Long-Baseline Neutrino Experiment Analysis Techniques

PhysStat-ν

Christopher Backhouse

California Institute of Technology

September 19, 2016

Overview

- LBL oscillation physics
- T2K analysis techniques
- NOvA analysis techniques
- Can we form 1D frequentist intervals for δ_{CP} with good coverage?

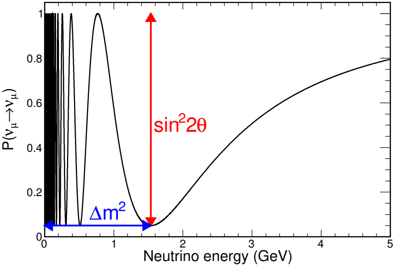
Apologies to KamLAND, MINOS, OPERA, DUNE, HyperK...

All opinions are my own, and do not reflect the views of either collaboration

LBL oscillation physics

ν_μ survival probability

- Two flavor approx. works well here
- $P_{\mu\mu} \approx 1 - \sin^2 2\theta_{23} \sin^2\left(\frac{\Delta m^2_{32} L}{4E}\right)$
- $\theta_{23} \approx 45^\circ \rightarrow$ almost all ν_μ expected to disappear at oscillation max.



$\nu_\mu \rightarrow \nu_e$ transition probability

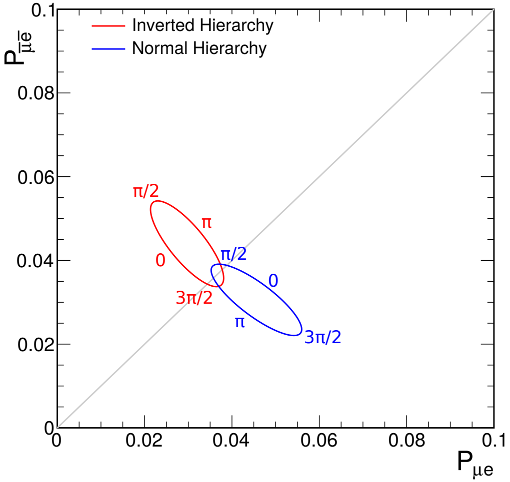
- $P_{\mu e} \approx \sin^2 2\theta_{13} \sin^2 \theta_{23} \sin^2\left(\frac{\Delta m^2_{32} L}{4E}\right) + f(\text{sign}(\Delta m^2_{32})) + f(\delta_{CP})$
- θ_{13} only 8.5° degrees, most ν_μ go to ν_τ instead
- Look for deviations due to hierarchy (matter effects) and CP-violation

$\times 2$ **for antineutrinos**

Principle of the ν_e measurement

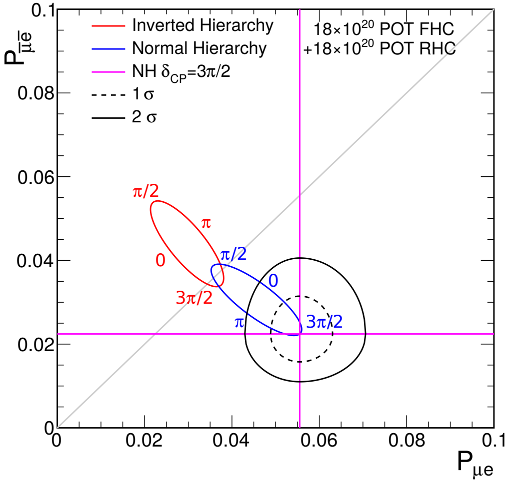
- To first order, NOvA measures $P(\nu_\mu \to \nu_e)$ and $P(\bar{\nu}_\mu \to \bar{\nu}_e)$ evaluated at 2GeV

- These depend differently on sign(Δm^2_{32}) and δ_{CP}



Principle of the ν_e measurement

- To first order, NOvA measures $P(\nu_\mu \to \nu_e)$ and $P(\bar{\nu}_\mu \to \bar{\nu}_e)$ evaluated at 2GeV
- These depend differently on sign(Δm^2_{32}) and δ_{CP}
- Ultimately constrain to some region of this space

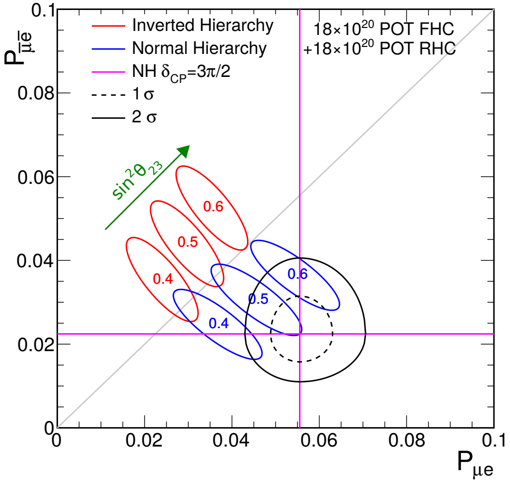


Principle of the ν_e measurement

- To first order, NOvA measures $P(\nu_\mu \to \nu_e)$ and $P(\bar{\nu}_\mu \to \bar{\nu}_e)$ evaluated at 2GeV
- These depend differently on $\text{sign}(\Delta m^2_{32})$ and δ_{CP}
- Ultimately constrain to some region of this space
- P also $\propto \sin^2\theta_{23}$

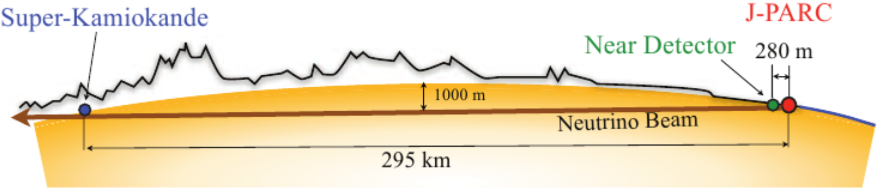
  $<$ 0.5: "lower octant"

  $>$ 0.5: "upper octant"



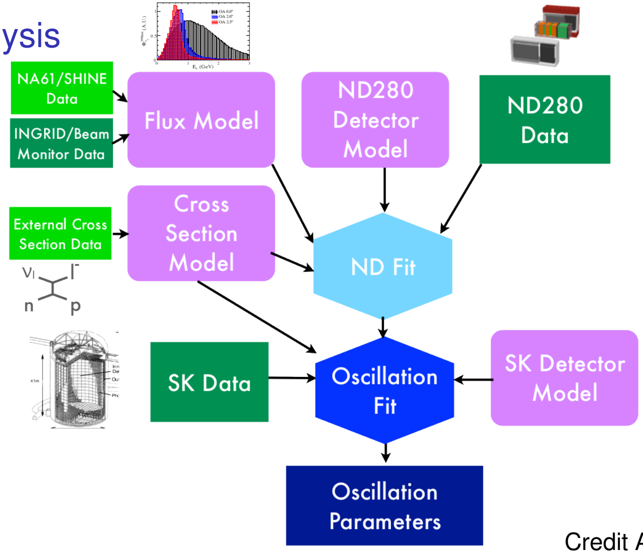
T2K

T2K overview



- $\nu_\mu \to \nu_e \quad \nu_\mu \to \nu_e \quad \bar{\nu}_\mu \to \bar{\nu}_\mu$ and $\bar{\nu}_\mu \to \bar{\nu}_e$
- Cross-section and flux constraints from Near Detector (ND280) and external experiments (NA61/SHINE)

T2K analysis

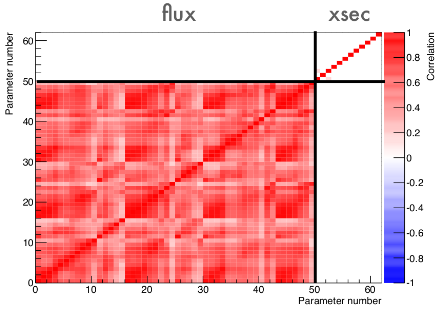


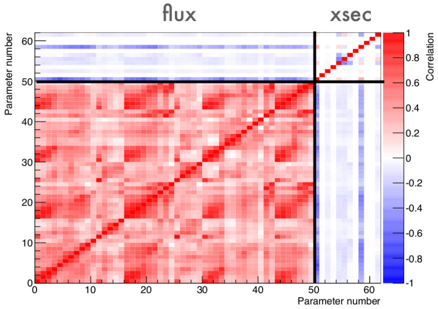
Credit Asher Kaboth

- Constrain parameters in xsec/flux model using ND280 and external data
- Appropriate if model knobs fully cover possibilities in reality

Error matrix

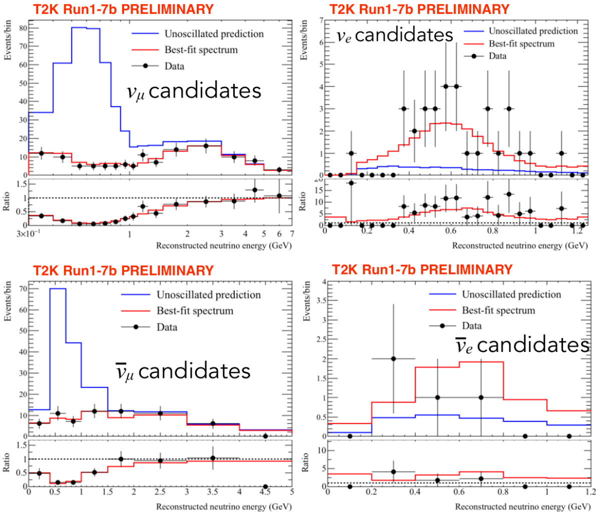
- Correlation matrix constrained by fit to ND data
- See upcoming VALOR talks





- Use of a correlation matrix appropriate if parameter measurements are gaussian

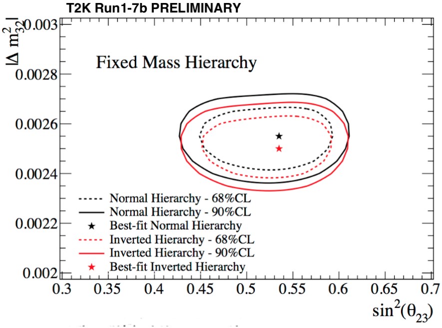
T2K FD data

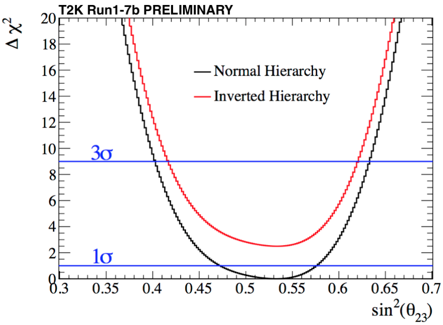


Multiple analysis approaches

- Frequentist $\Delta\chi^2$ fit
  - Profile over systematics
- Bayesian lhood fit
- Bayesian MCMC, simultaneous with ND

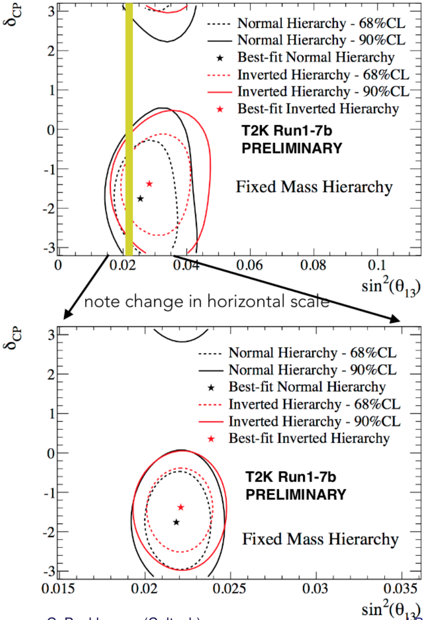
T2K results





- This parameter pairing dominated by ν_μ survival
- Bread-and-butter contour in frequentist stats, gaussian limit

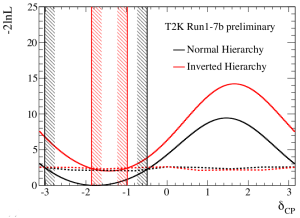
T2K results



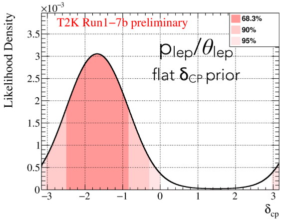
- Fixed gaussian $\Delta\chi^2_{\text{crit}}$ (“up value”)
- Analyze each hierarchy independently
- Some gain from including external reactor θ_{13} constraint

T2K δ_{CP} ranges

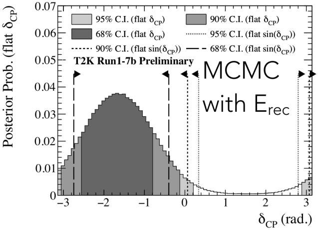
FC 90% C.L. crit. values



Bayes, prior flat in δ



Bayes, priors compared

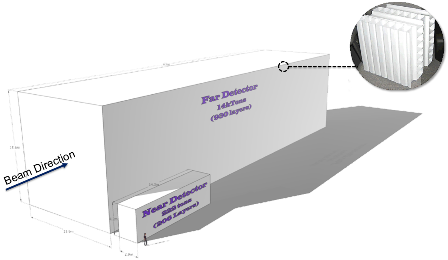


- Results from different approaches similar, not identical
- Maximal θ_{23} and minimal sensitivity to hierarchy help consistency?

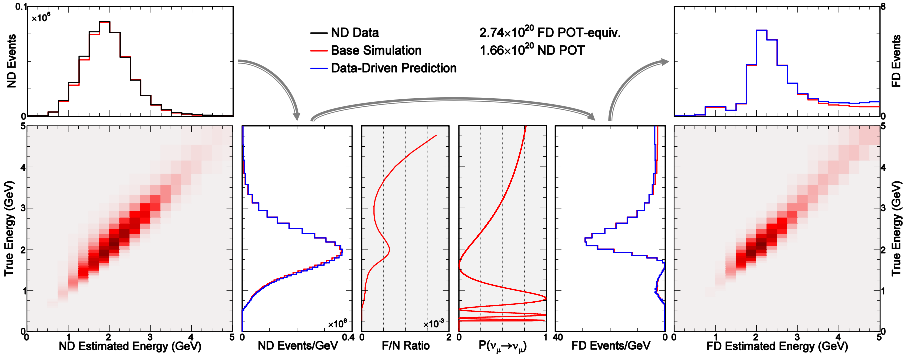
NOvA overview

- $\nu_\mu \to \nu_\mu$ and $\nu_\mu \to \nu_e$ channels
- $\bar{\nu}_\mu \to \bar{\nu}_\mu$ and $\bar{\nu}_\mu \to \bar{\nu}_e$ soon

- ND and FD are functionally identical

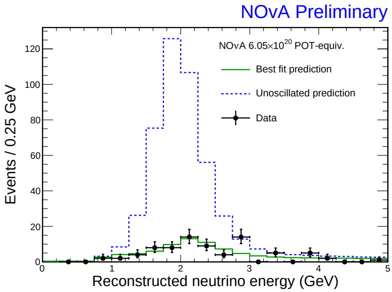


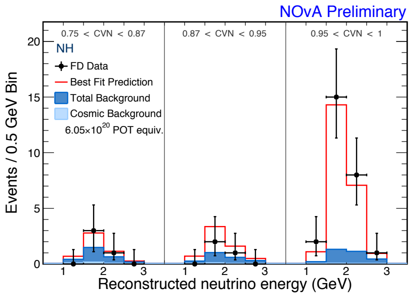
NOvA FD prediction

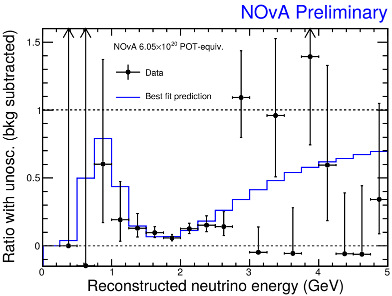


- "Extrapolate" ND data to FD prediction (via plenty of Monte Carlo)
- Assess systematics by varying MC and pushing through the whole chain
- Still some hand tweaking of parameters based on ND observations
- Should be more robust against unknown unknowns

NOvA data







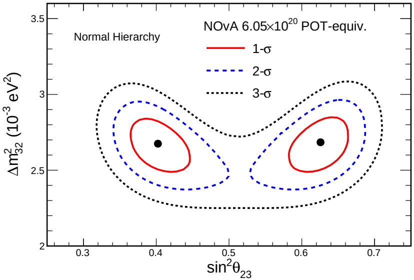
- Log-likehood fit

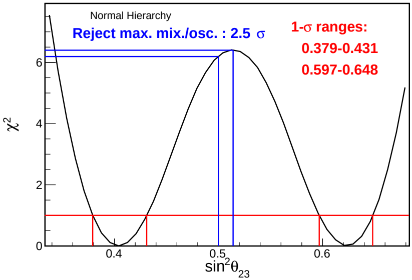
$$\mathcal{L}(N|\lambda) = \frac{\lambda^N e^{-\lambda}}{N!}$$

$$\Delta\chi^2 = -2\ln\frac{\mathcal{L}(N|\lambda)}{\mathcal{L}(N|N)}$$

$$= 2\left(\lambda - N + N\ln\frac{N}{\lambda}\right)$$

NOvA ν_μ results

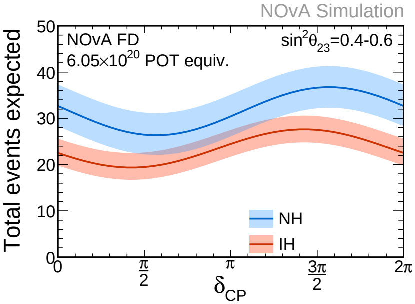




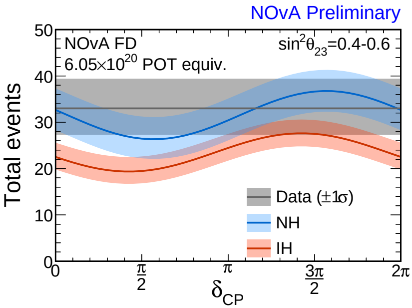
- Constant $\Delta\chi^2_{\text{crit}}$ shown here
- Systematic parameters profiled over
- FC corrections have minimal impact

- Prefer non-maximal mixing, at what sig. exactly do we reject maximal?
- Evaluate FC experiments at $\sin^2\theta_{23} = 0.5$, best fit Δm^2 given this θ_{23}
- Slightly increase rejection power

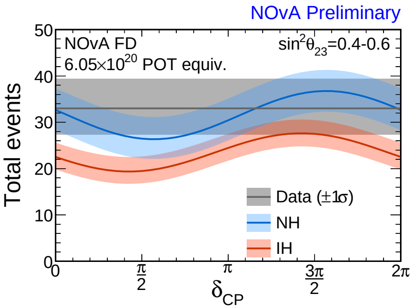
NOvA ν_e results

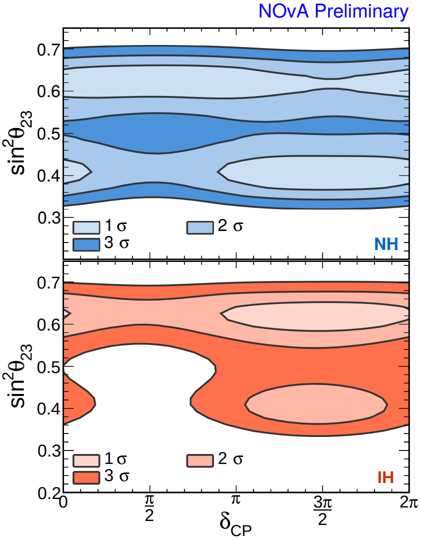


NOvA ν_e results



NOvA ν_e results





- Lots of interesting parameter correlations
- Extracted δ_{CP} conclusions depend on what you do with the other parameters

An interesting case study

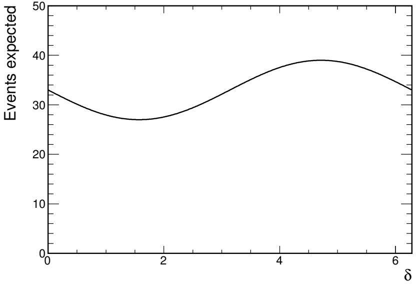
Coverage

- Frequentist coverage means: "if the true value of parameter *x* is *A*, 68% of experiments will include *A* in their confidence interval for *x*"
- FC procedure achieves this almost tautologously by throwing mock experiments at each *A* and finding the $\Delta\chi^2_{\text{crit}}$ that would have included that *A* in 68% of the experiments

- In the presence of a parameter *y* not displayed on the plot (a "nuisance parameter")
- Want correct coverage *no matter the true value of that parameter*

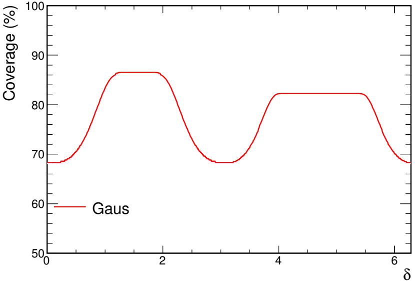
- Obviously impossible in general, infinite array of possible values for *y*, all requiring different critical values in principle
- But *e.g.* for two gaussian variables profiling over *y* gives correct coverage, even without invoking FC corrections
- So how does it work out in practice for our experiment?

The toy



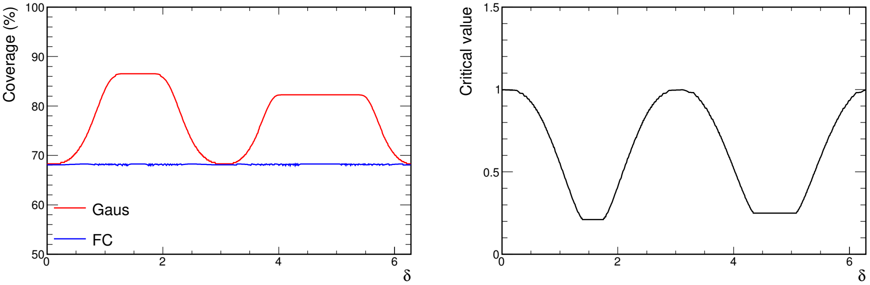
- Model δ_{CP} behaviour, neglect hierarchy and octant
- Expected number of events $= 33 - 6 \sin \delta$
- Throw experiments as gaussian numbers $N \pm \sqrt{N}$
- Eliminates complications from discontinuous event counts
- Can run full set of experiments in seconds

Results



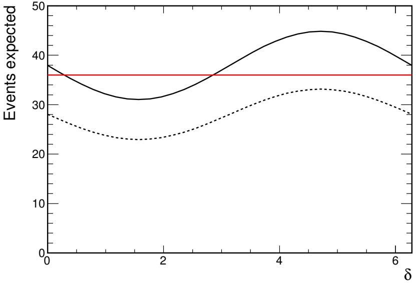
- Construct confidence intervals for many mock expts, evaluate coverage
- “Gaus” ($\Delta\chi^2_{\text{crit}} = 1$) works far from extremes
- *i.e.* when χ^2_{best} will be zero
- Signficantly overcovers elsewhere

Results



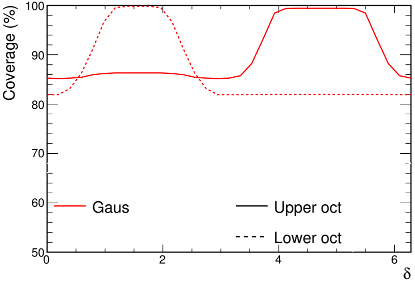
- Construct confidence intervals for many mock expts, evaluate coverage
- "Gaus" ($\Delta\chi^2_{\text{crit}} = 1$) works far from extremes
- *i.e.* when χ^2_{best} will be zero
- Signficantly overcovers elsewhere, big FC correction required
- Correct FC coverage, as expected

Upgraded toy



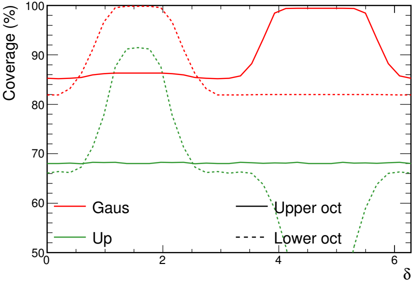
- Number expected $= (0.8 \text{ or } 1.2)(33 - 6\sin\delta)$
- Modelled after octant
- People are more willing to separate results by hierarchy, but want θ_{23} to be "profiled out"
- Goal is to make correct intervals in δ independent of true octant
- Red line shows one example experiment

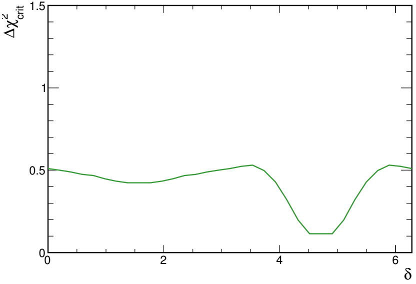
Critical value strategies



- "Gaus" ($\Delta\chi^2_{\rm crit} = 1$) heavily overcovers in all cases

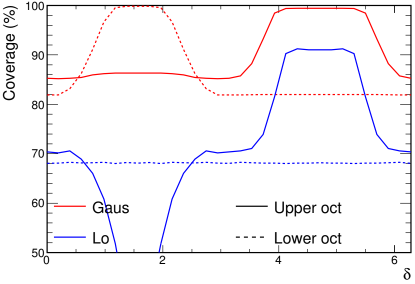
Critical value strategies

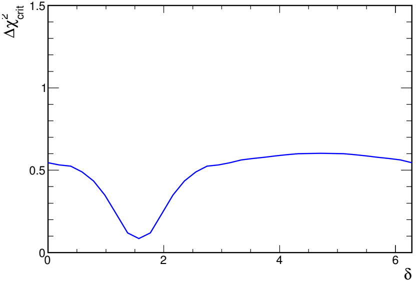




- "Gaus" ($\Delta\chi^2_{\text{crit}} = 1$) heavily overcovers in all cases
- "Up" throws all FC experiments from the upper octant
- Obviously perfect for upper octant, still very bad for lower

Critical value strategies





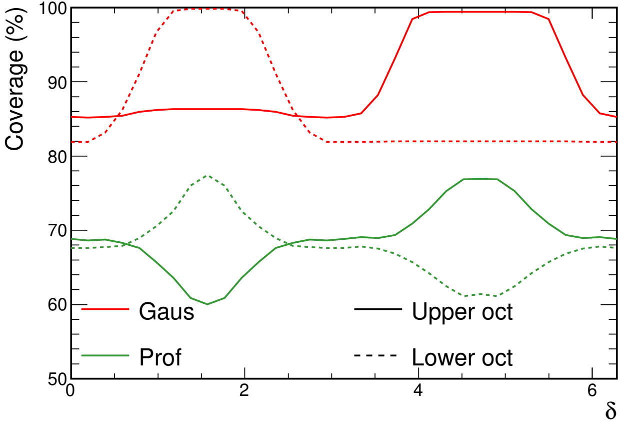
- "Lo" throws all experiments from the lower octant
- See how the necessary $\Delta\chi^2_{\text{crit}}$ differs from "Up"

“Profile” method

- How can we possibly satisfy the needs of both true octants?

- A possible loophole: allow $\Delta\chi^2_{\text{crit}}$ to depend on the observed data
- For each δ throw experiments in the octant the data favour
- Still will sometimes use $\Delta\chi^2_{\text{crit}}$ for the wrong octant, but may be rare enough?
- Call this method “Prof”

Critical value strategies



- Coverage properties are better, still not good enough to make people comfortable

Crazy ideas

- One can of course always guarantee no undercoverage by using the largest $\Delta\chi^2_{\text{crit}}$ for any true value of the suppressed variable
- Substantially understating the power of the experiment is not popular

Crazy ideas

- One can of course always guarantee no undercoverage by using the largest $\Delta\chi^2_{\text{crit}}$ for any true value of the suppressed variable
- Substantially understating the power of the experiment is not popular

- In this very specific case one could balance the competing needs of lower and upper octant by carefully picking the two ends of a range in *N* that you'll accept for each δ
- Not generic
- Gives up all of the benefits of using $\Delta\chi^2$ as the ordering criterion

Pragmatism

- No satisfactory way to “integrate out” hierarchy or octant possible
- Continue to plot four curves

- Problem really stems from large impact and bimodality of θ_{23}
- Studies beyond the scope of this toy show profiling over θ_{23} but constrained within a particular octant works much better
- For other parameters approximation that $\Delta\chi^2_{\text{crit}}$ does not depend on them is far better

- ν_μ contours much better behaved
- θ_{23} bimodal, but so degenerate it doesn’t matter

Conclusion

- Variety of ways to incorporate ND / external constraints
- Mix of Bayesian and frequentist approaches to set limits
- Starting to want to accept/reject specific points as well as provide a range
- Convolutions of oscillation formulae can provide interesting torture tests