Inclusive Rates and Spectra of the Lambda, Cascade, and Omega Hyperons at BaBar

Andrew L. Chien

Stanford Linear Accelerator Center
Stanford University
Palo Alto, CA 94309

SLAC-R-899

Prepared for the Department of Energy
under contract number DE-AC02-76SF00515

Printed in the United States of America. Available from the National Technical Information Service, U.S. Department of Commerce,5285 Port Royal Road, Springfield, VA 22161.

Inclusive Rates and Spectra of the Lambda, Cascade, and Omega Hyperons at BaBar

A dissertation submitted
to the graduate division of the
University of California, Los Angeles
in partial satisfaction for the degree
Doctor of Philosophy in Physics

by

Andrew Lawrence Chien

2008

SLAC-R-899
May 2008

Inclusive Rates and Spectra of the Lambda, Cascade, and Omega Hyperons at BaBar

Andrew L. Chien

Stanford Linear Accelerator Center, Stanford University, Palo Alto CA 94309

Abstract

We employ Runs 1-4 off-peak data sample (about 21.5 fb^{-1}) to produce the current world-best spectra and production rates measurements for three strangely-flavored baryons: the Λ hyperon, the cascade hyperon, and the Ω hyperon. These improved measurements shall enable theoretical and phenomelogical workers to generate more realistic models for the hadronization process, currently one of the unresolved problem areas in the standard model of particle physics. This analysis was conducted using codes from release 16 series.

We report the production rate at 10.54 GeV for the Λ as 0.0900 ± 0.0006(stat.) ± 0.0039(sys.) per hadronic event. Our measured production rate at the same energy for the cascade hyperon is 0.00562 ± 0.00013(stat.) ± 0.00045(sys.) per hadronic event, while that for the Ω hyperon is 0.00027 ± 0.00004(stat.) ± 0.0008(sys.) per hadronic event. The spectral measurements for the respective particles also constitute current world-best measurements.

Contents

1 Introduction 1

2 Some historical and theoretical issues relating to hadronization 3

2.1 Classification of hadronic states and analyses via perturbative field theories . 3
  2.1.1 Group theoretic classification schemes 3
  2.1.2 Description of strong interactions via gauge field theories 4
  2.1.3 Asymptotic freedom in QCD and its implications for the study of hadronization 5
2.2 Some issues related to phenomenological hadronization models 6
  2.2.1 The Field-Feynman model 7
  2.2.2 The relativistic string fragmentation model 9
  2.2.3 The origin of the idea to use strings to model strong interactions . . . 9
  2.2.4 Derivation of the string-based Lund Symmetric Fragmentation Function 12
2.3 Alternative fragmentation functions and models 14

3 The BaBar detector and related hardware and software triggers 16

3.1 The PEP-II B-factory 16
3.2 The BaBar detector 17
  3.2.1 The SVT 17
  3.2.2 The DCH 19
  3.2.3 The DIRC 22
  3.2.4 The EMC 24
  3.2.5 The IFR 24
3.3 L1 trigger system 25
  3.3.1 The DCT 26
  3.3.2 The EMT 27
  3.3.3 The GLT 27
3.4 L3 trigger system 27
3.5 OPR filters and offline reconstruction 29

4 Event selection 30

4.1 Event variables 31
4.2 'Preselection' cuts 33
4.3 Additional event cuts 35
4.4 Comparisons between Data sample and mixed MC sample 39
4.5 Estimated hadronic event purity in data sample 41

5 Tracking corrections for beampipe tracks 47
5.1 BaBar modeling of material interactions in the detectors 47
5.2 BaBar Tracking Group's DCH tracking efficiency corrections 48
5.3 Method for correcting $\pi^{\pm}$ MC/Data tracking differences 51
5.3.1 DCH relative efficiency 51
5.3.2 SVT relative efficiency 57
5.3.3 MC/Data absolute tracking efficiency comparison using τ 3+1 events 57
5.4 Method for correcting $p/\overline{p}$ MC/Data tracking differences 63

6 Hyperon selection and signal extraction 68
6.1 Overview 68
6.2 Λ^0 signal extraction 70
6.2.1 Cuts against Λ candidates 71
6.2.2 Λ efficiency matrix construction 74
6.2.3 Λ efficiency correction and signal fits 78
6.3 Ξ^- signal extraction 83
6.3.1 Cuts against Ξ candidates 85
6.3.2 Ξ efficiency matrix construction 88
6.3.3 Ξ efficiency correction and signal fits 91
6.4 Ω^- signal extraction 94
6.4.1 Cuts against Ω candidates 95
6.4.2 Ω efficiency matrix construction 97
6.4.3 Ω efficiency correction and signal fits 101

7 Hyperon spectra, errors, and comparisons 104
7.1 Λ^0 hyperon results 104
7.1.1 Preliminary Λ^0 and $\overline{\Lambda}^0$ spectra and statistical uncertainties 104
7.1.2 Propagated tracking and efficiency uncertainties 107
7.1.3 MC statistical error consistency check 110
7.1.4 Systematic errors from signal fits and momentum shifts 110
7.1.5 Systematic error contributions from separate $\Lambda^0/\overline{\Lambda}^0$ fits 114
7.1.6 Consistency check using angular bins 115
7.1.7 Consistency check using spectra from separate runs 119
7.1.8 Consistency check using SVT Λ^0 and $\overline{\Lambda}^0$ samples 119
7.1.9 Final Lambda spectrum and rate 123
7.1.10 Comparisons of Lambda spectrum with prior results and with Monte Carlo 126
7.2 Ξ^- hyperon results 130
7.2.1 Preliminary Ξ^- and $\overline{\Xi}^+$ spectra and statistical uncertainties 130
7.2.2 Propagated tracking and efficiency uncertainties 131
7.2.3 MC statistical error consistency check 132
7.2.4 Consistency check using separate $\Xi^-/\overline{\Xi}^+$ fits 135
7.2.5 Consistency check using SVT Ξ^- and $\overline{\Xi}^+$ samples 135
7.2.6 Final Xi spectrum and rate 137
7.2.7 Comparisons of Xi spectra with prior results and with Monte Carlo . 142
7.3 Ω^- hyperon results 142
7.3.1 Preliminary Ω^- and $\overline{\Omega}^+$ spectra and statistical uncertainties 142

7.3.2 Propagated tracking and efficiency uncertainties 146
7.3.3 Consistency check using separate $\Omega^-/\overline{\Omega}^+$ fits 148
7.3.4 Final Omega spectrum and rate 151
7.3.5 Comparisons of Omega rate with prior results and with Monte Carlo . 154

8 Summary and outlook **156**

List of Figures

3.1 Longitudinal cross-section of the BaBar detector. 18
3.2 SVT layers' layout. 19
3.3 Scematic representation of the DCH's longitudinal cross-section. 20
3.4 Schematic wire layout in the first four DCH superlayers. 21
3.5 Schematic representation of the DIRC's fused silica bars and the imaging region. 23
3.6 A longitudinal cross-section of the EMC. Each rectangle in the figure represents one of the 56 crystal rings. 25
3.7 A graphical illustration of L1 trigger primitives. 28

4.1 BaBar run number vs. number of events used for the current event selection studies. 31
4.2 Total event energy vs. R_2, MC mix and data after preselection 34
4.3 Total event energy vs. sphericity, MC mix and data after preselection 34
4.4 Total event energy vs. R_2, τ MC events 36
4.5 Total event energy vs. sphericity, $2-\gamma$ MC 36
4.6 Total event energy vs. R_2, MC mix vs. Data after all cuts 37
4.7 Total event energy vs. sphericity, MC mix vs. Data after all cuts 37
4.8 Reconstructed Y-Z coordinates of MC and data events selected after all cuts. 38
4.9 Vertex-beam radial separation distances and z-coordinates of reconstructed MC and data event vertices 39
4.10 Number of 'good' charged tracks in event comparisons between MC and data 40
4.11 Event sphericity comparisons between MC and data 42
4.12 Event R_2 comparisons between MC and data 43
4.13 Event total energy comparisons between MC and data 44

5.1 Averaged track DCH efficiency using TTG method with relaxed DOCA cut . 50
5.2 MC and data k_{short} mass versus numbers of k_{short} candidates 52
5.3 Averaged pion DCH relative efficiency (transverse momentum variable) . . . 55
5.4 Averaged pion DCH efficiency (lab angle variable) 56
5.5 Averaged pion DCH efficiency (multiplicity variable) 56
5.6 Examples of $\pi^{\pm}$ data/MC correction factors and associated errors. 57
5.7 Averaged pion SVT efficiency (transverse momentum variable) 58
5.8 Averaged pion SVT efficiency (lab angle variable) 59
5.9 Averaged pion SVT efficiency (multiplicity variable) 60
5.10 Pion SVT efficiency correction factors 60
5.11 Total event energy of τ3+1-selected MC and data events 62

5.12 Signal and sidebands in Λ's invariant mass for purposes of subtracting the combinatoric background. 64
5.13 Transverse momentum of protons from Λ decays (MC, data, and sideband-subtracted data) 65
5.14 Averaged proton DCH efficiency (transverse momentum variable) 65
5.15 Averaged proton DCH efficiency (lab angle variable) 66
5.16 Averaged proton DCH efficiency (multiplicity variable) 66

6.1 Beampipe Λ^0 invariant mass 72
6.2 Effect of cuts with respect to daughter SVT/DCH hits and proton momentum on off-peak data Λ^0 candidate acceptance. 73
6.3 Effect of cut on candidate pointing angle and the combined effect of all enumerated cuts with respect to off-peak data Λ^0 candidate acceptance. 73
6.4 2-d histograms of reconstructed and generator MC beampipe Λ^0 75
6.5 Smoothed Λ^0 efficiency grid bin relative statistical errors 77
6.6 Number of bins integrated during Λ^0 efficiency grid smoothing process 78
6.7 Track-corrected and smoothed (also mixed by data event multiplicity fractions) MC Λ^0 and $\overline{\Lambda}^0$ 2-d efficiency grids 79
6.8 Propagated tracking correction error associated with each bin in the smoothed Λ^0 efficiency grid. 79
6.9 A graphical representation of MC Λ^0 momentum/labAngle lab frame to CM frame conversion. 80
6.10 Λ^0+$\overline{\Lambda}^0$ CM momentum-binned invariant mass plots and fits. 81
6.11 CM momentum variations in fitted $\Lambda^0 + \overline{\Lambda}^0$ mass and width 82
6.12 Combinations of Λ^0 and $\overline{\Lambda}^0$ with any other non-repetitive charged track in the event using a π mass hypothesis (after PID cut.) 85
6.13 Effects of dXY cut, 'right sign' cut, and Λ^0 mass cut on Ξ^- candidates . . . 86
6.14 Combined effect of all 'preliminary cuts' on the acceptance of Ξ^- candidates. 86
6.15 Effects of Λ^0 pointing angle cut, proton DCH-hit cut, and π hit cuts on Ξ^- candidates 87
6.16 Combined effect of all specified cuts on the acceptance of Ξ^-+$\overline{\Xi}^+$ candidates from Runs 1-3. 87
6.17 Histograms of reconstructed MC' beampipe' Ξ^- and $\overline{\Xi}^+$ candidates which are truth-associated with MC generator particles. 88
6.18 Histograms of MC generator Ξ^- and $\overline{\Xi}^+$ candidates that decayed inside the PEP-II beampipe. 89
6.19 'Beampipe' Ξ^- and $\overline{\Xi}^+$ efficiency matrices obtained from Monte Carlo. . . . 90
6.20 Relative statistical errors associated with each bin in 'beampipe' cascade hyperon efficinecy matrices 90
6.21 Ξ^- and $\overline{\Xi}^+$ efficiency matrices for the 'beampipe+SVT' Monte Carlo candidate sample. 91
6.22 Fitted invariant mass spectra of efficiency-corrected 'beampipe' Ξ^-+$\overline{\Xi}^+$ from off-peak data 92
6.23 Fitted invariant mass spectra of efficiency-corrected 'beampipe+SVT' cascade hyperons from Runs 1-3 off-peak data (CM momentum $\geq$ 2.50 GeV/c.) . . . 93
6.24 CM momentum variations in fitted $\Xi^- + \overline{\Xi}^+$ mass and width 93

6.25 Invariant mass distribution of reconstructed Ω^-+$\overline{\Omega}^+$ candidates after all 'preliminary' cuts (all off-peak Run2 data) 96
6.26 Results of cuts with resepect to Ω^- daughters' DOCA to reconstructed candidate decay vertex 97
6.27 Results of additional cuts on Ω^- candidates 97
6.28 Invariant mass distribution of Ω^- candidates after cutting against Λ^0's protonic daughters DCH hits. 98
6.29 Invariant mass distribution of Ω^- candidates after all cuts 98
6.30 Distribution of reconstructed MC Ω^- and $\overline{\Omega}^+$ candidates in labAngle and lab momentum variables 99
6.31 Distribution of selected generator MC Ω^- and $\overline{\Omega}^+$ candidates in labAngle and lab momentum variables. 99
6.32 Two-dimensional efficiency matrices constructed from MC Ω^- and $\overline{\Omega}^+$. . . . 100
6.33 Bin statistical erros associated with Ω^- and $\overline{\Omega}^+$ efficiency matrices. 100
6.34 Fitted and efficiency-corrected Ω^- and $\overline{\Omega}^+$ signals. 102
6.35 CM momentum variations in fitted Ω^- + $\overline{\Omega}^+$ mass and width 102

7.1 Preliminary spectrum of Λ^0 and $\overline{\Lambda}^0$ and associated bin statistical errors. . . . 107
7.2 Zone of 'perfectly' correlated errors due to tracking corrections combined with the independent Λ^0 efficiency statistics errors 109
7.3 Pull distribution calculated from Lambda efficiency matrices and the estimated efficiency bin statistical uncertainties (with 1.55 multiplicative factor.) 111
7.4 Fitted widths of Monte Carlo Λ^0+$\overline{\Lambda}^0$ mass signals versus CM momentum. . . 111
7.5 Percentages of reconstructed Monte Carlo Λ^0+$\overline{\Lambda}^0$ that shifted their momentum bin locations relative to their true momentum bins 112
7.6 Fitted Monte Carlo Λ^0+$\overline{\Lambda}^0$ invariant mass signals in p_{cm} bins. 113
7.7 Reconstructed and true Monte Carlo Λ^0+$\overline{\Lambda}^0$ momentum spectrum. 114
7.8 Comparison of the separate Λ^0 spectrum with the $\overline{\Lambda}^0$ spectrum 115
7.9 CM momentum-binned invariant mass plots and fits for off-peak data's Λ^0-only subsample. 117
7.10 CM momentum-binned invariant mass plots and fits for off-peak data's $\overline{\Lambda}^0$-only subsample. 118
7.11 Λ^0 + $\overline{\Lambda}^0$ spectra extracted from different $\cos(\theta_{cm})$ bins and from all angle bins 119
7.12 Λ^0 + $\overline{\Lambda}^0$ spectra extracted from different ϕ_{lab} bins and from all angle bins . . 120
7.13 Λ^0 + $\overline{\Lambda}^0$ spectra extracted from different BaBar Run periods 120
7.14 Cascade hyperon's angular deflection in BaBar's transverse plane as functions of particle $p_\perp$ and XY distance traveled. 122
7.15 Comparison of beampipe Λ^0 spectrum with spectra from Runs 1-4 SVT samples (separted by Runs) 122
7.16 Comparison of beampipe $\overline{\Lambda}^0$ spectrum with spectra from Runs 1-4 SVT samples (separated by Runs) 123
7.17 ISR-corrected inclusive continuum spectrum for Λ^0 plus $\overline{\Lambda}^0$ at 10.54 GeV . . 127
7.18 BaBar inclusive continuum spectrum for Λ^0 plus $\overline{\Lambda}^0$ at 10.54 GeV compared with CLEO and ARGUS spectra 128
7.19 BaBar inclusive continuum spectrum for Λ^0 plus $\overline{\Lambda}^0$ at 10.54 GeV compared with the spectra from the default UCLA generator and BaBar's SP5 generator 129

7.20 Preliminary spectrum of Ξ^0 and $\overline{\Xi}^+$ and associated bin statistical errors. . . 131
7.21 Zone of correlated errors due to tracking corrections, $\Xi^-/\overline{\Xi}^+$ efficiency bin statistics, and beampipe/SVT Λ discrepancy 134
7.22 Pull distribution calculated from Xi efficiency matrices and the estimated efficiency bin statistical uncertainties (with 1.32 multiplicative factor.) 135
7.23 Comparison of the separate Ξ^- and $\overline{\Xi}^+$ spectra with the combined Ξ+$\overline{\Xi}^+$ spectrum 136
7.24 Comparison bewteen the spectrum from our composite beampipe-SVT Ξ^- sample with the full beampipe+SVT Ξ^- spectrum 138
7.25 ISR-corrected inclusive continuum spectrum for Ξ^- plus $\overline{\Xi}^+$ at 10.54 GeV . 141
7.26 BaBar inclusive continuum spectrum for Ξ^- plus $\overline{\Xi}^+$ at 10.54 GeV compared with CLEO and ARGUS spectra 143
7.27 BaBar inclusive continuum spectrum for Ξ^- plus $\overline{\Xi}^+$ at 10.54 GeV compared with the spectra from the default UCLA generator and BaBar's SP5 generator 144
7.28 Preliminary spectrum of Ω^0 and $\overline{\Omega}^+$ and associated bin statistical errors. . . 145
7.29 Zone of correlated errors due to tracking corrections, $\Omega^-/\overline{\Omega}^+$ efficiency bin statistics, and beampip/SVT Λ discrepancy 147
7.30 Comparison of the separate Ω^- and $\overline{\Omega}^+$ spectra with the combined Ω+$\overline{\Omega}^+$ spectrum 148
7.31 Momentum-binned invariant mass plots for the Ω^- subsample and fits. 149
7.32 Momentum-binned invariant mass plots for the $\overline{\Omega}^+$ subsample and fits. 150
7.33 Momentum-binned invariant mass plots for the Ω^- subsample without any efficiency corrections. 151
7.34 Momentum-binned invariant mass plots for the $\overline{\Omega}^+$ subsample without any efficiency corrections. 152
7.35 BaBar inclusive continuum spectrum for Ω^- plus $\overline{\Omega}^+$ at 10.54 GeV compared with the spectra produced by the Pythia/Jetset and the UCLA generators. . 155

List of Tables

4.1 Data samples used in event selection studies. The number of events referenced in the table are the absolute number of events accepted after the specified sets of cuts . . . 31
4.2 SP5 Monte Carlo samples used in event selection studies . . . 32
4.3 Monte Carlo cross sections used in mixture and selection efficiencies taken from 4.2. . . . 33
4.4 Alternative Monte Carlo cross sections used in mixture. . . . 45
4.5 SP5 Monte Carlo process efficiencies with added vertex cuts . . . 45
4.6 Percent standard deviation betwen MC mixture and data . . . 46

5.1 True species of Monte Carlo tracks from k_{short} candidates. . . . 53
5.2 Estimated "ghost" tracks in Monte Carlo and data events selected using our 'Bhabha' criteria. . . . 53
5.3 Definitions of tracks' transverse momentum bins for studies of π's DCH relative efficiency. . . . 54
5.4 Definitions of tracks' cosine(LabAngle) bins for studies of π's DCH relative efficiency. . . . 54
5.5 Number of accepted data and MC events after application of τ 3+1 selection criteria. . . . 62

6.1 Fractions of MC $\mathit{\Lambda}^0$ and $\overline{\mathit{\Lambda}}^0$ rejected/accepted by cuts on efficiency tables. . . 83
6.2 Fractions of MC $\mathit{\Lambda}^0$ and $\overline{\mathit{\Lambda}}^0$ rejected by decay distance cut relative to number of accepted candidates. . . . 84
6.3 Fractions of MC $\mathit{\Xi}^-$ and $\overline{\mathit{\Xi}}^+$ rejected/accepted by cuts on efficiency tables. . 94
6.4 Fractions of MC $\mathit{\Xi}^-$ and $\overline{\mathit{\Xi}}^+$ rejected by decay distance cut relative to number of accepted candidates. . . . 95
6.5 Fractions of MC $\mathit{\Omega}^-$ and $\overline{\mathit{\Omega}}^+$ rejected/accepted by cuts on efficiency tables. . 103
6.6 Fractions of MC $\mathit{\Omega}^-$ and $\overline{\mathit{\Omega}}^+$ rejected by decay distance cut relative to number of accepted candidates. . . . 103

7.1 Fitted numbers of $\mathit{\Lambda}^0$ and $\overline{\mathit{\Lambda}}^0$ and associated statistical uncertainties in p_{CM} bins . . . 106
7.2 Systematic uncertainties in p_{CM} bins due to tracking and efficiency matrices' bin statsitics . . . 108
7.3 Fitted numbers from separate $\mathit{\Lambda}^0$ and $\overline{\mathit{\Lambda}}^0$ samples and associated statistical uncertainties in p_{CM} bins. . . . 116
7.4 Summary of the sources of error for $\mathit{\Lambda}^0/\overline{\mathit{\Lambda}}^0$ spectrum and production rate. . . 124

7.5 $\mathit{\Lambda}^0+\overline{\mathit{\Lambda}}^0$ CM momentum spectrum (presented as fractions of hadronic σ_{hadron} per unit x_p.) . . . 125
7.6 Fitted numbers of $\mathit{\Xi}^-$ and $\overline{\mathit{\Xi}}^+$ and associated statistical uncertainties in p_{CM} bins . . . 130
7.7 $\mathit{\Xi}$ systematic uncertainties in p_{CM} bins due to tracking, MC efficiency matrices' bin statsitics, and SVT/beampipe $\mathit{\Lambda}$ fractions . . . 133
7.8 Fitted numbers from separate $\mathit{\Xi}^-$ and $\overline{\mathit{\Xi}}^+$ samples and associated statistical uncertainties in p_{CM} bins. . . . 137
7.9 Summary of the sources of error for $\mathit{\Xi}^-/\overline{\mathit{\Xi}}^+$ spectrum and production rate. . 140
7.10 $\mathit{\Xi}^-+\overline{\mathit{\Xi}}^+$ CM momentum spectrum (presented as fractions of hadronic σ_{hadron} per unit x_p.) . . . 140
7.11 Fitted numbers of $\mathit{\Omega}^-$ and $\overline{\mathit{\Omega}}^+$ and associated statistical uncertainties in p_{CM} bins . . . 145
7.12 $\mathit{\Omega}$ systematic uncertainties in p_{CM} bins due to tracking, MC efficiency matrices' bin statsitics, and SVT/beampipe $\mathit{\Lambda}$ fractions . . . 146
7.13 Summary of the sources of error for $\mathit{\Omega}^-/\overline{\mathit{\Omega}}^+$ spectrum and production rate . 153
7.14 $\mathit{\Omega}^-+\overline{\mathit{\Omega}}^+$ CM momentum spectrum (presented as fractions of hadronic σ_{hadron} per unit x_p.) . . . 154

Chapter 1

Introduction

It can be argued that experimental investigators are guided in their pursuit of timely and topical results by theoretical fashion, by the necessity to avoid impossibly difficult goals, and by the ocherous prospect of revelatory discoveries. So it is not difficult for us to acknowledge that the wider scientific community considers the most pressing issues in the field of experimental high energy physics to consist of the expected discovery of the Higgs boson at the LHC (the Large Hadron Collider at CERN) and of the hoped-for detection of the tell-tale signs of supersymmetry, a symmetry that connects fermions and bosons. But while it may appear to an observer that high energy physicsts are poised to complete the verification of the standard model of particle physics by placing the Higgs mass in its proper taxonomic location, and then to press beyond that model, it is nevertheless far from the case that the standard model offers no further challenges for experimentalists and theorists alike.

One of the major areas within the standard model that remain beyond a so-called "complete theoretical understanding" is that of hadronization: the process by which mesons and baryons condense from a mixture of elementary quarks, gluons, and virtual particles during strongly interacting processes.

As a topic of theoretical investigation, hadronization saw much interest forty years ago; one strand of such theoretical inquiry (namely, the "dual" models,) when combined with ideas from supersymmetry, has even been offered as a possible theory for all forces: the superstring theory. Nevertheless, ever since the physics community has accepted Quantum Chromodynamics ("QCD") as the correct (effective) theory of strong interaction during the nineteen seventies [1], the attention of theorists has largely shifted away from hardonization despite some contiued work in areas such as lattice field theory and on the topic of quark confinement. Still, the continued absence of analytical tools that would allow for a complete description of hadronization, in concert with the pressing need to coax interesting results from the hadronic end products produced at collider facilities, had prompted the development of phenomenological hadronization models by various workers.

The development and application of such models by high energy experimentalists and theorists (in the form of iterative Monte Carlo models that simulate the formation of hadronic final states during collisions) remains a continuing task. Some of the critical parameters that must be simulated by the existing Monte Carlo models include the total production rate of individual hadronic states and their momentum spectrums. Different phenomenological models may be tuned to model such parameters in a more or less natural way; more importantly, if techniques which would allow for a complete description of hadronization is ever to emerge, then such experimental parameters must be calculable from the first instance using such hoped-for techniques.

The present work details the measurements of the inclusive production rates and spectra of three strangely-flavored baryon states during the hadronization process (as part of the $\mathbf{e^+e^-} \rightarrow \gamma \rightarrow$ **anything** reaction.) The three particles include the spin 1/2 Λ particle (the lightest of the strange baryons and part of the spin one-half light baryon octet,) the spin 1/2 doubly-strange Ξ^- baryon, and the spin 3/2 triply-strange Ω^- baryon. These measurements shall exceed the accuracy which has been achieved by past collaborations at comparable energies [2], [3].

This dissertation is divided into six parts. In the first part (chapter 2,) we look at certain aspects associated with the historical development of hadronization models. In the second part (chapter 3,) we briefly describe the SLAC (Stanford Linear Accelerator Center) PEP-II electron-positron storage ring, the BaBar detector, and the attendant software triggers which produced the experimental data employed herein. In the third part (chapter 4,) we describe and examine the criteria and cuts employed to select an appropriate sample of collision events for our current studies. In the fourth part (chapter 5,) we examine the necessary correction factors that must be applied to the Monte Carlo simulation of proton and pion detector efficiencies at the BaBar detector. In the fifth part (chapter 6,) we examine the cuts and corrections which were employed to extract and fit the baryonic signals. And finally in chapter 7, we survey our results and their attendant uncertainties, as well as compare the results with those obtained by past collaborations and by some of the hadronization models.

Chapter 2

Some historical and theoretical issues relating to hadronization

Hadronic states interact more readily amongst themselves compared to particles that interact via the electromagnetic force and to hadronic decay processes; so it is not surprising that the underlying particle exchange which mediates hadronic collisional processes has become known among physicists as the strong force. As an illustration of the differing strengths of the strong and electromagnetic ("EM") interactions, Perkins [4] for one contrasts the magnitude of the hadronic pion-nucleon cross section ($\propto 10^4$ mb,) with that of the EM-mediated $\gamma \ + \ p \ \rightarrow \ \pi^0 \ + \ p$ process (on the order of 10^1 mb.) Another characteristic discriminant between the strong and EM forces is the effective range of these interactions: strongly interacting processes are typified by their relatively short range (this scale is oft identified with the "size" of a hadron,) whereas the electromagnetic interactions, proceeding via either on-shell or virtual photons, have an infinite range for the former and a range defined by the Heisenberg uncertainty principle for the latter. Experimental evidence for the limited range of the strong interaction (on the order of one fermi or 10^{-15} m,) include the observation of wide frequency widths for hadronic resonances and the apparent short life-times of these resonant states from which one can infer the effective range of the force using either the uncertainty principle or the hadronic compton wavelength.

2.1 Classification of hadronic states and analyses via perturbative field theories

Early attempts at describing the hadrons, and by implication the strong force, included Gell-Mann and Zweig's 1964 effort using an $\mathbf{SU(3)}$ (this is a special case of $\mathbf{SU(N)}$, the unitary group of transformations with unit determinant that can be realized upon the space of all traceless N × N matrices by $(N^2 - 1)$ traceless generating matrices) quark-flavor symmetry to organize the hadronic states into various multiplets. We now briefly look at how these early efforts led to the development of QCD.

2.1.1 Group theoretic classification schemes

Some examples of the early Gell-Mann/Zweig classification scheme include: the pseudoscalar light meson nonet (which contains the pions, kaons, η, and $\eta^{'}$) and the spin 1/2 light baryon octet (containing the nucleons, the $\mathit{\Lambda}$, and the $\mathit{\Xi}$. The mesons are hypothesized to be composites of a quark and an antiquark; so in the flavor $\mathbf{SU(3)}$ symmetry classification scheme, meson states can be grouped in terms of the following fundamental group representation: $3 \bigotimes \overline{3}$. These nine states can be further broken down in terms of their properties under the $\mathbf{SU(3)}$ transformation amongst the three light quark states (u, d, and s). This can be most easily seen using the method of Young Tableaux: we represent the conjugate $\overline{3}$ group with the set of three antisymmetric states formed from the $3 \bigotimes 3 = 6 \bigoplus \overline{3}$ representation, then using Young Tableaux, we see that $3 \bigotimes \overline{3} = 1 \bigoplus 8$. That is, the nonet (any nine physical states that transform amongst themselves under some group transformation and which constitutes an irreducible representation space characterized by some quantum number; singlet, doublet, octet, decimet are all examples of 'multiplets') breaks into an octet and a singlet (note that the physical states are mixtures of states in the fundamental representation.) Similarly, the light quark baryon octet results from combinations of three quarks ($3 \bigotimes 3 \bigotimes 3 = 1 \bigoplus 8 \bigoplus 8 \bigoplus 10$.) Note that the $\mathit{\Omega}^-$ baryon is part of the spin 3/2 decimet that results from the three quark combination.

Despite the apparent success and utility of the above classification scheme, quark flavors did not seem a likely underpinning with which to construct a quantum gauge theory of the strong interaction since it is not a unique characteristic of the strong interaction; different flavors of quarks also can be organized into $\mathbf{SU(2)}$ isopin doublets for purposes of describing weak interactions. Further, quark flavor symmetry amongst hadrons is only approximate as can be seen from the differing masses of the states in the same flavor multiplet. But there is a quality which is apparently unique to the strong force: during the 1960's, workers in the field began to suspect the existence of a "color" charge that operates in strongly interacting processes. One of the motivating observations for the introduction of the concept of color is the existence of hadronic states with spin and flavor combinations that otherwise would violate the spin-statistics theorem in the absence of some hitherto unknown quantum number (for example, the Δ^{++} resonance with spin 3/2 and isospin 3/2 composed of three s-wave u quarks is a completely symmetrical state in the absence of the color charge.) The introduction of the color quantum number in concert with the additional postulate that all hadronic states are color singlets ensured the continued validity of the spin-statistics theorem and the preservation of the Pauli exclusion principle. [1]

2.1.2 Description of strong interactions via gauge field theories

With the success of the electroweak unification (e.g. the Weinberg-Salam-Glashow-Iliopoulous-Maiani model) which is based on the postulate that the fundamental electroweak interaction can be described via a Lagrangian density that remains locally invariant under an $\mathbf{SU(2)_L} \bigotimes \mathbf{U(1)_Y}$ gauge transformation, some theorists focused their efforts on the construction of a fundamental theory of strong interaction based on an $\mathbf{SU(3)}$ color gauge symmetry of the Lagrangian density. The continuous transformation group employed in this effort was $\mathbf{SU(3)}$ because baryons were postulated to consist of color triplets and also because the alternative $\mathbf{SO(3)}$ group would imply the existence of unobserved diquark states; incidentally, the $\mathbf{U(3)}$ group also was rejected as an alternative due to the absence of long range strong interactions. [1]

The fruit of these efforts ripened in the early 1970's with the introduction of the renormalizable quantum theory of the strong force (QCD,) now regarded as a part of the standard model of particle physics. The Lagrangian density, the form of the covariant derivative, and that of the field strength tensor are well known (see e.g. §8.1 in [1].) Nevertheless, and in spite of the many successes of QCD which include predictions of three jet events in $e^+e^- \rightarrow$ hadron processes and the QCD corrections to the "Bjorken scaling" behavior of lepton-hadron scattering structure functions, there remain many unresolved issues which result in part from the inadequacies of QCD perturbation analysis.

2.1.3 Asymptotic freedom in QCD and its implications for the study of hadronization

One of the most important predictions of QCD is that the strong coupling constant grows in magnitude with decreasing energy scale, in contrast to the $\mathbf{U(1)} \times \mathbf{SU(2)}$ sector of the standard model where the coupling decreases in mangitude with decreasing energy scale; this presents a clear problem for perturbation methods which depend in large part on the decreasing importance of higher order expansion terms (enforced by higher orders of the relatively small coupling in these terms.)

We can invoke Wilson's Renormalization Group (RG) approach to examine briefly the behavior of QCD's coupling constant (theories with coupling constants that behave in this manner are generally referred to as "asymptotically free".) At its core, the RG method employs the functional integral representation of a renormalizable quantum field theory's Green function and explicitly integrates out the high momenta components' contribution to the amplitude of the process under consideration, thus effectively cutting off high energy divergences of the theory; the fundamental fields and the parameters of the theory are then rescaled and the resulting Lagrangian density is compared to the original Lagrangian. In effect, the Renormalization Group construes the parameters of renormalizable field theories as scale dependent quantities and the Lagrangian density's variations with energy scale then describes a flow within the space of all possible Lagrangians. [5] The RG transformation can be written as a differential equation of the Green function: the "Callan-Symanzik" equation.

As an example, we can apply the Callan-Symanzik equation to the total annihilation cross section of the $e^+e^- \rightarrow \gamma \rightarrow$ *anything* process. Due to the fact that the total cross section of a physical process is a directly measurable quantity at a defnite energy scale Q, there is no field-strength renormaliztion[1] component and so the differential equation obeyed by this cross section is simply (see §§12.2, 12.3, 17.2 in [5]):

$$\left[M\frac{\partial}{\partial M} + \beta(g)\frac{\partial}{\partial g}\right]\sigma(s, M, \alpha_s) = 0 \tag{2.1}$$

where $s \equiv Q^2$, M is the shifted mass, g is the shifted coupling, α_s is proportional to the shifted strong coupling constant g^2, and $\beta(g)$ is the coupling dependent beta function defined by

$$\beta(g) \ \equiv \ \frac{M}{\delta M}\,\delta g \tag{2.2}$$

The beta function can be interpreted as the rate of change in the renormalized coupling at some mass scale M (with respect to energy scale variations) that corresponds to the theory's

---

[1] Field-strength renormalization is the probability for some elementary field operator to create or annhilate a zero-momenta one-particle eigenstate of the theory's Hamiltonian. It is a quantity that can be calcuated after taking into account unobservable one-particle irreducible diagrams that modify an external field line. See §7.1 in [5].

bare coupling. The solution to the Callan-Symanzik equation can be easily obtained using the hydrodynamic-bacteriology analogy due to Sidney Coleman (see §12.3 in [5].) According to this analogy, the solution of the beta function consists of the following:

$$\beta(\overline{g}) = \frac{\partial \overline{g}}{\partial[\log(Q/M)]} \tag{2.3}$$

The coupling dependent beta function is a calculable quantity on an order by order basis using perturbation theory by employing renormalization counterterms. For $\mathbf{SU(N)}$ non-abelian theories to lowest order in the perturbation expansion, the solution of the beta function has the following form:

$$\beta(g) = -\frac{g^3}{(4\pi)^2}\left[\frac{11}{3}C_2(G) - \frac{4}{3}n_f C(r)\right] \tag{2.4}$$

In the above expression, $C_2(G)$ is the constant associated with the quadratic Casimir operator [2] in the adjoint representation of the relevant symmetry group, n_f is the number of fermion species in the given representation, and $C(r)$ is a constant that depends on the irreducible representation of the generator matrices; note that $C(r)$ is actually calculated from the trace of any two generator matrices of the irreducible representation. (See §§16.5-16.7 and eq. 15.78 in [5].) For our SU(3) color gauge theory in the fundamental representation, $C_2(G) = 3$ and $C(r) = 1/2$. Now we can use the beta function to solve for the coupling (defined as $\overline{g}$ at the renormalization scale M) as a function of the bare coupling of the theory: g. The result, written in terms of $\alpha_s \equiv g^2/4\pi$, is

$$\overline{\alpha}_s(Q) = \frac{\alpha_s}{1 + (\alpha_s/2\pi)(11 - [2/3]n_f)log(Q/M)} \tag{2.5}$$

The above expression tells us that in the standard model where we have six flavors of approximately massless quarks, the coupling decreases with increasing energy scale. Now it has been determined experimentally that QCD perturbation theory is applicable when the scale of the energy exchange is about 1 GeV and greater (see p.552 in [5].) This state of affairs presents a problem for the study of hadron formation from interactions of quark and gluons since such processes occur below 1 GeV. Note that 1 GeV corresponds to a characteristic distance on the order of 10^{-15} m, the effective range of the strong interaction; it is also the characteristic scale of hadronic mass.

2.2 Some issues related to phenomenological hadronization models

In the absence of techniques that can precisely address the process of hadron formation (i.e. "hadronization",) and in view of the need to model such processes in order to better extract various quantities from collisional products detected by high luminosity beam facilities, somewhat ad-hoc descriptions of hadronization which incorporate kinematics and more or less reasonable assumptions about the dynamics have been devised through the years. Some of the more important efforts in this phenomenological description of hardronization includes the Field-Feynman model and the Lund model (this is clearly not an exhaustive list.)

---

[2] The Casimir operator is an invariant for any simple Lie algebra and it commutes with all other members in a Lie Group. For example, in $\mathbf{SO(2)}$, the squared momentum operator P^2 is an quadratic opeator that commutes with all the momentum generators, and therefore has a unique eightvalue on the space of states. [6]

2.2.1 The Field-Feynman model

In their 1976 paper on the subject, Field and Feynman note with some frustration that observations of collisions between relativistic hadrons, collisions that theorists had hoped to involve direct interactions between the partonic constituents and thereby elucidate the short distance behavior of the strong force, revealed that these process appeared to be much more complex than first thought (for example, the $p_\perp$ dependence of the scaled cross sections was predicted to be approximately $p_\perp^{-4}$ with possible log corrections but were in fact observed to be closer to $p_\perp^{-8}$.) The possible fragmentation of the parent hadrons in the collisional process also made it difficult for workers to compare theoretical predictions regarding the behaviors of the strong force to experiments. [7] According to these authors, progress in the field appears possible only if we were to take some phenomelogical model and slowly and methodically test it against all experimental evidence. To that end, Field and Feynman constructed their own fragmentation model which is based on the picture that as a quark-antiquark pair separates, the strong color field between them builds up until additional pairs of quark-antiquark pairs are produced which breaks down this field; eventually, these additional quarks and antiquarks combine into color singlets to make hadrons, which in turn decay to give the final observed light mesons and baryons.

This basic picture apparently was first suggested by Bjorken to explain the failure by various experiments to observe free quarks in deep-inelastic scatterings, and was analyzed by Casher, Kogut, and Susskind using QED in one space and one time dimensions. [8]. In the latter paper, the authors argued that the ordinary femionic current for QED $j^\mu \equiv e\overline{\psi}\gamma^\mu\psi = \partial_\nu F^{\nu\mu}$ can be redefined in such a way that it satisfies the Klein-Gordon equation for a massive boson (this mechanism was due to work by Schwinger):

$$\left(\Box + m^2\right) j^\nu = 0 \tag{2.6}$$

The solution for the underlying fermionic field in this theory then can be written in terms of a quantity (interpreted as a dipole density) that is related to this massive bosonic current. The result of this ‘trick’ is that if there are any external ‘scalar photon’ that weakly couples to this simplified QED via the operator $s(x) = \overline{\psi}(x)\psi(x)$ (the interacting theory's ferminoic field can be written in terms of the scalar photon and the free field theory's fermionic field, thus the operator $s(x)$ is equivalent to the interaction Hamiltonian of the interacting theory) such that the matrix element of the $e^+e^- \rightarrow$ anything process in this simplified theory is given by the following time-ordered expression:

$$S(x) = i\langle 0|Ts(x)s(0)|0\rangle \tag{2.7}$$

then this matrix element can be rewritten in terms of the apparently massive bosonic propagator (an induced dipole density) rather than in terms of the underlying fermionic propagator. This induced dipole density has the following form:

$$\phi(x) = e\theta(t+z)\theta(t-z) \; - \; e\Delta_R(m^2, x^2) \tag{2.8}$$

where e is the theory's coupling, t and z are the time and space parameters respectively, and Δ_R is the retarded commutator of the dipole density. According to Casher, Kogut, and Susskind, the retarded commutator is well approximated by $\theta(t+z)\theta(t-z)$ near the light cone ($x^2 < m^2$,) and by $\exp\left(-m|x|\right)$ for large $|x|$. Physically, this implies that in a fermionic annihilation process where the initial state consists of a free fermion-antifermion pair, the initial bosonic "dipole density" is negligible; as the initial particle pair separates, the dipole density midway between them obtains the value of the coupling and at still later time, the dipole density everywhere in the line connecting the pair approaches a constant value equal to the coupling. In other words, polarized pairs of fermions are created out the vacuum in the line between the original fermions and "annihilate" them leaving only the final observed bosonic states. Furthermore, the dipole density field is described by a coherent state so that the resulting bosons are Poisson distributed on the rapidity axis, with the result that the momentum space density of the states is unity.

In addition to incorporating the above picture into the hadronization process, the Field-Feynman model further assumes that the original quarks always find themselves within hadrons at the higher end of the hadron momentum distribution while hadrons with smaller fractions of the total event momentum gradually become independent of quark flavors (as a result of the fact that the color field is flavor-independent,) and eventually become uniform in rapidity (the so-called rapidity plateau suggested by Casher, Kogut, and Susskind.) Finally, the Field-Feynman model assumes that the new quark pairs are equally likely to be $u\overline{u}$ and $d\overline{d}$ from isospin symmetry, but that strange quark production is suppressed due to its higher mass (charm quarks and additional flavors were disregarded in this model.) See p.2596, [7].

In a subsequent paper published the next year ([9],) Field and Feynman gave explicit parameters to their meson hadronization model by defining a probability function $f(\eta)$ which gives the probability that the remaining energy in a jet of mesons produced by an original fast-moving quark is a fraction η of the initial energy of the quark, after $1 - z_1$ fraction of the initial energy has been utilized to produce a single new meson, where $z_1 \equiv \zeta_1/W_0$ is the new meson's momentum fraction and W_0 is the total energy of the initial quark. The $f(\eta)$ function determines the momentum distribution of the mesons that are produced in the jet. This function satisfies the following normalization relation:

$$\int_0^1 f(\eta)\, d\eta \;=\; 1 \tag{2.9}$$

Then, if $F(z)\, dz$ represents the probability of finding a meson with fractional momentum z in dz in the quark jet, the new function $F(z)$ satisfies the following recursive integral equation:

$$F(z) = f(1-z) + \int_z^1 f(\eta)\, F\left(\frac{z}{\eta}\right)\, \frac{d\eta}{\eta} \tag{2.10}$$

The $f(1-z)$ term occurs in the above expression because the new meson may be the first one produced in the jet with probability $f(1-z)\, dz$; the second term arises from the fact that if the meson is not the first one produced by the jet, the jet has η fraction of energy left with probability $f(\eta)\, d\eta$ and in this remaining jet, the meson's production probability is given by $F(z/\eta)\ (dz/\eta)$. See eq. 2.10. In essence, this model further assumes that once a jet begins to fragment into parts, the parts continue to fragment independently of any other jet of hadrons.

The form of equation 2.10 involves only differences in the z rapidity ($-\log z$) so a formal solution can be obtained using Fourier transform or moments analysis as detailed on pp.6-7 of [9]. At this point, Field and Feynman chose an explicit expression for $f(\eta)$ in the form of powers of η plus a small constant due to its simplicity and its ability to generally reproduce some of the behaviors of meson momentum spectrums; this is the so-called Field-Feynman fragmentation function (with parameter a):

$$f(\eta) = 1 - a + 3a\eta^2 \tag{2.11}$$

Solving equation 2.10 using moments analysis with the chosen form for $f(\eta)$, they obtained the following expression for $F(z)$:

$$zF(z) = \frac{3}{3-2a} + \frac{3az^2}{(2a-1)} + \frac{2a(2a^2-3a-2)z^{3-2a}}{(3-2a)(2a-1)} \tag{2.12}$$

The above method, with its assumption that there's such a thing as a "primary" meson which carries most of the energy and momentum of the original quark, the recursive integral probability equation, and an assumed fragmentation function, is tailor-made for Monte Carlo simulations; indeed, it may have been created with Monte Carlo simulation in mind. Thus the Field-Feynman fragmentation function (and the attendant fragmentation model) may be thought of as one of the progenetors of a series of subsequent Monte Carlo fragmentation functions designed to simulate various aspects of hadronization.

2.2.2 The relativistic string fragmentation model

An alternative to the Field-Feynman fragmentation function is the Lund symmetric fragmentation function, (hereafter "LSFF") developed by a group from the University of Lund in the early 1980's. The practical importance of the LSFF lies in the fact that it is central to the Pythia e^+e^- Monte Carlo hadronization generator, a generator that is widely used by international high energy collaborations at present; the LSFF is also important in that is a fully relativistic and it logically connects the fragmentation of the original quark with the original antiquark. Just as in the Field-Feynman model, the Lund model assumes that the basic dynamics of hadronization consists of the fragmentation of a narrow tube of color field due to the production of "dipoles" of $q\overline{q}$ pairs along the the tube as it stretches between the initial pair of quark-antiquark; the fragmented jets then further fragment independently of other jets and hadrons in accordance with the recursive integral integration 2.10. [10] Three jets events with the radiation of an additional hard gluon is also treated as a 'kink' in the string.

The Lund group then derived the LSFF (employed in place of the Field-Feynman fragmentation function) using the massless relativistic string as a model for the color field between the initial quark pair. [11]. For this reason, the Lund model is often referred to as a "string fragmentation model".

2.2.3 The origin of the idea to use strings to model strong interactions

It is of some historical interest that the idea of using a relativistic string to model hadron formation is related to developments in the so-called "dual models" of strong interaction, developed during the 1960's and early 1970's, prior to the wide acceptance of QCD.

According to Veneziano, the essence of dual theories consists of an S-matrix[3] based description of strong interactions wherein $\langle \mathrm{f}|S|\mathrm{i}\rangle \equiv S_{fi}$ for a process that proceeds from an initial state 'i' to a final state 'f' with probability $|S_{fi}|^2$. In this scheme, the S-matrix must satisfy a set of general postulates which is sufficiently constraining so as to allow for an

---

[3] The S-matrix is the amplitude for some given scattering process with defined initial states and final states; the probability calculated from the S-matrix, after multiplying over phase space factors, gives the differential cross section for that process. See §4.5 in [5].

approximate though unique description of particle exchange between strongly interacting external states but without employing perturbation methods. This is in contrast to the case in field theory where one starts from a Lagrangian density written in terms of some fundamental fields and then proceed to solve the dynamical problem in an approximate manner by adding together the first order gauge field exchange term with additional smaller (due to powers of the coupling) perturbation terms. [12]

The postulates necessary to describe the strong interaction S-matrix includes the following: the S-matrix obeys Poincare invariance, the S-matrix results in the most likely interactions among particles with definite momentum being either no interaction or interaction in clusters ("cluster decomposition",) the S-matrix is unitary and thus conserves probability, the S-matrix is analytic in particle momenta with singularities only on the real axis and not on the "physical" (complex) sheet, a physical process' S-matrix can be analytically continued into another "connected" process' S-matrix using the same analytic function (e.g. a t-channel scattering process in Mandelstram space can be analytically continued into a s-channel scattering process; this is called the "crossing symmetry" requirement,) the S-matrix must have certain specified analytic structure in the high energy limit.

However, the above postulates by themselves are inadequate to specify a unique solvable model; what is needed is yet another postulate that would further constrain the possible theories. It turns out that, for dual models, the additional necessary postulate is the following: that the S-matrix continues to satisfy the previously identified postulates in the one-particle exchange (one particle exchanged at a time, but many different possible exhange particles are allowed) approximation.

The one-particle exchange approximation was needed by theorists to make headway in the construction of dual models due in part to the fact that the unitarity condition couples scattering amplitudes with an ever increasing number of external states such that if intermediate states with multiple particles make important contributions to the S-matrix, then the resulting set of equations become overly complicated for analysis. By assuming that one-particle exhannge is a good first order approximation to the full S-matrix (equivalent to the assumption of narrow resonance widths,) the equations then become greatly simplified and allows for solutions. In view of this, we see that dual resonance models also involve an inevitable approximation although in this case the approximation is made immediately at the S-matrix level unlike in the case for Lagrangian field theories.

In the one particle exchange scenario, the scattering amplitudes for an n-point function can be written as a sum of terms involving isolated poles[4] and their residues. We also need to impose the requirement that the scattering amplitude must satisfy crossing symmetry; this means that if an amplitude has resonances in the s and t channels, then the amplitude written in the s-channel with s-channel poles must be equal to the amplitude written in the t-channel with t-channel poles.

If one restricts oneself further to "planar" models (i.e. models which has no resonances in some channels,) then, for an n-point function, it is possible to start from the so-called Mittag-Leffer expansion of the amplitude in the s and t channel poles, impose crossing symmetry on the amplitude, and finally conclude that the set of dual model generators must act as generators of the conformal transformation and satisfy the Virasoro algebra (see pp.217-225 in [12]):

---

[4] The isolated poles correspond to the one-particle states viewed as analytic functions of their external momenta (and to correlation functions via the LSZ-reduction formula.) They also correspond to the Breit-Wiger resonance levels in the spectrum of the interaction under consideration See §§4.5,7.1,7.2 in [5].

$$[L_n, L_m] = (m-n)L_{n+m} + \frac{D}{12}(m^3 - m)\delta_{n+m} \tag{2.13}$$

The Virasoro algebra, along with certain other conditions, ensure that there are no ghosts (states with negative norms) in the model. This process involves defining vertex and mass operators in the n-poiint function that correspond to the $n-3$ Mandelstam variables which simultaneously experience exchanges at the same resonance pole. After imposing the crossing symmetry requirement, the equation involving equivalence of the amplitudes under crossing then can be expressed explicitly in terms of an $\mathbf{O(2,1)}$ conformal symmetry by rewriting the Mandelstam poles as follows:

$$\frac{1}{\alpha}\frac{1}{M - s_i} = \int_0^1 x^{\alpha(-s_i+M)-1}\, dx \ \equiv -i\int_0^\infty \exp\left[-i\alpha\tau(s_i - M)\right] d\tau \tag{2.14}$$

where s_i is the i'th Mandelstam variable, α is a parameter with the dimension of inverse squared mass, and τ is the parameter for the conformal transformation. Finally, generators of the conformal transformation can be defined and then written in terms of the raising and lowering operators of the harmonic oscillator. The excitation spectrum of the operators later was found by various workers to involve no ghosts only if $D \leq 26$ in the generalized beta-function model and $D \leq 10$ in the Neveu-Schwarz model where D is the constant in eq. 2.13.

An important development with respect to the various extant dual resonance models was that the excitation spectrums of these models subsequently were recognized to be similar to that of an excited relativistic string in space-time. See e.g. [13], [14]. Some theorists then attempted to further elucidate the string-dual resonance connection by examining the behavior of the relativistic string in detail.

Let's take a moment to examine the relativistic string following the 1973 Goddard paper [14]. First we shall pick orthonormal coordinates τ and σ to prameterize a point on the string which must have its end points moving at the speed of light so that we can write any point on the string as $x^\mu = x^\mu(\sigma, \tau)$. If the surface swept out by the string in space time must move at or below the speed of light, then this surface must satisfy the following condition written in terms of differential form $dF^{\mu\nu}$ (see [14]:

$$-dF^{\mu\nu}dF_{\mu\nu} = (d\sigma\ d\tau)^2\left[\left(\frac{\partial x}{\partial\sigma}\frac{\partial x}{\partial\tau}\right)^2 - \left(\frac{\partial x}{\partial\sigma}\right)^2\left(\frac{\partial x}{\partial\tau}\right)^2\right] \geq 0 \tag{2.15}$$

The differential surface defined by the string in space time then is given by

$$dA = \left[-dF^{\mu\nu}dF_{\mu\nu}\right]^{1/2} \tag{2.16}$$

We can define the action in terms of this surface:

$$S \ = \ -\frac{1}{2\pi\alpha\hbar c^2}\int_{\tau_i}^{\tau_f}\int_0^\pi dA \tag{2.17}$$

In equation 2.17, α is a constant with dimension of $\mathbf{energy^{-2}}$ (it is in fact the same constant used in equation 2.14.) Using the principle of least action, and if we take the τ parameter to be the time t in some reference frame such that the $x = x(t, \sigma)$ is now a three vector, then we can write the equation of motion and the boundary conditions of the string as follows:

$$\frac{\partial}{\partial t}\left[\frac{T_0}{c^2}\frac{\vec{v}_\perp}{\sqrt{1-[v_\perp^2/c^2]}}\frac{ds}{d\sigma}\right] = \frac{\partial}{\partial\sigma}\left[T_0\sqrt{1-\frac{v_\perp^2}{c^2}}\frac{\partial x}{\partial s}\right] \tag{2.18}$$

$$(v_\perp)_1 = (v_\perp)_2 = c \qquad \frac{d\sigma_1}{dt} = \frac{d\sigma_2}{dt} = 0 \tag{2.19}$$

In 2.18, $T_0 \equiv (1/2\pi\alpha\hbar c)$ is the rest tension of the string, ds is a length element along the string, and $v_\perp^2$ is the squared speed of the string element in the direction perpendicular to that element. Equation 2.18 is a wave equation with an effective tension $T = T_0\sqrt{1-(v_\perp^2/c^2)}$ and a mass per unit length given by $T_0/[c^2\sqrt{1-(v_\perp^2/c^2)}]$. It is interesting to note that a straight string with length $2a$ and undergoing rigid rotation with angular velocity ω has the following ratio between total angular momentum and the square of its total energy:

$$\frac{J}{E^2} = \alpha\hbar \tag{2.20}$$

The above expression is just the leading linear Regge trajectory. From this point on, one can define momentum and position generators and then proceed to obtain the spectrum of the relativistic string; but we shall not pursue this exposition any further.

The relativistic string, as formulated by Goddard et. al. was subsequently treated in the classicial limit by Artru and Mennessier [15] for application to modeling hadronic interactions (when applied as a theory of fundamental interactions after incorporating ideas from supersymmetry, the string model is now known as superstring theory.) At this point, we also note that the Lund group itself has referenced its use of relativistic string dynamics in fragmentation studies as a sort of application of the dual resonance model in its classical limit (e.g. in the abstract of [10].)

2.2.4 Derivation of the string-based Lund Symmetric Fragmentation Function

We now return to a discussion of the Lund fragmentation fuction. At the heart of the Lund string model lies the assumption that as the relativistic string is pulled apart by the receding pair of quark and antiquark, new $q\overline{q}$ pairs are produced at certain space-time vertices which are causally disconnected both from each other and from the original $q\overline{q}$ pair; therefore all pair production vertices in the string must be treated in an equivalent manner with the result that there is no such thing as a primary vertex or a primary hadron. The quarks and antiquarks from the newly created-pairs are then combined with other quarks and antiquarks to form color singlets that decay to final state observed hadrons. In terms of the iterative process employed in Monte Carlo simulations, the probability of finding a certain final state hadron should be the same whether the iteration procedure commences from the original quark end of the string or from the antiquark end as a result of the causal independence of the pair production vertices.

Specifically, the Lund group used the so-called "yo-yo" relativistic sting model (due to Artru et al.) with a string tension κ in one space and one time dimension. The differential probability that a new $q\overline{q}$ production vertex with momentum fraction z_1 is produced by the string is given by

$$dP \;=\; f(z_1, s_0)\, dz_1 \tag{2.21}$$

where $f(z_1, s_0)$ is the distribution of the first iterative meson produced by the combination of, say, an antiquark coming from the new production vertex with the original quark, and where s_0 is the squared mass of the original $q\overline{q}$ pair. From the assumption of independent fragmentation, the next vertex (with momentum fraction z_2) then must have a differential probability $dP \;=\; f(z_2, s_1)\; dz_2$, where s_1 is the squared mass in the remnant system. At this point, the Lund group further assumes the presence of a central rapidity plateau (z-rapidity: $y \equiv -\log z$ in the distribution of the final state mesons, as did Field and Feynman, such that a new vertex in this rapidity region has a differential probability that is dependent only on its proper time ($\kappa\tau^2 \equiv \Gamma$):

$$dP \;=\; C_0\; H(\Gamma)\; d\Gamma\; dy \qquad \text{with} \qquad \int_0^\infty d\Gamma\; H(\Gamma) \;=\; 1 \tag{2.22}$$

Here, C_0 is a constant equal to the number of vertices per unit rapidity. The production of a new meson in this central region depends on the presence of an existing production vertex with the above differential probability, as well as the presence of a new production vertex (let's say the $i+1$'th vertex counting from the "positive" quark end of the string) with differential probability $f(z_+)\; dz_+$, where $z_+ \equiv \kappa x_+/(E+P)$ is the momentum fraction of the new vertex in "forward" light cone coordinates, and where the function $f(z_+)$ is no longer dependent on s_0 due to the assumption that in the central rapidity region this function is independent of the mass. Then the total differential probability for the new meson is given by the following expression:

$$dP \;=\; H(\Gamma_i)\; f(z_+)\; d\Gamma_i\; dz_+ \tag{2.23}$$

But under the Lund group's assumptions, the meson should have the same differential probability if we consider the new vertex in an iterative manner from the original antiquark end (given in terms of the "backward" light cone coordinates.) Then rewriting $\Gamma_\pm$ in terms of the momentum fractions, and taking care to write the differential lengths in a Lorentz invariant manner, they arrived as the following expression:

$$z_+\; f(z_+)\; H\left(m^2 \frac{1-z_-}{z_+ z_-}\right) \;=\; z_-\; f(z_-)\; H\left(m^2 \frac{1-z_+}{z_+ z_-}\right) \tag{2.24}$$

This expression can be solved by turning the multiplicative factors into additive ones through the definitions $z\; f(z) \equiv \exp\left[g(z)\right]$ and $H(\Gamma) \equiv \exp\left[h(\Gamma)\right]$, and then by differentiating it with respect to $z_\pm$; the result is that the $g(z_\pm)$ terms drop out and we're left with

$$\frac{d}{d\Gamma_i} h(\Gamma_i) + \Gamma_i \frac{d^2}{d\Gamma_i^2} h(\Gamma_i) = \frac{d}{d\Gamma_{i+1}} h(\Gamma_{i+1}) + \Gamma_{i+1} \frac{d^2}{d\Gamma_{i+1}^2} h(\Gamma_{i+1}) \tag{2.25}$$

The above expression can be generally true only if each side is independently equal to some constant. Since the general solution for $h(\Gamma)$ from separately examining each side of this expresion is clearly given by $h(\Gamma) = A \log \Gamma + \sum_j B_j \Gamma^j$, we find:

$$h(\Gamma) = (a) \log \Gamma + (-b)\Gamma + c \tag{2.26}$$

where a, b and c are constants. Thus $H(\Gamma) = C\Gamma^a \exp\left(-b\Gamma\right)$. Using this form for $H(\Gamma)$ then allows us to solve for $zf(z) = \exp\left[g(z)\right]$ with the final result (in terms of momentum fractions):

$$f(z) \;=\; \frac{N}{z}\,(1-z)^{a}\;\exp\left(\frac{-bm^{2}}{z}\right) \tag{2.27}$$

This is the LSFF with a flavor-dependent free parameter "a", a flavor-independent "b", and a multiplicative normalization constant N. Thus the momentum distribution of a meson with mass m produced by the combination of a quark from some vertex with flavor α and an antiquark at another vertex with flavor β is given by

$$f^{m}_{\alpha\beta}(z) \;=\; \frac{N^{\beta}_{\alpha}}{z}\;z^{a_{\alpha}}\;\left(\frac{1-z}{z}\right)^{a_{\beta}}\;\exp\left(\frac{-bm^{2}}{z}\right) \tag{2.28}$$

The Lund group further references the apparent similarity between the exponential term (with the squared mass factor being proportional to an area in 1+1 space time) in the LSFF and the gauge invariant Wilson Loop. However, since the LSFF is an actual probability density, while the Wilson Loop generally is the trace of a comparator between fields at different points in space time under a global *phase* rotation (see §§15.1-15.4 in [5],) the validity of such a comparison between the two is far from clear.

While the LSFF is central to the Lund fragmentation model in the sense that, first and foremost, a hadronization model must model the momenta of final state hadrons more or less correctly, many additional pieces also are needed in order to construct a practical iterative model that can be implemented in the form of a Monte Carlo algorithm. Some of these pieces include the mechanism for creating new $q\overline{q}$ pairs in the string (in the Lund model, the transition from the vacuum to a state with a new $q\overline{q}$ pair is pictured as the pair's wavefunctions tunneling through a linear potential and analyzed using WKB methods with an adjustable parameter, see pp. 69-70 in [10],) gluon fragmentation (modeled as a kink or transverse excitation on the linear string which arises by using methods from perturbative QCD with specialized treatments for infrared divergences, see §4.4 in [10],) the suppression of the rate of vector mesons vis-a-vis pseudoscalars (usually achieved by setting a ratio parameter to 1 : 1; the Lund group analyzed the effect of spin-spin interaction on the tunneling of mesons from inside the vacuum and concluded that spin has relatively little effect on the ratio of vector meson to pseudoscalar meson so that this ratio should generally be smaller than what one would expect from simple spin counting. See §3.3 in [10]) and the production of baryon-antibaryon pairs (using either the production of diquark-antiquark color-singlet pairs in the string via tunneling followed by quark rearrangement with color flow to produce the baryons, or a mechanism called "popcorn" mesons wherein additional $q\overline{q}$ pairs are produced from the string and combines with non-singlet quarks or antiquarks to form both baryons and mesons.) These topics are covered in detail in [10] and in [16].

2.3 Alternative fragmentation functions and models

There are of course other implementations of iterative Monte Carlo hadronization models. For example, there is the UCLA model which substitutes the LSFF with a new fragmentation function that simultaneously determines the particles' momenta and suppress the rate of high mass hadrons (this fragmentation function also incorporates Clebsch-Gordon coefficients and constant terms called "knitting factors" that gives the probability of the hadron's spatial wavefunction.) [17] Another way in which the UCLA model differs from the Lund model is in the way that it derives the exponential term in the LSFF; whereas in the LSFF the exponential term is derived by using the principle of "left-right" string symmetry (in one space dimension,) the presence of a similar exponential in the UCLA fragmentation function results from the assumption that the fragmentation process should have an exponential space-time dependence (ostensibly suggested by lattice QCD theories and a Minkowski area law.) The UCLA fragmentation function has the following form:

$$f(z, m_h) = \frac{N_h C_h^2}{(4\pi)^2} \frac{(1-z)^a}{z} \left(1 - \frac{m_h^2}{Sz}\right)^a e^{-bm_h^2/z} \tag{2.29}$$

In the above expression, N_h is a 'knitting factor' that describes the combination of constituent quarks to form the hadron's spatial wave function, C_h represents the Clebsch-Gordon coefficients (describes the spin and flavor combinations of the quarks), z is the energy-momentum fraction of the hadron in the "positive" light-cone variable, m_h is the hadron's mass, a, b are two arbitrary parameters, and S is the squared center of mass energy of the system. We also note that the UCLA fragmentation model uses 6 free parameters in contrast to the Lund model which employes at least 17 free parameters.

Alternative fragmentation functions obtained by using mechanisms other than string dynamics also exist. These include the Peterson function, the Kartvelishvili function, the Collins-Spiller function, and the BCFY function. For a survey of the aforementioned fragmentation functions, see the relevant discussion in [18]. Note that there is also the Weber cluster fragmentation model [19] implemented in the HERWIG generator; this model calculates the parton shower using perturbative QCD and then simply groups the resulting partons into color singlets.

The present work is essentially an experimental determination of three baryonic data points during the hadronization process with greater accuracy than has been achieved to date. It is our hope that the results herein will allow for more accurate comparisons between data and Monte Carlo fragmentation models, thereby furthering the understanding of the hadronization process. In addition, if future progress is made in non-perturbative hadronization theories (either string-based or otherwise,) our results also will serve as an experimental benchmark. But since we will not examine the Monte Carlo and theoretical models ourselves, we shall not consider these matters further in this dissertation.

Chapter 3

The BaBar detector and related hardware and software triggers

In order to perform the measurements detailed herein, we use the datasets obtained by the BaBar detector located at the Stanford Linear Accelerator Center ("SLAC") between the years 1999 and 2005.

The BaBar collaboration relies upon SLAC's PEP-II e^+e^- storage rings to generate electron-positron collisions at the BaBar detector's interaction point ("IP".) The neutral and charged products of those collision events that fall within the acceptance volume of each of the subdetectors generate signals in the various systems; these signals then are processed and analyzed by BaBar's Level 1 ("L1") hardware trigger system. The L1 trigger was designed to help determine whether the information associated with a given event is to be saved for further study or to be discarded; events that survive the L1 trigger are passed on the Level 3 ("L3") sofware trigger system for further characterization. Note that currently there is no Level 2 trigger layer at BaBar; this level is reserved for possible implementation in the future if the need arises for further reductions in the event acceptance frequency.

At this point, we briefly consider the hardware and software which enable the collection of the relevant dataset. A more detailed treatment of the BaBar detector hardware and software is contained in [20].

3.1 The PEP-II B-factory

The PEP-II storage ring system receives and stores electron and positron bunches from the beams generated by the Stanford Linear Collider ("SLC") via a dedicated bypass line; beam injections from the SLC occur on cycles that average 40-50 minutes. The stored electron beam travels in a clockwise manner (viewed from above the storage ring) in the high energy ring and has an average energy of 9 GeV while the positron beam resides in the independent low energy ring and travels in a counterclockwiese direction at an average energy of 3.1 GeV.

The differing energies of the electron and positron beams and the asymmetric nature of the resulting collisions were designed to better achieve the primary goal of the collaboration: the systematic study of CP-violations in the decay of B mesons to CP eigenstates at a relatively high luminosity of 3.0×10^{33} cm^{-2} s^{-1} and above [5]. The head-on collisions of the electron

---

[5] In the standard model with three quark generations and a single Higgs field, the phenomenon of CP-violation is due entirely to the presence of an irremovable complex phase in the Cabibo-Kobayashi-Masakawa

and positron beams lead to a Lorentz boost of $\beta\gamma = 0.56$ such that in the center of mass frame, the e^+e^- collision enregy is at 10.58 GeV, just above the $\Upsilon(4s)$ resonance. In addition, about 12% of the data is taken at a center of mass energy of 10.54 GeV in order to facilitate studies of the continuum background present in the B-meson decays. [20].

The section of the beamline that lies near the collision IP are flanked by a series of magnets designed to separate the electron and positron beams in the horizontal plane prior to their anticipated head-on collisions. In addition, the beam collisional axis is offset from the global z-axis of the BaBar detector by about 20 mrad in the horizontal plane in order to minimize beam perturbations due to the solenoidal magnetic field. The beamline near the IP (and including the IP) is enclosed by a water-cooled beryllium cylinder (the "beampipe") of 2.7868 cm outer radius and approximately 2.4972 cm inner radius (water channels are embedded within layers of the beampipe.) The total thickness of the central beampipe section at normal incidence corresponds to about 1.06% of a radiation length. (See e.g. p.19 in [20] and p.13 in [22].)

3.2 The BaBar detector

The BaBar detector consists of several layers of subdetectors that surround the IP. The detector's design was optimized to maximize the geometric acceptance for the boosted $\Upsilon(4s)$ decays (the primary process of interest for CP-violation studies at BaBar) by shifting the IP relative to the entire detector by 0.37 m in the direction of the higher energy electron beam. The detector consists of six parts: the Silicon Vertex Tracker ("SVT",) the Drift Chamber ("DCH",) the Detector of Internally Reflected Cherenkov light ("DIRC",) the Electromagnetic Calorimeter ("EMC",) the Instrumented Flux Return ("IFR",) and finally the superconducting solenoid (not a subdetector) which generates the 1.5 T magnetic field used to determine the charge and momentum of charged particles. See figure 3.1.

The BaBar collaboration has defined a right-handed orthogonal coordinate system for use with the BaBar detector [23]. In this system, the positive z-axis is parallel to the solenoidal magnetic field and in the direction of the higher energy electron beam, the positive y-axis points vertically upward, and therefore the positive x-axis points in the direction away from the center of the PEP-II ring. The origin of the BaBar coordinates is nominally defined to be coincident with the IP (though the actual production vertex of an event usually is offset from the origin on the order of a few millimeters in the xy plane.) Since the beam-beam collisional axis is rotated by about 20 mrad (about the y-axis,) the boost axis is not coincident with the global BaBar z-axis.

We now take a brief look at each of the five subdetectors as well as at the trigger system.

3.2.1 The SVT

The design goal of the SVT is the precise reconstruction of charged particle trajectories and decay vertices near the IP. In particular, for purposes of CP assymetry measurements, the SVT must provide better than 80 μm resolution for the reconstruction of B-decays' mean z-vertex; similarly, the required resolution in the x-y plane must be on the order of 100 μm

---

mass mixing matrix. The fact that all CP-violating processes occur as a result of this single phase under the minimal standard model thereby strongly constrains the pattern of CP-violations in B-meson decays. BaBar was conceived to study this decay pattern in order to verify and to better constrain the CKM phase value [21].

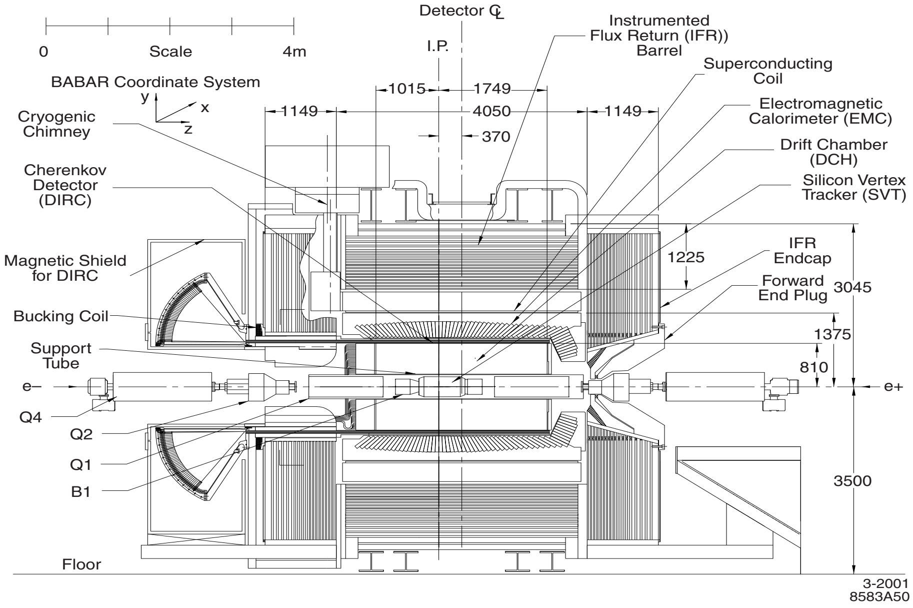


Figure 3.1: Longitudinal cross-section of the BaBar detector.

(this is determined by decay processes such as $B^0 \to D^+D^-$ in which the D vertices typically are separated in the x-y plane by $\sim$ 275 μm; see p.33 in [20].)

The physical structure of the SVT consists of five detector layers each mounted with double-sided silicon strip sensors (for a total of ten distinct layers.) The strips are organized in modules. In layers 1, 2, 3, 4, and 5, the strip sensors are organized into 6, 6, 6, 16, and 18 separate modules respectively; layer 1 is located at an average radial distance of 3.3 cm from the BaBar origin, layer 2 is at about 4.0 cm, layer 3 is at about 5.9 cm, layer 4 is located at distances from 9.1 cm to 12.7 cm (larger variations than inner layers due to this layer's "arched" architecture,) and layer 5 ranges from 11.4 cm to 14.6 cm. See figure 3.2 for a schematic cross-sectional (in the x-y plane) representation of the SVT modules.

According to the BaBar collaboration's published description, the strips on opposite side of each layer are oriented orthogonally to each other with the so-called ϕ *strips* oriented parallel to the beam and the z *strips* oriented in a transverse direction relatively to the beam axis. Further, the modules in layers 1, 2, and 3 lie flat along the beam while layers 4 and 5 have arched modules along the beam. The support tube for the SVT structure is mounted on the PEP-II ring supports so that the SVT can move relative to the rest of the BaBar detector. The SVT's readout electronics are mounted outside the active detector volume. See pp. 33-35 in [20].

The SVT sensors themselves consist of 300 μm thick silicon strip devices built on n-type substrates with p^+ and n^+ strips mounted on opposite sides; they are AC-coupled to the electronics through integrated decoupling capacitors. These sensors are essentially p-n junctions operated at reverse bias with contact regions depleted of mobile charge carriers.

When a charged particle passes through a sensor, it deposits energy in the sensor's silicon atoms via EM interaction which ionizes the atoms; the ionized charge is then swept by the electric field present in the sensor (thus producing either an electronic current or hole current) and collected at the integrated electrodes. This results in a pulse of signal which is amplified, digitized, and buffered by integrated electronic circuits (the "ATOM" chips, an acronym for

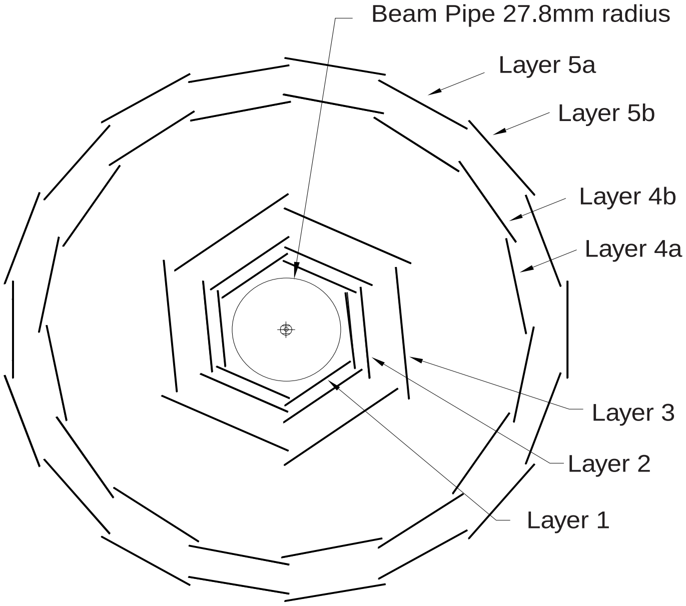


Figure 3.2: SVT layers' layout.

A Time Over threshold Machine; 128 sensor channels per ATOM chip.) If the total charge exceeds some predetermined threshold and a L1 trigger *Accept* signal is received by the ATOM chips, then the electrical signals are multiplexed and sent into fiber optic readout channels that connect to the data acquisition system. A time correction is applied to each signal pulse; cluster-finding algorithms then are run so as to group together temporally-consistent as well as spatially-consistent signals into clusters with position information. These are the so-called SVT "hits". There are a total of about 150,000 signal sensor channels in the SVT.

For additional descriptions of the sensors, associated electronics, alignment, and performance of this subdetector, please see §5 in [20]. At this point we will just note that the z-hit and ϕ-hit resolutions for all layers and all track angles are better than 50 μm based on studies conducted prior to 2002. The SVT also provides for up to 10 measurements of dE/dx for a given track.

3.2.2 The DCH

The DCH is intended to detect and measure charged particles' momenta and energy with a high degree of precision. The measurements from the DCH complement those obtained from the SVT. The performance goals of the DCH include the following: charged track spatial resolution better than 140 μm averaged over a drift cell, the acquisition of a charged track's dE/dx energy loss information to allow for particle identification ("PID") of low momentum tracks, the acquisition of up to 40 measurements of a track's spatial location (both in the transverse and longitudinal directions relative to the beam) etc. See p.43-44 in [20].

The DCH consists of two concentric cylinders each 2.80 meters in length. The inner cylinder is made of 1 mm thick beryllium (0.28% of a radiation length) with a raduis of 23.6 cm while the outer cylinder is composed of two layers of 1.66-mm-thick carbon fiber laminated onto a 6-mm-thick honeycomb core; the radius of the outer cylinder is about 80.9 cm. The forward aluminum endplate has a thickness of 12 mm beyond a raduis of 46.9 cm (thickness

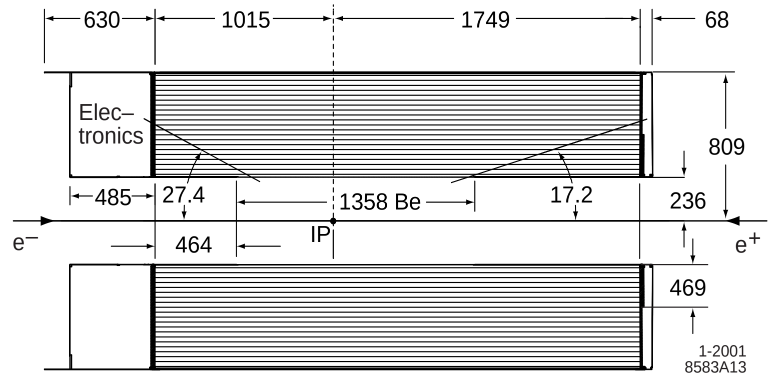


Figure 3.3: Scematic representation of the DCH's longitudinal cross-section.

of 24 mm at radial distances less than 46.9 cm) while the rear aluminum endplate has a thickness of 24 mm. The differing thickness of the foward endplate resulted from an effort to minimize material in the foward (boosted) direction. For the same reason, the readout electronics are mounted entirely on the rear endplate. See figure 3.3.

The main detection components of this subdetector are composed from a total of 28,768 wires strung between the front and rear endplates. There are mainly two types of wires, the "sense wires" (maintained at some specified high voltage, typically 1960 V) and the "field wires" (maintained at ground potential.) The wires are divided into 7,104 "drift-cells" which in turn are organized into 40 cylindrical layers (the layers are grouped by four and again divided into 10 superlayers.) Note that there are also some field-shaping "guard" wires at the boundaries of the superlayers (set at a potential of 340 V) and "clearing" wires near the edges of the DCH radial boundaries (these wires are set at 825 V.)

Each drift-cell is hexagonal in shape and has a dimension that is approximately 11.9 mm by 19.0 mm. A cell is defined by a sense wire (made of gold-plated tungsten-rhenium) located at the center of the cell and six field wires (made of gold-plated aluminum) located at the vertices of the abstract hexagon. All cell layers in a particular superlayer are designated as either an axial (A) layer, an U-stereo layer, or a V-stereo layer. Sequential layers are staggered by half a cell width. The stereo layers provide information on the z-location of the track hits. From superlayers 1 through 10, the arrangement is as follows: AUVAUVAUVA. See figure 3.4. The so-called stereo angles of the U, V layer wires varies between ±45 mrad. and ±76 mrad. This entire wire-layer-superlayer setup is enveloped by a gas that permeates the DCH and consisting of an 80:20 mixture of helium and isobutane maintained at an over pressure of 4 mbar.

The basic detection scenario is as follows: charged tracks ionize the gas molecules during their traverse through the detector; the resulting charge "drift" toward the sense wire due to the electric field maintained by the potential difference between the sense wire and the ground wire. These ionized charge, plus any additional charge generated by the avalanche effect, then is collected by the sense wire. Finally, the signal pulse (with avalanche gain) experienced by the sense wire is read by the detector electronics associated with the given drift cell which is mounted on the rear endplate of the detector.

The electronics include integrated amplifier/digitizer boards which take both time measurements as well as the total deposited charge in the relevant DCH channels and store the information in trigger latency buffers for 12.9 μs. If a L1 trigger *Accept* signal is initiated for

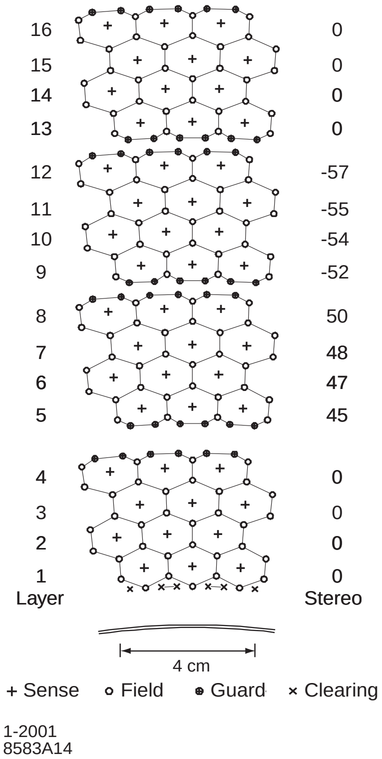


Figure 3.4: Schematic wire layout in the first four DCH superlayers.

the event, then the information contained in the buffer is transferred to the readout buffer. Drift chamber specific algorithms then convert the raw spatial and temporal information into drift times, total charge, and signal quality status (after correcting the time and charge information for time offsets and gas gains etc.)

The charge, drift time, and spatial information from DCH "hits" which form an identified charged track (identified by either trigger codes or by off-line software codes) can be used to compute the energy lost (dE/dx) by the track. The BaBar collaboration reported that the dE/dx resolution achieved using DCH information from Bhabha events was 7.5%, close to the expected resolution of 7.0%.[6] The DCH is also primarily responsible for charged track's p_t measurements. The collaboration reported that the transverse momentum resolution derived using cosmic muons is well represented by the following linear function (see p. 57 in [20]):

$$\frac{\sigma_{p_t}}{p_t} = [(0.13 \pm 0.01)\% \cdot p_t] + (0.45 \pm 0.03)\% \tag{3.2}$$

---

[6] Acquisition of dE/dx information enables PID via the standard Bethe-Bloch equation:

$$-\frac{dE}{dx} \;=\; \frac{Kz^2Z}{A\beta^2}\left[\frac{1}{2}\log\left(\frac{2m_ec^2\beta^2\gamma^2T_{max}}{I^2}\right) \;-\; \beta^2 \;-\; \frac{\delta}{2}\right] \tag{3.1}$$

where K is a constant, z is the charge number of the incident particle, Z is the atomic number of the absorber, A is the atomic mass of the absorber, T_{max} is the maximum kinetic energy that can be imparted to a free electron during a single collision, I is the mean excitation energy in eV, and δ is a density correction factor. A restricted energy loss rate per unit distance also can be found in [24]. A semiclassical derivation of an approximate energy loss expression and modifications due to density effects can be found in §§13.1-13.4 in [25].

In the above expression, the track's transverse momentum (p_t) is given in units of GeV/c.

3.2.3 The DIRC

The primary purpose of the DIRC is to enable charged particle PID using Cherenkov photons emitted by the particle as part of its ionization energy loss during its traversal through silica bars. Specifically, it was intended to provide π/K separation of 4σ or greater from pion Cherenkov threshold up to about 4.2 GeV/c (PID of particles with momentum below 700 MeV/c relies upon DCH and SVT dE/dx measurments.)

When an ultrarelativistic particle travels through certain media, it loses energy via electromagnetic radiation. The phenomenon is related to the coherent response of the material with respect to the particle's radiation (the so-called "density effect".) If there's little absorption of the emitted radiation by the material under consideration and if, in addition, the particle is moving with speeds greater than the phase velocity of light (at some given frequency) in that medium, then the radiation escapes to infinity and we observe the particle emitting photons while it moves through the material [25]. (In contrast, within the DCH and SVT, the particle energy loss is mostly absorbed by the material and then rerelased via ionization of the gas and silicon respectively.) The emitted photons will be arranged in a ring pattern with a characteristic angle defined by the familiar Cherenkov equation:

$$cos(\theta_c) = \frac{1}{\beta n_c} \tag{3.3}$$

where n_c is the relevant medium's index of refraction and θ_c is the Cherenkov angle as defined by the direction of an arc segment of the ring with respect an axis defined by the particle's direction of motion.

One can estimate the number of photons emitted by a particle as part of its loss of energy using a semiclassical argument. Starting from the classical radiative energy loss equation (e.g. eq. 13.79 in [25],) we can write the differential energy loss per unit distance per unit frequency of the particle as follows:

$$\frac{d^2E}{dx\,d\omega} = \frac{(ze)^2}{c^2}\omega\left(1 - \frac{1}{\beta^2\epsilon(\omega)}\right) \tag{3.4}$$

where ze is the incident particle's charge and $\epsilon(\omega)$ is the frequency-dependent dielectric constant of the medium. The above equation is obtained by considering the density effect (following Fermi's reasoning) in the case of particle energy loss (energy lost by a particle and transferred to an approximately harmonically bound charge.) The particle energy loss can be expressed by a dE/dx expression which then is transformed into a frequency integral and applied to the case where the impact parameter of the particle is consistent with little or no absorption by the travsersed medium. See §§13.4-13.5 in [25]. Now substituting the relation $E_{tot} = N\hbar\omega$ (where N is the total number of emitted photons) and defining $E_0 \equiv \hbar\omega$, equation 3.4 becomes:

$$\frac{d^2N}{dE_0\,dx} = \frac{\alpha z^2}{\hbar c}\left(1 - \frac{1}{\beta^2\epsilon(\omega)}\right) \tag{3.5}$$

where $\alpha \equiv e^2/\hbar c$. Assuming $\mu \sim 1$ so that $n^2(\omega) \simeq \epsilon(\omega)$, and then making use of equation 3.3, we obtain the following:

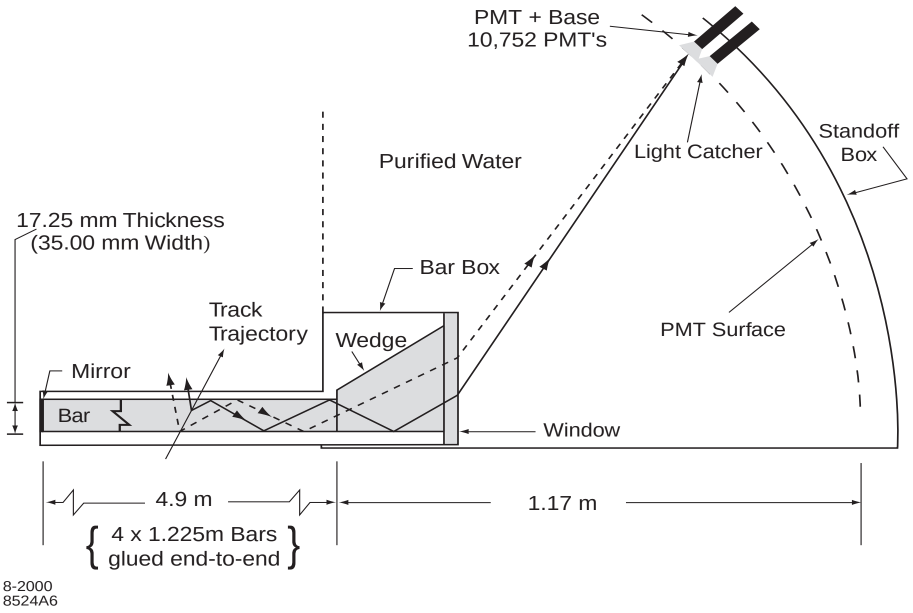


Figure 3.5: Schematic representation of the DIRC's fused silica bars and the imaging region.

$$\frac{d^2N}{dE_0\,dx} \simeq \frac{\alpha z^2}{\hbar c}\,\sin^2(\theta_c) \tag{3.6}$$

The above expression, after integrating over the Cherenkov photon energy range and the distance traveled by the particle, gives an estimate of the number of expected Cherenkov photons (with characteristic angle θ_c) emitted by a relativistic particle with charge ze traveling through some radiator medium.

The DIRC makes use of the above principles to detect the total energy lost by a particle while it traverses the DIRC's active volume. In particular, the DIRC employs synthetic fused silica bars arranged in the shape of a 12-sided polygonal barrel. The silica bars serve as both radiator material as well as photon pipes which transport the Cherenkov photons to the detecting elements of DIRC (photomultiplier tubes) via total internal reflection. The silica bars have mirrored surfaces on one of their longitudinal ends while they are buttressed by a wedge (made of the same fused silica and glued onto the end of the bars) on the "readout" end. 10 mm-thick fused silica "windows" are glued onto other side of the wedges to serve as interface and seal between the bars and the"standoff box". The standoff box is filled with purified water and 10,752 photomultiplier tubes (divided into 12 sectors) are installed on its far curved walls. Figure 3.5 contains a schematic representation of the DIRC (note that a 6 mrad angle on the bottom surface of the wedge is not shown in figure 3.5.)

According to the BaBar collaboration, the DIRC performace closely matches the design goals for π/K separation. In particular, the Cherenkov angle resolution for a single photon is about 10.2 mrad; this was esimated by fitting the signal peak from the measured versus the expected Cherenkov angle for single photons stemming from di-muon events. However, discounting correlated systematic errors, the overall resolution on a Cherenkov angle with N single photon measurements scale as

$$\sigma_{c,track} = \frac{\sigma_{c,single}}{\sqrt{N}} \tag{3.7}$$

where $\sigma_{c,single}$ is the single photon resolution. So for typical track measurements between 20 and 65 photons, the resolution is better than 2.3 mrad. For a more complete discussion of the DIRC's electronics, feature extraction algorithms, calibration methods, data analysis, detector efficiency, and performance, please consult §8 in [20].

3.2.4 The EMC

The EMC is designed to measure electromagnetic showers (cascade electron bremsstrahlung and photon pair production processes) that result from high energy electrons, photons, and neutral pions (and incidentally other charged and neutral hadrons.) It is a total-absorption calorimeter composed of segmented arrays of thallium-doped cesium iodide crystals. The detector readout consists of photodiodes which are matched to the spectrum of the scintillation light. According to the Collaboration, this subdectector's energy resolution is parameterized by the following expression:

$$\frac{\sigma_E}{E} = \frac{a}{[E(GeV)]^{1/4}} \oplus b \tag{3.8}$$

where E is the energy of a photon and σ_E is the RMS error on the photon energy measured in GeV. The detector's position resolution (in terms of angular variables) is given by

$$\sigma_\theta = \sigma_\phi = \frac{c}{\sqrt{E(GeV)}} + d \tag{3.9}$$

Under ideal conditions, a and b are on the 1-2% level while $c \approx 3$ mrad and $d \approx 1$ mrad. The actual performace is degraded from the ideal due to inert material, gaps, electronic limtations, and background contamination.

The physical layout of the EMC consists of a cylindrical barrel and a forward endcap. The barrel contains 5760 crystals arranged in 48 distinct rings, each ring with 120 identical crystals. The endcap holds another 820 crystals arranged in 8 rings. Figure 3.6 contains a schematic representation of the EMC's longitudinal cross-section (only the top half of the detector is represented in the figure.)

In the present study, we use information from the EMC only as part of a cut to select multihadronic events. Specifically, we demand that an event should be accepted for analysis only if there are more than three reconstructed single EMC "bumps" (a "bump" represents a cluster of contiguous crystals with valid signals) and if four or fewer charged tracks have been identified by other subdetectors. Since this is a rather loose cut using EMC information on the event selection level and since the EMC does not otherwise impact the analysis, we will not consider this detector any further; more information on the EMC can be found in §9 in [20].

3.2.5 The IFR

The IFR, in addition to serving as the flux-return region for the solenoidal magnetic field, also is intended to obtain signals that enable muon and neutral hadron identifications. For these purposes, the steel that serves as the magnet's flux return simultaneously acts as a muon filter and hadron absorber.

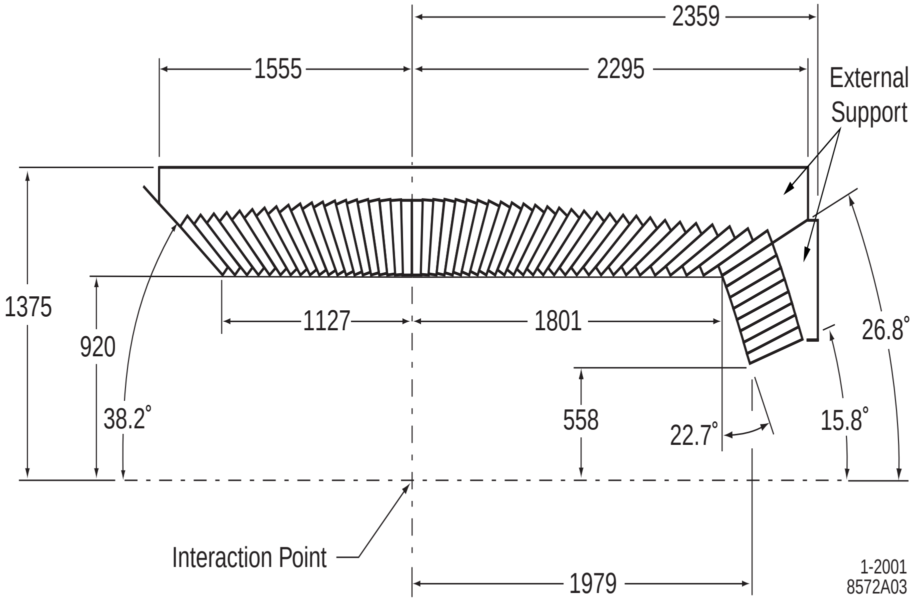


Figure 3.6: A longitudinal cross-section of the EMC. Each rectangle in the figure represents one of the 56 crystal rings.

The steel are segmented into 18 plates increasing in thickness from 2 cm for the inner nine plates to 10 cm for the outermost plates. Single gap resistive plate chambers ("RPC") with two coordinate readout were chosen as the active detecting elements; the RPCs are installed in the gaps of the steel segments. The RPCs detect particle passage from inonizing particles via capacitive readout strips. The time resolution on the readout typically is from 1-2 ns while the position resolution depends on the segment (the collaboration reported that values of a few mm is achievable.)

We do not use the information from the IFR in this analysis and so, in the interest of brevity, we shall not continue discussing this subdectector. More information can be found regarding the IFR in §10 in [20].

3.3 L1 trigger system

Signals from the various subdetectors are analyzed in the first instance by the L1 trigger system. As of 2002, BaBar's L1 trigger system was configured to have an output rate in the 1 kHz range (up to 2 kHz) during normal operations at a designed luminosity level of 3 $nb^{-1}s^{-1}$. It is designed to be over 99% efficient for $B\overline{B}$ events and at least 95% efficient for continuum events.

The two primary trigger modules are the drift chamber trigger ("DCT") and the electromagnetic calorimeter trigger ("EMT";) these modules receive raw data from the DCH and EMC front end electronics about 2 μs after an e^+e^- collision. There is also the instrumented flux return trigger ("IFT") which is used for triggering $\mu^+\mu^-$ and cosmic ray events (these events are used mostly for diagnostic purposes.) Each of these L1 triggers generates information packets called trigger *primitives* which summarize the position and transverse momentum information of particle tracks detected in the DCH and the deposited energy from electromagnetic showers in the EMC. These trigger primitives are then sent to the global trigger ("GLT") every 134 ns.

The GLT processes the various trigger primitives with previously defined masks or prescales to form additional logical triggering objects which are passed onto the Fast Control and Timing System ("FCTS".) If a valid "trigger" is received by the FCTS, then it issues a so-called *L1 Accept* signal to begin data readout from the memory buffers of the various subsystems. The entire process, from the receipt of raw data input by the triggers to the beginning of data readout after the issuance of a *L1 Accept* signal, takes about 11 μs, less than the 12.8 μs limit of the subdectecors' front end electronics' buffer capacity [26].

3.3.1 The DCT

The input data to the DCT consists of one bit of information for each of the 7,104 drift cells in the DCH; these bits contain information on the estimated drift times for signals obtained by the sense wire in each of the DCH cells.

Three distinct code segments then are executed by DCT. First, the Track Segment Finder ("TSF") estimate the ϕ positions and the drift time information obtained by a group of cells in a superlayer and outputs a candidate track segment; these position and time information are then passed to the Binary Link Tracker ("BLT") where the various track segments are linked to form complete tracks. In parallel with the execution of the BLT, transverse momentum (P_T) Discriminator ("PTD") are run; the PTD uses phi position information from track segments in the axial superlayers to search for tracks above some preset p_t threshold.

The TSF code modules group together the drift cells into 1,776 overlapping eight-cell *pivot groups* (consisting of a set of contiguous cells that span all four layers in a particular superlayer; each pivot group has only a single cell in layer 3 which is deemed the *pivot cell* and the shape of the group is designed so that valid track segments can be produced only for relatively straight tracks originating from near the IP. See [20] and [27].

The TSF samples DCH cell signals every 268 ns; after each time increment, the signal state in each cell in the pivot group is represented by a 2-bit counter so that a 16-bit value is formed for the entire pivot group. The 16-bit value is associated with a look-up table which designates the quality of a candidate track segment. When an acceptable track segment from the pivot group is found by the code, the segment quality is compared with the track quality from the same pivot group during the three previous clock increments. The highest weighted (i.e. the best) segment is output to the BLT at a rate of once every 134 ns [20].

If the output from the TSF indicates that a cell from a pivot group has a valid signal that contibutes to a track segment, then that cell signal is designated as a "hit". The BLT then uses the segment hit information from the various TSFs and attempt to link them into complete tracks, starting from the innermost superlayer and proceeding radially outward in a sequential manner. The DCH segments hits are mapped onto the DCH geometry in terms of 320 supercells [20] (32 sectors in ϕ and 10 radial superlayers.) The completed tracks are also given certain designations based upon their geometric characteristics. These track characteristics are then examined by the L1 system to form logical objects: the DCT trigger *primitives.*

Proceeding at the same time as the BLT code modules, the PTD code modules take relatively high quality TSF track segment information from axial superlayers and attempt to determine if that segment is consistent with a track with some specified minimum P_t value. The code starts with *seed* segments from superlayers A7 and A10 and apply search algorithms with respect to the other axial superlayer segments. See p. 94 in [20]. Results from the PTDs also contribute to producing the DCT trigger primitives for the event which are then sent to the GLT.

3.3.2 The EMT

The EMT divides the EMC into 280 *towers* (7 distinct towers in the θ direction and 40 distinct towers in the ϕ direction.) In the EMC barrel region, each tower consists of 24 EMC crystals while in the endcap region each tower consists of wedges of 19-22 crystals each. All energy deposited in the various crystals in a given tower is summed and then sent to the EMT once every 269 ns provided the energy in each crystal exceeds a threshold set at 20 MeV.

The energy data from the EMC towers are then further summed into 40 *phi*-sections (summed over various ranges of θ) and compared to thresholds for the various EMT's logical trigger primitives. The time of energy disposition for the summed energy is also estimated using a finite impulse response digital filter. The EMT information is then encoded within the EMT trigger primitives and sent onto the GLT. See pp.94-95 in [20].

There is also a short discussion of the L1 IFR trigger on p.95 in [20].

3.3.3 The GLT

There are a total of eight level one DCT and EMT trigger primitives. The global trigger uses these eight primitives (plus additional information from the IFT) to form global logical trigger objects which are passed onto the FCTS. The FTCS then uses the global objects to decide whether to issue a *L1 Accept* for the given event.

An example of a GLT trigger object is the so-called “BM”. This consists of the identification of a “B” track from the DCT-BLT (a B track represents a short track reaching to DCH superlayer 5 with a track threshold of 120 MeV/c) in a given event simultaneously with the idenfitication of an ”M” energy cluster from the EMT (an M energy cluster is an energy cluster in one of the 40 ϕ sectors summed over all θ and that exceeds a threshold energy of 100 MeV.) Figure 3.7 contains a graphical illustration of the L1 trigger primitives created by Jason McFall.

Combinations of various trigger objects along with a specified minimum number for each is then employed to determine if an event should be accepted at the L1 level. The collaboration estimated in 2002 that, based on a luminosity of 2.2×10^{33} cm^{-2} s^{-1}, the combined L1 trigger absolute efficiency (from all different combinations of GLT trigger objects that lead to a *L1 Accept*) for $c\overline{c}$ events is 99.9% while that for *uds* events is 98.2% (these include acceptance losses estimated using Monte Carlo simulations.)

3.4 L3 trigger system

The level 3 trigger system consists of software that refines the L1 classification of a given event; the L3 also provides for a set of event selection filters. Note that the L3 software can access the complete set of raw data from a given event as well as the information output from the L1 triggers. Another design goal of the L3 system is to further reduce the accepted event rate from 1-2 kHz (at the L1 level) to under 120 Hz.

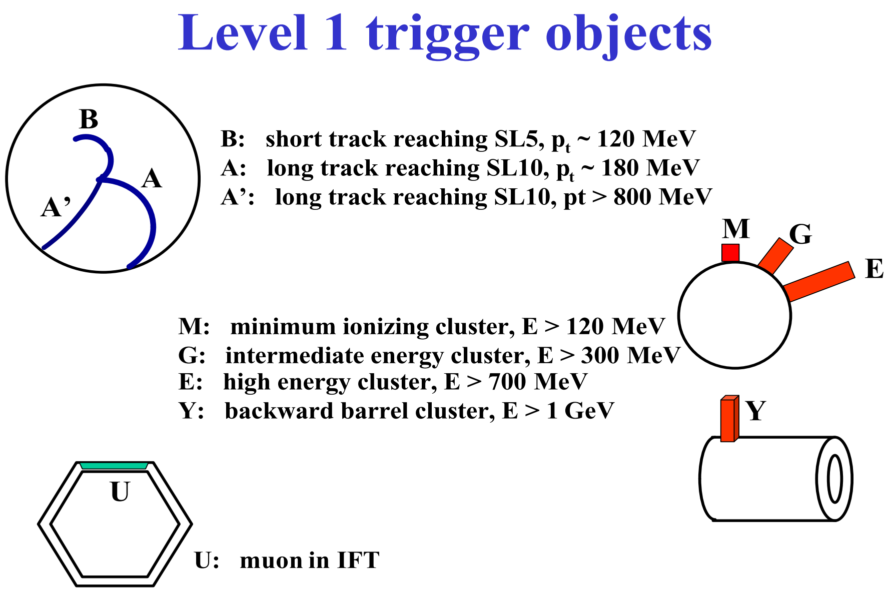


Figure 3.7: A graphical illustration of L1 trigger primitives.

The basic element of the L3 trigger is the "trigger line" with a unique name; each trigger line may be viewed as a decision-path with respect to a given collision event. The L3 first forms "input lines" by applying the logical OR operator with respect to a number of L1 triggers; if an input line has the logical value of "true", then certain specified software scripts are run which typically include the L3Dch and/or L3Emc software reconstruction tools plus a software filter. Finally, the L3 calculates "output lines" by evaluating the logical OR with respect to a number of script flags (with optional veto of certain script flags.) Event logging rate is reduced from the L1 level by requiring the prescaling of events that pass certain L3 trigger lines (e.g. events which are consistent with the characteristic of a Bhabha event in some given polar angle bins are logged once every n'th event; the actual event reduction step takes place at a software level that comes after the L3 system.)

The L3Dch script includes software modules that refine the identification and specification of DCH tracks; it uses TSF outputs to estimate the ϕ position where a track crosses a DCH superlayer and it uses DCH raw data to estimate the electron drift time for signals obtained by individual wires. Then it applies pattern-recognition code with specified search order and search trees to identifiy tracks (this step includes an attempt at the elimination of "ghost tracks" as well as attempts at fitting track segments' ϕ intercepts using least square minimization.) For further information on the L3 track finding code, please see [26] and [27]. The parameters associated with each of the idenfitifed tracks are called DCH "digis" data units.

The L3Emc script converts EMC energy deposit data above a noise threshold of 0.5 MeV from each of the active crystals into EMC digis data units. During the event runs in 2000, the average number of digis per accepted L1 event is 1,100, a number that is dominated by electronic noise and beam induced background. See p.22 in [26]. The EMC digis are created for each EMC clusters by grouping energy deposits in adjacent crystals. The software then calculates various quantities for each group of crystals that are part of a cluster. In particular, a cluster centroid and an average cluster time are calculated as the energy weighted sum of the crystals' center positions and signal peak times. See [26].

L3 also calculates a host of "physics filters"; these are essentially logical "true" or "false" designations assigned to an event based on a prescribed set of cuts. As an example, the "L3THadronicAFilter" is set to "true" if the following criteria is satisfied: there is a minimum of three "good" tracks in the event, the minimum track p_T is 0.5 GeV/c in the cms, the maximum track p_{cm} is 4.5 GeV/c, the minimum invariant mass of good tracks is 2.236 GeV/c^2, and the maximum 2nd Fox-Wolfram moment (this quantitiy will be defined in chapter 4) calculated using "good" tracks in the event is 0.95 (a "good track" in this context is a track with maximum distance of approach to the IP in the xy-plane of 1.0 cm, maximum distance of approach to the IP along the z-axis of 7.0 cm, and a minimum p_T of 0.25 GeV/c.) This set of cuts was valid during the year 2000. Information on other L3 filters is contained in [26].

3.5 OPR filters and offline reconstruction

Both the L1 and L3 are "online" systems in the sense that they are run in tandem with the e^+e^- collisions;. In addition to these systems, there are also additional software filters running in the so-called Online Prompt Reconstruction system ("OPR") which provides for a more detailed event reconstruction as well as outputs data for various collision monitoring systems. There are two types of PR level offline filters, the "DigiFilters" and the "BGFilters".

The DigiFilters use only information from the L1 and L3 triggers to select events; they are intended to remove excess calibration events such as the Bhabha events. Events which pass the DigiFilters are passed to the BGFilters where they undergo classification after additional DCH track finding routines and EMC cluster finding routines have been run (events may be classified as 'multihadron' events or $\tau\overline{\tau}$ events, etc.) The event information is then written to xtc files. For a complete discussion of the configuration of the various offline filters during 1999-2000 (Run1 period,) please consult p.47 et seq. in [26].

Events which are to undergo full event reconstruction are selected based upon the BGFilter classifications. During 2000, about 35% of the events on the xtc files underwent full reconstruction. See p.44 in [26]. Events which have been selected for full reconstruction are sent to the *Logging Manager* ("LM") running on a dedicated server via TCP/IP protocols. The LM writes the data to RAID storage arrays; a single data file is created for each data *run* (typically 2-3 hours of data acquisition.) The data files are archived on tapes. Within 8 hours after data acquisition, the data files are retrieved for full event reconstruction.

Full reconstruction employs raw detector data to construct and compute more complicated structures and quantities from the given event [28]. The reconstruction software code evolves frequently; a consistent set of code acquires a specific codename (e.g. the current Lambda and cascade hyperon analysis is conducted using the so-called "analysis-26" set of code; this code release is also known as 16.1.5.) The results from full reconstruction, along with the run configurations, are stored in object-oriented databases. See pp.18,109 in [20]. In addition to the Objectivity database, there currently exists (as of 2007) another event storage format called *Computing Model 2* ("CM2") which enables users to store additional calculated quantities and selection information. Under CM2, the data may be retrieved for analysis using the ROOT format.

Chapter 4

Event selection

We now endeavor to select a sample of data events which is well-modeled by Monte Carlo simulations. Such a selection would enable us to use the Monte-Carlo light quark and charmed quark samples to confidently correct for the number of data hyperons that were lost due to detector acceptance, software inefficiencies, and explcit cuts. In addition, the presence of a data sample that is well-modeled by the simulation would allow us to obtain a relatively precise estimate of the total number of quark pair-production events in our data; this estimate then would constitute the denominator of the hyperon rate estimates.

We immediately limit ourselves to the off-resonance BaBar data runs gathered from 1999 until 2004[7]. The off-resonance runs involve collisions at a center of mass energy of 10.54 GeV; therefore such a selection cleanly eliminates all the "background" $B\overline{B}$ events at the expense of accepting only about 12% of all BaBar data collected during the relevant period. However, the low fraction of off-resonance (aka 'off-peak') data is balanced by the high luminosity of PEP-II. After the off-resonance cut, we shall be working with a total integrated luminosity of about 21.5 fb^{-1} based upon BaBar Data Quality Group's estimates[8].

For the current event selection studies, we further restrict ourselves to a subsample of the off-resonance data collision events collected during BaBar Run 1 and Run 2 periods. Figure 4.1 contains a graphical representation of the number of off-resonance data events used for our current event selection studies versus BaBar's internal recordkeeping run number.

Our Monte Carlo ("MC") sample includes simulated light quark events, charmed events, bhabha events, muon and τ lepton events, and 2-photon events. All of these MC events are part of the BaBar database and were produced by a set of software codes that is known collectively as Simulation Production 5 ("SP5";) they also have been subjected to BaBar's full detector simulation. The light quark and charm events were generated by the PYTHIA-JETSET Monte Carlo generator developed by the Lund group (implemented via the EvtGen package.) The JETSET generator used a charmed cross section of about 1.35 nb and an uds cross section of about 2.11 nb[9]. Simulated Bhabha events were generated by the "bhwide"

---

[7] This time period corrsponds to BaBar Run 1 through Run 4 data collection periods.

[8] The Data Quality Group has given an 'official' estimate of the off-peak luminosity value for datasets 'skimmed' using release 16.xxx code series. According to a presentation given by Toyoko Orimoto addressing the topic of Physics Validation during BaBar collaboration meeting held on February 24, 2006, the Data Quality Group monitors muon pair-production rates from OPR to obtain an estimate of luminosity for both on-peak and off-peak data. B-counting is used to monitor on-peak luminosity values. These luminosity figures are cross-referenced with each other and with luminosity ratios from other physics processes; see pp.26-28 in [29].

[9] This is according to Anders Ryd's (CalTech) March 14, 2002 email to Colin Jes-

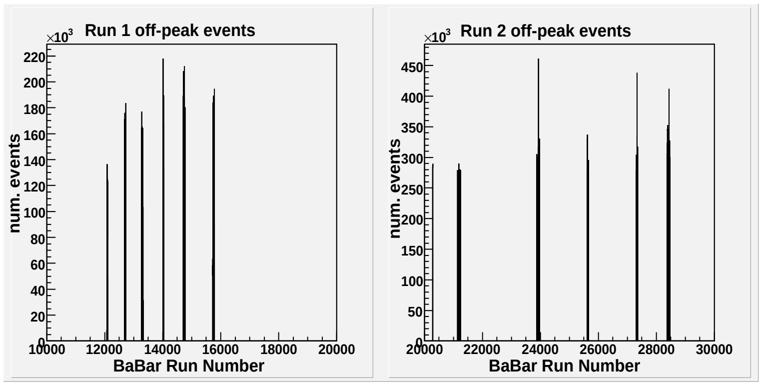


Figure 4.1: BaBar run number vs. number of events used for the current event selection studies.

Runs	Luminosity (fb^{-1})	Number of Events		
		all (no cuts)	‘preselection’	selected (all cuts)
Off-resonance Run 1	1.20	15,000,000	4,645,748	3,479,500
Off-resonance Run 2	1.20	15,000,000	4,548,230	3,355,800
Total	2.40	30,000,000	9,193,978	6,835,300

Table 4.1: Data samples used in event selection studies. The number of events referenced in the table are the absolute number of events accepted after the specified sets of cuts. Definitions for ‘preselection’ cuts and ‘all’ cuts are detailed in §§4.2, 4.3 herein.

event generator package, mu and tau lepton events were generated by the “koralb” and/or ”kk2f” packages. The two-photon events were generated by the “gamgam” package; note that only the $2\gamma \rightarrow 4\pi$ and $2\gamma \rightarrow p^+p^-$ channels are used in our event selection studies.

The total number of data events that we used in the present study is listed in table 4.1; the numbers and types of MC events used here is listed in table 4.2.

Note that in the tables, ‘luminosity’ refers to an integrated luminosity. The integrated luminosity for the continuum data is estimated using the ‘offical’ BaBar Data Quality Group’s estimates while the integrated luminosities for the various MC processes were calculated using Jetset 7.4’s event generator (see pp.74-75 in [21].)

Among other things, we shall compare the data events with a mixture of various types of MC events in order to identify useful selection cuts. After implementing the full set of selection cuts, we hope to acquire a set of surviving data events which generally resembles those from the Monte Carlo mixture that have survived the same set of cuts.

4.1 Event variables

The set of quantities which we will use to construct the event selection cuts include the following:

---

sop from SLAC’s radiative penguin analysis working group; this email is referenced at “http://www.slac.stanford.edu/BFROOT/www/Physics/Tools/generators/”.

MC process	Luminosity (fb^{-1})	Number of Events		
		all	'preselection'	selected
uds	1.531	3,200,000	2,903,702	2,505,798
charm	1.385	1,800,000	1,751,544	1,536,119
bhabha	$\sim 8.75 \times 10^{-4}$	350,000	99	94
μ	0.155	180,000	370	241
τ	1.034	972,000	227,304	30,228
$\gamma\gamma$	~ 0.330	330,000	90,568	18,193

Table 4.2: SP5 Monte Carlo samples used in event selection studies. The procedure by which these samples will be weighted and mixed with each other is described in §4.2 herein.

- *Total Event Energy* ... This quantity is the sum of the energy from all charged tracks and all neutral energy clusters from the EMC which have not been matched to any charged tracks.

- *2nd Fox-Wolfram Moment* ... In order to calculate this quantity, we use reconstructed charged tracks and recontructed neutral energy clusters and calculate the following:

  $$H_L = \sum_{i,j} \frac{|\vec{p}_i|\,|\vec{p}_j|}{E^2}\,P_L(\cos(\theta_{ij})) \qquad \text{for } L = 0, 2 \tag{4.1}$$

  In the above expression, the indices i, j span all reconstructed charged tracks and unassociated EMC clusters in the event, P_L is L'th Legendre polynomial, θ_{ij} is the angle between particle/cluster i and j, and E is the total visible energy in the event. The second (normalized) Fox-Wolfram moment ("R_2") then is given by the ratio H_2/H_0. For a back-to-back two-jet event, $R_2 \approx 1$.

- *Sphericity* ... This quantity is designed to give users a sense of whether the set of tracks and clusters in the event appear more or less 'spherical'. We first construct the sphericity tensor:

  $$S^{\alpha\beta} = \frac{\sum_i p_i^{\alpha} p_i^{\beta}}{\sum_j |\vec{p}_j|^2} \tag{4.2}$$

  In the above expression, the sums over indices i, j are with respect to the individual charged tracks and unassociated EMC clusters while the indices α and β represent the momentum components. Next, we diagonalize the matrix formed from the tensor elements and obtain the three eigenvalues with the following definitions: $\lambda_1 \geq \lambda_2 \geq \lambda_3$ (also, $\lambda_1 + \lambda_2 + \lambda_3 = 1$). Then the sphericity of the event is given by

  $$S = \frac{3}{2}\,(\lambda_2 + \lambda_3) \tag{4.3}$$

  This quantity basically sums the squared transverse momentum of the event relative to the event axis; for a back-to-back two jet event, $S \approx 0$ while for an isotropic event, $S \approx 1$.

- *Charged Tracks* ... This is just the total number of charged tracks in the event that have been reconstructed by the pattern recognition codes.

MC process	Cross Section (nb)	Preselection efficiency	Final selected efficiency
uds	2.09	90.7%	78.3%
charm	1.30	97.3%	85.3%
bhabha	~ 40	0.03%	0.03%
μ	1.16	0.21%	0.13%
τ	0.94	23.1%	3.1%
$\gamma\gamma$	~ 1.0	27.4%	5.5%

Table 4.3: Monte Carlo cross sections used in mixture and selection efficiencies taken from 4.2.

- *EMC bumps* ... This is the number of contiguous EMC clusters which have not been associated to any charged tracks in the event.

4.2 'Preselection' cuts

Despite the L1 triggers and the L3 filters, significant numbers of Bhabha and muon events remain in the total data sample. In fact, sample contamination due to these types of leptonic events tend to overwhelm the shapes of various plotted quantities in the data, thereby hindering comparisons between data and MC. To combat this problem, we initially implement a reduced set of event cuts which preferentially remove Bhabha and muon pair events. In the present study, we will refer to this set of cuts as the 'preselection' cuts. These cuts consist of the following:

- $R_2 < 0.98$.
- The event must have at least three charged tracks.
- The event must have a sphericity value greater than 0.02.
- We further require that the event must have *at least* five charged tracks *or, in the alternative* the number of unmatched single EMC bumps must be greater than three. The more severe cut on the number of charged tracks is designed to further reduce leptonic events since some of these events may have 'ghost' tracks or background tracks from beam effects and cosmic showers, thereby boosting their apparent charged track multiplicity beyond two. However, in order to retain hadronic events with low charged multiplicity but many neutrals, we implemented the more severe charged track cut in tandem (via a logical 'OR') with the cut on unassociated EMC bumps.

After implementing the preselection cuts, we venture a comparison between the surviving data and MC mixture samples. The MC mixture is produced by weighing the theoretical cross sections of the various processes by their respective MC preselection efficiencies. Table 4.3 lists the cross sections and efficiencies for the relevant MC processes.

Figure 4.2 offers a comparison between the surviving events in the data and MC mixture after the preselection cuts; the events are plotted in the two dimensional space given by total event energy on the x-axes and R_2 on the y-axes while the number of surviving events in each 2-d bin is encoded by its color; note that the number of surviving events in the MC mix has been normalized to the total number of surviving data events. Figure 4.3 contains a similar comparison between data and MC mixture plotted in energy-sphericity space.

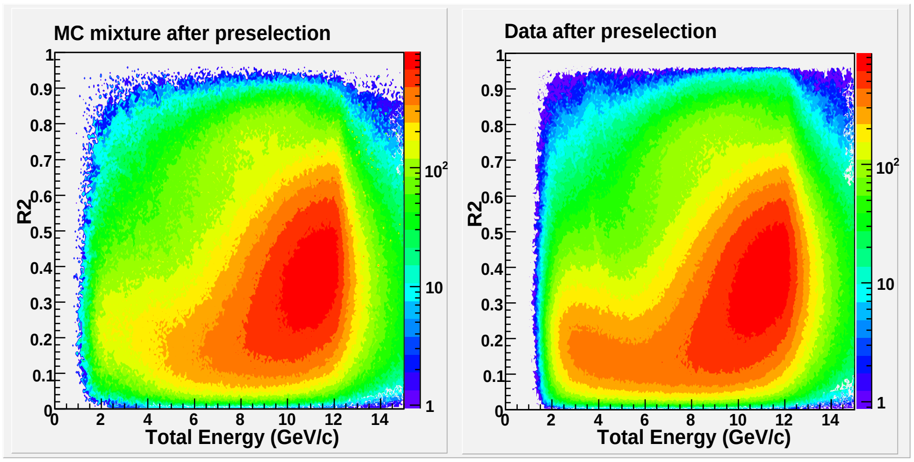


Figure 4.2: Total event energy vs. R_2 for normalized MC mix (left plot) and data (right plot.)

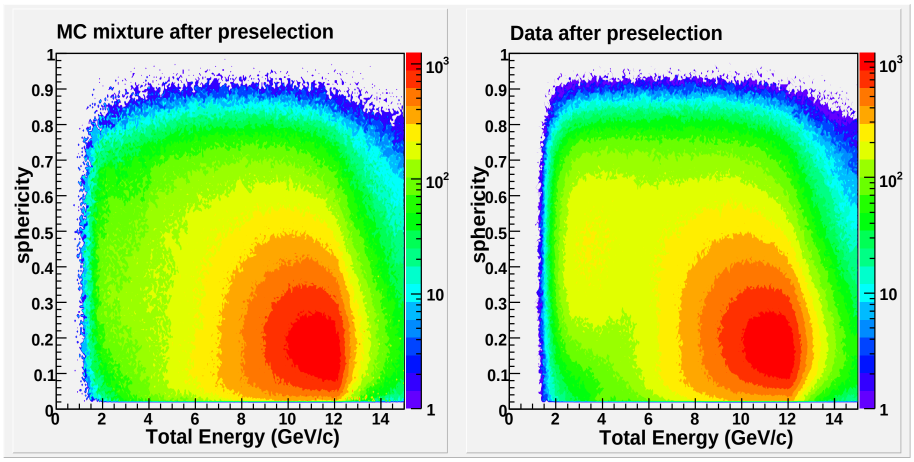


Figure 4.3: Total event energy vs. event sphericity for normalized MC mix (left plot) and data (right plot.)

In figure 4.2, the major difference between the MC mixture and data lies in the area between 2 GeV and 4 GeV in total energy and 0.05 to 0.3 in the R_2 variable. A much more prominent ridge is present in the surviving data events than in the MC mixture; this ridge is mainly due to 2-photon events which are not very well simulated in the MC. Such a difference between data and MC 2-γ events after the preselection cuts is not unexpected since we're using only two discrete 2-photon decay channels to represent all possible 2-γ channels. 2-γ events are also the likely culprits responsible for the observed differences between MC and data in figure 4.3 (in energy-sphericity space, there's a relative enhancement of data events at 2-4 GeV and sphericity values between 0.25-0.7. We shall seek to remedy this discrepancy between the data and MC by cutting away the unwanted 2-photon data and MC events as much as possible in the energy-sphericity space.

There is also another less prominent difference between MC and data in the high R_2 (0.6-0.8) and medium energy (about 6-10 GeV) region: the MC mixture seems to have a small excess in the number of surviving events in this region than the data. This excess is likely due to slightly more MC τ events surviving our preselection cuts. We will attempt to minimize this source of MC and data difference by cutting the τ leptonic events in the two dimensional energy-R_2 space.

In the next section, our goal is to minimize the differences between the data and MC samples by implementing additional cuts to reduce the number of 2γ and τ events. This will have the added benefit of producing a relatively pure sample of hadronic events with small uncertainties in the leptonic and 2γ contamination.

4.3 Additional event cuts

In order to reduce the number of τ-lepton and $2-\gamma$ events in our sample, we institute two-dimensional linear cuts in both energy-R_2 space and in energy-sphericity space; we further introduce a total energy cut at 5 GeV. These cuts are in addition to the cuts referenced in the previous subsection. The 5 GeV energy cut eliminates a large number of 2-γ events with relatively little impact of hadronic events; hadronic events that are eliminated by this cut also are more likely to have significant initial state radiation ("ISR") which reduce their usefulness for studies of continuum particle spectrums.

The details of the 2-d linear cuts are as follows:

- *Energy-R_2 Linear Cut* . . . We require all accepted events must satisfy the following condition:

  $$E_{tot} > \left[\left(\frac{6.0}{0.4} \cdot R_2 \right) - \frac{0.1}{3.0} \right] \tag{4.4}$$

  where E_{tot} is the total charged plus neutral EMC energy in the event and where R_2 is the second normalized Fox-Wolfram moment. This cut is designed to reduce τ-lepton events.

- *Energy-Sphericity Linear Cut* . . . We require all accepted events also must satisfy the following condition:

  $$E_{tot} > \left[\left(\frac{3.0}{0.4} \cdot S \right) + 2.5 \right] \tag{4.5}$$

  Again, E_{tot} is the total charged plus neutral EMC energy in the event, while S is the sphericity calculated for the event. This cut is designed to further reduce $2-\gamma$ events.

- *Total Energy Cut* . . . Additional two-photon events can be eliminated by requiring all events must have total energy greater than 5 GeV/c. This a fairly standard cut for hadronic analyses (see e.g. [30].)

Figure 4.4 displays the effect of the Energy-R_2 cut on MC τ events which have survived the preselection cuts. The peak event area after preselection has high R_2 (greater than 0.6) and medium total energy (6-10 GeV) as shown in the left plot; this peak is almost entirely removed by the linear cut.

However, the Energy-R_2 cut and the total energy cut by themselves let pass substantial amount of $2-\gamma$ events. In order to further reduce these type of events, we instituted the

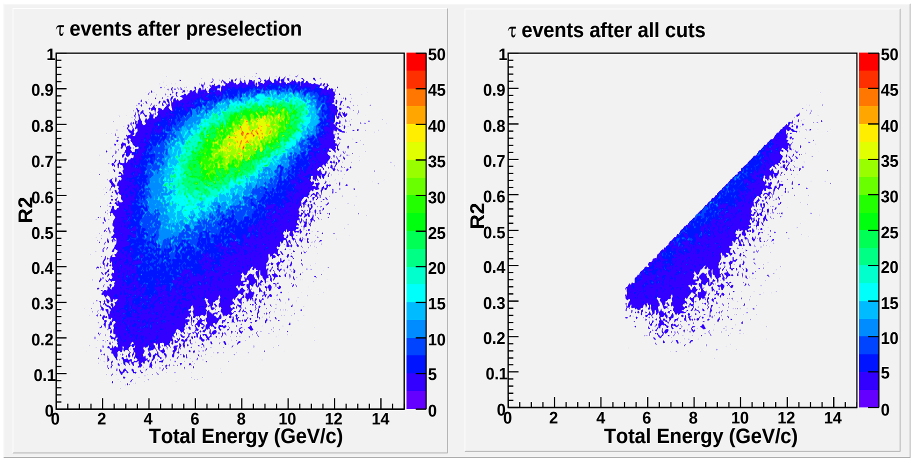


Figure 4.4: Total event energy vs. R_2 for MC $\tau\overline{\tau}$ events that survived preselection (left plot) and after all cuts (right plot.)

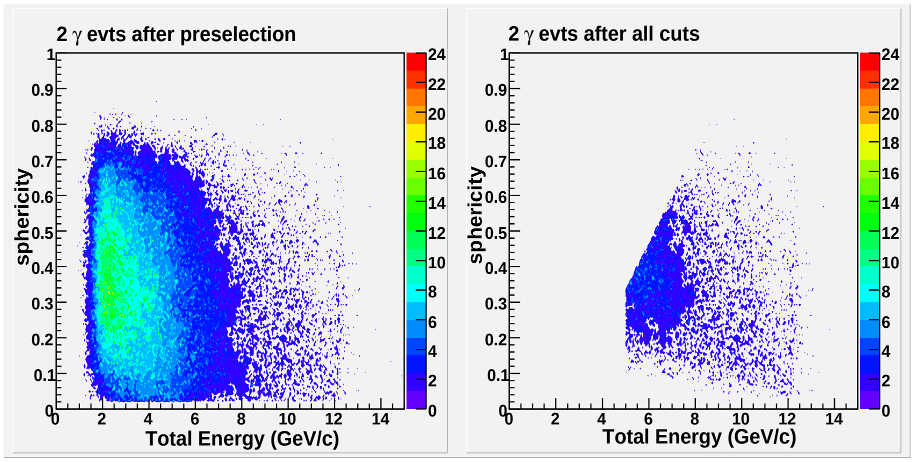
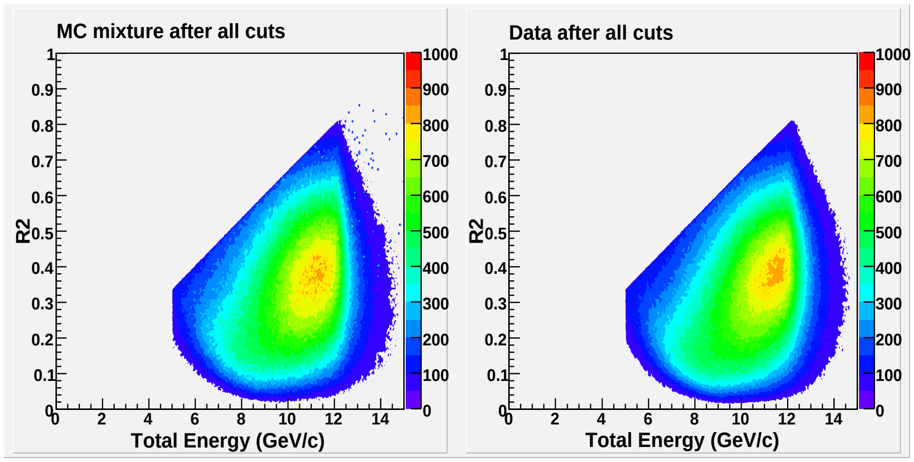


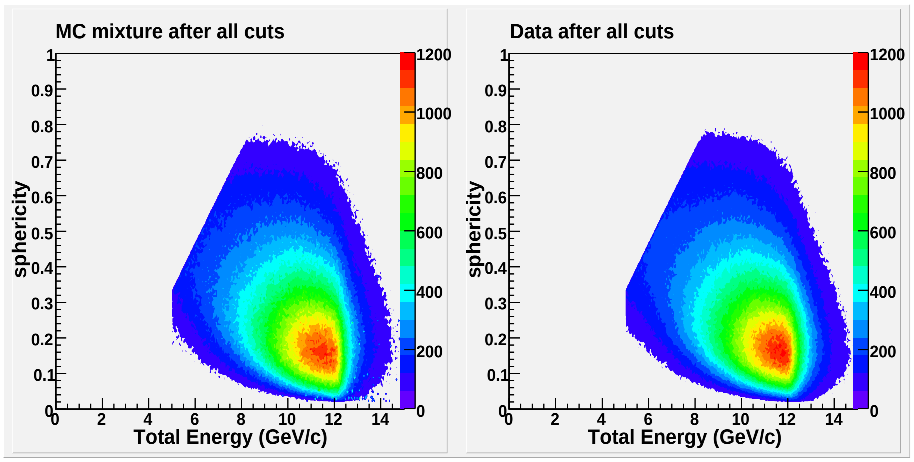
Figure 4.5: Total event energy vs. sphericity for MC $2-\gamma$ events that survived preselection (left plot) and after all cuts (right plot.) The MC sample includes only the 4π and p^+p^- channels.



MC mixture after all cuts

Data after all cuts

Figure 4.6: Total event energy vs. R_2 for MC mixture events that survived all cuts (left plot,) and data events that survived all cuts (right plot.)



MC mixture after all cuts

Data after all cuts

Figure 4.7: Total event energy vs. sphericity for MC mixture events that survived all cuts(left plot) and data events that survived all cuts (right plot.)

Energy-Sphericity linear cut. The effect of this cut is displayed on a sample of MC $2-\gamma$ events ($4\ \pi$ and p^+p^- channels only) in figure 4.5.

After these additional cuts, we're able to reduce the MC τ efficiency from about 23% after preselection to a 3% level while the MC $2-\gamma$ efficiency (for the specified channels) went from about 27% after preselection to the 5-6% level (see table 4.3.) The effect of the linear cuts in energy-R_2 and energy-sphericity spaces can be illustrated by the better agreement between the normalized MC mixture and the data sample in these variables (see Figures 4.6 and 4.7.)

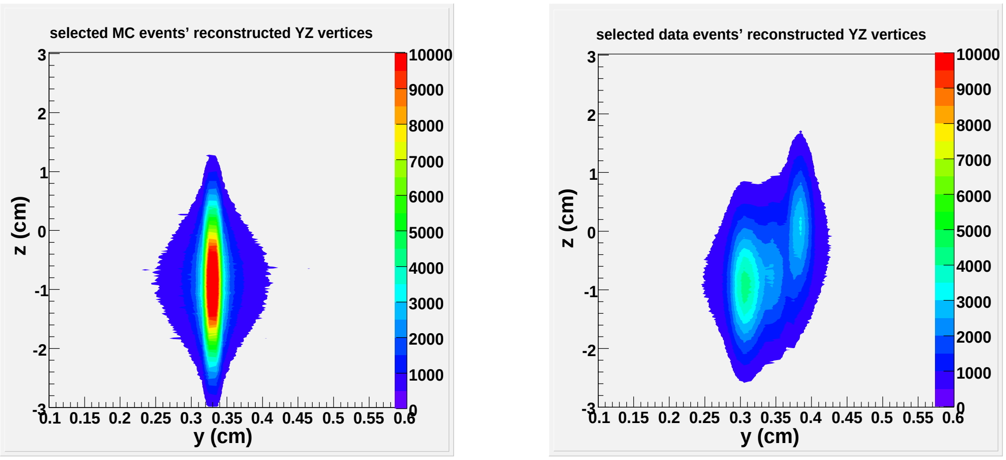


Figure 4.8: Reconstructed Y-Z coordinates of MC and data events selected after all cuts.

Figure 4.8 contains color contour maps of the reconstructed event vertices in BaBar's y-z plane; the plot on the left is that for the mixed MC sample while that on the right is for our data sample. The vertices have been estimated using only charged tracks found in the events and then applying a software routine (the "LeastChiVertexer") to iteratively fit the tracks to a common event vertex by minimizing the χ^2 difference between the tracks and the candidate vertex. The mixed MC sample (total number of MC events has been normalized to that from the data sample) has a single peak in the y-z plane reflecting the fact that all events from the MC sample employed the same detector conditions. The data sample on the other hand appears to have three distinct peaks due to changes in detector condition over the course of three years (2000-2002.)

Figure 4.9 compares the reconstructed event z-coordinates and the event radial (x-y plane) separation from the beam between the data and MC samples. In the plot on the left side of the figure, the blue spheres represent the data z event coordinates while the red line details the same variable from the mixed MC sample; the slight off-set in the z-coordinate peak between MC and data samples is due to the varying detector conditions in the data sample. In the plot on the right in the same figure, the red triangles represent the radial (xy) separation between the event vertex and the beam while the blue line represents the vertex-beam radial separation distance in the MC sample; since this quantity reflects a *difference* between two detector coordinate quantities, the agreement between MC and data is much better than MC-data comparison plots of single coordinate quantity. Note also the black spheres in the plot on the right: these represent the vertex-beam radial separation distance in the data after only the preselection cuts; the apparent 'bump' in the preselection data at radial distance of about 2.5 cm is due to events where the beam, as a result of one of the magnets near the

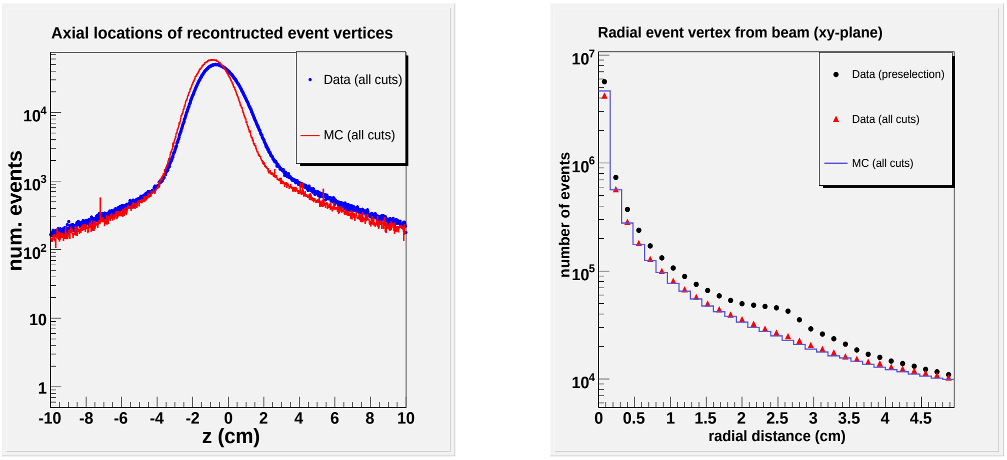


Figure 4.9: (left) Reconstructed z coordinates of MC and data events selected after all cuts. (right) Radial (xy) vertex distance from beam of reconstructed MC and data events.

interaction region, interacts with the beampipe.

4.4 Comparisons between Data sample and mixed MC sample

It is somewhat inconvenient to gauge the agreement (or disagreement) between the MC mixture and data using two-dimensional plots; for this reason, we include single variable comparisons between the MC mix and data in figures 4.10, 4.11, 4.12, and 4.13.

Figure 4.10 compares the number of 'good' charged tracks found in the MC and data events. In the present context, a 'good' track is one with at least 5 SVT hits and 10 DCH hits so as to allow adequate track momentum measurements. In the top left plot (after preselection) and in the bottom left plot (after all event cuts,) data is represented by black triangles while the MC mixture (normalized to total number of data events) is represented by the blue line. The two plots on the right are logarithmic plots of the shapes of the individual components of the MC mixture after preselection and after all cuts. The number of good charged tracks is the only one where the MC and data sample appear to agree better after the preselection cuts than after all cuts. Although the MC mixture reproduces the data variable shape fairly well after all cuts have been applied to the samples, the MC/data ratios in the lower track number bins tend to greater than one; on the other hand, in track number bins > 7, this ratio generally are less than one. The standard deviation between MC and data in this variable as a percentage of the total number of data events[10] is relatively large: $\sim 3.0\%$. However, this level of percent deviation varies by no more than about 0.2% when we apply more stringent vertex cuts and different MC light and charm quark cross sections (these

---

[10] The explicit percent standard deviation caculation is as follows:

$$PSD = \left[\frac{\sqrt{\sum_i \left(N_i^{mc} - N_i^{data}\right)^2}}{N_{tot}^{data}} \right] \tag{4.6}$$

where the index i spans all the bins, N_i^{data} is the number of data events in the i'th bin, N_i^{mc} is the number of mixed MC events in the bin, and N_{tot}^{data} is the total number of data events.

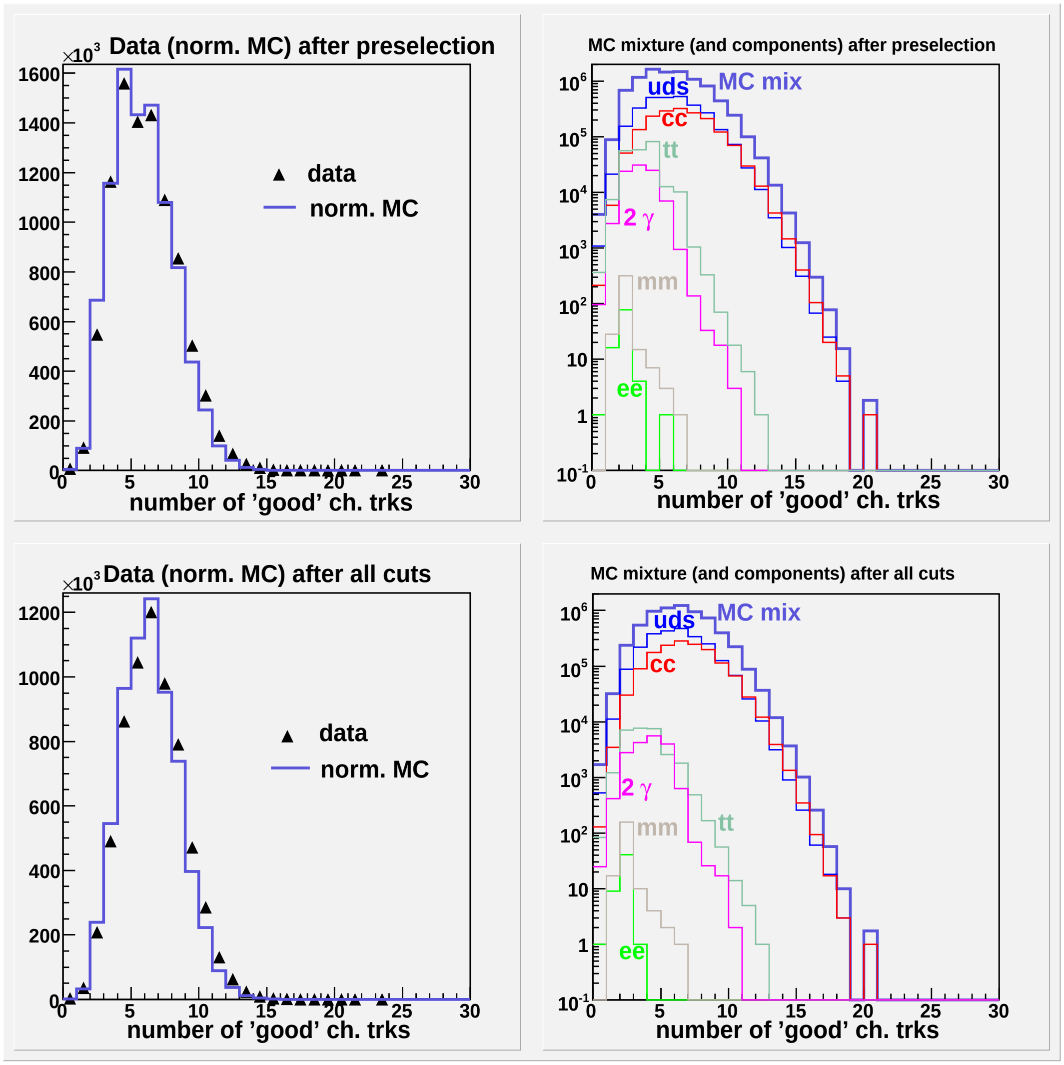


Figure 4.10: 'Good' charged tracks comparisons between MC mixture and data. Proceeding clockwise starting from top left plot: data vs. MC mix after preselection; MC componenets after preselection; MC components after all cuts; data vs. MC mix after all cuts. Note that 'mm' denotes $\mu\overline{\mu}$ events and 'tt' denotes $\tau\overline{\tau}$ events.

variations will be discussed in more detail in the next subsection; in particular see table 4.6.) Hence it is more likely that the observed differences between the MC and data samples in the number of event charged tracks is due to the MC sample's inability to reproduce data well rather than to contamination by beam-gas events or to an underestimation of hadronic events relative to leptonic and $2-\gamma$ events.

Figure 4.11 compares the distribution of the event sphericity in the MC and data samples. The top left plot in the figure contains the data and normalized MC samples' event sphericity distribution after the preselection cuts while the bottom left plot in the figure contains the corresponding distributions after all cuts have been applied to the samples. Once again, the blue lines represent the normalized MC sample while the black triangles represent the data sample. The effects of the linear cuts against $2-\gamma$ and τ events are reflected by the much better agreement between the shapes of the MC and data samples in the lower left plot. The percent standard deviation between data and MC is about 0.2% while variations in cuts and cross sections produce a maximal difference of another 0.03%. See table 4.6.

Figure 4.12 compares the distribution of the event R_2 in the MC and data samples. Again, the top left plot in the figure is for the case after application of the preselection cuts while the bottom left plot relates to the case after all cuts. The plots on the right side of the figure detail the contributions of the various MC components to the MC mixture. The selected MC mixture reproduced the shape and behavior of the selected data; the percentage standard deviation of the MC from the data is about 0.5% with maximal variations (from variations in cuts and cross sections) of 0.2%. See table 4.6.

Finally, Figure 4.13 compares the distribution of the summed charged and neutral event energy in the two samples. Note that in the upper left plot in the figure (MC and data sample energy distribution after application of only the preselection cuts,) there exists a prominent shoulder in the data (2-4 GeV region) that is not well-reproduced by the MC; this feature is due mainly to $2-\gamma$ events. After the application of the additional linear cuts and the energy cut, this shoulder has completly disappeared in both the data and MC and the two samples agree quite well in their general shape. The percentage standard deviation between the two samples is 0.3% with maximal additional variation of $< 0.1\%$. See table 4.6.

4.5 Estimated hadronic event purity in data sample

Using the calculated MC cross sections and selection efficiencies from table 4.3, the estimated hadronic event purity of the selected MC sample is 96.6%. We shall multiply the total of selected data events in this analysis by the event purity estimate in order to obtain an estimate of the number of hadronic events in our full sample; this will serve as the denominator in the estimate of the various hyperon rates. An alternative way to estimate the total number of hadronic events in a sample is to multiply the integrated luminosity of the full sample by the estimated hadronic cross sections (after correcting the cross sections using comparisons between the data sample a mixed MC sample, see e.g. [30].) We have opted to use the $(number \cdot purity)$ approach at present.

In order to examine the level of uncertainties in the 96.6% purity estimate, we have opted to vary both the selection cuts and the quark cross sections. The resulting variations in the purity estimates, in addition to the differences between the MC and data samples, should allow us to arrive at a reasonable estimate of the level of uncertainty in the purity.

First, we vary the light quark and charm quark cross sections' contribution to the MC

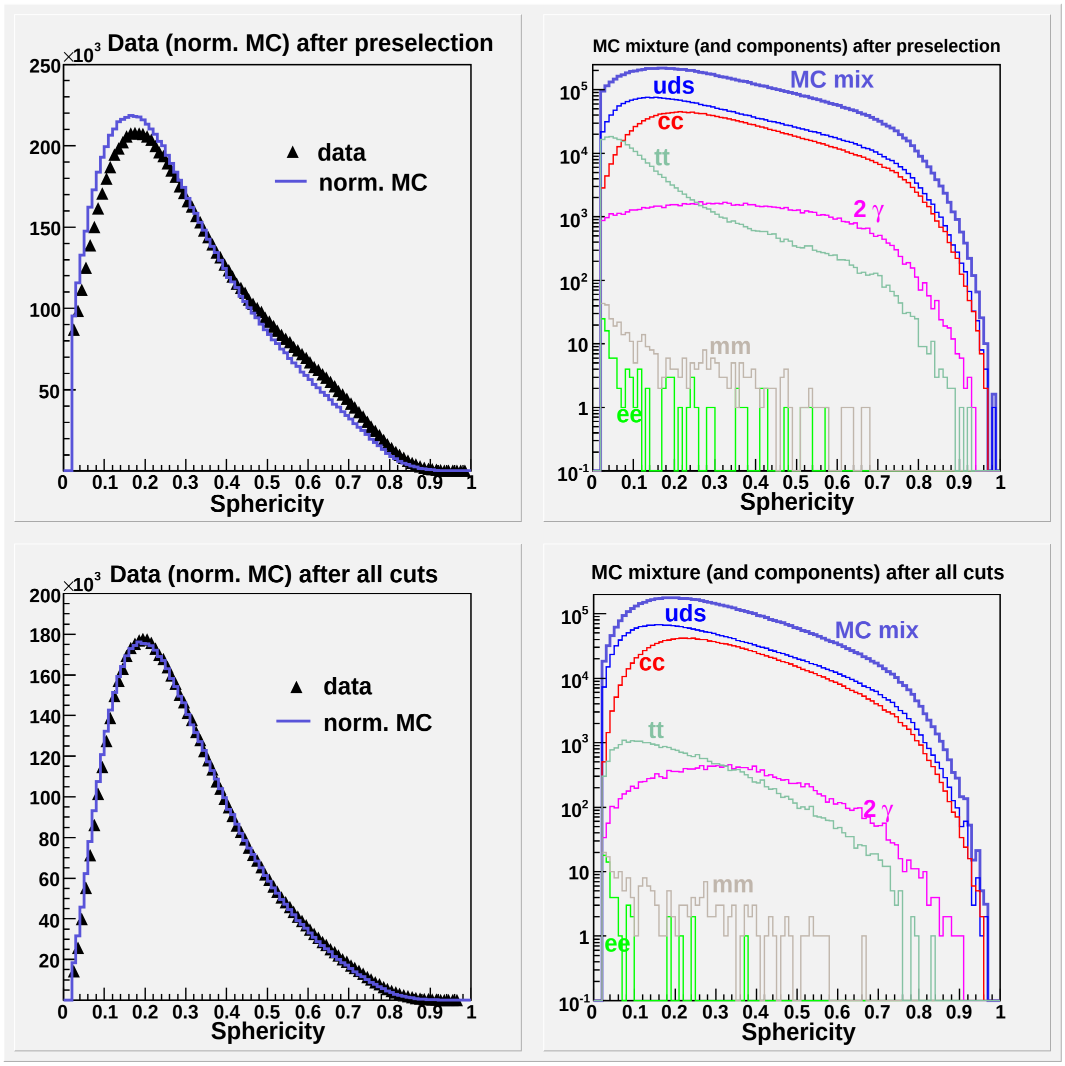


Figure 4.11: Event sphericity comparisons between MC mixture and data. Proceeding clockwise starting from top left plot: data vs. MC mix after preselection; MC componenets after preselection; MC components after all cuts; data vs. MC mix after all cuts.

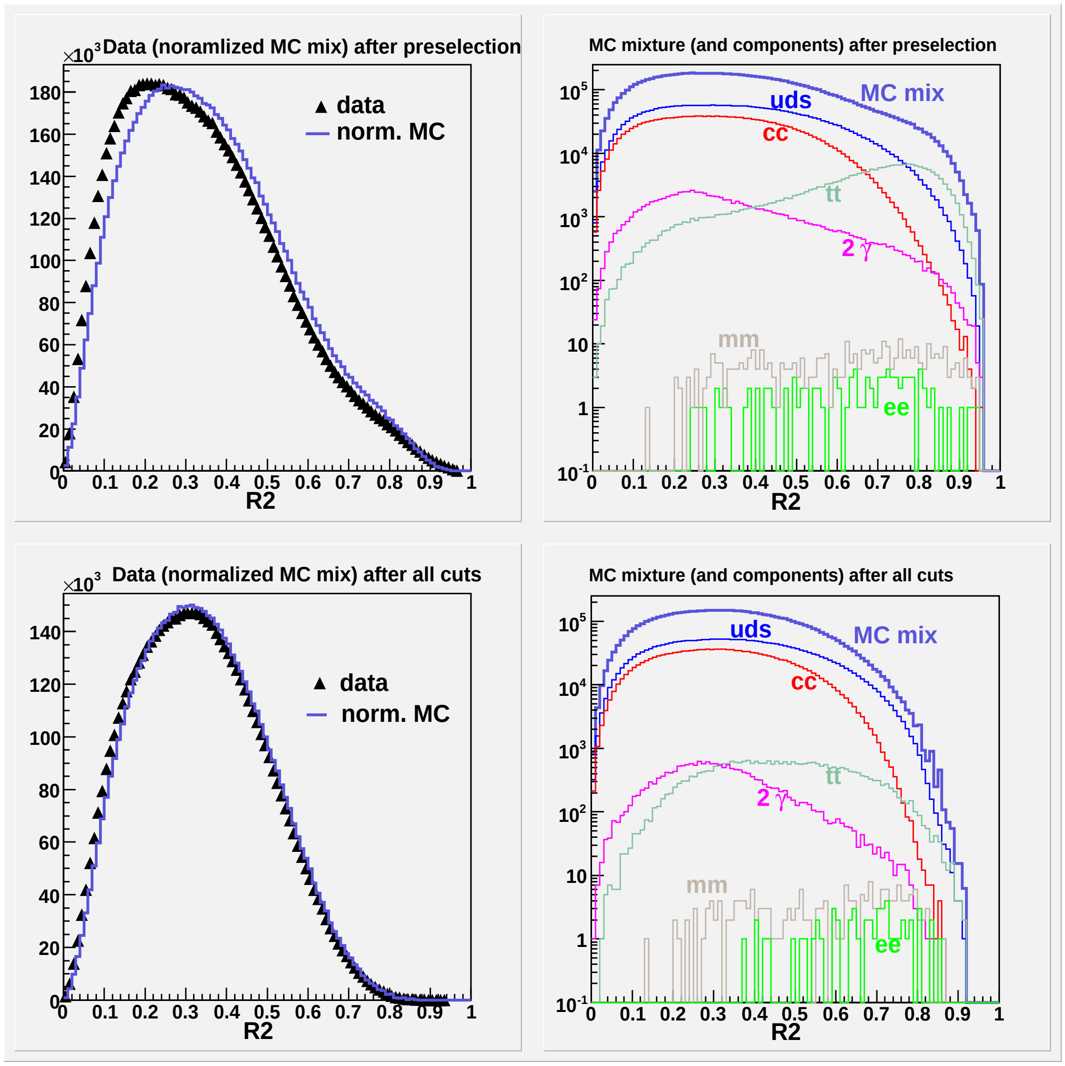


Figure 4.12: Event R_2 comparisons between MC mixture and data. Proceeding clockwise starting from top left plot: data vs. MC mix after preselection; MC componenets after preselection; MC components after all cuts; data vs. MC mix after all cuts.

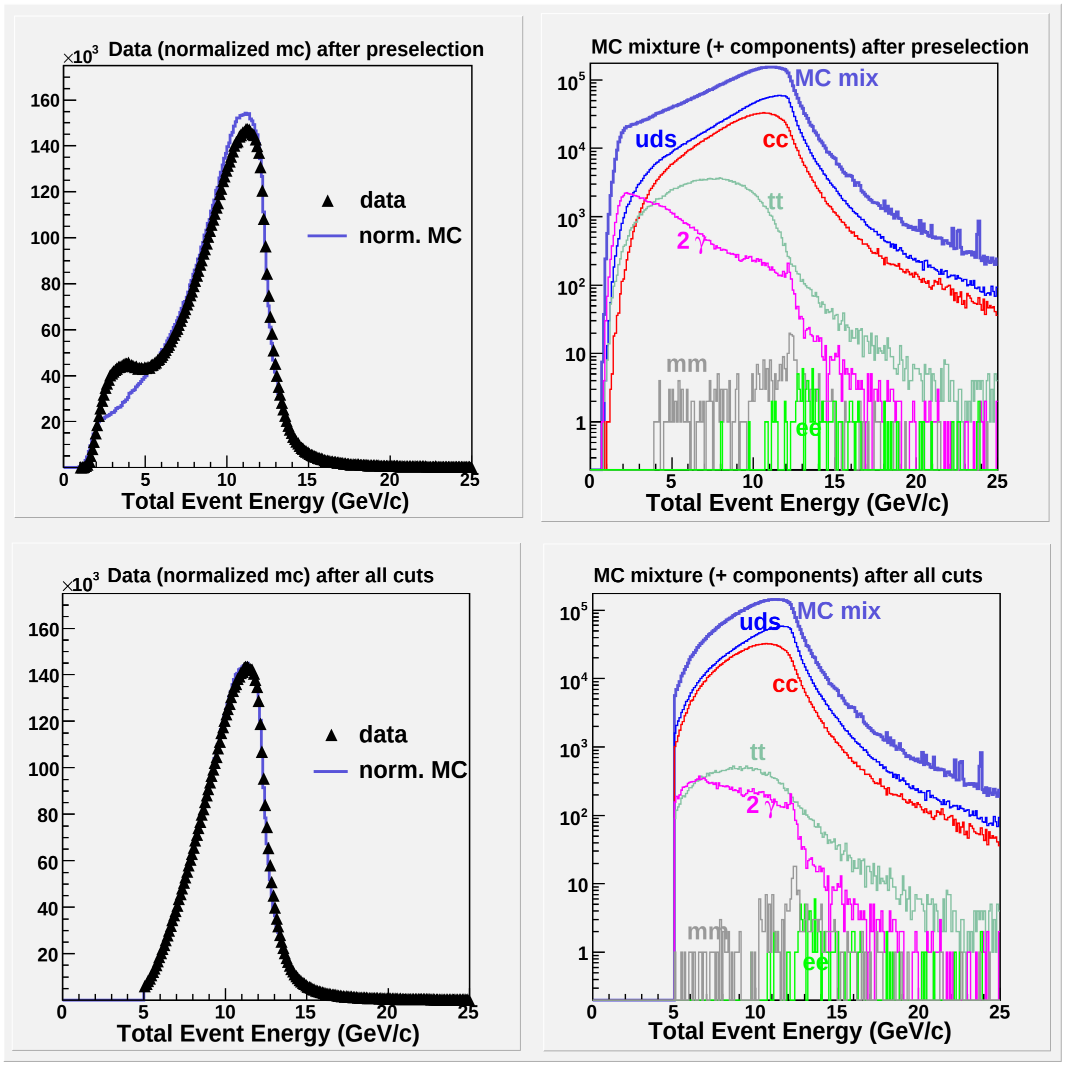


Figure 4.13: Event total energy comparisons between MC mixture and data. Proceeding clockwise starting from top left plot: data vs. MC mix after preselection; MC componenets after preselection; MC components after all cuts; data vs. MC mix after all cuts.

MC process	Cross Section (nb)
uds	2.11
charm	1.35
bhabha	~ 40
μ	1.16
τ	0.94
$\gamma\gamma$	~ 1.0

Table 4.4: Alternative Monte Carlo cross sections used in mixture.

MC process	# Events all	All + z-cut Efficiency	All + z-cut + xy-cut Efficiency
uds	3,200,000	75.36%	61.10%
charm	1,800,000	82.52%	66.05%
bhabha	350,000	0.03%	0.02%
μ	180,000	0.13%	0.12%
τ	972,000	2.77%	1.82%
$\gamma\gamma$	330,000	5.22%	4.33%

Table 4.5: SP5 Monte Carlo process efficiencies with added vertex cuts

sample. Table 4.3 contains the light quark cross section and the charmed cross section which we employed to 'mix' our MC sample; the cross sections are 2.09 nb and 1.30 nb respectively and are taken from the "official" BaBar figures referenced in [21]. As an alternative, we use the quark cross section numbers employed in SP4 MC production [30], here reproduced in table 4.4.

Using the cross section values in table 4.4 with the process efficiencies detailed in table 4.3 results in an estimated event purity of 96.8%.

Our event selection cuts did not explictly include cuts to reject beam-gas and other beam related events. To examine the effect of additional beam-gas cuts on purity estimates, we instituted two *separate* sets of additional event cuts:

- The absolute value of the reconstructed event z-coordinate must be less than 5 cm from the BaBar coordinate origin.
- The absolute value of the reconstructed event z-coordinate must be less than 5 cm from the BaBar coordinate origin, and *in addition* the radial (xy) separation between the reconstructed event vertex and the beam must be less than 0.5 cm. (This is a much more stringent set of cuts.)

The MC process efficiencies of the above cuts (in addition to our full set of event cuts) are contained in table 4.5.

Despite the changes in selection efficiencies due to the added vertex cuts, the event shapes between the mixed MC sample and the data sample varied little. The maximal variations in the percent standard deviation between MC and data due to the additional cuts on the

Cut Type	Total Energy	# Charge Tracks	2nd F-W Moment	Event Sphericity
All Cuts	0.299%	2.97%	0.570%	0.290%
All + z-cut	0.309%	2.91%	0.565%	0.269%
All + z/xy-cuts	0.318%	3.00%	0.557%	0.297%
$\Delta\ \sigma_{udsc}$	0.376%	2.78%	0.389%	0.295%

Table 4.6: Percent standard deviations between MC mixture and data in the relevant event variables categorized by the type of cut applied.

vertices are 0.02% in total event energy, 0.06% in number of charged tracks in event, 0.02% in event R_2, and $< 0.1\%$ in event sphericity. This suggests that beam-gas and other beam-effect events contaminate our selected sample by $< 0.1\%$ (see table 4.6.)

The variation in quark cross section resulted in a purity estimate variation of about 0.2% while additional event vertex cuts suggests that contamination by beam-effect events is less than 0.1%. Since these variations are from two completely independent sources, we can add them by quadrature; after rounding off to the nearest 0.1%, we obtain an estimated purity uncertainty of 0.2%. We take as our central purity value the figure arrived at using the official BaBar cross sections; thus we quote the purity as $96.6\% \pm 0.2\%$.

Chapter 5

Tracking corrections for beampipe tracks

In the next section we will use any two oppositely charged tracks from an event to form composite neutral objects. This is the fundamental "vertexing" step of the Λ hyperon analysis since the Λ signal should be present in the invariant mass plot of the neutral combinations that result from oppositely-charged track combinations.

Our initial focus will be to detect Λ hyperons that originate near the event vertex and which decayed inside the detector's active volume such that the daughter particles have deposited sufficient energy to enable their detection. The relevant decay process is $\Lambda \longrightarrow p^+\pi^-$ (and its conjugate.)

When compared with most of the so-called resonant states that result from QCD processes, Λ hyperons have very long decay lengths [11]. Thus significant numbers of these neutral particles (produced either during the fragmentation process itself or from decays of higher mass resonant states) will not decay until they have reached the SVT and even the DCH. However, the interaction between Λ and detector materials is less well understood than comparable interactions between stable particles (e.g. p, π, and k) and material; this renders the simulation of Λ material interaction less trustworhty and complicates any contemplated corrections for differences between such simulations and the data. Fortunately for us, the integrated luminosity of BaBar's off-peak dataset is sufficiently large as to provide a statistically rich sample of Λ's that have decayed within the PEP-II beampipe across the entire relevant momentum range; if we limit ourselves to such "beampipe" Λ's, we then would only need to tackle the easier problem of correcting for the inadequacies that are present in the MC modeling of interactions between detector materials and Λ's stable decay daughters (p, π, k etc.)

5.1 BaBar modeling of material interactions in the detectors

The current detector simulation code employed at BaBar in based on the object-oriented C++ Geant4 toolkit (the name of the relevant package is BOGUS, an acronym for "BaBar Object-oriented Geant4-based Unified Simulation".) Simulations of detector material, track energy deposition, particle transport[12], and particle interaction with material all are implemented

---

[11] The Particle Data Group gives the following value: $c\tau \simeq 7.89\ cm$. See [24]

[12] BaBar currently uses its own customized transportation code based on Kalman filters.

within this framework.

Some of the models used to simulate hadron-nucleon interactions in physics detectors employed by Geant4 are the so-called "parametrized models". Geant4 uses a low energy parameterized model for incident particle energy in the range from 1 GeV to 25 GeV (see Chapter 20 in [31].) According the Geant4 physics manual, under the paratermized models only the first hadron-nucleon collision is simulated in detail while secondary interactions are simulated by generating additional hadrons and then treating them as secondaries from the initial collision (the numbers, types, and distributions of the extra hadrons are determined by functions which were fitted to experimental data in order to reproduce general trends in hadron-nucleus collisions.)

Geant4's low energy parameterized model have been applied to the pions, kaons, nucleons, and the hyperons. It first calculates a mean number of hadrons produced in a hadron-nucleus collision given by

$$N_m \;=\; C(s)\, A^{1/3}\, N_{ic} \tag{5.1}$$

where A is the atomic mass of the nucleuss, N_{ic} is the approximate number of hadrons generated in the initial collision, and $C(s)$ is a function that depends only on the center of mass energy. If the function $C(s)$ includes the energy-dependent cross section of the interaction, then the above equation can be interpreted as the number of target nucleons excited by the initial collision. Once the mean number of hadrons (N_m) has been calculated, the total number of hadrons in the intra-nuclear cascade is sampled from a Poisson distribution about the mean.

BOGUS' implementation of Geant4 during BaBar simulation productions, including the production of SP5 events, specifically employs the low energy parameterized model to simulate proton-proton and proton-antiproton interactions [32].

5.2 BaBar Tracking Group's DCH tracking efficiency corrections

BaBar's Tracking Group ("TTG") pioneered the SVT-based measurement and correction of the DCH tracking efficiency using tracks that satisfy BaBar's "GoodTrackLoose" definition. GoodTrackLoose constitutes a subset of tracks that satisfies the "GoodTrackVeryLoose" definition which in turn is a subset of "ChargedTracks" in the event (this latter grouping includes any reconstructed charge track in the event.) GoodTrackVeryLoose includes any ChargedTrack that satisfies the following set of requirements:

- Track has momentum < 10 GeV.
- Track's distance of closest approach ("DOCA") in BaBar's xy plane to the BaBar origin is < 1.5 cm.
- Track's z-doca to $z = 0$ is < 10 cm.

Any track that satisfies the GoodTrackVeryLoose defnintion, and that further satisfies the following two additional requirements, is then classified as GoodTrackLoose.

- Track's transverse momentum is > 100 MeV.

- Track has 12 or more DCH hits.

TTG makes use of the fact that the DCH and the SVT can independently reconstruct tracks in order to measure both the MC and data tracks' DCH relative efficiencies. TTG first defines a raw DCH track relative efficiency as follows:

$$\epsilon_{raw} = \frac{\text{Number of GoodTrackLoose with } N_{SVT} \text{ SVT hits}}{\text{Number of GoodTrackVeryLoose with } N_{SVT} \text{ SVT hits}} \tag{5.2}$$

where $N_{SVT} \geq 10$ [33] (according to the TTG, N_{SVT} must be at least 8 to ensure that the track could have been found by the SVT alone.) The raw efficiency however is not the best measure for the tracking efficiency of tracks that pass the GoodTrackLoose definition. BaBar's track reconstruction routine observes the following sequence (see [33], see also [34]):

- DCH track-finding code is run and it produces an initial list of charged tracks; the DCH hits associated with the tracks found after this step are then marked as "used" and are unavailable for further track reconstruction usage.

- The initial list of tracks are extended into the SVT and SVT hits consistent with the track trajectory are then marked as "used" and added to the track. This step is performed by the code module titled "TrkSvtHitAdder".

- The SVT track-finding code is run over the remaining SVT hits which have not been marked as "used".

- The SVT tracks found from the previous step are exteneded into the DCH; consistent and "unused" DCH hits then are added to these additional tracks. This step is perfomred by the "TrkDchHitAdder" code module.

The "TrkSvtHitAdder" bias the observed DCH track efficiency upwards while the "TrkDchHitAdder" both bias the observed efficiency upwards as well as results in an actual increase in the track efficiency. (See discussion on p.3 in [33].) The TTG evaluated the effects of these code modules by running MC and data samples with and without them and then comparing the results. The changes in DCH efficiencies due to "TrkSvtHitAdder" is always less than 0.5% over differing DCH high voltage seetings while the efficiency changes due to "TrkDchHitAdder" is always less than 0.02% (see. p. 17 in [33].)

There is also the issue involving the so-called "SVT ghost tracks". Certain random combinations of noise hits in the SVT may be reconstructed by the software as a track (there are also ghost tracks in the DCH resulting from the assignment of energy deposition into neighboring DCH cells resulting in the "doubling" of a track.) The TTG estimated the effect of ghost tracks by selecting Bhabha events and matching two tracks to EMC clusters in such events; then any additional tracks in such events are assumed to be a ghost track. The measured DCH efficiency then is given by

$$\epsilon = \frac{\text{\# of GoodTrackLoose with } N_{SVT} \text{ SVT hits}}{(\text{\# of GoodTrackVeryLoose with } N_{SVT} \text{ SVT hits}) - (R_g \cdot N_{evt})} \tag{5.3}$$

where R_g is the number of estimated ghost tracks in Bhabha events divided by the total number of such Bhabha events, and where N_{evt} is the number of hadronic events used in the tracking analysis.

The TTG examined the MC and data DCH relative efficiencies in six P_T bins, six "lab angle" bins (lab angles are the angles that tracks make with respect to BaBar's z-axis,) three ϕ angle bins (angles that the tracks make with respect to BaBar's x-axis,) and five track multiplicity bins ("multiplicity" denotes the number of charged tracks in a given event,) for a total of 540 bins. The ratio of data efficiency to MC efficiency then gives the correction factors.

Here, we intend to extend the TTG's results to allow for particle species-specific comparisons between MC and data DCH relative efficiencies for all tracks that originate within the BaBar beampipe. As a baseline comparison, we employed the TTG's track definitions and computed the DCH relative efficiencies over all tracks from Runs 1 through 4 off-peak data *except* that in our case we've relaxed the cut on the tracks' XY-DOCA to the event vertex from 1.5 cm to 2.4 cm in order to encompass almost the entire two-dimensional transverse area of the beampipe. Figure 5.1 compares our baseline results using all tracks from the off-peak dataset (and divided into positively and negatively charged tracks) with the TTG results; they are both displayed in the track transverse momentum variable.

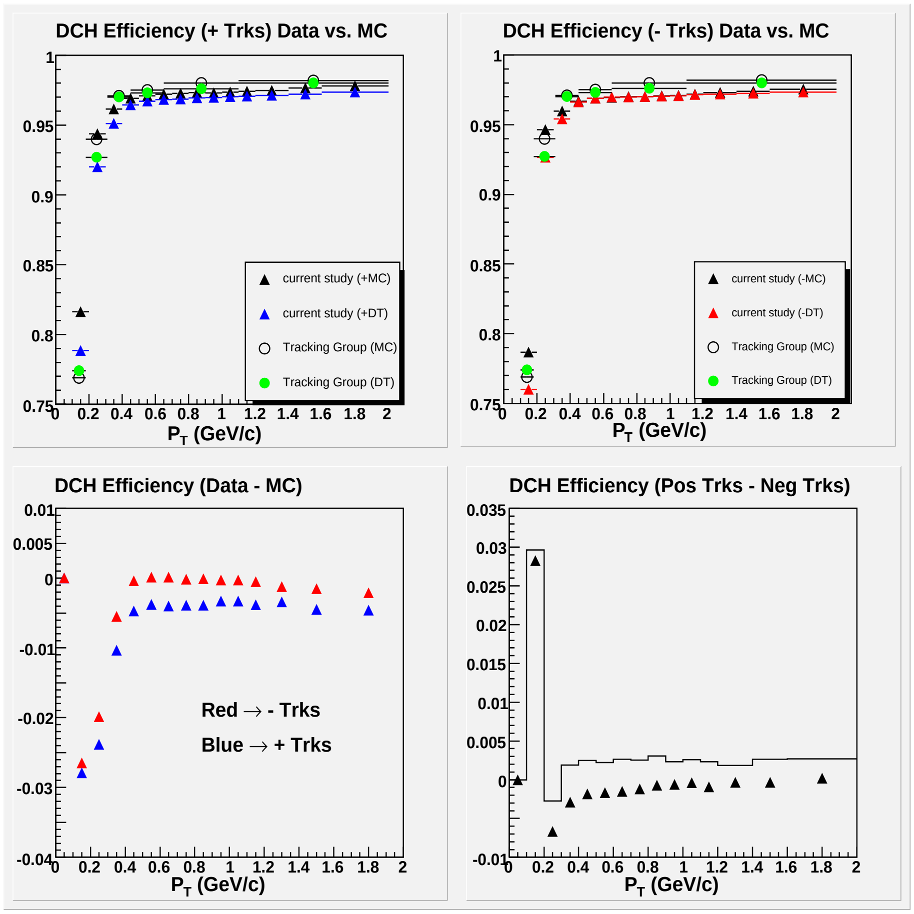


Figure 5.1: (Top left) Averaged positively-charged tracks' DCH relative efficiency in P_T bins; comparison of MC and data in current study and TTG study. (Top right) Similar comparison for negatively charged tracks. (Bottom left) Data/MC DCH efficiency difference for + and - tracks in current study. (Bottom right) Efficiency charge asymmetry for data (triangles) and MC (histogram) in the current study

Note that the DCH relative efficiencies found in the current study is generally slightly lower than the TTG results for $P_T > 0.3$ GeV/c. While some of this difference may be due to our relaxed XY-DOCA cut on the tracks (possible loss due to acceptance issues,) it is likely that most of this difference between the TTG results and our results for all tracks is due to the differing data and MC samples; the TTG employed MC and data taken mostly at the $\Upsilon(4s)$ resonance while we used only off-peak data and MC without $B\overline{B}$ event contributions.

5.3 Method for correcting $\pi^{\pm}$ MC/Data tracking differences

There are several drawbacks present in the TTG's DCH tracking correction factors that preclude their direct application in the present analysis.

First and foremost, the TTG method does not take into account the species of the charged particle; since pions are the most prevalent hadronic end-products in the selected collision events, the TTG correction factors are dominated by the effects due to pions but with significant contributions from other hadronic species. This state of affairs is rather inconvenient for our purposes since the first step in our baryon analysis requires us to properly correct for the Λ's proton and pion track efficiencies.

A second problem with the TTG correction factors is that they have been derived using relatively tight cuts on the number of SVT hits and on the XY track DOCA to the event's primary vertex. In particular, we want to restrict our Λ analysis to particles that decayed to protons and pions *within* the beampipe (located at a XY-radial distance of about 2.5 cm in BaBar coordinates.) However, the TTG corrections apply to tracks with at least 10 SVT hits and that have decayed within 1.5 cm from the event primary vertex. In order to maximize the number of Λ's that can be used in the present analysis, we ideally would like to relax the TTG's SVT hit and DOCA restrictions.

To address the first of the above identified problems with the TTG correction factors (i.e. the mixing of different track species,) we shall adapt the basic TTG correction method to analyze the MC/data differences of known pions and protons but *without* using dE/dx-based particle identification (which would generate additional sources of systematic errors.) For the case of pions, we shall do this by using tracks which can be vertexed with oppositely charged tracks that result in an neutral composite object with an invariant mass near 497 MeV; the resulting mass peak will consist mostly of k_{short}'s that decayed via the $k_{short} \rightarrow \pi^+\pi^-$ pathway.

5.3.1 DCH relative efficiency

We select the pion tracks from our data and MC event samples [13] by applying the following set of cuts [14] on oppositely charged pairs of tracks which successfully formed common vertices (using the iterative vertexing code which minimizes the chi-square statistic of tracks' positions and momenta; see infra. Chapter 6.2):

---

[13] We used all off-peak data from Runs 1-4 and 25,000,000 SP5 MC udsc events with the light quark uds events constituting 61.7% of this total and the rest being charmed events (to approximate the [2.09 nb/1.30 nb] ratio of uds to charm events at BaBar energy.)

[14] the 2.4 cm cut in the XY-plane of the detector addresses the second of the TTG's tracking correction problems in relation to the present analysis and allows us to subsequently accept a maximal number of track-corrected beampipe Λ's.

- The common tracks' vertex is more than 4 mm from the BaBar origin in the XY-plane of the detector.
- The common track vertex is less than 2.4 cm in the XY-plane from the BaBar origin (this selects k_{short}'s which decayed inside the beampipe.
- The tracks' three-dimensional DOCAs to the common vertex are less than 0.5 mm.
- The tracks' DOCAs to the event primary vertex is at least 0.5 mm.
- The tracks' refitted (using the new common decay vertex) transverse momentum is greater than 100 MeV.
- The newly vertexed candidate object has an invariant mass within 5 MeV from the PDG k_{short} mass of 497.648 MeV.

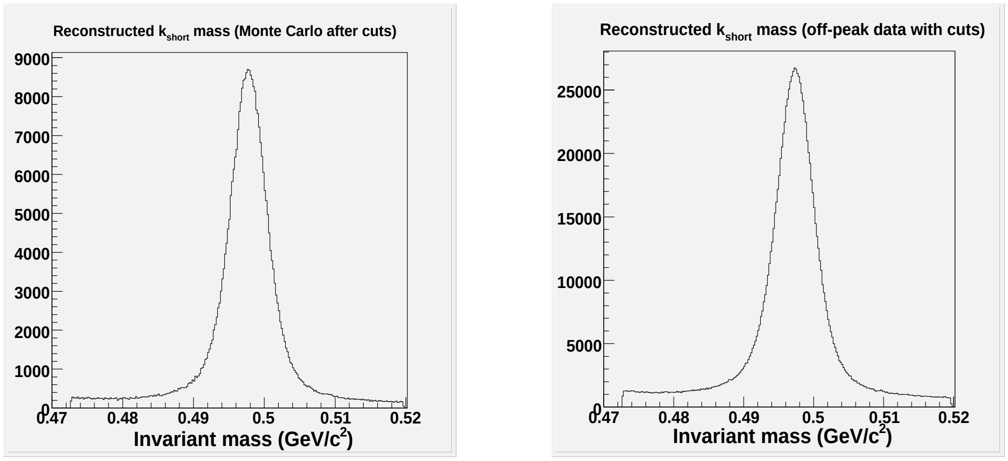


Figure 5.2: (Left plot) Number of successfully vertexed oppositely-charged MC tracks versus invariant mass of the vertexed candidate object (after cuts.) (Right plot) Analogous plot for Runs 1-4 off-peak data.

Figure 5.2 plots the invariant mass of the MC and off-peak data k_{short} candidates versus the total number of candidates after all of the above-referenced cuts *except* the mass cut.

It is evident from the figures that there remains some background under the k_{short} mass peak; however, since pion is the most prevalent species in the detector, we expect that much of the background consists of pion tracks as well. In fact, according to the MC, the pion purity in our sample of selected k_{short} is better than 99%. See table 5.1. The off-peak data sample selection, assuming similar proportions of track species as in the MC k_{short} sample and considering the slightly increased combinatoric background in the invariant mass plot, then has an estimated pion purity of better than 97%.

With the above-referenced selection of k_{short} tracks, we then calculated the following corrected efficiency following the TTG's general method:

$$\epsilon = \frac{\#\text{ of } k_s\ \textit{pion}\text{ trks with} \geq 8\text{ SVT hits}}{(\#\text{ of } k_s\ \textit{pion}\text{ trks with} \geq 8\text{ SVT hits} + \geq 12\text{ DCH hits}) - (R_g \cdot N_{evt})} \tag{5.4}$$

True track species of MC k_{short} candidate	Without k_{short} mass cut	With k_{short} mass cut
π from k_{short}	92.1%	98.2%
other π	4.3%	1.0%
p	2.2%	0.5%
k	0.2%	$> 0.1\%$
lepton	1.1%	$\sim 0.2\%$

Table 5.1: True species of Monte Carlo tracks from k_{short} candidates.

	Bhabha MC Events	Runs 1-3 off-peak Data Events
# Bhabha events	130,286	1,121,030
est. ghost tracks	1,322	22,140
R_g	1.0%	2.0%

Table 5.2: Estimated "ghost" tracks in Monte Carlo and data events selected using our 'Bhabha' criteria.

where R_g once again represents a ratio of ghost tracks estimated using Bhabha events and N_{evt} is the number of events used in our tracking analysis. Note that estimates of R_g is obtained by using the following selection for Bhabha events:

- Second Fox-Wolfram moment of the event (calculated from neutral plus charged energies) is ≥ 0.98 OR *in the alternative* the event sphericity calculated using charged and neutral energies is ≤ 0.004.
- The total number of charged tracks in the event is between 2 and 5 inclusive.
- Two of the charged tracks in the event are matched to two distinct continuous EMC clusters.
- The total event energy is within 200 MeV of the nominal total lab beam energy of 12.1 GeV.

Any additional unmatched tracks in the event then is considered a ghost track. See p.6 in [33]. We ran over Runs 1-3 off-peak dataset and 25,000,000 udsc MC events. Table 5.2 lists the various quantities associated with the selected MC (generated Bhabha events) and selected data Bhabha events.

We then calculate the data and MC DCH relative efficiencies using 4 LabAngle bins, 20 tranverse momentum bins, and 5 event charged track multiplicity bins, for a total of 400 bins. The π^+ and π^- track efficiencies were calculated separately. Tables 5.3 and 5.4 list the transverse momentum and lab angle bin intervals that we employed in the π DCH tracking efficiency studies.

In addition, the five event multiplicity intervals are defined as follows: 3-5, 6-8, 9-11, 12-14, ≥ 15.

P_T bin definitions (in GeV/c)		
Low Edge	High Edge	Bin Center
0.00	0.10	0.05
0.10	0.20	0.15
0.20	0.30	0.25
0.30	0.40	0.35
0.40	0.50	0.45
0.50	0.60	0.55
0.60	0.70	0.65
0.70	0.80	0.75
0.80	0.90	0.85
0.90	1.00	0.95
1.00	1.10	1.05
1.10	1.20	1.15
1.20	1.40	1.30
1.40	1.60	1.50
1.60	2.00	1.80
2.00	2.40	2.20
2.40	3.00	2.70
3.00	4.00	3.50
4.00	5.00	4.50
5.00	10.00	7.50

Table 5.3: Definitions of tracks' transverse momentum bins for studies of π's DCH relative efficiency.

Cosine(LabAngle) bin definitions		
Low Edge	High Edge	Bin Center
-0.7441	0.0408	-0.3517
0.0408	0.4300	0.2354
0.4300	0.7350	0.5825
0.7350	0.9171	0.8261

Table 5.4: Definitions of tracks' cosine(LabAngle) bins for studies of π's DCH relative efficiency.

Figure 5.3 contains plots of the averaged MC and data DCH efficiencies of π^+ and π^- in bins of transverse momentum; the efficiencies plotted here have been averaged over the event multiplicity and track lab angle bins. For comparison purposes, the averaged DCH efficiencies of MC and data (in bins of P_T) obtained by the TTG in [34] have been superimposed on the relevant plots. Recall that we have imposed a track transverse momentum cut at 100 MeV/c, so the first bins on all these plots are invariably zero. It is evident from these averaged efficiencies that in both the data and MC, the DCH efficiency values that we obtained in the current studies for pions are higher than the TTG's values which have significant contaminations from proton, kaon, and lepton species. In addition, the differences between MC and data is small in the sense that, except for the transverse momentum bins less than 300 MeV/c, the (data-mc) difference is generally on the order of half a percent or less with the MC generally slightly more optimistic than the data. However, significant differences exist between MC and data that rise to almost 4% for the 0.1-0.2 GeV/c bin. An analogous situtaion exists in the charge asymmetry differences between the π^+ and π^- efficiencies though this behavior is apparently well simulated by the Monte Carlo sample.

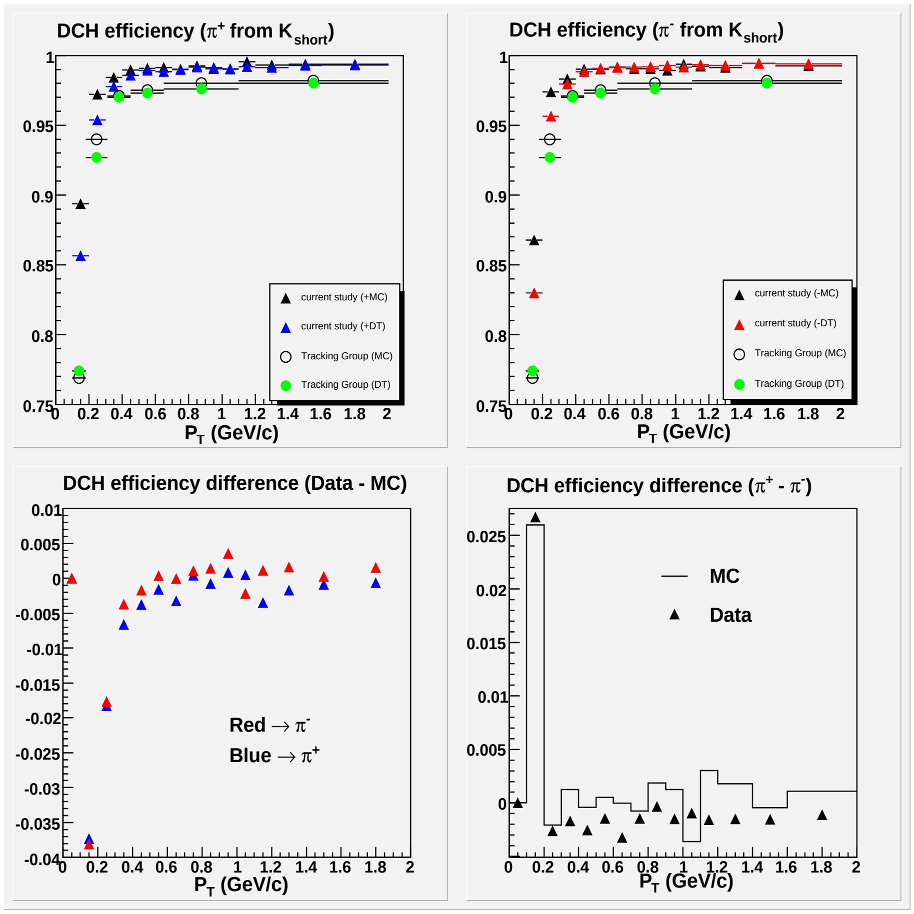


Figure 5.3: (Top left) Averaged π^+ DCH relative efficiency in transverse momentum bins. (Top right) Averaged π^- DCH relative efficiency in transverse momentum bins. (Bottom left) Differences between data and MC efficiencies obtained by the current study ($\epsilon_{data} - \epsilon_{mc}$.) (Bottom right) Charge asymmetries in the MC and data DCH efficiencies.

Figures 5.4 and 5.5 again contain the averaged DCH relative efficiencies for MC and data; however, in these plots the relevant display variables are track lab angle and event charged multiplicity respectively.

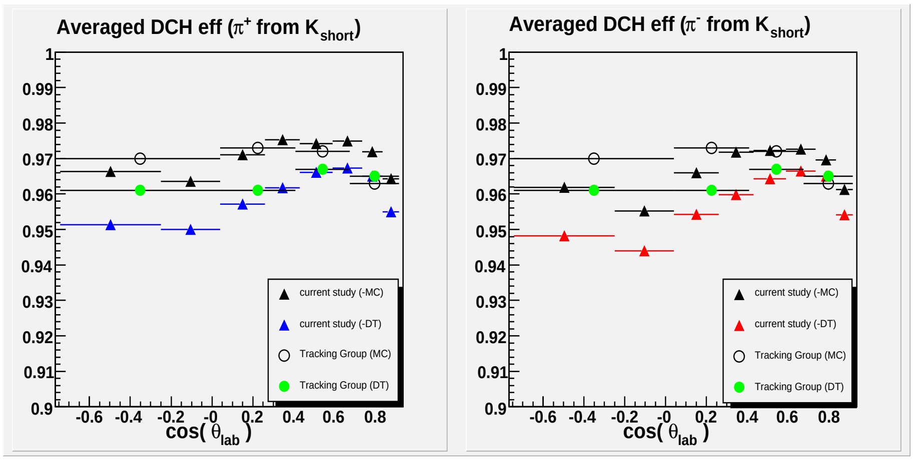


Figure 5.4: Averaged pion DCH relative efficiency in lab angle bins.

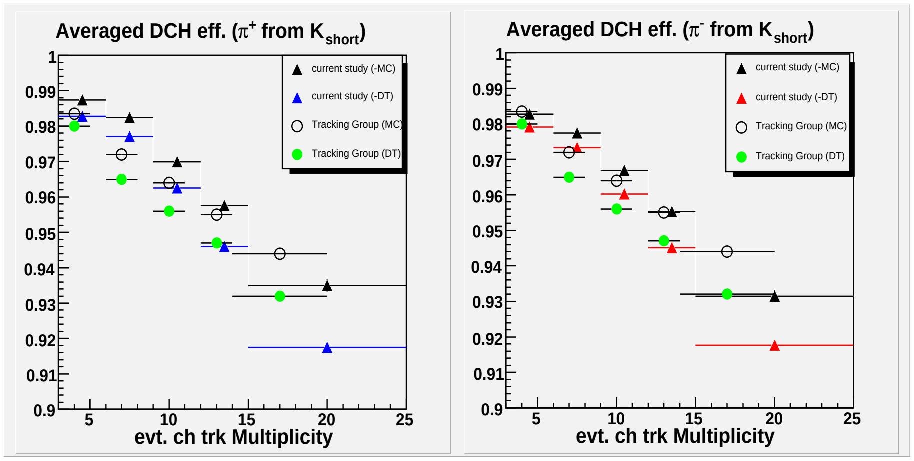


Figure 5.5: Averaged pion DCH relative efficiency in multiplicity bins.

The statistical uncertainty for each of the 400 DCH relative efficiency bins is assigned by assuming that the probability that a track which satisfies the requisite conditions for inclusion in the denominator of the efficiency ratio (e.g. a track with 8 SVT hits) will also satisfy the condition for inclusion in the numerator of that ratio (e.g. 8 SVT hits + 12 DCH hits) is given by the binomial probability distribution (see. pp. 26-28 in [35]); thus, the assigned bin statistical error is given by $\sqrt{N\epsilon(1-\epsilon)}$ where N is the number of tracks that contributes to the denominator of the bin efficiency and where ϵ is the estimated DCH relative bin efficiency.

The pion track correction for any relevant bin then is obtained by taking the ratio $\epsilon_{data}/\epsilon_{mc}$ for that bin and the assigned bin error is just the propagated error of the MC and data bin binomial errors.

Figure 5.6 displays some examples of the pion correction factors (and their associated binomial errors) that we obtained via the above-described procedure; note that all correction factors in these figures are for cosine lab angle bin values between 0.735 and 0.9171 and event multiplicity bin with 6-8 charged tracks.

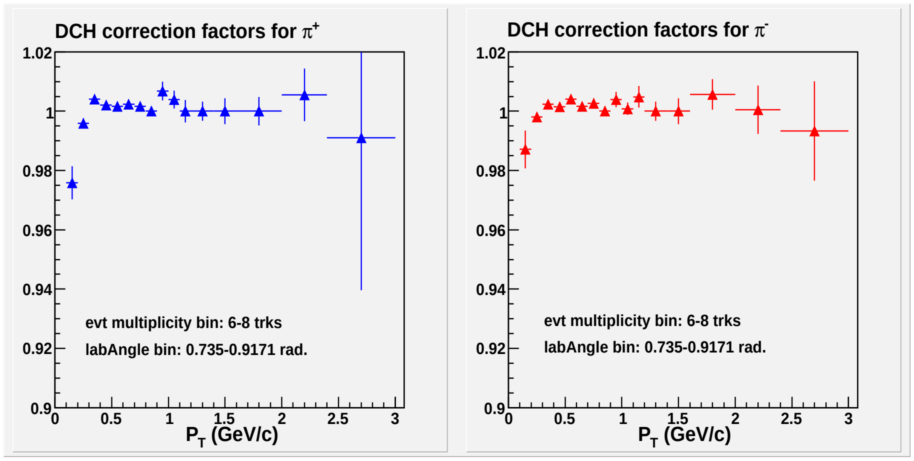


Figure 5.6: Examples of $\pi^{\pm}$ data/MC correction factors and associated errors.

5.3.2 SVT relative efficiency

We also make use of the SVT's and DCH's independent abilities to reconstruct tracks to examine the pion relative efficiency of the SVT. By necessity, this procedure examines the SVT's efficiency to reconstruct the subset of tracks which were already reconstructed in the DCH.

This procedure is completely analogous to the method by which we examined DCH's pion relative efficiency. We use the same track selection criteria detailed in the previous section with respect to the DCH efficiency *except* the SVT and DCH hit requirements. The relative SVT efficiency then is defined by the following ratio in each of the relevant bin:

$$\epsilon_{SVT} = \frac{\text{\# of selected trks with } \geq 8 \text{ SVT hits } \textit{and } \geq 12 \text{ DCH hist}}{\text{\# of selected trks with } \geq 12 \text{ DCH hits}} \tag{5.5}$$

Following are the averaged SVT pion relative efficiency displayed in the transverse momentum (5.7,) track cosine lab angle (5.8,) and event charged multiplicity variables (5.9.) The bin erros (when visible) again are the binomial errors calculated in the same manner as detailed for the case of the DCH relative efficiency.

For purposes of correcting for the 'beampipe' pions' SVT relative efficiencies, we will use our average MC and data efficiency values in the event multiplicity variable. Figure 5.10 contains a graphical presentation of our SVT correction factors for $\pi^{\pm}$; the error bars therein are the propagated binomial errors for the data and MC relative efficiencies.

5.3.3 MC/Data absolute tracking efficiency comparison using τ 3+1 events

Since the aforementioned DCH and SVT efficiency correction methods address only the efficiencies of DCH and SVT (and the efficiencies of their attendant software) *relative* to the other stand-alone detector, it is desirable that we also have a method to investigate the absolute charged track efficiency of the BaBar detector. If we have in our hand such a method,

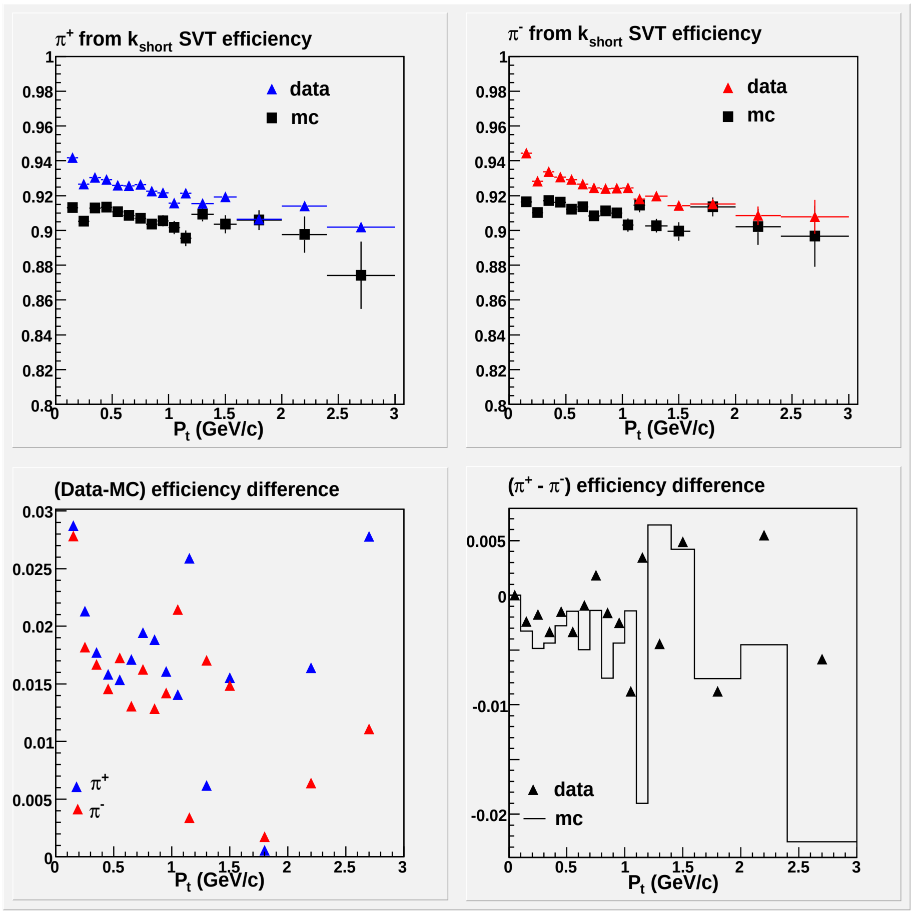


Figure 5.7: Averaged pion SVT relative efficiency in transverse momentum bins.

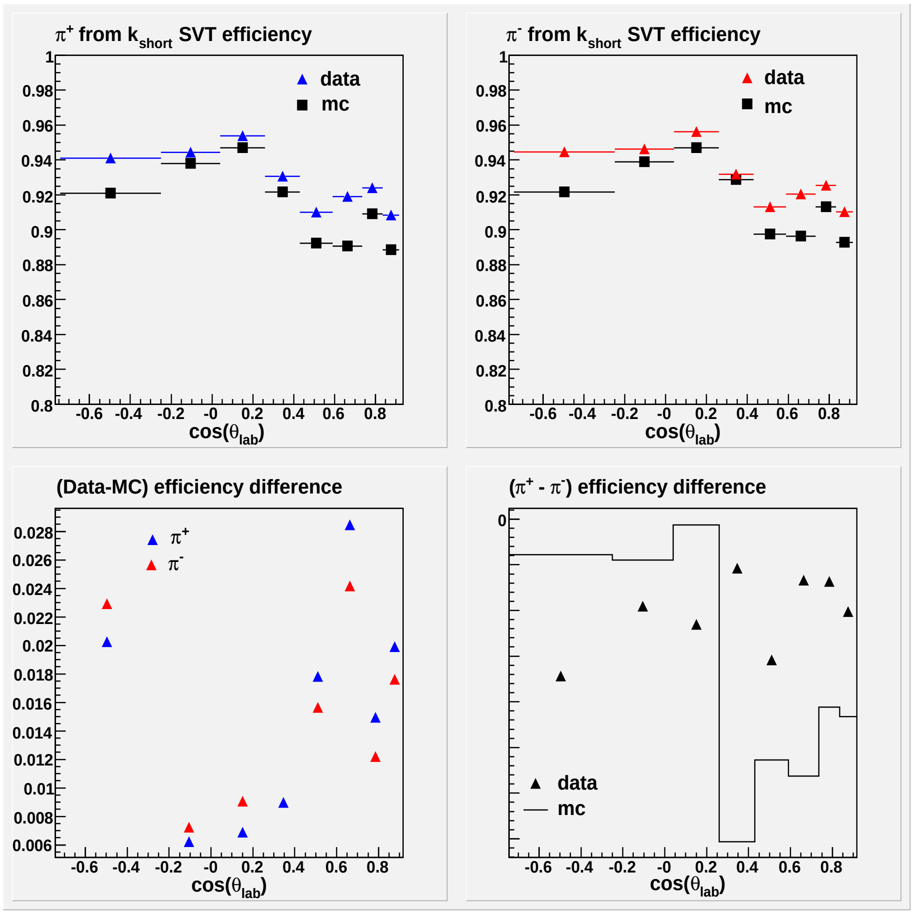


Figure 5.8: Averaged pion SVT relative efficiency in cosine lab angle bins.

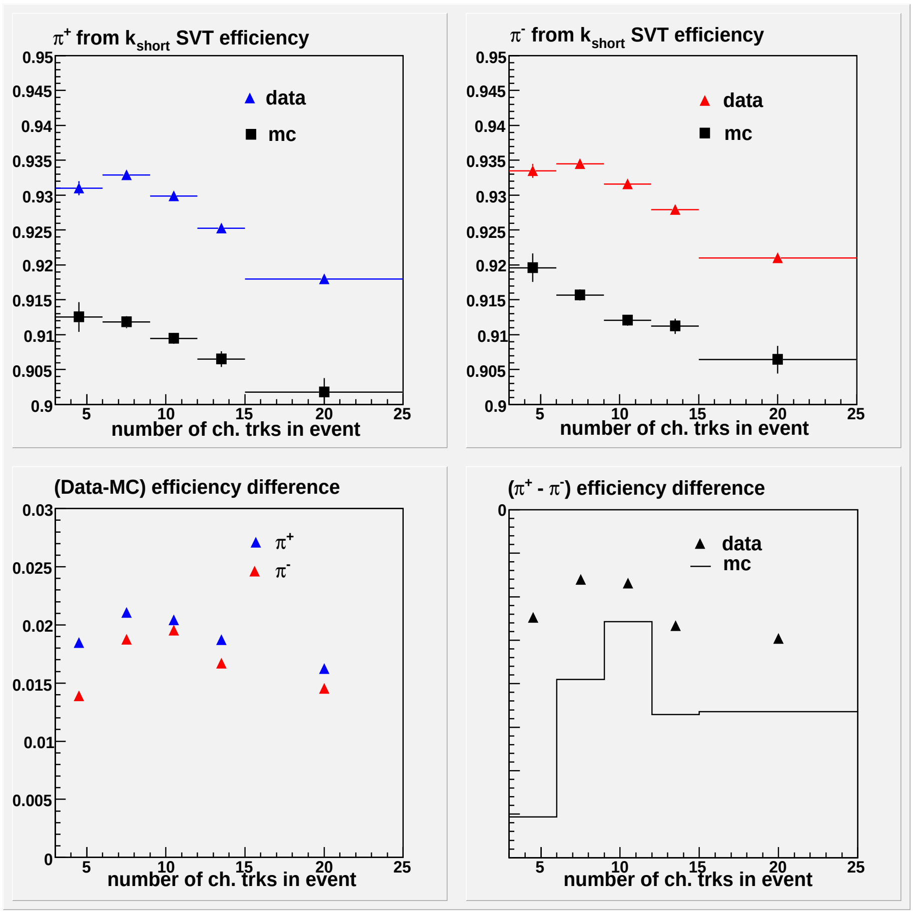


Figure 5.9: Averaged pion SVT relative efficiency in multiplicity bins.

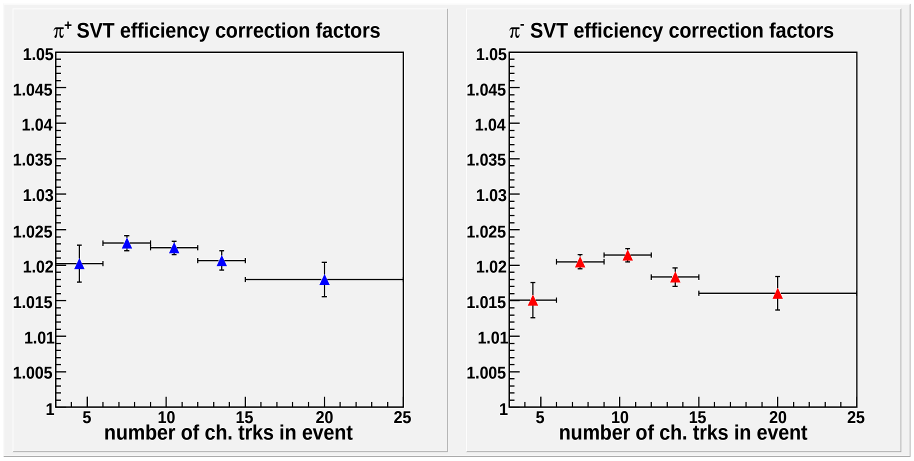


Figure 5.10: Pion SVT efficiency correction factors given in multiplicity bins.

then we can determine whether our combined data/MC DCH+SVT correction factors have failed to account for some of the differences between MC and data charged tracks.

The TTG has proposed to use the so-called τ 3+1 type events to investigate the absolute efficiency differences between data and the Monte Carlo. This type of events involve the following process $e^+e^- \to \tau^+\tau^-$ wherein one of the tau leptons further undergoes $\tau^\pm \to 3$ hadrons $+\ \nu_\tau$. The utility of these events for purposes of the investigation of absolute detector efficiency differences between MC and data lies in their well-defined final state characteristics; the τ decays are constrained to have odd multiplicity final states due to charge conservation.

Both the TTG and the current study seek to isolate samples of τ 3+1 events by identifying events with a lepton and at least two pion tracks. The event selection criteria that we used in the current study are as follows:

- Number of charged tracks in the event must be between three and five inclusive.
- Number of accepted k_{short} candidates accepted in the event must be zero. (We use the same k_{short} criteria as before; namely candidates produced by vertexing oppositely charged tracks with invariant mass less than 5 MeV from the PDG mass for k_{short}.)
- Number of accepted photon conversion candidates must be zero. (We count the number of vertexed oppositely charged tracks with invariant mass less than 50 MeV as an *accepted* photon conversion candidate.)
- Total neutral energy in the event must be less than 2.5 GeV.
- All tracks in the event must satisfy BaBar's GoodTrackLoose definition.
- Preciesely one track in the event must be identified as loosely consistent with either an electron or muon by BaBar's standard likelihood-based electron (PidLHElectronSelector) and muon (PidMuonMicroSelector) particle identification routines.
- The identified lepton track's angle must be more than 120 degrees from any other tracks in the event.
- The lepton track's momentum must be between 20% and 80% of total beam energy.
- At least two of the tracks in the event are consistent with pion hypothesis using BaBar's global likelihood particle identification selector (they must also be inconsistent with electron and muon identities using the aforementioned electron and muon selectors.)
- The total charge in the event must be between +1 and -1 inclusive.
- The total energy of the three idenfitifed tracks must be $< 80\%$ of $\sqrt{s}$.
- The missing transverse momentum of the event must be > 300 MeV and the cosine lab angle of the missing momentum must be < 0.80.

The absolute efficiency multiplied by acceptance of the detector then is given by equation 5.6:

$$\epsilon \times A \ = \ \frac{N_4}{N_3 + N_4 + N_5} \tag{5.6}$$

Event Type	$\epsilon \times A$	# of Accepted Events		
		N_3	N_4	N_5
τ MC	83.3%	45	408	37
off-peak data	86.1%	345	4,848	437

Table 5.5: Number of accepted data and MC events after application of τ 3+1 selection criteria.

where N_x is number of events with x number of charged tracks. Thus N_3 is the number of selected events where no additional tracks other than the lepton and two pion tracks have been found, N_4 is the number of selected events with one additional track beyond the identified tracks, etc. Note that N_5 is present in the denominator of the efficiency times acceptance quantity in order to account for the presence of ghost tracks and loopers (slow tracks that form multiple loops in the detector due to the magnetic field.) We did not segregate our selected event samples into electron and muon events, nor did we segregate out events wherein the two pions come from the $\rho(770)$ resonance decay.

We ran our τ 3+1 selection routine over 1,000,000 SP5 $\tau^+\tau^-$ MC events as well as over our entire off-peak dataset for Runs 1 through 4. Figure 5.11 displays the total energy of selected MC and data events. Table 5.5 gives the MC and data results for absolute track efficiency multiplied by acceptance in the selected events.

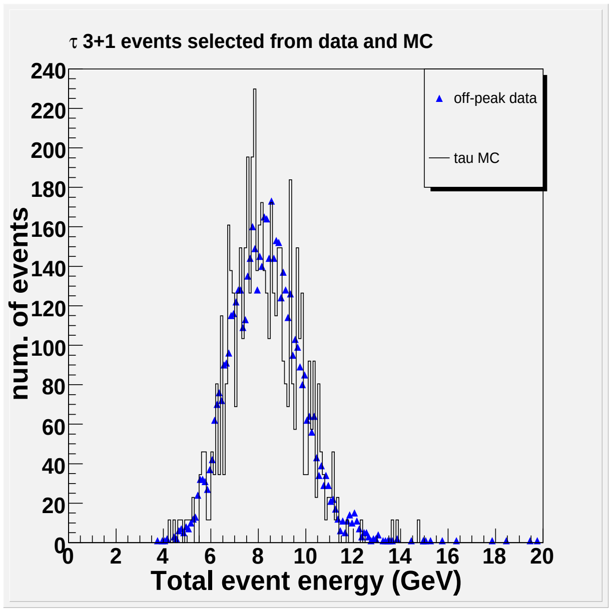


Figure 5.11: Total event energy of off-peak data and MC events selected for τ 3+1 characteristics. Note that the number of selected τ MC events have been normalized to the total area under the data histogram.

The TTG's τ 3+1 study using SP5 monte carlo events was more detailed than the current study in that it segregated the various τ decay channels and considered additional factors such as Bhabha event contamination and DCH high voltage effects; in addition the TTG study had greater τ statistics. The TTG found that the absolute efficiency times acceptance for tracks satisfying GoodTrackLoose criteria are generally between 85% and 89% with MC generally more optimistic than data [34] (this latter fact tends to agree with TTG's SVT/DCH-based relative efficiencies.)

In our τ study, the MC sample tends to be less optimistic than data with regard to the efficiency of GoodTrackLoose; although this differs from the TTG results, it does agree with our own DCH/SVT based efficiencies for events with charged mutliplicity between 3 and 5 tracks inclusive; this difference is likely due to our limitation to off-peak data events. In the 3-5 charged multiplicity region, our combined SVT/DCH relative efficiency from data is about 91.4% while that from MC is about 90.0%. This difference contrasts with the data/MC difference of about 2.8% using our τ selections. Although the statistics from the τ events are relatively poor for both MC and data in our study, we will nevertheless employ this information and assign an 1.4% flat systematic error to account for possible additional differences between our MC and data track efficiencies.

5.4 Method for correcting $p/\overline{p}$ MC/Data tracking differences

In order to correct for efficiency differences between MC and data for protons in our analysis without using dE/dx based particle identification, we need to isolate a sample of protons that resemble those that come from Λ decays. The simplest thing to do then is to perform the efficiency study using tracks that vertex to candidates with invariant mass near the Λ mass. In short, we will seek to use Λ candidates to self-correct the relative tracking efficiencies of their protonic daughters.

The Λ candidates are initially composed by vertexing (with χ^2 minimization code which will be briefly described in Chapter 6.2) tracks with opposite charge assignments using proton and pion track hypotheses and then taking care not to vertex a track with itself. Then, for purposes of obtaining a relatively pure sample of protons, we apply the following set of selection cuts to our Λ sample:

- The unfitted [15] three dimensional DOCA of the faster track (relative to the other track) to the candidate decay vertex must be less than 0.1 mm.
- The unfitted three dimensional DOCA of the slower track to the candidate decay vertex must be less than 0.75 mm.
- The composite candidate's decay vertex must be separated from the event primary vertex by at least 3 mm in BaBar's XY-plane.
- The composite candidate's decay vertex must be less than 2.40 cm in BaBar's XY-plane. *This constrains the candidate to be inside the beampipe.*
- The composite candidate's invariant mass is more than 10 MeV from PDG mass for k_{short} when both daughter tracks used the pion mass hypothesis.
- The composite candidate has total momentum greater than 350 MeV. *This ensures that the faster track of the Λ is always the proton; see p.36 in [30].*

---

[15] Once two tracks have successfully vertexed to a composite candidate, each of the tracks are refitted using the composite decay vertex as the tracks' origin. Here, we want the tracks' DOCA information prior to the refitting step.

- Both of the two daughter tracks have total momentum gerater than 100 MeV.
- The invariant mass of the composite candidate is within 3.0 MeV of the Λ mass.

After the aforementioned selection cuts, the resulting Λ candidates still retain substantial background in both data and MC as evidenced in 5.12. We employ a mass sideband subtraction method to account for the background contamination in our Λ-derived proton sample.

We use the same selection criteria as before, but now instead of the mass window of 6.0 MeV centered on the Λ's mass, we define two regions at $1.100 - 1.106$ GeV and $1.126 - 1.132$ GeV. We then count the number of proton candidates in some particular P_T-LabAngle-Multiplicity bin accepted from within the Λ mass window; if there is an accepted proton candidate in the same bin from one of the sideband regions, we would *subtract* the numbers in that bin by a factor of 0.5 (the 0.5 factor accounts for the presence of two sideband regions with relatively linear background shapes and approximately equidistant from the Λ mass.) Figure 5.13 displays the effect of the sideband subtraction in the proton P_T variable for both MC and data; note that in the case of MC (on the left plot in 5.13) we simultaneously displayed the histogram for the P_T of true protons from the Monte Carlo Lambda decays.

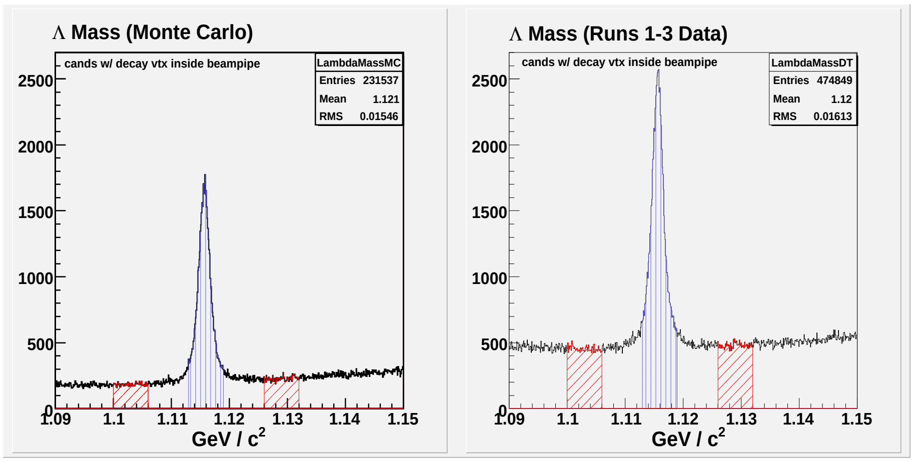


Figure 5.12: Signal and sidebands in Λ's invariant mass for purposes of subtracting the combinatoric background.

The actual DCH relative efficiencies for the protons are derived in precisely the same manner as that for the pions from k_{short}'s:

$$\epsilon = \frac{\#\text{ of } \Lambda\, proton \text{ trks with} \geq 8 \text{ SVT hits}}{(\#\text{ of } \Lambda\, proton \text{ trks with} \geq 8 \text{ SVT hits} + \geq 12 \text{ DCH hits}) - (R_g \cdot N_{evt})} \tag{5.7}$$

The proton relative efficiencies use the same binnings that we employed for k_{short} pions. Figures 5.14, 5.15, and 5.16 display the averaged proton relative efficiencies in the transverse momentum, cosine lab angle, and event multiplicity variables. Note that in the P_T variable, the data/MC differences in relative efficiencies are almost always greater for antiprotons than for protons; this is an expected behavior in that the antiproton cross section with matter is generally less well known than the proton-matter cross section.

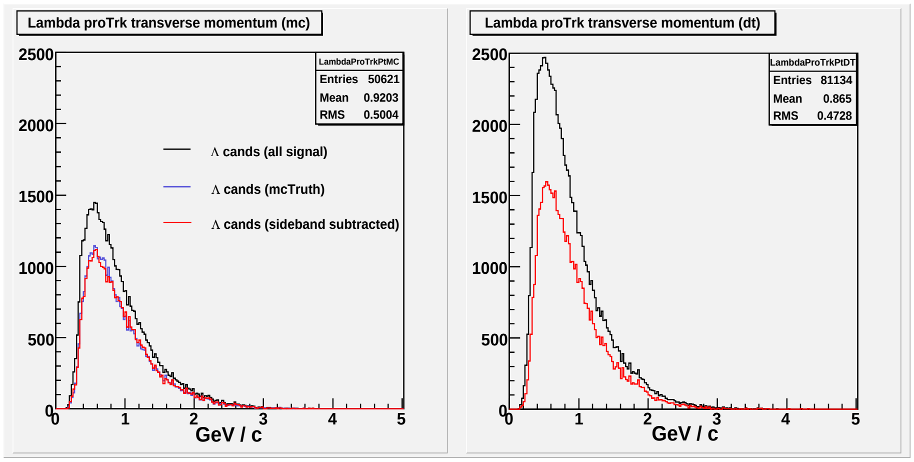


Figure 5.13: Transverse momentum of protons from Λ decays (MC, data, and sideband-subtracted data)

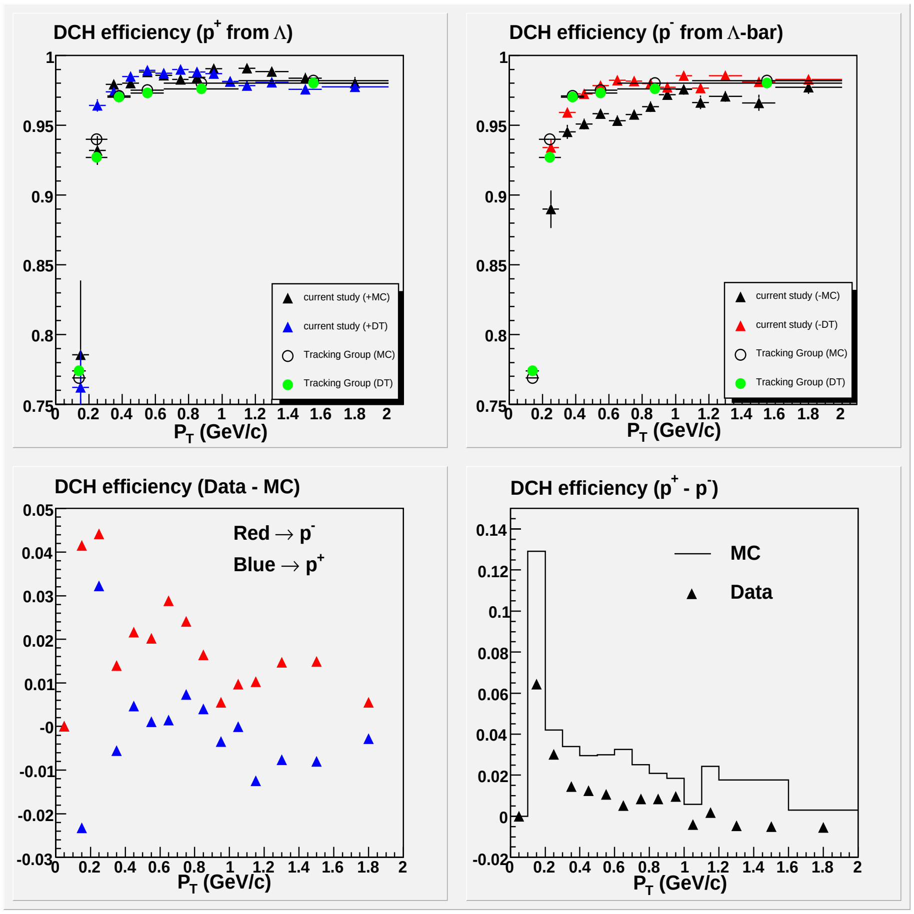


Figure 5.14: Averaged proton DCH relative efficiency in transverse momentum bins.

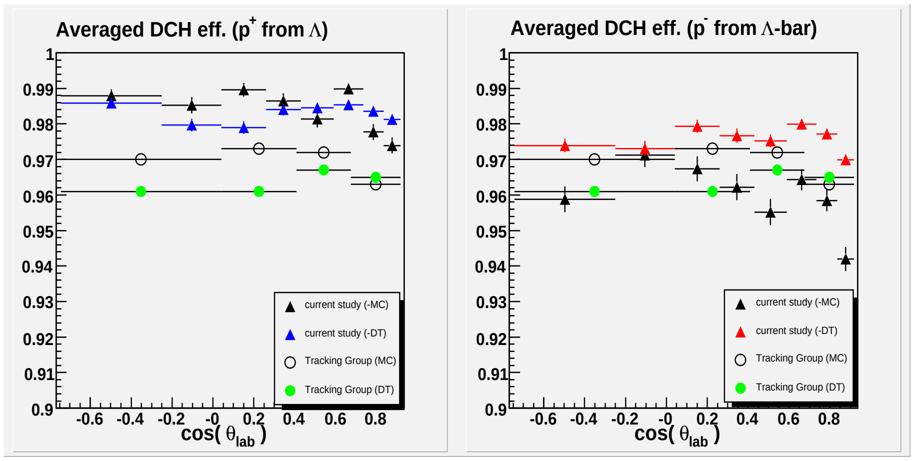


Figure 5.15: Averaged proton DCH relative efficiency in lab angle bins.

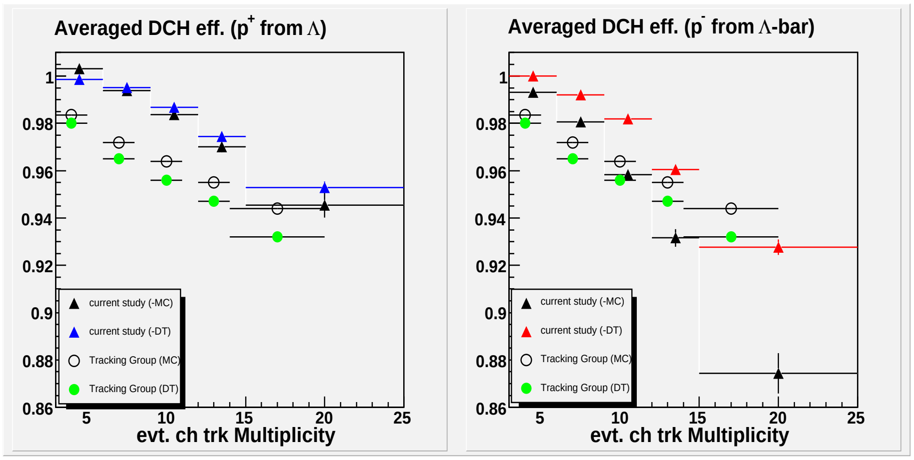


Figure 5.16: Averaged proton DCH relative efficiency in multiplicity bins.

The systematic errors for the proton efficiencies and correction factors are calculated using binomial errors just as in the case for π from k_{short}. Finally, studies show no corrections are needed for proton SVT relative efficiencies.

Chapter 6

Hyperon selection and signal extraction

We now describe the procedures by which we select and extract signals for the Λ^0, Ξ^-, and Ω^- hyperons. We then proceed to describe the efficiency correction and signal fitting methods for each of these particle states.

6.1 Overview

Initially, we note that one of the common ways to make high energy measurements involves the following process: we start with some theoretical expression that connects a quantity under consideration (e.g. the CKM matrix elements involved in CP-violation studies) with other more directly measureable signal quantities, we then construct significance-level based selectors using the probability distribution functions ("PDFs") of the signals and 'backgrounds' in the measured quantities in order to maximize the signal-to-noise ratio (this approach is consistent with both the Bayesian and the frequentist statistical principles,) the accepted signals then produce a best-estimate of the desired quantity with an error margin determined in part by confidence level cuts.

Alas, the above procedure is impractical for our current analysis for three reasons. First, hadronization cannot be calculated by precise theoretical means at present; thus we have no fundamental theoretical expression relating the particles' spectra to other measurable quantities. Second, the PDF of particle momentum is precisely what we seek to measure here; we cannot use the PDFs' derived from phenomelogical Monte Carlo models to measure the particles' true PDF distributions themselves. Third, our present analysis is fundamentally a number counting exercise: we seek to admit a maximal number of signal candidates in a relatively unbiased manner and this precludes usage of devices such as the Fisher linear discriminant and artificial neural-network-based selectors (see e.g. pp.51-57 in [35].)

In view of this state of affairs, our programme for obtaining information on the production rates and momentum spectra of the Λ^0, Ξ^-, and Ω^- hyperons shall conform to the following more transparent general outline:

First, we compose, select, and cut upon Λ^0 hyperon candidates which have decayed via the $\Lambda^0 \rightarrow p^+\pi^-$ channel while they were still inside the PEP-II beampipe; such a selection ensures that we can use our tracking correction tables for protons and pions that originate from within the PEP-II beampipe (the production of these correction tables were described in the previous chapter.) The Lambda candidates are composed using pairs of oppositely charged tracks from the same event through the application of a χ^2 minimization based 'vertexer'. Several cuts also are used to reduce the background relative to the Lambda signal. Then each of the $\Lambda^0/\overline{\Lambda}^0$ candidate that survived all the selection cuts is efficiency-corrected using a two-dimensional efficiency matrix where the relevant matrix variables are the candidates' CM momenta (here "CM" refers to the PEP-II beams' center of mass reference frame) and the labAngles (i.e. cosine angles with respect to the BaBar z-axis in the lab frame) defined by the momenta. Particle and antiparticle states are corrected using separate efficiency matrices.

These efficiency matrices shall be binned in such a manner as to allow for efficiency corrections to proceed with a minimum of bias due to the PEP-II beams' lab frame boost and yet permit the matrices to retain sufficient bin statistics to generate acceptably small MC statistical errors.

The efficiency-corrected number of $\Lambda^0/\overline{\Lambda}^0$ candidates then are plotted with respect to their invariant mass with the proviso that these invariant mass plots are further binned in terms of the candidates' CM momenta. After this step, we proceed to fit each of the momentum-binned mass signals and apply additional corrections (e.g. corrections for the particle's decay distance cut.) Next, we convert the number of fitted $\Lambda^0/\overline{\Lambda}^0$ in each CM momentum bin into a fractional hadronic cross section and thereby produce the $\Lambda^0/\overline{\Lambda}^0$ hyperon's momentum spectrum in units of fractional hadronic cross section. Finally, the individual cross sections in the various CM momentum bins are summed and then divided by the estimated total number of multihadronic events to produce an estimated production rate for the Λ^0.

All composed Λ^0 candidates, including background combinatorics [16], that fall within a loose invariant mass window between 1.10 and 1.13 GeV/c^2 and which have one and only one decay daughter that is consistent with a proton particle identification ("PID") hypothesis [17] are then used to futher compose Ξ^- and Ω^- candidates.

The Ξ^- candidates are produced by using the χ^2 minimization vertexer once again to compose the accepted Λ^0 candidates with all appropriately charged tracks (the Λ^0 candidates' masses are set to the true PDG Λ^0 mass while the tracks are given the pion mass hypothesis.) Where possible (that is, for Ξ^- candidates with momenta less than 3.0 GeV/c,) we select a Ξ^- candidate only if its Λ^0 decay daughter also had decayed within the PEP-II beampipe; however, for Ξ^- candidates that reside in momentum regions where the statistics for beampipe-Λ^0 decay daughters is too low or non-existent, we also select those Ξ^- candidates where the Λ^0 daughters have decayed within the inner SVT (i.e. $<$ 10 cm in BaBar's transverse plane.)

The Ω^- candidates are composed using a similar procedure as that used for the Ξ^- candidates; the main difference between the two lies in the fact that the extra charged track that is used (along with the Λ^0) to compose the higher order decay is imparted with the kaon mass hypothesis rather than the pion mass hypothesis. A further difference between the Ω^- and Ξ^- analyses lies in their differing selection methods; due to the low rate and efficiency

---

[16] False combinations of random tracks with invariant mass near the true particle mass

[17] We employ a PID selector developed by BaBar's Inclusive Hadronic Analysis Group that makes use of Log-Likelihood-Differences. This selector uses the dE/dx information from the DCH as well as the fitted Cherenkov angle information from the DIRC to calculate the log-likelihood differences between different species hypotheses; since the $p-K$ and $\pi-K$ separation sensitivity of the DCH decreases rapidly above 0.5 GeV/c and 0.8 GeV/c respectively, and the DIRC PID software is inefficient near Cherenkov threshold (1.5 GeV/c for protons and 1.0 GeV/c for kaons,) a momentum-dependent DCH vs. DIRC log-likelihood-difference linear cut is employed by this selector to maximize the overall PID efficiency. See pp.18-41 in [30].

of the Ω^-, there are very few detected Ω^- candidates where the Λ^0 daughters also decayed within the beampipe; hence, we accept Ω^-'s where their Λ^0 decay daughters decayed within the inner SVT throughout the entire Ω^- momentum range.

The production rates and the momentum spectra of both the Ξ^- hyperon and the Ω^- hyperon are generated in an entirely analgous manner as those for the Λ^0 hyperon.

6.2 Λ^0 signal extraction

We compose Λ^0 candidates by applying a χ^2 minimization "vertexer" software routine on two oppositely charged tracks (one track with proton mass hypothesis and the other track with pion mass hypothesis) and thereby creating a composite neutral candidate.

The χ^2 vertexer is a BaBar software routine that initially assumes the decay vertex of the purported neutral candidate to be the mid-point in the line that joins the points of closest approach between the two tracks under consideration. During the vertexing process, the underlying charged tracks' position and momenta are given in the so-called X-P representation (an alternative track representation used during tracking reconstruction is the so-called helix representation.) The software then proceeds to estimate the new vertex and momentum of the composite neutral candidate as constraints on the individual charged tracks' position and momenta. The χ^2 vertex estimation process is based on a generalized linear least square method using the Lagrange Undetermined Multiplier technique.

For the usual χ^2 minimization procedure where we seek to estimate N quantities λ_i with $i = 1, 2, ..., N$ which can be written as linear functions of θ_j such that $\lambda_i = \sum_{j=1}^{m} A_{ij}\theta_j$, the χ^2 statistic can be defined in matrix notation as follows:

$$\chi^2 = (\vec{y} - \vec{\lambda})^T \, \tilde{V}^{-1} \, (\vec{y} - \vec{\lambda}) \tag{6.1}$$

$$\chi^2 = (\vec{y} - \tilde{A}\vec{\theta})^T \, \tilde{V}^{-1} \, (\vec{y} - \tilde{A}\vec{\theta}) \tag{6.2}$$

where the N quantities λ_i have been written as a vector $\vec{\lambda}$, $\vec{y}$ is a vector of the measured values associated with $\vec{\lambda}$, $\tilde{V}$ is the matrix of the covariance errors, and where $\tilde{A}$ is the matrix of the linear coefficients that link λ_i with the parameters θ_j. See pp.95-98 in [35].

The probability distribution function ("PDF") for obtaining N measurements of y_i (assuming normal distribution) is the joint PDF of N gaussians and the log-likelihood function of this distribution is proportional to the sum of the N arguments of the N exponentials which is just the χ^2 function by definition. Minimization of the χ^2 statistic with respect to the θ parameters then maximizes the probability and allows us to solve for our least square best estimates for the θ parameter:

$$\hat{\theta} = \left(\tilde{A}^T \, \tilde{V}^{-1} \, \tilde{A}\right)^{-1} \, \tilde{A}^T \, \tilde{V}^{-1} \, \vec{y} \tag{6.3}$$

An example of the application of least square fitting using the χ^2 minimization method would be fitting the parameters of a reconstructed charged track. In this hypothetical case, the measured track hits y_i are functions of the unknown track parameters θ_i; given an initial guess for the track parameters and the estimated covariance matrix, we then obtain the least square best estimates $\hat{\theta}$ by minimizing the χ^2 statistic.

The vertexing of two charged tracks to form a neutral candidate on the other hand requires us to constrain the two individual charged tracks to come from a common vertex; we must also require that the four momenta of the tracks should be consistent with the four momentum of the composite object. The addition of such constraints can be incorporated into the χ^2 minimization method by use of Lagrange undetermined multipliers.

The BaBar vertexing software code assumes that the constraint functions can be made linear; in particular, if the unknown vertex and the composite candidates momentum are represented as the elements of a vector $\vec{\xi}$ and if we define $(\vec{y} - \vec{\lambda}) \equiv \vec{\delta}$ (where $\vec{\lambda}$ represents the parameters of the individual charged tracks , then a new χ^2 function with undetermined multipliers is given as follows:

$$\chi^2 = \vec{\delta}^T \, \tilde{V}^{-1} \, \vec{\delta} \; + \; 2 \, \vec{\mu}^T \left(\tilde{A}\vec{\xi} + \tilde{B}\vec{\delta} + \vec{c} \right) \tag{6.4}$$

where $\tilde{A}$ now represents the linear coefficient matrix that relates the composite vertex and composite momentum to the constraint function, $\tilde{B}$ represents the linear coefficient matrix that relates the $(\vec{y} - \vec{\lambda})$ quantity to the constraint function, μ is the vector of the Langrage undetermined multipliers, and $\vec{c}$ is the vector of the constraint values at the initial values of the track parameters and the first-guess value of the composite vertex and momenta [36]. Minimization of this χ^2 function then gives the estimated composite vertex and momentum vector as follows:

$$\vec{\xi} \; = \; - \left(\tilde{A}^T \tilde{G}_B \tilde{A} \right)^{-1} \tilde{A}^T \tilde{G}_B \vec{c} \tag{6.5}$$

where

$$\tilde{G}_B = \left(\tilde{B} \tilde{V}^{-1} \tilde{B}^T \right)^{-1} \tag{6.6}$$

The χ^2 minimization vertexing code is performed iteratively until the change in successive χ^2 values is less than 0.005 or until a maximum of six iterations [36].

The charged tracks are then refitted using the newly composed vertex as an additional constraint; this step is accomplished by the application of a software routine named "TreeFitter", which is a Least χ^2 minimization routine based on the Kalman filter [18].

We immediately limit the accepted neutral candidates to those with a reconstructed decay vertex located less than 2.4 cm in the XY-plane from the BaBar origin (this sample is the so-called "beampipe Λ^0 sample".) The number of off-peak data beampipe Λ^0 candidates as a function of the candidates' invariant mass is plotted in Figure 6.1. Though there is significant background, the Λ^0 mass peak near 1.115-1.116 GeV/c is clearly visible.

6.2.1 Cuts against Λ candidates

In order to reduce the combinatoric background under the Λ^0 mass peak, we institute the following additional cuts against our beampipe candidates sample:

- The distance between the candidate decay vertex and the event primary vertex must be > 2.0 mm in BaBar's XY-plane (we'll refer to this cut as the "dXY cut".) Since Λ^0 has a relatively long lifetime, this cut effectively eliminates a substantial portion of the combinatoric background coming from the event vertex.

---

[18] Kalman filter is a least χ^2 fitting method that estimates a new set of parameters based on an older set of parameter estimates and new/additional measurements. See e.g. [37]. (Note that the Kalman filter is also used in BaBar track fitting routines.)

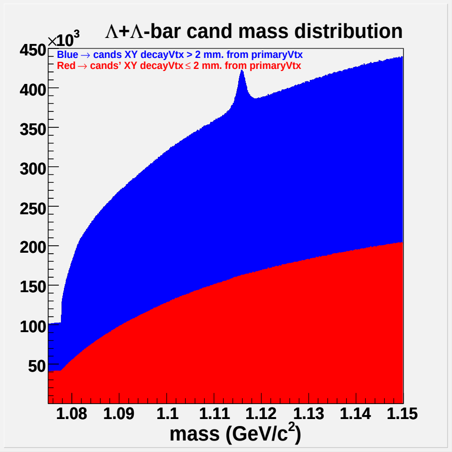


Figure 6.1: Number versus invariant mass of composite Λ^0 candidates in off-peak data formed from oppositely charged tracks with decay vertices inside the beampipe ("beampipe Λ^0 sample".) Those candidates which were rejected by the additional 2mm transverse decay vertex cut have been colored red in the plot while those accepted by the same cut are colored blue.

- The candidate's purported proton daughter must have transverse momentum greater than 100 MeV/c. This cut discriminates against pions masquerading as protons.

- The cosine angle between the candidate momentum vector and the ray connecting the event primary vertex with the candidate decay vertex (we'll refer to this angle as the "pointing angle") must be > 0.0. This constraint requires the reconstructed neutral candidate's momentum to be in the same hemisphere as the vector defined by the candidate's reconstructed decay vertex with respect to the primary event vertex.

- Both the purported proton and pion daughters of the neutral candidate must have at least 8 hits in the SVT and 12 hits in the DCH. This cut ensures that accepted candidates have daughters with well-measured momenta which can be efficiency-corrected by our data/MC tables developed in the previous chapter.

In addition to the above enumerated cuts, for those candidates with center of mass momenta greater than or equal to 3.0 GeV/c, we increase the dXY acceptance threshold from 2 mm to 5 mm. We also require that the daughter tracks' DOCAs to the reconstructed candidate's vertex must be less than 3 mm (the tracks' DOCAs are calculated using the original track parameters that existed prior to the vertexing step.)

There is a final complication, the efficiency for the very first CM momentum bin (0.00-0.25 GeV/c) is very low, while for the last bin, there are very few particles and antiparticles; for these bins, the signals are completely masked by background unless we apply PID. Thus we require that the protonic daughters of candidates in these bins must not be consistent with either the pion or kaon hypotheses (using the Log-Likelihood difference selector described briefly in footnote 17 herein and in [30].)

The effect of these individual cuts on off-peak data composite candidates can be seen in Figure 6.1,Figure 6.2, and in Figure 6.3; the combined result of these cuts are reflected in the right-hand plot in Figure 6.3.

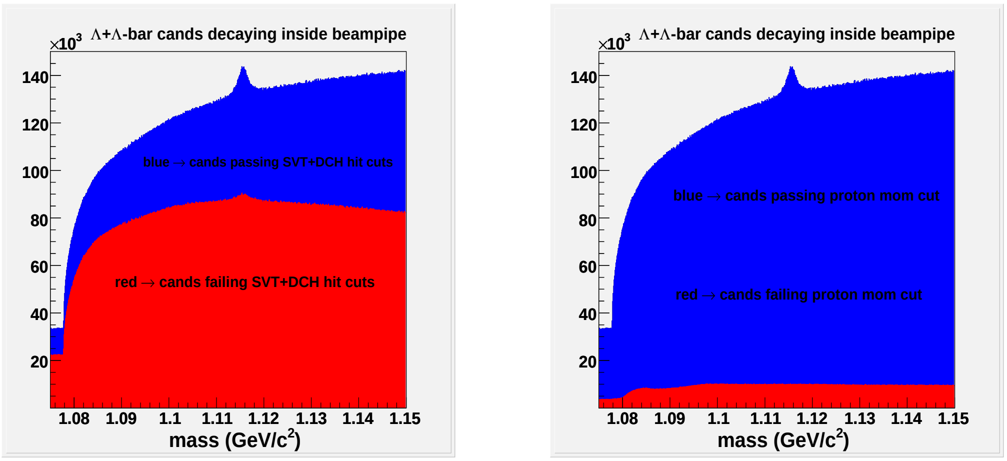


Figure 6.2: Effect of cuts with respect to daughter SVT/DCH hits and proton momentum on off-peak data Λ^0 candidate acceptance. Candidates rejected by the relevant cuts are colored red while those accepted by the same cuts are colored blue.

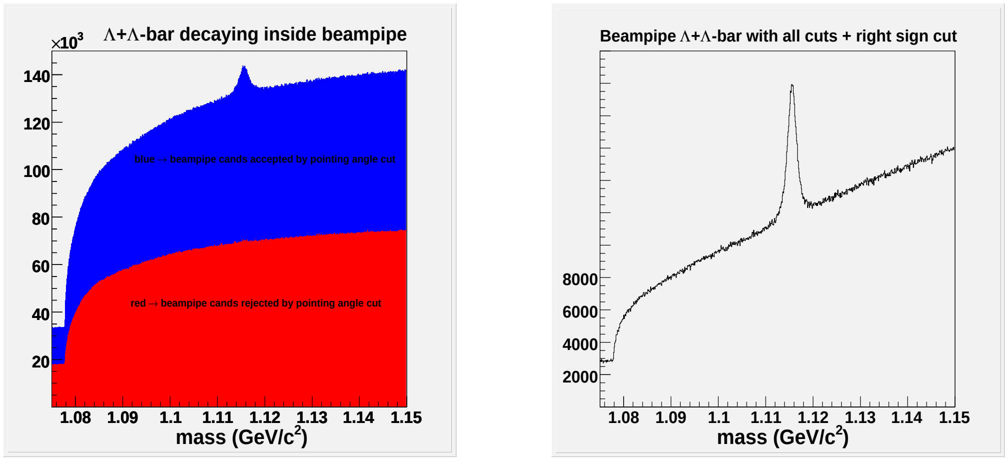


Figure 6.3: Effect of cut on candidate pointing angle and the combined effect of all enumerated cuts with respect to off-peak data Λ^0 candidate acceptance. Candidates rejected by this cut are colored red while those accepted by the cut are colored blue.

6.2.2 Λ efficiency matrix construction

We employ the same Λ^0 composition routine and cuts described in the previous section on our Monte Carlo event sample (this consists of 25 million MC events where pairs of 'u', 'd', 's', or 'c' quarks in proper relative ratios at BaBar energies are produced from the virtual photon, corresponding to an integrated luminosity of $\sim 7\, fb^{-1}$.) Accepted MC beampipe Λ^0 candidates with invariant mass between 1.108183 GeV/c^2 and 1.123183 GeV/c^2 (a relatively loose mass window with total width of 15 MeV/c^2) and which have been identified as true Λ^0s according to information from the MC generators [19] are then collected as a first step towards the construction of Λ^0 efficiency matrices.

The recontructed and accepted MC candidates which have been associated to true generator Λ^0 and $\overline{\Lambda}^0$ are divided into particles and antiparticles and stored in separate two-dimensional histograms. The 2-d histograms consist of 100 × 100 bins (for a total of 10,000 total bins) in the lab-frame momentum and labAngle [20] variables of the candidates. The 100 momentum bins are all uniform width bins that span candidate lab momentum between 0.0 GeV/c and 8.0 GeV/c (thus each momentum bin width is 0.08 GeV/c) while the 100 labAngle bins have uniform widths over the range between -1.0 and 1.0 inclusive (each labAngle bin width is 0.02.) In addition, MC generator Λ^0 and $\overline{\Lambda}^0$ that decayed inside the BaBar beampipe via the $\Lambda^0 \rightarrow p^+\pi^-$ channel (and its conjugate) also are gathered into separate 2-d histograms using the same bin definitions as before. These histograms form the bases for constructing the two dimensional efficiency grids for the beampipe Λ^0 sample. Figure 6.4 displays the above-referenced 2-d histograms for reconstructed MC Λ^0 and generator MC Λ^0 that decayed within the beampipe.

The reason for instituting the 2-d histograms based on candidate lab momentum and labAngle is that we want to construct MC efficiencies that are relatively unbiased in these variables. Our Lambda cuts are mostly based on fairly well known quantities such as the Lambda decay length, the boost due to the PEP-II beams, and the $p_\perp$ distribution of Lambda decays; therefore, in general the MC simulation of the Lambda decay process reproduces the data quite well. In contrast, the Lambda momentum spectrum and the detector geometry, acceptance, and material interaction are believed to be not as well simulated by the MC (indeed, the current work seeks to obtain an independent measurement of the spectrum.) The two-dimensional momentum/labAngle binning of the efficiency matrix then minimizes the model-dependent biases introduced into our analysis by MC-derived efficiency values.

We now describe the construction of the Λ^0 efficiency matrix; the efficiency matrix for $\overline{\Lambda}^0$ is constructed in a completely analogous manner.

First, in order to account for possible differences between MC and off-peak data with respect to the relative distribution of the charged track mutliplicity event variable, the 2-d histograms for reconstructed MC Λ^0 ("RecoMC") and for MC generator Λ^0 ("GenMC") are each divided into five constitutent histograms for the following event multiplicity bins: 3-5,6-

---

[19] A 'true' generator Λ^0 is one where one decay daughter is associated with a true generator proton and another one is associated with a true generator pion, and where both of the proton and pion daughters come from the same generator Λ^0 particle.

[20] What we shall from now refer to as 'labAngle' is in fact a cosine angle calculated as follows using a candidate's BaBar's lab-frame momentum:

$$\text{labAngle} \equiv \frac{p_z}{\sqrt{p_\perp^2 + p_z^2}} \tag{6.7}$$

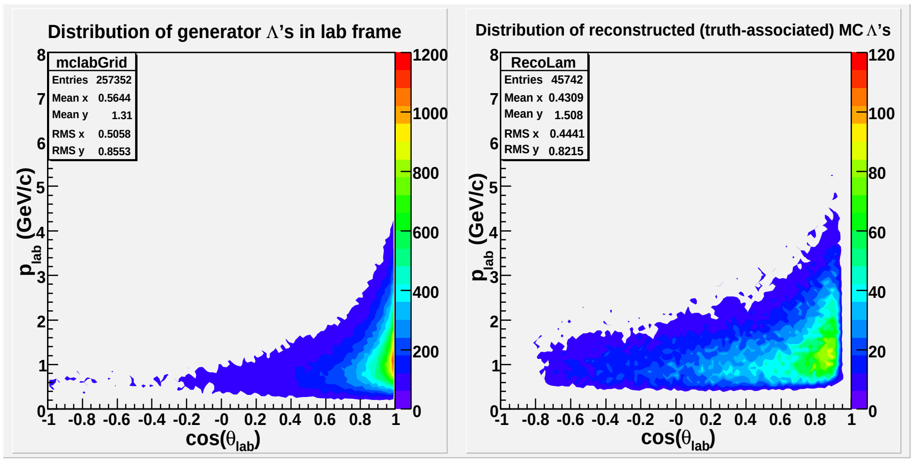


Figure 6.4: (Left) Two dimensional histogram of MC generator Λ^0 that decayed inside the beampipe. (Right) Two dimensional histogram of reconstructed and Truth-associated MC Λ^0 that decayed inside the beampipe. Note the change in scale between the two plots.

8,9-11,12-14,$\geq$ 15. Also note that during the construction of the multiplicity-binned RecoMC histograms, we instituted the Data/MC tracking correction factors described in Chapter 5 herein: for each detected and Truth-associated RecoMC candidate, we obtain the momentum and labAngle of the MC proton and pion daughters and then look up the appropriate p^+ and π^- correction factor for the relevant momentum, labAngle, and event multiplicity bin. The weighted entry for each detected RecoMC candidate then is obtained by multiplying the candidate's protonic correction factor by its pionic correction factor. The error associated with the correction factors are propagated and saved in separate 2-d histograms (with the same momentum and labAngle bin definitions.)

The multiplicity-binned 2-d constituent histograms are weighted by the relative event multiplicity fractions obtained from the complete set of Runs 1-4 off-peak data and mixed with each other. This results in new 2-d RecoMC and GenMC histograms which conform with off-peak data's multiplicity ratios.

Next, the raw efficiency matrix for the Λ^0 is constructed simply by dividing the contents of our new RecoMC histogram by the new GenMC histogram *bin by bin* for the entire set of 10,000 bins. For any given bin, the result of the calculation gives our initial estimate of the efficiency for reconstructing a beampipe Λ^0 that falls within the relevant bin's lab momentum and labAngle range.

However, the raw efficiency matrix for Λ^0 is subject to statistical fluctuations; if the number of particles in our sample of beampipe Λ^0 approaches ∞ and if the number of bins in the matrix also approaches ∞, then we'd expect the raw efficiency matrix to be a smooth function of both momentum and labAngle. One way to circumvent the 'jaggedness' in the efficiency matrix due to statistics is to deliberately smooth out the efficiency variations between neighboring bins by selectively enlarging our bin definitions.

The smoothing procedure is an iterative one (up to a maximum of five iterations.) For a given bin in the two dimensional lab momentum/labAngle grid, it usually conforms to the following:

- If the relative statistical uncertainty [21] is less than 3.50%, then the bin efficiency is considered sufficient for smoothing purposes and no more iterations are needed.
- Else, if the bin's relative statistical uncertainty is greater than 3.50% and the bin is not near the edge of the grid (to be defined later,) then we add the contents from the bin's eight neighboring bins (this defines a square centered on our original bin's central values) and recalculate the relative statistical uncertainty using summed values from all 9 bins.
- If the recalculated relative statistical uncertainty from the previous step is less than 3.50% then we stop further iterations; otherwise we proceed to add all bins within a square with five bins on a side and centered on our original bin's center value (we're now summing 25 bins) and recalculate the uncertainty using the summed values.
- If the recalculated relative statistical uncertainty from the previous step continues to be greater than 3.50%, we proceed to integrate the contents from a square grid with 49 bins centered on our original bins, then 81 bins and so on.

However, if the bin under consideration has a central labAngle value $\leq$ -0.92, then no additional bins with labAngle values lower than the center bin are added to the sum; similarly, if the bin's central labAngle value is $\geq +0.96$, then no additional bins with labAngle values greater than the center bin are added to the sum. In addition, if the center bin's central momentum value is ≤ 0.24 GeV/c, then no additional bins with momentum values less than the center bin are added to the sum. Finally, the above-described smoothing procedure ceases when it reaches a kinked cutoff line defined by the following two equations (smoothing occurs only for bins with central momenta *less* than the momentum defined by the kinked line):

$$(p\ +\ 0.6)\ =\ \left(\frac{2.25-1.75}{0.2-(-0.7441)}\right)\cos\left(\theta_{lab}\right)-(1.75+0.39408) \qquad \cos\left(\theta_{lab}\right)<0.2 \tag{6.9}$$

$$(p\ +\ 0.6)\ =\ \left(\frac{4.5-2.25}{0.9171-0.2}\right)\cos\left(\theta_{lab}\right)-(1.62247) \qquad \cos\left(\theta_{lab}\right)\geq 0.2 \tag{6.10}$$

where p is the value of the Λ^0 candidate's lab moomentum and where $\cos\left(\theta_{lab}\right)$ is the candidate's labAngle (i.e. the cosine of the angle defined by a particle's momentum vector with respect to the BaBar z-axis.) Note that for bins with central momenta greater than that defined by the 'kinked line', both the generator and reconstructed Λ^0 and $\overline{\Lambda}^0$ numbers are very low; however, the particles in this region have high lab momentum and generally have fairly high detection efficiencies. Thus in order to avoid throwing away real detected particles, we performed a simple averaging of bins in this region: the efficiencies for all bins with central

---

[21] The relative bin statisitical uncertainty is calculated as follows:

$$\sigma_{rel.stat.}=\left(\frac{\sigma_{reco}}{N_{gen}}\right)\left(\frac{1}{\epsilon}\right) \tag{6.8}$$

where N_{reco} is the total number of RecoMC candidates in the bin(s), N_{gen} is the total number of GenMC candidates in the bin(s), $\epsilon \equiv N_{reco}/N_{gen}$ is the interim raw efficiency of the bin(s), and $\sigma_{reco}=\sqrt{N\epsilon(1-\epsilon)}$ is the estimated standard deviation of the total number of candidates in the bin(s) assuming that this number conforms to a binomial distribution.

momentum values greater than that defined by the ‘kinked line’ are indiscriminantly summed into two large bins, one for labAngle bin values ≥ 0.2 and another for labAngle bin values < 0.2.

Finally, in the forward labAngle and low lab momentum region where the number of available MC generator and reconstructed particles are relatively plentiful, we perform the smoothing bin integration over only a maximum of 9 bins (1 integration iteration only;) this cutoff in integration iterations speeds up the speed of the integration code considerably. The labAnlge-momentum region where we implement a maximum of only one integration iteration is defined by the following equation:

$$(p \ + \ 0.4) \ \leq \ \left(\frac{2.25 - 0.5}{0.9171 - 0.2}\right) \cos\left(\theta_{lab}\right) - (0.5 - 0.00192) \tag{6.11}$$

The 2-d histogram in Figure 6.5 represents the calculated statistical error for each bin in the smoothed Λ^0 efficiency grid. The 2-d histogram in Figure 6.6 is a contour plot where the color of each bin represents the total number of bins integrated during the Λ^0 efficiency grid smoothing process that centered on the relevant bin.

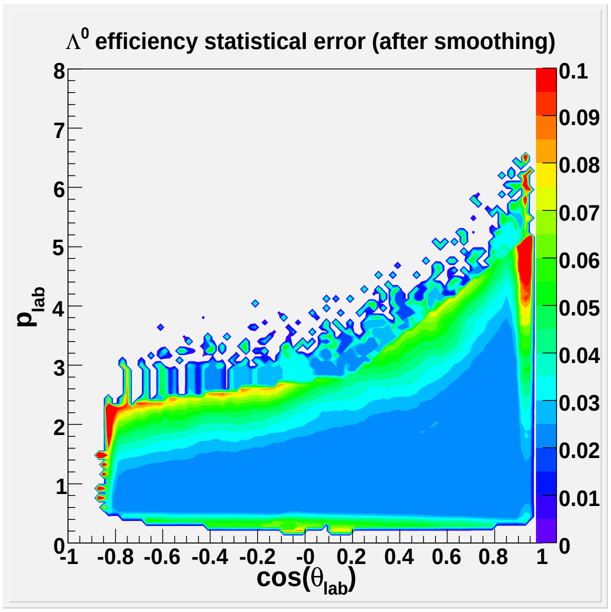


Figure 6.5: Relative statistical error associated with each bin in the smoothed Λ^0 efficiency grid.

After application of the initial smoothing routine described above, there generally remains some residual efficiency ‘spikes’ in certain bins; hence we apply a second level of smoothing in which we remove such spikes by interpolating neighboring (smoothed) bins.

The secondary smoothing routine is applied to bins that are ‘far’ from the grid edges (as defined above) and that have central momenta less than the above-refeneced ‘kinked line’. The routine begins by comparing a relevant bin’s efficiency with those for its nearest neighbors and next-nearest neightbors along the labAngle and momentum directions; if the efficiency difference between the bin and any of its nearest neighbors exceeds 10% OR if the efficiency difference between the bin and any of its next-nearest neighbors exceeds 20%, then

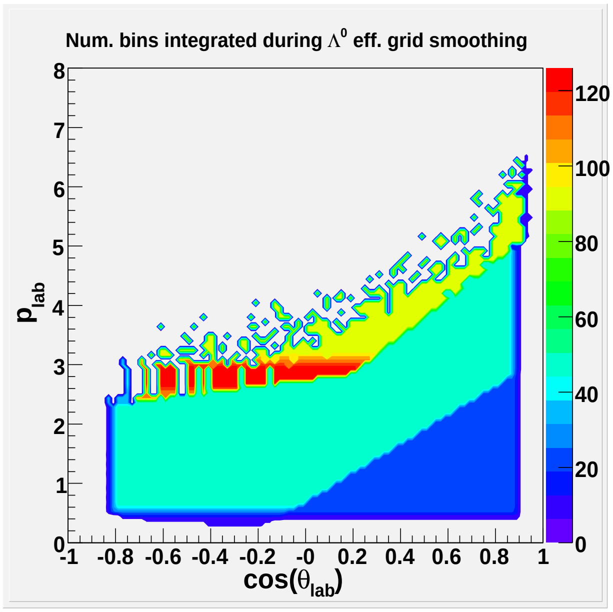


Figure 6.6: Number of bins integrated (centered on location of each bin) during the Λ^0 efficiency grid smoothing process.

efficiency interpolation is performed. Efficiency interpolation involves averaging the bin's two nearest neighbor efficiencies (neighbors in the grid direction in which the predefined efficiency difference limit has been exceeded,) *plus* its two next-nearest neighbor efficiencies, *plus* its two third-nearest neighbor efficiencies. The propagated tracking error for an interpolated bin is also interpolated in an analogous manner. Figure 6.7 displays the final results of the smoothing procedures for Λ^0 and $\overline{\Lambda}^0$.

The propagated error in each smoothed bin that is associated with the Data/MC tracking corrections for a candidate's decay daughters is also saved to a separate two-dimensional histogram; these errors will be accessed bin-by-bin during the efficiency correction process. Note that the propagated tracking errors are represented as an error on the efficiency. Figure 6.8 plots the two-dimensional histogram of the propagated tracking correction errors associated with the various bins in the smoothed Λ^0 efficiency grid.

6.2.3 Λ efficiency correction and signal fits

We run over all off-peak data collected during the period that spans BaBar Runs 1 through 4. The Λ^0 and $\overline{\Lambda}^0$ candidates from the off-peak dataset are composed from pairs of charged tracks using the same χ^2 based vertexer that we employed to compose the MC candidates; the background in the particle and antiparticle mass spectra are reduced using the same cuts described in section 6.1.1 herein. All accepted candidates that decayed inside the beampipe then are efficiency-corrected using the appropriate two-dimensional beampipe efficiency matrices described in section 6.1.2; the efficiency correction is performed on a candidate-by-candidate basis (including all residual background candidates) and the efficiency-corrected number of Λ^0 and $\overline{\Lambda}^0$ candidates then are binned with respect to the candidates' momenta in the PEP-II beams' CM reference frame and plotted using the candidates' invaraint mass. The momen-

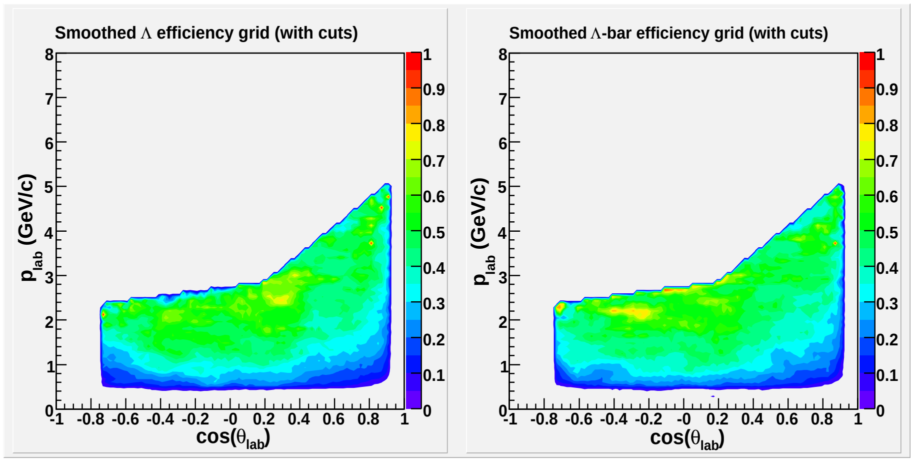


Figure 6.7: Track-corrected and smoothed (also mixed by data event multiplicity fractions) MC Λ^0 and $\overline{\Lambda}^0$ 2-d efficiency grids

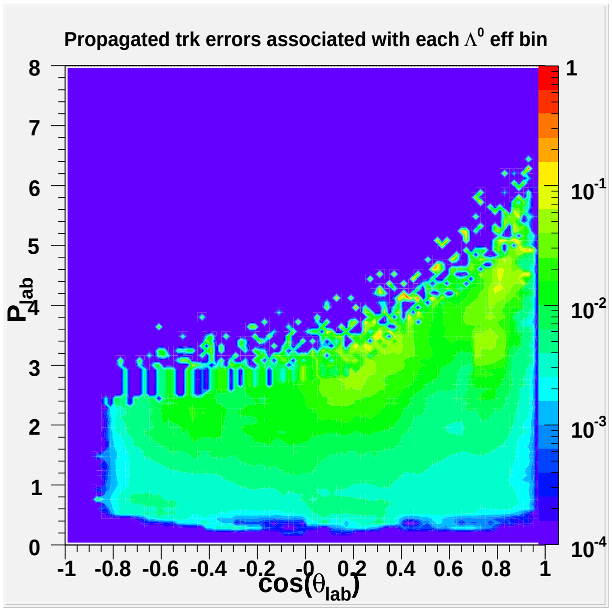


Figure 6.8: Propagated tracking correction error associated with each bin in the smoothed Λ^0 efficiency grid.

tum bins have uniform widths of 0.25 GeV/c for CM momentum range between 0.0 GeV/c to 4.25 GeV/c; a single bin encompasses all candidates within the momentum range going from 4.25 GeV/c up to the momentum limit of about 5.24 GeV/c (so there are a total of 18 momentum bins.)

Recall that BaBar's lab frame is forward-boosted and rotated with respect the CM frame of the PEP-II beams. The relationship between the candidates' momenta and labAngle in the two reference frames can be displayed as a banded contour where each band (plotted in the lab frame variables) represents an approximately iso-momentum or iso-labAngle level in BaBar's CM frame. See Figure 6.9.

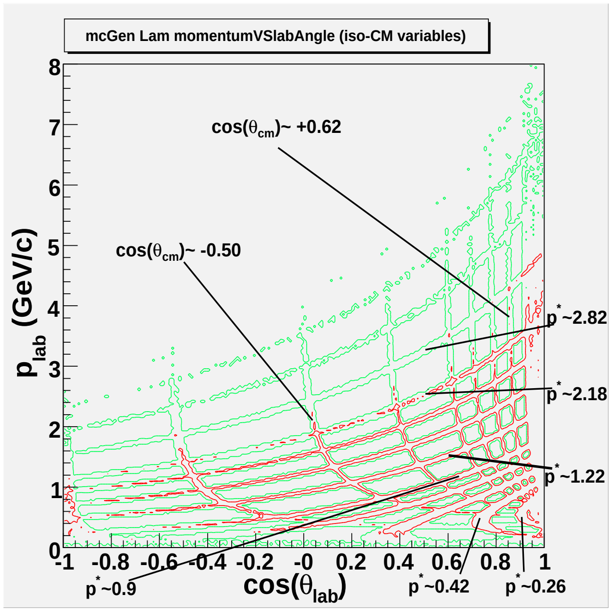


Figure 6.9: A graphical representation of MC Λ^0 momentum/labAngle lab frame to CM frame conversion.

The resulting momentum-binned $\Lambda^0/\overline{\Lambda}^0$ invariant mass plots are fitted using a second-order polynomial background function and a double gaussian signal function. The software routine responsible for the various fits is ROOT's implementation of the fortran-based MINUIT fitting package. We employ the binned log-likelihood method [22] to fit the combined background and signal functions to the measured mass distributions. Figure 6.10 displays examples of the momentum-binned invariant mass plots and the resulting signal+background fits for the case where we've compacted the Λ^0 histograms with the $\overline{\Lambda}^0$ histograms.

The shifts in the fitted signal invariant mass and the variations in the widths are contained in Figure 6.11.

---

[22] According to the code of the ROOT function TH1::Fit in the class TH1, the default fitter is the TFitter class which calls TMinuit (ROOT implementation of the MINUIT package.) When the log-likelihood fitting procedure is called to fit a binned ROOT histogram, the static function 'H1FitLikelihood' is defined in TFitter which forms the sum of the logarithms of the individual bins' likelihoods (defined as the predicted Poisson probability for the number of entries in a bin given the mean predicted value from the fitting function.) H1FitLikelihhod is then minimized using the TMinuit class. See http://root.cern.ch/root/html/404/src/TFitter.cxx.html, http://root.cern.ch/root/html/404/src/TH1.cxx.html#TH1:Fit, and p.72 in [35].

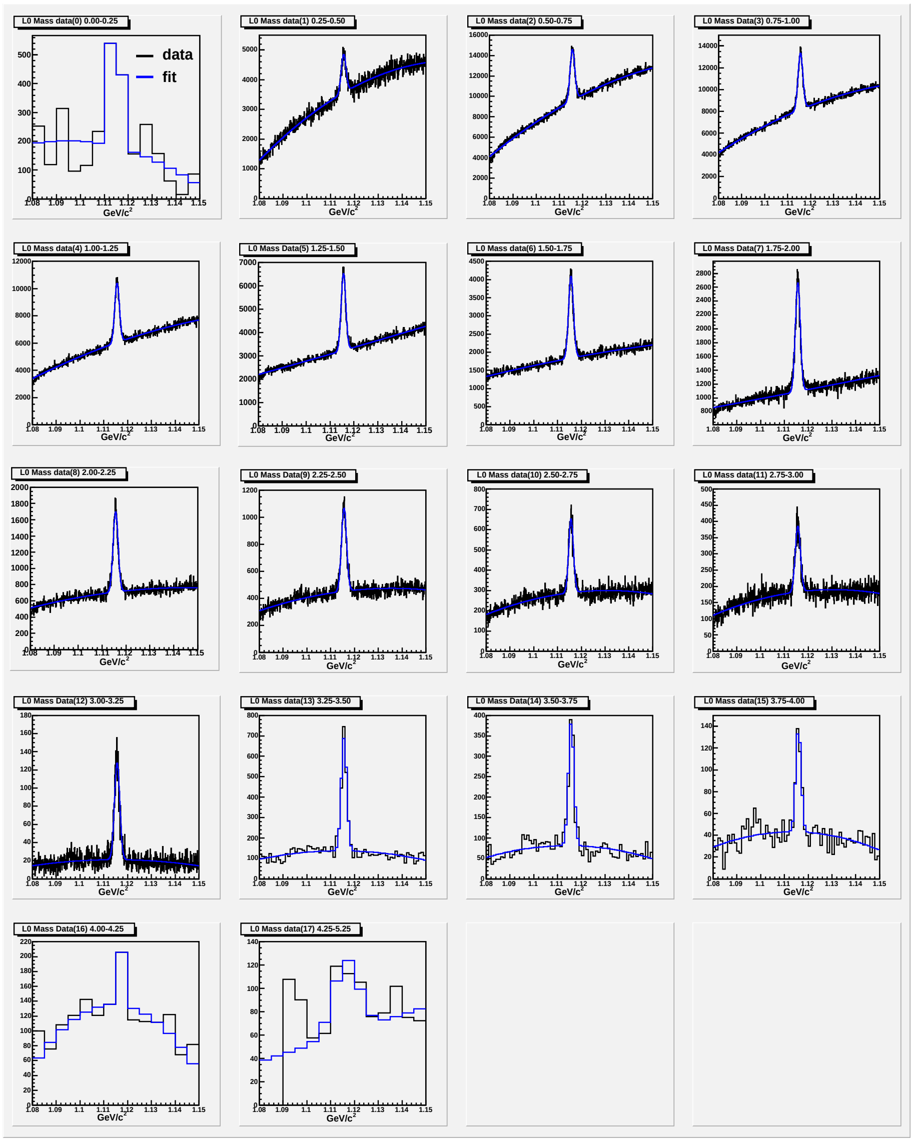


Figure 6.10: Λ^0+$\overline{\Lambda}^0$ CM momentum-binned invariant mass plots and fits.

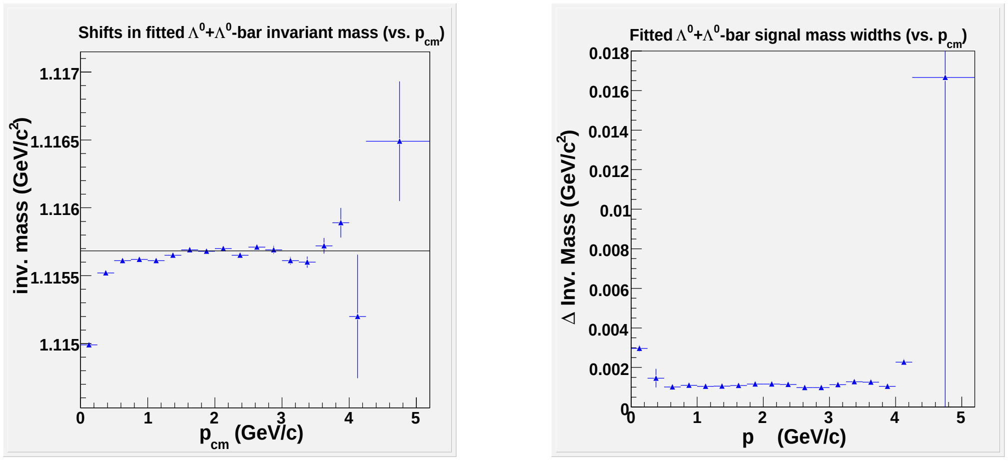


Figure 6.11: (Left) Shift in Λ^0+$\overline{\Lambda}^0$ fitted mass versus CM momentum (note the black line represents the PDG value for the Λ. (Right) Fitted widths of Λ^0+$\overline{\Lambda}^0$ signals versus CM momentum. The quoted widths are one-half the full-widths at half-maximum from single gaussian fits of the signals.

The fits of the invariant mass signals give us the efficiency-corrected Λ^0 and/or $\overline{\Lambda}^0$ that decayed within the beampipe via $\Lambda^0 \to p^+\pi^-$ and its conjugate. However, there are certain regions in the efficiency matrices where the MC efficiencies are very low; these regions were excluded from the momentum-binned efficiency correction step. We need to account to these excluded regions in order to obtain a fully efficiency-corrected beampipe charged-decay numbers. The excluded regions include all efficiency bins where

- Λ^0 or $\overline{\Lambda}^0$ labAngle < -0.7441.
- Λ^0 or $\overline{\Lambda}^0$ labAngle > +0.9171.
- For Λ^0 or $\overline{\Lambda}^0$ candidates with CM momentum < 0.25 GeV/c, all efficiency bins with lab momentum < 0.2 GeV/c are excluded.
- For Λ^0 or $\overline{\Lambda}^0$ candidates with CM momentum $\geq$ 0.25 GeV/c, all efficiency bins with lab momentum < 0.5 GeV/c are excluded.
- For Λ^0 or $\overline{\Lambda}^0$ candidates with CM momentum < 0.25 GeV/c or $\geq$ 3.0 GeV/c, all efficiency bins where the efficiency is $\leq$ 1.0% are excluded.
- For Λ^0 or $\overline{\Lambda}^0$ candidates with CM momentum $\geq$ 0.25 GeV/c and < 3.0 GeV/c, all efficiency bins where the efficiency is $\leq$ 10.0% are excluded.

The relative number of MC generator Λ^0 and $\overline{\Lambda}^0$ that are rejected and accepted due to the above cuts are listed in table 6.1. We employ these fractions to fully efficiency-correct (or extrapolate) the total number of beampipe particles and antiparticles that decayed via the charged channels.

We also need to correct for all the particles and antiparticles which were cut by our decay-distance based cut which limited our samples to candidates with reconstructed vertices inside

Candidate CM momentum Bins (GeV/c)	Fraction of MC Λ^0 + $\overline{\Lambda}^0$ (Rejected by efficiency cuts / Accepted)
0.00-0.25	120.6703
0.25-0.50	4.2606
0.50-0.75	0.8934
0.75-1.00	0.6112
1.00-1.25	0.4448
1.25-1.50	0.4141
1.50-1.75	0.4321
1.75-2.00	0.4392
2.00-2.25	0.4478
2.25-2.50	0.4541
2.50-2.75	0.4631
2.75-3.00	0.4364
3.00-3.25	0.4484
3.25-3.50	0.5000
3.50-3.75	0.4752
3.75-4.00	0.4360
4.00-4.25	0.4000
4.25-5.15	1.1923

Table 6.1: Fractions of MC Λ^0 and $\overline{\Lambda}^0$ rejected/accepted by cuts on efficiency tables.

the PEP-II beampipe (note that we're still dealing only with Λ^0 and $\overline{\Lambda}^0$ that decayed via the charged channels.) This correction/extrapolation is again performed using MC generator particles grouped by CM momentum. Table 6.2 gives the fractions of generator particles and antiparticles that were cut by our decay distance cut relative to the number accepted in CM momentum bins.

The Λ^0 and/or $\overline{\Lambda}^0$ spectra displayed in the next chapter are obtained after applying all the above corrections/extrapolations and then further corrected using the branching fraction for the charged channel decay currently cited by the PDG.

6.3 Ξ^- signal extraction

We begin composition of the cascade hyperon (i.e the Ξ^-) by requiring all Λ^0 and $\overline{\Lambda}^0$ candidates that fall within a very loose invariant mass window (between 1.075 and 1.15 GeV/c^2) in any accepted collision event should be vertexed with all charged tracks (excluding tracks used to compose the Λ candidate) using the pion mass hypothesis. The vertexing step again employs BaBar's χ^2 minimization vertexer described in chapter 6.2 herein. This same basic vertexing procedure will be employed to compose Ω^- hyperon candidates (but the additional charged track then will be given the kaon mass hypothesis instead;) however, due to the intrinsic scarcity of the Ω^- hyperons, PID must be used to cut down on the background. For the sake of convenience and in order to treat the cascade and Ω^- hyperons on the same footing as much as possible, we also imposed the PID cut during the composition of the cascade hyperon candidate. Specifically, we require that the purported protonic daughter of

Candidate CM momentum Bins (GeV/c)	Fraction of MC $\Lambda^0 + \overline{\Lambda}^0$ (Rejected by decay distance cut / Accepted)
0.00-0.25	19.1514
0.25-0.50	2.8954
0.50-0.75	2.0136
0.75-1.00	2.5599
1.00-1.25	3.0526
1.25-1.50	3.6897
1.50-1.75	4.5788
1.75-2.00	5.3678
2.00-2.25	6.1314
2.25-2.50	6.8499
2.50-2.75	7.6113
2.75-3.00	8.0985
3.00-3.25	8.7425
3.25-3.50	9.3080
3.50-3.75	10.1853
3.75-4.00	10.1311
4.00-4.25	10.4500
4.25-5.15	13.7692

Table 6.2: Fractions of MC Λ^0 and $\overline{\Lambda}^0$ rejected by decay distance cut relative to number of accepted candidates.

the Λ^0, which in turn is used to compose the Ξ^-, must be identified only as a proton by the global likelihood PID selector described in [30].

At the initial stage of the composition, no cuts other than the ones described above were applied to the Ξ^- candidates. Figure 6.12 displays the invariant mass spectrum of the Ξ^- + $\overline{\Xi}^+$ candidates from Run4 off-peak data obtained after this stage.

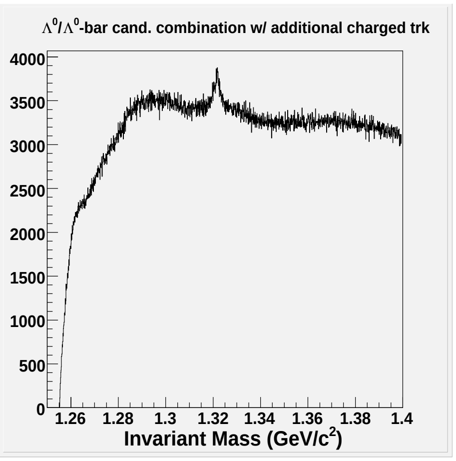


Figure 6.12: Combinations of Λ^0 and $\overline{\Lambda}^0$ with any other non-repetitive charged track in the event using a π mass hypothesis (after PID cut.)

We now proceed to impose a number of aditional cuts in order to abate the not insignificant background that is clearly evident in Figure 6.12.

6.3.1 Cuts against Ξ candidates

The additional cascade hyperon cuts include the following:

- The Ξ^- candidate's dXY (XY distance between decay vertex and event vertex) must be > 2.0 mm.
- The Λ^0 daughter's dXY also must be > 2.0 mm.
- For Ξ^- candidates' with CM momentum up to about 2.50 GeV/c, we further require both the Ξ^- and its Λ^0 daughter must have decayed within the PEP-II beampipe (we will refer to these cascade hyperons as the 'beampipe' sample.) This cut ensures that in the momentum region where the Ξ^- is expected to have the largest cross section, we can reliably correct for differences between the particles' data and MC efficiencies using our beampipe corrections for protons and pions.
- The proton and two pion charge combination must be consistent with either the $\Xi^- \rightarrow \Lambda^0 + \pi^- \rightarrow p^+ + \pi^- + \pi^-$ or with $\overline{\Xi}^+ \rightarrow \overline{\Lambda}^0 + \pi^+ \rightarrow p^- + \pi^+ + \pi^+$.
- A much tighter Λ^0 mass window with total width of 5 MeV/c^2 (from 1.1132 GeV/c^2 to 1.1182 GeV/c^2.)

We'll refer to the above-described set of Ξ^- cuts as the 'preliminary' cuts. The effect of some of the sinlge cuts listed above on Run4 off-peak data vis-a-vis Figure 6.12 are given in Figure 6.13. The combined effects of all the preliminary cuts is displayed in Figure 6.14.

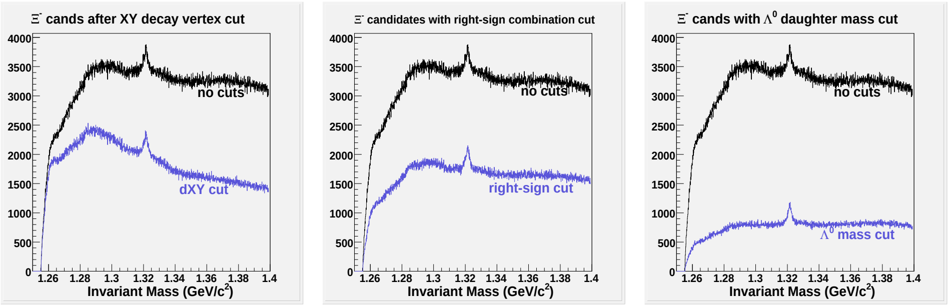


Figure 6.13: (Left) Effect of Λ^0 and Ξ^- dXY cuts on the acceptance of Ξ^- candidates; (center) effect of 'right sign' combination requirement; (right) effect of tight Λ^0 invariant mass cut.

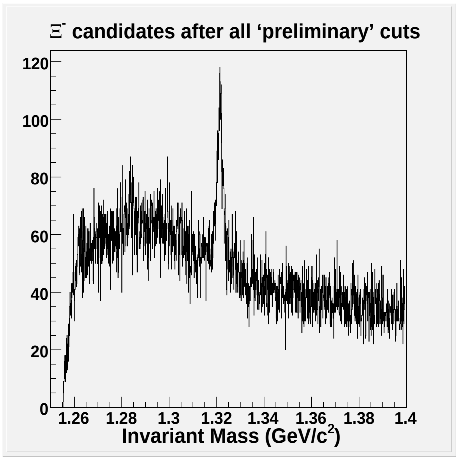


Figure 6.14: Combined effect of all 'preliminary cuts' on the acceptance of Ξ^- candidates.

The combinatoric background in the Ξ^- invariant mass spectrum remains overly prominent after the preliminary cuts. Additional cuts detailed below further reduces this background:

- The Ξ^- candidate's purported Λ^0 daughter must have pointing angle (recall this is a cosine angle) between 0.0 and 0.9999 (this corresponds to an angle of between 0.81^o and 90^o.)

- The Λ^0 daughter's own protonic daughter must have 12 or more DCH hits.

- The extra π that is used (along with the Λ^0) to compose the Ξ^- must have at least 6 SVT hits as well as at least 12 DCH hits.

- The Λ^0 daughter's momentum vector, projected backward towards the Ξ^- decay vertex, must have a DOCA of less than 3 mm. from that vertex.

- The additional π used to compose the Ξ^- must have a DOCA of less than 3 mm. from the Ξ^- decay vertex (the DOCA must be measured using the π track quantities resident prior to fitting the π track as coming from the decay vertex.)

- The Λ^0 daughter's decay vertex in the XY plane must have a distance at least 2 mm in excess of the Ξ^-'s radial decay distance.

Figure 6.15 displays the individual effects of the proton DCH-hit cut, π hit cuts, and Λ^0 pointing angle cut on the Ξ^- invariant mass distribution relative to the effects of the preliminary cuts only. Finally Figure 6.16 displays the combined effect of all aforementioned cuts on the Ξ^- invariant mass distribution.

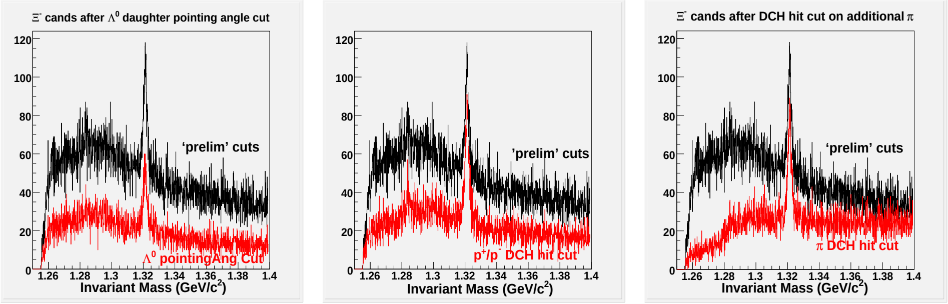


Figure 6.15: (Left) Effect of Λ^0 pointing angle cuts on the acceptance of Ξ^- candidates; (center) effect of Λ^0's protonic daughter DCH hit requirement; (right) effect of extra π's hit requirements.

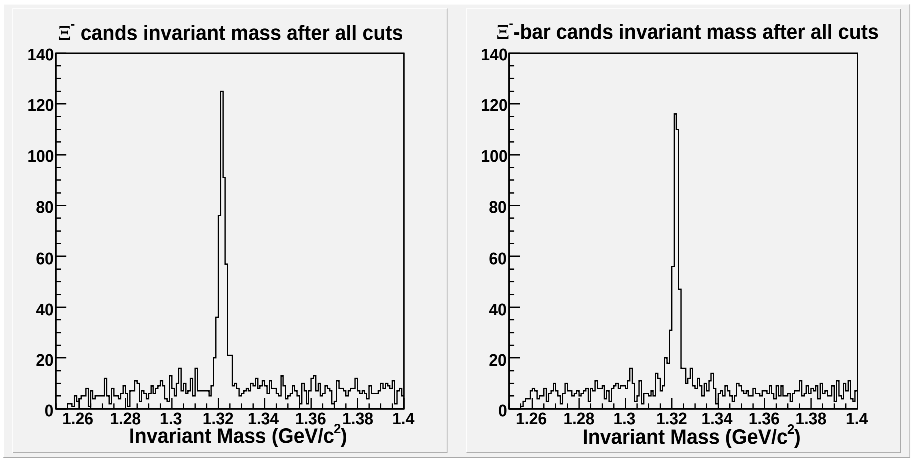


Figure 6.16: Combined effect of all specified cuts on the acceptance of $\Xi^-+\overline{\Xi}^+$ candidates from Runs 1-3.

6.3.2 Ξ efficiency matrix construction

We construct the Ξ^- and $\overline{\Xi}^+$ efficiency matrices in an analogous manner to the Λ^0 and $\overline{\Lambda}^0$ efficiency matrices. The same background cuts elaborated in the previous subsection are applied to reconstructed Ξ^- and $\overline{\Xi}^+$ candidates from the MC. Those MC candidates that survive these cuts and which are deemed to be true Ξ^- or $\overline{\Xi}^+$ MC particles [23] are then saved into separate 2-d histograms (again in the candidates' labAngle and lab momentum variables using 50 × 50 uniform-width bins.)

For candidates with CM momenta less than 2.50 GeV/c, we restrict our selected sample only to those with a daughter Λ^0 that has also decayed within the PEP-II beampipe; this selection greatly reduces the total number of MC generator particles that can be used to produce the efficiency matrices and for this reason we increase the total number of MC 'udsc' events employed for these purposes to over 200 million, corresponding to an integrated luminosity of $\sim 60\ fb^{-1}$. The 2-d histograms for the reconstructed MC Ξ^- and $\overline{\Xi}^+$ (Ξ's with CM momenta ≥ 2.50 GeV/c were reconstructed but not used in this part of the analysis) are displayed in Figure 6.17. Analogous histograms for the MC generator Ξ^- and $\overline{\Xi}^+$ that decayed within the PEP-II beampipe (these form the denominator of the effincies) are also displayed in Figure 6.18.

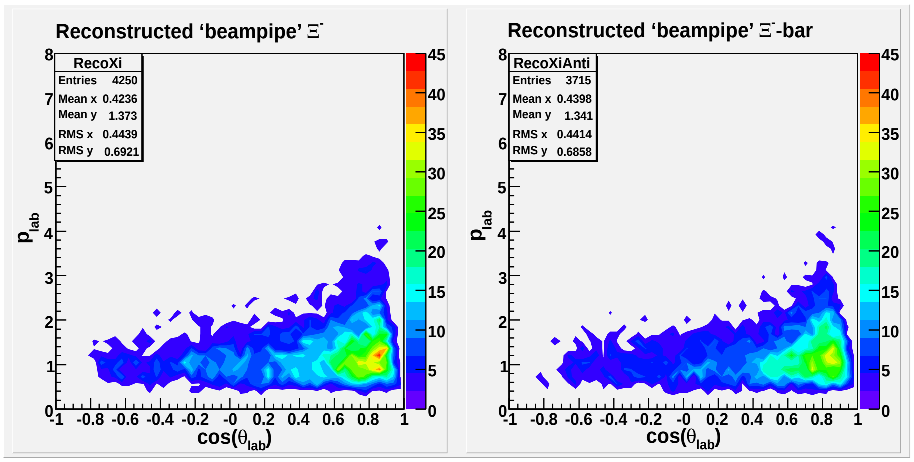


Figure 6.17: Histograms of reconstructed MC' beampipe' Ξ^- and $\overline{\Xi}^+$ candidates which are truth-associated with MC generator particles.

Note that the construction of efficiency matrices for candidates with CM momenta ≥ 2.50 GeV/c (candidates with momenta less than 2.50 GeV/c were reconstructed but not used in this part of the analysis,) we use both the so-called 'beampipe' cascade particles as well as particles where the Lambda daughter decayed inside the inner SVT (< 12 cm in the radial XY BaBar coordinate;) since the statistics is relatively good for this subsample, we use only 25 million udsc MC events.

The efficiency matrices are constructed by dividing the number of reconstructed cascade hyperon particles in each of the two-dimensional labAngle/lab-momentum bins and dividing

---

[23] A true MC Ξ^- or $\overline{\Xi}^+$ particle is deemed to have been found if the Lambda daughter is a 'true' Lambda (as explained in footnote 19) that came from the decay of a cascade hyperon; in addition, the additional π also must be a true MC pion that came from the same cascade hyperon.

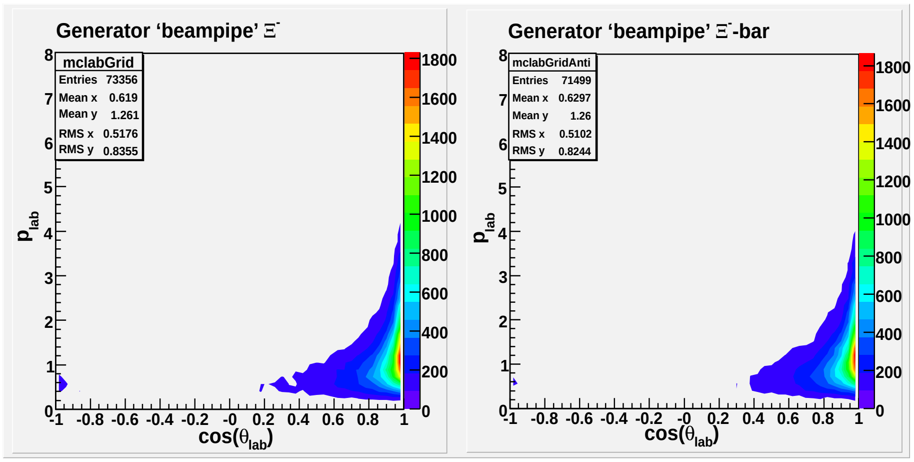


Figure 6.18: Histograms of MC generator Ξ^- and $\overline{\Xi}^+$ candidates that decayed inside the PEP-II beampipe.

it by the number of generator MC particles found in the same bin. As in the case for Λ^0 and $\overline{\Lambda}^0$, the matrices are first divided into five multiplicity bins (using the same multiplicity bin definitions for the case of Lambdas,) and then remixed using multiplicity fractions obtained from off-peak data as weights. We proceed to smooth the statistical jaggedness present in the multiplicity-remixed efficiency matrices constructed in the above manner.

Due to the relative scarcity of the Ξ^- in relation to the Λ^0, we further compact the bin definitions of the efficiency matrices *prior* to the smoothing procedure; the compaction involves reducing the 50 × 50 bins in the labAngle (-1.0 to 1.0) and lab momentum (0.0 to 8.0 GeV/c) variables into 25 × 25 equally-spaced bins: this is accomplished by simple bin integration of both the numerator (RecoMC) and denominator (Generator MC) 2-d matrices and then recomputing the multiplicity-mixed efficiency matrices.

After the reduction from 2,500 to 625 total bins, we use the exact same smoothing algorithm that was employed for the Lambda (described in chapter 6.2.2 herein) *with the proviso* that only bin integrations up to and including the two nearest neighbors are now allowed (i.e. only a maximum of two integration iteration is permitted.) The resulting smoothed Ξ^- and $\overline{\Xi}^+$ efficiency matrices are displayed in Figure 6.19.

Note that unlike the case for the Lambda hyperon, we did not directly include the tracking corrections for the final state protons and pions during the cascade hyperon's efficiency calculations. This is because if we include tracking corrections at this stage, the relative scarcity of reconstructed cascade candidates may further enhance the relative variations between the bins; instead, the tracking corrections will be applied to each reconstructed *data* candidate during the efficiency-correction stage.

The relative statistical uncertainty associated with each (smoothed) efficiency bin again is calculated by assuming that the statistical variation in the bin's numerator follows a binomial distribuion (see equation 6.8 infra.) Figure 6.20 displays the statistical error associated with the Ξ^- and $\overline{\Xi}^+$ efficiency bins.

In addition, in order to obtain sufficient data points for candidates with CM momenta greater than 2.50 GeV/c, we extended our Ξ selection to include candidates where the Λ^0 has decayed inside the inner SVT ($<$ 12 cm in the XY plane from the BaBar origin.) We

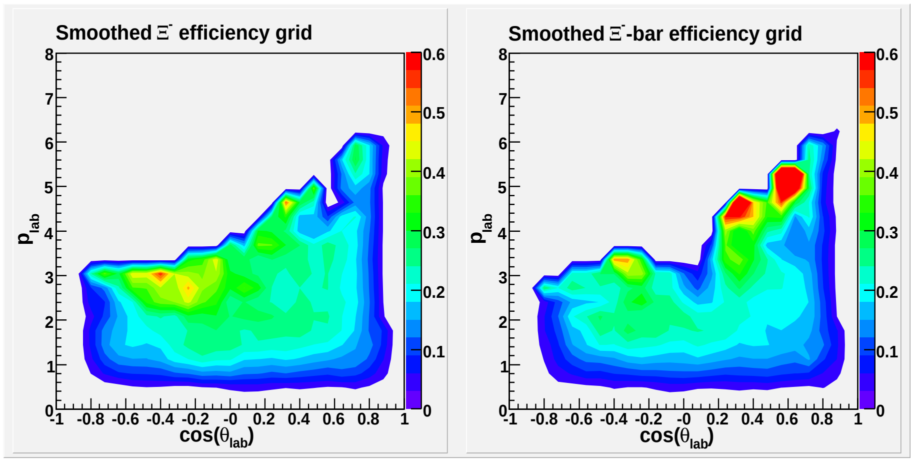


Figure 6.19: ‘Beampipe’ Ξ^- and $\overline{\Xi}^+$ efficiency matrices obtained from Monte Carlo.

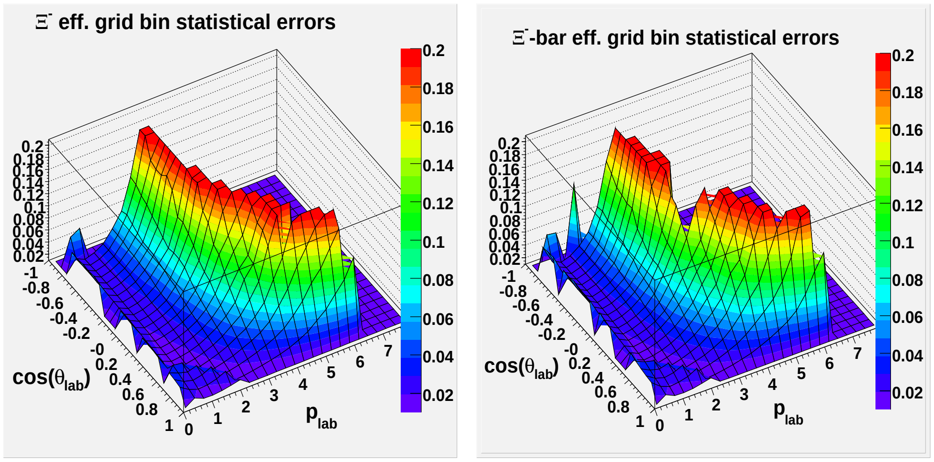


Figure 6.20: (Left) Relative statistical errors associated with each bin in beampipe Ξ^-’s efficiency matrix. (Right) Relative statistical errors associated with each bin in beampipe $\overline{\Xi}^+$’s efficiency matrix.

repeat the exact same steps elucidated above in the construction of efficiency matrices for the beampipe+SVT sample of cascade hyperons. The resulting efficiency matrices for this combined samples are displayed in Figure 6.21.

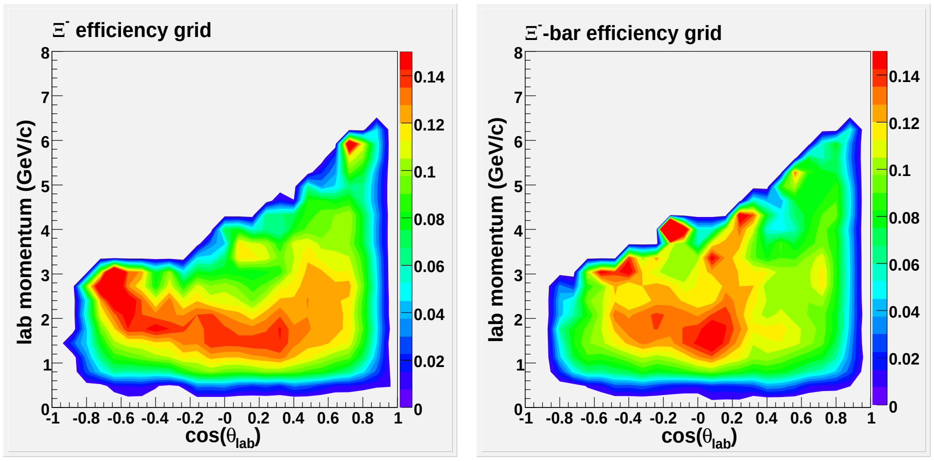


Figure 6.21: Ξ^- and $\overline{\Xi}^+$ efficiency matrices for the 'beampipe+SVT' Monte Carlo candidate sample.

6.3.3 Ξ efficiency correction and signal fits

The Ξ^- and $\overline{\Xi}^+$ data signals are obtained from BaBar's off-peak data from Runs 1 through 4 (Runs 1-3 for $p \geq 2.5$ GeV/c.) The same cuts and compositions routines that were used for the MC efficiency computations again were applied during the data runs. Efficiency-correction of the number of data candidates using the efficiency matrices described in the previous section is applied on a candidate-by-candidate basis; the tracking corrections for the final state proton and two pions that result from the cascade hyperon's charged decay are applied at this point. The efficiency-corrected number of accepted Ξ^- and $\overline{\Xi}^+$ are plotted in accordance with the candidates' invariant mass and binned with respect to their CM momentum in bins of 0.25 GeV/c bin-width (from 0.0 GeV/c up to 4 GeV/c;) an additional bin covers all candidates with CM momentum greater than 4 GeV/c (so there is a total of 17 CM momentum bins.)

The momentum-binned Ξ^- and $\overline{\Xi}^+$ invariant mass plots are fitted using a second-order polynomial background function and a signal function that convolutes a Breit-Wiger function with a gaussian function (this accounts for Ξs from resonant decays and the detector resolution.) The actual fitting is performed using ROOT's implementation of the MINUIT package.

Figure 6.22 displays the fitted 'beampipe' cascade hyperon signal and background invariant mass plots for the case where particles and antiparticles have been compacted into the same histograms. Although in Figure 6.22 the displayed bins for the beampipe sample include candidates with momenta up to 3.25 GeV/c, we use the beampipe candidates exclusively only up to a momentum of 2.50 GeV/c; for the momentum region $\geq 2.50 GeV/c$, we combine the beampipe and SVT Ξ samples. Figure 6.23 displays our fits for the beampipe+SVT sample (CM momentum bins $\geq$ 2.50 GeV/c.)

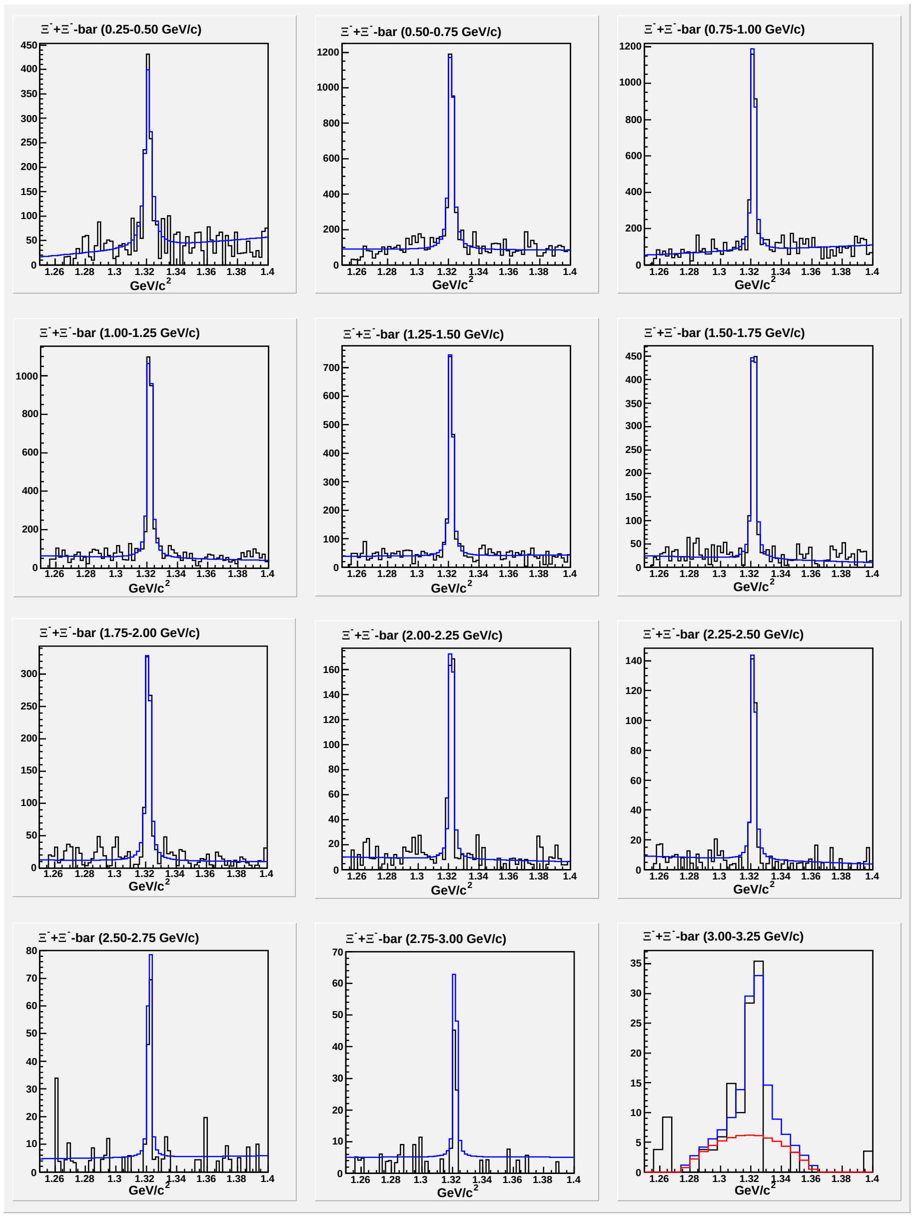


Figure 6.22: Fitted invariant mass spectra of efficiency-corrected 'beampipe' $\Xi^{-}+\overline{\Xi}^{+}$ from Runs 1-4 off-peak data (candidates with CM momenta up to 3.25 GeV/c.) Note the superimposed red histogram for the 3.00-3.25 GeV/c bin; this represents the fitted background in that bin. We only use those bins with momenta $<$ 2.50 GeV/c.

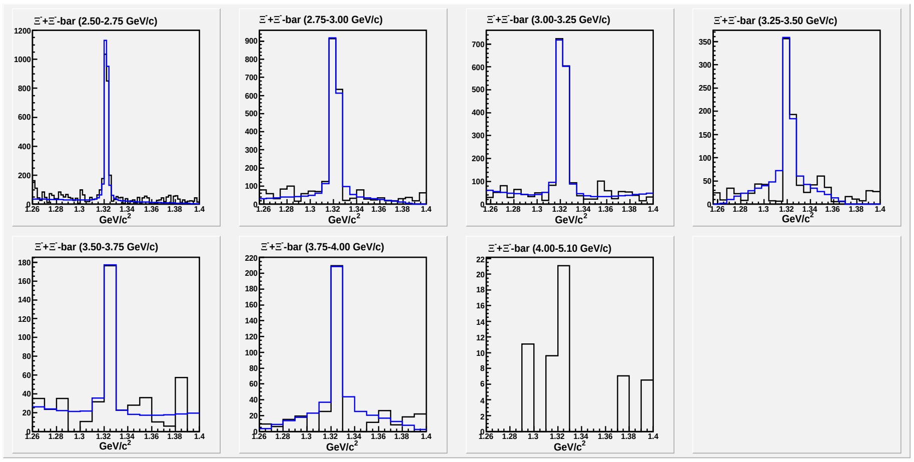


Figure 6.23: Fitted invariant mass spectra of efficiency-corrected 'beampipe+SVT' cascade hyperons from Runs 1-3 off-peak data (CM momentum $\geq$ 2.50 GeV/c.)

The shifts in the fitted signal invariant mass and the variations in the widths are contained in Figure 6.24.

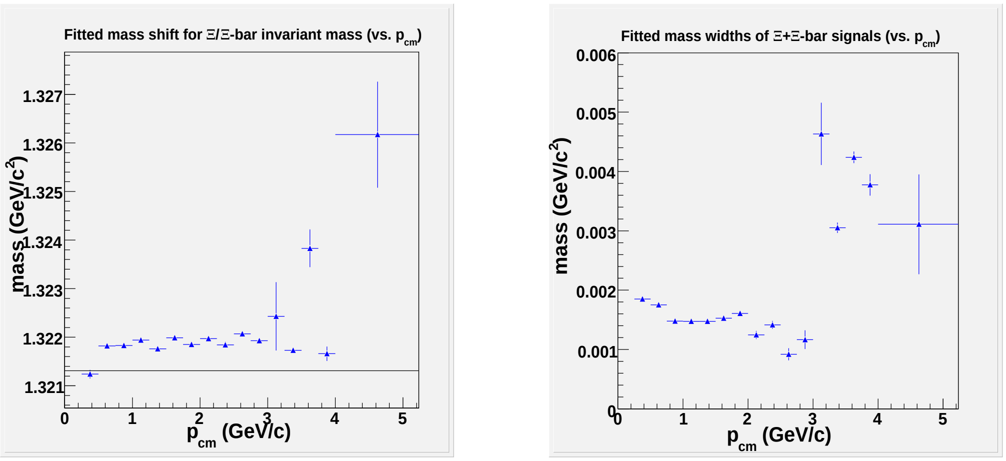


Figure 6.24: (Left) Shift in $\Xi^- + \overline{\Xi}^+$ fitted mass versus CM momentum (note the black line represents the PDG value for the Ξ. (Right) Fitted widths of $\Xi^- + \overline{\Xi}^+$ signals versus CM momentum. The quoted widths are one-half the full-widths at half-maximum from single gaussian fits of the signals.

A number of cuts against the efficiency matrices themselves also are instituted to eliminate bins with very small efficiencies or very few reconstructed and generated MC particles. These cuts include the following:

- Ξ^- or $\overline{\Xi}^+$ labAngle < -0.7441.
- Ξ^- or $\overline{\Xi}^+$ labAngle > +0.9171.

Candidate CM momentum Bins (GeV/c)	Fraction of MC $\Xi^- + \overline{\Xi}^+$ (Rejected by efficiency cuts / Accepted)
0.25-0.50	4.7156
0.50-0.75	1.4875
0.75-1.00	1.1526
1.00-1.25	1.2093
1.25-1.50	1.0682
1.50-1.75	0.9347
1.75-2.00	1.0487
2.00-2.25	1.1591
2.25-2.50	1.2273
2.50-2.75	0.2637
2.75-3.00	0.3746
3.00-3.25	0.3147
3.25-3.50	0.2961
3.50-3.75	0.2910
3.75-4.00	0.3202
4.00-5.10	2.3750

Table 6.3: Fractions of MC Ξ^- and $\overline{\Xi}^+$ rejected/accepted by cuts on efficiency tables.

- All efficiency bins where the Ξ^- or $\overline{\Xi}^+$ candidate lab momentum < 0.5 GeV/c are excluded.
- All efficiency bins where the efficiency is $\leq 4.0\%$ are excluded.
- All efficiency bins where the efficiency statistical error is $\geq 15.0\%$ are excluded.

We need to correct for the candidates rejected by these efficiency-table based cuts using Monte Carlo. Table 6.3 lists the relative number of generator Ξ^- and $\overline{\Xi}^+$ rejected and accepted due to the efficiency-based cuts. We also need to correct for all those candidates that were rejected by our decay distance cuts (candidates with momenta < 2.50 GeV/c are rejected if their Λ daughter decayed > 2.4 cm in BaBar's XY coordinate via the charged channel, while candidates with momenta ≥ 2.50 GeV/c are rejected if their Λ daughter decayed more than 12 cm in BaBar's XY coordinate via the charged channel.) Table 6.4 gives the relative fractions of cascade hyperons that were accepted/rejected by the decay distance cuts.

The Ξ^- and $\overline{\Xi}^+$ spectra displayed in chapter 7 include all the above corrections as well as the Λ and Ξ branching fraction corrections.

6.4 Ω^- signal extraction

The composition, efficiency matrix construction, and signal fits for the Ω^- and $\overline{\Omega}^+$ hyperons are similar in outline to those for Ξ^- and $\overline{\Xi}^+$. Therefore, we will focus only on the procedural differences between the Ω and the cascade hyperons.

Candidate CM momentum Bins (GeV/c)	Fraction of MC $\Xi^- + \overline{\Xi}^+$ (Rejected by decay distance cut / Accepted)
0.25-0.50	5.0554
0.50-0.75	4.8704
0.75-1.00	7.0014
1.00-1.25	10.8465
1.25-1.50	13.8583
1.50-1.75	17.1996
1.75-2.00	23.0974
2.00-2.25	30.1227
2.25-2.50	38.3712
2.50-2.75	0.4615
2.75-3.00	0.5495
3.00-3.25	0.6382
3.25-3.50	0.7169
3.50-3.75	0.7232
3.75-4.00	0.9888
4.00-5.10	2.3056

Table 6.4: Fractions of MC Ξ^- and $\overline{\Xi}^+$ rejected by decay distance cut relative to number of accepted candidates.

The only difference in the composition step between the Ω^- and the Ξ^- lies in the mass hypothesis of the additional charged track which is vertexed with the Λ candidate; whereas the Ξ^- requires a π^- mass hypothesis for the additional track, the Ω^- requires a k^- mass hypothesis for that same track.

6.4.1 Cuts against Ω candidates

The following list details the cuts (proton PID on one of the Λ^0 daughters is assumed throughout) that we use to reduce the background which would otherwise completely drown out the Omega signal:

- The Ω^- candidate's dXY (XY distance between decay vertex and event vertex) must be > 2.0 mm.
- The Ω^- candidate's XY decay vertex must be ≤ 10.0 cm in BaBar coordinates.
- The Λ^0 daughter's dXY also must be > 2.0 mm and < 12.0 cm.
- The proton, pion, and kaon charge combination must be consistent with either the $\Omega^- \rightarrow \Lambda^0 + k^- \rightarrow p^+ + \pi^- + k^-$ or with $\overline{\Omega}^+ \rightarrow \overline{\Lambda}^0 + k^+ \rightarrow p^- + \pi^+ + k^+$.
- The k^- track must not have been identified either as a proton or a pion by the Global Likelihood based PID selector.
- A Λ^0 mass window with total width of 5 MeV/c^2 (from 1.1132 GeV/c^2 to 1.1182 GeV/c^2.)

- Ω^- candidates whose Λ^0 daughters decayed inside the PEP-II beampipe must have pointing angles (recall these are cosine angles) > 0.9986 (this corresponds to an angle of $\sim 3.03^o$;) Ω^- candidates whose Λ^0 daughters decayed outside the PEP-II beampipe must have pointing angles > 0.94 (about $\sim 19.9^o$.)

- The Λ^0 daughter's protonic daughter must have 12 or more DCH hits.

- The Λ^0 daughter's mmomentum vector, projected backward towards the Ω^- decay vertex, must have a DOCA of less than 3 mm. from that vertex.

- The k^- used to compose the Ω^- must have a DOCA of less than 3 mm. from the Ω^- decay vertex (the DOCA must be measured using the k^- track quantities resident prior to fitting the π track as coming from the decay vertex.)

- The Λ^0 daughter's decay vertex in the XY plane must have a distance at least 2 mm in excess of the Ω^-'s radial decay distance.

We'll refer to the first five of the above listed cuts as the 'preliminary' cuts on Ω^- candidates. The effect of the preliminary cuts on the reconstructed $\Omega^-+\overline{\Omega}^+$ invariant mass distribution (Run2 off-peak data) is displayed in Figure 6.25. Note that a major difference between the Ξ^- and Ω^- selection criteria results from the fact that, due to the extremely small production cross section of the Ω^- hyperon, it is impractical to restrict our sample to those particles where the Λ^0 daughters also have decayed inside the PEP-II beampipe (as was done for Ξ^- candidates with CM momenta < 2.50 GeV/c.) This necessarily increases the systematic errors associated with the sample of constituent tracks.

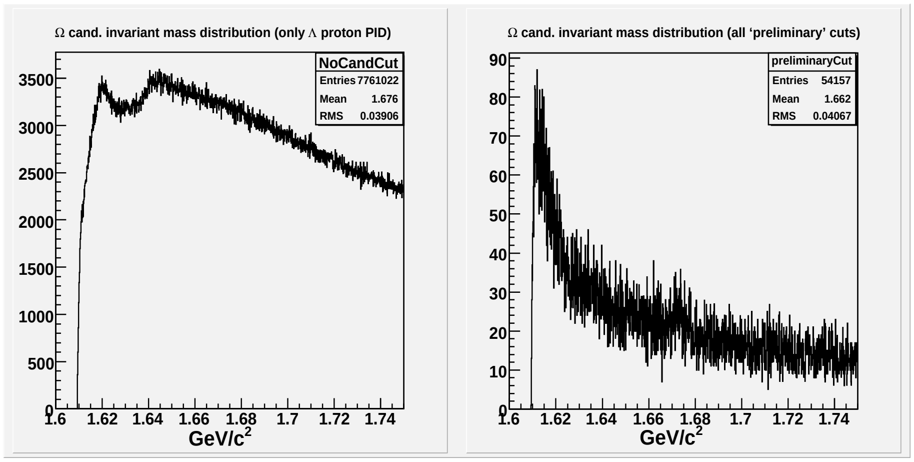


Figure 6.25: (Left) Invariant mass distribution of reconstructed $\Omega^-+\overline{\Omega}^+$ candidates with only Λ proton PID cut (all off-peak Run2 data.) (Right) Same distribution after all Ω preliminary cuts.

The individual effects of the additional cuts on the invariant mass distribution of candidates that survived the preliminary set of cuts are displayed in the following set of Figures. Figure 6.26 exhibits the effects of the three dimensional DOCA cuts of the candidates' Λ and k daughters. Figure 6.27 shows the effects of the cut on the difference between Ω^-'s dXY (cut on the XY displacement of a candidate's decay vertex from the primary event vertex) and its Λ^0 daughter's dXY, and the effects of the Ω^- pointing angle cut. Figure 6.28 lays out the result after the application of the DCH hit cut on the candidates' Λ^0 daughters. And finally, Figure 6.29 displays the combined result of all cuts on the Ω^- and $\overline{\Omega}^+$ candidates separately.

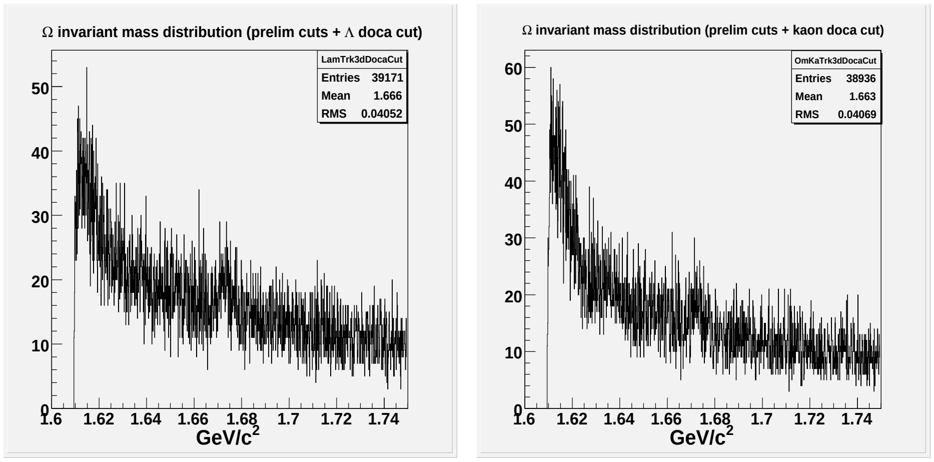


Figure 6.26: (Left) Invariant mass distribution of Ω^- candidates after cutting against Λ^0 daughter's DOCA to reconstructed Ω^- vertex. (Right) Invariant mass distribution of Ω^- candidates after cutting against k^- track DOCA to reconstructed Ω^- vertex.

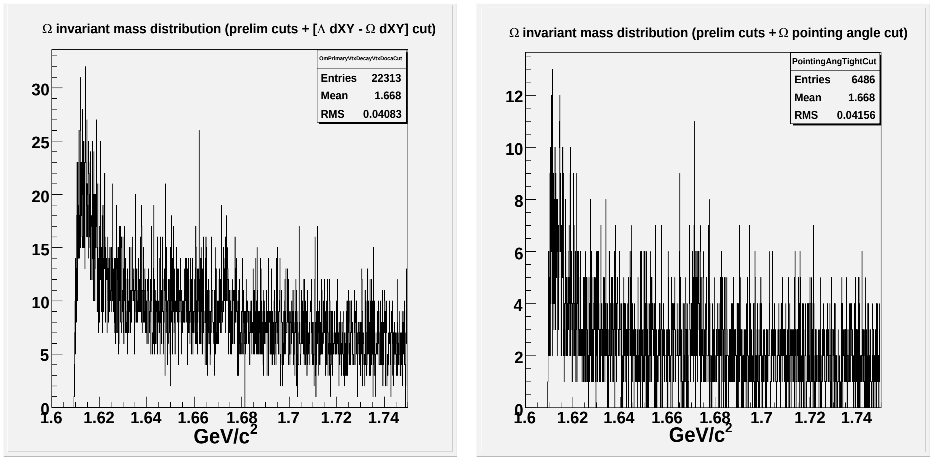


Figure 6.27: (Left) Invariant mass distribution of Ω^- candidates after cutting against the difference between Ω^-'s dXY and its Λ^0 daughter's dXY. (Right) Invariant mass distribution of Ω^- candidates after cutting against the candidate's pointing angle.

6.4.2 Ω efficiency matrix construction

The cuts listed in the previous section are applied to over 200 million MC udsc events (equivalent to an integrated luminosity of $\sim 60\ fb^{-1}$.) MC Ω^- and $\overline{\Omega}^+$ candidates that are associated

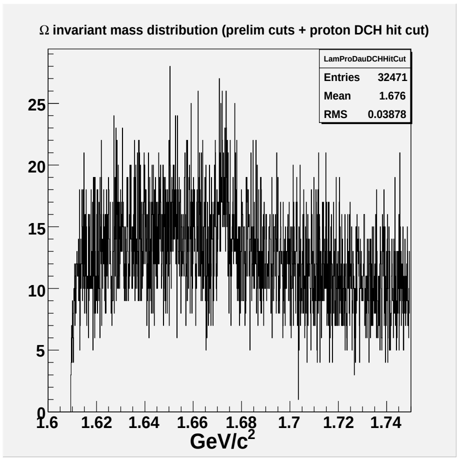


Figure 6.28: Invariant mass distribution of $\varOmega^{-}$ candidates after cutting against $\varLambda^{0}$'s protonic daughters DCH hits.

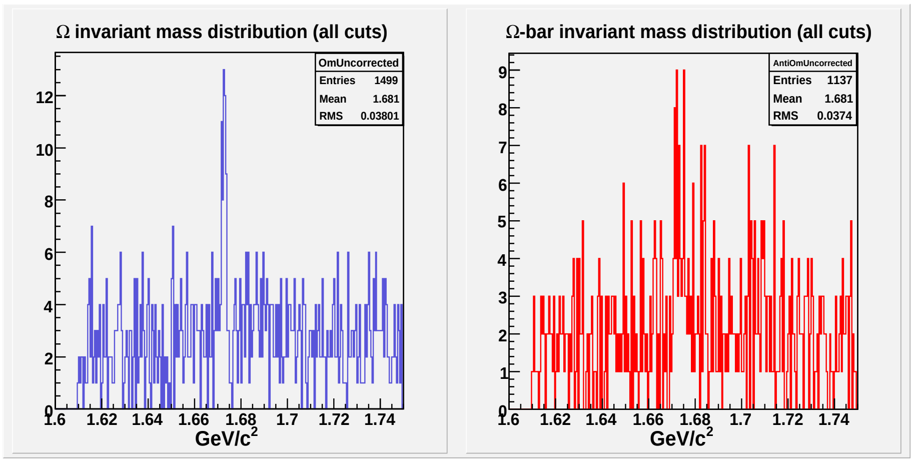


Figure 6.29: (Left) Invariant mass distribution of $\varOmega^{-}$ candidates after all cuts. (Right) Invariant mass distribution of $\overline{\varOmega}^{+}$ candidates after all cuts.

with true generator particles (again, 'truth' denotes each of the iterim and final state daughters has been identified with a particle of the correct species and momentum which came from the correct mother particle) are saved to separate two-dimensional histograms in the particle's labAngle and lab momentum variables. Incidentally, we use histograms with 50 uniform widthed bins in each of the variables. Figure 6.30 displays the reconstructed and truth-associated Ω^- and $\overline{\Omega}^+$ 2-d histograms.

All Omegas from the Monte Carlo generator, and which decayed via the appropriate decay channels and within the range accepted by the Omega decay distance cut, are also saved to similarly-sized two-dimensional histograms. Figure 6.31 contains histograms showing the distributions of generator Ω^- and $\overline{\Omega}^+$ in the labAngle and lab momentum variables.

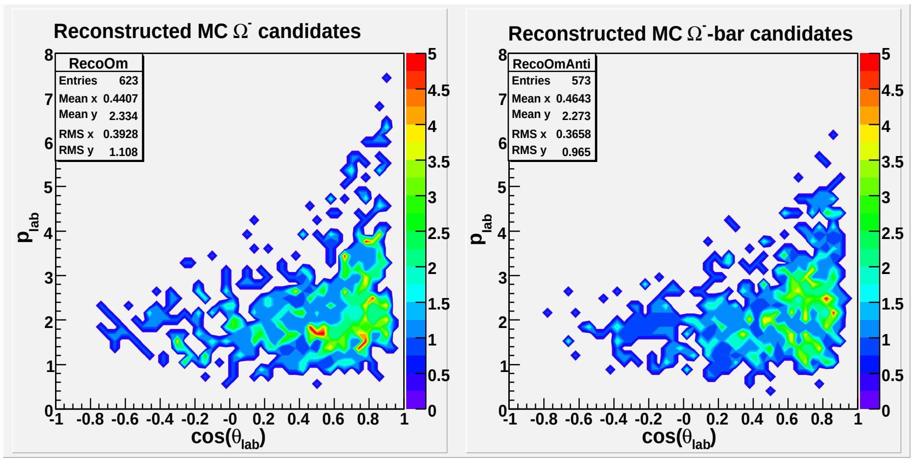


Figure 6.30: Distribution of reconstructed MC Ω^- and $\overline{\Omega}^+$ candidates in labAngle and lab momentum variables

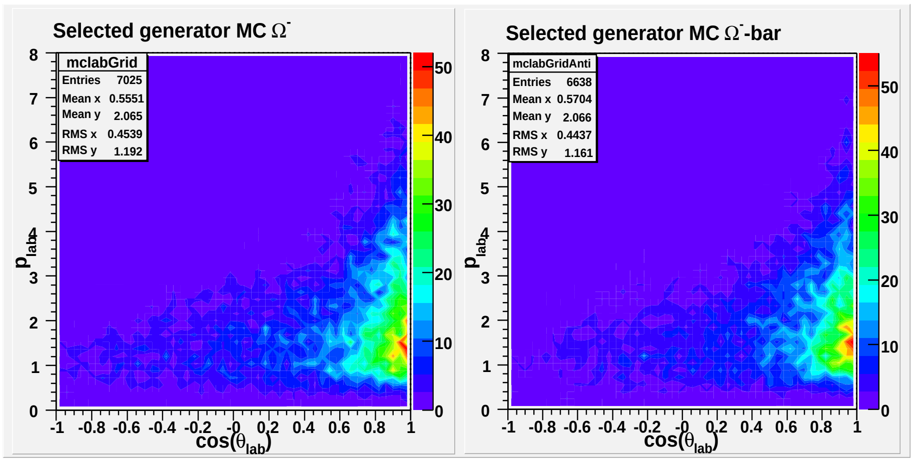


Figure 6.31: Distribution of selected generator MC Ω^- and $\overline{\Omega}^+$ candidates in labAngle and lab momentum variables.

The construction of the efficiency matrices again involves the bin-by-bin division of the number of reconstructed/truth-associated candidates by the number of generator particles; the construction process begins only after the matrices have been reweighted by the relative proportion of event charged-track multiplicity ratios obtained from off-peak data.

Due to the relative rarity of Omega events in our selected sample of candidates, we immediately increased the size of the two-dimensional bins by going from grids of 50 × 50 bins to grids of 10 × 10 bins (so the labAngle width of each bin is 0.2 while the lab momentum width of each bin is 0.8 GeV/c.) No additional smoothing is performed on the bin efficiencies except comparison with and averaging using each bin's nearest neighbors. The efficiency construction process and statistical error estimation are otherwise identical to those applied for the cascade hyperons.

The efficiency matrices and the associated relative statistical uncertainty in each bin (using binomial errors) are displayed in Figures 6.32 and 6.33.

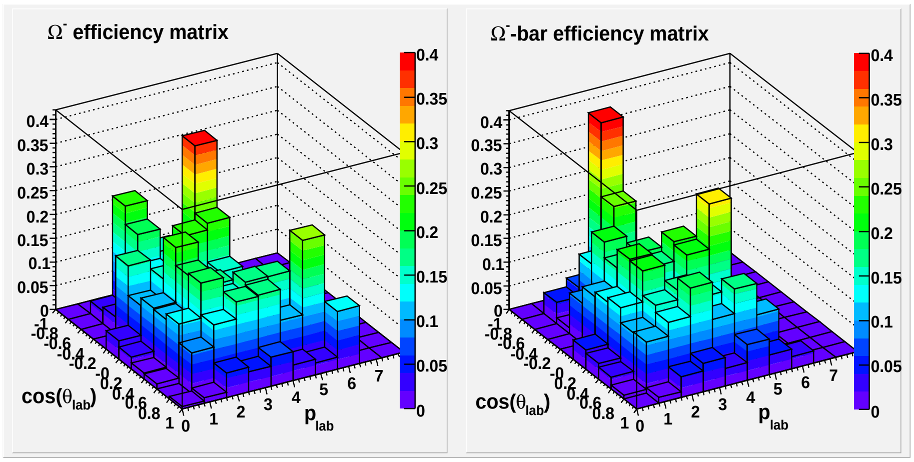


Figure 6.32: Two-dimensional efficiency matrices constructed from MC Ω^- and $\overline{\Omega}^+$.

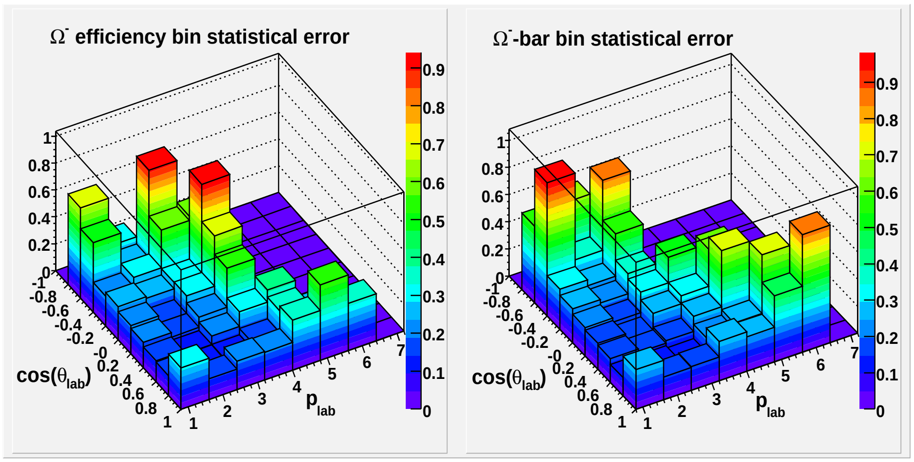


Figure 6.33: Bin statistical erros associated with Ω^- and $\overline{\Omega}^+$ efficiency matrices.

6.4.3 Ω efficiency correction and signal fits

The Ω^- and $\overline{\Omega}^+$ data signals again are obtained using BaBar's entire off-peak data set from Runs 1 through 3 (Run 4 data have been excluded due to issues to be discussed in the next chapter.) The same cuts and composition routines that were used for the MC efficiency computations again were applied to the data runs. Efficiency-correction using the 2-d efficiency matrices is applied on a candidate-by-candidate basis while the tracking corrections for the final state proton and pions are applied at this stage. Note that since we have as yet not obtained non-PID'ed kaon tracking correction factors, we performed no kaon track corrections in our analysis; in any case, this issue is complicated by the fact that most all of our accepted Omega candidates have Λ daughters that decayed outside the PEP-II beampipe. Finally, we note that we have imposed no explicit track hit cuts on the kaon candidate tracks; our studies of the pion tracks also revealed that no corrections were needed when one use very loose cuts for pion tracks and we can expect this to apply to the kaon tracks as well. In fact, since the MC/data differences in our PID cut on the kaon will overwhelm the tracking differences, if any, there appears to be no need to assign a separate systematic uncertainty due to kaon tracking; a large kaon PID error should take care of tracking Data/MC differences as well.

The efficiency-corrected number of accepted Ω^- and $\overline{\Omega}^+$ are plotted in accordance with the candidates' invariant mass and binned with respect to their CM momenta in bins of 0.50 GeV/c bin-width from 0.0 GeV/c up to 5 GeV/c (practically, we have signals only up to 4.0 GeV/c.)

The momentum-binned Ω^- and $\overline{\Omega}^+$ invariant mass plots are fitted using either a first or second-order polynomial background function (depending on the level of difficulty in fitting the oft-jagged backgrounds) and a signal function that convolutes a Breit-Wiger function with a gaussian function. Again, the actual fitting is performed using ROOT's implementation of the MINUIT package.

Figure 6.34 displays the fitted Omega hyperon signal and background invariant mass plots for the case where particles and antiparticles have been compacted into the same histograms.

The shifts in the preliminary fitted signal invariant mass (the plots displayed in Figure 6.34 consists of refined fits where PDG mass constrains have been applied in certain bins after comparisons with other bins and with fits using mass plots from separate particle and antiparticle subsamples) and the variations in the widths are contained in Figure 6.35.

Cuts against the Omega efficiency matrices again were instituted to eliminate bins with very small efficiencies or very large statistical errors. These cuts include:

- Ω^- or $\overline{\Omega}^+$ labAngle $<$ -0.80.
- Ω^- or $\overline{\Omega}^+$ labAngle $>$ +0.80.
- All efficiency bins where the Ω^- or $\overline{\Omega}^+$ candidate lab momentum $<$ 0.80 GeV/c or $\geq$ 5.00 GeV/c are excluded.
- All efficiency bins where the efficiency is $\leq$ 4.0% are excluded for candidates with momentum $>$ 0.5 GeV/c, while all efficiency bins with efficiency $\leq$ 1.0% are exluded for candidates with momentum $\leq$ 0.5 GeV/c.
- All efficiency bins where the efficiency statistical error is $\geq$ 35.0% are excluded for candidates with momentum $>$ 0.5 GeV/c, while all efficiency bins with statistical error $\geq$ 50.0% are exluded for candidates with momentum $\leq$ 0.5 GeV/c.

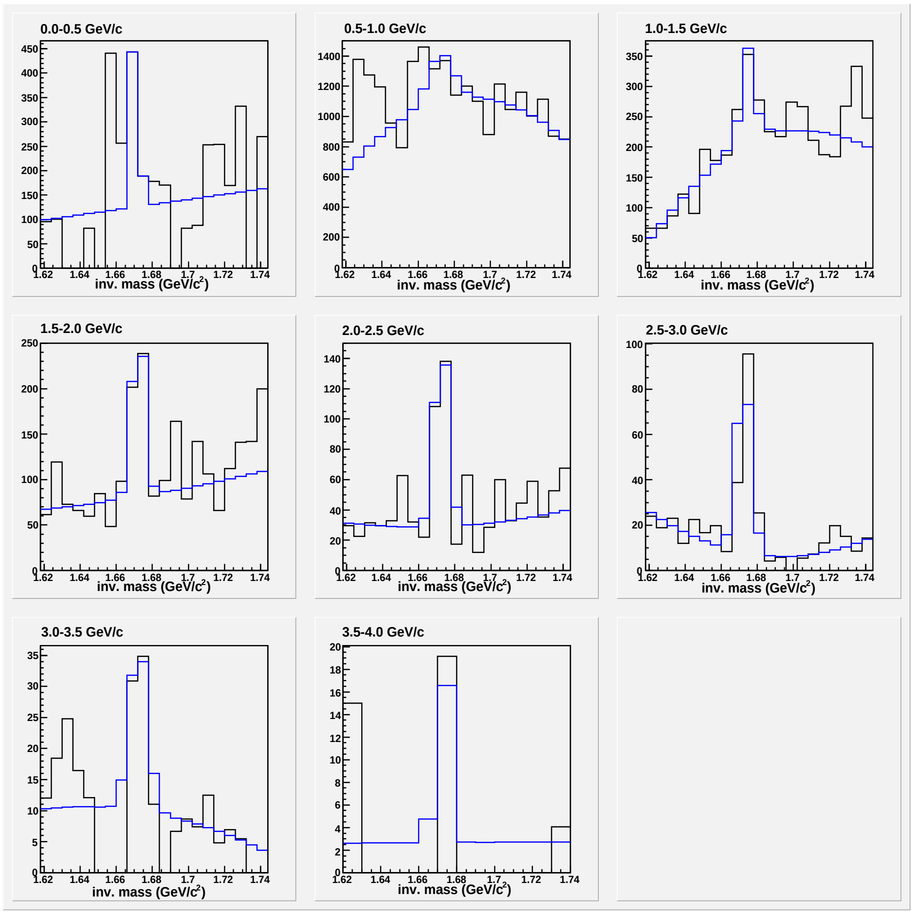


Figure 6.34: Fitted and efficiency-corrected Ω^- and $\overline{\Omega}^+$ signals.

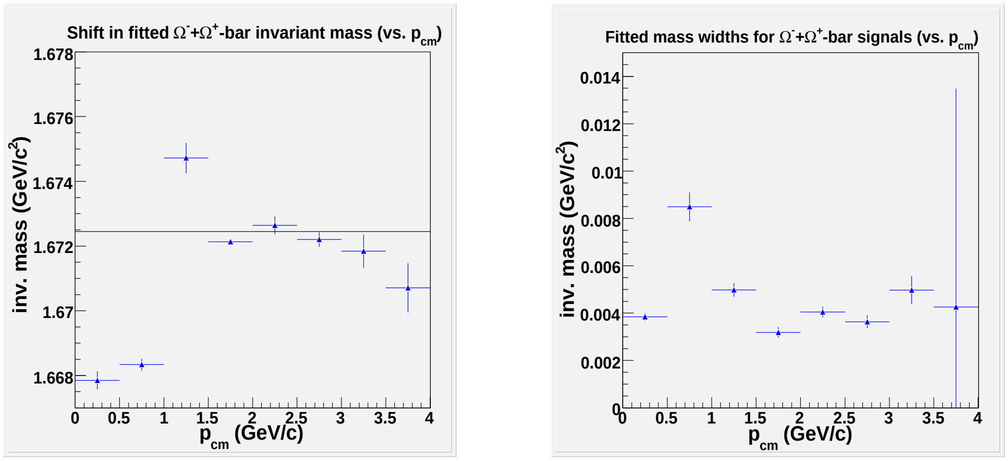


Figure 6.35: (Left) Shift in $\Omega^-+\overline{\Omega}^+$ fitted mass versus CM momentum (note the black line represents the PDG value for the Ω. (Right) Fitted widths of $\Omega^-+\overline{\Omega}^+$ signals versus CM momentum. The quoted widths are one-half the full-widths at half-maximum from single gaussian fits of the signals.

Candidate CM momentum Bins (GeV/c)	Fraction of MC $\Omega^- + \overline{\Omega}^+$ (Rejected by efficiency cuts / Accepted)
0.00-0.50	0.3191
0.50-1.00	0.2691
1.00-1.50	0.8384
1.50-2.00	0.8289
2.00-2.50	0.8561
2.50-3.00	0.5726
3.00-3.50	0.9615
3.50-4.00	6.4444

Table 6.5: Fractions of MC Ω^- and $\overline{\Omega}^+$ rejected/accepted by cuts on efficiency tables.

Candidate CM momentum Bins (GeV/c)	Fraction of MC $\Omega^- + \overline{\Omega}^+$ (Rejected by decay distance cut / Accepted)
0.00-0.50	0.0
0.50-1.00	0.0
1.00-1.50	0.0101
1.50-2.00	0.0374
2.00-2.50	0.0863
2.50-3.00	0.0684
3.00-3.50	0.1731
3.50-4.00	0.3333

Table 6.6: Fractions of MC Ω^- and $\overline{\Omega}^+$ rejected by decay distance cut relative to number of accepted candidates.

We must correct for the number of candidates rejected by these efficiency-table based cuts using MC. Table 6.5 lists the relative number of generator Ω^- and $\overline{\Omega}^+$ rejected and accepted by the efficiency-based cuts. We also need to correct for candidates that were rejected by our decay distance cuts (dXY < 2 mm *and* XY decay vertices ≥ 10 cm in BaBar coordinates.) Table 6.6 gives the fractions of Omegas that were accepted/rejected by the decay distance cuts.

Chapter 7

Hyperon spectra, errors, and comparisons

We now shall consider the momenta spectra of the three hyperons obtained using the methods described in chapter six. Integration of the spectral bins, after converting from cross sections to number of events and dividing by the total estimated number of multihadronic events, then provides us with estimates of the visible production rate for each of the particles. Further extrapolations of bins which have been excluded from the data analysis, if needed, then provide the continuum inclusive production rates for the particles.

7.1 Λ^0 hyperon results

We use the signal fits of the Lambda (particles plus antiparticles) invariant mass plots in CM momentum bins (see Figure 6.10) to construct the spectrum. Recall that for the very first CM momentum bin (CM momentum 0.00-0.25 GeV/c,) as well as for the last bin (4.25-5.15 GeV/c,) the signals were obtained with the *additional* PID cut. This additional cut was necessary in order to extract any signals at all from the backgrounds resident in these bins; kinematic cuts on our Lambda efficiency maps, when combined with the fact that true Lambdas with CM momenta less than 0.25 GeV/c are constrained to a region with very forward labAngle values and low lab momentum (see Figure 6.9) and that there are very few high momentum Lambdas in absolute terms, result in low signal to noise ratios for these bins without the application of a PID cut.

7.1.1 Preliminary Λ^0 and $\overline{\Lambda}^0$ spectra and statistical uncertainties

The preliminary Λ and $\overline{\Lambda}^0$ spectrum obtained using the fits described in §6.2 herein is given in terms of the total number of particles + antiparticles per CM momentum bin. This can be converted to a fractional (fraction of total hadronic cross section) cross section per unit momentum fraction using the following expression:

$$\frac{1}{\sigma_{\mathrm{h}}}\left(\frac{\Delta\sigma}{\Delta x_p}\right) = \frac{N_{bin}}{N_{\mathrm{MH}}}\left(\frac{1}{\mathcal{F}}\right)\left(\frac{1}{\Delta x_p}\right) \tag{7.1}$$

where N_{bin} is the total number of candidates obtained from the signal fit and after application of all corrections, N_{MH} is the total estimated number of hadronic events in our sample,

$\mathcal{F}$ is the Lambda charged fraction (the fraction of Lambdas that decay via the charged decay channel,) and $\Delta x_p \equiv (\Delta p_{bin}/5.15)$ is the bin width of the particle's CM momentum bin expressed as a momentum fraction of the maximal available kinetic energy to a Lambda or anti-Lambda particle.

We estimate the statistical uncertainty associated with the fully efficiency-corrected total number of Λ and $\overline{\Lambda}^0$ in a particular CM bin in the following manner:

- Since we're dealing with the process of counting the number of discrete candidates in some given bin, we assume that the statistical variations in the number of detected candidates follows a Poisson distribution. At this stage, we may construct a confidence interval for this distribution following the procedure described in §9.4 in [35]. However, since most bins in our case have large numbers of signal candidates and since it is well known that the gaussian distribution is the large number limit of the Poisson distribution, we elect to use the simplest possible statistical error estimate. Thus we take the one standard deviation Poisson estimate for the bin number variation ($\sqrt{N_{bin}}$) as the approximate 68% gaussian confidence interval.

- The number of signal candidates in a bin was efficiency-corrected using the efficiency matrices. The actual error estimate need to take into account such efficiency corrections. For the very first bin (0.00-0.25 GeV/c,) the average efficiency (only about 3 bins from the efficiency matrices contribute to the first CM momentum bin) is about 3%; for all other CM momentum bins, we use an average efficiency figure of 28.9% (obtained from an overall fit of all efficiency-corrected signal candidates and an overall fit of all *uncorrected* signal candidates. This average efficiency figure is dominated by bins with lower momenta (near 1 GeV/c.) Bins with higher momenta tend to have higher efficiencies (though much fewer candidates;) while this may be viewed as an overestimate of the statistical errors for higher momentum bins, such a procedure nevertheless has the desirable effect of increasing our approximate confidence intervals for bins with fewer candidates.

- During the calculation of the bin statistical uncertainties, we also take into account the correction ratios due to our kinematic cuts and the decay-distance cut detailed in Tables 6.1 and 6.2. This is accomplished by increasing the error sizes (dividing by the correction ratios.)

The bin-by-bin results for the number of fitted signal particles and antiparticles (after partial/full efficiency corrections) and the associated bin statistical uncertainties are given in Table 7.1.

In Table 7.1, the 'fitted' number of particles and antiparticles refer to the fitted bin signals after application of the 2-d efficiency matrix corrections *only*, while the 'fully corrrected' numbers refer to the total estimated bin numbers after all corrections, including corrections for all kinematic cuts and the decay distance cut, have been applied. Also note that the total fitted number of candidates that we obtained without any momentum binning, and using only efficiency matrix corrections, is 660,636 [24]. This is within 1.13% of our total using binned fits. The estimated total statistical uncertainty given in Table 7.1 was obtained by summing the individual statistical errors from the various bins by quadrature; it is appropriate to sum

---

[24] Fit of the unbinned signal without *any* efficiency corrections gives an estimate of 191,187.

Candidate p_{CM} bins (GeV/c)	Fitted $\Lambda^0 + \overline{\Lambda}^0$	Fully corrected	% bin stat. uncertainty
0.00-0.25	612	86,183	23.34%
0.25-0.50	32,814	267,629	1.03%
0.50-0.75	128,762	503,075	0.52%
0.75-1.00	138,150	576,245	0.50%
1.00-1.25	107,656	484,169	0.57%
1.25-1.50	81,537	416,144	0.65%
1.50-1.75	58,732	353,032	0.77%
1.75-2.00	41,481	282,363	0.91%
2.00-2.25	26,213	198,672	1.15%
2.25-2.50	16,612	137,943	1.44%
2.50-2.75	9,122	82,772	1.95%
2.75-3.00	5,559	53,004	2.49%
3.00-3.25	3,010	30,675	3.39%
3.25-3.50	1,614	17,442	4.63%
3.50-3.75	890	10,374	6.24%
3.75-4.00	267	3,093	11.38%
4.00-4.25	72	853	21.92%
4.25-5.15	171	2,729	14.22%
$\sum$ (0.00-5.15)	653,274	3,506,397	0.64%

Table 7.1: Fitted numbers of Λ^0 and $\overline{\Lambda}^0$ and associated statistical uncertainties in p_{CM} bins; note that the numbers in the second column entitled "Fitted $\Lambda^0 + \overline{\Lambda}^0$" are the fitted bin numbers after corrections using *only* the 2-d efficiency matrices described in §6.2.2.

the statistical uncertainties by quadrature since the individual bin statistics are generally not correlated with each other and each bin contributes independently to the final numerical sum.

Figure 7.1 is a graphical representation of our above results after converting the absolute numbers of fully efficiency-corrected particles and antiparticles to fractional hadronic cross sections; the associated bin statistical errors have been plotted separately in order to render them visible.

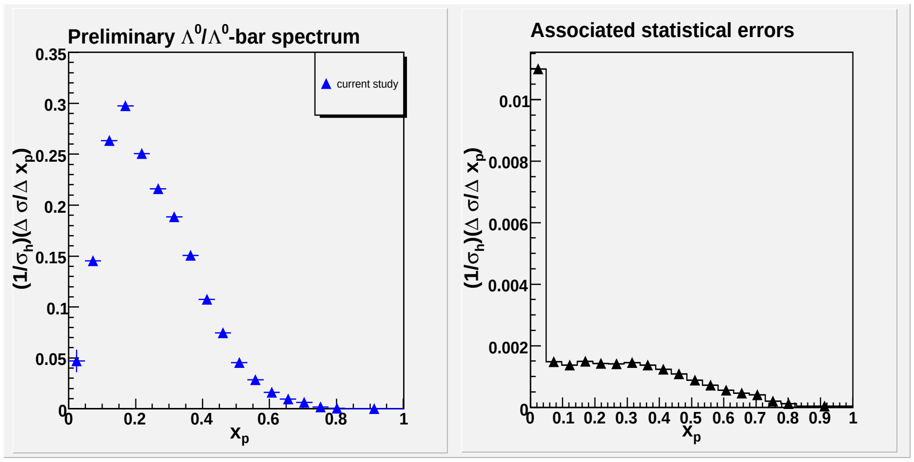


Figure 7.1: (Left) Preliminary spectrum of $\mathit{\Lambda}^0$ and $\overline{\mathit{\Lambda}}^0$. (Right) Associated bin statistical errors.

7.1.2 Propagated tracking and efficiency uncertainties

The propagated tracking correction errors (both proton and pion DCH errors as well as pion SVT errors) obtained during our construction of the Lambda and anti-Lambda efficiency matrices were associated with each of the efficiency bins; the Monte Carlo statistical errors were similarly associated with each of the bins (see §6.2.2 herein.) During our efficiency correction procedure, whenever we accessed the efficiency information from a particular bin we simultaneously access such bin error information.

The DCH correction error, SVT correction error, and efficiency statistical error are individually added [25] to the relevant bin efficiency fraction in order to obtain alternate bin efficiency values which in turn are employed to efficiency-correct the detected candidate. Thus, after correcting all detected data candidates using our efficiency matrices, we obtain, in addition to our primary efficiency-corrected mass signal plots (in candidate CM momentum bins,) alternative mass signal plots which represent variations due to DCH tracking uncertainties, SVT tracking uncertainties, and the efficiency bin statistical uncertainties. These alternative plots are fitted in the same manner as described in §6.2.3 herein. We take the difference between the fitted signals from §7.1.1 and the fitted numbers from our alternative p_{CM}-binned mass plots as the individual systematic error contributions from tracking corrections and from efficiency-matrix statistics.

---

[25] Recall that the tracking efficiency errors themselves were expressed as efficiency quantities (see §6.2.2;) on the other hand the efficiency statistical error is expressed as a percent error so in this case we 'add' this error to the efficiency by multiplying the bin efficiency by (1.0 + percent bin error).

Cand p_{cm} (GeV/c)	% errors DCH trk corr.	% errors SVT trk corr.	% errors eff. stat.
0.00-0.25	-	-	6.86%
0.25-0.50	0.44%	0.20%	1.64%
0.50-0.75	1.73%	0.14%	1.84%
0.75-1.00	1.64%	0.12%	2.28%
1.00-1.25	1.64%	0.14%	2.49%
1.25-1.50	1.72%	0.12%	2.69%
1.50-1.75	0.92%	2.61%	0.37%
1.75-2.00	1.90%	3.51%	3.18%
2.00-2.25	2.31%	0.12%	3.79%
2.25-2.50	2.83%	0.12%	4.63%
2.50-2.75	3.36%	0.11%	5.48%
2.75-3.00	5.25%	0.12%	8.13%
3.00-3.25	4.13%	0.12%	7.90%
3.25-3.50	4.29%	0.12%	5.98%
3.50-3.75	3.37%	0.12%	4.45%
3.75-4.00	0.87%	0.16%	6.11%
4.00-4.25	4.05%	4.00%	2.67%
4.25-5.15	-	-	82.84%

Table 7.2: Systematic uncertainties in p_{CM} bins due to tracking and efficiency matrices' bin statsitics. All uncertainties are reported as percentages of the total estimated numbers of signal particles in that bin as reported in 7.1.

At this point, we note that the tracking correction uncertainty is at its base a binomial statistical error and is propagated to the Lambda spectrum in accordance with the Lambda decay daughter's momentum; hence the tracking uncertainties' contributions to the Lambda error is correlated with Lambda momentum bins. However, in any particular Lambda momentum bin, many different daughter momenta values contribute more or less incoherently to the error so that the relative error in each Lambda momentum bin should be approximately the same. Thus this error source can be treated as a so-called normalization error. A normalization error does not affect the Lambda spectrum shape and is simply added as a component to the total Lambda rate's errors.

The uncertainties due to MC efficiency grid's bin statistics on the other hand should be entirely independent of the candidates' momenta. This source of error can be added in quadrature when computing the total error of the Lambda production rate.

The DCH and SVT tracking errors as well as the MC efficiency map statistical errors for each Lambda momentum bin is listed in Table 7.2. Note that for the first and last bins, only the combined errors are reported due to low statistics; since the statistical errors are expected to dominate in these bins, we quote the combined errors as statistical errors. Further note that the uncertainties due to efficiency bin statistics have been increased by a factor of 1.55 during the efficiency-correction process for reasons to be discussed in the next subsection.

All the reported percentage errors in Table 7.2 are the actual calculated percentage differences between the fully corrected numbers of particles and anti-particles in the relevant bins (see Table 7.1) and the alternative fitted numbers for the same bins after adding the specified uncertainty to the efficiency fractions. Note that the relative SVT tracking errors are fairly constant except for two bins between 1.50 GeV/c and 2.00 GeV/c and for the 4.00-4.25 GeV/c bin; the higher errors in these bins may be due entirely to fitting problems. On the other hand, the DCH tracking errors appear to be fairly constant at low momentum and then gradually increase above 2.00 GeV/c; this does not entirely align with our expectation that this should be a normalization error. Nevertheless, since the variations in these errors do not exceed the 1.5% level until we reach momenta over 2.50 GeV/c, and since the relative cross section represented by the momentum region over 2.50 GeV/c is small, we shall continue to treat the DCH tracking error as a normalization error.

Figure 7.2 is a graphical representation of the combined effect (per individual momentum bin) of the errors reported in Table 7.2; the range of the uncertainties is given as an "error band". Note that the first and last bins have been omitted from this figure.

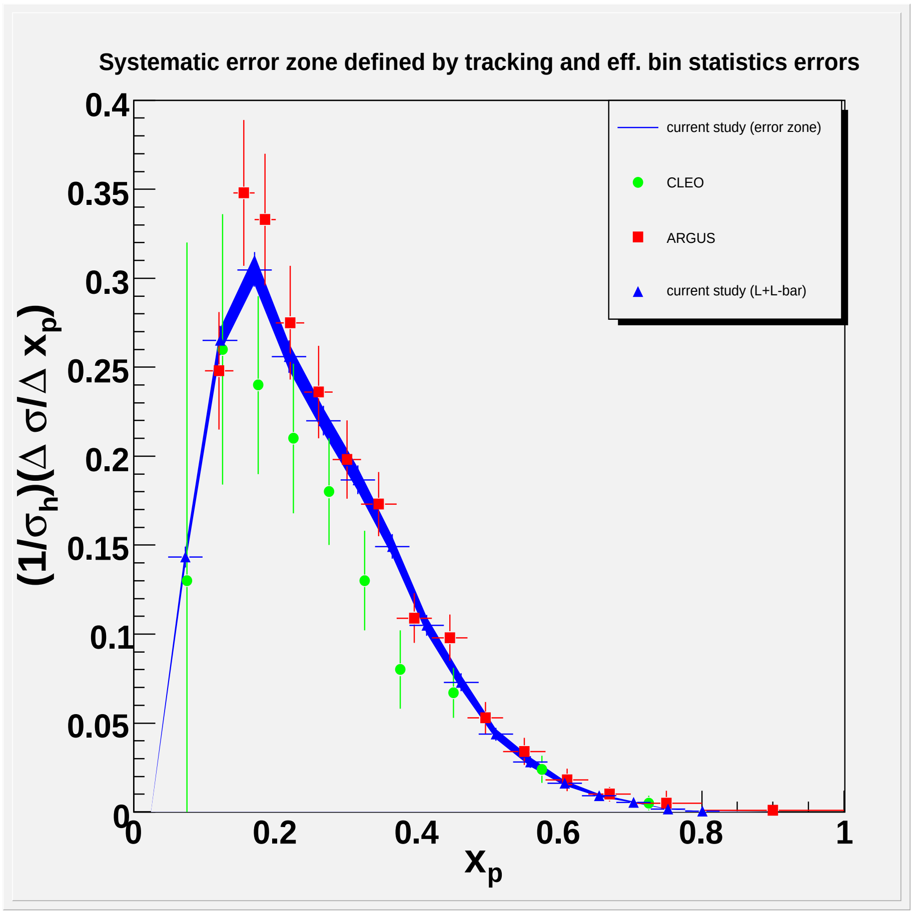


Figure 7.2: Zone of 'perfectly' correlated errors due to tracking corrections combined with the independent Λ^0 efficiency statistics errors. CLEO and ARGUS collaborations' results for Λ^0 have been superimposed for comparison purposes.

7.1.3 MC statistical error consistency check

We want to cross check the assigned bin statistical uncertainties associated with the Lambda efficiency matrices.

One way to perform such a cross check in a timely manner is to generate a set of MC generator Lambdas and anti-Lambdas that is independent of the set of MC particles that was employed to construct the efficiency matrices. We then quickly simulate the detector's effects on the new sample of particles and antiparticles by multiplying each generator particle by its efficiency obtained from our original efficiency matrices. Next, we construct a new set of efficiency matrices using the fractions of 'detected' particles and antiparticles from the new sample and compare them with the original sample. With the new efficiency matrices in hand, we can compare the efficiency bin contents of the new and old matrices and compute the 'pull' for any bin using the following expression:

$$\text{pull} \equiv \frac{\epsilon_{new} - \epsilon_{old}}{\sigma_{stat}} \tag{7.2}$$

where ϵ_{new} and ϵ_{old} are the bin efficiencies using the new and old efficiency matrices respectively, and where σ_{stat} is the associated bin statistical uncertainty from our original efficiency matrices. We can plot the pull values calculated from the various bins; if the bin efficiencies and the bin statistical errors have been correctly estimated by our matrix construction proceudre then the pull values should be normally distributed with a standard deviation of one and centered on zero.

We performed the above analysis by generating 1,000,000 new MC events using the default Jetset 7.4 (Pythia 6.2) generator which employs a default production rate of about 0.14 Lambda per hadronic event; this gave us about 44,000 generator Lambdas (not including anti-Lambdas) that decayed via the charged channel. We used a gaussian function to fit the resulting pull distribution. The fit reveals that we may be underestimating bin statistical uncertainties by a factor of 1.55 and that the mean efficiency may also be slightly lower than the 'true' efficiency by about 60% of a standard deviation (i.e. the bin efficiency uncertainties.) These numbers are not unexpected in that during the efficiency matrix construction procedure, we're integrating over bins where, on average, the efficiency gradient (measured from the center bin) is negative so that we tend to slightly underestimate the bin efficiency and statitical uncertainties (using binomial bin errors.)

In light of the above, we've increased all statistical errors during efficiency-correction procedure by a factor of 1.55; this reduces the efficiency underestimation to about 40% of a standard devation. The resulting pull distribution (and the gaussian fit of the distribution) after application of the 1.55 multiplicative factor is presented in Figure 7.3.

7.1.4 Systematic errors from signal fits and momentum shifts

We need to assign errors for the binned signal fitting procedure as well as for the possibility that a reconstructed particle with some true momentum value is incorrectly assigned to a neighboring momentum (this is possible due to detector inefficiencies, incorrect energy-loss corrections at the tracking reconstruction stage, etc.)

In so far as the errors due to momentum bin shifts can be corrected using the Monte Carlo, there's no need to perform extra work; the definition of our efficiency grids (see infra. §6.2.2) effectively corrects for this shift perfectly. If a reconstructed candidate that has been found in a certain momentum bin really should be located in a neighboring bin, then the efficiencies of

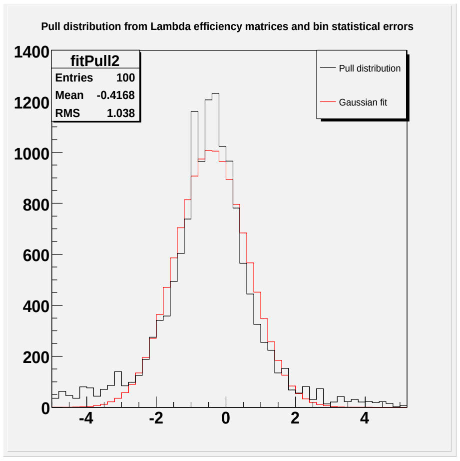


Figure 7.3: Pull distribution calculated from Lambda efficiency matrices and the estimated efficiency bin statistical uncertainties (with 1.55 multiplicative factor.)

both bins have already taken this shift into account and the true MC number in the bins will be obtained during the efficiency correction procedure. Nevertheless, there may be residual momentum shift errors due to differences in the momentum resolution between MC and data candidates. Since the invariant masses of the particle candidates were calculated from their energies and momenta, the momentum resolution should be proportional to the invaraiant mass resolution. We can compare the fitted signal mass widths from data and Monte Carlo. In chapter 6, Figure 6.11 displayed the fitted signal widths from our off-peak data; Figure 7.4 displays the equivalent plot for our Monte Carlo (after running the reconstruction code over our MC sample.)

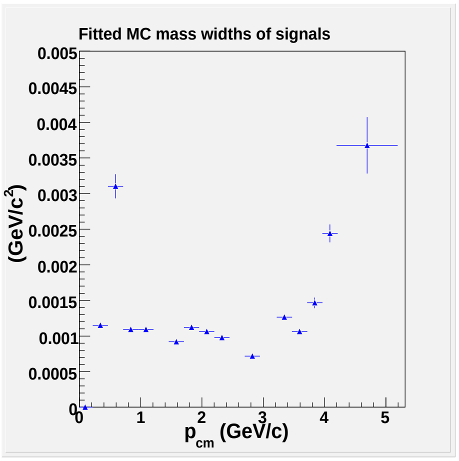


Figure 7.4: Fitted widths of Monte Carlo Λ^0+$\overline{\Lambda}^0$ mass signals versus CM momentum.

The Monte Carlo signal mass widths generally model the data well; this is particularly true in the momentum region from about 0.50 GeV/c up to about 3.00 GeV/c where both the MC and data have widths in the range 1.0-1.2 MeV/c^2 (except for bins with relatively poor MC fits.) This gives us some confidence that the momentum resolution in the data is well-modeled by the MC. Figure 7.5 exhibits the percentages of reconstructed Lambda candidates that shifts to a neighboring lab momentum bin. We take one-third of the largest net shift in any single lab momentum bin (other than the first) as a flat systematic error due to momentum shifts. According to the left plot in Figure 7.5, this is about 0.5%. In short, although momentum bin shifts is correlated with momentum, we're treating it as an independent error due to its small size.

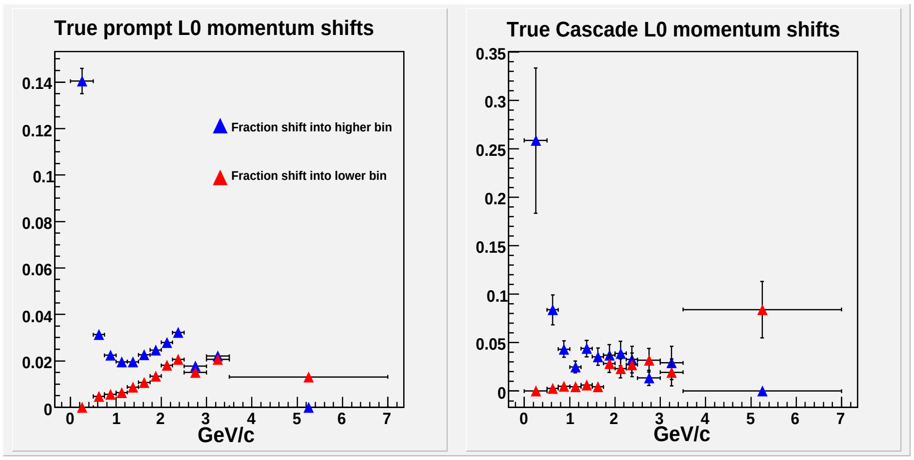


Figure 7.5: Percentages of reconstructed Monte Carlo Λ^0+$\overline{\Lambda}^0$ that shifted their lab momentum bin locations relative to their true momentum bins. Here, 'prompt' referes generally to candidates that decayed inside the beampipe, which is the sample that's of interest here.

The fitting errors can be assigned fairly easily by running our reconstruction and fitting software routines over our original sample of Monte Carlo Lambdas and anti-Lambdas. Using the MC events, we reconstruct the momentum spectrum for the patricles plus antiparticles and then compare the results with the true Monte Carlo spectrum for the generated particles and antiparticles. The differences between the reconstructed spectrum and the generated spectrum then gives us an idea of the robustness of our fitting procedure.

At first sight it would seem that the errors due to signal fitting should not be correlated with particle momentum; however, such an assumption may not be entirely on the mark due to the presence of momentum-dependent background shapes. So this would suggest that we should assign the errors due to fitting by momentum bins. However, since the background shapes varies greatly between data and MC, a momentum-dependent error based on studying the differences between reconstructed and true numbers of particles in the Monte Carlo may be misleadingly precise. Instead, we opt to use *twice* the total percentage standard deviation between the reconstructed spectral bin numbers and the true spectral bin numbers as a *flat* estimate for the fitting error across all bins; thus, we attribute a relative error of 2.0% to our fitting procedure. Figure 7.6 shows the fits from which we reconstructed the *Monte Carlo* Lambda momentum spectrum. The reconstructed spectrum itself is displayed in Figure 7.7 (for comparison purposes, the generator particles' spectrum is also displayed on the same plot.) We're again treating this potentially momentum dependent error as an independent error.

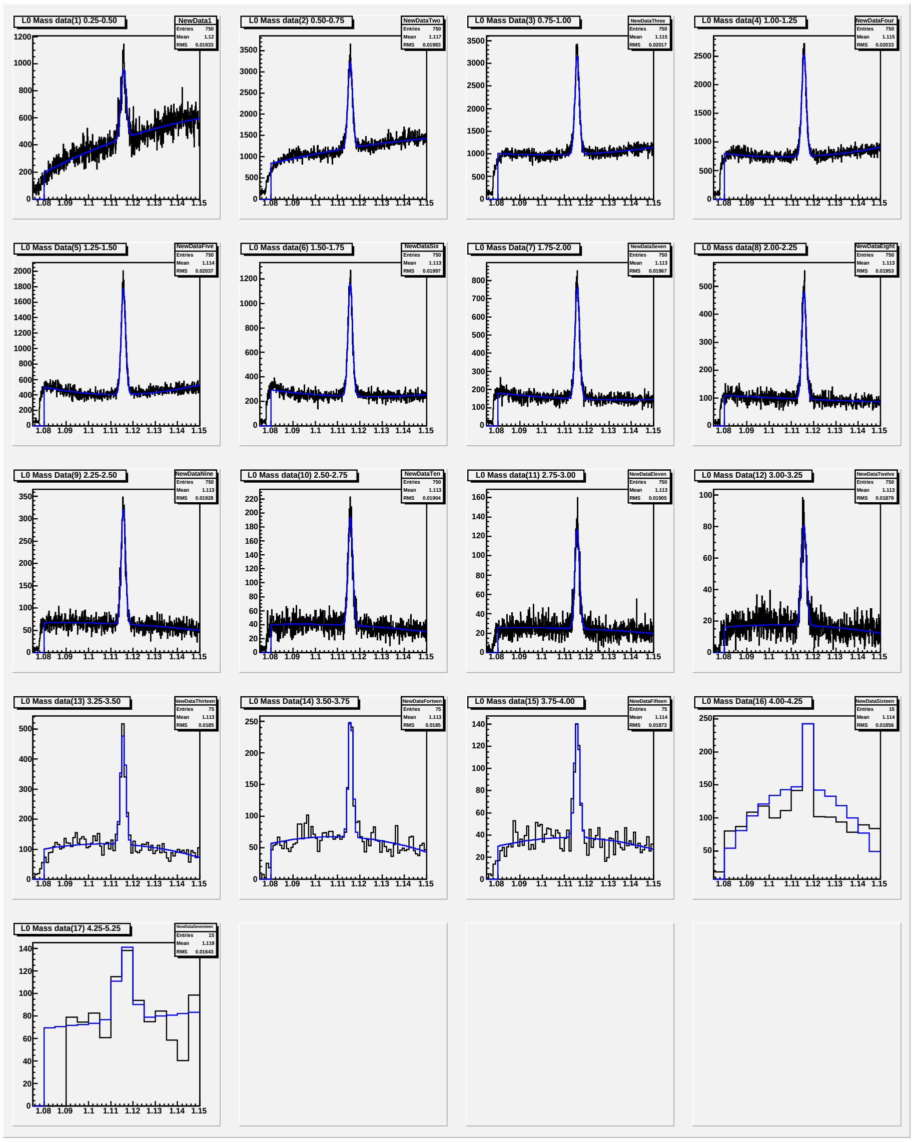


Figure 7.6: Fitted Monte Carlo Λ^0+$\overline{\Lambda}^0$ invariant mass signals in p_{cm} bins.

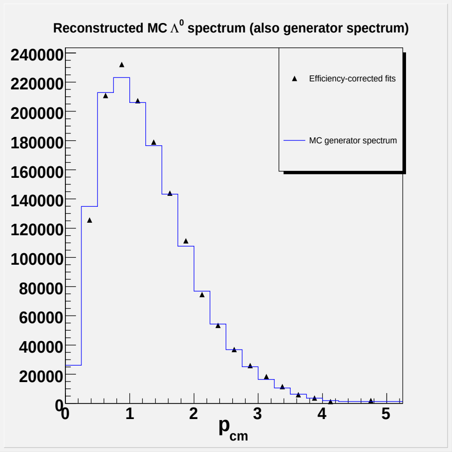


Figure 7.7: Reconstructed and true Monte Carlo Λ^0+$\overline{\Lambda}^0$ momentum spectrum.

7.1.5 Systematic error contributions from separate $\Lambda^0/\overline{\Lambda}^0$ fits

When considering the Λ^0 and the $\overline{\Lambda}^0$ samples separately, there does not appear to be any theoretical reasons which would suggest the presence of any difference between the particle and the antiparticle in terms of their momenta spectra and their production rates. We can make use of this state of affairs to check the consistency of our previously-derived spectrum which includes both the patricle and the antiparticle.

We ran our data efficiency-correction software routine for the Λ^0 and the $\overline{\Lambda}^0$ separately and then produced separate particle and antiparticle momentum spectra using the same procedures described in §6.2.3. The results are displayed in Figure 7.8 where we've converted momentum to momentum fraction and the absolute numbers to fractional hadronic cross setions. Note that in this figure, the fractional cross sections of the particle spectrum and the antiparticle spectrum have been scaled up by two to allow for comparisons with the combined particle plus antiparticle spectrum (also displayed in the same figure.)

A close inspection of Figure 7.8 reveals that for candidate CM momentum less than 1.25 GeV/c ($x_p \leq 0.2427$,) the $\overline{\Lambda}^0$ subsample's fractional cross sections exceed the corresponding fractional cross sections from the Λ^0 subsample; however, for CM momenta greater than 1.25 GeV/c ($x_p \geq 0.2427$,) the reverse is true (except for the 2.75-3.00 GeV/c bin (x_p between 0.5340 and 0.5825) where the antiparticle exceeds the particle number by 0.7%.) One possible source for such momentum-correlated discrepancy amongst the spectra may be due to the Monte Carlo decay distance ratios which we've used to correct for those Lambdas that decayed outside the PEP-II beampipe; if such MC ratios for 'slow' $\overline{\Lambda}^0$ is too high relative to the comparable ratios for the Λ^0 (e.g. if the MC expects more of the antiparticles to suffer annihilation interactions with material than reality,) than this may explain the apparent excess of the antiparticles relative to the particles below the 1.25 GeV/c threshold.

We elect to view the difference between the separate particle and antiparticle spectra as a real error that must be accounted for as part of our systematic uncertainties. Since this source of systematic uncertainties clearly is correlated with the candidate momentum, we must propagate it separately from other sources of pseudo-independent errors.

Table 7.3 gives the explicit numbers of fully efficiency-corrected and segragated Λ^0 and

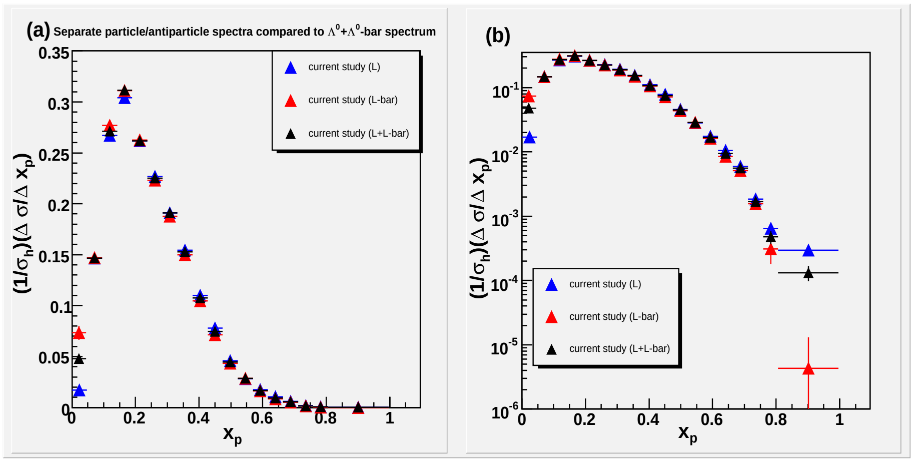


Figure 7.8: (a) Comparison of the separate Λ^0 spectrum with the $\overline{\Lambda}^0$ spectrum (also included is the spectrum from the combined particles + antiparticles fits.) (b) Same as plot in (a) but now the fractional cross sections are displayed in log-scale to render higher x_p data points visible.

$\overline{\Lambda}^0$ in the various spectral bins. Note that in this table also, and as part of the full correction, factors of 2 already have been applied to account for the fact that each of the spectra constitutes a half sample.

We also display the fits to the momentum-binned invariant mass signals which generated the numerical entries (after further corrections for decay distance cut, efficiency-bin based cuts etc.) in Table 7.3. Figure 7.9 displays the efficiency-corrected background and signals (corrections by the 2-d efficiency grids only) of the particle-only invariant mass plots from our off-peak data sample; our fits of the background plus signal have been superimposed upon the data. Figure 7.10 displays the equivalent plots for our antiparticle-only subsample. Figures 7.9 and 7.10 should be compared with Figure 6.10 which displays the combined particle plus antiparticle invariant mass plots and their respective fits. Further note that for the first bin (0.00-0.25 GeV/c) the actual fitted signal was 102 for the Λ^0 and 553 for the $\overline{\Lambda}^0$.

7.1.6 Consistency check using angular bins

We also checked the variations in the Λ^0 plus $\overline{\Lambda}^0$ momentum spectrum when we restrict our candidate acceptance to specified angular bins.

We define the following six azimuthal angle bins (given in terms of $\cos(\theta_{cm})$ where θ_{cm} is the angle defined by the candidate particle's CM momentum with respect to the BaBar z-axis): (-0.80) to (-0.65), (-0.65) to (-0.50), (-0.50) to (-0.33), (-0.33) to (-0.15), (-0.15) to (+0.06), (+0.06) to (+0.30). We also define three ϕ angle bins (given in terms of ϕ_{lab} where the lab ϕ angles are given in radians and are measured in the XY-plane with respect to the BaBar x-axis): (- π rad.) to (-1.575 rad.), (-1.575 rad.) to (-0.3937 rad.), (-0.3937 rad.) to (+π rad.).

The efficiency-correction procedure for each angular bin proceeds exactly the same as before except for the fact that we need to also correct for candidates that fall within the angular bins which were excluded. Figure 7.11 and Figure 7.12 display the fully-corrected

Cand. p_{CM} bins (GeV/c)	Fully corrected $\mathit{\Lambda}^0$	% bin stat. err. ($\mathit{\Lambda}^0$)	Fully corrected $\overline{\mathit{\Lambda}}^0$	% bin stat. err. ($\overline{\mathit{\Lambda}}^0$)
0.00-0.25	31,544	38.58%	132,936	18.79%
0.25-0.50	271,004	1.02%	272,050	1.02%
0.50-0.75	493,934	0.52%	512,180	0.51%
0.75-1.00	562,559	0.51%	575,462	0.50%
1.00-1.25	483,806	0.57%	485,123	0.57%
1.25-1.50	419,551	0.65%	412,793	0.65%
1.50-1.75	353,491	0.77%	346,961	0.77%
1.75-2.00	285,762	0.91%	277,020	0.92%
2.00-2.25	203,404	1.13%	193,848	1.16%
2.25-2.50	144,019	1.41%	132,055	1.47%
2.50-2.75	84,830	1.92%	80,612	1.97%
2.75-3.00	52,267	2.51%	52,637	2.50%
3.00-3.25	32,526	3.29%	29,870	3.43%
3.25-3.50	19,369	4.39%	15,626	4.89%
3.50-3.75	11,035	6.04%	9,403	6.55%
3.75-4.00	3,401	10.84%	2,908	11.72%
4.00-4.25	1,185	18.59%	569	26.83%
4.25-5.15	2,203	15.82%	32	100%
$\sum$ (0.00-5.15)	3,455,890	0.43%	3,532,084	0.74%

Table 7.3: Fitted numbers from separate $\mathit{\Lambda}^0$ and $\overline{\mathit{\Lambda}}^0$ samples and associated statistical uncertainties in p_{CM} bins. Factors of two have been applied to all bin numbers to correct for usage of particle/antiparticle only samples. Total errors are obtained by adding bin errors by quadrature.

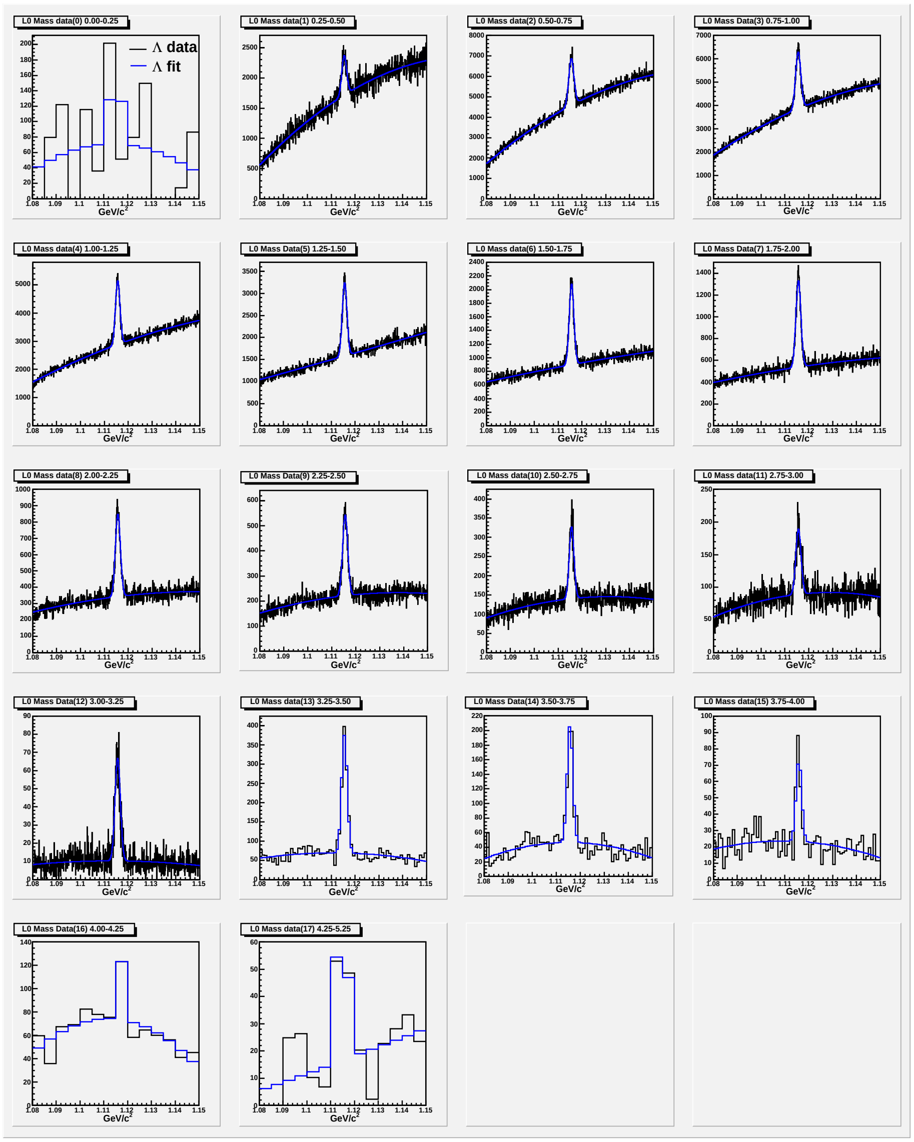


Figure 7.9: CM momentum-binned invariant mass plots and fits for off-peak data's Λ^0-only subsample.

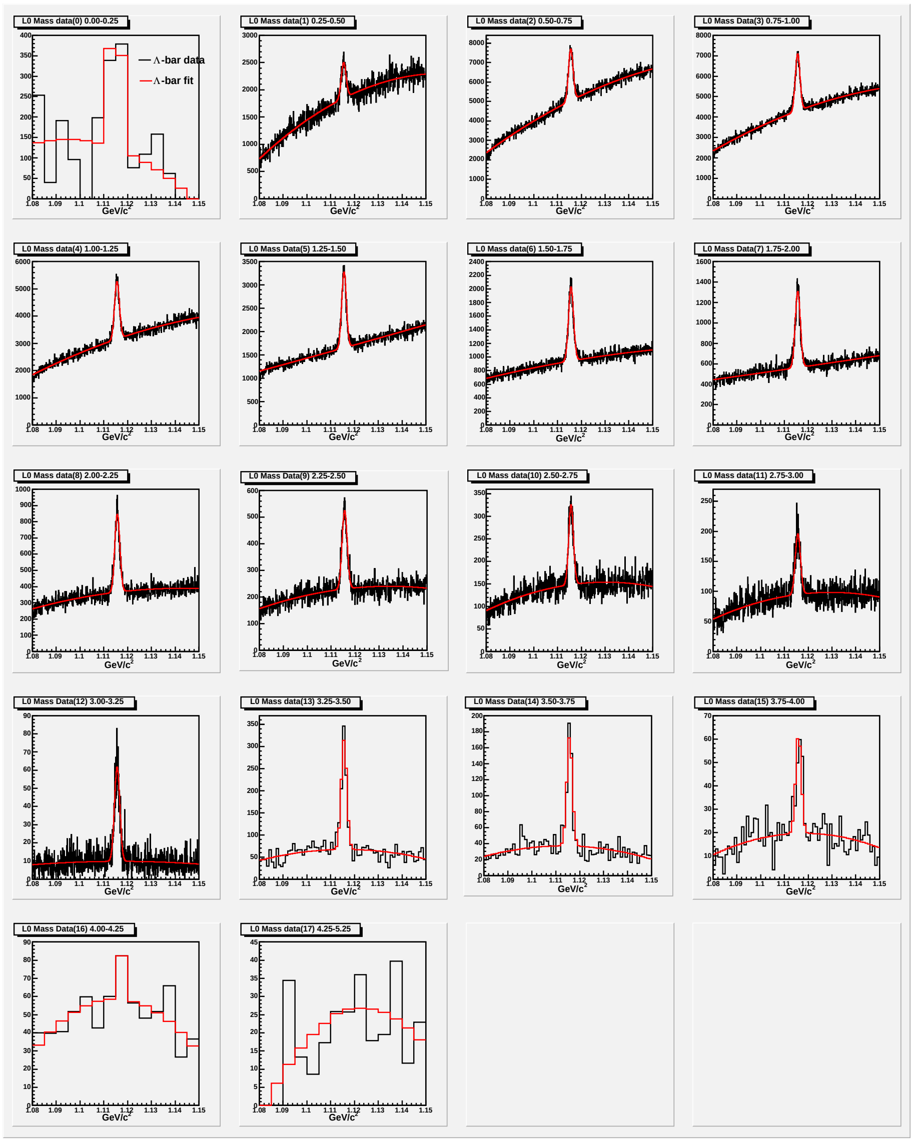


Figure 7.10: CM momentum-binned invariant mass plots and fits for off-peak data's $\overline{\Lambda}^0$-only subsample.

spectra for each of the $cos(\theta_{cm})$ and ϕ_{lab} bins respectively. Note that the bin error bars represent both the increased statistical errors as well as the estimated bin systematic errors.

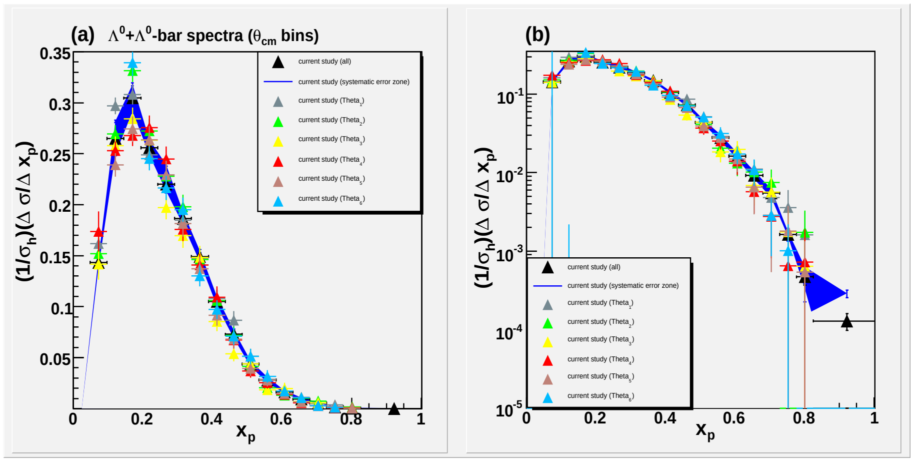


Figure 7.11: (a) Comparison of the spectra from different $\cos(\theta_{cm})$ bins with the combined spectrum for all bins. (b) Same as plot (a) but in log-scale. $\text{Theta}_1 \equiv +0.06\ to\ +0.30$, $\text{Theta}_2 \equiv -0.15\ to\ +0.06$, $\text{Theta}_3 \equiv -0.33\ to\ -0.15$, $\text{Theta}_4 \equiv -0.50\ to\ -0.33$, $\text{Theta}_5 \equiv -0.65\ to\ -0.50$, $\text{Theta}_6 \equiv -0.80\ to\ -0.65$.

The angular-binned spectra agree with our overall spectrum after the overall spectrum's systematic error zone has been increased to take into account variations due to the particle-only spectrum and anti-particle-only spectrum.

7.1.7 Consistency check using spectra from separate runs

We also would like to see how the spectrum varies over time. The simplest thing to do is to just perform our prior analysis with respect to the off-peak data by BaBar Run periods. In this manner, we obtained four combined particle plus antiparticle spectra for Run1, Run2, Run3, and Run4 periods. Figure 7.13 displays the results by BaBar Run periods.

The above figure appears to show that our reconstructed spectrum agrees best with the individual spectra from the Run2 and Run4; this is not unexpected since about 75% to 80% of our total off-peak data sample comes from these two run periods. Overall, the various spectra agree with each other within statistical plus systematic uncertainties. The exception to this is for the Run1 and Run3 data points in the 0.25-0.50 GeV/c bin; we attribute these discrepancies to the small statistics in this bin for Run1 and Run3 and discount their significance.

7.1.8 Consistency check using SVT Λ^0 and $\overline{\Lambda}^0$ samples

We limited our reconstruction to include only Λ^0 and $\overline{\Lambda}^0$ candidates that apprear to have decayed inside the PEP-II beampipe. However, since the Lambda has a mean life on the order of 10^{-10} seconds, a significant portion of true Lambda hyperons decayed outside the beampipe, in particular , the region between 4 cm. and 10 cm. in BaBar's XY-plane (located entirely within the inner portion of the SVT) should contain significant numbers of reconstructed particles and antiparticles.

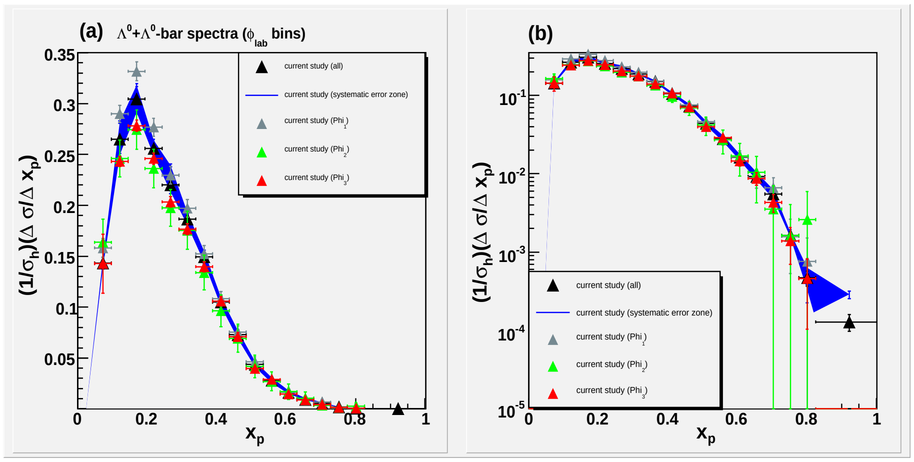


Figure 7.12: (a) Comparison of the spectra from different ϕ_{lab} bins with the combined spectrum for all bins. (b) Same as plot (a) but in log-scale. $\phi_1 = -0.3937$ rad. *to* $+\pi$ rad., $\phi_2 = -1.575$ rad. *to* -0.3937 rad., $\phi_3 = -\pi$ rad. *to* -1.575 rad..

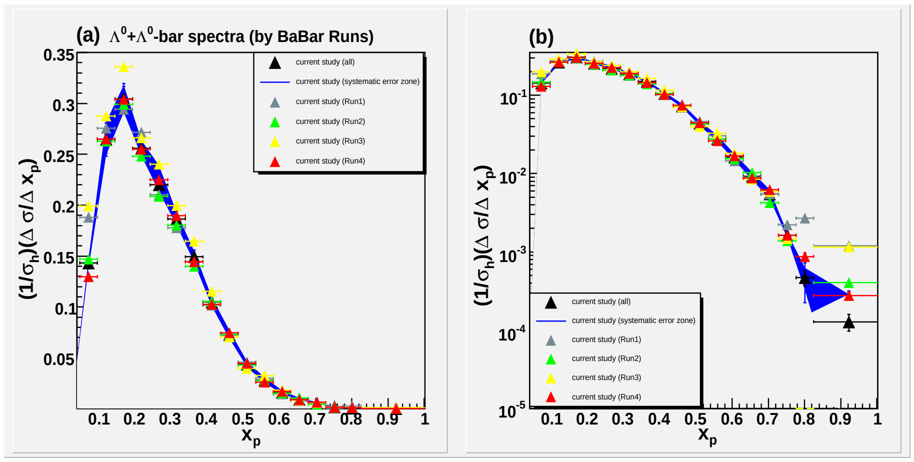


Figure 7.13: (a) Comparison of the spectra from different BaBar Run periods with the combined spectrum for all four Runs. (b) Same as plot (a) but in log-scale.

We can conduct our analysis using only this 'SVT sample' of reconstructed Λ^0 and $\overline{\Lambda}^0$ candidates; the resulting spectra gives us another means to check the consistency of our results from the beampipe sample.

In order to analyze the SVT sample of candidates, we constructed new efficiency maps using the Monte Carlo in which we accepted only reconstructed candidates and generator particles that decayed within the aforementioned inner SVT region. The new efficiency maps use the same grid definitions that we used to construct the beampipe sample efficiency grids. For higher momentum particles, these definitions should be adequate; however, Lambda hyperons with lower momenta ($\sim$ 0.75 GeV/c and less) are more likely to have been decay daughters of cascade hyperons and since the cascade hyperon can be deflected by the IFR magnetic field, the efficiency grid labAngle definitions based on the Lambda's momentum vector can be distorted with respect to BaBar's detector directions. We can calculate the deflection angle (in the lab frame) of the cascade hyperon as a function of its decay vertex:

$$\Delta\phi = \mathcal{A}\left[\left(\frac{eB}{m_{\Xi}}\right)\frac{E_{\perp}}{p_{\perp}c}\right] \tag{7.3}$$

where $\mathcal{A}$ is the total arclength traveled by the cascade hyperon, B represents the 1.5 Tesla IFR induction field, m_{Ξ} is the mass of the cascade hyperon, $E_{\perp}$ and $p_{\perp}$ are respectively the cascade's transverse energy and transverse momentum. Figure 7.14 displays six plots of the cascade hyperon's deflection angle (in degrees) as functions of decay distance (in meters) where each plot represents a different given particle momentum. We see from this figure that the total angular deflection out to 10 cm. in the transverse plane for cascade hyperons with transverse momenta gerater than 1 GeV/c is no more than about 3 degrees; but the total angular deflections rise to about 12^o at particle transverse momenta of 0.25 GeV/c. But since the BaBar lab reference frame is boosted in the forward direction and slightly rotated, we see that low transverse lab momentum particles contribute mostly to CM momentum values less than 1 GeV/c (see Figure 6.9.) For this reason, we expect some distortions in the SVT spectra with respect to the beampipe spectrum at low CM momentum due to the manner that the SVT efficiency grids are constructed at this time [26].

Figure 7.15 compares the separate Run1, Run2, Run3, and Run4 SVT Λ^0-only spectra with our beampipe spectrum (with particles and anti-particles combined.) The SVT Λ^0 spectra from Runs 1 throught 3 generally agree with our beampipe spectrum except at the expected low momentum region. However, the Run4 SVT Λ^0 spectrum appears to be relatively low.

It is unclear at this point whether the consistently lower spectrum of the Run4 SVT Λ^0 sample is a real effect. It is possible that the Monte Carlo simply does not model Run 4 Lambda candidates that decayed in the inner SVT very well; this matter merits further study. However, since the Run-separated beampipe spectra generally agree with each other

---

[26] There is also the additional angular distortion due to the angle defined by the Lambda daughter's momentum vector with respect to the cascade hyperon's momentum at its decay point. The maximum possible values for this angle can be quite large for cascade hyperons with momenta under 1 GeV/c; it goes from about 7.8^o at 1 GeV/c to over 15^o at 0.25 GeV/c. This angle can be calculated using the approximate epxression:

$$\theta_{max} \simeq \tan^{-1}\left(\frac{p_{max}}{\gamma\beta E_{max}}\right) \tag{7.4}$$

where p_{max} and E_{max} are the maximum possible momentum and energy for the daughter Lambda, and where γ and β are the boost and relative speed from the lab frame to the cascade hyperon's rest frame. This expression can be derived as an exercise (see exercise 11.20 in [25].)

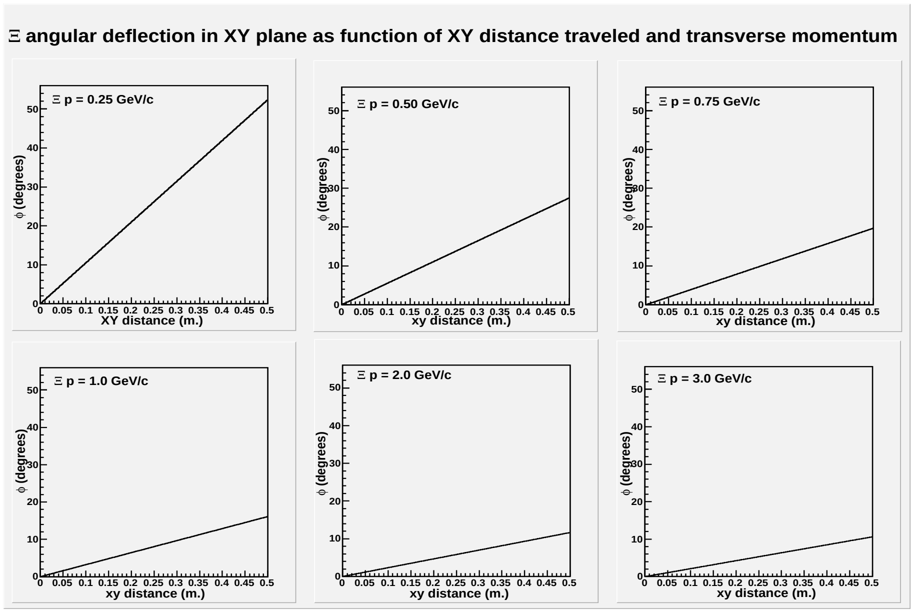


Figure 7.14: Cascade hyperon's angular deflection in BaBar's transverse plane as functions of particle $p_{\perp}$ and XY distance traveled.

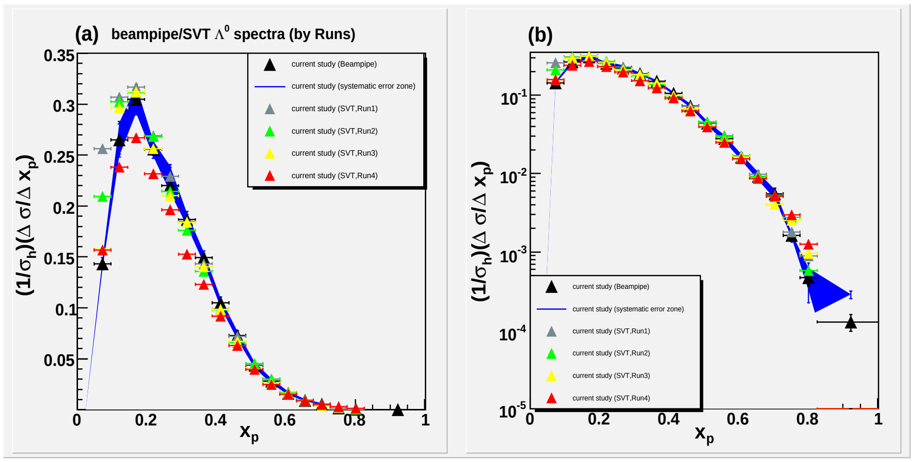


Figure 7.15: (a) Comparison of the spectrum from beampipe Λ^0 sample with spectra from Runs 1-4 SVT samples (separated by Runs.) (b) Same as plot (a) but in log-scale.

and since we believe that we understand the various effects associated with the beampipe sample, we shall defer further analysis of the Run4 SVT sample for a later date.

Figure 7.16 compares the separate Run1, Run2, Run3, and Run4 SVT $\overline{\Lambda}^0$-only spectra with our beampipe spectrum.

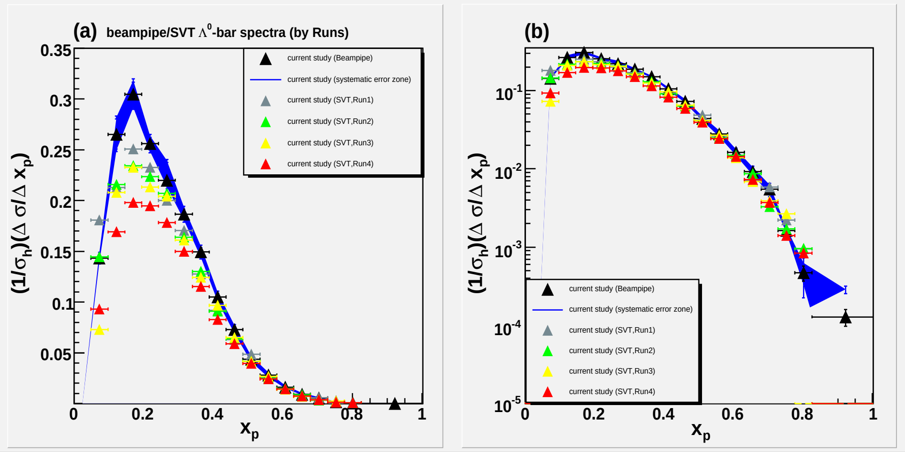


Figure 7.16: (a) Comparison of the spectrum from beampipe $\overline{\Lambda}^0$ sample with spectra from Runs 1-4 SVT samples (separted by Runs.) (b) Same as plot (a) but in log-scale.

The SVT samples' anti-Lambda spectra are clearly inconsistent with the beampipe sample spectrum (though it is interesting to note that the Run4 SVT $\overline{\Lambda}^0$ spectrum is relatively lower than its counterpart spectra from Run1, Run2, and Run3, analogous to the state of affairs for the SVT Λ^0 spectra.) This discrepancy is not well understood at present although it appears to persist through most of the CM momentum range. It is possible that there is a real depletion of the antiparticles that is not modeled by the Monte Carlo. Again, we defer further analysis of this issue until a later date. At present, we shall quote our Lambda rate and spectra results based on raw data coming from only the self-consistent beampipe Λ^0 and $\overline{\Lambda}^0$ sample.

7.1.9 Final Lambda spectrum and rate

There are a few residual sources of uncertainties that we must address. First, we note that for particle momentum-fraction (x_p) under 0.045 (corresponding to a Lambda CM momentum of about 0.23 GeV/c at BaBar,) the intrinsic detector efficiency and our selection cuts resulted in no signal candidates; in this momentum region we're essentially using Monte Carlo models to extrapolate the number of candidates by relying upon the reconstructed number of candidates from the sliver of momentum between 0.23 GeV/c and 0.25 GeV/c. Such extrapolations entail additional errors which we normally must address. However, since we have increased all the momentum bin error margins in order to account for experimentally derived particle/antiparticle differences, and since such an error increase for the first momentum bin is so enormous that it swamps differences between differing MC models, we have not included additional errors for the implicit MC extrapolation.

There is also PDG's published error on the Λ's lifetime. This error was introduced into our calculations during corrections for particles/antiparticles that decayed outside the beampipe;

Error Source	Type	Average error size (% of σ_{Λ})
Statistical error	independent	1.42%
MC efficiency error	independent	2.98%
Λ^0 PDG mean lifetime error	independent	0.76%
Event purity error	independent	0.20%
Cand. p-shifts error	correlated/independent	0.53%
Fitting errors	correlated/independent	2.00%
DCH tracking error	normalization	2.14%
SVT tracking error	normalization	0.93%
Error for $\Lambda^0/\overline{\Lambda}^0$ diff. (no 1st bin)	correlated	1.87%
Error for $\Lambda^0/\overline{\Lambda}^0$ diff. (all bins)	correlated	3.48%

Table 7.4: Summary of the sources of error for $\Lambda^0/\overline{\Lambda}^0$ spectrum and production rate. Errors labeled as "correlated/independent" are those which may be correlated with Lambda momentum but which we are nevertheless treating as if they are independent errors for the sake of simplicity due to their relative small size. Note that the average error sizes for the independent errors are larger than their contributions to the total production rate error since they will be summed quadratically for that purpose.

we explicitly add this as a source of systematic error.

Finally, there is the systematic error due to the event purity estimate which we have discussed in chapter 4 herein. Table 7.4 summarizes all sources of errors for the $\Lambda^0 + \overline{\Lambda}^0$ spectrum .

In the table, we have classified each error as one of the following types: independent, correlated, and normalization. Also note that according to table 7.4, the total relative error attributable to the particle/antiparticle differences drop from almost 3.5% to 1.9% when we remove the data point from the very first momentum bin (0.00-0.25 GeV/c.)

We now present a momentum-bin-by-momentum-bin account of the continuum inclusive fractional cross sections and errors of the Λ^0 plus $\overline{\Lambda}^0$ spectrum. The fractional cross sections per unit x_p [27] are given as fractions of the estimated total hadronic cross section. See Table 7.5.

Further note that the fractional cross sections per unit x_p given in Table 7.5 have been corrected for the presence of initial state radiation ("ISR") in the collision events by comparing the generated Λ^0 plus $\overline{\Lambda}^0$ spectra with and without ISR Pythia/Jetset Monte Carlo.

The total statistical error quoted in Table 7.5 are summed from the momentum binned statistical errors in quadrature. On the other hand the quoted total systematic error is obtained as follows: we perform a simple sum of the correlated errors from the various momentum bins (here, it's just the extra errors from the $\Lambda^0/\overline{\Lambda}^0$ comparisons) while the various sources of independent errors from the bins are summed quadratically; subsequently, the total correlated error, the total independent error, and the normalization errors are summed with each other by quadrature to produce the quoted total systematic error of 0.003897. The systematic error quoted for each individual momentum bin on the other hand represents the

---

[27] x_p is the momentum of a particle presented as a fraction of the available energy.

p_{CM} bins (GeV/c)	x_p Range	$\Lambda^0+\overline{\Lambda}^0$ $(\Delta\sigma/\Delta x_p)$	Stat. Err. $(\Delta\sigma/\Delta x_p)$	Sys. Err. $(\Delta\sigma/\Delta x_p)$
0.00-0.25	0.0000-0.0485	0.047068	0.010985	0.030033
0.25-0.50	0.0485-0.0971	0.143313	0.001471	0.003994
0.50-0.75	0.0971-0.1456	0.264411	0.001370	0.009514
0.75-1.00	0.1456-0.1942	0.297613	0.001488	0.011801
1.00-1.25	0.1942-0.2427	0.250481	0.001419	0.008359
1.25-1.50	0.2427-0.2913	0.215950	0.001406	0.007726
1.50-1.75	0.2913-0.3398	0.188343	0.001445	0.005324
1.75-2.00	0.3398-0.3884	0.150703	0.001375	0.006500
2.00-2.25	0.3884-0.4369	0.107491	0.001234	0.005394
2.25-2.50	0.4369-0.4854	0.074719	0.001078	0.005052
2.50-2.75	0.4854-0.5340	0.045230	0.000880	0.002921
2.75-3.00	0.5340-0.5825	0.028608	0.000713	0.002442
3.00-3.25	0.5825-0.6311	0.016348	0.000554	0.001666
3.25-3.50	0.6311-0.6796	0.009790	0.000453	0.001249
3.50-3.75	0.6796-0.7282	0.006378	0.000397	0.000676
3.75-4.00	0.7282-0.7767	0.001752	0.000199	0.000209
4.00-4.25	0.7767-0.8252	0.000533	0.000117	0.000211
4.25-5.15	0.8252-1.0000	0.000440	0.000063	0.000567
$\sum$ (0.00-5.15)	$\sum$ (0.0-1.0)	0.089967	0.000574	0.003897

Table 7.5: $\Lambda^0+\overline{\Lambda}^0$ CM momentum spectrum (presented as fractions of hadronic σ_{hadron} per unit x_p.) Note that the total statistical error (x_p 0.0-1.0) was obtained by a quadrature sum. The total systematic error (x_p 0.0-1.0) contribution from independent errors also are summed quadratically while the contributions from correlated errors are summed linearly; the total systematic includes the normalization errors.

quadratic sum of the correlated error in that bin with the various independent errors in the same bin.

Figure 7.17 displays the final ISR-corrected continuum Lambda spectrum using the information from Table 7.5. The error bars on the data points represent the statistical errors only (though they are generally not visible except in the first bin and the last three bins) while the region between the two blue lines represent the systematic error zone.

The total Lambda (particle plus antiparticle) continuum production rate is just the sum of the momentum-binned fractional cross sections. We quote this rate (i.e. the mean number of Λ^0 and $\overline{\Lambda}^0$ produced per continuum event at $\sqrt{s} = 10.54$ GeV) as: $0.0900 \pm 0.0006(\text{stat.}) \pm 0.0039(\text{sys.})$

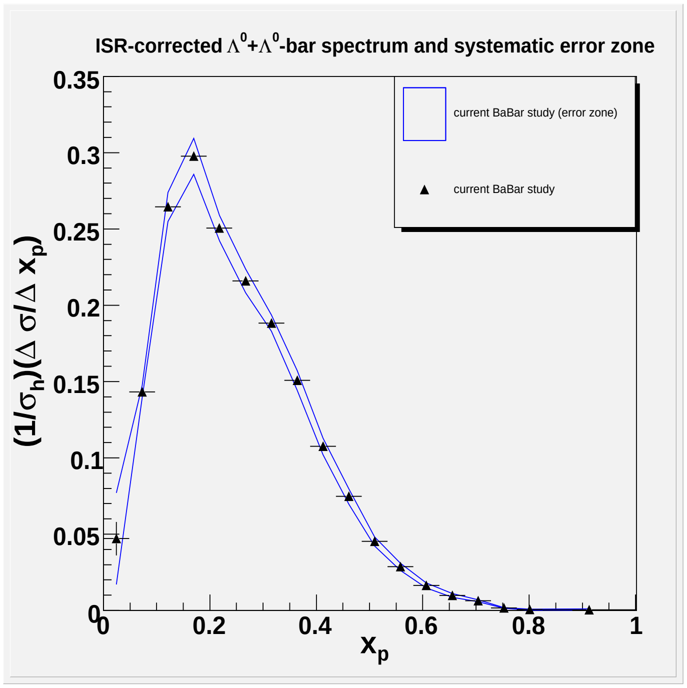
We emphasize that the level of systematic measurement error for our total production rate is 4.3%; the scale of this error is due to our decision to use our actual measurement for the first x_p bin ($x_p < 0.485$) which results in a more model-independent result for the total rate. The scale of the systematic error can be reduced if we attempt to extrapolate our spectrum down to the first bin using either Monte Carlo generators or logarithmic fits of our measured spectrum. However, the former method is obviously model-dependent while the latter method involves discarding the real signal from the first bin; for these reasons we will not employ extrapolations for the Lambda hyperon at present. We note in passing that the total production rate and errors *excluding the first bin* is $0.0878 \pm 0.0002(\text{stat.}) \pm 0.0029(\text{sys.})$ Thus if we exclude the first bin, the systematic error on the production rate falls to about 3.3% of the total rate for the x_p region extending from 0.0485 to 1.000.

7.1.10 Comparisons of Lambda spectrum with prior results and with Monte Carlo

Now we briefly compare our current results from BaBar with the continuum Lambda spectra obtained by the ARGUS and CLEO collaborations about 20 years ago. Figure 7.18 displays our current result for the Lambda hyperon in comparison with the CLEO (green spheres) and the ARGUS (red squares) data points; the error bars attached to the CLEO and ARGUS data points include both the statistical and systematic errors (added in quadrature.)

We observe that our current Lambda spectrum agrees with that obtained by the ARGUS collaboration to within common errors. The inclusive production rate in the continuum obtained by the ARGUS Collaboration [2] was $0.092 \pm 0.003(\text{stat.}) \pm 0.008(\text{sys.})$, which also agrees with our result of $0.0900 \pm 0.0006(\text{stat.}) \pm 0.0039(\text{sys.})$. On the other hand, the spectrum obtained by the CLEO Collaboration is generally lower than our spectrum even though, after accounting for the relatively large errors associated with CLEO's data points, the results are consistent with one another. The continuum inclusive production rate quoted by CLEO [3] is 0.066 ± 0.010 where the statistical and systematic errors have been added in quadrature. Based on our current studies, it would appear that the current world average (with ARGUS and CLEO results) for the Lambda inclusive production rate near 10 GeV is low.

Figure 7.19 displays our current result for the Lambda hyperon in comparison with the spectra generated by BaBar's custom SP5 generator (using the Pythia/Jetset generators) and the default UCLA generator. It appears that the default UCLA generator under-predicts while the BaBar implementation of Pythia over-predicts the continuum inclusive Λ production rate at BaBar energy. The default UCLA spectrum also appears to be harder than the data at high x_p.



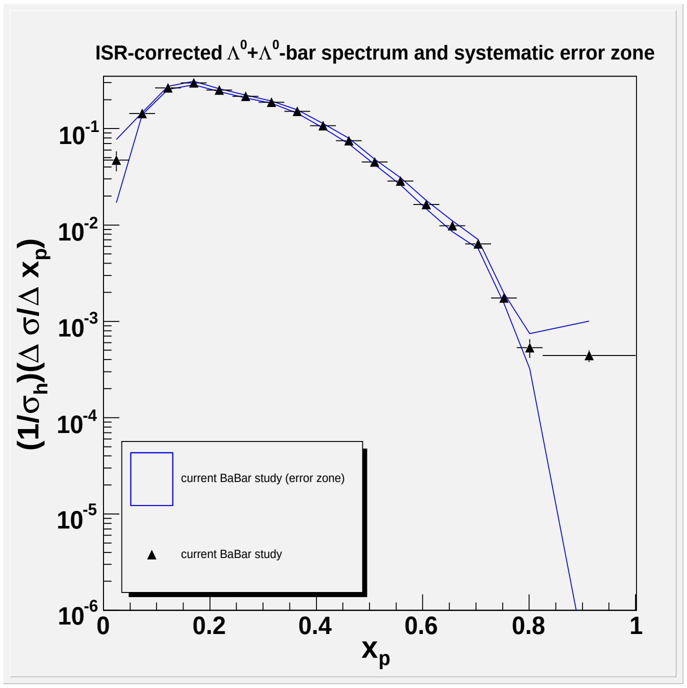


Figure 7.17: (Top) ISR-corrected inclusive continuum spectrum for Λ^0 plus $\overline{\Lambda}^0$ at CM energy 10.54 GeV. (Bottom) Same as top plot but in log-scale.

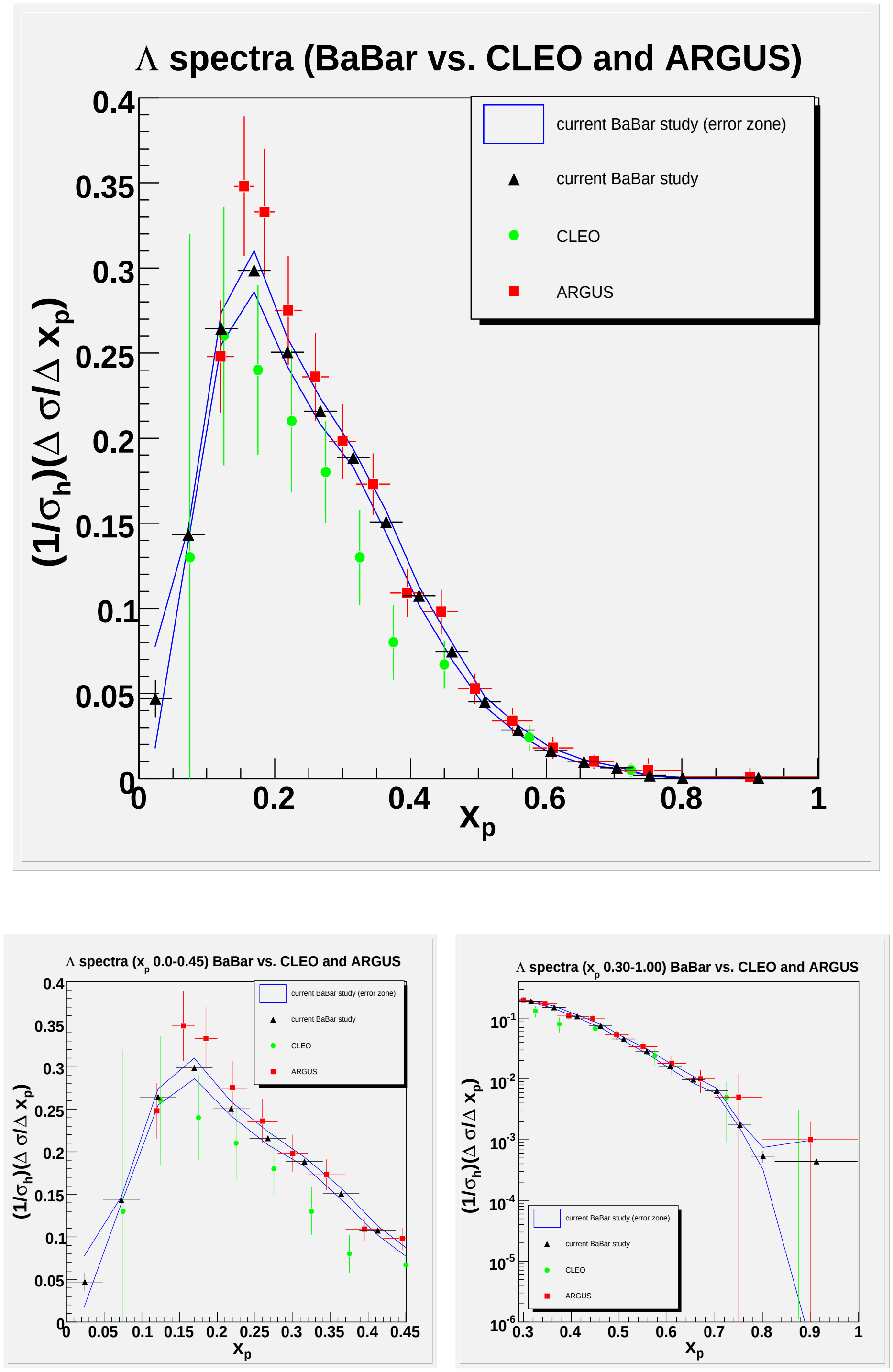


Figure 7.18: (Top) BaBar inclusive continuum spectrum for Λ^0 plus $\overline{\Lambda}^0$ at 10.54 GeV compared with CLEO and ARGUS spectra. (Bottom) Partial views of same spectra.

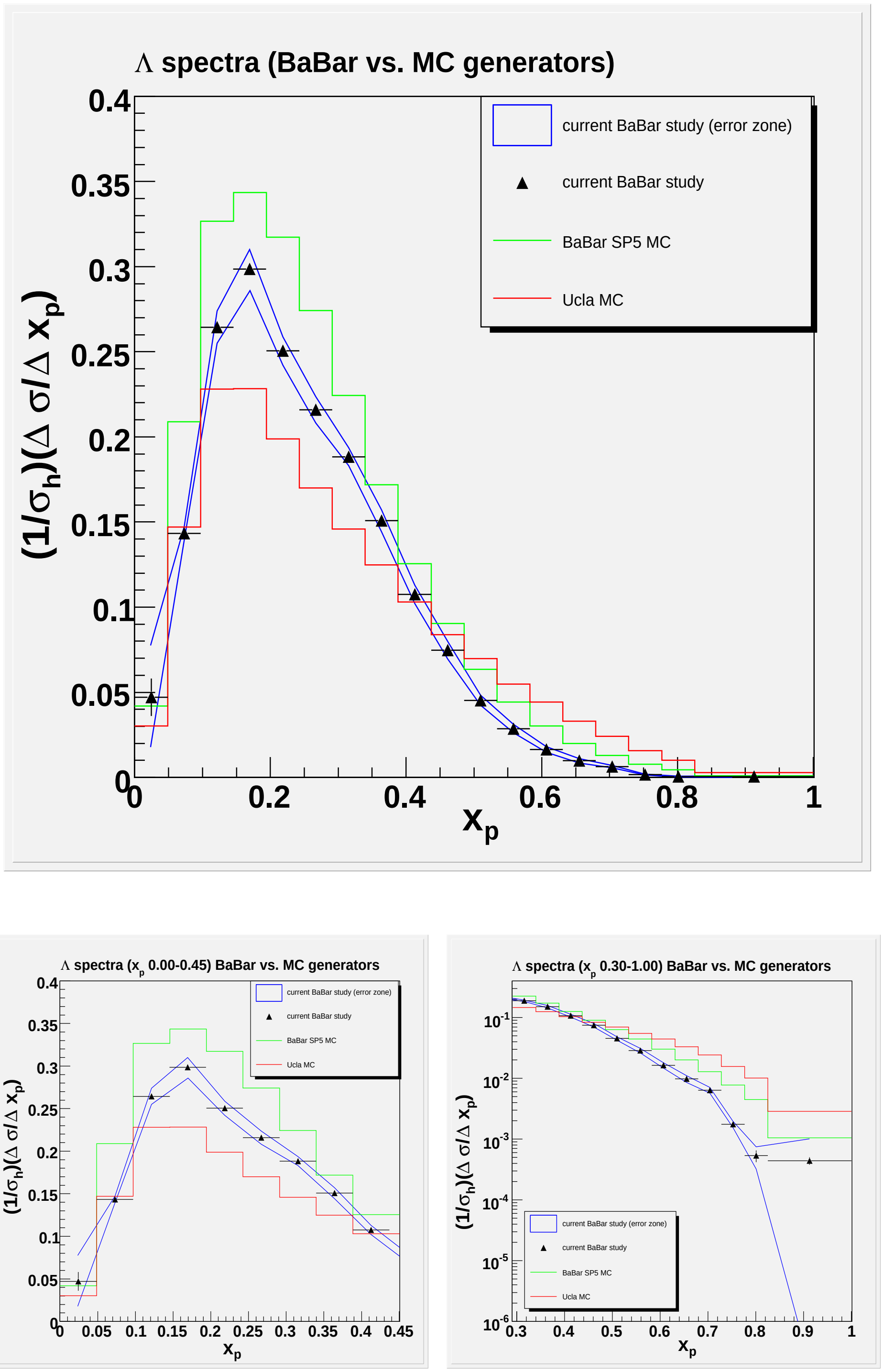


Figure 7.19: (Top) BaBar inclusive continuum spectrum for Λ^0 plus $\overline{\Lambda}^0$ at 10.54 GeV compared with the spectra from the default UCLA generator and BaBar's SP5 (Phythia) generator. (Bottom) Partial views of same spectra.

Candidate p_{CM} bins (GeV/c)	Fitted raw signals	Partially corrected #'s	Fully corrected	% bin stat. uncertainty
0.25-0.50	65	1,335	14,379	12.40%
0.50-0.75	254	3,127	23,008	6.27%
0.75-1.00	295	2,693	24,652	5.82%
1.00-1.25	339	2,822	36,844	5.43%
1.25-1.50	174	1,491	23,764	7.58%
1.50-1.75	173	1,184	22,655	7.60%
1.75-2.00	139	851	21,399	8.48%
2.00-2.25	73	408	13,171	11.70%
2.25-2.50	63	328	13,316	12.60%
2.50-2.75	282	2,490	4,296	5.95%
2.75-3.00	193	1,611	3,100	7.20%
3.00-3.25	162	1,421	2,775	7.86%
3.25-3.50	67	586	1,180	12.21%
3.50-3.75	19	193	389	22.94%
3.75-4.00	24	250	577	20.41%
4.00-5.10	1	12	68	100%
$\sum$ (0.00-5.10)	2,323	20,802	205,555	2.39%

Table 7.6: Fitted numbers of Ξ^- and $\overline{\Xi}^+$ and associated statistical uncertainties in p_{CM} bins.

7.2 Ξ^- hyperon results

We use the signal fits of the Xi (particles plus antiparticles) invariant mass plots in CM momentum bins (see Figures 6.22 and 6.23) to construct the spectrum.

7.2.1 Preliminary Ξ^- and $\overline{\Xi}^+$ spectra and statistical uncertainties

Recall that for the momentum region < 2.50 GeV/c, our invariant mass signals for Ξ candidates include only those for which the particles' Λ daughters also have decayed within the PEP-II beampipe; on the other hand, for momentum region $\geq$ 2.50 GeV/c, our signals include candidates for which the Λ daughters decayed inside the inner part of the SVT (between 4 and 12 cm. in BaBar coordinates) as well as inside the PEP-II beampipe.

The statistical uncertainties associated with the number of signal candidates that we detected in each momentum bin generally are estimated in the same way as for the Λ hyperon (see §7.1.1;) however, in the case for the Ξ, the number of acutal signals candidates without *any* efficiency corrections were fitted on an momentum-bin-by-momentum-bin basis. Thus, we can estimate momentum correlated bin efficiencies for the Ξ hyperon.

Our results for the number of fitted signal particles and antiparticles (raw, or after either partial or full efficiency corrections) and the associated bin statistical uncertainties are given in Table 7.6. Once again, the overall statstical uncertainty of about 2.4% is obtained by quadratically summing the estimated statistical uncertainties from the individual bins. Also note that the raw number from the last bin (4.00-5.10 GeV/c) was obtained by visual inspection instead of via software.

The explicit expression that we used for estimating the statistical error in any individual CM momentum bin in Table 7.6 consists of the following Poisson standard deviation expression:

$$\sigma_{\text{stat}} = \sqrt{N_{\text{noCorr}}}\left(\frac{N_{\text{fullCorr}}}{N_{\text{noCorr}}}\right) \tag{7.5}$$

where N_{noCorr} is the number of raw signal candidates in the bin (column 2 in Table 7.6) and N_{fullCorr} is the fully efficiency-corrected and fitted number of candidates given in column 4 of Table 7.6. Figure 7.20 displays a graphical representation of the information conained in Table 7.6 after converting from absolute numbers of efficiency-corrected candidates to fractional hadronic cross sections per unit x_p. The percent errors given in column 5 of the table are calculated using the usual expression: $(\sigma_{\text{stat}}/N_{\text{fullCorr}}) = (1/\sqrt{N_{\text{noCorr}}})$.

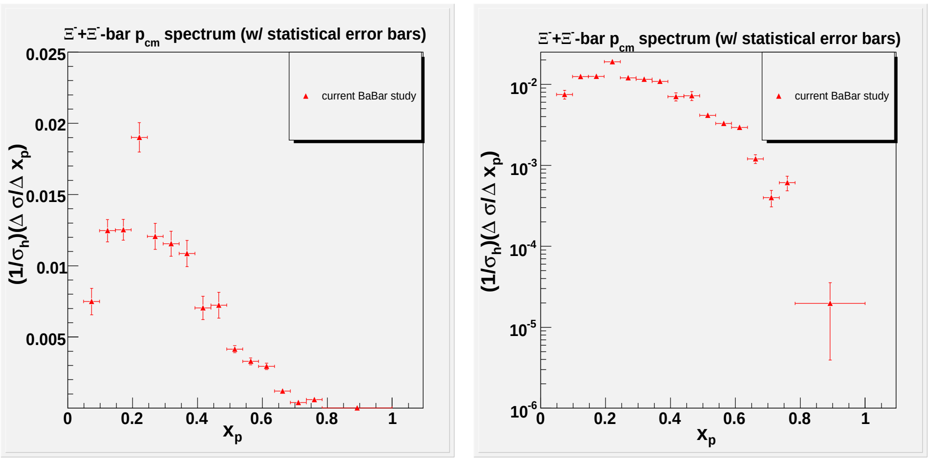


Figure 7.20: (Left) Preliminary spectrum of Ξ^- and $\overline{\Xi}^+$. (Right) Same as left plot but now in log-format.

7.2.2 Propagated tracking and efficiency uncertainties

The statistical errors associated with each bin in the 2-d Ξ^- and $\overline{\Xi}^+$ MC efficiency matrices as well as the tracking correction errors for the proton and the two pion tracks that together make up a cascade hyperon are propagated during the efficiency-correction process of each accepted data event.

Alternative efficiency-corrected number of candidates (treating the aforementioned errors as real corrections) are produced and the resulting signals are fitted. The differences between the alternative signals and our primary signals then give us estimates of the propagated systematic errors associated with tracking and the MC statistical errors. This is analogous to the procedure that we used to estimate such systematic uncertainties for the Lambda hyperon measurement.

There is an additional complication for cascade hyperon signals with CM momenta greater than 2.50 GeV/c. Recall that we accepted cascade candidates in this momentum region if their Lambda daughers decayed either within the PEP-II beampipe *or* within the inner SVT. These expanded criteria for candidate acceptance were necessary in order to gain better statistics for the higher momentum region; however, they also raised the problem of the apparent Lambda beampipe/SVT spectra discrepancy that we disucssed in §7.1.8. In the event that the beampipe/SVT Λ discrepancy represents a genuine loss of efficiency on the part of the SVT Lambda sample relative to the beampipe sample, we include the relative SVT/beampipe fractions as an additional source of error for the Ξ^- and $\overline{\Xi}^+$ candidates.

Thus at the time that we perform efficiency-correction on an individual cascade hyperon candidate from the off-peak data sample (Runs 1-3 only,) if such a candidate's Lambda daughter is adduced to have decayed between 4.0 cm. and 12 cm. in the BaBar detector's transverse plane (in BaBar coordinates,) then an additional error component is assigned to the cascade hyperon's detection efficiency based upon its Λ daughter's CM momentum (the scale of the error is equal to the relative SVT/beampipe Λ^0 or $\overline{\Lambda}^0$ fraction from the relevant CM momentum bin and for a specified BaBar Run period.) In order to further control the scale of this error, we excluded Run 4 cascade candidates with momenta $\geq$ 2.50 GeV/c from our analysis since the Run 4 SVT Lambda and anti-Lambda spectra displayed significantly greater levels of differences from the beampipe spectrum than the corresponding Runs 1-3 spectra (see Figures 7.15, 7.16.) This additional source of error dominates the tracking and MC statistical errors for cascade candidates with momenta $>$ 2.50 GeV/c.

We present the above-described sources of systematic errors in Table 7.7. In the table, the bins with momenta $<$ 2.50 GeV/c contain the combined tracking and MC statistics error only. The tracking errors are expected to be of the "normalization" type (as described in §§7.1.2, 7.1.9) while the cascade MC statistic is exepected to be an "independent" source of error. Thus the tracking errors should not affect the Ξ^- and $\overline{\Xi}^+$ spectrum and will be added to the systematic error of the cascade hyperon's production rate. Due to the size of the beampipe/SVT Lambda discrepancy factors, the bins in Table 7.7 with momenta $\geq$ 2.50 GeV/c essentially list only the Lambda correction source of error; for these bins, the scales of the associated tracking and MC statistics error have been set to 7%, the size of the tracking/statistics error for the 2.25-2.50 GeV/c bin.

Figure 7.21 presents the propagated systematic error information contained in Table 7.7 as a graphical error zone.

7.2.3 MC statistical error consistency check

We performed a check on the cascade hyperon's assigned systematic errors due to MC efficiency grids' statistics in the same manner that we performed the Lambda MC statistical errors check as described in §7.1.3. As in the case for the Lambda hyperon, we fed a new set of generated Pythia/Jetset events through our 2-d efficiency grids for the Ξ^- and the $\overline{\Xi}^+$ to simulate the effects of the detector on the generated events. Then new 2-d efficiency matrices are constructed using these events; the calculated bin statistical uncertainties for the new matrices then are compared with the old matrices to evaluate their consistency or lack thereof.

The consistency of the bin statistical uncertainties are plotted in terms of pulls and fitted with a single gaussian. The best fit (on a relatively small sample of bins) occurs when we increase the average calculated bin uncertainties by a factor of 1.32. However, in view of the fact that the Lambda efficiency bin statistical errors were increased by a larger factor of 1.55 (see §7.1.3,) we opted to take a more cautious approach with respect to the Xi; thus we multiplied the average calculated efficiency bin statistical uncertainties for the Ξ and $\overline{\Xi}^+$ by the same 1.55 factor as well. Figure 7.22 displays the calculated pull and the fit for the Ξ's

Cand p_{cm} (GeV/c)	Fully corrected Numbers (errors added)	% errors Trk. + MC + Λ
0.25-0.50	15,036	4.57%
0.50-0.75	24,171	5.05%
0.75-1.00	25,256	2.45%
1.00-1.25	38,280	3.90%
1.25-1.50	25,084	5.63%
1.50-1.75	23,688	4.56%
1.75-2.00	22,506	5.17%
2.00-2.25	13,914	5.64%
2.25-2.50	14,250	7.01%
2.50-2.75	5,241	22.01%
2.75-3.00	3,742	20.73%
3.00-3.25	3,605	29.91%
3.25-3.50	1,409	19.45%
3.50-3.75	508	30.57%
3.75-4.00	718	24.40%
4.00-5.10	102	50.00%
$\sum$ (0.25-5.10)	217,510	5.82%

Table 7.7: Ξ systematic uncertainties in p_{CM} bins due to tracking, MC efficiency matrices' bin statsitics, and SVT/beampipe Λ fractions. The percent uncertainties are reported as percentages of the total estimated numbers of fully-corrected signal particles in that bin as reported in 7.6.

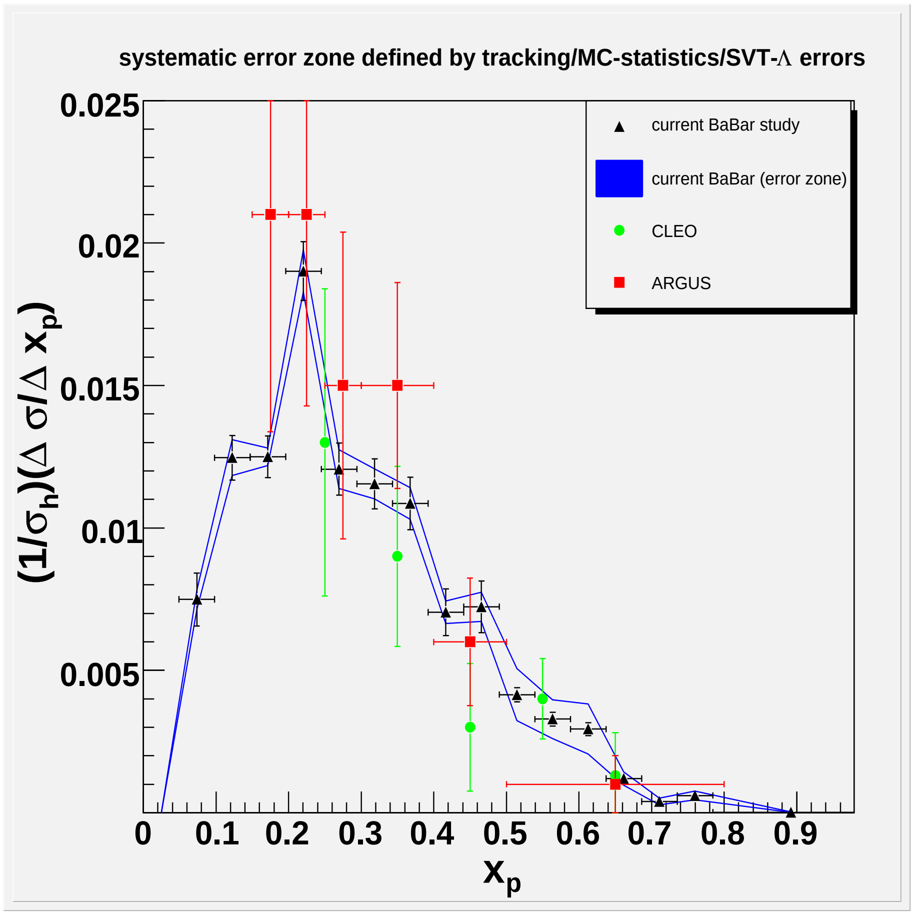


Figure 7.21: Zone of correlated errors due to tracking corrections, $\Xi^-/\overline{\Xi}^+$ efficiency bin statistics, and beampipe/SVT Λ discrepancy; in units of fractional hadronic cross section. CLEO and ARGUS collaborations' results for $\Xi^0/\overline{\Xi}^+$ have been superimposed for comparison purposes.

2-d efficiency statistical uncertainties.

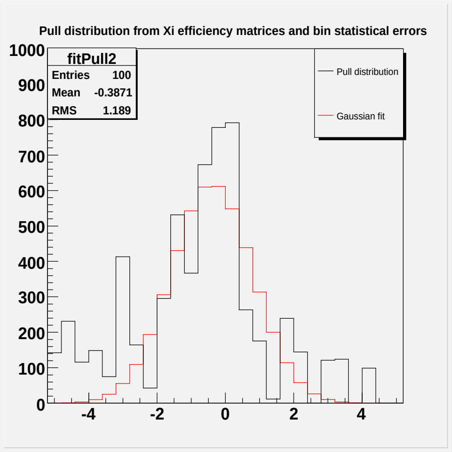


Figure 7.22: Pull distribution calculated from Xi efficiency matrices and the estimated efficiency bin statistical uncertainties (with 1.32 multiplicative factor.)

7.2.4 Consistency check using separate $\Xi^-/\overline{\Xi}^+$ fits

As in the case for the Lambda hyperon, we can use the particle and antiparticle subsamples as consistency checks on the stability of our results. We separately fit the Ξ^- and the $\overline{\Xi}^+$ subsamples; the fully efficiency-corrected particle and antiparticle numbers then are multiplied by factors of two in order to simulate the spectrum of the full particle+antiparticle samples. The fitted results are listed in Table 7.8.

Figure 7.23 compares the separate Ξ^- and $\overline{\Xi}^+$ spectra detailed in the Table 7.8 with the combined $\Xi^- + \overline{\Xi}^+$ spectrum. The error bars on all data points in the figure represent only the estimated statistical uncertainties while the error zone includes only the tracking error, MC statistics error, and (for momentum bins $\geq$ 2.50 GeV/c) the beampipe/SVT Lambda discrepancy error.

The figure shows that the separate particle and antiparticle spectra agree with our combined spectrum within common errors even though the figure is displaying only a limited set of systematic errors. Since the spectra generally agree within statistical plus systematic error margins, we do not increase the systematic error scale for our combined spectrum in order to include the data points from the separate particle/antiparticle spectra as we did for the case of the Lambda.

7.2.5 Consistency check using SVT Ξ^- and $\overline{\Xi}^+$ samples

We also can compare our spectrum with an alternative spectrum derived by taking into account cascade candidates with Lambda daughters that decayed within the inner SVT (4 cm. to 12 cm. in the BaBar reference frame.) This comparison of course is relevant only for the momentum region $<$ 2.50 GeV/c since our spectrum already employs this ‘alternative’ sample of reconstructed candidates if they have momenta $\geq$ 2.50 GeV/c. We also note that this alternative spectrum was efficiency-corrected using a separate set of 2-d efficiency grids

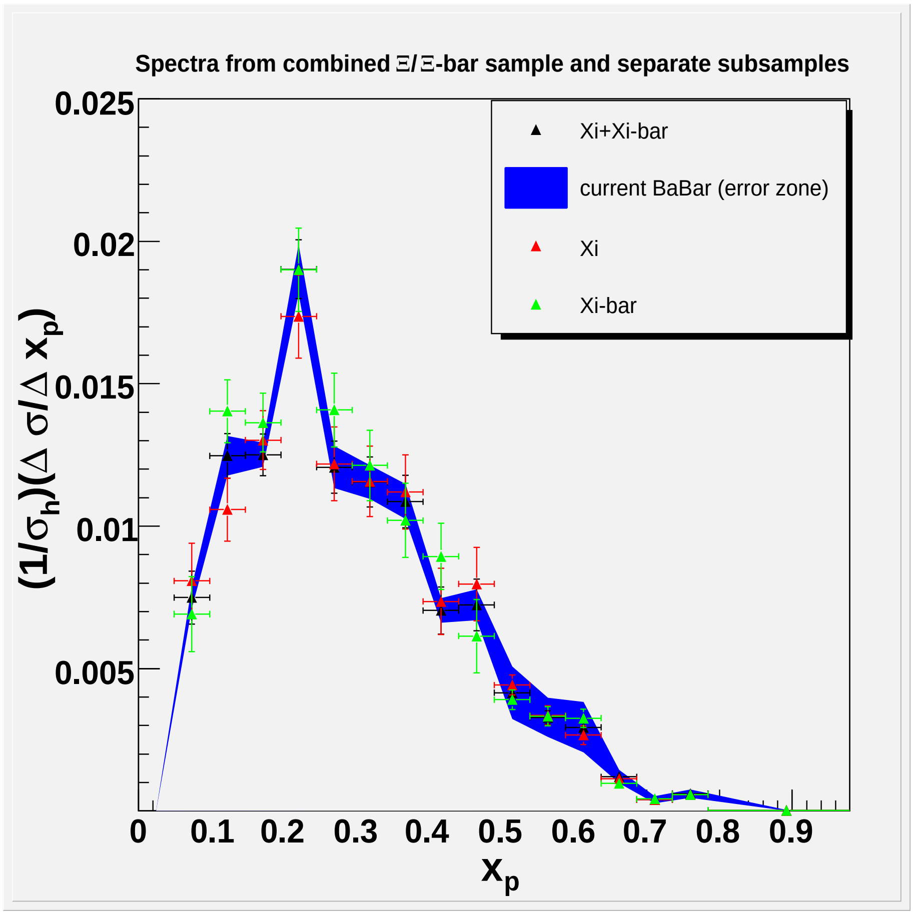


Figure 7.23: Comparison of the separate Ξ^- and $\overline{\Xi}^+$ spectra with the combined Ξ+$\overline{\Xi}^+$ spectrum

Cand. p_{CM} bins (GeV/c)	Fully corrected Ξ^-	Fully corrected $\overline{\Xi}^+$	% bin stat. err. (avg. $\Xi^-/\overline{\Xi}^+$)
0.25-0.50	15,503	13,263	17.54%
0.50-0.75	19,529	25,896	8.87%
0.75-1.00	25,675	26,889	8.23%
1.00-1.25	33,638	36,815	7.68%
1.25-1.50	23,985	27,712	10.72%
1.50-1.75	22,705	23,800	10.75%
1.75-2.00	22,057	20,098	12.00%
2.00-2.25	13,762	16,719	16.55%
2.25-2.50	14,668	11,304	17.82%
2.50-2.75	4,578	4,045	8.42%
2.75-3.00	3,159	3,135	10.18%
3.00-3.25	2,515	3,074	11.11%
3.25-3.50	1,103	947	17.28%
3.50-3.75	391	415	32.44%
3.75-4.00	540	545	28.87%
4.00-5.10	57	64	100%
$\sum$ (0.25-5.10)	203,866	214,721	11.3%

Table 7.8: Fitted numbers from separate Ξ^- and $\overline{\Xi}^+$ samples and associated statistical uncertainties in p_{CM} bins.

than the one employed for the beampipe cascade sample. The alternative efficiency grids were displayed in Figure 6.21.

Figure 7.24 compares the two referenced spectra. The two spectra agree remarkly well considering the fact that a similar comparison for the Lambda hyperons produced a significantly greater level of discrepancy. Nevertheless, there are three apparent points of disagreement bewteen the two cascade hyperon spectra located at momenta range 1.00-1.25 GeV/c (x_p between 0.1961 and 0.2451,) 1.75-2.00 GeV/c (x_p between 0.3431 and 0.3922,) and at 2.25-2.50 GeV/c (x_p between 0.4412 and 0.4902.) However, after we add in all sources of systematic errors, only the data points in momentum range 1.00-1.25 GeV/c still differ beyond common errors. We view this point of disagreement as another source of error for our primary spectrum. Therefore, we increase the systematic error for this momentum bin so that the two spectra agree to within common error. We consider this additional source of error to be correlated with momentum.

7.2.6 Final Xi spectrum and rate

There are additional sources of systematic errors. We did not conduct separate momentum shift and MC fitting studies for the cascade hyperons. Instead we rely on the estimated errors obtained for the Lambda hyperon. We use the 0.53% momentum bin shift error that we previously used for the Lambda. We also will use the same 2% Lambda signal fitting error for the cascade hyperon even though we would expect the cascade fits to be more accurate than the Lambda due to the use of PID to reduce the background. There are, in addition, the hadronic event purity error, the Lambda PDG lifetime error, and the Xi PDG lifetime

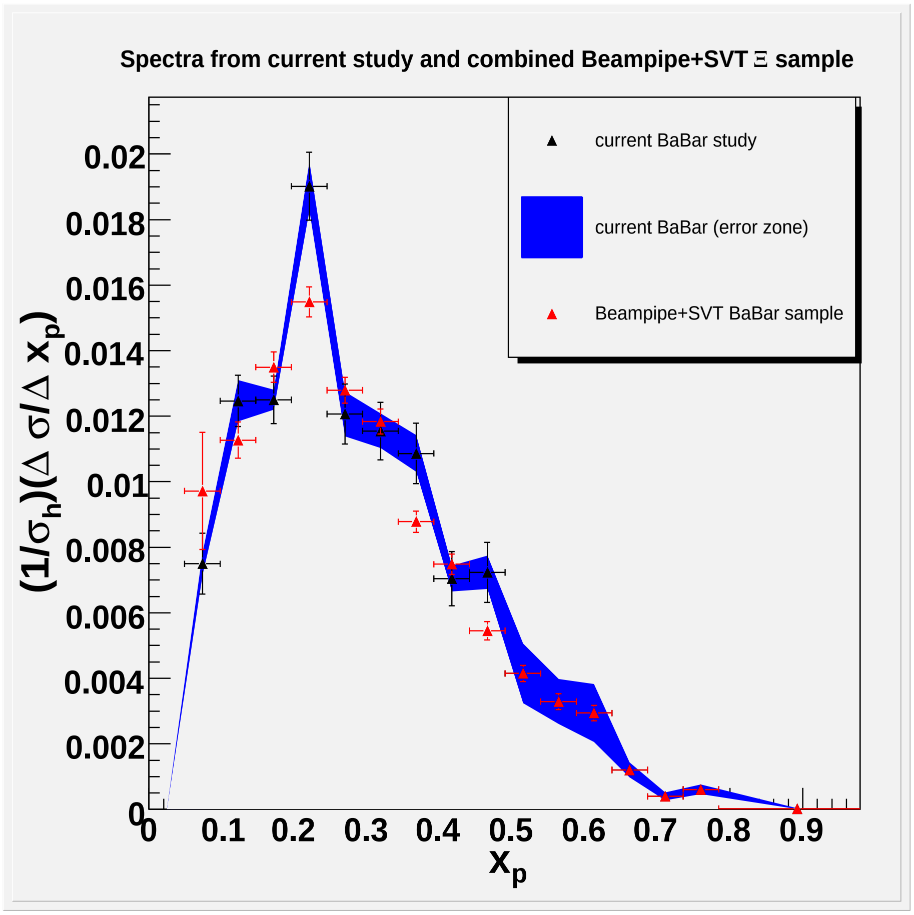


Figure 7.24: Comparison bewteen the spectrum from our composite beampipe-SVT Ξ^- sample with the full beampipe+SVT Ξ^- spectrum

error; the latter two errors contribute during the analysis step when we correct for cascade signal candidates which decayed outside our acceptance volume using the Monte Carlo.

There also is the issue of particle identification. We used PID on the Lambdas' perceived protonic daughters to reconstruct the cascade candidates in order to reduce the background. The PID performace of any software routine invariably will show differences when it is applied to data rather than to a MC sample. This issue has been studied by the Inclusive Hadronic group at SLAC for the global likelihood PID selector that we employed in the cascade analysis [30]. However, the SLAC study pertained to the so-called SP4 Monte Carlo sample rather than to the SP5 sample. In view of this fact, and in lieu of conducting an extended study of this selector's performance with respect to SP5, we take the largest momentum-binned proton PID efficiency difference between SP4 and data as a uniform systematic error. Since PID is expected to be perfectly correlated with particle momentum (i.e. more or less uniform relative size across momentum bins) and since we're treating this error as uniform, this is a normalization error.

Finally we must address the issue of first bin extrapolation which was employed to obtain a total estimated production rate for the cascade hyperon. The efficiency cuts and the intrinsic detector efficiency combined to produce no signal for CM momentum bin 0.00-0.25 GeV/c. We used the Pythia generator's relative production statistics for the momentum region 0.00-0.75 GeV/c to extrapolate the expected number of signal candidates produced in the unobserved region. As an estimate of the inherent error due to the model dependence of such an extrapolation, we also used the UCLA generator to produce an alternative number of unobserved cascade signals in the 0.00-0.25 GeV/c region. This is considered as a source of systematic error. Further, the statistical and systematic errors in the 0.25-0.75 GeV/c region of our data sample also were included as additional sources of systematic error for the extrapolation. The total size of these systematic sources of error due to the extrapolation is pegged at 0.3% of the total cascade hyperon cross section. This is treated as a momentum-correlated source of error.

Table 7.9 lists the various sources of errors (statistical and systematic) and provides estimated average sizes for each of these identified errors.

Table 7.10 contains a momentum-binned account of the continuum inclusive differential cross section and errors of the Ξ^- plus $\overline{\Xi}^+$ spectrum. The Xi differential cross section per unit x_p are given as fractions of the estimated total hadronic cross section. In addition, the fractional cross sections per unit x_p given in the following table have been corrected for the presence of ISR as was done for the case of the Lambda hyperon. Also note that the total systematic error in any particular bin includes only the independent and momentum-correlated errors *from the same bin* summed in quadrature.

Figure 7.25 displays the final ISR-corrected continuum cascade hyperon spectrum using the information from Table 7.10. The error bars on the data points represent the statistical errors only, while the region between the two blue lines represents the estimated systematic error zone.

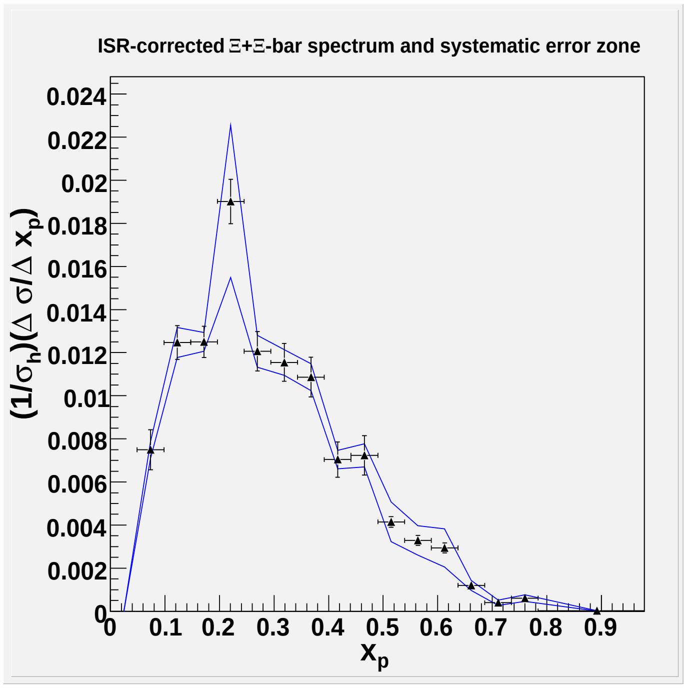
The total inclusive production rate for the cascade hypeon is obtained from Table 7.10 by summing the bins' cross sectional contributions to the total rate and including the model-dependent extrapolation of the first x_p bin (0.000-0.049). The total statistical error in the production rate is obtained from the same table by summing the bins' statistical error contributions in quadrature. The production rate's total systematic error is obtained by first summing the independent sources of errors by themselves in a quadratic manner and then summing the correlated errors separately in a linear manner; after obtaining the separate

Error Source	Type	Average error size (% of σ_{Ξ})
Statistical error	independent	7.90%
MC efficiency error	independent	7.48%
Λ^0 PDG mean lifetime error	independent	0.76%
Ξ^- PDG mean lifetime error	independent	0.92%
Event purity error	independent	0.20%
Cand. p-shifts error	corr./indep.	0.53%
Fitting errors	corr./indep.	2.00%
DCH+SVT tracking error	normalization	3.05%
PID error	normalization	3.50%
Beampipe/SVT Λ^0 difference	correlated	3.21%
First bin extrapolation	correlated	$> 1\%$

Table 7.9: Summary of the sources of error for $\Xi^-/\overline{\Xi}^+$ spectrum and production rate. Errors labeled as "corr./indep." are those which may be correlated with Xi momentum but which we are nevertheless treating as if they are independent errors. Again, the average error sizes for the independent errors are larger than their contributions to the total production rate error.

p_{CM} bins (GeV/c)	x_p Range	Ξ^-+$\overline{\Xi}^+$ ($\Delta\sigma/\Delta x_p$)	Stat. Err. ($\Delta\sigma/\Delta x_p$)	Sys. Err. ($\Delta\sigma/\Delta x_p$)
0.25-0.50	0.0490-0.0980	0.007494	0.000930	0.000680
0.50-0.75	0.0980-0.1471	0.012467	0.000782	0.001092
0.75-1.00	0.1471-0.1961	0.012501	0.000728	0.000591
1.00-1.25	0.1961-0.2451	0.019017	0.001033	0.003767
1.25-1.50	0.2451-0.2941	0.012065	0.000915	0.001166
1.50-1.75	0.2941-0.3431	0.011547	0.000878	0.000918
1.75-2.00	0.3431-0.3922	0.010861	0.000921	0.000969
2.00-2.25	0.3922-0.4412	0.007040	0.000824	0.000681
2.25-2.50	0.4412-0.4902	0.007233	0.000911	0.000860
2.50-2.75	0.4902-0.5392	0.004148	0.000247	0.000918
2.75-3.00	0.5392-0.5882	0.003288	0.000237	0.000686
3.00-3.25	0.5882-0.6373	0.002939	0.000231	0.000882
3.25-3.50	0.6373-0.6863	0.001203	0.000147	0.000236
3.50-3.75	0.6863-0.7353	0.000398	0.000091	0.000122
3.75-4.00	0.7353-0.7843	0.000609	0.000124	0.000149
4.00-5.10	0.7843-1.0000	0.000020	0.000016	0.000010
$\sum$ (0.25-5.10)	$\sum$ (0.049-1.000)	0.005534	0.000132	0.000437

Table 7.10: Ξ^-+$\overline{\Xi}^+$ CM momentum spectrum (presented as fractions of hadronic σ_{hadron} per unit x_p.) Once again, the total statistical error was obtained by a quadrature sum. The total systematic error contribution from independent errors also are summed quadratically while the contributions from correlated errors are summed linearly; the total systematic includes the normalization errors.



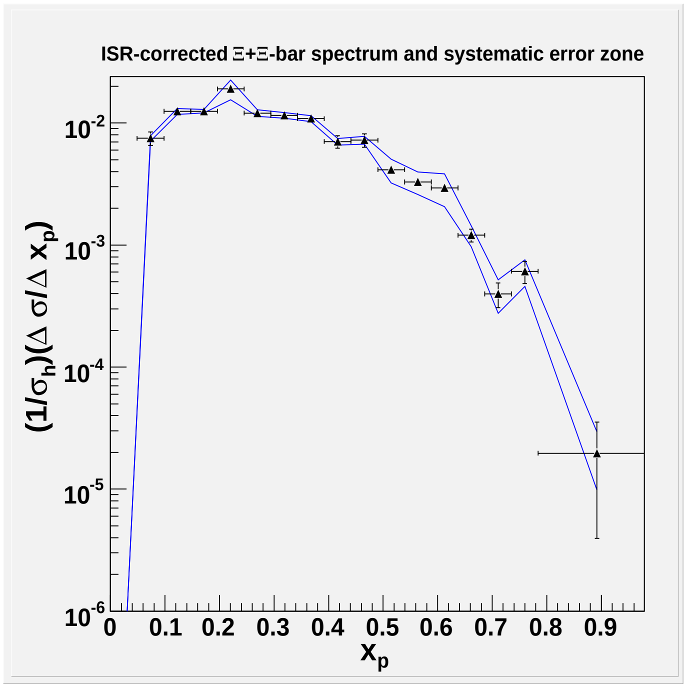


Figure 7.25: (Top) ISR-corrected inclusive continuum spectrum for Ξ^- plus $\overline{\Xi}^+$ at CM energy 10.54 GeV. (Bottom) Same as top plot but in log-scale.

total correlated error and total independent error, they then are summed quadratically along with the total normalization error to produce the final systematic error figure of 0.000437. We note that the extrapolated fractional cross section contribution of the first x_p bin to the total production rate is 0.000197; this is about 3.5% of our total rate. The extrapolated first bin's systematic error (treated as momentum-correlated) contributes 0.000017 to the rate uncertainty; this represents about 3.8% of the total systematic uncertainty in the rate.

We quote the following as the inclusive continuum production rate for the cascade hyerpon at 10.54 GeV: $0.00562 \pm 0.00013(\text{stat.}) \pm 0.00045(\text{sys.})$

7.2.7 Comparisons of Xi spectra with prior results and with Monte Carlo

The two previous measurements of the continuum inclusitve production rate of the cascade hyperon near 10 GeV were conducted by CLEO and ARGUS. CLEO's result was 0.005 ± 0.001 while ARGUS gave $0.0067 \pm 0.0006(\text{stat.}) \pm 0.0007(\text{sys.})$ (see [2], [3].) We observe that our current result falls between the numbers quoted by CLEO and ARGUS and agree with both numbers within errors.

Figure 7.26 displays the spectrum obtained in the current study (with both statistical and systematic errors) superimposed with the published spectrum obtained by both the CLEO and ARGUS collaborations.

It is evident that except for one single CLEO data point located at about $x_p \simeq 0.45$, all the CLEO and ARGUS momentum data points agree with the current spectrum. This is not an altogether unexpected result considering that the CLEO and ARGUS momentum-binned cross section per unit x_p results are associated with relatively large total errors. Our total momentum-binned errors are in all cases smaller than the prior results' errors.

We have also included a comparison of the inclusive cascade hyperon momentum spectra generated by the default Pythia/Jetset and UCLA generators (Figure 7.27.) It appears that the default Pythia cross section is significantly higher than our current result though its higher momentum behavior is in general agreement with our data. On the other hand the UCLA generator's default cascade hyperon spectrum more closely resembles our result in terms of total rate though its high momentum behavior is much harder than the current result. Of course, modifications of the various parameters in the generators may alter these observations.

7.3 Ω^- hyperon results

We now proceed to discuss the Omega hyperon's inclusive continuum spectrum and total production rate.

7.3.1 Preliminary Ω^- and $\overline{\Omega}^+$ spectra and statistical uncertainties

We first present the number of fully-corrected Ω^- plus $\overline{\Omega}^+$ obtained in each CM momentum bin using the fits described in §6.4.3. See Table 7.11.

Figure 7.28 displays a graphical representation of the information contained in Table 7.11 after converting from absolute numbers of efficiency-corrected candidates to fractional hadronic cross sections per unit x_p. The percent errors given in column 5 of the table are calculated in the same manner as for the cascade hyperon (see §7.2.1.)

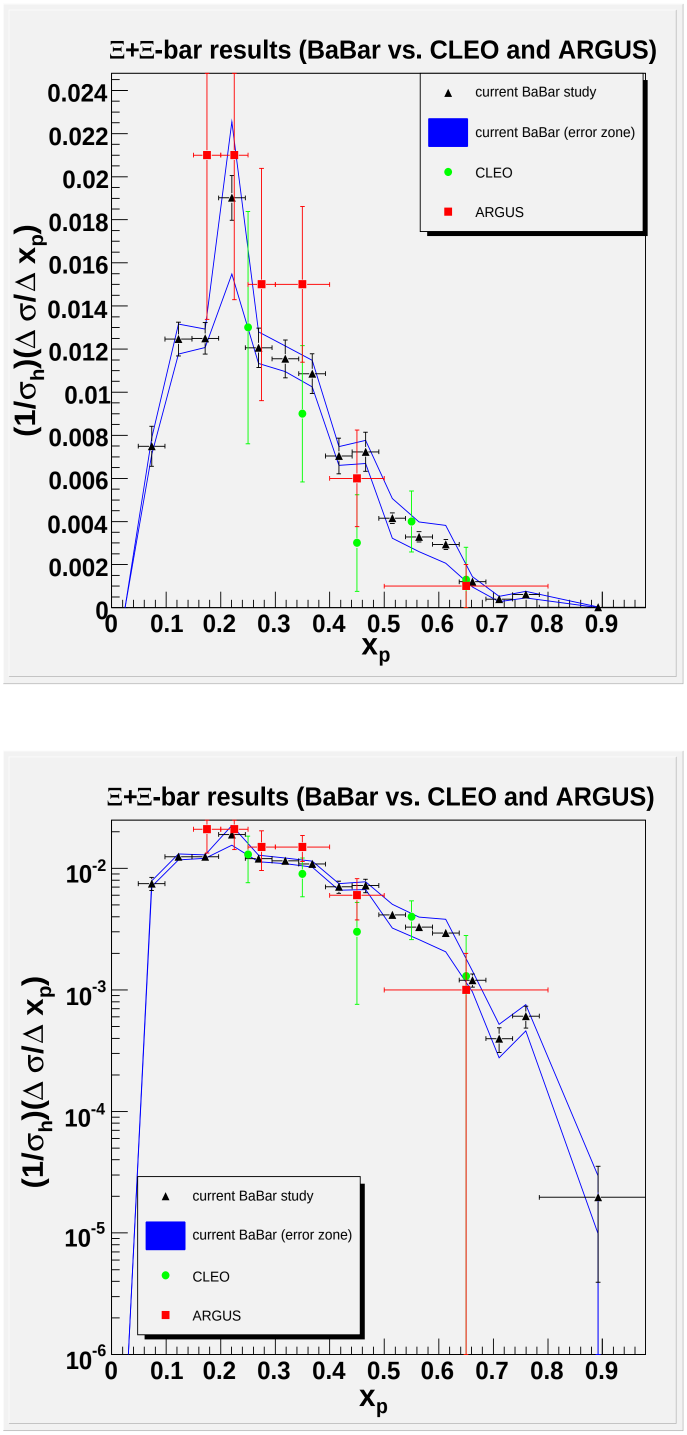


Figure 7.26: (Top) BaBar inclusive continuum spectrum for Ξ^- plus $\overline{\Xi}^+$ at 10.54 GeV compared with CLEO and ARGUS spectra. (Bottom) Log plots of the same spectra.

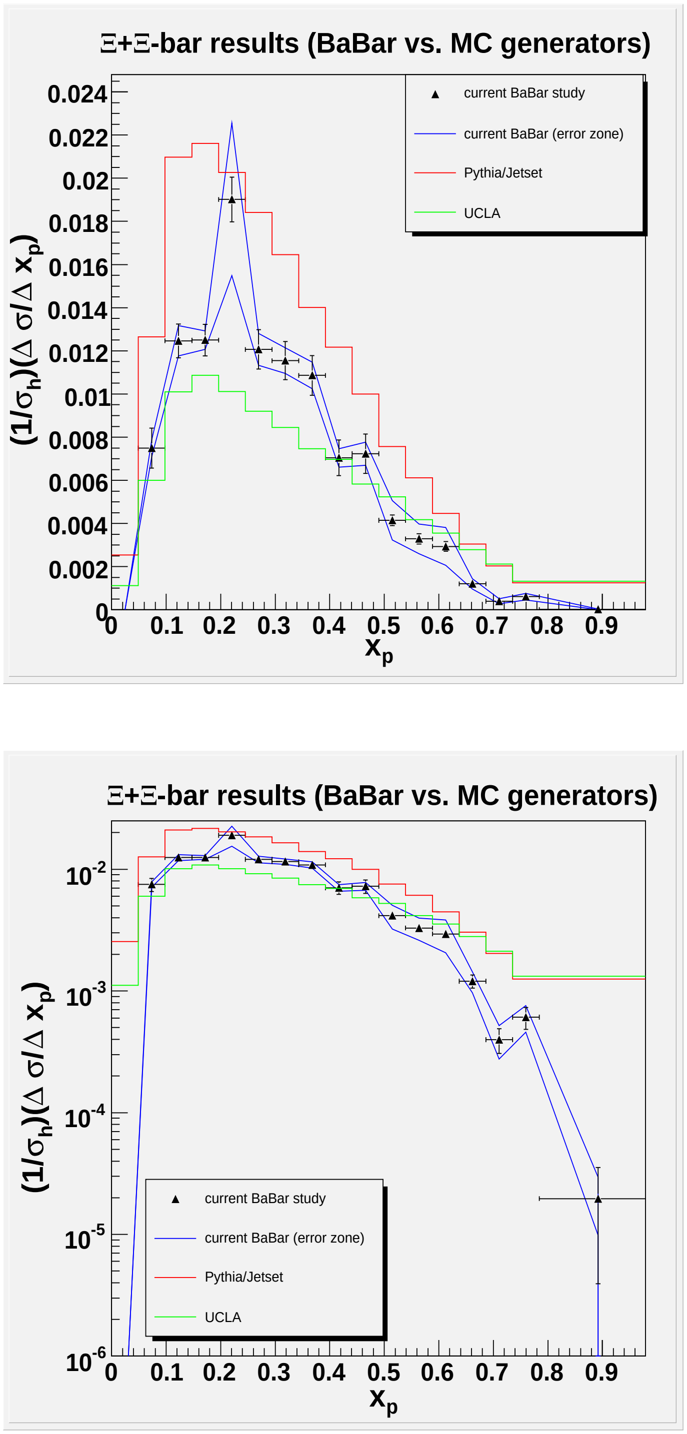


Figure 7.27: (Top) BaBar inclusive continuum spectrum for Ξ^- plus $\overline{\Xi}^+$ at 10.54 GeV compared with the spectra from the default UCLA generator and BaBar's SP5 (Phythia) generator. (Bottom) Log plots of the same spectra.

Candidate p_{CM} bins (GeV/c)	Raw signals	Partially corrected #'s	Fully corrected	% bin stat. uncertainty
0.00-0.50	4	380	501	50.0%
0.50-1.00	9	970	1,231	33.3%
1.00-1.50	31	347	641	18.0%
1.50-2.00	37	300	560	16.4%
2.00-2.50	26	208	404	19.6%
2.50-3.00	8	138	226	35.4%
3.00-3.50	6	57	122	40.8%
3.50-4.00	2	16	124	70.7%
4.00-5.00	0	0	0	–
$\sum$ (0.00-4.00)	123	2,416	3,809	13.8%

Table 7.11: Fitted numbers of Ω^- and $\overline{\Omega}^+$ and associated statistical uncertainties in p_{CM} bins.

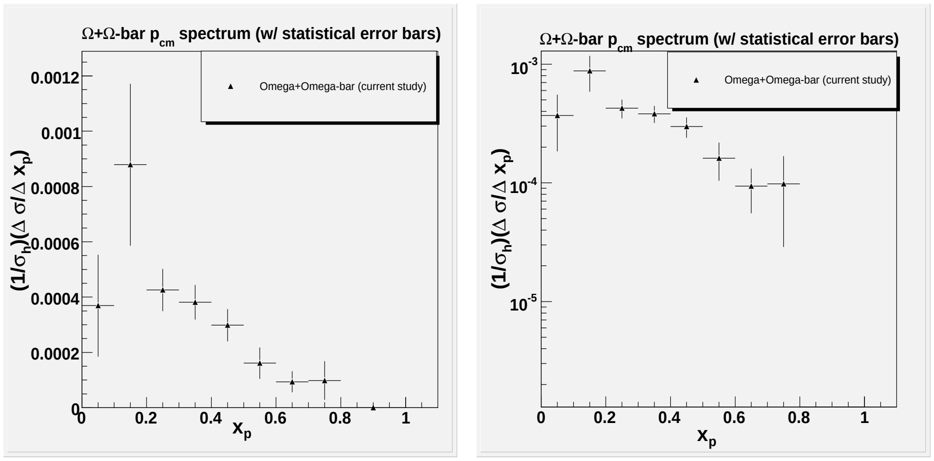


Figure 7.28: (Left) Preliminary spectrum of Ω^- and $\overline{\Omega}^+$. (Right) Same as left plot but now in log-format.

Cand p_{cm} (GeV/c)	Fully corrected Numbers (errors added)	% errors Trk. + MC + Λ
0.00-0.50	954	90.4%
0.50-1.00	949	22.9%
1.00-1.50	834	30.1%
1.50-2.00	644	15.0%
2.00-2.50	478	18.3%
2.50-3.00	228	1.0%
3.00-3.50	156	27.9%
3.50-4.00	163	31.5%
4.00-5.00	0	–
$\sum$ (0.00-4.00)	4,406	15.7%

Table 7.12: Ω systematic uncertainties in p_{CM} bins due to tracking, MC efficiency matrices' bin statsitics, and SVT/beampipe Λ fractions. The percent uncertainties are reported as percentages of the difference between the column two numbers and the total estimated numbers of fully-corrected signal particles in that bin as reported in Table 7.11.

7.3.2 Propagated tracking and efficiency uncertainties

For the Ω^- and $\overline{\Omega}^+$ hyperons, we propagated the tracking errors, errors on MC statistics, and the error due to the beampipe/SVT Lambda discrepancy in nearly the same manner as that for the cascade hyperons (see §7.2.2.) The major difference between the Omega and the cascade hyperons lies in the fact that we have to use the 'SVT Omega sample' for the entire center of mass momentum range; recall that for the case of the cascade hyperon, we used 'beampipe' cascade candidates for the CM momentum region below 2.50 GeV/c while we employed the beampipe + SVT cascade sample for the momentum region between 2.50 and 5.10 GeV/c.

Here, the term 'SVT Omega sample' refers to Ω^- and $\overline{\Omega}^+$ candidates which decayed either within the PEP-II beampipe or within the inner SVT (4-10 cm. in BaBar's XY-plane) *and* whose Lambda daughters also decayed within the inner SVT (< than 12 cm. in BaBar's XY-plane.) The 'beampipe Omega sample' then refers to Ω^- and $\overline{\Omega}^+$ candidates which decayed entirely within the PEP-II beampipe in addition to the fact that these candidates' Lambda daughters also decayed within the beampipe.

For the Omega hyperons we must employ the SVT sample for the entire Omega momentum range due to the intrinsic scarcity of the Omega particles: there simply is insufficient numbers of beampipe Omegas to allow us to perform momentum-binned measurements.

As in the case for the cascade hyperon, the tracking, MC statistics, and beampipe/SVT Lambda systematic sources of errors were incorporated into the momentum-binned invariant mass plots of the Omega during efficiency-correction and the results were fitted to extract an alternative spectrum; the difference between the alternative spectrum and our primary spectrum for the Omega then defines a partial systematic error zone. We present the alternative Omega spectrum which incorporates the above-described sources of systematic error in column two of Table 7.12. Figure 7.29 displays the error zone extracted from the information in Tables 7.11 and 7.12.

Note that in Figure 7.29, the beampipe/SVT Lambda error and the MC statistics error

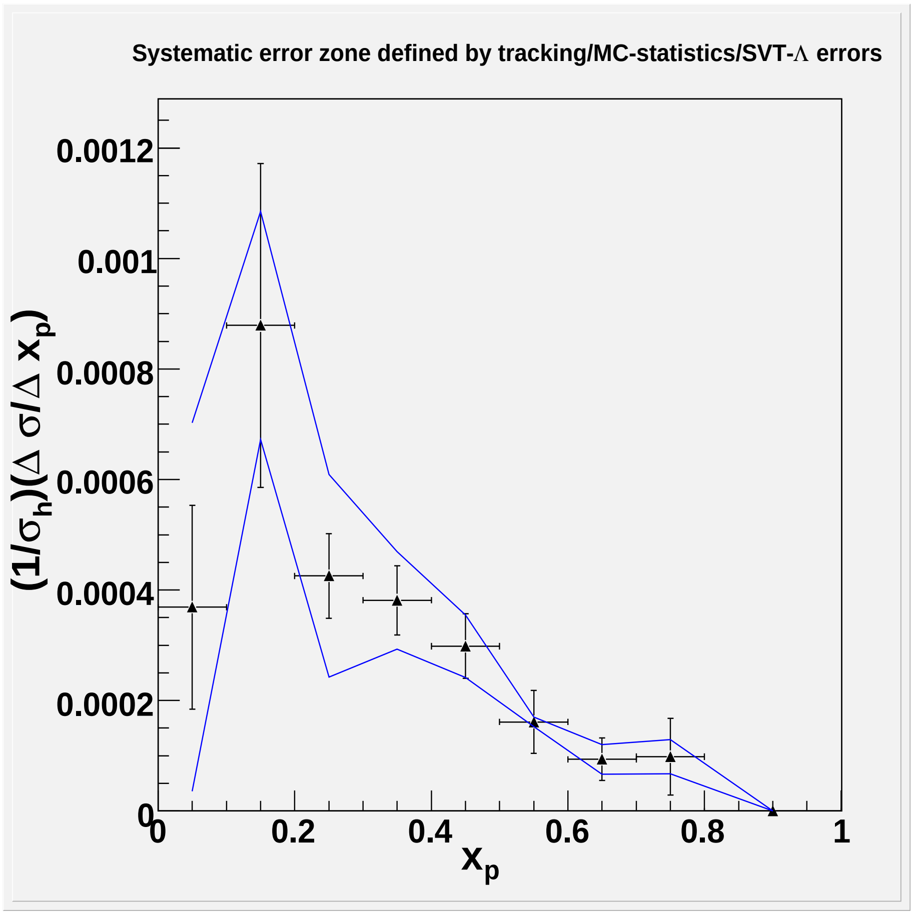


Figure 7.29: Zone of correlated errors due to tracking corrections, $\Omega^-/\overline{\Omega}^+$ efficiency bin statistics, and beampip/SVT Λ discrepancy; in units of fractional hadronic cross section.

dominate all other errors [28]; in fact, though there are additional sources of systematic error that contribute to the final Omega systematic error zone, the zone displayed in this figure (aside from the two data points in the momentum range 1.0-2.0 GeV/c corresponding to an x_p range of 0.2-0.4) is visually indistinguishable from the final Omega error zone due to the sizes of all other errors relative to the large beampipe/SVT Lambda error.

7.3.3 Consistency check using separate $\Omega^-/\overline{\Omega}^+$ fits

We also fitted the Ω^- and $\overline{\Omega}^+$ subsamples separately as a consistency check. Figure 7.30 displays the spectra of Ω^-, $\overline{\Omega}^+$, and the combined sample.

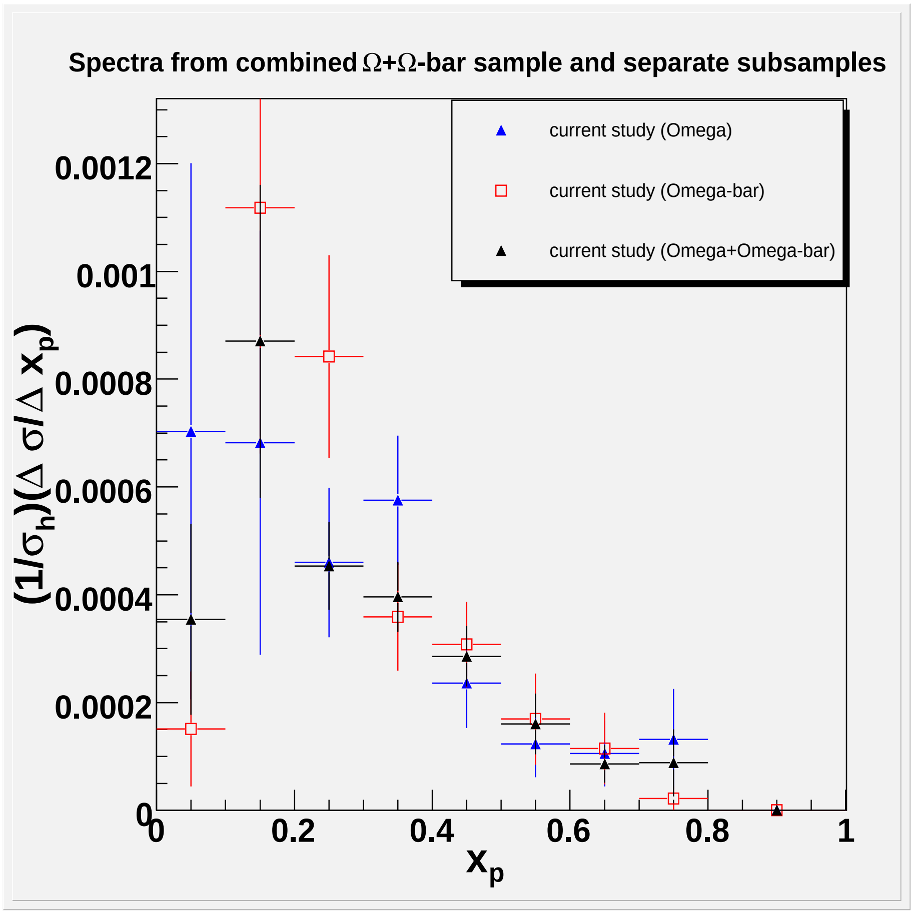


Figure 7.30: Comparison of the separate Ω^- and $\overline{\Omega}^+$ spectra with the combined Ω+$\overline{\Omega}^+$ spectrum

The particle and antiparticle numbers used in Figure 7.30 were obtained from the fits displayed in Figures 7.31 and 7.32.

In order to enable a better view of the presence of possible signals within certain bins that

---

[28] In the production rate, the beampipe/SVT Lambda error also dominates the MC statistics error since the former is momentum-correlated and is summed linearly while the latter is an independent error and is summed quadratically.

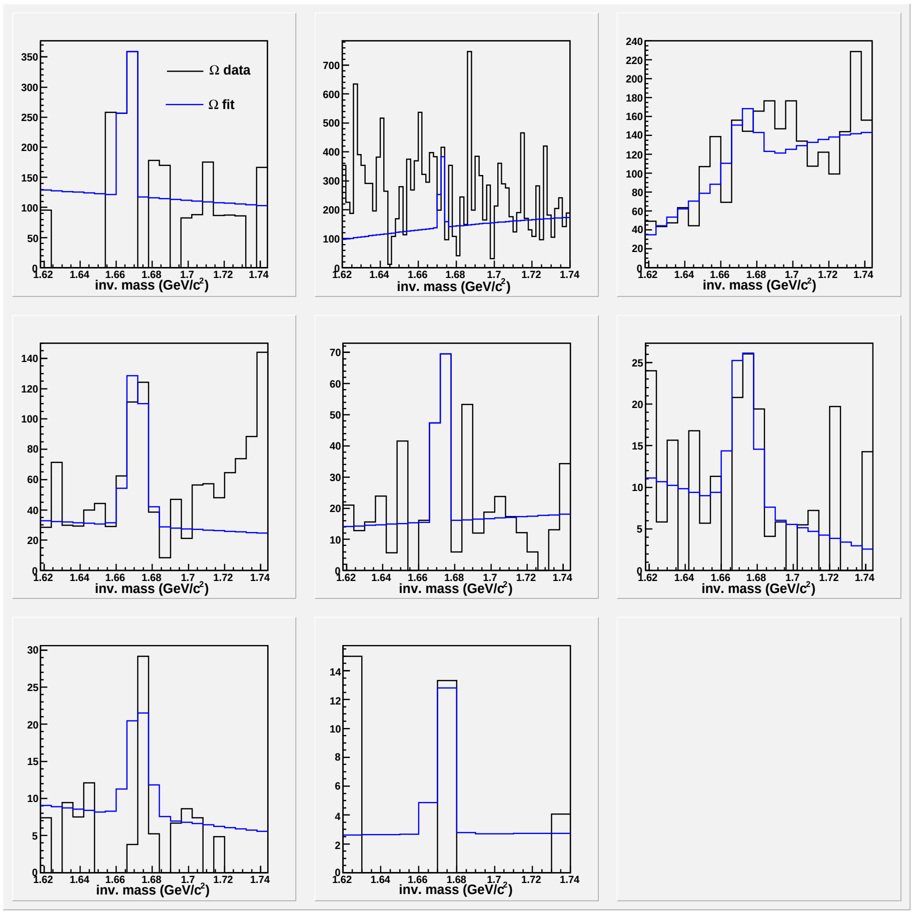


Figure 7.31: Momentum-binned invariant mass plots for the Ω^- subsample and fits.

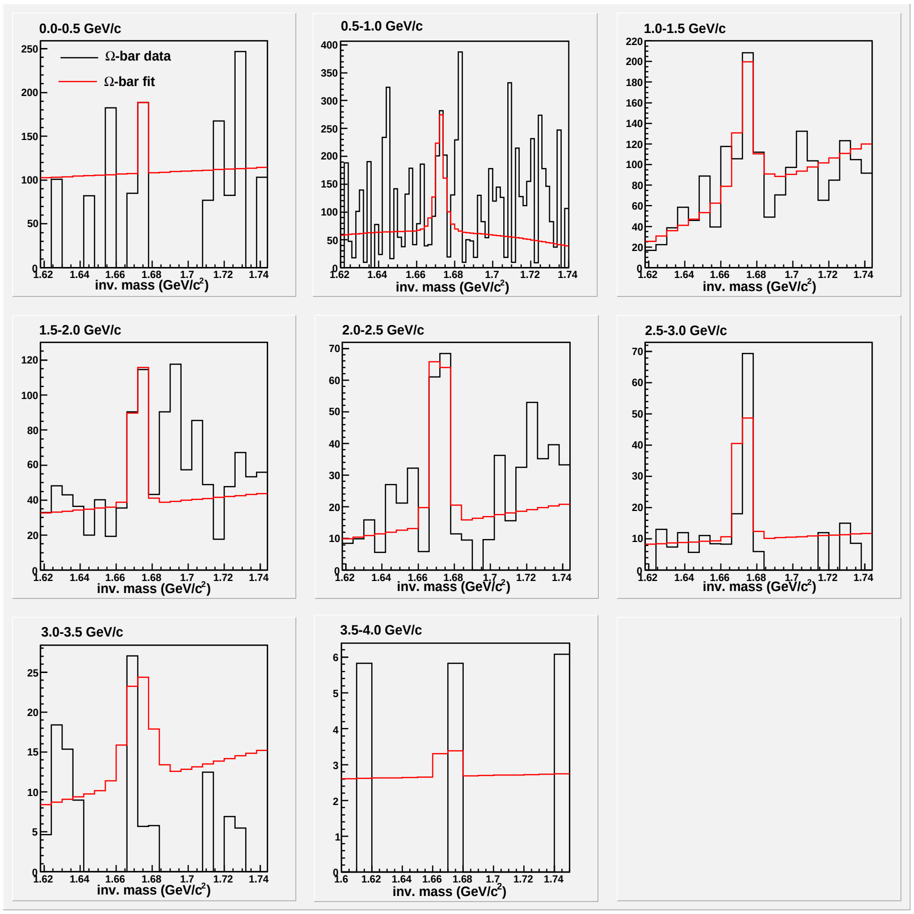


Figure 7.32: Momentum-binned invariant mass plots for the $\overline{\Omega}^+$ subsample and fits.

is free of the often distracting effects of efficiency corrections, we also display the momentum-binned mass plots of the Ω^- and $\overline{\Omega}^+$ subsamples without any efficiency corrections. These are contained in Figures 7.33, 7.34.

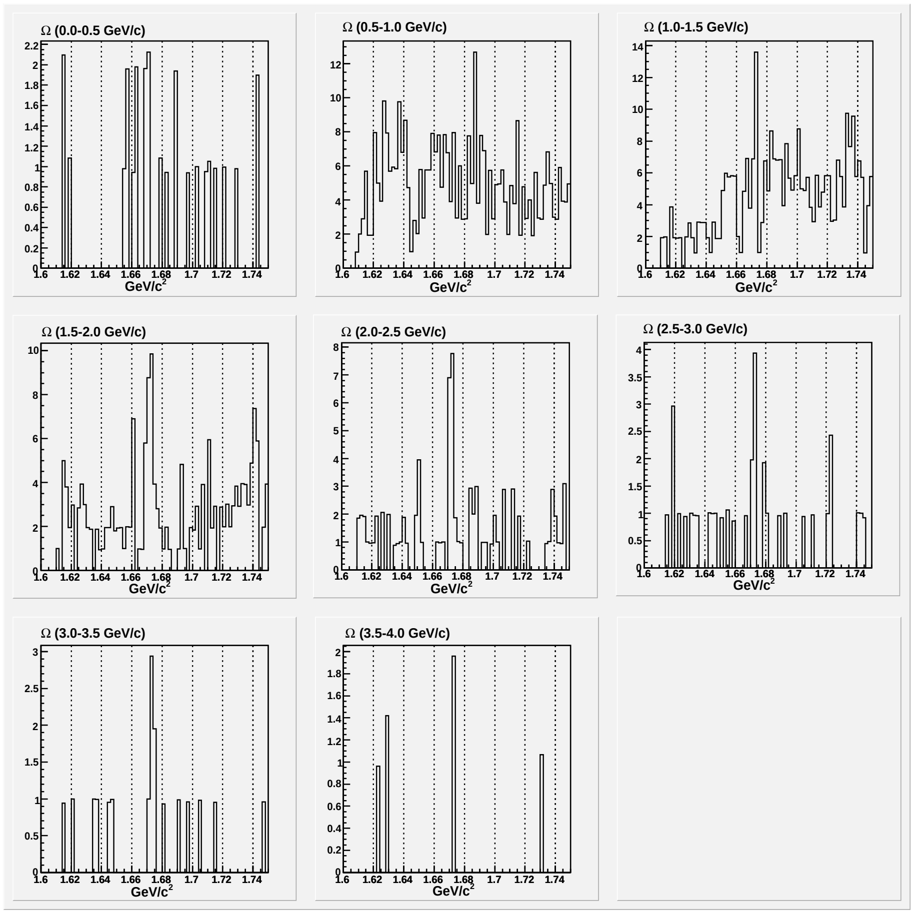


Figure 7.33: Momentum-binned invariant mass plots for the Ω^- subsample without any efficiency corrections.

We observe that only two data points in Figure 7.30 (the $\overline{\Omega}^+$ data point at $x_p = 0.25$ and the Ω data point at $x_p = 0.35$) appear not to agree with the other spectra within statistical error margins. We shall increase the systematic error zone to encompass the discrepancies between our primary spectrum and the spectra of the subsamples.

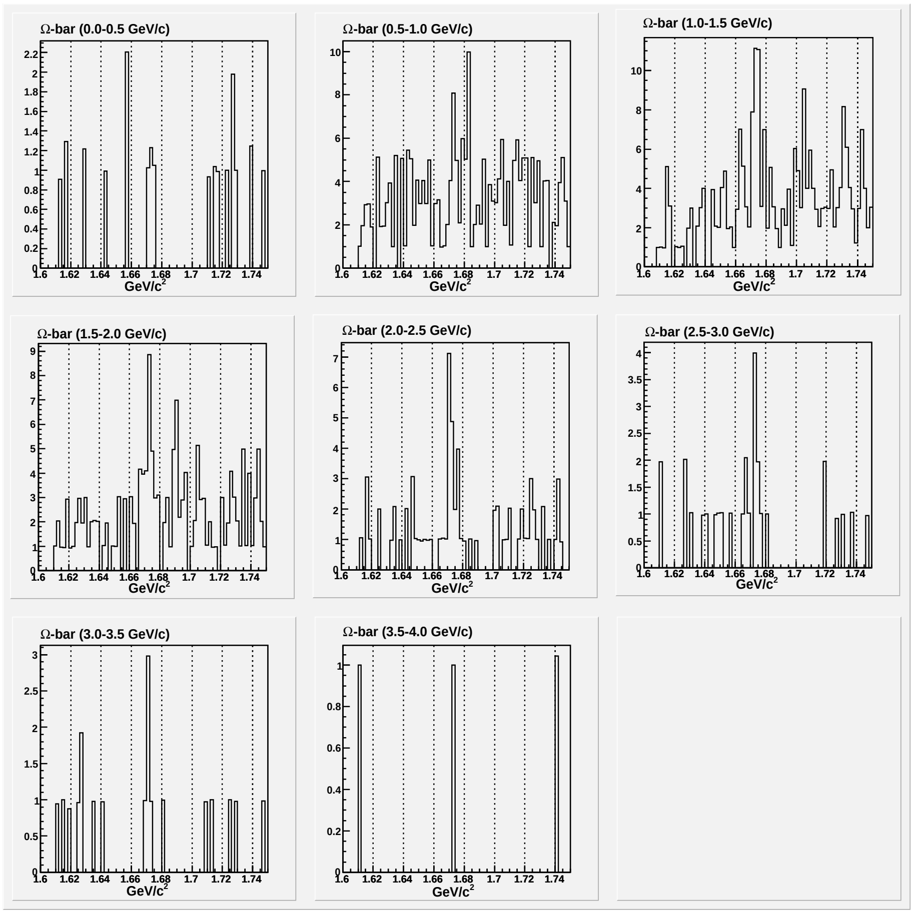
7.3.4 Final Omega spectrum and rate

The relative scarcity of the Omega hyperon makes it impractical to perform additional checks such as the check on the estimated MC statistical errors, beampipe/SVT sample difference check, momentum bin shifts, etc. Instead, we rely on the Lambda and the cascade hyperon studies for obtaining estimates on these sources of systematic errors.

There is an additional source of systematic error that is unique to the Omega hyperon. In order to extract visible signals, we imposed the additional PID cut on the kaon candidate [29].

---

[29] We require that the Global Likelihood selector does *not* identify the kaon candidate either as a pion or as


Ω-bar (0.0-0.5 GeV/c)

Ω-bar (0.5-1.0 GeV/c)

Ω-bar (1.0-1.5 GeV/c)

Ω-bar (1.5-2.0 GeV/c)

Ω-bar (2.0-2.5 GeV/c)

Ω-bar (2.5-3.0 GeV/c)

Ω-bar (3.0-3.5 GeV/c)

Ω-bar (3.5-4.0 GeV/c)

Figure 7.34: Momentum-binned invariant mass plots for the $\overline{\Omega}^+$ subsample without any efficiency corrections.

Error Source	Type	Average error size (% of σ_{Ω})
Statistical error	independent	13.8%
MC efficiency error	independent	18%
Λ^0 PDG mean lifetime error	independent	0.8%
Ω^- PDG mean lifetime error	independent	1.3%
Event purity error	independent	0.2%
Cand. p-shifts error	corr./indep.	0.5%
Fitting errors	corr./indep.	5.0%
DCH+SVT tracking error	normalization	3.1%
proton PID error	normalization	3.5%
kaon PID error	normalization	2.9%
Beampipe/SVT Λ^0 difference	correlated	15%
Extra error from spectra discrepancies	independent	5%
Last bin extrapolation	correlated	$< 1\%$

Table 7.13: Summary of the sources of error for $\Omega^-/\overline{\Omega}^+$ spectrum and production rate. Errors labeled as "corr./indep." are those which may be correlated with Omega momentum but which we are nevertheless treating as if they are independent errors. The average error sizes for the independent errors are larger than their contributions to the total production rate error.

As is the case for proton PID, we did not conduct separate studies on the MC/data PID disagreements for our SP5 Monte Carlo sample. Instead, we rely on the maximum PID data/MC difference for kaon misidentification rates estimated for SP4 Monte Carlo given in [30]. Note that the data/MC differences in kaon misidentication rates given in [30] were estimated using kaons from the $D^* \to D^0\pi \to K\pi\pi$ decay process [30].

We list the various sources of errors and their estimated average sizes in Table 7.13. We estimate the error contribution due to the statistics of the MC efficiency grids is about 18% (using the average statistical errors in accepted bins displayed in Figure 6.33.) We use the average DCH and SVT tracking errors estimated for the cascade hyperon for the Omega as well since the cascade and Omega tracking acceptance criteria were identical.

Table 7.14 contains a momentum-binned account of the continuum inclusive differential cross section and errors of the Ω^- plus $\overline{\Omega}^+$ spectrum. The Omega differential cross sections per unit x_p are given as fractions of the estimated total hadronic cross section and have been corrected for the presence of ISR in the events. Again, the total systematic error in any particular bin includes only the independent and momentum-correlated errors *from the same bin* summed in quadrature.

Figure 7.35 graphically displays the information contained in Table 7.14.

The Omega production rate is obtained from the total in Table 7.14 plus extrapolations for the last momentum bin (x_p 0.8-1.0) where we obtained zero signal candidates from data. We perform the extrapolation using the Pythia generator. The extrapoloated fractional cross sectional contribution to the production rate from the last momentum bin is 0.000004 (the

---

a proton.

[30] The data/MC difference for a real kaon misidentified as a pion is about 2.5% while the difference for a real kaon misidentified as a proton is about 1.5%

p_{CM} bins (GeV/c)	x_p Range	$\Omega^-+\overline{\Omega}^+$ $(\Delta\sigma/\Delta x_p)$	Stat. Err. $(\Delta\sigma/\Delta x_p)$	Sys. Err. $(\Delta\sigma/\Delta x_p)$
0.00-0.50	0.00-0.10	0.000369	0.000185	0.000333
0.50-1.00	0.10-0.20	0.000879	0.000293	0.000200
1.00-1.50	0.20-0.30	0.000426	0.000076	0.000181
1.50-2.00	0.30-0.40	0.000381	0.000063	0.000085
2.00-2.50	0.40-0.50	0.000298	0.000059	0.000077
2.50-3.00	0.50-0.60	0.000161	0.000057	0.000029
3.00-3.50	0.60-0.70	0.000094	0.000038	0.000026
3.50-4.00	0.70-0.80	0.000098	0.000069	0.000031
$\sum$ (0.00-4.00)	$\sum$ (0.00-0.80)	0.000271	0.000038	0.000082

Table 7.14: $\Omega^-+\overline{\Omega}^+$ CM momentum spectrum (presented as fractions of hadronic σ_{hadron} per unit x_p.) The total statistical error was obtained by a quadrature sum. The total systematic error contribution from independent errors also are summed quadratically while the contributions from correlated errors are summed linearly; the total systematic includes the normalization errors.

extrapolated fractional cross section per unit x_p for this bin is 0.000018,) while the systematic error contribution due to the extrapolation is estimated at 0.000008 (in units of fractional hadronic cross section.) However, since this extrapolation error is added quadratically with the rest of the systematic errors when calculating the total production rate systematic error, the total systematic error level quoted in Table 7.14 is not altered by the extrapolation (after accounting for the relevant significant digits.)

We quote the Ω^- plus $\overline{\Omega}^+$ production rate at BaBar energies to be the following: $0.00027 \pm 0.00004(stat.) \pm 0.00008(sys.)$.

7.3.5 Comparisons of Omega rate with prior results and with Monte Carlo

Neither the CLEO collaboration nor the ARGUS collaboration reported a momentum spectrum for the Omega hyperon near 10 GeV. Thus, the present study represents the first measurement of the Omega momentum spectrum at this energy.

The ARGUS did report a total production rate for the Omega: $0.00072 \pm 0.00036(stat.) \pm 0.00013(sys.)$ Our current analysis produced a production rate that is less than half the ARGUS value; in that sense, there appears to be a discrepancy between the two measurements. However, since the total ARGUS production rate error is about 53% and our current measurement, while more accurate, still has an error rate of 33

We also compare our current measured Omega momentum spectrum with the spectra produced by the default Pythia/Jetset and the UCLA generators (the 'default' generator parameters are the best fit values considering all prior results for mesons and baryons at several different energies.) See Figure 7.35.

Based on this figure and the associated table, and also taking into account the large error bars for the two lowest x_p bins, we can say that our current measurement does not disagree with the spectra generated by either of these models although currently the default Pythia model agrees slightly better with our current measurement.

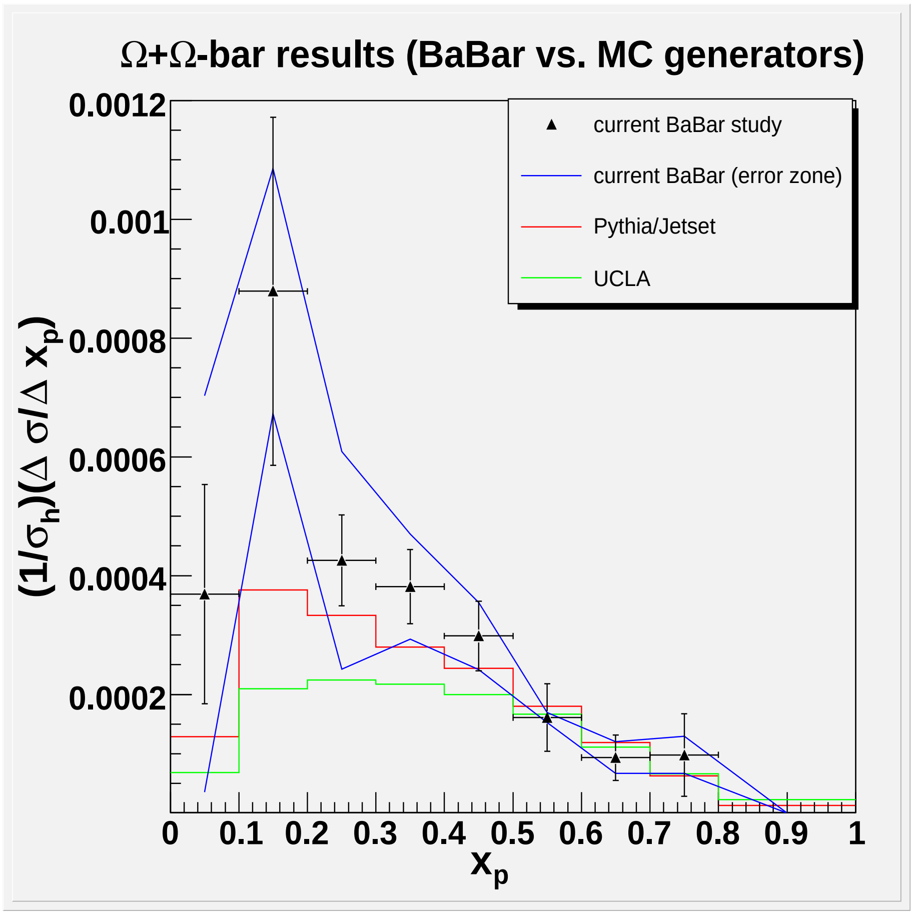


Figure 7.35: (Top) BaBar inclusive continuum spectrum for Ω^- plus $\overline{\Omega}^+$ at 10.54 GeV compared with the spectra produced by the default Pythia/Jetset and the UCLA generators.

Chapter 8

Summary and outlook

We analyzed the momentum spectra and production rates of the Lambda hyperon, the cascade hyperon, and the Omega hyperon using part of the currently available BaBar data collected at $\sqrt{s} = 10.54$ GeV.

Despite the fact that we've used less than half of the available off-peak BaBar data, and despite the presence of some unresolved issues that increased the current systematic error estimates for the cascade and the Omega hyperons, our present study already has produced world-best measurments of the continuum momentum spectra and production rates for these three baryons near 10 GeV. Our measured spectra have more data points as well as smaller bin statistical and systematic errors than the spectra produced by either the CLEO or the ARGUS collaborations. In fact, the Omega momentum spectrum is the first such measurement near 10 GeV. We hereby summarize the production rates measured by this present study:

$$\Lambda^0 + \overline{\Lambda}^0 \longrightarrow 0.0900 \pm 0.0006(\text{stat.}) \pm 0.0039(\text{sys.}) \text{ per hadronic event}$$

$$\Xi^- + \overline{\Xi}^+ \longrightarrow 0.00562 \pm 0.00013(\text{stat.}) \pm 0.00045(\text{sys.}) \text{ per hadronic event}$$

$$\Omega^- + \overline{\Omega}^+ \longrightarrow 0.00027 \pm 0.00004(\text{stat.}) \pm 0.00008(\text{sys.}) \text{ per hadronic event}$$

The current world average inclusive production rates for the three hyperons are: 0.080 ± 0.007 for the Λ^0, 0.0059 ± 0.0007 for the cascade hyperon, and 0.0007 ± 0.0004 for the Ω^- [38] (the systematic and statistical errors have been combined by quadrature in the world averages.) We immediately notice that our current results all agree with the current world averages within common error margins. In the case of Omega however, this is not particularly significant since the current world average value for the production rate has a 53% error margin.

Our current Lambda production rate measurement is 1.3 standard deviation (12%) above the previous world average and has a total measurement error that is 50% of that quoted for the world average. Our cascade hyperon's total measurement error on the production rate is 0.3 standard deviation (5%) below the previous world average with a total error that is 67% of the world average error. Our Omega production rate is more than 50% below the previous world average (however, this is only about one standard deviation lower than the world average due to the large uncertanties associated with the world average;) we also have reduced the absolute error on the Omega hyperon production rate to about 22% of the previous world average error rate.

If we were to include our current measurements with the prior measurements, then the new world averages for the inclusivve production rates of the Lambda, cascade, and Omega hyperons from e^+e^- collision events would be:

$$\Lambda^0 + \overline{\Lambda}^0 \quad \longrightarrow \quad 0.087 \pm 0.003 \qquad \text{per hadronic event}$$

$$\Xi^- + \overline{\Xi}^+ \quad \longrightarrow \quad 0.0057 \pm 0.0004 \qquad \text{per hadronic event}$$

$$\Omega^- + \overline{\Omega}^+ \quad \longrightarrow \quad 0.00030 \pm 0.00009 \qquad \text{per hadronic event}$$

It is clear that our results would dramatically reduce the world average error margins for all three of these hyperons. In addition, our results for the Lambda, cascade, and Omega production rates would be the dominating contributions to the revised world averages. The improved rates and the precision on the Lambda and cascade hyperon numbers should help modelers to more closely simulate the data and possibly enhance our understanding of baryon formation. In particular, since our new Omega number would reduce the relevant world average production rate number by more than 50% and in addition would reduce the absolute error level on the Omega number by a factor of four, this measurement would force the various hadronization generators to conform to a much tighter set of constraints as they attempt to model the physics of strange baryon production.

Despite the improvements in errors that we have achieved in the current analysis, it is clear that additional work using all available BaBar data can further refine the current measurements and reduce the associated errors. Some areas for further work include the following non-exhaustive list:

- Update the analysis software to enable analysis of off-peak BaBar data from the Run 5, 6, and 7 run periods.
- Resolve the apparent discrepancy between the Lambda hyperon spectrum from the ‘beampipe’ and the SVT samples; the successful resolution of this issue would enable the inclusion of more data (even without Runs 5-7) for the cascade and Omega hyperon measurements.
- Significantly enhance the number of Monte Carlo events used in the analysis and employ events generated by an updated MC Simulation Production code.
- Employ the Lambda hyperon to conduct a study of momentum-correlated MC/data differences in the Global Likelihood PID selector.
- Employ kaons from the $D^* \rightarrow D^0\pi \rightarrow K\pi\pi$ decay process to study the MC/data differences in kaon tracking efficiencies.
- Use on-peak BaBar data (collected near the $\Upsilon(4s)$ resonance at 10.58 GeV) to increase the hyperon statistics in high momentum bins (generally > 2.5 GeV/c.)
- Explore the possibility of using different momentum bin widths to improve the statistics in certain momentum bins.

- Explore whether using fits of the particle efficiency rates (in lieu of using our current grid smoothing procedure) would improve the MC statistical uncertainties. etc.

In addition, techniques similar to those employed in the present study can be applied to obtain the spectra and production rates for additional baryons such as the $\Sigma^{(1385)}$ and the $\Xi^{(1530)}$.

Bibliography

[1] Chris Quigg. *Gauge Theories of the Strong, Weak, and Electromagnetic Interactions.* Addison-Wesley, Reading, Massachusetts, 1983.

[2] ARGUS Collaboration and H. Allbrecht et. al. Hyperon production in e^+e^- annihilation at 10 GeV center of mass energy. *Z. Phys. C*, 39:177–186, 1988.

[3] CLEO Collaboration and S. Behrends et. al. Inclusive hadron production in upsilon decays and in nonresonant electron-positron annihilation at 10.49 GeV. *Phys. Rev. D*, 31:2161–2191, 1985.

[4] Donald H. Perkins. *Introduction to High Energy Physics.* Addison-Wesley, Reading,Massachusetts, 1972.

[5] Michael E. Peskin and Daniel V. Schroeder. *An Introduction to Quantum Field Theory.* Addison-Wesley, Reading, Massachusetts, 1995.

[6] Wu-Ki Tung. *Group Theory in Physics.* World Scientific, Singapore, 1985.

[7] R.D. Field and R.P. Feynman. Quark elastic scattering as a source of high-transverse-momentum mesons. *Phys. Rev. D*, 15:2590–2616, 1977.

[8] A. Casher, J. Kogut, and Lenoard Susskind. Vacuum polarization and the quark-parton puzzle. *Phys. Rev. Lett.*, 31:792–795, 1973.

[9] R.D. Field and R.P. Feynman. A parametrization of the properties of quark jets. *Nucl. Phys. B*, 136:1–76, 1978.

[10] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand. Parton fragmentation and string dynamics. *Phys. Rep.*, 97:31–145, 1983.

[11] B. Andersson, G. Gustafson, and B. Söderberg. A general model for jet fragmentation. *Z. Phys. C*, 20:317–329, 1983.

[12] G. Veneziano. An introduction to dual models of strong interactions and their physical motivations. *Phys. Rep.*, 9 C:199–242, 1974.

[13] L. Susskind. Harmonic-oscillator analogy for the veneziano model. *Phys. Rev. Lett.*, 23:545–547, 1969.

[14] P. Goddard, J. Goldstone, C. Rebbi, and C.B. Thorn. Quantum dynamics of a massless relativistic string. *Nucl. Phys.*, B56:109–135, 1973.

[15] X. Artru and G. Mennessier. String model and multiproduction. *Nucl. Phys.*, B70:93–115, 1974.

[16] T. Sjöstrand, L. Lönnblad, S. Mrenna, and P. Skands. *Pythia 6.2: Physics and Manual.* hep-ph/0108264, Lund, Sweden, 2002.

[17] S. Chun and C. Buchanan. A simple plausible path from QCD to successful prediction of $e^+e^- \rightarrow$ hadronization data. *Phys. Rep.*, 292:239–417, 1998.

[18] B. Hartfiel. *Search for popcorn mesons in events with two charmed baryons.* PhD thesis, UCLA, 2005.

[19] G. Marchesini and B.R. Webber. Simulation of QCD jets including soft gluon interference. *Nucl. Phys.*, B238:1–29, 1984.

[20] B. Auber et al. The BABAR detector. *Nucl. Instr. and Meth.*, A 479:1–116, 2002.

[21] P.F. Harrison and H.R. Quinn, editors. *The BaBar Physics Book; Physics at an Assymetric B-Factory.* Report of the BaBar Physics Workshop and SLAC, 1998.

[22] B. Dunwoodie, C. Hast, J.P. Lees, and P. Robbe. Study of material interactions with gamma conversions and protons. *Internal BaBar document*, BAD #106:1–43, 1999.

[23] *Internal BaBar document*, Babar Note #230:1, 1999.

[24] S. Eidelman et al. Review of particles physics. *Phys. Lett.*, B 592:1, 2004.

[25] J.D. Jackson. *Classical Electrodynamics, 2nd ed.* John Wiley and Sons, New York, New York, 1975.

[26] R. Bartoldus et al. Trigger and filter documentation for Run1. *Internal BaBar document*, BAD #194:1–84, 2002.

[27] A. Dvoretskii and A. Ryd. Level 3 trigger tracking. *Internal BaBar document*, BAD #301:1–68, 2001.

[28] B. Jacobsen. Structure of the BaBar reconstruction software. *Internal BaBar document*, BaBar Reconstruction Note #1:1–7, 1996.

[29] Data Quality Group. DQG monitoring. *Internal BaBar document*, BAD #1592:1–32, 2007.

[30] T. Hadig et al. Production of charged $\pi^\pm$, $K^\pm$ and $p/\overline{p}$ in e^+e^- annihilations at $\sqrt{s} = 10.54\ GeV$ and in hadronic $\Upsilon(4s)$ decays. *Internal BaBar document*, BAD #610:1–97, 2003.

[31] Geant4 physics reference manual. December 2004.

[32] Private communication from Dennis H. Wright. June 2006.

[33] M. Bona et al. Report of the tracking efficiency task force for 2001. *Internal BaBar document*, BAD #324:1–48, 2001.

[34] T. Allmendinger et al. Tracking efficiency studies in release 12 and 14. *Internal BaBar document*, BAD #867:1–101, 2004.

[35] Glen Cowan. *Statistical Data Analysis.* Oxford University Press, Oxford, Great Britain, 1998.

[36] Vertexing and Composition Tools Group. The BaBar Vertexing. *Internal BaBar document*, BAD #102:1–83, 2005.

[37] G. Welch and G. Bishop. A introduction to the Kalman filter. *University of North Carolina-Chapel Hill, Dept. of Computer Science*, TR 95-041:1–16, 2006.

[38] Particle Data Group. Review of Particle Physics. *J. Phys G. Nucl. Part. Phys.*, 33:1, 2006.