Alignment of the ATLAS Inner Detector – Run I Experience

(a quick journey to the world of geometry and linear algebra)

> Pawel Brückman de Renstrom Institute of Nuclear Physics P.A.N. Kraków

LPNHE seminar, 12 December 2013









Basic track fit (linearization) $\pi = (d, z, \varphi, \vartheta, Q/p_T), \quad \vec{e} \equiv \vec{e}(\pi)$ $\chi^2 = \mathbf{r}^T \mathbf{V}^{-1} \mathbf{r}, \quad r_i \equiv (\vec{e} - \vec{m}) \bullet \hat{k}$ Track fit 6 DoF $\mathbf{r}(\pi) = \mathbf{r}_0 + \frac{\partial \mathbf{r}}{\partial \pi} (\pi - \pi_0)$ linear exp. aro. seed $\frac{d\chi^2}{d\pi} = 0$ minimization condition $\vec{r}_i \equiv \vec{m}_i - \vec{e}_i(\pi, a)$

$$\pi - \pi_0 = \left(\frac{\partial \mathbf{r}}{\partial \pi}^T \mathbf{V}^{-1} \frac{\partial \mathbf{r}}{\partial \pi}\right)^{-1} \frac{\partial \mathbf{r}}{\partial \pi}^T \mathbf{V}^{-1} \mathbf{r}_0$$

LPNHE seminar, 12 Dec 2013

P. Brückman IFJ, PAN

Basic track fit

$$\pi = (d, z, \varphi, \vartheta, Q / p_T)$$

CAUTION: Lots of simplifications in the above. In reality at least two more effects need to be accounted for: A)Multiple Coulomb Scattering (track deflects at every intersected material) B)Energy Loss (particle losses energy for ionisation - changes momentum)

$$\pi - \pi_0 = \left(\frac{\partial \mathbf{r}}{\partial \pi}^T \mathbf{V}^{-1} \frac{\partial \mathbf{r}}{\partial \pi}\right)^{-1} \frac{\partial \mathbf{r}}{\partial \pi}^T \mathbf{V}^{-1} \mathbf{r}_0$$

LPNHE seminar, 12 Dec 2013

What do we need it for?

p7



Example of one of the first measurements issue of ATLAS tracking:

invariant mass,
primary event vertex.
K_s⁰ decay vertex,



P. Brückman IFJ, PAN

$H \rightarrow ZZ^* \rightarrow 4I$ candidate



P. Brückman IFJ, PAN

How does it work? *We cannot see trajectories, only their "footmarks". *Must "reconstruct" what happened.



How does it work?

We cannot see trajectories, only their "footmarks".
Must "reconstruct" what happened.

*Each "footmark" gets "captured" independently.
*In order to reconstruct a track one must complete the puzzle:



*Only one hipothesis is correct!

Alignment - what is it all about?

Parameters of charged particles reconstructed from space-point measurements:



Alignment is supposed to bring us from A to B!

P. Brückman IFJ, PAN

LPNHE seminar, 12 Dec 2013

When did alignment became relevant?

- 1. Detector is composed of more than one sensitive element,
- 2. Intrinsic resolution is better than the assembly precisision and mechanical stability.

Ad1: The Inner Detector of ATLAS consists of ~360,000 elements.

Ad2: Initial knowledge of their positions is one to two orders of magnitude worse than the resolution.



ATLAS Pixel detector









ATLAS Semiconductor Tracker (SCT)

4 concentric cylinders
2x9 discs in the forward region

ATLAS Transition Radiation Tracker (TRT)



Straws: - 350,000 proportional drift tubes, 4mm in diameter, arranged in * 96 barrel sectors * 2x20 end-cap wheels

ATLAS tracking system (ID)

silicon + gaseous devices



Pixels (Si pads):

- 1744 modules (10,464 par's)
- Pixel size: 50 μ m × 400 μ m
- Resolution : 10 μ m × 115 μ m

SCT (Si ministrip):

- 4088 modules (24,528 par's)
- Strip size: 80 µm × 12 cm
- Resolution: 17 μ m × 580 μ m

TRT (gas proportional):

- 350,048 straws (701,696 par's)
- Size: 4 mm × 71/39 cm
- Resolution: 130 µm

LPNHE seminar, 12 Dec 2013







Track-based alignment a closer look

Track fit and alignment

How to get from fitted tracks to alignment corrections? Two basic philosophies:
1. Local: after fitting tacks attempt is made to match detector positions accordingly (inherently iterative)



1. Global: a simultaneous optimization (fit) of both track parameters and detector element positions is performed

Idea of the Global χ^2 approach

$$\mathbf{r}(\pi) = \mathbf{r}_{\mathbf{0}} + \frac{\partial \mathbf{r}}{\partial \pi} (\pi - \pi_0) + \frac{\partial \mathbf{r}}{\partial a} (a - a_0) \qquad \frac{d\chi^2}{d\pi} = \frac{d\chi^2}{da} \mathbf{0}$$

simultaneous fit of all tracks and alignment parameters N+n*k pars! Impossible to solve !!! \otimes $\frac{d\mathbf{r}}{da}$ Way out!: Fold the track fit in. Solve for the alignment only: $\mathbf{r}(\pi) = \mathbf{r_0} + \left(\frac{\partial \mathbf{r}}{\partial \pi} \frac{d\pi}{da} + \frac{\partial \mathbf{r}}{\partial a}\right)(a - a_0)$

$$a - a_0 = \left(\sum_{\text{tracks}} \frac{d\mathbf{r}}{da}^T \mathbf{V}^{-1} \frac{d\mathbf{r}}{da}\right)^{-1} \sum_{\text{tracks}} \frac{d\mathbf{r}}{da}^T \mathbf{V}^{-1} \mathbf{r}_0$$

"Locality ansatz" in the Global χ^2 approach

Having fitted the track, one satisfies:

$$\mathbf{0} = \left(\frac{\partial \mathbf{r}}{\partial \pi}^T \mathbf{V}^{-1} \frac{\partial \mathbf{r}}{\partial \pi}\right)^{-1} \frac{\partial \mathbf{r}}{\partial \pi}^T \mathbf{V}^{-1} \mathbf{r}_0$$

 $\frac{d\chi^2}{d\pi} = \frac{d\chi^2}{da}0$

Most importantly, residuals not explicitly dependent on alignment parameters drop out. Only "actual" residuals survive:

$$a - a_0 = \left(\sum_{\text{tracks}} \frac{d\mathbf{r}}{da}^T \mathbf{V}^{-1} \frac{d\mathbf{r}}{da}\right)^{-1} \sum_{\text{tracks}} \frac{\partial \mathbf{r}}{\partial a}^T \mathbf{V}^{-1} \mathbf{r}_0$$

Idea of the Global χ^2 approach



Idea of the Local χ^2 approach

$$\mathbf{r}(\pi) = \mathbf{r}_{\mathbf{0}} + \frac{\partial \mathbf{r}}{\partial \pi} (\pi - \pi_0) + \frac{\partial \mathbf{r}}{\partial a} (a - a_0) \qquad \qquad \frac{d\chi^2}{da} \mathbf{0}$$

Fit of alignment parameters ignoring the correlations via tracks. Numerically a lot easier. Problem breaks down to local (n=6) equations . Requires multiple iterations over the full reconstruction !

$$\mathbf{r}(\pi) = \mathbf{r}_0 + \left(\frac{\partial \mathbf{r}}{\partial \pi} \frac{\partial \mathbf{r}}{\partial a} + \frac{\partial \mathbf{r}}{\partial a}\right)(a - a_0)$$

$$a - a_0 = \left(\sum_{\text{tracks}} \frac{\partial \mathbf{r}}{\partial a}^T \mathbf{V}^{-1} \frac{\partial \mathbf{r}}{\partial a}\right)^{-1} \sum_{\text{tracks}} \frac{\partial \mathbf{r}}{\partial a}^T \mathbf{V}^{-1} \mathbf{r}_0$$

Example: uniform expansion Let's try applying the two alignment philosophies...



Local approach (tracks are frozen within one alignment pass) Iteration 1:



Local approach (tracks are frozen within one alignment pass) Iteration 2:



Local approach (tracks are frozen within one alignment pass) Iteration 2:



Local approach (tracks are frozen within one alignment pass) Iteration 3:



Local approach (tracks are frozen within one alignment pass) Iteration 3:



Local approach (tracks are frozen within one alignment pass) Iteration 4:



Local approach (tracks are frozen within one alignment pass) Iteration 4:



Local approach (tracks are frozen within one alignment pass) Iteration 5:



Local approach (tracks are frozen within one alignment pass) **Iteration 5**:



p36

LPNHE seminar, 12 Dec 2013
Global vs local approach (a simple 1D cartoon)

Example: uniform expansion Now let's see what the global approach does...



Global vs local approach (a simple 1D cartoon)

Global approach (tracks allowed to refit within one alignment pass - all correlations retained!) Iteration 1:



Global vs local approach (a simple 1D cartoon)

Global approach (tracks allowed to refit within one alignment pass - all correlations retained!) Iteration 1:



Things to remember:

Local and Global approaches should asymptotically give the same result.

Local methods are not numerically demanding. It may take them a lot of iterations to converge, though.

Larger the system is more beneficial it is to use the global approach (potentially slower convergence of local methods for large systems) but become numerically challenging.

Where is the difficulty?

Two very different aspects of the same alignment task:

Efficiency of track reconstruction Good track fit

Track reconstruction free from systematics Fine track parameter resolution Quality vertexing "easy"

highly nontrivial !!!

Matrix M generally singular (at least ill-conditioned). Needs special treatment. Formally, the most elegant – diagonalization.



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 may directly lead to biases on physics (systematic effects).

LPNHE seminar, 12 Dec 2013



Example: cosmic alignment with Global χ^2

• Possible trap: Do not try to exploit all apparent information:



• Alignment quality the same for -1500!!!

LPNHE seminar, 12 Dec 2013





Singular (and weak) modes need to be removed from the solution. What can we do for very large systems? soft-mode-cut ... and use fast solvers) $\mathbf{M}X = Y$, $\mathbf{U}\mathbf{M}\mathbf{U}^{\mathrm{T}}\mathbf{U}X = \mathbf{U}Y \Longrightarrow \mathbf{D}X_{D} = Y_{D}$ $X = \sum_{i} E_{i} \frac{1}{\lambda} Y_{D}$ $\mathbf{M} \rightarrow \mathbf{M} + \kappa \mathbf{l}, \quad \mathbf{D} \rightarrow \mathbf{D} + \kappa \mathbf{l}, \quad \lambda_i \rightarrow \lambda_i + \kappa$ 10^{5} **Eigen-spectrum** 10^{-3} 10^{2} 10 10 10 10 10 1000 2000 \cap 3000 4000 5000 6000 LPNHE seminar. 12 Dec 2013 P. Brückman IFJ p46



Singular (and weak) modes need to be removed from the solution. What can we do for very large systems? soft-mode-cut (...and use fast solvers)

In practice, it is useful to normalize the soft-cuts to the expected uncertainty. The cut-off can be tuned to each DoF individually. This is actually used in ATLAS.

$$\mathcal{M}_{ij} = \frac{d\chi^2}{d\frac{\alpha}{\sigma_i} d\frac{\alpha}{\sigma_j}}, \qquad \mathcal{M}_{ij} \longrightarrow \sigma(\alpha_i) \sigma(\alpha_j) \mathcal{M}_{ij}$$

$$(\sigma(\alpha_i)\sigma(\alpha_j)\mathcal{M}_{ij}+1)\frac{X_j}{\sigma(\alpha_j)} = \sigma(\alpha_i)Y_i \implies (\mathcal{M}_{ij}+\mathcal{D}(\frac{1}{\sigma(\alpha_i)^2}))X_j = Y_i$$

Here, $\mathcal{D}(\frac{1}{\sigma(\alpha_i)^2})$ denotes a diagonal matrix with $\frac{1}{\sigma(\alpha_i)^2}$ elements on the diagonal. In summary, the scaled soft mode cut is realized by adding the diagonal matrix \mathcal{D} to the original matrix \mathcal{M} .

LPNHE seminar, 12 Dec 2013

Solving the alignment problem

In the most general case diagonalisation is the approach:
Allows to control statistical significance of individual modes
Full covariance matrix readily available
Memory-demanding
Time consuming (~N³)
Can be used for problems < O(10,000)</p>

Sparse problems (usually the case) can be tackled using fast solvers (Gaussian elimination - MA27, Numerical norm minimization GMRES)):

Much faster (<N²) and less memory-demanding
 Require preconditioning to remove weak modes
 No direct error control - indirectly using soft-cuts

Local method does not present any numerical challenge (except for large number of iterations).

LPNHE seminar, 12 Dec 2013

To remember:

Global approach results in a large system of linear equations which (usually) correlates all DoF's of the system.

□ It is challenging to solve but provides the optimal answer in (quasi) single go.

One must take special care not to introduce artificial deformations due to wrong statistical treatment!

Strategy adopted by ATLAS:

The procedure consists of alignment at different levels of granularity - baginning with large structures and getting down to individual sensitive elements.
 Heavily rely on the Beam Spot constraint.

The Global method is used for systems not exceeding the size of Pixel +SCT+TRT@L2 ~35,000 DoF's (diagonalization or "fast solver") TRT straw-level alignment (currently 2 DoF's per straw => ~700,000 parametes) uses the Local method.

Level 1 (4 alignable structures) Pixel Barrel SCTECC SCT Barrel **4**SCT Layers Level 2 (31 alignable structures) **3 Pixels Discs 3 Pixels Layer** SCT barrel modules Level 3 (5832 alignable structures) Pixels module

Distortions considered so far

Charge-antisymmetric momentum bias (sagitta):

$$q/p_{\rm T} \longrightarrow q/p_{\rm T} + \delta_{\rm sagitta}$$
 or $p_{\rm T} \longrightarrow p_{\rm T}(1 + qp_{\rm T} \, \delta_{\rm sagitta})^{-1}$.
 $p \longrightarrow p(1 + qp_{\rm T} \, \delta_{\rm sagitta})^{-1}$.

Charge-symmetric momentum bias (radial):

 $\begin{aligned} r &\longrightarrow (1 + \epsilon_{\text{radial}}(\phi, \eta))r. \\ p_{\text{T}} &\longrightarrow p_{\text{T}}(1 + 2\epsilon_{\text{radial}}) & \text{for small } \epsilon_{\text{radial}}. \\ p_{\text{Z}} &\longrightarrow p_{\text{Z}}(1 + \epsilon_{\text{radial}}). \end{aligned}$

Bias on the Impact Parameter (XY or Z)

 $t \longrightarrow t + \delta d_0,$ $\Delta \phi[\text{rad}] = \frac{\delta d_0}{r},$

LPNHE seminar, 12 Dec 2013







Track parameter constraints in GX

* The generic expression for an arbitrary constraint on fitted track parameters takes the form:

track fit:

$$\delta \pi = -J^{-1} \left(\mathbf{E}^T V^{-1} r(\pi_0, a) + S^{-1} (\pi_0 - x) \right)$$
$$J \equiv \mathbf{E}^T V^{-1} \mathbf{E} + S^{-1}$$

alignment fit:

$$\delta a = -\mathcal{M}^{-1}$$
$$\sum_{tracks} \left(\frac{\partial r^T}{\partial a_0} W r_0 - \frac{\partial r^T}{\partial a_0} V^{-1} E J^{-1} S^{-1} (\pi_0 - x) \right)$$

 M is the second derivative matrix constructed in the standard way but using the constrained covariance matrix of tracks (J).
 The "locality ansatz" preserves the constraint because:

$$0 = J^{-1} \Big[EV^{-1}r + S^{-1}(\pi_0 - x) \Big]$$

$$\delta a = -\mathbf{M}^{-1} \sum \frac{\partial r}{\partial a} \Big[(V^{-1} - V^{-1}EJ^{-1}EV^{-1})r - V^{-1}EJ^{-1}S^{-1}(\pi_0 - x) \Big]$$

LPNHE seminar, 12 Dec 2013

* "Full" vertex constraint. The idea of the vertex constraint in the GX sense is to leave its position fully free in the fit:

$$\frac{dr}{da} = \frac{\partial r}{\partial a} + E\frac{d\pi}{da} + F\frac{db}{da}$$

$$\frac{d\pi}{da} = -J^{-1}E^{T}V^{-1}\left(\frac{\partial r}{\partial a} + F\frac{db}{da}\right)$$

$$\frac{db}{da} = -\left(\sum_{tracks}^{ev} F^{T}WF\right)^{-1}\left(\sum_{tracks}^{ev} F^{T}W\frac{\partial r}{\partial a}\right) \quad (\text{vertex refit})$$

Track parameters are re-parameterrised!
The final expression is analogous to the "baseline one" but the covariance matrix of the measurements takes a new (more complex) form:

$$\delta a = -\underbrace{\left(\sum_{tracks} \frac{\partial r^{T}}{\partial a_{0}} X \frac{\partial r}{\partial a_{0}}\right)^{-1}}_{\mathcal{M}} \underbrace{\sum_{tracks} \frac{\partial r^{T}}{\partial a_{0}} X r(\pi_{0}, b_{0}, a_{0})}_{\mathcal{V}} X \equiv W - WFM_{b}^{-1} \left(\sum_{tracks}^{ev} F^{T}W\right)$$

Run I timeline

→ 2009:

→ Alignment infrastructure validation

→ Alignment with cosmics & 900GeV run

→ 2010 (~50pb⁻¹):

- → Dedicated "ID calibration" stream and new alignment code
- → Alignment with 7TeV collision data
- → Wire to wire TRT alignment
- → Use of pixel module deformations

→ 2011(~5fb⁻¹):

→ Implementation of alignment at Tier-0 calibration loop

- \rightarrow ID to B-field alignment
- → Detected time dependent movements. Run-by-run alignment monitoring.
- → Use of E/P constraint, $Z \rightarrow \mu\mu$ as cross check
- → New Pixel clustering (NN) introduced

→ 2012 (~20fb⁻¹):

- → Use of Z → $\mu\mu$ constraint, E/P as cross check
- → Use of Impact Parameters biases as constraints
- → Advanced alignment split in periods
- → Detailed studies of momentum scale
- → Update the alignment constants (if needed) for the first T0 reprocessing

P. Brückman IFJ, PAN

LPNHE seminar, 12 Dec 2013

The first alignment using the cosmic-ray muons (2009)







ATLAS (CO) (CO)

Reconstruction of kinematic parameters



Observation of clear side bowing of pixel staves (collision data 2009)



FJ, PAN

2010 ($\sqrt{s}=7$ TeV)

o 2010 was the first year of performing the alignment under 'real' working conditions Careful understanding of detector behaviour!

<u>×</u>10³

100

80

60

40

20

Hits on tracks / 12 μm

LPNHE se



2010 (√s=7 TeV)



LPNHE seminar, 12 Dec 2013

ID-BF tilt: what to expect?

➤ Radial distortions give: $p_T \rightarrow p_T(1+\delta)$ - charge symmetric Δm/m ~ δ [p_T^2 (1-cos(Δφ))]/m² ~ δ



 $R = R_0 + \sin \varphi^* \alpha^* z, \qquad z = R_0^* \cot \theta$ $R = R_0 (1 + \sin \varphi^* \alpha^* \cot \theta)$

Which mimicks exactly the "radial" deformation proportional to $\cot(\theta)$.





B-field rotation fit

	ECC	ECA
K _s	0.67 +/- 0.02	0.53 +/- 0.02
J/Psi (cal. data)	0.48 +/- 0.04	0.52 +/- 0.04
J/Psi (cal. MC)	0.52 +/- 0.03	0.58 +/- 0.03

Rotate the B-field by +0.55 mrad around X



2011 (√s=7 TeV)



2011 (Vs=7 TeV)





Run-by-run L1 alignment done at Tier 0 allows to monitor any gross movemets of ID structures.

LPNHE seminar, 12 Dec 2013

2012 (1/s=8 TeV)



Run-by-run L1 alignment can trigger quasi real-time geometry update for the TierO bulk reconstruction.

LPNHE seminar, 12 Dec 2013

2012 (√s=8 TeV)



Impact of Level 1 re-alignment on sagitta distortions

P. Brückman IFJ, PAN

2012 (1/s=8 TeV)



Correcting the Impact Parameter biases using constrained alignment.

P. Brückman IFJ, PAN

LPNHE seminar, 12 Dec 2013

2012 ($\sqrt{s=8 \text{ TeV}}$) Summary of observed systematics

p70





Summary of the residual misalignment (with any outstanding detector effects entangled)

Detector	Pull σ before	Pull σ after	C (µm)
Pixel Barrel X	1.18	1.00	4.5
Pixel Barrel Y	1.13	1.00	33.5
Pixel End-cap X	1.06	1.00	3.6
Pixel End-cap Y	1.00	0.99	0
SCT Barrel	1.08	1.00	8.1
SCT End-cap	1.06	1.00	8.7
TRT Barrel	0.96	0.96	0.0
TRT End-cap	0.96	0.96	0.0

2015 - the Outlook

Integration of the IBL - will be decisive for Beam Spot, vertexing, etc.

Improve on run-by-run procedure. Possibly extend to Level 2 alignment.

Understand better the remaining sagitta and overall momentum scale bias - finetune.

Out-of-the-plane deformations may be included in the alignment procedure.

The Isertable B Layer (IBL)

14 staves of mixed technology will be installed directly on the reduced beam pipe.

Average radius 3.3 cm, 64 cm long.
 Each stave consists of 12 planar pixel

modules and 8 3D sensors.

 \Box CO₂ - cooling system

□ low radiation lenth ~1.9% X₀ together with carbon-fibre support.

Will have to be aligned as a separate entity, due to mechanical independence time dependence important!




Conclusions

Alignment is an indispensable element of modern experiments but potentially hazardous. LHC Run I allowed to gather vast experience with the detector and achieve final performance exceeding original requirements. Good quality track fit was the easy part of the game (although involved solving linear systems with O(10-100)k parameters. Most of the effort was spent on understanding and eliminating systematic deformations. Run II will see integration of another subsystem (IBL), and further finetuning.

Thank you!

Documentation

ATL-INDET-PUB-2005-002, ATL-INDET-PUB-2007-009 ATLAS-CONF-2011-012, ATLAS-CONF-2012-141, ATL-COM-SOFT-2012-006 (going public imminently) ATL-COM-INDET-2013-033 (going public imminently)



ATLAS NOTE July 27, 2012



Common Framework Implementation for the Track-Based Alignment of the ATLAS Detector

P. Brückman de Renstrom¹, R. Harrington² S-C. Hsu³, S. Kaiser⁴, D. Kóllar⁵, V. Lacuesta Miquel⁶, S. Martí i García⁷, A. Morley⁸, J. Schieck⁹, J. Wang¹⁰

LPNHE seminar, 12 Dec 2013

p75

P. Brückman IFJ, PAN