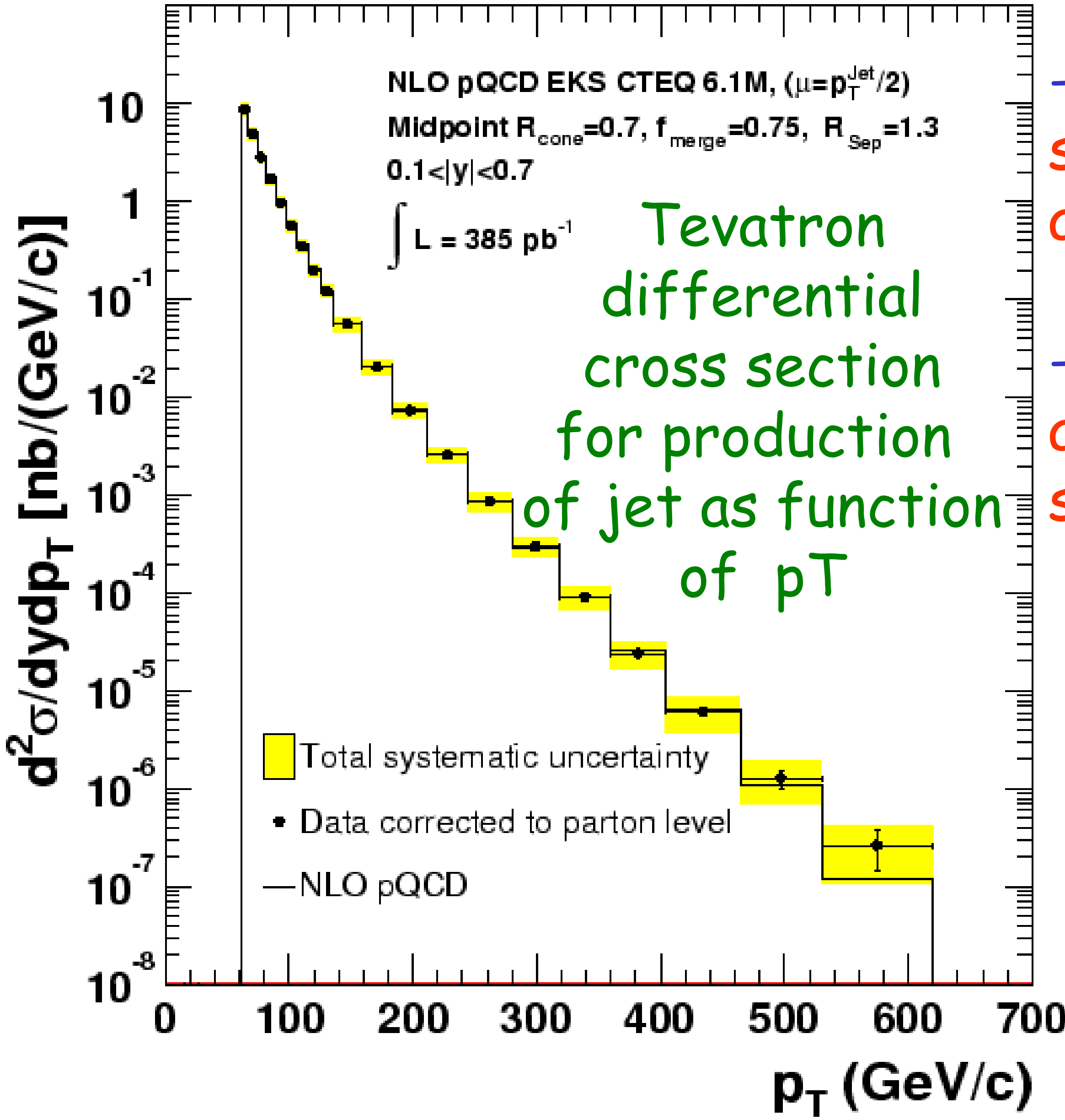
Binned Cross Section Measurements



- Last lecture: cross sections, luminosity and Monte Carlo intro

-Today: how to measure a differential cross section …
  - Acceptances
  - Binning choices
  - More corrections
  - Systematic effects

Reminder: Cross Sections

A crucial link between experimental and theoretical worlds!

$$\sigma = \frac{N_{obs} - N_{b/g}}{L \cdot A_{cc}}$$

... where efficiency ε absorbed in A_{cc}

- A cross section should be uniquely defined in terms of configurations of **observable** particles… i.e. in terms of observed hadrons, not Partons or `Leading Order'!… theories develop and improve; data are fixed by nature!

- Definitons in terms of **Lorentz invariants** preferred!

$N_{b/g}$ and A_{cc} can be obtained from Monte Carlo models, but should always check that our analysis is ROBUST …
i.e. relatively insensitive to the details of the Monte Carlo!!!

Reminder: Monte Carlo Simulations

- 3 steps: 1) Event **generator** models underlying physics
  … outputs a list of `stable' particles
  2) **Simulation** of detector response to particles
  3) **Reconstruction** identical to that of data

- If all steps are done well, MC output should closely model the data, but crucially, we have full info about **parton** and **hadron** levels as well as **reconstructed** level!

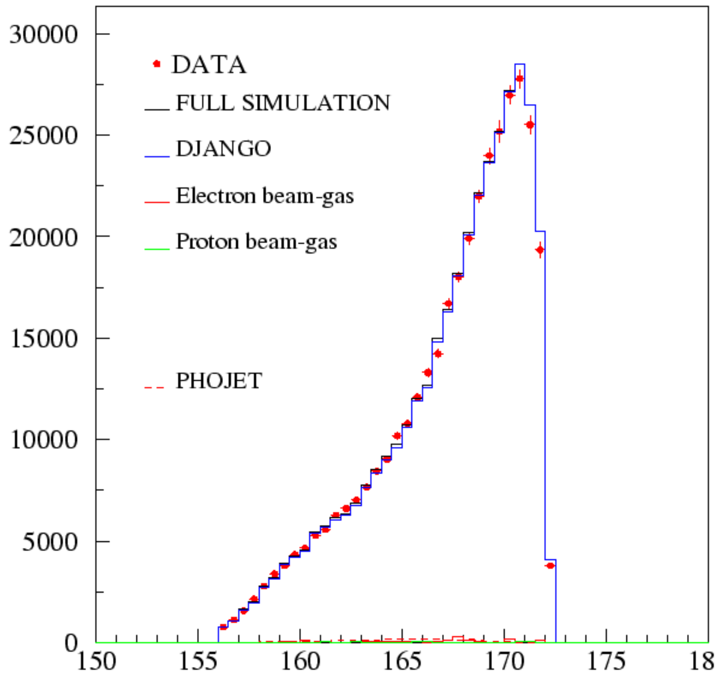
  … Can answer questions about our measurement. e.g:
  - What was my acceptance and efficiency?
  - What fraction of my `electrons' are really photons?

Answers are (slightly) affected if underlying generator is poor model of physics. Test with `control distributions'

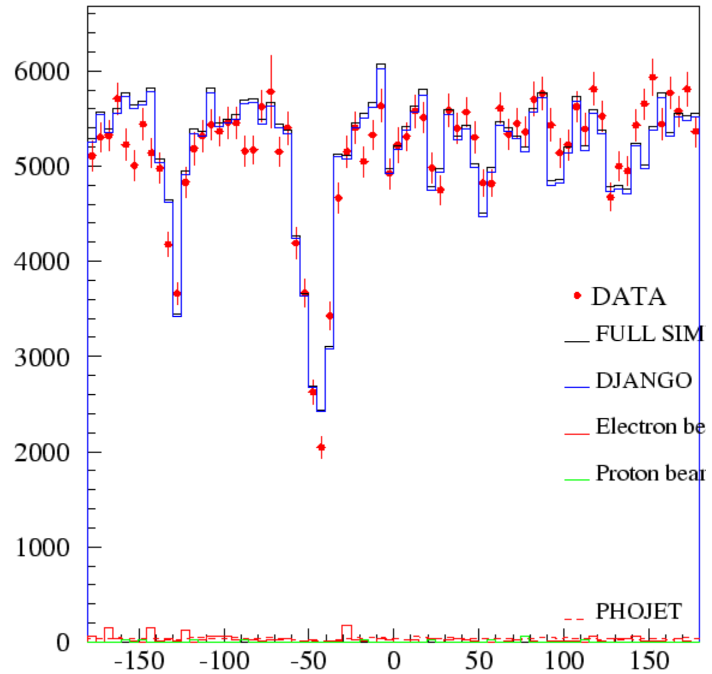
e.g. Control Distributions of reconstructed data (H1)

… compare data before any corrections with rec level MC …

Scattered electron θ

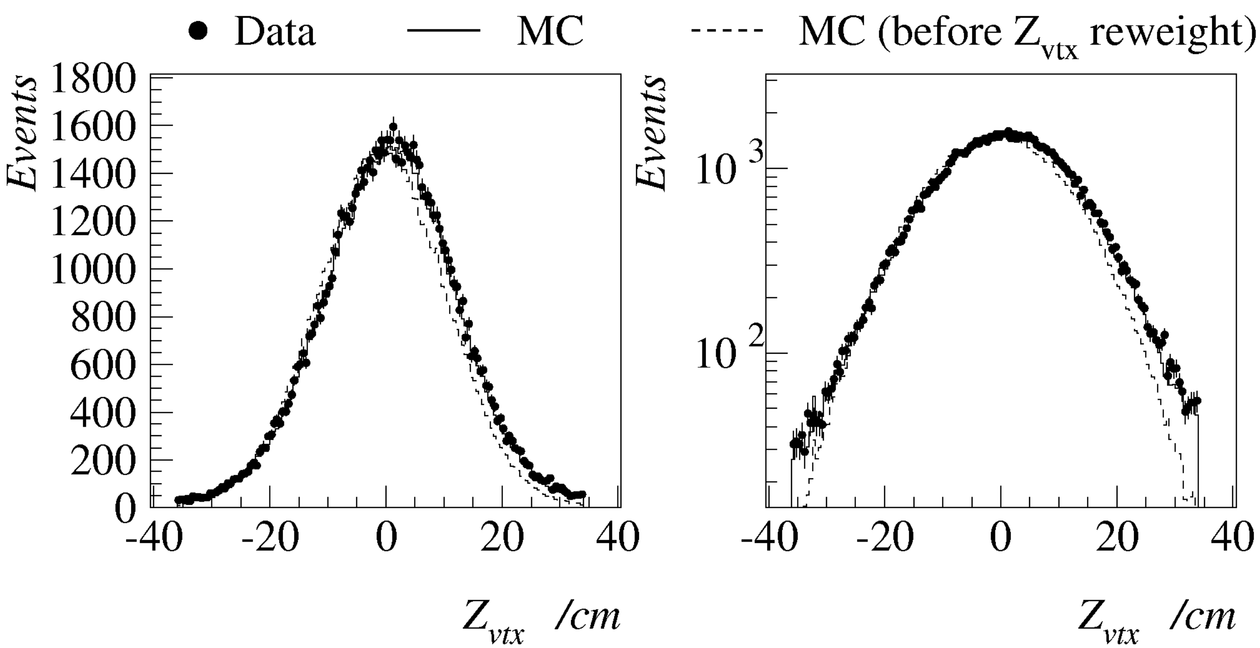


Scattered electron φ

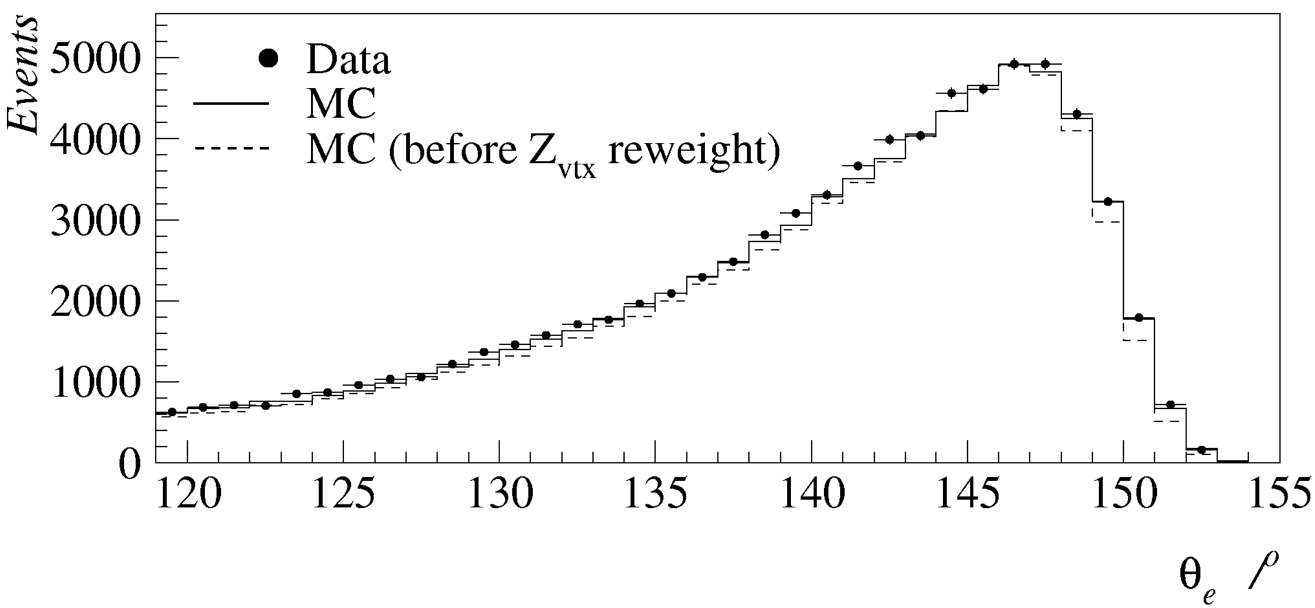


- Good description by DJANGO DIS Monte Carlo simulation ☺
- Big holes in φ due to dead bits of detector – but described by simulation
- Background Monte Carlo (PHOJET) gives tiny contribution
- Reassuring results – happy with model - use to correct for inefficiencies such as the holes!

Bad Control Distributions and their Influence



In this H1 example, distribution of simulated z position of interaction vertex is slightly wrong due to a mistake in steering parameters



Fixing vertex distribution results in much better description of an important phsysics distribution – electron scattering angle.

What is a Differential Cross Section

Often we are interested in the dependence of a cross section on a variable (e.g. transverse energy of a jet, pseudorapidity, momentum transferred …)

Work in terms of `differential cross sections e.g. $d\sigma / dE_t$

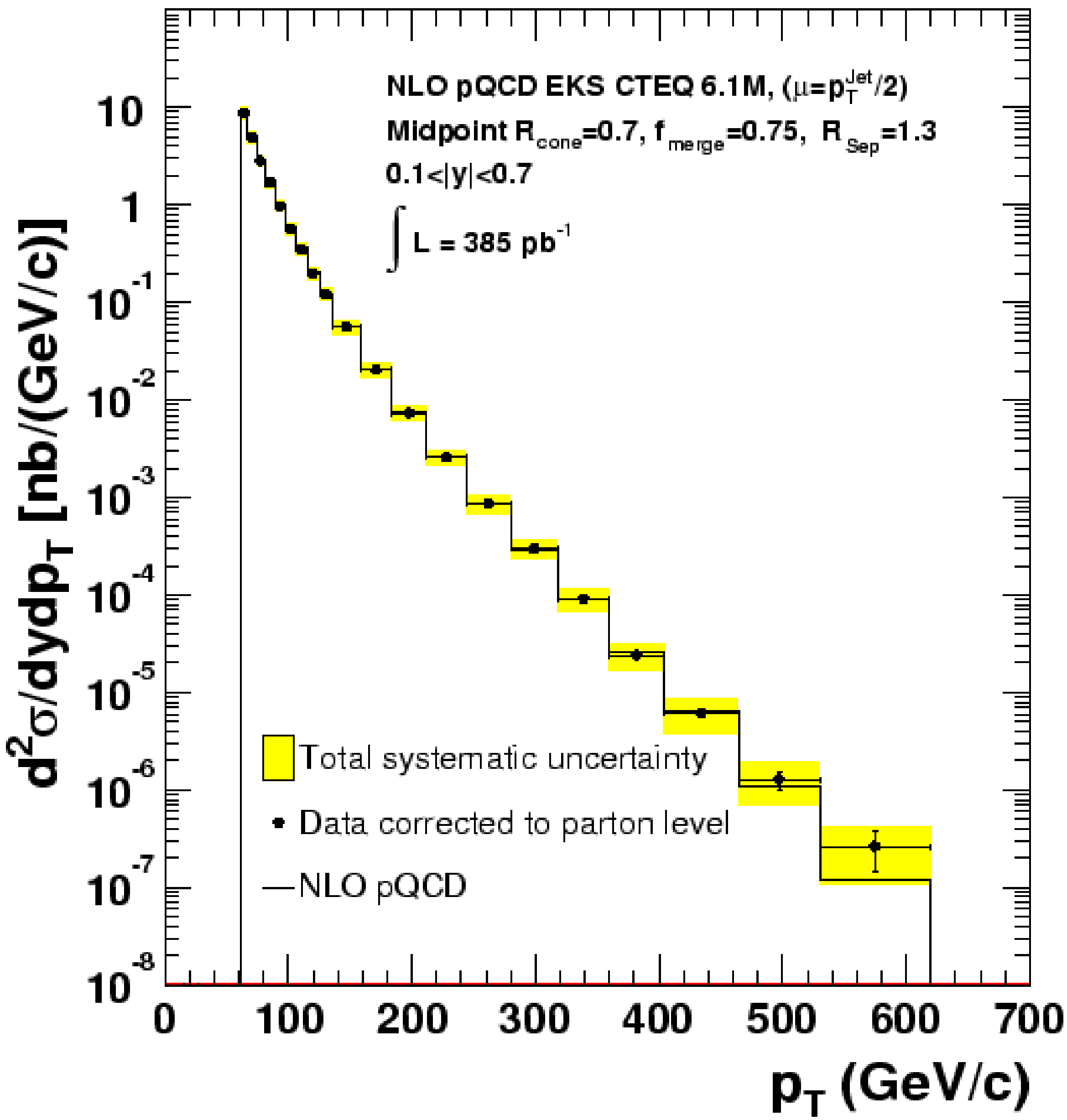
Definition: $$\frac{d\sigma}{dx} = \lim_{\Delta x \to 0} \frac{\Delta\sigma}{\Delta x}$$ such that $$\int_{-\infty}^{\infty} \frac{d\sigma}{dx}\, dx = \sigma_{tot}$$

… measure a differential cross section at a single fixed point in x by creating a bin of width Δx about that point …
… modification to cross section formula …

$$\frac{d\sigma}{dx} = \frac{N_{obs} - N_{b/g}}{L \,.\, A_{cc} \,.\, \Delta x}$$

… where Δx is the width of our bin in x
(still have to worry about exactly which x value we measured at)

An Example Differential Cross Section



In this example, a jet cross section is measured differentially in the jet transverse momentum p_T and the jet rapidity y

… note the horizontal error bars indicating that the data point could correspond to any point in the bin … (pessamistic!!!)

`Generalised' or `Smeared' Acceptance

Recall: $$\sigma = \frac{N_{obs} - N_{b/g}}{L\,.\varepsilon\,.\,A_{cc}}$$

Introduce smeared acceptance A_{cc} (calculated from Monte Carlo), which corrects for 'everything' in binned cross section measurements:

1. Finite selection efficiency / acceptance: some events end up in no bin at all.

2. Finite resolution: events end up in wrong bin (migrations).

Calculating A_{cc} with an MC can be very easy!..

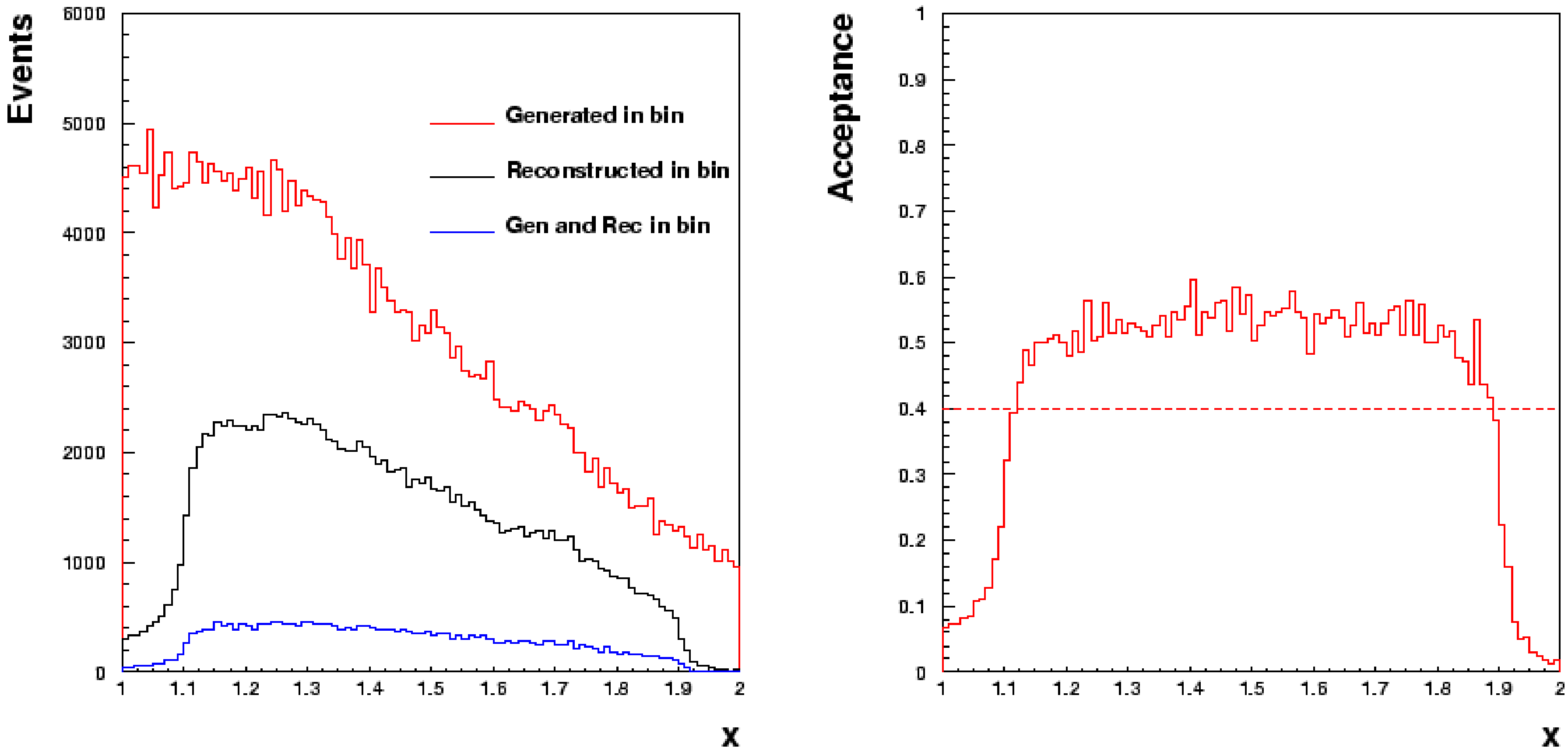
$\boxed{A_{cc}=N_{rec}/N_{gen}}$

N_{rec} = # Events reconstructed (found) in a bin (wherever they were generated)

N_{gen} = # Events generated in a bin

This simple treatment is OK if resolutions are well behaved (Gaussian) and bin choices are sensible (migrations not too large) … otherwise more complex 'unfolding' needed.

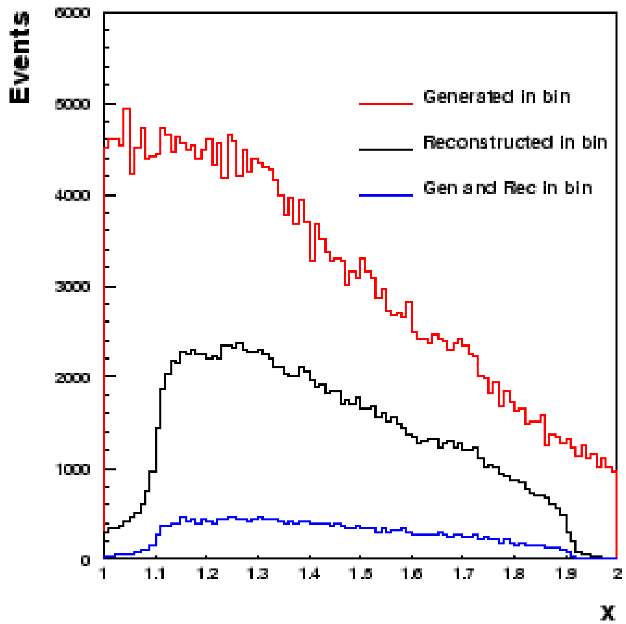
A Real Acceptance Example

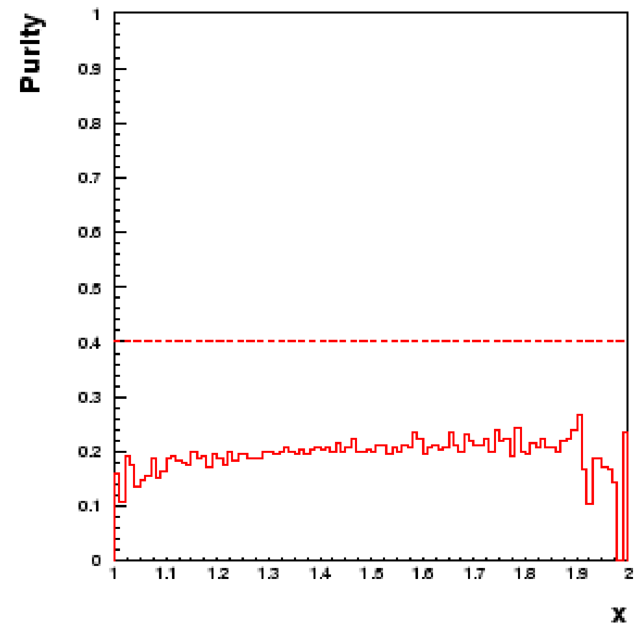


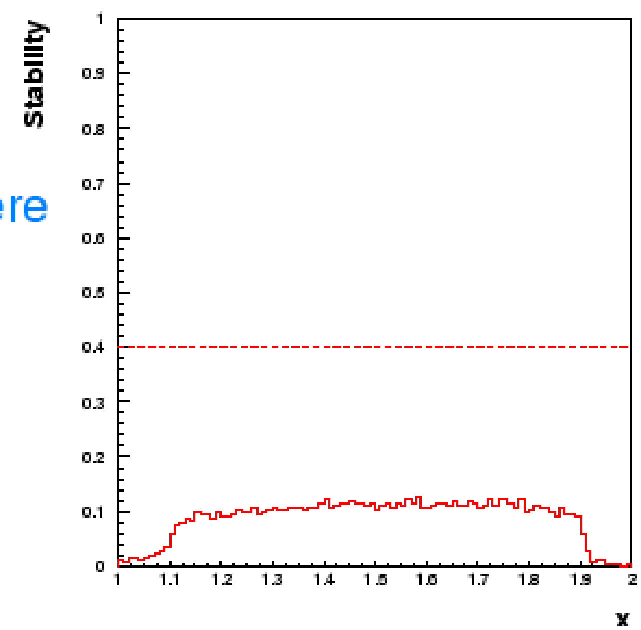
Here acceptance $\sim 50\%$ for $1.1 \lesssim x \lesssim 1.9 \ldots \rightsquigarrow$ can make measurement

$\ldots$ but is acceptance the full story? $\ldots$ here very few events are reconstructed and generated in same bin $\ldots$ we need many fewer bins!

Bin Purity and Stability







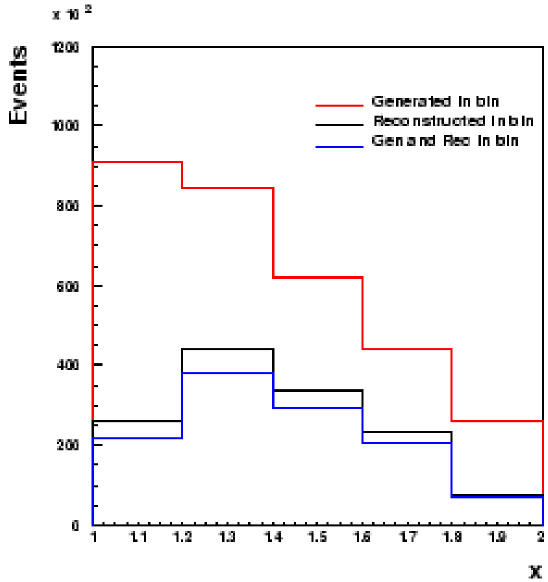
Purity $= \frac{\text{gen \& rec in bin}}{\text{rec in bin}}$ . . . tells us what fraction of events reconstructed in a bin belong there

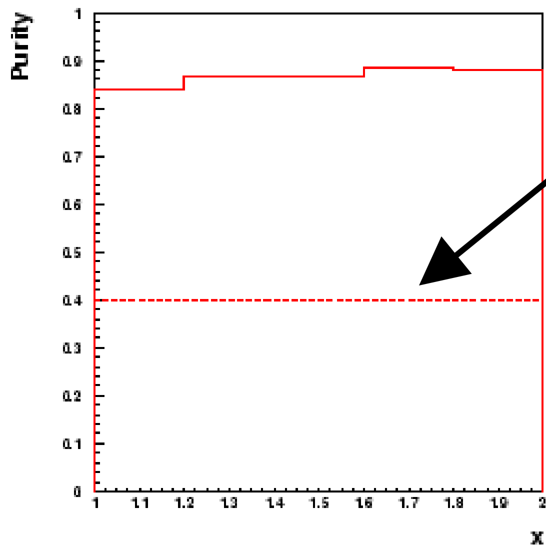
Stability $= \frac{\text{gen \& rec in bin}}{\text{gen in bin}}$ . . . tells us what fraction of events generated in a bin stay there

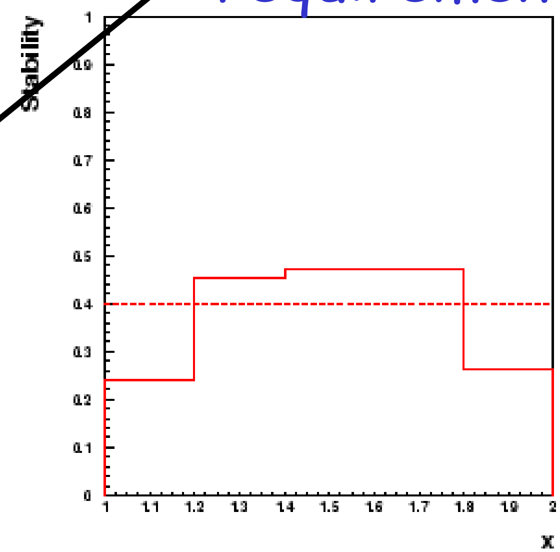
. . . Acceptance $= \frac{\text{Stability}}{\text{Purity}}$

. . . Same data, bigger bin sizes

40% is a typical requirement





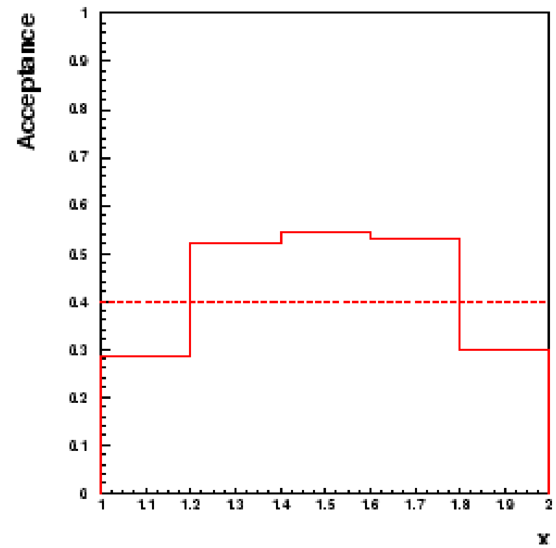


With bigger bin sizes, purity and stability improve whereas acceptance is unchanged.

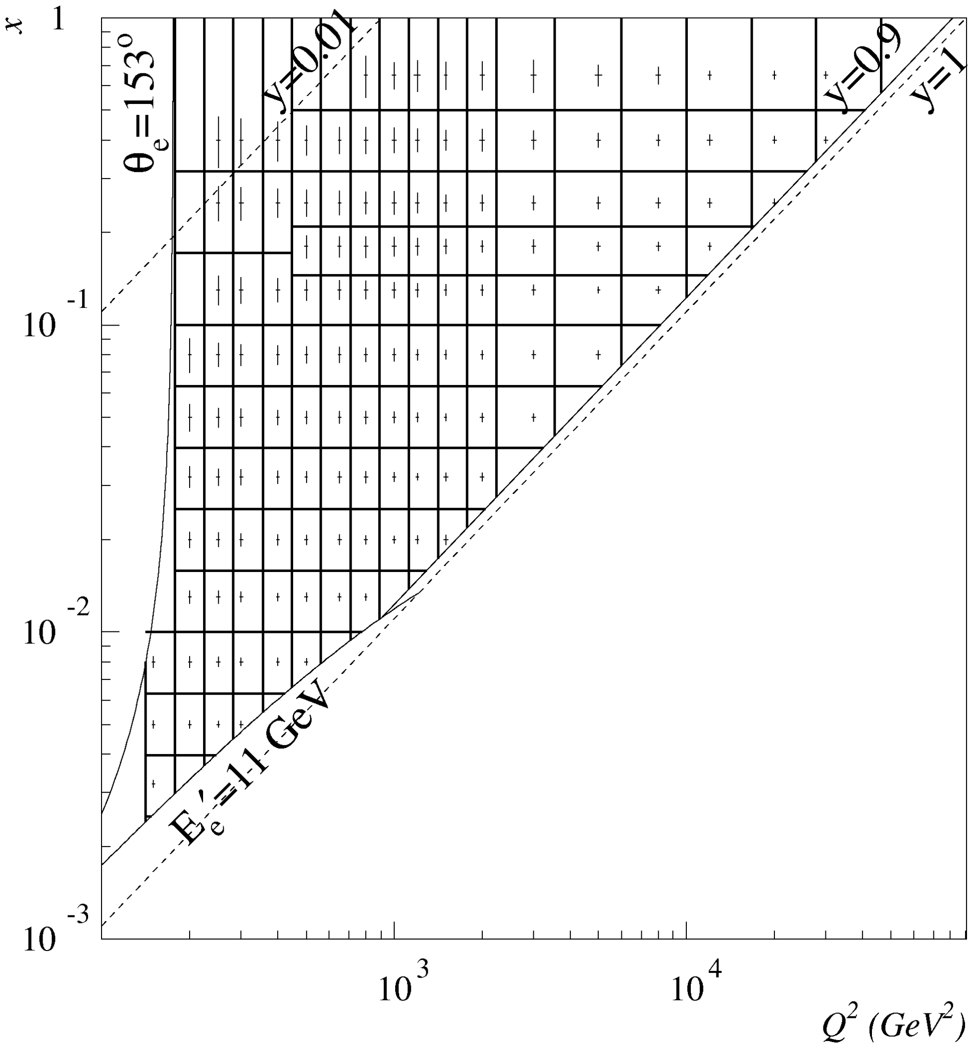
Measurement region is determined by acceptance
. . . must be high and stable across bin

Binning choices usually defined by purity
e.g. require purity > 0.68 . . . i.e. bin width > resolution

Exception: if data statistics are limited



e.g. Binning Choice from H1 (double differential)



Example from a 2 dimensional measurement (x and Q^2 in DIS)

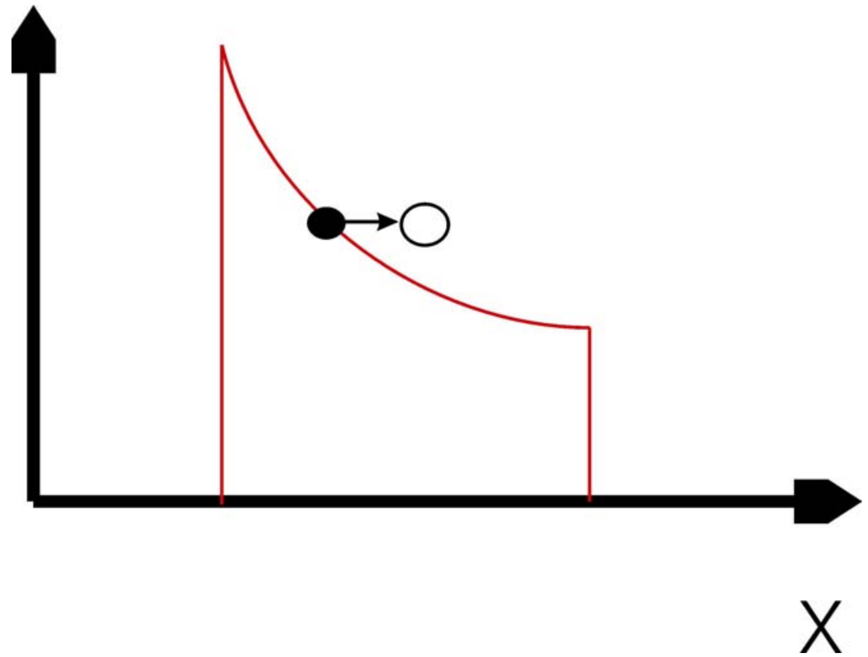
Crosses indicate resolution calculated in each bin, always smaller than bin size

Optimal measurement required variable bin sizes due to resolution and available statistics (decreasing as Q^2 increases)

All acceptances, purities and stabilities (as calculated from Monte Carlo models) are well above 40% …

More corrections with MC: Finite Bin Size Effects

• The cross section we measure is for the `centre of gravity' of the bin.
• For a cross section which falls fast, this can be very different from the cross section at the bin centre.
• Plotting at the bin centre without correction is a **very** common mistake
• One solution is to show horizontal error bars (as in Tevatron example)



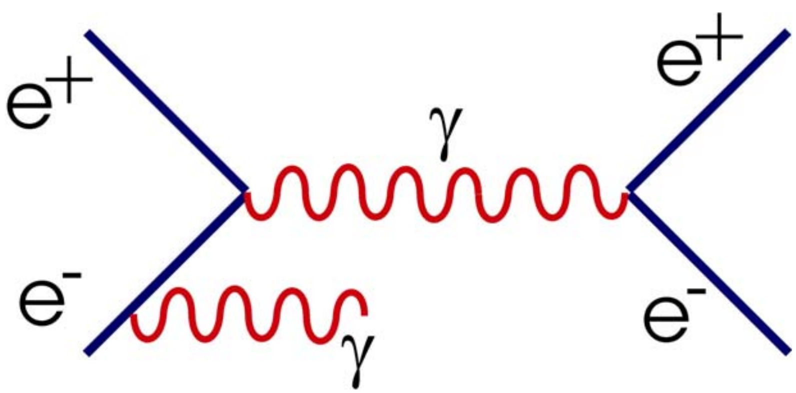
Better to plot at the bin centre, x_c, and make a **`bin centre correction'** …. back to the Monte Carlo again to calculate as …

$$\text{BCC} = \frac{\text{Cross section at bin centre}}{\text{Cross section averaged over bin}}$$

Using MC to deal with Radiative Effects

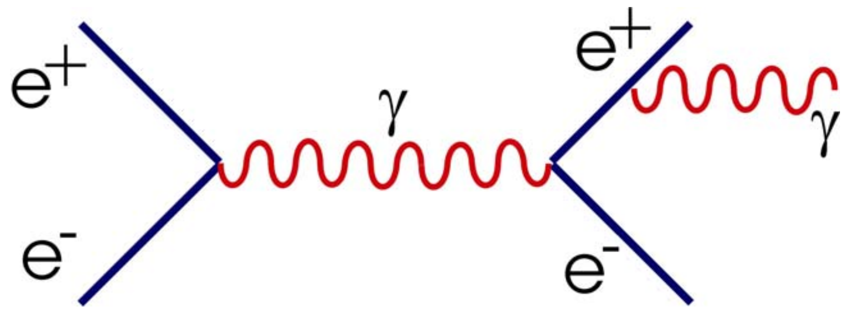
- For many measurements we want to correct for higher order QED effects such as initial state photon radiation.
- Whether we do this or not is a question of definition, but theorists probably don't want to have the radiative effects included.
- They want to work with a **BORN Level Cross Section**

1) Initial state radiation (ISR) effectively lowers beam energy of some events

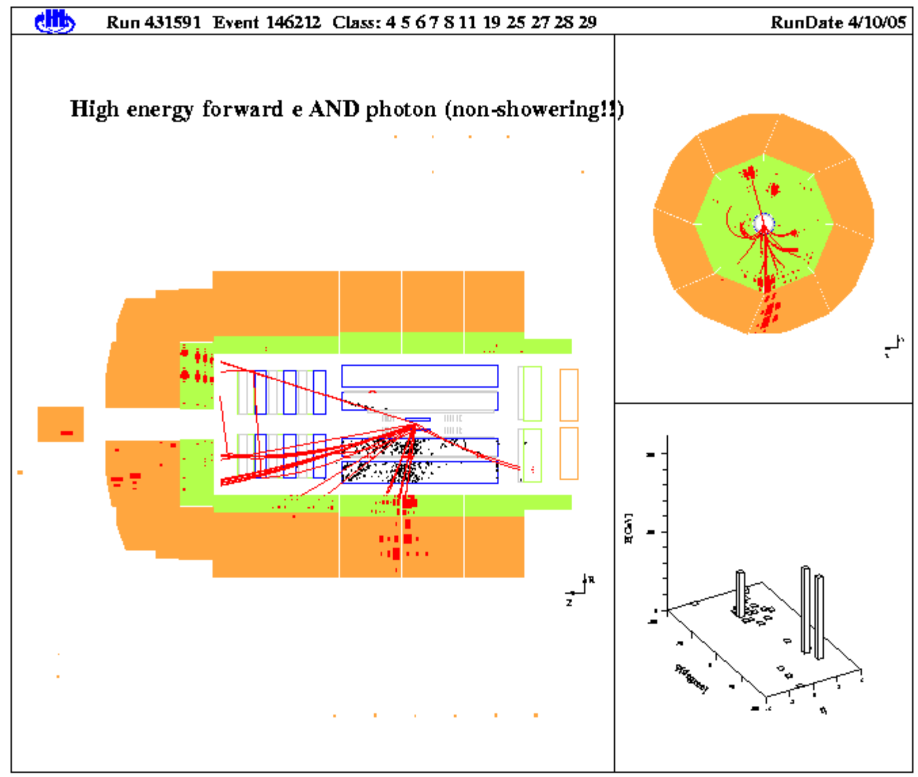


Photon emitted usually at very small angle to beam electron – sometimes detected e.g. in small angle calorimeter ….. but usually not

Radiative Effects II: Final State Radiation



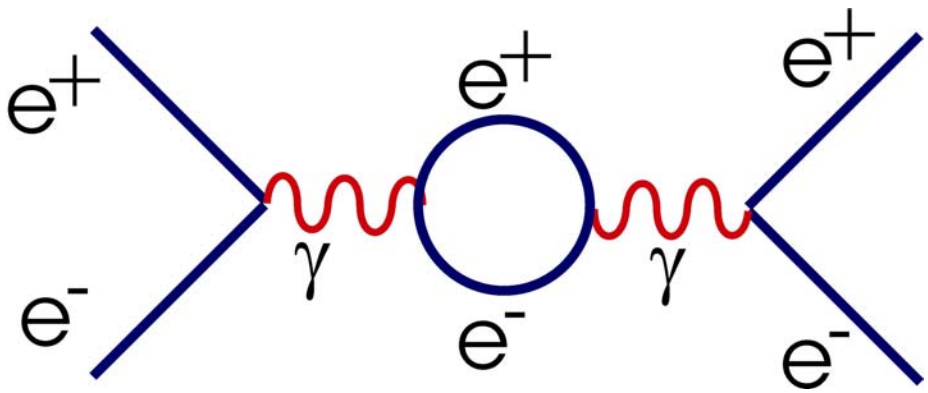
ep example



• As for ISR, final state photon radiation is usually nearly collinear with radiating electron – in fact you will usually reconstruct it in the same calorimeter cluster – in which case no correction needed!

• Still need small corrections for wide angle emissions

Radiative Effects III



Third effects is `virtual loop corrections' – effectively just a change in photon coupling

Monte Carlo models to the rescue again, as long as it has an interface to a reliable QED model! Correct for all 3 effects with a factor …..

$$\frac{\text{Predicted cross section with ISR, FSR and loops switched off in MC}}{\text{Predicted cross section with ISR, FSR and loops simulated in MC}}$$

Radiative corrections can be large e.g. if measurement involves photons or is very sensitive to the beam energy

What about the Uncertainties?

• We now have all ingredients needed to obtain the basic result for a (binned) differential cross section ☺

• However, this is not really a '*measurement*' until we also assess the errors … which is usually most of the work ☹

• Future lectures will deal with **statistical errors** (always calculable in principle, but often done incorrectly) and **systematic errors** (generally a poorly understood subject, even by professionals, but still almost[?] always quantifiable in principle

… first, some basic error theory, which applies to both …

Reminder of Mean, Standard Deviation ...

Mean / Average of a distribution

$\mu(x) = \langle x \rangle = \bar{x} = \sum_x p_x \cdot x$ ... or $\bar{x} = \int P(x) \cdot x dx$ if continuous distribution

Mean of a derived quantity

$\overline{f(x)} = \sum_x p_x \cdot f(x)$ e.g. If $f(x) = 2x$, $\overline{f(x)} = \sum_x p_x \cdot 2x = 2\bar{x}$

Variance V

Average squared deviation from mean used for spread (average deviation from mean is zero)

$V(x) = \overline{(x - \bar{x})^2} = \sum_x p_x (x - \bar{x})^2 = \overline{x^2} - (\bar{x})^2$

Standard Deviation

Often used in preference to Variance because it has the same units as x

$\sigma(x) = \sqrt{V(x)} = \sqrt{\frac{1}{N}\Sigma(x - \bar{x})^2}$

- For most purposes, assume that measurements follow a Gaussian distribution about the true value of a variable.
- By the `error' on a measurement, we mean 1σ (Gaussian)