Find a signal

$$
\begin{aligned}
\mathbf{N_{obs}(p_T, \phi, \eta, z)} &= B + \sigma A \epsilon \int \mathcal{L} dt \\
&= B(p_T, \phi, \eta, z, \mathcal{L}) + \\
&\int \sigma(p_T^0, \phi^0, \eta^0) \times \mathcal{L}(t, z) \times A(p_T^0, \phi^0, \eta^0, z^0) \times \\
&\qquad R(p_T^0, \phi^0, \eta^0, z^0; p_T, \phi, \eta, z, \mathcal{L}) \times \\
&\qquad \epsilon(p_T, \phi, \eta, z, \mathcal{L}) dp_T^0 d\phi^0 d\eta^0 dt
\end{aligned}
$$

$\sigma(X_\alpha ...)$ is the true cross section as a function of true variables $X_a ...$.

$\mathcal{L}(t, z)$ is the luminosity as a function of time.

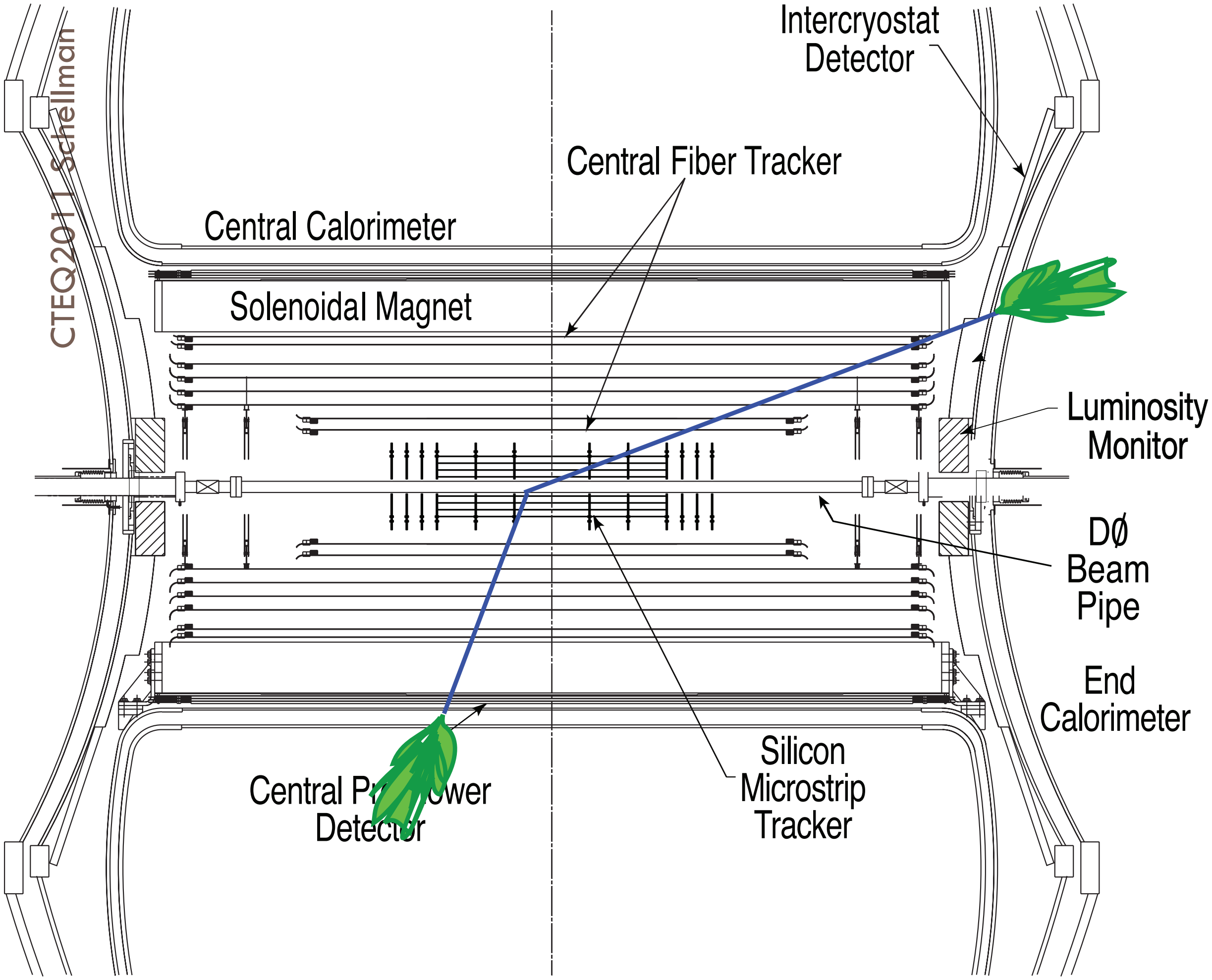
$A(X_\alpha ..)$ is the geometrical Acceptance as a function of true variables.

$R(X_\alpha ..., \mathcal{L}; X_a ...)$ is the Resolution function which smears true $X_\alpha ...$ to detected $X_a ...$

$\epsilon(X_a ..., \mathcal{L})$ is the probability that a particle is actually detected by a physical detector.

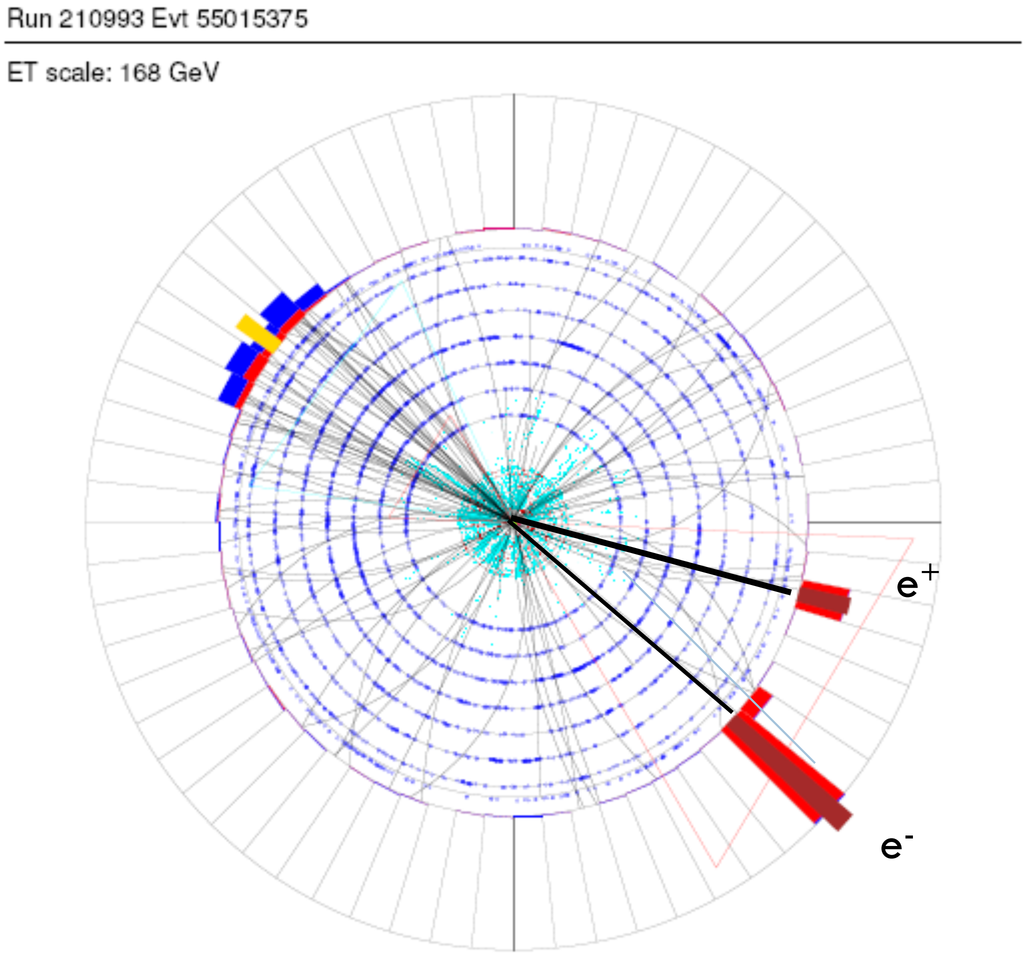
$B(X_a ..., \mathcal{L})$ is the background

Look for Z’s In our detector



- Expect 2 “electrons”
- PT ~ MZ/2
- Calorimeter energy
  - Shower looks like an electron
- Track momentum
  - Tracks come from the same point

Z→ee



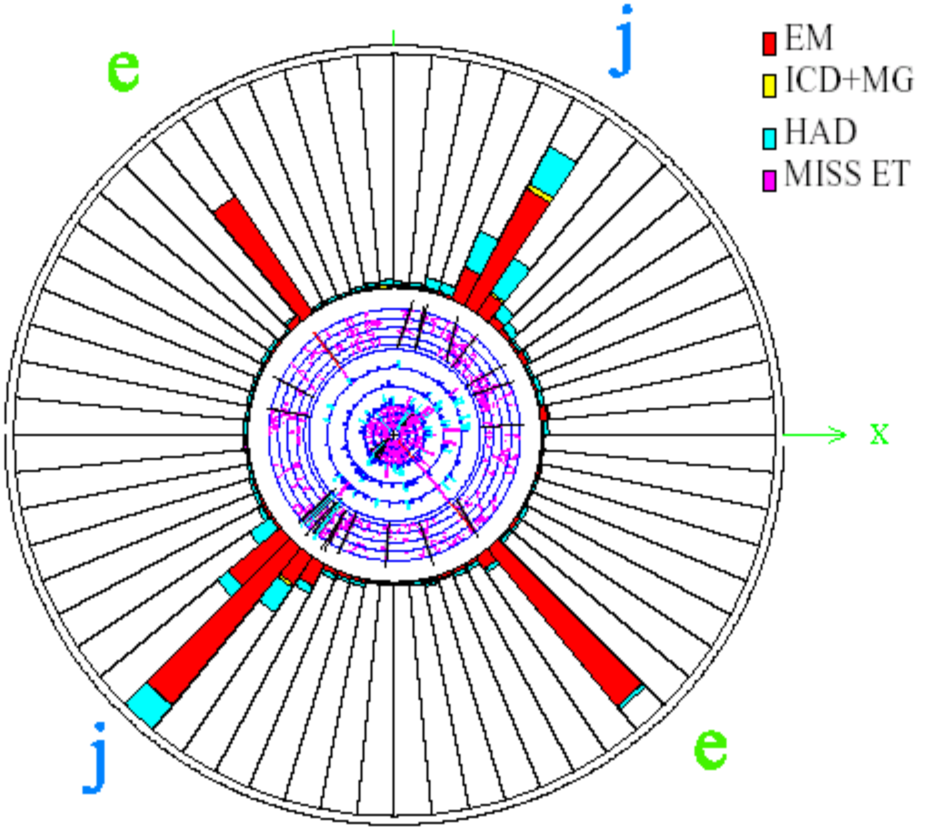
This Z0 has high transverse momentum!

A Z^0 -> e^+e^- + jjevent

M_{ee} ~ 91 GeV/c^2

Electrons have a track and EM energy

Jets have many tracks and EM+Hadron energy



End view

What do we expect

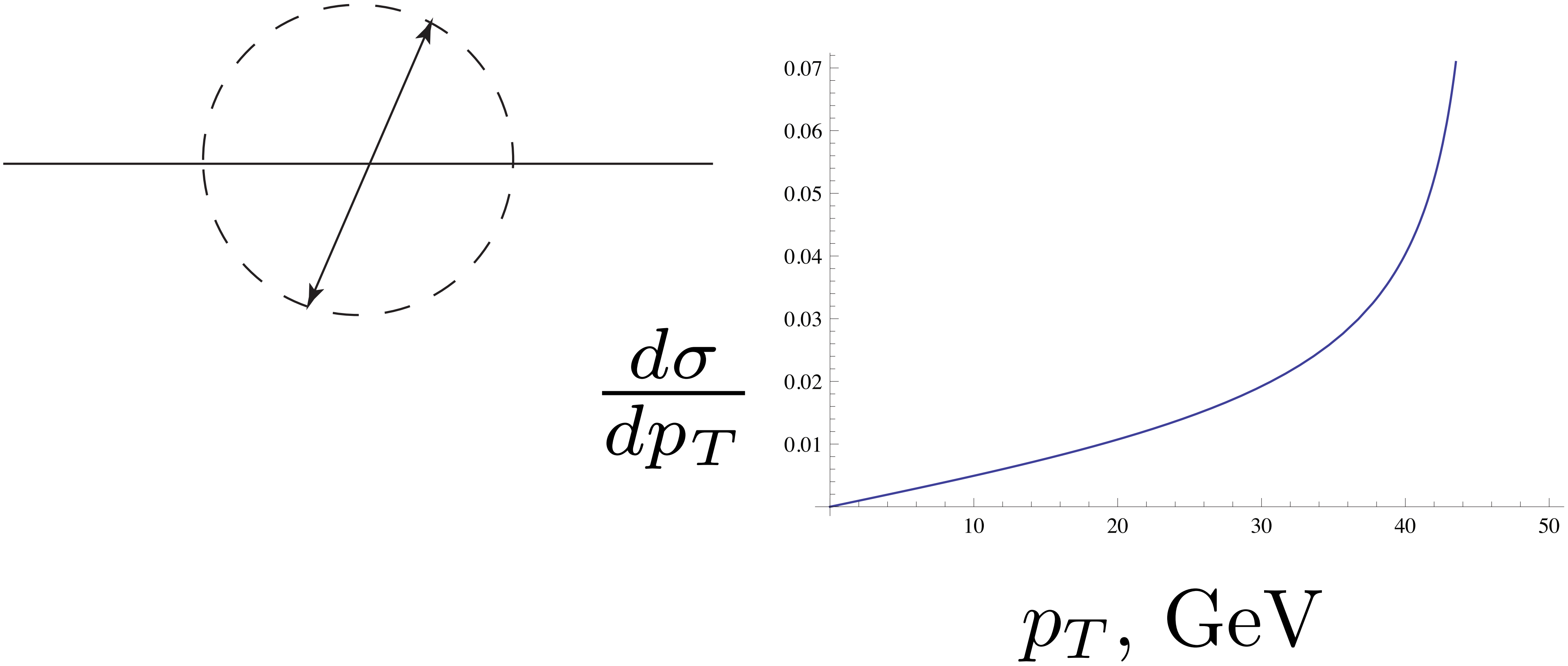
$$\frac{d\sigma(f\bar{f} \to e^- e^+)}{d\cos\theta^*} = (1/128\pi\hat{s})[(|A_{LL}|^2 + |A_{RR}|^2)(1+\cos\theta^*)^2$$

$$+ (|A_{LR}|^2 + |A_{RL}|^2)(1-\cos\theta^*)^2] \quad .$$

$$\frac{d\sigma(gg \to gg)}{d\cos\theta^*} \propto \frac{1}{\sin^4 \frac{\theta}{2}^*}$$

The leptons are more isotropic than the backgrounds

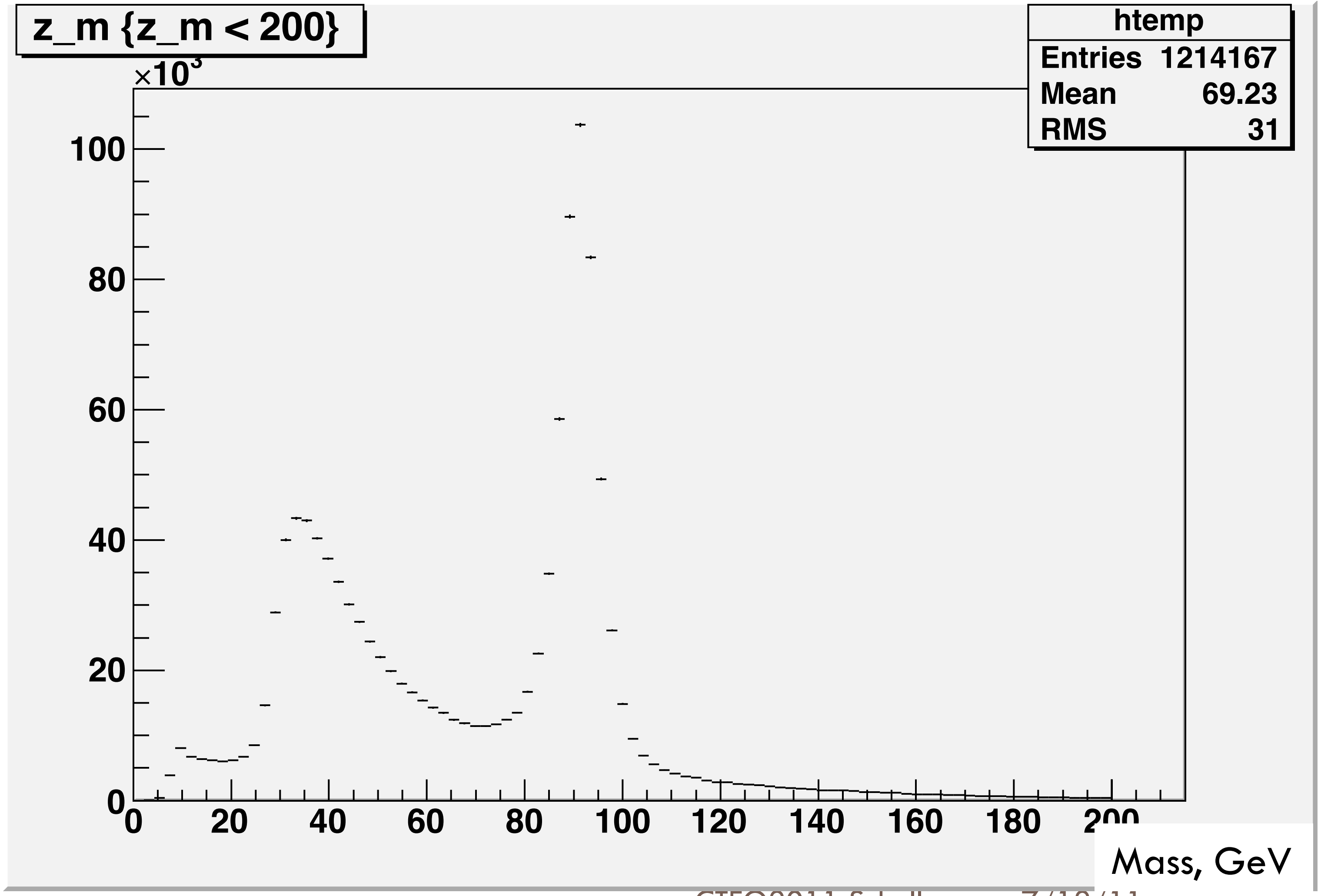
Jacobian peak



Now let's do some data analysis

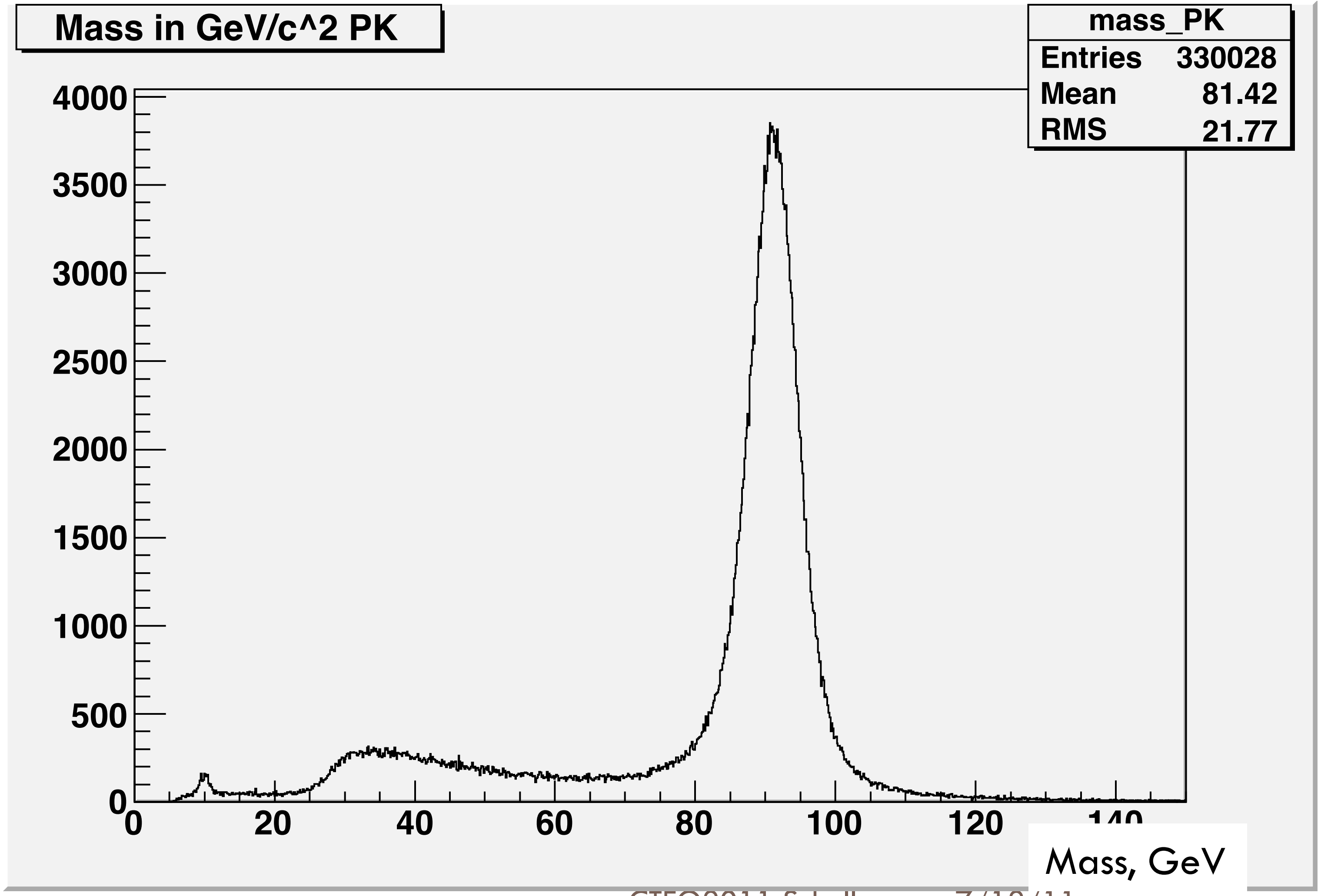
- I happen to have a di-electron root file lying around.

Invariant mass of raw electron pairs

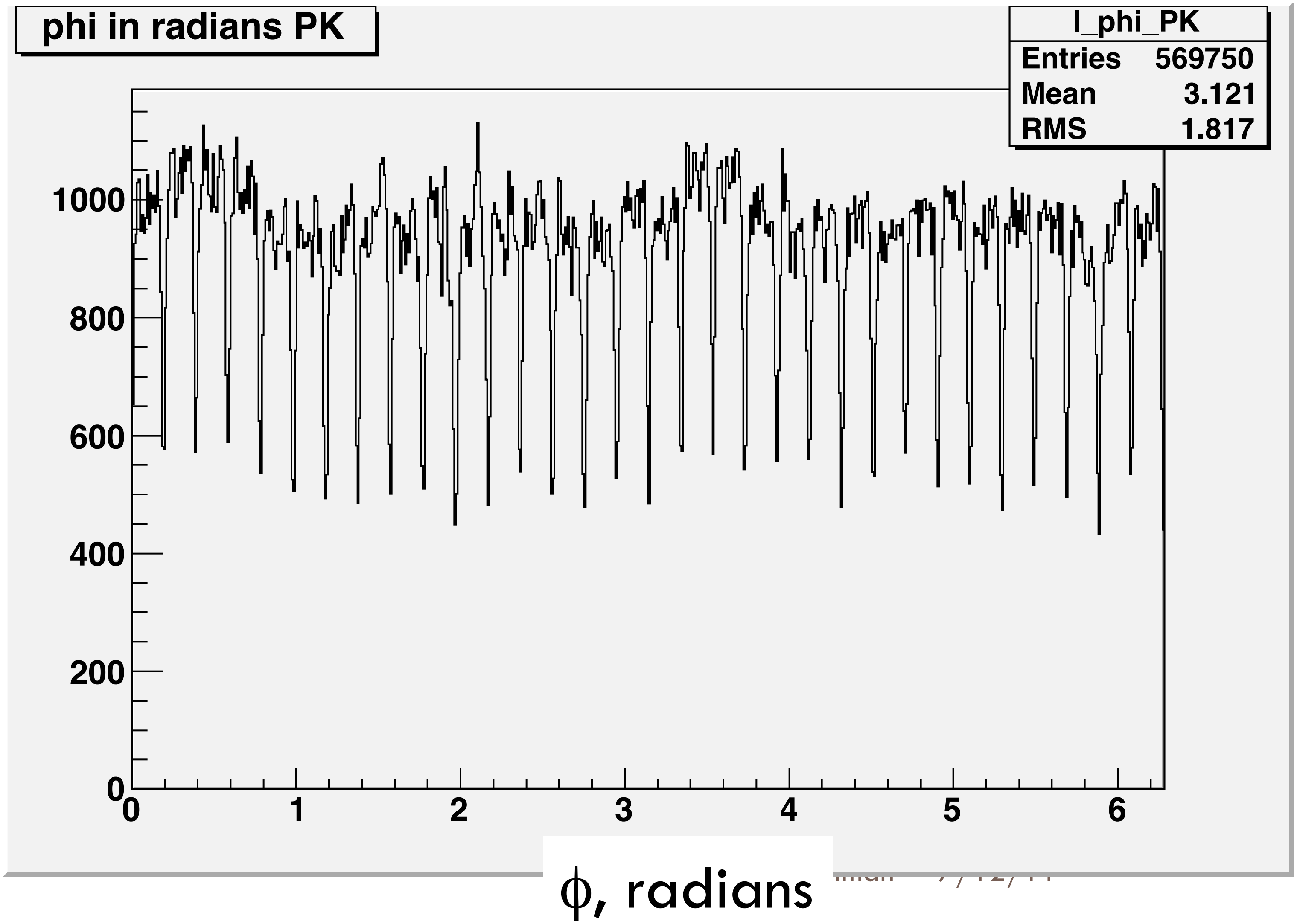


Mass, GeV

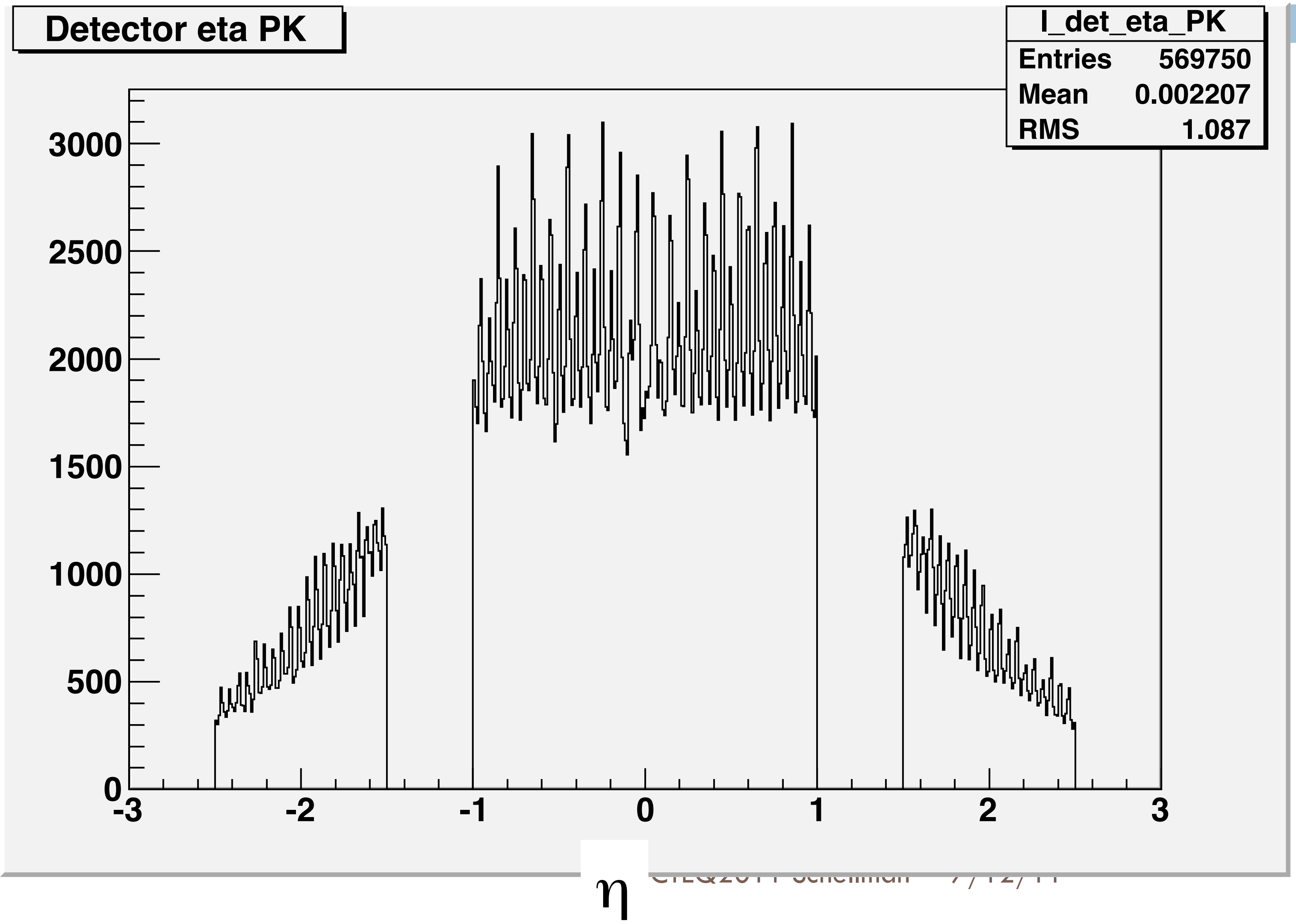
After quality cuts



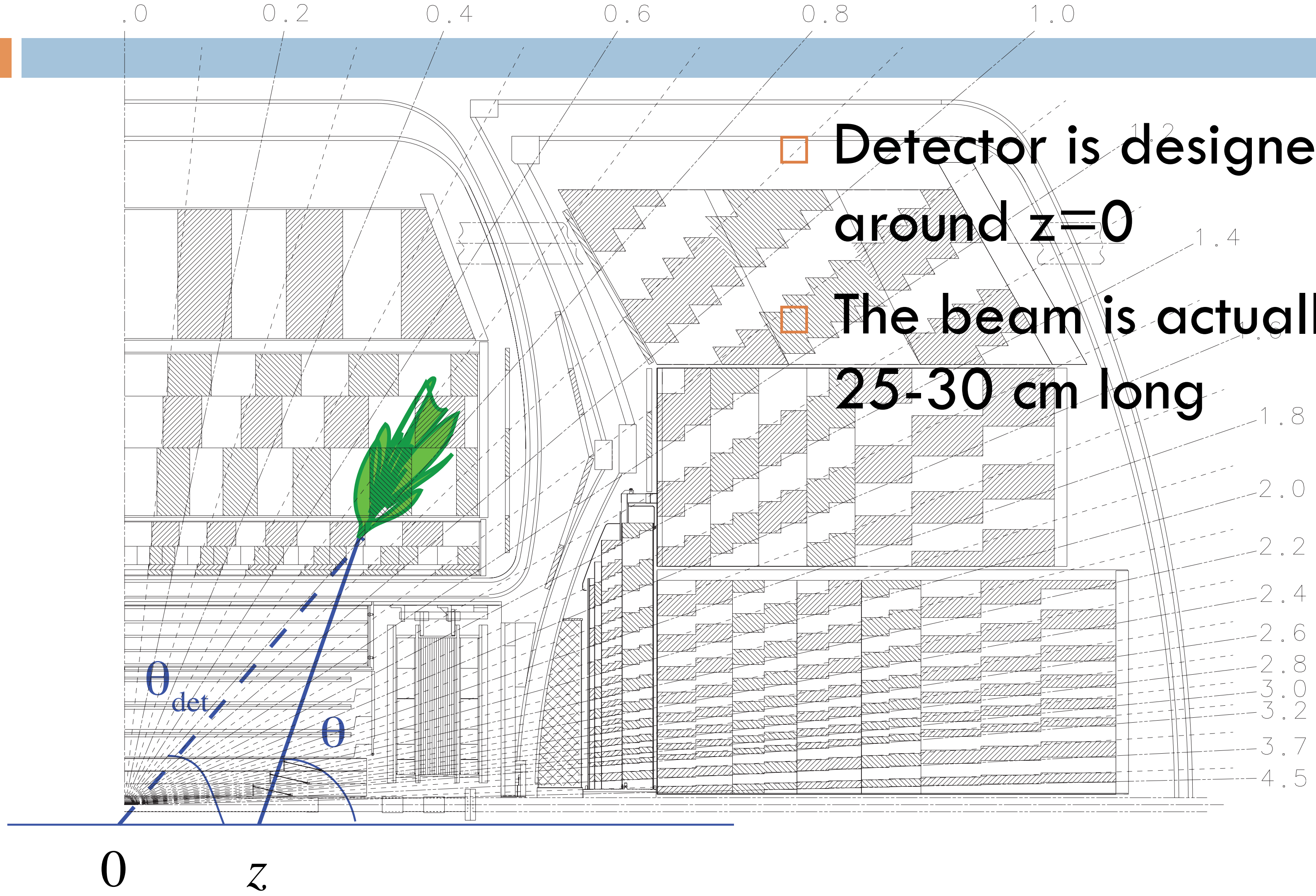
Phi


phi in radians PK
l_phi_PK
Entries 569750
Mean 3.121
RMS 1.817
1000
800
600
400
200
0
0
1
2
3
4
5
6
φ, radians

Detector η

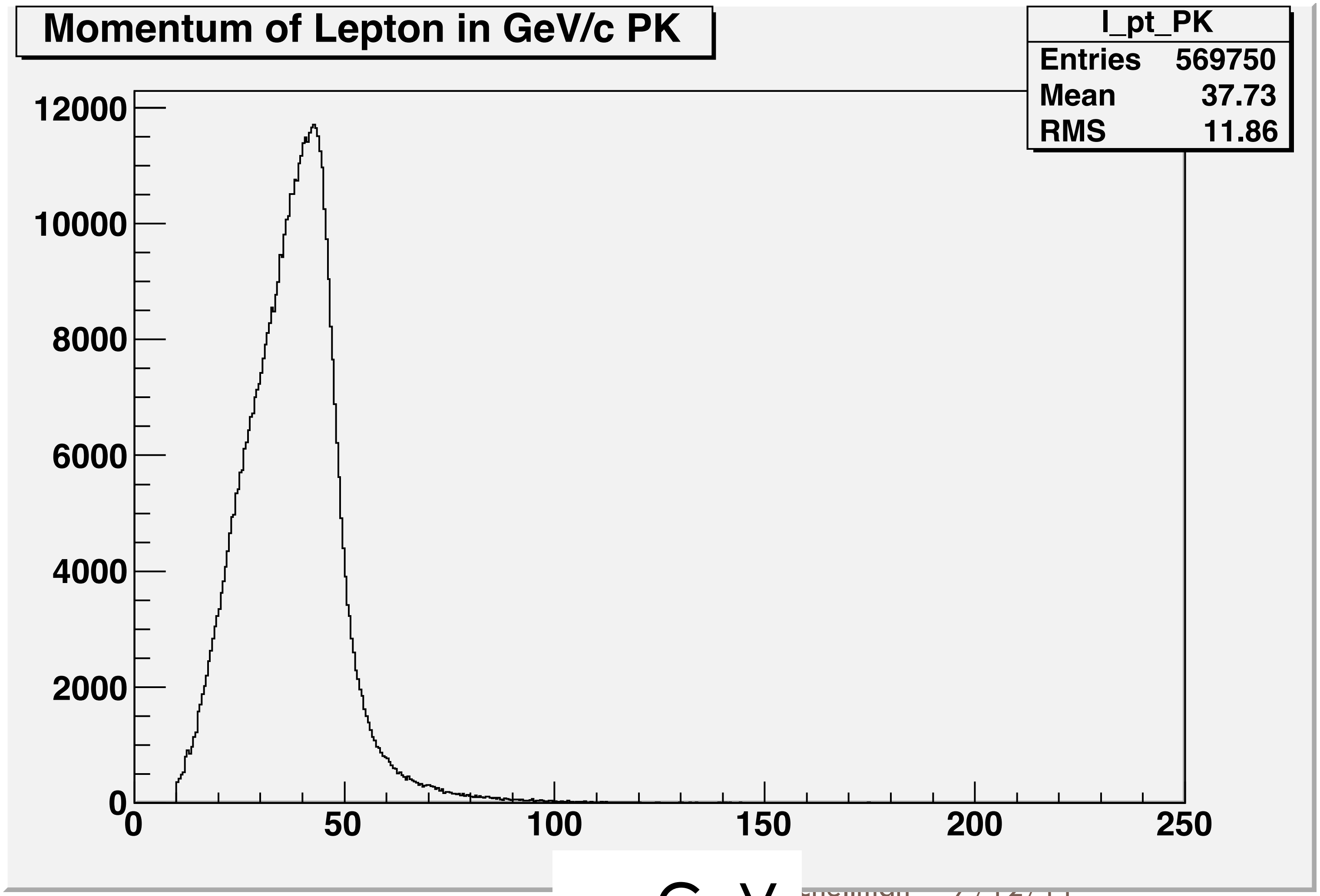

Detector eta PK
l_det_eta_PK
Entries 569750
Mean 0.002207
RMS 1.087
3000
2500
2000
1500
1000
500
0
-3
-2
-1
0
1
2
3
η

Detector and physics coordinates



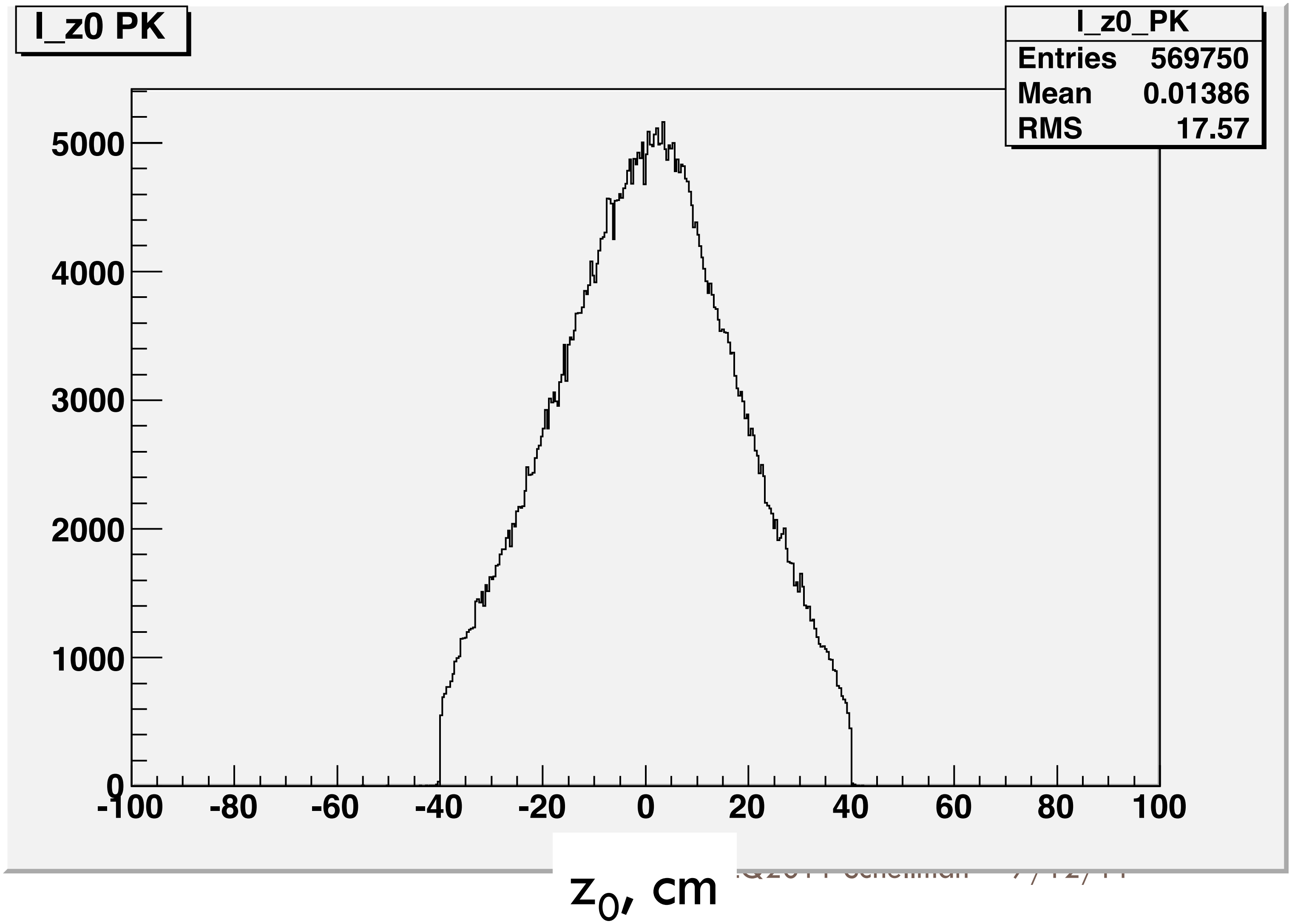
- Detector is designed around z=0
- The beam is actually 25-30 cm long

PT of leptons

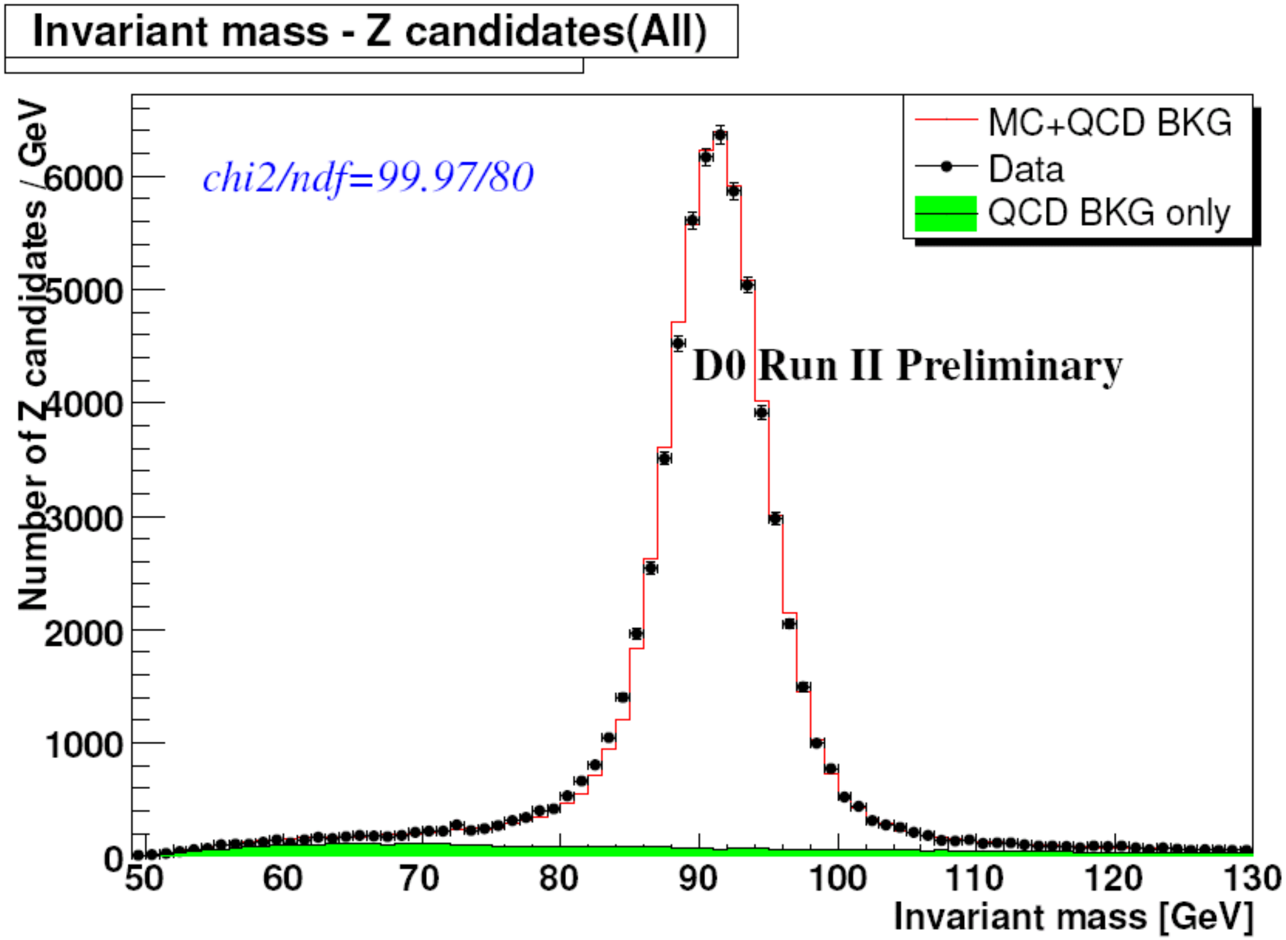


p_T, GeV

Z position of vertex in cm



The signal


Invariant mass - Z candidates(All)
MC+QCD BKG
Data
QCD BKG only
chi2/ndf=99.97/80
D0 Run II Preliminary
Number of Z candidates / GeV
0
1000
2000
3000
4000
5000
6000
50
60
70
80
90
100
110
120
130
Invariant mass [GeV]

Luminosity

$$\begin{aligned} N_{obs}(p_T, \phi, \eta, z) &= B + \sigma A \epsilon \int \mathcal{L} dt \\ &= B(p_T, \phi, \eta, z, \mathcal{L}) + \\ &\int \sigma(p_T^0, \phi^0, \eta^0) \times \mathcal{L}(t, z) \times A(p_T^0, \phi^0, \eta^0, z^0) \times \\ &\quad R(p_T^0, \phi^0, \eta^0, z^0; p_T, \phi, \eta, z, \mathcal{L}) \times \\ &\quad \epsilon(p_T, \phi, \eta, z, \mathcal{L}) dp_T^0 d\phi^0 d\eta^0 dt \end{aligned}$$

$\sigma(X_\alpha ...)$ is the true cross section as a function of true variables $X_a ...$.

$\mathcal{L}(t, z)$ is the luminosity as a function of time.

$A(X_\alpha ..)$ is the geometrical Acceptance as a function of true variables.

$R(X_\alpha ..., \mathcal{L}; X_a ...)$ is the Resolution function which smears true $X_\alpha ...$ to detected $X_a ...$

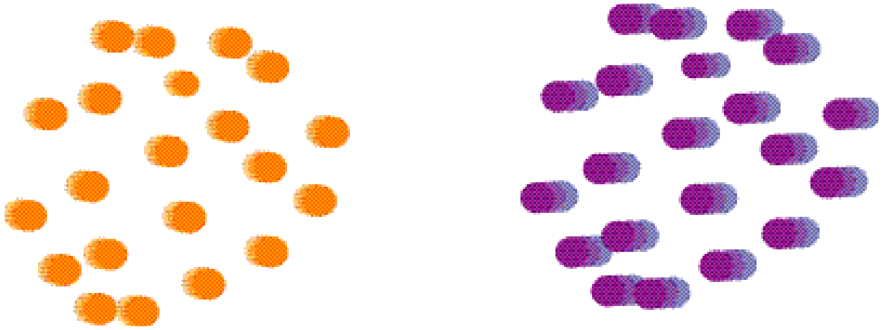
$\epsilon(X_a ..., \mathcal{L})$ is the probability that a particle is actually detected by a physical detector.

$B(X_a ..., \mathcal{L})$ is the background

Luminosity

Luminosity is a measure of how often protons/antiprotons get close enough to interact

$$L = f\,\frac{n_1 n_2}{4\pi\, s_x s_y}$$



f= beam crossing frequency 396 nsec

n= protons/bunch 10^{11}

s = transverse beam size = 0.0001 m

L ~ 10^{32} crossings/cm^2/sec

Typical Cross Sections

Total proton/antiproton cross section is 7×10^{-30} m^2

Unit of Barns (b) = 10^{-28} m^2

$$\sigma(p\bar{p} \Rightarrow X) \approx 70\, mb$$

Run II L ~ 10^{32} crossings/cm^2/sec

N/sec ~ σL = 7×10^{6}/sec

> 3 interactions per beam crossing!

Cross Section for top production:

$$\sigma(p\bar{p} \Rightarrow t\bar{t} + X) \approx 5\, pb$$

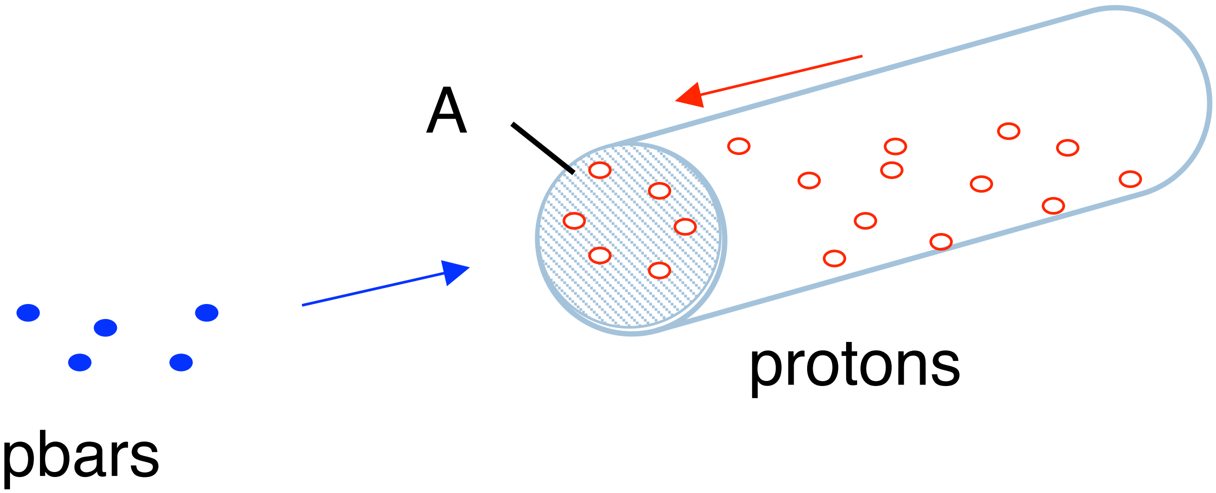
This is around $1/10^{10}$ of total

N/sec ~ σL = 5×10^{-4}/sec

A couple were created/hour but we only saw a small %

Luminosity

$$\text{Rate} = \mathcal{L}\,\sigma_{int}$$



$$P_{int} = N_{prot}\sigma_{int}/A$$

$$\mathcal{L} = f_{rev} N_{pbar}\, N_{prot}/A$$

Not easy to do this

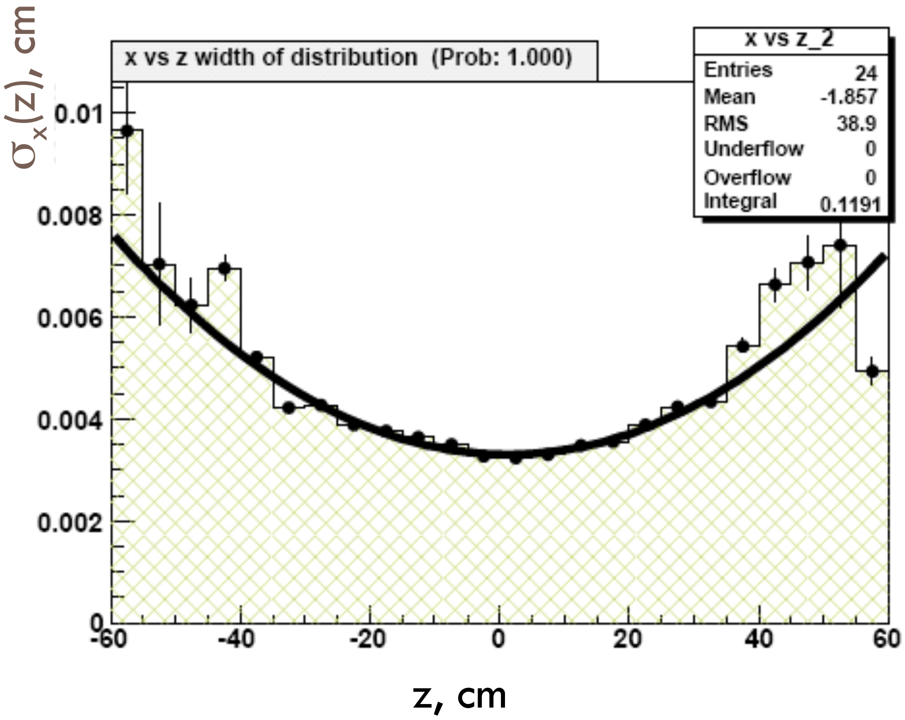
- Hard to measure currents accurately
- the beam density depends on the intrinsic beam lengths and widths and on the local beam optics.

$$\mathcal{L} = 2\, f_{rev} \iiiint \rho_1\, \rho_2 \;dx\, dy\, dz\, d(ct)$$

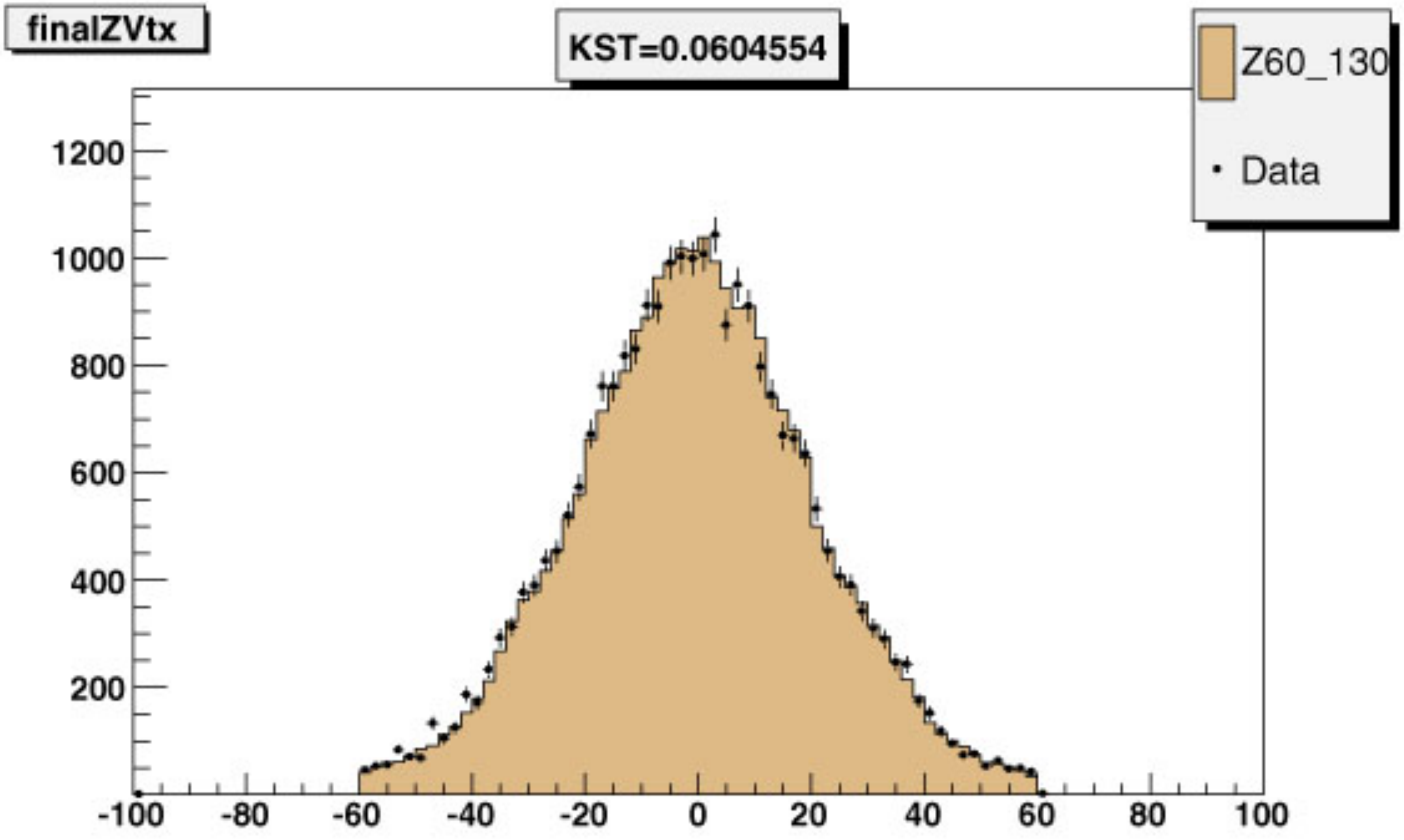
$$\begin{aligned}\rho(x,y,z,ct) = {} & \frac{N_1}{\sqrt{2\pi}\,\sigma_x} \exp[-\,(x+\Delta x/2)^2/\,2\,\sigma_x^2] \\ & \frac{1}{\sqrt{2\pi}\,\sigma_y} \exp[-\,(y+\Delta y/2)^2/\,2\sigma_y^2] \\ & \frac{1}{\sqrt{2\pi}\,\sigma_z} \exp[\,-(z+ct-ct_0)^2/\,2\sigma_z^2]\end{aligned}$$

You can measure the beam spot size as a function of distance along the beam to get $\sigma_x(z)$, $\sigma_y(z)$

- Also need estimates of beam current
- σ_z can be estimated from beam timing or N(z)
- Estimates are good to ~10% at the Tevatron.
- The beam position and size vary with time as the beam "heats up" and "cools down"



$\mathcal{L}(z)$ distribution for beam collisions

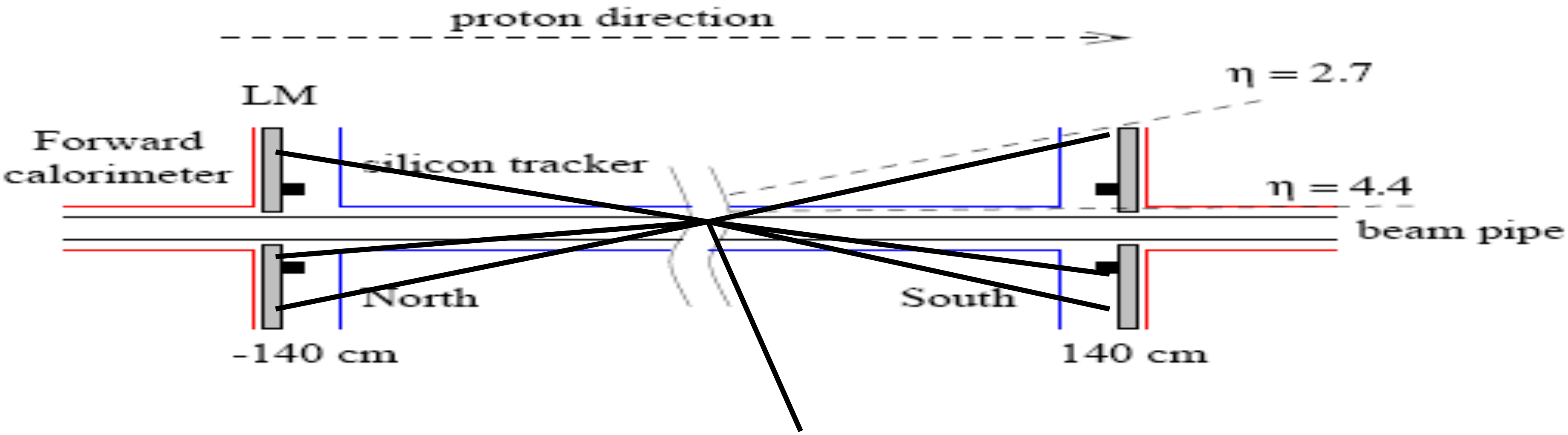


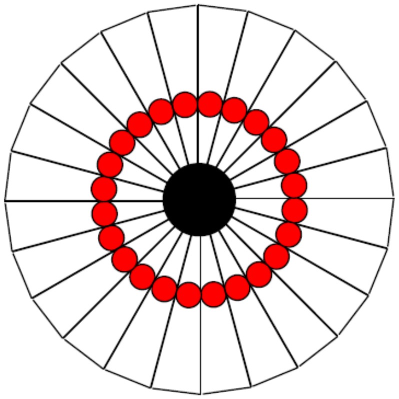
z(vtx) , cm

Alternate method

- Measure the total inelastic proton antiproton cross section in a special run.
- Count inelastic collisions as you run your regular experiment.
- Use those to get an estimate of your integrated luminosity.

D0 Luminosity System





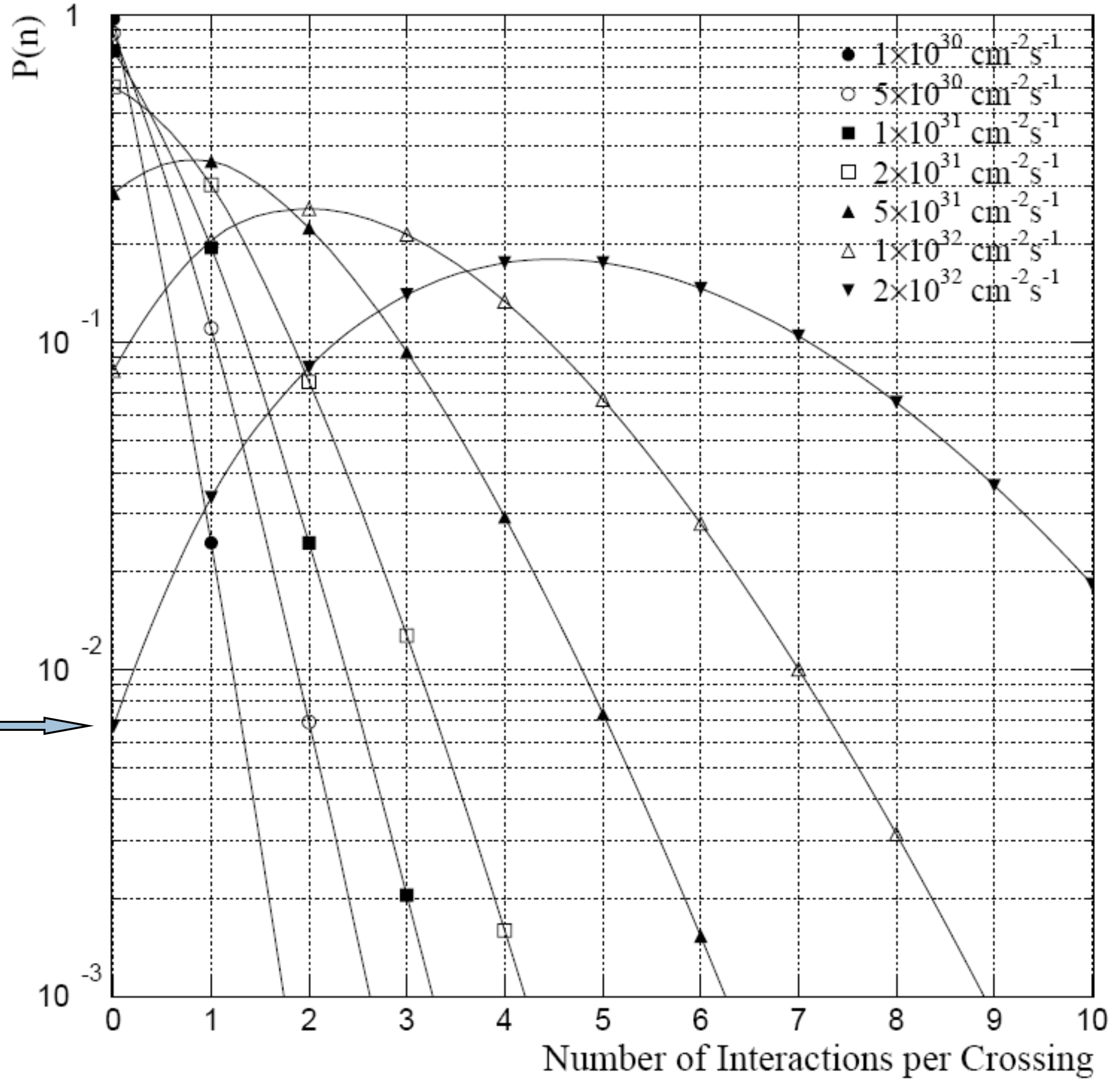
Count fraction, *P(0)*, of beam crossings without inelastic interactions.

.

Method

- Measure μ, the average number of interactions/crossing
- Apply acceptance and efficiency corrections and then compare to the inelastic pbar p total cross section to get the luminosity integrated over the 396 nsec crossing time.
- Typical μ are 1-10 at the Tevatron

Interactions/crossing



Typical running at the Tevatron

Luminosity errors

- The luminosity measurement has 3 sources of error
  - The inelastic cross section is only known to 3-4%
  - The luminosity detector is not perfect… 3-4%
  - Bookkeeping – you have to match the luminosity data stream up with your real data.
    - What happens if you lose a data tape? Flag data as bad?
    - We track this by matching "luminosity blocks", 1 minute periods of data.

So far we

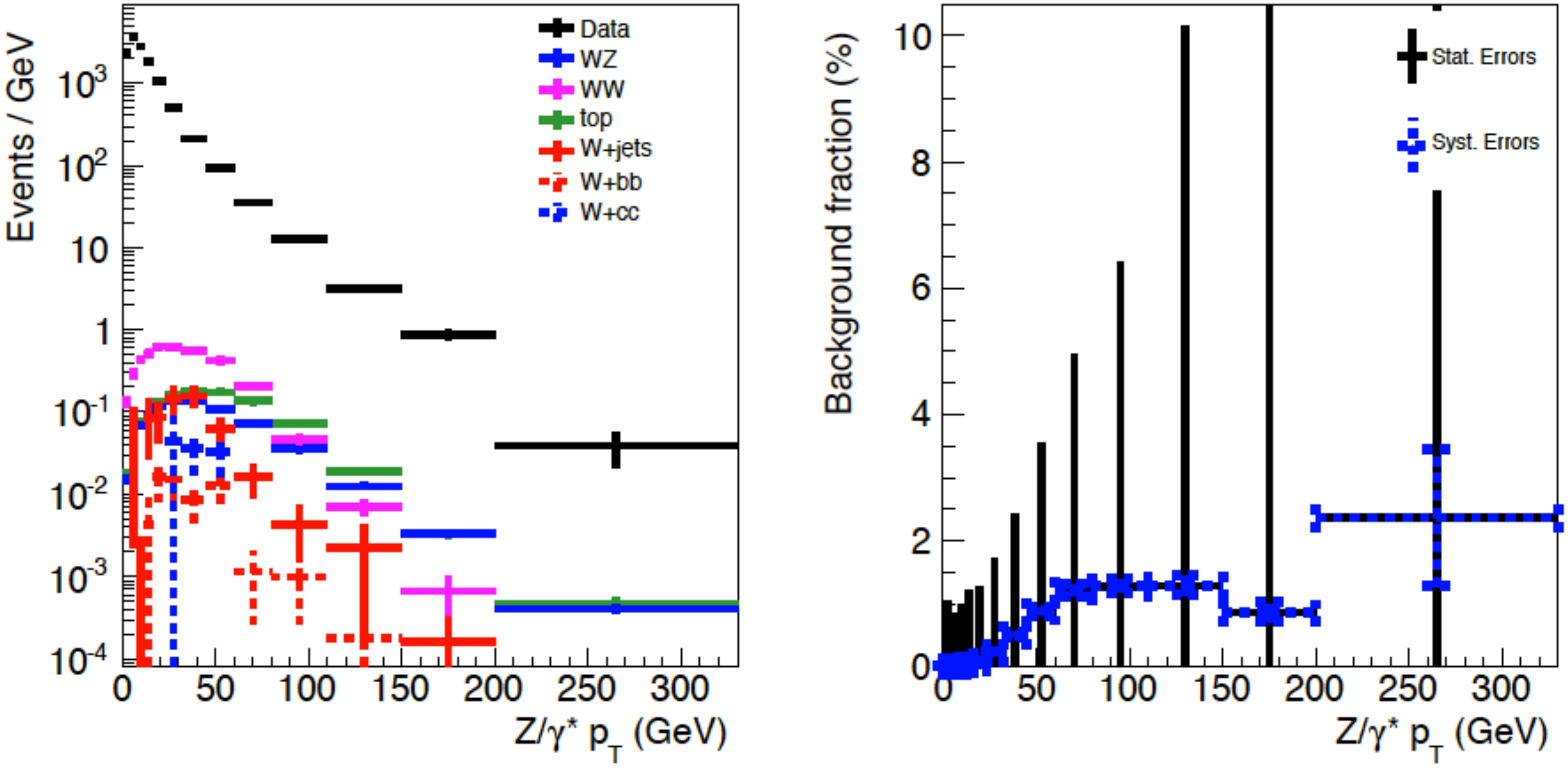
- Have a signal
- Have a normalization
- How do we find backgrounds?
- How do we correct for detector efficiency and resolution?
- How do we correct for Acceptance?

Estimating backgrounds

- Simulate and subtract
  - Requires good modeling
  - Hard to normalize absolutely
- Simulate and fit
  - Template method
- Vary cuts – matrix method

Simulated backgrounds

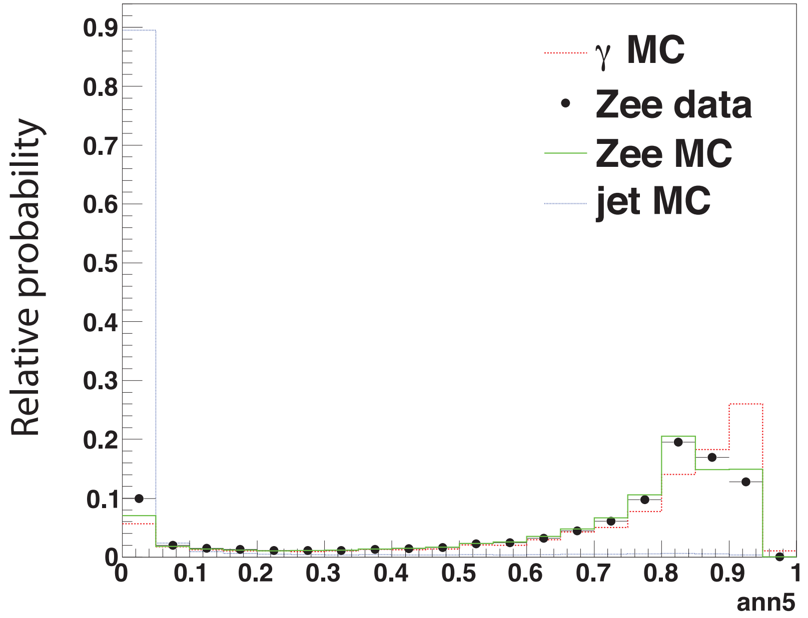
Muon channel

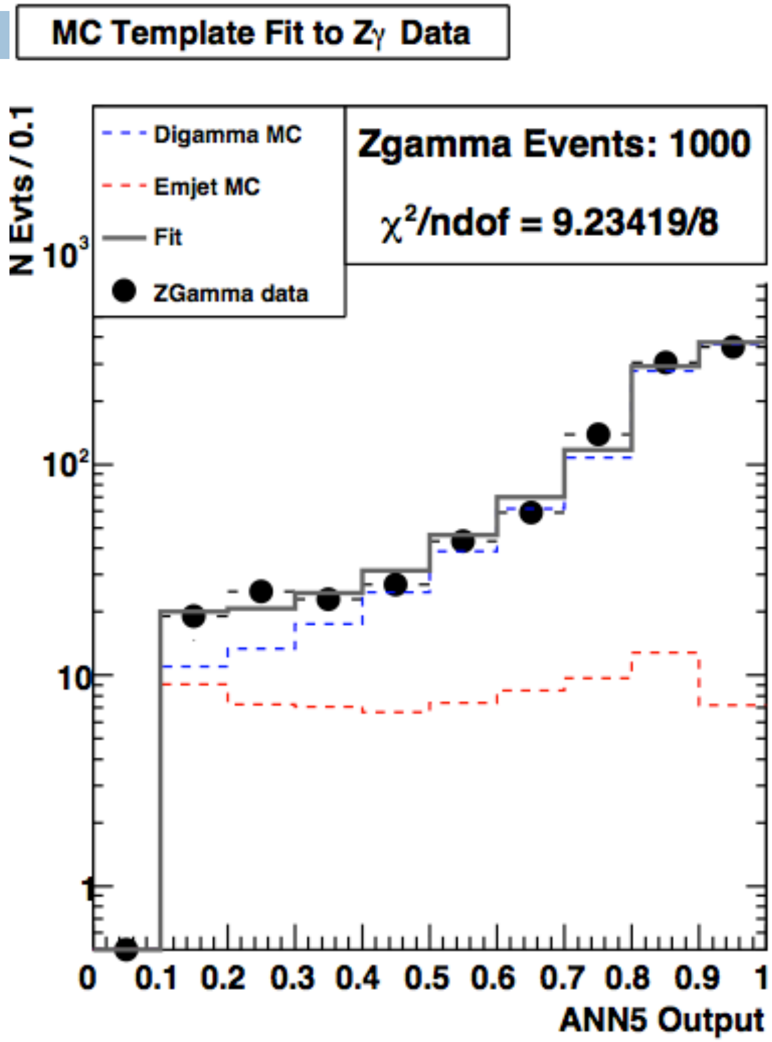


Background from template fit

Do this for Z+γ signal

Neural network discriminant for EM vs jet





Vary cuts – matrix method

Vary cuts

$$N_{loose} = \epsilon_{loose}\ N_e + f_{loose}\ N_{QCD}$$

$$N_{tight} = \epsilon_{tight}\ N_e + f_{tight}\ N_{QCD}$$

ϵ is efficiency for signal $\approx$ 70-100%.

f is probability to create a fake $\approx$ 0.1-10%

$$N_e \approx \frac{f_{tight} N_{loose} - f_{loose} N_{tight}}{\epsilon_{loose} f_{tight} - \epsilon_{tight} f_{loose}}$$

Efficiency

$$
\begin{aligned}
\mathbf{N_{obs}}(\mathbf{p_T},\phi,\eta,\mathbf{z}) &= B+\sigma A\epsilon\int \mathcal{L}\mathbf{dt}\\
&= B(p_T,\phi,\eta,z,\mathcal{L})+\\
&\int \sigma(p_T^0,\phi^0,\eta^0)\times\mathcal{L}(\mathbf{t},\mathbf{z})\times A(p_T^0,\phi^0,\eta^0,z^0)\times\\
&\quad R(p_T^0,\phi^0,\eta^0,z^0;p_T,\phi,\eta,z,\mathcal{L})\times\\
&\quad \epsilon(\mathbf{p_T},\phi,\eta,\mathbf{z},\mathcal{L})dp_T^0d\phi^0d\eta^0dt
\end{aligned}
$$

$\sigma(X_\alpha...)$ is the true cross section as a function of true variables $X_a....$

$\mathcal{L}(t,z)$ is the luminosity as a function of time.

$A(X_\alpha..)$ is the geometrical Acceptance as a function of true variables.

$R(X_\alpha...,\mathcal{L};X_a...)$ is the Resolution function which smears true $X_\alpha...$ to detected $X_a...$

$\epsilon(\mathbf{X_a}...,\mathcal{L})$is the probability that a particle is actually detected by a physical detector.

$B(X_a...,\mathcal{L})$ is the background

Efficiency

- First pass
  - Define an "acceptance region" where you would expect your detector to work
    - (say $|\eta|<3$, $p_T > 10$ GeV)
  - Take a particle level MC like pythia
  - Reweight kinematics to reflect best knowledge
  - Trace particles through detector
  - Overlay noise and interactions from other events
- Count how often you reconstruct the particle

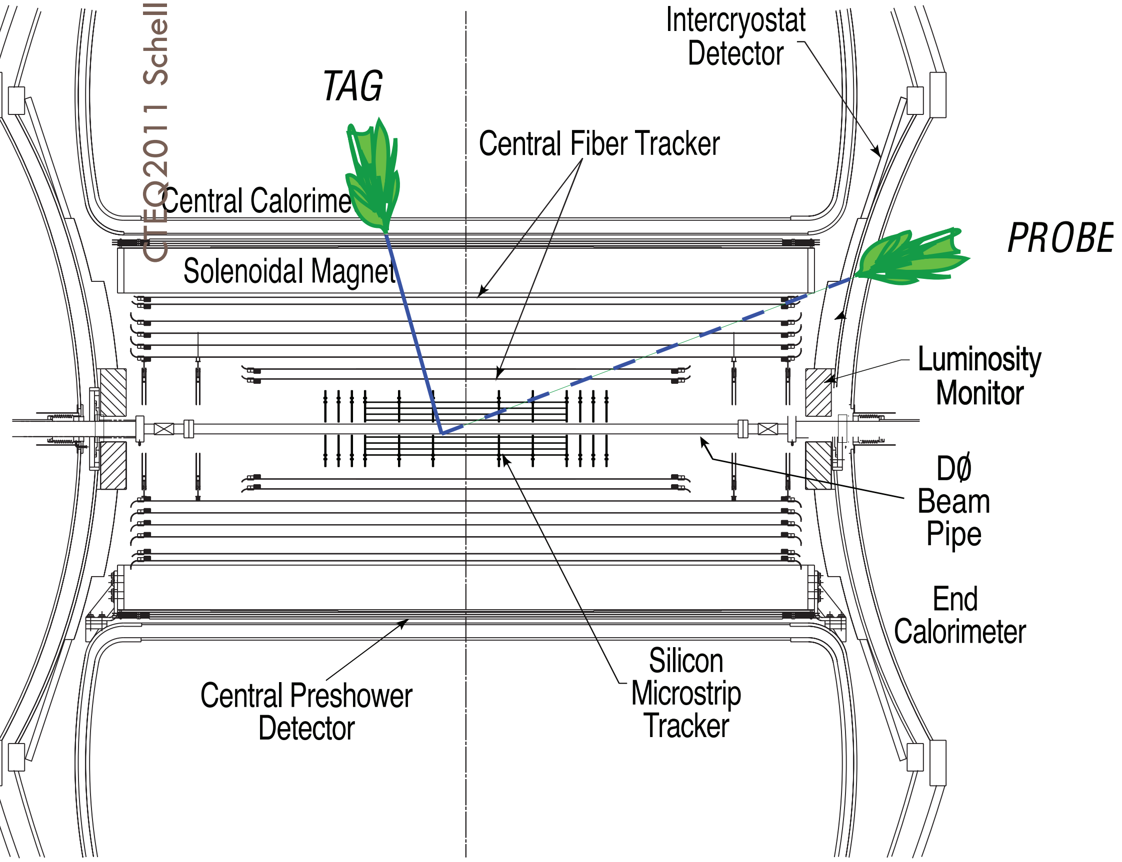
The raw simulation isn't good enough

- Use scale factors

- Find two ways of measuring something
- Compare the two in data and in simulation
- Correct by the ratio

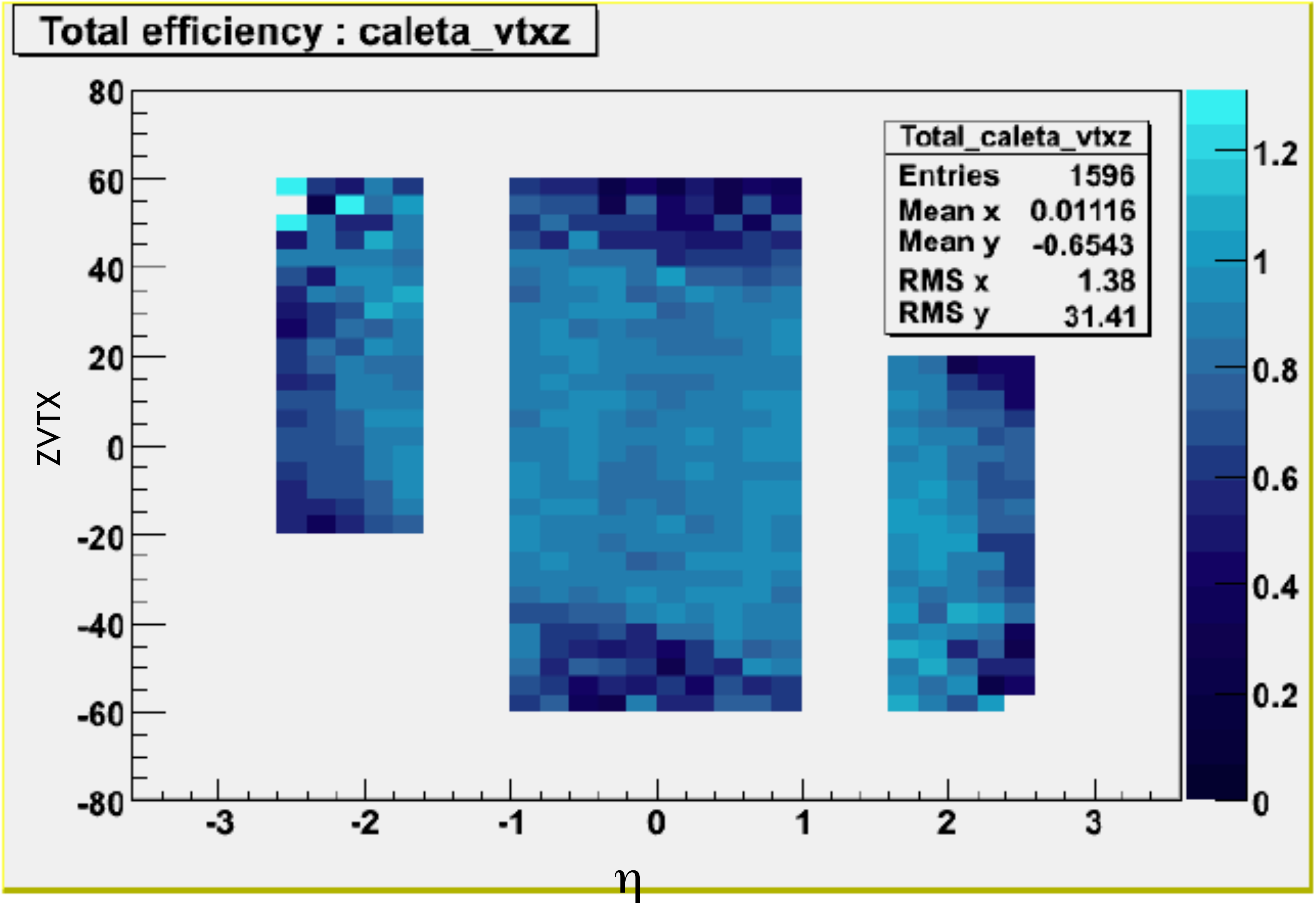
$$\epsilon(p_T, z) = \frac{(P \cdot T)_{data}/T_{data}}{(P \cdot T)_{sim}/T_{sim}} \epsilon_{sim}(p_T, z)$$

Tag and probe

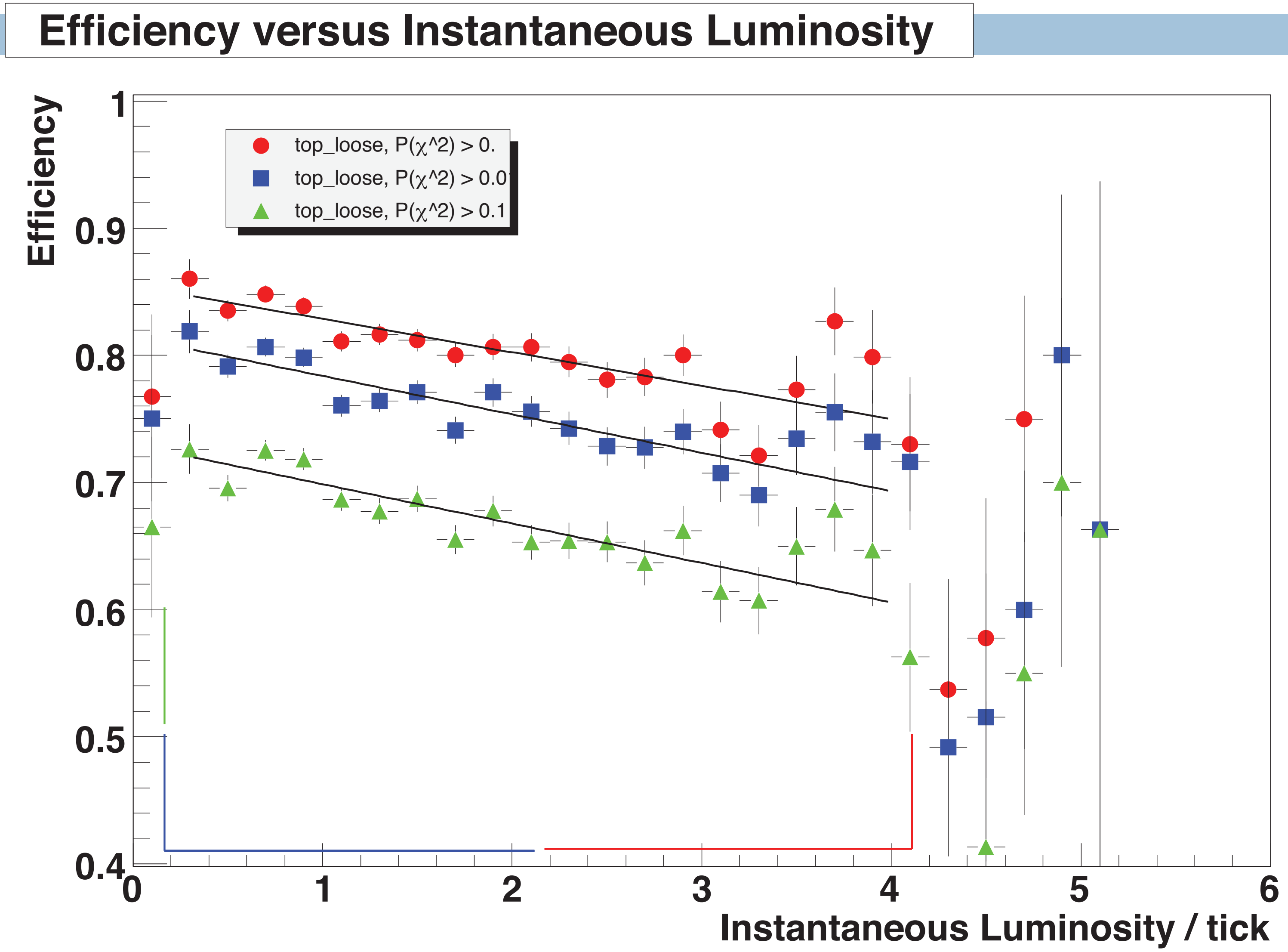


- Find a really good electron (TAG)
- Find a track which might be from a Z decay (PROBE)
- Did you find a shower

- Do the same for a shower (PROBE) that might have a track


Total efficiency : caleta_vtxz
Total_caleta_vtxz
Entries 1596
Mean x 0.01116
Mean y -0.6543
RMS x 1.38
RMS y 31.41
ZVTX
η
80
60
40
20
0
-20
-40
-60
-80
-3
-2
-1
0
1
2
3
1.2
1
0.8
0.6
0.4
0.2
0

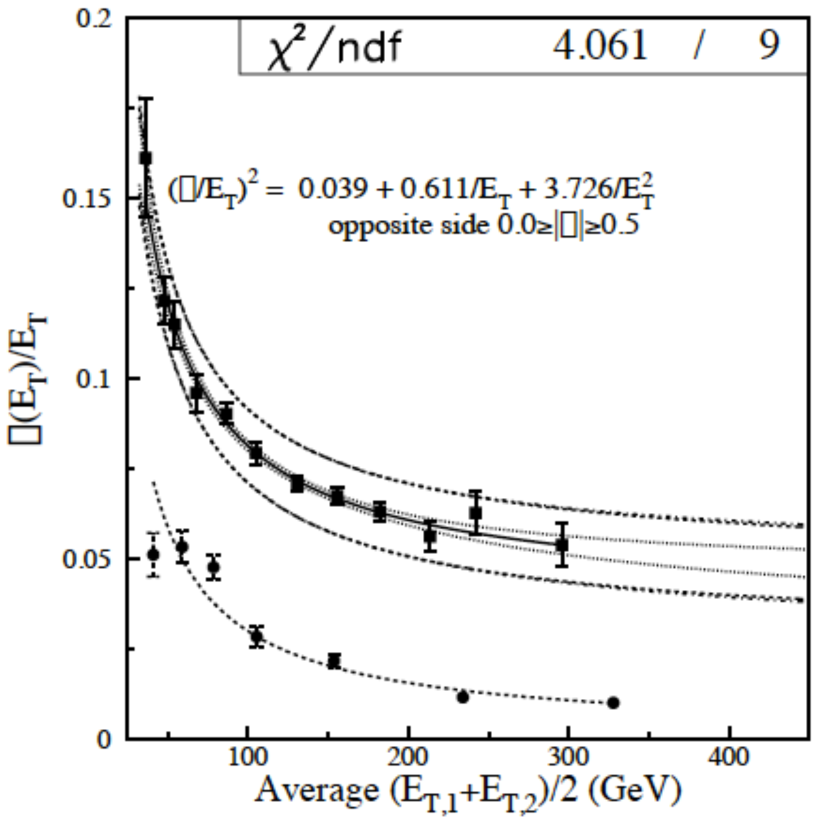
Typical electron efficiencies



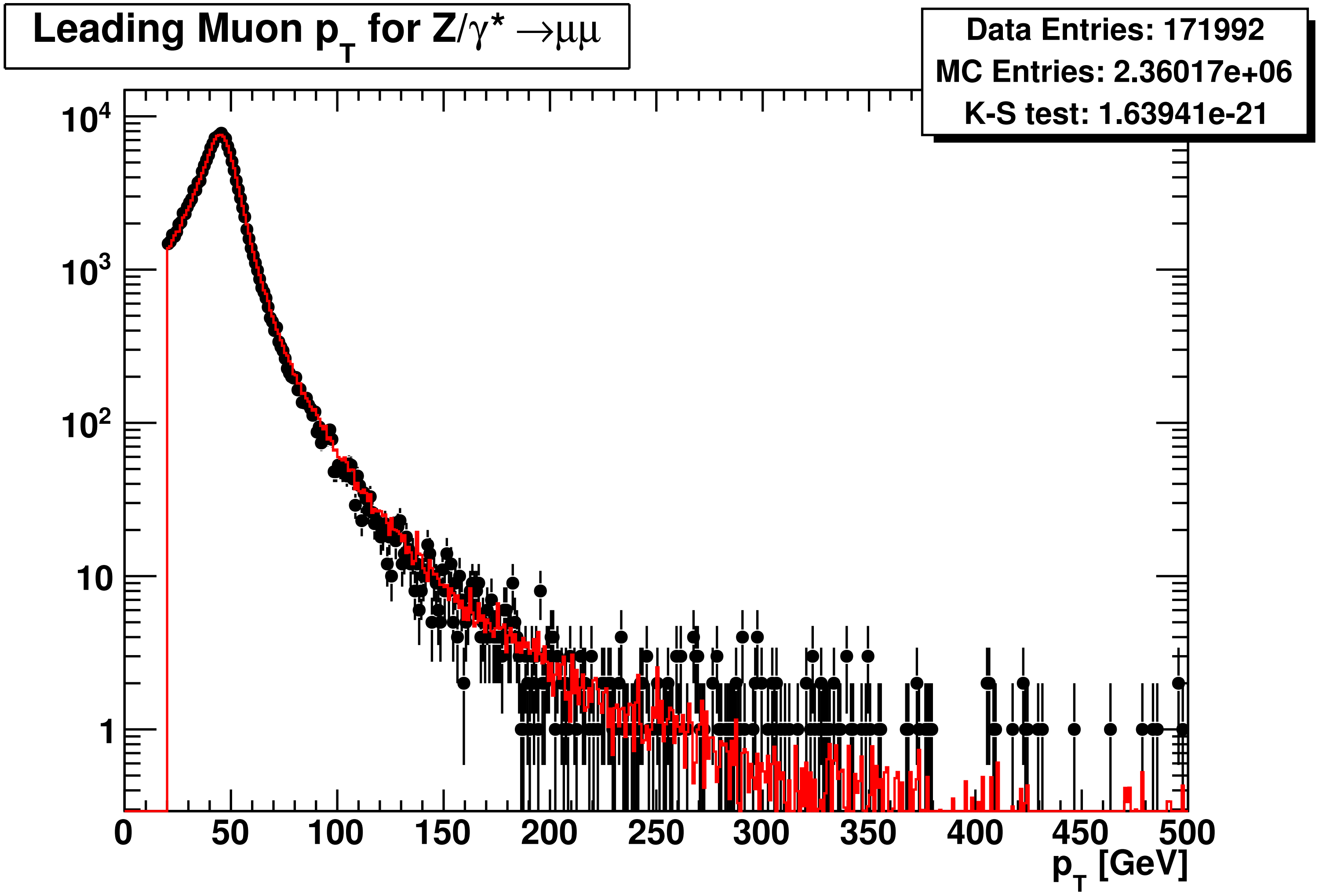
Resolutions

- Step 1 – Simulate
- Step 2 – Cry in frustration
- Step 3 – use the data
  - Tune to the Z peak width or
  - Compare energy of back to back jets

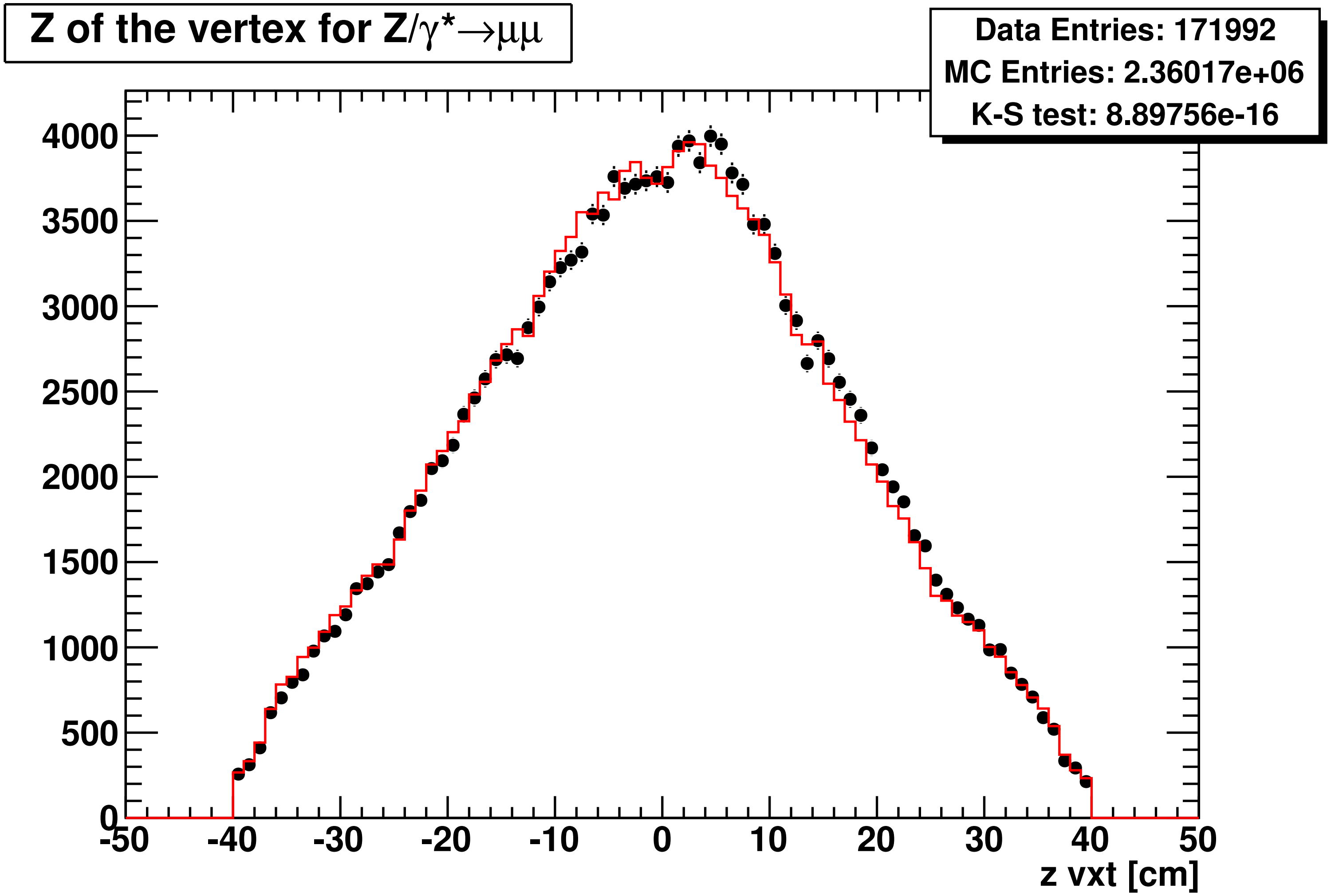
$$A = \frac{E_T^1 - E_T^2}{E_T^1 + E_T^2}$$



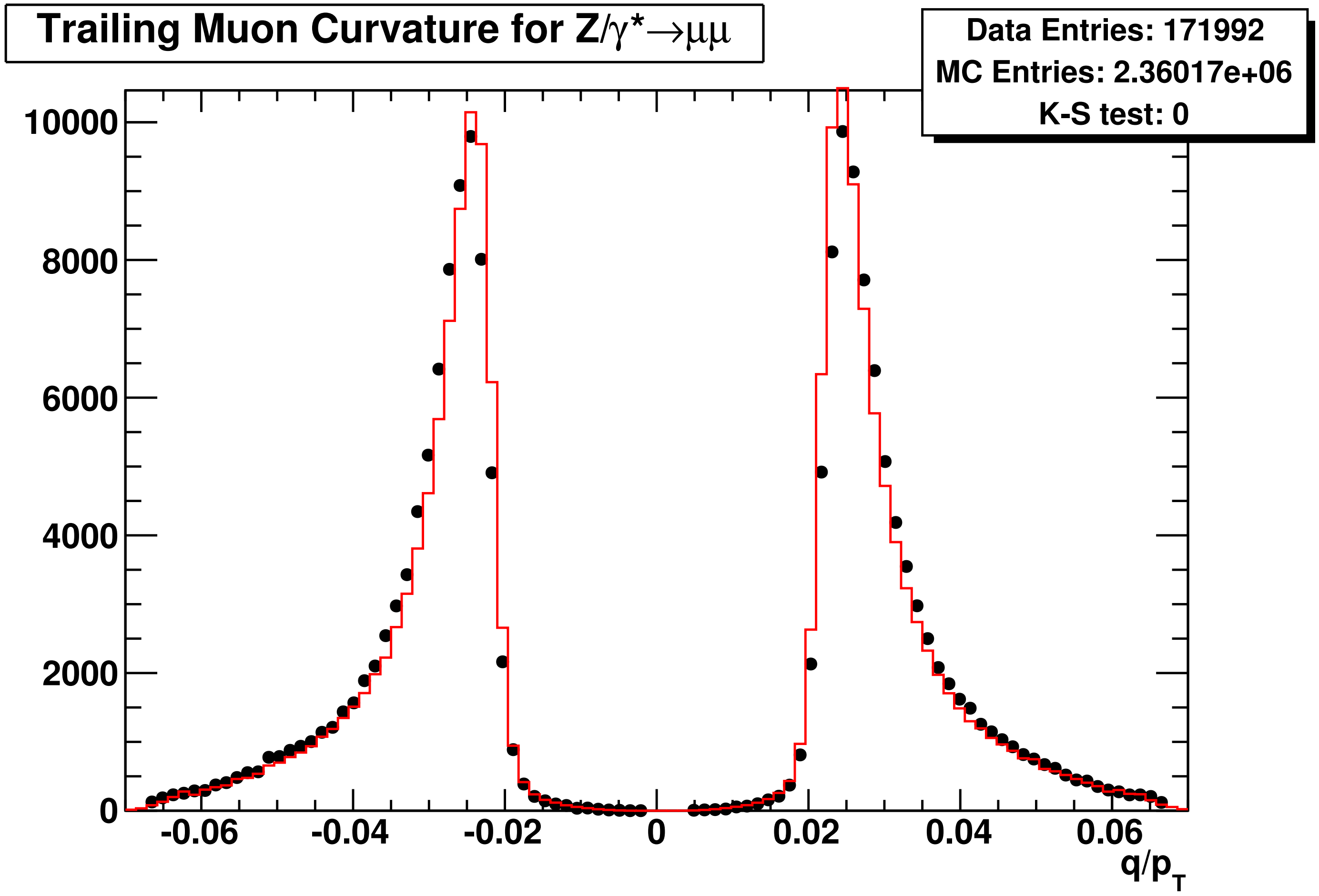
Check the tuned full detector simulation



Z vtx check


Z of the vertex for Z/γ*→μμ
Data Entries: 171992
MC Entries: 2.36017e+06
K-S test: 8.89756e-16
4000
3500
3000
2500
2000
1500
1000
500
0
-50
-40
-30
-20
-10
0
10
20
30
40
50
z vxt [cm]

Curvature check



- Ok, looks like we understand our efficiencies and resolutions pretty well in our simulation

- Correct data for backgrounds and efficiency
- Now on to unfolding the resolution effects

Unfolding the resolution effects

$$\begin{aligned}\mathbf{N_{obs}(p_T,\phi,\eta,z)} &= B+\sigma A\epsilon\int \mathbf{\mathcal{L}dt}\\ &= B(p_T,\phi,\eta,z,\mathcal{L})+\\ &\int \sigma(p_T^0,\phi^0,\eta^0)\times\mathcal{L}(\mathbf{t},\mathbf{z})\times A(p_T^0,\phi^0,\eta^0,z^0)\times\\ &\quad \mathbf{R}(\mathbf{p_T^0},\phi^0,\eta^0,\mathbf{z^0};\mathbf{p_T},\phi,\eta,\mathbf{z},\mathcal{L})\times\\ &\quad \epsilon(\mathbf{p_T},\phi,\eta,\mathbf{z},\mathcal{L})dp_T^0 d\phi^0 d\eta^0 dt\end{aligned}$$

$\sigma(X_\alpha ...)$ is the true cross section as a function of true variables X_a....

$\mathcal{L}(t,z)$ is the luminosity as a function of time.

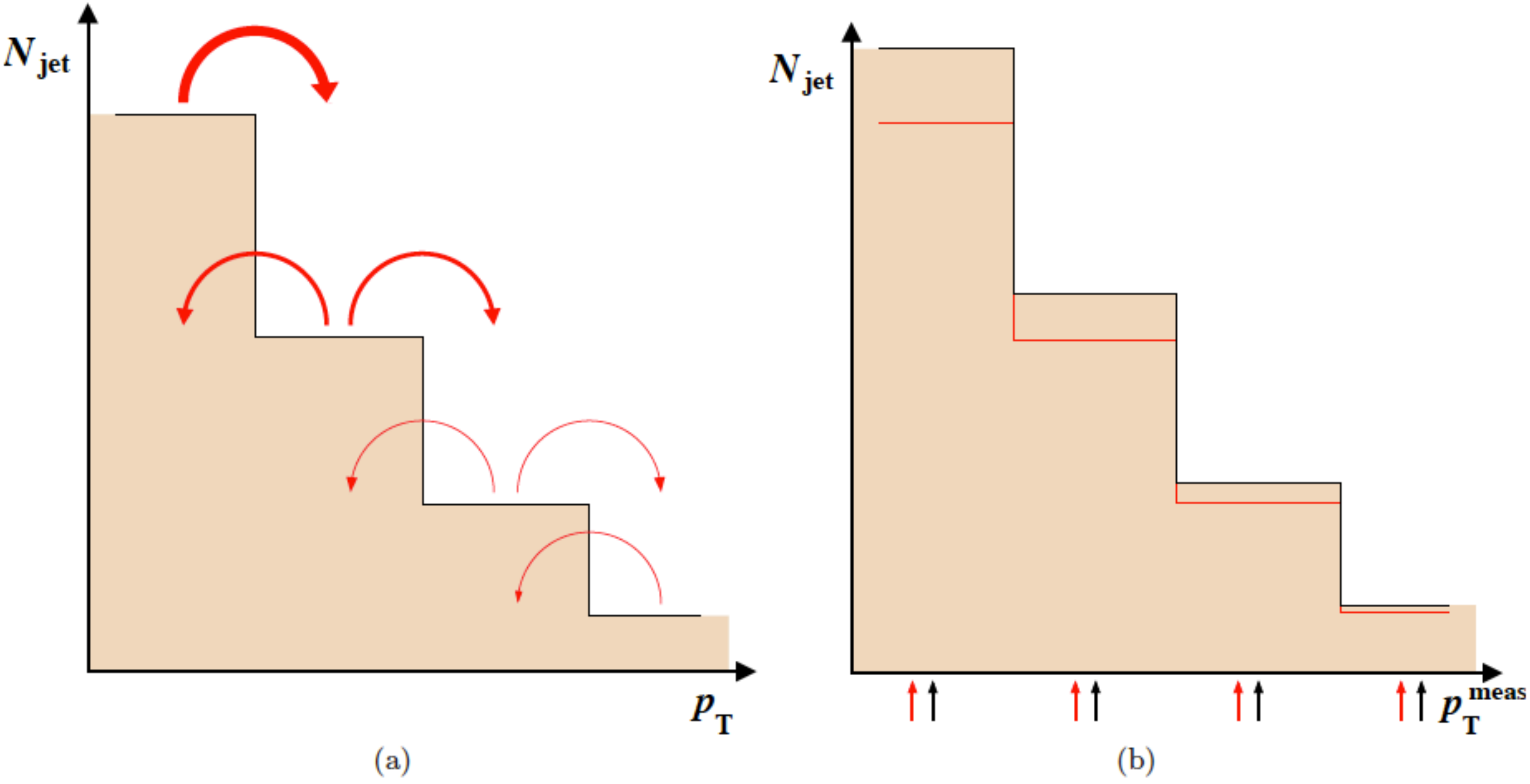
$A(X_\alpha ..)$ is the geometrical Acceptance as a function of true variables.

$\mathbf{R}(\mathbf{X}_\alpha ...,\mathcal{L};\mathbf{X_a}...)$ is the Resolution function which smears true $X_\alpha ...$ to detected $X_a ...$

$\epsilon(\mathbf{X_a}...,\mathcal{L})$is the probability that a particle is actually detected by a physical detector.

$B(X_a ...,\mathcal{L})$ is the background

Resolution Smearing



(a) (b)

Unfolding method 1 - matrix

- Invert the resolution matrix

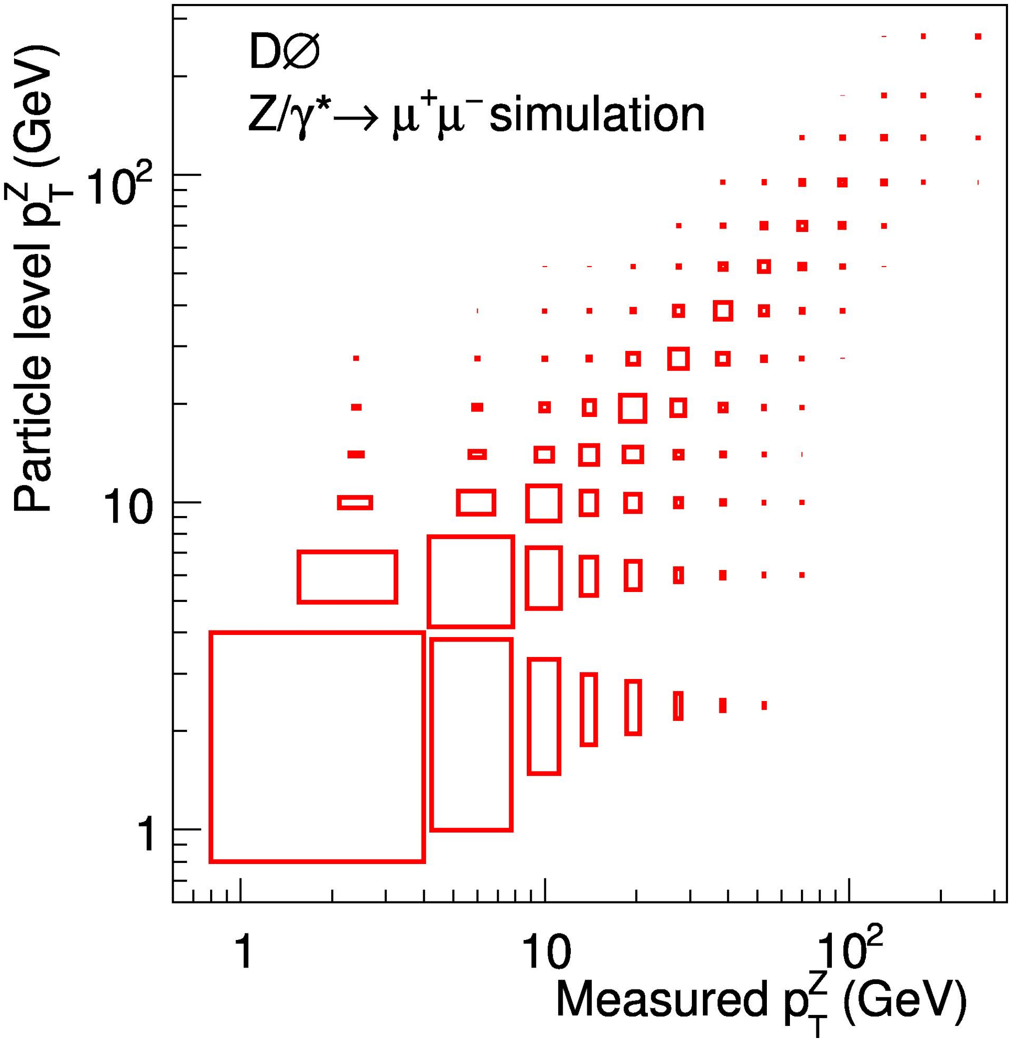
$$N_a = B_a + \epsilon_a R_{a\alpha} \; A_\alpha \; \sigma_\alpha \int \mathcal{L} dt$$

$$\text{Invert}$$

$$\sigma_\alpha = \frac{1}{A_\alpha} \; R^{-1}\alpha a \; \frac{(N_a - B_a)}{\epsilon_a} \int \mathcal{L} dt$$

- Common to regularize the matrix
  - Improved stability
  - Bias towards smoothing the function

Smearing Matrix R(X_α->X_a)



Unfolding method 2 - Ansatz

- Convolute a trial unsmeared function with the resolution as a function of (pt, η)

$$f(p_T, \eta) = N_0 \left(\frac{p_T}{100\ \mathrm{GeV/c}} \right)^{-\alpha} \left(1 - \frac{2 p_T \cosh(y_{\mathrm{min}})}{\sqrt{s}} \right)^{\beta} \exp\left(-\gamma p_T\right).$$

- Fit this convoluted “smeared” function to your data
- Correct data by the ratio of the unsmeared to smeared “ansatz” function.

$$R(X_\alpha \to X_a) = \delta_{\alpha a} \frac{\mathrm{Smeared}(X_a)}{\mathrm{Unsmeared}(X_\alpha)}$$

- This works if a simple functional form can fit your data well and the resolution function is well understood.

Ansatz fit

$$R(X_\alpha \to X_a) = \delta_{\alpha a}\frac{\text{Smeared}(X_a)}{\text{Unsmeared}(X_\alpha)}$$

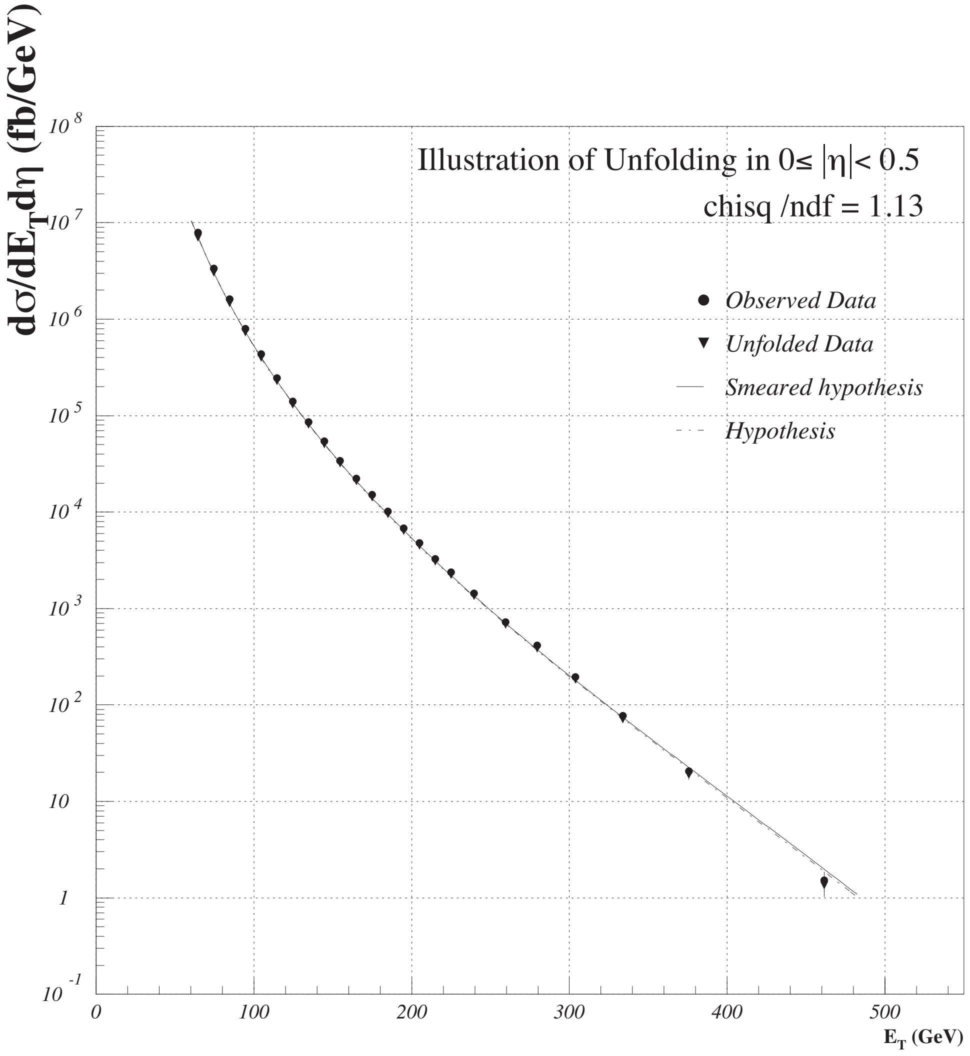
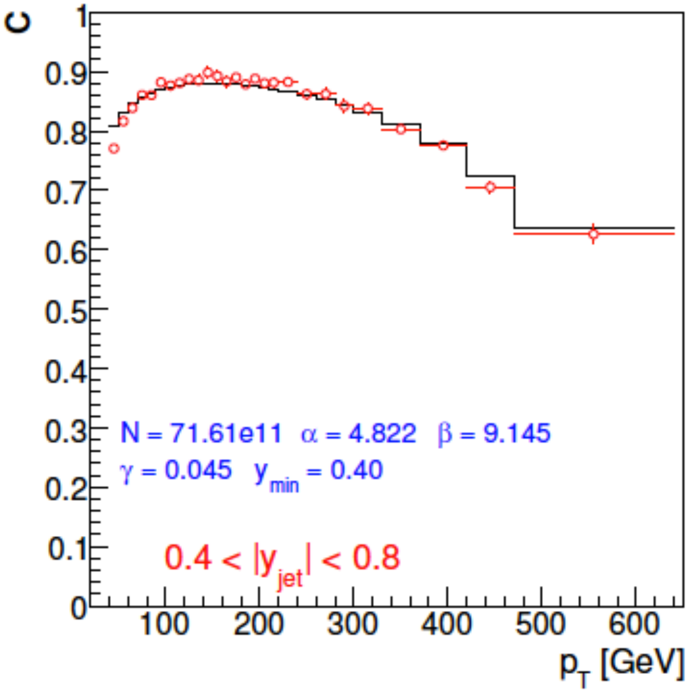


Figure 49: *Illustration of the data unfolding procedure in the rapidity region* $|\eta| < 0.5$.



Now for the final step: Acceptance

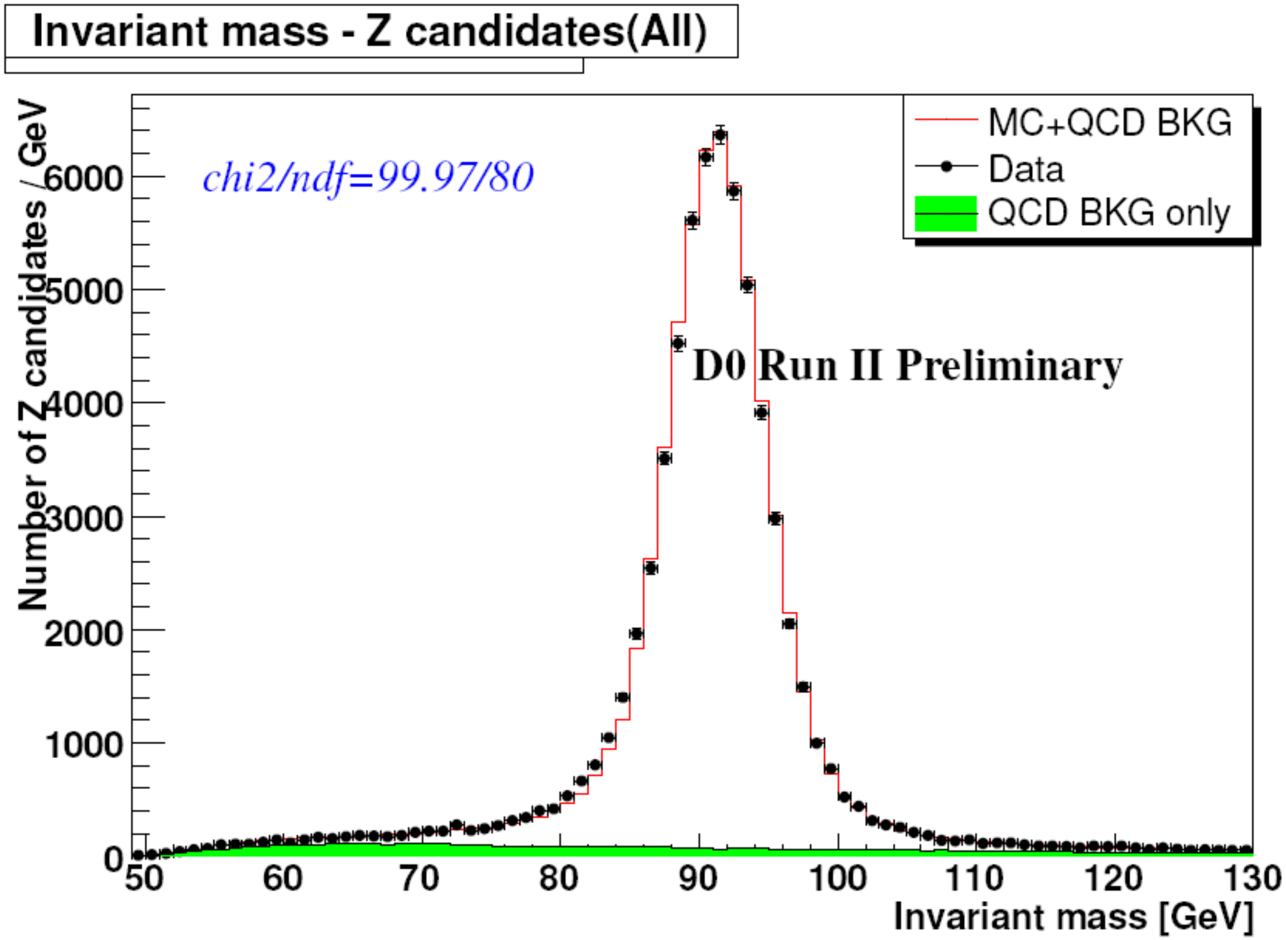
- Correct from the region we cover to the rest of phase space
- Why do this instead of including in the "efficiency"?

  - This correction is big – factor of 3 for Z production

$\varepsilon \sim 40\%,\ A \sim 30\%$

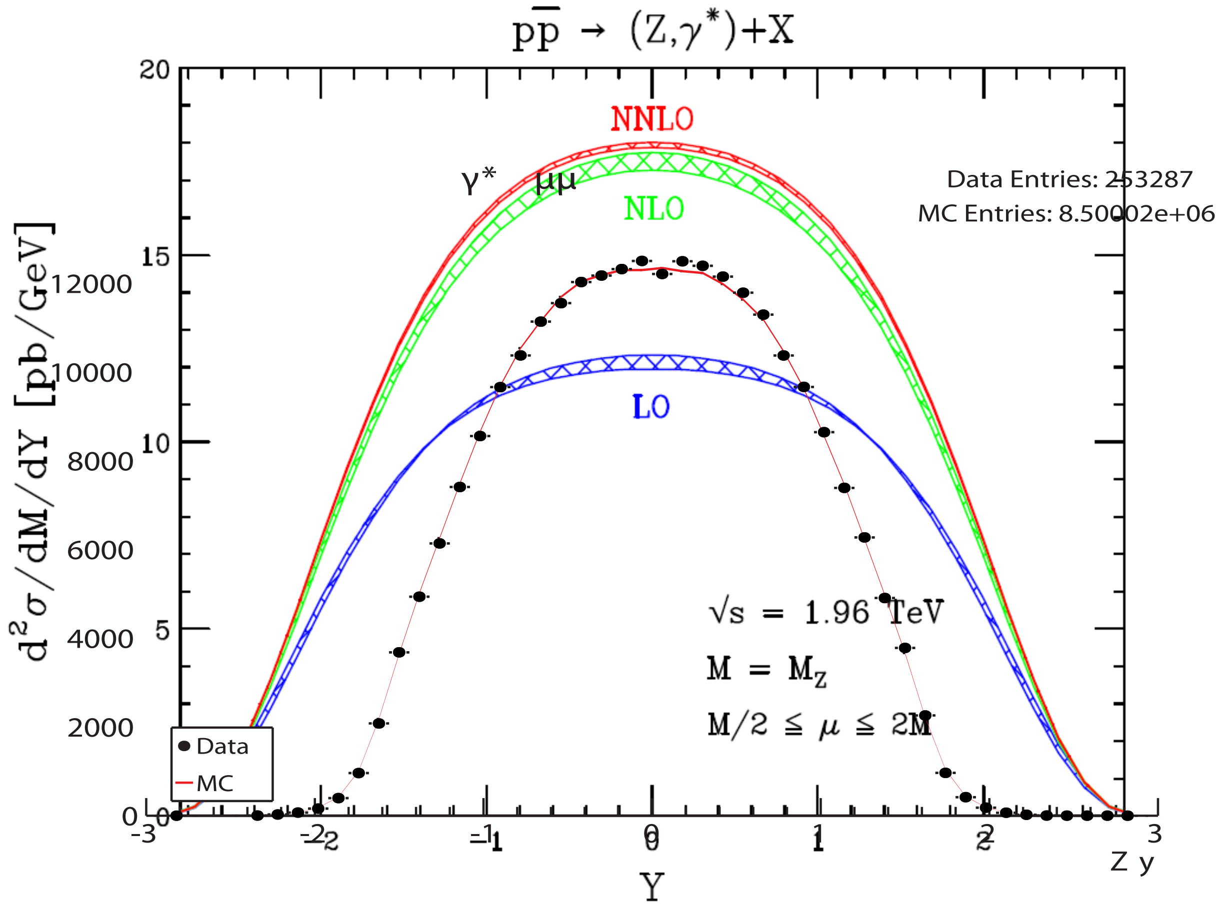
  - Only depends on generator level quantities
    - Can use the best NNLO simulations

D0 1 fb^{-1} sample



63,000 Z^0

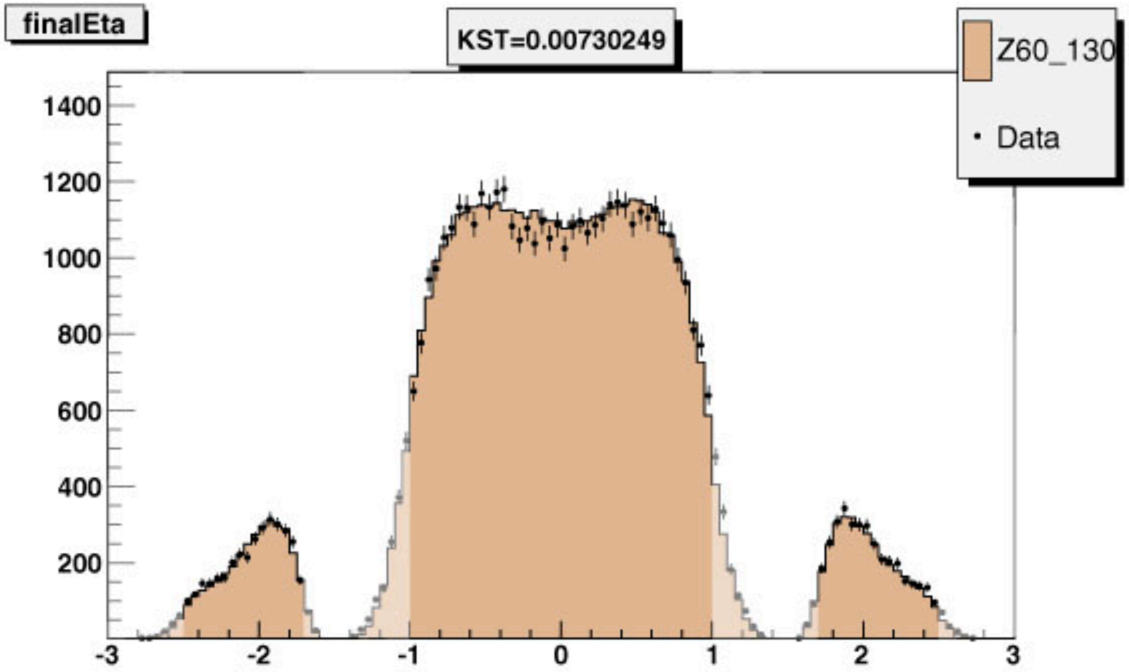
Z Rapidity raw and generated



rapidity

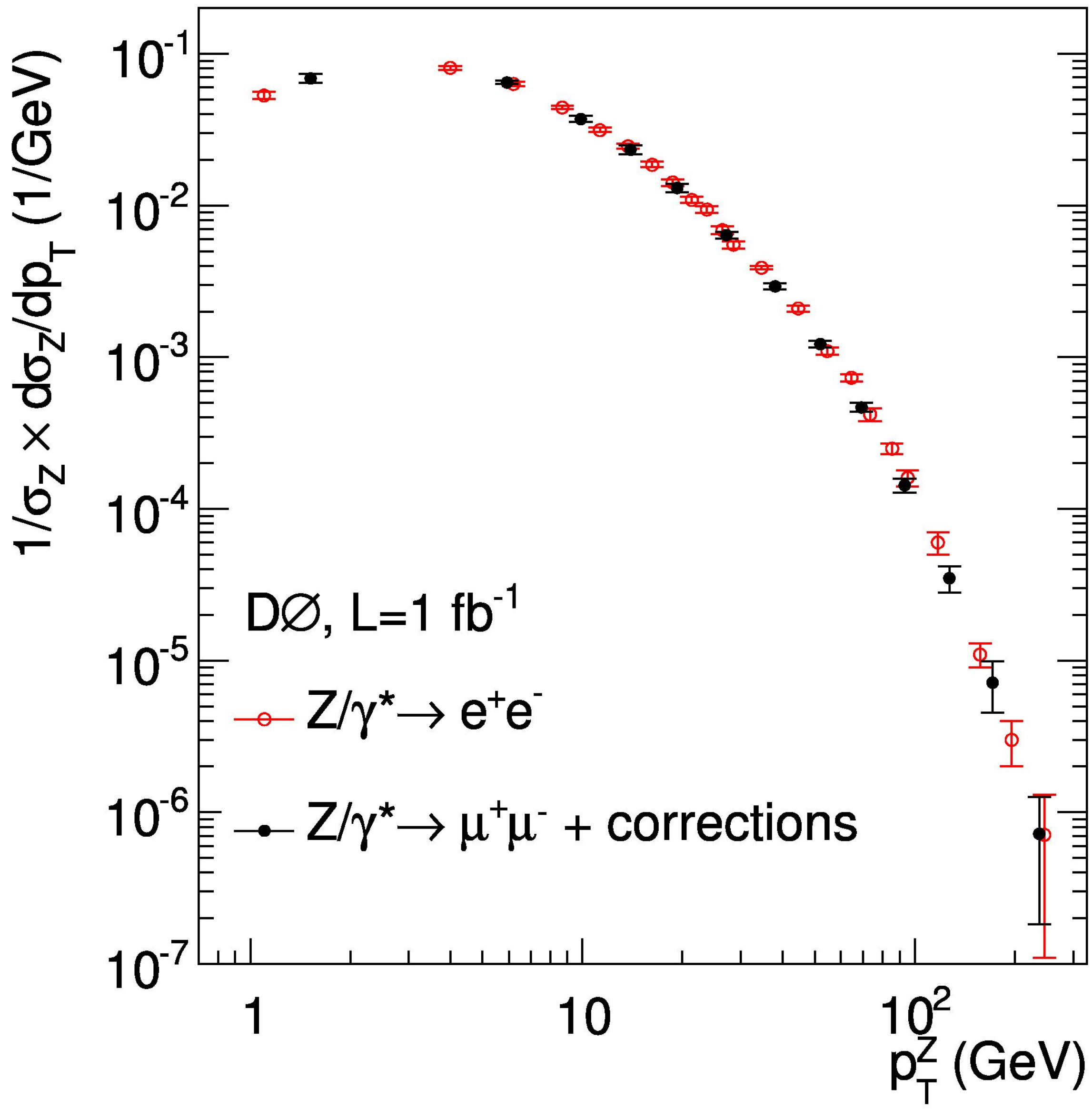
We actually define acceptance as $|\eta_e| < 1$ or $1.5 < |\eta_e| < 2.4$ and $p_T > 25$ GeV

- ☐ Cut cross section in electron rapidity and pt

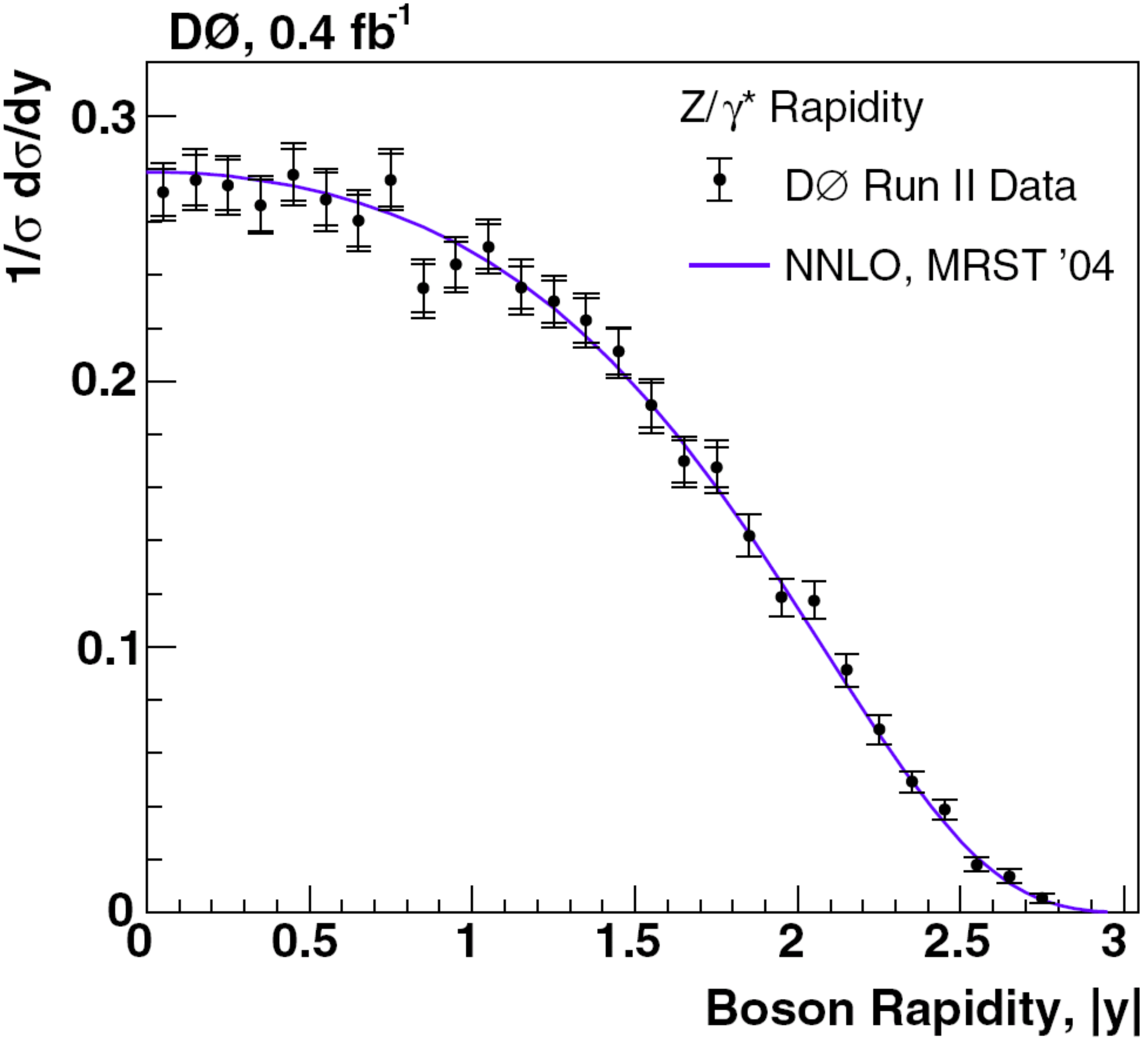


electron rapidity

Fully corrected normalized distributions



PHYSICAL REVIEW D **76**, 012003 (2007)



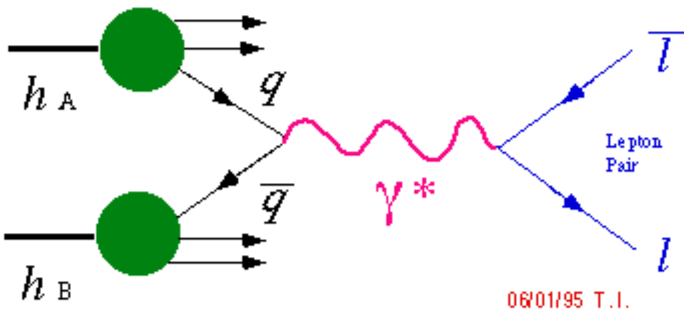
Theory

$$\frac{d\sigma^{h_1 h_2}}{dQ^2\, dp_T^2\, dy\, d\Omega^*} = \sum_{a,b} \int dx_1\, dx_2\, f_a^{h_1}(x_1, \mu_F^2)\, f_b^{h_2}(x_2, \mu_F^2)\, \frac{s\, d\hat{\sigma}_{ab}}{dQ^2\, dt\, du\, d\Omega^*} \left(x_1 P_1, x_2 P_2, \alpha_s(\mu_R^2)\right)$$

Parton density functions

Hard cross section

The Drell-Yan Process



Higher order QCD is important

gq→Zq, qq→Zg etc.

Cross section*B(Z→ee)
LO ~ 180 pb
NLO ~ 250 pb
NNLO ~ 260 pb

Z→ ee Errors

Statistical	0.62%
Preselection efficiency	0.85%
Radiative corrections	<0.5%
ID eff stat error	0.4%
Tag-probe bias	0.3%
Noise corrections	0.22%
Vertex z	0.6%
Cut variations	1.5%
Total systematic error	2.0%
PDF error on sigma(tot) Luminosity	+1.3% -1.7% 6.1%

More fun

- Radiative corrections
- Calibration
- Jet definitions
- B-tagging

- My list of things to check is around 80 items long