Experimental Lecture #6
Statistics

D. Jason Koskinen

NBIA PhD School: Neutrinos Underground and in the Heavens
June 23-27, 2014

Niels Bohr Institutet

The Niels Bohr International Academy

centre for particle physics Discovery

Statistical Tests

Statistical Tests

- Chi-squared tests

Statistical Tests

- Chi-squared tests

$$\chi^2 = \sum \frac{(Observed - Expected)^2}{Expected}$$

Statistical Tests

- Chi-squared tests

$$\chi^2 = \sum \frac{(Observed - Expected)^2}{Expected}$$

- Many different permutations for a Figure Of Merit (FOM), and a quick modification of χ^2 is a nice tool to have when seeing new results

$$\chi^2 = \sum \frac{(Observed - Expected)^2}{Expected}$$

Statistical Tests

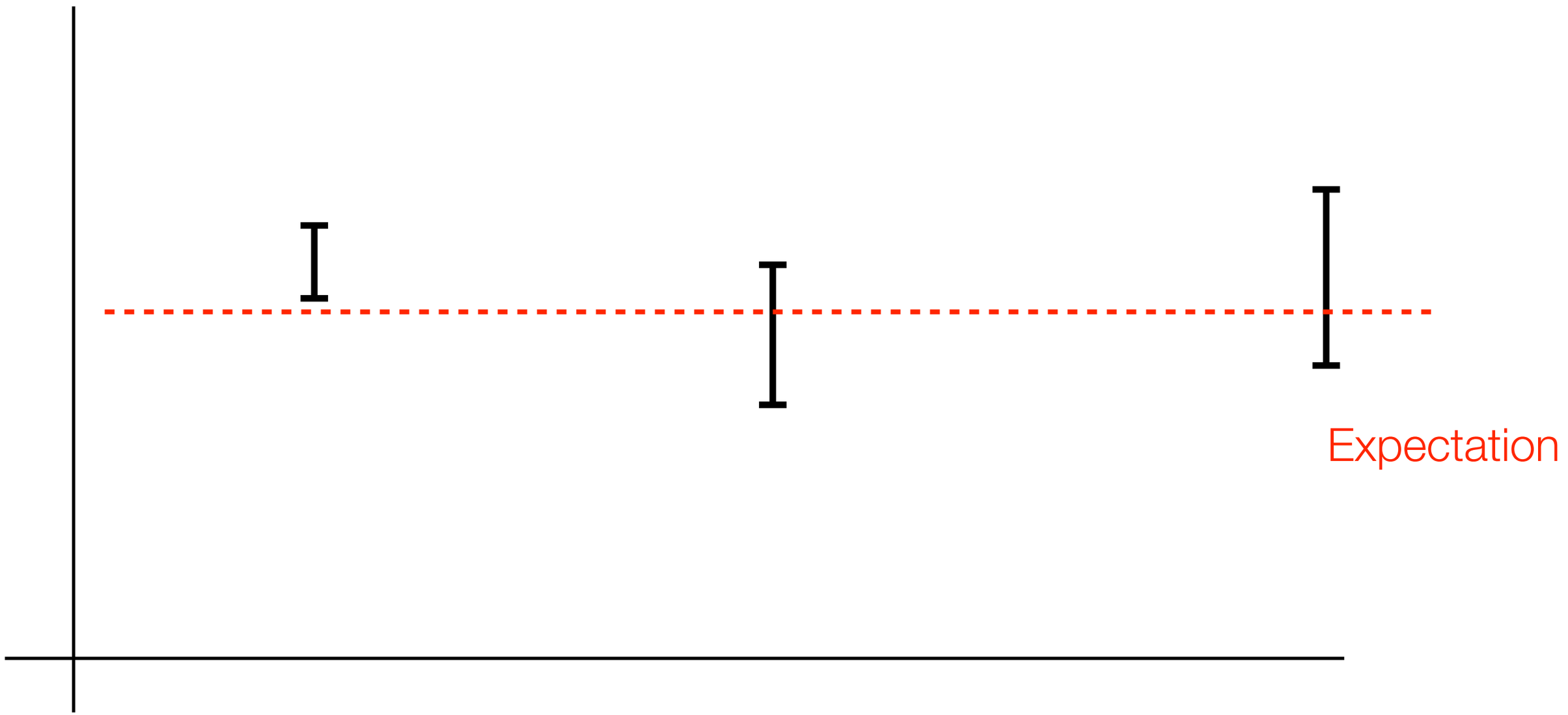
- Chi-squared tests

$$\chi^2 = \sum \frac{(Observed - Expected)^2}{Expected}$$

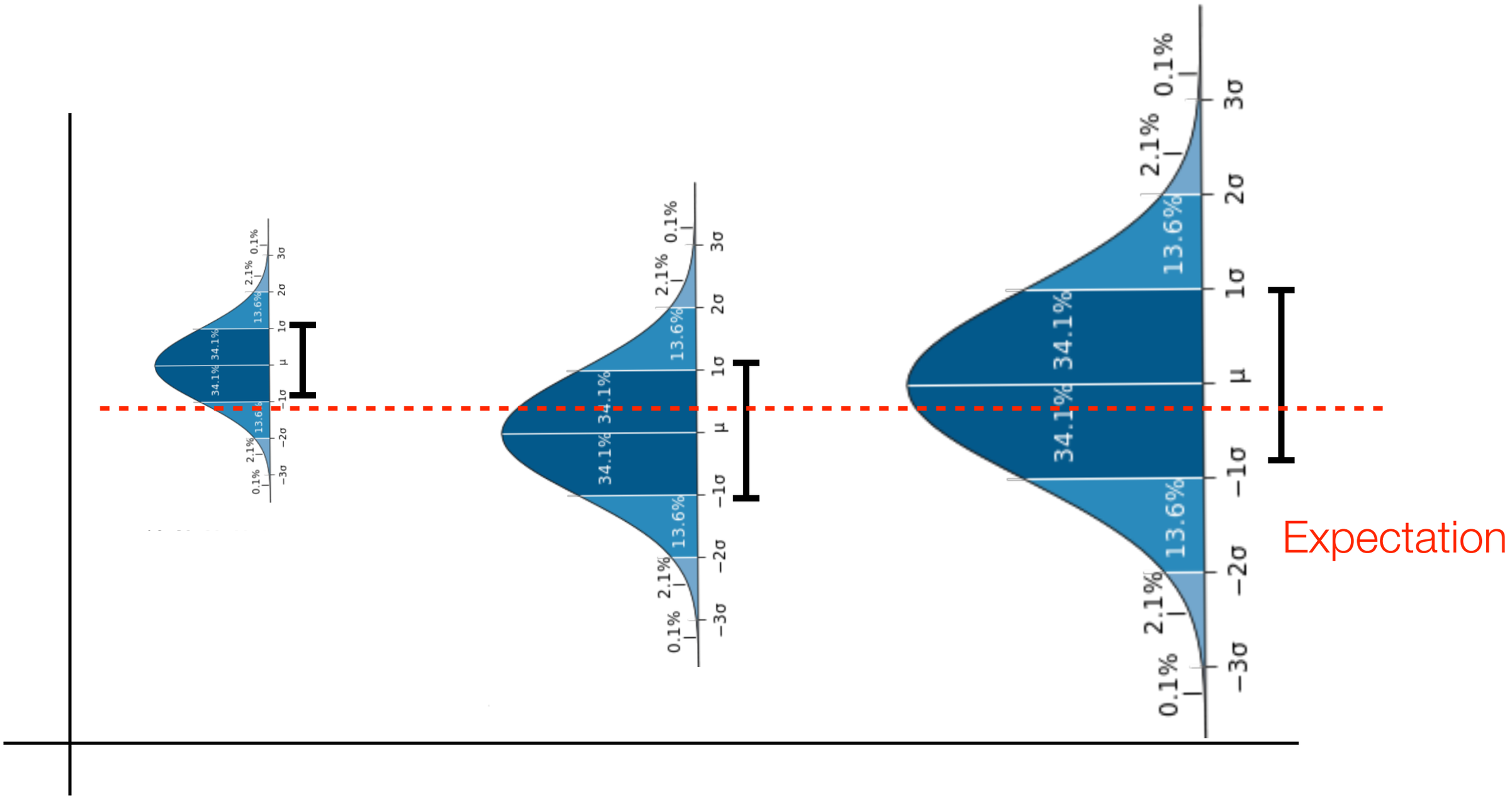
- Many different permutations for a Figure Of Merit (FOM), and a quick modification of χ^2 is a nice tool to have when seeing new results

$$\chi^2 = \sum \frac{(Observed - Expected)^2}{Expected + \sigma^2_{systematics}}$$

Basic Reduced Chi-Square

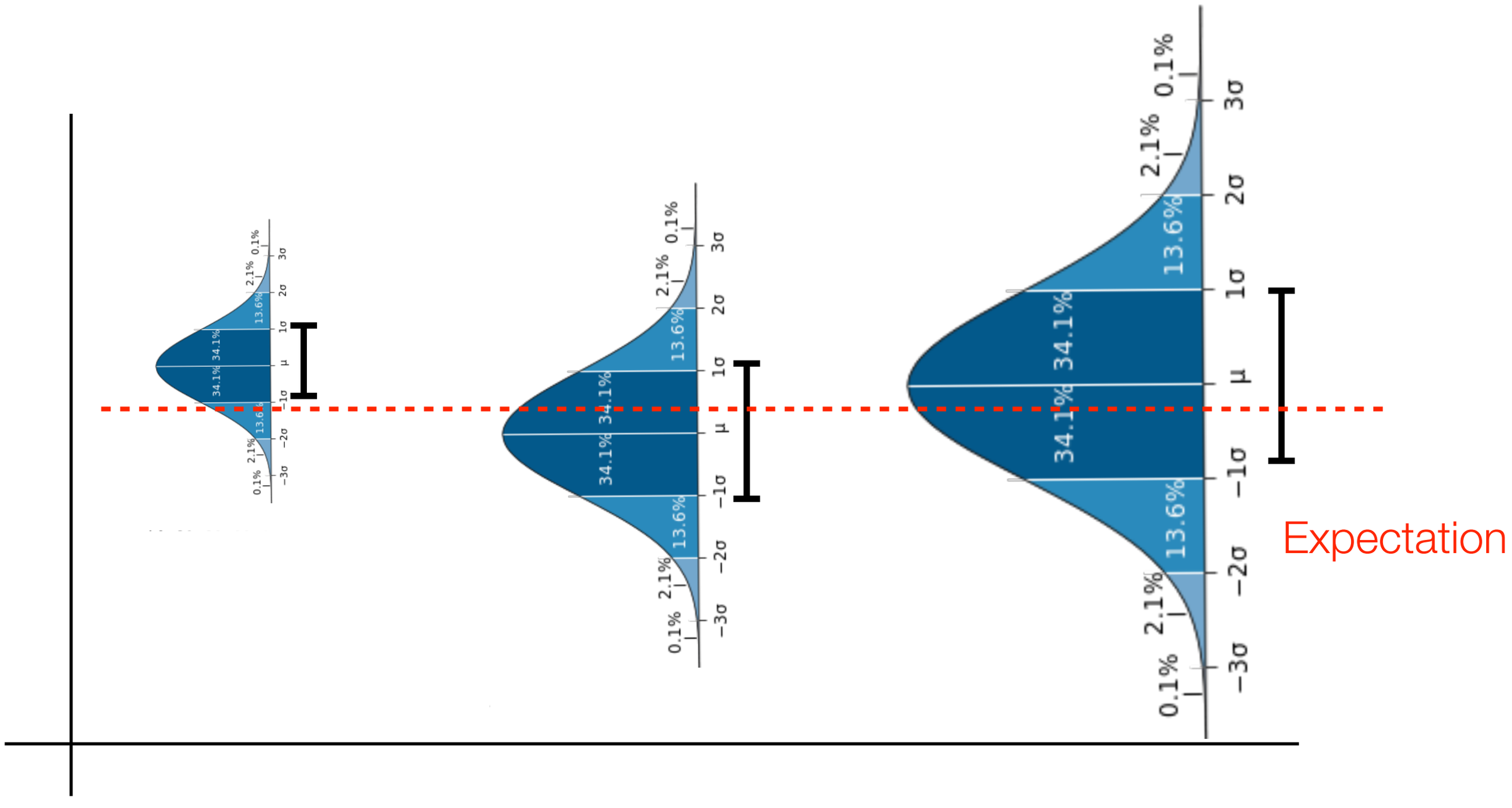


Basic Reduced Chi-Square



Basic Reduced Chi-Square

$$\chi^2_{reduced} = \chi^2 / D.O.F.$$



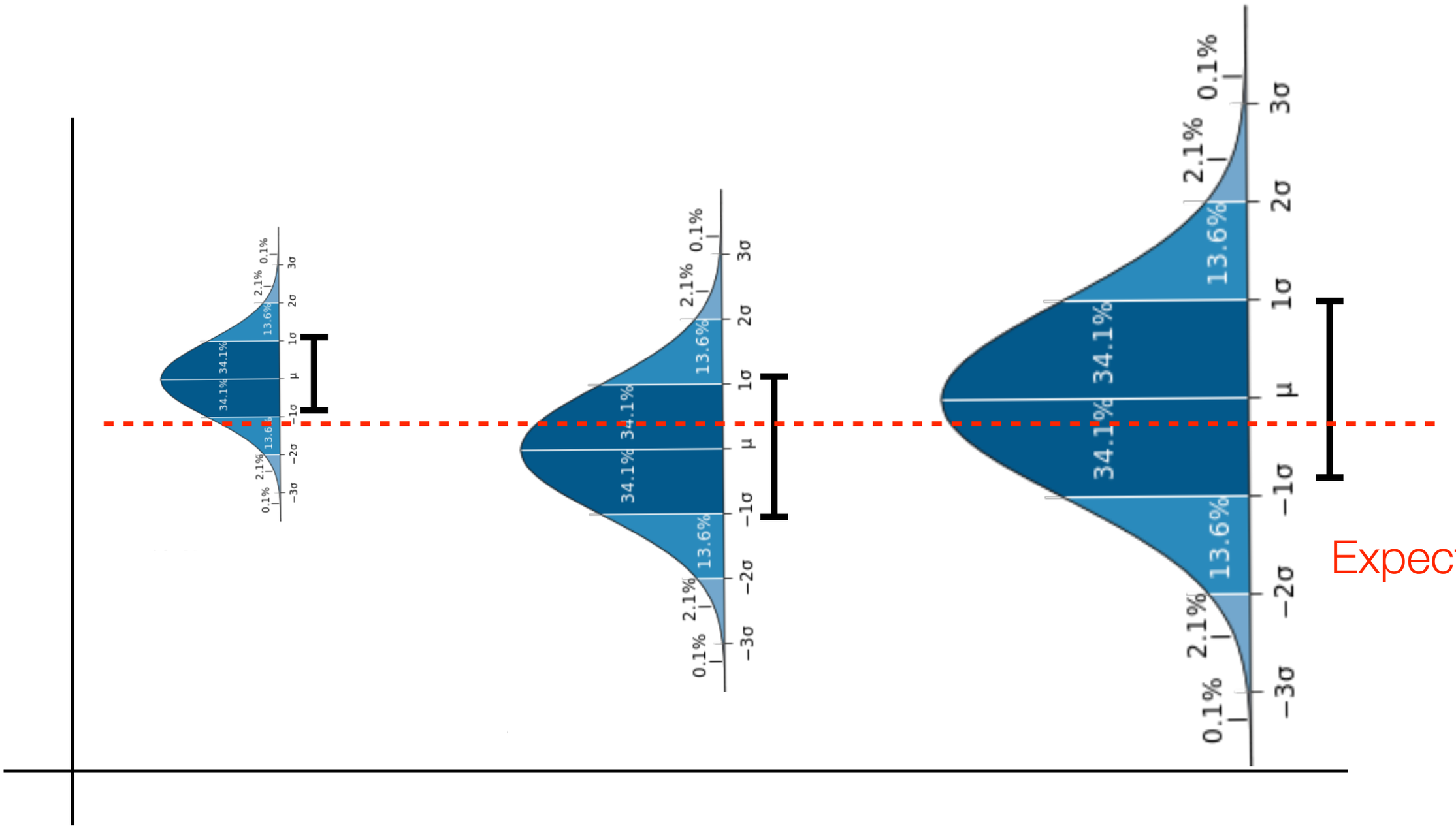
Basic Reduced Chi-Square

$$\chi^2_{reduced} = \chi^2 / D.O.F.$$

$$\chi^2_{reduced} \ll 1$$

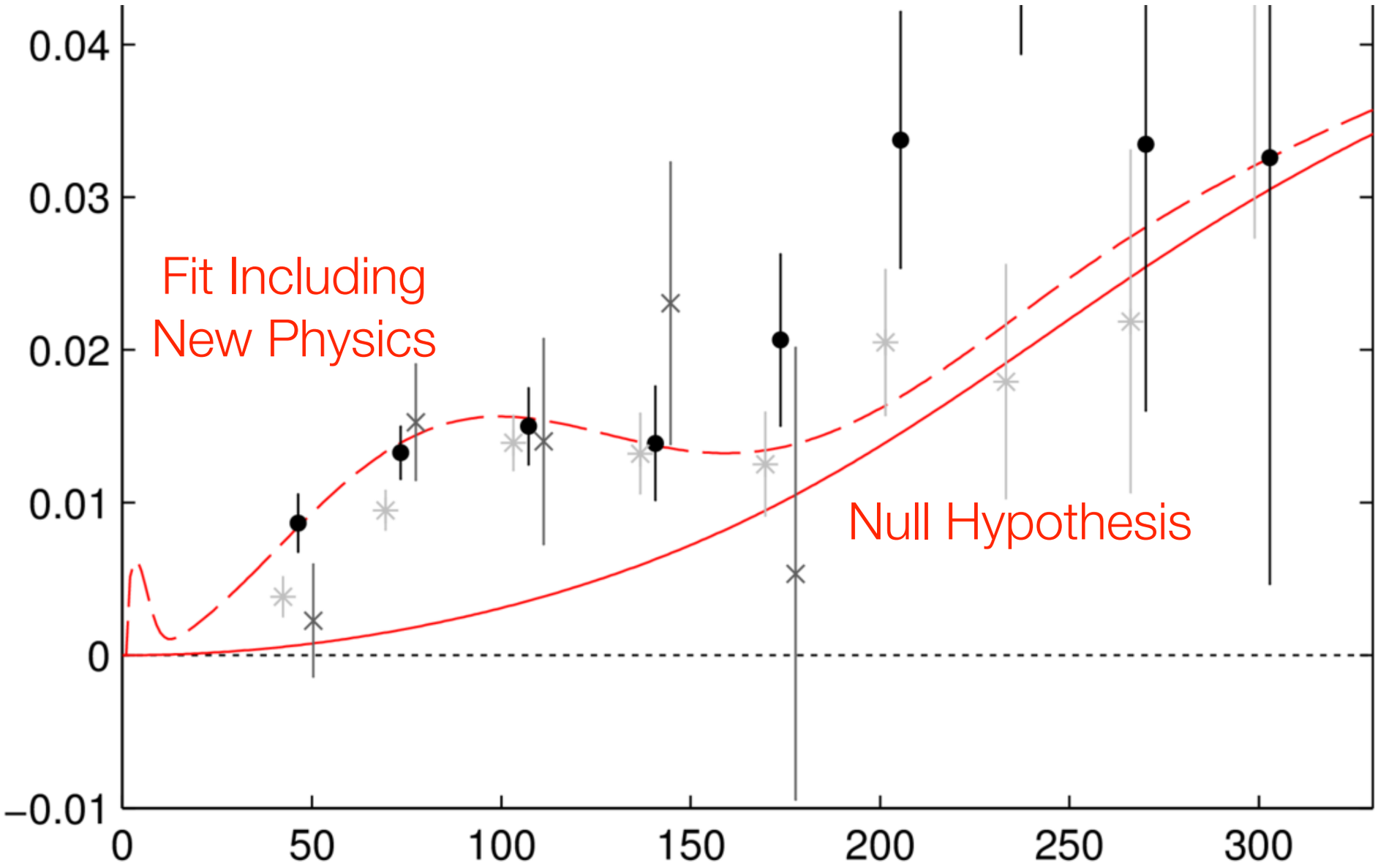
$$\chi^2_{reduced} \approx 1$$

$$\chi^2_{reduced} \gg 1$$

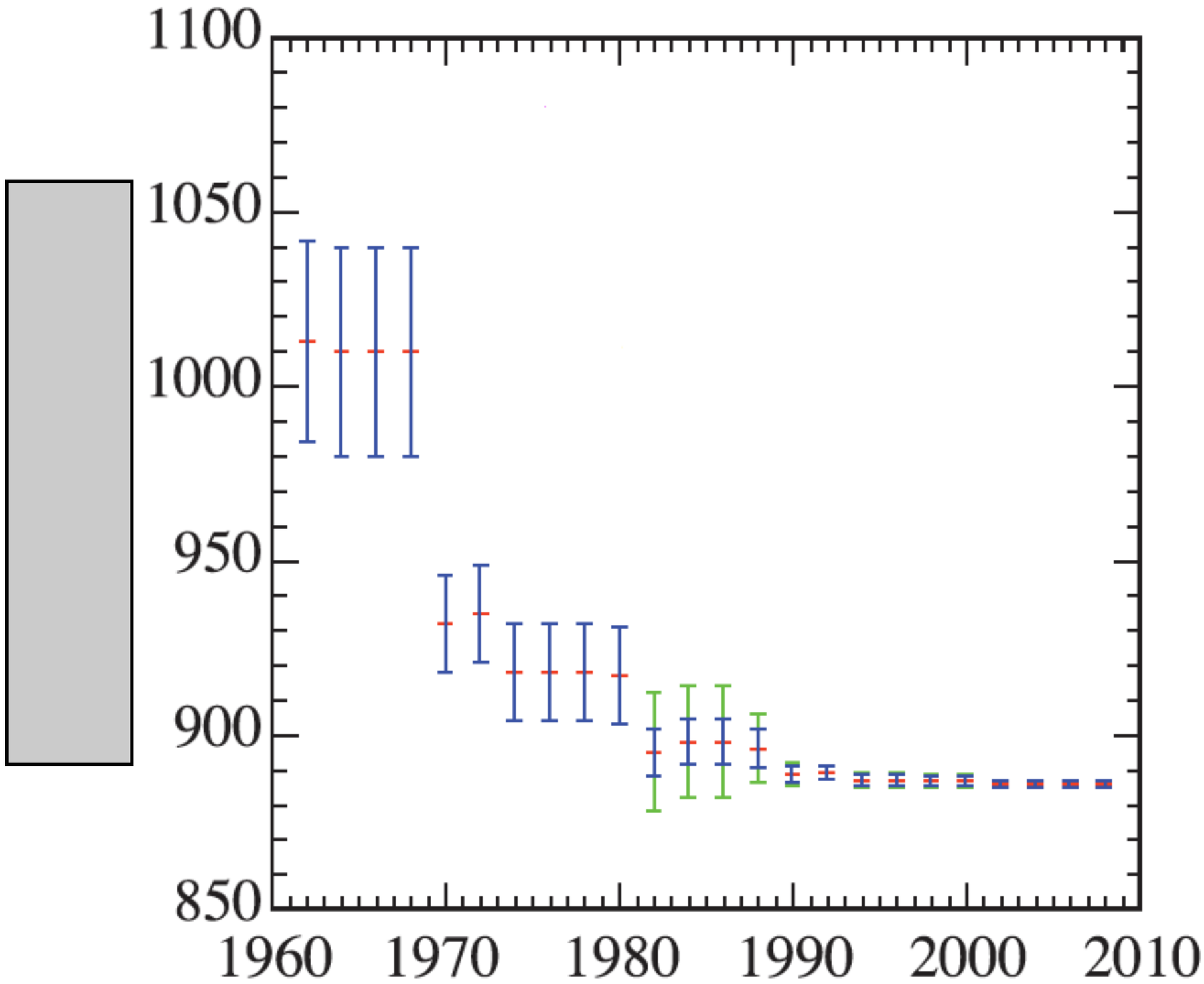


Chi-By-Eye

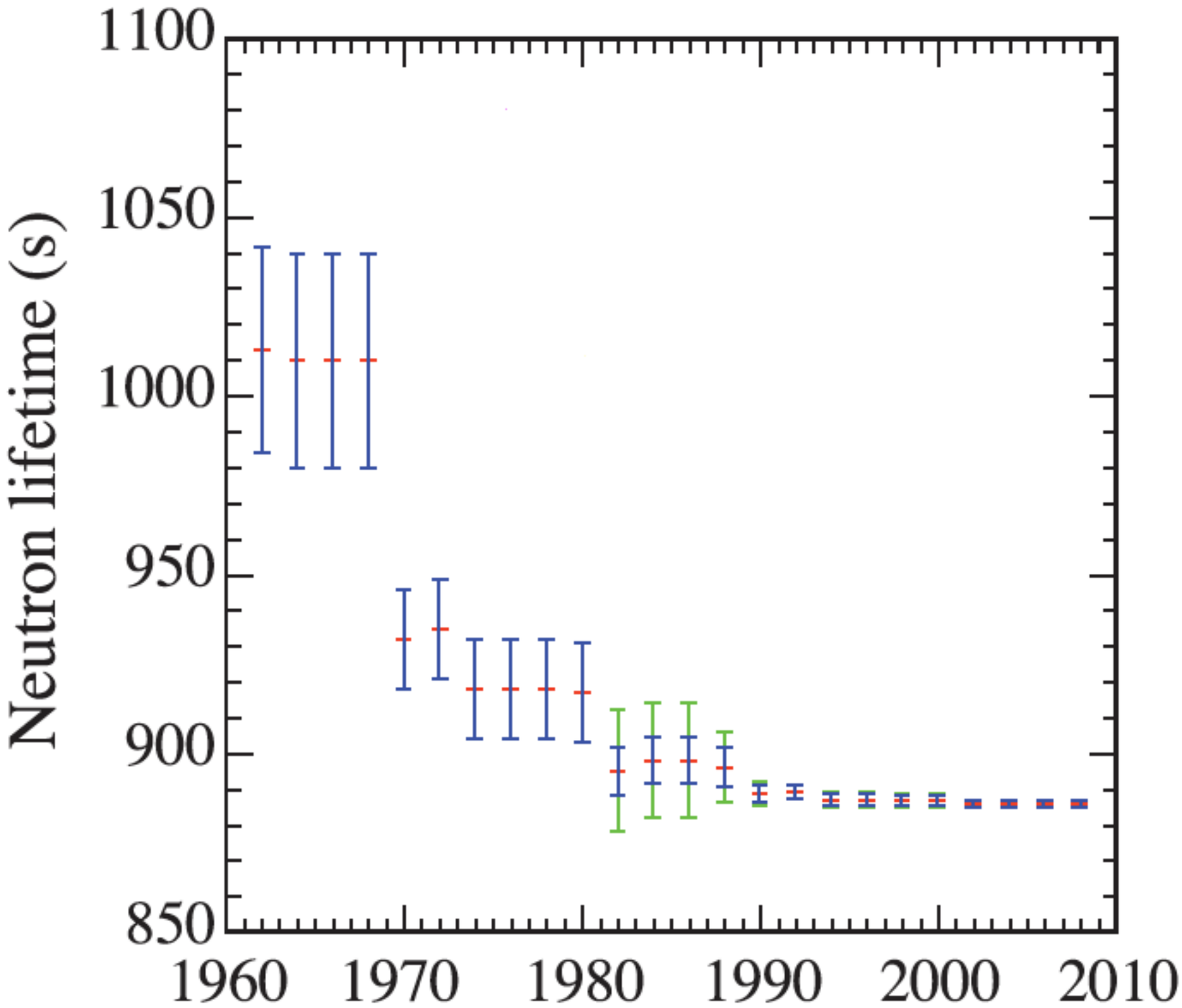
$$\chi^2 = \sum \frac{(Observed - Expected)^2}{Expected + \sigma^2_{systematics}}$$



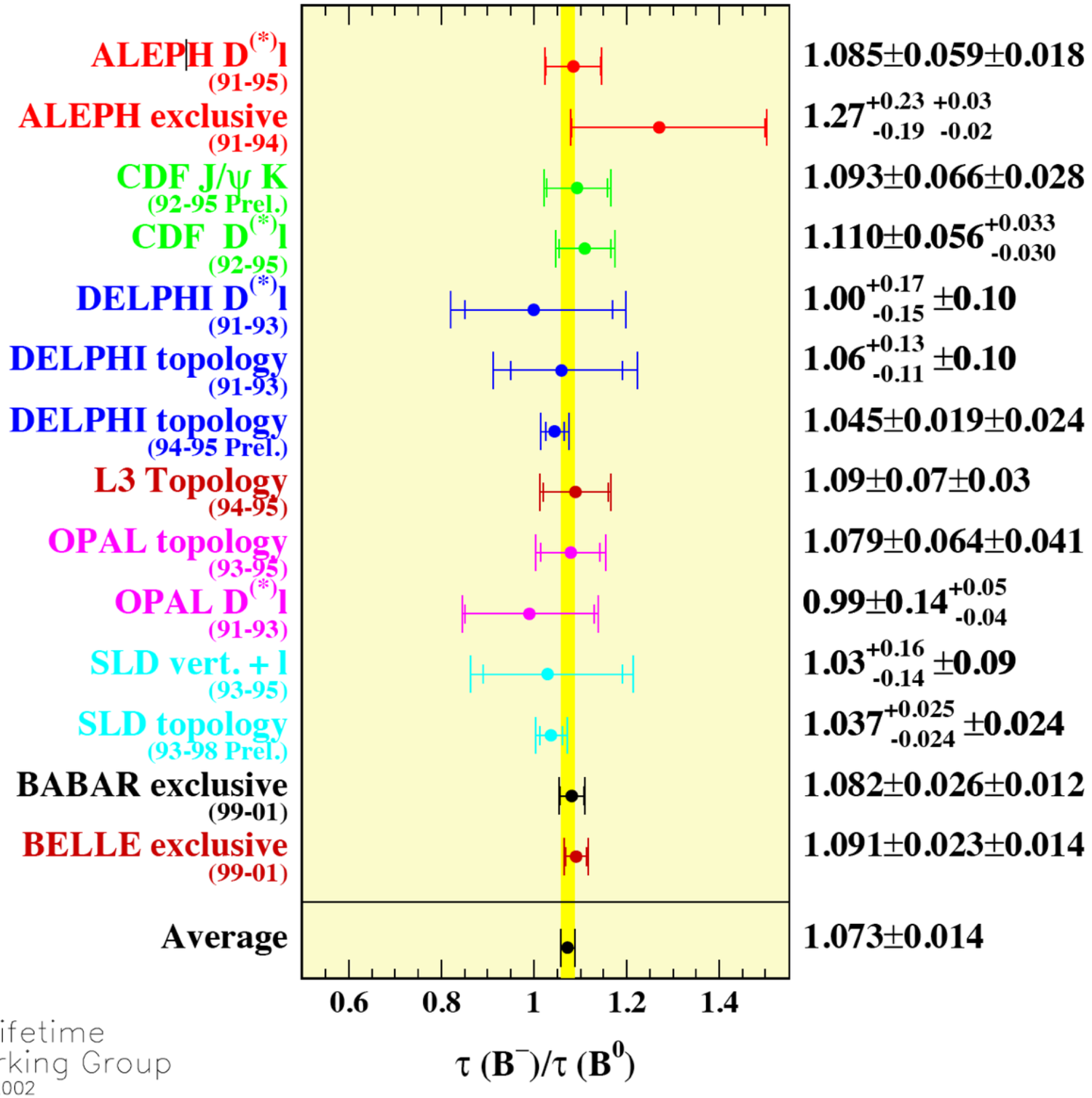
Bias



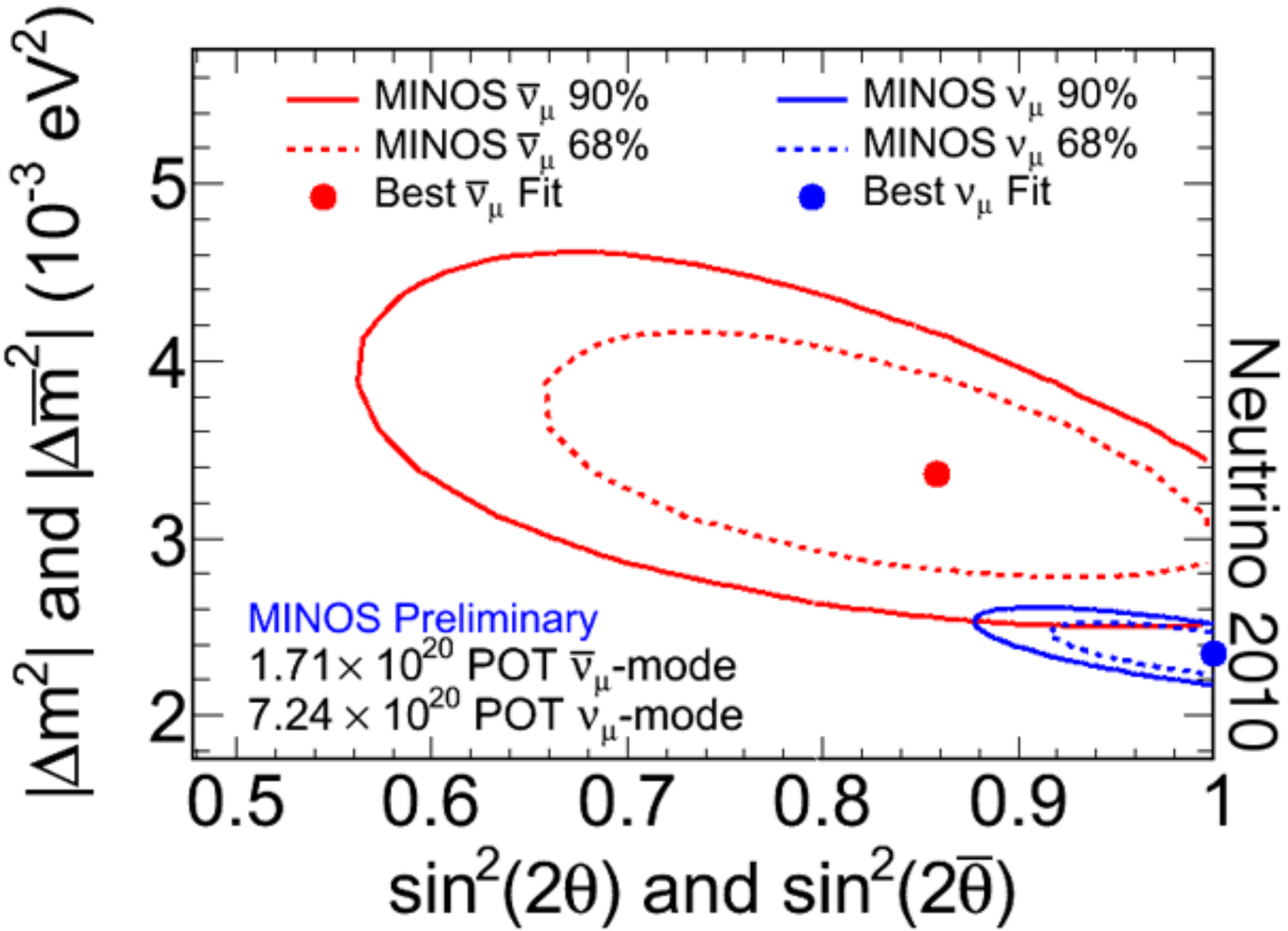
Bias


Neutron lifetime (s)
1100
1050
1000
950
900
850
1960
1970
1980
1990
2000
2010

B-Mesons

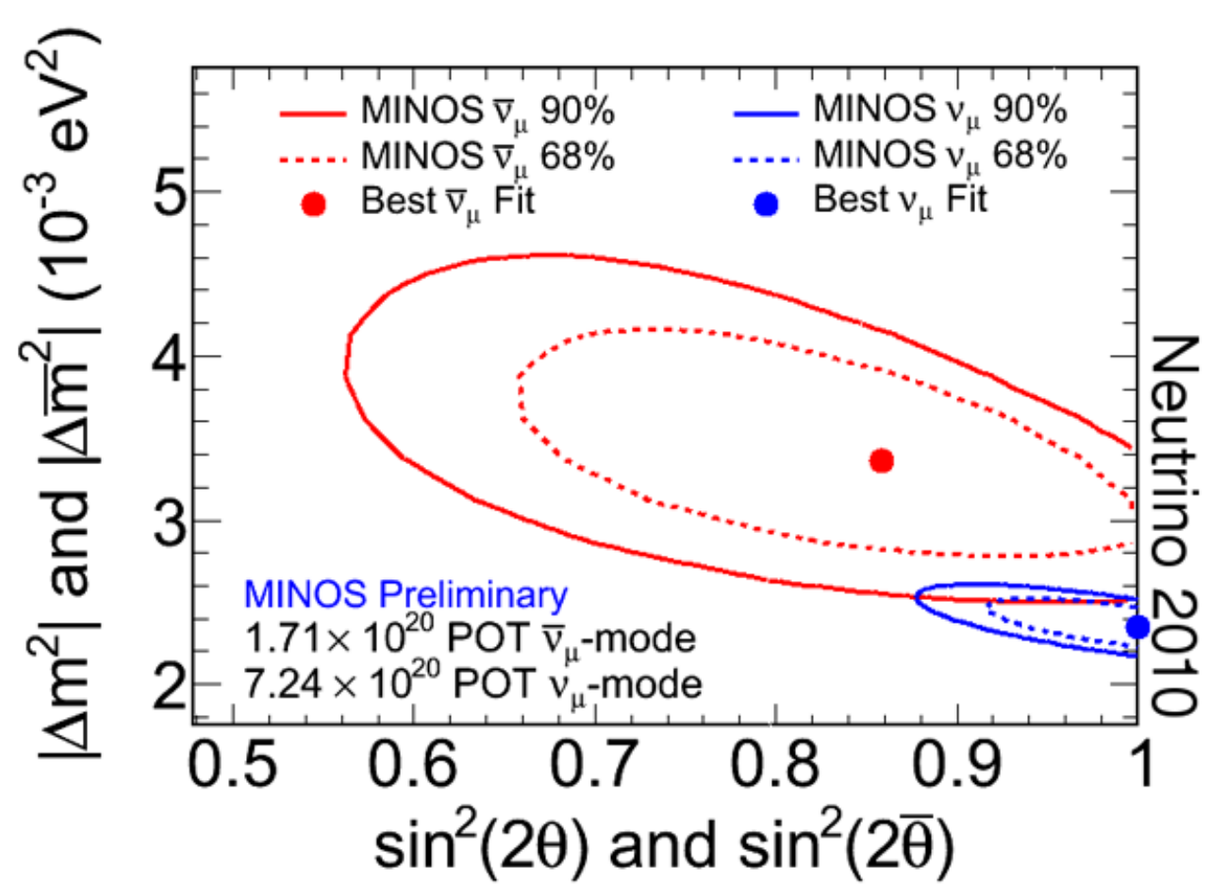

ALEPH D(*)l (91-95)
1.085±0.059±0.018
ALEPH exclusive (91-94)
1.27 +0.23 -0.19 +0.03 -0.02
CDF J/ψ K (92-95 Prel.)
1.093±0.066±0.028
CDF D(*)l (92-95)
1.110±0.056 +0.033 -0.030
DELPHI D(*)l (91-93)
1.00 +0.17 -0.15 ±0.10
DELPHI topology (91-93)
1.06 +0.13 -0.11 ±0.10
DELPHI topology (94-95 Prel.)
1.045±0.019±0.024
L3 Topology (94-95)
1.09±0.07±0.03
OPAL topology (93-95)
1.079±0.064±0.041
OPAL D(*)l (91-93)
0.99±0.14 +0.05 -0.04
SLD vert. + l (93-95)
1.03 +0.16 -0.14 ±0.09
SLD topology (93-98 Prel.)
1.037 +0.025 -0.024 ±0.024
BABAR exclusive (99-01)
1.082±0.026±0.012
BELLE exclusive (99-01)
1.091±0.023±0.014
Average
1.073±0.014
0.6
0.8
1
1.2
1.4
τ (B⁻)/τ (B⁰)
B Lifetime Working Group
July 2002

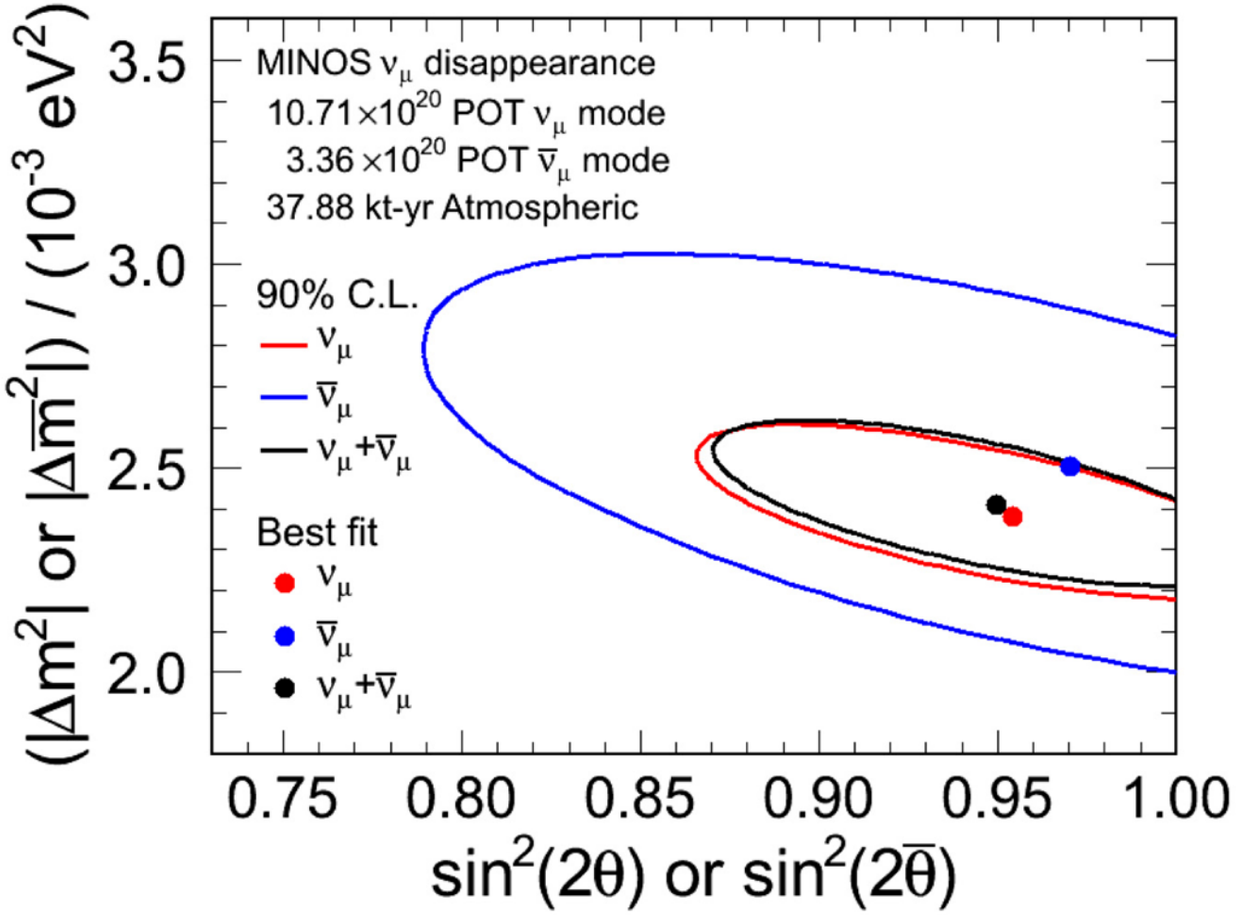
Too Different?



Now in Sync

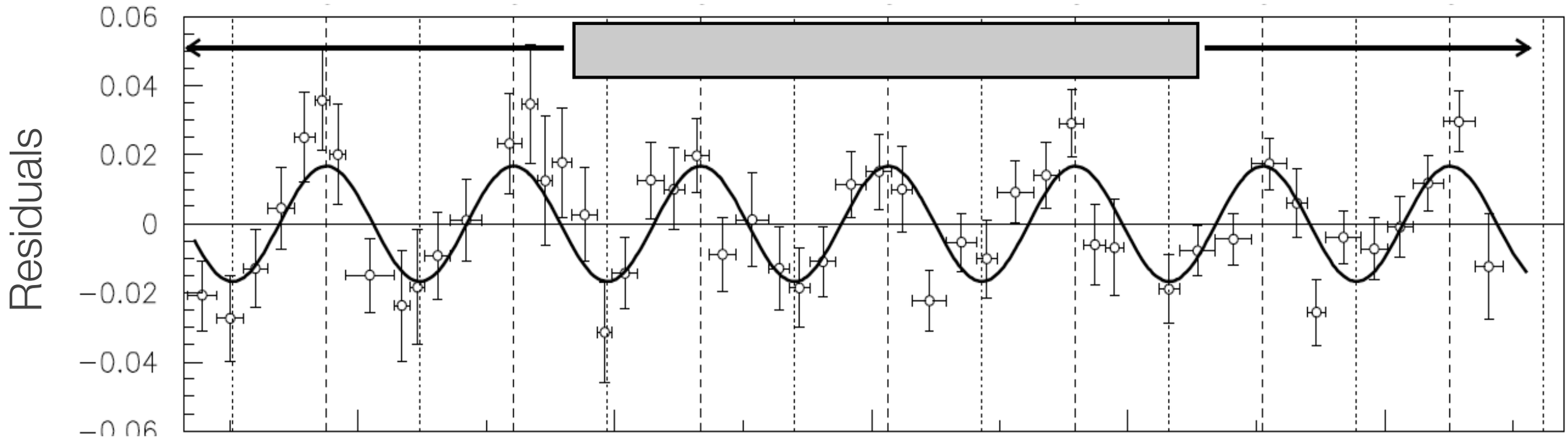
- At what point should we start thinking that a 2-3σ will persist or grow?





Enough Information?

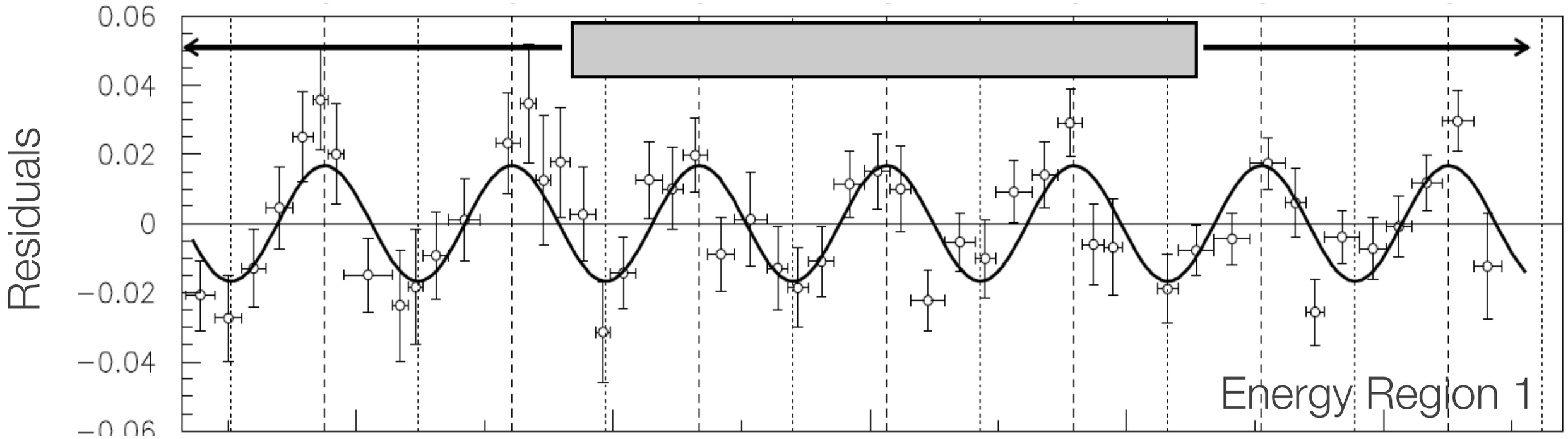
- Null hypothesis in following experiment is no modulation of signal with time
- Error bars are experimental (systematic)



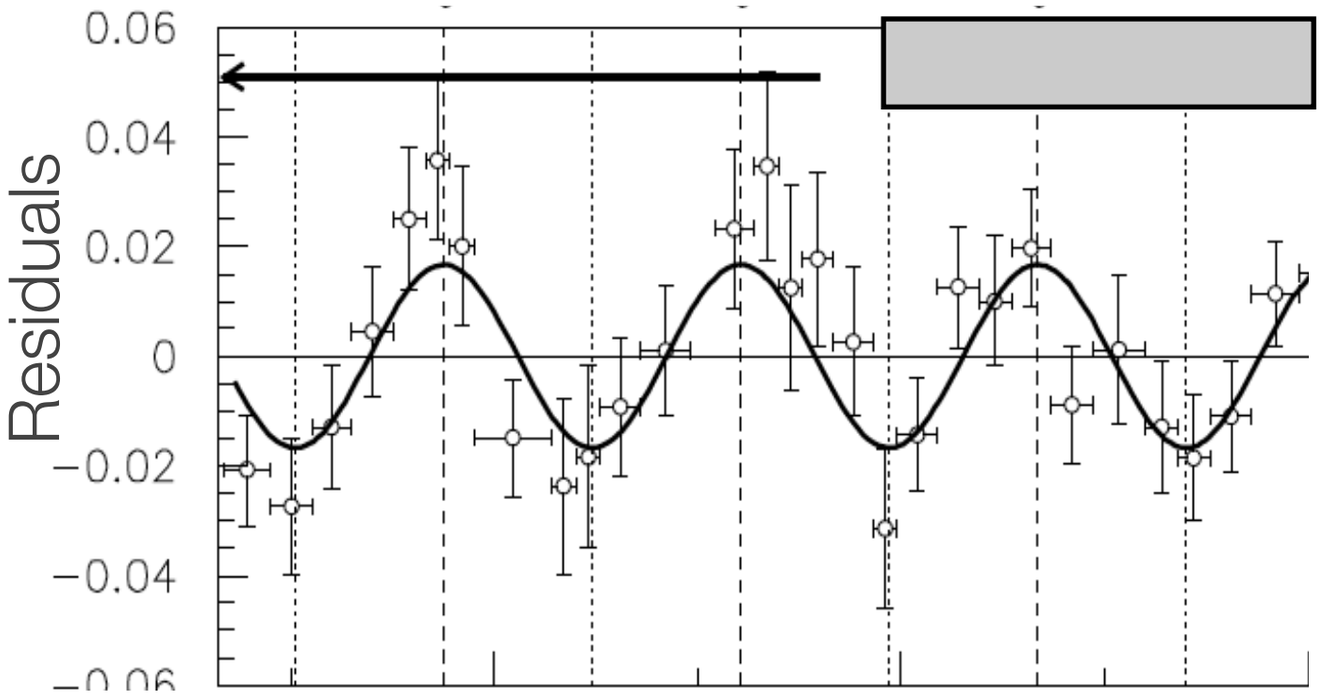
arXiv:1308.5109

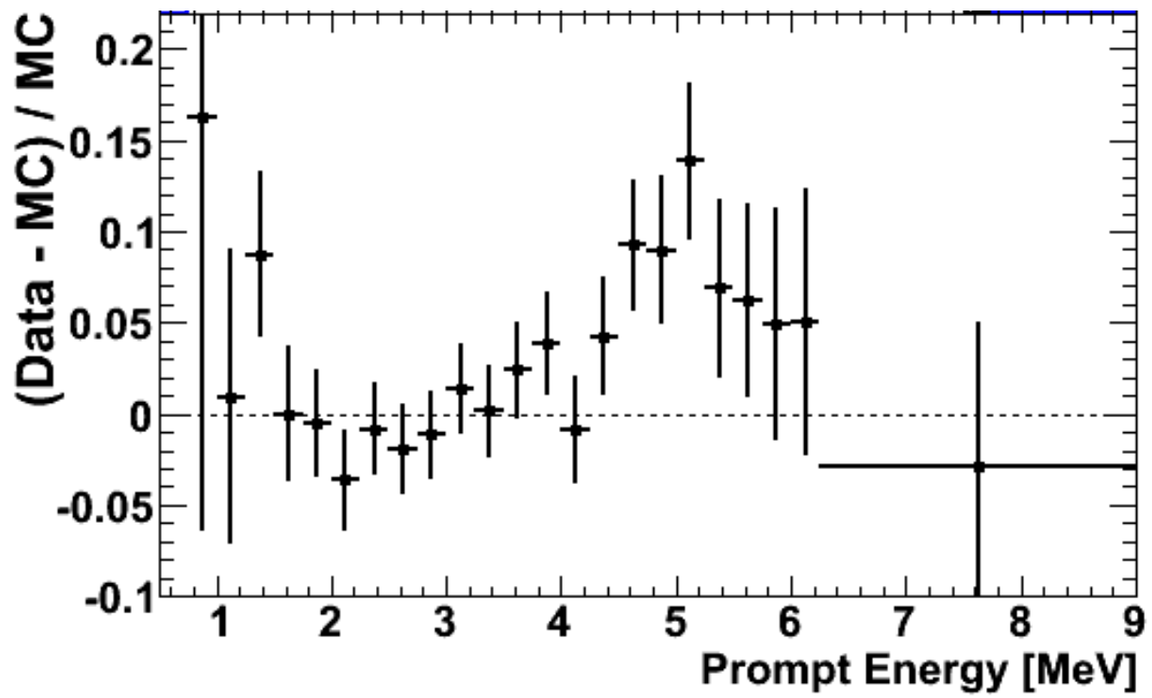
Enough Information?

	Experimental Setup A	Experimental Setup A+B
Energy Region 1	$A=(0.0167\pm0.0022) \rightarrow 7.6\ \sigma$ C.L. $\chi^2/\text{d.o.f.} = 52.3/49$	$A=(0.0179\pm0.0020) \rightarrow 9.0\ \sigma$ C.L. $\chi^2/\text{d.o.f.} = 87.1/86$
Energy Region 2	$A=(0.0122\pm0.0016) \rightarrow 7.6\ \sigma$ C.L. $\chi^2/\text{d.o.f.} = 41.4/49$	$A=(0.0135\pm0.0015) \rightarrow 9.0\ \sigma$ C.L. $\chi^2/\text{d.o.f.} = 68.2/86$
Energy Region 3	$A=(0.0096\pm0.0013) \rightarrow 7.4\ \sigma$ C.L. $\chi^2/\text{d.o.f.} = 29.3/49$	$A=(0.0110\pm0.0012) \rightarrow 9.2\ \sigma$ C.L. $\chi^2/\text{d.o.f.} = 70.4/86$

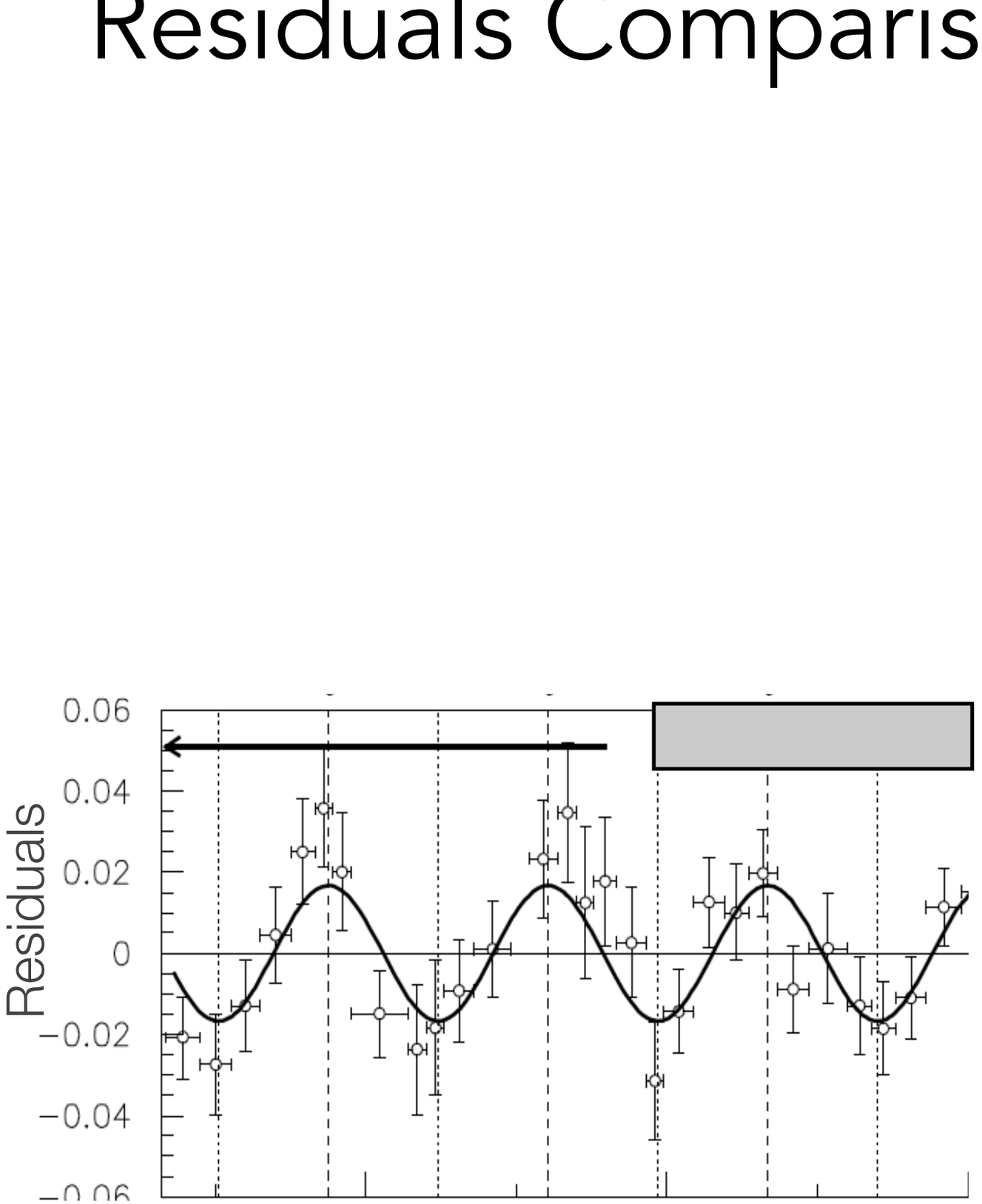


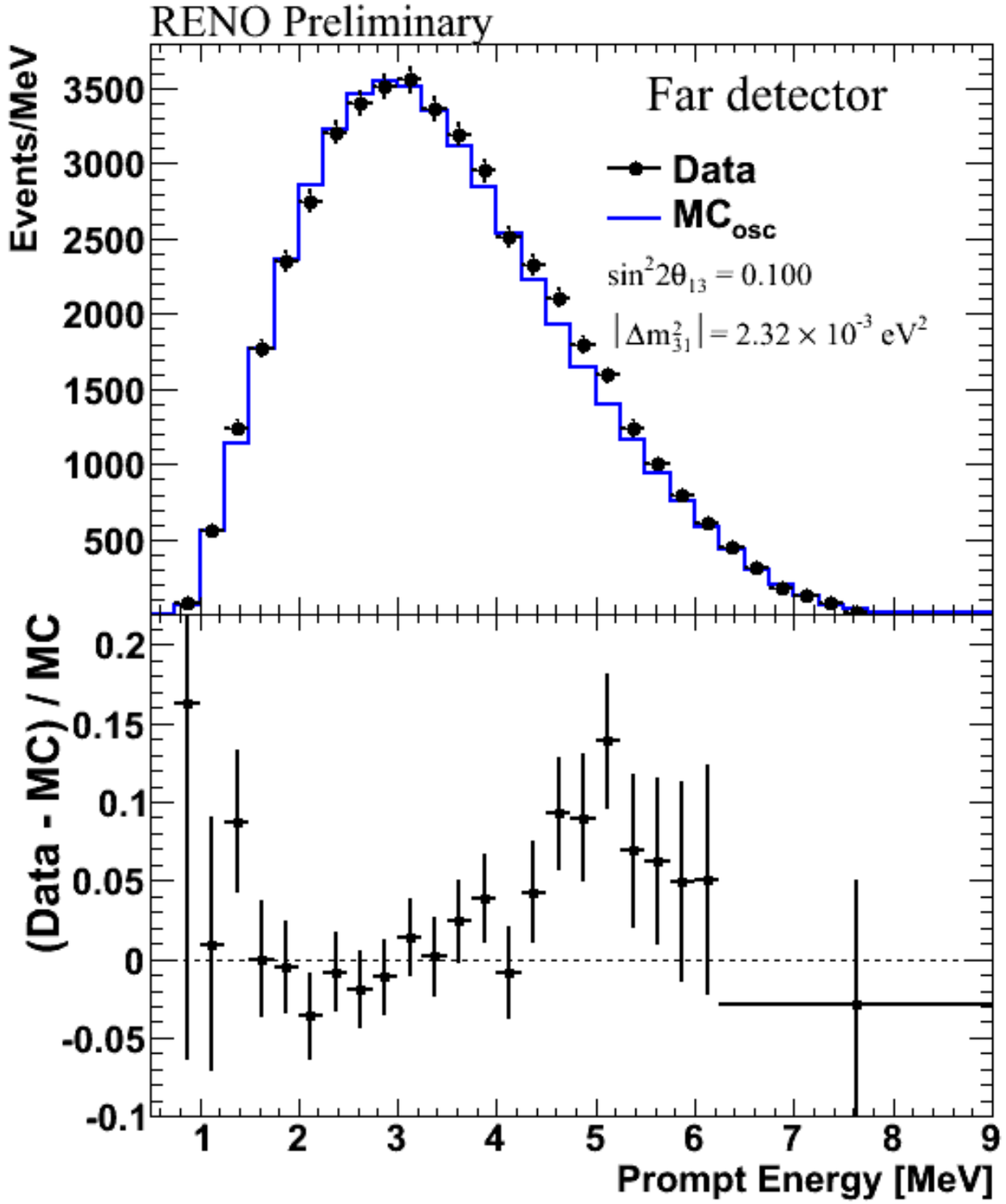
Residuals Comparison





Residuals Comparison





P-Value

P-Value

- Many different ways to think about statistics… and even more ways to argue

P-Value

- Many different ways to think about statistics… and even more ways to argue

- Given the observed data, for a given expectation (generally the null hypothesis) how often should I expect to observe data that is at least as aberrant/extreme as the actual observed data
  - P-value is affected by the amount of data examined
  - Be mindful of pre-trial p-value versus post-trial p-value

P-Value

- Many different ways to think about statistics… and even more ways to argue

- Given the observed data, for a given expectation (generally the null hypothesis) how often should I expect to observe data that is at least as aberrant/extreme as the actual observed data
  - P-value is affected by the amount of data examined
  - Be mindful of pre-trial p-value versus post-trial p-value

This one time I flipped a coin and it came up ‘heads’ 500 consecutive times

P-Value

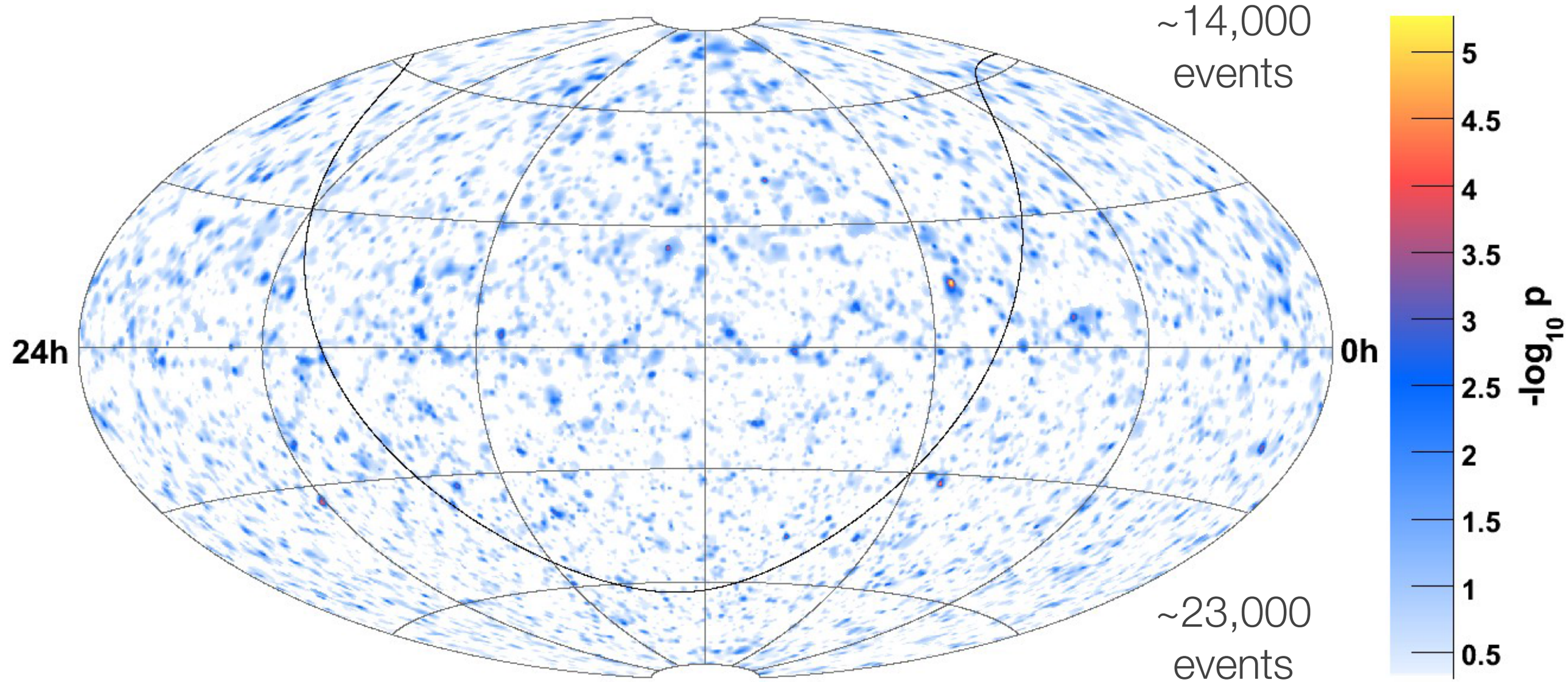
- Many different ways to think about statistics… and even more ways to argue

- Given the observed data, for a given expectation (generally the null hypothesis) how often should I expect to observe data that is at least as aberrant/extreme as the actual observed data
  - P-value is affected by the amount of data examined
  - Be mindful of pre-trial p-value versus post-trial p-value

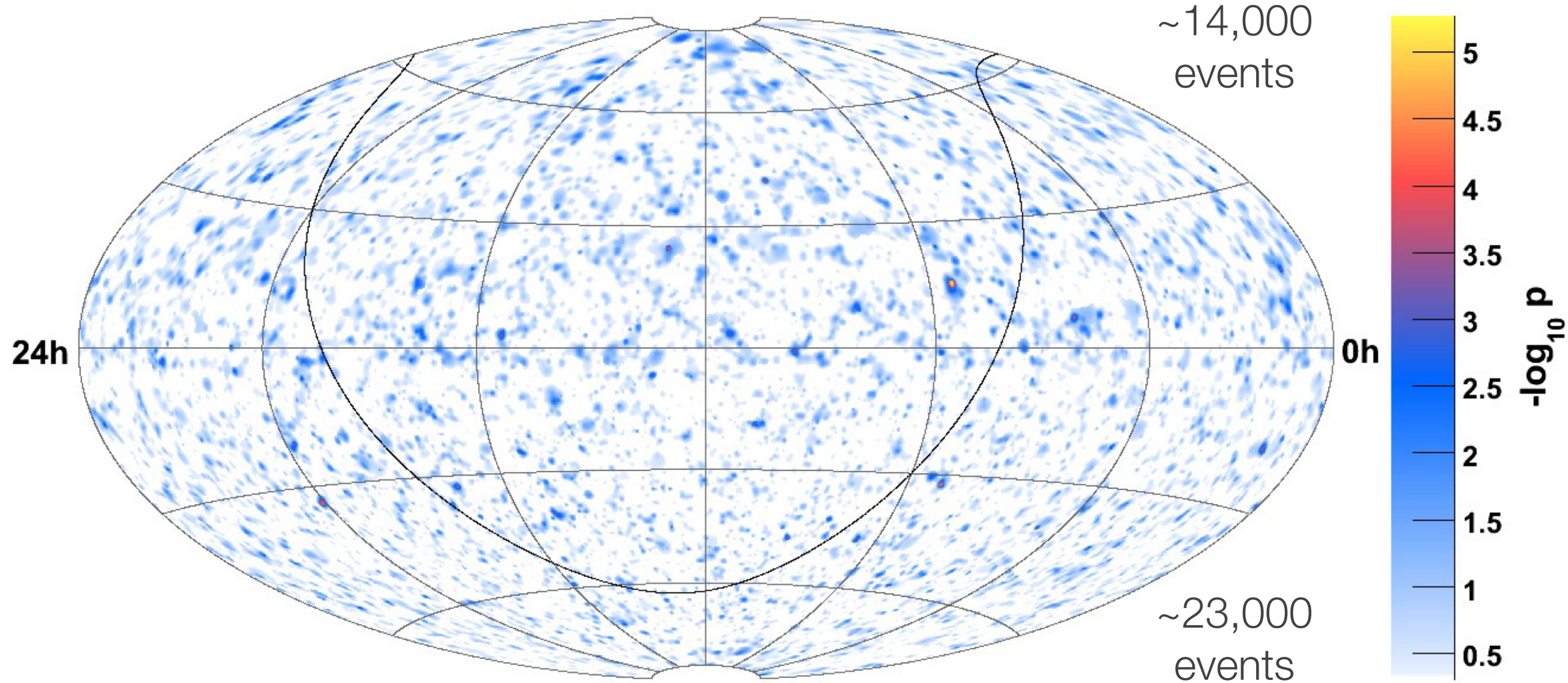
This one time I flipped a coin and it came up ‘heads’ 500 consecutive times

This one time I flipped a coin 7 billion times and at one point it came up ‘heads’ 500 consecutive times

Lots of data

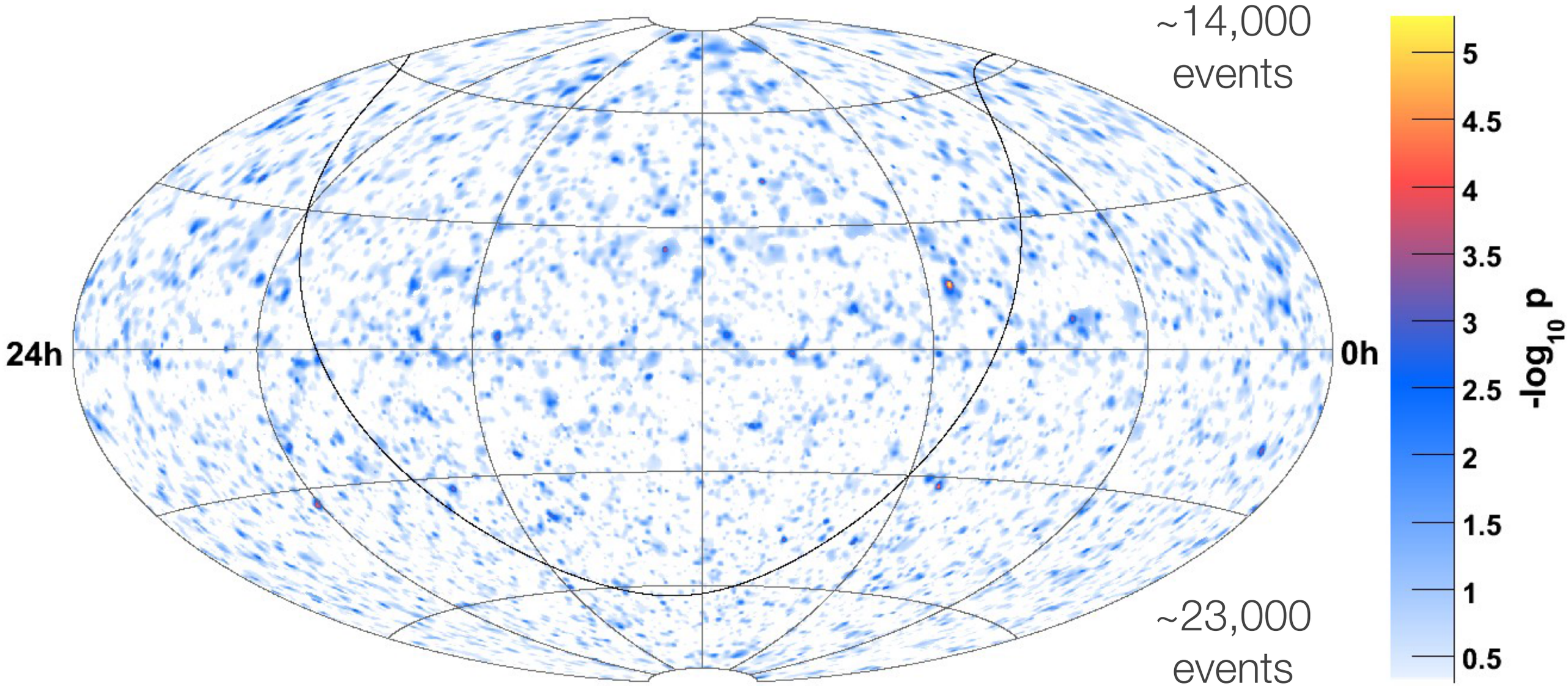


Lots of data



pre-trial $-\log_{10}$(p-value) = 5.28

Lots of data



pre-trial $-\log_{10}$(p-value) = 5.28

post-trial (p-value) = 18%

Statistical

Statistical

- OPERA has 3 events
  - If stat. uncertainty goes as sqrt(number of events) how is the OPERA significance > 3 sigma?

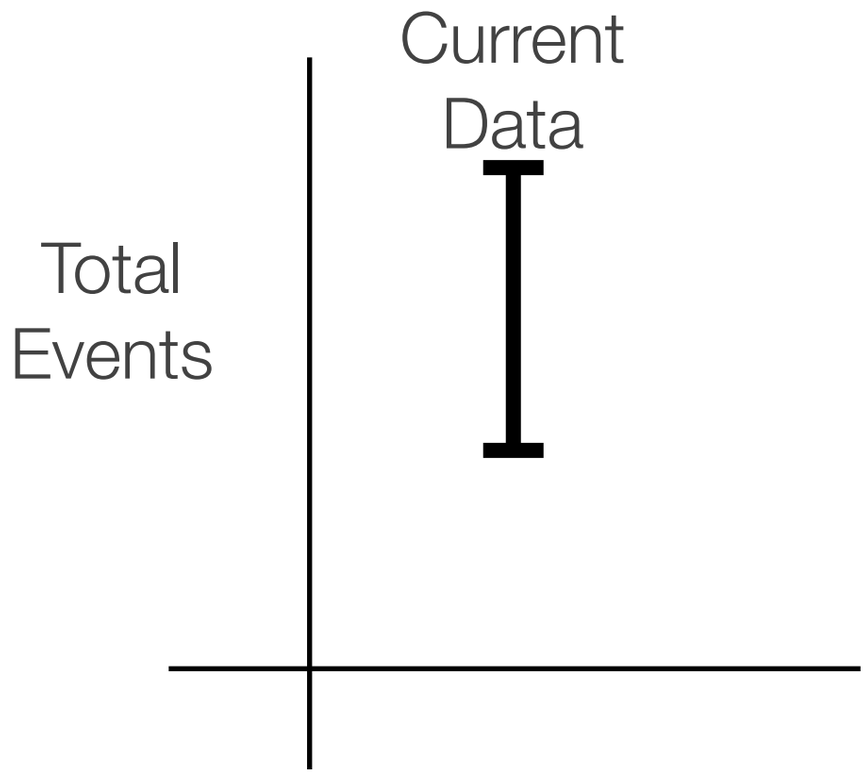
Statistical

- OPERA has 3 events
  - If stat. uncertainty goes as sqrt(number of events) how is the OPERA significance > 3 sigma?
- OPERA has 1.78 expected events and 0.184 background
  - But we can only see ‘whole’ events
  - What does 1.78 expected mean in the context of observing 3?
  - What if only 1 event had been observed?

Statistical

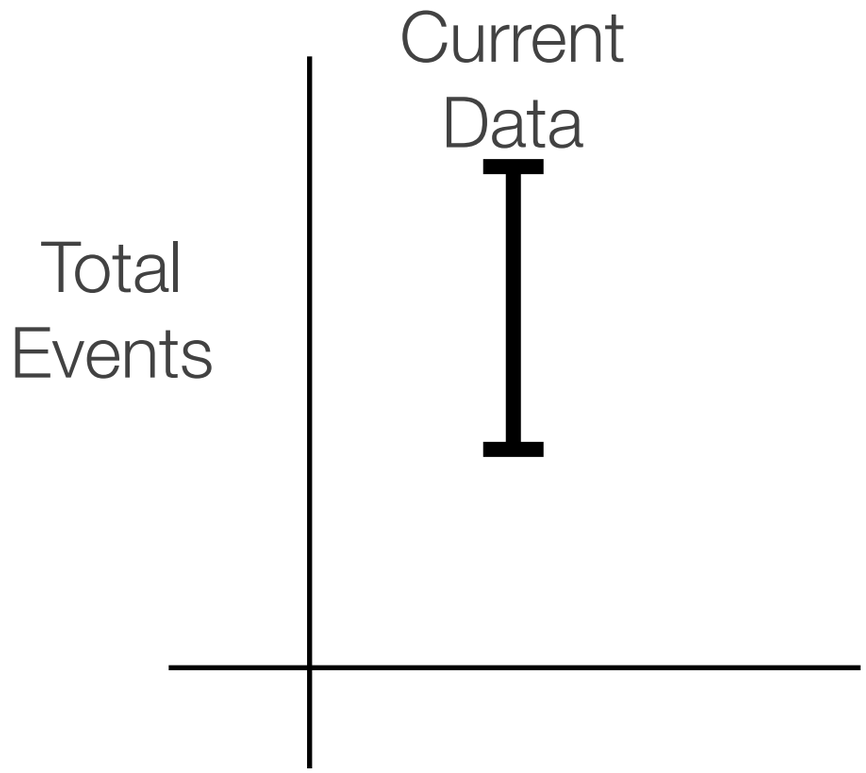
- OPERA has 3 events
  - If stat. uncertainty goes as sqrt(number of events) how is the OPERA significance > 3 sigma?
- OPERA has 1.78 expected events and 0.184 background
  - But we can only see ‘whole’ events
  - What does 1.78 expected mean in the context of observing 3?
  - What if only 1 event had been observed?
- Some physics is binary
  - Neutrino mass hierarchy is either A or B and not B ± 10%
  - Many tests (chi-squared) have implicit assumptions of gaussian statistics

Biased Chi-squared scenario



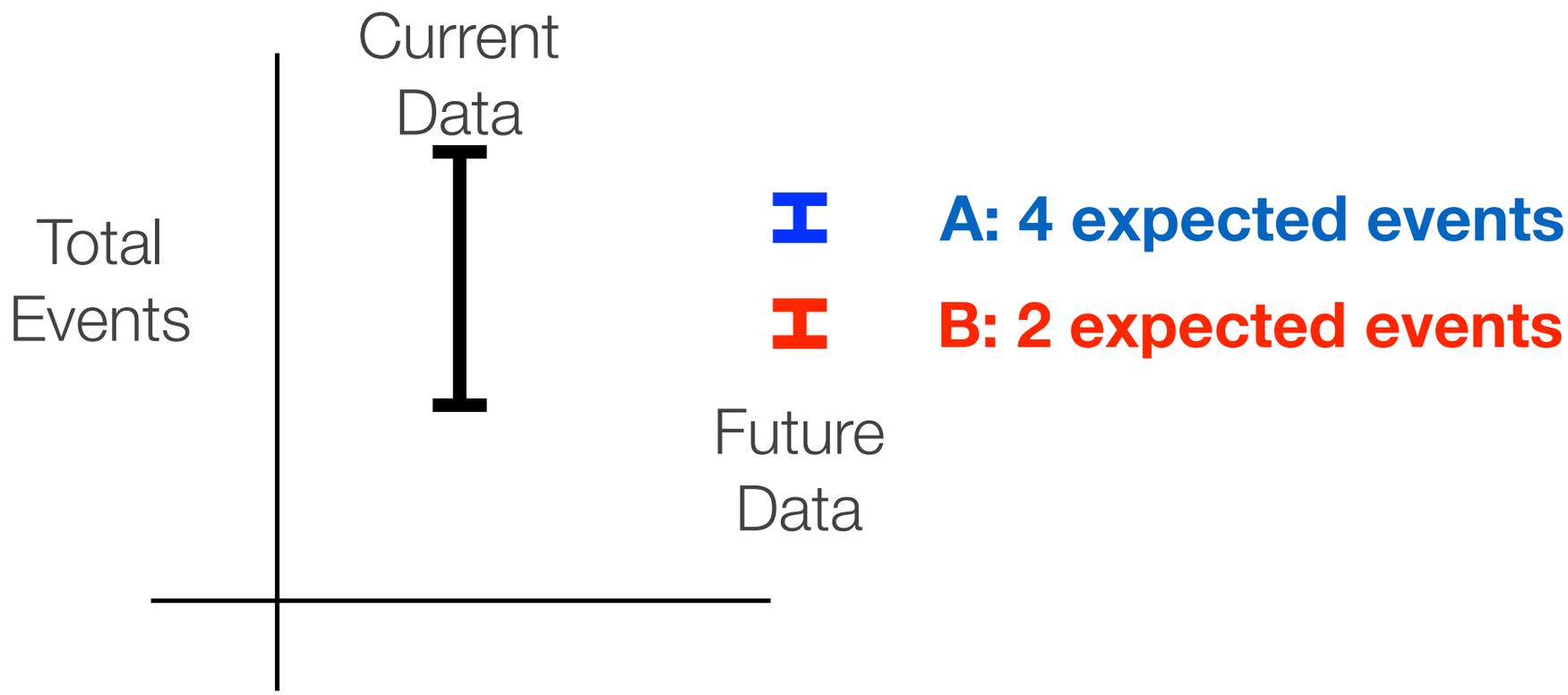
Biased Chi-squared scenario

$$\chi^2 = \sum \frac{(Observed - Expected)^2}{Expected}$$



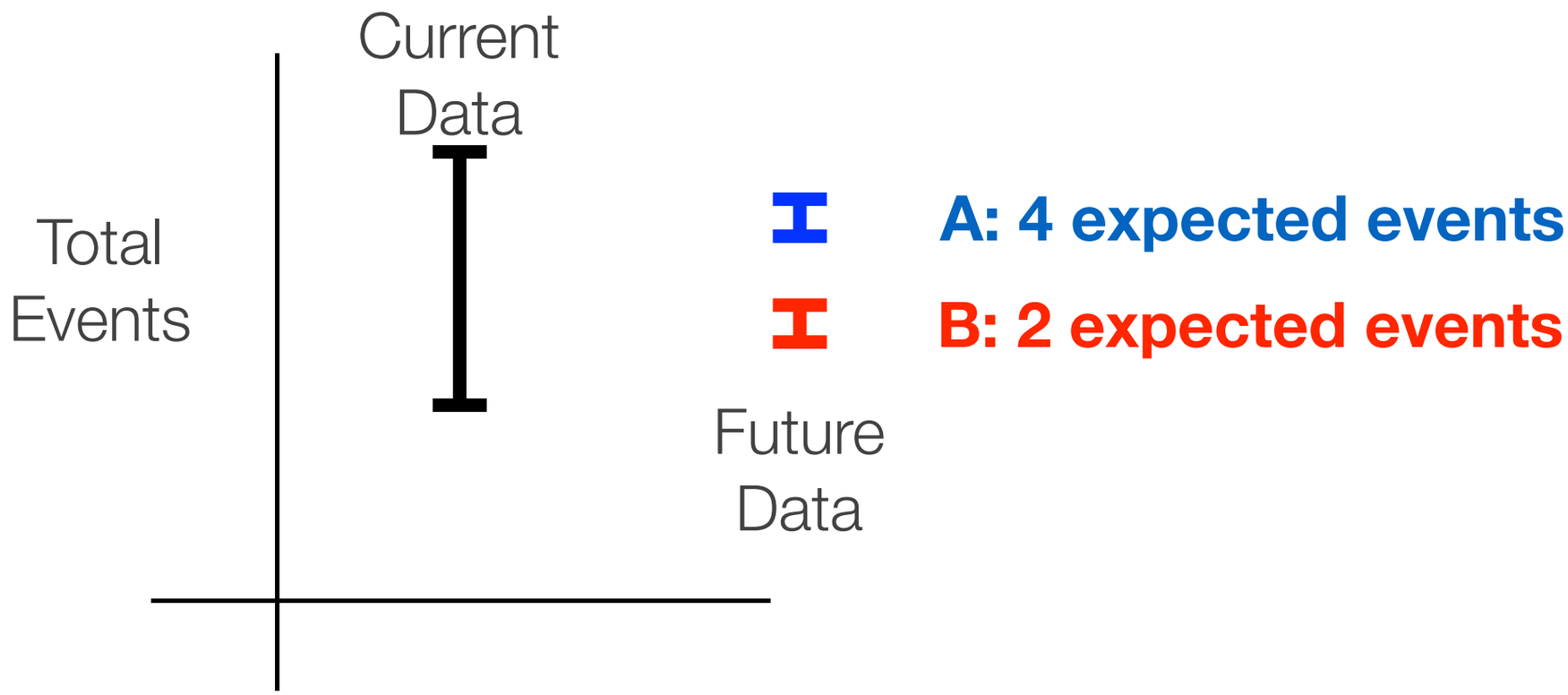
Biased Chi-squared scenario

$$\chi^2 = \sum \frac{(Observed - Expected)^2}{Expected}$$



Biased Chi-squared scenario

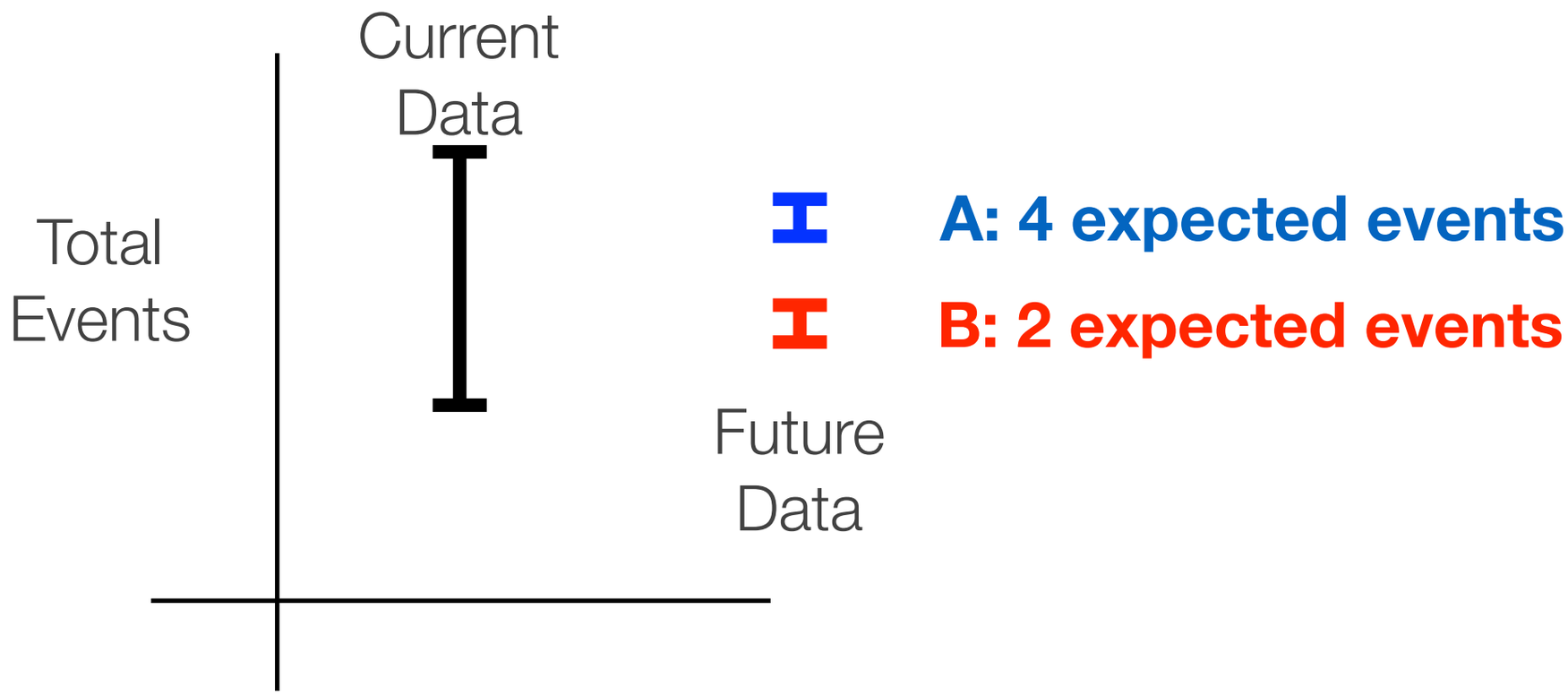
$$\chi^2 = \sum \frac{(Observed - Expected)^2}{Expected}$$



- In many experiments, *A priori*, the significance of A being physical reality and excluding B should be the same as B being physical reality and excluding A

Biased Chi-squared scenario

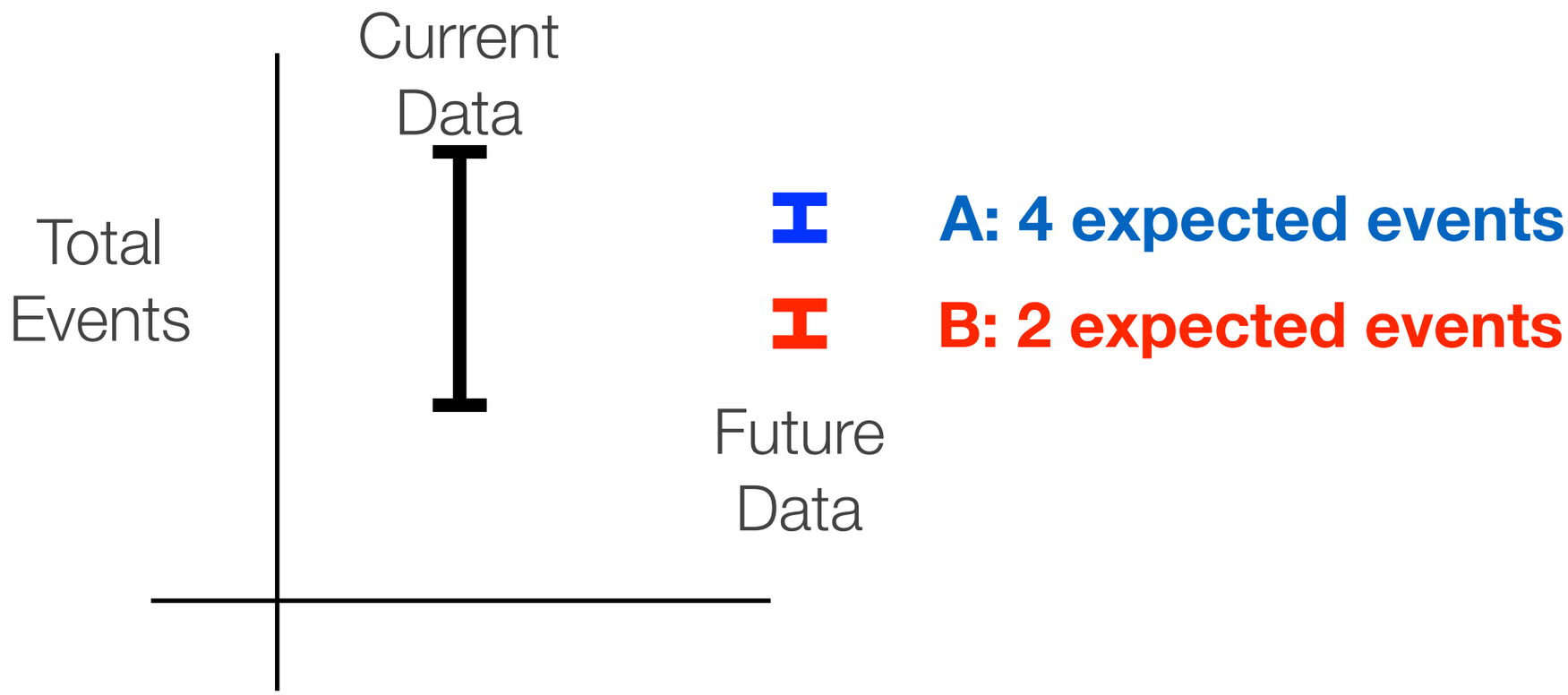
$$\chi^2 = \sum \frac{(Observed - Expected)^2}{Expected}$$



- In many experiments, *A priori*, the significance of A being physical reality and excluding B should be the same as B being physical reality and excluding A
  - A $\chi^2 = (4\text{-}2)^2/4 = 1$

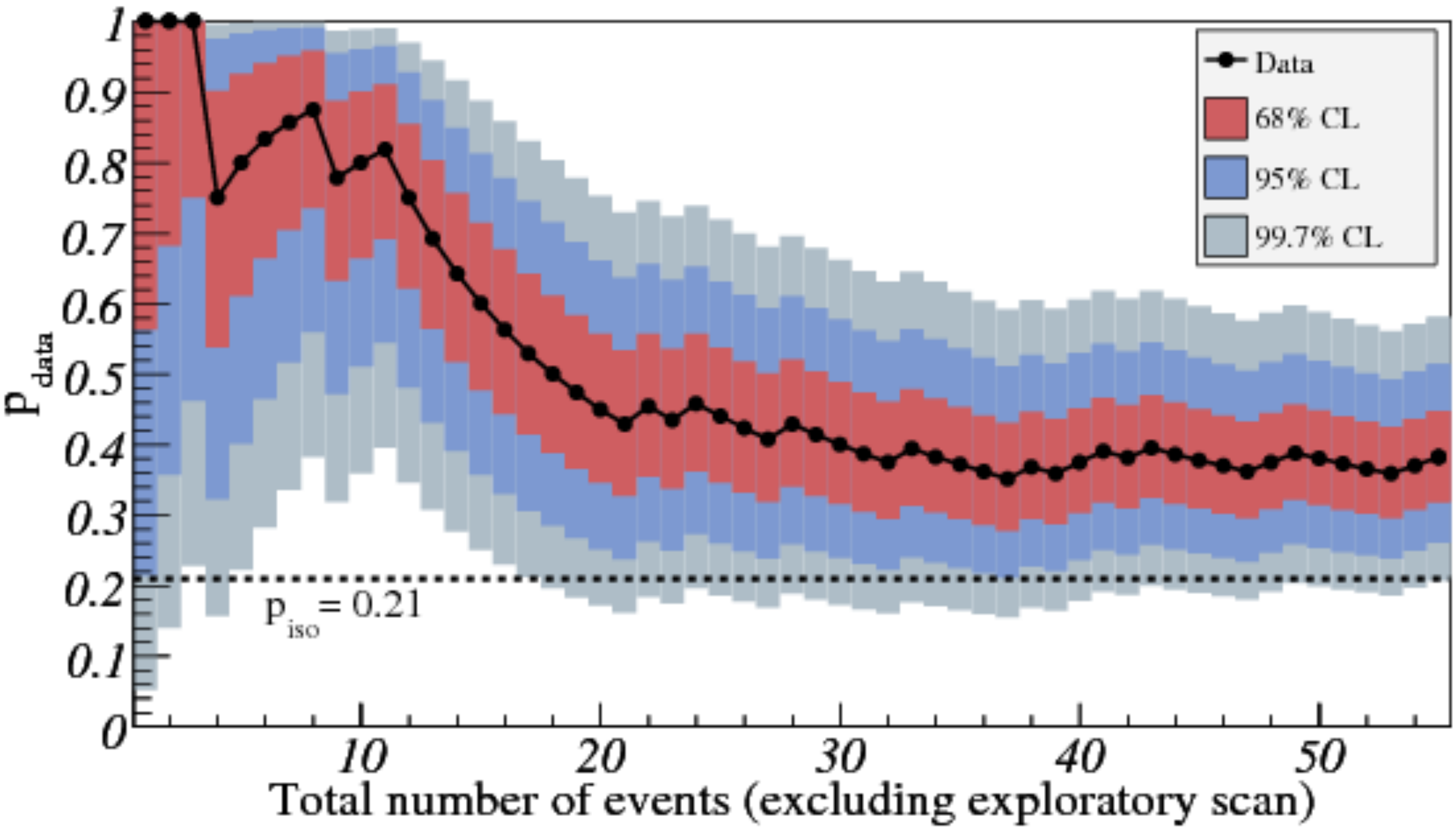
Biased Chi-squared scenario

$$\chi^2 = \sum \frac{(Observed - Expected)^2}{Expected}$$



- In many experiments, *A priori*, the significance of A being physical reality and excluding B should be the same as B being physical reality and excluding A
  - A $\chi^2 = (4\text{-}2)^2/4 = 1$
  - B $\chi^2 = (2\text{-}4)^2/2 = 2$

What?


Data
68% CL
95% CL
99.7% CL
p_{data}
$p_{iso} = 0.21$
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
10
20
30
40
50
Total number of events (excluding exploratory scan)

Wrap-Up

- Many, many statistical tests available
- Quick chi-squared test is a useful tool, but should NOT be used to draw definitive conclusions