Reparamaterizing track errors in DC06

Vladimir Gligorov, Eduardo Rodrigues

University of Glasgow

T-Rec Meeting

26 May 2008

A big thanks to Hugo and Jose for their help!

OVERVIEW

- Motivation
- Fitting procedure
- DC06 results
  - Best fit polynomials & Residuals
  - Impact on HLT selections
- Future work

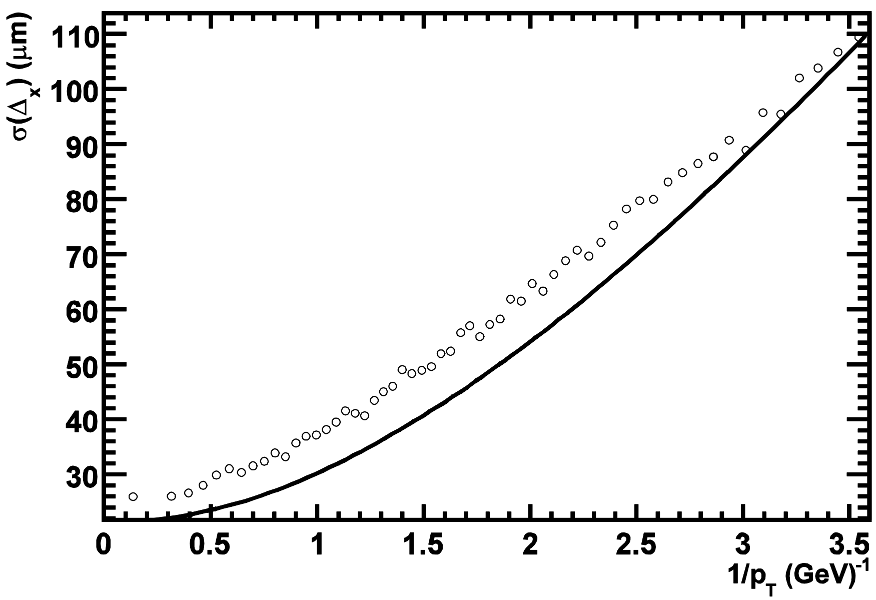
WHY ARE TRACK ERRORS IMPORTANT?

Without a correct calculation of track errors, using significance cuts in the trigger introduces inefficiencies

No time for a full fit in HLT2

- Must parameterize the track errors

The DC04 parametrization no longer works in DC06.



FITTING PROCEDURE

ERRORS AS A FUNCTION OF PT

- Follow the same procedure as used in Hugo's DC04 note*

- The track errors are binned as a function of $1/p_T$

- A polynomial is fitted to these errors, and used in the HLT to calculate entries in the 5x5 track covariance matrix

  - Errors in x and y are assumed uncorrelated

*** Ref: Hugo Ruiz, LHCb 2005-012**

CALCULATING THE ERRORS (DETAILS)

- Extrapolate the track to the same Z point as the primary vertex
- Calculate the X (or Y) deviation of the track from the PV (binned in p_T)
- In each bin, use an iterative procedure to estimate the core Gaussian width of the X(Y) deviation from the PV
  - Errors on the PV position are considered negligable, hence ignored

CALCULATING THE ERRORS (MORE DETAILS)

- For each bin, iterate as follows

    1. Compute RMS; reject tracks > 8xRMS from the mean

    2. Compute RMS; reject tracks > 7xRMS from the mean

    3. Compute RMS; perform Gaussian fit in ±1xRMS region; Reject tracks > 6σ from the mean

    4. Repeat 3, rejecting tracks > 5σ from the mean

    5. Repeat 3, rejecting tracks > 4σ from the mean

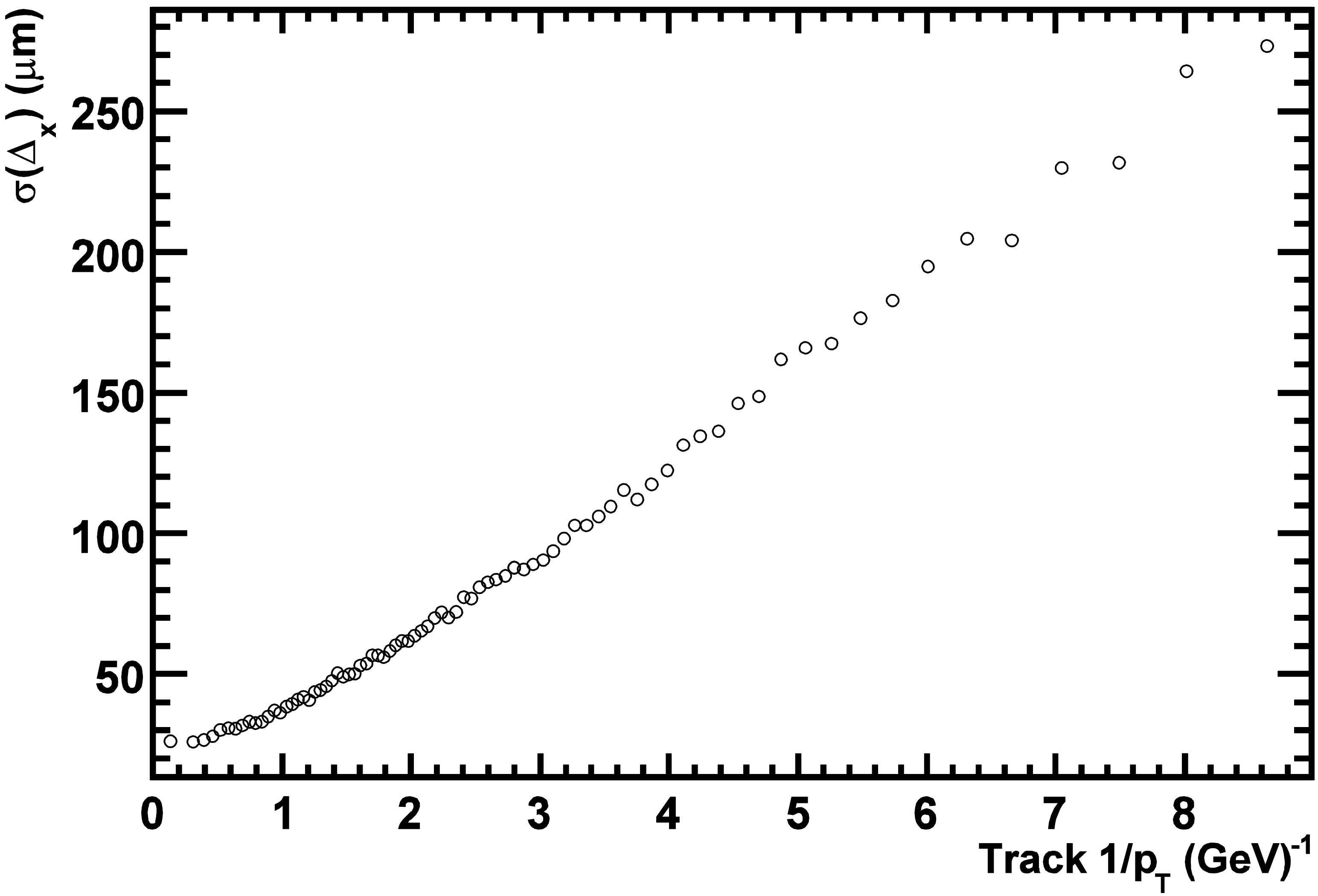
    6. Perform a final Gaussiaan fit in the ±3xRMS region

DC06 results

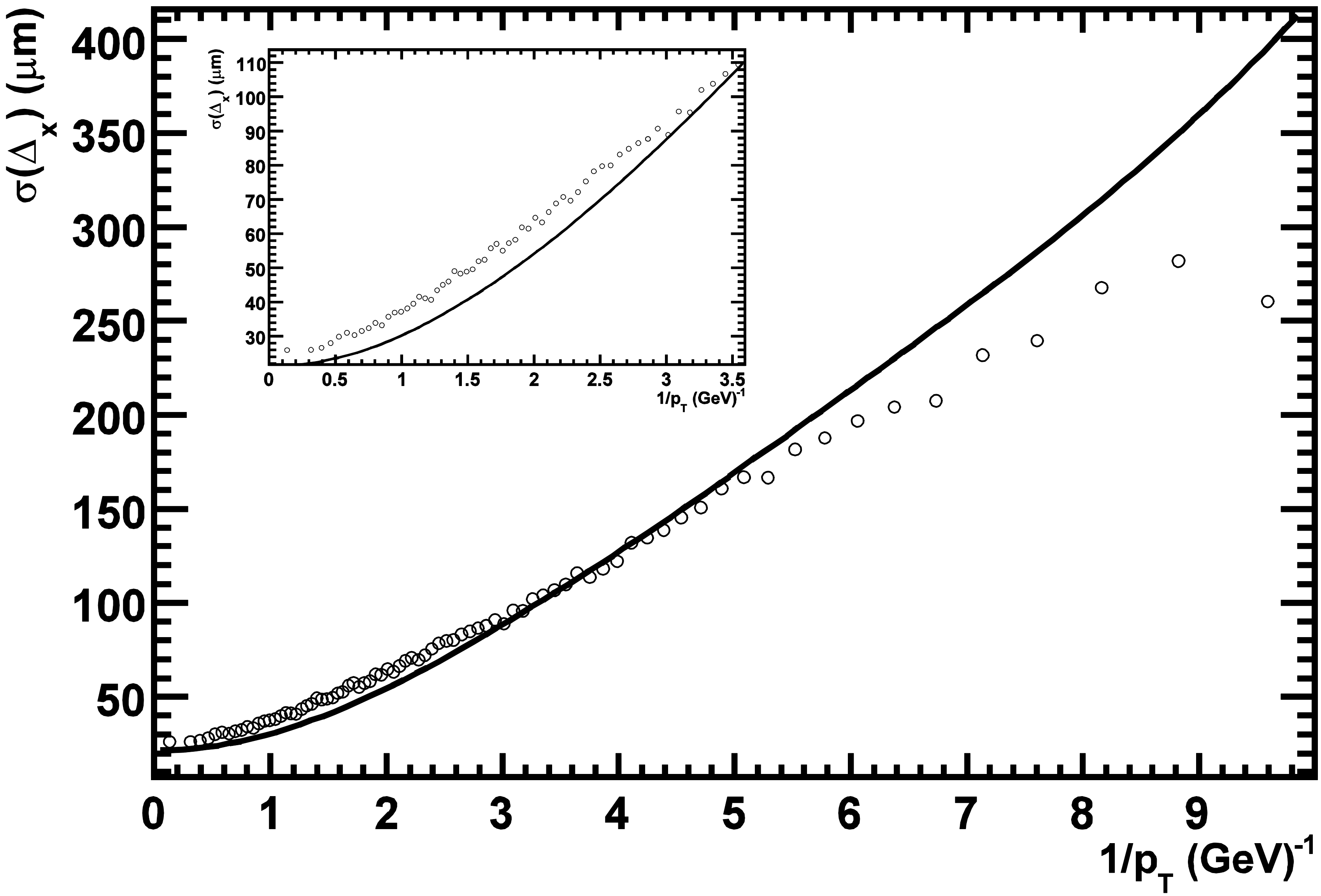
DATA SAMPLE

- DaVinci v19r12
- 20,000 L0 stripped minbias events
- Select events with 1PV only
- Make all particles as pions using StdNoPIDsPions
  - Only Long tracks used!
- No MC truth information used

Errors as a function of $1/p_T$



WITH THE OLD PARAMETRIZATION...

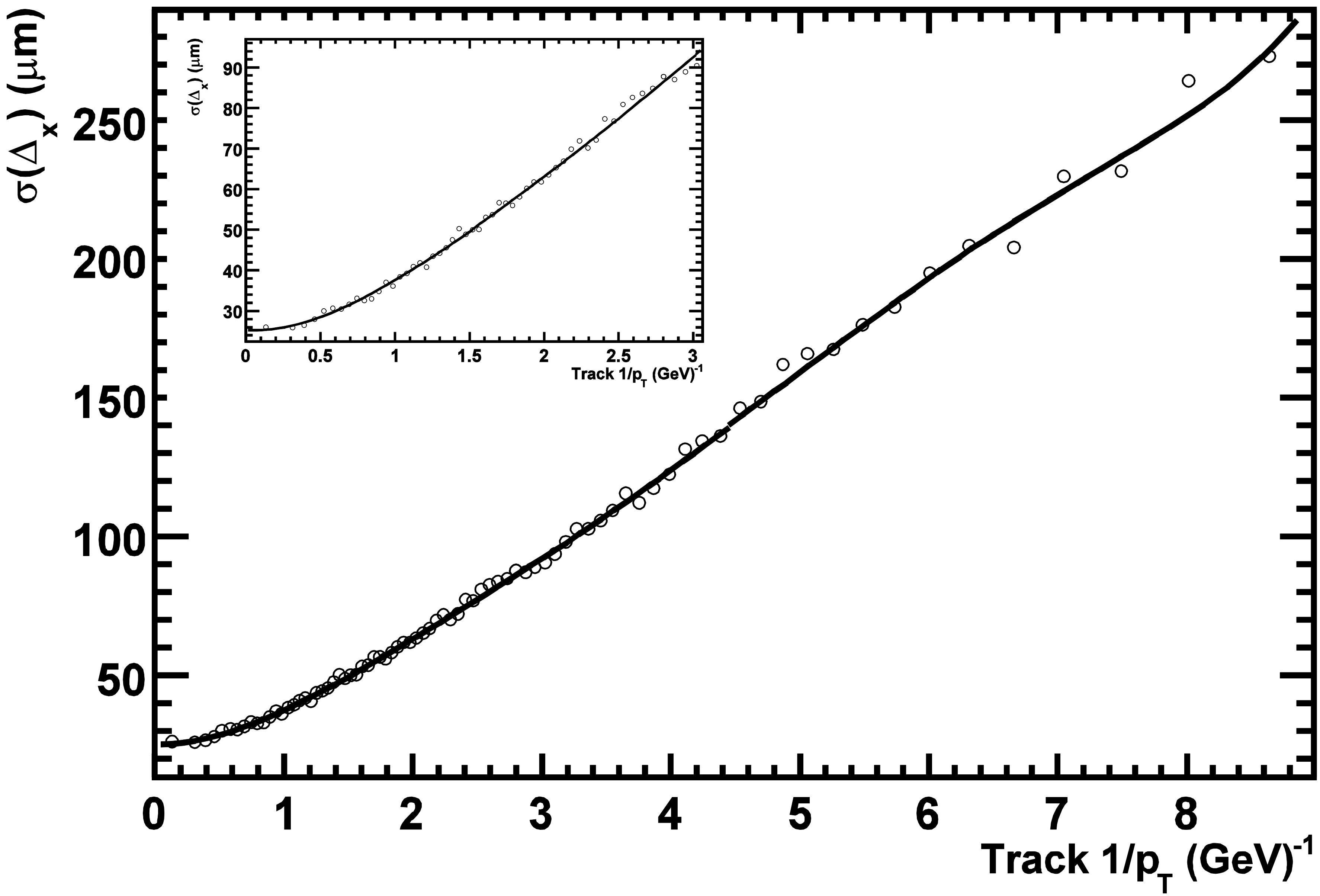


BEST FIT POLYNOMIALS

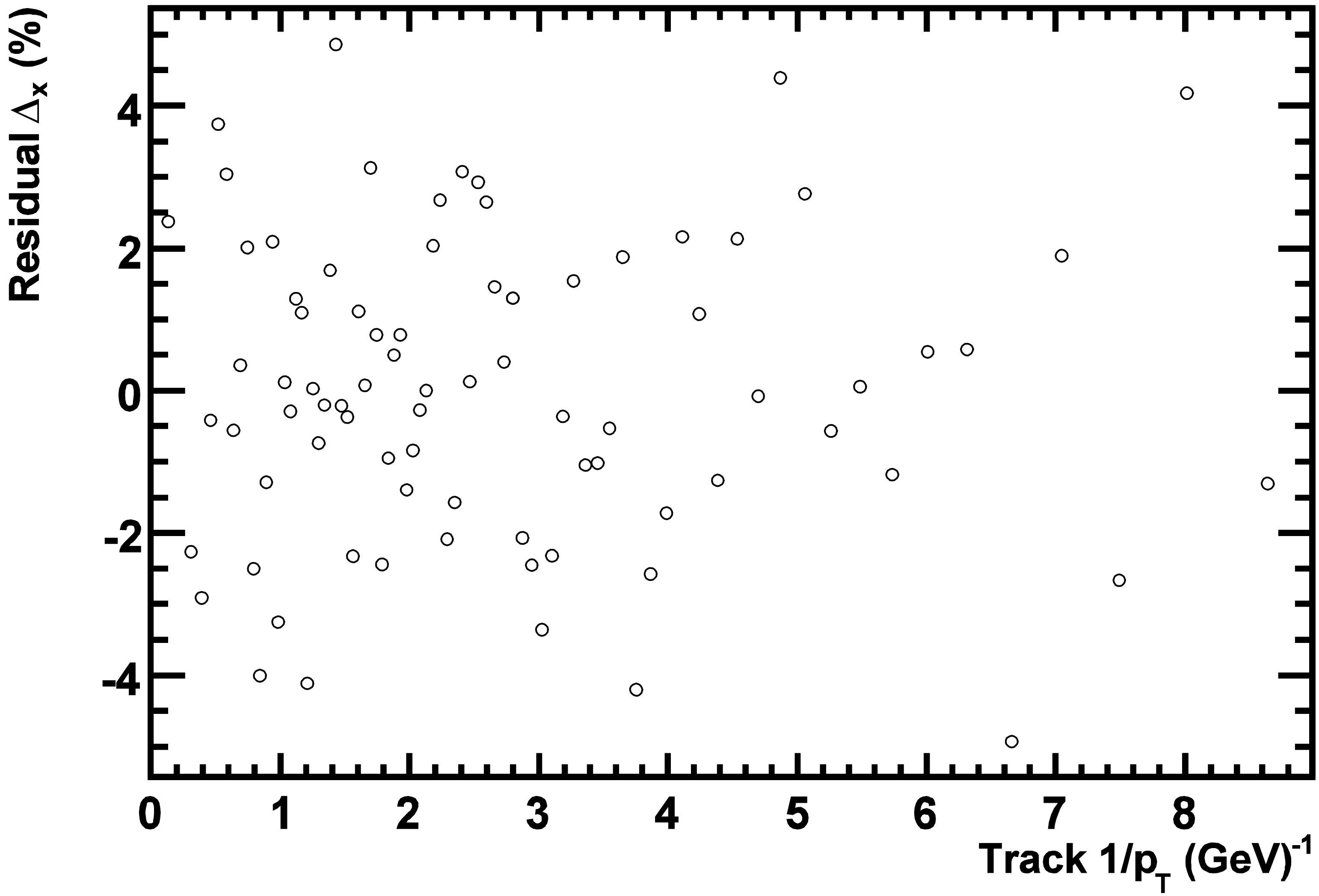
EXPLANATION

- Plots have been made with second to sixth degree polynomials
- Show the sixth degree fits here, the rest are in the backups
- Note that the first few bins are the most important since they contain the high p_T signal-like tracks
  - The first bin contains all tracks with $p_T > 10$ GeV and is hence espacially important

FITTING WITH A 6TH DEGREE POLY

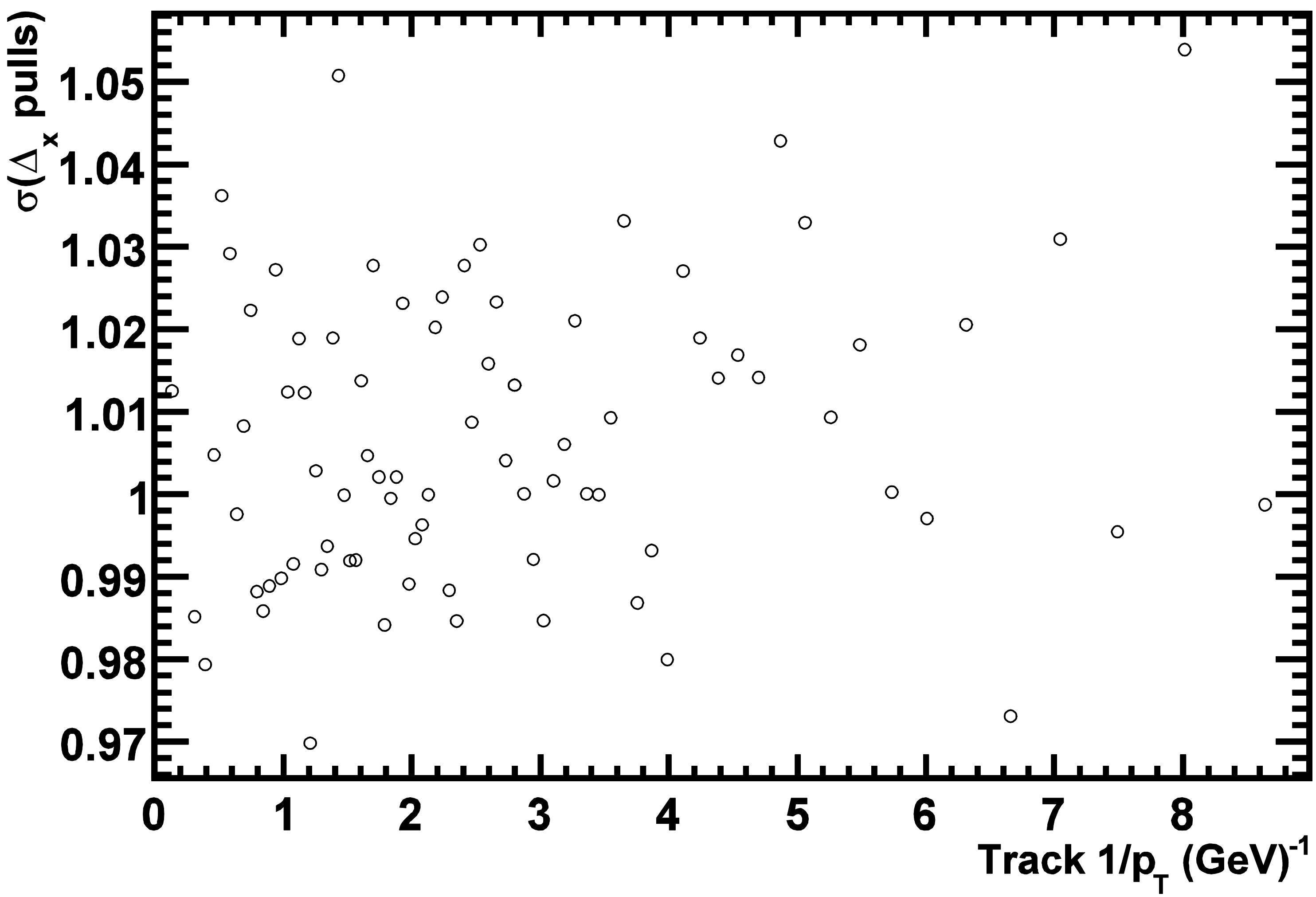


RESIDUALS WITH A 6TH DEG. POLY FIT



This plot shows the residual between the fitted and measured error for each bin of p_T, calculated at the midpoint of the bin.

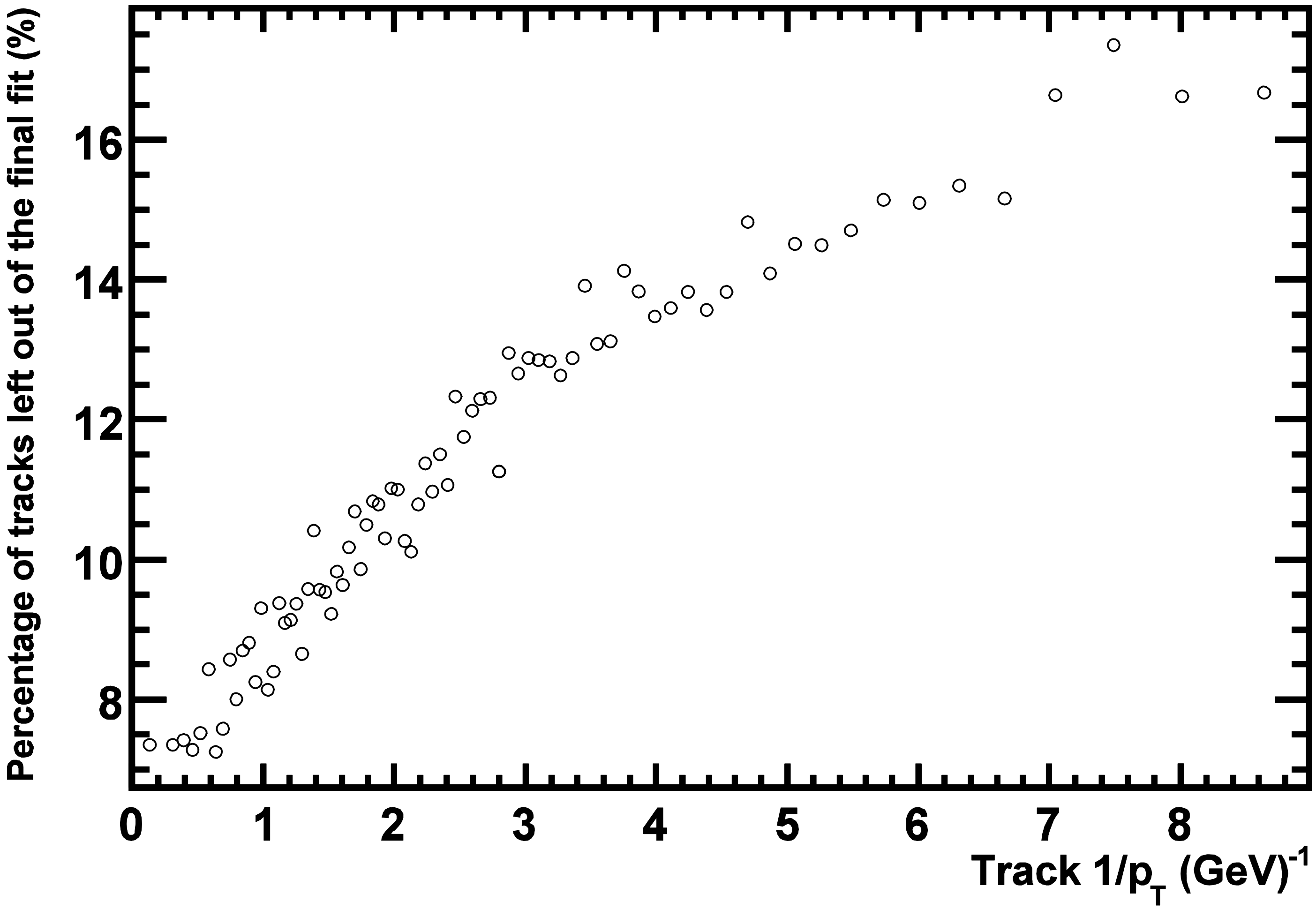
PULLS WITH A 6TH DEG. POLY FIT



This plot shows the width of the pull in bins of p_T. The pull in any one bin is computed by dividing the measured Δ_x for every track in that bin by the parameterized error. For perfect agreement the widths should be equal to 1.

HOW WELL IS THE ITERATIVE PROCEDURE ACTUALLY WORKING?

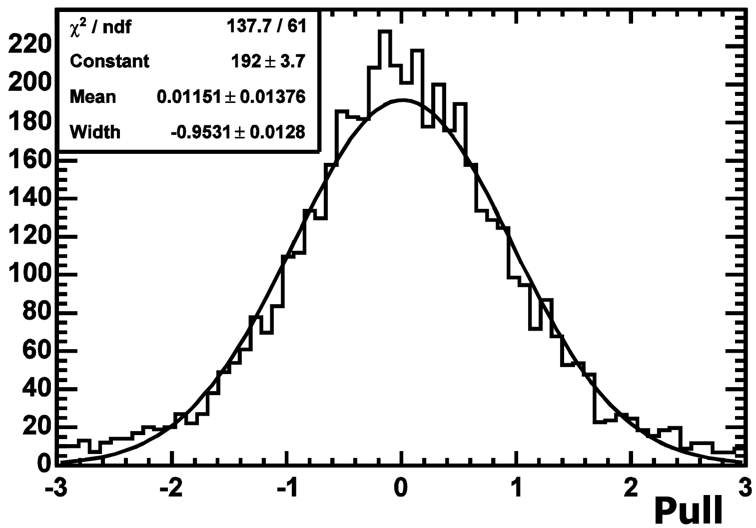
TRACKS OUT OF FIT



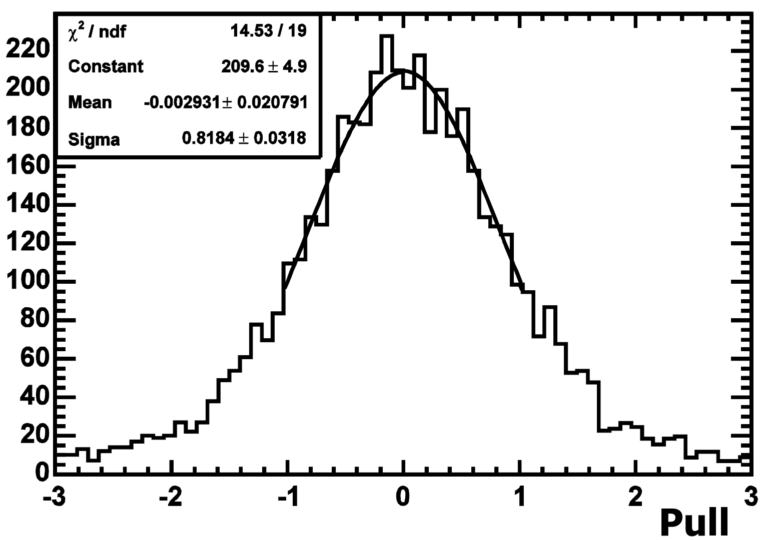
This plot shows the percentage of tracks removed by the iterative procedure before the final 3σ fit

FINAL 3σ GAUSSIAN FIT PULLS

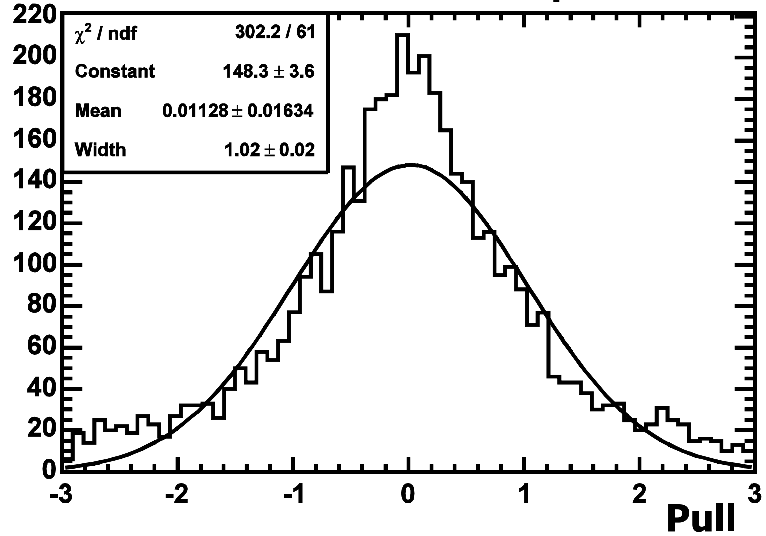
First bin of p_T



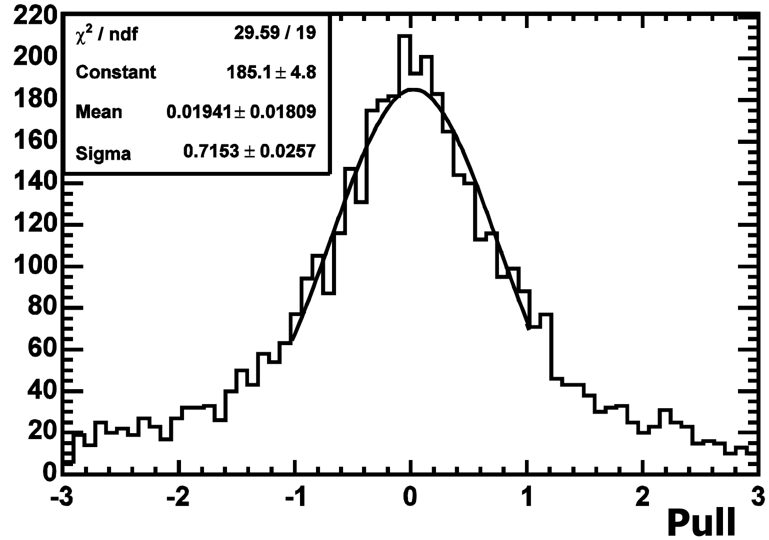
Refitted in 1σ region



Last bin of p_T



Refitted in 1σ region



COMMENT ON ITERATIVE PROCEDURE

- Even after iteratively rejecting tracks, we do not get single gaussians, especially at low p_T.
- Are we overestimating the errors (maybe)? How do you quote a single width for something which is a double gaussian anyway?

EFFECT OF NEW PARAM. ON HLT2 SEL.

Test the two parametrizations on a sample of $B_s \rightarrow D_s K$ events, selected by requiring all final state particles to have a p_T > 600 MeV

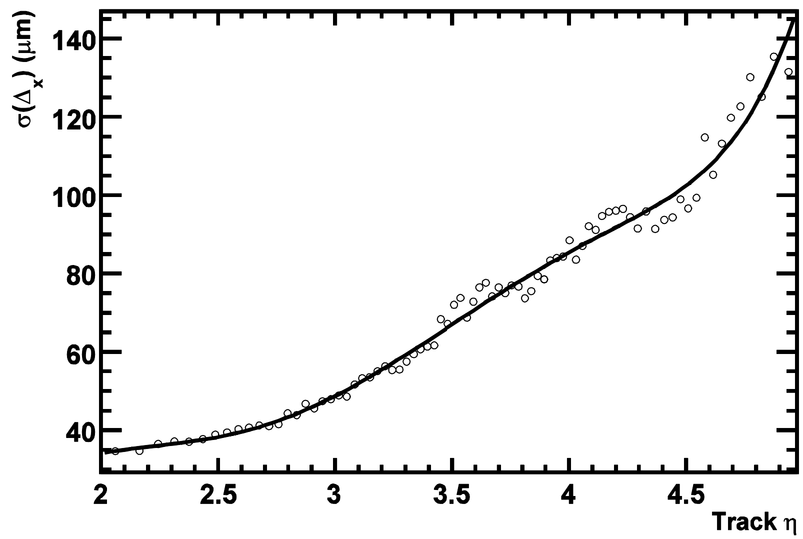
With DC06 Error parametrization	
Applied Cut	**HLT2 Eff. wrt. Offline**
IPS > 3 (all final state)	**85%**
B IPS < 4	**98%**
B flight significance > 8	**80%**

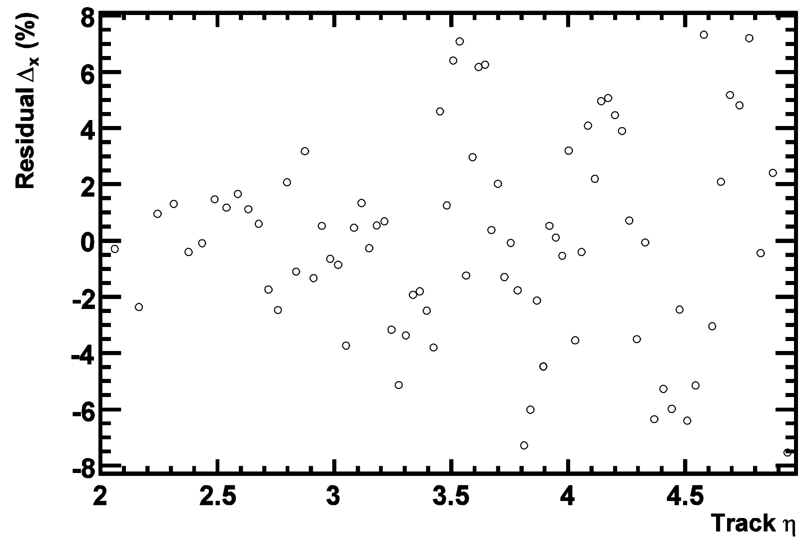
With DC04 Error parametrization	
Applied Cut	**HLT2 Eff. wrt. Offline**
IPS > 3 (all final state)	**91%**
B IPS < 4	**100%**
B flight significance > 8	**87%**

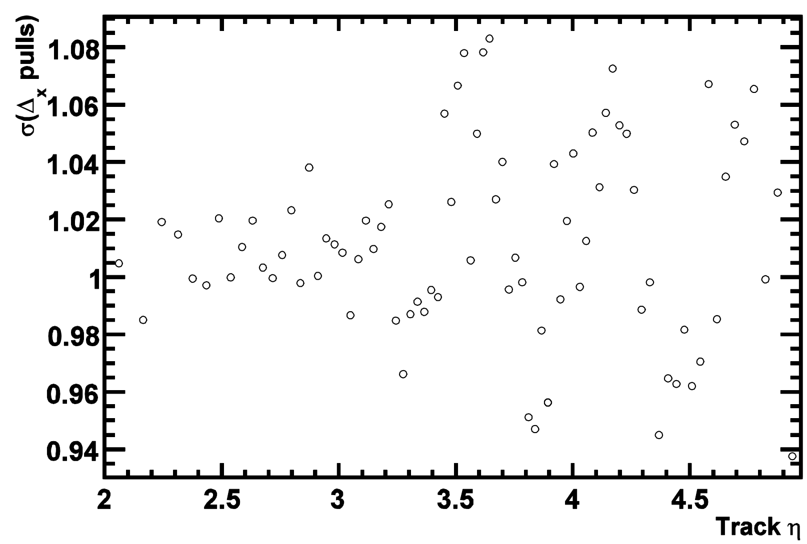
Unfortunately, the DC04 parametrization was underestimating the errors, so the relative efficiencies actually get worse, not better... on the other hand, the minbias rejection should get better as well.

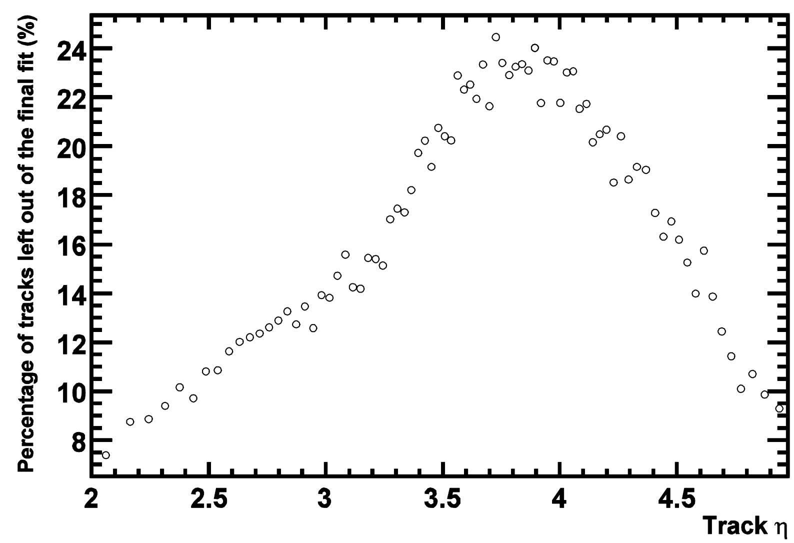
DEPENDANCE ON OTHER VARIABLES

Errors as a function of η

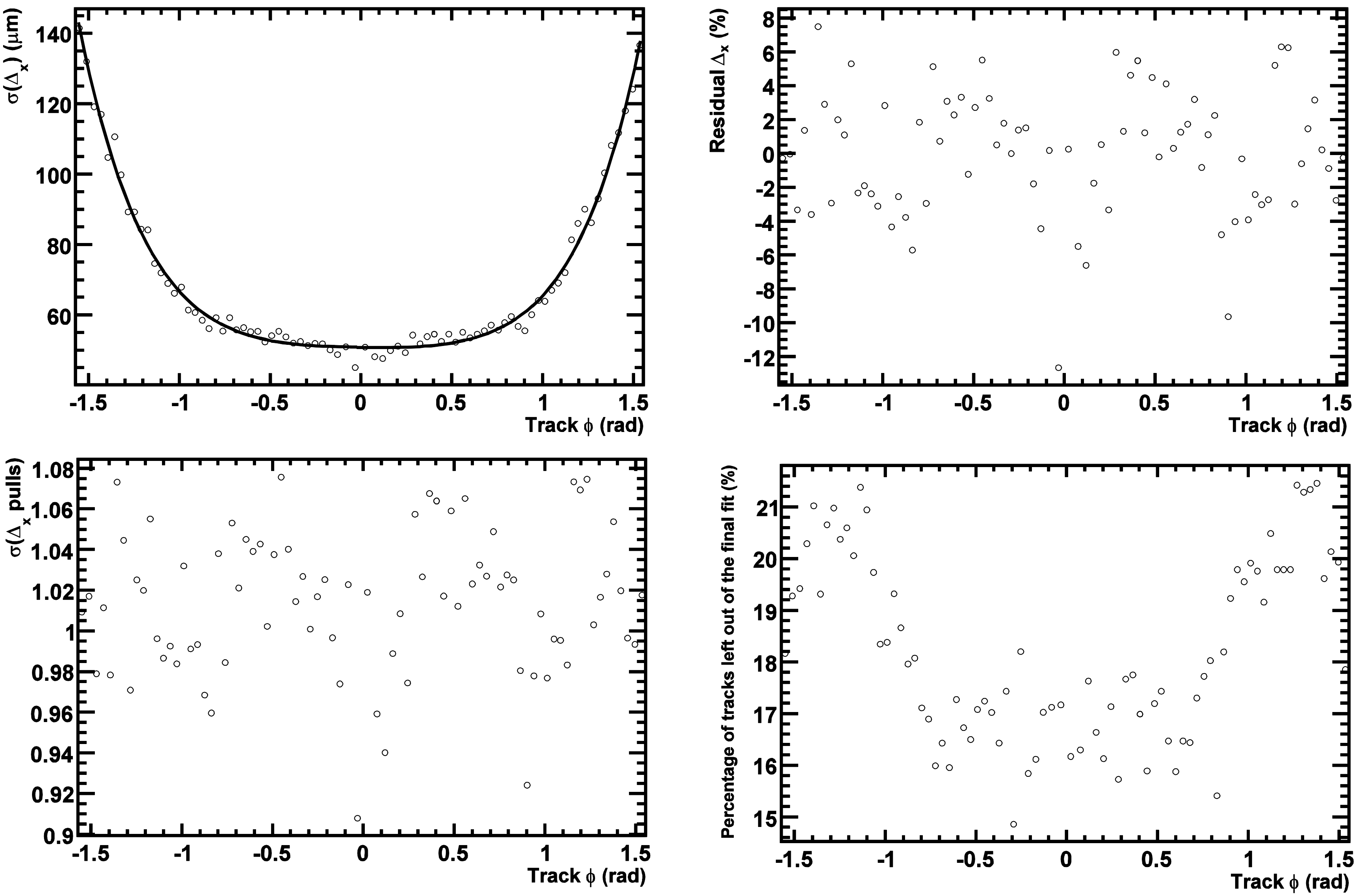






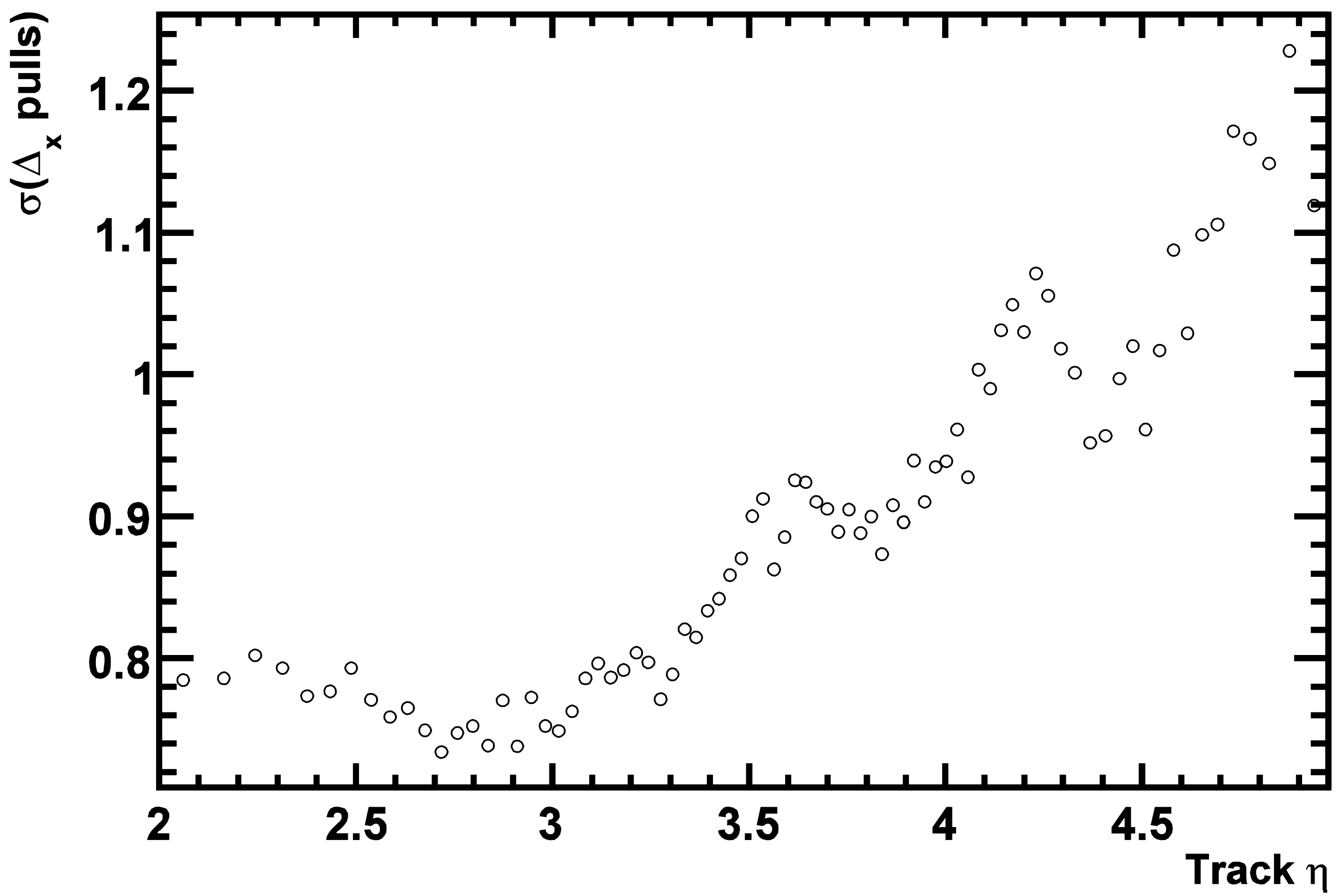


Errors as a function of φ



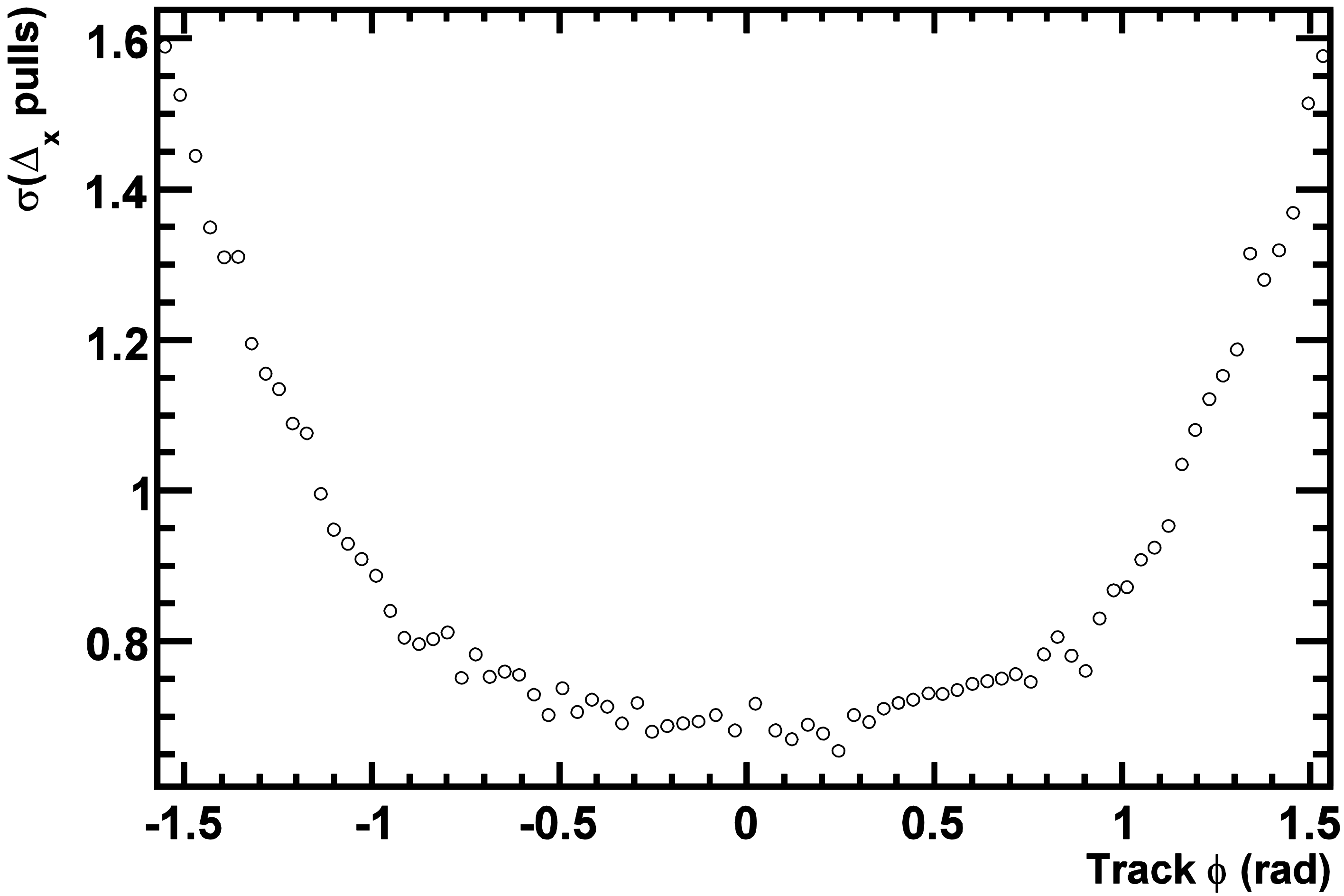
RESIDUAL DEPENDENCIES WHEN FITTING AGAINST P_T

RESIDUAL DEPENDENCY ON η



This plot shows the width of the pull in bins of **η** when the p_T parametrization is assumed. For perfect agreement the widths should be equal to one.

RESIDUAL DEPENDENCY ON φ



This plot shows the width of the pull in bins of **φ** when the p_T parametrization is assumed. For perfect agreement the widths should be equal to one.

TAKING RESIDUAL DEPENDENCIES INTO ACCOUNT

- As noted in the DC04 study, there are substantial correlations between the different residual correction
- Hence if we want an improvement on the p_T only correction, we would need a look-up table
- In first instance, consider the varibales p_T, phi, and eta; each split into 80 bins
  - The table then has 512000 entries (half if you assume perfect symmetry in phi)
  - Presents certain logistical difficulties... can it be implemented like magnetic field map?
  - Do we need it? Would need more than 20,000 tracks to do the fit for the table.

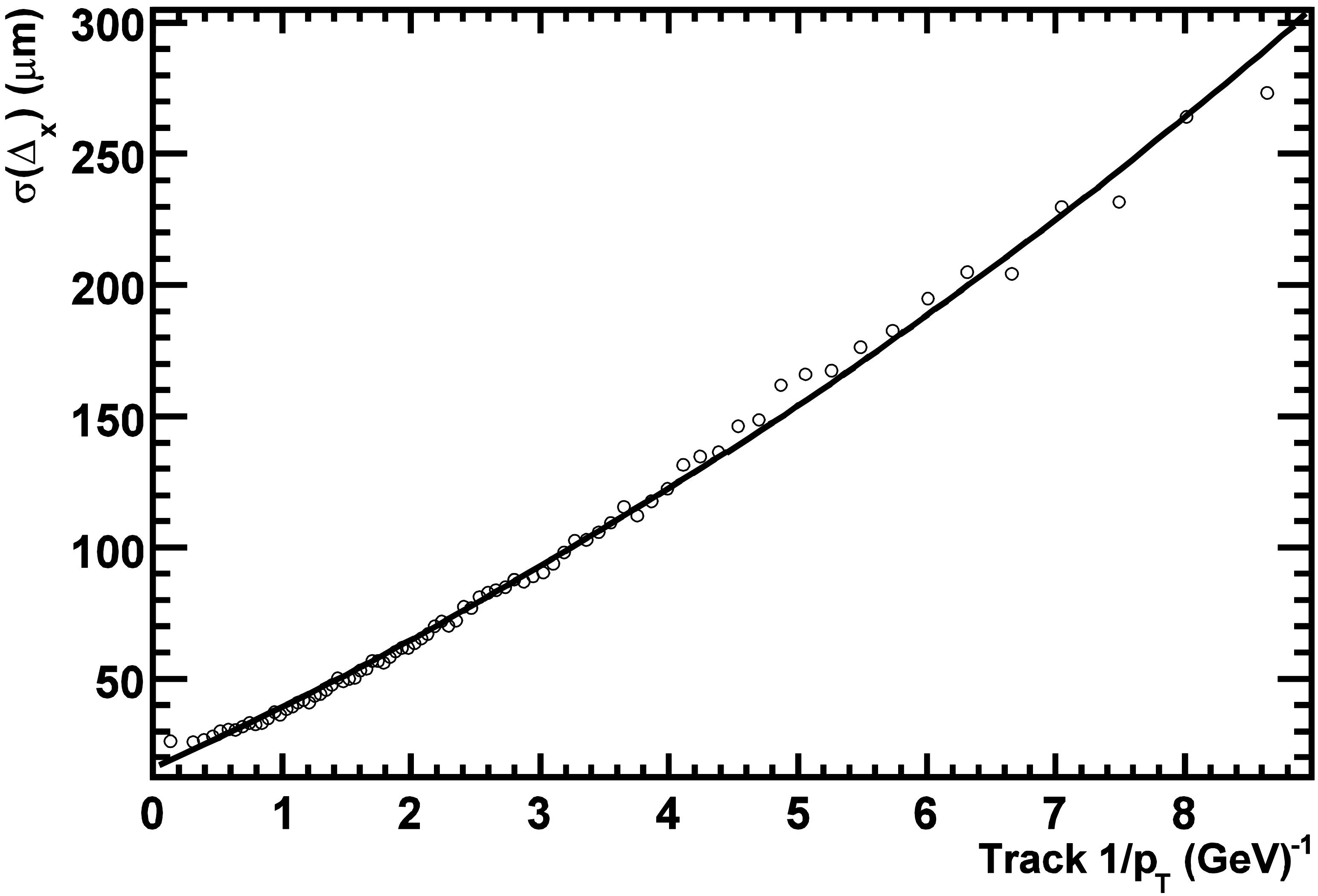
CONCLUSIONS AND FUTURE WORK

CONCLUSIONS AND FUTURE WORK

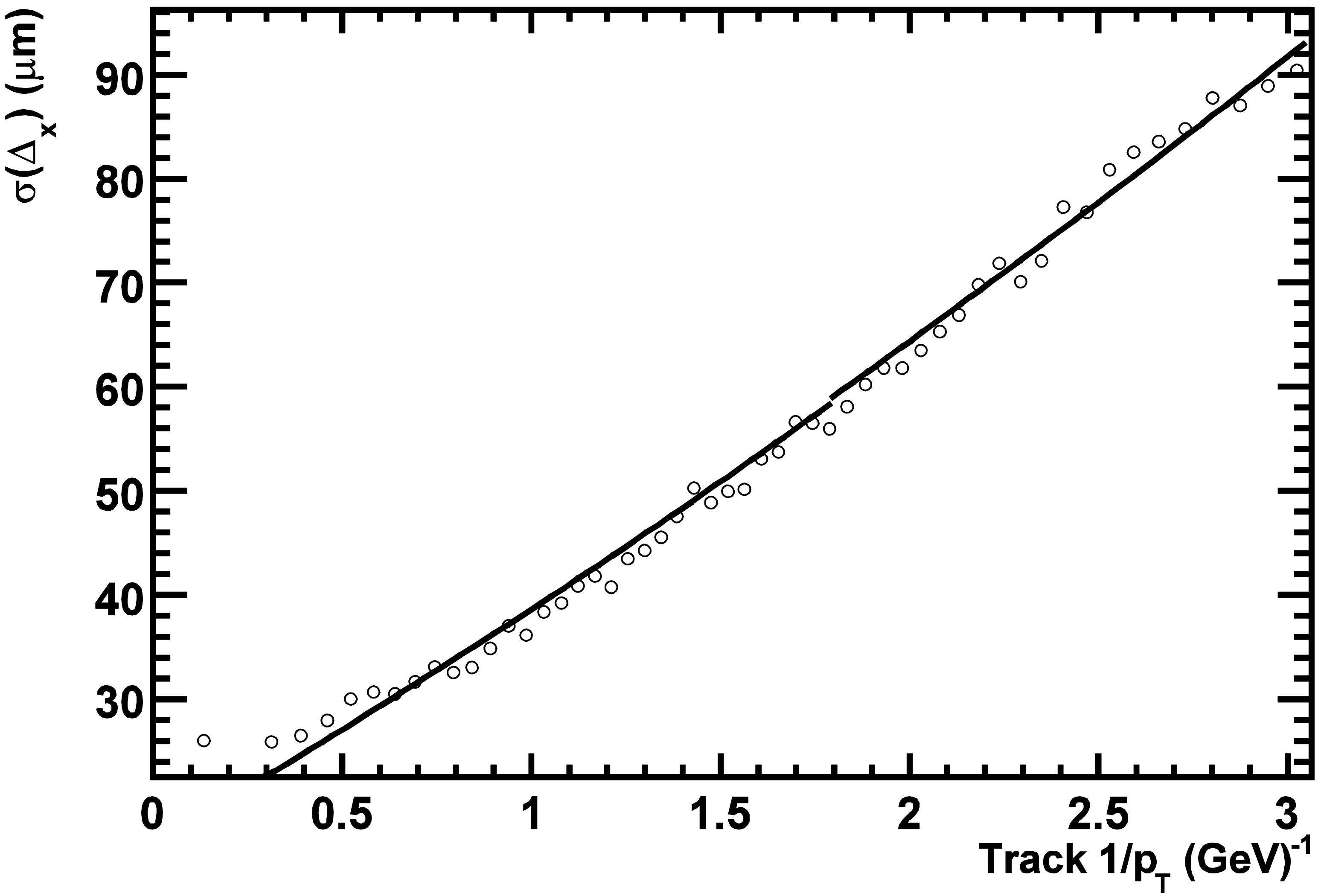
- The fitter is ready for release and public use
  - How do we envisage this to be used on real data? As a monitoring algorithm?
  - Do we want or need a look up table?
- Does the iterative procedure need refining?
  - Should switch to an unbinned fit for use with real data... sadly we now have a lot of time to work on this.
- Note detailing this work on the way soon

BACKUP

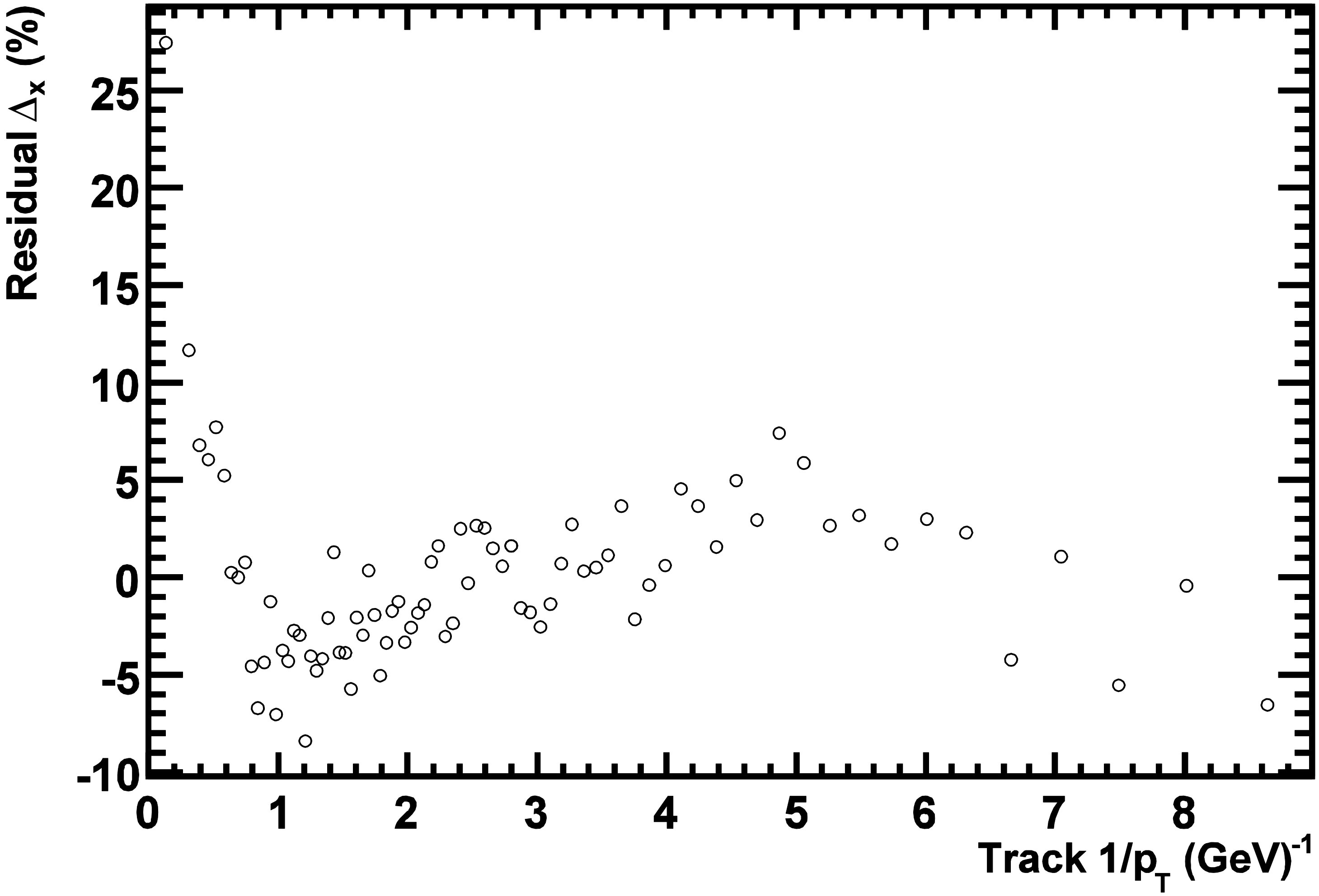
FITTING WITH A QUADRATIC



FITTING WITH A QUADRATIC (ZOOM)

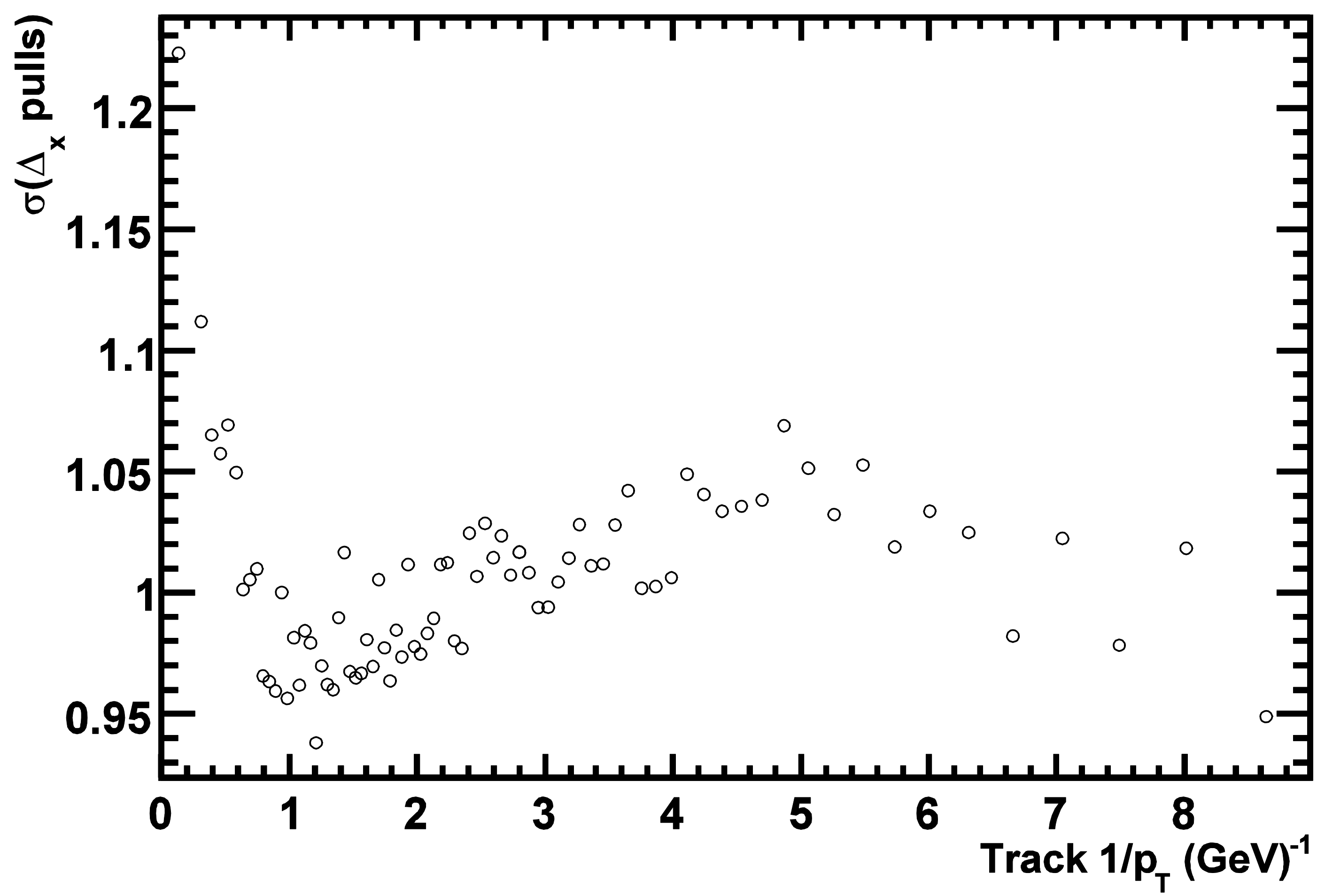


RESIDUALS WITH QUADRATIC FIT



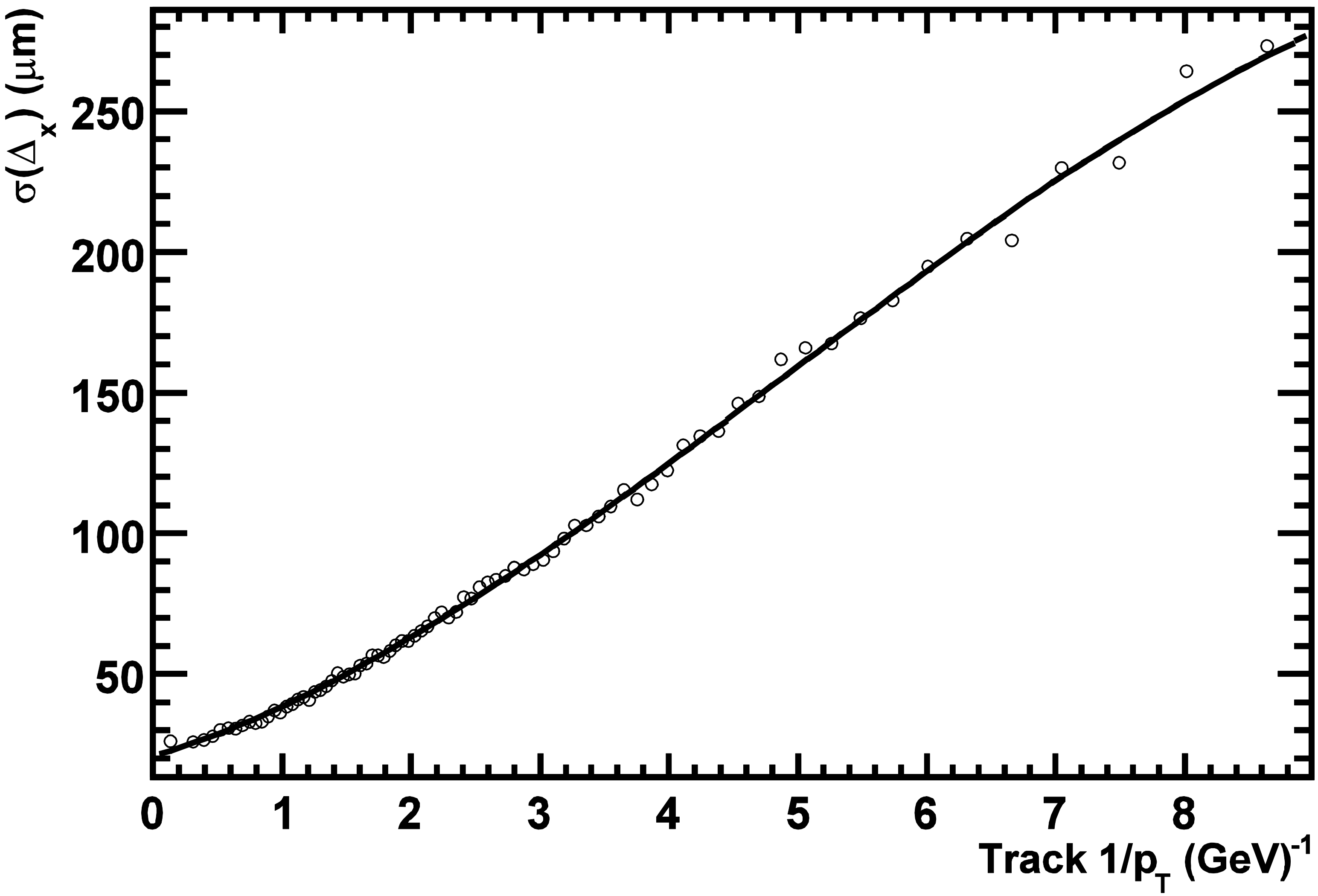
This plot shows the residual between the fitted and measured error for each bin of p_T, calculated at the midpoint of the bin.

PULLS WITH QUADRATIC FIT

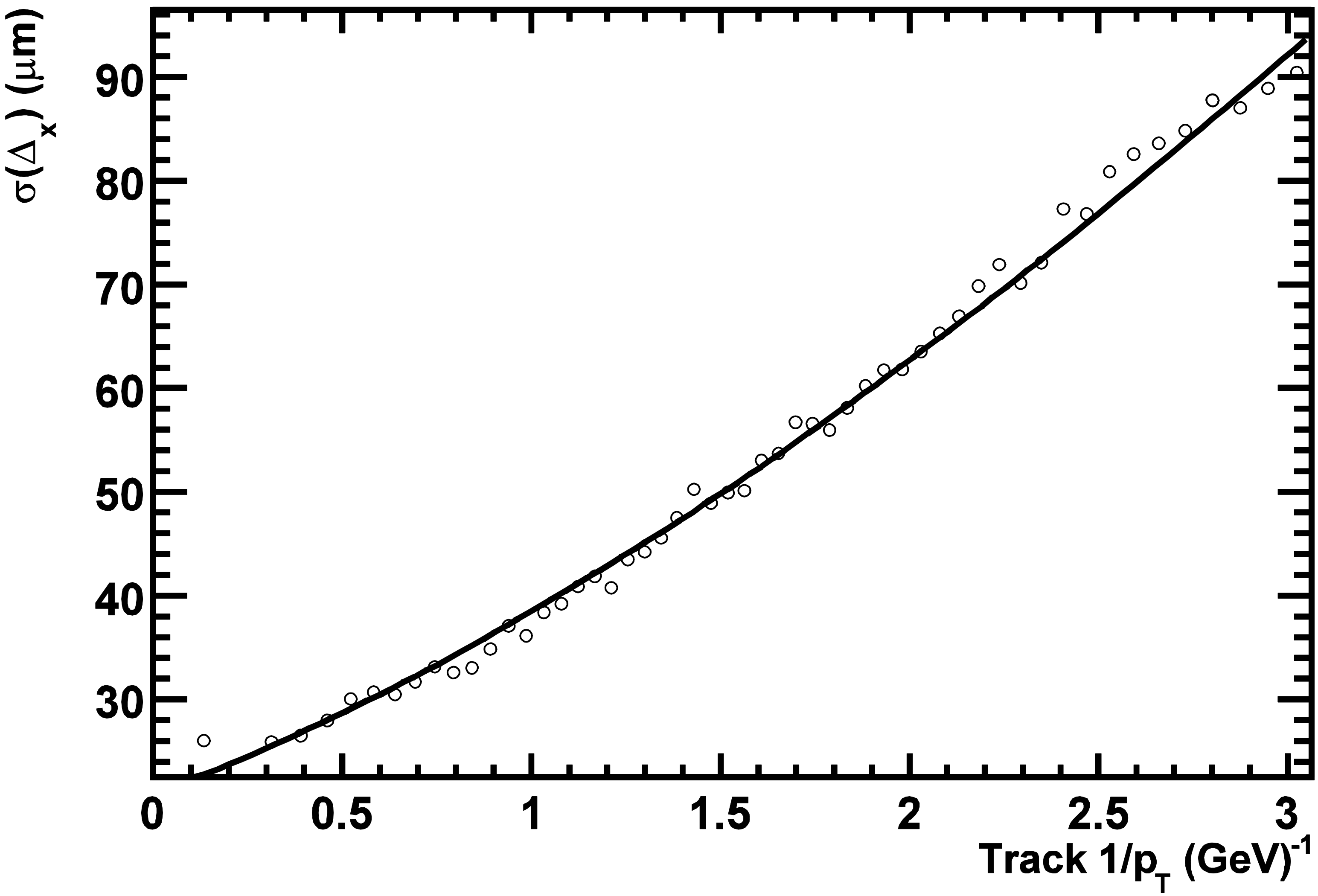


This plot shows the width of the pull in bins of p_T. The pull in any one bin is computed by dividing the measured Δ_x for every track in that bin by the parameterized error. For perfect agreement the widths should be equal to 1.

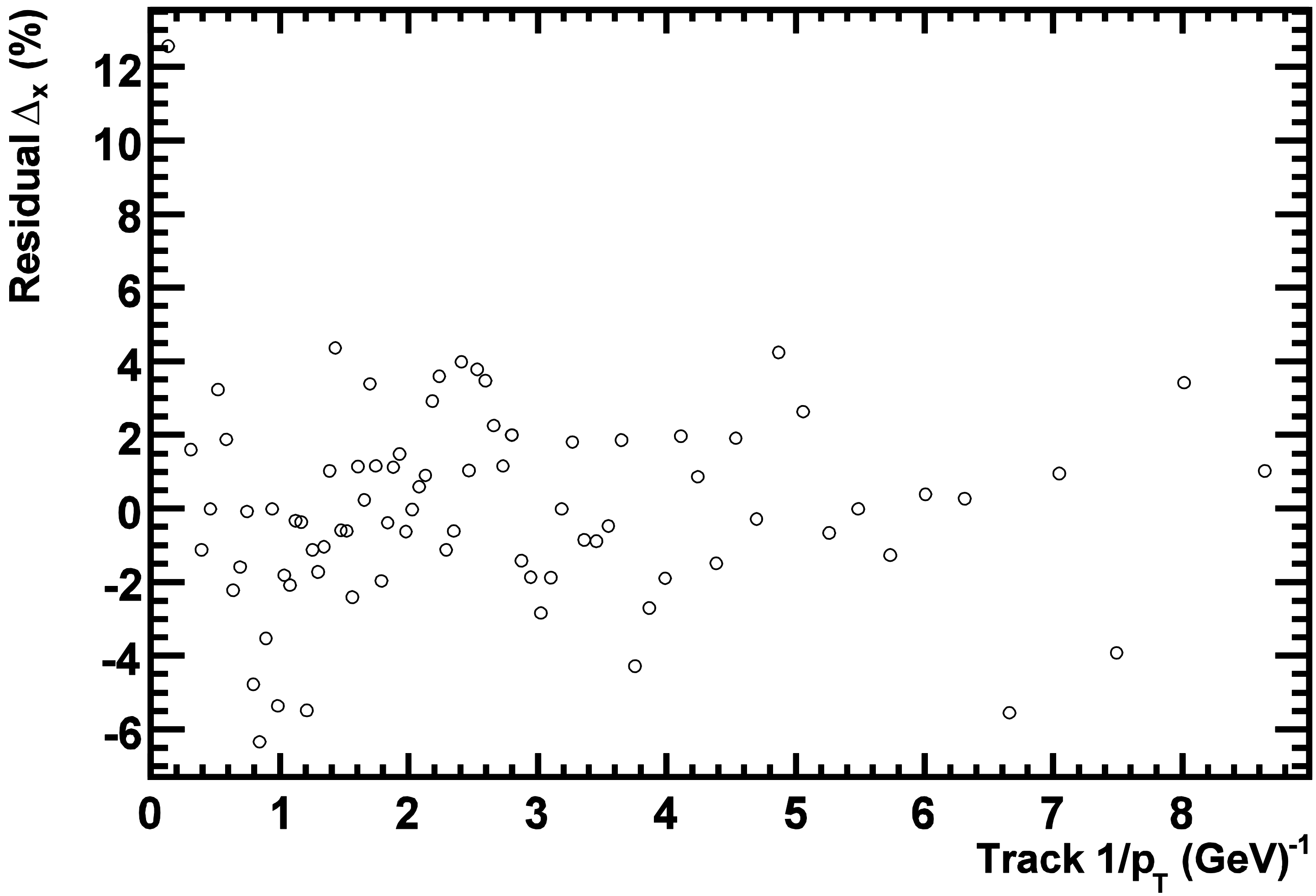
FITTING WITH A CUBIC



FITTING WITH A CUBIC (ZOOM)

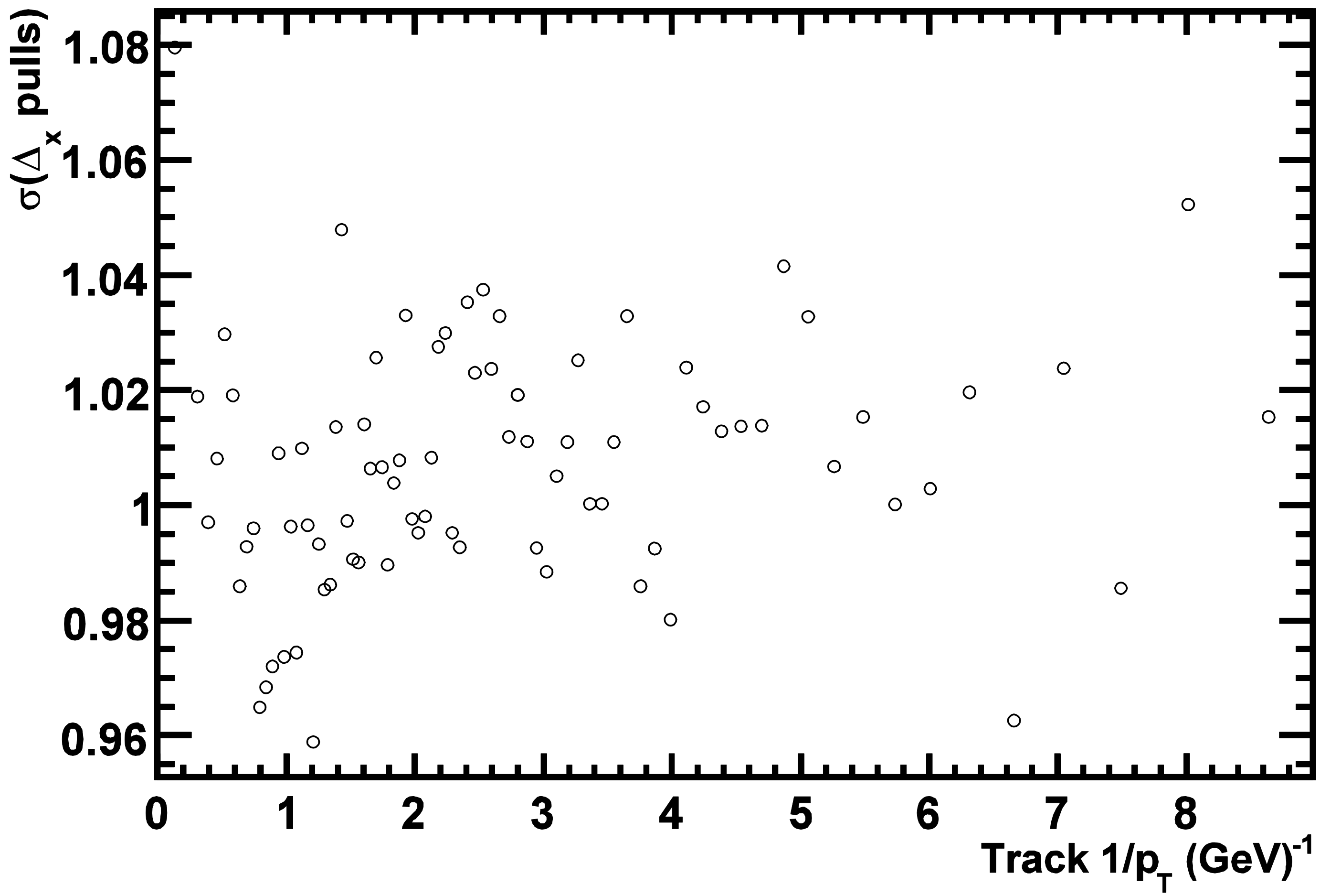


RESIDUALS WITH CUBIC FIT



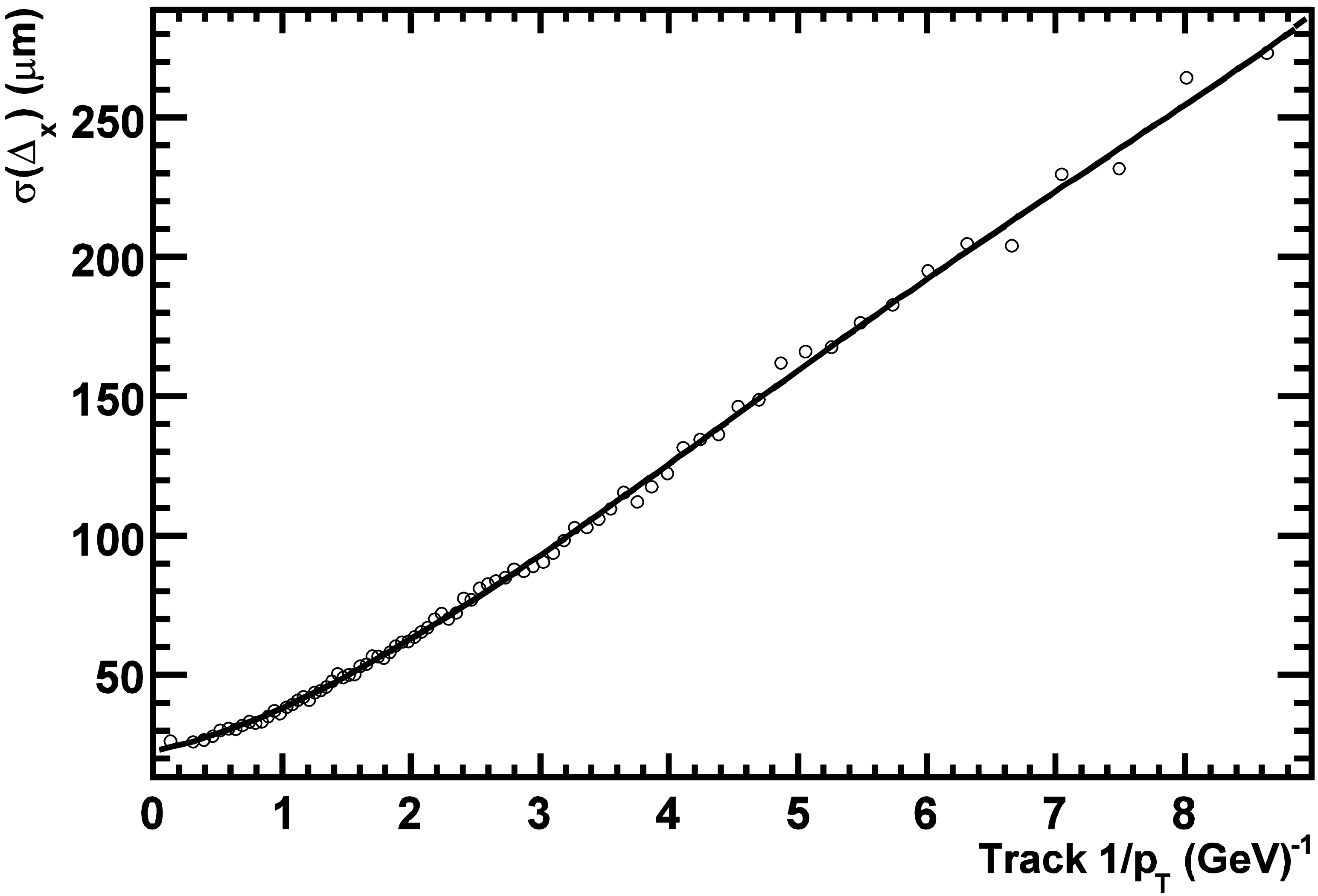
This plot shows the residual between the fitted and measured error for each bin of p_T, calculated at the midpoint of the bin.

PULLS WITH CUBIC FIT

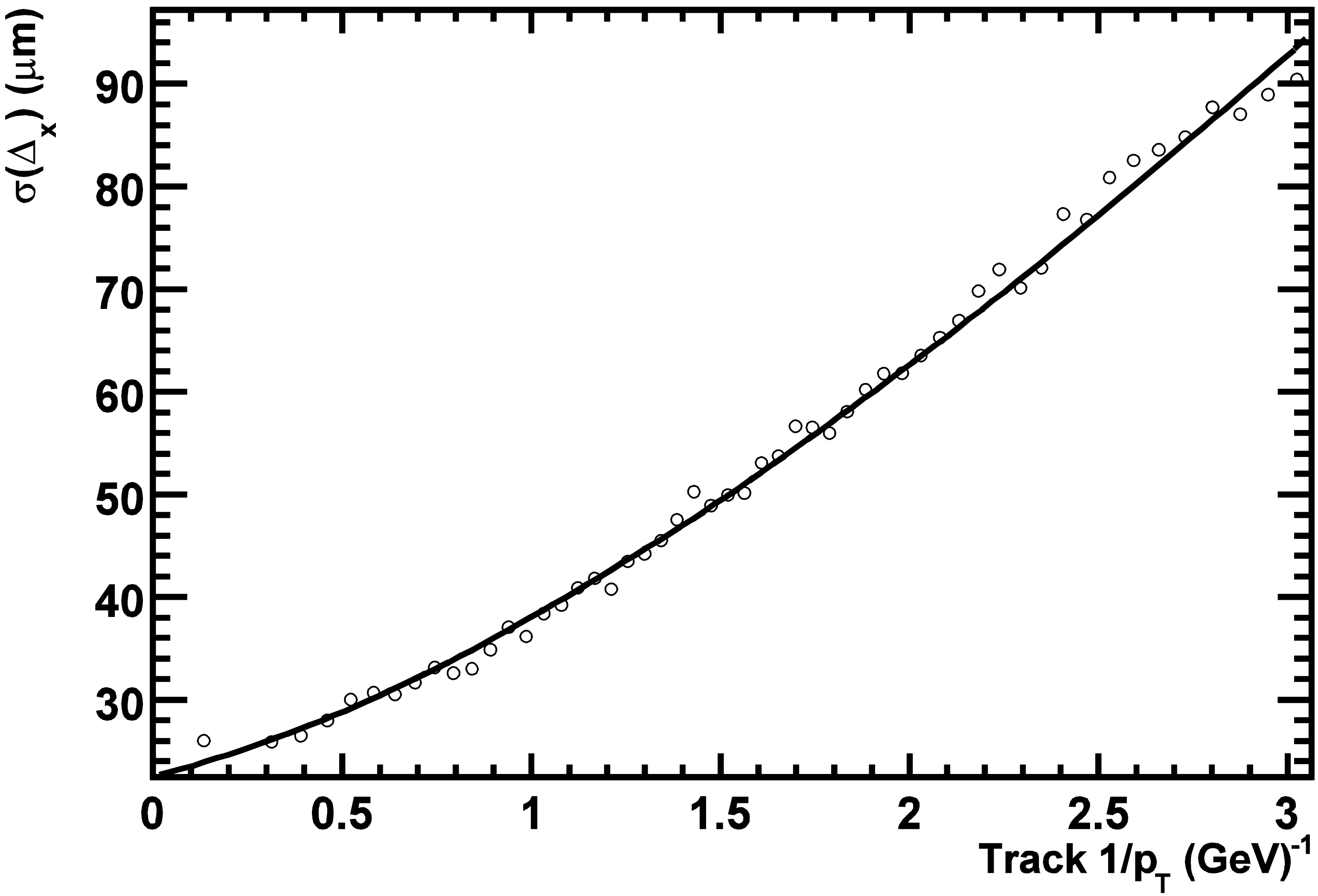


This plot shows the width of the pull in bins of p_T. The pull in any one bin is computed by dividing the measured Δ_x for every track in that bin by the parameterized error. For perfect agreement the widths should be equal to 1.

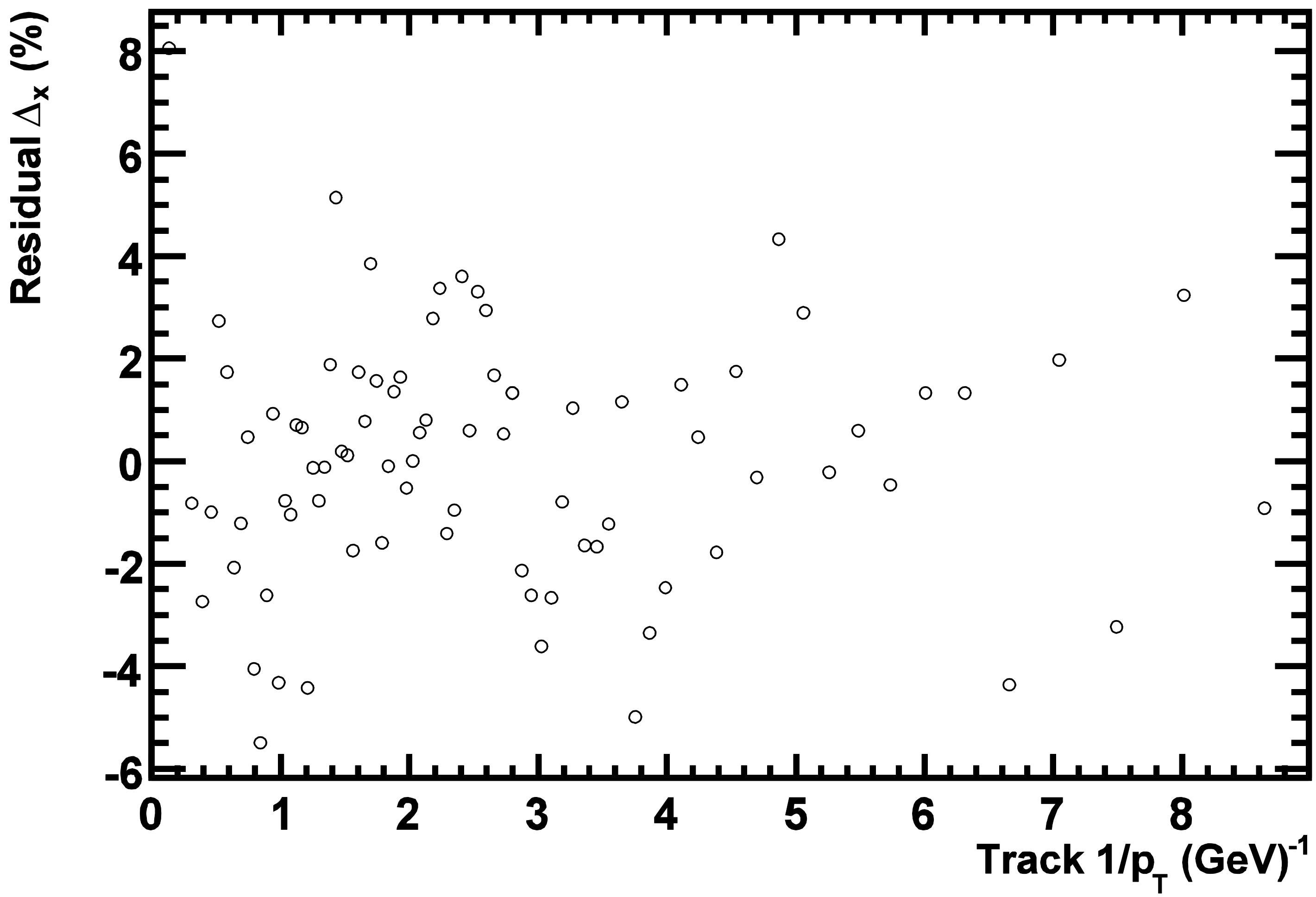
FITTING WITH A QUARTIC



FITTING WITH A QUARTIC (ZOOM)

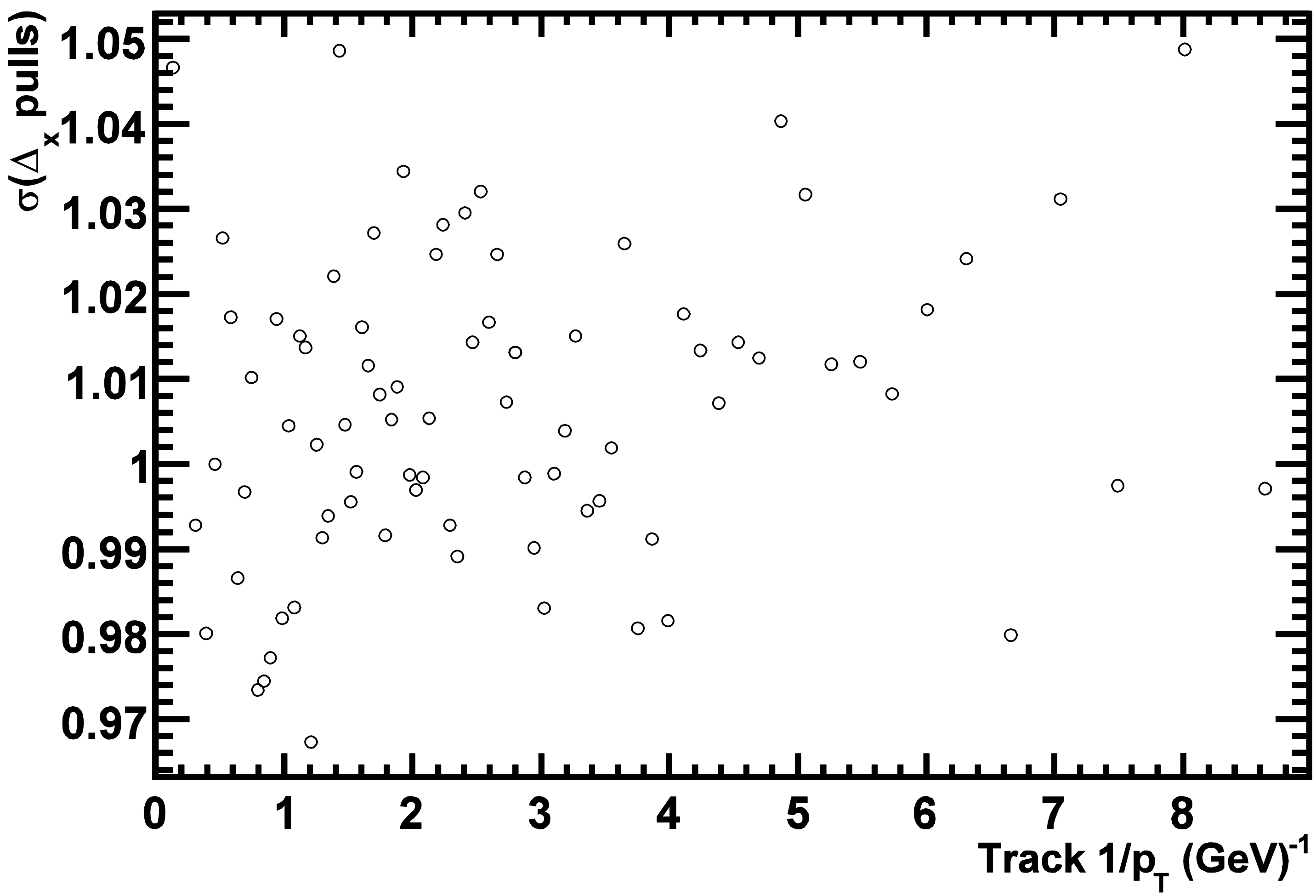


RESIDUALS WITH QUARTIC FIT



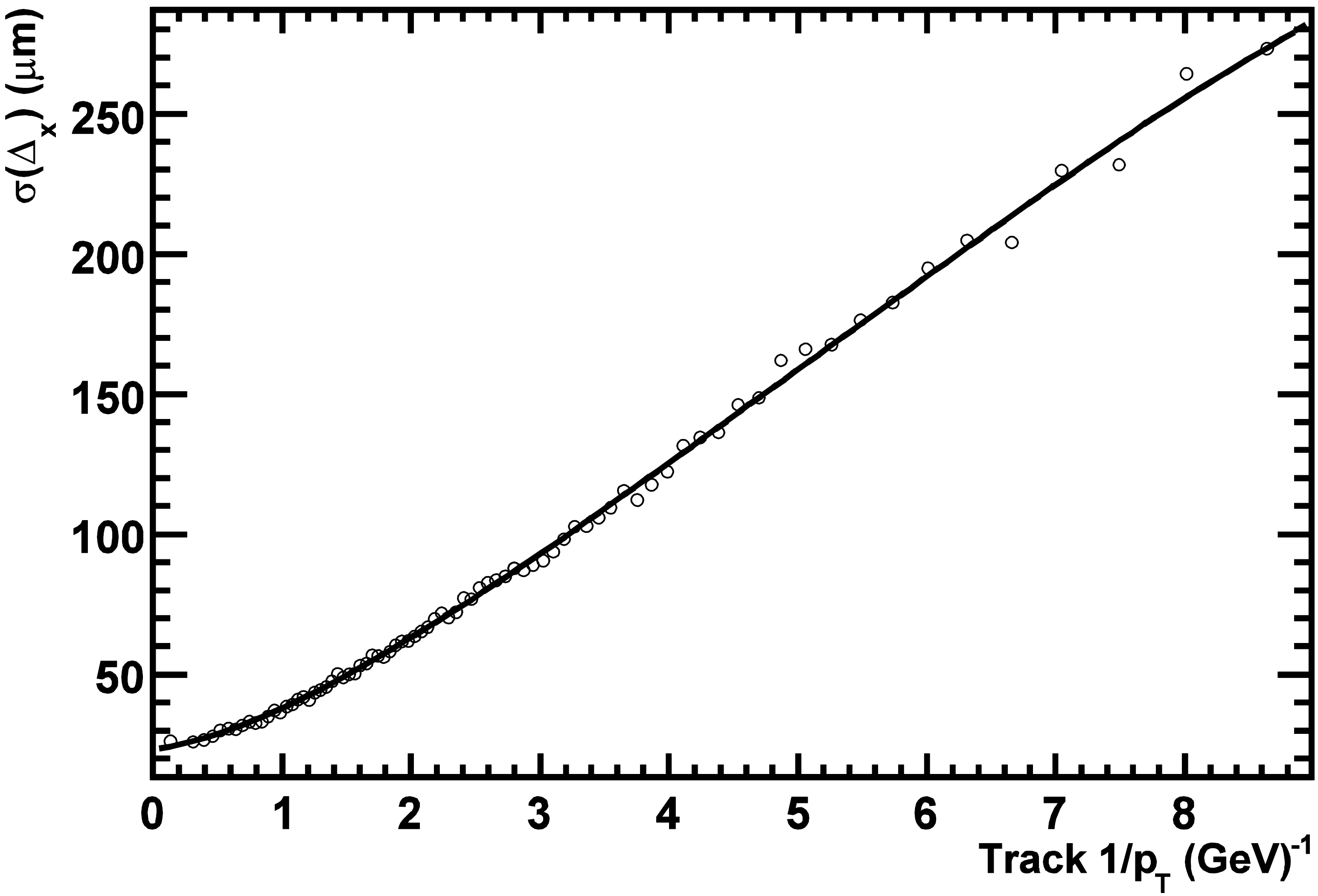
This plot shows the residual between the fitted and measured error for each bin of p_T, calculated at the midpoint of the bin.

PULLS WITH QUARTIC FIT

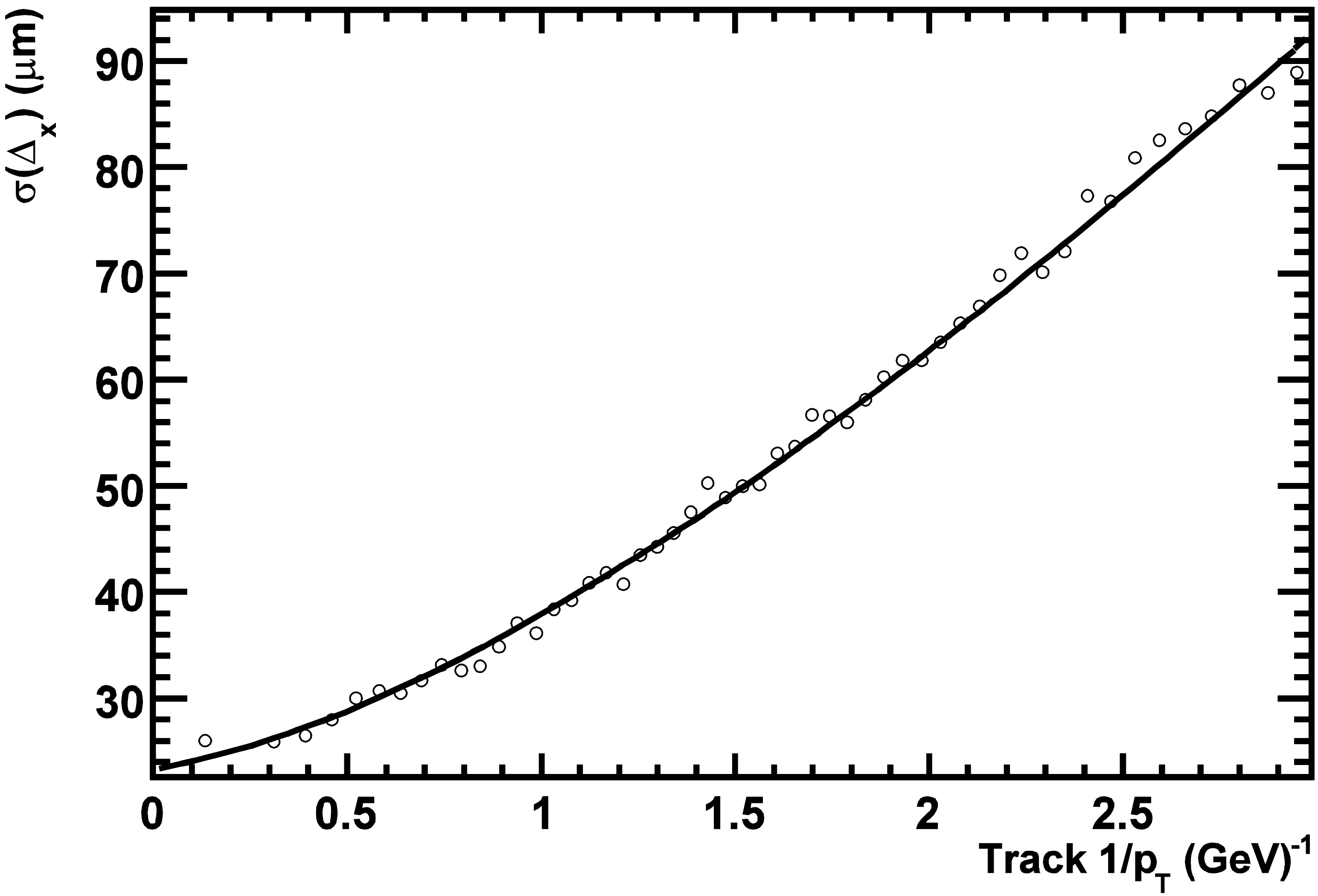


This plot shows the width of the pull in bins of p_T. The pull in any one bin is computed by dividing the measured Δ_x for every track in that bin by the parameterized error. For perfect agreement the widths should be equal to 1.

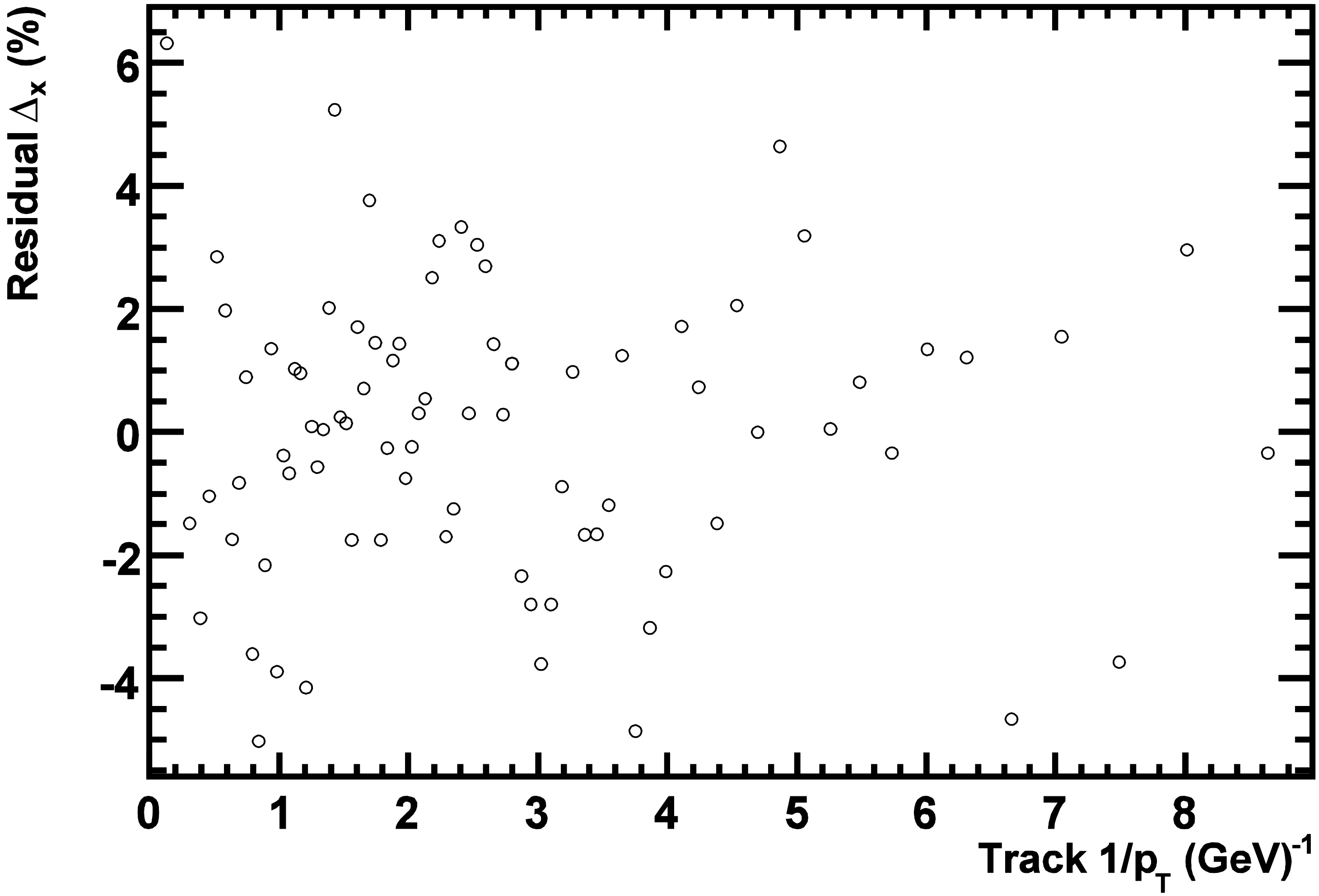
FITTING WITH A QUINTIC



FITTING WITH A QUINTIC (ZOOM)

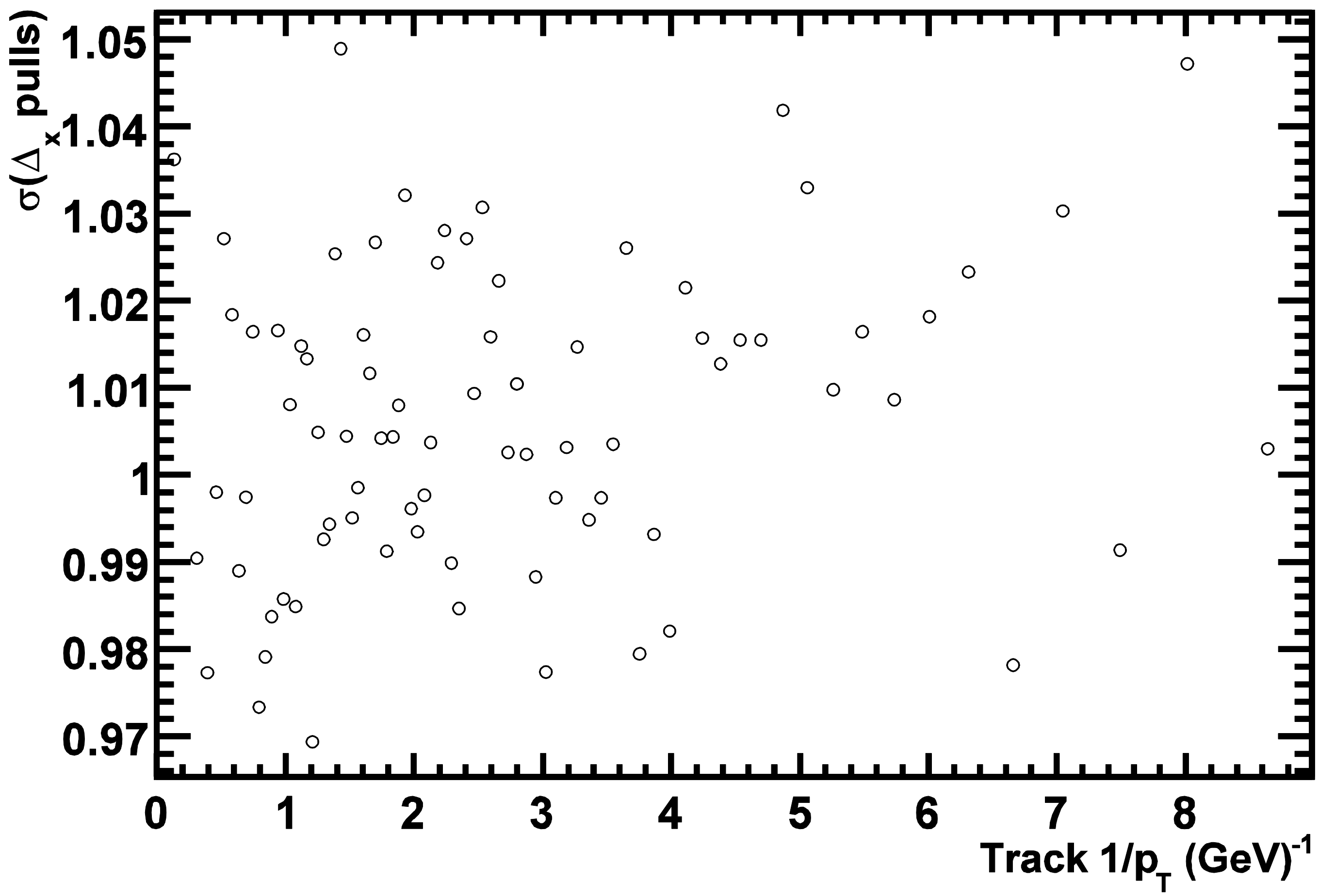


RESIDUALS WITH QUINTIC FIT



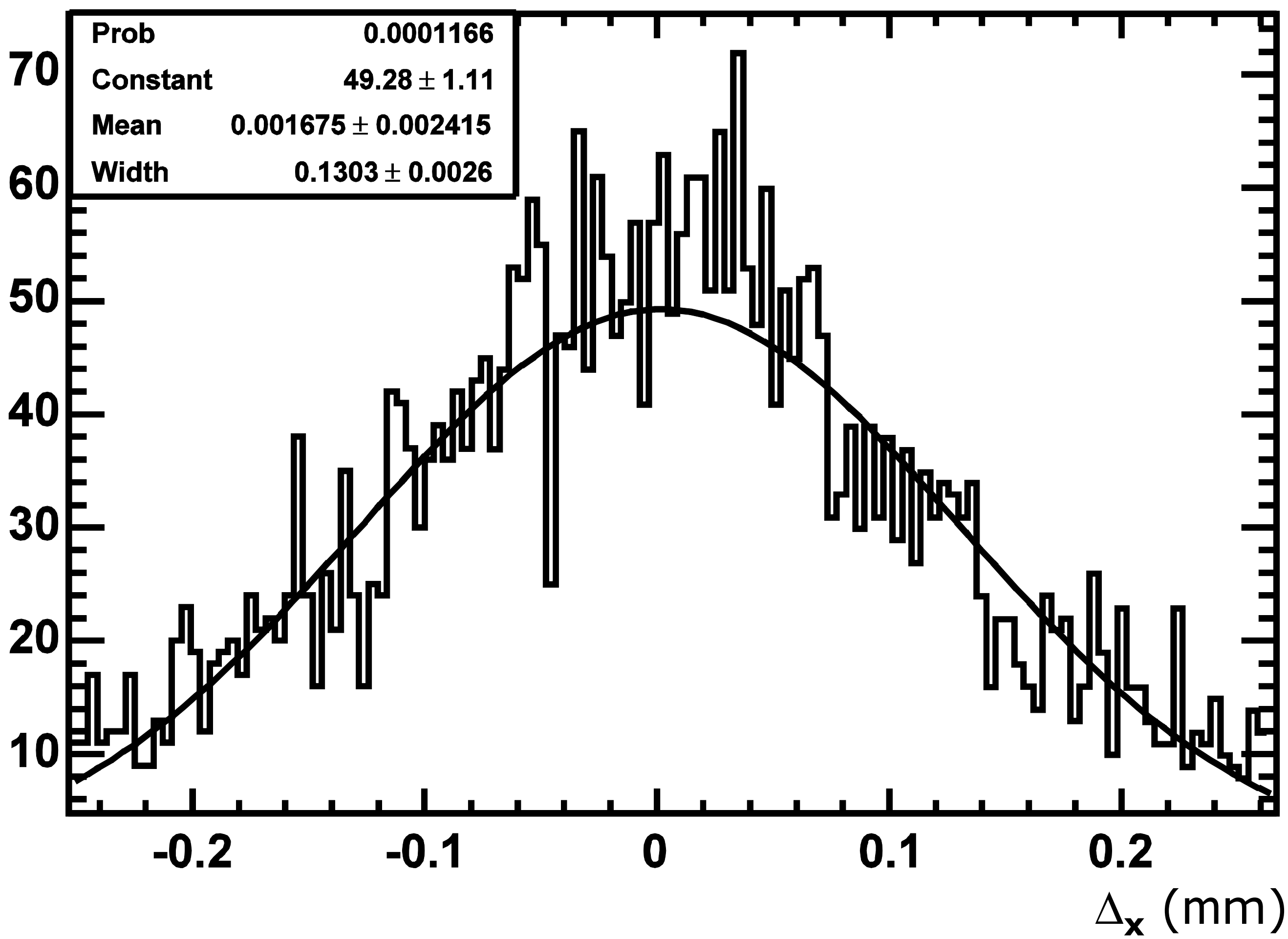
This plot shows the residual between the fitted and measured error for each bin of p_T, calculated at the midpoint of the bin.

PULLS WITH QUINTIC FIT



This plot shows the width of the pull in bins of p_T. The pull in any one bin is computed by dividing the measured Δ_x for every track in that bin by the parameterized error. For perfect agreement the widths should be equal to 1.

INTERMEDIATE 1σ GAUSSIAN FIT



DC04 RESULTS*

* Ref: Hugo Ruiz, LHCb 2005-012

The parametrization

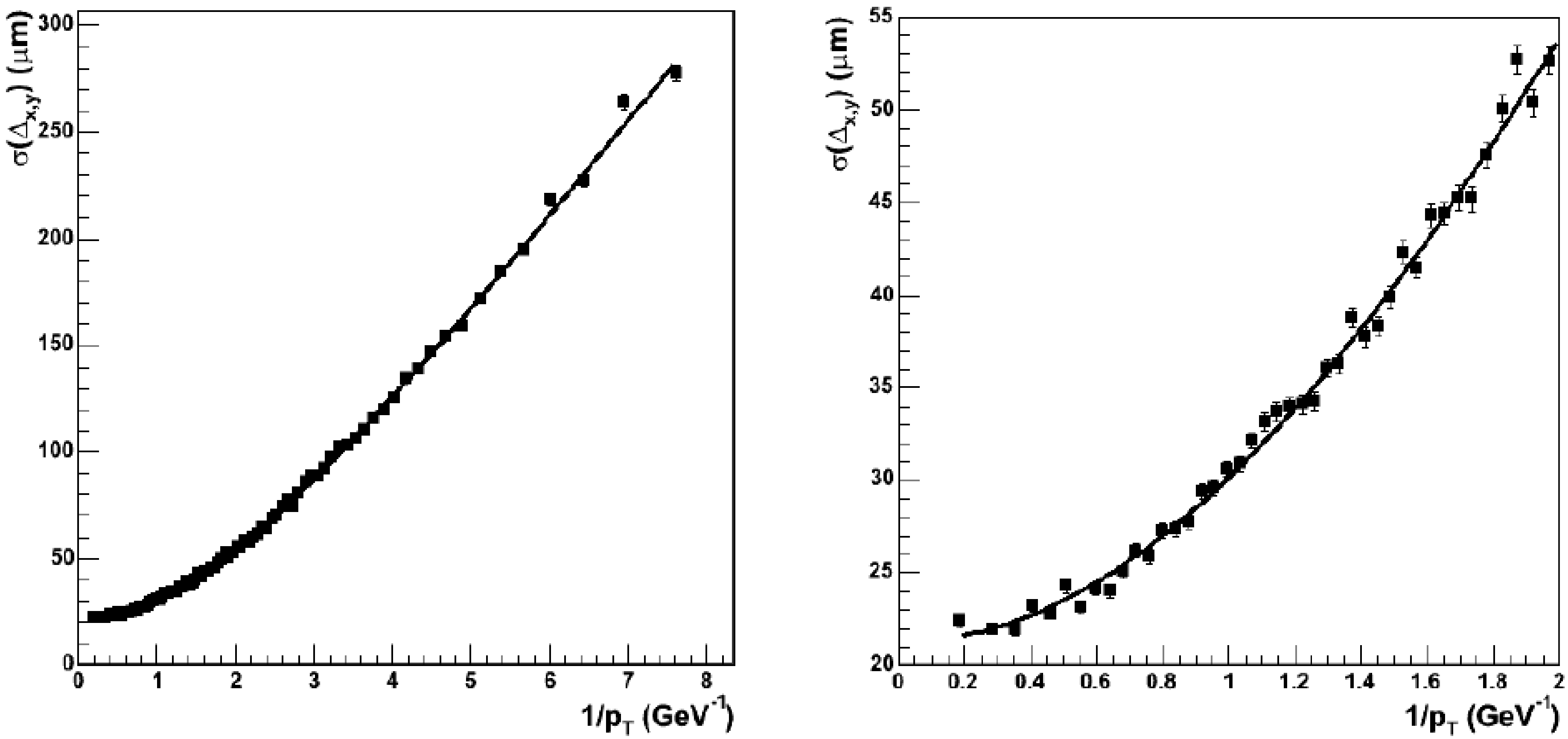


Figure 4: The best-fitting 4th-degree polynomial in two different ranges of $1/p_T$.

Residuals

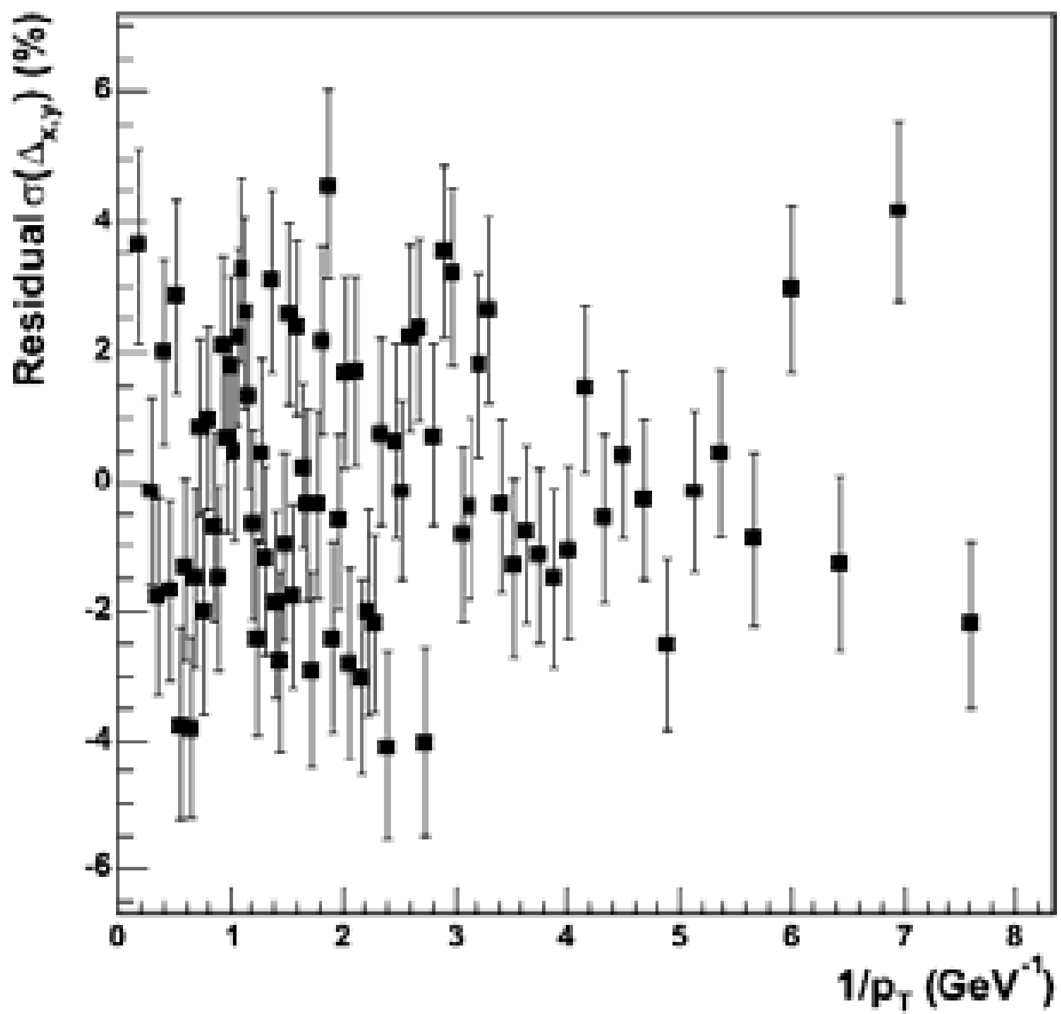


Figure 3: Residuals of the parameterization of the uncertainties

Tracks left out of the fit

- Because of the iterative procedure, some tracks are left out of the fit

  - The proportion varies with p_T

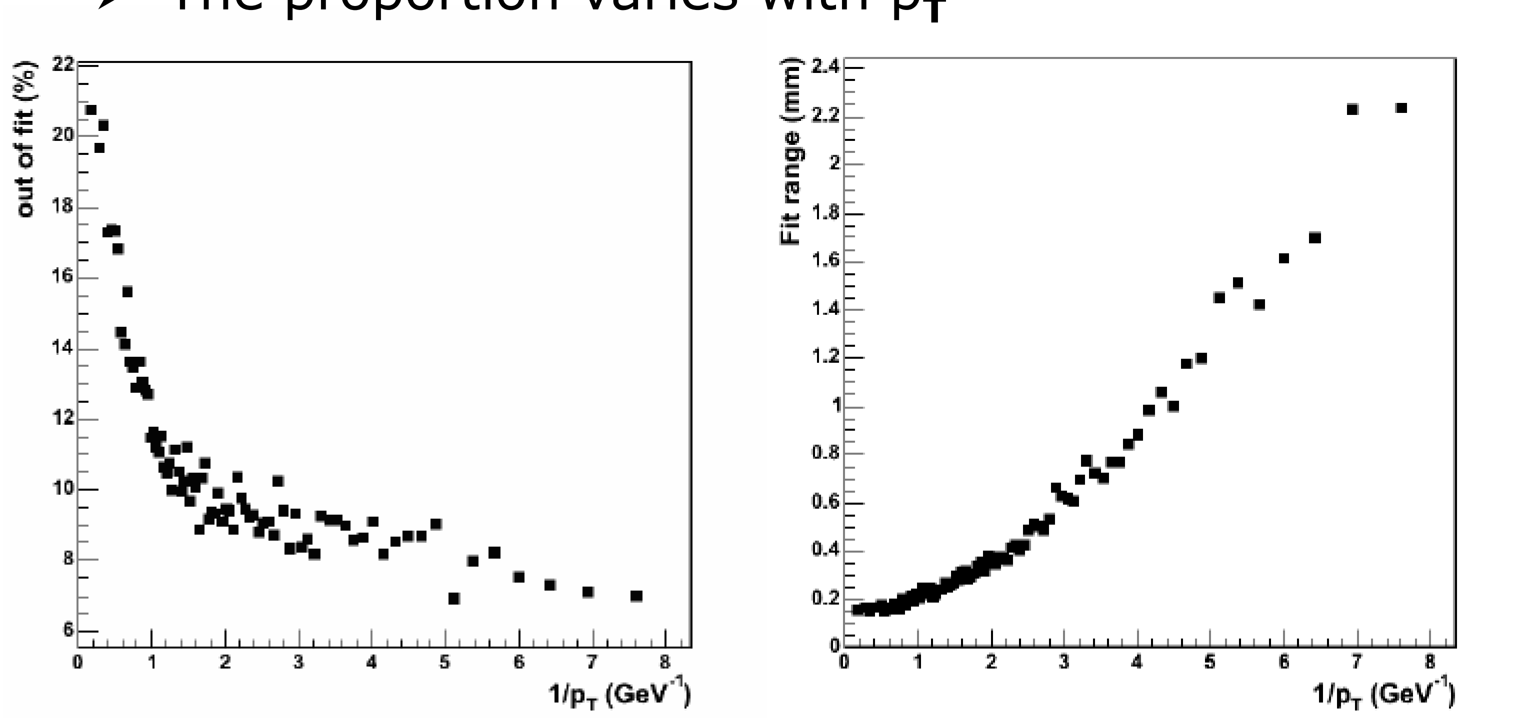


Figure 5: Left: fraction of tracks left out of the final fit due to the iterative process for fitting the σ in a $\pm 3\sigma$ region. Right: size of the region where the final fit is performed.