathrm{d}h(x)}{\mathrm{d}x}\right|_{x_0}$$

- which, with Newton-Raphson, gives

$$\hat{x} = x_0 - \left(\frac{\mathrm{d}^2\chi^2}{\mathrm{d}x^2}\right)^{-1}\frac{\mathrm{d}\chi^2}{\mathrm{d}x}$$

expression is just the same as for linear model

- note that the variance can be written as

$$\mathrm{var}\,(x) = 2\left(\frac{\mathrm{d}^2\chi^2}{\mathrm{d}x^2}\right)^{-1}$$

- we now need a *sensible starting point* and *iterations and* repeat the calculation of derivatives and x, until we are 'close enough'

The minimum chisquare fit

- define a track chisquare as

$$\chi^2 = \sum_{\text{hits i}} \left(\frac{m_i - h_i(x)}{\sigma_i} \right)^2$$

where

- m → measurement, σ → measurement error
- x → track parameters, usually 5
- h → measurement model

- we can also write this in a matrix notation

$$\chi^2 = r^T V^{-1} r$$

- r = m - h(x) → residual vector
- V → measurement covariance matrix (usually diagonal)

- how do we 'minimize' this chisquare?

Chisquare minimization for alignment

- suppose now, that we have a set of alignment parameters 'alpha'
- we would like to minimize a total chisquare

$$\chi^2 = \sum_{\text{tracks j}} \left(r^T V^{-1} r \right)_j$$

with respect to alpha and with repect to the track parameters

- two solutions:
  1. minimize for x and alpha simultaneously on large sample of tracks
     - unpractical, because too many parameters
  2. minimize every track to x first, then alpha on a large sample of tracks
     - keep track of dependence of x on alpha through total derivative

$$\frac{\mathrm{d}}{\mathrm{d}\alpha} = \frac{\partial}{\partial\alpha} + \frac{\partial x}{\partial\alpha}\frac{\partial}{\partial x}$$

- how do we calculate dx/dα?

How tracks change as function of alpha

- we can calculate the derivatives of x to alpha by requiring that the derivative of the track chisquare remains unchanged

$$0 = \frac{\mathrm{d}}{\mathrm{d}\alpha}\frac{\partial \chi^2}{\partial x} = \frac{\partial^2 \chi^2}{\partial \alpha \partial x} + \frac{\mathrm{d}x}{\mathrm{d}\alpha}\frac{\partial^2 \chi^2}{\partial x \partial x}$$

$$\Rightarrow \quad \frac{\mathrm{d}x}{\mathrm{d}\alpha} = -\frac{\partial^2 \chi^2}{\partial \alpha \partial x}\left(\frac{\partial^2 \chi^2}{\partial x \partial x}\right)^{-1}$$

- this yields "derivatives of chisquare wrt alpha,
  under condition that it remains minimal wrt to x"

$$\frac{\mathrm{d}\chi^2}{\mathrm{d}\alpha} = 2\sum_{\text{tracks}} \frac{\partial r}{\partial \alpha}^T V^{-1}\left(V - HCH^T\right)V^{-1} r$$

C = Cov(x)

$$\frac{\mathrm{d}^2\chi^2}{\mathrm{d}\alpha^2} = 2\sum_{\text{tracks}} \frac{\partial r}{\partial \alpha}^T V^{-1}\left(V - HCH^T\right)V^{-1} \frac{\partial r}{\partial \alpha}$$

- these are the derivatives for the 'global chisquare method' for alignment

An important observation: first derivative is 'local'

- looking at the first derivative, it seems as if each derivative on track depends on each hit on track

$$\frac{\mathrm{d}\chi^2}{\mathrm{d}\alpha} = 2 \sum_{\text{tracks}} \frac{\partial r}{\partial \alpha}^T V^{-1} \left(V - HCH^T \right) V^{-1} r$$

this matrix correlates derivatives for module 'i' with hits in module 'j'

- however, remember that the track chisquare is at its minimum

$$H^T V^{-1} r = 0 \quad \Rightarrow \quad \frac{\mathrm{d}\chi^2}{\mathrm{d}\alpha} = 2 \sum_{\text{tracks}} \frac{\partial r}{\partial \alpha}^T V^{-1} r$$

- so, the first derivative is local: the derivative for module 'i' only depends on residuals in module 'i'

- the 'global' method only distinguishes from a 'local' method by the off-diagonal terms in the 2nd derivative

Alignment Parameter Determination

As we have seen, we just need to linearize the ♍2 condition with respect to the alignment parameters. Starting from a initial value ♋$_0$ we have a linear system of M equations:

$$\frac{d^2\chi^2}{d\alpha^2}\Delta\alpha = -\frac{d\chi^2}{d\alpha}$$

$$\mathrm{Cov}(\alpha) = 2\left(\frac{d^2\chi^2}{d\alpha^2}\right)^{-1}$$

Iterations are needed if the residual are not linear w.r.t. ♋ and *x*

This is the ***global chi2 method*** for the alignment

Algorithms like MILLIPEDE [3] (used by CMS) use this kind of procedure

The ***local chi2 method*** solves the same system but ignoring the correlations among modules in the second derivatives.

- It is equivalent to decompose the matrix in small blocks.
- It needs however iterations even in a pure linear condition.

Alignment in ATLAS

For Silicon there are two packages where respectively the global and the local methods are implemented.

The TRT has a different (but equivalent) implementation of the global method.
In addition, the same package can align also the SCT and Pixel.
A simple flag can be set to ignore correlations and run in local mode

The muon system started to recycle these codes to have a global chi2 alignment as well.
For the future we foreseen to merge all these code and have one ALTAS implementation of the global chi2 method

More "theory" I could not cover:

- Robust Alignment: radically different approach, no χ^2 minimization involved !!
- Advanced course: how to implement external constraints (vertex, mass, etc..), treatment of multiple scattering, use of Kalman filter for alignment, biased vs unbiased residuals, etc...

ID Alignment Implementation in ATLAS

Alignment Implementation - I

ID alignment can be performed at different granularity (a.k.a Levels)

Big Structures:

- Highest impact on physics (and on trigger too!)
- Require less statistics
- Generally easy to perform
- Easy to monitor
- Expected to be fairly stable
- Small misalignment can lead to big bias

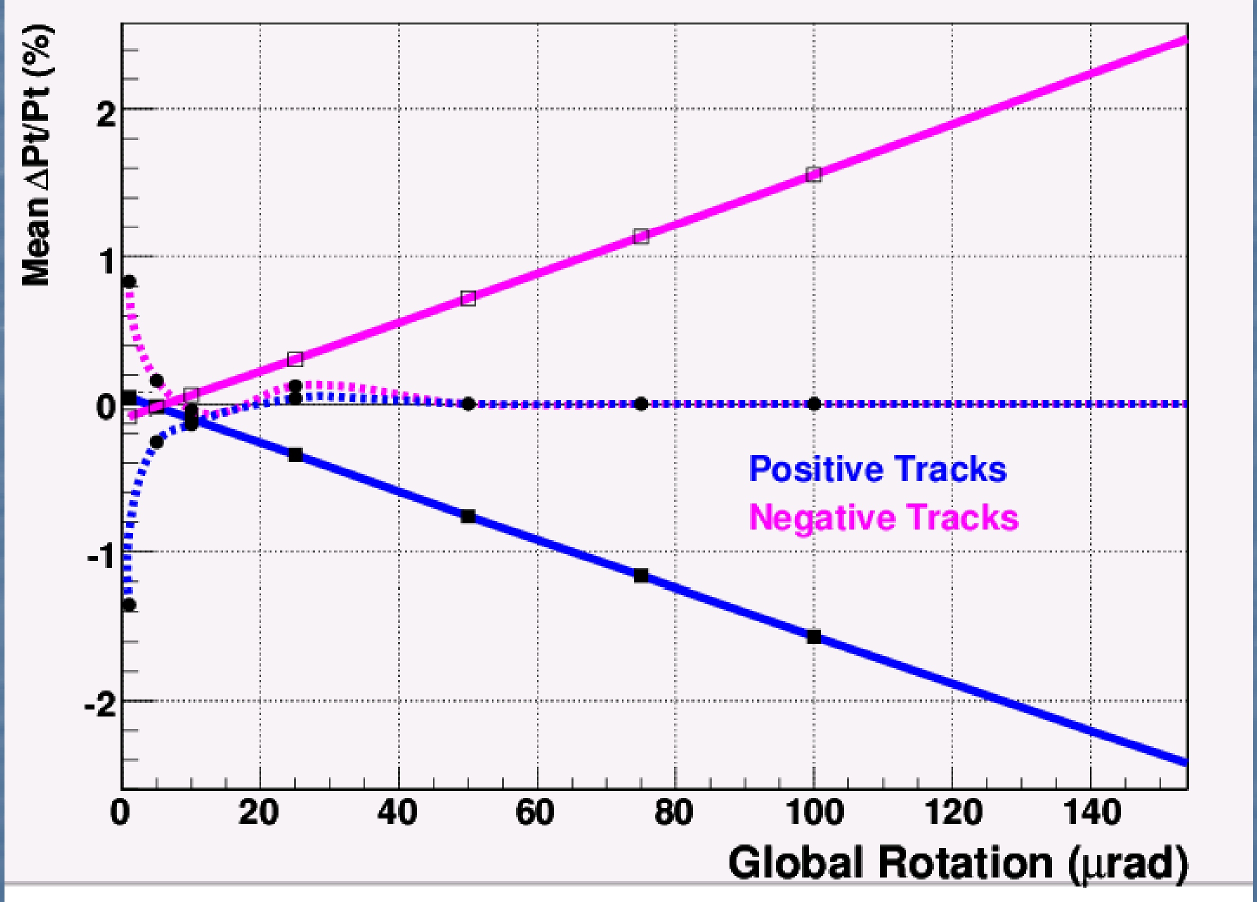
Small Structures:

- Impact on precision physics
- Require a lot of statistics
- Huge number of DoF
- Hard to monitor (need smart observables)
- Very sensitive to running condition (i.e. not stable)
- Difficult to find the their “truth” position (i.e. weak modes)

Example of Big Structure Bias

TRT Barrel Global Rotation w.r.t the Silicon

[Pt (Measured) – Pt (true)]/ Pt (true)



Note: 10μrad corresponds to a movement of ~10μm at the TRT radius

Alignment Implementation - I

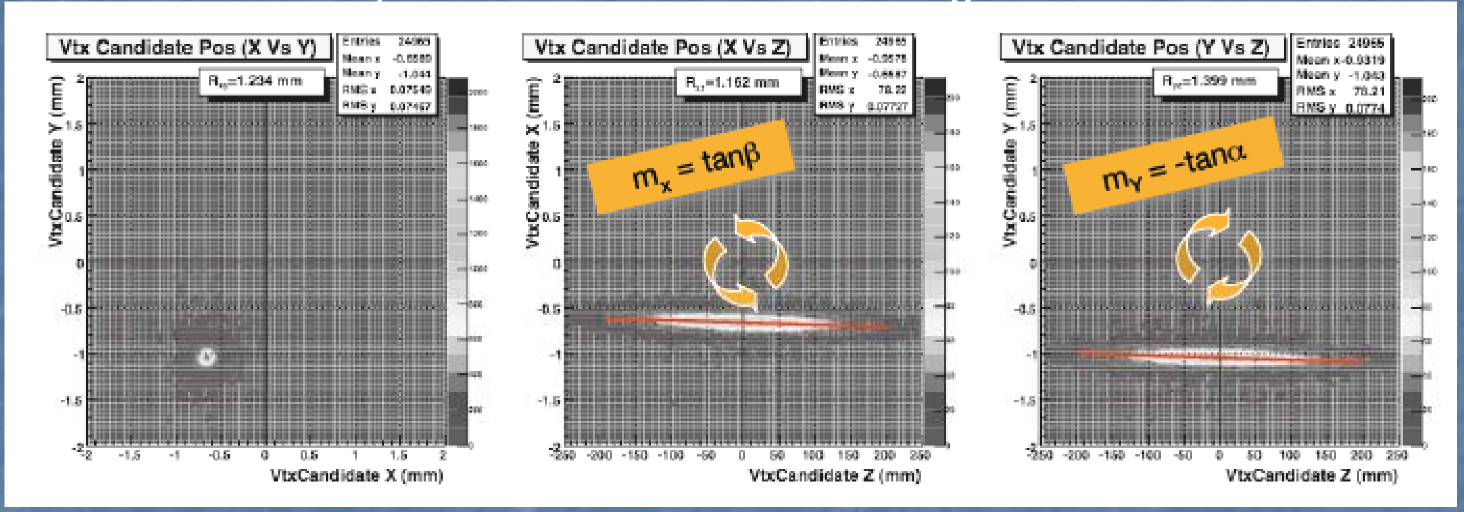
ID alignment can be performed at different granularity (a.k.a Levels)

The (in)famous alignment levels (i.e. detector granularity):

- **Level 1:** Subdetector Components (barrel/endcap, Si/TRT, …)
  - Non-gaussian impact parameter (vertex)
- **Level 2:** Layers/Disks(Si), Modules (TRT)
  - Produce momentum distortions (clocking/sagitta)
- **Level 3:** Modules (Si), Straws (TRT)
  - Almost no impact on residual distribution but essential for a correct track momentum and physics measurement reconstruction

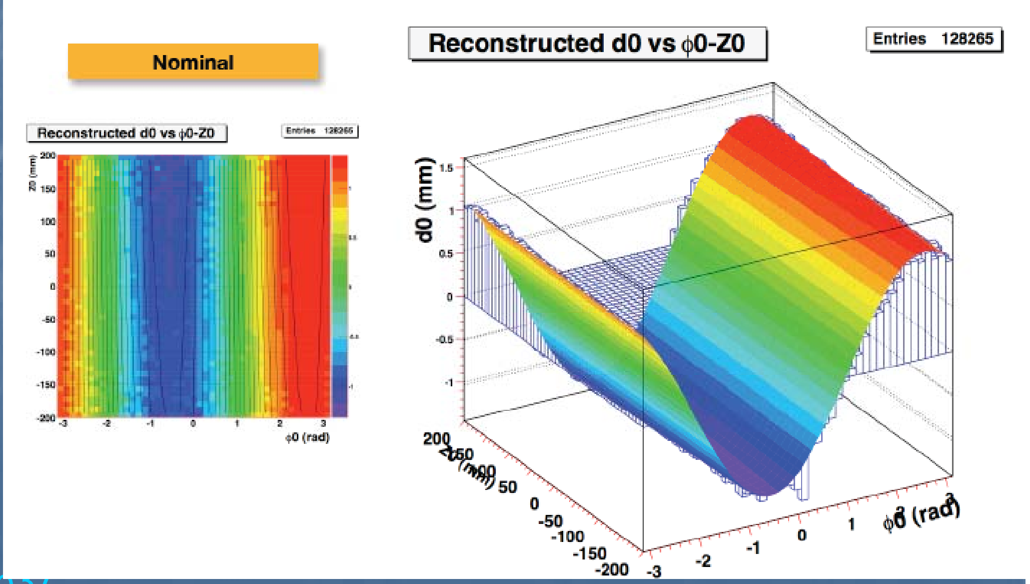
Beam Spot

Beam Position (and beam line) not expected to be centered in 0
Equivalent to a Level 1 misalignment



Not need sophisticated alignment algorithms

In ATLAS we do not center the detector on the beam spot position



Alignment Implementation - III

Recently added a new Level 2.5:

- Level 2 not enough to resolve all the global modes
- However Level 3 (full ~40K Dof) hard to solve
- There are physical substructures between Level 2 and Level 3
  - Pixel barrel staves, Pixel disk sectors, SCT barrel rows, etc...

Main Advantages

- Keep the matrix reasonably small (not need sophisticated solvers)
- Need less statistics than L3 (daily operation)
- Help to reduce bias before the full L3 alignment (enough for most of the physics)

Let's See How It Works

The Alignment Solution is:

$$\Delta\alpha = -\left(\frac{d^2\chi^2}{d\alpha^2}\right)^{-1}\frac{d\chi^2}{d\alpha}$$

The problematic part is the inversion of the second derivative

For ~1K parameters it can be handle by CLHEP or LAPACK.
For ~40K parameters we need to be smart

With parallel computing it has been shown that LAPACK can handle at least 10K parameters [4]

Unfortunately the CPU time grows as the third power of the matrix dimension

Let's See How It Works

The Alignment Solution is:

$$\Delta\alpha = -\left(\frac{d^2\chi^2}{d\alpha^2}\right)^{-1}\frac{d\chi^2}{d\alpha}$$

Several Strategies Investigated:

- Fast solver programs (MA27, MA57, PARADISO, ...)
  - Iterative method to minimize the distance

$$d = \left|\frac{d^2\chi^2}{d\alpha^2}\Delta\alpha + \frac{d\chi^2}{d\alpha}\right|$$

  - Quick (exploit the sparseness of the matrix). No error estimate.
- Full Inversion:
  - Time consuming. Error Estimate
- Diagonalization:
  - More time consuming. Error Estimate. Modes removal.

Diagonalization Output: Eigenvalues

The diagonalization is not the end of the story !

What we just do is rewrite the second derivative matrix A (symmetric):

$A \equiv \frac{1}{2}\frac{d^2\chi^2}{d\alpha^2}$ as $A = UDU^T$ with $D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$ $(U^{-1} = U^T)$

The inversion of A is now trivial: $UD^{-1}U^T$.

But this is also the covariance matrix for $\Delta\alpha$

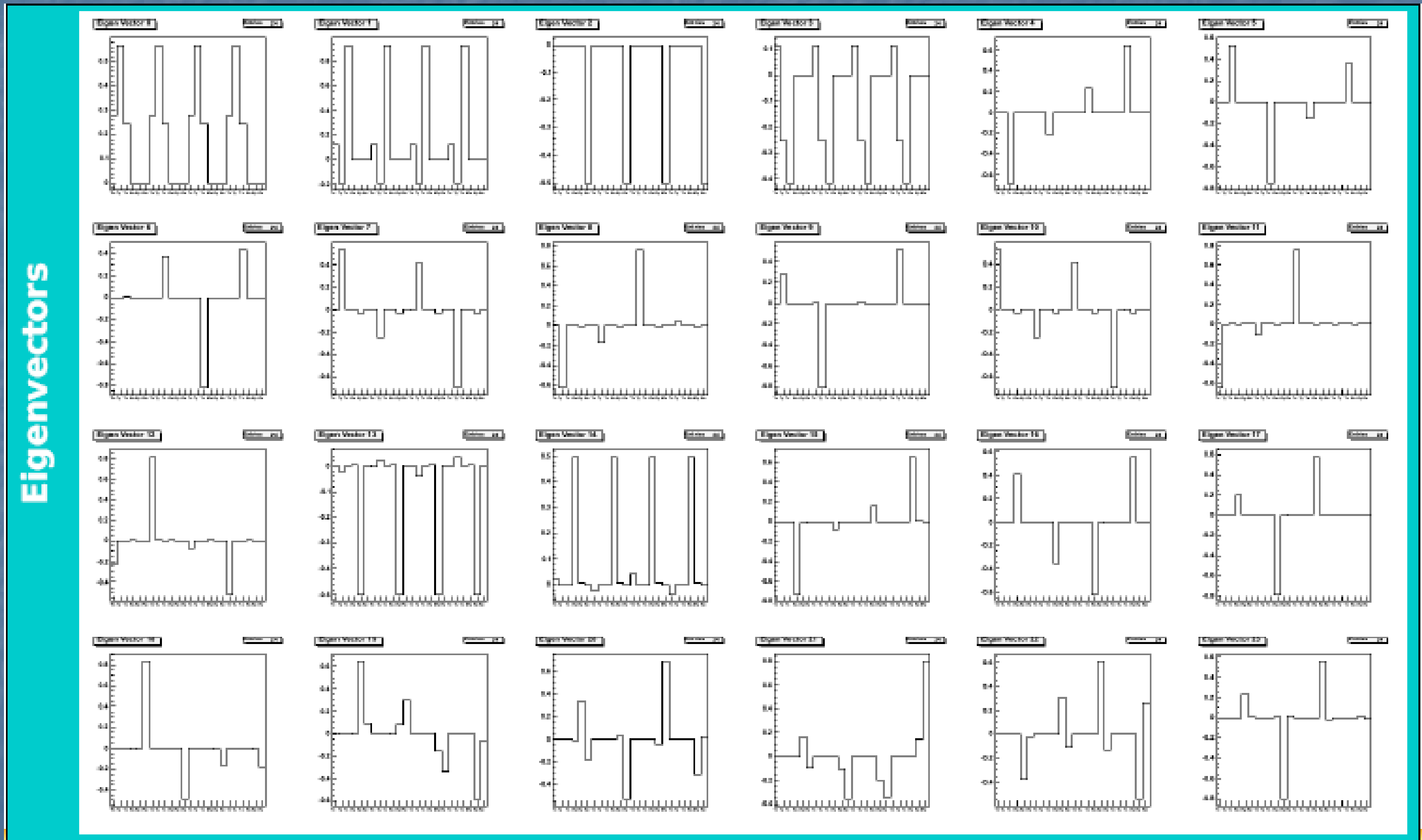
As a consequence the errors on ☟♋ are proportional to 1/⌦●

Conclusion: small eigenvalues lead to large statistical uncertainties in the alignment parameters

The associated eigenvectors are known as “Weak Modes”

Let's See How They Look Like

Alignment at Level2 of the SCT barrel (simulation): 4 SCT layers → 24 DoF

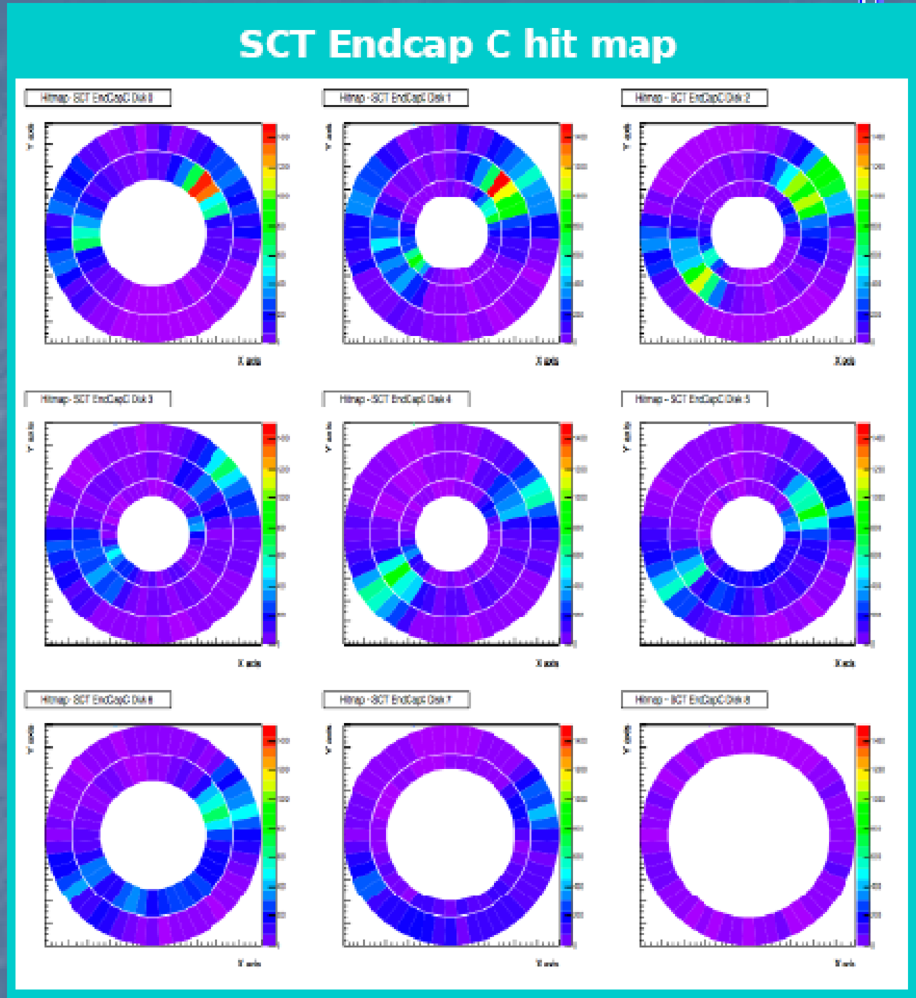


Where they come from

In general low eigenvalues comes from linear combination of alignment parameters that are poorly constrained

Structure in the detector with small statistics:
N_{hits} ❄ 0 implies ● ❄ 0

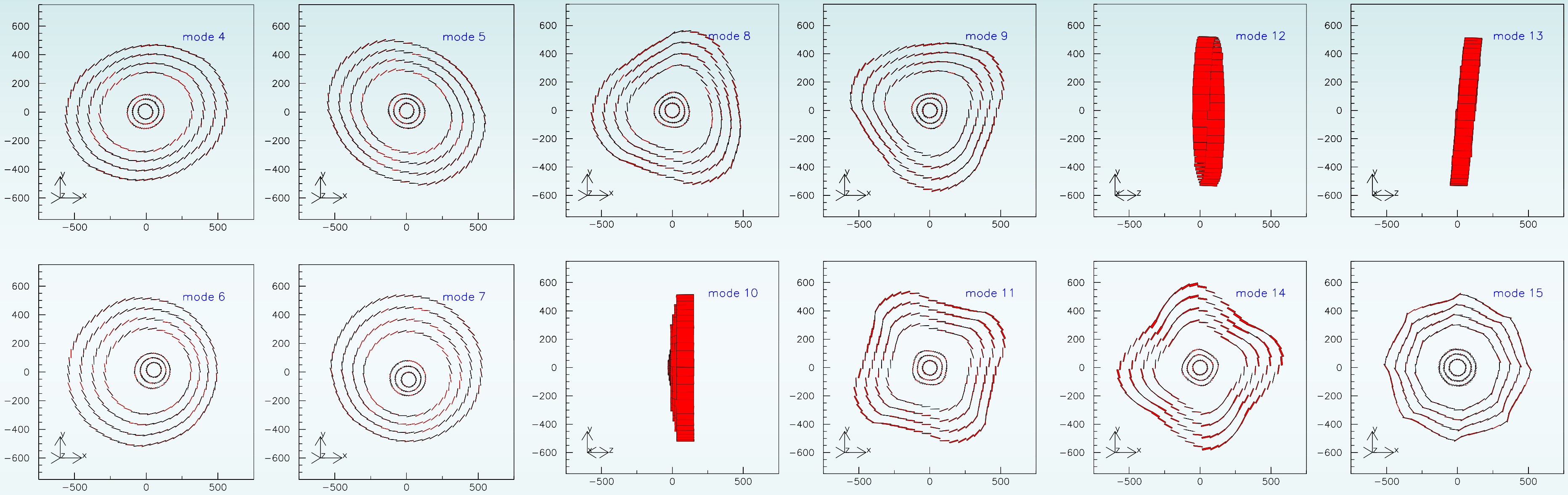
“Real” weak modes: configurations of the detector that leave the ♍2 unchanged (we’ll be back on that)



Exercise: What if A is not invertible? (i.e. rank A 🖫 dim A)

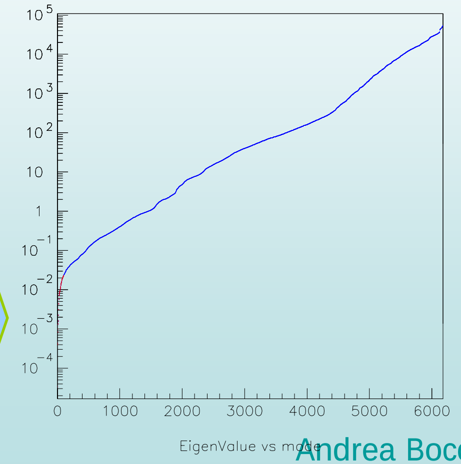
Example "lowest modes" in PIX+SCT as reconstructed by the χ^2 algorithm Global Freedom have been ignored

(only one Z slice shown)



➢The above "weak modes" contribute to the lowest part of the eigen-spectrum. Consequently they dominate the overall error on the alignment parameters.

➢More importantly, these deformations lead directly to biases on physics (systematic effects).



How to cope with them - I

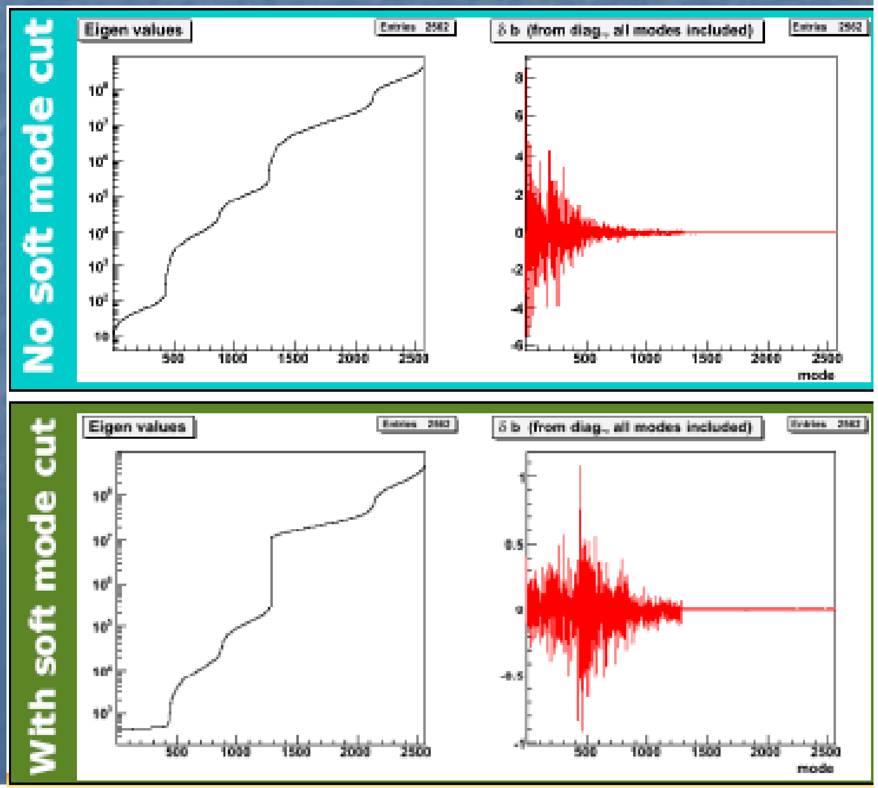
Mathematically, we can decide just to ignore them. However:

- Are (how?) we sure we will find a right solution ?
- How to decide the threshold?
- Anyhow the absolute values of ●’s are rather arbitrary (rescaling)

Possible approach (used in some simulated test):

- Constrain the movement within “reasonable” values
- Add a penalty term to the ♍2
- It keeps errors under control

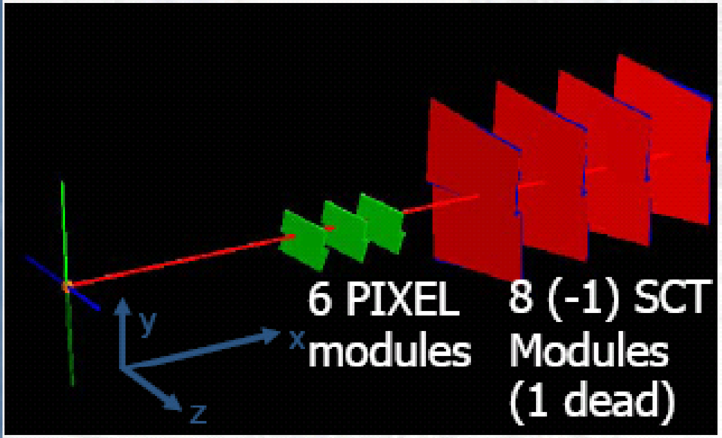
Example with real cosmics data
(very limited DoF)



And now some results...

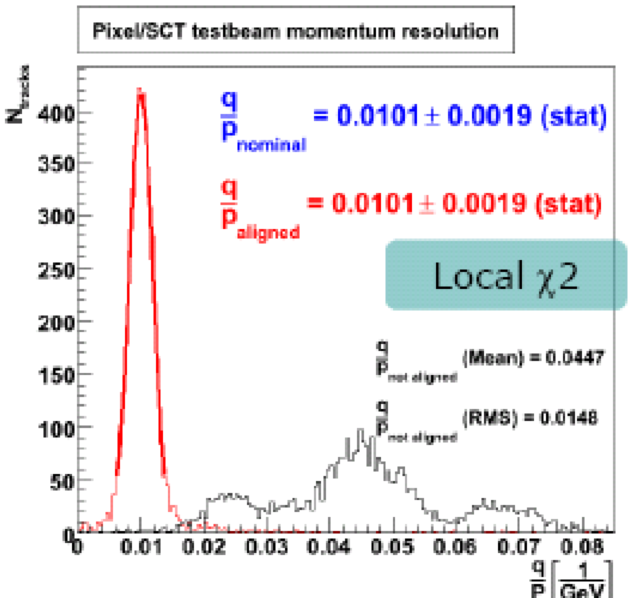
- Test Beam Data
- SR1 Cosmics
- Cosmics in the pit
- CSC Simulated data (full detector)

Combined Test Beam (2004)

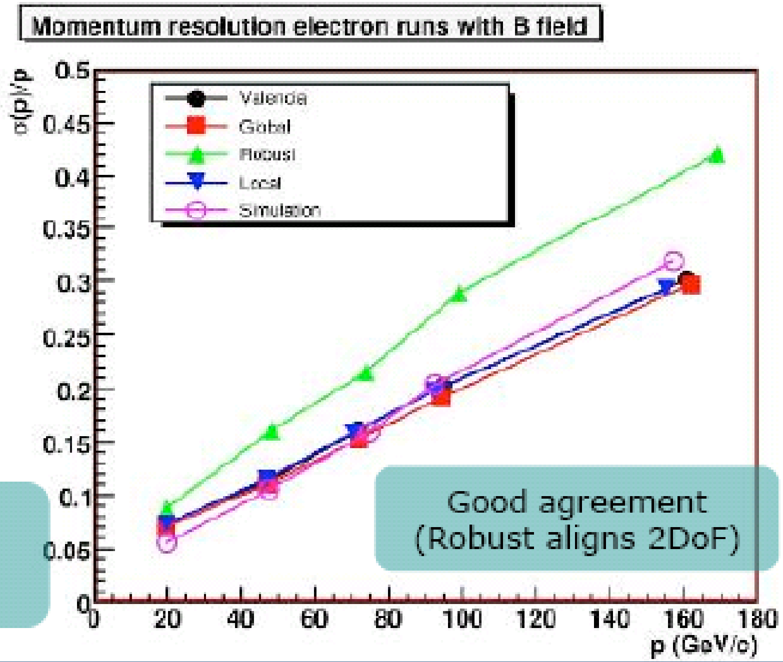


- First real data from Inner Detector!
- Large statistics of e^+/e^- and π (2-180 GeV) (O(10^5) tracks/module/E),
- Magnetic field on/off runs.
- Limited layout (systematic effects in modes)
- A good start to test algorithms for more realistic upcoming data!

100 GeV pions momentum spectrum

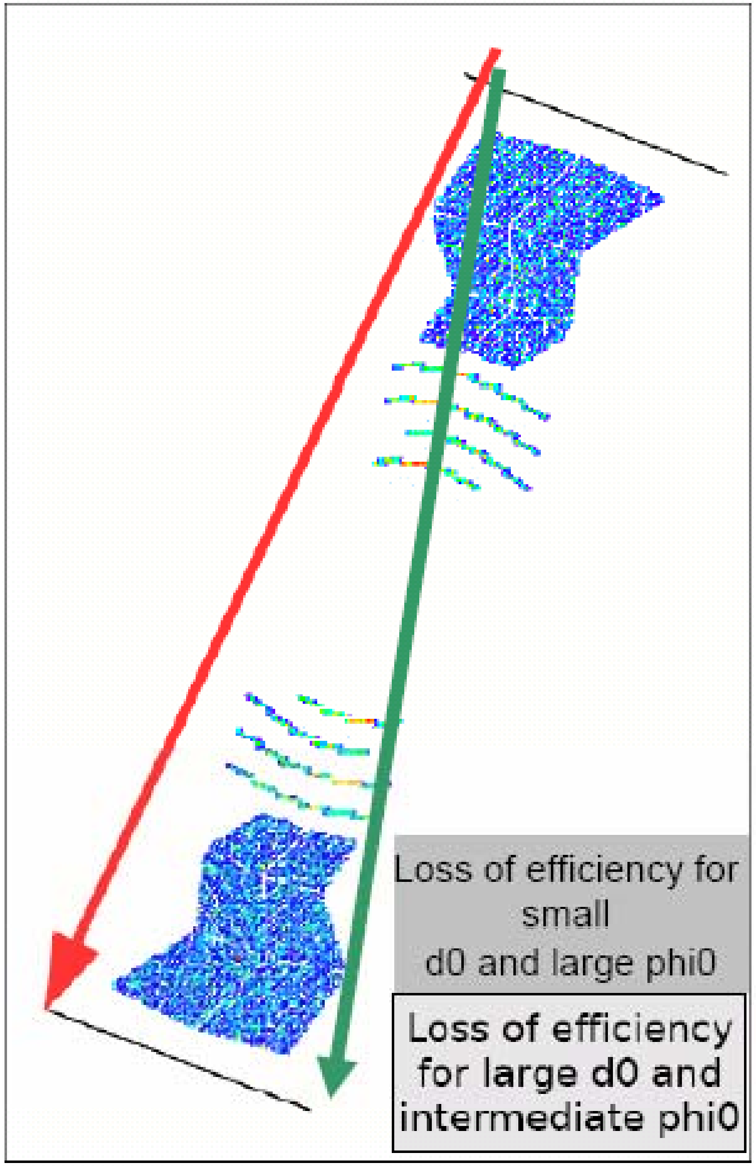


Momentum resolutions after alignment



Good agreement (Robust aligns 2DoF)

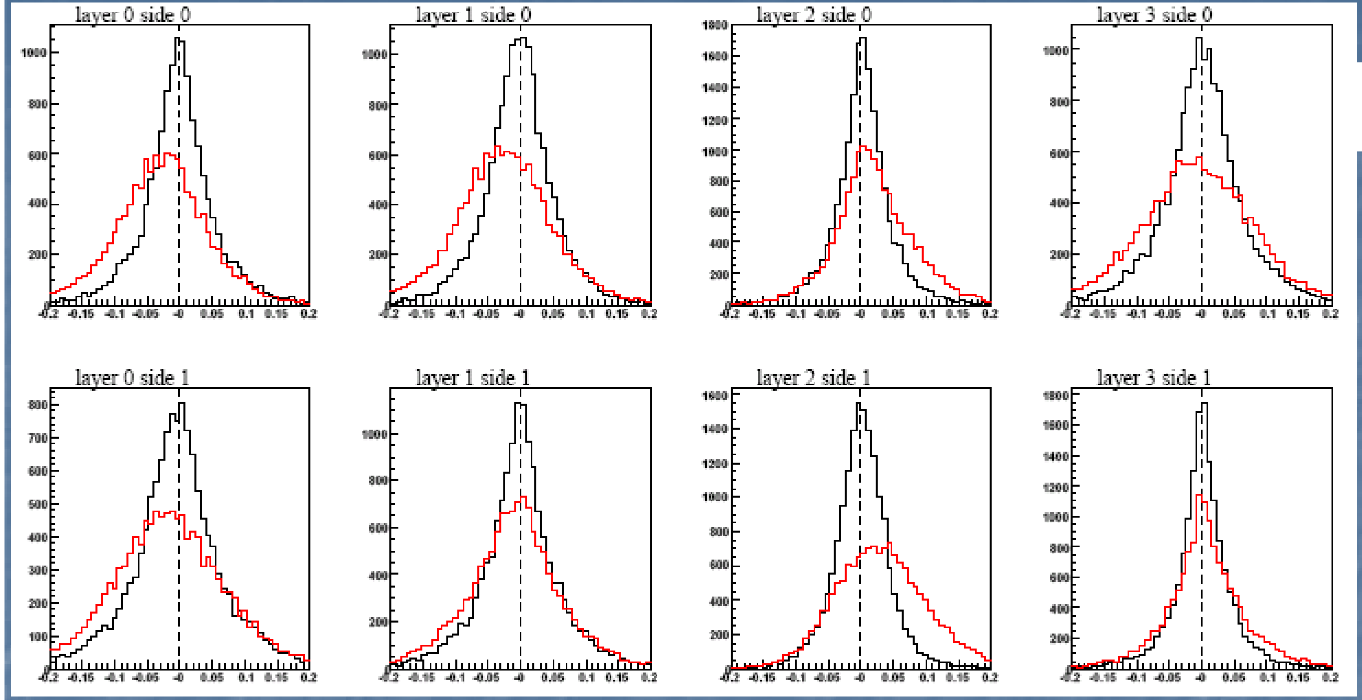
SR1 Cosmics Data



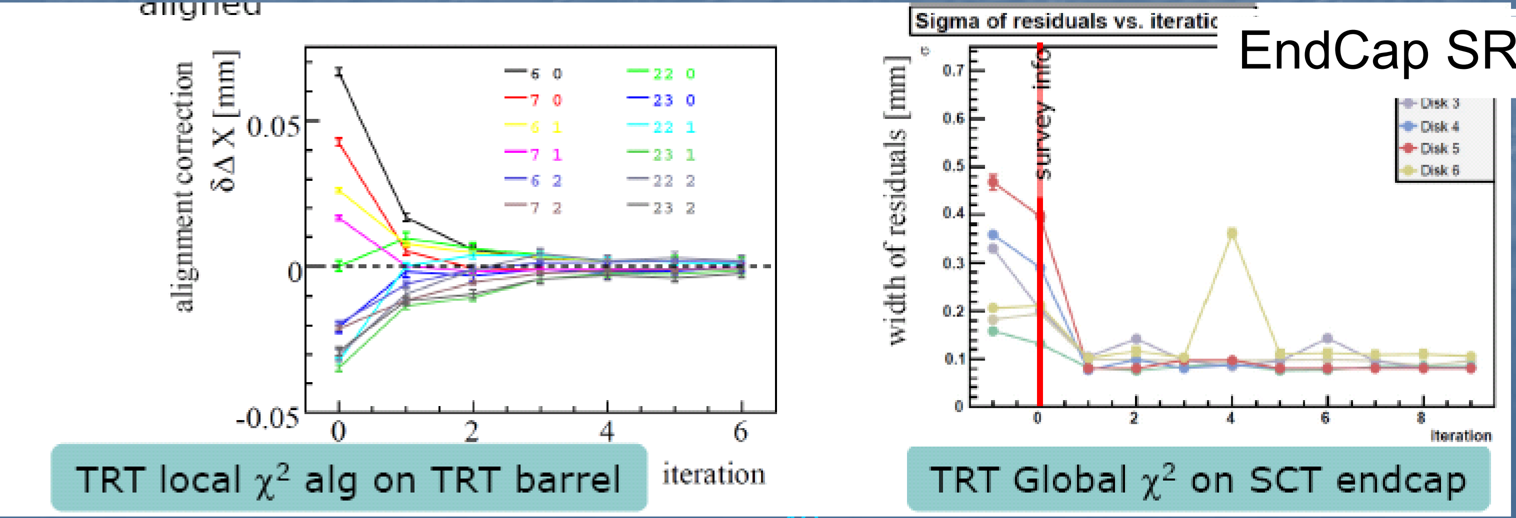
- Data-taking at surface (SR1) in spring 2006: ~400k events recorded!
- 22% of SCT Barrel: 467 instrumented modules
- 13% of TRT Barrel
- 3 (or 2) Scintillators to trigger 144cmx40cmx2.5cm
- Trigger allows ~ 0.3-0.4GeV cutoff for alignment studies
- No B-field! No momentum! MSC important ~<10 GeV, need to deal with large residuals

- All the algorithms have been tested with those data
- For the first time a combined TRT+SCT alignment was performed using both the global and the local χ2 methods

SR1 Cosmics Data



SCT Residuals



EndCap SR1 Cosmics

TRT local χ^2 alg on TRT barrel

TRT Global χ^2 on SCT endcap

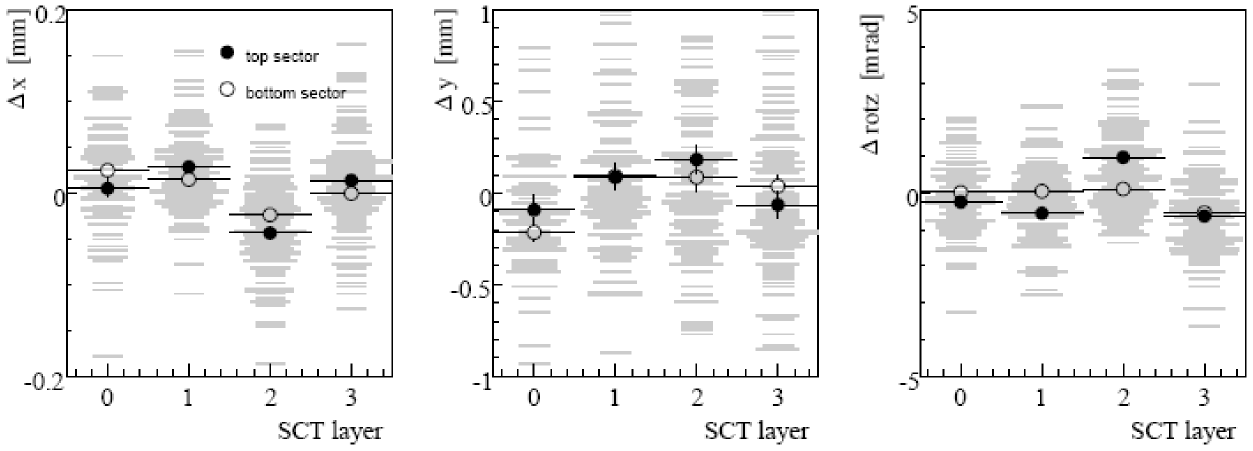
SR1 Cosmics Data

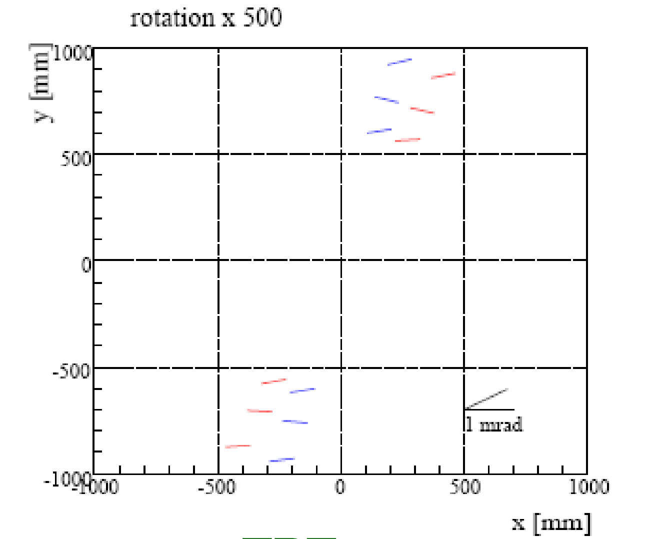
First measurement of the position of the ID substructures with tracks

Observed L1 misalignment between SCT and TRT

run	η_x [mm]	ϵ_y [mrad]	ϵ_z [mrad]
3007	-0.290 ± 0.001	0.277 ± 0.003	0.254 ± 0.002
3099	-0.289 ± 0.001	0.293 ± 0.003	0.226 ± 0.002

Module Displacements

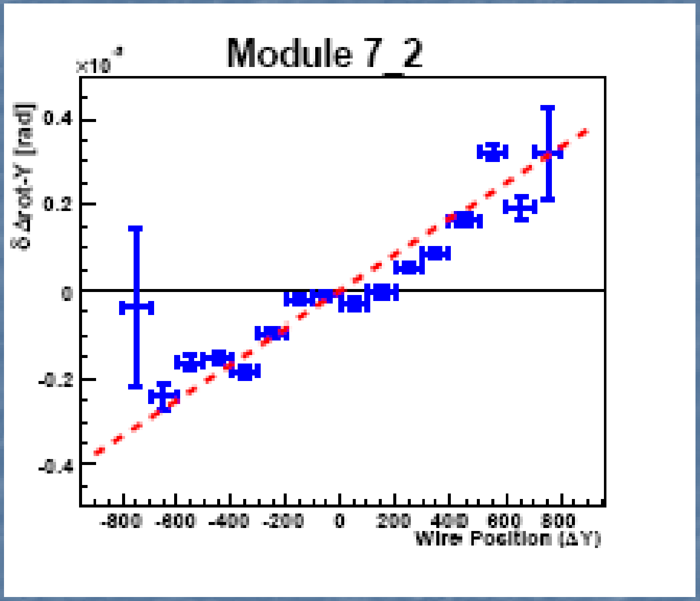




TRT

SR1 Cosmics Data

First Evidence of Module Deformation (deviation from the rigid body approx.)



Rotation as a function of z in the TRT (barrel)
TWIST!

Deformation expected also for silicon modules

Need a lot of data (or hardware devices like FSI in the SCT) to assess them

Cosmics in the Pit

Inner Detector in the pit since 2007. However a lot of hardware problems

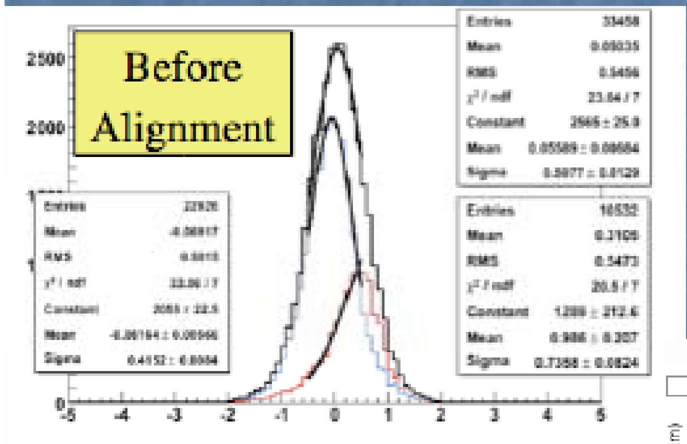
Several Milestone Runs with a cosmics trigger for commissioning
TRT always present (with increased instrumentation coverage)

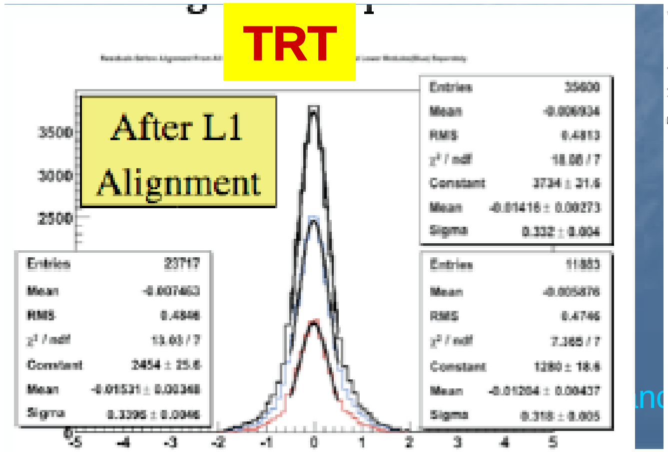
Before last week only one set of runs (M6) had both SCT and TRT

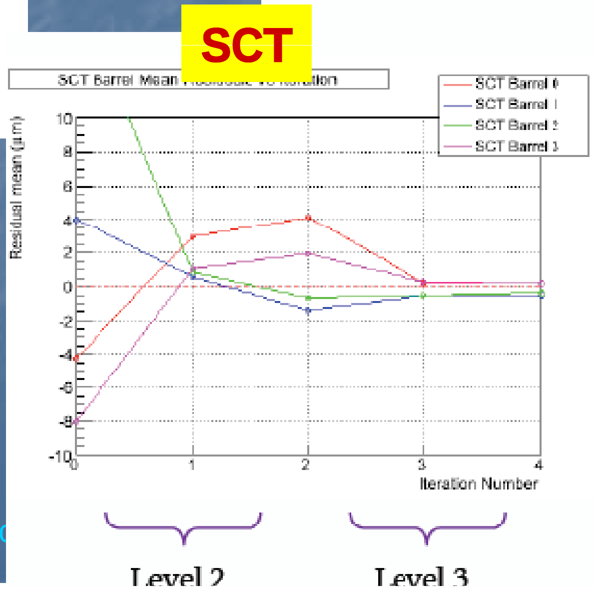
First Pixel+SCT combined run few days ago

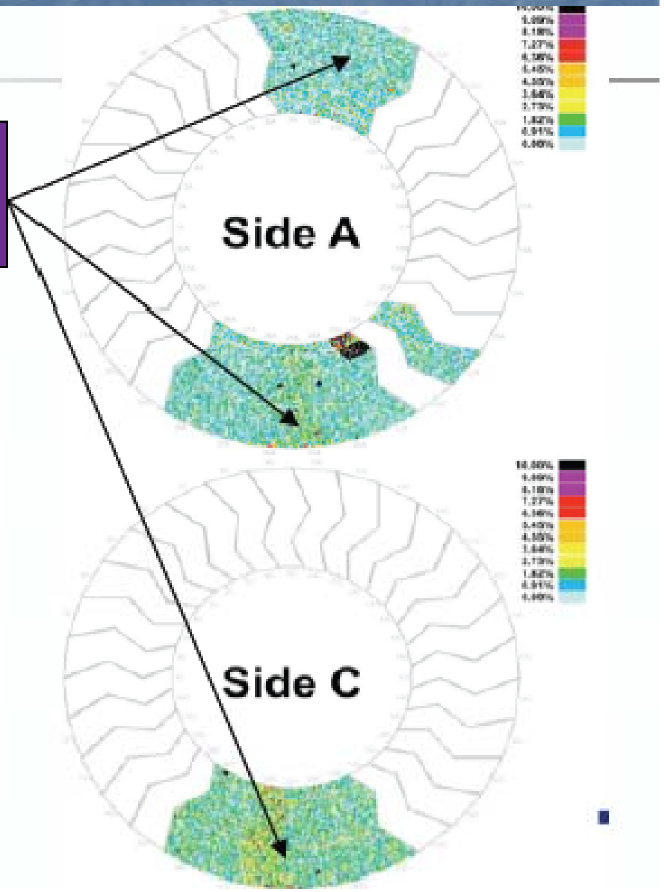
Alignment of M6 Cosmics data

- Limited statistics
- Limited part of the detector instrumented
- Not uniformity illumination of the instrumented parts
- Just few iterations at level 2 for the SCT
- TRT aligned after the SCT with full tracks (SCT+TRT)
- Constants found will be the starting point for day1









Full Scale Test of the ID Alignment (CSC)

- The Computing System Commissioning challenge:
  - Test the ATLAS software and computing infrastructure with an *as built* detector configuration
  - Huge sample of physics and calibration events simulated with
    - Uncalibrated and misaligned geometry
    - Inhomogeneous axial ID magnetic field
    - Distorted material

The Goal: Provide a good set of alignment constants for such "detector":

- Good residual distribution ⇒ Algorithms are working well
- D0 distribution ⇒ Vertex position correct
- Pt distribution ⇒ No clocking or sagitta effect
- No significant degradation of physics measurement ⇒ Z peak

The great challenge of the CSC exercise prompted a fast development and implementation of ideas as long as extensive tests of alignment algorithms, monitor and validation procedure

The "CSC" Detector

	LEVEL 1	X	Y	Z	α	β	γ
TRT	TRT Barrel	+1	+1	+1	0.20	-0.05	0
	TRT Endcap A	+2	-1	+2	-0.15	0.10	0
	TRT Endcap C	-2	+2	-3	-0.20	-0.15	0
SCT	SCT Barrel	+0.70	+1.20	-1.30	0.10	0.05	0.80
	SCT Endcap A	+2.10	-0.80	-1.80	-0.25	0.00	-0.50
	SCT Endcap C	-1.90	+2.00	-3.10	-0.10	0.05	0.40
Pixel	Whole	+0.60	+1.05	-1.15	-0.10	0.25	0.65

(displacements in **mm**; rotations in **mrad**)

L2	Layer	Systematic radial shift	Random shift in X,Y
TRT	Layer 0	+1.0 mm	0.2 mm
	Layer 1	-0.5 mm	0.1 mm
	Layer 2	+1.5 mm	0.3 mm

- Shifts are realistic !
  - Though may seem huge
  - Surveyed during assembly…

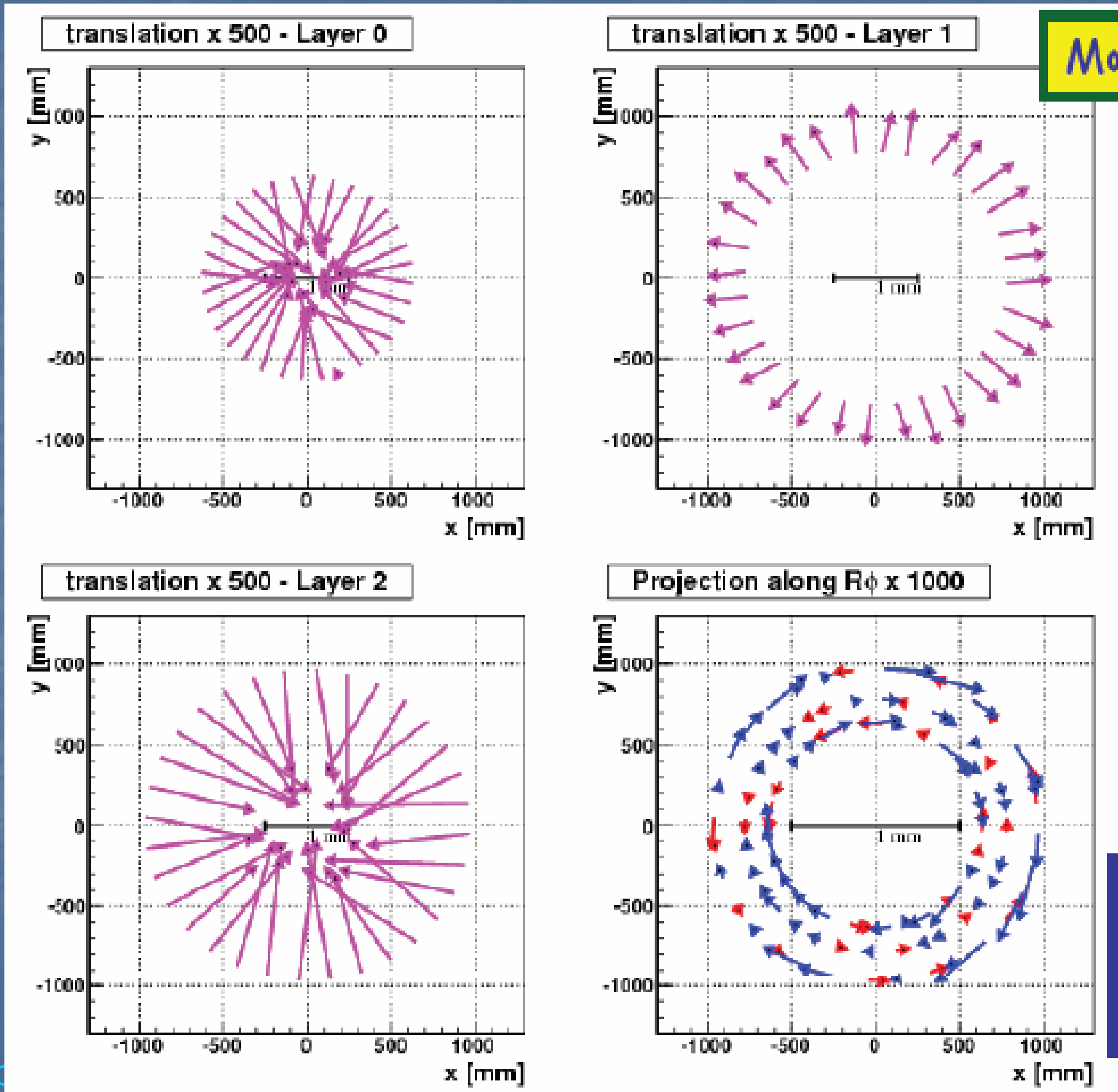
LEVEL2

	Layer/Disk	X	Y	Z	α	β	γ
Pixel Barrel	0	0.020	0.010	0	0	0	0.006
	1	-0.030	0.030	0	0	0	0.005
	2	-0.020	0.030	0	0	0	0.004
SCT Barrel	0	0	0	0	0	0	-0.001
	1	0.050	0.040	0	0	0	0.009
	2	0.070	0.080	0	0	0	0.008
	3	0.100	0.090	0	0	0	0.007
SCT Endcap A	1	0.050	0.040	0	0	0	-0.001
	2	0.010	-0.080	0	0	0	0
	3	-0.050	0.020	0	0	0	0.001
	4	-0.080	0.060	0	0	0	0.002
	5	0.040	0.040	0	0	0	0.003
	6	-0.050	0.030	0	0	0	0.004
	7	-0.030	-0.020	0	0	0	0.005
	8	0.060	0.030	0	0	0	0.006
	9	0.080	-0.050	0	0	0	0.007
SCT Endcap C	1	0.050	-0.050	0	0	0	0.008
	2	0	0.080	0	0	0	0
	3	0.020	0.010	0	0	0	0.001
	4	0.040	-0.080	0	0	0	-0.008
	5	0	0.030	0	0	0	0.003
	6	0.010	0.030	0	0	0	-0.004
	7	0	-0.060	0	0	0	0.004
	8	0.030	0.030	0	0	0	0.006
	9	0.040	0.050	0	0	0	-0.007

LEVEL3

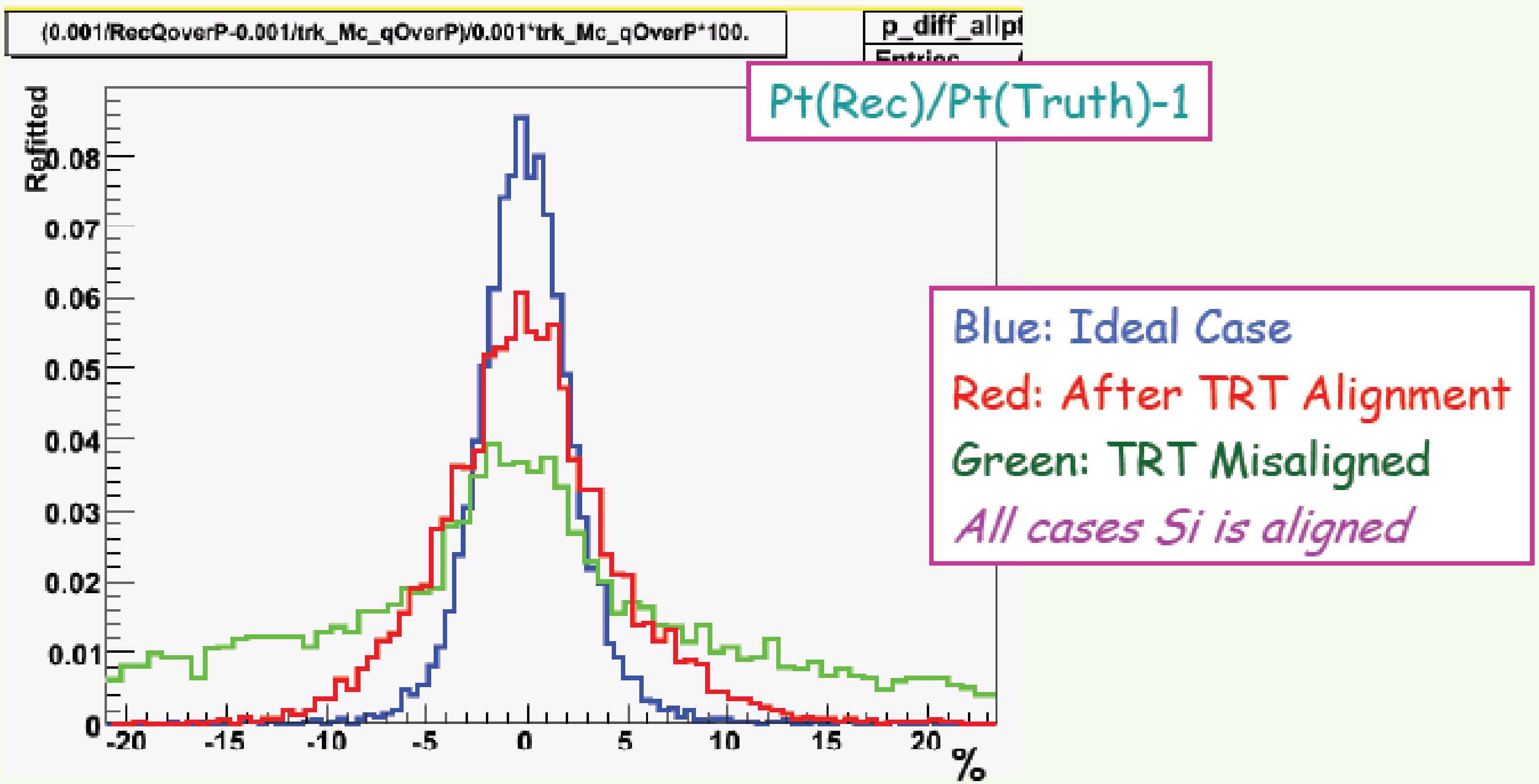
	x	y	z	α	β	γ
Pixel Barrel modules	0.030	0.030	0.050	0.001	0.001	0.001
Pixel Endcap modules	0.030	0.030	0.050	0.001	0.001	0.001
SCT Barrel modules	0.150	0.150	0.150	0.001	0.001	0.001
SCT Endcap modules	0.100	0.150	0.150	0.001	0.001	0.001

TRT L2 Module Displacements



Level 1 (TRT-Silicon)

After Global Only (L1)

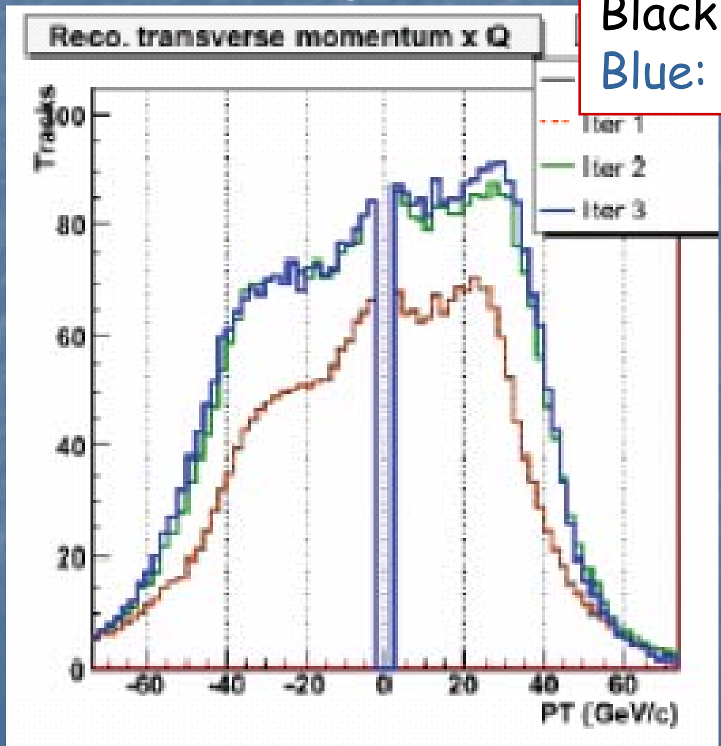


Centering the TRT and Silicon has a big impact in recovering the momentum (for Si-TRT tracks)

CSC Lesson 1

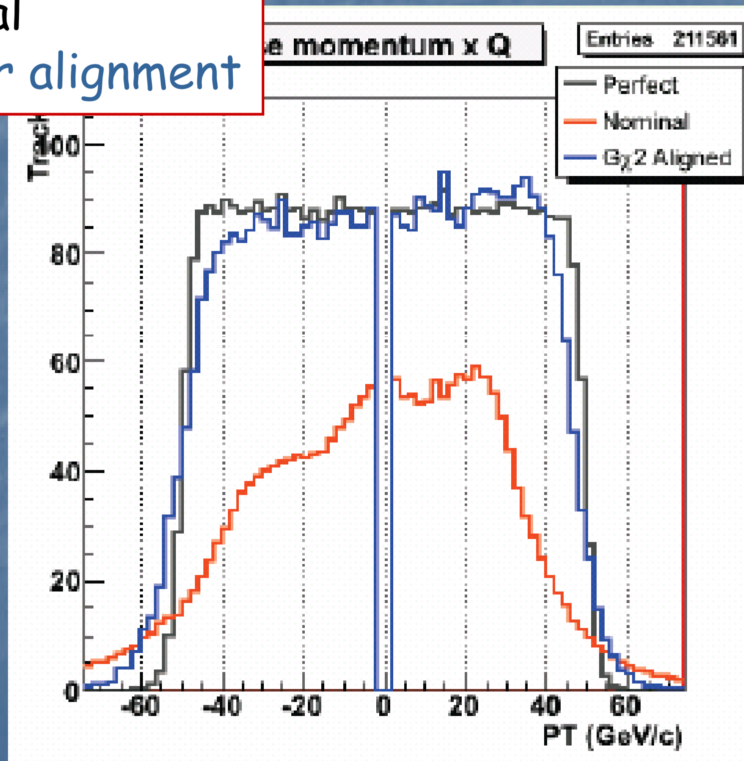
Misalignment at Level 3 has basically no impact on residuals but huge impact on physics (Bad news: we minimize the residuals!)

L3 Misaligned



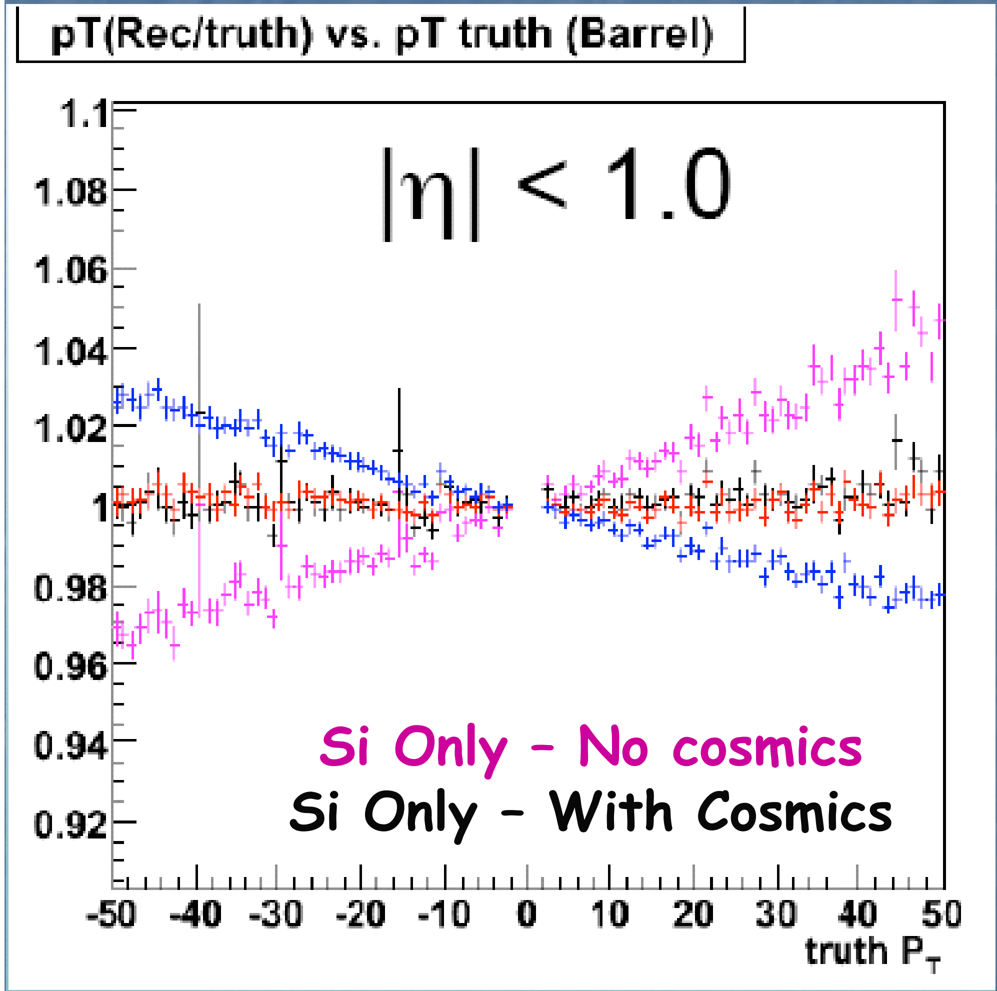
Black: ideal
Blue: after alignment

L3 Aligned



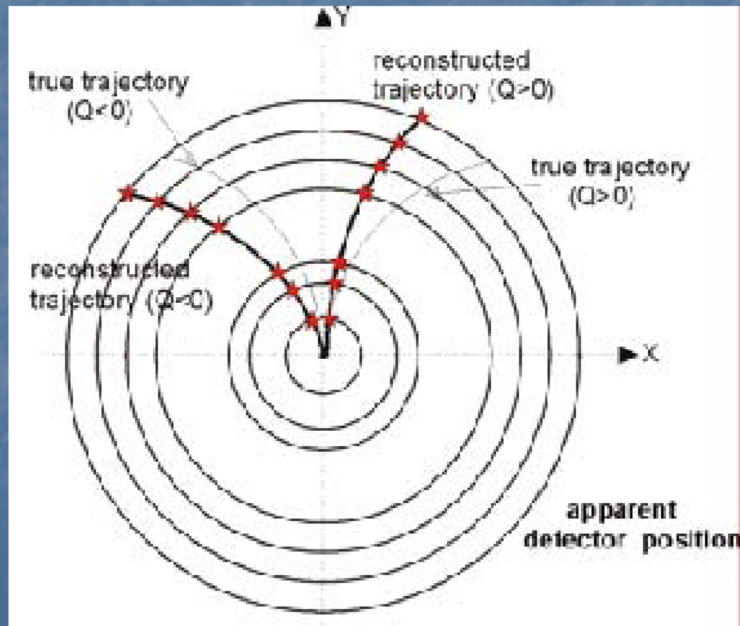
CSC Lesson 2

Tracks generated from the IP are not enough



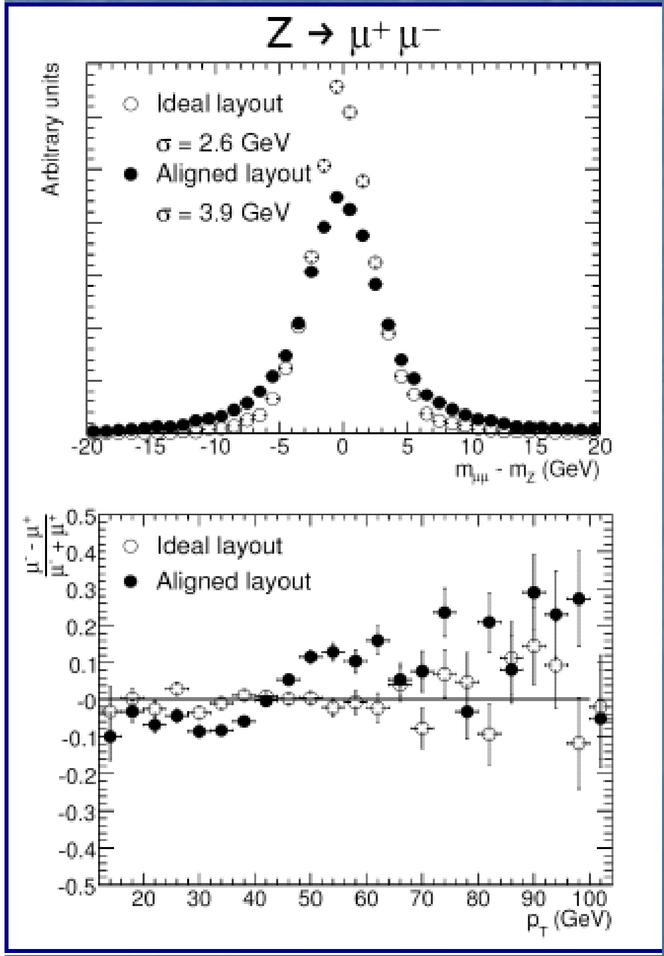
Use of cosmics data complementary to tracks from collisions

Helps to resolve positive/negative asymmetry bias (clocking effect)



CSC Lesson 3

Minimizing the residuals is not enough to recover the "correct" geometry



Even with cosmics the Z peak after alignment still wider than the "ideal" layout

Detector movements corresponding to poorly constrained degrees of freedom (a.k.a. **weak modes**) are not fully recovered

Weak Modes:
The Ultimate Challenge

What we mean by Weak Modes

The term “weak modes” is a little abused these days: it’s the main (only) issue for the alignment (not only ID)

I introduced the weak modes as the eigenvectors with “small” eigenvalues

- “small” eigenvalues lead to high uncertainties
- A (linear) combination of alignment parameters that are poorly constrained by the track sample used in the alignment
- Note: Nothing wrong with the “math”, it’s a feature of the track sample
- Note 2: For the same detector, different track samples can lead to different weak modes (it’s a good thing!)

Consequences of the presence of weak modes in the solution:

- The final detector position is not equivalent to the real one and physics observables are biased.

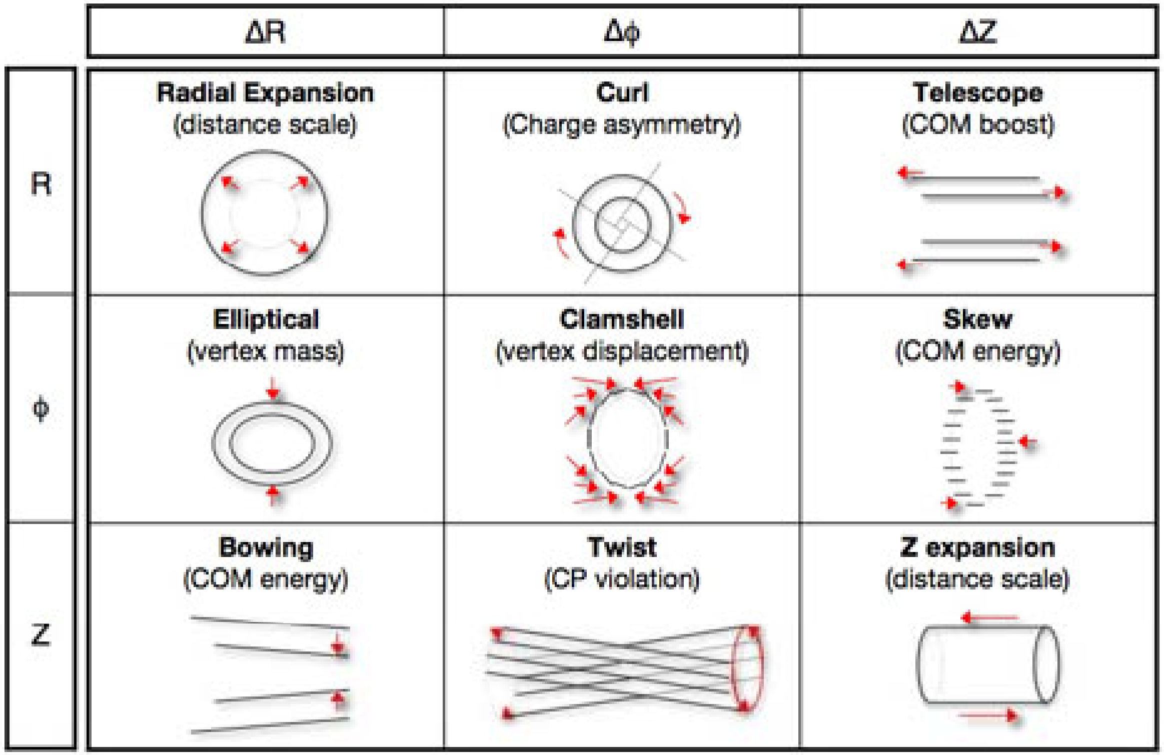
Sometime the actual detector configuration/deformation is called the weak mode

Weak Mode Configurations

Configurations corresponding to weak modes are hard “to move”
(poorly constrained implies poor convergence)

Important to identify the impact on physics of those modes:
- to give the “good enough” flag to the alignment procedure
- to take into account of misalignment systematic in the physics analysis

Big effort started to create ID geometries with such configurations in order to learn how to recognize them in the data (see Giorgio tutorial)



What It Can Be Done

By definition the weak modes are such movement that the track-based residual minimization (with collision tracks) cannot resolve

It is not question of math, there is no magic trick to get rid of them

The only option is the use of "extra" information that we generally call "external constraints"

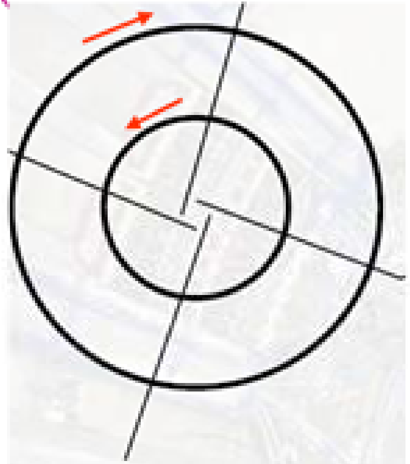
Cosmics Data:

- Very helpful to correct for clocking effects
- Already part of the alignment procedure
- Unfortunately not that useful in the endcaps
  - Beam gas/beam helo tracks can do a similar job

What It Can Be Done

Vertex Constraint:

- Very helpful to correlate different part of the detector
- Not used at the moment but implementation is in place (need some trick to make it right)
- Some redundancy with beam constraint
- It will make the matrix fully populate (not sparse)



"clocking"
$\delta\phi = \lambda + \beta / R$

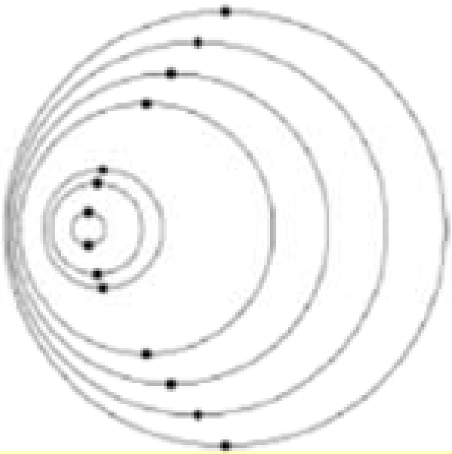
Pt Constrain from the TRT:

- Idea is to use the momentum information from the TRT as a constraint for silicon tracks
- Helps for momentum biasing weak modes (the hardest!)
- We have the tool for first test
- Need to assess the impact of an internal misalignment TRT
- It should come from free with once we have the new code to perform the ID alignment as a whole

What It Can Be Done

Mass constraint:

- Use known resonances ($Z \longrightarrow \mu\mu$, $J/Psi \longrightarrow \mu\mu$, etc...)
- Only second order dependence on track momentum
- Helps in configurations where cosmics cannot do much
- Best are resonances in different part of the detector (collimated/boosted resonances not very useful)
- Strongly dependent on the purity of the sample
  - Essential to understand the background
- Make matrix fully populated (not sparse)



Φ Dependent sagitta

Comparison B field ON/OFF:

- Cosmics with B field ON not as good as with B field OFF
- Need to assess the effect of ON/OFF switch
- Keep toroids ON and switch solenoid ON/OFF also fine
- Need more thinking on the possible gain and understand what is technically possible

The End

Further Reading

[1] A.Bocci and W. Hulsbergen, ATL-INDET-PUB-2007-009
[2] P. Bruckman et al, ATL-INDET-PUB-2005-002
[3] V. Blobel, Millepede, NIM A 566 (2006) 5
[4] M. Unel et al, Parralel computing, Contribution to CHEP2006

Backup

Alignment Strategy for the First Days

On the FDR2 exercise we proved we can provide good alignment constants in the given time:

- Still need some optimization for maximum speed

Before collisions use all the possible/usable tracks to have an idea of the "real" initial misalignment:

- Cosmics, beam halo, etc…
- Maybe using the finding as initial value?

When a statistical significant set of tracks from collisions is available:

- Bootstrap of the alignment constants
- Provide good tracks to the experiment so it is possible to use them to start understand and calibrate the rest of the detector
- Use the alignment monitoring to have an idea about the stability of the beam/alignment in real condition and react accordingly
- When confident enough and no other urgent issues, switch on the Level 3

Continue with a stable-condition standard 24-hours operation routine

Plans for the Second Pass Alignment

Data will be reprocessed after O(few months) at Tiers 1's with the most precise calibration/alignment available

At this stage the alignment group should provide the best alignment constants to correct all possible misalignment/distortions of the detector
(Those reprocessed data will be used for physics analyses)

The focus is now on how we can reach the ultimate alignment precision (and therefore the maximum potential from the Inner Detector)

A systematic study of the weak modes has already been started:

- Not all the weak modes are equal, some may impact the physics more than other
- How to identify the weak mode of our detector?

Need feedback from the physics/performance group

Center of Gravity (CoG)

- Use the CoG of the whole ID as a fixed reference for the alignment
  - i.e. the position in "space" of the CoG will not move
  - Flexibility in the definition of the CoG with the use of weights for individual sections/modules
    - Helps to reduce dependence on less stable components
  - Note: the exact form of fixed reference does not need to be fixed once for all

- CoG algorithm based on a pioneer implementation in the alignment monitor (Tobi G., using the ALEPH tecnique) for the CTB and generalized by Sergio

- Tests of the implementation done on CSC misalignment geometry

- CoG determination now part of the alignment chain and included in the FDR2 exercise
  - Basically a final movement to bring back the CoG at the starting (before alignment) position is added

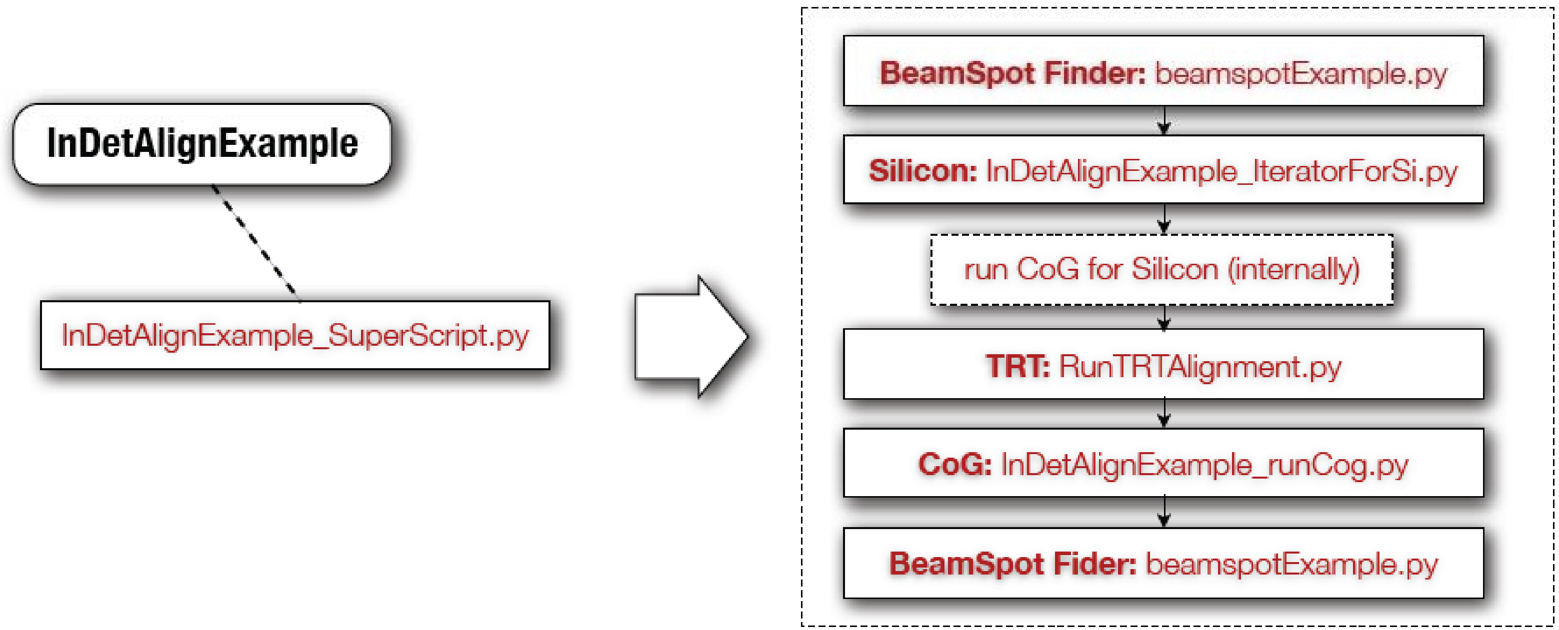
FDR2 experience

In general, the FDR2 was a success for the ID alignment group.

- FDR2 alignment was focused on implementation of the technical infrastructure.
- The primary goals were achieved:
  - ID alignment constants were produced and delivered on time.
    - Used for FDR2 bulk reprocessing.
  - Decent quality for a start up scenario (Nominal geometry).
    - Although it is true that still behind CSC constants.
- In general, access to data & input/output worked quite well (afs, castor, w0, CAF, etc...)
- Input data:
  - Calibstream and cosmic rays were used (main difference wrt the FDR1 exersice).
- ID Alignment software worked well with official FDR2 release: 14.1.0
  - Express stream was not available / used for us.
- Reports were given at the FDR2 daily meetings.
  - Very active participation also from InDetAlignmentMonitor in the express stream.

Alignment SuperScript

- First usage of the ID alignment superscript
  - Aim: Do the entire job (Si, TRT, BS, CoG, Db) pressing just one button
  - Didn't have enough time to fully test it before FDR2
    - Some small problems and human intervention needed (see Carlos talk at the FDR2 washup meeting on 23/6/08 http://indico.cern.ch/getFile.py/access?contribId=4&resId=0&materialId=slides&confId=35770)



FDR2 Experience

- Except for the few technical problems the whole chain worked well
- Robust alignment not ready for FDR2 but now implemented in the SuperScript as option
- Cosmics implemented as well in the chain
  - Since cosmics make the matrix not sparse we adopted the L2.5
- TRT alignment also worked and done after silicon
  - Used perfect calibration
  - Need to optimize the number of iteration as a function of statistics and convergence (number of L1 iterations vs. number of L2 iterations)
- Need to work on overall speed to be sure we can do a good job in the 12-hours limit
  - FDR2 used single particle, we need to test the speed on a more realistic ID alignment stream data sample
- For Beam Spot and ID monitoring consideration see next talks

Alignment Results

Disclaimer: the main goal of the FDR2 exercise was to test the ATLAS model of the everyday data operation and the computing model of data acquisition (including ES validation, update of db, sign-off procedures)

This kind of exercise is not the ideal test-bench to debug problems and issue of the alignment procedure/results

Unfortunately the validation of the alignment code in rel. 14 was not completed by the time of the FDR2 exercise

A change in the extrapolated position of the track in the SCT was introduced in rel 14 and not taken into proper account by the alignment code

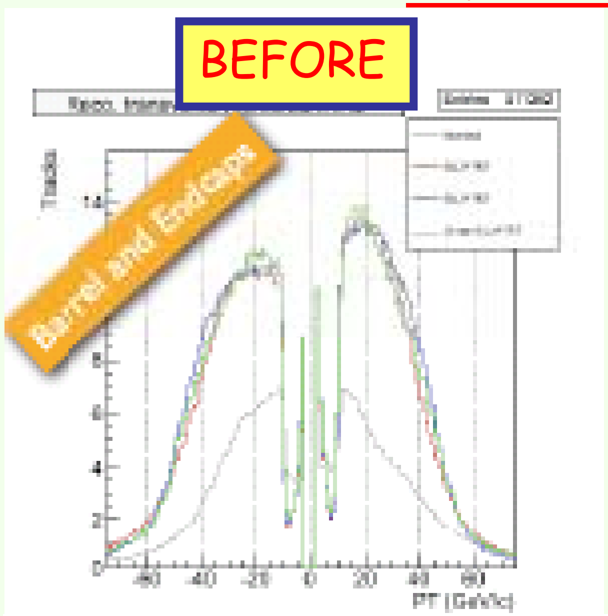
As a result the residuals calculation was not correct making the alignment virtually impossible to be properly performed (with more devastating effect in the endcap because of the peculiar geometry)

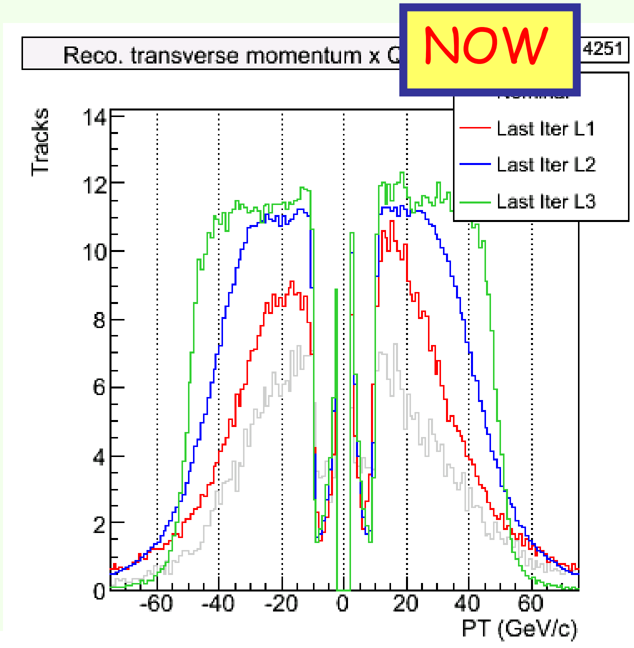
A fix has now been implemented and we are redoing the alignment of FDR2 data

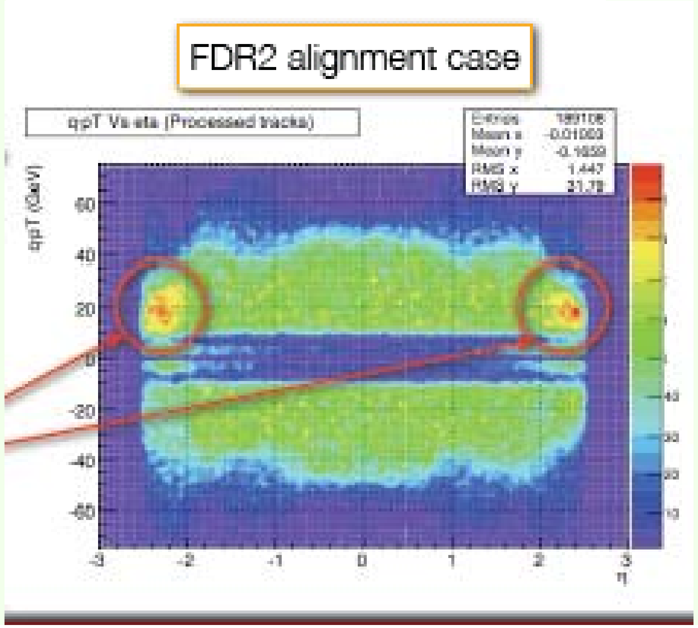
Recent Rerun of FDR2 Data

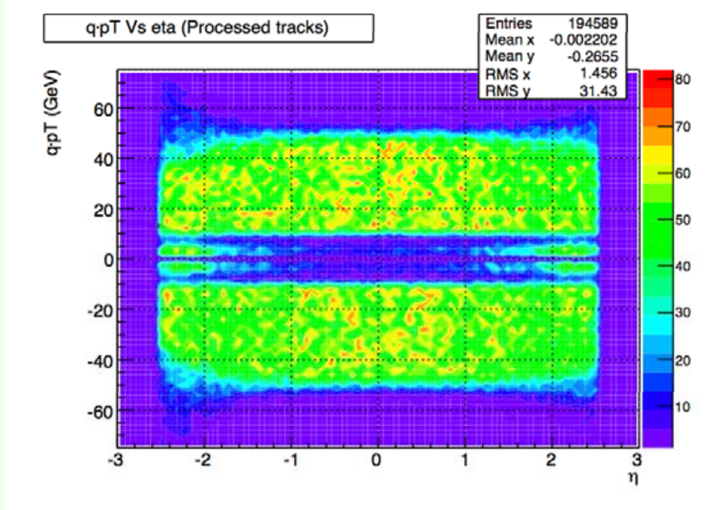
Thanks
Vicente & Ben!

… finished just before the coffee break…



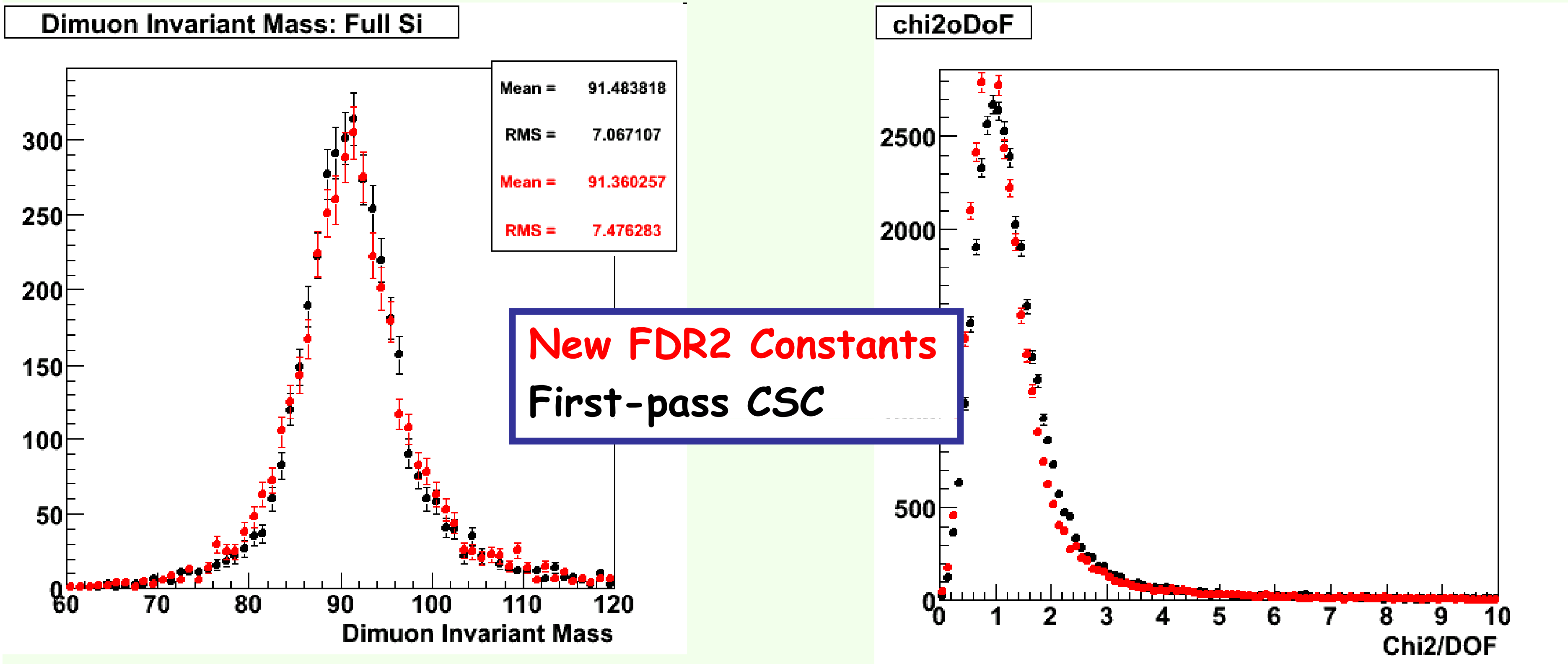






Recent Rerun of FDR2 Data

… finished just before the coffee break…



After the bug was fixed we are basically back to the CSC First-pass resolution

FDR2 Lessons and Conclusions

- The main aim of the exercise was achieved and the constants for the ID were produced in an automated way and delivered on time
- Access to the data on CAF worked much better than the Castor in FDR1
- Calibration stream and cosmics were used together in an automated way
- Lesson 0: Validate the code <u>before</u> using it "officially"
- Lesson 1: Optimized iterations/events to achieve the main goal: produce the constants on time (12-hours)
- Lesson2: Need to think if we want to use the "external" constraints (i.e. resonances) in the first pass
  - Can we use the ES in the CAF? We have time? How much can be parallelize? Enough statistic? It is really worth it?
- Lesson 3: No vertex constrain possible with the Alignment stream? Can this info to be added to the track? Or "flags" the tracks coming from the same vertex? Other ideas besides to use a different stream?

Preparation for Data Taking and Second Pass

- In recent months we were absorbed on the 24-hours operation and FDR2 exercise preparation
- It is time to think how to reach the ultimate alignment precision at the second pass reprocessing
- Main topics:
  - Systematic study of the weak modes
  - Usage of the “external” constraints
  - Help from the beam helo/beam gas tracks
  - Strategy on ultimate alignment including dataset
  - Determination and allocation of the resources outside CERN
- Answers cannot be done today of course, we need to discuss those points in details in the following months
  - Second pass reprocessing of the first data for the end 2008 ??