Measurement of the B_s Meson Oscillation Frequency with CDF II

Alessandro Cerri

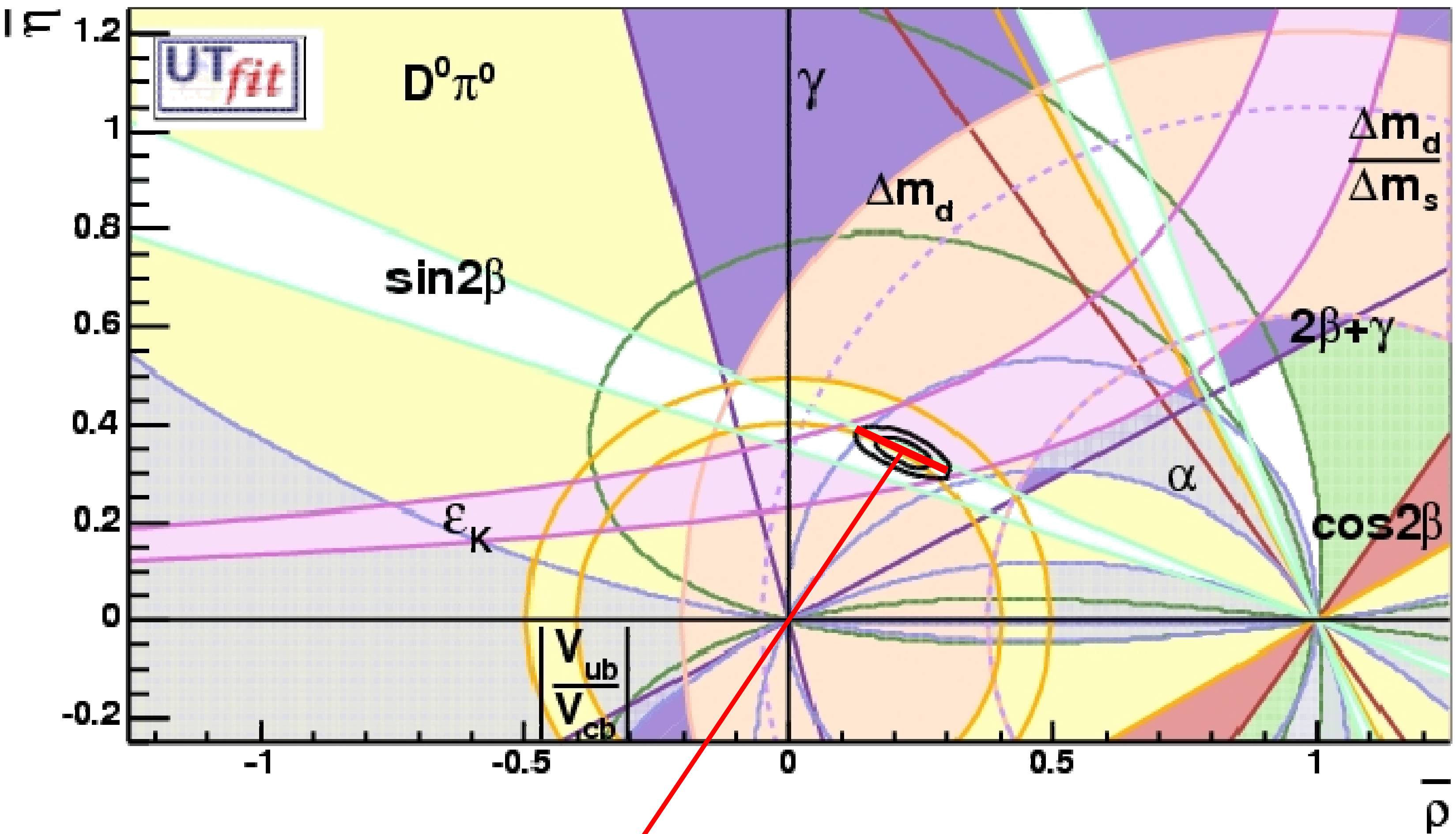
LBNL

May 11th 2006

Synopsis

- Introduction
- Quick reminder on the ingredients
- A sound statistical approach
- The D0 result
  - Sample and scans
  - Evidence of a signal?
- The CDF result
  - Samples
  - Results (subsamples & combined)
  - Implications on CKM/BSM
- What's next?
- Conclusions

Two months ago…

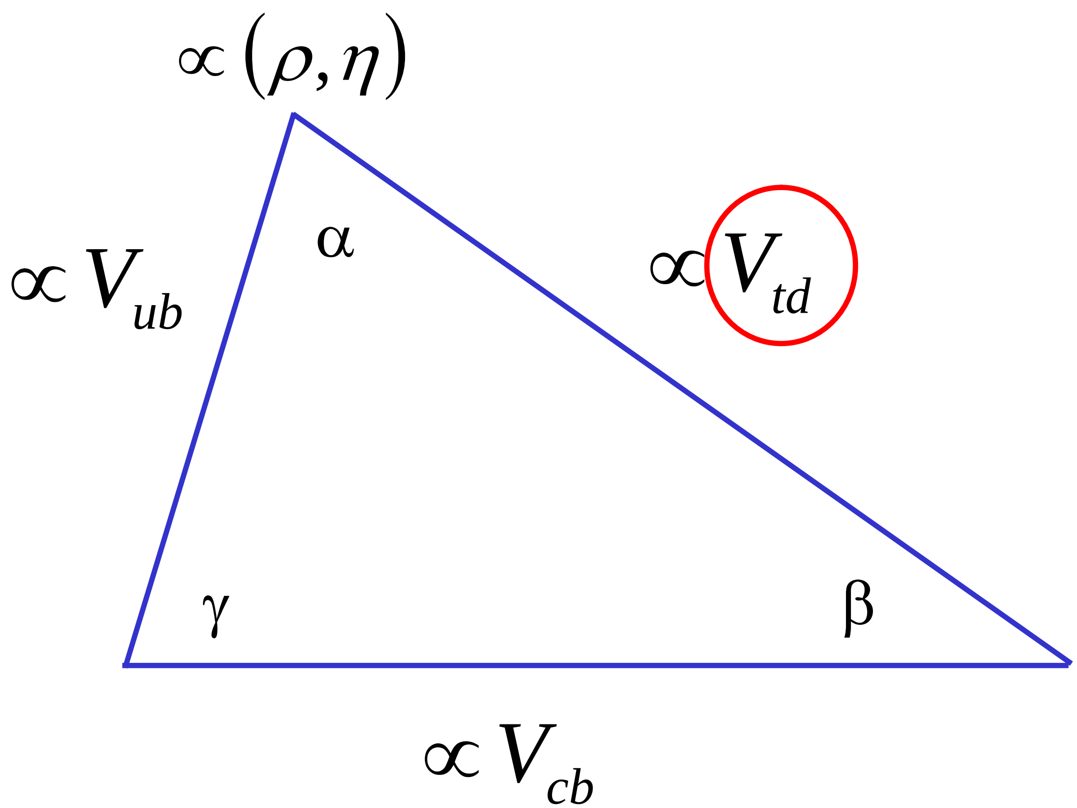


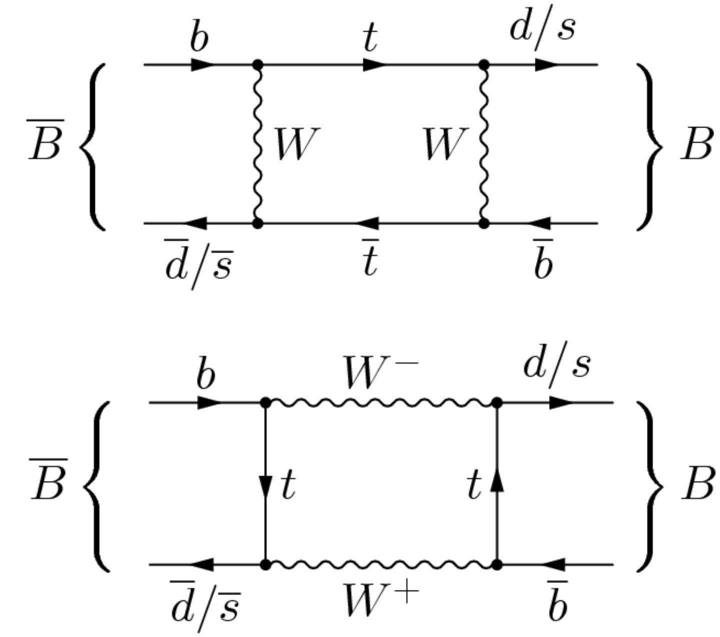
TeVatron contribution is critical!

What happened in the last two months?

- D0 came out with a result based on $1fb^{-1}$ (Moriond)
- CDF shortly afterwards released its latest greatest result (but not the last word) on $1fb^{-1}$
- Bottom line?
  - Evidence of a mixing 'signal'
  - Not enough statistical power for 'observation' (5σ)
  - If signal is there $\Delta m_s = 17.33^{+0.42}_{-0.21} \pm 0.07(\mathrm{syst})\mathrm{ps}^{-1}$
- I will focus mostly on how to read these results and what one should take out of them

Why so much interest around Δm_s?





- V_{td} is derived from mixing effects
- QCD uncertainty is factored out in this case resorting to the relative Bs/Bd mixing rate (V_{td}/V_{ts})
- Beyond the SM physics could enter in loops!

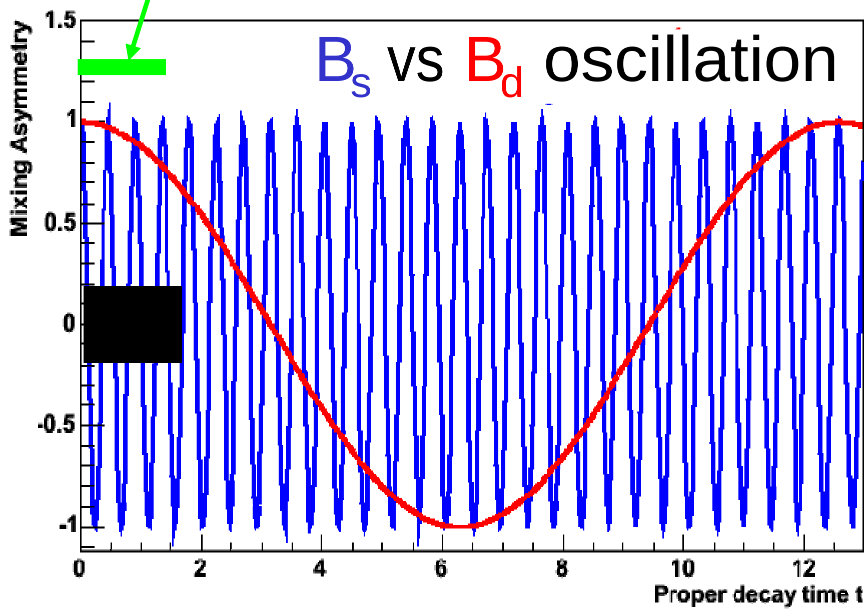
B_s Mixing 101

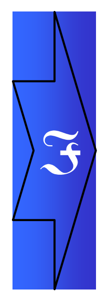
$$A = \frac{N_{unmix} - N_{mix}}{N_{unmix} + N_{mix}} \propto \cos(\Delta m\, t)$$

- $\Delta m_s >> \Delta m_d$
- Different oscillation regime → Amplitude Scan

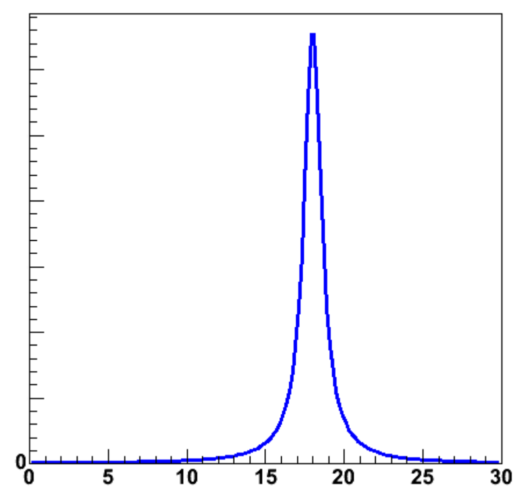
Perform a 'fourier transform' rather than fit for frequency

B lifetime





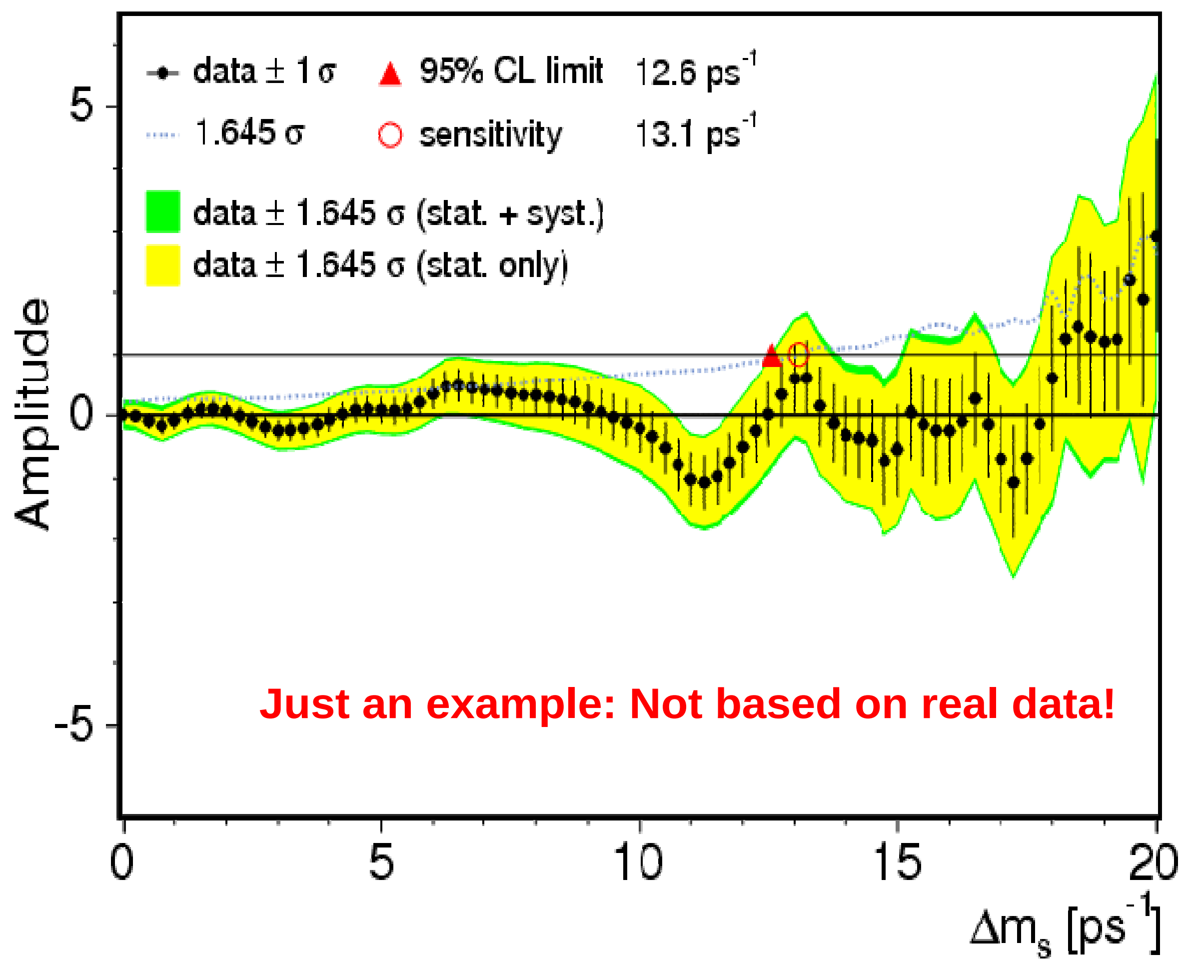
A



Δm_s [ps^{-1}]

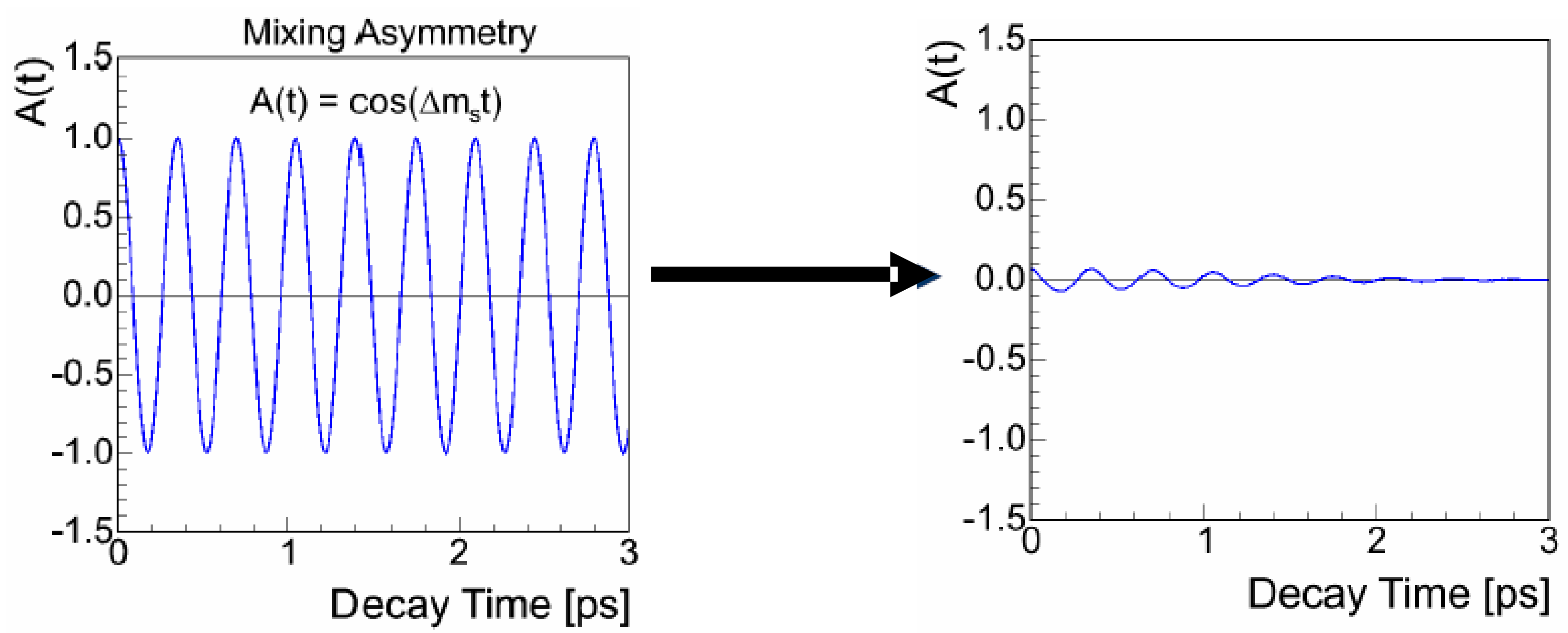
Amplitude Scan: introduction

- Mixing amplitude fitted for each (fixed) value of Δm
- On average every Δm value (except the true Δm) will be 0
- "sensitivity" defined for the average experiment [mean 0]
- The actual experiment will have statistical fluctuations
- Actual limit for the actual experiment defined by the systematic band centered at the measured asymmetry
- Combining experiments as easy as averaging points!



Is this an effective tool to search for a signal?

Mixing in the real world

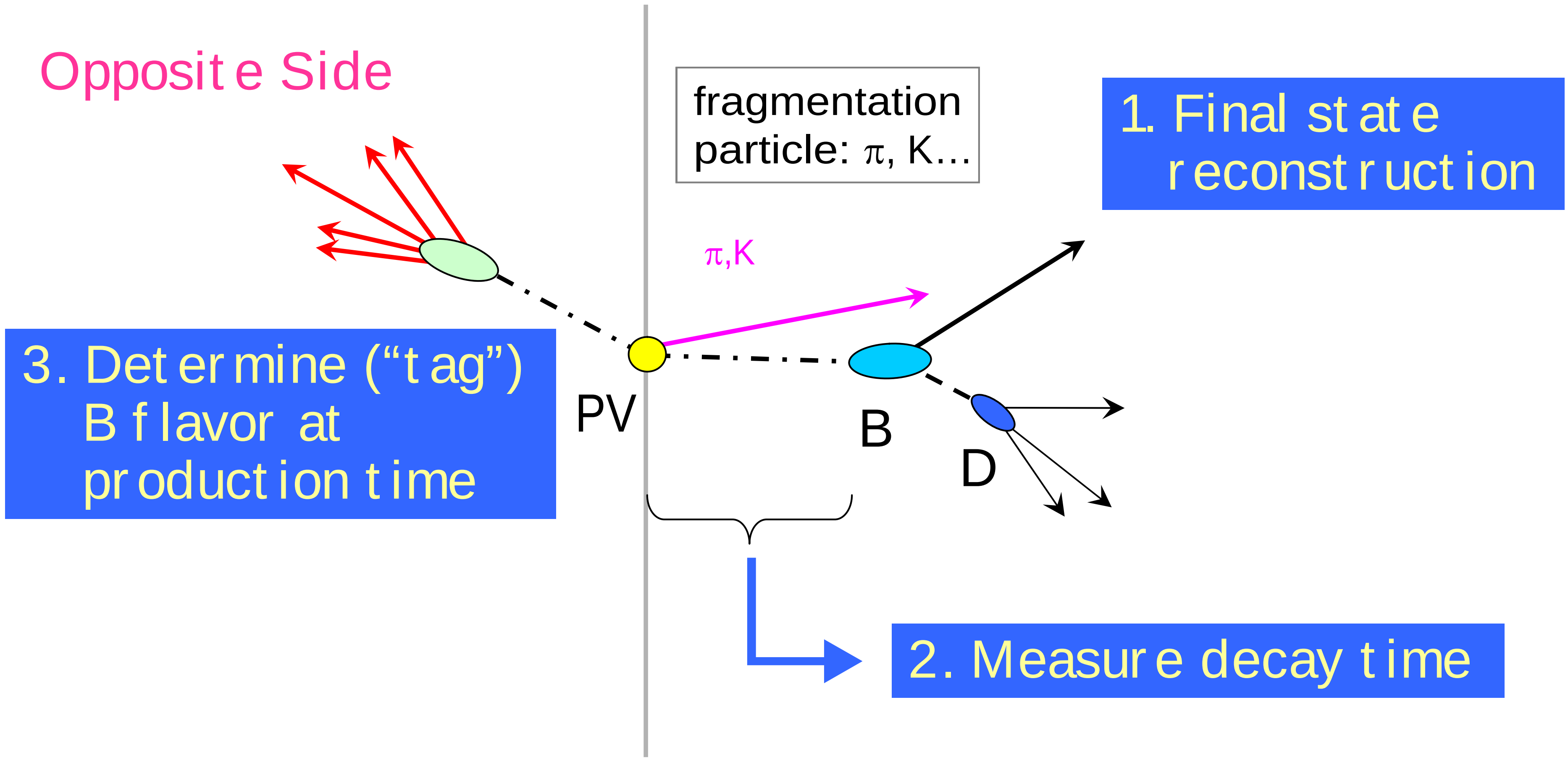


Flavor tagging power

Proper time resolution

$$\frac{1}{\sigma} = \sqrt{\frac{S\varepsilon D^2}{2}}\, e^{-\frac{(\Delta m_s \sigma_{ct})^2}{2}} \sqrt{\frac{S}{S+B}}$$

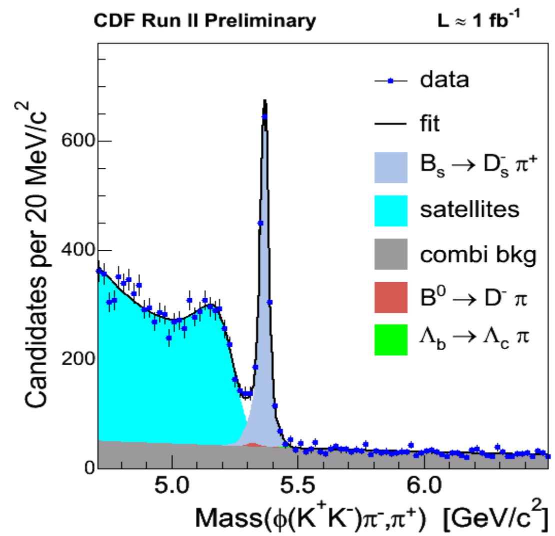
One Slide Summary: Mixing Measurements

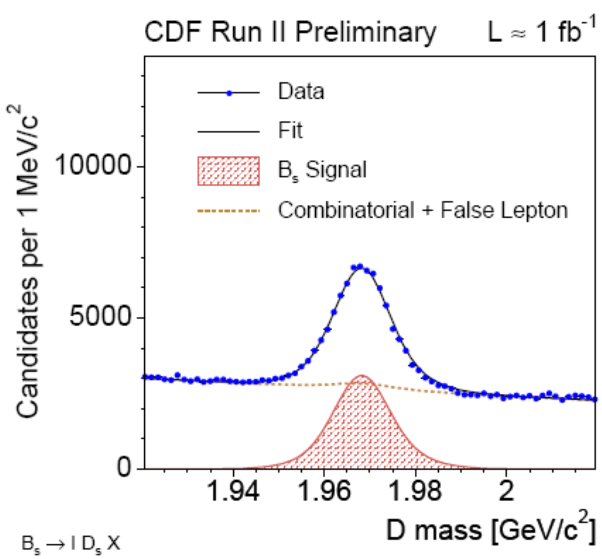


Signals

$$Significance = \sqrt{\frac{S\varepsilon D^2}{2}}e^{-\frac{(\Delta m_s \sigma_t)^2}{2}}\sqrt{\frac{S}{S+B}}$$

$B_s \to D_s\pi$	Yield	s/b
$D_s \to \phi\pi$	1600	~4:1
$D_s \to K^*K$	800	~2:1
$D_s \to \pi\pi\pi$	600	~1:1





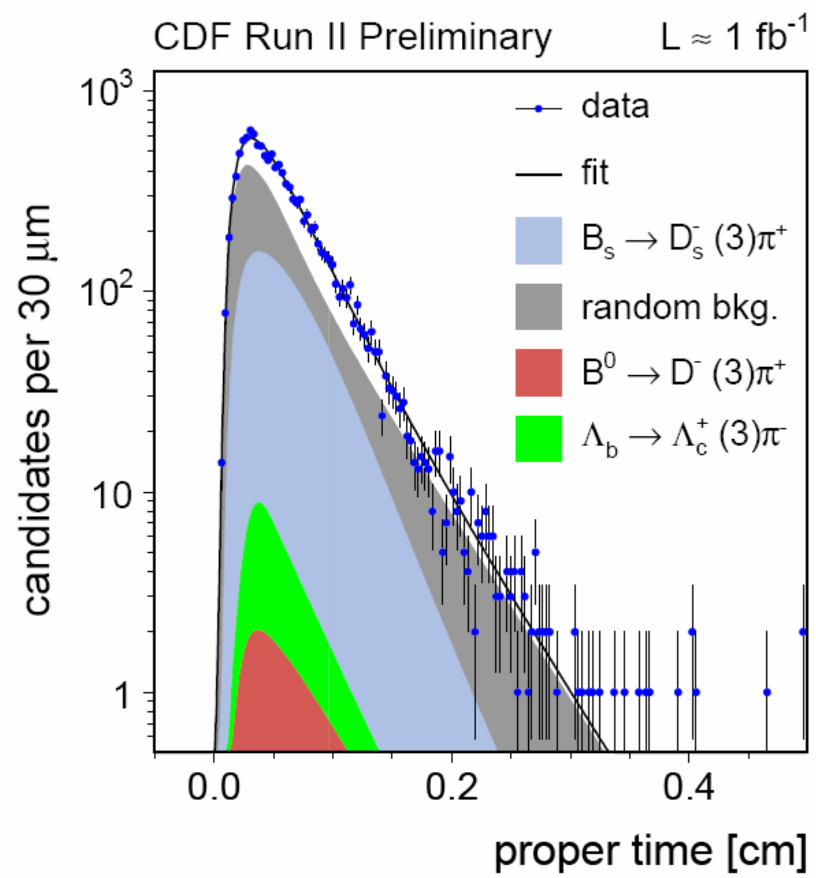
$B_s \to D_s l \nu$	Yield	s/b
$D_s \to \phi\pi$	32300	~2:1
$D_s \to K^*K$	10900	~1:2
$D_s \to \pi\pi\pi$	10100	~1:5

B_s Mixing Ingredients

Proper time reconstruction

$$Significance = \sqrt{\frac{S\varepsilon D^2}{2} e^{-\frac{(\Delta m_s \sigma_t)^2}{2}}} \sqrt{\frac{S}{S+B}}$$

Hadronic Lifetime Results



~3000 candidates

$$ct = \frac{L_{xy}}{\beta\gamma} = \frac{L_{xy} m_B}{P_t} \rightarrow \frac{\sigma_{ct}}{ct} = \frac{\sigma_{L_{xy}}}{L_{xy}} \oplus \frac{\sigma_{P_t}}{P_t}$$

Mode	Lifetime [ps] (stat. only)
$B^0 \to D^- \pi^+$	1.508 ± 0.017
$B^- \to D^0 \pi^-$	1.638 ± 0.017
$B_s \to D_s \pi (\phi\pi)$	1.538 ± 0.040

- World Average:

B^0 1.534 ± 0.013 ps^{-1}

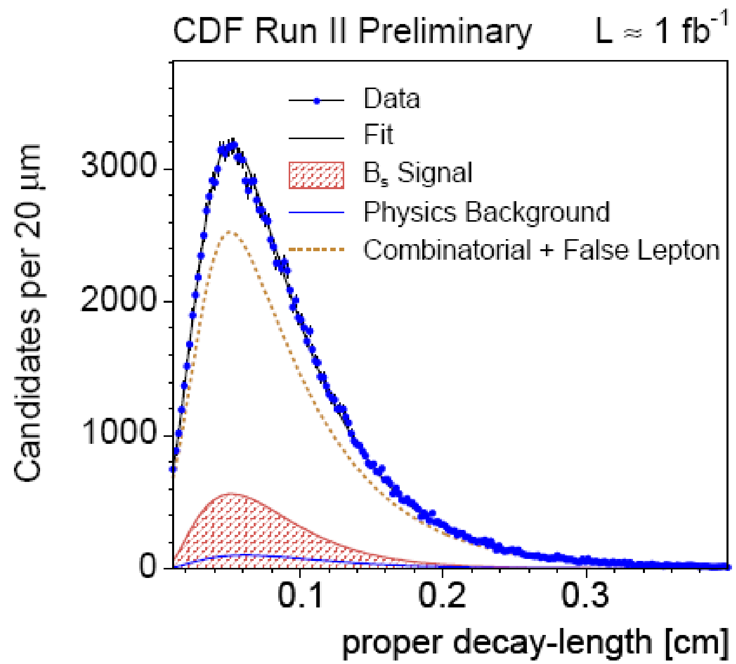
B^+ 1.653 ± 0.014 ps^{-1}

B_s 1.469 ± 0.059 ps^{-1}

Excellent agreement!

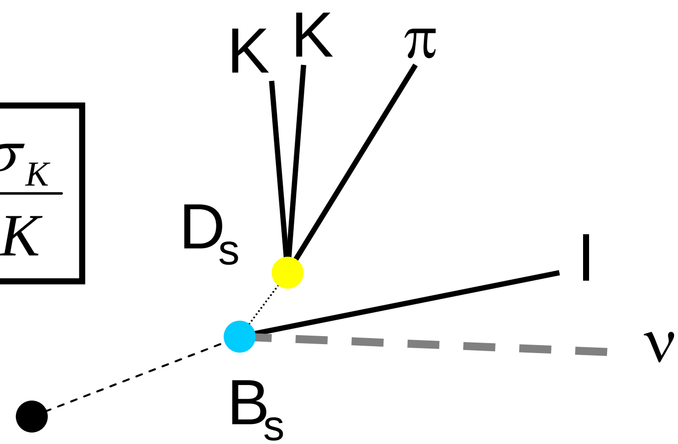
lD_s Lifetime Results

	Lifetime (ps)
Bs:Ds → $\phi\pi$	1.51 ± 0.04 stat. only
Bs:Ds → K*K	1.38 ± 0.07 stat. only
Bs:Ds → $\pi\pi\pi$	1.40 ± 0.09 stat. only
Bs combined	1.48 ± 0.03 stat. only



- lifetimes measured on first 355 pb^{-1}
- compare to World Average: Bs: (1.469 ± 0.059) ps

$$ct = \frac{L_{xy}\, m_B}{P_t^{VIS}} \left\langle \frac{P_t^{VIS}}{P_T} \right\rangle_{MC} \rightarrow \frac{\sigma_{ct}}{ct} = \frac{\sigma_{L_{xy}}}{L_{xy}} \oplus \frac{\sigma_{P_t}}{P_t} \otimes \frac{\sigma_K}{K}$$



B_s Mixing Ingredients

Flavor tagging

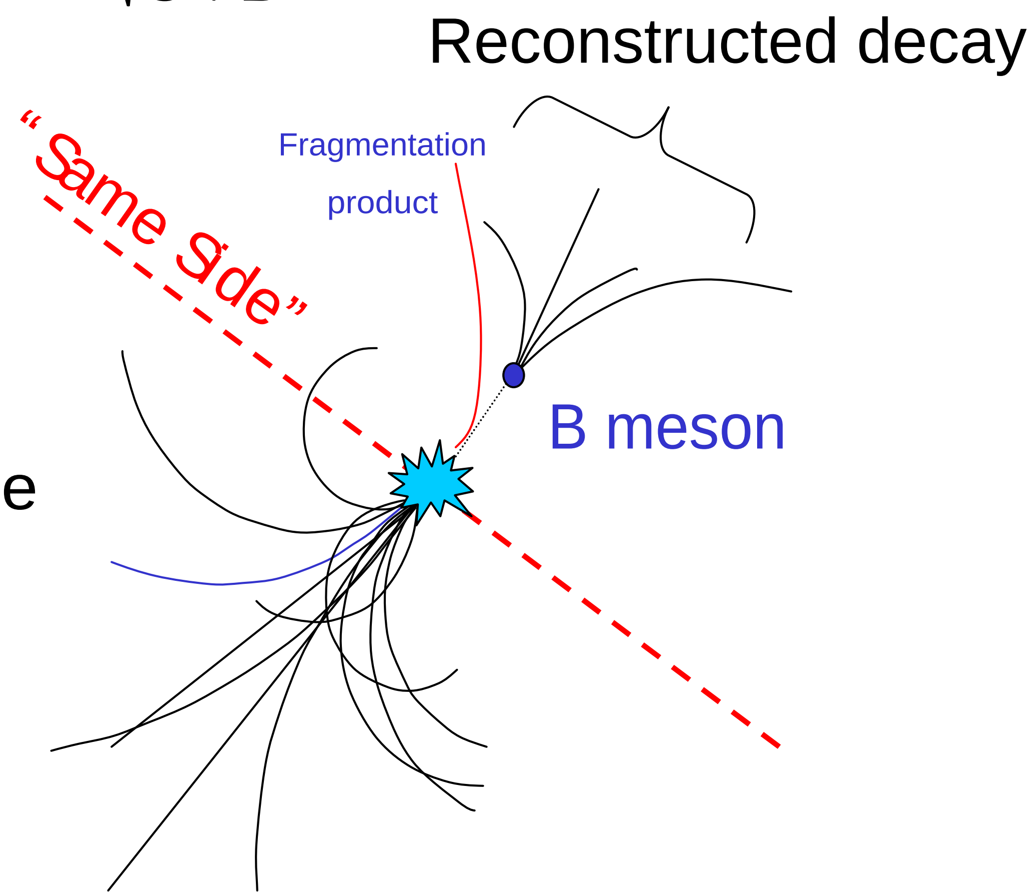
$$Significance = \sqrt{\frac{S\varepsilon D^2}{2}} e^{-\frac{(\Delta m_s \sigma_t)^2}{2}} \sqrt{\frac{S}{S+B}}$$

Flavor Tagging

$$Significance = \sqrt{\frac{S\varepsilon D^2}{2}}\, e^{-\frac{(\Delta m_s \sigma_t)^2}{2}} \sqrt{\frac{S}{S+B}}$$

$$D = \frac{N_{right} - N_{wrong}}{N_{right} + N_{wrong}}$$

Amplitude → D×Amplitude



Several methods, none is perfect !!!

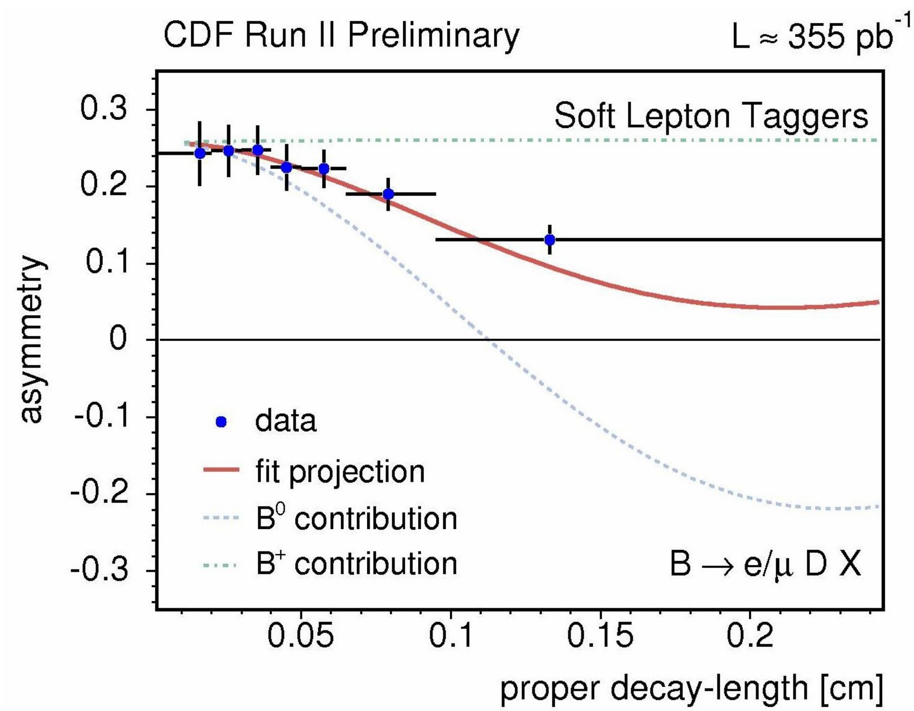
Unbinned Likelihood Δm_d Fits

B_d/B^+ samples used as guinea pigs:

- Validate fit implementation
- Characterize taggers

1. Semileptonic and hadronic samples are fit separately
2. A is fixed to 1
3. ε,D, Δm_d are measured!

semileptonic, lD^-, muon tag



hadronic: $\Delta m_d = 0.536 \pm 0.028$ (stat) ± 0.006 (syst) ps^{-1}

semileptonic: $\Delta m_d = 0.509 \pm 0.010$ (stat) ± 0.016 (syst) ps^{-1}

world average: $\Delta m_d = 0.507 \pm 0.004$ ps^{-1}

B_s Mixing: tagging performance

Tagger "calibration":

1. Tune tagger (selection cuts, algorithm details)
2. Measure performance (ε, D) on control samples

• Measured from B_d data (CDF)

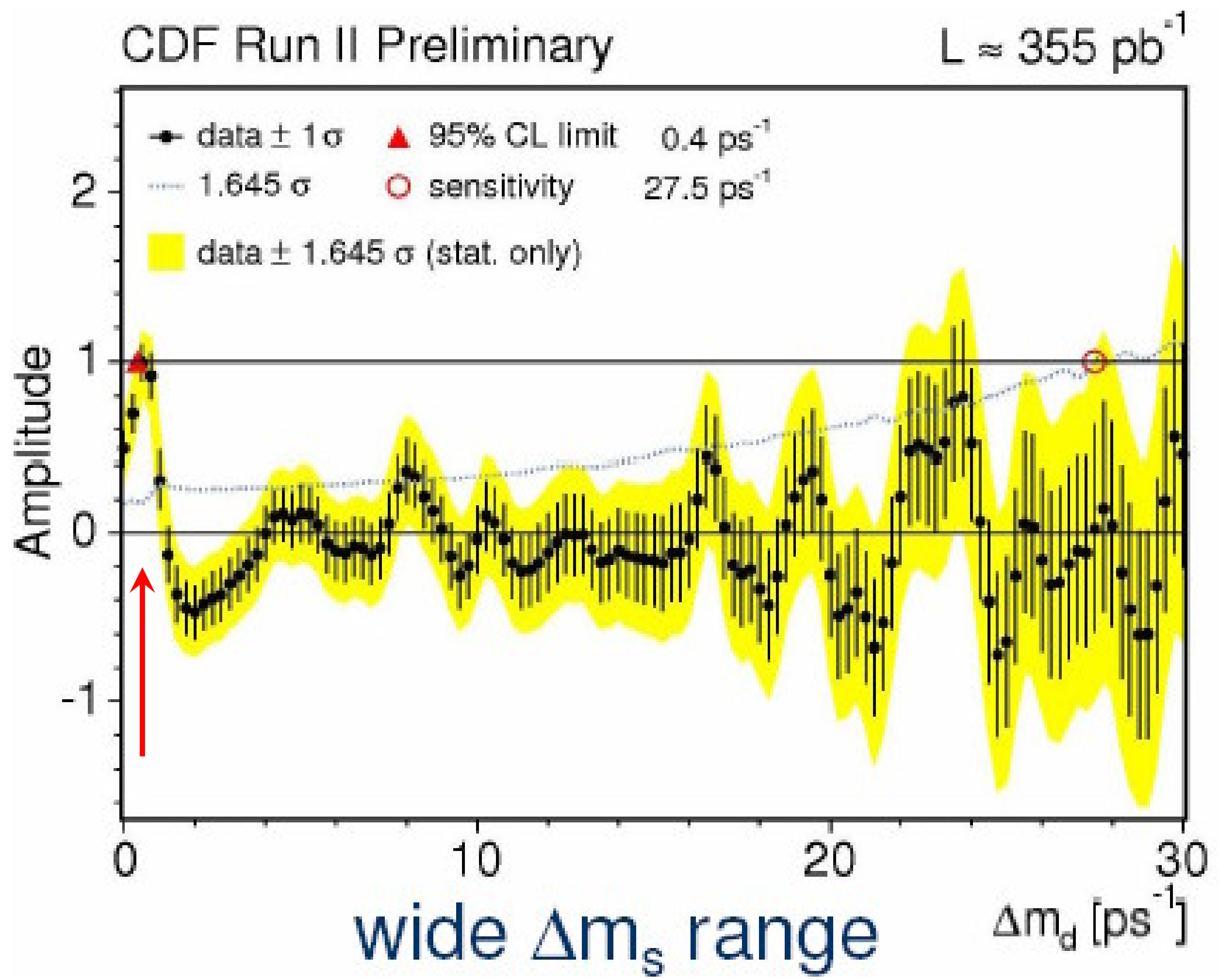
	εD^2 Hadronic (%)	εD^2 Semileptonic (%)
Muon	0.48 ± 0.06 (stat)	0.62 ± 0.03 (stat)
Electron	0.09 ± 0.03 (stat)	0.10 ± 0.01 (stat)
JQ/Vertex	0.30 ± 0.04 (stat)	0.27 ± 0.02 (stat)
JQ/Prob.	0.46 ± 0.05 (stat)	0.34 ± 0.02 (stat)
JQ/High p_T	0.14 ± 0.03 (stat)	0.11 ± 0.01 (stat)
Total OST	1.47 ± 0.10 (stat)	1.44 ± 0.04 (stat)
SSKT	3.42 ± 0.06 (stat)	4.00 ± 0.04 (stat)

CDF: ~5% of the Events are effectively used!

D0: ~2.5% of the events are effectively used!

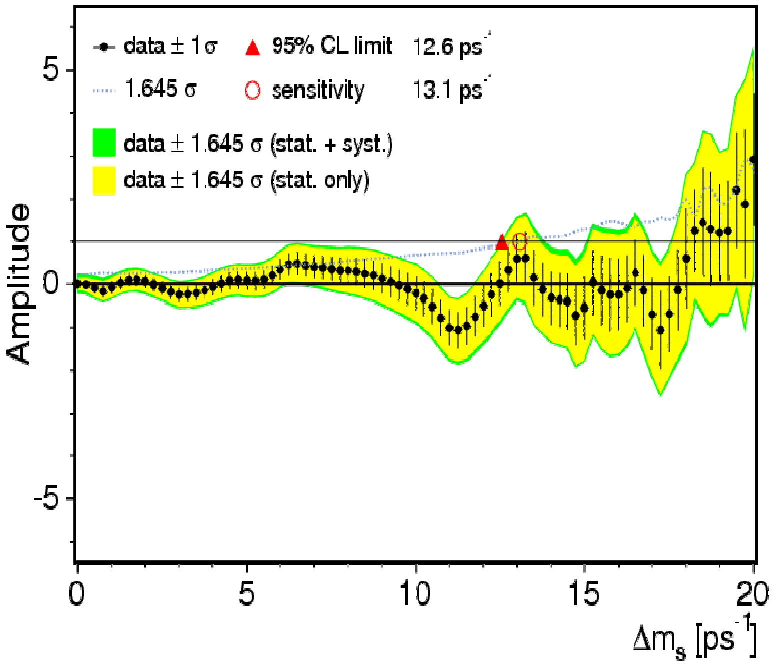
Amplitude Scan: signal?

- B_d mixing can be searched for too
- Signal is clearly visible both by CDF and D0
- Detailed features of the scan when signal is present can vary from one experiment to the other
- What happens when you see a signal?
    - See a peak
    - Details of the peak depend on the experiments properties
    - How do you define the significance of a signal?



Remember: this all becomes an academic exercise when statistics is large enough!

Amplitude Scan do and don't

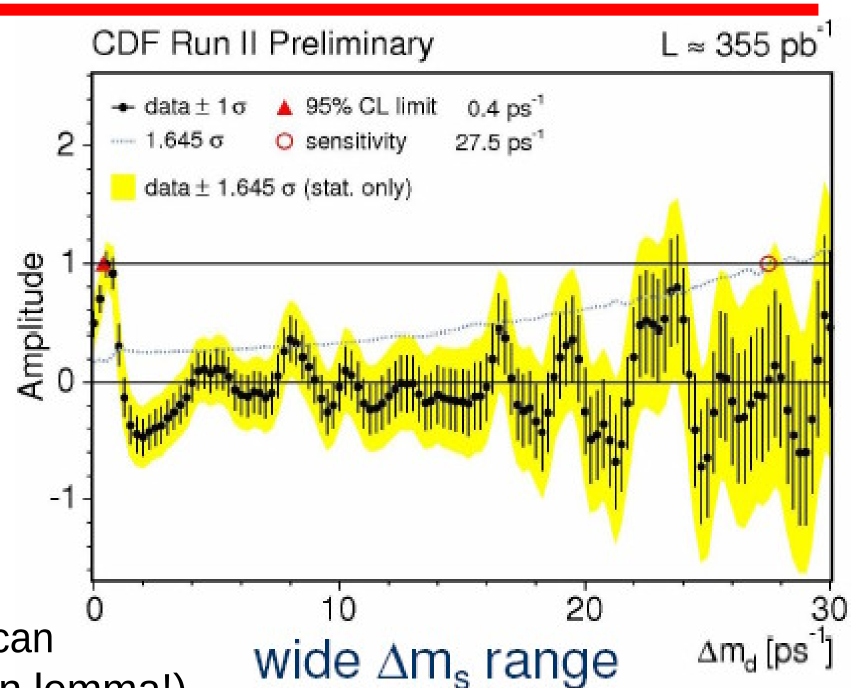


- Amplitude scan is helpful to:
  - Set a Δm limit
  - Combine experimental results
- It is not easy to measure mixing from it
- How does an evidence of a signal look like?
- What procedure should one follow if aiming at a measurement?
- These questions must be asked before performing the analysis!
- Otherwise lack of coverage is the punishment!

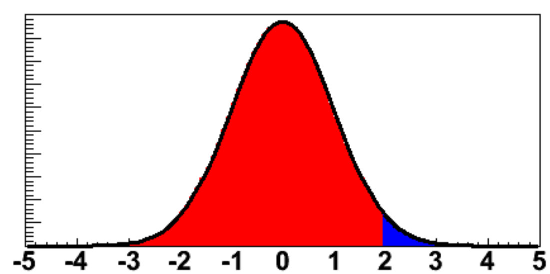
Remember:

- Not to confuse the individual significance of each A measurement with the overall significance of the 'feature'
- 'Discovery threshold' is an arbitrary cut on the probability for non-signal to produce the same features: nothing to do in general with how significant the value of a given parameter you measure is!

Neyman-Pearson



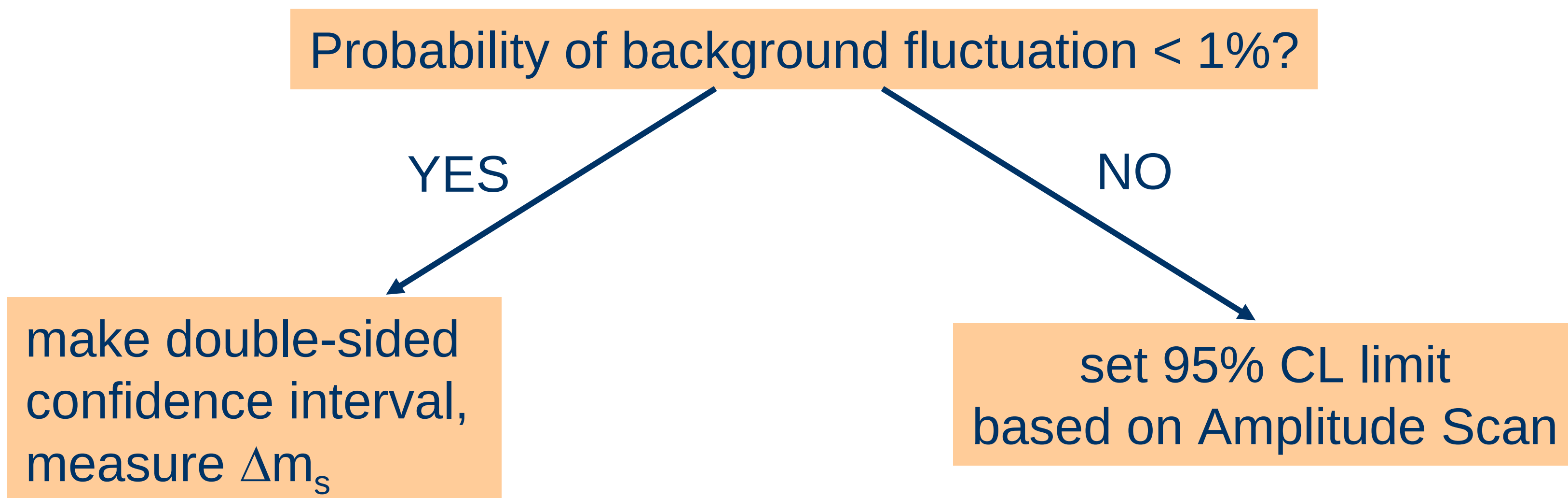
- Several ways of using your data
  - set a lower limit? Set an upper limit?
  - Obtain a two-sided bound?
  - Measure Δms?
- We want to discern between
  - H_0 = no signal
  - H_1 = mixing at a certain Δm value
- Neyman-Pearson test:
  - Pick an observable ξ, e.g.:
    - Significance of the highest peak in A-scan
    - Likelihood ratio (UMP! Neyman-Pearson lemma!)
  - Derive: $P(\xi|H_0)$ $P(\xi|H_1)$
  - Define:
    - Bands in ξ for rejecting H_0/H_1
    - $\Rightarrow$ Desired detection & false alarm probabilities
  - Open the box!



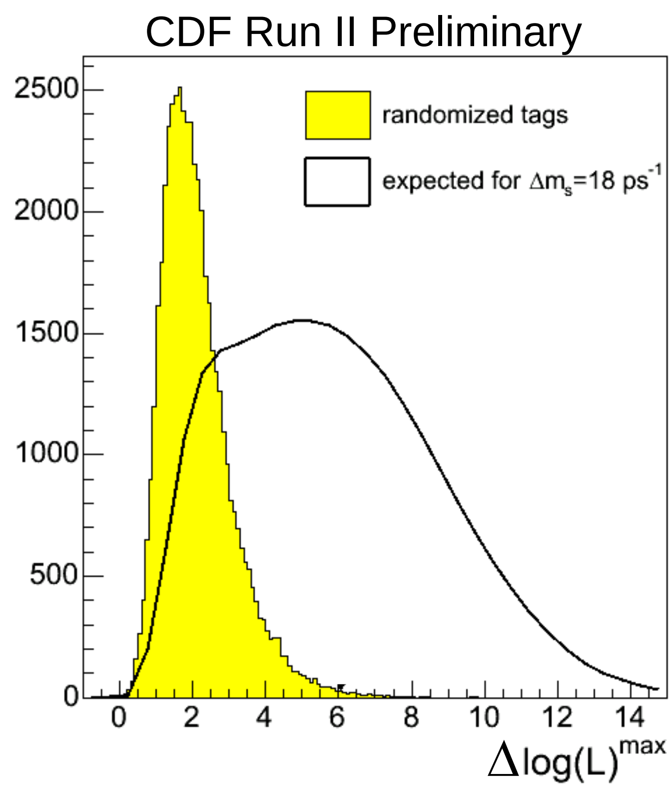
- Dangerous things:
  - Defining procedure (observable, probability thresholds and bands) after looking at your sample
  - Being confused about the procedure
  - Switch from one way of using data to another (limit vs measurement)

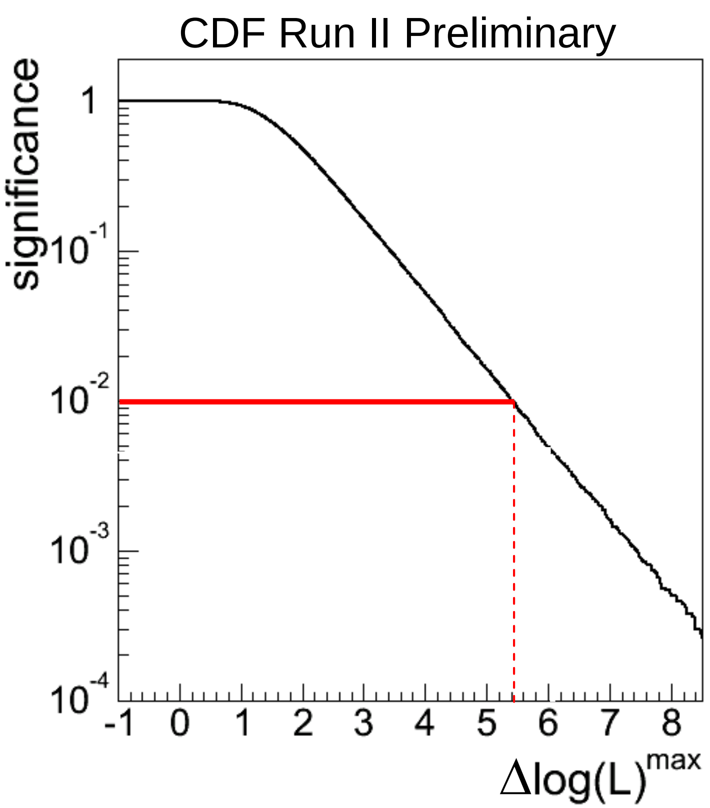
CDFs Choice of Procedure

- Decided upon before un-blinding $1fb^{-1}$ of data
- P-value: probability that observed effect is due to background (false alarm): 1% (should be $\sim 6 \cdot 10^{-7}$ [5σ] for a 'discovery')
- to be estimated using method defined in the next slide
- no search window to be used



Significance

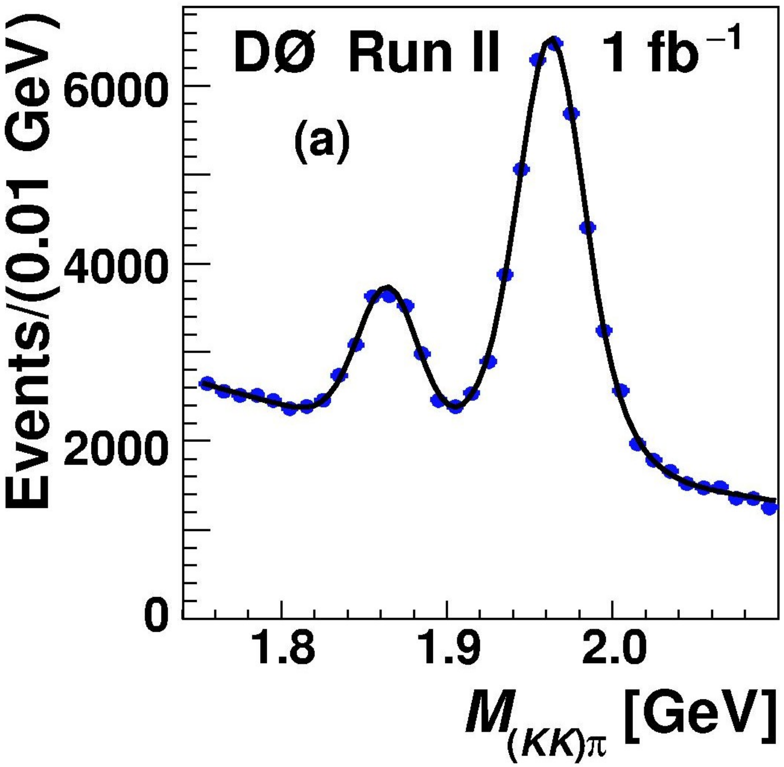


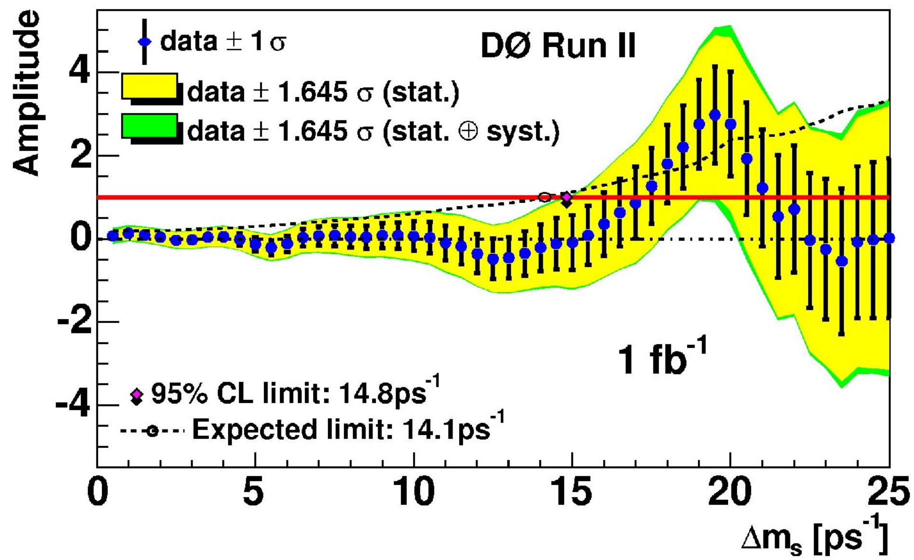


- Δlog(L) = log[L(A=1) / L(A=0)] $\rightarrow$ signal at likelihood's deepest "dip"
- more powerful discriminant than A/σ(A)
- probability of random tag fluctuations evaluated on data (with randomized tags) $\rightarrow$ checked that toy Monte Carlo gives same answer

B_s Mixing: D0 Result

Hep-ex/0603029





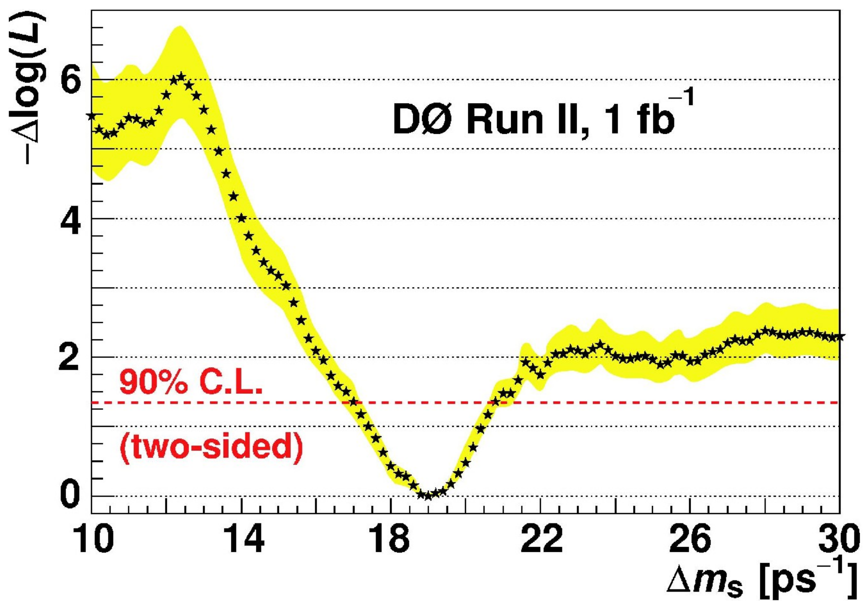
$\Delta m_s > 14.8$ ps^{-1} @ 95% CL

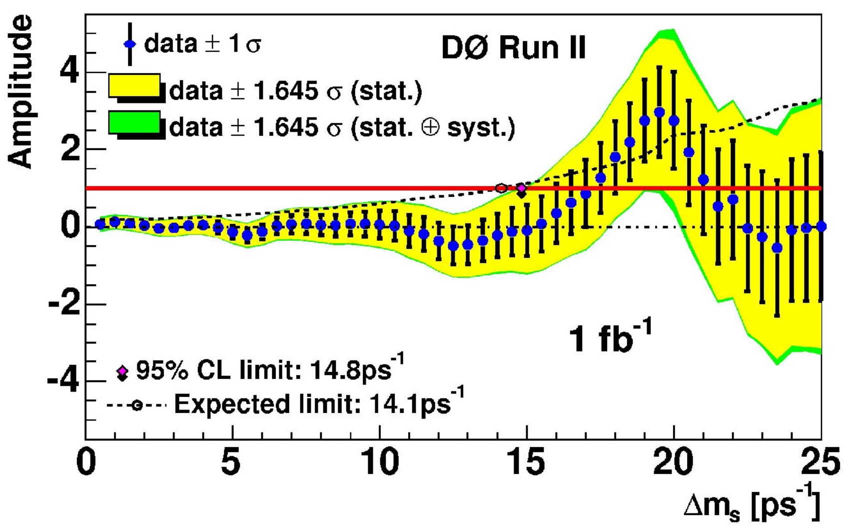
Sensitivity: 14.1 ps^{-1}

- 26700 lD_s candidates
- $\varepsilon D^2 \sim 2.5\%$

B_s Mixing: D0 Result

Very exciting: is this a mixing signal???





Pros	Cons
$\Delta m \approx 19$	$\Delta m \approx 19$
$A/\sigma_A \approx 2.5$	$(A-1)\ /\sigma_A \approx 1.6$
L has a nice dip	..but shallow
P(BCKGND)~5%	P(SIGNAL)~15%

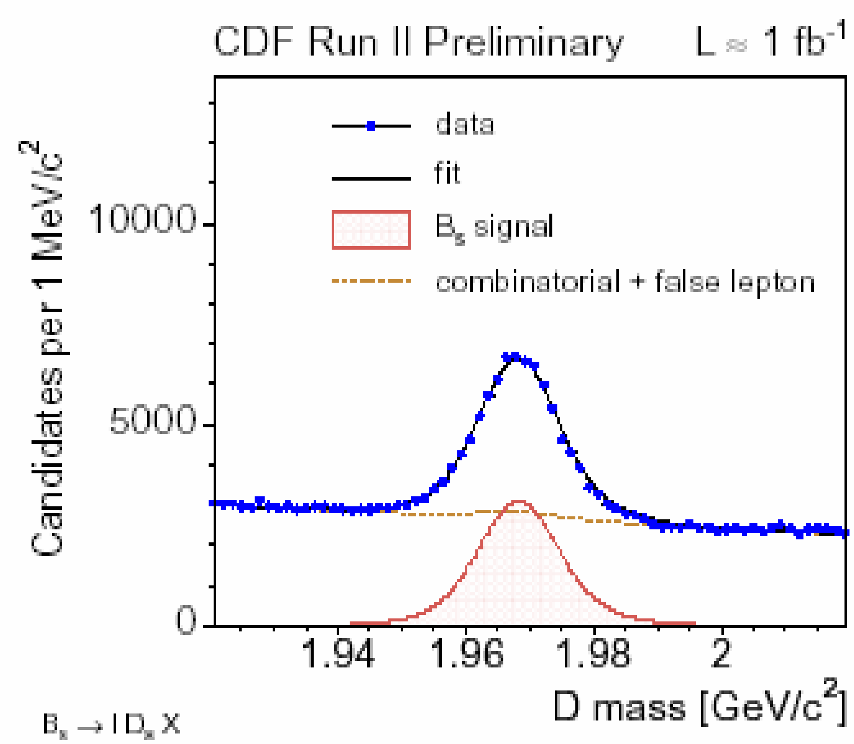
D0 PRL offers a set of possible choices:

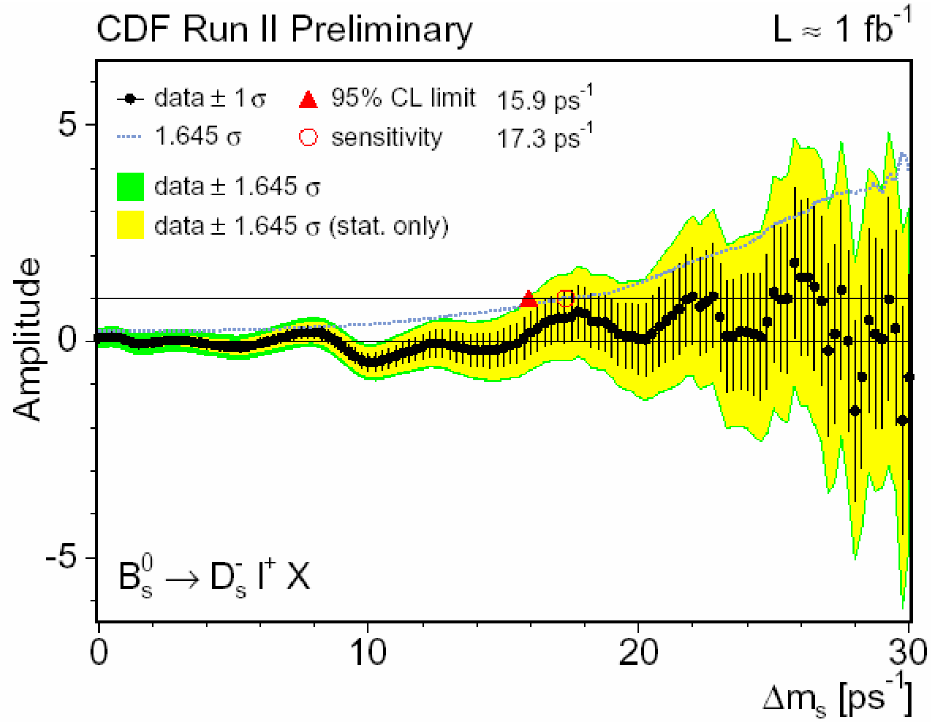
- Setting a limit?
  - upper?
  - Lower?
  - Two-sided?
- Default choice seems to be ‘two sided limit’

B_s Mixing: CDF semileptonic

http://www-cdf.fnal.gov/physics/new/bottom/060406.blessed-Bsmix/

$B_s \to D_s l \nu$	Yield	s/b
$D_s \to \phi\pi$	32300	~2:1
$D_s \to K^*K$	10900	~1:2
$D_s \to \pi\pi\pi$	10100	~1:5



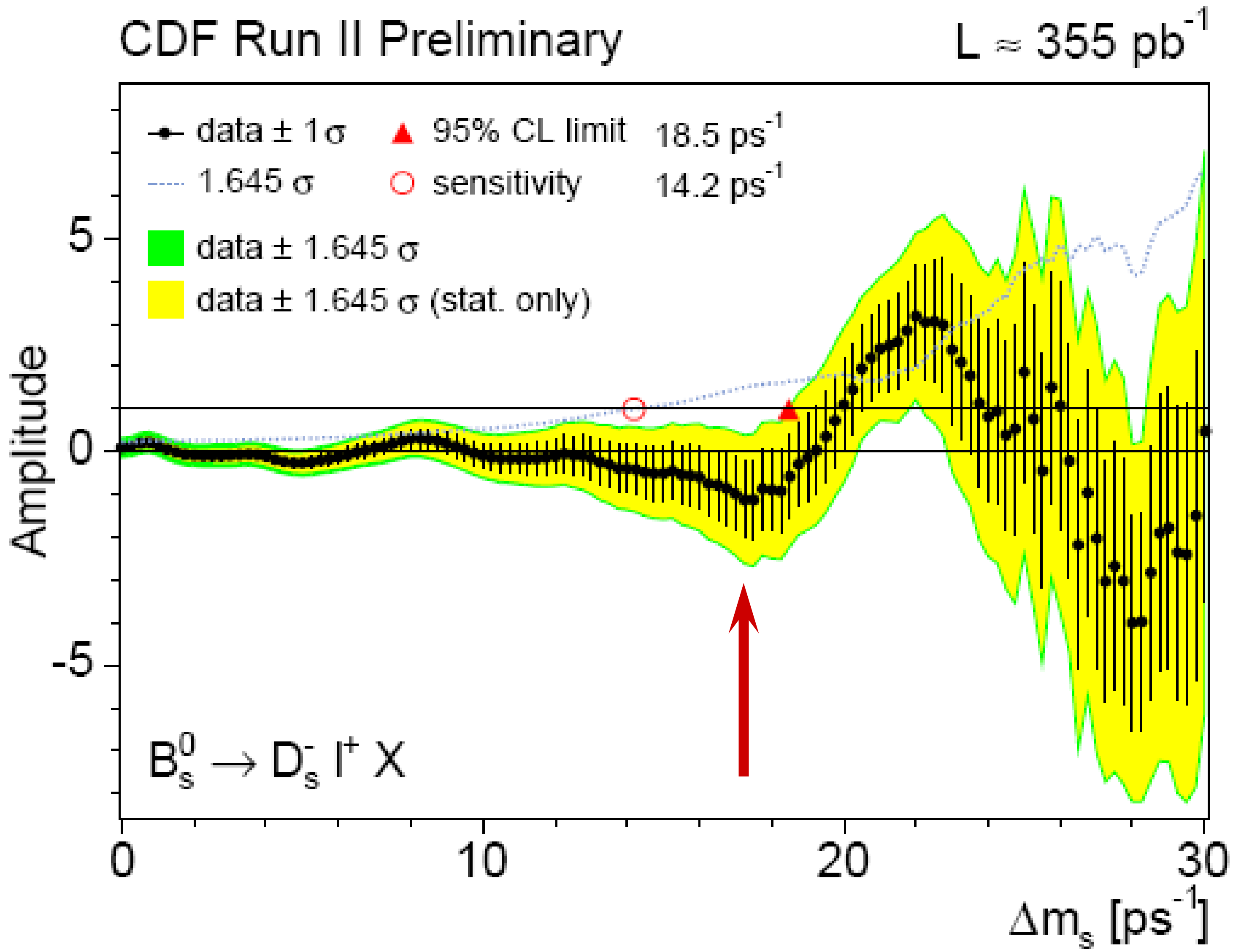


$\Delta m_s > 15.9\ \mathrm{ps}^{-1}$ @ 95% CL

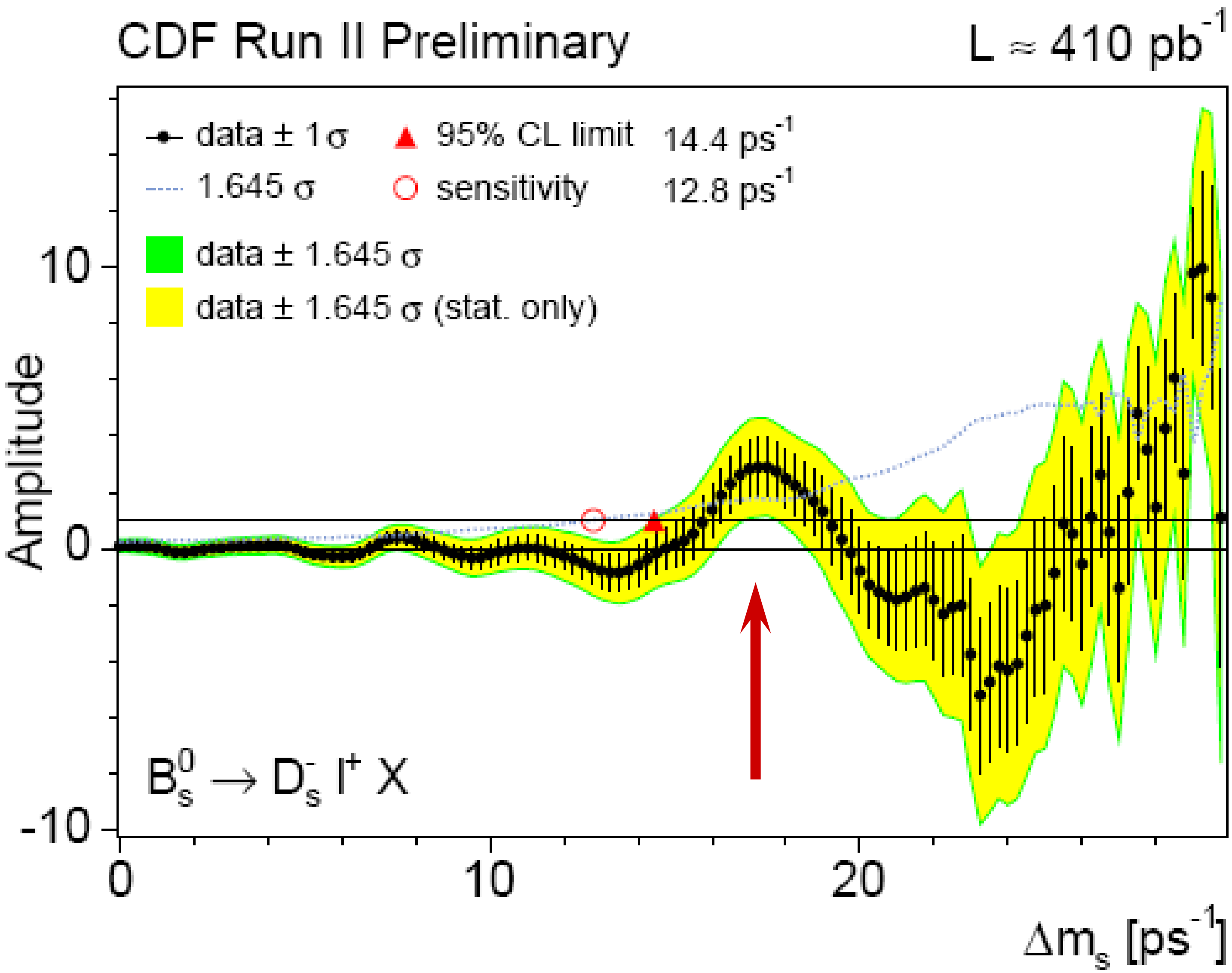
Sensitivity: $17.3\ \mathrm{ps}^{-1}$

Reach at large Δm_s limited by incomplete reconstruction (σ_{ct})!

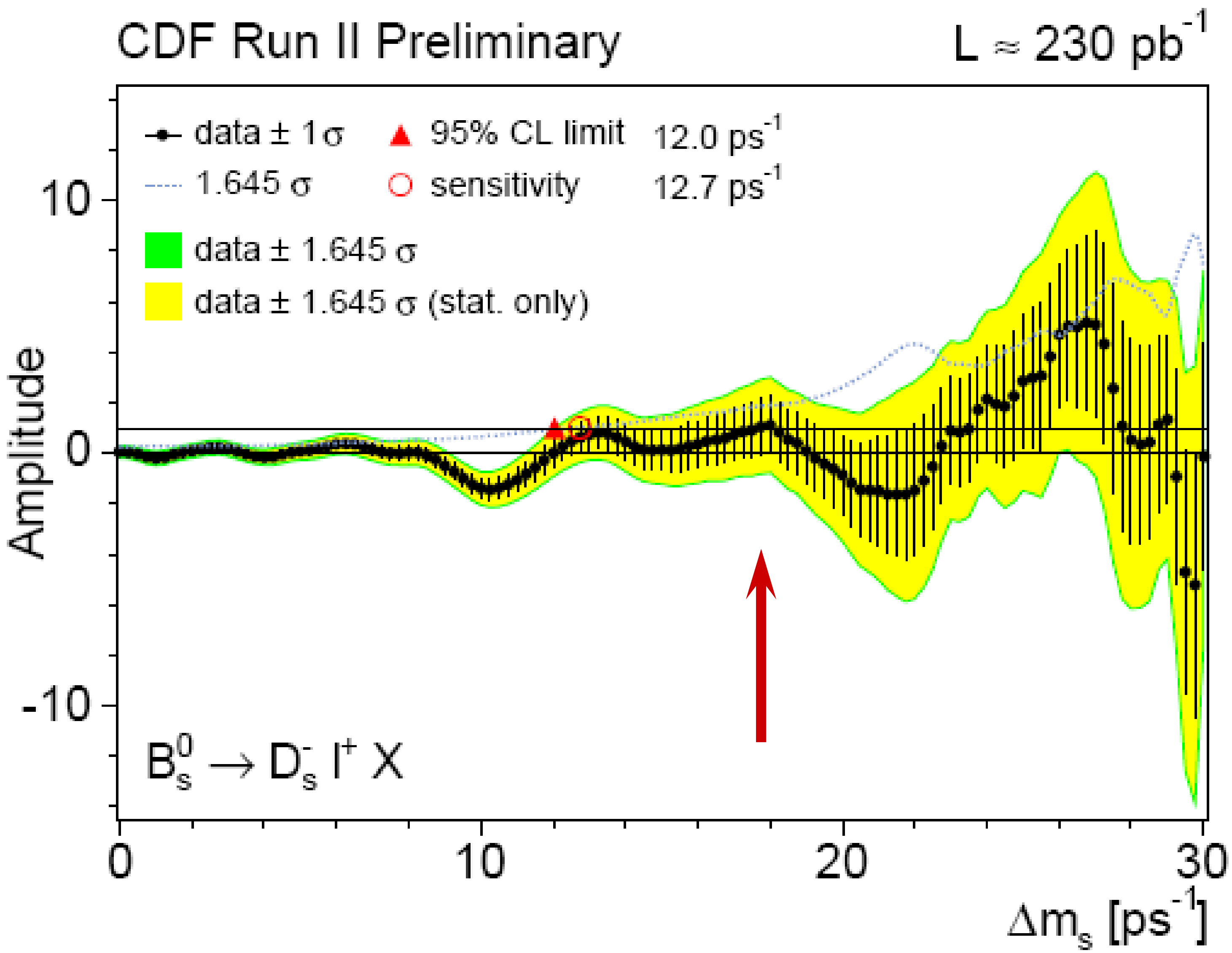
CDF Semileptonic Scan: Period 1



CDF Semileptonic Scan: Period 2



CDF Semileptonic Scan: Period 3

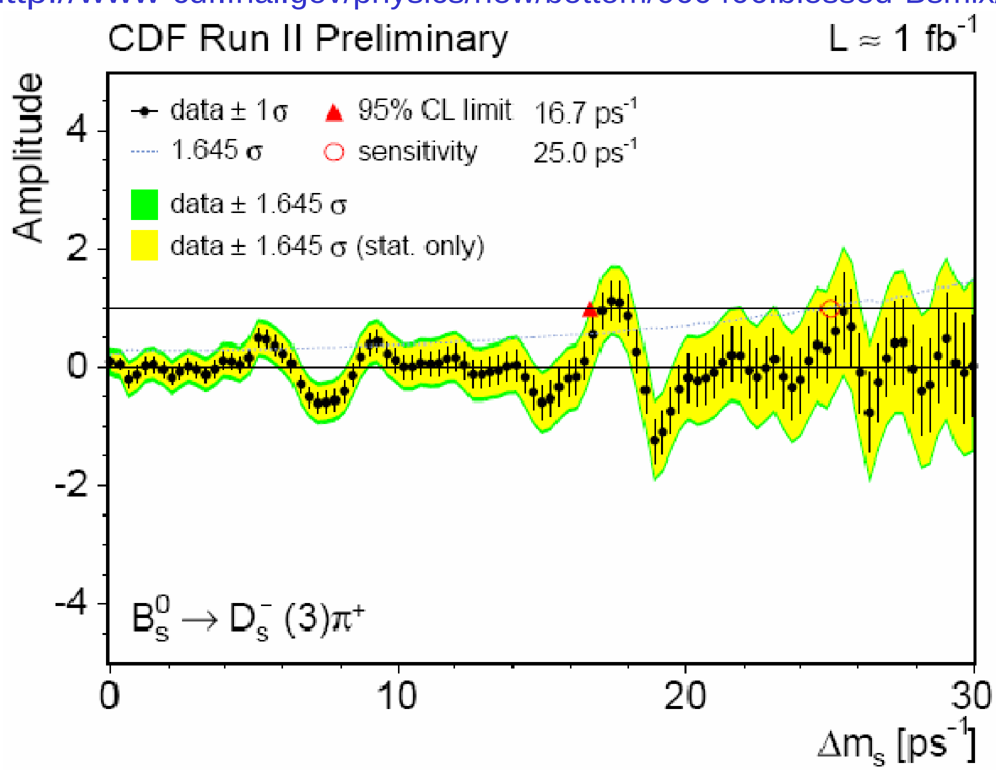


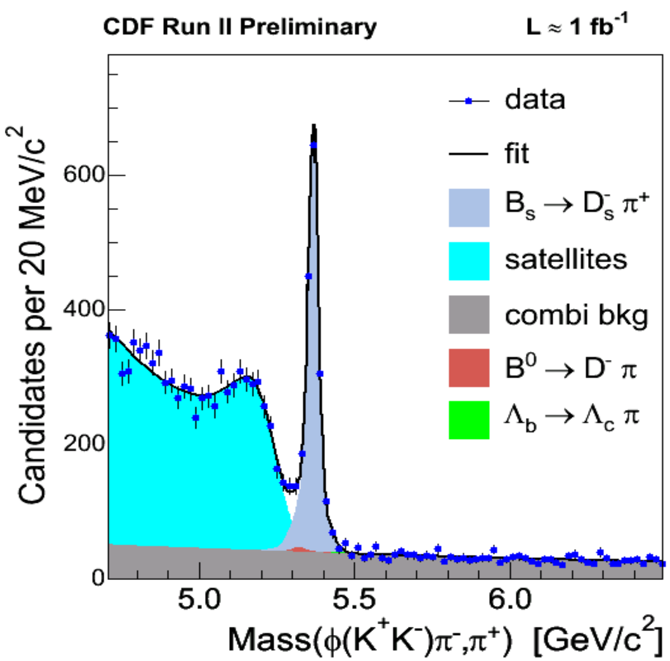
B_s Mixing: CDF hadronic

$B_s \to D_s\pi$	Yield	s/b
$D_s \to \phi\pi$	1600	~4:1
$D_s \to K^*K$	800	~2:1
$D_s \to \pi\pi\pi$	600	~1:1

Using also $B_s \to D_s\pi\pi\pi$ [~1/4 more statistics]

http://www-cdf.fnal.gov/physics/new/bottom/060406.blessed-Bsmix/



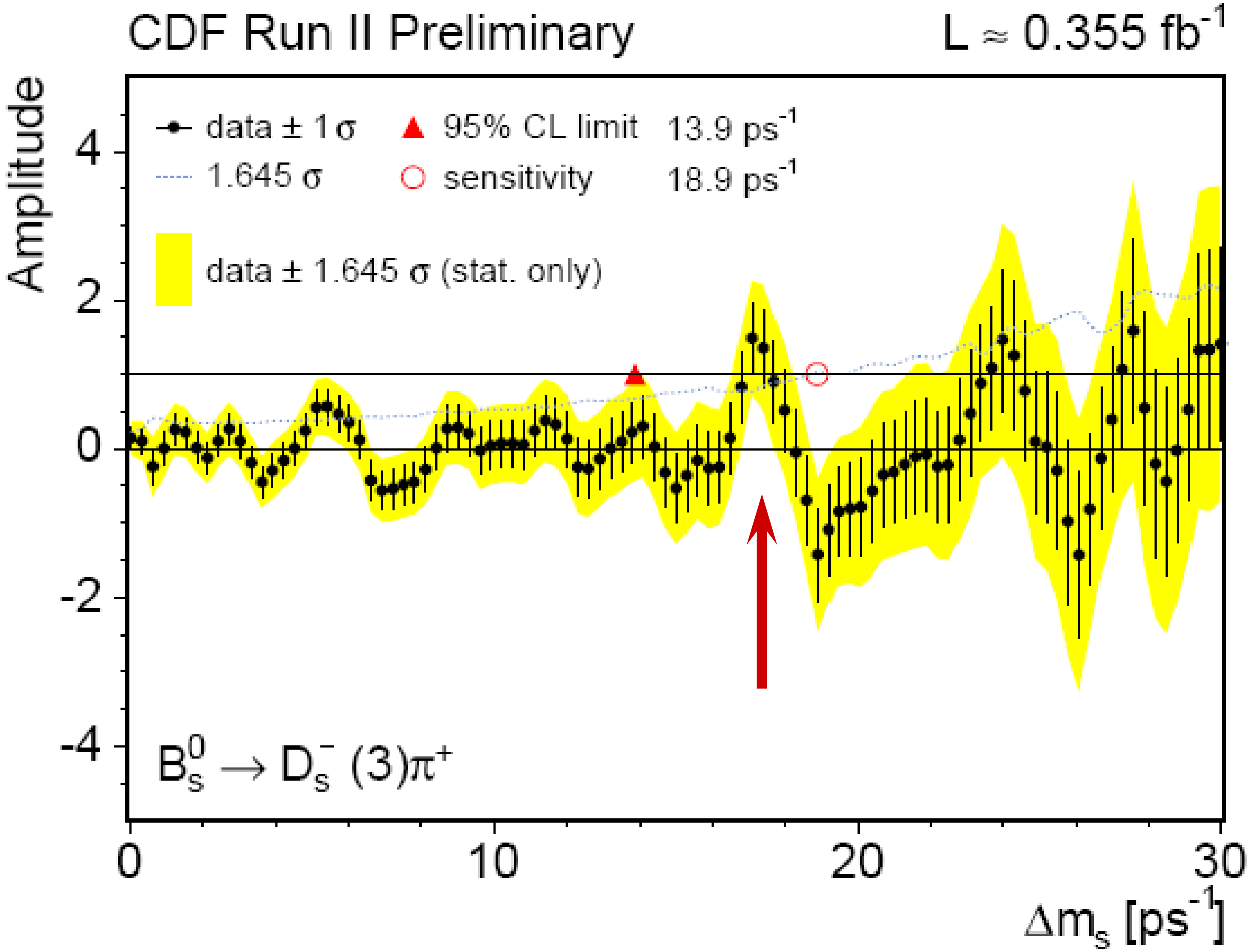


$\Delta m_s > 16.7\ ps^{-1}$ @ 95% CL

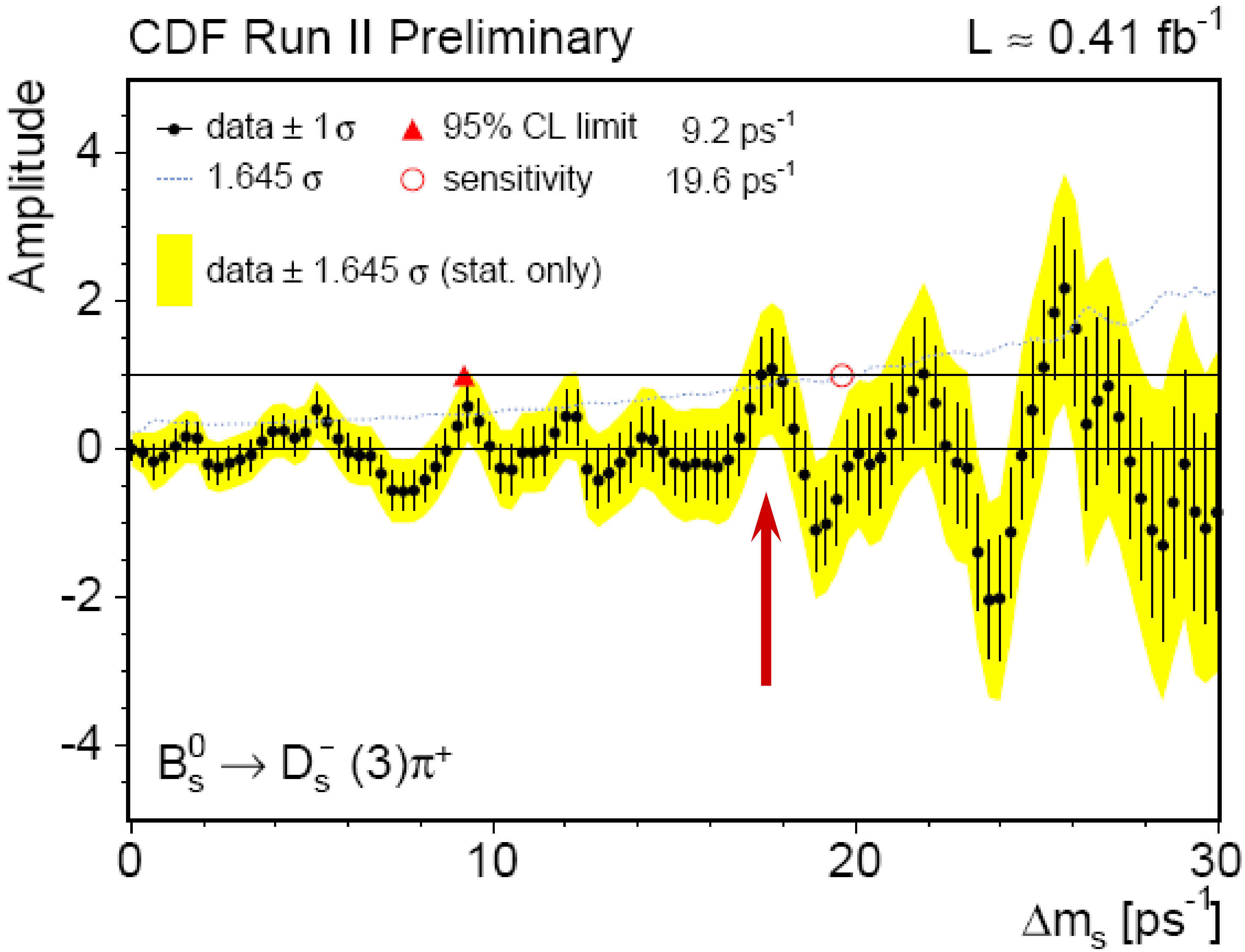
Sensitivity: $25.0\ ps^{-1}$

Is there something?

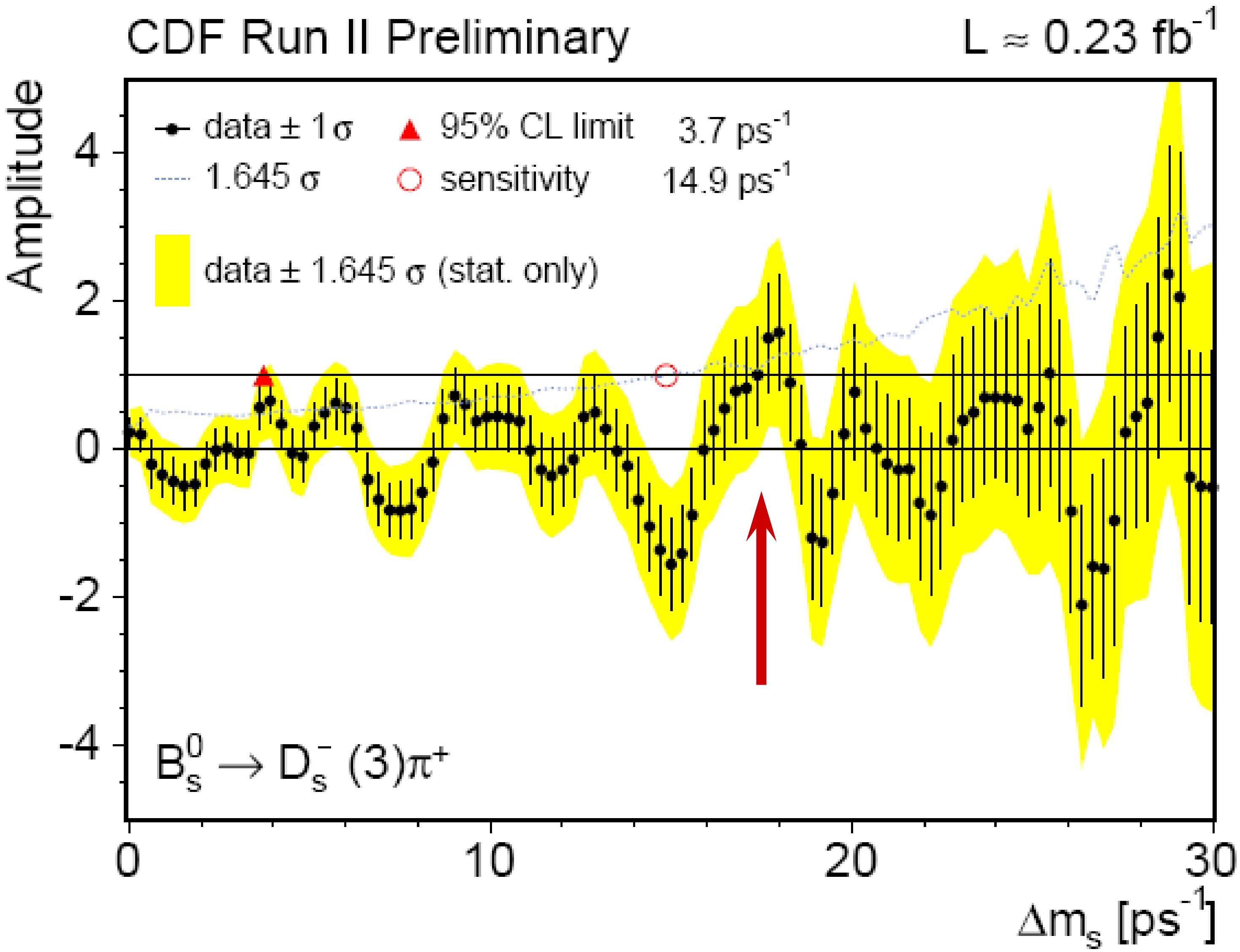
Amplitude Scan: Hadronic Period 1



Amplitude Scan: Hadronic Period 2

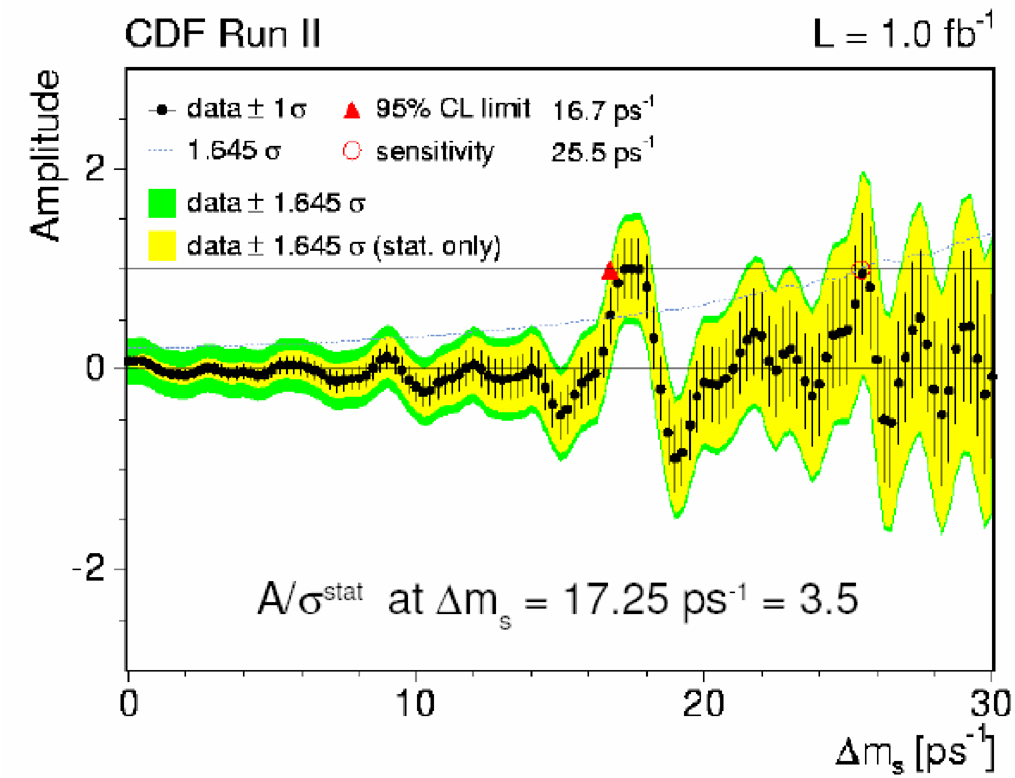


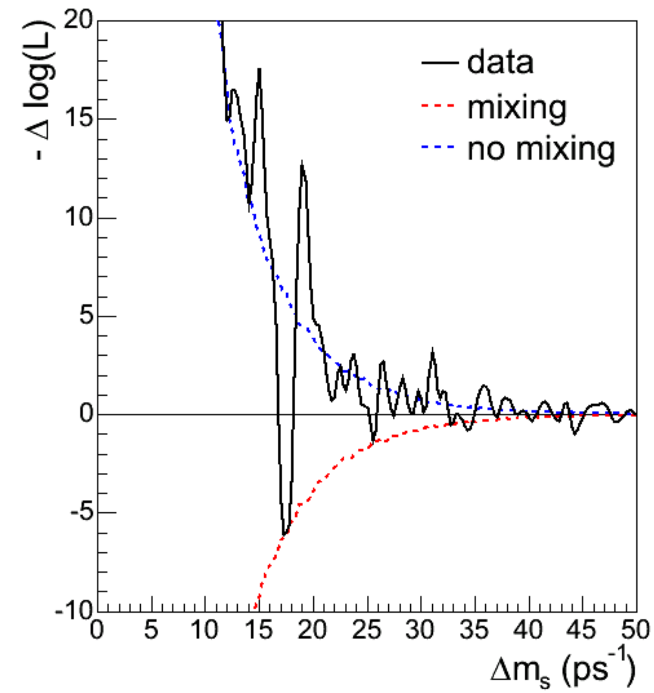
Amplitude Scan: Hadronic Period 3



B_s Mixing: combined CDF result

http://www-cdf.fnal.gov/physics/new/bottom/060406.blessed-Bsmix/





$\Delta m_s > 16.7$ ps^{-1} @ 95% CL

Sensitivity: 25.5 ps^{-1}

$\log[\mathcal{L}(A=1)/\mathcal{L}(A=0)]^{max} = 6.06$

Background has ~0.5% probability to mimic this!

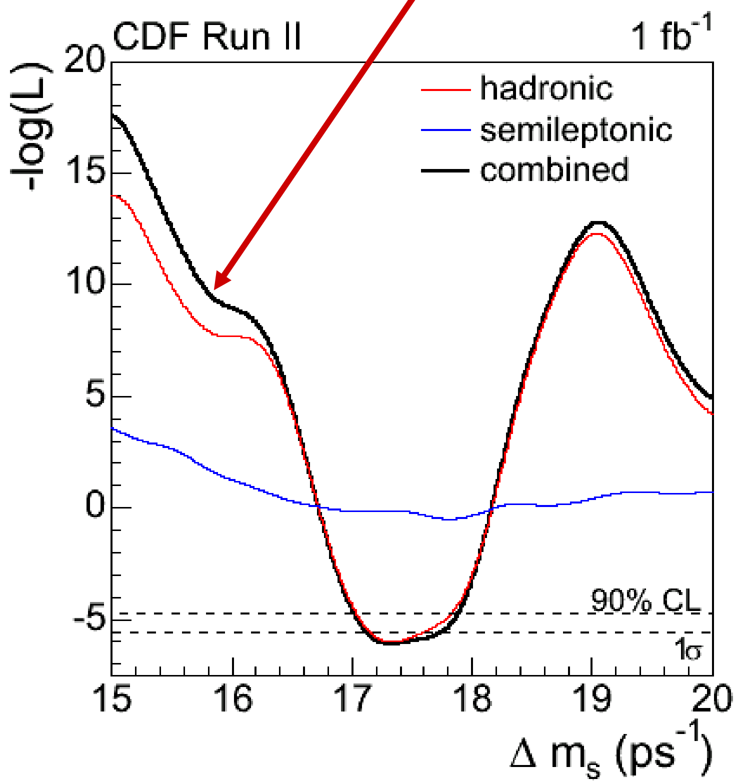
$$\Delta m_s = 17.33^{+0.42}_{-0.21} \pm 0.07(\mathrm{syst})\mathrm{ps}^{-1}$$

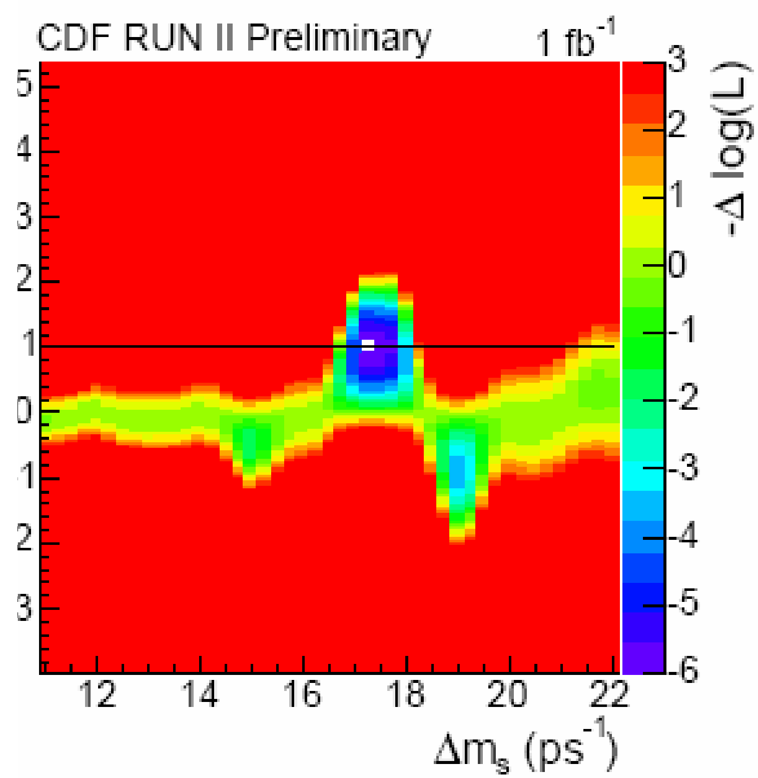
$\Delta m_s \in [17.00, 17.91]$ at 90 % CL

$\Delta m_s \in [16.94, 17.97]$ at 95 % CL

Likelihood Ratio

combined likelihoods from hadronic and semileptonic channels





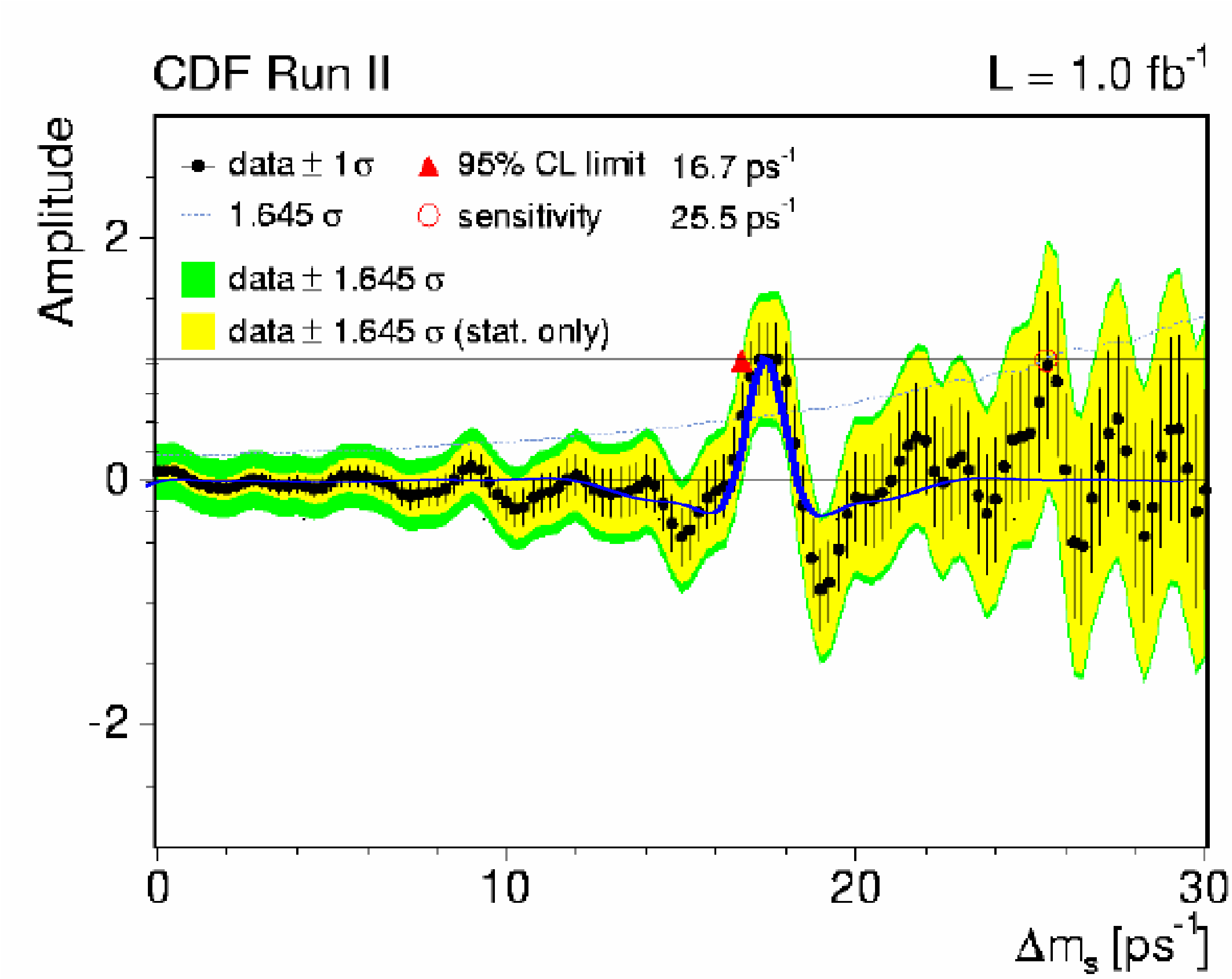
$\Delta m_s = 17.33\ ^{+0.42}_{-0.21}$ (stat) $\pm$ 0.07 (syst) ps^{-1}

Δm_s in [17.00, 17.91] ps^{-1} at 90% CL

Δm_s in [16.94, 17.97] ps^{-1} at 95% CL

the measurement is already very precise! (at 2.5% level)

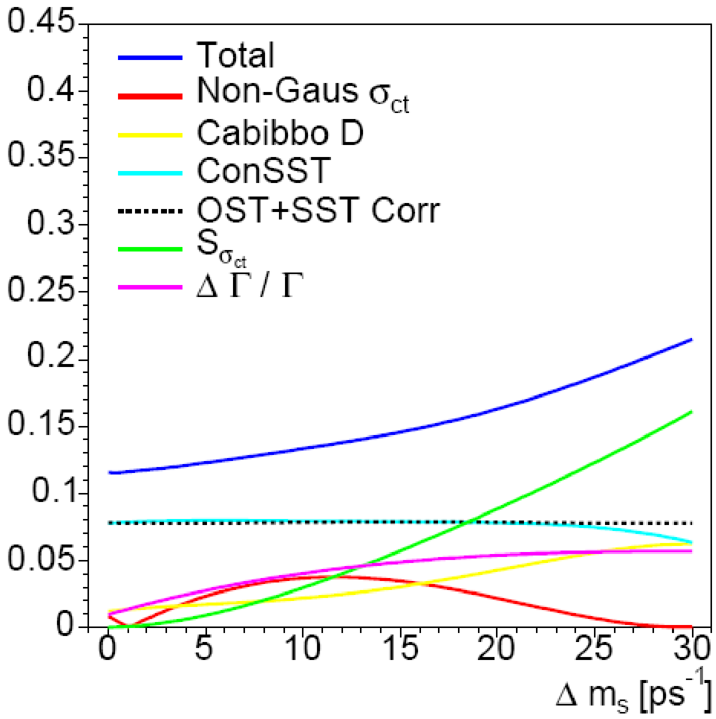
Why the undershoot?



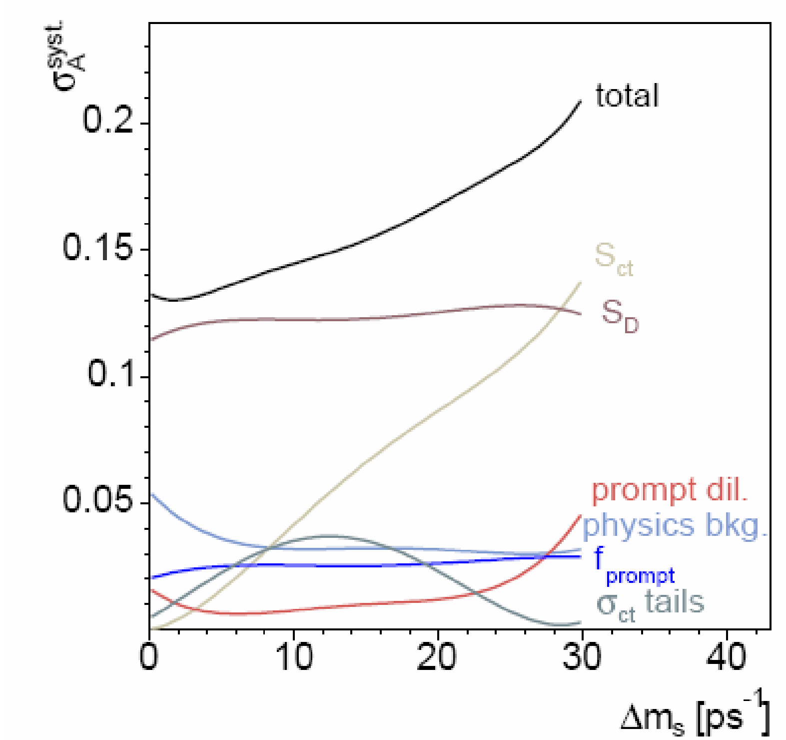
- Peculiarity of our ct-dependent efficiency!
- Does not matter if signal is not present (i.e. the only case where you use an amplitude scan!)
- CDFs amplitude scan can still be combined with the rest of the world for combined limit

Systematic Uncertainties I

Hadronic



Semileptonic



- related to absolute value of amplitude, relevant only when setting limits
  - cancel in A/σ_A, folded in confidence calculation for observation
  - systematic uncertainties are very small compared to statistical

Systematic Uncertainties II: Δm_s

- systematic uncertainties from fit model evaluated on toy Monte Carlo
- have negligible impact
- relevant systematic unc. from lifetime scale

	Syst. Unc
Fitting Model	$< 0.01 \mathrm{ps}^{-1}$
SVX Alignment	$0.04\ \mathrm{ps}^{-1}$
Track Fit Bias	$0.05\ \mathrm{ps}^{-1}$
PV bias from tagging	$0.02\ \mathrm{ps}^{-1}$
Total	$0.07\ \mathrm{ps}^{-1}$

All relevant systematic uncertainties are common between hadronic and semileptonic samples

Δm_s and V_{td}

$$\frac{\Delta m_s}{\Delta m_d} = \frac{m_{Bs}}{m_{Bd}} \xi^2 \frac{|V_{ts}|^2}{|V_{td}|^2}$$

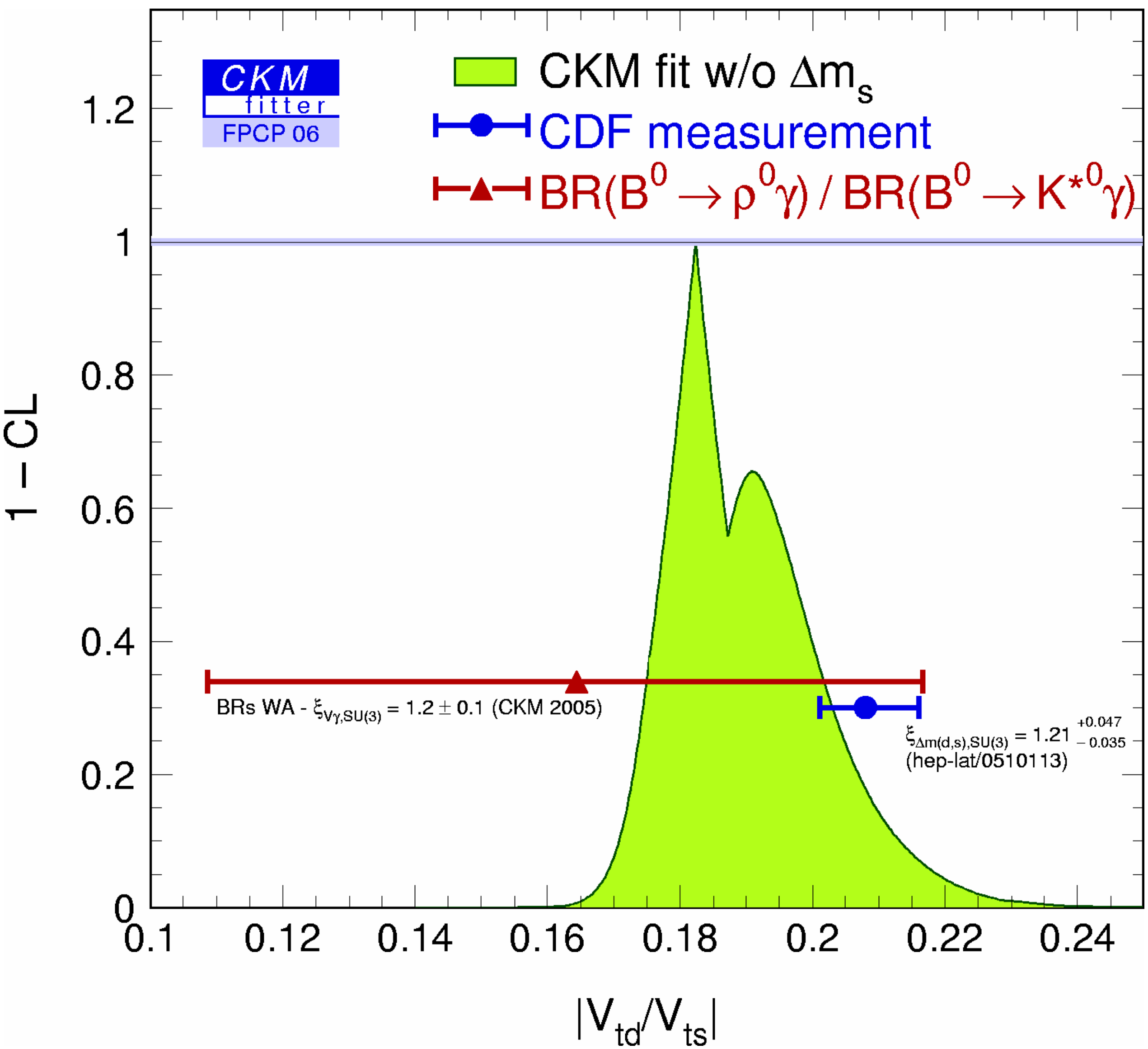
- inputs:
  - → $m(B^0)/m(B_s) = 0.9830$ (PDG 2006)
  - → $\xi = 1.21\,^{+0.47}_{-0.35}$ (M. Okamoto, hep-lat/0510113)
  - → $\Delta\, m_d = 0.507 \pm 0.005$ (PDG 2006)

$|V_{td}| / |V_{ts}| = 0.208\,^{+0.008}_{-0.007}$ (stat + syst)

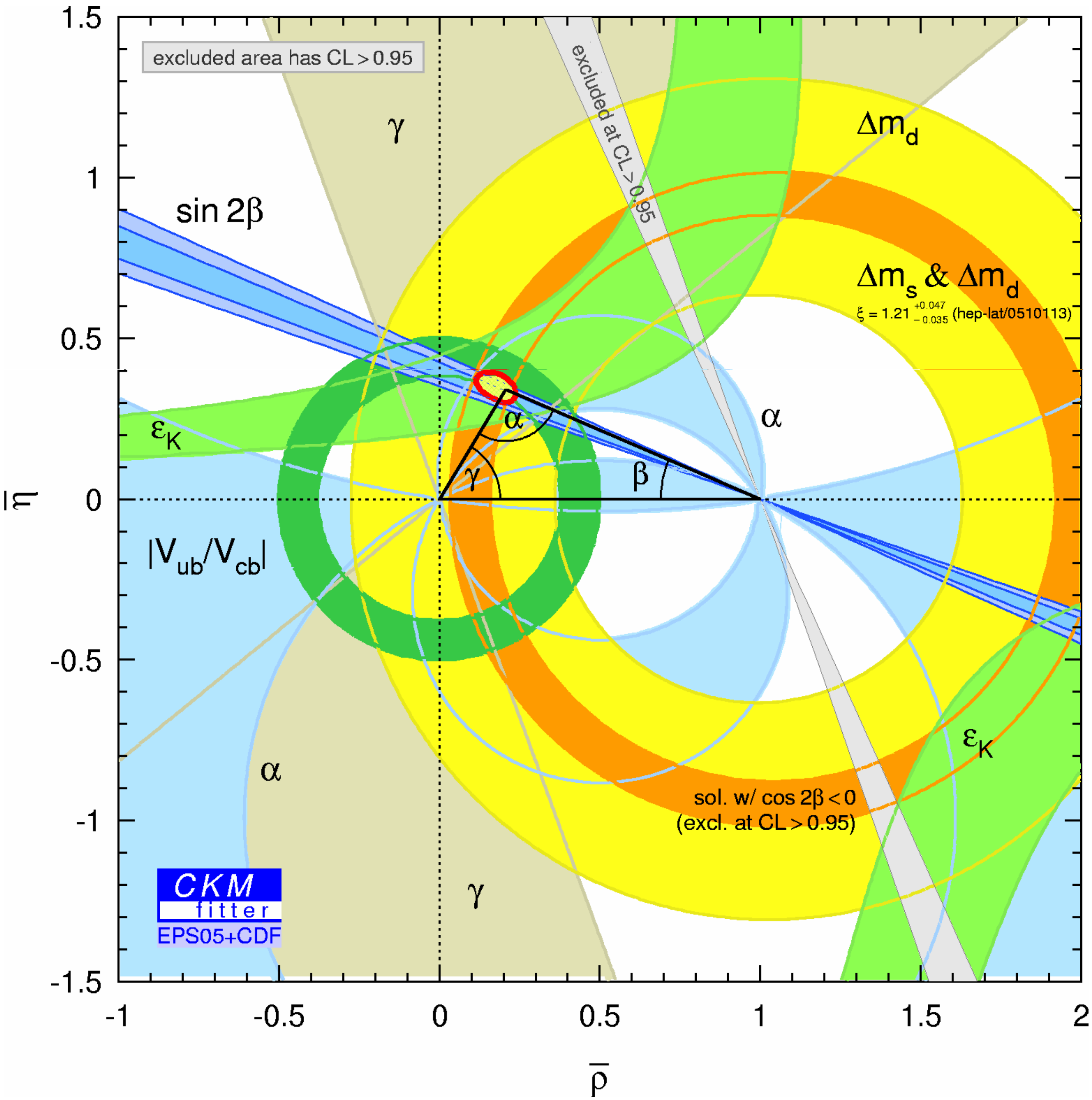
- compare to Belle $b \to s\gamma$ (hep-ex/050679):

$|V_{td}| / |V_{ts}| = 0.199\,^{+0.026}_{-0.025}\,(\text{stat})\,^{+0.018}_{-0.015}\,(\text{syst})$

Δm_s and V_{td}



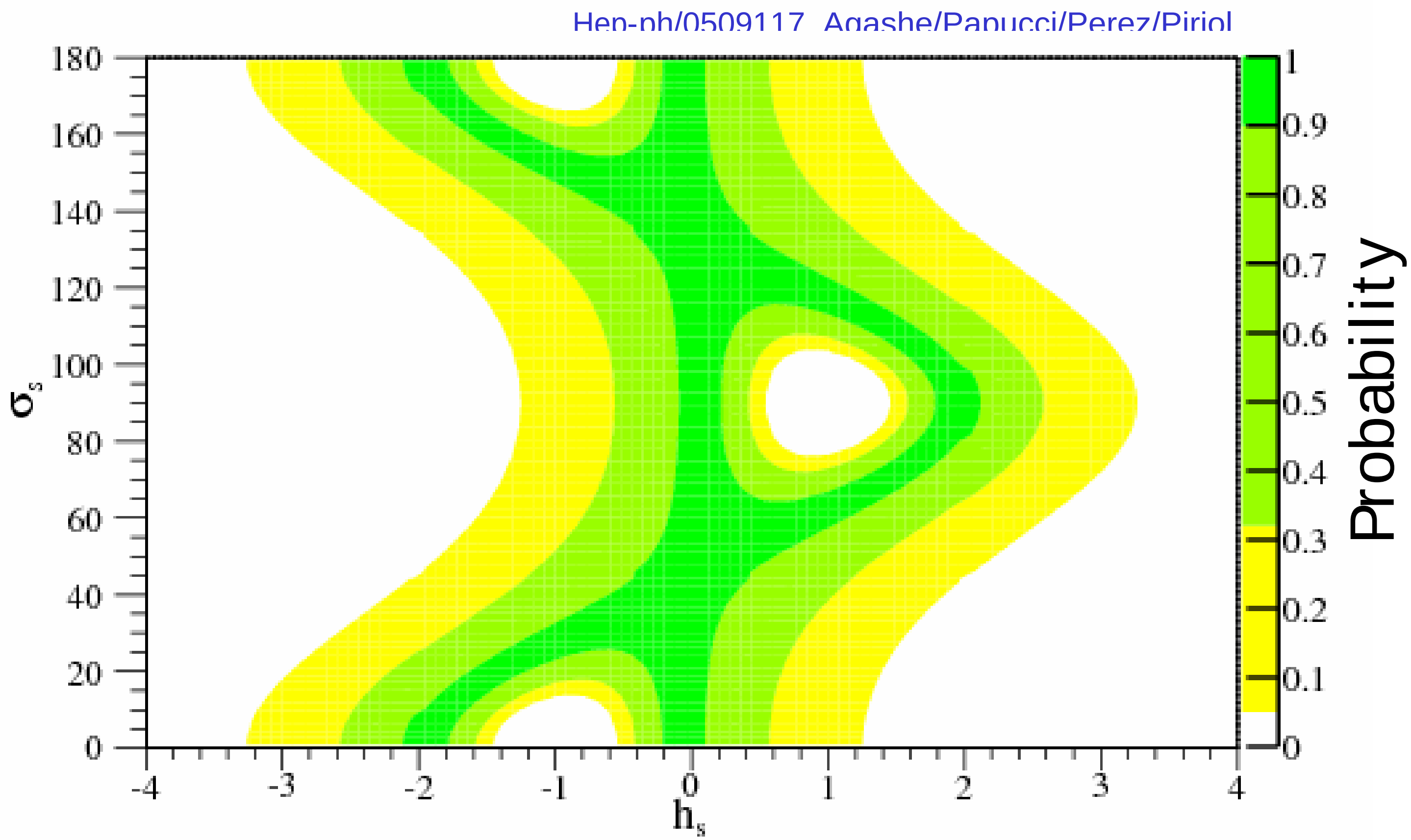
Δm_s & CKM



excluded area has CL > 0.95
γ
excluded at CL > 0.95
Δm_d
sin 2β
Δm_s & Δm_d
ξ = 1.21 +0.047 −0.035 (hep-lat/0510113)
α
ε_K
α
γ
β
|V_ub/V_cb|
α
sol. w/ cos 2β < 0
(excl. at CL > 0.95)
ε_K
γ
CKM fitter
EPS05+CDF
1.5
1
0.5
0
-0.5
-1
-1.5
-1
-0.5
0
0.5
1
1.5
2
η̄
ρ̄

Δms from Tevatron & BSM Limits

$$A_{SM} \rightarrow A_{SM}\left(1 + h_s e^{i\sigma_S}\right)$$



What's next?

- TeVatron samples will be frozen until summer – at the least
- Experiments will refine their analyses:
  - D0, [my guesses on] possible improvements:
    - More D_s modes
    - Include fully reconstructed hadronic decays
    - Improve taggers
  - CDF:
    - Improve tagger usage (we have been very draconian this round on what to/not to use)
    - Additional 'almost fully reconstructed' modes

B_s Mixing: Perspectives

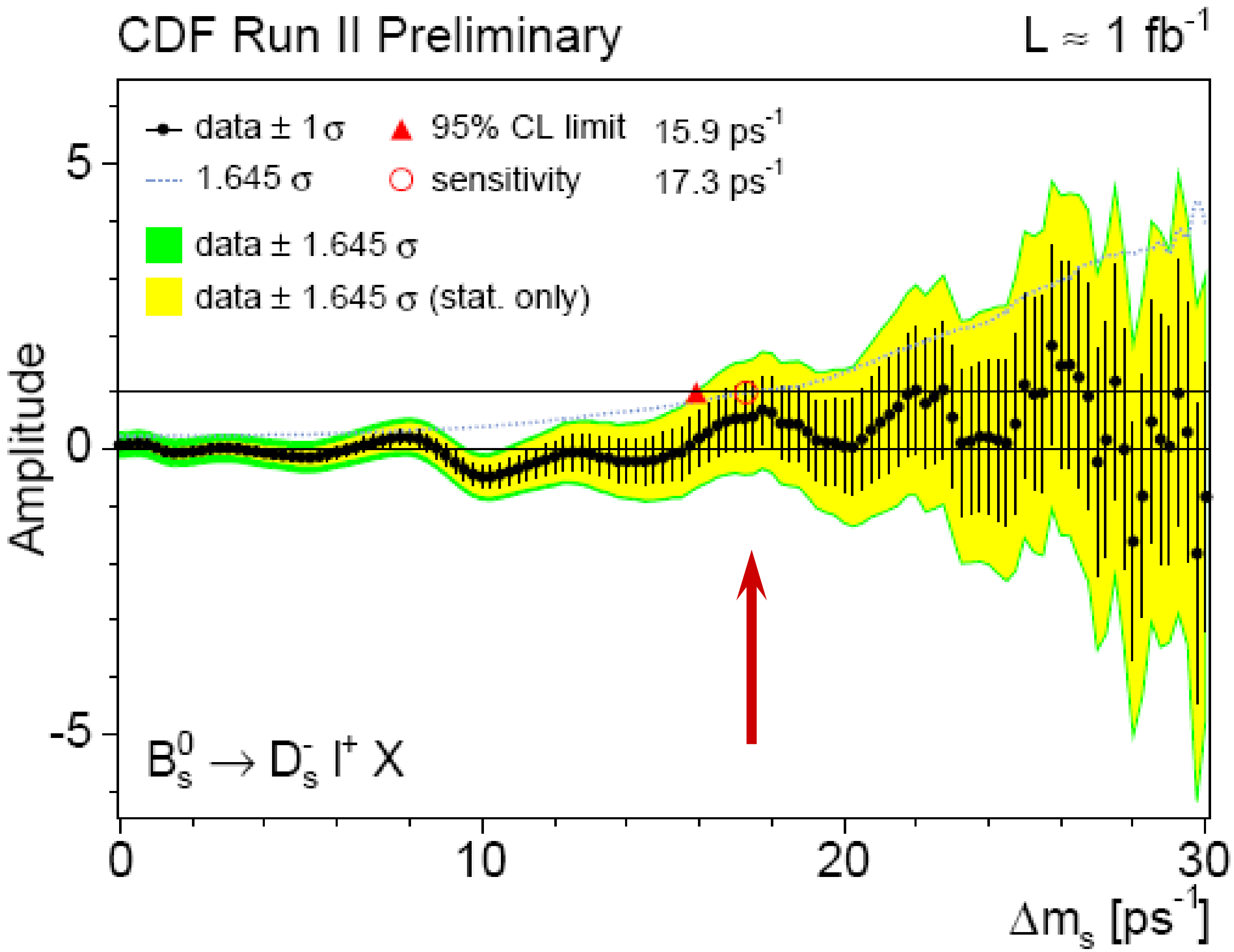
Exciting times ahead:

- 'Discovery' could be close
- B_s result has become an important complementary addition to the CKM mapping!
- ...soon we will improve our mixing sensitivity and move on to new frontiers:

$$B_s \to \psi\phi,\ B_s \to D_s K\ldots$$

Backup Slides

CDF Semileptonic Scan: Combined

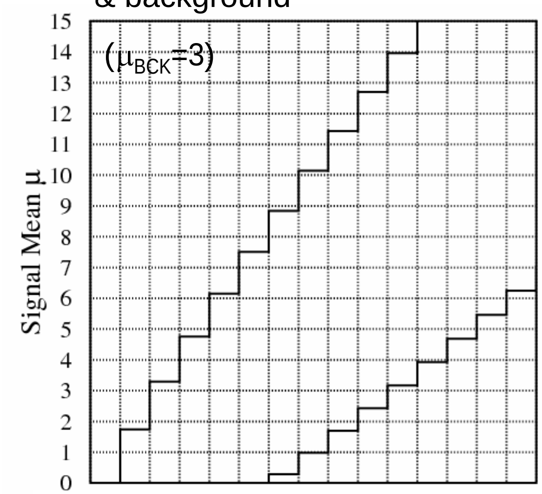


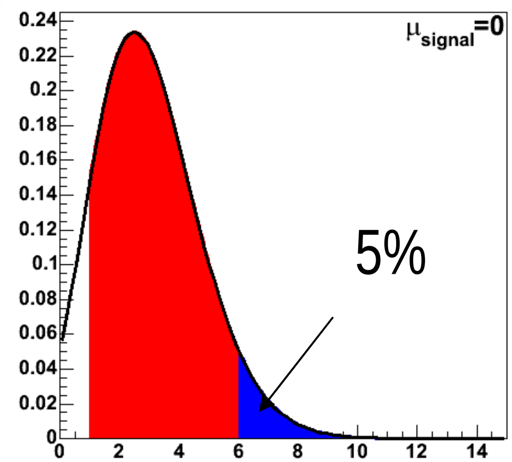
Neyman-Pearson

- Several ways of using your data
  - set a lower limit? Set an upper limit?
  - Obtain a two-sided bound?
  - Measure Δms?
- We want to discern between
  - H_0 = no signal
  - H_1 = mixing at a certain Δm value
- Neyman-Pearson test:
  - Pick an observable ξ, e.g.:
    - Significance of the highest peak in A-scan
    - Likelihood ratio (UMP! Neyman-Pearson lemma!)
  - Derive: $P(\xi|H_0)$ $P(\xi|H_1)$
  - Define:
    - Bands in ξ for rejecting H_0/H_1

    $\Rightarrow$ Desired detection & false alarm probabilities
  - Open the box!

Example: 90% CL poisson signal & background

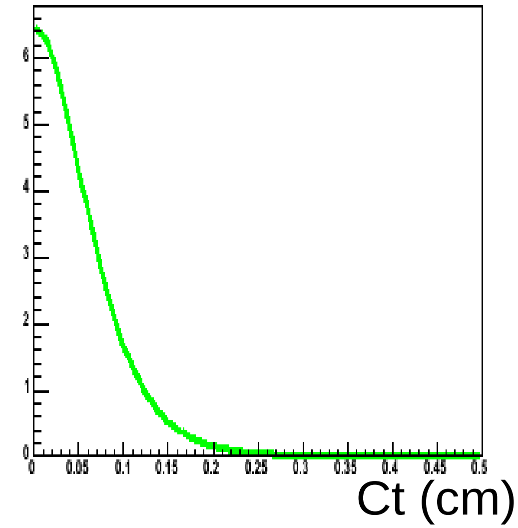


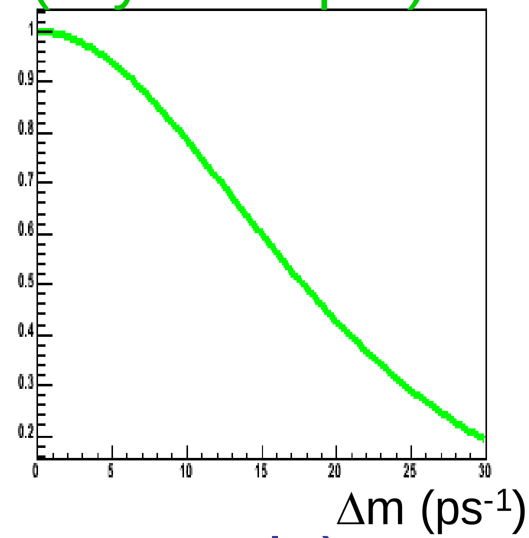


- Dangerous things:
  - Defining procedure (observable, probability thresholds and bands) after looking at your sample
  - Being confused about the procedure
  - Switch from one way of using data to another (limit vs measurement)

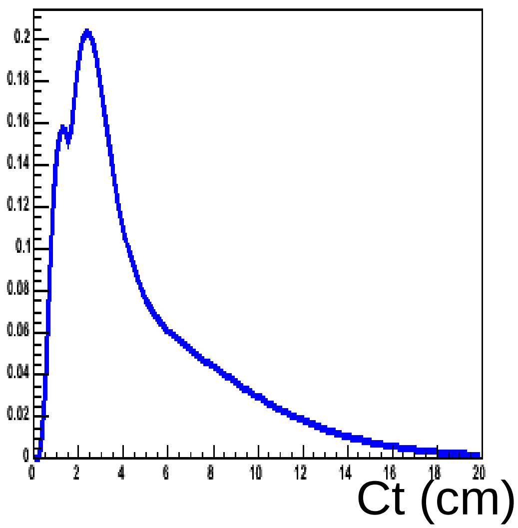
Mixing & Fourier Transforms

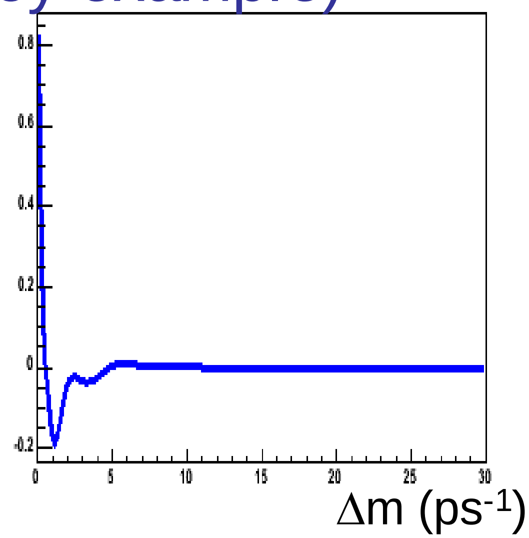
Ct Resolution (toy example)

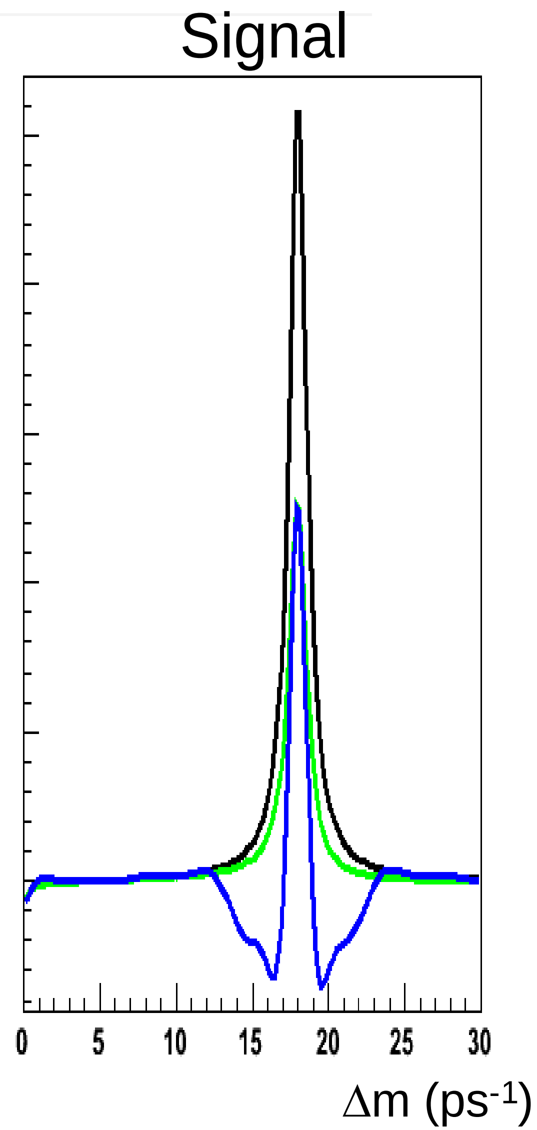




ε Curve (toy example)

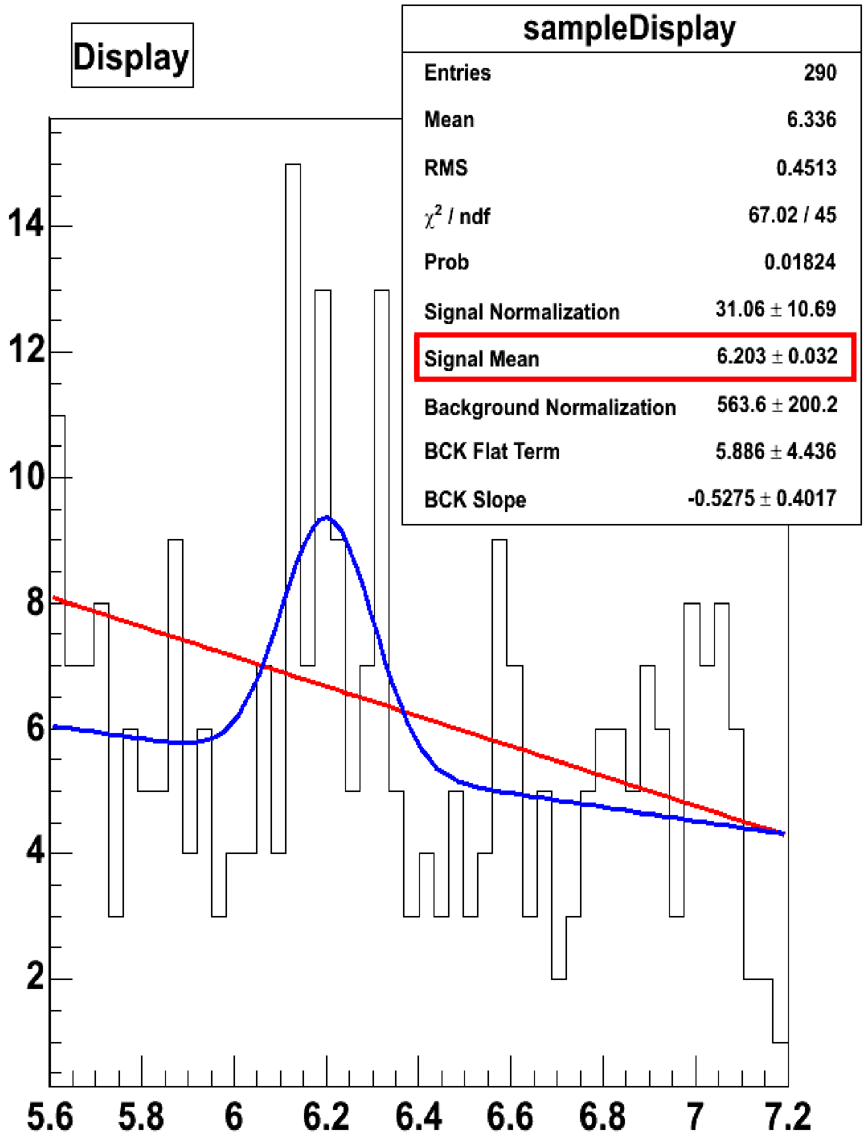






Analogy: searching for a peak

- Familiar problem with analogous issues:
  - Unknown mass (Δm)
  - Some knowledge of width
- Peak hunting is dangerous:
  - Easy to bias yourself from:
    - prior knowledge
    - Statistical fluctuations
  - Sensitivity depends on:
    - Binning (can go unbinned though, if mass model is robust)
    - Search window
- m can be measured pretty well on a statistical fluctuation!



B_s Mixing Ingredients

Proper time resolution

Flavor tagging

Signal-to-noise

Event yield

$$Significance = \sqrt{\frac{S\varepsilon D^2}{2}}\, e^{-\frac{(\Delta m_s \sigma_t)^2}{2}} \sqrt{\frac{S}{S+B}}$$

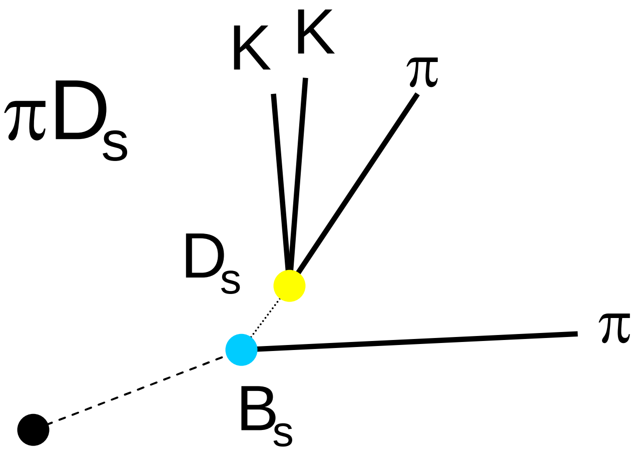
B_s Mixing Ingredients: σ_{ct}

Proper time resolution

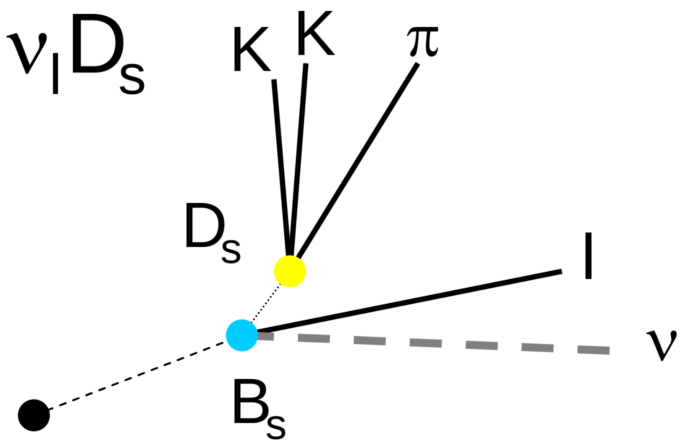
$$Significance = \sqrt{\frac{S\varepsilon D^2}{2}e^{-\frac{(\Delta m_s \sigma_t)^2}{2}}}\sqrt{\frac{S}{S+B}}$$

Proper time resolution

$B_s \to \pi D_s$



$B_s \to l \nu_l D_s$



~0.5%

$$ct = \frac{L_{xy}}{(\beta\gamma)} = \frac{L_{xy} m_B}{p_T} \quad \longrightarrow \quad \sigma_{ct} = \frac{m_B}{p_T}\sigma_{L_{xy}} \oplus ct\left(\frac{\sigma_{p_T}}{p_T}\right)$$

~15%

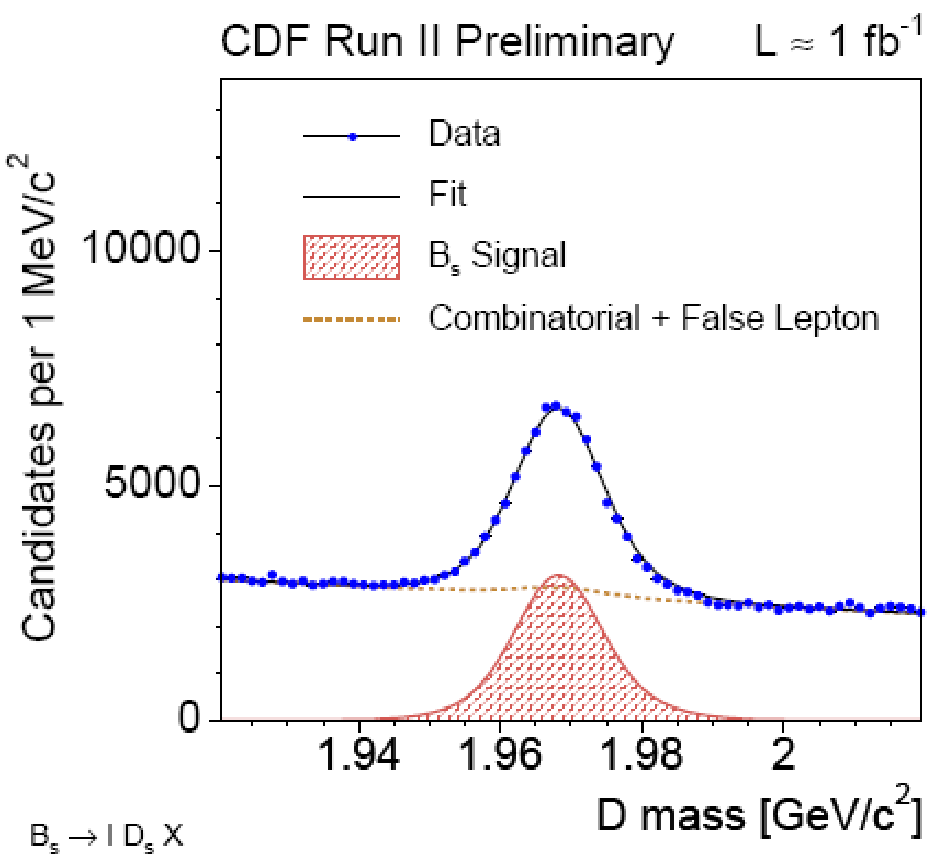
$$ct = \frac{m_B L_{xy}}{P_t(lD_s)} \cdot \left\langle \frac{P_t(lD_s)}{P_t(B_s)} \right\rangle_{mc} \quad \longrightarrow \quad \sigma_{ct} = \frac{m_B}{P_t}\sigma_{L_{xy}} \oplus ct\left(\frac{\sigma_{P_t}}{P_t}\right) \otimes \sigma_K$$

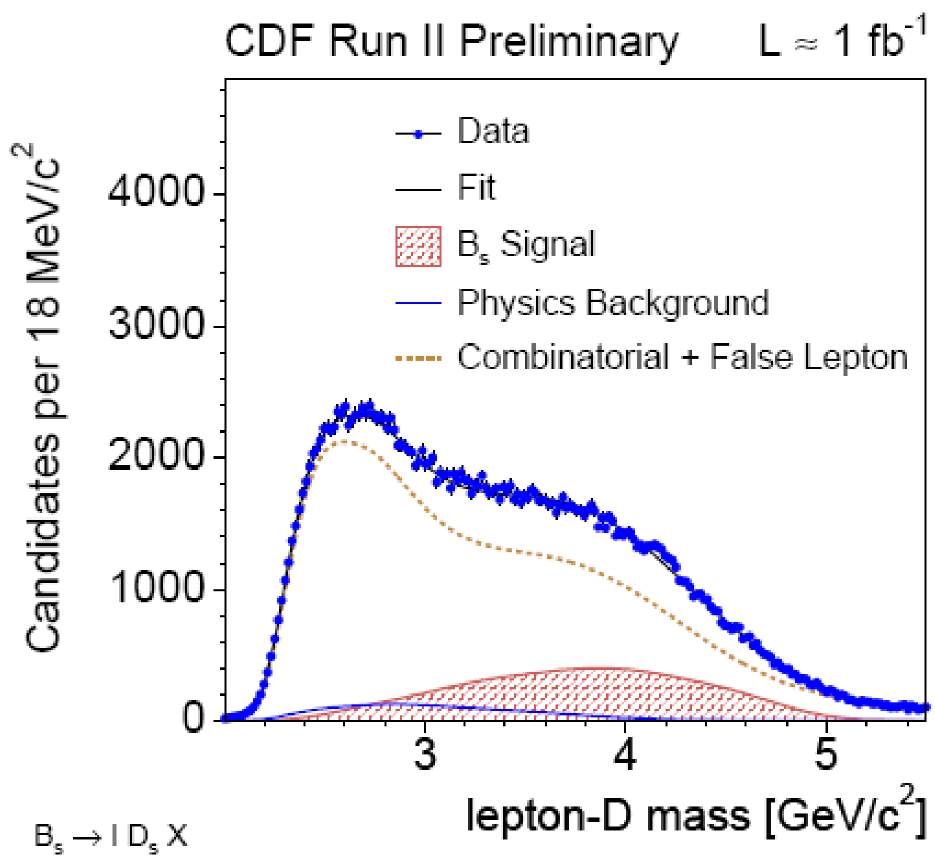
Semileptonic modes: momentum uncertainty

Fully reconstructed: Lxy uncertainty → improve reconstruction

Samples of B_s Decays

Semileptonic Samples: $D_s^- l^+ X$





~53 K events

$m(lD_s^-)$ distribution

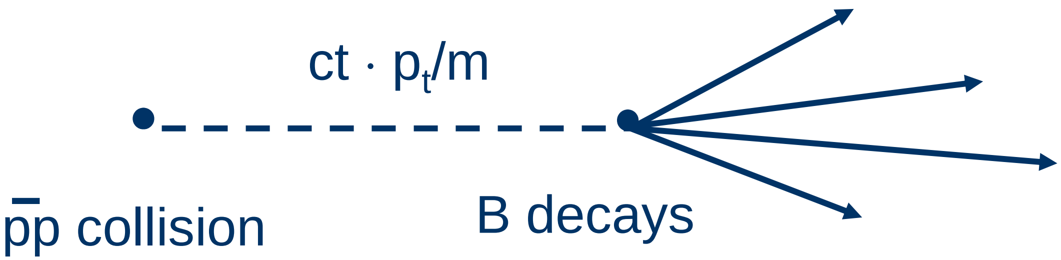
Signal Yield Summary: Semileptonic

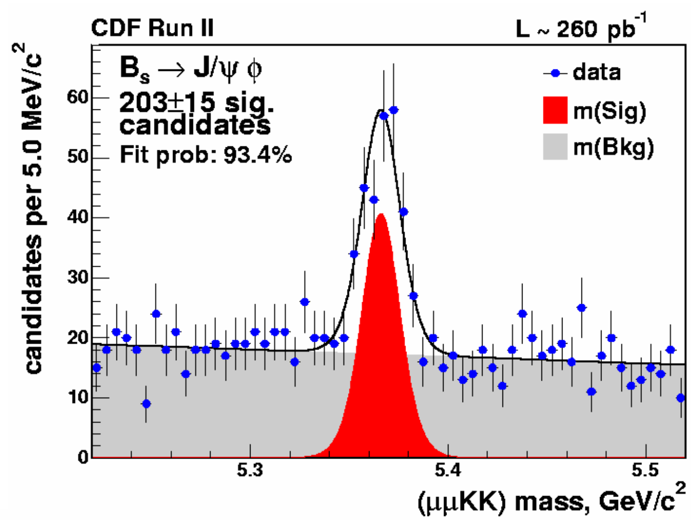
	muon	electron
lD_s: $D_s \to \phi\pi$	~ 24 K	~ 8 K
lD_s: $D_s \to K^*K$	~ 8 K	~ 3 K
lD_s: $D_s \to \pi\pi\pi$	~ 7.5 K	~ 2.5 K

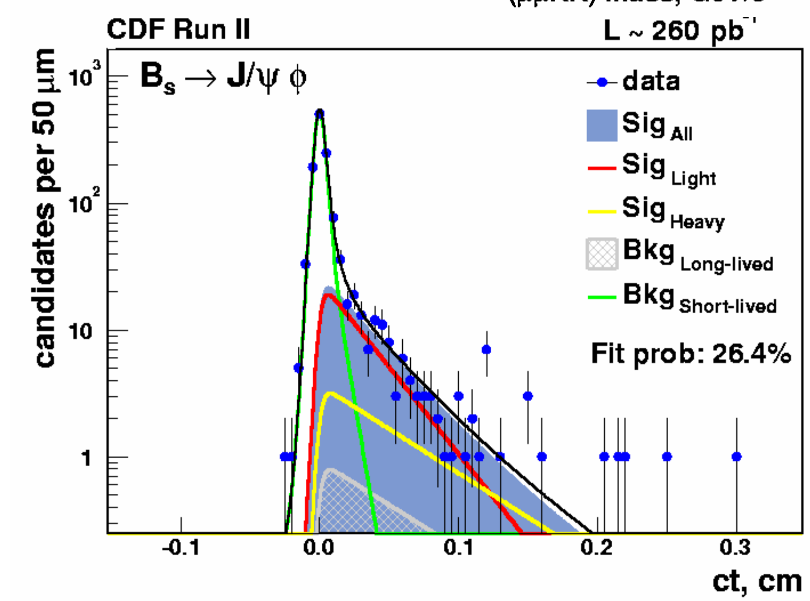
lD^0: $D^0 \to K\pi$	~ 400 K	~ 140 K
lD^{*-}: $D^0 \to K\pi$	~ 54 K	~ 21 K
lD^-: $D^- \to K\pi\pi$	~ 220 K	~ 80 K

B Lifetime Measurements

“Classic” B Lifetime Measurement







- reconstruct B meson mass, p_T, L_{xy}
- calculate proper decay time (ct)
- extract $c\tau$ from combined mass+lifetime **fit**
- signal probability:

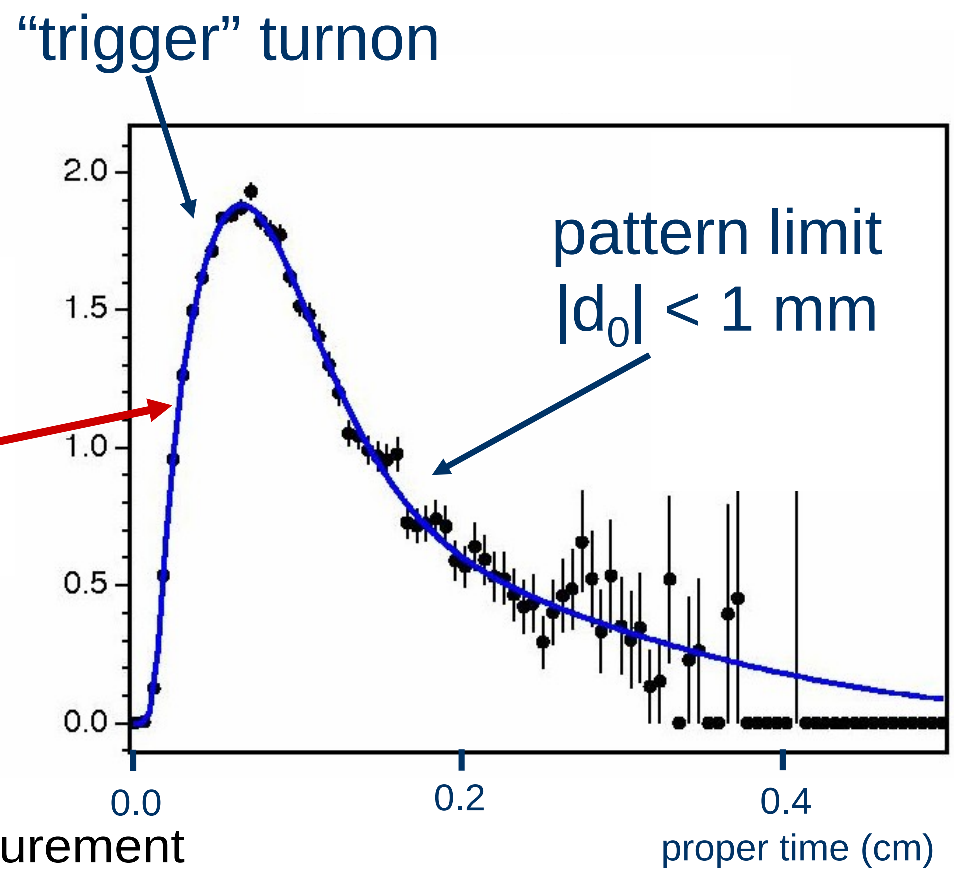
  $p_{signal}(\mathbf{t}) = e^{-t'/\tau} \otimes R(t', \mathbf{t})$

- background $p_{bkgd}(t)$ modeled from sidebands

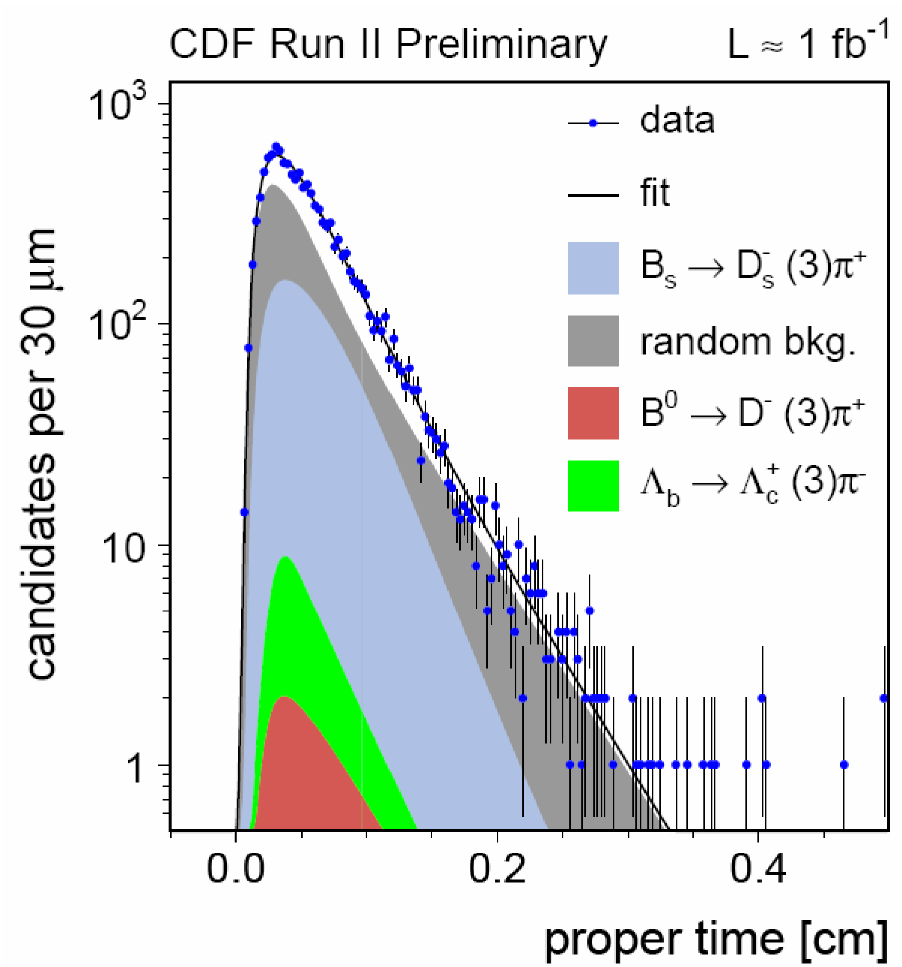
Hadronic Lifetime Measurement

- SVT trigger, event selection sculpts lifetime distribution
- correct for on average using efficiency function:

  $p = e^{-t'/\tau} - R(t', t) \otimes \varepsilon(t)$
- efficiency function shape contributions:
  - event selection, trigger
- details of efficiency curve
  - important for lifetime measurement
  - inconsequential for mixing measurement



Hadronic Lifetime Results



Mode	Lifetime [ps] (stat. only)
$B^0 \to D^- \pi^+$	1.508 ± 0.017
$B^- \to D^0 \pi^-$	1.638 ± 0.017
$B_s \to D_s \pi(\phi\pi)$	1.538 ± 0.040

- World Average:

B^0 1.534 ± 0.013 ps^{-1}

B^+ 1.653 ± 0.014 ps^{-1}

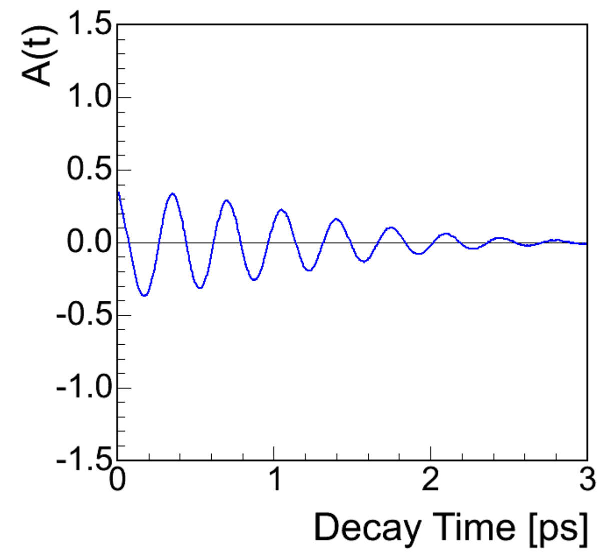
B_s 1.469 ± 0.059 ps^{-1}

Excellent agreement!

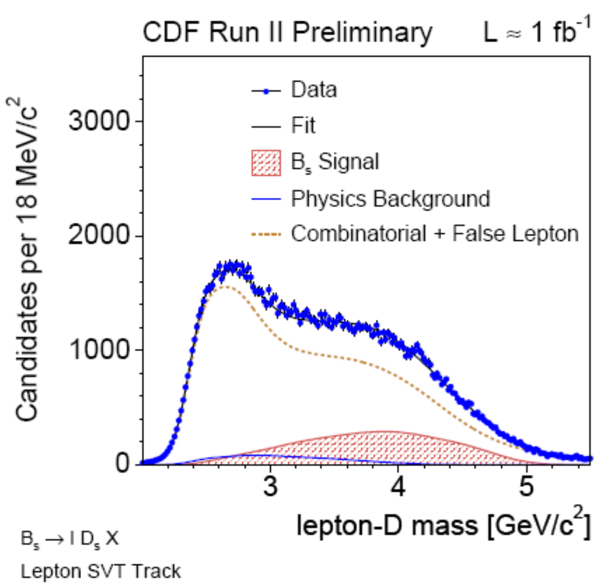
Semileptonic Lifetime Measurement

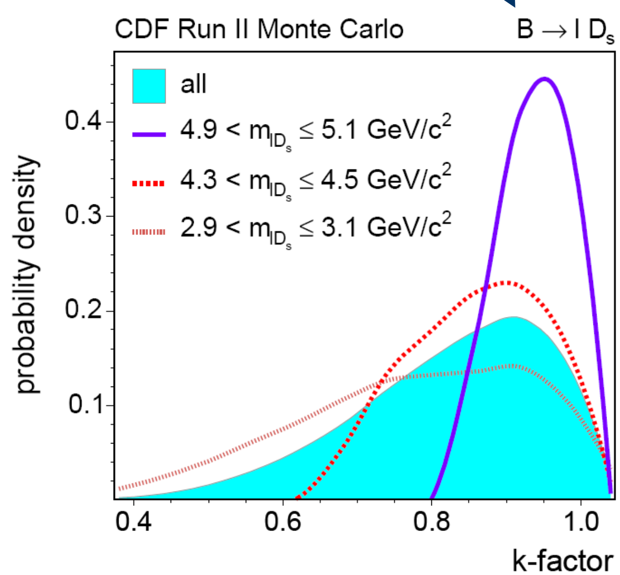
- neutrino momentum not reconstructed

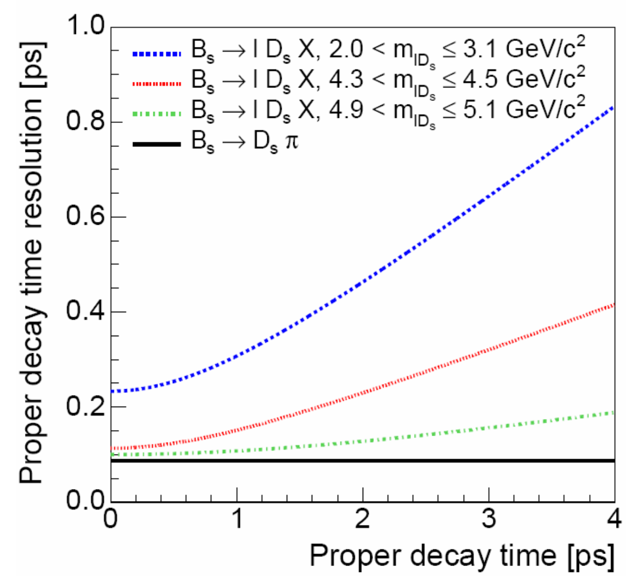
$$K = \frac{p_T(lD)}{p_T(B)} \cdot \frac{L(B)}{L(lD)}$$



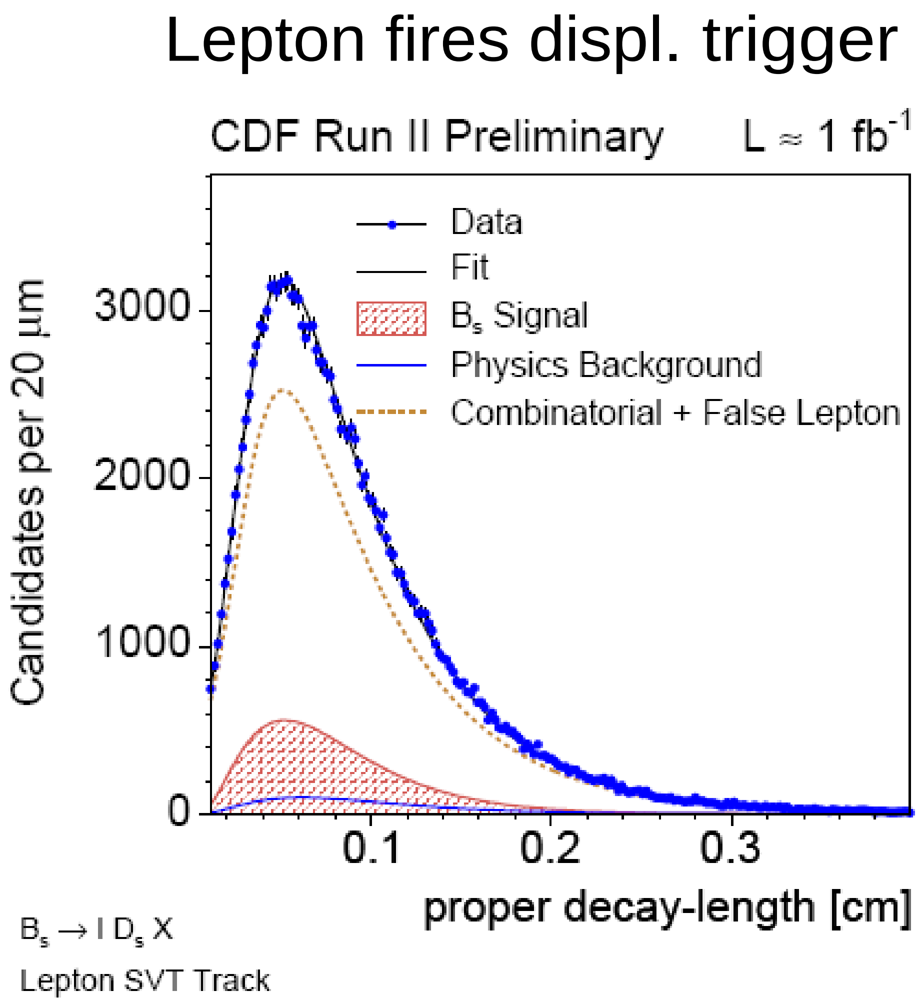
- correct for neutrino on average

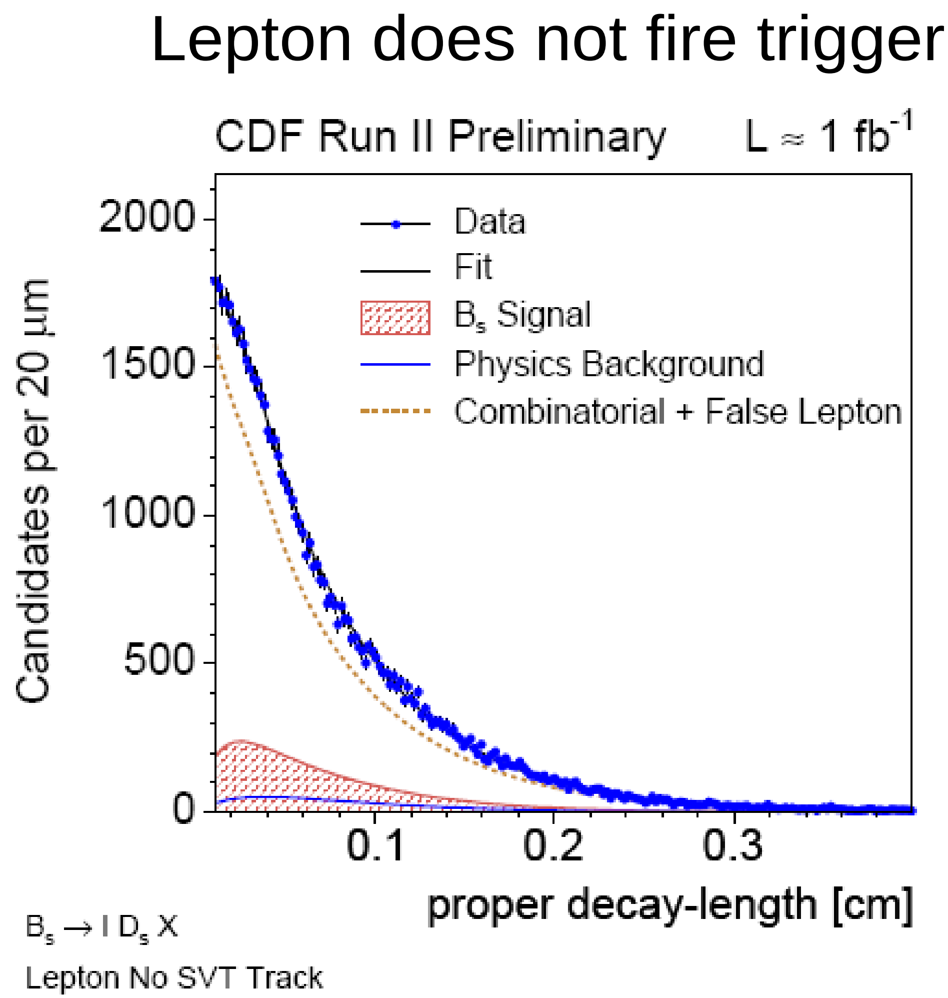






lD_s ct* Projections





Semileptonic Lifetime Results

	Lifetime (ps)
Bs: Ds $\rightarrow \phi\pi$	1.51 ± 0.04 stat. only
Bs: Ds $\rightarrow$ K*K	1.38 ± 0.07 stat. only
Bs: Ds $\rightarrow \pi\pi\pi$	1.40 ± 0.09 stat. only
Bs combined	1.48 ± 0.03 stat. only

- lifetimes measured on first 355 pb^{-1}
- compare to World Average: B_s: (1.469 ± 0.059) ps

Proper Time Resolution

Proper Time Resolution

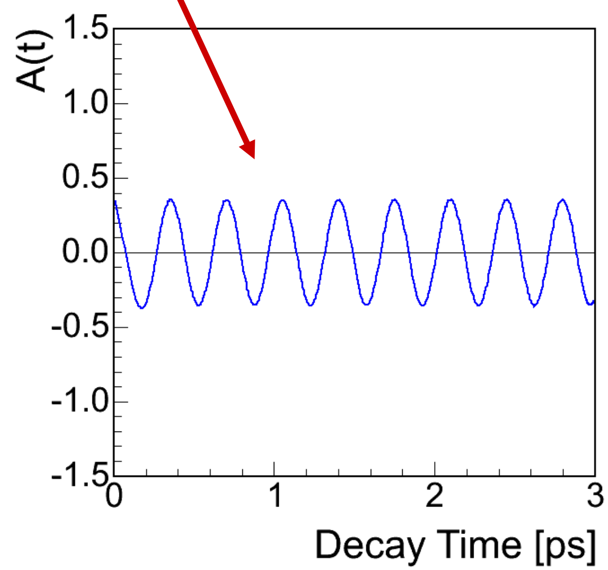
- Reminder,
  measurement
  significance: $\text{Signif} = \sqrt{\frac{N\epsilon D^2}{2}} e^{-\frac{(\Delta m_s \sigma_t)^2}{2}} \sqrt{\frac{S}{S+B}}$

- significant effect

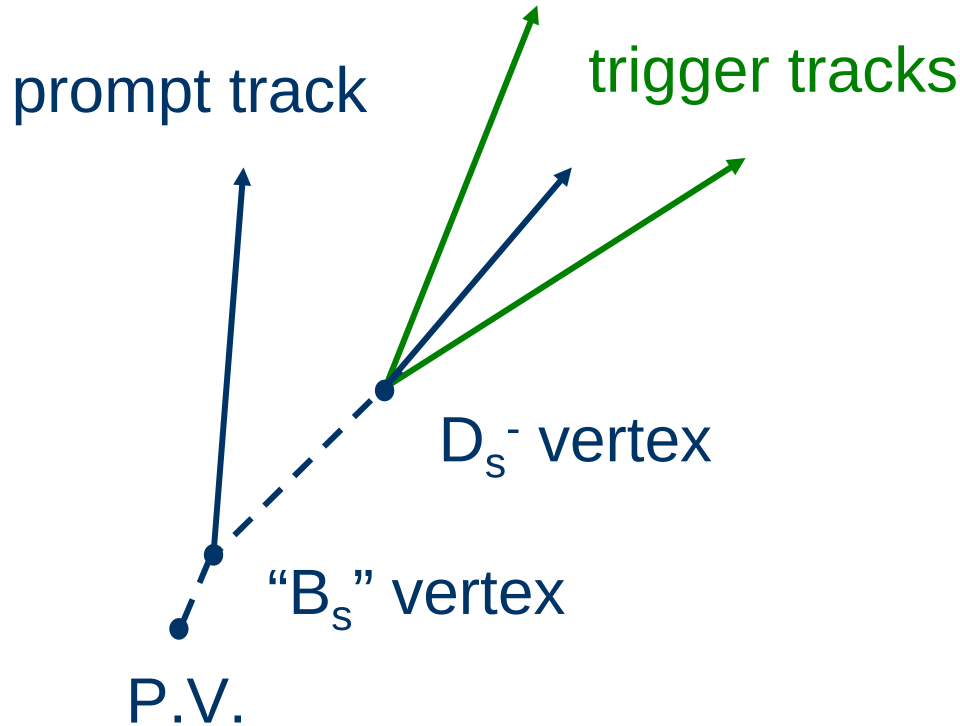
- fitter has to correctly account for it

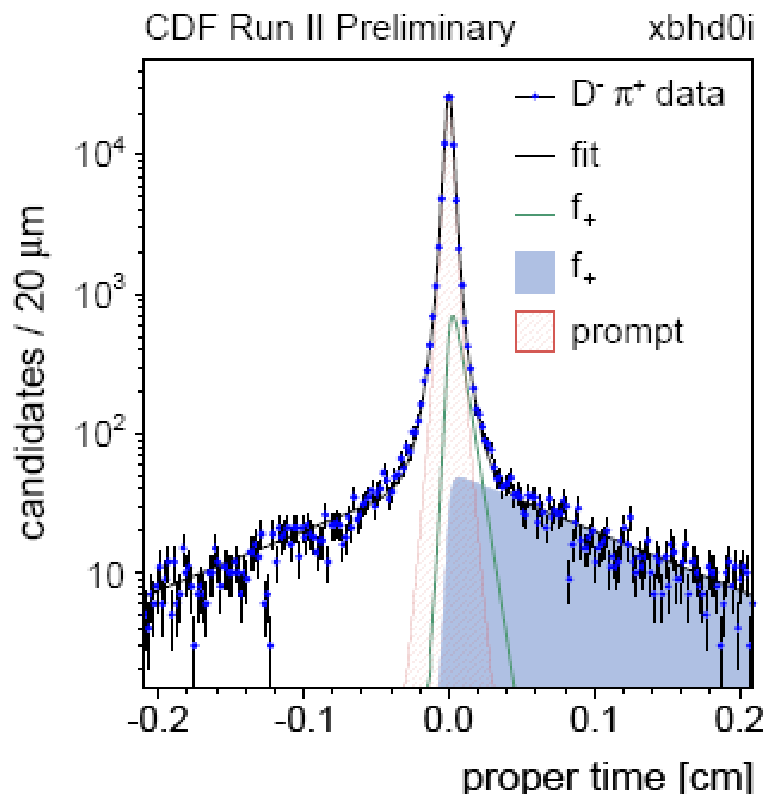
- lifetime measurements not very sensitive to resolution

- a dedicated calibration is needed!



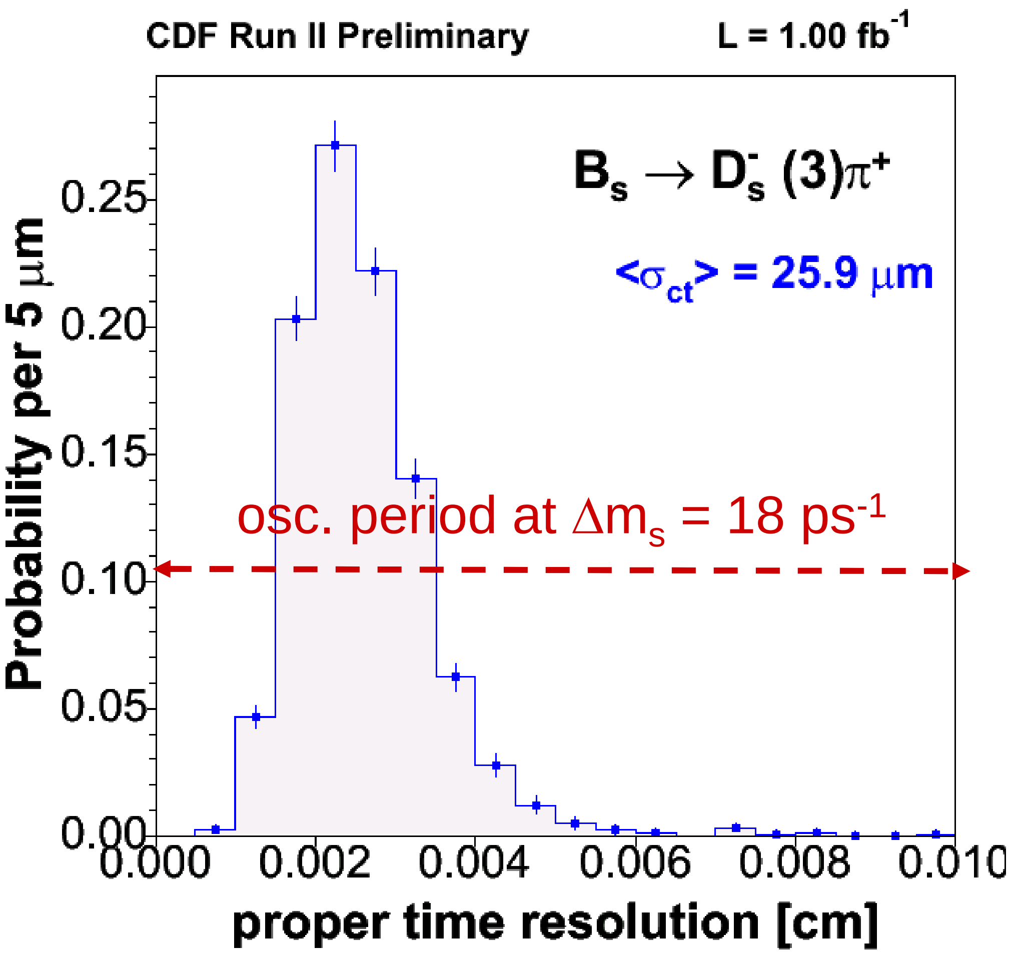
Calibrating the Proper Time Resolution





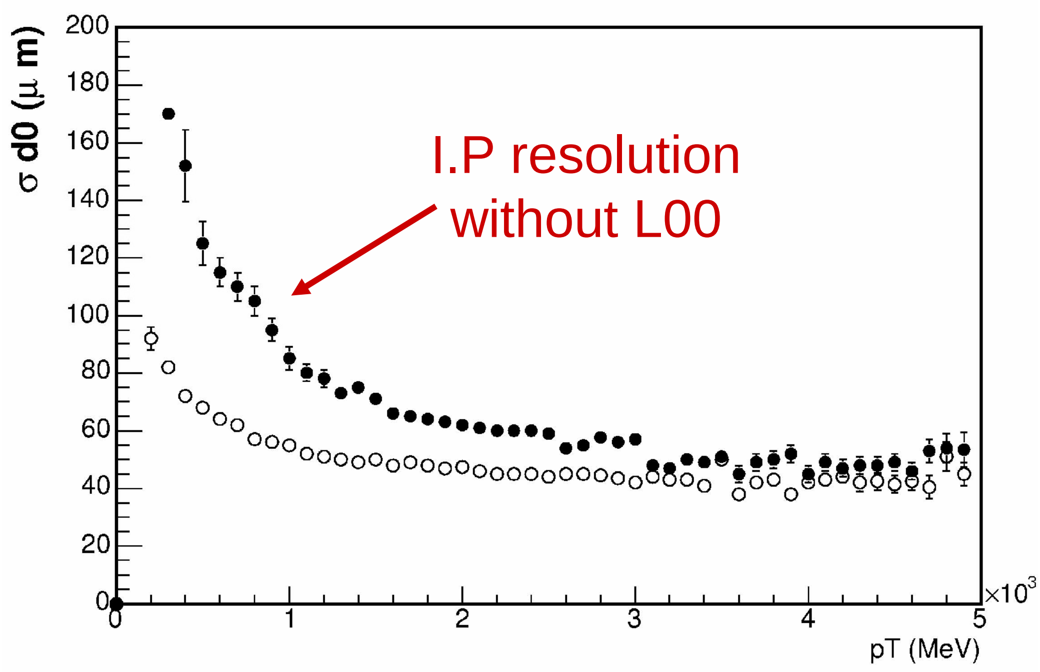
- utilize large prompt charm cross section
- construct "Bs-like" topologies of prompt D_s^- + prompt track
- calibrate ct resolution by fitting for "lifetime" of "Bs-like" objects

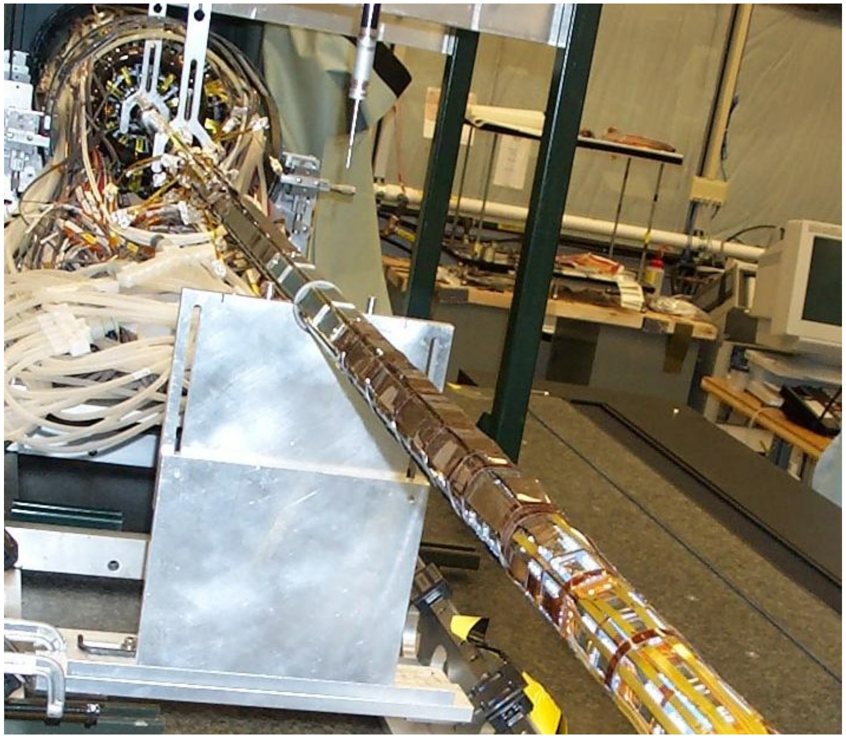
B_s Proper Time Resolution



- event by event determination of primary vertex position used
- average uncertainty $\sim 26\ \mu m$
- this information is used per candidate in the likelihood fit

Layer "00"

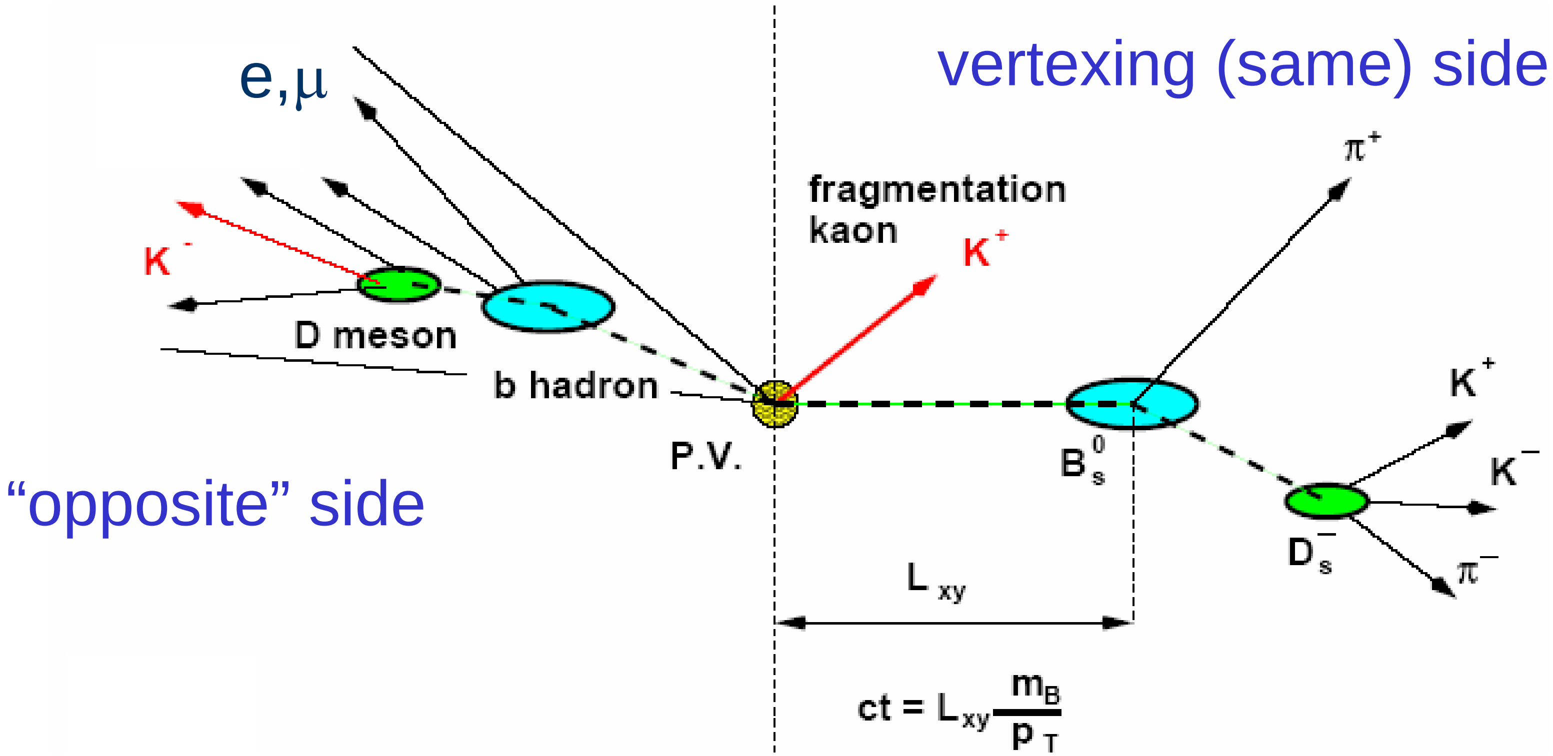




- layer of silicon placed directly on beryllium beam pipe
- radial displacement from beam ~1.5 cm
- additional impact parameter resolution, radiation hardness

Flavor Tagging

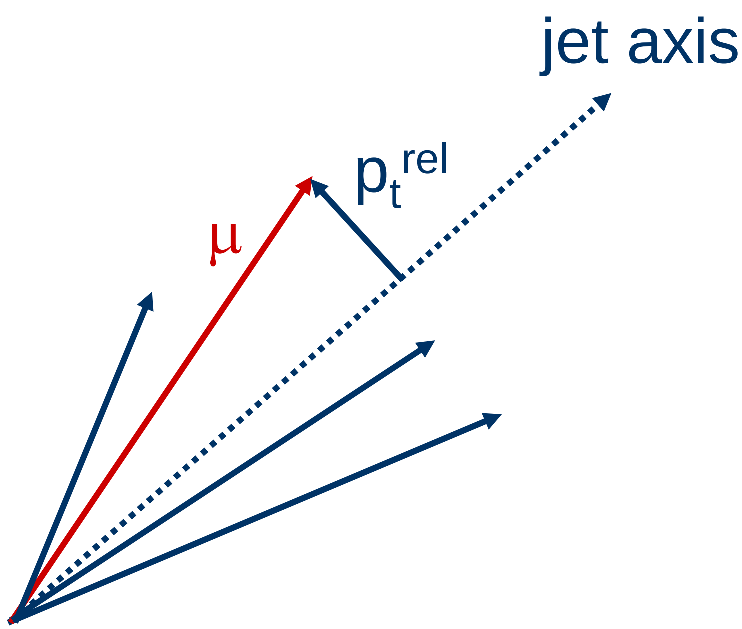
Tagging the B Production Flavor

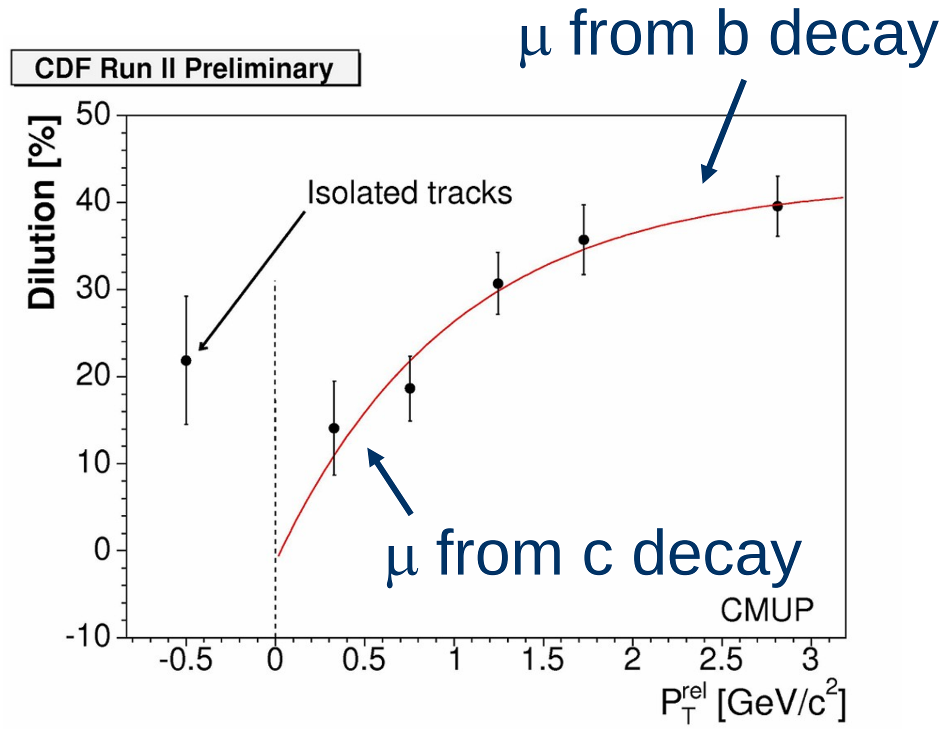


- use a combined same side and opposite side tag!
- use muon, electron tagging, jet charge on opposite side
- jet selection algorithms: vertex, jet probability and highest p_T
- particle ID based kaon tag on same side

Parametrizing Tagger Decisions

- use characteristics of tags themselves to increase their tagging power, example: muon tags

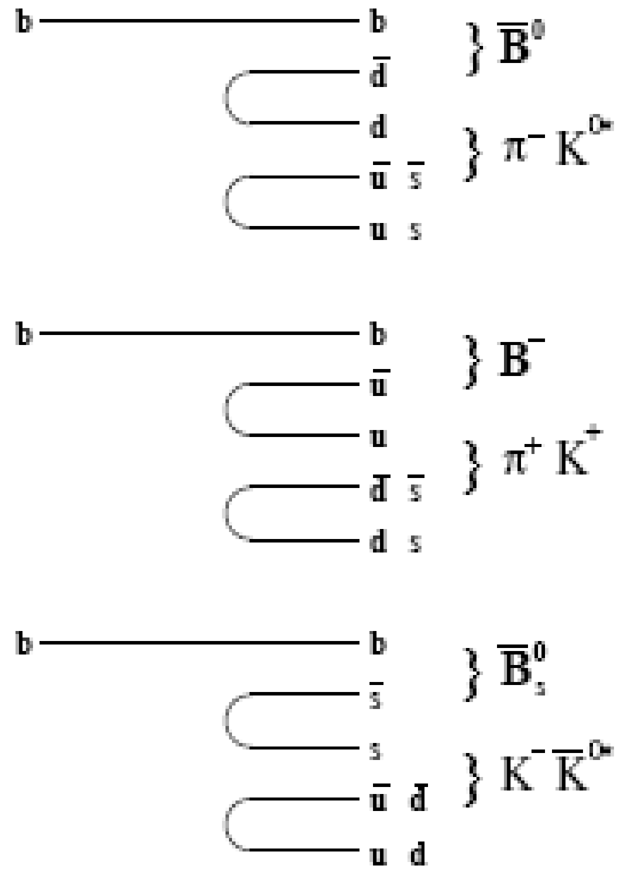




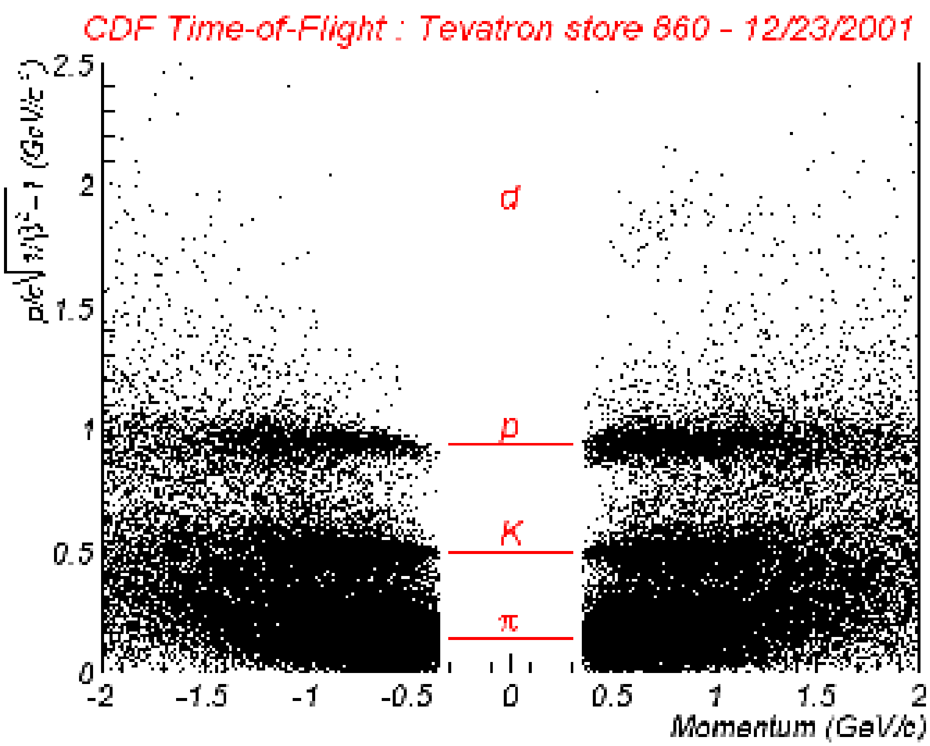
- tune taggers and parametrize event specific dilution
- technique in data works with opposite side tags

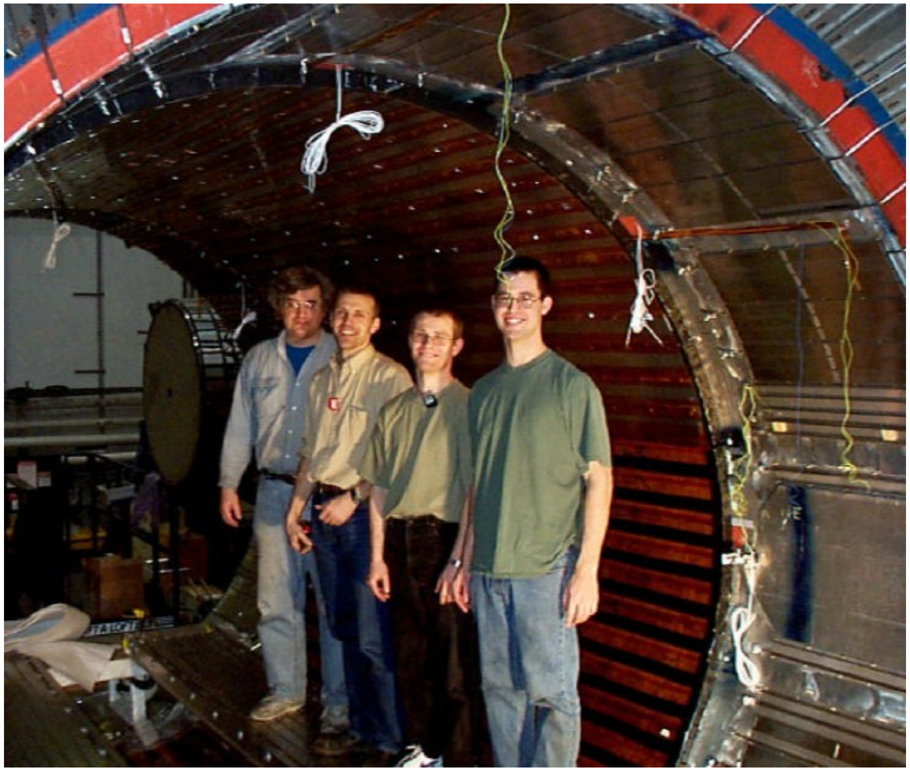
Same Side Kaon Tags

- exploit b quark fragmentation signatures in event
- B^0/B^+ likely to have a π^-/π nearby
- B_s^0 likely to have a K^+
- use TOF and COT dE/dX info. to separate pions from kaons
- problem: calibration using only B^0 mixing will not work
- tune Monte Carlo simulation to reproduce B^0, B^- distributions, then apply directly to B_s^0



Time Of Flight System

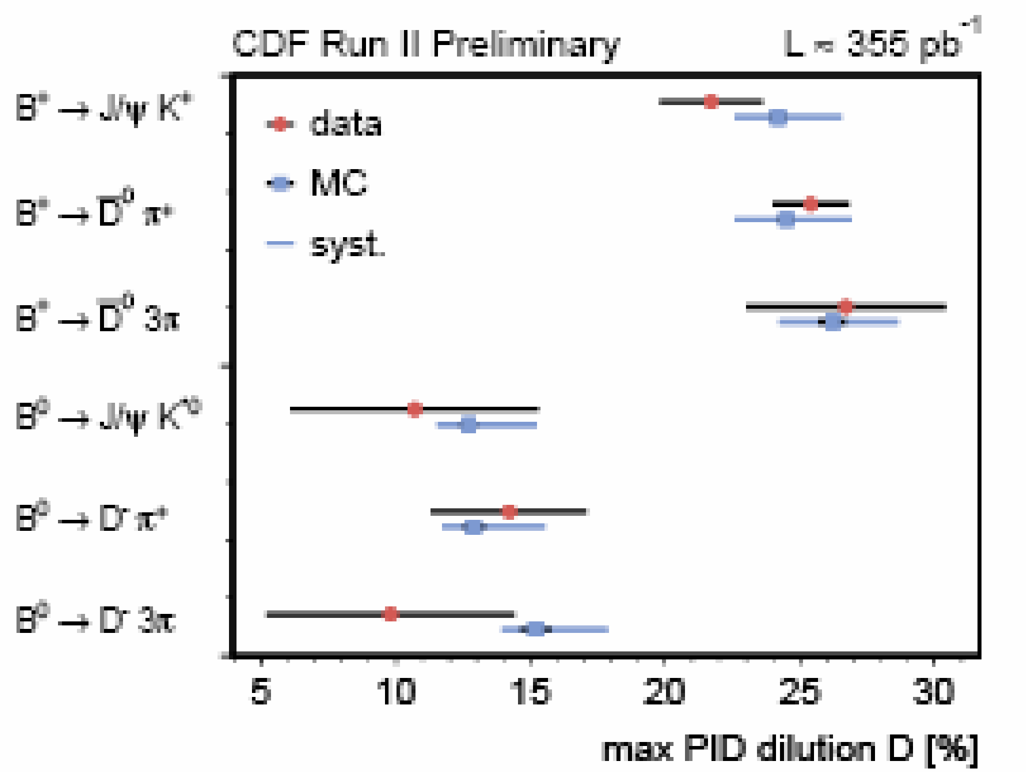




- timing resolution ~100 ps ! resolves kaons from pions up to p ~ 1.5 GeV/ c
- TOF provides most of the Particle ID power for SSKT

Calibrating SSKT

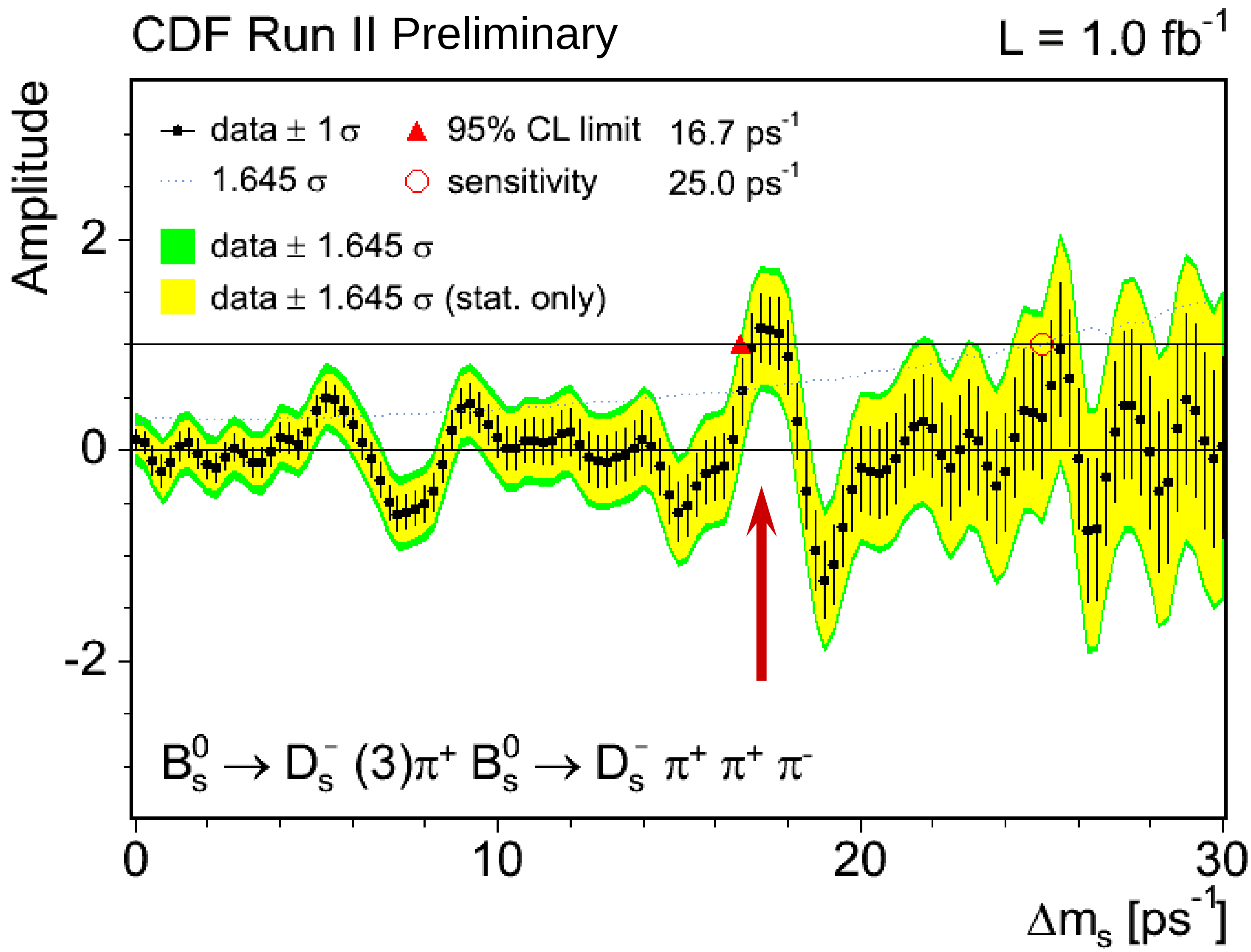
- Analogous to transfer scale factor in Opposite Side Tags
- Check dilution in light B meson decays



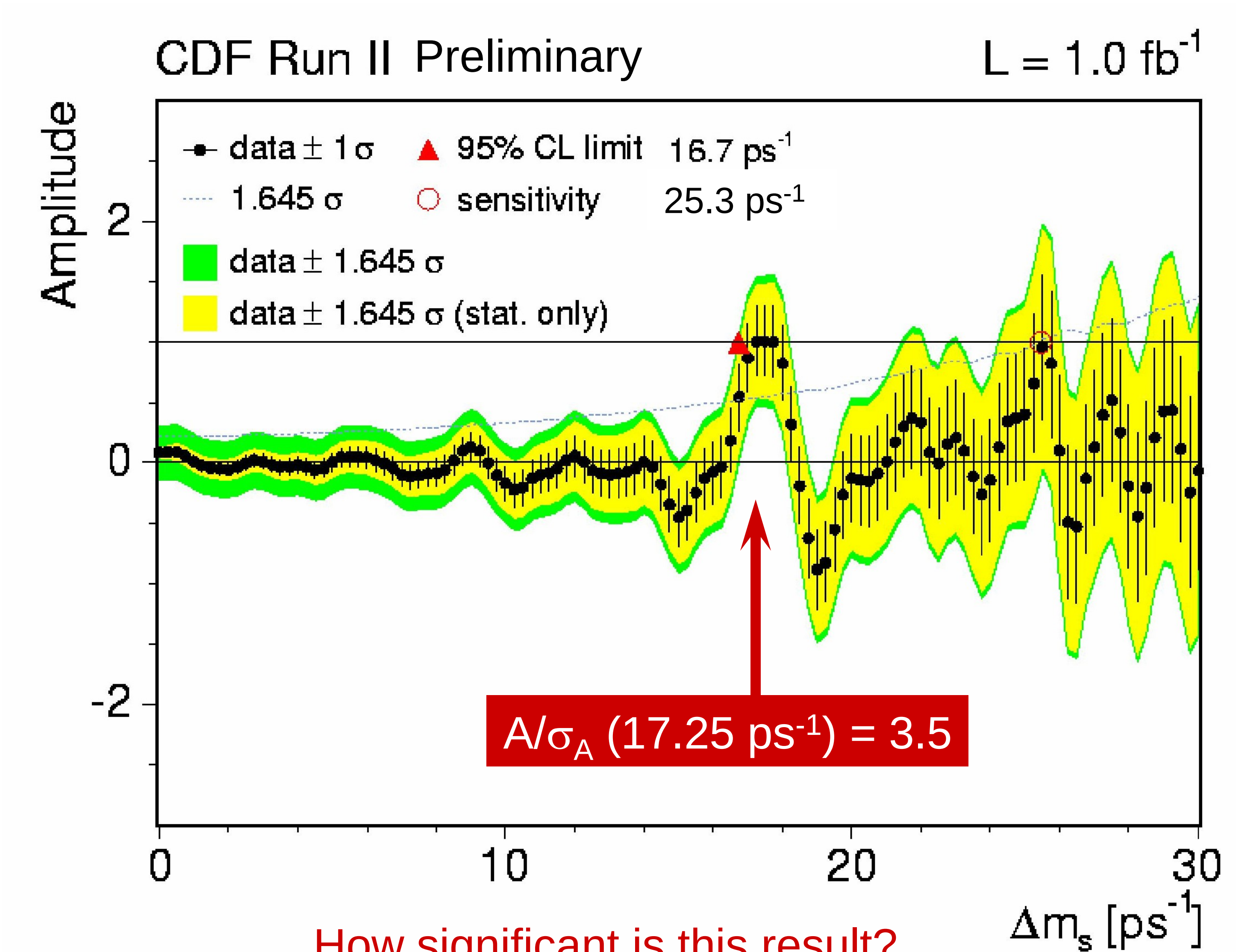
Data/MC agreement is the largest systematic uncertainty ! O(8%)

The Data

Hadronic Scan: Combined

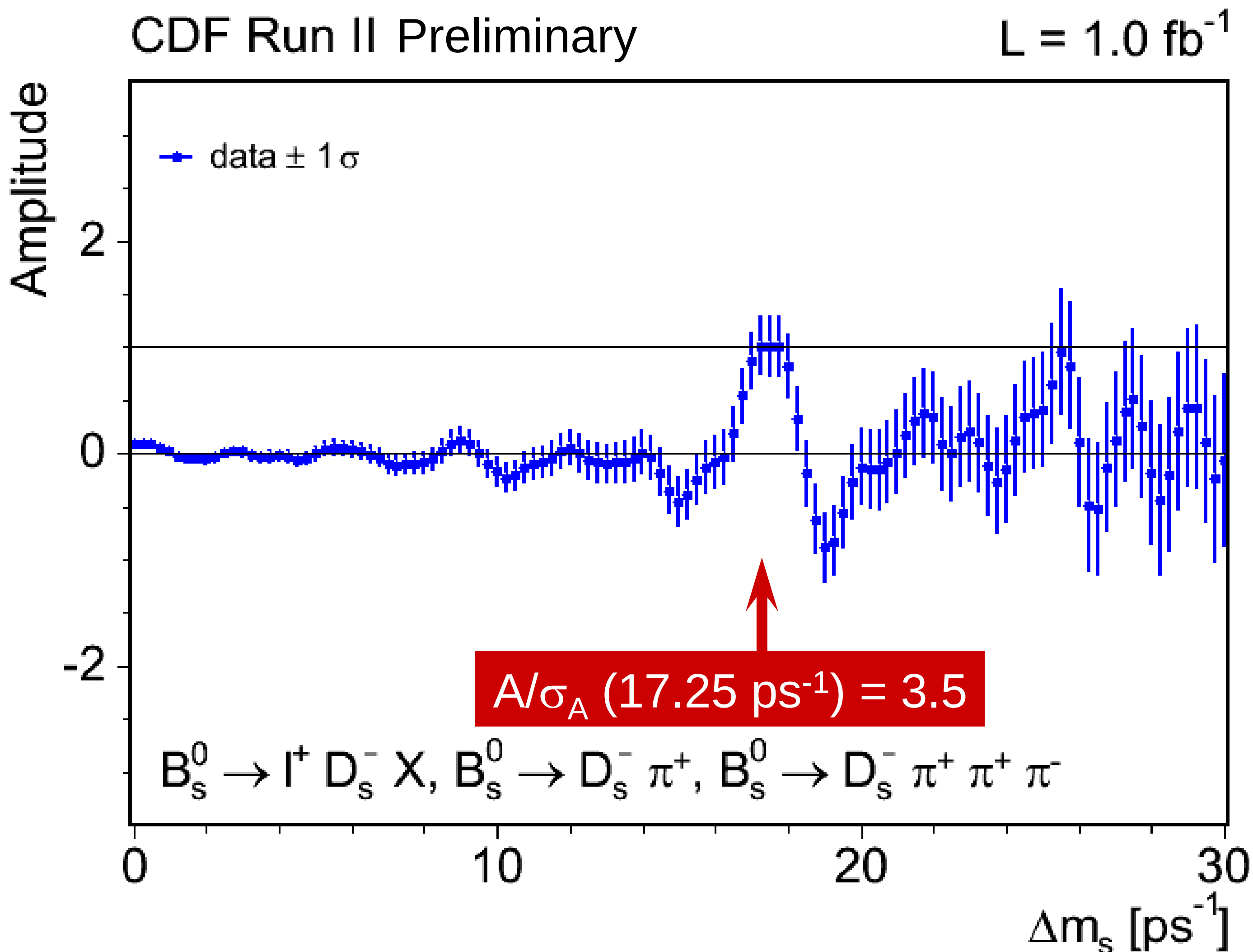


Combined Amplitude Scan



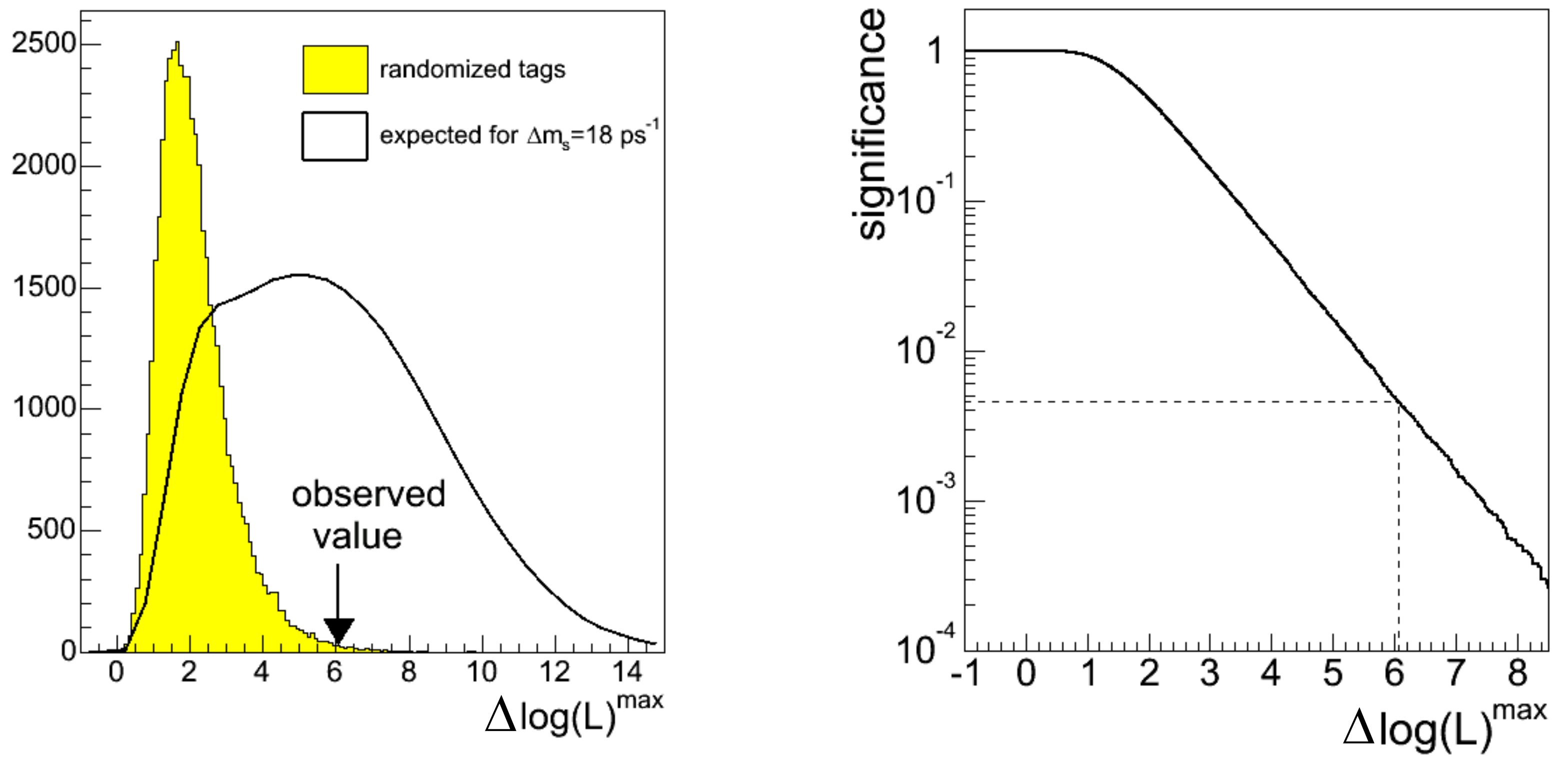
How significant is this result?

Combined Amplitude Scan



How significant is this result?

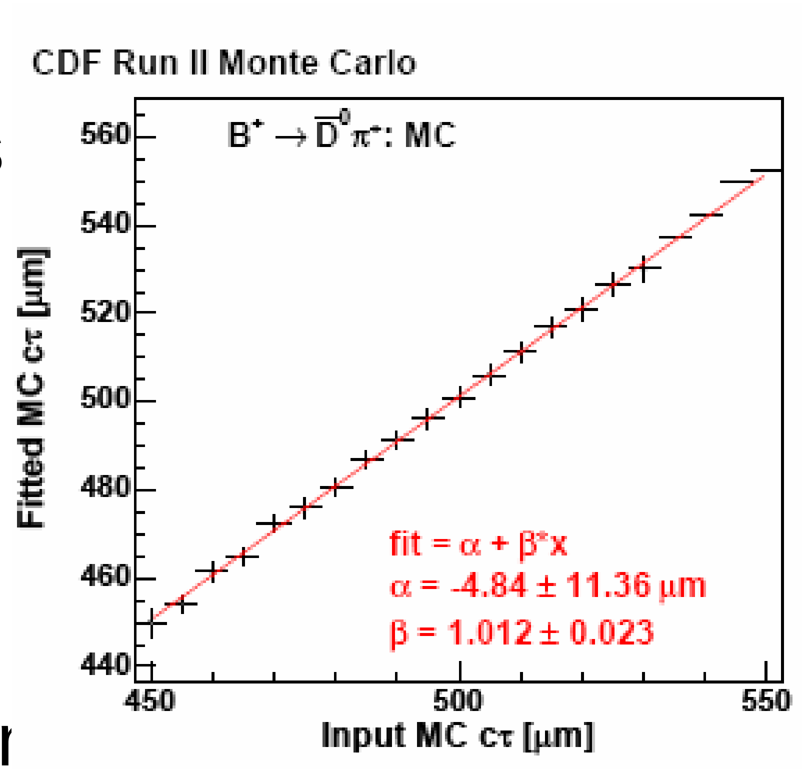
Likelihood Significance



- randomize tags 50 000 times in data, find maximum Δlog(LR)
- in 228 experiments, Δlog(LR) $\geq$ 6.06
- probability of fake from random tags = 0.5% $\Rightarrow$ measure Δm_s!

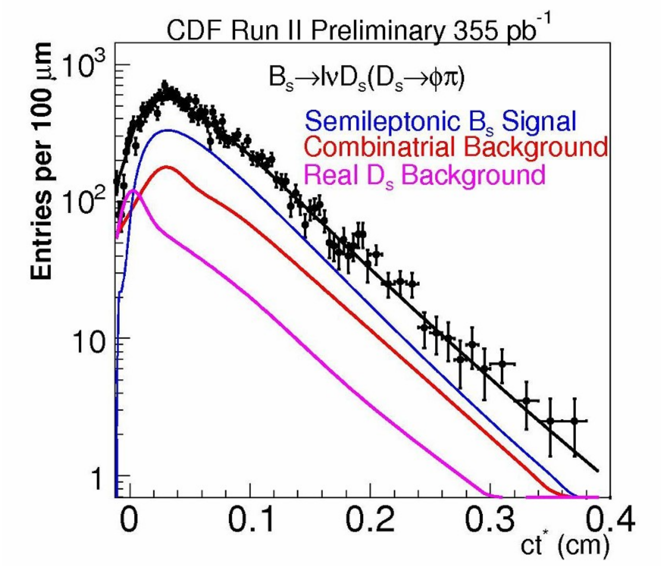
Does the MC bias the answer?

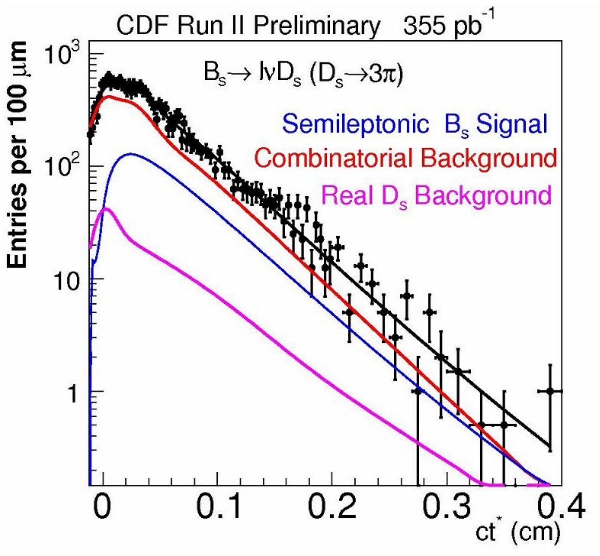
- efficiency function is derived from Monte Carlo
- the Monte Carlo is derived with an input lifetime
- does the input lifetime bias the fit outcome?
- test: fit many Monte Carlos with various input lifetimes
- derive efficiency function using one lifetime (500 µm)
- compare fit result to input lifetime
- observe no bias for ±50 µm
- measurement stat error ~7µı

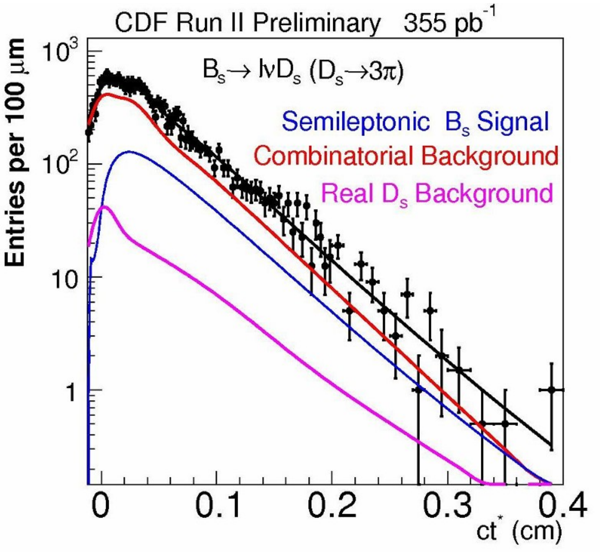


Semileptonic Lifetime Fits (Winter '05)

$D_s^- \to \phi\pi^-$ | $D_s^- \to K^*K$ | $D_s^- \to \pi^+\pi^-\pi^-$



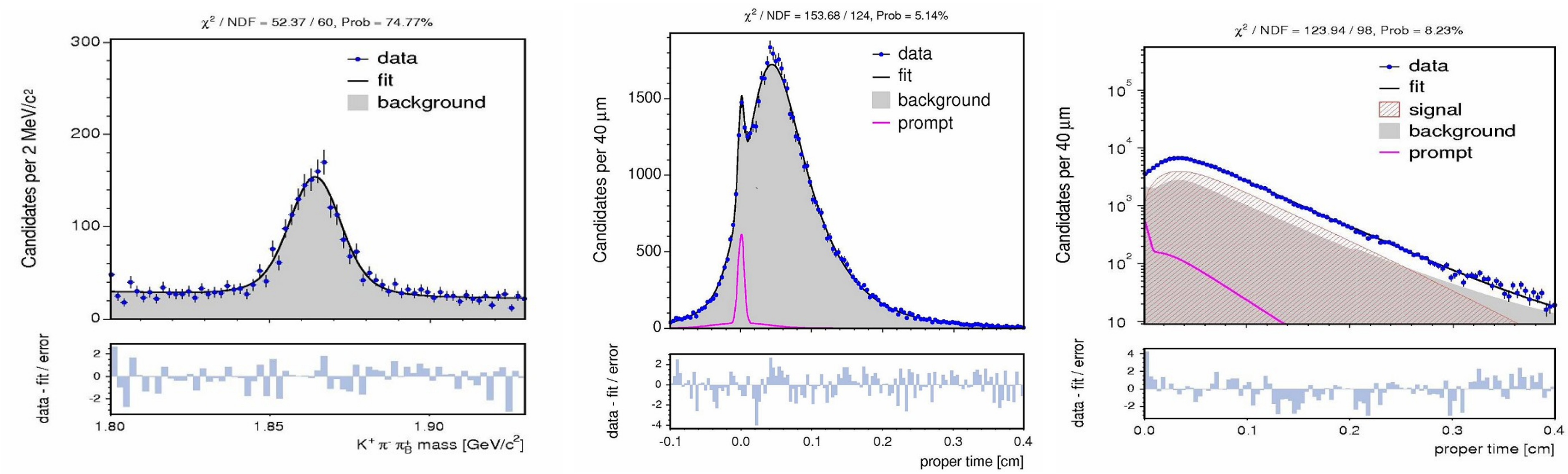




$c\tau = 455.9 \pm 11.9\ \mu m$ | $c\tau = 413.8 \pm 20.1\ \mu m$ | $c\tau = 422.6 \pm 25.7\ \mu m$

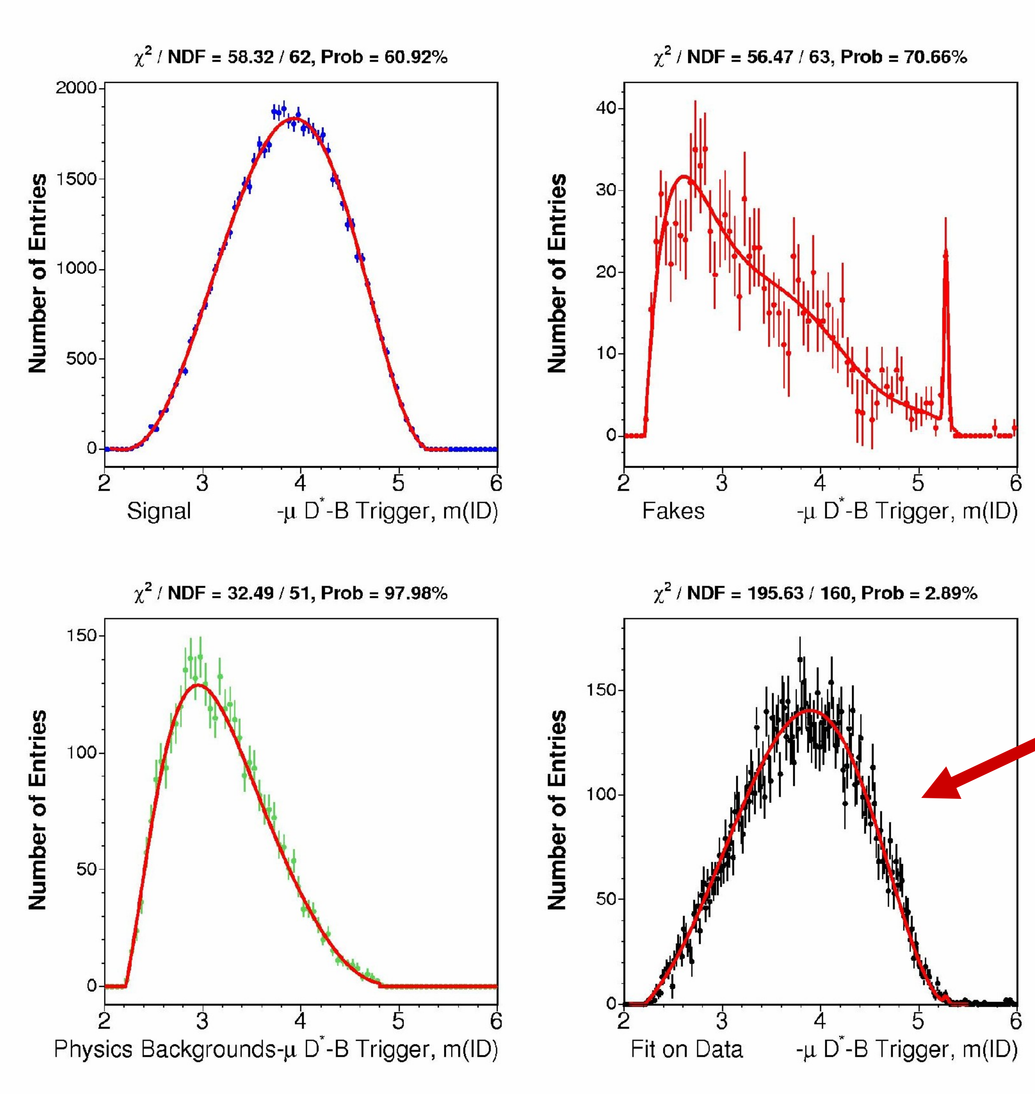
- B^0, B^+ lifetimes within 20 μm of world average values
- combined lD_s^- lifetime fit result: 445 ± 9.5 (stat) μm
- world average value: 438 ± 17 μm

“Prompt” Charm Background



- due to fake leptons, reconstruct some amount of prompt charm (D^-, D^0, D^{*-}) as B signal (in D mass signal region)
- can not disentangle from signal in any variable
- need to account for in lifetime, mixing fits
- extract shape from wrong-sign l^-D^- sample, use in fit

m(lD) fits

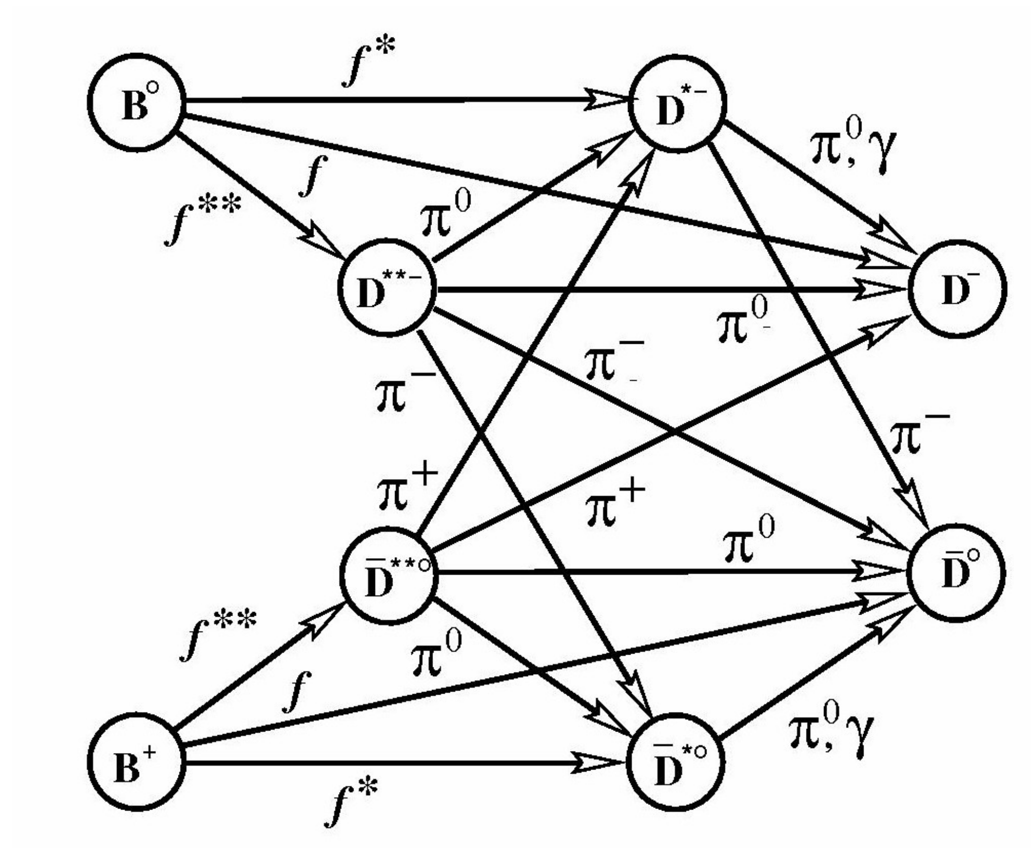


- signal distribution from Monte Carlo
- distribution for "fake" leptons from data
- physics background distribution from MC
- fit linear combination to sideband subtracted data to extract fractions

I.K.F1

Cross-Talk

- problem:
- lD^-, lD^0 are a mixture of B^+, B^0
- when fitting for lifetimes and mixing amplitude, account for this effect in fitter



Slide 82

I.K.F1 goes to backup

Ivan K Furic, 3/14/2005

Tagger Calibration

- taggers are parametrized in l+track sample
- kinematically different from final ($D_s\,\pi$, l+D_s^-)
- final tagger calibration:
- perform B^0 mixing fit in hadronic and semi-leptonic sample
- use per-event dilution, extract tagger scale factor:
- $p \sim ½\,[1 \S\, S_D\, D_i \cos(\,\Delta m_D\, t\,)]$
- use per-event corrected dilutions in $\Delta\, m_s$ **fit**
- for hadronic sample, final calibration in $D^{-/0}\pi$, $J/\psi\ K^{(*)}$
- for semileptonic sample, final calibration in $D^{-/0}\ l$, $D^{*-}\ l$

I.K.F2

Slide 83

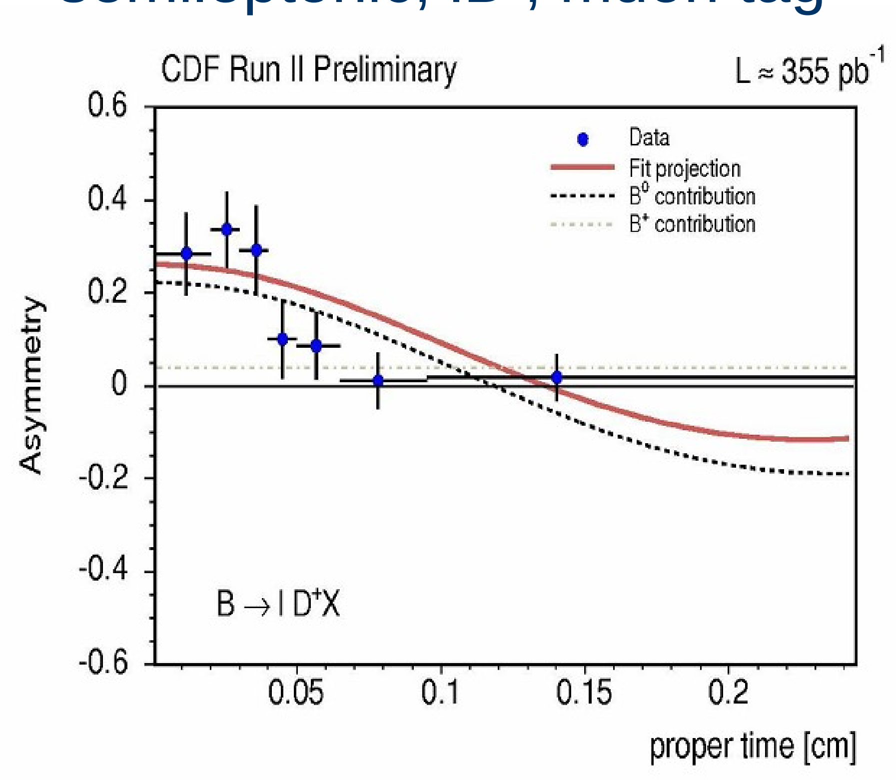
I.K.F2 move all this to backup

Ivan K Furic, 3/14/2005

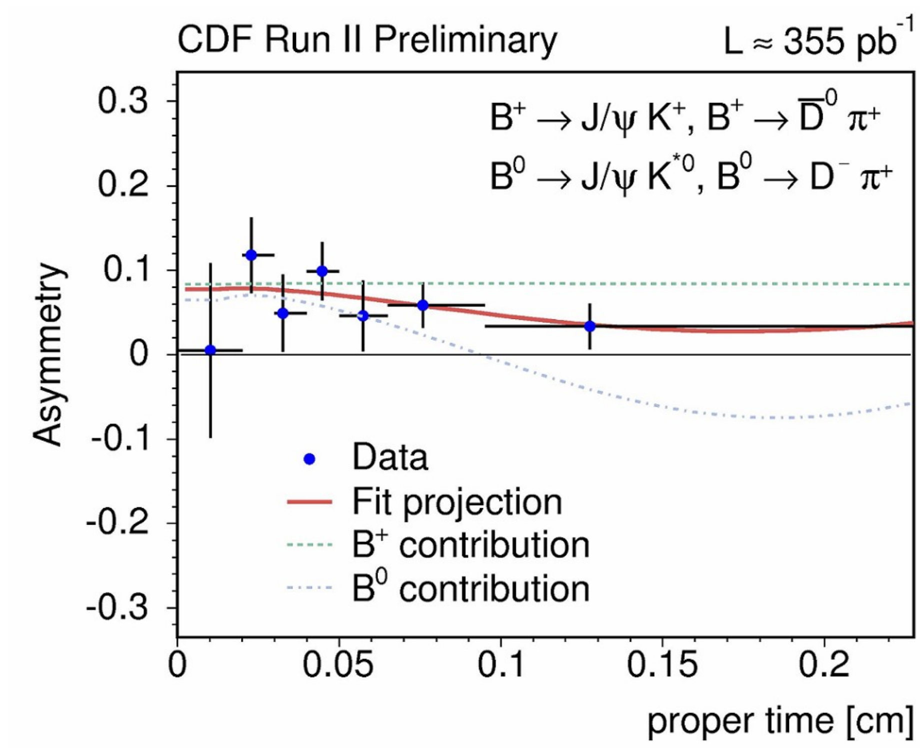
I.K.F3

$\Delta\ m_d$ Fits

semileptonic, lD^-, muon tag



hadronic, all channels, all tags



hadronic: $\Delta\ m_d = 0.503 \pm 0.063$ (stat) $\pm\ 0.015$ (syst) ps^{-1}

semileptonic: $\Delta\ m_d = 0.497 \pm 0.028$ (stat) $\pm\ 0.015$ (syst) ps^{-1}

I.K.F3

unbinned likelihood fit

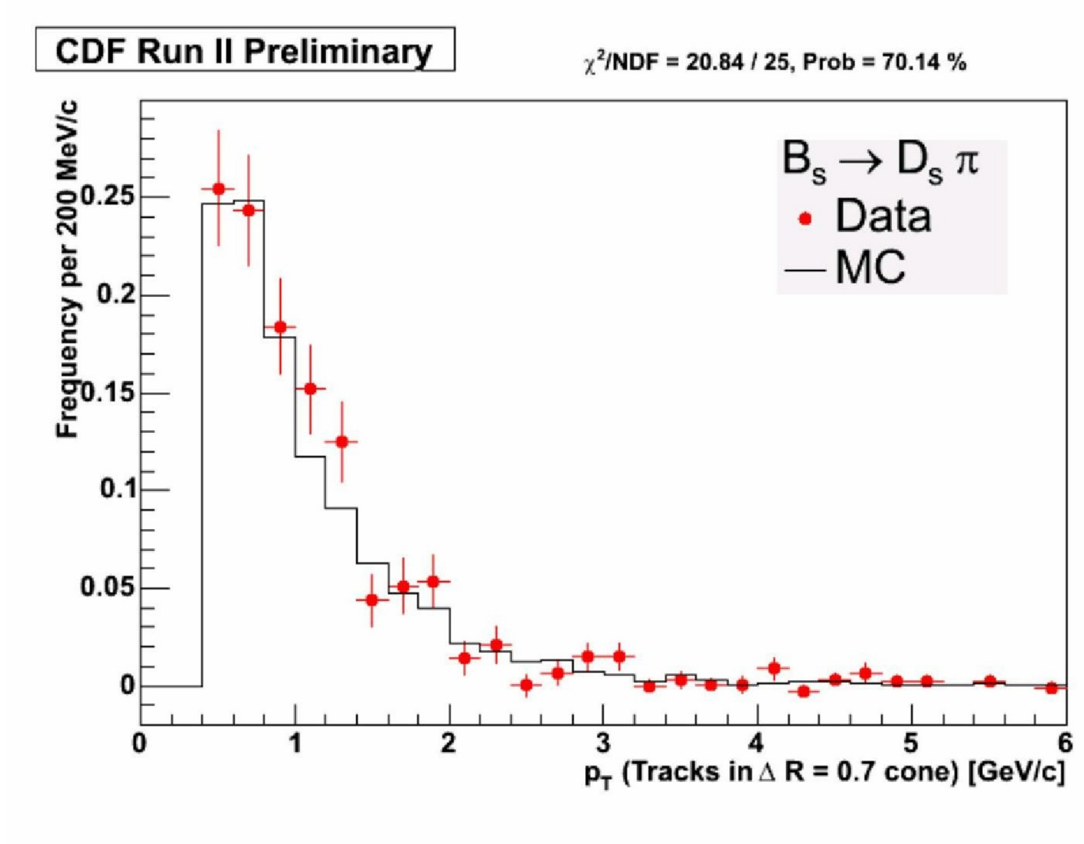
simultaneously measure

- tagger performance
- delta md

Ivan K Furic, 3/14/2005

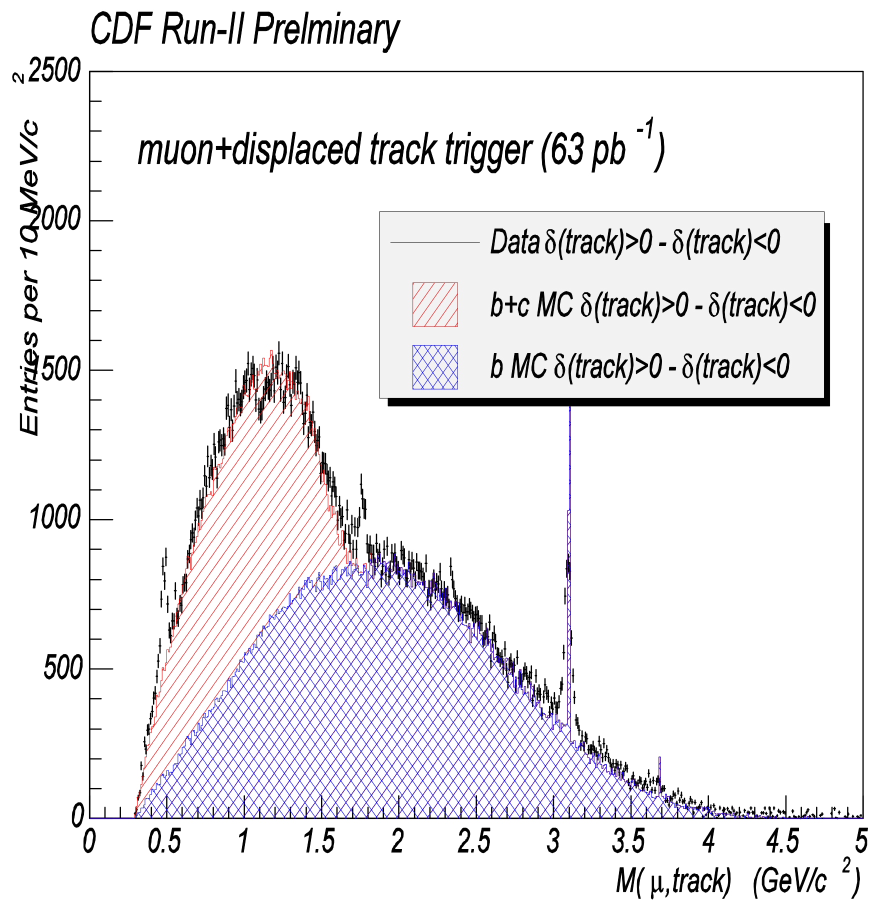
Kaon Tagging

- no straight way to determine tagger dilution from data unless B_s mixing is observed
- but we need to know the dilution to set the limit
- must use MC to measure dilution
- tune MC on B^0, B^+
- predict B_s



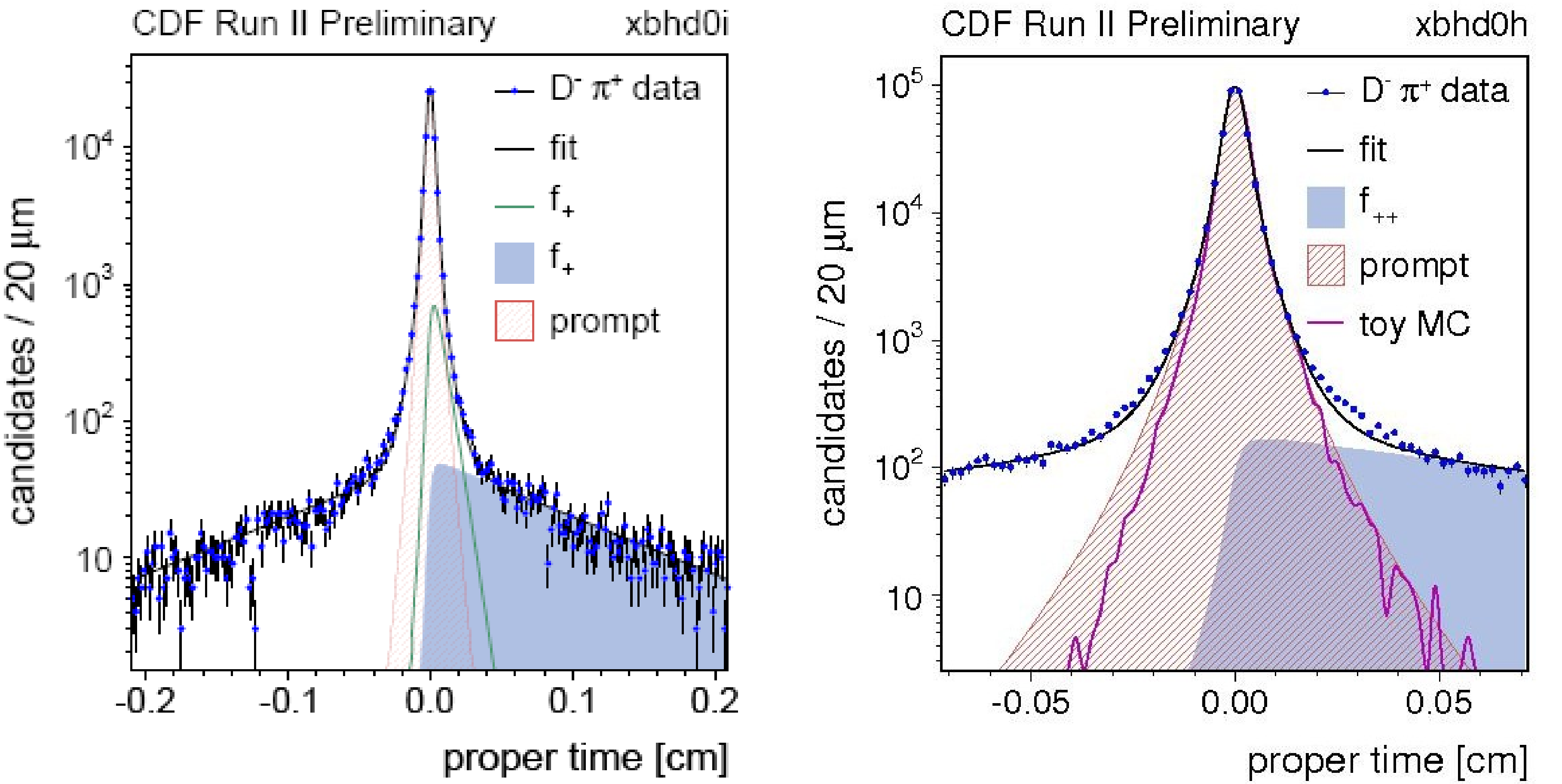
Calibrating Opposite Side Tags

- Statistical Power of the tag: εD^2
  - Tagging efficiency (ε)
  - Tagging dilution (D = 1-2w)
    - w = mistag rate
- "Binned Tagger"
  - Tag1: ε_1=50%, D_1 = 0.5
  - Tag2: ε_2=50%, D_2 = 0.1
  - $\langle D \rangle = (D_1 + D_2)/2 = 0.3$
  - $\langle D^2 \rangle = 0.36$
- Dividing events into different classes based on tagging power improves εD^2
- Calibration the tagger performance requires high statistics



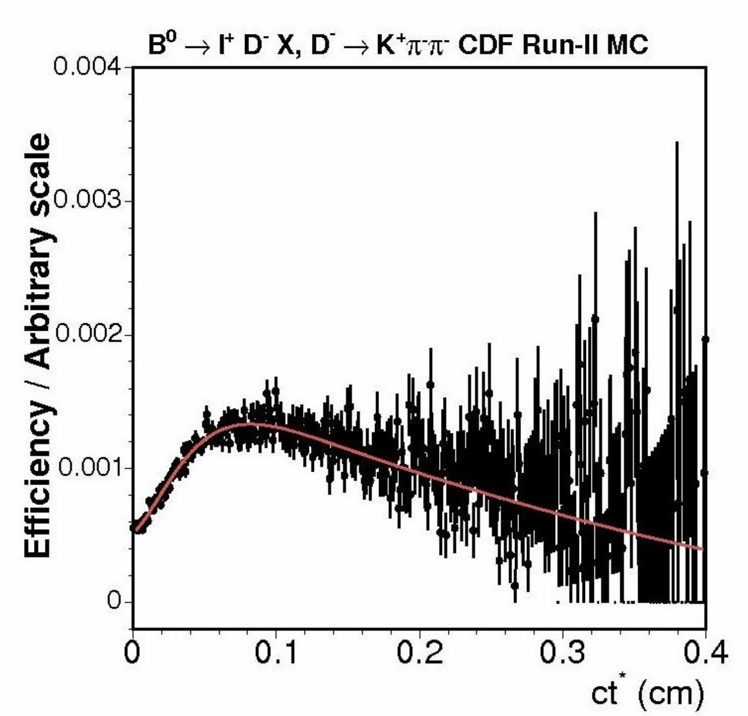
- inclusive B →track+lepton
- 1.4 M events of flavor specific B

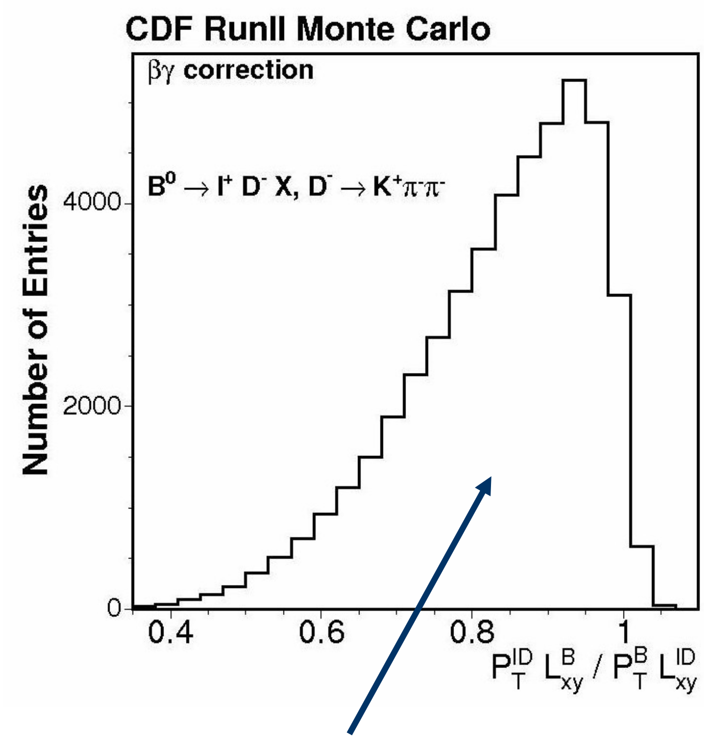
Non-Gaussian Tails



- amplitude corrected for effects of non-Gaussian tails
- correction derived from toy Monte Carlo, tuned to reproduce data

Lifetime Measurement: Semileptonic Subsample

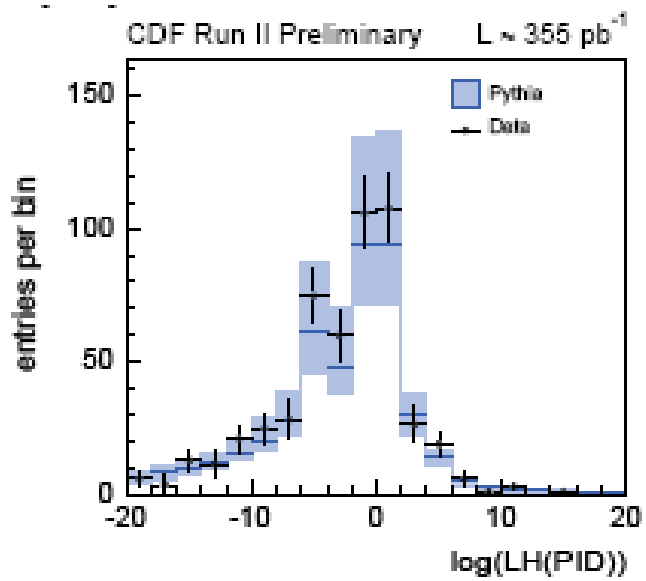




- in addition to SVT bias, correct for missing energy (K-factor)
- bin K-factor in l+D invariant mass to obtain narrow K-factor distributions

Calibrating SSKT (1)

- use combined PID likelihood, select most “kaon-like” track as tagging track
- parametrize dilution based on maximum PID likelihood value
- verify kinematic distributions (p_T, tagging track p_T, multiplicity, isolation) of light B mesons in Pythia simulation
- verify particle ID simulation
- test for dependences on:
  - fragmentation model
  - bb production mechanisms
  - detector/PID resolution
  - multiple interactions
  - pid content around B meson
  - data/MC agreement



- Final test: cross-check tagging power against high statistics light B decays

The Method

- We are looking for a periodic signal: Fourier space is the natural tool
  - Moser and Roussarie already mentioned this!
  - They use it to derive the most useful properties of A-scan
  - Amplitude approach is approximately equivalent to the Fourier transform

Amplitude from scan $\leftrightarrow$ Re[Fourier]

- Aim: move to Fourier transform based analysis
  - Computationally lighter
  - As powerful as A-scan
  - As is, no need *in principle* for measurements of D, ε etc. (however these ingredients add information and tighten the limit)
  - Will provide an alternate path to the A-scan result!

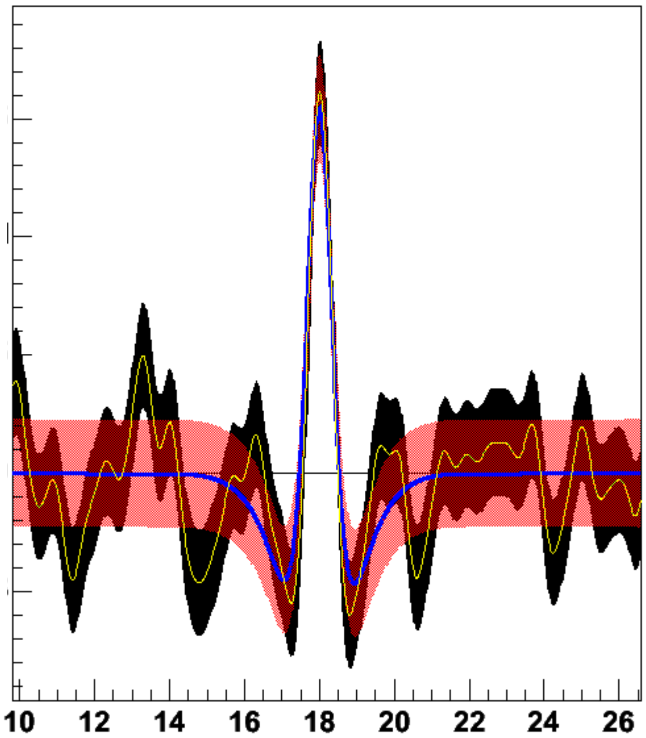
Dilution weighted transform

- Discrete Fourier transform definition
  - Given N measurements $\{t_j\}$ → $g(\omega) = \sum_{k=1}^{N} D_k e^{-i\omega t_k}$
- Properties:
  - A particular application of $g(\omega) = \sum_{k=1}^{N} w_k e^{-i\omega t_k}$ (CDF8054)
  - Average: $\langle g(\omega) \rangle = N \langle D \rangle \tilde{f}(\omega)$

    (f(t) is the parent distribution of $\{t_j\}$)
  - Corresponds to dilution-weighted Likelihood approach
  - Errors computed from data: $\sigma^2(\mathrm{Re}\, g(\omega)) \approx \frac{N}{2}\left(\langle D^2 \rangle + o\left(\frac{1}{N}\right)\right)$
    - NB: Errors can be calculated directly from the data!
    - $\Delta(\omega) \equiv g_{\mathrm{Un}Mix}(\omega) - g_{Mix}(\omega)$ behaves "as you'd expect"
    - While Δ and its uncertainty are fully data-driven, predicted Δ requires exactly the same ingredients as the amplitude scan fit

Properties of Δ…

- Re[Δ]
  a) contains information equivalent to the standard amplitude scan
  b) (Amplitude scan)≈Re[Δ]
- Re[F] and $\sigma_{Re[F]}$ can be computed directly from data!
- b) ⇒ Sensitivity is exactly:



$$\frac{\Delta(\omega=\Delta m_s)}{\sigma_\Delta}=\sqrt{N\varepsilon\langle D\rangle^2}\sqrt{\frac{S}{S+B}}e^{-\Delta m^2\sigma_{ct}^2/2}\sqrt{1+\frac{\sigma_D^2}{\langle D^2\rangle}}$$

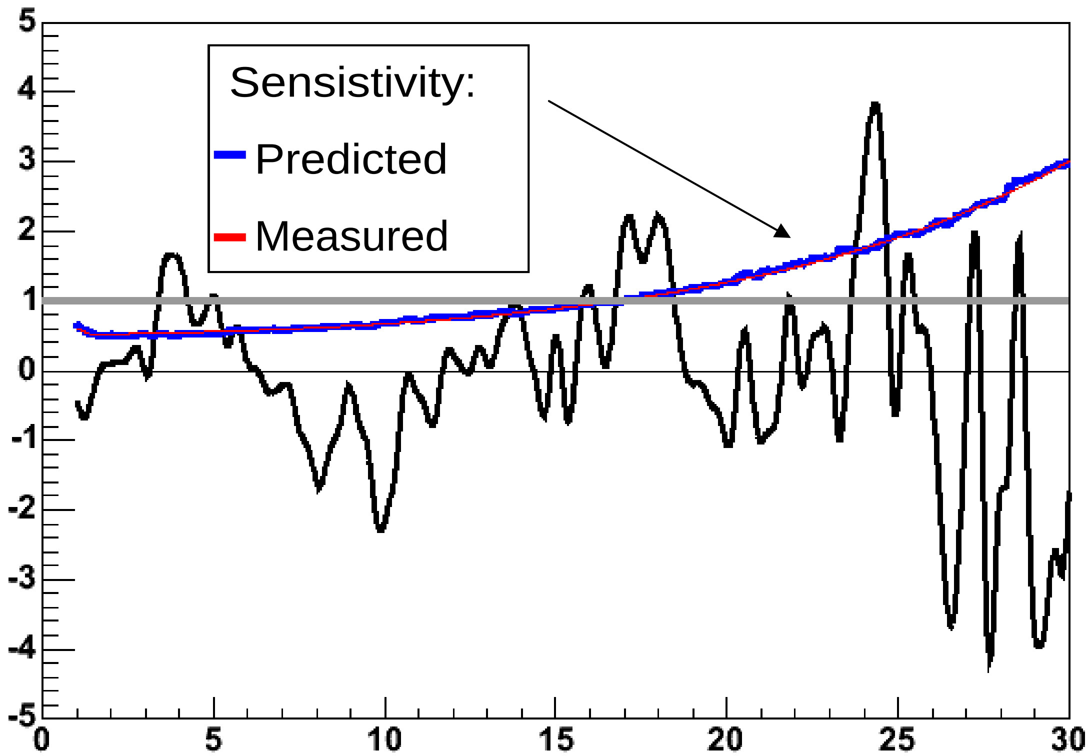
Can we reproduce the A-scan itself?

Toy Example

- 1000 toy events
- Δm_s=18
- S/B=2.
- εD_{signal}^2=1.6%
- εD_{back}^2=0.4%
- Background and signal parameterized according to standard analyses
- Histogrammed σ_{ct}
- Best knowledge on SF parameterization

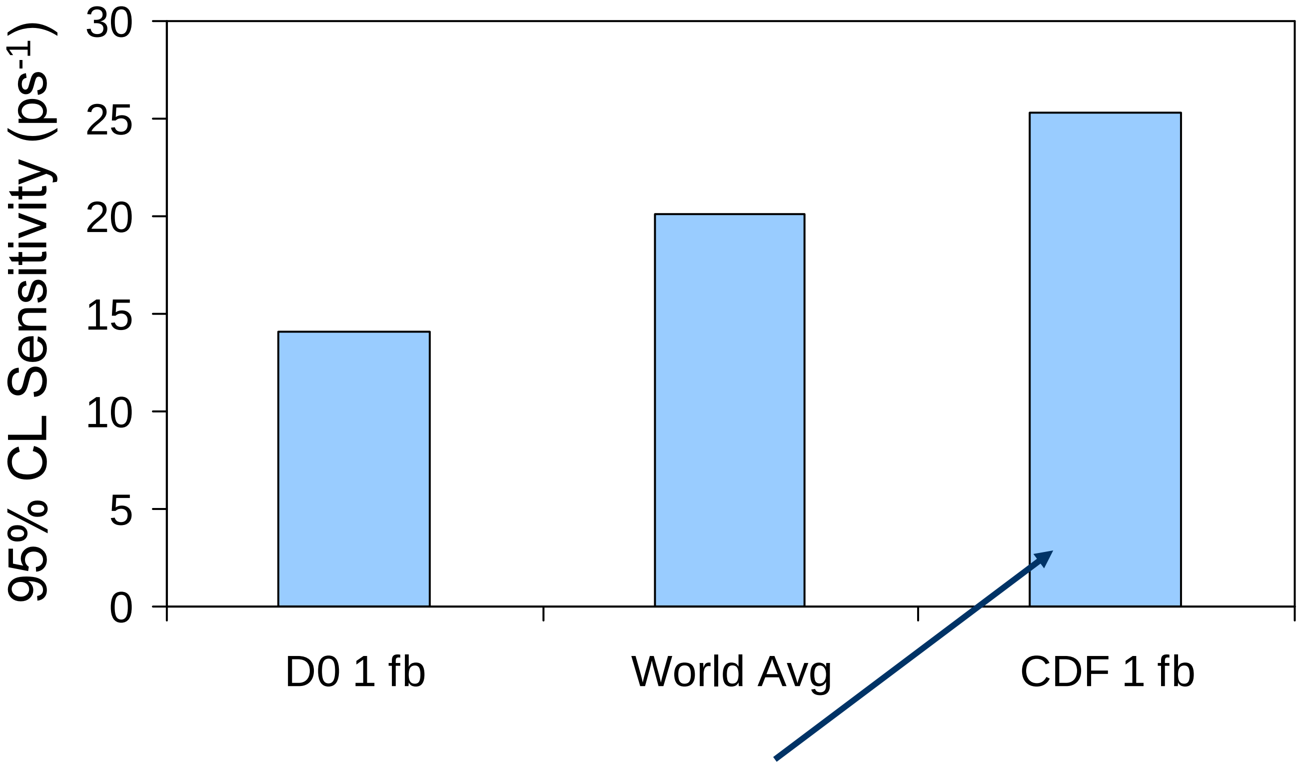
" A-scan" a\` la fourier

$$\frac{\Delta(\omega)}{pred.\Delta(\omega;\, \Delta m_s = \omega)}$$



No actual fit involved: this method allows to flexibly study systematics!

Measurement Sensitivity



- estimated from scan on "blinded" data (randomized tags)
- unusual situation – one single measurement more sensitive than the world average knowledge!