Advanced track reconstruction algorithms

Kalman filter limitations

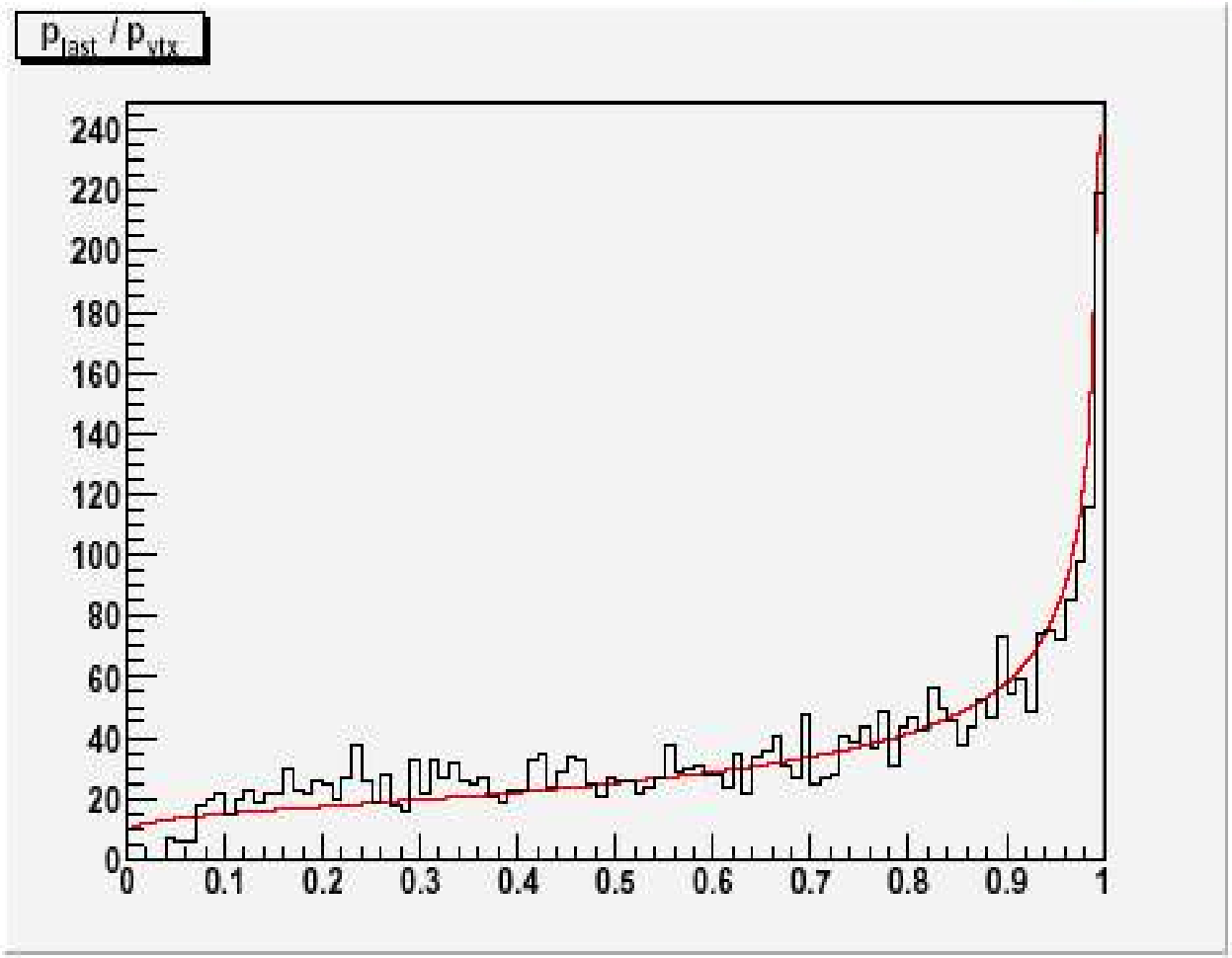
- The Kalman filter is an optimal estimator of track parameters for a purely Gaussian problem
- Reality is non-Gaussian, and the Kalman filter cannot take that into account
  - in reality it is not an optimal estimator
- The classical Kalman filter suffers from lack of robustness in the presence of noise (RecHits not from the track)

GSF

- The Gaussian Sum Filter was implemented to address the problem of non-Gaussian measurements and material effects.
  - it is impossible to treat in a computationally efficient way arbitrary probability density functions
  - but arbitrary functions can be approximated as sums of Gaussians
  - since the Kalman filter is linear the fit of a sum of Gaussians is a sum of Kalman filters

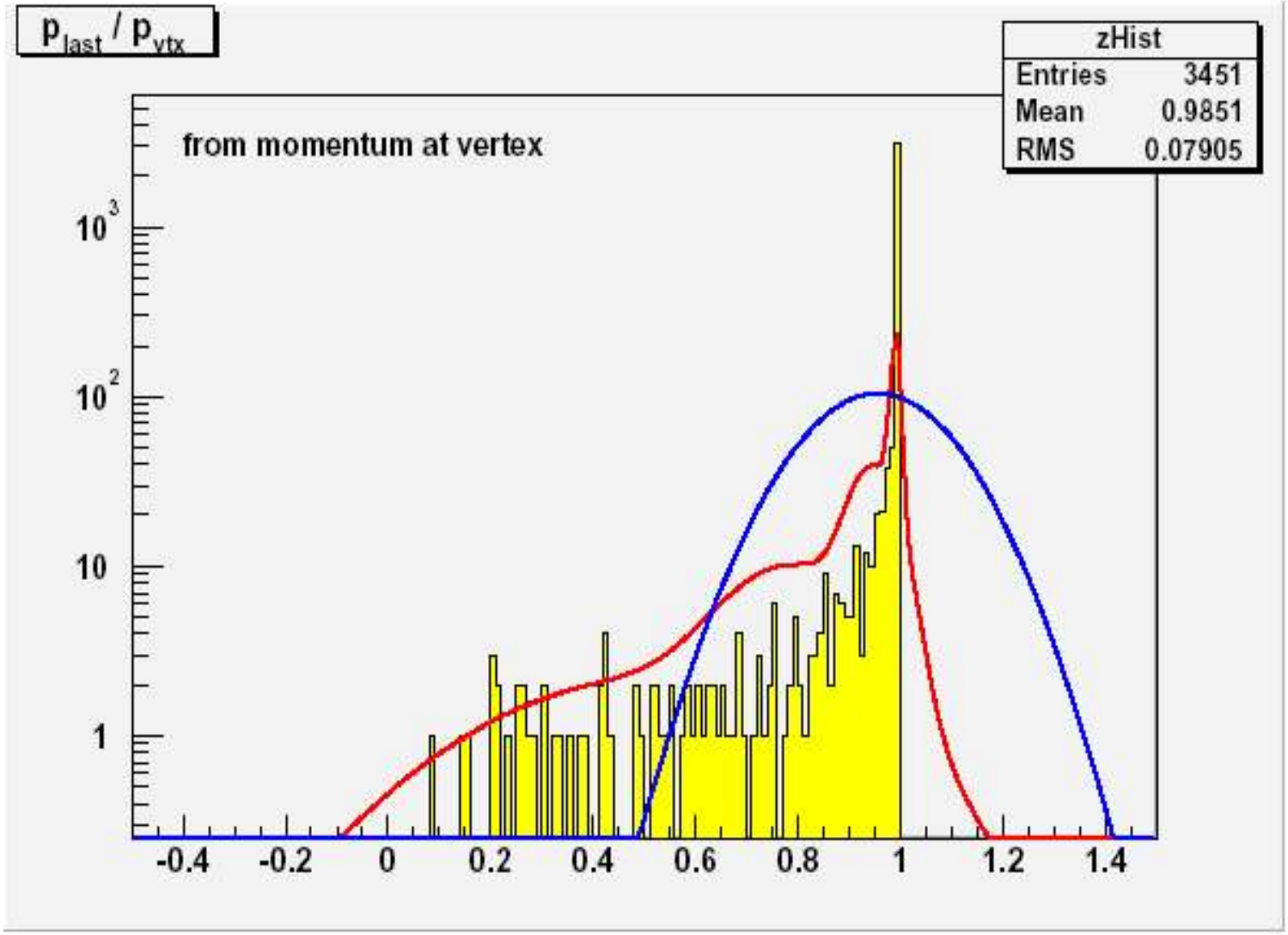
Electron energy loss

Observed distribution for full track and estimate from Bethe-Heitler using $\sum dx_i/x_0$.



Fractional energy loss distribution

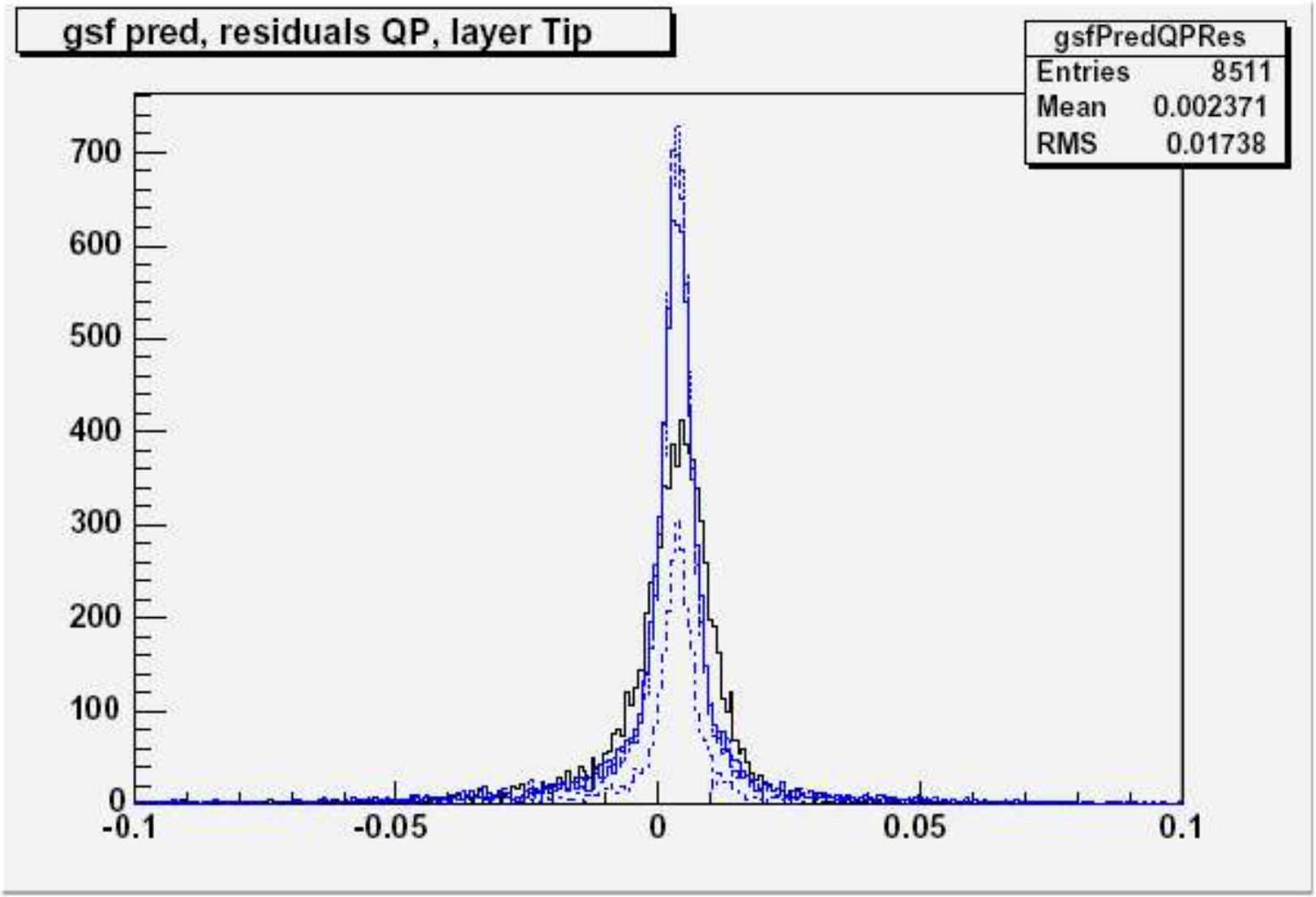
Observed distribution for single layer with KF (blue) and GSF (red) estimate.



Results at p_t = 10 GeV

Q/P residuals for full simulation:

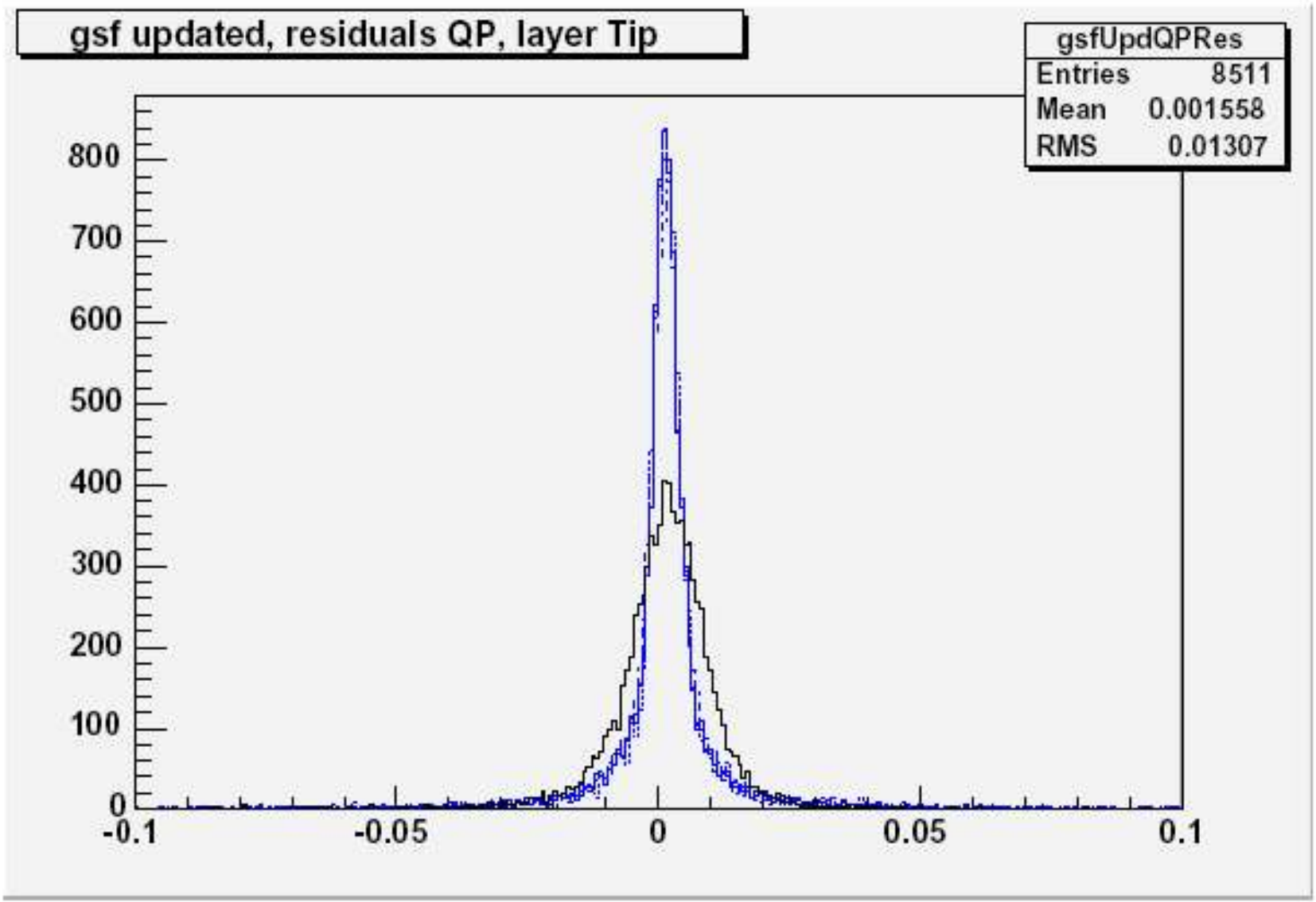
KF (black) vs. GSF (blue) with different number of components



Results at p_t = 10 GeV

Q/P residuals for full simulation with vertex constraint:

KF (black) vs. GSF (blue) with different number of components

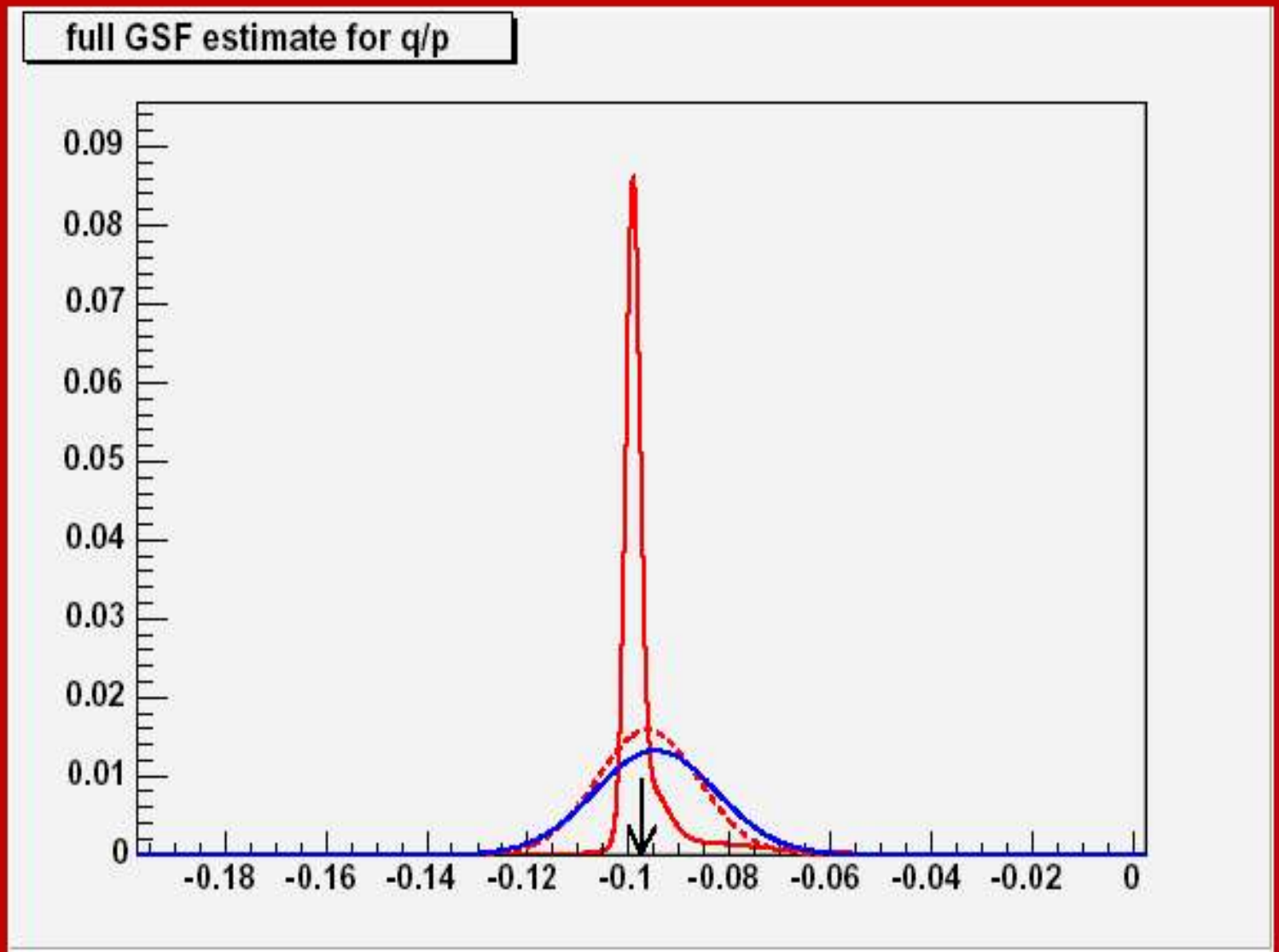


GSF performance for electrons

- The GSF gives significant improvement of resolution for the core part of the residual distribution
- The tails are not reduced
  - as a consequence the RMS is not reduced either
- The GSF is most effective for low energy electrons (a few tens of GeV), almost no gain for 100 GeV Pt.

Some events at p_t = 10 GeV

q/p at transverse impact parameter plane: KF (blue), GSF (red) and simulated value (black)



$z = 26\%$ (first event $\leq 30\%$)

Individual track PDF

- The GSF gives a complete 5D probability density function for each track (a sum of a configurable number of Gaussians)
  - A single gaussian approximation to this PDF loses a lot of information
- Wherever the track is used, the pdf should be used explicitly (in impact parameter probability, kinematic fit,...)
  - already used in vertex fit

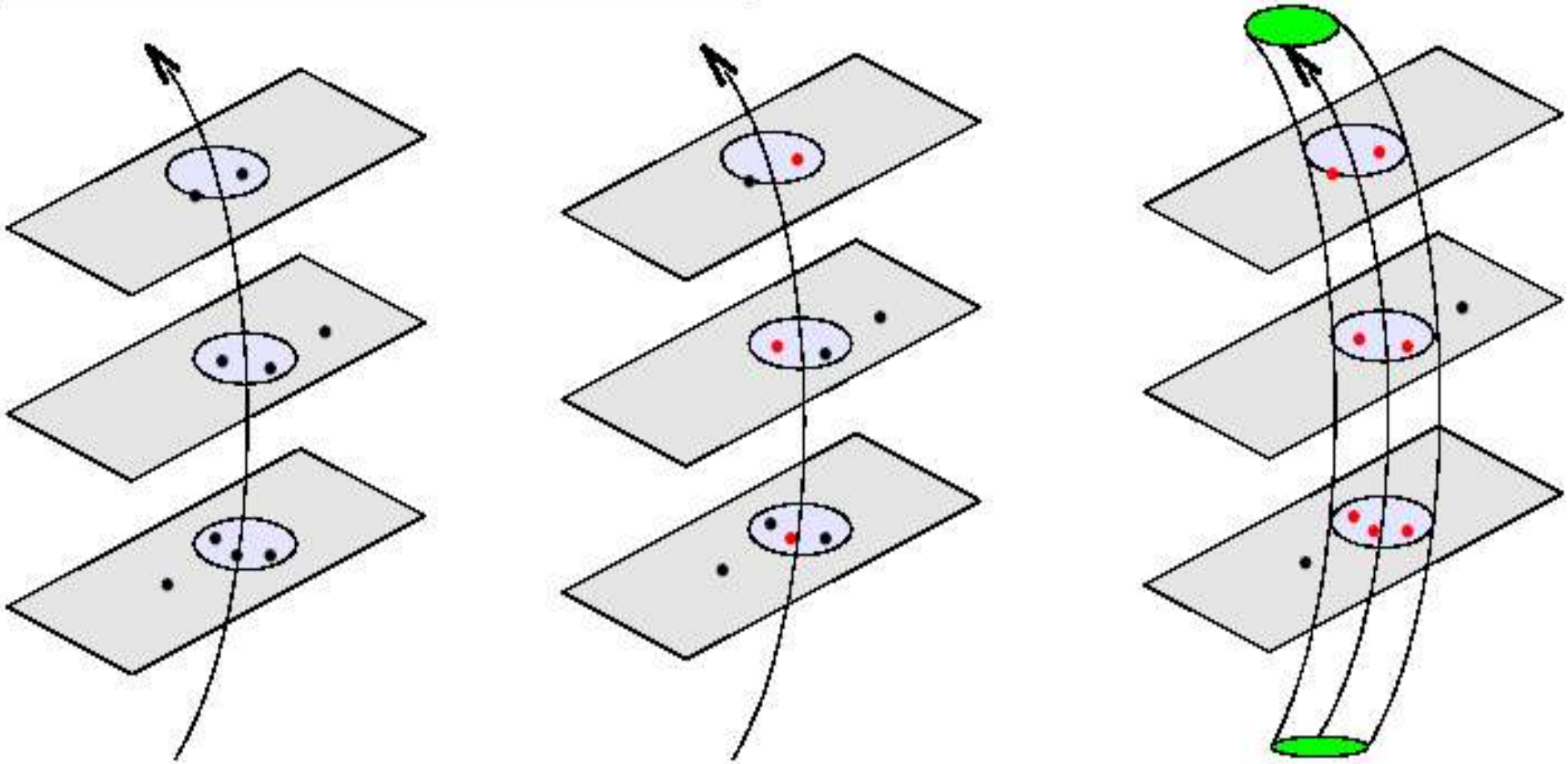
GSF status

- Currently the GSF is extensively validated for electrons, with multi-Gaussian energy loss
- The GSF treatment of multiple scattering is implemented, but the results are not very encouraging
- GSF treatment of multi-Gaussian RecHits is planned
  - may “solve” the problem of non-gaussian pixel RecHit residuals
  - track by track pdf should be better than global calibration of IP probability
    - and does not require channel-specific calibration

DAF

- The track reconstruction in dense environment is degraded by
  - increasing probability to assign wrong hits to the track
  - RecHit degradation due to contamination by other tracks or noise
- The Deterministic Annealing Filter was implemented to assign the hits to a track in an optimal way
  - soft assignment iterative method

Motivation: What is the DAF?



KF:
- **hard hit−track assignment** (1 or 0)
- explicit pattern recognition (combinatorics)

DAF:
- **soft hit−track assignment** (1...0, "assignment probability")
- explicit + implicit pattern recognition (combinatorics + annealing)

What do we expect? **better hit−to−track assignment −> better track parameters**

DAF performance

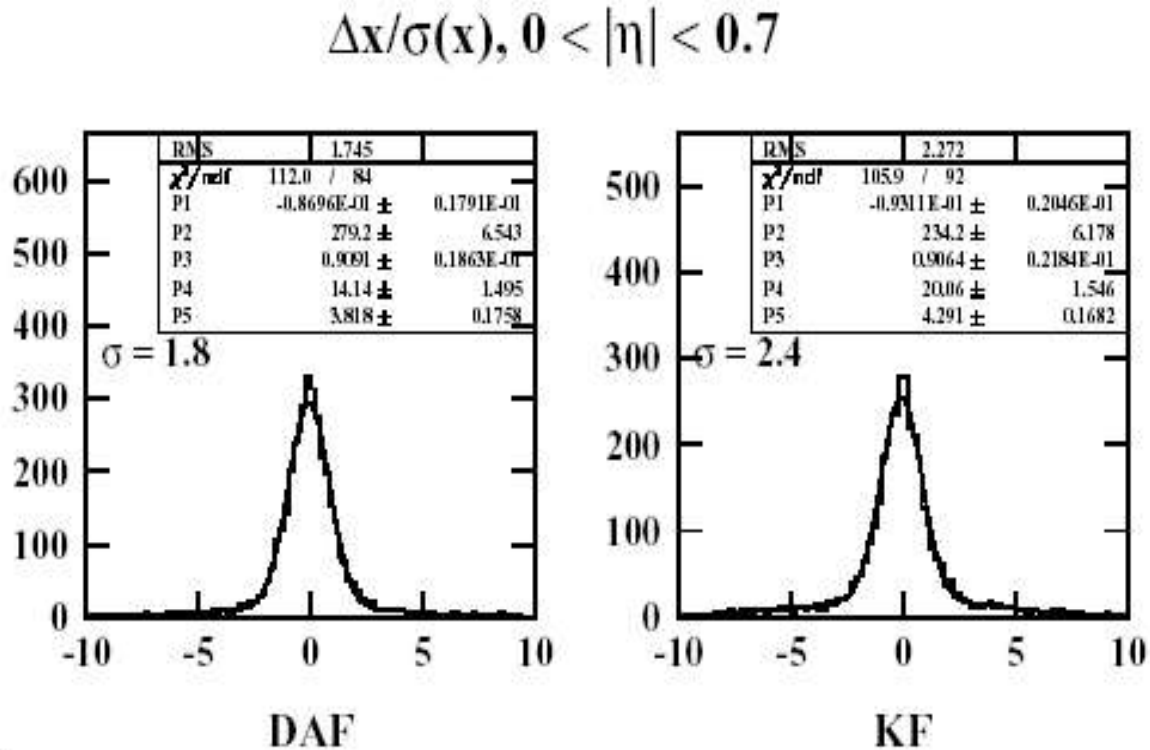
- For isolated tracks, even at high luminosity, the DAF does not provide a measurable improvement in track quality
- To see an effect one has to go to dense environments, e.g. core of a 200 GeV jet or 3-prong high Pt tau

Tracks in b−jets: example

Selection:

bbar Et: 200 GeV
|eta|: 0−0.7
RecTracks: pt > 15 GeV/c

example: transverse impact parameter pulls



$$\sigma^2 = \frac{\sigma_1 P2}{\sigma_1 P2 + \sigma_2 P4}\sigma_1^2 + \frac{\sigma_2 P4}{\sigma_1 P2 + \sigma_2 P4}\sigma_2^2$$

P1 = common mean

P2 = height Gauss1 P3 = σ_1

P4 = height Gauss2 P5 = σ_2

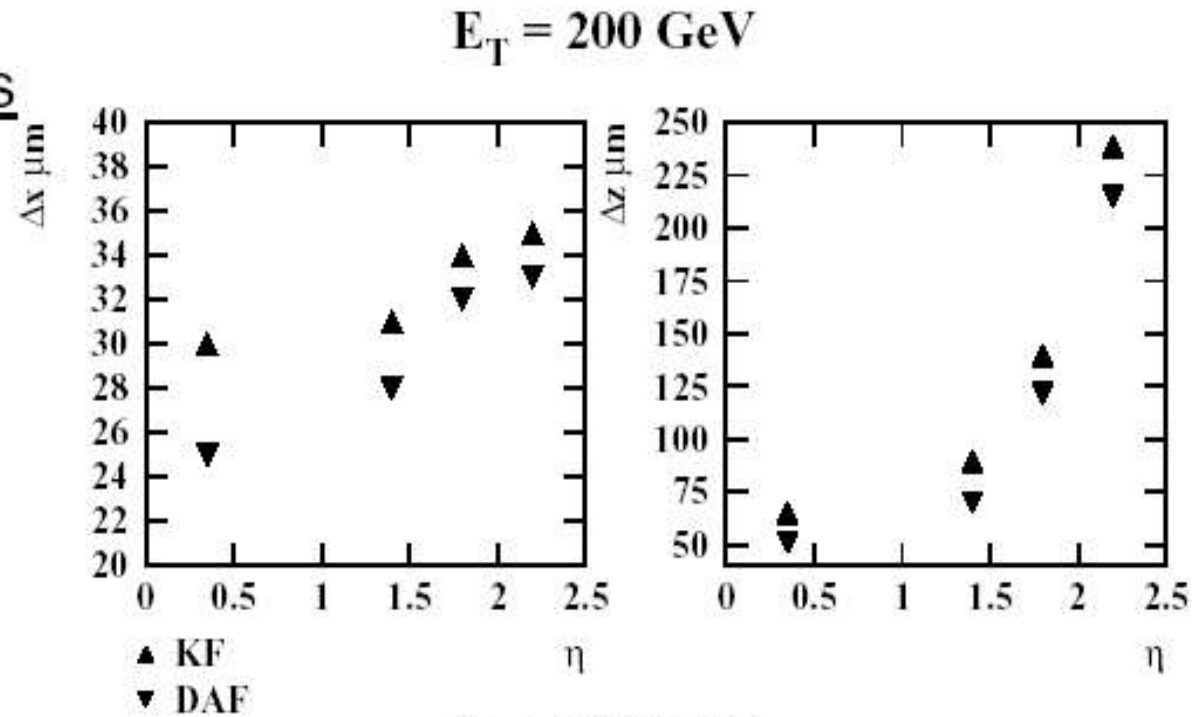
Difficult environment: high−pt tracks in high−Et b−jets (dense track environment)

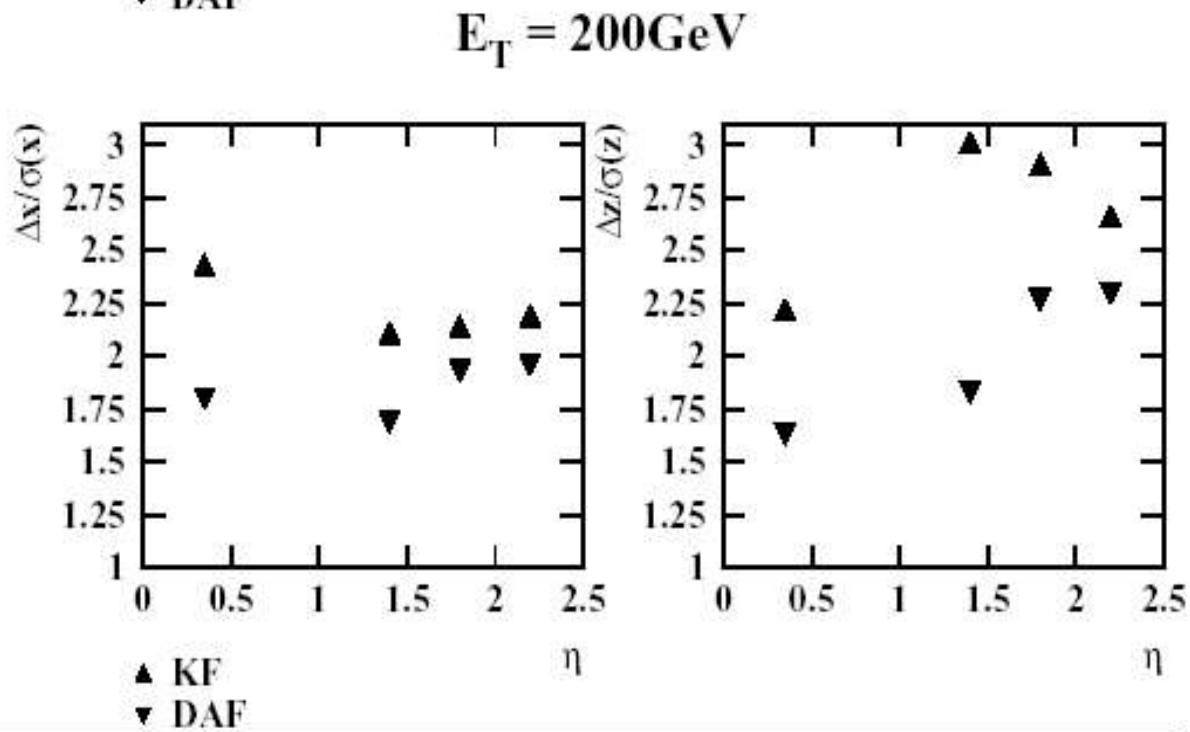
b−jets: impact parameters

Selection:

bbar Et:	200 GeV
|eta|:	0−0.7, 1.2−1.6, 1.6−2.0, 2.0−2.4
RecTracks:	pt > 15 GeV/c

DAF has better ip resolutions and pulls

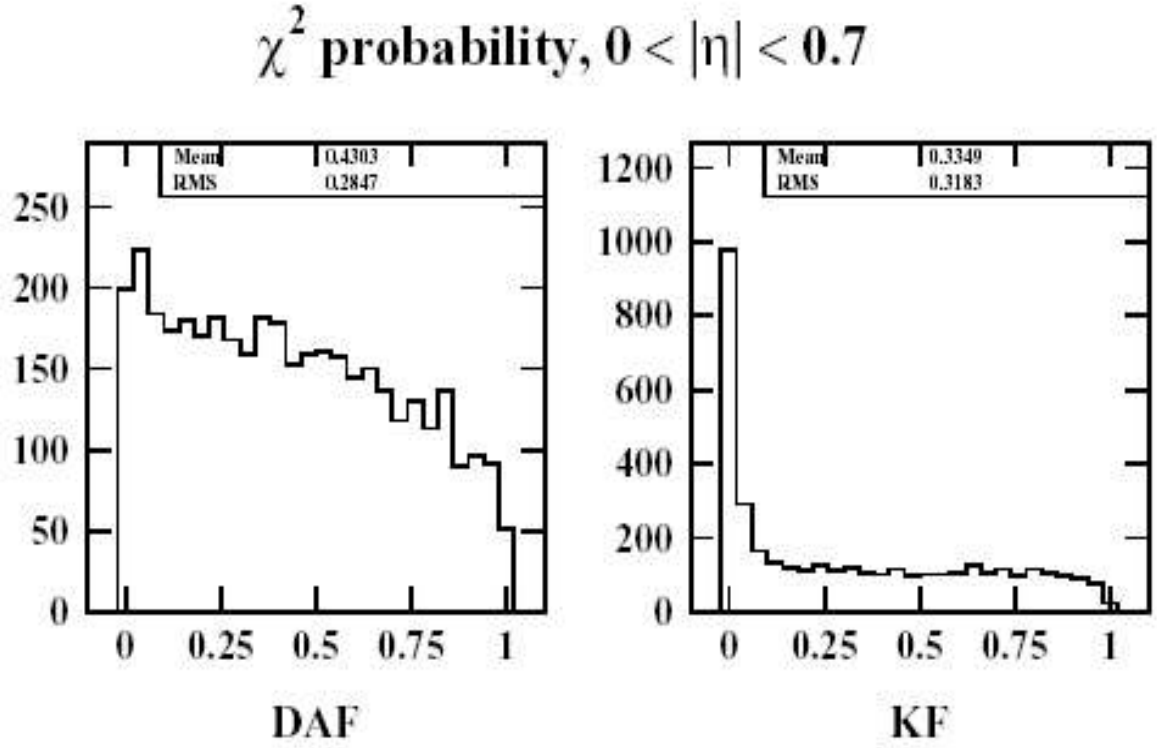




b−jets: track χ^2 probability

Selection:

bbar Et: 200 GeV
|eta|: 0−0.7
RecTracks: pt > 15 GeV/c



DAF has better track quality than KF

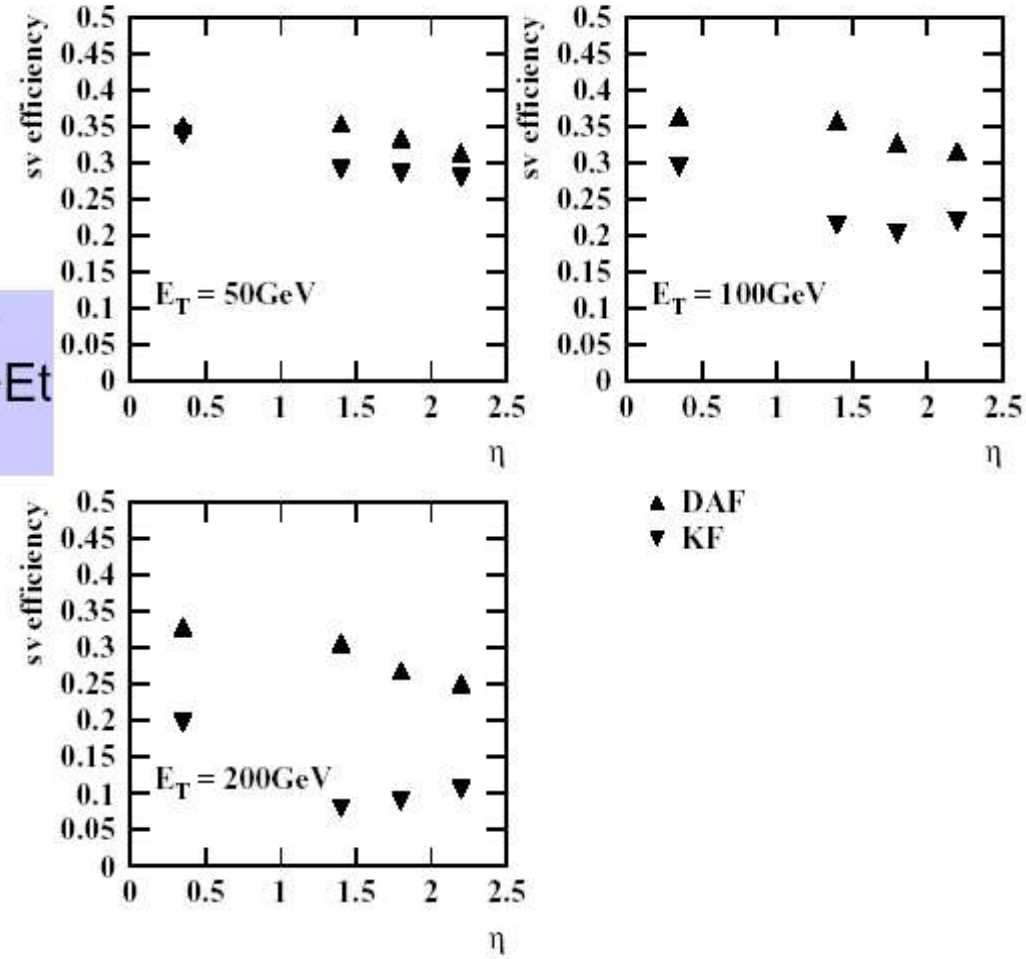
Old study: DAF was mature, but vertexing not should be repeated

b−jets: SV finding efficiency

using Pascal's .cpp

DAF: no change in SV−efficiency with jet−Et

KF: varies a lot



DAF efficiency

- The DAF needs an initial hit collection, which should contain ideally only one track, to operate
- Currently the hit collection is done around a Kalman filter track
  - the DAF efficiency is equal, but cannot be better, than the KF efficiency
- Better ways to seed the DAF are needed, but this is a serious project (Ph.D. student needed)

DAF status

- After Matthias Winkler finished his thesis the DAF has been maintained “as is”
  - just ported to each new version of ORCA
  - not yet a RecAlgorithm
- The DAF package should be significantly cleaned up
  - to make use of new functionality, like grouped measurements
  - to extract the MultiRecHits for use by other algorithms (GSF)
  - for RecAlgorithm

Other algorithms

- We are in a sad situation
  - All advanced algorithms we thought about are already implemented
    - the multi-track filter is super – but we don't have the track densities to need it, even in 500 GeV 3-prong tau
  - CMS is the only collaboration AFAIK with a GSF and a DAF implementation (but our friend Are Strandlie is now in Atlas...)
  - A possible bright spot is the idea of Christian to reconstruct secondary interactions to validate short tracks
    - seemed impossible at LHC, but now it may be doable
    - only way to improve efficiency for early interacting tracks

Summary

- We need to make the advanced algorithms default where appropriate, and use them for physics reconstruction
  - they are much slower, should only be used where they give a real advantage