Faculty of
Mathematics and Physics
University of Ljubljana

Data analysis

(from raw data to physics results)

- From raw data to summary data

  ("Raw data -> DST")
  track fitting
  momentum determination
  calorimetry (cluster reconstr.)
  particle identification (Cherenkov angle)

- Calibration

  tracking detectors
  data (RICH) and MC (tracking) calibration

- Analysis

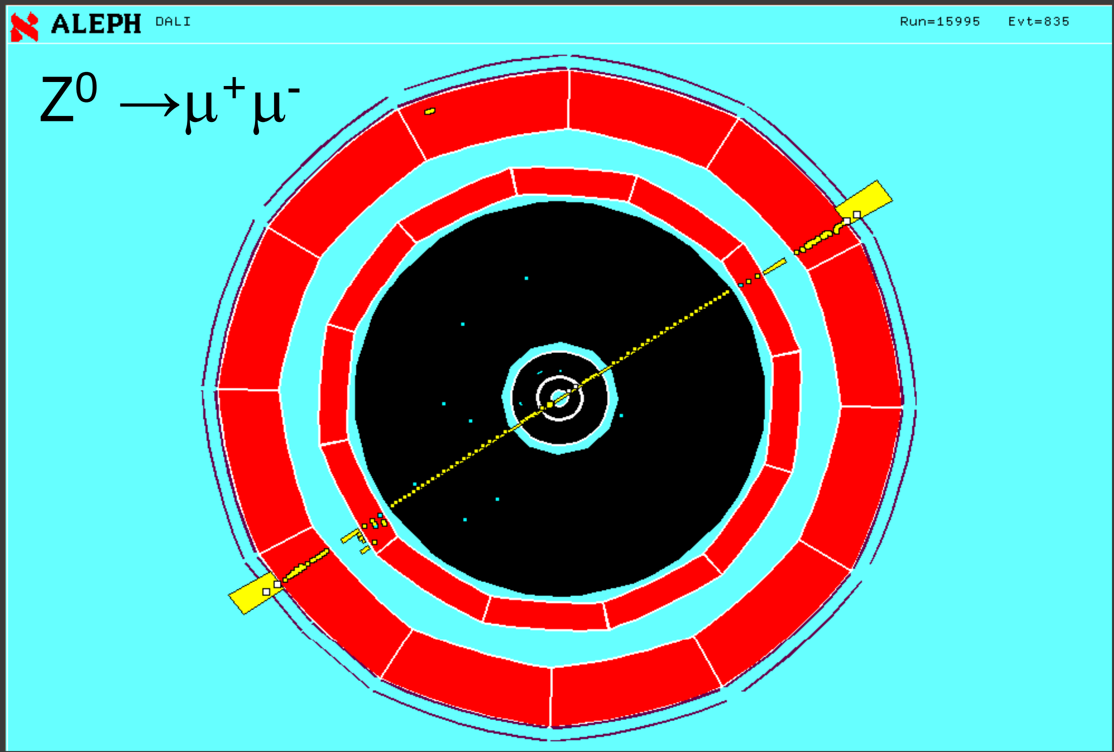
  stat. methods → other lectures
  jet reconstruction
  b-quark tagging
  flavour tagging
  fitting using kinematic constraints
  exclusive/inclusive channels
  neural networks → other lectures

From raw data to summary data

- Raw data: digitized record of detector electronic signals;

  directly used for graphical presentation;

detector part	signal value



for statistical analysis: need physics quantities
p, E, q, m, ...

- processed data, summary data, Data Summary Tape (DST)

From raw data to summary data

reconstruction

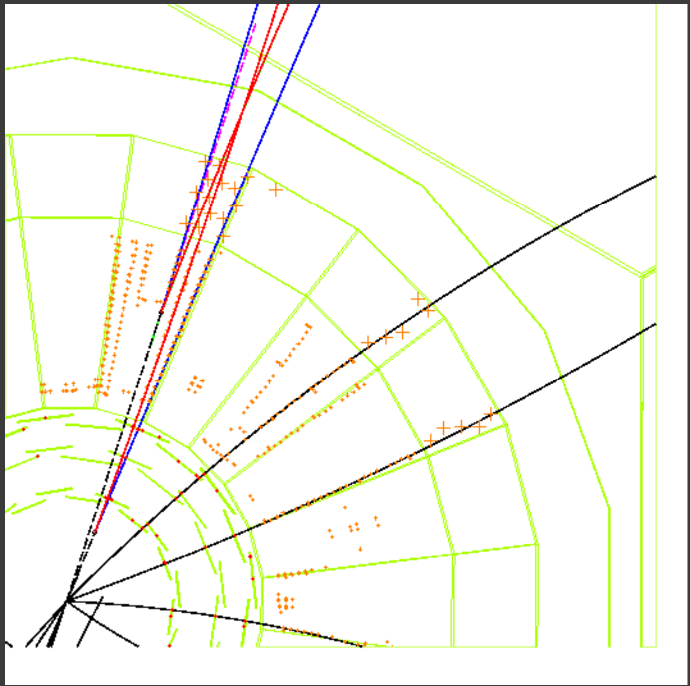
- Procedure of processing raw data to summary data: reconstruction

example: to conclude about $Z^0 \rightarrow \mu^+\mu^-$ decay one needs to

establish two tracks of corresponding **p**	determine small energy deposited in EM calorimeter(μ)	identify μ
↓	↓	↓
association of signals in tracking det. into tracks; track fitting; determination of **p**	association of signals in calorim. into clusters; association of clusters to tracks	hits in μ det.; association to tracks (different procedures for hadron ident.)

From raw data to summary data

track fitting

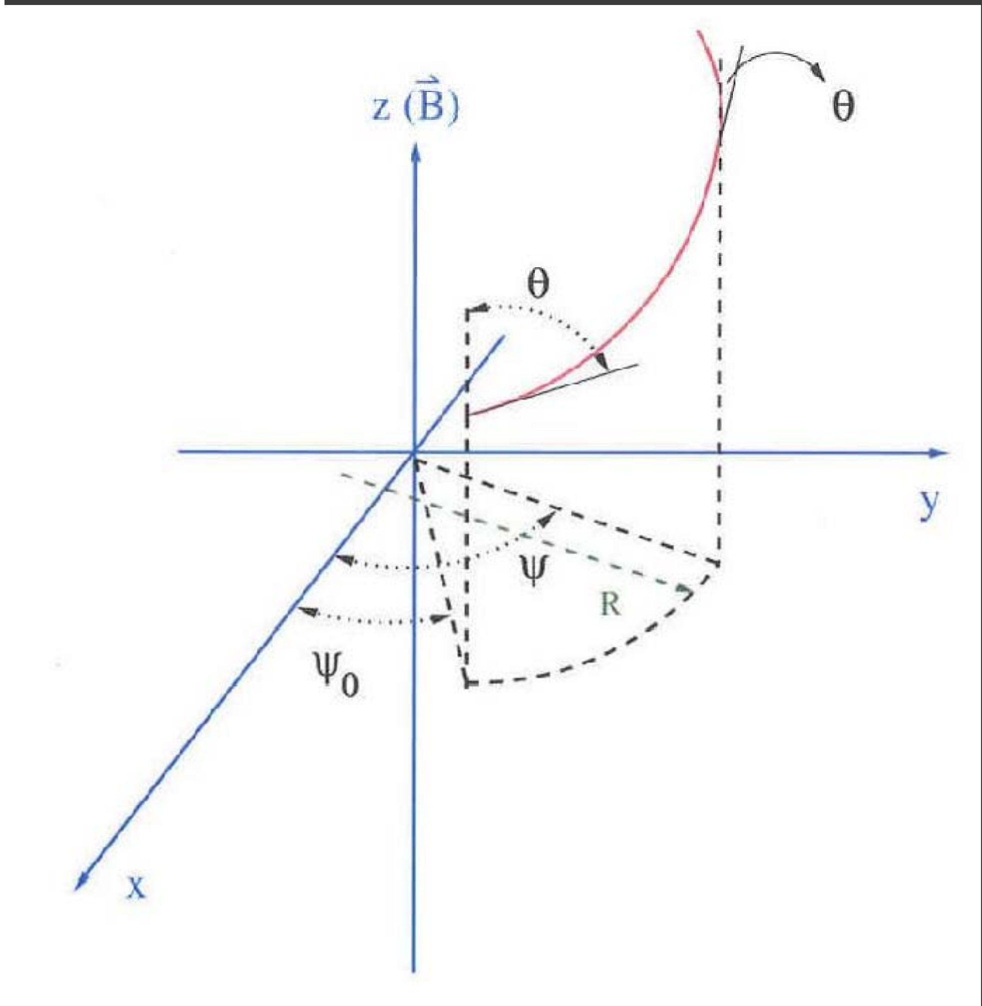


- charged track in **B** $\Rightarrow$ helix
- association of electronic signals in tracking detectors into groups - tracks
  pattern recognition
- fitting of helix parameters to associated hits
  track fitting

From raw data to summary data

helix

helix parameterization:



$$
\begin{aligned}
x &= x_0 + R(\sin\psi - \sin\psi_0) \\
y &= y_0 - R(\cos\psi - \cos\psi_0) \\
z &= z_0 + R\cot\theta(\psi - \psi_0)
\end{aligned}
$$

helix is parameterized with 5 parameters at chosen point, e.g.:

y_0, z_0, ψ_0, θ_0, $1/R$

($x_0=y_0/\tan\psi_0$)

From raw data to summary data
pattern recognition

- pattern recognition:
  high number of detector hits → difficult association with helix;
  transformation of transversal helix projection

  for parts of track (in most precise tracking detector) x_S, y_S, z_S, ψ_S, ϕ_S, $1/R_S$ available - TE, Track Element

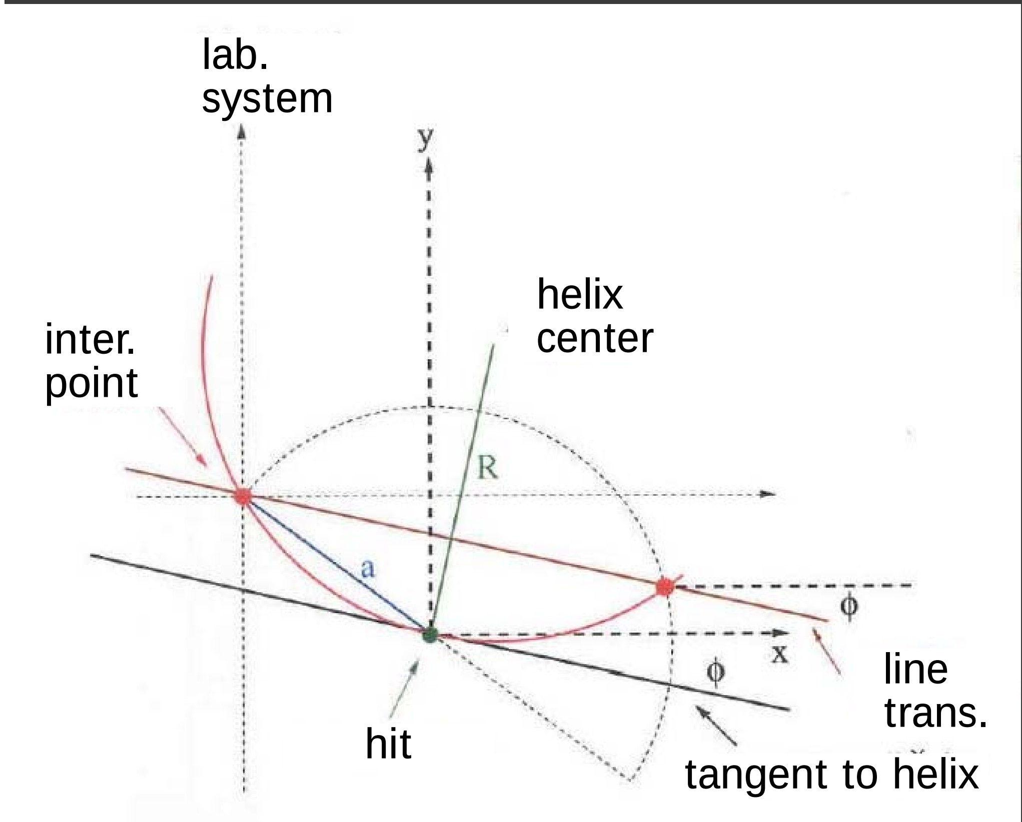
  for each other TE: calculate transformed point x', y';

  calculate ϕ' ($\angle(x', y')$ and int. point);
  check $|\phi' - \phi_S| < \alpha$;

  from $\Delta z = z - z'$ and ψ_S calculate ψ' $(\Delta z = R_S \cot\theta(\psi' - \psi_S))$;
  check $|\psi'(calculated) - \psi'(measured)| < \beta$

  join consistent TE's

From raw data to summary data
pattern recognition



transformation

$$x' = \frac{a^2 x}{x^2 + y^2}$$

$$y' = \frac{a^2 y}{x^2 + y^2}$$

circle $(x - x_c)^2 + (y - y_c)^2 = R^2$

line $y' = -\frac{x_c}{y_c} x' + \frac{a^2}{2y_c}$

$$\frac{dy}{dx}\Big|_{x=0,y=0} = -\frac{x_c}{y_c}$$

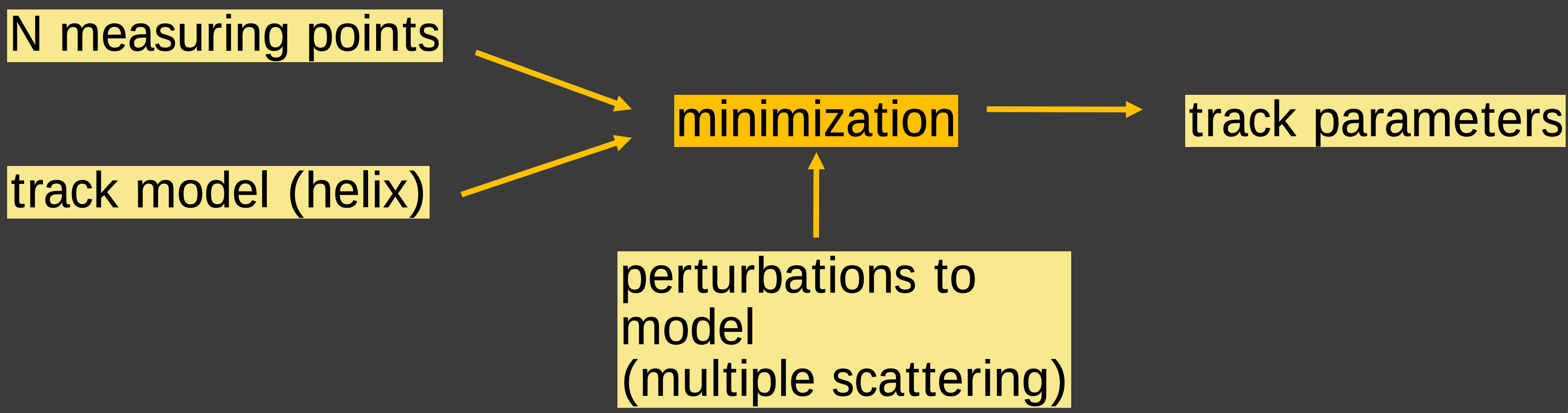
easier to check consistentcy of hits with straight line than with helix

From raw data to summary data

pattern recognition / track fitting

- algorithm properties:
  minimal number of loops;
  α, β determined individually for each sub-detector;
  using int. point - not applicable to secondary tracks;
  each TE can be associated to several tracks;
  additional info can be included (energy, direction, ...)

- track fit:
  from multiple TE's determine best helix parameters in chosen point
  (closest approach to int. point)



From raw data to summary data

track fitting

- Track fitting algorithms:
divided according to
track model usage, inclusion of model distortions (mult. scatt., energy losses)

Global Methods   Progressive Methods   Break Point Methods

- Global Methods:
simultaneous minimization of χ^2 of all measurement points;
mult. scatt. included in the error matrix

properties:
  all meas. points used simultaneously;
  simultaneous pattern recognition not possible
  (as opposed to Progressive methods);
  calculation expensive (NxN matrix inversion);

From raw data to summary data
track fitting

Global method - track model:
expected coordinate values

$$\begin{pmatrix} x^n_{\exp} \\ y^n_{\exp} \\ z^n_{\exp} \end{pmatrix} = \begin{pmatrix} x_0 + R_0^{-1}\left[\sin\psi_n - \sin\psi_0\right] \\ y_0 - R_0^{-1}\left[\cos\psi_n - \cos\psi_0\right] \\ z_0 + R_0^{-1}\cot\theta_0\left[\psi_n - \psi_0\right] \end{pmatrix}$$

5 free parameters: $\boldsymbol{p_0}=(y_0, z_0, \psi_0, \theta_0, 1/R)$
$(x_0=y_0/\tan\psi_0)$

N measured 3-dimensional points $\Rightarrow$ N 3-dimensional functions depending on 5 parameters $\boldsymbol{f}(\boldsymbol{p_0})$

global χ^2 minimization:

$$\chi^2(\vec{p}_0) = \left(\vec{f}(\vec{p}_0) - \vec{m}\right)^T \vec{C}^{-1}\left(\vec{f}(\vec{p}_0) - \vec{m}\right)$$

From raw data to summary data
track fitting

Global method - example:
straight line fit

model: $y_n = kx_n + y_0$
N meas. of y at x_n

$$\chi^2 = \sum_{n=1}^{N} \frac{(y_n - kx_n - y_0)^2}{\sigma_n^2}$$

minimization yields

$$k \sum_{n=1}^{N} \frac{x_n^2}{\sigma_n^2} + y_0 \sum_{n=1}^{N} \frac{x_n}{\sigma_n^2} - \sum_{n=1}^{N} \frac{y_n x_n}{\sigma_n^2} = 0$$

$$k \sum_{n=1}^{N} \frac{x_n}{\sigma_n^2} + y_0 \sum_{n=1}^{N} \frac{1}{\sigma_n^2} - \sum_{n=1}^{N} \frac{y_n}{\sigma_n^2} = 0$$

for $x_n = n\Delta x$ and $\sigma_n = \sigma \Rightarrow$

$$k = \frac{1}{\Delta x} \frac{N \sum n y_n - \sum n \sum y_n}{N \sum n^2 - (\sum n)^2}$$

N	$k\Delta x$	$\sigma_k \Delta x$
2	$y_2 - y_1$	$\sqrt{2}\sigma$
3	$(y_3 - y_1)/2$	$\sigma/\sqrt{2}$
4	$(3y_4 + y_3 - y_2 - 3y_1)/10$	$\sigma/\sqrt{5}$

From raw data to summary data
track fitting

- Progressive method (Kalman filter):
  parameters after n measurement points known;
  extrapolate (track model) to $(n+1)^{st}$ point;
  parameters after n+1 points = average of extrapolated and
  measured parameters at $(n+1)^{st}$ point;

  properties:
  enables simultaneous pattern recognition and track fitting;
  specific scattering regions inherently included;

From raw data to summary data

track fitting

Progressive method:

vector of parameters after n measurement points $\mathbf{p}_n^F$;
error matrix after n measurement points W_n;
vector of extrapolated parameters $\mathbf{p}_n^{Fe}$;
extrapolated error matrix W_n^e

$$W_n^e = D^T W_n D, \quad D = \frac{\partial \vec{p}}{\partial \vec{p}^e}$$

χ^2: sum of contribution from extrapolation and measurement:

$$\chi^2(\vec{p}_{n+1}) = \underbrace{\chi^2(\vec{p}_n^F)}_{\chi^2 \text{ from n points}} + \underbrace{\left[\vec{p}_{n+1} - \vec{p}_n^{Fe}\right]^T W_n^e \left[\vec{p}_{n+1} - \vec{p}_n^{Fe}\right]}_{\chi^2 \text{ from extrapolat.}} + \underbrace{\left[\vec{p}_{n+1} - \vec{p}_{n+1}^{izmer}\right]^T U \left[\vec{p}_{n+1} - \vec{p}_{n+1}^{izmer}\right]}_{\chi^2 \text{ of (n+1)}^{st} \text{ point}}$$

after minimization: set of equations for $\mathbf{p}_{n+1}^F$;
if χ^2 from extrapol. larger than chosen value for specific point
$\Rightarrow$ point not assigned to track

From raw data to summary data
track fitting

Progressive method - example:
straight line;
y_n^F and k_n^F after n measurement points;
extrapolation to $(n+1)^{st}$ point:

$$\vec{p}_n^{\,Fe} = \begin{pmatrix} y_n^F + k_n^F \Delta x \\ k_n^F \end{pmatrix} = \begin{pmatrix} y_n^{Fe} \\ k_n^{Fe} \end{pmatrix}$$

extrapolated error matrix:

$$D = \frac{\partial \vec{p}}{\partial \vec{p}^{\,e}} = \begin{pmatrix} \frac{\partial y_n}{\partial y_n^e} & \frac{\partial y_n}{\partial k_n^e} \\ \frac{\partial k_n}{\partial y_n^e} & \frac{\partial k_n}{\partial k_n^e} \end{pmatrix} = \begin{pmatrix} 1 & -\Delta x \\ 0 & 1 \end{pmatrix}$$

$$W_n^e = D^T W_n D$$

From raw data to summary data
track fitting

Progressive method - example:
straight line;
start with first point,
$y_1^F = y_1^{Fe} = y_1^m$,
$k_1^F = k_1^{Fe} = k_1^m = 0$

starting error matrix:

$$W_1 = \begin{bmatrix} 1/\sigma^2 & 0 \\ 0 & 0 \end{bmatrix},$$

$$\Rightarrow W_1^e = \begin{bmatrix} 1/\sigma^2 & -\Delta x/\sigma^2 \\ -\Delta x/\sigma^2 & \Delta x^2/\sigma^2 \end{bmatrix}$$

$$W_2 = W_1^e + U = \begin{bmatrix} 2/\sigma^2 & -\Delta x/\sigma^2 \\ -\Delta x/\sigma^2 & \Delta x^2/\sigma^2 \end{bmatrix}$$

N	$k^F \Delta x$	$\sigma_k^F \Delta x$
2	$y_2 - y_1$	$\sqrt{2}\sigma$
3	$(3y_3 - y_2 - 2y_1)/5$	$\sqrt{(14/25)}\sigma = 0.748\sigma$
4	$(30y_4 - y_3 - 18y_2 - 11y_1)/70$	0.524σ

etc.

From raw data to summary data

track fitting

Global method – multiple scattering:
error matrix:

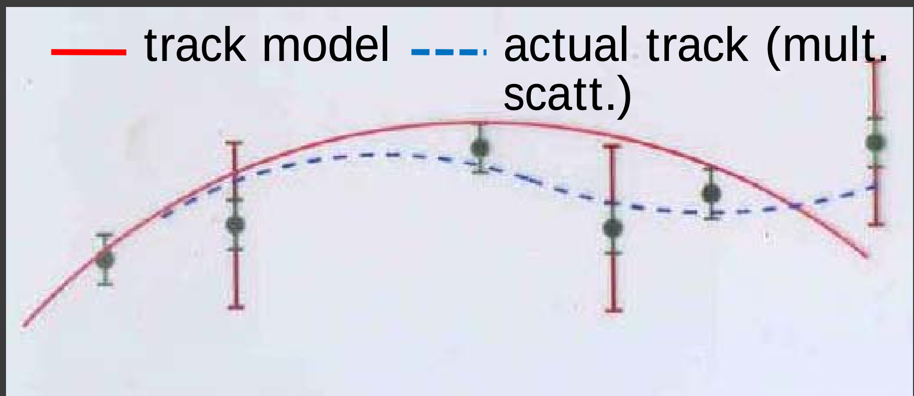
$$C_{ij} = \sigma_i \sigma_j \delta_{ij} + \overline{\epsilon_i^{\rm MS} \epsilon_j^{\rm MS}}$$

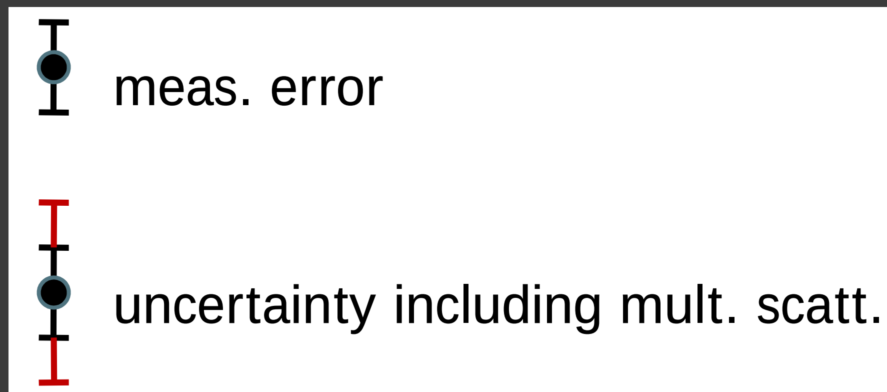
σ_i: uncertainty of ind. measurem.;
ϵ_i: contr. to uncertainty due to mult. scatt.
(Molièr formula:

$$\overline{\theta_i^{\rm MS}} = 0$$

$$\sqrt{\overline{(\theta_i^{\rm MS})^2}} = \frac{13{,}6\ MeV}{cp\beta}\sqrt{\frac{L}{X_0}}\left[1 + 0.038 \ln \frac{L}{X_0}\right]$$

)





distribution of $(y_{meas}-y_{fit})/\sigma_y$ ("pull") is a measure of understanding the effect of mult. scatt. rather than of understanding the meas. errors

From raw data to summary data
track fitting

- Progressive method – multiple scattering:
  mult. scatt. between n^{th} and $(n+1)^{st}$ point:

$$W_n^e = \left[\left[D^T W_n D \right]^{-1} + W_{\mathrm{MS}}^{-1} \right]^{-1}$$

  included in the error matrix extrapolation;

  using a corresponding mult. scatt. matrix W_{MS} one can include specifics of material between n^{th} and $(n+1)^{st}$ point

From raw data to summary data
track fitting

- Break points method:
appropriate for detectors with a limited number of regions with significant scattering;

scattering angles included in χ^2 as free parameters

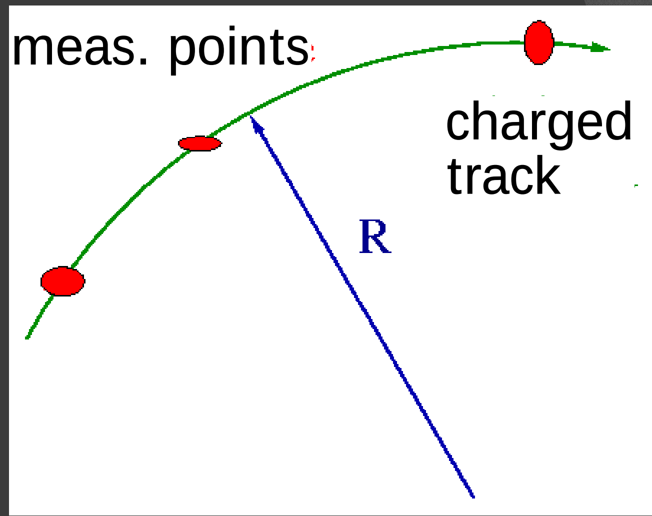
$\chi^2(p_n^F) \rightarrow \chi^2(p_n^F,\theta_n)$

From raw data to summary data
momentum measurement

Magnetic field:
$p_t = qBR$;
from curvature R one determines the transverse (w.r.t. **B**) component of **p**;
actual meas. is curvature R;

accuracy depends on:
# of meas. points;
spatial resolution of each point;
mag. field integral BL;
momentum p;

multiple scattering;

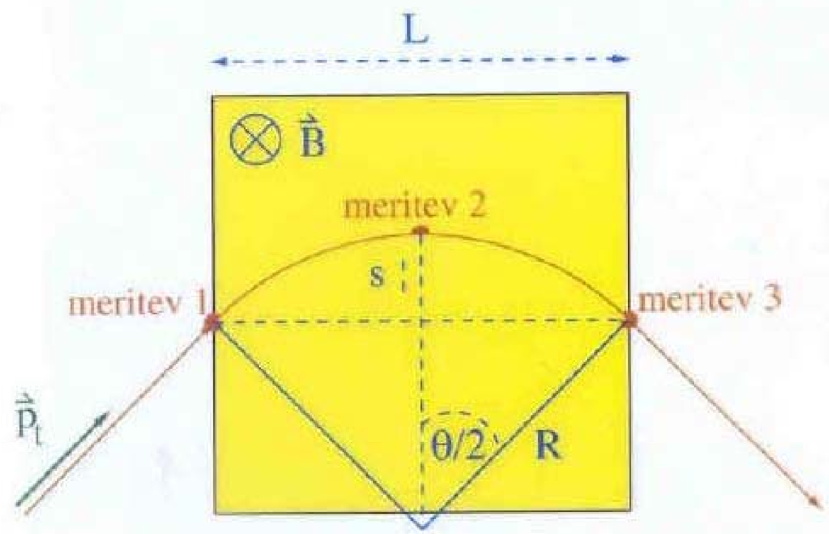


$$\frac{\sigma_{p_t}}{p_t} = \sqrt{a p_t^2 + b}$$

intrinsic resol. mult. scatt.

From raw data to summary data
momentum measurement

Example of momentum determination:



$$p_t = q\,B\,R \Rightarrow R = \frac{p_t}{qB}$$

$$\sin\frac{\theta}{2} \approx \frac{\theta}{2} = \frac{L}{2R}$$

$$\cos\frac{\theta}{2} = \frac{R-s}{R} \Rightarrow s \approx R\frac{\theta^2}{8}$$

$$s = \frac{B\,L^2\,q}{8p_t}$$

if s determined by
3 measurement points:

$$s = x_2 - \frac{x_1+x_3}{2}$$

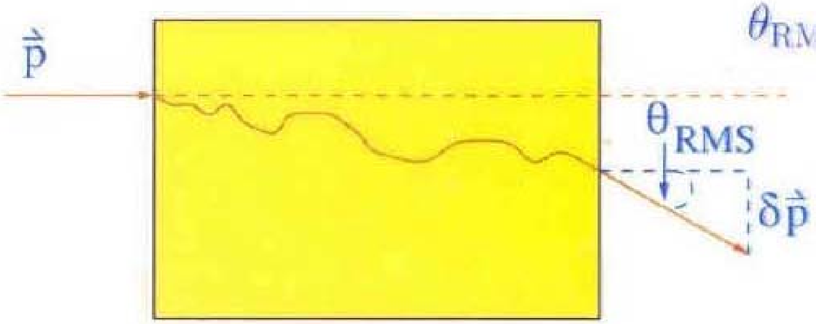
$$\frac{\sigma(p_t)}{p_t} = \frac{\sigma(s)}{s} = \frac{\sqrt{\frac{3}{2}}\sigma(x)8p_t}{BL^2q}$$

for N measurement points:

$$\frac{\sigma(p_t)}{p_t} = \frac{\sigma(x)8p_t}{BL^2q}\sqrt{\frac{720}{N+4}}$$

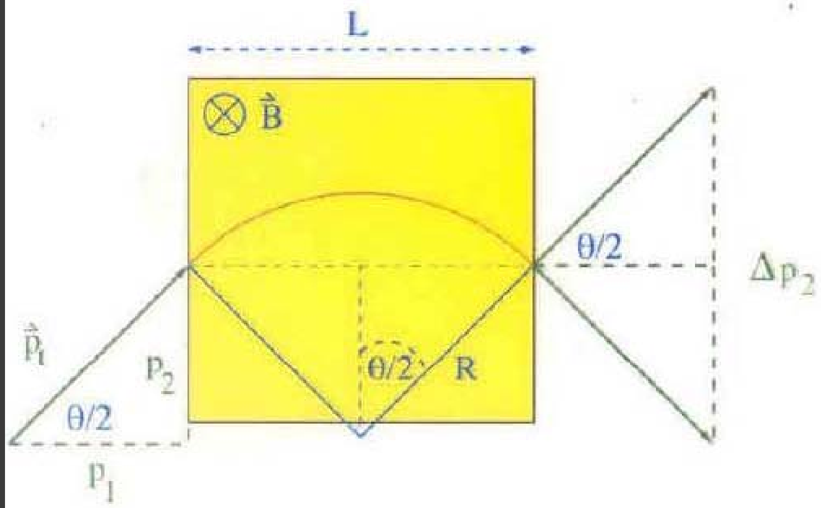
From raw data to summary data
momentum measurement

Multiple scattering:



$$\theta_{\rm RMS} \approx \frac{13.6\text{MeV}}{cp\beta}\sqrt{\frac{L}{X_0}}\left[1 + 0.038 \ln \frac{L}{X_0}\right]$$

$$\delta p = p \sin \theta_{\rm RMS}$$



$$\Delta p_2 = 2p_t \sin\frac{\theta}{2} \approx p_t\theta$$

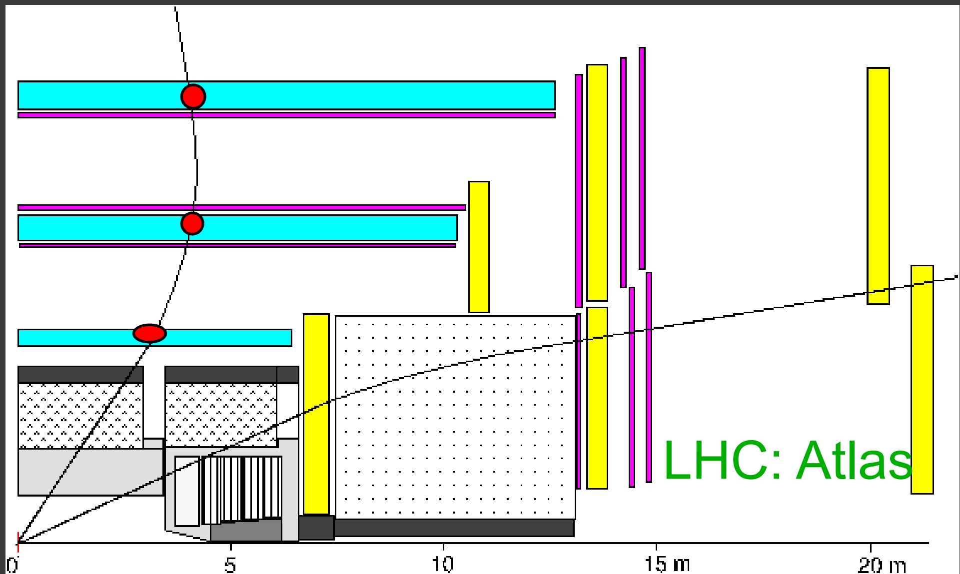
$$\frac{\sigma(p_t)}{p_t} = \frac{\sigma(\Delta p_2)}{\Delta p_2}$$

$$\frac{\sigma(p_t)}{p_t} \approx \frac{p_t \sin\theta_{\rm RMS}}{p_t\theta} \approx \frac{\theta_{\rm RMS}}{\theta}$$

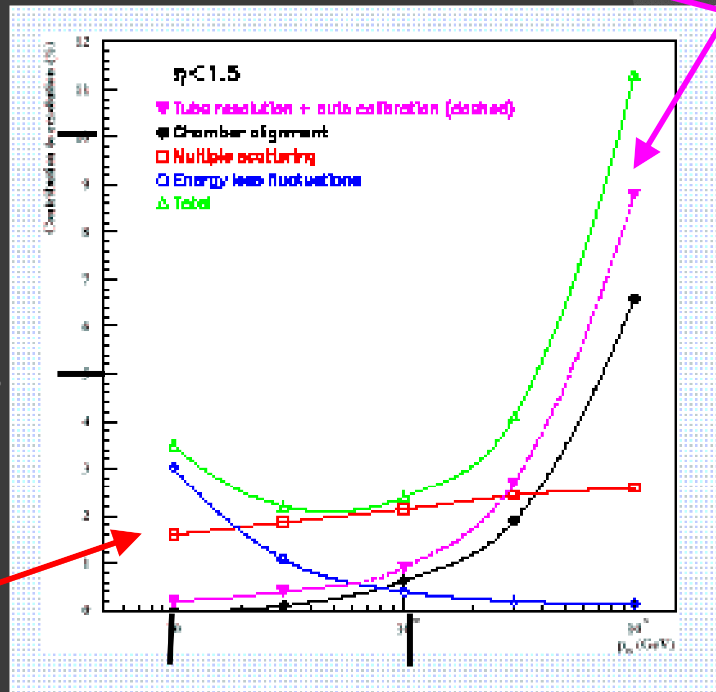
$$\frac{\sigma(p_t)}{p_t} \approx \frac{13.6\ \text{MeV}}{qB\sqrt{LX_0}}$$

From raw data to summary data
momentum measurement

Momentum meas. ATLAS (μ):



Muon Drift Tube chambers (MDT)
3 meas. points in barrel;
$\sigma(x)$ = 50 μm
L = 4 m
B = 1T (BL = 3 - 9 Tm)
1000 GeV μ from W',Z'
=> $\sigma(p)/p$ ~ 10%

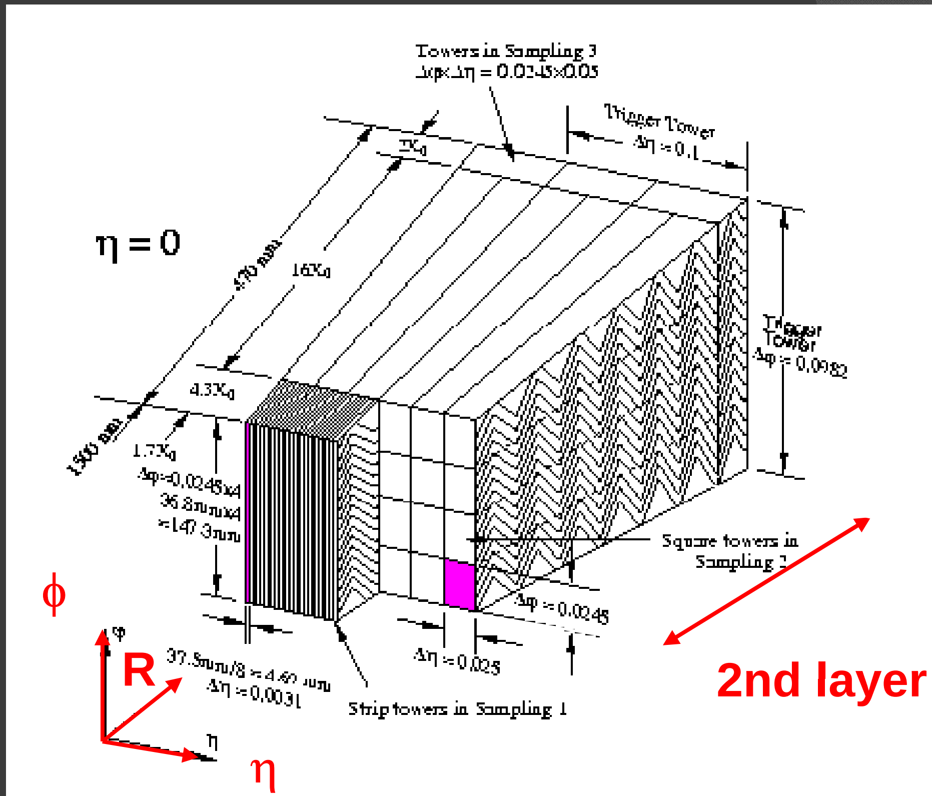


From raw data to summary data
calorimeter showers

- Calorimeters

are granulated
(composed of individual cells);
charged and neutral particles deposit energy in several cells;
to measure E of particle (or even hadronic jet) need method of associating individual cell energy deposit to particles ("clustering")

purpose of clustering:
improved signal/noise (considering correlations among cells);
separation of EM/hadronic showers;
search for isolated particles (e, γ, μ,...)



ATLAS LiAr EM calorim.:
accordion geom.;
3 layers in radial direction;
2nd layer: $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$
$\eta = -\ln(\mathrm{tg}\,\Theta/2)$ ($\Rightarrow$ 4 cm X 4 cm)

From raw data to summary data
calorimeter showers

- Reconstruction in few steps:
  - basic selection of cells
    - rejection of cells with known noise (online);
    - selection of cells with high signal (“seed”) and neighbouring cells with lower signal;
    - $E_i/E_{noise,i}$ > a (a=3,4,...) and $E_{i+1}/E_{noise,i+1}$>b (b=2,3,...);
  - association of cells into showers
    - several known alghoritms, e.g. Mulguisin alghoritm

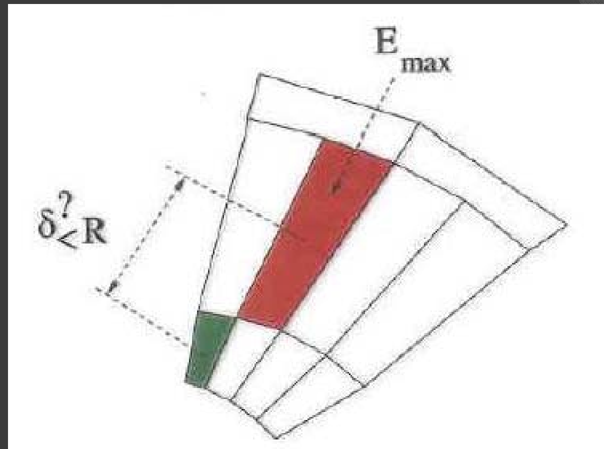
Hadron calorim.
for hadronic jets energy meas.;
precision of reconstruction reflects in invariant mass resolution

optimization of clustering depends on
luminosity;
process under study;

From raw data to summary data
calorimeter showers

Mulguisin algorithm:

- search for cell with largest E deposit – represents initial shower; dimension set to calorim. spatial resolution R_0;



- search for cell with 2nd largest E deposit;

- calculate distance between two cells;

  if smaller than shower dimension ⇒ assoc. cell to shower;

  (can calculate new shower center (weighted);
  new shower dimension can be set to max. dist. between shower center and each assoc. cell;)

  if larger than shower dimension ⇒ start of new shower with dimension R_0

- repeat until all cells taken into account;

From raw data to summary data
calorimeter showers

Mulguisin algorithm, implementations:

$$\begin{aligned} E_{k+1} &= E_k^{\text{pljusk}} + E_k^{\text{celica}} \\ \eta_{k+1}^{\text{pljusk}} &= \eta_k^{\text{pljusk}} \\ \phi_{k+1}^{\text{pljusk}} &= \phi_k^{\text{pljusk}} \\ R_{k+1}^{\text{pljusk}} &= R_0 \end{aligned}$$

celica → cell

pljusk → shower

size and direction of shower const.

$$\begin{aligned} E_{k+1} &= E_k^{\text{pljusk}} + E_k^{\text{celica}} \\ \eta_{k+1}^{\text{pljusk}} &= (E_k^{\text{pljusk}}\eta_k^{\text{pljusk}} + E_k^{\text{celica}}\eta_k^{\text{celica}})/(E_k^{\text{pljusk}} + E_k^{\text{celica}}) \\ \phi_{k+1}^{\text{pljusk}} &= (E_k^{\text{pljusk}}\phi_k^{\text{pljusk}} + E_k^{\text{celica}}\phi_k^{\text{celica}})/(E_k^{\text{pljusk}} + E_k^{\text{celica}}) \\ R_{k+1}^{\text{pljusk}} &= R_0 \end{aligned}$$

size of shower const., direct. recalculated at each step

$$\begin{aligned} E_{k+1} &= E_k^{\text{pljusk}} + E_k^{\text{celica}} \\ \eta_{k+1}^{\text{pljusk}} &= (E_k^{\text{pljusk}}\eta_k^{\text{pljusk}} + E_k^{\text{celica}}\eta_k^{\text{celica}})/(E_k^{\text{pljusk}} + E_k^{\text{celica}}) \\ \phi_{k+1}^{\text{pljusk}} &= (E_k^{\text{pljusk}}\phi_k^{\text{pljusk}} + E_k^{\text{celica}}\phi_k^{\text{celica}})/(E_k^{\text{pljusk}} + E_k^{\text{celica}}) \\ R_{k+1}^{\text{pljusk}} &= max(R_k, \delta) \end{aligned}$$

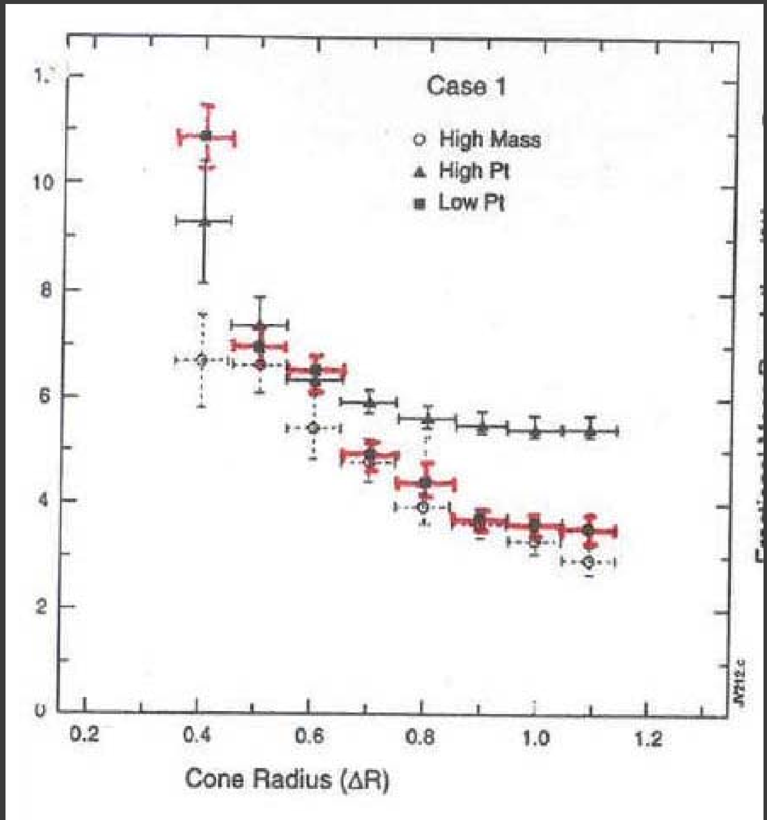
size and direct. of shower recalculated at each step

From raw data to summary data
calorimeter showers

Resolution on dijet invariant mass:

LHC simulation; individual event:

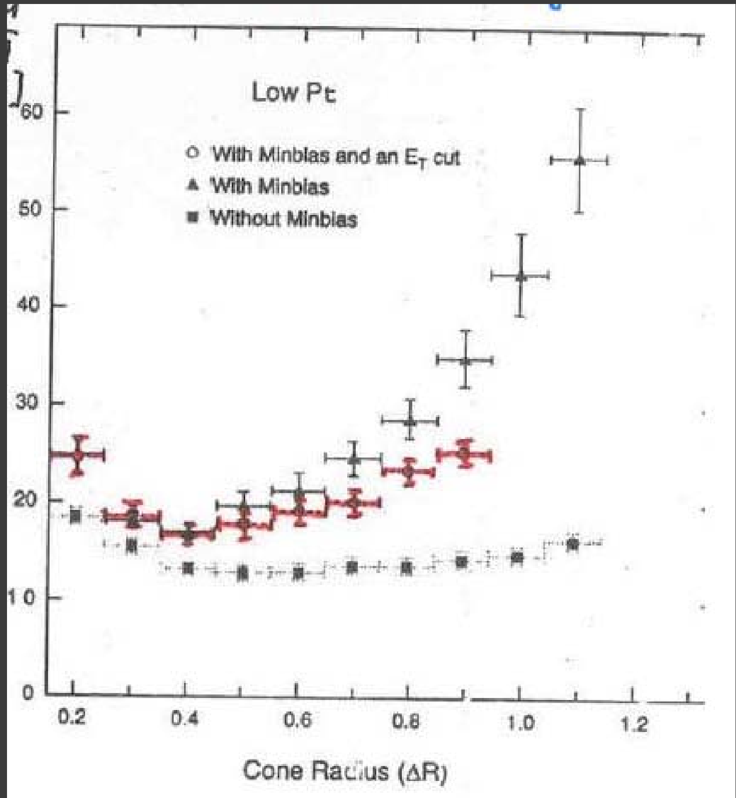
σ_M/M [%]



$\rightarrow$ increasing allowed shower size $\Rightarrow$ larger fraction of E reconstr. $\Rightarrow$ better resol.

event + 30 min. bias events:

σ_M/M [%]



$\leftarrow$ decreasing allowed shower size $\Rightarrow$ smaller fraction of E reconstr. $\Rightarrow$ worse resol.;

$\rightarrow$ increasing allowed shower size $\Rightarrow$ larger fract. of E from other events $\Rightarrow$ worse resol.

From raw data to summary data
particle identification

Hadron identification
most detectors use some variation of Cherenkov light detection;

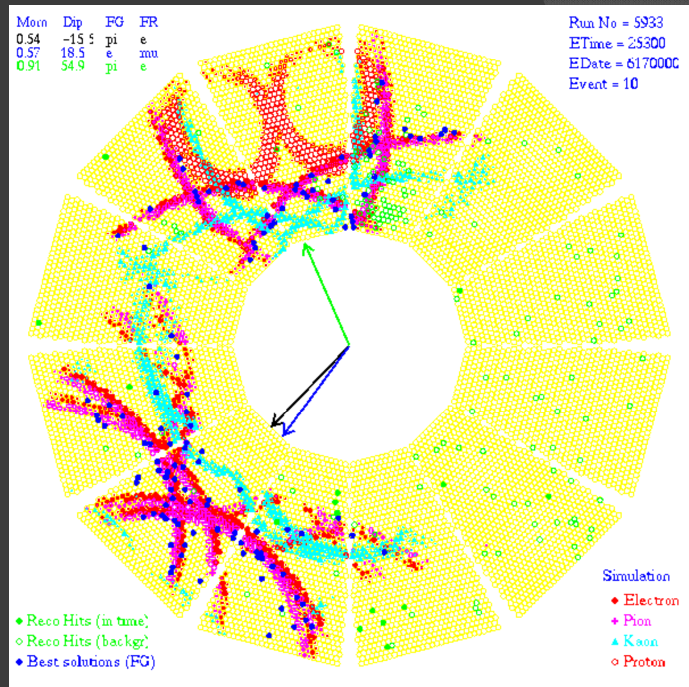
Cherenkov ring detectors:
photons in detector ⇒ radius of ring
⇒ Cherenkov angle ⇒ particle velocity ⇒ mass

large number of γ's – impossible to consider all combinations;

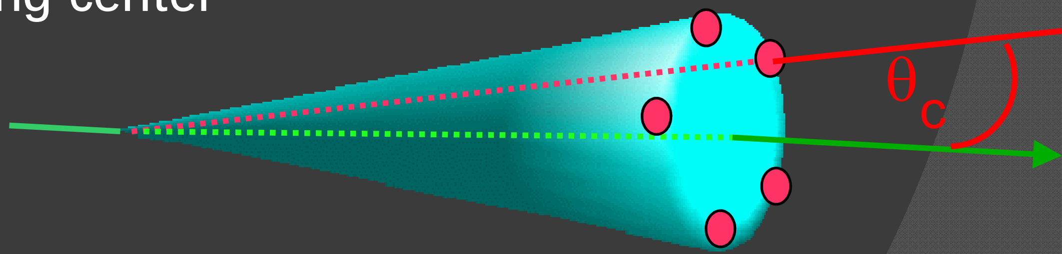
charged track (through geometry dependent equations) determines ring center;

consider only γ's consistent with ring center

$\theta_C = \theta_C^{exp}(m_i) \pm N\sigma_{\theta C}$



BaBar - DIRC



From raw data to summary data
particle identification

Cherenkov ring detectors

likelihood function:

$$\mathfrak{L}(N_{exp},N_{bg}) = \mathcal{P}_{\rm Poisson}\prod_{i=1}^{N}\mathcal{P}_i$$

$$\mathcal{P}_{\rm Poisson} = \frac{(N_{\rm exp}+N_{\rm bg})^N\, e^{-(N_{\rm exp}+N_{\rm bg})}}{N!}$$

$$\mathcal{P}_i = \frac{N_{\rm exp}}{N_{\rm exp}+N_{\rm bg}}\underbrace{\frac{1}{\sqrt{2\pi}\sigma_i}e^{\frac{-(\theta_i-\theta_C)^2}{2\sigma_i^2}}}_{\text{Gaussova porazdelitev }\gamma\text{ okoli }\theta_C(p_i,m_i)} + \frac{N_{\rm bg}}{N_{\rm exp}+N_{\rm bg}}\underbrace{\frac{\theta_i}{6\theta_C\sigma_i}}_{\text{enakomerna porazdelitev po povrsini detekt. x }\theta_i}$$

N_{exp}, $\theta_C = f(p_i,m_i)$

background:
uniform distrib. over detector surface

$\mathfrak{L}(N_{exp},N_{bg})/\mathfrak{L}(N_{exp}=0,N_{bg})$: measure of probability for set of γ's to originate from a particle with (p_i,m_i)

From raw data to summary data
particle identification

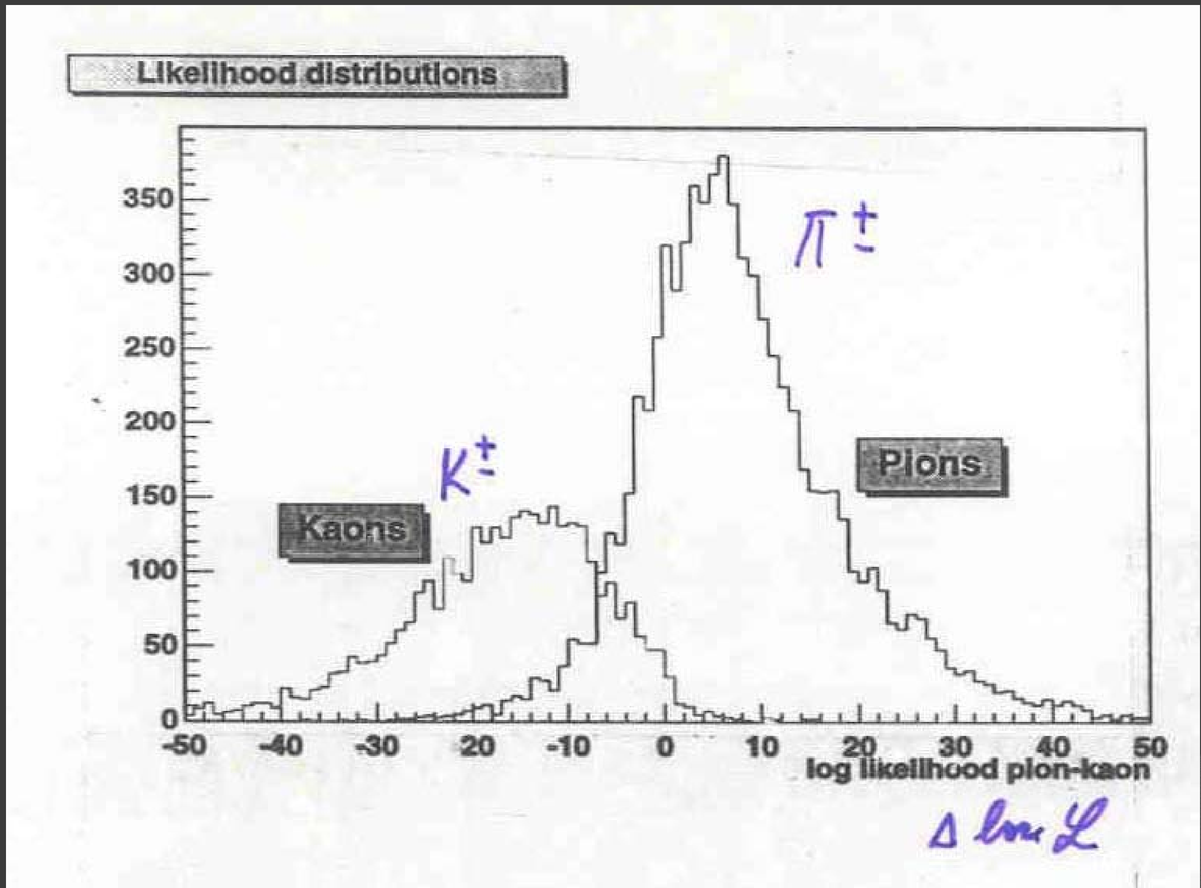
Cherenkov ring detectors

particle separation:

from $\mathcal{L}(N_{exp},N_{bg})/\mathcal{L}(N_{exp}=0,N_{bg}) \Rightarrow P(m_i)$

$P(m_i)/P(m_j)$ particle separation

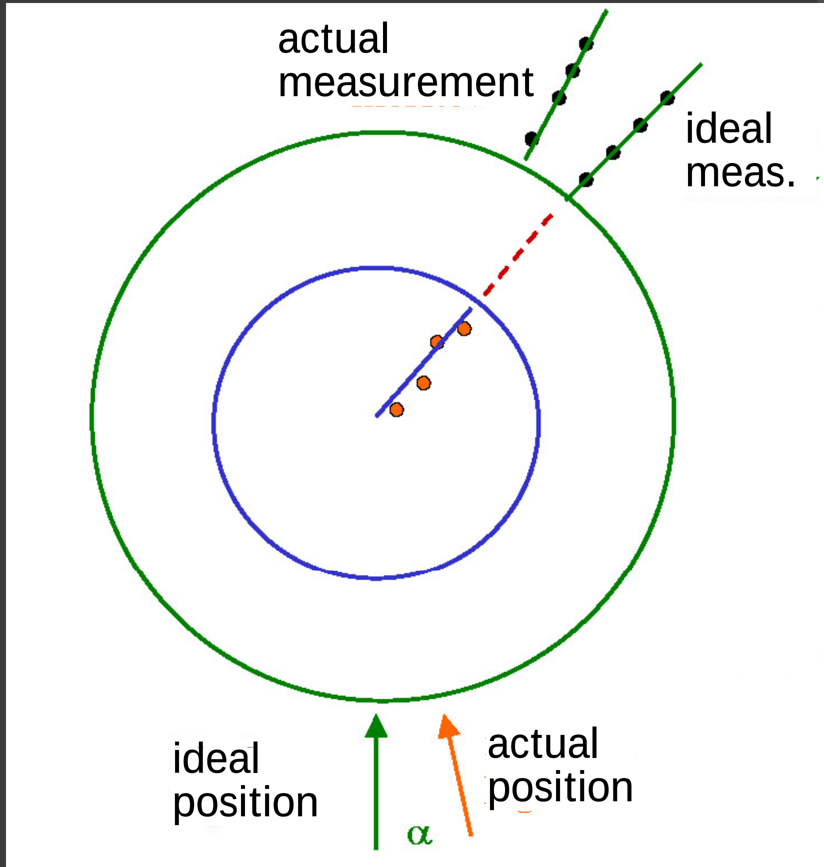
e.g. HERA-B $P(\pi, not\ K)/P(K, not\ \pi)$:



Calibration
Tracking detectors

- Tracking detectors calibration
  individual subdetectors must be properly inter-orineted, otherwise tracks distorted;

  for any calibration need
  sample (tracks, decays, ...) with precisely known detector response



Calibration

Tracking detectors

Description of detector (mis)alignment

position of individual subdetector w.r.t. reference (most precisely mechanically positioned detector)

described by set of small parameters α (translation, rotation, t-delay,...)

assume linear relation

$$\vec{q}^{\,meas} - \vec{q}^{\,ext} = S\vec{\alpha}$$

$\mathbf{q}^{meas}$: vector of measured coordinates
$\mathbf{q}^{ext}$: vector of extrapolated coord. (from the reference detector)
S: matrix depending on measuring coord., track model, detector geometry

simplest case:
α composed of 3 translations and 3 rotations
$\alpha = (\eta_x, \eta_y, \eta_z, \varepsilon_x, \varepsilon_y, \varepsilon_z)$

Calibration
Tracking detectors

Determination of position

minimization:

$$\chi^2 = \sum_{k \in \text{meas. points}} [\vec{q}_k^{\,meas} - \vec{q}^{\,ext} - S_k\vec{\alpha}]^T W_k^{-1} [\vec{q}_k^{\,meas} - \vec{q}^{\,ext} - S_k\vec{\alpha}]$$

result:

$$\Big(\sum_{k \in \text{meas. points}} S_k^T W_k^{-1} S_k\Big)\vec{\alpha} = \sum_{k \in \text{merske točke}} S_k^T W_k^{-1} (\vec{q}_k^{\,meas} - \vec{q}^{\,ext})$$

vector of displacements α

coordinates meas. in subdetector are corrected by α

Calibration

Tracking detectors

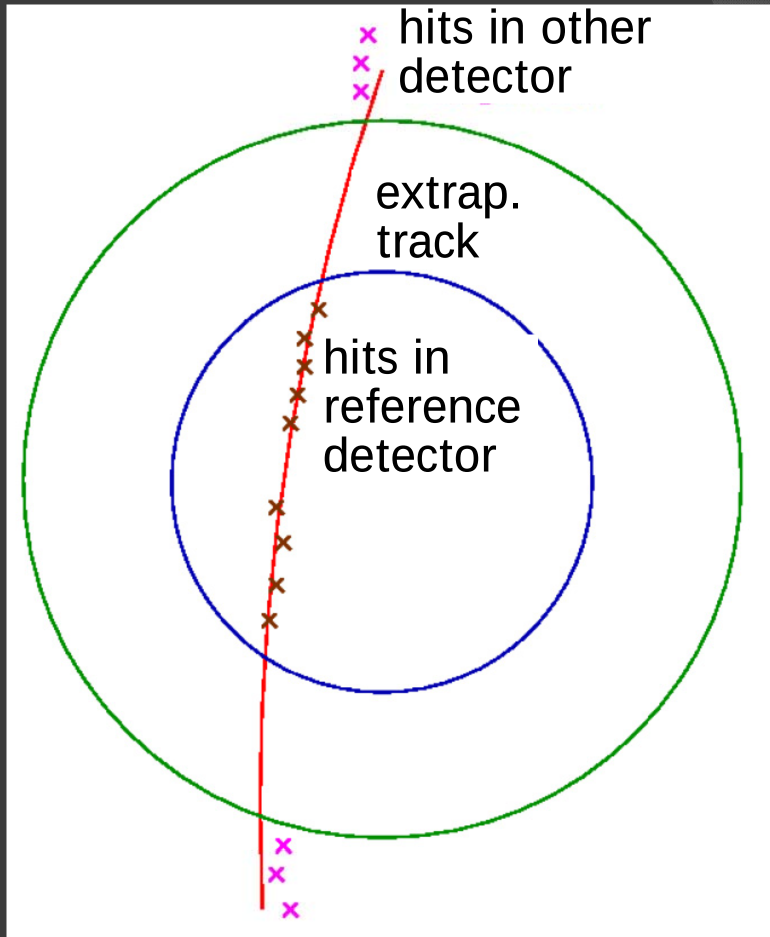
Appropriate sample

often cosmic rays;

other decays observed,

e.g. $Z^0 \rightarrow \mu^+\mu^-$ (LEP);

(needed also to check
the alignment method)

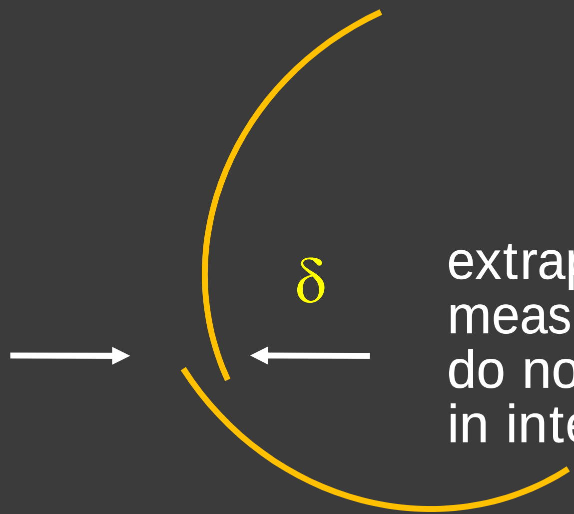


Calibration

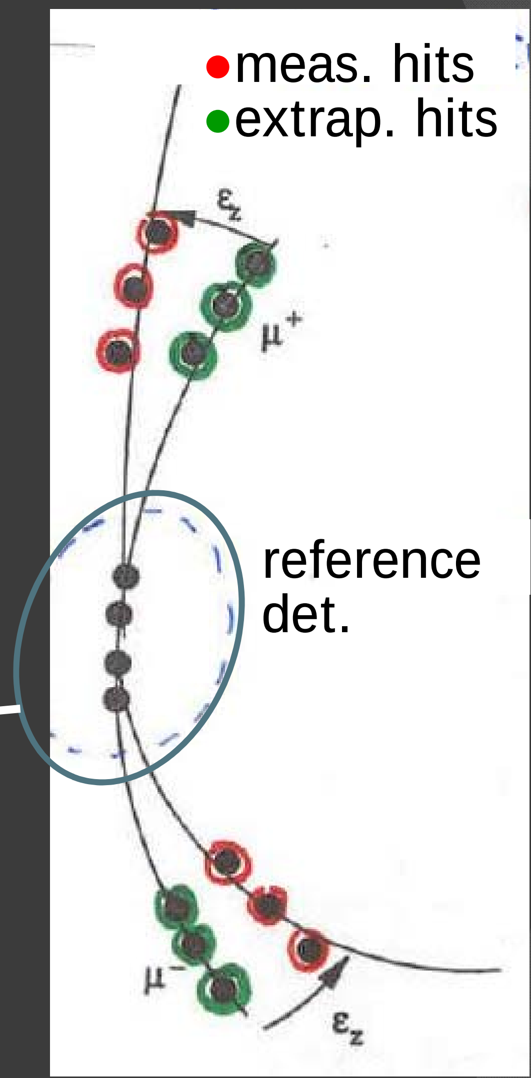
Tracking detectors

Appropriate sample

e.g. $Z^0 \rightarrow \mu^+\mu^-$ (LEP);



extrapolations of
meas. tracks
do not intersect
in interaction point



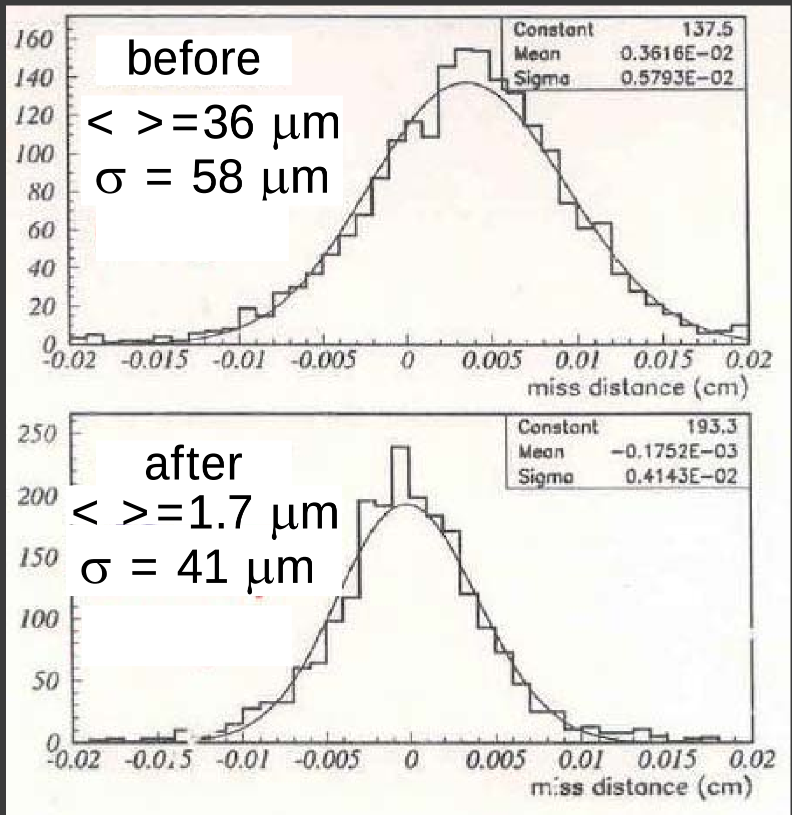
Calibration

Tracking detectors

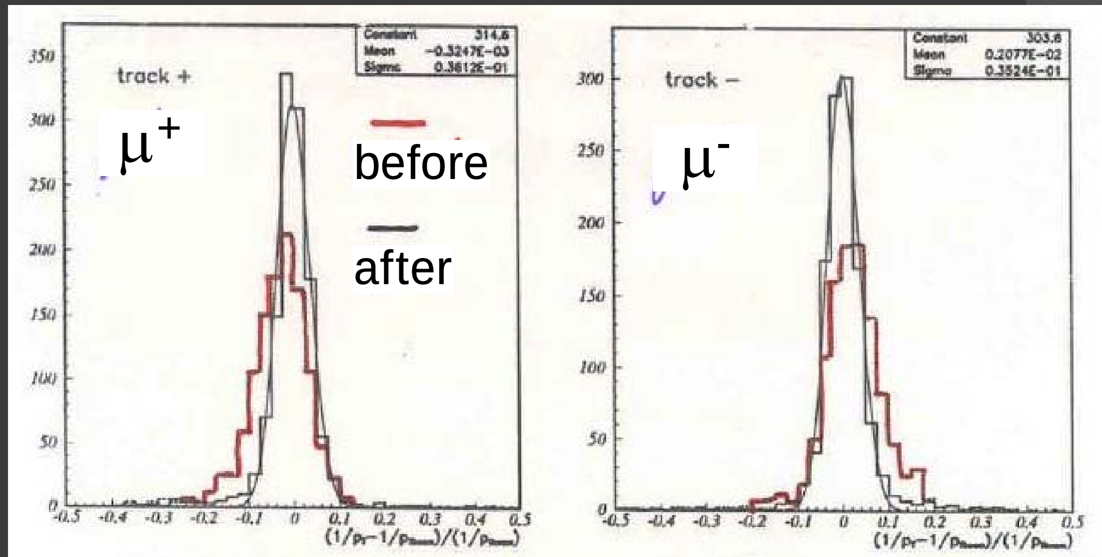
Example

Delphi detector at LEP

δ [cm]



$[(1/p_t)-(1/p_t^{ext})]/(1/p_t^{ext})$



Calibration

Data and simulation

- Calibration of data and MC simulation
  example of RICH (Delphi at LEP)

  sample with known detector response:

  $\cos\theta_C = 1/\beta n$

  tracks with p>6 GeV; even protons at p>6 GeV $\beta = 1 - 10^{-2}$

  sample yields value of n;

  expected error on θ_C, $\sigma(\theta_C)$ needed for fits;

  $(\theta_C^{meas} - \theta_C^{exp})/\sigma(\theta_c)$ ”pull” examined;

  pull distribution properties:

  for gaussian distribution of θ_C^{meas} :

  $< > = 0$

  $\sigma = 1$ if $\sigma \neq 1 \Rightarrow$ correct $\sigma(\theta_c)$

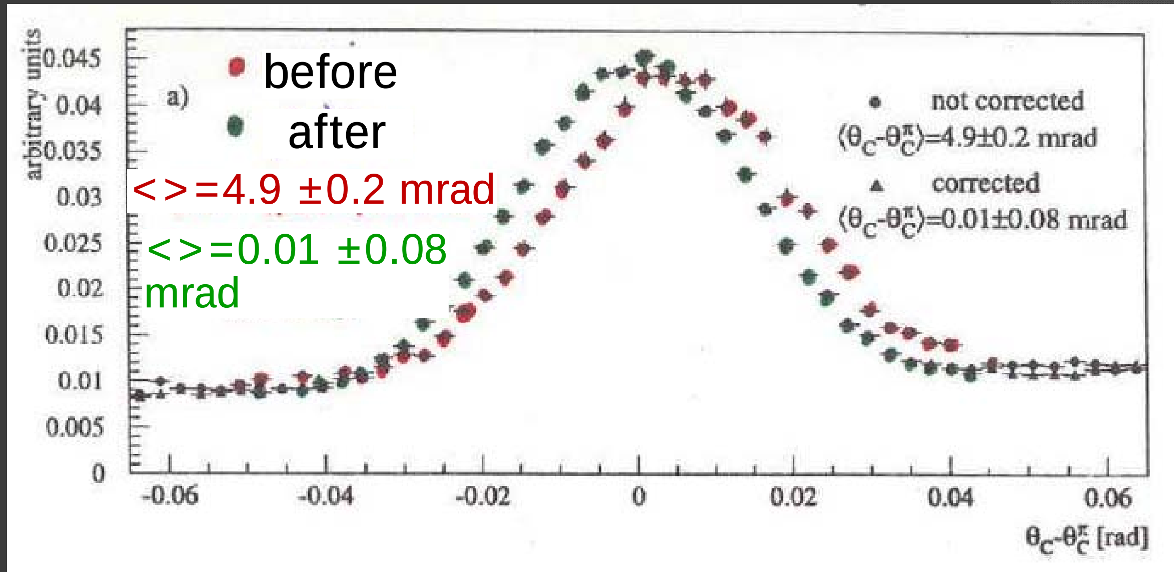
Calibration

Data and simulation

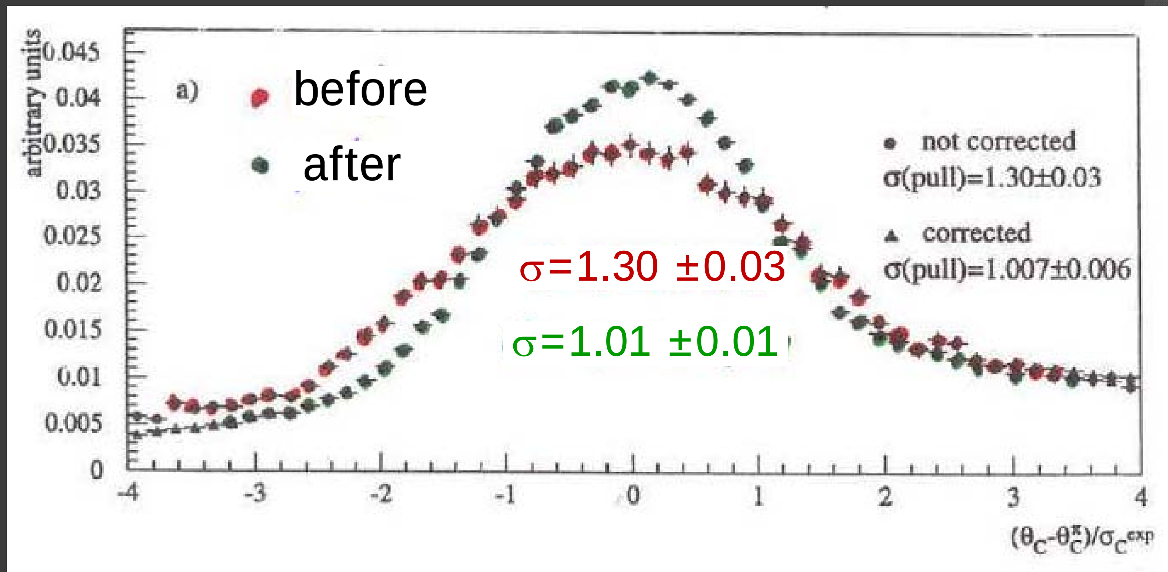
example of RICH
(Delphi at LEP)

same method and
corrections applied
to MC simulation
to match the data

$\theta_C^{meas} - \theta_C^{exp}$ [rad]



$(\theta_C^{meas} - \theta_C^{exp})/\sigma(\theta_c)$



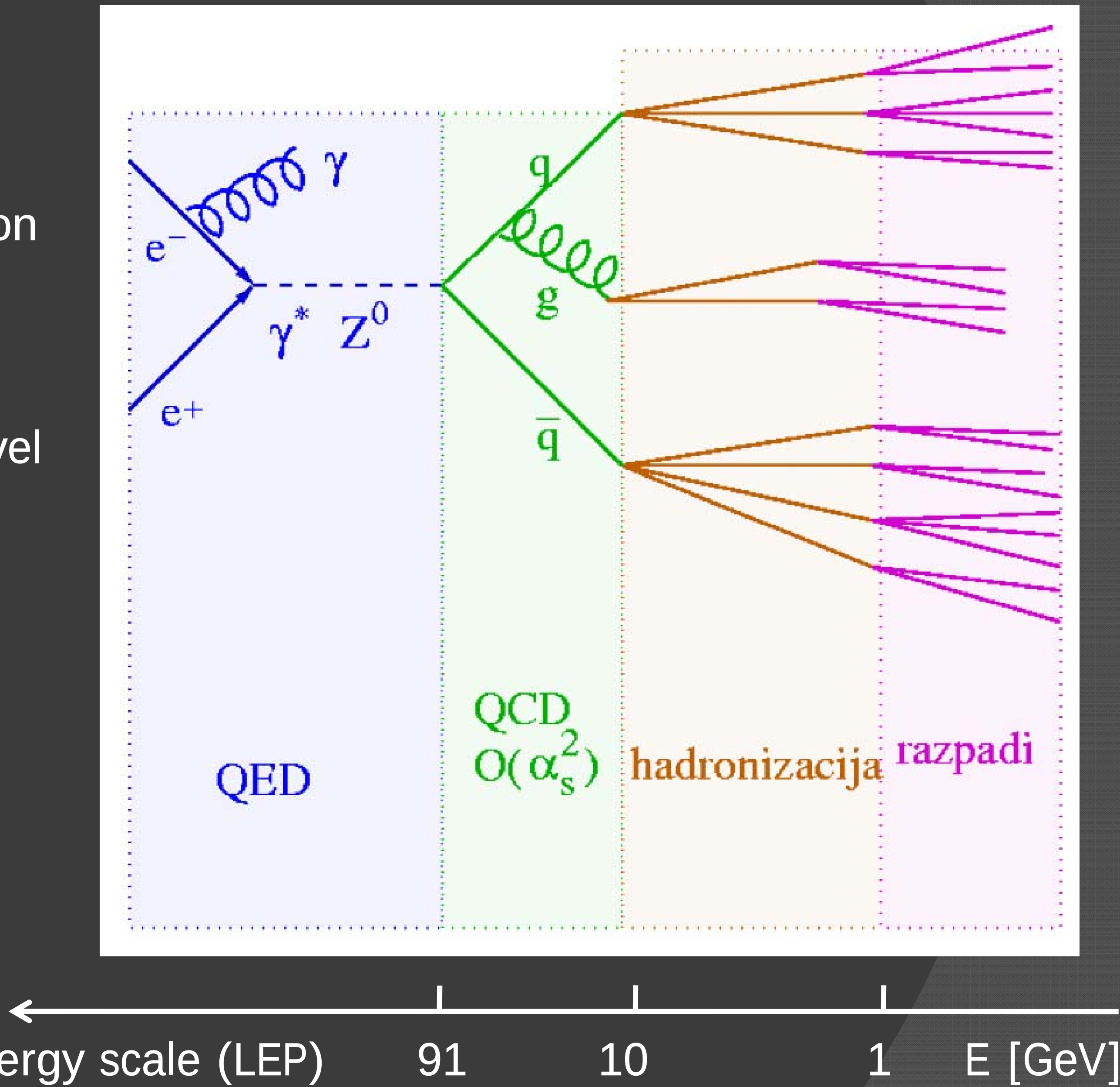
Analysis of data

Hadronic jets reconstruction

- quark production
  observable jets

  example of jet formation

  reconstructed jets ⇒
  interpretation of
  processes at parton level



Analysis of data

Hadronic jets reconstruction

jet reconstruction

experimental method
(assigning tracks to jets, calculation of energy, ….);
observables (# jets, # tracks in jet, angular distributions,…)
must be expressed in terms of theory parameters
in order to test predictions;

definition of jet

should be appropriate for exp. usage and theoretical calculations in order to confront theory & experiments

Algorithm for track association

resolution parameter

y_{cut};
if $y_{ij} < y_{cut}$ two tracks in same jet

combination of tracks

e.g. $\mathbf{p}_{jet} = \mathbf{p}_i + \mathbf{p}_j$;

Analysis of data

Hadronic jets reconstruction

jet reconstruction
algorithms

name	resolution parameter	combination	comment
JADE	$y_{\rm cut} = M_{ij}^2/E_{\rm mer}^2 = 2E_iE_j(1-\cos\theta_{ij})/E_{\rm mer}^2$	$p_k = p_i + p_j$	preserves E, **p**
p	$y_{\rm cut} = (p_i + p_j)^2/E_{\rm mer}^2$	$\vec{p}_k = \vec{p}_i + \vec{p}_j$ $E_k = \lvert\vec{p}_k\rvert$	preserves **p**
DURHAM	$y_{\rm cut} = 2\,\min(E_i^2, E_j^2)(1-\cos\theta_{ij})/E_{\rm mer}^2$	$p_k = p_i + p_j$	preserves E, **p**; resummable NLO log's

in Table: $p_i \rightarrow 4-momentum$
$\vec{p}_i \rightarrow 3-momentum$

higher order calculations in perturbative QCD performed for massless partons
⇒ resolution parameters calculated for massless partons;
summing two jets 4-momenta in general leads to a non-zero mass object (new jet);
several algorithms exist to avoid the problem

Analysis of data

Hadronic jets reconstruction

alghorithms comparison

for all alghorithms perturbative calculations exist to $\mathcal{O}(\alpha_s^2)$,
e.g. relative rate of n jets:

$$R_2 = 1 - A(y_{\mathrm{cut}})\frac{\alpha_s(\mu)}{2\pi} - [B(y_{\mathrm{cut}},\mu) + C(y_{\mathrm{cut}})]\left(\frac{\alpha_s(\mu)}{2\pi}\right)^2$$

$$R_3 = A(y_{\mathrm{cut}})\frac{\alpha_s(\mu)}{2\pi} + B(y_{\mathrm{cut}},\mu)\left(\frac{\alpha_s(\mu)}{2\pi}\right)^2$$

$$R_4 = C(y_{\mathrm{cut}})\left(\frac{\alpha_s(\mu)}{2\pi}\right)^2$$

using above predictions + models of hadronization
comparison of parton and hadron distributions

Analysis of data

Hadronic jets reconstruction

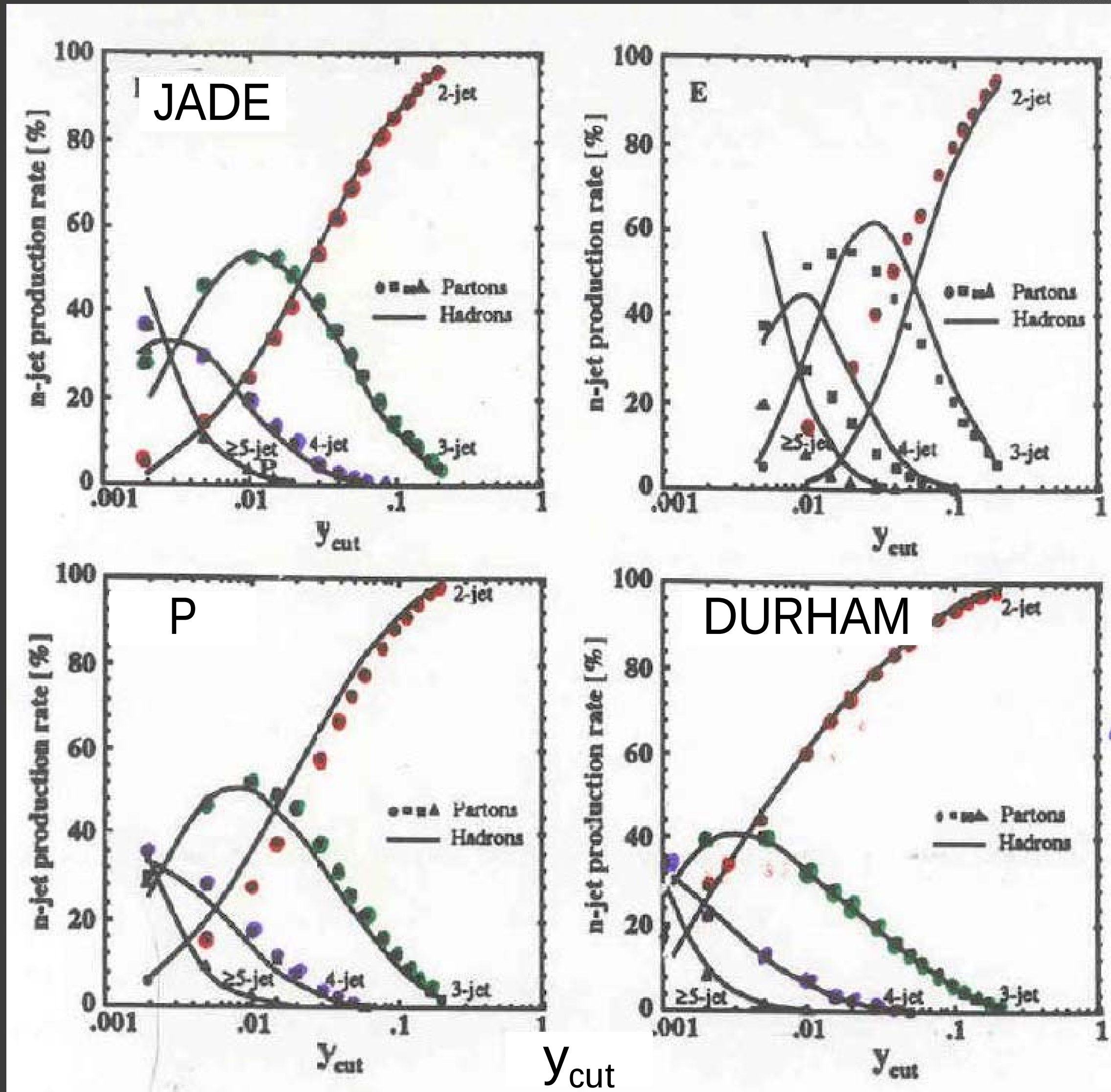
alghorithms comparison

n jet rates vs. y_{cut}

before hadronization (parton level)

- 2 jets
- 3 jets
- 4 jets

-- after hadronization



smallest hadronization corrections for JADE and DURHAM

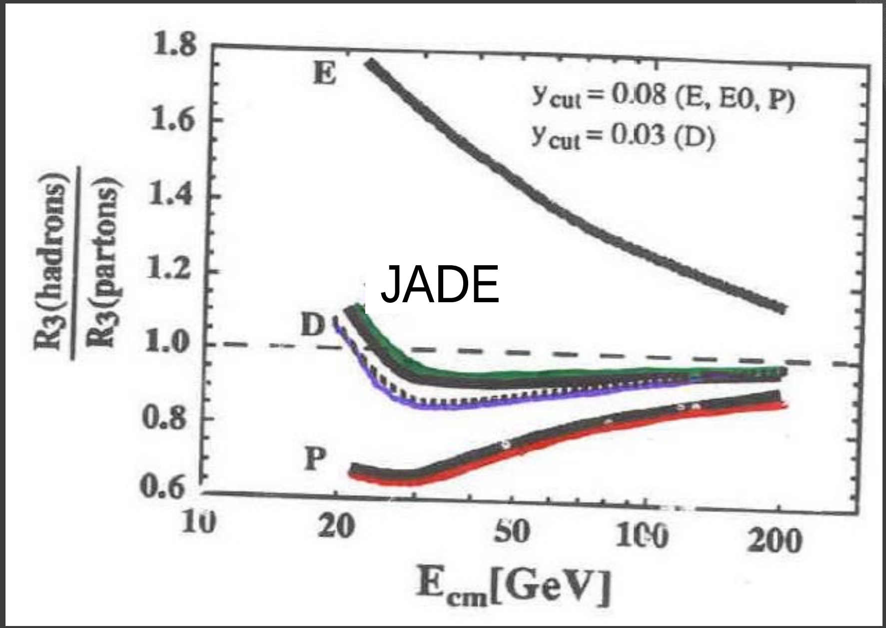
Analysis of data

Hadronic jets reconstruction

alghorithms comparison

n jet rates vs. E

smallest hadronization corrections for JADE



Analysis of data
Heavy quark tagging

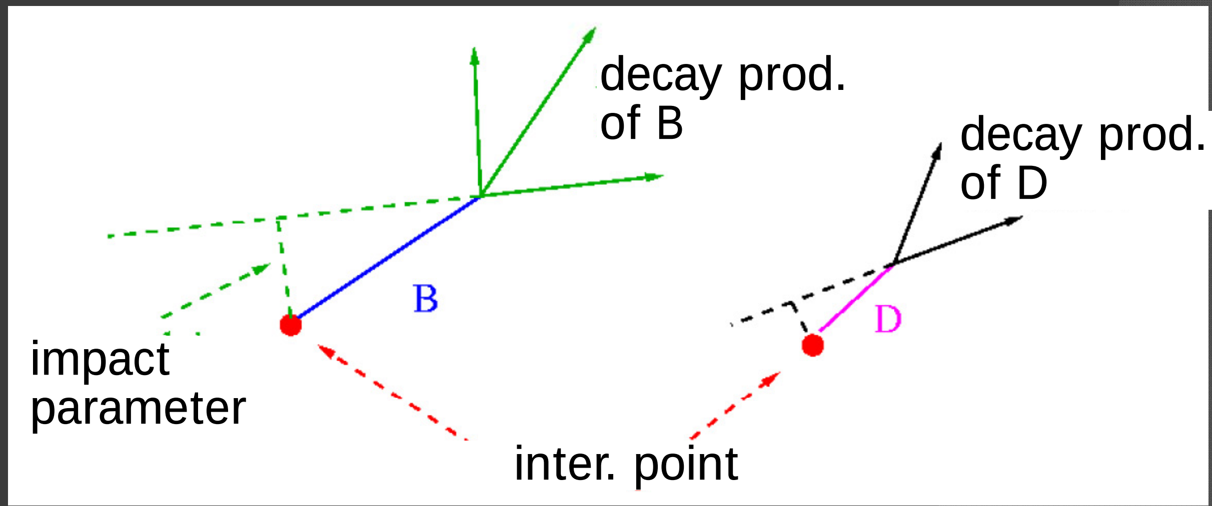
- Heavy (b) quark tagging

$H(m > 150 \text{ GeV}) \rightarrow b\bar{b} > 50\%$; CPV in B system; ....

try to discriminate b initiated jets from others;
use properties of hadrons composed of b quarks:

lifetime	mass	energy
1.6 ps	5.3 GeV	in fragmentation higher than for other hadrons

lifetime:
$\gamma c\tau_B > \gamma c\tau_D$ $\delta_B > \delta_D$



Analysis of data
Heavy quark tagging

- Heavy (b) quark tagging

  mass:
  example of
  rapidity for
  $B(M) \rightarrow X(E_1)Y(E_2)$;

$$\text{rapidity}: \quad y = \frac{1}{2}\ln\left[\frac{E+p_z}{E-p_z}\right]$$

$$M^2 \approx 2E_1E_2(1-\cos\theta)$$

$$E_1 \approx E_2 \approx E/2$$

$$\cos\theta \approx 1-\frac{2M^2}{E^2}$$

$$y \approx \frac{1}{2}\ln\left[\frac{1+\cos\theta/2}{1-\cos\theta/2}\right] \underbrace{\approx}_{M/E\ll 1} \frac{1}{2}\ln\left[\frac{4E^2}{M^2}-1\right]$$

large M $\Rightarrow$ small y;
average # of decay products higher for B than D
$\Rightarrow$ y even smaller

Analysis of data

Heavy quark tagging

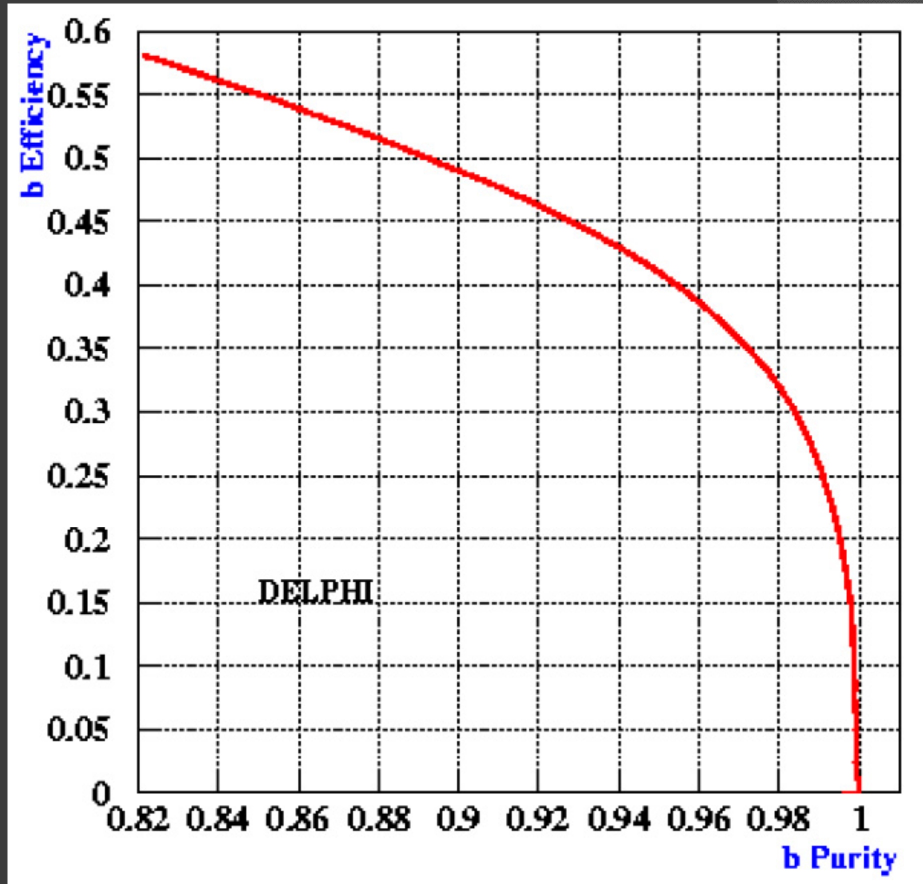
- Heavy (b) quark tagging

  combination of several discriminating variables into single one (likelihood ratio);

  example for Delphi at LEP;

  actual method for b-tagging depends on specific experimental conditions

$N_B(\text{selected})/N_B(\text{generated})$



$N_B(\text{selected})/N_{all}(\text{selected})$

Analysis of data
Summary

Path from electronic signal detection to result for measured physical quantities involves a number of steps

Each of those represents a specific problem and requires specific methods and solutions (some of those illustrated here)

Quality (correctness and accuracy) of the final results depends crucially on the quality of reconstruction of raw data