CP Physics with b quarks at e^+e^- Colliders

Lectures at the 1999 SLAC
Summer Institute

Patricia Rankin, University of Colorado

Acknowledgements

- I have had a lot of help and advice in preparing these lectures
- In particular I would like to thank
  - BaBar Collaboration / All contributors to the BaBar Physics Book (SLAC-R-504)
  - Peter Dornan
  - Bill Lockman
  - Hwanbae Park
  - Peter Rowson
  - Jim Smith
  - Dave Burke, Lance Dixon and Helen Quinn

Introduction

- Why study the B system?
  - Represents a Precision Frontier
  - Increasing data sets, increasing the information content of data sets with better particle id, better decay time resolution
    - Beginning to probe Rare Decays
      - The evidence for new physics will be signals which are larger than predicted
    - Opening a new window on CP Violation - one area in which the presence of new physics is strongly expected
      - Look for inconsistencies in the overall picture which develops

BEAUTY IS TRUTH, TRUTH BEAUTY
THAT IS ALL YE KNOW ON EARTH
AND ALL YE NEED TO KNOW

Keats,
Ode to a Grecian Urn, 1819

Outline of Lectures

- Choice of Machine Energy
  - Υ(4s), Υ(5s)
  - LEP/SLC
- Techniques/Measurements
  - Focus on data
    - » Available Results
      - Mixing Measurements
      - B decays to charmless final states
    - » Anticipated Results
      - BaBar, BELLE
        - J/ψ K, golden mode for measuring β
        - Inclusive measurements of $|V_{ub}|$
  - Tools
    - Tagging
    - Background Suppression
    - Extraction of CP asymmetry

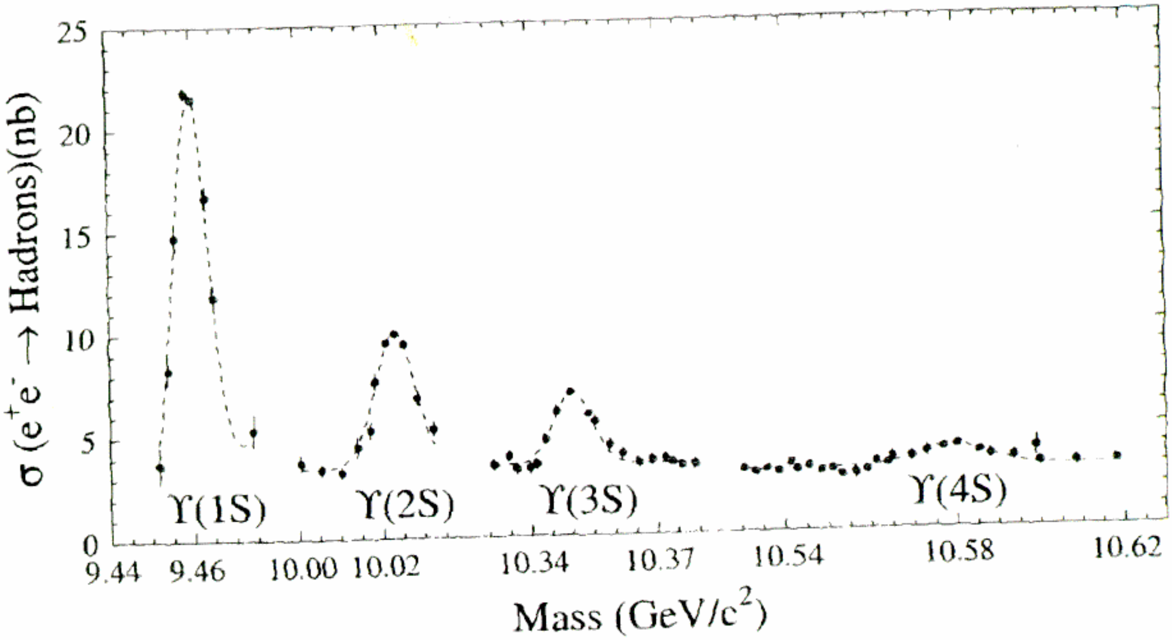
Goals/Requirements

- Goals
  - Clean measurements of CKM matrix parameters
  - Clean measurements of angles of unitarity triangle
- Requirements
  - Rule of thumb for uncertainty in sin2β, determined as the co-efficient of the sinΔmt term in the time dependent rate, is $\sigma=\sqrt{(3/N_{perfect})}$, where $N_{perfect}$ is the number of reconstructed, tagged events
    - » $N_{perfect}=100$ gives $\sigma=0.17$ (for a branching fraction of 2×10^{-5}, 50% tagging/reconstruction efficiency ⇒ 10^{7} events…)

Interesting Energies for e^+e^- colliders studying CP violation

- Running at the Υ(4s)
  - The Υ(4s) has a mass of 10.58 MeV/c^2 and can only decay into pairs of B_d mesons, $B^0\overline{B}^0$ or B^+B^-
    - No additional fragmentation hadrons
    - B's are produced almost at rest in Υ(4s) rest frame with $\beta=0.06$, travelling about $\beta\gamma c\tau=26\mu m$ before decaying in that frame.
  - $B\overline{B}$ pair produced in p state and will remain in this coherent state (even though neutral mesons can mix) until one decays
    - Must be able to study time dependence of decays to study CP violation through mixing
  - $\overline{B}B$ cross section is large - about 1.15nb, and continuum background is relatively low (about 3 times the signal)

Upsilon resonances



$e^+e^- \rightarrow$ (at 4s)	Cross-Section (nb)
$b\bar{b}$	1.05
$c\bar{c}$	1.30
$s\bar{s}$	0.35
$u\bar{u}$	1.39
$d\bar{d}$	0.35
$\tau^+\tau^-$	0.94
$\mu^+\mu^-$	1.16
e^+e^-	~40

Symmetric *v.* Asymmetric

- 26μm traveled by B’s in Υ(4s) rest frame does not allow for studies of time development
  - Can shift to running at threshold for $B\bar{B}^*$ production, after γ emission end up with $B\bar{B}$ pair in opposite CP state
    - » cross section estimates are about 1/7 of Υ(4s) cross section
- Solution is to use beams of different energies so that the Υ(4s) rest frame is boosted w.r.t. the lab frame increasing the spatial separation of the decays and making it measurable
- This makes a key CP measurement accessible at asymmetric Υ(4s) machines but not at symmetric Υ(4s) machines

- Asymmetric machines may have an advantage for other measurements
    - » separated vertex cuts can be used to reduce continuum backgrounds....
- BUT, may make some measurements harder
    - » Tracks concentrated in forward cone
- Acceptance could be higher at symmetric machines.......
    - » Reconstruction may be easier - e.g. for π^0's decaying to two photons
- Will learn how these effects trade off in coming months
- Integrated luminosity will be key

- Running at the Y(5s)
  - Y(5s) mass is 10.868 MeV/c^2 which is above threshold for production of $B_s\overline{B}_s$
- BUT
  - Small cross section at peak (< 0.3nb)
  - Six possible final states including $B\overline{B}$,$B\overline{B}^*$,$B^*\overline{B}^*$,$B_s\overline{B}_s^*$,$B_s^*\overline{B}_s^*$
  - Also background from production of $B(\overline{B}^*)$'s with a pion
  - Phase space prefers lighter modes
- So
  - predicted $B_s\overline{B}_s$ cross section less than 0.1nb
- Unlikely to soon (if ever) be practical to study CP violation at Y(5s)

◆ LEP/SLD

- The b-pair cross section at the Z^0 peak is about 6.7nb
- b's have a large boost (mean momentum about 35 GeV/c ⇒ decay lengths of 2-3mm)
- Produce all species of b hadrons, including B^0,B^+,B_s and b-baryons.
- The high boost greatly helps in studies of time evolution of mixing
  - $x_d = (\Delta m_d / \Gamma) \approx 0.7$ measured, x_s ?
  - B_s production rate about a quarter of B_d , oscillation rate expected to be about 20x
- Machine luminosities /integrated luminosities currently available don't give a high sensitivity for CP violation studies

Advantages/ Disadvantages

- Advantages
  - Clean environment
  - Good signal to noise ratios
    - Straightforward triggers (100 Hz rates)
    - Decays easy to identify
    - High efficiencies for reconstruction
    - At Y(4s) lots of kinematic constraints
    - Possible to do physics with neutral modes
      -
      -
- Disadvantages
  - Overall rates for rare modes…
    - nb cross sections not μb
      -
      -

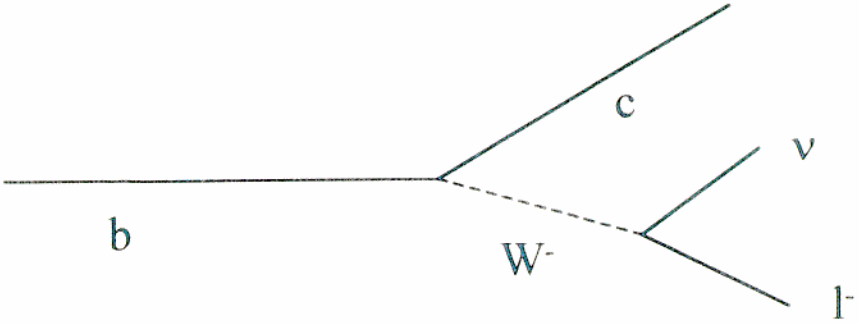
Potential Discussion topic ?

topic	Symmetric e^+e^- at the $\Upsilon(4S)$	Asymmetric e^+e^- at the $\Upsilon(4S)$	hadron collider
$\sin 2\beta$	−	+	+
α	?	+	+
Direct CP violation	+	+	+
γ	+	+	+
x_s	−	−	+
Absolute branching fractions B_d	+	+	−
Absolute branching fractions B_s	?	?	?
General properties of B_s decays	?	?	+
B_c physics	−	−	+
b-baryon physics	−	−	+
Rare exclusive $B_{u,d}$ decays with γ's	+	+	?
Rare exclusive $B_{u,d}$ decays with π^0's	+	+	?
Rare exclusive $B_{u,d}$ decays with l^+l^-	+	+	+
Rare inclusive $B_{u,d}$ decays with γ's	+	+	?
Rare inclusive $B_{u,d}$ decays with π^0's	+	+	−
Rare inclusive $B_{u,d}$ decays with l^+l^-	+	+	?
Very rare exclusive $B_{u,d}$ decays	−	−	?
Rare exclusive B_s decays with l^+l^-	−	−	+
Semileptonic decays ($B_{u,d} \to c$)	+	+	+
Semileptonic decays ($B_{u,d} \to u$)	+	+	?
Semileptonic decays ($B_s \to c$)	−	−	+
Semileptonic decays ($B_s \to u$)	−	−	?
Leptonic decays of $B_{u,d}$	+	+	−
Leptonic decays of D and D_s	+	+	−

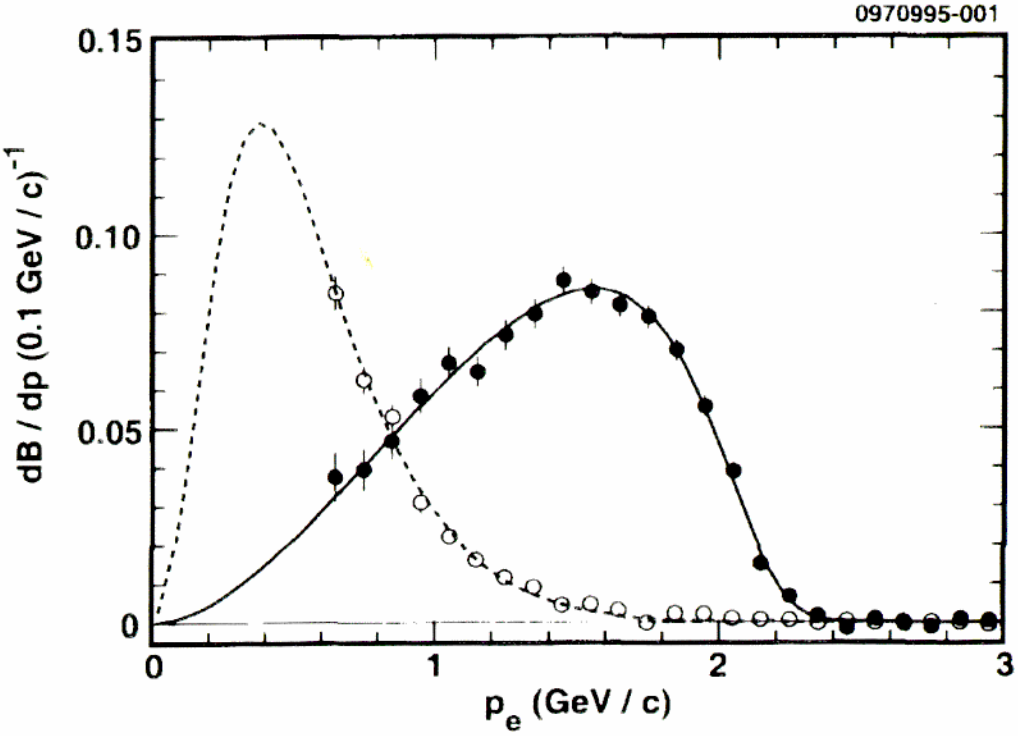
Report of the NSF Elementary Particle Physics
Special Emphasis Panel on B Physics , July 1998

Technique - Tagging (lepton)

- Can determine the flavor of the decaying b quark by
  - Looking at charge of lepton from semileptonic primary decay
    - » Br(B→Xlν) = (9.6 ±0.5±0.4)%



  - Complication - can also get leptons from semileptonic charm decays - these will have the opposite sign
  - leptons from primary decay have higher momentum (could reverse sign if momentum very soft ⇒secondary)



Spectrum of electrons from primary b decay (full circles) and secondary decay (open). (CLEO data, dilepton analysis, uses a fast momentum lepton to tag presence and flavor of a B. The second lepton is assumed to be from the primary decay of the other B if it has opposite sign to tag, if sign same as tag second lepton is assumed to come from secondary decay.

Data shown is corrected for wrong sign correlations (come from where?))

Technique - Tagging Charm/Kaon

- Can also use the flavor of the charm meson produced to tag the decay
  - Full reconstruction
    - » $B_d \rightarrow D^{*-} l^+$, $D^{*-} \rightarrow D^0 \pi^-$
      - Sign of soft pion flags flavor of D*
    - » $B_s \rightarrow D_s^- l^+$, $D_s \rightarrow \phi \pi^-$
  - Partial reconstruction
    - » On Y(4s) can use distinctive slow pion to infer D* momentum without fully reconstructing D daughter
- Can use flavor of Kaon produced in the charm decay
  - What about events with multiple strange particles?
  - What about Doubly Cabbibbo suppressed decays?

Technique - Thrust/Thrust Axis

- Thrust axis of an event $\widehat{T}$, is the direction which maximizes the sum of the longitudinal momenta of the particles (can use to separate event into two hemispheres)
- Thrust (T)

$$T = \frac{\sum_i \left|\hat{T} \bullet p_i\right|}{\sum_i \left|p_i\right|}$$

  - Thrust (2 ideal jets) = 1
  - $\text{T(isotropic event)} = \frac{\int d\Omega_p \left|p \cdot \hat{T}\right|}{\int d\Omega_p \left|p\right|} = \frac{1}{2}$

- Distribution of decay axis for candidate B and thrust axis for remainder of event can be used as a background discriminator between continuum and $B\overline{B}$ at Y(4s)
  - Roughly collinear for background (for two body B decays usually get a track from each of the two jets)
  - Uncorrelated for signal

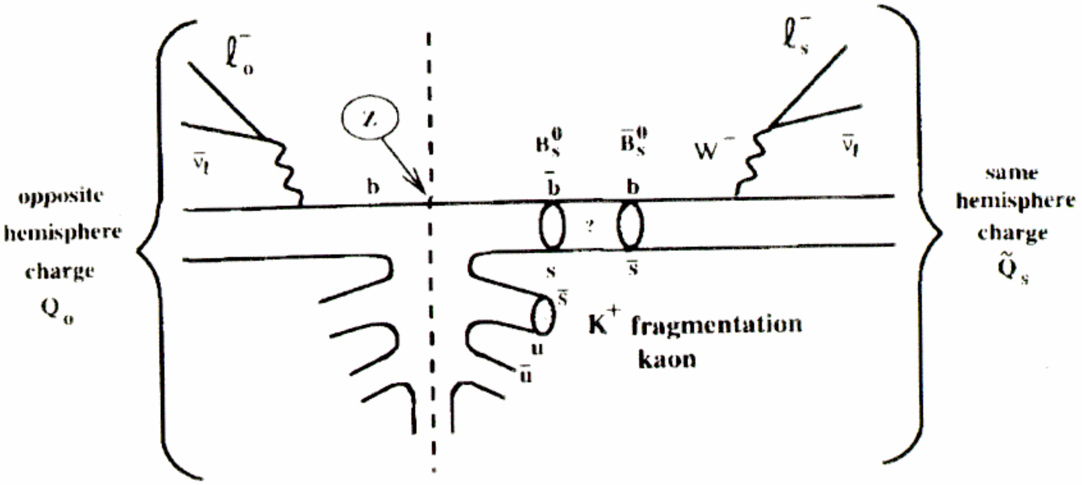
Technique - Tagging Jet Charge (Z^0)

- Uses tracks in hemisphere opposite that of the reconstructed B vertex
- Calculate an opposite hemisphere momentum weighted track charge defined as

$$Q_{opp} = \sum_i q_i \left| \vec{p}_i \bullet \hat{T} \right|^{\kappa}$$

  - q_i is charge of track i, p_i is momentum and T is thrust axis direction
  - κ is chosen to be 0.5 to maximize the separation between b and $\overline{b}$ quarks
- Since sum of charges of tracks from decay of B is zero, this variable gives information on charge of initial flavor state

- In hemisphere of B candidate look at nature and charge of most energetic fragmentation track produced in association
    - e.g $K^+ \Rightarrow B_s$
- In hemisphere opposite that of detected B look for indications of b or $\bar{b}$ production such as an energetic lepton

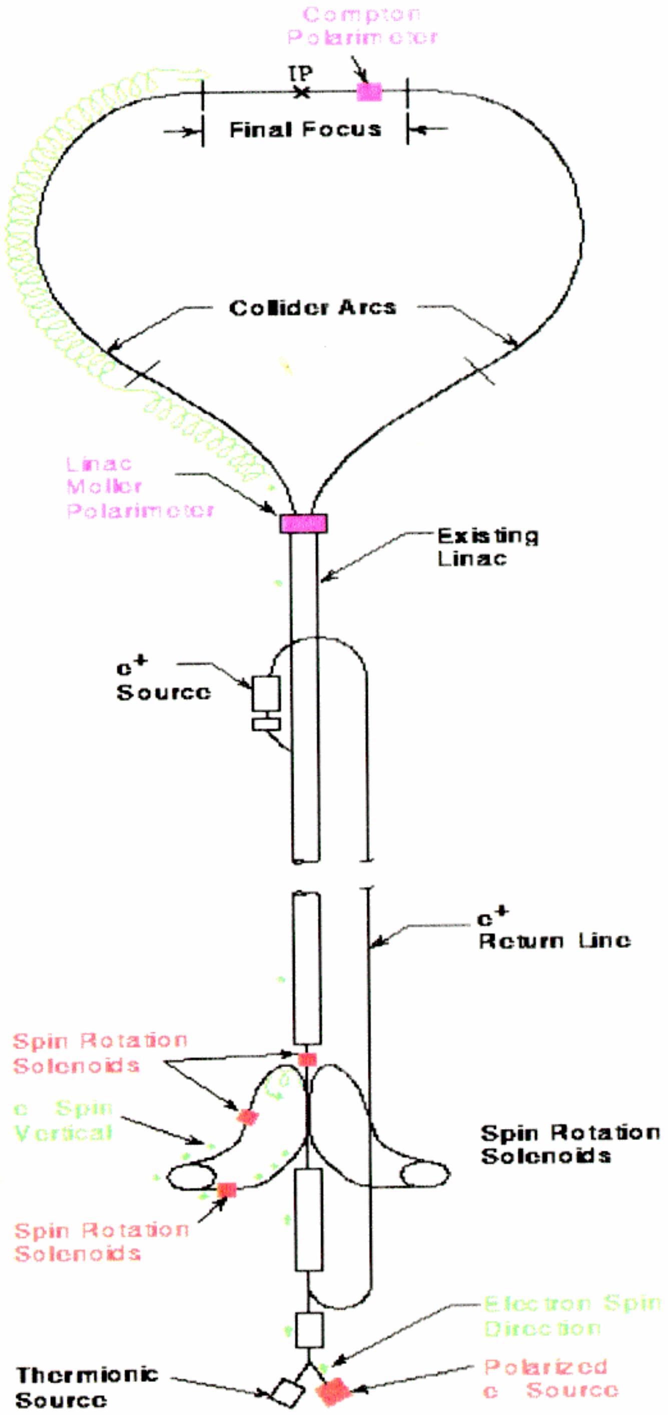


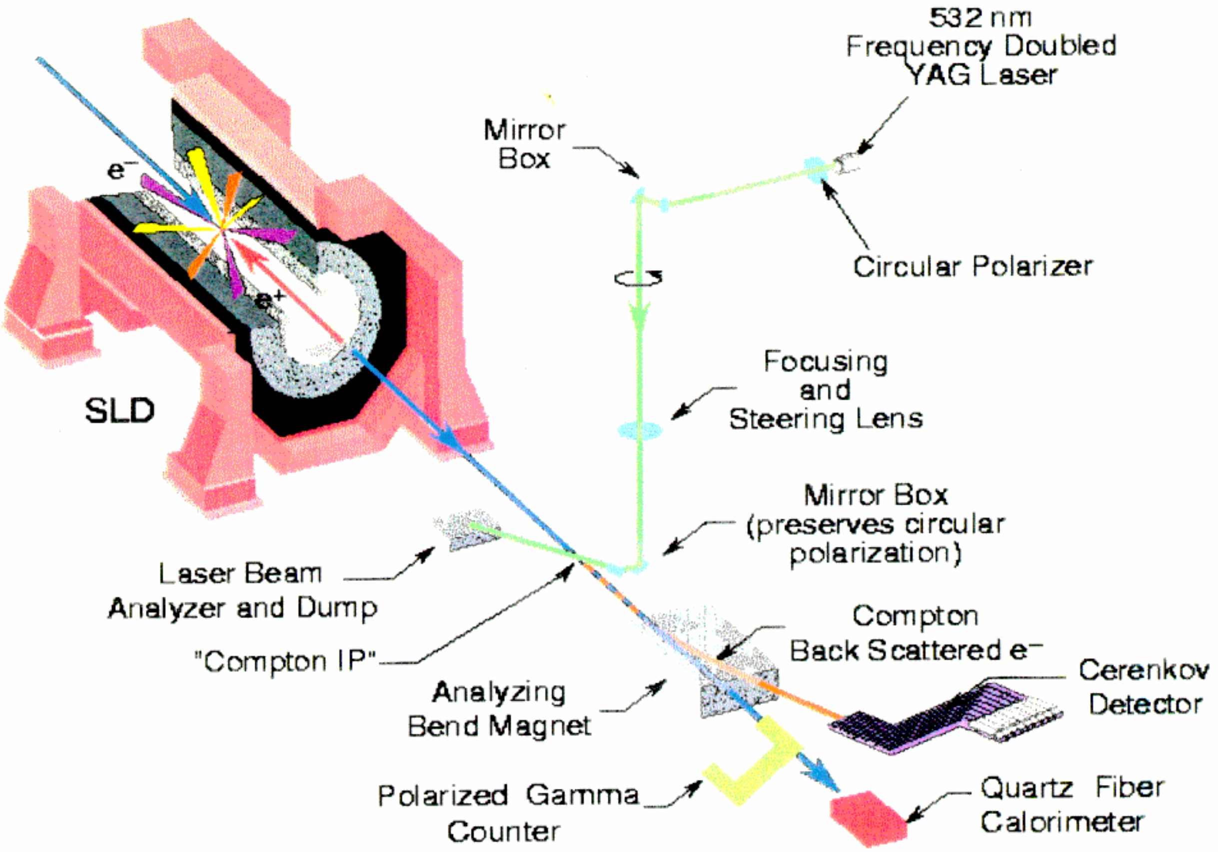
Technique - Tagging $A_{fb}(Z^0)$

- At SLC/SLD can use asymmetry in production which is enhanced by fact electron beam is polarized
  - The polarized forward-backward asymmetry at the Z^0

$$\widetilde{A}_{FB} = 2A_b \frac{A_e - P_e}{1 - A_e P_e} \frac{\cos\theta_T}{1+\cos^2\theta_T}$$

- » $A_b = 0.94$, $A_e=0.150$ (Standard model values)
- » P_e is electron beam longitudinal polarization
- » θ_t is the angle between the thrust axis and the electron beam direction (thrust is signed so that it points in same hemisphere as reconstructed vertex)
  - Gives an average correct tag probability of 0.74 for an electron polarization of 73%


Compton Polarimeter
IP
Final Focus
Collider Arcs
Linac Moller Polarimeter
Existing Linac
e+ Source
e+ Return Line
Spin Rotation Solenoids
e- Spin Vertical
Spin Rotation Solenoids
Spin Rotation Solenoids
Electron Spin Direction
Thermionic Source
Polarized e- Source

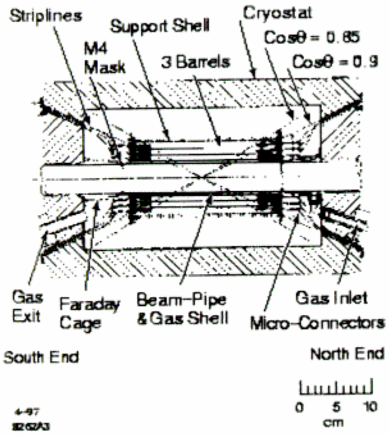


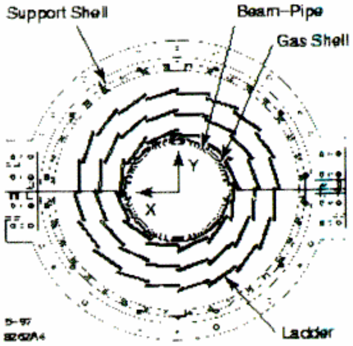
- High polarization strained lattice GaAs photocathode source.
- Extensively crosschecked, reliable, precision Compton polarimeter.

B Oscillations (Z^0)

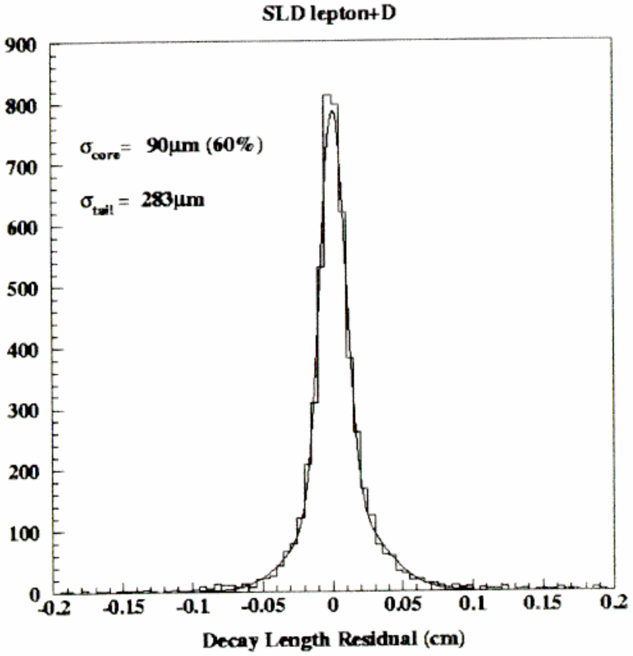
- Time Dependent
  - Aleph,Delphi,L3,Opal
    - » Each experiment has about 4 million hadronic Z events or 1 million pairs of B mesons
    - » Silicon vertex detectors give a resolution of around 300μm on decay lengths
    - » Particle ID
      - dE/dx Aleph,Opal
      - Ring Imaging Cerenkov (RICH) for Delphi

- SLD - Final event sample about 550K hadrons, ***BUT helped by***
  - Additional “asymmetry” tag
  - Linear collider provides small,stable beam spot (x,y,z) of about (1.5μm by 0.8μm by 0.7mm)
  - High precision 3D CCD based vertexing with very high signal to noise
    - » Decay length resolutions of around 50-100μm
  - PID - RICH





SLD/Decay length residuals



B mixing measurements

- These use the techniques we have just been looking at.
- Need to
  - Identify presence of B^0 meson
    - » Inclusive lepton analyses
      - Look for a high energy lepton with a large impact parameter
        - Tens of thousands of events, low purity (~10% for B_s)
    - » Look for an exclusive charm meson decay in addition, e.g.
      - Use lepton and meson to reconstruct B vertex, vertex and IP to get decay length
        - 100's of events, purities up to 50%
    - » Fully reconstruct the decay
      - DELPHI - tens of B_s

- Tag flavor of produced B meson
  - Divide event into two opposite hemispheres defined w.r.t the event thrust axis and look at each hemisphere
- Mistags due to
  - Experimental errors - particle mis-id, mis-assignment of tracks to vertices
  - Physics - opposite hemisphere $\overline{B}$ may also mix, could get lepton from secondary decay
- Can apply cuts to various distributions distinguishing signal and background, and to enhance purity of tag
      - » Not always optimal approach
        - Do not use all the available information such as shape of the distributions, correlations between variables ⇒ loss of efficiency
- Combine information to get an overall tag probability

Technique- Multivariate Analysis

- Various techniques have developed to use information in an event as efficiently as possible, these include
- Genetic Algorithms
  - Optimize a set of cuts
- Parameterized approach (PA)
  - Likelihood method which deals with shapes of cuts
- Fisher
  - Takes correlations into account using linear combinations of variables
- Neural Networks
  - Takes correlations into account using non-linear combinations

Genetic Algorithms

- Each cut is considered as a gene, each set of cuts is considered to be an individual
- Genetic algorithm searches for best solution from a population of solutions (Darwinian survival of the fittest) by calculating the fitness value
  - Set up by user - e.g. signal/noise ratio for individual
- Removes least fit from population
- Spawns a new population of descendants from survivors using three genetic operators
  - Reproduction
  - Crossover
  - Mutation
- On average, fitness of population improves
- Has been shown that can help HEP analyses which need statistical optimization

Parameterized Approach

- Uses relative likelihood
- Take two classes of events (A,B) to be distinguished between and the distributions of variables to be used to distinguish between them for each class
- Ideally, for N variables, the likelihood to belong to class A is given by the N-dimensional density distribution $g^A(x_1,\ldots..,x_n)$
- Can use the ratio of likelihoods to characterize the event

$$X_{PA} = \frac{g^A(x_1,\ldots\ldots,x_n)}{g^A(x_1,\ldots\ldots,x_n) + g^B(x_1,\ldots\ldots,x_n)}$$

- In practice, obtaining the N-dimensional density distributions can be hard so in PA can make the approximation that the discriminating variables are uncorrelated, i.e.

$$g^A(x_1,\ldots\ldots,x_n) = \prod_{i=1}^{n} g_i^A[x_i]$$

- Note
  - By construction X_{pa} tends to 1 for class A, 0 for class B events.
  - Can include variables which have little discriminating power without affecting usefulness of other variables
  - If no correlation exists between variables then X_{pa} is optimal
  - If there is correlation then lose information
  - Can improve by incorporating two-dimensional distributions for highly correlated pairs of variables

Fisher / Mahalanobis

- Linear Discriminant Analysis (LDA) or the Fisher method works by combining the N variables chosen to describe the events linearly
- Discrimination task is to find an axis in R^N space of the discriminating variables so that the two classes are maximally separated
- Need to supply
  - Mean value of variable over full sample and over each class ($\overline{x}, \overline{x}_A, \overline{x}_B$)
  - Total variance covariance matrix (T) which can be separated into two components related to "within-class" (W) and "between-class" (B) components
    - $T_{\mu\nu}=W_{\mu\nu}+B_{\mu\nu}$
    - W reflects dispersion of events relative to center of gravity of their own class
    - B represents distance of a class to the total center of gravity.

- Distance between projected points is a maximum along the direction defined by line between x_A, x_B, line segment $\overline{x}_A, \overline{x}_B$ is projection axis
- To discriminate events
  - Compare value of discriminating function for an event(x)

$$X_{FI} = \frac{\sqrt{n_A n_B}}{n} (\overline{x}_A - \overline{x}_B)^T W^{-1} x$$

  - with some threshold value

$$\theta_0 = \frac{\sqrt{n_A n_B}}{n} (\overline{x}_A - \overline{x}_B)^T W^{-1} (\overline{x}_A + \overline{x}_B)/2$$

  - n_A, n_B, n are number of events in each sample and total number
  - More usually, just calculate probability for X_{FI} to be in each class
  - Mahalanobis uses full covariance matrix not W

P. Rankin, SSI 1999

algorithms, perhaps

parallel-processing
ısed in a variety of
› solve more easily
ırse, our proverbial
ҳ we should be able

we try to make a
: task, consider the
onsisting of numer-
ern, yet collectively
ical of the kinds of
cognize. For exam-
focus strictly on the
: allow our perspec-
ıge of a commonly
see the image, it is
ure.
re we to program a
. The first thing our
or areas of interest
r the splotches into
vith one object. We
; line segments. We
onsistency, trying to
context of the other
rules describing the
: attributed to noise
ld attempt to isolate
oles and completing

ı profile, facing left,
ıtes the complexity
dog is illustrated as
e write a computer
e of the dog, which
h spots are simply

can see the dog in
scrimination? This
switching time of
ıan seven orders of
gical systems. This

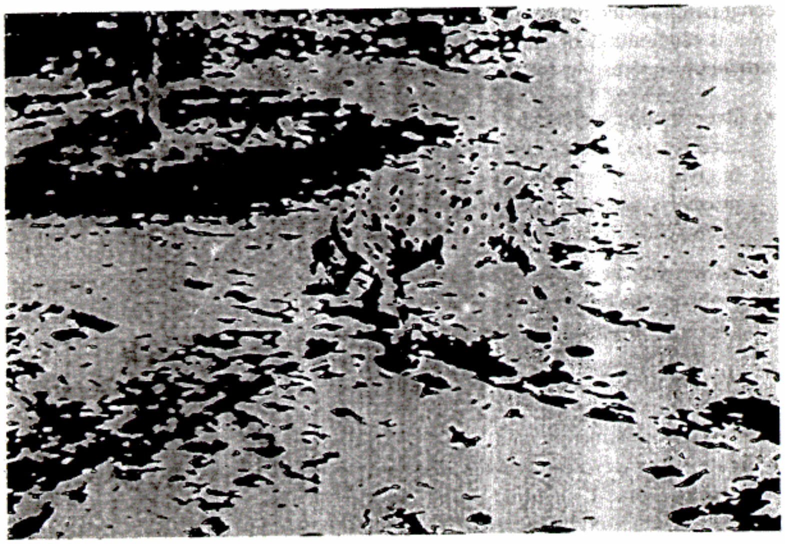


Figure 1.1 The picture is an example of a complex pattern. Notice how the image of the object in the foreground blends with the background clutter. Yet, there is enough information in this picture to enable us to perceive the image of a commonly recognizable object. *Source: Photo courtesy of Ron James.*

question is partially answered by the fact that the architecture of the human brain is significantly different from the architecture of a conventional computer. Whereas the response time of the individual neural cells is typically on the order of a few tens of milliseconds, the massive parallelism and interconnectivity observed in the biological systems evidently account for the ability of the brain to perform complex pattern recognition in a few hundred milliseconds.

In many real-world applications, we want our computers to perform complex pattern recognition problems, such as the one just described. Since our conventional computers are obviously not suited to this type of problem, we therefore borrow features from the physiology of the brain as the basis for our new processing models. Hence, the technology has come to be known as **artificial neural systems** (ANS) technology, or simply **neural networks**. Perhaps the models we discuss here will enable us eventually to produce machines that can interpret complex patterns such as the one in Figure 1.1.

In the next section, we will discuss aspects of neurophysiology that contribute to the ANS models we will examine. Before we do that, let's first consider how an ANS might be used to formulate a computer solution to a pattern-matching problem similar to, but much simpler than, the problem of

①

Neural Networks

- Can extend non linear approaches by using a non-linear one such as a neural network
- One architecture is the multilayer perceptron (MLP) using back propagation of error
  - Basic building block of network is neuron
    - » Associate N input variables x_k and a response , the output z_j
    - » Inputs are linearly combined according to some parameters called weights (w_{jk}) and a threshold term(θ_j) can be added to give an neuron activation signal (Z)

$$Z = \sum_{k=1}^{N} w_{jk} x_k + \theta_j$$

    - » Activation is simulated by evaluating a non-linear function a(x) at point Z, often a(x) is the sigmoid function

$$a(x) = \frac{1}{2}(1 + \tanh x)$$

- In the MLP architecture neurons are layered
  - Data directed from first layer of N inputs (discriminating variables) through N_h neurons in layers, to the last layer which gives response through hidden layers
  - connection weights between layers of neurons are determined by minimizing the error function

$$E = \frac{1}{2n} \sum_{p=1}^{n} (X_{NN} - t_i^P)^2$$

  - where X_{NN} is the output parameter defined as a function of input parameters (x_k) by

$$X_{NN} = a\left(\sum_{j=1}^{N_h} w_{1j} a\left(\sum_{k=1}^{N} w_{jk} x_k + \theta_j \right) + \theta_1 \right)$$

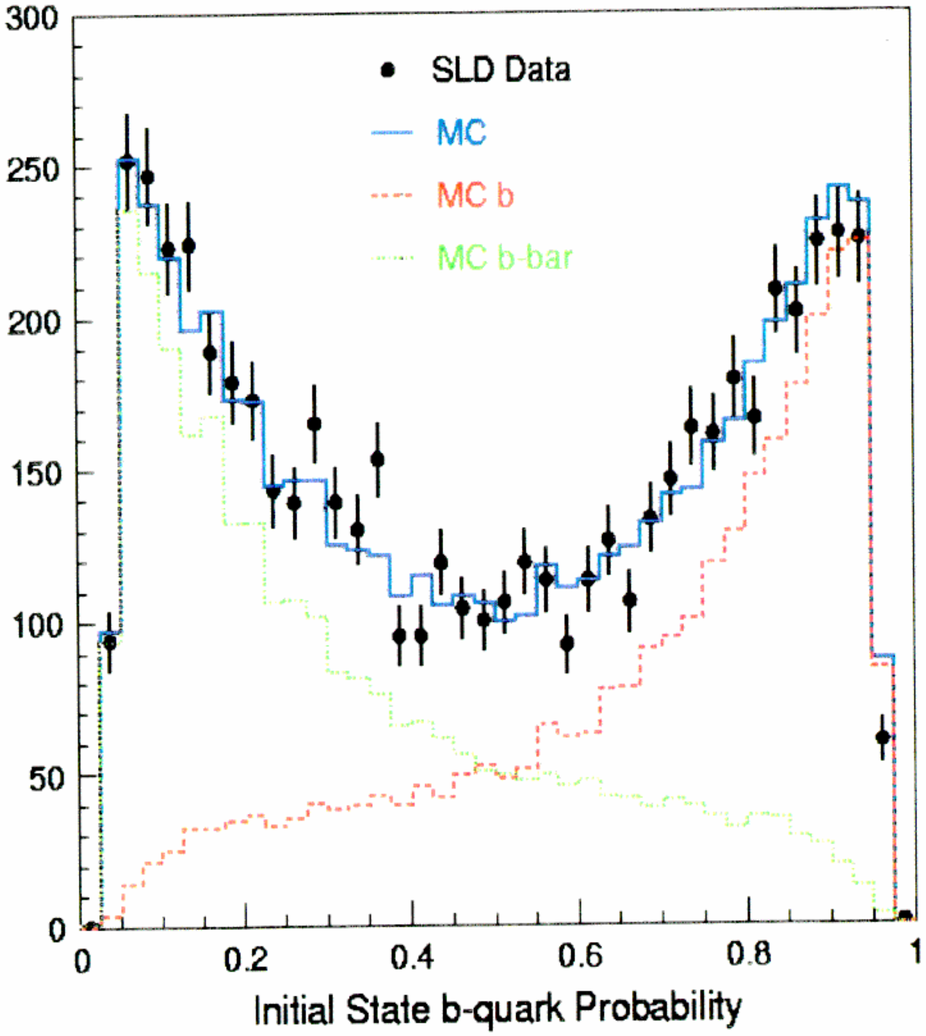
  - where n is number of events used for training and t^P is desired output value (1 for signal, 0 for background)

Something to think about

- Neural networks can adapt to unknown probability distributions even those with strong correlations
- Can use supervised or unsupervised training
  - Supervised training requires a set of signal events and a set of background events
    - How do you select these?
      - If you use Monte Carlo how can you be sure that samples are realistic?
      - If you use real data sets how do you select them?
  - How do you assign errors?
  - Unsupervised networks don’t need samples and could be useful if cannot get reliable training samples (e.g. of beam related backgrounds)

"The computer is claiming *its* intelligence is real, and *ours* is artificial."

SLD/tag probability for Δm_s analysis



- Tag flavor of decaying B meson
  - » Use charge of energetic, high p_t lepton
- Measure Decay time with error
  - » Calculate B boost using lepton, reconstructed charm meson, neutrino energy (estimated from missing energy, if possible)
    - LEP - decay time resolutions of order 0.3 ps (inclusive lepton), 0.15ps(charm-lepton), 0.06ps for DELPHI, exclusive B_s
    - SLD, decay time resolutions of order 0.06ps for $D_s l$ (charm- lepton) events

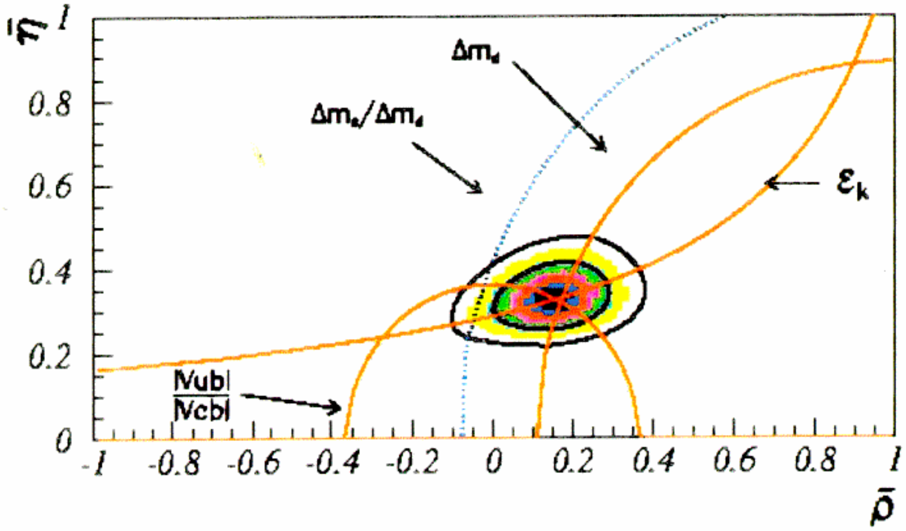
$$t = L / \gamma \beta c$$

$$\gamma \beta = P_B / M_B$$

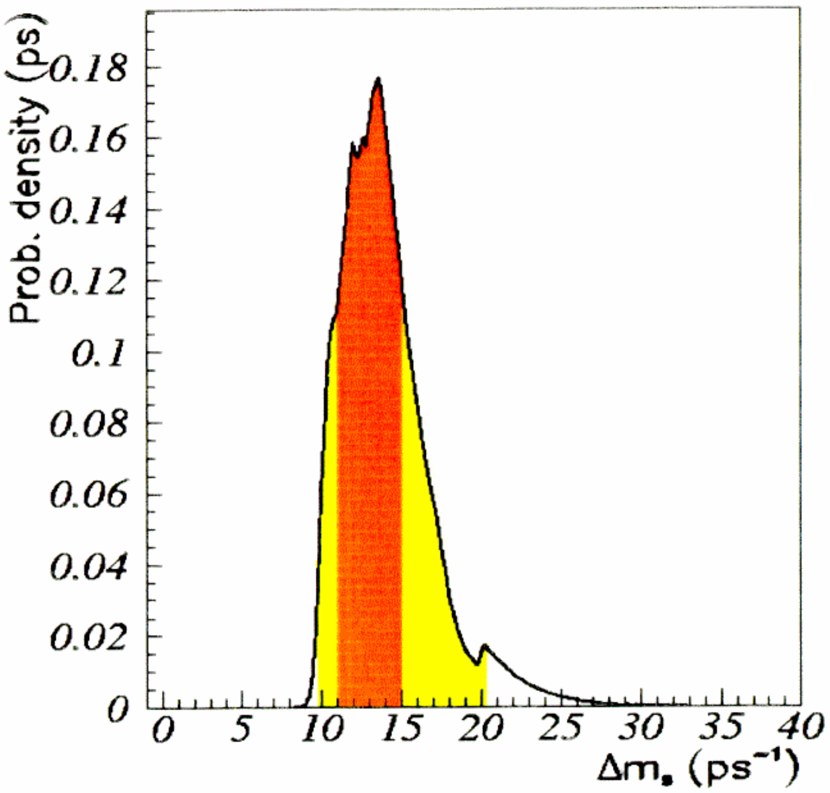
Preliminary

Standard Model fit to current data

P.Paganini et al., hep-ph/9711261 & hep-ph/9802289



Standard Model 'prediction' *including Δm_s analyses from LEP*



1. Apply the input vector, $\mathbf{x}_p = (x_{p1}, x_{p2}, \dots, x_{pN})^t$ to the input units.
2. Calculate the net-input values to the hidden layer units:

$$\text{net}^h_{pj} = \sum_{i=1}^{N} w^h_{ji} x_{pi} + \theta^h_j$$

3. Calculate the outputs from the hidden layer:

$$i_{pj} = f^h_j(\text{net}^h_{pj})$$

4. Move to the output layer. Calculate the net-input values to each unit:

$$\text{net}^o_{pk} = \sum_{j=1}^{L} w^o_{kj} i_{pj} + \theta^o_k$$

5. Calculate the outputs:

$$o_{pk} = f^o_k(\text{net}^o_{pk})$$

6. Calculate the error terms for the output units:

$$\delta^o_{pk} = (y_{pk} - o_{pk}) f^{o\prime}_k(\text{net}^o_{pk})$$

7. Calculate the error terms for the hidden units:

$$\delta^h_{pj} = f^{h\prime}_j(\text{net}^h_{pj}) \sum_k \delta^o_{pk} w^o_{kj}$$

   Notice that the error terms on the hidden units are calculated *before* the connection weights to the output-layer units have been updated.

8. Update weights on the output layer:

$$w^o_{kj}(t+1) = w^o_{kj}(t) + \eta \delta^o_{pk} i_{pj}$$

9. Update weights on the hidden layer:

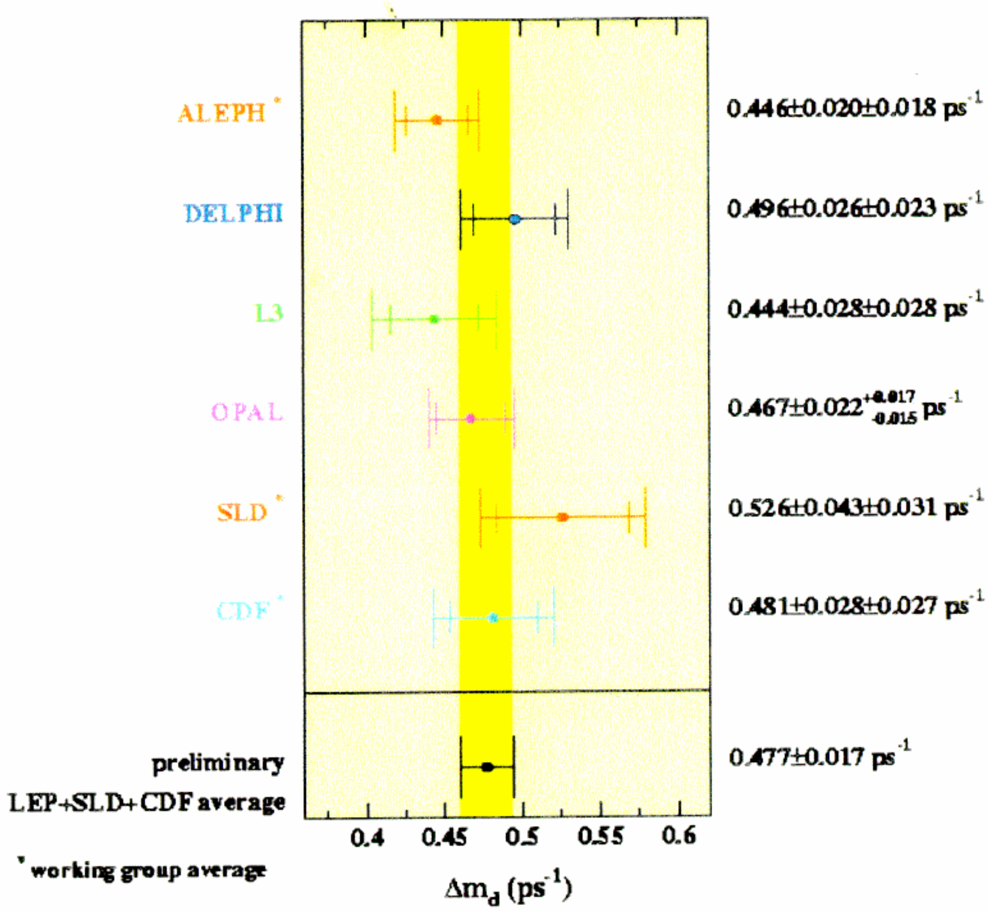
$$w^h_{ji}(t+1) = w^h_{ji}(t) + \eta \delta^h_{pj} x_i$$

   The order of the weight updates on an individual layer is not important.

   Be sure to calculate the error term

$$E_p = \frac{1}{2} \sum_{k=1}^{M} \delta^2_{pk}$$

since this quantity is the measure of how well the network is learning. Whe the error is acceptably small for each of the training-vector pairs, training c be discontinued.


ALEPH *
DELPHI
L3
OPAL
SLD *
CDF *
preliminary
LEP+SLD+CDF average
* working group average
0.446±0.020±0.018 ps^{-1}
0.496±0.026±0.023 ps^{-1}
0.444±0.028±0.028 ps^{-1}
0.467±0.022$^{+0.017}_{-0.015}$ ps^{-1}
0.526±0.043±0.031 ps^{-1}
0.481±0.028±0.027 ps^{-1}
0.477±0.017 ps^{-1}
0.4
0.45
0.5
0.55
0.6
Δm_d (ps^{-1})

Resolution/Reach

- For Δm_s have to deal with
  - small fraction of strange B mesons in b jets ($P_s \approx 10\%$)
  - large value of Δm_s
- For finite resolution on measured time σ_t, oscillation amplitude is damped by the quantity

$$\rho = \exp\left[-\frac{(\Delta m_s \sigma_t)^2}{2}\right]$$

  - So to access values of $\Delta m_s \approx 10\text{ps}^{-1}$, need to achieve $\sigma_t \approx 0.2$ ps
  - Can relate time resolution to resolution on B decay distance (σ_d) and on the B energy (σ_E)

$$\left(\frac{\sigma_t}{\tau}\right)^2 = \left(\frac{\sigma_d}{\langle d \rangle}\right)^2 + \left(\frac{t}{\tau}\right)^2 \left(\frac{\sigma_E}{E}\right)^2$$

    - » Second term, increases with t showing that need to work close to interaction point….
    - » Can access $\Delta m_s \approx 10\text{ps}^{-1}$ with $\sigma_d \sim 250\mu\text{m}$

Technique - Amplitude Fit

- Use a likelihood approach
  - include effect of detector smearing, mistags, selection efficiencies and dependence on oscillation frequency
- Amplitude fit
  - Time dependent mixing generates a *periodic signal*
    - suited to Fourier analysis
  - Find minimum of a negative Log likelihood distribution, constructed from individual probabilities evaluated for each event to observe measured decay time, separating mixed and unmixed candidates.
  - Time distribution introduces a quantity, A, usually termed the oscillation amplitude

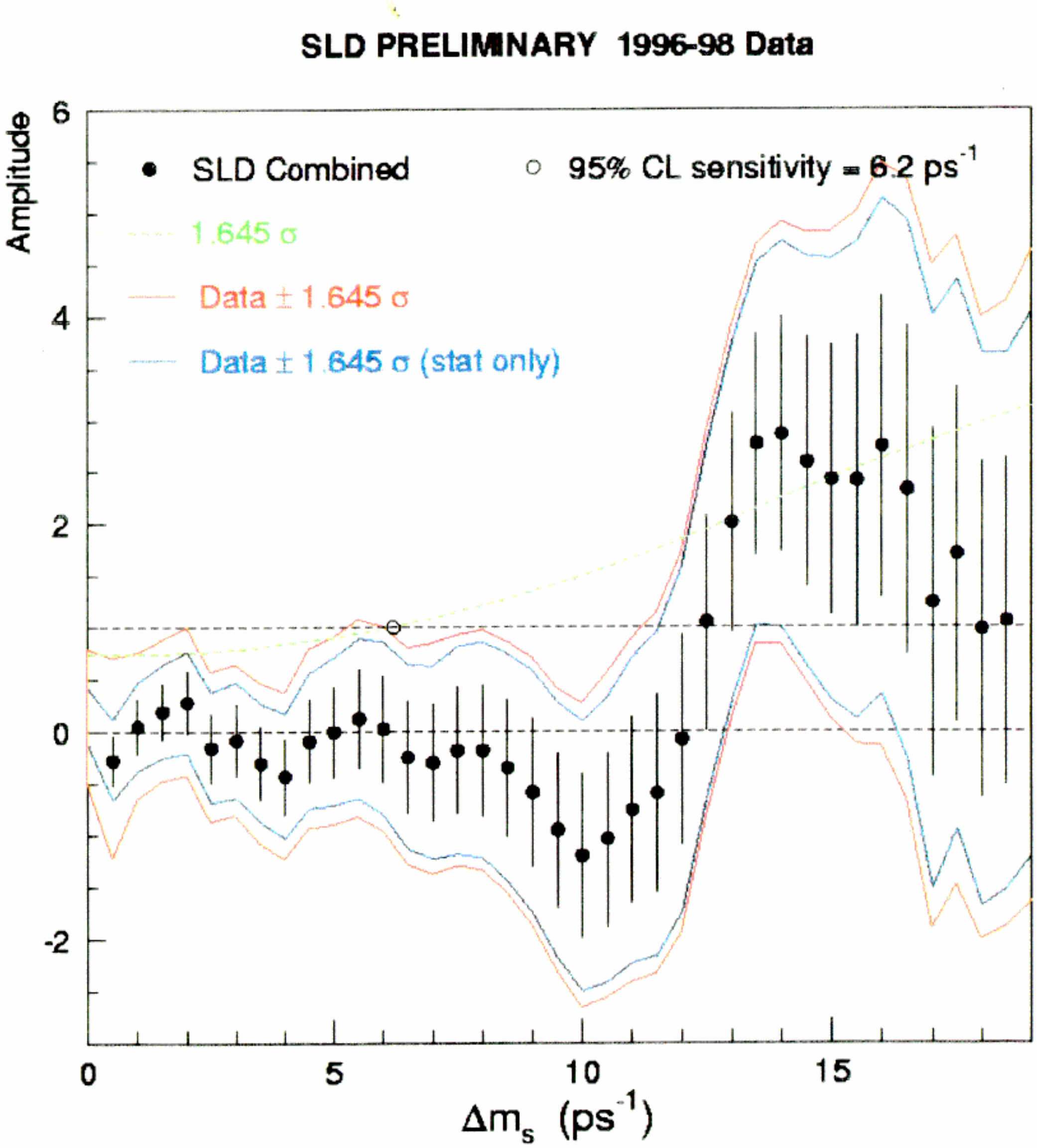
$$\mathrm{prob}(\mathrm{B}^0 \rightarrow \overline{\mathrm{B}}^0) = \frac{1}{2}\Gamma e^{-\Gamma t}(1 - A\cos(\Delta m \cdot t))$$

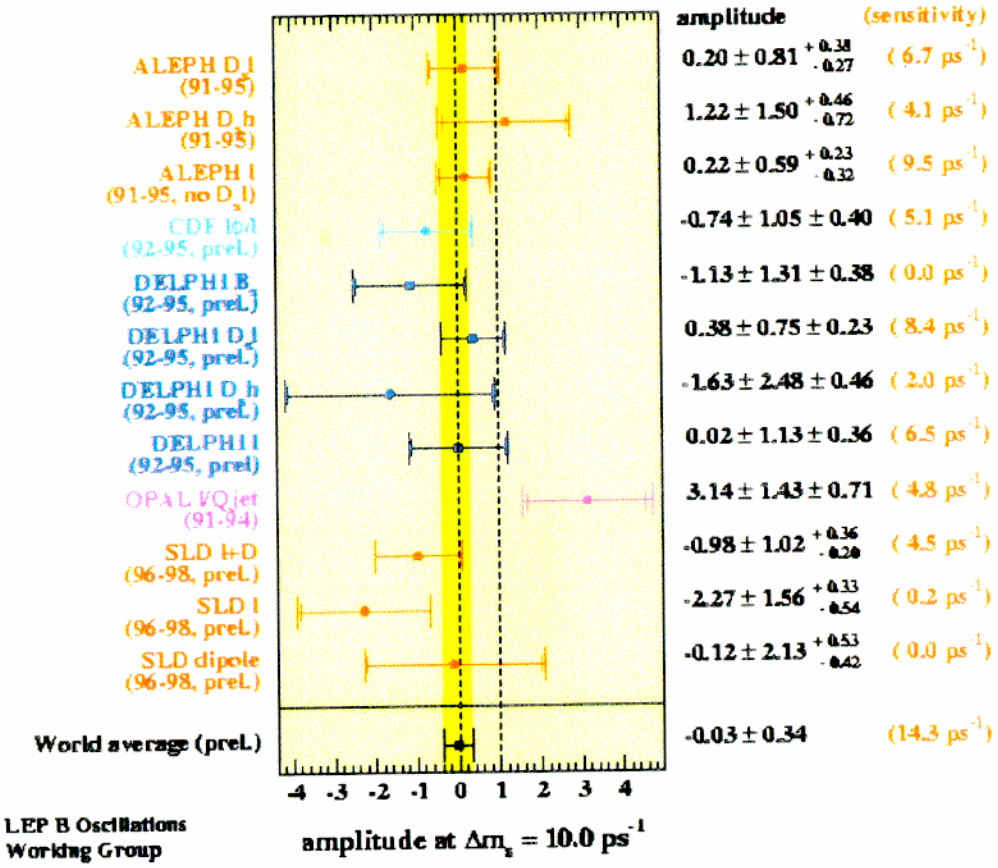
  - For each assumed value of Δm_s, fit for A
    - For a genuine signal expect A compatible with unity, within measurement errors.

Technique - combining data on Δm_s

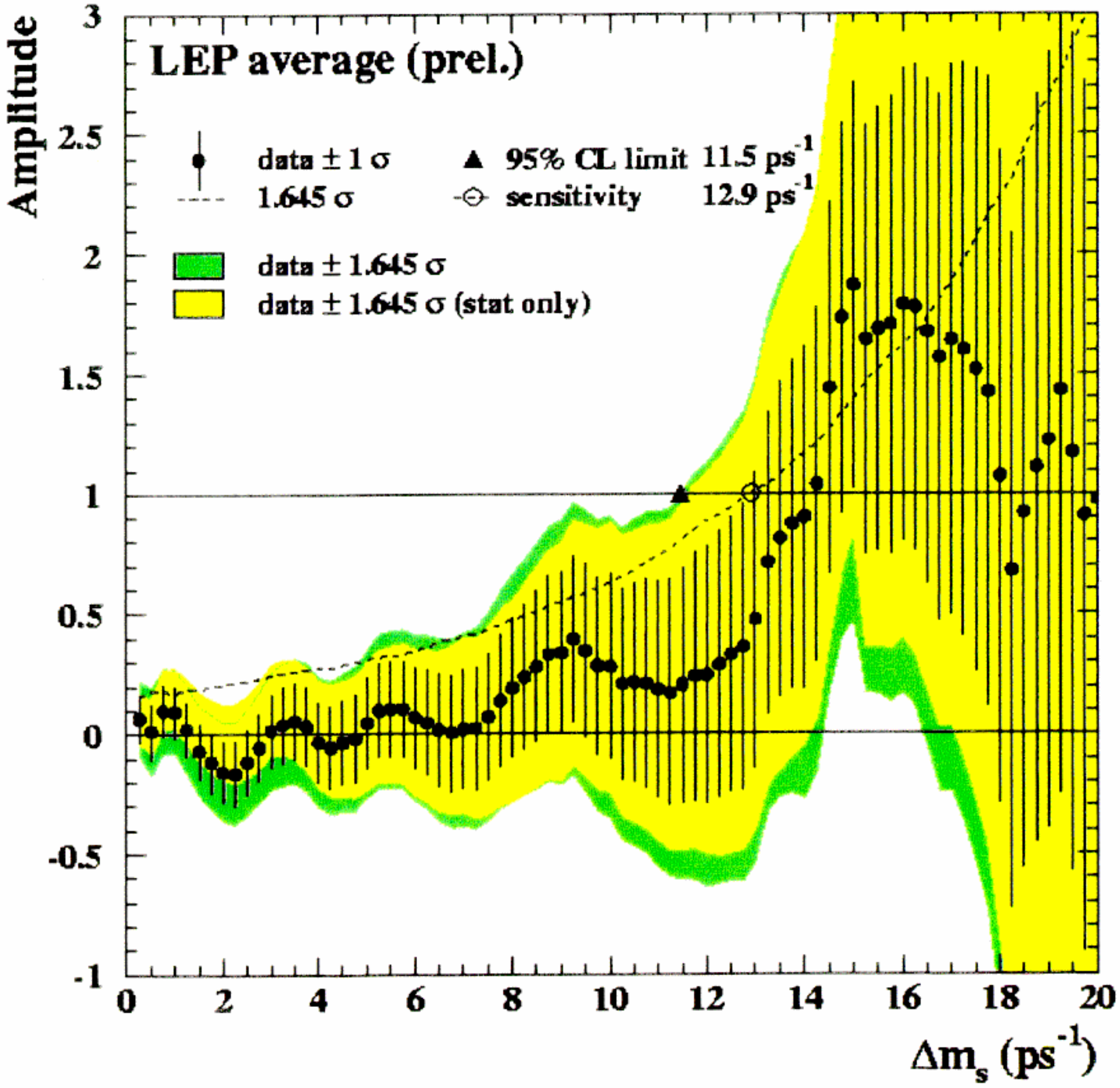
- If experiment has enough sensitivity to observe a signal expect that the variation of A *vs.* Δm_s will be a Breit-Wigner distribution with a maximum at A=1 at the exact value of the mass difference, and a full width equal to the inverse of the B_s meson lifetime.
- At present, only 95% c.l. limits have been derived
  - » Correspond to probability that a genuine signal (A=1) would give measured amplitude less than or equal to the one observed in less than 5% of cases.
- Can combine different measurements by averaging the various measurements at each Δm_s value
  - » Have taken correlations between systematic and statistical errors into account

Expect A=0 for no oscillations A=1 if frequency is true Δm_s



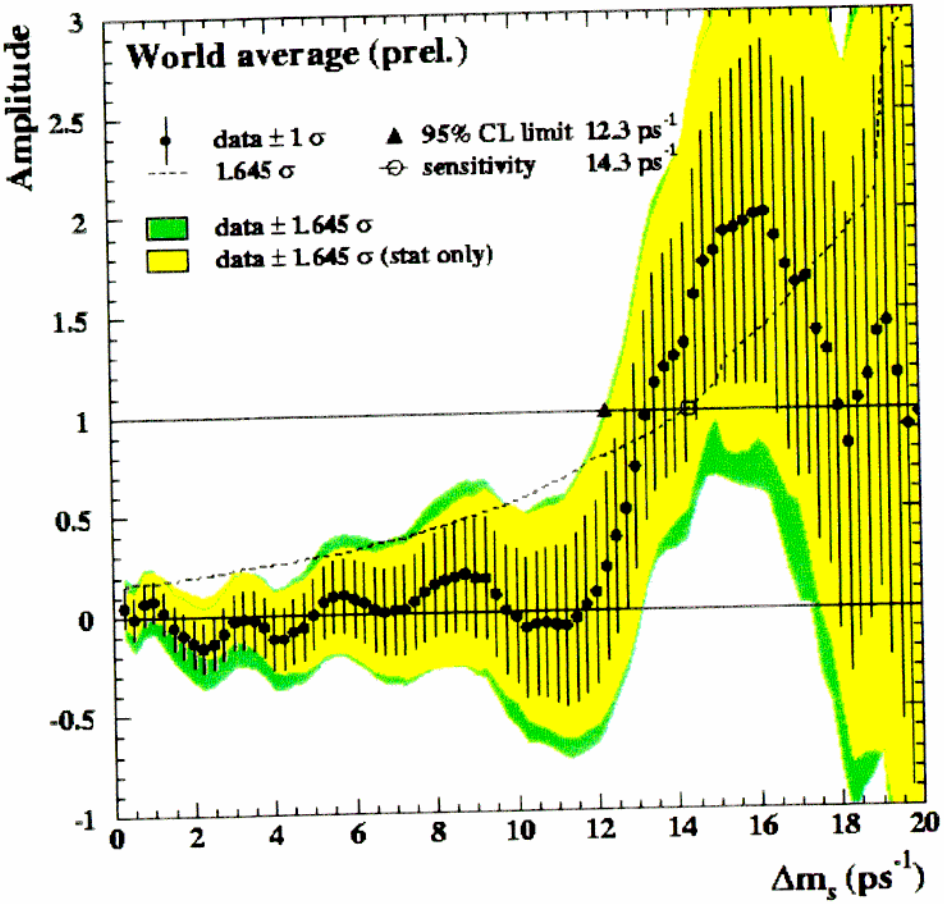

amplitude
(sensitivity)
ALEPH D_s l (91-95)
$0.20 \pm 0.81 ^{+0.38}_{-0.27}$ (6.7 ps^{-1})
ALEPH D_s h (91-95)
$1.22 \pm 1.50 ^{+0.46}_{-0.72}$ (4.1 ps^{-1})
ALEPH l (91-95, no D_s l)
$0.22 \pm 0.59 ^{+0.23}_{-0.32}$ (9.5 ps^{-1})
CDF lφ/l (92-95, prel.)
$-0.74 \pm 1.05 \pm 0.40$ (5.1 ps^{-1})
DELPHI B_s (92-95, prel.)
$-1.13 \pm 1.31 \pm 0.38$ (0.0 ps^{-1})
DELPHI D_s l (92-95, prel.)
$0.38 \pm 0.75 \pm 0.23$ (8.4 ps^{-1})
DELPHI D_s h (92-95, prel.)
$-1.63 \pm 2.48 \pm 0.46$ (2.0 ps^{-1})
DELPHI l (92-95, prel)
$0.02 \pm 1.13 \pm 0.36$ (6.5 ps^{-1})
OPAL l/Qjet (91-94)
$3.14 \pm 1.43 \pm 0.71$ (4.8 ps^{-1})
SLD l+D (96-98, prel.)
$-0.98 \pm 1.02 ^{+0.36}_{-0.20}$ (4.5 ps^{-1})
SLD l (96-98, prel.)
$-2.27 \pm 1.56 ^{+0.33}_{-0.54}$ (0.2 ps^{-1})
SLD dipole (96-98, prel.)
$-0.12 \pm 2.13 ^{+0.53}_{-0.42}$ (0.0 ps^{-1})
World average (prel.)
-0.03 ± 0.34 (14.3 ps^{-1})
-4 -3 -2 -1 0 1 2 3 4
LEP B Oscillations Working Group
amplitude at $\Delta m_s = 10.0\ ps^{-1}$

LEP preliminary



Combined
Preliminary result

- World Average (May 99)



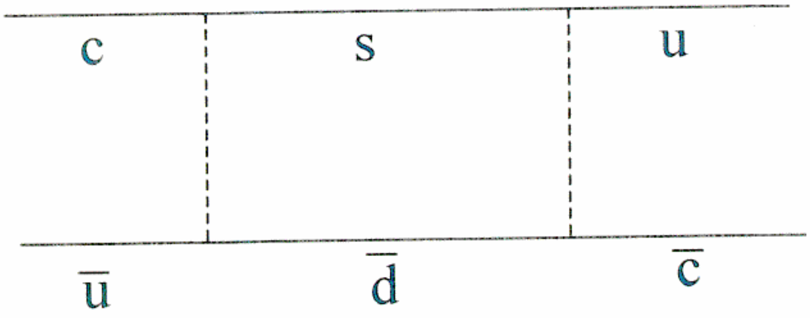
Results

- LEP/SLD have a preliminary lower limit of $\Delta m_s > 12.3\ \mathrm{ps}^{-1}$ at 95% CL
- Structure in vicinity of 13-17 ps-1 is *intriguing* - significance still being determined (How ? Are points independent ?)
- $\Delta m_d = 0.471 \pm 0.016 \mathrm{ps}^{-1}$ (preliminary, all data)
- SM predictions of these quantities have large uncertainty due to QCD correction factors ⇒ take ratio ...

$$\frac{\Delta m_d}{\Delta m_s} \cong \left(\frac{|V_{td}|}{|V_{ts}|}\right)^2 \cong \lambda^2 \cong 0.05$$

D^0 - $\overline{D}^0$ Mixing

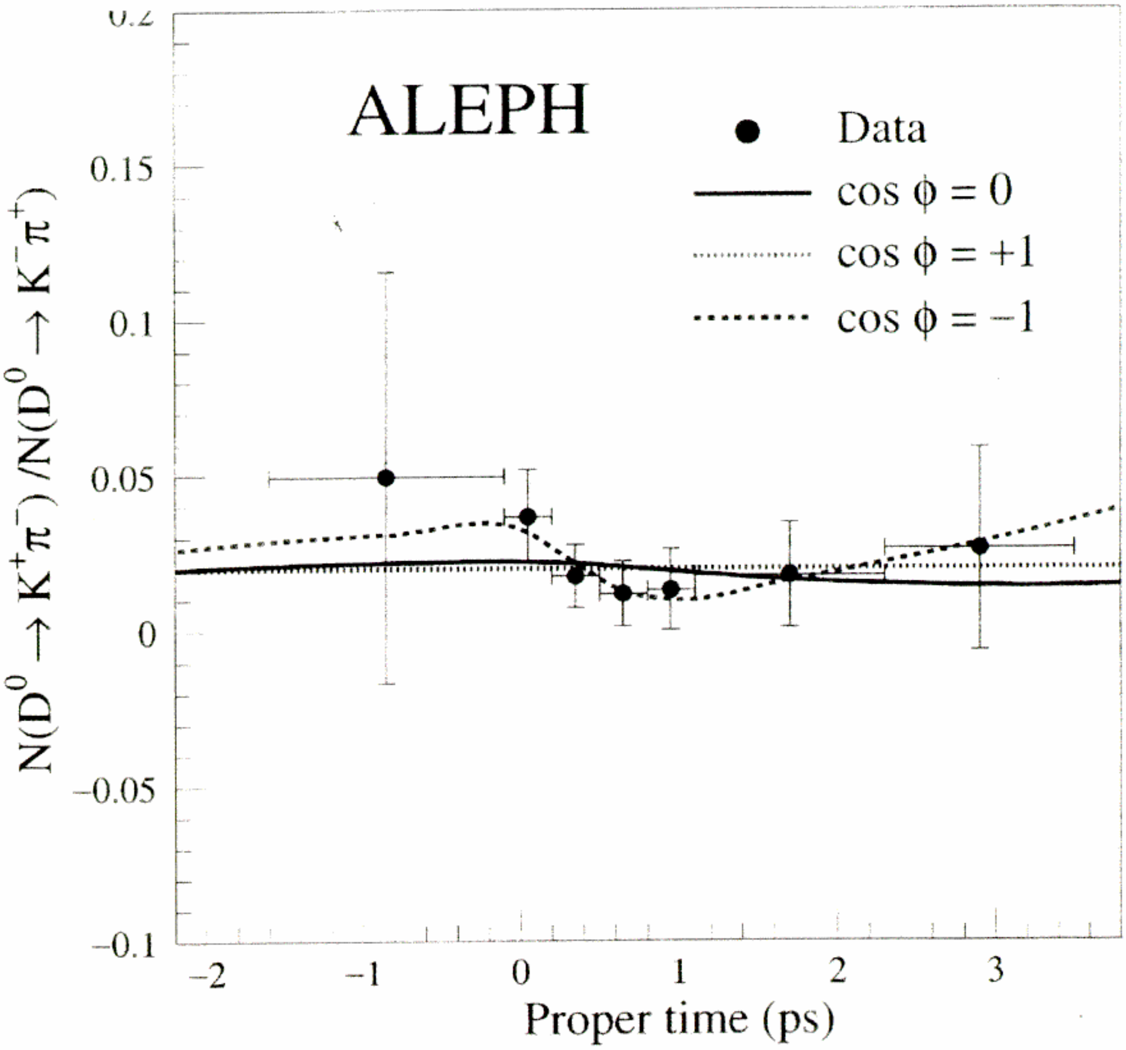
- Expected to be tiny ~10^{-7} (though rate disputed, some estimates ~10^{-4})
  - Could indicate new physics if large
    - » Some models which suppress CP violation in B system enhance D mixing
- Two ways which can get $\Delta C=-\Delta S$ transitions
  - Mixing

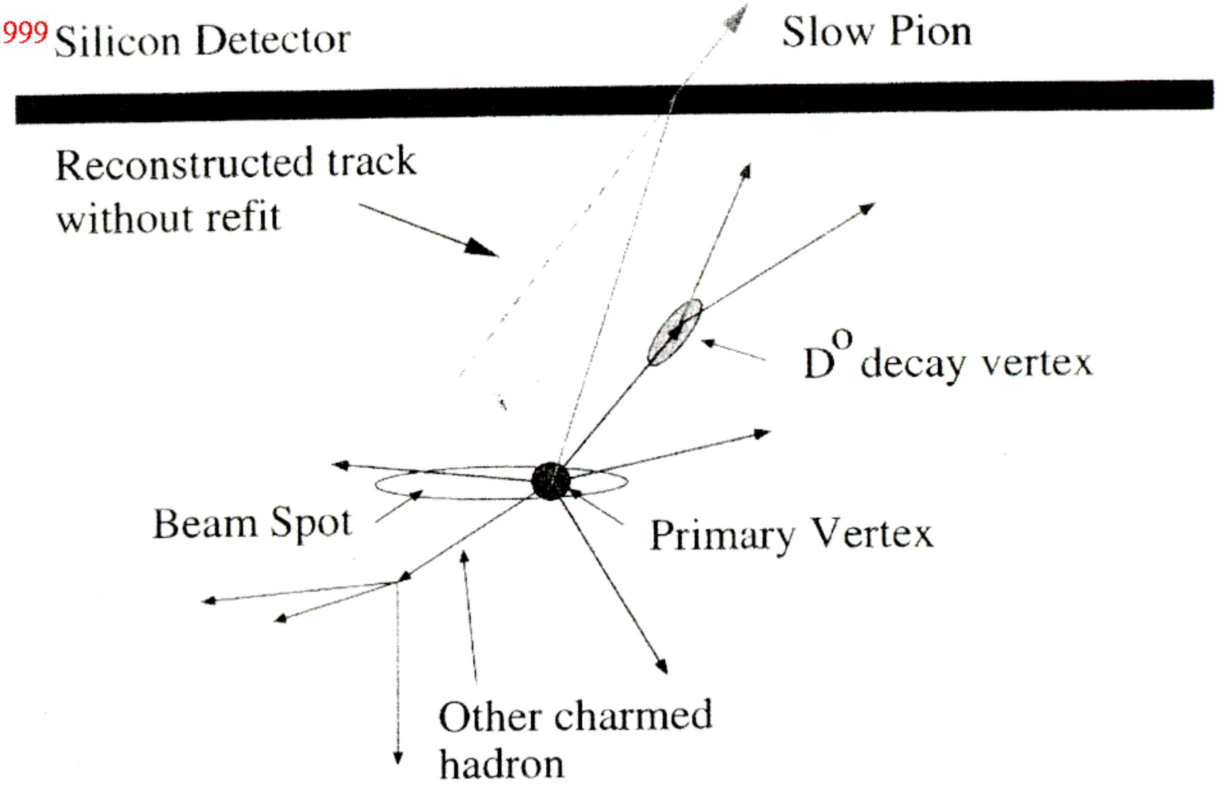


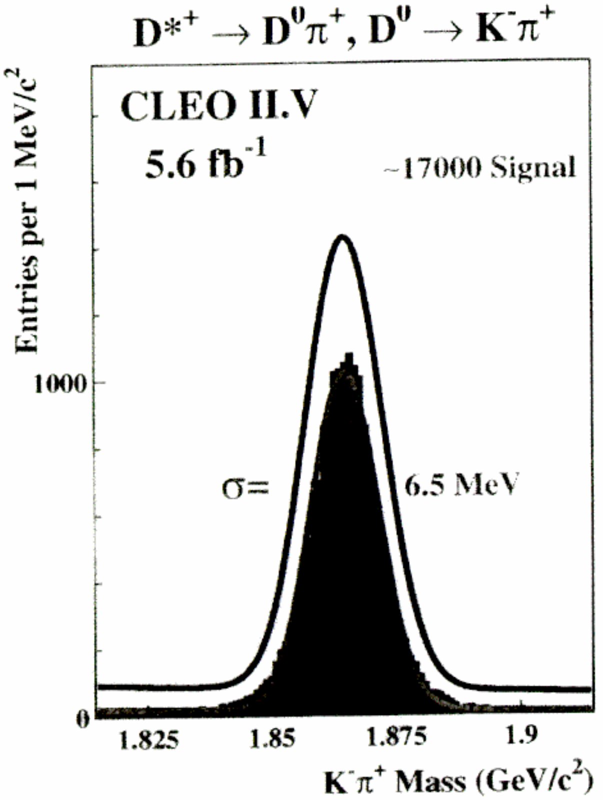
  - Doubly Cabibbo Suppressed decays
    - » $D^0 \rightarrow K^- \pi^+$ large
    - » $D^0 \rightarrow K^+ \pi^-$ small (about 1%)
- CLEO II set limit on sum of these of
  - $B(D^0 \rightarrow K^+\pi^-)/\ B(D^0 \rightarrow K^-\pi^+)=0.077\pm0.025\pm0.025$ which is $\approx 3 \tan^4(\theta_c)$ (used $1.8 fb^{-1}$)

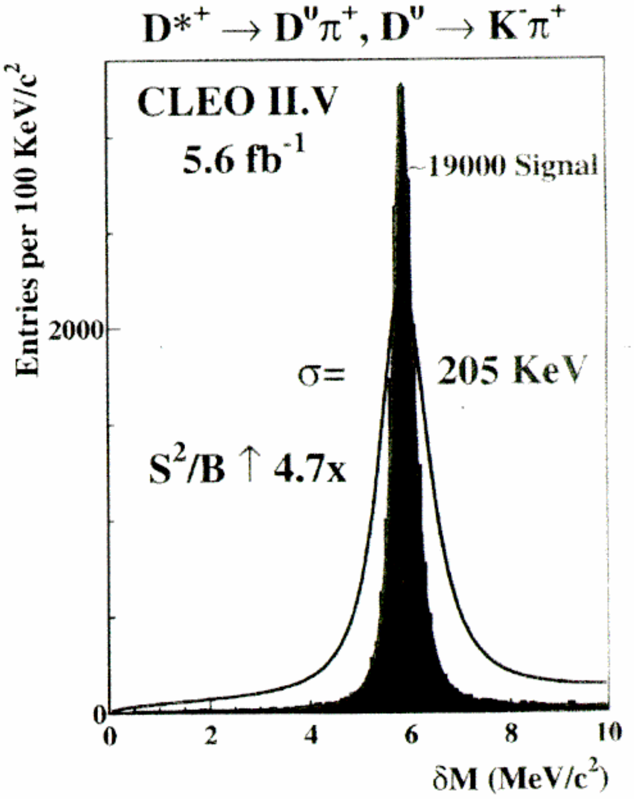
How do you separate these two effects?

- Time Dependence
  - DCSD give ratio of right/wrong sign decays at zero decay time
- Can study using
  - Coherent $D^0\bar{D}^0$ production at the $\Psi(3770)$ (Mark III)
  - OR
    - » Identify flavor at production using slow pion from decay $D^{*+} \rightarrow D^0\pi^+$
    - » Identify flavor at decay using
      - Semileptonic decay (unambiguous)
      - Hadronic decay (ambiguous due to DCSD, can fit to separate effects)


ALEPH
● Data
cos ϕ = 0
cos ϕ = +1
cos ϕ = −1
0.15
0.1
0.05
0
−0.05
−0.1
−2
−1
0
1
2
3
Proper time (ps)
N(D⁰ → K⁺π⁻) /N(D⁰ → K⁻π⁺)


Silicon Detector
Slow Pion
Reconstructed track
without refit
D0 decay vertex
Beam Spot
Primary Vertex
Other charmed
hadron


D*+ → D0π+, D0 → K-π+
CLEO II.V
5.6 fb-1
~17000 Signal
σ= 6.5 MeV
Entries per 1 MeV/c2
1000
0
1.825
1.85
1.875
1.9
K-π+ Mass (GeV/c2)


D*+ → D0π+, D0 → K-π+
CLEO II.V
5.6 fb-1
~19000 Signal
σ= 205 KeV
S2/B ↑ 4.7x
Entries per 100 KeV/c2
2000
0
0
2
4
6
8
10
δM (MeV/c2)