Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at LHC

W. Adam (HEPHY Vienna),
R. Frühwirth (HEPHY Vienna),
A. Strandlie (CERN),
T. Todorov (IReS Strasbourg)

Outline

- **Introduction**
- **Approximating the fractional energy loss distribution by Gaussian mixtures**
- **Reducing the number of components**
- **Results from simulated tracks**
- **Conclusions and outlook**

Introduction

For electrons the dominant contribution to energy loss is bremsstrahlung.

A well-known model of the bremsstrahlung energy loss is due to Bethe and Heitler:

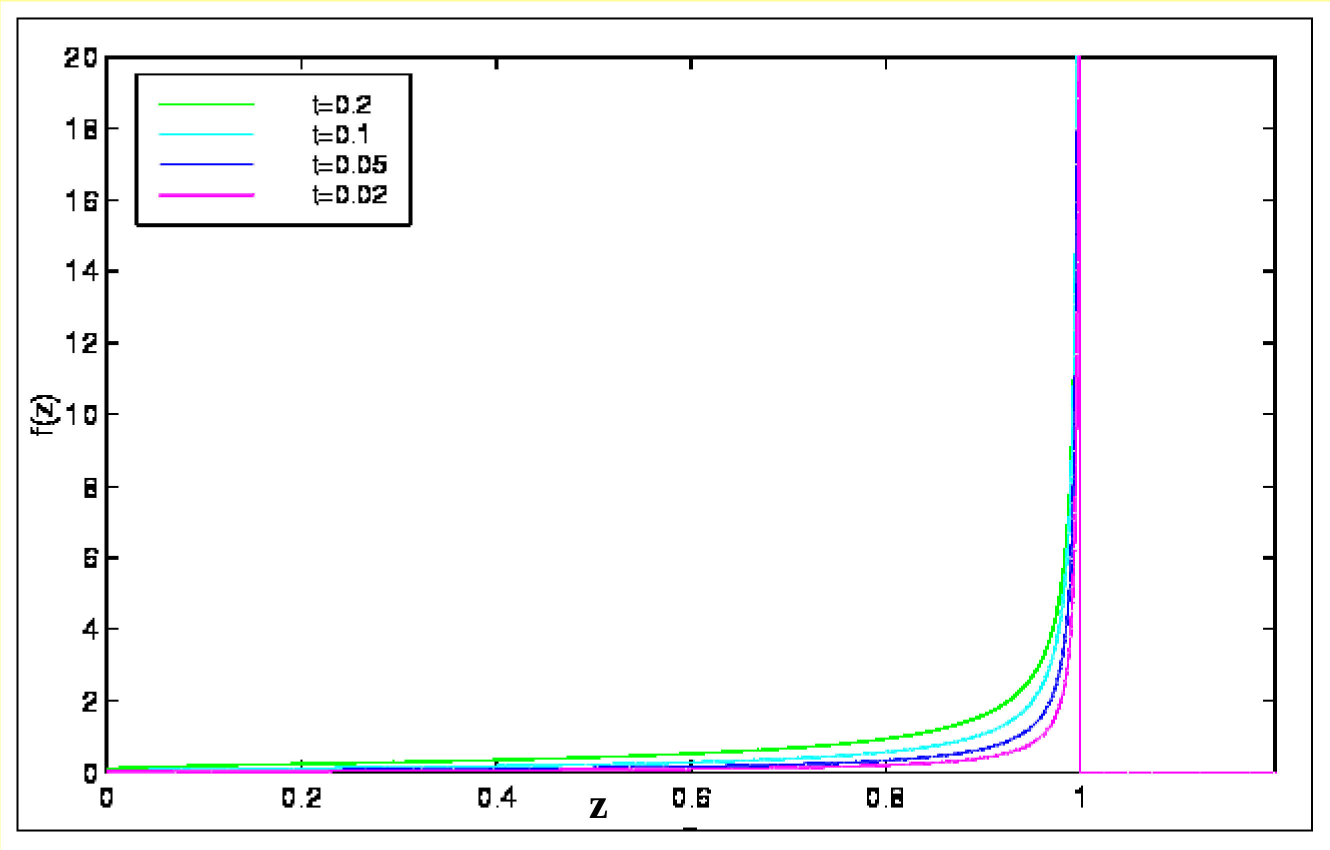
$$f(z) = \frac{[-\ln z]^{c-1}}{\Gamma(c)}, \quad \text{with} \quad c = t/\ln 2,$$

Fractional energy remaining

Path length in units of radiation length

Introduction

f(z) for different values of t



Are Strandlie, CERN

CHEP03, San Diego, 25.03.03

Introduction

The optimal treatment of radiative energy loss within the Kalman filter formalism is to correct the momentum part of the state vector with the mean value of energy loss and to add the variance of the energy loss distribution to the relevant part of the covariance matrix,

$$\mu = 2^{-c} = e^{-t}, \quad \sigma^2 = 3^{-c} - 4^{-c}.$$

NB!! Mean value and variance of f(z)!!

which should ensure unbiased estimates of track parameters and the associated uncertainties [1].

[1] D. Stampfer et al., Track fitting with energy loss. *Computer Physics Communications* 79 (1994) 157.

Introduction

- The Kalman filter is proved to be optimal only when all probability densities involved are Gaussian.
- Approximating the Bethe-Heitler distribution with a single Gaussian is a quite crude approximation.
- It is plausible that a non-linear estimator which takes the actual shape of the distribution better into account can do better.

Introduction

- **The basic idea is to model the Bethe-Heitler distribution as a Gaussian mixture g(z) instead of a single Gaussian, the different components modelling different degrees of hardness of the bremsstrahlung in the layer under consideration:**

$$g(z) = \sum_{i=1}^{N} g_i \, \varphi(z; \mu_i, \sigma_i)$$

PDF → $g(z)$

weight → g_i

single Gaussian → φ

Mean value of comp. i → μ_i

Standard deviation of comp. i → σ_i

Introduction

- **The natural description of the state vector then also becomes a Gaussian mixture.**
- **A non-linear generalization of the Kalman filter, the Gaussian-sum filter (GSF) [2, 3], is able to treat such state vectors appropriately.**
- **This algorithm has been implemented in the reconstruction software of the CMS tracker at CERN.**
- **The GSF resembles a set of Kalman filters running in parallel, each Kalman filter corresponding to one of the components of the state vector mixture.**

[2] R. Frühwirth, Track fitting with non-Gaussian noise. *Computer Physics Communications* 100 (1997) 1.

[3] R. Frühwirth and S. Frühwirth-Schnatter, On the treatment of energy loss in track fitting. *Computer Physics Communications* 110 (1998) 80.

Approximating mixture

Two kinds of distances between distribution and mixture have been minimized:

model

Gaussian mixture

Kullback-Leibler

$$D_{\mathrm{KL}} = \int_{-\infty}^{\infty} \ln[f(z)/g(z)]\, f(z)\, dz$$

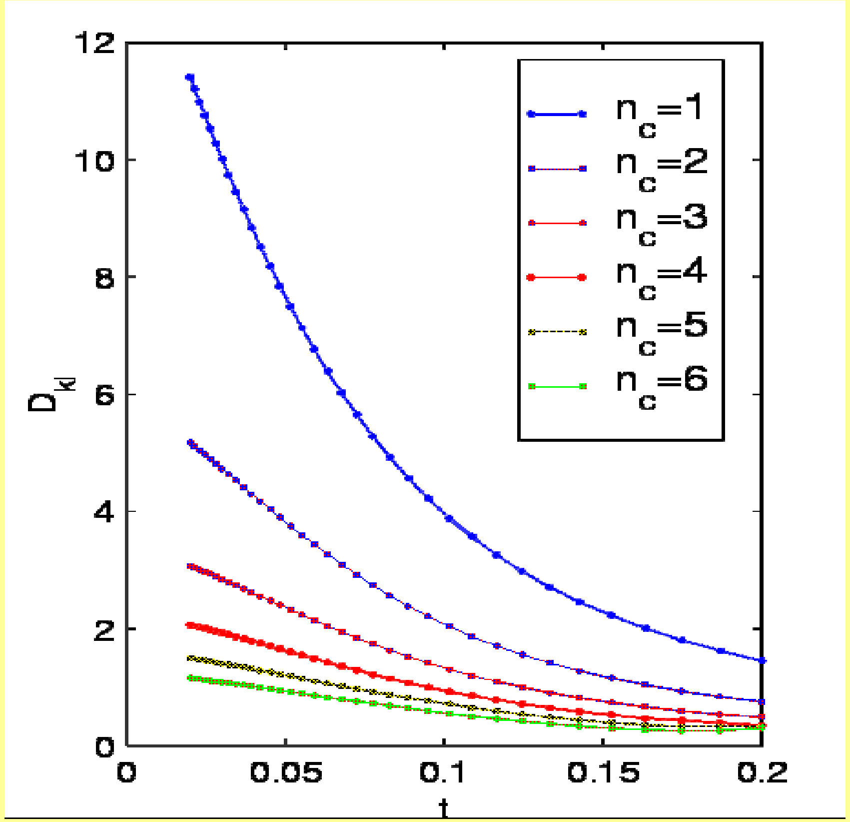
Cumulative distribution functions

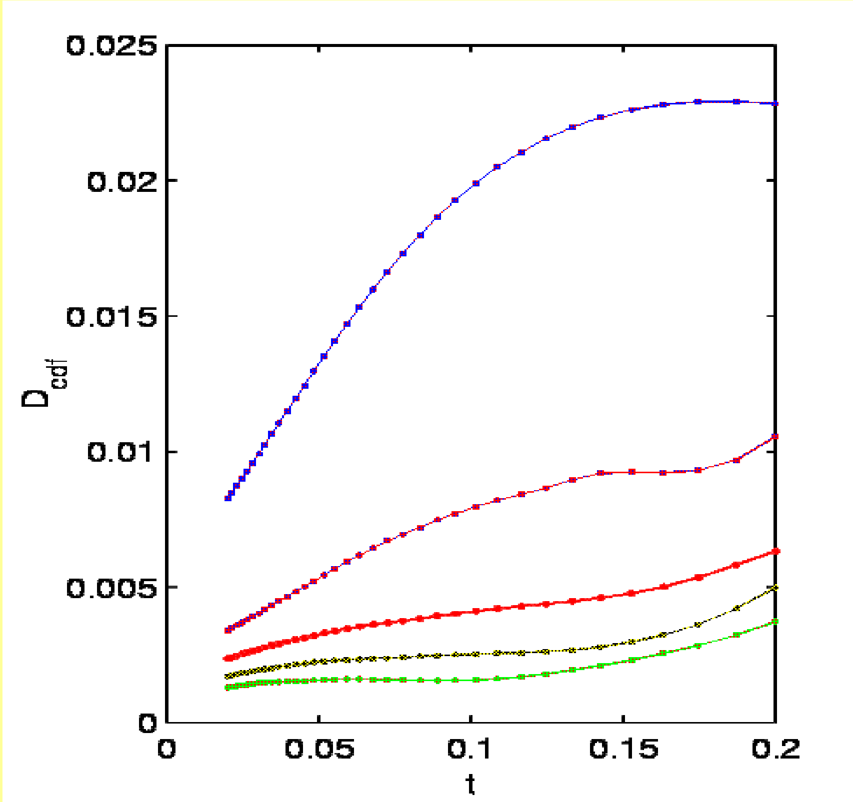
$$D_{\mathrm{CDF}} = \int_{-\infty}^{\infty} |F(z) - G(z)|\, dz$$

with respect to the mean, the variance and the weight of each of the components of the mixture.

Approximating mixture

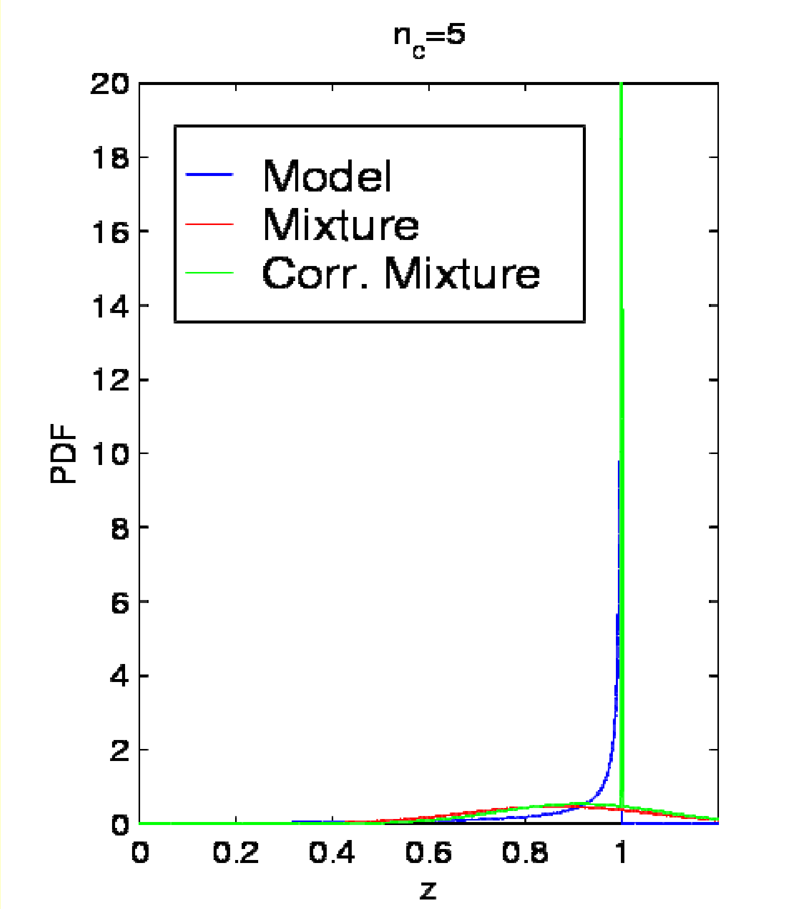
Distances as a function of path length

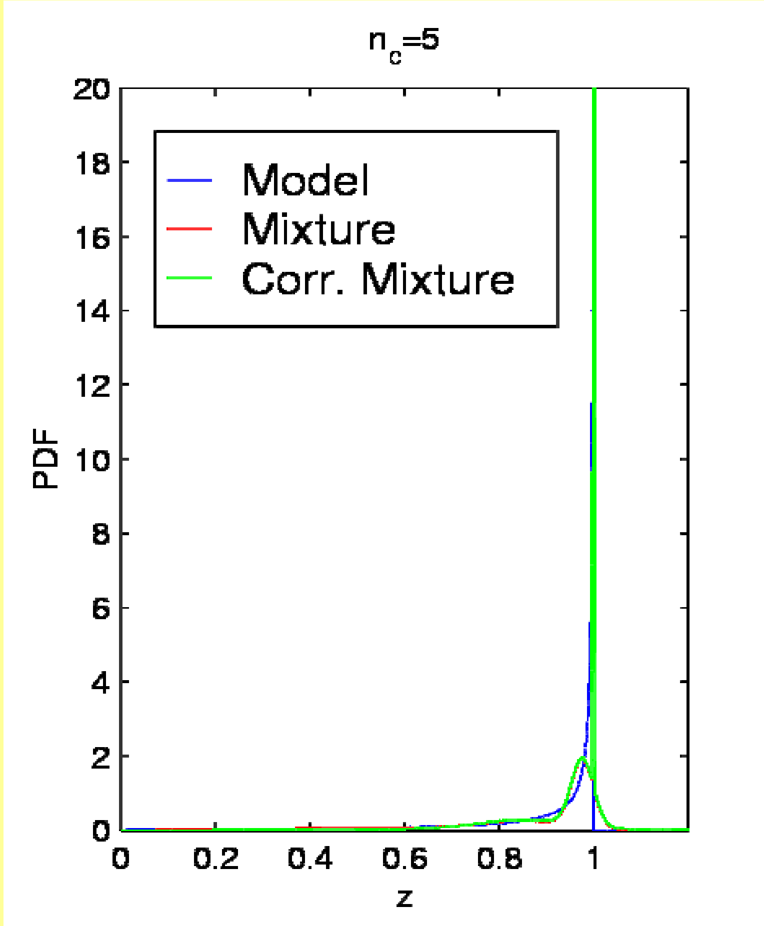




Approximating mixture

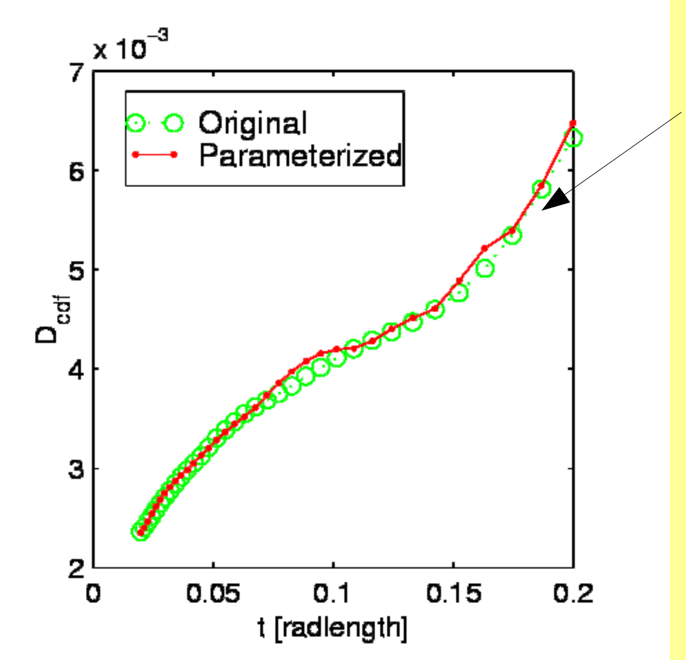
PDF for KL- and CDF- mixtures





Approximating mixture

- **Polynomials have been fitted to the weights, means and variances of the components as a function of the radiation length.**



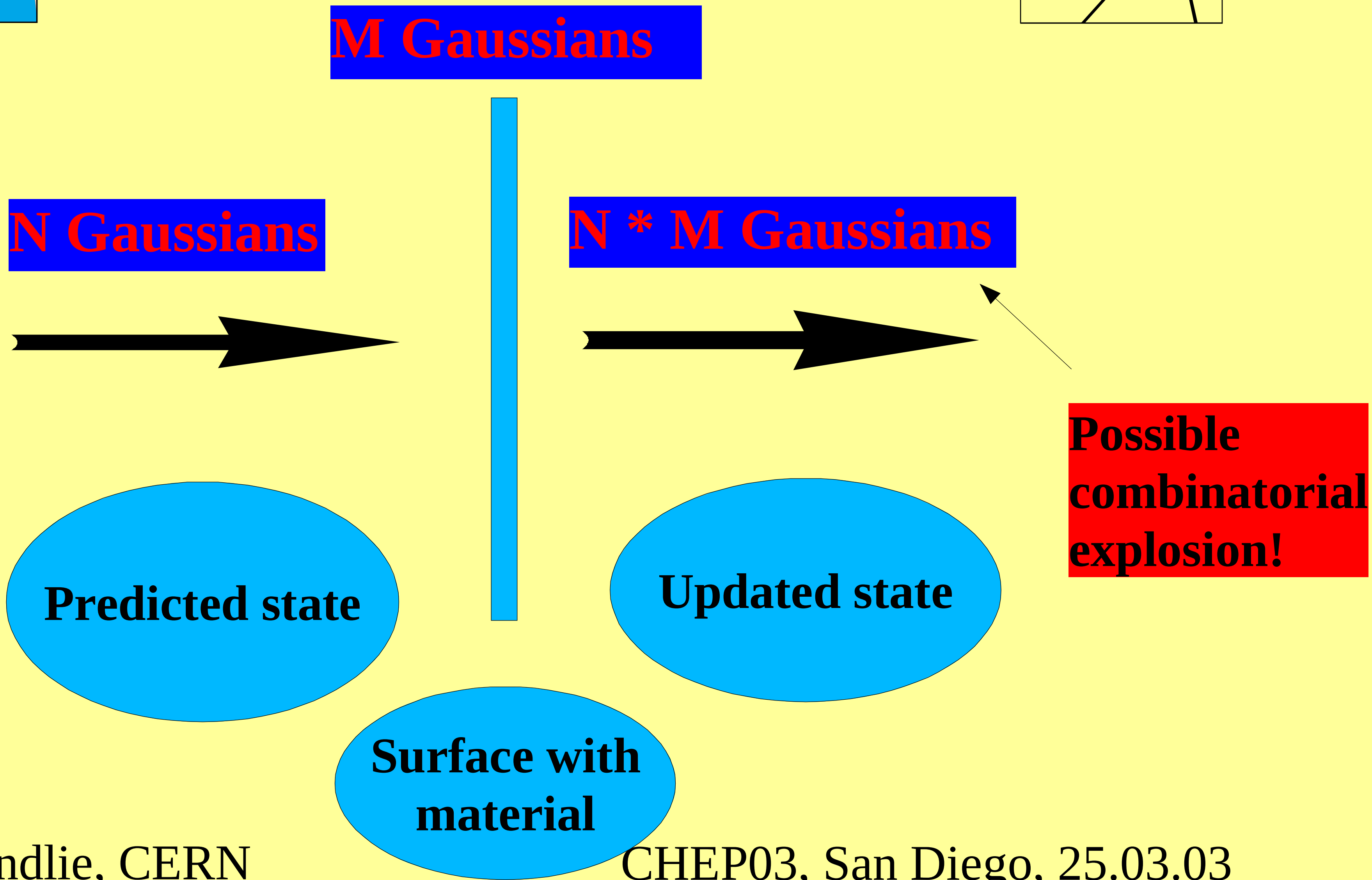
Comparison of distance from the model, using original and fitted mixture parameters.

Weights, means and variances calculated on the fly in CMS implementation of GSF. First and second moments are forced to those of the Bethe-Heitler distribution.

Component reduction

- Approximation of energy loss by a Gaussian mixture amounts to a convolution with the current state, which in general also is a Gaussian mixture.
- A strict application of the GSF algorithm quickly leads to a prohibitively large number of components due to the combinatorics involved each time a layer is traversed.
- The number of components must therefore repeatedly be reduced to a predefined maximum.
- The loss of information due to this procedure should be kept as small as possible.

Component reduction

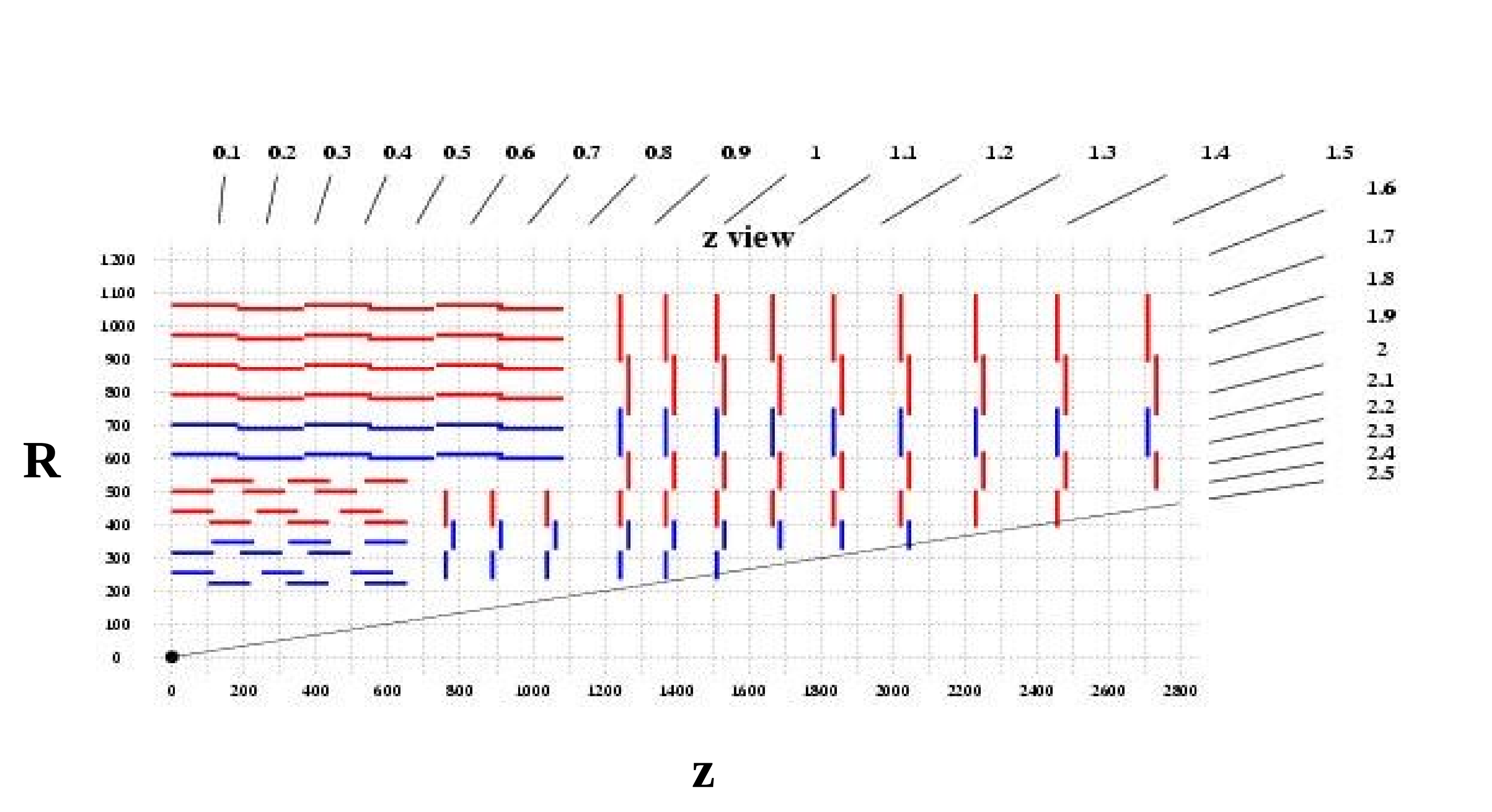


Component reduction

- **Two strategies have been evaluated:**
  1) **Keeping the components with the N largest weights.**
  2) **Clustering components which are close according to some metric in the five-dimensional parameter space of the tracks.**
- **The first approach is clearly inferior, since none of the first two moments of the estimated parameters turn out to be correctly described.**
- **In the second approach the adopted metric is the Kullback-Leibler distance defined earlier.**

Simulation studies

One quadrant of CMS tracker (not showing pixels)



Full-silicon tracker with pixel detectors, single-sided and double-sided silicon strip detectors.

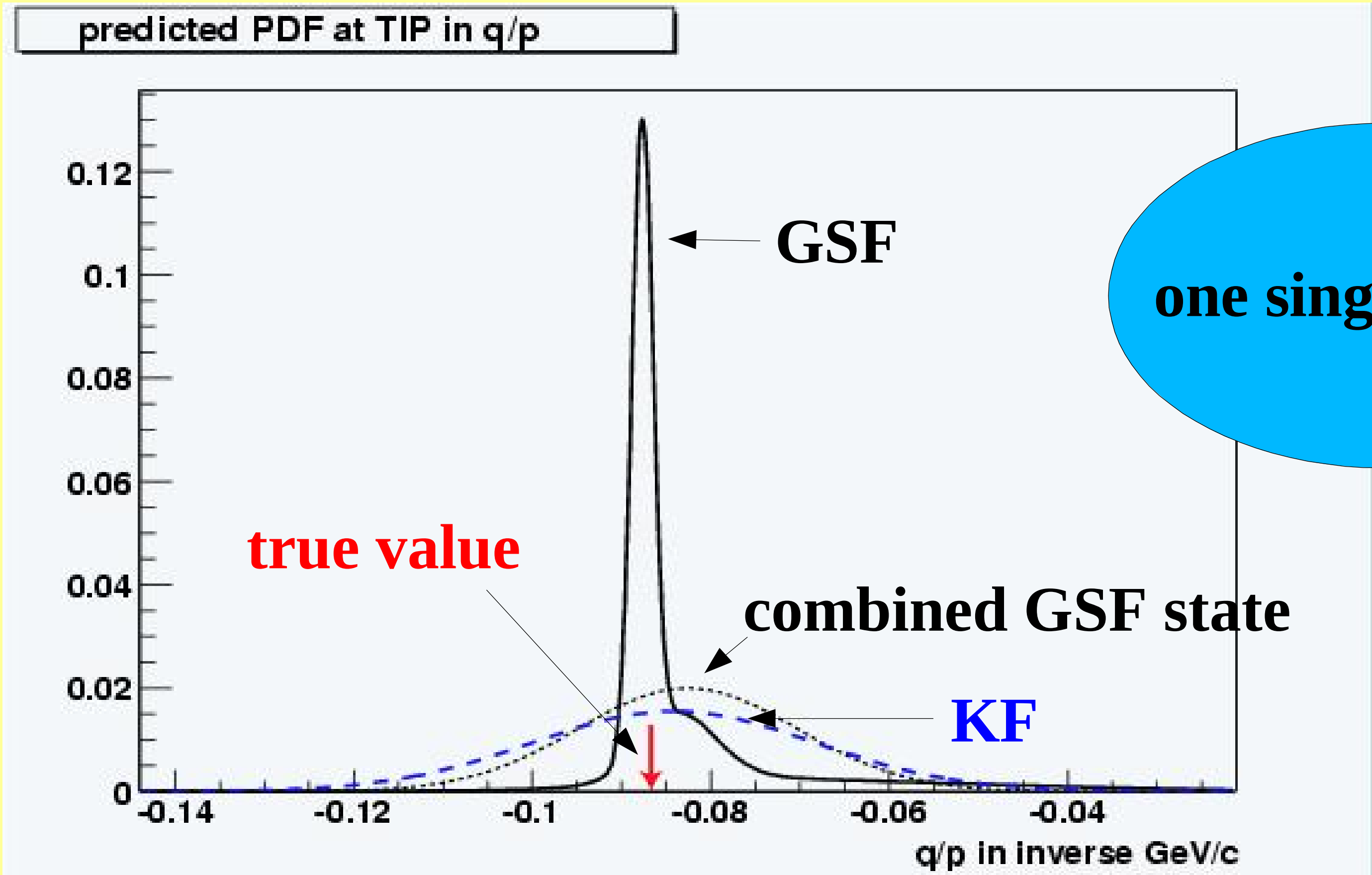
Simulation studies

- **Simplified simulation, bremsstrahlung simulated by sampling from Bethe-Heitler distribution. Multiple scattering and ionization energy loss have been turned off.**
- **10 GeV/c transverse momentum electrons in the barrel (|pseudo-rapidity| < 1).**
- **Full knowledge of amount and position of material.**
- **Considering estimates of charge over absolute value of momentum (q/p) at the transverse impact point (TIP), i.e. point of closest approach in transverse plane to vertex.**
- **Collecting reconstructed hits from knowledge of true hits constituting the track - no pattern recognition involved.**

Simulation studies

- **The implementation of the GSF in the CMS reconstruction software was quite straightforward, since most of the main building blocks of the standard Kalman filter could be re-used.**
- **The main extension as compared to the Kalman filter is the necessity of describing the state vector as a Gaussian mixture rather than a single Gaussian.**
- **Since C++ allows a state-vector class to be polymorphic, these complications can be hidden inside such a class - minimizing the need of modifying other basic components.**

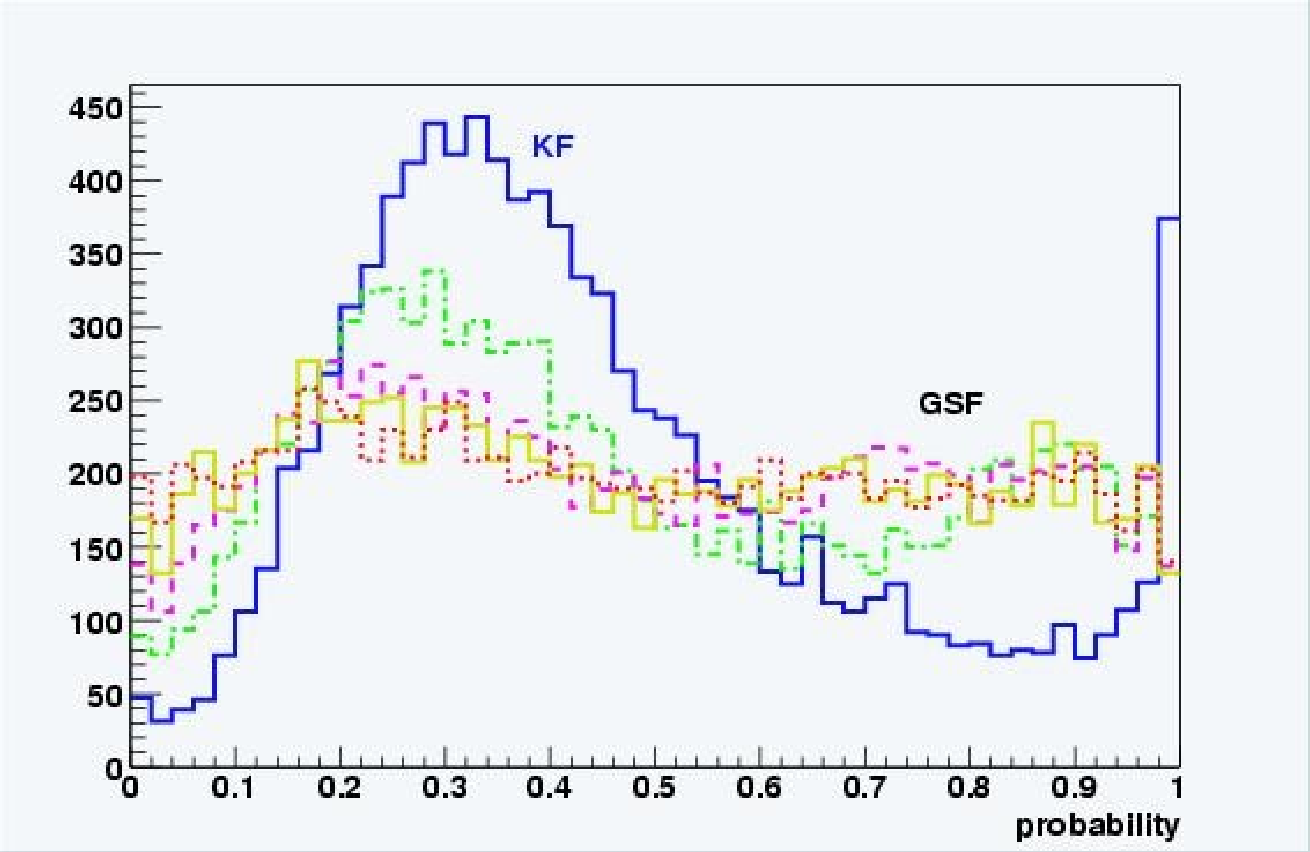
Simulation studies



one single track!

Simulation studies

Estimated q/p, simulating from mixture



Varying max. # of components:

6 : dashed- dotted

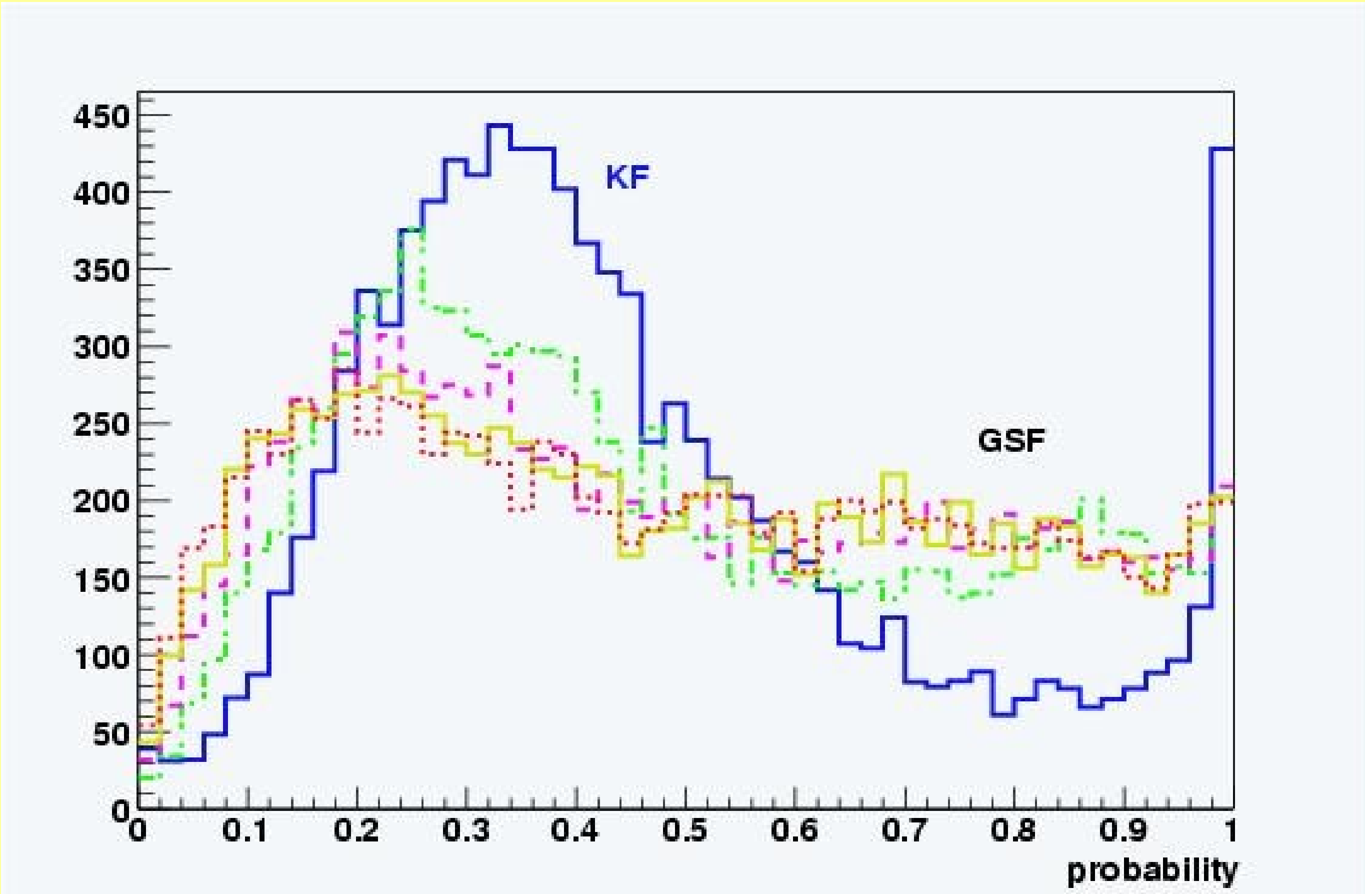
12 : dashed

18 : solid

36 : dotted

Simulation studies

Estimated q/p, simulating from Bethe-Heitler



Varying max. # of components:

6 : dashed- dotted

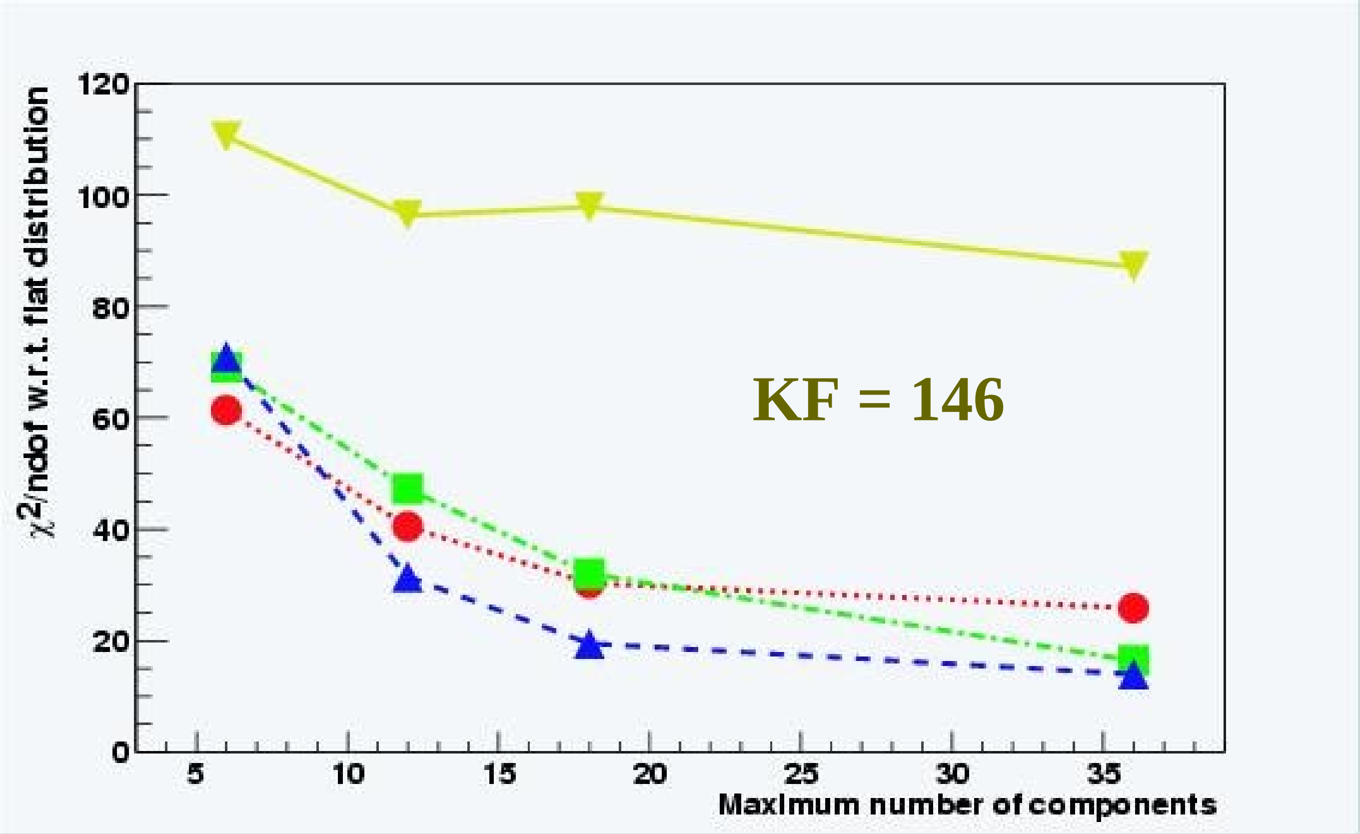
12 : dashed

18 : solid

36 : dotted

Simulation studies

Chisquare/ndof with respect to flat distribution



KL-mixture with
6 components.

<u>CDF-mixtures with:</u>

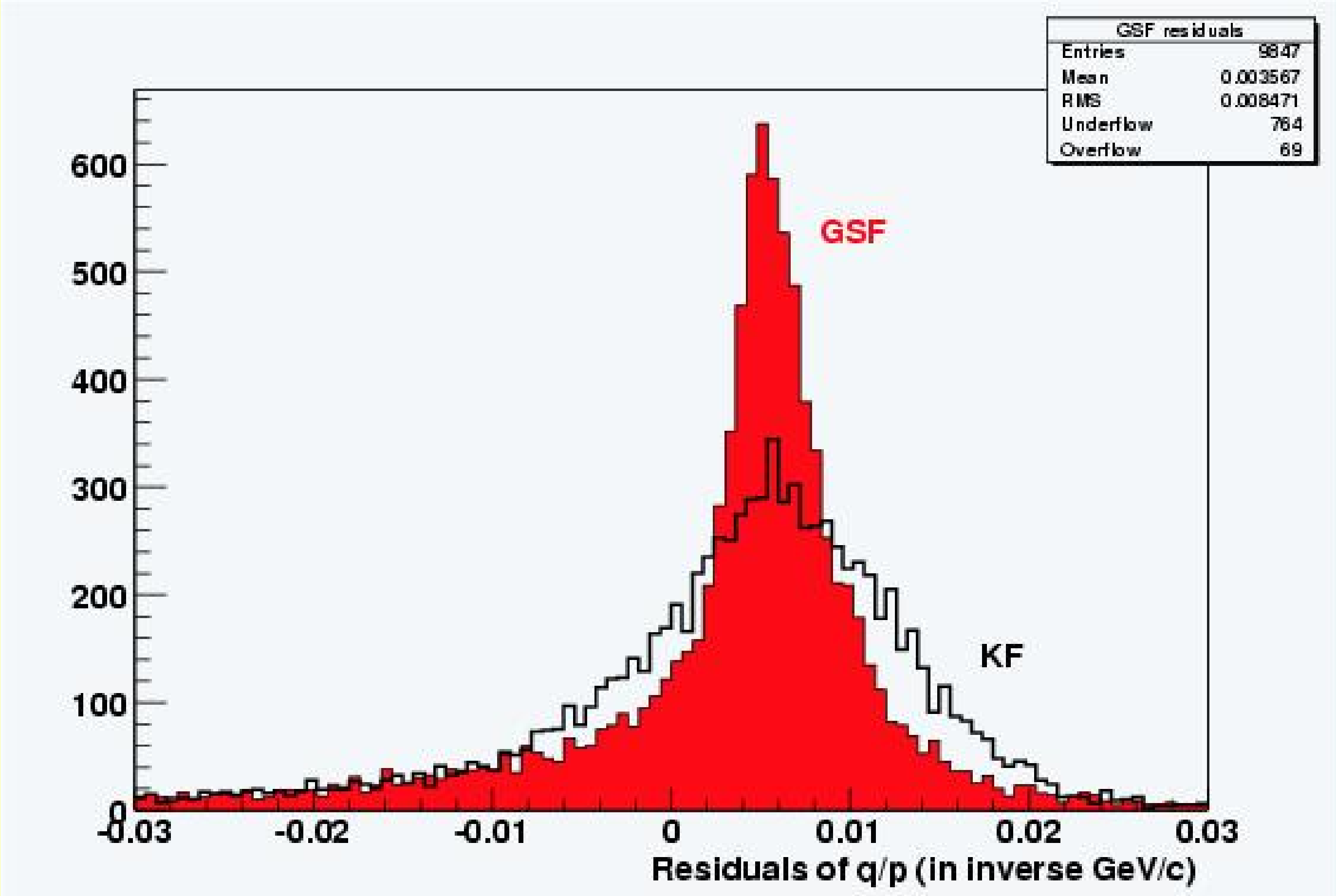
4 components

5 components

6 components

Simulation studies

Residuals (estimated q/p with respect to true value) at TIP



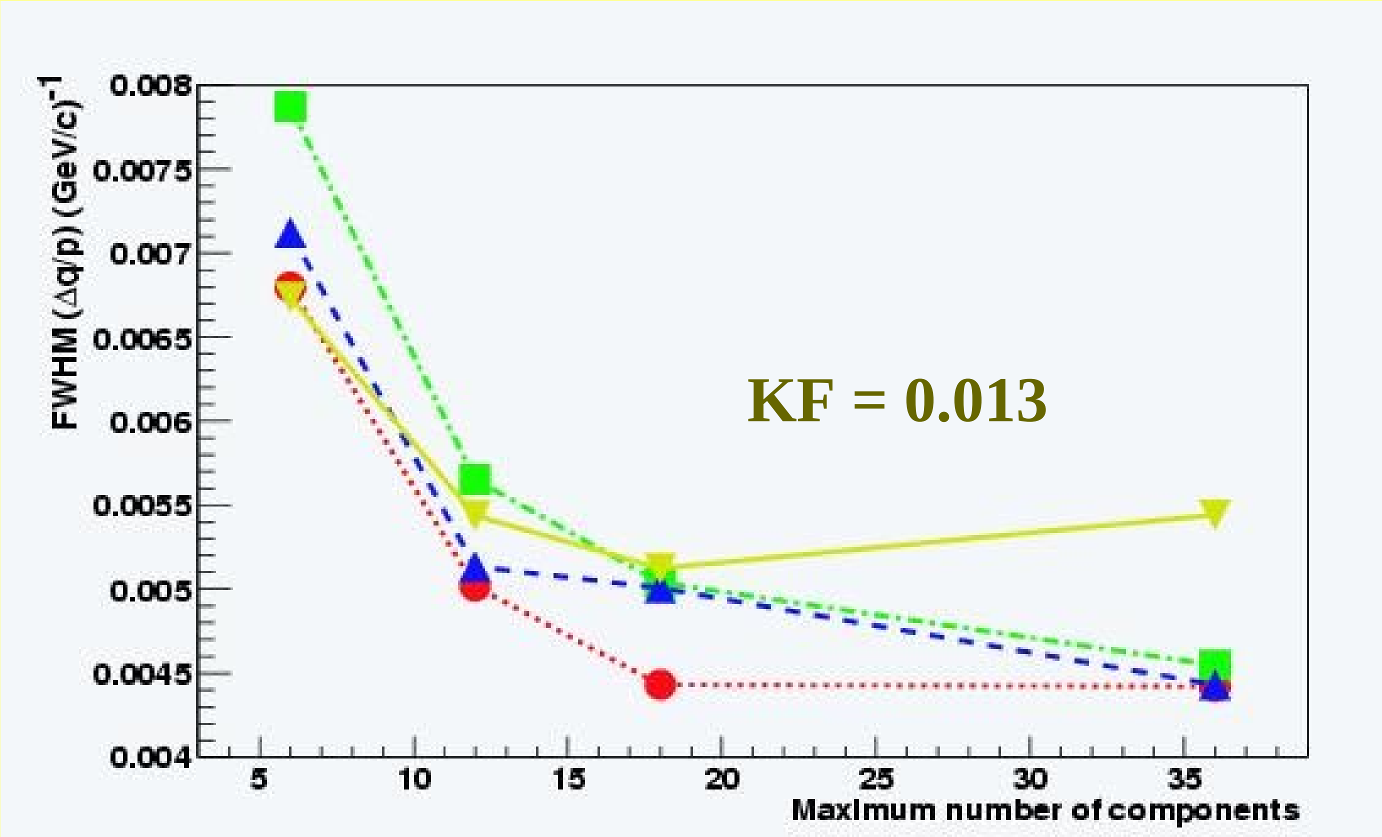
CDF-mixture with 6 components.

Maximum # of components: 12.

Long tails due to hard radiation in innermost layers exist for both methods.

Simulation studies

Full-width at half-maximum of residual distributions



KL-mixture with 6 components.

CDF-mixtures with:

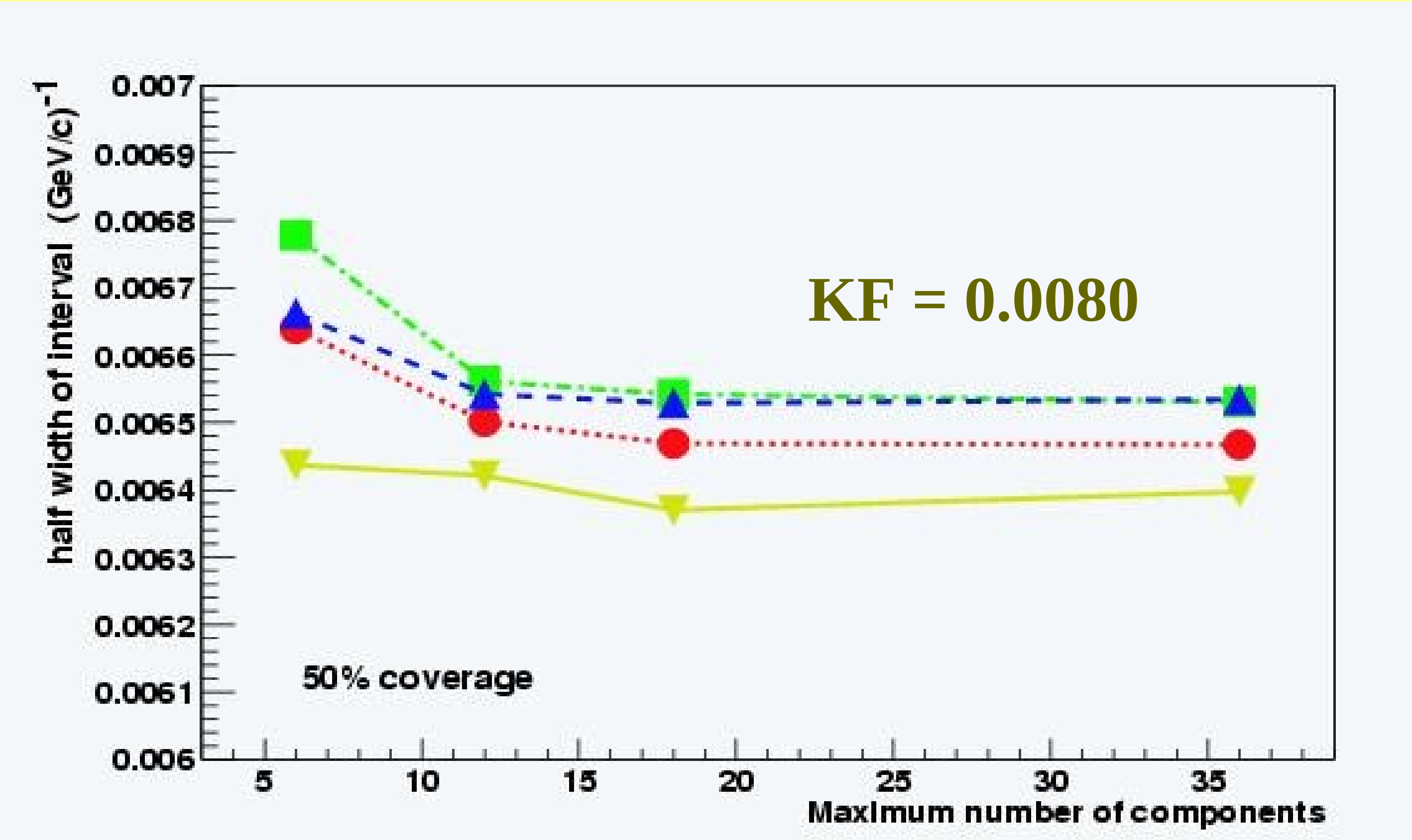
4 components

5 components

6 components

Simulation studies

Half-width of interval covering 50 %



KL-mixture with 6 components.

CDF-mixtures with:

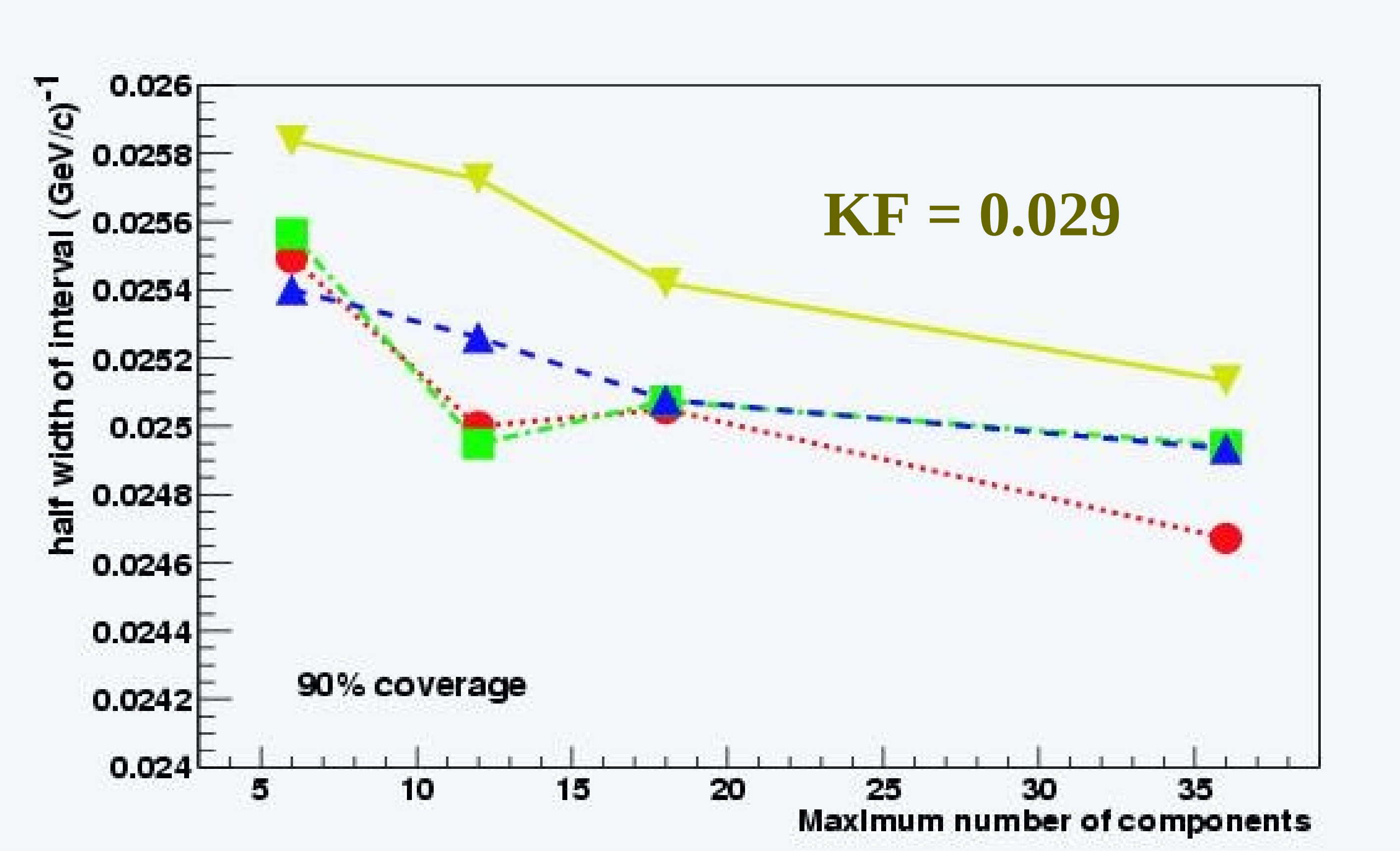
4 components

5 components

6 components

Simulation studies

Half-width of interval covering 90 %



KL-mixture with 6 components.

CDF-mixtures with:

4 components

5 components

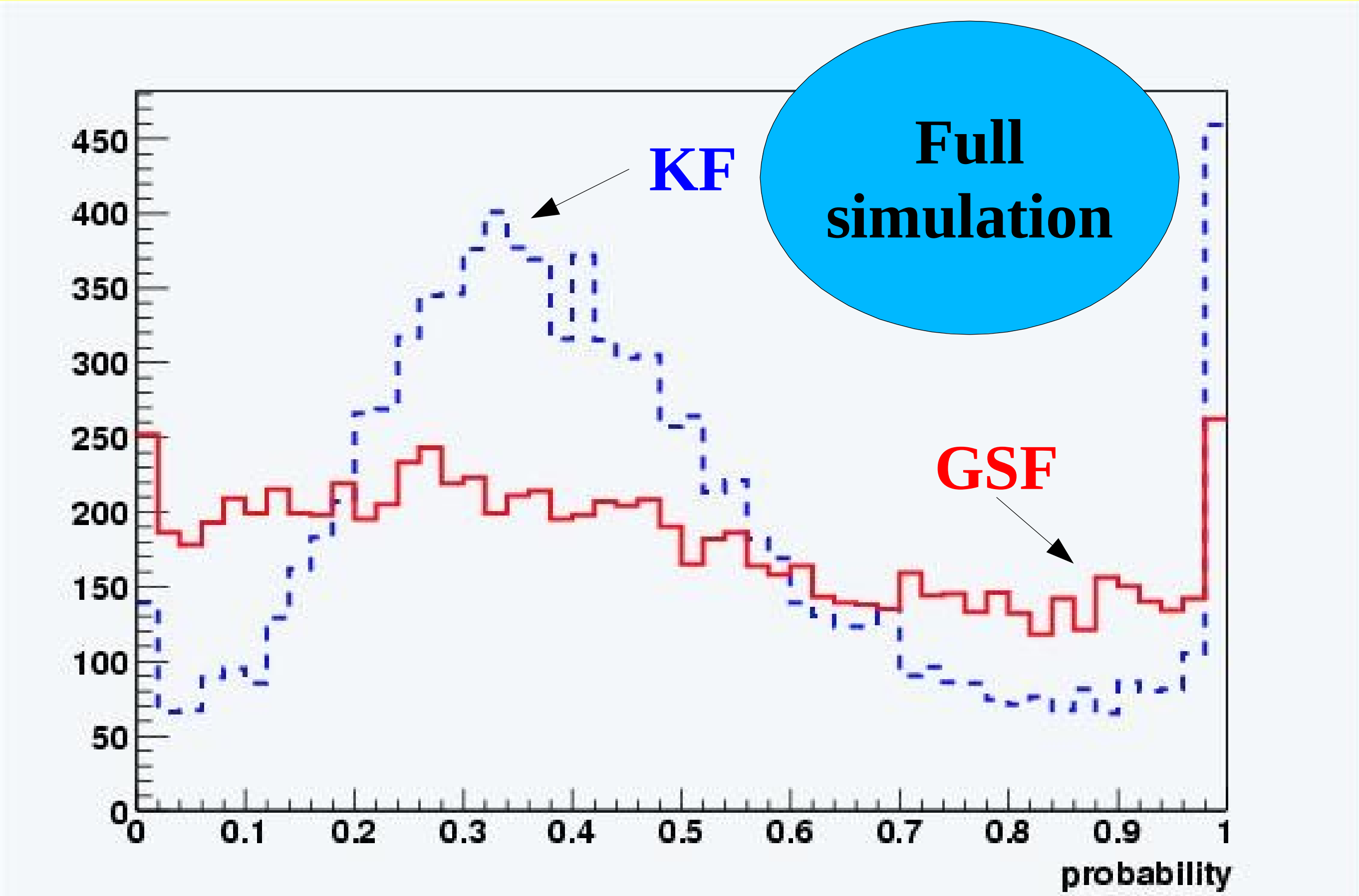
6 components

Simulation studies

- The GSF and the KF have also been run on tracks from a full simulation using the official CMS simulation program.
- Transverse momentum and eta range same as in simplifed simulation, but amount and spatial distribution of material different.
- Results from simplified and full simulation are thus not immediately quantitatively comparable.
- Qualitatively, however, one expects a similar relative behaviour between the GSF and the KF.

Simulation studies

Probability distribution, estimated q/p

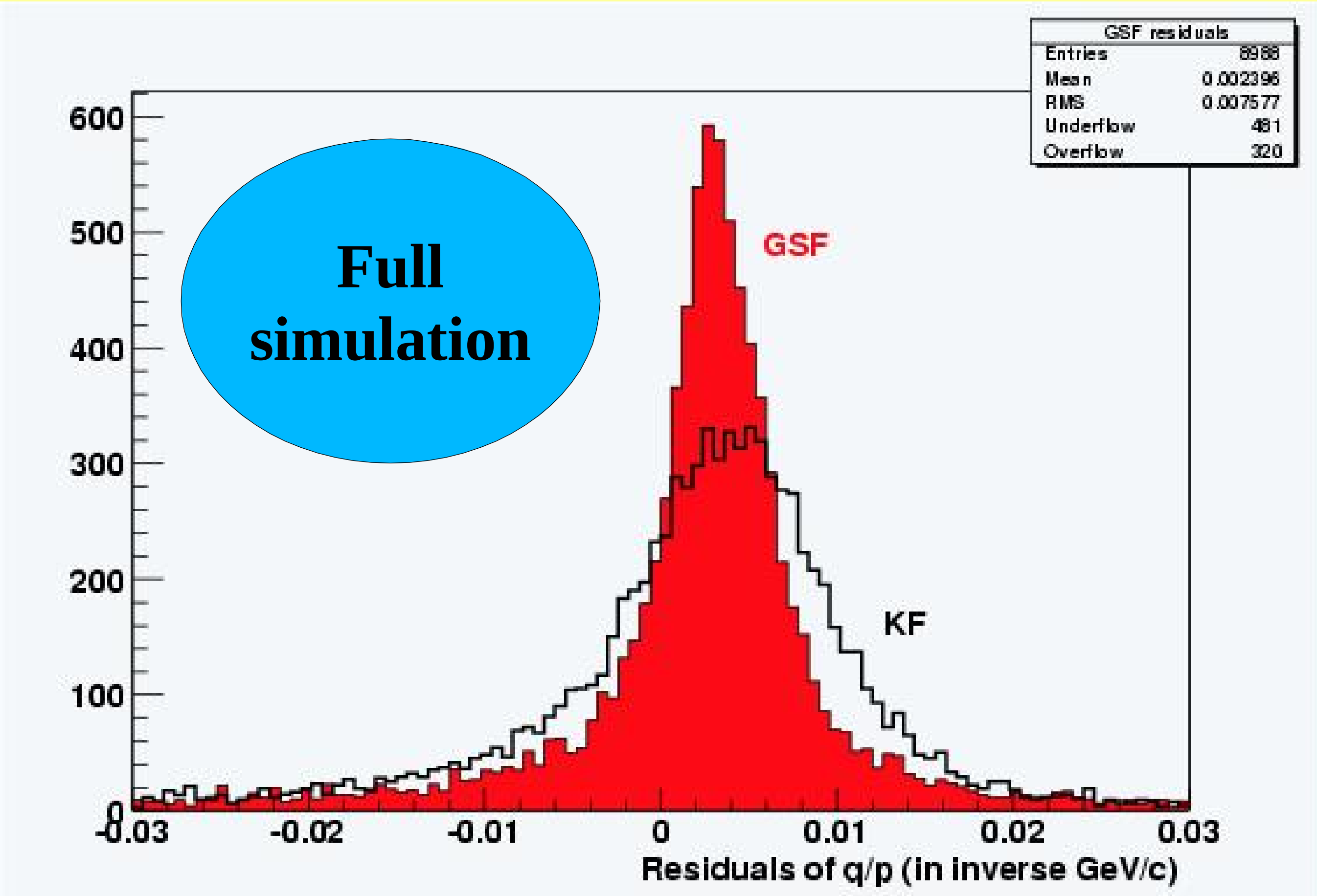


GSF:

CDF-mixture with 6 components, max. # of components: 12

Simulation studies

Residuals (estimated q/p with respect to true value) at TIP



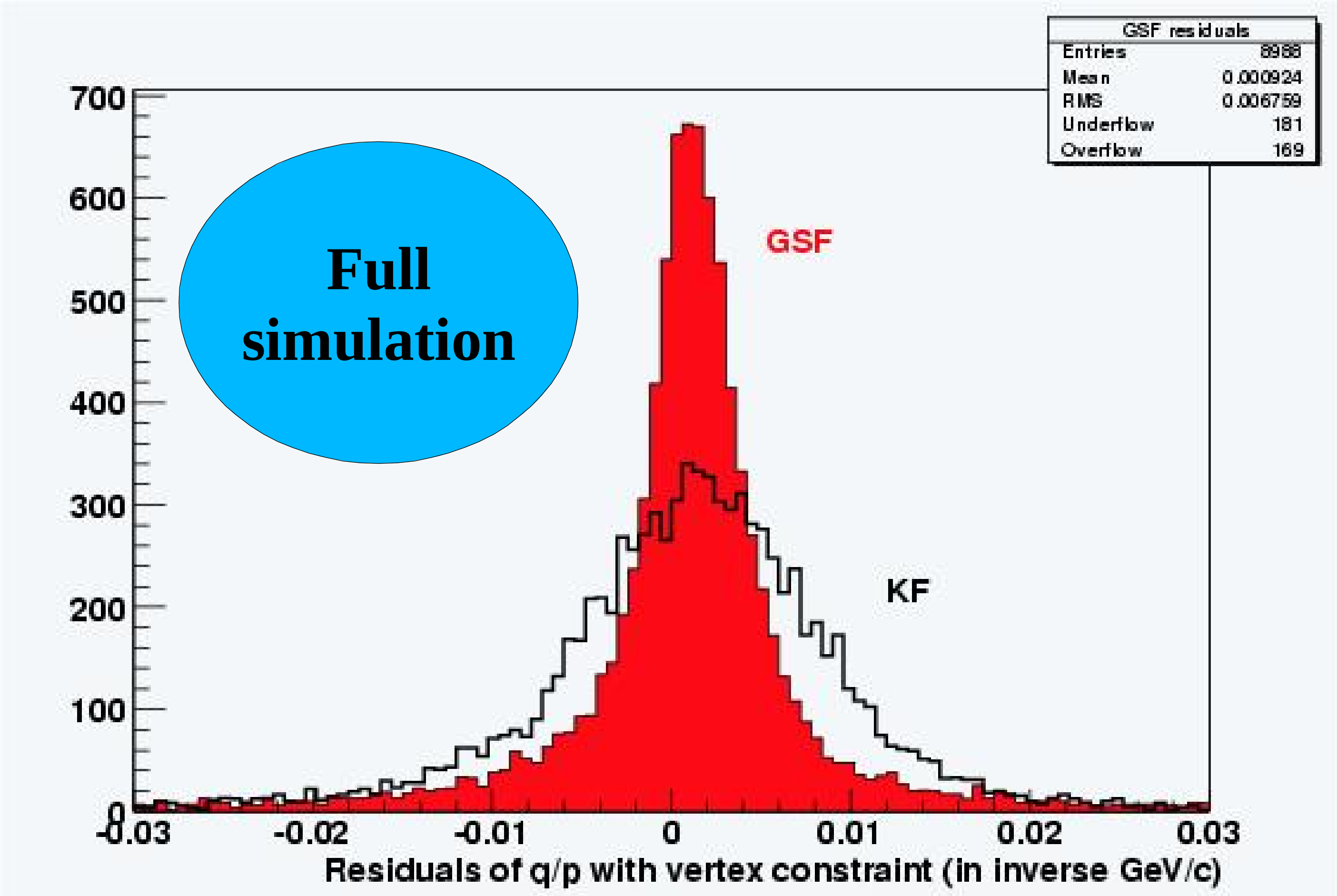
CDF-mixture with 6 components

Maximum # of components: 12

Qualitatively looking very similar to simplified simulation!

Simulation studies

Residuals at TIP, including vertex constraint



Vertex constraint allows to measure momentum in innermost layers.

Distribution less skew, mode closer to zero and reduced amount of tracks in the tails.

Conclusions / outlook

- GSF has been implemented in the reconstruction software of the CMS tracker.
- Momentum resolution for electrons significantly improved with respect to the Kalman filter.
- Quality of estimated track parameters significantly better than that of the Kalman filter.
- More systematic studies of tracks from full simulation needed.
- Nevertheless, safe to conclude that GSF yields substantial gain in precision as compared to the Kalman filter.

Conclusions / outlook

- **At the very high energies accessible at LHC, also muons will suffer from bremsstrahlung energy loss.**
- **With a critical energy of several hundred GeV, muon bremsstrahlung is not relevant for the CMS tracker, but indeed for the CMS muon detectors.**
- **In principle a GSF such as the one described herein can be used also for reconstruction of bremsstrahlung muons.**

Conclusions / outlook

- **The GSF can in principle be expected to give improvements with respect to the standard Kalman filter in all situations where prior knowledge of a non-Gaussian kind can be included in the reconstruction procedure.**
- **Possible examples:**
  - **non-Gaussian measurement errors,**
  - **non-Gaussian tails of multiple scattering.**
- **A GSF related to latter situation has already been implemented in CMS and is under evaluation.**