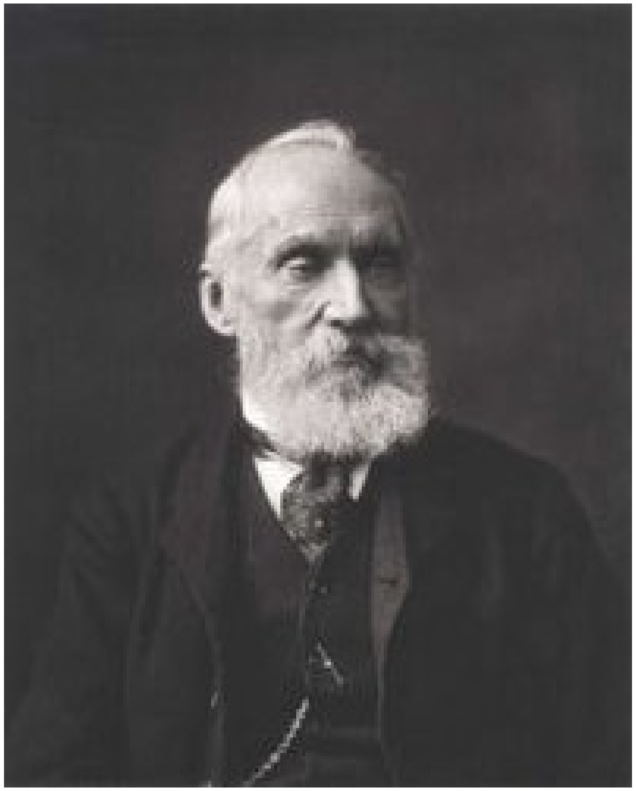
B_s^0 Mixing at CDF

Sinéad M. Farrington
University of Liverpool

University of Edinburgh Seminar October 2006

Lord Kelvin



"Science is bound, by the everlasting vow of honour, to face fearlessly every problem which can be fairly presented to it."

James Clerk Maxwell

“Aye, I suppose I could stay up that late.”

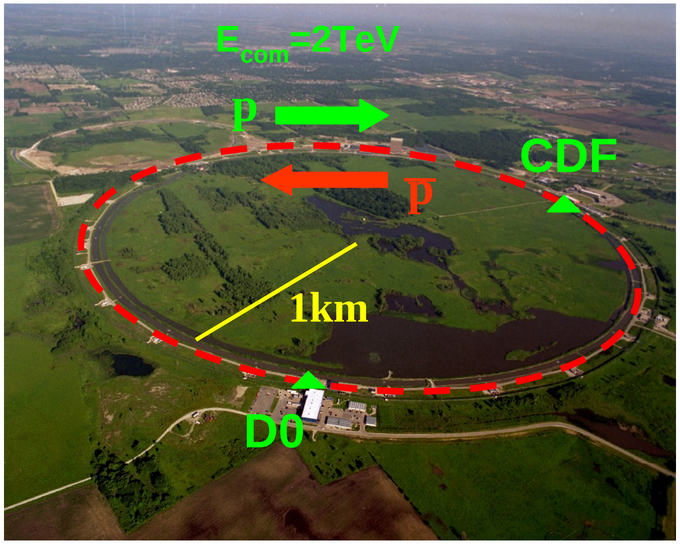
Observation of B_s Mixing

This year the phenomenon of mixing was observed for the first time in the B_s meson system

I shall:

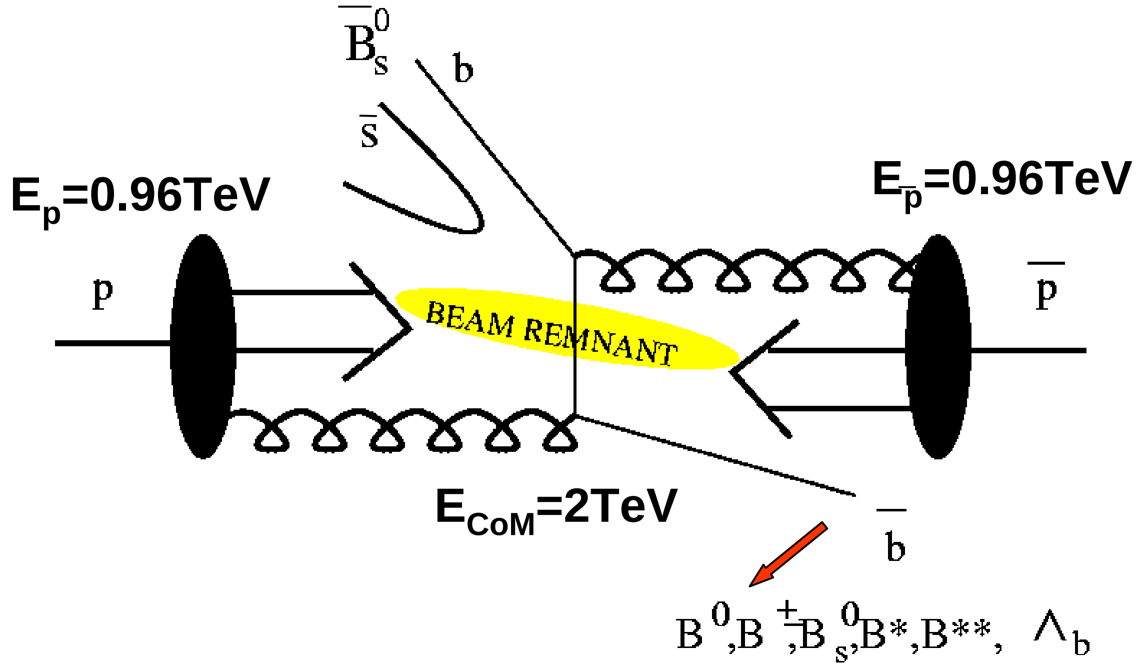
- Describe, in brief, the CDF experiment
- Explain why B_s mixing is interesting
- Explain the experimental method to measure it
- Present the experimental results
- Show how these are interpreted within the Standard Model

The Tevatron



Fermilab, Chicago 

Currently the world's highest energy collider

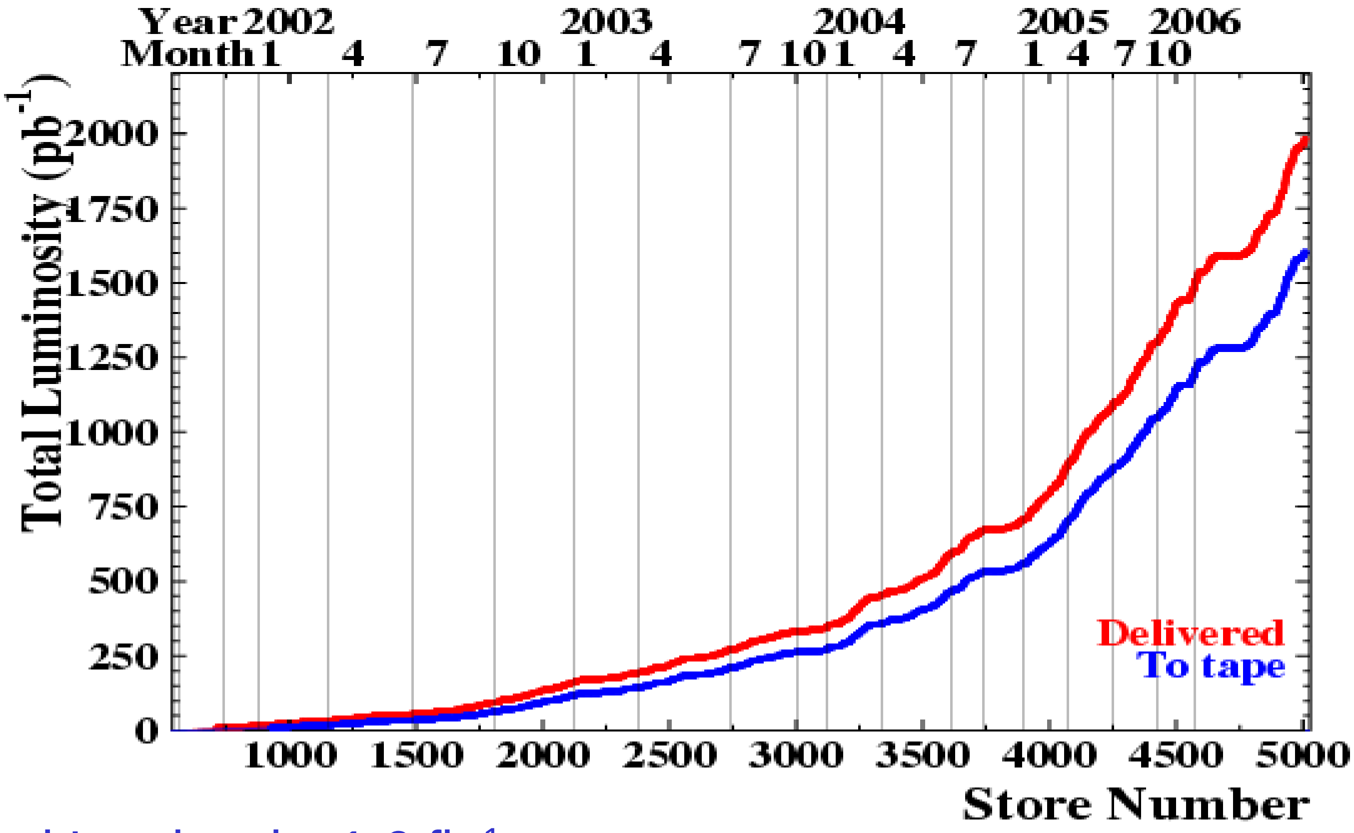


Hadron collisions can produce a wide spectrum of b hadrons (in a challenging environment)

B_s cannot be produced at the B factories since their Centre of Mass energy is below threshold (except for a special run by Belle)

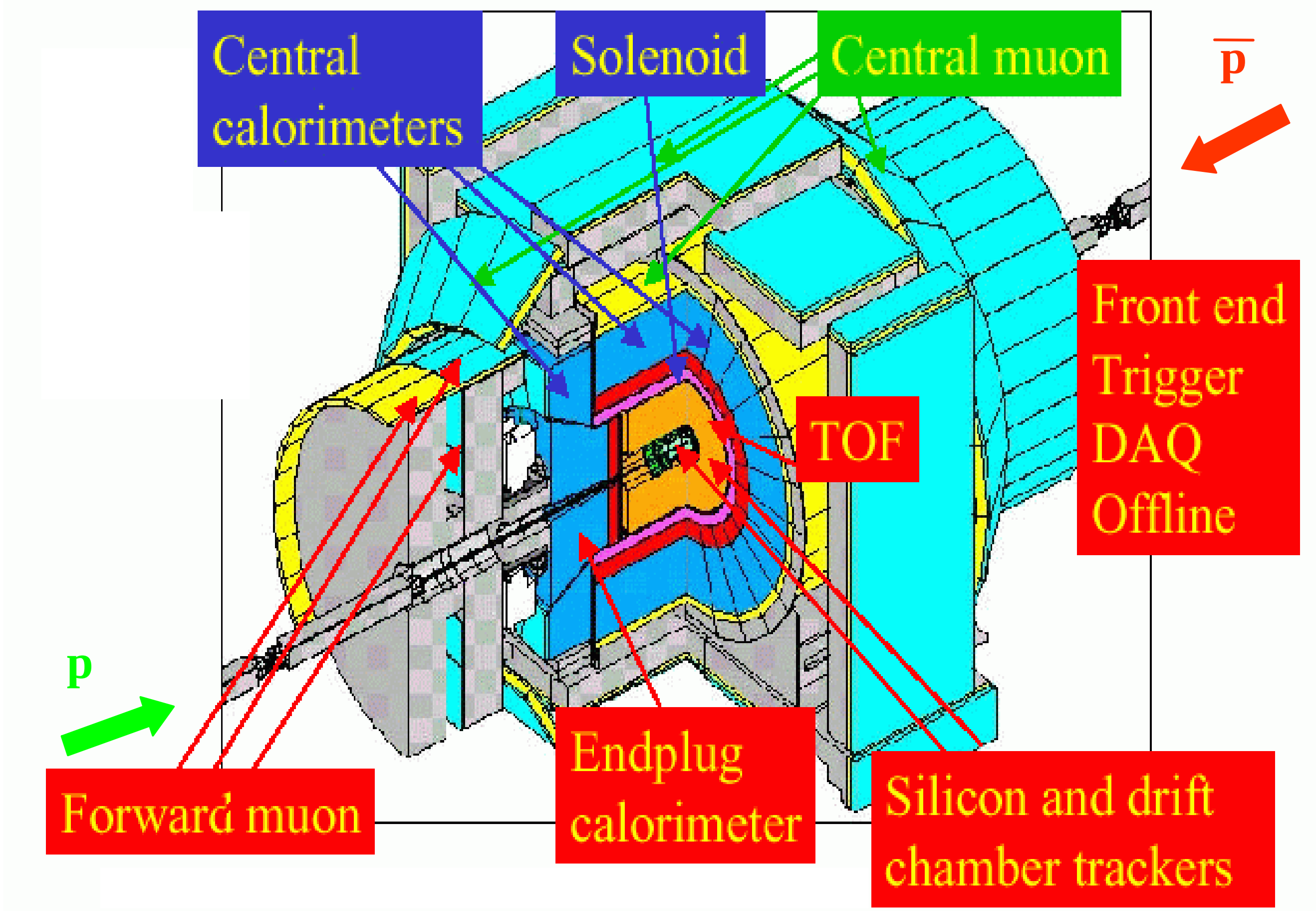
Tevatron Integrated Luminosity

Run I: 1992-1996 L= 0.1fb^{-1}
Major Upgrades 1996-2001
Run II: 2001-2006 L= 1.6 fb^{-1}



- Recorded Luminosity 1.6 fb^{-1}
- This analysis: Feb 2002 – Jan 2006: 1 fb^{-1}

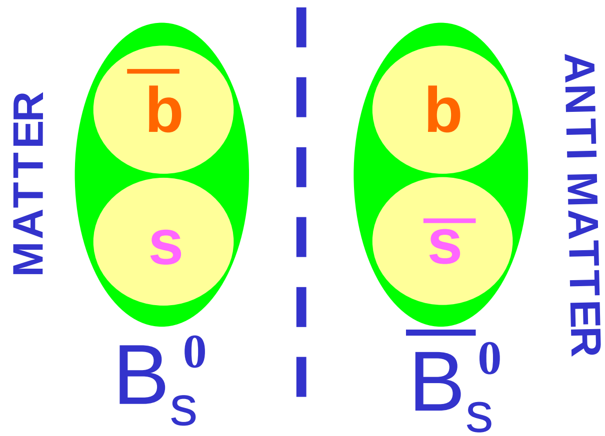
The CDF Detector and Triggers



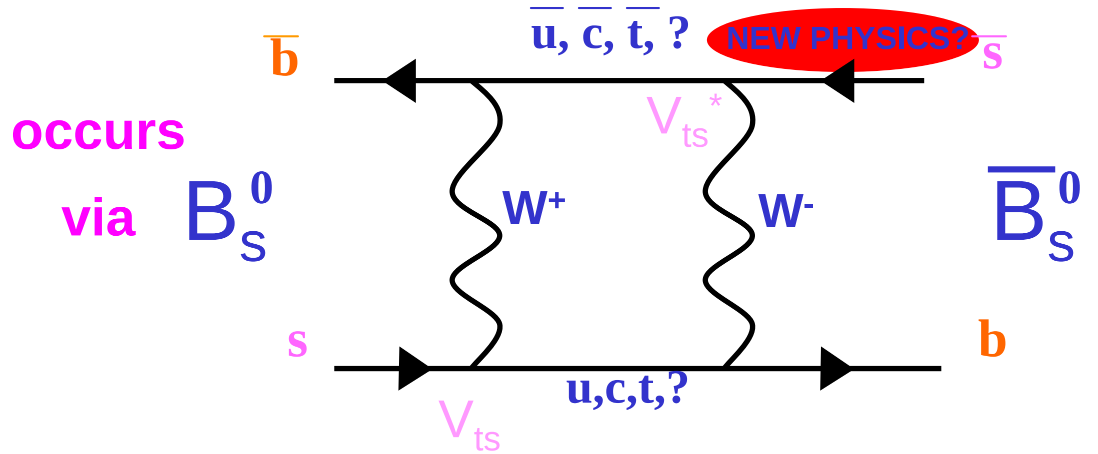
- $\sigma(b\bar{b}) << \sigma(p\bar{p})$ $\Rightarrow$ B events are selected with specialised triggers
- Displaced vertex trigger exploits long lifetime of B's
- Yields per pb^{-1} are ~3x those of Run I

B_s^0 Physics

Bound states:



Matter↔antimatter:



- Physical states, H and L, evolve as superpositions of B_s^0 and $\bar{B}_s^0$
- System characterised by 4 parameters:

  masses: m_H, m_L lifetimes: Γ_H, Γ_L ($\Gamma=1/\tau$)
- Predicted Δm_s around 20ps^{-1}

$$\Delta m_s = \frac{G_F^2 m_W^2 \eta S(m_t^2 / m_W^2)}{6\pi^2} m_{B_s} f_{B_s}^2 B_{B_s} \left|V_{ts}^* V_{tb}\right|^2$$

- No measurements of Δm_s have been made until now:

"I have no satisfaction in formulas unless I feel their numerical magnitude."

(Kelvin)

Why is Δm_s interesting?

1) Probe of New Physics
   - may enter in box diagrams

2) Measure CKM matrix element:

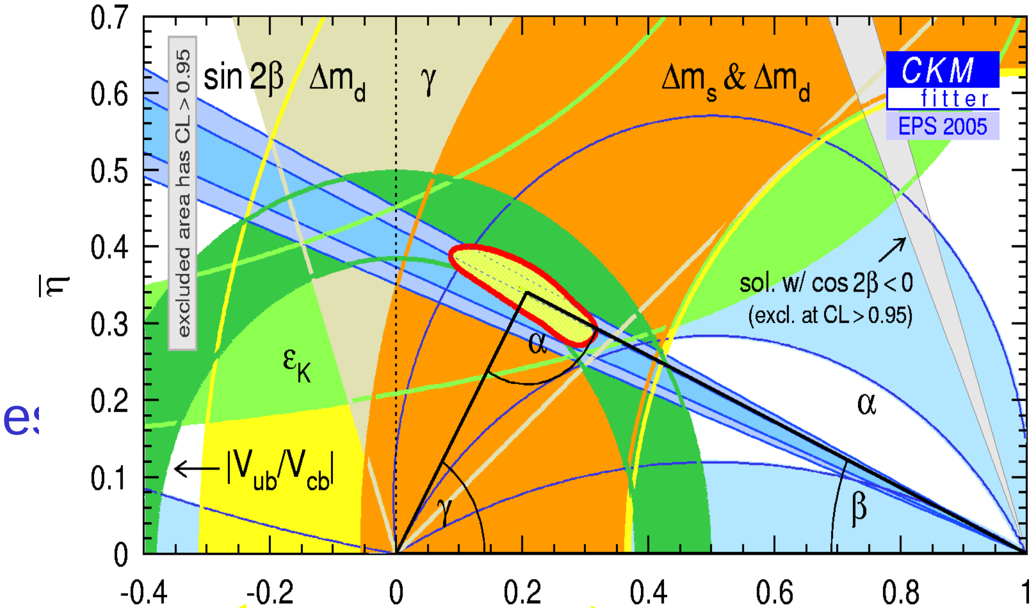
Δm_d known accurately from B factories

- V_{td} known to 15%
- Ratio $V_{td}/V_{ts} \propto \Delta m_d/\Delta m_s$ related by constants:

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

- ξ (from lattice QCD) known to ~4%

- So: measure Δm_s gives V_{ts}



$\bar{\rho}$

Lower limit on Δm_s

from Δm_d

from $\Delta m_d/\Delta m_s$

•CKM Fit result:

Δm_s: 18.3+6.5 (1σ) : +11.4 (2σ) ps^{-1}

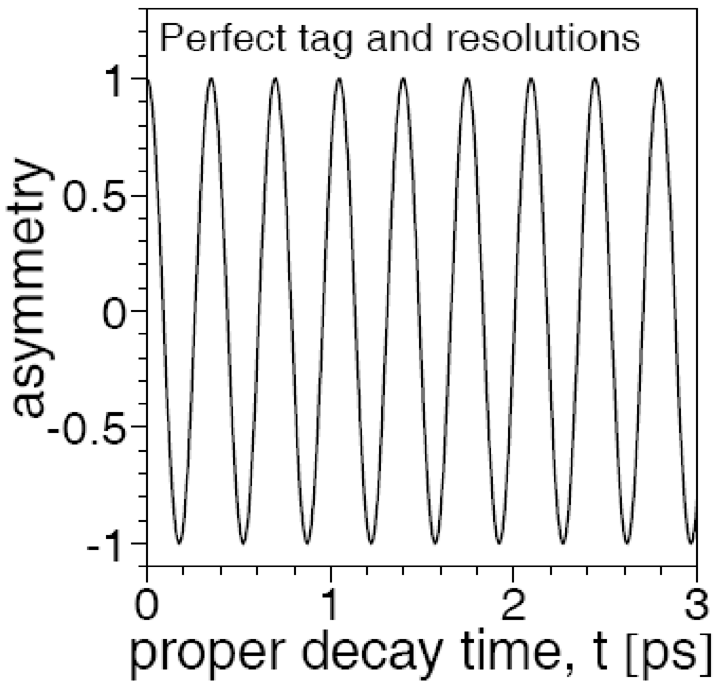
Standard Model Predicts rate of mixing, $\Delta m = m_H - m_L$, so Measure rate of mixing $\Rightarrow V_{ts}$ (or hints of NEW physics)

Measuring Δm_s

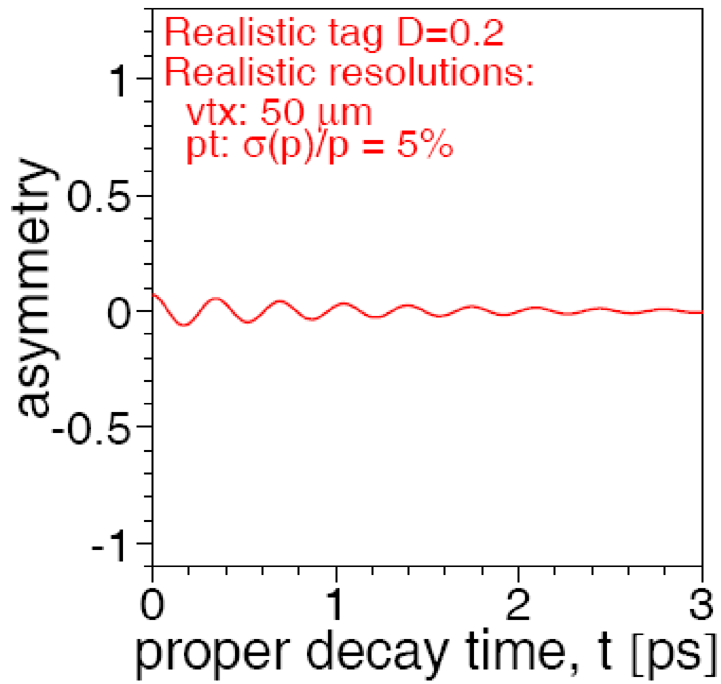
In principle: Measure asymmetry of number of matter and antimatter decays:

$$A(t) \equiv \frac{N(B_s^0 \to B_s^0)(t) - N(B_s^0 \to \overline{B}_s^0)(t)}{N(B_s^0 \to B_s^0)(t) + N(B_s^0 \to \overline{B}_s^0)(t)} \propto \cos(\Delta m t)$$

In practice: asymmetry is barely discernible after experimental realities:



Perfect resolutions



After momentum, time resolution, flavour tag power

Measuring Δm_s

So instead we employ two methods:

1: amplitude scan method

H. G. Moser, A. Roussarie, NIM **A384** (1997)

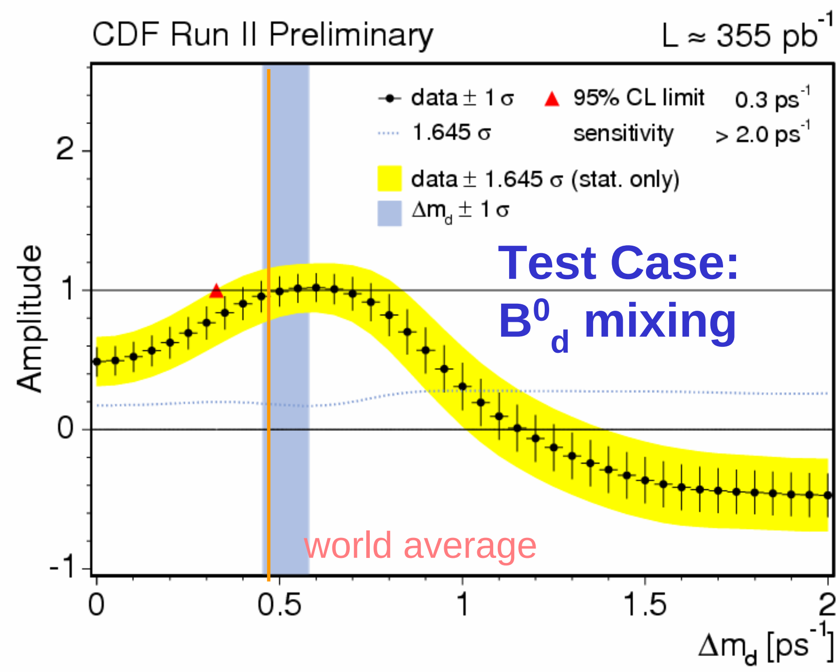
• Introduce Amplitude, A, to mixing probability formula

$$P_{unmix}^{B_s} = \frac{1}{2}\Gamma_{B_s} e^{-\Gamma_{B_s} t}\left(1 + A\cos \Delta m_s t\right)$$

$$P_{mix}^{B_s} = \frac{1}{2}\Gamma_{B_s} e^{-\Gamma_{B_s} t}\left(1 - A\cos \Delta m_s t\right)$$

- Evaluate A at each Δm point
- A=1 if evaluated at correct Δm
- This method facilitates limit setting before mixing signal observed

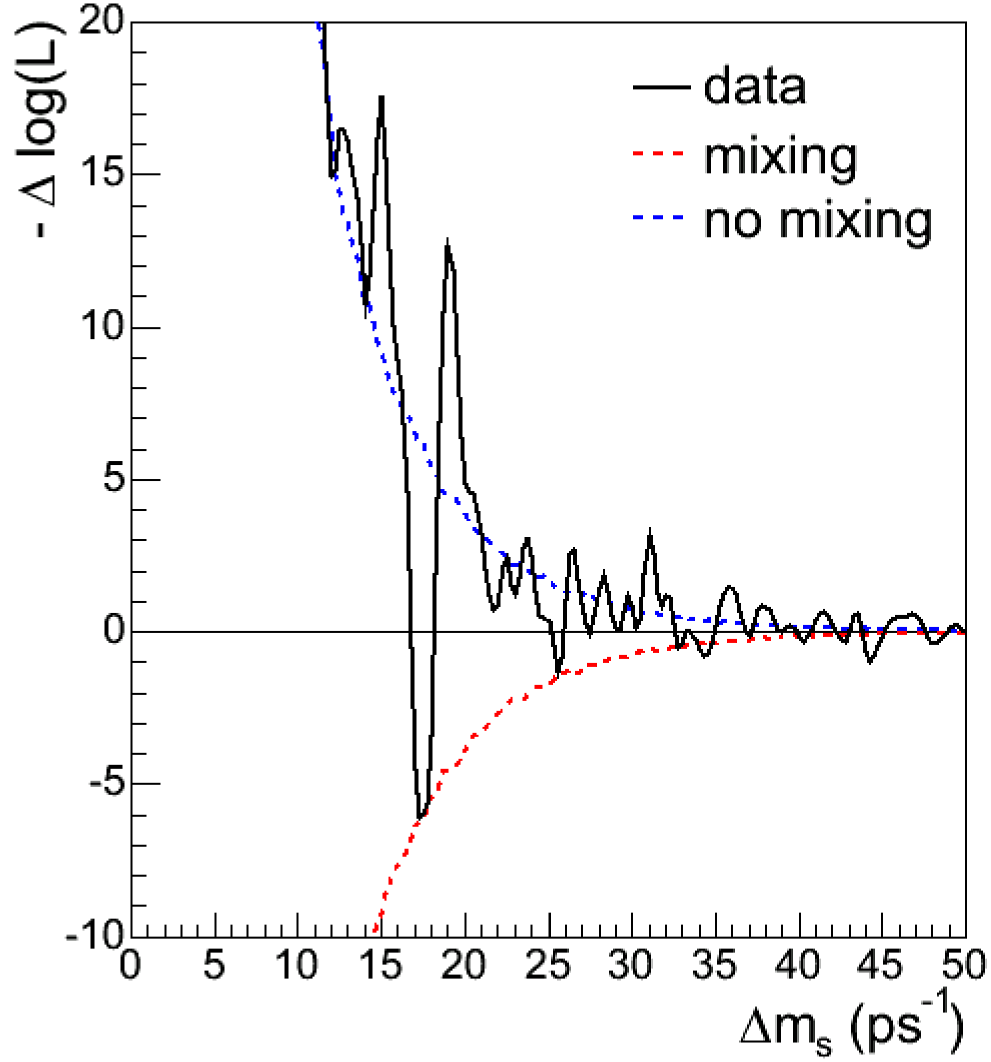
Mixing signal manifests itself as points in the plot which are most compatible with A=1



Measuring Δm_s

2: To establish the value of Δm_s, we evaluate the likelihood profile:

Log L(A=0)-Log L(A=1)



The Method

or

How do we get to the amplitude scan?

Mixing Ingredients

1) Signal samples
- semileptonic and hadronic modes

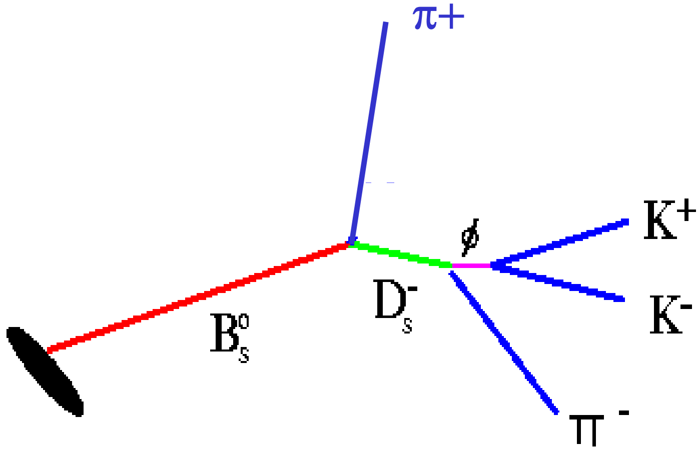
2) Time of Decay
- and knowledge of Proper decay time resolution

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

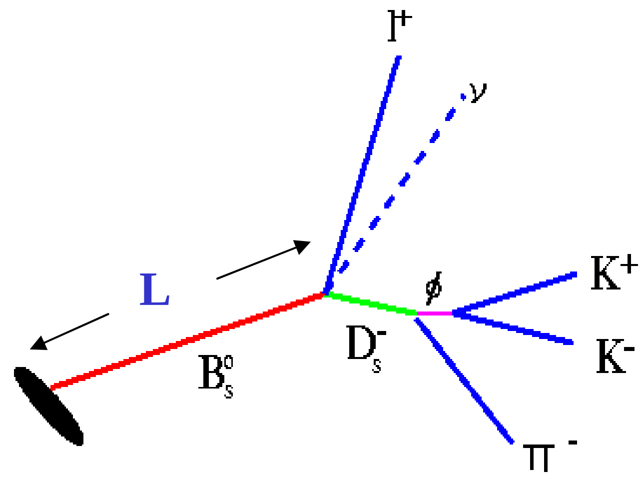
3) Flavour tagging
- opposite side (can be calibrated on B^0 and B^+)
- same side (cannot be calibrated on B^0 and B^+, used for the first time at CDF)

1) Signal Samples for B_s Mixing

Hadronic: fully reconstructed



Semileptonic: partially reconstructed

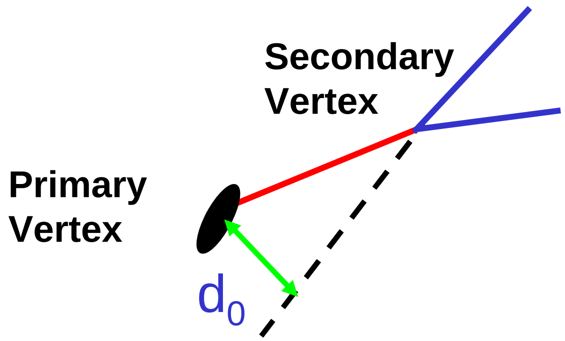


These modes are flavour specific: the charges tag the B at decay

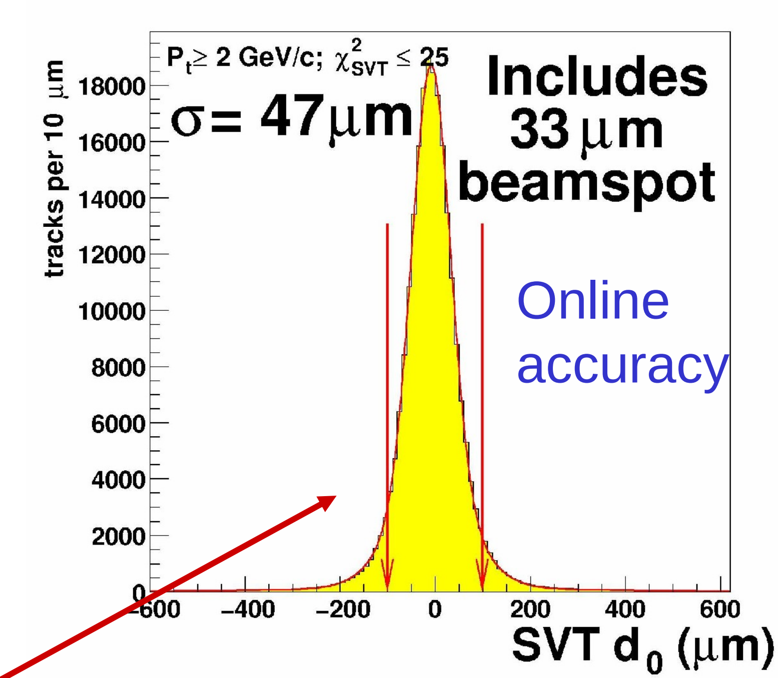
Crucial: Triggering using displaced track trigger
(Silicon Vertex Trigger)

Triggering On Displaced Tracks

- trigger $B_s \rightarrow D_s^- \pi$, $B_s \rightarrow D_s^- l^+$



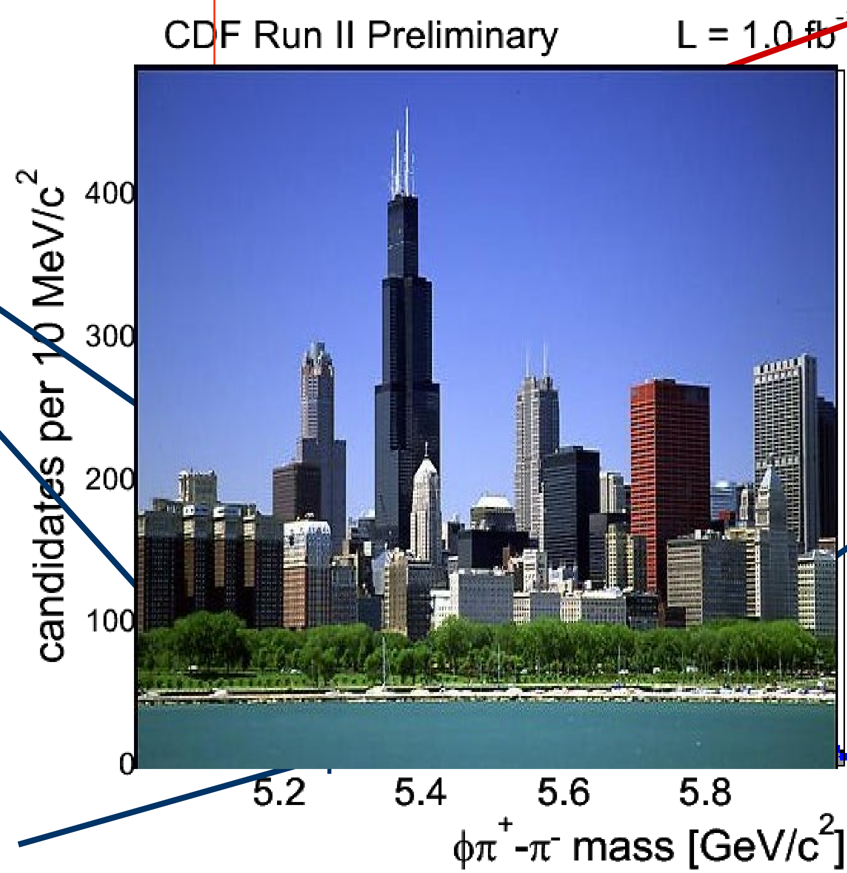
- trigger processes 20 TB /sec
- trigger requirement:
  - two displaced tracks:
    (p_T > 2 GeV/c, 120 μm<$|d_0|$<1mm)
- requires precision tracking in silicon vertex detector



Example Hadronic Mass Spectrum

Now we use the entire range, capitalising on satellites also

partially reconstructed B mesons (satellites)



signal

$B_s \rightarrow D_s \pi$,

$D_s \rightarrow \phi\pi$

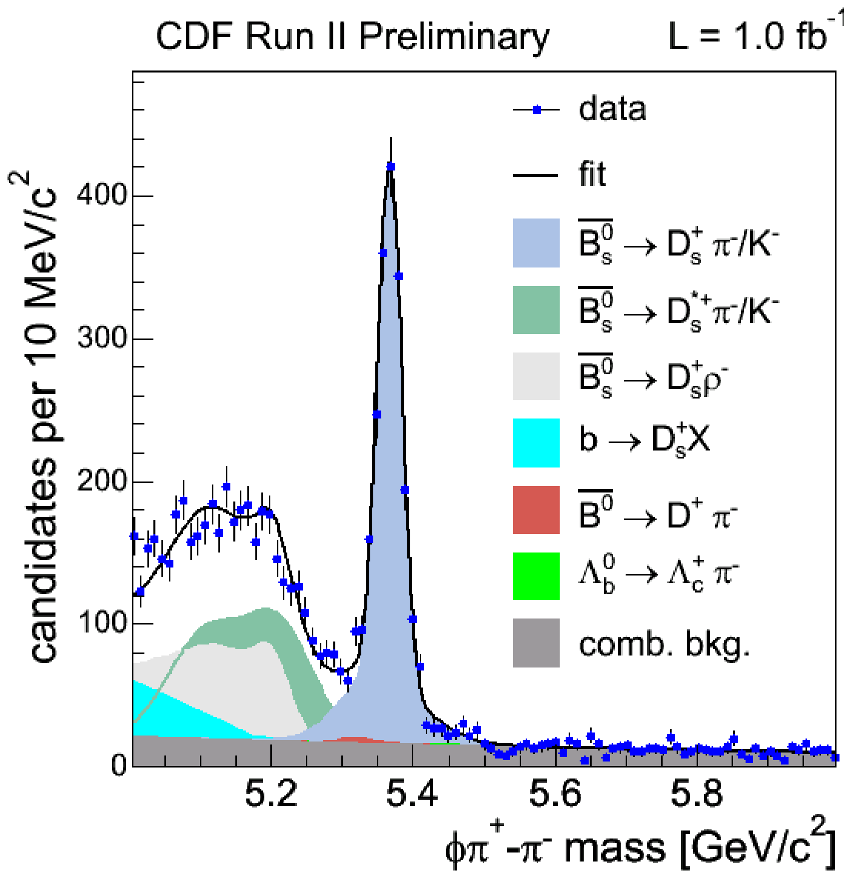
$\phi \rightarrow K^+K^-$

combinatorial background

$B^0 \rightarrow D^- \pi$ decays

Hadronic Signal Yields

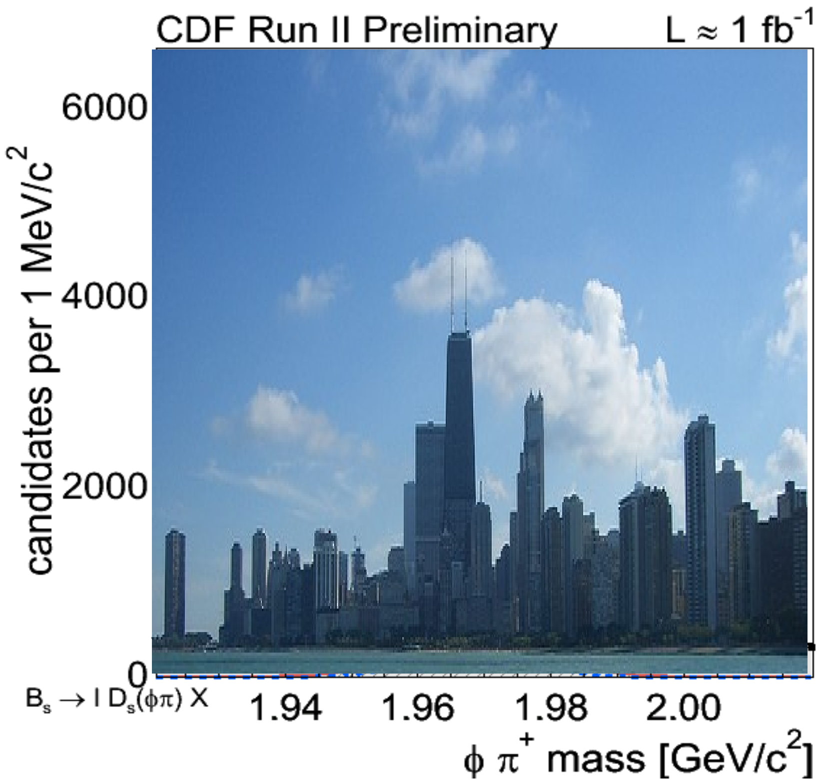
Decay Channel	Yield
$B_s \rightarrow D_s\pi$ ($\phi\pi$)	2000
Satellites	3100
$B_s \rightarrow D_s\pi$ ($K^* K$)	1400
$B_s \rightarrow D_s\ \pi$ (3π)	700
$B_s \rightarrow D_s3\pi$ ($\phi\ \pi$)	700
$B_s \rightarrow D_s3\pi$ (K^*K)	600
$B_s \rightarrow D_s3\pi$ (3π)	200
Total	8700



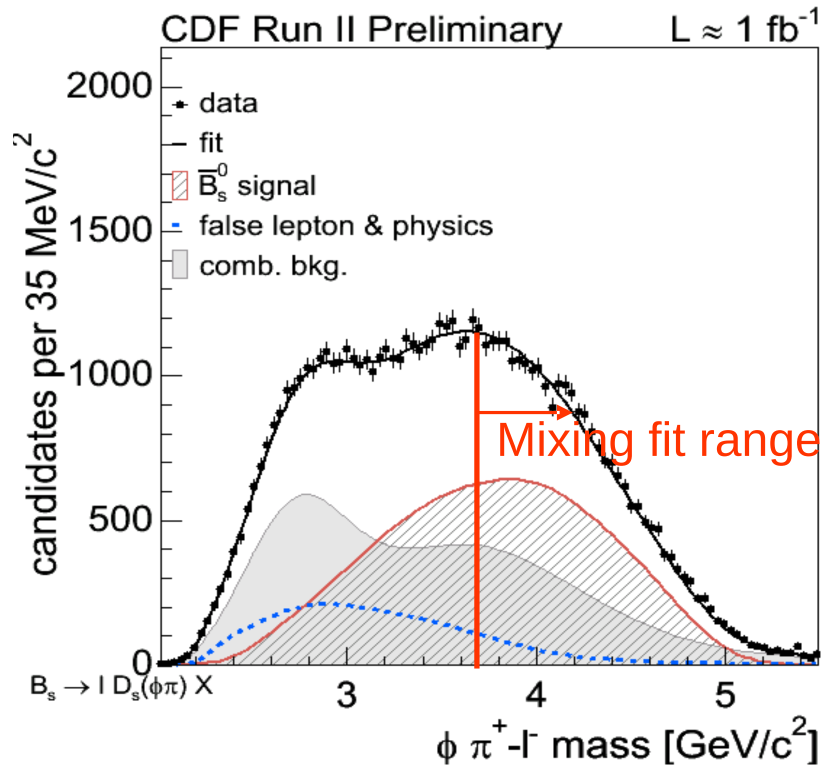
- Neural Network selection used in these modes
- Particle ID (dE/dx, Time of Flight) used to suppress backgrounds

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

Fully reconstructed D_s mesons:



B_s mesons not fully reconstructed:



Particle ID used; new trigger paths added → **61500 semileptonic candidates**

The candidate's $m(lD_s^-)$ is included in the fit: discriminates against "physics backgrounds" of the type $B^{0/+} \rightarrow D^+ D_s$

Summary of Yield changes since April 2006

$1fb^{-1}$ of data used in both analyses

What changed?

Hadronic modes:

- Added partially reconstructed “satellite” B_s decays
- Add Neural Net for candidate selection
- Used particle identification to eliminate background

Semileptonic Modes:

- Used particle identification to eliminate background
- Added new trigger path

Effective increase in statistics x2.5 from these changes

What do the candidates cost?: FECb

Tevatron Accelerator Value:	**$7M/year**
	($741M RPV at 70% spread over 25 years and 3 experiments)
CDF Detector Value:	**$0.8M/year**
	($95M total facilities RPV at 70% value)
Tevatron Operation to CDF:	**$48M/year**
	($120M/year at 40% of overall facilities)
CDF Operation:	**$5M/year**
Total CDF data	**$61M/year**
B Physics Program:	**$12M/year**
	(1/5 per physics group)

The B^0_sottom Line: **$** 850 Per B_s meson

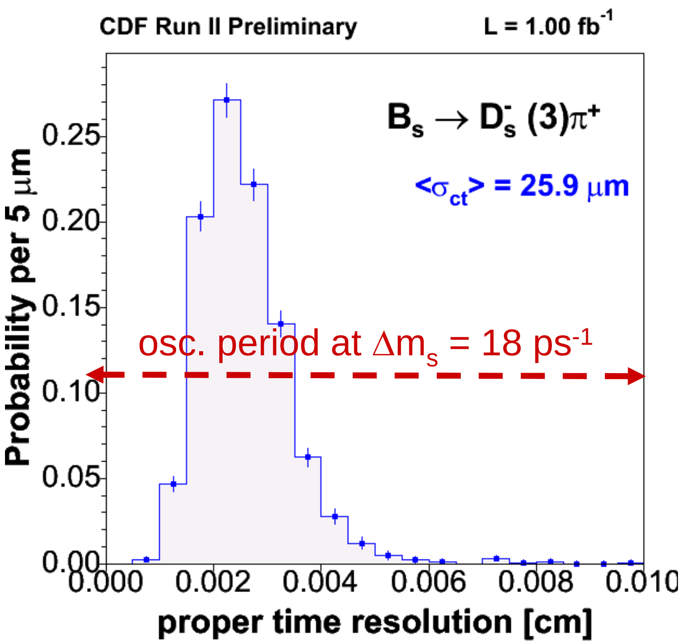
2) Time of Decay

- Reconstruct decay length by vertexing
- Measure p_T of decay products

$$ct = \frac{L}{\beta\gamma} = L\frac{m(B)}{p(B)} = \frac{L_{xy}m(B)}{p_T(lD)}$$

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

Proper time resolution:



Semileptonic:

$\sigma_{ct}^0 \approx 59\mu m$

$\sigma_p / p \approx 15\%$

Hadronic:

$\sigma_{ct}^0 \approx 30\mu m$

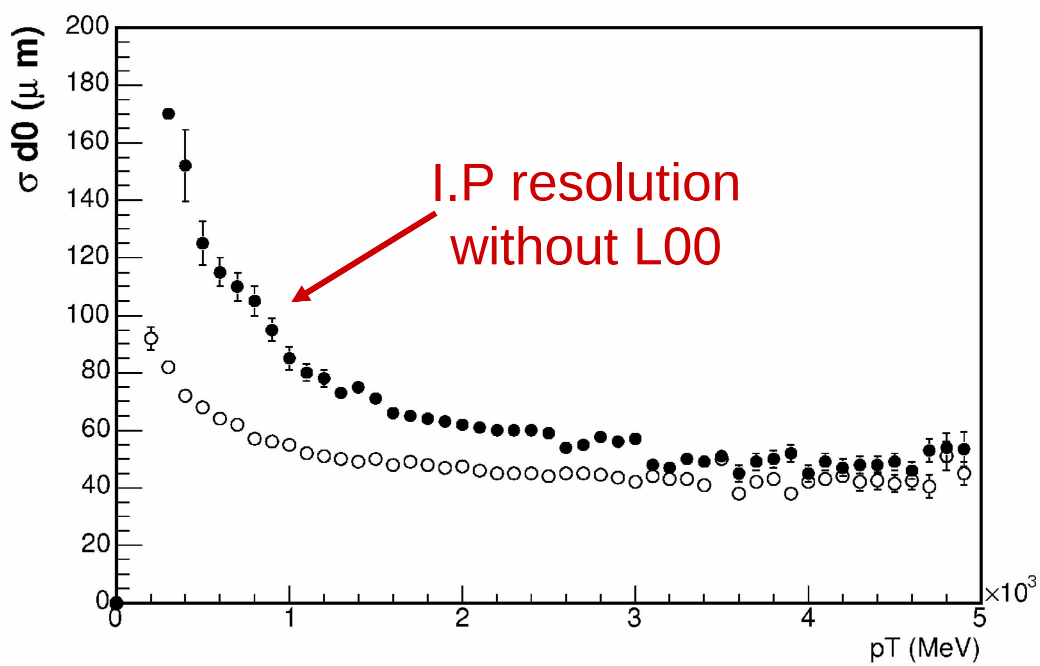
$\sigma_p / p \approx 0\%$

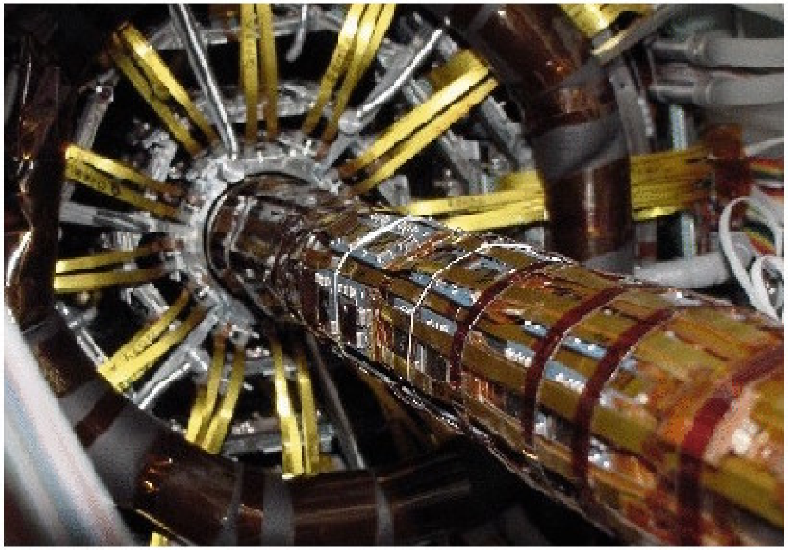
Crucial: Vertex resolution

(Silicon Vertex Detector, in particular Layer00 very close to beampipe)

Layer 00

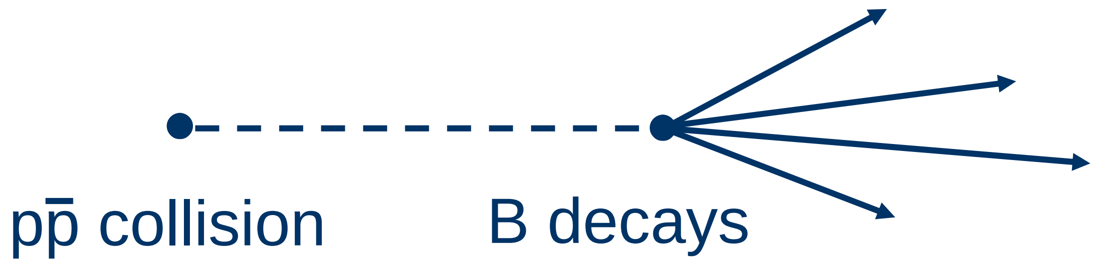
- So-called because we already had layer 0 when this device was designed!
- UK designed, built and (mostly) paid for this detector!





- **layer of silicon placed directly on beryllium beam pipe**
- **Radius of 1.5 cm**
- **additional impact parameter resolution**

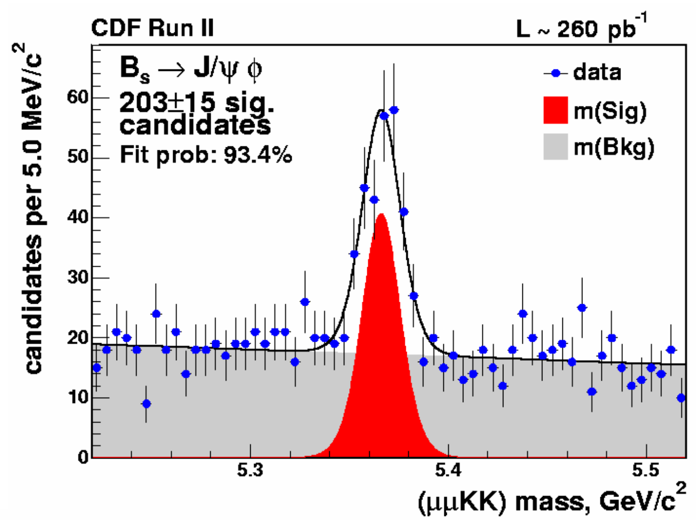
Classic B Lifetime Measurement

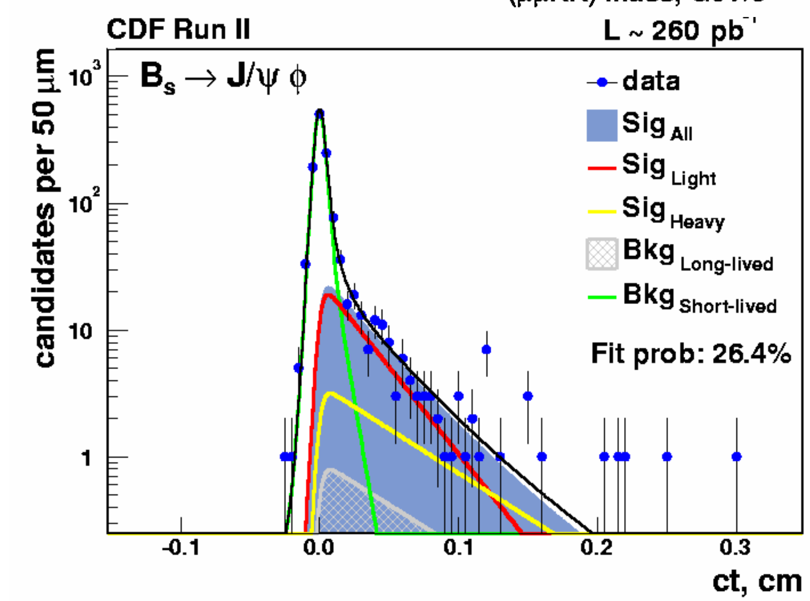


- reconstruct B meson mass, p_T, L_{xy}
- calculate proper decay time (ct)
- extract $c\tau$ from combined mass+lifetime fit
- signal probability:

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

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





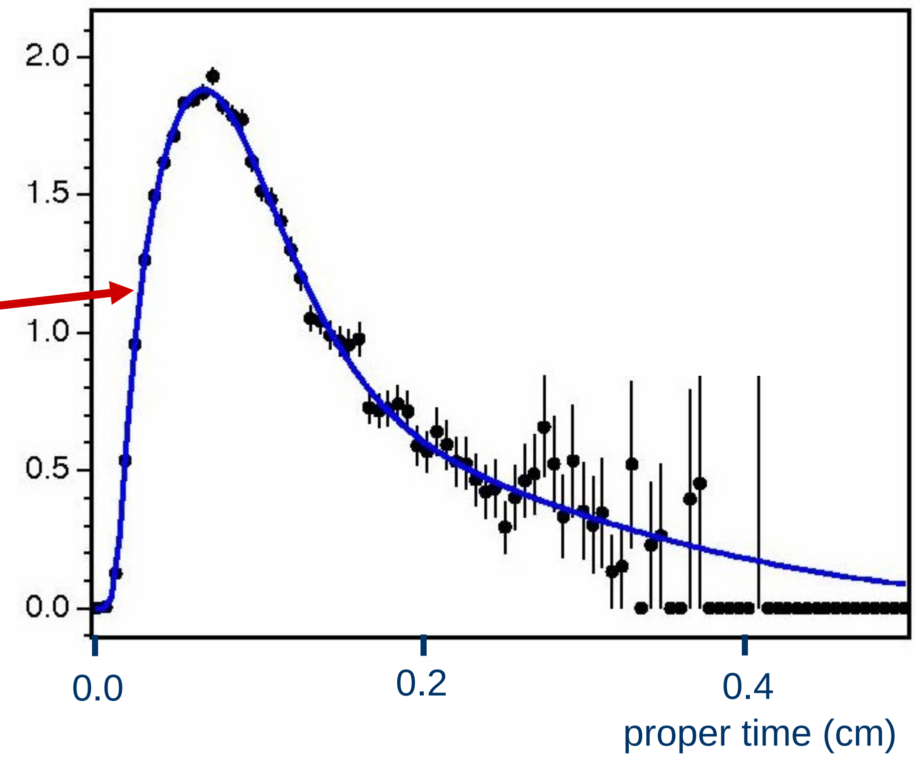
Hadronic Lifetime Measurement

- Displaced track trigger biases the lifetime distribution

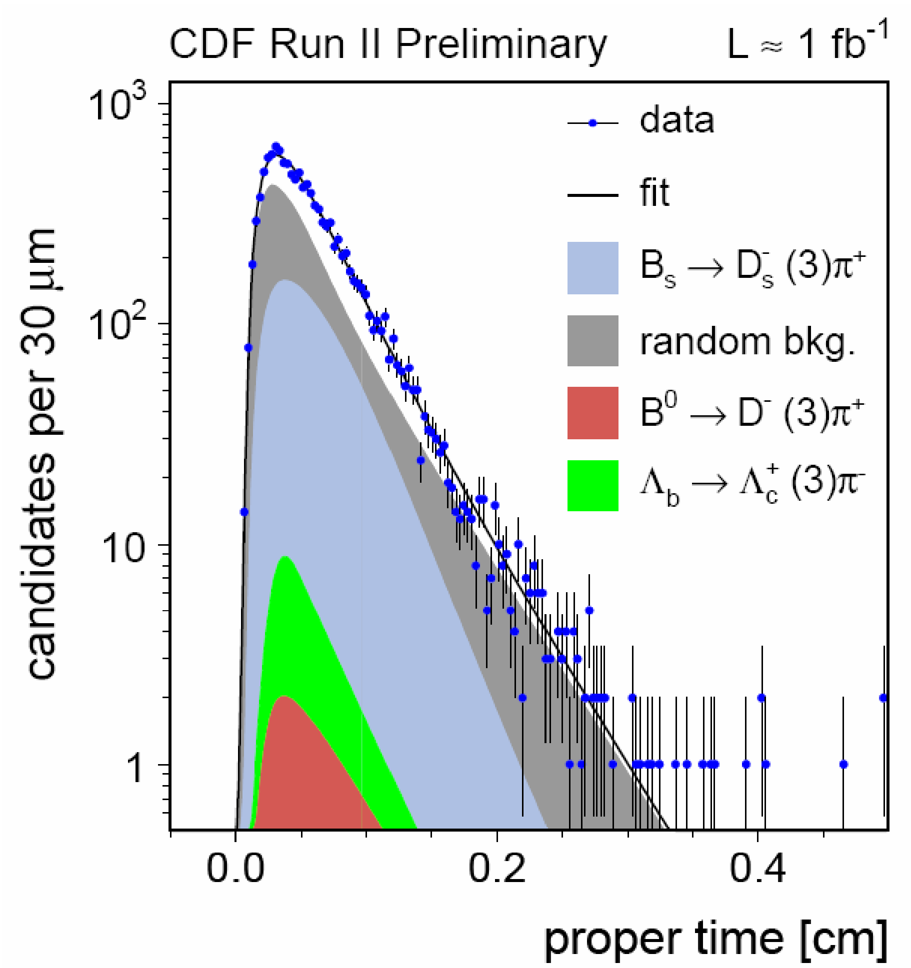
- Correct with an

  efficiency function derived from MC:

$$p = e^{-t'/\tau} \otimes R(t',t) \otimes \varepsilon(t)$$



Hadronic Lifetime Measurements



Mode	Lifetime (ps)
$B^0 \to D^- \pi^+$	1.508± 0.017
$B^- \to D^0 \pi^-$	1.638 ± 0.017
$B_s \to D_s \pi(\pi\pi)$	1.538 ± 0.040

Errors are statistical only

World Averages:

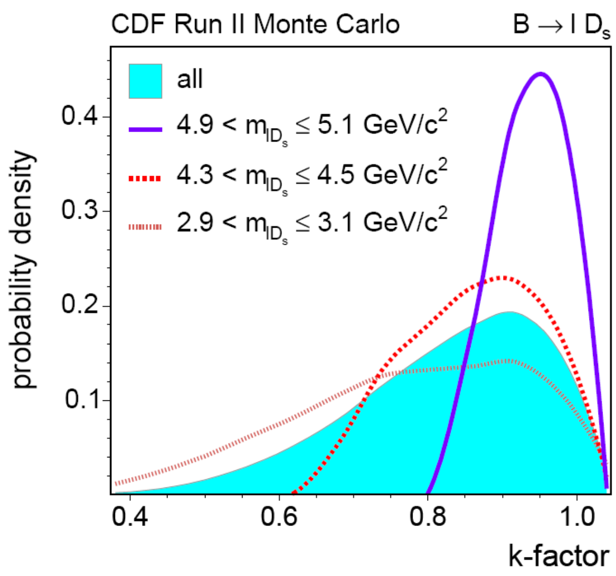
B^0 : 1.534 ± 0.013 ps
B^- : 1.653 ± 0.014 ps
B_s : 1.469 ± 0.059 ps

Good agreement in all modes

Semileptonic Lifetime Measurement

- neutrino momentum missing
- Correct with "K factor" from MC:

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



High m(lD) candidates have narrow K factor distribution: almost fully reconstructed events!

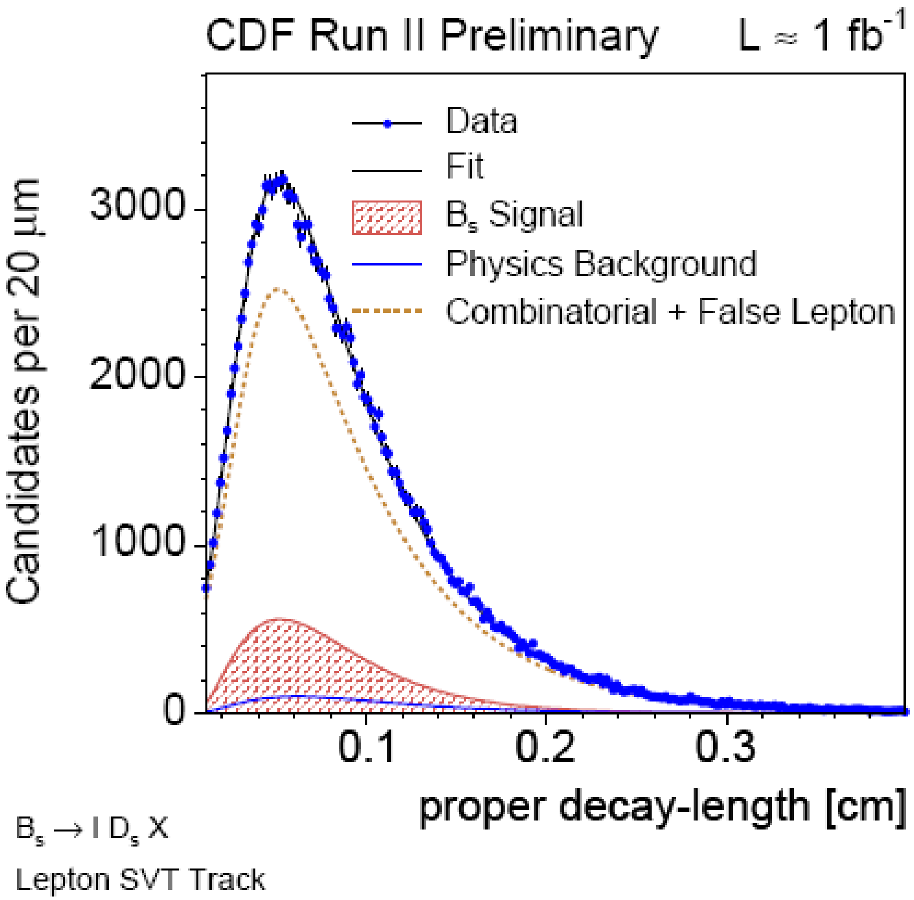
Capitalise on this by binning K factor in m(lD)

- Also correct for displaced track trigger bias as in hadronic case

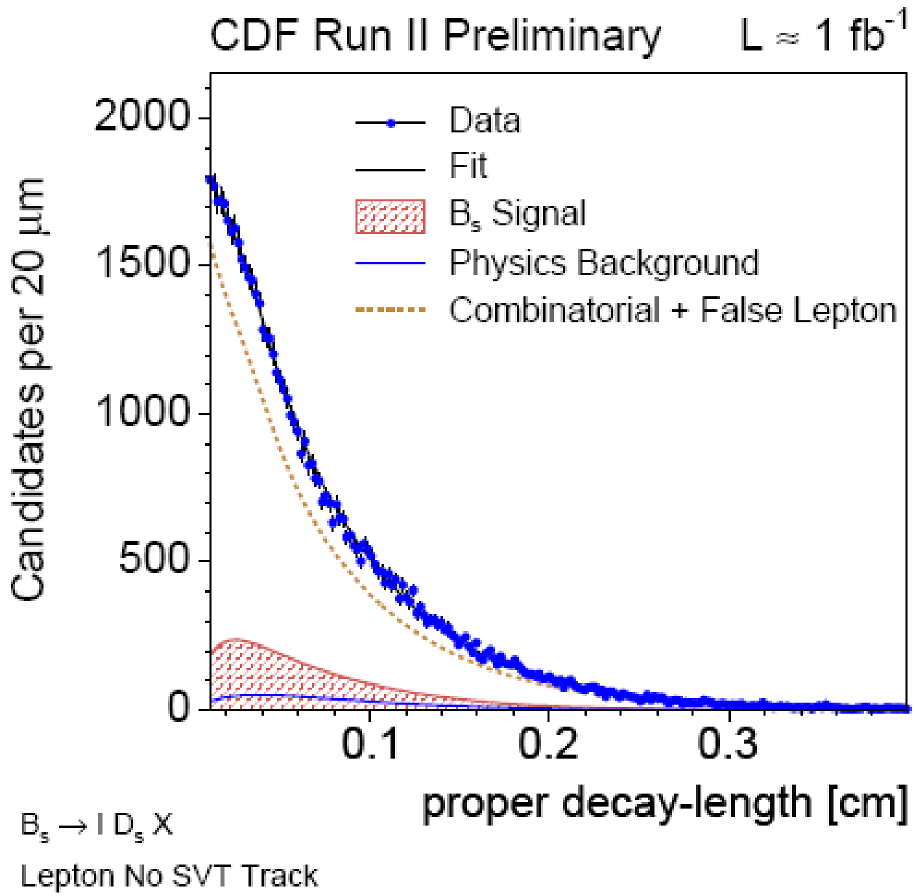
Lepton+D_s Lifetime Fits

Two cases treated separately:

Lepton is a displaced track:



Lepton is not a displaced track:



Semileptonic Lifetime Results

	Lifetime (ps)
B_s:$D_s \to \phi\pi$	1.51 ± 0.04
B_s:$D_s \to K^*K$	1.38 ± 0.07
B_s:$D_s \to \pi\pi\pi$	1.40 ± 0.09
B_s combined	1.48 ± 0.03

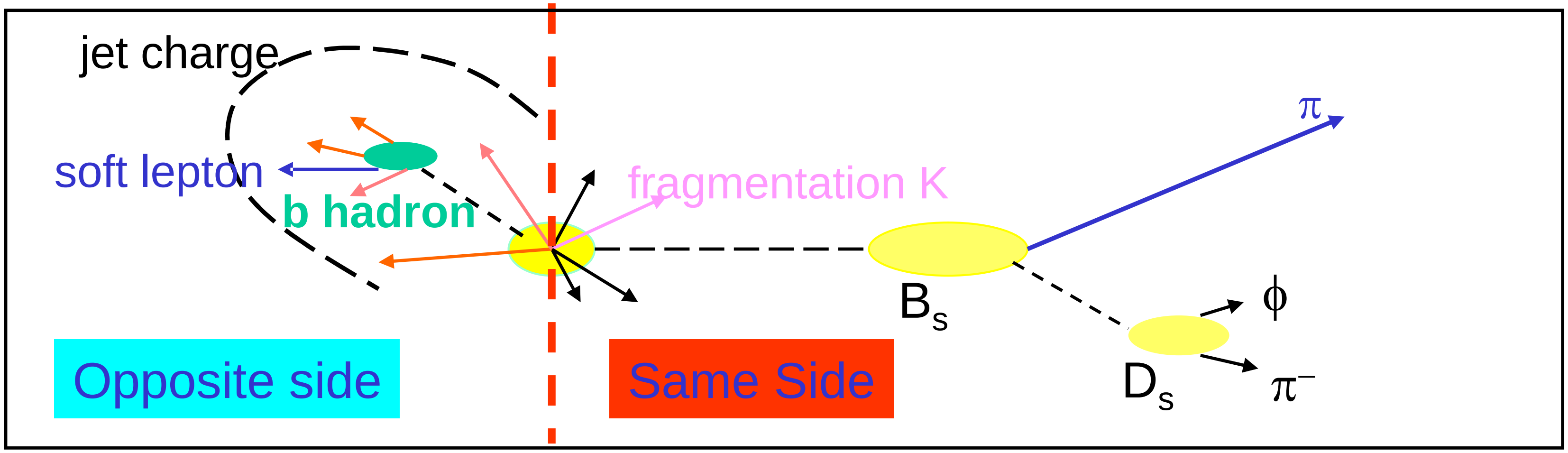
- Errors are statistical only
- Lifetimes measured on first 355 pb^{-1}
- Compare to World Average: B_s: (1.469 ± 0.059) ps

- All Lifetime results are consistent with world average
- Gives confidence in fitters, backgrounds, ct resolution

3) Flavour Tagging

To determine B flavour at production, use tagging techniques:

b quarks produced in pairs ⇒ only need to determine flavour of one of them



OPPOSITE SIDE
Soft Muon Tag } semileptonic BR ~10%
Soft Electron Tag
Jet charge tag sum of charges in jet
εD^2 = 1.82±0.04 % (semileptonic)
1.81±0.10 % (hadronic)

SAME SIDE
Same Side K Tag

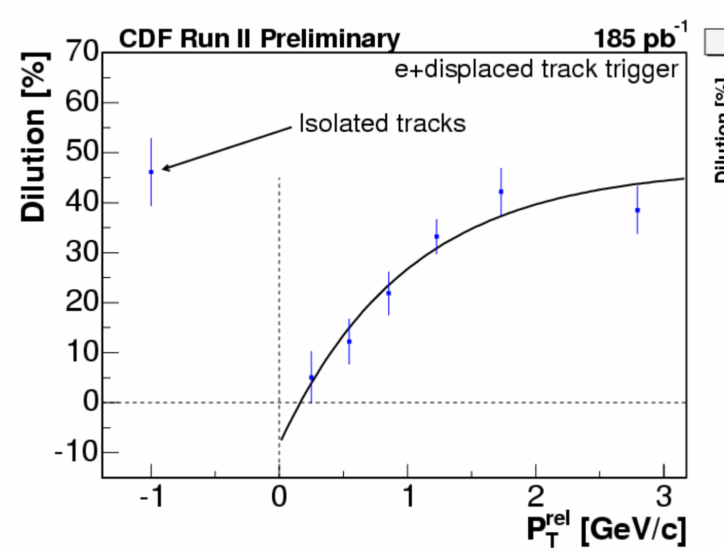
εD^2 = 4.8±0.04 %(semileptonic)
3.5±0.06 % (hadronic)

Figure of merit is εD^2 ε = efficiency (% events tagger can be applied)
D = dilution (% events tagger is correct)

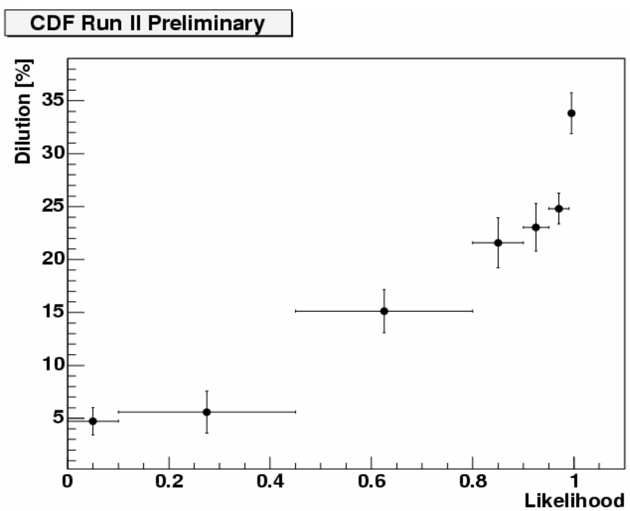
Crucial: Particle Identification (Time of Flight Detector)

Opposite Side Taggers

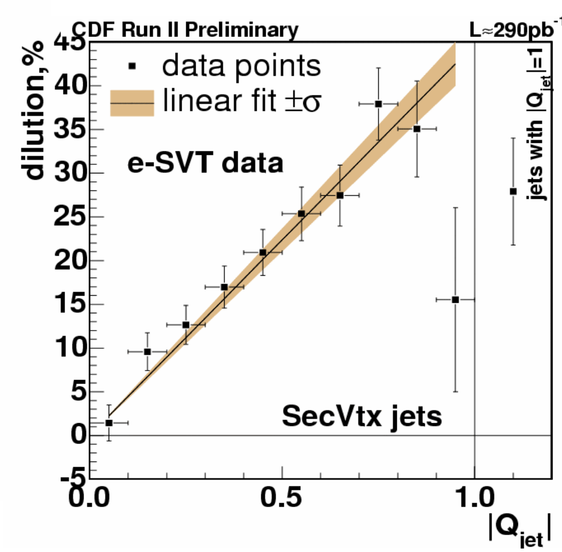
- Performance studied in high statistics inclusive lepton+SVT trigger
  - Enables calibration of taggers
  - Can also parameterise tagging dilution as function of variables:
  - Soft Lepton Tag: dilution parameterised as function of likelihood and p_t^{rel}
  - Jet Charge Tag: dilution parameterised as function of jet charge for a given jet



Soft Electron Tag



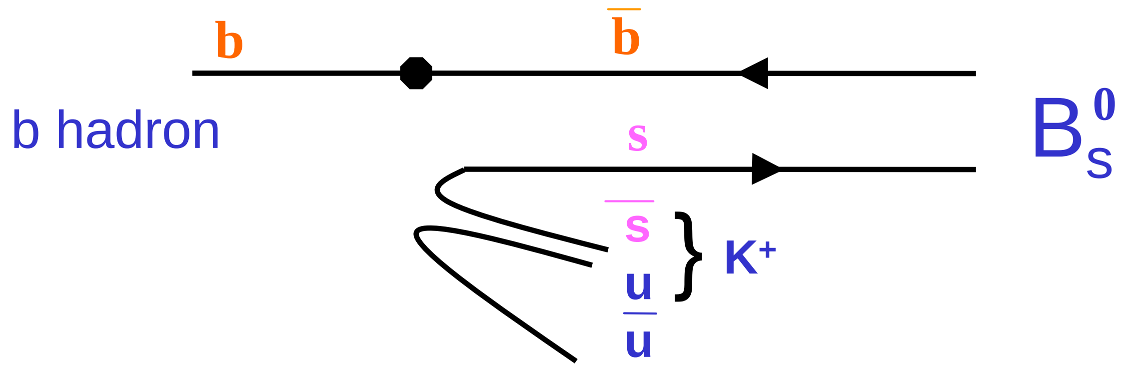
Soft Electron Tag



Jet Charge Tag

Same Side (Kaon) Tagger

- This is the first time this type of tagger has been implemented
- Principle: charge of B and K correlated



- Use TOF, dE/dx to select track
- Tagger εD^2 not measurable in data until B_s mixing frequency known

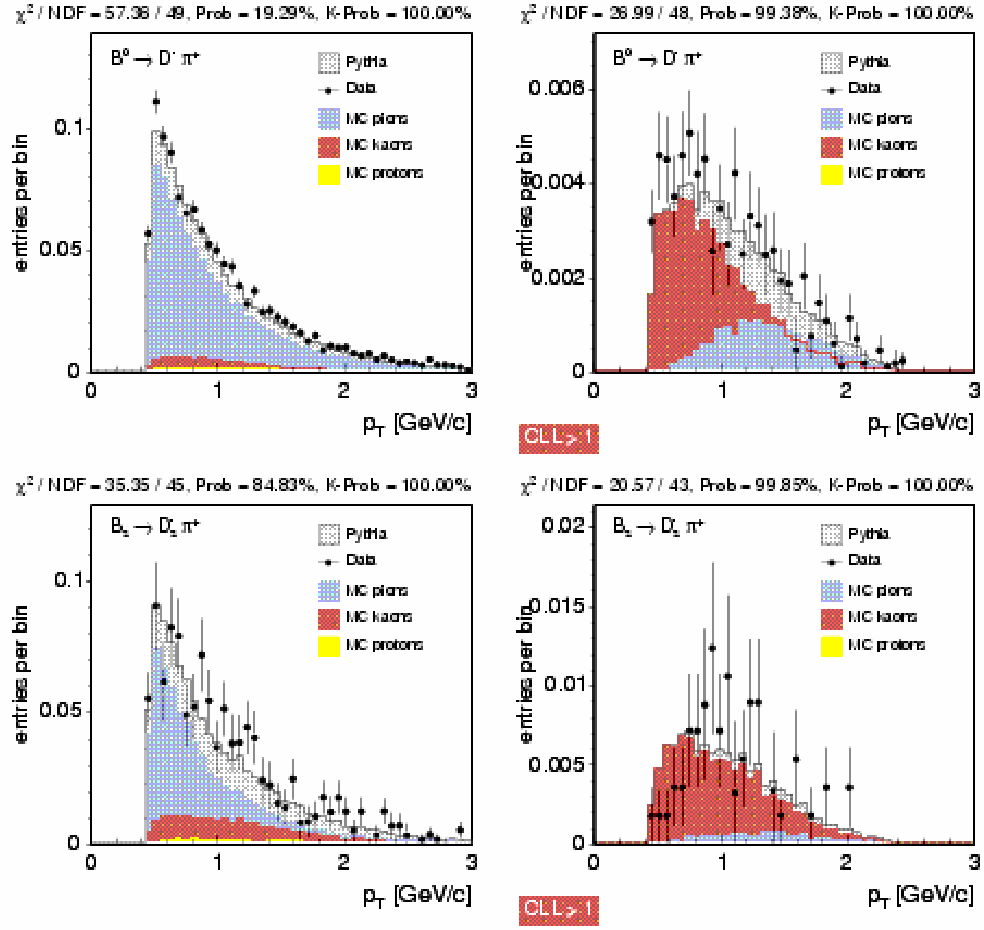
Same Side (Kaon) Tagger

CDF Public Note 8206

- If MC reproduces distributions well for B^0,B^+, then rely on it to extract tagger power in B_s (with appropriate systematic errors)
- High statistics B^0 and B^+ samples in which to make data/MC comparisons:

B^0_d

B^0_s



Kaon enhanced

- Systematics: production mechanism, fragmentation model, particle fraction around B, PID simulation, pile-up, MC/data agreement

Summary of Tagging changes since April 2006

What changed?

Opposite Side Taggers:

- Added new tagger: Opposite Side Kaon Tagger
- New method to combine opposite side tags
  - Before, it was hierarchical
  - Now combination is performed by neural net
  - Every tagger can contribute some power

Same Side Kaon Tagger:

- Neural Net used to incorporate kinematic information as well as particle identification

The Results

Put the 3 Ingredients Together

- Amplitude scan performed on B_s candidates
- Inputs for each candidate:
  - Mass
  - Decay time
  - Decay time resolution
  - Tag decisions
  - Predicted dilution
  - Mass(lepton+D) if semileptonic

- All elements are then folded into the amplitude scan

$$\frac{1}{\tau} e^{-t/\tau} \left(1 \pm ADS_D \cos\left(\Delta m t\right)\right)$$

"With three parameters, I can fit an elephant." (Kelvin)

A Priori Procedure

Decided upon before un-blinding the data:
(everything blinded so far by scrambling tagger decision)

- Find highest significant point on amplitude scan consistent with an amplitude of 1
- significance to be estimated using Δ(log Likelihood) method
- effectively infinite Δm_s search window to be used

Is probability for "signal" to be a fluctuation < 1%?

YES

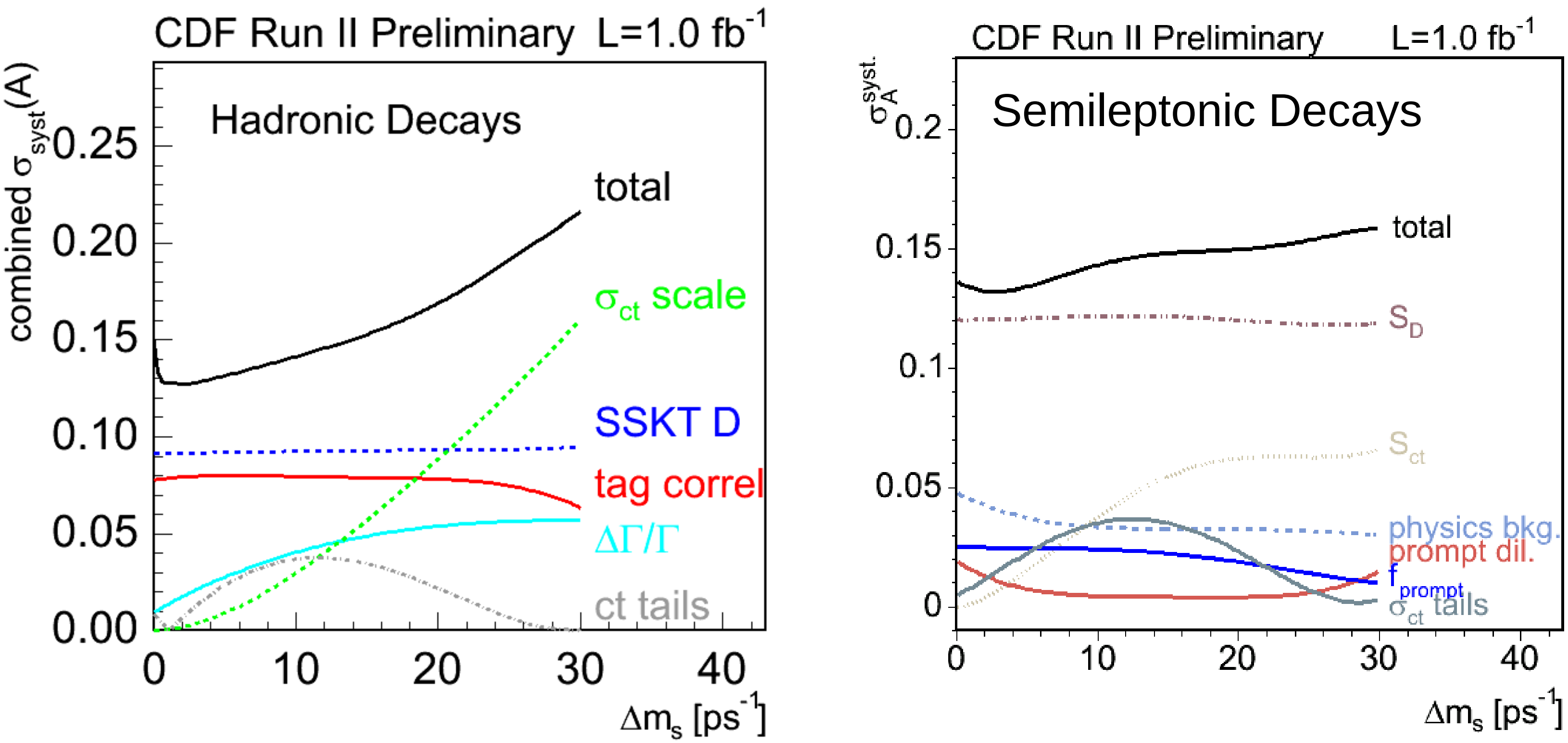
make double-sided confidence interval from Δlog(Likelihood)

Measure Δm_s

NO

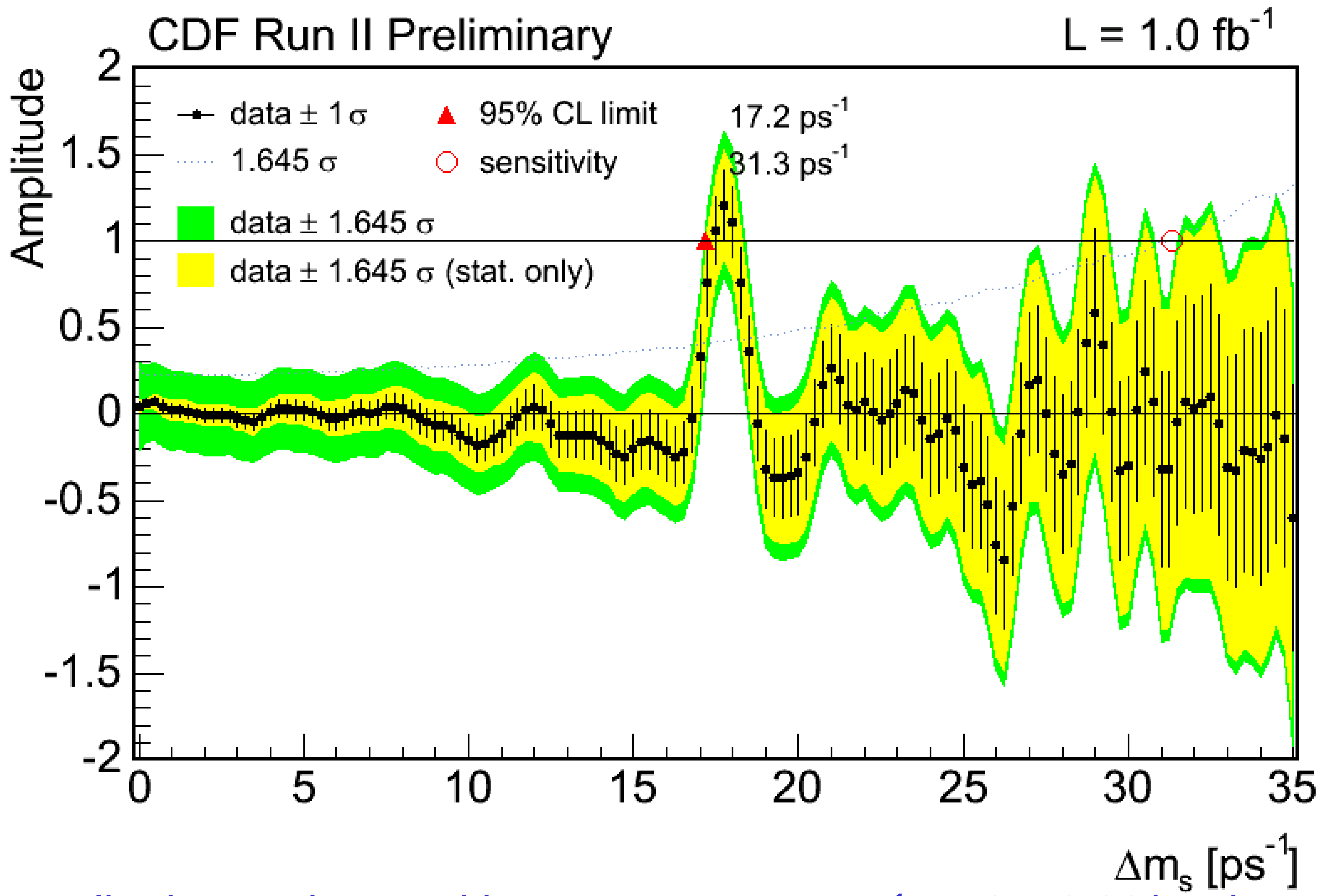
(Since we already had <1% probability in April we weren't expecting to follow this route in September with the improved analysis)

Systematic Uncertainties



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

Combined Amplitude Scan

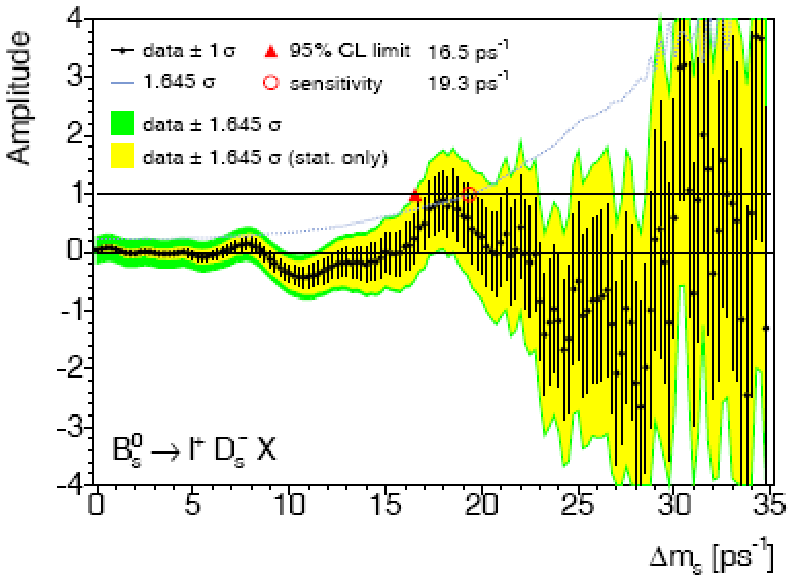


Amplitude consistent with 1 at Δm_s ~17.75ps^{-1}: 1.21±0.20(stat) (and inconsistent with 0)

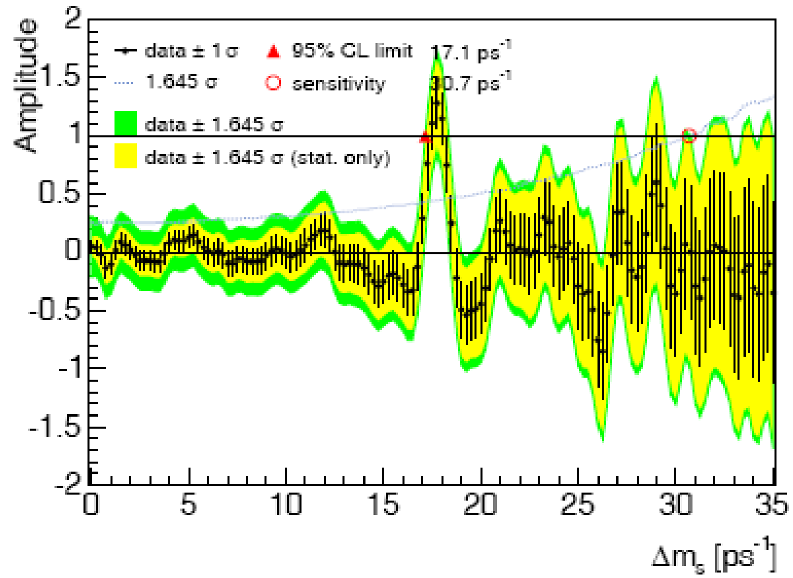
How significant is this result?

Separate Samples

Semileptonic



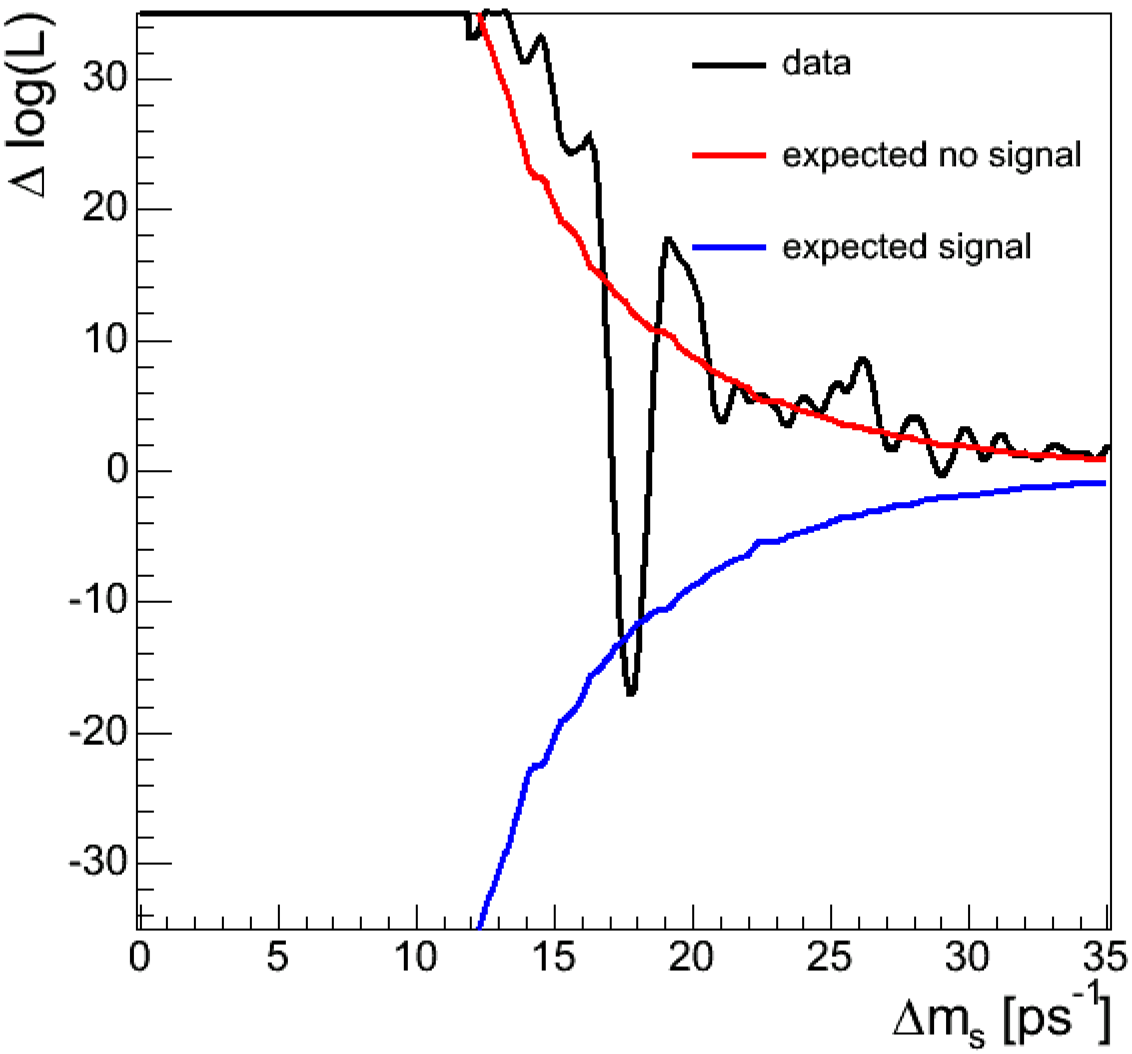
Hadronic



World best semileptonic analysis with sensitivity of 19.3ps-1

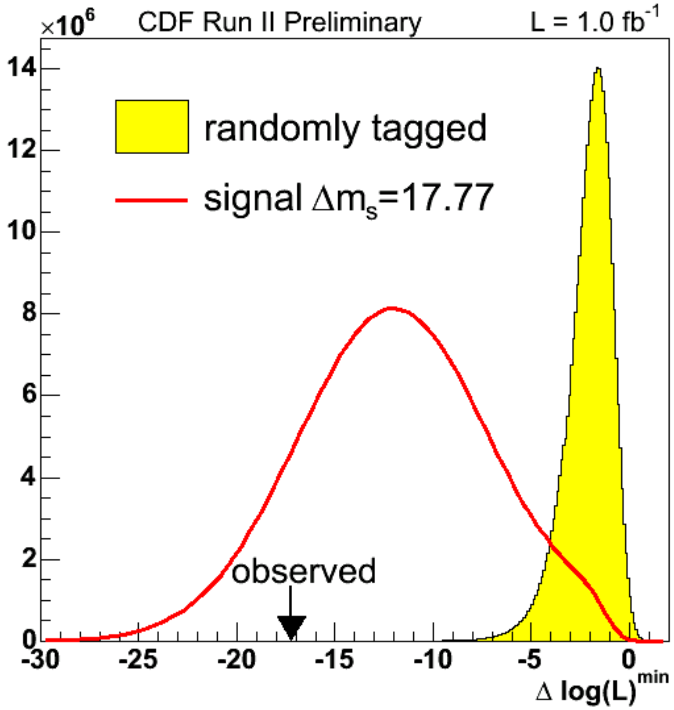
…but the hadronic analysis gives a clear signature of mixing even on its own!

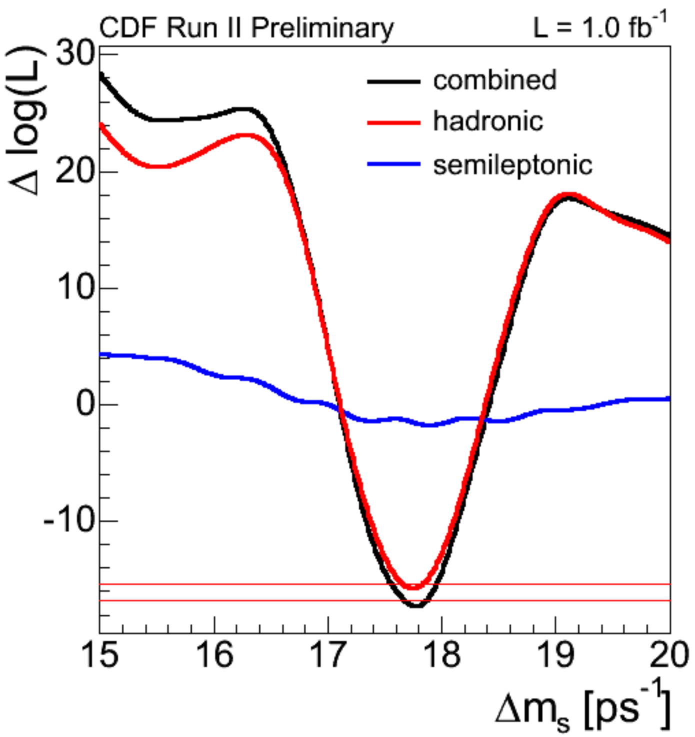
Likelihood Ratio Profile



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

Likelihood Significance





- probability of fake from random tags = $8\text{x}10^{-8}$ $\Rightarrow$measure Δm_s
- Equivalent to 5.4σ significance

Δm_s = 17.77±0.10(stat)±0.07(syst) ps^{-1}

Systematic Uncertainties on Δm_s

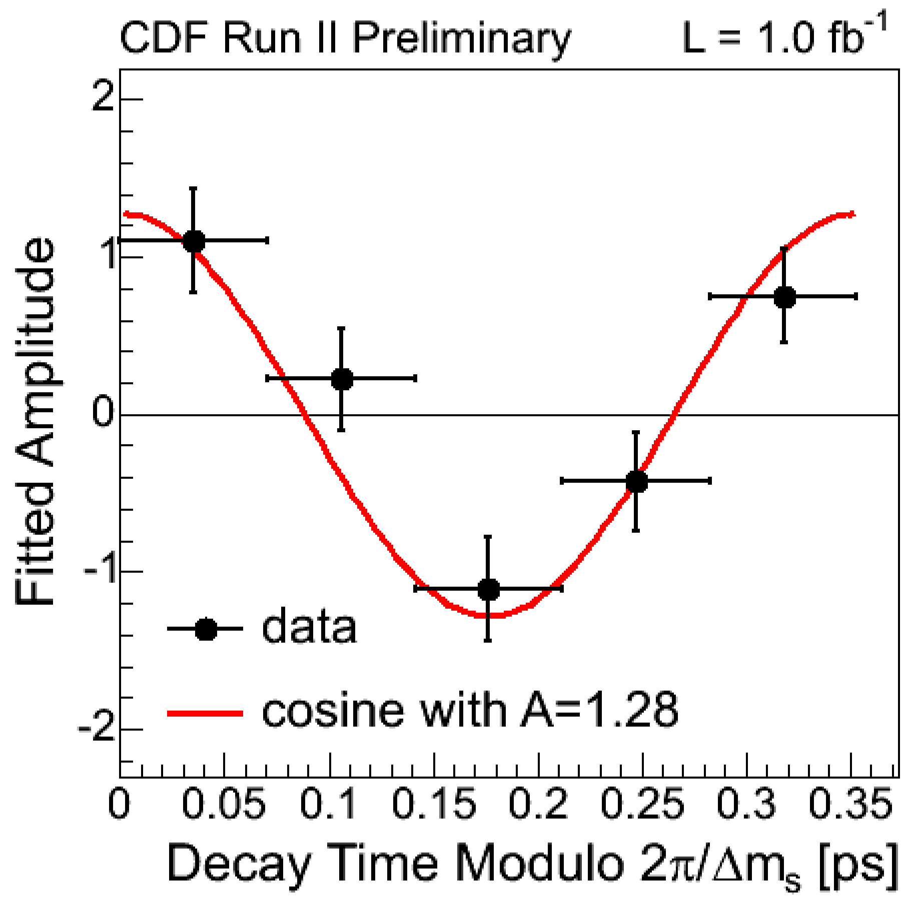
- Systematic uncertainties from fit model evaluated on toy Monte Carlo
- Have negligible impact
- Relevant systematic uncertainties are from lifetime scale

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

All systematic uncertainties are common between hadronic and semileptonic samples

Asymmetry

Oscillations folded modulo $2\pi/\Delta m_s$



$|V_{ts}| / |V_{td}|$

- Can extract V_{ts} value

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

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

$$|V_{td}| / |V_{ts}| = 0.199 \,^{+0.026}_{-0.025}\,(\text{exp}) \,^{+0.018}_{-0.016}\,(\text{theo})$$

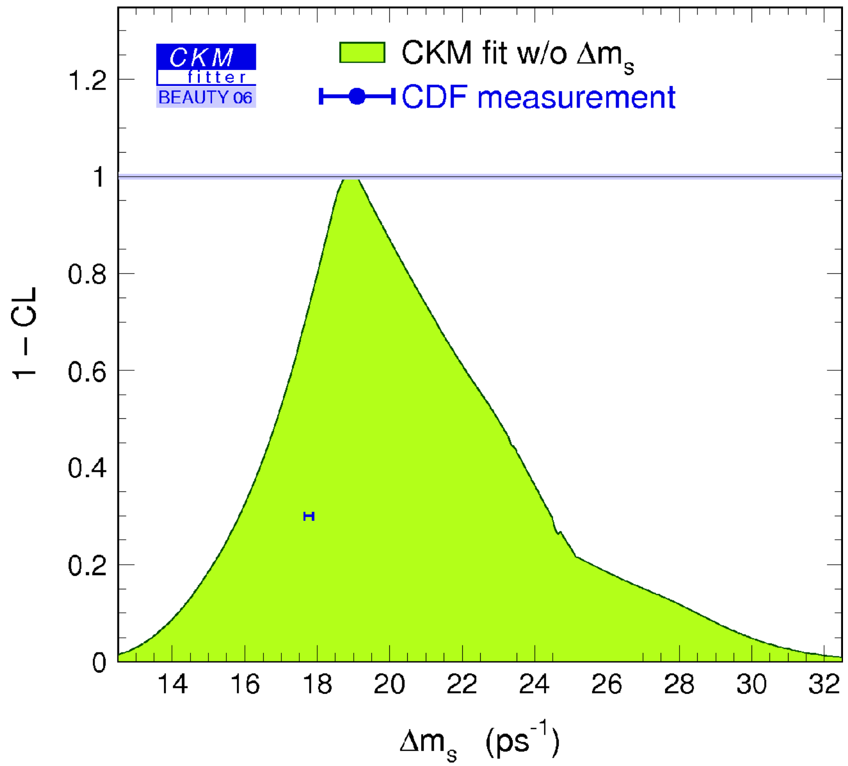
- our result:

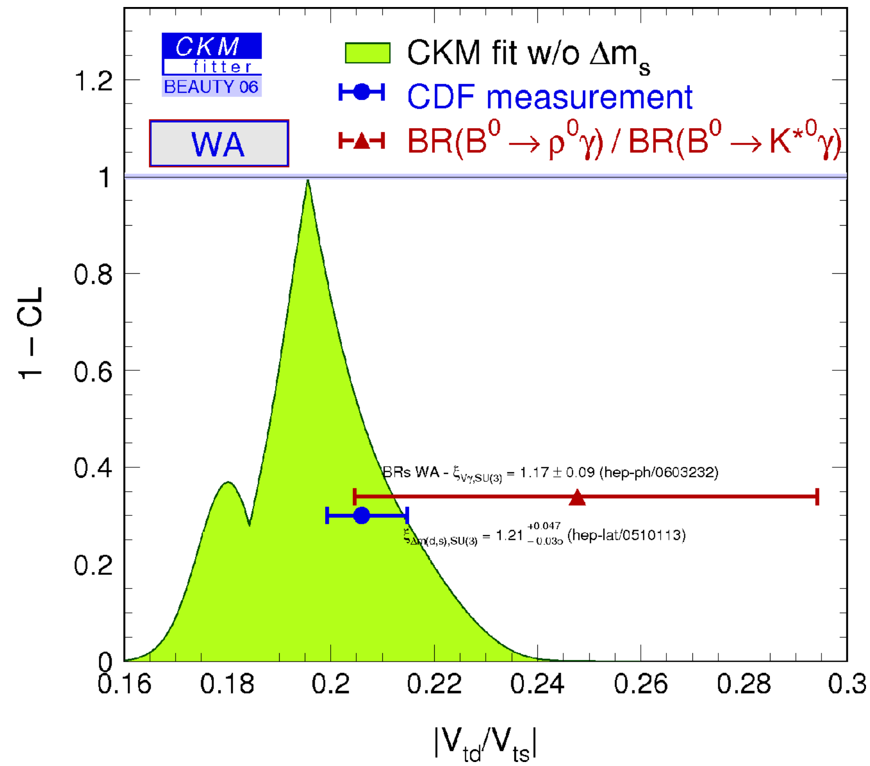
$$|V_{td}| / |V_{ts}| = 0.2060 \pm 0.0007\,(\text{exp}) \,^{+0.0081}_{-0.0060}\,(\text{theo})$$

- inputs:
  - $m(B^0)/m(B_s) = 0.9832$ (PDG 2006)
  - $\xi = 1.21 \,^{+0.05}_{-0.04}$ (Lattice 2005)
  - $\Delta m_d = 0.507 \pm 0.005$ (PDG 2006)

Interpretation of Results

Measurements compared with global fit (CKM fitter group) updated this month

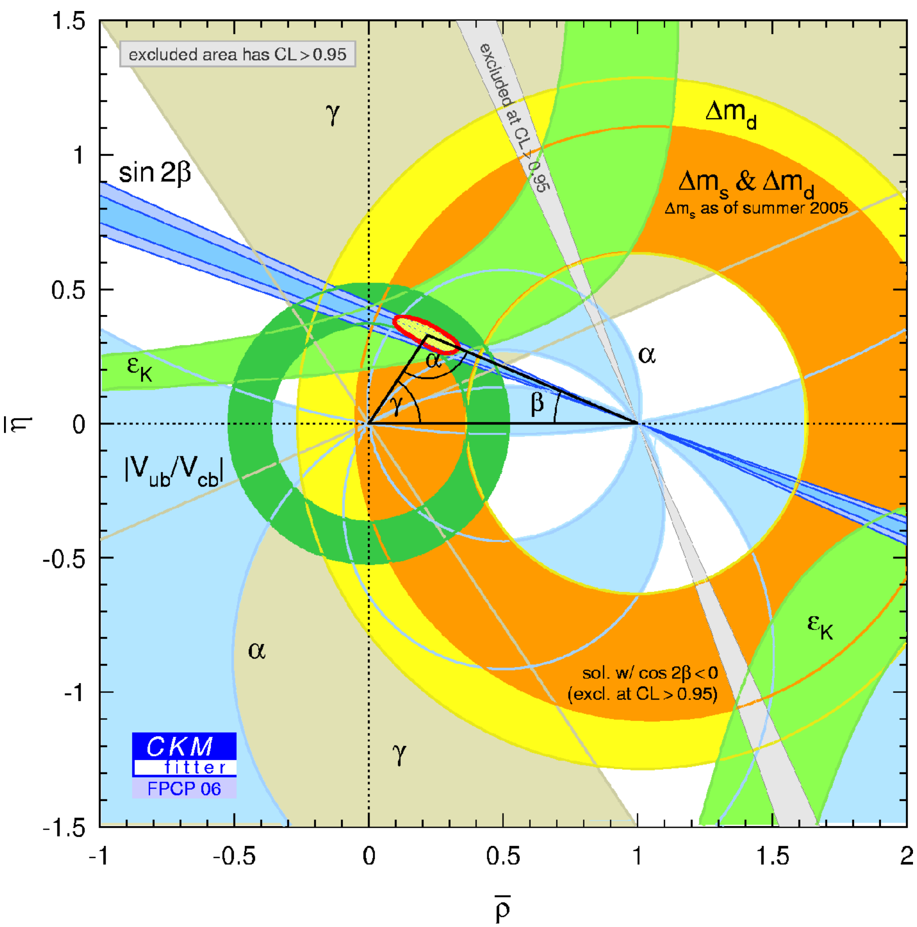


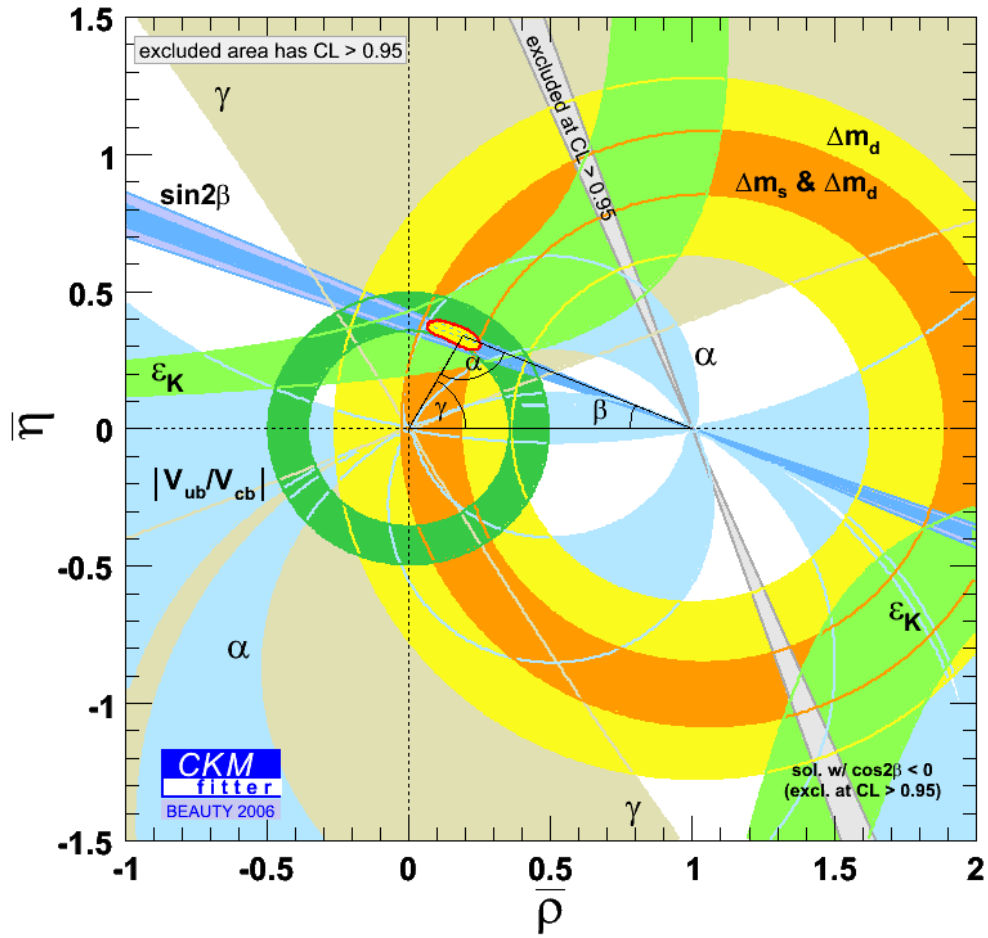


In excellent agreement with expectations

Interpretation of Results

This measurement decreases uncertainty on CKM triangle apex:





Easter 2006 → October 2006

Conclusions

- CDF has found a signature consistent with B_s - $\overline{B}_s$ oscillations
- Probability of this being a fluctuation is 8×10^{-8}
- Presented direct measurement of the B_s - $\overline{B}_s$ oscillation frequency:

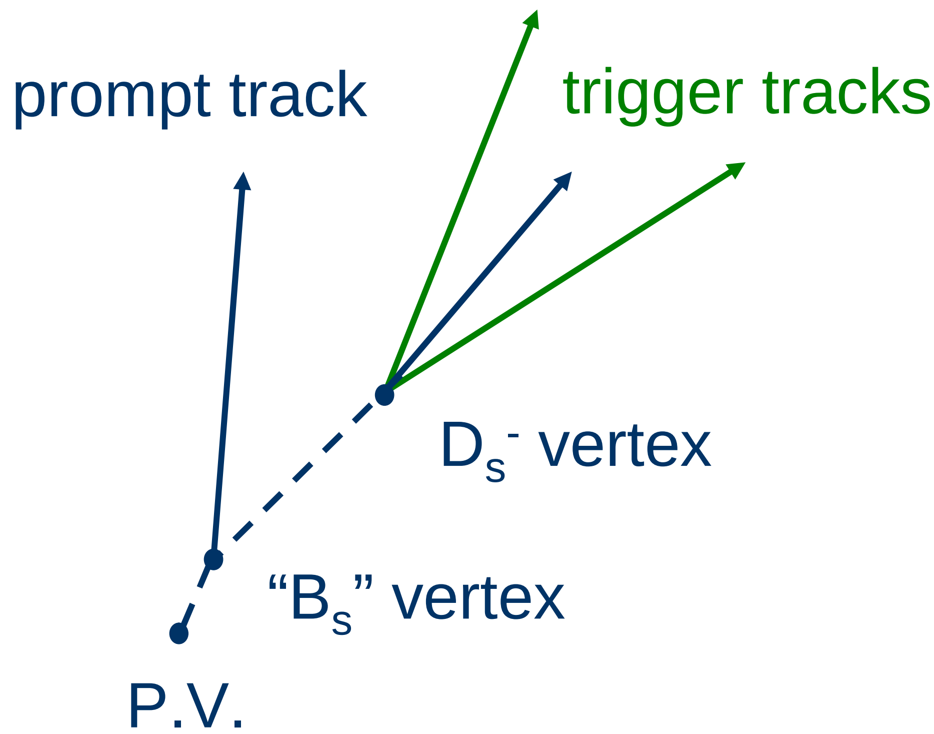
$$\Delta m_s = 17.77 \pm 0.10(\text{stat}) \pm 0.07(\text{syst})\ \text{ps}^{-1}$$

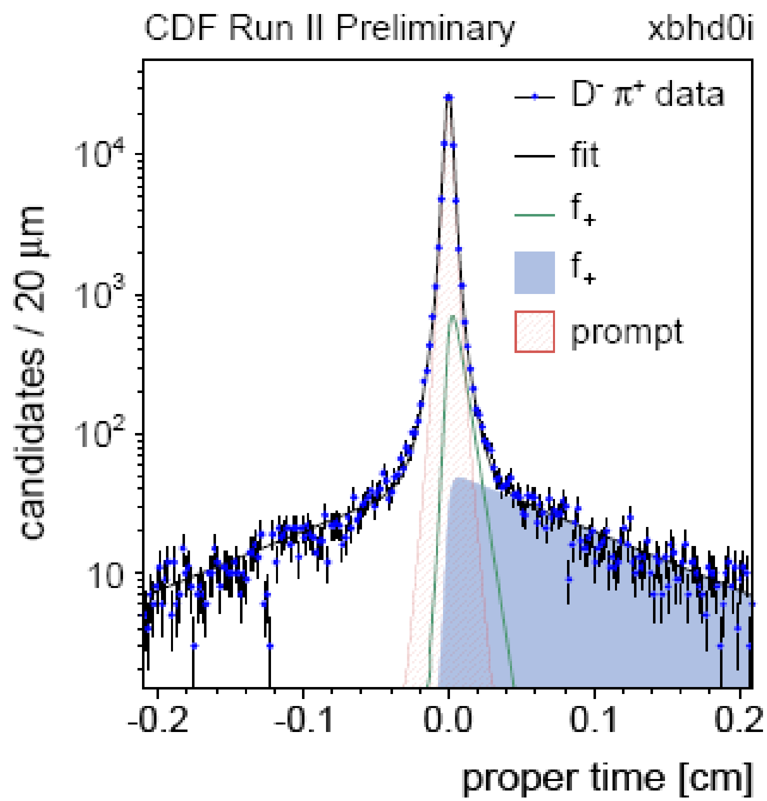
$$V_{ts} / V_{td} = 0.2060 \pm 0.0007\ (\text{exp})\ ^{+0.0081}_{-0.0060}\ (\text{theo})$$

"There is nothing more practical than a good theory."

Proper Time Resolution

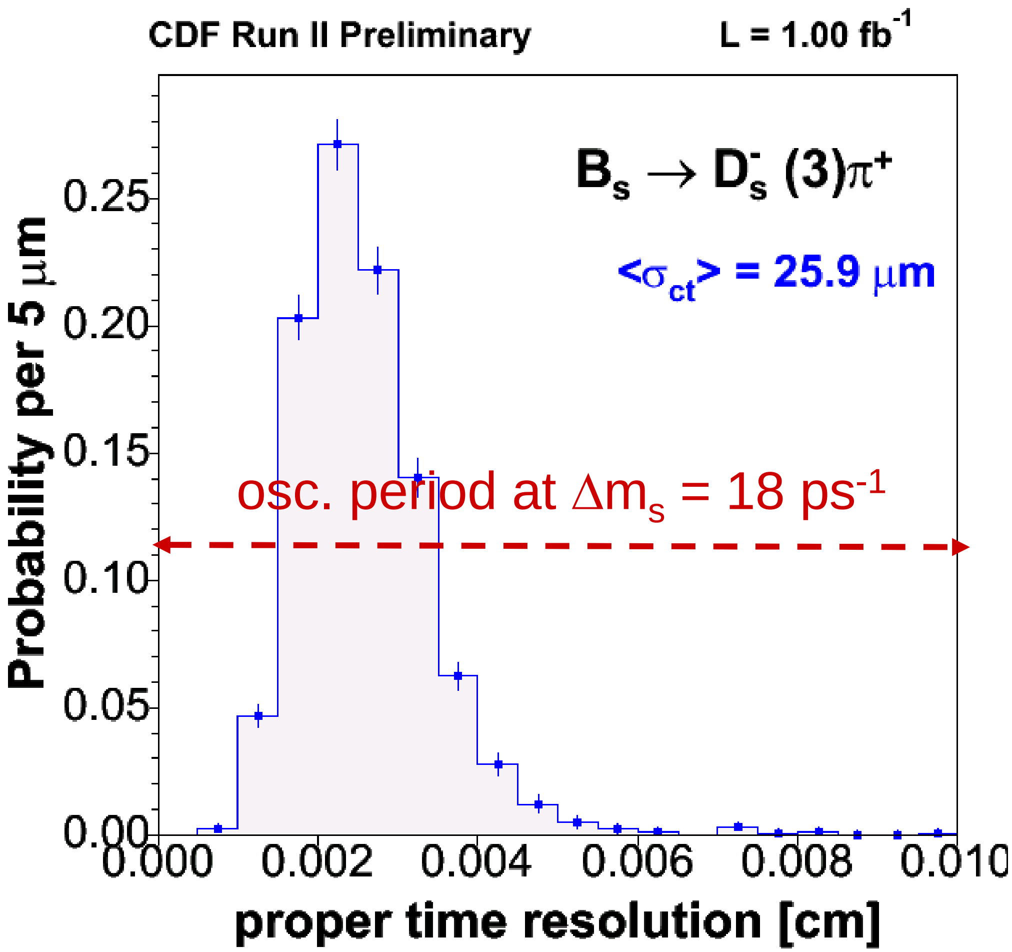
- Displaced track triggers also gather large prompt charm samples
- construct "B_s-like" topologies of prompt D_s^- + prompt track
- calibrate ct resolution by fitting for "lifetime" of "B_s-like" objects
  - expect zero lifetime by construction





Proper Time Resolution

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



- event by event determination of primary vertex position used
- average uncertainty

  ~ 26 μm

- this information is used per

semileptonic:

$\sigma_{ct}^0 \approx 59\,\mu m$

$\sigma_p / p \approx 15\%$

hadronic:

$\sigma_{ct}^0 \approx 30\,\mu m$

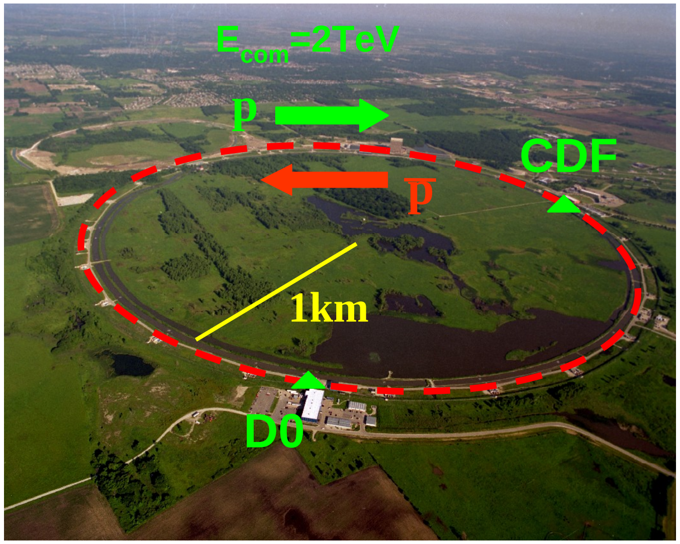
$\sigma_p / p \approx 0\%$

Performance of All Taggers

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

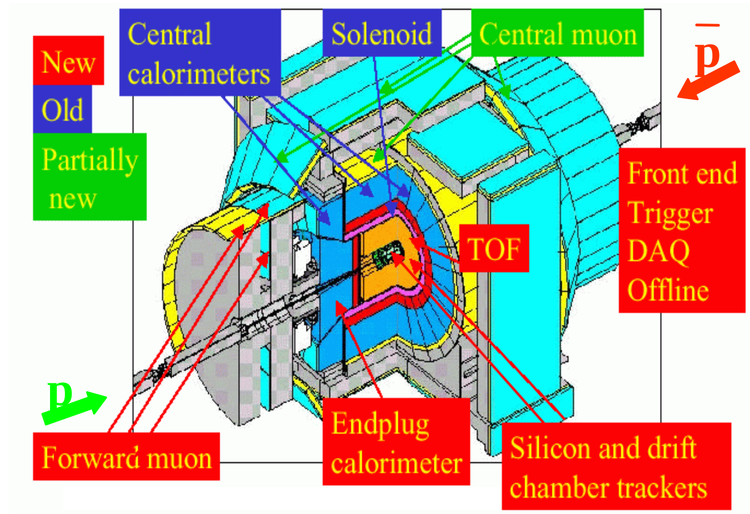
- **Errors are statistical only**
- **use exclusive combination of tags on opposite side**
- **same side and opposite side taggers are assumed to be independent**

The Tevatron and CDF



Fermilab, Chicago 

Currently the world's highest energy collider

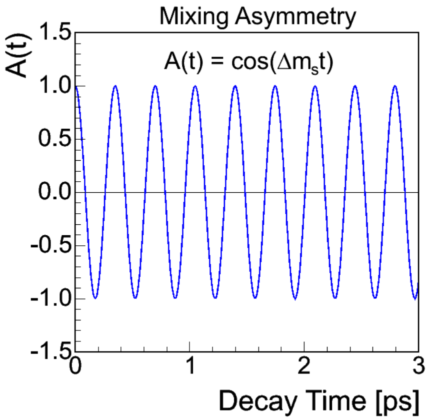


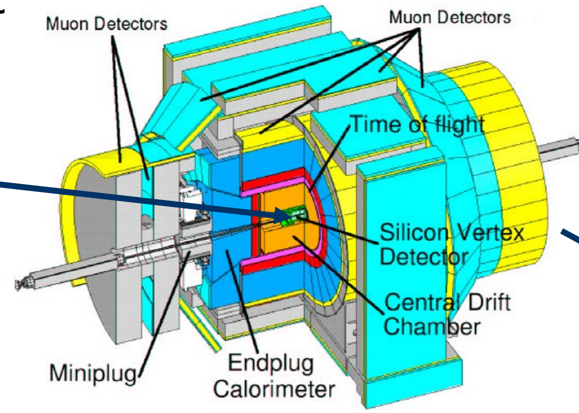
CDF Run I: 1992-1996 L= $0.1fb^{-1}$
Major Upgrades 1996-2001
CDF Run II: 2001-2006 L= $1fb^{-1}$

$p\bar{p}$ collisions can produce a wide spectrum of B hadrons in a challenging environment

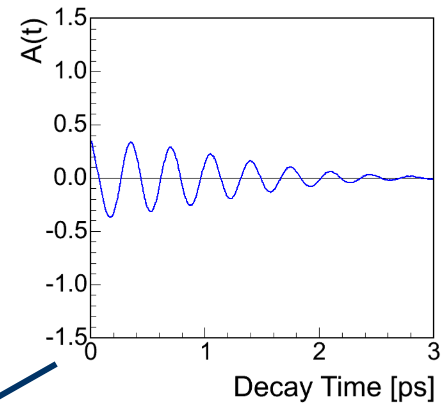
B_s cannot be produced at the B factories since Centre of Mass energy is below threshold

Real Measurement Layout





Data



momentum resolution
displacement resolution
flavor tagging power

scan for signal:

$A(\Delta m_s=[1\ldots30]\ ps^{-1})= ?$

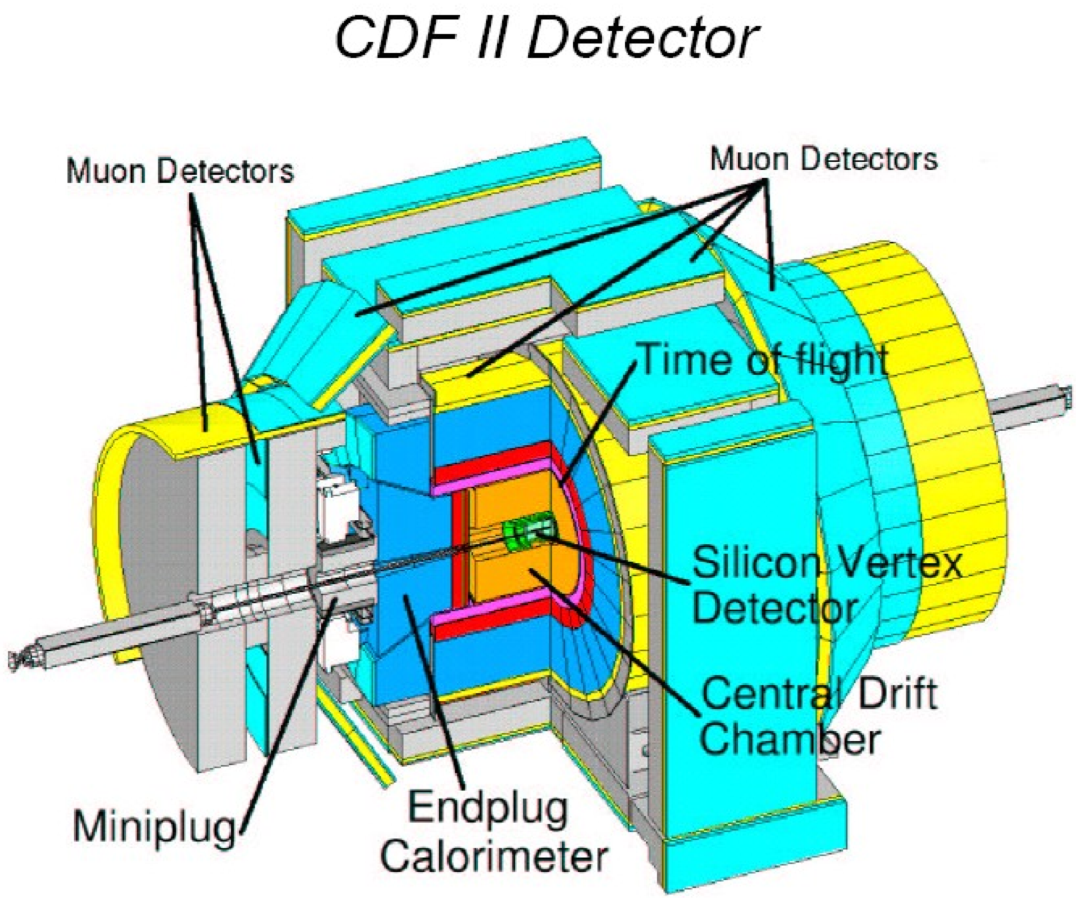
Unbinned Likelihood Fitter

measure frequency:

$\Delta\, m_s = ?$

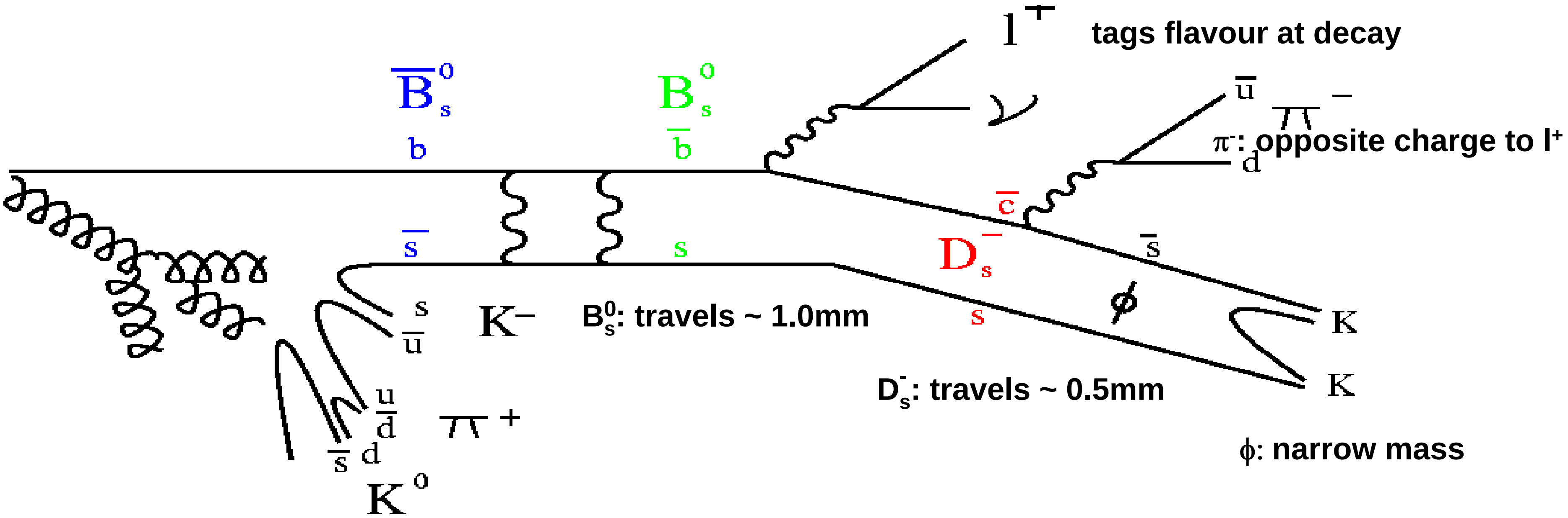
The CDFII Detector

- multi-purpose detector
- excellent momentum resolution $\sigma(p)/p<0.1\%$



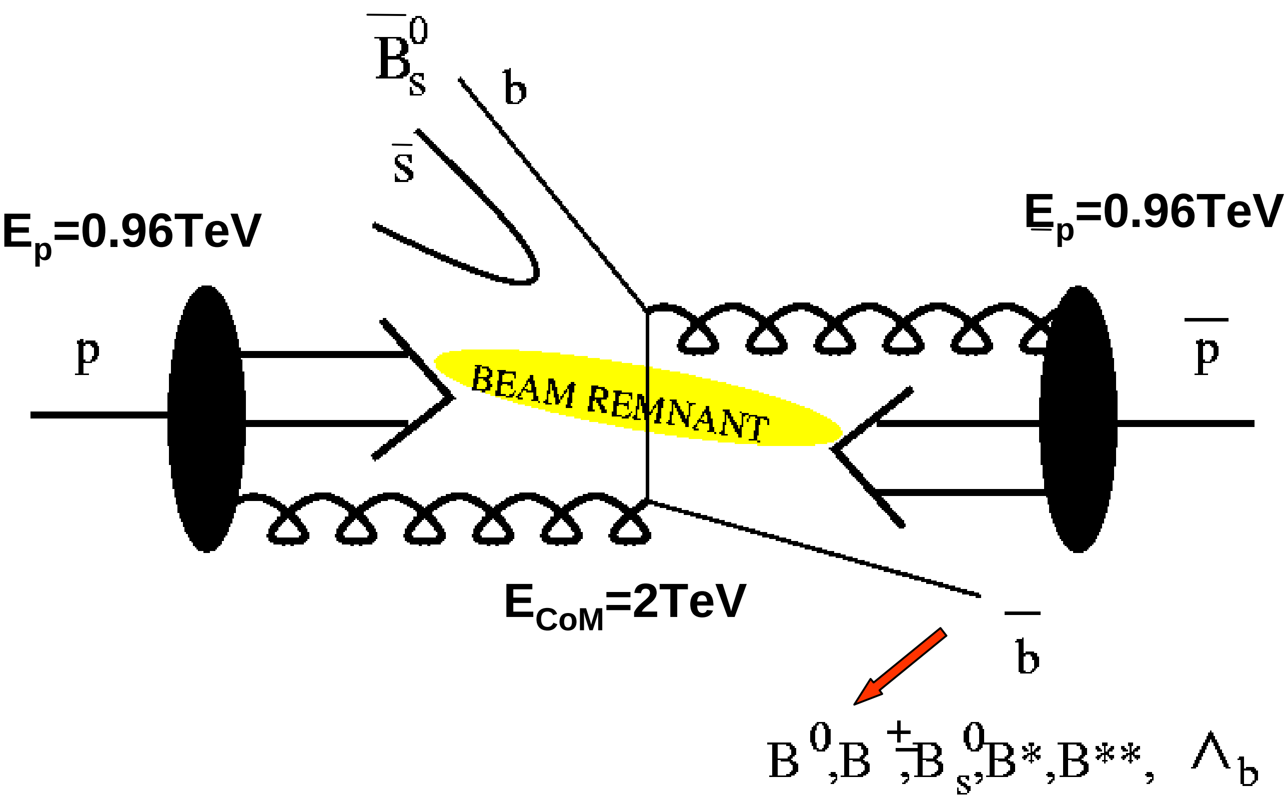
- Yield:
  - SVT based triggers
- Tagging power:

$B_s^0 - \overline{B}_s^0$ System



'neighbour' tags flavour at production

b Hadron Production at the Tevatron



Semileptonic Decay Fit Model

Unbinned maximum likelihood fit to cτ(B)

- **Background is parameterised by delta function and positive exp convoluted with Gaussian resolution:**

$$F_{bkg} = \left[(1 - f_+)\delta(t - \Delta_D) + \frac{f_+}{\tau_+}\exp\left(\frac{\Delta_E - t}{\tau_+}\right) \right] \otimes G(t, \sigma_G)$$

Free parameters: Δ_D Δ_E λ_+ f_+ σ_G

- **Signal: exp convoluted with Gaussian resolution, K factor distribution, P(K), and bias function, ε**

$$F_{sig} = N \frac{K}{c\tau}\exp\left(\frac{-Kt}{\tau}\right)\varepsilon(Kt) \otimes G(t, s\sigma_i) \otimes P(K)$$

- **Maximum likelihood function:**

$$L = \prod_i^{N_{sig}} \left[(1 - f_{bkg}) F_{sig}^i + f_{bkg} F_{bkg}^i \right] \cdot \prod_j^{N_{bkg}} F_{bkg}^j$$

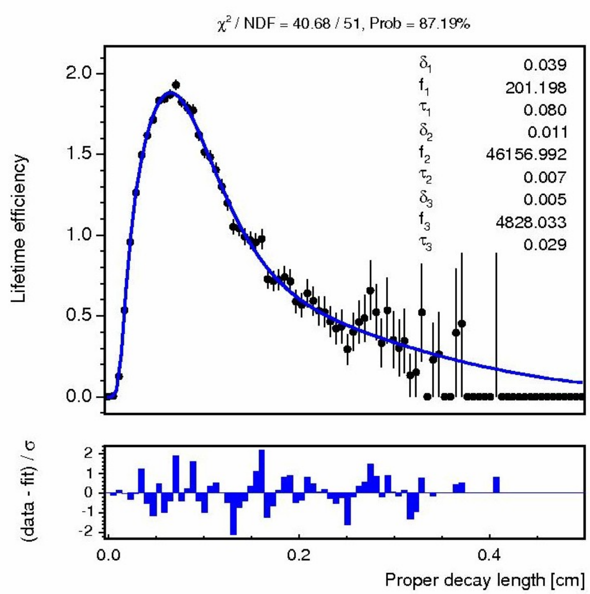
2) Time of Decay

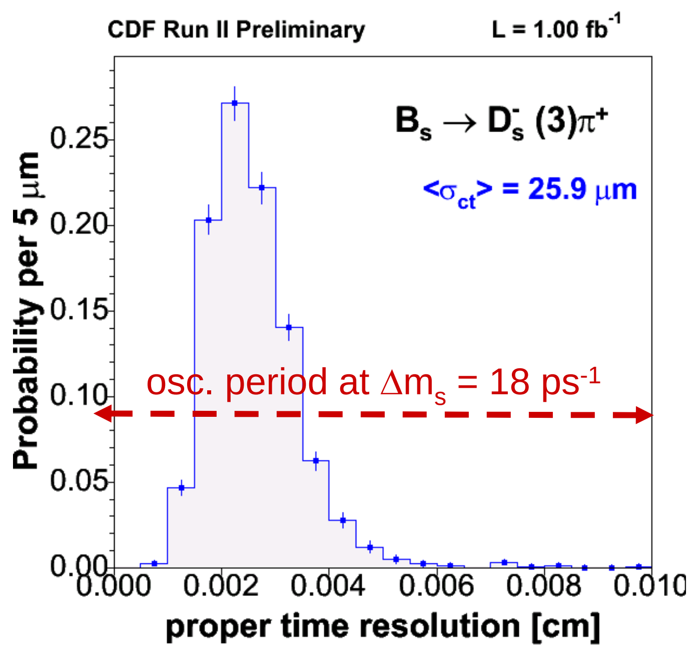
- Reconstruct decay length by vertexing
- Measure p_T of decay products

$$ct = \frac{L}{\beta\gamma} = L\frac{m(B)}{p(B)} = \frac{L_{xy}\, m(B)}{p_T(lD)}K$$

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

•Displaced Track Trigger imposes bias ⇒ correct with efficiency function





$\sigma_{ct}^{0} \approx 59\,\mu m$

$\sigma_p / p \approx 15\%$

$\sigma_{ct}^{0} \approx 30\,\mu m$

$\sigma_p / p \approx 0\%$

Crucial: Vertex resolution

(Silicon Vertex Detector, in particular Layer00 very close to beampipe)

$B_s^0 - \bar{B}_s^0$ System

Want to understand: - Average lifetime, Γ $= \frac{\Gamma_H + \Gamma_L}{2}$

- Lifetime difference, $\Delta\Gamma$ $= \Gamma_H - \Gamma_L$

- Rate of mixing, Δm $= m_H - m_L$

Current Status:	**Experiment**	**Theory**
$\Delta\Gamma/\Gamma$ (%)	**<0.29**	**≈0.15**
Δm (ps^{-1})	**>14.1**	**≈20**