8.882 LHC Physics

Experimental Methods and Measurements

Search Strategies and Observations

[Lecture 14, March 30, 2009]

Organizational Issues

Project 2

- due April 9

Project 3 is coming up quickly afterwards

- due May 2

Guest lecture planned on:

- Baysian versus frequentistic approach to statistics
- Michael Betancourt

Lecture Outline

Search Strategies and Observations

- introduction
- general methods
- analysis biases and how to avoid them
- two examples
  - resonance searches: $D^0 \rightarrow \mu^+\mu^-$
  - search for an oscillation frequency: B_s oscillations

Introduction

Different types of analysis in particle physics

- measurement
  - yields a number with an uncertainty for a given observable
  - that number should say something about the Standard Model (SM)
  - at best: could be inconsistent with the SM: something is wrong
  - at least: knowledge of SM improved, better test: *bread and butter*
- search
  - assume some physics model (usually new physics or even SM)
  - determine some characteristic observable to verify the model
  - analyze data concerning observable: is model supported by data?
    - yes: a signal was found, turn the search into a measurement
    - no: search yields a limit with a confidence level
  - at best: find physics beyond the Standard Model
  - at least: set a new limit for a given physics model

Typical Design of a Search

Physics assumption

- search for SM Higgs at mass 115 GeV
- search for *Z′* (mass larger then *Z* mass)
- new physics (NP) appears in tail of high transverse momenta

Choice of observable

- SM Higgs: best channel $H \rightarrow \gamma\gamma$, observable $m_{\gamma\gamma}$
- *Z′*: prime decay channel $Z' \rightarrow \mu^+\mu^-$, observable $m_{\mu\mu}$
- NP: all final states with identifiable objects (*j* – jet)
  - μμ, μμμ, μe, μμe, ee, eee, eeμ, *jj*, *jjj*, ...
  - use events falling into the high p_T portion

Typical Design of a Search

Optimization with Monte Carlo / data

- find quantity Q_S which defines sensitivity of analysis
  - classical optimization quantities: $S/\sqrt{B}$ or $S/\sqrt{S+B}$
  - $S/\sqrt{B}$ used for search without knowledge of cross section
    - careful, unnatural behavior at low S
  - $S/\sqrt{S+B}$ used when signal cross section is known
    - careful, it really optimizes the measurement if search successful
  - improved quantity: $S/(1.5+\sqrt{B})$
  - see www.cmsaf.mit.edu/twiki/bin/view/Class8882/WebHome
- find optimal set of cuts maximizing Q_S

Include systematic uncertainties

- in many cases the effect is small
- for large uncertainties statistical methods complicated: see www.cmsaf.mit.edu/twiki/bin/view/Class8882/WebHome

Typical Design of a Search

Look at fully optimized data analysis

- find a signal: determine significance
  - method reasonably straight forward
  - see examples in what follows
- conventions
  - observation or discovery at 5 standard deviations
  - evidence is to mark the turf, at about 3 standard deviations
  - evidence became more popular since experiments take so long to accumulate statistics to have an observation
- set a limit at a given confidence level (usually 95%)
  - once sensitivity is known confidence level straight forward
  - usually implies Poisson statistics: few events are found
  - see examples in what follows

Clever Hans Effect

Astonishing horse (real story)

- Hans von Osten in early 1900 could do math!
- Hans added up numbers, result communicated in number of pawns on the ground
- trainer cues? no, trainer was send outside, hmmm?
- 1907 Oskar Pfungst proposed: *result should be unknown to people in the room*: Hans lost all ability to add numbers
- horse sensed cues from people who knew the answer
- really a clever horse
- imagine how 'clever' we could be ....

*The Hawthorne Experiment**

Study of illumination** (optimal value for productivity)

- Study 1a: In the first experiment, there was no control group. The researchers experimented on three different departments; all showed an increase of productivity, whether illumination increased or decreased.
- Study 1b: A control group had no change in lighting, while the experimental group got a sequence of increasing light levels. Both groups substantially increased production, and there was no difference between the groups. This naturally piqued the researchers' curiosity.

* Hawthorne experiments have been intensely discussed. Read up on the web if you like.
** source is wikipedia

The Hawthorne Experiment

Study of illumination (optimal value for productivity)

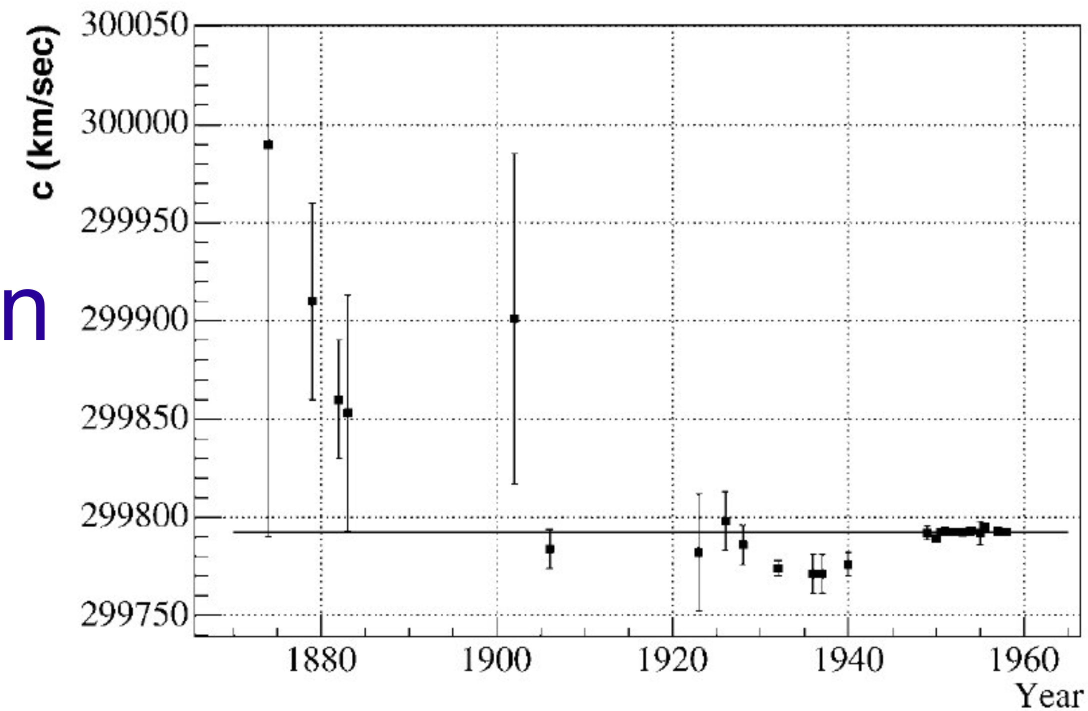
- Study 1c: The researchers decided to see what would happen if they decreased lighting. The control group got stable illumination; the other got a sequence of decreasing levels. Surprisingly, both groups steadily increased production until finally the light in experimental group got so low that they protested and production fell off.
- Study 1d: This was conducted on two women only. Their production stayed constant under widely varying light levels. It was found that if the experimenter said bright was good, they said they preferred the light; the brighter they believed it to be, the more they liked it. The same was true when he said dimmer was good. If they were deceived about a change, they said they preferred it. Researchers concluded that their preference on lighting level was completely subjective - if they were told it was good, they believed it was good and preferred it, and vice versa.

Effects in Physics

Measurements of the speed of light*

- measurements around 1930-40 are clearly off
- investigating these measurements it was concluded:
  - the investigator searches for the source or sources of errors, and continues to search until he gets a result close to the accepted value. *Then he stops!*

Particle physics, even recently has many many more examples



* from http://arjournals.annualreviews.org/doi/pdf/10.1146/annurev.nucl.55.090704.151521

Feynman's Short Version

"The first principle is that you must not fool yourself – and you are the easiest person to fool."

So, what can we do?

Blind Analysis ....

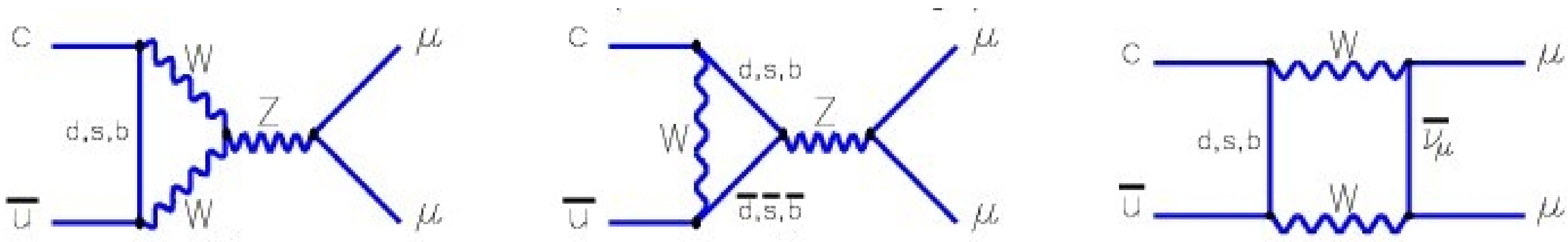
Dangerous traps in searches

- optimizing the analysis looking at data in signal region
  - small number of events, makes one vulnerable to statistical fluctuations: many examples in history
  - bias introduced through optimization procedure and human being
  - simple way out, do not even look at data: jargon blind analysis
  - device strategy where data in signal area are not used
  - usually use: data for background area and MC for signal area

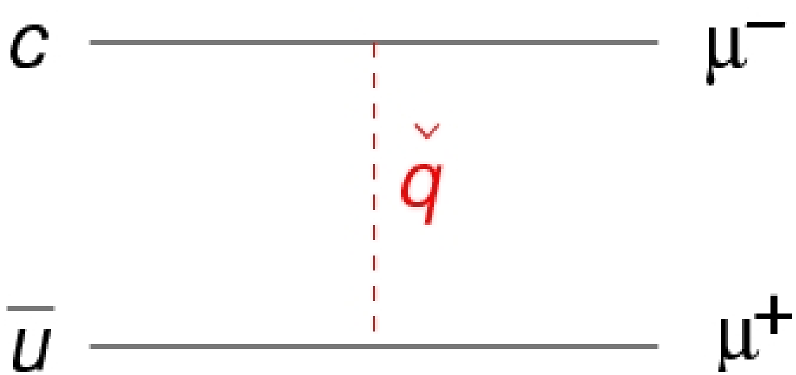
Search for Rare Decay: $D^0 \rightarrow \mu^+\mu^-$

Physics interest* (CDF analysis)

- reaction is a flavor changing neutral current reaction
- highly suppressed in Standard Model: rate exp. 10^{-13}
- SU(per)SY(mmetric) models can accommodate large branching ratio



(a) (b) (c)

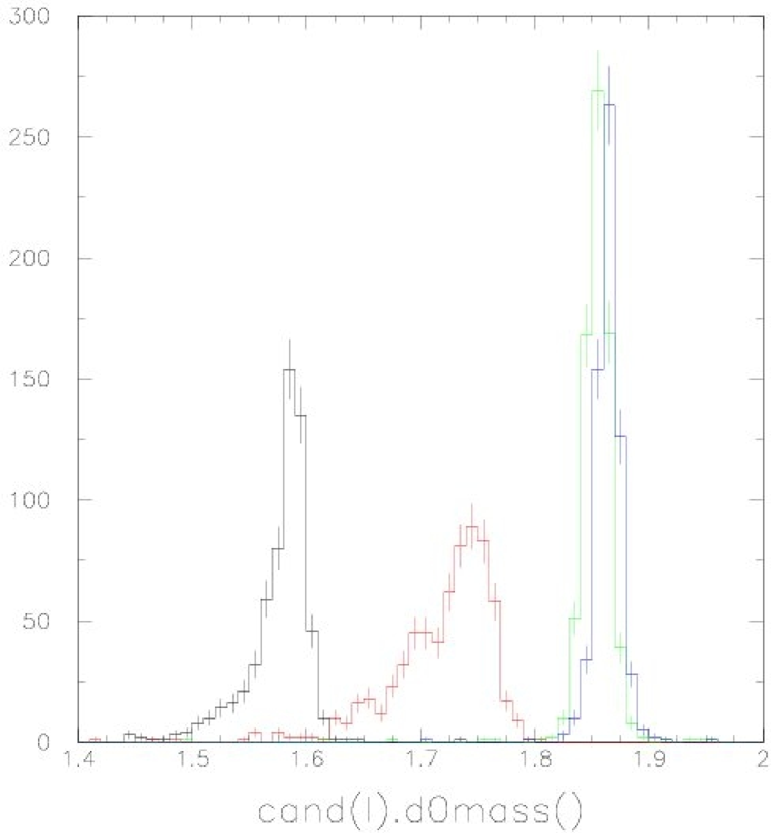


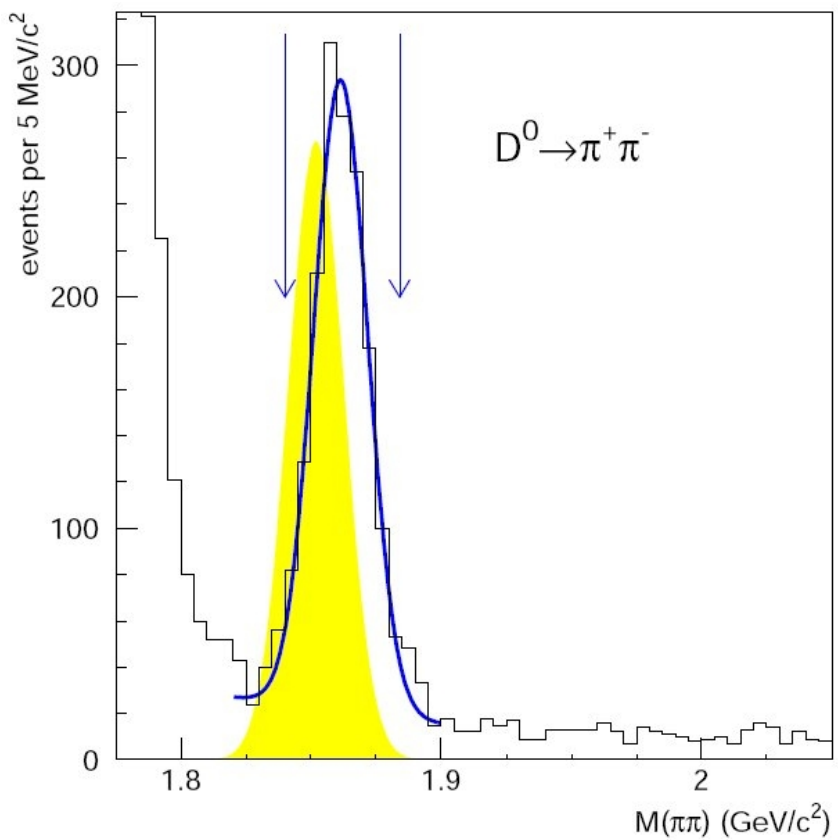
* from CDF note 6273

Search for Rare Decay: $D^0 \to \mu^+\mu^-$

Analysis outline

- use copiously produced: $D^0 \to K^- \pi^+$ and $D^0 \to \pi^- \pi^+$
- determine: acceptance, background, signal normalization
- signal looks like $D^0 \to \pi^- \pi^+$ with two muon stubs attached





Search for Rare Decay: $D^0 \to \mu^+\mu^-$

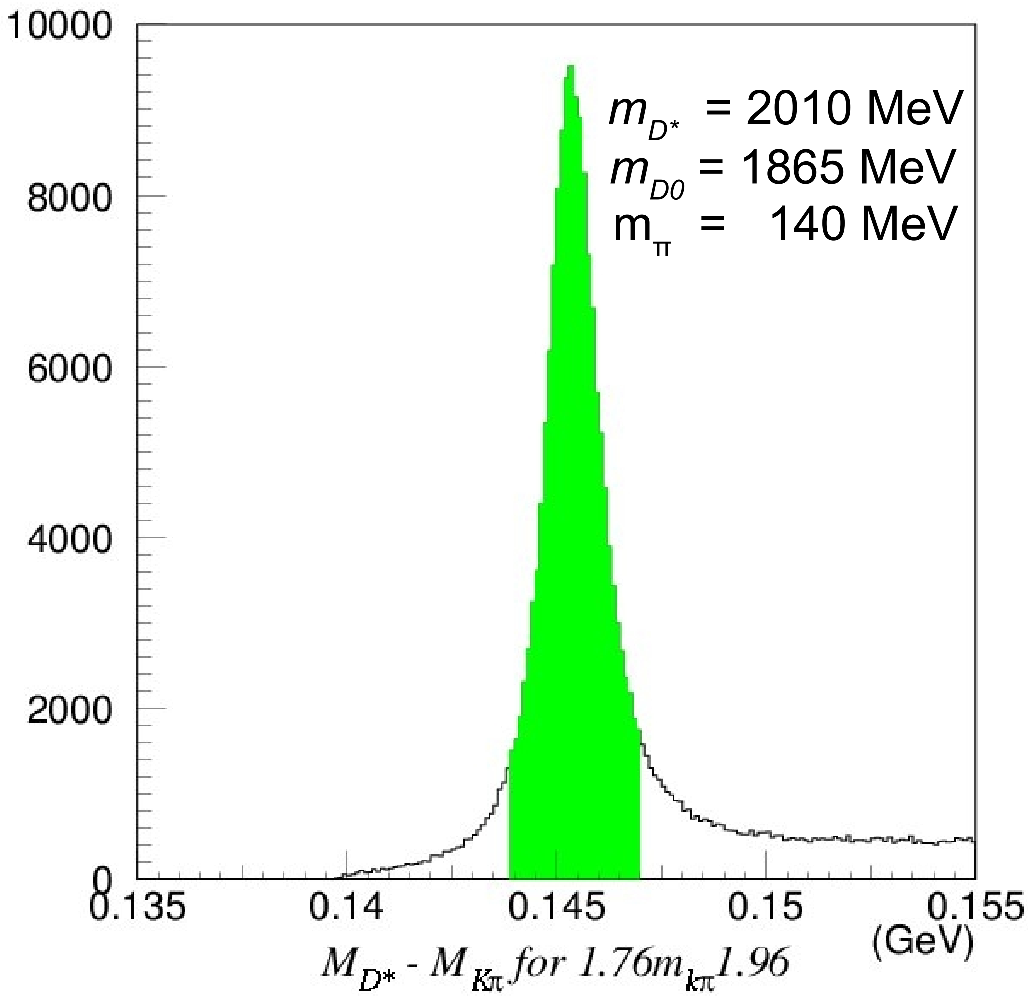
Analysis outline, *continued*

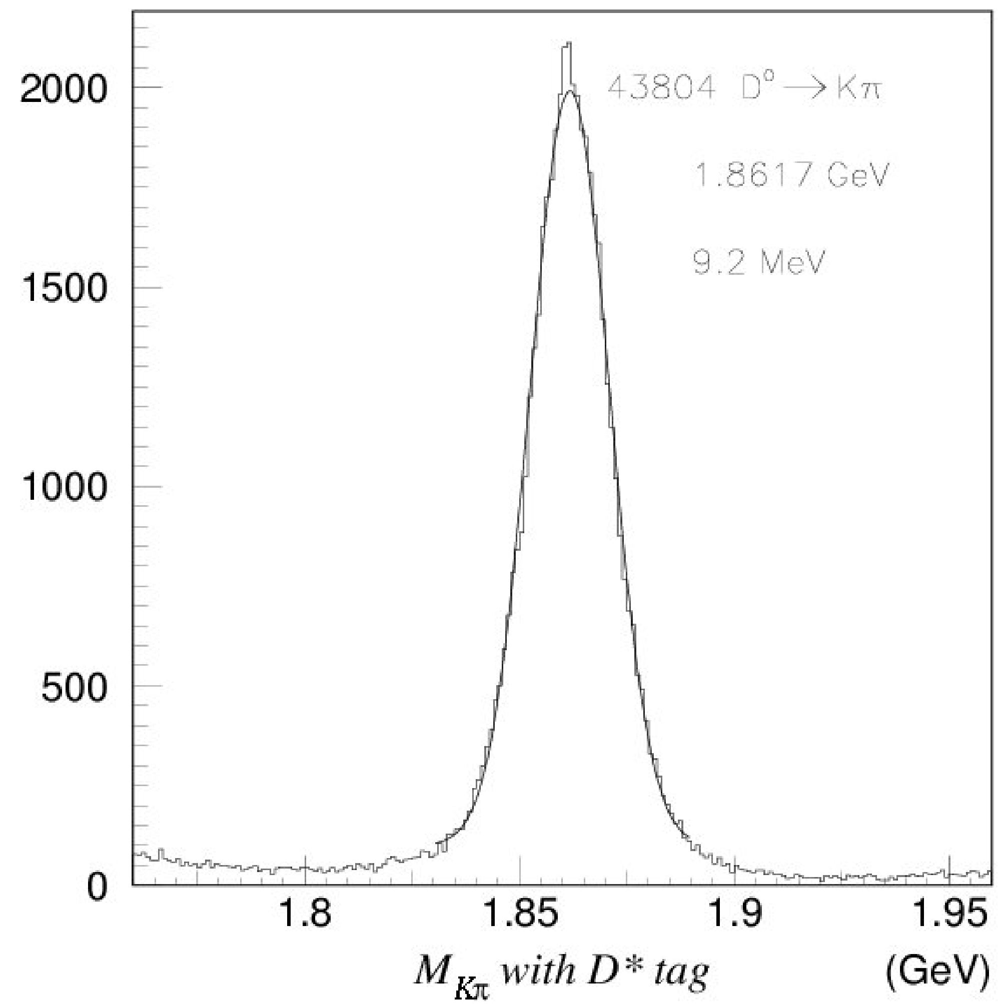
- branching ratio: $$Br(D^0 \to \mu^+\mu^-) = Br(D^0 \to \pi^+\pi^-)\frac{N(\mu\mu)\ \varepsilon(\pi\pi)}{N(\pi\pi)\ \varepsilon(\mu\mu)}$$
- limit: $$Br_{95\%CL}(D^0 \to \mu^+\mu^-) = Br(D^0 \to \pi^+\pi^-)\frac{N(\mu\mu)_{95\%CL}\ \varepsilon(\pi\pi)}{N(\pi\pi)\ \varepsilon(\mu\mu)}$$
- $N(\mu\mu)_{95\%CL}$ is upper limit of how many events could be there according to accumulated statistics we have and given number of events observed (covers 95% of cases)
    - often selection completely erases candidates, and mostly there are very few events
    - non-Gaussian statistics has to be applied, *Poisson* statistics: TMath::Poisson(n_{exp},n_{obs}) gives the probability to observe n_{obs} events when on average n_{exp} are expected
    - here $N(\mu\mu)_{95\%CL}$ is 3 events when zero events are observed (from TMath::Poisson(0,3) = 5%)

Search for Rare Decay: $D^0 \to \mu^+\mu^-$

Analysis outline, *continued*

- the D^* trick: $D^{*+} \to D^0\pi^+$ with very soft π^+
- very clean D^0 sample for studies (removes auto reflection)





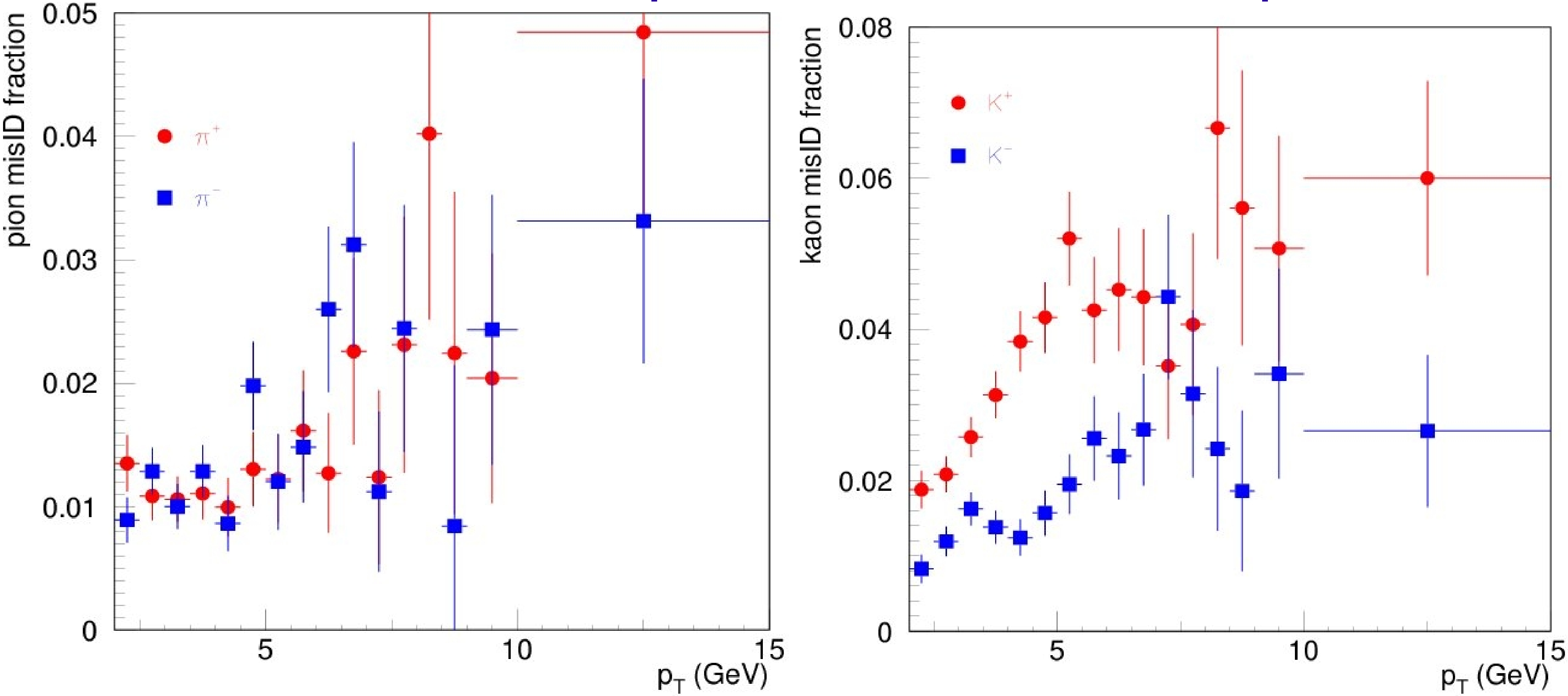
Search for Rare Decay: $D^0 \to \mu^+\mu^-$

Selection requirement optimization

- use selection quality: $Q_S = S/(1.5 + \sqrt{B})$
- blinding is implemented by not using the signal but a fake signal sample:
  - S is obtained from the $D^0 \to \pi^- \pi^+$ sample
  - B is derived from the $D^0 \to K^- \pi^+$ sample with both K and π have muons attached to them (no overlap with $D^0 \to \mu^+\mu^-$)
  - cuts are varied to maximize Q_S
- signal sample with $D^0 \to \mu^+\mu^-$ is never looked at

Search for Rare Decay: $D^0 \to \mu^+\mu^-$

Use our clean and copious $D^0 \to K^- \pi^+$ sample



Probabilities to attach muons to pions (left) or kaons (right) [decay in flight or punch through]

Search for Rare Decay: $D^0 \to \mu^+\mu^-$

background contributions

- combinatorial background: no real physics signal but simply random coincidence between muon stub and a track
- physics background: both pions decay to real muon or punch through the calorimeter and have a real muon signal (relevant only $D^0 \to \pi^-\pi^+$)

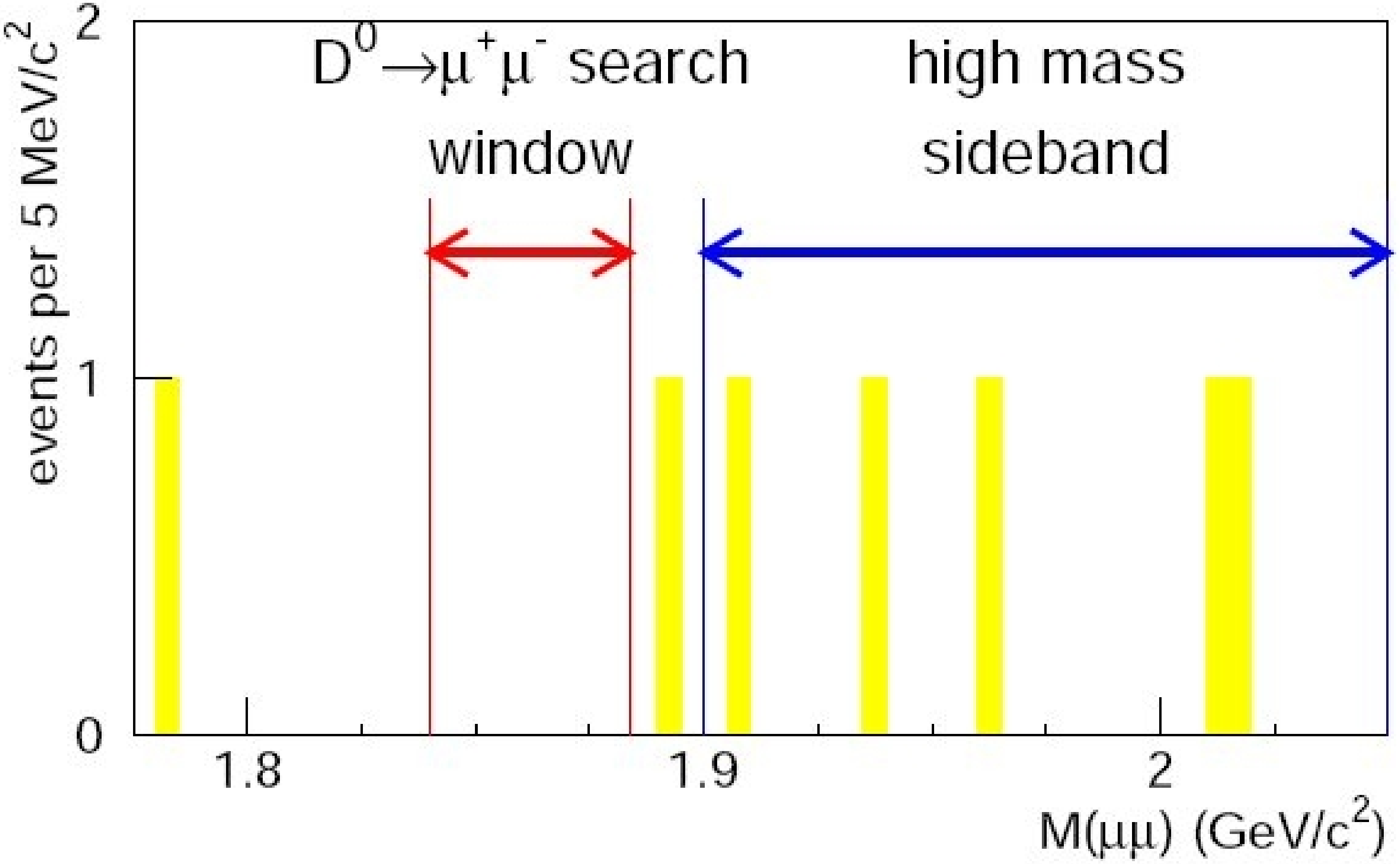
determine the background

- combinatorial background: use high mass side band of D^0 with both pions having muon stub attached (no physics contribution): 1.6 ± 0.7 events
- physics background: use $D^0 \to \pi^- \pi^+$ signal and multiply with muon attach probabilities for pions according to values measured in $D^0 \to K^- \pi^+$ sample: 0.22 ± 0.02 events
- total background: 1.8 ± 0.7 events

Search for Rare Decay: $D^0 \rightarrow \mu^+\mu^-$

Opening the box

- analysis with all systematic uncertainties is completed
- predicted number of events from background 1.8
- no events inside signal area



Search for Rare Decay: $D^0 \to \mu^+\mu^-$

Deriving the limit

- no signal is found
- derive upper limit on the branching fraction for $D^0 \to \mu^+\mu^-$
  - assume that there is no background (conservative)
  - for 95% CL we can exclude a maximum of 3 events

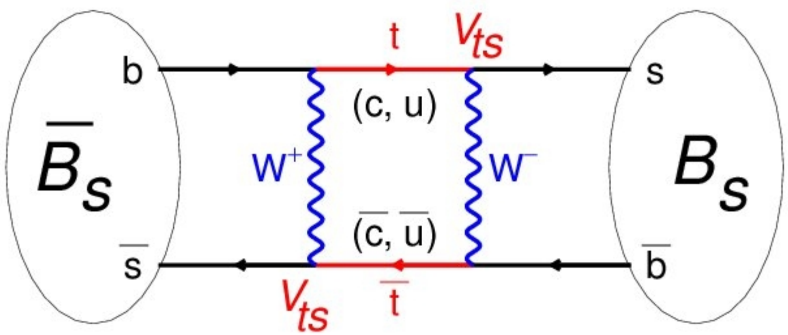
$$Br(D^0 \to \mu^+\mu^-) < 1.43 \cdot 10^{-3} \cdot \frac{3}{1412} \cdot 1.08 = 3.3 \times 10^{-6}$$

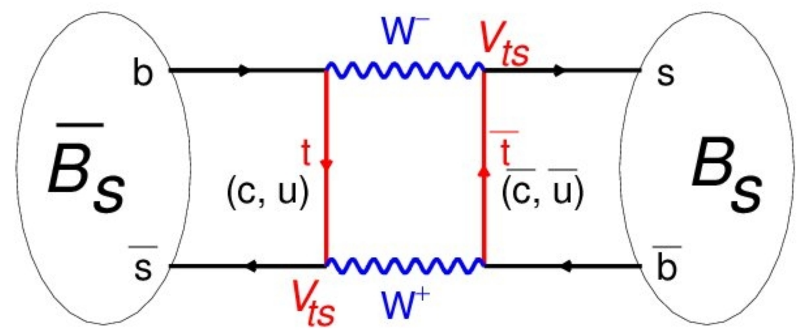
Also included are systematic uncertainties

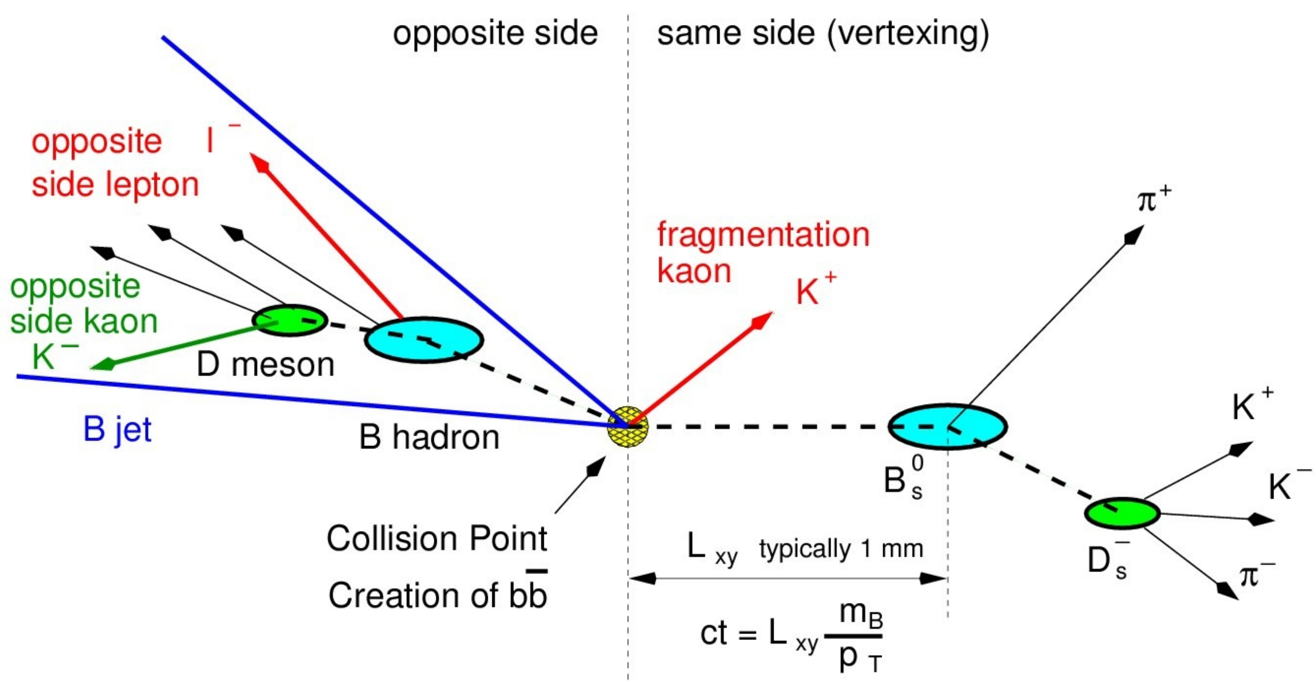
- all turned out to be negligible ($\delta S < 5\%$)
- formula for inclusion is: $\Delta N_{95\%CL} = 0.5 \left(N_{95\%CL} \frac{\delta S}{S} \right)^2$

B_s Oscillations

Neutral Mesons like K^0, B^0, B_s can instantaneously switch into their anti particles (higher order)



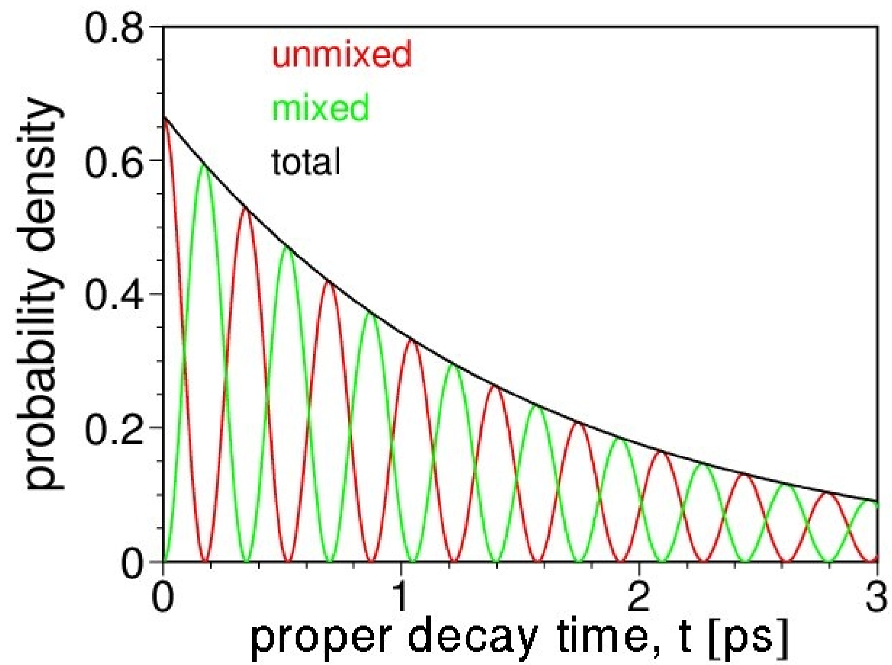




B_s Mixing Phenomenology

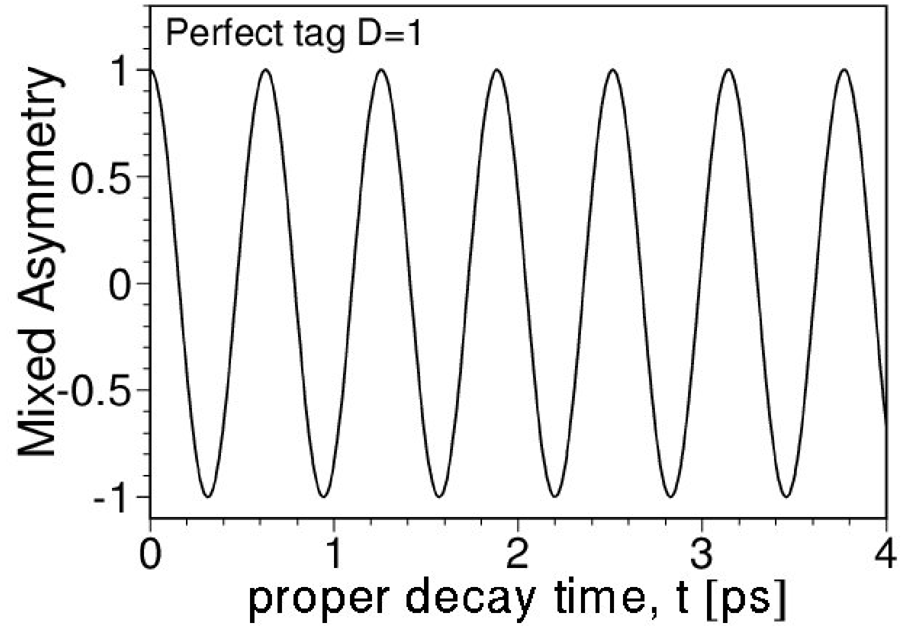
Behavior in proper time

$$P(t)_{B^0\to B^0} = \frac{1}{2\tau}e^{-t/\tau}(1+\cos\Delta m t)$$
$$P(t)_{B^0\to \overline{B^0}} = \frac{1}{2\tau}e^{-t/\tau}(1-\cos\Delta m t)$$



Determine asymmetry

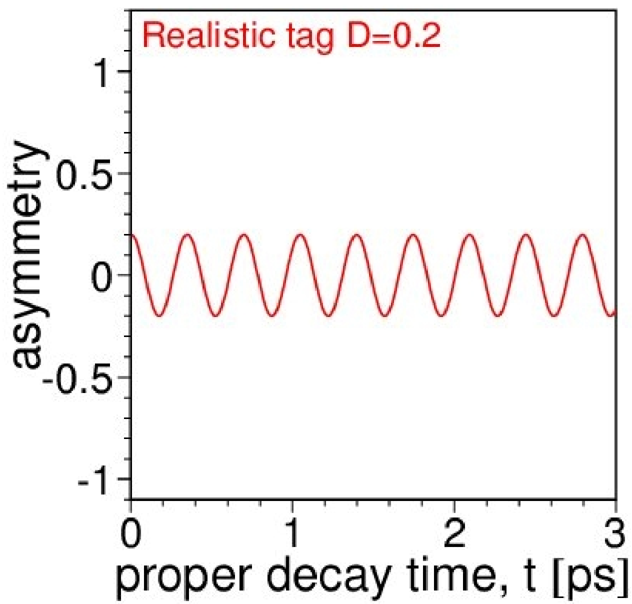
$$A_0(t) = \frac{N(t)_{unmixed} - N(t)_{mixed}}{N(t)_{unmixed} + N(t)_{mixed}} = \cos\Delta m\, t$$



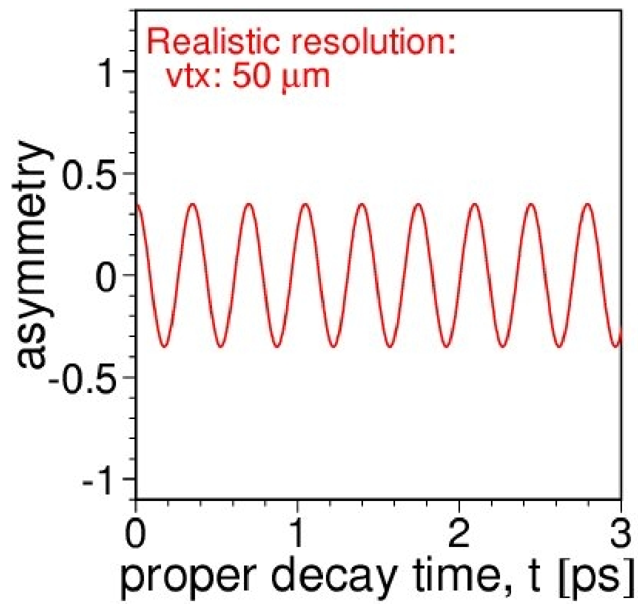
In a perfect world

What Do We See in the End?

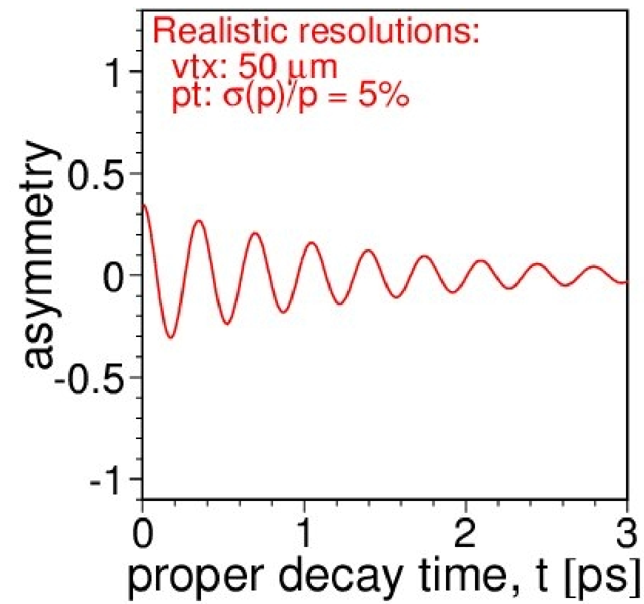
Flavor tagging



Vertex



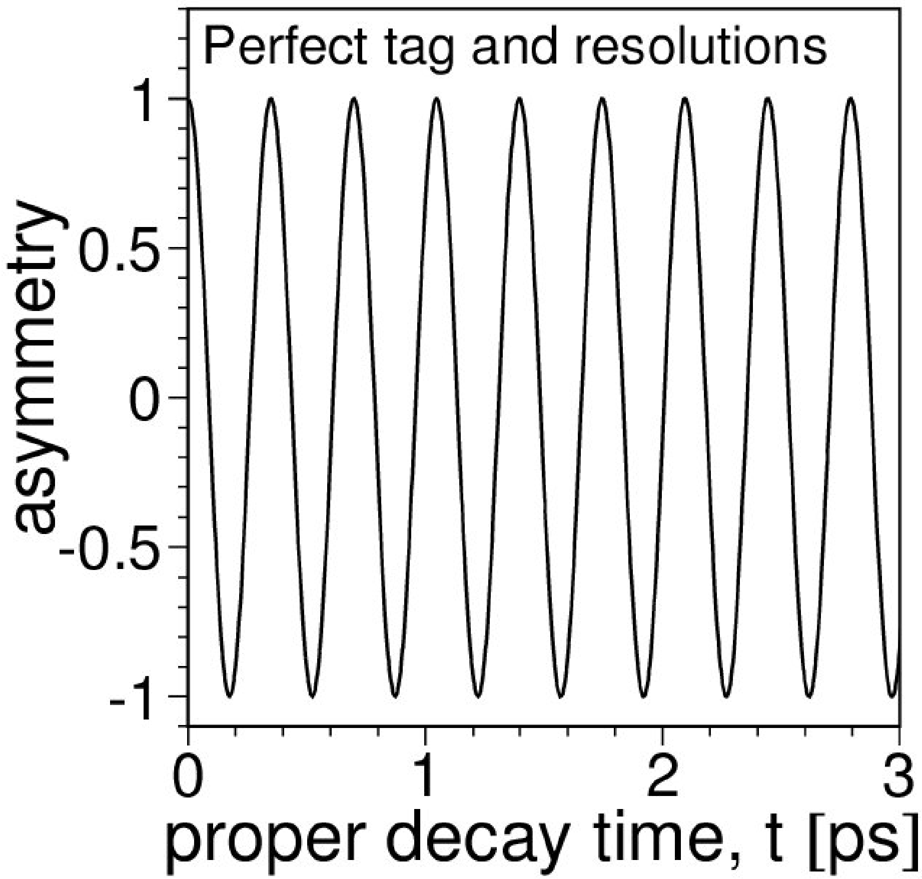
Vertex and Momentum

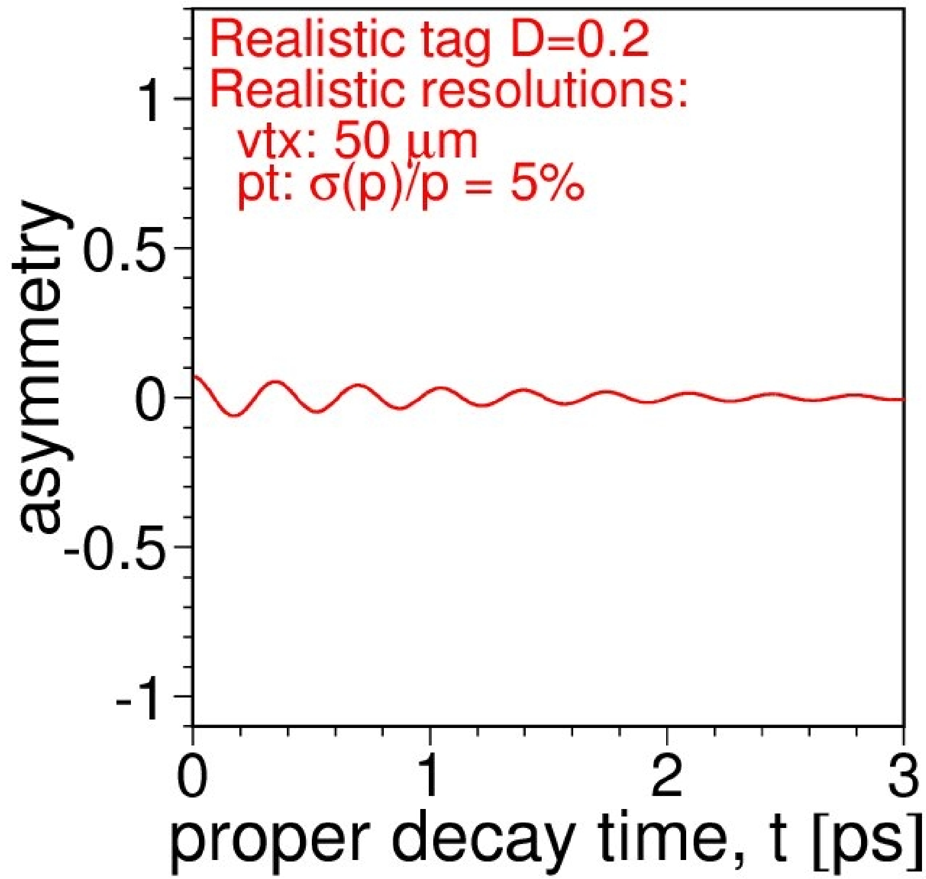


$$1/\sigma_A = \sqrt{\frac{n_S \varepsilon D^2}{2}} \sqrt{\frac{n_S}{n_S + n_B}} \exp\left(-\frac{(\Delta m_s \sigma_{ct})^2}{2}\right)$$

$$\sigma_{ct} = \sqrt{(\sigma_{ct}^0)^2 + \left(ct\, \frac{\sigma_p}{p}\right)^2}$$

Perfect World to the Real World



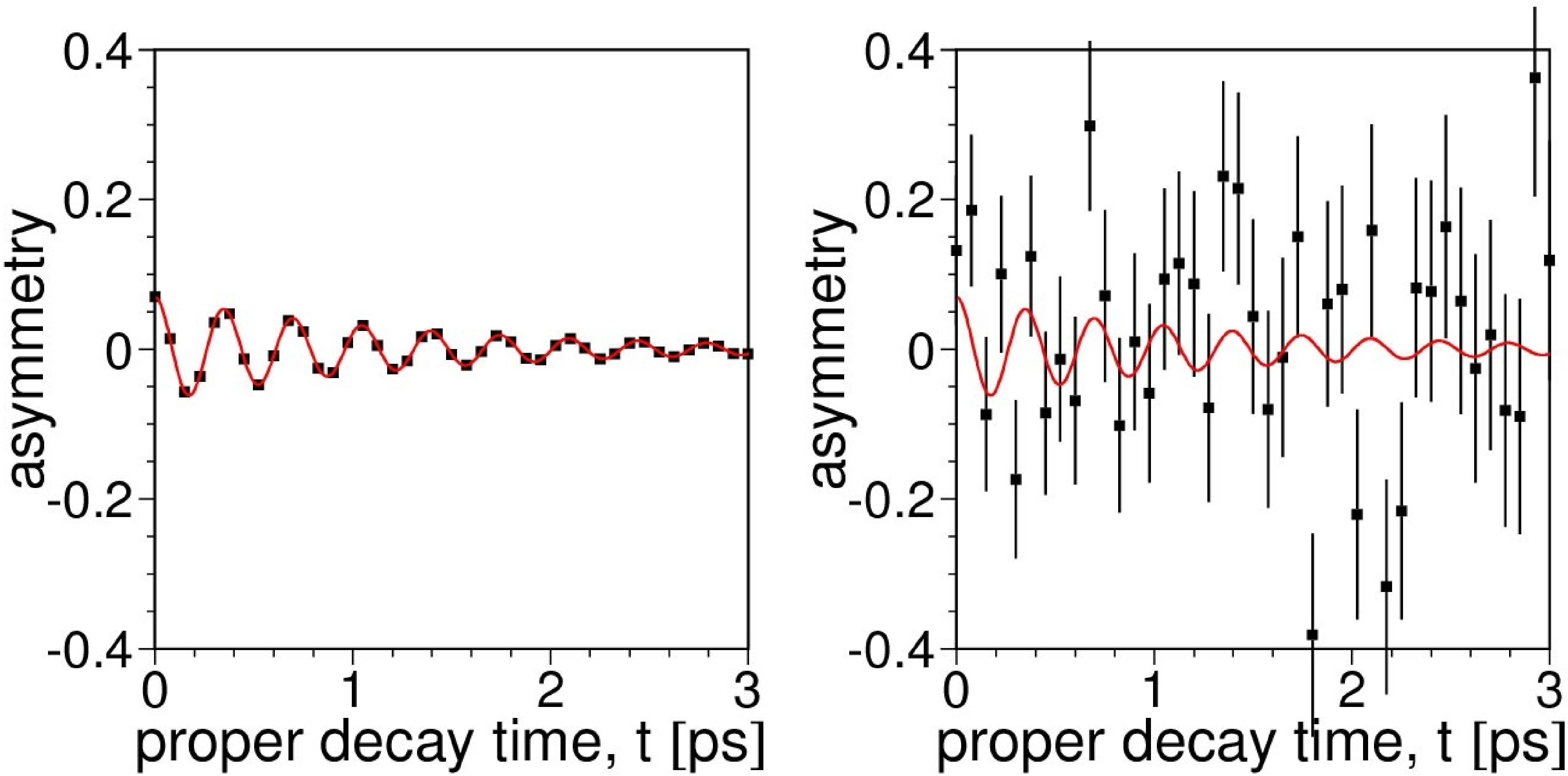


Unbinned likelihood fit: $p \sim exp(-t/\tau)(1 \pm AD \cos \Delta mt)$

- scan Δm_s for signal: determine amplitude, A
- measure Δm_s with $A = 1$

apply Fourier analysis

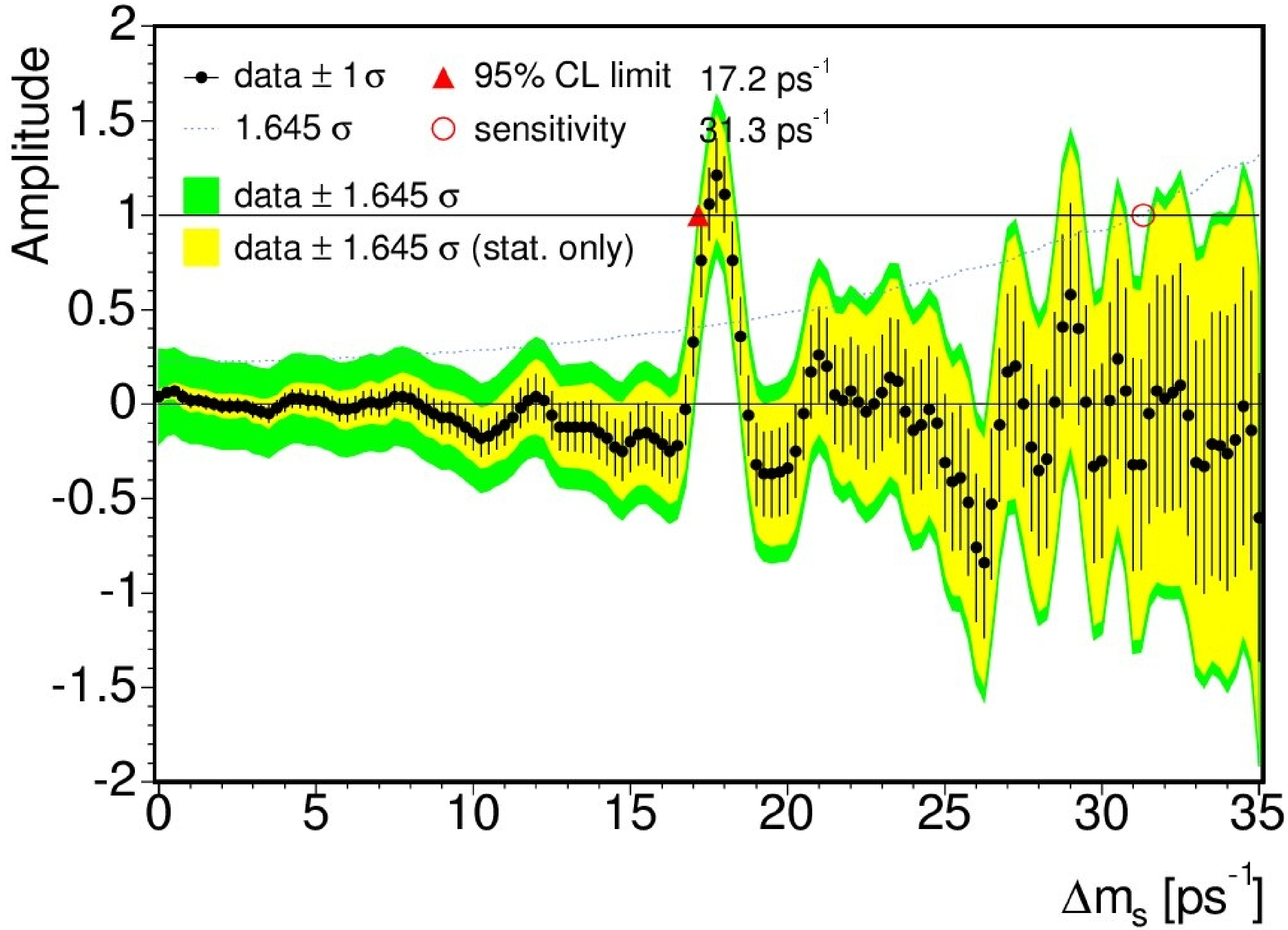
How Would a Signal Look Like?



Unbinned likelihood fit: $p \sim exp(-t/\tau)(1 \pm AD \cos \Delta m t)$

- scan Δm for signal: determine amplitude, A
- measure Δm_s with $A = 1$

Bs Mixing Signal



$A = 1.21 \pm 0.20(\text{stat})$ compatible with 1 for $\Delta m_s \sim 17.75\ \text{ps}^{-1}$

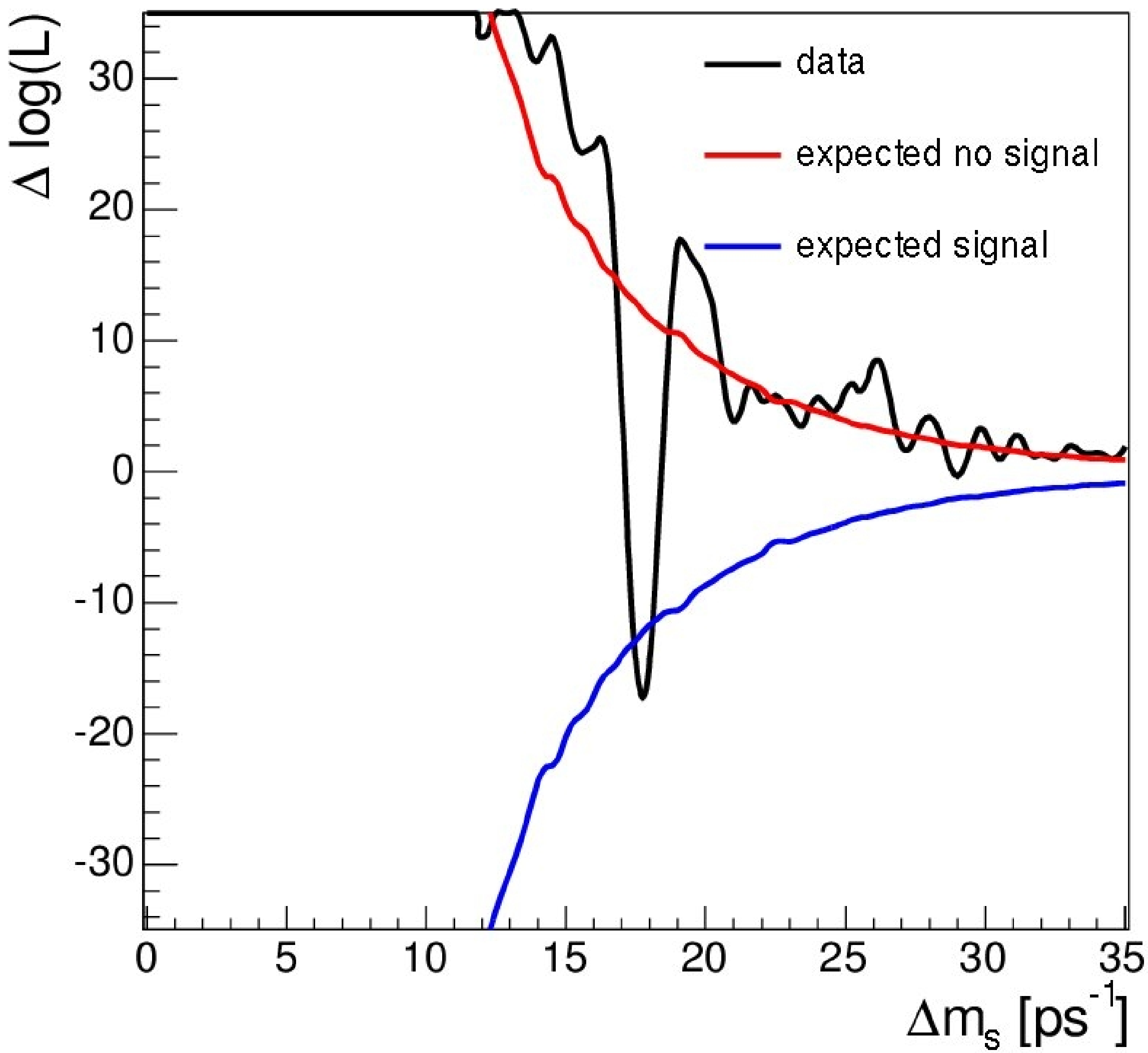
$A/\sigma_A(\Delta m_s = 17.75\ \text{ps}^{-1}) = 6.05$, but what is the p-value?

Likelihood Profile

Difference, $-\Delta \log(L)$

$\log(\mathcal{L}(A=1)) - \log(\mathcal{L}(A=0))$

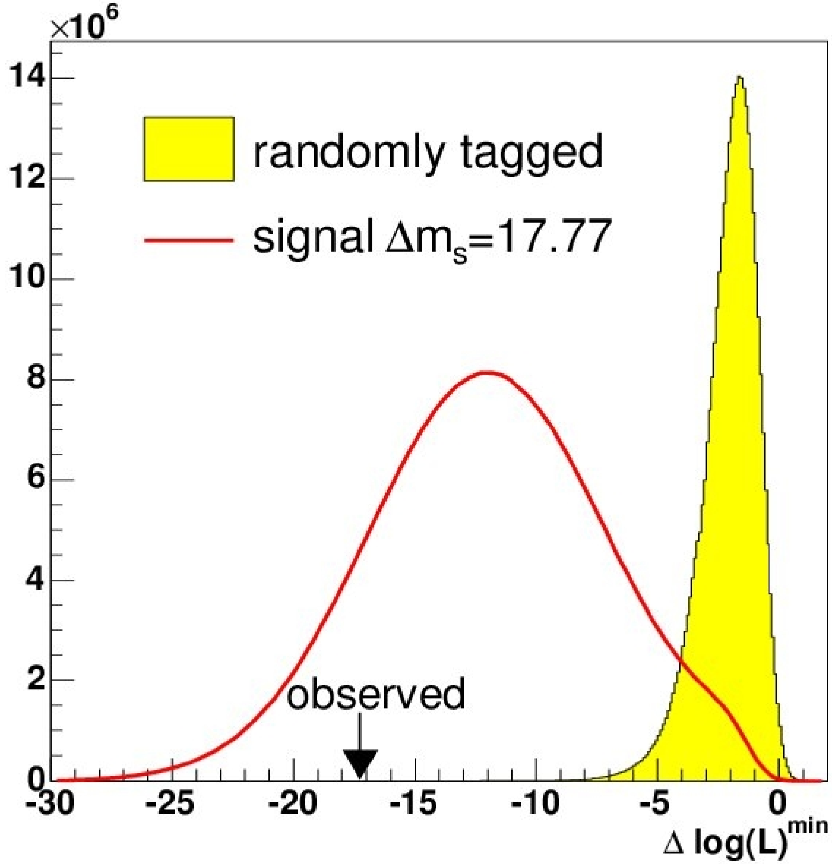
Minimum: −17.26

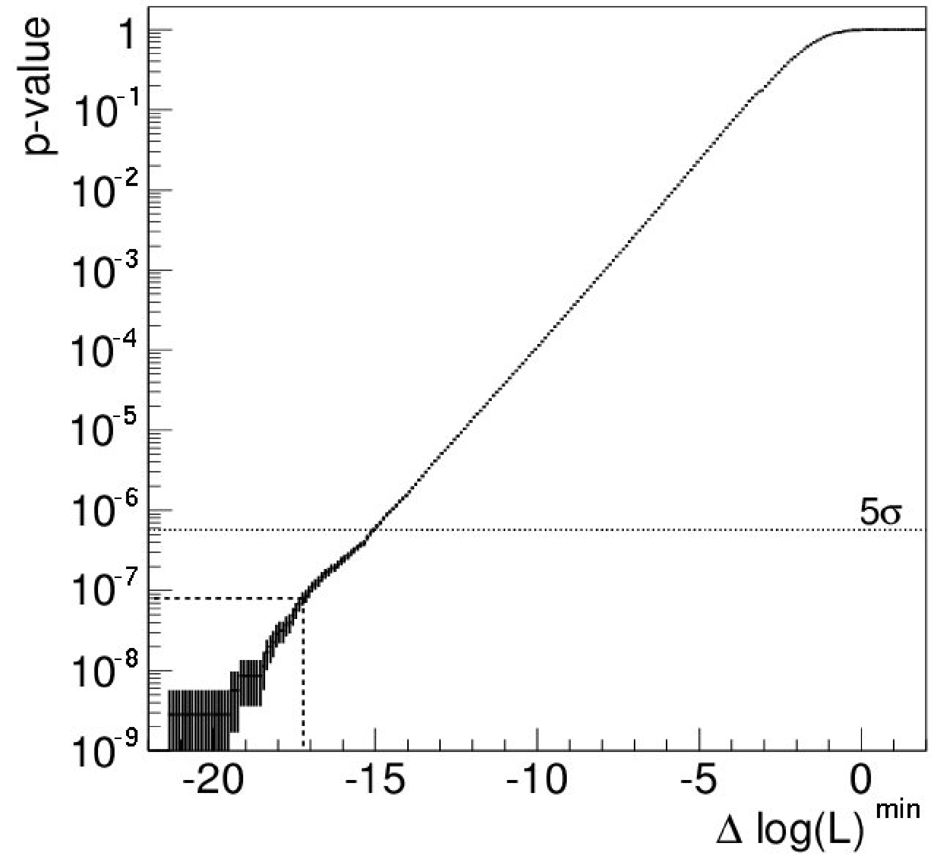


Key question:

How often can random tags produce a minimum at least as deep?

Likelihood Significance





28 trials out of 350 million

p-value $\approx 8 \times 10^{-8}$ corresponding to 5.4σ

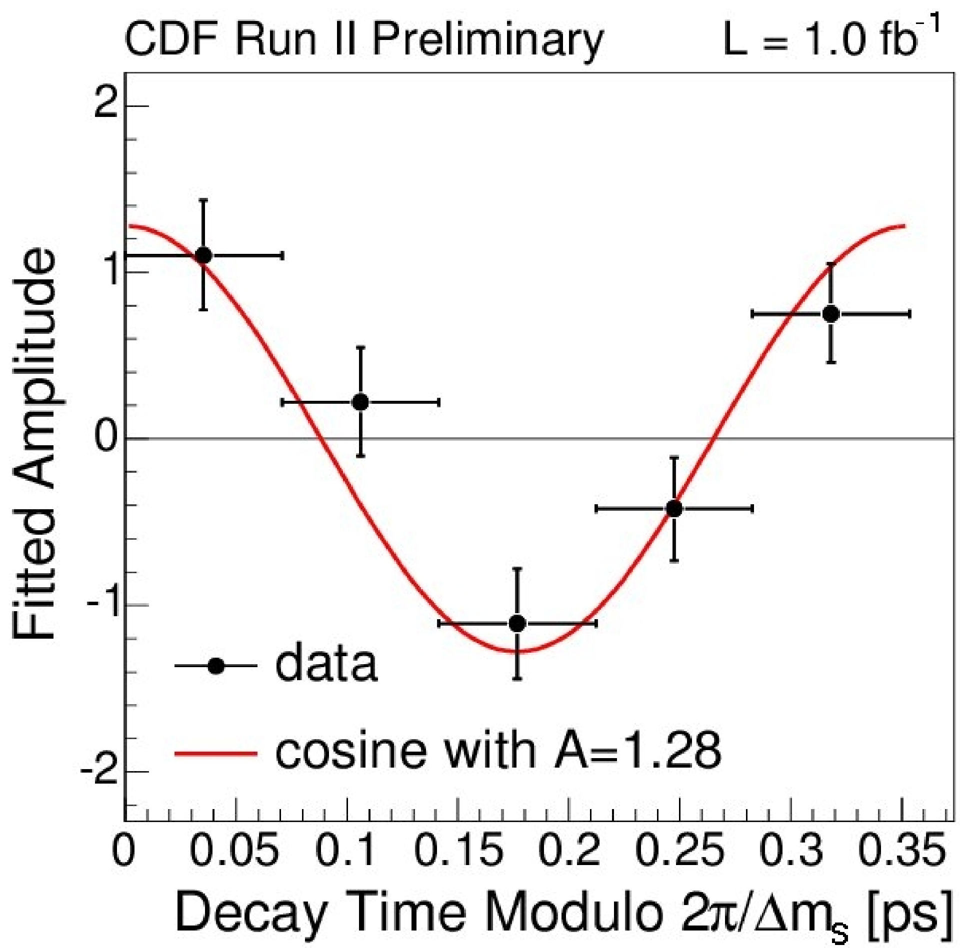
(5 standard deviations is $= 5.7 \times 10^{-7}$)

$\rightarrow$ passed observation criterion

Folding All Oscillations on Top

The eye is not good at Fourier transforms on the fly

- take events and fold them on top of each other with the measured frequency



Conclusion

Searches versus measurements

- search looks for a so far unobserved signal (most likely in a mass distribution but other distributions possible)
- searches start with some assumptions, they could be very general (more difficult) or very specific (gives more handles but is less general)
- signal is accepted as observed with five standard deviations, evidence around 3 standard deviations
- non-observation needs more work to evaluate carefully analysis sensitivity to be able to exclude a signal with some confidence level
- measurements analyze a known process and measure a quantity which is either a SM parameter or can be predicted in the Standard Model (test SM hypothesis)
- measurement has potential to find SM disagreement

Next Lecture

Efficiency and Acceptance