Frequentist Analysis
Feldman-Cousins Challenges

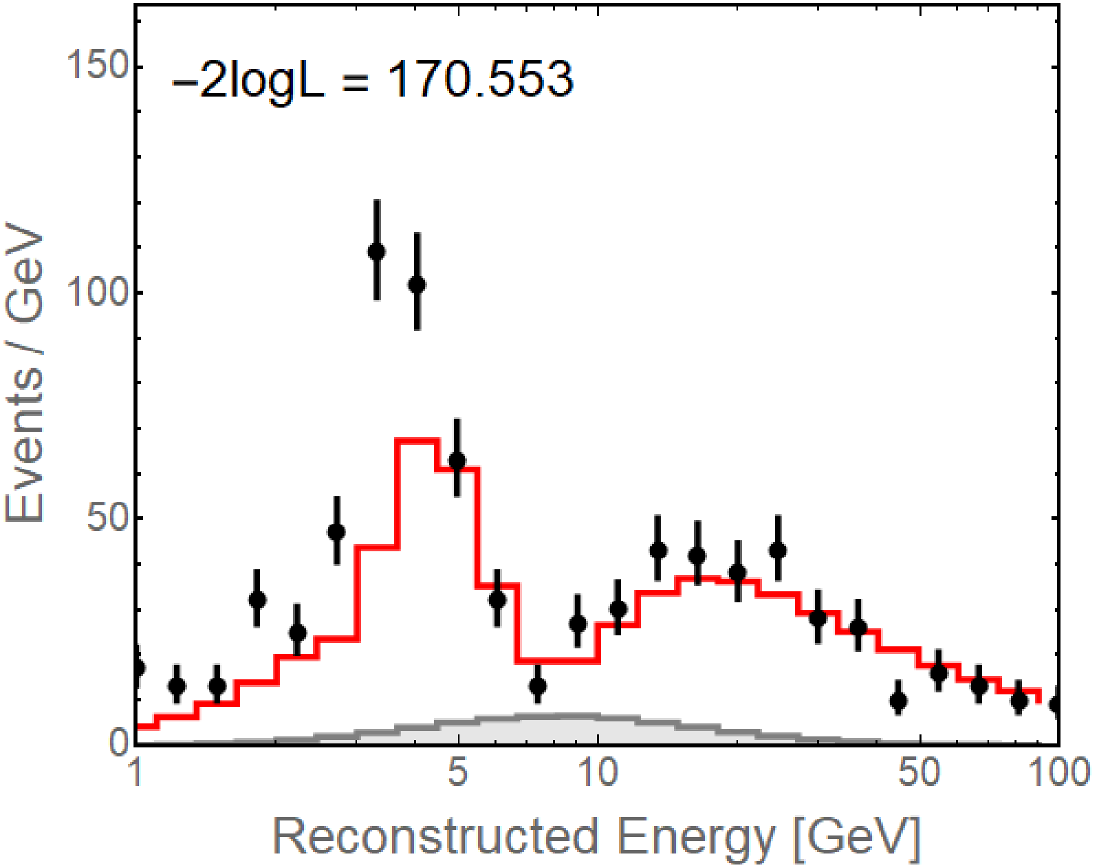
João Coelho

APC Laboratory

16 November 2023

Lets look at some data

- In general, our data doesn't agree with our predictions
- We quantify the disagreement by computing some metric
- Ideally we want a metric that is zero if data = prediction, and growing as prediction deviates from data



Log Likelihood-Ratio

- The natural choice is to use the likelihood P(data | prediction)
- In general our metric is:

$$\lambda(data, pred) = -2\log\left[\frac{P(data \mid pred)}{P(data \mid pred = data)^{*}}\right]$$

- If data is distributed as a Gaussian:

$$\lambda(d, p) = \chi^2 = \sum_i \frac{(p_i - d_i)^2}{\sigma_i^2}$$

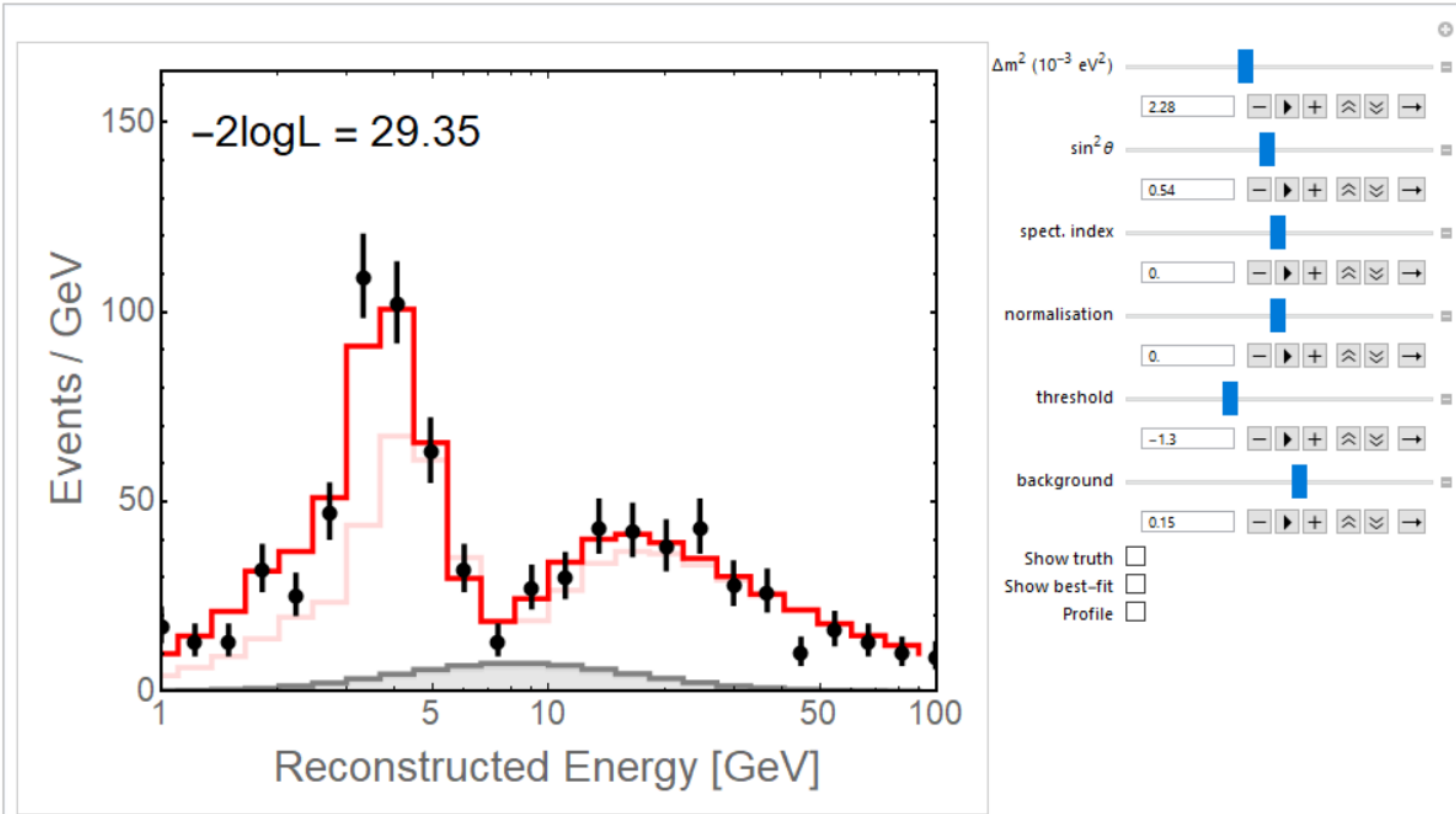
- For Poisson distributed data, this results in:

$$\lambda(d, p) = 2\sum_i p_i - d_i + d_i \; log(d_i/p_i)$$

* Technically pred = best possible prediction, which may not be =data

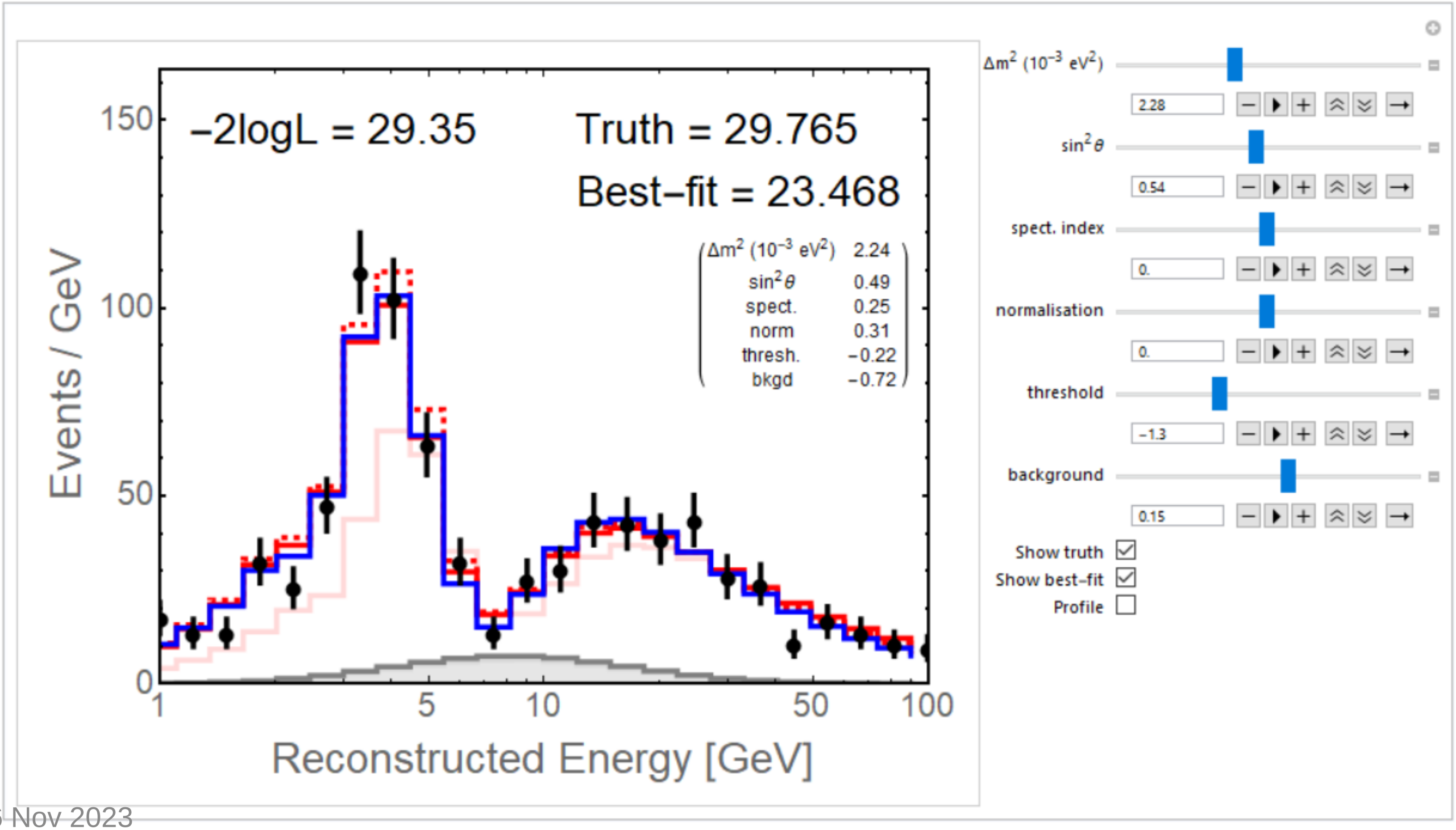
Now lets try to fix the prediction

- We can play around with multiple parameters to minimize -2logL



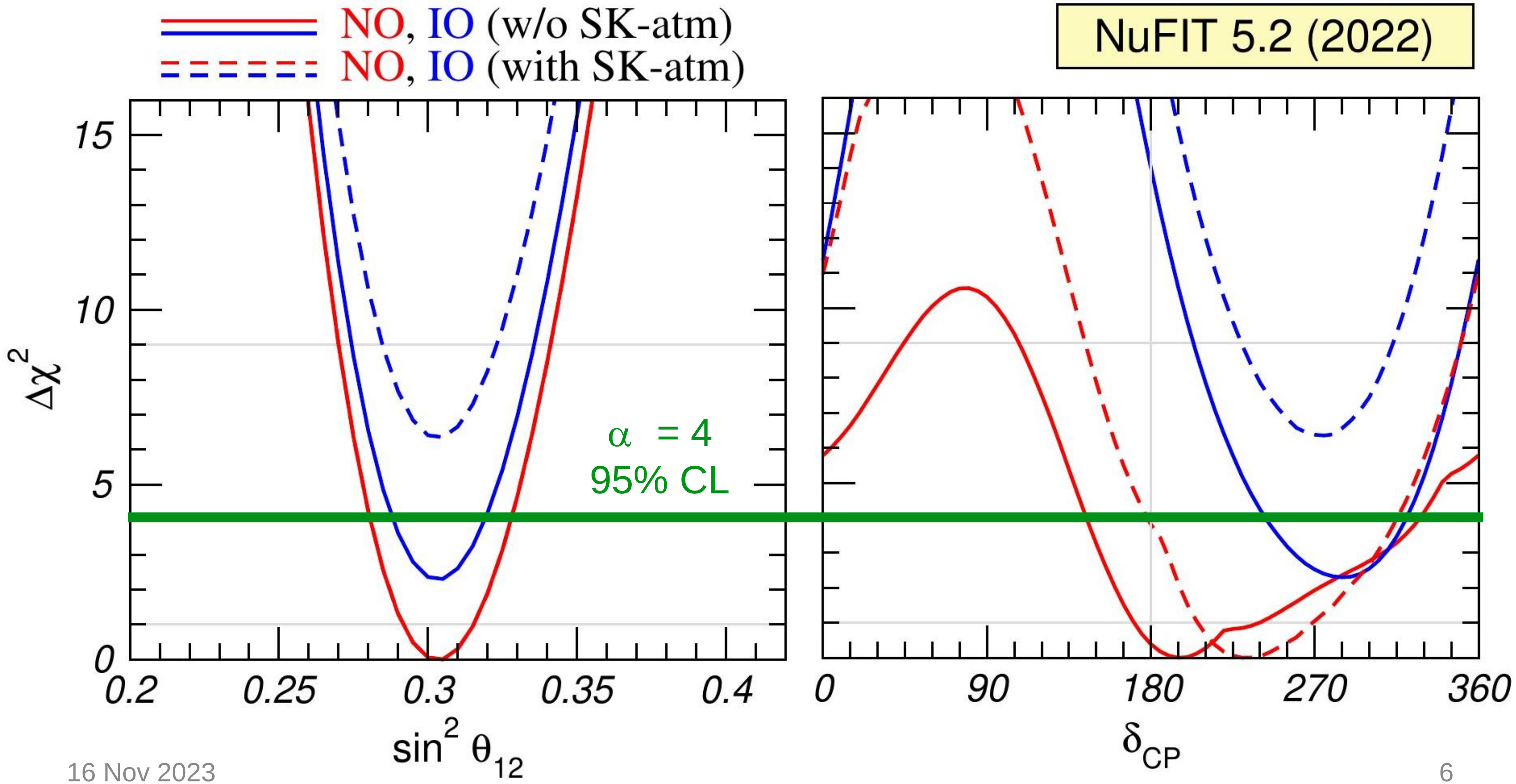
Now lets try to fix the prediction

- In practice, we use gradient descent and fit all parameters
- We can then build confidence regions around any parameter by considering what parameter values have $-2\Delta \log L < \alpha$



Building Confidence Intervals

- In practice, we use gradient descent and fit all parameters
- We can then build confidence regions around any parameter by considering what parameter values have $-2\Delta\log L < \alpha$



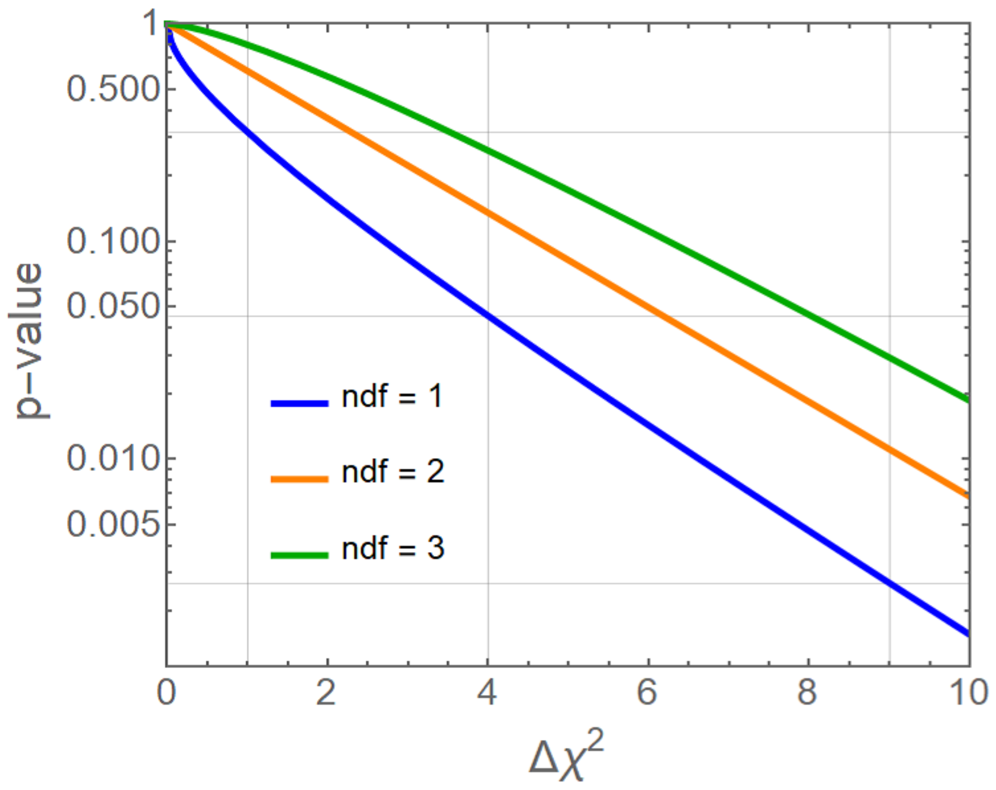
Wilks' Theorem

Theorem: *If a population with a variate x is distributed according to the probability function $f(x, \theta_1, \theta_2 \cdots \theta_h)$, such that optimum estimates $\bar{\theta}_i$ of the θ_i exist which are distributed in large samples according to (3), then when the hypothesis H is true that $\theta_i = \theta_{0i}$, $i = m + 1, m + 2, \cdots h$, the distribution of $-2 \log \lambda$, where λ is given by (2) is, except for terms of order $1/\sqrt{n}$, distributed like χ^2 with $h - m$ degrees of freedom.*

Translating:

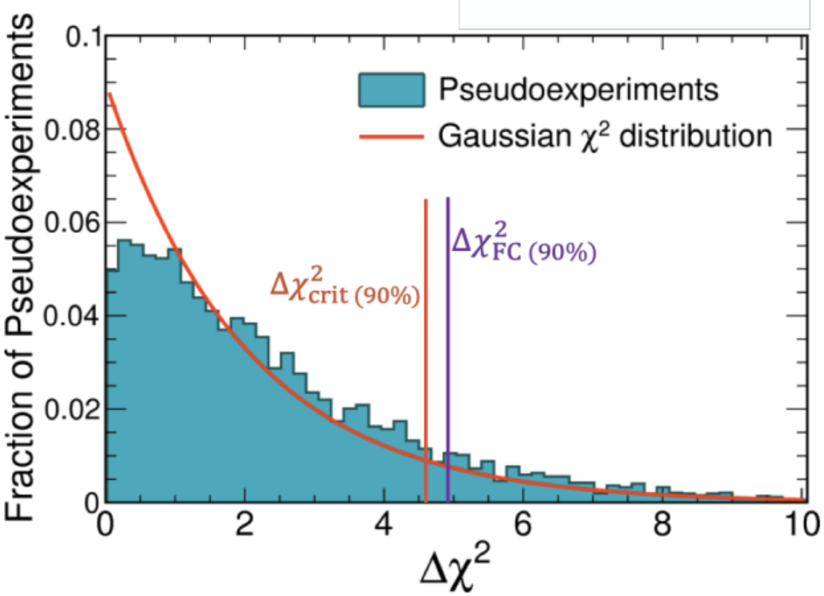
- In the limit of large samples, -2ΔlogL behaves as a χ^2 distribution

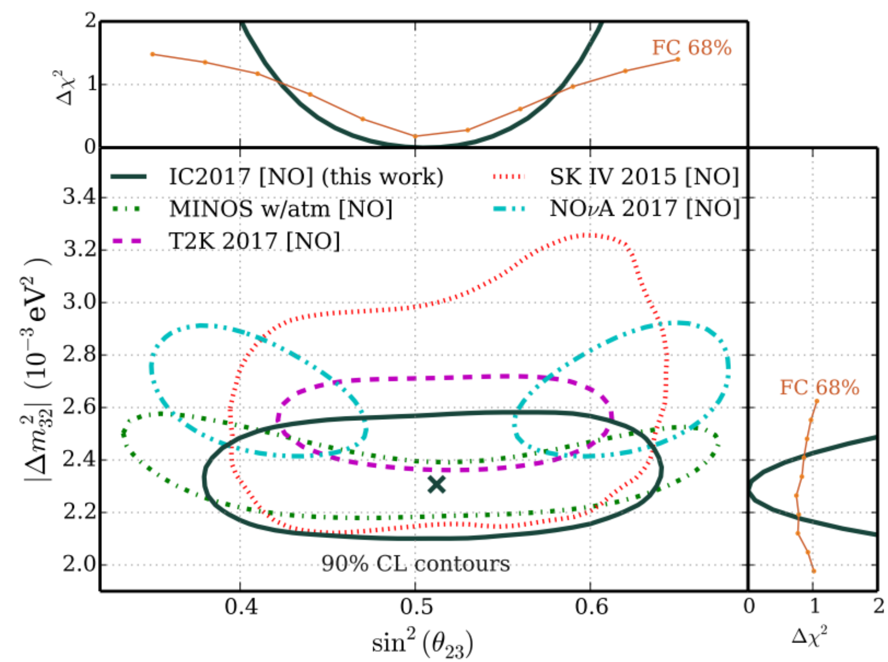
- Since we know the relationship between p-values and χ^2, we can convert easily from one to the other



Feldman-Cousins

- Sometimes Wilks' theorem fails. Typically near physical boundaries and/or when statistics are low
- The FC procedure is design to obtain CLs without assuming the -2ΔlogL distribution is χ^2 shaped
- Simulate many possible realizations of our experiment and plot their distribution
- Pick the -2ΔlogL that contains the fraction of realizations that you want





Nuisance Parameters

- One of the open questions about the FC procedure is how to deal with the nuisance parameters
- What counts as a different realization of our experiment?
- From a frequentist perspective, all parameters have a fixed true value and all pseudo-experiments should be simulated at those true values
- This works well for the parameters of interest, since we are anyway scanning them over different hypotheses

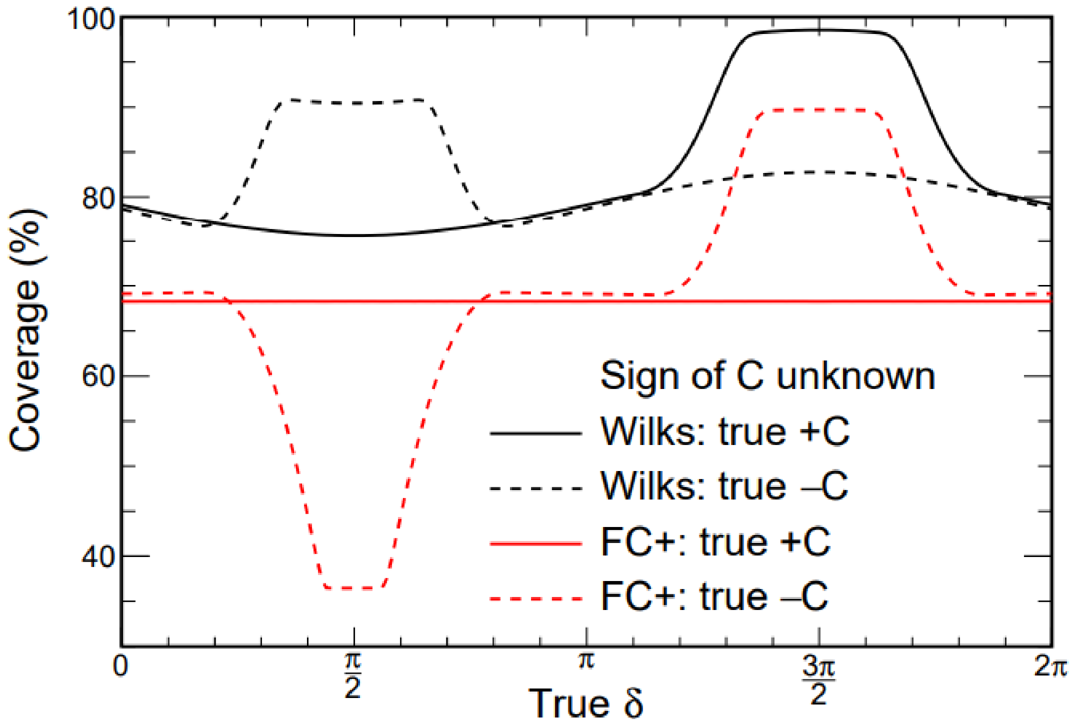
- But what true value should we assume for nuisance parameters?
  - Best knowledge before you ran your experiment?
  - Best knowledge after you ran your experiment?
  - Sample randomly from some prior distribution? (not really frequentist: Hybrid)
  - Posterior distribution? (not really frequentist: Hybrid)

- Maybe answer should be based on how well the results agree with our definition of confidence interval:
  - The interval whose construction would lead to the true value being included at the target fraction of the realizations

Nuisance Parameters

- NOvA study shows standard FC procedure can fail coverage if true value of nuisance parameters are different from the assumed valued

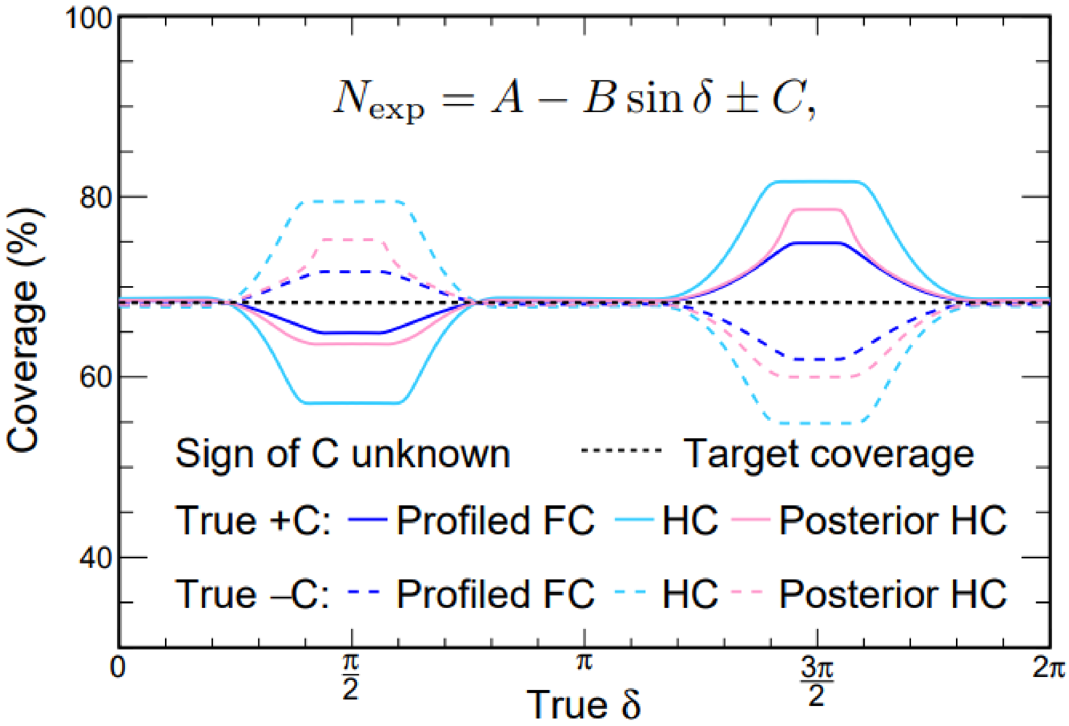
$$N_{\text{exp}} = A - B \sin \delta \pm C,$$



https://inspirehep.net/literature/2102110

Nuisance Parameters

- NOvA study shows standard FC procedure can fail coverage if true value of nuisance parameters are different from the assumed valued
- Hybrid methods do better in this case, at the expense of no perfect solution
- NOvA proposes to choose the post-fit value of nuisances at each value of the parameter of interest: Profiled FC



https://inspirehep.net/literature/2102110

Another Proposal

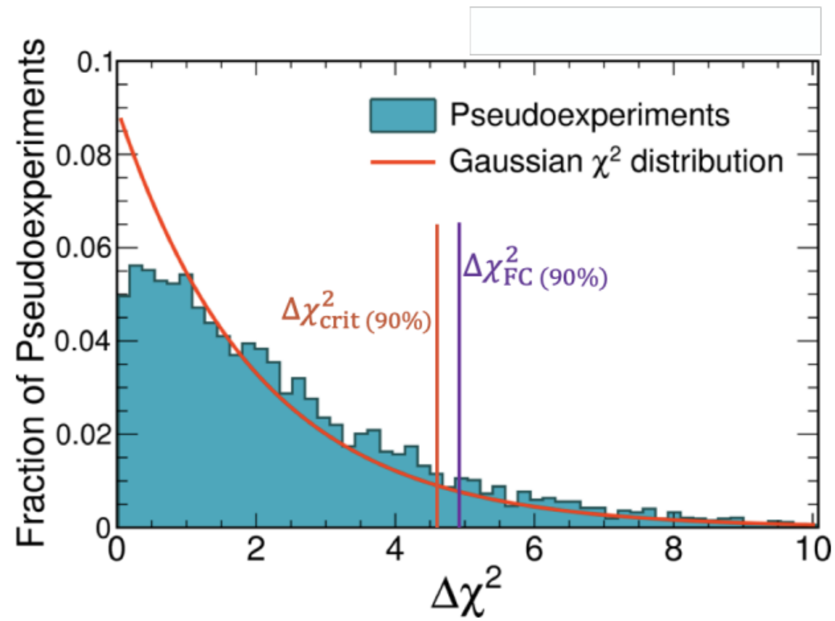
TASK: Study coverage of different methods

$$-2\Delta \log \mathcal{L} = \min_{\vec{\nu}} \sum_i \mathcal{L}(D_i(\vec{\theta}_T, \vec{\nu}_T) \,|\, M_i(\vec{\theta}_T, \vec{\nu})) - \min_{\vec{\theta},\vec{\nu}} \sum_i \mathcal{L}(D_i(\vec{\theta}_T, \vec{\nu}_T) \,|\, M_i(\vec{\theta}, \vec{\nu})) + \sum_j \frac{(\nu_j - \nu_j^0)^2}{\sigma_j^2}$$

Proposed sampling choices:

- D: Pseudo-data
  - **sampled from Poisson with mean value M**
- θ_T: True pars of interest, e.g. TauNorm
  - **Always kept fixed at test point**
- ν_T: True value of nuisance pars, e.g. θ_{13}
  - Should be fixed from freq. persp.
  - **Use best estimate, i.e. post-fit value at θ_T**
- ν^0: Mean value of our prior on ν_T
  - Represents external measurement, e.g. Daya Bay measurement of θ_{13}
  - **May be sampled as part of PE ?**
- σ: Std dev of our prior on ν_T
  - Represents uncertainty on external measurement
  - **Should be fixed at original value pre-fit**

Profiled FC



Similar, but not same as HC
Sample, ν^0 instead of ν_T

Conclusion

- Frequentist methods have been the bread and butter of particle physics statistical inference for years
- Usually liked because “no dependence on priors”

- However, interpreting the likelihood ratios can be difficult
- Feldman-Cousins procedure helps, but does not have a clear choice for dealing with nuisance parameters

- And beyond theoretical aspects, performing FC corrections typically involves performing thousands of fits at each tested point and **can be computationally prohibitive**

- Still very useful to be able to **provide both Bayesian and Frequentist results to the community**

Backup