Modern Methods of Data Analysis

Lecture XIII (21.01.08)

Contents:

- Blind Analysis

Introduction

Use of blind techniques in mainstream science began in 1930's, when double blinded procedure were introduced in medical research. Blinding is a basic too to prevent conscious and unconscious bias in research.

- Single-blind experiment "Pepsi Challenge":
  - test person are blinded, experimenter are not, can any bias appear through interaction with experimenter?
- Double-blind experiment:
  - both test person & experimenters are blind, introduced to achieve a higher standard of scientific rigour
  - The key for assignment to the experimental or control group is kept by a third party and not "unblinded" till the study is over (and analyzed).
  - Computer-controlled experiments are sometimes referred to as double-blind experiments, since software can be designed to not cause any bias

Blind Analysis in Physics

- *E. Rutherford, 1933*: “It seems to me that in some way it is regrettable that we had a theory of the positive electron before the beginning of the experiments. Blackett did everything possible not to be influenced by the theory, but the way of anticipating results must inevitably be influence to some extend by the theory. I would have liked it better if the theory had arrived after the experimental facts had been established.”
- 1977-81 a group of Standford physicists led by W. Fairbank, published a series of results on the search for fractional charge in ordinary matter. They claimed “unambiguously the existence of fractional charges of 1/3 e. Louis Alvarez subsequently proposed that “blind tests” be employed, in which a randomly chosen charge of value unknown to the experimenters, would be added to the data. Subsequent measurements by the Standford group did incorporate the blind test. After “unblinding” the new results, the tests did not confirm the original “discovery”.
- Blind analysis are commonly used since about 10 years ago (Babar experiment (SLAC) pushed forward its extensive use).

Definition of Bias

- An estimator is unbiased if $E[\hat{a}] = a$. An Estimator is biased if $E[\hat{a}] = a + b$. If bias vanishes for large data samples: The estimator is called asymptotically unbiased.

- Statistical bias can be corrected for
  - redo experiments + parameter estimate many times in toy experiments. Compute $b = E[\hat{a}] - a$; subtract bias from result in data.

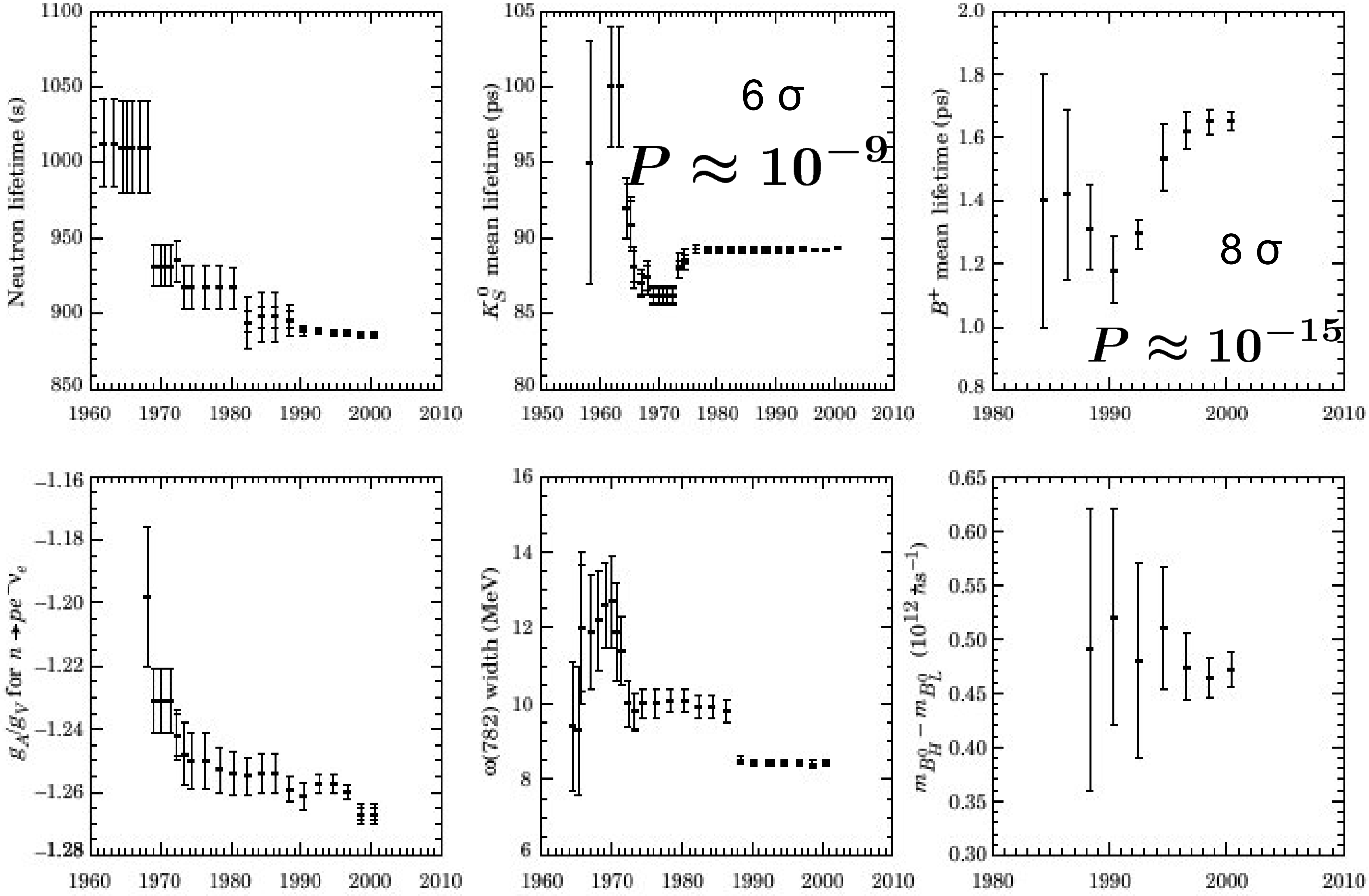
- Experimenter bias occurs when human behaviour enters the equation.

Typical Sources For Bias

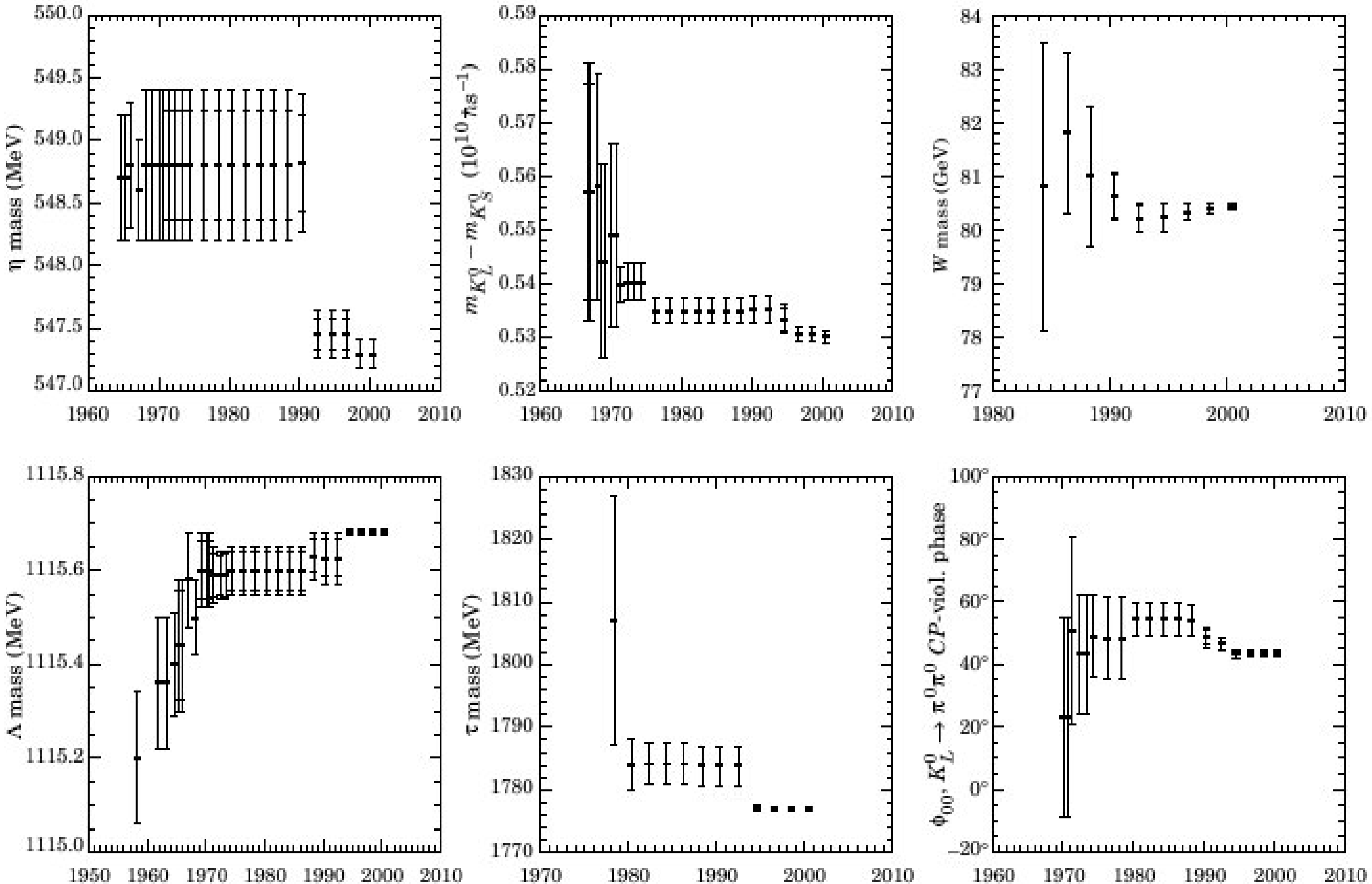
- Looking for bugs, when a result does not conform to expectation (and not looking when it does)
- Looking for additional sources of systematic uncertainty when a result does not conform. These check may lead to “corrections” that change the answer.
- Decide wether to publish, or to wait for more data
- Choosing to drop “strange” events (e.g. track is 2σ away from expectation)
- Data selection criteria (“cuts”) are unconsciously adjusted to bring the answer closer to a theoretical value or a previous measurement.
- Several competing analyses are performed using the same data. The physics group charged with making the decision chooses which is worthy of publication after learning the answer, unconsciously favouring the analyses that “come out right”.

In each case, the experimenter bias in unintentional – the experimenters normally know that these practices are objectionable, however in each example, the course of the analysis is unconsciously influenced by the knowledge of how the outcome is affected.

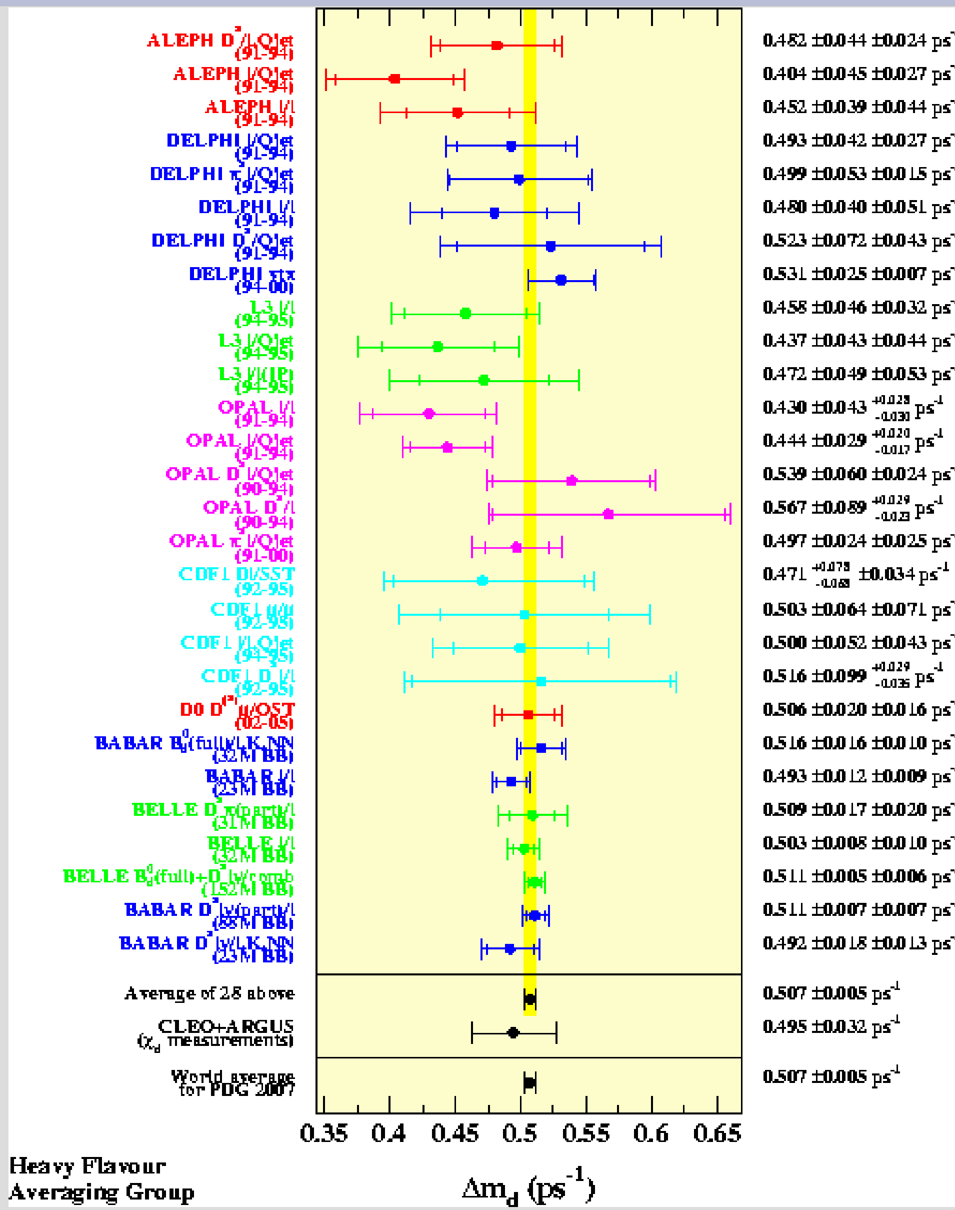
PDG: Experimenter Effect (I)



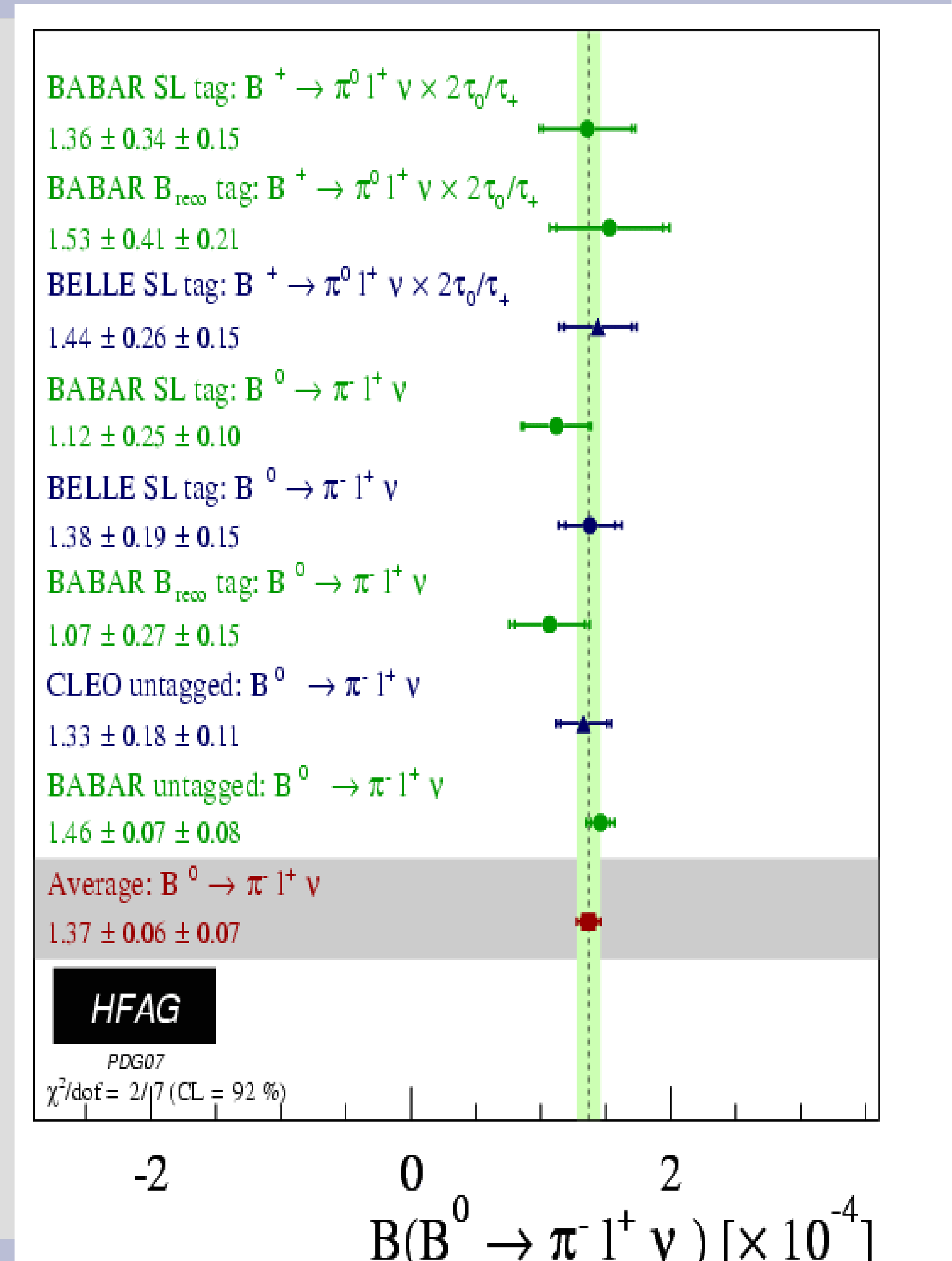
PDG: Experimenter Effect (II)



Do Experiments agree too well?



$$\chi^2 = 17,\ ndf = 27,\ P = 4\%$$



To Avoid Bias ...

- Plan your analysis before hand, do a complete test analysis on Monte Carlo. Write down list of checks (corrections) you plan to do in case your result in data is 8σ off and perform them in any case and don't perform any additional checks, when you are really far off.
  - However, if your Monte Carlo don't describe your data well, one like to develop the analysis instead on data.
  - Is one biased in performing checks when one knows the answer?

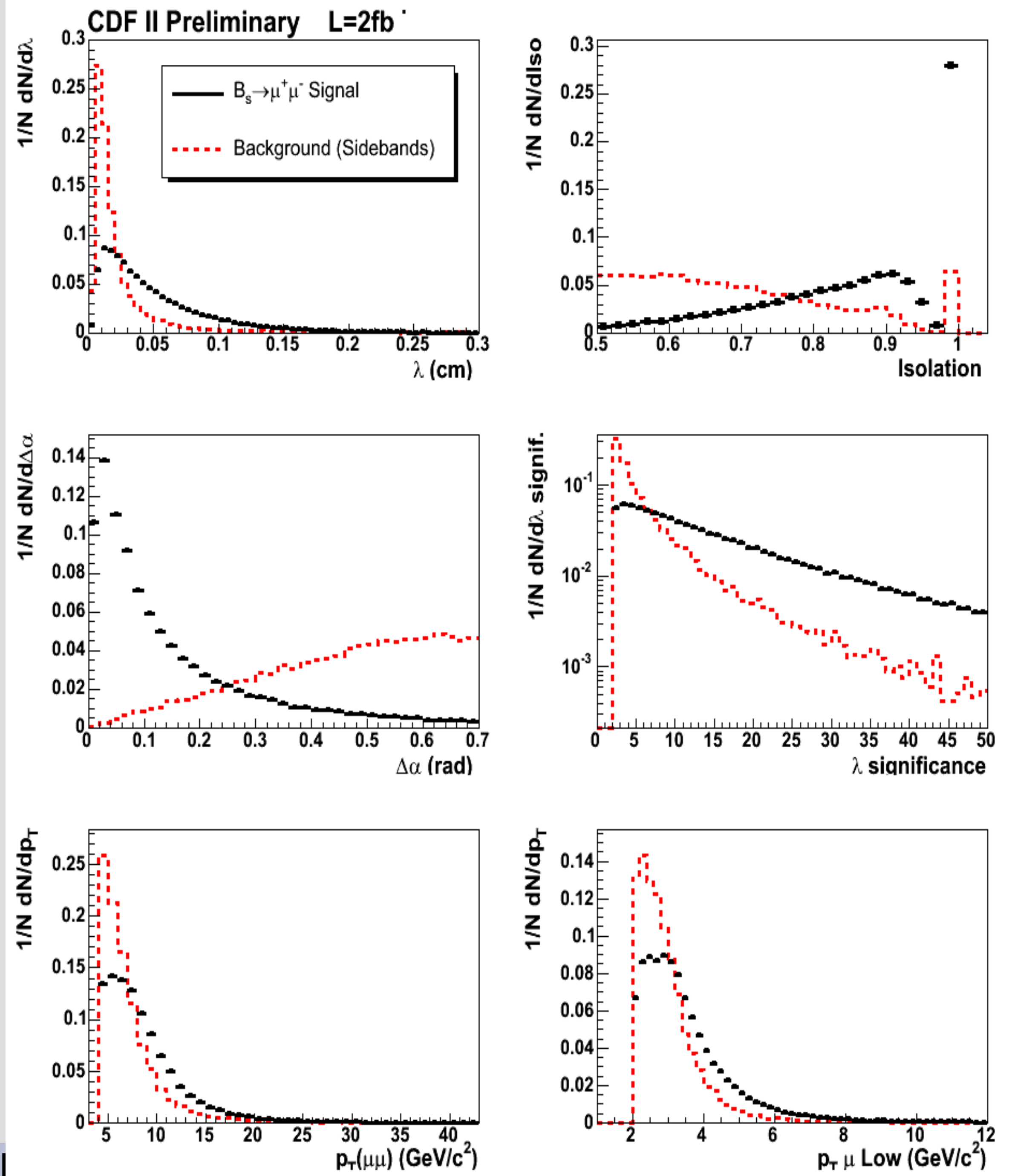
    => blinding data,
    this allows to perform complete analysis and checks on data without knowing the result

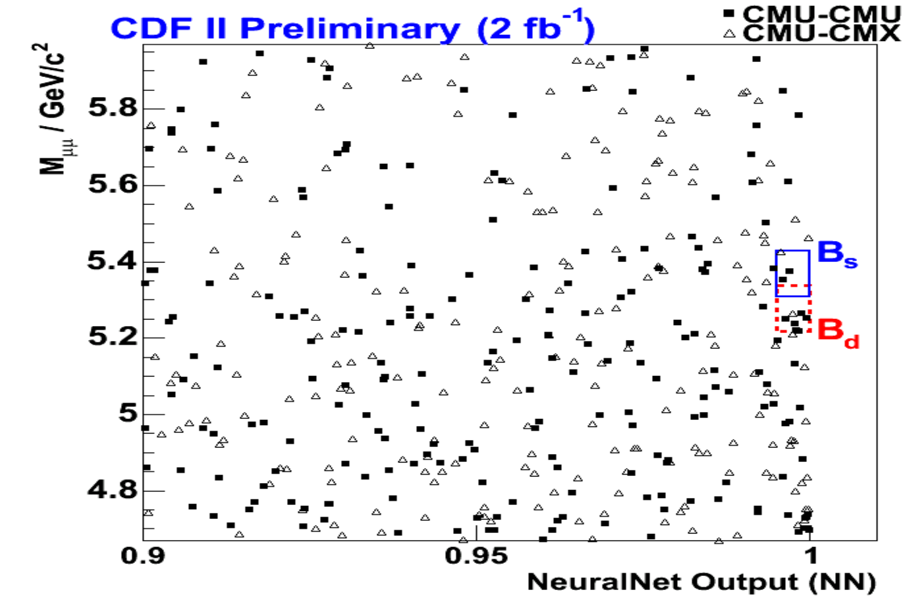
Blinding Techniques (I)

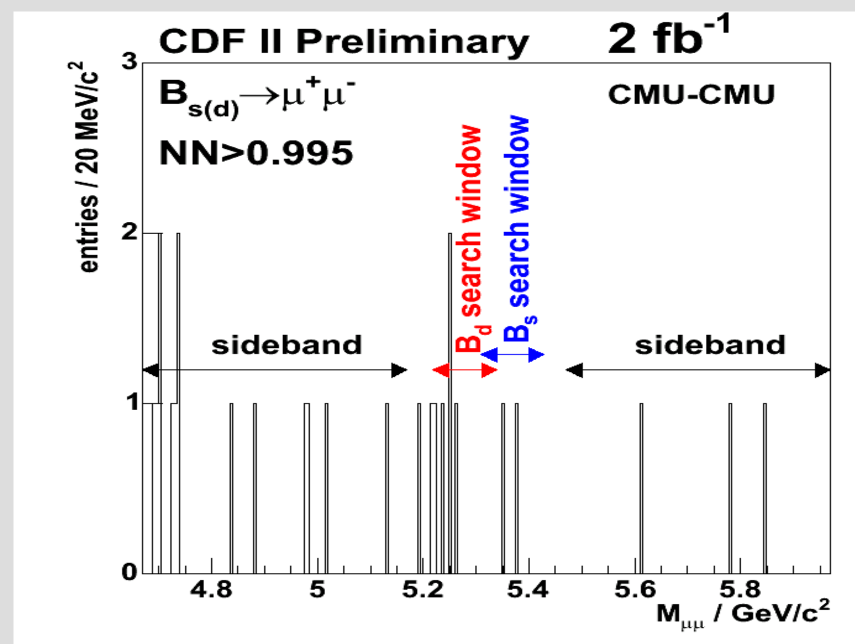
Only rather general advices available, the experimenters doing the analysis must ultimately decide, what is need and what is doable. The techniques are, in many cases, trivial

- Hiding the signal/answer
  - optimize cuts on MC signal and background (sidebands) from data. Don't look at signal data before final cuts are selected.
  - Standard for searches for rare decays (new particles)

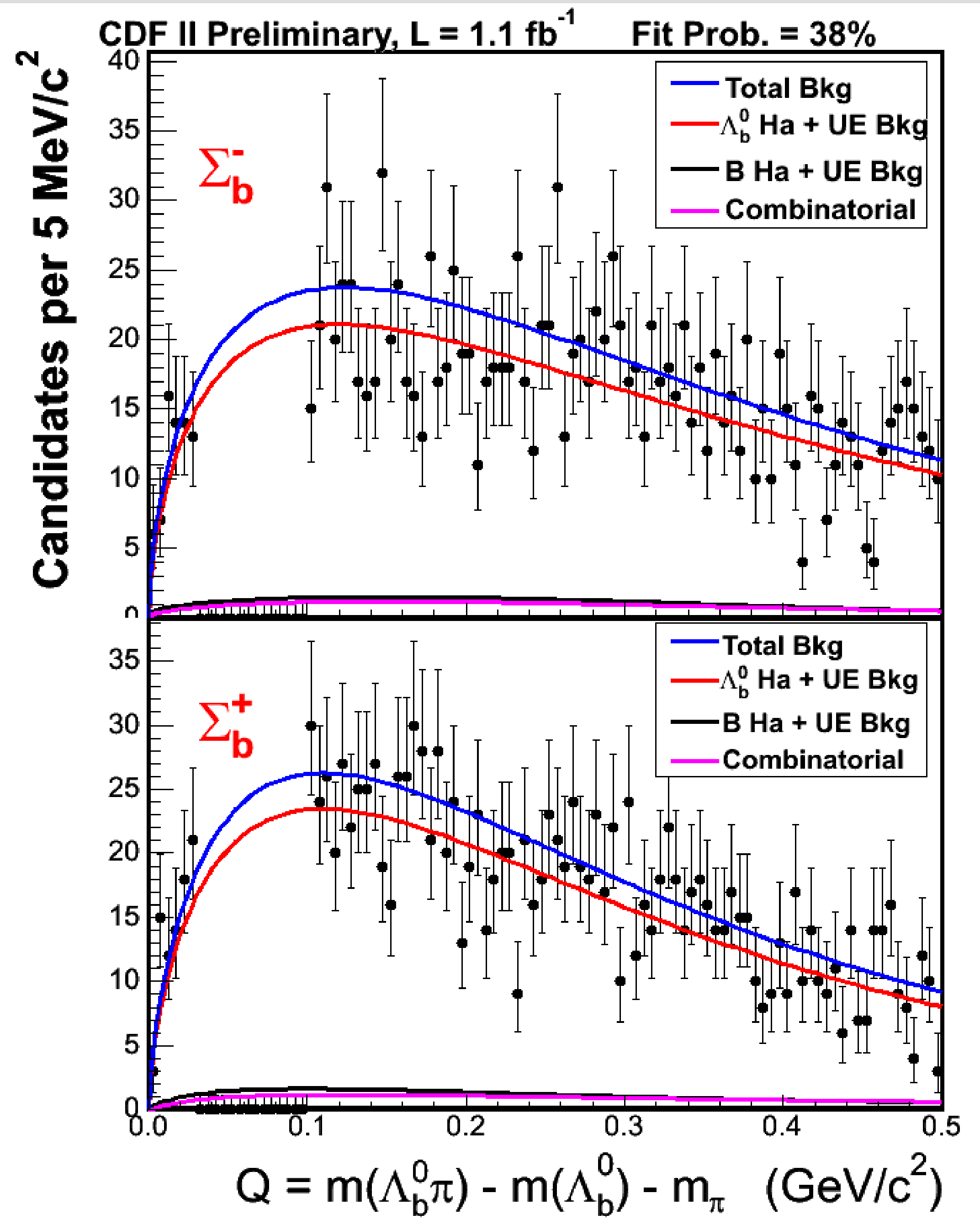
Example : $B_{s/d} \rightarrow \mu^+\mu^-$

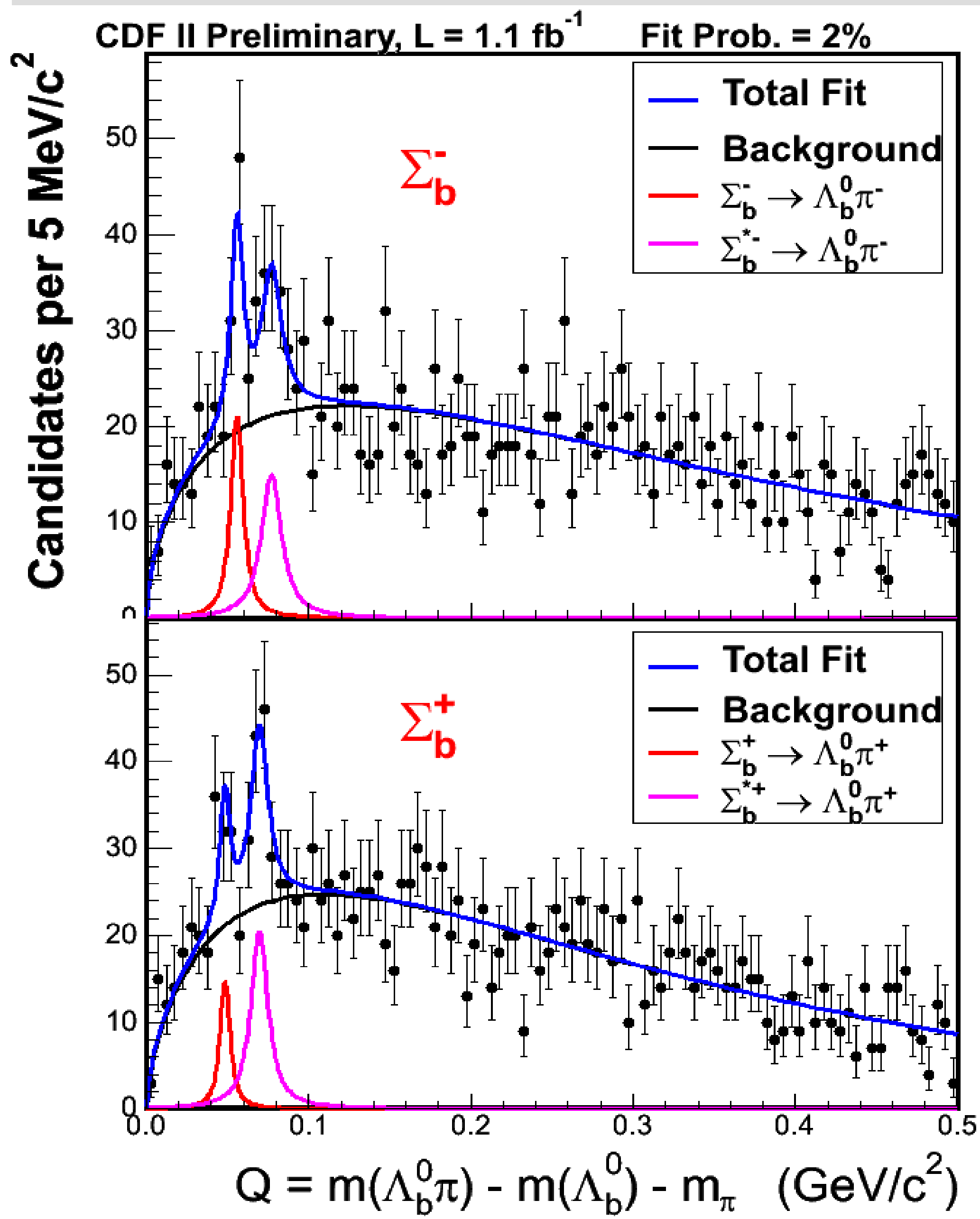






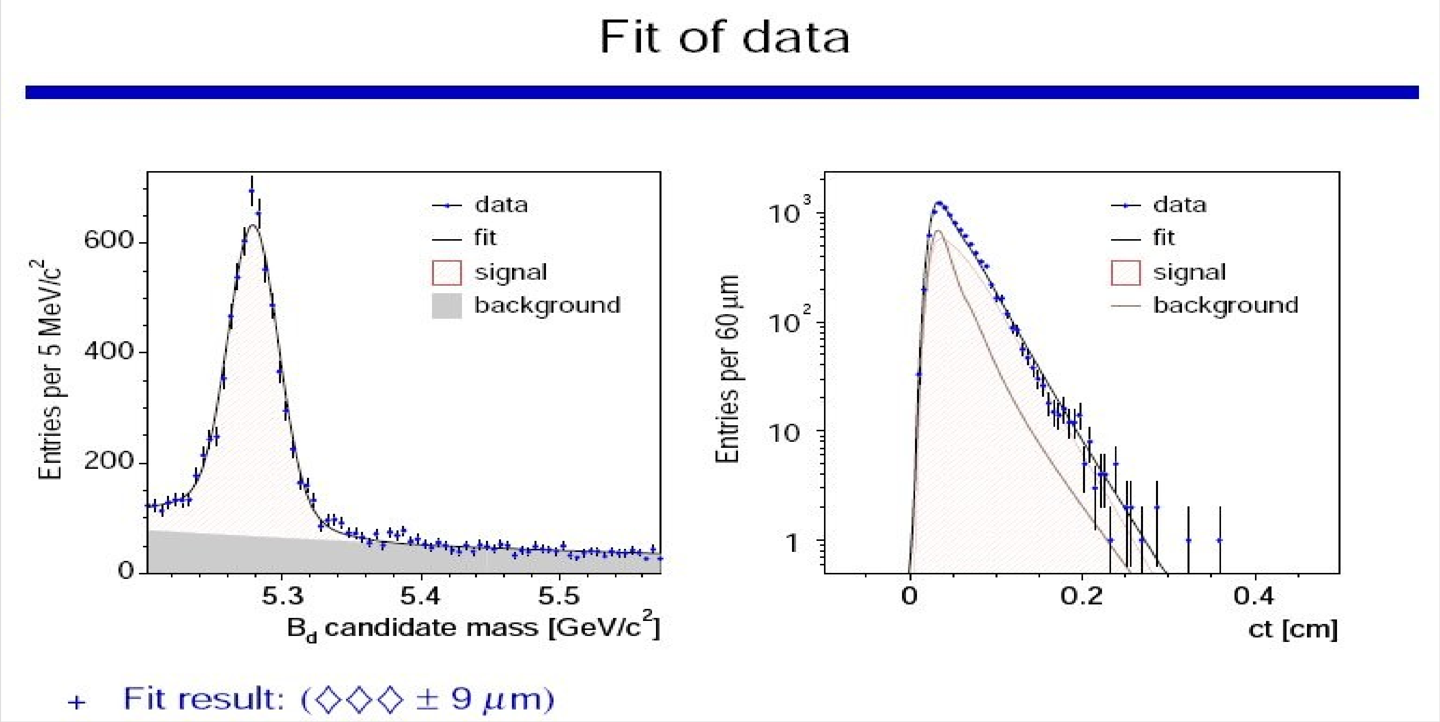
Example: $\Sigma_b \rightarrow \lambda_b(p\pi)\pi$





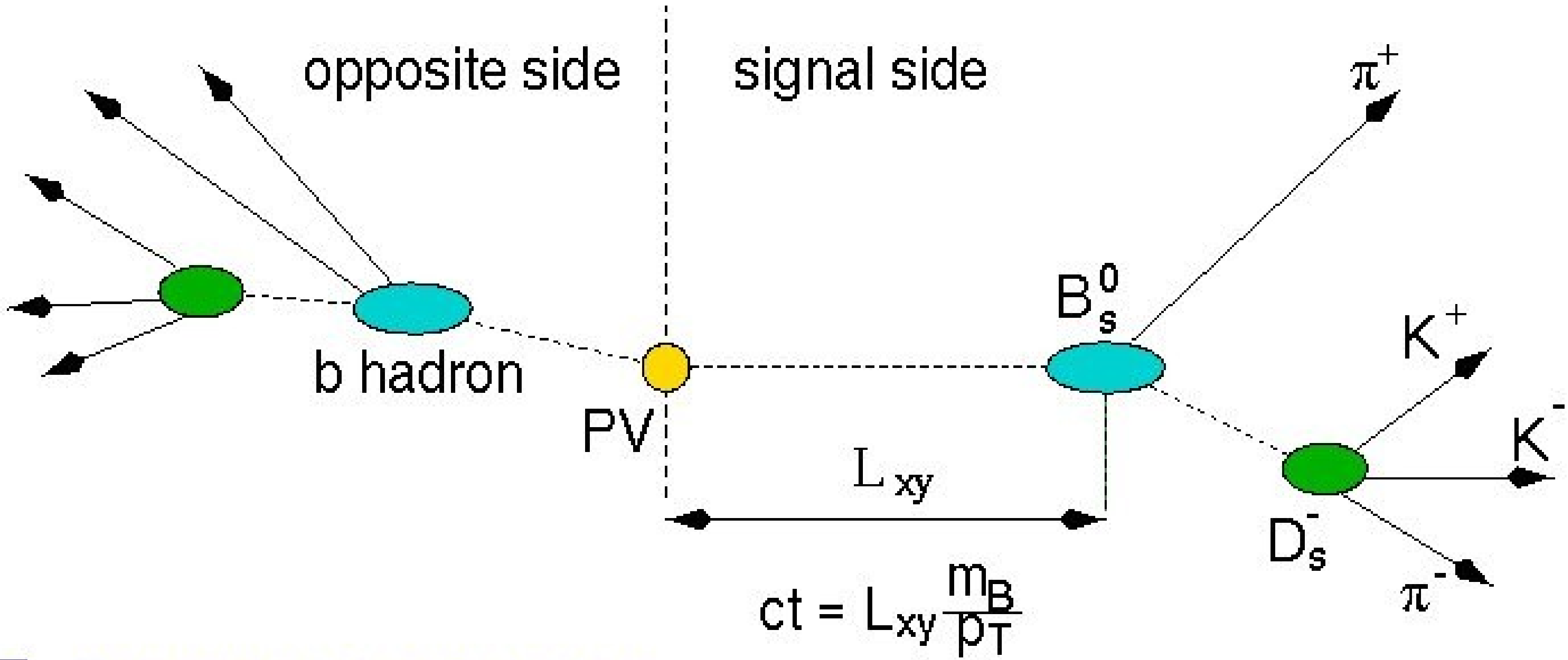
Blinding Techniques (II)

- Shifting the answer:
  - In some cases, it may be sufficient to shift the answer by adding a random (but fixed and unknown) offset to the answer. E.g. the fit for a lifetime don't fit for $\tau(B_d)$ but for $\tau(B_d) + \Delta$



- Advantage: different groups (using same offset, can compare their results. Systematics can be completely evaluated before unblinding!

Example: B_s Oscillation



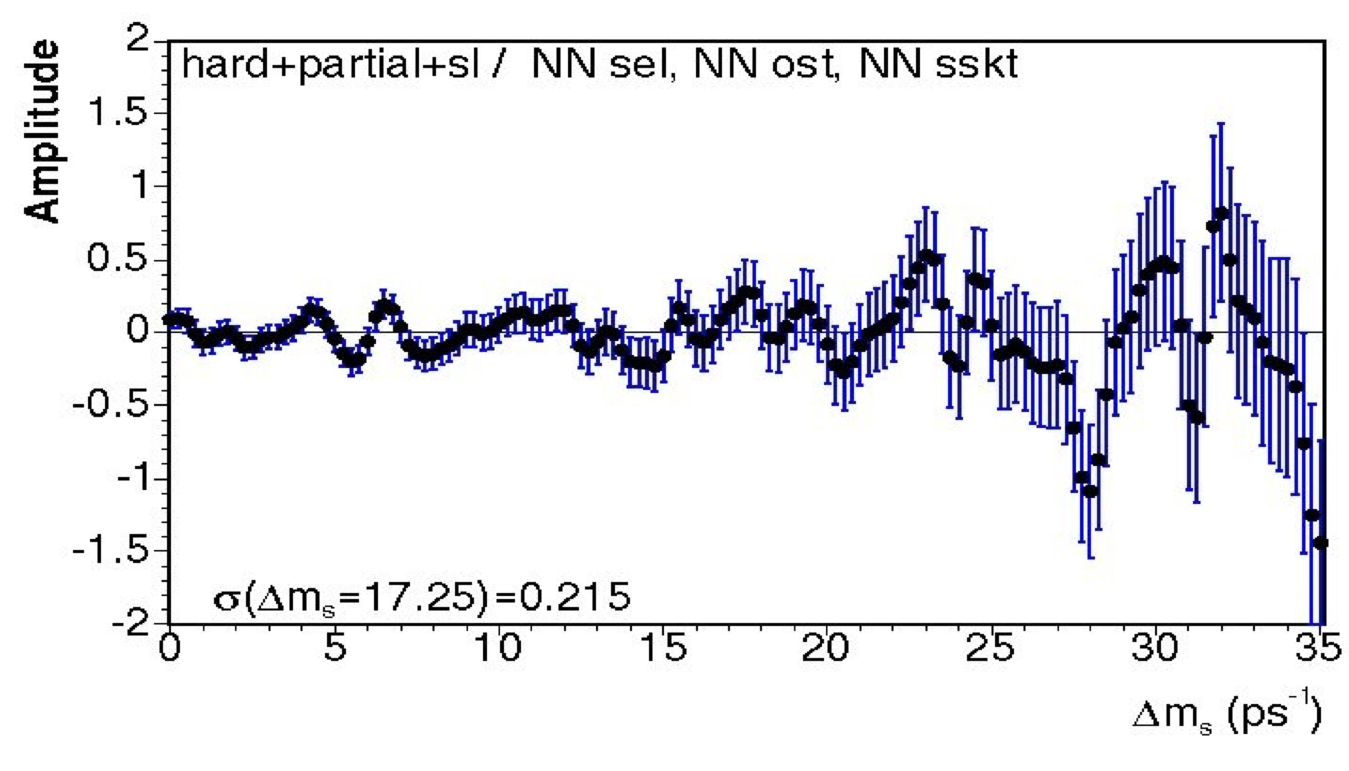
1) B_s Rekonstruktion/Selektion
2) Messung der Eigenzeit
3) Flavour tagging (Hauptproblem an Hadronen-Maschinen)

Asymmetriemessung als Funktion der Zeit:

$$\mathcal{A}(t) \equiv \frac{N(t)_{gemischt} - N(t)_{ungemischt}}{N(t)_{ungemischt} + N(t)_{gemischt}} = \mathcal{D}\cos(\Delta m_s t), \quad \mathcal{D} = 1 - 2P_{mistag}$$

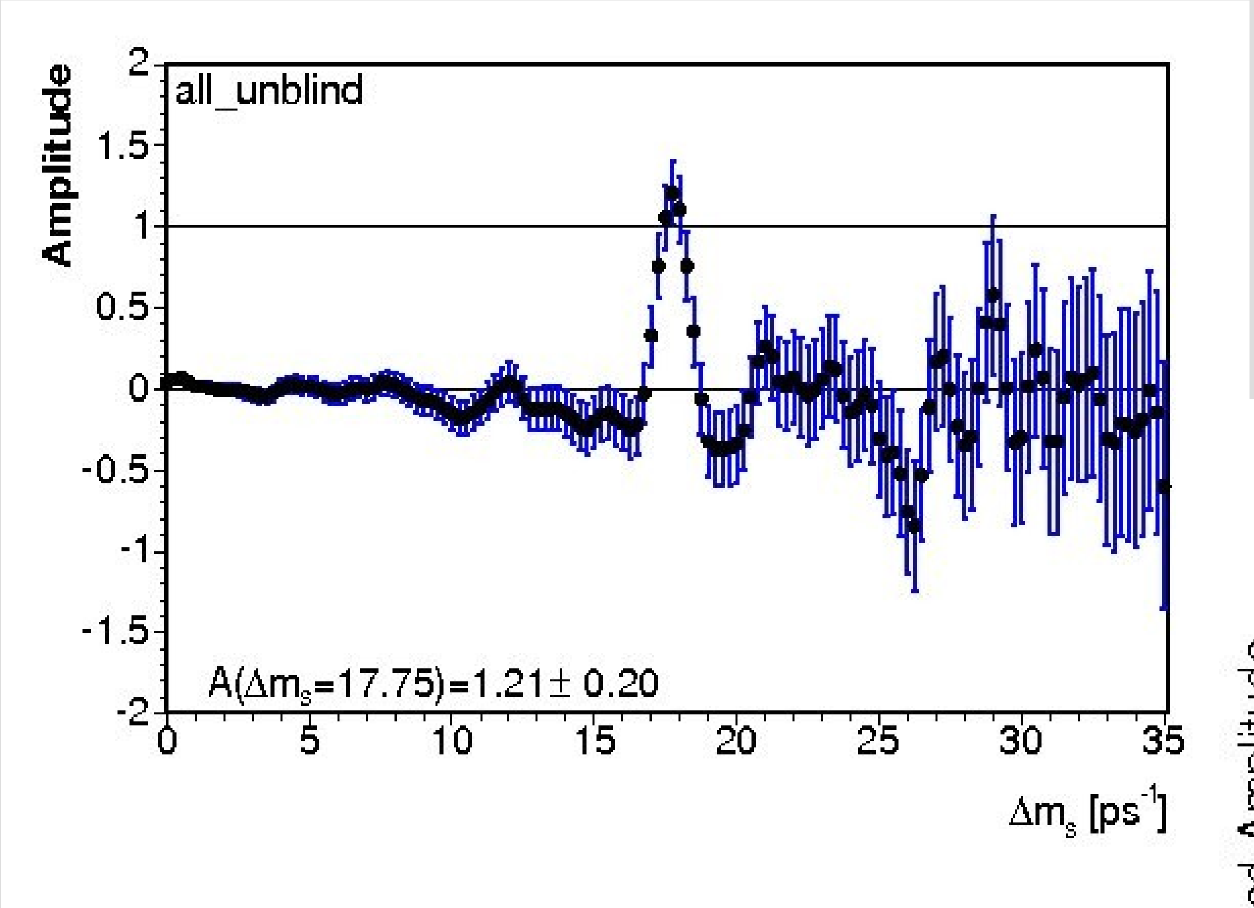
Example: B_s Oscillation

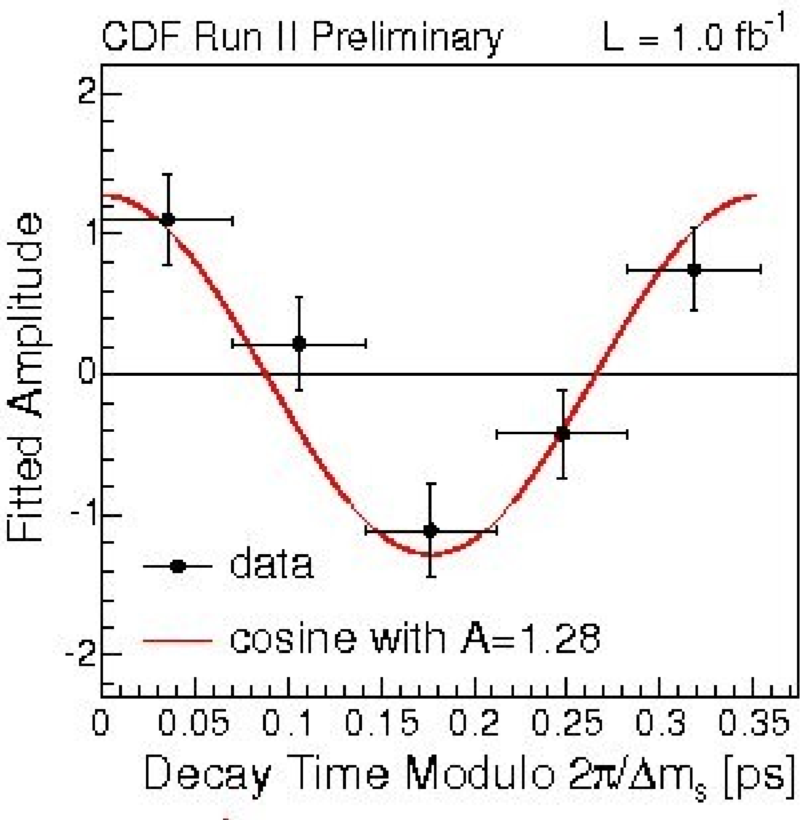
- Amplitude Scan: Fourier Transformation of frequency space – expect to be "1" at correct frequency and "0" elsewhere.
- Blinding: tagging decision multiplied by $(-1)^{event\ number}$
  Blinding method has to fit to analysis, this one special for all analysis using flavor tagging.
- Full evaluation of uncertainty done on blinded data.



Example: B_s Oscillation

- After unblinding: clear peak around $\mathbf{17.75\ ps^{-1}}$



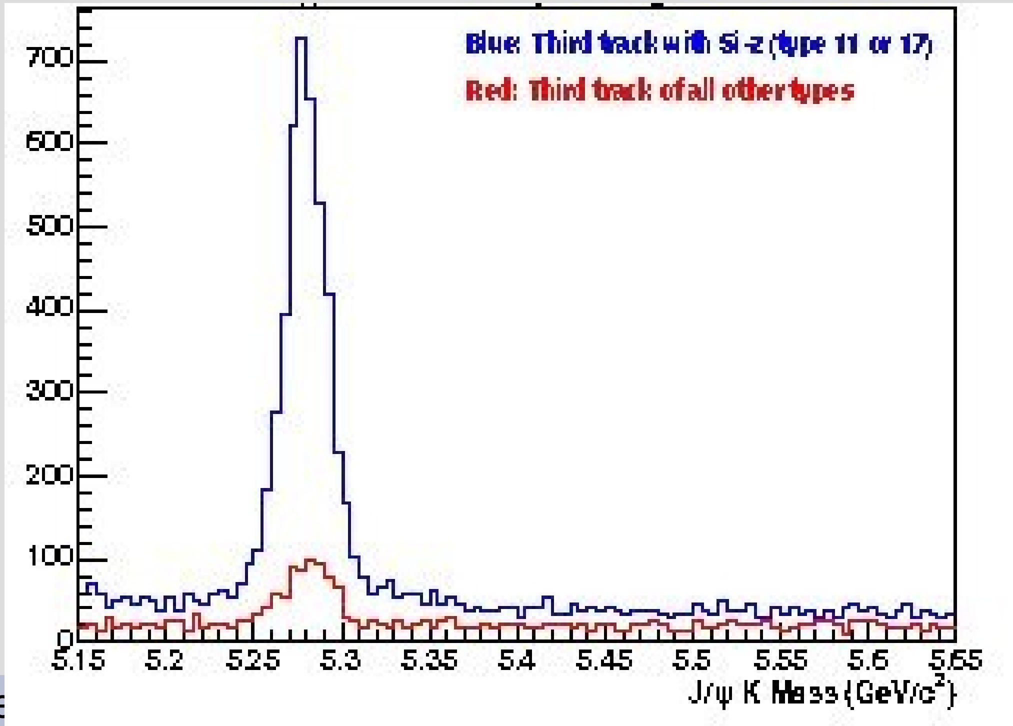


Summary of Blind Analysis

Blind analysis techniques are recommended as a way to reduce the chance for experimenter bias, thereby also reducing the rate of wrong answers. The fundamental strategy is to avoid konwing the answer until analysis procedure has been set. Since checks may lead to a change (or correction) of the procedure, they should be completed, or at least scheduled, before the answer is revealed. Despite the precautions, should a major (unanticipated, answer-changing) bug turn up after the answer is revealed, it must be explained in the publication, so that the reader can form their own opinion.

Example: $B_c \rightarrow J/\psi\pi$

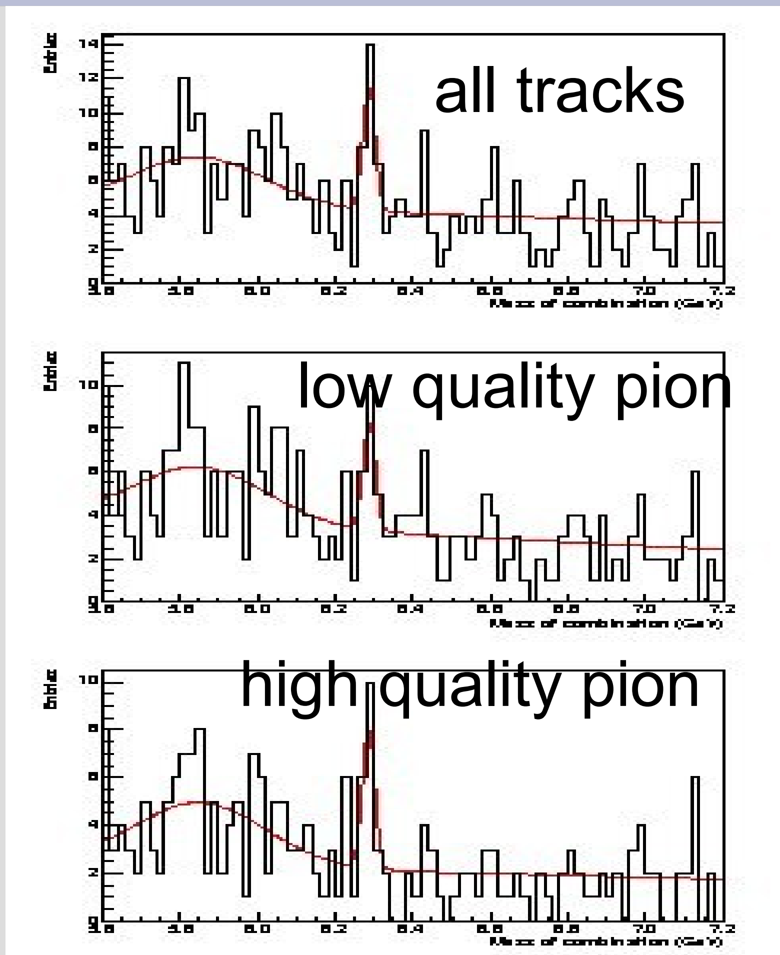
- Scan through mass region of fit (Gaussian + linear background). Width of Gaussian fixed from reference mode ($B^+ \rightarrow J/\psi K^+$) to same mass resolution.
- After unblinding, independent group discovered that 40% of all candidates in "signal" region are with a low quality pion, thus have must have a much broader mass resolution => Those events have to be a background fluctuation! The mass peak in the reference mode had only very few of those events.

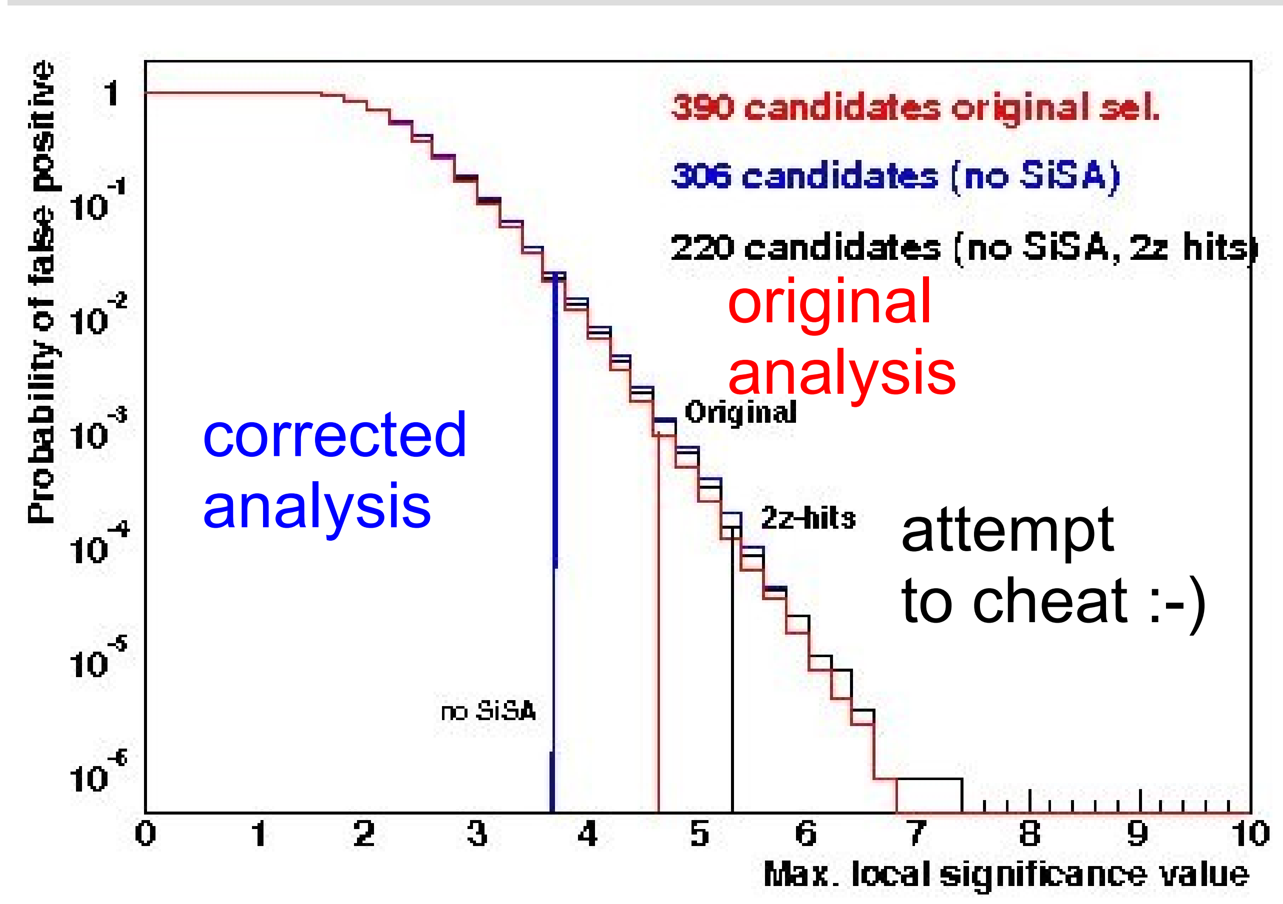


mass distribution of reference mode, including kaon tracks of all track types.

mass distribution of reference mode, including only kaon tracks of poor quality.

Example: $B_c \to J/\psi\pi$





- All events with low mass resolution were removed, new (lower) significance quoted and explicitly mentioned in the publication, that additional selection cut was introduced after unblinding.