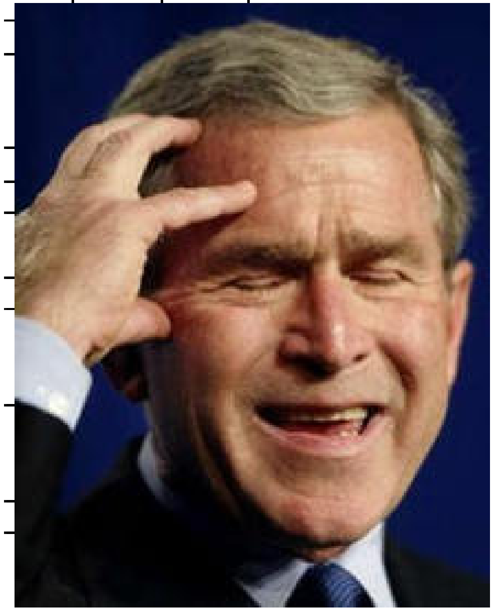
Errors: Propagation and Systematics

Last time:
- How to measure a (differential) cross section

Today's lecture:
- Experimental errors and how to propagate them
- Systematic Effects and Systematic Uncertainties

$x_{I\!P}$	Q^2 [GeV^2]	β	$x_{I\!P}\sigma_r^{D(3)}$	δ_{stat} [%]	δ_{sys} [%]	δ_{tot} [%]	δ_{unc} [%]	δ_{lar} [%]	δ_{ele} [%]	δ_{θ} [%]	δ_{noise} [%]	$\delta_{x_{I\!P}}$ [%]	δ_{β} [%]	δ_{bg} [%]	δ_{Plug} [%]	δ_{Q^2} [%]	δ_{spa} [%]
0.0003	3.5	0.17	0.0224	17.2	10.8	20.3	1.6	0.2	5.3	−1.7	−4.7	−0.6	−0.8	0.0	−0.0	3.0	1.7
0.0003	3.5	0.27	0.0262	7.6	6.4	9.9	1.6	0.2	1.3	2.7	−4.6	−0.4	−0.5	−0.1			
0.0003	3.5	0.43	0.0351	5.5	6.2	8.3	1.6	0.2	0.2	1.6	4.6	−1.6	−0.2	0.0			
0.0003	3.5	0.67	0.0443	5.5	11.7	13.0	1.6	0.3	−1.0	1.9	11.2	−0.3	0.7	−0.1			
0.0003	5.0	0.27	0.0392	11.9	7.7	14.2	1.6	0.4	3.8	1.4	4.7	−0.3	−0.8	0.0			
0.0003	5.0	0.43	0.0422	7.7	7.0	10.4	1.6	0.4	−0.8	1.5	4.9	−1.7	−0.1	0.0			
0.0003	5.0	0.67	0.0528	7.9	9.8	12.6	1.6	−0.2	−0.6	2.6	9.0	−0.2	0.7	0.0			
0.0003	6.5	0.43	0.0452	11.0	7.8	13.5	1.6	−0.8	−2.7	1.9	5.3	−1.2	0.1	0.0			
0.0003	6.5	0.67	0.0580	9.5	8.9	13.0	1.6	0.1	1.2	0.9	7.6	−0.8	0.9	0.0			
0.0003	8.5	0.43	0.0353	18.2	11.0	21.3	1.6	−0.5	6.5	1.3	5.5	−0.7	0.0	0.0			
0.0003	8.5	0.67	0.0570	11.7	9.5	15.0	1.6	0.2	−0.5	0.6	6.9	−1.6	0.8	0.0			
0.0003	12.0	0.67	0.0670	18.1	10.0	20.7	1.6	0.2	0.3	−1.4	8.5	−0.8	1.2	0.0			
0.001	3.5	0.05	0.0167	15.5	6.5	16.8	1.7	−0.7	1.0	−1.6	−2.6	−1.0	−2.2	0.0			
0.001	3.5	0.08	0.0189	8.2	6.9	10.7	1.7	0.5	2.5	1.8	−5.3	−0.9	−1.3	0.0			
0.001	3.5	0.13	0.0239	6.6	6.9	9.5	1.7	0.7	0.3	2.0	−5.8	−1.1	−1.2	0.0			
0.001	3.5	0.2	0.0239	6.2	6.9	9.3	1.7	0.5	0.7	0.7	−6.2	−1.2	−1.1	0.0			
0.001	3.5	0.32	0.0243	5.6	6.8	8.8	1.7	−0.1	−2.1	1.7	−5.7	−0.9	−0.9	0.0			
0.001	3.5	0.5	0.0281	5.5	4.5	7.1	1.7	0.2	−1.2	2.1	2.1	−0.2	−0.7	0.0			
0.001	3.5	0.8	0.0456	7.4	4.9	8.9	1.7	0.3	−1.4	1.3	3.0	1.5	0.8	−0.2			



Reminder of Mean, Standard Deviation …

Mean / Average of a distribution

$\mu(x) = \langle x \rangle = \quad \bar{x} = \sum_x p_x \cdot x$ … 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 $f(x) = 2x$, $\overline{f(x)} = \sum_x p_x \cdot 2x = 2\overline{x}$

Variance V

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

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

Standard Deviation

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

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

- 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)

Error Propagation

Single variable, linear function of x

Suppose $f(x) = ax + b$

...then $V(f(x)) = a^2 V(x)$

General function of a single variable

Can use a Taylor expansion to turn it into a linear function!

Expand about mean: $f(x) = f(x = \mu) + (x - \mu)\left(\frac{\partial f}{\partial x}\right)_{x=\mu}$

Then from linear case, $V(f(x)) = \left(\frac{\partial f}{\partial x}\right)^2 V(x)$

N.B. valid for small $x - \mu$ only! ...... 'tails' of distributions can cause problems!

Combination of two *independent* variables

Using the same arguments twice, $V(f(x,y)) = \left(\frac{\partial f}{\partial x}\right)^2 V(x) + \left(\frac{\partial f}{\partial y}\right)^2 V(y)$

...**independent errors add in quadrature**.

Always try to reduce problems to independent errors ... otherwise, there are 'covariance' terms

Example Error Propagation

Consider our 'standard' cross section formula

$$\sigma = \frac{N - N_{BG}}{\mathcal{L} \cdot A}$$

often, there will be statistical errors on all of N, N_{BG} and A

$$V(\sigma) = \left(\frac{\partial\sigma}{\partial N}\right)^2 V(N) + \left(\frac{\partial\sigma}{\partial N_{BG}}\right)^2 V(N_{BG}) + \left(\frac{\partial\sigma}{\partial A}\right)^2 V(A)$$

. . . where

$$\frac{\partial\sigma}{\partial N} = \frac{1}{\mathcal{L} \cdot A} \qquad \frac{\partial\sigma}{\partial N_{BG}} = \frac{-1}{\mathcal{L} \cdot A} \qquad \frac{\partial\sigma}{\partial A} = \frac{-(N - N_{BG})}{\mathcal{L} \cdot A^2}$$

$V(N) = N$,

$V(N_{BG}), V(A)$ come from MC statistical errors . . . see later.

This technique can be applied to all errors, stat or syst!

Systematics in the `Traditional' Physics Analysis

1. Devise cuts, get result
2. Do analysis for statistical errors
3. Make big table
4. Alter cuts by arbitrary amounts, put in table
5. Repeat step 4 until time/money exhausted
6. Add table in quadrature
7. Call this the systematic error
8. If challenged, describe it as 'conservative'

Systematics are not a precise science, but we can do better than this!

What is a Systematic Error?

Typing define: systematic error into google …

… nobody is really sure!

• `An error that results from a measurement method that is inherently wrong'

• `The component of the total error that is due to changes in the test method'

• `The difference of the mean of a series of measurements from the real value of the variable (which is usually unknown)'

• `A consistent error of the same size and sign due to a recurring cause'

Some of these definitions are just wrong. None really help!

Systematic Effects and Systematic Errors

Systematic effects are things that we try to identify and eliminate from our measurement.

Systematic errors are the uncertainties associated with this procedure … i.e. they are uncertainties, not mistakes!

A simple example of a systematic effect:
A steel rule is calibrated at 0^o C, but used in a warm lab…

- If not spotted, it is a mistake … Measurement is just wrong!
- If it is spotted and temperature is measured, we can make a correction…
- There is a systematic uncertainty associated with this correction – eg from uncertainty on temperature measurement and our knowledge of the expansion of steel on heating.

How to Evaluate Systematic Errors

• Everything measurement depends on (except raw number of events) is a potential source of systematic error

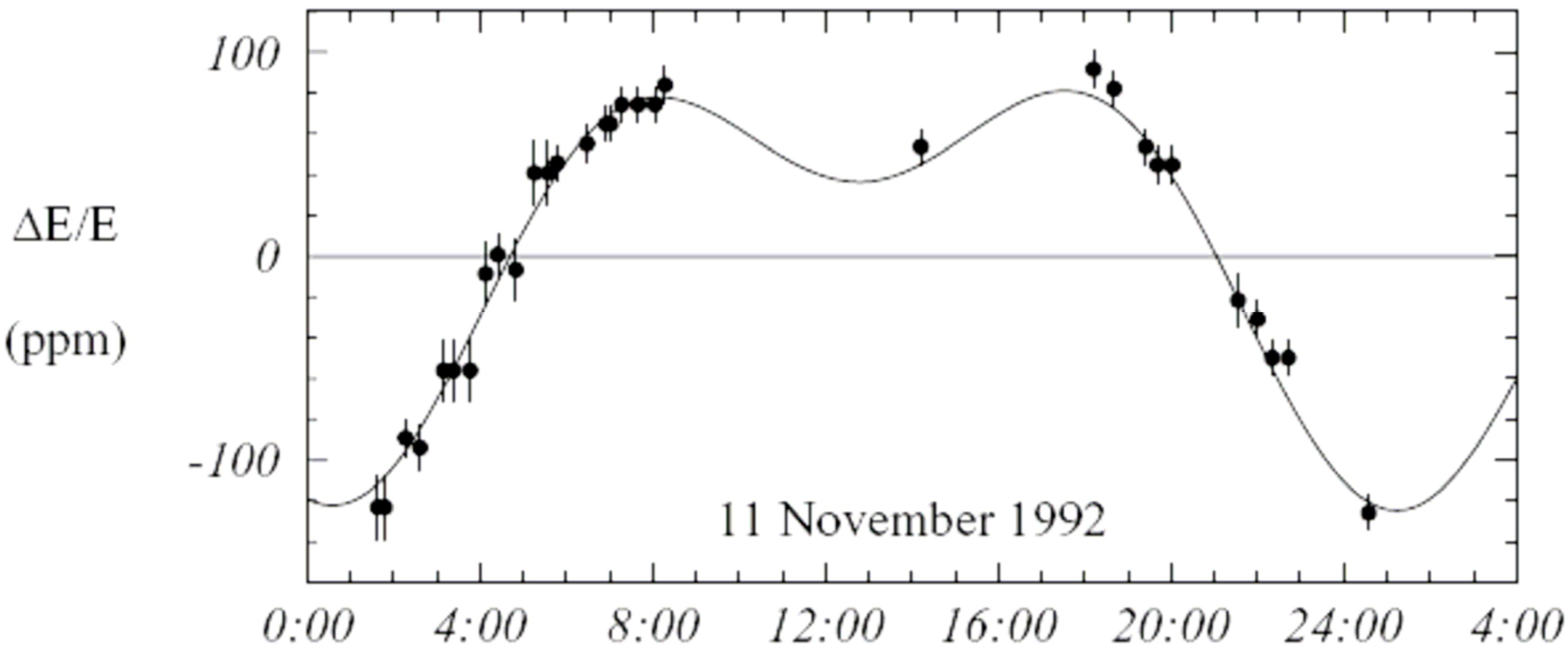
Some such uncertainties lead to repeated measurements being consistently biased (too high / low) in the same (unknown) way … e.g. the example with the steel rule … but not always!… **propagate** uncertainty through measurement to find out

• Common sense tells us which systematic errors we need to evaluate and which to neglect, based on the precision of the measurement … e.g. (usually!!!) …

- evaluate effect of possible calorimeter miscalibration
- ignore the effect of the phase of the moon

• Conservative syst. errors are better than no syst. errors
• 1σ syst. errors are best …. and can usually be evaluated!

In Case you thought I was joking ……



- Tidal effects caused variations in LEP circumference by 1mm
- This had a ~100 MeV systematic effect on the beam energy, which was not recognised for some time!
- Ultimately, many LEP measurements took the phase of the moon into account (though rarely the largest syst. error!)

Typical Sources of Systematic Error

General principle:

- Write down all possible systematic effects (everything data depend on)
- Select those which are likely to lead to non-negligible uncertainties
- Determine **1σ uncertainty on your treatment of the effect**
- Apply this shift and repeat the analysis.
- The resulting shift in the result is your systematic error.
- Add all systematic errors in quadrature to get the total syst. error

Experimental sources:

e.g. calorimeter calibration uncertainty, angular alignment uncertainty, trigger efficiency uncertainty, luminosity uncertainty …

Theoretical sources:

e.g. model used for acceptance corrections or background subtraction, input parameters such as a particle mass or branching ratio

Example 1: A Branching Ratio Uncertainty

Suppose we are measuring a cross section for J/ψ production
by detecting muons in the decay channel J/ψ -> μ+ μ-

Need to correct for the branching fraction … BR = Prob(J/ψ -> μ+ μ-)

Our cross section formula is modified to …

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

From particle data book, summarising current status of world knowledge, BR = 5.9 +/- 0.1%   (a 1 standard deviation error)

So evaluate central value for cross section using BR=0.059
To get 1σ systematic error, use BR=0.060 and BR=0.058.

Alternatively, we could use error propagation theory (see last time).

Example 2: A Model Dependence Uncertainty

For the same J/ψ measurement, we will also need to estimate the uncertainty on the calculated acceptance $A_{cc.}$

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

A_{cc} usually comes from a Monte Carlo simulation, which contains many approximations, phenomenological models and has probably been tweaked to match our data.

Standard Approach:
Get another Monte Carlo model with different approximations etc and hope that the difference between the two reflects the error!

Take $Acc = Acc_1 \pm |Acc_1 - Acc_2|$ if you prefer model 1

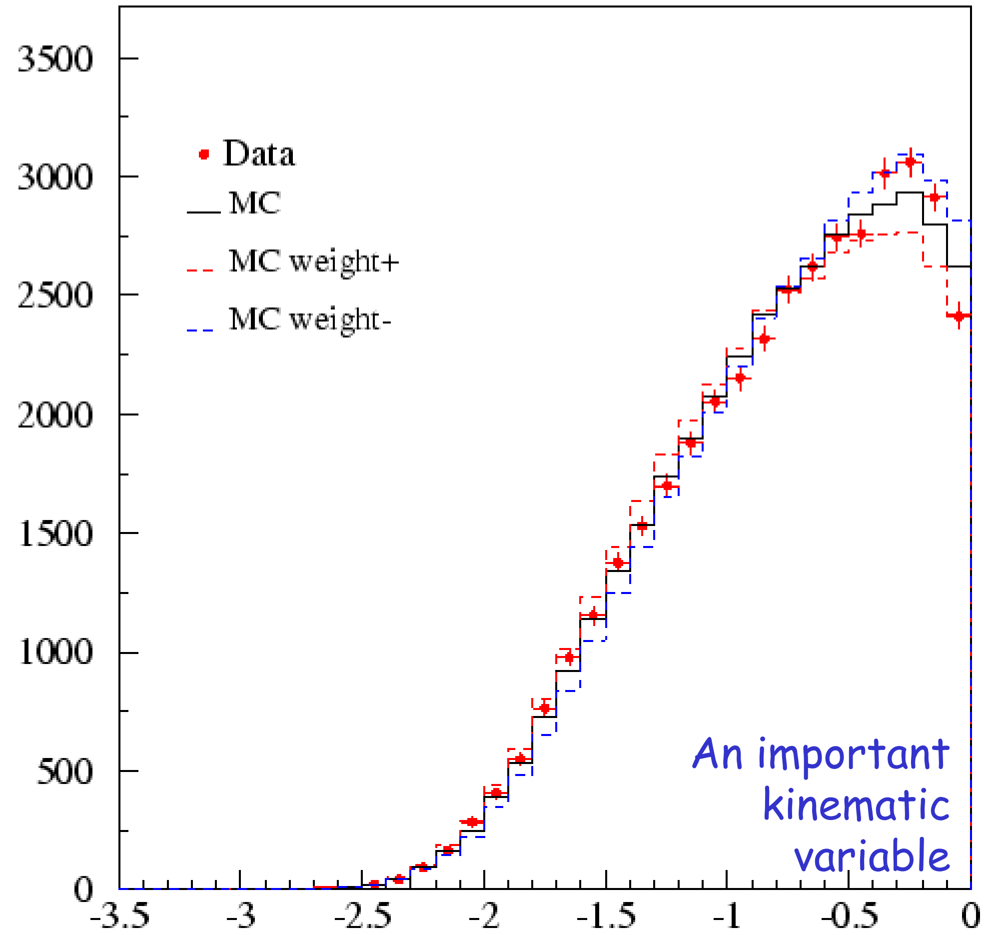
Take $Acc = \frac{1}{2}(Acc_1 + Acc_2) \pm |Acc_1 - Acc_2|/\sqrt{2}$ if they are equally rated

There are other (better?) ways

Better?… Model Dependence Error with 1 Model

Using just one MC model …

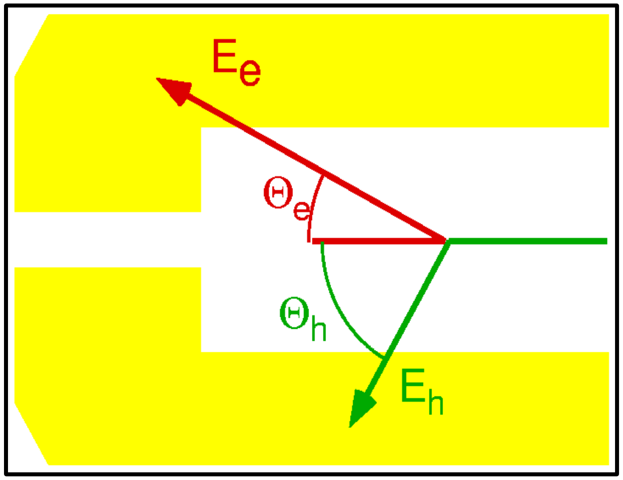
• Modify important distributions at the generator level by applying weights to each event depending on value of that variable.



• Choose weights such that after simulation, detector … control distributions are (just) still described.

• Look at changes to acceptance (and hence cross section) resulting from the reweights to get your systematic error.

Example 3: A detector calibration uncertainty



(H1, eq → eh)

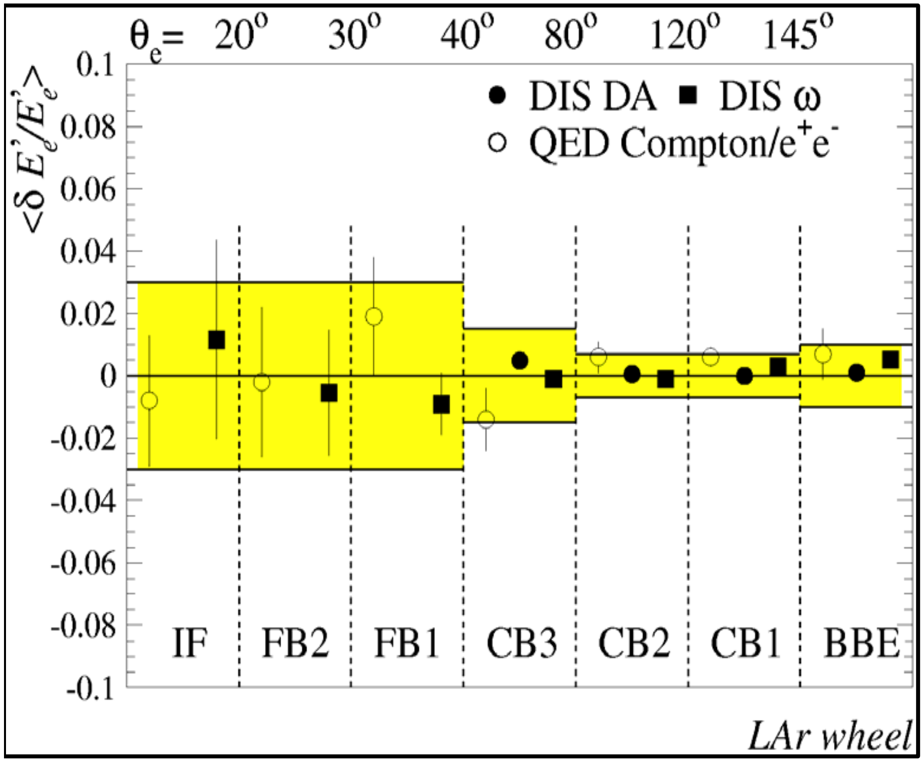
- We have 4 basic measurements (electron and hadron energies & angles)
- Only 2 of these are independent (2 degrees of freedom, corresponding to x and Q^2)

- We can predict E_e using any two of the other variables!

- EM calorimeter calibration done by comparing measured energy E_e with the energy predicted on the basis of the θ_e and θ_h measurements … the `double angle' (DA) method

- This calibration procedure gets rid of most of **systematic effect** on the E_e measurement due to detector understanding, but we still need to estimate the systematic error?

Error on EM Calorimeter calibration

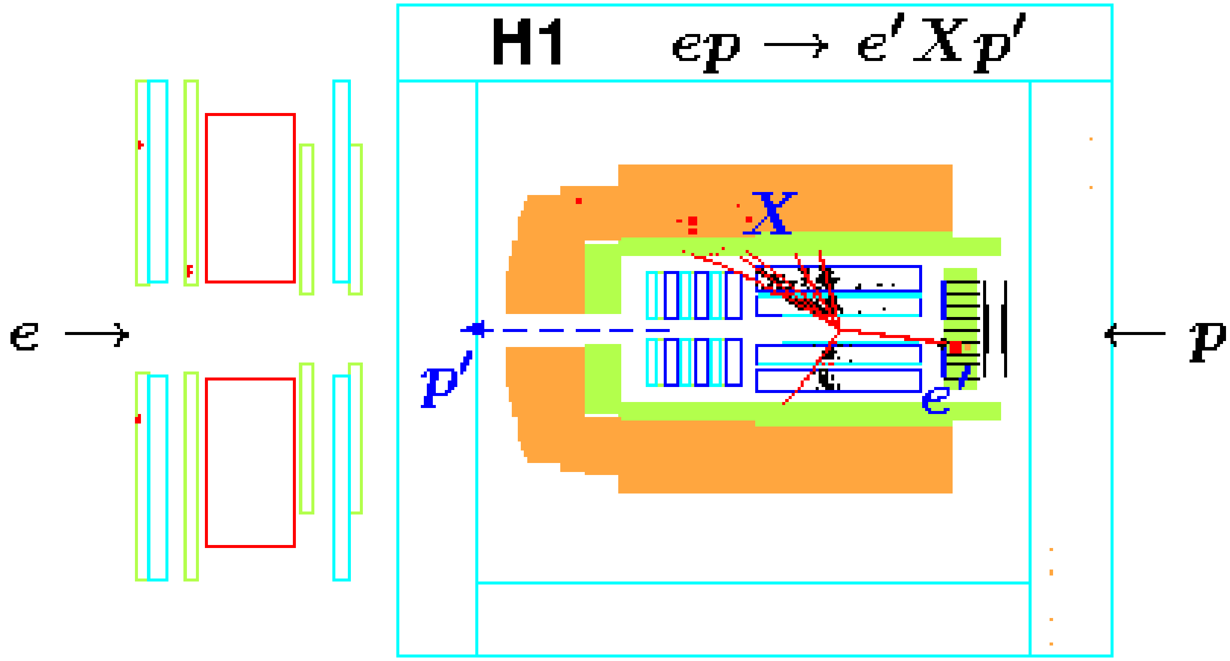


- To estimate size of **syst. uncertainty**, compare with calibration using other methods.
- Here the 0 of $\delta E/E$ is at default DA calibration point
- 'DIS DA' = double angle with a different sample, 'DIS ω' and 'QED Compton' are completely different methods.

- Methods agree within (roughly) yellow band, chosen by eye
- Yellow band → syst. uncertainty is 0.7 - 3.0% (θ dependent)
- Shift all electron energies in data (or MC) by this amount and repeat analysis to get **systematic error on measurement**
- This method relies on **redundancy** (unused measured info)

Systematics of a complete analysis ep-> eXp

Measurement of `diffractive DIS', when proton remains intact



- Measure x, Q^2 using scattered electron energy & angle
- Measure mass M_x of X using all observed hadrons
- Be sure that scattered proton is really a proton by demanding a signal in tagger inserted into the beampipe downstream (not shown)
- Need to estimate the systematic errors on $\sigma(x,Q^2,M_x)$?

Main Systematic Errors for Diffractive DIS

- Electromagnetic energy scale (scattered electron energy)
- Electromagnetic calorimeter alignment (electron angle)
- Hadronic energy scale (M_x measurement)
- Calorimeter noise subtraction (M_x measurement)
- Beampipe tagger efficiency (scattered proton intact)
- Model dependence of acceptance correction (several!)
- Model dependence of background subtraction (several!)
- Luminosity measurement
- Trigger efficiency
- Linking efficiency of charged track to electron
- …

Note that some errors (e.g. lumi, branching ratio) affect all data points equally, whilst others vary from point to point!

Critique of ` The traditional physics analysis'

1. Devise cuts, get result
2. Do analysis for statistical errors
3. Make big table
4. Alter cuts by arbitrary amounts, put in table
5. Repeat step 4 until time/money exhausted
6. Add table in quadrature
7. Call this the systematic error
8. If challenged, describe it as 'conservative'

This is not *completely* stupid - should always vary everything we can think of to check for mistakes and make sure our measurement is **robust** ..... but it is essential to distinguish between a **systematic *check*** and a **systematic *error***!