Track-based Alignment using a Kalman Filter Technique

R. Frühwirth and E. Widl

Institute for High-Energy Physics, Vienna, Austria

Outline

- Introduction
- Sequential updating
- Implementation and computational complexity
- Two-track fitter
- Examples
- Summary and outlook

Introduction

- Global alignment without having to solve a large system of linear equations
- Recursive approach, based on Kalman filter equations
- Update after each track
- Update is not restricted to modules crossed by the track
- Update is limited to modules with significant correlations
- Some bookkeeping required

Introduction

- Easy to use prior information form mechanical or laser alignment
- Easy to fix the position of reference detectors
- Method suitable for alignment relative to another detector
- Stil in the experimental phase

Sequential Updating

- Notation for alignment related objects:

N	total number of alignable detector units (modules)
$\boldsymbol{d}_{\mathrm{t}}$	vector of true alignment parameters
$\boldsymbol{d}_0$	expansion point of alignment parameters
$\boldsymbol{d}$	current estimate of alignment parameters
$\boldsymbol{d}_i$	subvector of alignment parameters of detector unit i
$\boldsymbol{D}$	covariance matrix of $\boldsymbol{d}$
$\boldsymbol{D}_{ij}$	submatrix of cross-correlations between detector units i and j
$\widehat{\boldsymbol{d}}$	updated estimate of alignment parameters
$\widehat{\boldsymbol{D}}$	covariance matrix of $\widehat{\boldsymbol{d}}$

Sequential Updating

- Notation for track related objects:

I	list of modules crossed by the current track
k	size of I
$\boldsymbol{m}$	observations of the current track
$\boldsymbol{V}$	covariance matrix of $\boldsymbol{m}$
$\boldsymbol{x}_{\mathrm{t}}$	true track parameters of the current track
$\boldsymbol{x}_0$	expansion point of track model
$\boldsymbol{x}$	predicted track parameters of the current track
$\boldsymbol{C}$	covariance matrix of $\boldsymbol{x}$
$\widehat{\boldsymbol{x}}$	updated track parameters of the current track

Sequential Updating

- The observations $\boldsymbol{m}$ depend on the track parameters $\boldsymbol{x}_{\rm t}$ via the track model f:

$$\boldsymbol{m} = \boldsymbol{f}(\boldsymbol{x}_{\rm t}) + \boldsymbol{\varepsilon}, \quad \mathsf{cov}(\boldsymbol{\varepsilon}) = \boldsymbol{V}$$

- $\boldsymbol{\varepsilon}$ contains the effects of the observation error and of multiple scattering. Energy loss is considered as deterministic and is included in the track model.
- Its variance-covariance matrix $\boldsymbol{V}$ can be assumed to be known.
- A preliminary track fit gives a provisional estimate $\boldsymbol{x}$ of the track parameters. The actual estimates of the alignment parameters are used at this stage.

Sequential Updating

- The observations also depend on the alignment parameters $\boldsymbol{d}_{\mathrm{t}}$.
- The track model is extended accordingly:

$$\boldsymbol{m} = \boldsymbol{f}(\boldsymbol{x}_{\mathrm{t}}, \boldsymbol{d}_{\mathrm{t}}) + \boldsymbol{\varepsilon}, \quad \mathsf{cov}(\boldsymbol{\varepsilon}) = \boldsymbol{V}$$

- First-order Taylor expansion at expansion points $\boldsymbol{d}_0$ and $\boldsymbol{x}_0$:

$$\boldsymbol{m} = \boldsymbol{c} + \boldsymbol{A}\boldsymbol{d}_{\mathrm{t}} + \boldsymbol{B}\boldsymbol{x}_{\mathrm{t}} + \boldsymbol{\varepsilon} = \boldsymbol{c} + \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \end{pmatrix} \begin{pmatrix} \boldsymbol{d}_{\mathrm{t}} \\ \boldsymbol{x}_{\mathrm{t}} \end{pmatrix} + \boldsymbol{\varepsilon}$$

- Jacobians:

$$\boldsymbol{A} = \partial \boldsymbol{m} / \partial \boldsymbol{d}_{\mathrm{t}} \big|_{\boldsymbol{d}_0}, \quad \boldsymbol{B} = \partial \boldsymbol{m} / \partial \boldsymbol{x}_{\mathrm{t}} \big|_{\boldsymbol{x}_0}$$

Sequential Updating

- Constant:

$$\boldsymbol{c} = f(\boldsymbol{x}_0, \boldsymbol{d}_0) - \boldsymbol{A}\boldsymbol{d}_0 - \boldsymbol{B}\boldsymbol{x}_0$$

- Expansion point $\boldsymbol{d}_0$: the nominal or the currently estimated module alignment.
- Expansion point $\boldsymbol{x}_0$: result of a preliminary track fit.
- The Kalman filter requires a prediction $\boldsymbol{x}$ of the track parameters, along with its variance-covariance matrix $\boldsymbol{C}$.
- The prediction has to be stochastically independent of the observations in the track.

Sequential Updating

- First case: independent prediction exists
  - External prediction from already aligned detector
  - External information form vertex or kinematical constraint
- Update equation of the Kalman filter:

$$\begin{pmatrix} \widehat{\boldsymbol{d}} \\ \widehat{\boldsymbol{x}} \end{pmatrix} = \begin{pmatrix} \boldsymbol{d} \\ \boldsymbol{x} \end{pmatrix} + \boldsymbol{K} \left(\boldsymbol{m} - \boldsymbol{c} - \boldsymbol{A}\boldsymbol{d} - \boldsymbol{B}\boldsymbol{x} \right)$$

Sequential Updating

- Gain matrix of the filter:

$$\boldsymbol{K} = \begin{pmatrix} \boldsymbol{D} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{C} \end{pmatrix} \begin{pmatrix} \boldsymbol{A}^T \\ \boldsymbol{B}^T \end{pmatrix} \underbrace{\left(\boldsymbol{V} + \boldsymbol{A}\boldsymbol{D}\boldsymbol{A}^T + \boldsymbol{B}\boldsymbol{C}\boldsymbol{B}^T \right)^{-1}}_{\boldsymbol{G}}$$

$$= \begin{pmatrix} \boldsymbol{D}\boldsymbol{A}^T\boldsymbol{G} \\ \boldsymbol{C}\boldsymbol{B}^T\boldsymbol{G} \end{pmatrix}$$

- Update decomposes:

$$\widehat{\boldsymbol{d}} = \boldsymbol{d} + \boldsymbol{D}\boldsymbol{A}^T\boldsymbol{G}\,(\boldsymbol{m} - \boldsymbol{c} - \boldsymbol{A}\boldsymbol{d} - \boldsymbol{B}\boldsymbol{x})$$

$$\widehat{\boldsymbol{x}} = \boldsymbol{x} + \boldsymbol{C}\boldsymbol{B}^T\boldsymbol{G}\,(\boldsymbol{m} - \boldsymbol{c} - \boldsymbol{A}\boldsymbol{d} - \boldsymbol{B}\boldsymbol{x})$$

Sequential Updating

- Case two: no independent prediction exists
- The prediction $\boldsymbol{x}_0$ gets zero weight in order not to bias the estimation
- Multiply $\boldsymbol{C}$ by a scale factor α and let α tend to infinity:

$$\begin{aligned}\boldsymbol{G} &= \lim_{\alpha \longrightarrow \infty} \left(\boldsymbol{V} + \boldsymbol{A}\boldsymbol{D}\boldsymbol{A}^T + \alpha \boldsymbol{B}\boldsymbol{C}\boldsymbol{B}^T \right)^{-1} \\ &= \boldsymbol{V}_D^{-1} - \boldsymbol{V}_D^{-1}\boldsymbol{B}(\boldsymbol{B}^T\boldsymbol{V}_D^{-1}\boldsymbol{B})^{-1}\boldsymbol{B}^T\boldsymbol{V}_D^{-1}\end{aligned}$$

with

$$\boldsymbol{V}_D = \boldsymbol{V} + \boldsymbol{A}\boldsymbol{D}\boldsymbol{A}^T$$

Sequential Updating

- Because of $\boldsymbol{GB} = \mathbf{0}$ the update equation of the alignment parameters can be simplified to

$$\widehat{\boldsymbol{d}} = \boldsymbol{d} + \boldsymbol{D}\boldsymbol{A}^T\boldsymbol{G}\,(\boldsymbol{m} - \boldsymbol{c} - \boldsymbol{A}\boldsymbol{d})$$

- The update of the covariance matrix can be calculated by linear error propagation:

$$\widehat{\boldsymbol{D}} = (\boldsymbol{I} - \boldsymbol{D}\boldsymbol{A}^T\boldsymbol{G}\boldsymbol{A})\,\boldsymbol{D}\,(\boldsymbol{I} - \boldsymbol{A}^T\boldsymbol{G}\boldsymbol{A}\boldsymbol{D}) + \boldsymbol{D}\boldsymbol{A}^T\boldsymbol{G}\boldsymbol{V}\boldsymbol{G}\boldsymbol{A}\boldsymbol{D}$$

- Both terms on the right hand side are positive definite, so the left hand side is guaranteed to be positive definite as well.

Sequential Updating

- The iteration needs starting values.
- Mechanical and laser alignment can be used for the starting values.
- Reference units can be fixed by giving them very small initial errors.

Implementation and computational complexity

- The current track crosses k detector units.
- The dimension $n = 2\,k$ of $\boldsymbol{m}$ is small, in the order of 30 for the CMS Inner Tracker.
- The matrix $\boldsymbol{B}$ is of size $n \times 5$ and is therefore small.
- The matrix $\boldsymbol{A}$ is a row of N blocks $\boldsymbol{A}_i$ of size $n \times m$, where m is the number of alignment parameters per detector unit (usually 6).
- For each track, only k out of these N blocks are different from zero:

$$\boldsymbol{A} = \big(\mathbf{0} \ldots \mathbf{0}\ \boldsymbol{A}_{i_1}\ \mathbf{0} \ldots \mathbf{0}\ \boldsymbol{A}_{i_2}\ \mathbf{0} \ldots\ldots \mathbf{0}\ \boldsymbol{A}_{i_k}\ \mathbf{0} \ldots \mathbf{0}\big)$$

Implementation and computational complexity

- The only large matrix in the update formulas is the product $\boldsymbol{D}\boldsymbol{A}^T$.
- It is a column of N blocks each of which has size $m \times n$.
- Complete computation of $\boldsymbol{D}\boldsymbol{A}^T$ would lead to an algorithm that scales with N^2.
- This is too slow for practical purposes.

Implementation and computational complexity

- Two alternatives:
  - Algorithm A: Compute only the blocks of the modules in the current track, neglecting all correlations.
  - Algorithm B: Compute the blocks of the modules having significant correlations with the modules in the current track.
- Algorithm A gives an unbiased estimate, but is suboptimal because of the missing correlations.
- Algorithm B is nearly optimal, but it has to be guaranteed that $\widehat{\boldsymbol{D}}$ is positive definite all the time. This problem is being studied, but there is not yet a foolproof solution.

Implementation and computational complexity

- Tradeoff between speed and precision.
- In order to keep track of the necessary updates, a list L_i is attached to each detector unit i, containing the detector units that have significant correlations with i.
- This list may contain only i itself in the beginning and grows as more tracks are processed.
- If there is prior knowledge about correlations, for instance because of mechanical constraints, it can be incorporated in the list and in the initial covariance matrix.

Implementation and computational complexity

- Update of the alignment parameters

1. Update the list L_i for every $i \in I$ (see below).
2. Form the list L of all detector units that are correlated with the ones crossed by the current track: $L = \bigcup_{i \in I} L_i$. The size of L should be much smaller than N.
3. For all $j \in L$ compute: $(\boldsymbol{D}\boldsymbol{A}^T)_j = \sum_{i \in I} \boldsymbol{D}_{ji}\boldsymbol{A}_i^T$. Each block $\boldsymbol{D}_{ji}$ is of size $m \times m$.
4. Compute: $\boldsymbol{A}\boldsymbol{D}\boldsymbol{A}^T = \sum_{i \in I} \boldsymbol{A}_i(\boldsymbol{D}\boldsymbol{A}^T)_i$.
5. Compute: $\boldsymbol{V}_D = \boldsymbol{V} + \boldsymbol{A}\boldsymbol{D}\boldsymbol{A}^T$ and $\boldsymbol{G}$. All matrices involved are of size $n \times n$.
6. Compute: $\boldsymbol{m}' = \boldsymbol{G}\left(\boldsymbol{m} - \boldsymbol{c} - \sum_{i \in I} \boldsymbol{A}_i\boldsymbol{d}_i\right)$.
7. For all $j \in L$ compute: $\widehat{\boldsymbol{d}}_j = \boldsymbol{d}_j + (\boldsymbol{D}\boldsymbol{A}^T)_j\boldsymbol{m}'$.

Implementation and computational complexity

- Update of the covariance matrix

For all $j, l \in L$ compute: $\widehat{\boldsymbol{D}}_{jl} = \boldsymbol{D}_{jl} + (\boldsymbol{D}\boldsymbol{A}^T)_j(\boldsymbol{G}\boldsymbol{V}_D\boldsymbol{G} - 2\boldsymbol{G})[(\boldsymbol{D}\boldsymbol{A}^T)_l]^T$

- The computational complexity of the parameter update is of the order $|L| \cdot |I|$.
- The computational complexity of the update of the covariance matrix is of the order $|L|^2$.
- Restricting the size of the lists L_i is of crucial importance.

Implementation and computational complexity

- Current proposal for building the lists L_i is based on the concept of a distance between two modules i and j.
- Define the following relation:

$$i \sim j \iff i \text{ and } j \text{ have been crossed by the same track}$$

- The relation "$\sim$" is symmetric, but not transitive.
- Define the distance $d(i, j)$ between detector units i and j by:

1. $d(i, i) = 0$
2. If $i \neq j$ and $i \sim i_1 \sim i_2 \sim \cdots \sim i_n \sim j$ is the shortest chain connecting i to j, the distance is $d(i, j) = n + 1$.

Implementation and computational complexity

- $i \sim j \iff d(i,j) = 1$
- d is a proper metrics:

> 1. $d(i,j) = 0$ if and only of $i = j$
> 2. $d(i,j) = d(j,i)$
> 3. $d(i,j) \leq d(i,k) + d(k,j)$ for all k

- The list L_i contains all modules k with $d(k,i) \leq d_{\max}$.

Implementation and computational complexity

- Alternative approaches conceivable:
  - Let the lists grow until the correlations have stabilized. Then, drop all correlations below an upper limit. Dynamic, adapts to the track sample used.
  - Determine "optimal" correlation structure from simulated data. Static, has to be done separately for every potential track sample (cosmics, beam halo, interactions).
- Detailed studies required.

Two-track fitter

- Use vertex- and mass-constrained track pairs to improve momentum resolution
- Examples: $Z \longrightarrow \mu^+\mu^-$, $J/\psi \longrightarrow \mu^+\mu^-$
- The five track parameters are replaced by nine decay parameters:
  - the position v of the decay vertex
  - the momentum p of the mother particle in the lab frame
  - the two decay angles (θ, φ) in the rest frame of the mother particle
  - the mass m of the mother particle

Two-track fitter

- The mass is constrained by adding a virtual observation of the mass (theoretical value plus width).
- The Jacobians w.r.t. the decay parameters are obtained by the chain rule:

$$\frac{\partial \boldsymbol{m}_i}{\partial(\boldsymbol{v}, \boldsymbol{p}, \theta, \varphi)} = \frac{\partial \boldsymbol{m}_i}{\partial \boldsymbol{x}_i} \cdot \frac{\partial \boldsymbol{x}_i}{\partial(\boldsymbol{v}, \boldsymbol{p}_i)} \cdot \frac{\partial(\boldsymbol{v}, \boldsymbol{p}_i)}{\partial(\boldsymbol{p}, \theta, \varphi)}$$

- Otherwise the formalism remains unchanged.

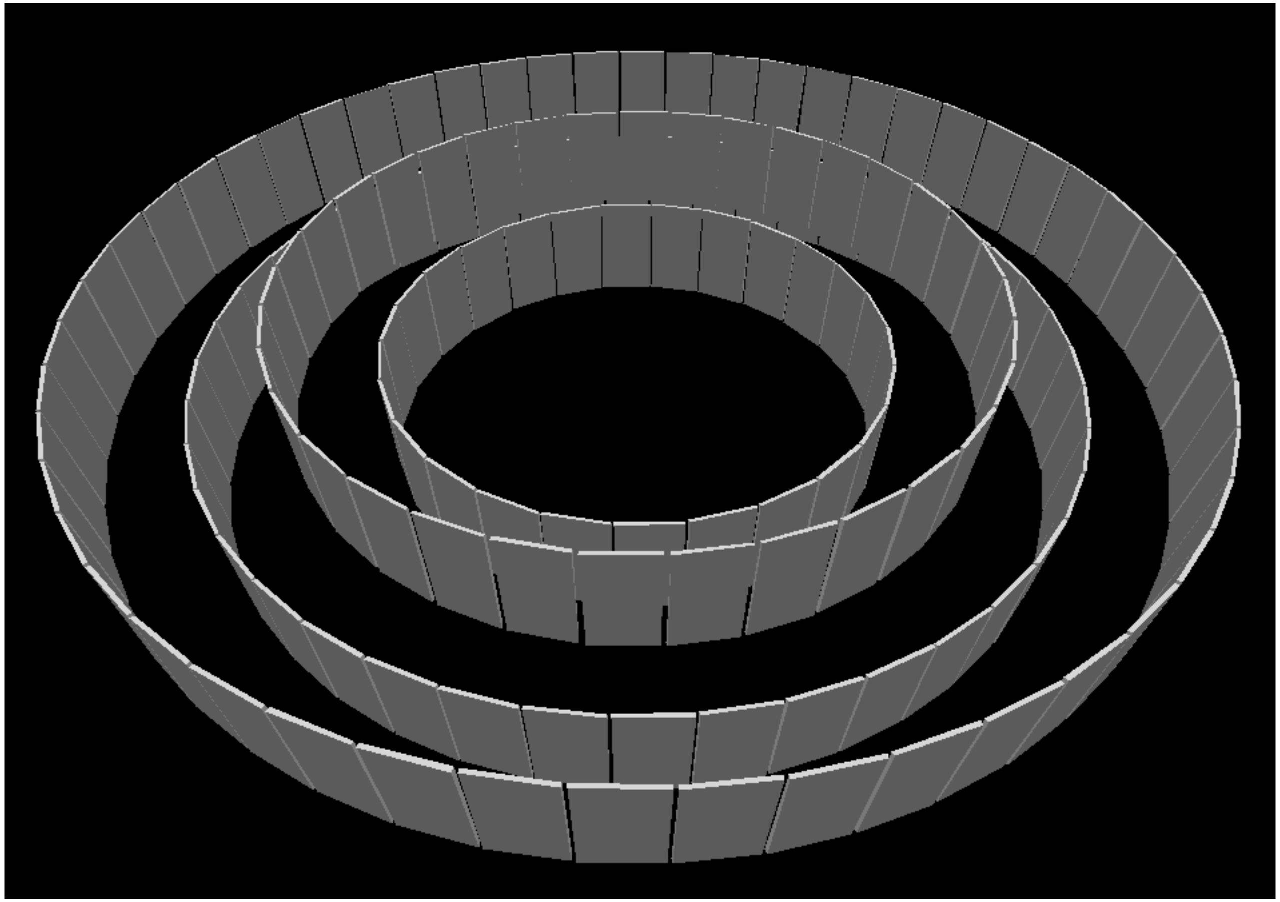
Examples

- Two setups, subsets of CMS Tracker
  - “Small wheel”: 3 Pixel layers (180 modules), 4 TIB layers (156 modules)
  - “Large wheel”: 3 Pixel layers (180 modules), 4 TIB layers (156 modules), 6 TOB layers (344 modules)
- Alignment of TIB and TOB relative to Pixel barrel

Examples

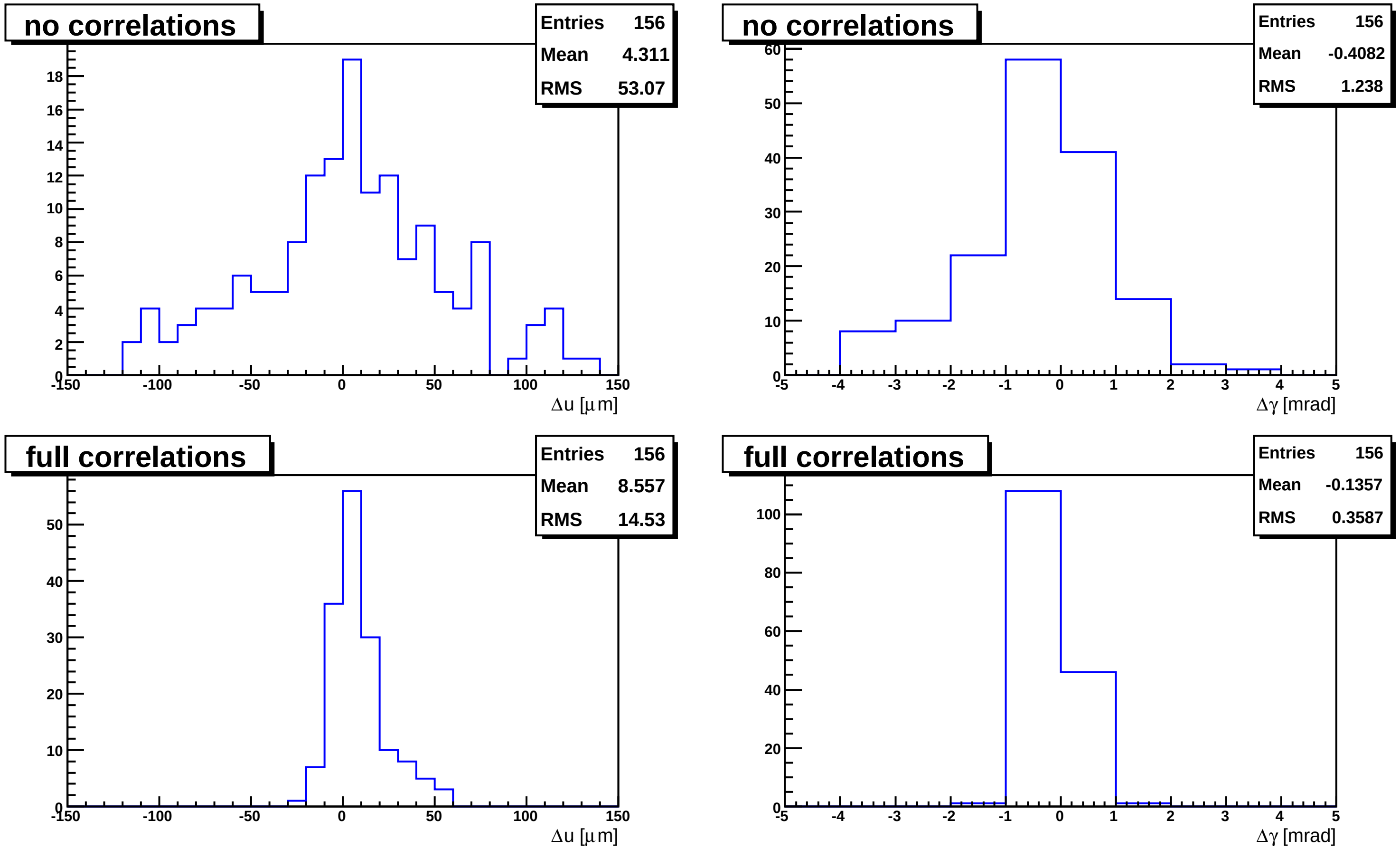
- Small wheel
- pixels are fixed, 156 TIB modules are aligned
- Misalignment:
  $\sigma(\Delta u) = 100\,\mu\text{m}, \sigma(\Delta v) = 100\,\mu\text{m}, \sigma(\Delta\gamma) = 5\,\text{mrad}$
- 25000 muon pairs from $Z \longrightarrow \mu^+\mu^-$ (50000 tracks)
  - No correlations
  - Full correlations

Examples



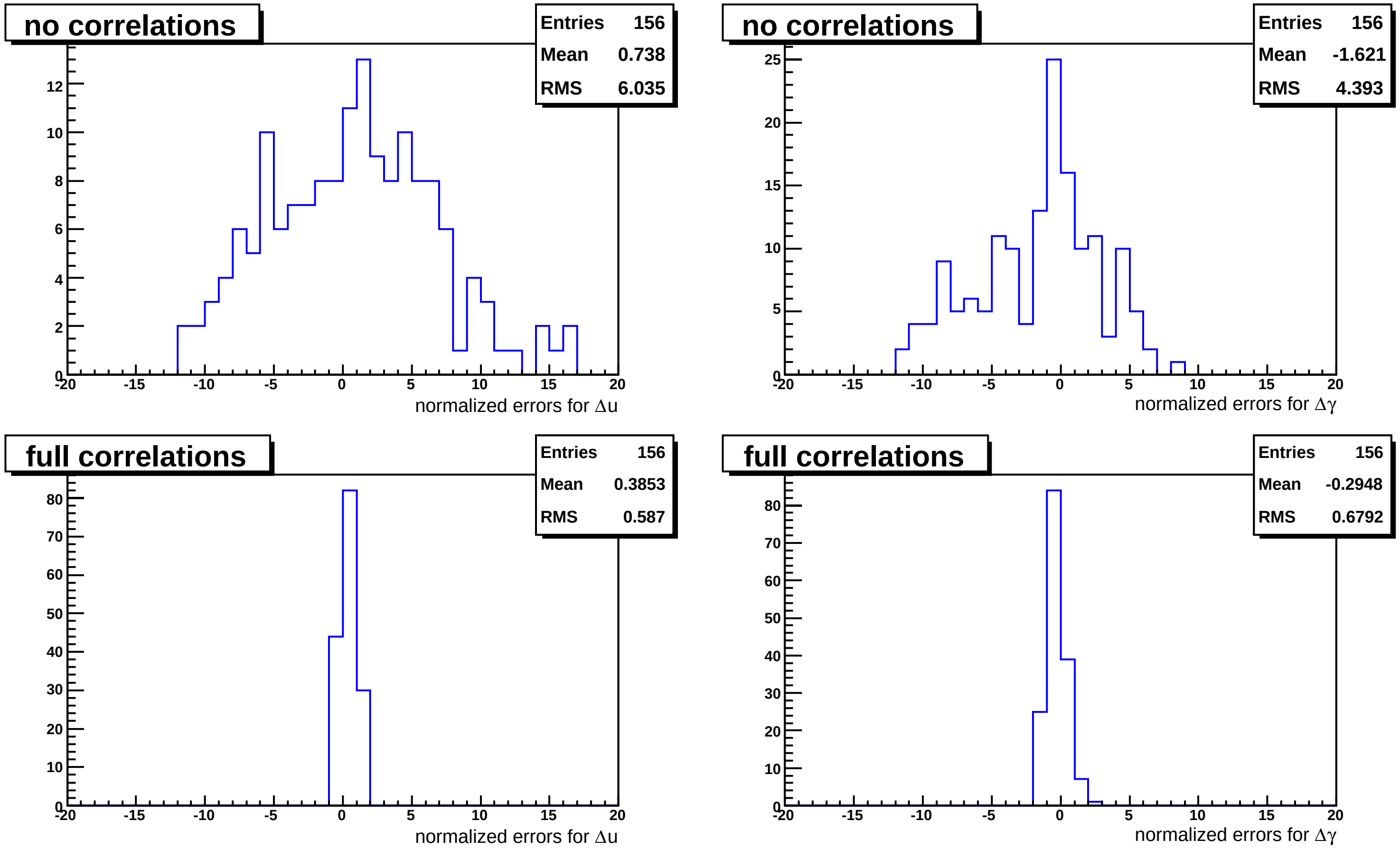
Small wheel with 4 TIB layers

Examples



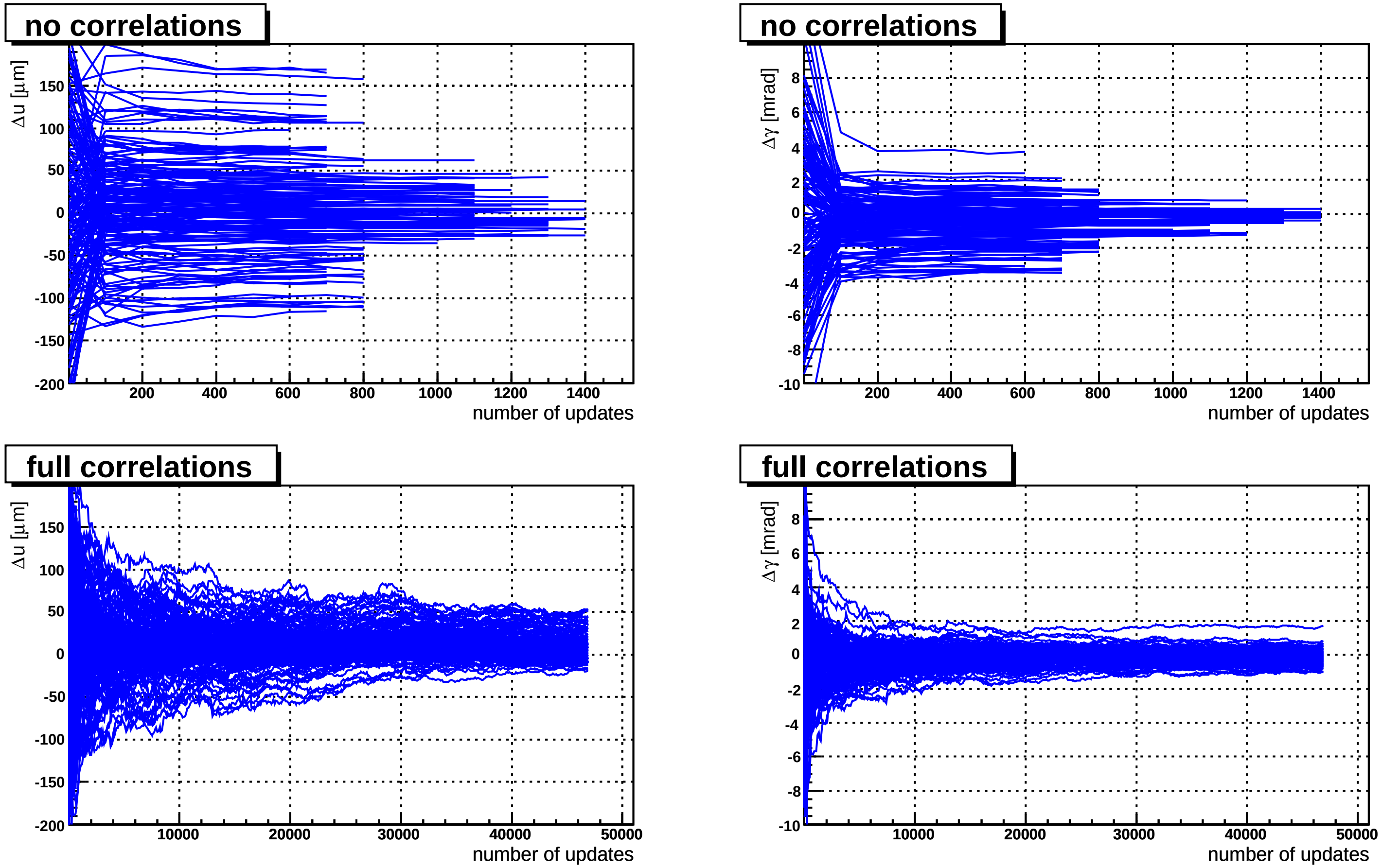
Residuals after alignment

Examples



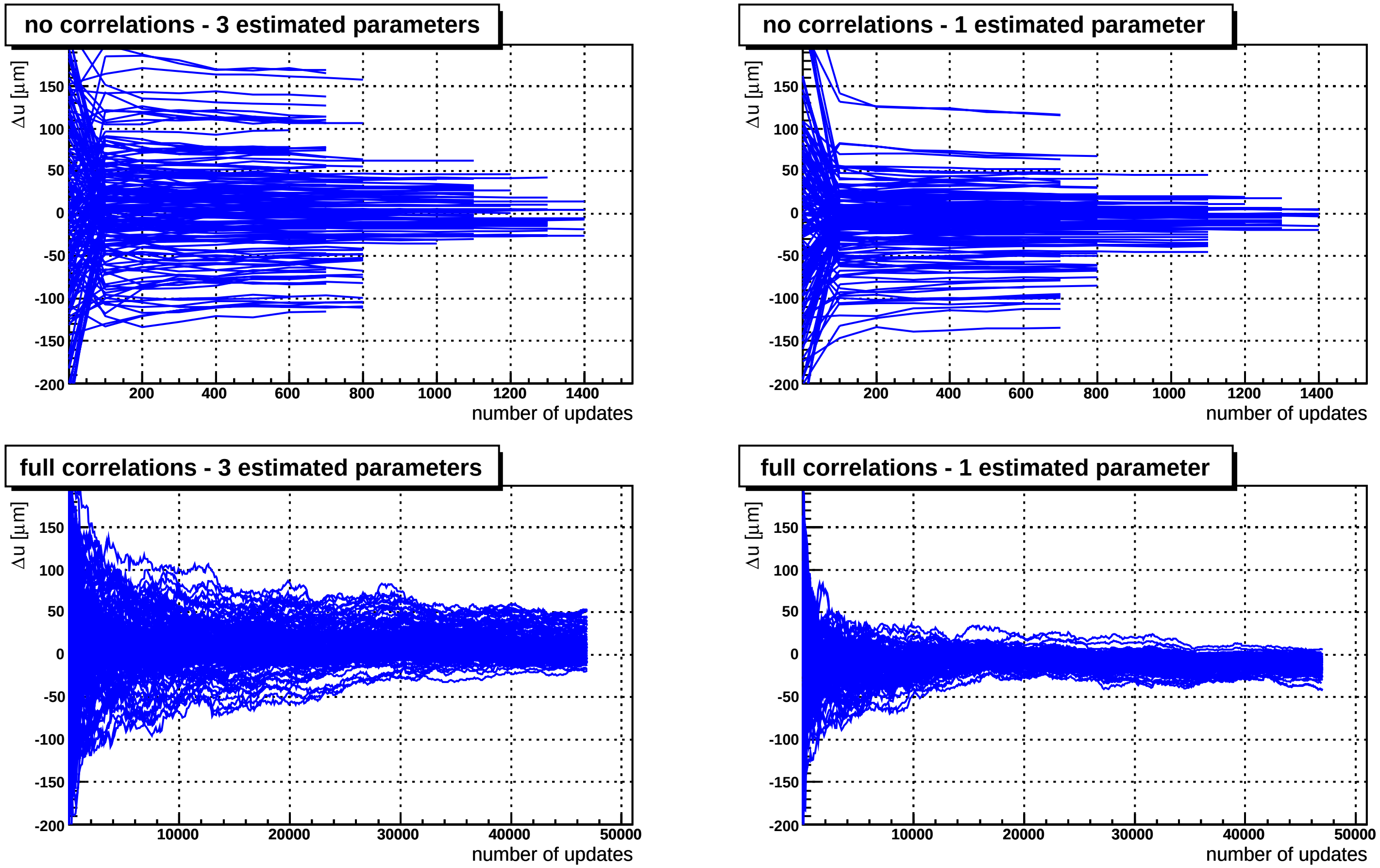
Standardized residuals after alignment

Examples



Evolution of residuals

Examples

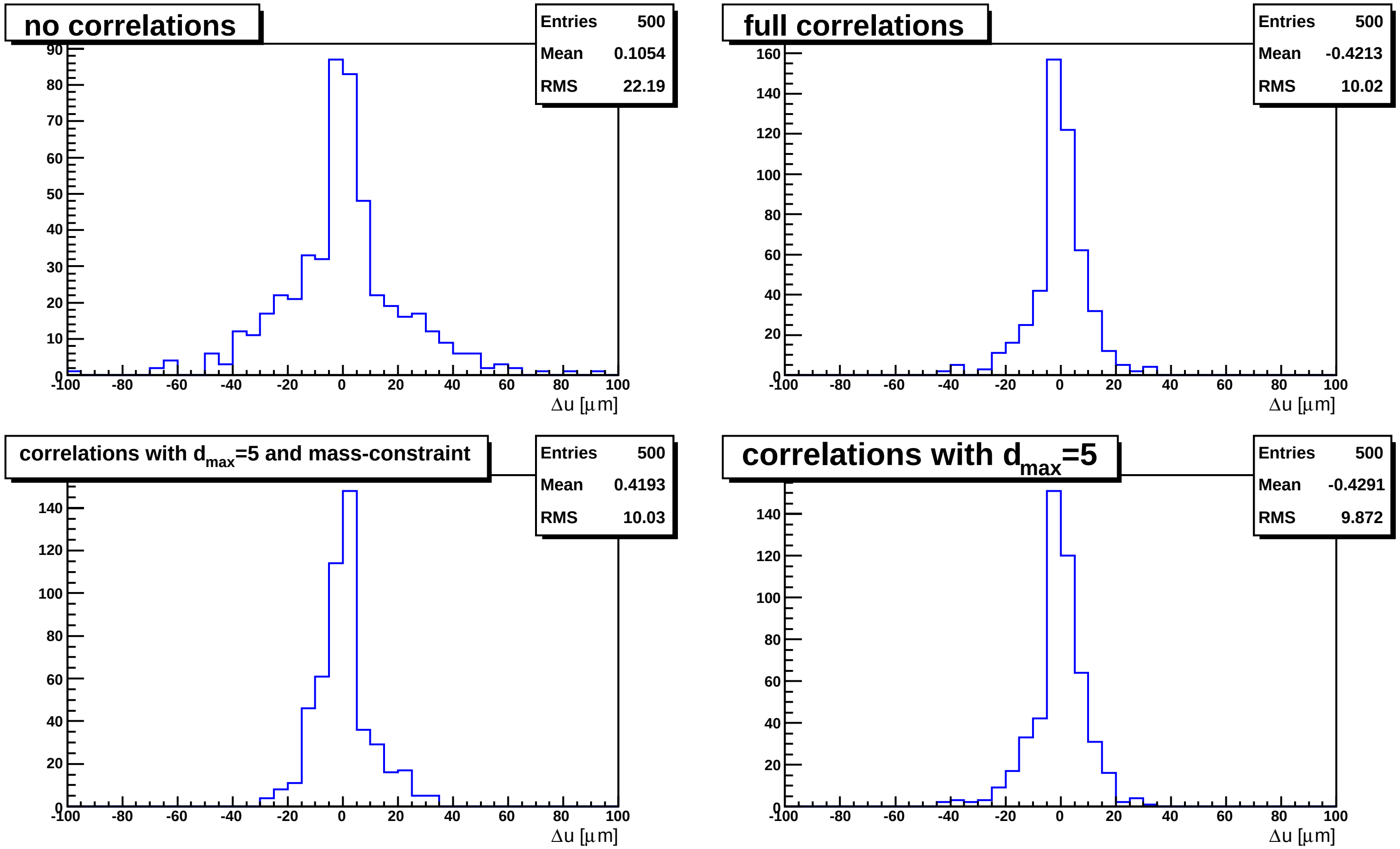


Evolution of residuals (1 or 3 alignment parameters)

Examples

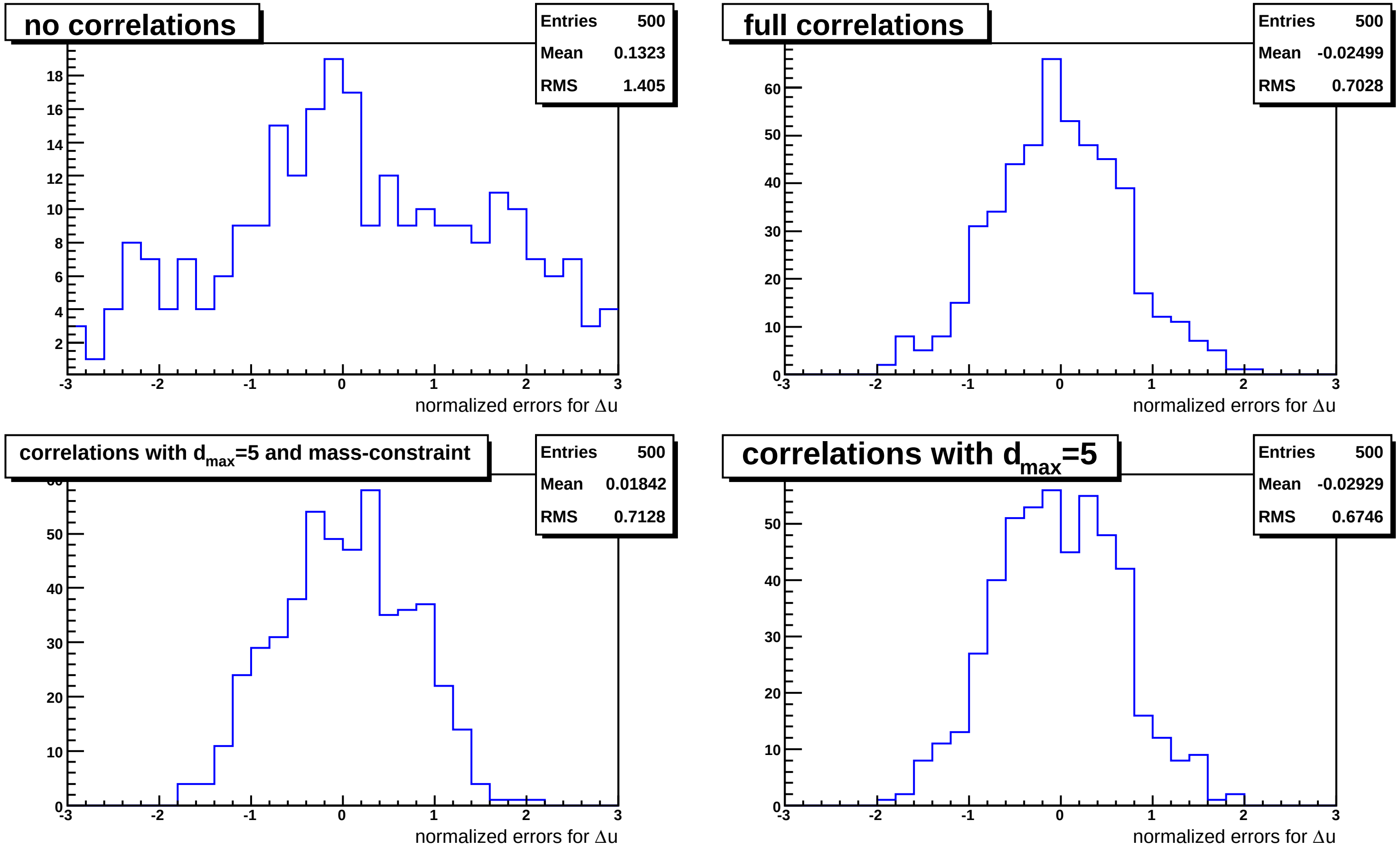
- Large wheel
- pixels are fixed, 500 TIB/TOB modules are aligned
- Misalignment:
  $\sigma(\Delta u) = 100\,\mu\text{m}, \sigma(\Delta v) = 100\,\mu\text{m}$
- 25000 muon pairs from $Z \longrightarrow \mu^+\mu^-$ (50000 tracks)
  - No correlations
  - Full correlations
  - Correlations up to $d_{\max} = 5$
  - Two-track fitter with correlations up to $d_{\max} = 5$

Examples



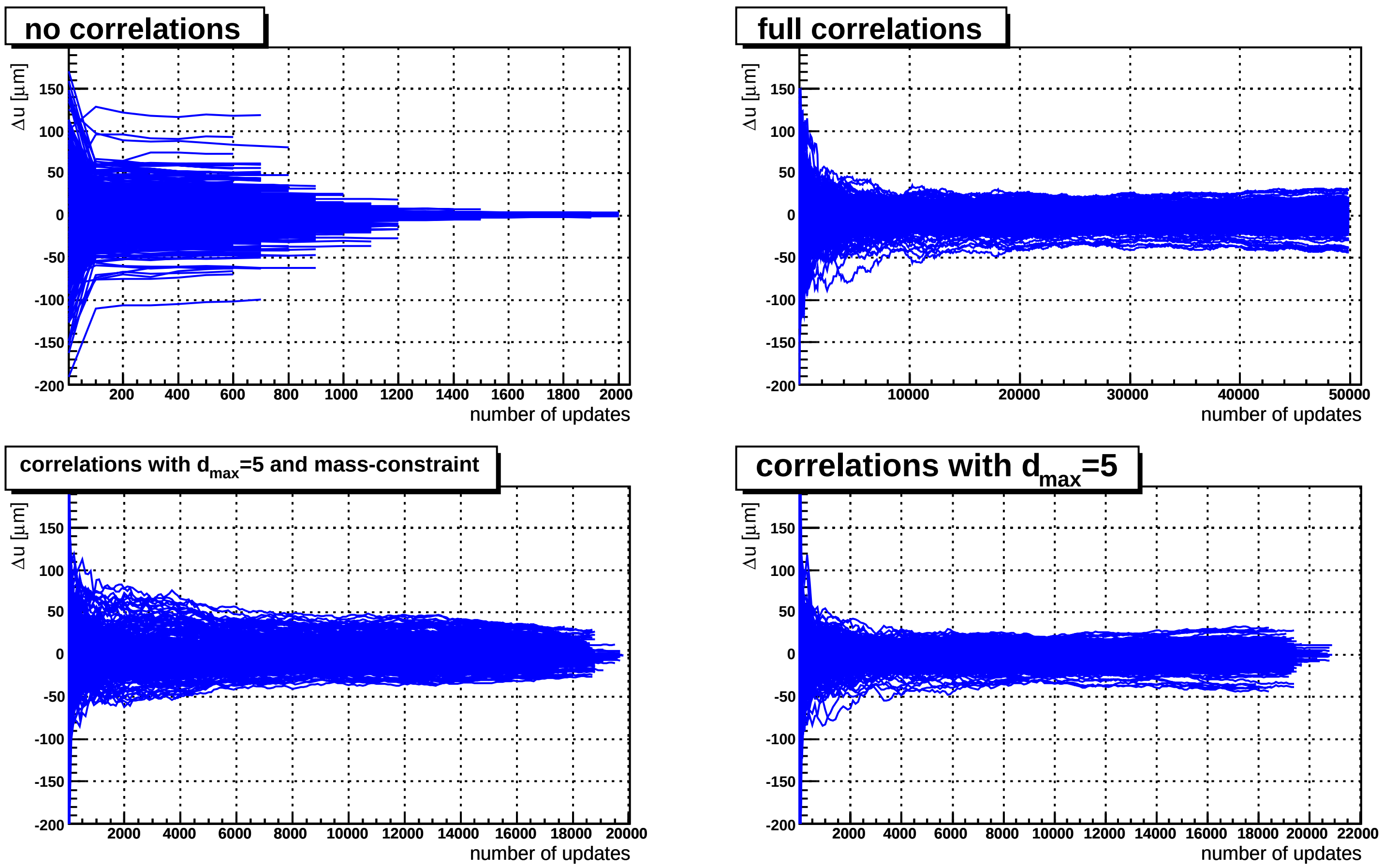
Residuals after alignment

Examples



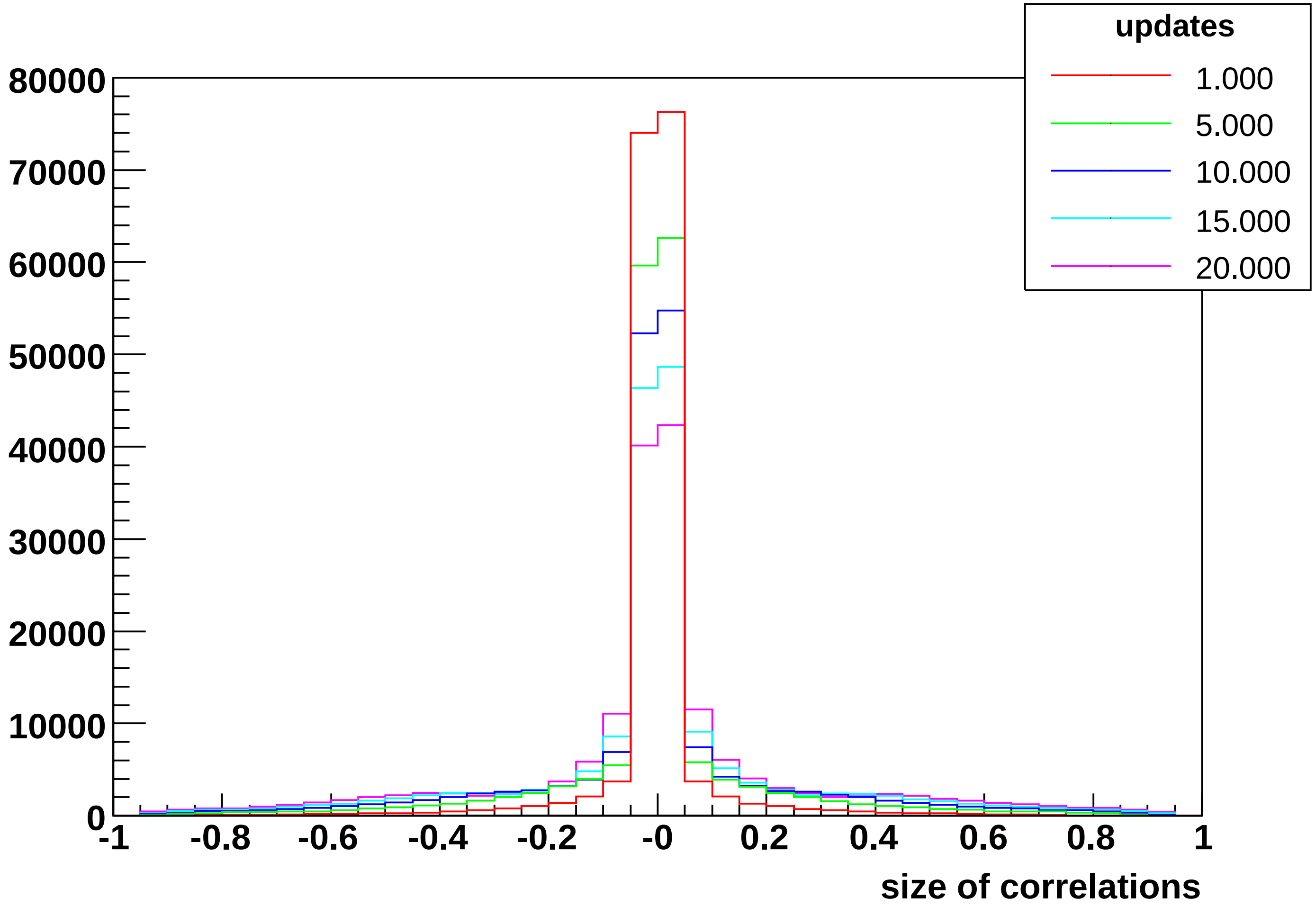
Standardized residuals after alignment

Examples



Evolution of residuals

Examples



Evolution of correlations

Examples

- Approximate total processing times
  - No correlations: 1.5 h
  - Full correlations: 39.5 h
  - Correlations up to $d_{\max} = 5$: 7.5 h
  - Two-track fitter with correlations up to $d_{\max} = 5$: 16 h
- Can be speeded up considerably by using pixel tracks as prediction (no update of pixel modules)

Summary and Outlook

- Kalman filter for sequential estimation of alignment constants
- Successful test on small-scale setups
- Advantages
  - No solution of large systems of equations
  - Depth of correlations can be taylored to setup
  - Errors of estimated alignment constants are always available
  - Can be used for stopping criterion

Summary and Outlook

- Disadvantages
  - Larger computational expense per track
  - More bookkeeping required
- Outlook
  - Extend to full set of angles and shifts
  - Study alternative approaches to correlation lists
  - Speed optimization
  - Large-scale examples

Acknowledgements

- We thank Wolfgang Adam (HEPHY Vienna) for technical support.
- We thank CERN for financial support to E. Widl.