A Gaussian-sum Filter for vertex reconstruction

T. Speer - University of Zurich
R. Frühwirth - HEPHY Vienna

Computing in High Energy Physics
Interlaken
30^{th} September 2004

Vertex reconstruction

- Standard tool for vertex reconstruction is the Kalman Filter (also implemented in the reconstruction software of the CMS experiment at LHC, CERN)
- The Kalman Filter is mathematically equivalent to a global least square minimization (LSM)
- If the model is linear and random noise is Gaussian:
  - ➢ LS estimators are unbiased and have minimum variance
  - ➢ Residuals and pulls of estimated quantities are also Gaussian
- For non-linear models or non-Gaussian noise, it is still the optimal linear estimator
- Non-Gaussian measurement errors degrade results!

The Gaussian-sum Filter

- Gaussian-sum Filter (GSF)

  Measurement error distributions modelled by mixture of Gaussians:
  - Main component of the mixture would describe the core of the distribution
  - Tails would be described by one or several additional Gaussians.
- First proposed by R. Frühwirth for track reconstruction

  (Computer Physics Communications 100 (1997) 1.)
- Successfully implemented in the CMS reconstruction software for electron track reconstruction:
  - Bethe-Heitler energy loss distribution modeled by a mixture of Gaussians
- GSF for vertex reconstruction now also implemented in the CMS reconstruction software.

The Gaussian-sum Filter for vertex reconstruction

- Track parameter error distributions modeled by a mixture of Gaussians
- Vertex State vector x , is also distributed according to a mixture of Gaussians
- Iterative procedure: estimate of the vertex is updated with one track at the time
- Add new track to vertex, each component of the Vertex State is updated with each component of the track (Combinatorial combination of all track components)
- The filter is a weighted sum of several Kalman Filters
  - GSF is implemented as a number of Kalman filters run in parallel
  - The weights of the components are calculated separately
  - Non-linear estimator: weights depend on the measurements
- The new Vertex State x_k is therefore distributed according to a mixtures of N_k ($= N_{track - k} * N_{vertex - k-1}$) Gaussians
  - The number of components increases exponentially (combinatorial explosion)
  - The GSF vertex filter shows little sensitivity to the number of components kept

Simulation

Simplified simulation in a fully controlled environment:

- Tracks generated at a common vertex
- No track reconstruction
- Track parameters are smeared according to known distributions:

  2 component Gaussian mixture:

  · Narrow component: 90 % Relative weight

  (Standard deviation of Impact parameter = 100μm)

  · Wide component: 10 % Relative weight

  Std dev. 10x larger (Impact parameter = 1000μm)

  ➻ Ratios of Standard deviation = 10

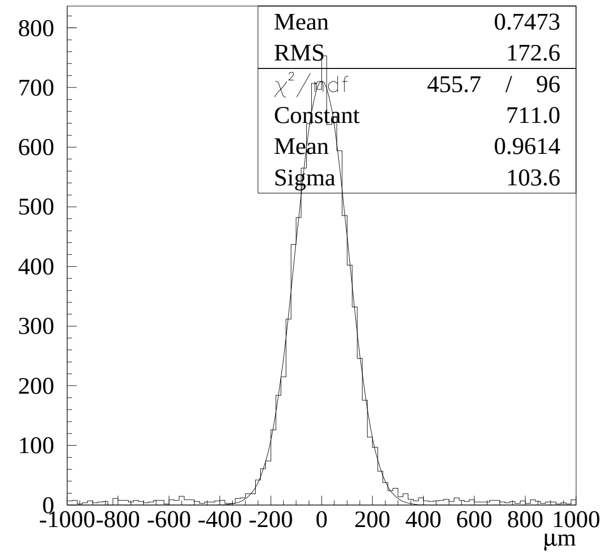
- For the Kalman Filter:

  · tracks smeared according to two-component mixture

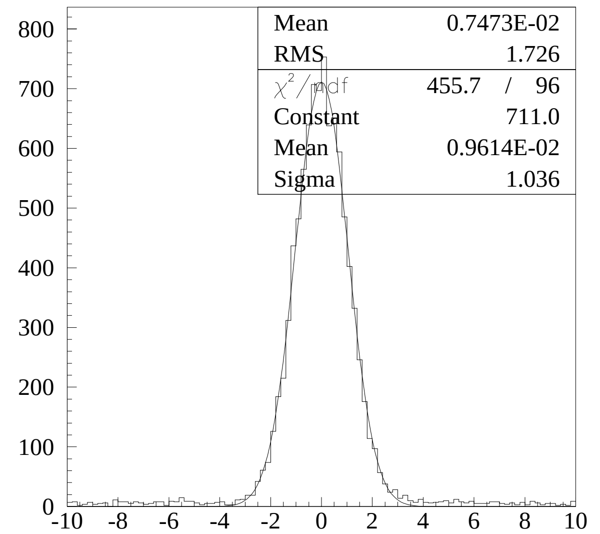
  · single component used in the fit:

  ➢ track parameter variance of dominating component

  ➢ estimated position independent of scaling of variance (but not position uncertainty or χ^2)



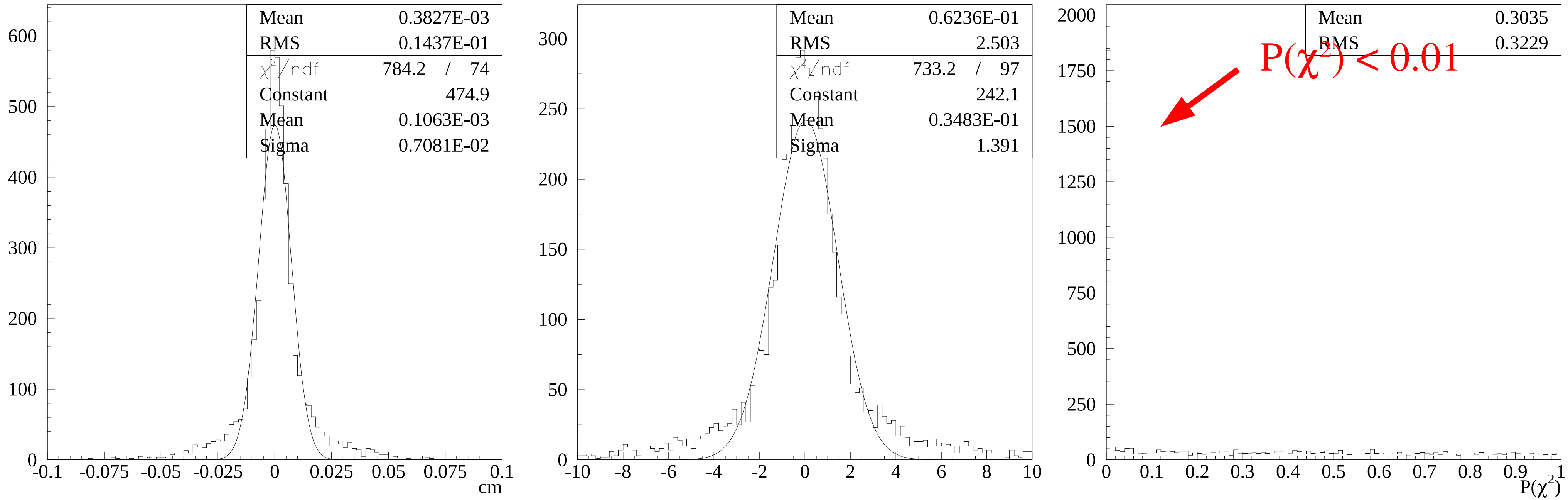
Transverse IP - residuals (μm)



Transverse IP - pulls for Kalman Filter

Kalman Filter fit

Four track-vertex fit with the Kalman Filter:



Residuals – y coord. (cm)

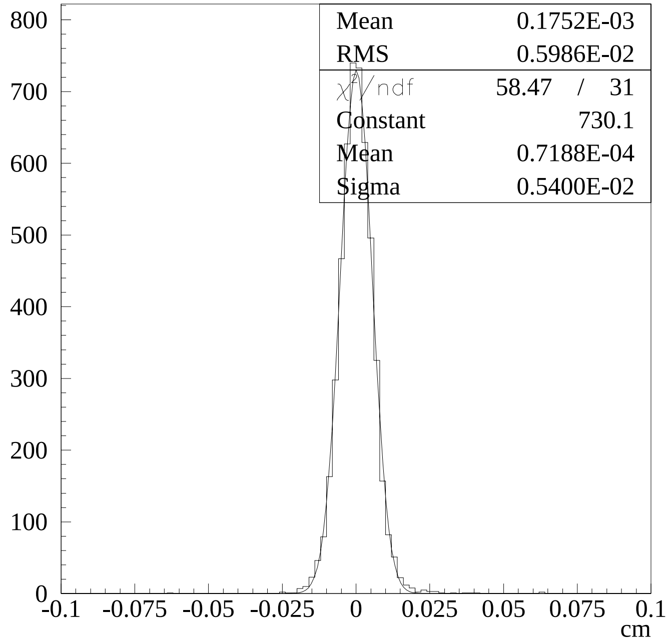
Pull – y coord.

$P(\chi^2)$

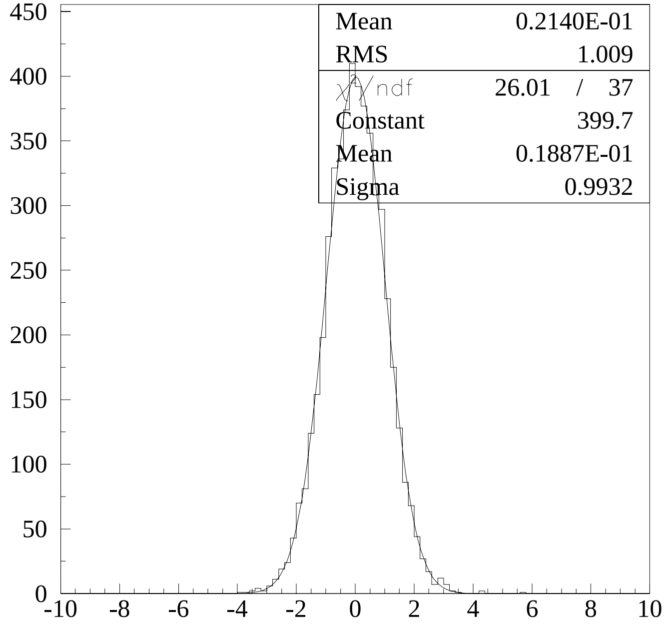
- Non-Gaussian tails in the distributions of residuals and pulls
- Large number of fits with $P(\chi^2) < 0.01$

Gaussian-sum Filter fit

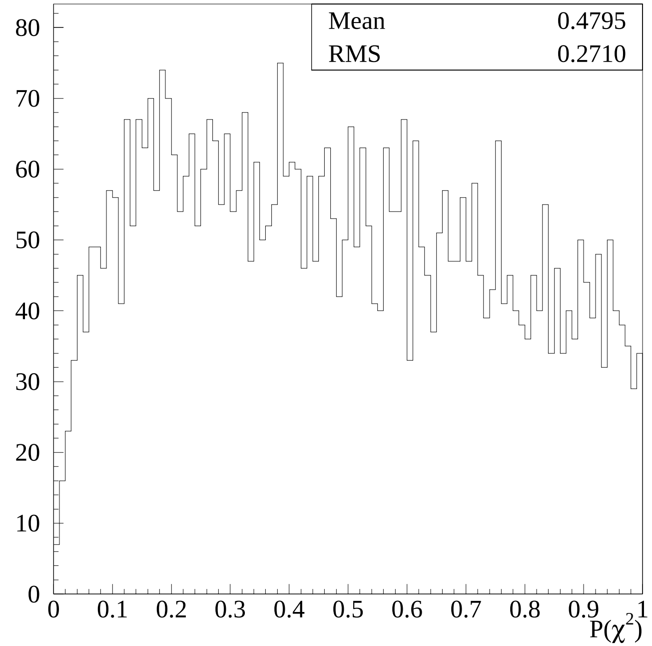
Four track-vertex fit with the GSF (using the full Gaussian mixture)



Residuals – *y* coord. (cm)



Pull – *y* coord.



$P(\chi^2)$

Residuals: smaller tails then with the Kalman Filter, smaller resolution

The remaining tails are due to events with several outliers.

No outliers in the pull distributions: error on the outliers correctly taken into account

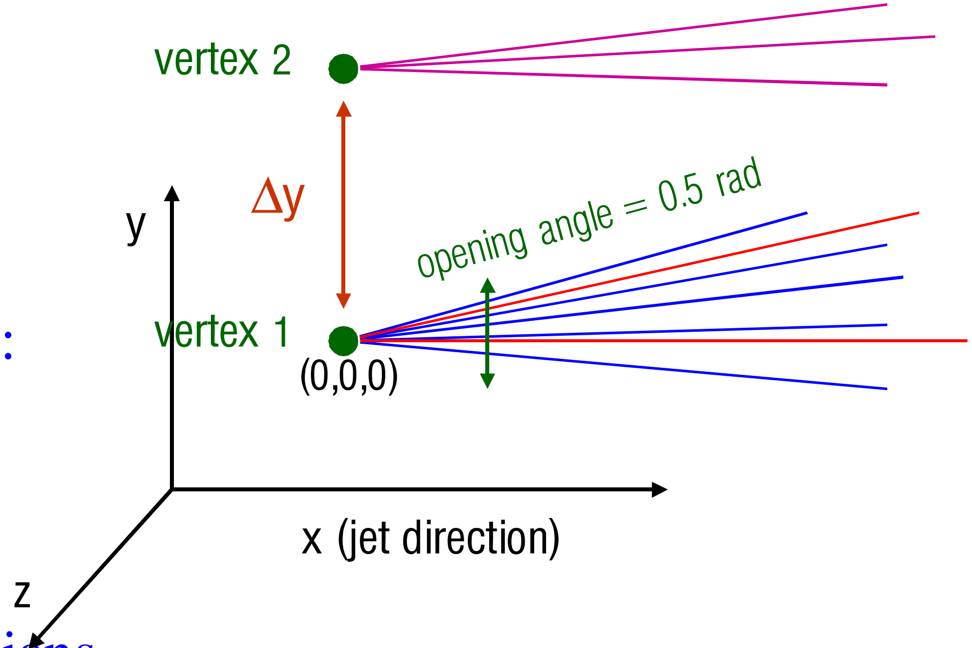
$P(\chi^2)$: dip at 0. - in early stages of the fit, bias towards components with a low χ^2

The filters needs several iterations (tracks) to stabilise and select the correct vertex component (combination of track components)

Robustness tests

- LS estimators:
  - Optimal when no outliers are present
  - Very sensitive to outliers
- What about the GSF?

- Type 1 outliers = mismeasured tracks:
  - $\sigma_{\text{outliers}}/\sigma_{\text{inliers}} \neq 1$
- Type 2 outliers = track from another vertex:
  - 4 tracks from main vertex
  - 1 track from second vertex

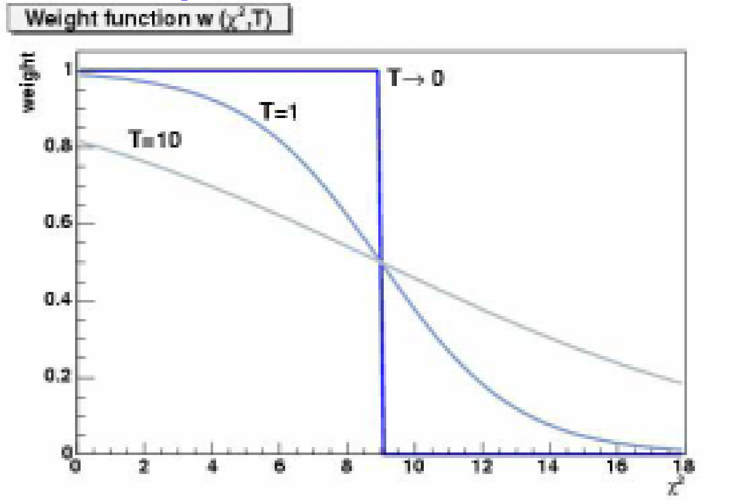


Tracks smeared with same mixture of 2 Gaussians

Robust filters: Adaptive Vertex fitter (AVF)

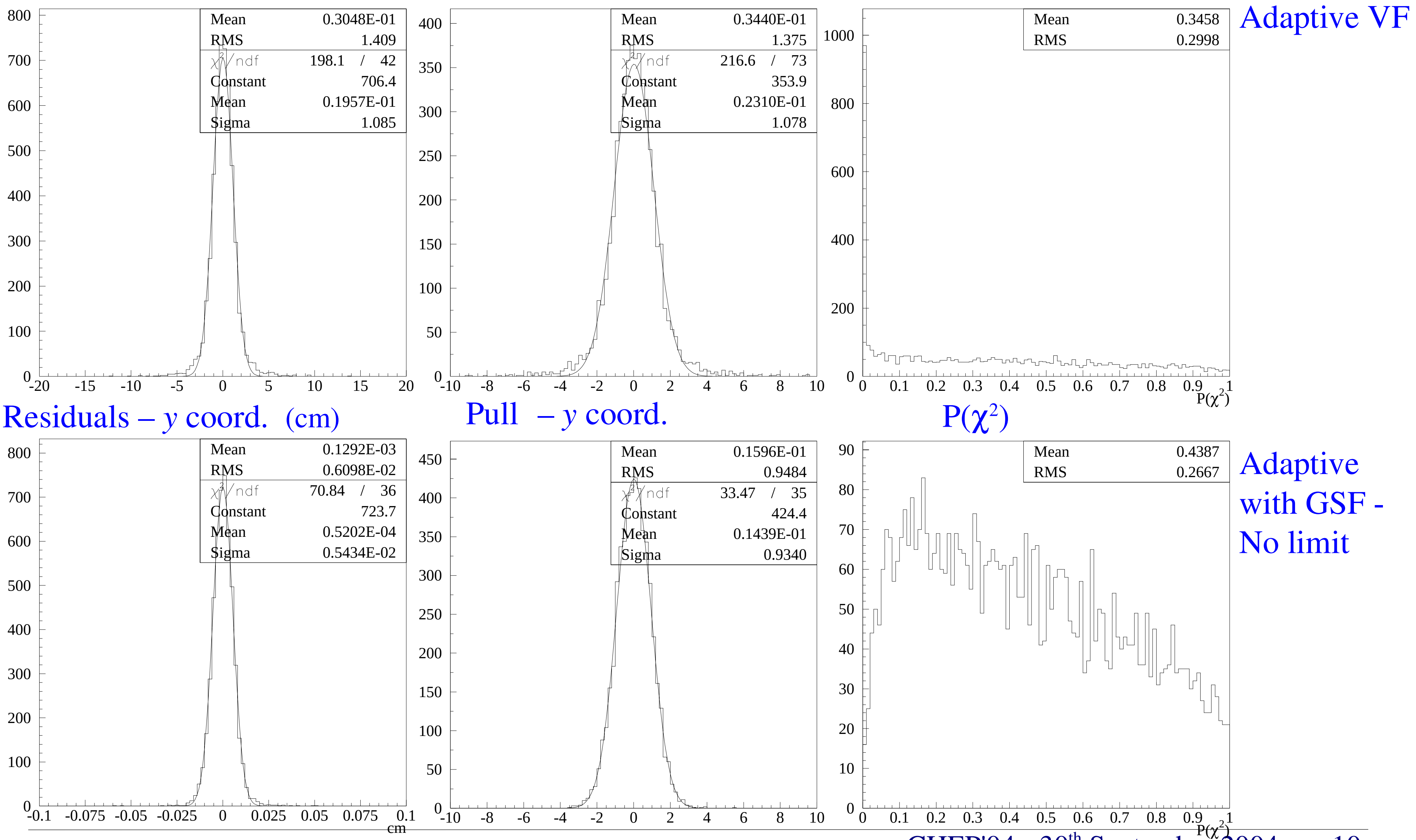
- Iterative, re-weighted LS fit
- More robust by down-weighting outliers:
  - ➢ Weight of a track depends on its distance to the vertex
- Adaptive VF very stable with high break-down point

$$w_i = \frac{1}{1 + \exp\left(\frac{\chi^2 - \chi_c^2}{2T}\right)}$$



- Default implementation: Adaptive Vertex fitter (AVF)
  - ➢ LS (Kalman) updator: sensitive only to the core of the distributions
- Adaptive Gaussian Sum Filter (A-GSF):
  - ➢ AVF with GSF updator: Gaussian mixture correctly taken into account!

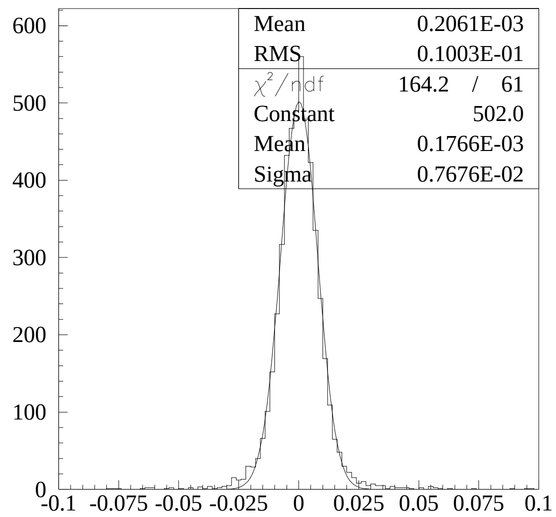
Robust filters – No outliers



Vertices without outliers

Filter	mean $P(\chi^2)$	Resolution	Pull	C(50%)	C(90%)	Rel.Eff.
Kalman	0.32	71	1.39	66	350	-
GSF	0.48	54	0.99	36	90	5.8
Adaptive	0.3	59	1.08	39	113	3.4
A-GSF	0.44	54	0.93	36	91	1.2

- 50% and 90% coverage: half-widths of the symmetric intervals covering 50% and 90% of the residual distribution (x-coordinate)
- Relative efficiency: ratio of the mean (3D) distances of the estimated vertex from its simulated position, for fits with the KVF and the considered filter
    - For KVF: estimated position independent of scaling of track parameter variance



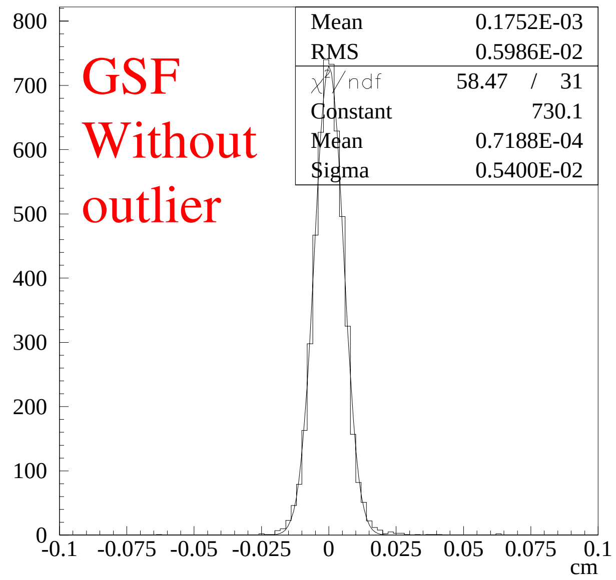
Residuals (cm)

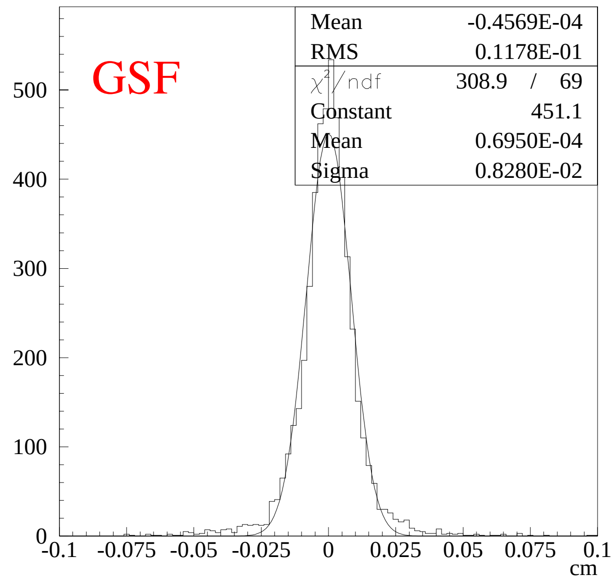
Type 1 outliers

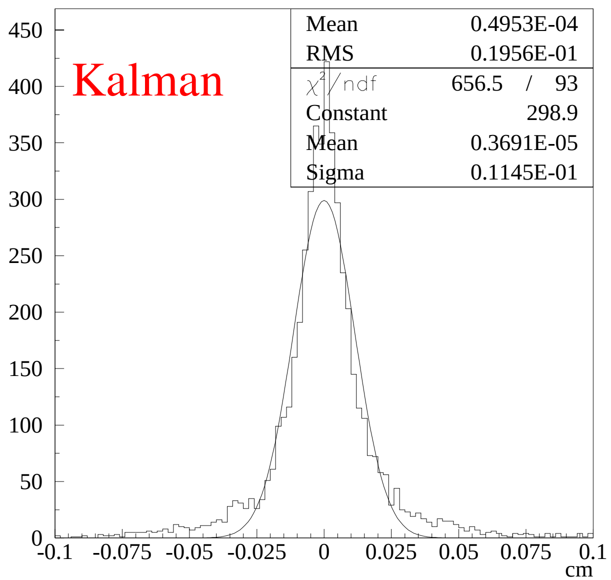
4 tracks from main vertex

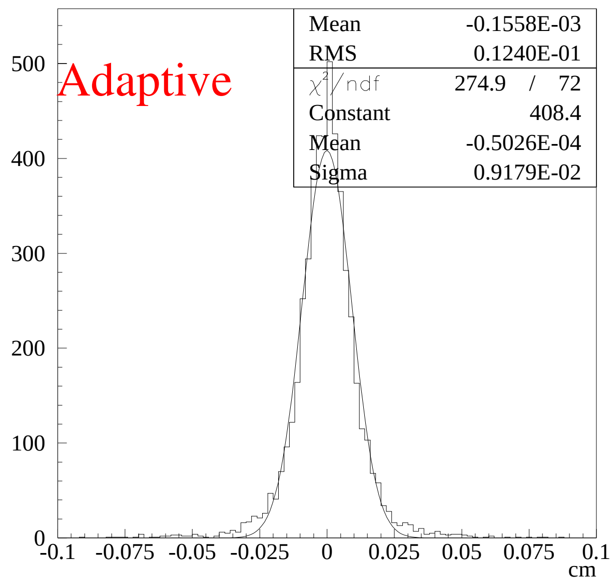
- 3 inliers (Filter sees mixture with correct covariance matrices)
- 1 outlier with $\sigma_o/\sigma_i = 3$

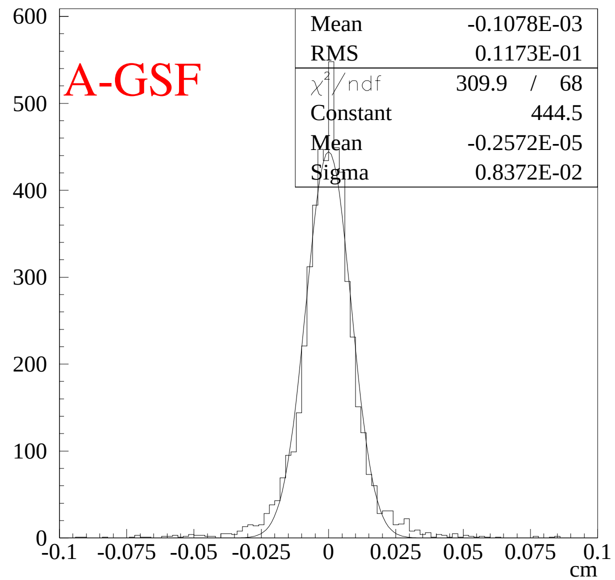
Residuals – *y* coordinate [cm]











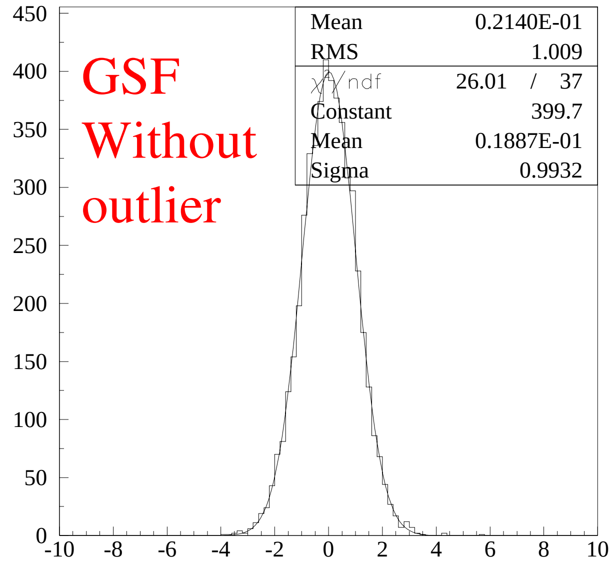
Type 1 outliers

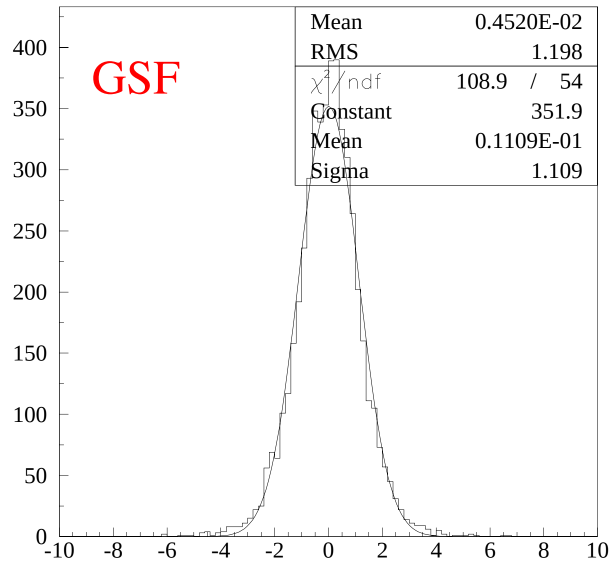
4 tracks from main vertex

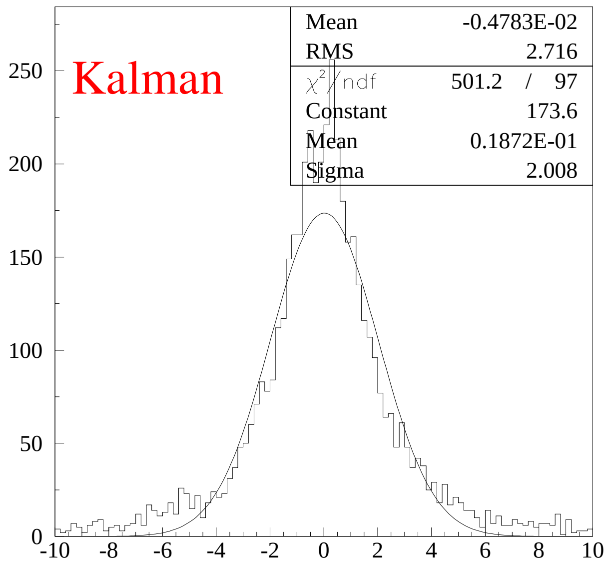
3 inliers

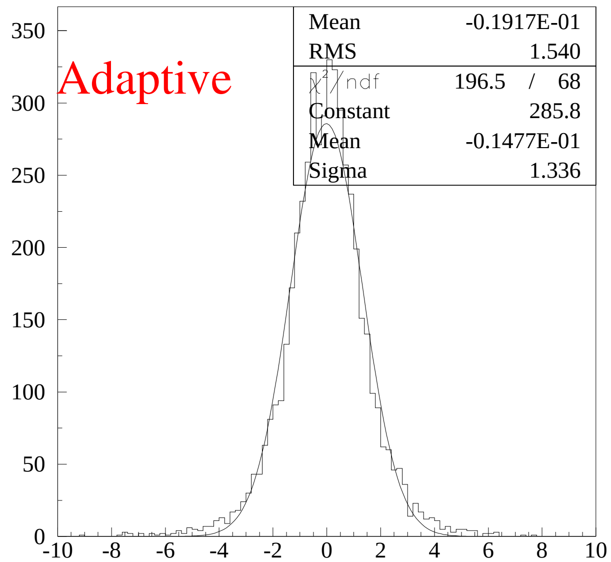
1 outlier with $\sigma_o/\sigma_\iota = 3$

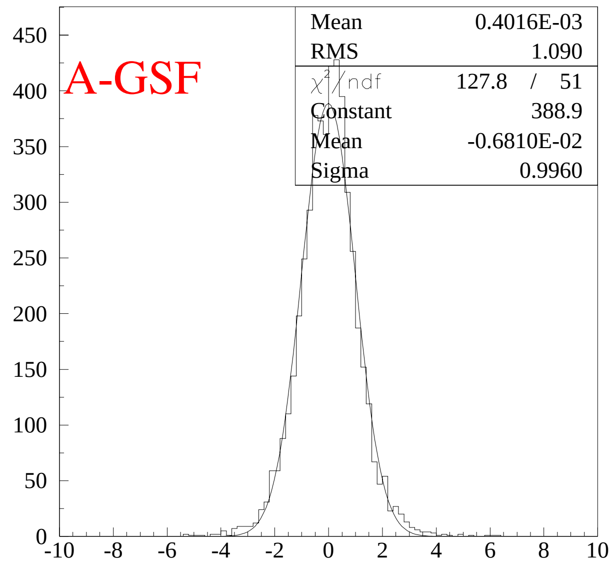
Pulls– *y* coordinate



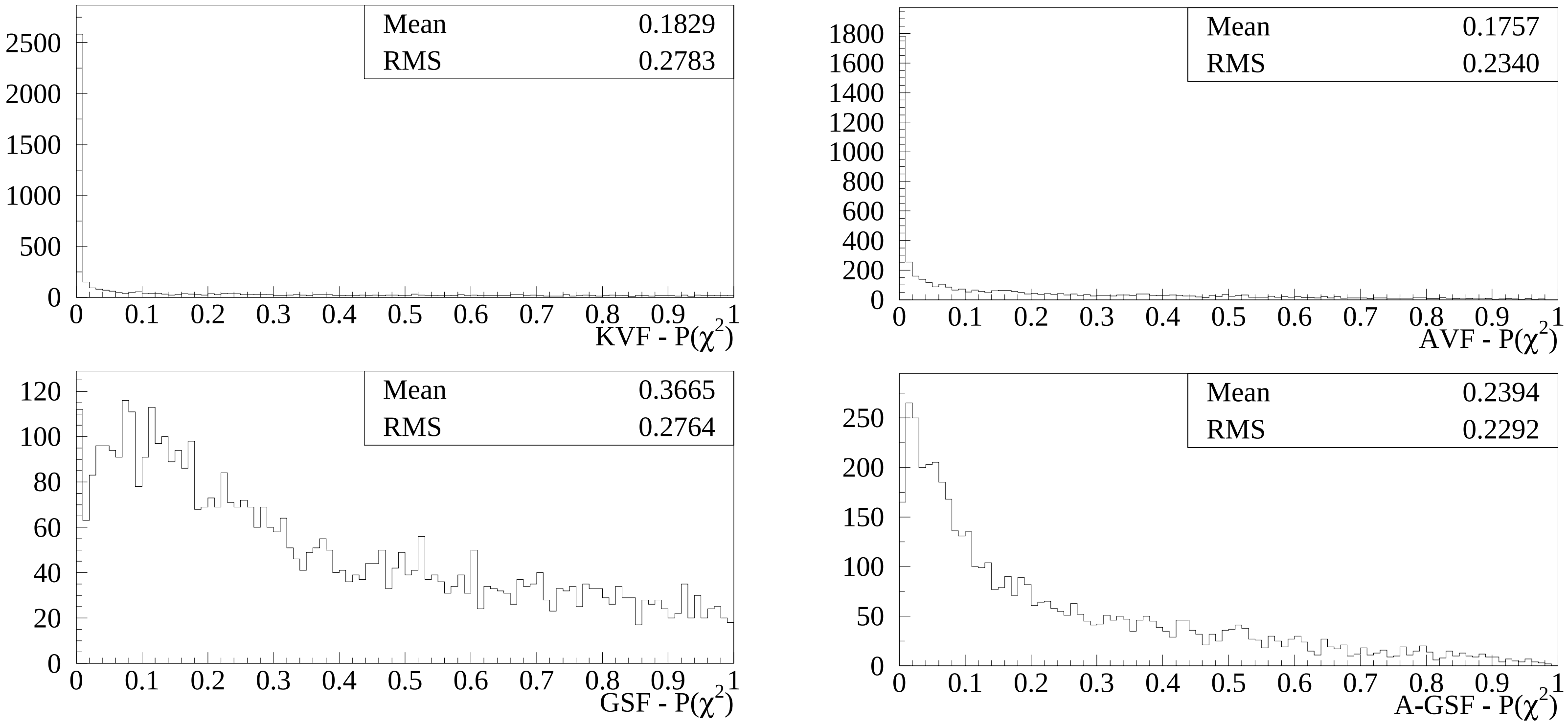








Type 1 outliers



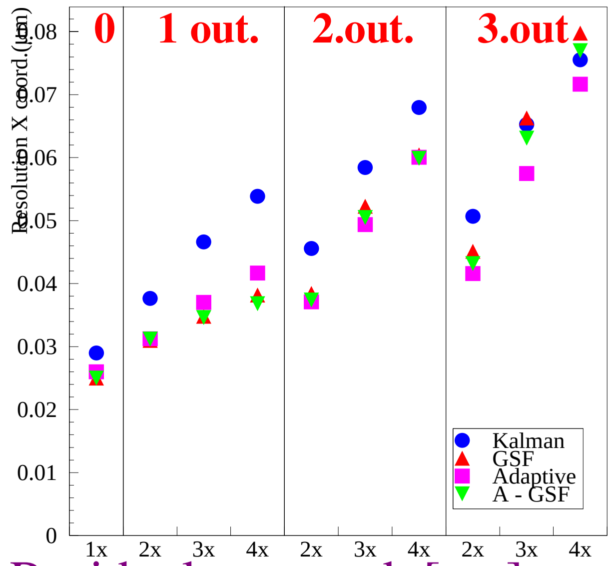
Filter	mean $P(\chi^2)$	Resolution	Pull	C(50%)	C(90%)	Rel.Eff.
Kalman	0.18	115	1.61	297	326	-
GSF	0.37	83	1.11	54	167	2.5
Adaptive	0.18	92	1.34	59	180	2
A-GSF	0.24	84	1	55	168	2.1

Type 1 outliers

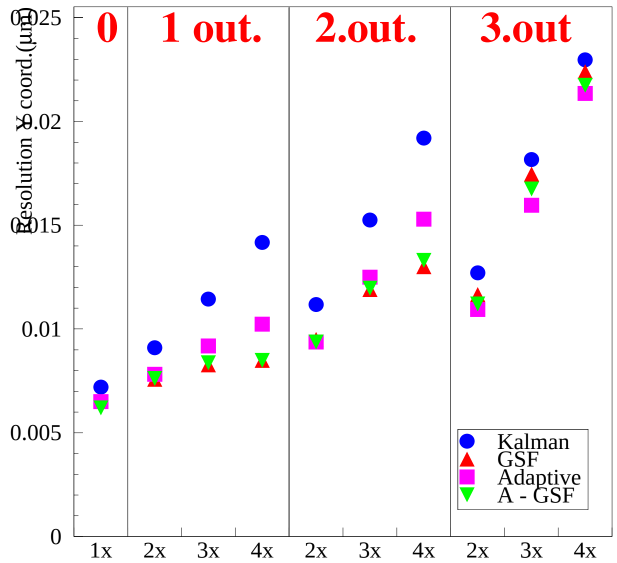
4 tracks from main vertex:

1, 2 or 3 outlier with

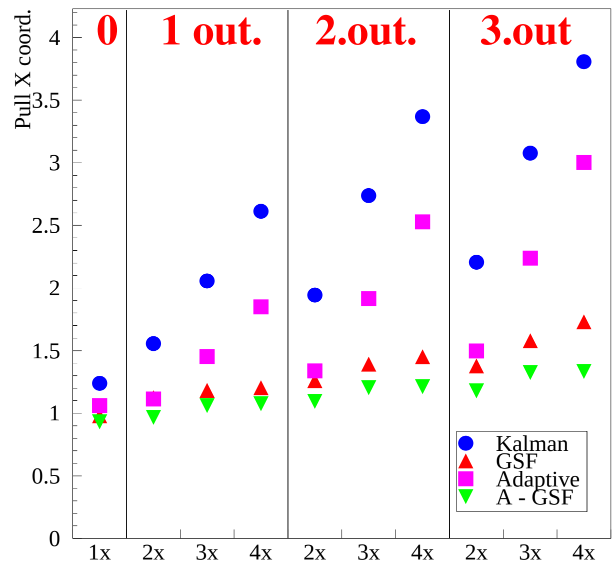
σ_o/σ_ι = 2, 3 or 4



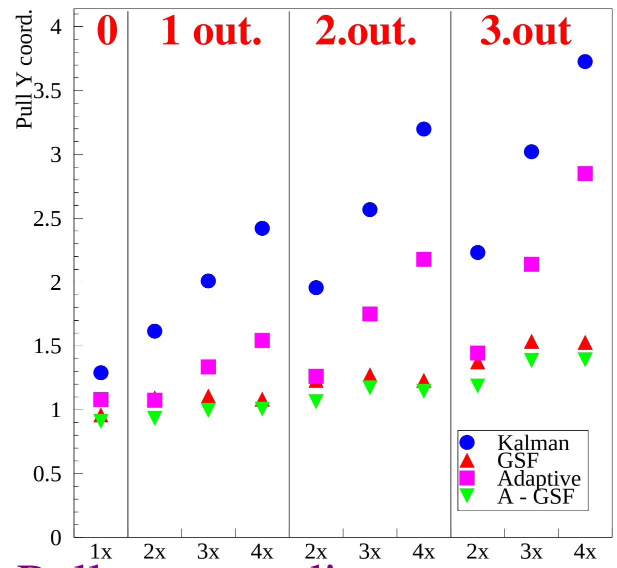
Residual– *x* coord. [cm]



Residual– *y* coord. [cm]



Pulls– *x* coordinate



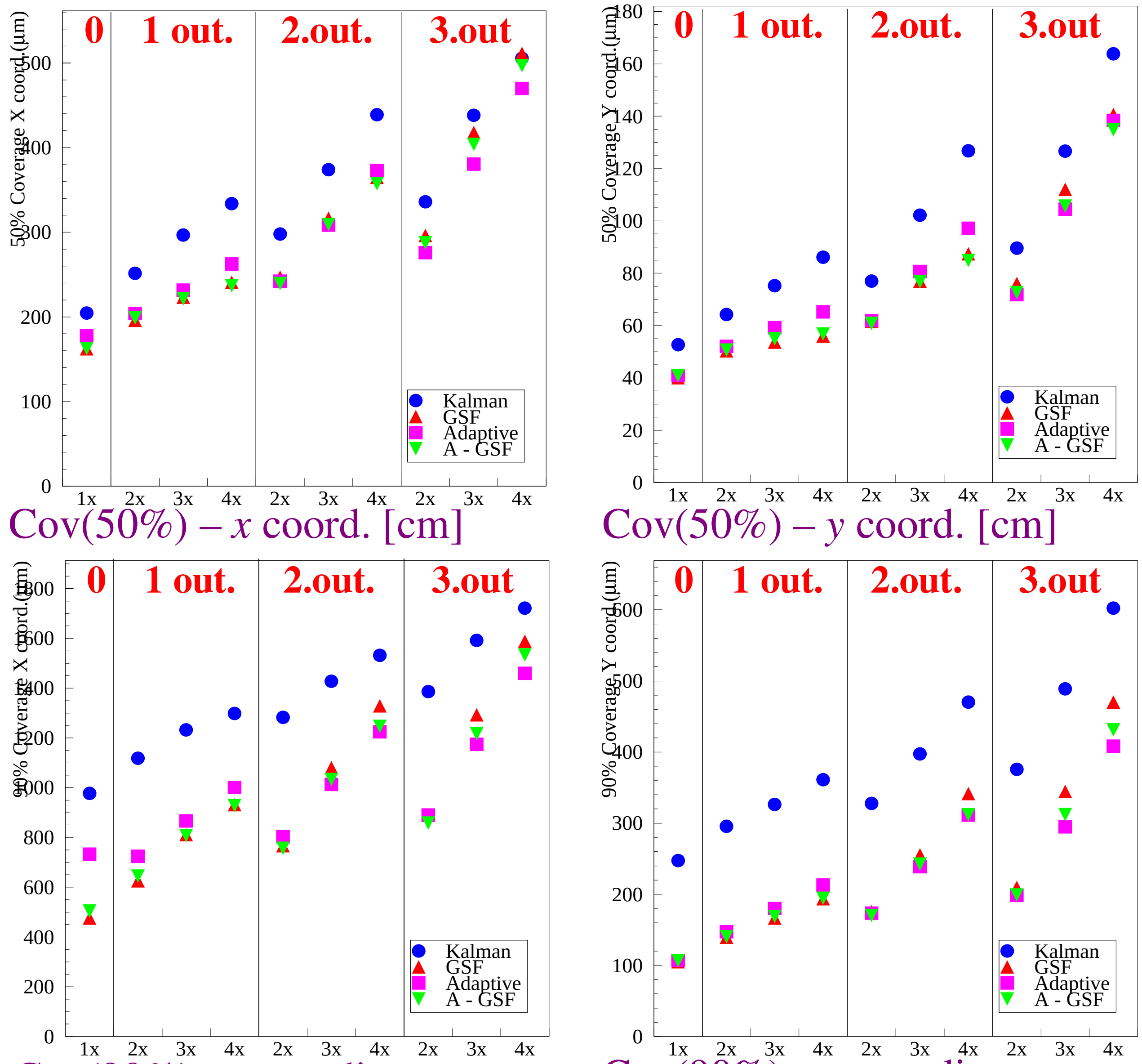
Pulls– *y* coordinate

Type 1 outliers

4 tracks from main vertex:

1, 2 or 3 outlier with

σ_o/σ_t = 2, 3 or 4



Cov(50%) – *x* coord. [cm]

Cov(50%) – *y* coord. [cm]

Cov(90%)– *x* coordinate

Cov(90%) – *y* coordinate

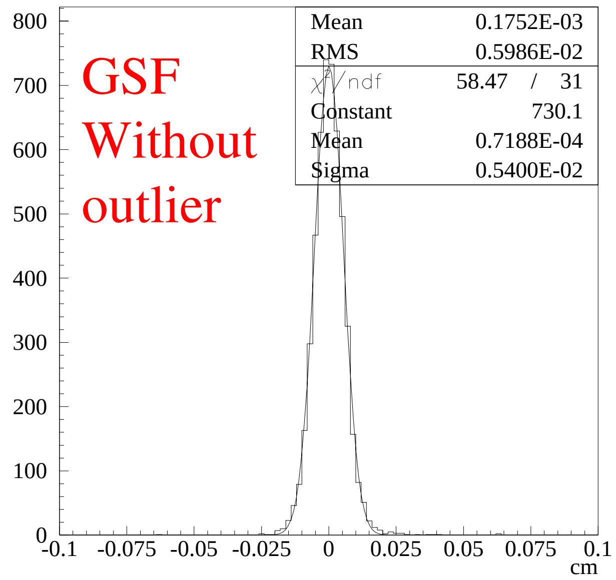
Type 2 outliers

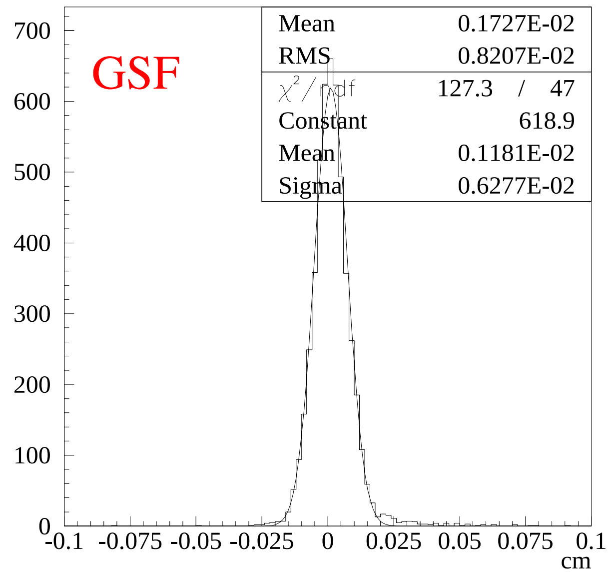
4 tracks from main vertex

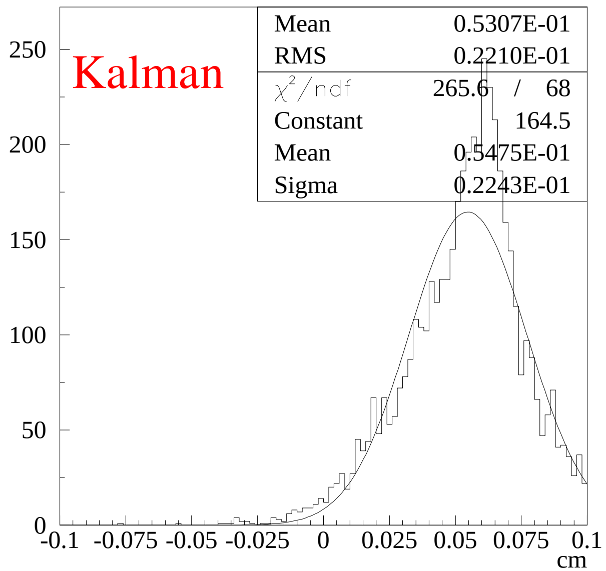
1 track from second vertex,

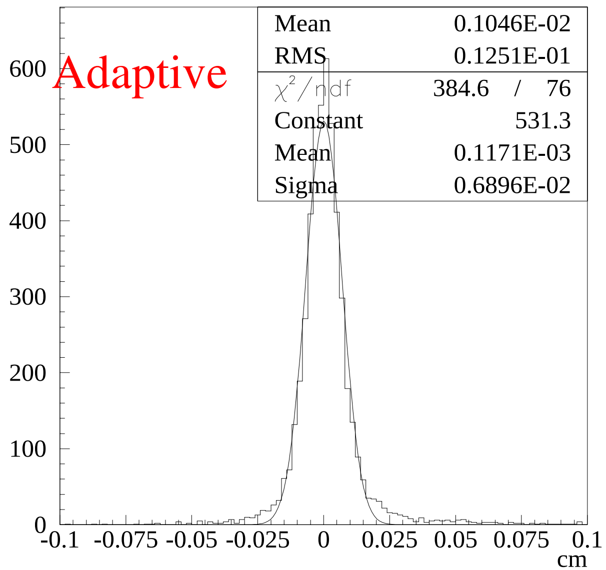
Δy = 3mm

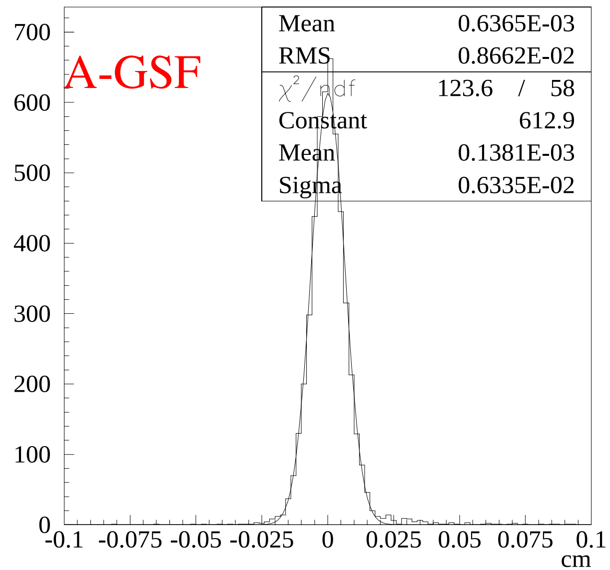
Residuals – *y* coordinate [cm]











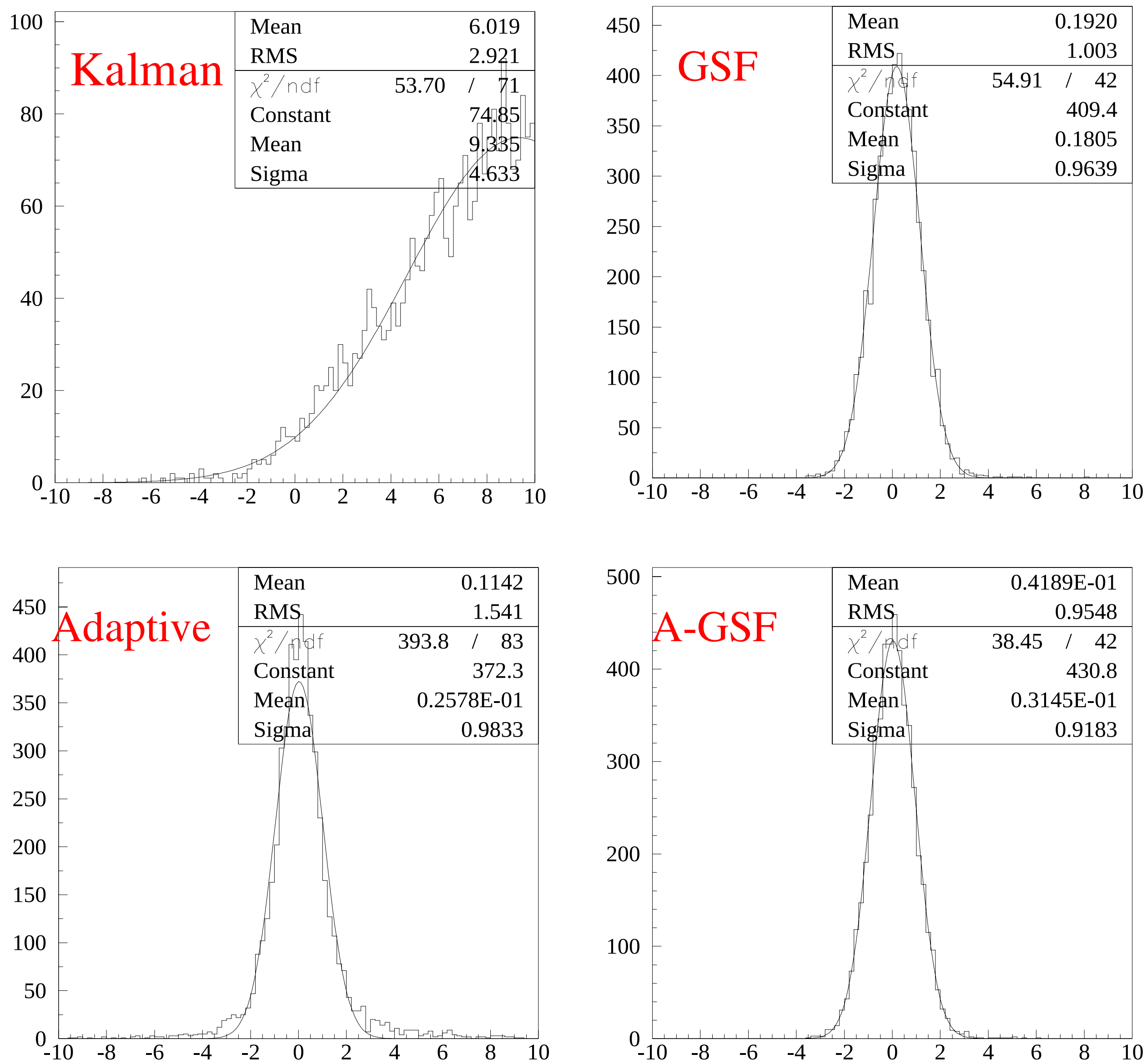
Type 2 outliers

4 tracks from main vertex

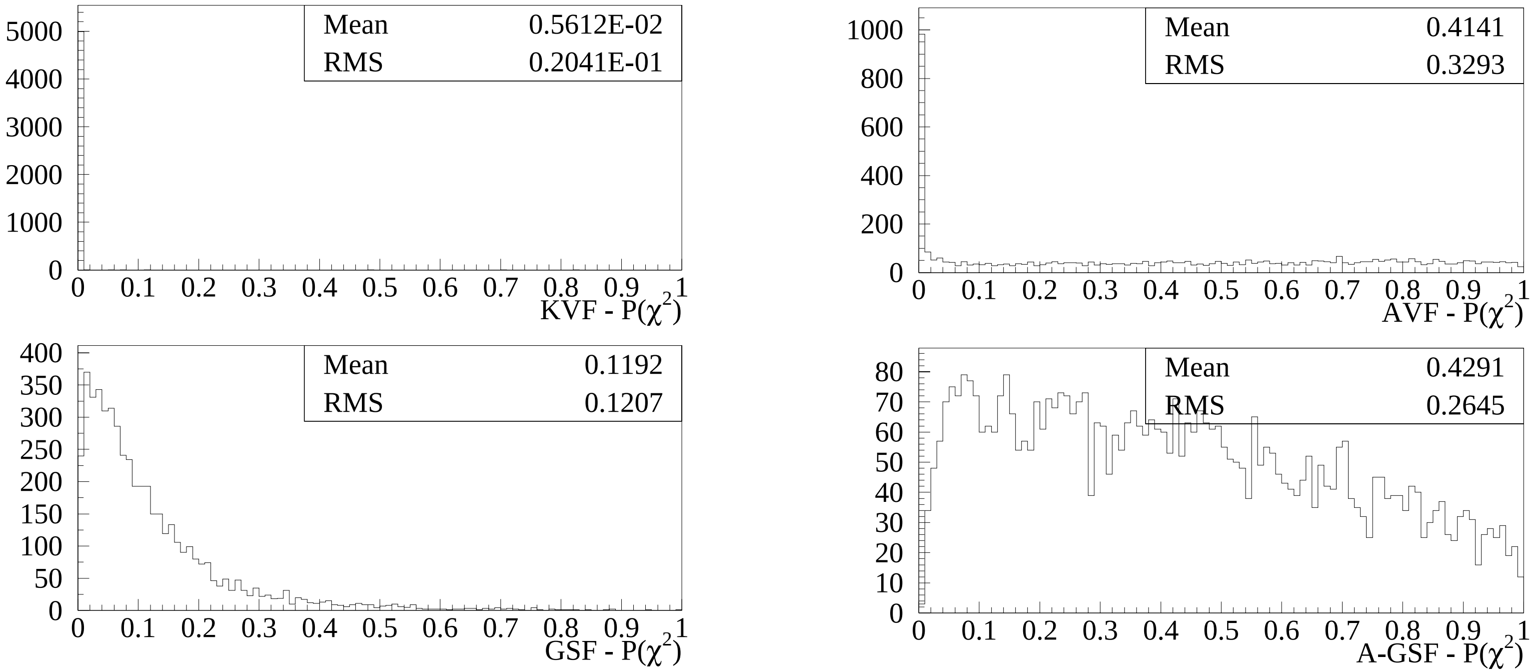
1 track from second vertex,

$\Delta y = 3\text{mm}$

Pulls – *y* coordinate [cm]



Type 2 outliers



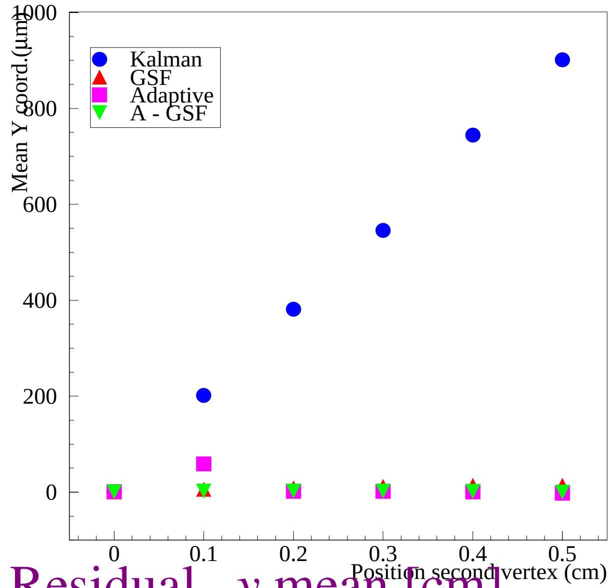
Filter	mean $P(\chi^2)$	Resolution	Pull	C(50%)	C(90%)	Rel.Eff.
Kalman	0.06	224 bias!	1.24	573	853	-
GSF	0.12	63	0.97	42	113	19.2
Adaptive	0.41	69	0.99	46	155	3.6
A-GSF	0.43	63	0.86	42	112	12.6

Type 2 outliers

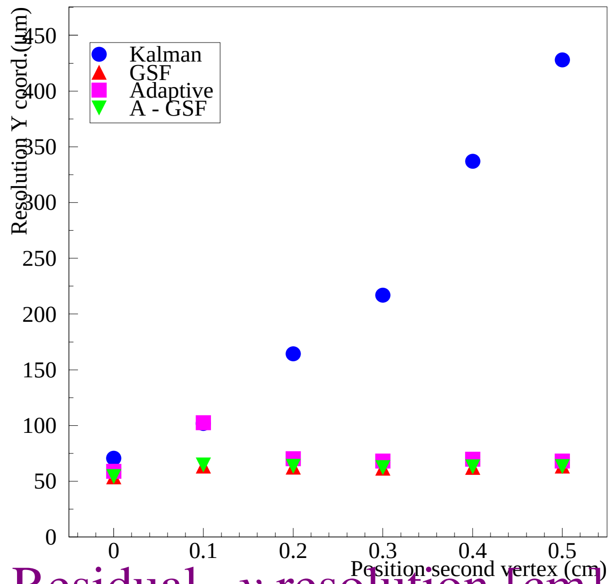
4 tracks from main vertex

1 track from second vertex,

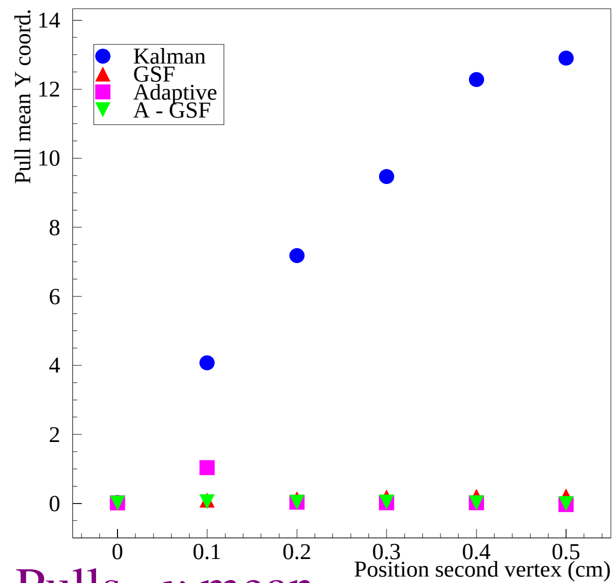
$\Delta y = 0 - 5$ mm



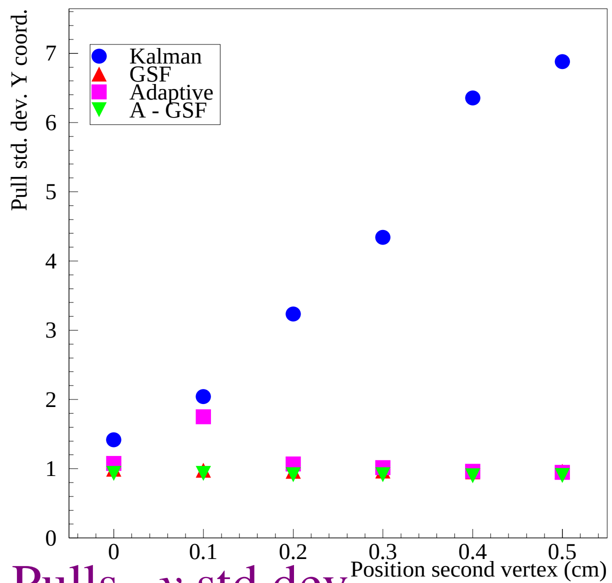
Residual– *y* mean [cm]



Residual– *y* resolution [cm]



Pulls– *y* mean



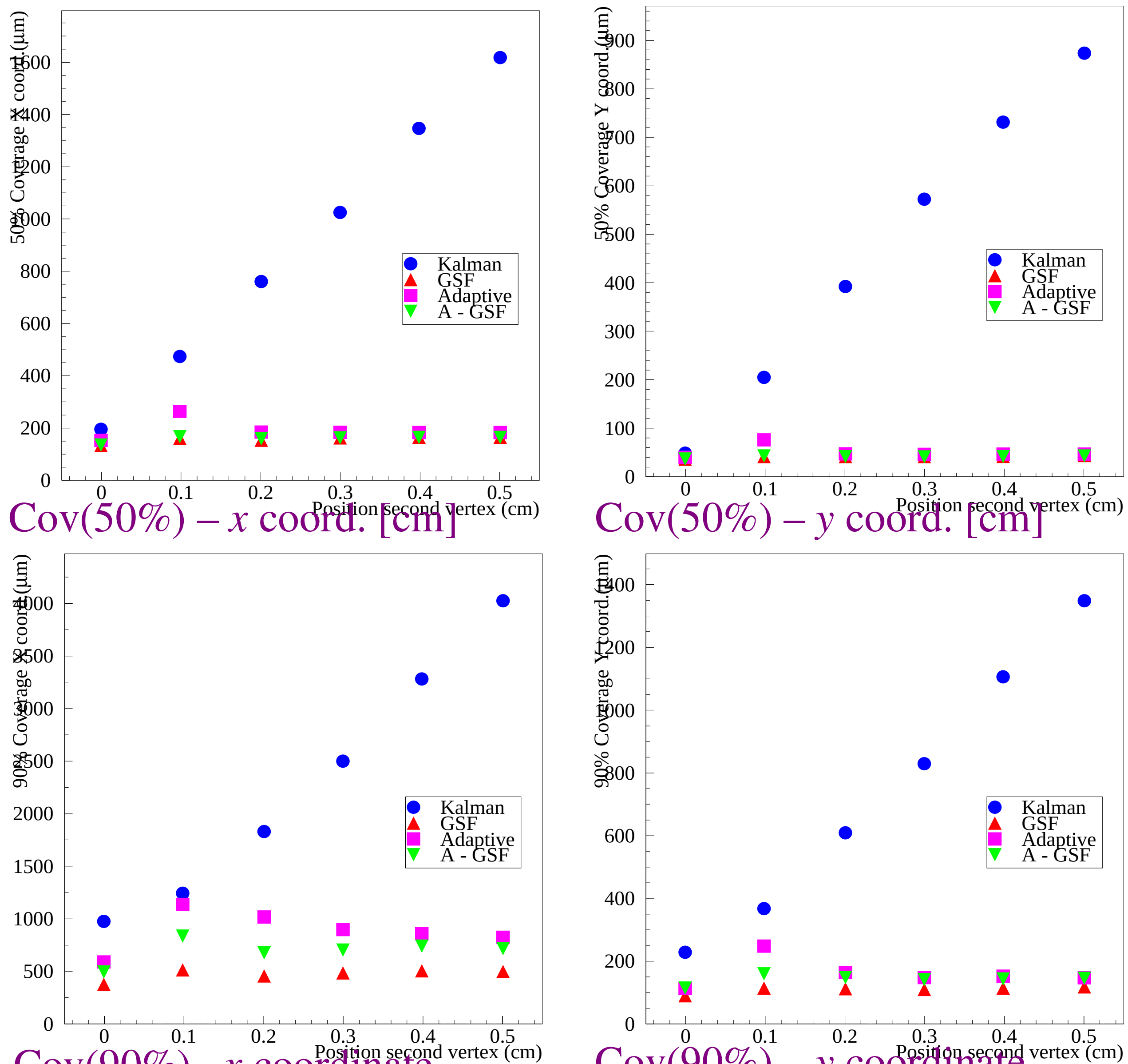
Pulls– *y* std.dev.

Type 2 outliers

4 tracks from main vertex

1 track from second vertex,

$\Delta y = 0 - 5$ mm



Cov(50%) – *x* coord. [cm]

Cov(50%) – *y* coord. [cm]

Cov(90%)– *x* coordinate

Cov(90%) – *y* coordinate

Conclusion

- A Gaussian-sum Filter for vertex reconstruction has been implemented in the CMS reconstruction software
- Shows an improvement of the resolution and error estimate of the fitted vertex and of the χ^2 of the fit with respect of the Kalman Filter when the track parameters residuals have non-Gaussian tails.
- Shows little sensitivity to the number of components kept during fit
  - ➢ A small number of components can be kept without degrading the fit too much.
- For electrons reconstructed with the GSF:
  - ➢ Allows to use the full mixture, and not only the single collapsed state.

Conclusion

- Results depend on the quality of the mixture describing the track
- Additional components improve robustness
  - In the presence of outliers, GSF residual, errors stable
- Adaptive-GSF:
  - Able to downweigh outliers
  - Uses full mixture
  - Most robust when tracks can be modelled by mixture of Gaussians