Search for B_s^0–$\overline{B_s^0}$ Oscillations with a Charge Dipole Technique at SLD

Dissertation
zur Erlangung des Doktorgrades
des Fachbereichs Physik
der Universität Hamburg

vorgelegt von
Julia Thom
aus Bonn

Hamburg
2001

Zusammenfassung

Das Thema dieser Arbeit ist die Studie der $B_s^0 - \overline{B_s^0}$ Oszillationen bei SLD. Die Studie gibt Aufschluß über noch ungenau gemessene Konstanten der Theorie der elektroschwachen Wechselwirkung. Die Studie bezieht sich auf B_s^0 Mesonen, die in 400,000 hadronischen Zerfällen des Z-Bosons von 1996 bis 1998 am SLC erzeugt wurden.

Zur Bestimmung der Oszillationsfrequenz muß die Ladung des Bottom-Quarks bei der Entstehung und beim Zerfall des B_s^0 Mesons gemessen werden. Außerdem muß die Zerfallszeit des B_s^0 Mesons bestimmt werden. Zur Bestimmung der Quarkladung bei der Entstehung wird die Asymmetrie von polarisierten Z-Boson Zerfällen in Bottom- und Anti-Bottom Quark Paare benutzt. Die Quarkladung beim Zerfall wird mit der Ladungsdifferenz zwischen inclusiv rekonstruierten B-Hadronen und D-Hadronen gemessen, und die Zerfallszeit wird mit den gemessenen Werten für Zerfallslänge und Energie des B-Hadrons bestimmt. Eine besondere Rolle fällt dabei dem hochauflösenden Silizium-Pixel-Detektor zu.

In der Studie ergab sich kein statistisch signifikantes Signal für eine bestimmte Oszillationsfrequenz. Die folgenden Werte der Oszillationsfrequenz Δm_s konnten mit 95% CL ausgeschlossen werden:

$$\begin{aligned} \Delta m_s &< 4.9 \text{ ps}^{-1}, \text{ und} \\ 7.9 \text{ ps}^{-1} < \Delta m_s &< 10.3 \text{ ps}^{-1}. \end{aligned}$$

Alle Ergebnisse von SLD, LEP und CDF wurden kombiniert. Damit kann Δm_s <14.6 ps^{-1} mit 95% CL ausgeschlossen werden. Das Ergebnis schränkt das Unitaritätsdreieck erheblich ein.

Abstract

We report a study of the time dependence of B_s^0–$\overline{B_s^0}$ mixing using a sample of 400,000 hadronic Z decays collected by the SLD experiment at the SLC between 1996 and 1998. The analysis determines the b hadron flavor at production by exploiting the large forward-backward asymmetry of polarized $Z \to b\overline{b}$ decays and uses additional information from the hemisphere opposite that of the reconstructed B decay. To determine the b hadron flavor at decay, the charge difference between inclusively reconstructed B and D decay vertices is used. The proper time of the decay is determined from the reconstructed B decaylength and a measure of the B hadron energy. The analysis takes advantage of the excellent vertexing efficiency and resolution of the pixel-based CCD Vertex Detector.

No significant signal for a particular value of the oscillation frequency was found. The following ranges of the oscillation frequency of B_s^0–$\overline{B_s^0}$ mixing are excluded at 95% CL from the analysis presented in this thesis:

$$\begin{aligned} \Delta m_s &< 4.9 \text{ ps}^{-1}, \text{ and} \\ 7.9 \text{ ps}^{-1} < \Delta m_s &< 10.3 \text{ ps}^{-1}. \end{aligned}$$

Combining results from SLD, LEP and CDF, Δm_s <14.6 ps^{-1} can be excluded at 95% CL. This limit puts significant constraints on the CKM unitarity triangle.

TABLE OF CONTENTS

LIST OF FIGURES 7

Chapter

1. Theoretical Motivation 1

1.1 Introduction 1
1.2 The Electroweak Interaction 2
1.3 The CKM Matrix 6
1.4 Determination of V_{ts} and V_{td} from B^0–$\overline{B^0}$ mixing 9

1.4.1 General discussion of heavy neutral meson systems . . 10
1.4.2 B^0–$\overline{B^0}$ mixing 12

1.5 CKM matrix elements from Tree Level Decays and Unitarity 17

1.5.1 V_{ud}, V_{us} and V_{cb} 17
1.5.2 V_{cd} and V_{cs} 18
1.5.3 $|V_{tb}|$ and $|V_{ub}|$ 19
1.5.4 Unitarity Constraints 20

1.6 CP Violation and New Physics in Mixing 21

1.6.1 New Physics in the B_s^0 system 23

2. Experimental Apparatus: the SLC and SLD 26

2.1 The SLC 26

2.1.1 Beam Transport 27
2.1.2 Polarized Electrons 28
2.1.3 Positrons 30
2.1.4 Acceleration and the Final Focus 30
2.1.5 The Beam Energy Measurement 31
2.1.6 The Electron Beam Polarization Measurement 32

2.2 The SLC Large Detector (SLD) 34

2.2.1 The Luminosity Monitor (LUM) 37
2.2.2 The Vertex Detectors (VXD2 and VXD3) 37
2.2.3 The Drift Chamber (CDC) 42
2.2.4 The Cherenkov Ring Imaging Detector (CRID) 45
2.2.5 The Liquid Argon Calorimeter (LAC) 48
2.2.6 The Warm Iron Calorimeter (WIC) 49

2.3 The SLD Event Trigger 50
2.4 The SLD Monte Carlo 53

3. Event Selection 55

3.1 Hadronic Event Selection 55
3.2 Heavy Flavor Tagging 56

3.2.1 IP Reconstruction 57
3.2.2 Neural Networks 58
3.2.3 Secondary Vertex Reconstruction 61
3.2.4 Track Attachment 64
3.2.5 VXD-Alone Tracking 64
3.2.6 Flavor Discrimination 69
3.2.7 Recovering Low Decay Length Vertices 72
3.2.8 Tracking Efficiency 73

4. Flavor Tagging and Proper Time Reconstruction 76

4.1 B flavor at production: the Initial State Tag 76

4.1.1 Polarized Forward-Backward Asymmetry 77
4.1.2 Jet Charge 78
4.1.3 Opposite Hemisphere Topological Tags 79
4.1.4 Combination of the Tags 79

4.2 B flavor at decay: Vertex Charge Dipole Analysis 81

4.2.1 Final State Tag Event Selection 83
4.2.2 Characterizing the event sample 92

4.3 Proper time reconstruction 96

4.3.1 Decay Length Reconstruction 97
4.3.2 Boost Reconstruction 99

4.4 Mixing Tag 102

4.4.1 Checking the Dipole Tagging Power 103

4.5 A Signal to Noise Estimate 105

5. Maximum Likelihood Analysis 111

5.1 B Hadron Physics Functions 111
5.2 Efficiency 114
5.3 Proper Time Resolution 116

5.3.1 Decay Length Resolution 117
5.3.2 Boost Resolution 120
5.3.3 Systematic Shifts in Proper Time Resolution 121

5.4 The Amplitude Fit 125

6. Results 136

6.1 B_s^0–$\overline{B_s^0}$ Mixing Results 136

6.1.1 B_d^0–$\overline{B_d^0}$ Mixing 141

6.2 Systematic Errors 142

6.2.1 B Hadron Lifetimes 143
6.2.2 Errors on the Proper Time Resolution 143
6.2.3 Initial and Final State Tag Purity 145

6.3 Does the Dipole Analysis work for B_s^0? 146
6.4 Setting a Limit on Δm_s 151
6.5 Combination of Mixing Analyses 153
6.6 Conclusions 156

APPENDICES

A. B decay modelling in the SLD Monte Carlo 162

A.1 The Composition of B Hadrons 162
A.2 The SLD B Decay Model 165

B. Ghost Track Algorithm 169

BIBLIOGRAPHY 173

LIST OF FIGURES

Figure Page

1. Unitarity Triangle 8

2. Box Graphs for B^0–$\overline{B^0}$ mixing 12

3. The probability for a B_s^0 decaying as a B_s^0, as a function of the proper time. The plots show $|g_\pm(t)|^2$ for different values of $\Delta\Gamma$ and Δm_s: the solid line was obtained with the values $\Delta m_s = 20$ ps^{-1} and $\Delta\Gamma = 0$ ps^{-1}, and the dashed line for the values $\Delta m_s = 0$ ps^{-1} and $\Delta\Gamma = 0.075$ ps^{-1}. The dotted line was calculated for $\Delta m_s = 0$ ps^{-1} and a very large assumed value of $\Delta\Gamma = 10\times$ 0.075 ps^{-1}. 16

4. Layout of the SLC. 27

5. Energy state diagram for (top) bulk GaAs and (bottom) the strained GaAs lattice. The encircled numbers denote the relative intensities of the transitions. The strained lattice breaks the degeneracy and allows a theoretical polarization of 100%. 29

6. SLC/SLD luminosity history. In the 1997-8 data run, the luminosity improved substantially and reached 20,000 Z^0/week. 31

7. SLD Wire Imaging Synchrotron Detectors (WISRD) 32

8. SLD Compton Polarimeter 34

9. SLD detector in isometric view 36

10. SLD detector in quadrant view 36

11. SLD Luminosity Monitor (LUM) 37

12. Two-CCD VXD ladder. 39

13. Schematic layout of VXD3 in the xy view. 39

14. Schematic layout of VXD3 in the z view. 39

15. Comparison of VXD2 and VXD3 in xy plane 40

16. Miss distance of tracks in $Z \to \mu^+\mu^-$ events, in the xy and z projections. 42

17. SLD Central Drift Chamber (CDC) 43

18. CDC drift cell layout 44

19. Field map of a CDC drift cell 44

20. Barrel CRID sector in axial and transverse views 46

21. The Cherenkov angle for the liquid (solid line) and gas radiators (dashed line) in the barrel CRID as a function of momentum for the three long lived hadronic particle species. 47

22. EM and HAD modules in the barrel LAC. 49

23. Cutaway view of a WIC section, including details of single layers and double layers. 50

24. Distributions of event selection variables: (a) Number of CDC tracks, (b) visible energy. The points (histogram) denote the data (Monte Carlo). The arrows indicate the regions which pass the cuts, which are applied cumulatively for each successive plot. A subset of the 1998 data is shown. 57

25. Muon-pair xy impact parameter to the IP. 59

26. Schematic diagram of a typical feed-forward neural network. 60

27. The track (a) and vertex (b) functions projected onto the xy plane (cm). Primary and secondary vertices are visible in (b). 63

28. Distributions of seed vertex selection variables: (a) distance from IP, (b) normalized distance from IP, (c) angle between flight direction and vertex momentum, (d) neural network output. The arrow indicates the accepted region. 65

29. Schematic illustration of the quantities used in the track-attachment procedure described in the text (not to scale). 66

30. Distributions of cascade track selection variables: (a) T, (b) L, (c) L/D, (d) α, (e) b/σ_b, (f) neural network output. The arrow indicates the accepted region. 67

31. Distributions of VXD-vector selection variables: (a) T, (b) L, (c) L/D, (d) α, (e) neural network output. The arrow indicates the accepted region. The probability to assign the correct charge to a vector based on its fitted curvature is shown in (f), both with and without the secondary vertex as a constraint. 68

32. Schematic illustration of the missing P_t determination. 70

33. Distributions of flavor discrimination variables: (a) M_{VTX}, (b) P_{VTX}, (c) $M_{VTX}-P_{VTX}$ correlation, (d) D_{VTX}, (e) N_{trk}, (f) neural network output. 71

34. The all-flavor MC B decay length distribution of standard NN selected events, events reconstructed with quality tracks and *udsc* background (hatched histograms). The points are the 96-98 SLD data. 73

35. The overall reconstructed charge of b-tagged events. The points are the data, the histogram is the all-flavor Monte Carlo, and the shaded histogram shows the fraction of true neutral B decays. 75

36. Distribution of the initial state b-quark probability using the polarized forward-backward asymmetry tag. Data (points) and Monte Carlo (histograms) showing the b and $\bar{b}$ components for the events selected in the charge dipole analysis. 78

37. Distribution of the computed initial state b-quark probability for data (points) and Monte Carlo (histograms) showing the b and $\bar{b}$ components for the events selected in the charge dipole analysis. 81

38. Distributions of a)B decay length, b)B to D distance, c)total B momentum and d)cosΘ of the thrust axis for data (points) and Monte Carlo (histograms). 84

39. Distributions of (a)Number of tracks attached to the primary vertex, (b)B vertex track multiplicity, (c)D vertex track multiplicity, (d)charge difference between B and D vertices, (e)number of tracks and (f)number of VXD-alone vectors for data (points) and Monte Carlo (histograms) in the charge dipole analysis. 86

40. Distribution of the charge dipole for data (points) and Monte Carlo (solid histogram). Also shown are the contributions from b hadrons containing a b quark (dashed histogram) or a $\bar{b}$ quark (dotted histogram). 87

41. The Charge Dipole correct tag probability for B_s^0 decays as a function of the charge dipole magnitude, as extracted from the Monte Carlo. The solid line is the result of a fit to the points. 88

42. Monte Carlo distributions of $M(D^\star) - M(D)$ for 4 different cases: a)neutral B decays, b)charged B decays, c)neutral B decays with m_D >1 GeV and d)charged B decays with m_D >1 GeV. 90

43. The distribution of $M(D^\star) - M(D)$ for 4 different cases in the 96-98 SLD data set: a)neutral B decays, b)charged B decays, c)neutral B decays with m_D >1 GeV and d)charged B decays with m_D >1 GeV. 91

44. The distribution of the relative b hadron fractions as a function of $\cos\Theta$. The functions are the result of fits. They are used to scale the average b hadron fractions. 93

45. Charge dipole correct tag probability as a function of reconstructed decay length in simulated B_s^0 decays. a)all decay modes, b)$B_s^0 \to \overline{D^0}X$, c)$B_s^0 \to D^-X$, d)$B_s^0 \to D_sX$, e)$B_s^0 \to c$ baryons, f)$B_s^0 \to D\overline{D}$. The functions are the fit results used to parametrize the charge dipole correct tag probability as a function of decay length. 96

46. The decay mode fractions as a function of the B_s^0 vertex mass. a)$B_s^0 \to \overline{D^0}X$, b)$B_s^0 \to D^-X$, c)$B_s^0 \to D_sX$, d)$B_s^0 \to$ baryons, e)$B_s^0 \to D\overline{D}$. 97

47. The reconstructed boost of events selected in the dipole analysis. The histogram denotes the Monte Carlo distribution, and the points show the reconstructed boost in the data. 101

48. The reconstructed proper time of events selected in the charge dipole analysis. The histogram denotes the Monte Carlo distribution, and data points show the reconstructed proper time in the data. 102

49. Distributions of $\cos\theta$ for the thrust axis direction signed by the product $(\delta Q \times P_e)$ for data (points) and Monte Carlo (histograms) in subsamples with $Q_{tot} = 0$ and $Q_{tot} = \pm 1$ for the charge dipole analysis. 103

50. The fraction of events tagged as mixed in the dipole sample, as a function of the reconstructed proper time. The histogram is taken from the MC, the data points are extracted from the 96-98 data. . 104

51. The fraction of events tagged as mixed in the dipole sample, as a function of the probability that an event was tagged incorrectly, shown in 4 bins of reconstructed proper time. The last plot is integrated over all time bins. The histograms are taken from the MC, the data points are extracted from the 96-98 data. 106

52. A plot of the analytical function of the mixed fraction assuming a typical set of experimental parameters and $\Delta m_s = 20ps^{-1}$, for $\sigma_L = 60\mu$m (SLD-like) and $\sigma_L = 200\mu$m (LEP-like). 110

53. The proper time distribution of the small contribution of background events in the SLD Monte Carlo. The function is the fit result used to parametrize the light flavor background component as a function of true proper time. 113

54. The event selection efficiency per b hadron in the dipole analysis (MC simulation), as a function of true proper time. The functions are the fit results used to parametrize the charge dipole efficiency as a function of true proper time. 115

55. σ_t as a function of t, plotted analytically using average values for the B_s^0 core fraction of σ_L and $\sigma_{\gamma\beta}/\gamma\beta$. 117

56. The widths of the Gaussian functions fitted to these MC distributions of $\frac{L_{rec}-L_{true}}{\sigma_{meas}}$ are the scale factors for correctly and incorrectly tagged events used to adjust the event-by-event decay length resolution. a) correctly tagged $B_s^0 \to D_s X$ decays, b) incorrectly tagged $B_s^0 \to D_s X$ decays, c) correctly tagged $B_s^0 \to D\overline{D}X$ decays, and d) incorrectly tagged $B_s^0 \to D\overline{D}X$ decays. Two Gaussian functions are fitted with 60% of the decays in the core and 40% of the events in the tail. 119

57. The widths of the Gaussians fitted to these MC distributions are the residuals for correctly (left) and incorrectly (right) tagged B_s^0 decays with one track in each of the B and D vertices. Two Gaussians are fitted with 60% of the decays in the core and 40% of the events in the tail. 120

58. The width of the relative boost residual as a function of the charged B energy for the 4 different b hadrons. The open squares represent the 60% core fraction of the decays. The dark squares represent the 40% tail fraction. The functions are the fit results used to parametrize the relative boost resolution as a function of the charged B energy. 121

59. The proper time residuals, $t_{rec} - t$, in 6 bins of true proper time, for the case of $B_s^0 \to D\overline{D}X$ decays. The data points are extracted from the Monte Carlo, and the histogram is the calculated resolution function. 124

60. The likelihood fits for A at four different values of Δm_s: a) 5 ps^{-1}, b) 10 ps^{-1}, c) 15 ps^{-1} and d) 20 ps^{-1}. 127

61. Amplitude fit to a pure Monte Carlo B_s^0 sample, selected by the dipole charge analysis. The Monte Carlo Δm_s input value is $10ps^{-1}$. . . . 128

62. Amplitude fit to a Monte Carlo $b\overline{b}$ sample, selected by the dipole charge analysis. The Monte Carlo Δm_s input value is $10ps^{-1}$. 129

63. Amplitude fit to a Monte Carlo $b\overline{b}$ sample, selected by the dipole charge analysis. The Monte Carlo Δm_s input value is $15ps^{-1}$. 130

64. Amplitude fit to the SLD Monte Carlo sample, selected by the dipole charge analysis. The Monte Carlo Δm_s input value is $200ps^{-1}$. Also shown are the amplidude A divided by σ_A and A/σ_A versus Δm_s. . 132

65. Amplitude fits for 12 data-size Monte Carlo samples, where the input value of Δm_s has been set to 200 ps^{-1}. 133

66. A/σ_A for the 12 independent MC experiments. A/σ_A is plotted for well-separated values of A atΔm_s=1 ps^{-1}, 5 ps^{-1}, 10 ps^{-1}, 15 ps^{-1} and 20 ps^{-1} only. The function is the result of a Gaussian fit to the histograms. 134

67. The mixed fraction for the four b hadron types as a function of reconstructed time. The points are the Monte Carlo, and the histogram represents the likelihood function for each b hadron type separately. The input B_s^0–$\overline{B_s^0}$ mixing frequency is $\Delta m_s = 10$ ps^{-1}. 135

68. Measured amplitude for the full 96-98 SLD data set using the charge dipole analysis described in this thesis. The 1σ error bars are statistical only. The dashed line is $1.645\sigma_A$. The Δm_s value for the intersection between this line and the A = 1 line represents the sensitivity of the experiment. 137

69. Measured log$\mathcal{L}$ fit for the full 96-98 SLD data set using the charge dipole analysis described in this thesis. 138

70. The mixed fraction as a function of reconstructed proper time, for dipole selected events. The datapoints are from the 96-98 data, the histogram is a representation of the likelihood function used to extract the Amplitude fit. The likelihood function was generated with $\Delta m_s = 25ps^{-1}$. 139

71. The mixed fraction as a function of reconstructed proper time. Shown are the 96-98 data selected by the dipole analysis (points) overlaid with four different likelihood functions, generated with (a)A=1.0 and Δm_s=25 ps^{-1}, (b)A=0.6 and Δm_s=7 ps^{-1}, (c)A=0.6 and Δm_s=12 ps^{-1} and (d)A=1.8 and Δm_s=17 ps^{-1}. 140

72. The amplitude fit with a cut at t_{rec}=0.2ps^{-1}. 141

73. The amplitude fit for the full data set 96-98 in the charge dipole analysis.Δm_s is fixed and log$\mathcal{L}$ is minimized for A at each value of Δm_d. 142

74. The systematic errors due to uncertainties in B_s^0 production fraction, resolution model, initial and final state tag, decay length resolution and relative boost resolution. The solid curves correspond to the positive variation of the parameter, while the dotted curves correspond to the negative variation. 147

75. The B_s^0 fraction output parameter. The points are the SLD data, the histogram is the Monte Carlo. The shaded histogram shows the true MC B_s^0 content in every bin. 148

76. The MC b fraction versus the B_s^0 fraction NN output parameter, for each b hadron separately. The functions are the result of a fit to the MC points. 149

77. The dipole distribution for a small event sample selected by a NN to enhance the fraction of B_s^0 decays. The points are the data, the histograms MC. Also shown are the contributions from b hadrons containing a b quark (dashed histogram) or a $\bar{b}$ quark (dotted histogram). 150

78. The dipole distribution for a small event sample of mixed and leptonically tagged decays. The points are the data, the histograms MC. Also shown are the contributions from b hadrons containing a b quark (darkly shaded histogram) or a $\bar{b}$ quark (lightly shaded histogram). 151

79. The amplitude fit for the full 96-98 SLD data set from the charge dipole analysis including systematic errors. The solid curve bordering on the lightly shaded region is $A + 1.645\sigma_A$ (stat) and the solid curve bordering on the darker region is $A + 1.645\sigma_A$ (stat+sys). The dashed curve is $1.645\sigma_A$(stat+sys) which represents the sensitivity of the experiment. 152

80. Measured amplitude as a function of Δm_s for the lepton+D, Ds+tracks, and charge dipole analyses combined. 155

81. Measured amplitude as a function of Δm_s for the CDF, LEP and SLD Amplitudes combined. 157

82. Comparison of σ_A at $\Delta m_s = 17ps^{-1}$ for the CDF, LEP and SLD experiments. 158

83. Current constraints on ρ and η. 161

CHAPTER 1

Theoretical Motivation

1.1 Introduction

Quantum chromodynamics (QCD), the electroweak theory and general relativity currently give an accurate description of all known phenomena in fundamental physics. No experimental facts are known that disagree with these theories. The Standard Model (SM) incorporates QCD and the electroweak theory. Its matter content consists of spin 1/2 quarks and leptons, and spin 1 gauge bosons. They are considered to be elementary, meaning that they have no constituents. There are six quark flavors (up, down, charm, strange, top and bottom) and in addition to its spin each of the quarks has a second degree of freedom, called color, which can have three different values. There are also six different leptons, e, ν_e, μ, ν_μ, τ and ν_τ. Forces between fermions due to the electroweark interaction are mediated by exchanges of four gauge bosons ($\gamma, W^+, W^-$and Z^0); and the forces due to the strong interaction, or QCD, are mediated by eight gauge bosons, the gluons. An in depth description of the SM can be found in [1], [2].

Together the 12 fermions are arranged in 3 families, ordered according to their masses, such that each family contains a pair of up- and down-type quarks, as well as one charged lepton and one neutrino.

A set of twenty-five fundamental constants are required in order to describe the particle masses and the strength of the interactions between them. For example, the electroweak interaction can be expressed in terms of three constants, the electric charge e, the weak charge g, and the mass of either the W or Z. The strengths of the quark flavor changing transitions can be described using a unitary 3×3 matrix, the Cabibbo-Kobayashi-Maskawa (CKM) matrix. Its elements can be described in terms of four constants.

The determination of the Standard Model constants is an important goal of particle physics today. It is conceivable that future measurements will give results that do not fit into the Standard Model framework and would question its validity.

This thesis is dedicated to putting tighter constraints on two poorly measured SM parameters, ρ and η, two of the four constants that determine the CKM quark transition matrix. They are related to the coupling of the top quark to lighter quarks. A very promising way to do this is to measure the rates for B^0–$\overline{B^0}$ mixing, i.e., the conversion of neutral B mesons into their anti-particles. Here, we will review how measurements of the B_s^0 and B_d^0 mixing rates can be combined to constrain the two SM parameters in question. A summary of how non-SM physics could affect the B^0–$\overline{B^0}$ mixing rates will also be given.

1.2 The Electroweak Interaction

The theory of electroweak interactions is based on an $SU(2)_L \times U(1)$ local gauge symmetry [3]. The i'th family of quarks consists of left-handed up and down type quarks, u_{L_i} and d_{L_i}, which transform as an $SU(2)_L$ doublet $\Psi_{L_i} = \begin{pmatrix} u_L \\ d_L \end{pmatrix}_i$, and right handed $SU(2)_L$ singlet quarks u_R^i and d_R^i. Similarly, each family of leptons

contains an $SU(2)_L$ doublet and a right handed $SU(2)_L$ singlet charged lepton. All of the fields are charged under the $U(1)$ symmetry. The requirement of local gauge invariance implies the existence of 4 massless spin 1 bosons, which are the mediators of the electroweak interactions. Three of them, the $W^+, W^-$and Z^0, gain mass due to electroweak symmetry breaking, whereas the photon remains massless.

Table 1: Standard Model fermions [4]

quarks	Q/e	mass[GeV]	leptons	Q/e	mass[GeV]	family
up	2/3	≈0.001-0.005	e	-1	0.005	1
down	-1/3	≈0.003-0.009	ν_e	0	$< 3 \cdot 10^{-6}$	
charm	2/3	≈1.15-1.35	μ	-1	0.106	2
strange	-1/3	≈0.075-0.17	ν_μ	0	$< 2 \cdot 10^{-4}$	
top	2/3	174.3±5.1	τ	-1	1.778	3
bottom	-1/3	≈4.0-4.4	ν_τ	0	< 0.018	

Table 2: The electroweak gauge bosons

boson	Q/e	mass[GeV]
γ	0	0
Z^0	0	91.19
$W^\pm$	± 1	80.42 ± 0.06

The fermions couple to the gauge bosons γ and Z via neutral currents and to the $W^\pm$ via charged currents. The transitions measured in this thesis are due to charged currents, so their structure is reviewed here. The mechanism of quark mass generation is also briefly discussed.

In the SM particle masses are generated via the Higgs-mechanism. A description can be found in [5]. In the simplest version a single $SU_L(2)$ doublet of scalar

fields $\Phi \equiv \begin{pmatrix} \Phi^+ \\ \Phi^0 \end{pmatrix}$ is introduced. The potential for Φ is arranged so that the neutral component acquires a vacuum expectation value, $< \Phi^0 >= \frac{v}{\sqrt{2}}$, leading to spontaneous symmetry breaking. Of the four degrees of freedom in Φ, three become the longitudinal components of the massive $W^{\pm}$ and Z^0, leaving a single physical neutral Higgs scalar H. Unfortunately, the Higgs field has as yet not been observed, so that a complete understanding of the mechanism for electroweak symmetry breaking remains one of the central problems in particle physics today.

The Lagrangian contains 'Yukawa' terms for the coupling of the quarks to Φ:

$$\mathcal{L}^{\text{Yukawa}} = -\left[h_U \overline{\Psi}_L \tilde{\Phi} u_R + h_D \overline{\Psi}_L \Phi d_R\right] + h.c., \tag{1.1}$$

where h_U and h_D are 3×3 Yukawa coupling matrices containing the dimensionless coupling strengths, and $\tilde{\Phi} \equiv i\sigma_3 \Phi^\star = \begin{pmatrix} \overline{\Phi}^0 \\ -\Phi^- \end{pmatrix}$. Substituting the vacuum expectation value for Φ^0 into $\mathcal{L}^{\text{Yukawa}}$ yields up and down quark mass matrices:

$$m_U = h_U \cdot \frac{v}{\sqrt{2}} \quad \text{and} \quad m_D = h_D \cdot \frac{v}{\sqrt{2}}. \tag{1.2}$$

The 3×3 mass matrices m_U and m_D can be transformed into real, diagonal matrices describing the physical mass eigenstates via bi-unitary transformations:

$$V_L^U m_U V_R^{U\dagger} \equiv \begin{pmatrix} m_u & 0 & 0 \\ 0 & m_c & 0 \\ 0 & 0 & m_t \end{pmatrix} \quad \text{and} \quad \mathrm{V_L^D m_D V_R^{D\dagger}} \equiv \begin{pmatrix} m_d & 0 & 0 \\ 0 & m_s & 0 \\ 0 & 0 & m_b \end{pmatrix}. \tag{1.3}$$

The Lagrangian for the charged current interactions of the $W^{\pm}$ with the quarks in the weak interaction basis can be written as

$$\mathcal{L}^{CC} = \frac{e}{\sqrt{2}\sin\Theta_W} \sum_{i=1}^{3} \overline{u}_L^i \gamma^\mu W_\mu^+ d_L^i + h.c., \tag{1.4}$$

where Θ_W is the weak mixing angle and e the electric charge. The three 'physical' up and down type quark mass eigenstates are given by

$$\overline{u}_L^{\mathrm{Phy^i}} = (V_L^U \overline{u}_L)^i \quad \text{and} \quad \mathrm{d_L^{Phy^i}} = (\mathrm{V_L^D d_L})^{\mathrm{i}}, \ \mathrm{i} = 1, 2, 3. \tag{1.5}$$

The physical up quarks are denoted by u,c,t (for up, charm, top), and the down quarks by d,s,b (for down, strange, bottom). The CKM matrix [6], [7] is defined as

$$V_{CKM} \equiv V_L^U V_L^{D^\dagger}. \tag{1.6}$$

It is unitary, but not diagonal. Finally we can rewrite the charged current Lagrangian in the physical quark basis as:

$$\mathcal{L}^{CC} = \frac{e}{\sqrt{2}\sin\Theta_W} (\overline{u}\ \overline{c}\ \overline{t})_L \gamma_\mu W^\mu V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix}_L . \tag{1.7}$$

The elements of the CKM matrix

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{1.8}$$

describe the "mixing"-strengths of the quarks, i.e. the strengths of the charged current transitions between up and down type quarks.[1] The elements of V_{CKM} are not predicted in the SM, but can in principle be determined by measuring the strength of transitions between quarks.

[1] Up to about two years ago, neutrino masses were assumed to vanish in the Standard Model, which excluded lepton flavor changing transitions. However, recent experimental results strongly suggest non zero neutrino masses and the existence of transitions between at least two lepton families. It is possible therefore, that a mixing matrix analogous to the CKM matrix exists in the lepton sector.

Table 3: Relative strengths of quark transitions

relative amplitude	transition	example of source of information
~ 1	u,d	Nuclear β decay
~ 1	c,s	$D^0 \to K^- e^+ \nu_l$
~ 0.22	u,s	$\overline{K}^0 \to \pi^+ e^- \overline{\nu}_e$
~ 0.22	c,d	$D^0 \to \pi^- e^+ \nu_l$
~ 0.04	c,b	$B \to X_c l \nu_l$
~ 0.04	t,s	$b \to s\gamma$
~ 1	t,b	$t \to bW$
~ 0.003	u,b	$B \to \pi l \nu$
~ 0.01	t,d	$B^0 - \overline{B}^0$ mixing

1.3 The CKM Matrix

A hierarchy of quark flavor-changing transitions is observed experimentally in the sense that transitions between the first and second families are about an order of magnitude stronger than transitions between the second and third. Transitions between the first and third families are even weaker. Table 3 summarizes the approximate observed strengths of the quark transitions and lists one source of information for each. A more detailed discussion will be given in section 1.5.

The CKM matrix is usually parametrized in some specific way. The purpose of the parametrizations is to incorporate the constraints of 3×3 unitarity. Some parametrizations also reflect experimental information, like the Wolfenstein parametrization, which is based on the flavor transition strength hierarchy. It is a parametrization in which unitarity only holds approximately. Other possible parametrizations can for example be based on the observed mass hierarchies.

In general a complex $N \times N$ matrix V is parametrized by $2N^2$ parameters. The

number of independent parameters is halved if unitarity is imposed:

$$\sum_j V_{ij} V_{kj}^* = \delta_{ik} \tag{1.9}$$

As the phases are arbitrary, $2N-1$ of them can be absorbed by phase rotations. Furthermore, a unitary matrix is a complex extension of an orthogonal matrix, therefore $\frac{1}{2}N(N-1)$ parameters need to be identified with rotation angles, leaving $\frac{1}{2}N(N-3)+1$ phases. Thus, the appearance of a complex phase requires $N > 2$.

In the case of the unitary CKM matrix with N=3, four free parameters are left after phase rotations and unitarity conditions are imposed. Three of them can be identified with the real Euler angles, leaving one phase. This phase allows for the accommodation of CP violation. A discussion follows in section 1.6.

In the Wolfenstein parametrization [8], three real parameters (λ, A, ρ) and a phase η representing CP violation, are arranged in the following way:

$$V_{CKM} = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + O(\lambda^4). \tag{1.10}$$

The small parameter λ serves as an expansion parameter: transitions between the first and second families are of order λ, between the second and third of order λ^2, and between the first and third of order λ^3. λ and A are identified with V_{us} and $\frac{|V_{cb}|}{\lambda^2}$, which are determined from the precisely measured semileptonic K and B decays [4].

$$\lambda = |V_{us}| = 0.2205 \pm 0.0018, \tag{1.11}$$

$$A = \frac{|V_{cb}|}{\lambda^2} = 0.824 \pm 0.075. \tag{1.12}$$

The parameters ρ and η, on the other hand, are not very well known.

A way to illustrate the unitarity relations between the different matrix elements is via a triangle in the complex (ρ, η) plane. Fig. 1 gives the full description of the

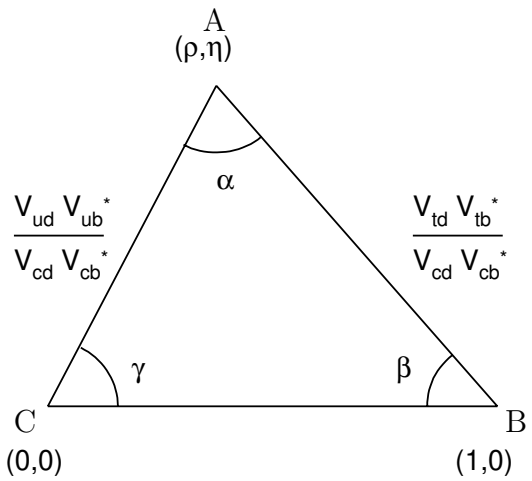


Figure 1: Unitarity Triangle in the Wolfenstein Parametrization.

CKM matrix in the Wolfenstein approximation. This 'unitarity triangle' is obtained by rescaling one of the nine unitarity relations

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0, \tag{1.13}$$

by a factor $|V_{cd}V_{cb}^*|$ so that the base is normalized to unity. The lengths CA and BA are equal to $\left|\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right|$ and $\left|\frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*}\right|$, respectively.

One can achieve a greater level of precision in a similar parametrization with redefined parameters $(\tilde{\rho}, \tilde{\eta})$, where higher orders of λ are included and unitarity holds to an excellent approximation. With

$$\tilde{\rho} \equiv \rho(1 - \frac{\lambda^2}{2}) \quad \text{and} \quad \tilde{\eta} \equiv \eta(1 - \frac{\lambda^2}{2}), \tag{1.14}$$

one obtains

$$V_{CKM} = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\tilde{\rho}-i\tilde{\eta}) & -A\lambda^2 + \frac{1}{2}A(1-2\rho)\lambda^4 - i\eta A\lambda^4 & 1 \end{pmatrix}. \tag{1.15}$$

The corresponding lengths CA and BA in the complex $(\tilde{\rho}, \tilde{\eta})$ plane are equal to

$$R_b \equiv \sqrt{\tilde{\rho}^2 + \tilde{\eta}^2} = (1 - \frac{\lambda^2}{2})\frac{1}{\lambda}\left|\frac{V_{ub}}{V_{cb}}\right| \quad \text{and} \tag{1.16}$$

$$R_t \equiv \sqrt{(1-\tilde{\rho})^2 + \tilde{\eta}^2} = (1 - \frac{\lambda^2}{2})\frac{1}{\lambda}\left|\frac{V_{td}}{V_{cb}}\right|. \tag{1.17}$$

In order to determine if the weak hadronic decays are correctly described in the SM it is essential to over-determine the CKM parameters $\tilde{\rho}$ and $\tilde{\eta}$ by making as many independent measurements as possible. In the language of the unitarity triangle this means that the bands in the $(\tilde{\rho}, \tilde{\eta})$ plane extracted from all measurements should overlap at the apex of the triangle.

In the next section, the possibility of obtaining a constraint on $\tilde{\rho}$ and $\tilde{\eta}$ via measurements of the B^0–$\overline{B^0}$ mixing rates is discussed.

1.4 Determination of V_{ts} and V_{td} from B^0–$\overline{B^0}$ mixing

Flavor-changing neutral current processes that are generated at the one loop level in the SM play a crucial role in the determination of the CKM matrix elements V_{ts} and V_{td}. However, the extraction of these elements is complicated by the presence of hadronic uncertainties. New Physics can also make significant contributions to one loop processes. This will be discussed in section 1.6.

Examples of such processes are rare decays of neutral mesons, B_d^0, B_s^0, K^0 and D^0, and neutral meson mixing, which we discuss here.

1.4.1 General discussion of heavy neutral meson systems

Consider a beam of neutral mesons P^0. It is described by the wave function

$$|\Psi(t)>=\Psi_1(t)|P^0>+\Psi_2(t)|\overline{P^0}>, \tag{1.18}$$

where t is the proper time. P^0 and $\overline{P^0}$ are flavor eigenstates. The wave function evolves according to the time dependent Schroedinger equation

$$i\frac{d}{dt}\begin{pmatrix}\Psi_1\\ \Psi_2\end{pmatrix}=\begin{pmatrix}R_{11} & R_{12}\\ R_{21} & R_{22}\end{pmatrix}\begin{pmatrix}\Psi_1\\ \Psi_2\end{pmatrix} \tag{1.19}$$

with

$$R=M-\frac{i}{2}\Gamma. \tag{1.20}$$

M and Γ are 2×2 hermitian matrices known as the mass and decay matrices. They describe the dispersive and absorptive parts of P^0–$\overline{P^0}$ mixing. Hermiticity implies $M_{12}=M_{21}^*$ and $\Gamma_{12}=\Gamma_{21}^*$, however R itself is not hermitian. M_{ij} and Γ_{ij} can in principle be calculated from second order perturbation theory. CPT invariance guarantees that $M_{11}=M_{22}$ and $\Gamma_{11}=\Gamma_{22}$.

The flavor eigenstates P^0 and $\overline{P^0}$ are not the physical mass eigenstates. The latter are obtained by solving the eigenvalue problem for R. The two eigenstates of R may be labeled as P_a and P_b. In the B^0–$\overline{B^0}$ system it is customary to choose a=H and b=L for heavy and light states. In the K^0 system, where the lifetimes of K_a and $\overline{K}_b$ are widely different, the labels L and S (for long and short lifetime) are used to distinguish the two physical eigenstates.

The heavy and light mass eigenstates can be written as a quantum mechanical superposition of flavor eigenstates:

$$|P_H>=p|P^0>+q|\overline{P^0}>, \tag{1.21}$$

$$|P_L >= p|P^0 > -q|\overline{P^0} > . \tag{1.22}$$

Since R is not hermitian their eigenvalues are complex:

$$\mu_H = m_H - \frac{i}{2}\Gamma_H \tag{1.23}$$

$$\mu_L = m_L - \frac{i}{2}\Gamma_L, \tag{1.24}$$

where $m_{H,L}$ and $\Gamma_{H,L}$ denote the masses and widths of the physical states P_H and P_L. Using the convention

$$\Delta\mu \equiv \mu_H - \mu_L = \Delta m - \frac{i}{2}\Delta\Gamma \tag{1.25}$$

$$\Delta m \equiv m_H - m_L \tag{1.26}$$

$$\Delta\Gamma \equiv \Gamma_H - \Gamma_L, \tag{1.27}$$

the eigenvalue problem

$$\det\ |\mathrm{M} - \frac{\mathrm{i}}{2}\Gamma - \mu| = 0 \tag{1.28}$$

leads to the condition

$$\Delta\mu = 2\sqrt{(M_{12}^* - i\Gamma_{12}^*/2)(M_{12} - i\Gamma_{12}/2)}. \tag{1.29}$$

The real and imaginary parts of this equation give

$$\Delta m^2 - (\Delta\Gamma^2/4) = 4(|M_{12}|^2 - \frac{1}{4}|\Gamma_{12}|^2) \ \text{ and}$$

$$\Delta m\Delta\Gamma = 4Re(M_{12}\Gamma_{12}^*). \tag{1.30}$$

Finally, using these two equations one can write down the mass and decay width differences in terms of M_{12} and Γ_{12}:

$$\Delta m = \sqrt{2}\sqrt{|M_{12}|^2 - \frac{1}{4}|\Gamma_{12}|^2 + \sqrt{(|M_{12}|^2 - \frac{1}{4}|\Gamma_{12}|^2)^2 + [Re(M_{12}\Gamma_{12}^*)]^2}}, \tag{1.31}$$

$$\Delta\Gamma = 2\sqrt{2}\left(\sqrt{|M_{12}|^2 - \frac{1}{4}|\Gamma_{12}|^2 + [Re(M_{12}\Gamma_{12}^*)]^2} - (|M_{12}|^2 - \frac{1}{4}|\Gamma_{12}|^2)\right)^{\frac{1}{2}}. \tag{1.32}$$

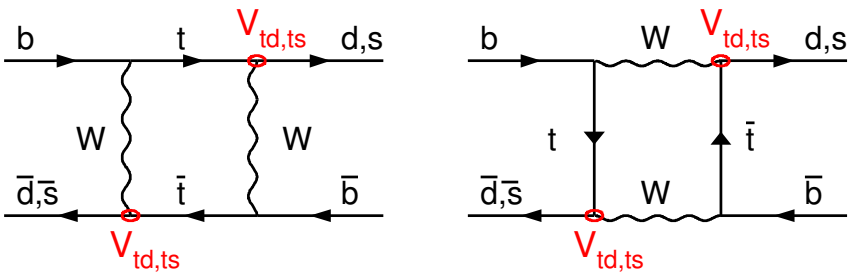


Figure 2: Box Graphs for B^0–$\overline{B^0}$ mixing. Shown is the dominant contribution due to the exchange of a top quark.

Solving for the eigenvalues yields

$$\frac{q}{p} = \frac{-\Delta\mu}{2(M_{12} - i\Gamma_{12}/2)} = \pm\sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}}. \tag{1.33}$$

In the next section, we apply the above to the case of B^0–$\overline{B^0}$ mixing in the SM.

1.4.2 B^0–$\overline{B^0}$ mixing

The matrix element M_{12} can be obtained from the exchange of all virtual intermediate states, with the dominant contribution due to the exchange of a top quark (see Fig. 2). On the other hand, Γ_{12} is given by the exchange of all possible real intermediate states (i.e. states that both the B^0 and the $\overline{B^0}$ mesons can decay into). The extraction of M_{12} and Γ_{12} is summarized here, for detailed discussions see for example [9], [10].

At the quark level the amplitude $\mathcal{M}$ for $\overline{b}s \to \overline{s}b$ or $\overline{b}d \to \overline{d}b$ transitions is given by

$$\mathcal{M} = -g^4 |V_{tb}V_{tq}^*|^2 m_t^2 \frac{i\pi^2}{M_W^4} A\left(\frac{m_t^2}{M_W^2}\right), \tag{1.34}$$

where $q = d$ (for the B_d^0 system) or $q = s$ (for the B_s^0 system). g is the weak

coupling constant, m_t is the top quark mass and M_W is the W mass. $A\left(\frac{m_t^2}{M_W^2}\right)$ is the Inami-Lim function [11], given by

$$A(x) = [1 - \frac{3}{4}\frac{x+x^2}{(1-x)^2} - \frac{3}{2}\frac{x^2}{(1-x)^2}\ln x] \tag{1.35}$$

It follows from the integral over loop momenta.

The corresponding effective Hamiltonian is

$$H_{eff} = \frac{G_F^2}{2\pi^2}|V_{tb}V_{tq}^*|^2 m_t^2 A\left(\frac{m_t^2}{M_W^2}\right)(\overline{q}_L\gamma_\mu b_L)(\overline{q}_L\gamma^\mu b_L), \tag{1.36}$$

where the Fermi constant G_F is given (at lowest order) by

$$\frac{G_F}{\sqrt{2}} = \frac{g^2}{8M_W^2}. \tag{1.37}$$

Note that the coefficient of the 4-quark operator in H_{eff} is a purely short distance quantity, which can be calculated perturbatively. M_{12} is given by the matrix element $M_{12} = \frac{1}{m_{B_q}} < B^0|H_{eff}|\overline{B^0} >$, yielding

$$M_{12}^q \approx \frac{G_F^2}{12\pi^2} f_{B_q}^2 \mathcal{B}_{B_q} m_{B_q} \eta_{QCD} |V_{tb}V_{tq}^*|^2 A\left(\frac{m_t^2}{M_W^2}\right) m_t^2. \tag{1.38}$$

f_{B_q} is the B_q decay constant and $\mathcal{B}_q$ is the 'bag parameter' which parameterizes the non-perturbative matrix element of the 4-quark operator. η_{QCD} is the perturbative QCD correction.

A perturbative estimate of Γ_{12}^q gives

$$\Gamma_{12}^q \approx \frac{G_F^2}{8\pi} f_{B_q}^2 \mathcal{B}_q m_{B_q} m_b^2 \eta_{QCD} \left[(V_{tb}V_{tq}^*)^2 + V_{cb}V_{cq}^* V_{tb}V_{tq}^* O\left(\frac{m_c^2}{m_b^2}\right) + (V_{cb}V_{cq}^*)^2 O\left(\frac{m_c^4}{m_b^4}\right)\right]. \tag{1.39}$$

Note that $|M_{12}^q| \gg |\Gamma_{12}^q|$ because of the top mass dependence of M_{12}. Together with Eq.(1.31) this implies

$$\Delta m \approx 2|M_{12}| \tag{1.40}$$

and therefore

$$\Delta m \approx \frac{G_F^2}{6\pi^2} f_{B_q}^2 \mathcal{B}_{B_q} m_{B_q} \eta_{QCD} |V_{tb} V_{tq}^*|^2 A \left(\frac{m_t^2}{M_W^2}\right) m_t^2. \tag{1.41}$$

The ratio of mass and decay width differences becomes

$$\frac{\Delta\Gamma}{\Delta m} \sim O\left(\frac{m_b^2}{m_t^2}\right). \tag{1.42}$$

Measurements of Δm_s and Δm_d in principle allow the extraction of the magnitudes of the CKM matrix elements V_{td} and V_{ts}. Especially interesting, as mentioned earlier, is the matrix element V_{td}, which is related to $\tilde{\rho}$ and $\tilde{\eta}$, i.e., $V_{td}^2 = (A\lambda^3(1-\tilde{\rho}-i\tilde{\eta}))^2$. Δm_d is already measured with very good precision, and more measurements are expected in the future. The current world average is [12]

$$\Delta m_d = 0.487 \pm 0.014 \text{ ps}^{-1}. \tag{1.43}$$

The main problem for the extraction of V_{td} is the large uncertainty in the determination of $f_{B_q}\sqrt{\mathcal{B}_{B_q}}$. Currently, lattice QCD calculations [13] give a $20-30\%$ uncertainty. It is advantageous to extract the ratio V_{ts}/V_{td} from the ratio $\frac{\Delta m_s}{\Delta m_d}$, as many theoretical uncertainies are common to B_s^0 and B_d^0 mixing, and cancel [13]:

$$\frac{\Delta m_d}{\Delta m_s} = (1.16 \pm 0.05)^2 \left|\frac{V_{td}}{V_{ts}}\right|^2. \tag{1.44}$$

To obtain a rough estimate of the value of Δm_s compared to Δm_d in the SM consider the CKM matrix in powers of λ,

$$V_{CKM} \approx \begin{pmatrix} 1 & \lambda & \lambda^3 \\ \lambda & 1 & \lambda^2 \\ \lambda^3 & \lambda^2 & 1 \end{pmatrix}, \tag{1.45}$$

so that

$$\Delta m_s \propto m_t^2 |V_{tb} V_{ts}^*|^2 \sim (\lambda^2 m_t)^2 \text{ and} \tag{1.46}$$

$$\Delta m_d \propto m_t^2 |V_{tb} V_{td}^*|^2 \sim (\lambda^3 m_t)^2. \tag{1.47}$$

This illustrates that (in the SM) one expects Δm_s to be about a factor $\frac{1}{\lambda^2} \sim 20$ larger than Δm_d. Such a large mixing frequency is experimentally very hard to measure and requires, among other things, the ability to resolve fast oscillations in the detector.

In order to extract an expression that relates to measurable quantities, one calculates the time dependence of the mass eigenstates. The exponential evolution of $P_H(t)$ and $P_L(t)$ is given by

$$|P_H(t)>= e^{-i\mu_H t}|P_H>= e^{-im_H t}e^{-\Gamma_H t/2}|P_H> \tag{1.48}$$

$$|P_L(t)>= e^{-i\mu_L t}|P_L>= e^{-im_L t}e^{-\Gamma_L t/2}|P_L>, \tag{1.49}$$

where t refers to the time measured in the rest frame of the decaying particle. Using Eqs. (1.21) and (1.22) the time dependence of the flavor eigenstates follows:

$$|P^0(t)>= g_+(t)|P^0> + \frac{q}{p} g_-(t)|\overline{P^0}> \tag{1.50}$$

$$|\overline{P^0}(t)>= \frac{q}{p} g_-(t)|P^0> + g_+(t)|\overline{P^0}>, \tag{1.51}$$

where

$$g_\pm(t) \equiv \frac{1}{2}(e^{-i\mu_H t} \pm e^{-i\mu_L t}). \tag{1.52}$$

The above equation describes how a beam produced at $t = 0$ as a flavor eigenstate $|P^0>$ evolves in time. The two flavor eigenstates oscillate into each other with time dependent probabilities proportional to $|g_\pm(t)|^2$. Note that for the B_s^0 system the ratio $\frac{q}{p}$, as given in Eq.(1.33), is 1 to a very good approximation, because $\arg(V_{tb}V_{ts}^\star) \approx 0$ (i.e. mixing induced CP asymmetries vanish). The deviation of $\frac{q}{p}$ from 1 signifies the amount of mixing induced CP violation.

One defines

$$\Gamma \equiv \frac{\Gamma_H + \Gamma_L}{2}, \tag{1.53}$$

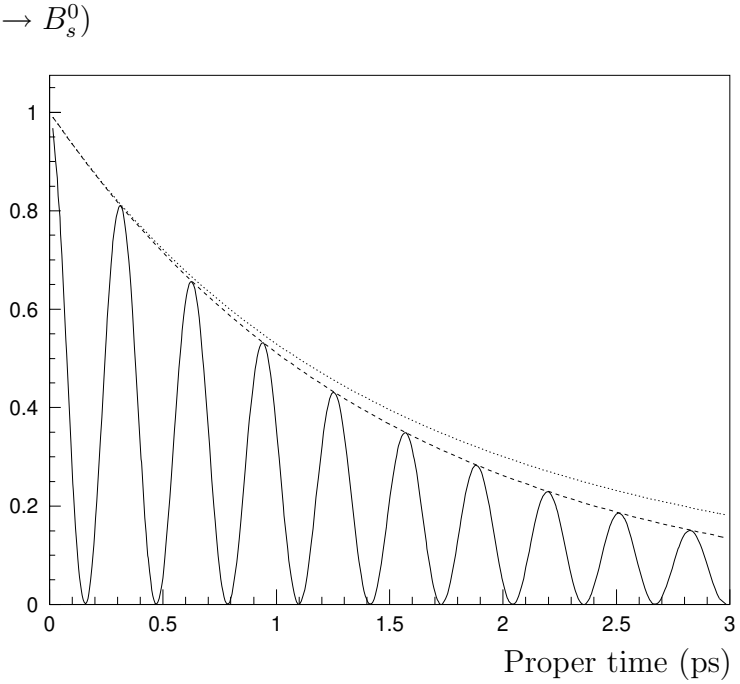


Figure 3: The probability for a B_s^0 decaying as a B_s^0, as a function of the proper time. The plots show $|g_\pm(t)|^2$ for different values of $\Delta\Gamma$ and Δm_s: the solid line was obtained with the values $\Delta m_s = 20$ ps^{-1} and $\Delta\Gamma = 0$ ps^{-1}, and the dashed line for the values $\Delta m_s = 0$ ps^{-1} and $\Delta\Gamma = 0.075$ ps^{-1}. The dotted line was calculated for $\Delta m_s = 0$ ps^{-1} and a very large assumed value of $\Delta\Gamma = 10\times$ 0.075 ps^{-1}.

so that

$$|g_\pm(t)|^2 = \frac{e^{-\Gamma t}}{2}[\cosh\frac{\Delta\Gamma}{2}t \pm \cos\Delta m t]. \tag{1.54}$$

This time dependent quantity can be determined experimentally. Fig 3 shows this mixing probability as a function of proper time t for the case of B_s^0–$\overline{B_s^0}$ mixing only. All experimental effects, for example detector effects and background composition, which rapidly wash out the mixing amplitude, have been left out. The plots show $|g_\pm(t)|^2$ for different values of $\Delta\Gamma$ and Δm_s: the solid line was obtained with the values $\Delta m_s = 20$ ps^{-1} and $\Delta\Gamma = 0$ ps^{-1}, and the dashed line for the values $\Delta m_s = 0$ ps^{-1} and $\Delta\Gamma = 0.075$ ps^{-1} [4]. The dotted line was calculated for $\Delta m_s = 0$ ps^{-1} and a very large assumed value of $\Delta\Gamma = 10\times$ 0.075 ps^{-1}. As expected, $\Delta\Gamma$

only has a marginal effect on the exponential decay. However, $\Delta\Gamma$ plays a role in the search for new physics in B_s^0–$\overline{B_s^0}$ mixing. For experimental purposes, one uses

$$|g_{\pm}(t)|^2 = \frac{e^{-\Gamma t}}{2}[1 \pm \cos \Delta m_s t]. \tag{1.55}$$

To conclude, it is possible to extract the ratio V_{ts}/V_{td} from $\frac{\Delta m_s}{\Delta m_d}$ with relatively low theoretical uncertainties of about 5%. B_d^0–$\overline{B_d^0}$ mixing has been measured with good precision, but B_s^0–$\overline{B_s^0}$ mixing is expected to be an order of magnitude faster and experimentally hard to resolve. The B_s^0–$\overline{B_s^0}$ mixing rate will be measured as a function of proper time.

1.5 CKM matrix elements from Tree Level Decays and Unitarity

In this section current results from direct measurements of the CKM matrix elements and other constraints are reviewed, in decreasing order of precision. The resulting upper and lower limits on the CKM elements will be discussed [4]. The accuracy of the measurements varies greatly. For example, V_{ud} is determined up to 0.1%, whereas measurements sensitive to V_{tb} suffer from large errors of about 15%.

1.5.1 V_{ud}, V_{us} and V_{cb}

The errors on V_{ud}, V_{us} and V_{cb} are the smallest because large data samples are available for decays into light hadrons. Concerning V_{cb}, the heavy quark effective theory (HQET) provides a nearly model-independent treatment of B semileptonic decays to charmed mesons resulting in small theoretical errors.

$|V_{ud}|$ has been determined from the comparison between different nuclear beta decays. Small theoretical uncertainties in this measurement arise from radiative

corrections. $|V_{ud}|$ can also be determined from neutron decays, its value depending on the free neutron lifetime and other parameters, so that the errors on those measurements are larger. The current world average for $|V_{ud}|$ is [4]

$$|V_{ud}| = 0.9735 \pm 0.0008. \tag{1.56}$$

$|V_{us}|$ is known from the semileptonic decays of strange particles, for example from the $\overline{K}^0 \to \pi^+ e^- \overline{\nu}_e$ decay rate. Theoretical uncertainties come from the calculation of the $K \to \pi$ form factor. The value is [14]

$$|V_{us}| = \lambda = 0.2205 \pm 0.0018. \tag{1.57}$$

Semileptonic hyperon decays can also be used to extract $|V_{us}|$ and give consistent results.

Inclusive semileptonic $B \to X_c l\nu$ decays can be calculated in HQET. This allows $|V_{cb}|$ to be extracted from the semileptonic branching ratio. In the exclusive method (for example $B^0 \to D^* l\nu$) the differential decay rate can be determined from HQET, and depends on the D^* form factor $F_D(q)$. Combining all LEP data on exclusive and inclusive decays the average value of $|V_{cb}|$ is [15]

$$|V_{cb}| = 0.0402 \pm 0.0019. \tag{1.58}$$

The new CLEO inclusive result is somewhat different: $|V_{cb}| = 0.0462 \pm 0.0032$ [16].

1.5.2 V_{cd} and V_{cs}

The direct determinations of V_{cd} and V_{cs} are rather poor and are affected by considerable theoretical uncertainties. Information can be extracted from deep inelastic neutrino scattering in reactions like $\nu_\mu d \to \mu^- c$ and $\nu_\mu s \to \mu^- c$ (CDHS [17],

CCFR [18] collaborations). The resulting value of V_{cs} depends on assumptions about the strange-quark density in the parton sea. Average values are given by [4]

$$|V_{cd}| = 0.224 \pm 0.016, \tag{1.59}$$

$$|V_{cs}| \geq 0.59. \tag{1.60}$$

Another source of information for V_{cs} is the exclusive semileptonic decay $D \to \overline{K} e^+ \nu_e$. The procedure is similar to the one mentioned for V_{cb}: one compares the experimental value for the width of the decay with the following expression that follows from the standard weak interaction amplitude:

$$\Gamma(D \to \overline{K} e^+ \nu_e) = |f_+^D(0)|^2 |V_{cs}|^2 (1.54 \times 10^{11} s^{-1}). \tag{1.61}$$

Here $f_+^D(q^2)$ with $q = p_D - p_K$ is the form factor relevant to the decay. The result depends on the calculation of $f_+^D(0)$, which is the main source of error. The result is $f_+^D(0) = 0.7 \pm 0.1$ [19], [20], which yields, together with the experimental information,

$$|V_{cs}| = 1.04 \pm 0.16. \tag{1.62}$$

1.5.3 $|V_{tb}|$ and $|V_{ub}|$

A first rough measurement of $|V_{tb}|$ from $t \to bW$ decays at CDF [21] gives $V_{tb} = 0.99 \pm 0.15$. There are no tree level results for the other top quark couplings V_{td} and V_{ts}. Because of very small branching ratios for such tree level top quark decays it is presently only possible to determine these elements in loop induced processes like B^0 mixing (as discussed earlier), or radiative $b \to s\gamma$ decays. However, there are experiments [22] which intend to measure V_{td} and V_{ts} from extremely rare $K^+ \to \pi^+ \nu\overline{\nu}$ and $K_L^0 \to \pi^0 \nu\overline{\nu}$ decays. In the SM these decays are mediated by box

and electroweak penguin diagrams. The ratio $|V_{td}|/|V_{ts}|$ is directly related to the length R_t of the unitarity triangle.

Measurements of $|V_{ub}|$ have been obtained from the end-point of the lepton energy spectrum in $B \to X_u e \nu_e$ decays. The theoretical model dependence can be reduced by using exclusive decays, but the final result still contains a large error. Usually the ratio [23]

$$|V_{ub}|/|V_{cb}| = 0.09 \pm 0.025 \tag{1.63}$$

is quoted. Future measurements will also focus on the hadronic and charged lepton-neutrino invariant mass spectra.

The measurement of $|V_{ub}|/|V_{cb}|$ corresponds to one of the sides of the unitarity triangle:

$$R_b = (1 - \frac{\lambda^2}{2})\frac{1}{\lambda}|V_{ub}|/|V_{cb}|. \tag{1.64}$$

1.5.4 Unitarity Constraints

The assumption of 3×3 unitarity of the CKM matrix allows one to further constrain the matrix elements. One starts with the 90% confidence level values for the following elements:

$$0.219 \leq V_{us} \leq 0.226 \tag{1.65}$$

$$0.038 \leq V_{cb} \leq 0.042 \tag{1.66}$$

$$0.002 \leq V_{ub} \leq 0.005. \tag{1.67}$$

Using the unitarity relations

$$|V_{ud}| = \sqrt{1 - |V_{us}|^2 - |V_{ub}|^2}, \tag{1.68}$$

$$|V_{tb}| = \sqrt{1 - |V_{cb}|^2 - |V_{ub}|^2} \tag{1.69}$$

one then obtains

$$0.9742 \leq |V_{ud}| \leq 0.9757 \tag{1.70}$$

$$0.9990 \leq |V_{tb}| \leq 0.9993. \tag{1.71}$$

Interestingly, $|V_{tb}|$ is the most tightly constrained matrix element, even though its direct measurement has a 15% uncertainty. A discrepancy between the highly constrained value of $|V_{tb}|$ and future direct measurements would be a signal for New Physics.

Using the unitarity conditions on the remaining 4 elements one arrives at the following 90% confidence bounds [4]:

$$|V| = \begin{pmatrix} 0.9742 - 0.9757 & 0.219 - 0.226 & 0.002 - 0.005 \\ 0.219 - 0.225 & 0.9734 - 0.9749 & 0.037 - 0.043 \\ 0.0034 - 0.014 & 0.035 - 0.043 & 0.9990 - 0.9993 \end{pmatrix} \tag{1.72}$$

To summarize, the SM parameters A and λ have been measured in tree level decays. The associated errors are at the few per cent level. On the other hand, constraints on the parameters ρ and η from tree level decays are plagued by much larger errors.

Not mentioned so far are measurements of the amount of CP violation, i.e. the magnitude of the complex phase η .

1.6 CP Violation and New Physics in Mixing

CP stands for the combination of discrete operations of charge conjugation C, transforming particles into antiparticles, and parity transformation P, taking $\vec{r}$ to $-\vec{r}$. Note that the orientation of the particle spin relative to its direction of flight

is switched. The combination CPT (T symmetry standing for time reversal) is an exact symmetry in any local quantum field theory.

The expression "CP symmetry breaking", or "CP violation", relates to a difference in the dynamical behavior of particles and antiparticles. CP violation is an expected consequence of the SM with three quark generations, as discussed in section 1.3.1. In fact, CP is broken in any theory that has complex coupling constants in the Lagrangian which cannot be removed by any choice of phase redefinition of the fields in the theory. An example is the introduction of complex couplings in the Yukawa terms, as in the SM electroweak theory. CP violation has been observed in the neutral kaon system [24], and in the decays of B_d^0 mesons [25], [26]. In the next few years further tests are expected at the B-factories, and at the Tevatron.

The possible manifestations of CP violation can be classified in the following way:

- CP violation in decay, which occurs in both charged and neutral decays, when the amplitude for a decay and its CP conjugate process have different magnitudes;

- CP violation in mixing, which occurs when the two neutral mass eigenstates cannot be chosen to be CP eigenstates;

- CP violation in the interference between decays with and without mixing, which occurs in decays into final states that are common to B^0 and $\overline{B}^0$. It can occur in combination with the other two types. CP violation in B_d^0 decays allows the determination of $\sin 2\beta$ [27], one of the angles of the unitarity triangle.

CP violation is required for an explanation of the observed asymmetry between the number of baryons and anti-baryons in the universe (baryogenesis). In a large

class of models CP violation and baryogenesis are electroweak in origin. However, the amount of CP violation in the SM is not sufficient to obtain a large enough baryon asymmetry.

1.6.1 New Physics in the B_s^0 system

The B_s^0 system is a prime candidate for the discovery of non-SM physics. The B_s^0–$\overline{B_s^0}$ mixing amplitude is a CKM-suppressed fourth order weak interaction process. Therefore it is very sensitive to the intervention of New Physics (NP). Moreover, in the SM the mixing induced CP asymmetries in B_s^0 decays almost vanish because the mixing phase $\phi^{SM} = \arg(\mathrm{M_{12}}) = \arg(\mathrm{V_{tb}V_{ts}^*}) \approx 0$. This is different than for the B_d^0 system, where the SM predicts a large CP violating effect. Any sizable mixing induced CP violating effects in B_s^0 decays would therefore have to come from non-SM processes, which makes the B_s^0 system especially interesting. For example, new flavor-changing interactions in supersymmetric models can give large new contributions to the mixing amplitude with $O(1)$ CP-violating phases that are unrelated to the phase of the CKM matrix (see, e.g., Refs. [28], [29]).

NP effects on the mixing amplitude M_{12} can be parametrized as

$$M_{12} = |M_{12}|e^{i\phi} = M_{12}^{SM}(1 + ae^{i\Theta}). \tag{1.73}$$

Note that $\phi \approx \phi^{NP}$, as explained above. The parameters a and Θ denote the relative magnitude and phase of the NP contribution, i.e. $a = |M_{12}^{NP}/M_{12}^{SM}|$ and $\Theta = \arg(\mathrm{M_{12}^{NP}/M_{12}^{SM}})$. So far experimental constraints do not exclude $a > 1$ and Θ can have any value.

Three of the many models that can give NP contributions to B_s^0–$\overline{B_s^0}$ mixing are briefly mentioned here.

- SUSY models where squark-gluino box diagrams contribute to B_s^0–$\overline{B_s^0}$ mixing.

$a > 1$ is allowed and Θ is arbitrary [29].

- Left-Right-Symmetric models, with a $SU(2)_L \times SU(2)_R \times U(1) \times CP$ symmetry. As an example, the particular model described in [30] predicts $a \approx 1$ and arbitrary Θ. The measured Δm_s value could be compatible with the SM, and NP could only be confirmed by a measurement of the CP violating phase.

- Fourth generation models with large contributions from t' box diagrams allow $a > 1$ and arbitrary Θ [31], [32].

Note that even if Δm_s is found to be consistent with SM predictions this would not rule out large NP contributions to the mixing amplitude. However, a measurement of CP violation in mixing would provide an unambiguous signal for NP.

A promising testing ground for NP induced CP violation in B_s^0–$\overline{B_s^0}$ mixing are decays into CP eigenstates f in the $b \to c\overline{c}s$ chain, for example the hadronic decay $B_s^0 \to D_s^{(*)^+} D_s^{(*)^-}$ (mostly CP-even [33]). The observable is the time dependent CP asymmetry $a_{CP}(t)$

$$a_{CP}(t) \equiv \frac{\Gamma(\overline{B_s}(t) \to f) - \Gamma(B_s(t) \to f)}{\Gamma(\overline{B_s}(t) \to f) + \Gamma(B_s(t) \to f)}, \tag{1.74}$$

which can be calculated to be [34]

$$a_{CP}(t) = \frac{\sin\phi \sin(\Delta m t)}{\cosh(\Delta\Gamma t/2) - |\cos\phi| \sinh(|\Delta\Gamma| t/2)}. \tag{1.75}$$

Another example is the decay $B_s^0 \to \Psi\Phi$, which is an admixture of CP-even and CP-odd components. It is experimentally easier to detect, but requires an angular analysis to disentangle the components. Denoting the linear polarization amplitudes with $A_0(t)$, $A_\parallel(t)$ and $A_\perp$, the decay rates are given by

$$|A_0(t)|^2 = |A_0(0)|^2 e^{-\Gamma t} \left[\cosh\frac{\Delta\Gamma t}{2} - |\cos\phi| \sinh\frac{|\Delta\Gamma| t}{2} + \sin\phi \sin(\Delta m t)\right], \tag{1.76}$$

$$|A_{\parallel}(t)|^2 = |A_{\parallel}(0)|^2 e^{-\Gamma t}\left[\cosh\frac{\Delta\Gamma t}{2} - |\cos\phi|\sinh\frac{|\Delta\Gamma|t}{2} + \sin\phi\sin(\Delta m t)\right], \quad (1.77)$$

$$|A_{\perp}(t)|^2 = |A_{\perp}(0)|^2 e^{-\Gamma t}\left[\cosh\frac{\Delta\Gamma t}{2} + |\cos\phi|\sinh\frac{|\Delta\Gamma|t}{2} - \sin\phi\sin(\Delta m t)\right]. \quad (1.78)$$

CP asymmetries or a sizeable $\sin\phi$ from a fit of the decay rates $|A_{0,\parallel,\perp}(t)|^2$ would provide an unambiguous signal of NP contributions to B_s^0–$\overline{B_s^0}$ mixing. Δm_s needs to be resolvable to measure ϕ.

A possible test for the presence of NP independently of Δm_s is a precise measurement of $\Delta\Gamma$. NP contributions in most cases reduce $\Delta\Gamma$ [35]. This can be understood intuitively by considering that CP violation would allow both mass eigenstates to decay into CP even final states, thereby lowering the SM predicted width difference between them. Another possibility would be a fit to the time dependent decay width into CP even eigenstates, for example $B_s^0 \to D_s^{(*)^+} D_s^{(*)^-}$. In the SM $\Gamma(t) \propto e^{-i\Gamma_S t}$, whereas in the presence of new sources of CP violation $\Gamma(t) \propto e^{-i\Gamma_S t} + e^{-i\Gamma_L t}$, i.e. the decay width could only be fitted with two exponentials.

CHAPTER 2

Experimental Apparatus: the SLC and SLD

The Stanford Linear Collider (SLC) at SLAC is dedicated to the study of Z^0 boson physics. Electron- and positron bunches are sent down a two mile long linear accelerator before they are split and kicked into two arcs, forcing the beams to collide. The beams are accelerated to about 46 GeV each, resulting in a head-on collision at the Z-boson energy.

Construction of the SLC began in 1983 and concluded in 1987. The first physics results from the SLC were reported by the Marc II Collaboration in 1989 and 1990. The physics program of the SLC Large Detector (SLD) began with an engineering run in 1991 and ended in the spring of 1998. The research program includes electroweak and QCD studies, as well as heavy quark physics, e.g., B_s^0–$\overline{B_s^0}$ mixing. Physics studies take advantage of SLC's polarized electron beams and the extremely small and stable beam spot size. Many studies, including the B_s^0–$\overline{B_s^0}$ mixing analysis, also rely on the very high resolution of the SLD vertex detector.

2.1 The SLC

Figure 4 shows an outline of the SLC with its two mile long linear accelerator (linac) for both electrons and positrons, and the adjoining arcs and collision area.

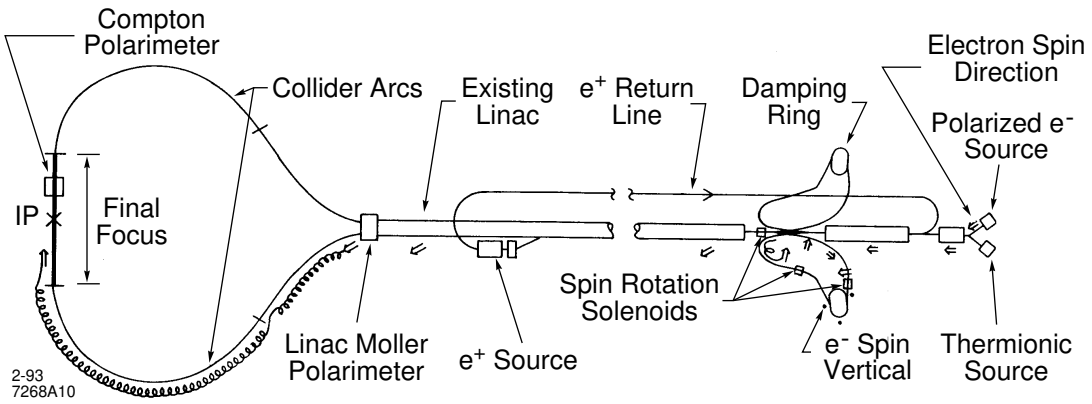


Figure 4: Layout of the SLC.

The SLC operates with a very low repetition rate, at a 120 Hz machine cycle, because of klystron power limits. In the final focus sections, both beams are focussed by superconducting quadrupole magnets before colliding head-on in a micron-sized interaction point (IP) inside the SLD. After passing through the IP each spent bunch is absorbed into a beam dump behind the collision area.

2.1.1 Beam Transport

Each accelerator cycle begins at the electron source with the production of two bunches containing about 5×10^{10} electrons each. These bunches, which are accumulated with a 178 MHz RF field, are accelerated to 50MeV, enter the linac and are further accelerated. Then they are fed into the North Damping Ring, where their emittance is reduced by synchroton radiation. This is essential to get small beam spots. The lost energy is restored by an accelerating cavity. The electrons are stored for one machine cycle (=8.3 ms), then fed back into the linac. About 2/3 down the linac the first electron bunch is peeled off and diverted into a target in order to produce positrons, while the other bunch travels to the IP. Positrons are produced by pair-production in the positron target, packed into bunches with

the same density as electron bunches, sent via a return line back to the beginning of the accelerator and into the South Damping ring. As they have a larger energy spread than electrons, a positron bunch stays in the damping ring for two machine cycles (=16.6ms) and awaits the next accelerator cycle.

2.1.2 Polarized Electrons

The SLC electron gun produces polarized electrons. A strained-lattice GaAs photocathode is exposed to circularly polarized laser light from a Nd:YAG-pumped Ti sapphire laser. It excites electron transitions into longitudinally-polarized states in the conduction band of the photocathode. Figure 5 shows the electron transitions between the relevant energy levels.

The electron bunch is then accelerated by a 30 kV electric field before it is injected into the linac. The 1992 physics run made use of a bulk GaAs photocathode whose theoretical maximum polarization of 50% is due to the degeneracy in the valence band as shown in Figure 5. Approximately 22% average polarization was achieved. For the 1993 physics run, a strained lattice photocathode consisting of a 300 nm layer of GaAs deposited on a GaAsP substrate was used. The name for this cathode derives from the difference in the lattice spacings of the two materials, which leads to a strain on the GaAs layer. This strain breaks the degeneracy in the valence band, yielding a theoretical maximum polarization of 100%. About 63% average polarization was achieved. For the 1994-5 physics run, the thickness of the epitaxial layer was further reduced to 100 nm, improving the average polarization to 77%. This type of photocathode was retained for the 1996 (1997-1998) physics run, and resulted in an average polarization of 77% (73%).

The electron polarization needs to be preserved in the Damping Ring and transported undisturbed to the IP. Unless the spin vector is vertical, precessions dilute

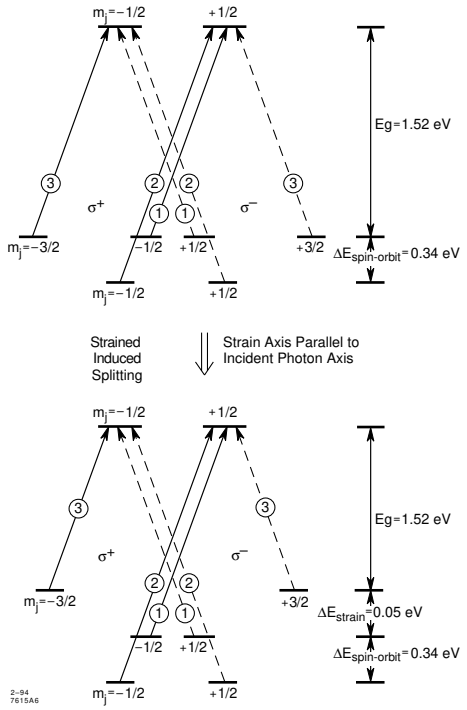


Figure 5: **Energy state diagram for (top) bulk GaAs and (bottom) the strained GaAs lattice. The encircled numbers denote the relative intensities of the transitions. The strained lattice breaks the degeneracy and allows a theoretical polarization of 100%.**

the polarization. Thus it is essential to turn the originally longitudinal spin into a perpendicular spin, along the direction of the magnetic field, to preserve the spin state. This is achieved by passing the electrons through a solenoidal field just as they enter the ring. The electrons are accelerated with their spin pointing up through the whole linac, before passing through a series of "spin bumps" in the arcs, which turn the spin back into the longitudinal direction. SLC is running with flat beams ($\epsilon_x = 10 \times \epsilon_y$), because the lower emittance in the vertical plane increases the luminosity, i.e. the reaction rates per unit cross-section, and reduces the backgrounds in the SLD detector.

2.1.3 Positrons

Electron bunches that are deflected from the linac at an energy of 30 GeV are used to produce (unpolarized) positrons. The electron beam is directed onto a 6 X_0 tungsten-rhenium alloy target. The relativistic electrons initially produce Bremsstrahlung ($e^- \rightarrow e^-\gamma$), followed by pair production $\gamma \rightarrow e^+e^-$. Electrons and positrons repeat the process, so that a shower develops. Positrons with energies between 2 and 20 MeV are captured in an accelerator structure imbedded in a strong solenoidal field. They are transported back to the front-end of the linac in a separate line, and used in the next acceleration cycle.

2.1.4 Acceleration and the Final Focus

Radio frequency cavities powered by klystrons provide an accelerating field of 17 MeV/m. A series of 240 klystrons power the 2 miles of cavities. The particles are accelerated to an energy of 0.9 GeV higher than their energy at the IP. There are no more RF cavities in the arcs, and the extra energy is used up as synchroton radiation, until the beam energy reaches 45.6 GeV just before the final focus region, thus setting the collision energy to the mass of the Z^0.

The repetition rate of a linear collider is much lower than that of a storage ring. It has to be compensated by an extraordinarily small beam size. The final focus (FF), an array of superconducting magnets reduces the beams to micron size. The FF consists of three superconducting quadrupole magnets, the first of which is located 1.5 meters in front of the IP. At the IP the beam spot dimension is about 1mm along the beam direction (z), and 0.8μm perpendicular to the beam direction (xy). The SLC constantly improved the luminosity since it started producing Z^0 in 1989. By 1998 SLC achieved production rate peaks of well above 200 Z^0/hour

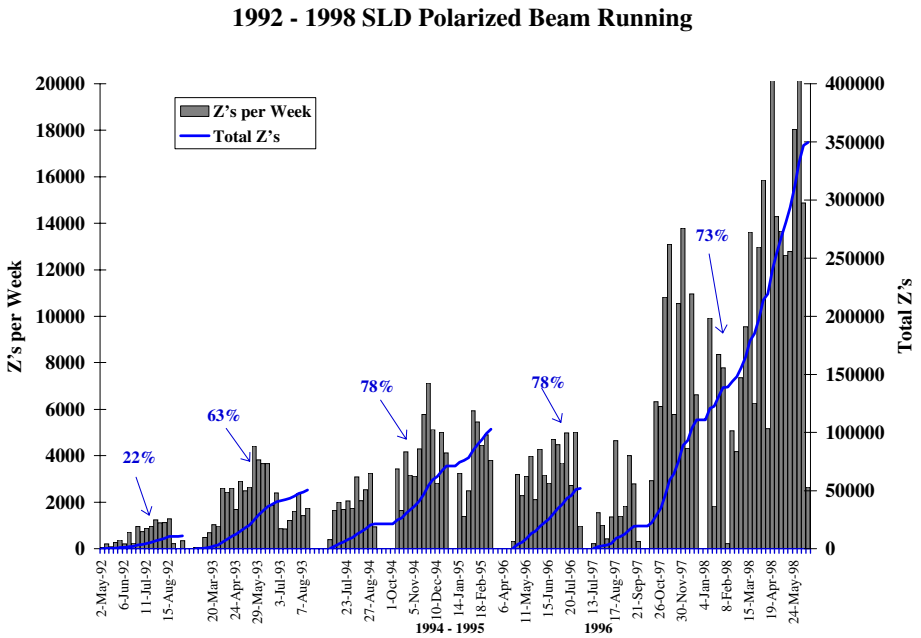


Figure 6: SLC/SLD luminosity history. In the 1997-8 data run, the luminosity improved substantially and reached 20,000 Z^0/week.

($\mathcal{L} = 2 \times 10^{30} cm^{-2} s^{-1}$). The SLC luminosity history is summarized in Figure 6.

2.1.5 The Beam Energy Measurement

After the IP, before the beams are dumped, a spectrometer is used to measure the energy of both electron and positron beams. It is called "The Wire Imaging Synchrotron Detector" (WISRD) [36], and works essentially by bending the beam and observing the angle of synchroton radiation with respect to the initial beam direction. The beam is deflected horizontally, then vertically by a bend magnet, then

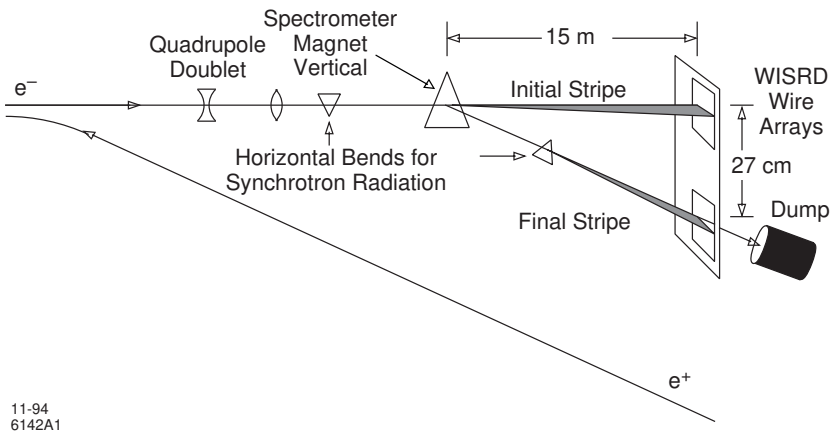


Figure 7: Schematic layout of the WISRD. The deflection of the beam is inversely proportional to its energy.

horizontally again. The horizontal bends produce two radiation swaths that can be measured using the two WISRD spectrometers. The vertical distance between the two stripes is inversely proportional to the beam energy, which can be extracted using the integrated field of the bend magnet and the distance to the detector. During the 1997-8 SLD data taking period, a Z^0 peak scan was performed which allowed the WISRD energy spectrometer to be calibrated against the precise Z^0 mass measurement obtained at LEP. The resulting luminosity weighted mean center-of-mass energy for the 1997-8 data run was determined to be $E_{cm} = 91.237 \pm 0.029$ GeV [37]. Figure 7 shows a schematic view of the WISRD.

2.1.6 The Electron Beam Polarization Measurement

Many of the SLD physics studies, for example the study of B_s^0–$\overline{B_s^0}$ mixing discussed in this thesis, depend on the knowledge of the electron beam polarization at the IP. The polarization at SLD is measured using the asymmetry in the Compton scattering cross sections for the $J_z = 3/2$ and $J_z = 1/2$ combinations of circularly

polarized light and the longitudinally polarized electron beam. The differences in scattering probability are observable as an asymmetry in the scattered electron energy spectrum. The unknown electron beam polarization P_e can be extracted from

$$A_{meas} = \frac{\sigma_{J_z=\frac{3}{2}} - \sigma_{J_z=\frac{1}{2}}}{\sigma_{J_z=\frac{3}{2}} + \sigma_{J_z=\frac{1}{2}}} = a_d P_e P_\gamma A_c(E), \tag{2.1}$$

where P_γ is the (measured) laser polarization, E is the (measured) scattered electron energy, and A_c is the (calculated) Compton scattering asymmetry. A_{meas} is the observed asymmetry, and a_d is the (calculated) analyzing power of the detector [38].

The errors on the beam polarization measurement come primarily from uncertainties in the net circular polarization of the laser and the analyzing power of the detector. The polarimeter system is shown in Figure 8. A beam of 2.33 eV photons is provided by a frequency-doubled YAG laser. The beam is transported down into the SLC to collide with the electron beam, 33 meters downstream of the SLC IP. The electron energy is then measured by a spectrometer consisting of an analyzing bend magnet and a multi-channel Cherenkov detector.

Two additional detectors, the Polarized Gamma Counter (PGC) and the Quartz Fiber Calorimeter (QFC), are used to measure the electron beam polarization. They work by observing the scattered photons rather than the electrons, as shown in Figure 8.

Both devices provide cross checks of the Compton Polarimeter result with a precision of better than 1%.

Finally, we expect the polarization of the positron beam to be zero. The measurement of the positron beam polarization with a Møller Polarimeter [39] in the SLAC End Station A yielded $P_{e^+} = (-0.02 \pm 0.07)\%$, i.e. it is consistent with zero.

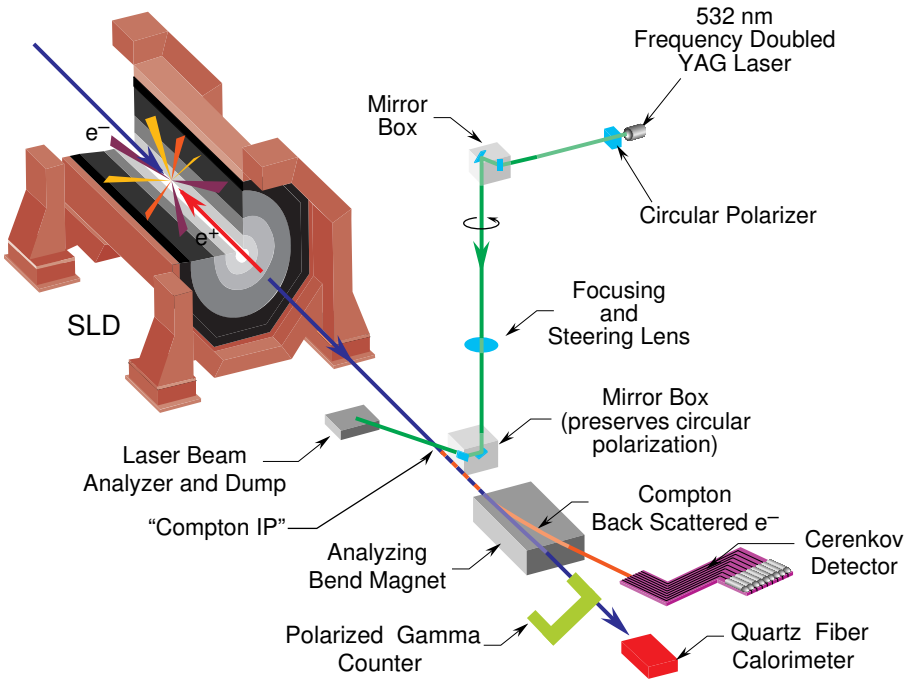


Figure 8: Schematic view of the Compton Polarimeter.

2.2 The SLC Large Detector (SLD)

The SLD measures the properties of collisions at the IP. It is a cylindrical device with a radius of about 4.5 meters and a length of 10 meters. It combines excellent tracking, calorimetry and particle identification. Two endcaps, which can be moved for access to the interior of the detector barrel, close off the detector at both ends. Figures 9 and 10 show the layout of the SLD. Each layer of the detector contains a separate sub-detector used to record different aspects of the collision. The sub-division into layers is similar for the barrel and the endcaps. The narrow beampipe and its support structures allow for a 98% solid angle coverage. The detector elements, from the IP outward, are a CCD-based vertex detector (VXD2 1992-96, VXD3 97-98) and a Drift Chamber (CDC) for extremely precise particle tracking, the Cherenkov Ring Imaging Detector (CRID) for charged particle identification, the Liquid Argon Calorimeter (LAC) for electromagnetic and hadronic calorimetry,

the solenoid magnet which produces a 0.6 Tesla axial magnetic field, in which the trajectories of charged particles bend, and the Warm Iron Calorimeter (WIC) used for muon identification. Table 7 summarizes the performance of each subsystem of the SLD detector. All subsystems have optical fiber transmission cable connections for readout.

Table 4: SLD performance.

Detector sub-system	performance
Vertex Detector (VXD3)	segmentation: $20\mu m \times 20\mu m$ Impact parameter resolution: $10\mu m$ Two track separation: $40\mu m$ Solid Angle coverage: 97%
Drift Chamber System	spatial resolution: $< 100\mu m$ Momentum Resolution: $\sigma(1/p)$ measurement limit: 1.3×10^{-3}(GeV/c)$^{-1}$ $\sigma(p)/p$ Coulomb scattering limit: 1×10^{-2} Two track separation 1mm
Particle Identification	e/π 1×10^{-3} μ/π (above 1 GeV) 2×10^{-3} K/π (up to 30 GeV) 1×10^{-3} K/p (up to 50 GeV) 1×10^{-3}
Calorimetry	EM calorimeter energy resolution $\sigma_E/E = 15\%/\sqrt{E(GeV)}$ Hadronic calorimeter energy resolution $\sigma_E/E = 60\%/\sqrt{E(GeV)}$ segmentation(EM): 33×36 mrad segmentation(Had): 66×72mrad angular resolution(EM): 5 mrad angular resolution(Had): 10 mrad Solid Angle coverage (EM): $\geq 99\%$ Solid Angle coverage (Had): 97%

The following discussion presents more details on each sub-system, especially the tracking system, which is important for the B_s^0–$\overline{B_s^0}$ mixing analysis presented here. Only the "barrel" of the detector will be described. Depending on the sub-system, the barrel covers tracking to $\cos\theta = 0.8$. The coordinate system used is

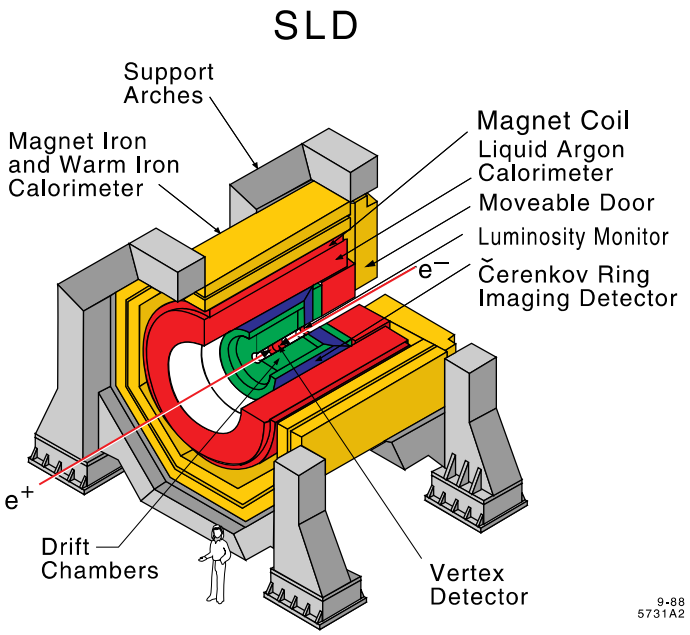


Figure 9: **The SLD in isometric view showing the cylindrical layout of the sub-detectors. The Luminosity Monitor and end-caps are not shown.**

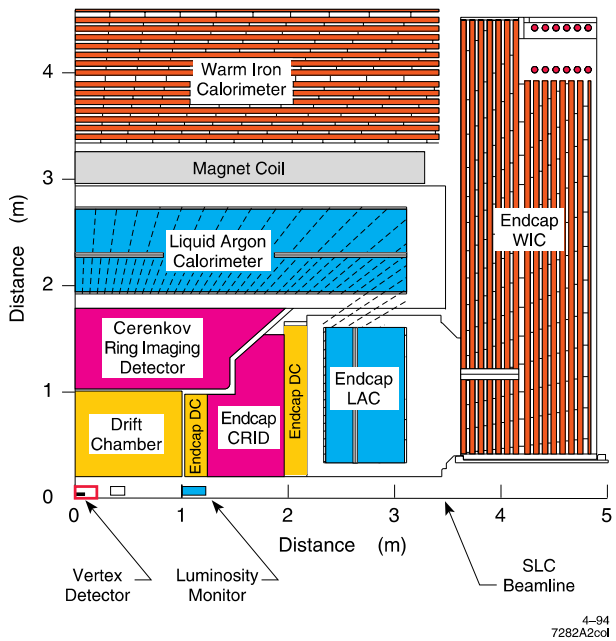


Figure 10: The SLD in quadrant view. The dimensions are indicated.

the usual right handed polar system, with z-axis pointing north in the direction of the positron beam, and x and y axes in the plane perpendicular to the beams. The vertical y axis points upwards. ϕ runs from 0 to 2π, with $\phi = 0$ pointing west, and θ runs from 0 (north) to π (south).

2.2.1 The Luminosity Monitor (LUM)

The SLD Luminosity Monitor (LUM) [40] measures the luminosity delivered by the SLC using small-angle Bhabha scattering ($e^+e^- \rightarrow e^+e^-$), whose rate is well known and which is easy to identify. The Bhabha events are measured 1 m downstream from the IP with a silicon-tungsten calorimeter module, the Luminosity Monitor/Small Angle Trigger (LMSAT), providing polar angle coverage from 28 mrad to 68 mrad. Electromagnetic showers develop in the 23 layers of alternating silicon detectors and tungsten radiator plates. The ionization signal is detected after amplification. The depth of the LUM is 21 radiation lengths. The LUM is shown in Figure 11. An energy resolution of about 3% at 50 GeV has been achieved.

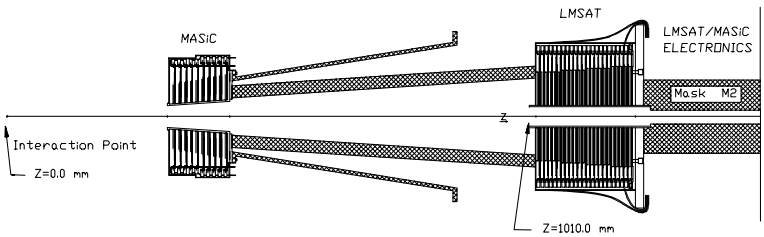


Figure 11: The SLD luminosity monitor, showing the MASiC (not used) and the LMSAT.

2.2.2 The Vertex Detectors (VXD2 and VXD3)

Vertex detectors are located just outside the beampipe, and, because of the highest track density in that area, have to have the highest segmentation.

The SLD Vertex Detector (VXD) [41] detects the ionization deposited by particles passing through it. Silicon pixel Charged Coupled Devices (CCD's) are used as the medium. SLD is the first collider experiment to use the CCD technology for a vertex detector. It allows for extremely high tracking and vertexing resolution, as well as high track reconstruction efficiency.

After a prototype "VXD1", the 120M pixel VXD2 was installed in 1992, and a much improved version, the VXD3, in December 1995. Its improved resolution and larger solid angle coverage greatly enhance the heavy flavor program. CCD-based vertex detectors need certain conditions that are met by the SLC: long intervals between beam crossings and a relatively low background environment as provided by an e^+e^- machine like the SLC. The goal is to reach optimal coverage of the IP area, as close to the IP as possible, with the largest tracking lever arm possible. The small SLC beam pipe allows detectors to be placed very close to the interaction region, at 2.8 cm, allowing for very high resolution heavy quark studies.

Only a small fraction of the SLD data was taken with the VXD2, most of it taken with the upgraded detector, VXD3. In this section the VXD3 detector is discussed, and VXD2 mentioned briefly.

VXD3 uses CCDs with an active area measuring 80×16 mm^2. Each individual pixel is $20 \times 20 \mu$m^2, for a total of 4000×800 pixels. Two CCDs are mounted on a beryllium substrate to form a ladder, as shown in Figure 12. The two CCDs overlap by $\sim$1 mm to allow their relative alignment using charged tracks. The ladders were mounted onto a series of three concentric beryllium annuli in a shingled arrangement, so that complete azimuthal coverage was obtained, with $\sim 500\mu$m overlap between adjacent ladders for alignment purposes. Views of the VXD3 ladder layout transverse and parallel to the beamline are shown in Figures 13 and 14. Figure 15 compares the ladder layout of VXD2 and VXD3.

Figure 12: Two-CCD VXD ladder.

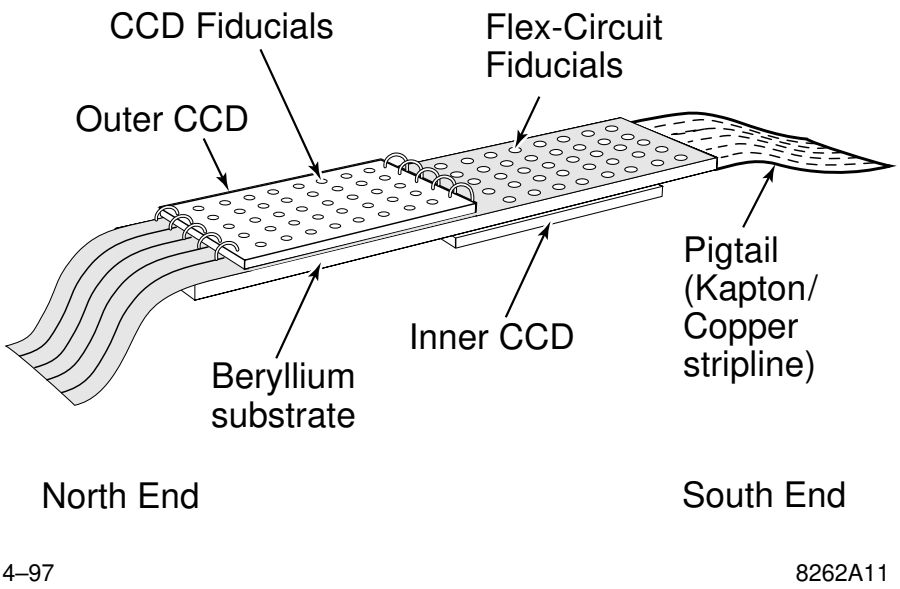


Figure 13: Schematic layout of VXD3 in the xy view.

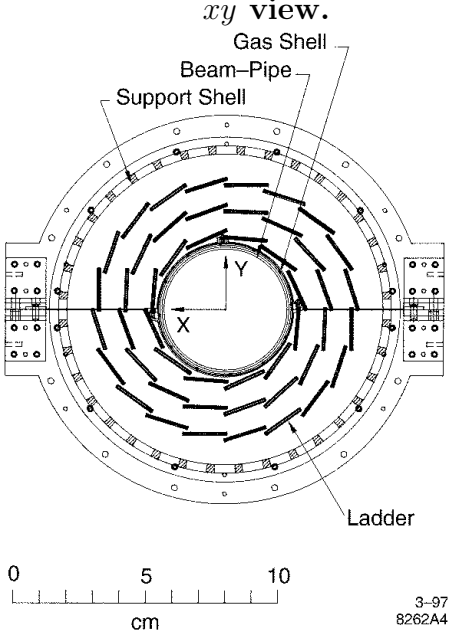
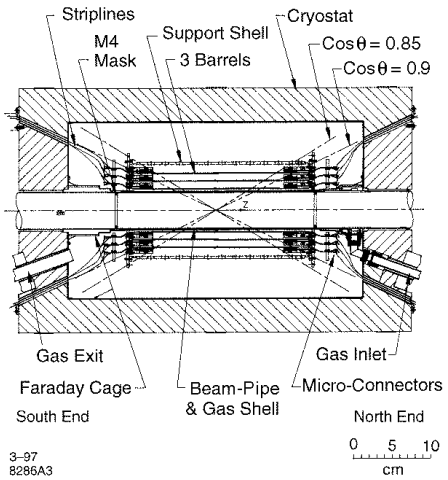


Figure 14: Schematic layout of VXD3 in the z view.



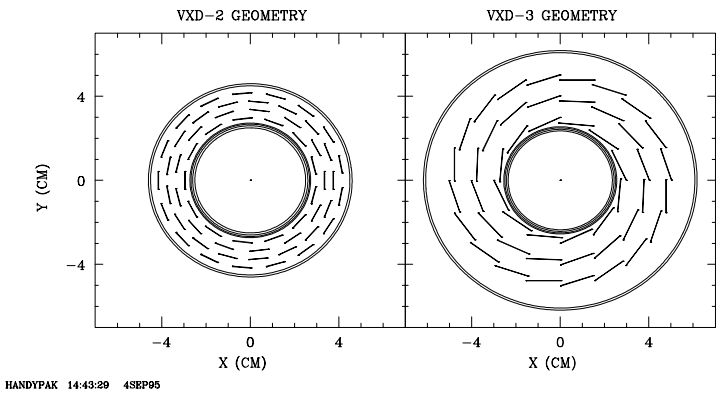


Figure 15: Comparison of VXD2 and VXD3, viewed in the xy plane. VXD3 features overlapping layers and a longer lever arm.

The layer one radius is 28 mm, fixed by the radius of the beam-pipe. The layer three radius of 48.3 mm was determined by requiring three-layer acceptance for tracks with $|\cos\theta| \leq 0.85$, to match the effective acceptance of the central drift chamber. The layer two radius of 38.2 mm provides acceptable lever arm for tracks with a missing hit on layer one or three, and provides a third space point for enhanced pattern recognition at high $|\cos\theta|$ and for self-tracking purposes. The device is 159mm long. A total of 48 ladders are used, for a total number of 307 million pixels. Each layer contributes a radiation thickness of only $0.4\%X_0$, reducing the tracking errors caused by multiple scattering.

The bias voltage applied to the CCDs was ∼10 V, producing a depletion zone of around 20 μm thickness. A minimum-ionizing particle produces about 1200 electron/hole pairs. The charge which is collected in each pixel is read out in a doubly-serial manner. Each CCD is divided into four quadrants, with one output amplifier for each located at the four corners of the CCD. The entire dectector is read out in 0.2 seconds or about 26 beam crossings at 120 Hz. Because the pixel occupancy is $< 10^{-4}$ background pileup hits is easily rejected.

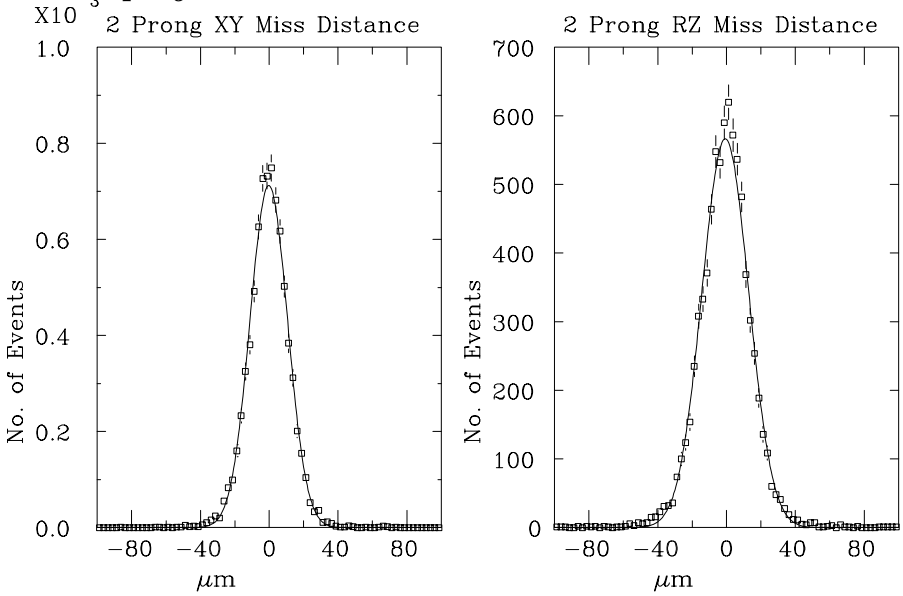
The full assembly is mounted in a nitrogen gas cryostat and maintained at

a temperature of around 185 K. It is necessary to operate the detector at low temperatures to reduce lattice defects caused by radiation damage, which can cause trapping centers to develop in the silicon. Because CCDs are read out serially, a trapping center affects not just nearby pixels, but can impact the charge transfer out of all pixels behind it in that column.

The CCD height profiles and the relative positions of the ladders on the detector were optically surveyed at room temperature to a precision of about 10 μm. In order to achive the desired precision of about 4 μm an internal alignement using tracks was necessary. Even though translations, rotations and distortions in CCD shape are corrected after the optical survey, higher order systematics due to changes after installation, cooling etc. exist. Fixing the deviations at the CCD edges to 0, 3 additional parameters are needed to correct these higher order systematics. The procedure used to determine the internal alignment parameters works as follows:

Since the true track trajectory is unknown it is necessary to identify hits on at least three CCDs associated with a track reconstructed in CDC. Only the measured momentum from the CDC track is used to constrain the track extrapolation inside the VXD3. Good quality tracks are selected with a momentum of at least 1 GeV. In general, the track is constrained to pass through two of the CCD hits and a corresponding residual with respect to the third CCD hit is measured. Groups of residuals are classified ("shingle"-residual: a track passes through the overlapping shingle region of two neighboring CCDs. A vector is fixed at the hit on one of the shingle ladders and the furthest away hit on a different layer. The residual is determined from the distance of the second shingle ladder to the fixed vector. "Doublet"-residual: calculated using a vector fixed at one of the hits through overlapping CCD's on the same ladder. "Triplet"-residual: calculated using a vector fixed at a layer 3 and layer 1 hit.)

Figure 16: Miss distance of tracks in $Z \to \mu^+\mu^-$ events, in the xy and z projections.



For each type of residual and each combination of CCDs data files are compiled containing the residuals as a function of the dip angle and angle ϕ. These data files are fitted to polynominals, to determine the parameters. After a combination of all parameters using weights for the different classes, the final alignment parameters are extracted. From the results, a consistent one-hit resolution of about 4μm in xy is obtained.

The track position resolution at the IP can be estimated from the miss distance of the two tracks in $Z \to \mu^+\mu^-$ decays, shown in Figure 16. The widths indicate resolutions of 7.7 μm in xy and 9.6 μm in z [42].

2.2.3 The Drift Chamber (CDC)

The Central Drift Chamber (CDC) [43] is the primary SLD tracking device. The CDC cylinder is 2m long, and its inner and outer radii are 20 cm and 1 m, respectively. It is immersed in a uniform solenoidal field of 0.6 T. The inner and

outer walls consist of an aluminum sheet-Hexcell fiberboard laminate, entailing $1.8\%X_0$ ($1.6\%X_0$) for the inner (outer) wall. The endplates are constructed from aluminum. The drift cells in the CDC are arranged in 10 staggered superlayers, as shown in Figure 17. The sense wires in the superlayers are either axial or have a 41 mrad stereo angle with respect to the beam axis. The wire layout – 8 sense wires, 18 guard wires, and 25 field wires – for a cell, measuring about 6 cm by 5 cm high, is displayed in Figure 18. The sense wires consist of 25 μm gold-coated tungsten, and the guard and field wires are made of 150 μm gold-coated aluminum.

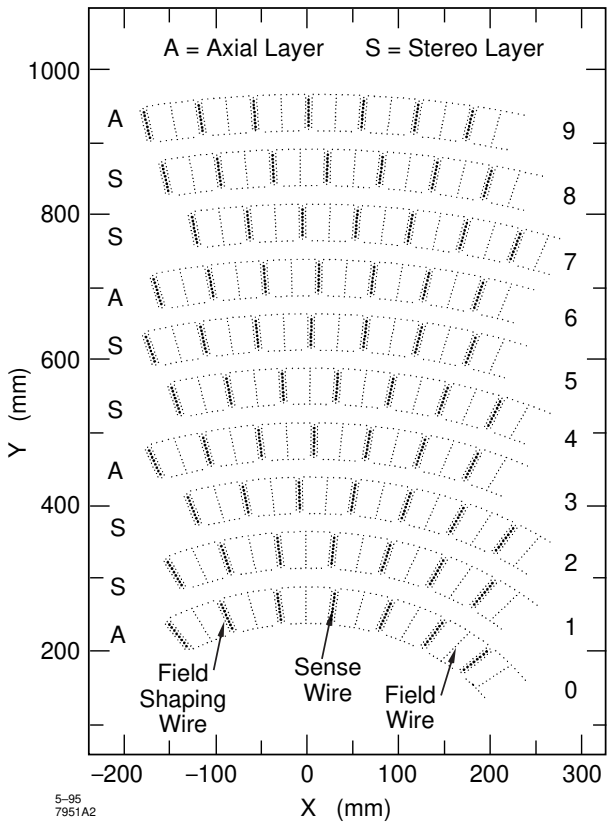


Figure 17: Endplate view of the CDC, showing the arrangement of the axial (A) and stereo (S) superlayers.

The gas in the CDC consists of 75% CO_2, 21% Ar, 4% Isobutane, and 0.2% H_2O. Properties of this mixture are low drift velocity and low diffusion, improving the

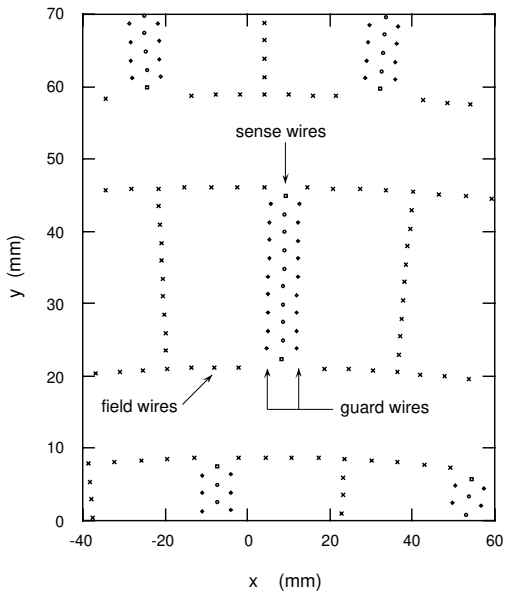


Figure 18: Schematic view of the CDC cell layout.

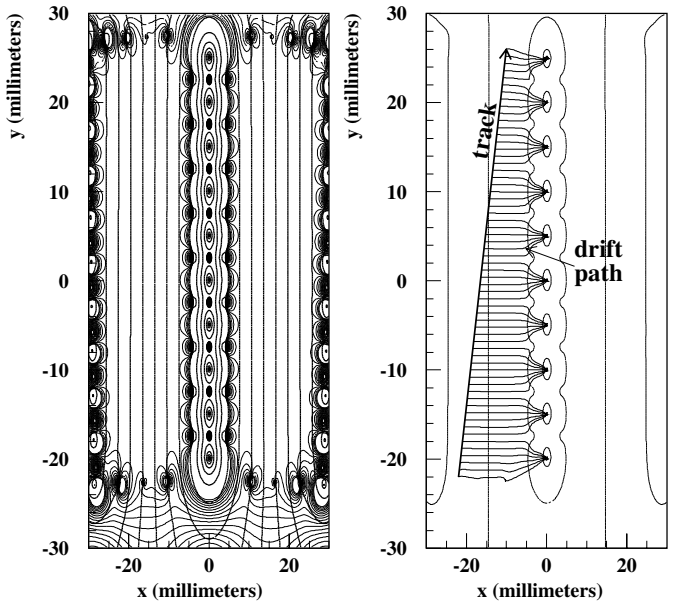


Figure 19: Field map for a CDC cell, showing (left) the field configuration for a cell and (right) the drift paths of liberated electrons caused by a charged particle traversing the cell.

spatial resolution. The addition of H_2O helps diminish the effects of wire aging [44]. The average drift field of 0.9 kV/cm leads to a drift velocity of 7.9 μm/ns in the gas.

When a charged track passes through the cells, a trail of ions is produced in the drift gas. The drift field directs the freed electrons toward the sense wires, as shown in Figure 19. The hit position is given by the wire address, drift time information, and charge division along the wire. Together with information from the other sense wires in the same cell, a "vector hit" is formed. Since the sense wires in one cell are not staggered, an ambiguity in the form of a mirror image arises. A pattern recognition program combines the vector hits from adjacent cells into track candidates. The individual wire hit information is then used to perform a detailed fit, using the more precise z information from the stereo layers, and taking into account electric and magnetic field variations, and energy loss of the track. Details of the track reconstruction algorithms can be found in [45]. After alignment [46], the momentum resolution for the CDC has been determined to be $(\sigma_{p_\perp}/p_\perp)^2 = 0.010^2 + (0.0050 p_\perp)^2$ where $p_\perp$ is the track momentum transverse to the beam axis in GeV/c. The first term denotes the uncertainty arising from multiple scattering, and the second term the measurement error [43],[47]. The momentum resolution improves to $(\sigma_{p_\perp}/p_\perp)^2 = 0.0095^2 + (0.0026 p_\perp)^2$ when the vertex detector hit information is added to the fit [48].

2.2.4 The Cherenkov Ring Imaging Detector (CRID)

The barrel portion of the Cherenkov Ring Imaging Detector (CRID) [49] allows efficient charged particle identification over a wide momentum range. The operating principle of the CRID is based on the Cherenkov effect. Charged particles traversing a dielectric medium with a velocity exceeding the velocity of light in that

medium emit a coherent wave front of Cherenkov photons. The emission angle, θ_C, to the track is given by $\cos\theta_C = \frac{1}{\beta n}$ where n denotes the index of refraction and $\beta = v/c$, with the particle velocity v.

The Cherenkov photons are imaged on an array of Time Projection Chambers (TPCs), filled with C_2H_6 gas and an admixture of tetrakis-dimethylamino-ethylene (TMAE). Photons with an energy greater than 5.4 eV ionize TMAE which releases photo-electrons into the drift gas. These photoelectrons drift to the instrumented ends of the TPCs. The barrel CRID, shown in Figure 20, utilizes two Cherenkov radiators.

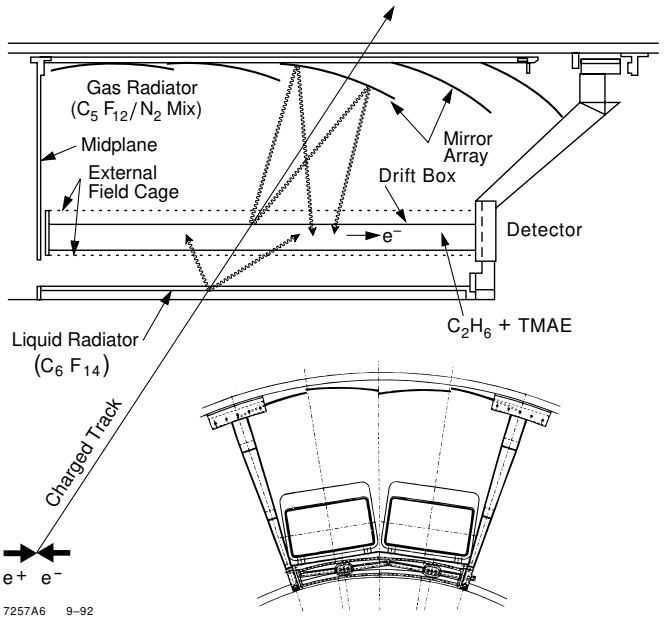
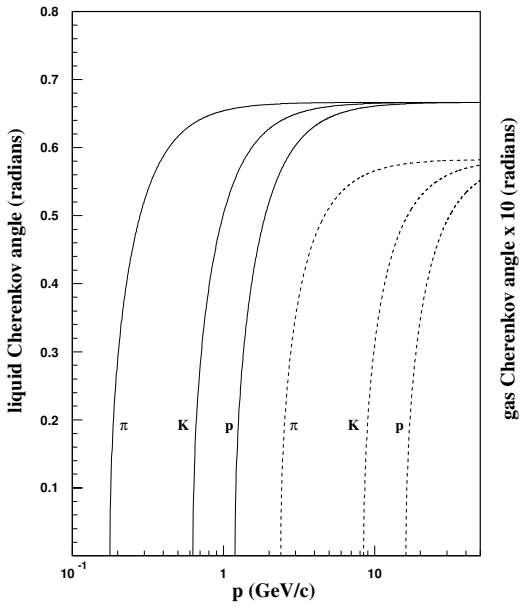


Figure 20: The barrel CRID, showing one sector in (top) axial view and (bottom) transverse view.

The liquid C_6F_{14} radiator is contained in 40 quartz-windowed trays of 1 cm liquid thickness. The gaseous radiator, consisting of 85% C_5F_{12} and 15% N_2, is 40 cm thick. Cherenkov photons from the liquid radiator pass directly into the

Figure 21: **The Cherenkov angle for the liquid (solid line) and gas radiators (dashed line) in the barrel CRID as a function of momentum for the three long lived hadronic particle species.**



TPC, whereas the photons from the gas radiator are focused by an array of 400 spherical mirrors into rings on the TPCs.

The origin of each photo-electron can be inferred from the drift time, the wire address, and the charge division on the wire. The measured diameter of the ring of reconstructed photoelectrons allows the velocity of the particle to be determined. Together with the momentum measurement provided by the CDC, the five charged particle candidates $e^{\pm}, \mu^{\pm}, \pi^{\pm}, K^{\pm}, p/\bar{p}$ can be distinguished over much of the momentum range. Figure 21 illustrates the Cherenkov angle curves for $\pi^{\pm}$, $K^{\pm}$ and $p/\bar{p}$ as a function of momentum for the two radiators.

The momentum thresholds for the 5 charged particle candidates $e^{\pm}$, $\mu^{\pm}$, $\pi^{\pm}$, $K^{\pm}$ and $p/\bar{p}$ are summarized in Table 5.

Table 5: Momentum thresholds for the barrel CRID radiators in GeV/c.

Radiator	$e^{\pm}$	$\mu^{\pm}$	$\pi^{\pm}$	$K^{\pm}$	$p/\bar{p}$
C_5F_{12}/N_2	0.009	1.811	2.393	8.463	16.084
C_6F_{14}	0.001	0.134	0.177	0.628	1.193

2.2.5 The Liquid Argon Calorimeter (LAC)

The Liquid Argon Calorimeter (LAC) [50] provides energy measurements of charged and neutral particles. The barrel extends radially from 1.8 m to 2.9 m, providing coverage between 35^o and 145^o in polar angle. The endcaps extend the coverage down to 8^o and 172^o. About 98% of the solid angle is covered by the LAC. The LAC is a lead-argon sampling calorimeter consisting of lead plates immersed in liquid argon. The lead plates induce particle showers and collect the charge caused by charged particles ionizing argon as they pass through it. The lead consists of alternating layers of grounded plates and tiles held at -2 kV, separated by plastic spacers and immersed in liquid argon. Several tiles from adjacent layers form projective towers, covering about 33 mrad in polar angle and azimuth. These are connected to amplifiers to measure the charge deposited. The LAC is segmented into four radial layers denoted EM1, EM2, HAD1, and HAD2 for their primary role: in the EM sections, the 2.0 mm thick lead plates are separated by 2.75 mm Ar gaps. The HAD sections consist of 6.0 mm thick lead plates with identical Ar gap size. The first two layers, EM1 and EM2, contain most of the energy from electromagnetic showers in 21 radiation lengths of material. The outer two layers, HAD1 and HAD2, extend the containment to hadronic showers in 2 absorptions lengths of material. The EM and HAD sections combined contain 49 radiation lengths and 2.8 absorptions lengths of material. Figure 22 shows a schematic view

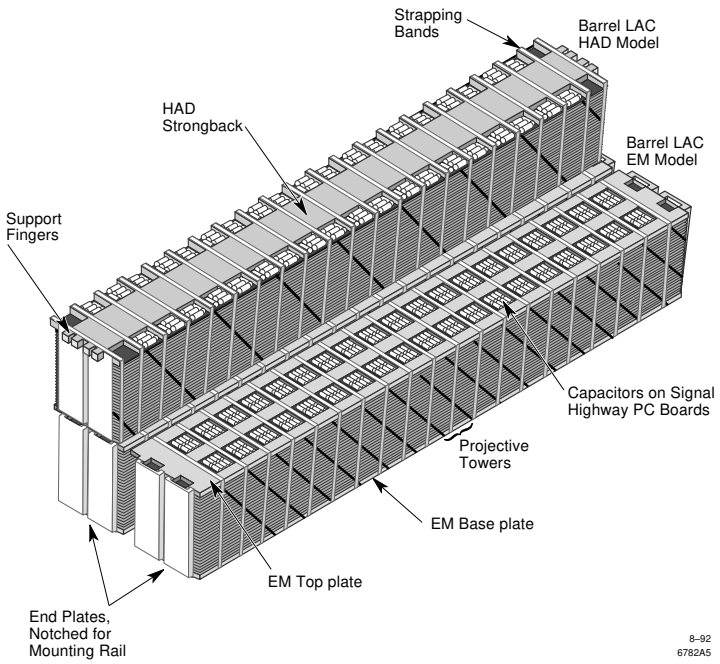


Figure 22: EM and HAD modules in the barrel LAC.

of the LAC modules. About 99% of the energy of a 45 GeV electron is contained in the EM sections, and 90-95% of the total energy of a hadronic Z^0 decay [51] is contained by the entire LAC. The energy resolution [50, 48] for the LAC is approximately $15\%/\sqrt{E}$ for electromagnetic showers and $60\%/\sqrt{E}$ for hadronic showers, with the energy E given in GeV.

2.2.6 The Warm Iron Calorimeter (WIC)

The Warm Iron Calorimeter (WIC) [52] serves multiple roles. Primary roles are the flux return for the solenoid and the structural support for the SLD, also it has also been instrumented to provide muon identification and additional calorimetry information. The intent to use the WIC for containing the 5-10% energy from hadronic showers leaking through the LAC was not realized due to problems in the calibration of its energy response. The WIC is constructed from 18 layers of Iarocci streamer tubes [53] contained in 3.2 cm gaps between 5 cm thick steel plates.

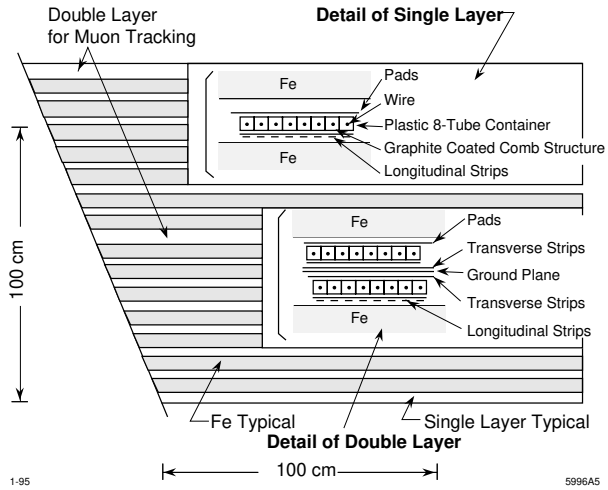


Figure 23: Cutaway view of a WIC section, including details of single layers and double layers.

Figure 23 illustrates a section of the WIC. The Iarocci tubes are long, rectangular plastic extrusions with central copper-beryllium anode wires, and filled with 88% CO_2, 9.5% isobutane, and 2.5% Ar. The tubes feature external copper cathode readouts, square readout pads for calorimetric measurements, and long strips for muon tracking, arranged in separate, perpendicular arrays.

2.3 The SLD Event Trigger

The 8.3 ms SLC beam crossing time simplifies the design of the trigger which decides if data should be recorded to tape. The data acquisition at the SLD utilizes a FASTBUS architecture. Data from the various detector sub-systems reach "slave" modules, which depend on the sub-system being read out via fiber optics connections. The slave modules process data from the sub-systems, as summarized in Table 6.

The data from these slave modules is pooled in ALEPH event builder (AEB)

Table 6: The slave modules for the different detector sub-systems.

Detector sub-system	Slave module
VXD2 and VXD3	Vertex Data Acquisition (VDA)
CDC and CRID	Waveform Sampling Module (WSM)
LAC	Calorimeter Data Module (CDM)
WIC	Digital Readout Module (DRM)

modules [54] which assemble complete data events by combining the data from each sub-system. A dedicated trigger AEB performs the data triggering. The trigger criteria are given in more detail in Refs. [55], [56] and are summarized below. Except for the Bhabha trigger, all sub-systems of the detector are written out when a trigger occurs.

- Energy trigger: requires at least 8 GeV of total deposited energy in the EM and/or HAD calorimeter towers in the LAC. Only towers above the threshold of 60 (120) ADC counts for EM (HAD) towers, corresponding to 246 MeV (1.298 GeV), contribute.

- Charged Track trigger: based on a pattern map of the cells that might be hit as a charged track of momentum greater than 250 MeV/c passes through the CDC [57]. A hit cell is defined as a cell where pulses consistent with hits from tracks fire at least 6 of the possible 8 wires. Only events containing two tracks passing through at least 9 superlayers of the CDC and lying roughly 120^o apart are passed by this trigger.

- HAD trigger: combines the first two triggers

- Muon trigger: requires a combination of a charged track in the CDC and hits in opposite WIC octants, as this is the expected signature for a $Z^0 \rightarrow \mu^+\mu^-$

event.

- WAB trigger: enables the recording of wide angle Bhabha events by requiring two charged back-to-back tracks in the CDC. Even short track pieces, dismissed by the tracking trigger, may satisfy this requirement.

- Bhabha trigger: requires at least 12.5 GeV of total energy in both the north and south EM2 sections of the LUM. Background splashes on the luminosity monitor do not trigger detector readout.

- Random trigger: occurs every 20 seconds. This trigger writes out data for background studies, independent of the status of other triggers.

The software controlling data taking at the SLD runs on a Digital Equipment Corporation VAX/VMS computer cluster. Data is written to tape via fiber optic connections to a silo located in the SLAC computing center. A typical SLD event is roughly 250-300 kbytes. About 40%, 25%, and 25% in size is contributed by the CRID, the drift chambers, and the VXD, respectively. Background conditions, varying with the tuning of the SLC, have a strong impact on the event size. Typical trigger rates were about 0.2 Hz (0.5 Hz) for the 1993-6 (1997-8) data taking periods under low background conditions. It is estimated that the combined efficiency for the three hadronic triggers (Energy, Tracking, HAD) exceeds 96% [55] for accepting hadronic events.

2.4 The SLD Monte Carlo

Modern Particle Physics experiments rely on detailed Monte Carlo simulations to model both underlying physical processes and the detector response. The simulation process consists of two steps. The generation of simulated events starts with the production of generator level events, followed by a full detector simulation of the SLD detector response.

The JETSET 7.4 [58] event generator, based on the LUND fragmentation model with parameters tuned to hadronic e^+e^- annihilation data [59], was used to generate the 1994-8 SLD Monte Carlo. JETSET 6.3 was used for the 1993 SLD Monte Carlo, with similar tuning. For c and b quarks, the Peterson fragmentation function [74] was used, and the B hadron decay model [48],[61] developed at CLEO was tuned to ARGUS [62] and CLEO [63] data. In the next step, acceptance, efficiency, and resolution effects are modeled in a detailed Monte Carlo simulation of the detector. The SLD uses the standard GEANT 3.21 [64] package which tracks particles through the various sub-detectors, accounting for e.g. the magnetic field, scattering, and energy loss. The description of the showering in calorimeters is based on EGS4 [65] and GEANT GHEISHA [66] for electromagnetic and hadronic interactions, respectively. Random trigger data are overlaid with the results of the simulation to simulate both, beam-related backgrounds and electronic noise. Finally, simulated data and real data are processed by the same reconstruction package.

The detailed simulations of tracking [45], calorimetry [51],[40], and CRID response [67] give confidence that the detector response is properly modeled. Inclusive distributions of single-particle and event-topology observables in hadronic events are found to be well described by the SLD simulation [56].

The simulated $Z^0 \to q\,\overline{q}$ events are generated using JETSET 7.4 [58]. The B meson decays are simulated using the CLEO B decay model [68] tuned to reproduce the spectra and multiplicities of charmed hadrons, pions, kaons, protons and leptons as measured at the Υ(4S) by ARGUS and CLEO [69]. Semileptonic decays of B mesons follow the ISGW model [70] including 23% D^{**} production. The branching fractions of the charmed hadrons are tuned to the existing measurements [73]. The b-quark fragmentation follows the Peterson *et al.* parameterization [74]. Finally, the SLD detector is simulated using GEANT 3.21 [75].

CHAPTER 3

Event Selection

In this chapter the identification of hadronic Z decays and b-tagging, the prerequisites of the B_s^0–$\overline{B_s^0}$ mixing analyses, are discussed. The procedure is to first flag hadronic events with triggers and filters, then to tag long lived heavy flavor decays by searching for events that have vertices of tracks displaced from the interaction point (IP). The last step is to discard Z$\rightarrow c\overline{c}$ events using a Neural Net based on the vertex mass (assuming that all tracks are pions and correcting for neutral decay products), the total charged track momentum, the flight distance of the vertex and the track multiplicity.

3.1 Hadronic Event Selection

The SLD triggers that are used to flag hadronic events are the Energy, Track and Hadron triggers, described in section 2.3. The rejection of backround is improved by a filter that requires either a >1 GeV/c track in the drift chamber or energy deposition in the LAC inconsistent with beam-gas or SLC muon background events. The combination of the trigger system and this filter is estimated to be 92% efficient for hadronic Z decays [80].

To enhance the fraction of hadronic Z decays the charged track multiplicity and their total energy is taken into account. At least seven CDC tracks are required per event, where a well-measured track has momentum transverse to the beam direction $p_{perp} > 0.2$ GeV/c and passes within 5 cm of the IP at the point of closest approach to the beam axis. In addition there must be at least three CDC tracks with hits in the vertex detector. The total energy measured using CDC tracks must exceed 18 GeV. Detector acceptance is also considered by evaluating the event thrust axis, $\vec{T}$, which is intended to provide information about the energy flow direction of the event. The thrust axis is located by maximizing the following quantity T,

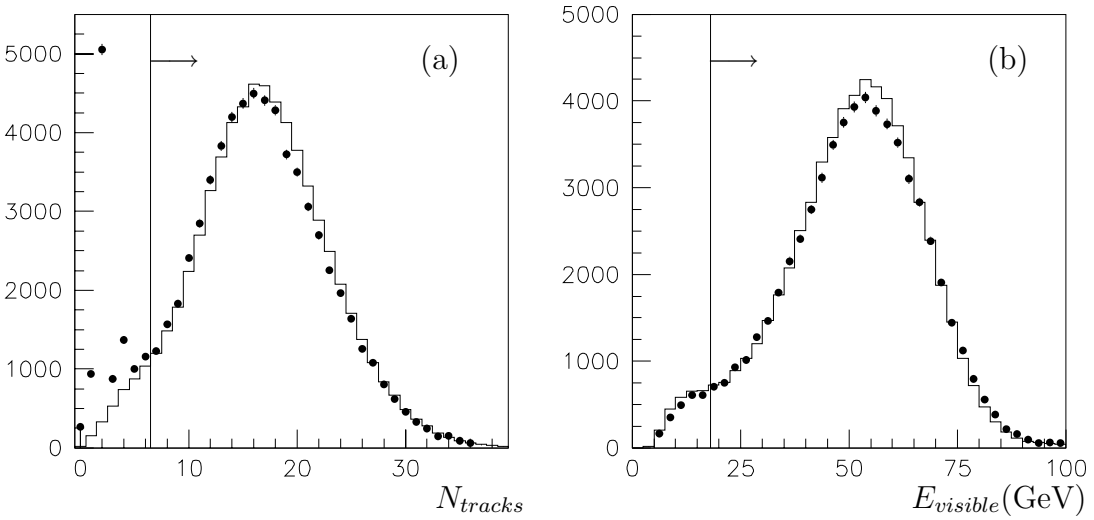
$$T = \frac{\sum^i |\hat{T} \cdot \vec{E}_i|}{\sum^i |\vec{E}_i|} \tag{3.1}$$

where $\vec{E}_i$ represents the energies and directions of final state particles determined by calorimeter clusters. The discussed quantities are shown in Figure 24. A total of 228712 events from the SLD 1996-98 data pass this hadronic event selection. Background, predominately due to τ pairs, is estimated at $< 0.1\%$.

3.2 Heavy Flavor Tagging

In a next step long-lived heavy hadrons are selected. The strategy is to look for vertices of tracks displaced from the IP. First, the IP has to be reconstructed using charged tracks. Then a secondary vertex is determined with a topological vertexing technique. Since the mass of the Z is significantly greater than the mass of the quark pairs it decays into, the selected events can be divided into two hemispheres. For example, in $Z^0 \to b\,\overline{b}$ events a b hadron decay is expected in each hemisphere (except in the case of hard gluon radiation). In each hemisphere there will usually be tracks originating from three vertices: the fragmentation tracks from the primary vertex

Figure 24: **Distributions of event selection variables: (a) Number of CDC tracks, (b) visible energy. The points (histogram) denote the data (Monte Carlo). The arrows indicate the regions which pass the cuts, which are applied cumulatively for each successive plot. A subset of the 1998 data is shown.**



at the IP, tracks from the secondary b hadron decay and tracks from the tertiary d hadron decay.

3.2.1 IP Reconstruction

The common technique for locating the IP is, for each event, to fit all tracks consistent with the luminous region to one vertex. The SLC final focus optics give an IP which is approximately $1\mu m \times 2\mu m$ in the directions transverse to the beam and about 700μm in the longitudinal direction. The beam feedback mechanism used to stabilize collisions indicates that the transverse movement in the IP is smaller than 6μm over periods of 20-30 events. Therefore, the transverse location of the IP can be determined more accurately by averaging over such event sequences. A trial IP location is chosen, and all tracks with VXD hits that pass within 3σ of the trial IP are fit to a common vertex. The new IP position is taken as the new trial

IP and the procedure is repeated until it converges.

As the longitudinal (z) bunch size is much larger, the best estimate of the z coordinate is obtained event-by-event with hadronic decays. The point of closest approach to the transverse IP position is determined for each track with a VXD hit. The z coordinate at that point is called z_{POCA}. Tracks with xy impact parameters less than 500μm that pass within 3σ of the IP are then selected and the z coordinate of the IP is obtained from the median of the z_{POCA} distribution from each event.

The transverse beam position error can be determined on average from the data using decays of the Z into muon pairs. The impact parameter σ_{imp} of such high momentum tracks is plotted in Figure 25. The width of the impact parameter distribution is 8.2μm. This distribution is the convolution of the IP error distribution with the high momentum track error distribution. The track error distribution σ_{track} is independently measured from the miss distance between the two muons. The width of the distribution is about 7.6μm. The IP measurement errors in the xy direction can then be estimated as $\sigma_{IP} = \sqrt{8.2^2 - 7.55^2} = 3.2\mu m$.

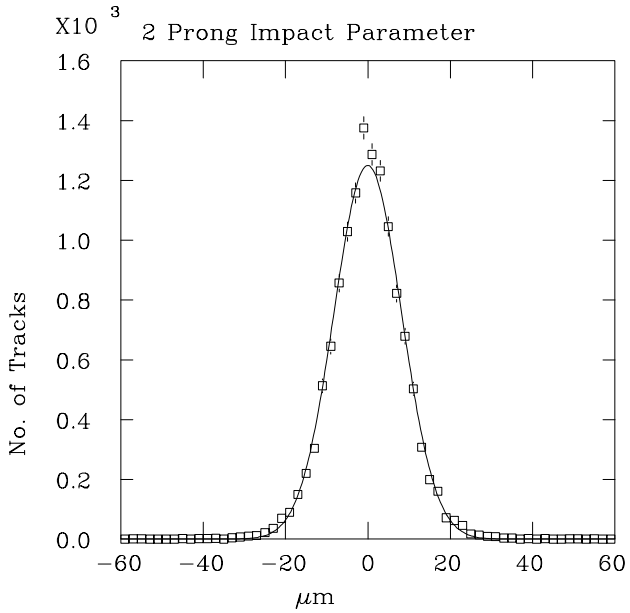
The longitudinal IP position error has to be estimated from MC studies. On average it is 20μm for $b\overline{b}$ events.

In the next section, a general description of Neural Networks is given, as they are used repeatedly throughout the analysis.

3.2.2 Neural Networks

A common problem in particle physics is how to select a subsample of events with certain properties, for example: how to find tracks originating from the B decay vertex, as opposed to fragmentation tracks from the IP. The usual procedure is to apply a cut on one or more of the sensitive variables x_i. One possibility is to select tracks within a certain transverse momentum range. In general, the x_i

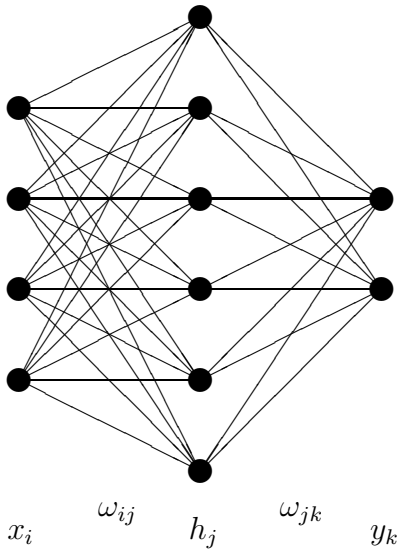
Figure 25: Muon-pair xy impact parameter to the IP.



are correlated, so that a simple cut ignores information from other variables that could improve the selection. A way around that are cuts in 2 or 3 dimensional space, which can be impractical to handle. Neural Networks (NN) are a method to combine the information contained in the x_i to find signal regions. The procedure is essentially an iterative process that, using a 'training sample' of MC, finds the probability for any track being a 'signal track' as a function of the x_i. In other words, the MC training sample is used to 'learn' how the x_i are correlated with each other and with the signal parameter [81]. Neural Networks can be constructed using a variety of architectures and learning algorithms, but the most common are feed-forward networks trained using back-propagation. The discussion here will be restricted to this type.

A typical feed-forward network architecture is shown in Figure 26. The basic elements of a Neural Network are nodes and links. In a feed-forward network the nodes are arranged in layers as shown, with links only between nodes in adjacent

Figure 26: Schematic diagram of a typical feed-forward neural network.



layers. The first layer of nodes corresponds to the input variables x_i, with one node for each variable. Next come any number of hidden layers, with any number of nodes in each. In practice networks with more than two hidden layers are rarely used, and one hidden layer with one more node than in the input layer proves sufficient for most problems. The last layer contains the output node(s) (it is possible to construct a NN with a reduced set of variables y_k instead of only one output node. More than one output node may be in place if the network is to be used to sort the items into more categories than just signal and background).

Feed-forward networks are evaluated layer by layer, so that the input variables x_i are propagated step-by-step through the network to the output layer. The value of a node h_j in the first hidden layer is given by:

$$h_j = g(\sum_i \omega_{ij} x_i + \theta_j), \tag{3.2}$$

where ω_{ij} is the weight assigned to each link between the input and hidden layers, θ_j is the threshold for each hidden node h_j, and $g(z)$ is the node activation function. For purposes as described above, the activation function is usually:

$$g(z) = \frac{1}{2}(1 + \tanh z) = \frac{1}{1 + e^{-2z}}. \tag{3.3}$$

If there are more hidden layers they are evaluated in turn, replacing the input node values x_i with the appropriate h_j values from the previous layer. The output layer values y_k are given by:

$$y_k = g(\sum_j \omega_{jk} h_j + \theta_k), \tag{3.4}$$

where now the ω_{jk} are the weights for the links between the last hidden and output layers, and θ_k are the output node thresholds. ω and θ are free parameters of the network, which are determined by training it against test samples. The training samples are collections of items for which both the input patterns x_i and the desired network outputs t_k are specified. The samples are drawn from the Monte Carlo, which provides good modeling of the data distributions. The training procedure involves minimization of an error measure, usually a mean square error:

$$E = \frac{1}{2N_p} \sum_p \sum_k (y_k^{(p)} - t_k^{(p)})^2, \tag{3.5}$$

where N_p is the number of patterns (items) in the training set, and the (p) denotes the observed and target network outputs for a particular input pattern $x_i^{(p)}$. Training is therefore analogous to performing a χ^2 fit for the parameters w.

3.2.3 Secondary Vertex Reconstruction

The method for reconstructing secondary vertices is to search for space points in 3D where track density functions overlap [84]. Each track is parameterized by

a Gaussian probability density tube $f(\vec{r})$ with a width equal to the uncertainty in the track position at its point of closest approach (POCA) to the IP, $\vec{r}_0$,

$$f(\vec{r}) = \exp\left\{-\frac{1}{2}\left[\left(\frac{x-(x_0+y^2\kappa)}{\sigma_1}\right)^2 + \left(\frac{z-(z_0+y\tan\lambda)}{\sigma_2}\right)^2\right]\right\} \tag{3.6}$$

The first term is a parabolic approximation to the track circular trajectory in the xy plane, where κ is a function of the track charge and transverse momentum and of the SLD magnetic field. The second term represents the track linear trajectory in the rz plane, where λ is the track dip angle from the vertical. The σ parameters are the uncertainties in the track positions after extrapolation to $\vec{r}_0$ for the two projections.

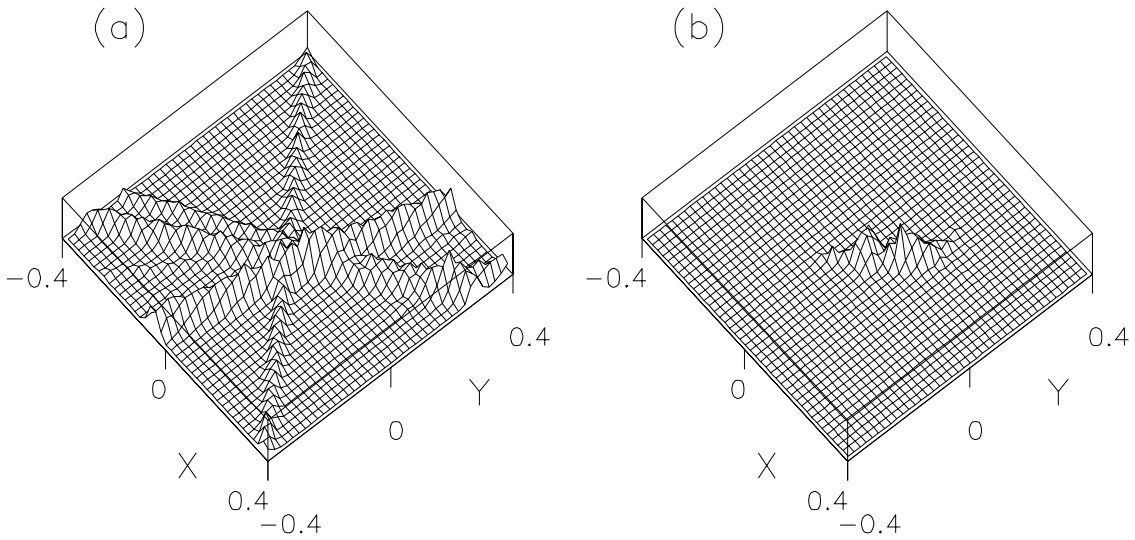
The function $f_i(\vec{r})$ is formed for each track under consideration and used to construct a vertex probability function $V(\vec{r})$. Also included is $f_0(\vec{r})$, a $7\times7\times20\mu$m $(x\times y\times z)$ Gaussian ellipsoid, centered at the IP position, reflects the uncertainty in the position of the IP. A smooth, continuous function is desired so that its maxima may be found. The requirements result in the form:

$$V(\vec{r}) = \sum_i^N f_i(\vec{r}) - \frac{\sum_i^N f_i^2(\vec{r})}{\sum_i^N f_i(\vec{r})}, \tag{3.7}$$

where N is the number of tracks. The track and vertex functions are shown in Figure 27 for a simple MC event hemisphere. Secondary vertices are found by searching for local maxima in $V(\vec{r})$ that are well-separated from the peak at the IP position. The tracks whose density functions contribute to a local maximum are then identified as originating from a secondary vertex (SV).

A loose set of cuts is applied to tracks used for secondary vertex reconstruction. Tracks are required to have ≥ 3 VXD hits and $p_{perp} > 250$ MeV. Tracks with 3D impact parameter > 3 mm or consistent with originating from a γ conversion or

Figure 27: The track (a) and vertex (b) functions projected onto the xy plane (cm). Primary and secondary vertices are visible in (b).



from K^0, or Λ^0 decay are also removed. The event is divided into two hemispheres using the thrust axis, and the vertexing procedure is performed in each using only the tracks in that hemisphere.

The identified vertices are required to be within a radius of 2.3 cm of the center of the beam pipe to remove false vertices from interactions with the detector material. A mass cut of $|M_{VTX} - M_{K^0}| < 0.015$ GeV removes any K^0 decays that survived the track cuts. The remaining vertices are then passed through a simple neural network to further improve the background rejection. A subset of the 1998 Monte Carlo was used for training purposes. The input variables are the flight distance from the vertex to the IP, that distance normalized by its error, and the angle between the flight direction and the total momentum vector of the vertex. These quantities are shown in Figure 28, along with the output of the neural network, for a separate Monte Carlo subsample. Vertices with $NN_{vtx} > 0.7$ are retained. At least one good secondary vertex is found in 72.7% of bottom, 28.2% of charm, and

0.41% of light quark hemispheres in the Monte Carlo. Around 16% of the bottom hemispheres have more than one good secondary vertex.

3.2.4 Track Attachment

Due to the cascade nature of B decays, not all secondary tracks originate from the same space point. In order to recover information about the true B vertex location and for better reconstruction of the b hadron charge, a process of reconsidering discarded tracks and reattaching tracks to the secondary vertex has been developed [82]. It uses a second neural network.

The network is trained to accept only tracks which come from a B or D decay, and to reject tracks from the IP or from strange particle decays. To optimize the charge reconstruction, any track not already part of a good SV and with ≥ 2 VXD hits is tried, including those which were removed from the SV-finding procedure.

The first four inputs of the NN are defined at the point of closest approach of the track to the axis joining the secondary vertex to the IP. They are the transverse distance from the track to that axis (T), the distance from the IP along that axis to the POCA (L), that distance divided by the flight distance of the SV from the IP (L/D), and the angle of the track to the IP-SV axis (α). The last input is the 3D impact parameter of the track to the IP normalized by its error (b/σ_b). These quantities are shown schematically in Figure 29. The distributions are shown in Figure 30, along with the neural network output. Tracks with $NN_{trk} > 0.6$ are added to the list of secondary vertex tracks.

3.2.5 VXD-Alone Tracking

Because the SLD tracking system is not 100% efficient, not all of the heavy hadron charged decay products will be found even for perfect secondary vertex

Figure 28: Distributions of seed vertex selection variables: (a) distance from IP, (b) normalized distance from IP, (c) angle between flight direction and vertex momentum, (d) neural network output. The arrow indicates the accepted region.

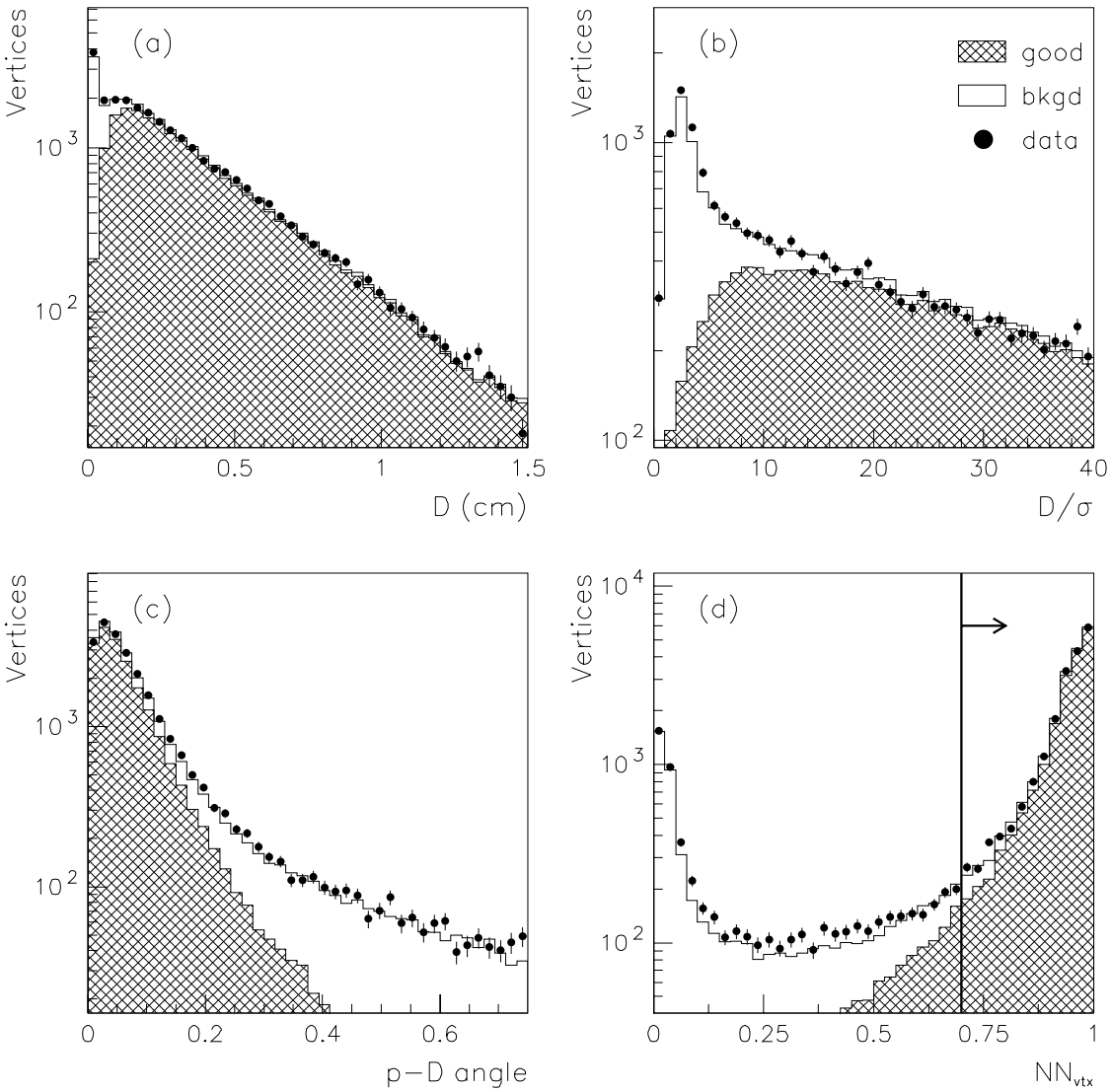
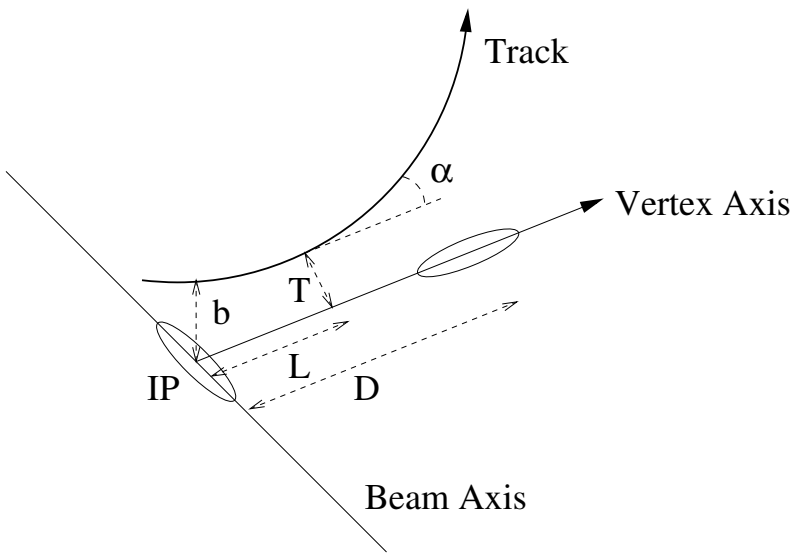


Figure 29: Schematic illustration of the quantities used in the track-attachment procedure described in the text (not to scale).



reconstruction. The MC indicates that ∼90% of the charged decay products of a heavy hadron produce a VXD-linked track. Part of the inefficiency can be recovered by using tracks reconstructed in the vertex detector alone. Vectors of VXD hits not associated with a CDC track are used, where there is at least one hit on each of the three VXD layers. Including these vectors raises the tracking efficiency to ∼97% for heavy hadron decay products. A trial track is constructed from each vector using a helix fit to the VXD hits. A third neural network is used to select the vectors consistent with the heavy hadron decay chain, using as inputs the first four variables used for the track attachment NN. These quantities are shown in Figure 31, along with the neural net output. The normalized impact parameter is not used since the momentum determination is poor and the error can't be reliably calculated. Vectors with $NN_{out} > 0.5$ are added to the secondary vertex track list.

Once a vector has been attached to a secondary vertex, the helix fit is repeated but with the vertex as an additional point to improve the curvature determination.

Figure 30: Distributions of cascade track selection variables: (a) T, (b) L, (c) L/D, (d) α, (e) b/σ_b, (f) neural network output. The arrow indicates the accepted region.

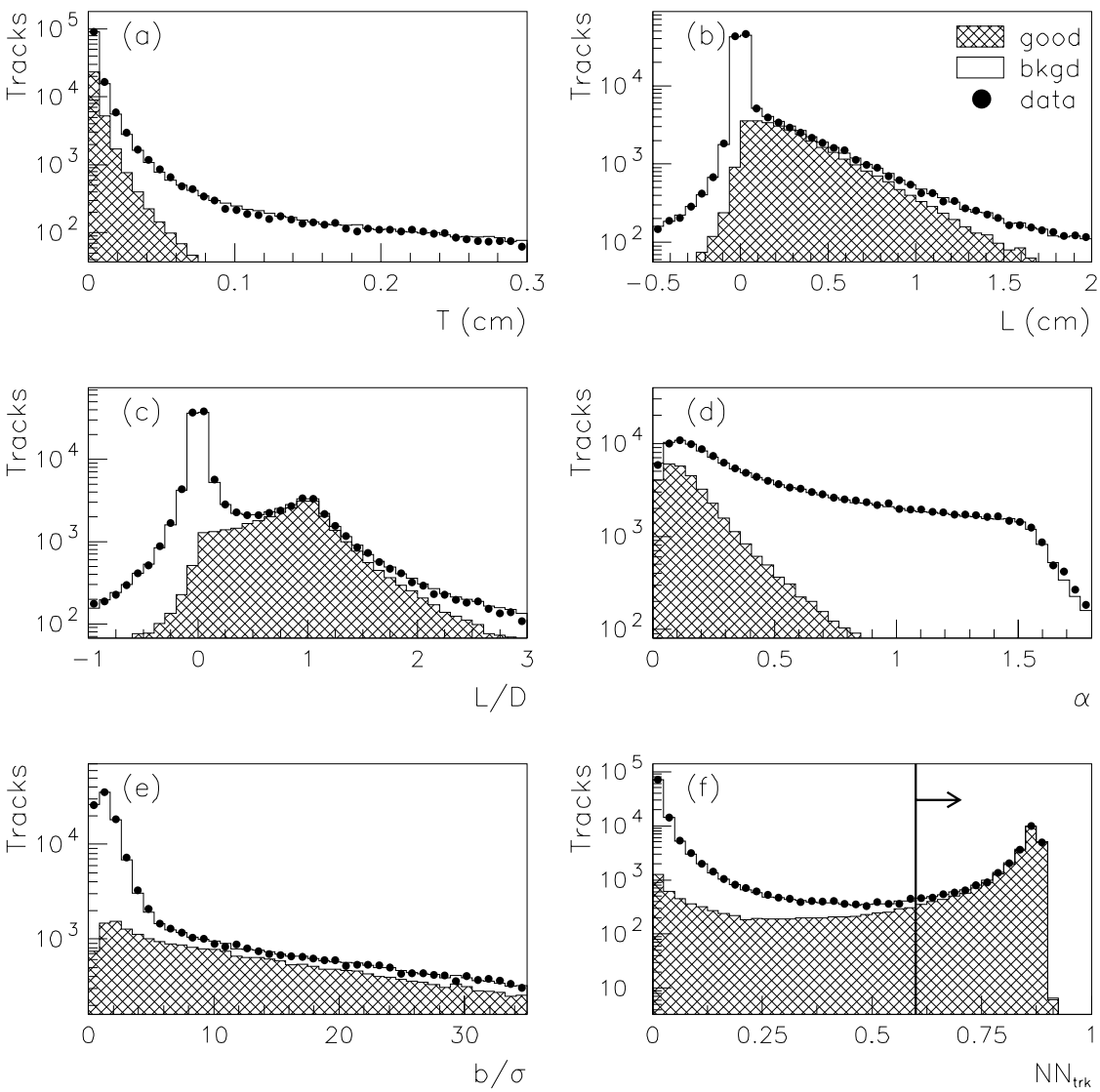
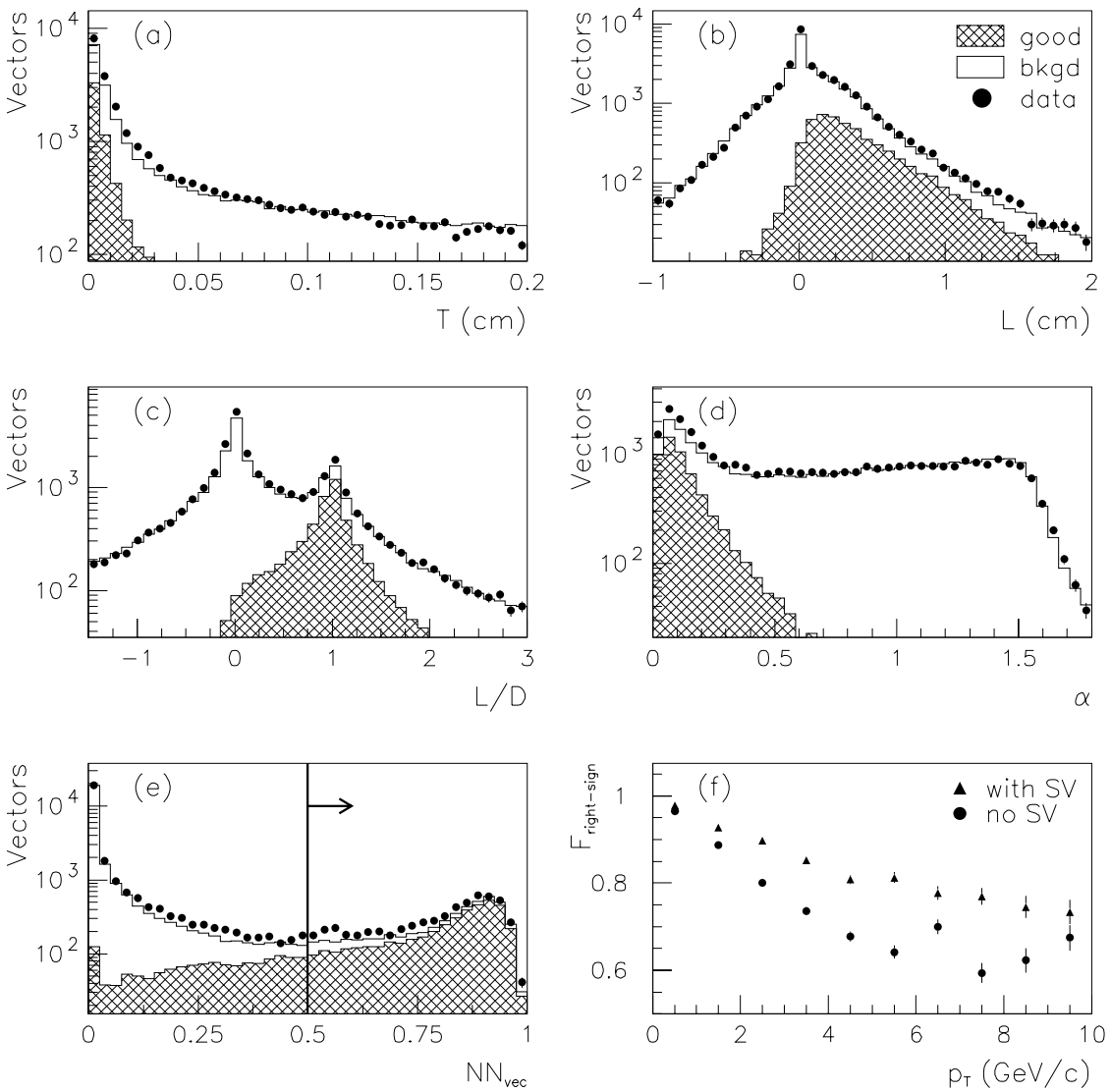


Figure 31: Distributions of VXD-vector selection variables: (a) T, (b) L, (c) L/D, (d) α, (e) neural network output. The arrow indicates the accepted region. The probability to assign the correct charge to a vector based on its fitted curvature is shown in (f), both with and without the secondary vertex as a constraint.



The charge of the underlying particle is correctly found from this fit for $\sim 85\%$ of the attached vectors. Figure 31f shows the correct-sign probability for attached vectors as a function of the true momentum of the underlying MC particle, both with and without the secondary vertex constraint. At this point each hemisphere is associated with a list of tracks, including tracks that were already used to fit a secondary vertex, attached tracks and attached VXD-alone tracks.

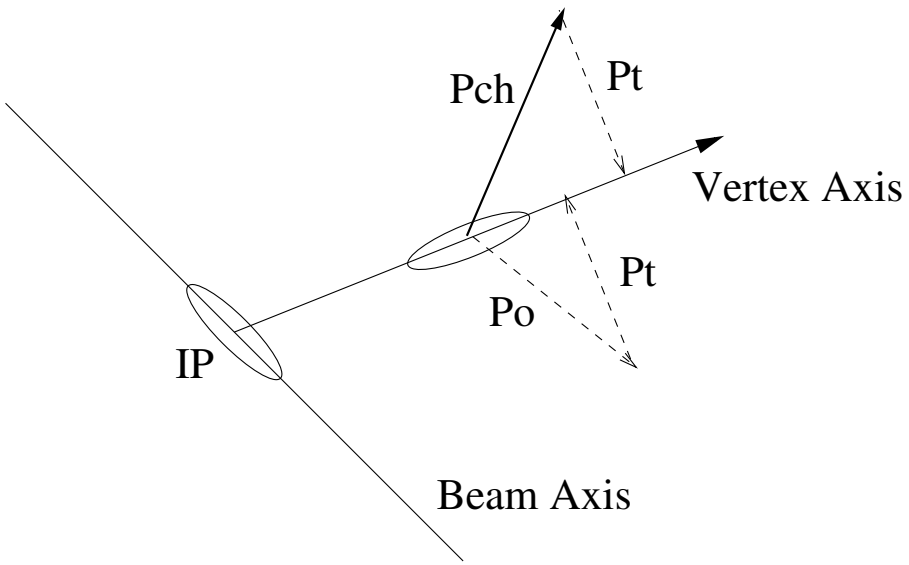
3.2.6 Flavor Discrimination

From the list of secondary vertex tracks and attached tracks or VXD-alone tracks several signatures can be computed that help to discriminate between bottom, charm and light hemispheres [82]. The first three do not use the VXD-alone tracks because of their poor momentum resolution. A vertex position for the hemisphere as a whole is determined by fitting all SV and cascade tracks to a common vertex. The first signature is the vertex mass. Each track is assigned the mass of a charged pion and the invariant mass M_{ch} of the selected tracks is calculated. This mass can be partially corrected for the unknown contribution from neutral decay products using the vertex transverse momentum. The minimum amount of momentum P_t^{min} required to align the charged momentum $\vec{P}_{ch}$ of the vertex with its flight direction $\vec{D}$ from the IP to within errors is found, as shown in Figure 32. The charged mass is then corrected to obtain the P_t-corrected vertex mass

$$M_{VTX} = \sqrt{M_{ch}^2 + P_t^{min^2}} + |P_t^{min}|. \tag{3.8}$$

The magnitude of P_t is constrained to be $\leq M_{ch}$ to prevent fake vertices from resulting in large masses through this correction. The second and third signatures are the magnitudes of the charged momentum $|P_{ch}|$ and flight distance $|D|$ of the

Figure 32: Schematic illustration of the missing P_t determination.

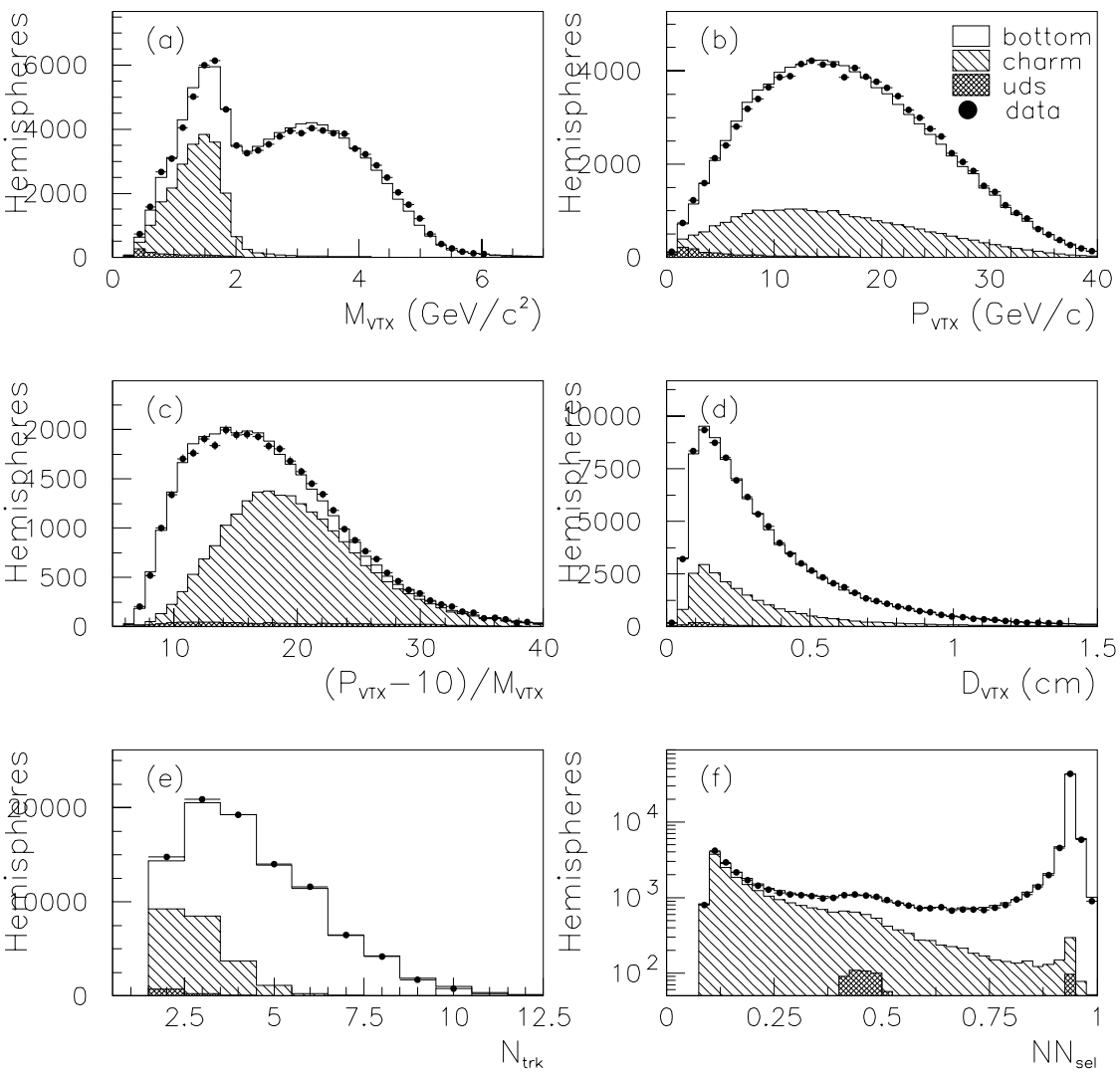


total hemisphere vertex. The last signature is just the selected track multiplicity, now with the VXD-alone tracks included. These four signatures are used as inputs for a neural network trained to distinguish bottom/charm/light hemispheres. The four inputs and the neural network output are shown in Figure 33. Also shown is the correlation between M_{vtx} and P_{ch} for bottom and charm. This correlation in the low M_{vtx} region provides a mechanism for selecting pure samples of charm hemispheres.

This last selection neural network is trained to put charm hemispheres near zero, bottom hemispheres near one, and light-flavor background near one-half. This allows a simple selection of bottom hemispheres by specifying an lower limit for the NN output.

A cut requiring the total P_t-corrected mass of all tracks in the vertex to be >2 GeV/c^2 in the event hemisphere achieves an efficiency of 62% and a b-purity of 98% (from all-flavor Monte Carlo). In the 96-98 SLD data 73812 hemispheres pass

Figure 33: Distributions of flavor discrimination variables: (a) M_{VTX}, (b) P_{VTX}, (c) $M_{VTX} - P_{VTX}$ correlation, (d) D_{VTX}, (e) N_{trk}, (f) neural network output.



this cut.

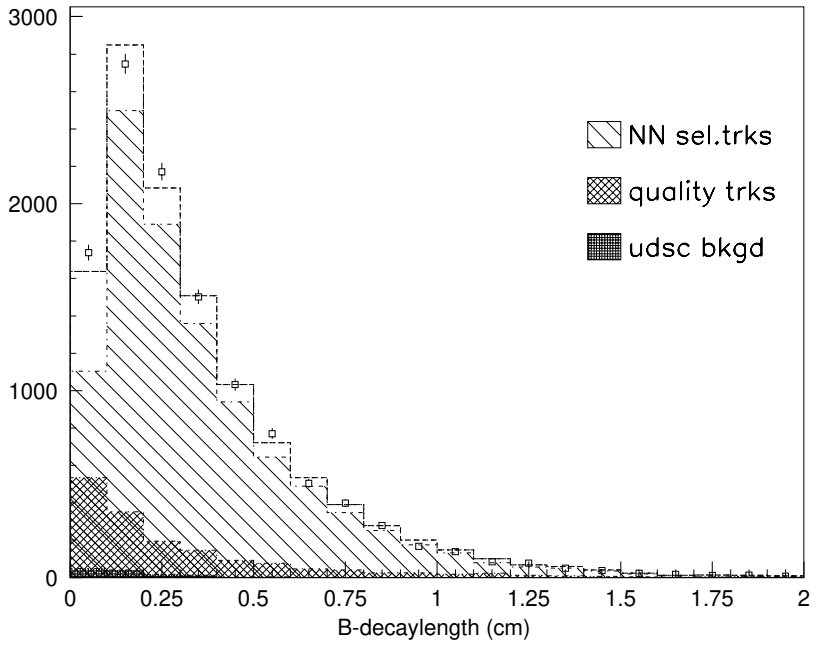
3.2.7 Recovering Low Decay Length Vertices

In the process of IP and light flavor background discrimination, events with low proper time are likely to be discarded, as they may contain IP or light flavor tracks. For example, in the all flavor MC the average true B_s^0 decay length is 0.3 cm, whereas the average decay length of b-tagged B_s^0 mesons is 0.4 cm, indicating a bias toward higher decay length. In order to recover some low proper time events, a different criterion for track selection can be used to find displaced high-mass vertices. This is only done if the high-purity NN b-tagging procedure fails to find a secondary vertex in the event hemisphere, but finds such a vertex in the opposite hemisphere. The criterion for track selection, so-called quality track selection, requires tracks to have $\geq$23 CDC hits, momentum transverse to the beam axis $p_\perp > 250$MeV/c, a $\chi^2/d.o.f$ of the CDC track fit < 8, and a $\chi^2/d.o.f$ of the CDC/VXD fit < 8. In addition, the radius of the first CDC hit is required to be < 39.0 cm, 2-D impact parameter in xy <1 cm, at least 2 VXD hits, and a distance of closest approach to the IP in z < 1.5 cm. The topological vertexing method described earlier [84] is used to find secondary vertices, and the P_t-corrected mass is then required to be > 2 GeV/c^2 in the event hemisphere to tag $Z \rightarrow b\overline{b}$ events. Note that there is no procedure to reject IP or light flavor tracks. However, the requirement of a standard NN-selected B-vertex in the opposite hemisphere keeps the light flavor background at the 1% level overall. The average decay length of vertices selected in this way is 0.2 cm, indicating that this procedure recovers hemispheres containing B decays with low decay lengths.

Using this procedure in addition to the standard NN b tagging method adds 35% more b-tagged data events. The net number of b-tagged events is 99844.

Fig. 34 shows the decay length distribution of standard NN events, quality tracks and background events.

Figure 34: The all-flavor MC B decay length distribution of standard NN selected events, events reconstructed with quality tracks and $udsc$ background (hatched histograms). The points are the 96-98 SLD data.



3.2.8 Tracking Efficiency

The tracking efficiency in the default SLD Monte Carlo appears to be somewhat higher than what can be inferred from the data, therefore some tracks in the Monte Carlo are discarded to produce a better match. The rate at which tracks should be 'tossed' has been estimated using two methods.

The first method selects a sample of hemispheres tagged with $M_{VTX} > 2$ GeV/c^2, and tunes the track-tossing rate to best match the Q_{VTX} distribution

to that observed in the data. The width of the Q_{VTX} distribution is used due to its insensitivity to the generated b hadron decay charged particle multiplicity. This procedure first calculates Q_{VTX} using only the full VXD+CDC tracks, and finds that 1.5% of the tracks should be discarded from the Monte Carlo. This is done by 'unlinking' the track, i.e., breaking the VXD to CDC link, leaving VXD-vectors. To determine what should be done with these vectors, the Q_{VTX} distributions are compared again, using the remaining tracks plus any newly attached VXD-vectors. From this it is found that 75% of the leftover VXD-vectors from the unlinked tracks should themselves be discarded.

The second method to estimate the track-tossing rate compares the fraction of CDC tracks which have associated VXD hits. The rate at which Monte Carlo linked tracks must be considered unlinked (and for tagging purposes, discarded) is found by demanding that this linked fraction match that observed in the data. An advantage of this method is that it can be performed in bins of ϕ and $\cos\theta$ to correct for any local efficiency variations. No significant such effects are seen, and the overall unlinking rate of 1.5% is in excellent agreement with the first method. It is not possible with this method to determine whether the VXD part of an unlinked track should be retained as a VXD-vector, so the 75% tossing rate found from the first method is used.

Applying the tracking efficiency corrections the charge reconstruction of b-tagged events is shown in Fig. 35. The achieved charge reconstruction purity is 83%. The plot demonstrates SLD's ability to select neutral b hadrons, due to the high tracking efficiency and the high b tagging purity.

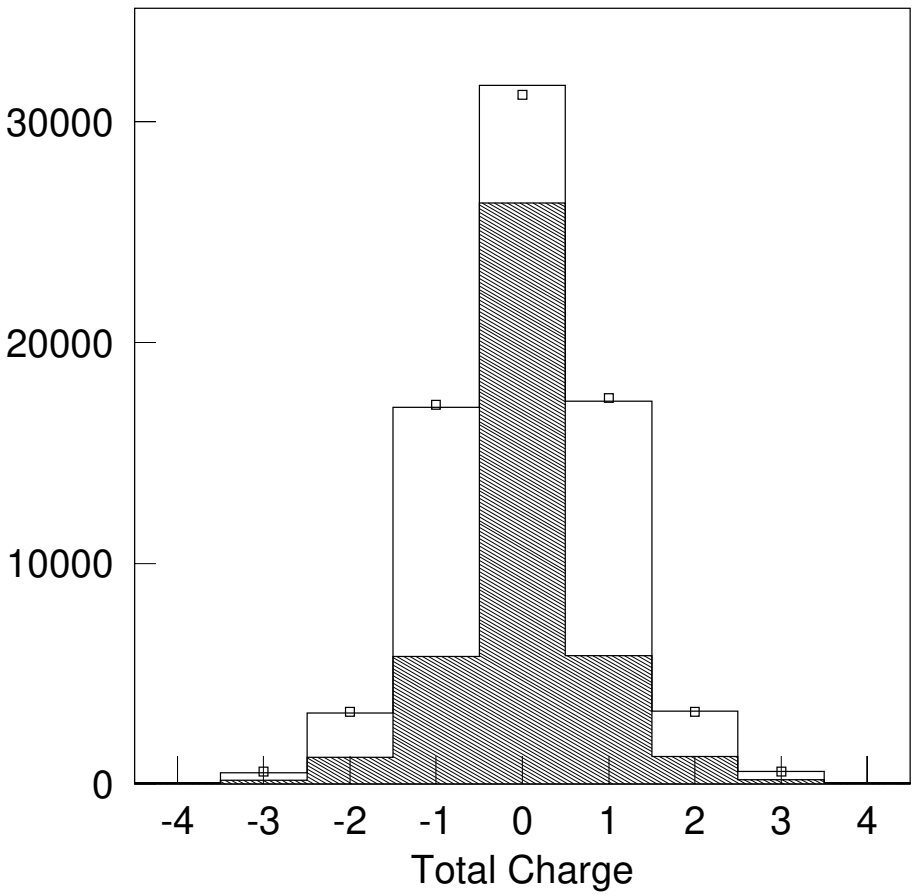


Figure 35: **The overall reconstructed charge of b-tagged events. The points are the data, the histogram is the all-flavor Monte Carlo, and the shaded histogram shows the fraction of true neutral B decays.**

CHAPTER 4

Flavor Tagging and Proper Time Reconstruction

In this chapter the strategy to analyze time dependent B_s^0–$\overline{B_s^0}$ mixing is discussed. First, the B flavor has to be determined at production and at decay, to determine if mixing occurred or not. Secondly, the B decay proper time has to be reconstructed. Final state tagging and proper time reconstruction will be discussed below in section 4.2 on the charge dipole analysis. These three ingredients will determine a time dependent mixing probability. Finally, a method to estimate the signal to noise ratio of the experiment using Fourier analysis is discussed.

4.1 B flavor at production: the Initial State Tag

The initial state B flavor is determined with a combination of six tags of differing strength. Some of them use information about the opposite hemisphere of the decay hemisphere. These tags, in rough order of importance, are the polarized forward-backward asymmetry, jet charge, vertex charge, charge dipole, lepton and kaon tags. The charge dipole initial state tag is identical to the charge dipole final state tag, which will be discussed in detail in section 4.2.

4.1.1 Polarized Forward-Backward Asymmetry

The differential cross section for polarized $e^+e^- \to Z \to b\overline{b}$ has a large forward-backward asymmetry, i.e., for polarized Z decays the outgoing quark is produced preferably along the direction opposite to the spin of the Z. The polarized forward-backward asymmetry $\tilde{A}_{FB}$ can be described by

$$\tilde{A}_{FB}(\cos\theta_T) = 2A_b \, \frac{A_e - P_e}{1 - A_e P_e} \, \frac{\cos\theta_T}{1 + \cos^2\theta_T} \, , \tag{4.1}$$

where $A_b = 0.935$ and $A_e = 0.150$ (Standard Model values), P_e is the electron beam longitudinal polarization ($P_e > 0$ for right handed electron beam), and θ_T is the angle between the thrust axis and the electron beam direction (the thrust axis is signed such that it points in the same hemisphere as the reconstructed B vertex). Thus, left- (right-)polarized electrons tag b ($\bar{b}$) quarks in the forward hemisphere, and $\bar{b}$ (b) quarks in the backward hemisphere. Averaged over the acceptance, this yields a correct tag probability of 72% for an average electron polarization of $P_e = 73\%$. The probability for correctly tagging a b quark at production is expressed as

$$P_A(\cos\theta_T) = \frac{1 + \tilde{A}_{FB}(\cos\theta_T)}{2} \, . \tag{4.2}$$

Fig. 36 shows the distributions for data and Monte Carlo tagged using the polarized forward-backward asymmetry, as selected in the charge dipole analysis (the description of the analysis follows in the next section). It indicates the clear separation between b and $\overline{b}$ quarks and good agreement between data and Monte Carlo.

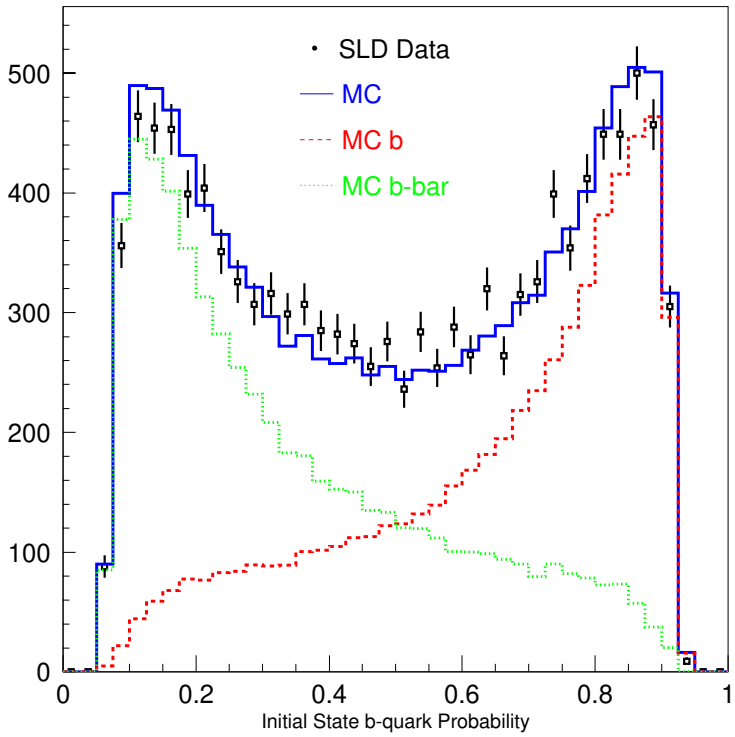


Figure 36: **Distribution of the initial state b-quark probability using the polarized forward-backward asymmetry tag. Data (points) and Monte Carlo (histograms) showing the b and $\bar{b}$ components for the events selected in the charge dipole analysis.**

4.1.2 Jet Charge

A jet charge technique is used in addition to the polarized forward-backward asymmetry. For this tag, tracks in the hemisphere opposite that of the reconstructed vertex are selected. These tracks are required to have momentum transverse to the beam axis $p_\perp > 0.15$ GeV/c, total momentum $p < 50$ GeV/c, impact parameter in the plane perpendicular to the beam axis $\delta < 2$ cm, distance between the primary vertex and the track at the point of closest approach along the beam axis $\Delta z < 10$ cm, and $|\cos\theta| < 0.90$. With these tracks, an opposite hemisphere

momentum-weighted track charge is defined as

$$Q_{opp} = \sum_i q_i \left| \vec{p}_i \cdot \hat{T} \right|^{\kappa} \ , \tag{4.3}$$

where q_i is the electric charge of track i, $\vec{p}_i$ its momentum vector, $\hat{T}$ is the thrust axis direction, and κ is a coefficient chosen to be 0.5 to maximize the separation between b and $\overline{b}$ quarks. This technique yields an average correct tag probability of 66% and is independent of the polarized forward-backward asymmetry tag.

4.1.3 Opposite Hemisphere Topological Tags

Finally, the tag is further enhanced by the addition of other flavor-sensitive quantities from the hemisphere opposite that of the selected vertex. For this purpose, the inclusive topological vertexing technique mentioned earlier is used. The sensitive variables are:

- total track charge and charge dipole of a topologically reconstructed vertex.
- the charge of a kaon identified in the Cherenkov Ring Imaging Detector.A $K^-(K^+)$ signifies that the parent hadron had a $b(\overline{b})$ quark. If more than one kaon is associated with the B decay chain, the total charge is used.
- charge of a lepton originating from the B vertex. The charge correlation between b quark and lepton follows the chain $\overline{b} \rightarrow W^+ \rightarrow l^+$.

4.1.4 Combination of the Tags

The various tags are combined using a series of neural networks to form an overall initial state tag characterized by a b quark probability P_i [83]. Neural Networks are of advantage in this situation, because of the strong correlations

between the tags. For practical reasons three separate networks are trained first for vertex charge, charge dipole and lepton/kaon/jet charge (these last three tags are combined).

Table 7: Input parameters for a Neural Net combining several initial state tags.

Tag	input parameters	average tag purity	efficiency
vertex charge	P_t corrected mass vertex charge vertex decay length number of VXD-alone tracks	74%	52%
charge dipole	B,D charge difference P_t corrected mass B decay length vertex charge	68%	29%
lepton,kaon, jet charge	Q_{jet} vertex charge B decay length lepton charge (weighted by transverse momentum) sum of kaon charges (weighted by total momentum of kaons)	67%	20%

A fourth net is trained to combine the results of the three charge tag neural nets and returns the b-quark probability P_i as the output. Finally, the result of the combination net is analytically combined with the independent b quark probability from the polarization tag. The overall tagging purity is 78%, however the information is used on an event-by-event basis, so that a big fraction of the events have the benefit of a very high initial state tag purity. The combined tag is 100% efficient. Fig. 37 shows the P_i distributions for data and Monte Carlo as selected in the charge dipole analysis (the description of the analysis follows in the

next section). It indicates the clear separation between b and $\overline{b}$ quarks and good agreement between data and Monte Carlo. The tagging power has improved due to the addition of the jet charge tag and the opposite hemisphere tags.

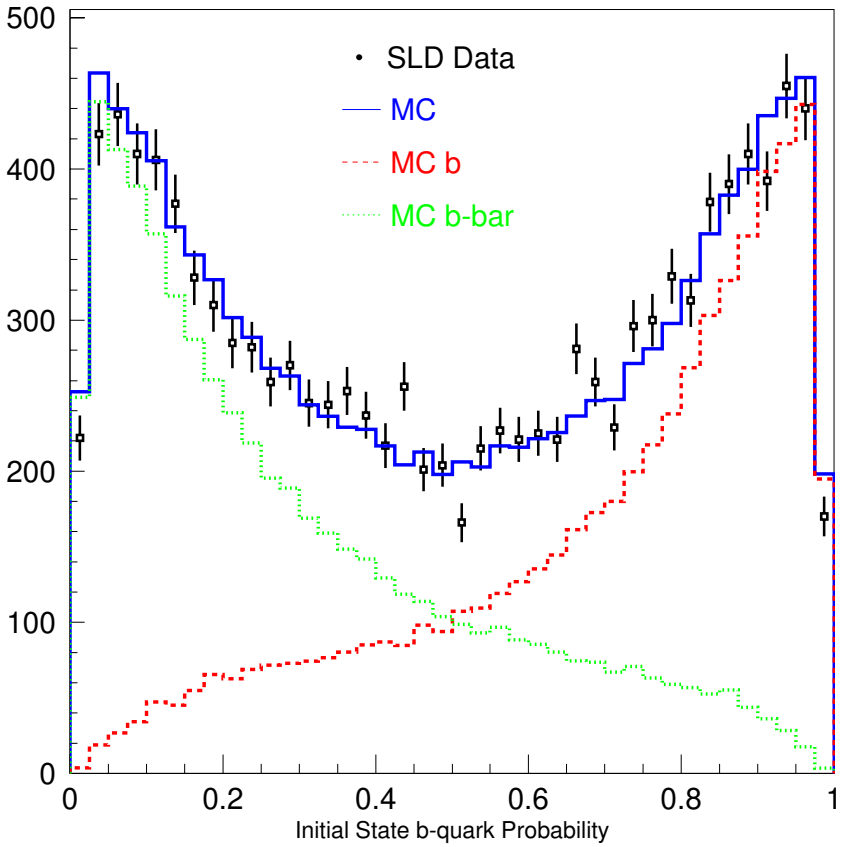


Figure 37: Distribution of the computed initial state b-quark probability for data (points) and Monte Carlo (histograms) showing the b and $\overline{b}$ components for the events selected in the charge dipole analysis.

4.2 B flavor at decay: Vertex Charge Dipole Analysis

The charge dipole analysis aims at selecting decays with distinct B and D vertices and tags the B^0 or $\overline{B^0}$ decay flavor based on the charge difference between them. This analysis technique was first developed at SLD and relies extensively on

the excellent resolution of the vertex detector. It is the most inclusive and therefore the most efficient final state tag at SLD.

The algorithm used to identify primary, secondary and tertiary vertices ('ghost track algorithm') is described in detail in appendix B. It is optimized for this analysis. The vertex finding process relies on the kinematic fact that the boost of the B decay system carries the cascade charm decay downstream from the B decay vertex. In general, any secondary and tertiary vertices will be located along the direction of the jet. First, the best estimate of the straight line from the IP directed along the B decay chain is found. This line is promoted to the status of a track ('ghost track') by assigning it a finite width. Secondly, the selected tracks are vertexed with the ghost track and the IP to build up the decay chain along the ghost direction. The ghost track algorithm operates on any set of input tracks.

The charge difference between the reconstructed secondary and tertiary vertices is the charge dipole flavor tag, and the reconstructed decay length of the secondary vertex is used to reconstruct the B decay proper time.

Once an event passes the b tag, the same tracks that make up secondary vertices are used to feed into the ghost track algorithm. Note that there are two groups of tracks, as described in sections 3.2.6 and 3.2.7: the large group of standard NN selected secondary tracks (seed tracks, attached tracks and VXD-alone tracks), and the small group of quality tracks originating from low decay length B decays. The ghost-track algorithm operates on these groups of tracks and arranges them into one or several vertices. Hemispheres in which three vertices are found – the primary (IP), a secondary and a tertiary – are used for the charge dipole analysis. The secondary vertex is identified as the B decay vertex and the tertiary as the cascade charm decay. As well as improving the purity and efficiency of the dipole reconstruction (by requiring the vertices be consistent with a single line of flight),

the ghost track algorithm has the additional advantage of allowing the reconstruction of 1-prong vertices (vertices with only one track attached to them), including the topology consisting of 1-prong B and D decays.

In the case of quality track hemispheres VXD-alone tracks are attached to the B decay chain only after vertices have been located, to improve the overall charge reconstruction. The attachment criteria rely on the variables $\tilde{T}_i$ and $\tilde{L}_i$ for each track i, as defined below. A vertex axis is formed by a straight line joining the IP to a vertex combining both secondary and tertiary tracks. The 3-D distance of closest approach of the track to the vertex axis, $\tilde{T}_i$, and the distance from the IP along the vertex axis to this point, $\tilde{L}_i$, are calculated. Track i is attached to the vertex if $\tilde{T}_i < 0.1$ cm, $\tilde{L}_i > 0.025$ cm and $0.25 < \tilde{L}_i/\tilde{L} < 2$ (where $\tilde{L}$ is the distance between the IP and the combined vertex). An average of 0.2 VXD-alone tracks is added per decay, to be compared with an average of 5.0 VXD+CDC tracks per decay (for $M > 2$ GeV/c^2). The VXD-alone tracks are attached to either the B or D vertex according to their longitudinal displacement $\tilde{L}_i$: tracks with $\tilde{L}_i < L_B + 0.5(L_D - L_B)$ are attached to the B vertex and all others are attached to the D vertex; L_B (L_D) is the distance between the IP and the B (D) vertex.

4.2.1 Final State Tag Event Selection

For the charge dipole analysis hemispheres containing both a secondary and a tertiary vertex are selected. This requirement rejects $\sim$71% of the b-tagged hadronic events, leaving a sample of 29208 event hemispheres. Furthermore, the total track charge Q_{tot} (from both secondary and tertiary vertices) is required to be zero, and B and D charges are required to be different from each other. This is meant to enhance the fraction of B_s^0 decays in the sample and to increase the quality of the charge difference reconstruction for neutral B decays. Another 28% of the sample

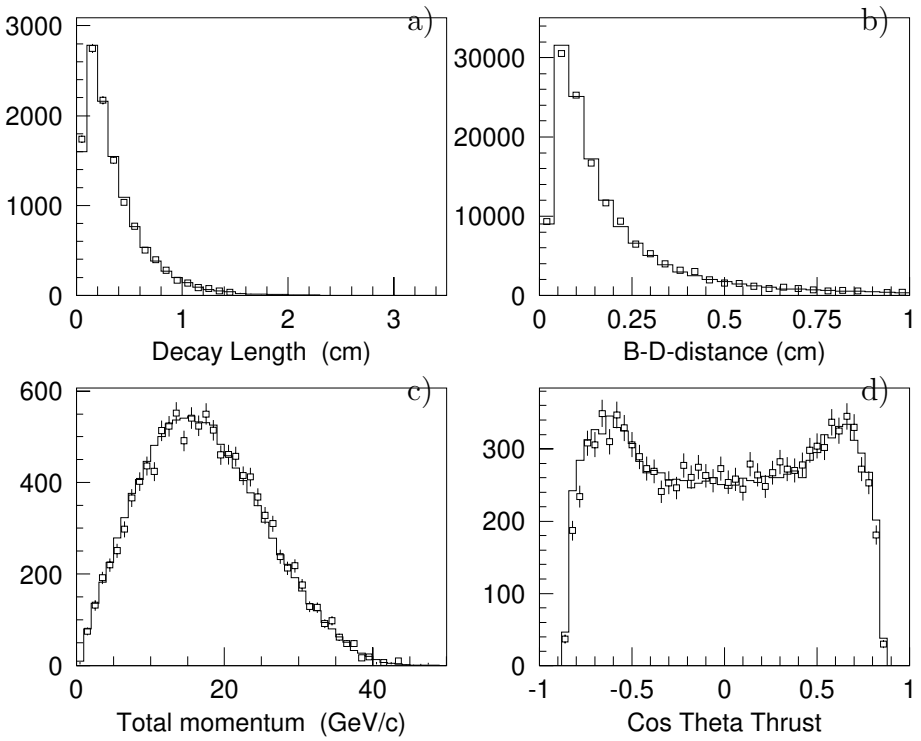


Figure 38: Distributions of a)B decay length, b)B to D distance, c)total B momentum and d)cosΘ of the thrust axis for data (points) and Monte Carlo (histograms).

is rejected in this way. As mentioned in the previous section, the vertex closer to the IP is labelled 'B'(secondary) and that further away (tertiary) is labelled 'D'. Fig. 38 demonstrates that basic distributions of the B vertex decay length, B to D distance, total B momentum and the distribution of the $\cos\theta_{thrust}$ in the SLD data are modelled well in the Monte Carlo, except for small discrepancies at large $|\cos\Theta|$.

A 'charge dipole' is defined as $\delta Q \equiv D_{BD} \times SIGN(Q_D - Q_B)$, where D_{BD} is the distance between the two vertices and Q_B (Q_D) is the charge of the B (D) vertex. Positive (negative) values of δQ tag $\overline{B^0}$ (B^0) decays. Further vertex selection cuts are:

- A minimal separation between the B and D vertex to avoid decays with

incorrectly attached secondary and tertiary tracks: $250\,\mu\text{m} < D_{BD} < 1$ cm. This cut rejects another 12% of the events.

- A mass cut on the D vertex mass is applied to restrict the reconstruction of fake D vertices: D vertex mass < 2.0 GeV/c^2(assuming all tracks are pions). Only 4% of the events are removed (each previous cut is applied in the quoted percentages of rejected events).

- A cut on events with negative decay lengths (i.e.the B vertex is reconstructed behind the IP): B vertex decay length $L_B > 0$. This cut removes only 1% of the events.

- The ghost track width (see appendix B) is required to be $< 300\,\mu$m. Very large ghost track errors allow random track attachment to the vertices. This cut only removes 0.5% of the events.

- The cosine of the angle between the straight line connecting the IP and the B vertex, and the nearest jet axis direction is required to be < 0.9. This removes another 2.5% of the events.

- The decay is rejected if any attached VXD-alone track has $p_\perp > 4$ GeV/c, since in that case the charge is not reliably reconstructed. Only very few decays (0.1%) are rejected in this way.

Applying all of the dipole final state tag cuts, a sample of 13587 decays is selected from the original sample of 99844 decays with a b-tag in the 1996-98 data. Figure 39 shows distributions of the IP, B and D vertex track multiplicities, as well as the distance and charge difference between B and D vertices in the selected sample. Good agreement between data and MC is obtained. The slight discrepancy in the

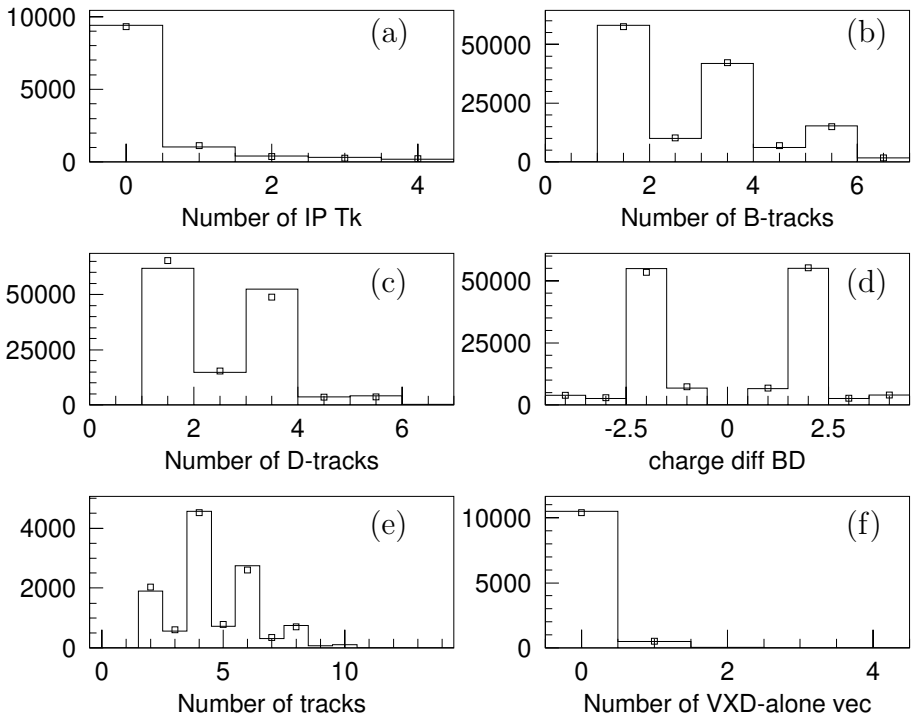


Figure 39: Distributions of (a)Number of tracks attached to the primary vertex, (b)B vertex track multiplicity, (c)D vertex track multiplicity, (d)charge difference between B and D vertices, (e)number of tracks and (f)number of VXD-alone vectors for data (points) and Monte Carlo (histograms) in the charge dipole analysis.

D vertex track multiplicity, due to a deficit in 1-prong D decays in the SLD Monte Carlo, was determined to have negligible impact on the analysis.

Fig. 40 displays the distribution of the charge dipole δQ for the data sample. It indicates the separation between b hadrons containing b or $\bar{b}$ quarks in the MC. The Charge Dipole correct tag probability, P_f, as extracted from the SLD Monte Carlo can be plotted as a function of the charge dipole (see Fig. 41). As already shown in Fig. 40, the purity of the tag increases from ∼50% at $|\delta Q|$=0 cm to ∼80% at $|\delta Q|$=0.2 cm. The track assignment improves as B and D vertices become more clearly separated (i.e., the distance between the vertices becomes much larger than the vertex resolution). Beyond $|\delta Q|$=0.2 cm no improvement of the tagging purity

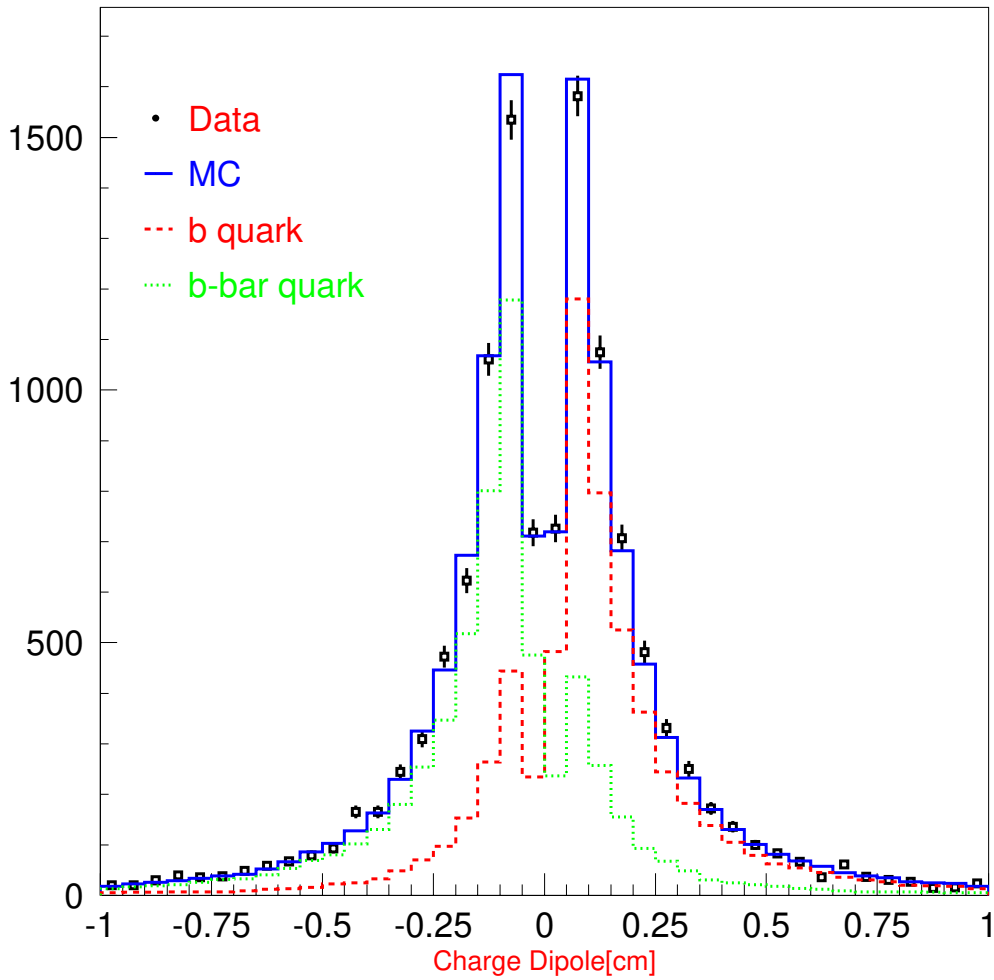


Figure 40: **Distribution of the charge dipole for data (points) and Monte Carlo (solid histogram). Also shown are the contributions from b hadrons containing a b quark (dashed histogram) or a $\bar{b}$ quark (dotted histogram).**

with increasing $|\delta Q|$ is observed. This is due to the $B \to D\overline{D}$ background which limits, independently of B and D vertex separation, the tagging purity. Double charm events will be discussed in detail in the next section. The function used to fit the charge dipole correct tag probability is described using two parameters a and b:

$$f(|\delta Q|) = 0.5 + a\ \frac{(1 - e^{\ b|\delta Q|})}{(1 + e^{\ b|\delta Q|})}. \tag{4.4}$$

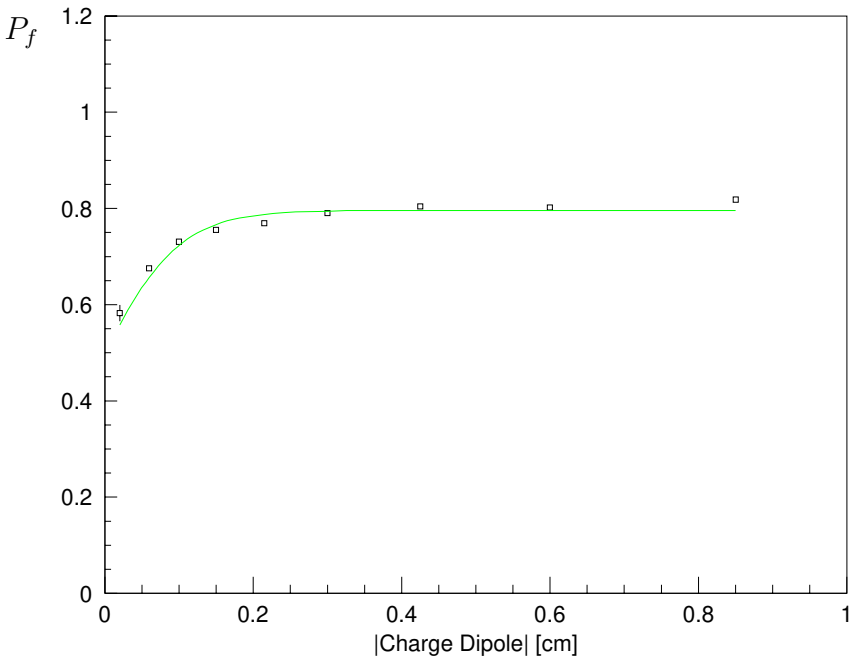


Figure 41: **The Charge Dipole correct tag probability for B_s^0 decays as a function of the charge dipole magnitude, as extracted from the Monte Carlo. The solid line is the result of a fit to the points.**

Note that all of the Charge Dipole cuts also enhance the fraction of B_d^0 decays in the sample. In order to suppress the dominant contribution of B_d^0 mesons, events with inclusively reconstructed $D^\star$ decays are discarded, as one expects a

large branching ratio for $B_d^0 \to D^{\star -} X$. The corresponding branching ratio for $B_s^0 \to D^{\star -} X$ is $< 1\%$.

The $D^{\star +} \to D^0 \pi^+$ decays can be reconstructed inclusively using slow pion candidates from either the secondary or the tertiary vertex. Fig. 42 shows the mass difference between D vertices with an associated slow pion and D vertices without one, as simulated in the Monte Carlo. 4 different cases are shown: neutral (charged) B decays and neutral (charged) B decays with a D mass cut at m_D >1 GeV. A $D^\star$ peak can be seen in the mass distribution of neutral B decays at 0.14 GeV $< M(D^\star) - M(D) <$ 0.16 GeV. Fig. 43 shows the corresponding distributions in the 1996-98 SLD data. Requiring there are no slow pion candidates with $M(D^\star) - M(D)$ <0.16 GeV and $M(D)$ >1 GeV, increases the relative fraction of B_s^0 events in the dipole event sample by 4%. Only 2.5% B_s^0 decays are removed, mostly from $B_s^0 \to \overline{D^0} X$ decays. The cut removes another 4.6% of the charge dipole data sample, reducing its size to 12960 decays.

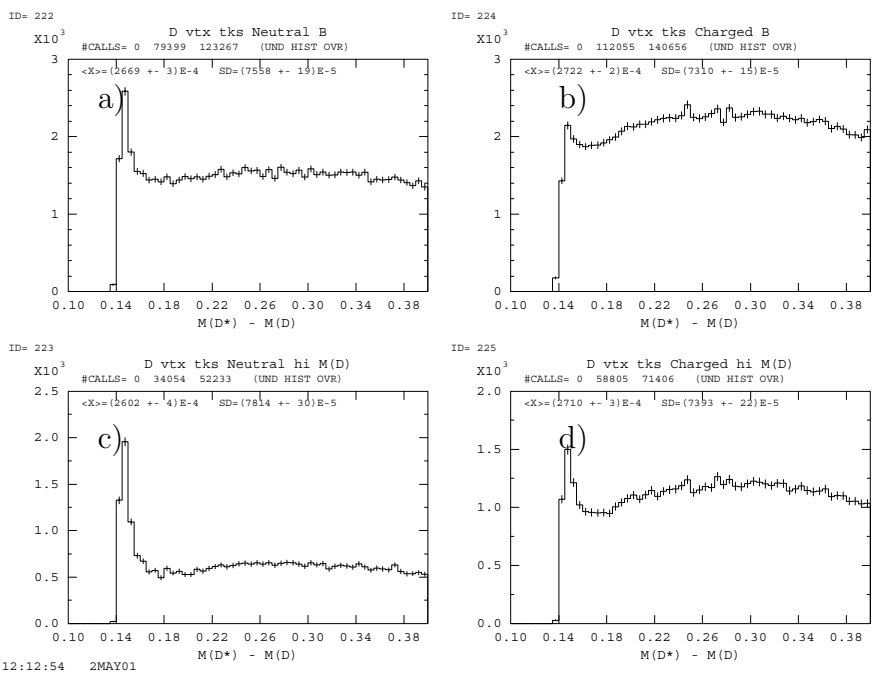


Figure 42: **Monte Carlo distributions of $M(D^\star)-M(D)$ for 4 different cases: a)neutral B decays, b)charged B decays, c)neutral B decays with m_D >1 GeV and d)charged B decays with m_D >1 GeV.**

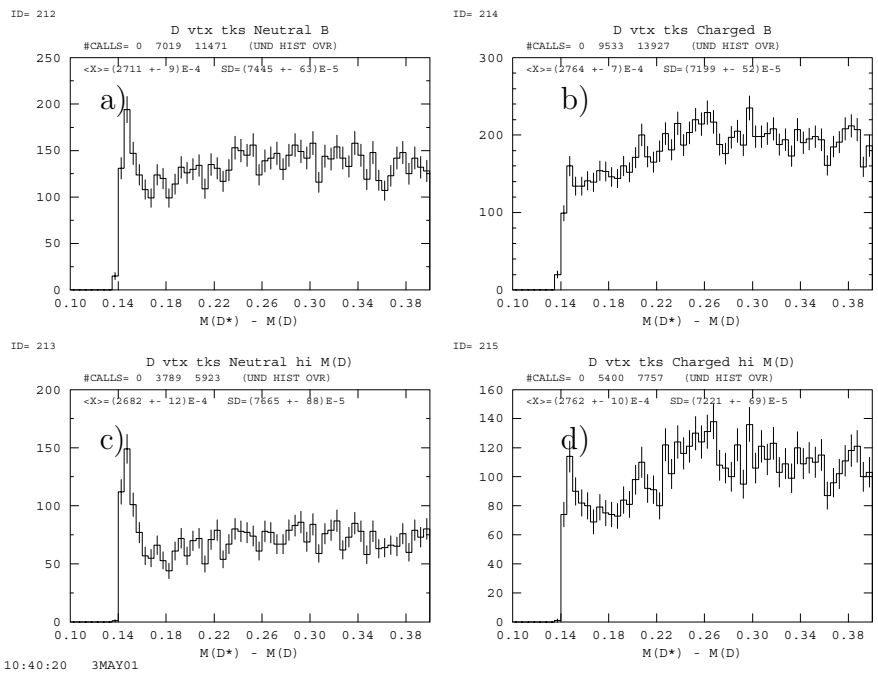


Figure 43: **The distribution of $M(D^\star) - M(D)$ for 4 different cases in the 96-98 SLD data set: a)neutral B decays, b)charged B decays, c)neutral B decays with m_D >1 GeV and d)charged B decays with m_D >1 GeV.**

4.2.2 Characterizing the event sample

The charge dipole event sample is composed of many different decays because of its inclusive nature. It is important to understand which decays are contained in the event sample to be able to describe it accurately. In the analysis, the important groups of events are identified and their behavior is parametrized. First, the fractions of different b hadrons in the event sample are determined from the Monte Carlo. The dipole cuts boost the B_s^0 fraction from its assumed production value of 9.8% to 16% due to the selection of neutral events and the cut on inclusively reconstructed $D^\star$ decays. All b hadron fractions are summarized in table 8 [76]. A detailed discussion of b hadron production fractions can be found in appendix A1. Note that the largest fraction of hadrons in the sample are B_d^0 mesons. It will

Table 8: b hadron fractions in the Monte Carlo at production and after the dipole selection.

Hadron Type	production fraction	fraction in dipole sample
B_s^0	9.8%	16.0%
B_d^0	39.9%	58.3%
B^+	39.9%	18.3%
Baryons	10.3%	7.4%

be important later on to demonstrate that the dipole tag is indeed sensitive to the oscillations of the small fraction of B_s^0 mesons.

The fraction of the different b hadron species in the event sample is not constant over the whole range of acceptance. Beyond $|\cos\Theta| \approx 0.7$ the fraction of neutral b hadrons in the selected event sample decreases while the background of misreconstructed B^+ mesons increases. This is due to the fact that at the edge of the detector acceptance tracking inefficiencies are more likely. The functions used

to scale the average b hadron fractions are the results of fits to the distributions shown in Fig. 44. They are parametrized in terms of three parameters a, b and c:

$$f_b = a + \frac{b}{c}\ \mathbf{e}^{-|\cos\Theta|/c}. \tag{4.5}$$

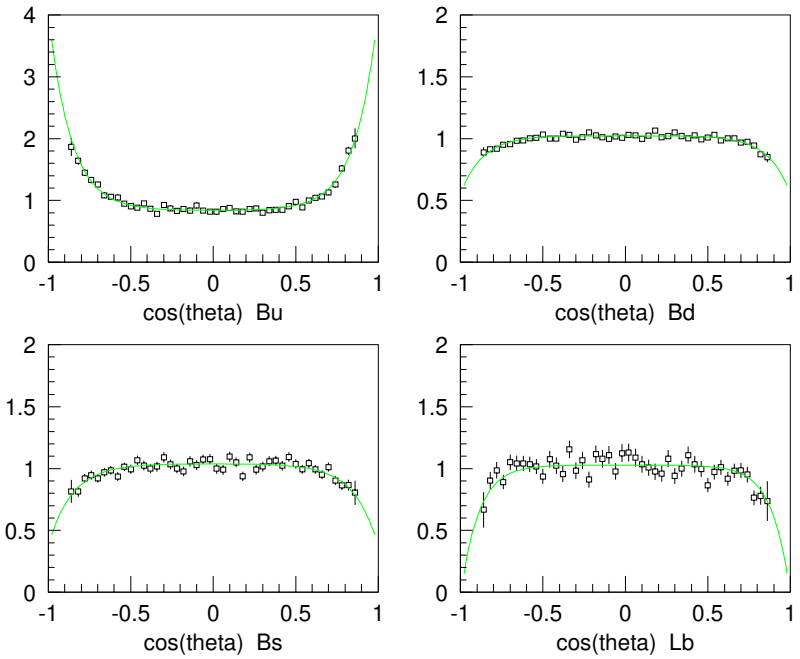


Figure 44: The distribution of the relative b hadron fractions as a function of $\cos\Theta$. The functions are the result of fits. They are used to scale the average b hadron fractions.

As discussed above, the dipole selection is optimized to tag events with a $B \to D$ cascade structure, where the total charge of all tracks in the event is required to be zero, and the individual charges of B and D vertices are different from each other. This is tailored to select events of the type $B_d^0 \to D^- X$ or $B_s^0 \to D_s^- X$. These decays have a very high tagging purity (or low mistag rate). Other decay modes, like $B \to \overline{D^0} X$, which are expected not to have a charged dipole, are most likely to have been mistagged, due to tracking inefficiencies or incorrectly attached tracks.

The decay mode with the worst tagging purity corresponds to $B \to D\overline{D}$ decays. Very often one of the D vertices is mistaken for a B, and the charge assignment is random. Double charm events therefore have no tagging power.

Table 9: **B_s^0 decay mode fractions and average mistag probabilities for the charge dipole analysis.**

Decay mode	fraction	average mistag probability
$B_s^0 \to \overline{D^0}X$	10.2%	(38.2±0.8)%
$B_s^0 \to D^-X$	6.1%	(7.0±0.6)%
$B_s^0 \to D_sX$	59.5%	(9.1±0.2)%
$B_s^0 \to$charmed baryon X	2.9%	(30.3±0.2)%
$B_s^0 \to D\overline{D}X$	21.2%	(46.9±0.5)%

Table 10: **B_d^0 decay mode fractions and average mistag probabilities for the charge dipole analysis.**

Decay mode	fraction	average mistag probability
$B_d^0 \to \overline{D^0}X$	32.0%	(26.6±0.2)%
$B_d^0 \to D^-X$	41.7%	(6.7±0.1)%
$B_d^0 \to D_sX$	1.7%	(8.6±0.7)%
$B_d^0 \to$charmed baryon X	3.7%	(21.3±0.7)%
$B_d^0 \to D\overline{D}X$	20.9%	(43.6±0.3)%

Because of these differences the mistag probability is parameterized separately for B_d^0 and B_s^0 decays into five different final states: D^0X, D^+X, D_sX, charmed baryon X, and $D\overline{D}X$ (this last category also incorporates charmonium production, i.e. it includes all $b \to c\overline{c}s$ decays). For example, the mistag probability is only 9% for $B_s^0 \to D_sX$ decays but 47% for $B_s^0 \to D\overline{D}X$ decays. Similarly, the mistag probability of B_d^0 decays is only 7% for the $B_d^0 \to D^-X$ decay channel but 44% for $B_d^0 \to D\overline{D}X$ decays. The mistag probabilities are extracted from the MC simulation. Tables 9 and 10 show all decay modes of the B_s^0 and B_d^0 mesons that

are modelled, as well as their respective fractions and average mistag probabilities extracted from the Monte Carlo. Averaged over all decay modes, the charge dipole mistag probabilities for B_s^0 and B_d^0 are quite similar: $(20.6 \pm 0.2)\%$ for the B_s^0 and $(21.3 \pm 0.2)\%$ for the B_d^0 meson. The largest decay mode fractions are $B_s^0 \to D_s X$ and $B_d^0 \to D^- X$ with similar average mistag probabilities.

The tagging power of charged B mesons and baryons does not need to be parametrized in such great detail, because these background decays do not contribute to the mixing signal. The average mistag probabilities are $(29.1 \pm 0.3)\%$ for $B^\pm$ and $(30.3 \pm 0.4)\%$ for baryons.

The correct tag probability of the decay modes could also depend on the B decay length L, because a B decaying very close to the IP is in general harder to reconstruct than one that decays far away, as IP tracks may be included in the B or even the D vertex. Therefore the correct tag probability is parameterized as a function of decay length, as shown in Fig. 45 for the low decay length range of 0-2 mm. Note that most decay modes exhibit no decay length dependence, indicating excellent tagging strength even at distances $< 200\mu$m.

Another effect that needs to be accounted for is the dependence of the decay mode fractions on the P_t-corrected B vertex mass. For example, the fraction of $B_s^0 \to D_s X$ decays increases with the vertex mass, whereas the fraction of B_s^0 decays into $D\overline{D}X$ final states decreases. This is due to the fact that missing or incorrectly attached tracks result in a low B vertex mass, so that a mistake in the B vertex reconstruction is more likely. Figure 46 shows the fractions of the five decay modes as a function of the B reconstructed vertex mass. The function is the result of a third degree polynomial fit to the decay mode fractions between M >2 GeV and M <6 GeV. For M <2 GeV and M >6 GeV the fractions are kept constant.

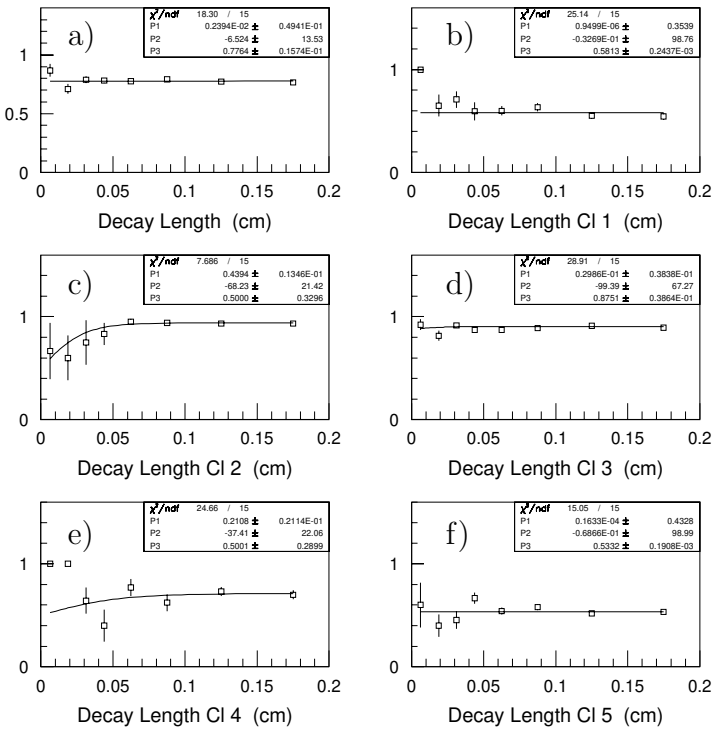


Figure 45: **Charge dipole correct tag probability as a function of reconstructed decay length in simulated B_s^0 decays. a)all decay modes, b)$B_s^0 \to \overline{D^0}X$, c)$B_s^0 \to D^-X$, d)$B_s^0 \to D_sX$, e)$B_s^0 \to c$ baryons, f)$B_s^0 \to D\overline{D}$. The functions are the fit results used to parametrize the charge dipole correct tag probability as a function of decay length.**

4.3 Proper time reconstruction

The proper time of the B decay is determined from the reconstructed B decay length and the relativistic boost, $\gamma\beta = \frac{p_B}{m_B}$, where the B mass is assumed to be 5.3 GeV and the B momentum p_B is calculated from the reconstructed B energy. The decay proper time is then computed from $t_{rec} = \frac{L}{\beta\gamma c}$.

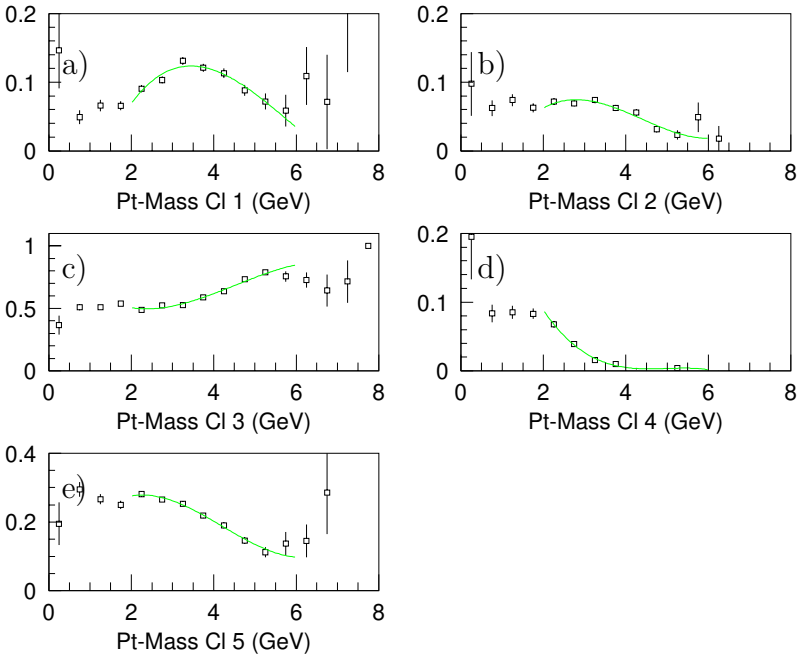


Figure 46: **The decay mode fractions as a function of the B_s^0 vertex mass. a)$B_s^0 \to \overline{D^0}X$, b)$B_s^0 \to D^-X$, c)$B_s^0 \to D_sX$, d)$B_s^0 \to$** baryons**, e)$B_s^0 \to D\overline{D}$.**

4.3.1 Decay Length Reconstruction

The general procedure to locate the B decay vertex and thereby the B decay length has been described in section 4.2.1. Some aspects will be pointed out here in more detail.

After the initial B vertex position has been located using the ghost track algorithm, an additional piece of information using a virtual D track can be used to increase its precision, if at least two tracks were attached to the tertiary vertex. In this procedure all tertiary tracks are fit to a common vertex, and a resultant D track with finite width is created. Its direction is given by the net momentum of the tertiary tracks. The computation of track width and direction have been

done using the Kalman Fitter method, described in [86]. The angular uncertainty of the resulting D track is scaled with the mass of the D vertex, giving the D track a higher angular uncertainty the more the reconstructed D mass differs from the true D mass, thereby accounting for misreconstruction of the D track direction due to missing tracks. This virtual D track is then included in a vertex fit with all of the tracks associated with the B vertex which results in a B vertex location with associated error ellipse. The advantage of creating such a D track is that a realistic error is assigned to 1-prong B vertices. Also, the vertex resolution improves with the addition of the D track.

For events with only one track in the tertiary vertex, it is not useful to create such a virtual track as the D direction is poorly determined. If the secondary vertex has more than one track associated with it, the loss of the virtual D track results in slightly worse decay length resolution. If both B and D vertices have only one track associated with them, large uncertainties in vertex finding and proper time determination are to be expected. As discussed before, the vertex finding procedure produces a ghost track along the B decay chain, and in principle the B decay length can be reconstructed from the intersection point of ghost track and B track, $L_{ghost+B}$. The ghost track and the secondary track generally have a rather low angle between them, producing an elongated error ellipse along the decay direction. A general bias towards lower decay lengths is observed. Another possibility is to use the intersection point of the B track and the D track, L_{B+D}, which suffers from the opposite problem, a bias toward larger decaylength. The method chosen here is to average L_{B+D} and $L_{ghost+B}$, which produces a reconstructed decay length closest to the true decay length.

A discussion of the decay length resolution follows in section 5.3.1.

4.3.2 Boost Reconstruction

Two different methods of determining the boost are combined to achieve the optimal boost reconstruction for selected dipole events. These two methods are the mass method [87], where the B energy is calculated using the known invariant mass of the B vertex, and the maximum missing mass method [88] which, in a certain mass limit, uses additional kinematic information to constrain neutral and ν energy.

Method 1: The B energy E_B, is the sum of neutrino energy E_ν, charged track energy E_{ch} and neutral particle energy E_{neut}: $E_B = E_{ch} + E_\nu + E_{neut}$. E_{ch} is the energy of all charged tracks associated with the secondary B decay chain assuming pion mass, and E_{neut} is determined from electromagnetic calorimeter clusters that can be associated with the B decay but are not associated with any charged track. To test if these neutral clusters come from the decay of the B rather than from fragmentation tracks, the clusters are tested for compatibility with the B invariant mass. Each neutral cluster is added to the secondary vertex while keeping the direction of the vertex momentum unchanged. The process starts with the cluster closest to the secondary vertex flight direction. The cluster also has to pass a minimum energy cut of E $>$ 3 GeV. The invariant mass of the vertex is then recalculated, and if the mass of the new combination is closer to m_B=5.3 GeV than before, the cluster is considered to come from a neutral particle originating from the B decay. This procedure continues as long as the B mass remains below 5.3 GeV. Any other clusters are considered unassociated with the B decay.

To determine the energy of the escaped neutrinos, E_ν, the total energy in the jet containing the B, E_{jet}, is considered: $E_\nu = E_{jet} - E_{visible}$. E_{jet} is calculated from the Z mass. For events with two jets a $Z \rightarrow b\overline{b}$ event is assumed and E_{jet} is

fixed to $m_Z/2$. For a three jet event, the energy in each jet is calculated using the jet directions:

$$E_{jet}^{1,2,3} = \frac{sin\Theta_{23,13,12}}{sin\Theta_{12} + sin\Theta_{13} + sin\Theta_{23}} m_Z, \tag{4.6}$$

where Θ_{ij} is the angle between jets i and j. In the few cases of four jets the event is divided into two hemispheres and treated as a two jet event. The visible energy $E_{visible}$ is simply all of the energy which is measured in the calorimeter for the event hemisphere.

Method 2: This method determines E_ν and E_{neut} by calculating the missing invariant mass, M_0, and the missing longitudinal momentum, P_{0l}^2, to reconstruct the energy of all particles not associated with the vertex (neutrinos and other neutral particles):

$$E_0 = E_\nu + E_{neut} = M_0^2 + p_{0t}^2 + p_{0l}^2. \tag{4.7}$$

The transverse neutral momentum p_{0t} is identical to the measured total charged track momentum transverse to the B flight direction (by momentum conservation). One kinematic constraint on the two other quantities comes from the fact that $m_B^2 = E_B^2 - p_B^2$, where the total momentum of the secondary vertex is $p_B = p_{chl} + p_{0l}$, where p_{chl} is the longitudinal momentum of the vertex-associated charged tracks (the component along the vertex flight direction). An additional constraint comes from the mass of the B vertex:

$$\sqrt{M_{ch}^2 + P_t^2} + \sqrt{M_0^2 + P_t^2} \leq M_B. \tag{4.8}$$

The equality holds in the limit where both p_{0l} and p_{chl} vanish in the B rest frame. It sets an upper bound on the missing mass:

$$M_{0max}^2 = m_B^2 - 2M_B\sqrt{M_{ch}^2 + P_t^2} + M_{ch}^2. \tag{4.9}$$

It appears from MC studies that for small M^2_{0max}, $M^2_0 \approx M^2_{0max}$. With the transverse charged mass $M^2_{ch\perp} = M^2_{ch} + P^2_t$ one can then calculate P_{0l}:

$$P_{0l} = \frac{m_B - M_{ch\perp}}{M_{ch\perp}} P_{chl}. \tag{4.10}$$

Finally, using Eq.4.7, $E_0 = E_\nu + E_{neut}$ can be computed and, adding the energy of all charged tracks, E_B is obtained.

Methods 1 and 2 are combined linearly in five different bins of reconstructed energy such that the optimal B energy determination for dipole selected events is chosen. The B boost is then determined by

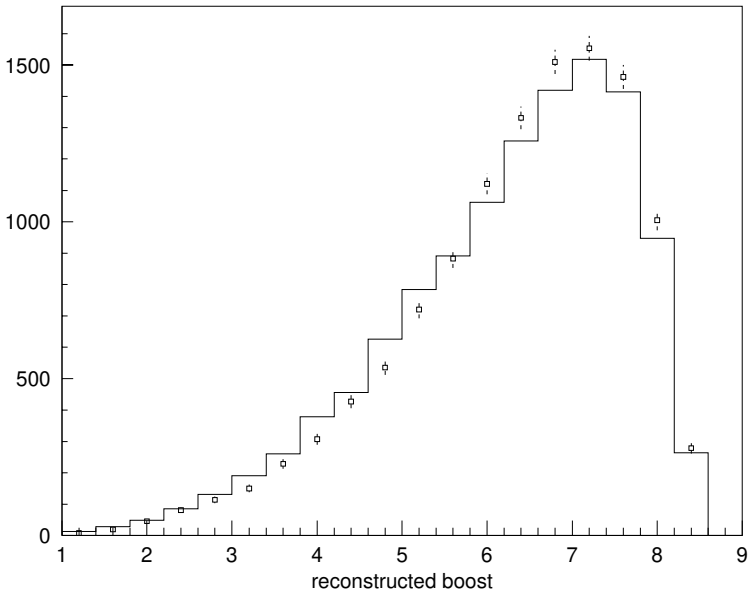


Figure 47: The reconstructed boost of events selected in the dipole analysis. The histogram denotes the Monte Carlo distribution, and the points show the reconstructed boost in the data.

$$\beta\gamma = \frac{p_B}{m_B} = \frac{\sqrt{E^2_B - m^2_B}}{m_B}, \tag{4.11}$$

where m_B is assumed to be 5.3 GeV. The distribution of the reconstructed boost is shown in Fig. 47 for data and Monte Carlo simulation. Note that the Monte Carlo boost distribution does not reflect the measured information about the b quark fragmentation function [88], so that the small disagreement between data and Monte Carlo observed in the plot is expected. The reconstructed proper time is shown in Fig. 48. For a discussion of the boost resolution see section 5.3.2.

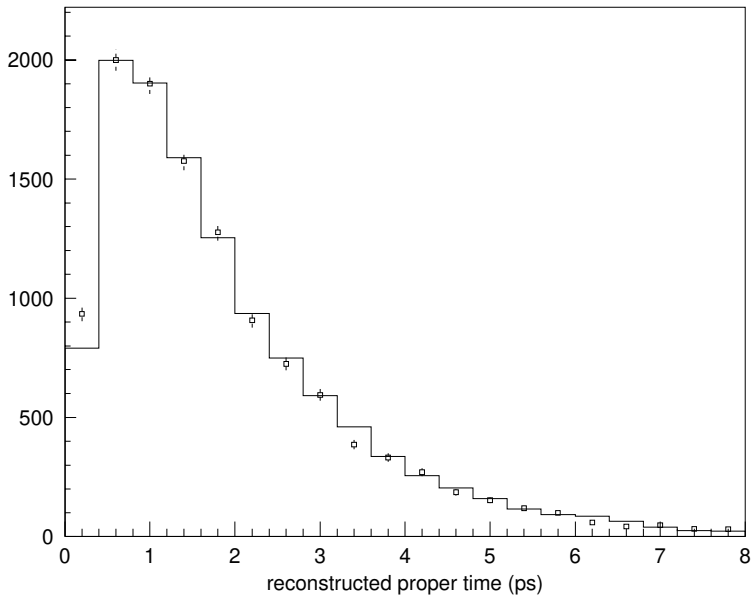


Figure 48: **The reconstructed proper time of events selected in the charge dipole analysis. The histogram denotes the Monte Carlo distribution, and data points show the reconstructed proper time in the data.**

4.4 Mixing Tag

It is now possible to tag individual events as mixed or unmixed based on the initial and final state tags: decays are tagged as mixed or unmixed if the product

$(P_i - 0.5) \times (P_f - 0.5)$ is smaller or greater than 0, respectively.

4.4.1 Checking the Dipole Tagging Power

It is important to check the dipole tagging power directly with data. Such a test is provided by the polarized forward-backward asymmetry shown in Fig. 49. This figure shows good agreement between data and Monte Carlo indicating that the tagging power is adequately modelled by the simulation. The asymmetry is

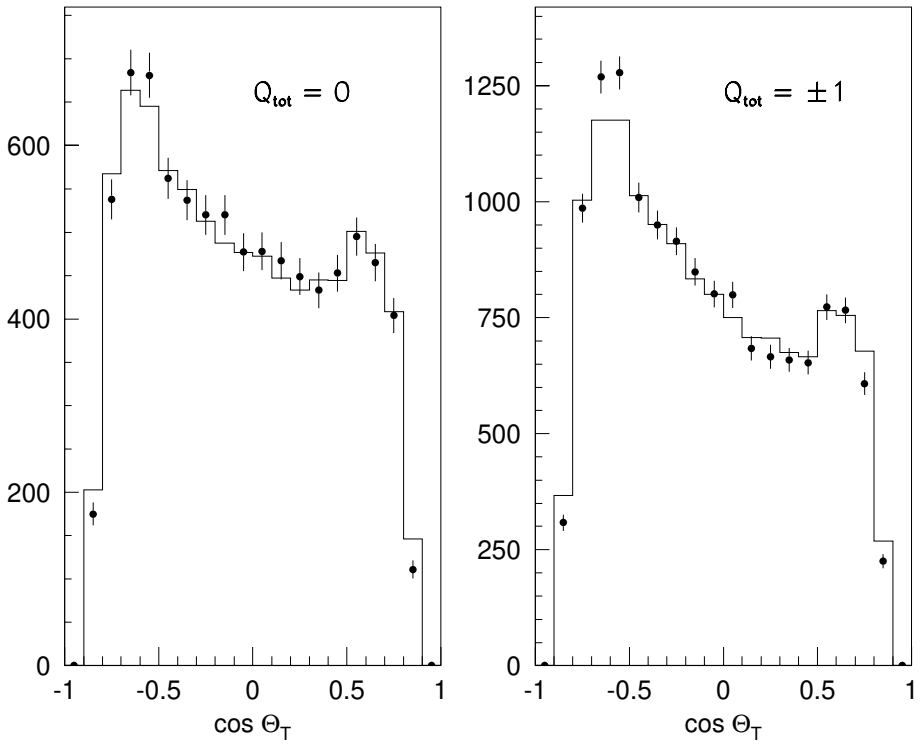


Figure 49: Distributions of $\cos\theta$ for the thrust axis direction signed by the product $(\delta Q \times P_e)$ for data (points) and Monte Carlo (histograms) in subsamples with $Q_{tot} = 0$ and $Q_{tot} = \pm 1$ for the charge dipole analysis.

diluted by both initial and final state mistags and by B^0–$\overline{B^0}$ mixing. The dilution due to mixing can be reduced by selecting vertices with total charge $Q_{tot} = \pm 1$, in which case a stronger asymmetry is observed.

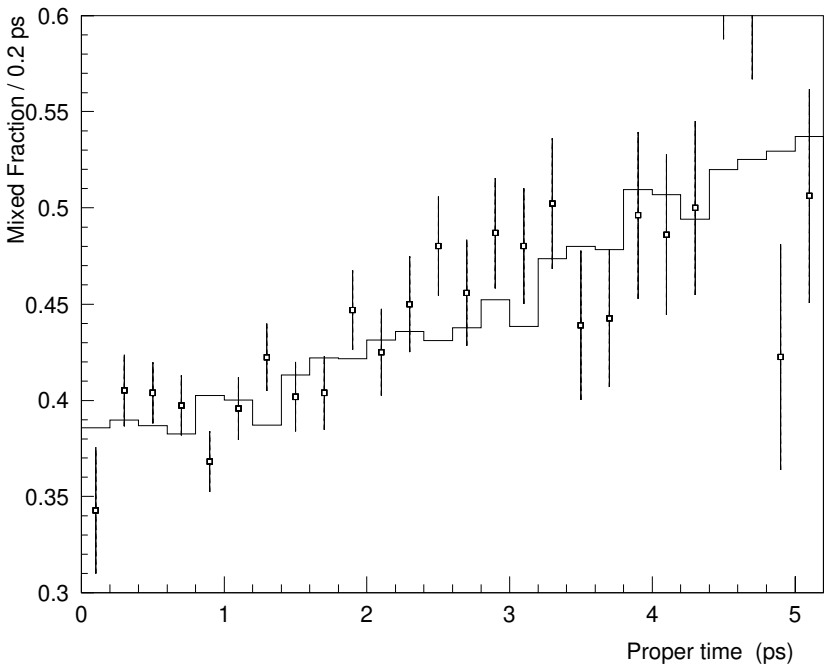


Figure 50: The fraction of events tagged as mixed in the dipole sample, as a function of the reconstructed proper time. The histogram is taken from the MC, the data points are extracted from the 96-98 data.

Another way of checking the tagging power is to compare the mixed fraction, which is the number of events tagged as mixed divided by the total number of events, in data and Monte Carlo. Fig.50 is a plot of the mixed fraction as a function of reconstructed proper time. The MC is plotted as a histogram, and the data as points. The mixed fraction is expected to rise as a function of time due to the slow B_d^0 mixing. The rapid B_s^0 oscillations are invisible on a plot of this type. A typical sign for an overestimated MC tagging power would be a higher mixed fraction in the data than the MC at low proper time, which is not observed.

Another interesting comparison between the tagging power in the data and the Monte Carlo is shown in Fig.51. It shows the mixed fraction as a function of

the computed mistag rate, in 4 bins of reconstructed proper time. The last plot is integrated over all time bins. In the first bin of proper time, at low mistag probability the mixed fraction is expected to be low, as the average number of mixed events due to B_d^0 mixing is small at small reconstructed proper times. The mixed fraction is expected to rise with the mistag rate, eventually reaching a mixed fraction of 50% (no mixing information) with a mistag fraction of 50%. With increasing proper time the contribution of B_d^0 mixing becomes larger. The general agreement of data and Monte Carlo mixed fraction shows that the Monte Carlo is describing the tagging power correctly in all bins of proper time.

All plots show good agreement between data and MC. The tagging ability is well-modelled in the MC and the selected events are sensitive to mixing.

4.5 A Signal to Noise Estimate

It is interesting to ask how well one can hope to measure a B_s^0–$\overline{B_s^0}$ mixing signal with the charge dipole event sample (or any other experiment), given the set of experimental parameters associated with it.

It is clear that the signal to noise ratio, or 'sensitivity', will depend on the B_s^0 fraction in the event sample, on the number of events selected, on the mistag rate and on the time resolution in the detector. The B_s^0 fraction $f_{B_s^0}$ and mistag rate w are multiplicative factors that effectively reduce the amplitude of the mixing signal.

Starting with $P_m(t)$, the probability for a neutral B meson to mix into its antiparticle before decaying, as it was obtained in chapter 1 from quantum mechanics, is

$$P_m(t) = \Gamma \frac{e^{-\Gamma t}}{2}[1 - \cos \Delta m t], \tag{4.12}$$

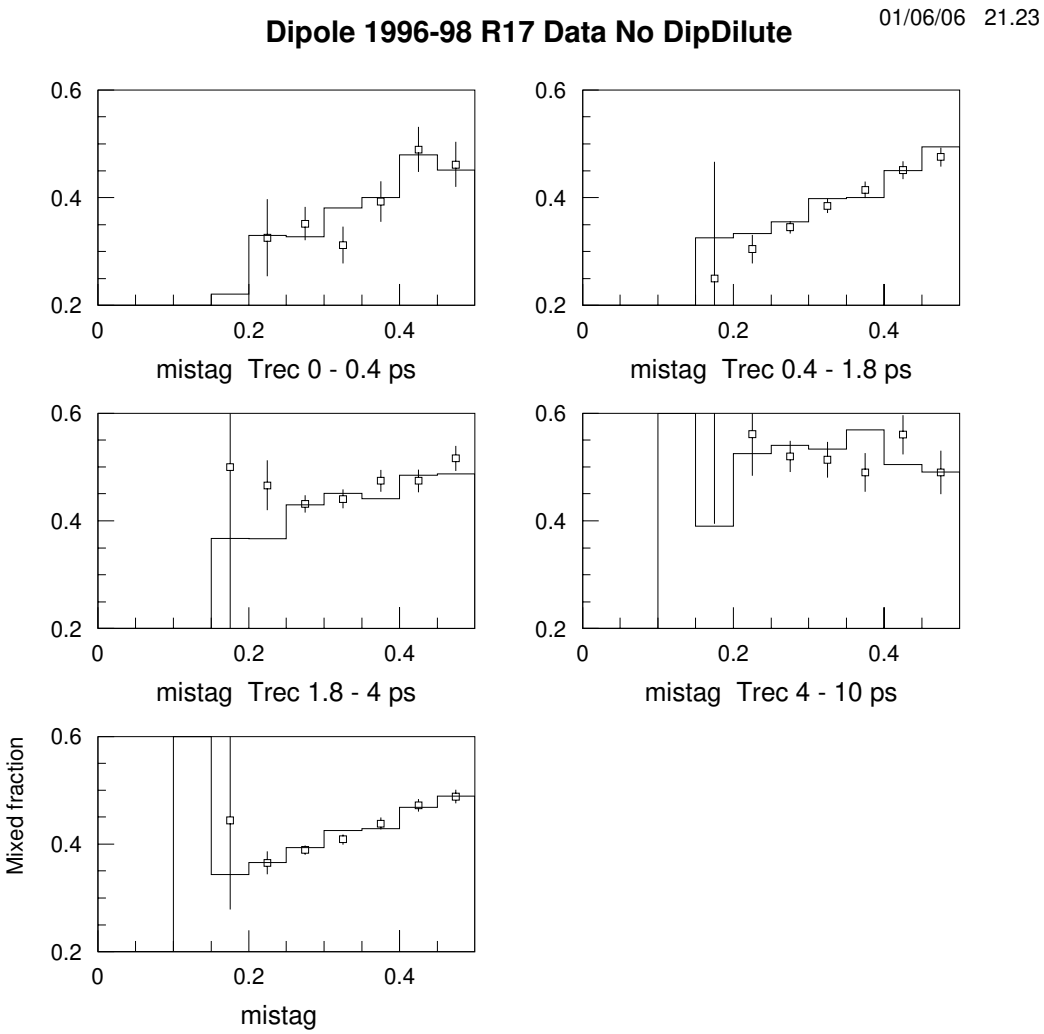


Figure 51: **The fraction of events tagged as mixed in the dipole sample, as a function of the probability that an event was tagged incorrectly, shown in 4 bins of reconstructed proper time. The last plot is integrated over all time bins. The histograms are taken from the MC, the data points are extracted from the 96-98 data.**

where Γ is the decay width of the B meson, t is the true proper time of the B decay and Δm the mass difference between the two mass eigenstates. Similarly, the probability $P_u(t)$, that a neutral B meson does not mix and decays as it was produced, is

$$P_u(t) = \Gamma \frac{e^{-\Gamma t}}{2}[1 + \cos \Delta m t]. \tag{4.13}$$

Taking into account the experimental factors w and $f_{B_s^0}$, P_m becomes

$$P_m^{exp}(t) = \Gamma f_{B_s^0} \frac{e^{-\Gamma t}}{2}[1 - (1 - 2w) \cos \Delta m t]. \tag{4.14}$$

The factor $(1 - 2w)$ is derived by considering that either a mixed state was tagged correctly or an unmixed state was mistagged.

To quantify the size of a periodic signal it is of advantage to perform a Fourier analysis of the frequency spectrum [89]. The expected Fourier transform (FT) of the mixing signal in a data sample of n events is,

$$FT(\Delta m) \approx \frac{n f_{B_s^0}(1 - 2w)}{2} \left(\frac{\Gamma^2}{\Gamma^2 + (\Delta m - \Delta m^{true})^2} + \frac{\Gamma^2}{\Gamma^2 + (\Delta m + \Delta m^{true})^2} \right) + n(1 - f_{B_s^0})(1 - 2w) \frac{\Gamma^2}{\Gamma^2 + \Delta m^2}. \tag{4.15}$$

The second term of the function can be neglected because $\Delta m > \Gamma$. It can be assumed that the last term (from the non-mixing B) can be subtracted in the final analysis. Therefore, we expect a peak in the frequency spectrum at $\Delta m = \Delta m^{true}$ with a Breit-Wigner shape of width 2Γ:

$$FT(\Delta m) \approx \frac{n f_{B_s^0}(1 - 2w)}{2} \left(\frac{\Gamma^2}{\Gamma^2 + (\Delta m - \Delta m^{true})^2} \right). \tag{4.16}$$

To estimate the effect of the time resolution of the detector, σ_t, each measured time t is smeared with Gaussian errors, given the resolution σ_t:

$$P_m^{exp}(t) \rightarrow \hat{P}_m^{exp}(t) = \int_0^\infty P_m^{exp}(t) \frac{1}{\sqrt{2\pi}\sigma_t} e^{-\frac{(t-t')}{2\sigma_t^2}} dt'. \tag{4.17}$$

To simulate this effect in the FT analysis one can use the convolution theorem. It states that the FT of the convolution of two functions is proportional to the product of the individual FTs of these functions:

$$\mathrm{FT}[\hat{\mathrm{P}}_{\mathrm{m}}^{\mathrm{exp}}(\mathrm{t})] = \sqrt{2\pi}\mathrm{FT}[\mathrm{P}_{\mathrm{m}}^{\mathrm{exp}}(\mathrm{t})] * \mathrm{FT}[\mathrm{g}(\sigma_{\mathrm{t}})], \tag{4.18}$$

where $g(\sigma_t)$ is a Gaussian function of width σ_t. The proper time is computed from the measured decay length L and the boost $\gamma\beta$:

$$t = \frac{L}{\gamma\beta c}. \tag{4.19}$$

Therefore, the proper time resolution σ_t is

$$\sigma_t = \sqrt{\left(\frac{\sigma_L}{\gamma\beta c}\right)^2 + \left(\frac{\sigma_{\gamma\beta} t}{\gamma\beta}\right)^2}. \tag{4.20}$$

The FT of a Gaussian is simply another Gaussian, but the calculation is complicated by the time dependence in the boost component of the resolution. The approximate result is given by

$$\mathrm{FT}[\mathrm{g}(\sigma_{\mathrm{t}})] = \mathrm{e}^{-\frac{\sigma_{\mathrm{L}}^2 \Delta \mathrm{m}^2}{2}} \mathrm{r}(\Delta \mathrm{m}^2, \sigma_{\mathrm{p}}). \tag{4.21}$$

The first term is a Gaussian of width $1/\sigma_L$ and describes the damping of the FT peak due to the decay length resolution σ_L. The expected peak amplitude decreases with Δm_s.

The second term accounts for the boost resolution and can be approximated by

$$r(\Delta m^2, \sigma_p) = \sqrt{\pi} Y e^{Y^2} ERFC(Y) \tag{4.22}$$

with

$$Y = \frac{1}{\sqrt{2}} \frac{\Gamma}{\sigma_p \Delta m}, \qquad ERFC(Y) = \frac{2}{\sqrt{\pi}} \int_Y^\infty e^{-t^2} dt. \tag{4.23}$$

This term also results in a reduction of the FT peak as the frequency increases.

In conclusion, a measurement of the Fourier amplitude becomes more difficult as the frequency increases, as faster oscillations are harder to resolve given a limited time resolution in the detector. The expected size of the signal, or FT peak at the true frequency, is

$$FT(\Delta m_s) = \frac{n}{2} f_{B_s^0}(1-2w) e^{-\frac{\sigma_L^2 \Delta m^2}{2}} r(\Delta m^2, \sigma_p). \tag{4.24}$$

The noise is given by $\sqrt{\frac{n}{2}}$. Therefore, the sensitivity S is

$$S = \sqrt{\frac{n}{2}} f_{B_s^0}(1-2w) e^{-\frac{\sigma_L^2 \Delta m^2}{2}} r(\Delta m^2, \sigma_p). \tag{4.25}$$

This is an important equation. It shows that the sensitivity depends linearly on the square root of the number of events, mistag probability and B_s^0 fraction, whereas the proper time resolution appears in the exponent. This means that at high oscillation frequencies the proper time resolution becomes the limiting factor of a measurement. As an example, Fig. 52 shows the difference in observable mixing signal in two simulated experiments, both with the typical LEP/SLD relative boost resolution of 10%, a mistag rate of 25% and Δm_s=20 ps^{-1}, but different decay length resolutions: $\sigma_L = 60\mu$m (SLD-like) and $\sigma_L = 200\mu$m (LEP-like). At high frequencies the SLD analysis will dominate due to the high resolution of the vertex detector, whereas at the low Δm_s end, the LEP experiments dominate the measurement due to the advantage in statistics. Both experiments suffer from severe damping of the oscillating signal at proper times $>1.5\text{ps}^{-1}$ due to the boost resolution. The interplay between boost and decaylength resolution will be discussed in detail later on.

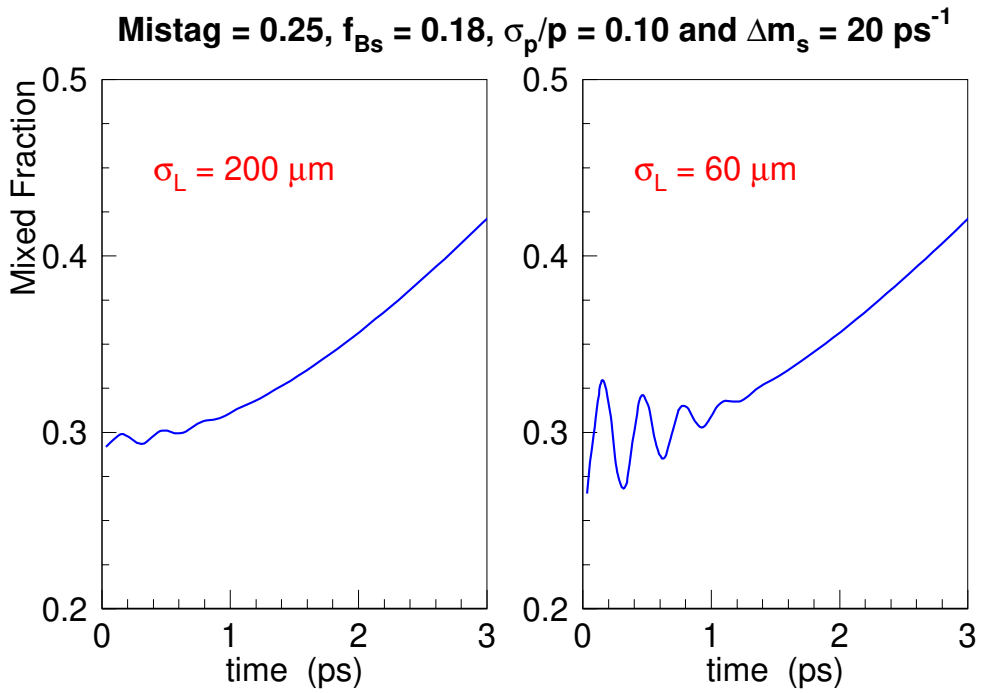


Figure 52: **A plot of the analytical function of the mixed fraction assuming a typical set of experimental parameters and $\Delta m_s = 20ps^{-1}$, for $\sigma_L = 60\mu$m (SLD-like) and $\sigma_L = 200\mu$m (LEP-like).**

CHAPTER 5

Maximum Likelihood Analysis

In this chapter, the analysis procedure used to extract limits on the B_s^0 oscillation frequency Δm_s from the full 96-98 SLD data set will be discussed. An amplitude fit to the frequency spectrum of the mixing signal is introduced.

5.1 B Hadron Physics Functions

The likelihood function for $b\bar{b}$ events is composed of terms describing the probability for observing a mixed or unmixed event as a function of proper time. B^+, B_d^0, B_s^0, b baryons and background events have to be treated separately.

Based on the experimental probability density function $\hat{P}_m^{exp}(t)$ introduced in the last chapter, an experimental function F that includes the effect of B^+, B_d^0, B_s^0, b baryons or background events can be derived. A discussion of proper time resolution, selection efficiency and other effects will follow in the next sections. The function F_{mix} describing the probability for a decay with proper time t to be tagged

as mixed is:

$$
\begin{aligned}
F_{mix} = & \Gamma_{B^+} f_{B^+} \frac{e^{-\Gamma_{B^+} t}}{2} w_u \\
& + \Gamma_{B^0_d} f_{B^0_d} \frac{e^{-\Gamma_{B^0_d} t}}{2} [(1 - (1 - 2w_d) \cos \Delta m_d t)] \\
& + \Gamma_{B^0_s} f_{B^0_s} \frac{e^{-\Gamma_{B^0_s} t}}{2} [(1 - (1 - 2w_s) \cos \Delta m_s t)] \\
& + \Gamma_{\Lambda} f_{\Lambda} \frac{e^{-\Gamma_{\Lambda_B} t}}{2} w_{\Lambda} \\
& + f_{udsc} \frac{1}{2} F_{udsc}(t), \qquad (5.1)
\end{aligned}
$$

where f_j represents the fraction of each b hadron type (see table 8) and light quark background (f_{udsc}=0.1%), ($j = u, d, s$, *baryon*, and *udsc* correspond to B^+, B^0_d, B^0_s, b baryon, and *udsc* background). $F_{udsc}(t)$ is a function describing the proper time distribution of the small contribution of background events. To parametrize this contribution, a sum of Gaussian and exponential functions is used. The distribution is shown in Fig.53.

$w_{u,d,s,\Lambda}$ are the mistag rates of the different b hadrons, as described in the previous chapter. It was discussed in the last chapter that w depends on the decay modes. In order to give an accurate description of the data sample these different decay modes must be treated separately in F_{mix}. Also recall that the net mistag probability depends on both the initial state mistag w^I and final state mistag probability w^F. For example, a B^+ can only be in the mixed sample if a mistake either in the initial state or the final state tag was made, but not both: $w = w^I(1 - w^F) + w^F(1 - w^I)$.

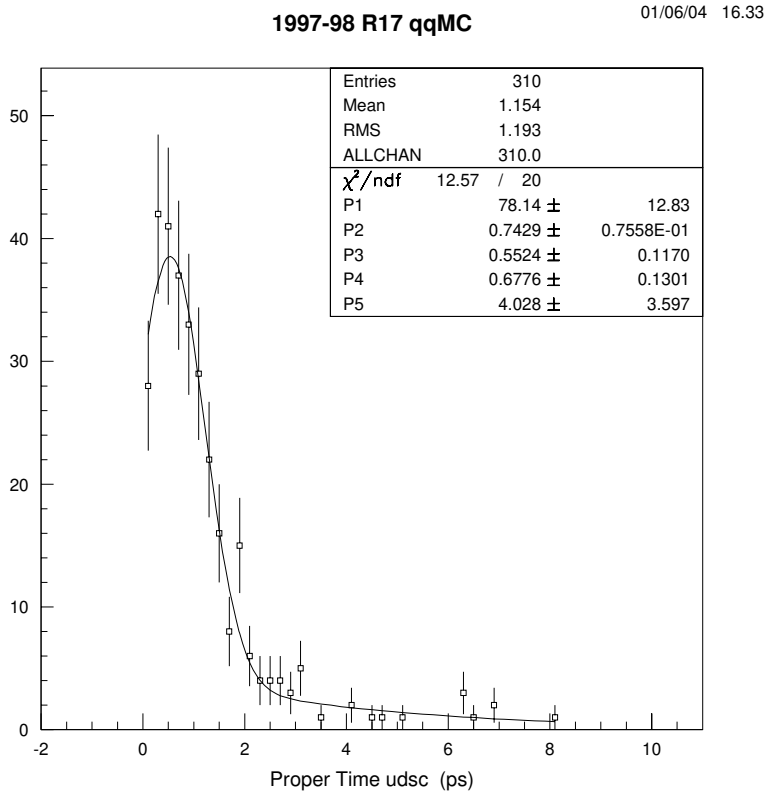


Figure 53: The proper time distribution of the small contribution of background events in the SLD Monte Carlo. The function is the fit result used to parametrize the light flavor background component as a function of true proper time.

The complete physics function describing the time development of the probability for a decay to be in the mixed sample is then expressed as

$$\begin{aligned}
F_{mixed}(t) =& \Gamma_{B^+} f_{B^+} \frac{e^{-\Gamma_{B^+} t}}{2} \left[w^I (1 - w_u^F) + (1 - w^I) w_u^F \right] \\
&+ \Gamma_{B_d^0} f_{B_d^0} \frac{e^{-\Gamma_{B_d^0} t}}{2} \left(\sum_{k=1}^{5} g_{dk} \left[(1 - w_{dk}^F)(1 + [2w^I - 1] \cos \Delta m_d t) + w_{dk}^F (1 - [2w^I - 1] \cos \Delta m_d t) \right] \right) \\
&+ \Gamma_{B_s^0} f_{B_s^0} \frac{e^{-\Gamma_{B_s^0} t}}{2} \left(\sum_{k=1}^{5} g_{sk} \left[(1 - w_{sk}^F)(1 + [2w^I - 1] \cos \Delta m_s t) + w_{sk}^F (1 - [2w^I - 1] \cos \Delta m_s t) \right] \right) \\
&+ \Gamma_{\Lambda} f_{\Lambda} \frac{e^{-\Gamma_{\Lambda_B} t}}{2} \left[w^I (1 - w_{baryon}^F) + (1 - w^I) w_{baryon}^F \right] \\
&+ f_{udsc} \frac{1}{2} F_{udsc}(t). \qquad (5.2)
\end{aligned}$$

A similar expression for the probability $F_{unmixed}$ to observe a decay tagged as

unmixed is obtained by replacing the initial state mistag rate w^I by $(1 - w^I)$.

The quantities w^F_{dk} and w^F_{sk} are the final state mistag probabilities for B^0_d and B^0_s, with the index $k = 1, ..., 5$ representing the five different decay final states: D^0X, D^+X, D_sX, charmed baryon X, and $D\overline{D}X$; g_{dk} and g_{sk} are the fractions of B^0_d and B^0_s decays into each of the above final states (see tables 9 and 10).

Several of the quantities in Eq. (5.2) can be determined on an event by event basis, as discussed in the last chapter. The initial state mistag probability w^I depends on $\cos\theta$ of the thrust axis, the electron beam polarization, as well as several quantities from the opposite hemisphere: jet charge, vertex charge, kaon charge, lepton charge and dipole charge. Recall that the final state mistag probabilities w^F_{jk} are parametrized as functions of the reconstructed decay length to take into account the degradation of the charge dipole tag close to the IP, and that the average b hadron type depends on $\cos\Theta$. Finally, the decay final state fractions g_{dk} and g_{sk} are parametrized as a function of the overall B vertex mass.

The functions $F_{mix}(t)$ and $F_{unmix}(t)$ are functions of the true B decay proper time t. They need to be transformed into functions of the reconstructed time t_{rec} as observed in the data. This will be done by taking into account all effects that depend on true proper time and integrating over it.

5.2 Efficiency

One important effect when considering proper time dependencies is that caused by the event selection cuts. It introduces a considerable bias, because at low proper time the IP is not very well separated from any secondary vertex. The effect is a time dependent b hadron selection efficiency. It is parametrized with a time-dependent function $\epsilon(t)$ for each b hadron type, shown in Fig.54. It is defined as the

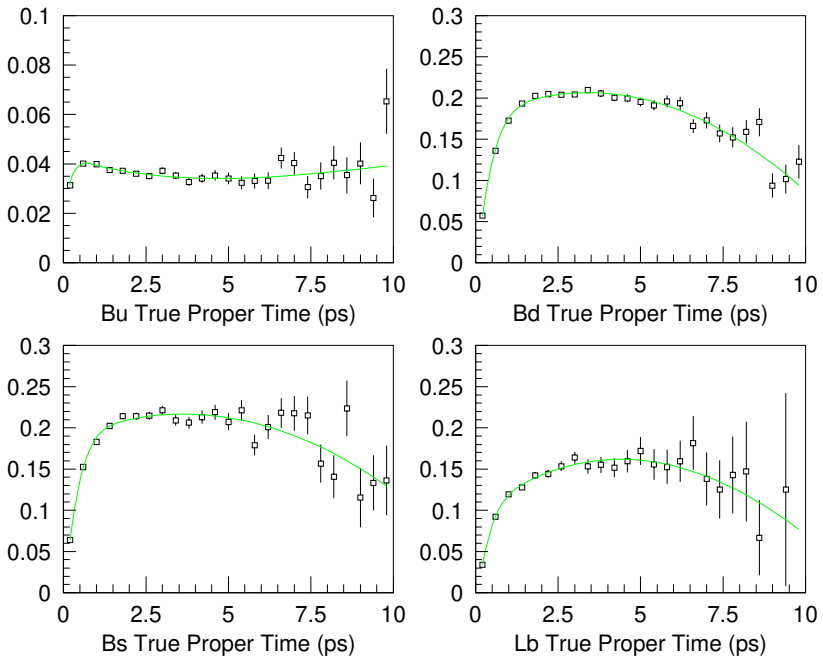


Figure 54: **The event selection efficiency per b hadron in the dipole analysis (MC simulation), as a function of true proper time. The functions are the fit results used to parametrize the charge dipole efficiency as a function of true proper time.**

fraction of B^+, B^0_d, B^0_s and b baryons in the $b\overline{b}$ Monte Carlo that pass the analysis cuts. All of the efficiency functions fall off at low proper time, but are fairly flat otherwise. At very high true proper time the selection efficiency decreases. This is due to a cut in the event selection removing K^0 and Λ candidates. Recall that for the B^0_s–$\overline{B^0_s}$ mixing analyses the high proper time region ($> 1.5\text{ps}^{-1}$) is less important, as the sensitivity for an oscillating signal decreases rapidly beyond that point due to the boost resolution. For example, the efficiency for B^0_s decays can be parametrized using 4 parameters a, b, c and d,

$$\epsilon(t) = a\,\frac{1-e^{bt}}{1+e^{bt}} + ct + dt^2. \tag{5.3}$$

Note that the contamination of B^+ at low proper time is considerable. However, with increasing proper time the fraction of neutral b hadrons becomes larger.

5.3 Proper Time Resolution

To convert the functions F_{mixed} and $F_{unmixed}$ into real experimental functions they are convoluted with a proper time resolution function $\mathcal{R}(t, t_{rec})$. This is necessary because to be able to relate true proper time and reconstructed time, information about the time resolution σ_t is needed. The reconstructed proper time t_{rec} is computed from the measured decay length L and the boost $\gamma\beta$:

$$t_{rec} = \frac{L}{\gamma\beta c} \tag{5.4}$$

with calculated proper time resolution σ_t

$$\sigma_t = \sqrt{\left(\frac{\sigma_L}{\gamma\beta c}\right)^2 + \left(\frac{\sigma_{\gamma\beta} t}{\gamma\beta}\right)^2}. \tag{5.5}$$

At very low proper time the decay length resolution dominates, whereas at larger times, the boost resolution component limits σ_t. This can be seen in Fig. 55, where Eq.5.5 is plotted for average values of the core component of σ_L and $\sigma_{\gamma\beta}/\gamma\beta$ in B_s^0 decays. The proper time resolution σ_t is almost constant up to values of $t \approx 0.3$ps. For higher values of t, σ_t rises linearly with t.

Given these residuals one can form a Gaussian resolution function relating true and reconstructed proper times:

$$G(t, t_{rec}) = \frac{1}{\sqrt{2\pi}\sigma_t} e^{-\frac{1}{2}((t-t_{rec})/\sigma_t)^2} \tag{5.6}$$

For both decay length and boost reconstruction, the resolution can be estimated decay-by-decay. This is preferable to using resolution averages, because a fraction

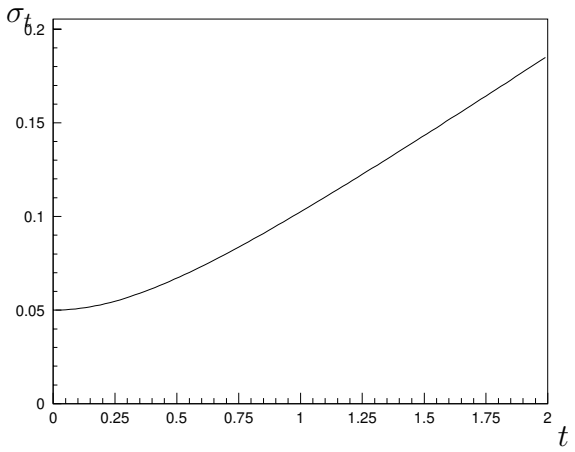


Figure 55: **σ_t as a function of t, plotted analytically using average values for the B_s^0 core fraction of σ_L and $\sigma_{\gamma\beta}/\gamma\beta$.**

of the events then have the benefit of a very high resolution and the sensitivity increases since well-reconstructed decays tend to have better decay length and boost resolution as well as lower mistag.

In the following sections, the two components of the proper time resolution, decay length and boost resolution, are discussed.

5.3.1 Decay Length Resolution

For decays with more than one track in either the B or D vertex, the resolution σ_L is estimated on an event-by-event basis from the B vertex fit and IP position measurement errors, combined to yield an uncertainty σ_{meas} along the flight direction.

As an example, Fig. 56 shows the distribution of the normalized decay length residuals, $\frac{L_{meas}-L_{true}}{\sigma_{meas}}$, for the case of correctly (incorrectly) tagged $B_s^0 \to D_s X$ and $B_s^0 \to D\overline{D}X$ events. Note that the latter distributions have very large asymmetric tail. The bias towards larger reconstructed decay length is due to the fact that

in $B_s^0 \to D\overline{D}X$ decays, as discussed earlier, very often one of the D decays is misreconstructed as a B vertex, while the true B decay is missed. This results in a bias towards large reconstructed decay length.

Two Gaussians are fitted to the 60% core fractions and the 40% tail fractions of the events. The widths s_1 and s_2 of the two Gaussians for correctly (incorrectly) tagged $B_s^0 \to D_s X$ decays are 0.91(1.35) and 1.96(4.45). The parameters s_1 and s_2 are used to appropriately scale σ_{meas}, and a 60% core resolution $\sigma_{L1} = s_1 \times \sigma_{meas}$ and a 40 % tail resolution $\sigma_{L2} = s_2 \times \sigma_{meas}$ is determined. The offsets in the distributions are addressed later, in the discussion of systematic shifts. As before, the charged B meson and b baryon scale factor parametrizations are extracted by averaging over all decay modes, whereas for B_s^0 and B_d^0 decays the scale factors are extracted separately for each of the decay modes.

For decays with only one track in both B and D vertices the resolution is extracted from the overall decay length residual distributions in the simulation. The difference in decay length resolution between right and wrong charge dipole tags motivates treating those separately in the likelihood function. Fig. 57 shows the average decay length residual for B_s^0 events with only one track in each of the B and D vertices, plotted separately for correctly and incorrectly tagged events. The average decay length resolution for all B_s^0 decays with right (wrong) charge dipole tag can be parameterized by the sum of two Gaussians of widths $\sigma_{L1} = 76\,\mu$m (112 μm) and $\sigma_{L2} = 311\,\mu$m (450 μm), where the first Gaussian represents 60% of the decays. In the case of B_d^0 and B_s^0 decays, the right- and wrong-tag decay length resolutions are estimated separately for each of the five decay final states.

As for the question why we trust the Monte Carlo modelling of the resolution, it is tuned to the data using impact parameter distributions and τ 3-prong decay vertices. This is discussed in section 6.2.2.

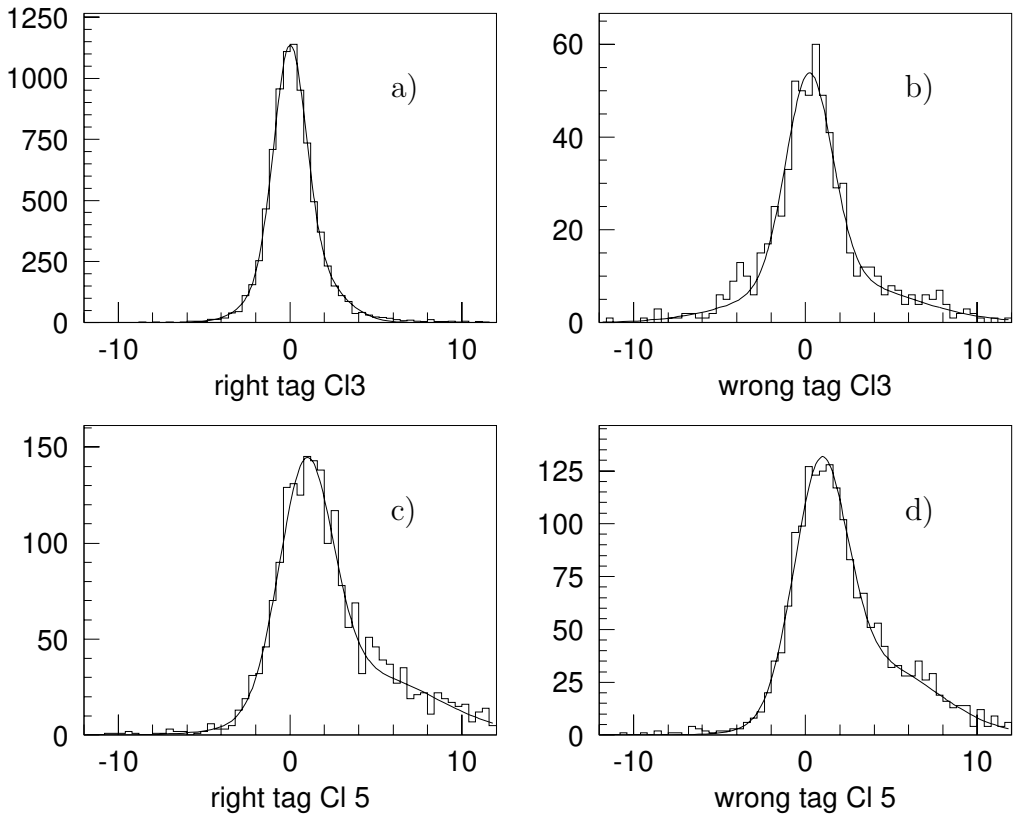


Figure 56: **The widths of the Gaussian functions fitted to these MC distributions of $\frac{L_{rec}-L_{true}}{\sigma_{meas}}$ are the scale factors for correctly and incorrectly tagged events used to adjust the event-by-event decay length resolution. a) correctly tagged $B_s^0 \to D_s X$ decays, b) incorrectly tagged $B_s^0 \to D_s X$ decays, c) correctly tagged $B_s^0 \to D\overline{D}X$ decays, and d) incorrectly tagged $B_s^0 \to D\overline{D}X$ decays. Two Gaussian functions are fitted with 60% of the decays in the core and 40% of the events in the tail.**

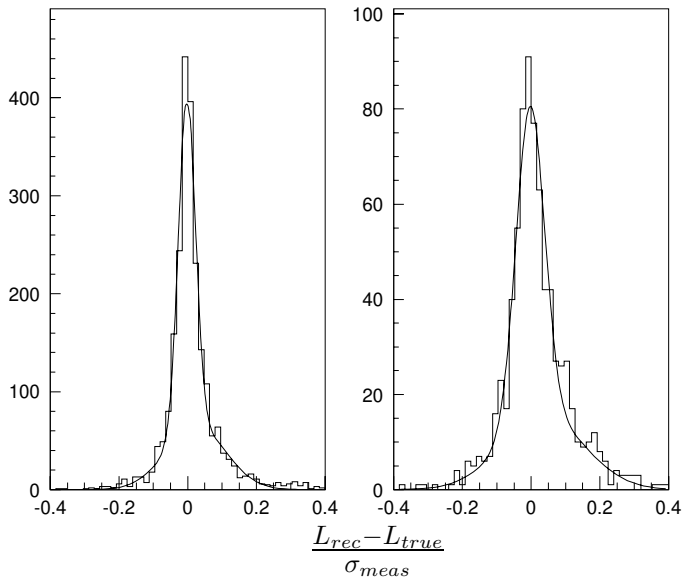


Figure 57: The widths of the Gaussians fitted to these MC distributions are the residuals for correctly (left) and incorrectly (right) tagged B_s^0 decays with one track in each of the B and D vertices. Two Gaussians are fitted with 60% of the decays in the core and 40% of the events in the tail.

5.3.2 Boost Resolution

The relative boost resolution $\sigma_{\gamma\beta}/\gamma\beta$ is parametrized with the sum of two Gaussians using the MC simulation. Considering all selected B_s^0 decays, the widths of the two Gaussians are $\sigma_{B1} = 0.07$ and $\sigma_{B2} = 0.21$, where the first Gaussian represents 60% of the decays. However, the analysis takes into account the strong dependence of the resolution on the total charged track energy in each decay. Recall that the boost is determined as the sum of charged track energy $E_B^{\pm}$ and neutral energy. The charged track energy $E_B^{\pm}$ can be determined accurately, because the CDC and the vertex detector provide precise tracking determination. By comparison, the LAC information used to determine the neutral component is relatively imprecise due to problems such as track-cluster association and cluster overlap [87], and the

fact that it is very difficult to determine the origin of neutral particles. Therefore, the strong dependence of the relative boost resolution on $E_B^{\pm}$ is to be expected. Fig. 58 shows the width of the relative boost residual as a function of the charged B energy for the four different b hadrons. Using this information, $\sigma_{\gamma\beta}/\gamma\beta$ can be estimated on an event-by-event basis.

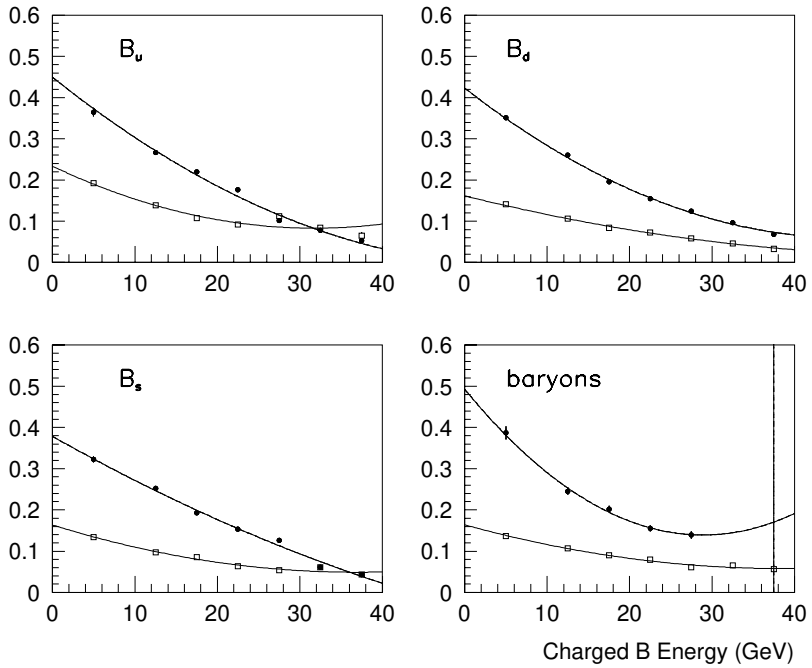


Figure 58: The width of the relative boost residual as a function of the charged B energy for the 4 different b hadrons. The open squares represent the 60% core fraction of the decays. The dark squares represent the 40% tail fraction. The functions are the fit results used to parametrize the relative boost resolution as a function of the charged B energy.

5.3.3 Systematic Shifts in Proper Time Resolution

Once estimates of the decay length and boost resolutions are made per event, the resolution function can be derived. The resolution function can in principle be

parametrized by the sum of four Gaussians:

$$G(t, t_{rec}) = \sum_{i=1}^{2} \sum_{j=1}^{2} f_{ij} \frac{1}{\sigma_{ij}(t)\sqrt{2\pi}} e^{-\frac{1}{2}((t-t_{rec})/\sigma_{ij}(t))^2} \tag{5.7}$$

where the index $i = 1(2)$ corresponds to the core (tail) component of the decay length resolution scale factor, and the index $j = 1(2)$ similarly corresponds to the core (tail) component of the relative boost resolution $\frac{\sigma_{\gamma\beta j}}{\gamma\beta}$. The various fractions are $f_{11} = 0.36$, $f_{12} = f_{21} = 0.24$, and $f_{22} = 0.16$. This produces a Gaussian function centered at $t = t_{rec}$ and therefore does not account for any systematic shifts in the reconstruction.

Decays reconstructed within 200μm of the IP or at low reconstructed boost have worse proper time resolution and suffer from asymmetric tails. In addition, offsets in the proper time residuals are observed in the Monte Carlo. This is mostly due to double charm decays with one of the D vertices misreconstructed as a B vertex with a long decay length. To account for this effect, an estimate of the proper time shift is extracted for each event from fits to the proper time residuals in bins of true proper time. A Gaussian was found to yield a good fit to the symmetric core fraction of 60% of the events, and a Novosibirsk function (described below) was used to fit the 40% tail fraction, which is distributed asymmetrically around zero. The Novosibirsk function is used to fit Gaussian-like distributions with asymmetric tails. In addition to a width and a mean, $\sigma(t)$ and μ, as in the case of a Gaussian function, the Novosibirsk function uses a third parameter P_{tail} which parametrizes the width of the asymmetric tail. Both the Gaussians that parametrize the core fraction of the events and the Novosibirsk functions for the tail fractions have time dependent offsets $\mu^c(t)$ and $\mu^t(t)$, respectively. The Novosibirsk function is given

by:

$$f_{Novo}(t;\sigma(t),\mu^t(t),P_{tail}) =$$
$$\frac{1}{(\sqrt{2\pi}\sigma(t))}exp\left(-\frac{1}{2}\frac{\log\left[1.0+P_{tail}(t_{rec}-t-\mu^t(t))\frac{\sinh(1.18P_{tail})}{(1.18\sigma P_{tail})}\right]^2}{P_{tail}}-\frac{1}{2}P_{tail}^2\right) \quad (5.8)$$

and the Gaussian is given by

$$f_{Gauss}(t;\sigma(t),\mu^c(t)) = \frac{1}{(\sqrt{2\pi}\sigma)}exp\left(-\frac{1}{2}\left(\frac{(t_{rec}-t-\mu^c(t))}{\sigma(t)}\right)^2\right). \quad (5.9)$$

f_{Novo} and f_{Gauss} are then combined to form a resolution function that includes the systematic shifts due to the offsets:

$$\begin{aligned} G(t,t_{rec}) &= 0.36 f_{Gauss}(t;\sigma_{11}(t),\mu^c(t)) \\ &+ 0.24 f_{Gauss}(t;\sigma_{12}(t),\mu^c(t)) \\ &+ 0.24 f_{Novo}(t;\sigma_{21}(t),\mu^t(t),P_{tail}) \\ &+ 0.16 f_{Novo}(t;\sigma_{22}(t),\mu^t(t),P_{tail}). \end{aligned} \quad (5.10)$$

This is done separately for decays involving one or two charm particles, as the effect is smaller for the former. Mistagged and correctly tagged events are also treated separately, as the effect is stronger for mistagged events.

The results of the event-by-event resolution parametrization can be checked in the Monte Carlo. Fig.59 shows the proper time residuals, $t_{rec}-t$, in 6 bins of true proper time, for the case of $B_s^0 \to D\overline{D}X$ decays. The points are extracted from the Monte Carlo, and the histogram is the calculated resolution function with the corrected reconstructed proper time. The plots indicate that the resolution function adequately describes the resolution and biases in the simulated proper time reconstruction.

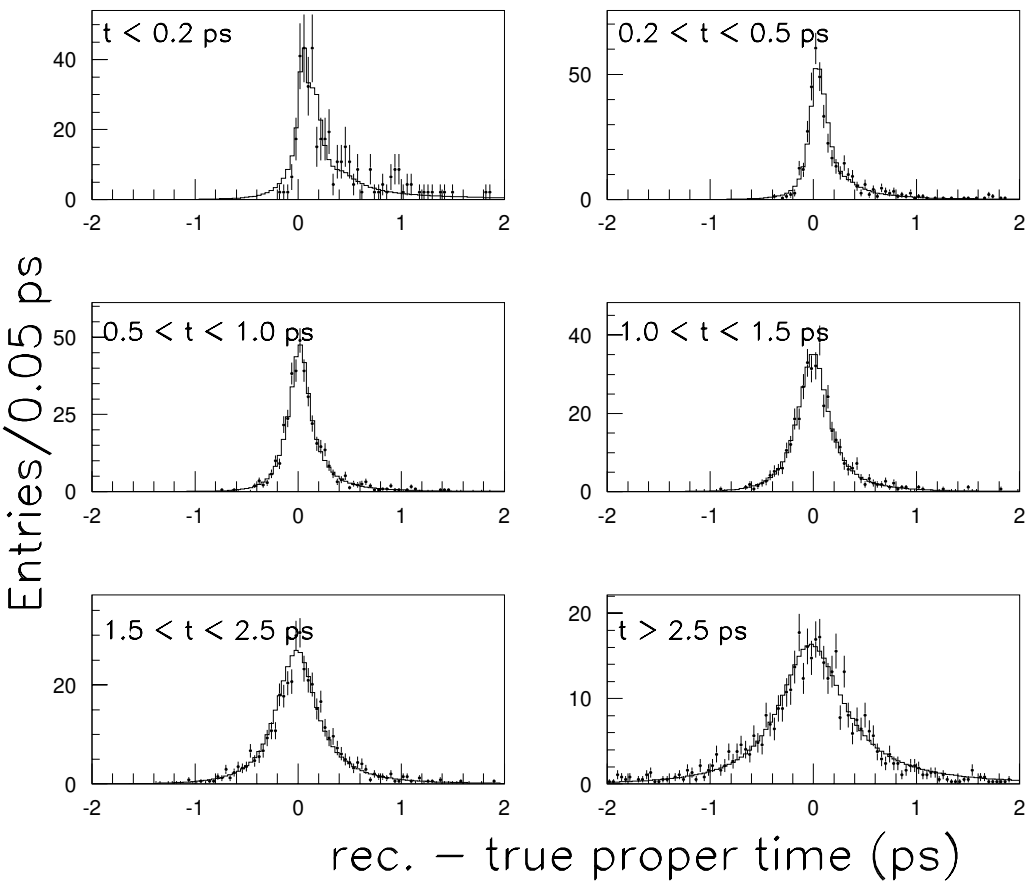


Figure 59: **The proper time residuals, $t_{rec} - t$, in 6 bins of true proper time, for the case of $B_s^0 \to D\overline{D}X$ decays. The data points are extracted from the Monte Carlo, and the histogram is the calculated resolution function.**

Given the vertex efficiency ϵ and the resolution function $G(t, t_{rec})$, we now transform F_{mix} and F_{unmix} into functions of the reconstructed proper time t_{rec}:

$$P_{mix}(t_{rec}) = \int_0^\infty F_{mix}(t)\epsilon(t)G(t, t_{rec})dt, \tag{5.11}$$

and

$$P_{unmix}(t_{rec}) = \int_0^\infty F_{unmix}(t)\epsilon(t)G(t, t_{rec})dt. \tag{5.12}$$

To summarize, $P_{mix}(t_{rec})$ and $P_{unmix}(t_{rec})$ describe the expected reconstructed proper time distributions of the charge dipole tag that would agree or disagree respectively with the initial state tag. Finally, the likelihood function is given by

$$\mathcal{L} = \prod_{events} [P_{mix}(t_{rec}) + P_{unmix}(t_{rec})] \tag{5.13}$$

5.4 The Amplitude Fit

One of the ways to search for a periodic signal is to perform a Fourier analysis of the frequency spectrum. This technique is called the amplitude fit [89]. The procedure is to modify the B_s^0 oscillatory terms in the likelihood function by introducing the parameter A:

$$\frac{1}{2}(1 \pm cos\Delta m_s t) \rightarrow \frac{1}{2}(1 \pm Acos\Delta m_s t). \tag{5.14}$$

For fixed values of Δm_s the log$\mathcal{L}$ function is minimized with respect to the parameter A, which is the normalized Fourier amplitude of the mixing signal for a given frequency Δm_s. The Fourier analysis of an oscillation signal was introduced in section 4.5 (see Eq.4.16). In this context, the amplitude A is equivalent to a normalized Fourier amplitude $A(\Delta m_s)$:

$$A(\Delta m_s) = \frac{d(\Delta m_s)}{a(\Delta m_s)}, \tag{5.15}$$

where $d(\Delta m_s)$ represents the Fourier transform of the data and $a(\Delta m_s)$ is the expected peak amplitude. Using Eq.4.16 one obtains

$$A(\Delta m_s) = \frac{F(\Delta m_s^{true})}{F(\Delta m_s)} \frac{\Gamma^2}{\Gamma^2 + (\Delta m_s - \Delta m_s^{true})^2}. \tag{5.16}$$

The expected amplitude distribution retains the Breit-Wigner shape with $A = 1$ at $\Delta m = \Delta m^{true}$. For $|\Delta m - \Delta m^{true}| > 2\Gamma$, A is expected to be $\approx$0. This is true in the limit of good proper time resolution. As studies with toy Monte Carlo at LEP suggest, the Breit-Wigner shape can be distorted considerably if the resolution smearing is large. For reference on this subject, see [93]. The statistical uncertainty in A is

$$\sigma(A) = \frac{\sigma_d}{a(\Delta m_s)}, \tag{5.17}$$

where σ_d is the statistical noise in the FT of the data. Since $a(\Delta m_s)$ decreases with increasing Δm_s due to the proper time resolution, we expect the uncertainty on A to grow as a function of Δm_s.

The likelihood fits for A are generally nearly parabolic at the minima. As an example, Fig. 60 shows the likelihood fits for A at four different values of Δm_s: 5 ps^{-1}, 10 ps^{-1}, 15 ps^{-1} and 20 ps^{-1}. Shown is the MC with an input mixing frequency of Δm_s=10 ps^{-1}.

The $\pm 1\sigma$ errors, σ_A, are easily estimated by the $\Delta \log \mathcal{L} = 1/2$ range about the minimum. If no measurement of Δm_s can be made, a lower limit can be determined by the frequency at which $A + 1.645\sigma_A$ crosses the $A = 1$ line, and one can exclude any value of Δm_s for which $A + 1.645\sigma_A < 1$. Systematic errors must also be included and will be discussed later on. The intrinsic sensitivity to set a 95% confidence limit on Δm_s is determined by the frequency at which $1.645\sigma_A = 1$.

To conclude, the amplitude fit method is a procedure in which for each oscillation frequency a hypothetical amplitude is fitted. It combines the method of a

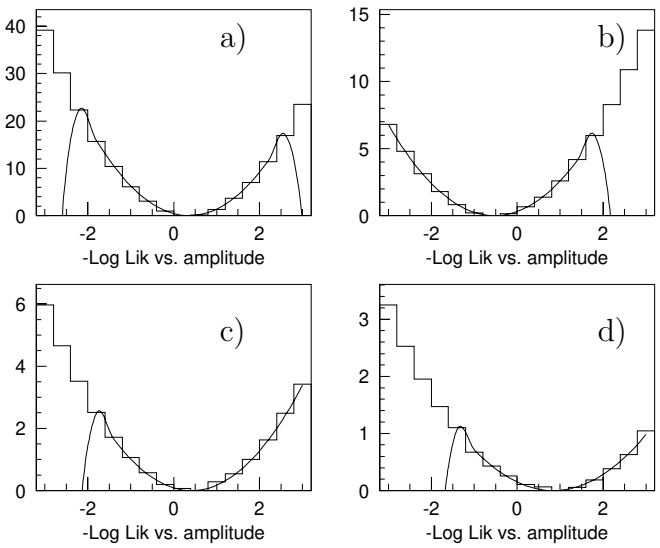


Figure 60: **The likelihood fits for A at four different values of Δm_s: a) 5 ps^{-1}, b) 10 ps^{-1}, c) 15 ps^{-1} and d) 20 ps^{-1}.**

Fourier analysis and a 'classical' likelihood fit. It can be shown that likelihood referenced to infinity and amplitude fit method are mathematically equivalent [89], however, the amplitude fit method offers the possibility to combine the results of independent analyses by averaging the different amplitude spectra taking into account the errors. In the case of likelihood methods, a similar combination is only possible when using reference to infinity. A combination is not possible when using likelihood differences from the minimum. This would need a calibration of the total likelihood which appears to be not feasible.

Before progressing to an amplitude fit to the data, it is important to check if the complicated fit mechanism works correctly. For this reason it has been tested on the Monte Carlo in various cases. For example, Fig. 61 shows the results of the amplitude fits performed on a pure B_s^0 Monte Carlo sample as selected by the charge dipole analysis. The fit clearly indicates a signal at the Monte Carlo input

value of Δm_s= 10ps^{-1}. The amplitude curve reaches a peak at $A = 1$. The error bars on the amplitude measurements grow, as expected, as a function of Δm_s. The negative fluctuation of the amplitude at $\approx$ 6 ps^{-1} is at the 3σ level. Note that deviations of two or three σ are not uncommon in amplitude plots, as will be seen. The result of the fit is a value of $\Delta m_s = 10.324 \pm 0.114$ps^{-1} at A=1.

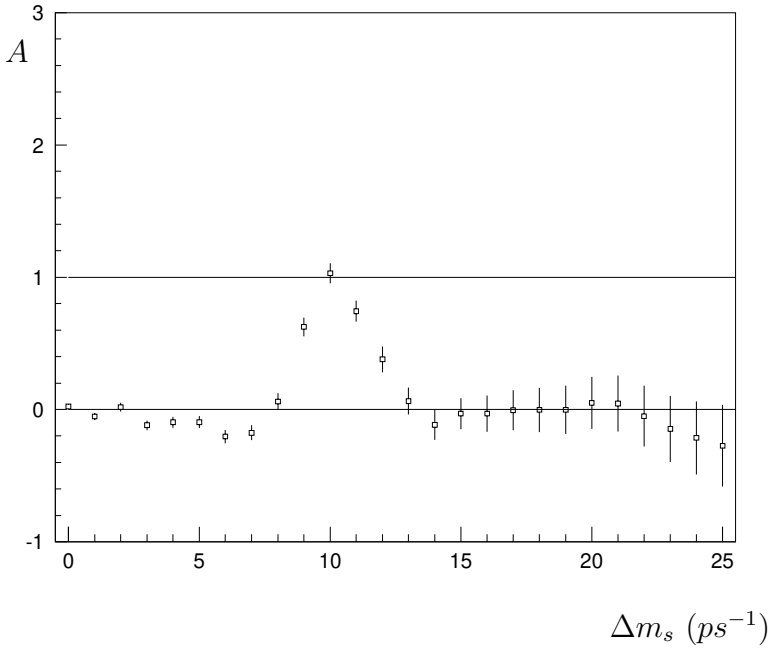


Figure 61: **Amplitude fit to a pure Monte Carlo B_s^0 sample, selected by the dipole charge analysis. The Monte Carlo Δm_s input value is 10ps^{-1}.**

Fig. 62 shows a plot of the amplitude fit performed on $b\bar{b}$ Monte Carlo, which has $\sim$10 times more events than the 1996-98 data set. Again, the plot of the amplitude indicates a signal at the Monte Carlo input value of Δm_s= 10ps^{-1}.

It is interesting to change the Monte Carlo input value to Δm_s=15 ps^{-1}, a more realistic value, to observe if the faster oscillation can still be observed in the amplitude fit to the Monte Carlo (see Fig. 63). At Δm_s = 15ps^{-1} a value of A=0.698$\pm$ 0.294 is observed, consistent with 1. The central value of the amplitude

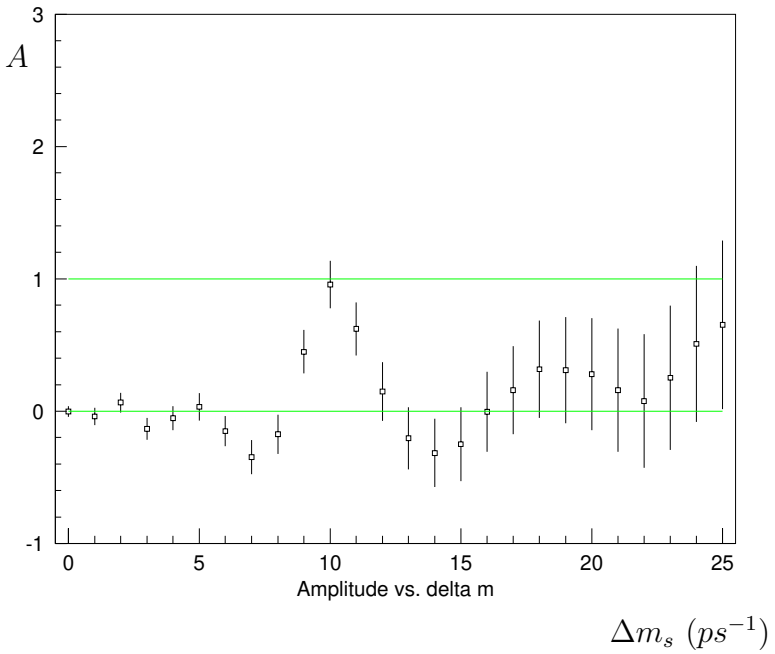


Figure 62: **Amplitude fit to a Monte Carlo $b\bar{b}$ sample, selected by the dipole charge analysis. The Monte Carlo Δm_s input value is $10ps^{-1}$.**

continues to rise to $\Delta m_s = 16\text{ps}^{-1}$ and decreases beyond that point. The amplitude fluctuates 2.5σ low at $\Delta m_s = 7\text{ps}^{-1}$.

Fig. 64 shows the amplitude fit in the SLD Monte Carlo, as the input value of Δm_s is changed from $10ps^{-1}$ to $200ps^{-1}$. Also shown are the amplitude A divided by σ_A and A/σ_A versus Δm_s. A Gaussian fit to the distribution of A/σ_A has a width consistent with 1 and a mean consistent with 0. This signifies that σ_A covers the range of fluctuations around an amplitude of $A = 0\text{ps}^{-1}$ and is well modelled. In other words, the fact that A/σ_A versus Δm_s does not fluctuate beyond values much higher than $A/\sigma_A \approx 1$ signifies that the amplitude fit does not introduce any unexpected fluctuations due to mismodelling of the Monte Carlo.

It is also possible to divide the Monte Carlo into data-size samples and to run the amplitude fit on them. Fig. 65 shows 12 small Monte Carlo samples, with a

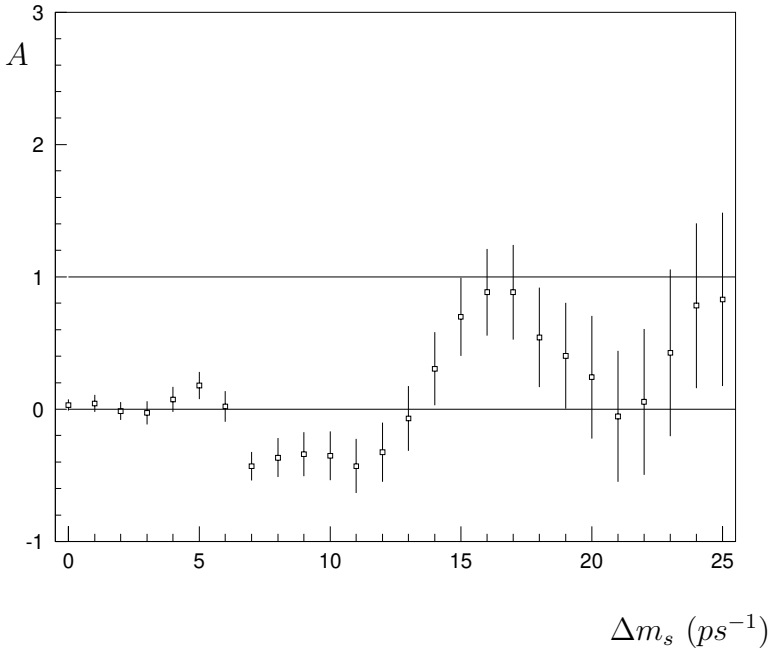


Figure 63: **Amplitude fit to a Monte Carlo $b\bar{b}$ sample, selected by the dipole charge analysis. The Monte Carlo Δm_s input value is $15ps^{-1}$.**

size comparable to the data sample (the data sample contains 11462 events, the 12 small MC samples contain 8700 events each). The input value of Δm_s has again been set to 200 ps^{-1}. These samples are statistically uncorrelated. Fig. 66 shows a plot of A/σ_A for the 12 experiments. Since the amplitude points are correlated, A/σ_A is plotted for well-separated values of A at Δm_s=1 ps^{-1}, 5 ps^{-1}, 10 ps^{-1}, 15 ps^{-1} and 20 ps^{-1} only. The fluctuations are Gaussian, with a width of 1.114±0.140 and a mean compatible with 0. This signifies that the errors on A cover the range of fluctuations around an amplitude of $A = 0$ps^{-1} and are well modelled.

A useful test of the likelihood function is to study the match of mixed fraction, i.e., the fraction of events tagged as mixed, in the MC and the likelihood function. Fig. 67 shows the mixed fraction for the four b hadron types with their respective likelihood functions, as a function of reconstructed proper time. Note the slow

oscillation in the B_d^0 mesons (the SLD MC input B_d^0–$\overline{B_d^0}$ mixing frequency is Δm_d = 0.484 ps^{-1}), and the rapid oscillations in the B_s^0 mesons (the SLD MC input B_s^0–$\overline{B_s^0}$ mixing frequency is Δm_s = 10 ps^{-1}). As expected, the mixed fraction in B^+ and b baryons is flat as a function of time. Note the small increase of the mixed fraction for B^+ at very small proper time, due to the lower tagging power in that region. The match between Monte Carlo and likelihood function is excellent and gives confidence in the accuracy of the likelihood function.

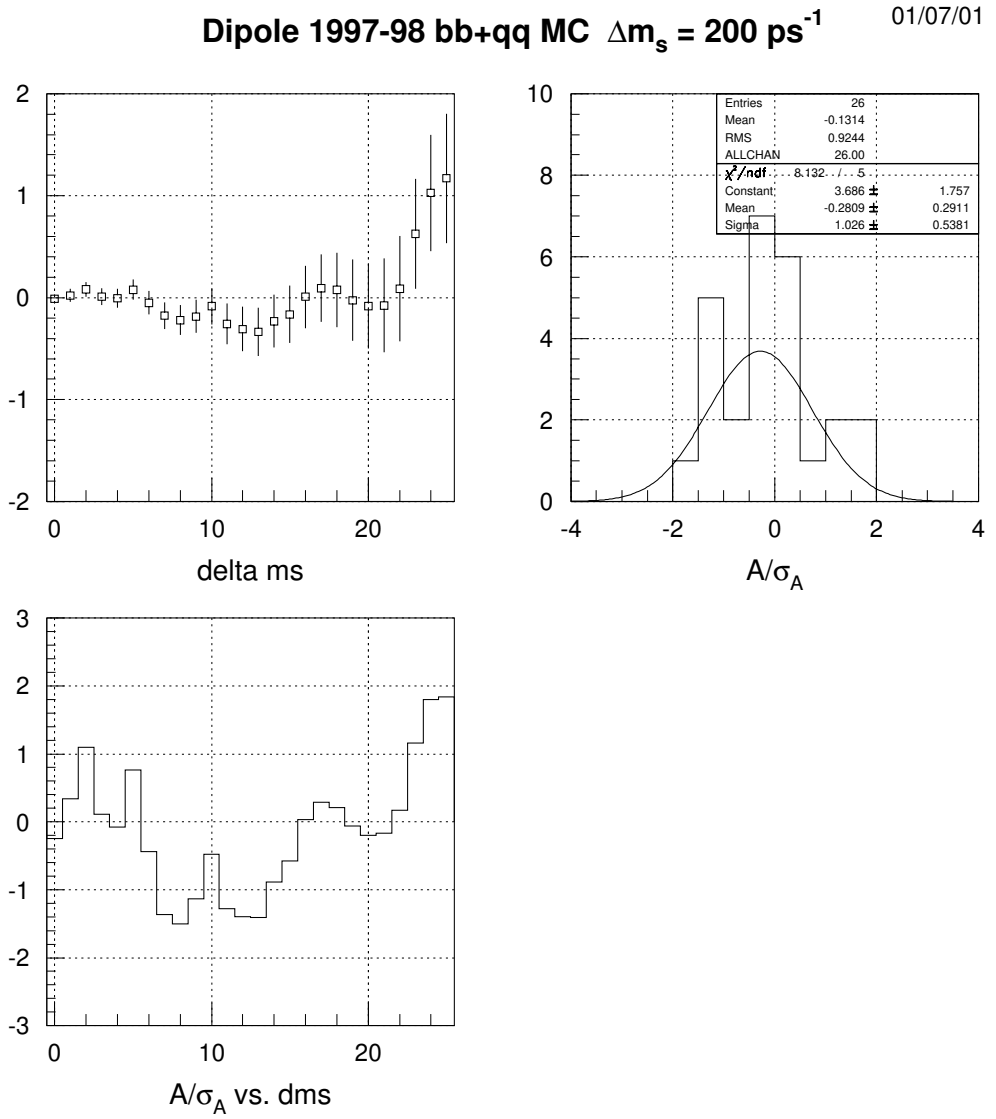


Figure 64: **Amplitude fit to the SLD Monte Carlo sample, selected by the dipole charge analysis. The Monte Carlo Δm_s input value is $200ps^{-1}$. Also shown are the amplidude A divided by σ_A and A/σ_A versus Δm_s.**

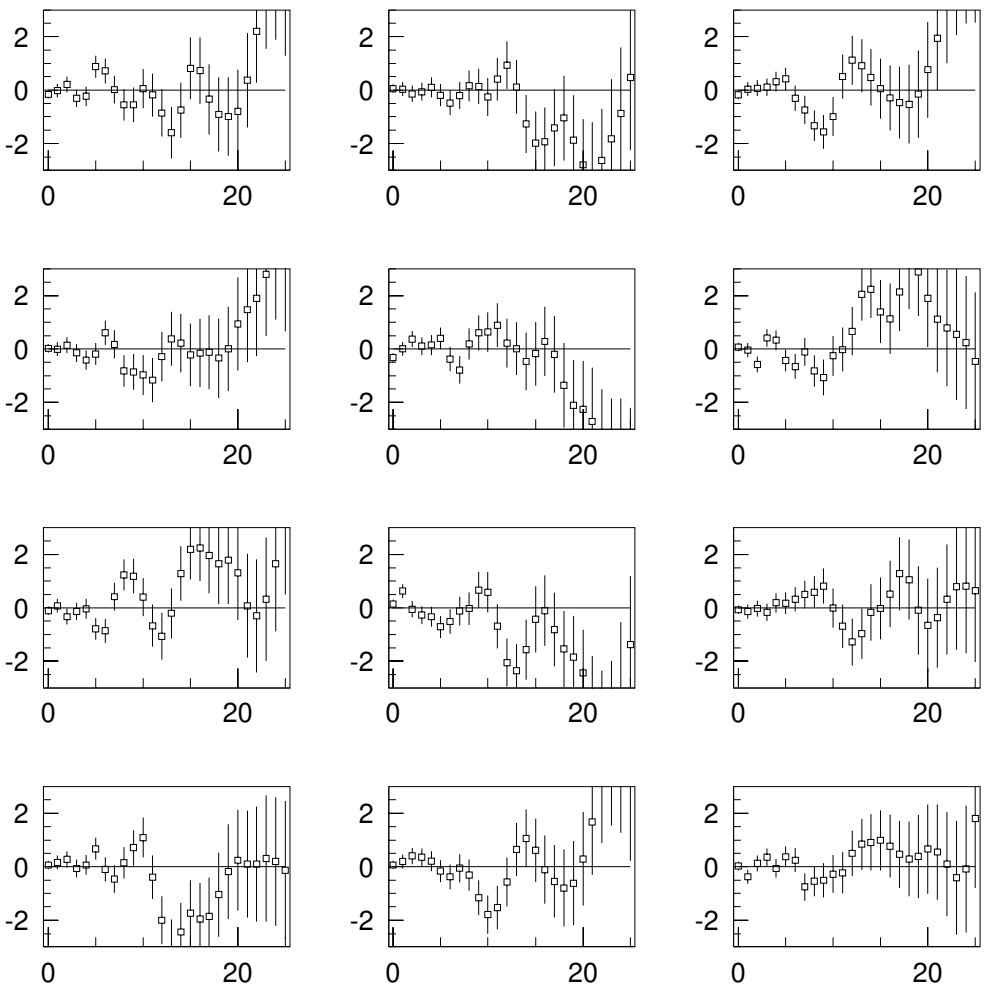


Figure 65: **Amplitude fits for 12 data-size Monte Carlo samples, where the input value of Δm_s has been set to 200 ps^{-1}.**

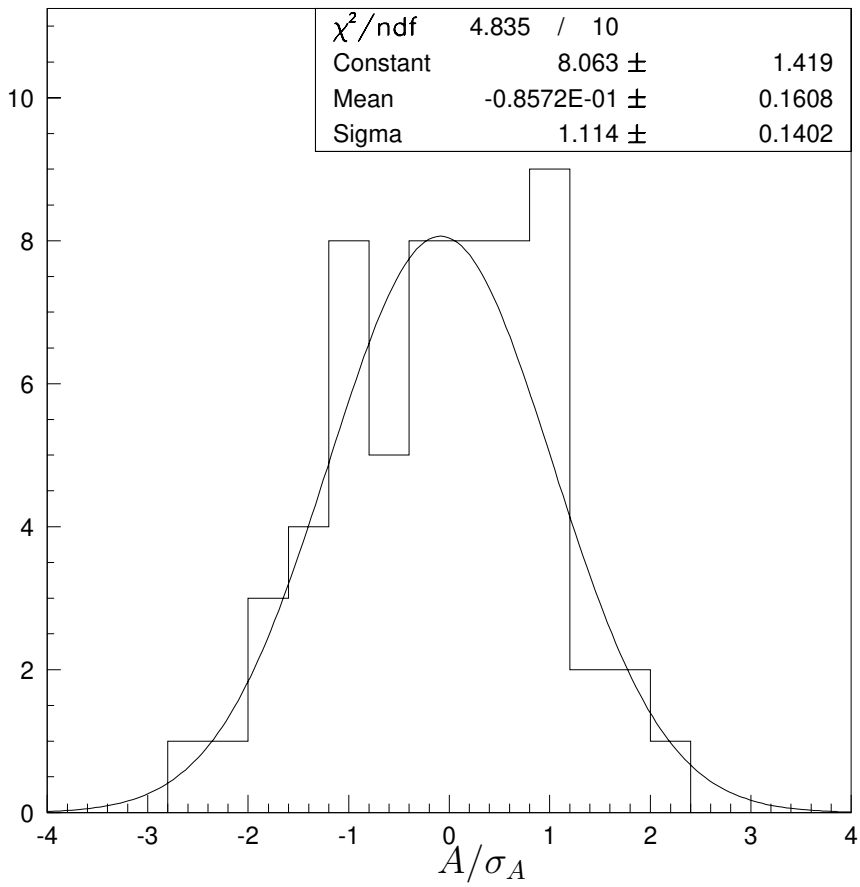


Figure 66: **A/σ_A for the 12 independent MC experiments. A/σ_A is plotted for well-separated values of A atΔm_s=1 ps^{-1}, 5 ps^{-1}, 10 ps^{-1}, 15 ps^{-1} and 20 ps^{-1} only. The function is the result of a Gaussian fit to the histograms.**

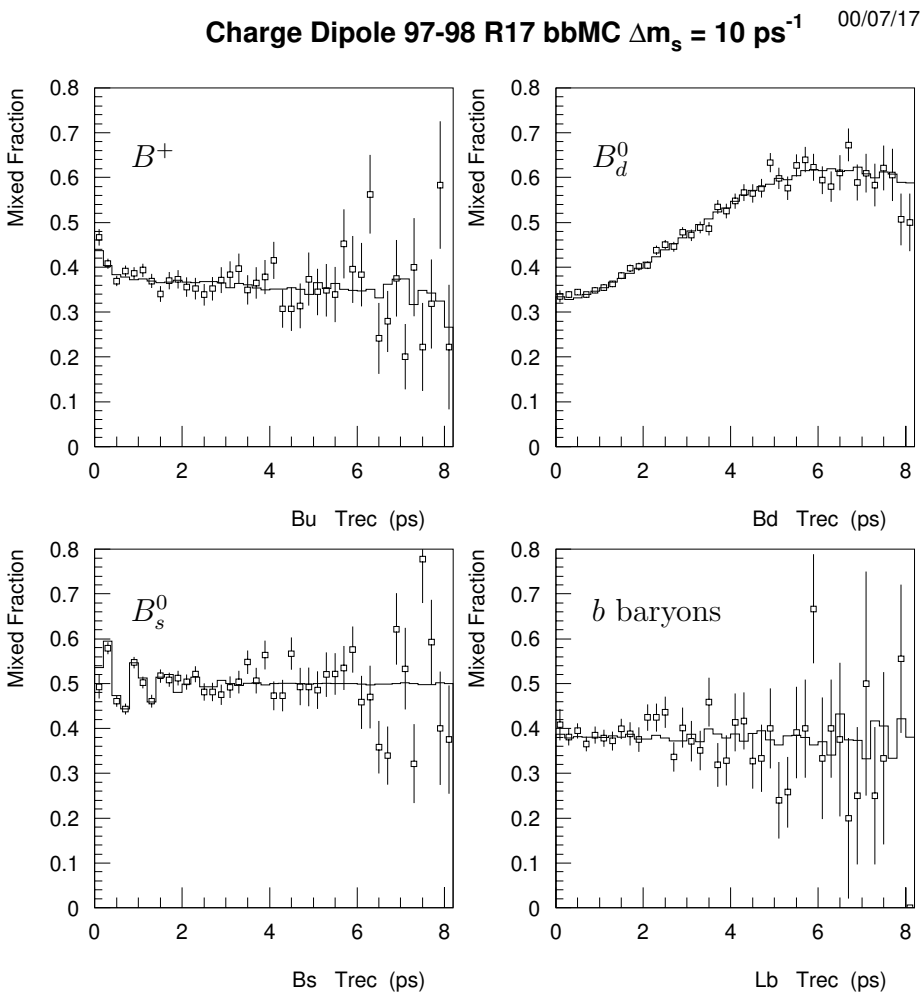


Figure 67: **The mixed fraction for the four b hadron types as a function of reconstructed time. The points are the Monte Carlo, and the histogram represents the likelihood function for each b hadron type separately. The input B_s^0–$\overline{B_s^0}$ mixing frequency is $\Delta m_s = 10$ ps^{-1}.**

CHAPTER 6

Results

In this chapter, the results of the amplitude fit method will be presented for the B_s^0–$\overline{B_s^0}$ mixing analysis. The systematic uncertainties on the measurement will also be examined. In addition, several consistency checks including the measurement of the B_d^0–$\overline{B_d^0}$ mixing frequency with the amplitude fit method will be presented. Finally, the implications of this measurement on the determination of the CKM matrix elements will be discussed.

6.1 B_s^0–$\overline{B_s^0}$ Mixing Results

Fig. 68 shows the amplitude versus Δm_s for the charge dipole analysis using the full 96-98 SLD data. From the approximately 400k hadronic events selected during this period, 11462 dipole events were selected. This number excludes 1498 events that were selected by the other two SLD B_s^0–$\overline{B_s^0}$ mixing analyses, to keep the analyses statistically uncorrelated. The procedure of removing these overlap events is discussed later, in section 6.4. The amplitude errors are statistical only. The sensitivity is represented by the dashed curve. The plot shows that all values of the amplitude are consistent with $A = 0$ ps^{-1}. The largest deviation from $A = 0$, although statistically insignificant, occurs around $\Delta m_s = 17$ ps^{-1}. Even

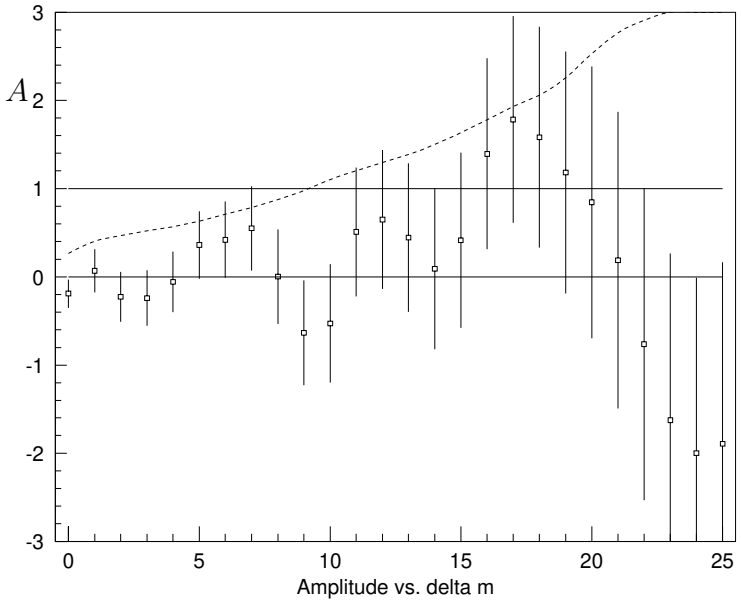


Figure 68: Measured amplitude for the full 96-98 SLD data set using the charge dipole analysis described in this thesis. The 1σ error bars are statistical only. The dashed line is $1.645\sigma_A$. The Δm_s value for the intersection between this line and the A = 1 line represents the sensitivity of the experiment.

less significant bumps are observed at $\Delta m_s = 7$ ps^{-1} and at $\Delta m_s = 12$ ps^{-1}. A discussion of excluded regions has to be postponed until systematic errors have been taken into account. Fig. 69 shows the likelihood fit performed on the same sample, where the amplitude has been set to 1. Instead of minimizing -log$\mathcal{L}$ for A at different values of Δm_s, here the -log$\mathcal{L}$ is minimized directly for Δm_s. It exhibits the same structure as the amplitude fit. The most significant minimum occurs at $\Delta m_s = 17$ ps^{-1}. Two other local minima are observed as in the amplitude fit, at $\Delta m_s = 7$ ps^{-1} and at $\Delta m_s = 12$ ps^{-1}.

It is important to check if the likelihood function used to extract the amplitude fit is a good representation of the data. For this purpose, it is instructive to study

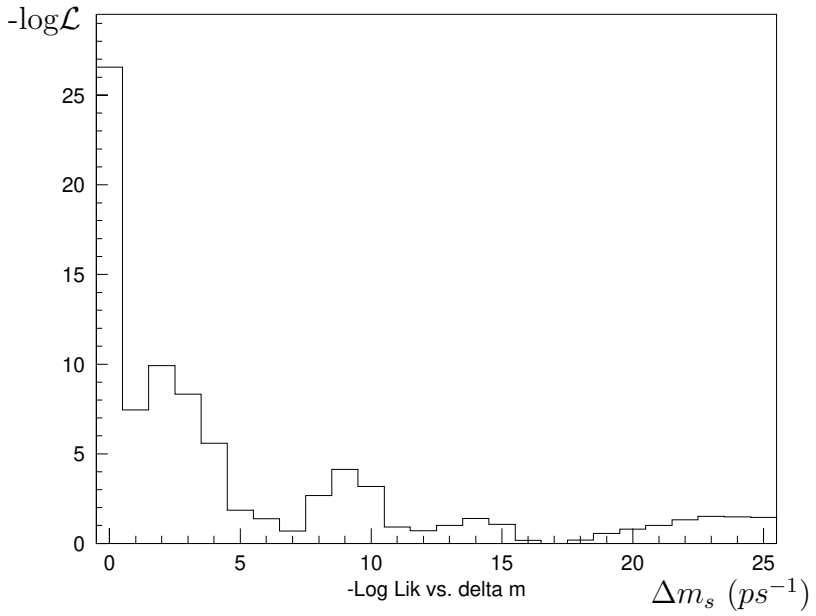


Figure 69: Measured log$\mathcal{L}$ fit for the full 96-98 SLD data set using the charge dipole analysis described in this thesis.

Fig. 70. It shows the mixed fraction as a function of reconstructed proper time, for the charge dipole selected data events. The mixed fraction is the ratio of events tagged as mixed to the full sample of events. The data points are the 96-98 data, and the histogram is a representation of the likelihood function used to extract the amplitude fit. The likelihood function was generated with the very high frequency of Δm_s=25ps^{-1}, so that no oscillations (other than the B_d^0 oscillations) can be seen in the plot. The likelihood function appears to be a very good description of the data. Note that the data points fluctuate above and below the likelihood function. Every fluctuation observed in the data mixed fraction corresponds to a fluctuation in the amplitude plot. This is demonstrated in Fig. 71, where the data mixed fraction is overlaid with four different likelihood functions, generated with Δm_s=25 ps^{-1}, Δm_s=7 ps^{-1}, Δm_s=12 ps^{-1} and Δm_s=17 ps^{-1}. In each of the three

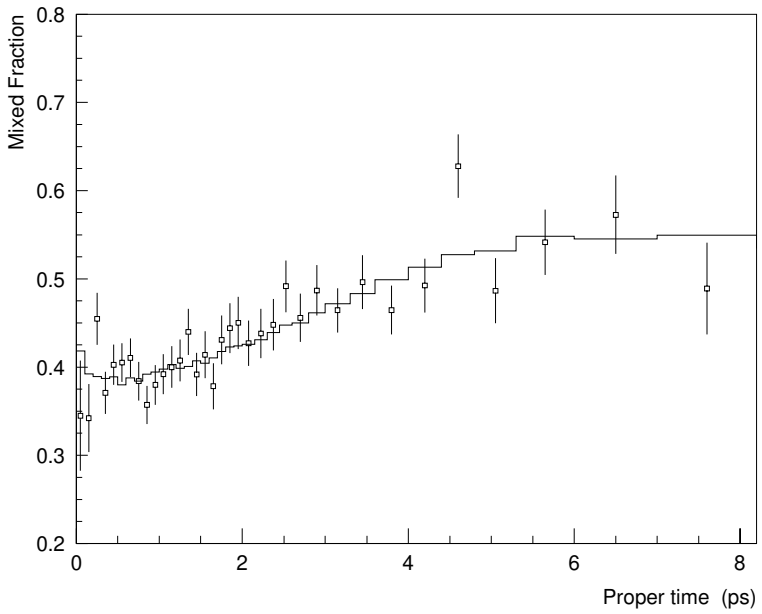


Figure 70: The mixed fraction as a function of reconstructed proper time, for dipole selected events. The datapoints are from the 96-98 data, the histogram is a representation of the likelihood function used to extract the Amplitude fit. The likelihood function was generated with $\Delta m_s = 25ps^{-1}$.

last cases, the likelihood function fits a certain aspect of the data fluctuations.

This shows that the amplitude fits in the high Δm_s region are especially sensitive to the structure of the data in the very low reconstructed proper time region, where the uncertainties on the reconstruction are the highest. As a check on how much of the amplitude fit structure is due to events reconstructed at the very low end of the proper time, the amplitude fit was repeated for events with $t_{rec} > 0.2\text{ps}^{-1}$ only. Fig. 72 shows the amplitude fit with a cut at t_{rec}=0.2ps^{-1}, excluding the very low proper time region. The figure shows a small change in structure in the high Δm_s region, whereas it stays the same for values of Δm_s lower than $12ps^{-1}$. This means that the effect of these events on the shape of the amplitude fit is not significant. Also, σ_A stays the same.

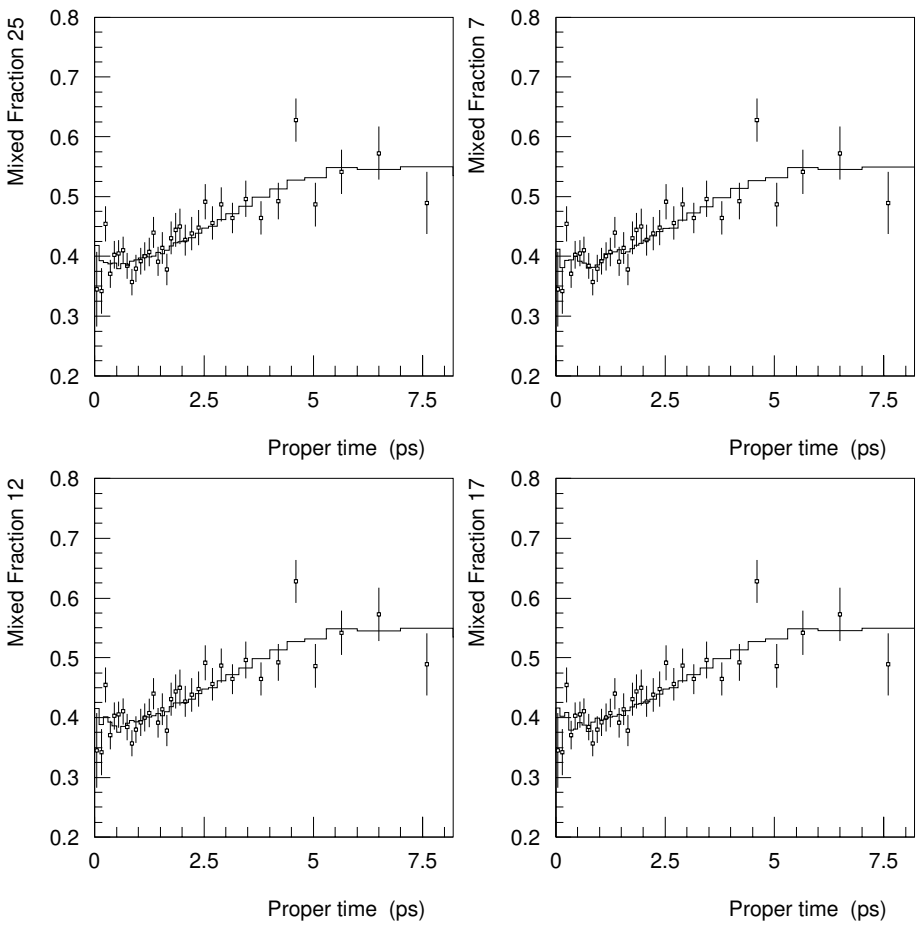


Figure 71: **The mixed fraction as a function of reconstructed proper time. Shown are the 96-98 data selected by the dipole analysis (points) overlaid with four different likelihood functions, generated with (a)A=1.0 and Δm_s=25 ps^{-1}, (b)A=0.6 and Δm_s=7 ps^{-1}, (c)A=0.6 and Δm_s=12 ps^{-1} and (d)A=1.8 and Δm_s=17 ps^{-1}.**

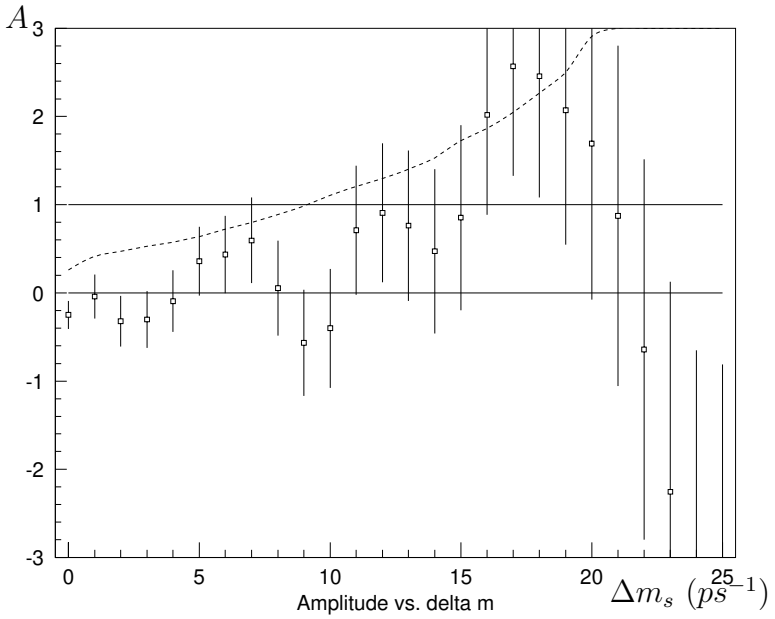


Figure 72: The amplitude fit with a cut at t_{rec}=0.2ps^{-1}.

6.1.1 B_d^0–$\overline{B_d^0}$ Mixing

B_d^0–$\overline{B_d^0}$ mixing provides a good check of the B_s^0–$\overline{B_s^0}$ mixing results, as the data analysis is essentially identical. The difference in the fitting procedure is that Δm_s is fixed and -log$\mathcal{L}$ is minimized with respect to Δm_d. The B_d^0 mixing frequency has been measured within 3%, which gives a well measured calibration point.

Fitting for the B_d^0–$\overline{B_d^0}$ mixing frequency yields $\Delta m_d = 0.537 \pm 0.030$ ps^{-1} (statistical error only). This value is within 1.8σ of the latest world average value of 0.485 ± 0.015 ps^{-1} [90]. Fig. 73 shows the amplitude fit for the full 96-98 data set in the charge dipole analysis. The measured amplitude near the world average is A = 0.940± 0.064 at $\Delta m_d = 0.48$ ps^{-1}, consistent with the expectation of A = 1.

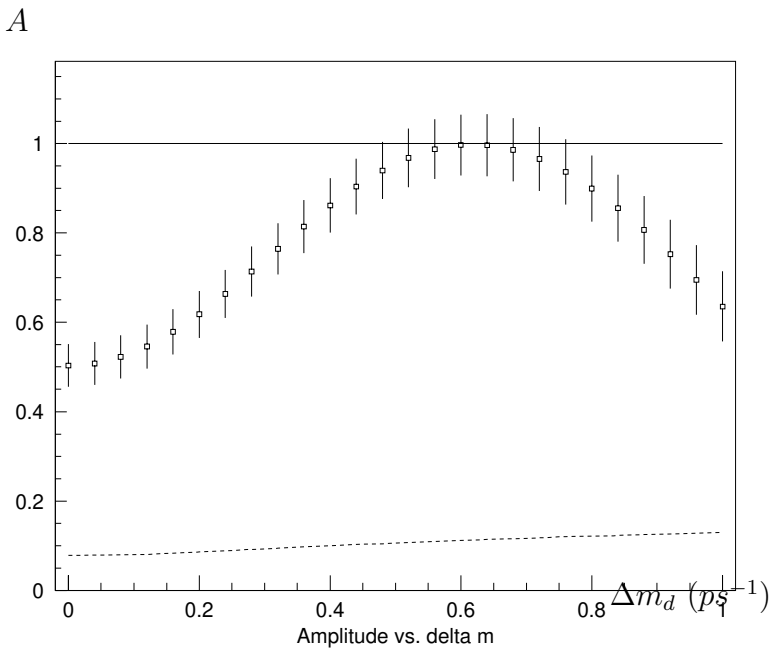


Figure 73: The amplitude fit for the full data set 96-98 in the charge dipole analysis.Δm_s is fixed and log$\mathcal{L}$ is minimized for A at each value of Δm_d.

6.2 Systematic Errors

In order to estimate the size of systematic errors introduced by physics modelling, detector resolution and other uncertainties, the effect of varying the input parameters on the amplitude is studied. In all of the current B_s^0–$\overline{B_s^0}$ mixing analyses, the systematic errors on the amplitude are much smaller than the statistical ones.

The systematic error σ_A^{syst} is determined using the following formula [89]:

$$\sigma_A^{syst} = A^{new} - A^{nom} + (1 - A^{nom})\frac{\sigma_A^{new} - \sigma_A^{nom}}{\sigma_A^{nom}}. \tag{6.1}$$

A^{nom} and σ_A^{nom} are the measured amplitude and its error with the nominal set of input parameters, A^{new} and σ_A^{new} are the new amplitude and error obtained after one or more input parameters are varied. The systematic error has to be calculated

for every value of Δm_s considered, and depends on both the amplitude and the statistical uncertainty.

The main source of systematic error is our limited knowledge of the exact composition of the selected data sample, i.e., the errors on the production fractions of the b hadrons, which originate in uncertainties on measurements of branching ratios and theoretical assumptions (see Appendix A). For this analysis, the following uncertainties on the b hadron production fractions were used[76]:

$$f_{B_s^0} = (9.8 \pm 1.2)\%$$

$$f_{b-baryon} = (10.3 \pm 1.8)\%$$

$$f_{B_d^0} = f_{B^+} = (39.9 \pm 1.1)\%$$

Other sources of systematic error that were considered for this analysis are discussed in the next sections.

6.2.1 B Hadron Lifetimes

Since the analysis is highly sensitive to any time dependence of relevant parameters like the sample composition, the effect of b hadron lifetimes has to be taken into account, in addition to the errors on the time integrated quantities like the b hadron fractions.

The current world averages and their errors used to estimate the systematic errors are summarized in table 11 [92].

6.2.2 Errors on the Proper Time Resolution

The proper time resolution σ_t determines the rate at which the uncertainty on the individual amplitude measurements grows as a function of Δm_s. Recall that

Table 11: b hadron lifetimes.

b hadron	lifetime
B_d^0	(1.542 ± 0.029) ps
B_s^0	(1.464 ± 0.057) ps
B^+	(1.656 ± 0.025) ps
b baryon	(1.208 ± 0.051) ps

the proper time resolution is a function of proper time itself. For example, between 0.2 and 0.5 ps the average proper time resolution is 0.06 ps for the 60% core fraction of the events, and 0.18 ps for the 40% tail fraction of the events (see Fig. 59 in the last chapter). For proper times > 2.5 ps, the average proper time resolution is 0.32 ps for the 60% core fraction of the events, and 0.89 ps for the 40% tail fraction of the events. For errors on the proper time resolution, the two components of σ_t, σ_L and $\sigma_{\gamma\beta}$, are considered separately.

As already discussed, σ_L is estimated on an event-by-event basis, by scaling tracking errors according to fits to the core- and tail fractions. The Monte Carlo tracking has been tuned to the track impact parameter distribution in the data. The amount of systematic error that needs to be assigned to the decay length resolution was estimated from τ pair events [77], in particular the 3-prong τ decays, which can be selected such that backgrounds become negligible. The error on the vertex resolution is measured using the width of the negative tail of the 3-prong τ vertex decay length distribution. In the MC, the width of the vertex resolution is $202 \pm 8\mu$m and in the data it is $207 \pm 14\mu$m. The difference is $5 \pm 16\mu$m, so for the purpose of estimating the systematic error on σ_L an uncertainty of 10% was assumed.

To estimate the error on $\sigma_{\gamma\beta}$, on the other hand, MC and data were studied

for the amount of unassociated cluster energy in the electromagnetic calorimeter. Comparing the data with $q\overline{q}$ MC the average neutral energy was found to be 15% too low in the data. As the calorimetry component of the observed B energy is about 20%, the uncertainty in the B energy due to calorimetry is about 3%. Combining the error in quadrature with the 3% uncertainty due to tracking the total uncertainty becomes 4%. A very conservative error estimate of 20% was used for the analysis. In addition, the uncertainties about the resolution model were included by lowering the core fractions in the parametrization of systematic shifts in proper time reconstruction from 60% to 55%.

6.2.3 Initial and Final State Tag Purity

The uncertainty on the initial state tag purity can be estimated in the data by comparing the results of the tag in both hemispheres: from the measured number of events with a a) b in both hemispheres, b) $\overline{b}$ in both hemispheres and c) b in one and a $\overline{b}$ in the other hemisphere, the correct tag probability can be determined analytically. It is $(70.2 \pm 0.75)\%$ for b-quarks and $(70.6 \pm 0.75)\%$ for $\overline{b}$-quarks. The corresponding Neural Net correct tag probability outputs are 70.5% and 70.7%, respectively. In order to account for uncertainties in the initial state tagging the above errors were applied.

Systematic uncertainties in the charge dipole final state mistag modelling are obtained by varying the final state mistag probability derived from the MC simulation by 2.5%. This is a conservative estimate taken from comparisons between the tagging power in the data and Monte Carlo. The average mixed fraction is 0.427±0.003 in the data and 0.433±0.001 in the MC. Therefore the difference in mixed fraction is 0.006±0.001. From that the difference in mistag fraction can be calculated to be 0.012±0.002, i.e. the difference in average mistag fraction between

data and MC is 1.2%. As a further test on the charge dipole final state tagging power, see the next section.

Not mentioned in the discussion so far are two minor sources of systematic error: the light quark background, which is at the 1% level, and the effect of uncertainties due to the assumed value of Δm_d.

Light quark background events are concentrated at small proper times. To be conservative, the light quark fraction was varied by $\pm 20\%$, motivated by the maximum amount of disagreement found between data and Monte Carlo in the distributions of reconstructed B decay proper time.

The effect of incorrect values of Δm_d on the Δm_s amplitude fit is negligible but taken into account by applying the following value: $\Delta m_d = 0.480 \pm 0.020$ ps^{-1}.

Fig. 74 shows the impact of each of the systematic errors on the amplitude. The solid curves correspond to the positive variation of the parameter, while the dotted curves correspond to the negative variation. The largest systematic effects are due to the uncertainty in the B_s^0 production fraction and boost resolution. Note that all systematic errors are very small compared to the statistical errors.

6.3 Does the Dipole Analysis work for B_s^0?

It is important to test if the dipole analysis, which is done on a data sample of mostly B_d^0 decays, works for B_s^0 decays. As the MC information used to model B_s^0 decays is incomplete and relies on theoretical assumptions it would be interesting to probe the dipole strength in a sample that is clearly enhanced in B_s^0 decays. Naturally, the statistics in such a sample is bound to be very low.

One possibility to select a sample enhanced in B_s^0 decays is to use a Neural Net. The network was trained to select B_s^0 decays. The input parameters that were used

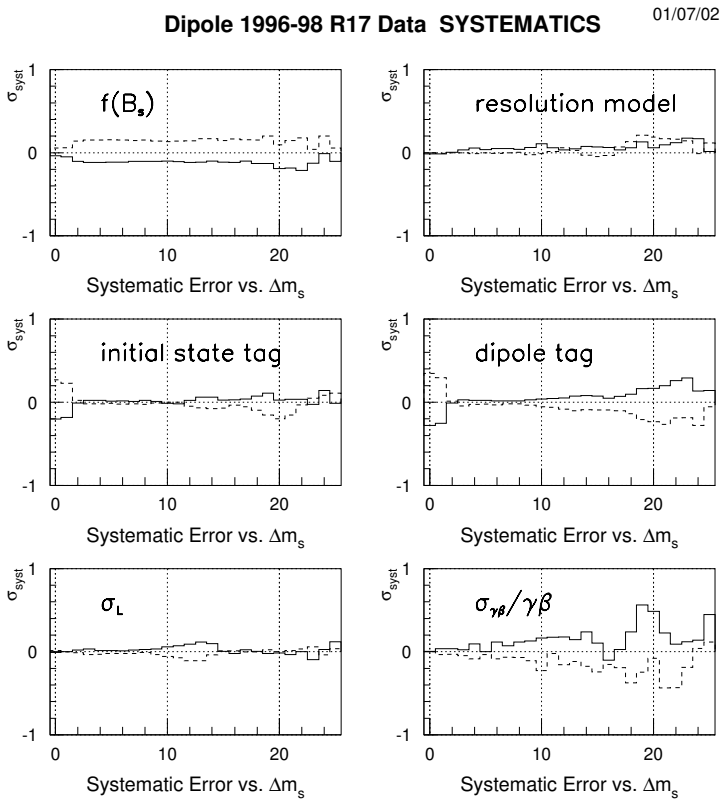


Figure 74: **The systematic errors due to uncertainties in B_s^0 production fraction, resolution model, initial and final state tag, decay length resolution and relative boost resolution. The solid curves correspond to the positive variation of the parameter, while the dotted curves correspond to the negative variation.**

include

- the number of kaons identified by the CRID originating from the tertiary vertex (pointing to a $D_s \to K^+K^-X$ decay).
- the number of identified fragmentation kaons (pointing to B_s^0 production).
- the charge correlation between a lepton from the secondary vertex and a kaon from the tertiary vertex.
- the decay length of the B vertex, B and D vertex masses and multiplicities.

- ϕ candidates from the tertiary vertex, pointing to a $B_s^0 \rightarrow D_s X \rightarrow \phi X$ decay. ϕ candidates were reconstructed combining any two oppositely charged tracks from the tertiary vertex. Assuming kaon masses for the two tracks, the invariant mass m_ϕ was computed and the candidate closest to m_ϕ=1.019 GeV and within the mass range of 1GeV< m_{KK} <1.04 GeV was selected.

The output parameter is the B_s^0 fraction in the sample. Figure 75 shows the Neural Net output for the Monte Carlo (histogram) and the data (points). Data and MC are in reasonable agreement.

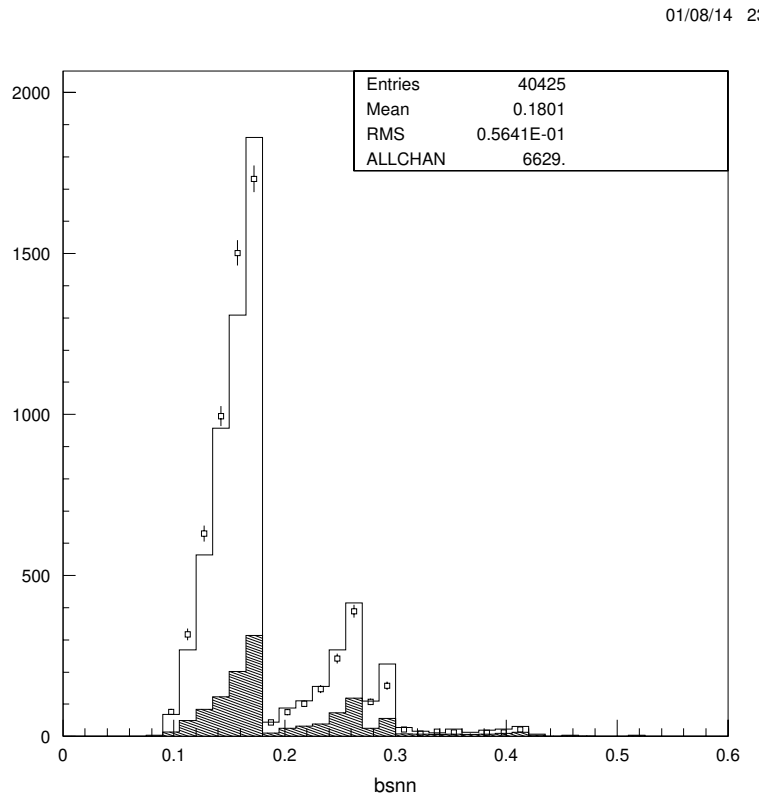


Figure 75: The B_s^0 fraction output parameter. The points are the SLD data, the histogram is the Monte Carlo. The shaded histogram shows the true MC B_s^0 content in every bin.

Fig. 76 shows the b hadron fractions as a function of the Neural Net output, for

B^+, B_d^0, B_s^0 and b baryons separately. The functions are the result of a fit to the MC points.

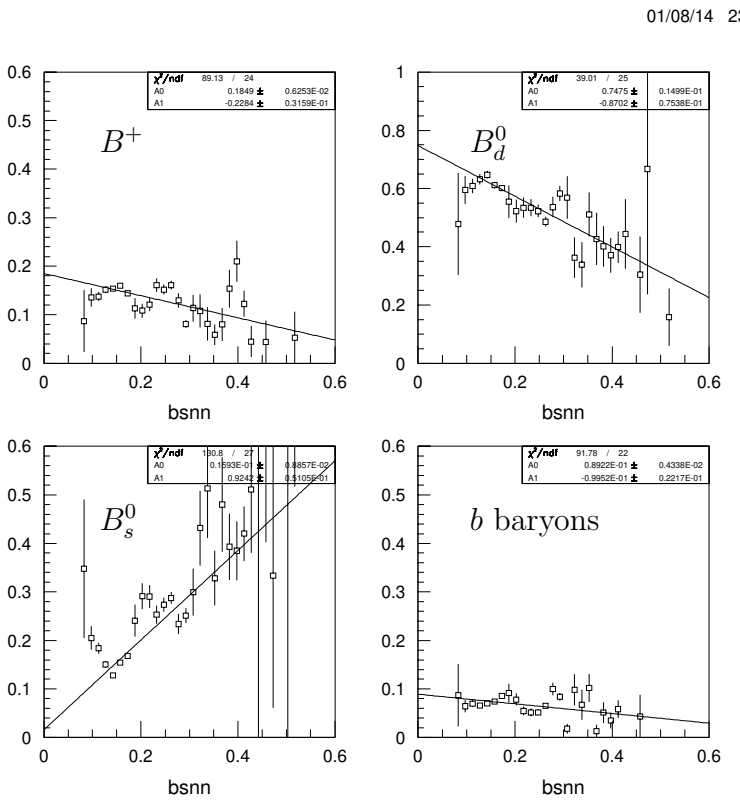


Figure 76: **The MC b fraction versus the B_s^0 fraction NN output parameter, for each b hadron separately. The functions are the result of a fit to the MC points.**

With a NN cut excluding events below 0.35 the average B_s^0 fraction is 40%. 159 data events satisfy this requirement. Fig. 77 shows the dipole separation in data and MC, demonstrating that the dipole tag is sensitive to the B_s^0 component.

The dipole strength can also be tested using the events selected by the lepton+D analysis as a control sample. Events tagged by the lepton+D analysis are well modelled in the Monte Carlo and have a low mistag fraction. For the dipole test, mixed lepton+D events at proper time $<2\text{ps}^{-1}$ were selected. The B_s^0 fraction in this sample is high, as the slowly mixing B_d^0 decays are tagged as mixed generally at

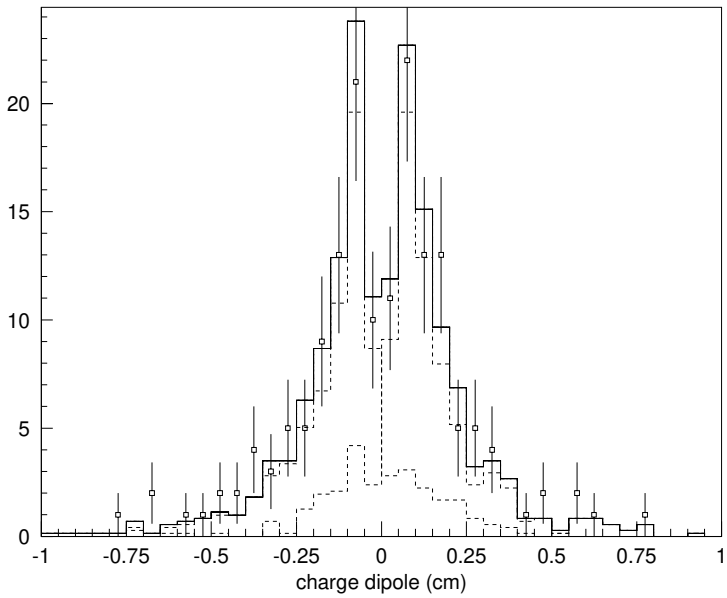


Figure 77: The dipole distribution for a small event sample selected by a NN to enhance the fraction of B_s^0 decays. The points are the data, the histograms MC. Also shown are the contributions from b hadrons containing a b quark (dashed histogram) or a $\bar{b}$ quark (dotted histogram).

higher proper times. MC studies show that the average B_s^0 fraction in this sample is 28%. The semileptonic tag purity is 88% in the selected sample. If available, the dipole tag agrees with the lepton tag (94±1)% of the time in the $q\overline{q}$ Monte Carlo, and (95±2)% in the data, demonstrating that the dipole tag works correctly for a control sample enriched in B_s^0 decays.

In the data, the control sample contains 139 events. Fig. 78 shows its dipole distribution for data and MC, demonstrating the dipole separation.

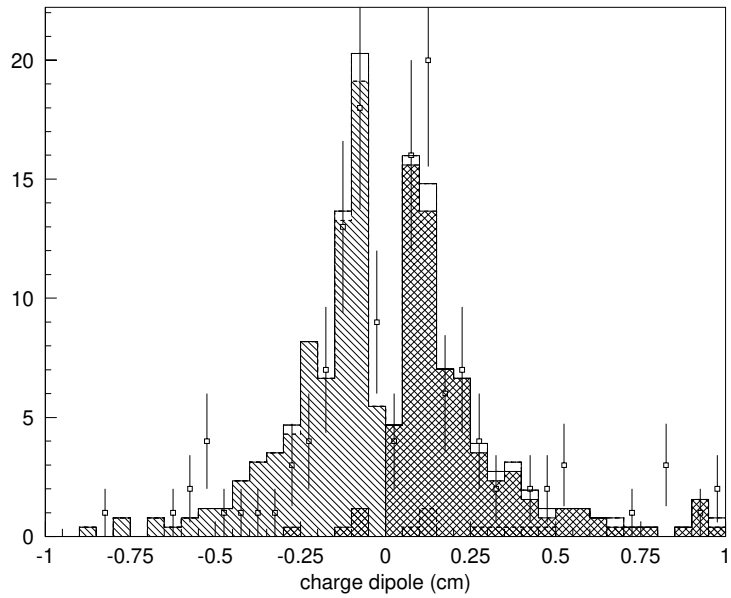


Figure 78: **The dipole distribution for a small event sample of mixed and leptonically tagged decays. The points are the data, the histograms MC. Also shown are the contributions from b hadrons containing a b quark (darkly shaded histogram) or a $\bar{b}$ quark (lightly shaded histogram).**

6.4 Setting a Limit on Δm_s

Fig. 79 shows the amplitude fit for the full 96-98 SLD data set as selected by the charge dipole analysis including systematic errors. The lightly shaded region includes $A \pm 1.645\sigma_A$ (stat) and the darker region is $A \pm 1.645\sigma_A$ (stat+sys). The dashed curve is $1.645\sigma_A$(stat+sys) which represents the sensitivity of the experiment.

The amplitude fit consists of measured values of the amplitude and its uncertainty as a function of Δm_s. The probability density of the amplitude at a given Δm_s is represented by a single Gaussian centered at A with width σ_A. As A is

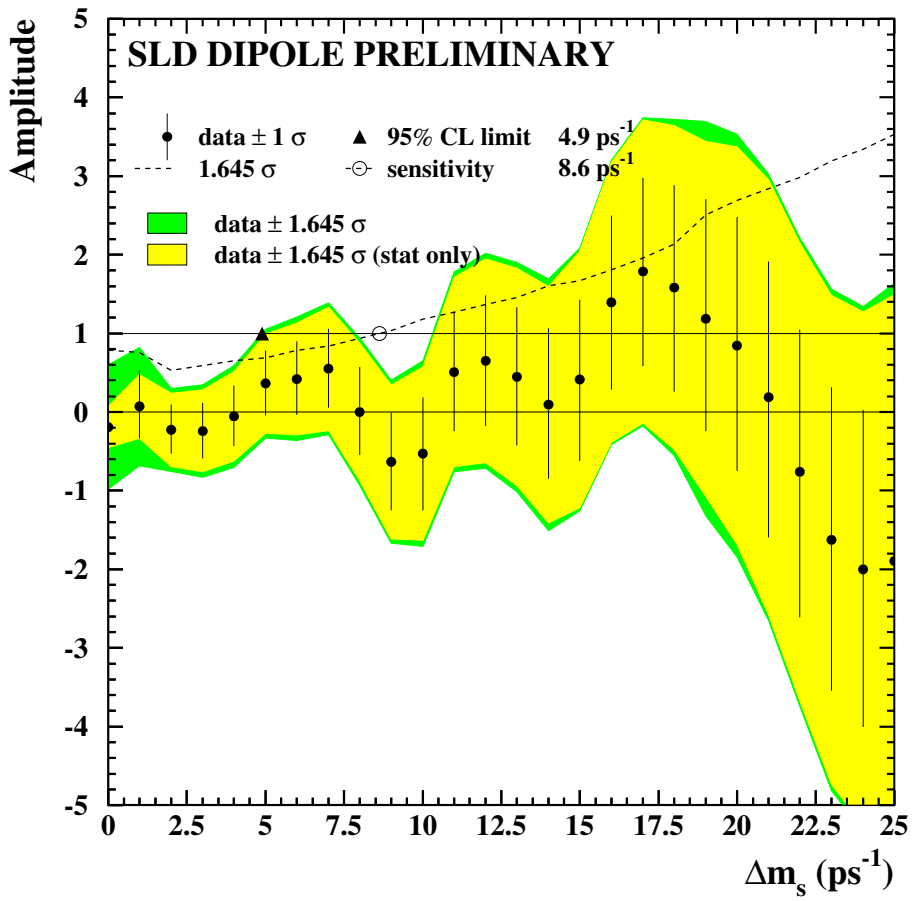


Figure 79: The amplitude fit for the full 96-98 SLD data set from the charge dipole analysis including systematic errors. The solid curve bordering on the lightly shaded region is $A + 1.645\sigma_A$ (stat) and the solid curve bordering on the darker region is $A + 1.645\sigma_A$ (stat+sys). The dashed curve is $1.645\sigma_A$(stat+sys) which represents the sensitivity of the experiment.

simply the normalized Fourier peak of the oscillating signal, the amplitude is expected to be consistent with 1 at the true value of Δm_s and 0 far away from the true value. Therefore, a given value of Δm_s can be excluded at the 95% Confidence Level if [91]

$$\int_1^\infty g(A_m, \sigma_A) dA \ < 5\%. \tag{6.5}$$

This is equivalent to excluding Δm_s at the 95% CL if

$$A_m + 1.645\sigma_A \ < \ 1, \tag{6.6}$$

i.e. in 5% of equivalent experiments one would expect the true value to be excluded. Applying this technique to Fig. 79 the following region of Δm_s can be excluded at the 95% confidence level:

$$\begin{aligned} \Delta m_s \ &< \ 4.9 \text{ ps}^{-1}, \ \text{ and} \\ 7.9 \text{ ps}^{-1} < \ \Delta m_s \ &< \ 10.3 \text{ ps}^{-1}. \end{aligned} \tag{6.7}$$

The sensitivity of the charge dipole analysis in the 96-98 SLD data is 8.6 ps^{-1}.

6.5 Combination of Mixing Analyses

The charge dipole analysis is one of three analyses performed at SLD. All of the SLD analyses use the same initial state tagging technique. The final state flavor tagging and the proper time reconstruction are different for each one of the analyses. The three analyses use discrete sets of events, so that no event is used twice. To determine how to distribute events that were tagged by more than one analysis, the analyses are ranked in order of sensitivity per event. Each analysis is then required to remove those events already selected by a more sensitive analysis.

The two other SLD final state tags used to obtain an SLD average are the 'lepton+D' and the 'D_s+ tracks' analyses. The lepton+D analysis [101] uses the charge of leptons produced in semileptonic B decays to tag the final state b flavor. Topological vertexing is used to partially reconstruct a D vertex downstream of the selected lepton. This vertex is required to have a reconstructed mass < 1.95 GeV. The lepton is selected using a Neural Network. The input parameters include the lepton P_t with respect to the D vertex line of flight. The B decay position is located by intersecting the D vertex momentum with the lepton. A double Gaussian fit to the decay residual yields a core width of 54μm and tail width of 213μm with a 60% core fraction. As in the dipole analysis, the B vertex charge is required to be zero to boost the fraction of B_s^0 decays to 16%. The mistag rate for B_s^0 decays is 4%. The same boost reconstruction algorithm as in the dipole analysis is used. The efficiency achieved is about four times lower than that of the dipole analysis.

The third SLD B_s^0 mixing analysis is the D_s+ tracks analysis [102]. It exclusively reconstructs $\overline{D_s} \to K^{\star 0} K^-$ and $D_s \to \phi\pi^-$ decays, thereby boosting the B_s^0 fraction to 38% overall. The D_s candidates are then intersected with any secondary tracks, to reconstruct a B vertex. The mistag rates are 13% for D_s+ hadrons and 5% for D_s+ leptons. Only 361 events are found in the 96-98 SLD data, however the very high decay length resolution of $50\mu m$ (core) and $151\mu m$ (tail) makes this analysis competitive at high Δm_s. The combined SLD amplitude fit is shown in Figure 80.

The measured values are consistent with $A = 0$ for the whole range of Δm_s up to 25 ps^{-1}. No evidence is found for a preferred value of the mixing frequency. The following range of B_s^0–$\overline{B_s^0}$ oscillation frequencies is excluded at 95% C.L.

$$\Delta m_s < 11.1 ps^{-1}. \tag{6.8}$$

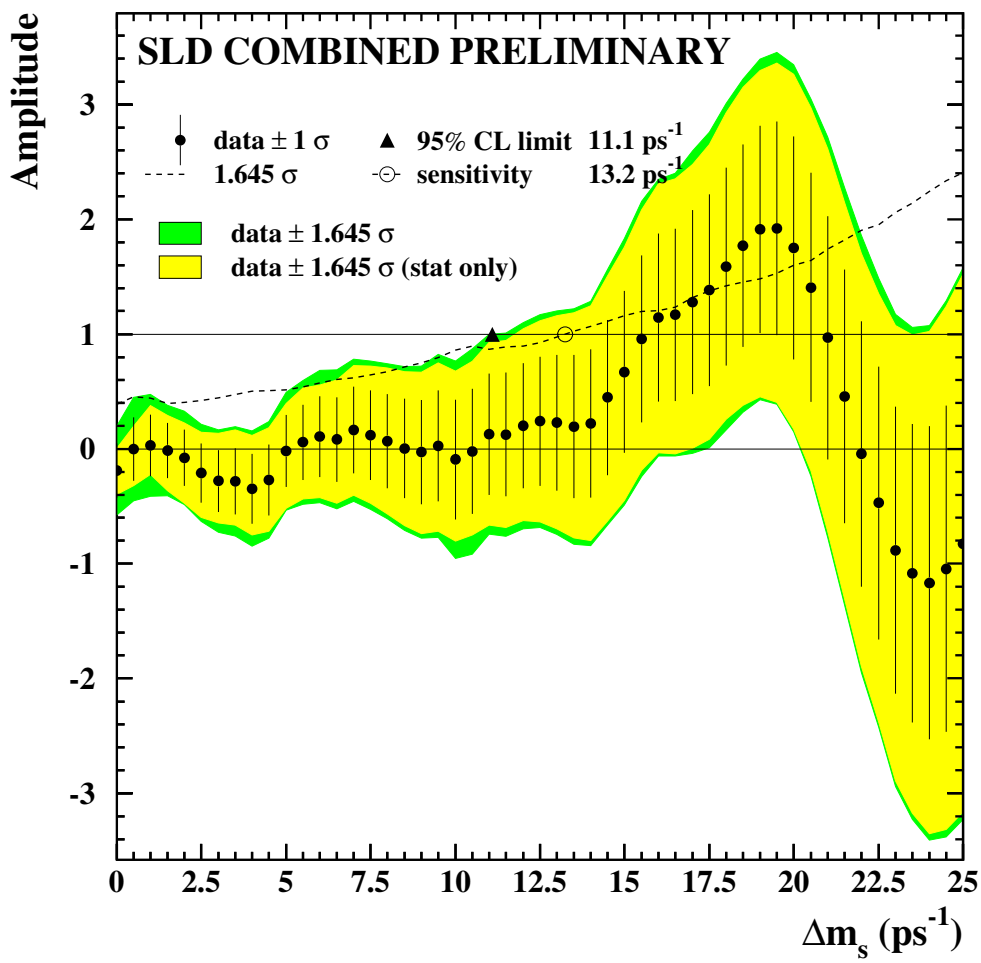


Figure 80: Measured amplitude as a function of Δm_s for the lepton+D, Ds+tracks, and charge dipole analyses combined.

The combined sensitivity to set a 95% C.L. lower limit is found to be at a Δm_s value of 13.2 ps^{-1}. These results are preliminary.

6.6 Conclusions

A combined world average has been compiled by the B Oscillations Working Group from the latest LEP, CDF and SLD results which includes the work presented in this thesis [92]. The averaging procedure takes into account common systematic uncertainties as well as differing parameter values assumed in the various analyses. The combined amplitude plot is shown in Fig. 81. The 95% confidence level allowed region is:

$$\Delta m_s > 14.6\text{ps}^{-1}, \tag{6.9}$$

and the combined sensitivity is 18.3 ps^{-1}. An upper value on the ratio $\left|\frac{V_{td}}{V_{ts}}\right|$ can be extracted, using Eq.1.44, and the current world average of Δm_d. The result is

$$\left|\frac{V_{td}}{V_{ts}}\right| < 0.165. \tag{6.10}$$

The very small statistical error at low Δm_s is mostly due to the LEP analyses. Note that the SLD analyses dominate the structure of the combined amplitude fit beyond $\Delta m_s \approx 16\text{ps}^{-1}$ because of the much better decay length resolution of the SLD vertex detector. This can be seen in a comparison of the different SLD and LEP B_s^0–$\overline{B_s^0}$ mixing analyses at Δm_s=17ps^{-1}, see Fig.82.

The greatest deviation of the amplitude fit from 0 is at Δm_s=17 ps^{-1} with 2.6 standard deviations. Toy Monte Carlo studies determined that the probability of this deviation to be consistent with a fluctuation is $\approx$ 3% [93], too large to conclude at this point that the structure originates from a signal at Δm_s=17 ps^{-1}.

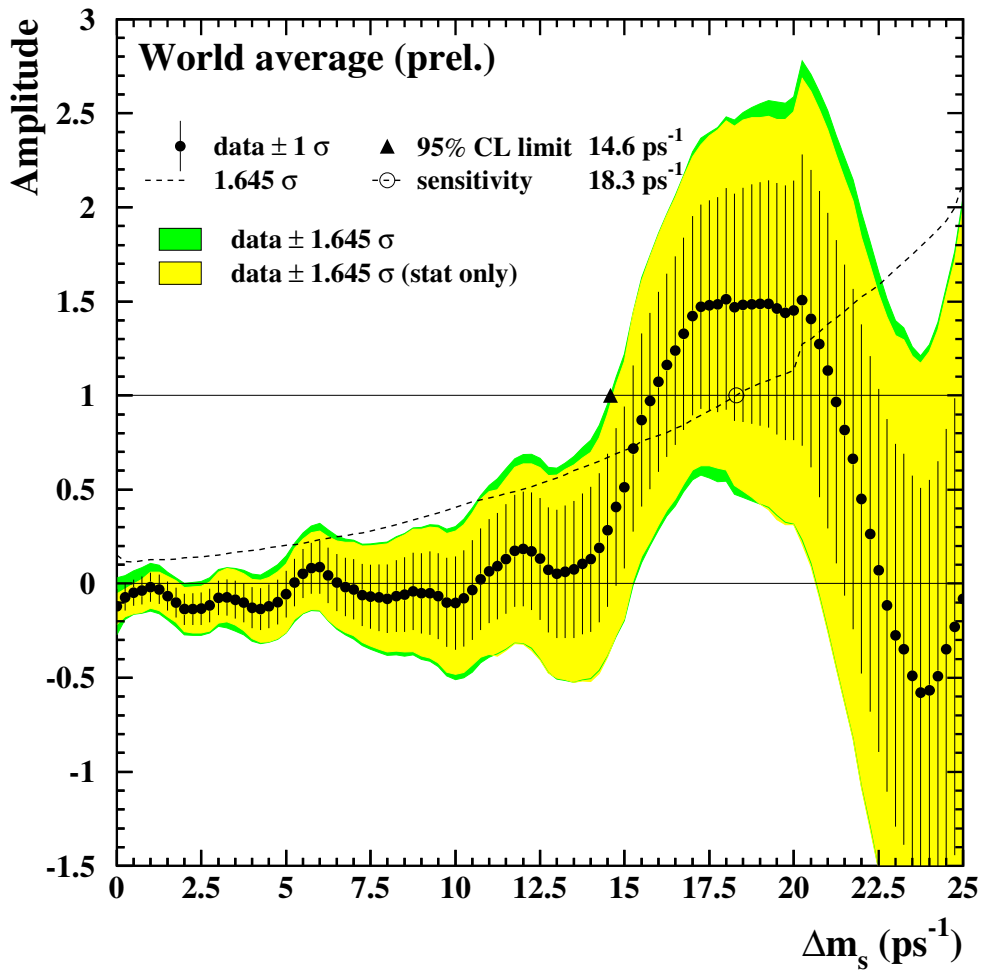


Figure 81: Measured amplitude as a function of Δm_s for the CDF, LEP and SLD Amplitudes combined.

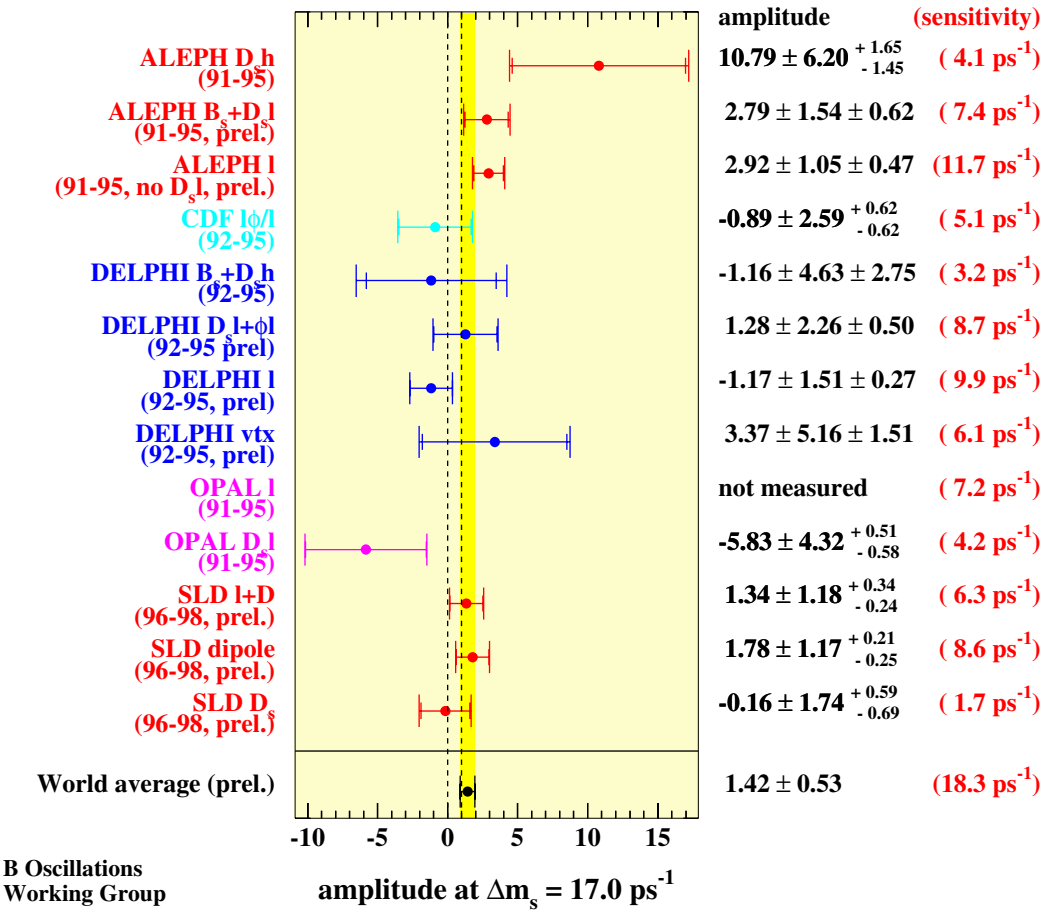


Figure 82: Comparison of σ_A at $\Delta m_s = 17ps^{-1}$ for the CDF, LEP and SLD experiments.

The current experimental measurements of B_d^0 mixing, B_s^0 mixing and sin2β and the current theoretical constraints and uncertainties have been included in global fits to constrain the ρ and η parameters of the CKM matrix in the Wolfenstein parametrization. The resulting constraints on the unitarity triangle are [103]

$$\begin{aligned} \rho &= 0.218 \pm 0.038 \quad \text{and} \\ \eta &= 0.316 \pm 0.040, \end{aligned} \tag{6.11}$$

shown in Fig. 83. For details on the analysis, see [104].

The bands are the $\pm 1\sigma$ allowed regions determined by the measurement of Δm_d, ϵ_K, sin2β and $|V_{ub}/V_{cb}|$. The dashed curve represents the current lower limit on Δm_s at 14.6 ps^{-1}. The portion of the $\rho - \eta$ plane to the left of the dashed curve is excluded by the B_s^0 mixing studies. The apex of the SM unitarity triangle is significantly constrained.

In conclusion, SLD has made significant progress towards a measurement of B_s^0–$\overline{B_s^0}$ mixing. Despite a factor of ∼40 less statistics compared to the LEP experiments, SLD is producing a highly competitive result.

In the next few years, a measurement of B_s^0–$\overline{B_s^0}$ mixing is expected at the Tevatron. For example, the CDF experiment can observe a 5σ signal up to $\Delta m_s \approx$ 40 ps^{-1} using exclusively reconstructed $B_s^0 \to D_s^- \pi^+$ decays (Run II, 2 fb^{-1}) [105]. Combined with the expected precision on the value of Δm_d of better than 1% from measurements at the B factories and at the Tevatron, a measurement of $|V_{td}|$ with a precision better than 5% will be possible. Also, within a few years sin2β is expected to be measured with a precision better than 7% at the B factories. Measurements of Δm_s and sin2β will be the most powerful constraints on the Unitarity Triangle due to their precision and the fact that they are nearly orthogonal. A confirmation of the fundamental Standard Model predictions or a discovery of New Physics is

within reach.

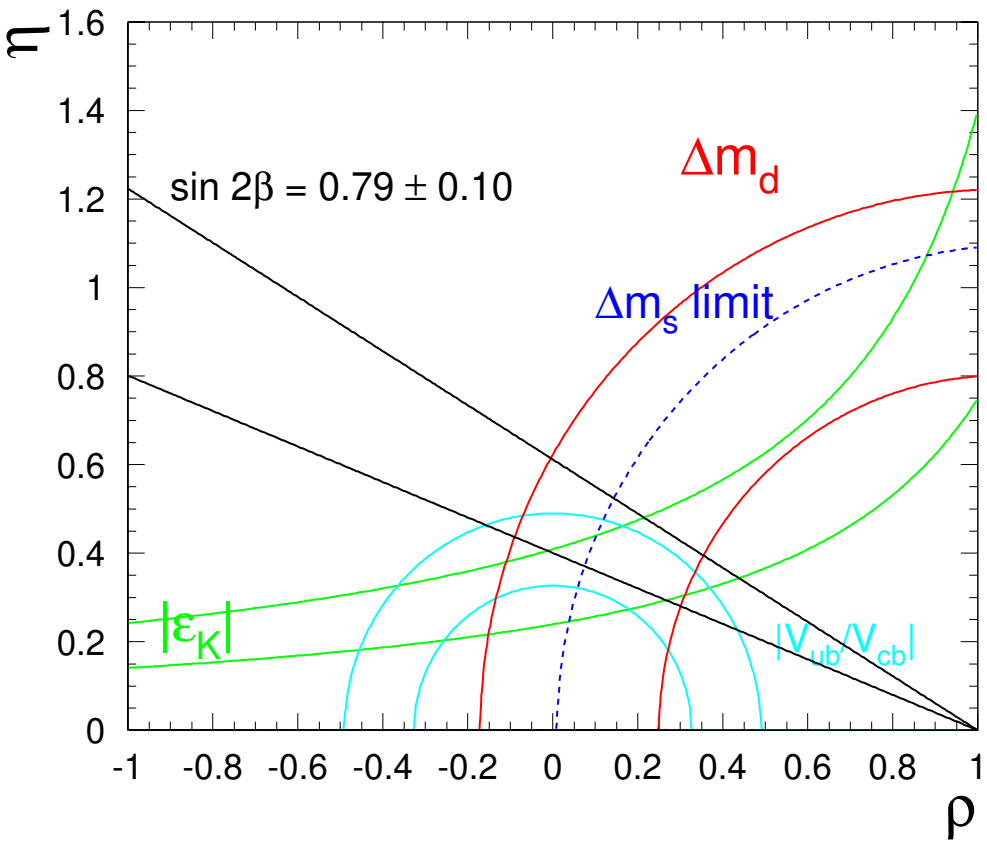


Figure 83: Current constraints on ρ and η.

APPENDIX A

B decay modelling in the SLD Monte Carlo

In this chapter, a summary of the B decay modelling at SLD is given. First, the current knowledge about the production of the different species of B hadrons is summarized, and then the decay model of B hadrons, especially B_s^0 mesons, used in the SLD Monte Carlo is discussed.

For a B_s^0 mixing analysis it is essential to correctly model the composition and decay structure of B hadrons, i.e., to produce the correct rate of visible vertices and the correct number of tracks coming from the vertices, as well as realistic B hadron decay length distributions and kinematics.

A.1 The Composition of B Hadrons

The default composition of B hadrons produced in the SLD Monte Carlo is 41% B^+, 41% B_d^0, 11% B_s^0 and 7% b baryons. The B^0–$\overline{B^0}$ mixing analyses use the latest world average of the B hadron fractions to scale these production rates.

The measurements of branching ratios and the theoretical assumptions used to obtain the B hadron fractions and their estimated systematical errors are summarized here.

The B_s^0 and b baryon fractions $f_{B_s^0}$ and f_{bary} are estimated from measurements of branching fraction products using channels with characteristic signatures. $f_{B_d^0}$ and f_{B^+} are then calculated assuming that $f_{B_s^0} + f_{B_d^0} + f_{B^+} + f_{bary} = 1$ (B_c mesons are neglected), and $f_{B_d^0} = f_{B^+}$, which is motivated by the CLEO measurements of the ratio of $B_d^0\overline{B_d^0}$ to B^+B^- decays at the Υ(4S) and the near equality of the B_d^0 and B^+ masses [4].

Direct information on production rates is available from measurements of branching fraction products using channels with characteristic signatures. For a detailed description, see [76]. At CDF and LEP, the $\overline{B_s^0}$ production rate has been evaluated using events with a D_s^+ accompanied by a lepton of opposite sign in the final state. Since the rate of these events is given by $f_{B_s^0} \times BR(B_s^0 \to D_s^- l^+ \nu X)$, it is necessary to evaluate $BR(B_s^0 \to D_s^- l^+ \nu X)$. This has been done assuming, from SU(3) symmetry, that the partial semileptonic decay widths into D, $D^\star$ and $D^{\star\star}$ final states are the same for all b mesons. Also, isospin symmetry is assumed, and known semileptonic branching ratios of the B_d^0 can, using the lifetime ratio $\tau(B_s^0)/\tau(B_d^0)$, be used to extract semileptonic branching ratios of the B_s^0 mesons. At LEP, the result is $f_{B_s^0} = (12.2^{+4.5}_{-3.1})\%$.

In a similar way, the fraction of b baryons is estimated from the measured production rates of $\Lambda_c^+ l^-$ [98] and $\Xi^- l^-$ [99] final states, yielding, respectively, $f_{\Lambda_b^0} = (11.6^{+4.6}_{-3.1})\%$ and $\Xi_b^- = (1.1^{+0.6}_{-0.4})\%$. The value for BR($\Lambda_b^0 \to \Lambda_c^+ X l^- \overline{\nu_l}$) has been obtained considering that there could be only one Λ_c^+ produced in every decay. Similar considerations have been applied to Ξ_b semileptonic decays. The semileptonic decay width $\Gamma(\Xi_b \to \Xi_c X l^- \overline{\nu_l})$ has been taken to be equal to $\Gamma(\Lambda_b^0 \to \Lambda_c^+ X l^- \overline{\nu_l})$. The total b baryon production rate is then: $f_{b-baryon} = (13.7^{+4.8}_{-3.2})\%$. The result has been combined with the direct measurement of $f_{b-baryon} = (10.2 \pm 2.8)\%$ from the number of protons in b events [100]. Finally, the CDF Collaboration has

measured the production rate of b baryons relative to non-strange B mesons using $\Lambda_c^+ e^-$ final states.

Additional information on the production rates can be obtained from measurements of the time-integrated mixing probability of b hadrons. For an unbiased sample of semileptonic b hadron decays with fractions $g_{B_d^0}$ and $g_{B_s^0}$ of B_d^0 and B_s^0 mesons, this mixing probability is equal to:

$$\overline{\chi} = g_{B_s^0}\chi_s + g_{B_d^0}\chi_d \tag{A.1}$$

where χ_d is the the time integrated mixing probability for B_d^0 mesons. As already mentioned, the semileptonic width is assumed to be the same for all b-hadron species, implying $g_i = f_i R_i$, where $R_i = \frac{\tau_i}{\tau_b}$ are the lifetime ratios. This leads to the relation

$$f_{B_s^0} = \frac{1}{R_s}\frac{(1+r)\overline{\chi} - (1 - f_{b-bary}R_{b-bary})\chi_d}{(1+r)\chi_s - \chi_d} \tag{A.2}$$

where $r = R_u/R_d = \tau(B^+)/\tau(B_d^0)$. This is used to extract another determination of $f_{B_s^0}$ from the b baryon fraction, the lifetime ratio average, the world average value of $\chi_d = 0.181 \pm 0.007$ and the $\overline{\chi}$ average of 0.1194 ± 0.0043 [92]. This new estimate of $f_{B_s^0} = (10.1 \pm 1.4)\%$, is then combined with the b-hadron fractions from direct measurements, taking into accound correlations. The final average values and errors are[76]:

$$f_{B_s^0} = (9.8 \pm 1.2)\%$$

$$f_{b-baryon} = (10.3 \pm 1.8)\%$$

$$f_{B_d^0} = f_{B^+} = (39.9 \pm 1.1)\%$$

A.2 The SLD B Decay Model

The SLD B decay model is a combination of the JETSET [58] and CLEO [68] decay models. In addition, specific decay kinematics can be adjusted. If nothing else is specified, a random phase space sampling technique is used. For the decays of the B^+, B_d^0 and B_s^0 the CLEO decay tables were used, whereas b baryons and charmed hadrons follow the JETSET decay model. The known exclusive branching ratios do not make up a complete table of all possible decays, however it is possible to deduce a number of unknown branching ratios from measured ones, using theoretical assumptions like isospin symmetry and factorization (the assumption, for example, that the virtual $W^{\pm}$ particle decays and fragments into hadrons independently of the fragmentation of the remaining quarks from the decaying particle).

Another assumption made is that since vector mesons have three polarizations, the branching fraction to the vector meson is three times the branching fractions for the corresponding pseudoscalar meson. In case the vector meson branching fraction has not been measured, this rule is used to obtain relative rates.

Even after using these assumptions to obtain additional information, the measured and deduced exclusive branching fractions sum up to only 60%. The remaining 40% have to be modelled using inclusive channels of B to D decays. Two kinds of models can be used:

- the B decays to a D and a number of hadrons ($q\overline{q}$ pairs are popped out of the vacuum, hadronized and given kinetic energy until the phase space of the B is used up). This method typically produces relatively low D energy, and a large number of low momentum light hadrons. The rate of producing $D^{\star}$ mesons is adjusted to match the rates measured by CLEO.

- the B decays into a higher mass charm resonance and then into 2-body modes

like $D\pi$ or DK, thereby shifting the D meson to higher energies. The mass of the resonance controls the produced D energy. The remaining quarks in the B decay are hadronized using the first method.

The models are combined with the goal of reproducing measured distributions like the D momentum and multiplicity, $K^{\pm}$ and $\pi^{\pm}$ momentum, and B charged track multiplicities [77]. A typical problem with this procedure is that a low charged multiplicity is observed in the distributions measured at SLD, so that the tracks obtain rather high kinetic energy. This causes disagreement at the low end of the measured $K^{\pm}$ and $\pi^{\pm}$ momentum distributions, and a compromise has to be found. Still, this procedure is preferable to model dependent assumptions about the decays.

The resulting inclusive B decay mode composition is shown in table 12. "B" refers to (B^+, B^0_d, B^0_s and b baryons). For a more detailed discussion of the SLD B decay model, see [77].

Table 12: B Decay Model Composition.

mode	branching ratio
BR($B \to D^0 X$)	0.452
BR($B \to D^{\pm} X$)	0.154
BR($B \to D_s X$)	0.075
BR($B \to 2DX$)	0.178
BR($B \to c$ baryonX)	0.118
BR($B \to J/\Psi X$)	0.023

In the following, only the decay model of B^0_s mesons is discussed, as the composition of the B^0_s decays is of particular interest for the B^0_s–$\overline{B^0_s}$ mixing analyses. Unfortunately, little direct experimental information is available for B^0_s decays. Therefore, the MC simulation is bound to be somewhat model-dependent. It is for

this reason that all information related to the B_s^0 Monte Carlo used in the analysis will have to be tested carefully against the available data.

For the SLD B_s^0 decay Monte Carlo isospin symmetry was assumed so that the B_s^0 decays could basically be modelled according to tabulated measured modes of B_d^0 mesons, including the Isgur-Wise model for semileptonic decays. All D_s production has been tuned by CLEO to match the measured momentum spectrum. Both "upper vertex" and "lower vertex" D_s production is included, (upper vertex refers to the D_s production from the $W^{\pm}$. In the case of lower vertex production only one of the quarks from the $W^{\pm}$ recombines to produce a D_s meson). The SLD B_s^0 decay table breaks up into the following groups of decays:

- 24.7% semileptonic decays: $b \to c\ (l\nu)$.

  This includes 10.8% $D_s X e\nu$, 10.8% $D_s X \mu\nu$, 2.7% $D_s X \tau\nu$ and 0.4% $\Lambda_c l\nu$.

- 9.1% $b \to$ charm baryons.

- 18% $b \to c\ (s\overline{c})$ (upper vertex D_s production).

  A major source of double charm production.

- 0.8% $b \to c$ (lower vertex D_s production).

- 0.6% $b \to c\ (d\overline{c})$,

  modelled based on results reported in [78].

- 6.8% $b \to c\ (d\overline{u})$. (exclusive channels)

- 38.3% $b \to c\ (d\overline{u})$ and 5% $b \to c\ (s\overline{u})$ ("inclusive unknowns"),

  modelled using a combination of methods as explained above.

The charge dipole B_s^0–$\overline{B_s^0}$ mixing analysis, relies on the $b \to c$ cascade structure. In particular, it tags the flavor of the B_s^0 by reconstructing the charges of secondary

(B) and tertiary (D) vertices. This works best in the case of single charm decays, e.g. $B_s^0 \to D_s X$. On the other hand, double charm decays like $B_s^0 \to D_s \overline{D_s} X$ are problematic, as one of the D vertices is easily mistaken for a B vertex. Table 13 displays the branching ratios of single and double charm decays of the B_s^0 meson, as modelled in the SLD Monte Carlo. Not included in this table are the contributions from decays into c baryons and the inclusive unknowns.

Table 13: Inclusive mode composition of the SLD B_s^0 decay MC.

Mode	branching ratio
BR($B_s^0 \to D_s^{(*)} D_s^{(*)} X$)	0.03
BR($B_s^0 \to D_s^{(*)} D^{(*)} X$)	0.14
BR($B_s^0 \to D^{(*)} D^{(*)} X$)	0.004
BR($B_s^0 \to D_s^{(*)} X$)	0.26
BR($B_s^0 \to D_s^{**} X$)	0.05
BR($B_s^0 \to D^{(*)} X$)	0.003

A P P E N D I X B

Ghost Track Algorithm

Monte Carlo studies show that in B decays producing a single D meson the cascade D decays on average $4200\,\mu$m from the IP, while the intermediate B vertex is displaced on average only $46\,\mu$m transversely from the line joining the IP to the D decay vertex. This kinematic stretching of the B decay chain into an approximately straight line is exploited by the ghost track algorithm. It has two stages and operates on any given list of tracks. First, the best estimate of the straight line from the IP directed along the B decay chain is found. This line is promoted to the status of a track ("ghost track") by assigning it a finite width. Secondly, the selected tracks are vertexed with the ghost track and the IP to build up the decay chain along the ghost direction. Both stages are now described in more detail.

Given a set of tracks in a hadronic jet or hemisphere a new track G is created with the properties that it is a straight line from the IP directed along the jet or thrust axis and has a constant resolution width of $25\,\mu$m in both xy and z. For each track i a vertex is formed with track G and the vertex location $\mathbf{r_i}$, fit χ^2_i and L_i are determined (L_i is the longitudinal displacement from the IP of $\mathbf{r_i}$ projected onto the direction of track G). This is calculated for each of the tracks and the

summed χ^2 is formed:

$$\chi^2_S = \sum_i \begin{array}{ll} \chi^2_i & \mathrm{L}_i \geq 0.0 \\ (2\chi^2_{0i} - \chi^2_i) & \mathrm{L}_i < 0.0 \end{array} \tag{B.1}$$

where χ^2_{0i} is the χ^2_i of track i to track G determined at $\mathrm{L}_i = 0$ rather than at the best fit vertex location. The aim is to construct this quantity, χ^2_S, such that when the direction of G is varied the minimum of χ^2_S provides the best estimate of the B decay direction. If the initial direction is a relatively long way from the B line of flight, some or all of the decay tracks may vertex with G with a negative value of L_i. In this case the $2\chi^2_{0i} - \chi^2_i$ term above helps to push track G towards the B flight path as χ^2_S is minimized. This first minimization using equation B.1 is designed for this purpose. (Note that the contribution of each track as χ^2_S is minimized changes in a continuous manner even if L_i changes sign since $\chi^2_i = \chi^2_{0i}$ at $\mathrm{L}_i = 0$.)

The value of χ^2_S is recalculated as track G is rotated (about the pivot at the IP) incrementally in ever decreasing angular steps $\delta\theta$ and $\delta\phi$ until the minimum is found within the required precision (< 0.1 mrad, i.e. within 1 μm at 1 cm from the IP). The width of track G is set such that the maximum $\chi^2_i = 1.0$ for all tracks with $\mathrm{L}_i > 0$ (if this is less than 25μm, it is restored to 25μm). The track G is now consistent with all potential B decay candidate tracks ($\mathrm{L}_i > 0$) at the level $\chi^2_i \leq 1.0$. In other words, the new width of G measures the degree to which the tracks conform to a straight line decay chain. A second iteration in $\delta\theta$,$\delta\phi$ now takes place with the summed χ^2 redefined as:

$$\chi^2_S = \sum_i \begin{array}{ll} \chi^2_i & \mathrm{L}_i \geq 0.0 \\ \chi^2_{0i} & \mathrm{L}_i < 0.0 \end{array} \tag{B.2}$$

which is not sensitive to any spurious background track with a negative value of L_i which might otherwise perturb the direction of track G. After finding the new

minimum of χ^2_S the width of G is again recalculated such that $\chi^2_i \leq 1.0$ for all tracks i with $L_i > 0$. Again this width is required to be at least 25 μm. Track G is now directed along the best guess of the B decay line of flight and has a width such that it is consistent with potential B decay tracks in the jet, track G is now called the "ghost" track.

The second stage of the algorithm begins by defining a fit probability for a set of tracks to form a vertex with each other and with the ghost track (or IP). This probability then measures the likelihood of the set of tracks both belonging to a common vertex *and* being consistent with the ghost track (or IP) and hence forming a part of the B decay chain. These probabilities are determined from the fit χ^2 which is in turn determined algebraically from the parameters of the selected tracks and the ghost track (or the $7 \times 7 \times 30 \mu m^3$ ellipsoid assumed for the IP). The earlier requirement that each $L_i > 0$ track makes a $\chi^2_i \leq 1.0$ with the ghost track has the effect that the fit probabilities have the desired property of having an approximately flat distribution from 0.0 to 1.0 for genuine vertices, independent of both multiplicity and decay length. This property also relies on the choice of the number of degrees of freedom as 2N−2 (or 2N) when fitting N tracks together with the ghost track (or IP). Fake vertices peak at probability close to 0.0.

For a set of N tracks, there are initially N+1 candidate vertices (N secondary vertices made up of only one track and the ghost track, and a bare IP). A matrix of track i – track j associations is constructed to store the calculated probabilities of each candidate vertex pair fitted together with the ghost track. A further column and row is added to the matrix to store the probabilities of each track fit with the IP ellipsoid. The upper triangle of the matrix (i.e. the ij $(i < j)$ elements) stores the probabilities while the lower triangle (initialized with ij $(i > j)$ elements set to 0.0) indelibly records which tracks (and IP) have been assigned together in a common

vertices as the algorithm progresses. Once the upper triangle has been filled, the highest probability in the matrix table is found and the corresponding candidate vertex pair are from then on tied together in a new candidate vertex for all future computations by flagging the corresponding lower triangle elements of the matrix with non-zero values. The upper triangle of the matrix is now refilled taking into account the associations that have so far been made, the new maximum probability is found, and the corresponding subset of the tracks and IP is tied together. At each iteration of combining the maximum probability matrix element contributors, the number of candidate vertices decreases by one. The iterations continue until the maximum probability is less than 1%. At this point the tracks and IP have been divided into unique subsets by the associations thereby defining topological vertices.

BIBLIOGRAPHY

[1] W. N. Cottingham, *An Introduction to the Standard Model of Particle Physics* Cambridge Univ. Press (1998).

[2] J. Donoghue, E. Golowich, B. R. Holstein (Massachusetts U., Amherst). *Dynamics of the Standard Model* Published in Cambridge Monogr.Part.Phys.Nucl.Phys.Cosmol.2:1-540, 1992.

[3] S.L. Glashow, J. Iliopoulus, and L. Maiani, *Phys. Rev.* **D2** 1285 (1970).

[4] D.E. Groom et al. *The European Physics Journal* **C15** 1 (2000).

[5] C. Quigg, *Gauge Theories of the Strong, Weak, and Electromagnetic Interactions* Frontiers in Physics, Lecture Note Series 56 (1983).

[6] N. Cabibbo, *Phys. Rev. Lett.* 10 (1963) 531.

[7] M. Kobayashi and K.Maskawa, *Prog. Theor. Phys* 49 (1973) 652.

[8] L. Wolfenstein, *Phys. Rev. Lett.* 51 (1983) 1945.

[9] A. J. Buras, M. Jamin, P. H. Weisz, *Nucl. Phys.* **B347** 491 (1990).

[10] A. J. Buras, R. Fleischer, *Phys. Lett.* **B360** 138 (1995).

[11] T. Inami and C.S. Lim, *Prog. Theor. Phys. 65* 297 (1981)

[12] B Oscillations Working Group, see http://www.cern.ch/LEPBOSC

[13] JLQCD Collaboration (N. Yamada et al.), KEK-CP-113, hep-lat/0110087 (2001).

[14] H. Leutwyler and M. Roos, *Z. Phys.* **C25** 91 (1984).

[15] R. Poling, *Int. J. Mod. Phys.* **A15S1** 53 (2000).

[16] D. G. Cassel, *Presented at the XX International Symposium on Lepton and Photon Interactions at High Energies* July 23-28, 2001, Rome (to be published).

[17] H. Abramowicz et al. *Z. Phys.* **C15** 19 (1982)

[18] S.A. Rabinowitz et al. *Phys. Rev. Lett. 70* 134 (1993)

[19] T.M. Aliev et al. *Yad. Fiz.* **40** 823 (1984).

[20] M. Bauer, B. Stech, and M. Wirbel *Z. Phys.***C29** 637 (1985).

[21] M. Narain, *Presented at Symposium on Heavy Flavors* Santa Barbara 1997. Published in *Santa Barbara 1997, Heavy Flavor Physics* 379.

[22] G. Redlinger, *Talk at KAON99* Chicago, June 1999.

[23] Z. Ligeti, hep-ph/9908432 *Presented at KAON99* Chicago, June 1999.

[24] J.H. Christenson, J.W. Cronin, V.L. Fitch and R. Turlay, *Phys. Rev. Lett.* **13** 138 (1964).

[25] G. Hamel de Monchenault *Presented at the International Europhysics Conference on High Energy Physics* July 12-18, 2001, Budapest (to be published).

[26] S. Olsen *Presented at the XX International Symposium on Lepton and Photon Interactions at High Energies* July 23-28, 2001, Rome (to be published).

[27] *The BABAR Physics Book* SLAC-R-504 (1998).

[28] R. Aleksan, B. Kayser, D. London, *Phys. Rev. Lett.* **73**18 (1994).

[29] G. Barenboim, M. Raidal DESY 99-022

[30] P. Ball, R. Fleischer CERN-TH/99-377

[31] Y. Nir *Phys. Lett.* **B 327** 85 (1994).

[32] T. Banks, Y. Grossman, E. Nardi and Y. Nir *Phys. Rev.* **D 52** 5319 (1995).

[33] R. Aleksan, A. Le Yaouanc, L. Olivier, O. Pene and J.C. Raynal *Phys. Lett.* **B316** 567 (1993).

[34] I. Dunietz, R. Fleischer, U. Nierste hep-ph/0012219

[35] Y. Grossman hep-ph/9603244

[36] J. Kent et al., SLAC-PUB-4922 (1989).

[37] P. C. Rowson et al., SLD Note 264 (1999).

[38] M. Fero et al., SLD Physics Note 50 (1996).

[39] H. R. Band, G. Mitchell, R. Prepost and T. R. Wright *Nucl. Instr. Meth.* **A400** 24 (1997).

[40] K. T. Pitts, Ph.D. Thesis, University of Oregon, Eugene (1994), SLAC-R-446.

[41] C. J. S. Damerell et al., *Nucl. Instr. Meth.* **A288** (1990) 236.
C. J. S. Damerell et al., *Nucl. Instr. Meth.* **A400** (1997) 287.
K. Abe et al., SLAC-PUB-7385 (1997).

[42] T. Abe et al., SLAC-PUB-8239 (1999).
Presented at the 8^{th} *International Workshop on Vertex Detectors* VERTEX99, Netherlands, June 1999.

[43] M. D. Hildreth et al., *Nucl. Instr. Meth.* **A367** (1995) 111.

[44] J. P.Venuti and G. B. Chadwick, *IEEE Trans. Nucl. Sci.* **NS-36** (1989) 595.

[45] H. Neal, Ph.D. Thesis, Stanford University, Stanford, CA (1995), SLAC-R-473.

[46] D. C. Williams, Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA (1994), SLAC-R-445.

[47] M. D. Hildreth et al., *IEEE Trans. Nucl. Sci.* **NS-42** (1995) 451.

[48] SLD Collaboration, K. Abe et al. *Phys. Rev.* **D53** (1996) 1023.

[49] K. Abe et al. *Nucl. Inst. Meth.* **A343** (1994) 74.

[50] D. Axen et al.*Nucl. Inst. Meth.* **A328** (1993) 472.

[51] S. Gonzalez, Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA (1994), SLAC-R-439.

[52] A. C. Benvenuti et al.*Nucl. Inst. Meth.* **A276** (1989) 94.

[53] E. Iarocci *Nucl. Inst. Meth.* **A217** (1983) 30.

[54] A. Marchioro et al. *IEEE Trans. Nucl. Sci.* **NS-34** (1987) 133.

[55] P. N. Burrows et al., SLD Note 229 (1993).

[56] SLD Collaboration, K. Abe et al.*Phys. Rev.* **D51** (1995) 962.

[57] T. Himel, Ph.D Thesis, SLAC-Report-223 (1979).

[58] T. Sjöstrand, Comp. Phys. Comm. **82**, 74 (1994).

[59] P. N. Burrows, *Z. Phys.* **C41** (1988) 375.
OPAL Collaboration, M. Z. Akrawy et al., *Z. Phys.* **C47** (1990) 505.
P. N. Burrows and H. Masuda, SLD Note 36 (1995).

[60] C. Peterson et al., *Phys. Rev.* **D27** (1983) 105.

[61] SLD Collaboration, K. Abe et al.,*Phys. Rev. Lett.* **79** (1997) 590.

[62] ARGUS Collaboration, H. Albrecht et al., *Z. Phys.* **C54** (1992) 13.
ARGUS Collaboration, H. Albrecht et al., *Z. Phys.* **C58** (1993) 191.

[63] M. Thulasidas, Ph.D. Thesis, Syracuse University, Syracuse, NY (1993).

[64] R. Brun et al., GEANT3, CERN-DD/EE/84-1 (1987).
CERN Application Software Group, GEANT 3.21 Program, CERN Program Library (1994).

[65] W. Nelson, H. Hirayama, D. W. O. Rogers, SLAC-R-265 (1985).

[66] H. C. Fesefeldt, Simulation of Hadronic Showers, Physics and Applications, Technical Report PITHA 85-02, III Physikalisches Institut, Aachen, Germany (1985).

[67] T. J. Pavel, Ph.D. Thesis, Stanford University, Stanford, CA (1997), SLAC-R-0491.

[68] CLEO B decay model provided by P. Kim and the CLEO Collaboration.

[69] B. Barish *et al.*, Phys. Rev. Lett. **76**, 1570 (1996); H. Albrecht *et al.*, Z. Phys. **C58**, 191 (1993); H. Albrecht *et al.*, Z. Phys. **C62**, 371 (1994); P. Avery *et al.*, CLEO CONF 96-28, July 1996; L. Gibbons *et al.*, Phys. Rev. **D56**, 3783 (1997); T.E. Coan *et al.*, CLNS 97/1516; CLEO Collab., CLEO CONF 97-27, Aug. 1997; M. Zoeller, Ph.D. Thesis, SUNY Albany, 1994; X. Fu *et al.*, Phys. Rev. Lett. **79**, 3125 (1997); D. Gibaut *et al.*, Phys. Rev. **D53**, 4734 (1996).

[70] N. Isgur, D. Scora, B. Grinstein, and M.B. Wise, Phys. Rev. **D39**, 799 (1989).

[71] K. Abe *et al.*, Phys. Rev. **D53**, 1023 (1996).

[72] K. Abe *et al.*, Nucl. Inst. and Meth. **A400**, 287 (1997).

[73] Particle Data Group, Phys. Rev. **D54**, Part I (1996).

[74] C. Peterson *et al.*, Phys. Rev. **D27**, 105 (1983).

[75] R. Brun *et al.*, Report No. CERN-DD/EE/84-1, 1989.

[76] The LEP collaborations: ALEPH, DELPHI, L3 and OPAL, the LEP Electrowek Working Group and the SLD Heavy Flavor Group. Combined results on b-hadron production rates and decay properties, June 26, 2001. CERN-EP/2001-050.

[77] B Decay Charm Counting via Topological Vertexing. Dissertation by Aaron Chu, August 2001, SLD. To be published.

[78] The CLEO Collaboration, E.Lipeles et al., "Study of the decays $B^0 \to D^{(\star)+} D^{(\star)-}$", CLEO CONF 97-26.

[79] The CLEO Collaboration, "Exclusive reconstruction of $\overline{B} \to D^{(\star)}\overline{D}^{(\star)}K^-$, Phys.Rev. D62(2000) 032005, hep-ex/0002065.

[80] P. N. Burrows et al. Estimate of Combined triggering and Selection Efficiency for Hadronic events. Technical report, SLD, 1993. SLD Note 229

[81] B.Denby, "Neural Networks in high energy physics: A ten year perspective", Comput. Phys. Commun. 119 (1999) 219

[82] T. R. Wright, Ph.D. Thesis, University of Wisconsin-Madison, (2001)

[83] T. B. Moore, SLD Collaboration Meeting (2001)

[84] D. J. Jackson, Nucl. Inst. and Meth. **A388**, 247 (1997).

[85] D. J. Jackson, SLD Collaboration Meeting (2000).

[86] R. Luchsinger, C. Grab, "Vertex reconstruction by means of the method of Kalman filtering", Comput. Phys. Commun. 76 (1993) 263

[87] T. Moore, Ph.D. Thesis, Yale University (2000).

[88] D. Dong, Ph.D. Thesis, MIT (1999), SLAC-R-550.

[89] H.-G. Moser and A. Roussarie, Nucl. Inst. and Meth. **A384**, 491 (1997).

[90] See A. Golutvin, *Heavy Flavour Physics*, summary talk presented at ICHEP 2000.

[91] DELPHI Collaboration. Search for B_s^0–$\overline{B_s^0}$ oscillations. Phys. Lett. B, 414:382, 1997.

[92] The LEP B Oscillations Working Group. Combined Results on B^0 Oscillations for the summer 2001 conferences, July 2001.

[93] The B_s^0 Oscillation Amplitude Analysis, G. Boix, D. Abbaneo. CERN-EP-99-103, Jul 1999.
JHEP 9908:004,1999
e-Print Archive: hep-ex/9909033

[94] D. Abbaneo *et al.* (LEP Heavy Flavor Steering Group), *Combined Results on b Hadron Production Rates, Lifetimes, Oscillations and Semileptonic Decays*, CERN-EP-2000-096, SLAC-PUB-8492, March 2000.

[95] K. Abe *et al.* (SLD Collaboration), *Measurement of the B^+ and B^0 Lifetimes using Topological Vertexing at SLD*, SLAC-PUB-8206, July 1999, contributed paper # 477 to EPS-HEP99.

[96] K. Abe *et al.*, Phys. Rev. Lett. **80**, 660 (1998).

[97] K. Abe *et al.*, Phys. Rev. **D56**, 5310 (1997).

[98] R. Barate *et al.*, ALEPH Collaboration, Eur. Phys. J. C2 197 (1998).

[99] D. Buskulic *et al.*, DELPHI Collaboration, Phys. Lett. B394 449 (1996).

[100] R. Barate *et al.*, ALEPH Collaboration, Eur. Phys. J. C5 205 (1998).

[101] The SLD Collaboration, SLAC-PUB-8568, August 2000, hep-ex/0011041. Contributed to the XXXth ICHEP 200, Osaka, Japan, July 2000.

[102] C. S. Lin, Ph.D. Thesis, University of Massachusetts Amherst (2001).

[103] See the fit results in http://parodi.home.cern.ch/parodi/ckm/ckm.html.

[104] M. Ciuchini, G.D'Agostini, E. Franco, V. Lubicz, G. Martinelli, F. Parodi, P. Roudeau, A. Stocchi, LAL 00-77, ROME1-1307/00, hep-ph/0012308 (2000).

[105] *B-Physics at the Tevatron,* Draft of the Workshop Report (2001), to be published as a Fermilab Report.

Acknowledgments

Special thanks to Rolf Heuer and Albrecht Wagner, for taking me on as a PhD student. I would like to thank John Jaros, Stephane Willocq and Su Dong for their help, advice and support. Stephane has been a great Mixing Group Leader. Thanks to Dave Jackson, this work profited from many of his and Stephane's ideas.

I would like to thank my parents Inge and Dieter, my partner Alex and my sisters Hanna and Corinna, for their support, understanding and good spirits.

Thanks to Tom Wright, Cheng-Ju Lin, Aaron Chu, Thomas Moore, Chris Damerell and Vera Lüth for help and useful conversations, and to Nicolo, Jodi, Hyejoo, Danning, Peter, Carsten, Gavin and Sean for their companionship. I thank all of the SLD and SLC physicists and technical staff for making the experimental program successful.