Observation of $B_s^0 - \overline{B_s^0}$ Oscillations

A. Belloni
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

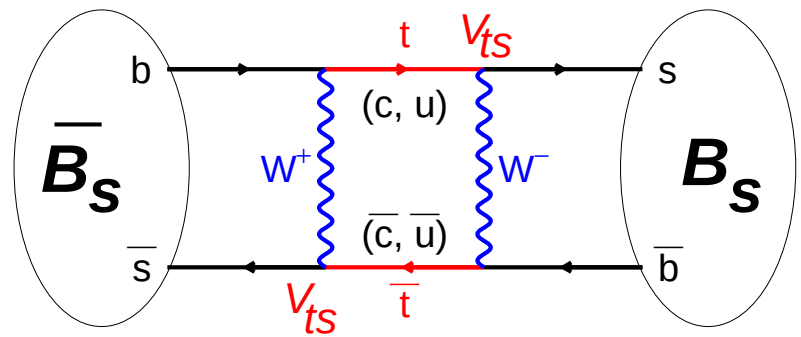
On behalf of the CDF Collaboration

Beauty 2006, Oxford UK
September 25-29, 2006

Theoretical predictions for Δm

In SM B^0/B_s mix through box diagram:

$$\Delta m_q \propto m_{B_q} \hat{B}_{B_q} f_{B_q}^2 \left| V_{tb} V_{tq}^* \right|^2$$
$$q = s, d$$

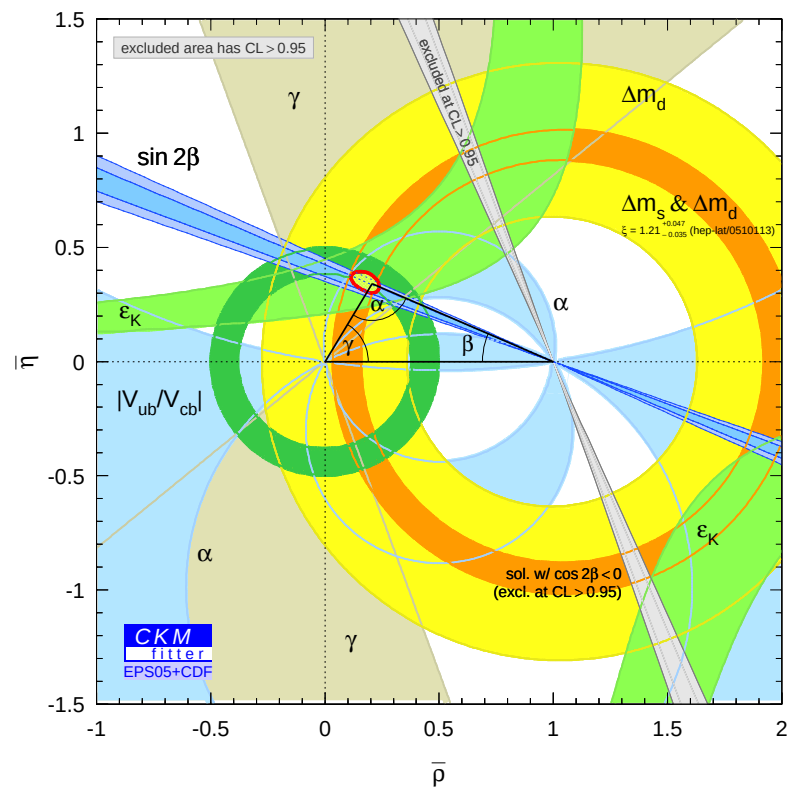


Most uncertainties cancel in ratio:

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

with $\xi = 1.21^{+0.047}_{-0.035}$

(Okamoto, Lattice 2005)



+ measure $\frac{\Delta m_s}{\Delta m_d} \Rightarrow$ find $\frac{|V_{ts}|}{|V_{td}|}$ to ~ 3-4%
+ Δm_d measured to high precision:

$$\Delta m_d = 0.507 \pm 0.005 \text{ ps}^{-1}$$

+ Phys. Rev. Lett. **97**, 062003 (2006):

$$\Delta m_s = 17.31^{+0.33}_{-0.18} \pm 0.07 \text{ ps}^{-1}$$

+ Signif. of Δm_s measurement: 3σ

A bit of history: a 20-year effort

1987

- First evidence of B^0 mixing from UA1 C. Albajar *et al.*, PLB **186**, 247 (1987)
- Argus observes B^0 mixing $\Rightarrow$ UA1 result points to large B_s mixing ($m_t > 50$ GeV/c^2) H. Albrecht *et al.*, PLB **192**, 245 (1987)

1989

- CLEO confirms Argus result M. Artuso *et al.*, PRL **62**, 2233 (1989)

1990s

- Inclusive measurements of *B* mixing from LEP establish B_s mixing D. Buskulic *et al.*, PLB **313**, 498 (1993)

1993

- First time dependent measurement of Δm_d from Aleph
- First lower limit on Δm_s from Aleph: $\Delta m_s > 12 \cdot 10^{-4}$ eV/c^2 CDF 06: 0.011 eV/c^2 D. Buskulic *et al.*, PLB **322**, 441 (1994)

1999

- CDF Run I result on Δm_s: $\Delta m_s > 5.8$ ps^{-1} F. Abe *et al.*, PRL **82**, 3576 (1999)

2000s

- D0 first result on Δm_s: $\Delta m_s > 5.0$ ps^{-1}
- CDF Run II first result on Δm_s: $\Delta m_s > 7.9$ ps^{-1}
- D0 reports interval: $\Delta m_s \in [17, 21]$ ps^{-1} at 90% CL PRL **97**, 021802 (2006)
- CDF Run II first measurement: $\Delta m_s = 17.31^{+0.33}_{-0.18} \pm 0.07$ ps^{-1} PRL **97**, 062003 (2006)

Δm_s *Search at CDF*

1999

- Run I: $\Delta m_s > 5.8\ \mathrm{ps}^{-1}$, sensitivity: $5.1\ \mathrm{ps}^{-1}$ F. Abe *et al.*, PRL **82**, 3576 (1999)

Feb 2005

- First Run II: $\Delta m_s > 7.9\ \mathrm{ps}^{-1}$, sensitivity: $8.4\ \mathrm{ps}^{-1}$

Oct 2005

- $\Delta m_s > 8.6\ \mathrm{ps}^{-1}$, sensitivity: $13.0\ \mathrm{ps}^{-1}$

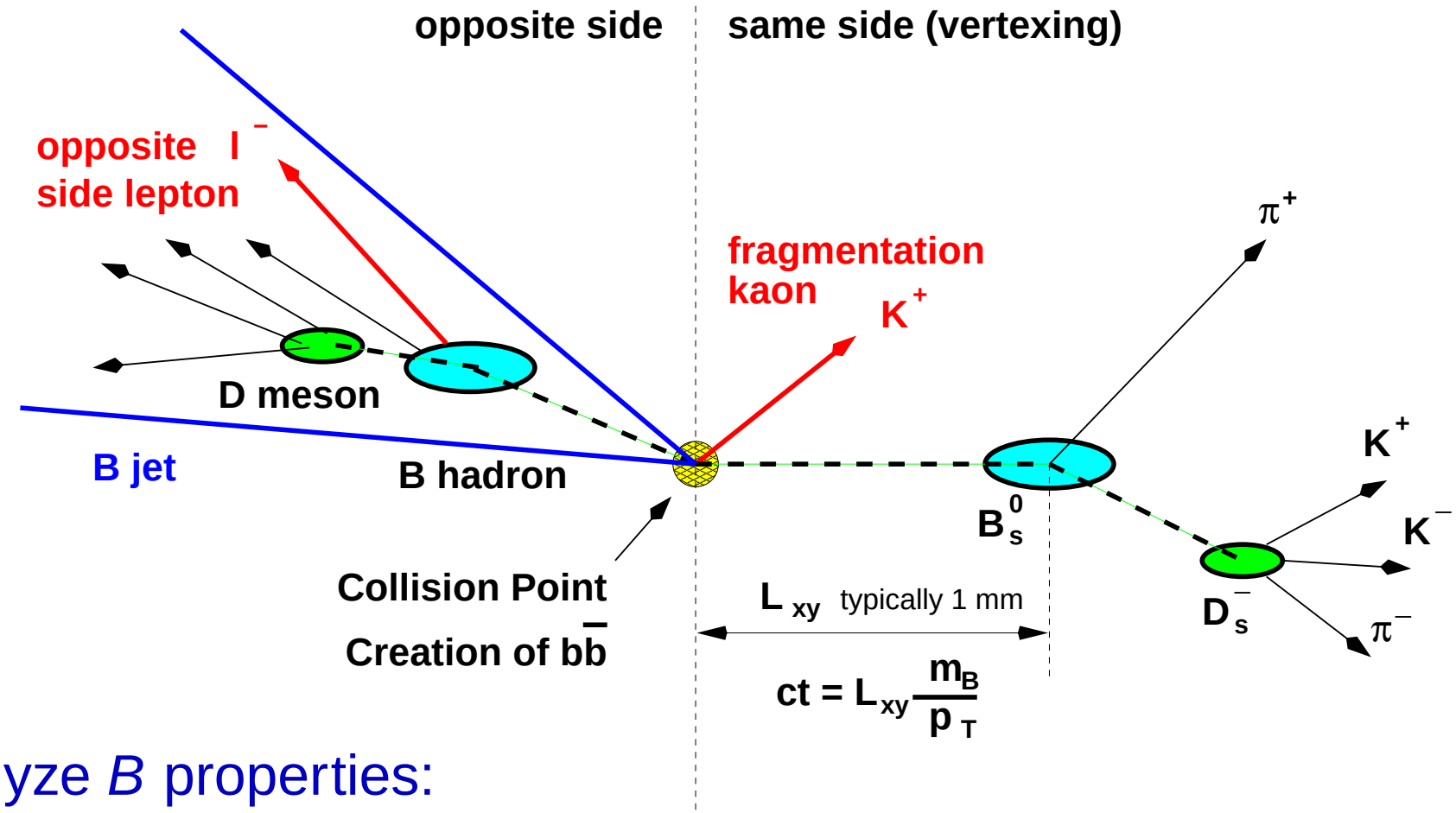
Apr 2006

- First measurement: $\Delta m_s = 17.31^{+0.33}_{-0.18} \pm 0.07\ \mathrm{ps}^{-1}$, sensitivity: $25.8\ \mathrm{ps}^{-1}$

PRL **97**, 062003 (2006)

Significance not sufficient to claim DEFINITIVE observation!

Understanding Single B_s^0 Meson Candidate



Analyze *B* properties:

- \+ flavor at decay: from final state
- \+ proper decay length *ct*: in *B* rest frame
- \+ flavor at production: from flavor tagging

Fourier analysis

Two domains for oscillation fit

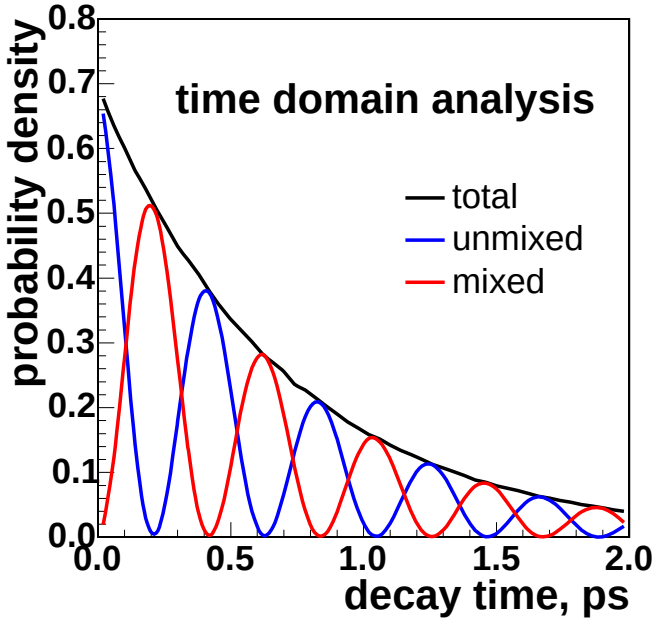
Time domain:

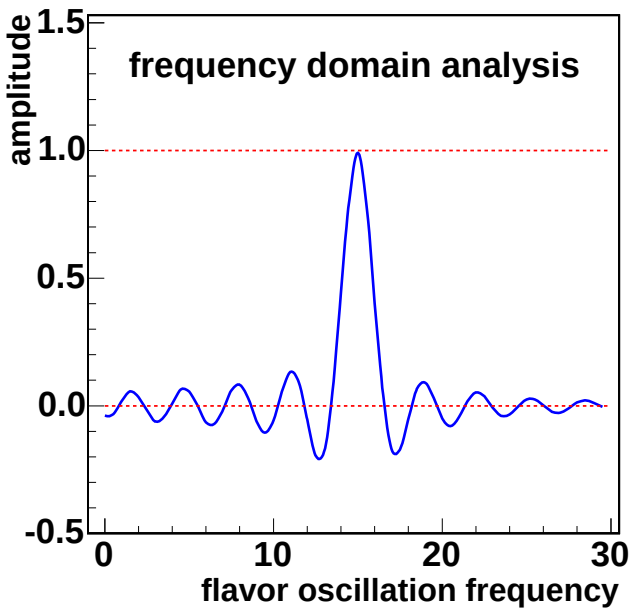
+ fit for Δm_s in $P(t) \sim (1 \pm D \cos \Delta m_s t)$

  dilution $D = 1 - 2w$, w = mistag prob.

Frequency domain: amplitude scan

+ introduce amplitude:
  $P(t) \sim (1 \pm \mathcal{A} D \cos \Delta m_s t)$
+ fit $\mathcal{A}$ for different Δm_s and obtain frequency spectrum
+ true $\Delta m_s \Rightarrow \mathcal{A} = 1$, else $\mathcal{A} = 0$
+ traditionally used for B_s^0 mixing search
+ easy to combine experiments





Significance of observation

Statistical power of analysis:

$$\frac{1}{\sigma_{\mathcal{A}}} = \sqrt{\frac{S}{S+B}} \cdot e^{-\frac{\Delta m_S^2 \sigma_{ct}^2}{2}} \cdot \sqrt{S\frac{\varepsilon D^2}{2}}$$

Signal purity & statistics

- Hadronic modes: $B_s \to D_s(3)\pi$, $D_s^*\pi$, $D_s\rho$
- Semileptonic modes: $B_s \to D_s\ell\nu X$

Proper decay time resolution

- Fully reconstructed modes provide better accuracy

Tagging power

- Effective statistics scale with εD^2
  - efficiency ε = probability to have a tag
  - dilution $D = 1 - 2 \cdot w$, w: mistag probability

CDF II Detector — Critical components for Δm_s

Reconstruct B_s signal: low background, high statistics

- + Trigger system: displaced Track Trigger A. Annovi, Th. S1
- + Central Outer Tracker (COT): momentum resolution
- + Silicon tracker (SVX/ISL/L00): vertex identification
- + Muon Chambers and Calorimeters: lepton identification

Proper decay time resolution:

- + L00 (mounted on beam pipe): 10-20% improvement in the impact parameter resolution of tracks

b-flavor tagging: separation of kaons versus pions

- + dE/dx information from the COT
- + Time-Of-Flight detector:
  - + significantly improves Same Side Kaon Tagging

B_s reconstruction

Golden hadronic sample

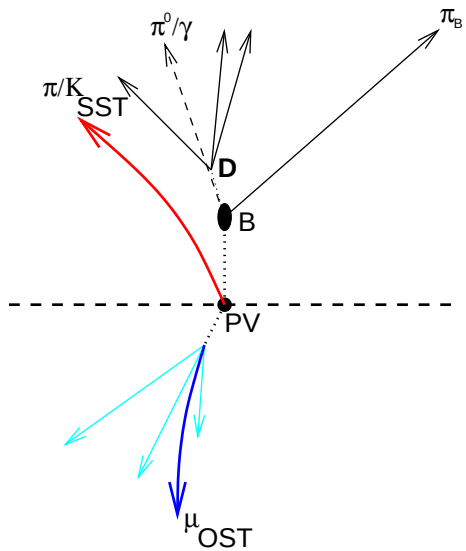
Fully reconstructed peak

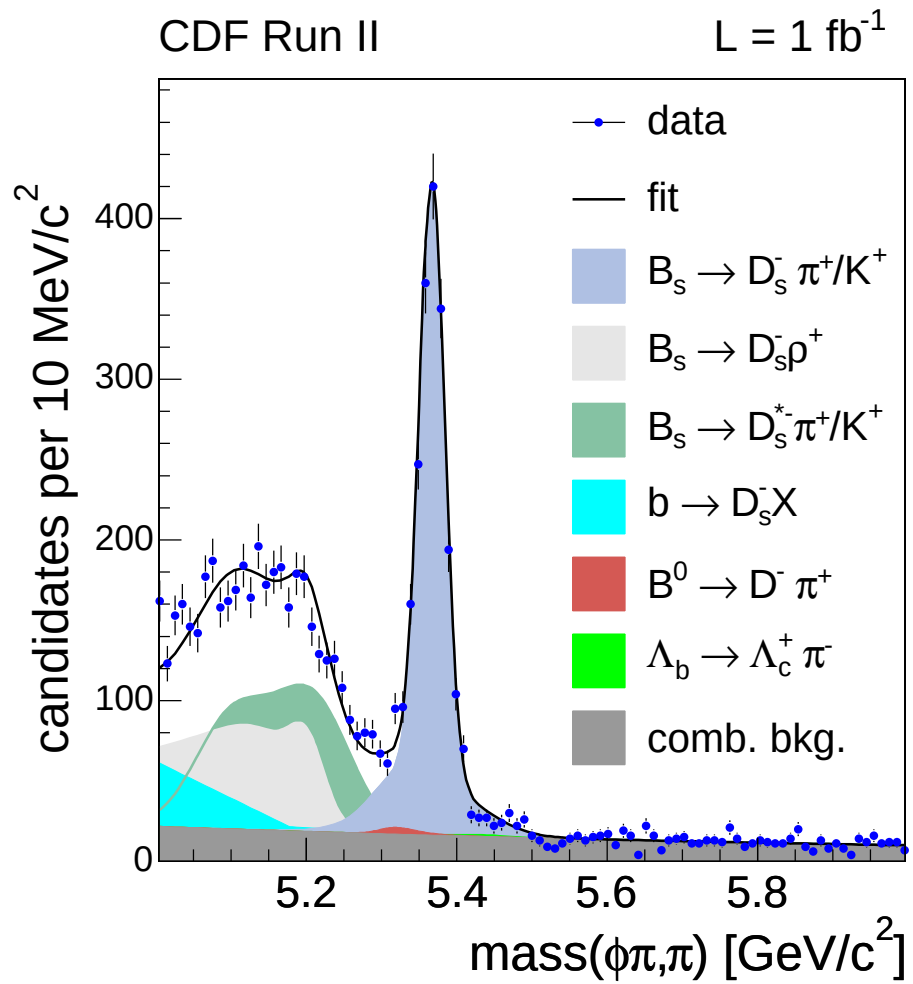
+ 2000 $B_s \to D_s^- \pi^+, D_s^- \to \phi \pi^-$ candidates

New!

Partially reconstructed signals

+ $\approx$ *double* signal yield
+ soft γ or π^0 lost, fraction of p_T reconstructed $\sim$ 96%
+ nearly fully reconstructed, critical attribute for σ_{ct}





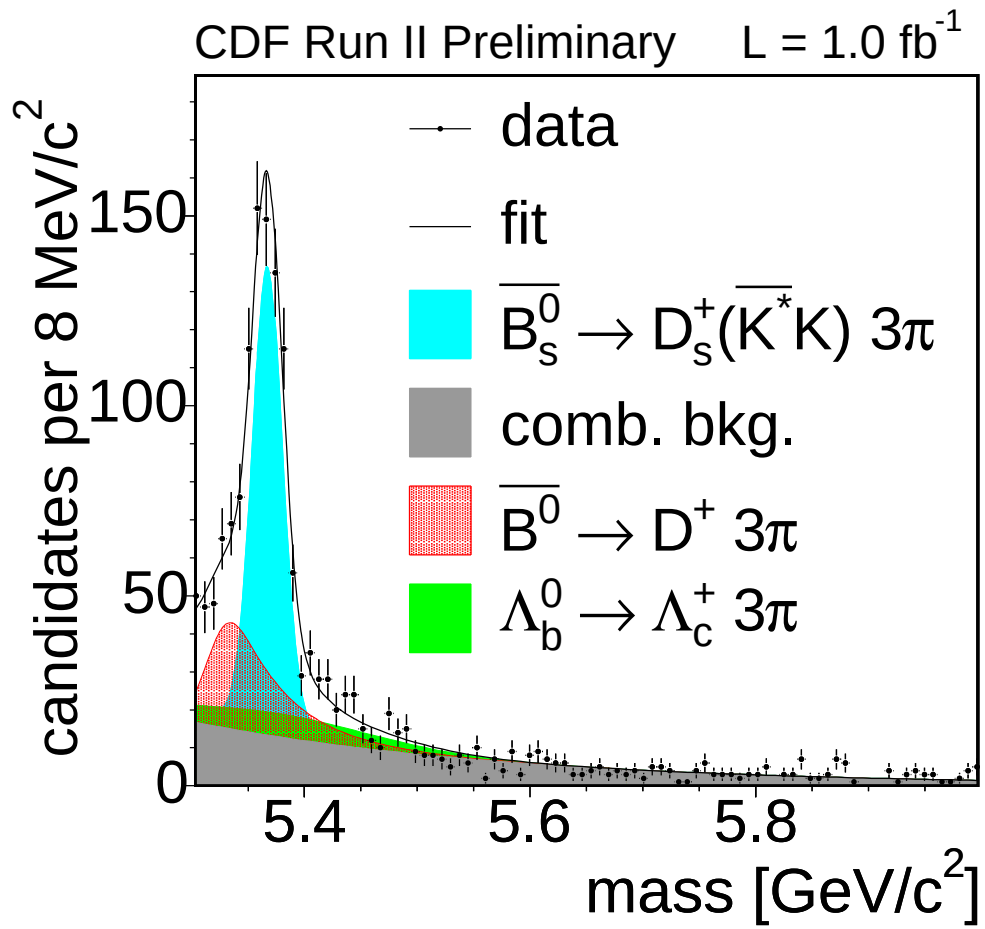
Why "golden mode"?

+ Few tracks (4), narrow ϕ

$\Longrightarrow$ Low comb. bkg; Great $\mathcal{S}/\sqrt{\mathcal{S}+\mathcal{B}}$

Other fully hadronic modes

B_s Decay	Signal	April
$D_s(\phi\pi)\pi$	2000	1600
Partially rec.	3100	—
$D_s(K^*K)\pi$	1400	800
$D_s(3\pi)\pi$	700	600
$D_s(\phi\pi)3\pi$	700	500
$D_s(K^*K)3\pi$	600	200
$D_s(3\pi)3\pi$	200	—
Total	8700	3600



Neural Network Selection

- \+ Improved background rej.
- \+ PID used for B^0 suppression in $B_s \to D_s(K^*K)(3)\pi$

Semileptonic Decays

Introduced particle-id into selection

- + Reduce D^+ reflection in K*K
- + Combinatorial background dominated by pions

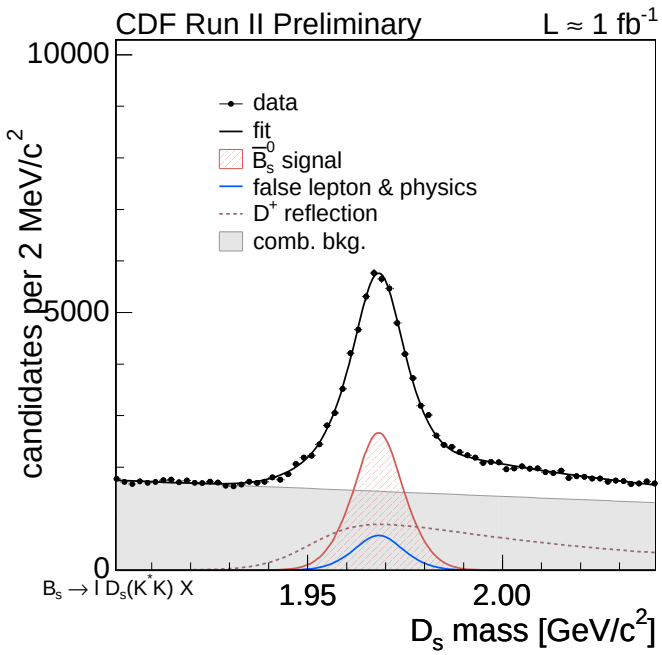
New trigger paths added

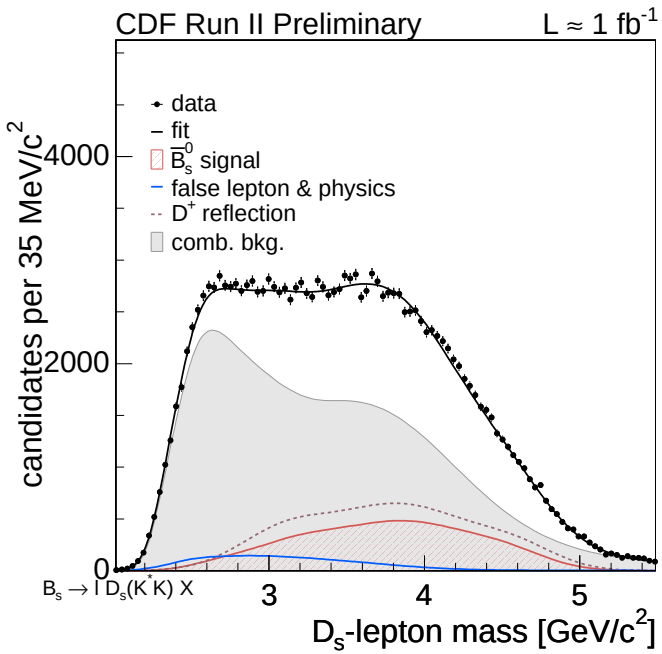
Signal yield 61,500 (was 37,000)

- + S/B improves 100% in $\phi\pi$ and K^*K

Sensitivity driven by high ℓD_s mass candidates

- + Exploited as fit input





Proper decay time

Proper Decay Time Reconstruction

Fully reconstructed decays:

$$+ \quad ct = L_{xy}^{B} \frac{m_B}{p_T^B}$$

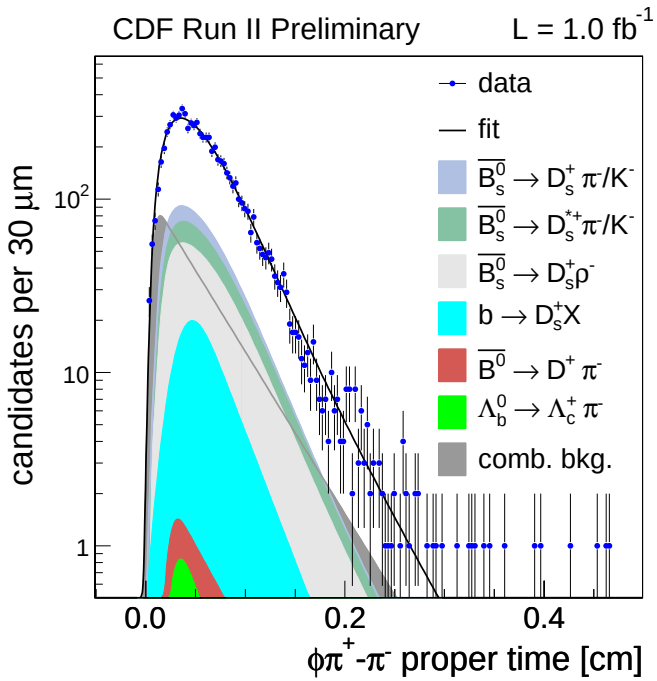
Semileptonic & partially reconstructed decays:

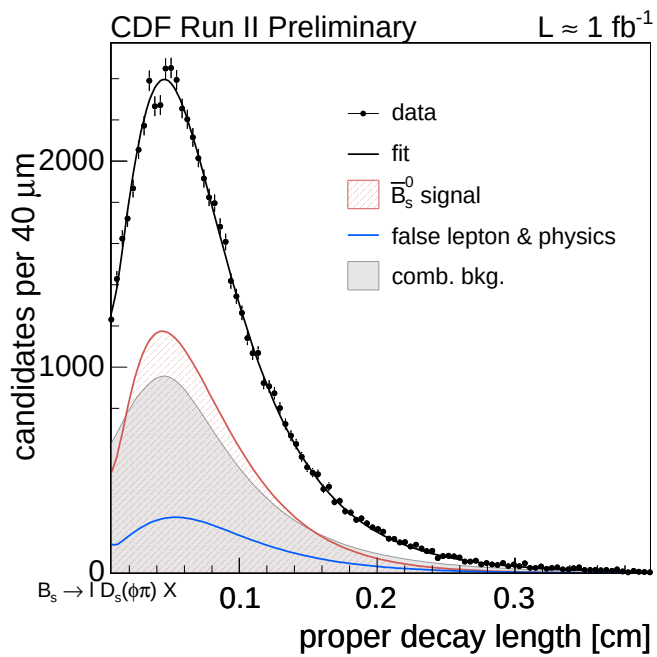
$$+ \quad ct = L_{xy}^{\ell D} \frac{m_B}{p_T^{\ell D}} \cdot k$$

$$k = \left\langle \frac{p_T^{\ell D}}{p_T^B} \frac{L_{xy}^B}{L_{xy}^{\ell D}} \right\rangle$$ from MC

Model from well studied $B^{+/0}$ semileptonic decays

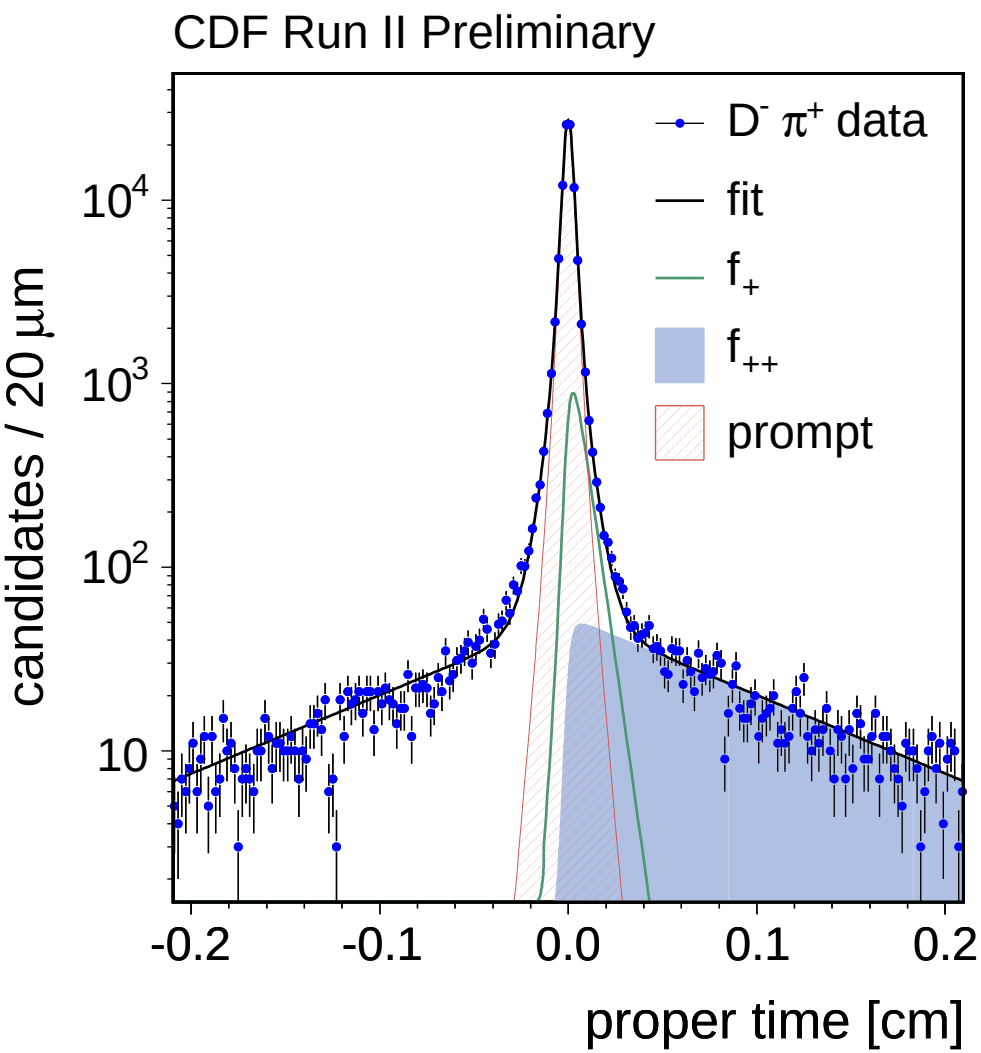
Lifetime measurements consistent with PDG





Proper Decay Time Resolution (I)

Critical aspect of the analysis, limiting factor at high Δm_s



Need reference for measuring resolution

- \+ PV natural choice

Use prompt D^+ and track

- \+ large sample of prompt D^+
- \+ most tracks from PV
- \+ same topology as signal
- \+ measure of *ct* resolution

Parametrize σ_{ct} with kinematic of candidate

Calibrated on our data

Proper Decay Time Resolution (II)

$$\sigma_{ct} = \sigma_{ct}^0 \oplus ct \cdot \frac{\sigma_p}{p}$$

Fully rec. $B_s \to D_s(3)\pi$:

- $\sigma_{ct}^0 \sim 26\ \mu m$ (87 fs)
- $\sigma_p/p < 1\%$

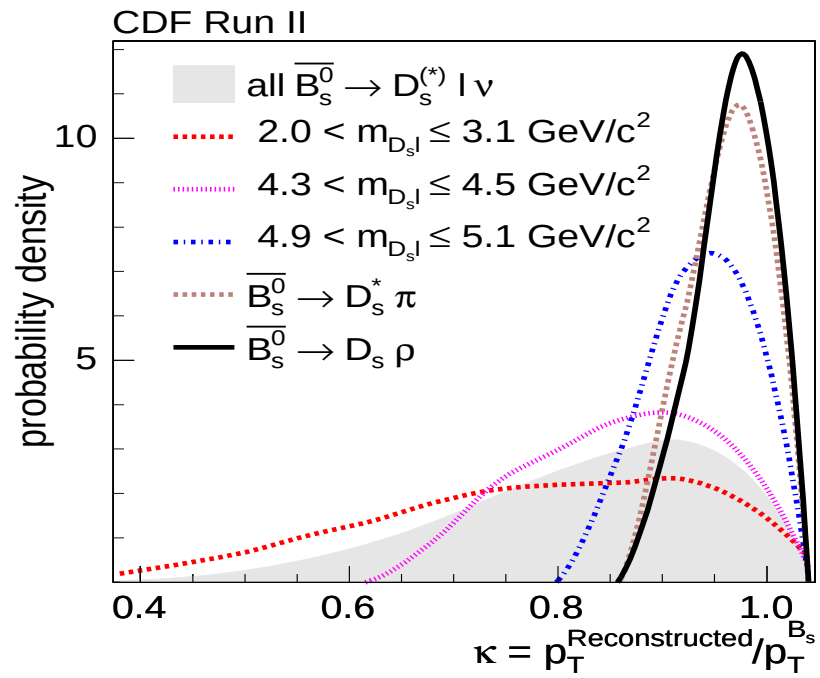
Partially rec. $B_s \to D_s^* \pi, D_s \rho$:

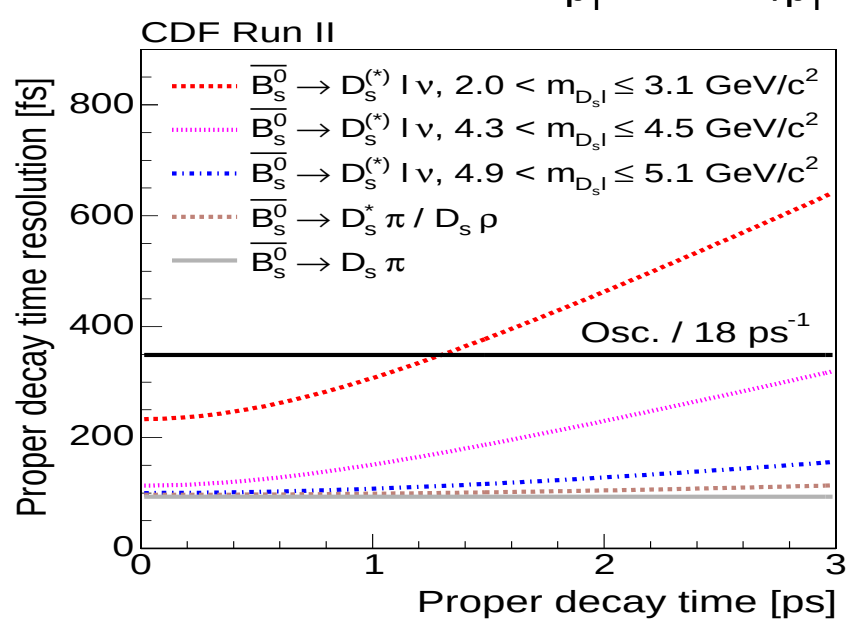
- $\sigma_{ct}^0 \sim 29\ \mu m$ (97 fs)
- $\sigma_p/p \sim 2\%$

Semileptonic $B_s \to D_s \ell X$:

- $\sigma_{ct}^0 \sim 30\text{-}70\ \mu m$ (100-230 fs)
- $\sigma_p/p \sim 3\text{-}20\%$

Variation of k-factor with $m(\ell D)$ significantly improves decay time resolution in semileptonic





Flavor tagging

b-Flavor Tagging – Opposite Side Tag

Soft Lepton tagger

- + look for $B \to \ell X$ decays on OS
- + lepton charge indicates *b*-flavor

  Mistag due to $\bar{b} \to \bar{c} \to \ell^- X$ and oscillation

Jet Charge tagger

- + look for jet or secondary vertex on OS
- + jet charge indicates *b*-flavor

Opposite Side Kaon tagger New!

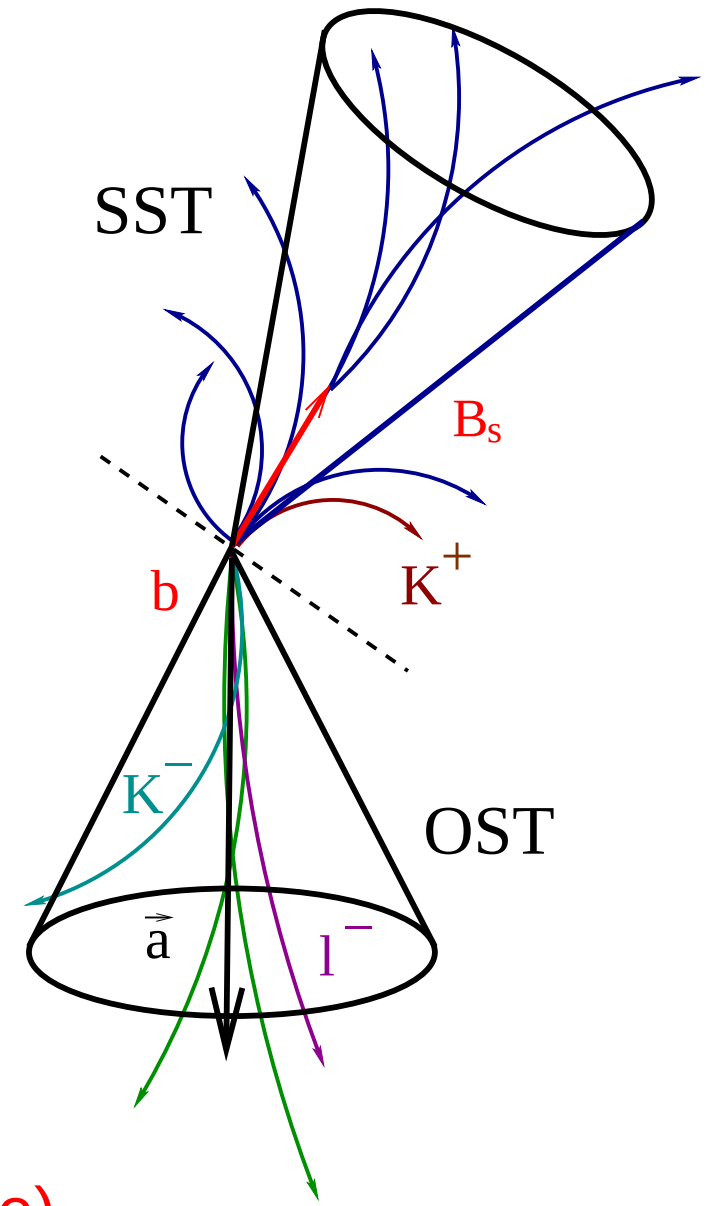
- + look for strange decays on OS
- + kaon charge indicates *b*-flavor

Neural Network combines taggers New!

OST's perform identically in $B_{u,d,s}$:

calibrated in high statistics B^+/B^0 data

$\varepsilon D^2 = 1.8\%$ (+20% relative increase)



b-Flavor Tagging – Same Side Tag

Most powerful tagger available

- 2-3 times more effective than combined OST

Exploits charge correlation between *b* and fragmentation tracks

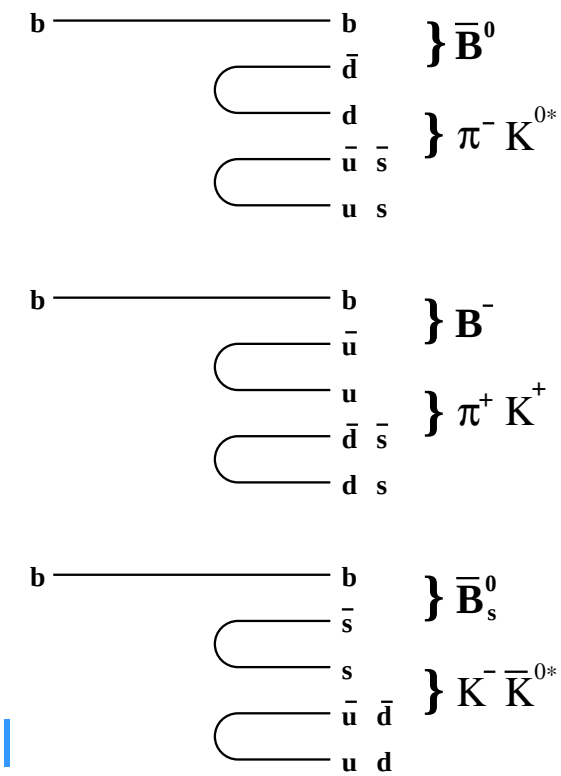
- B^-, $\overline{B^0}$ likely to have a π^+, π^- nearby
- $\overline{B_s}$ likely to have a K^-

Neural Network separates kaons and pions

- TOF and COT dE/dX crucially important
- Kinematic of candidate provide additional separating power

Unlike OST, SSKT is different for B^0, B^+ and B_s

- SST: need to rely on MC simulation
- Data and MC thoroughly compared



OST and SST combined independently

$\varepsilon D^2 = 3.5\%$ (hadr.): +0%

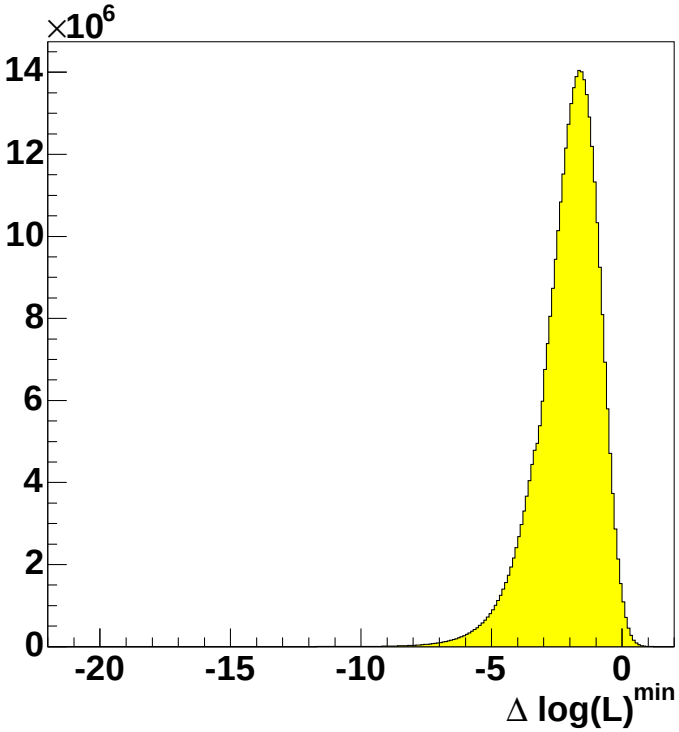
$\varepsilon D^2 = 4.8\%$ (semil.): +9%

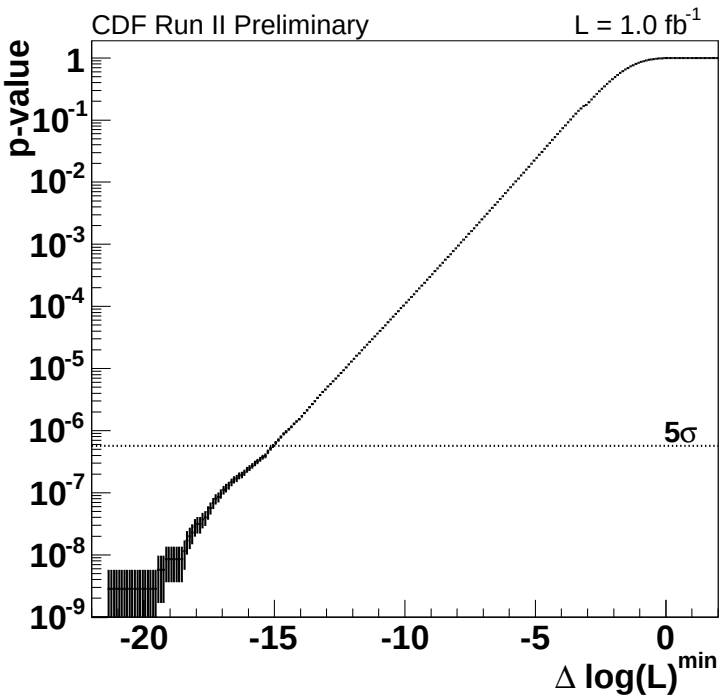
Oscillation analysis

Before looking at data

Estimate significance of oscillation signal

- p-value: probability that background fluctuation produces observed effect
- Estimate p-value using $-\Delta \log L = -\log[L(\mathcal{A}=1)/L(\mathcal{A}=0)]$: more powerful than $\mathcal{A}/\sigma_{\mathcal{A}}$
- Probability of random tag fluctuation evaluated on data
- No search windows

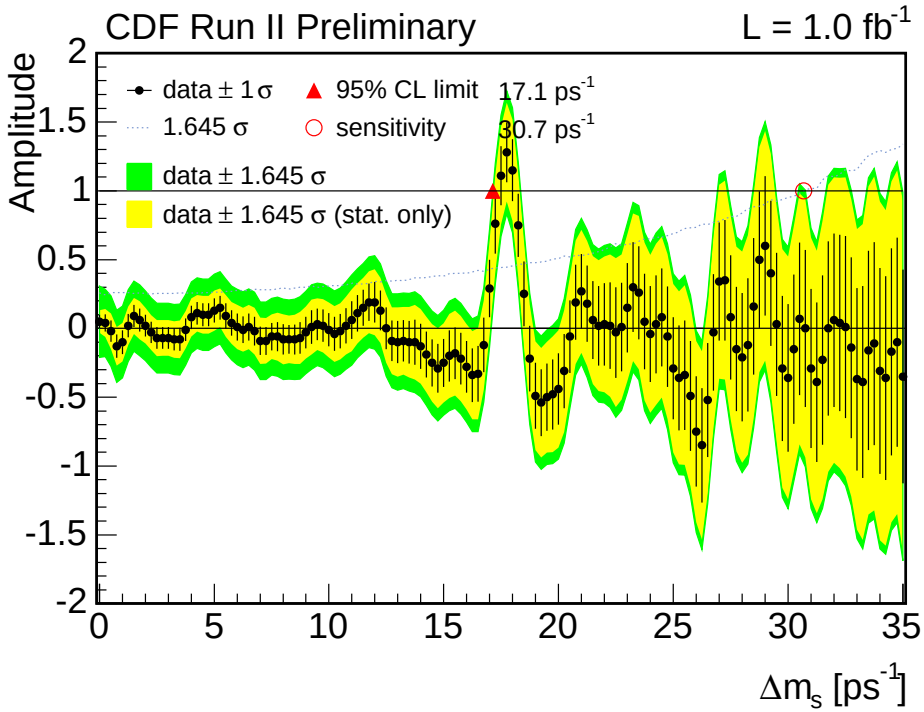




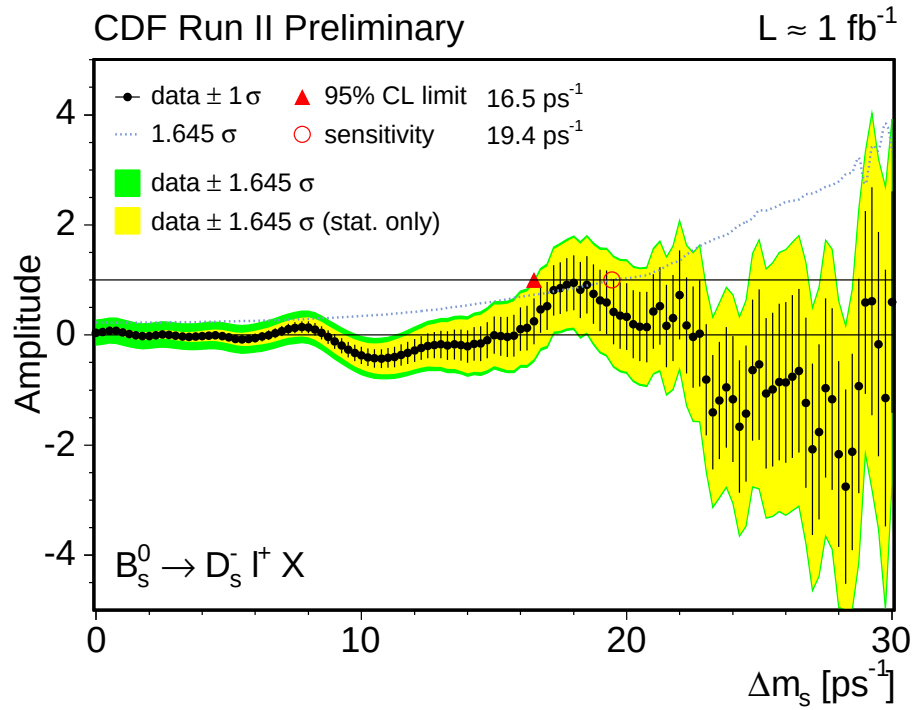
The Data

Amplitude Scans

Hadronic



Semileptonic



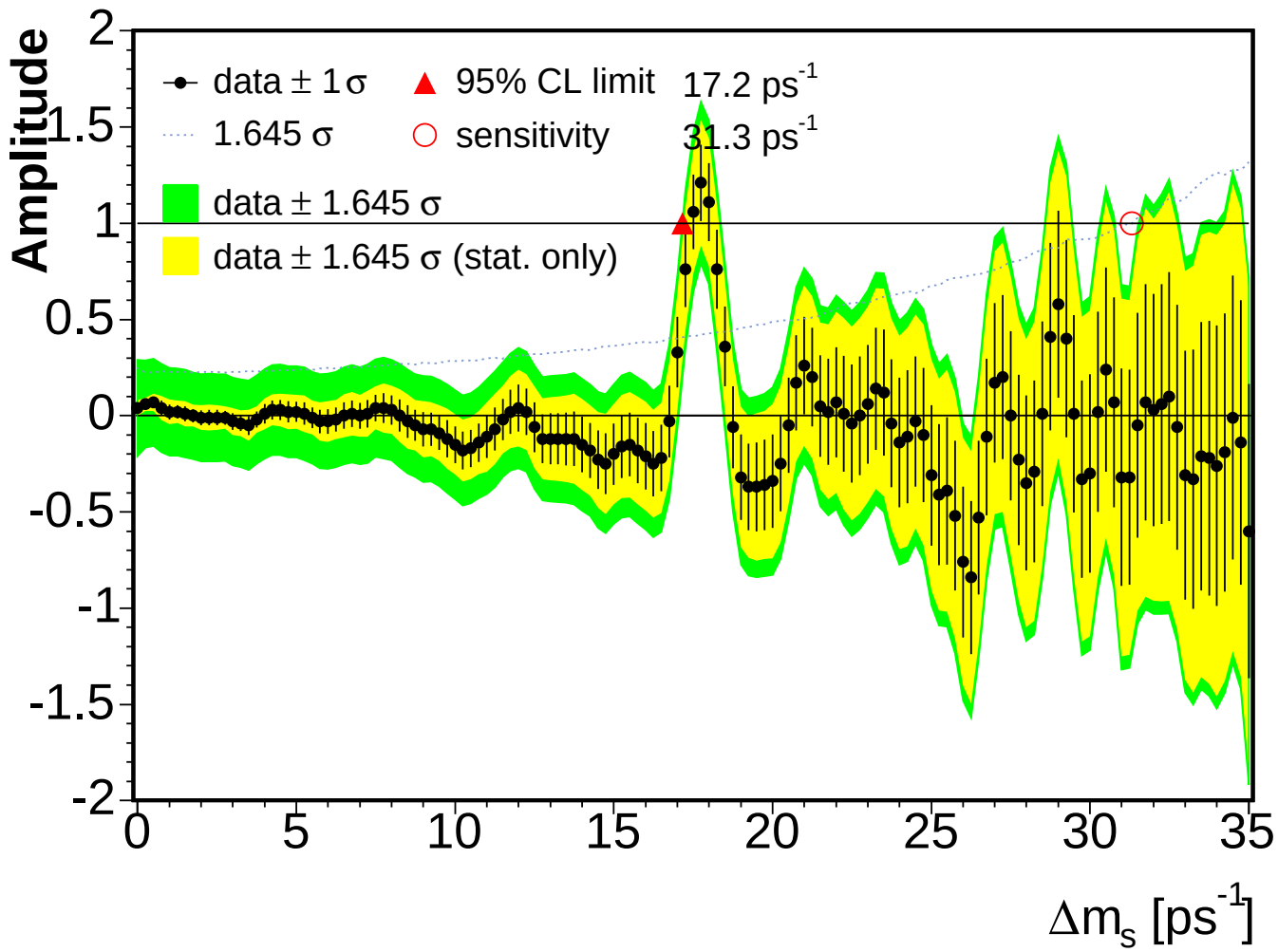
$\mathcal{A} = 1.28 \pm 0.22$

$\mathcal{A}/\sigma_{\mathcal{A}} \sim 2$

Best (hadronic) and 2nd best sensitivity!

$\mathcal{A}$ compatible with 1 for $\Delta m_s \sim 17.75$ ps^{-1}!

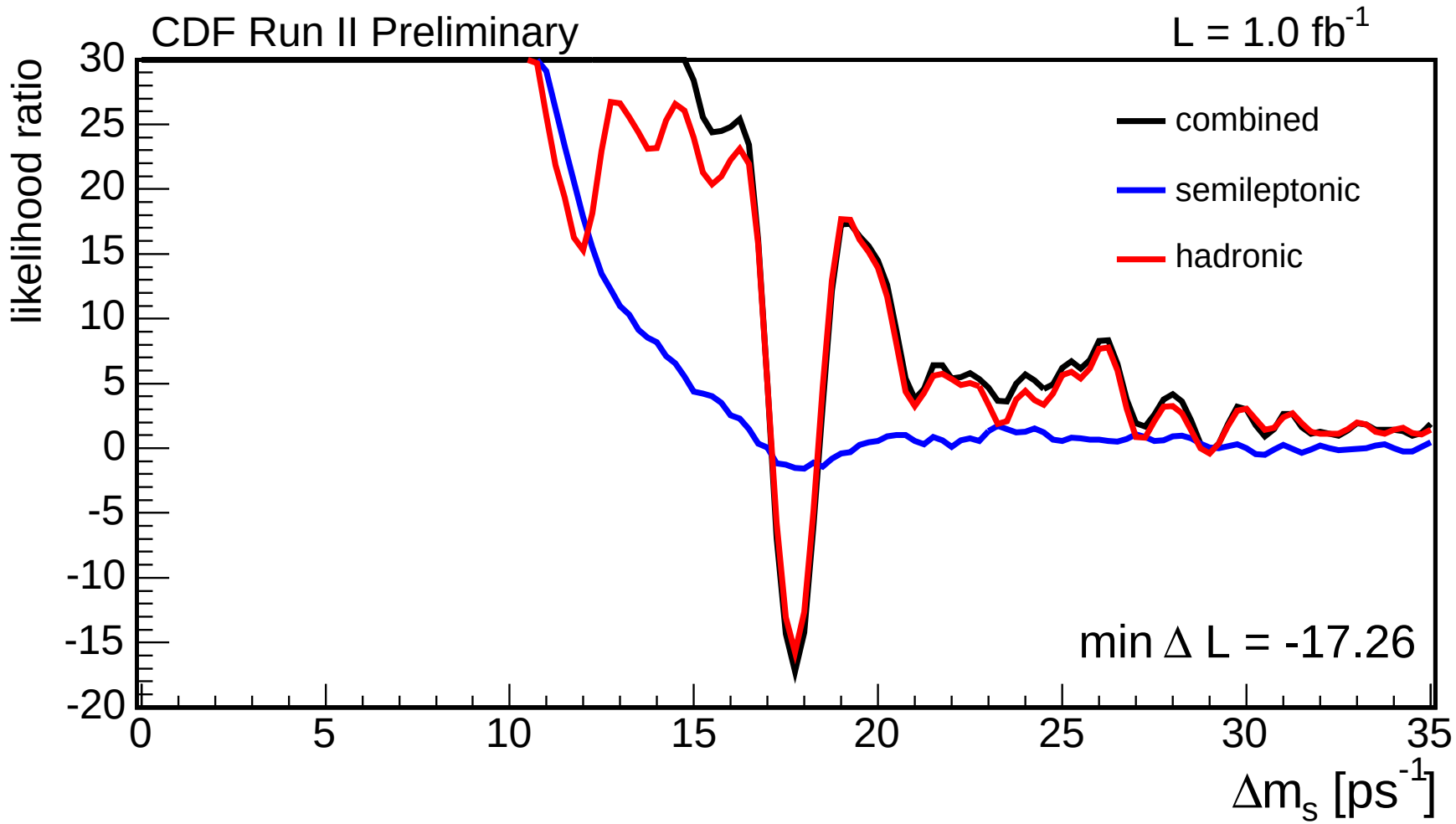
Combined Amplitude Scan



$$\mathcal{A}(\Delta m_s = 17.75\ \mathrm{ps}^{-1}) = 1.21 \pm 0.20$$

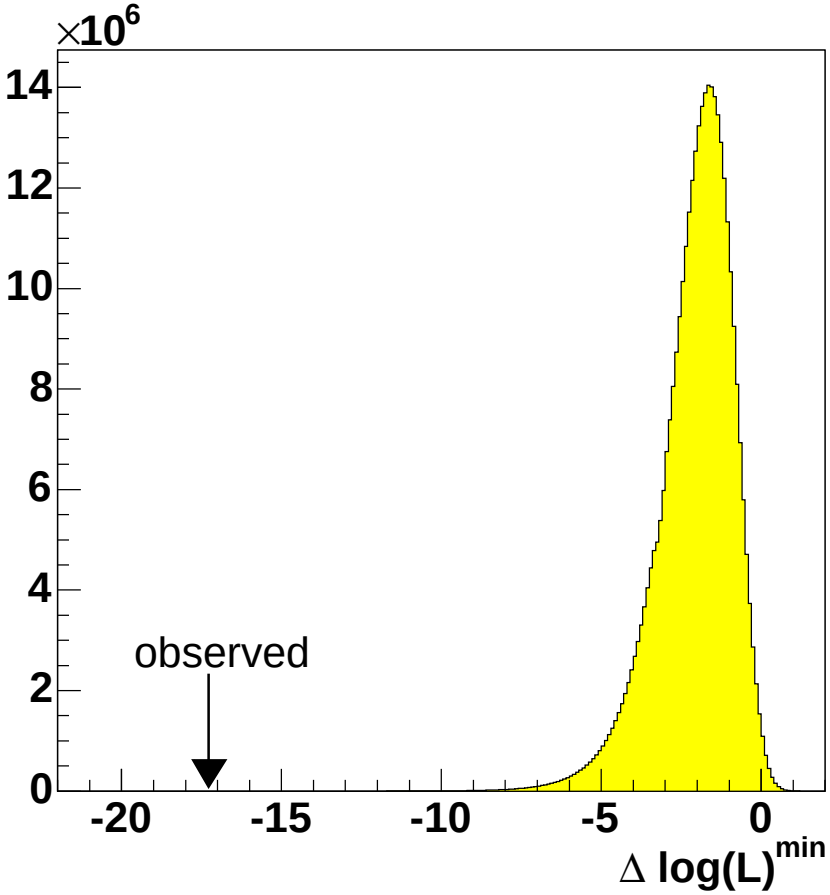
$$\mathcal{A}/\sigma_{\mathcal{A}}(\Delta m_s = 17.75\ \mathrm{ps}^{-1}) = 6.05,$$ but what is the p-value?

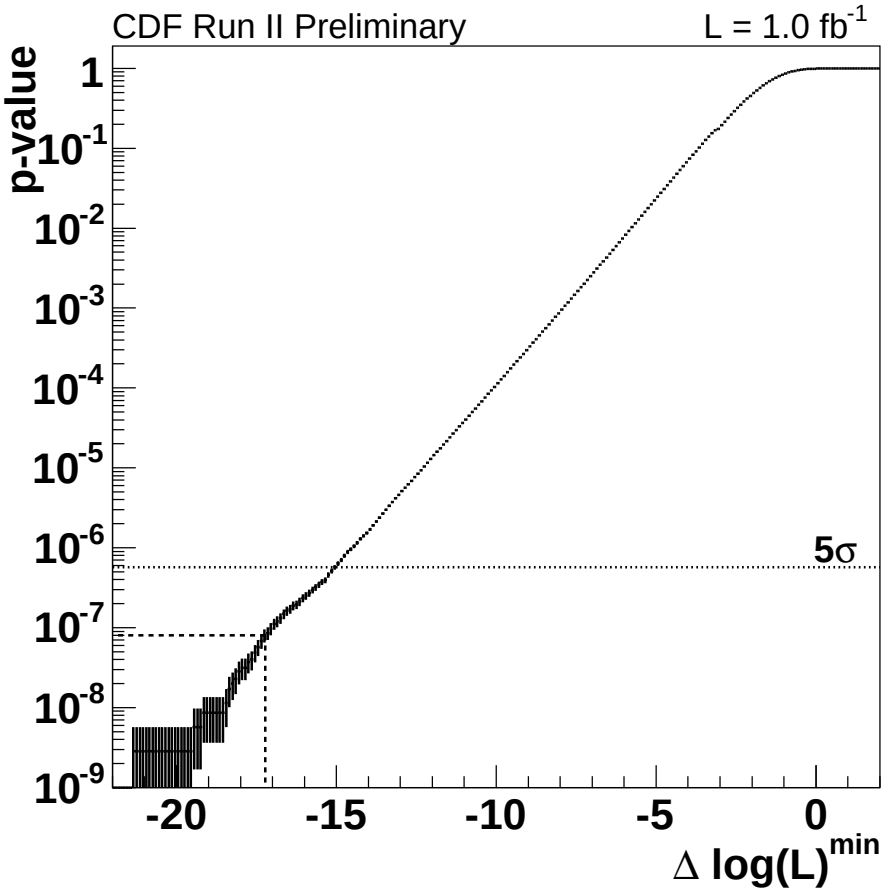
Likelihood Profile



How often can random tags produce a likelihood minimum this deep?

Likelihood Significance





p-value $= 8 \cdot 10^{-8}$ 28 trials with $\min \Delta \log L < -17.26$ / $3.5 \cdot 10^8$ generated

Very small probability of being background fluctuation

$\Rightarrow$ <u>We are in 5σ land!</u> $\Leftarrow$

Δm_s Measurement

Limits

- $[17.56, 17.96]\text{ps}^{-1}$ @ 90% C.L.
- $[16.51, 18.00]\text{ps}^{-1}$ @ 95% C.L.

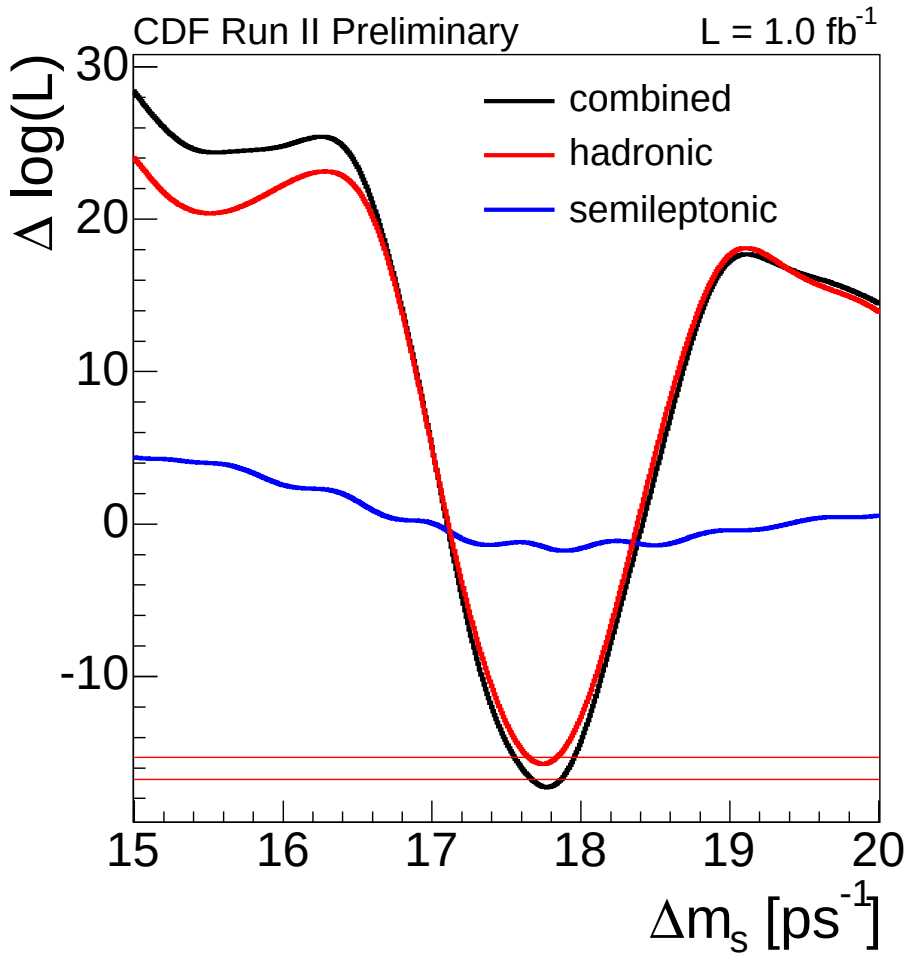
Consistent with SM

- $18.3\ ^{+6.5}_{-1.5}\ \text{ps}^{-1}$ EPS 2005

Agrees with 1st measurement

- $17.31\ ^{+0.33}_{-0.18} \pm 0.07\ \text{ps}^{-1}$

PRL **97**, 062003 (2006)



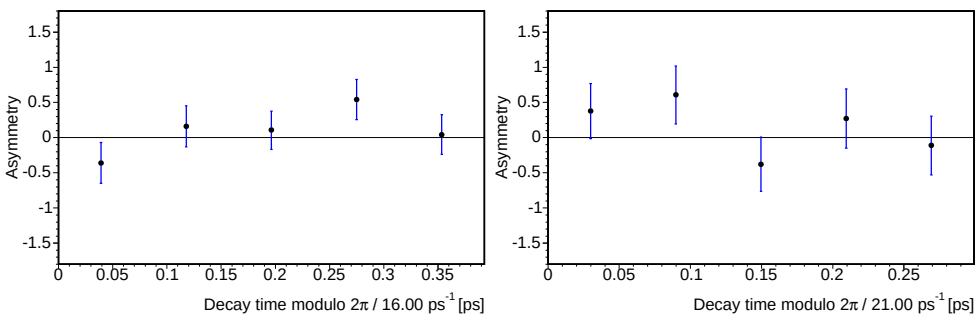
$$\Delta m_s = 17.77 \pm 0.10(stat.) \pm 0.07(syst.)\ \text{ps}^{-1}$$

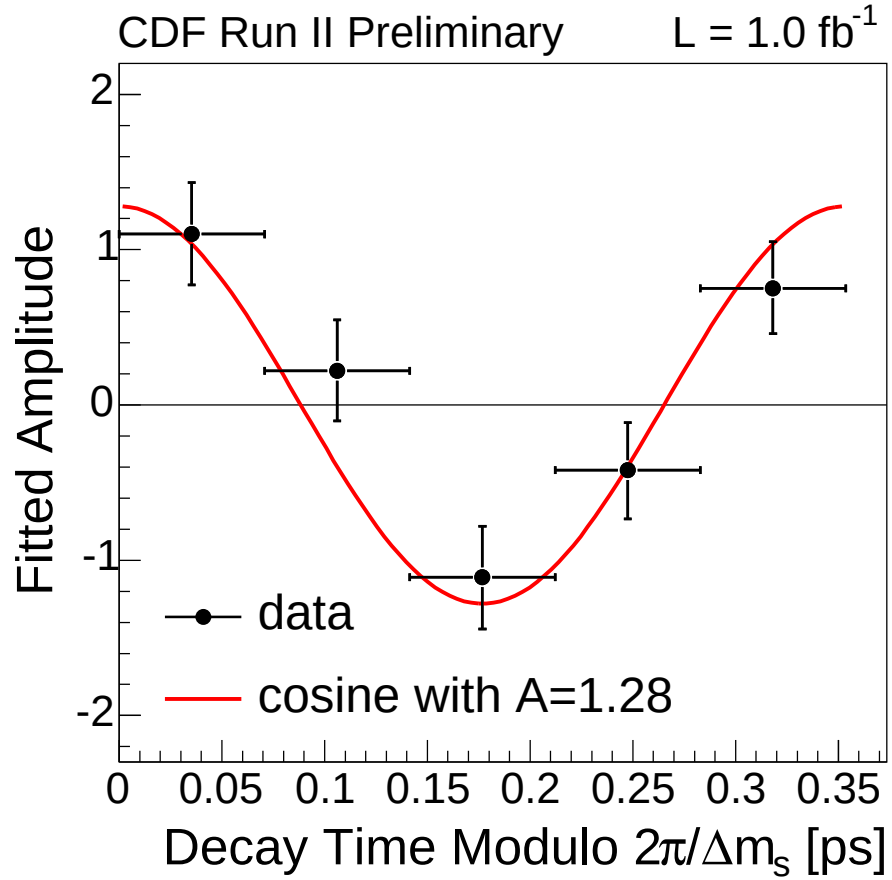
Systematic dominated by the *ct* scale, any other effect very small

Asymmetry plot

- Oscillations folded modulo $2\pi/\Delta m_s$
- Fitter view of data

Fitted $\mathcal{A}$ closely matches result of amplitude scan





$\Delta m_s = 16, 21\ \mathrm{ps}^{-1}$

$\Rightarrow$ no oscillations!

$\chi^2 = 4.77$

NDF = 5(3) $\rightarrow$ Prob = 44%(19%)

A=0 : $\chi^2 = 30.5$

Measurement of CKM parameters

Ratio $\Delta m_s/\Delta m_d$ closely related to parameters of CKM matrix:

$$+ \quad \frac{\Delta m_s}{\Delta m_d} = \frac{m_{B_s}}{m_{B_d}} \xi^2 \frac{|V_{ts}|^2}{|V_{td}|^2}$$

Ingredients with their contribution to the uncertainty:

- $+ \quad \Delta m_d = 0.507 \pm 0.005\ \mathrm{ps}^{-1}$: 1% W.-M. Yao *et al.* 2006
- $+ \quad \Delta m_s = 17.77 \pm 0.10 \pm 0.07\ \mathrm{ps}^{-1}$: $< 1\%$ CDF 2006
- $+ \quad m_{B_d}/m_{B_s} = 0.98390$: $< 0.1\%$ CDF 2006
- $+ \quad \xi = 1.21\ ^{+0.047}_{-0.035}$: 3-4% Okamoto, Lattice 2005

Use $\Delta m_d/\Delta m_s$ ratio to infer:

$$|V_{td}|/|V_{ts}| = 0.2060 \pm 0.0007(exp.)\ ^{+0.0081}_{-0.0060}(th.)$$

Compare to Belle $b \to d\gamma$: $|V_{td}|/|V_{ts}| = 0.199\ ^{+0.026}_{-0.025}(exp.)\ ^{+0.018}_{-0.015}(th.)$

PRL **96**, 221601 (2006)

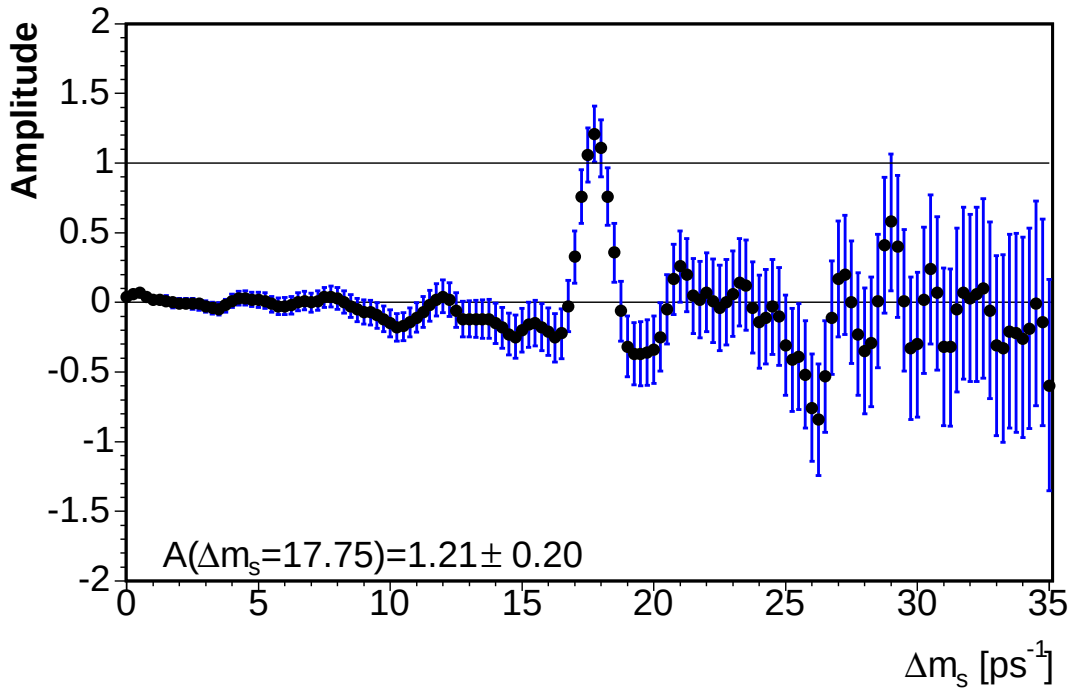
Ratio is no longer limited by experimental measurements

Summary

+ Historic observation of $B_s^0 - \bar{B}_s^0$ oscillations
+ Probability of random tag fluctuation $\sim 8 \cdot 10^{-8}$: $> 5\sigma$!
+ Most precise measurement of oscillation frequency:

$$\Delta m_s = 17.77 \pm 0.10(stat.) \pm 0.07(syst.)\ \mathrm{ps}^{-1}$$

$$|V_{td}|/|V_{ts}| = 0.2060 \pm 0.0007(exp.)\ ^{+0.0081}_{-0.0060}(th.)$$



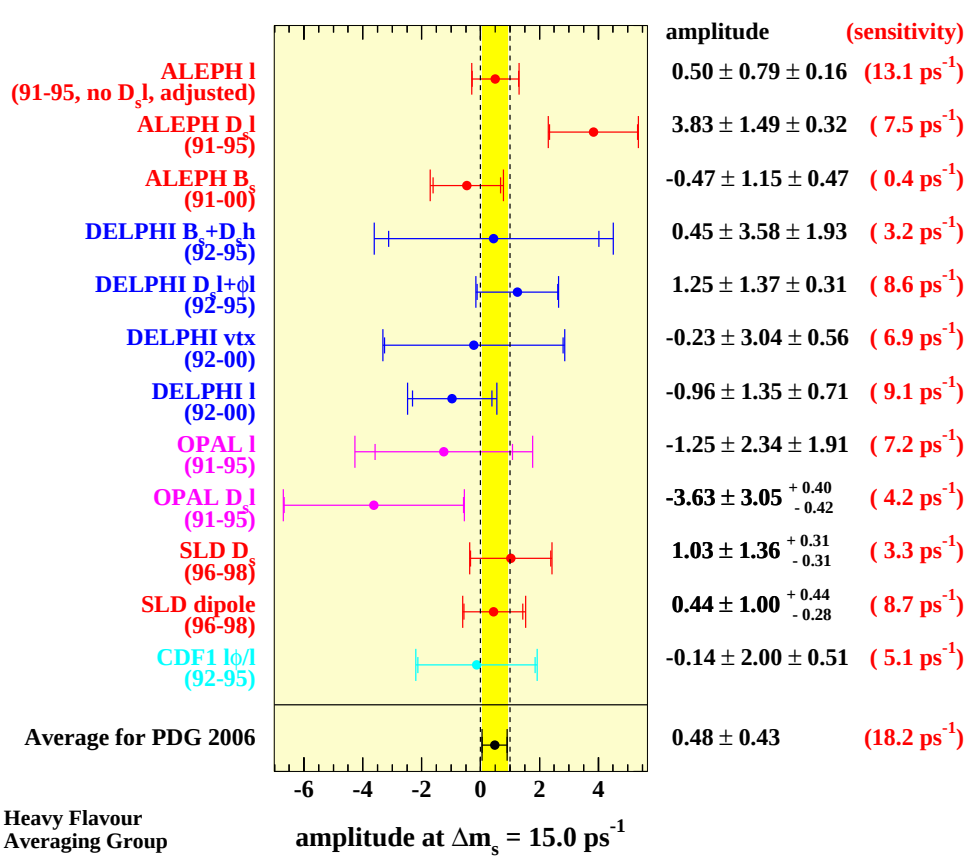
Submitted to PRL! `hep-ex/0609040`

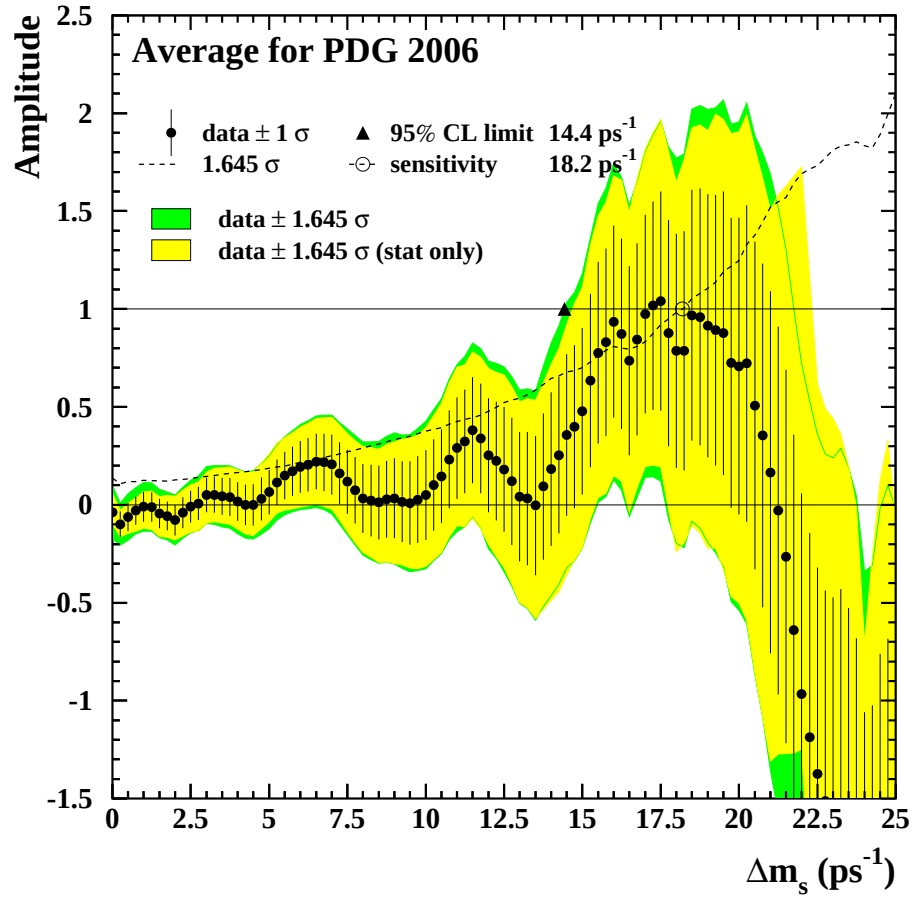
Back Up

Status of published results on Δm_s (as of Jan 06)

LEP + SLD + CDF Run I

$\Delta m_s > 14.4\ \text{ps}^{-1}$ 95% CL





$A(\Delta m_s = 15\ \text{ps}^{-1}) = 0.48 \pm 0.43$

Sensitivity: $18.2\ \text{ps}^{-1}$

Amplitude method: H-G. Moser, A. Roussaire, NIM **A384**, 491 (1997)

CKM Matrix

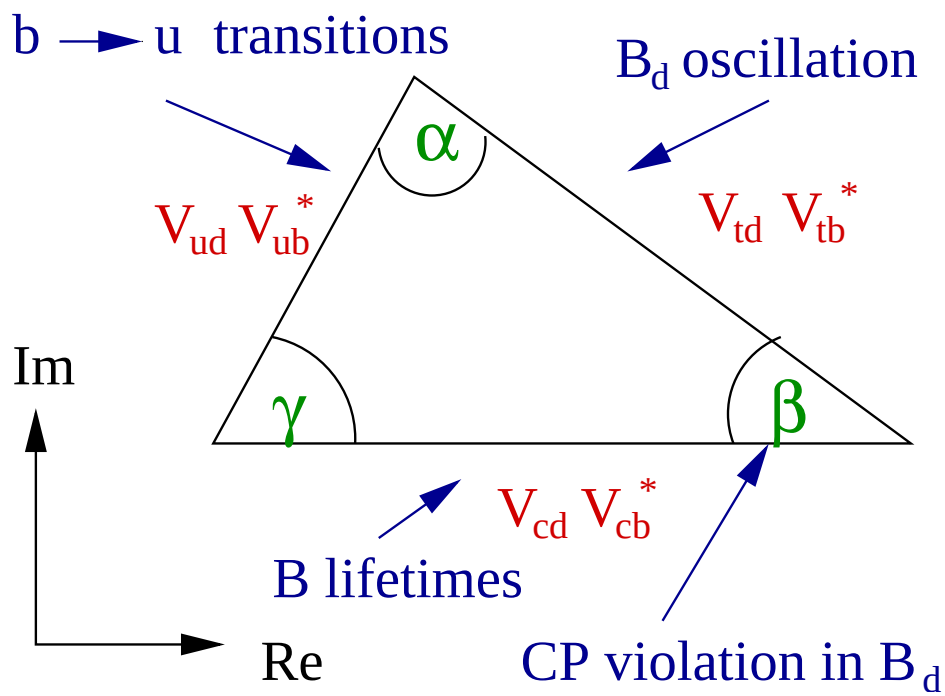
What is the origin of flavor symmetry breaking?
$\rightarrow$ quark mixing, CKM matrix

quark mass eigenstates
$\neq$ weak interaction eigenstates

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}$$

$V * V^{\dagger} = 1$

CKM elements not predicted by SM



Goal: Measure sides/angles of CKM triangle in all possible ways

B-Physics at Hadron Colliders

Large production rates

+ $\sigma(p\bar{p} \to bX, |y| < 0.6) \approx 18\mu b$
+ 10^3 higher than at $\Upsilon(4S)$

Heavy and excited B states currently uniquely at Tevatron:

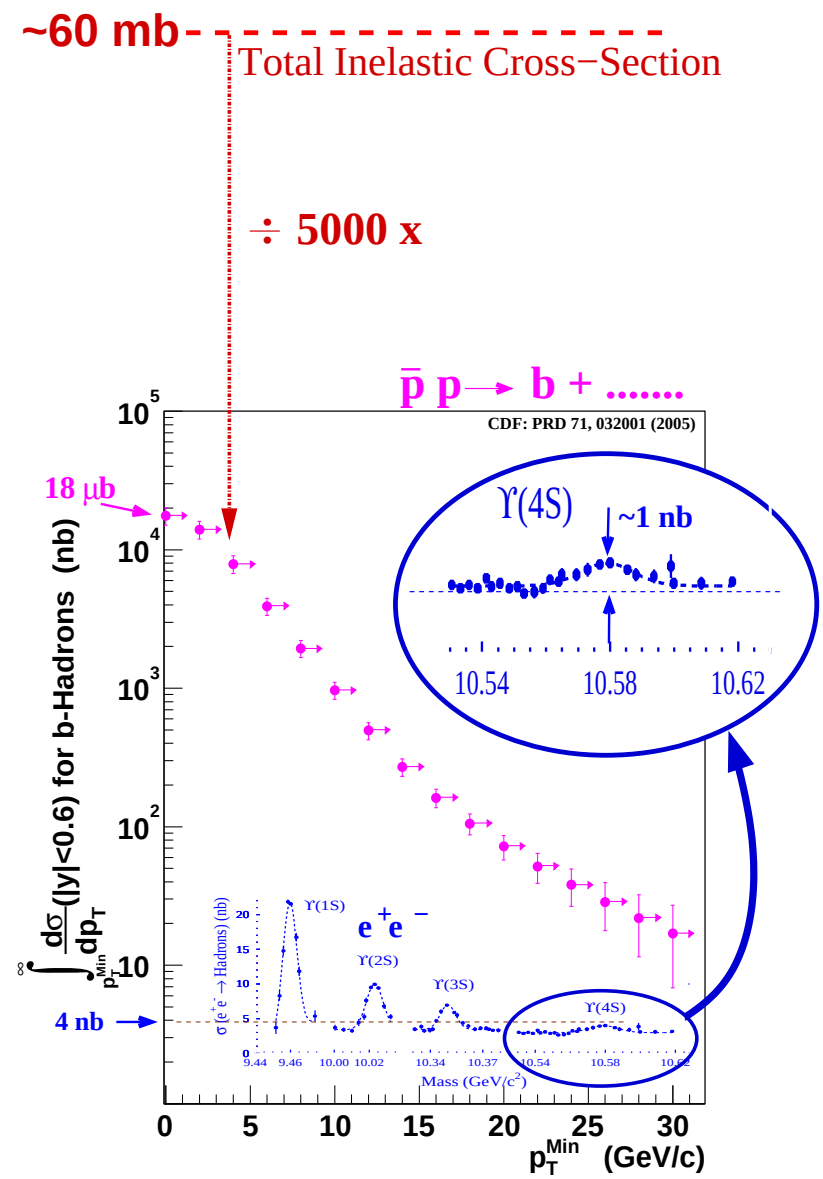
+ $B_s, B_c, \Lambda_b, \Xi_b, B^{**}, B_s^{**}, \ldots$

QCD background is 10^3 higher than signal

+ Triggers are critical

Event signature polluted by many fragmentation tracks

+ High precision vertex tracker
+ Dedicated reconstruction algorithms



B Triggers

B decays to $J/\psi \rightarrow \mu^+\mu^- \rightarrow$ Di-muon trigger

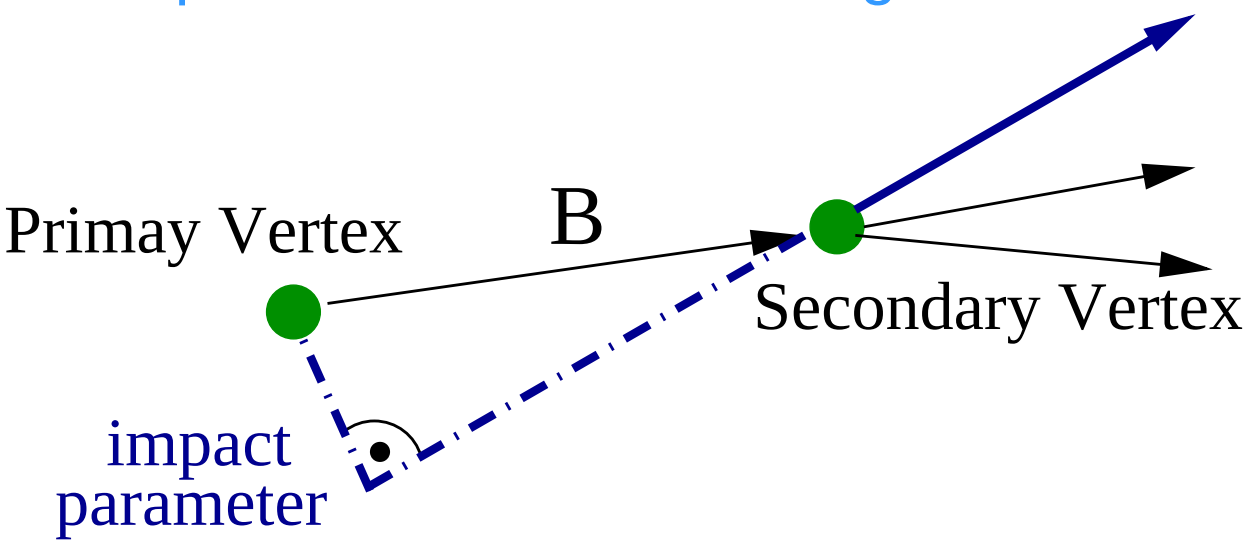
- \+ easy trigger
- \+ clean signature

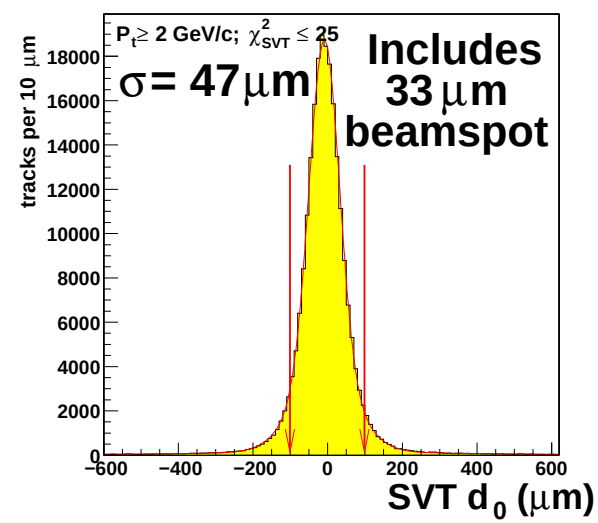
Semileptonic *B* decays $\rightarrow$ lepton + displaced track trigger

- \+ large branching ratio
- \+ missing momentum (neutrino & neutrals)

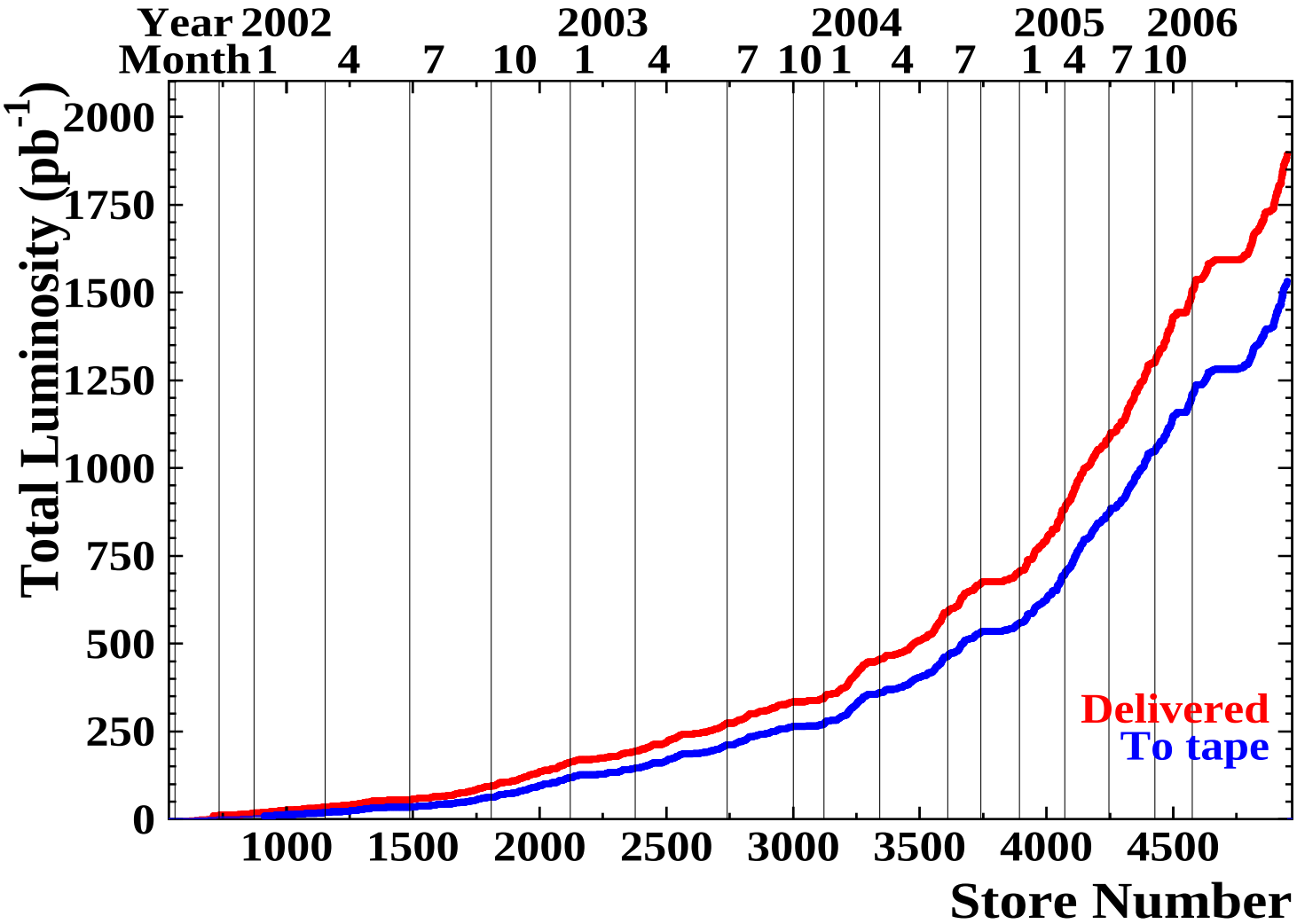
Fully hadronic *B* decays $\rightarrow$ Two Track trigger

- \+ requires displaced track trigger
- \+ requires fast online tracking





Data



$\mathcal{L}$(tape) ~ 1.5 fb^{-1}

$\mathcal{L}$(used in analysis) ~ 1.0 fb^{-1}

B_s *signal samples*

Fully reconstructed modes: $B_s \to D_s(\pi\pi)\pi, D_s \to \phi\pi/K^*K/3\pi$

- + relatively small branching ratios
- + all daughters reconstructed
- + "Two Displaced Track Trigger" used to collect those events

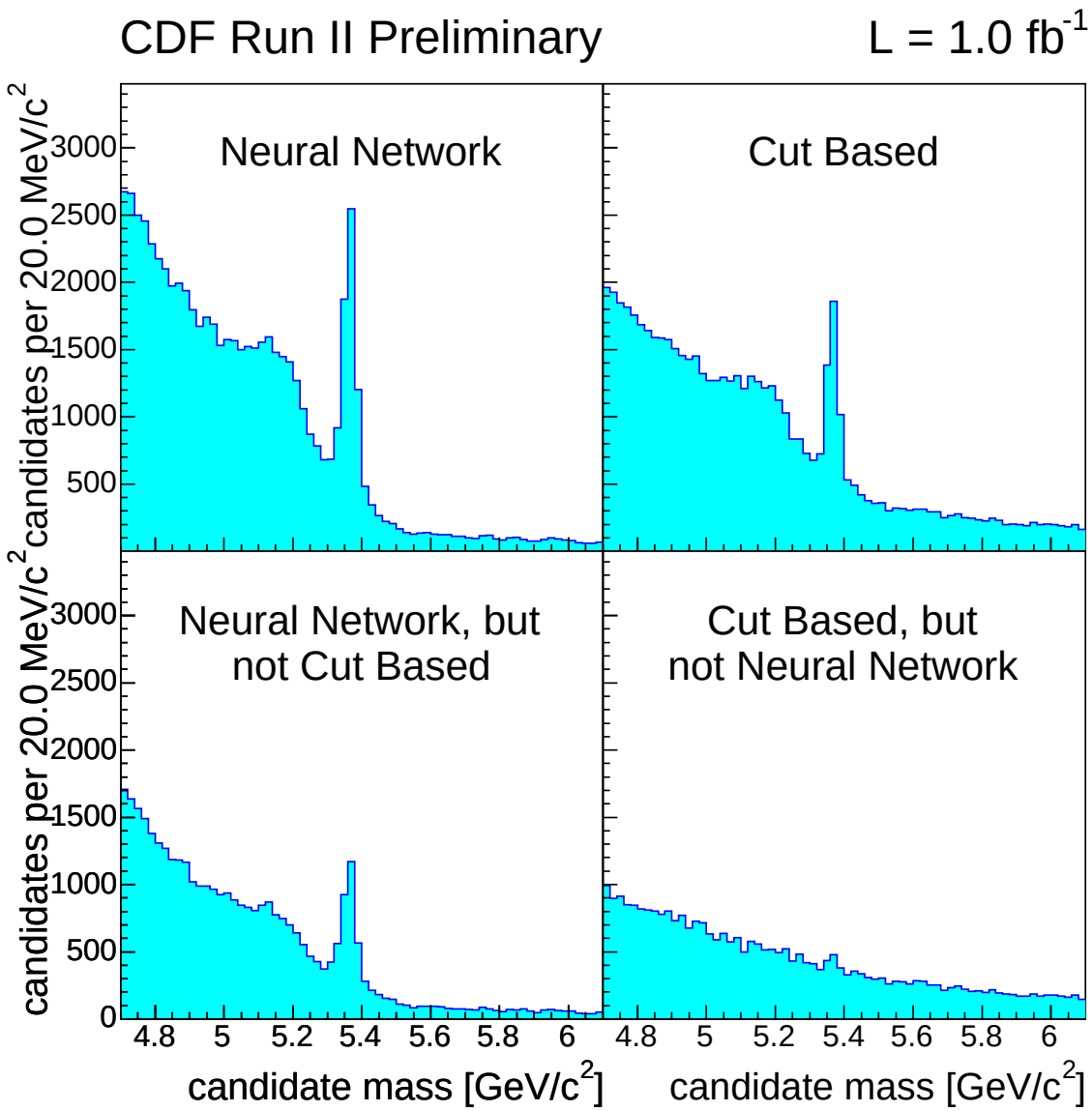
Partially reconstructed modes: $B_s \to D_s{}^*\pi/D_s\rho, D_s \to \phi\pi$ New!

- + *almost* completely reconstructed (soft track lost)
- + same trigger as fully reconstructed modes

Semileptonic modes: $B_s \to D_s\ell X, D_s \to \phi\pi/K^*K/3\pi$

- + large branching ratios
- + missing momentum (neutrino & neutrals)
- + "Two Displaced Track Trigger" & "lepton + Displaced Track Trigger" used to collect those events

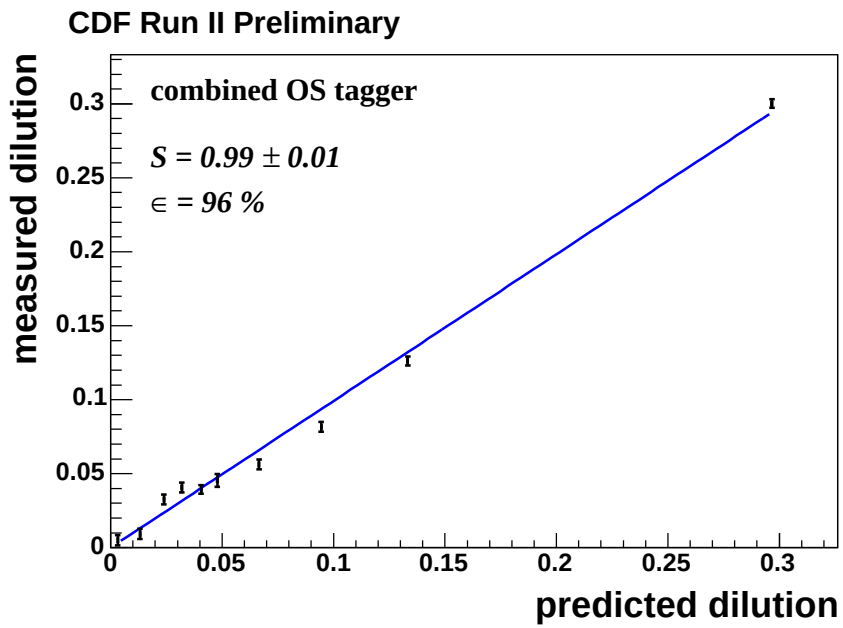
Performance of Neural Network selection



Opposite Side Taggers

Combined Tagger:

- + Soft Muon Tagger
- + Soft Electron Tagger
- + Jet Charge Tagger
- + Opposite Side Kaon Tagger



Neural Network uses very well understood OST taggers

Dilution is proportional to NN output

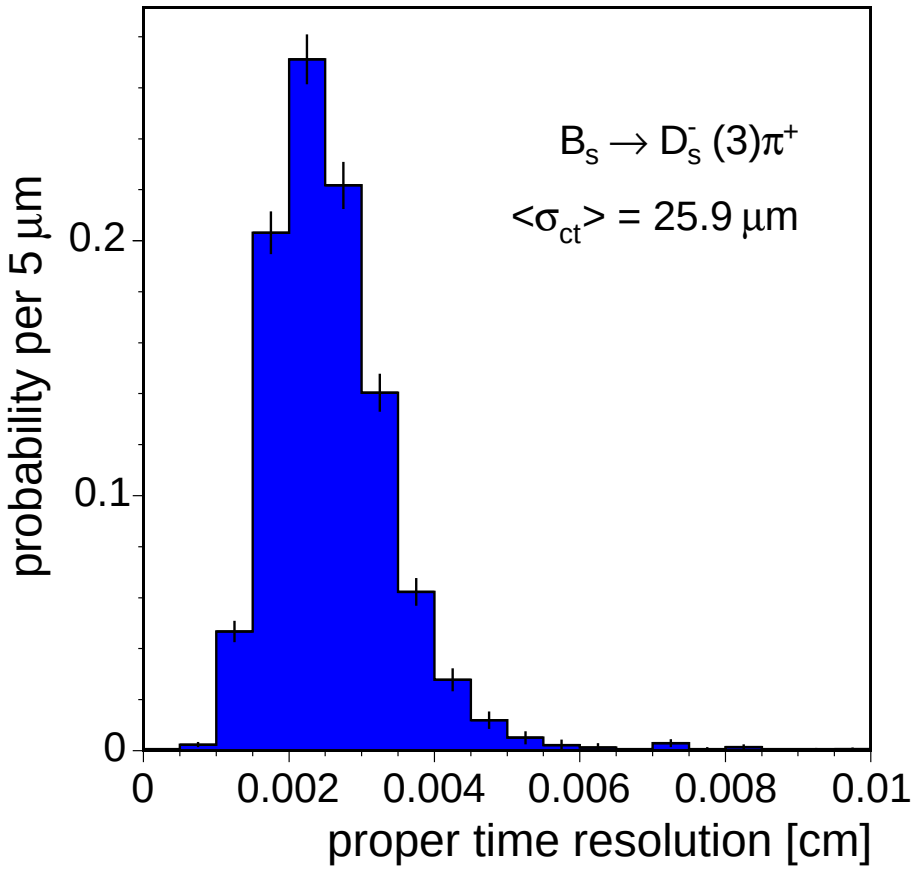
Combined tagger calibrated in our high statistics $\ell + SVT$ samples

Overall scale factor measured on $B^{0/+}$ candidates to take care a possible overall (small) shift

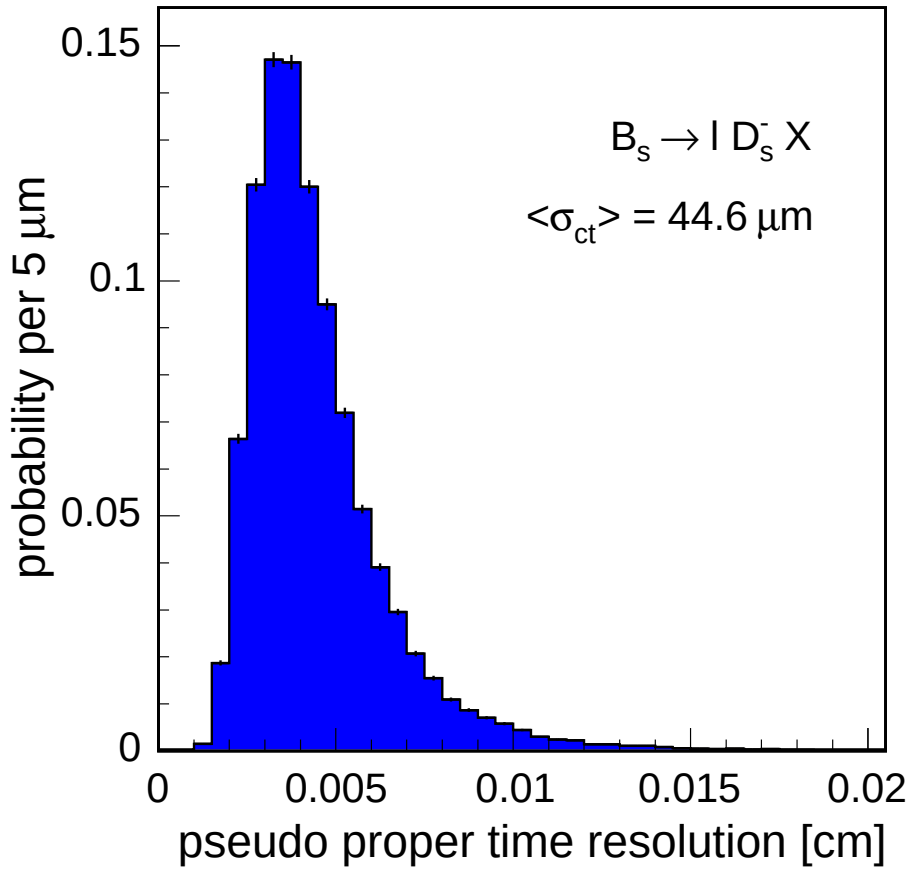
Improvement of 15% w.r.t. hierarchical combination

Proper Decay Time Resolution - σ_{ct}^{hadr} and σ_{ct}^{semil}

Critical aspect of the analysis, limiting factor at high Δm_s



Hadronic (fully rec.)



Semileptonic

One oscillation at $\Delta m_s = 18\ ps^{-1}$: $\sim 100\ \mu m$

$\Rightarrow$ can resolve 4 oscillations with hadronic resolution!

Semileptonic plot w/o p_T-dependent error

Same Side Kaon Tagger (I)

Input variables

- CLL, ΔR, p_T, p_L^{rel}, p_T^{rel}
- boolean var. set true when all tag cands have same charge

NN Training

- Signal: RS kaons
- Background: WS kaons, pions and protons

Tag decision: charge of track with max NN output

Dilution parametrized as function of NN output

Improved performance in Monte Carlo

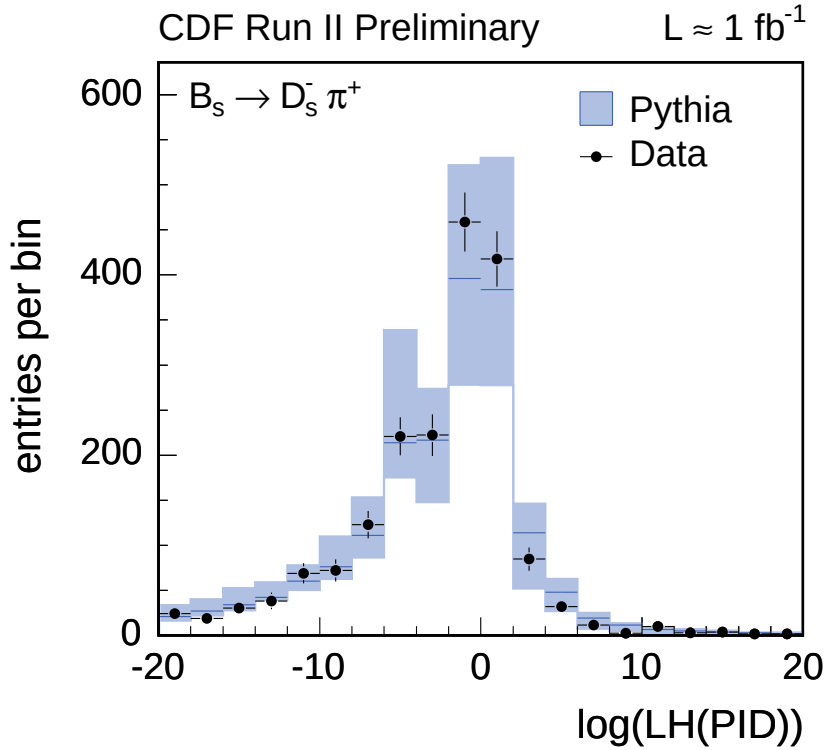
OLD	NEW
$S_D = 98.8 \pm 2.5$ %	$S_D = 99.2 \pm 2.1$ %
$S_D\sqrt{< D^2 >} = 28.5 \pm 0.7$ %	$S_D\sqrt{< D^2 >} = 30.2 \pm 0.7$ %

Same Side Kaon Tagger (II)

Small discrepancies covered by systematics

Systematic studies cover:

- + Fragmentation Model
- + $b\bar{b}$ Production Mechanisms
- + B^{**} content
- + Detector/PID resolution
- + Multiple interactions
- + PID content around B
- + Data/MC agreement



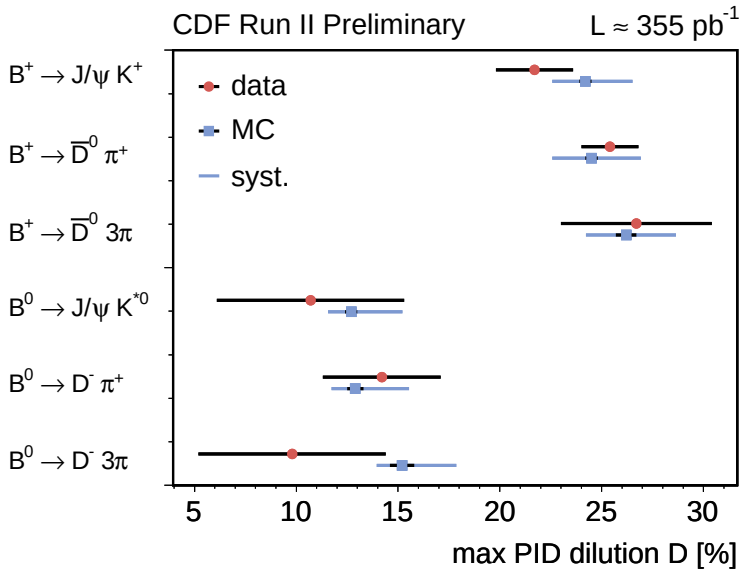
TOF & dE/dx are used for particle identification

Same Side Kaon Tagger (III)

Have to rely on Monte Carlo for prediction of SS(K)T performance for B_s decays!

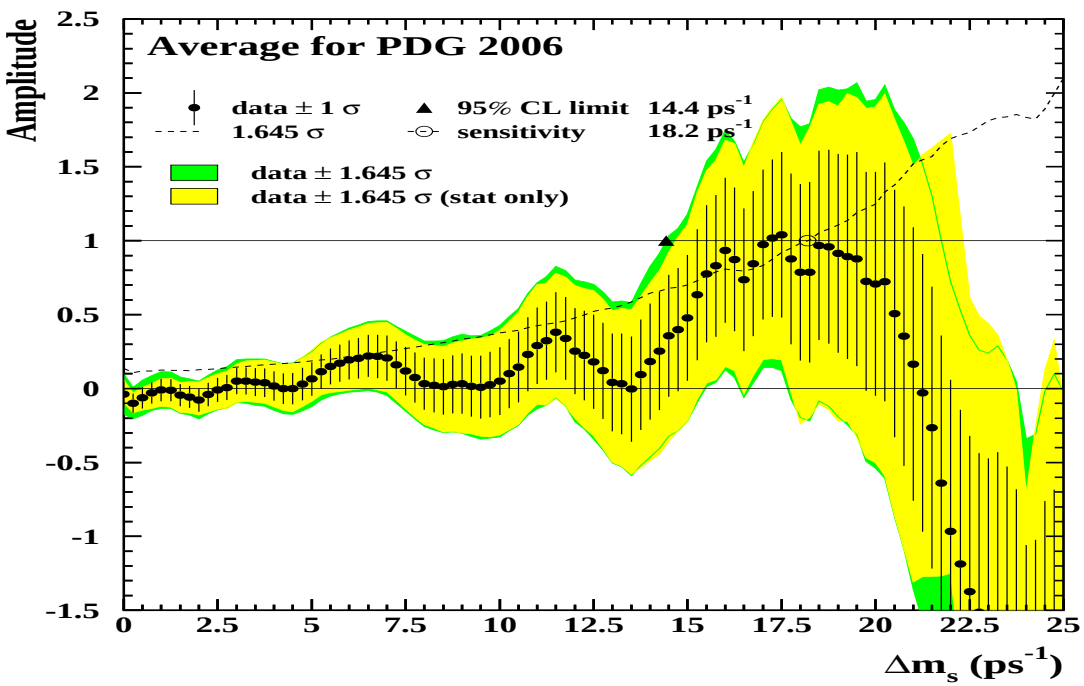
- + Extensive data/MC comparisons on all tagging related quantities
- + Different tagging algorithms probe different aspects of the fragmentation

Very good agreement in high statistics B^+ and B^0 modes in all checks!



Amplitude scan notations

- amplitude error bars from unbinned likelihood fit
- yellow: 1.645σ around data points defines 95% CL region
- Δm values where $\mathcal{A} + 1.645\sigma < 1$ are excluded at 95% CL
- dashed line: 1.645σ as a function of Δm
- sensitivity: the Δm where $1.645\sigma = 1$
- on average, we expect to observe mixing within sensitivity



Crucial Test of the fitter: B_d Mixing

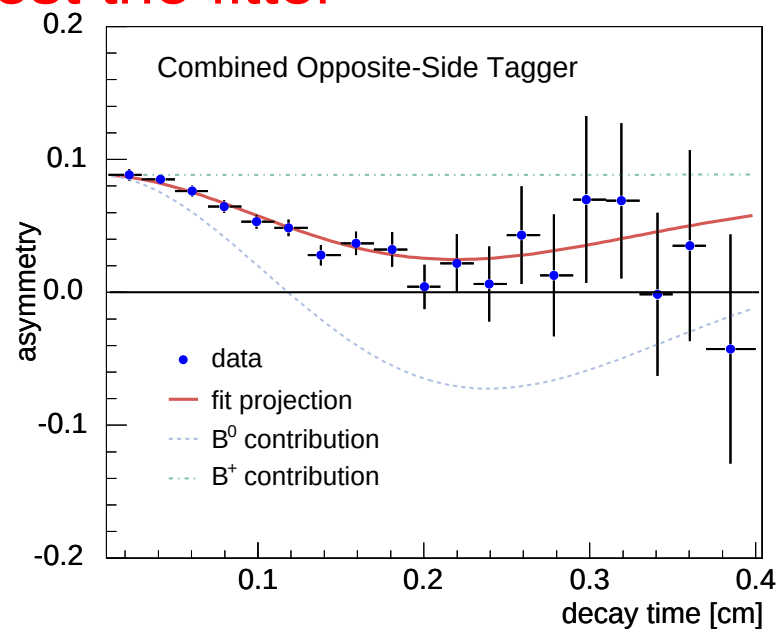
To set limits on Δm_s, knowledge of tagger performance is crucial

- measure tagging dilution in kinematically similar B^0/B^+ samples (for OST)

Δm_d and Δm_s fits are very complex

- combining several B flavor and several decay modes
- combining several taggers
- mass and lifetime templates for various backgrounds

Δm_d measurement is very important to test the fitter



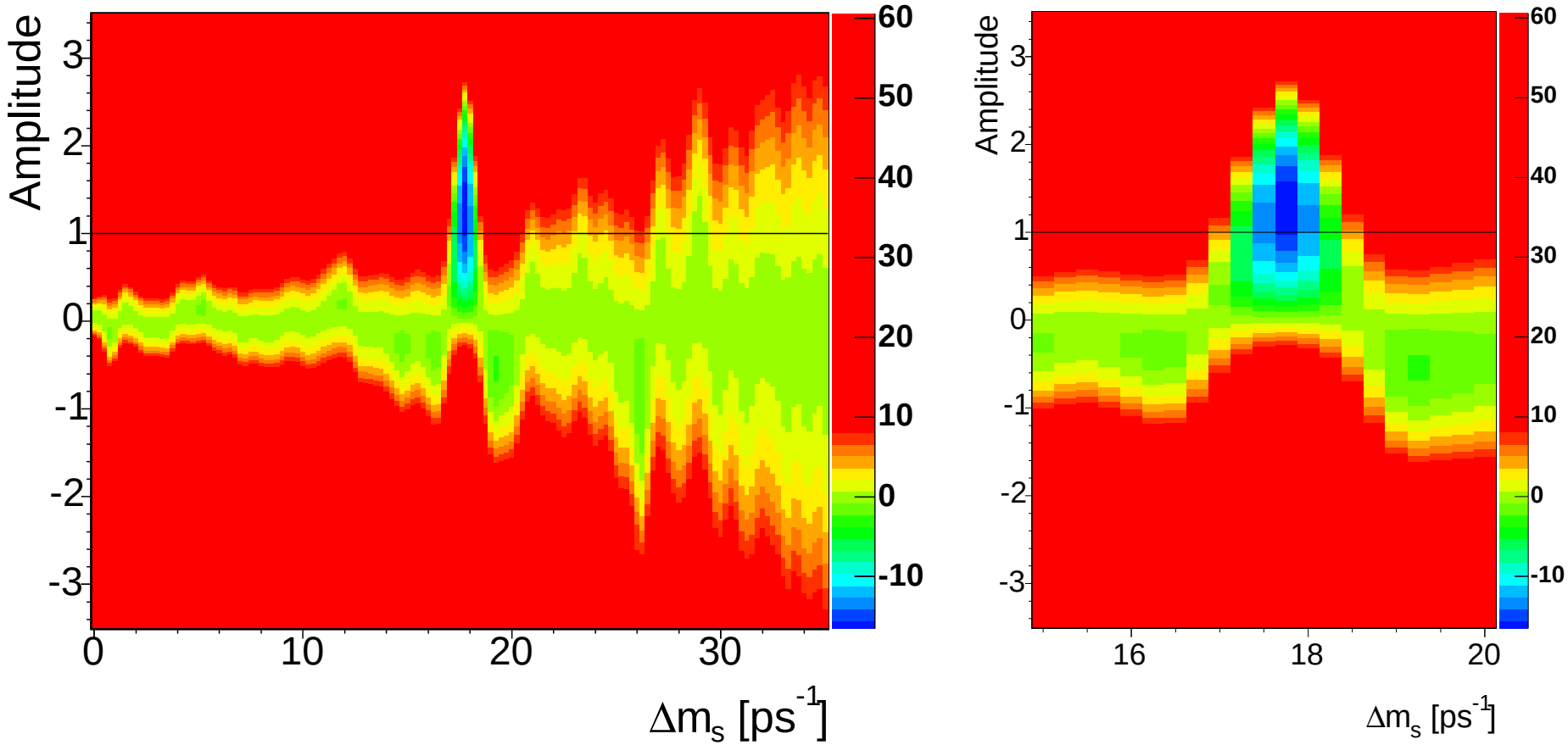
CDF measurement:

- $\Delta m_d = 0.509 \pm 0.010 \pm 0.016\ \mathrm{ps}^{-1}$

World average:

- $\Delta m_d = 0.507 \pm 0.005$

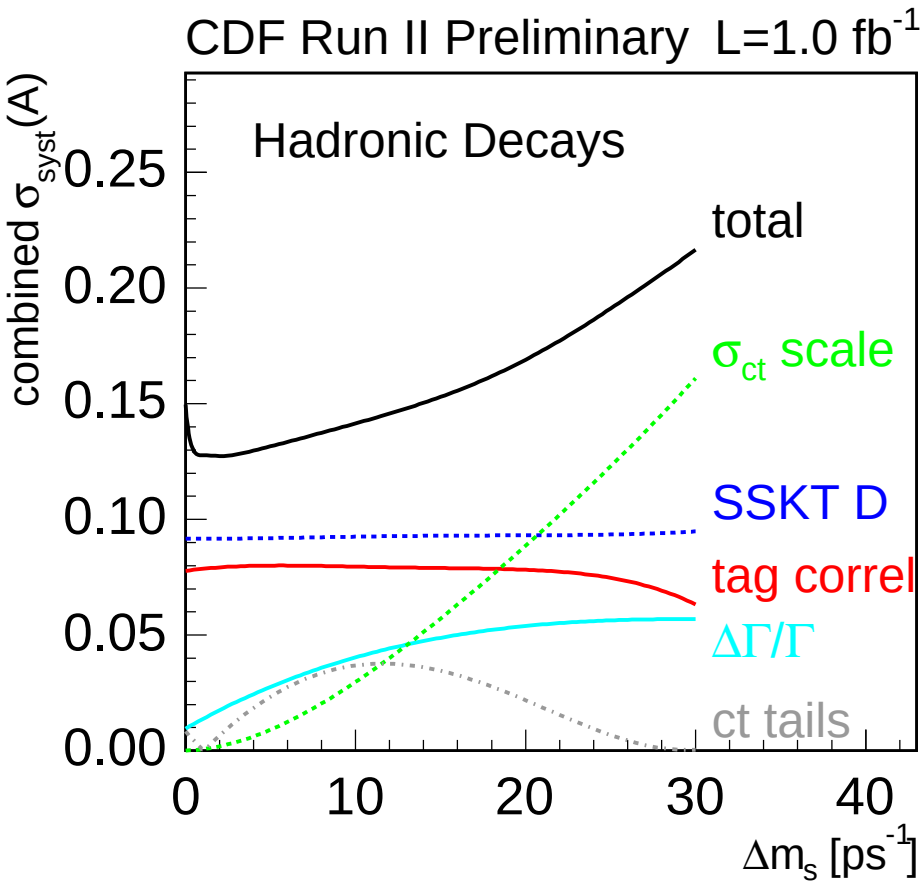
Amplitude vs. Δm_s



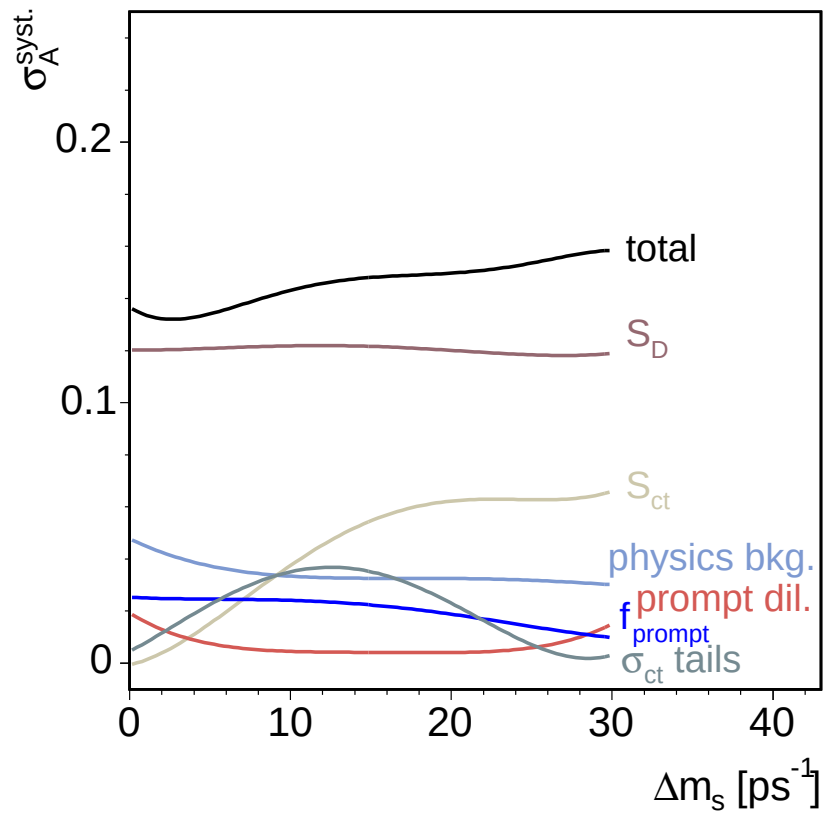
Clear deep in $\mathcal{A} \sim 1$ and $\Delta m_s \sim 17.75$ ps^{-1}

Systematic Uncertainties on $\mathcal{A}$

Hadronic



Semileptonic



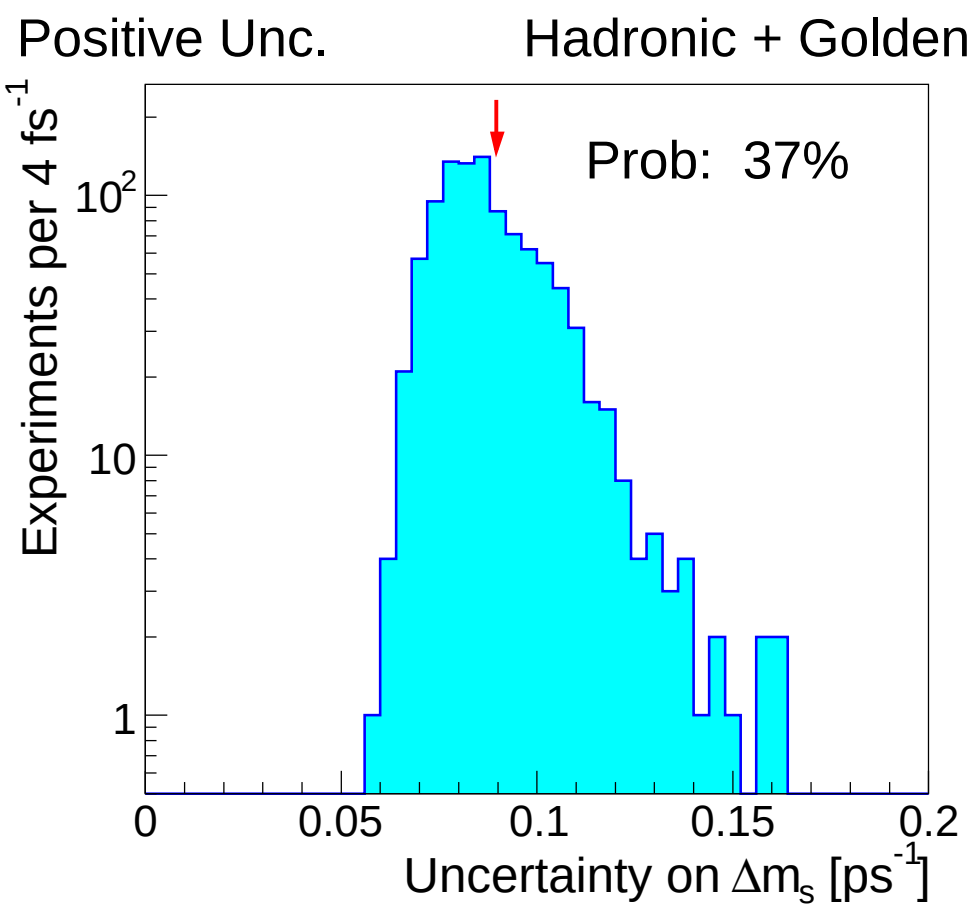
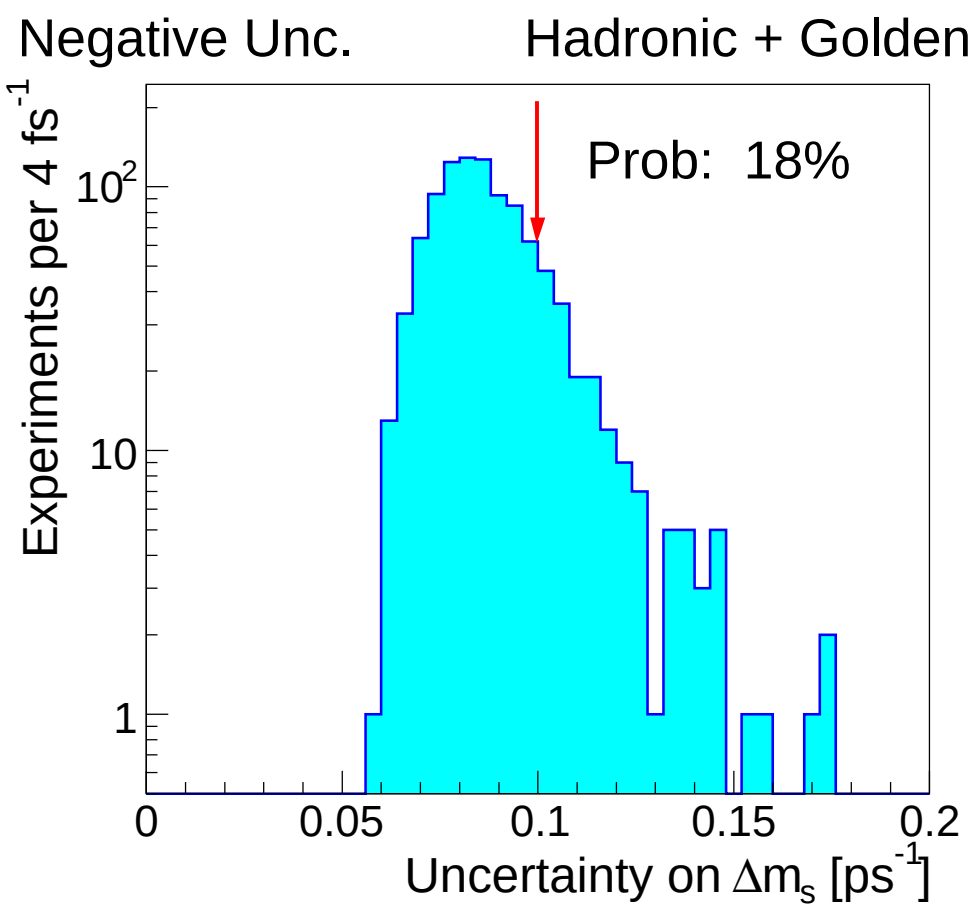
Systematic uncertainties ~0.15-0.20 at high Δm_s

Analysis limited by statistics

Systematic Uncertainties on Δm_s

Source	Value (ps^{-1})
Silicon detector alignment	0.04
Track fit bias	0.05
Primary vertex bias	0.02
Hadronic k-factors	0.03
Amplitude scan systematic effects	< 0.01
Total	0.07

All relevant systematic uncertainties:

+ related to ct scale
+ common between hadronic and semileptonic samples

Probability of Δm_s uncertainty