Lecture 9:

Useful Tools for Experimental Design

- Effective Contributions to Uncertainties and "Pulls" Analysis
- Blind Analysis
- Bifurcated Side-Band Analysis
- Statistical Optimisation
- A Note on Redundancy & Calibration

Separating Contributions of Systematic Uncertainties

Systematic uncertainties are often handled by "floating" them as free or constrained (priors!) nuisance parameters within the likelihood fit that are then marginalised over when extracting the parameters of interest. But we also want to make clear the separate contributions from systematic and statistical uncertainties due to their different natures (lecture 5).

We can assess the impact of statistical uncertainties alone by simply fixing the systematic nuisance parameters to their nominal values and measuring the shape of the likelihood. This can then be compared to the likelihood with systematics floating to determine their impact.

It is often useful to show this in terms of the equivalent 1-sigma Gaussian uncertainties:

$\sigma_{tot} \equiv$ total equiv. Gaussian uncertainty with floated systematics

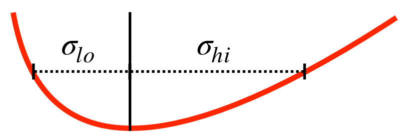
$\sigma_{stat} \equiv$ equiv. Gaussian uncertainty with fixed systematics

Then treating these as if we had independent Gaussian uncertainties:

$$\sigma_{tot}^2 = \sigma_{stat}^2 + \sigma_{sys}^2 \quad \rightarrow \quad \sigma_{sys}^2 = \sigma_{tot}^2 - \sigma_{stat}^2$$

$$\rightarrow \quad X \pm \sigma_{stat} \pm \sigma_{sys}$$

where X is the result obtained from the combined fit



More typically, the likelihood will not have a symmetrical shape in the region of the maximum. In this case, a better approximation can be obtained by quoting different upper and lower Gaussian equivalent uncertainties using the same approach:

$$X \;{}^{+\sigma_{hi}}_{-\sigma_{lo}}\,(stat)\;{}^{+\sigma_{hi}}_{-\sigma_{lo}}\,(sys)$$

“Pulls” Analysis

More generally, the result itself may well have shifted as a result of propagating the systematic uncertainties if it appreciably alters the shape of the likelihood. The significance of the shift in terms of ‘standard deviations’ due to systematic uncertainties can be quantified by defining the systematic “pull”:

$$g_{sys} \equiv \frac{X(total) - X(stat\ only)}{\sigma_{sys}}$$

$$= \frac{X(total) - X(stat\ only)}{\sqrt{\sigma_{tot}^2 - \sigma_{stat}^2}}$$

← difference in the determined parameter of interest due to the inclusion of floating systematics

← appropriate (‘hi’ or ‘lo’) equivalent Gaussian contribution to the total uncertainty.

Pulls can be separately assessed for individual systematics to show their impact and check for consistency.

Pulls Decomposition in a Constrained Fit

Assume we have made a measurement of some quantity, $x_m \pm \sigma_m$, that has been combined with an independent constraint, $x_c \pm \sigma_c$ (perhaps from a calibration or a separate measurement etc.), to obtain an improved fit estimate of $x_f \pm \sigma_f$

For Gaussian uncertainties:

$$x_f = \frac{x_m w_m + x_c w_c}{w_m + w_c} \qquad where \quad w_m = \frac{1}{\sigma_m^2}, \; w_c = \frac{1}{\sigma_c^2}$$

$$\sigma_f^2 = \sigma_m^2 \left(\frac{w_m}{w_m + w_c} \right)^2 + \sigma_c^2 \left(\frac{w_c}{w_m + w_c} \right)^2$$

$$= \sigma_m^2 \left(\frac{1/\sigma_m^2}{1/\sigma_m^2 + 1/\sigma_c^2} \right)^2 + \sigma_c^2 \left(\frac{1/\sigma_c^2}{1/\sigma_m^2 + 1/\sigma_c^2} \right)^2$$

$$= \frac{1/\sigma_m^2 + 1/\sigma_c^2}{\left(1/\sigma_m^2 + 1/\sigma_c^2\right)^2} = \frac{1}{\left(1/\sigma_m^2 + 1/\sigma_c^2\right)} = \frac{\sigma_m^2 \sigma_c^2}{\sigma_m^2 + \sigma_c^2}$$

$$\sigma_m^2 = \frac{\sigma_f^2 \sigma_c^2}{(\sigma_c^2 - \sigma_f^2)}$$

Pulls Decomposition in a Constrained Fit

Assume we have made a measurement of some quantity, $x_m \pm \sigma_m$, that has been combined with an independent constraint, $x_c \pm \sigma_c$ (perhaps from a calibration or a separate measurement etc.), to obtain an improved fit estimate of $x_f \pm \sigma_f$

For Gaussian uncertainties:

$$x_f = \frac{x_m w_m + x_c w_c}{w_m + w_c} \qquad x_f - x_c = \frac{x_m w_m + x_c w_m}{w_m + w_c}$$

$$\sigma_{fc}^2 = \sigma_m^2 \left(\frac{w_m}{w_m + w_c} \right)^2 + \sigma_c^2 \left(\frac{w_m}{w_m + w_c} \right)^2$$

$$= (\sigma_m^2 + \sigma_c^2) \left(\frac{1/\sigma_m^2}{1/\sigma_m^2 + 1/\sigma_c^2} \right)^2 = \frac{\sigma_c^4}{\sigma_m^2 + \sigma_c^2}$$

$$\sigma_m^2 = \frac{\sigma_f^2 \sigma_c^2}{(\sigma_c^2 - \sigma_f^2)}$$

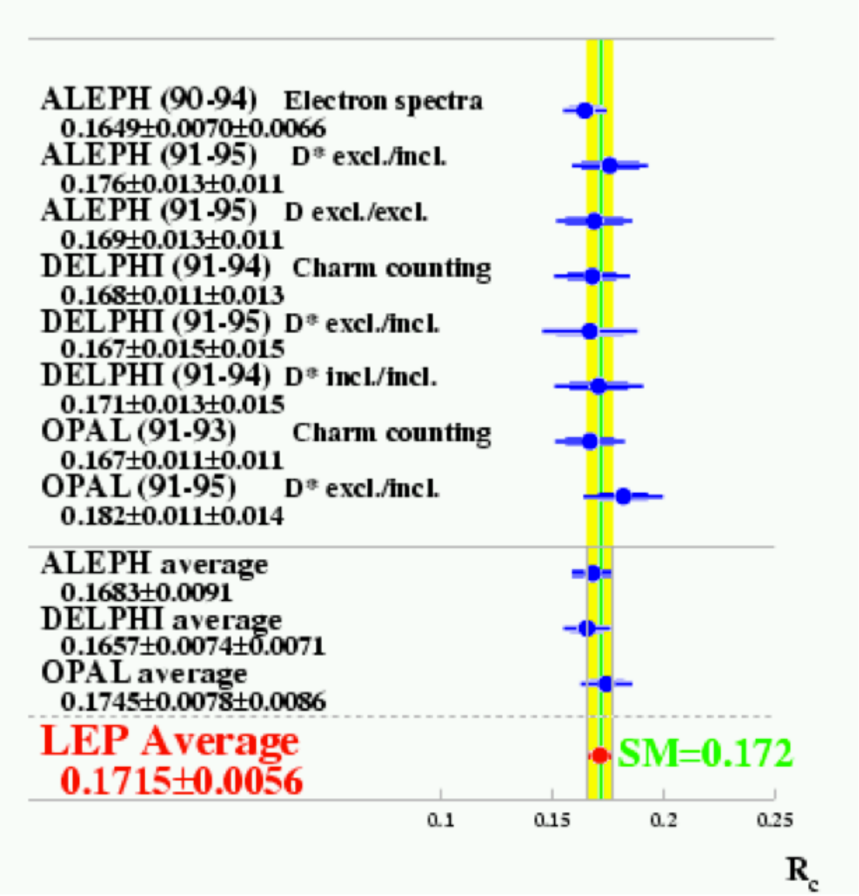
$$= \frac{\sigma_c^4 (\sigma_c^2 - \sigma_f^2)}{\sigma_f^2 \sigma_c^2 + \sigma_c^2 (\sigma_c^2 - \sigma_f^2)} = \frac{\sigma_c^4 (\sigma_c^2 - \sigma_f^2)}{\sigma_c^4} = \sigma_c^2 - \sigma_f^2$$

$$g_c = \frac{x_f - x_c}{\sqrt{\sigma_c^2 - \sigma_f^2}}$$

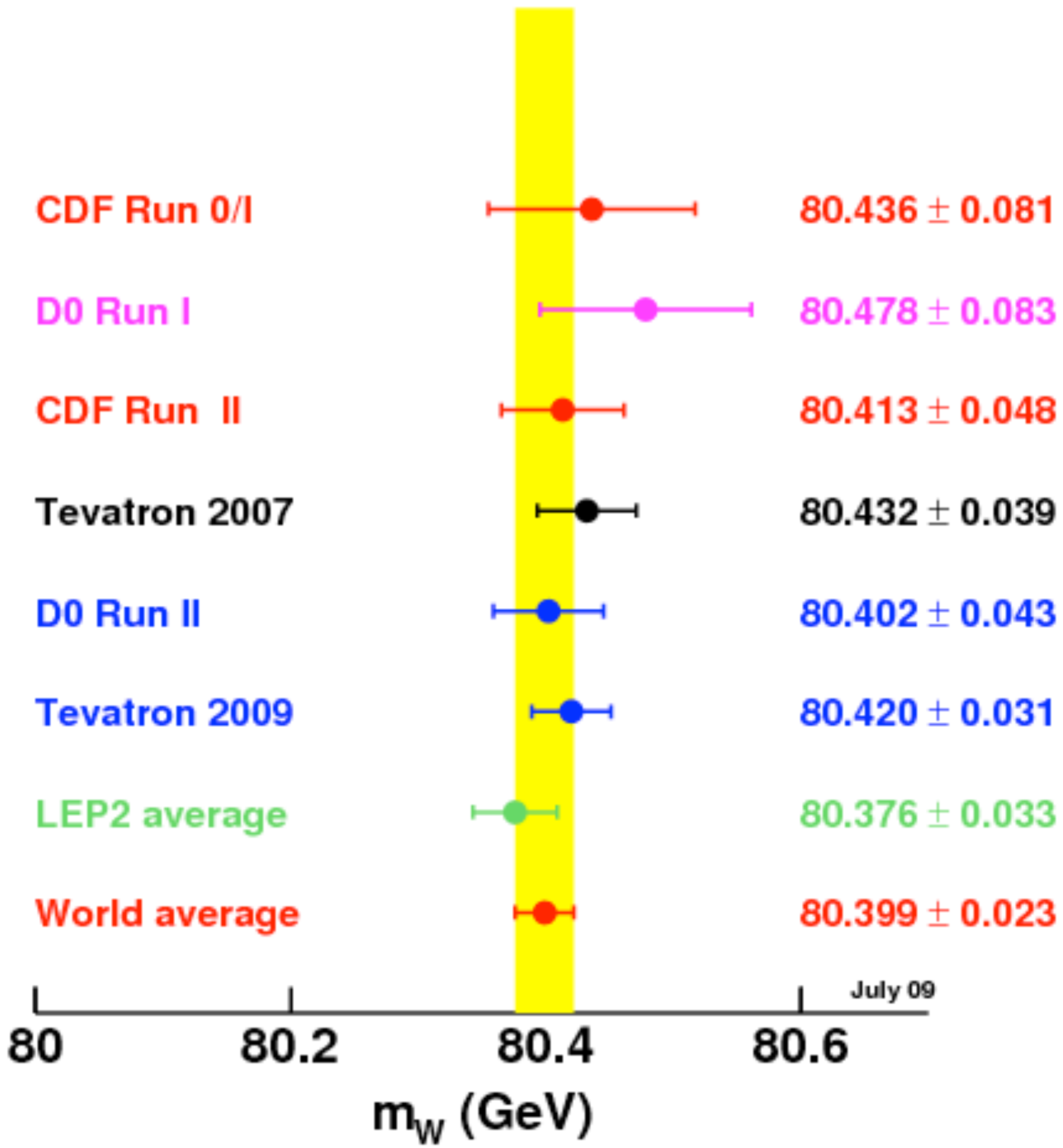
and similarly: $g_m = \dfrac{x_f - x_m}{\sqrt{\sigma_m^2 - \sigma_f^2}}$

So we can also separate the effective contributions of measurement and constraint to the final fit result

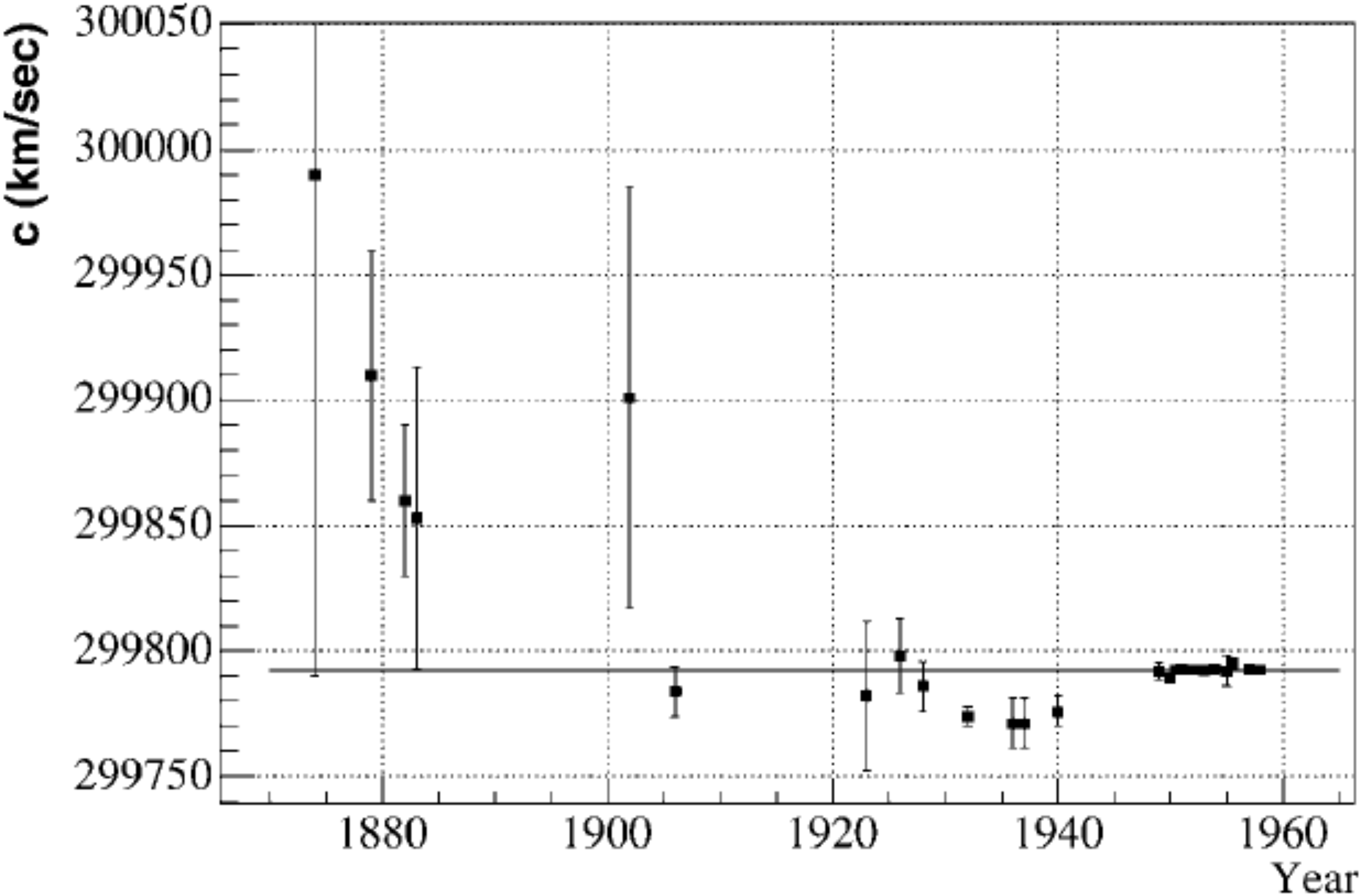

Blindness

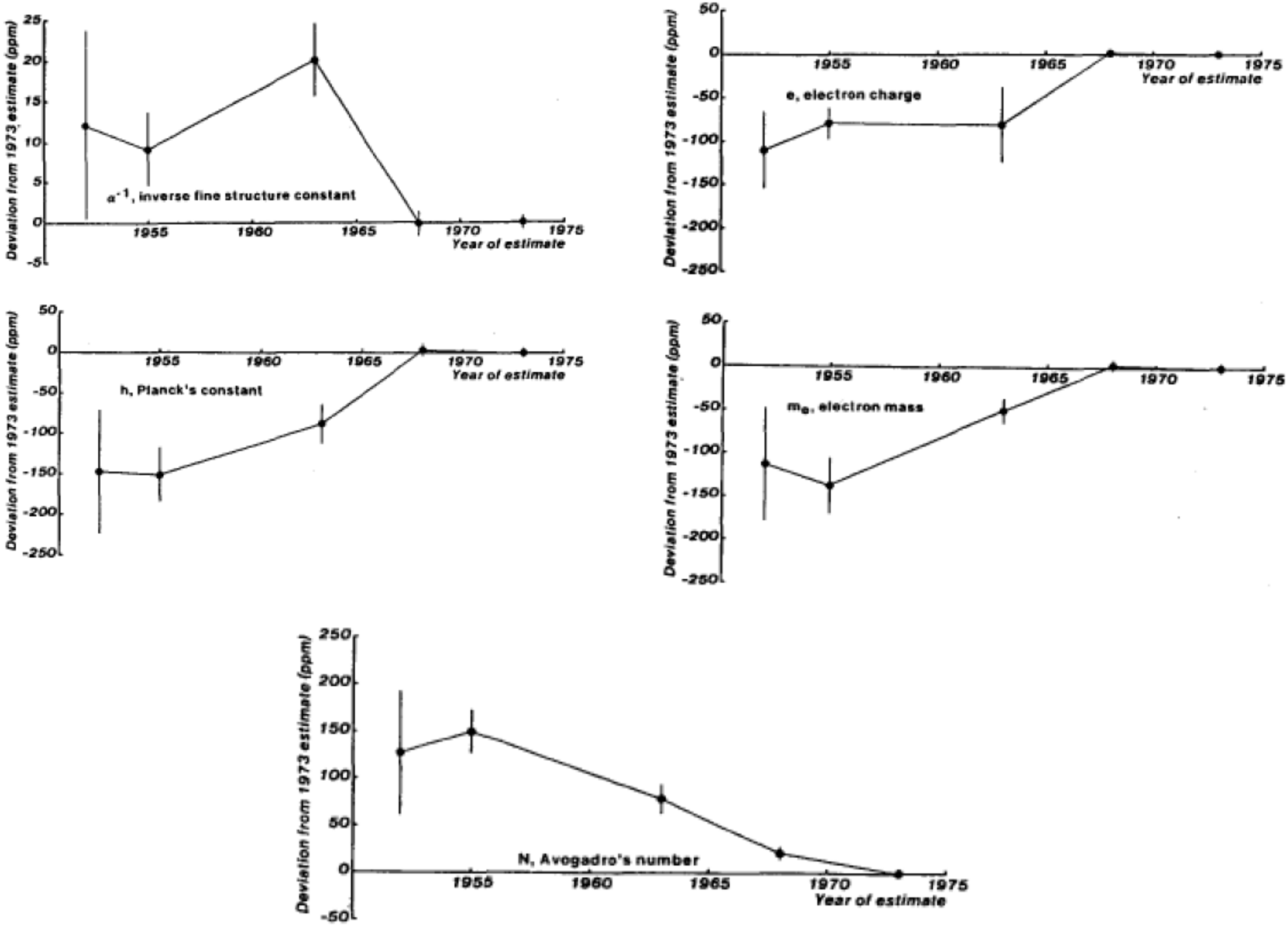


ALEPH (90-94) Electron spectra
0.1649±0.0070±0.0066
ALEPH (91-95) D* excl./incl.
0.176±0.013±0.011
ALEPH (91-95) D excl./excl.
0.169±0.013±0.011
DELPHI (91-94) Charm counting
0.168±0.011±0.013
DELPHI (91-95) D* excl./incl.
0.167±0.015±0.015
DELPHI (91-94) D* incl./incl.
0.171±0.013±0.015
OPAL (91-93) Charm counting
0.167±0.011±0.011
OPAL (91-95) D* excl./incl.
0.182±0.011±0.014
ALEPH average
0.1683±0.0091
DELPHI average
0.1657±0.0074±0.0071
OPAL average
0.1745±0.0078±0.0086
LEP Average
0.1715±0.0056
SM=0.172
0.1
0.15
0.2
0.25
R_c

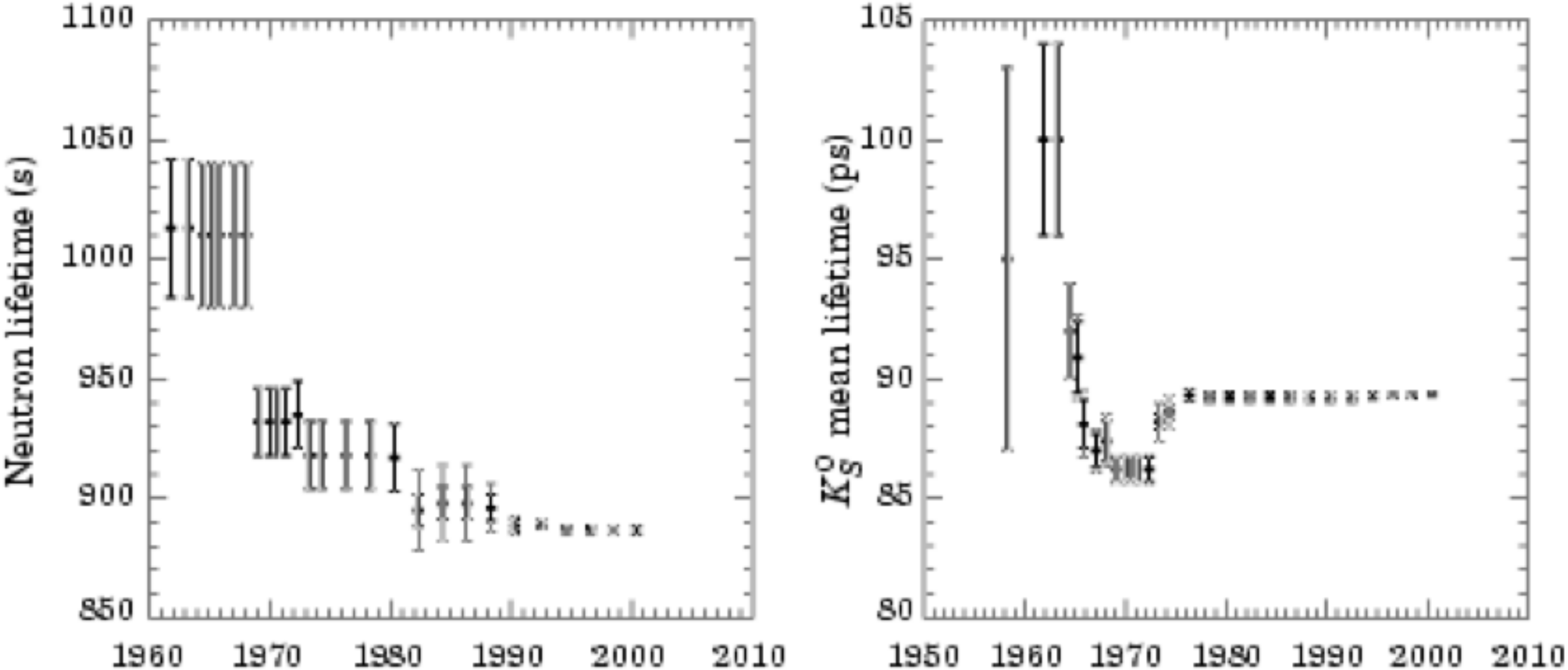


CDF Run 0/I
80.436 ± 0.081
D0 Run I
80.478 ± 0.083
CDF Run II
80.413 ± 0.048
Tevatron 2007
80.432 ± 0.039
D0 Run II
80.402 ± 0.043
Tevatron 2009
80.420 ± 0.031
LEP2 average
80.376 ± 0.033
World average
80.399 ± 0.023
July 09
80
80.2
80.4
80.6
m_W (GeV)


c (km/sec)
300050
300000
299950
299900
299850
299800
299750
1880
1900
1920
1940
1960
Year

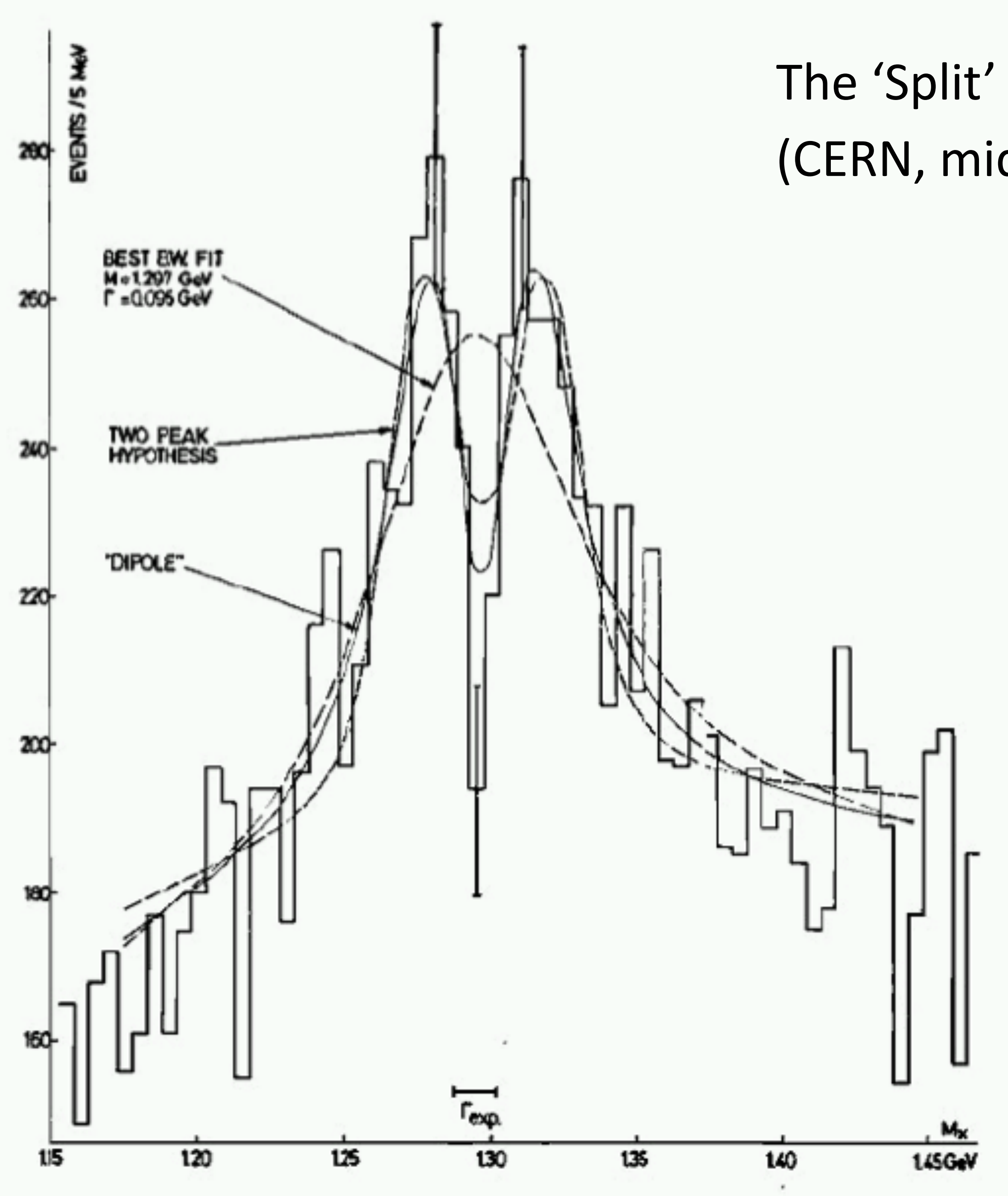


Deviation from 1973 estimate (ppm)
Year of estimate
α^{-1}, inverse fine structure constant
e, electron charge
h, Planck's constant
m_e, electron mass
N, Avogadro's number



Neutron lifetime (s)
1100
1050
1000
950
900
850
1960
1970
1980
1990
2000
2010
K_S^0 mean lifetime (ps)
105
100
95
90
85
80
1950
1960
1970
1980
1990
2000
2010

The ‘Split’ A_2 Meson (CERN, mid 1960’s)

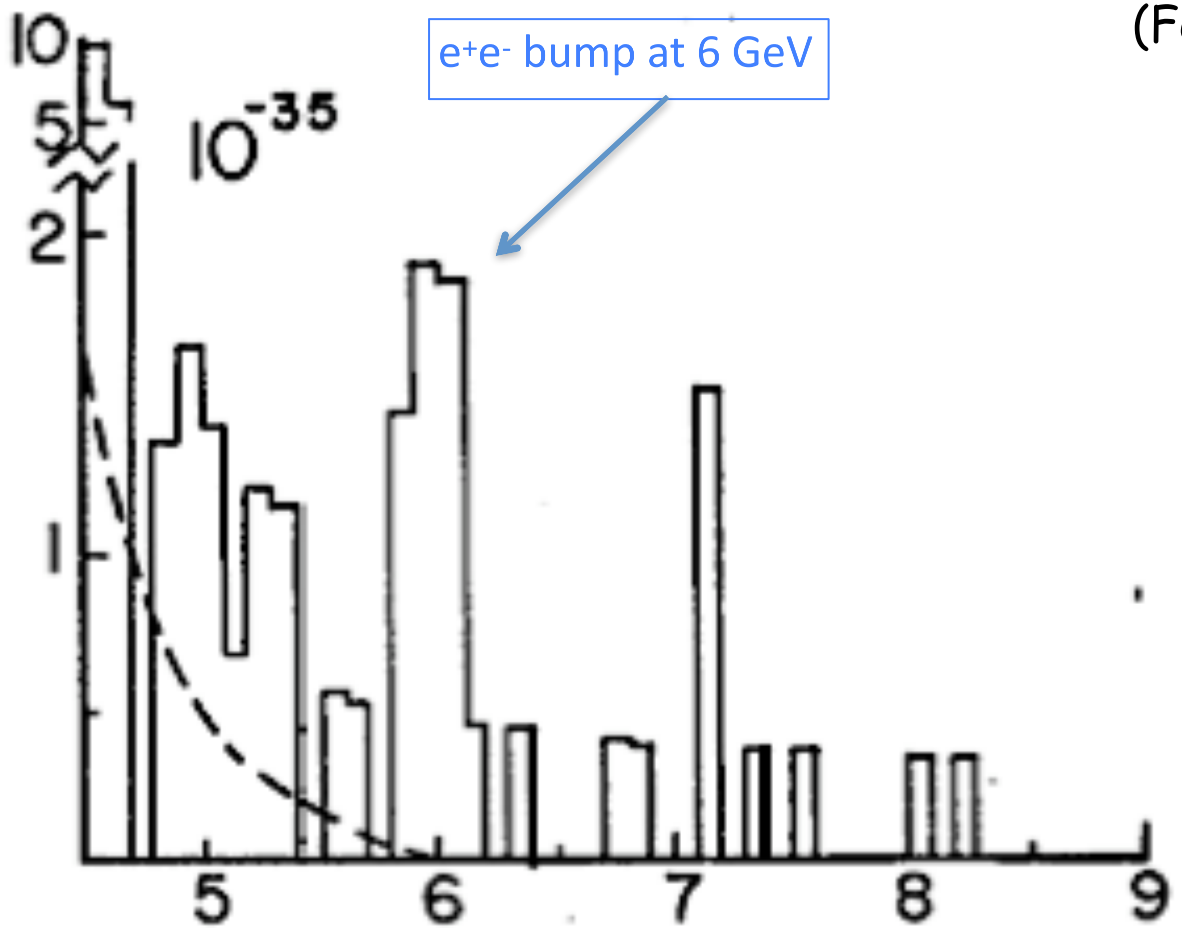


biased data selection

ϒ

"Oops-Leon"

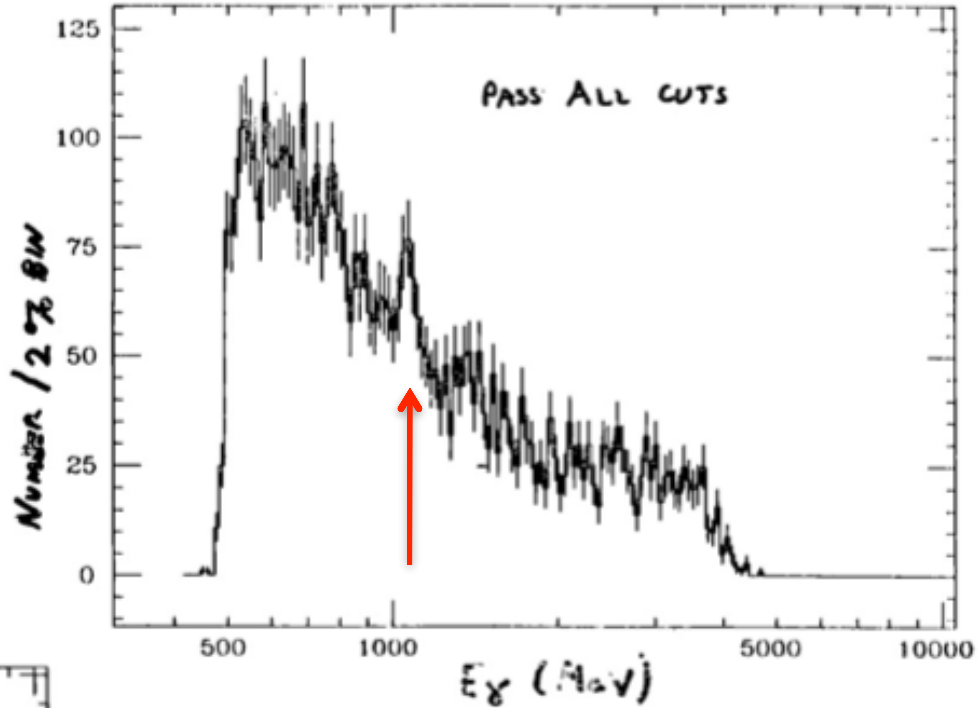
(Fermilab, 1976)

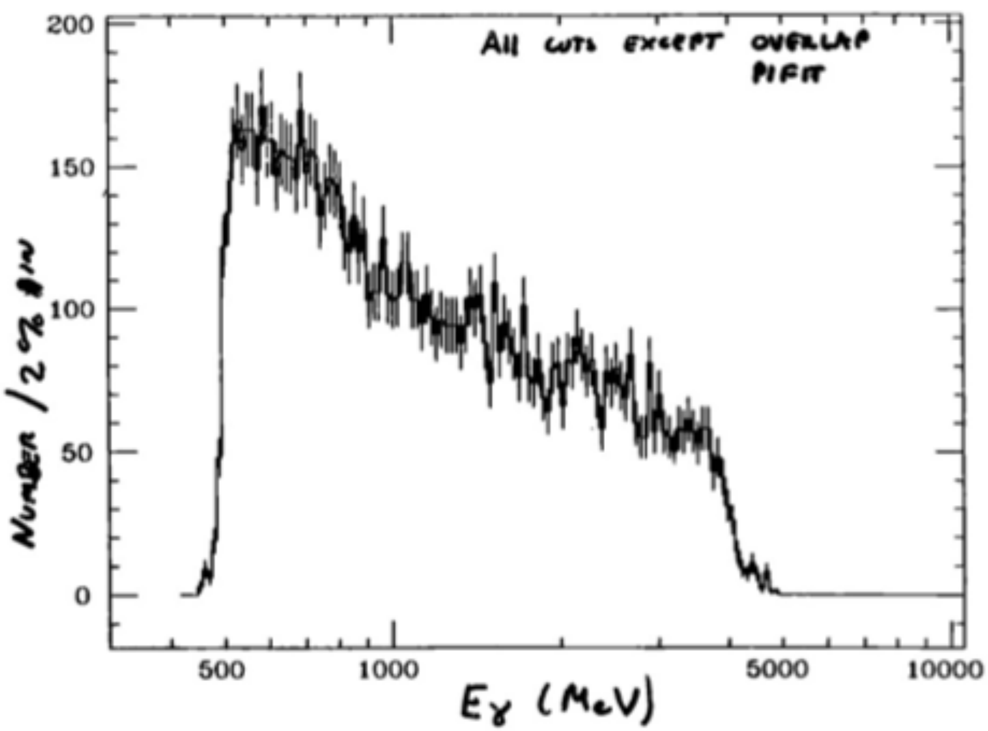


e+e- bump at 6 GeV

Statistical fluctuation

The ζ Particle
(DESY, 1984)





Biased data cuts

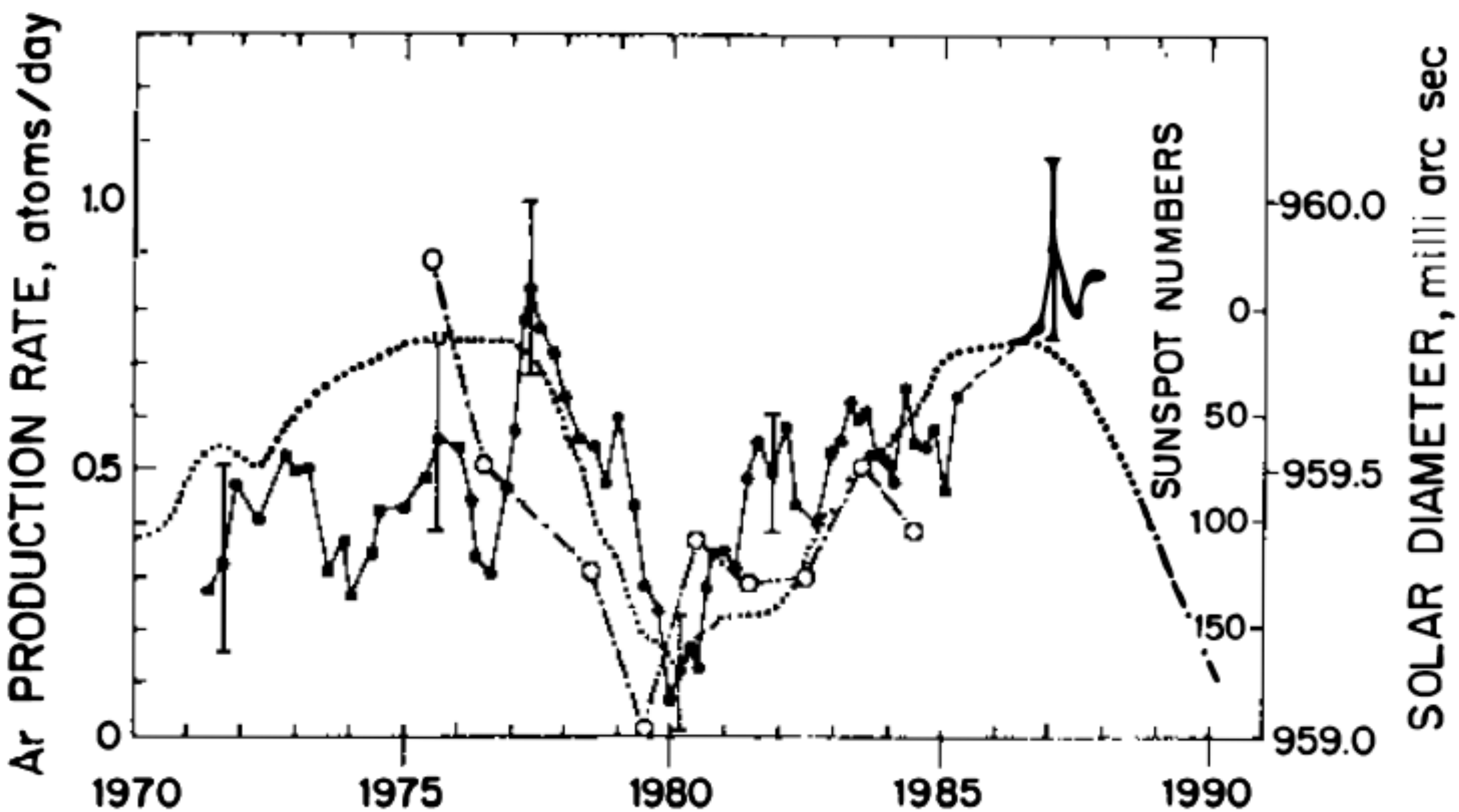
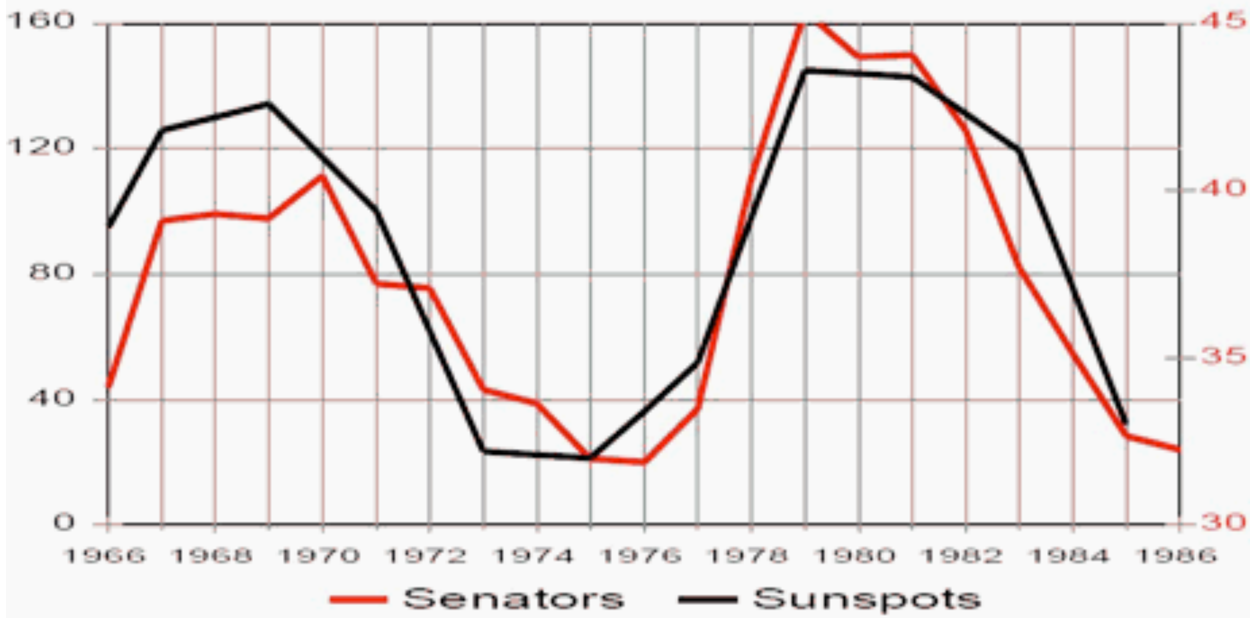


Figure 5 Plots of five-point running average of ^{37}Ar production and smoothed sunspot numbers against time in years (from 130). Solid circles, ^{37}Ar production; dotted curve, sunspot numbers; open circles, solar diameter.

REPUBLICAN SENATORS CAUSE SUNSPOTS
OR MAYBE SUNSPOTS CAUSE REPUBLICAN SENATORS?



CERN, 2011

NewScientist

Surprise LHC blip hints at Higgs – again

22:49 22 July 2011

"...The combined statistical significance, taking all three types of excess reported by ATLAS into account, is 2.8 sigma, slightly below the 3 sigma threshold (equivalent to a 1-in-370 chance of being due to a fluke) that a measurement must pass to count as "evidence" for something new: only 5 sigma data, equivalent to a 1-in-1.7 million chance of being due to a fluke, gains "discovery" status.

The other main detector at the LHC, called CMS, has found an excess in a similar range, between 130 and 150 GeV, reports Nature. The size of that excess is roughly 2 sigma, writes physicist Adam Falkowski on the Resonances blog.

If all this sounds a tad familiar, rewind back to April, when four physicists claimed to have found hints of the Higgs in ATLAS data in a study abstract leaked online. A subsequent official analysis by the collaboration of 700 physicists who run ATLAS concluded that result was an error. Unlike that claim, the new excesses have been vetted by the ATLAS and CMS collaborations respectively."

guardian.co.uk

Higgs boson signals fade at Large Hadron Collider

Cern scientist says he sees 'no striking evidence of anything that could resemble a discovery' in hunt for Higgs boson

Ian Sample
guardian.co.uk, Monday 22 August 2011 17.10 BST
Article history

Bias

and Experimental Design

"Blind" Analysis Techniques

Goal: To remove the ability to unconsciously tune on statistical fluctuations and/or adjust analyses towards a particular outcome by hiding the final result until the full analysis (incl. assessment of uncertainties) is fixed.

At which point you then "open the box" and take what life brings you!

Rules of the Game

- **Agree on an appropriate blindness scheme in advance**
- **Make sure no one breaks it**
- **Agree on the criteria necessary to "open the box"**
- **State the blindness scheme up front in any publication**
- **Agree to show exactly what results from box-opening and then justify any alterations**

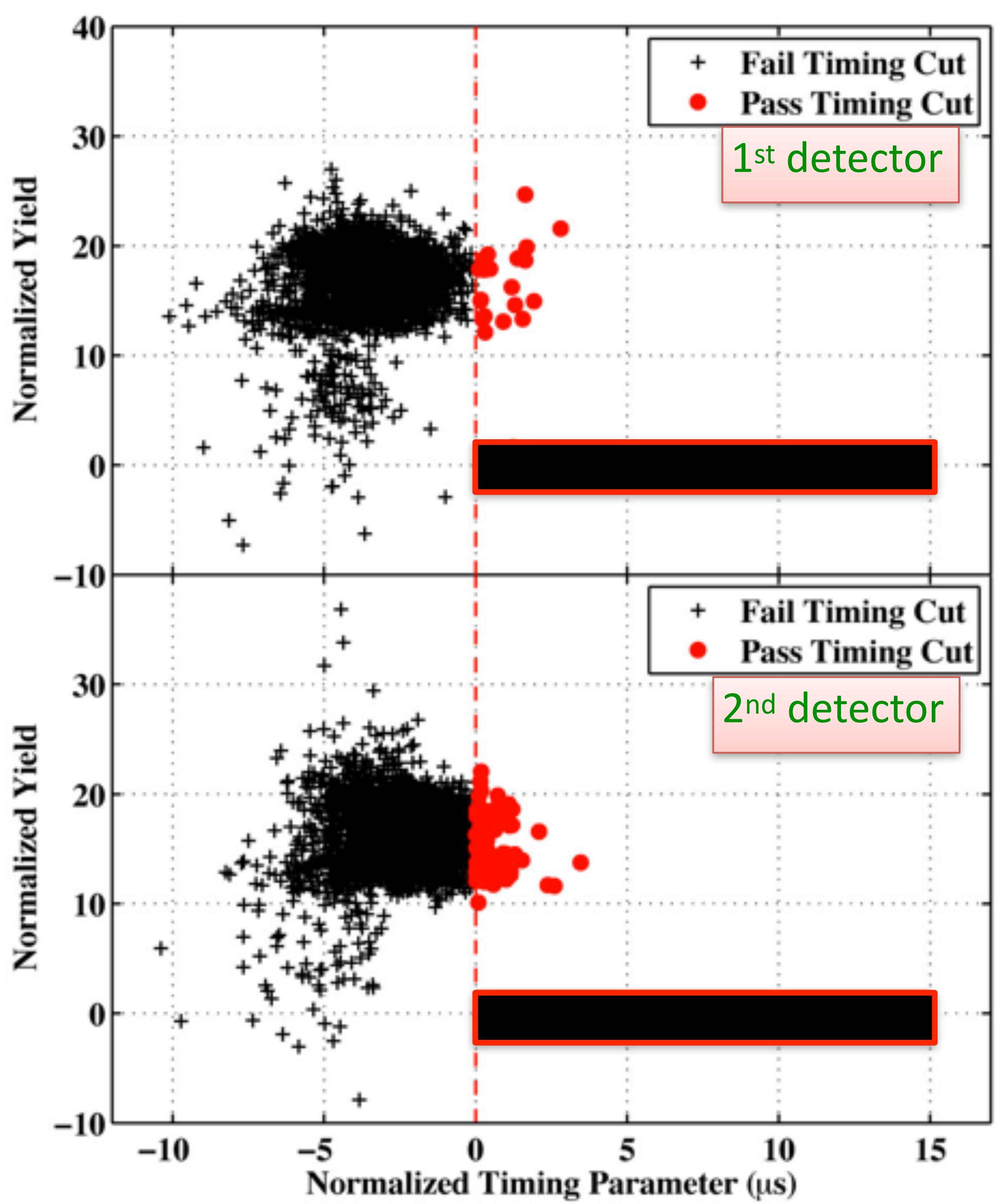
Signal Box Method

CDMS results on search for Dark Matter (Dec, 2009)

Expected summed background in both detectors: **0.9 ± 0.2**

RESULTS:

[redacted]



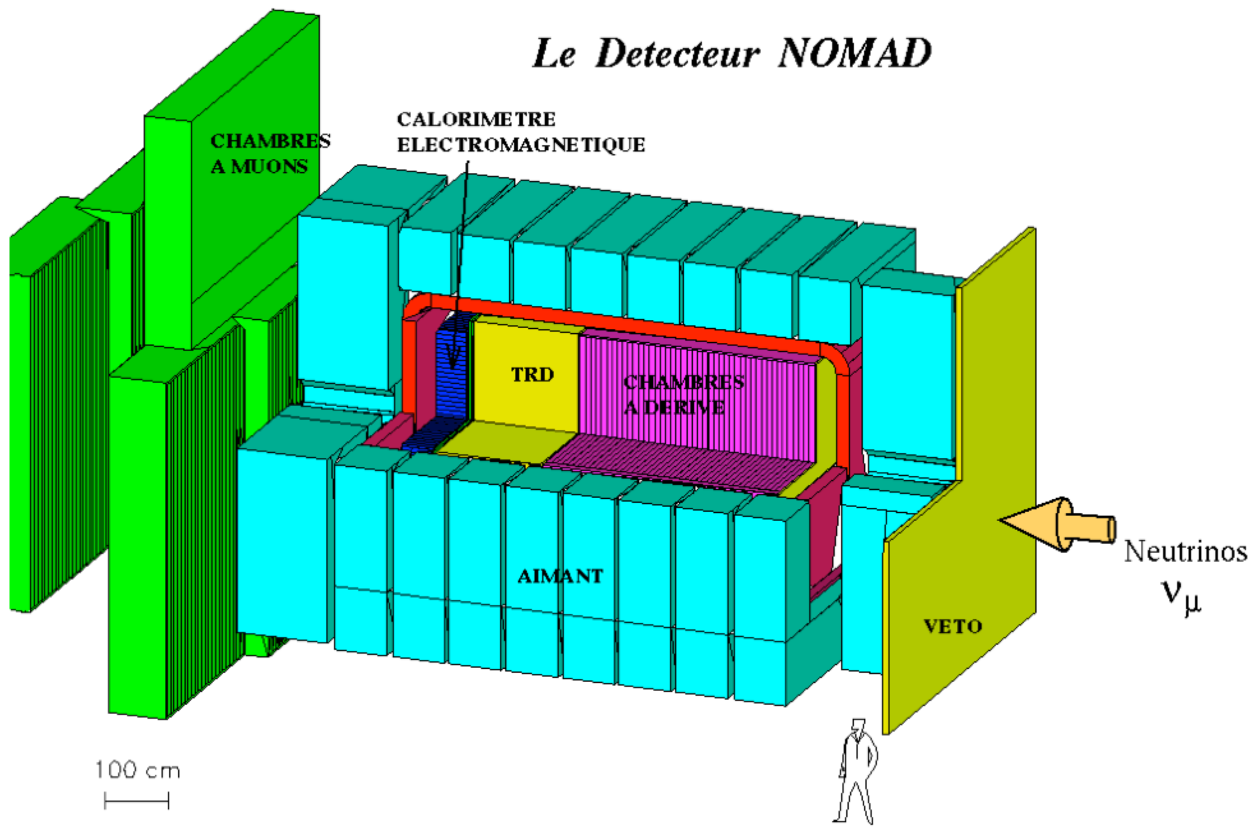
Divided Data Sample

NOMAD Search for

ν_μ - ν_τ oscillations

(Feb, 1999)

Used 20% of data to confirm background predictions and define search window, then impose signal box method on remaining 80% of the data

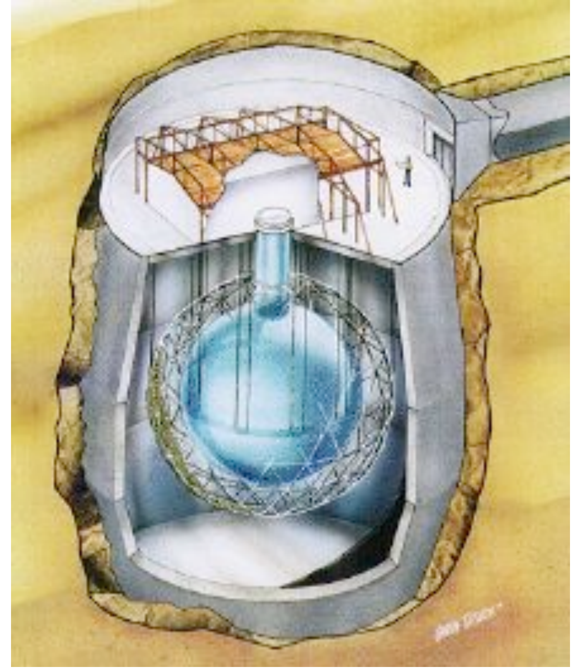


Expected background in signal box: **6.5 ± 1.1**

RESULTS:

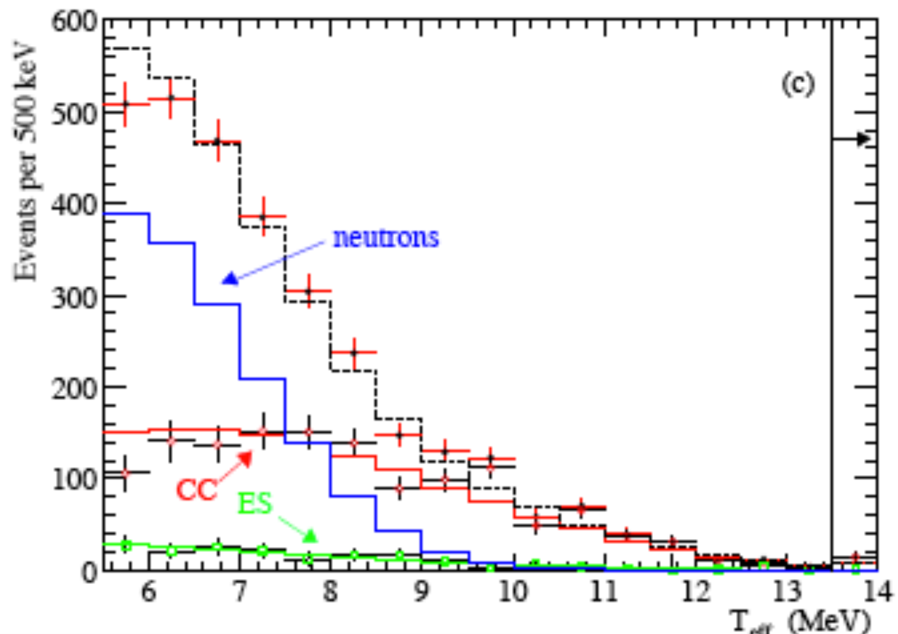
[redacted]

Hidden Parameters



SNO Measurement of total solar neutrino flux (Sept, 2003)

Excluded a hidden fraction of the final data set (unknown flux normalisation), included hidden admixture of tagged background neutrons, scaled simulation NC cross section by hidden factor

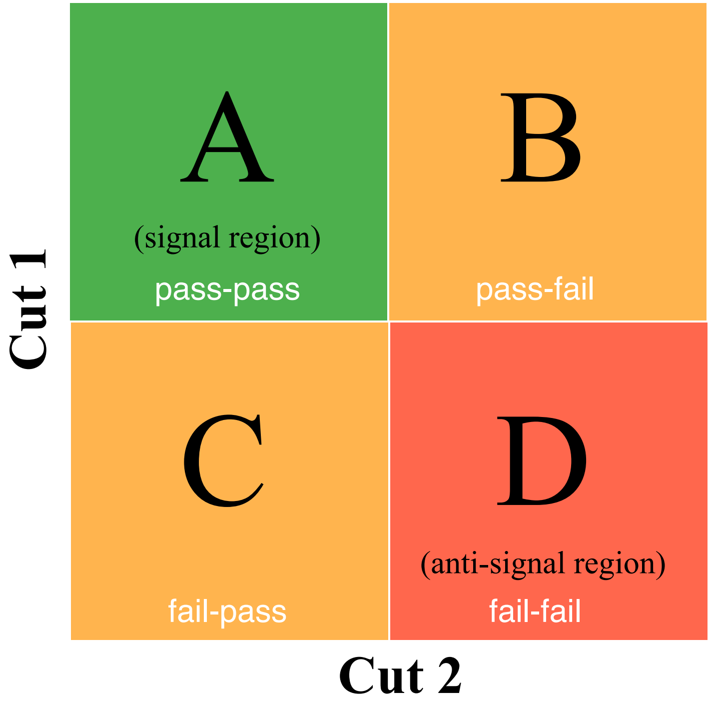


RESULTS:

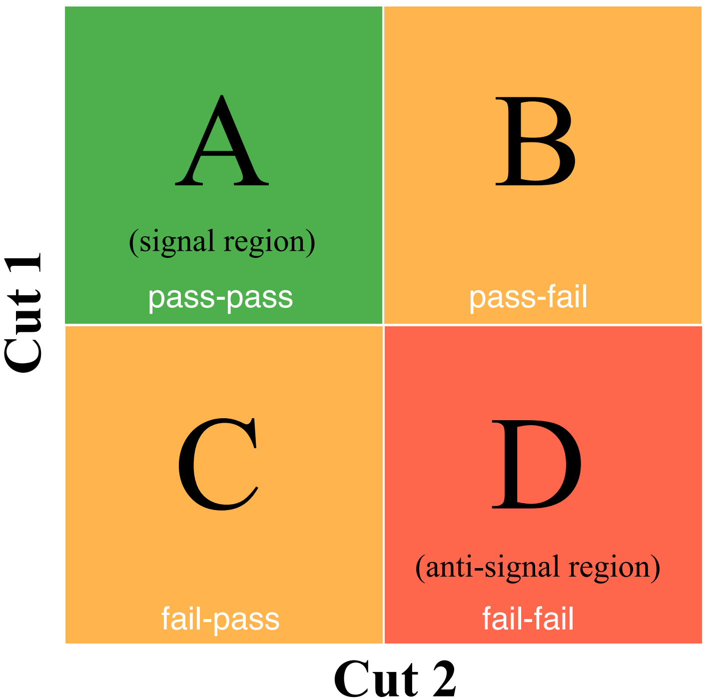
[redacted]

Bifurcated Side-Band Analysis*

Assume we have a data set with a total number of signal S and a total number of background B. Further assume that we have two independent parameters (for example, energy and fiducial volume) that can be used to cut out some number of unknown background while maintaining high signal efficiency (based on simulations of the signal). We wish to estimate the background contamination in the signal region:



*Generalisation of Adler *et al.*, PRL 79, 12 1997 and Nix *et al.*, NIM A615, 2, 2010 to account for signal efficiencies



Take the efficiency of retaining signal from each cut in the signal region to be **ε_1** and **ε_2**, respectively. Similarly, take the fractions of background rejected by each cut in this region to be **r_1** and **r_2**, respectively.

$$N_A = S\epsilon_1\epsilon_2 + Br_1r_2 \qquad \equiv s + b$$

$$N_B = S\epsilon_1(1-\epsilon_2) + Br_1(1-r_2)$$

$$N_C = S\epsilon_2(1-\epsilon_1) + Br_2(1-r_1)$$

$$N_D = S(1-\epsilon_1)(1-\epsilon_2) + B(1-r_1)(1-r_2)$$

To simplify the algebra a bit, let's redefine variables:

$$n_A \equiv \frac{N_A}{\epsilon_1\epsilon_2} = S + B\left(\frac{r_1r_2}{\epsilon_1\epsilon_2}\right)$$

$$n_C \equiv \frac{N_C}{\epsilon_2(1-\epsilon_1)} = S + B\left(\frac{r_2(1-r_1)}{\epsilon_2(1-\epsilon_1)}\right)$$

$$n_B \equiv \frac{N_B}{\epsilon_1(1-\epsilon_2)} = S + B\left(\frac{r_1(1-r_2)}{\epsilon_1(1-\epsilon_2)}\right)$$

$$n_D \equiv \frac{N_D}{(1-\epsilon_1)(1-\epsilon_2)} = S + B\left(\frac{(1-r_1)(1-r_2)}{(1-\epsilon_1)(1-\epsilon_2)}\right)$$

$$n_A - S = B\left(\frac{r_1 r_2}{\epsilon_1\epsilon_2}\right) \quad n_B - S = B\left(\frac{r_1(1-r_2)}{\epsilon_1(1-\epsilon_2)}\right) \quad n_C - S = B\left(\frac{r_2(1-r_1)}{\epsilon_2(1-\epsilon_1)}\right) \quad n_D - S = B\left(\frac{(1-r_1)(1-r_2)}{(1-\epsilon_1)(1-\epsilon_2)}\right)$$

$$(n_C - S)(n_B - S) = (n_A - S)(n_D - S)$$

$$n_C n_B - n_C S - S n_B + S^2 = n_A n_D - n_A S - S n_D + S^2$$

$$S = \frac{n_A n_D - n_C n_B}{n_A + n_D - n_C - n_B}$$

re-expanding:

$$S = \frac{N_A N_D - N_C N_B}{N_A(1-\epsilon_1)(1-\epsilon_2) + N_D\epsilon_1\epsilon_2 - N_C\epsilon_1(1-\epsilon_2) - N_B\epsilon_2(1-\epsilon_1)}$$

$$s = S\epsilon_1\epsilon_2 \qquad\qquad b = N_A - S\epsilon_1\epsilon_2$$

Do not need to know details about r_1 and r_2 !

$$S = \frac{N_A N_D - N_C N_B}{N_A(1-\epsilon_1)(1-\epsilon_2) + N_D\epsilon_1\epsilon_2 - N_C\epsilon_1(1-\epsilon_2) - N_B\epsilon_2(1-\epsilon_1)}$$

$$s = S\epsilon_1\epsilon_2 \qquad\qquad b = N_A - S\epsilon_1\epsilon_2$$

note: as $\epsilon_1, \epsilon_2 \to 1$ $\qquad b \to \frac{N_B N_C}{N_D}$

Do not need to even look in the signal region !

So, for large efficiencies, the variance in the estimated background contamination, **b**, is approximately:

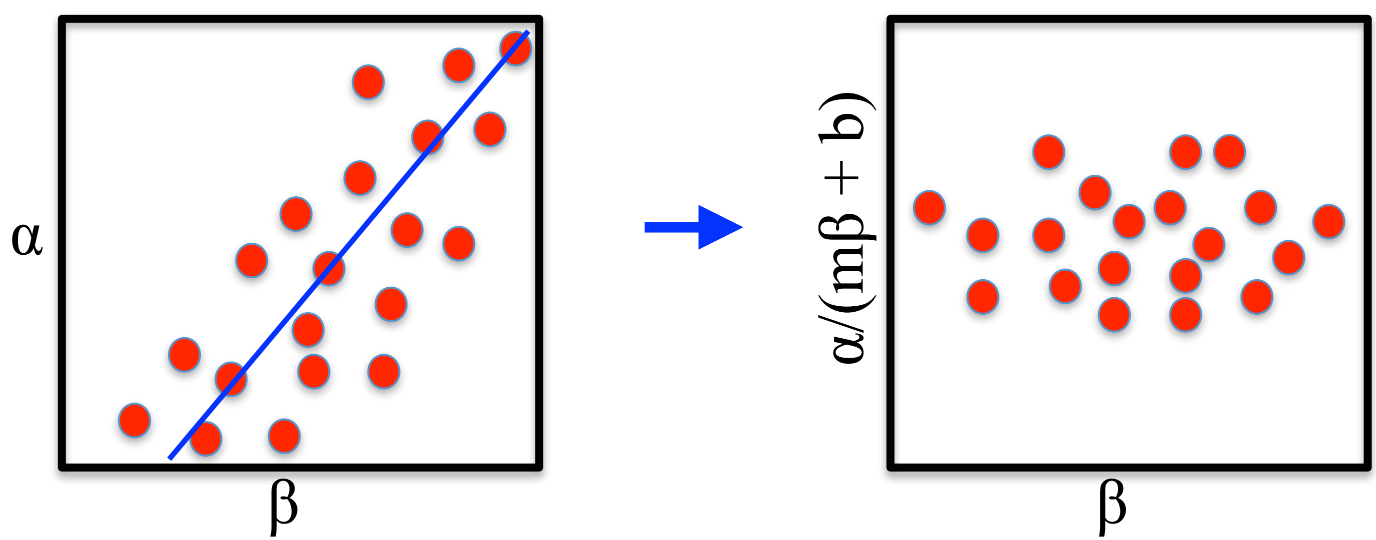
$$\sigma^2_{var} \simeq N_B\left(\frac{N_C}{N_D}\right)^2 + N_C\left(\frac{N_B}{N_D}\right)^2 + N_D\left(\frac{N_B N_C}{N_D^2}\right)^2$$

Could first use tight cuts with high efficiency for 1st order look, then loosen cuts in pre-determined way once box is opened to better evaluate signal and background contamination

Remember, this assumes cut parameters are uncorrelated! Note that a mixed background model can inadvertently produce correlations if, for example, both **r1** and **r2** are notably different between background components: then a particular cut value could favour a particular background, which could then produce a correlated rejection for the second cut.

In general, should look for possible correlations by plotting one cut parameter versus another, for example, in the anti-signal cut region (*i.e.* box D).

If a correlation is present, you may be able to redefine your parameters to remove this to first order. For example:



Alternatively, we can first define the background model as the sum of various components. Now assume that we can decompose these into a set of backgrounds that are **well-modelled and/or sub-dominant,** plus a background with the highest uncertainty that we most wish to evaluate:

$$\sum_i B_i r_1^i r_2^i + B r_1 r_2$$

background we most want to evaluate

Then, similar to before, we can define the following quantities:

$$\eta_A \equiv \frac{1}{\epsilon_1 \epsilon_2}\left(N_A - \sum_i B_i r_1^i r_2^i\right) = S + B\left(\frac{r_1 r_2}{\epsilon_1 \epsilon_2}\right)$$

$$\eta_B \equiv \frac{1}{\epsilon_1(1-\epsilon_2)}\left(N_B - \sum_i B_i r_1^i (1-r_2^i)\right) = S + B\left(\frac{r_1(1-r_2)}{\epsilon_1(1-\epsilon_2)}\right)$$

$$\eta_C \equiv \frac{1}{\epsilon_2(1-\epsilon_1)}\left(N_C - \sum_i B_i r_2^i (1-r_1^i)\right) = S + B\left(\frac{r_2(1-r_1)}{\epsilon_2(1-\epsilon_1)}\right)$$

$$\eta_D \equiv \frac{1}{(1-\epsilon_1)(1-\epsilon_2)}\left(N_D - \sum_i B_i (1-r_1^i)(1-r_2^i)\right) = S + B\left(\frac{(1-r_1)(1-r_2)}{(1-\epsilon_1)(1-\epsilon_2)}\right)$$

$$S = \frac{\eta_A \eta_D - \eta_C \eta_B}{\eta_A + \eta_D - n_C - n_B}$$

Does not depend on knowing the rejection factors for the unknown background!

$$S = \frac{\eta_A \eta_D - \eta_C \eta_B}{\eta_A + \eta_D - n_C - n_B}$$

as $\epsilon_1, \epsilon_2 \rightarrow 1$

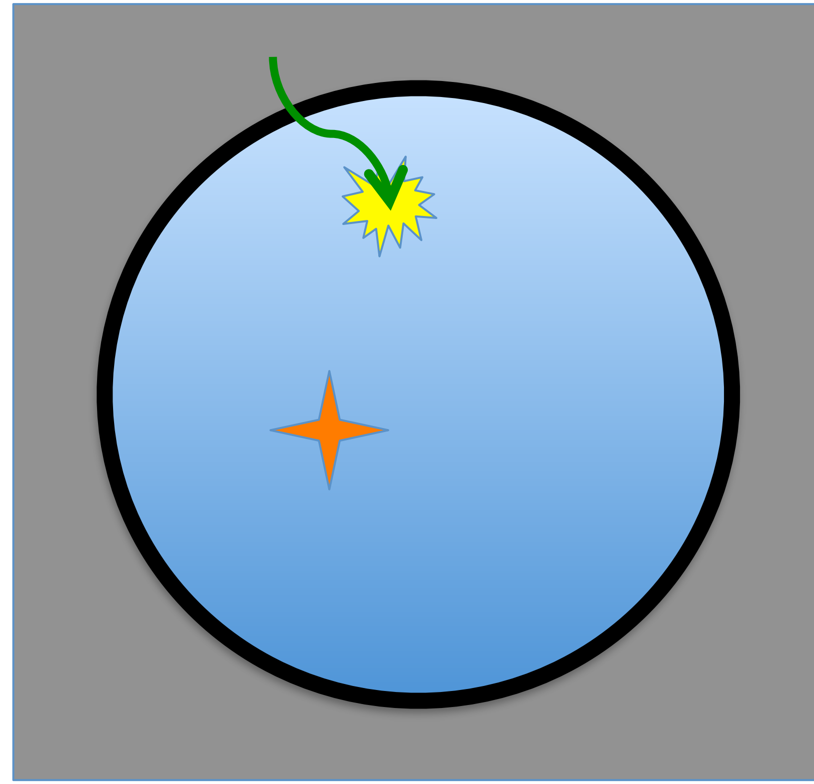
$$b \rightarrow \frac{\left[N_B - \sum_i B_i r_1^i (1 - r_2^i)\right] \left[N_C - \sum_i B_i r_2^i (1 - r_1^i)\right]}{\left[N_D - \sum_i B_i (1 - r_1^i)(1 - r_2^i)\right]}$$

Again, no need to even look in the signal region !

Statistical Optimisation

Working Backwards

Assume that both the signal and background levels are proportional to the detector mass, M, and running time, T. Find an expression for the maximum background level that can be tolerated to achieve a 3σ detection as a fraction of the expected signal for a given model. How does the sensitivity change as a function of M and T ?



$$B = fS$$

$$1\sigma = \sqrt{B} = \sqrt{fS}$$

under H0

Thus, for a 3σ signal: $$3\sqrt{fS} = S$$

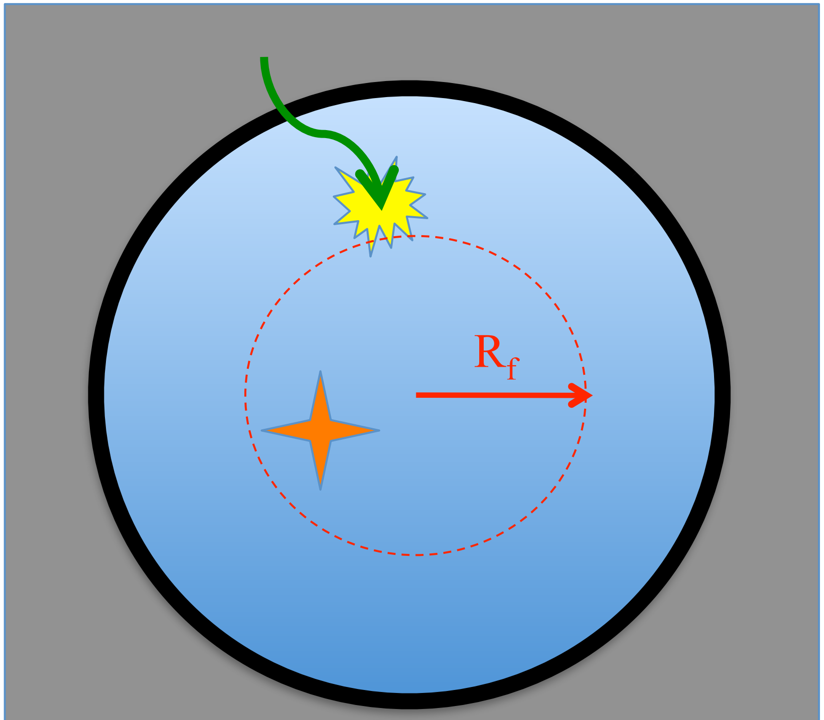
(able to tolerate more background for larger signal) $$f = \frac{S}{9}$$

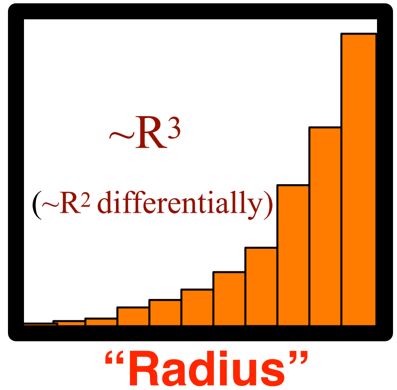
or $$B = \frac{S^2}{9}$$

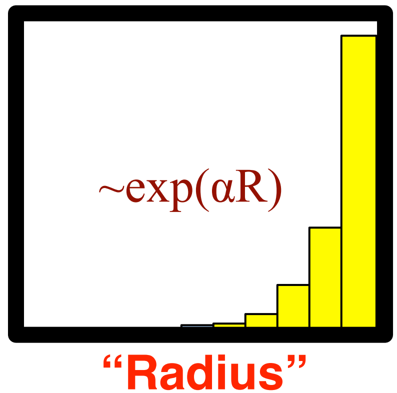
$$\text{Significance } (\sigma\text{'s}) = \frac{S}{\sqrt{B}}$$

$$= \frac{\alpha MT}{\sqrt{f\alpha MT}} \propto \sqrt{MT}$$

Example of Statistical Optimisation







Assume that we are in the "large N" limit and expected the number of counts to be dominated by background events.

We wish to exclude the worst of the background by choosing a radius to define a "fiducial volume," within which will look for an excess of events as evidence of a signal.

What choice of fiducial radius will give the best sensitivity for the search?

$$\frac{S}{\sqrt{B}} \sim \frac{R^3}{\sqrt{\exp(\alpha R)}} = R^3 e^{-\alpha R/2}$$

maximise:

$$3R^2 e^{-\alpha R/2} - \frac{\alpha}{2} R^3 e^{-\alpha R/2} = 0$$

$$3R^2 = \frac{\alpha}{2} R^3 \qquad R = \frac{6}{\alpha}$$

From the plot, it looks like backgrounds fall by ~1/e when R changes by 10% of the detector radius... so α ~10

$$R_f = 0.6 R_d$$

VHE γ-Ray Astronomy

Assuming an angular resolution characterised by a Gaussian, what is the optimal angular bin radius to maximise signal sensitivity?

$$S \sim \int_0^a \theta \exp(-\frac{\theta^2}{2\sigma^2}) d\theta$$

$$= \sigma^2 \left[1 - \exp(-\frac{a^2}{2\sigma^2})\right]$$

$$B \sim \int_0^a \theta d\theta = \frac{a^2}{2}$$

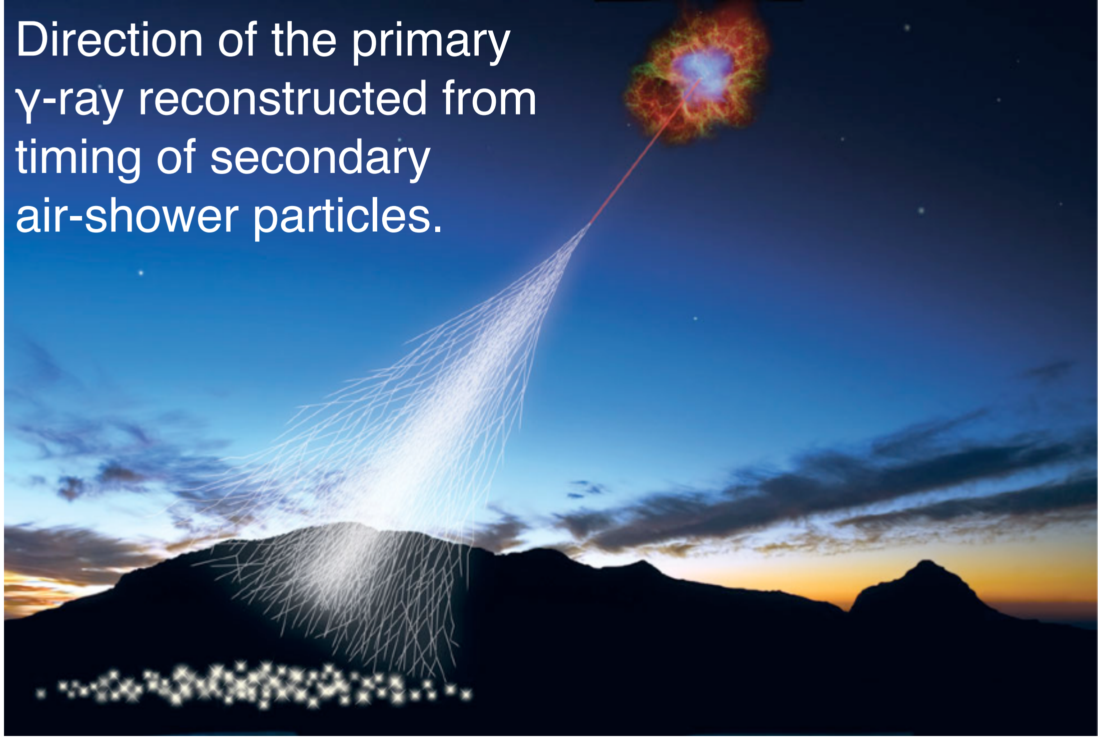
$$\frac{S}{\sqrt{B}} \sim \left[1 - \exp(-\frac{a^2}{2\sigma^2})\right] a^{-1}$$

Maximise:

$$-\frac{1}{a^2}\left[1 - \exp(-\frac{a^2}{2\sigma^2})\right] + \frac{1}{\sigma^2}\exp(-\frac{a^2}{2\sigma^2}) = 0$$

$$\exp(-\frac{a^2}{2\sigma^2}) = \left(1 + \frac{a^2}{\sigma^2}\right)^{-1}$$

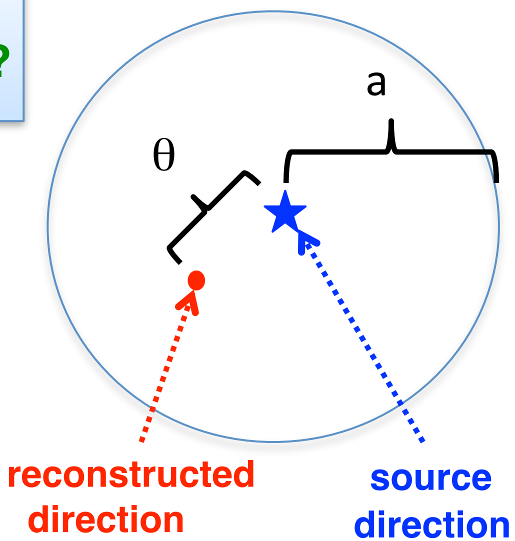
Numerically: $\frac{a}{\sigma} \simeq 1.58$



Direction of the primary γ-ray reconstructed from timing of secondary air-shower particles.

What if the expected background is very small?

Sky View:



A Brief Note On

Redundancy

&

Calibrations:

Sudbury Neutrino Observatory (SNO)

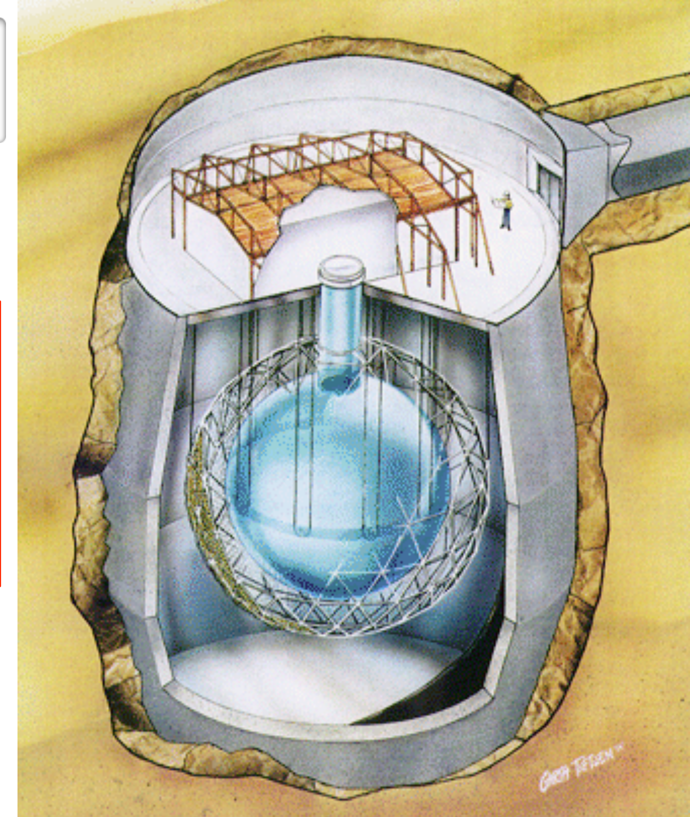
3 Different Operational Phases

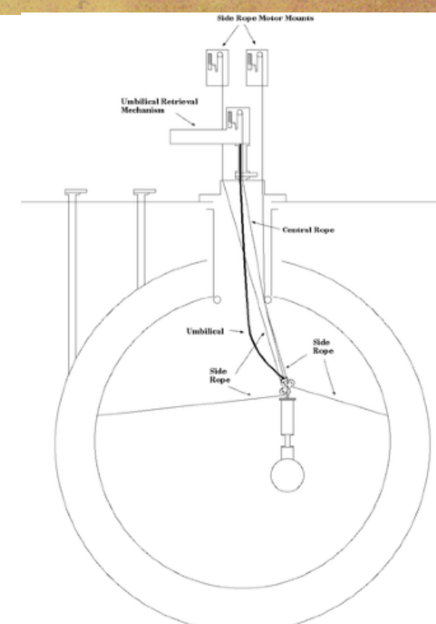
Found that estimated systematic uncertainty in possible position-dependent energy resolution was larger for the 2nd phase, which should have performance at least as good as 1st phase(?!)

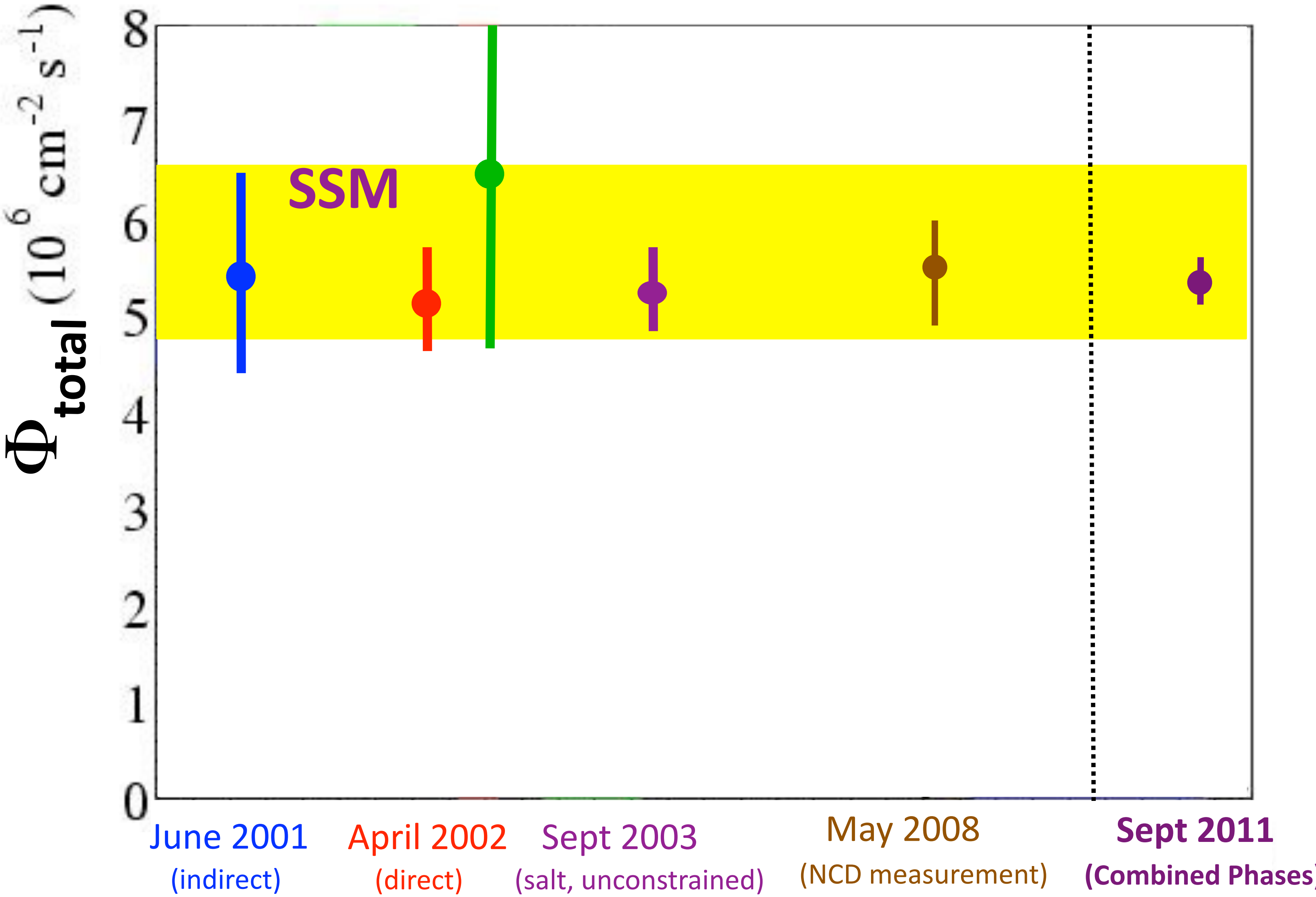
Realised that fewer calibrations had been done in 1st phase, so there was less data to compare!

If you don't look, you don't see!!

(Some groups seem to have elevated this to a strategy for getting small errors!)







3 Experimental Techniques,
at Least 2 Analyses/Technique + Combined Cross-checks