Tracking with the Hough transform

Felix Fehr
LPC Clermont-Ferrand

CALICE Collaboration Meeting
UT Arlington
March 10th, 2010

Calorimeter for ILC

Motivation

- Develop MIP ID and tracking algorithm
  - to check data quality, in particular to study impact of overlay events (e.g. π+e) on electron analysis
  - to address tracking capabilities of calorimeter (can provide benchmark for further studies)
  - for calibration purposes

- Algorithm should be
  - robust, fast, and not too complicated
  - applicable for Si-W ECAL and potentially also be useable for HCAL without too many modifications

The Hough transform

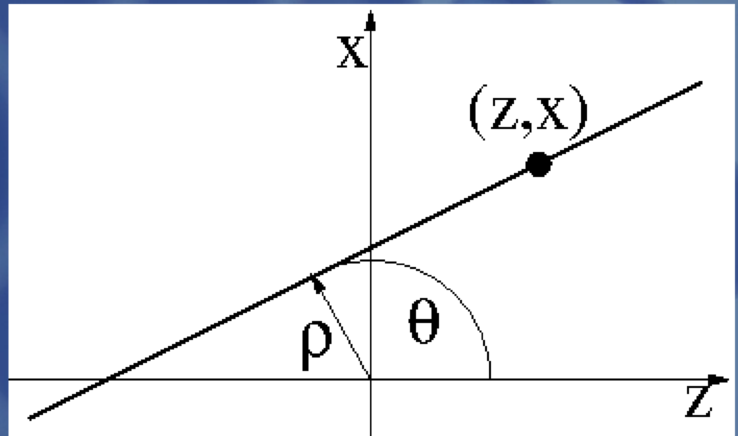
- Identifies general curves/surfaces
  - robust, fast, yet simple algorithm
  - Paul V.C. Hough, US Patent, 1962
- Look at (x,z) and (y,z) projections: straight lines in 2D
- Definition of Hough space

$$\rho = x \sin\theta + z \cos\theta$$

ρ : distance of line to origin

θ : angle between foot of normal and z-axis

(analogue definitions for (y,z))



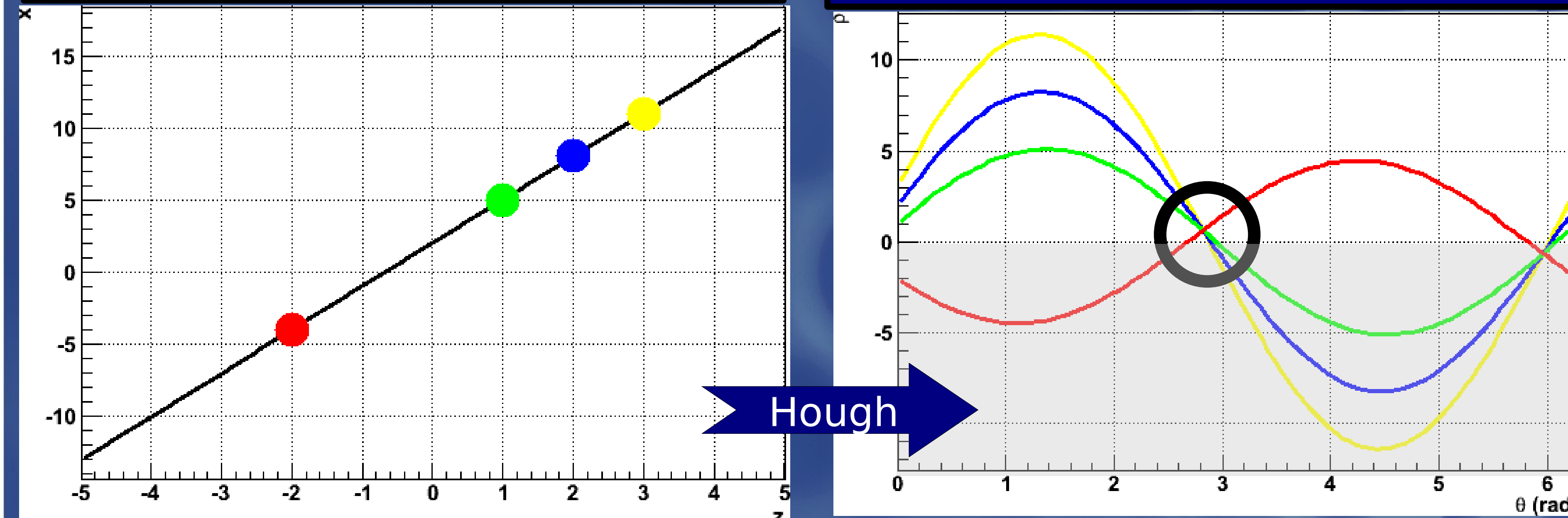
$$\rho \in [0, \infty]$$

$$\theta \in [0, 2\pi]$$

Principle of the HT

Feature space

Hough parameter space



- Hough transform of a point given by sinoidal function
- In Hough space, transforms of points on a line intersect at line parameters

Hit filtering and weighting

- Select MIP compatible hits: select cells with no more than 1.5 of average energy induced by MIP

- Identifiy isolated hits (no more than 2.0 MIPs within r=2cm in same layer)

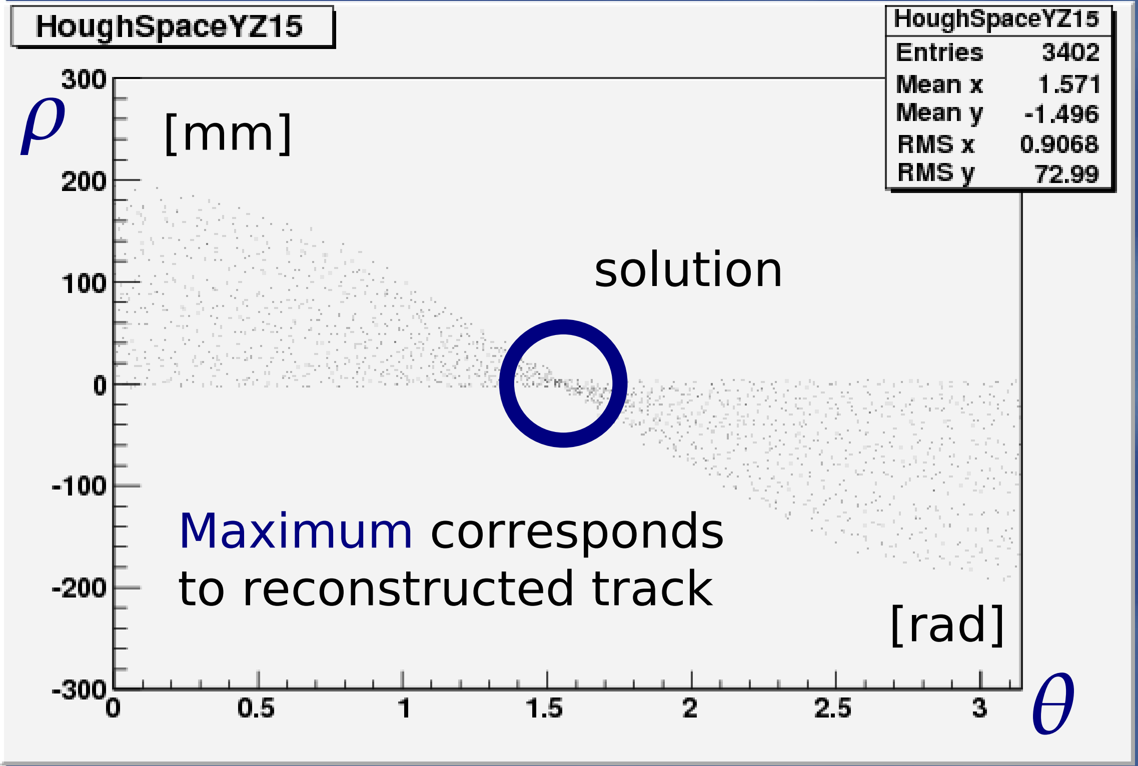
- Amplify MIP signals with layer dependent weights:

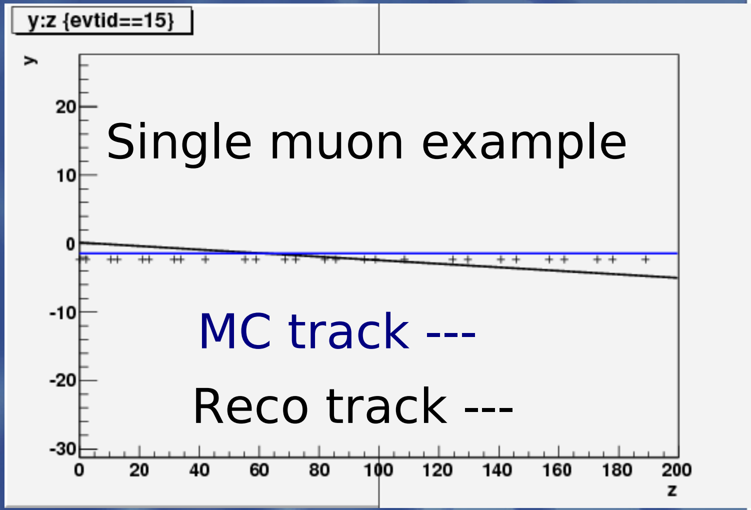
Hit type	Layer 1-5	Layer 6-25	Layer 26-30
isolated	4	3	4
isolated and connected	3	2	3
isolated and connected to isolated	2	1	2

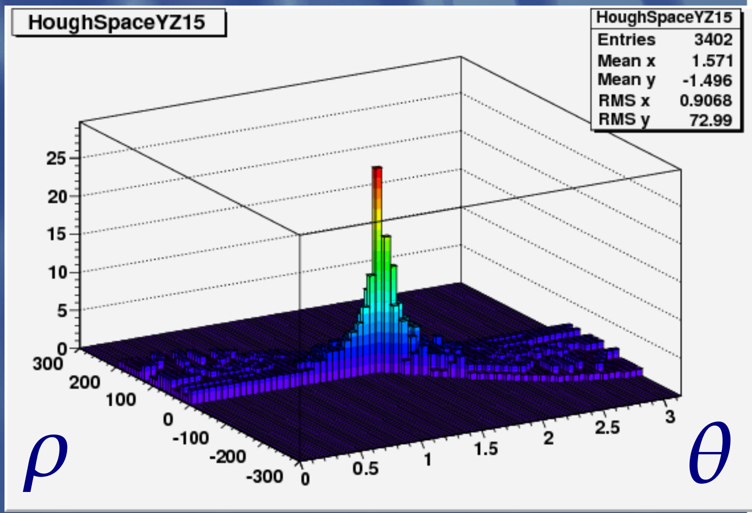
Transformation & peak detection

For each hit:

vary $\theta \in [0, 2\pi]$ and calculate ρ



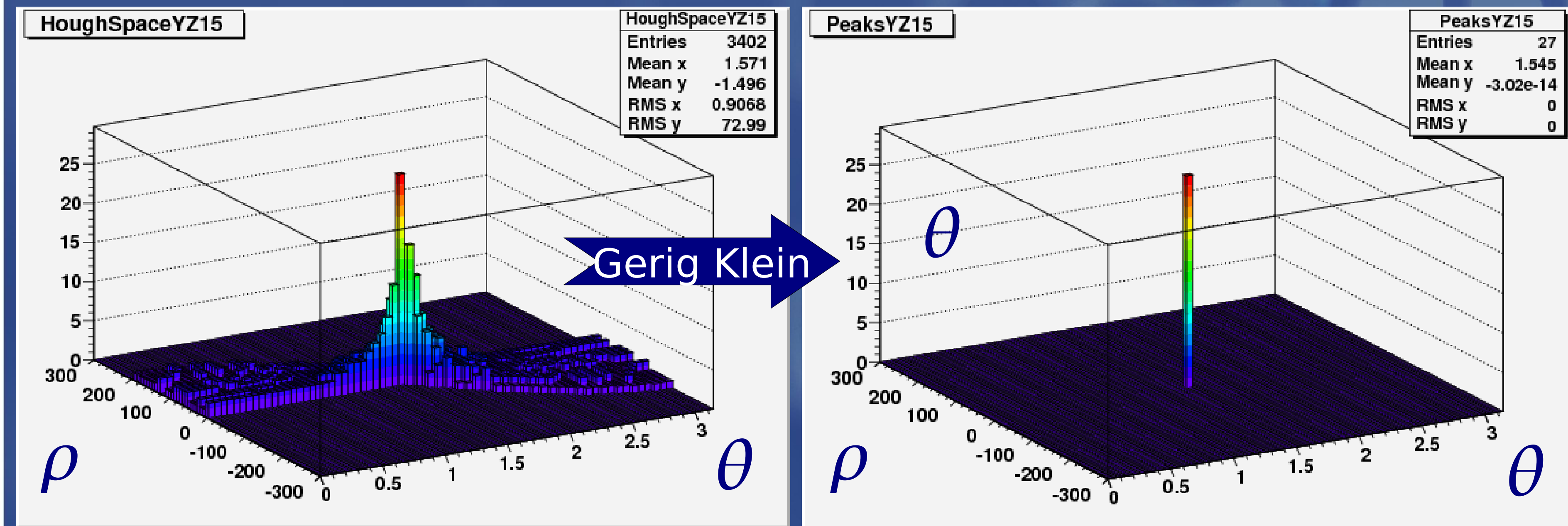




$$\rho = y \sin\theta + z \cos\theta$$

Gerig and Klein post-processing

useful for 'busy' events



- allows each hit only one vote:
  each hit (y,z) votes for the maximum in (ρ, θ) with

$$\rho = y \sin\theta + z \cos\theta$$

Generalisation of post-processing

Basic idea: allow each hit (x,y,z) to vote for only one (!) $(\rho_x, \theta_x, \rho_y, \theta_y)$ combination

- combines 2D x 2D information, yields track parameters
- exploits fact that signals in (ρ_x, θ_x) also peak in (ρ_y, θ_y): each hit is associated with only one 3D track
- improves suppression of ghost solutions
- seems to be a new approach

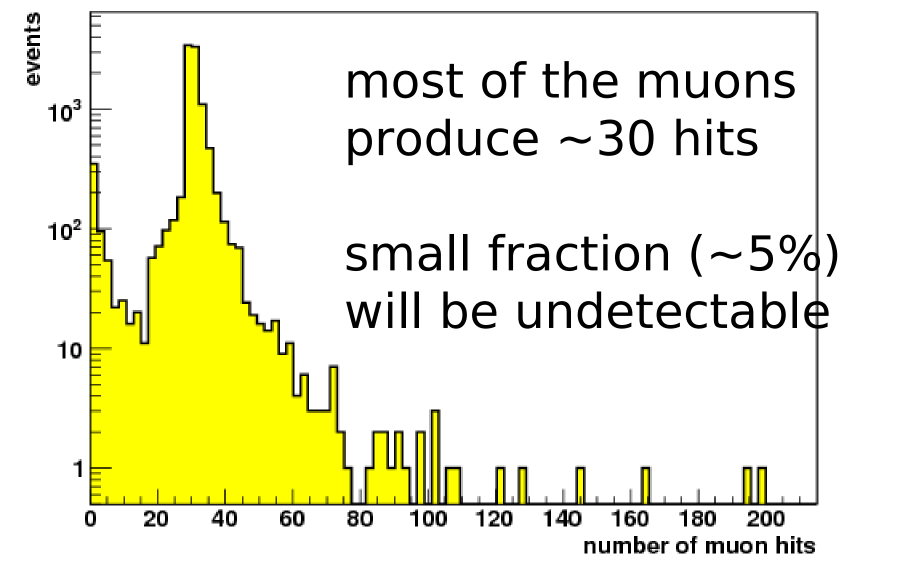
Maxima in $(\rho_x, \theta_x, \rho_y, \theta_y)$ detected with cluster algorithm

Final chi2 fit of 3D track to selected hits (not essential, used as a cross-check)

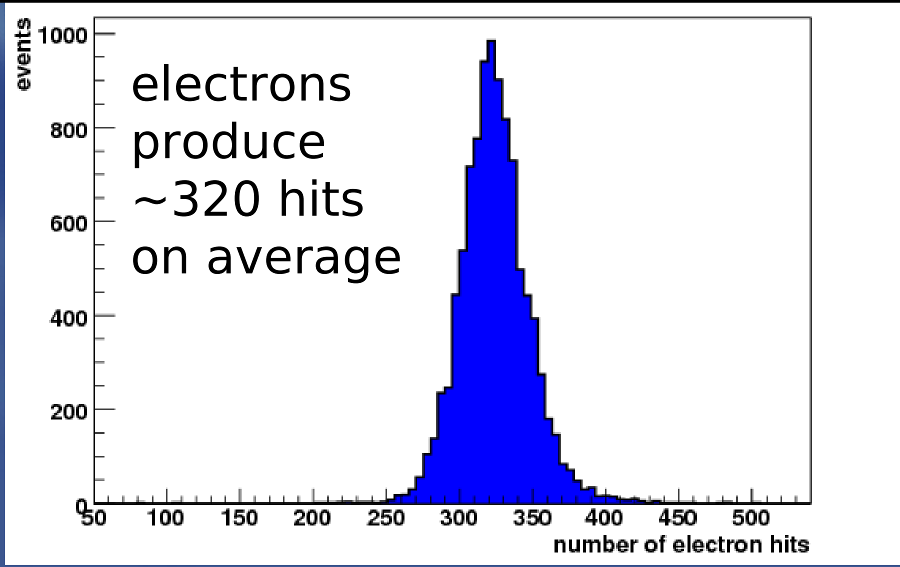
The overlay sample

- 30 GeV electrons in SiW ECAL overlayed with MIPs
  - 10.000 events (1 e + 1 MIP per event) generated
  - simulation of electron / MIP separately with Mokka
  - event merging with LCIO tools
  - full hit info available (parent ID stored)

Number of hits due to muon



Number of hits due to electron

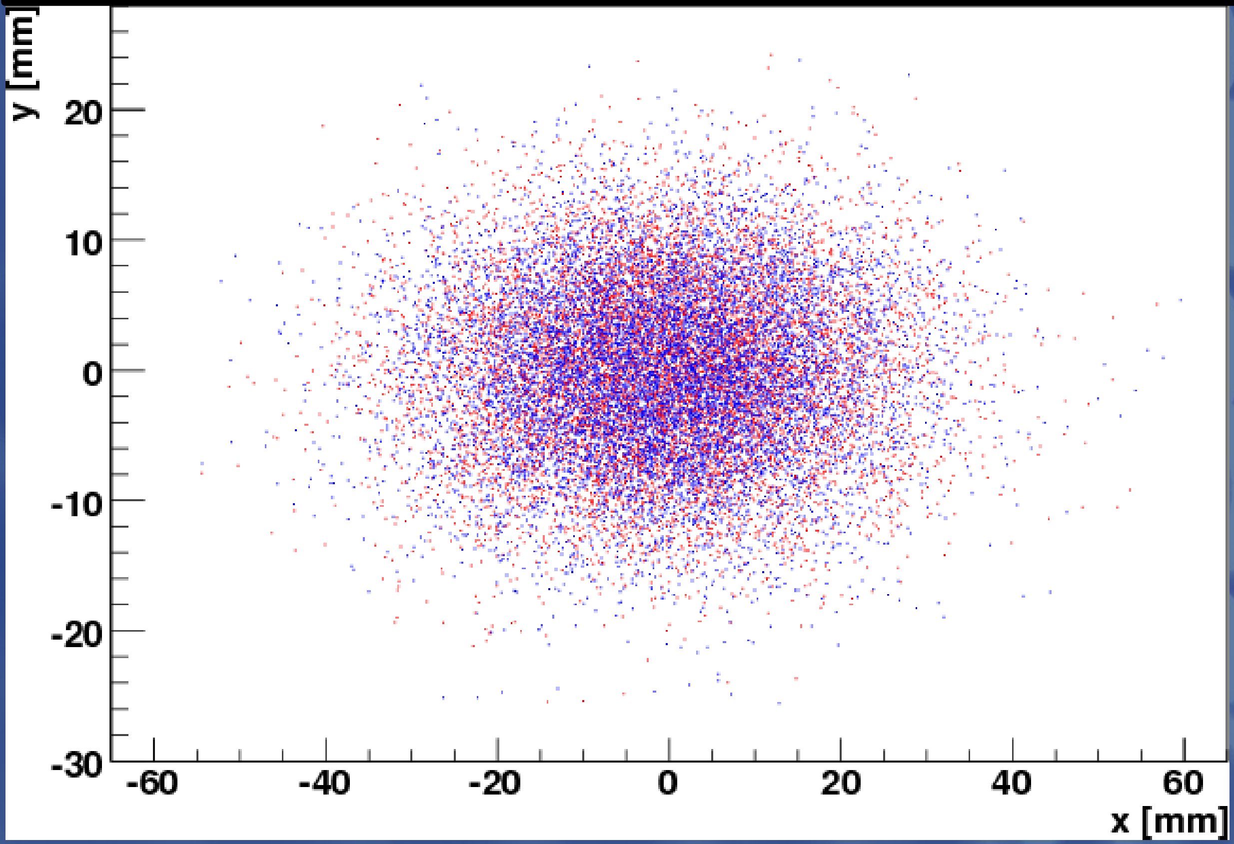


"Beam profile"

generated beam profile should reflect experimental conditions

reconstruction can be quite challenging if MIP and shower are very close

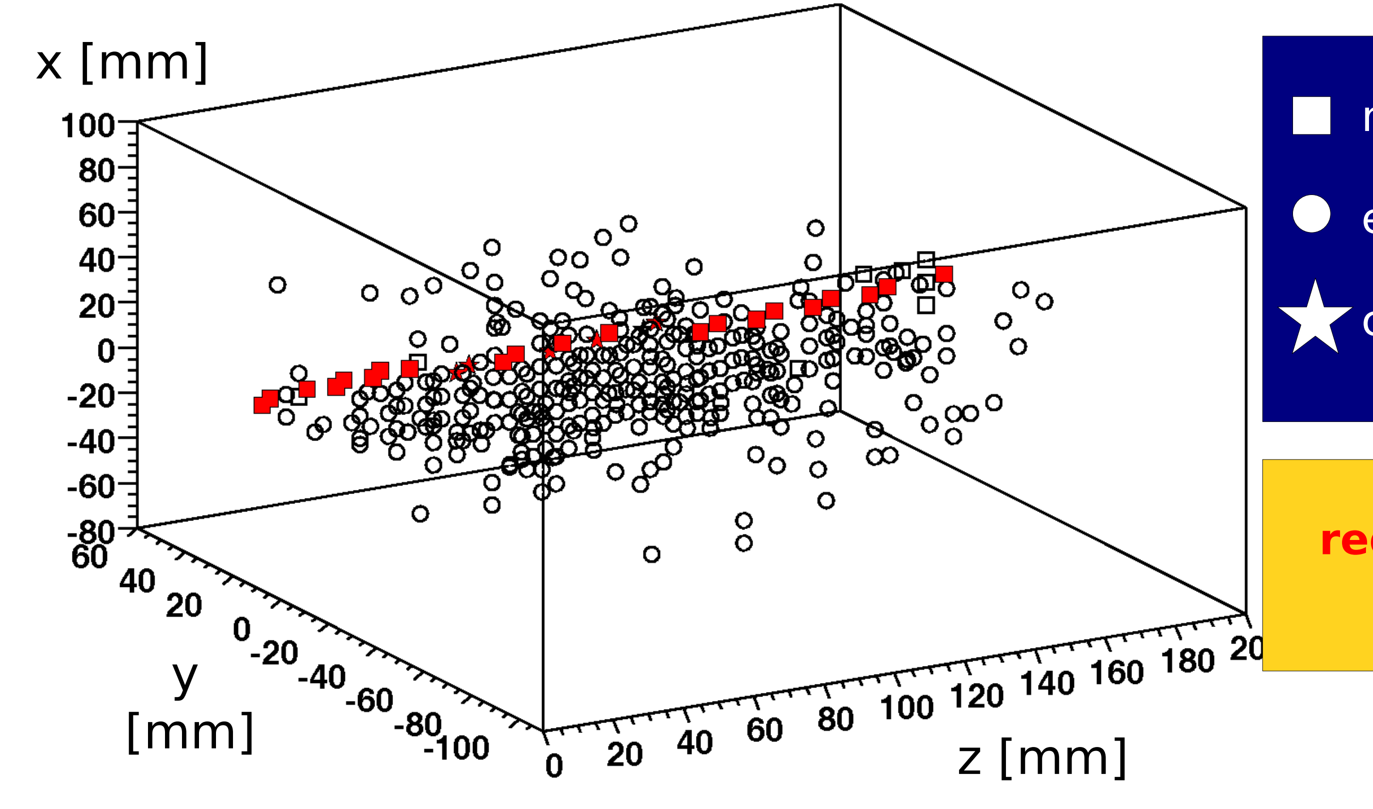
X-Y profile of beams at front face of ECAL



Electrons Muons

Example event

Simulated overlay event in SiW ECAL prototype



muon-induced hits

electron-induced hits

cells with mu/e signals

red: hits selected by algorithm

muon track nicely identified

Detection efficiency

- Overall (event-wise) detection efficiency defined by

$$E = \frac{num.\,of\ detected\ MIP\ overlays}{num.\,of\ generated\ events}$$

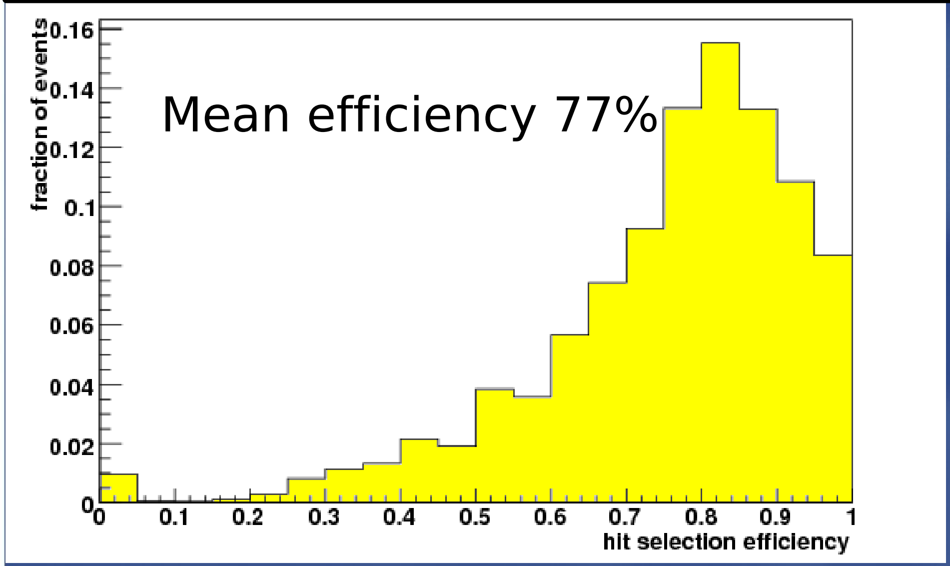
- Detection efficiency w/o any requirements (remember 5% of MIPs are almost undetectable)

E = 75 %

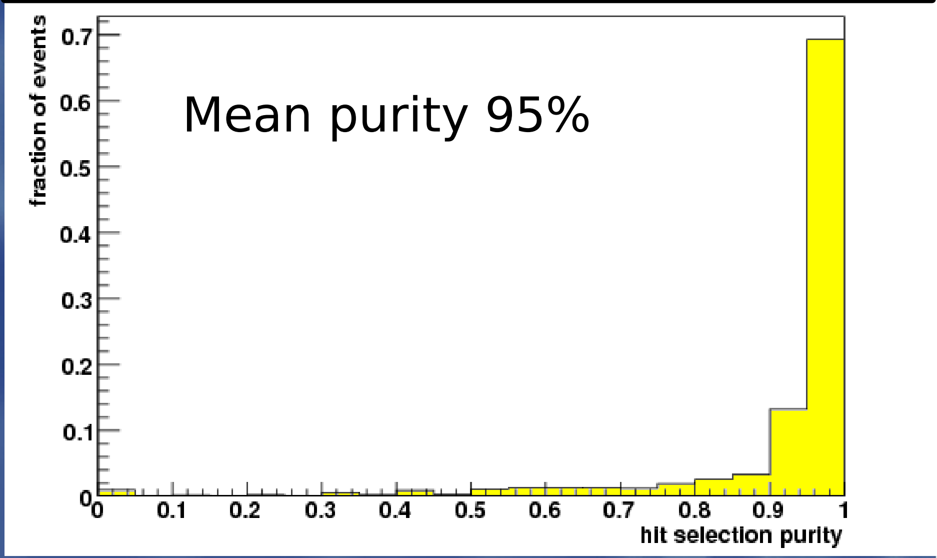
Hit selection efficiency and purity

$$\epsilon = \frac{num.\ of\ selected\ MIP\ cells}{true\ num.\ of\ cells\ with\ MIP\ signals} \qquad p = \frac{num.\ of\ selected\ MIP\ cells}{num.\ of\ selected\ cells}$$

Hit selection efficiency



Hit selection purity



Very high purity, good efficiency

Shower / MIP distance

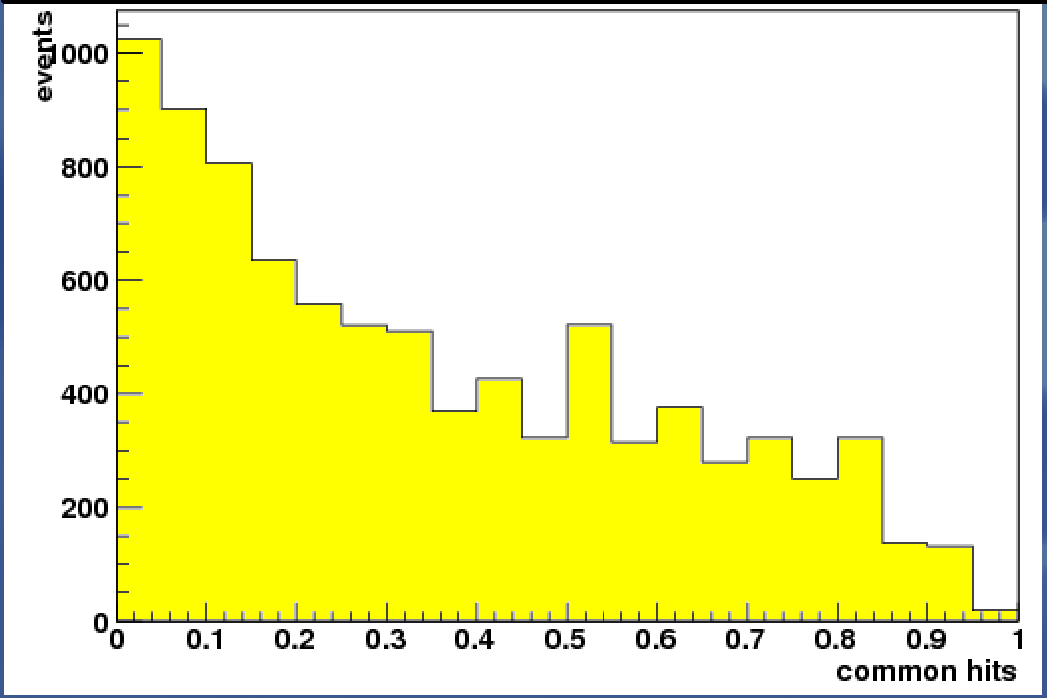
- MIP ID / tracking quality depends on shower / MIP distance
- use fraction of common cells (or proximity) to describe performance:

$$\kappa = \frac{num.\,of\ cells\ with\ MIP \wedge shower\ signals}{num.\,of\ cells\ with\ MIP\ signals}$$

MIP and shower completely separated at k = 0

MIP completely hidden in shower if k = 1

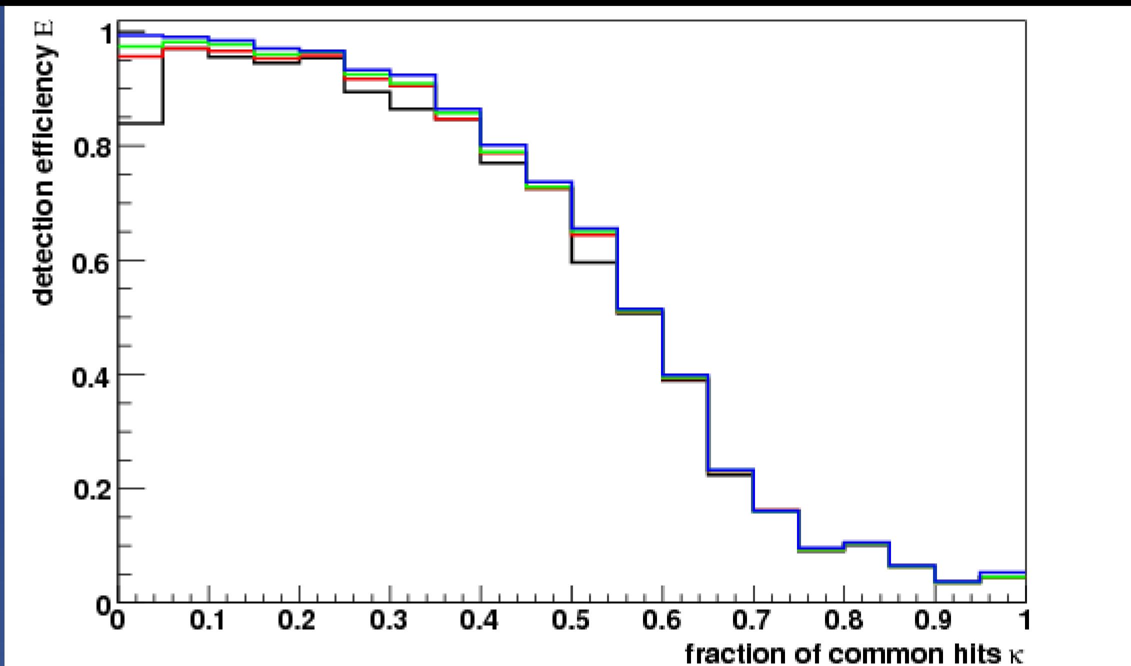
Distribution of fraction of common cells



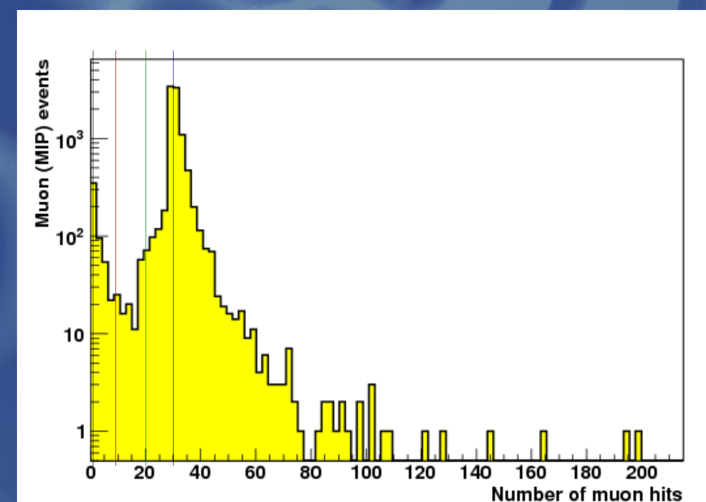
Detection efficiency

$$E = \frac{num.\ of\ detected\ MIP\ overlays}{num.\ of\ generated\ events}$$

Detection efficiency vs. proximity



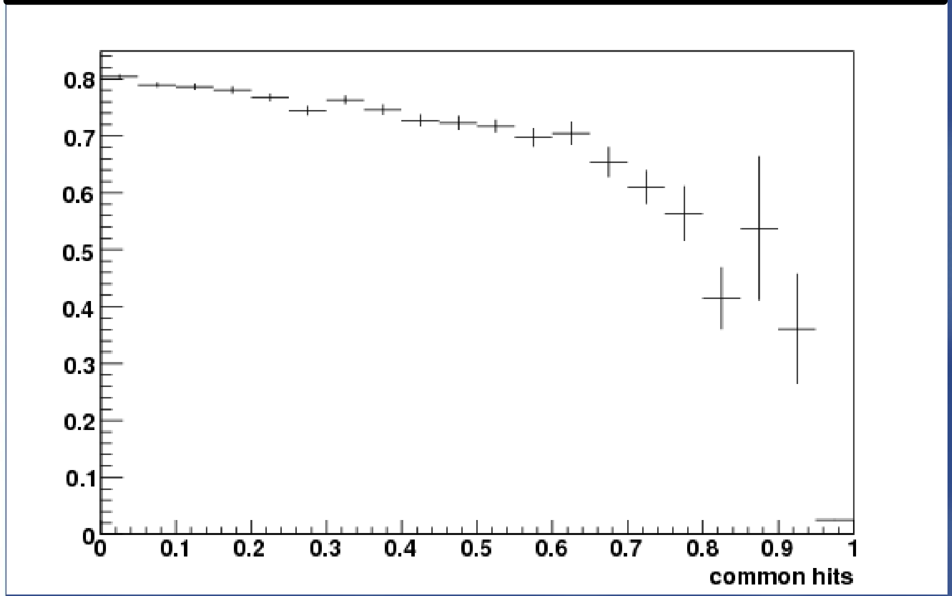
Number of MIP hits
>= 0 >= 10
>= 20 >= 30



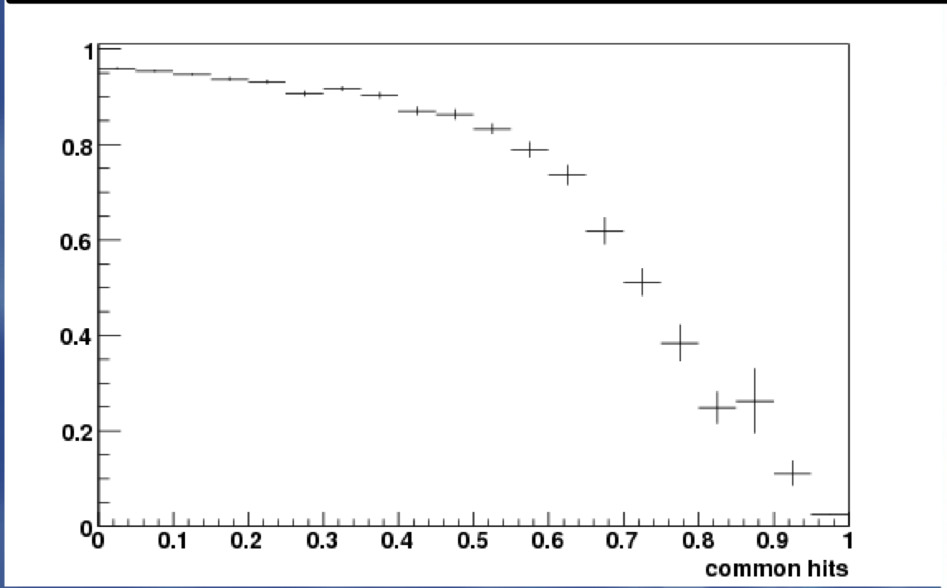
- At 50 % contamination still 70 % detection efficiency
- Dip at k=0 caused by undetectable muons (nhit < 10)

Hit selection efficiency and purity

Average efficiency vs. proximity



Average purity vs. proximity



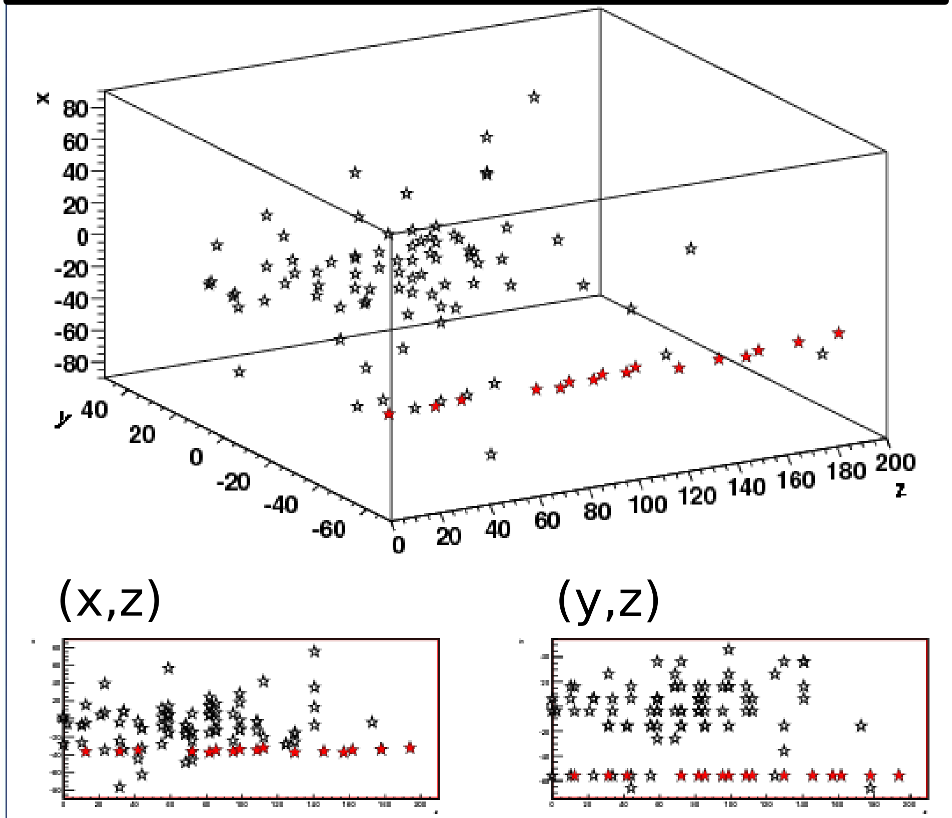
$$\epsilon = \frac{num.\ of\ selected\ MIP\ cells}{true\ num.\ of\ cells\ with\ MIP\ signals}$$

$$p = \frac{num.\ of\ selected\ MIP\ cells}{num.\ of\ selected\ cells}$$

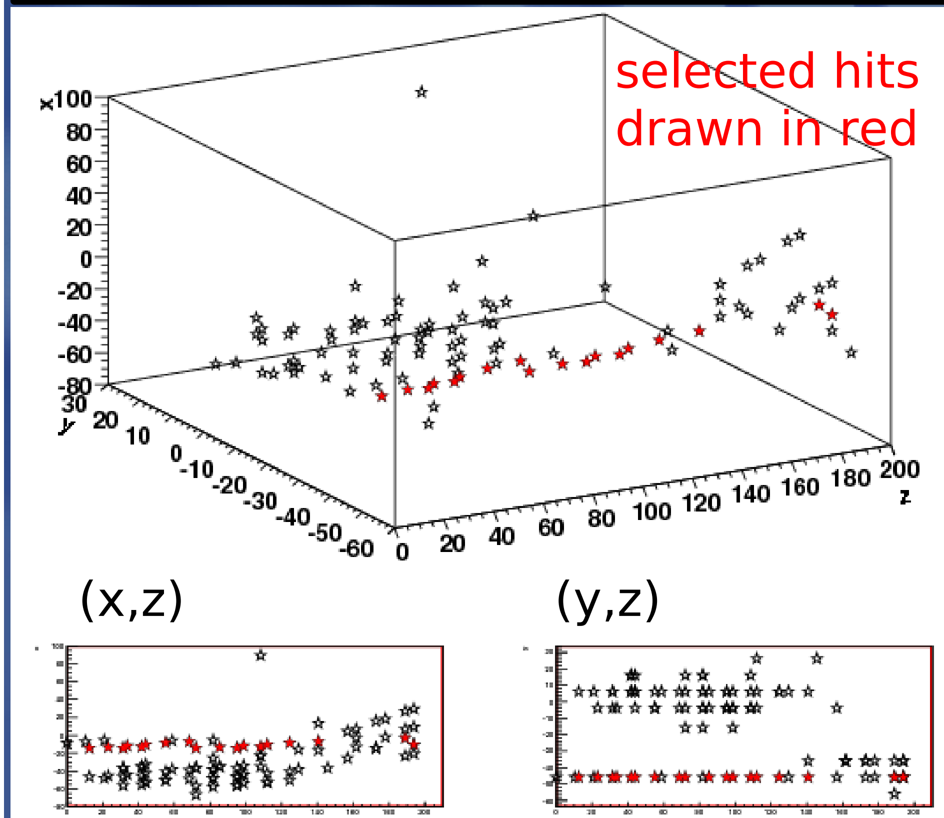
Application to test beam data

- Algorithm can now be used to clean test beam data:

Electron + MIP event



Electron + pion event



Summary and outlook

- **Tracking algorithm based on Hough transform:**
  - ~ 75% detection efficiency for overlay sample
  - high hit selection purities (> 90%) can be achieved while keeping good hit selection efficiency (~ 75%)
  - fraction of common hits useful for characterisation

- **Status / plans:**
  - work on algorithm (almost) finished
  - polish plots by adding more statistics
  - Analysis Note (~10-15 pages) in preparation will include a study of overlay impact on e-analysis