Bayesian Unfolding

Nadia Drenska

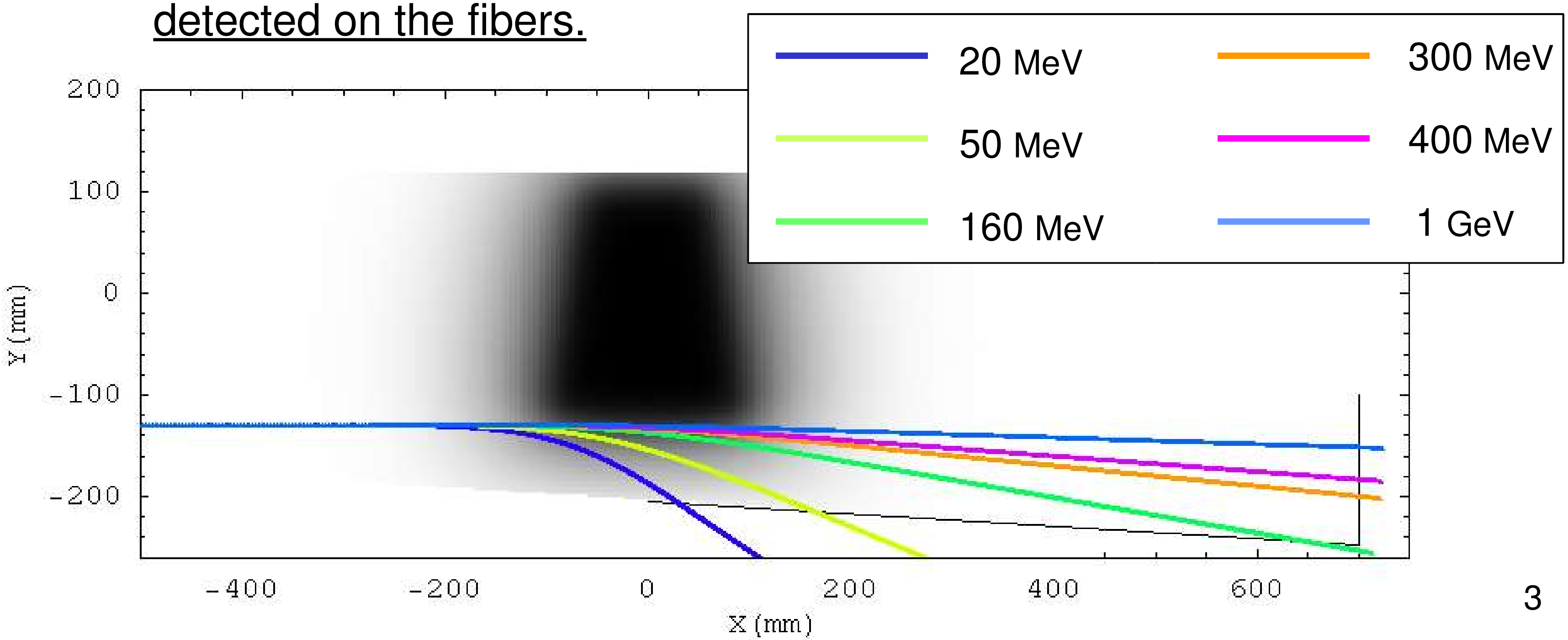
Based on:
“A multidimensional unfolding method based on Bayes' Theorem”, Giulio D'Agostini

Outline

- Problem definition
- Bayesian Unfolding: Theoretical background
- Application

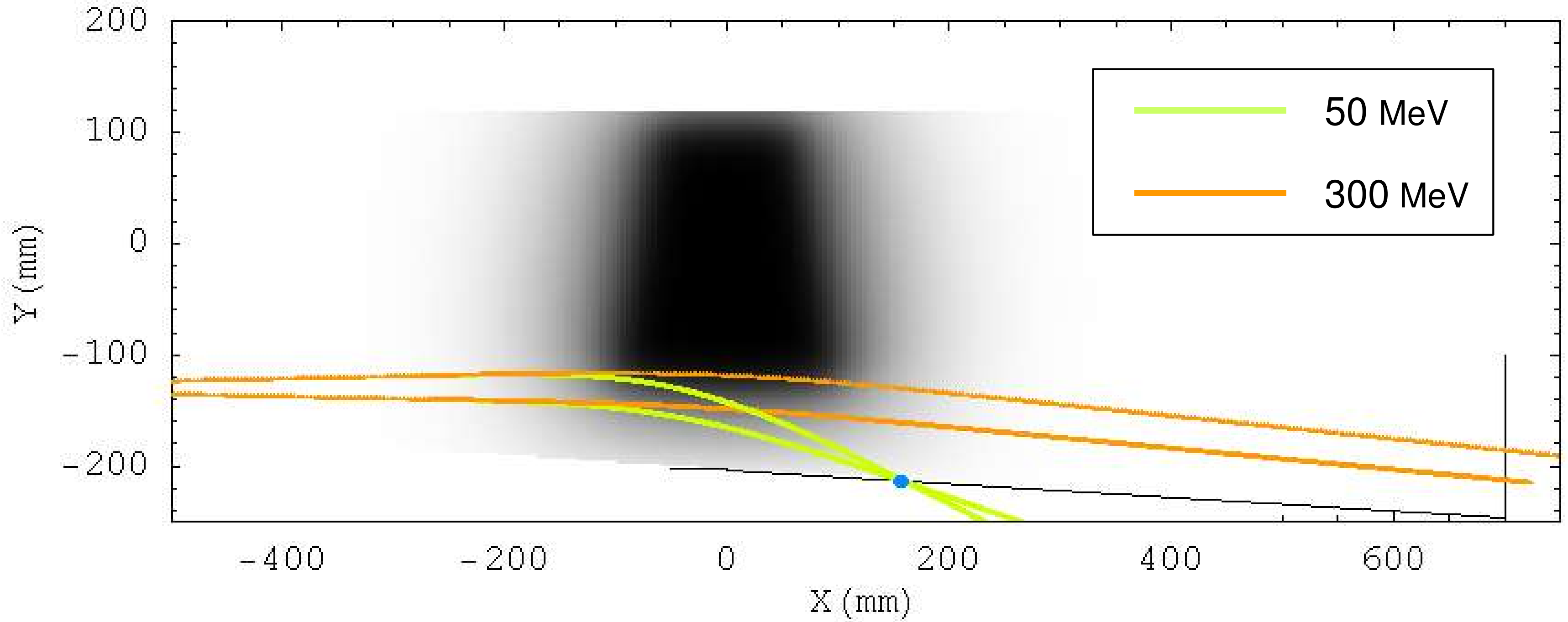
Problem definition

- Laser Plasma-Acceleration: electron beam production;
- Deviation of the beam using a magnet dipole;
- Scintillating fibers as electron position detector.
- <u>Estimation of the energy spectrum starting from the charge distribution detected on the fibers.</u>

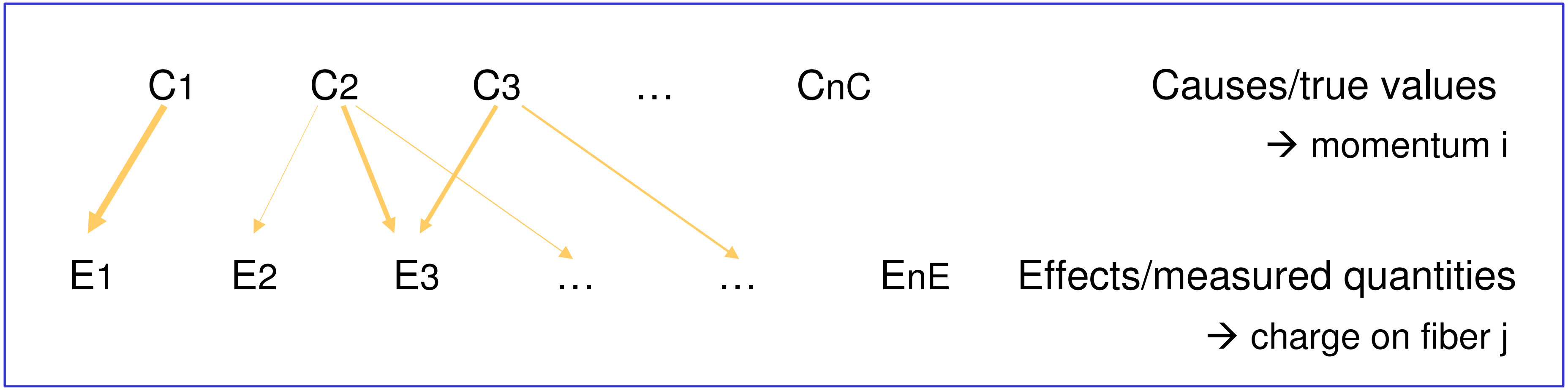


Angular divergence and focusing

- Pb: initial angular divergence $\Delta\theta$ = 2mrad
- Partial solution: focusing



Bayesian Unfolding (1)



Bayesian Unfolding (2)

Hp:

1- Initial probability P(Ci) ‘known’

2- Conditional probability P(Ej|Ci) known

Th:

$$P(C_i|E_j) = \frac{P(E_j|C_i)P(C_i)}{\sum_{l=1}^{n_C} P(E_j|C_l)P(C_l)}$$

P(Ci) → knowledge of P(Ci) increases with the number of observation

P(Ej|Ci) → calculated with MC

Bayesian Unfolding(3)

$$P(C_i|E_j)=\frac{P(E_j|C_i)P(C_i)}{\sum_{l=1}^{n_C} P(E_j|C_l)P(C_l)}$$

$A-\ P(E_j|C_i)\equiv S$ Smearing Matrix S

$B-\ \sum_{i=1}^{n_C} P(C_i)=1$

$C-\ \sum_{i=1}^{n_C} P(C_i|E_j)=1$

$D-\ \sum_{j=1}^{n_E} P(E_j|C_i)¿\epsilon_i¿1$ Efficiency

Bayesian Unfolding(4)

N_{obs} = number of experimental observations

$\underline{n}(E) = \{n(E_1), n(E_2), \ldots, n(E_{n_E})\}$ → distribution of frequencies

The expected number of events to be assigned to each of the causes and only due to the observed events:

$$n(C_i)|_{obs} = \sum_{j=1}^{n_E} n(E_j) P(C_i|E_j)$$

Best estimate of the true number of events

$$n(C_i) = \frac{1}{\epsilon_i} n(C_i)|_{obs}$$

Bayesian Unfolding (5)

From these unfolded events we can estimate:

the true total number of events $N_{true} = \sum_{i=1}^{n_C} n(C_i)$

the final probabilities of the causes $P(C_i) \equiv P(C_i | \underline{n}(E)) = \frac{n(C_i)}{N_{true}}$

overall efficiency $\epsilon = \frac{N_{obs}}{N_{true}}$

Loop

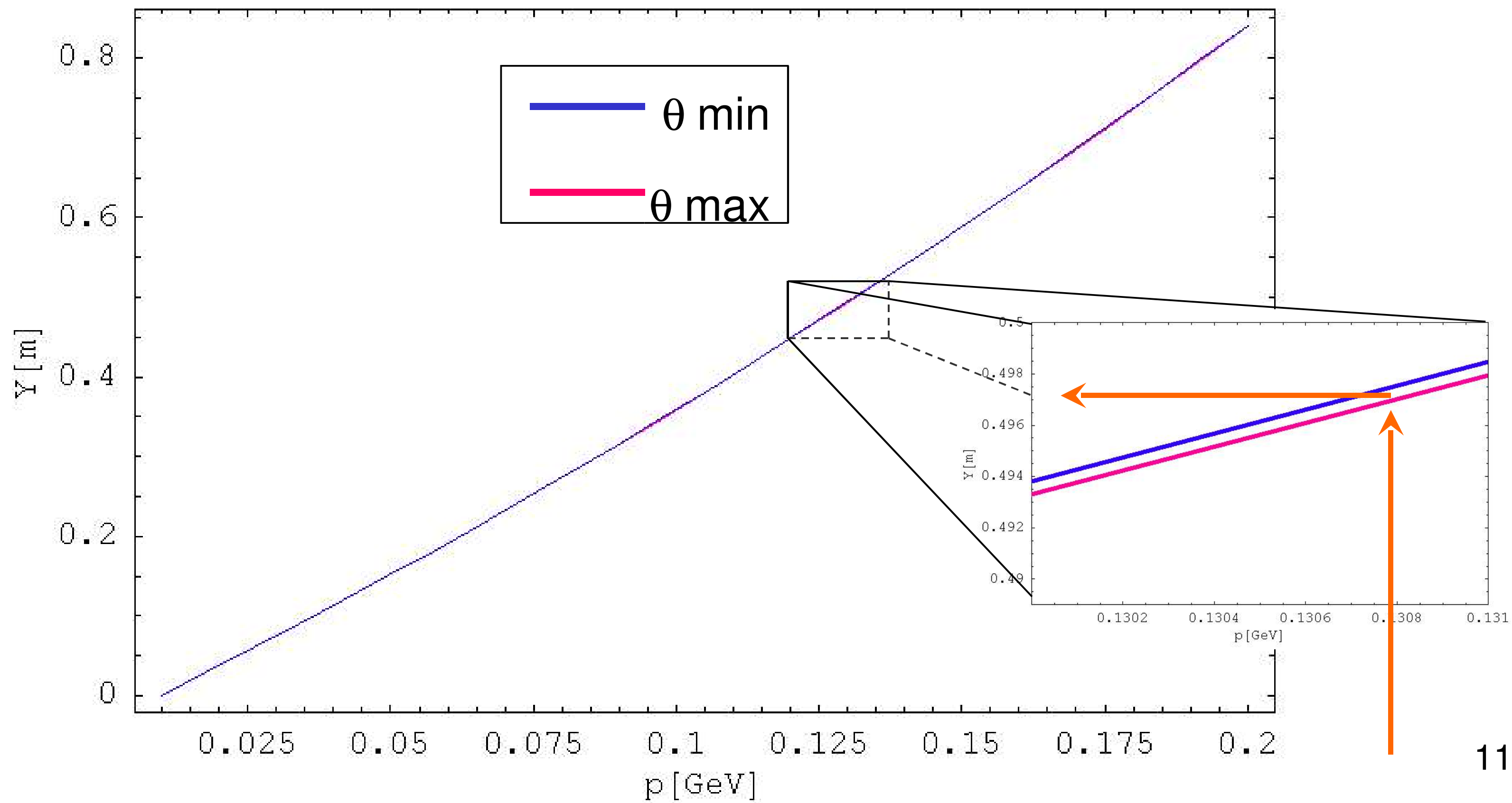
1-->chose the initial distribution of $P_0(C)$ from the best knowledge of the process under study (uniform distribution to start)
2-->calculate n(C) and P(C)
3--> make a χ^2 comparison between n(C) and $n_0(C)$
4-->replace $P_0(C)$ by P(C) and $n_0(C)$ by n(C) and start again; if after the second iteration the value of χ^2 is small enough stop the iteration otherwise go to step 2.

$$n(C_i) = \sum_{j=1}^{n_E} M_{ij}\, n(E_j) \quad where \quad M_{ij} = \frac{P(E_j|C_i)\, P_0(C_i)}{\left[\sum_{l=1}^{n_E} P(E_l|C_i)\right]\left[\sum_{l=1}^{n_C} P(E_j|C_l)\, P_0(C_l)\right]}$$

Calibration - low

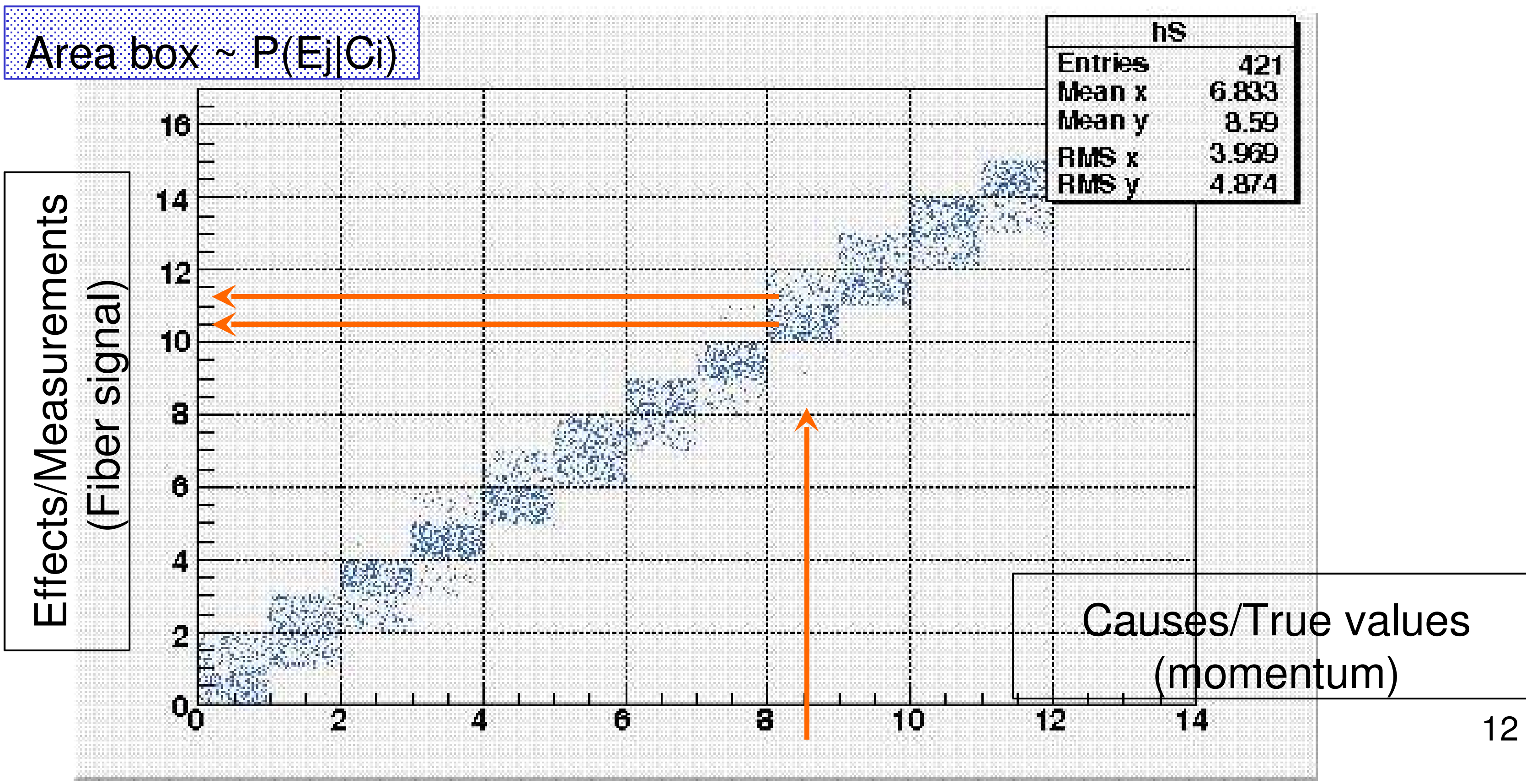
Direct problem

Momentum pi —— numerical integration eq. motion ——> Position yi

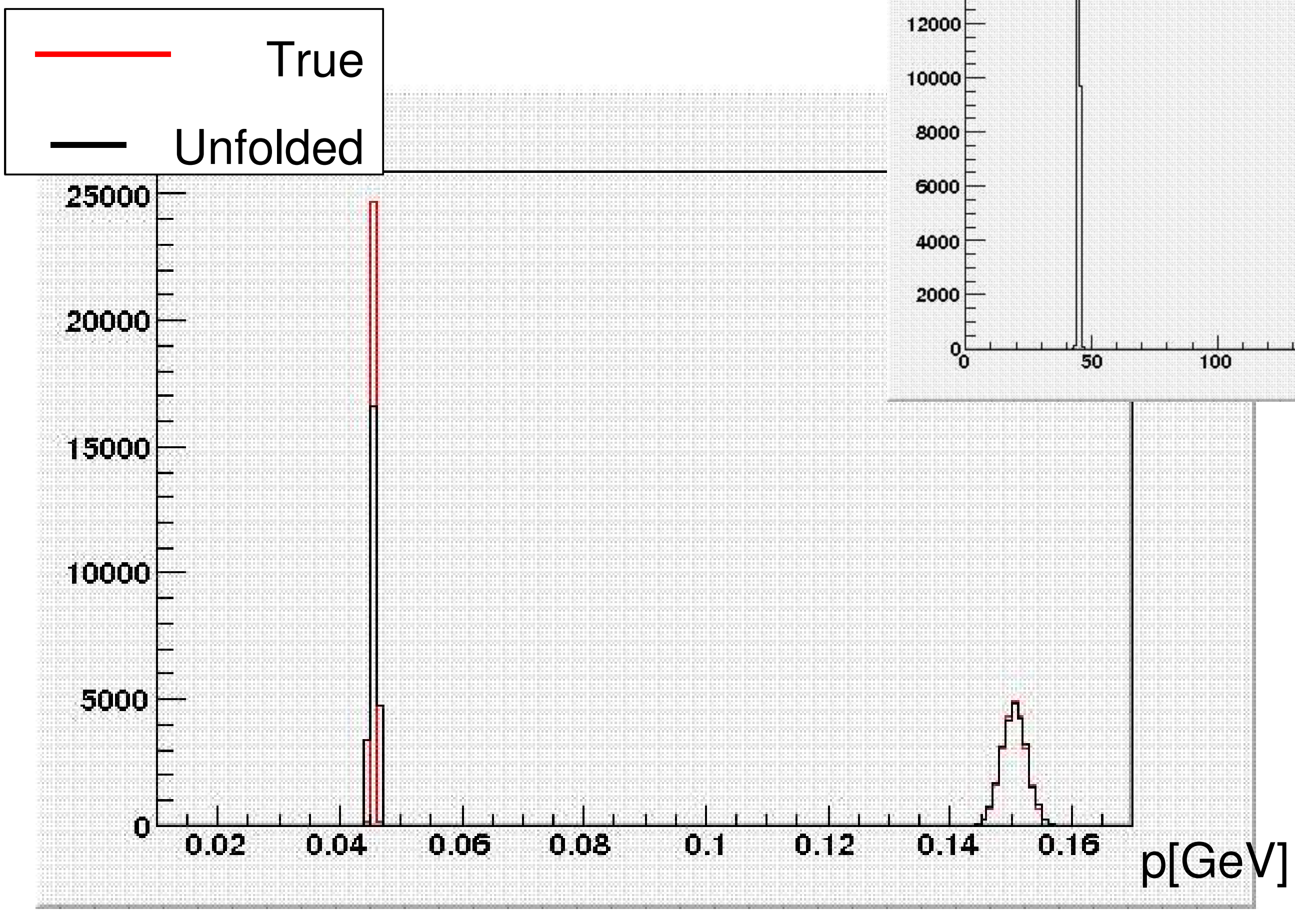


Smearing Matrix- low

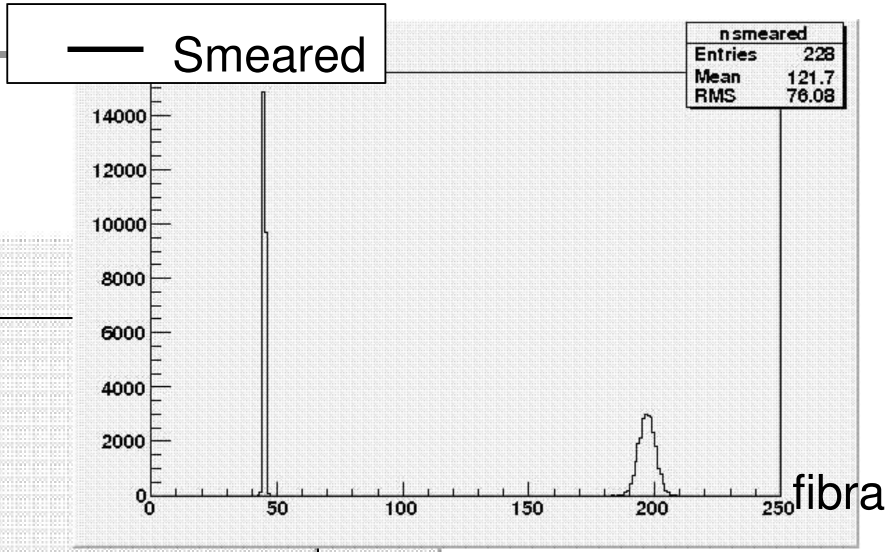
- Random: pi uniform, θ gauss
- Interpolation using the calibration curve



Results- low



True
Unfolded
25000
20000
15000
10000
5000
0
0.02
0.04
0.06
0.08
0.1
0.12
0.14
0.16
p[GeV]

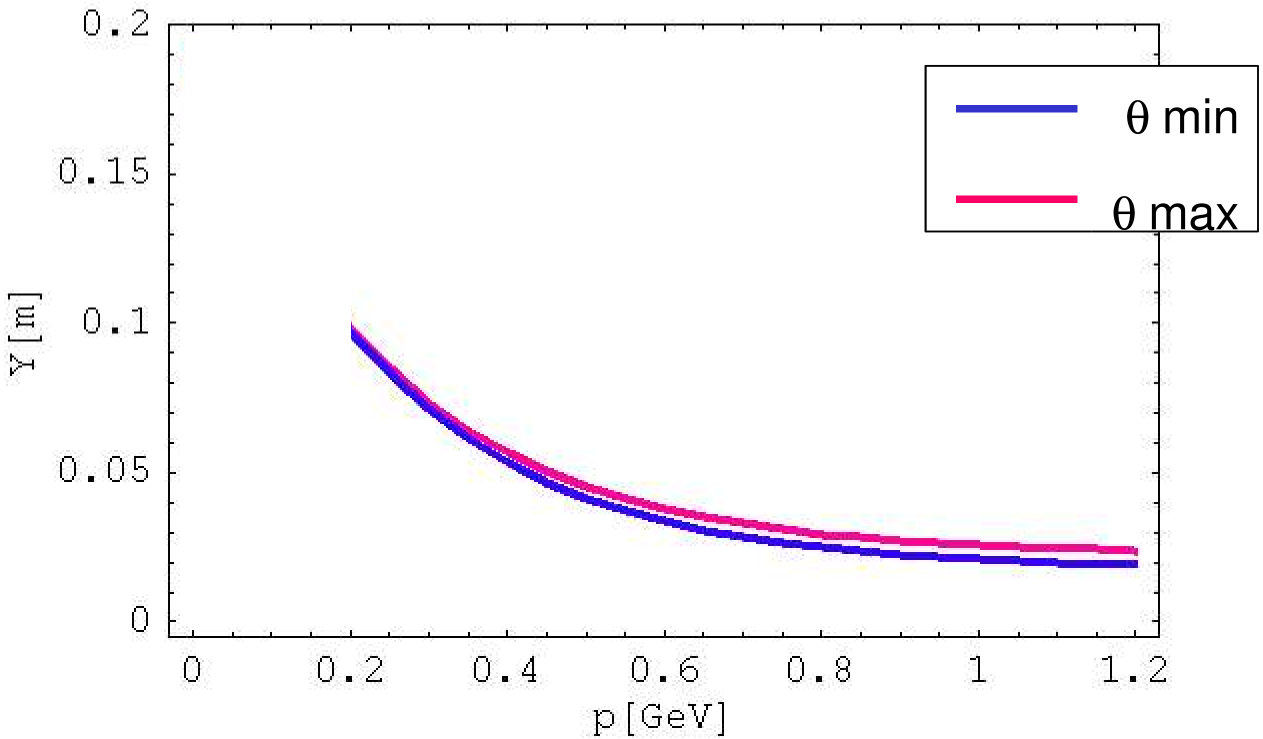


Smeared
nsmeared
Entries 228
Mean 121.7
RMS 76.08
14000
12000
10000
8000
6000
4000
2000
0
0
50
100
150
200
250
fibra

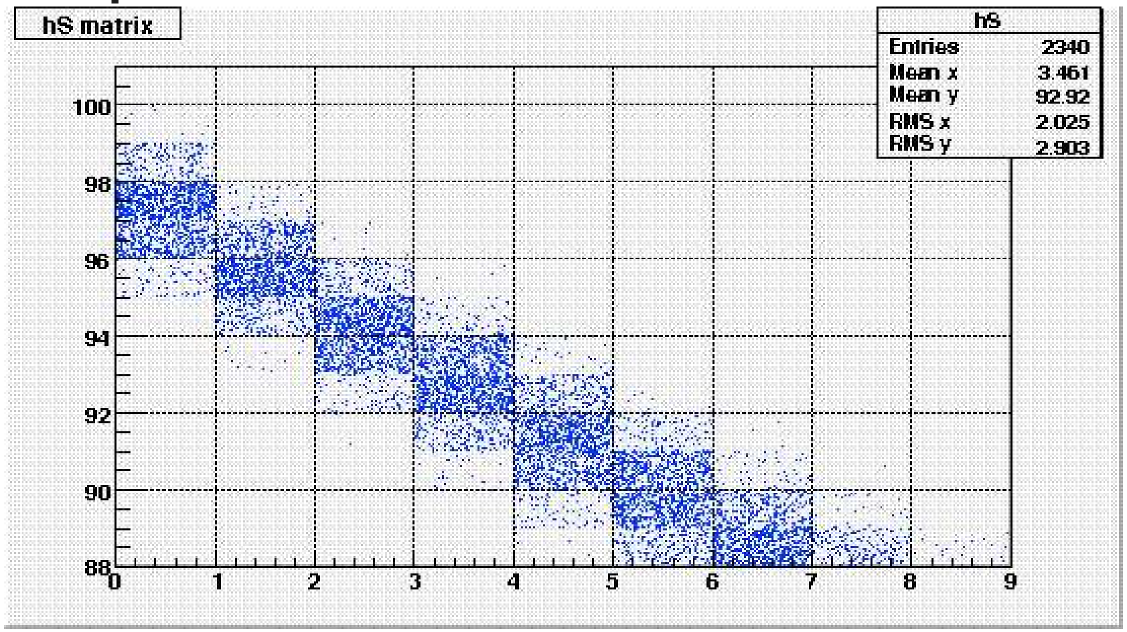
Calibration - high

Direct problem

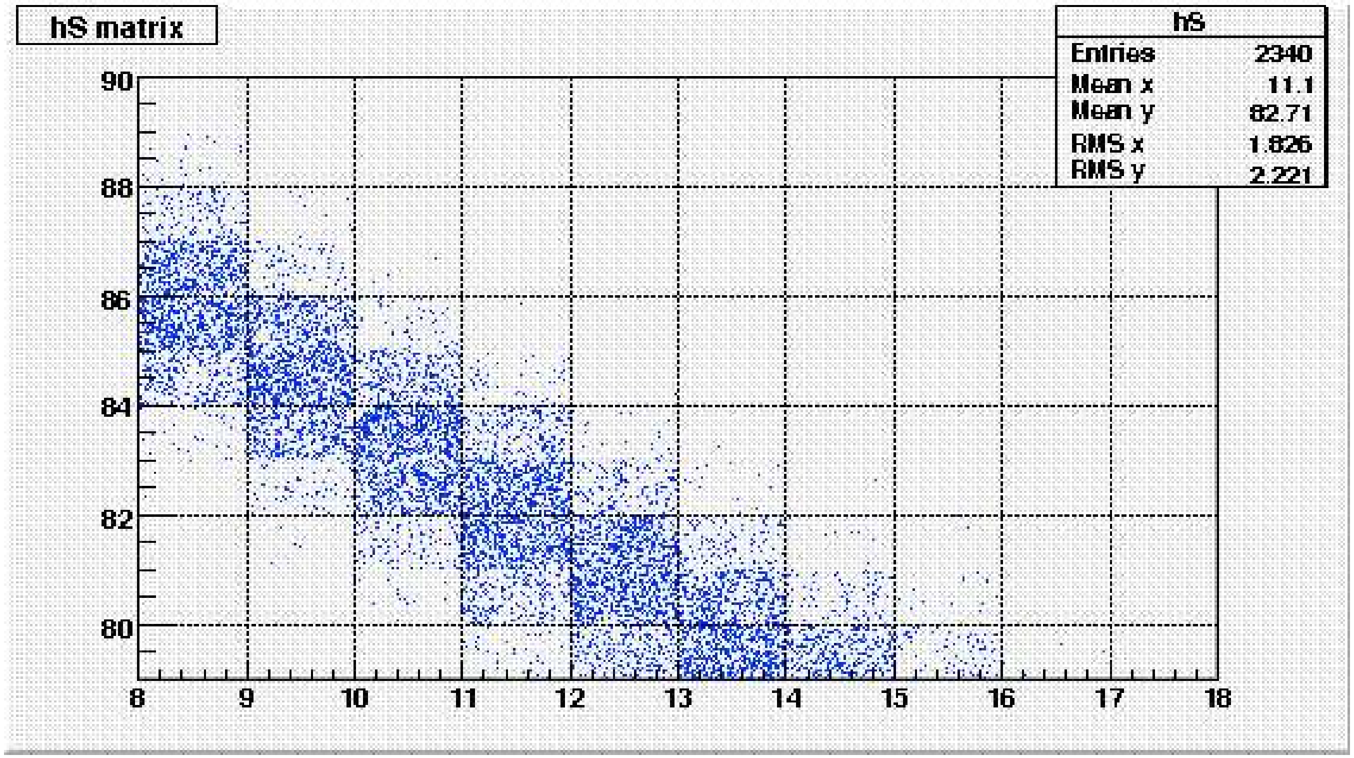
Momentum pi —numerical integration→ Position yi



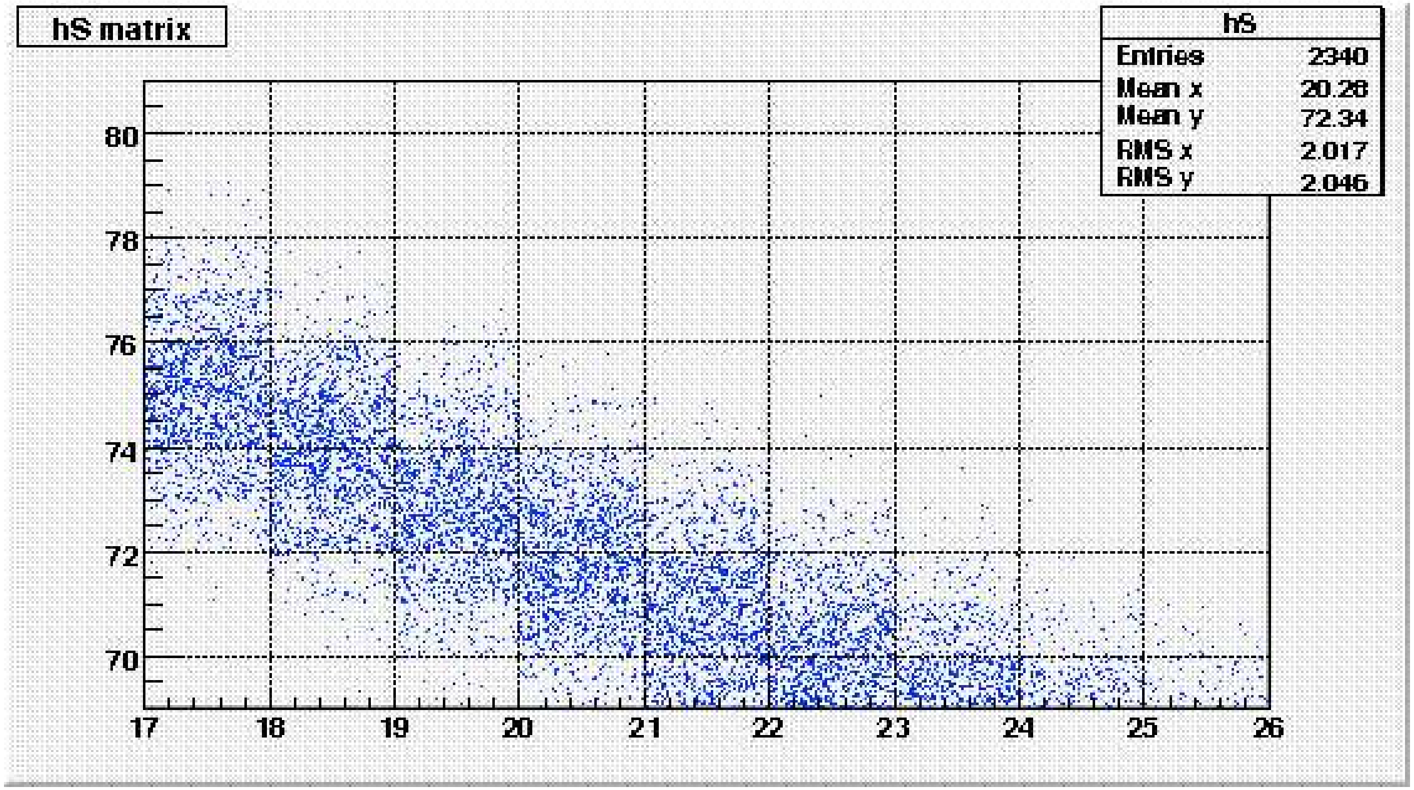
Smearing Matrix- high



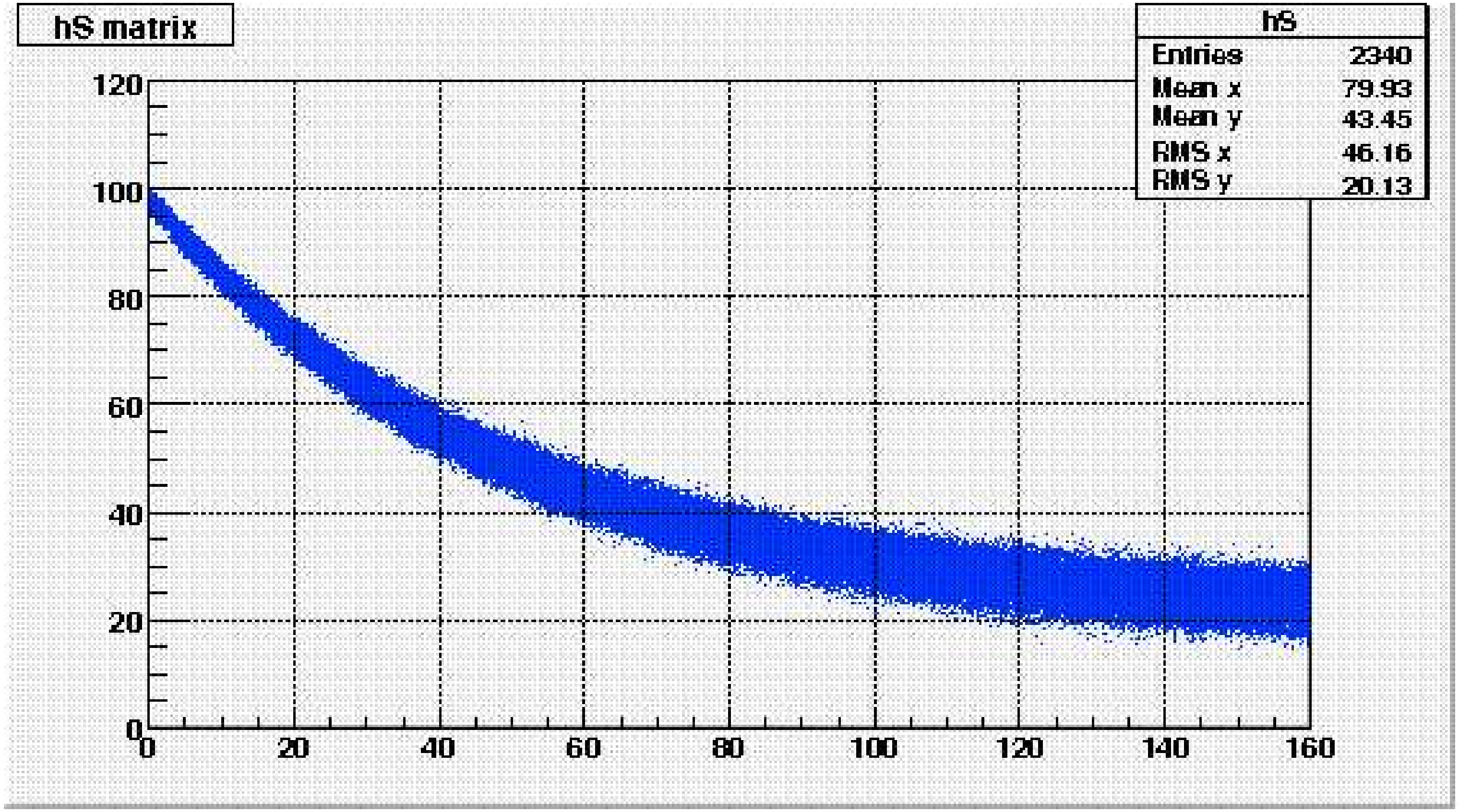
Smearing Matrix - high



Smearing Matrix - high



Complite Smearing Matrix - high



Results - high